Metody algebry liniowej w fizyce

materiały do wykładu dla kierunku Fizyka Techniczna WFiIS AGH

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Kraków 2023

References

[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 1,2*, Wydawnictwo Naukowe PWN, 2004.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Metody algebry liniowej w fizyce
Wykład nr 1
Data: ....

TEMAT: *Metody dowodzenia twierdzeń. Działania i wybrane struktury algebraiczne*

1.1 Notacja

Parą uporządkowaną (a, b) nazywamy zbiór $\{\{a\}, \{a, b\}\}$ podzbiorów zbioru $\{a, b\}$. Dwie pary $(a, b), (c, d)$ są równe wtedy i tylko wtedy, gdy $a = c \wedge b = d$. *Iloczynem kartezjańskim* zbiorów A i B nazywamy zbiór

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

Przykład 1.1.1. Wyznacz $A \times B$, $B \times A$, B^2.

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

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

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

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

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

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

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 1.1.3. $\prod_{i=0}^{6}(9 - j) = 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3$
$\prod_{i=1}^{n} i = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (-1) \cdot n = n!$
$\underbrace{x \cdot x \cdot \ldots \cdot x}_{n\ razy} = \prod_{i=1}^{n} x$

1.2 Metody dowodzenia twierdzeń

Twierdzenie	założenia $\Rightarrow$ teza
	$p \Rightarrow q$, gdzie p, q są zdaniami

Dowód to wykazanie prawdziwość implikacji $p \Rightarrow q$.

Dowód wprost
Przeprowadzając dowód wprost, dowodzimy prawdziwości tezy bezpośrednio poprzez dedukcję z założeń twierdzenia.

Przykład 1.2.1. Udowodnimy, że dla dowolnego $n \in \mathbb{N}$ liczba $n^5 - 5n^3 + 4n$ jest podzielna przez 5. Przekształcamy wyrażenie.

$$n^5 - 5n^3 + 4n = n(n^4 - 5n^2 + 4) = n(n^2 - 1)(n^2 - 4) = (n - 2)(n - 1)n(n + 1)(n + 2)$$

Otrzymaliśmy iloczyn pięciu kolejnych liczb naturalnych. Zatem jedna z ich jest podzielna przez 5. Stąd iloczyn jest podzielny przez 5.

Dowód nie wprost (łac. reductio ad absurdum – sprowadzenie do sprzeczności, łac. contradictio in contrarium – zaprzeczenie przeciwieństwa)
Przeprowadzając dowód nie wprost, z założenia o nieprawdziwości tezy wyprowadzamy sprzeczność z przyjętymi założeniami.

Przykład 1.2.2. Niech $a \in \mathbb{Z}$. Jeśli a^2 jest liczbą parzystą, to wówczas a również jest liczbą parzystą.

Przypuśćmy, że a jest liczbą nieparzystą. Wówczas istnieje $k \in \mathbb{Z}$ takie, że $a = 2k+1$. Stąd $a^2 = (2k+1)^2 = 4k^2+4k+1 = 2(2k^2+2k)+1$. Zatem a^2 jest liczbą nieparzystą, co przeczy założeniu twierdzenia. To oznacza, że nasze początkowe przypuszczenie jest fałszywe i a jest liczbą parzystą.

Przykład 1.2.3. Udowodnimy, że $\sqrt{2}$ jest liczbą niewymierną.

Przypuśćmy, że $\sqrt{2}$ jest liczbą wymierną.
Wówczas istnieją $a, b \in \mathbb{Z}$, $b \neq 0$ względnie pierwsze i takie, że $\sqrt{2} = \frac{a}{b}$.
Stąd wynika, że $2 = \frac{a^2}{b^2}$ lub równoważnie $2b^2 = a^2$.
Zatem a^2 jest liczbą parzystą. Stąd wynika, że również a jest liczbą parzystą, czyli istnieje $k \in \mathbb{Z}$ takie, że $a = 2k$.
Stąd $2b^2 = 4k^2$ oraz $b^2 = 2k^2$. Zatem b^2 jest liczbą parzystą, a więc również b jest liczbą parzystą.
Ułamek $\frac{a}{b}$ nie jest nieskracalny.
Otrzymaliśmy sprzeczność. Zatem $\sqrt{2}$ jest liczbą niewymierną.

Dowód indukcyjny
Dowód indukcyjny jest to dowód w którym wykorzystujemy **zasadę indukcji matematycznej**.

Twierdzenie 1.2.4 (Zasada indukcji matematycznej)**.** Niech $\Phi(n)$ będzie zdaniem logicznym dla $n \in \mathbb{N}_0$ oraz niech $n_0 \in \mathbb{N}_0$. Jeśli zachodzą warunki

i) $\Phi(n_0)$, ($\Phi(n_0)$ jest prawdziwe)

ii) $\forall n \in \mathbb{N}_0, n \geq n_0 \quad \Phi(n) \Rightarrow \Phi(n+1)$, (prawdziwa jest implikacja)

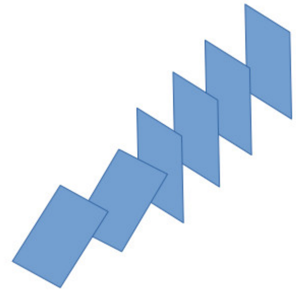
to wówczas $\forall n \in \mathbb{N}_0, n \geq n_0 \quad \Phi(n)$. ($\forall n \in \mathbb{N}_0$ zdanie $\Phi(n)$ jest prawdziwe)

Jeśli jakieś twierdzenie, w którym mowa o liczbach naturalnych,
jest prawdziwe dla określonej liczby naturalnej n_0 oraz
jeżeli dla każdej liczby naturalnej n z założenia, że twierdzenie jest prawdziwe dla n wynika,
że jest ono prawdziwe dla liczby następnej $n+1$,
to twierdzenie jest prawdziwe dla każdej liczby naturalnej nie mniejszej niż n_0.

Kostki domina

Czy wszystkie kostki się przewrócą?

Jeśli przewrócono pierwszą kostkę oraz
kostki są ustawione tak,
że upadek którejkolwiek z nich przewraca następną,
to wszystkie kostki się przewrócą.



Przykład 1.2.5. Sprawdźmy poniższe równości.

$$\begin{aligned}
&1^3 &&= \tfrac{1}{4}\cdot 1^2\cdot 2^2\\
&1^3+2^3 &&= \tfrac{1}{4}\cdot 2^2\cdot 3^2\\
&1^3+2^3+3^3 &&= \tfrac{1}{4}\cdot 3^2\cdot 4^2\\
&1^3+2^3+3^3+4^3 &&= \tfrac{1}{4}\cdot 4^2\cdot 5^2
\end{aligned}$$

Udowodnimy, że dla dowolnej liczby naturalnej $n \in \mathbb{N}$ prawdziwa jest równość

$$\sum_{i=1}^{n} i^3 = 1^3 + 2^2 + 3^3 + \ldots + n^3 = \frac{1}{4}n^2(n+1)^2.$$

I. Warunek początkowy

Sprawdziliśmy już, że dla $n_0 = 1$ wzór jest prawdziwy.

II. Krok indukcyjny (prawo przekazywania)

Niech n oznacza liczbę naturalną. Chcemy uzasadnić, że jeśli wzór jest prawdziwy dla n, to jest on prawdziwy dla $n+1$. Innymi słowy chcemy uzasadnić, że z założenia (tzw. założenie indukcyjne)

$$\sum_{i=1}^{n} i^3 = 1^3 + 2^2 + 3^3 + \ldots + n^3 = \frac{1}{4}n^2(n+1)^2$$

wynika teza

$$\sum_{i=1}^{n+1} i^3 = 1^3 + 2^2 + 3^3 + \ldots + n^3 + (n+1)^3 = \frac{1}{4}(n+1)^2(n+2)^2.$$

Dowód przedstawiamy poniżej.

$$L = \sum_{i=1}^{n+1} i^3 = (n+1)^3 + \sum_{i=1}^{n} i^3 \overset{z.ind.}{=} (n+1)^3 + \frac{1}{4}n^2(n+1)^2 =$$

$$= (n+1)^2\left(n+1+\frac{1}{4}n^2\right) = \frac{1}{4}(n+1)^2(4n+4+n^2) = \frac{1}{4}(n+1)^2(n+2)^2 = P$$

Zatem $L = P$, implikacja jest prawdziwa.

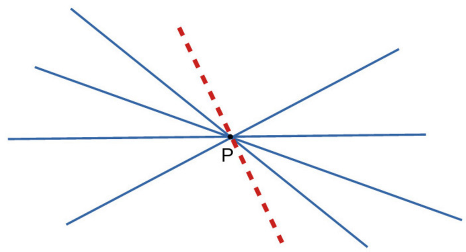
III. Wniosek

Na mocy zasady indukcji matematycznej wzór jest prawdziwy dla każdej liczby naturalnej, tj. $\forall n \in \mathbb{N} \quad \sum_{i=1}^{n} i^3 = \frac{1}{4}n^2(n+1)^2$.

Przykład 1.2.6. Udowodnimy, że n różnych prostych na płaszczyźnie, przechodzących przez ustalony punkt P, dzieli płaszczyznę na $2n$ części.

I. Warunek początkowy, dla $n_0 = 1$
Jedna prosta dzieli płaszczyznę na $2n_0 = 2$ części.

II. Krok indukcyjny
Niech $n \in \mathbb{N}$. Załóżmy, że n prostych przechodzących przez ustalony punkt P, dzieli płaszczyznę na $2n$ części. Chcemy uzasadnić, że wówczas $n+1$ prostych dzieli płaszczyznę na $2(n+1) = 2n+2$ części.



Poprowadźmy dodatkową prostą. W miejsce dwóch obszarów powstają cztery. A zatem mamy $2n+2$ obszary.

III. Wniosek: Na mocy zasady indukcji matematycznej twierdzenie jest prawdziwe dla dowolnej liczby prostych.

Przykład błędnego rozumowania indukcyjnego można znaleźć tutaj.

1.3 Działania wewnętrzne i ich własności

Niech A będzie zbiorem niepustym.

Definicja 1.3.1. 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 1.3.2. 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 1.3.3. 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 1.3.4. 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 1.3.5. Jeżeli element neutralny istnieje to jest jedyny.

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

Przykład 1.3.6. 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 1.3.7. 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 1.3.8. 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 $-2 - x$.

Twierdzenie 1.3.9. Jeżeli działanie wewnętrzne jest łączne oraz posiada element neutralny, to każdy element posiada co najwyżej jeden element symetryczny.

1.4 Podstawowe struktury algebraiczne

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

Definicja 1.4.1. i) Parę $(G, *)$ nazywamy *półgrupą*, jeżeli działanie jest łączne.

ii) 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, grupie przemiennej (abelowej).*

Przykład 1.4.2. Oznaczmy $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$.

	$(\mathbb{N}, +)$	$(\mathbb{Z}, \cdot)$	$(\mathbb{Q}, +)$	$(\mathbb{Q}, \cdot)$	$(\mathbb{R}^*, +)$	$(\mathbb{R}^*, \cdot)$
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.	półrupa przemienna z el. neutr.	grupa abelowa	półrupa przemienna z el. neutr.		grupa abelowa

Przykład 1.4.3. 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: półgrupa przemienna z elementem neutralnym

Definicja 1.4.4. 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

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

Definicja 1.4.5. Zespół $(K, \circ, *)$ złożony ze zbioru K zawierającego co najmniej dwa elementy i określonych w nim działań wewnętrznych $\circ : K \times K \to K$, $* : K \times K \to K$ nazywamy *ciałem*, jeśli

- $(K, \circ)$ jest grupą abelową (z elementem neutralnym $e_\circ$),
- $(K \setminus \{e_\circ\}, *)$ jest grupą abelową,
- działanie $*$ jest rozdzielne względem działania $\circ$.

Zatem 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.

Zatem w ciele można zdefiniować operację *dzielenia* w sposób następujący:

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

Przykład 1.4.6. 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) $(\mathbb{Z}/p\mathbb{Z}, +_p, \cdot_p)$, 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 (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

$\mathbb{Z}/4\mathbb{Z}$ 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

Przykład 1.4.7. Zbiór $\mathbb{R}^2$ wraz z działaniami dodawania i mnożenia zdefiniowanymi poniżej ma strukturę ciała.

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

Definicja 1.4.8. Zdefiniowane powyżej ciało nazywamy *ciałem liczb zespolonych* i oznaczamy symbolem $\mathbb{C}$. Elementy tego ciała nazywamy *liczbami zespolonymi.*

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Metody algebry liniowej w fizyce
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: $(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0) \in \Omega$, $(x_1, 0) \cdot (x_2, 0) = (x_1x_2, 0) \in \Omega$

przemienność:
$$(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0) = (x_2 + x_1, 0) = (x_2, 0) + (x_1, 0)$$
$$(x_1, 0) \cdot (x_2, 0) = (x_1x_2, 0) = (x_2x_1, 0) = (x_2, 0) \cdot (x_1, 0)$$

łączność:
$$[(x_1, 0) + (x_2, 0)] + (x_3, 0) = (x_1 + x_2 + x_3, 0) = (x_1, 0) + [(x_2, 0) + (x_3, 0)]$$
$$[(x_1, 0) \cdot (x_2, 0)] \cdot x_3, 0 = (x_1x_2x_3, 0) = (x_1, 0) \cdot [(x_2, 0) \cdot (x_3, 0)]$$

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:
$$[(x_1, 0) + (x_2, 0)] \cdot (x_3, 0) = (x_1 + x_2, 0) \cdot (x_3, 0) = ((x_1 + x_2)x_3, 0) =$$
$$= (x_1x + 3 + x_2x_3, 0) = [(x_1, 0) \cdot (x_3, 0)] + [(x_2, 0) \cdot (x_3, 0)]$$

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 $i^2 = -1$.

$$\begin{array}{ll} (x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2) & (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2) \\ (x_1, y_1) \cdot (x_2, y_2) = (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1) & (x_1 + iy_1) \cdot (x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1) \end{array}$$

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

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

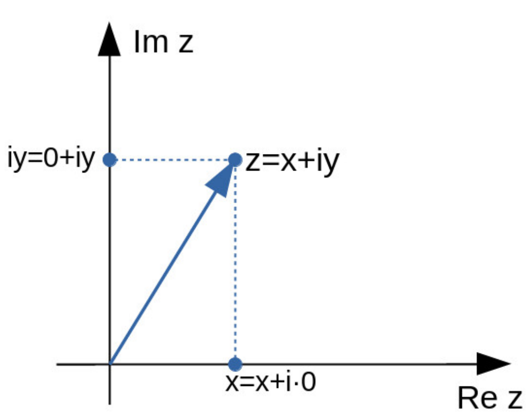
$\frac{2+3i}{2-5i} = \frac{(2+3i)(2+5i)}{(2-5i)(2+5i)} = \frac{4+10i+6i+15i^2}{4-25i^2} = \frac{-11+16i}{29} = -\frac{11}{29} + \frac{16}{29}i$

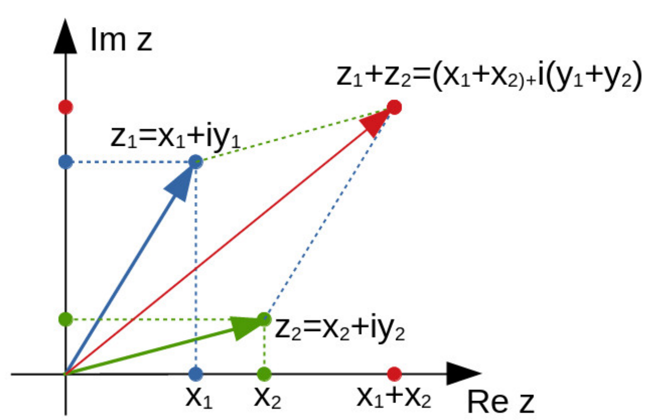
Uwaga 2.1.2. W ciele $\mathbb{C}$ nie można określić porządku liniowego.

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

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

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

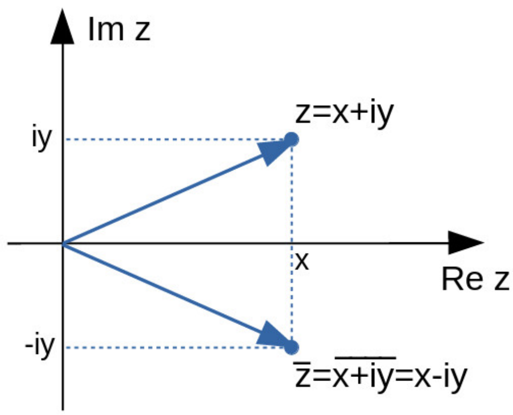


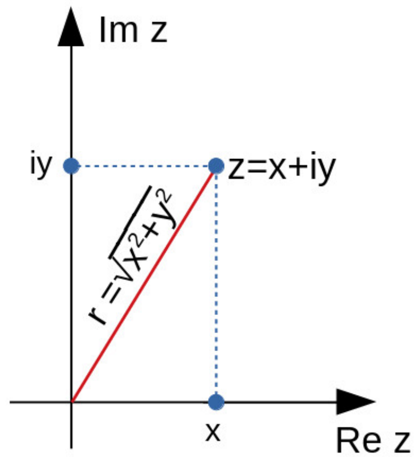


Definicja 2.1.3. Niech $z = x + iy \in \mathbb{C}$, $x, y \in \mathbb{R}$.

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 $|z|$.





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) $\overline{\left(\frac{z_1}{z_2}\right)} = \frac{\bar{z}_1}{\bar{z}_2}$, 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) $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$

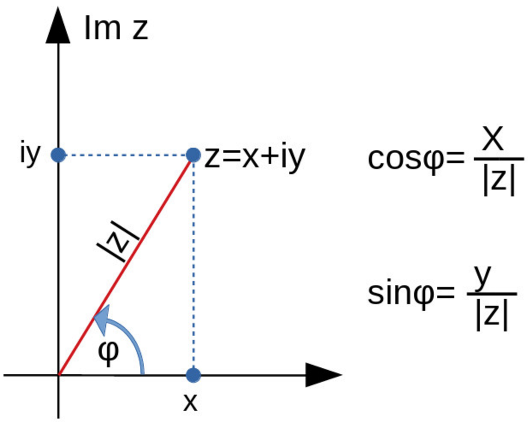
x) $|z_1 + z_2| \leq |z_1| + |z_2|$ (nierówność trójkąta) $\Big||z_1| - |z_2|\Big| \leq |z_1 - z_2|$

2.2 Postać trygonometryczna i wykładnicza liczby zespolonej

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

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



Dowolny taki kąt nazywamy *argumentem* liczby z. Ten z argumentów liczby zespolonej, który leży w przedziale $[0, 2\pi)$, 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. 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.

$$z = 7 \qquad\qquad z = -i \qquad\qquad z = -\sqrt{27} - 3i$$

$$z = 7(1 + 0i) \qquad\qquad |z| = \sqrt{0^2 + (-1)^2} = 1 \qquad\qquad |z| = \sqrt{27 + 9} = 6$$

$$z = 1\,(0 + (-1)\cdot i) \qquad\qquad z = 6\left(-\frac{\sqrt{3}}{2} - \frac{1}{2}i\right)$$

$$\begin{cases} \cos\varphi = 0 \\ \sin\varphi = -1 \end{cases} \qquad\qquad \begin{cases} \cos\varphi = -\frac{\sqrt{3}}{2} \\ \sin\varphi = -\frac{1}{2} \end{cases}$$

$$\varphi = 0 + 2k\pi \qquad\qquad \varphi = -\frac{\pi}{2} + 2k\pi \qquad\qquad \varphi = \frac{7}{6}\pi + 2k\pi, \qquad k \in \mathbb{Z}$$

$$\arg z = 0 \qquad\qquad \arg z = \frac{3}{2}\pi \qquad\qquad \arg z = \frac{7}{6}\pi$$

$$z = 7(\cos 0 + i\sin 0) \qquad -i = 1(\cos\tfrac{3}{2}\pi + i\sin\tfrac{3}{2}\pi) \qquad -\sqrt{27} - 3i = 6(\cos\tfrac{7}{6}\pi + i\sin\tfrac{7}{6}\pi)$$

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) $z_1 \cdot z_2 = |z_1| \cdot |z_2| \cdot \big(\cos(\alpha+\beta) + i\sin(\alpha+\beta)\big)$,

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*

Przykład 2.2.3. Oblicz $\left(\frac{1+i\sqrt{3}}{1-i}\right)^{20}$.

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

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

$$\left(\frac{1+i\sqrt{3}}{1-i}\right)^{20} = 2^{10}\left(\cos(\tfrac{140}{12}\pi) + i\sin(\tfrac{140}{12}\pi)\right) = 2^{10}\left(\cos(\tfrac{35}{3}\pi) + i\sin(\tfrac{35}{3}\pi)\right) =$$
$$= 2^{10}\left(\cos(12\pi - \tfrac{\pi}{3}) + i\sin(12\pi - \tfrac{\pi}{3})\right) = 2^{10}\left(\cos\tfrac{\pi}{3} - i\sin\tfrac{\pi}{3}\right) = 2^{10}\left(\tfrac{1}{2} - \tfrac{\sqrt{3}}{2}i\right) =$$
$$= 2^9(1 - \sqrt{3}i)$$

Twierdzenie 2.2.4 (Własności argumentu)**.** Niech $z, z_1, z_2 \in \mathbb{C}$.
Wówczas prawdziwe są następujące równości.

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

ii) $\arg(\frac{z_1}{z_2}) = \arg z_1 - \arg z_2 + 2k\pi$, $k \in \mathbb{Z}$ $(k = 0$ 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 2.2.5. $\arg i = \frac{\pi}{2}$ $\arg(-1) = \pi$ $\arg(-i) = \frac{3}{2}\pi$
$i = (-1)\cdot(-i) \;\Rightarrow\; \arg i = \pi + \frac{3}{2}\pi + 2k\pi = \frac{5}{2}\pi + 2k\pi$ dla pewnego $k \in \mathbb{Z}$, tj. $k = -1$

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

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

najpiękniejszy wzór w matematyce

Twierdzenie 2.2.7. Niech $\alpha, \beta \in \mathbb{R}$. Wówczas:

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

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

Wniosek 2.2.8. Niech $z = re^{i\varphi}$, $z_1 = r_1 e^{i\alpha}$, $z_2 = r_2 e^{i\beta}$ będą liczbami zespolonymi w postaci wykładniczej. Wówczas:

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

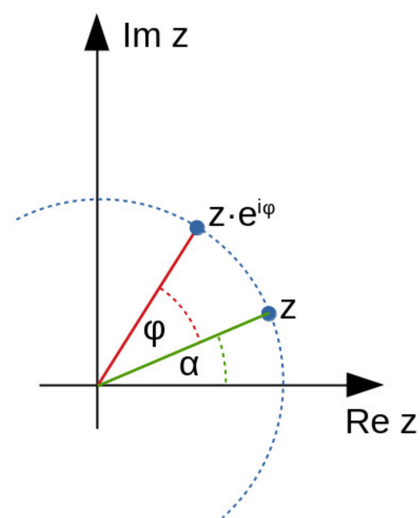
ii) $z^k = r^k e^{ik\varphi}$, dla $k \in \mathbb{Z}$,

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

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 :

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

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



2.3 Pierwiastkowanie, równania wielomianowe

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

Definicja 2.3.1. Każdą liczbę $z \in \mathbb{C}$ spełniająca 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 wykładnicza

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

$|w| = 10$
$w = 8 + 6i = 10(\frac{4}{5} + i\frac{3}{5})$
$\varphi = \arcsin\frac{3}{5} + 2k\pi$

$|z| = \sqrt{10}$
$\varphi_k = \frac{1}{2}\arcsin\frac{3}{5} + k\pi$

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

II 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 $z = x$, $x^2 \neq 8 + 6i$)
$x = \frac{3}{y}$, $\frac{9}{y^2} - y^2 = 8$
$y^4 + 8y^2 - 9 = 0$, $\Delta = 100$
$y^2 = 1$, $y = \pm 1$

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

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

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

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

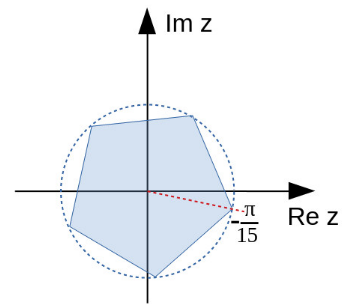
$$\sqrt[n]{w} = \{z \in \mathbb{C} : z^n = w\} = \{z_0, z_1, \ldots, z_{n-1}\}.$$

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

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

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

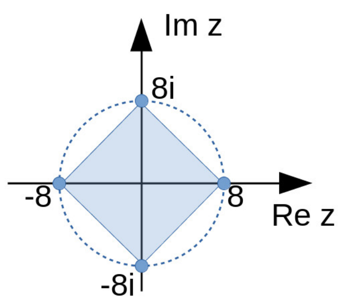


Liczby $z_0, z_1, \ldots, z_{n-1}$ będące rozwiązaniami równania $z^n = w$ stanowią wierzchołki n-kąta foremnego, wpisanego w koło o środku $z = 0$ i promieniu $\sqrt[n]{r}$.

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 $W(z) = (z - z_0)P(z)$.

Definicja 2.3.9. Niech $k \in \mathbb{N}$. 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.

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

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

Twierdzenie 2.3.14. Niech $k \in \mathbb{N}$. 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. 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 $r \in \mathbb{R}$, $r > 0$.

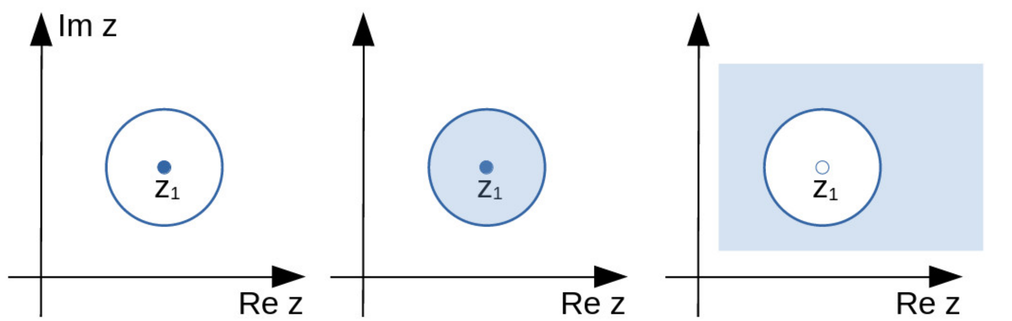
1) $|z_1 - z_2|$ odległość z_1 od z_2

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

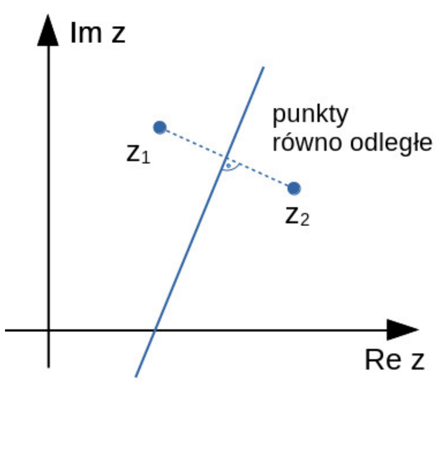
2) $|z - z_1| = r$ równanie okręgu o środku z_1 i promieniu r

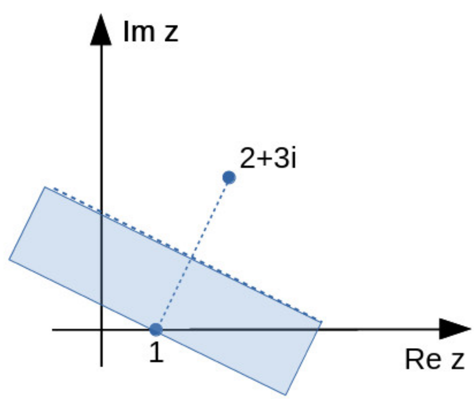
$r = |(x - x_1) + (y - y_1)i| = \sqrt{(x - x_1)^2 + (y - y_1)^2} \;\Rightarrow\; r^2 = (x - x_1)^2 + (y - y_1)^2$

$|z - z_1| = r$ okrąg $\qquad |z - z_1| \leq r$ koło $\qquad |z - z_1| \geq r$ zewnętrze koła



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

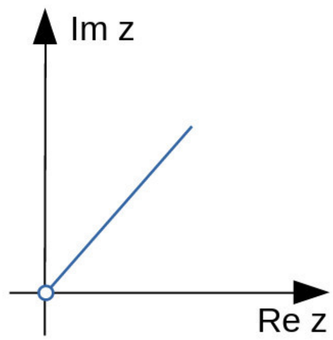




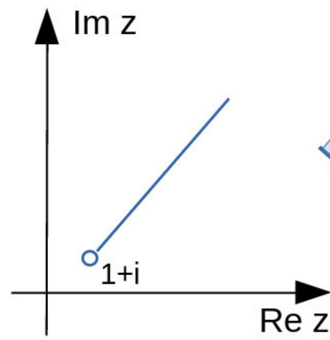
4) Zbiór $\{z \in \mathbb{C}: \ \arg z = \varphi\}$, gdzie $\varphi \in \mathbb{R}$ ustalony, to półprosta.

Przykład 2.4.1.

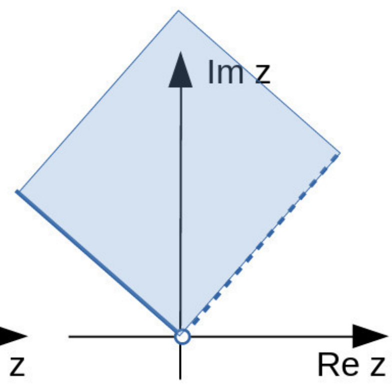
$\arg z = \frac{\pi}{4}$



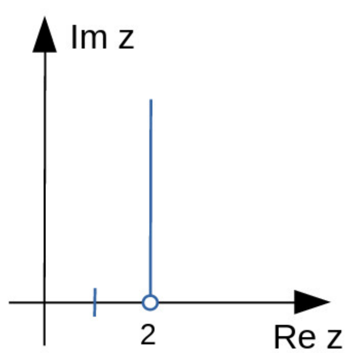
$\arg(z - 1 - i) = \frac{\pi}{4}$



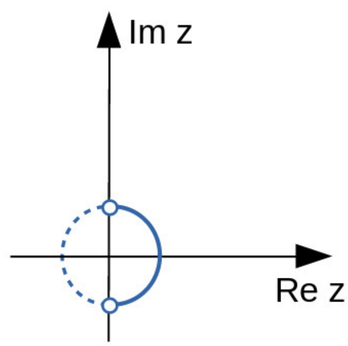
$\frac{\pi}{4} < \arg z \leq \frac{3}{4}\pi$



$\arg(z - 2) = \frac{\pi}{2}$



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



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

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

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

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Metody algebry liniowej w fizyce
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

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

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

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 $A \in M_{m\times n}(K)$, $A = [a_{ij}]$
  $C = A^T$, $C = [c_{ij}] \in M_{n\times m}(K)$, $c_{ij} = a_{ji}$

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.

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

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

$D = AB \in M_2(\mathbb{R})$ schemat Falka
$$\begin{array}{ccc|cc} & & & 0 & 1 \\ & & & 2 & 2 \\ & & & \text{-1} & 1 \\ \hline 1 & 2 & 3 & \color{red}{1} & 8 \\ 4 & 5 & 6 & 4 & 20 \end{array} \qquad \color{red}{1} \color{black}{= 1 \cdot 0 + 2 \cdot 2 + 3 \cdot (-1)}$$

$G = BA \in M_3(\mathbb{R})$
$$\begin{array}{cc|ccc} & & 1 & 2 & 3 \\ & & 4 & 4 & 6 \\ \hline 0 & 1 & 4 & 5 & 6 \\ 2 & 2 & 10 & 14 & 18 \\ \text{-1} & 1 & 3 & 3 & 3 \end{array} \qquad AB \neq BA$$

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

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

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

Twierdzenie 3.1.10. 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 3.1.11. 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 3.1.12. $\mathrm{tr}\begin{bmatrix} 1 & 0 & 0 \\ 2 & -7 & 0 \\ 0 & 1 & 3 \end{bmatrix} = 1 - 7 + 3 = -3$

Własności śladu macierzy

Twierdzenie 3.1.13. 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 3.1.14. 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 3.1.15. Macierz kwadratową $A = [a_{ij}] \in M_n(K)$ nazywamy macierzą:

a) *symetryczną*, gdy $A = A^T$

b) *antysymetryczną*, gdy $A = -A^T$

Twierdzenie 3.1.16. 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 3.1.17. $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 3 & 0 & 7 \end{bmatrix}$ macierz symetryczna

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

Definicja 3.1.18. Macierz utworzoną z macierzy B_{ij}, dla $1 \leq i \leq m, 1 \leq j \leq n$ postaci

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

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

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

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

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

$$B = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 1 & -6 \\ -3 & -6 & 8 \end{bmatrix} \quad \det B = (-1)^{1+1} a_{11} \det B_{11} + (-1)^{1+2} a_{12} \det B_{12} + (-1)^{1+3} a_{13} \det B_{13} =$$

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

Metoda Sarrusa

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

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

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

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

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

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

(rozwinięcie względem i-tego wiersza)

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

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

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

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

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, a ? dowolnymi macierzami odpowiednich wymiarów, to wówczas $\det A = \det B_1 \cdot \det B_2 \cdot \ldots \cdot \det B_n$.

Wniosek 3.2.10. 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
	$\lambda \cdot w_i,\ \lambda \in K, \lambda \neq 0$	pomnożenie wiersza przez liczbę
	$w_i + \lambda \cdot w_j,\ \lambda \in K$	dodanie do w_i wielokrotności w_j

Przykład 3.2.11.

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

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

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

3.3 Macierz odwrotna

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 $A^{-1} = \frac{1}{\det A} \cdot D^T$.

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

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

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

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.

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

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

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

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

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

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) $\left(A^r\right)^{-1} = \left(A^{-1}\right)^r$

Twierdzenie 3.3.11. Jeśli macierz kwadratowa A jest macierzą blokowo-diagonalną postaci $A = \begin{bmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \dots & \dots & \ddots & \dots \\ 0 & 0 & \dots & B_k \end{bmatrix}$, to A jest odwracalna wtedy i tylko wtedy, gdy odwracalne są macierze $B_1, B_2, \dots, B_k$. Wówczas $A^{-1} = \begin{bmatrix} B_1^{-1} & 0 & \dots & 0 \\ 0 & B_2^{-1} & \dots & 0 \\ \dots & \dots & \ddots & \dots \\ 0 & 0 & \dots & B_k^{-1} \end{bmatrix}$.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Metody algebry liniowej w fizyce
Wykład nr 4
Data: ....

TEMAT: *Układy równań liniowych*

4.1 Układy równań 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.

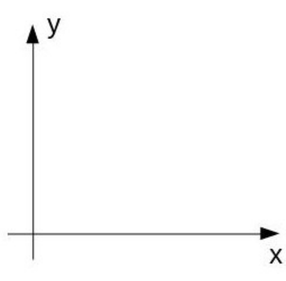
$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2 \\ \ldots & \ldots & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m \end{cases}, \quad a_{ij}, b_i \in K,\ i \in \{1, 2, \ldots, m\}, j \in \{1, 2, \ldots, n\}$$

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

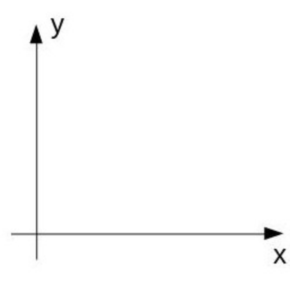
Rozwiązaniem układu nazywamy dowolny ciąg $(\gamma_1, \gamma_2, \ldots, \gamma_n) \in K^n$ spełniający ten układ. Układ nie posiadający rozwiązania nazywamy *układem sprzecznym*, układ posiadający dokładnie jedno rozwiązanie - *układem oznaczonym*, zaś układ posiadający nieskończenie wiele rozwiązań - *układem nieoznaczonym*.

Przykład 4.1.1.

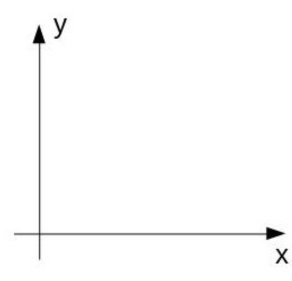
$$\begin{cases} x - y & = & 3 \\ 2x + y & = & 0 \end{cases} \qquad \begin{cases} x + y & = & 5 \\ 2x + 2y & = & 10 \end{cases} \qquad \begin{cases} x + y & = & 0 \\ 2x + 2y & = & 10 \end{cases}$$



układ oznaczony



układ nieoznaczony



układ sprzeczny

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

B, gdzie

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

macierz współczynników — kolumna niewiadomych — kolumna wyrazów wolnych

Przykład 4.1.2.

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

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.

Dowód. $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A^{-1}B = \frac{1}{\det A} D^T B = \frac{1}{\det A}[D_{ji}] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$

Stąd $x_j = \frac{1}{\det A} \sum_{i=1}^{n} b_i D_{ij} = \frac{1}{\det A} \begin{vmatrix} a_{11} & \dots & a_{1,j-1} & b_1 & a_{1,j+1} & \dots & a_{1n} \\ a_{21} & \dots & a_{2,j-1} & b_2 & a_{2,j+1} & \dots & a_{2n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{n,j-1} & b_n & a_{n,j+1} & \dots & a_{nn} \end{vmatrix}.$ $\square$

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

<u>Metoda eliminacji</u>:

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

<u>Rozwiązanie równania macierzowego</u>:

$X=A^{-1}B$, $A^{-1}=?$
$\det A=16$, $D=\begin{bmatrix} 4 & -2 \\ 2 & 3 \end{bmatrix}$, $A^{-1}=\frac{1}{16}\begin{bmatrix} 4 & 2 \\ -2 & 3 \end{bmatrix}$, $X=\frac{1}{16}\begin{bmatrix} 4 & 2 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 5 \\ 1 \end{bmatrix}$
$X=\begin{bmatrix} x \\ y \end{bmatrix}=\frac{1}{16}\begin{bmatrix} 22 \\ -7 \end{bmatrix}$, $\quad x=\frac{11}{8}$, $y=-\frac{7}{16}$

<u>Wzory Cramera</u>:

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

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 Twierdzenie Kroneckera-Capellego

Rząd macierzy

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.

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

Przykład 4.3.3.

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

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

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

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

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

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.

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.

$$A = \begin{bmatrix} 2 & 2 & 2 & 6 & 3 & 0 \\ 0 & 6 & 6 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 9 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 2 & 9 \\ 0 & 1 & -7 \\ 0 & 0 & 0 \\ 0 & 0 & 15 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & 2 & 2 & 1 \\ 0 & 0 & 6 & 1 \\ 0 & 0 & 7 & 0 \end{bmatrix}$$

4 schodki   3 schodki   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. $\square$

Przykład 4.3.8.

$$D = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 5 & 7 & 10 \\ 4 & 6 & 8 & 9 & 3 \\ 4 & 7 & 7 & 8 & 3 \end{bmatrix}$$

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

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.

Przykład 4.3.11.

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

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

$\det A = 0$ oraz $\begin{vmatrix} 1 & -1 \\ 4 & 1 \end{vmatrix} \neq 0$, zatem $r(A) = 2$

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

$r(U) = 3 \neq r(A) \Rightarrow$ 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 4.3.12.

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

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

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

$$\begin{vmatrix} -1 & 2 \\ 0 & 7 \end{vmatrix} \ne 0 \text{ minor niezerowy}$$

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

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

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 4.3.13.

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

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

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

Przykład 4.3.14.

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

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

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

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

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

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

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

Przykład 4.3.16. 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 4.3.17. 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 4.3.18.

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

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

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

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

Przykład 4.3.19. Określ ilość rozwiązań układu w zależności od parametru $p \in \mathbb{R}$.

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

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

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

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.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Metody algebry liniowej w fizyce
Wykład nr 5
Data: ....

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

5.1 Wektory w przestrzeni

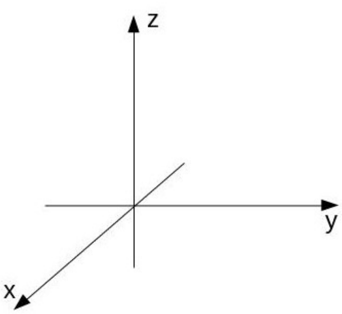
$\mathbb{R}^3 = \{(x, y, z) : x, y, z \in \mathbb{R}\}$ możemy interpretować jako:

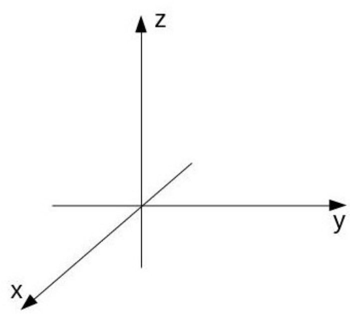
- zbiór punktów $P = (x, y, z)$, gdzie x, y, z to współrzędne punktu
- zbiór wektorów zaczepionych w początku układu współrzędnych $\vec{a} = \overrightarrow{OP} = [x, y, z]$, gdzie x, y, z to współrzędne wektora
- zbiór wektorów swobodnych $\vec{a}$. Wektor swobodny to zbiór wszystkich wektorów zaczepionych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

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

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

$\vec{u} + \vec{v} = [u_x + v_x, u_y + v_y, u_z + v_z]$	suma wektorów
$\lambda \cdot \vec{u} = [\lambda u_x, \lambda u_y, \lambda u_z]$	iloczyn wektora przez skalar
$-\vec{u} = [-u_x, -u_y, -u_z]$	wektor przeciwny do $\vec{u}$
$\vec{u} - \vec{v} = \vec{u} + (-\vec{v}) = \vec{u} + (-1) \cdot \vec{v}$	różnica 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 $\hat{\imath} = [1,0,0], \hat{\jmath} = [0,1,0], \hat{k} = [0,0,1]$ wersory osi układu współrzędnych. Wówczas zapis $\vec{u} = [u_x, u_y, u_z]$ oznacza $\vec{u} = u_x\hat{\imath} + u_y\hat{\jmath} + u_z\hat{k}$. Ponadto $|\vec{u}| = \sqrt{u_x^2 + u_y^2 + u_z^2}$.

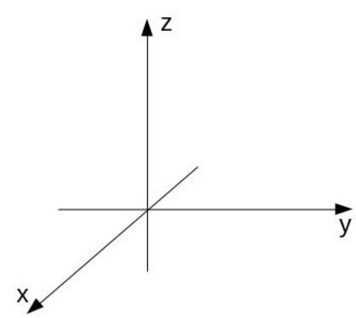
Własności długości: $|\vec{u}| = 0 \Leftrightarrow \vec{u} = \vec{\mathbf{0}}, \quad |\lambda \cdot \vec{u}| = |\lambda| \cdot |\vec{u}|, \quad |\vec{u} + \vec{v}| \leq |\vec{u}| + |\vec{v}|$.

Wersorem niezerowego wektora $\vec{u}$ nazywamy wersor o tym samym kierunku i zwrocie co $\vec{u}$. Oznaczamy go $\hat{u}$. Oczywiście $\hat{u} = \frac{1}{|\vec{u}|} \cdot \vec{u}$.

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 $\cos\alpha = \frac{u_x}{|\vec{u}|}$, $\cos\beta = \frac{u_y}{|\vec{u}|}$, $\cos\gamma = \frac{u_z}{|\vec{u}|}$ nazywamy *cosinusami kierunkowymi* wektora $\vec{u}$.



Iloczyn skalarny

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

Definicja 5.1.1. *Iloczynem skalarnym* dwóch niezerowych wektorów $\vec{u}, \vec{v}$ nazywamy liczbę $|\vec{u}| \cdot |\vec{v}| \cdot \cos\measuredangle(\vec{u}, \vec{v})$. Oznaczamy ją symbolem $\vec{u} \circ \vec{v}$. Gdy jeden z wektorów jest zerowy, przyjmujemy, że iloczyn jest równy 0.

Twierdzenie 5.1.2. 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 5.1.3. $[1,4,0] \circ [2,-1,1] = 1 \cdot 2 + 4 \cdot (-1) + 0 \cdot 1 = -2 < 0$
Cosinus kąta między wektorami jest ujemny, zatem kąt jest rozwarty.

Twierdzenie 5.1.4 (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) $(\vec{u} + \vec{v}) \circ \vec{w} = (\vec{u} \circ \vec{w}) + (\vec{v} \circ \vec{w})$

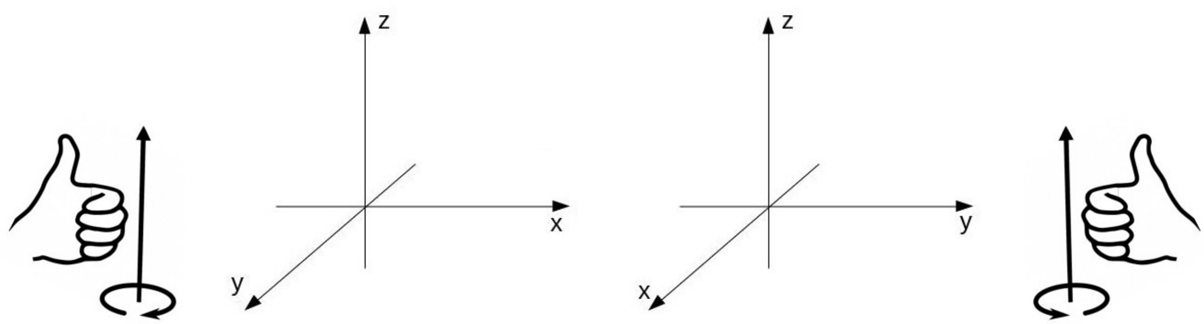
v) $|\vec{u} \circ \vec{v}| \leq |\vec{u}| \cdot |\vec{v}|$

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

Układ współrzędnych

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

układ lewoskrętny układ prawoskrętny



Definicja 5.1.5. Uporządkowana trójka wektorów $(\vec{u}, \vec{v}, \vec{w})$ 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 zawarte są te wektory. Piszemy wówczas $\vec{u} \parallel \vec{v}$.

Definicja 5.1.6. *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 $(\vec{u}, \vec{v}, \vec{w})$ jest zgodna z orientacją układu współrzędnych.

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

Przykład 5.1.7. $\hat{\imath} \times \hat{\imath} = \hat{\jmath} \times \hat{\jmath} = \hat{k} \times \hat{k} = 0$
$\hat{\imath} \times \hat{\jmath} = \hat{k}, \quad \hat{\jmath} \times \hat{k} = \hat{\imath}, \quad \hat{k} \times \hat{\imath} = \hat{\jmath}, \quad \hat{\jmath} \times \hat{\imath} = -\hat{k}, \quad \hat{k} \times \hat{\jmath} = -\hat{\imath}, \quad \hat{\imath} \times \hat{k} = -\hat{\jmath}$

Twierdzenie 5.1.8. Jeśli $\vec{u} = [u_x, u_y, u_z]$, $\vec{v} = [v_x, v_y, v_z]$, to wówczas

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

Przykład 5.1.9. Niech $\vec{u} = [1, 2, -3], \vec{v} = [3, 4, 5]$. Korzystamy z twierdzenia Laplace'a

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

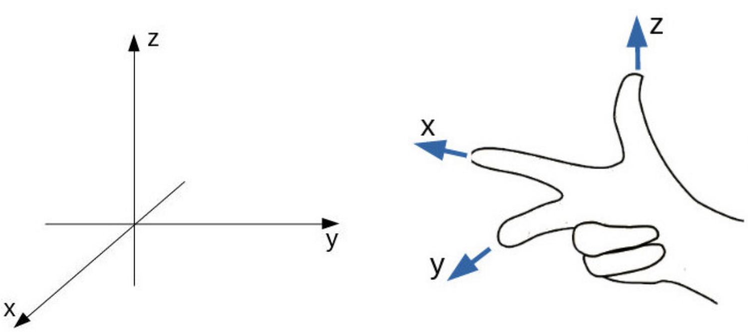
lub metody Sarrusa

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

Twierdzenie 5.1.10 (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) $\vec{u} \times \vec{v} = \vec{\mathbf{0}} \Leftrightarrow \left(\vec{u} \parallel \vec{v} \vee \vec{u} = \vec{\mathbf{0}} \vee \vec{v} = \vec{\mathbf{0}}\right)$

Reguła prawej dłoni:



Uwaga 5.1.11. Pole równoległoboku rozpiętego na wektorach $\vec{u}, \vec{v}$ równe jest $|\vec{u} \times \vec{v}|$.

Iloczyn mieszany

Definicja 5.1.12. 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 $(\vec{u}\times\vec{v})\circ\vec{w}$. Oznaczamy ją symbolem $\big(\vec{u}, \vec{v}, \vec{w}\big)$.

Twierdzenie 5.1.13. 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

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

Uwaga 5.1.14. Objętość równoległościanu rozpiętego na wektorach $\vec{u}, \vec{v}, \vec{w}$ równa jest $|\big(\vec{u}, \vec{v}, \vec{w}\big)|$.

Dowód. Niech $\alpha = \measuredangle(\vec{u}\times\vec{v}, \vec{w})$. Ponieważ $V = P_p \cdot d = |\vec{u}\times\vec{v}| \cdot h$ 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 5.1.15. 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.

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

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

Punkty są współpłaszczyznowe (komplanarne).

Twierdzenie 5.1.16 (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) $\big(\vec{u}, \vec{v}, \vec{w}\big) = -\big(\vec{v}, \vec{u}, \vec{w}\big) = \big(\vec{v}, \vec{w}, \vec{u}\big)$

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

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

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

5.2 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 $\vec{n} = [A, B, C] \neq \vec{\mathbf{0}}$ ustalonym wektorem. Wówczas zbiór

$$\pi = \{P = (x, y, z) \in \mathbb{R}^3 : \overrightarrow{P_0P} \perp \vec{n}\}$$

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: $\pi : A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$
Równanie ogólne: $\pi : Ax + By + Cz + D = 0,\ D = -Ax_0 - By_0 - Cz_0$

Przykład 5.2.1. $P_0 = (1, 2, 5), \vec{n} = [1, -1, 3]$, $P_0 \in \pi$, $\vec{n} \perp \pi$, $\pi = ?$

$\pi : 1 \cdot (x - 1) + (-1) \cdot (y - 2) + 3 \cdot (z - 3) = 0$ równanie normalne

$\pi : x - y + 3z - 14 = 0$ równanie ogólne

Równanie parametryczne płaszczyzny

Niech $P_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$ będzie ustalonym punktem, zaś $\vec{a} = [a_x, a_y, a_z] \neq \vec{\mathbf{0}}$, $\vec{b} = [b_x, b_y, b_z] \neq \vec{\mathbf{0}}$ ustalonymi wektorami niewspółliniowymi, tj. $\vec{a} \nparallel \vec{b}$. Wówczas zbiór

$$\pi = \{P = (x, y, z) \in \mathbb{R}^3 : \quad \exists t, s \in \mathbb{R} \quad \overrightarrow{P_0P} = t\vec{a} + s\vec{b}\}$$

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: $\pi : \begin{cases} x = x_0 + t \cdot a_x + s \cdot b_x \\ y = y_0 + t \cdot a_y + s \cdot b_y \\ z = z_0 + t \cdot a_z + s \cdot b_z \end{cases}, t, s \in \mathbb{R}$

Przykład 5.2.2. Napisz równanie parametryczne płaszczyzny π przechodzącej przez punkty $A = (1, 1, 4)$, $B = (2, 5, 4)$ i równoległej do osi Oy.

$$\overrightarrow{AB} = [1, 4, 0] \parallel \pi, \quad \hat{j} = [0, 1, 0] \parallel Oy \Rightarrow \hat{j} \parallel \pi, \quad A \in \pi$$

$$\pi : \begin{cases} x = 1 + 1 \cdot t + 0 \cdot s &= 1 + t \\ y = 1 + 4 \cdot t + 1 \cdot s &= 1 + 4t + s \\ z = 4 + 0 \cdot t + 0 \cdot s &= 4 \end{cases}, \ t, s \in \mathbb{R}$$

Inne równania płaszczyzny

Równanie postaci $\pi: \frac{x}{a}+\frac{y}{b}+\frac{z}{c}=1, \quad gdzie \quad a,b,c \in \mathbb{R}\setminus\{0\}$, opisuje płaszczyznę przechodząca przez punkty $(a,0,0),(0,b,0),(0,0,c)$. Jest to tzw. *równanie odcinkowe* płaszczyzny.

Równanie płaszczyzny przechodzącej przez trzy niewspółliniowe punkty $P_1=(x_1,y_1,z_1), P_2=(x_2,y_2,z_2), P_3=(x_3,y_3,z_3)$ ma postać

$$\pi: \begin{vmatrix} x-x_1 & y-y_1 & z-z_1 \\ x_2-x_1 & y_2-y_1 & z_2-z_1 \\ x_3-x_1 & y_3-y_1 & z_3-z_1 \end{vmatrix} = 0.$$

Istotnie, ponieważ $\overrightarrow{P_1P_2}=[x_2-x_1,y_2-y_1,z_2-z_1]$, $\overrightarrow{P_1P_3}=[x_3-x_1,y_3-y_1,z_3-z_1] \parallel \pi$ oraz $\vec{n}=\overrightarrow{P_1P_2}\times\overrightarrow{P_1P_3} \perp \pi$, zatem $\pi=\{P=(x,y,z): \overrightarrow{P_1P} \perp \vec{n}\}=\{P=(x,y,z): [x-x_1,y-y_1,z-z_1]\circ\vec{n}=0\}$.

5.3 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 $\vec{a}=[a_x,a_y,a_z]\neq\vec{\mathbf{0}}$ ustalonym wektorem. Wówczas zbiór

$$l=\{P=(x,y,z)\in\mathbb{R}^3:\ \overrightarrow{P_0P}\parallel\vec{a}\}$$

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 $l: \begin{cases} x=x_0+t\cdot a_x \\ y=y_0+t\cdot a_y \\ z=z_0+t\cdot a_z \end{cases}$, $t\in\mathbb{R}$ nazywamy *równaniem parametrycznym* prostej l.

Rugując z każdego z powyższych równań parametr t otrzymujemy równanie postaci $l: \frac{x-x_0}{a_x}=\frac{y-y_0}{a_y}=\frac{z-z_0}{a_z}$, które nazywamy *równaniem kierunkowym* prostej l.

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 $A_1^2+B_1^2+C_1^2\neq0$, $A_2^2+B_2^2+C_2^2\neq0$ będą dwiema nierównoległymi płaszczyznami. Ich częścią wspólną jest prosta $l=\pi_1\cap\pi_2$.

$$P\in l \quad\Leftrightarrow\quad (P\in\pi_1 \wedge P\in\pi_2)$$

Równanie krawędziowe: $l: \begin{cases} A_1x+B_1y+C_1z+D_1=0 \\ A_2x+B_2y+C_2z+D_2=0 \end{cases}$.

Przykład 5.3.1. Napisz 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 $l : \begin{cases} x = 2 - t \\ y = 3 + 13t \\ z = 1 + 19t \end{cases}, \ t \in \mathbb{R}$.

5.4 Wzajemne położenie płaszczyzn i prostych

Wzajemne położenie płaszczyzn

Niech $\pi_1 : \ A_1x + B_1y + C_1z + D_1 = 0$, $\vec{n_1} = [A_1, B_1, C_1] \neq \vec{\mathbf{0}}$,
$\pi_2 : \ A_2x + B_2y + C_2z + D_2 = 0$, $\vec{n_2} = [A_2, B_2, C_2] \neq \vec{\mathbf{0}}$.

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 $A = \begin{bmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{bmatrix}$, $U = \begin{bmatrix} A_1 & B_1 & C_1 & -D_1 \\ A_2 & B_2 & C_2 & -D_2 \end{bmatrix}$.

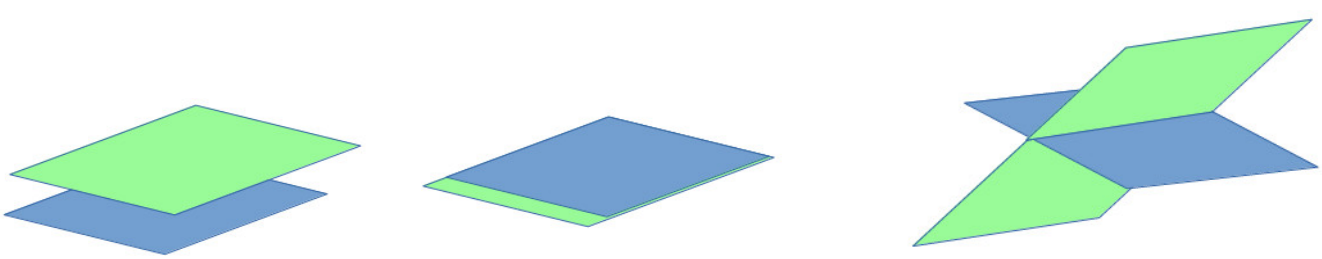
Płaszczyzny mogą być równoległe. $\quad \pi_1 \| \pi_2 \ \Leftrightarrow \ n_1 \| n_2 \ \Leftrightarrow \ \vec{n_1} \times \vec{n_2} = \vec{\mathbf{0}}$

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.

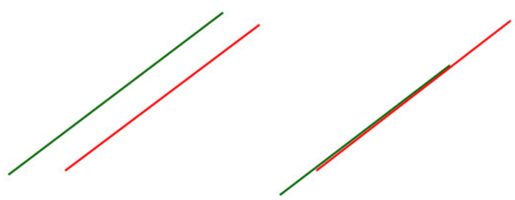
$$\pi_1 \perp \pi_1 \ \Leftrightarrow \ n_1 \perp n_2 \ \Leftrightarrow \ \vec{n_1} \circ \vec{n_2} = 0$$



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.

Proste mogą być równoległe. $\quad l \parallel k \Leftrightarrow \vec{a} \parallel \vec{b} \Leftrightarrow \vec{a} \times \vec{b} = \vec{\mathbf{0}}$.
Wówczas albo $l = k$, albo $l \cap k = \varnothing$.

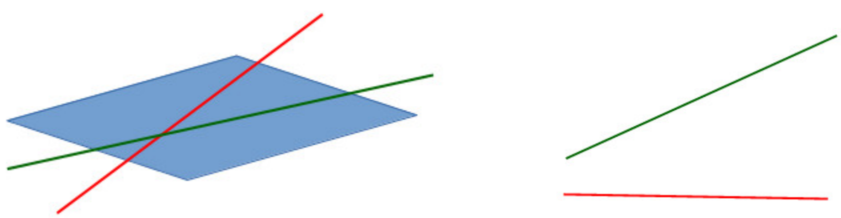


Gdy $l \nparallel k$, możliwe są dwie sytuacje.

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 $l \cap k = \varnothing$.

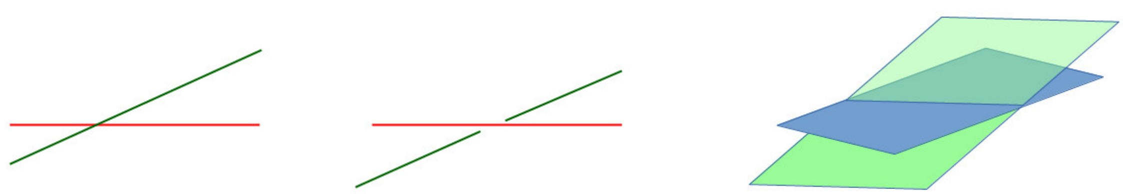
Zatem proste l i k są skośne wtedy i tylko wtedy gdy $\left(\overrightarrow{P_1P_2}, \vec{a}, \vec{b}\right) \neq 0$.



Kąty

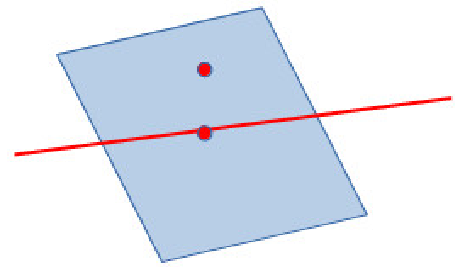
Definicja 5.4.1. i) *Kątem między dwiema prostymi* nazywamy kąt ostry (lub prosty, gdy proste są prostopadłe) między odpowiednio zwróconymi wektorami kierunkowymi tychże prostych.

ii) *Kątem między dwiema płaszczyznami* nazywamy kąt ostry (lub prosty, gdy płaszczyzny są prostopadłe) między odpowiednio zwróconymi wektorami normalnymi tychże płaszczyzn.



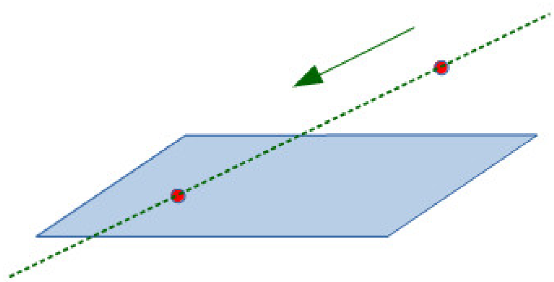
Definicja 5.4.2. i) *Rzutem prostokątnym punktu P na płaszczyznę π* nazywamy punkt $P' \in \pi$ taki, że $PP' \perp \pi$.

ii) *Rzutem prostokątnym punktu P na prostą l* nazywamy punkt $P' \in l$ taki, że $PP' \perp l$.

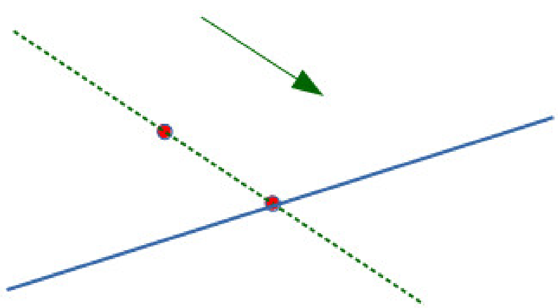


Można zdefiniować *rzut ukośny* w kierunku zadanego wektora.

Rzut punktu P na płaszczyznę π w kierunku wektora $\vec{v} \nparallel \pi$:

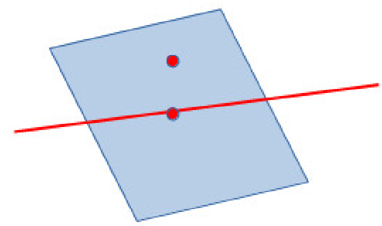


Rzut punktu P na prostą l w kierunku wektora $\vec{v}$, o którym zakładamy, że należy do płaszczyzny zawierającej P oraz l:



Przykład 5.4.3. 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 $k \perp \pi$, $P \in k$.

$\vec{u} = [2, 0, 1] \parallel \pi, \quad \vec{v} = [1, 3, 1] \parallel \pi, \quad A = (2, 1, 3) \in \pi$

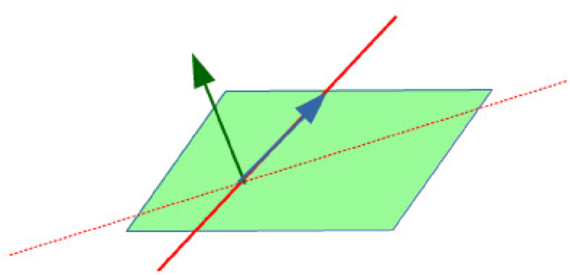
$$\vec{n} = \vec{u} \times \vec{v} \perp \pi, \quad \vec{n} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 2 & 0 & 1 \\ 1 & 3 & 1 \end{vmatrix} = [-3, -1, 6], \quad k \perp \pi \Rightarrow k \parallel \vec{n}$$

$$k : \begin{cases} x = 4 - 3t \\ y = 5 - t \\ z = -3 + 6t \end{cases}, \ t \in \mathbb{R}, \quad \pi : \ -3(x-2) - (y-1) + 6(z-3) = 0$$

$\pi : \ 3x + y - 6z + 11 = 0 \quad \{P'\} = k \cap \pi = ?$

$3(4 - 3t) + 5 - t - 6(-3 + 6t) = 0 \Rightarrow t = 1, \quad P' = (1, 4, 3)$

Definicja 5.4.4. *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.



Przykład 5.4.5. 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 $\vec{a} = [-1, 0, 2] \parallel l, \quad \vec{n} = [3, 1, 1] \perp \pi. \quad$ Oznaczmy $\beta = \measuredangle(\vec{n}, \vec{a})$, $\alpha = \frac{\pi}{2} - \beta$.

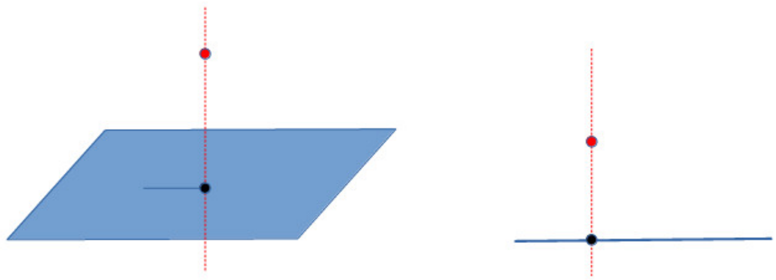
Obliczamy $\cos\beta = \frac{|\vec{n} \circ \vec{a}|}{|\vec{n}| \cdot |\vec{a}|} = \frac{|-3+0+2|}{\sqrt{9+1+1} \cdot \sqrt{1+4}} = \frac{1}{\sqrt{55}}, \quad \alpha = \frac{\pi}{2} - \arccos\frac{1}{\sqrt{55}}$,

albo $\sin\alpha = \cos\beta \Rightarrow \alpha = \arcsin\frac{1}{\sqrt{55}}$

Odległości

Definicja 5.4.6. 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 $d(P, l)$.



Wzór na odległość punktu od płaszczyzny

Niech $P_0 = (x_0, y_0, z_0)$ oraz $\pi: \; Ax + By + Cz + D = 0$, $\vec{n} = [A, B, C] \neq \vec{\mathbf{0}}$. 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 $k: \begin{cases} x = x_0 + At \\ y = y_0 + Bt \\ z = z_0 + Ct \end{cases}$

$\{P_0'\} = k \cap \pi = ?$

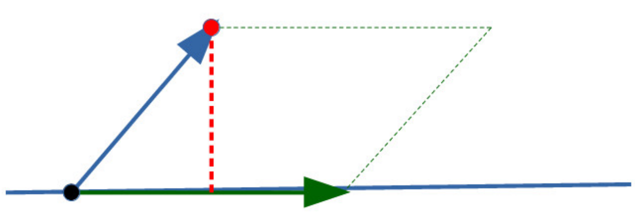
$A(x_0 + At) + B(y_0 + Bt) + C(z_0 + Ct) + D = 0, \quad t = -\frac{Ax_0 + By_0 + Cz_0 + D}{A^2 + B^2 + C^2}$

$d(P, \pi) = |P_0P_0'| = \sqrt{(At)^2 + (Bt)^2 + (Ct)^2} = |t|\sqrt{A^2 + B^2 + C^2} = \frac{|Ax_0 + By_0 + Cz_0 + D|}{\sqrt{A^2 + B^2 + C^2}}$

Wzór na odległość punktu od prostej

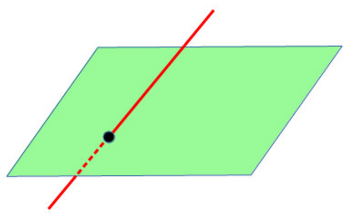
Niech $P_0 = (x_0, y_0, z_0)$ oraz niech l będzie prostą przechodząca przez punkt P_1 o wektorze kierunkowym $\vec{a}$. Wówczas

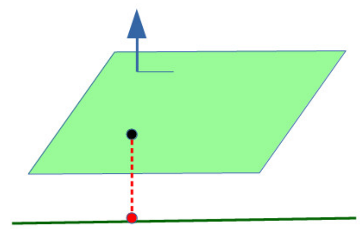
$$d(P_0, l) = \frac{|\overrightarrow{P_1P_0} \times \vec{a}|}{|\vec{a}|}.$$



Odległość prostej od płaszczyzny

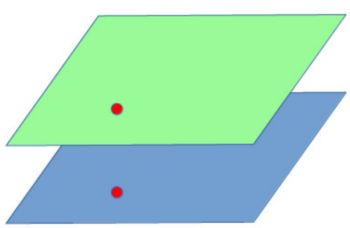
Jeśli prosta l nie przecina płaszczyzny π, to wówczas odległością prostej l od płaszczyzny π nazywamy odległość dowolnego punkty prostej od płaszczyzny.





Definicja 5.4.7. *Odległością dwóch płaszczyzn (prostych) równoległych* nazywamy odległość dowolnego punktu jednej z nich od drugiej.

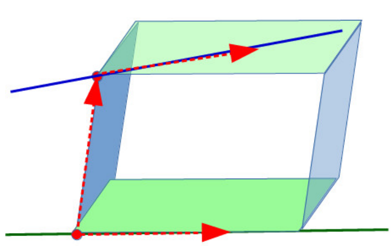
Można wykazać, że dla płaszczyzn równoległych $\pi_1 : \ Ax + By + Cz + D_1 = 0$, $\pi_2 : \ Ax + By + Cz + D_2 = 0$, $\vec{n_2} = [A, B, C] \neq \vec{\mathbf{0}}$ zachodzi wzór $d(\pi_1, \pi_2) = \frac{|D_2 - D_1|}{|\vec{n}|}$.



Definicja 5.4.8. *Odległością dwóch prostych skośnych* nazywamy odległość dwóch płaszczyzn równoległych zawierających te proste.

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

$$d(k, l) = \frac{|\left(\overrightarrow{P_1 P_2}, \vec{a}, \vec{b}\right)|}{|\vec{a} \times \vec{b}|}.$$



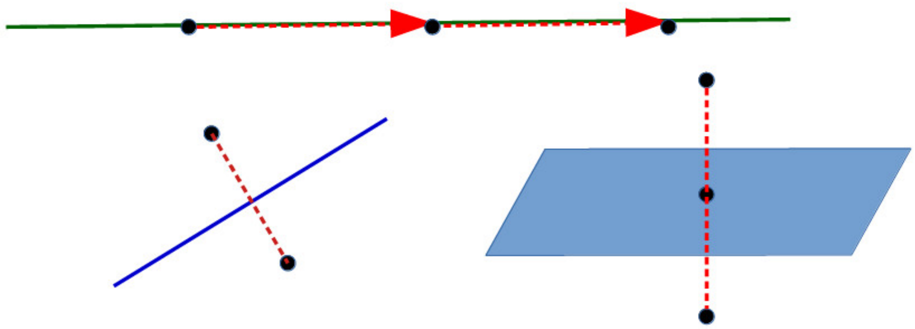
Symetrie

Definicja 5.4.9. 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 $\overrightarrow{PA} \perp \pi$.



Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Metody algebry liniowej w fizyce
Wykład nr 6
Data: ....

TEMAT: *Przestrzenie liniowe*

6.1 Przestrzenie liniowe i ich podprzestrzenie

Niech $K = (K, +, \cdot)$ będzie ciałem, 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 $u = [u_1, \dots, u_n], v = [v_1, \dots, v_n] \in \mathbb{R}^n$ oraz $\alpha \in \mathbb{R}$ definiujemy
$u \oplus v = [u_1 + v_1, \dots, u_n + v_n]$, $\alpha \odot u = [\alpha u_1, \dots, \alpha u_n]$.

ii) $(K^n, \oplus, K, \odot)$, gdzie $K = (K, +, \cdot)$ to dowolne ciało
Dla $u = [u_1, \dots, u_n], v = [v_1, \dots, v_n] \in K^n$ oraz $\alpha \in K$ definiujemy
$u \oplus v = [u_1 + v_1, \dots, u_n + v_n]$, $\alpha \odot u = [\alpha \cdot u_1, \dots, \alpha \cdot u_n]$.

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 $V = (C, +, K, \cdot)$.

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

$$\forall u_1, u_2 \in U \quad \forall \alpha, \beta \in K \quad (\alpha \odot u_1) \oplus (\beta \odot u_2) \in U.$$

Dowód. Implikacja z prawa na lewo jest oczywista. Implikacja w druga stronę wynika z faktu, że U jest podgrupą grupy V. $\square$

Uwaga 6.1.7. Niech $V = (V, \oplus, K, \odot)$ będzie przestrzenią liniową. Wówczas $U = \{\mathbf{0}\}$ jest podprzestrzenią liniową. Nazywamy ją *podprzestrzenią trywialną*. Podobnie $U = V$ jest podprzestrzenią liniową V. Podprzestrzenie te nazywamy *podprzestrzeniami niewłaściwymi.*

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) $V = (\mathbb{R}^{[0,1]}, +, \mathbb{R}, \cdot)$, $U = (\mathcal{C}([0,1], \mathbb{R}), +, \mathbb{R}, \cdot)$

U jest podprzestrzenią liniową V, bowiem suma funkcji ciągłych jest funkcją ciągłą oraz iloczyn funkcji ciągłej przez liczbę jest funkcja ciągłą.

ii) $V = (\mathbb{R}[x], +, \mathbb{R}, \cdot)$, $U = \{f \in \mathbb{R}[x] : \deg f \ - \ parzysty\}$

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 $f + g \notin U$.

iii) $V = (\mathbb{R}^4, +, \mathbb{R}, \cdot)$, $U = \{(x, y, z, t) \in \mathbb{R}^4 : \ 2x + z - 3t = 0 \wedge y = 0\}$

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 $\alpha u_1 = (\alpha x_1, 0, \alpha(3t_1 - 2x_1), \alpha t_1) = (\alpha x_1, 0, 3\alpha t_1 - 2\alpha x_1, \alpha t_1) \in U$

iv) $V = (\mathbb{R}[x], +, \mathbb{R}, \cdot)$, $U = \mathbb{R}_n[x] := \{p \in \mathbb{R}[x] : \ \deg p \leq n\}$.
Przyjmujemy, że $\deg \mathbf{0} = -\infty$. Wówczas U jest podprzestrzenią liniową V.

Podprzestrzenie wektorowe $\mathbb{R}^2$ i $\mathbb{R}^3$

Jedynymi nietrywialnymi podprzestrzeniami liniowymi $\mathbb{R}^2$ są proste przechodzące przez $(0, 0)$.
Jedynymi nietrywialnymi podprzestrzeniami liniowymi $\mathbb{R}^3$ są płaszczyzny i proste przechodzące przez $(0, 0)$.

6.2 Liniowa niezależność wektorów

Niech $V = (V, +, K, \cdot)$ będzie przestrzenią liniową nad ciałem K. Niech $\alpha_1, \ldots \alpha_m \in K$, $v_1, \ldots, v_m \in V$. Niech $W \neq \varnothing$ będzie podzbiorem zbioru V.

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 $\{v = \alpha_1 w_1 + \ldots + \alpha_k w_k \,:\, \alpha_1, \ldots \alpha_k \in K;\, w_1, \ldots, w_k \in W;\, k \in \mathbb{N}\}$, 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 $\mathrm{lin}\{w_1, \ldots, w_m\} = \{\alpha_1 w_1 + \ldots + \alpha_m w_m : \alpha_1, \ldots \alpha_m \in K\}$.

Twierdzenie 6.2.2. 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) podprzestrzeń V zawierająca zbiór W.

Wniosek 6.2.3. 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 $\mathrm{lin}\{v_1, v_2, \ldots, v_m\} = \mathrm{lin}\{u, v_2, \ldots, v_m\}$

Definicja 6.2.4. Elementy zbioru W nazywamy *generatorami* przestrzeni $\mathrm{lin} W$, zaś podprzestrzeń $\mathrm{lin} W$ nazywamy podprzestrzenią *generowaną* przez zbiór W.

Przykład 6.2.5. Wersory $\hat{i} = (1, 0)$ oraz $\hat{j} = (0, 1)$ generują przestrzeń $\mathbb{R}^2$, bowiem dla dowolnego $\vec{u} = (u_x, u_y) \in \mathbb{R}^2$ mamy $\vec{u} = u_x(1, 0) + u_y(0, 1) = u_x \hat{i} + u_y \hat{j}$.

Niech $V = (V, +, K, \cdot)$ będzie przestrzenią liniową nad ciałem K.

Definicja 6.2.6. 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.*

Twierdzenie 6.2.7. 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.

Twierdzenie 6.2.8. 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.

Wniosek 6.2.9. Jeśli wektory $v_1, \ldots, v_k \in \mathbb{R}^n$ generują $\mathbb{R}^n$, to $k \geq n$.

Przykład 6.2.10. Czy wektory $u = (1,1,-1)$, $v = (2,1,0)$, $w = (5,2,2)$ generują przestrzeń $\mathbb{R}^3$?

Sprawdzamy, czy dla dowolnego $b = (x,y,z) \in \mathbb{R}^3$ istnieją $\alpha, \beta, \gamma \in \mathbb{R}$ takie, że $b = \alpha u + \beta v + \gamma w$.

$(x,y,z) = \alpha(1,1,-1) + \beta(2,10) + \gamma(5,2,2) = (\alpha + 2\beta + 5\gamma, \alpha + \beta + 2\gamma, -\alpha + 2\gamma)$

$$\begin{bmatrix} 1 & 2 & 5 \\ 1 & 1 & 2 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

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 $\mathbb{R}^3$.

Przykład 6.2.11. 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 6.2.12. 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.

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 6.2.13. 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 $k \leq n$.

Twierdzenie 6.2.14. 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.

Definicja 6.2.15. Niech $f_1, \ldots, f_n \in \mathcal{C}^{n-1}(\mathbb{R})$, $n \geq 2$. Macierz $W(x)$ postaci

$$W_{f_1,\ldots,f_n}(x) = \begin{bmatrix} f_1(x) & f_2(x) & \ldots & f_n(x) \\ f_1'(x) & f_2'(x) & \ldots & f_n'(x) \\ \ldots & & \ldots & \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \ldots & f_n^{(n-1)}(x) \end{bmatrix}$$

nazywamy *macierzą Wrońskiego* układu funkcji $f_1, \ldots, f_n$, a jej wyznacznik *wrońskianem.*

Twierdzenie 6.2.16. Niech $f_1, \ldots, f_n \in \mathcal{C}^{n-1}(\mathbb{R})$, $n \geq 2$. Jeśli wrońskian układu funkcji $f_1, \ldots, f_n$ nie zeruje się tożsamościowo na $\mathbb{R}$, tzn. $\exists x_0 \in \mathbb{R} : \det W_{f_1,\ldots,f_n}(x_0) \neq 0$, to funkcje $f_1, \ldots, f_n$ są liniowo niezależne w przestrzeni $\mathcal{C}(\mathbb{R})$.

Przykład 6.2.17. Zbadaj, czy funkcje $f_1(x) = 1$, $f_2(x) = \sin x$, $f_3(x) = \cos x$ tworzą układ liniowo niezależny w $\mathcal{C}(\mathbb{R})$.

$$\det W(x) = \begin{vmatrix} 1 & \sin x & \cos x \\ 0 & \cos x & -\sin x \\ 0 & -\sin x & -\cos x \end{vmatrix} = -\cos^2 x - \sin^2 x = -1 \neq 0$$

Tak, tworzą układ liniowo niezależny.

6.3 Baza i wymiar przestrzeni liniowej

Niech $V = (V, +, K, \cdot)$ będzie przestrzenią liniową nad ciałem K, zaś $b_1, \ldots, b_n \in V$.

Definicja 6.3.1. Układ wektorów $\mathcal{B} = \{b_1, \ldots, b_n\}$ nazywamy *bazą* przestrzeni V jeśli jest on liniowo niezależny oraz $V = \text{lin}\{b_1, \ldots, b_n\}$.

Uwaga 6.3.2. Baza przestrzeni wektorowej jest maksymalnym (w sensie relacji inkluzji) układem liniowo niezależnym w tej przestrzeni.

Przykład 6.3.3. Baza przestrzeni $\mathbb{R}^n$

Układ wektorów $\{e_1, \ldots, e_n\}$ stanowi bazę przestrzeni $\mathbb{R}^n$.

$$e_1 = (1, 0, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$$

Wniosek 6.3.4. 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 6.2.9 oraz 6.2.13. □

Przykład 6.3.5. 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 6.2.3).

Twierdzenie 6.3.6. i) Każda przestrzeń wektorowa różna od $\{\mathbf{0}\}$ posiada bazę.

ii) Wszystkie bazy danej przestrzeni wektorowej skończenie wymiarowej 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.

Definicja 6.3.7. 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 6.3.8. 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.

Przykład 6.2.10 - 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 $\mathbb{R}^3$.

Wniosek 6.3.9. 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 6.2.8 oraz 6.3.8 iii). $\square$

Wniosek 6.3.10. 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 6.3.11. 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) $(\mathbb{R}^n, +, \mathbb{R}, \cdot)$

$$\mathcal{B}_k^n = \Big(e_1, \ldots, e_n\Big), \text{ gdzie } e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$$

2) $(\mathbb{R}[x], +, \mathbb{R}, \cdot)$

$$\mathcal{B}_k^n = \Big(1, x, x^2, x^3, \ldots\Big), \quad \dim \mathbb{R}[x] = \infty, \quad p(x) = a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2 + \ldots + a_s \cdot x^s$$

3) $(\mathbb{R}_n[x], +, \mathbb{R}, \cdot)$, $\quad \mathbb{R}_n[x] = \{f \in \mathbb{R}[x] : \deg f \leq n\}$

$$\mathcal{B}_k^n = \Big(1, x, x^2, \ldots, x^n\Big), \quad \dim \mathbb{R}[x] = n + 1$$
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) $(M_{m\times n}(\mathbb{R}), +, \mathbb{R}, \cdot)$

$$\mathcal{B}_k^n = \Big(E_{11}, E_{12}, \ldots, E_{1n}, E_{21}, \ldots, E_{mn}\Big), \text{ gdzie } E_{kl} = [e_{ij}^{kl}], \text{ zaś } e_{ij}^{kl} = \begin{cases} 1 & (i, j) = (k, l) \\ 0 & (i, j) \neq (kl) \end{cases}$$
$\dim M_{m\times n}(K) = m \cdot n$

Przykład 6.3.12. 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 6.3.13. Niech V będzie przestrzenią liniową, zaś W jej podprzestrzenią. Wówczas

i) $W \neq V \Rightarrow \dim W < \dim V$,

ii) $\dim W = \dim V < \infty \Rightarrow W = V$.

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 6.3.14. 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 $v = [\alpha_1, \ldots \alpha_n]_{\mathcal{B}}$.

Przykład 6.3.15. $\mathbb{R}^3 = (\mathbb{R}^3, +, \mathbb{R}, \cdot)$, $\mathcal{B} = \mathcal{B}^3_k$ - 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.
UMOWA: Piszemy $b'_1 = (4, 2, 1)$ zamiast $b'_1 = [4, 2, 1]_{\mathcal{B}^3_k}$.
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

$$\begin{bmatrix} 4 & -5 & 1 \\ 2 & 2 & 3 \\ 1 & 3 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} -3 \\ 15 \\ 7 \end{bmatrix}.$$

$$\left[\begin{array}{ccc|c} 4 & -5 & 1 & -3 \\ 2 & 2 & 3 & 15 \\ 1 & 3 & 0 & 7 \end{array}\right] \xrightarrow{w_3 \leftrightarrow w_1} \left[\begin{array}{ccc|c} 1 & 3 & 0 & 7 \\ 2 & 2 & 3 & 15 \\ 4 & -5 & 1 & -3 \end{array}\right] \xrightarrow[w_3 - 4w_1]{w_2 - 2w_1} \left[\begin{array}{ccc|c} 1 & 3 & 0 & 7 \\ 0 & -4 & 3 & 1 \\ 0 & -17 & 1 & -31 \end{array}\right] \xrightarrow{-4 \cdot w_2} \left[\begin{array}{ccc|c} 1 & 3 & 0 & 7 \\ 0 & 16 & -12 & -4 \\ 0 & -17 & 1 & -31 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 7 \\ 0 & 16 & -12 & -4 \\ 0 & -1 & -11 & -35 \end{array}\right] \xrightarrow[w_3 \leftrightarrow w_2]{-1 \cdot w_2} \left[\begin{array}{ccc|c} 1 & 3 & 0 & 7 \\ 0 & 1 & 11 & 35 \\ 0 & 16 & -12 & -4 \end{array}\right] \xrightarrow{w_3 - 16w_2} \left[\begin{array}{ccc|c} 1 & 3 & 0 & 7 \\ 0 & 1 & 11 & 35 \\ 0 & 0 & -188 & -564 \end{array}\right] \xrightarrow{-\frac{1}{188} \cdot w_3}$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 7 \\ 0 & 1 & 11 & 35 \\ 0 & 0 & 1 & 3 \end{array}\right] \xrightarrow{w_2 - 11w_3} \left[\begin{array}{ccc|c} 1 & 3 & 0 & 7 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right] \xrightarrow{w_1 - 3w_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right] \Rightarrow v = [1, 2, 3]_{\mathcal{B}'}$$

Niech V będzie przestrzenią liniową, zaś $\mathcal{B} = (b_1, \dots, b_n)$ oraz $\mathcal{B}' = (b'_1, \dots, b'_n)$ jej dwiema bazami. Wówczas istnieją skalary $\alpha_{ij} \in K$, $i, j \in \{1, 2, \dots, n\}$ takie, że

$$\begin{array}{l} b'_1 = \alpha_{11}b_1 + \alpha_{21}b_2 + \ldots + \alpha_{n1}b_n \\ b'_2 = \alpha_{12}b_1 + \alpha_{22}b_2 + \ldots + \alpha_{n2}b_n \\ \ldots \\ b'_n = \alpha_{1n}b_1 + \alpha_{2n}b_2 + \ldots + \alpha_{nn}b_n \end{array}.$$

Definicja 6.3.16. 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* 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 $\mathcal{B} = (b_1, \dots, b_n)$ oraz $\mathcal{B}' = (b'_1, \dots, b'_n)$ jej dwiema bazami. Niech $P = P_{\mathcal{B}\to\mathcal{B}'}$.

Twierdzenie 6.3.17. Niech $v \in V$, $v = [x_1, \dots, x_n]_{\mathcal{B}} = [x'_1, \dots, x'_n]_{\mathcal{B}'}$. 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 6.3.15 - ciąg dalszy
$\mathbb{R}^3 = (\mathbb{R}^3, +, \mathbb{R}, \cdot)$, $\mathcal{B} = \mathcal{B}^3_k$ - 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 $\mathcal{B}'$.
$P = P_{\mathcal{B}^3_k\to\mathcal{B}'} = \begin{bmatrix} 4 & -5 & 1 \\ 2 & 2 & 3 \\ 1 & 3 & 0 \end{bmatrix}$, $X' = P^{-1}X =?$.
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 6.3.18. 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 6.3.19. 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) $P_{\mathcal{B}\to\mathcal{B}'} \cdot P_{\mathcal{B}'\to\mathcal{B}''} = P_{\mathcal{B}\to\mathcal{B}''}$.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Metody algebry liniowej w fizyce
Wykład nr 7
Data: ....

TEMAT: *Przekształcenia liniowe*

7.1 Odwzorowania liniowe i ich 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.*

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 <u>nie</u> 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.

4) $\varphi : \mathbb{R}^3 \to \mathbb{R}^2$, $\varphi(x,y,z) = (x - y + z, 2y + z + 1)$

Odwzorowanie <u>nie</u> 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 $\varphi(x,y) = (ax + by, cx + dy)$.

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 $\varphi(x,y,z) = (ax + by + cz, dx + ey + fz, gx + hy + jz)$.

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.

Przykład 7.1.9. Odwzorowanie $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 2x + 5$ nie jest liniowe, bowiem $f(0) = 5 \neq 0$.

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 przestrzeni V.

Przykład 7.1.13. Podaj wzór odwzorowania liniowego φ, jeśli

$$\varphi : \mathbb{R}_2[x] \to \mathbb{R}_1[x], \quad \varphi(x^2 + x) = 6x + 10, \quad \varphi(x - 1) = 4, \quad \varphi(2x) = 8.$$

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 $\varphi(V)$.

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 $\mathbf{0}_W \in \mathrm{Im}\varphi$. □

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

$$r(\varphi) + \dim \mathrm{Ker}\varphi = \dim V.$$

Wniosek 7.2.7. Dla dowolnego $\varphi \in \mathcal{L}(V, W)$ mamy $\dim \mathrm{Im}\varphi \leq \dim V$.

Przykład 7.2.8.

$$\varphi : \mathbb{R}^4 \to \mathbb{R}^3, \quad \varphi(x, y, z, t) = (x + y + z + 2t, x - y + z + 6t, x + y - z - 4t)$$

Wyznacz jądro oraz obraz φ, ich bazy i wymiary. Podaj własności φ.

$$\varphi(x, y, z, t) = (0, 0, 0) \Leftrightarrow \begin{cases} x + y + z + 2t = 0 \\ x - y + z + 6t = 0 \\ x + y - z - 4t = 0 \end{cases}$$

$$\left[\begin{array}{rrrr|r} 1 & 1 & 1 & 2 & 0 \\ 1 & -1 & 1 & 6 & 0 \\ 1 & 1 & -1 & -4 & 0 \end{array}\right] \xrightarrow[w_3 - w_1]{w_2 - w_1} \left[\begin{array}{rrrr|r} 1 & 1 & 1 & 2 & 0 \\ 0 & -2 & 0 & 4 & 0 \\ 0 & 0 & -2 & -6 & 0 \end{array}\right] \xrightarrow[-\frac{1}{2}w_3]{-\frac{1}{2}w_2} \left[\begin{array}{rrrr|r} 1 & 1 & 1 & 2 & 0 \\ 0 & 1 & 0 & -2 & 0 \\ 0 & 0 & 1 & 3 & 0 \end{array}\right] \xrightarrow{w_1 - w_3}$$

$$\left[\begin{array}{rrrr|r} 1 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -2 & 0 \\ 0 & 0 & 1 & 3 & 0 \end{array}\right] \xrightarrow{w_1 - w_2} \left[\begin{array}{rrrr|r} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -2 & 0 \\ 0 & 0 & 1 & 3 & 0 \end{array}\right] \Rightarrow \begin{cases} x = -t \\ y = 2t \\ x = -3t \\ t \in \mathbb{R} \end{cases}$$

$\text{Ker}\varphi = \{(-t, 2t, -3t, t),\ t \in \mathbb{R}\} = \text{lin}\{(-1, 2, -3, 1)\}$
Układ $\{(-1, 2, -3, 1)\}$ jest bazą $\text{Ker}\varphi$ oraz $\dim \text{Ker}\varphi = 1$.
Zatem φ nie jest monomorfizmem. Ponadto
$r(\varphi) = \dim \mathbb{R}^4 - \dim \text{Ker}\varphi = 4 - 1 = 3 = \dim \mathbb{R}^3$, zatem φ jest epimorfizmem.
Stąd $\text{Im}\varphi = \mathbb{R}^3$. Można to również sprawdzić bezpośrednim rachunkiem.
$\varphi(x, y, z, t) = x(1, 1, 1) + y(1, -1, 1) + z(1, 1, -1) + t(2, 6, -4)$
$\text{Im}\varphi = \text{lin}\{(1, 1, 1), (1, -1, 1), (1, 1, -1), (2, 6, -4)\}$

$$r \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 2 & 6 & -4 \end{bmatrix} = r \begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \\ 0 & 4 & 6 \end{bmatrix} = 3 \Rightarrow \text{Im}\varphi = \text{lin}\{(1, 1, 1), (0, -2, 0), (0, 0, -2)\}$$

Układ $\{(1, 1, 1), (0, -2, 0), (0, 0, -2)\}$ jest bazą przestrzeni $\text{Im}\varphi$.
$\text{Im}\varphi \subseteq \mathbb{R}^3 \wedge \dim \text{Im}\varphi = 3 = \dim \mathbb{R}^3 \Rightarrow \text{Im}\varphi = \mathbb{R}^3$.

Twierdzenie 7.2.9. 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 7.2.10. 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.

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.

Twierdzenie 7.3.2. i) Jeśli $f \in \mathcal{L}(U, V)$ oraz $g \in \mathcal{L}(V, W)$, to wówczas $g \circ f \in \mathcal{L}(U, W)$.

ii) Jeśli $f \in \mathcal{L}(V, W)$ jest bijekcją, to $f^{-1} \in \mathcal{L}(W, U)$.

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.

Grupa $Aut_K(V)$ bywa też oznaczana symbolem $GL(V)$ i nazywana *pełną* lub *ogólną grupą liniową przestrzeni liniowej V*.

Przykład 7.3.4. Odwzorowanie $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$, dane wzorem
$\varphi(x_1, x_2, x_3) = (x_1 + x_2 + x_3, x_2 + x_3, x_1 + x_2)$ jest automorfizmem przestrzeni $\mathbb{R}^3$.

7.4 Reprezentacja macierzowa odwzorowania liniowego

Niech V oraz W będą przestrzeniami liniowymi skończenie wymiarowymi nad 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

$$\varphi(1,0,0) = (3,0),\ \varphi(0,1,0) = (0,2),\ \varphi(0,0,1) = (0,1), \qquad M_\varphi(\mathcal{B}_k^3, \mathcal{B}_k^2) = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3x \\ 2y + z \end{bmatrix}$$

b) bazy $\mathcal{B} = \big(b_1 = (1,2,0), b_2 = (1,1,1), b_3 = (0,0,1)\big)$, $\mathcal{C} = \big(c_1 = (1,2), c_2 = (0,1)\big)$

$$\begin{aligned}
&\varphi(b_1) = \varphi(1,2,0) = (3,4) = [\alpha_1, \beta_1]_{\mathcal{C}}, && (3,4) = \alpha_1 c_1 + \beta_1 c_2 = (\alpha_1, 2\alpha_1 + \beta_1), \\
& && \Rightarrow\ \alpha_1 = 3, \beta_1 = -2,\ \varphi(b_1) = [3,-2]_{\mathcal{C}} \\
&\varphi(b_2) = \varphi(1,1,1) = (3,3) = [\alpha_2, \beta_2]_{\mathcal{C}}, && (3,3) = \alpha_2 c_1 + \beta_2 c_2 = (\alpha_2, 2\alpha_2 + \beta_2) \\
& && \Rightarrow\ \alpha_2 = 3, \beta_2 = -3,\ \varphi(b_2) = [3,-3]_{\mathcal{C}} \\
&\varphi(b_3) = \varphi(0,0,1) = (0,1) = [\alpha_3, \beta_3]_{\mathcal{C}}, && (3,4) = \alpha_3 c_1 + \beta_3 c_2 = (\alpha_3, 2\alpha_3 + \beta_3) \\
& && \Rightarrow\ \alpha_3 = 0, \beta_3 = 1,\ \varphi(b_3) = [0,1]_{\mathcal{C}}
\end{aligned}$$

$$M_\varphi(\mathcal{B}, \mathcal{C}) = \begin{bmatrix} 3 & 3 & 0 \\ -2 & -3 & 1 \end{bmatrix}$$

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. Wyznacz macierz odwzorowania liniowego $f : \mathbb{C}_1[z] \to M_2(\mathbb{C})$, $f(\alpha z + \beta) = \alpha A + \beta I_2$, gdzie $A = \begin{bmatrix} 2+i & 1 \\ 3 & 4-i \end{bmatrix}$, 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 7.4.5. Niech V oraz W będą przestrzeniami liniowymi nad 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. 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 7.4.6. Istnieje wzajemnie jednoznaczna odpowiedniość między odwzorowaniami liniowymi a macierzami (przy ustalonych bazach). Macierz odwzorowania liniowego φ w pełni opisuje to odwzorowanie, można zatem badać macierz, zamiast odwzorowania.

Wniosek 7.4.7. Rząd macierzy A przekształcenia liniowego φ nie zależy od wyboru baz $\mathcal{B}_V, \mathcal{B}_W$. Ponadto $\operatorname{rank}(\varphi) = \operatorname{rank} A$.

Wniosek 7.4.8. Przyjmijmy oznaczenia jak w twierdzeniu 7.4.5. Wówczas

i) φ jest epimorfizmem $\Leftrightarrow$ $r(A) = m$,

ii) φ jest monomorfizmem $\Leftrightarrow$ $r(A) = n$.

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 $\varphi(1,2,3) = (3 \cdot 1, 2 \cdot 2 + 2) = (3,7)$
$(1,2,3) = 1(1,2,0) + 3(0,0,1) = 1 \cdot b_1 + 2 \cdot b_3 = [1,0,3]_\mathcal{B}$
$A' \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 0 \\ -2 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$
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 7.4.9. 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 7.4.10. Jeśli $\dim V = \dim W$, 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 7.4.11. Niech $f \in \mathcal{L}(V,W)$ będzie izomorfizmem oraz niech $A = M_f(\mathcal{B}_V, \mathcal{B}_W)$. Wówczas $A^{-1} = M_{f^{-1}}(\mathcal{B}_W, \mathcal{B}_V)$.

Twierdzenie 7.4.12. 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

$$B \cdot A = M_{g \circ f}(\mathcal{B}_U, \mathcal{B}_W).$$

Przykład 7.4.13. 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 wyznacz wzór odwzorowania

$\varphi = 2h^{-1} \circ h^{-1} \circ (f+g)$ i oblicz $\varphi(1,2,3)$.

$f : \mathbb{R}^3 \to \mathbb{R}^2 \Rightarrow M_f := M_f(\mathcal{B}_k^3, \mathcal{B}_k^2) \in M_{2\times 3}(\mathbb{R}),\ M_f = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & 1 \end{bmatrix}$

$g : \mathbb{R}^3 \to \mathbb{R}^2 \Rightarrow M_g := M_g(\mathcal{B}_k^3, \mathcal{B}_k^2) \in M_{2\times 3}(\mathbb{R}),\ M_g = \begin{bmatrix} 1 & 0 & -3 \\ 1 & 1 & 0 \end{bmatrix}$

$M_{f+g} := M_{f+g}(\mathcal{B}_k^3, \mathcal{B}_k^2) \in M_{2\times 3}(\mathbb{R}), \quad M_{f+g} = M_f + M_g = \begin{bmatrix} 2 & -1 & -2 \\ 1 & 3 & 1 \end{bmatrix}$

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

$M_{h^{-1}} := M_{h^{-1}}(\mathcal{B}_k^2, \mathcal{B}_k^2) \in M_2(\mathbb{R}), \quad M_{h^{-1}} = \big(M_h\big)^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}$

$\varphi : \mathbb{R}^3 \to \mathbb{R}^2 \Rightarrow M_\varphi := M_\varphi(\mathcal{B}_k^3, \mathcal{B}_k^2) \in M_{2\times 3}(\mathbb{R}),$

$M_\varphi = 2M_{h^{-1}}^2 M_{f+g} = \frac{2}{9}\begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} 2 & -1 & -2 \\ 1 & 3 & 1 \end{bmatrix} = \frac{2}{9}\begin{bmatrix} 3 & -5 & -5 \\ 3 & 16 & 7 \end{bmatrix}$

$\varphi(x,y,z) = ?$

$M_\varphi \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{2}{9}\begin{bmatrix} 3x-5y-5z \\ 3x+16y-7z \end{bmatrix} \Rightarrow \varphi(x,y,z) = (\frac{2}{3}x - \frac{10}{9}y - \frac{10}{9}z, \frac{2}{3}x + \frac{32}{9}y - \frac{14}{9}z)$

$\varphi(1,2,3) = ?$

$M_\varphi \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \frac{2}{9}\begin{bmatrix} 3 & -5 & -5 \\ 3 & 16 & 7 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \frac{2}{9}\begin{bmatrix} -22 \\ 56 \end{bmatrix} \Rightarrow \varphi(1,2,3) = (-\frac{44}{9}, \frac{112}{9})$

7.5 Zmiana macierzy odwzorowania liniowego przy zmianie 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

$$A' = Q^{-1}AP.$$

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 $\mathcal{B} = \big(b_1 = (1,2,0), b_2 = (1,1,1), b_3 = (0,0,1)\big)$, $\mathcal{C} = \big(c_1 = (1,2), c_2 = (0,1)\big)$.

$A = M_\varphi(\mathcal{B}_k^3, \mathcal{B}_k^2) = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \end{bmatrix}, \quad A' = M_\varphi(\mathcal{B}, \mathcal{C}) = ?$

$P = P_{\mathcal{B}_k^3 \to \mathcal{B}} = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, Q = P_{\mathcal{B}_k^2 \to \mathcal{C}} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \quad Q^{-1} = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$

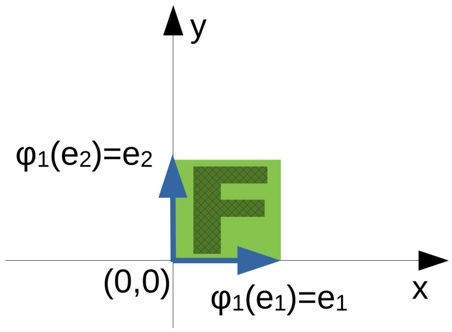
$$A' = Q^{-1}AP = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 0 \\ -2 & -3 & 1 \end{bmatrix}$$

Uwaga 7.5.2. 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 $\mathcal{B}$.

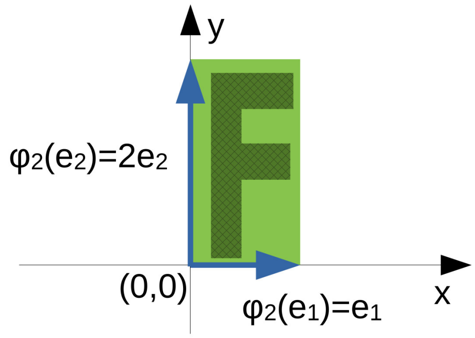
Dowód. Niech $\mathcal{B} = (b_1, \ldots, b_n)$, $\mathcal{B}' = (b'_1, \ldots, b'_n)$. Wówczas

$$\begin{cases} \mathrm{id}(b'_1) = b'_1 = a_{11}b_1 + \ldots + a_{n1}b_1 \\ \ldots \\ \mathrm{id}(b'_n) = b'_1 = a_{n1}b_1 + \ldots + a_{nn}b_n \end{cases}, \text{ skąd } P_{\mathcal{B}\to\mathcal{B}'} = M_{\mathrm{id}}(\mathcal{B}', \mathcal{B}). \quad \square$$

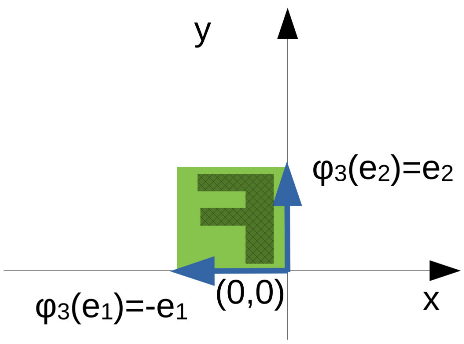
Przykład 7.5.3. Endomorfizmy przestrzeni $\mathbb{R}^2$



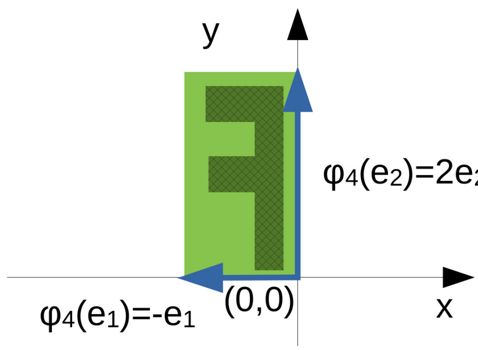
identyczność $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$



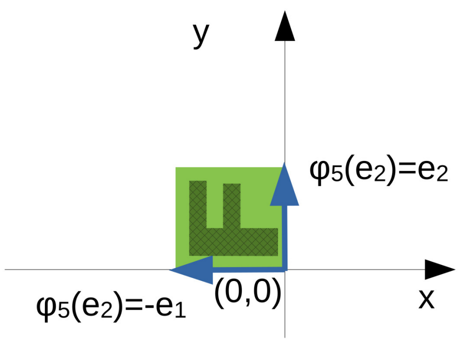
rozciąganie $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$



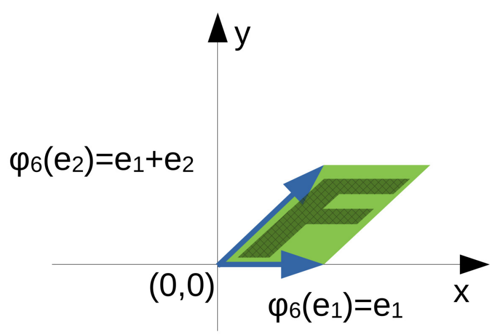
odbicie (symetria osiowa) $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$



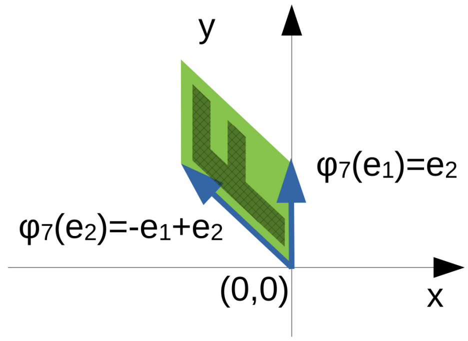
rozciąganie i odbicie $\begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$



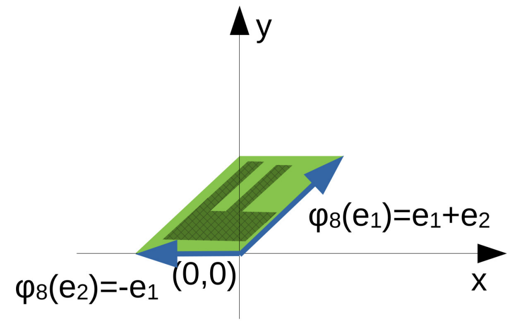
obrót (rotacja) o kąt $\frac{\pi}{2}$ $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$



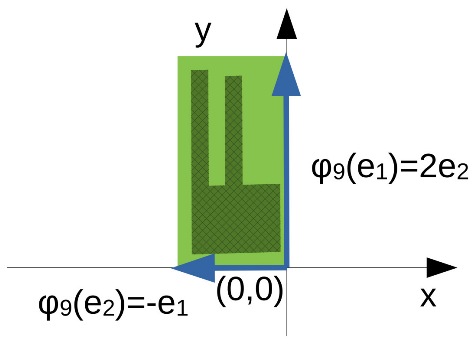
powinowactwo ścinające (ang. shear) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$



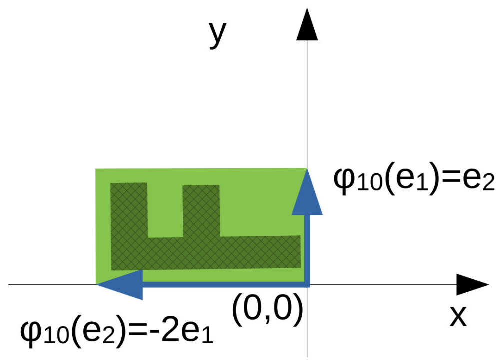
powinowactwo ścinające i rotacja $\begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$



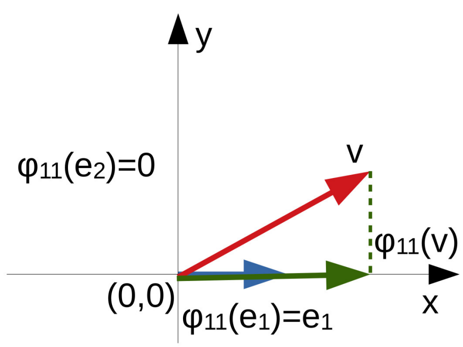
rotacja i powinowactwo ścinające $\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$



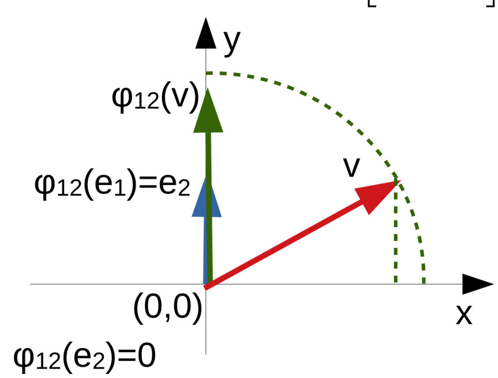
rotacja i rozciąganie $\begin{bmatrix} 0 & -1 \\ 2 & 0 \end{bmatrix}$



rozciąganie i rotacja $\begin{bmatrix} 0 & -2 \\ 1 & 0 \end{bmatrix}$



rzutowanie (projekcja) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$



projekcja i rotacja $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Metody algebry liniowej w fizyce
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) $\forall \alpha \in \mathbb{R} \quad \forall f, g \in End(V) \quad \alpha \cdot (f \circ g) = (\alpha \cdot f) \circ g = f \circ (\alpha \cdot g)$

Dowód. i), ii) Wynikają z twierdzeń 7.3.1 oraz 7.3.2. iii) Wynika z odpowiednich związków dla macierzy odwzorowań $f, g, f \circ g, \alpha f, \alpha g$. $\square$

Definicja 8.1.2. Podprzestrzeń liniową U przestrzeni V nazywamy *niezmienniczą względem endomorfizmu* $\varphi \in End(V)$ lub krótko φ*-niezmienniczą*, jeżeli

$$\varphi(U) \subset U, \quad tzn. \quad \forall u \in U \; \varphi(u) \in U.$$

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, $\varphi \in End(V)$ dowolne, $U = \mathrm{Ker}\varphi$

Niech $u \in U$, wówczas $\varphi(u) = \mathbf{0}_V$. Oczywiście $\mathbf{0}_V \in U$, bowiem $\varphi(\mathbf{0}_V) = \mathbf{0}_V$.
Zatem U jest φ-niezmiennicza.

Znaczenie podprzestrzeni niezmienniczych

Dla dowolnego endomorfizmu $\varphi : V \to V$ i dla dowolnej podprzestrzeni $U \subset V$, mamy $\varphi(U) \subset V$. Gdy U jest φ-niezmiennicza mamy $\varphi(U) \subset U$, zatem restrykcja $\varphi|_U : U \to U$, czyli $\varphi|_U \in End(U)$.

Jeśli istnieje właściwa podprzestrzeń niezmiennicza $U \subset V$, to w odpowiednio dobranej bazie macierz A operatora φ ma prostszą postać. Bierzemy dowolną bazę $(c_1, c_2, \ldots, c_k)$ przestrzeni U i uzupełniamy ją do bazy przestrzeni V. Z warunku $\varphi(c_i) \in U$ dla $i = 1, 2, \ldots, k$ wynika, że $A = \left[\begin{array}{c|c} A_1 & B \\ \hline 0 & A_2 \end{array}\right]$, gdzie $A_1 \in M_k(K), A_2 \in M_{n-k}(k), B \in M_{k\times(n-k)}(K), 0 \in M_{(n-k)\times k}$. Ponadto A_1 to macierz $\varphi|_U : U \to U$.

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.

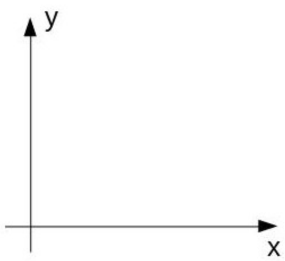
ii) Problem polegający na wyznaczeniu dla danego endomorfizmu φ wartości własnych i odpowiadających im wektorów własnych nazywamy *zagadnieniem własnym* dla endomorfizmu φ.

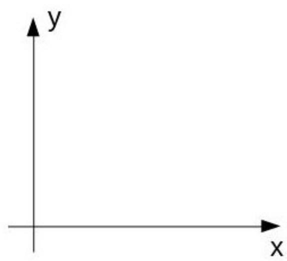
iii) Zbiór wszystkich wartości własnych operatora liniowego φ oznaczamy symbolem $Spec(\varphi)$ i nazywamy *widmem* bądź *spektrum* tego operatora.

Przykład 8.1.5. Niech $V = \mathbb{C}^\infty(\mathbb{R}, \mathbb{R})$ 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 $\frac{d}{dx}$. Istotnie, niech $\lambda \in \mathbb{R}$ będzie dowolne. Zdefiniujmy
$g_\lambda : \mathbb{R} \to \mathbb{R}$, $g_\lambda(x) = a \cdot e^{\lambda x}$, gdzie $a \in \mathbb{R} \neq \{0\}$. 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 $Spec(\varphi) = \mathbb{R}$.

Uwaga 8.1.6. Każdy wektor własny odpowiada dokładnie jednej wartości własnej.

Przykład 8.1.7. 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. 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 $v = (0, v_y)$





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 8.1.8. 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 8.1.9. Przestrzeń wektorową E_λ nazywamy *podprzestrzenią własną* endomorfizmu φ, odpowiadającą wartości własnej λ.

Uwaga 8.1.10. Na mocy uwagi 8.1.6 otrzymujemy $\lambda_1 \neq \lambda_2 \Rightarrow E_{\lambda_1} \cap E_{\lambda_2} = \{\mathbf{0}_V\}$. Zatem zamiast badać endomorfizm $\varphi \in End(V)$, możemy badać jego restrykcje $\varphi|_{E_{\lambda_i}} \in End(E_{\lambda_i})$ na podprzestrzenie niezmiennicze E_{λ_i}, gdzie $\lambda_i \in Spec(\varphi)$.

Twierdzenie 8.1.11. 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 8.1.12. 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

$$(A - \lambda I)X = \mathbf{0}, \quad \text{gdzie} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Twierdzenie 8.1.13. Wartości własne endomorfizmu $\varphi \in End(V)$ nie zależą od wyboru bazy przestrzeni V.

Definicja 8.1.14. *Wielomianem charakterystycznym* endomorfizmu $\varphi \in End(V)$ nazywamy wielomian $\chi_\varphi \in K[t]$ postaci $\chi_\varphi(t) = \det(A - tI)$, gdzie A jest reprezentacją macierzową odwzorowania φ w pewnej bazie przestrzeni V. Równanie $\chi_\varphi(t) = 0$ nazywamy *równaniem charakterystycznym* tego endomorfizmu. Pierwiastki wielomianu χ_φ nazywamy *pierwiastkami charakterystycznymi* odwzorowania φ.

Uwaga 8.1.15.
i) Pierwiastki charakterystyczne wielomianu χ_φ należące do ciała K to wartości własne endomorfizmu φ.

ii) Na mocy twierdzenia 8.1.13 wielomian χ_φ nie zależy od wyboru bazy przestrzeni V.

iii) Jeśli $\dim V = n$, to wówczas $\deg \chi_\varphi = n$.

Niech $A \in M_n(K)$, gdzie $K = \mathbb{R}$ lub $K = \mathbb{C}$.

Definicja 8.1.16.
i) *Wielomianem charakterystycznym* macierzy A nazywamy wielomian $\chi_A \in K[t]$ postaci $\chi_A(t) = \det(A - tI)$. Równanie $\chi_A(t) = 0$ 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 8.1.17. Wartości i wektory własne endomorfizmu $\varphi \in End(\mathbb{R}^n)$ (lub $\varphi \in End(\mathbb{C}^n)$) są identyczne z wartościami i wektorami własnymi macierzy $A \in M_n(\mathbb{R})$ (odpowiednio $A \in M_n(\mathbb{C})$), będącej reprezentacją macierzową odwzorowania φ w bazie kanonicznej przestrzeni $\mathbb{R}^n$ (odpowiednio $\mathbb{C}^n$).

Niech V będzie przestrzenią liniową n-wymiarową. Niech $\varphi \in End(V)$ oraz $\lambda \in Spec(\varphi)$.

Definicja 8.1.18. 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 8.1.19. Niech $\varphi \in End(V)$, $\lambda \in Spec(\varphi)$ oraz niech A będzie reprezentacją macierzową φ w pewnej bazie przestrzeni V. Wówczas

i) $1 \leq \dim E_\lambda \leq k_\lambda$,

ii) $\dim E_\lambda = \dim V - \operatorname{rank}(A - \lambda I)$.

Przykład 8.1.20. 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.

$$A = M_\varphi(\mathcal{B}_k^3, \mathcal{B}_k^3) = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 0 \\ -2 & -2 & -1 \end{bmatrix}$$

$$\chi_A(t) = \chi_\varphi(t) = \det(A - tI) = \begin{vmatrix} 1-t & 2 & 0 \\ 0 & 2-t & 0 \\ -2 & -2 & -1-t \end{vmatrix} = (1-t)(2-t)(-1-t)$$

$Spec(\varphi) = \{\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = -1\}$, $k_1 = k_2 = k_3 = 1$ widmo proste

$E_{\lambda_3} = E_{-1} = ?$

$$(A - \lambda_3 I)X = \mathbf{0}_{\mathbb{R}^3} \Leftrightarrow (A + I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 2 & 2 & 0 \\ 0 & 3 & 0 \\ -2 & -2 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Leftrightarrow \begin{cases} 2x + 2y &= 0 \\ 3y &= 0 \\ -2x - 2y &= 0 \end{cases} \Rightarrow x = y = 0,\ z \in \mathbb{R}$$

Wektory własne odpowiadające $\lambda_3 = -1$ są postaci $v = (0, 0, z)$, gdzie $z \in \mathbb{R} \setminus \{0\}$.

$E_{-1} = \{(0, 0, z) : z \in \mathbb{R}\} = \operatorname{lin}\{(0, 0, 1)\}$ oraz $\dim E_{-1} = 1$

Alternatywnie, korzystając z twierdzenia 8.1.19 ii), możemy obliczyć

$$\dim E_{-1} = \dim \mathbb{R}^3 - r(A + I) = 3 - r\begin{bmatrix} 2 & 2 & 0 \\ 0 & 3 & 0 \\ -2 & -2 & 0 \end{bmatrix} = 3 - 2 = 1.$$

Tak naprawdę, wiemy to z góry, bowiem na mocy twierdzenia 8.1.19 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 $\dim E_{\lambda_1} = \dim E_{\lambda_2} = 1$

Twierdzenie 8.1.21. Niech V będzie skończenie wymiarową przestrzenią liniową nad ciałem K. Niech $A \in M_n(K)$, $\lambda \in Spec(A)$ oraz niech $v = (v_1, \ldots, v_n) \in V$ będzie wektorem własnym odpowiadającym wartości własnej λ. Wówczas prawdziwe są następujące stwierdzenia.

i) $\lambda^k \in Spec(A^k)$ dla każdego $k \in \mathbb{N}, k \geq 2$, oraz v jest wektorem własnym odpowiadającym λ^k.

ii) $c \cdot \lambda \in Spec(c \cdot A)$ dla każdego $c \in K$, oraz v jest wektorem własnym odpowiadającym $c \cdot \lambda$.

iii) $p(\lambda) \in Spec\big(p(A)\big)$ dla każdego wielomianu $p \in K[X]$, oraz v jest wektorem własnym odpowiadającym $p(\lambda)$.

iv) Jeśli A jest nieosobliwa oraz $\lambda \neq 0$, to $\frac{1}{\lambda} \in Spec(A^{-1})$ oraz v jest wektorem własnym odpowiadającym $\frac{1}{\lambda}$.

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

vi) A jest odwracalna $\Leftrightarrow 0 \notin Spec(A)$.

Twierdzenie 8.1.22. 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 <u>rzeczywiste</u>, 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.

8.2 Diagonalizacja

Twierdzenie 8.2.1. Wektory własne operatora liniowego $\varphi \in End(V)$ odpowiadające różnym wartościom własnym są liniowo niezależne.

Twierdzenie 8.2.2. Niech V będzie rzeczywistą (zespoloną) przestrzenią liniową n-wymiarową oraz niech $\varphi \in End(V)$.

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) 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 8.2.3. 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 8.2.4. 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 K (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.

Przykład 8.2.5. Niech $\varphi \in End(\mathbb{R}^2)$ oraz $A = M_\varphi(\mathcal{B}_k^2, \mathcal{B}_k^2) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in M_2(\mathbb{R})$.

$\chi_A(t) = \begin{vmatrix} -t & 1 \\ -1 & -t \end{vmatrix} = t^2 + 1 \neq 0.$

Wielomian charakterystyczny nie posiada pierwiastków rzeczywistych, zatem $Spec(A) = \varnothing$ i nie istnieje baza $\mathbb{R}^2$ złożona z wektorów własnych φ.

Uwaga 8.2.6. Z zasadniczego twierdzenia algebry wynika, że każdy operator liniowy na zespolonej przestrzeni liniowej ma wektory własne.

Macierz diagonalizująca

Definicja 8.2.7. 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 8.2.8 (o niezmiennikach macierzy podobnych)**.** Jeżeli macierze A i B są podobne, to wówczas

i) $r(A) = r(B)$,

ii) $\det A = \det B$

iii) $\mathrm{tr}(A) = \mathrm{tr}(B)$

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 8.2.9. Macierz $A \in M_n(K)$ nazywamy *diagonalizowalną*, gdy jest podobna do macierzy diagonalnej, tzn. istnieje macierz <u>nieosobliwa</u> $P \in GL_n(K)$ taka, że macierz $P^{-1}AP$ jest diagonalna. Mówimy wówczas, że macierz P *diagonalizuje* macierz A.

Wniosek 8.2.10. Macierz $A \in M_n(K)$ jest diagonalizowalna wtedy i tylko wtedy, gdy istnieje baza przestrzeni K^n złożona z wektorów własnych A.

Z uwagi na związek między widmem endomorfizmu a widmem macierzy wszystkie twierdzenia udowodnione dla endomorfizmu są prawdziwe dla macierzy.

Przykład 8.2.11. Czy $A = \begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix} \in M_2(\mathbb{R})$ jest diagonalizowalna?

$\chi_A(t) = \begin{vmatrix} 2-t & -3 \\ 1 & 1-t \end{vmatrix} = (2-t)(1-t) + 3 = t^2 - 3t + 5 \neq 0.$
Wielomian charakterystyczny nie posiada pierwiastków rzeczywistych, zatem $Spec(A) = \varnothing$ i macierz rzeczywista A nie jest diagonalizowalna.

Przykład 8.2.12. Czy $A = \begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix} \in M_2(\mathbb{C})$ jest diagonalizowalna?

$\chi_A(t) = \begin{vmatrix} 2-t & -3 \\ 1 & 1-t \end{vmatrix} = (2-t)(1-t) + 3 = t^2 - 3t + 5.$
Wielomian charakterystyczny to wielomian zespolony o współczynnikach rzeczywistych. Posiada zatem dwa sprzężone pierwiastki zespolone.
Mamy $\chi_A(t) = (t - \frac{3-\sqrt{11}i}{2})(t - \frac{3+\sqrt{11}i}{2})$. Macierz zespolona A jest diagonalizowalna.

Przykład 8.2.13. 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
$\lambda_1 = -1, v_1 = (1,1,1)$, $\lambda_2 = 1, v_1 = (0,1,1)$, $\lambda_3 = 2, v_1 = (0,0,1)$.
Widmo jest proste, a zatem na mocy twierdzenia 8.2.2 i), operator φ jest diagonalizowalny.

Przykład 8.2.14. Czy $\varphi \in End(\mathbb{R}^3)$ dany wzorem
$\varphi(x,y,z) = (3x + 8z, 3x - y + 6z, -2x - 5z)$ jest diagonalizowalny?

$$A = M_\varphi(\mathcal{B}_k^3, \mathcal{B}_k^3) = \begin{bmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{bmatrix}$$

$$\chi_A(t) = \begin{vmatrix} 3-t & 0 & 8 \\ 3 & -1-t & 6 \\ -2 & 0 & -5-t \end{vmatrix} = (-1-t)(-1)^4 \begin{vmatrix} 3-t & 8 \\ -2 & -5-t \end{vmatrix} =$$

$= -(1+t)(1+2t+t^2) = -(1+t)^3$

$Spec(\varphi) = \{\lambda_1 = -1\}$, $k_1 = 3$

$$E_{\lambda_1} = E_{-1} = \{v = (x,y,z) :\ (A+I) \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\}$$

$$\dim E_{-1} = 3 - r(A+I) = 3 - r \begin{bmatrix} 4 & 0 & 8 \\ 3 & 0 & 6 \\ -2 & 0 & -4 \end{bmatrix} = 3 - 2 = 1 \neq k_1 = 3.$$

Endomorfizm φ nie jest diagonalizowalny, bowiem nie istnieje baza $\mathbb{R}^3$ złożona z wektorów własnych φ.

Przykład 8.2.15. Wyznacz wszystkie wartości własne endomorfizmu f i określ ich krotności algebraiczne. Wyznacz odpowiadające im podprzestrzenie własne i wymiar tychże podprzestrzeni. Czy f jest diagonalizowalny? Jeśli tak, podaj bazę wektorów własnych, macierz D endomorfizmu f w tejże bazie oraz macierz diagonalizującą P.

$$f \in End(M_2(\mathbb{R})), \quad f(A) = A + A^T$$

$$\mathcal{B} = \left(E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right) \text{ baza } M_2(\mathbb{R})$$

$$f(E_{11}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$f(E_{12}) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$f(E_{21}) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$f(E_{22}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

$$A = M_f(\mathcal{B}, \mathcal{B}) = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

$$\chi_f(t) = \det(A - tI) = (2-t) \begin{vmatrix} 1-t & 1 & 0 \\ 1 & 1-t & 0 \\ 0 & 0 & 2-t \end{vmatrix} = (2-t)^2 \begin{vmatrix} 1-t & 1 \\ 1 & 1-t \end{vmatrix} = (2-$$

$t)^2\big((1-t)^2 - 1\big) = (2-t)^2(t^2-2t) = -t(2-t)^3$

$Spec(f) = \{\lambda_1 = 0, \lambda_2 = 2\}$, $\quad k_1 = 1, k_2 = 3$, $\dim E_0 = 1$, $1 \leq \dim E_2 \leq 3$

$E_{\lambda_1} = E_0 = \mathrm{Ker}(f) = \{B \in M_2(\mathbb{R}) :\ B + B^T = O\}$

Niech $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, gdzie $a, b, c, d \in \mathbb{R}$.

$$B + B^T = O \Leftrightarrow \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \Rightarrow \begin{cases} 2a & = 0 \\ b + c & = 0 \\ 2d & = 0 \end{cases}$$

Zatem $B = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$ oraz $E_0 = \left\{ \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix} : b \in \mathbb{R} \right\} = \operatorname{lin} \left\{ \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\}$

$E_{\lambda_2} = E_2 = \operatorname{Ker}(f - 2 \cdot \mathrm{id}_{M_2(\mathbb{R})}) = \{B \in M_2(\mathbb{R}) : B + B^T = 2B\}$

$B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = [a, b, c, d]_{\mathcal{B}}$, $O_{2\times 2} = [0, 0, 0, 0]_{\mathcal{B}}$

$$(A - 2I) \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = O \Leftrightarrow \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow b = c, \quad a, b, d \in \mathbb{R}$$

$$E_2 = \left\{ \begin{bmatrix} a & b \\ b & d \end{bmatrix} : a, b, d \in \mathbb{R} \right\} = \operatorname{lin} \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ są liniowo niezależne $\Rightarrow \dim E_2 = 3$

Baza wektorów własnych $\mathcal{C} = \left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right)$

$$D = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \ P_{\mathcal{B}\to\mathcal{C}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

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 $v_1 \dots, v_n$ odpowiadające wartościom własnym $\lambda_1, \dots, \lambda_n$ (niekoniecznie różnym), tworzą bazę $\mathcal{C} = (v_1 \dots, v_n)$ przestrzeni V, to wówczas dla dowolnego $r \in \mathbb{N}$

$$M_{\varphi^r}(\mathcal{B}, \mathcal{B}) = PD^rP^{-1}, \quad \text{gdzie} \quad P = P_{\mathcal{B}\to\mathcal{C}}, \quad D = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & & \dots & \lambda_n \end{bmatrix}.$$

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 $\varphi^{101}(1, 2, 3)$.

$$\varphi : \mathbb{R}^3 \to \mathbb{R}^3, \quad \varphi(x, y, z) = (4x - 2y + 2z, 2x + 2z, y + z - x)$$

$$A = M_\varphi(\mathcal{B}_k^3, \mathcal{B}_k^3) = \begin{bmatrix} 4 & -2 & 2 \\ 2 & 0 & 2 \\ -1 & 1 & 1 \end{bmatrix}$$

$$\chi_A(t) = \begin{vmatrix} 4-t & -2 & 2 \\ 2 & -t & 2 \\ -1 & 1 & 1-t \end{vmatrix} \overset{w_2+2w_3}{=} \begin{vmatrix} 4-t & -2 & 2 \\ 0 & 2-t & 4-2t \\ -1 & 1 & 1-t \end{vmatrix} \overset{k_3+k_2}{=} \begin{vmatrix} 4-t & -2 & 0 \\ 0 & 2-t & 6-3t \\ -1 & 1 & 2-t \end{vmatrix} \overset{k_2+k_1}{=}$$

$$\begin{vmatrix} 4-t & 2-t & 0 \\ 0 & 2-t & 6-3t \\ -1 & 0 & 2-t \end{vmatrix} = (4-t)(2-t)^2 - 3(2-t)^2 = (2-t)^2(1-t)$$

$Spec\varphi = \{\lambda_1 = 1 \lambda_2 = 2\}$, $\quad k_1 = 1,\, k_2 = 2, \quad \dim E_1 = 1, \quad 1 \le \dim E_2 \le 2$

φ będzie diagonalizowalny, jeśli $\dim E_2 = 2 = k_2$.

$$\dim E_2 = 3 - r(A - 2I) = 3 - r \begin{bmatrix} 2 & -2 & 2 \\ 2 & -2 & 2 \\ -1 & 1 & 1 \end{bmatrix} = 3 - 1 = 2,$$

zatem φ jest diagonalizowalny.

Wyznaczymy bazę złożoną z wektorów własnych. Rozważmy $\lambda_1 = 1$.

$$(A - I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 3 & -2 & 2 \\ 2 & -1 & 2 \\ -1 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\left[\begin{array}{rrr|r} 3 & -2 & 2 & 0 \\ 2 & -1 & 2 & 0 \\ -1 & 1 & 0 & 0 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & -1 & 0 & 0 \\ 3 & -2 & 2 & 0 \\ 2 & -1 & 2 & 0 \end{array}\right] \xrightarrow[w_3 - 2w_1]{w_2 - 3w_1} \left[\begin{array}{rrr|r} 1 & -1 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right]$$

$$\begin{cases} x - y & = 0 \\ y + 2z & = 0 \end{cases} \Rightarrow x = y = -2z, z \in \mathbb{R}$$

$E_1 = \{(-2z, -2z, z) : z \in \mathbb{R}\} = \operatorname{lin}\{(-2, -2, 1)\}$

Rozważmy $\lambda_2 = 2$.

$$(A - 2I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 2 & -2 & 2 \\ 2 & -2 & 2 \\ -1 & 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$-x + y - z = 0 \Rightarrow z = y - x,\ x, y \in \mathbb{R}$

$E_2 = \{(x, y, y - x) : x, y \in \mathbb{R}\} = \operatorname{lin}\{(1, 0, -1), (0, 1, 1)\}$

Baza wektorów własnych $\mathcal{C} = \big(v_1 = (-2, -2, 1), v_2 = (1, 0, -1), v_3 = (0, 1, 1)\big)$

$$D = M_\varphi(\mathcal{C}, \mathcal{C}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \text{ macierz diagonalizująca } P := P_{\mathcal{B}_k^3 \to \mathcal{C}} = \begin{bmatrix} -2 & 1 & 0 \\ -2 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

Obliczymy $\varphi^{101}(1, 2, 3)$ na dwa różne sposoby.

METODA I: Wykorzystamy macierz D. Niech $v := (1, 2, 3) = [a, b, c]_{\mathcal{C}}$.

$$\text{Wówczas } \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = P \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \text{ skąd } \begin{bmatrix} a \\ b \\ c \end{bmatrix} = P^{-1}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

$$\text{Obliczamy } P^{-1} = \begin{bmatrix} 1 & -1 & 1 \\ 3 & -2 & 2 \\ 2 & -1 & 2 \end{bmatrix} \text{ oraz } v = [2, 5, 6]_{\mathcal{C}}.$$

$$D^{101}\begin{bmatrix}2\\5\\6\end{bmatrix}=\begin{bmatrix}1&0&0\\0&2^{101}&0\\0&0&2^{101}\end{bmatrix}\begin{bmatrix}2\\5\\6\end{bmatrix}=\begin{bmatrix}2\\5\cdot 2^{101}\\6\cdot 2^{101}\end{bmatrix}$$

$\varphi^{101}(v)=[2,5\cdot 2^{101},6\cdot 2^{101}]_{\mathcal{C}}=2v_1+5\cdot 2^{101}v_2+6\cdot 2^{101}v_3=$
$=(-4+5\cdot 2^{101},-4-6\cdot 2^{101},2+2^{101})$

METODA II: Obliczamy

$$A^{101}\begin{bmatrix}1\\2\\3\end{bmatrix}=PD^{101}P^{-1}\begin{bmatrix}1\\2\\3\end{bmatrix}=\begin{bmatrix}-2&1&0\\-2&0&1\\1&-1&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&2^{101}&0\\0&0&2^{101}\end{bmatrix}\begin{bmatrix}1&-1&1\\3&-2&2\\2&-1&2\end{bmatrix}\begin{bmatrix}1\\2\\3\end{bmatrix}=$$

$$\begin{bmatrix}3\cdot 2^{101}-2&2-2^{102}&2^{102}-2\\2^{102}-2&2-2^{101}&2^{102}-2\\1-2^{101}&-1+2^{101}&1\end{bmatrix}\begin{bmatrix}1\\2\\3\end{bmatrix}=\begin{bmatrix}-4+5\cdot 2^{101}\\-4-6\cdot 2^{101}\\2+2^{101}\end{bmatrix}$$

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) $\quad a_1=r_1,a_2=r_2,\ldots,a_k=r_k,$ gdzie $r_1,r_2,\ldots,r_k\in K$, $k\geq 1$.
ii) $\quad a_n=p_1a_{n-1}+p_2a_{n-2}+\ldots+p_ka_{n-k},$ 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 $a_1 = 0, a_2 = 8$, $a_n = 2a_{n-1} + 3a_{n-2}$.

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 $\det(A_2 - tI) = \begin{vmatrix} 2-t & 3 \\ 1 & -t \end{vmatrix} = t^2 - 2t - 3 = (t-3)(t+1)$.

$Spec(A_2) = \{-1, 3\}$ 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.

Przykład 8.4.2. Niech $A = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix}$. Wiemy, że $\chi_A(t) = t^2 - 2t - 3$.

Na mocy powyższego twierdzenia $A^2 - 2A + 3I = 0$. Sprawdzimy to bezpośrednim 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 $\chi_A(t) = c_n t^n + c_{n-1} t^{n-1} + \ldots + c_1 t + c_0$ będzie wielomianem charakterystycznym macierzy A.

i) Jeśli A jest odwracalna, to wówczas
$A^{-1} = -\frac{1}{c_0}(c_n A^{n-1} + c_{n-1} A^{n-2} + \ldots + c_2 A + c_1 I)$

ii) $A^n = -\frac{1}{c_n}(c_{n-1}A^{n-1} + \ldots + c_1 A + c_0 I)$

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
$\chi_A(t) = \det(A - tI) = \begin{vmatrix} 1-t & 4 \\ 2 & 3-t \end{vmatrix} = t^2 - 4t - 5 \quad \Rightarrow \quad A^2 - 4A - 5I = 0$
$A^{-1}(A^2 - 4A - 5I) = A - 4I - 5A^{-1} = 0 \quad \Rightarrow \quad A^{-1} = \frac{1}{5}(A - 4I) = \frac{1}{5}\left(\begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} - \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}\right)$
$A^{-1} = \frac{1}{5}\begin{bmatrix} -3 & 4 \\ 2 & -1 \end{bmatrix}$

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 $\chi_B(t) = \det(B - tI) = \begin{vmatrix} -t & 2 & 6 \\ 2 & -8-t & -26 \\ -2 & 2 & 8-t \end{vmatrix} \overset{w_1 - w_3, w_2 + w_3}{=} \begin{vmatrix} -t+2 & 0 & -2+t \\ 0 & -6-t & -18-t \\ -2 & 2 & 8-t \end{vmatrix}$
$\overset{k_3 - 3k_2}{=} \begin{vmatrix} -t+2 & 0 & -2+t \\ 0 & -6-t & 2t \\ -2 & 2 & 2-t \end{vmatrix} \overset{w_3 + w_1}{=} \begin{vmatrix} -t+2 & 0 & -2+t \\ 0 & -6-t & 2t \\ -t & 2 & 0 \end{vmatrix} =$
$= (2-t)(6+t)t - 4t(2-t) = -t^3 + 4t.$
Na mocy twierdzenia Cayleya-Hamiltona $B^3 = 4B$, skąd
$B^9 = 4^3 B^3 = 4^4 B = 256 \begin{bmatrix} 0 & 2 & 6 \\ 2 & -8 & -26 \\ -2 & 2 & 8 \end{bmatrix}.$

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}.

Oznaczmy $p(t) = t^5 + t^3$. Obliczamy $\chi_C(t) = (1-t)^2$. Dzielimy wielomian p przez wielomian χ_C (na ogół z resztą). Zatem istnieją wielomiany q_1, r_1 takie że $p(t) = q_1(t)\chi_C(t) + r_1(t)$ oraz $\deg(r_1) < \deg(\chi_C)$. Wykonując dzielenie, otrzymujemy $q_1(t) = t^3 + 2t^2 + 4t + 6, r_1(t) = 8t - 6$.

Na mocy twierdzenia Cayleya-Hamiltona
$p(C) = \chi_C(C)q_1(C) + r_1(C) = r_1(C) = 8C - 6I = 8\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - 6\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 8 \\ 0 & 2 \end{bmatrix}.$

Oznaczmy $w(t) = t^{50}$. Istnieją wielomiany q_2, r_2 takie że $w(t) = q_2(t)\chi_C(t) + r_2(t)$ oraz $\deg(r_2) < \deg(\chi_C) = 2$. Oznaczmy $r_2(t) = at + b$, $a, b \in \mathbb{R}$. Różniczkując obustronnie otrzymujemy $w'(t) = q_2'(t)\chi_C(t) + q_2(t)\chi_C'(t) + r_2'(t)$. Do układu równań $\left\{\begin{array}{rcl} t^{50} & = & q_2(t)(1-t)^2 + at + b \\ 50t^{49} & = & q_2'(t)(1-t)^2 - 2q_2(t)(1-t) + a \end{array}\right.$ podstawiamy $t = 1$.

Stąd $\left\{\begin{array}{rcl} 1 & = & a + b \\ 50 & = & a \end{array}\right.$. Zatem $r_2(t) = 50t - 49$ oraz $C^{50} = 50C - 49I = \begin{bmatrix} 1 & 50 \\ 0 & 1 \end{bmatrix}$.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Metody algebry liniowej w fizyce
Wykład nr 9
Data: ....

TEMAT: *Przestrzenie euklidesowe*

9.1 Przestrzenie euklidesowe

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 (iloczynem wewnętrznym)*, jeżeli spełnia ona następujące warunki:

i) $\forall u, v, w \in V \quad \forall \alpha, \beta \in \mathbb{R} \quad s(\alpha u + \beta v, w) = \alpha s(u, w) + \beta s(v, w)$,

ii) $\forall u, v \in V \quad s(u, v) = s(v, u)$,

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 euklidesową.* Bywa ona oznaczana symbolem E. Zamiast $s(u, v)$ będziemy również pisać $u \circ v$ lub $\langle u, v \rangle$.

Przykład 9.1.2. 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
$\forall u = (x_1, \ldots, x_n), v = (y_1, \ldots, y_n) \in \mathbb{R}^n \quad u \circ v = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \ldots + x_n y_n$
Przestrzeń euklidesową $(\mathbb{R}^n, \circ)$ oznaczamy symbolem $\mathbb{E}^n$.

2) $V = \mathcal{C}([a, b], \mathbb{R}), \quad \langle ., . \rangle : V \times V \to \mathbb{R}, \quad \forall f, g \in V \quad \langle f, g \rangle = \int_a^b f(x) g(x) dx$

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) $V = \mathbb{R}_n[x], \quad \langle ., . \rangle : \mathbb{R}_n[x] \times \mathbb{R}_n[x] \to \mathbb{R}$,
$\forall p, q \in V \quad \langle p, q \rangle = \sum_{i=0}^{n} p(x_i) q(x_i)$, 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 $p(-1) = p(0) = p(1) = 0$. Nie oznacza to jeszcze, że p jest wielomianem zerowym.

Niech $p(x) = a_0 + a_1x + \ldots + a_nx^n$. Załóżmy, że $\langle p, p\rangle = 0$. Wówczas $p(x_0) = p(x_1) = \ldots = p(x_n) = 0$, skąd otrzymujemy $\begin{bmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ & \ldots & & \ddots & \\ 1 & x_n & x_n^2 & \ldots & x_n^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$.

Wyznacznik główny W tego układu, to *wyznacznik Vandermonde'a*.
$W = \prod_{0\le i<j\le n}(x_j - x_i) \neq 0$, bowiem $x_i \neq x_j$ dla $i \neq j$. Zatem jest to układ oznaczony jednorodny, jego jedynym rozwiązaniem jest $a_0 = a_1 = \ldots = a_n = 0$, co oznacza, że p jest wielomianem zerowym.

4) $V = M_{m\times n}(\mathbb{R})$, $\quad \langle ., .\rangle : M_{m\times n}(\mathbb{R}) \times M_{m\times n}(\mathbb{R}) \to \mathbb{R}$,
$\forall A, B \in M_{m\times n}(\mathbb{R}) \quad \langle A, B\rangle = \operatorname{tr}(AB^T)$

Na mocy twierdzenia 3.1.13, mówiącego o własnościach śladu macierzy, można wywnioskować, że jest to iloczyn skalarny.

Twierdzenie 9.1.3. Niech (V, s) będzie przestrzenią euklidesową. Wówczas

i) $\forall u, v, w \in V \quad \forall \alpha, \beta \in \mathbb{R} \quad s(u, \alpha v + \beta w) = \alpha s(u, v) + \beta s(u, w)$,

ii) $\forall v \in V \quad s(v, \mathbf{0}_V) = 0$,

iii) $\forall u, v \in V \quad \Big(s(u, v)\Big)^2 \le s(u, u) \cdot s(v, v)$ $\quad$ *nierówność Schwarza*

Niech $V = (V, +, \mathbb{R}, \cdot)$ będzie przestrzenią liniową.

Definicja 9.1.4. Funkcję $||.|| : V \to \mathbb{R}$ nazywamy *normą*, jeżeli spełnia ona następujące warunki:

i) $\forall v \in V \quad ||v|| \ge 0 \quad \wedge \quad ||v|| = 0 \Leftrightarrow v = \mathbf{0}_V$,

ii) $\forall v \in V \quad \forall \alpha \in \mathbb{R} \quad ||\alpha v|| = |\alpha| \cdot ||v||$,

iii) $\forall u, v \in V \quad ||u + v|| \le ||u|| + ||v||$. $\qquad$ tzw. warunek trójkąta

Liczbę $||v|| \ge 0$ nazywamy *normą (lub długością) wektora* v. Parę $(V, ||.||)$ nazywamy *przestrzenią unormowaną*.

Twierdzenie 9.1.5. 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.*

Przykład 9.1.6. 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 $\mathbb{R}^n$.

Wniosek 9.1.7. Każda przestrzeń euklidesowa jest przestrzenią unormowaną.

9.2 Układy ortogonalne

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 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 $u(x) = 2, v(x) = 5 - x$.

$$\begin{aligned}
&\langle u, v\rangle = u(0)v(0) + u(1)v(1) = 2 \cdot 5 + 2 \cdot 4 = 18\\
&||u|| = \sqrt{\langle u, u\rangle} = \sqrt{[u(0)]^2 + [u(1)]^2} = \sqrt{2^2 + 2^2} = 2\sqrt{2}\\
&||v|| = \sqrt{\langle v, v\rangle} = \sqrt{[v(0)]^2 + [v(1)]^2} = \sqrt{5^2 + 4^2} = \sqrt{41}\\
&\measuredangle(u, v) = \arccos \tfrac{18}{2\sqrt{2}\cdot\sqrt{41}} = \arccos \tfrac{9}{\sqrt{82}}
\end{aligned}$$

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 $\{v_1, \dots v_n\} \subset V$ nazywamy *układem ortogonalnym*, jeżeli

$$\forall i, j \in \{1, 2, \dots, n\} \quad i \neq j \;\Rightarrow \langle v_i, v_j\rangle = 0.$$

ii) Układ wektorów nazywamy *układem ortonormalnym*, jeśli jest układem ortogonalnym i każdy wektor jest unormowany, czyli

$$\forall i, j \in \{1, 2, \dots, n\} \quad \langle v_i, v_j\rangle = \begin{cases} 0 & ; \; i \neq j \\ 1 & ; \; i = j \end{cases}.$$

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.

Twierdzenie 9.2.5 (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 9.2.6. Układ ortogonalny nie zawierający wektora zerowego jest liniowo niezależny.

Wniosek 9.2.7. 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 9.2.8. Bazę przestrzeni euklidesowej, która jest układem ortogonalnym (ortonormalnym), nazywamy *bazą ortogonalną (ortonormalną)* tej przestrzeni.

Przykład 9.2.9. W przestrzeni $\mathbb{R}^n$ ze standardowym iloczynem skalarnym bazą ortogonalną jest baza kanoniczna $e_1=(1,0,\dots,0), e_2=(0,1,\dots,0),\dots,e_n=(0,0,\dots,1)$.

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 $v_x = v\circ\hat{i}$, $v_y = v\circ\hat{j}$, $v_z = v\circ\hat{k}$.
Z dokładnością do znaku skalary v_x,v_y,v_z to długości rzutów ortogonalnych wektora v na osie Ox,Oy,Oz odpowiednio, zaś 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.

Twierdzenie 9.2.10. Niech $\mathcal{B} = \big(b_1,b_2,\dots,b_n\big)$ będzie bazą ortogonalną przestrzeni euklidesowej $(V,\langle .,.\rangle)$. Niech $v\in V$, $v=[\alpha_1,\alpha_2,\dots,\alpha_n]_{\mathcal{B}}$. Wówczas

$$\forall i\in\{1,2,\dots,n\} \quad \alpha_i = \frac{\langle v,b_i\rangle}{||b_i||^2}.$$

Ponadto

$$\forall u,w\in V \quad \langle u,w\rangle = \frac{\langle u,b_1\rangle\langle w,b_1\rangle}{||b_1||^2} + \frac{\langle u,b_2\rangle\langle w,b_2\rangle}{||b_2||^2} + \dots \frac{\langle u,b_n\rangle\langle w,b_n\rangle}{||b_n||^2}.$$

Wniosek 9.2.11. Jeśli $\mathcal{B} = \big(b_1, b_2, \ldots, b_n\big)$ 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

$$\forall u, w \in V \quad \langle u, w \rangle = \langle u, b_1 \rangle \langle w, b_1 \rangle + \langle u, b_2 \rangle \langle w, b_2 \rangle + \ldots + \langle u, b_n \rangle \langle w, b_n \rangle.$$

Wniosek 9.2.12. Niech $\mathcal{B} = \big(b_1, b_2, \ldots, b_n\big)$ 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 $\langle u, w \rangle = \alpha_1\beta_1 + \alpha_2\beta_2 \ldots + \alpha_n\beta_n$.

9.3 Metoda ortogonalizacji Grama-Schmidta

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 $\mathrm{lin}\{b_1, \ldots b_m\} = \mathrm{lin}\{c_1, \ldots c_m\}$.

Wniosek 9.3.2. i) Każda skończenie wymiarowa przestrzeń euklidesowa rożna 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

Przykład 9.3.3. Rozważmy przestrzeń $\mathbb{E}^3$, tj. $\mathbb{R}^3$ ze standardowym iloczynem skalarnym. Dana jest baza $\mathcal{B} = \big(b_1 = (1, -2, 0), b_2 = (5, 5, 1), b_3 = (5, 4, 4)\big)$ przestrzeni $\mathbb{R}^3$. Dokonamy ortogonalizacji bazy $\mathcal{B}$.

Zauważmy, że baza $\mathcal{B}$ nie jest ortogonalna, bowiem $b_1 \circ b_2 = 5 - 10 + 0 = -5 \neq 0$.
Niech $\mathcal{C} = (c_1, c_2, c_3)$ oznacza poszukiwaną bazę ortogonalną.

I KROK: Niech $c_1 := b_1 = (1, -2, 0)$. Wówczas oczywiście $\mathrm{lin}\{c_1\} = \mathrm{lin}\{b_1\}$.

II KROK: Aby zagwarantować, że $\mathrm{lin}\{c_1, c_2\} = \mathrm{lin}\{b_1, b_2\}$, poszukujemy c_2 w postaci $c_2 = b_2 + \alpha c_1$, dla pewnego $\alpha \in \mathbb{R}$. Dobierzemy α w taki sposób, by $c_2 \circ c_1 = 0$.
Obliczamy $c_2 \circ c_1 = (b_2 + \alpha c_1) \circ c_1 = b_2 \circ c_1 + \alpha \cdot (c_2 \circ c_1)$.
Aby $c_2 \circ c_1 = 0$, wystarczy przyjąć $\alpha = -\frac{b_2 \circ c_1}{c_1 \circ c_1}$.
W naszym przykładzie $0 = b_2 \circ c_1 + \alpha \cdot (c_2 \circ c_1) = (5, 5, 1) \circ (1, -2, 0) + \alpha \cdot (1, -2, 0) \circ (1, -2, 0) = -5 + 5\alpha$, skąd $\alpha = 1$. Zatem $c_2 = b_2 + c_1 = (6, 3, 1)$.

III KROK: Aby zagwarantować, że $\mathrm{lin}\{c_1, c_2, c_3\} = \mathrm{lin}\{b_1, b_2, b_3\}$, poszukujemy c_3 w postaci $c_3 = b_3 + \beta_1 c_1 + \beta_2 c_2$, dla pewnych $\beta_1, \beta_2 \in \mathbb{R}$. Dobierzemy β_1, β_2 w taki sposób, by $c_3 \circ c_1 = 0$ oraz $c_3 \circ c_2 = 0$.

Mamy $0 = c_3 \circ c_1 = (b_3 + \beta_1 c_1 + \beta_2 c_2) \circ c_1 = b_3 \circ c_1 + \beta_1 (c_1 \circ c_1) + \beta_2 (c_2 \circ c_1) = b_3 \circ c_1 + \beta_1 ||c_1||^2$, skąd $\beta_1 = -\frac{b_3 \circ c_1}{||c_1||^2}$. Analogicznie $0 = c_3 \circ c_2 = (b_3 + \beta_1 c_1 + \beta_2 c_2) \circ c_2 = b_3 \circ c_2 + \beta_1 (c_1 \circ c_2) + \beta_2 (c_2 \circ c_2) = b_3 \circ c_2 + \beta_2 ||c_2||^2$, skąd $\beta_2 = -\frac{b_3 \circ c_2}{||c_2||^2}$.

W naszym przykładzie $\begin{cases} 0 = b_3 \circ c_1 + \beta_1 ||c_1||^2 = (5,4,4) \circ (1,-2,0) + \beta_1 (1+4+0) = -3 + 5\beta_1 \\ 0 = b_3 \circ c_2 + \beta_2 ||c_2||^2 = (5,4,4) \circ (6,3,1) + \beta_2 (36+9+1) = 46 + 46\beta_2 \end{cases}$.
Skąd $\beta_1 = \frac{3}{5}$, $\beta_2 = -1$.
Zatem $c_3 = b_3 + \frac{3}{5} c_1 - c_2 = (5,4,4) + \frac{3}{5}(1,-2,0) - (6,3,1) = (-\frac{2}{5}, -\frac{1}{5}, 3)$.
$\mathcal{C} = \big(c_1 = (1,-2,0), c_2 = (6,3,1), c_3 = (-\frac{2}{5}, -\frac{1}{5}, 3)\big)$ jest bazą ortogonalną.

Ponieważ $||c_1|| = \sqrt{5}$, $||c_2|| = \sqrt{46}$, $||c_3|| = \sqrt{\frac{46}{5}}$, zatem bazą ortonormalną jest
$\mathcal{C}' = \left(c_1' = \frac{c_1}{||c_1||} = \frac{1}{\sqrt{5}} \cdot (1,-2,0), c_2' = \frac{c_2}{||c_2||} = \frac{1}{\sqrt{46}} \cdot (6,3,1), c_3' = \frac{c_3}{||c_3||} = \sqrt{\frac{5}{46}} \cdot (-\frac{2}{5}, -\frac{1}{5}, 3)\right)$.

Wniosek 9.3.4. 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.

$$c_1 := b_1 \qquad c_2 := b_2 - \frac{\langle b_2, c_1 \rangle}{||c_1||^2} c_1 \qquad c_3 := b_3 - \frac{\langle b_3, c_1 \rangle}{||c_1||^2} c_1 - \frac{\langle b_3, c_2 \rangle}{||c_2||^2} c_2 \quad \ldots \quad c_n := b_n - \sum_{i=1}^{n-1} \frac{\langle b_n, c_i \rangle}{||c_i||^2} c_i$$

Przykład 9.3.5. Rozważmy przestrzeń $\mathbb{R}_2[x]$ z iloczynem skalarnym $\langle p, q \rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$. Dokonamy ortogonalizacji bazy $\mathcal{B} = \big(1, x, x^2\big)$.

Niech $b_1 = 1, b_2 = x, b_3 = x^2$. Zauważmy, że 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 $c_1 := b_1 = 1$.

II KROK: Poszukujemy $c_2 = b_2 + \alpha c_1$, $\alpha \in \mathbb{R}$. Dobierzmy α tak, by $c_2 \circ c_1 = 0$.
Obliczamy $0 = c_2 \circ c_1 = (b_2 + \alpha c_1) \circ c_1 = b_2 \circ c_1 + \alpha \cdot (c_2 \circ c_1) = (-1 \cdot 1 + 0 \cdot 1 + 1 \cdot 1) + \alpha(1+1+1) = 3\alpha$.
Skąd $\alpha = 0$. Zatem $c_2 = b_2 = x$. (Mogliśmy to zauważyć wcześniej.)

III KROK: Poszukujemy c_3 w postaci $c_3 = b_3 + \beta_1 c_1 + \beta_2 c_2$, $\beta_1, \beta_2 \in \mathbb{R}$. Dobierzmy β_1, β_2 tak, by $c_3 \circ c_1 = 0$ oraz $c_3 \circ c_2 = 0$.
Mamy $0 = c_3 \circ c_1 = (b_3 + \beta_1 c_1 + \beta_2 c_2) \circ c_1 = b_3 \circ c_1 + \beta_1 (c_1 \circ c_1) + \beta_2 (c_2 \circ c_1) = b_3 \circ c_1 + \beta_1 ||c_1||^2 = (1 \cdot 1 + 0 \cdot 1 + 1 \cdot 1) + \beta_1 \cdot 3$, skąd $\beta_1 = -\frac{2}{3}$.
Analogicznie $0 = c_3 \circ c_2 = (b_3 + \beta_1 c_1 + \beta_2 c_2) \circ c_2 = b_3 \circ c_2 + \beta_1 (c_1 \circ c_2) + \beta_2 (c_2 \circ c_2) = b_3 \circ c_2 + \beta_2 ||c_2||^2 = (1 \cdot (-1) + 0 \cdot 0 + 1 \cdot 1) + \beta_2 \cdot ((-1)^2 + 0^2 + 1^2) = 2\beta_2$, skąd $\beta_2 = 0$.
Zatem $c_3 = b_3 - \frac{2}{3} c_1 = x^2 - \frac{2}{3}$.
$\mathcal{C} = \big(c_1 = 1, c_2 = x, c_3 = x^2 - \frac{2}{3}\big)$ jest bazą ortogonalną.

$||c_1|| = \sqrt{1+1+1} = \sqrt{3}$, $||c_2|| = \sqrt{1+0+1} = \sqrt{2}$, $||c_3|| = \sqrt{\frac{1}{9}+\frac{4}{9}+\frac{1}{9}} = \frac{\sqrt{6}}{3}$
Bazą ortonormalną jest układ wektorów
$\mathcal{C}' = \left(c_1' = \frac{c_1}{||c_1||} = \frac{\sqrt{3}}{3}, c_2' = \frac{c_2}{||c_2||} = \frac{\sqrt{2}}{2}x, c_3' = \frac{c_3}{||c_3||} = \frac{\sqrt{6}}{2}(x^2 - \frac{2}{3}) = \frac{\sqrt{6}}{2}x^2 - \frac{\sqrt{6}}{3}\right)$.

9.4 Rzut ortogonalny na podprzestrzeń

Niech $(V, \langle .,. \rangle)$ będzie przestrzenią euklidesową, zaś $S \subset V$ dowolnym podzbiorem.

Definicja 9.4.1. 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 9.4.2. Niech V będzie przestrzenią euklidesową.

i) Jeśli $U \subset V$ jest podprzestrzenią liniową przestrzeni V, to wówczas $U^\perp$ również jest podprzestrznią liniową.

ii) Dla dowolnego wektora $u \in V$ zbiór $\{u\}^\perp$ jest podprzestrzenią liniową przestrzeni V.

Przykład 9.4.3. 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 9.4.4. 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 9.4.5. 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 9.4.6. 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 $w \perp U$?

$w \perp U = \mathbb{R}_1[x] = \text{lin}\{1, x\} \Leftrightarrow w \perp 1 \wedge w \perp x$
$\langle w, 1 \rangle = \int_0^1 (6x^2 - 6x + 1)dx = 2x^3 - 3x^2 + x\Big|_0^1 = 0$
$\langle w, x \rangle = \int_0^1 (6x^3 - 6x^2 + x)dx = \frac{3}{2}x^4 - 2x^3 + \frac{1}{2}x^2\Big|_0^1 = 0 \quad$ Zatem $w \perp U$.

Własności dopełnienia ortogonalnego

Twierdzenie 9.4.7. Niech V będzie przestrzenią euklidesową, zaś U, U_1, U_2 jej podprzestrzeniami liniowymi. Wówczas:

i) $U_1 \subset U_2 \Rightarrow U_2^{\perp} \subset U_1^{\perp}$,

ii) $U^{\perp} = (\mathrm{lin} U)^{\perp}$,

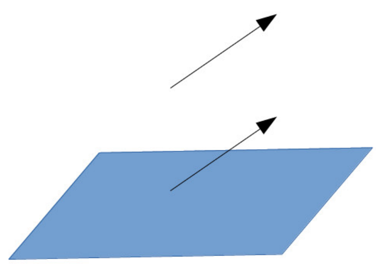
iii) $U \subset (U^{\perp})^{\perp}$.

Niech V będzie przestrzenią euklidesową, zaś U jej podprzestrzenią liniową.

Definicja 9.4.8. 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.



Twierdzenie 9.4.9 (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 $\mathcal{B} = \big(b_1, b_2, \ldots, b_k\big)$ 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 $\mathcal{B} = \big(b_1, b_2, \ldots, b_k\big)$ jest bazą ortogonalną podprzestrzeni U, wówczas rzut ten określony jest wzorem

$$u = \pi(v) = \frac{\langle v, b_1 \rangle}{||b_1||^2} b_1 + \frac{\langle v, b_2 \rangle}{||b_2||^2} b_2 + \ldots + \frac{\langle v, b_k \rangle}{||b_k||^2} b_k.$$

iii) Jeśli $\mathcal{B} = \big(b_1, b_2, \ldots, b_k\big)$ jest bazą ortonormalną podprzestrzeni U, wówczas

$$u = \pi(v) = \langle v, b_1\rangle b_1 + \langle v, b_2\rangle b_2 + \ldots + \langle v, b_k\rangle b_k.$$

Lub równoważnie, jeśli A jest macierzą, której kolumnami są kolejne wektory bazy ortonormalnej $\mathcal{B}$, to wówczas AA^T jest reprezentacją macierzową projekcji π_U.

Macierz $\begin{bmatrix} \langle b_1, b_1\rangle & \langle b_1, b_2\rangle & \ldots & \langle b_1, b_k\rangle \\ \langle b_2, b_1\rangle & \langle b_2, b_2\rangle & \ldots & \langle b_2, b_k\rangle \\ \ldots & & \ldots & \\ \langle b_k, b_1\rangle & \langle b_k, b_2\rangle & \ldots & \langle b_k, b_k\rangle \end{bmatrix}$ występującą w powyższym twierdzeniu nazywamy *macierzą Grama* układu wektorów $\big(b_1, b_2, \ldots, b_k\big)$.

Wniosek 9.4.10. Niech V będzie przestrzenią euklidesową, zaś U jej podprzestrzenią liniową. Niech $\mathcal{B} = \big(b_1, b_2, \ldots, b_k\big)$ 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 9.4.11. W przestrzeni euklidesowej $\mathbb{E}^4$ wyznaczymy rzut ortogonalny wektora $v = (1, 1, 1, 0)$ na podprzestrzeń $U = \mathrm{lin}\{(2, 1, 1, 2), (1, 1, -3, 0)\}$.

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
$u := \pi_U(v) =?, \quad w := v - u \perp U \Leftrightarrow w \perp b_1 := (2, 1, 1, 2) \wedge w \perp b_2 := (1, 1, -3, 0)$
$u \in U$, zatem istnieją $\alpha, \beta \in \mathbb{R}$ takie, że $u = \alpha b_1 + \beta b_2$. 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
$u = [\alpha, \beta]_{\mathcal{B}} = [\frac{2}{5}, -\frac{1}{11}]_{\mathcal{B}} = \frac{2}{5}b_1 - \frac{1}{11}b_2 = \big(\frac{4}{5}, \frac{2}{5}, \frac{2}{5}, \frac{4}{5}\big) - \big(\frac{1}{11}, \frac{1}{11}, \frac{-3}{11}, 0\big) = \big(\frac{39}{55}, \frac{17}{55}, \frac{37}{55}, \frac{4}{5}\big)$

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 9.2.10 mamy
$u = \Big[\frac{\langle u, b_1\rangle}{||b_1||^2}, \frac{\langle u, b_2\rangle}{||b_2||^2}\Big]$. Ale $\langle v, b_i\rangle = \langle u, b_i\rangle$, zatem $u = \Big[\frac{\langle v, b_1\rangle}{||b_1||^2}, \frac{\langle v, b_2\rangle}{||b_2||^2}\Big]$. 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 $AA^T \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{39}{55} \\ \frac{17}{55} \\ \frac{37}{55} \\ \frac{4}{5} \end{bmatrix}$.

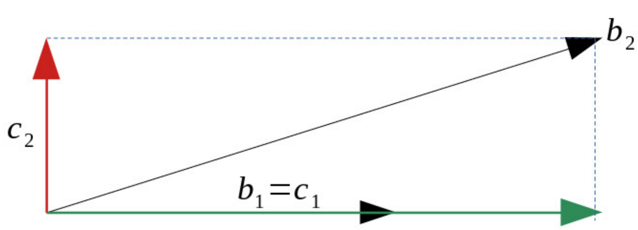
Interperatacja geometryczna metody Grama-Schmidta
Rozważmy przestrzeń $\mathbb{E}^3$ i jej bazę $\mathcal{B} = (b_1, b_2, b_3)$. Niech $\mathcal{C} = (c_1, c_2, c_3)$ będzie szukaną bazą ortogonalną.

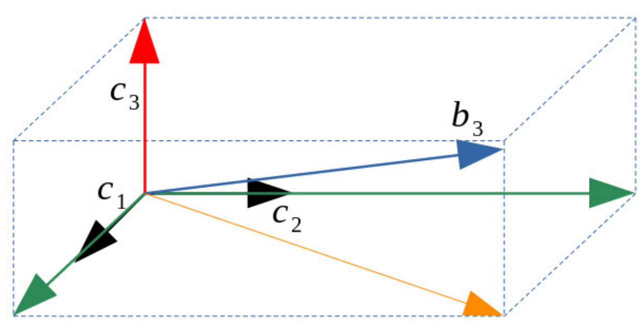
KROK I:

$$c_1 := b_1$$

KROK II:

$$c_2 := b_2 - \frac{\langle b_2, c_1\rangle}{||c_1||^2} c_1 = b_2 - \pi_{\mathrm{lin}\{c_1\}}(b_1)$$





KROK III:

$$c_3 := b_3 - \frac{\langle b_3, c_1\rangle}{||c_1||^2} c_1 - \frac{\langle b_3, c_2\rangle}{||c_2||^2} c_2 = b_3 - \pi_{\mathrm{lin}\{c_1\}}(b_3) - \pi_{\mathrm{lin}\{c_2\}}(b_3) = b_3 - \pi_{\mathrm{lin}\{c_1, c_2\}}(b_3)$$

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.

Definicja 9.5.2. Niech $n \in \mathbb{N}$. Macierz $A \in M_n(\mathbb{R})$ nazywamy *macierzą ortogonalną*, jeśli $A^T A = A^T A = I_n$.

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 (względem standardowego iloczynu skalarnego w $\mathbb{R}^n$). 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 $1 = \det I_n = \det(AA^T) = (\det A)^2$. $\square$

Niech $(V, \langle ., . \rangle)$ będzie przestrzenią euklidesową, skończenie wymiarową. Rozpatrujemy normę zadaną przez iloczyn skalarny.

Definicja 9.5.6. Operator liniowy $\varphi \in End(V)$ nazywamy

i) *ortogonalnym*, jeśli jego macierz w pewnej bazie ortonormalnej przestrzeni V jest macierzą ortogonalną,

ii) *izometrią liniową*, jeśli zachowuje on odległość, tzn. spełniony jest warunek

$$\forall u, v \in V \quad ||\varphi(u) - \varphi(v)|| = ||u - v||.$$

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 9.5.7. 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 $v \in \mathrm{Ker}\varphi$, to $0 = ||\mathbf{0}_V|| = ||\varphi(v)|| = ||v||$, zatem $v = \mathbf{0}_V$. 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 9.5.8. 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

Wniosek 9.5.9. 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.

Przykład 9.5.10. 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.

Uwaga 9.5.11. Jeśli rozpatrujemy niestandardowy iloczyn skalarny, baza kanoniczna nie musi być ortonormalna. Nie zawsze zatem wystarczy badać macierz w bazie kanonicznej.

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 $\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}$.

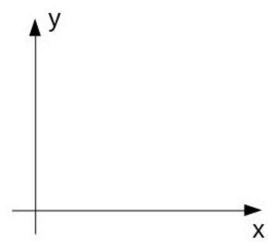
$\boxed{\mathbf{n = 1}}$

$\varphi : \mathbb{R} \to \mathbb{R}$ 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 $\varphi_2(x) = -x$.

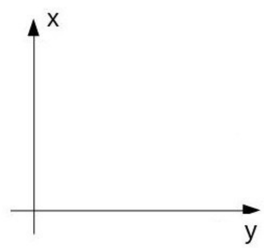
$\boxed{\mathbf{n = 2}}$

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.

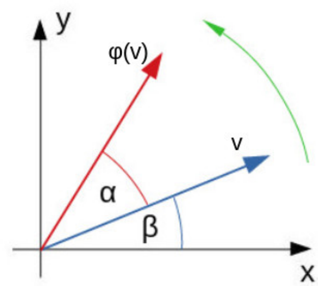
Podejście algebraiczne
Z założenia $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ jest linowe, zatem możemy rozważyć jego macierz w bazie standardowej $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \in M_2(\mathbb{R})$. Wiemy, że macierz A jest ortogonalna oraz $\det A = \pm 1$.
Z równości $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = A^T A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} a^2+b^2 & ab+cd \\ ab+cd & c^2+d^2 \end{bmatrix}$
otrzymujemy układ równań $\begin{cases} a^2+b^2 = 1 \\ c^2+d^2 = 1 \\ ab+cd = 0 \end{cases}$.
Punkty $(a,b), (c,d)$ leżą na okręgu jednostkowym, zatem istnieje taki kąt α, że $a = \cos\alpha$ oraz $b = \sin\alpha$. Wówczas na mocy warunku ortogonalności $ab+cd = 0$ mamy, że $c = \cos\left(\frac{\pi}{2}+\alpha\right) = -\sin\alpha$, $d = \sin\left(\frac{\pi}{2}+\alpha\right) = \cos\alpha$ lub też $c = \cos\left(\frac{3}{2}\pi+\alpha\right) = \sin\alpha$, $d = \sin\left(\frac{3}{2}\pi+\alpha\right) = -\cos\alpha$. Podsumowując,

$$A = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \quad \textit{lub} \quad A = \begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix} = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}.$$

1) W pierwszym przypadku $\det A = a^2 + b^2 = 1$. 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 R_α lub $R(\alpha)$.

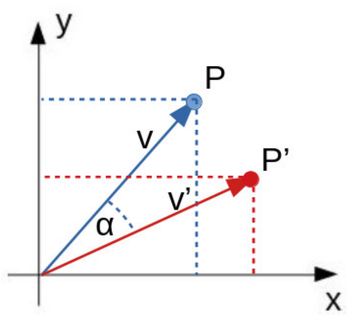


Macierzą obrotu w kierunku zgodnym z ruchem wskazówek zegara (obrotu o kąt $-\alpha$) będzie

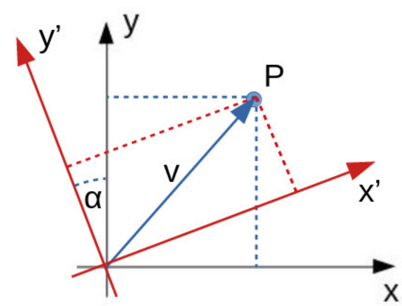
$$A^{-1} = A^T = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}.$$

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 9.5.12. 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.

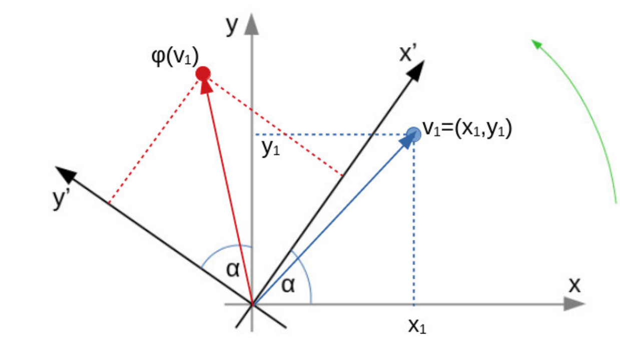


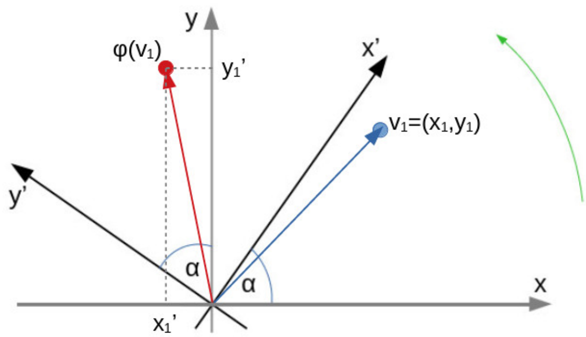
Punkt (wektor) są nieruchome. Obracamy układ współrzędnych, tzn. dokonujemy zmiany bazy. W nowej bazie punkt ma nowe współrzędne.

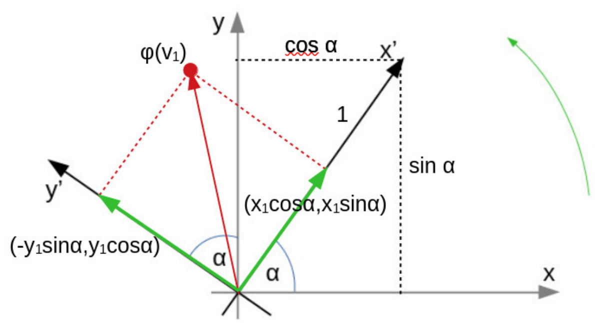
Współrzędne punktu P' w “starej bazie” (rysunek po lewej) są takie same, jak współrzędne punktu P w “nowej bazie” (rysunek po prawej).

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 9.5.13.

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$ identyczność, tj. rotacja o kąt miary 0

$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix}$ rotacja o kąt miary $\frac{\pi}{2}$

$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ -y \end{bmatrix}$ rotacja o kąt miary π

$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ -x \end{bmatrix}$ rotacja o kąt miary $\frac{3}{2}\pi$

Przykład 9.5.14. Dana jest macierz $A = \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix} \in M_2(\mathbb{R})$.
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 przeciwnym do ruchu wskazówek zegara o kąt $\frac{4}{3}\pi$ lub (równoważnie) obrót w kierunku zgodnym z ruchem wskazówek zegara o kąt $\frac{2}{3}\pi$.

Przykład 9.5.15. 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 $(m+1)x + (m-1)y = 0$.

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 $y = \operatorname{tg}\beta \cdot x$. 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 $\varphi(v) = c_1b_1 - c_2b_2$.

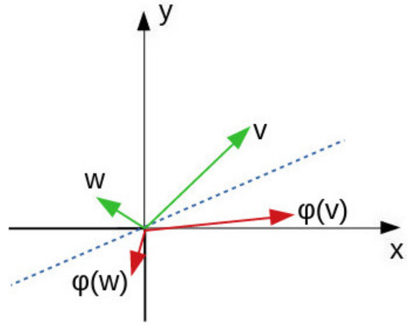
$$\varphi(v)^T = c_1\begin{bmatrix} \cos\beta \\ \sin\beta \end{bmatrix} - c_2\begin{bmatrix} -\sin\beta \\ \cos\beta \end{bmatrix} = \begin{bmatrix} c_1\cos\beta + c_2\sin\beta \\ c_1\sin\beta - c_2\cos\beta \end{bmatrix} = \begin{bmatrix} \cos\beta & \sin\beta \\ \sin\beta & -\cos\beta \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

Stąd

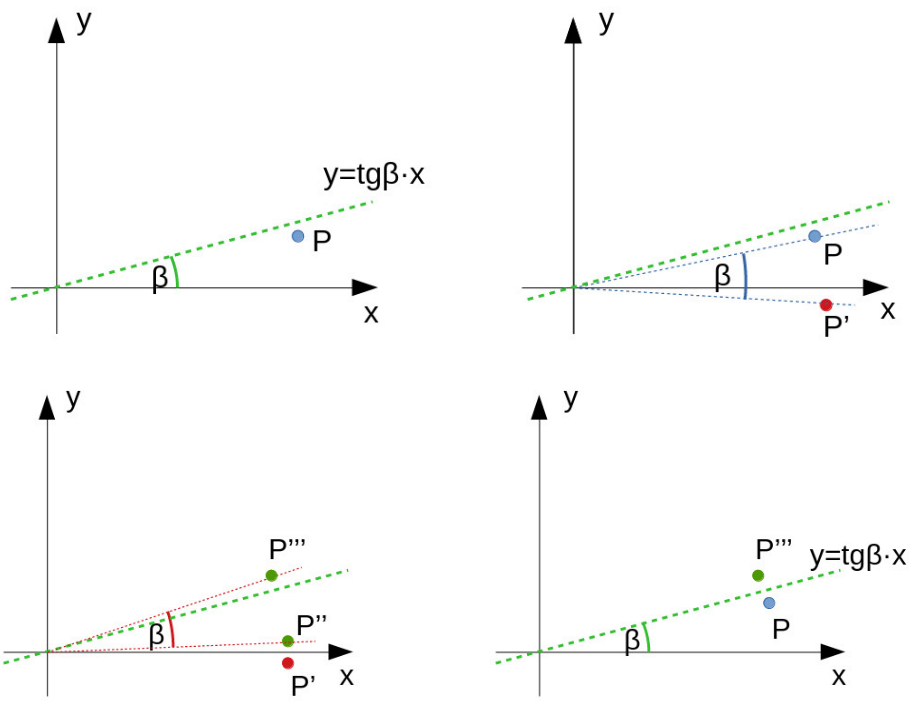
$$\varphi(v)^T = \begin{bmatrix} \cos\beta & \sin\beta \\ \sin\beta & -\cos\beta \end{bmatrix}\begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix}^{-1} v^T =$$

$$= \begin{bmatrix} \cos\beta & \sin\beta \\ \sin\beta & -\cos\beta \end{bmatrix}\begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} v^T = \begin{bmatrix} \cos 2\beta & \sin 2\beta \\ \sin 2\beta & -\cos 2\beta \end{bmatrix} v^T$$

Czyli $\alpha = 2\beta$ i macierz $\begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix}$ reprezentuje odbicie względem prostej o równaniu $y = \operatorname{tg}\frac{\alpha}{2} \cdot x$.



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 $-\beta$.

$$\begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix}$$

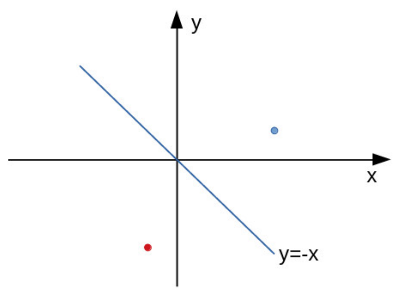
4) Otrzymujemy

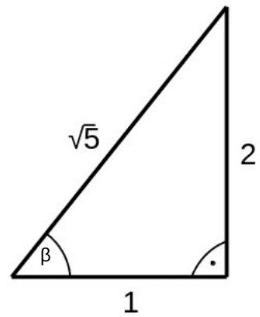
$$\begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} =$$

$$= \begin{bmatrix} \cos\beta & \sin\beta \\ \sin\beta & -\cos\beta \end{bmatrix} \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} = \begin{bmatrix} \cos 2\beta & \sin 2\beta \\ \sin 2\beta & -\cos 2\beta \end{bmatrix}.$$

Wniosek 9.5.16. 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 9.5.17. 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 9.5.18. Wyznaczymy macierz izometrii liniowej płaszczyzny będącej symetrią ortogonalną względem prostej $y = 2x$.

METODA I: Jeśli $\mathrm{tg}\beta = 2$, 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 9.5.19. 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 9.5.20. Dana jest macierz $A = \begin{bmatrix} -\frac{3}{5} & -\frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{bmatrix} \in M_2(\mathbb{R})$.
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 9.5.21. 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 $A^{-1} = A$.

$\boxed{\textbf{n} - \textbf{dowolne}}$

Twierdzenie 9.5.22. Dla dowolnego $n \in \mathbb{N}$ macierz obrotu w przestrzeni $\mathbb{R}^n$ jest macierzą ortogonalną o wyznaczniku równym 1. I odwrotnie, jeśli macierz $A \in M_n(\mathbb{R})$ jest macierzą ortogonalną o wyznaczniku równym 1, to A jest macierzą obrotu.

Istnieje zatem wzajemnie jednoznaczna korespondencja pomiędzy obrotami a macierzami ortogonalnymi o wyznaczniku równym 1.

Twierdzenie 9.5.23. Dla ustalonego $n \in \mathbb{N}$ zbiór macierzy ortogonalnych stopnia n o wyznaczniku równym 1 wraz z działaniem mnożenia macierzy tworzy grupę. Grupę tę nazywamy *specjalną grupą ortogonalną* lub *grupą obrotów właściwych* i oznaczamy symbolem $SO(n)$.

Wniosek 9.5.24. Złożenie rotacji jest rotacją. Ponadto odwzorowanie odwrotne do rotacji jest rotacją.

Twierdzenie 9.5.25.

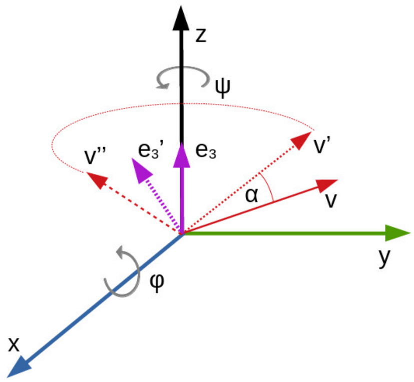
i) Grupa $SO(2)$ jest abelowa.

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

Przykład 9.5.26. Niech $A = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$.
Ponieważ $\det A = \det B = 1$, są to macierze obrotów. Mamy

$$AB = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix} \neq BA = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Uwaga 9.5.27. Składaniu odwzorowań odpowiada mnożenie reprezentujących je macierzy (w pewnej ustalonej bazie). Oznacza to, że osie obrotu traktujemy jako *niezmienne*, z góry zadane. Przykładowo, jeśli $\varphi, \psi \in End(\mathbb{R}^3)$ to odpowiednio rotacja o kąt γ wokół prostej l i rotacja o kąt α wokół prostej k, zaś $A, B \in SO(3)$ to reprezentujące je macierze (w bazie kanonicznej $\mathcal{B}_k^3$), to wówczas AB reprezentuje złożenie $\varphi \circ \psi$. Dany punkt $P_0 = (x_0, y_0, z_0)$ zostanie najpierw obrócony wokół osi k. Otrzymamy punkt P_0'. Kolejno P_0' zostanie obrócony wokół osi l (l nie została obrócona, pozostaje niezmienna). Otrzymamy punkt P_0''. Mnożąc $AB \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$ wyznaczymy współrzędne punktu P_0'' w bazie $\mathcal{B}_k^3$.



k - oś Ox, l - oś Oz

Orientacja rzeczywistej przestrzeni wektorowej

Niech V będzie skończenie wymiarową przestrzenią euklidesową. Niech $\mathcal{B}, \mathcal{B}'$ to dwie różne uporządkowane bazy przestrzeni V. Każda baza wyznacza orientację przestrzeni V. Niech $P = P_{\mathcal{B}\to\mathcal{B}'}$. Bazy $\mathcal{B}, \mathcal{B}'$ wyznaczają *tę samą orientację* lub są *zgodnie zorientowane*, jeśli $\det P > 0$. Gdy $\det P < 0$, bazy są *przeciwnie zorientowane.*

Własność posiadania tej samej orientacji jest relacją równoważności w zbiorze reprerów bazowych przestrzeni V. Istnieją dokładnie dwie klasy abstrakcji. Mówimy o *orientacji dodatniej i ujemnej.* Wybór orientacji polega na wyborze repera bazowego (jego klasy abstrakcji). Przyjmujemy, że baza kanoniczna $\mathbb{R}^n$ jest zorientowana dodatnio.

Definicja 9.5.28. Niech V będzie skończenie wymiarową przestrzenią euklidesową. Mówimy, że automorfizm $\varphi \in Aut(V)$ *zachowuje orientację*, jeśli dla dowolnej bazy $(b_1, b_2, \ldots, b_n)$ przestrzeni V zorientowanej dodatnio (ujemnie), baza $(\varphi(b_1), \varphi(b_2), \ldots, \varphi(b_n))$ również jest zorientowana dodatnio (ujemnie). W przeciwnym wypadku mówimy, że φ *odwraca orientację.*

Oznacza to w szczególności, że w $\mathbb{R}^2$ i $\mathbb{R}^3$ układy prawoskrętne przechodzą na układy prawoskrętne. Łatwo zauważyć, że φ zachowuje orientację (odwraca orientację) wtedy i tylko wtedy, gdy macierz odwzorowania φ ma *dodatni (ujemny) wyznacznik.*

Twierdzenie 9.5.29. Niech $(V, \langle ., . \rangle)$ będzie przestrzenią euklidesową n wymiarową. Dla każdego operatora ortogonalnego $\varphi \in End(V)$ istnieje taka baza ortonormalna przestrzeni V, w której macierz operatora jest macierzą blokowo-diagonalną postaci

$$\begin{bmatrix} \cos\alpha_1 & -\sin\alpha_1 & & & & & \\ \sin\alpha_1 & \cos\alpha_1 & & & & & \\ & & \ddots & & & & \\ & & & \cos\alpha_r & -\sin\alpha_r & & \\ & & & \sin\alpha_r & \cos\alpha_r & & \\ & & & & & I_k & \\ & & & & & & -I_l \end{bmatrix}, \quad k + l + 2r = n.$$

Rotacje Givensa

Macierz ortogonalna postaci

$$P_{ij} = \begin{array}{c} \begin{array}{ccccccccc} & & k_i & & & k_j & & & \end{array} \\ \begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & c & & & s & & & \\ & & & 1 & & & & & \\ & & & & \ddots & & & & \\ & & -s & & & c & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{bmatrix} \end{array},$$

gdzie $c^2 + s^2 = 1$ reprezentuje rotację w płaszczyźnie Ox_ix_j w $\mathbb{R}^n$. Rotacje te nazywamy *rotacjami Givensa.*

$$P_{ij} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ cx_i + sx_j \\ \vdots \\ -sx_i + cx_j \\ \vdots \\ x_n \end{bmatrix} \begin{array}{c} \\ \\ w_i \\ \\ w_j \\ \\ \\ \end{array}.$$

Jeśli $x_i^2 + x_j^2 \neq 0$, to przyjmując $c = \frac{x_i}{\sqrt{x_i^2+x_j^2}}$, $s = \frac{x_j}{\sqrt{x_i^2+x_j^2}}$, otrzymujemy

$$P_{ij} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ \sqrt{x_i^2 + x_j^2} \\ \vdots \\ 0 \\ \vdots \\ x_n \end{bmatrix} \begin{array}{c} \\ \\ w_i \\ \\ w_j \\ \\ \\ \end{array}.$$

Zatem poprzez obrót w płaszczyźnie możemy *wyzerować wybrane współrzędne.*

Wniosek 9.5.30. Dla dowolnego wektora $x \in \mathbb{R}^n$ i dla dowolnego $i \in \{1, 2, \ldots, n\}$ istnieje macierz $P \in O(n)$ taka, że $Px^T = ||x||e_i^T$, gdzie $P = P_{in}P_{i,n-1} \ldots P_{i,i+1}P_{i,i-1} \ldots P_{i2}P_{i1}$

Przykład 9.5.31 (Konstruowanie bazy ortonormalnej zawierającej dany wersor). Wykorzystując rotacje Givensa, skonstruujemy bazę ortonormalną przestrzeni $\mathbb{R}^3$, zawierającą wersor $x = \frac{1}{\sqrt{3}}(1,1,1)$.

Obliczamy $\frac{x_1}{\sqrt{x_1^2+x_2^2}} = \frac{x_2}{\sqrt{x_1^2+x_2^2}} = \frac{1}{\sqrt{2}}$ oraz $P_{12}x^T = \frac{1}{\sqrt{3}}\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{3}}\begin{bmatrix} \sqrt{2} \\ 0 \\ 1 \end{bmatrix}$.

Niech $w = (w_1, w_2, w_3) = \frac{1}{\sqrt{3}}(\sqrt{2}, 0, 1)$. Wówczas $\frac{w_1}{\sqrt{w_1^2+w_3^2}} = \sqrt{\frac{2}{3}}$, $\frac{w_3}{\sqrt{w_1^2+w_3^2}} = \frac{1}{\sqrt{3}}$ oraz

$$P_{13}P_{12}x^T = \frac{1}{\sqrt{3}}\begin{bmatrix} \sqrt{\frac{2}{3}} & 0 & \sqrt{\frac{1}{3}} \\ 0 & 1 & 0 \\ -\sqrt{\frac{1}{3}} & 0 & \sqrt{\frac{2}{3}} \end{bmatrix}\begin{bmatrix} \sqrt{2} \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = e_1^T.$$

Macierz $P = P_{13}P_{12}$ jest ortogonalna, zatem $x^T = P^{-1}e_1^T = P^Te_1^T = P_{12}^TP_{13}^Te_1^T$.

Kolumny macierzy $P^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \sqrt{\frac{2}{3}} & 0 & -\sqrt{\frac{1}{3}} \\ 0 & 1 & 0 \\ \sqrt{\frac{1}{3}} & 0 & \sqrt{\frac{2}{3}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{bmatrix}$

tworzą bazę ortonormalną. W pierwszej kolumnie znajdują się współrzędne wektora x.

$\boxed{\mathbf{n = 3}}$

Wniosek 9.5.32. Dla każdego operatora ortogonalnego $\varphi \in End(\mathbb{R}^3)$ istnieje taka baza ortonormalna przestrzeni $\mathbb{R}^3$, w której macierz operatora jest macierzą postaci $\begin{bmatrix} A & 0 \\ 0 & \pm 1 \end{bmatrix}$, gdzie A jest reprezentacją macierzową izometrii liniowej płaszczyzny $\mathbb{R}^2$.

Niech $\mathcal{B} = (b_1, b_2, b_3)$ będzie bazą ortonormalną taką, że $M_\varphi(\mathcal{B}, \mathcal{B}) = \begin{bmatrix} A & 0 \\ 0 & \pm 1 \end{bmatrix}$. Możliwe są następujące przypadki.

i) Jeśli $\varphi(b_3) = b_3$, zaś A jest macierzą rotacji, wówczas φ jest *obrotem wokół osi Ob_3*.

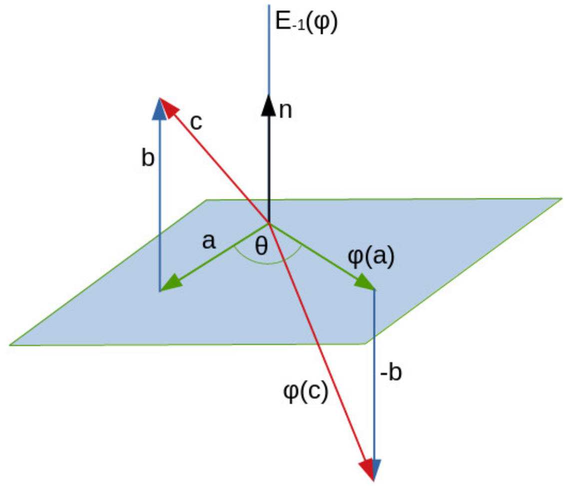
ii) Jeśli $\varphi(b_3) = b_3$, zaś A jest macierzą odbicia względem prostej l leżącej w płaszczyźnie Ob_1b_2, wówczas φ jest *odbiciem względem płaszczyzny* zawierającej b_3 i prostą l.

W szczególności, gdy b_2 leży na prostej l wówczas $M_\varphi(\mathcal{B}, \mathcal{B}) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ i φ jest odbiciem względem płaszczyzny Ob_2b_3.

iii) Jeśli $\varphi(b_3) = -b_3$, zaś A jest macierzą rotacji na płaszczyźnie Ob_1b_2, wówczas φ jest *złożeniem obrotu wokół osi Ob_3 i odbicia względem płaszczyzny prostopadłej do osi Ob_3*.

iv) Jeśli $\varphi(b_3) = -b_3$, zaś A jest macierzą odbicia, wówczas można tak dobrać bazę, by macierz odwzorowania φ miała postać $\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Wówczas φ jest *obrotem o kąt π wokół osi Ob_3*.

Gdy $\det M_\varphi(\mathcal{B},\mathcal{B}) > 0$, mamy do czynienia z obrotem, a gdy $\det M_\varphi(\mathcal{B},\mathcal{B}) < 0$, mamy do czynienia z odbiciem lub złożeniem obrotu wokół prostej i odbicia względem płaszczyzny ortogonalnej do osi obrotu (kolejność składania nie ma znaczenia). Rozróżnienie między odbiciem a złożeniem obrotu i odbicia możliwe jest poprzez analizę wartości własnych φ (które nie zależą od wyboru bazy). Dla odbicia wartościami własnymi są $1, 1, -1$, zaś dla złożenia obrotu i odbicia wartości własne to $-1, -1, -1$ lub -1 jest jedyną rzeczywistą wartością własną.



Jeśli -1 jest jedyną rzeczywistą wartością własną, zaś n odpowiadającym jej wektorem własnym, wówczas $\varphi(n) = -n$. Ponadto dla dowolnego wektora u leżącego w płaszczyźnie prostopadłej do n i przechodzącej przez początek układu współrzędnych wektor $\varphi(u)$ otrzymujemy poprzez obrót wektora u o kąt θ. Dowolny wektor c jest sumą wektora b równoległego do n i wektora a prostopadłego do n. Mamy $\varphi(c) = \varphi(a+b) = \varphi(a)+\varphi(b) = \varphi(a) - b$.

Odbicie φ jest inwolucją, tzn. $\varphi \circ \varphi = \mathrm{id}$. Równoważnie, jeśli A jest macierzą odbicia φ, to wówczas $A^2 = I$ lub inaczej $A^{-1} = A$.

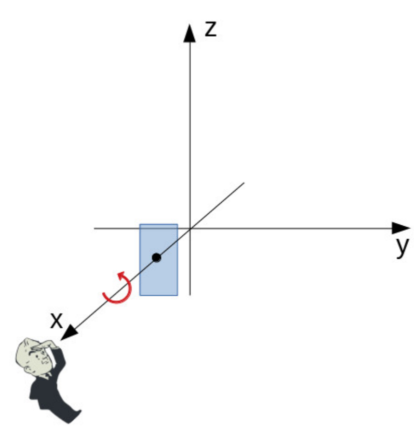
Jeśli $AA^T = I$, $\det A = 1$ oraz $A^2 = I$, to mamy do czynienia z rotacją o kąt π i macierz A jest symetryczna, tj. $A = A^T$.

Obroty podstawowe (elementarne) w $\mathbb{R}^3$

Zakładamy, że mamy do czynienia z układem prawoskrętnym. Rozpatrujemy standardowy iloczyn skalarny.

I. **Obrót o kąt α względem osi Ox (w płaszczyźnie Oyz)** jest reprezentowany przez macierz

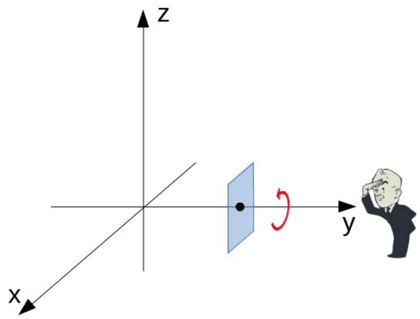
$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}.$$



$$R_x(\alpha)\begin{bmatrix}1\\0\\0\end{bmatrix}=\begin{bmatrix}1\\0\\0\end{bmatrix}, R_x(\alpha)\begin{bmatrix}0\\1\\0\end{bmatrix}=\begin{bmatrix}0\\\cos\alpha\\\sin\alpha\end{bmatrix}, R_x(\alpha)\begin{bmatrix}0\\0\\1\end{bmatrix}=\begin{bmatrix}0\\-\sin\alpha\\\cos\alpha\end{bmatrix}$$

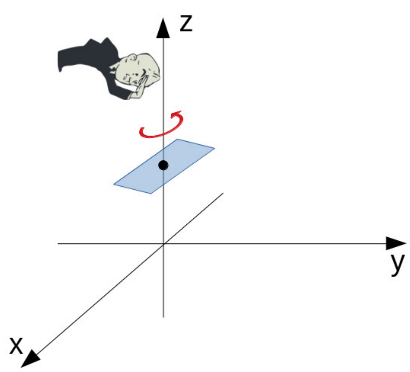
II. **Obrót o kąt β względem osi Oy (w płaszczyźnie Oxz)** jest reprezentowany przez macierz

$$R_y(\beta)=\begin{bmatrix}\cos\beta & 0 & \sin\beta\\ 0 & 1 & 0\\ -\sin\beta & 0 & \cos\beta\end{bmatrix}.$$



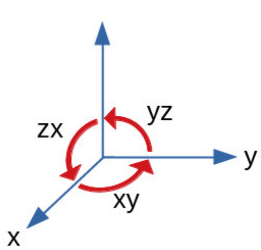
III. **Obrót o kąt γ względem osi Oz (w płaszczyźnie Oxy)** jest reprezentowany przez macierz

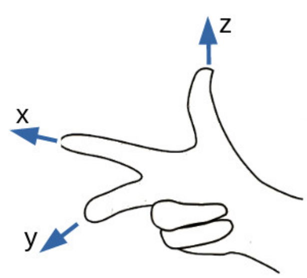
$$R_z(\gamma)=\begin{bmatrix}\cos\gamma & -\sin\gamma & 0\\ \sin\gamma & \cos\gamma & 0\\ 0 & 0 & 1\end{bmatrix}.$$



Powyższe trzy przykłady są przykładami *rotacji planarnej*, tj. endomorfizmu przestrzeni $\mathbb{R}^n$ będącego rotacją w pewnej płaszczyźnie $W \subset \mathbb{R}^n$ i pozostawiającego wektory należące do $W^\perp$ niezmienione.

Uwaga 9.5.33. Odmienny układ znaków w macierzy $R_y(\beta)$ wynika z faktu, że obrót odbywa się w płaszczyźnie zgodnie z ruchem wskazówek zegara (widok przy osi obrotu skierowanej ku nam).

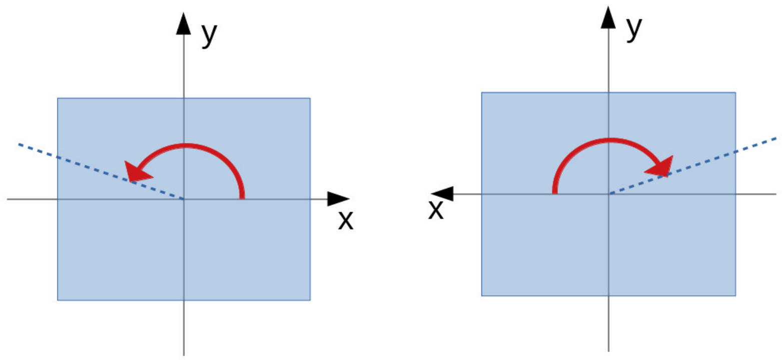




Obroty w $\mathbb{R}^3$ o dowolny kąt wokół dowolnej osi przechodzącej przez początek układu współrzędnych

Symbolem $R_u(\alpha)$ będziemy oznaczać macierz rotacji w $\mathbb{R}^3$ wokół osi wyznaczonej przez wektor u o kąt α, stosując przy tym regułę prawej dłoni W szczególności

$$R_u(\alpha) = R_{-u}(-\alpha).$$



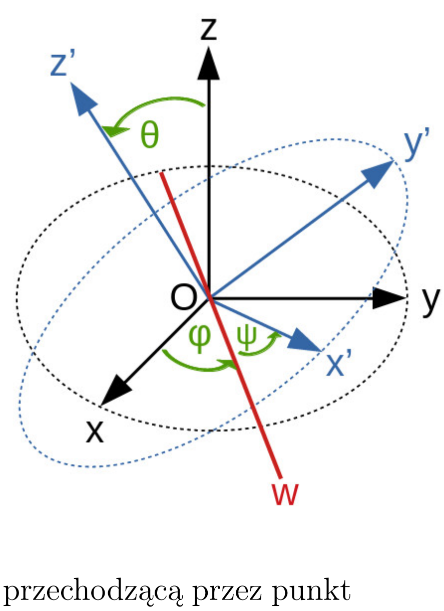
widok z dodatniej półosi Oz widok z ujemnej półosi Oz

Rotacja *nie zmienia orientacji* układu współrzędnych, a zatem *zachowuje iloczyn wektorowy*, tzn. długości i zwroty wektorów, jak również kąty między nimi pozostają niezmienione. Inaczej mówiąc, dla dowolnej rotacji φ i dla dowolnych wektorów v, w zachodzi

$$\varphi(v \times w) = \varphi(v) \times \varphi(w).$$

Kąty Eulera

Dowolnie zorientowany układ współrzędnych $Ox'y'z'$ można otrzymać z danego układu $Oxyz$ (o tej samej skrętności) przez złożenie trzech obrotów wokół osi układu.



PRZYPADEK I:
Obrót nie zachowuje zwrotu ani kierunku osi Oz (osie Oz i Oz' nie są równoległe).

Każdemu obrotowi układu współrzędnych w $\mathbb{R}^3$, nie zachowującemu zwrotu ani kierunku osi Oz, można wzajemnie jednoznacznie przypisać uporządkowaną trójkę kątów $(\varphi, \psi, \theta) \in [0, 2\pi) \times [0, 2\pi) \times (0, \pi)$.

Aby dokonać obrotu układu współrzędnych $Oxyz$ tak, by otrzymać nowy układ $Ox'y'z'$ postępujemy następująco:

i) Wyznaczamy prostą w prostopadłą do płaszczyzny Ozz' i przechodzącą przez punkt O. Prosta w jest nazywana *osią węzłów*.

ii) Rozpatrujemy kąty Eulera.

φ – kąt mierzony od osi Ox do osi węzłów w kierunku wyznaczonym osią Oz
θ – kąt mierzony od osi Oz do Oz' w kierunku wyznaczonym osią węzłów
ψ – kąt mierzony od osi węzłów do osi Ox' w kierunku wyznaczonym osią Oz'

iii) Wykonujemy w podanej kolejności trzy obroty.

1) Obracamy układ wokół osi Oz o kąt φ. Oś Ox staje się osią węzłów.
Istnieją dwa takie obroty o kąty różniące się o π, nadające osi Ox przeciwne zwroty. Wybieramy zwrot zgodny ze zwrotem iloczynu wektorowego wersorów osi Oz i Oz' (przyjmując go za zwrot osi węzłów).

2) Obracamy układ wokół osi węzłów (nowa oś Ox) o kąt θ. Oś Oz pokrywa się z osią Oz'. Zauważmy, że będzie to obrót o kąt z zakresu $(0, \pi)$.

3) Obracamy układ wokół osi Oz' o kąt ψ. Oś węzłów pokrywa się z osią Ox' , oś Oy pokrywa się z osią Oy'.

Macierz A rotacji będącej złożeniem trzech powyższych obrotów ma postać

$$A = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{bmatrix}.$$

PRZYPADEK II:
Osie Oz i Oz' są równoległe (identyczne lub o przeciwnych zwrotach).

Płaszczyzna Ozz' i linia węzłów nie są wówczas jednoznacznie określone. Możliwe jest przekształcenie osi Oz na oś Oz' poprzez obrót wokół dowolnej prostej przechodzącej przez punkt O i leżącej w płaszczyźnie $Oxy = Ox'y'$. Będzie to obrót o kąt 0 gdy osie są identyczne lub o kąt π, gdy mają przeciwny zwrot. Stąd $\theta = 0$ lub $\theta = \pi$. Ponadto suma lub różnica kątów φ i ψ wyznacza jednoznacznie ustawienie osi Ox' oraz Oy'.

Uwaga 9.5.34. Niekiedy zachodzi potrzeba rozkładu danego obrotu na złożenie trzech obrotów podstawowych. Możliwe są inne trzykrotne złożenia, niż to przedstawione powyżej. Wybór najczęściej podyktowany jest zastosowaniami. Jeśli znamy macierz danego obrotu, możliwe jest odczytanie z niej kątów Eulera. Zobacz tutaj i tutaj.

Obliczanie osi i kąta obrotu na podstawie macierzy

Na mocy twierdzenia 3.1.13 dla dowolnego $n \in \mathbb{N}$ oraz dla dowolnych $A, B \in M_n(\mathbb{R})$ zachodzi równość $\mathrm{tr}(AB) = \mathrm{tr}(BA)$. Jeśli zatem P jest macierzą nieosobliwą stopnia n, to wówczas $\mathrm{tr}(P^{-1}AP) = \mathrm{tr}(PP^{-1}A) = \mathrm{tr}A$.

Twierdzenie 9.5.35. Dla dowolnego obrotu $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ liczba 1 jest wartością własną. Ponadto jeśli $\varphi \neq \mathrm{id}_{\mathbb{R}^3}$, to wówczas wektor własny odpowiadający wartości własnej 1 jest wektorem kierunkowym osi obrotu.

Przykład 9.5.36. W $\mathbb{R}^3$ rozpatrujemy standardowy iloczyn skalarny. Niech $A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ będzie macierzą endomorfizmu $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ w bazie kanonicznej.
Jest to izometria liniowa, ponieważ macierz $A \in O(3)$. Istotnie

$$AA^T = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = I = A^T A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Obliczamy $\det A = 1$. Mamy zatem do czynienia z obrotem. Znajdziemy oś obrotu i kąt obrotu.

<u>Oś obrotu</u>
Wiemy, że 1 jest wartością własną φ.

$$(A - I) \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow x = y = z$$

Zatem $E_1 = \mathrm{lin}\{(1,1,1)\}$ i osią obrotu jest prosta $x = y = z$.

<u>Kąt obrotu</u>
METODA I: Zauważmy, że $(1,1,1) \circ (x,y,z) = 0$, gdy $x+y+z = 0$. Zatem $\pi_1 : x+y+z = 0$ jest płaszczyzną prostopadłą do osi obrotu. Ponadto jeśli $(1,1,1) \circ (x,y,z) = 0$, to $(1,1,1) \circ \varphi(x,y,z) = (1,1,1) \circ (z,x,y) = z+x+y = 0$. Zatem φ jest rotacją w płaszczyźnie π_1. Weźmy $w = (1,1,-2) \in \pi_1$. Obliczamy

$$\cos \measuredangle(w, \varphi(w)) = \frac{w \circ \varphi(w)}{|w| \cdot |\varphi(w)|} = \frac{(1,1,-2) \circ (-2,1,1)}{\sqrt{6} \cdot \sqrt{6}} = \frac{-3}{6} = -\frac{1}{2}$$

Zatem $\measuredangle(w, \varphi(w))$ ma miarę $\frac{2}{3}\pi$ lub $-\frac{2}{3}\pi$. A ponieważ $w \times \varphi(w) = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 1 & 1 & -2 \\ -2 & 1 & 1 \end{vmatrix} =$ $(3,3,3)$, więc mamy do czynienia z rotacją o kąt $\frac{2}{3}\pi$.

METODA II: A jest macierzą obrotu o pewien kąt α, zatem $\boxed{\mathrm{tr} A = 1 + 2\cos\alpha}$ (ponieważ ślad nie zależy od wyboru bazy). Obliczamy $\mathrm{tr} A = 0$, skąd $1 + 2\cos\alpha = 0$. Czyli $\cos\alpha = -\frac{1}{2}$ jak powyżej.

Uwaga 9.5.37. Dokonując obrotu zgodnie z regułą prawej dłoni, należy właściwe wybrać kąt, tak by odpowiadał wybranemu wersorowi osi obrotu. Jeśli u to wersor osi obrotu, zaś

w dowolny wektor, to wektory $u, w \times \varphi(w)$ są równoległe. W szczególności pokrywają się (gdy mają ten sam zwrot) lub kąt między nimi ma miarę π radianów (gdy maja przeciwne zwroty). Aby rozstrzygnąć, z którym przypadkiem mamy do czynienia, możemy obliczyć iloczyn skalarny $u \circ (w \times \varphi(w))$. Wynik dodatni wskazuje ma kąt 0 (i obrót przeciwnie do ruchu wskazówek zegara, tj. kąt dodatni), zaś wynik ujemny na kąt π (i obrót przeciwnie do ruchu wskazówek zegara, tj. kąt ujemny). Obrót w płaszczyźnie zawierającej w i $\varphi(w)$ odbywa się przeciwnie do ruchu wskazówek zegara, jeśli patrzymy tak, że wektor $w \times \varphi(w)$ skierowany jest ku nam.

Jak znaleźć macierz rotacji, gdy dane są oś i kąt obrotu?

Twierdzenie 9.5.38 (**Formuła Rodriguesa dla rotacji**). Niech v będzie dowolnym wektorem w $\mathbb{R}^3$, zaś u wersorem wskazującym oś obrotu φ o kąt θ (zgodnie z regułą prawej dłoni). Wówczas

$$\varphi(v) = v\cos\theta + (u \times v)\sin\theta + u(u \circ v)(1 - \cos\theta).$$

Formułę Rodriguesa dla rotacji można również zapisać w postaci macierzowej. Niech $u = (u_x, u_y, u_z), v = (v_x, v_y, v_z)$. Wówczas

$u \times v = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{vmatrix} = (u_y v_z - u_z v_y, u_z v_x - u_x v_z, u_x v_y - u_y v_x)$ oraz

$$\begin{bmatrix} u_y v_z - u_z v_y \\ u_z v_x - u_x v_z \\ u_x v_y - u_y v_x \end{bmatrix} = \begin{bmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}.$$

Przyjmujemy oznaczenie $S_u = \begin{bmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{bmatrix}$. Wówczas $u \times v = S_u v^T$ oraz $u \times (u \times v) = S_u(S_u v^T) = S_u^2 v^T$.

Ponadto zauważmy, że $u(u \circ v) = (v \circ u)u = (vu^T)u = v(u^T u)$ oraz $u(u \circ v) = v - v_\perp = v + u \times (u \times v)$.

Zatem $\boxed{R_u(\theta) = I\cos\theta + S_u \sin\theta + u^T u(1 - \cos\theta)}$ lub $\boxed{R_u(\theta) = I + S_u \sin\theta + S_u^2(1 - \cos\theta).}$

Przykład 9.5.39. Niech $u = \frac{1}{\sqrt{3}}(1, 1, 1)$ oraz $\theta = \frac{2}{3}\pi$. Znajdziemy macierz $R_u(\theta)$.

Obliczamy $\cos\theta = -\frac{1}{2}, \sin\theta = \frac{\sqrt{3}}{2}$, $uu^T = \frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ oraz

$S_u = \frac{1}{\sqrt{3}}\begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}$. Stąd

$$R_u(\theta) = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix} + \begin{bmatrix} 0 & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} + \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Nasz rachunek możemy sprawdzić tutaj.

Przykład 9.5.40. Macierz

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}$$

reprezentuje złożenie rotacji o kąt α względem osi Ox (w płaszczyźnie Oyz) z odbiciem względem płaszczyzny Oyz.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Metody algebry liniowej w fizyce
Wykład nr 10
Data: ....

TEMAT: *Przestrzenie unitarne*

10.1 Definicja przestrzeni unitarnej i podstawowe własności

Niech $V = (V, +, \mathbb{C}, \cdot)$ będzie przestrzenią liniową. Dla dowolnego $\alpha \in \mathbb{C}$, jeśi $\alpha = x + iy$, $x, y, \in \mathbb{R}$, to wówczas $\overline{\alpha} = x - iy$.

Definicja 10.1.1. Funkcję $s : V \times V \to \mathbb{C}$ nazywamy *iloczynem skalarnym (iloczynem hermitowskim)*, jeżeli spełnia ona następujące warunki:

i) $\forall u, v, w \in V \quad \forall \alpha, \beta \in \mathbb{R} \quad s(\alpha u + \beta v, w) = \alpha s(u, w) + \beta s(v, w)$,

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. i) $\forall v \in V \ s(v, v) \in \mathbb{R}$, zatem warunek $s(v, v) \geq 0$ ma sens.

ii) 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 $\forall u = (x_1, \ldots, x_n), v = (y_1, \ldots, y_n) \in \mathbb{C}^n \quad s(u, v) = \sum_{i=1}^{n} x_i \overline{y_i} = x_1 \overline{y_1} + \ldots + x_n \overline{y_n}$.

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

$$\langle (z_i)_{i\in\mathbb{N}}, (w_i)_{i\in\mathbb{N}} \rangle = \sum_{n=1}^{\infty} z_i \overline{w_i}.$$

3) Niech $\mathcal{C}([a,b],\mathbb{C}) = \{f : [a,b] \to \mathbb{C} \,|\, f-\text{ciągła}\}$. 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

$$\langle f, g \rangle = \int_a^b f(t)\overline{g(t)}dt.$$

4) Zdefiniujemy iloczyn skalarny w przestrzeni macierzy $M_n(\mathbb{C})$. Niech $A, B \in M_n(\mathbb{C})$, $A = [a_{ij}], B = [b_{ij}]$.

$$\langle A, B \rangle = \sum_{i,j=1}^{n} a_{ij}\overline{b_{ij}} = \mathrm{tr}(A^T\,\overline{B})$$

Definicja 10.1.5. Macierz sprzężona 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.

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 $A \in M_n(\mathbb{C})$

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 10.1.8. 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 $\{v_1, \dots v_n\} \subset V$ nazywamy *układem ortogonalnym*, jeżeli $\forall i, j \in \{1, 2, \dots, n\} \quad i \neq j \;\Rightarrow \langle v_i, v_j\rangle = 0$. Układ wektorów nazywamy *układem ortonormalnym*, jeśli jest układem ortogonalnym i każdy wektor jest unormowany, czyli

$$\forall i, j \in \{1, 2, \dots, n\} \quad \langle v_i, v_j\rangle = \begin{cases} 0 & ; \quad i \neq j \\ 1 & ; \quad i = j \end{cases}.$$

Każdy ortonormalny układ wektorów jest liniowo niezależny. Jeśli przestrzeń wektorowa jest skończenie wymiarowa, to każdy ortonormalny układ wektorów można uzupełnić do bazy ortonormalnej.

Jeśli $\mathcal{B} = \big(b_1, b_2, \dots, b_n\big)$ 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 $\mathcal{B} = \big(b_1, b_2, \dots, b_n\big)$ jest bazą przestrzeni unitarnej $(V, \langle ., .\rangle)$, zaś $v, w \in V$, $v = [\alpha_1, \alpha_2, \dots, \alpha_n]_{\mathcal{B}}$, $w = [\beta_1, \beta_2, \dots, \beta_n]_{\mathcal{B}}$, to wówczas baza $\mathcal{B}$ jest ortonormalna wtedy i tylko wtedy gdy $\langle u, w\rangle = \alpha_1\overline{\beta_1} + \alpha_2\overline{\beta_2} \dots + \alpha_n\overline{\beta_n}$.

10.2 Macierze hermitowskie i unitarne

Definicja 10.2.1. *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 10.2.2. 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 10.2.3. 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 $\mathrm{tr} A^* = \overline{\mathrm{tr} A}$.

vi) Jeśli $n = m$ oraz $\alpha \in Spec(A)$, to $\overline{\alpha} \in Spec(A^*)$.

Definicja 10.2.4. Macierz kwadratową A nazywamy macierzą

i) *hermitowską*, jeśli $A = A^*$,

ii) *antyhermitowską*, jeśli $A = -A^*$,

iii) *normalną*, jeśli $AA^* = A^*A$,

iv) *unitarną*, jeśli $AA^* = A^*A = I$ lub równoważnie $A^{-1} = A^*$.

Uwaga 10.2.5. 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 10.2.6.
$A = \begin{bmatrix} 1 & 1-i \\ 1+i & 2 \end{bmatrix}$, $A^* = A$, $AA^* = A^2 = \begin{bmatrix} 3 & 3-3i \\ 3+3i & 6 \end{bmatrix} = 3A$
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 10.2.7. 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 10.2.8. 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 unitarnych* 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 Unitarna diagonalizacja macierzy hermitowskich

Twierdzenie 10.3.1. i) Wartości własne macierzy hermitowskiej (lub rzeczywistej symetrycznej) są liczbami rzeczywistymi.

ii) Wektory własne odpowiadające różnym wartościom własnym macierzy hermitowskiej (lub rzeczywistej symetrycznej) są wzajemnie ortogonalne.

ii) Każdej wartości własnej o krotności algebraicznej k macierzy hermitowskiej (lub rzeczywistej symetrycznej) odpowiada k liniowo niezależnych wektorów własnych.

Twierdzenie 10.3.2 (Twierdzenie spektralne dla macierzy hermitowskich)**.** Dla każdej macierzy hermitowskiej $A \in M_n(\mathbb{C})$ istnieje macierz diagonalna $D \in M_n(\mathbb{R})$ oraz macierz unitarna $P \in U(n)$ takie, że

$$P^{-1}AP = P^*AP = D.$$

Oznacza to, że dla danego operatora liniowego $\varphi \in End(V)$ (reprezentowanego przez macierz A) istnieje baza ortonormalna przestrzeni $(V, \langle ., . \rangle)$, w której macierz operatora φ jest diagonalna (i elementy na diagonali są liczbami rzeczywistymi). Mówimy wówczas, że macierz A jest *unitarnie diagonalizowalna*.

Wniosek 10.3.3 (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*.

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 10.3.4. 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 $\hat{c_1} = \frac{c_1}{|c_1|} = (\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0)$, $\hat{c_2} = \frac{c_2}{|c_2|} = \sqrt{\frac{5}{6}}(-\frac{1}{5}, \frac{2}{5}, 1)$.

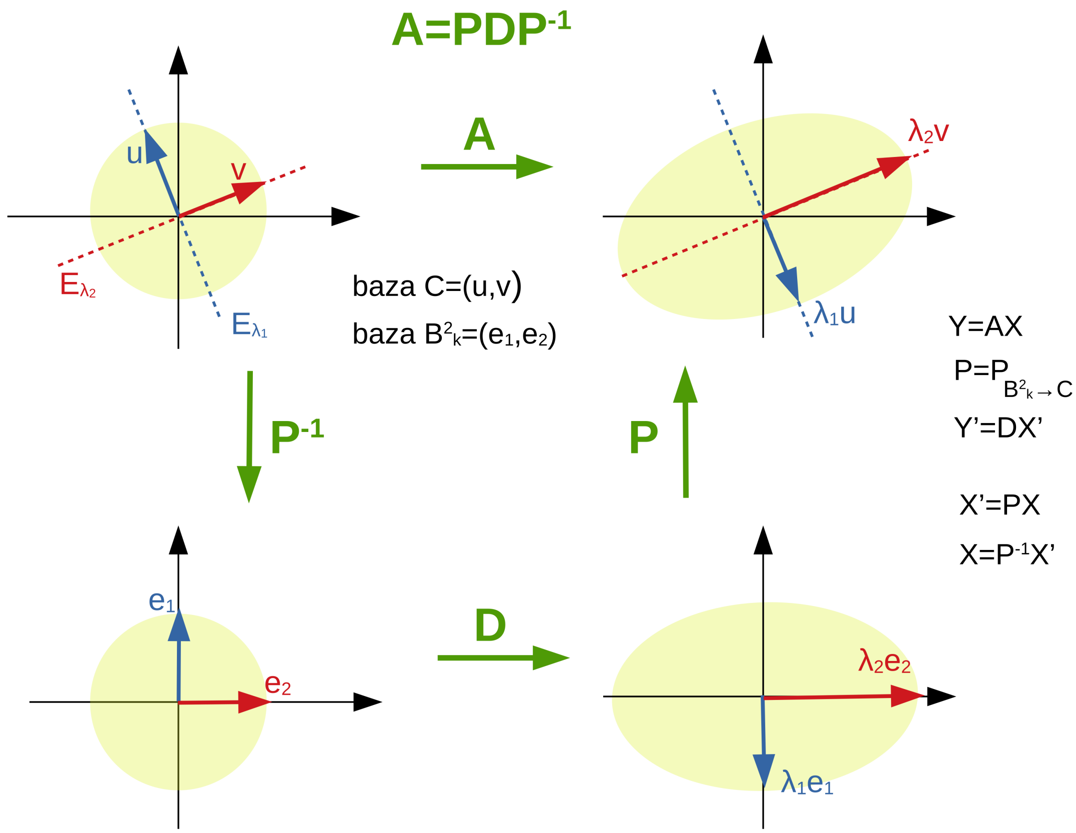
Rozważamy bazę $(\hat{u}, \hat{c_1}, \hat{c_2})$. Macierz diagonalizująca ma postać $P = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{5}} & -\frac{1}{5\sqrt{30}} \\ -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{\sqrt{5}}{\sqrt{6}} \end{bmatrix}$.

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 10.3.5. i) Dla dowolnej macierzy $A \in M_n(\mathbb{R})$ i dla dowolnych wektorów $u, v \in V$ prawdziwa jest równość $\langle Au^T, v^T\rangle = \langle u^T, A^Tv^T\rangle$, gdzie u^T, v^T oznaczają odpowiednie wektory kolumnowe. W przypadku macierzy symetrycznej otrzymujemy równość postaci $\langle Au^T, v^T\rangle = \langle u^T, Av^T\rangle$

ii) Ortogonalna diagonalizacja macierzy symetrycznej może być widziana jako *rotacja osi układu współrzędnych* w taki sposób, *aby były one równoległe do wektorów własnych.* Zobacz tutaj.

iii)



Uwaga 10.3.6. 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 10.3.7. 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.

$\det(A - tI) = \begin{vmatrix} \cos\alpha - t & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha - t & 0 \\ 0 & 0 & 1-t \end{vmatrix} = (1-t)\big((\cos\alpha - t)^2 + \sin\alpha^2\big) =$

$(1-t)(1 - 2t\cos\alpha + t^2) = \mathbf{0} \quad \Rightarrow \quad t = 0 \quad \vee \quad 1 - 2t\cos\alpha + t^2 = 0$

$\Delta = 4\cos^2\alpha - 4 = 4i^2\sin^2\alpha, \quad t = \cos\alpha \pm i\sin\alpha$

$Spec(A) = \{\lambda_1 = 1, \lambda_2 = \cos\alpha + i\sin\alpha, \lambda_3 = \cos\alpha - i\sin\alpha\}$

$$(A-\lambda_1 I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos\alpha - 1 & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha - 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} x(\cos\alpha - 1) + y\sin\alpha = 0 \\ -x\sin\alpha + (\cos\alpha - 1)y = 0 \end{cases}$$

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 $\quad x\sin\alpha + \frac{(\cos\alpha-1)^2}{\sin\alpha}x = \frac{\sin^2\alpha + (\cos\alpha-1)^2}{\sin\alpha}x = \frac{2\cos\alpha - 2}{\sin\alpha}x = 0 \Rightarrow x = 0$.
W konsekwencji $y = 0$, zaś $z \in \mathbb{C}$ dowolne. Wybieramy wektor własny $u = (0, 0, 1)$.

$$(A-\lambda_2 I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -i\sin\alpha & \sin\alpha & 0 \\ -\sin\alpha & -i\sin\alpha & 0 \\ 0 & 0 & 1 - \cos\alpha - i\sin\alpha \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

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 $v = (1, i, 0)$.

$$(A-\lambda_3 I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} i\sin\alpha & \sin\alpha & 0 \\ -\sin\alpha & i\sin\alpha & 0 \\ 0 & 0 & 1 - \cos\alpha + i\sin\alpha \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

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

$$D = P^{-1}AP = \frac{1}{2}\begin{bmatrix} 0 & 0 & \sqrt{2} \\ 1 & -i & 0 \\ -1 & -i & 0 \end{bmatrix}\begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & -1 \\ 0 & i & i \\ \sqrt{2} & 0 & 0 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha + i\sin\alpha & 0 \\ 0 & 0 & \cos\alpha - i\sin\alpha \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{i\alpha} & 0 \\ 0 & 0 & e^{-i\alpha} \end{bmatrix}.$$

Twierdzenie 10.3.8. Macierz jest unitarnie diagonalizowalna wtedy i tylko wtedy, gdy jest macierzą normalną.

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 $UU^* = U^*U = I$
$\det U = \pm 1$	$\|\det U\| = 1$
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.

10.4 Równoczesna diagonalizacja pary macierzy

Jeśli macierze $A, B \in M_n(\mathbb{C})$ są *równocześnie diagonalizowalne*, tzn. istnieje $P \in Gl_n(\mathbb{C})$ taka, że macierze $P^{-1}AP$ oraz $P^{-1}BP$ są diagonalne, 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 10.4.1. Jeżeli macierze hermitowskie (lub unitarne) $A, B \in M_n(\mathbb{C})$ komutują, to istnieje macierz unitarna $P \in U(n)$ oraz macierze diagonalne $D_A, D_B \in M_n(\mathbb{R})$ takie, że $P^{-1}AP = D_A$ oraz $P^{-1}BP = D_B$.

Przykład 10.4.2. Niech $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$.
Są to rzeczywiste macierze symetryczne, a zatem każda z nich jest ortogonalnie diagonalizowalna.
$AB = \begin{bmatrix} 8 & 7 \\ 7 & 8 \end{bmatrix} = BA$, więc istnieje $P \in O(3)$ taka, że macierze $P^{-1}AP$, $P^{-1}BP$ są diagonalne.
Znajdujemy wartości własne i wektory własne macierzy A.
$\chi_A(t) = \det(A - tI) = \begin{vmatrix} 2-t & 1 \\ 1 & 2-t \end{vmatrix} = (2-t)^2 - 1 = t^2 - 4t + 3 = (t-1)(t-3)$
$Spec(A) = \{\lambda_1 = 1, \lambda_2 = 3\}$ widmo proste

$(A - \lambda_1 I)\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x + y = 0$
Wybieramy wektor własny $u = (1, -1)$ i normujemy go $\hat{u} = \frac{1}{\sqrt{2}}(1, -1)$.

$(A - \lambda_2 I)\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x - y = 0$
Wybieramy wektor własny $v = (1, 1)$ i normujemy go $\hat{v} = \frac{1}{\sqrt{2}}(1, 1)$.

Stąd $P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. Ponadto $P^{-1} = P^T = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ oraz $P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$.
Obliczamy $P^{-1}BP = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}$.
Widzimy, że $Spec(B) = \{\lambda_1 = 1, \lambda_2 = 5\}$.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Metody algebry liniowej w fizyce
Wykład nr 11
Data: ....

TEMAT: *Formy kwadratowe*

11.1 Definicja formy kwadratowej

Przypomnienie: Każdą macierz kwadratową D można przedstawić w postaci sumy macierzy symetrycznej i antysymetrycznej $D = B + C$, gdzie $B = \frac{1}{2}(D + D^T)$, $C = \frac{1}{2}(D - D^T)$, $B = B^T$, $C = -C^T$.

Obserwacja: $X^T DX = X^T BX$
Istotnie $X^T DX = X^T(B + C)X = X^T BX + X^T CX$ oraz
$X^T CX = X^T \frac{D-D^T}{2} X = \frac{1}{2}\left(X^T DX - X^T D^T X\right) = \frac{1}{2}\left(X^T DX - (X^T D^T X)^T\right) = \frac{1}{2}\left(X^T DX - X^T DX\right) = O$

Niech $A \in M_n(\mathbb{R})$ będzie macierzą symetryczną, tzn. $A = A^T$.
Na mocy powyższej obserwacji możemy przyjąć to założenie bez straty dla ogólności.

Definicja 11.1.1. Funkcję $\gamma : \mathbb{R}^n \to \mathbb{R}$ daną wzorem $\gamma(x) = X^T AX$, gdzie $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ dla dowolnego $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, nazywamy *formą kwadratową*. Macierz symetryczną A nazywamy *macierzą formy kwadratowej* γ.

Uwaga 11.1.2.

i) Jeśli $A = [a_{ij}]$, to wówczas $\gamma(x) = \sum_{i,j=1}^n a_{ij}x_ix_j$.

ii) γ jest wielomianem n zmiennych jednorodnym stopnia 2.

iii) $\forall x \in \mathbb{R}^n \ \forall \lambda \in \mathbb{R} \quad \gamma(\lambda x) = \lambda^2\gamma(x)$

Dowód. i) $[x_1 \ldots x_n] \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = [x_1 \ldots x_n] \begin{bmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \ldots \\ a_{n1}x_1 + \ldots + a_{nn}x_n \end{bmatrix} =$
$a_{11}x_1^2 + a_{12}x_1x_2 + \ldots + a_{nn}x_n^2$
ii) Na mocy *i*) każdy jednomian γ jest stopnia 2.
iii) Niech $x = (x_1, \ldots, x_n)$, $\lambda \in \mathbb{R}$. Wówczas na mocy i) mamy $\gamma(\lambda x) = \gamma(\lambda x_1, \ldots, \lambda x_n) = \sum_{i,j=1}^n a_{ij}(\lambda x_i)(\lambda x_j) = \lambda^2 \sum_{i,j=1}^n a_{ij}x_ix_j$. $\square$

Przykład 11.1.3. Poniższe odwzorowania to formy kwadratowe.

i) $\gamma : \mathbb{R}^3 \to \mathbb{R}$, $\gamma(x_1, x_2, x_3) = 2x_1^2 + 3x_1x_2 + 4x_1x_3 + x_2^2 + x_3^3$

$$\gamma(x_1, x_2, x_3) = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 2 & \frac{3}{2} & 2 \\ \frac{3}{2} & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

ii) Niech $\Omega \subset \mathbb{R}^n$ będzie obszarem, zaś $f : \Omega \to \mathbb{R}$ dowolną funkcją taką, że $f \in \mathcal{C}^2(\Omega)$. Niech $P_0 \in \Omega$. Różniczka rzędu drugiego funkcji f w punkcie P_0

$$d^2_{P_0} f : \mathbb{R}^n \to \mathbb{R}, \quad d^2_{P_0} f(h_1, \dots, h_n) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(P_0) h_i h_j$$

jest formą kwadratową. Jej macierz ma postać

$$H_f(P_0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(P_0) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(P_0) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(P_0) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(P_0) & \frac{\partial^2 f}{\partial x_2^2}(P_0) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(P_0) \\ \dots & & \dots & \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(P_0) & \frac{\partial^2 f}{\partial x_n \partial x_2}(P_0) & \dots & \frac{\partial^2 f}{\partial x_n^2}(P_0) \end{bmatrix}.$$

Jest to tak zwana *macierz Hessego*. Jej wyznacznik nazywamy *hesjanem*.

11.2 Określoność formy kwadratowej

Dla dowolnej formy kwadratowej γ mamy $\gamma(\mathbf{0}) = 0$.

Definicja 11.2.1. Formę kwadratową $\gamma : \mathbb{R}^n \to \mathbb{R}$, $\gamma(x) = X^T A X$ nazywamy

i) *dodatnio określoną*, gdy $\forall x \in \mathbb{R}^n \setminus \{\mathbf{0}\} \quad \gamma(x) > 0$,

ii) *ujemnie określoną*, gdy $\forall x \in \mathbb{R}^n \setminus \{\mathbf{0}\} \quad \gamma(x) < 0$,

iii) *dodatnio półokreśloną*, gdy $\forall x \in \mathbb{R}^n \quad \gamma(x) \geq 0$,

iv) *ujemnie półokreśloną*, gdy $\forall x \in \mathbb{R}^n \quad \gamma(x) \leq 0$,

v) *nieokreśloną*, jeśli przyjmuje wartości zarówno dodatnie jak i ujemne.

Terminologia ta przenosi się na macierze. Macierz symetryczna A jest dodatnio określona, gdy forma kwadratowa $\gamma(x) = X^T A X$ jest dodatnio określona itd.

Przykład 11.2.2. i) $\gamma : \mathbb{R}^3 \to \mathbb{R}$, $\gamma(x_1, x_2, x_3) = 2x_1^2 + x_2^2 + x_3^2$ jest dodatnio określona

ii) $\gamma : \mathbb{R}^3 \to \mathbb{R}$, $\gamma(x_1, x_2, x_3) = -x_1^2 - 4x_2^2 - x_3^2$ jest ujemnie określona

iii) $\gamma : \mathbb{R}^3 \to \mathbb{R}$, $\gamma(x_1, x_2, x_3) = x_1^2 - x_2^2 + x_3^2$ jest nieokreślona, bowiem $\gamma(1, 0, 1) = 2 > 0$, zaś $\gamma(0, 1, 0) = -1 < 0$.

iv) $\gamma : \mathbb{R}^3 \to \mathbb{R}$, $\gamma(x_1, x_2, x_3) = 2x_1^2 + x_3^2$ jest dodatnio półokreślona. $\forall x_2 \in \mathbb{R} \;\; \gamma(0, x_2, 0) = 0$

Badanie określoności formy kwadratowej

Definicja 11.2.3. Niech $A = [a_{ij}] \in M_n(\mathbb{R})$.

i) *Minorem głównym* stopnia k macierzy A nazywamy wyznacznik dowolnej macierzy powstałej przez skreślenie $n - k$ wierszy i kolumn o tych samych indeksach.

ii) *Minorem wiodącym głównym* stopnia k macierzy A nazywamy wyznacznik macierzy powstłej przez skreślenie $n-k$ ostatnich wierszy i kolumn. Oznaczamy go symbolem Δ_k.

Symbolem $D_{i_1 \ldots i_k}$ oznaczamy minor główny stopnia k, powstały przez skreślenie wierszy i kolumn poza tymi o indeksach $i_1 < \ldots < i_k$.

Przykład 11.2.4. $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ -1 & -2 & -3 & -4 \\ -5 & -6 & -7 & -8 \end{bmatrix}$

minory wiodące główne: $\Delta_1 = D_{11} = 1, \;\; \Delta_2 = D_{12} = \begin{vmatrix} 1 & 2 \\ 5 & 6 \end{vmatrix} = -4,$

$\Delta_3 = D_{123} = \begin{vmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ -1 & -2 & -3 \end{vmatrix} = 0, \;\; \Delta_4 = D_{1234} = \det A = 0$

minory główne: $D_{14} = \begin{vmatrix} 1 & 4 \\ -5 & -8 \end{vmatrix} = 12, \;\; D_{24} = \begin{vmatrix} 6 & 8 \\ -6 & -8 \end{vmatrix} = 0$

Twierdzenie 11.2.5 (Kryterium Sylvestera)**.** Niech $A = [a_{ij}] \in M_n(\mathbb{R})$ będzie macierzą symetryczną.

i) A jest dodatnio określona wtedy i tylko wtedy, gdy jej minory wiodące główne są dodatnie, tzn. $\forall k \in \{1, 2, \ldots, n\} \; \Delta_k > 0$.

ii) A jest ujemnie określona wtedy i tylko wtedy, gdy jej minory wiodące główne parzystego stopnia są dodatnie, a nieparzystego ujemne,
tzn. $\forall k \in \{1, 2, \ldots, n\} \; (-1)^k \Delta_k > 0$.

Bez dowodu.

Przykład 11.2.6. i) $A = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -3 \end{bmatrix}$ jest ujemnie określona.

$$\Delta_1 = -2 < 0, \quad \Delta_2 = \begin{vmatrix} -2 & 1 \\ 1 & -1 \end{vmatrix} = 1 > 0, \quad \Delta_3 = \det A = -3 < 0.$$

ii) $B = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$ jest półokreślona ujemnie, gdyż $\gamma(x_1, x_2) = -x_2^2 \leq 0$.

Uwaga 11.2.7. Z warunku $\forall k \in \{1, 2, \ldots, n\}$ $\Delta_k \geq 0$ nie wynika dodatnia półokreśloność A.

Twierdzenie 11.2.8. Niech $\gamma : \mathbb{R}^n \to \mathbb{R}$, $\gamma(x) = X^T A X$, gdzie $A = A^T$.

i) Forma kwadratowa γ jest dodatnio półokreślona wtedy i tylko wtedy, gdy wszystkie minory główne macierzy A są nieujemne,
tzn. $\forall\, 1 \leq k \leq n \quad \forall\, 1 \leq i_1 < \ldots < i_k \leq n \quad D_{i_1 \ldots i_k} \geq 0$.

ii) Forma kwadratowa γ jest ujemnie półokreślona wtedy i tylko wtedy, gdy minory główne macierzy A przyjmują następujące znaki
$\forall\, 1 \leq k \leq n \quad \forall\, 1 \leq i_1 < \ldots < i_k \leq n \quad (-1)^k D_{i_1 \ldots i_k} \geq 0$.

Bez dowodu.

Przykład 11.2.9. Forma kwadratowa $\gamma : \mathbb{R}^3 \to \mathbb{R}$,
$\gamma(x_1, x_2, x_3) = x_1^2 + x_3^2 + 2x_1x_3$ jest dodatnio półokreślona.

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \ \Delta_1 = 1, \ \Delta_2 = \Delta_3 = 0,$$

$D_2 = 0$, $D_3 = 1$, $D_{13} = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0$, $D_{23} = \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix} = 0$
Oczywiste bowiem $\gamma(x_1, x_2, x_3) = (x_1 + x_3)^2 \geq 0$.

Uwaga 11.2.10. i) Jeśli $\Delta_2 = D_{12} < 0$, to macierz jest nieokreślona.

ii) Rzeczywista macierz symetryczna $A = [a_{ij}] \in M_n(\mathbb{R})$ ma n rzeczywistych wartości własnych (licząc z krotnościami).

Dowód. i) $\Delta_2 = D_{12}$ to minor stopnia 2. Jeśli $D_{12} = (-1)^2 D_{12} < 0$, to na mocy kryterium Sylvestera, macierz nie jest dodatnio określona ani ujemnie określona. Na mocy twierdzenia 11.2.8 nie jest półokreślona dodatnio ani półokreślona ujemnie.
ii) Dowód można znaleźć w [4]. □

Twierdzenie 11.2.11 (Kryterium wartości własnych)**.** Niech $\gamma : \mathbb{R}^n \to \mathbb{R}$, $\gamma(x) = X^TAX$, gdzie $A = A^T$ oraz niech $Spec(A) = \{\lambda_1, \ldots, \lambda_n\}$. Wówczas forma kwadratowa γ jest

i) dodatnio określona wtedy i tylko wtedy, gdy $\forall\, i \in \{1, 2, \ldots n\} \quad \lambda_i > 0$,

ii) ujemnie określona wtedy i tylko wtedy, gdy $\forall\, i \in \{1, 2, \ldots n\} \quad \lambda_i < 0$,

iii) dodatnio półokreślona wtedy i tylko wtedy, gdy $\forall\, i \in \{1, 2, \ldots n\} \quad \lambda_i \geq 0$,

iv) ujemnie półokreślona wtedy i tylko wtedy, gdy $\forall\, i \in \{1, 2, \ldots n\} \quad \lambda_i \leq 0$,

iv) nieokreślona wtedy i tylko wtedy, gdy $\exists\, i, j \in \{1, 2, \ldots n\} \quad \lambda_i > 0 \wedge \lambda_j < 0$.

Dowód. Niech $\lambda \in Spec(A)$ oraz niech v będzie wektorem własnym odpowiadającym λ. Oczywiście $v \neq \mathbf{0}_V$, więc $|v| \neq 0$. Obliczamy $v^TAv = \gamma(v) = v^T\lambda v = \lambda v^Tv = \lambda|v|^2$. Zatem $\lambda(v) > 0 \Leftrightarrow \lambda > 0$, co dowodzi podpunktu i). Analogicznie w pozostałych przypadkach. □

Przykład 11.2.12. $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ jest dodatnio półokreślona.

$$\chi_A(t) = \begin{vmatrix} 1-t & 0 & 1 \\ 0 & -t & 0 \\ 1 & 0 & 1-t \end{vmatrix} = -t(1-t)^2 + t = t\big(1 - (1-t)^2\big) = t(2t - t^2) = t^2(2-t)$$

$Spec(A) = \{\lambda_1 = \lambda_2 = 0, \lambda_3 = 2\}$

Contents