Fundusze Europejskie
Wiedza Edukacja Rozwój

Rzeczpospolita Polska

Unia Europejska
Europejski Fundusz Społeczny

DU DOSKONAŁY UNIWERSYTET

# Algebra liniowa z geometrią

Klaudiusz Wójcik

Kraków, 2023

# Spis treści

# Kilka uwag i wskazówek dla studentów

Wykłady na UJ nie są obowiązkowe i otrzymacie ode mnie notatki do wykładu. Mimo to, gorąco zachęcam Was do **chodzenia na wykłady** – nie tylko moje :). Notatki nigdy nie zawierają wszystkiego. Wykład to kontakt z żywym człowiekiem, pozwalający przekazać mniej formalne informacje i intuicje. Na wykładzie dowiecie się co jest naprawdę istotne, a co jest tylko detalem technicznym. Nie wyczytacie tego z notatek.

Kolekcjonujcie i analizujcie możliwie dużo **konkretnych przykładów**. Są one równie ważne jak definicje i twierdzenia. Starajcie się każdą definicje i każde twierdzenie widzieć w kontekście przykładów. Nie bójcie się rachunków na prostych przykładach. Wyrabiają one intuicje i często zawierają w sobie zalążki formalnych dowodów.

O ile to możliwe starajcie się myśleć geometrycznie i **robić rysunki**. Wiele pojęć matematycznych ma znaczenie geometryczne pozwalające na lepsze ich zrozumienie.

W trakcie wykładu zaproponuję Wam serię zadań. **Spróbujcie je rozwiązać samodzielnie**. Nie bójcie się dyskusji o nich z koleżankami i kolegami - to zawsze może pomóc. Jeśli się to nie uda, poproście o pomoc prowadzących ćwiczenia. **Ostrzeżenie:** prowadzący ćwiczenia łatwo rozpoznają czy próbowaliście rozwiązać problemy czy poszliście na łatwiznę.

Zaproponuję Wam również **serię testów typu prawda czy fałsz**. Powinniście je rozwiązać **samodzielnie**. Liczba sukcesów w odpowiedziach da Wam wyobrażenie o stopniu zrozumienia prezentowanego na wykładach materiału.

**Wszystko w Waszych rękach**. My jesteśmy tu po to, aby Wam pomóc. Reszta należy do Was.

U To jest przykład **Uwagi** - będą one przewijały się w tekście i powinniście je wnikliwie przemyśleć.
W tekście pojawiają się Definicje, Twierdzenia, Lematy, Wnioski i Przykłady. Definicje nadają nazwy obiektom ważnym dla rozważanej teorii. Twierdzenia opisują szczególnie ważne własności wcześniej zdefiniowanych pojęć. Z logicznego punktu widzenia, Lematy są tym samym co Twierdzenia, ale ich rola jest pomocnicza - często dowody Twierdzeń są poprzedzone serią Lematów. Wnioski są zazwyczaj szczególnymi przypadkami lub prostymi konsekwencjami Twierdzeń i Definicji podkreślającymi ich użyteczność. Ważna wskazówka: starajcie się analizować założenia Twierdzeń. Zadawajcie sobie pytania typu: a może jakieś założenie nie jest potrzebne, aby zaszła teza? Analizujcie proste konkretne przykłady. Często dają one odpowiedź dlaczego założenia twierdzeń są konieczne i nie można ich pominąć.

**Przykład 0.1.** Zobrazuję powyższą uwagę konkretnym przykładem. Rozważmy Twierdzenie:

Suma liczb całkowitych jest liczbą całkowitą.

Chyba się zgodzimy, że jest to zdanie prawdziwe, czyli Twierdzenie. Powinniście sobie zadać pytanie:

Co się stanie jak liczby całkowite zastąpimy innym zbiorem liczbowym?

Czy prawdziwe są stwierdzenia

Suma liczb wymiernych jest liczbą wymierną.

albo

Suma liczb niewymiernych jest liczbą niewymierną.

Przemyślcie Swoje odpowiedzi!

A teraz przykład **Ostrzeżenia** – będą to szczególnie ważne Uwagi.

Jeśli Wniosek z Twierdzenia nie jest całkiem oczywisty w mojej opinii, to będę podawał jego uzasadnienie (dowód). Jeżeli Wniosek jest podany bez dowodu, to **powinniście dobrze przemyśleć** dlaczego uznałem, że nie ma potrzeby podawać dodatkowych argumentów jego prawdziwości, gdyż jest on bezpośrednią konsekwencją Definicji albo bardziej ogólnego Twierdzenia.

## Część I

# Algebra liniowa z geometrią 1

# Rozdział 1

# Pojęcia wstępne

SŁOWA KLUCZOWE ROZDZIAŁU
ciało liczb rzeczywistych $\Diamond$ ciało liczb zespolonych $\Diamond$ wzór de Moivre'a $\Diamond$ zasadnicze twierdzenie algebry $\Diamond$ ciało $\mathbb{Z}_p$ $\Diamond$ grupa $\Diamond$ grupa $\mathbb{S}^1$ $\Diamond$ grupa pierwiastków zespolonych z jedynki

## 1.1 Zbiory i liczby rzeczywiste

Pojęcia zbiór nie definiujemy. *Zbiór* jest kolekcją pewnych obiektów. Dla przykładu, zbiór

$$A = \{1, 2, 3\}$$

składa się z elementów $1, 2, 3$ będących liczbami naturalnymi. Napis $1 \in A$ oznacza, że liczba naturalna 1 jest *elementem* zbioru $A$. Natomiast napis $4 \notin A$ stwierdza, że 4 nie jest elementem zbioru $A$. Zbiory $A$ i $B$ są *równe*, gdy składają się z tych samych elementów. Zbiory

$$\{1, 2, 3\} = \{3, 2, 1\} = \{1, 2, 3, 1\}$$

są równe. Zapis $A \subset B$ oznacza, że zbiór $A$ jest *podzbiorem* zbioru $A$, czyli każdy element zbioru $A$ jest elementem zbioru $B$. Dla przykładu,

$$\{1, 2\} \subset \{-1, 0, 1, 2\}, \quad \{1, 8\} \not\subset \{-1, 0, 1, 2\}.$$

**Przykład 1.1.** Ważne zbiory liczbowe:

- zbiór liczb naturalnych $\mathbb{N} = \{1, 2, 3, 4, 5, \dots\}$;
- zbiór liczb całkowitych $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$;
- zbiór liczb wymiernych $\mathbb{Q} = \left\{\frac{m}{n} : m \in \mathbb{Z},\ n \in \mathbb{Z} \setminus \{0\}\right\}$;
- zbiór liczb rzeczywistych $\mathbb{R}$.

Zachodzą zawierania

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R},$$

ale żadne z nich nie jest równością. Zbiór liczb rzeczywistych $\mathbb{R}$ składa się z liczb wymiernych $\mathbb{Q}$ i liczb niewymiernych $\mathbb{R} \setminus \mathbb{Q}$. Przykładowo, liczby $\sqrt{2}$, $\pi$, $3 - \sqrt{3}$ są niewymierne.

*Zbiór pusty* $\emptyset$, to zbiór który nie zawiera żadnych elementów. Przykładowo,

$$\begin{aligned}
\emptyset &= \{n \in \mathbb{N} : n^2 = -1\} \\
&= \{x \in \mathbb{R} : x \cdot 0 \neq 0\} \\
&= \{m \in \mathbb{Z} : m^3 = 2\} \\
&= \{m \in \mathbb{N} : 3.01 \leqslant m < 3.99\}.
\end{aligned}$$

TEST $\longrightarrow$ Oceń czy poniższe zdania są prawdziwe:

- Zbiory $A = \{x \in \mathbb{N} : x^2 = 16\}$ i $B = \{x \in \mathbb{Z} : x^2 = 16\}$ są równe;
- Suma liczby wymiernej $x$ i niewymiernej $y$ jest liczbą niewymierną;
- Suma $x + y$ i iloczyn $x \cdot y$ liczb wymiernych jest liczbą wymierną;
- Suma $x + y$ liczb niewymiernych jest liczbą niewymierną;
- Iloczyn $x \cdot y$ liczb niewymiernych jest liczbą niewymierną;
- Iloczyn $x \cdot y$ liczb: wymiernej $x$ i niewymiernej $y$ jest liczbą niewymierną;
- Suma $x + y$ liczb: wymiernej $x$ i niewymiernej $y$ jest liczbą niewymierną.

$\longleftarrow$

Liczby rzeczywiste możemy dodawać i mnożyć. Wynikiem tych operacji jest ponownie liczba rzeczywista. Bardziej formalnie oznacza to, że określone są dwie funkcje (operacje): dodawanie liczb rzeczywistych

$$+ : \mathbb{R} \times \mathbb{R} \ni (x, y) \longmapsto x + y \in \mathbb{R}$$

oraz ich mnożenie

$$\cdot : \mathbb{R} \times \mathbb{R} \ni (x, y) \longmapsto x \cdot y \in \mathbb{R}.$$

Jeśli poprawnie rozwiązaliście Test to zauważycie, że dodawanie i mnożenie jest również *dobrze określonym* działaniem w zbiorze liczb wymiernych. Powyższe zdanie oznacza tyle, że *suma* $x+y$ *i iloczyn* $x \cdot y$ *liczb wymiernych jest liczbą wymierną.* Nie jest tak z liczbami niewymiernymi. Działania te wyprowadzają nas ze zbioru liczb niewymiernych. Przykładowo, liczby $\sqrt{2}$ i $-\sqrt{2}$ są niewymierne, ale ich suma $\sqrt{2}+(-\sqrt{2}) = 0$ i iloczyn $\sqrt{2}\cdot(-\sqrt{2}) = -2$ są liczbami wymiernymi (nawet całkowitymi).

Wróćmy do działań dodawania i mnożenia w zbiorze liczb rzeczywistych. *Zakładamy*, że dla dowolnych liczb rzeczywistych $a, b, c \in \mathbb{R}$ zachodzą warunki (aksjomaty):

(i) $(a + b) + c = a + (b + c)$

(ii) $a + b = b + a$

(iii) $a + 0 = a$

(iv) istnieje taka liczba rzeczywista $b$, że $a + b = 0$

(v) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$

(vi) $a \cdot b = b \cdot a$

(vii) $a \cdot 1 = a$

(viii) jeżeli $a \neq 0$, to istnieje taka liczba rzeczywista $a^{-1}$ , że $a \cdot a^{-1} = 1$

(ix) $a \cdot (b + c) = a \cdot b + a \cdot c$

Algebraicy mówią wtedy krótko, że zbiór liczb rzeczywistych $\mathbb{R}$ z działaniami $+$ oraz $\cdot$ jest *ciałem.* Oznacza to, że trójka $(\mathbb{R}, +, \cdot)$ (zbiór $\mathbb{R}$ i dwa działania w nim) spełnia warunki (i)–(ix).

Liczby rzeczywiste możemy ze sobą porównywać. Dla dowolnych $x, y \in \mathbb{R}$ zachodzi jeden z warunków

$$x = y \quad \text{lub} \quad x < y \quad \text{lub} \quad y < x.$$

Relacja $<$ jest zgodna z działaniami w tym sensie, że zachodzą warunki:

(1) jeśli $a < c$ i $b < d$, to $a + b < c + d$,

(2) jeśli $a < b$ i $c > 0$, to $a \cdot c < b \cdot c$.

Dla liczby rzeczywistej $x \in \mathbb{R}$ definiujemy jej *wartość bezwzględną* wzorem

$$|x| = \begin{cases} x, & \text{gdy } x \geqslant 0, \\ -x, & \text{gdy } x \leqslant 0. \end{cases}$$

Liczby rzeczywiste możemy interpretować jako punkty na prostej. Liczba nieujemna $|x|$ jest odległością punktu $x$ od punktu 0. Dla liczb rzeczywistych $x, y \in \mathbb{R}$ zachodzi nierówność

$$|x + y| \leqslant |x| + |y|.$$

## 1.2 Liczby zespolone

Wprowadzimy teraz ważne uogólnienie ciała liczb rzeczywistych, czyli ciało liczb zespolonych.

Zbiór *liczb zespolonych* $\mathbb{C}$ definiujemy jako zbiór uporządkowanych par liczb rzeczywistych, czyli

$$\mathbb{C} := \{(x, y) : x, y \in \mathbb{R}\}.$$

Podobnie jak liczby rzeczywiste możemy interpretować jako punkty na prostej, tak elementy zbioru $\mathbb{C}$ są punktami na płaszczyźnie.

Na razie mamy pewien zbiór tzn. zbiór punktów na płaszczyźnie. Aby mówić o ciele liczb zespolonych musimy w zbiorze $\mathbb{C}$ wprowadzić dwa działania $+$ oraz $\cdot$ spełniające warunki (i)–(ix). Zrobimy to wykorzystując działania określone w zbiorze liczb rzeczywistych. Jest to typowy dla matematyki mechanizm tworzenia nowych struktur w oparciu o struktury już istniejące.

**Definicja 1.1** (Suma i iloczyn liczb zespolonych)**.** Sumę i iloczyn liczb zespolonych $(x_1, y_1)$, $(x_2, y_2) \in \mathbb{C}$ określamy wzorami

$$(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2), \tag{1.1}$$

$$(x_1, y_1) \cdot (x_2, y_2) := (x_1 \cdot x_2 - y_1 \cdot y_2, x_1 \cdot y_2 + y_1 \cdot x_2). \tag{1.2}$$

(!) Wzory (1.1) i (1.2) wymagają pewnego komentarza. W formule (1.1) znak $+$ z lewej strony równości jest nowo definiowanym pojęciem dodawania liczb zespolonych przy pomocy, wcześniej zdefiniowanej, formuły z prawej strony: jest to więc znak dodawania punktów na płaszczyźnie. Po prawej stronie znak $+$ jest już wcześniej znaną operacją dodawania liczb rzeczywistych. Mamy tu więc do czynienia z pewną kolizją oznaczeń: znak $+$ z prawej i lewej strony oznacza inne operacje, zdefiniowane w innych zbiorach. Musimy się do tego przyzwyczaić i domyślać się z kontekstu o jakie działanie chodzi. Alternatywą byłoby oznaczanie dodawania liczb zespolonych jakimś innym symbolem niż $+$ np. $\heartsuit$. Moglibyśmy wtedy formułę (1.1) zapisać jako

$$(x_1, y_1) \heartsuit (x_2, y_2) = (x_1 + x_2, y_1 + y_2).$$

Tego typu mnożenie oznaczeń nie wydaje się jednak celowe. Zachęcam do analogicznego przeanalizowania znaczenia symboli $\cdot$, $\pm$ we wzorze (1.2).

Dla liczby zespolonej $z = (x, y) \in \mathbb{C}$ wygodnie jest wprowadzić następującą terminologię. Liczby rzeczywiste $x$ oraz $y$ nazywamy odpowiednio *częścią rzeczywistą* i *częścią urojoną* liczby zespolonej $z$. Będziemy używać oznaczeń

$$\operatorname{Re} z := x, \quad \operatorname{Im} z := y.$$

**Definicja 1.2** (Sprzężenie i moduł liczby zespolonej)**.** Dwie liczby zespolone $(x_1, y_1)$ oraz $(x_2, y_2)$ są *równe*, wtedy i tylko wtedy, gdy $x_1 = x_2$ oraz $y_1 = y_2$, tzn. zarówno ich części rzeczywiste, jak i urojone są równe. Dla liczby zespolonej $z = (x, y) \in \mathbb{C}$ definiujemy liczbę przeciwną $-z := (-x, -y)$ oraz

$$\overline{z} := (x, -y) \qquad \textit{sprzężenie liczby zespolonej}$$

$$|z| := \sqrt{x^2 + y^2} \qquad \textit{moduł liczby zespolonej.}$$

Moduł liczby zespolonej $z = (x, y)$ jest odległością punktu $(x, y)$ od początku układu współrzędnych $(0, 0)$.

Zachodzą własności

$$\overline{z_1 + z_2} = \overline{z}_1 + \overline{z}_2, \qquad \overline{z_1 \cdot z_2} = \overline{z}_1 \cdot \overline{z}_2$$
$$|z_1 + z_2| \leqslant |z_1| + |z_2|, \qquad |z_1 \cdot z_2| = |z_1||z_2|$$
$$\operatorname{Re} z = \frac{z + \overline{z}}{2}, \qquad \operatorname{Im} z = \frac{z - \overline{z}}{2i}.$$

**Przykład 1.2.** Uzasadnimy, że

$$|z + w| \leqslant |z| + |w|, \quad z, w \in \mathbb{C}.$$

Zauważmy, że

$$\begin{aligned}
|z + w|^2 &= (z + w)(\overline{z + w}) \\
&= (z + w)(\overline{z} + \overline{w}) \\
&= |z|^2 + z \cdot \overline{w} + \overline{z} \cdot w + |w|^2 \\
&= |z|^2 + |w|^2 + 2 \operatorname{Re}(z \cdot \overline{w}) \\
&\leqslant |z|^2 + |w|^2 + 2|z \cdot w| \\
&= (|z| + |w|)^2.
\end{aligned}$$

**Wniosek 1.1.** *Dla dowolnej liczby zespolonej $z = (x, y) \in \mathbb{C}$ zachodzi wzór*

$$z \cdot \overline{z} = (|z|^2, 0).$$

*Dowód.* Dowód polega na bezpośrednim rachunku. Mamy

$$\begin{aligned}
z \cdot \overline{z} &= (x, y) \cdot (x, -y) \\
&= (x \cdot x - y \cdot (-y), x \cdot (-y) + y \cdot x) \\
&= (x^2 + y^2, 0) \\
&= (|z|^2, 0).
\end{aligned}$$
□

**Twierdzenie 1.1.** *Dla dowolnych liczb zespolonych $z, w, q$ mamy*

*(i)* $(z + w) + q = z + (w + q)$

*(ii)* $z + w = w + z$

*(iii)* $z + (0, 0) = z$

*(iv)* $z + (-z) = (0, 0)$

*(v)* $(z \cdot w) \cdot q = z \cdot (w \cdot q)$

*(vi)* $z \cdot w = w \cdot z$

*(vii)* $z \cdot (1, 0) = z$

*(viii) jeżeli $z \neq 0$, to istnieje jedyna taka liczba zespolona $z^{-1}$, że $z \cdot z^{-1} = (1, 0)$*

*(ix)* $z \cdot (w + q) = z \cdot w + z \cdot q$

*W szczególności, trójka $(\mathbb{C}, +, \cdot)$ jest ciałem.*

*Dowód.* Udowodnimy dla przykładu punkty (viii) i (ix) pozostawiając dowody pozostałych punktów jako ćwiczenie. Zaczniemy od dowodu punktu (ix). Niech $z = (x_1, y_1)$, $w = (x_2, y_2)$ i $q = (x_3, y_3)$ będą dowolnymi liczbami zespolonymi. Wtedy

$$\begin{aligned}
z \cdot (w + q) &= (x_1, y_1) \cdot \big((x_2, y_2) + (x_3, y_3)\big) \\
&= (x_1, y_1) \cdot (x_2 + x_3, y_2 + y_3) \\
&= \big(x_1 \cdot (x_2 + x_3) - y_1 \cdot (y_2 + y_3), x_1 \cdot (y_2 + y_3) + y_1 \cdot (x_2 + x_3)\big) \\
&= (x_1 \cdot x_2 - y_1 \cdot y_2 + x_1 \cdot x_3 - y_1 \cdot y_3, x_1 \cdot y_2 + y_1 \cdot x_2 + x_1 \cdot y_3 + y_1 \cdot x_3) \\
&= (x_1 \cdot x_2 - y_1 \cdot y_2, x_1 \cdot y_2 + y_1 \cdot x_2) + (x_1 \cdot x_3 - y_1 \cdot y_3, x_1 \cdot y_3 + y_1 \cdot x_3) \\
&= z \cdot w + z \cdot q.
\end{aligned}$$

Zajmiemy się teraz dowodem punktu (ix). Niech $z = (x, y) \neq (0, 0)$ będzie dowolną, niezerową liczbą zespoloną. Zauważmy, że $|z|^2 = x^2 + y^2 \neq 0$. Szukamy takiej liczby zespolonej $z^{-1}$, że $z \cdot z^{-1} = (1, 0)$. Ponieważ

$$z \cdot \overline{z} = (|z|^2, 0),$$

więc wystarczy przyjąć, że

$$z^{-1} = \left( \frac{x}{|z|^2}, \frac{-y}{|z|^2} \right) = \frac{\overline{z}}{|z|^2}. \qquad \square$$

Liczby zespolone postaci $(x, 0)$ będziemy utożsamiać z liczbami rzeczywistymi, tzn. będziemy pisać $(x, 0) = x$.

**Definicja 1.3** (Jednostka urojona)**.** Liczbę zespoloną

$$i := (0, 1)$$

nazywamy *jednostką urojoną*.

Zauważmy, że

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

Stąd

$$i^2 = -1, \quad i^3 = -i, \quad i^4 = 1,$$

czyli

$$i^{4k} = 1, \quad i^{4k+1} = i, \quad i^{4k+2} = -1, \quad i^{4k+3} = -i, \quad k \in \mathbb{N}.$$

Każdą liczbę zespoloną $z = (x, y)$ możemy zapisać w *postaci algebraicznej*:

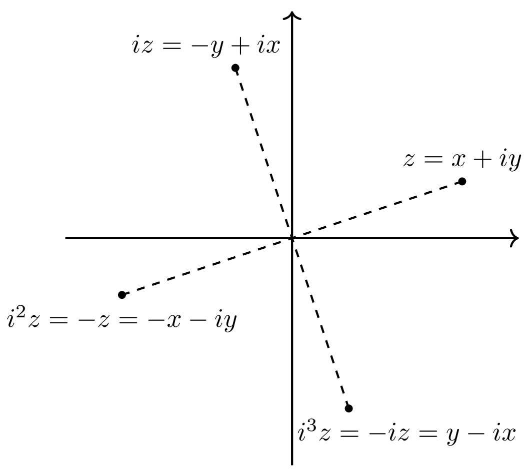
$$z = (x, y) = (x, 0) + (0, y) = (x, 0) + (0, 1) \cdot (y, 0) = x + iy.$$

Dodawanie i mnożenie liczb zespolonych $z_1 = x_1 + iy_1$ oraz $z_2 = x_2 + iy_2$ w postaci algebraicznej przyjmuje postać

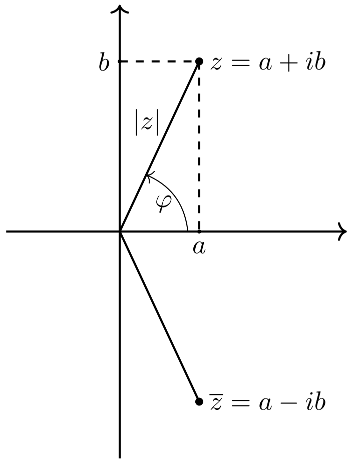
$$\begin{aligned} z_1 + z_2 &= (x_1 + x_2) + i(y_1 + y_2), \\ z_1 \cdot z_2 &= (x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2). \end{aligned}$$

W tej konwencji zapisu liczby zespolone mnożymy tak samo jak liczby rzeczywiste pamiętając, że $i^2 = -1$. Dla przykładu,

$$\begin{aligned} (1 + i\sqrt{2})(4 + 3i) &= 4 + 4i\sqrt{2} + 3i + 3\sqrt{2}i^2 \\ &= 4 - 3\sqrt{2} + i(4\sqrt{2} + 3). \end{aligned}$$




Rysunek 1.1: Mnożenie przez $i$ jest obrotem o $\pi/2$.




Rysunek 1.2: Liczba zespolona $z$, jej argument, moduł i liczba zespolona $\overline{z}$ sprzężona do $z$.

**Przykład 1.3.** Zapiszemy liczbę $\frac{1}{1+3i}$ w postaci $a + bi$. Mamy

$$\begin{aligned}\frac{1}{1+3i} &= \frac{1}{1+3i} \cdot \frac{1-3i}{1-3i} \\ &= \frac{1-3i}{1+3^2} \\ &= \frac{1-3i}{10}.\end{aligned}$$

Każdą liczbę zespoloną $z = x + iy \neq 0$ można zapisać w postaci

$$z = x + iy = |z| \left( \frac{x}{|z|} + i \frac{y}{|z|} \right).$$

Ponadto

$$\frac{x}{|z|} = \cos\varphi, \quad \frac{y}{|z|} = \sin\varphi, \tag{1.3}$$

gdzie $\varphi$ jest kątem między między osią rzeczywistą a półprostą o początku w punkcie $(0, 0)$ przechodzącą przez punkt $(x, y)$. Otrzymujemy stąd *postać trygonometryczną* liczby zespolonej:

$$z = |z|(\cos\varphi + i \sin\varphi). \tag{1.4}$$

Każdą liczbę $\varphi$ spełniającą równania (1.3) nazywamy *argumentem* liczby zespolonej $z \neq 0$ i oznaczamy $\arg(z)$.

Argument $\arg(z)$ nie jest wyznaczony jednoznacznie. Jeżeli $\varphi$ jest pewnym argumentem $z$, to każdy inny argument $z$ jest postaci

$$\varphi + 2k\pi,$$

dla pewnego $k \in \mathbb{Z}$. Jeśli $z = 0$, to równość (1.4) zachodzi dla dowolnej liczby rzeczywistej $\varphi$. *Argumentem głównym* liczby zespolonej $z \neq 0$ nazywamy liczbę rzeczywistą $\varphi \in (-\pi, \pi]$ określoną równaniami (1.3). Oznaczamy go przez $\mathrm{Arg}\,(z)$.

Z powyższy rozważań wynika, że liczba zespolona $z \neq 0$ jest jednoznacznie wyznaczona przez swój moduł i argument. Wynika stąd, że dwie liczby zespolone są równe wtedy i tylko wtedy, gdy mają równe moduły i ich argumenty różnią się o całkowitą wielokrotność $2\pi$.

**Przykład 1.4.** Zapiszemy liczbę $z = 1+i\sqrt{3}$ w postaci trygonometrycznej. Mamy $|z| = \sqrt{4} = 2$ oraz

$$\cos\theta = \frac{1}{2}, \quad \sin\theta = \frac{\sqrt{3}}{2}.$$

Stąd $\theta = \frac{\pi}{3}$, czyli

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

**Lemat 1.1.** *Dla danych w postaci trygonometrycznej liczb zespolonych*

$$z_1 = |z_1|(\cos\phi + i\sin\phi), \qquad z_2 = |z_2|(\cos\psi + i\sin\psi)$$

*zachodzi wzór*

$$z_1 \cdot z_2 = |z_1||z_2|(\cos(\phi+\psi) + i\sin(\phi+\psi)).$$

*Dowód.* Bezpośredni rachunek pokazuje, że

$$\begin{aligned} z_1 \cdot z_2 &= |z_1|(\cos\phi + i\sin\phi) \cdot |z_2|(\cos\psi + i\sin\psi) \\ &= |z_1||z_2|(\cos\phi\cos\psi - \sin\phi\sin\psi + i(\sin\phi\cos\psi + \cos\phi\sin\psi)) \\ &= |z_1||z_2|(\cos(\phi+\psi) + i\sin(\phi+\psi)). \end{aligned}$$ □

**Twierdzenie 1.2** (Wzór de Moivre'a)**.** *Dla liczby zespolonej* $z = |z|(\cos\phi + i\sin\phi)$ *oraz dla* $n \in \mathbb{N}$ *zachodzi wzór*

$$z^n = |z|^n(\cos n\phi + i\sin n\phi).$$

*W szczególności, jeśli* $|z| = 1$*, to*

$$(\cos\phi + i\sin\phi)^n = \cos n\phi + i\sin n\phi.$$

*Dowód.* Zastosujemy indukcję względem $n$. Wzór zachodzi dla $n = 1$. Zakładamy teraz, że jest on prawdziwy dla pewnej liczby naturalnej $n$. Udowodnimy, że zachodzi on również dla $n+1$. Mamy

$$\begin{aligned} z^{n+1} &= z \cdot z^n \\ &= |z|(\cos\phi + i\sin\phi) \cdot |z|^n(\cos n\phi + i\sin n\phi) \\ &= |z|^{n+1}(\cos(n+1)\phi + i\sin(n+1)\phi). \end{aligned}$$ □

**Przykład 1.5.** Obliczymy $(1+i\sqrt{3})^5$. Ponieważ $1 + i\sqrt{3} = 2\left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right)$, więc

$$\begin{aligned} (1+i\sqrt{3})^5 &= 2^5\left(\cos\frac{5\pi}{3} + i\sin\frac{5\pi}{3}\right) \\ &= 32\left(\frac{1}{2} - i\frac{\sqrt{3}}{2}\right). \end{aligned}$$

**Definicja 1.4** (Pierwiastek z liczby zespolonej)**.** Niech $n \in \mathbb{N}$. Liczbę zespoloną $w$ nazywamy pierwiastkiem stopnia $n$ z liczby zespolonej $z$, jeśli $w^n = z$.

Niech $w = |w|(\cos\psi + i\sin\psi)$ będzie pierwiastkiem stopnia $n$ z liczby $z = |z|(\cos\phi + i\sin\phi) \neq 0$. Ze wzoru de Moivre'a wynika, że wtedy

$$|w|^n(\cos n\psi + i\sin n\psi) = |z|(\cos\phi + i\sin\phi).$$

Stąd $|w|^n = |z|$ oraz

$$n\psi = \phi + 2k\pi,$$

dla pewnego $k \in \mathbb{Z}$. Wynika stąd, że dla pewnego $k \in \mathbb{Z}$ mamy

$$w = \sqrt[n]{|z|}\left(\cos\left(\frac{\phi + 2k\pi}{n}\right) + i\sin\left(\frac{\phi + 2k\pi}{n}\right)\right). \tag{1.5}$$

Łatwo sprawdzamy, że istnieje dokładnie $n$ różnych pierwiastków stopnia $n$ z liczby $z \neq 0$ danych wzorami

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

U Na pierwiastki $w_k$ stopnia $n$ z niezerowej liczby zespolonej $z = |z|(\cos\phi + i\sin\phi)$ możemy popatrzeć następująco. Niech

$$u = \sqrt[n]{|z|}\left(\cos\left(\frac{\phi}{n}\right) + i\sin\left(\frac{\phi}{n}\right)\right)$$

będzie oczywistym pierwiastkiem. Dla pierwiastka stopnia $n$ z jedynki:

$$\omega = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$$

mamy

$$w_k = u\omega^k, \quad k = 0, \ldots, n-1,$$

czyli pierwiastkami stopnia $n$ z liczby $z \in \mathbb{C}$ są

$$u,\ u\omega,\ u\omega^2,\ \ldots,\ u\omega^{n-1}.$$

**Przykład 1.6.** Wyznaczymy pierwiastki zespolone stopnia trzeciego z liczby

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

Dla $k = 0$ mamy

$$w_0 = \sqrt[3]{2}\left(\cos\frac{\pi}{9} + i\sin\frac{\pi}{9}\right).$$

Dla $k = 1$ otrzymujemy

$$\begin{aligned} w_1 &= \sqrt[3]{2}\left(\cos\frac{\frac{\pi}{3} + 2\pi}{3} + i\sin\frac{\frac{\pi}{3} + 2\pi}{3}\right) \\ &= \sqrt[3]{2}\left(\cos\frac{7\pi}{9} + i\sin\frac{7\pi}{9}\right), \end{aligned}$$

a dla $k = 2$ mamy

$$\begin{aligned} w_2 &= \sqrt[3]{2}\left(\cos\frac{\frac{\pi}{3} + 4\pi}{3} + i\sin\frac{\frac{\pi}{3} + 4\pi}{3}\right) \\ &= \sqrt[3]{2}\left(\cos\frac{13\pi}{9} + i\sin\frac{13\pi}{9}\right). \end{aligned}$$

**Przykład 1.7.** Pierwiastkami zespolonymi stopnia 4 z jedynki są liczby

$$1, \quad i, \quad -1, \quad -i.$$

U Ponieważ pierwiastki $w_k$ dla $k = 0, 1, \ldots, n-1$ mają moduły równe $\sqrt[n]{|z|}$ i kolejne argumenty różnią się o $\frac{2\pi}{n}$, więc $w_k$ są wierzchołkami $n$-kąta foremnego wpisanego w okrąg o środku w punkcie $(0, 0)$ i promieniu $\sqrt[n]{|z|}$.

**Przykład 1.8** (Pierwiastki zespolone wielomianu rzeczywistego)**.** Rozważmy równanie wielomianowe o współczynnikach rzeczywistych $a_0, \ldots, a_n \in \mathbb{R}$:

$$a_n z^n + \ldots + a_1 z + a_0 = 0, \quad a_n \neq 0.$$

Uzasadnimy, że jeżeli liczba zespolona $z$ jest pierwiastkiem tego równania, to jest nim również $\overline{z}$. Nie jest to prawda, gdy współczynniki $a_0, \ldots, a_n$ są zespolone, ale nie są rzeczywiste. Jeśli

$$a_n z^n + \ldots + a_1 z + a_0 = 0,$$

to

$$\begin{aligned} 0 = \overline{0} &= \overline{a_n z^n + \ldots + a_1 z + a_0} \\ &= \overline{a_n} \cdot \overline{z}^n + \ldots + \overline{a_1} \cdot \overline{z} + \overline{a_0} \\ &= a_n \overline{z}^n + \ldots + a_1 \overline{z} + a_0. \end{aligned}$$

**Twierdzenie 1.3** (Zasadnicze twierdzenie algebry)**.** *Każdy wielomian*

$$w(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0, \quad a_i \in \mathbb{C}$$

*o współczynnikach zespolonych ma n pierwiastków zespolonych (liczonych z krotnościami)* $z_1, \ldots, z_n$*. Ponadto,*

$$w(z) = (z - z_1)(z - z_2)\ldots(z - z_n).$$

Pierwszy dowód twierdzenia 1.3 podał Gauss w 1799 roku.

U

Rozkład wielomianu rzeczywistego nad $\mathbb{R}$

Z zasadniczego twierdzenia algebry wynika, że wielomian o współczynnikach zespolonych $p(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0$ ma rozkład na iloczyn wielomianów stopnia pierwszego

$$p(z) = (z - z_1)\ldots(z - z_n),$$

gdzie $z_1, \ldots, z_n \in \mathbb{C}$ są pierwiastkami $p$. Zastanówmy się, co możemy powiedzieć o rozkładzie wielomianu $p(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0$ o współczynnikach rzeczywistych. Jak wiemy, jeśli liczba zespolona $z_1 = a + ib$ z $b \neq 0$ jest jego pierwiastkiem, to jest nim również liczba sprzężona $z_2 = a - ib$. Zauważmy, że

$$\begin{aligned}(z - z_1)(z - z_2) &= (z - (a + ib))(z - (a - ib)) \\ &= z^2 - 2az + a^2 + b^2,\end{aligned}$$

jest wielomianem stopnia 2 o współczynnikach rzeczywistych nie posiadającym pierwiastków rzeczywistych. Wynika stąd, że $p$ możemy rozłożyć na iloczyn wielomianów stopnia jeden i wielomianów nierozkładalnych stopnia dwa o współczynnikach rzeczywistych.

**Przykład 1.9.** Można sprawdzić, że wielomian $w(z) = z^4 - 5z^2 - 10z - 6$ ma pierwiastki zespolone $-1, 3, -1 + i, -1 - i$. Jako wielomian o współczynnikach zespolonych rozkłada się on na czynniki stopnia pierwszego:

$$w(z) = (z + 1)(z - 3)(z + 1 + i)(z + 1 - i).$$

Jeśli chcemy, go rozłożyć na iloczyn wielomianów o współczynnikach rzeczywistych, to otrzymujemy

$$w(z) = (z + 1)(z - 3) \underbrace{(z^2 + 2z + 2)}_{=(z+1+i)(z+1-i)}.$$

**Przykład 1.10** (Pierwiastki zespolone równania kwadratowego)**.** Rozważmy równanie kwadratowe

$$az^2 + bz + c = 0, \quad a \neq 0$$

o współczynnikach rzeczywistych. Jeśli $\Delta = b^2 - 4ac < 0$, to ma ono dwa sprzężone rozwiązania zespolone dane wzorami

$$z_1 = \frac{-b - i\sqrt{|\Delta|}}{2a}, \quad z_2 = \frac{-b + i\sqrt{|\Delta|}}{2a}.$$

Przykładowo, dla

$$z^2 + z + 1 = 0$$

mamy $\Delta = -3$ oraz

$$z_1 = \frac{-1 - i\sqrt{3}}{2}, \quad z_2 = \frac{-1 + i\sqrt{3}}{2}.$$

U

Pierwiastki równania kwadratowego

Rozważmy równanie kwadratowe

$$az^2 + bz + c = 0, \quad a \neq 0$$

o współczynnikach zespolonych $a, b, c \in \mathbb{C}$. Dla $\Delta = b^2 - 4ac$ możemy je zapisać w postaci

$$a\left[\left(z + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a^2}\right] = 0,$$

co jest równoważne z równością

$$\left(z + \frac{b}{2a}\right)^2 = \frac{\Delta}{4a^2},$$

czyli

$$(2az + b)^2 = \Delta.$$

Wystarczy więc znaleźć pierwiastki kwadratowe z liczby zespolonej $\Delta$.
Zauważmy, że jeśli $z_1, z_2 \in \mathbb{C}$ są pierwiastkami równania $az^2 + bz + c = 0$, to

$$az^2 + bz + c = a(z - z_1)(z - z_2)$$

oraz zachodzą wzory Viète'a

$$z_1 + z_2 = -\frac{b}{a}, \quad z_1 \cdot z_2 = \frac{c}{a}.$$

**Przykład 1.11.** Rozwiążemy równanie

$$4z^2 + 4iz + (-13 - 16i) = 0, \quad z \in \mathbb{C}.$$

Jest ono równoważne z równaniem $z^2 + iz + \dfrac{-13 - 16i}{4} = 0$, czyli

$$\left(z + \frac{i}{2}\right)^2 + \frac{1}{4} - \frac{13}{4} - 4i = 0.$$

Stąd

$$\left(z + \frac{i}{2}\right)^2 = 3 + 4i.$$

Znajdziemy pierwiastki stopnia 2 z liczby $3 + 4i$. Jeśli $(a + ib)^2 = 3 + 4i$, to

$$a^2 - b^2 = 3, \quad 2ab = 4.$$

Z drugiego równania $a$ i $b$ są różne od zera i $b = \frac{2}{a}$. Z pierwszego równania mamy

$$a^2 - \frac{4}{a^2} = 3,$$

czyli

$$a^4 - 3a^2 - 4 = 0.$$

Stąd

$$(a^2 + 1)(a^2 - 4) = 0,$$

czyli $a = \pm 2$. W konsekwencji, $b = \pm 1$. Mamy więc, że

$$z + \frac{i}{2} = 2 + i \quad \text{lub} \quad z + \frac{i}{2} = -2 - i,$$

czyli

$$z = 2 + \frac{i}{2} \quad \text{lub} \quad z = -2 - \frac{3i}{2}.$$

## 1.3 Grupy i ciała

Wprowadzimy teraz pewne podstawowe struktury algebraiczne. Ten fragment będzie nieco bardziej abstrakcyjny i wymaga od Was szczególnej uwagi. Ten trud opłaci się w Waszym dalszym matematycznym życiu. Podejmijcie go! Robimy to nie tylko z miłości do abstrakcji, ale przede wszystkim dlatego, że wprowadzenie i używanie tego języka jest po prostu wygodne.

Niech $X$ będzie niepustym zbiorem.

**Definicja 1.5** (Działanie wewnętrzne w zbiorze)**.** *Działaniem wewnętrznym w zbiorze* $X$ nazywamy dowolne odwzorowanie

$$\circ : X \times X \ni (x, y) \longmapsto x \circ y := \circ(x, y) \in X$$

przyporządkowujące uporządkowanej parze elementów zbioru $X$ element zbioru $X$.

**Przykład 1.12.** Dodawanie i mnożenie liczb rzeczywistych (zespolonych) są działaniami wewnętrznymi w zbiorze $\mathbb{R}$ ($\mathbb{C}$). Ponieważ $2 - 3 = -1$, więc odejmowanie nie jest działaniem wewnętrznym w zbiorze liczb naturalnych $\mathbb{N}$. Dzielenie $x \circ y = \frac{x}{y}$ nie jest działaniem wewnętrznym w zbiorze liczb rzeczywistych $\mathbb{R}$, bo nie jest zdefiniowane dla pary $(2, 0)$ – nie wolno dzielić przez 0.

**Definicja 1.6** (Grupa i grupa abelowa)**.** Niech $\circ : X \times X \longrightarrow X$ będzie działaniem w zbiorze $X$. Para $(X, \circ)$ jest *grupą*, jeśli zachodzą warunki

(G1) dla dowolnych $x, y, z \in X$ zachodzi

$$(x \circ y) \circ z = x \circ (y \circ z), \quad \text{łączność działania,}$$

(G2) istnieje taki $e \in X$, że dla dowolnego $x \in X$ zachodzi

$$e \circ x = x = x \circ e, \quad \text{istnieje element neutralny,}$$

(G3) dla dowolnego $x \in X$ istnieje taki $x' \in X$, że

$$x \circ x' = x' \circ x = e, \quad \text{istnieje element odwrotny,}$$

Jeżeli ponadto

(G4) dla dowolnych $x, y \in X$ mamy

$$x \circ y = y \circ x, \quad \text{przemienność działania,}$$

to $(G, \circ)$ nazywamy *grupą przemienną* lub *abelową*.

U

Element neutralny $e$ w grupie $G$ jest wyznaczony jednoznacznie. Rzeczywiście, jeśli $e, e^* \in G$ są elementami neutralnymi, to

$$e = e \circ e^* = e^*.$$

Podobnie, element odwrotny $x'$ do $x$ jest wyznaczony jednoznacznie. Jeśli $x', x''$ są odwrotne do $x$, to

$$\begin{aligned} x' &= x' \circ e \\ &= x' \circ (x \circ x'') \\ &= (x' \circ x) \circ x'' \\ &= e \circ x'' \\ &= x''. \end{aligned}$$

TEST Oceń które ze zdań jest prawdziwe

- $(\mathbb{N}, +)$ jest grupą abelową;
- $(\mathbb{Z}, +)$ jest grupą abelową;
- $(\mathbb{Q}, +)$ jest grupą abelową;
- $(\mathbb{R}, +)$ jest grupą abelową;
- $(\mathbb{C}, +)$ jest grupą abelową;
- $((0, +\infty), +)$ jest grupą abelową;
- $(\mathbb{N}, \cdot)$ jest grupą abelową;
- $(\mathbb{Z}, \cdot)$ jest grupą abelową;
- $(\mathbb{Q}, \cdot)$ jest grupą abelową;
- $(\mathbb{R}, \cdot)$ jest grupą abelową;
- $(\mathbb{C}, \cdot)$ jest grupą abelową;
- $(\mathbb{Z} \setminus \{0\}, \cdot)$ jest grupą abelową;
- $(\mathbb{Q} \setminus \{0\}, \cdot)$ jest grupą abelową;
- $(\mathbb{R} \setminus \{0\}, \cdot)$ jest grupą abelową;
- $(\mathbb{C} \setminus \{0\}, \cdot)$ jest grupą abelową.

$\longleftarrow$

**Przykład 1.13.** Niech

$$G = \{x + y\sqrt{5} : x, y \in \mathbb{Q},\ x^2 - 5y^2 = 1\}.$$

Pokażemy, że $G$ z mnożeniem liczb rzeczywistych jest grupą abelową. Musimy przede wszystkim pokazać, że jest to działanie wewnętrzne w $G$. Niech $x + y\sqrt{5}, u + w\sqrt{5} \in G$. Oznacza to, że $x, y, u, w \in \mathbb{Q}$ oraz

$$x^2 - 5y^2 = 1 = u^2 - 5w^2.$$

Mamy

$$(x + y\sqrt{5})(u + w\sqrt{5}) = \underbrace{xu + 5yw}_{\in\mathbb{Q}} + \underbrace{(xw + yu)}_{\in\mathbb{Q}}\sqrt{5}.$$

Musimy pokazać, że $(xu + 5yw)^2 - 5(xw + yu)^2 = 1$. Sprawdzamy, że

$$\begin{aligned}(xu + 5yw)^2 - 5(xw + yu)^2 &= x^2u^2 + 10xuyw + 25y^2w^2 - 5x^2w^2 - 10xwyu - 5y^2u^2 \\ &= x^2u^2 + 25y^2w^2 - 5x^2w^2 - 5y^2u^2 \\ &= \underbrace{(x^2 - 5y^2)}_{=1}\underbrace{(u^2 - 5w^2)}_{=1} \\ &= 1.\end{aligned}$$

Mnożenie liczb rzeczywistych jest łączne i przemienne, więc warunki (G1) i (G4)są spełnione. Ponadto, $1 \in G$, więc zachodzi warunek (G2). Zauważmy, że jeśli $x + y\sqrt{5} \in G$, to z warunku $x^2 - 5y^2 = 1$ wynika, że $x \neq 0$. Ponadto $x + y\sqrt{5} \neq 0$, bo wtedy $\sqrt{5}$ byłby liczbą wymierną, gdy $y \neq 0$. Łatwo sprawdzamy, że $x - y\sqrt{5} \in G$ oraz

$$(x + y\sqrt{5})(x - y\sqrt{5}) = x^2 - 5y^2 = 1,$$

czyli warunek (G3) zachodzi.

**Przykład 1.14** (Grupa $\mathbb{S}^1$)**.** Poprawne rozwiązanie testu 1.3 prowadzi nas do wniosku, że $(\mathbb{C} \setminus \{0\}, \cdot)$ jest grupą abelową. Oznacza to, że działanie mnożenia

$$\cdot : \mathbb{C} \setminus \{0\} \times \mathbb{C} \setminus \{0\} \ni (z, w) \longmapsto z \cdot w \in \mathbb{C} \setminus \{0\}$$

spełnia warunki (G1)–(G4)z definicji grupy. Zbiór $\mathbb{C} \setminus \{0\}$ jest płaszczyzną z usuniętym początkiem układu współrzędnych. Rozważmy okrąg jednostkowy o środku w początku układu

$$\mathbb{S}^1 := \{z \in \mathbb{C} : |z| = 1\}.$$

Oczywiście $\mathbb{S}^1 \subset \mathbb{C} \setminus \{0\}$. Zauważmy, że jeżeli $z_1,\ z_2 \in \mathbb{S}^1$, to $z_1 \cdot z_2 \in \mathbb{S}^1$, gdyż

$$|z_1 \cdot z_2| = |z_1||z_2| = 1.$$

Oznacza to, że iloczyn liczb zespolonych ze zbioru $\mathbb{S}^1$ jest ponownie liczbą należącą do $\mathbb{S}^1$. Wynika stąd, że mnożenie liczb zespolonych jest działaniem wewnętrznym w zbiorze $\mathbb{S}^1$ i możemy je traktować jako odwzorowanie

$$\cdot : \mathbb{S}^1 \times \mathbb{S}^1 \longrightarrow \mathbb{S}^1.$$

W takim razie może $(\mathbb{S}^1, \cdot)$ jest również grupą?

U $\boxed{\text{Podgrupa}}$

Przypuśćmy, że $(G, \circ)$ jest grupą i $H \subset G$ jest takim jej niepustym podzbiorem, że $x \circ y \in H$ dla $x, y \in H$. Oznacza to, że $\circ$ jest również działaniem wewnętrznym w $H$. To jeszcze za mało, aby $(H, \circ)$ było grupą. Rozważmy prosty przykład grupy liczb całkowitych $G = \mathbb{Z}$ z działaniem dodawania liczb całkowitych $\circ = +$. Wtedy jest to również działanie wewnętrzne w zbiorze liczb naturalnych $H = \mathbb{N} \subset G = \mathbb{Z}$. Zauważmy, że $(\mathbb{N}, +)$ nie jest grupą. Przede wszystkim dodawanie nie ma elementu neutralnego w $\mathbb{N}$, bo $0 \notin \mathbb{N}$. W takim razie rozważmy podzbiór $\mathbb{N}_0 = \{0\} \cup \mathbb{N} \subset \mathbb{Z}$. Oczywiście, dodawanie liczb całkowitych jest również działaniem wewnętrznym w $\mathbb{N}_0$ i nie mamy już problemu z elementem neutralnym. Pojawia się jednak nowy problem, gdyż przykładowo, element $1 \in \mathbb{N}_0$ nie ma elementu odwrotnego (przeciwnego) w zbiorze $\mathbb{N}_0$, bo $-1 \notin \mathbb{N}_0$.

Z naszych rozważań wynika, że podzbiór $H \subset G$ dziedziczy strukturę grupy z $(G, \circ)$, jeśli spełnione są warunki:

- $x \circ y \in H$ dla wszystkich $x, y \in H$,
- $e \in H$, gdzie $e \in G$ jest elementem neutralnym dla $\circ$,
- $x' \in H$, gdzie $x' \in G$ jest elementem odwrotnym do $x \in H$.

Wtedy rzeczywiście $(H, \circ)$ jest grupą. Algebraicy mówią wtedy, że $(H, \circ)$ jest *podgrupą* grupy $(G, \circ)$.

Wróćmy do okręgu $\mathbb{S}^1$. Zauważmy, że element neutralny 1, dla mnożenia liczb zespolonych, należy do zbioru $\mathbb{S}^1$. Ponadto, dla $z \in \mathbb{S}^1$ mamy $z^{-1} = \overline{z} \in \mathbb{S}^1$, bo $|\overline{z}| = |z|$. Ponieważ warunki (G1)–(G4)zachodzą w zbiorze $\mathbb{C} \setminus \{0\}$, więc tym bardziej są spełnione w mniejszym zbiorze $\mathbb{S}^1$. Wynika stąd, że para $(\mathbb{S}^1, \cdot)$ jest również grupą abelową. Jest ona podgrupą grupy $(\mathbb{C} \setminus \{0\}, \cdot)$.

**Przykład 1.15** (Grupa zespolonych pierwiastków z jedynki)**.** Zdefiniujemy teraz jeszcze pewne ważne podgrupy grupy $(\mathbb{S}^1, \cdot)$. Będą one składały się ze skończonej liczby elementów. Ustalmy dowolną liczbę naturalną $n \geqslant 1$. Rozważmy zbiór $G_n$ wszystkich pierwiastków stopnia $n$ z liczby 1, czyli

$$G_n = \{\epsilon_k : k = 0, 1, \dots, n-1\},$$

gdzie

$$\epsilon_k = \cos\left(\frac{2k\pi}{n}\right) + i \sin\left(\frac{2k\pi}{n}\right).$$

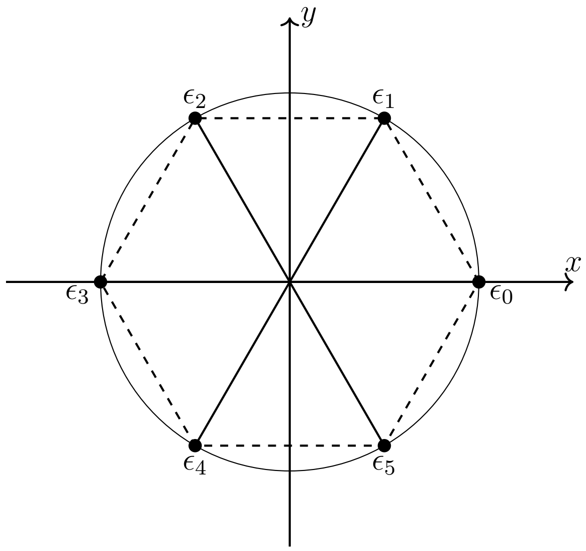
Zauważmy, że dla $z, w \in G_n$ mamy

$$(z \cdot w)^n = z^n \cdot w^n = 1 \cdot 1 = 1.$$

Wynika stąd, że mnożenie liczb zespolonych jest działaniem wewnętrznym w zbiorze $G_n$. Ponadto, $1 \in G_n$ oraz $z^{-1} = \overline{z}$, więc $z^{-1} \in G_n$, czyli $(G_n, \cdot)$ jest grupą abelową.

Podamy teraz abstrakcyjną definicję algebraicznej struktury ciała z którą zetknęliśmy się już w przypadku liczb rzeczywistych i zespolonych. Rozważmy zbiór $\mathbb{F}$ w którym określone są dwa działania wewnętrzne

$$+ : \mathbb{F} \times \mathbb{F} \longrightarrow \mathbb{F}, \quad \cdot : \mathbb{F} \times \mathbb{F} \longrightarrow \mathbb{F}.$$




Rysunek 1.3: Pierwiastki zespolone szóstego stopnia z jedynki. Stanowią one grupę abelową z działaniem mnożenia liczb zespolonych.

Zgodnie z przyjętym (całkowicie umownie) oznaczeniem będziemy je nazywali dodawaniem i mnożeniem, chociaż nie muszą one mieć nic wspólnego ze znanymi nam działaniami arytmetycznymi w zbiorach liczbowych.

**Definicja 1.7** (Ciało)**.** Trójka $(\mathbb{F}, +, \cdot)$ jest *ciałem*, jeśli zachodzą warunki

(F1) para $(\mathbb{F}, +)$ jest grupą abelową z elementem neutralnym $0 \in \mathbb{F}$ dla działania $+$;

(F2) para $(\mathbb{F} \setminus \{0\}, \cdot)$ jest grupą abelową z działaniem $\cdot$;

(F3) $x \cdot (y + z) = x \cdot y + x \cdot z$ dla dowolnych $x, y, z \in \mathbb{F}$.

! W punkcie (F2) w sposób niejawny zakładamy, że iloczyn elementów różnych od 0 jest różny od 0, bo mnożenie $\cdot$ jest z założenia działaniem wewnętrznym w zbiorze $\mathbb{F} \setminus \{0\}$ (definicja grupy). Z drugiej strony dla dowolnego $x \in \mathbb{F}$ zachodzi równość

$$0 \cdot x = 0,$$

bo

$$0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x.$$

**Przykład 1.16** (Rozszerzenie ciała $\mathbb{Q}$ o $\sqrt{2}$)**.** Rozważmy zbiór

$$\mathbb{Q}(\sqrt{2}) := \left\{ a + b\sqrt{2} : a, b \in \mathbb{Q} \right\}.$$

Elementy zbioru $\mathbb{Q}(\sqrt{2})$ są liczbami rzeczywistymi, czyli $\mathbb{Q}(\sqrt{2}) \subset \mathbb{R}$. Pokażemy, że struktura ciała $(\mathbb{R}, +, \cdot)$ „dziedziczy się” na zbiór $\mathbb{Q}(\sqrt{2})$ z działaniami dodawania i mnożenia liczb rzeczywistych zawężonymi do zbioru $\mathbb{Q}(\sqrt{2})$. Zauważmy najpierw, że

$$0 = 0 + 0\sqrt{2} \in \mathbb{Q}(\sqrt{2}), \qquad 1 = 1 + 0\sqrt{2} \in \mathbb{Q}(\sqrt{2}),$$

czyli elementy neutralne działań należą do zbioru $\mathbb{Q}(\sqrt{2})$. Ponadto, działania są wewnętrzne w $\mathbb{Q}(\sqrt{2})$, bo

$$(a + b\sqrt{2}) + (a_1 + b_1\sqrt{2}) = \underbrace{a + a_1}_{\in\mathbb{Q}} + \underbrace{(b + b_1)}_{\in\mathbb{Q}}\sqrt{2} \in \mathbb{Q}(\sqrt{2}),$$

oraz

$$(a + b\sqrt{2}) \cdot (a_1 + b_1\sqrt{2}) = \underbrace{aa_1 + 2bb_1}_{\in\mathbb{Q}} + \underbrace{(ab_1 + ba_1)}_{\in\mathbb{Q}}\sqrt{2} \in \mathbb{Q}(\sqrt{2}).$$

Zauważmy, że

$$(a+b\sqrt{2})+(-a-b\sqrt{2})=0,$$

więc element odwrotny do $a+b\sqrt{2}$ względem działania $+$, czyli $-a-b\sqrt{2}$ jest elementem zbioru $\mathbb{Q}(\sqrt{2})$.

Załóżmy, że $a+b\sqrt{2}\neq 0$. Wtedy $a-b\sqrt{2}\neq 0$. Stąd

$$a^2-2b^2=(a+b\sqrt{2})(a-b\sqrt{2})\neq 0,$$

czyli

$$(a+b\sqrt{2})\cdot\left(\frac{a-b\sqrt{2}}{a^2-2b^2}\right)=1,$$

więc

$$\frac{a-b\sqrt{2}}{a^2-2b^2}=\frac{a}{a^2-2b^2}-\frac{b}{a^2-2b^2}\sqrt{2}\in\mathbb{Q}(\sqrt{2})$$

jest elementem odwrotnym do $a+b\sqrt{2}$. Z powyższych rozważań wynika, że $(\mathbb{Q}(\sqrt{2}),+,\cdot)$ jest ciałem.

**Przykład 1.17.** Rozważmy zbiór

$$X=\{a+b\sqrt[3]{5}:a,b\in\mathbb{Q}\}.$$

Uzasadnimy, że nie jest on ciałem z dodawaniem i mnożeniem liczb rzeczywistych. Problem mamy z mnożeniem, bo nie jest ono działaniem wewnętrznym w zbiorze $X$. Pokażemy bowiem, że $(\sqrt[3]{5})^2\notin X$. Przypuśćmy, że $(\sqrt[3]{5})^2=a+b\sqrt[3]{5}$ dla pewnych $a,b\in\mathbb{Q}$. Mnożąc przez $\sqrt[3]{5}$ dostajemy, że

$$\begin{aligned}5&=a\sqrt[3]{5}+b(\sqrt[3]{5})^2\\&=a\sqrt[3]{5}+ab+b^2\sqrt[3]{5}\\&=ab+\sqrt[3]{5}(a+b^2).\end{aligned}$$

Ponieważ $\sqrt[3]{5}$ jest liczbą niewymierną, więc $a+b^2=0$ i $ab=5$, czyli $-b^3=5$, więc $\sqrt[3]{5}=-b\in\mathbb{Q}$. Otrzymana sprzeczność pokazuje, że $(\sqrt[3]{5})^2\notin X$.

**Przykład 1.18** (Ciało $\mathbb{Z}_2$)**.** Niech $\mathbb{Z}_2=\{0,1\}$. Definiujemy dodawanie $+$ w zbiorze $\mathbb{Z}_2=\{0,1\}$ przez tabelę wyników tego działania

| $+$ | 0 | 1 |
|---|---|---|
| 0 | 0 | 1 |
| 1 | 1 | 0 |

Analogicznie definiujemy mnożenie $\cdot$ w zbiorze $\mathbb{Z}_2=\{0,1\}$ poprzez tabelę

| $\cdot$ | 0 | 1 |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 0 | 1 |

Sprawdzamy łatwo, że tak zdefiniowana trójka $(\mathbb{Z}_2,+,\cdot)$ jest ciałem.

U

Ciało $\mathbb{Z}_p$

Niech $\mathbb{Z}_3=\{0,1,2\}$. Definiujemy działania $+$ oraz $\cdot$ w zbiorze $\mathbb{Z}_3$ przez

| $+$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0 | 1 | 2 |
| 1 | 1 | 2 | 0 |
| 2 | 2 | 0 | 1 |

| $\cdot$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | 2 |
| 2 | 0 | 2 | 1 |

Sprawdzamy łatwo, że $(\mathbb{Z}_3, +, \cdot)$ jest ciałem.

Dla liczby całkowitej $x \in \mathbb{Z}$ i liczby naturalnej $n \geqslant 1$ przez $[x]_n$ oznaczamy resztę z dzielenia liczby $x$ przez $n$. Zauważcie, że działania w $\mathbb{Z}_3$ możemy zapisać następująco:

$$x + y = [x + y]_3, \quad x \cdot y = [x \cdot y]_3, \quad x, y \in \mathbb{Z}_3.$$

Nic nie stoi na przeszkodzie, aby podobnie zdefiniować działania w zbiorze

$$\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$$

dla liczby naturalnej $n \geqslant 1$. Przyjmujemy, że

$$x + y = [x + y]_n, \quad x \cdot y = [x \cdot y]_n, \quad x, y \in \mathbb{Z}_n.$$

Trójka $(\mathbb{Z}_n, +, \cdot)$ nie zawsze jest ciałem. Elementem neutralnym dodawania jest 0. Zauważmy, że przykładowo w $\mathbb{Z}_6$ mamy

$$2 \cdot 3 = [6]_6 = 0.$$

Oznacza to, że $\mathbb{Z}_6$ nie jest ciałem. Analogiczny argument pokazuje, że $\mathbb{Z}_n$ nie jest ciałem, gdy $n$ nie jest liczbą pierwszą. Można sprawdzić, że (ćwiczenie) $(\mathbb{Z}_n, +, \cdot)$ jest ciałem wtedy i tylko wtedy, gdy $n$ jest liczbą pierwszą.

U Ciało skończone nie musi być równe $\mathbb{Z}_p$ dla liczby pierwszej $p$. Poniższe tabelki definiują ciało 4-elementowe:

| + | 0 | 1 | a | b |
|---|---|---|---|---|
| 0 | 0 | 1 | a | b |
| 1 | 1 | 0 | b | a |
| a | a | b | 0 | 1 |
| b | b | a | 1 | 0 |

| $\cdot$ | 0 | 1 | a | b |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | a | b |
| a | 0 | a | b | 1 |
| b | 0 | b | 1 | a |

$(\mathbb{Z}_4, +, \cdot)$ nie jest ciałem, bo $2 \cdot 2 = 0$.

## 1.4 Macierze $M_{2\times 2}(\mathbb{F})$

Niech $\mathbb{F}$ będzie ciałem. Przez *macierz* $A = [a_{ij}]$ wymiaru 2 rozumiemy tablicę

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad a_{ij} \in \mathbb{F}$$

mającą dwa wiersze i dwie kolumny. Zbiór wszystkich takich macierzy będziemy oznaczać przez $M_{2\times 2}(\mathbb{F})$. Powiemy, że $A = [a_{ij}], B = [b_{ij}] \in M_{2\times 2}(\mathbb{F})$ są *równe* ($A = B$), jeśli $a_{ij} = b_{ij}$ dla $i, j = 1, 2$.

W zbiorze $M_{2\times 2}(\mathbb{F})$ definiujemy działanie dodawania

$$+ : M_{2\times 2}(\mathbb{F}) \times M_{2\times 2}(\mathbb{F}) \longrightarrow M_{2\times 2}(\mathbb{F})$$

wzorem

$$A + B = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}.$$

Sprawdzamy łatwo, że $(M_{2\times 2}(\mathbb{F}), +)$ jest grupą abelową. Elementem neutralnym dodawania jest *macierz zerowa*

$$0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

a elementem przeciwnym do $A$ jest macierz

$$-A := \begin{bmatrix} -a_{11} & -a_{12} \\ -a_{21} & -a_{22} \end{bmatrix}.$$

W zbiorze $M_{2\times 2}(\mathbb{F})$ definiujemy działanie mnożenia

$$\cdot : M_{2\times 2}(\mathbb{F}) \times M_{2\times 2}(\mathbb{F}) \longrightarrow M_{2\times 2}(\mathbb{F}).$$

Formuła na $A \cdot B$ jest nieco bardziej skomplikowana i w dalszym ciągu wyjaśni się dlaczego taka właśnie. Definiujemy

$$A \cdot B = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}.$$

Sprawdzamy łatwo, że dla macierzy

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

oraz dowolnej macierzy $A \in M_{2\times 2}(\mathbb{F})$ mamy

$$IA = AI = A.$$

Oznacza to, że $I$ jest elementem neutralnym mnożenia macierzy. Naturalne jest pytanie, czy trójka $(M_{2\times 2}(\mathbb{F}), +, \cdot)$ jest ciałem. Bezpośredni rachunek pokazuje, że spełnione są następujące warunki:

- $(A + B) + C = A + (B + C)$
- $A + 0 = 0 + A = A$,
- $A + (-A) = 0$,
- $A + B = B + A$,
- $(AB)C = A(BC)$,
- $AI = IA = A$,
- $A(B + C) = AB + BC$.

Z definicją ciała mamy jednak kilka problemów. Po pierwsze dla niezerowej macierzy $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ mamy

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

czyli mnożenie macierzy nie jest działaniem wewnętrznym w zbiorze macierzy niezerowych. Nie mamy więc do czynienia z ciałem. Ponadto, niezerowa macierz $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ nie posiada elementu odwrotnego względem mnożenia macierzy. Rzeczywiście, dla dowolnej macierzy $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ mamy

$$AB = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ a & b \end{bmatrix} \neq I.$$

Mnożenie macierzy nie jest też przemienne, bo przykładowo

$$\underbrace{\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}_{B} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\underbrace{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}}_{A} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

**Przykład 1.19.** Dla macierzy $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \in M_{2\times 2}(\mathbb{F})$ mamy

$$AB = B, \quad BA = 0.$$

U Możecie, a nawet powinniście zapytać, dlaczego mnożenie macierzy definiujemy tak dziwną formułą. Moglibyśmy przecież, przez analogię do definicji dodawania macierzy, przyjąć że

$$A \cdot B = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} \\ a_{21}b_{21} & a_{22}b_{22} \end{bmatrix}.$$

Przecież nie jest to takie złe mnożenie. Ma element neutralny, czyli macierz $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Jest ono oczywiście łączne i przemienne. Oczywiście mamy problem z istnieniem elementu odwrotnego do macierzy niezerowej, bo jeśli tylko w macierzy $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ któryś z wyrazów $a, b, c, d$ jest równy zero, to $A$ nie posiada macierzy odwrotnej. Ale przecież podobne problemy mieliśmy z wcześniej zdefiniowanym mnożeniem. Powód takiego zdefiniowania mnożenia macierzy jest nieco głębszy i poznamy go nieco później.

Naturalne jest pytanie, dla jakich macierzy $A$ istnieje taka macierz $B \in M_{2\times 2}(\mathbb{F})$, że

$$AB = BA = I.$$

Okazuje się, że takie macierze $A$ można łatwo scharakteryzować. W tym celu dla macierzy $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ definiujemy element ciała $\mathbb{F}$ wzorem

$$\det A = ad - bc$$

i nazywamy *wyznacznikiem* macierzy $A$.

**Wniosek 1.2.** *Dla* $A, B \in M_{2\times 2}(\mathbb{F})$ *mamy*

$$\det(AB) = \det A \cdot \det B$$

*oraz* $\det I = 1$.

*Dowód.* Sprawdzamy to bezpośrednim rachunkiem. Oczywiście zachodzi równość $\det I = 1$. Dla macierzy $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ oraz $B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$ mamy

$$\begin{aligned} \det(AB) &= \det \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix} \\ &= (ae + bg)(cf + dh) - (ce + dg)(af + bh) \\ &= aedh + bgcf - cebh - dgaf \\ &= (ad - bc)(eh - fg) \\ &= \det A \cdot \det B. \end{aligned}$$

□

**Wniosek 1.3.** *Zachodzą warunki:*

- $\det(A_1 \ldots A_n) = \det A_1 \cdot \ldots \cdot \det A_n$
- $\det(A^n) = (\det A)^n$
- $\det \begin{bmatrix} za & zb \\ zc & zd \end{bmatrix} = z^2 \cdot \det \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad z \in \mathbb{F}$
- $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = -\det \begin{bmatrix} b & a \\ d & c \end{bmatrix}$

- $\det \begin{bmatrix} za & b \\ zc & d \end{bmatrix} = z \cdot \det \begin{bmatrix} a & b \\ c & d \end{bmatrix}$
- $\det \begin{bmatrix} a+e & b \\ c+f & d \end{bmatrix} = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \det \begin{bmatrix} e & b \\ f & d \end{bmatrix}$

Jeśli $A = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{C})$ jest macierzą zespoloną, to definiujemy

$$\overline{A} = \begin{bmatrix} \overline{a} & \overline{b} \\ \overline{c} & \overline{d} \end{bmatrix} \in M_{2\times 2}(\mathbb{C}).$$

Wtedy

- $\det \overline{A} = \overline{\det A}$,
- $\overline{(\overline{A})} = A, \quad \overline{AB} = \overline{A}\,\overline{B}$,
- $\overline{A+B} = \overline{A} + \overline{B}, \quad \overline{zA} = \overline{z}\overline{A}$,
- jeśli $\det A \neq 0$, to $(\overline{A})^{-1} = \overline{A^{-1}}$

**Twierdzenie 1.4.** *Dla macierzy* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{F})$ *następujące warunki są równoważne*

*(1)* $\det A \neq 0$*,*

*(2) istnieje taka macierz* $B \in M_{2\times 2}(\mathbb{F})$*, że*

$$AB = BA = I.$$

*Piszemy wtedy, że* $B = A^{-1}$*.*

*Dowód.* (1) wynika z (2), bo wtedy

$$\det A \cdot \det B = \det(AB) = \det I = 1,$$

czyli $\det A \neq 0$.

Jeśli zachodzi warunek (1), to definiujemy $B$ wzorem

$$B = \frac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Sprawdzamy łatwo, że $AB = BA = I$. □

**Wniosek 1.4.** *Niech* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{F})$*. Jeśli* $\det A \neq 0$*, to*

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Dla macierzy $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{F})$ definiujemy jej macierz *transponowaną*

$$A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

oraz dla $z \in \mathbb{F}$ określamy

$$z \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} za & zb \\ zc & zd \end{bmatrix}.$$

Bezpośredni rachunek pokazuje, że

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

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

(iii) $(z \cdot A)^T = z \cdot A^T$,

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

(v) $\det A^T = \det A$.

Ponadto, jeśli $A$ jest odwracalna, to $A^{-1}A = I$, czyli $A^T(A^{-1})^T = I$, więc $A^T$ jest też odwracalna oraz $(A^T)^{-1} = (A^{-1})^T$.

**Przykład 1.20** (Macierzowa interpretacja ciała $\mathbb{C}$)**.** Wskażemy teraz podzbiór zbioru $M_{2\times 2}(\mathbb{R})$, który pod względem algebraicznym bardzo przypomina ciało liczb zespolonych. Definiujemy

$$M_{\mathbb{C}} := \left\{ A \in M_{2\times 2}(\mathbb{R}) : A = \begin{bmatrix} x & -y \\ y & x \end{bmatrix}, \quad x, y \in \mathbb{R} \right\}.$$

Łatwo sprawdzamy, że jeśli $A = \begin{bmatrix} x & -y \\ y & x \end{bmatrix}, B = \begin{bmatrix} u & -w \\ w & u \end{bmatrix} \in M_{\mathbb{C}}$, to

$$A + B = \begin{bmatrix} x+u & -(y+w) \\ y+w & x+u \end{bmatrix} \in M_{\mathbb{C}}$$

oraz

$$AB = \begin{bmatrix} \operatorname{Re}(x+iy)(u+iw) & -\operatorname{Im}(x+iy)(u+iw) \\ \operatorname{Im}(x+iy)(u+iw) & \operatorname{Re}(x+iy)(u+iw) \end{bmatrix} \in M_{\mathbb{C}},$$

więc dodawanie i mnożenie macierzy jest działaniem wewnętrznym w $M_{\mathbb{C}}$.

Można sprawdzić bezpośrednim rachunkiem, że trójka $(M_{\mathbb{C}}, +, \cdot)$ jest ciałem. My, aby to zauważyć, zastosujemy nieco inne podejście. W tym celu definiujemy

$$f : \mathbb{C} \ni x + iy \longmapsto \begin{bmatrix} x & -y \\ y & x \end{bmatrix} \in M_{\mathbb{C}}.$$

Jest to bijekcja oraz $f(0) = 0$, $f(1) = I$. Ponadto,

$$f(z_1 + z_2) = f(z_1) + f(z_2),$$

oraz

$$f(z_1 z_2) = f(z_1) f(z_2),$$

dla dowolnych $z_1, z_2 \in \mathbb{C}$.

Bijekcja $f$ o powyższych własnościach pozwala na identyfikację struktur algebraicznych $(\mathbb{C}, +, \cdot)$ oraz $(M_{\mathbb{C}}, +, \cdot)$. Algebraicy mówią wtedy, że te ciała są *izomorficzne.* Z algebraicznego punktu widzenia ciała izomorficzne są nierozróżnialne.

Sprawdzamy poprzez bezpośredni rachunek, że $f$ ma dodatkowe własności:

- dla macierzy $J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ mamy

$$J^2 = -I \quad \text{oraz} \quad J^4 = I,$$

- $f(i) = J$,
- $f(z^n) = (f(z))^n$,

- $f(\overline{z}) = (f(z))^T$, gdzie
$$\begin{bmatrix} x & -y \\ y & x \end{bmatrix}^T := \begin{bmatrix} x & y \\ -y & x \end{bmatrix},$$
- $\det f(z) = |z|^2$,
- $f(1/z) = (f(z))^{-1} = \dfrac{1}{|z|^2}(f(z))^T$.

Macierzowa interpretacja ciała $\mathbb{C}$

$$z = x + iy \longmapsto A = \begin{bmatrix} x & -y \\ y & x \end{bmatrix} = x \cdot I + y \cdot J \qquad \overline{z} = x - iy \longmapsto A^T = \begin{bmatrix} x & y \\ -y & x \end{bmatrix}$$

$$(x+iy)(x-iy) = x^2 + y^2 \longmapsto \begin{bmatrix} x & -y \\ y & x \end{bmatrix} \begin{bmatrix} x & y \\ -y & x \end{bmatrix} = (x^2+y^2) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{aligned} (x+iy)(a+ib) &= xa - yb + i(xb + ya) \\ &\longmapsto \begin{bmatrix} x & -y \\ y & x \end{bmatrix} \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} xa - yb & -(xb+ya) \\ xb + ya & xa - yb \end{bmatrix} \end{aligned}$$

Dla $x^2 + y^2 \neq 0$

$$z^{-1} = \frac{1}{x+iy} = \frac{x-iy}{x^2+y^2} \longmapsto \begin{bmatrix} x & -y \\ y & x \end{bmatrix}^{-1} = \frac{1}{x^2+y^2} \begin{bmatrix} x & y \\ -y & x \end{bmatrix}$$

$$i^n = \cos\frac{n\pi}{2} + i \sin\frac{n\pi}{2} \longmapsto \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^n = \begin{bmatrix} \cos\frac{n\pi}{2} & -\sin\frac{n\pi}{2} \\ \sin\frac{n\pi}{2} & \cos\frac{n\pi}{2} \end{bmatrix}$$

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i \sin n\theta \longmapsto \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^n = \begin{bmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{bmatrix}$$

Dla $z = x + iy = |z|(\cos\theta + i\sin\theta)$

$$z^n = |z|^n(\cos n\theta + i \sin n\theta) \longmapsto \begin{bmatrix} x & -y \\ y & x \end{bmatrix}^n = |z|^n \begin{bmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{bmatrix}$$

**Przykład 1.21** (Grupy izomorficzne)**.** Zobrazujemy pojęcie izomorfizmu grup. Rozważmy grupę abelową
$$G_4 = \{1,\, i,\, -1,\, -i\} = \{1,\, i,\, i^2,\, i^3\}$$
pierwiastków zespolonych 4 stopnia z jedynki. Działaniem grupowym jest mnożenie liczb zespolonych. Zbiór
$$G_4' = \{I,\, J,\, -I,\, -J\} = \{I,\, J,\, J^2,\, J^3\}$$
jest grupą abelową z działaniem mnożenia macierzy. Zbiór reszt z dzielenia liczby całkowitej przez 4
$$\mathbb{Z}_4 = \{0,\, 1,\, 2,\, 3\}$$
jest grupą z działaniem „dodawania reszt“. Działania w tych grupach są opisane tabelkami:

| · | **1** | i | -1 | -i |
|---|---|---|---|---|
| **1** | **1** | i | -1 | -i |
| i | i | -1 | -i | **1** |
| -1 | -1 | -i | **1** | i |
| -i | -i | **1** | i | -1 |

| · | **I** | J | -I | -J |
|---|---|---|---|---|
| **I** | **I** | J | -I | -J |
| J | J | -I | -J | **I** |
| -I | -I | -J | **I** | J |
| -J | -J | **I** | J | -I |

| + | **0** | 1 | 2 | 3 |
|---|---|---|---|---|
| **0** | **0** | 1 | 2 | 3 |
| 1 | 1 | 2 | 3 | **0** |
| 2 | 2 | 3 | **0** | 1 |
| 3 | 3 | **0** | 1 | 2 |

Powyższe tabelki pokazują, że struktura algebraiczna wszystkich trzech grup jest identyczna. Jest ona taka sama jak dla grupy 4-elementowej $G = \{e, a, b, c\}$ z działaniem $\circ$ opisanym tabelką

| $\circ$ | **e** | a | b | c |
|---|---|---|---|---|
| **e** | **e** | a | b | c |
| a | a | b | c | **e** |
| b | b | c | **e** | a |
| c | c | **e** | a | b |

**Przykład 1.22** ($M_{2\times2}(\mathbb{Z}_2)$)**.** Zbiór $M_{2\times2}(\mathbb{Z}_2)$ składa się z $2^4$ elementów. Macierz $A \in M_{2\times2}(\mathbb{Z}_2)$ jest odwracalna, gdy $\det A \neq 0$. W ciele $\mathbb{Z}_2$ oznacza to, że $\det A = 1$. Mamy 6 macierzy spełniających ten warunek:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

Podzbiory $M_{2\times2}(\mathbb{F})$ dostarczają wielu przykładów grup.

**Przykład 1.23** (Grupa $\mathrm{GL}_2(\mathbb{F})$)**.** Rozważmy zbiór macierzy odwracalnych

$$\mathrm{GL}_2(\mathbb{F}) = \{A \in M_{2\times2}(\mathbb{F}) : \det A \neq 0\}.$$

Jeśli $A, B \in \mathrm{GL}_2(\mathbb{F})$, to

$$\det(AB) = \det A \cdot \det B \neq 0,$$

czyli $AB \in \mathrm{GL}_2(\mathbb{F})$. Ponadto, $I \in \mathrm{GL}_2(\mathbb{F})$ oraz $A^{-1} \in \mathrm{GL}_2(\mathbb{F})$ dla $A \in \mathrm{GL}_2(\mathbb{F})$, bo

$$\begin{aligned} 1 &= \det I \\ &= \det(AA^{-1}) \\ &= \det A \cdot \det A^{-1}, \end{aligned}$$

czyli $\det A^{-1} \neq 0$. Oznacza to, że $\mathrm{GL}_2(\mathbb{F})$ z działaniem mnożenia macierzy jest grupą. Nie jest to grupa abelowa, bo przykładowo

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

Definiujemy ponadto zbiór

$$\mathrm{SL}_2(\mathbb{F}) = \{A \in M_{2\times2}(\mathbb{F}) : \det A = 1\} \subset \mathrm{GL}_2(\mathbb{F}).$$

Jest on również grupą (nieabelową) z mnożeniem macierzy.

**Przykład 1.24** (Macierze symplektyczne)**.** Definiujemy zbiór macierzy *symplektycznych* jako

$$\mathrm{Symp}(2) := \{A \in M_{2\times2}(\mathbb{R}) : A^T J A = J\}, \qquad J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Sprawdzimy, że jest to grupa z mnożeniem macierzy. Można to zrobić sprawdzając warunki definiujące grupę. W tym celu należy sprawdzić, że

- jeśli $A, B \in \mathrm{Symp}(2)$, to $AB \in \mathrm{Symp}(2)$,
- $I \in \mathrm{Symp}(2)$,
- jeśli $A \in \mathrm{Symp}(2)$, to $A$ jest odwracalna oraz $A^{-1} \in \mathrm{Symp}(2)$.

Zrobimy to nieco inaczej. Bezpośredni rachunek pokazuje, że

$$A^T J A = \begin{bmatrix} 0 & -\det A \\ \det A & 0 \end{bmatrix}.$$

Wynika stąd, że $A \in \mathrm{Symp}(2)$ wtedy i tylko wtedy, gdy $\det(A) = 1$. Oznacza to, że

$$\mathrm{Symp}(2) = SL_2(\mathbb{R})$$

**Przykład 1.25** (Macierze górnie trójkątne)**.** Macierz $A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \in M_{2\times 2}(\mathbb{F})$ nazywamy *górnie trójkątną*. Zauważmy, że

$$\begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \begin{bmatrix} x & y \\ 0 & z \end{bmatrix} = \begin{bmatrix} ax & ay + bz \\ 0 & dz \end{bmatrix},$$

czyli iloczyn macierzy górnie trójkątnych jest macierzą górnie trójkątną. Macierz $A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ jest odwracalna wtedy i tylko wtedy, gdy $\det A = ad \neq 0$. Macierz odwrotna jest wtedy równa

$$A^{-1} = \begin{bmatrix} \frac{1}{a} & -\frac{b}{ad} \\ 0 & \frac{1}{d} \end{bmatrix},$$

czyli jest górnie trójkątna. Wynika stąd, że macierze górnie trójkątne o niezerowym wyznaczniku tworzą grupę z mnożeniem macierzy. Jest ona podgrupą grupy $\mathrm{GL}_2(\mathbb{F})$. Podobnie, macierze górnie trójkątne o wyznaczniku równym 1 (tzn. $d = 1/a$) tworzą podgrupę grupy $\mathrm{SL}_2(\mathbb{F})$. Ponadto macierze górnie trójkątne spełniające warunek $a = d = 1$ są również podgrupą $\mathrm{SL}_2(\mathbb{F})$.

**Przykład 1.26** (Macierze ortogonalne)**.** Definiujemy zbiór macierzy *ortogonalnych* przez

$$O(2) = \{A \in M_{2\times 2}(\mathbb{R}) : A^T A = I\}.$$

Zauważmy, że

$$\begin{aligned} 1 &= \det I \\ &= \det(A^T A) \\ &= \det(A^T) \cdot \det A \\ &= (\det A)^2, \end{aligned}$$

więc $\det A = \pm 1$. Oznacza to, że $A$ jest odwracalna, czyli istnieje taka macierz $A^{-1}$, że $A^{-1}A = AA^{-1} = I$. Mnożąc równość $A^T A = I$ z prawej strony przez $A^{-1}$, otrzymujemy, że $A^T = A^{-1}$.

Jeśli $A, B \in O(2)$, to

$$\begin{aligned} (AB)^T(AB) &= (B^T A^T)(AB) \\ &= B^T \underbrace{(A^T A)}_{=I} B \\ &= B^T B \\ &= I, \end{aligned}$$

czyli $AB \in O(2)$. Oczywiście $I \in O(2)$. Ponieważ $A^{-1} = A^T$ dla $A \in O(2)$ oraz

$$(A^T)^T A^T = AA^T = I,$$

czyli $A^{-1} \in O(2)$. Wynika stąd, że $O(2)$ jest grupą i podgrupą grupy $\mathrm{GL}_2(\mathbb{R})$. Nie jest ona abelowa. Przyglądniemy się jej bliżej. Niech

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in O(2).$$

Jeśli $\det A = 1$, to

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

czyli z równości $A^T = A^{-1}$ mamy

$$a = d, \quad b = -c.$$

Macierz $A$ ma więc postać

$$A = \begin{bmatrix} a & -c \\ c & a \end{bmatrix}$$

dla pewnych takich liczb $a, c \in \mathbb{R}$, że $a^2 + c^2 = \det A = 1$. Wynika stąd, że

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

dla pewnego $\theta \in \mathbb{R}$.

Jeśli $\det A = -1$, to

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

więc równość $A^T = A^{-1}$ oznacza, że

$$a = -d, \quad b = c.$$

Macierz $A$ ma więc postać

$$A = \begin{bmatrix} a & c \\ c & -a \end{bmatrix}$$

dla pewnych takich liczb $a, c \in \mathbb{R}$, że $a^2 + c^2 = -\det A = 1$. Wynika stąd, że

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

dla pewnego $\theta \in \mathbb{R}$.

Zbiór

$$SO(2) = \{A \in O(2) : \det A = 1\}$$

jest grupą z mnożeniem macierzy. Składa się ona z wszystkich macierzy postaci

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad \theta \in \mathbb{R}.$$

Sprawdzamy łatwo, że jest to grupa abelowa. Jest ona izomorficzna z grupą $(\mathbb{S}^1, \cdot)$ poprzez

$$f : \mathbb{S}^1 \ni \cos\theta + i\sin\theta \longmapsto \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \in SO(2).$$

U Macierze ortogonalne i symplektyczne możemy traktować jako szczególny przypadek ogólniejszej konstrukcji. Mianowicie, dla ustalonej macierzy $S \in M_{2\times 2}(\mathbb{R})$ rozważmy

$$G_S = \left\{A \in GL_2(\mathbb{R}) : A^T S A = S\right\}.$$

Jest to grupa z mnożeniem macierzy. Dla $S = I$ otrzymujemy $O(2)$, a dla $S = J$ grupę $\mathrm{Symp}(2)$.

**Przykład 1.27** (Macierze symetryczne)**.** Zbiór rzeczywistych macierzy symetrycznych

$$\mathrm{Sym}_2(\mathbb{R}) = \left\{A \in M_{2\times 2}(\mathbb{R}) : A = A^T\right\}$$

jest grupą z działaniem dodawania macierzy.

Dla macierzy $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{F})$ definiujemy jej *ślad* wzorem

$$\operatorname{tr} A = a + d.$$

Sprawdzamy łatwo, że

(i) $\operatorname{tr}(A + B) = \operatorname{tr} A + \operatorname{tr} B$,

(ii) $\operatorname{tr}(z \cdot A) = z \operatorname{tr} A$,

(iii) $\operatorname{tr}(AB) = \operatorname{tr}(BA)$,

(iv) $\operatorname{tr}(A^T) = \operatorname{tr} A$.

U Niech $A \in M_{2\times 2}(\mathbb{F})$. Mówimy, że macierz $B \in M_{2\times 2}(\mathbb{F})$ jest *podobna* do $A$, jeśli istnieje taka macierz odwracalna $S \in M_{2\times 2}(\mathbb{F})$, że $B = S^{-1}AS$. Wtedy

$$\begin{aligned} \det B &= \det(S^{-1}AS) \\ &= \underbrace{\det S^{-1}}_{=\frac{1}{\det S}} \cdot \det A \cdot \det S \\ &= \det A \end{aligned}$$

oraz

$$\begin{aligned} \operatorname{tr} B &= \operatorname{tr}(S^{-1}AS) \\ &= \operatorname{tr}((S^{-1}A)S) \\ &= \operatorname{tr}(S(S^{-1}A)) \\ &= \operatorname{tr}((SS^{-1})A) \\ &= \operatorname{tr} A. \end{aligned}$$

**Przykład 1.28** (Macierze o śladzie równym 0)**.** Definiujemy zbiór

$$T_0(\mathbb{F}) = \{A \in M_{2\times 2}(\mathbb{F}) : \operatorname{tr} A = 0\}.$$

Jest to grupa abelowa z działaniem dodawania macierzy.

**Przykład 1.29** (Macierze hamiltonowskie)**.** Rozważmy zbiór

$$H_2 = \{A \in M_{2\times 2}(\mathbb{R}) : (JA)^T = JA\}.$$

Ponieważ $J^T = -J$ więc $(JA)^T = JA$ wtedy i tylko wtedy, gdy

$$JA + A^T J = 0.$$

Takie macierze są nazywane *macierzami hamiltonowskimi*. Pełnią one ważną rolę w teorii równań różniczkowych i mechanice klasycznej. Uzasadnimy, że $H_2$ jest grupą z dodawaniem macierzy. Zamiast sprawdzać warunki z definicji grupy zauważmy, że

$$JA + A^T J = \begin{bmatrix} 0 & -\operatorname{tr} A \\ \operatorname{tr} A & 0 \end{bmatrix}.$$

Oznacza to, że $H_2 = T_0(\mathbb{R})$, czyli jest grupą.

**Przykład 1.30** (Macierze unitarne)**.** Dla macierzy zespolonej $A = \begin{bmatrix} z_1 & w_1 \\ z_2 & w_2 \end{bmatrix} \in M_{2\times 2}(\mathbb{C})$ definiujemy

$$A^* = \begin{bmatrix} \overline{z}_1 & \overline{z}_2 \\ \overline{w}_1 & \overline{w}_2 \end{bmatrix}$$

Sprawdzamy łatwo, że

- $A^* = (\overline{A})^T$,
- $\det A^* = \overline{\det A}$,
- $(A^*)^* = A, \quad (zA + B)^* = \overline{z}A^* + B^*$,
- $(AB)^* = B^*A^*$,
- $\det A^* = \overline{\det A}$.

Rozważmy zbiór *macierzy unitarnych*

$$U(2) := \{A \in M_{2\times 2}(\mathbb{C}) : A^*A = I\}.$$

Warunek $A^*A = I$ oznacza, że $A^* = A^{-1}$ oraz $|\det A| = 1$. $U(2)$ stanowi grupę z mnożeniem macierzy.

Jeśli $\det A = 1$, to

$$A^{-1} = \begin{bmatrix} w_2 & -w_1 \\ -z_2 & z_1 \end{bmatrix}.$$

Z równości $A^* = A^{-1}$ otrzymujemy, że

$$z_1 = \overline{w}_2, \quad z_2 = -\overline{w}_1.$$

Macierz $A$ ma więc w tym przypadku postać

$$A = \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix}$$

dla pewnych takich liczb zespolonych $z, w \in \mathbb{C}$, że

$$|z|^2 + |w|^2 = \det A = 1.$$

Grupa

$$\mathrm{SU}(2) = \{A \in U(2) : \det A = 1\}$$

składa się więc z macierzy postaci

$$\begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix}, \quad \text{gdzie} \quad |z|^2 + |w|^2 = 1.$$

**Przykład 1.31** (Macierze hermitowskie)**.** Zbiór zespolonych *macierzy hermitowskich*

$$\mathrm{Herm}_2(\mathbb{C}) = \{A \in M_{2\times 2}(\mathbb{C}) : A = A^*\}$$

jest grupą z działaniem dodawania macierzy. Przyglądniemy się jej bliżej. Niech $A = \begin{bmatrix} z_1 & w_1 \\ z_2 & w_2 \end{bmatrix} \in \mathrm{Herm}_2(\mathbb{C})$. Ponieważ

$$A^* = \begin{bmatrix} \overline{z}_1 & \overline{z}_2 \\ \overline{w}_1 & \overline{w}_2 \end{bmatrix},$$

więc

$$z_1 = \overline{z}_1, \quad w_2 = \overline{w}_2, \quad w_1 = \overline{z}_2.$$

Wynika stąd, że $A \in \mathrm{Herm}_2(\mathbb{C})$ jeśli $A$ ma postać

$$A = \begin{bmatrix} a & \overline{c} \\ c & d \end{bmatrix}, \quad \text{gdzie} \quad a, d \in \mathbb{R}, \quad c \in \mathbb{C}.$$

## 1.5 Pewne ważne macierze

Rozważymy pewne specjalne macierze, które będą pojawiały się w dalszym ciągu wykładu. Zaczniemy od pewnego ogólnego rezultatu zwanego twierdzeniem Cayleya–Hamiltona.

**Lemat 1.2** (Twierdzenie Cayleya–Hamiltona)**.** *Dla macierzy $A \in M_{2\times 2}(\mathbb{F})$ zachodzi równość*

$$A^2 - \operatorname{tr}(A) \cdot A + \det(A) \cdot I = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

*Dowód.* Zastosujemy brutalną siłę, czyli bezpośredni rachunek. Niech $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Mamy

$$\begin{aligned} A^2 - (a+d) \cdot A + (ad - bc) \cdot I &= (A - a \cdot I)(A - d \cdot I) - (bc) \cdot I \\ &= \begin{bmatrix} 0 & b \\ c & d-a \end{bmatrix} \begin{bmatrix} a-d & b \\ c & 0 \end{bmatrix} - \begin{bmatrix} bc & 0 \\ 0 & bc \end{bmatrix} \\ &= \begin{bmatrix} bc & 0 \\ c(a-d) + (d-a)c & bc \end{bmatrix} - \begin{bmatrix} bc & 0 \\ 0 & bc \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

□

U Twierdzenie Cayleya–Hamiltona pozwala wyznaczyć macierz odwrotną do macierzy nieosobliwej $A$. Mnożąc równość $A^2 - \operatorname{tr}(A)A + \det(A)I = 0$ przez $A^{-1}$ otrzymujemy, że

$$A^{-1} = \frac{1}{\det(A)}\big(\operatorname{tr}(A) \cdot I - A\big)$$

Ponadto $\operatorname{tr}(A^2) = (\operatorname{tr}(A))^2 - 2\det(A)$.

U Wzory Viète'a Dla macierzy $A \in M_{2\times 2}(\mathbb{C})$ rozważmy równanie kwadratowe

$$p_A(z) = z^2 - \operatorname{tr}(A)\, z + \det(A) = 0.$$

Ma ono dwa pierwiastki zespolone $\lambda_1, \lambda_2 \in \mathbb{C}$, czyli

$$\begin{aligned} z^2 - \operatorname{tr}(A)\, z + \det(A) &= (z - \lambda_1)(z - \lambda_2) \\ &= z^2 - (\lambda_1 + \lambda_2) z + \lambda_1 \lambda_2. \end{aligned}$$

Wynika stąd, że

$$\operatorname{tr}(A) = \lambda_1 + \lambda_2, \quad \det(A) = \lambda_1 \lambda_2.$$

Liczby $\lambda_1, \lambda_2$ będziemy nazywać *wartościami własnymi* macierzy $A$.
Z twierdzenia Cayleya–Hamiltona mamy

$$A^2 - (\lambda_1 + \lambda_2) \cdot A + (\lambda_1 \lambda_2) \cdot I = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

czyli

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = (A - \lambda_1 \cdot I)(A - \lambda_2 \cdot I) = (A - \lambda_2 \cdot I)(A - \lambda_1 \cdot I).$$

**Przykład 1.32.** Niech $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \in M_{2\times 2}(\mathbb{C})$. Korzystając z twierdzenia Cayleya–Hamiltona obliczymy $A^5 + A^3$. Rozważmy wielomian $w(z) = z^5 + z^3$. Dzieląc go przez wielomian charakterystyczny $p_A(z) = z^2 - 2z + 1 = (z-1)^2$ macierzy $A$, otrzymujemy, że

$$w(z) = (z^3 + 2z^2 + 4z + 6)p_A(z) + (8z - 6).$$

Z twierdzenia Cayleya–Hamiltona mamy

$$A^5 + A^3 = (A^3 + 2A^2 + 4A + 6I)\underbrace{(A - I)^2}_{=0} + (8A - 6I) = 8A - 6I,$$

czyli

$$A^5 - A^3 = 8\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - 6\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 8 \\ 0 & 2 \end{bmatrix}.$$

**Przykład 1.33.** Korzystając z twierdzenia Cayleya–Hamiltona obliczymy $A^{1000}$, gdzie $A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$. Wielomian charakterystyczny $p_A(z) = z^2 - 3z + 2$ ma dwa pierwiastki $\lambda_1 = 1$ i $\lambda_2 = 2$. Dzieląc wielomian $w(z) = z^{1000}$ przez $p_A$ otrzymujemy, że

$$w(z) = q(z)p_A(z) + r(z),$$

gdzie $r(z) = r_1 z + r_0$ jest resztą z dzielenia. Współczynniki $r_1$ i $r_0$ możemy wyznaczyć podstawiając $\lambda_1$ i $\lambda_2$ za $z$. Otrzymujemy, że

$$r_1 + r_0 = \lambda_1^{1000} = 1, \quad 2r_1 + r_0 = \lambda_2^{1000} = 2^{1000}.$$

Rozwiązaniem układu są

$$r_1 = 2^{1000} - 1, \quad r_0 = 2 - 2^{1000}.$$

Z twierdzenia Cayleya–Hamiltona wynika, że

$$\begin{aligned} A^{1000} &= r(A) \\ &= (2^{1000} - 1)\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} + (2 - 2^{1000})\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2^{1000} - 1 \\ 0 & 2^{1000} \end{bmatrix}. \end{aligned}$$

**Przykład 1.34** (Wartości własne macierzy symetrycznych i hermitowskich)**.** Jeśli $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix} \in M_{2\times 2}(\mathbb{R})$ jest macierzą symetryczną, to $A$ ma rzeczywiste wartości własne. Rzeczywiście, dla równania

$$z^2 - (a + d)z + ad - b^2 = 0$$

mamy

$$\begin{aligned} (a + d)^2 - 4(ad - b^2) &= a^2 + 2ad + d^2 - 4ad + 4b^2 \\ &= (a - d)^2 + 4b^2 \\ &\geqslant 0. \end{aligned}$$

Również, jeśli

$$A = \begin{bmatrix} a & \overline{c} \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{C}), \quad a, d \in \mathbb{R}, \quad c \in \mathbb{C}$$

jest macierzą hermitowską, to $A$ ma rzeczywiste wartości własne. Dla równania

$$0 = z^2 - (a + d)z + ad - c\overline{c} = z^2 - (a + d)z + ad - |c|^2$$

mamy

$$\begin{aligned} (a + d)^2 - 4(ad - |c|^2) &= a^2 + 2ad + d^2 - 4ad + 4|c|^2 \\ &= (a - d)^2 + 4|c|^2 \\ &\geqslant 0. \end{aligned}$$

U Twierdzenie Cayleya–Hamiltona pozwala łatwo obliczać potęgi macierzy $A$. Rozważmy macierz $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Wtedy $A^2 - 5A - 2I = 0$. Powiedzmy, że chcemy znaleźć $A^4$. Mamy

$$\begin{aligned} A^3 &= A\underbrace{(5A + 2I)}_{=A^2} \\ &= 5A^2 + 2A \\ &= 5(5A + 2I) + 2A \\ &= 27A + 10I \end{aligned}$$

oraz

$$\begin{aligned} A^4 &= A^3 A \\ &= (27A + 10I)A \\ &= 27A^2 + 10A \\ &= 27(5A + 2I) + 10A \\ &= 145A + 54I. \end{aligned}$$

Analogicznie pokazuje się, że $A^n = a \cdot A + b \cdot I$ dla pewnych $a, b \in \mathbb{F}$

**Przykład 1.35** (Macierz nilpotentna)**.** Powiemy, że macierz $A \in M_{2\times 2}(\mathbb{C})$ jest *nilpotentna*, jeśli $A^n = 0$ dla pewnego $n \in \mathbb{N}$. Pokażemy najpierw, że $A^2 = 0$. Możemy założyć, że $n \geqslant 2$, bo $A = 0$ dla $n = 1$. Zauważmy, że

$$0 = \det A^n = (\det A)^n,$$

czyli $\det A = 0$. Niech $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Wtedy

$$A^2 = \begin{bmatrix} a^2 + bc & ab + bd \\ ac + cd & bc + d^2 \end{bmatrix}.$$

Ponieważ, $\det A = ad - bc$, więc

$$A^2 = \begin{bmatrix} (a+d)a & (a+d)b \\ (a+d)c & (a+d)d \end{bmatrix} = \operatorname{tr}(A)A.$$

Stąd

$$0 = A^n = (\operatorname{tr}(A))^{n-1}A,$$

czyli $\operatorname{tr} A = a + d = 0$. Z twierdzenia Cayleya–Hamiltona wynika, że

$$A^2 = 0.$$

Warunek ten oznacza, że

$$a^2 + bc = b(a + d) = c(a + d) = bc + d^2 = 0.$$

Wiemy, że $a + d = 0$, czyli $a = -d$. Przyglądnijmy się równaniu $a^2 + bc = 0$. Mamy dwa przypadki ze względu na $c$:

- jeśli $c = 0$, to $a = 0$ i $A = \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix}$ dla $b \in \mathbb{C}$.
- jeśli $c \neq 0$, to $b = -\frac{a^2}{c}$ oraz

$$A = \begin{bmatrix} a & -\frac{a^2}{c} \\ c & -a \end{bmatrix} = \frac{1}{c}\begin{bmatrix} ac & -a^2 \\ c^2 & -ac \end{bmatrix}, \quad a, c \in \mathbb{C},\ c \neq 0.$$

**Przykład 1.36** (Inwolucja)**.** Powiemy, że $A \in M_{2\times 2}(\mathbb{C})$ jest *inwolucją*, jeśli $A^2 = I$. Wynika stąd, że $(\det A)^2 = 1$, czyli $\det A = \pm 1$. Ponadto, $A^{-1} = A$. Ze wzoru na $A^{-1}$ mamy, że

$$\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \det(A) \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Jeśli $\det A = 1$, to $a = d$ oraz $b = c = 0$. Ponadto, $1 = ad - bc = a^2$ czyli $a = \pm 1$. Stąd, $A = I$ lub $A = -I$.

Jeśli $\det A = -1$, to $a = -d$ i $-1 = ad - bc = -a^2 - bc$, czyli $a^2 + bc = 1$. Mamy dwa przypadki względem $b$:

- jeśli $b = 0$, to $a = \pm 1$ i $c$ jest dowolne; wtedy

$$A = \begin{bmatrix} -1 & 0 \\ c & 1 \end{bmatrix} \quad \text{lub} \quad A = \begin{bmatrix} 1 & 0 \\ c & -1 \end{bmatrix}.$$

- jeśli $b \neq 0$, to $c = \frac{1-a^2}{b}$ oraz

$$A = \begin{bmatrix} a & b \\ \frac{1-a^2}{b} & -a \end{bmatrix}, \quad a \in \mathbb{C}, \quad b \in \mathbb{C} \setminus \{0\}.$$

**Przykład 1.37** (Macierz idempotentna)**.** Macierz $A \in M_{2\times 2}(\mathbb{C})$ nazywamy *idempotentną*, jeśli $A^2 = A$. Ponieważ $(\det A)^2 = \det A$, więc $\det A = 1$ lub $\det A = 0$. Jeśli $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, to $A^2 = A$ oznacza, że

$$a^2 + bc = a, \quad b(a + d - 1) = c(a + d - 1) = (a - d)(a + d - 1) = 0.$$

Jeśli $a + d - 1 \neq 0$, to $b = c = 0$, $a = d$ i $a^2 = a$. Wtedy $A = 0$ lub $A = I$.

Jeśli $a + d - 1 = 0$, to warunki redukują się do $a^2 + bc = a$.

Mamy dwa przypadki względem $b$:

- jeśli $b \neq 0$, to $c = \frac{a^2-a}{b}$ oraz

$$A = \begin{bmatrix} a & b \\ \frac{a-a^2}{b} & 1 - a \end{bmatrix}, \quad a \in \mathbb{C},\ b \in \mathbb{C} \setminus \{0\}.$$

- jeśli $b = 0$, to albo $a = 0$ albo $a = 1$ i wtedy

$$A = \begin{bmatrix} 0 & 0 \\ c & 1 \end{bmatrix} \quad \text{lub} \quad \begin{bmatrix} 1 & 0 \\ c & 0 \end{bmatrix}, \quad c \in \mathbb{C}.$$

**Przykład 1.38** (Macierz Grama)**.** Dla macierzy rzeczywistej $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{R})$ jej *macierz Grama* jest zdefiniowana jako $A^T A$. Stąd

$$A^T A = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{bmatrix}.$$

Zauważmy, że

- $A^T A$ jest symetryczna,
- $A^T A$ ma rzeczywiste wartości własne,
- $\operatorname{tr} A^T A = a^2 + b^2 + c^2 + d^2$,
- $\operatorname{tr} A^T A = 0$ wtedy i tylko wtedy, gdy $A = 0$.

Przyglądnijmy się wyznacznikowi macierzy $A^T A$. Mamy

$$\begin{aligned}\det A^T A &= \det A^T \cdot \det A \\ &= (\det A)^2 \geqslant 0.\end{aligned}$$

**Przykład 1.39** (Rzeczywista macierz normalna)**.** Rozważmy taką macierz rzeczywistą $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{R})$, że

$$A^T A = AA^T.$$

Nazywamy ją macierzą *normalną*. Oznacza to, że

$$\begin{bmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{bmatrix} = \begin{bmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{bmatrix},$$

czyli

$$c^2 = b^2, \quad ab+cd = ac+bd.$$

Mamy dwie możliwości:

- jeśli $c = b$, to $A$ jest symetryczna,
- jeśli $b = -c$, to $c(d-a) = c(a-d)$; wtedy albo $c = 0$ i $A$ jest symetryczna albo $c \neq 0$ i wtedy $a = d$ oraz $A$ ma postać

$$A = \begin{bmatrix} a & -c \\ c & a \end{bmatrix}.$$

# Rozdział 2

# Przestrzeń wektorowa

| |
|---|
| SŁOWA KLUCZOWE ROZDZIAŁU<br>przestrzenie wektorowe $\mathbb{R}^n$ i $\mathbb{C}^n$ ◊ przestrzeń wektorowa $\mathbb{F}^n$ ◊ kombinacja liniowa ◊ liniowa niezależność ◊ baza ◊ baza standardowa $\mathbb{F}^n$ ◊ iloczyn skalarny w $\mathbb{R}^n$ ◊ norma w $\mathbb{R}^n$ ◊ nierówność Cauchy'ego–Schwarza ◊ wektory ortogonalne ◊ proste i płaszczyzny w $\mathbb{R}^3$ ◊ iloczyn hermitowski |

## 2.1 Płaszczyzna $\mathbb{R}^2$ jako przestrzeń wektorowa

Punkt na płaszczyźnie $\mathbb{R}^2$ utożsamiamy z jego współrzędnymi kartezjańskimi $(x_1, x_2)$, czyli liczbą zespoloną $z = x_1 + ix_2$. Wprowadzimy teraz w zbiorze $\mathbb{R}^2$ pewną nową strukturę algebraiczną, strukturę przestrzeni wektorowej. Geometrycznie punkt $P = (x_1, x_2)$ na płaszczyźnie będziemy interpretowali jako wektor będący strzałką o początku w punkcie $0 = (0, 0)$ i o końcu w punkcie $P$. W tej interpretacji będzie dla nas wygodne zastosowanie nieco innej konwencji zapisu współrzędnych punktu. Zamiast pisać $(x_1, x_2)$ będziemy stosować zapis kolumnowy $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, czyli pierwsza współrzędna nad drugą. Dla oszczędności papieru wprowadzamy oznaczenie

$$[x_1, x_2]^T := \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Napis $(x_1, x_2)$ oznacza, że mamy na myśli liczbę zespoloną, a napis sugeruje $[x_1, x_2]^T$, że myślimy o wektorze na płaszczyźnie.

**Definicja 2.1** (Płaszczyzna $\mathbb{R}^2$).
Płaszczyzną rzeczywistą nazywamy zbiór

$$\mathbb{R}^2 := \Big\{[x_1, x_2]^T : x_1, x_2 \in \mathbb{R}\Big\}.$$

Jego elementy będziemy nazywać *wektorami.*

Zbiór $\mathbb{R}^2$ możemy łatwo wyposażyć w dodatkową strukturę algebraiczną. Wektory będziemy dodawać i mnożyć je przez liczbę rzeczywistą.

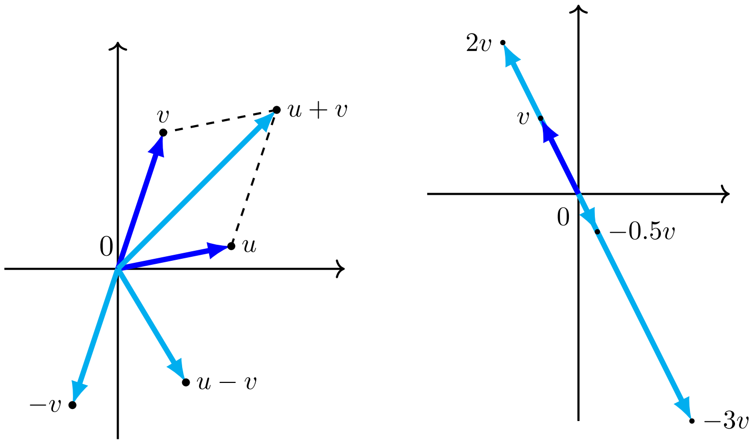
**Definicja 2.2** (Dodawanie wektorów i mnożenie ich przez skalar).
Niech $x = [x_1, x_2]^T, y = [y_1, y_2]^T \in \mathbb{R}^2$. Definiujemy ich *sumę* przez

$$x + y = [x_1, x_2]^T + [y_1, y_2]^T = [x_1 + y_1, x_2 + y_2]^T.$$

W zapisie kolumnowym mamy

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}$$

Powyższy wzór określa *dodawanie* $+ : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ punktów (wektorów) na płaszczyźnie. Przypisuje ono dwóm punktom płaszczyzny pewien punkt na płaszczyźnie.




Rysunek 2.1: Dodawanie wektorów i mnożenie ich przez liczbę rzeczywistą. Suma wektorów $u+v$ jest przekątną równoległoboku $R$ rozpiętego na wektorach $u$ i $v$. Różnicę wektorów $u-v$ otrzymujemy dodając do wektora $u$ wektor $v$ pomnożony przez liczbę $-1$ tzn. $u-v = u+(-1)v$. Długość wektora $u-v$ jest równa długości drugiej przekątnej równoległoboku $R$. Mnożąc niezerowy wektor $v$ przez wszystkie liczby rzeczywiste, otrzymujemy prostą na płaszczyźnie. Dokładniej, zbiór $\{t \cdot v : t \in \mathbb{R}\}$ jest prostą przechodzącą przez 0 i równoległą do wektora $v$.

Niech $a \in \mathbb{R}$ i $x = [x_1, x_2]^T \in \mathbb{R}^2$. Definiujemy *mnożenie wektora przez liczbę* rzeczywistą $a$ wzorem

$$a \cdot [x_1, x_2]^T = [a \cdot x_1, a \cdot x_2]^T.$$

Równoważnie,

$$a \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a \cdot x_1 \\ a \cdot x_2 \end{bmatrix}$$

Zdefiniowane powyżej mnożenie definiuje odwzorowanie $\cdot : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$. Przypisuje ono liczbie rzeczywistej i punktowi na płaszczyźnie pewien punkt na płaszczyźnie.

## 2.2 Definicja przestrzeni wektorowej

Naturalne jest pytanie: dlaczego mielibyśmy się ograniczać wyłącznie do płaszczyzny $\mathbb{R}^2$? Przecież podobną konstrukcję możemy powtórzyć w przestrzeni $\mathbb{R}^3$, gdzie położenie punktu jest opisane trzema współrzędnymi w miejsce dwóch. A może rozważyć pięć współrzędnych? Lepiej zróbmy to od razu ogólnie dla dowolnej liczby naturalnej $n$.

**Definicja 2.3** (Przestrzeń wektorowa $\mathbb{R}^n$)**.**
Niech $n \in \mathbb{N}$. Definiujemy zbiór

$$\mathbb{R}^n := \Big\{[x_1, \dots, x_n]^T : x_1, \dots, x_n \in \mathbb{R}\Big\},$$

gdzie

$$[x_1, \dots, x_n]^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Analogicznie do płaszczyzny $\mathbb{R}^2$ definiujemy działanie dodawania wektorów $+ : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ oraz ich mnożenie przez liczbę rzeczywistą $\cdot : \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ wzorami

$$x + y = [x_1, \dots, x_n]^T + [y_1, \dots, y_n]^T = [x_1 + y_1, \dots, x_n + y_n]^T$$

$$t \cdot [x_1, \dots, x_n]^T = [t \cdot x_1, \dots, t \cdot x_n]^T, \quad t \in \mathbb{R}$$

*Początek układu współrzędnych*, czyli *wektor zerowy* $[0, \ldots, 0]^T$ będziemy oznaczać przez 0. Zazwyczaj nie prowadzi to do nieporozumień – z kontekstu wynika czy 0 oznacza zero jako liczbę rzeczywistą czy raczej wektor zerowy w $\mathbb{R}^n$.

$\mathbb{R}^n$ – AKSJOMATY PRZESTRZENI WEKTOROWEJ

**Lemat 2.1.** *Dla dowolnych wektorów $x, y, z \in \mathbb{R}^n$ i liczb rzeczywistych $a, b \in \mathbb{R}$ mamy*

*(1)* $(x + y) + z = x + (y + z)$ *(łączność dodawania wektorów)*

*(2)* $x + y = y + x$ *(przemienność dodawania wektorów)*

*(3)* $x + 0 = x$

*(4) istnieje (jedyny) taki wektor $x'$, że $x + x' = 0$* $(x' = (-1) \cdot x)$

*(5)* $a \cdot (x + y) = a \cdot x + a \cdot y$ *(rozdzielność)*

*(6)* $(a + b) \cdot x = a \cdot x + b \cdot x$ *(rozdzielność)*

*(7)* $(a \cdot b) \cdot x = a \cdot (b \cdot x)$ *(łączność)*

*(8)* $1 \cdot x = x$.

*Dowód.* Dowód jest bardzo prosty. Jak się dobrze przyglądniemy, to zauważymy, że wszystkie własności (1)–(8) wynikają bezpośrednio z własności działań w zbiorze liczb rzeczywistych. Wystarczy je zastosować do każdej współrzędnej z osobna. □

Początki teorii przestrzeni wektorowych związane są z Williamem Rowanem Hamiltonem (1805–1865) oraz Hermannem Guntherem Grassmannem (1809–1877)

Mamy więc do czynienia z pewnym zbiorem $V (= \mathbb{R}^n)$ i dwoma działaniami

$$+ : V \times V \longrightarrow V, \quad \cdot : \mathbb{R} \times V \longrightarrow V$$

spełniającymi warunki (1)–(8).

**Definicja 2.4** (Przestrzeń wektorowa nad ciałem $\mathbb{R}$)**.**

Rozważmy trójkę $(V, +, \cdot)$ z opisanymi powyżej działaniami

$$+ : V \times V \longrightarrow V, \quad \cdot : \mathbb{R} \times V \longrightarrow V.$$

Powiemy, że $(V, +, \cdot)$ jest *przestrzenią wektorową nad ciałem* $\mathbb{R}$, (bo mnożymy elementy zbiory $V$ przez liczby rzeczywiste) jeśli zachodzą powyższe warunki (1)–(8).

Z powyższych rozważań wynika, że $(\mathbb{R}^n, +, \cdot)$ jest przykładem przestrzeni wektorowej nad ciałem $\mathbb{R}$. Możecie zapytać po co nam jakieś przestrzenie wektorowe. Przecież mamy już wektory w $\mathbb{R}^n$, które dodajemy i mnożymy przez liczbę rzeczywistą. Czy to nie wystarczy? Otóż nie. Popatrzcie na funkcję $x^2 + \sin(x)$ (przepraszam ortodoksyjnych matematyków za ten zapis funkcji bez podania dziedziny i przeciwdziedziny). Przecież to też jest suma pewnych obiektów (w tym wypadku funkcji), tylko teraz dodajemy funkcje zamiast liczb lub wektorów z $\mathbb{R}^n$. Funkcje, podobnie jak wektory możemy mnożyć przez liczby. Pewnie zetknęliście się z funkcjami typu $3x^4$ lub $-2\cos(x)$.

**Przykład 2.1** (Przestrzeń $F(\mathbb{R}, \mathbb{R})$ funkcji $f : \mathbb{R} \to \mathbb{R}$)**.** Niech $F(\mathbb{R}, \mathbb{R})$ będzie zbiorem wszystkich funkcji $f : \mathbb{R} \longrightarrow \mathbb{R}$. W zbiorze $F(\mathbb{R}, \mathbb{R})$ mamy naturalne dodawanie $+ : F(\mathbb{R}, \mathbb{R}) \times F(\mathbb{R}, \mathbb{R}) \longrightarrow F(\mathbb{R}, \mathbb{R})$ dane wzorem

$$(f + g)(x) = f(x) + g(x),$$

oraz mnożenie przez skalar $\cdot : \mathbb{R} \times F(\mathbb{R}, \mathbb{R}) \longrightarrow F(\mathbb{R}, \mathbb{R})$ zdefiniowane formułą

$$(a \cdot f)(x) = a \cdot f(x).$$

Trójka $(F(\mathbb{R}, \mathbb{R}), +, \cdot)$ spełnia warunki (1)–(8), więc jest przestrzenią wektorową nad ciałem liczb rzeczywistych.

**Przykład 2.2** (Przestrzeń wielomianów)**.** Niech $P$ będzie zbiorem wszystkich wielomianów o współczynnikach rzeczywistych. Z naturalnymi działaniami dodawania $+$ i mnożenia przez liczbą $\cdot$, trójka $(P, +, \cdot)$ jest przestrzenią wektorową nad ciałem liczb rzeczywistych.

U Jak pewnie zauważyliście, przez cały czas podkreślam, że rozważane dotąd przestrzenie wektorowe $V$ są przestrzeniami wektorowymi nad ciałem $\mathbb{R}$. Może moglibyśmy rozważać przestrzenie wektorowe nad innymi ciałami? Poniższy przykład pokazuje, że możemy to robić. Co więcej jest to bardzo naturalne, a naturalne konstrukcje bardzo w matematyce lubimy.

**Przykład 2.3** (Przestrzeń wektorowa $\mathbb{C}^n$)**.** Przez analogię do zbioru $\mathbb{R}^n$ możemy zdefiniować zbiór $\mathbb{C}^n$ jako

$$\mathbb{C}^n := \Big\{[z_1, \ldots, z_n]^T : z_1, \ldots, z_n \in \mathbb{C}\Big\},$$

gdzie

$$[z_1, \ldots, z_n]^T = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}.$$

W zbiorze $\mathbb{C}^n$ mamy naturalnie zdefiniowane działanie dodawania

$$+ : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}^n$$

oraz ich mnożenie przez liczbę zespoloną

$$\cdot : \mathbb{C} \times \mathbb{C}^n \longrightarrow \mathbb{C}^n.$$

Są one dane wzorami

$$z + w = [z_1, \ldots, z_n]^T + [w_1, \ldots, w_n]^T = [z_1 + w_1, \ldots, z_n + w_n]^T,$$

$$t \cdot [z_1, \ldots, z_n]^T = [t \cdot z_1, \ldots, t \cdot z_n]^T, \quad t \in \mathbb{C}.$$

Łatwo sprawdzamy, że tak określona trójka $(\mathbb{C}^n, +, \cdot)$ spełnia warunki (1)–(8). Jedyna różnica polega na tym, że wektory (elementy zbioru $\mathbb{C}^n$) mnożymy teraz przez liczby zespolone, a nie liczby rzeczywiste. Możemy, więc stwierdzić, że $(\mathbb{C}^n, +, \cdot)$ jest przestrzenią wektorową nad ciałem liczb zespolonych. Możecie powiedzieć, że przecież to nie koniec. Taką samą konstrukcje możemy przeprowadzić rozważając zbiór $\mathbb{Q}^n$ i jego elementy mnożyć przez liczby wymierne. Otrzymamy przestrzeń wektorową nad ciałem liczb wymiernych $\mathbb{Q}$. Prowadzi nas to do ogólnej definicji przestrzeni wektorowej nad dowolnym ciałem $\mathbb{F}$.

**Definicja 2.5** (Przestrzeń wektorowa nad ciałem $\mathbb{F}$)**.** Niech $V$ będzie zbiorem i niech $\mathbb{F}$ będzie ciałem. Załóżmy, że określone są działania: $+ : V \times V \longrightarrow V$ (dodawanie wektorów) oraz $\cdot : \mathbb{F} \times V \longrightarrow V$ (mnożenie przez skalar). Powiemy, że trójka $(V, +, \cdot)$ jest *przestrzenią wektorową nad ciałem* $\mathbb{F}$, jeśli dla dowolnych wektorów $x, y, z \in V$ oraz skalarów $a, b \in \mathbb{F}$ zachodzą warunki:

(1) $(x + y) + z = x + (y + z)$

(2) $x + y = y + x$

(3) istnieje taki wektor $0 \in V$, że $x + 0 = x$

(4) istnieje taki wektor $x'$, że $x + x' = 0$

(5) $a \cdot (x + y) = a \cdot x + a \cdot y$

(6) $(a + b) \cdot x = a \cdot x + b \cdot x$

(7) $(a \cdot b) \cdot x = a \cdot (b \cdot x)$

(8) $1 \cdot x = x$.

Zauważmy, że pierwsze cztery warunki dotyczą wyłącznie dodawania wektorów i oznaczają, że $(V, +)$ jest grupą abelową. W dalszym ciągu napis $V \in \mathrm{Vekt}_{\mathbb{F}}$ będzie oznaczał, że $V$ jest przestrzenią wektorową nad ciałem $\mathbb{F}$.

U Wektor zerowy $0 \in V$ jest wyznaczony jednoznacznie. Jeśli $0_1, 0_2 \in V$ spełniają warunek (3), to $0_1 = 0_1 + 0_2 = 0_2$.

U Rozważmy przestrzeń wektorową $V \in \mathrm{Vekt}_{\mathbb{F}}$ nad ciałem $\mathbb{F}$. Uzasadnimy, że dla $v \in V$ oraz $0 \in \mathbb{F}$ (elementu neutralnego dodawania w ciele $\mathbb{F}$) zachodzi równość

$$0 \cdot v = 0 \in V.$$

Możecie powiedzieć, że to przesada, aby uzasadnić takie rzeczy. Przecież $0 \in \mathbb{F}$ pomnożone przez „cokolwiek" daje 0. Tak jest, gdy „cokolwiek" jest elementem ciała $\mathbb{F}$. W naszym przypadku $0 \cdot v$ oznacza, że mnożymy $0 \in \mathbb{F}$ przez wektor $v \in V$, natomiast 0 po prawej stronie jest wektorem, a nie elementem ciała $\mathbb{F}$.

Z własności (6) i definicji 0 jako elementu neutralnego dla dodawania w ciele $\mathbb{F}$ wynika, że

$$\begin{aligned} 0 \cdot v &= (0 + 0) \cdot v \\ &= 0 \cdot v + 0 \cdot v, \end{aligned}$$

więc $0 \cdot v = 0$.

U Niech $x \in V$. Wektor $x'$ spełniający warunek $x + x' = 0$ jest wyznaczony jednoznacznie. Rzeczywiście, jeśli $x + x' = 0 = x + x''$, to

$$\begin{aligned} x' &= x' + 0 \\ &= x' + (x + x'') \\ &= (x' + x) + x'' \\ &= 0 + x'' \\ &= x'' \end{aligned}$$

Wektor $x'$ nazywamy wektorem przeciwnym do $x$ i oznaczamy przez $-x$. Uzasadnimy, że $-x = (-1) \cdot x$.

$$\begin{aligned} x + (-1) \cdot x &= (1 + (-1)) \cdot x \\ &= 0 \cdot x \\ &= 0, \end{aligned}$$

czyli $x' = -x = (-1) \cdot x$.
Analogicznie pokazujemy, że
$a \cdot 0 = 0$ oraz $-a \cdot v = (-a) \cdot v$

**Przykład 2.4** (Przestrzeń wektorowa $(\mathbb{F}^n, +, \cdot)$)**.** Jeśli $\mathbb{F}$ jest dowolnym ciałem, to naturalnie określona trójka $(\mathbb{F}^n, +, \cdot)$ jest przestrzenią wektorową nad ciałem $\mathbb{F}$. W szczególności, $(\mathbb{F}, +, \cdot)$ jest przestrzenią wektorową nad ciałem $\mathbb{F}$.

**Przykład 2.5** (O co chodzi z tymi ciałami?)**.** Rozważmy ponownie przestrzeń wektorową $(\mathbb{C}^n, +, \cdot)$. Mnożenie $\cdot$ jest mnożeniem wektora przez liczbę zespoloną. Oznaczmy je chwilowo przez $\cdot_{\mathbb{C}}$. Zauważmy, że w zbiorze $\mathbb{C}^n$ mamy też dobrze określone mnożenie $\cdot_{\mathbb{R}}$ przez liczbę rzeczywistą. Trójki $(\mathbb{C}^n, +, \cdot_{\mathbb{C}})$ oraz $(\mathbb{C}^n, +, \cdot_{\mathbb{R}})$ są różnymi obiektami matematycznymi. $(\mathbb{C}^n, +, \cdot_{\mathbb{C}})$ jest przestrzenią wektorową nad ciałem $\mathbb{C}$, a $(\mathbb{C}^n, +, \cdot_{\mathbb{R}})$ jest przestrzenią wektorową nad ciałem $\mathbb{R}$. Co za różnica? Otóż zasadnicza. Pierwsze dwa elementy trójek są identyczne tzn. zbiór $\mathbb{C}^n$ i określone w nim działanie dodawania wektorów $+$. Aby dostrzec różnicę, rozważmy wektor $v = [1 + 2i, 1 + 2i]^T \in \mathbb{C}^2$. Zauważmy, że w przestrzeni $(\mathbb{C}^n, +, \cdot_{\mathbb{C}})$ nad ciałem $\mathbb{C}$ zachodzi równość

$$\begin{bmatrix} 1 + 2i \\ 1 + 2i \end{bmatrix} = (1 + 2i) \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

W przestrzeni $(\mathbb{C}^n, +, \cdot_\mathbb{R})$ nad ciałem $\mathbb{R}$ taka równość nie zachodzi. Co więcej, zapis

$$(1+2i) \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

nie ma sensu w tej przestrzeni, bo dopuszczamy tylko mnożenie wektorów z $\mathbb{C}^2$ przez liczby rzeczywiste.

**Przykład 2.6.** Rozważmy przestrzeń wektorową

$$\mathbb{Z}_3^5 = \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3$$

oraz wektory

$$u = [2,2,0,1,2]^T, \quad v = [1,2,2,2,1]^T.$$

Wtedy

$$u + v = [0,1,2,0,0]^T, \quad 2 \cdot u = [1,1,0,2,1]^T.$$

**U** Uważny czytelnik zauważył z pewnością, że działania w przestrzeni funkcyjnej $F(\mathbb{R},\mathbb{R})$ są zdefiniowane dzięki temu, że potrafimy dodawać wartości dwóch funkcji z $F(\mathbb{R},\mathbb{R})$ oraz mnożyć wartości funkcji z $F(\mathbb{R},\mathbb{R})$. Jest to możliwe, bo przeciwdziedzina funkcji z $F(\mathbb{R},\mathbb{R})$, czyli $\mathbb{R}$ jest ciałem. Nie jest natomiast ważne, że ich dziedzina $\mathbb{R}$ ma strukturę ciała. Możemy więc rozważać bardziej ogólne funkcyjne przestrzenie wektorowe. Dokładniej, jeśli $X$ jest dowolnym niepustym zbiorem i $\mathbb{F}$ jest ciałem, to zbiór

$$F(X, \mathbb{F})$$

wszystkich funkcji $f : X \longrightarrow \mathbb{F}$ jest przestrzenią wektorową z działaniami

$$(f+g)(x) = f(x) + g(x), \quad (a \cdot f)(x) = a \cdot f(x), \quad a \in \mathbb{F}.$$

Przykładowo, $F([-1,1],\mathbb{R})$ jest rzeczywistą przestrzenią wektorową. Zwróćcie uwagę, że powyżej zdefiniowane działania, nie są działaniami wewnętrznymi w zbiorze $F(\mathbb{R},[-1,1])$, bo $\sin, \cos \in F(\mathbb{R},[-1,1])$, ale $\sin + \cos \notin F(\mathbb{R},[-1,1])$ oraz $2\sin \notin F(\mathbb{R},[-1,1])$.

Niech $\mathbb{F}$ będzie ciałem. W przestrzeni wektorowej $(\mathbb{F}^n, +, \cdot)$ wyróżniamy $n$ wektorów danych przez

$$e_1 = [1,0,\ldots,0]^T,\ e_2 = [0,1,\ldots,0]^T,\ \ldots,\ e_n = [0,\ldots,0,1]^T$$

Zauważmy, że każdy wektor $v = [v_1, \ldots, v_n]^T \in \mathbb{F}^n$ możemy zapisać jako

$$\begin{aligned}
[v_1, \ldots, v_n]^T &= [v_1, \ldots, 0]^T + \ldots + [0, \ldots, v_n]^T \\
&= v_1 \cdot [1, \ldots, 0]^T + \ldots + v_n \cdot [0, \ldots, 1]^T \\
&= v_1 \cdot e_1 + \ldots + v_n \cdot e_n.
\end{aligned}$$

Oznacza to, że zachodzi równość

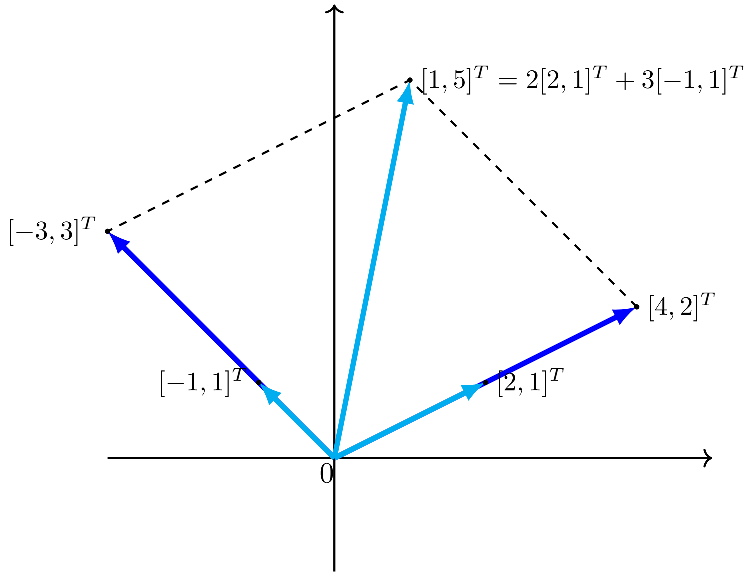
$$v = v_1 \cdot e_1 + \ldots + v_n \cdot e_n, \quad v_i \in \mathbb{F}.$$

Mówimy wtedy, że wektor $v$ jest *kombinacją liniową wektorów* $e_1, \ldots, e_n$. Zdefiniujemy to pojęcie bardziej ogólnie, co przyda nam się w przyszłości.

**Definicja 2.6** (Kombinacja liniowa wektorów)**.**
Niech $(V, +. \cdot)$ będzie przestrzenią wektorową nad ciałem $\mathbb{F}$. Dla wektorów $v_1, \ldots, v_k \in V$ oraz skalarów $a_1, \ldots, a_k \in \mathbb{F}$, wektor

$$a_1 \cdot v_1 + \ldots + a_k \cdot v_k \in V$$

nazywamy *kombinacją liniową wektorów* $v_1, \ldots, v_k \in V$.




Rysunek 2.2: Przykładowa kombinacja liniowa wektorów $[2,1]^T$ i $[-1,1]^T$.

**Przykład 2.7.** Niech $(V,+,\cdot)$ będzie przestrzenią wektorową nad ciałem $\mathbb{F}$. Ustalmy wektory $v_1,\ldots,v_k \in V$ i rozważmy zbiór

$$\operatorname{span}\{v_1,\ldots,v_k\} = \{a_1\cdot v_1 + \ldots + a_k\cdot v_k : a_1,\ldots,a_k \in \mathbb{F}\}$$

wszystkich możliwych kombinacji liniowych wektorów $v_1,\ldots,v_k$. Działania $+$ oraz $\cdot$ są działaniami wewnętrznymi w zbiorze $\operatorname{span}\{v_1,\ldots,v_k\}$ oraz $(\operatorname{span}\{v_1,\ldots,v_k\},+,\cdot)$ jest przestrzenią wektorową nad ciałem $\mathbb{F}$.

Załóżmy, że $V=\mathbb{R}^2$. Mamy

(a) $\operatorname{span}\{0\} = \{0\}$,

(b) jeśli $0 \neq v \in \mathbb{R}^2$, to $\operatorname{span}\{v\} = \{a\cdot v : a \in \mathbb{R}\}$ jest prostą przechodzącą przez 0 i równoległą do wektora $v$.

**Lemat 2.2.** *Rozważmy przestrzeń wektorową* $(\mathbb{F}^n,+,\cdot)$. *Wtedy każdy wektor* $v\in\mathbb{F}^n$ *można jednoznacznie zapisać w postaci kombinacji liniowej*

$$v = v_1\cdot e_1 + \ldots + v_n\cdot e_n, \quad v_i \in \mathbb{F}.$$

*Dowód.* Wiemy już, że dla dowolnego wektora $v\in\mathbb{F}^n$ istnieją takie skalary $v_1,\ldots,v_n\in\mathbb{F}$, że $v = v_1\cdot e_1 + \ldots + v_n\cdot e_n$. Pokażemy, że takie przedstawienie wektora $v$ w postaci kombinacji liniowej wektorów $e_1,\ldots,e_n$ jest wyznaczone jednoznacznie. Przypuśćmy, że

$$v = v_1\cdot e_1 + \ldots + v_n\cdot e_n = x_1\cdot e_1 + \ldots + x_n\cdot e_n.$$

Wtedy

$$\begin{aligned} 0 = [0,\ldots,0]^T &= v + (-1)\cdot v \\ &= (v_1 - x_1)\cdot e_1 + \ldots + (v_n - x_n)\cdot e_n \\ &= [v_1 - x_1,\ldots,v_n - x_n]^T, \end{aligned}$$

czyli $v_1 = x_1,\ldots,v_n = x_n$. □

**Definicja 2.7** (Baza standardowa $\mathbb{F}^n$)**.**

Wektory $e_1,\ldots,e_n$ nazywamy *bazą standardową* przestrzeni $\mathbb{F}^n$.

U Zauważmy, że jednoznaczność przedstawienia wektora $v$ jako kombinacji liniowej wektorów $e_1, \ldots, e_n$ wynika z prawdziwości następującej implikacji:

$$x_1 \cdot e_1 + \ldots + x_n \cdot e_n = 0 \quad \Rightarrow \quad x_1 = \ldots = x_n = 0.$$

Jest to więc ważna własność wektorów $e_1, \ldots, e_n$. Warto więc ją jakoś nazwać i zbadać.

**Definicja 2.8** (Liniowa niezależność)**.**

Niech $V$ będzie przestrzenią wektorową nad ciałem $\mathbb{F}$. Wektory $v_1, \ldots, v_k \in V$ nazywamy *liniowo niezależnymi*, jeśli dla $x_1, \ldots, x_k \in \mathbb{F}$ zachodzi implikacja

$$x_1 \cdot v_1 + \ldots + x_k \cdot v_k = 0 \quad \Rightarrow \quad x_1 = \ldots = x_k = 0.$$

**Przykład 2.8.** Niech $v, w \in \mathbb{R}^2$. Przypuśćmy, że wektory $v, w$ są liniowo zależne. Wtedy istnieją takie liczby rzeczywiste $a, b \in \mathbb{R}$, że $a \neq 0$ lub $b \neq 0$ oraz $a \cdot v + b \cdot w = 0$. Jeśli $a \neq 0$, to $v = -\frac{b}{a} \cdot w$, czyli wektor $v$ leży na prostej $\{tw : t \in \mathbb{R}\}$.

**Przykład 2.9.** Wektory

$$v_1 = [1, 1, 0]^T, \quad v_2 = [1, 0, 1]^T$$

są liniowo niezależne w $\mathbb{F}$. Przypuśćmy bowiem, że $x_1 \cdot v_1 + x_2 \cdot v_2 = 0$ dla pewnych $x_1, x_2 \in \mathbb{F}$. Ponieważ $x_1 \cdot v_1 + x_2 \cdot v_2 = [x_1 + x_2, x_1, x_2]^T$, więc

$$x_1 + x_2 = 0, \quad x_1 = 0, \quad x_2 = 0,$$

czyli $v_1$ i $v_2$ są liniowo niezależne.

**Definicja 2.9** (Baza przestrzeni wektorowej)**.**

Niech $V$ będzie przestrzenią wektorową nad ciałem $\mathbb{F}$. Wektory $v_1, \ldots, v_k \in v$ nazywamy *bazą przestrzeni wektorowej* $V$, jeśli dla dowolnego wektora $v \in V$ istnieją jednoznacznie wyznaczone takie skalary $x_1, \ldots, x_k \in \mathbb{F}$, że

$$v = x_1 \cdot v_1 + \ldots + x_k \cdot v_k.$$

**Przykład 2.10.** Podmieńmy wektor $e_3 = [0, 0, 1]^T$ w bazie dla $\mathbb{F}^3$ przez wektor $u = [0, 1, 1]^T$. Pokażemy, że wektory $e_1$, $e_2$, $u$ tworzą bazę $\mathbb{F}^3$. Rozważmy dowolny wektor $[a, b, c]^T \in \mathbb{F}^3$. Musimy uzasadnić, że istnieją jedyne takie skalary $x_1$, $x_2$, $x_3 \in \mathbb{F}^3$, że $x_1 \cdot e_1 + x_2 \cdot e_2 + x_3 \cdot u = [a, b, c]^T$. Zwróćmy uwagę, że

$$x_1 \cdot e_1 + x_2 \cdot e_2 + x_3 \cdot u = [x_1, x_2 + x_3, x_3]^T.$$

Równość $[x_1, x_2 + x_3, x_3]^T = x_1 \cdot e_1 + x_2 \cdot e_2 + x_3 \cdot u = [a, b, c]^T$ jest więc równoważna z układem równań

$$x_1 = a, \quad x_2 + x_3 = b, \quad x_3 = c.$$

Ma on jednoznaczne rozwiązanie

$$x_1 = a, \quad x_2 = b - c, \quad x_3 = c,$$

czyli $e_1$, $e_2$, $u$ są bazą $\mathbb{F}^3$.

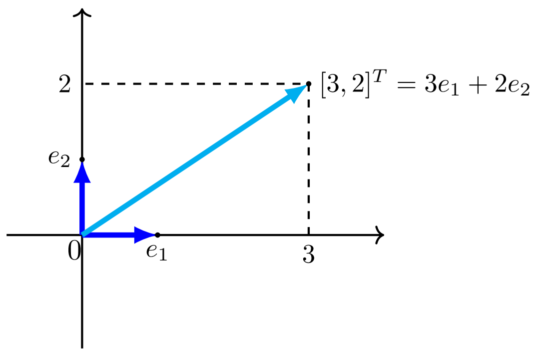
**Przykład 2.11.** Wektory $v_1 = [1, 0, 0]^T$, $v_2 = [0, 1, 1]^T$ są liniowo niezależne, bo równość

$$\begin{aligned}
[0, 0, 0]^T &= x_1 \cdot v_1 + x_2 \cdot v_2 \\
&= x_1 \cdot [1, 0, 0]^T + x_2 \cdot [0, 1, 1]^T \\
&= [x_1, x_2, x_2]^T,
\end{aligned}$$

oznacza, że $x_1 = x_2 = 0$. Nie tworzą one jednak bazy $\mathbb{F}^3$, bo ich kombinacje liniowe $x_1 \cdot v_1 + x_2 \cdot v_2$ mają postać $[x_1, x_2, x_2]^T$, więc przykładowo, wektor $[1, 2, 3]^T$ nie jest kombinacją liniową $v_1$ i $v_2$.

Każdy wektor $[x_1, x_2, x_3]^T \in \mathbb{F}^3$ jest kombinacją liniową wektorów $e_1$, $e_2$, $e_3$, $[0, 1, 1]^T$, bo

$$[x_1, x_2, x_3]^T = x_1 \cdot e_1 + x_2 \cdot e_2 + x_3 \cdot e_3 + 0 \cdot [0, 1, 1]^T.$$




Rysunek 2.3: Baza standardowa $e_1$, $e_2$ na płaszczyźnie. Każdy wektor na płaszczyźnie można jednoznacznie zapisać jako kombinację liniową wektorów $e_1$, $e_2$.

Wektory $e_1$, $e_2$, $e_3$, $[0,1,1]^T$ nie tworzą jednak bazy $\mathbb{F}^3$, bo przykładowo

$$\begin{aligned}[0,1,1]^T &= 0 \cdot e_1 + 1 \cdot e_2 + 1 \cdot e_3 + 0 \cdot [0,1,1]^T \\ &= 0 \cdot e_1 + 0 \cdot e_2 + 0 \cdot e_3 + 1 \cdot [0,1,1]^T,\end{aligned}$$

czyli przedstawienie nie jest jednoznaczne. Wynika to z faktu, że wektory $e_1$, $e_2$, $e_3$, $[0,1,1]^T$ są liniowo zależne, bo

$$0 \cdot e_1 + 1 \cdot e_2 + 1 \cdot e_3 - 1 \cdot u = 0.$$

**Przykład 2.12** (Przestrzeń wektorowa macierzy $M_{2\times 2}(\mathbb{F})$)**.** Rozważmy zbiór macierzy $M_{2\times 2}(\mathbb{F})$. Jest w nim określone działanie dodawania macierzy

$$+ : M_{2\times 2}(\mathbb{F}) \times M_{2\times 2}(\mathbb{F}) \longrightarrow M_{2\times 2}(\mathbb{F})$$

dane wzorem

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}$$

oraz mnożenie macierzy przez skalar

$$\cdot : \mathbb{F} \times M_{2\times 2}(\mathbb{F}) \longrightarrow M_{2\times 2}(\mathbb{F})$$

zdefiniowane wzorem

$$t \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ta & tb \\ tc & td \end{bmatrix}, \quad t \in \mathbb{F}.$$

Trójka $(M_{2\times 2}(\mathbb{F}), +, \cdot)$ jest przestrzenią wektorową nad ciałem $\mathbb{F}$. Ponadto, dla $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{F})$ mamy

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \cdot \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \cdot \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d \cdot \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Powyższym przedstawienie jest jednoznacznie wyznaczone przez macierz $A$, czyli macierze

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

tworzą bazę przestrzeni $M_{2\times 2}(\mathbb{F})$.

## 2.3 Iloczyn skalarny i norma w $\mathbb{R}^n$

Wprowadzimy teraz w przestrzeni wektorowej $(\mathbb{R}^n, +, \cdot)$ pewną dodatkową geometryczną strukturę, zadaną przez euklidesowy iloczyn skalarny.

**Definicja 2.10** (Euklidesowy iloczyn skalarny w $\mathbb{R}^n$)**.**
*Iloczynem skalarnym wektorów* $v = [v_1, v_2]^T, u = [u_1, u_2]^T \in \mathbb{R}^2$ nazywamy liczbę rzeczywistą

$$(v|u) := v_1 \cdot u_1 + v_2 \cdot u_2.$$

Analogicznie, iloczyn skalarny wektorów

$$v = [v_1, \ldots, v_n]^T, \quad u = [u_1, \ldots, u_n]^T \in \mathbb{R}^n$$

definiujemy jako liczbę rzeczywistą

$$(v|u) := v_1 \cdot u_1 + \ldots + v_n \cdot u_n$$

U Iloczyn skalarny jest też często oznaczany symbolem $\langle v, u\rangle$. My będziemy używali oznaczenia $(v|u)$, bo w tej konwencji nierówność $(v|u) > 0$ jest czytelniejsza niż napis $\langle v, u\rangle > 0$.

Iloczyn skalarny jest więc odwzorowaniem

$$(\cdot|\cdot) : \mathbb{R}^n \times \mathbb{R}^n \ni (v, u) \longmapsto (v|u) \in \mathbb{R}.$$

Przypisuje ono parze wektorów $(v, u)$ liczbę rzeczywistą $(v|u)$.

**Lemat 2.3.** *Dla dowolnych wektorów $u, v, w \in \mathbb{R}^n$ i skalarów $a, b \in \mathbb{R}$ zachodzą następujące własności*

*(1) (dwuliniowość)*

$$(a \cdot u + b \cdot v|w) = a(u|w) + b(v|w),$$
$$(u|a \cdot v + b \cdot w) = a(u|v) + b(u|w).$$

*(2) (symetryczność)*

$$(u|v) = (v|u)$$

*(3) (dodatnia określoność)*

$$(u|u) \geqslant 0$$

*przy czym $(u|u) = 0$ wtedy i tylko wtedy, gdy $u = 0$.*

**Definicja 2.11** (Norma euklidesowa w $\mathbb{R}^n$)**.**
*Normą (długością) wektora* $u = [u_1, \ldots, u_n]^T \in \mathbb{R}^n$ nazywamy liczbę

$$\|u\| := \sqrt{(u|u)} = \sqrt{u_1^2 + \ldots + u_n^2} \geqslant 0$$

U Zauważmy, że wektor $[\cos\theta, \sin\theta]^T$ ma normę jeden i każdy wektor o normie jeden można zapisać w tej postaci dla pewnego $\theta$. W szczególności, dowolny wektor $v \in \mathbb{R}^2$ można zapisać w postaci

$$v = \|v\|[\cos\theta, \sin\theta]^T.$$

**Lemat 2.4** (Tożsamość równoległoboku)**.** *Dla dowolnych wektorów $u, v \in \mathbb{R}^n$ zachodzi równość*

$$\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$$

*Dowód.* Bezpośredni rachunek pokazuje, że

$$\begin{aligned}\|u+v\|^2+\|u-v\|^2 &= (u+v|u+v)+(u-v|u-v)\\ &= (u|u)+(u|v)+(v|u)+(v|v)\\ &+ (u|u)-(u|v)-(v|u)+(v|v)\\ &= 2(u|u)+2(v|v)=2(\|u\|^2+\|v\|^2).\end{aligned}$$

□

**Lemat 2.5** (Wzór polaryzacyjny)**.** *Zachodzi wzór polaryzacyjny*

$$(u|v)=\tfrac{1}{4}(\|u+v\|^2-\|u-v\|^2).$$

*Dowód.* Mamy

$$\begin{aligned}\|u+v\|^2-\|u-v\|^2 &= (u+v|u+v)-(u-v|u-v)\\ &= (u|u)+(u|v)+(v|u)+(v|v)\\ &- ((u|u)-(u|v)-(v|u)+(v|v))\\ &= 4(u|v).\end{aligned}$$

□

**Lemat 2.6** (Twierdzenie Pitagorasa)**.** *Niech $u, v \in \mathbb{R}^n$ będą wektorami. Wówczas,*

$$(u,v)=0 \Leftrightarrow \|u-v\|^2=\|u\|^2+\|v\|^2.$$

*Dowód.* Zauważmy, że

$$\begin{aligned}\|u-v\|^2 &= (u-v|u-v)\\ &= \|u\|^2-2(u|v)+\|v\|^2,\end{aligned}$$

czyli teza zachodzi. □

**Twierdzenie 2.1** (Nierówność Cauchy'ego–Schwarza)**.** *Dla dowolnych wektorów $u$ i $v \in \mathbb{R}^n$ zachodzi nierówność*

$$|(u|v)| \leqslant \|u\|\|v\|.$$

*Dowód.* Możemy założyć że wektory $u$ i $v$ są niezerowe. Dla dowolnego $t \in \mathbb{R}$ mamy

$$\begin{aligned}0 &\leqslant \|u+t\cdot v\|^2\\ &= (u+t\cdot v|u+t\cdot v)\\ &= \|u\|^2+2t(u|v)+t^2\|v\|^2,\end{aligned}$$

czyli z własności funkcji kwadratowych wynika, że

$$4((u|v))^2-4\|u\|^2\|v\|^2 \leqslant 0,$$

co kończy dowód. □

**Wniosek 2.1** (Własności normy)**.** *Dla dowolnych $u, v \in \mathbb{R}^n$ i $a \in \mathbb{R}$ mamy*

*(1)* $\|u\| \geqslant 0$ *oraz* $\|u\| = 0$ *wtedy i tylko wtedy, gdy* $u = 0$*;*

*(2)* $\|a\cdot u\| = |a|\cdot\|u\|$*;*

*(3) (nierówność trójkąta)*

$$\|u+v\| \leqslant \|u\|+\|v\|$$

*Dowód.* Pierwsza własność wynika bezpośrednio z definicji normy. Dla dowodu drugiej zauważamy, że

$$\begin{aligned}\|a \cdot u\|^2 &= (a \cdot u | a \cdot u) \\ &= a^2 (u|u) \\ &= a^2 \|u\|^2.\end{aligned}$$

Uzasadnimy nierówność trójkąta. Stosując nierówność Cauchy'ego–Schwarza, mamy

$$\begin{aligned}\|u + v\|^2 &= \|u\|^2 + 2(u|v) + \|v\|^2 \\ &\leqslant \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\ &= (\|u\| + \|v\|)^2,\end{aligned}$$

co kończy dowód. □

**Definicja 2.12** (Wektory ortogonalne)**.**
Wektory $u, v \in \mathbb{R}^n$ są *ortogonalne (prostopadłe)* wtedy i tylko wtedy, gdy

$$(u|v) = 0.$$

W szczególności, wektor zerowy jest prostopadły do każdego wektora, więc również do samego siebie.

U Pokażemy, że dla dowolnych wektorów $u, v \in \mathbb{R}^2$ zachodzi równość

$$(u|v) = \|u\|\|v\| \cos \angle(u, v),$$

gdzie $\angle(u, v)$ jest kątem między wektorami $u$ i $v$. W szczególności,

$$|(u|v)| \leqslant \|u\|\|v\|.$$

Jeśli któryś z wektorów $u$ lub $v$ jest zerowy, to powyższa równość jest prawdziwa. Załóżmy więc, że wektory $u$ i $v$ są niezerowe. Niech wektor $q$ będzie rzutem prostopadłym wektora $u$ na prostą rozpiętą przez wektor $v$. Wtedy $q = c \cdot v$ dla pewnego $c \in \mathbb{R}$. Stąd wektor $w = u - c \cdot v$ jest prostopadły do wektora $v$. Gdy kąt $\angle(u, v)$ jest ostry, to $c \geqslant 0$ oraz

$$\cos \angle(u, v) = \frac{\|c \cdot v\|}{\|u\|} = |c| \frac{\|v\|}{\|u\|}.$$

Jeśli kąt $\angle(u, v)$ jest rozwarty, to $c < 0$ oraz

$$\cos \angle(u, v) = \cos(\pi - \angle(u, -v)) = -\cos \angle(u, -v) = -|c| \frac{\|v\|}{\|u\|}.$$

W obydwu przypadkach

$$\cos \angle(u, v) = c \frac{\|v\|}{\|u\|}.$$

Ponieważ wektor $w = u - c \cdot v$ jest prostopadły do $v$, więc

$$0 = (u - c \cdot v, v) = (u|v) - c\|v\|^2,$$

czyli

$$c = \frac{(u|v)}{\|v\|^2}.$$

Stąd

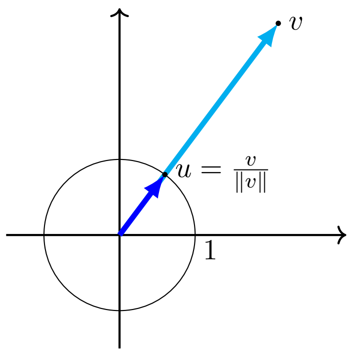
$$\cos \angle(u, v) = \frac{(u|v)}{\|v\|^2} \frac{\|v\|}{\|u\|} = \frac{(u|v)}{\|v\|\|u\|},$$

co kończy dowód.
Możemy na to spojrzeć jeszcze inaczej. Wektory $u$ i $v$ możemy zapisać w postaci trygonometrycznej:

$$u = |u|(\cos\theta, \sin\theta)^T, \quad v = |v|(\cos\psi, \sin\psi)^T.$$

Wtedy

$$(u|v) = |u||v|(\cos\theta\cos\psi + \sin\theta\sin\psi) = |u||v|\cos(\theta - \psi).$$




Rysunek 2.4: Normalizacja niezerowego wektora $v$. Wektor $u = \frac{v}{\|v\|}$ leży na okręgu jednostkowym.

Z nierówności Cauchy'ego–Schwarza wynika, że dla niezerowych wektorów $u, v \in \mathbb{R}^n$ mamy

$$-1 \leqslant \frac{(u|v)}{\|u\|\|v\|} \leqslant 1, \quad u, v \neq 0.$$

Istnieje więc jednoznacznie wyznaczony taki kąt $0 \leqslant \angle(u, v) \leqslant \pi$, że

$$\cos \angle(u, v) = \frac{(u|v)}{\|u\|\|v\|}.$$

**Definicja 2.13** (Kąt między wektorami)**.**
*Kątem* $\angle(u, v)$ *pomiędzy niezerowymi wektorami* $u, v \in \mathbb{R}^n$ nazywamy jedyny taki kąt $0 \leqslant \angle(u, v) \leqslant \pi$, że

$$\cos \angle(u, v) = \frac{(u|v)}{\|u\|\|v\|}.$$

**Definicja 2.14** (Wektor jednostkowy)**.**
Wektor $u \in \mathbb{R}^n$ nazywamy *jednostkowym*, gdy $\|u\| = 1$.

**Przykład 2.13.** Jeśli wektor $v \in \mathbb{R}^n$ jest niezerowy, to wektor $u = \frac{v}{\|v\|}$ jest wektorem jednostkowym.

**Przykład 2.14.** Niech $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{R})$. Niech $a_1 = [a, c]^T$ i $a_2 = [b, d]^T \in \mathbb{R}^2$ będą wektorami utworzonymi z kolumn macierzy $A$. Macierz Grama ma postać

$$A^T A = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} (a_1|a_1) & (a_1|a_2) \\ (a_1|a_2) & (a_2|a_2) \end{bmatrix}.$$

W szczególności, $A \in O(2)$ jest ortogonalna tzn. $A^T A = I$ wtedy i tylko wtedy, gdy

$$\|a_1\| = \|a_2\| = 1, \quad (a_1|a_2) = 0,$$

czyli kolumny $A$ są wektorami jednostkowymi i są ortogonalne do siebie.

**U** Metryka euklidesowa

Korzystając z normy możemy zdefiniować funkcję

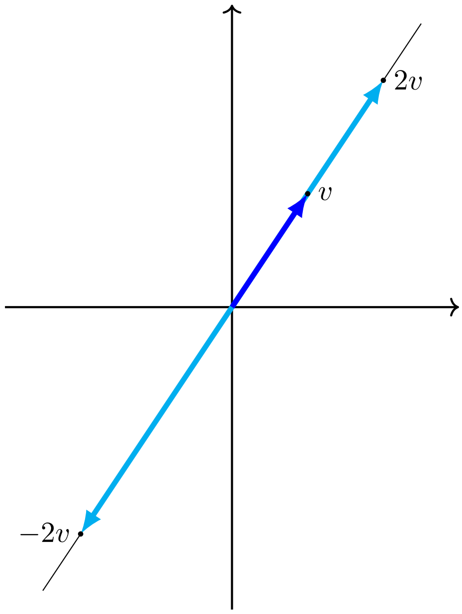
$$d : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R},$$

wzorem

$$d(v, w) = \|v - w\| = \sqrt{(v_1 - w_1)^2 + \ldots + (v_n - w_n)^2}.$$

Z własności normy, dla dowolnych $v, w, u \in \mathbb{R}^n$ mamy

(i) $d(v, w) \geqslant 0$ oraz $d(v, w) = 0$ wtedy i tylko wtedy, gdy $v = w$,




Rysunek 2.5: Mnożąc niezerowy wektor $v$ przez wszystkie liczby rzeczywiste $t$, otrzymujemy prostą.

(ii) $d(v, w) = d(w, v)$,

(iii) $d(v, w) \leqslant d(v, u) + d(u, w)$.

Przykładowo,

$$\begin{aligned} d(v, w) &= \|v - w\| \\ &= \|(v - u) + (u - w)\| \\ &\leqslant \|v - u\| + \|u - w\| \\ &= d(v, u) + d(u, w) \end{aligned}$$

Funkcję $d$ nazywamy *metryką euklidesową* w $\mathbb{R}^n$. Liczbę $d(v, w)$ nazywamy *odległością* między $v$ i $w$.

## 2.4 Proste i płaszczyzny

Rozważmy *niezerowy* wektor $v = [v_1, v_2]^T \in \mathbb{R}^2$. Jak wiemy, zbiór wszystkich wektorów postaci $t \cdot v$ dla $t \in \mathbb{R}$ jest prostą równoległą do wektora $v$ i przechodzącą przez punkt 0. Prostą równoległą do $v$ i przechodzącą przez punkt $[a_1, a_2]^T$ możemy opisać jako zbiór punktów postaci $[a_1, a_2]^T + t \cdot v$ $(t \in \mathbb{R})$.

**Definicja 2.15** (Równanie parametryczne prostej)**.**
Równanie parametryczne prostej $l$ przechodzącej przez $[a_1, a_2]^T$ i równoległej do niezerowego wektora $v = [v_1, v_2]^T$ ma postać

$$[x_1, x_2]^T = [a_1, a_2]^T + t \cdot [v_1, v_2]^T, \quad t \in \mathbb{R}^2$$

W skrócie, $x = a + tv$.

U

Definicja oznacza, że współrzędne punktu $[x_1, x_2]^T$ leżącego na prostej $l$, spełniają, dla pewnego skalara $t$, układ równań

$$\begin{cases} x_1 = a_1 + tv_1 \\ x_2 = a_2 + tv_2 \end{cases}$$

Ponieważ wektor $v$ jest niezerowy, więc któraś z jego współrzędnych $v_1$ lub $v_2$ jest różna od zera. Załóżmy przykładowo, że $v_1 \neq 0$. Możemy wtedy z pierwszego równania wyliczyć $t$, otrzymując, że $t = \dfrac{x_1 - a_1}{v_1}$. Wstawiając do drugiego równania, otrzymujemy, że

$$x_2 - a_2 - \frac{x_1 - a_1}{v_1} v_2 = 0,$$

czyli

$$-v_2(x_1-a_1)+v_1(x_2-a_2)=0.$$

Jest to *równanie ogólne* prostej $l$ przechodzącej przez punkt $[a_1,a_2]^T$ i równoległej do wektora $v$.

**Definicja 2.16** (Równanie ogólne prostej)**.**
Równanie ogólne prostej na płaszczyźnie ma postać

$$Ax_1+Bx_2+C=0, \quad A^2+B^2>0.$$

U Zauważmy, że prosta o równaniu ogólnym

$$Ax_1+Bx_2+C=0$$

przechodzi przez początek układu współrzędnych $0=[0,0]^T$ wtedy i tylko wtedy, gdy $C=0$. Ma ona wtedy równanie $Ax_1+Bx_2=0$, które możemy zapisać przy pomocy iloczynu skalarnego następująco:

$$\big([A,B]^T|[x_1,x_2]^T\big)=0.$$

Oznacza to, że składa się ona ze wszystkich wektorów $[x_1,x_2]^T$ prostopadłych do wektora $[A,B]^T$.

**Wniosek 2.2.** *Prosta o równaniu ogólnym*

$$A(x_1-a_1)+B(x_2-a_2)=0, \quad A^2+B^2>0$$

*przechodzi przez punkt* $[a_1,a_2]^T$ *i jest prostopadła do wektora* $[A,B]^T$*.*

Analogicznie, dla niezerowego wektora $v=[v_1,\dots,v_n]^T\in\mathbb{R}^n$ i punktu $a=[a_1,\dots,a_n]^T$ *równanie parametryczne prostej* równoległej do $v$ i przechodzącej przez $a$ ma postać

$$x=a+t\cdot v, \quad t\in\mathbb{R}.$$

Przyglądnijmy się przypadkowi $n=3$, czyli prostym w przestrzeni $\mathbb{R}^3$. W postaci parametrycznej jest ona zadana układem równań

$$\begin{cases} x_1=a_1+tv_1 \\ x_2=a_2+tv_2 \\ x_3=a_3+tv_3 \end{cases}.$$

Jeżeli wszystkie liczby $v_1\neq 0$ dla $i=1,2,3$, to eliminując $t$ z powyższego układu, otrzymujemy dwa równania

$$\frac{x-a_1}{v_1}=\frac{x-a_2}{v_2}=\frac{x-a_3}{v_3},$$

nazywane *równaniami kanonicznymi* prostej w przestrzeni. Jeżeli np. $v_3=0$ a $v_1$, $V_2\neq 0$, to równania przyjmują postać

$$\frac{x-a_1}{v_1}=\frac{x-a_2}{v_2}, \quad x_3=a_3.$$

Gdy np. $v_2=v_3=0$ i $v_1\neq 0$, to równanie ma postać

$$x_2=a_2, \quad x_3=a_3.$$

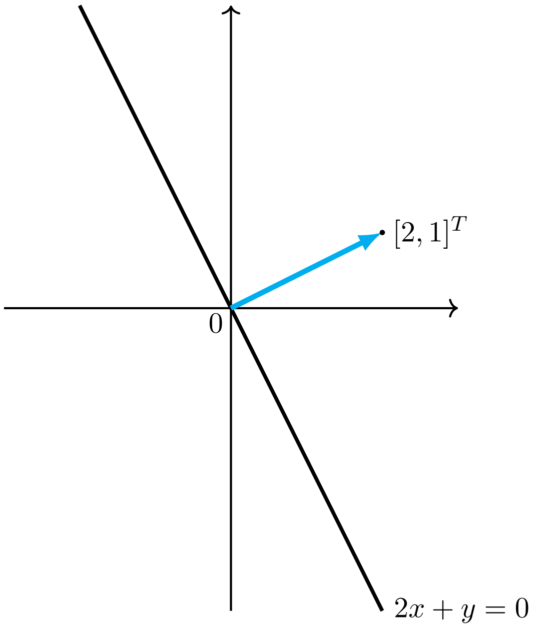
**Przykład 2.15.** Zastanówmy się jaki podzbiór $P$ przestrzeni $\mathbb{R}^3$ jest opisany przez równanie

$$2x_1-x_2+3x_3=0.$$

Możemy je zapisać równoważnie, wykorzystując iloczyn skalarny, jako

$$([2,-1,3]^T|[x_1,x_2,x_3]^T)=0,$$

czyli zbiór $P$ składa się z wszystkich wektorów $[x_1,x_2,x_3]^T$ w $\mathbb{R}^3$ prostopadłych do wektora $[2,-1,3]^T$. Oznacza to, że $P$ jest płaszczyzną prostopadłą do wektora $[2,-1,3]^T$ i zawierającą początek układu $0=[0,0,0]^T$.




Rysunek 2.6: Prosta $2x + y = 0$ jest prostopadła do wektora $[2, 1]^T$.

**Definicja 2.17** (Równanie ogólne płaszczyzny w $\mathbb{R}^3$)**.**
Równanie ogólne płaszczyzny w przestrzeni $\mathbb{R}^3$ ma postać

$$Ax_1 + Bx_2 + Cx_3 + D = 0, \quad A^2 + B^2 + C^2 > 0.$$

W szczególności, równanie

$$A(x_1 - a_1) + B(x_2 - a_2) + C(x_3 - a_3) = 0$$

opisuje płaszczyznę przechodzącą przez punkt $[a_1, a_2, a_3]^T$ i prostopadłą do wektora $[A, B, C]^T$.

**Przykład 2.16.** Rozważmy układ równań liniowych

$$\begin{cases} x - 2y + 3z = 1 \\ x + y + z = 1 \end{cases}$$

Równania układu opisują dwie nierównoległe płaszczyzny. Zbiorem rozwiązań układu jest prosta $L$ będąca ich częścią wspólną.

Wyznaczymy dokładniej prostą $L$. Odejmując pierwsze równanie od drugiego otrzymujemy układ równoważny

$$\begin{cases} x - 2y + 3z = 1 \\ 3y - 2z = 0 \end{cases} \iff \begin{cases} x - 2y + 3z = 1 \\ y = \frac{2}{3}z \end{cases}$$

Stąd

$$\begin{cases} x = 1 - \frac{5}{3}z \\ y = \frac{2}{3}z \end{cases}$$

Zbiór rozwiązań układu jest dany jako zbiór

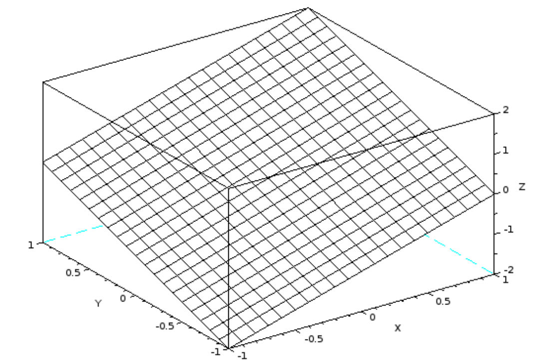
$$\Big\{[1 - 5/3z, 2/3z, z]^T : z \in \mathbb{R}\Big\}.$$

Zauważmy, że

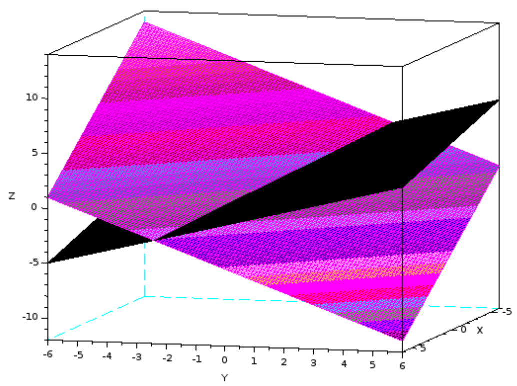
$$[1 - 5/3z, 2/3z, z]^T = [-5/3, 2/3, 1]^T z + [1, 0, 0]^T.$$

Rozwiązaniem układu jest prosta $L$ o równaniu parametrycznym

$$[-5/3, 2/3, 1]^T z + [1, 0, 0]^T, \quad z \in \mathbb{R}.$$




Rysunek 2.7: Płaszczyzna $x + y - z = 0$.




Rysunek 2.8: Płaszczyzny $x - 2y + 3z = 1$, $x + y + z = 1$ i ich przecięcie, będące prostą w $\mathbb{R}^3$.

# Rozdział 3

# Układy równań liniowych

SŁOWA KLUCZOWE ROZDZIAŁU
układ równań liniowych ◊ dozwolone operacje ◊ macierz rozszerzona układu ◊ eliminacja Gaussa ◊ postać schodkowa ◊ podprzestrzeń $\operatorname{span}\{a_1, \ldots, a_k\}$ ◊ układy jednorodne ◊ wyznacznik ◊ wzory Cramera

Rozważmy układ równań

$$\begin{cases} x_1 + x_2 + x_3 = 0 \\ x_1 - x_2 + x_3 = 0 \end{cases}$$

Każde z równań układu opisuje pewną płaszczyznę w $\mathbb{R}^3$. Szukamy więc punktów leżących na obydwu płaszczyznach. Obydwie płaszczyzny zawierają punkt 0. Pierwsza z nich jest prostopadła do wektora $[1, 1, 1]^T$, a druga jest prostopadła do wektora $[1, -1, 1]^T$. Nie są one równoległe, więc ich przecięciem jest pewna prosta w $\mathbb{R}^3$. Naszym najbliższym celem będzie nauczenie się wyznaczania zbioru rozwiązań układu równań liniowych.

**Definicja 3.1** (Układ równań liniowych)**.**
Niech $\mathbb{F}$ będzie ciałem. *Układem równań liniowych* nazywamy układ równań postaci:

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n & = b_1 \\ \qquad\vdots & \quad\vdots \\ a_{k1}x_1 + \ldots + a_{kn}x_n & = b_k \end{cases} \tag{3.1}$$

Skalary $a_{ij} \in \mathbb{F}$ nazywamy *współczynnikami* układu, a $b_i \in \mathbb{F}$ *wyrazami wolnymi*. Jest to układ $k$ równań z $n$ niewiadomymi $x_1, \ldots, x_n$. *Zbiór rozwiązań* układu składa się ze wszystkich wektorów $[x_1, \ldots, x_n]^T \in \mathbb{F}^n$, których współrzędne spełniają wszystkie równania tego układu. Mówimy, że układ

- jest *sprzeczny*, gdy zbiór rozwiązań jest pusty;
- *ma rozwiązanie*, jeśli zbiór rozwiązań jest niepusty;
- *ma nieskończenie wiele rozwiązań*, gdy zbiór rozwiązań jest nieskończony;
- *ma dokładnie jedno rozwiązanie*, jeśli zbiór rozwiązań jest jednoelementowy.

U Dwa układy równań są *równoważne*, jeżeli mają takie same zbiory rozwiązań.
Istnieją pewne operacje, które możemy wykonać na układzie równań (3.1) otrzymując układ równoważny.

DOZWOLONE OPERACJE NA RÓWNANIACH UKŁADU

(I) możemy dwa równania układu zamienić miejscami;

(II) możemy obydwie strony któregoś z równań pomnożyć przez dowolny niezerowy skalar;

(III) możemy równanie pomnożone przez dowolny skalar dodać do innego równania.

Punkty (I) i (II) są oczywiste. Punkt (III) wynika z faktu, że $[x_1, \ldots, x_n]^T$ spełnia

$$\begin{aligned} a_{i1}x_1 + \ldots + a_{in}x_n &= b_i \\ a_{j1}x_1 + \ldots + a_{jn}x_n &= b_j \end{aligned}$$

wtedy i tylko wtedy, gdy spełnia

$$\begin{aligned} a_{i1}x_1 + \ldots \quad + a_{in}x_n &= b_i \\ (a_{j1} + ca_{i1})x_1 + \ldots + (a_{jn} + ca_{in})x_n &= b_j + cb_i. \end{aligned}$$

**Definicja 3.2** (Dozwolone operacje)**.**
Operacje (I)–(III) nazywamy dozwolonymi operacjami na równaniach układu równań liniowych.

**Przykład 3.1.** Szczególnie łatwe do rozwiązania są układu w postaci „schodkowej". Przykładowo, w układzie równań (w ciele $\mathbb{R}$)

$$\begin{cases} 3x_1 + 2x_2 + \phantom{2}x_3 = 1 \\ \phantom{3x_1 + 2}x_2 - \phantom{2}x_3 = 2 \\ \phantom{3x_1 + 2x_2 +} 2x_3 = 4 \end{cases}$$

ostatnie równanie oznacza, że $x_3 = 2$. Wtedy z drugiego równania $x_2 = 4$. W konsekwencji, z pierwszego równania otrzymujemy, że $x_1 = -3$. Zbiór rozwiązań układu składa się z jednego wektora $[-3, 4, 2]^T$.

## 3.1 Eliminacja Gaussa

Opiszemy na przykładzie *metodę eliminacji Gaussa*, pozwalającą sprowadzić dowolny układ do postaci schodkowej. Rozważmy układ równań (w ciele $\mathbb{R}$)

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 1 \\ -x_1 - x_2 + 0x_3 + 0x_4 + x_5 = -1 \\ -2x_1 - 2x_2 + 0x_3 + 0x_4 + 3x_5 = 1 \\ 0x_1 + 0x_2 + x_3 + x_4 + 3x_5 = 3 \\ x_1 + x_2 + 2x_3 + 2x_4 + 4x_5 = 4 \end{cases}$$

Wszystkie informacje o rozważanym układzie możemy wygodnie zakodować przy pomocy macierzy (tabeli):

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 \\ -2 & -2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 1 & 3 \\ 1 & 1 & 2 & 2 & 4 \end{bmatrix}, \quad \begin{bmatrix} A \mid b \end{bmatrix} = \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 & -1 \\ -2 & -2 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 1 & 3 & 3 \\ 1 & 1 & 2 & 2 & 4 & 4 \end{array}\right].$$

W wierszach macierzy $\begin{bmatrix} A \mid b \end{bmatrix}$ wypisujemy współczynniki i wyrazy wolne poszczególnych równań. Macierz $A$ nazywamy *macierzą układu*, a macierz $\begin{bmatrix} A \mid b \end{bmatrix}$ jego macierzą *rozszerzoną*. Dozwolone operacje (I)–(III) mają swój odpowiednik w postaci dozwolonych operacji na wierszach macierzy:

Dozwolone operacje na wierszach macierzy układu równań liniowych

(I) możemy dwa wiersze macierzy zamienić miejscami (permutacja);

(II) możemy wiersz pomnożyć przez dowolną niezerową liczbę rzeczywistą (skalowanie);

(III) możemy wiersz pomnożony przez dowolną liczbę rzeczywistą dodać do innego wiersza.

Zaczynamy od tego, że szukamy wiersza w którym pierwszy wyraz jest niezerowy. W naszym przypadku jest to wiersz pierwszy. Jeśli pierwszy niezerowy wyraz jest różny od 1, to używamy skalowania (II), aby uzyskać 1 na pierwszej pozycji.

U Jeżeli w pierwszym wierszu na pierwszej pozycji jest zero, to szukamy innego wiersza z niezerowym wyrazem na pierwszej pozycji. Jeśli taki istnieje, to używamy pierwszej operacji (I) i zamieniamy go miejscami z wierszem pierwszym. Może się zdarzyć, że wszystkie wiersze mają na pierwszej pozycji 0. Wtedy powtarzamy procedurę szukając wiersza z niezerowym wyrazem na drugiej pozycji i iterujemy (powtarzamy) poprzednią procedurę.

W pierwszym kroku używamy pierwszego wiersza i operacji (III) do wprowadzenia zer w pierwszej kolumnie pod jedynką w pierwszym wierszu. Dodając pierwszy wiersz do drugiego otrzymujemy macierz:

$$\left[\begin{array}{ccccc|c} \mathbf{1} & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ \boxed{-2} & -2 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 1 & 3 & 3 \\ 1 & 1 & 2 & 2 & 4 & 4 \end{array}\right]$$

Następnie, pierwszy wiersz pomnożony przez 2 dodajemy do wiersza trzeciego

$$\left[\begin{array}{ccccc|c} \mathbf{1} & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 2 & 2 & 5 & 3 \\ 0 & 0 & 1 & 1 & 3 & 3 \\ \boxed{1} & 1 & 2 & 2 & 4 & 4 \end{array}\right]$$

Kończymy ten krok, odejmując pierwszy wiersz od wiersza piątego

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & \mathbf{1} & 1 & 2 & 0 \\ 0 & 0 & \boxed{2} & 2 & 5 & 3 \\ 0 & 0 & \boxed{1} & 1 & 3 & 3 \\ 0 & 0 & \boxed{1} & 1 & 3 & 3 \end{array}\right].$$

W drugim kroku używamy analogicznie drugiego wiersza do eliminacji niezerowych elementów w trzeciej kolumnie i wierszach od trzeciego do piątego. W efekcie otrzymujemy

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 3 \\ 0 & 0 & 0 & 0 & \boxed{1} & 3 \\ 0 & 0 & 0 & 0 & \boxed{1} & 3 \end{array}\right].$$

W ostatnim kroku używamy trzeciego wiersza do eliminacji jedynek w piątej kolumnie oraz czwartym i piątym wierszu. Otrzymujemy macierz:

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Ostatnie dwa wiersze nie niosą ze sobą żadnej informacji, więc możemy je pominąć, otrzymując w efekcie macierz

$$\left[\begin{array}{ccccc|c} \mathbf{1} & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & \mathbf{1} & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 3 \end{array}\right],$$

której odpowiada układ równań

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 1 \\ x_3 + x_4 + 2x_5 = 0 \\ x_5 = 3 \end{cases}$$

Wstawiamy teraz $x_5 = 3$ do trzeciego równania i wyliczamy w nim $x_3$, otrzymując

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 1 \\ x_3 = -x_4 - 6 \\ x_5 = 3 \end{cases}$$

Cofamy się teraz do pierwszego równania i wyliczamy $x_1$ po wstawieniu wartości za $x_3$ i $x_5$, otrzymując

$$\begin{cases} x_1 = -x_2 + 4 \\ x_3 = -x_4 - 6 \\ x_5 = 3 \end{cases}$$

$x_2$ i $x_4$ pełnią tu rolę parametrów. Zbiór rozwiązań jest dany przez

$$\left\{[-x_2 + 4, x_2, -x_4 - 6, x_4, 3]^T : x_2, x_4 \in \mathbb{R}\right\}.$$

Przykładowo, dla $x_2 = x_4 = 0$ jednym z rozwiązań jest $[4, 0, -6, 0, 3]^T$. Zbiór rozwiązań jest nieskończony, bo dowolnie wybranym $x_2$ i $x_4$ odpowiada jakieś rozwiązanie.

**Przykład 3.2.** Przyglądnijmy się układom trzech równań z trzema niewiadomymi. Dla macierzy $A$ z niezerową pierwszą kolumną postać schodkowa musi być wtedy jednej z postaci:

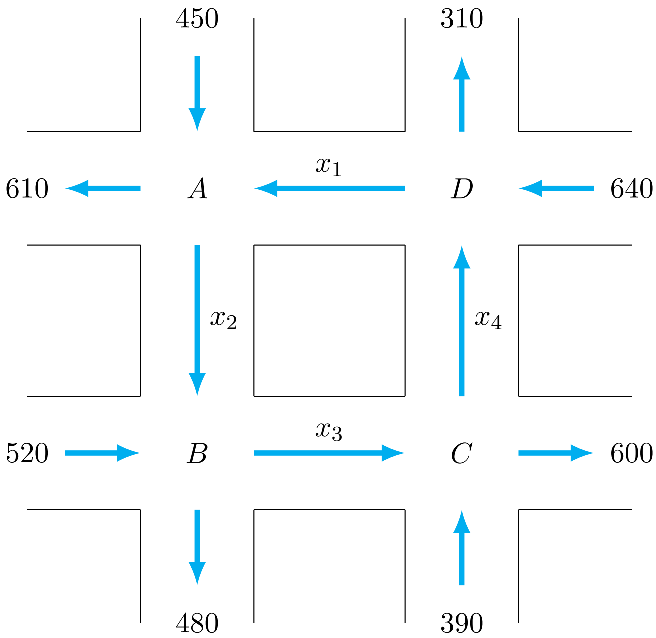
$$\left[\begin{array}{ccc|c} \mathbf{1} & * & * & * \\ 0 & \mathbf{1} & * & * \\ 0 & 0 & \mathbf{1} & * \end{array}\right], \quad \left[\begin{array}{ccc|c} \mathbf{1} & * & * & * \\ 0 & \mathbf{1} & * & * \\ 0 & 0 & 0 & \mathbf{0} \end{array}\right], \quad \left[\begin{array}{ccc|c} \mathbf{1} & * & * & * \\ 0 & 0 & \mathbf{1} & * \\ 0 & 0 & 0 & \mathbf{0} \end{array}\right],$$

$$\left[\begin{array}{ccc|c} \mathbf{1} & * & * & * \\ 0 & \mathbf{1} & * & * \\ 0 & 0 & 0 & \neq \mathbf{0} \end{array}\right], \quad \left[\begin{array}{ccc|c} \mathbf{1} & * & * & * \\ 0 & 0 & \mathbf{1} & * \\ 0 & 0 & 0 & \neq \mathbf{0} \end{array}\right], \quad \left[\begin{array}{ccc|c} \mathbf{1} & * & * & * \\ 0 & 0 & 0 & \neq \mathbf{0} \\ 0 & 0 & 0 & \mathbf{0} \end{array}\right], \quad \left[\begin{array}{ccc|c} \mathbf{1} & * & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{0} \end{array}\right]$$

W pierwszym przypadku układ ma dokładnie jedno rozwiązanie. W przypadku drugim $x_1$ i $x_2$ możemy uzależnić od $x_3$, które pełni rolę parametru. Zbiór rozwiązań jest nieskończony, bo $x_3$ może być dowolne. W trzecim przypadku $x_3$ jest jednoznacznie wyznaczone i możemy uzależnić $x_1$ od $x_2$ pełniącego rolę parametru. Ponownie zbiór rozwiązań jest nieskończony. W czwartym, piątym i szóstym przypadku układ jest sprzeczny. W ostatnim układzie zbiorem rozwiązań jest płaszczyzna.

U

Zredukowana postać schodkowa macierzy $A \in M_{m\times n}(\mathbb{F})$ jest wyznaczona jednoznacznie. Uzasadnimy ten fakt, stosując indukcję względem $n$, czyli liczby kolumn macierzy $A$. Dla $n = 1$ jest to oczywiste. Niech $n > 1$. Rozważmy macierz $A'$ otrzymaną z $A$ przez skreślenie ostatniej kolumny. Każdy ciąg operacji elementarnych, który sprowadza $A$ do zredukowanej postaci schodkowej, sprowadza również macierz $A'$ do zredukowanej postaci schodkowej. Z założenia indukcyjnego wynika więc, że jeśli $B$ i $C$ są zredukowanymi postaciami schodkowymi macierzy $A$, to $B$ i $C$ mogą się różnić jedynie ostatnią kolumną. Przypuśćmy, że $B \neq C$. Z definicji zredukowanej postaci schodkowej wynika, że na przykład $n$-ta kolumna $B$ jest niezerowa, a $n$-ta kolumna $C$ jest zerowa. Rozważmy taki wektor $x$, że $Bx = 0$. Równoważnie $Cx = 0$, bo dozwolone operacje nie zmieniają zbioru rozwiązań. Stąd również $(B - C)x = 0$, czyli $x_n = 0$, bo $n - 1$-pierwszych kolumn macierzy $B - C$ jest zerowych, a $n$-ta jest niezerowa. Prowadzi to do sprzeczności, bo $n$-ta kolumna macierzy $C$ jest zerowa, więc $x_n$ w takim wektorze $x$, że $Cx = 0$ może być dowolne.




Rysunek 3.1: Natężenie ruchu w sieci skrzyżowań.

**Przykład 3.3.** Znajdziemy wielomian stopnia 3 przechodzący przez punkty

$$(x_1, y_1) = (1, 3), \quad (x_2, y_2) = (2, -2), \quad (x_3, y_3) = (3, -5), \quad (x_4, y_4) = (4, 0).$$

Szukany wielomian ma postać

$$w(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3.$$

Jego współczynniki $a_0, a_1, a_2, a_3$ muszą spełniać układ równań

$$\begin{cases} a_0 + a_1 x_1 + a_2 x_1^2 + a_3 x_1^3 = y_1 \\ a_0 + a_1 x_2 + a_2 x_2^2 + a_3 x_2^3 = y_2 \\ a_0 + a_1 x_3 + a_2 x_3^2 + a_3 x_3^3 = y_3 \\ a_0 + a_1 x_4 + a_2 x_4^2 + a_3 x_4^3 = y_4 \end{cases}$$

o macierzy rozszerzonej

$$\left[\begin{array}{cccc|c} 1 & x_1 & x_1^2 & x_1^3 & y_1 \\ 1 & x_2 & x_2^2 & x_2^3 & y_2 \\ 1 & x_3 & x_3^2 & x_3^3 & y_3 \\ 1 & x_4 & x_4^2 & x_4^3 & y_4 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 3 \\ 1 & 2 & 4 & 8 & -2 \\ 1 & 3 & 9 & 27 & -5 \\ 1 & 4 & 16 & 64 & 0 \end{array}\right] \dashrightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & -5 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right].$$

Stąd

$$a_0 = 4, \quad a_1 = 3, \quad a_2 = -5, \quad a_3 = 1,$$

czyli

$$w(x) = 4 + 3x - 5x^2 + x^3.$$

**Przykład 3.4.** Rozważmy system skrzyżowań dwóch zbiorów dróg jednokierunkowych pokazany na rysunku 3.1. Liczby przy strzałkach oznaczają godzinną średnią liczbę pojazdów wkraczających i opuszczających skrzyżowanie w godzinach szczytu. Ponieważ liczba pojazdów wjeżdżających na skrzyżowanie jest równa liczbie pojazdów z niego wyjeżdżających, więc otrzymujemy układ równań

$$\begin{cases} x_1 + 450 = x_2 + 610 \\ x_2 + 520 = x_3 + 480 \\ x_3 + 390 = x_4 + 600 \\ x_4 + 640 = x_1 + 310 \end{cases}$$

**Przykład 3.5.** Rozwiążemy następujący układ równań nad $\mathbb{Z}_3$:

$$\begin{cases} x_1 + 2x_2 + x_3 = 0 \\ x_1 + x_3 = 2 \\ x_2 + 2x_3 = 1 \end{cases}$$

Jest on równoważny z układem

$$\begin{cases} x_1 + 2x_2 + x_3 = 0 \\ 3x_1 + 4x_2 + 3x_3 = 2 \\ x_2 + 2x_3 = 1 \end{cases}$$

czyli

$$\begin{cases} x_1 + 2x_2 + x_3 = 0 \\ 0x_1 + x_2 + 0x_3 = 2 \\ x_2 + 2x_3 = 1 \end{cases}$$

Stąd $x_2 = 2$, więc $2 + 2x_3 = 1$, czyli $x_3 = 1$ oraz $x_1 = 1$.

ELIMINACJA GAUSSA – PODSUMOWANIE

Jest to proces sprowadzania macierzy do postaci schodkowej. Możemy używać trzech dozwolonych operacji na wierszach:
permutacja wierszy – skalowanie wierszy – dodawanie wiersza do innego wiersza

POSTAĆ SCHODKOWA MACIERZY

- jeśli wiersz jest niezerowy, to pierwszym wyrazem niezerowym (liderem) jest 1
- jeśli wiersz $k$-ty jest niezerowy, to liczba wiodących wyrazów zerowych w wierszu $k+1$ jest większa niż w wierszu $k$-tym,
- jeśli są wiersze zerowe, to są one poniżej wierszy niezerowych.

$$\begin{bmatrix} \mathbf{1} & * & * & * \\ 0 & \mathbf{1} & * & * \\ 0 & 0 & 0 & \mathbf{1} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{1} & * & * & * \\ 0 & \mathbf{1} & * & * \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} \mathbf{1} & * & * & * \\ 0 & \mathbf{1} & * & * \\ 0 & 0 & \mathbf{1} & * \\ 0 & 0 & 0 & \mathbf{1} \end{bmatrix}$$

ZREDUKOWANA POSTAĆ SCHODKOWA MACIERZY

- macierz ma postać schodkową,
- pierwszy niezerowy element w wierszu jest jedynym niezerowym elementem w jego kolumnie.

$$\begin{bmatrix} \mathbf{1} & 0 & * & 0 \\ 0 & \mathbf{1} & * & 0 \\ 0 & 0 & 0 & \mathbf{1} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{1} & 0 & * & * \\ 0 & \mathbf{1} & * & * \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} \mathbf{1} & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & \mathbf{1} \end{bmatrix}$$

RZĄD MACIERZY

$\operatorname{rank} A$ = liczba niezerowych wierszy w postaci schodkowej macierzy $A$ = liczba kolumna zawierających lidera.

$$\operatorname{rank} \begin{bmatrix} \mathbf{1} & 0 & * & 0 \\ 0 & \mathbf{1} & * & 0 \\ 0 & 0 & 0 & \mathbf{1} \end{bmatrix} = 3, \quad \operatorname{rank} \begin{bmatrix} \mathbf{1} & 0 & * & * \\ 0 & \mathbf{1} & * & * \\ 0 & 0 & 0 & 0 \end{bmatrix} = 2, \quad \operatorname{rank} \begin{bmatrix} \mathbf{1} & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & \mathbf{1} \end{bmatrix} = 4$$

Zbiór rozwiązań

Dla $k \times n$-macierzy $A$ oraz $b \in \mathbb{R}^k$ mamy

- układ ma dokładnie jedno rozwiązanie $\Leftrightarrow$ $\operatorname{rank} A = \operatorname{rank} \left[A \mid b\right] = n$
- układ jest sprzeczny $\Leftrightarrow$ $\operatorname{rank} A < \operatorname{rank} \left[A \mid b\right]$
- układ ma nieskończenie wiele rozwiązań $\Leftrightarrow$ $\operatorname{rank} A = \operatorname{rank} \left[A \mid b\right] < n$

## 3.2 Kilka ważnych obserwacji

Dla układu równań (3.1) wprowadzamy następujące oznaczenia: macierz

$$A := \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kn} \end{bmatrix} \tag{3.2}$$

nazywamy *macierzą główną* układu, natomiast macierz

$$\left[\ A \mid b\ \right] := \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ \vdots & & & \vdots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kn} & b_k \end{array}\right],$$

*macierzą rozszerzoną* układu. Macierz $A$ ma $k$ wierszy i $n$ kolumn. Przez $M_{k\times n}(\mathbb{F})$ oznaczamy zbiór wszystkich macierzy $A$ postaci (3.2), gdzie wszystkie *wyrazy* macierzy $a_{ij}$ ($i = 1, \dots k$, $j = 1, \dots, n$) są elementami ciała $\mathbb{F}$.

Kolumny macierzy $A$ oraz prawą stronę układu możemy interpretować jako wektory w $\mathbb{F}^k$:

$$a_1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{k1} \end{bmatrix}, \ a_2 = \begin{bmatrix} a_{12} \\ \vdots \\ a_{k2} \end{bmatrix}, \ \dots, \ a_n = \begin{bmatrix} a_{1n} \\ \vdots \\ a_{kn} \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_k \end{bmatrix}.$$

Będziemy wtedy pisali, że

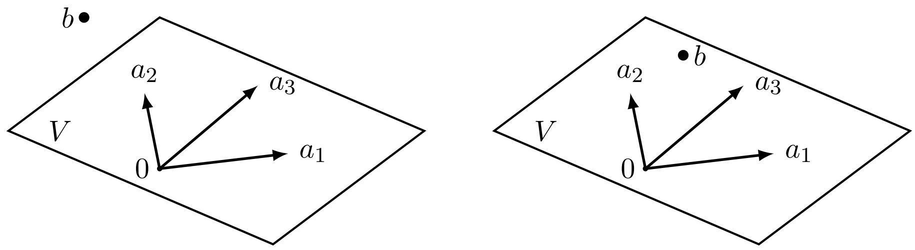
$$A = \left[a_1 \mid \dots \mid a_n\right], \quad \left[A \mid b\right] = \left[a_1 \mid \dots \mid a_n \mid b\right].$$

Zauważmy, że wektor $x = [x_1, \dots, x_n]^T$ jest rozwiązaniem układu (3.1) wtedy i tylko wtedy, gdy

$$x_1 \cdot \begin{bmatrix} a_{11} \\ \vdots \\ a_{k1} \end{bmatrix} + \dots + x_n \cdot \begin{bmatrix} a_{1n} \\ \vdots \\ a_{kn} \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_k \end{bmatrix},$$

czyli

$$x_1 \cdot a_1 + \dots + x_n \cdot a_n = b.$$




Rysunek 3.2: Kolumny macierzy $A = [a_1|a_2|a_3]$ leżą w jednej płaszczyźnie $V = \operatorname{span}\{a_1, a_2, a_3\} \subset \mathbb{R}^3$. Po lewej: $b \notin V$, więc $Ax = b$ nie ma rozwiązań. Po prawej: $b \in V$ i układ ma rozwiązanie. Nie jest ono jednoznaczne. Wektor $b$ jest kombinacją liniową kolumn $a_1$, $a_2$, ale jest również kombinacją kolumn $a_1$ i $a_3$ oraz kolumn $a_2$, $a_3$.

**Konwencje macierzowe zapisu układu równań**

W konwencji wierszy

$$Ax = \begin{bmatrix} - & w_1 & - \\ - & w_2 & - \\ & \vdots & \\ - & w_k & - \end{bmatrix} \begin{bmatrix} | \\ x \\ | \end{bmatrix} = \begin{bmatrix} (w_1^T|x) \\ (w_2^T|x) \\ \vdots \\ (w_k^T|x) \end{bmatrix}$$

W konwencji kolumn

$$Ax = \left[ \ a_1 \ \middle| \ a_2 \ \middle| \ \ldots \ \middle| \ a_n \ \right] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = x_1 \cdot a_1 + \ldots + x_n \cdot a_n.$$

**Układ równań w zapisie macierzowym**

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \vdots \\ a_{k1}x_1 + \ldots + a_{kn}x_n = b_k \end{cases} \Leftrightarrow Ax = b$$

$$\Leftrightarrow \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & & & \vdots \\ a_{k1} & a_{k2} & \ldots & a_{kn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix}$$

$$\Leftrightarrow x_1 \cdot a_1 + \ldots + x_n \cdot a_n = b$$

$$\Leftrightarrow \ b_1 = (w_1^T|x), \ \ldots, \ b_k = (w_k^T|x)$$

Układ równań (3.1) będziemy zapisywać w skrócie jako $Ax = b$. Będziemy mówili, że układ $Ax = b$ ma rozwiązanie, jeśli zbiór jego rozwiązań jest niepusty. Powiemy, że ma on jednoznaczne rozwiązanie, jeśli zbiór jego rozwiązań składa się z jednego wektora.

**Wniosek 3.1.** *Układ równań $Ax = b$ ma rozwiązanie wtedy i tylko wtedy, gdy wektor $b$ jest kombinacją liniową kolumn $a_1, \ldots, a_n$ macierzy $A$.*

Dla wektorów $a_1, \ldots, a_n \in \mathbb{F}^k$ naturalnym więc jest rozważenie zbioru

$$\operatorname{span}\{a_1, \ldots, a_n\} = \Big\{ x_1 \cdot a_1 + \ldots + x_n \cdot a_n : x_1, \ldots, x_n \in \mathbb{F} \Big\}$$

wszystkich możliwych kombinacji liniowych wektorów $a_1, \ldots, a_n$.

**Wniosek 3.2.** *Układ równań liniowych $Ax = b$ ma rozwiązanie wtedy i tylko wtedy, gdy $b \in \operatorname{span}\{a_1, \ldots, a_n\}$.*

**Definicja 3.3** (Układ jednorodny)**.**
Układ równań (3.1) nazywamy *jednorodnym*, gdy

$$b_1 = \ldots = b_k = 0.$$

**Wniosek 3.3.** *Jeżeli układ (3.1) jest jednorodny, to $x_1 = \ldots = x_n = 0$ jest jednym z jego rozwiązań.*

Wektor zerowy $x = 0$ jest **zawsze** jednym z rozwiązań układu **jednorodnego** $Ax = 0$. Ciekawe jest pytanie o to, kiedy układ jednorodny $Ax = 0$ nie ma innych rozwiązań. Kiedy $Ax = 0$ ma jednoznaczne rozwiązanie? Aby tak było, musi zachodzić implikacja: jeśli $x_1 \cdot a_1 + \ldots + x_n \cdot a_n = 0$ dla pewnych $x_i \in \mathbb{F}$, to $x_1 = \ldots = x_n = 0$. Innymi słowy, jeśli kombinacja liniowa wektorów $a_1, \ldots, a_n$ jest wektorem zerowym, to wszystkie jej współczynniki skalarne są równe 0. Oznacza to liniową niezależność wektorów $a_1, \ldots, a_n$.

**Wniosek 3.4.** *Układ jednorodny $Ax = 0$ ma jednoznaczne rozwiązanie wtedy i tylko wtedy, gdy kolumny $a_1, \ldots, a_n$ macierzy $A$ są liniowo niezależne.*

**Wniosek 3.5.** *Układ $Ax = b$ ma co najwyżej jedno rozwiązanie wtedy i tylko wtedy, gdy kolumny $a_1, \ldots, a_n$ macierzy $A$ są liniowo niezależne.*

*Dowód.* Zauważmy, że wektory $v = [v_1, \ldots, v_n]^T, w = [w_1, \ldots, w_n]^T \in \mathbb{F}^n$ są rozwiązaniami układu $Ax = b$ wtedy i tylko wtedy, gdy wektor $v - w$ jest rozwiązaniem układu jednorodnego $Ax = 0$. $\square$

## 3.3 Układy dwóch równań z dwoma niewiadomymi

Zajmiemy się teraz dokładniej układem dwóch równań liniowych (w ciele $\mathbb{R}$) z dwoma niewiadomymi

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \tag{3.3}$$

Skojarzone z nim macierze: główna i rozszerzona to odpowiednio

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \begin{bmatrix} A & b \end{bmatrix} = \left[\begin{array}{cc|c} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{array}\right].$$

Przyglądnijmy się możliwym przypadkom.
**Przypadek 1.** Zakładamy, że

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

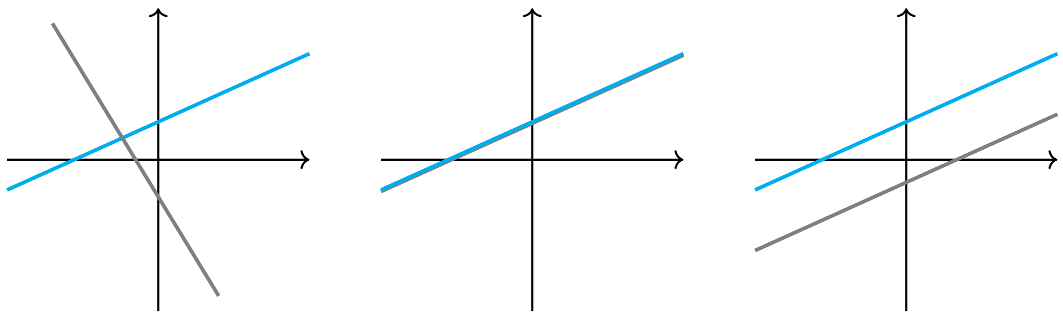
jest macierzą zerową. Układ (3.3) przyjmuje postać

$$\begin{cases} 0 = b_1 \\ 0 = b_2 \end{cases}$$

Jeżeli któraś z liczb $b_1$ lub $b_2$ jest różna od zera, to układ jest sprzeczny, czyli jego zbiór rozwiązań jest zbiorem pustym. Jeśli $b_1 = b_2 = 0$, to zbiorem rozwiązań jest cała płaszczyzna $\mathbb{R}^2$.

**Przypadek 2.** Zakładamy, że któryś z wierszy macierzy $A$, powiedzmy drugi jest zerowy (czyli $a_{21} = a_{22} = 0$), a pierwszy jest niezerowy. Układ (3.3) ma wtedy postać

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ 0 = b_2 \end{cases}$$



Rysunek 3.3: Dwie proste na płaszczyźnie: albo przecinają się w jednym punkcie albo są równe albo nie mają punktów wspólnych.

Jeśli $b_2 \neq 0$, to układ jest sprzeczny. Dla $b_2 = 0$ redukuje się on do jednego równania

$$a_{11}x_1 + a_{12}x_2 = b_1$$

opisującego prostą na płaszczyźnie, bo założyliśmy, że $a_{11}^2 + a_{12}^2 > 0$. W tym przypadku zbiorem rozwiązań układu jest więc prosta na płaszczyźnie.

**Przypadek 3.** Zakładamy, że obydwa wiersze macierzy $A$ są niezerowe. Wtedy obydwa równania opisują proste na płaszczyźnie. Szukamy więc punktów wspólnych dwóch prostych $l_1$ i $l_2$ na płaszczyźnie. Zachodzi jedna z trzech możliwości:

(i) Proste są równe ($l_1 = l_2$). Wtedy zbiorem rozwiązań jest prosta $l_1$.

(ii) Proste $l_1$ i $l_2$ są równoległe, ale różne. Wtedy układ jest sprzeczny.

(iii) Proste $l_1$ i $l_2$ nie są równoległe. Wtedy przecinają się w jednym punkcie. Zbiór rozwiązań układu jest jednoelementowy, czyli układ ma dokładnie jedno rozwiązanie.

**Definicja 3.4** (Wyznacznik)**.**
Wyznacznik macierzy

$$A = \begin{bmatrix} a_1 \mid a_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in M_{2\times 2}(\mathbb{F})$$

definiujemy jako skalar

$$\det(A) := a_{11}a_{22} - a_{12}a_{21} \in \mathbb{F}.$$

U Dla niezerowego skalara $a \in \mathbb{F}$ przez $\frac{1}{a}$ będziemy oznaczać element odwrotny $a^{-1}$ do $a$ względem mnożenia w ciele $\mathbb{F}$.

**Twierdzenie 3.1** (Wzory Cramera)**.** *Układ równań (3.3) ma dokładnie jedno rozwiązanie wtedy i tylko wtedy, gdy* $\det(A) \neq 0$. *Jest ono wtedy dane wzorami Cramera*

$$x_1 = \frac{\det \begin{bmatrix} b \mid a_2 \end{bmatrix}}{\det \begin{bmatrix} a_1 \mid a_2 \end{bmatrix}}, \quad x_2 = \frac{\det \begin{bmatrix} a_1 \mid b \end{bmatrix}}{\det \begin{bmatrix} a_1 \mid a_2 \end{bmatrix}}.$$

*Dowód.* Załóżmy, że $\det(A) = a_{11}a_{22} - a_{12}a_{21} \neq 0$. Wtedy $a_{11} \neq 0$ lub $a_{21} \neq 0$ tzn. wektor $a_1$ nie może być zerowy. Załóżmy, że $a_{11} \neq 0$. Dowód w przypadku $a_{21} \neq 0$ jest analogiczny. Z pierwszego równania $x_1 = \dfrac{b_1 - a_{12}x_2}{a_{11}}$. Po podstawieniu do drugiego równania otrzymujemy równoważny z (3.3) układ równań

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ (a_{11}a_{22} - a_{12}a_{21})x_2 = b_1a_{21} - b_2a_{11} \end{cases} \tag{3.4}$$

Jego rozwiązaniem jest jeden punkt

$$x_1 = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{12}a_{21}}, \quad x_2 = \frac{b_1a_{21} - b_2a_{11}}{a_{11}a_{22} - a_{12}a_{21}},$$

co kończy dowód tej implikacji.

Zakończymy dowód pokazując, że jeżeli $\det(A) = 0$, to układ (3.3) jest albo sprzeczny albo jego zbiór rozwiązań jest nieskończony. Zachodzi jedna z dwóch możliwości

(a) $a_1 = [0, 0]^T$;

(b) $a_1 \neq [0, 0]^T$.

W przypadku (a) układ (3.3) przyjmuje postać

$$\begin{cases} a_{12}x_2 = b_1 \\ a_{22}x_2 = b_2 \end{cases}$$

Jest on albo sprzeczny (nie istnieje rozwiązanie $x_2$) albo ma nieskończenie wiele rozwiązań (istnieje rozwiązanie $x_2$ i wtedy $x_1$ jest dowolną liczbą rzeczywistą).

W przypadku ($b$), jeśli przykładowo gdy $a_{11} \neq 0$, to na mocy poprzedniej części dowodu układ (3.3) jest równoważny z układem (12.3), czyli

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ 0 = b_1a_{21} - b_2a_{11} \end{cases}$$

Jest on albo sprzeczny, gdy $b_1a_{21} - b_2a_{11} \neq 0$ albo ma nieskończenie wiele rozwiązań spełniających $a_{11}x_1 + a_{12}x_2 = b_1$, gdy $b_1a_{21} - b_2a_{11} = 0$. □

**Twierdzenie 3.2.** *Wektory $a_1, a_2 \in \mathbb{F}^2$ są liniowo niezależne wtedy i tylko wtedy, gdy*

$$\det \left[a_1 \mid a_2\right] \neq 0.$$

*Wówczas dla dowolnego $b \in \mathbb{F}^2$ istnieją jednoznacznie wyznaczone takie skalary $x_1, x_2 \in \mathbb{F}$, że*

$$x_1 \cdot a_1 + x_2 \cdot a_2 = b.$$

*Dowód.* Liniowa niezależność wektorów $a_1$ i $a_2$ jest z definicji równoważna z faktem, że równość $x_1 \cdot a_1 + x_2 \cdot a_2 = 0$ implikuje, że $x_1 = x_2 = 0$, czyli układ równań

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = 0 \\ a_{21}x_1 + a_{22}x_2 = 0 \end{cases}$$

ma dokładnie jedno rozwiązanie $x_1 = x_2 = 0$. Ponieważ $x_1 = x_2 = 0$ jest pewnym rozwiązaniem tego układu, więc teza wynika z twierdzenia 3.1. □

**Wniosek 3.6.** *Niech $a_1, a_2, a_3 \in \mathbb{F}^2$. Wtedy wektory $a_1, a_2, a_3$ są liniowo zależne.*

*Dowód.* Przeprowadzimy dowód niewprost. Przypuśćmy, że wektory $a_1$, $a_2$, $a_3$ są liniowo niezależne. Z definicji liniowej niezależności wynika, że wtedy również wektory $a_1, a_2$ są liniowo niezależne. Z twierdzenia 3.2 istnieją takie skalary $x_1, x_2 \in \mathbb{F}$, że

$$x_1 \cdot a_1 + x_2 \cdot a_2 = a_3,$$

co prowadzi do sprzeczności z liniową niezależnością wektorów $a_1$, $a_2$, $a_3$, bo

$$x_1 \cdot a_1 + x_2 \cdot a_2 + (-1) \cdot a_3 = 0.$$ □

U Podamy pewną geometryczną interpretację wyznacznika na płaszczyźnie $\mathbb{R}^2$. Załóżmy, że wektory $e_1 = [1, 0]^T$ i $u = [u_1, u_2]^T$ są liniowo niezależne. Wtedy $\det \left[e_1 \mid u\right] = u_2 \neq 0$. Zauważmy, że

- $\det \begin{bmatrix} e_1 \mid u \end{bmatrix} > 0$ wtedy i tylko wtedy, gdy $u_2 > 0$ wtedy i tylko wtedy, gdy najkrótszy kąt obrotu od wektora $e_1$ do wektora $u$ jest skierowany przeciwnie do ruchu wskazówek zegara;
- $\det \begin{bmatrix} e_1 \mid u \end{bmatrix} < 0$ wtedy i tylko wtedy, gdy $u_2 < 0$ wtedy i tylko wtedy, gdy najkrótszy kąt obrotu od wektora $e_1$ do wektora $u$ jest skierowany zgodnie z ruchem wskazówek zegara.

Po co nam wyznacznik?

Dla macierzy

$$A = \begin{bmatrix} a_1 \mid a_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

poniższe warunki są równoważne:

- $\det A \neq 0$,
- kolumny $a_1$, $a_2$ są liniowo niezależne,
- każdy wektor $b \in \mathbb{R}^2$ można jednoznacznie zapisać w postaci kombinacji liniowej $b = x_1 a_1 + x_2 a_2$,
- dla każdego wektora $b \in \mathbb{R}^2$ układ równań $Ax = b$ ma dokładnie jedno rozwiązanie.

# Rozdział 4

# Baza i wymiar

| |
|---|
| Słowa kluczowe rozdziału<br>generowanie ♢ liniowa niezależność ♢ baza ♢ wymiar ♢ postać schodkowa ♢ podprzestrzeń span$\{a_1, \ldots, a_k\}$ ♢ przestrzeń wektorowa $M_{m\times n}(\mathbb{F})$ ♢ podprzestrzeń wektorowa |

## 4.1 Generowanie i liniowa niezależność

**Definicja 4.1** (Podprzestrzeń generowana przez wektory)**.**
Niech $(V, +, \cdot)$ będzie przestrzenią wektorową nad ciałem $\mathbb{F}$. *Kombinacją liniową wektorów* $v_1, \ldots, v_k \in V$ nazywamy dowolny wektor $v$ postaci

$$v = x_1 \cdot v_1 + \ldots + x_k \cdot v_k, \quad x_1, \ldots, x_k \in \mathbb{F}.$$

Zbiór span $\{v_1, \ldots, v_k\}$ oznacza zbiór wszystkich wektorów, będących kombinacjami liniowymi wektorów $v_1, \ldots, v_k$, czyli

$$\operatorname{span}\{v_1, \ldots, v_k\} = \{x_1 \cdot v_1 + \ldots + x_k \cdot v_k : x_1, \ldots, x_k \in \mathbb{F}\}.$$

Nazywamy go *podprzestrzenią generowaną przez wektory* $v_1, \ldots, v_k$. Powiemy, że wektory $v_1, \ldots, v_k$ generują przestrzeń wektorową $V$, jeśli

$$\operatorname{span}\{v_1, \ldots, v_k\} = V.$$

**Przykład 4.1.** Niech $v = [0, 1, 1]^T, u = [1, 1, 1]^T \in \mathbb{Z}_2^3$. Wyznaczymy span $\{v, u\} \subset \mathbb{Z}_2^3$. Mamy 4 możliwe kombinacje liniowe:

- $0 \cdot u + 0 \cdot v = 0,$
- $1 \cdot u + 0 \cdot v = u,$
- $0 \cdot u + 1 \cdot v = v,$
- $1 \cdot u + 1 \cdot v = [0, 1, 1]^T + [1, 1, 1]^T = [1, 0, 0]^T.$

Stąd

$$\operatorname{span}\{v, u\} = \{0, v, u, [1, 0, 0]^T\}.$$

**Definicja 4.2** (Liniowa niezależność wektorów)**.** Wektory $v_1, \ldots, v_k \in V$ są *liniowo niezależne*, jeśli dla skalarów $x_1, \ldots, x_k \in \mathbb{F}$ zachodzi implikacja

$$x_1 \cdot v_1 + \ldots + x_k \cdot v_k = 0 \Longrightarrow x_1 = \ldots = x_k = 0.$$

Oznacza to, że jeśli kombinacja liniowa wektorów $v_1, \ldots, v_k$ jest wektorem zerowym, to wszystkie jej skalarne współczynniki są równe zero.

**Przykład 4.2.** Rozważmy przestrzeń wektorową $V$ wszystkich funkcji $f : \mathbb{R} \to \mathbb{R}$. Pokażemy, że funkcje $f = \sin(x)$ i $g = \cos(x)$ są liniowo niezależne. Przypuśćmy, że $x_1 \cdot f + x_2 \cdot g = 0$ dla pewnych $x_1, x_2 \in \mathbb{R}$. Oznacza to, że dla dowolnego $x \in \mathbb{R}$ mamy

$$x_1 \cdot \sin(x) + x_2 \cdot \cos x = 0.$$

Podstawiając w tej równości $x = 0$, otrzymujemy, że $x_2 = 0$. Wstawiając $x = \pi/2$ dostajemy, że $x_1 = 0$.

**Przykład 4.3.** Pokażemy, że wektory

$$v = [1, 1, 3]^T, \quad u = [2, 0, 2]^T, \quad w = [4, 3, 0]^T \in \mathbb{Z}_5^3$$

są liniowo zależne w $\mathbb{Z}_5^3$. Zauważmy, że

$$v + u = [3, 1, 0]^T$$

oraz

$$3 \cdot (v + u) = [4, 3, 0]^T = w,$$

czyli $w$ jest kombinacją liniową wektorów $v, u$.

Z drugiej strony, wektory $v, u, w \in \mathbb{R}^3$ są liniowo niezależne nad $\mathbb{R}$. Rzeczywiście, równość

$$x \cdot v + y \cdot u + z \cdot w = 0, \quad x, y, z \in \mathbb{R}$$

oznacza, że

$$x + 2y + 4z = x + 3z = 3x + 2y = 0,$$

więc $x = y = z = 0$.

**Definicja 4.3** (Baza przestrzeni wektorowej)**.**

Niech $V$ będzie przestrzenią wektorową nad ciałem $\mathbb{F}$. Wektory $v_1, \ldots, v_n \in V$ stanowią bazę przestrzeni wektorowej $V$, jeśli spełnione są warunki:

(B1) wektory $v_1, \ldots, v_n$ *generują* przestrzeń $V$, czyli

$$\operatorname{span}\{v_1, \ldots, v_n\} = V,$$

(B2) wektory $v_1, \ldots, v_n$ są *liniowo niezależne.*

Warunek (B1) oznacza, że każdy wektor $v \in V$ można przedstawić jako kombinację liniową wektorów $v_1, \ldots, v_n$, a warunek (B2) gwarantuje, że takie przedstawienie jest jednoznaczne.

**Przykład 4.4.** Wektory

$$e_1 = [1, 0, \ldots, 0]^T,\ e_2 = [0, 1, 0, \ldots, 0]^T, \ldots,\ e_n = [0, \ldots, 0, 1]^T$$

stanowią bazę przestrzeni wektorowej $\mathbb{F}^n$. Rzeczywiście, są one liniowo niezależne, bo jeśli

$$\begin{aligned}[0, \ldots, 0]^T &= x_1 \cdot e_1 + \ldots + x_n \cdot e_n \\ &= [x_1, \ldots, x_n]^T,\end{aligned}$$

to $x_1 = \ldots = x_n = 0$. Sprawdzimy warunek (B2). Niech $v = [x_1, \ldots, x_n]^T \in \mathbb{F}^n$ będzie dowolnym wektorem. Wtedy

$$\begin{aligned}v &= [x_1, \ldots, x_n]^T \\ &= [x_1, 0, \ldots, 0]^T + \ldots + [0, \ldots, 0, x_n]^T \\ &= x_1 \cdot e_1 + \ldots + x_n \cdot e_n,\end{aligned}$$

czyli $v \in \operatorname{span}\{e_1, \ldots, e_n\}$.

**Wniosek 4.1.** *Wektory $v_1, \ldots, v_n$ stanowią bazę dla $V$ wtedy i tylko wtedy, gdy dla każdego wektora $v \in V$ istnieją jednoznacznie wyznaczone takie skalary $x_1, \ldots, x_k \in \mathbb{F}$, że*

$$v = x_1 \cdot v_1 + \ldots + x_n \cdot v_n.$$

*Dowód.* Załóżmy najpierw, że $v_1, \ldots, v_n$ tworzą bazę $V$. Niech $v \in V$ będzie dowolnym wektorem. Z warunku (B2) wynika, że istnieją takie skalary $x_1, \ldots, x_n \in \mathbb{F}$, że $v = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$. Musimy pokazać, że są one wyznaczone jednoznacznie. Przypuśćmy więc, że

$$v = y_1 \cdot v_1 + \ldots + y_n \cdot v_n,$$

gdzie $y_i \in \mathbb{F}$. Wtedy

$$\begin{aligned} 0 &= v - v \\ &= (x_1 \cdot v_1 + \ldots + x_n \cdot v_n) - (y_1 \cdot v_1 + \ldots + y_n \cdot v_n) \\ &= (x_1 - y_1) \cdot v_1 + \ldots + (x_n - y_n) \cdot v_n. \end{aligned}$$

Ponieważ wektory $v_1, \ldots, v_n$ są liniowo niezależne, więc

$$x_1 - y_1 = \ldots = x_n - y_n = 0,$$

co kończy dowód pierwszej implikacji.

Załóżmy teraz, że dla każdego wektora $v \in V$ istnieją jednoznacznie wyznaczone takie skalary $x_1, \ldots, x_k \in \mathbb{F}$, że

$$v = x_1 \cdot v_1 + \ldots + x_n \cdot v_n.$$

Wtedy oczywiście zachodzi warunek (B2). Pokażemy liniową niezależność wektorów $v_1, \ldots, v_n$. Przypuśćmy, że $x_1 \cdot v_1 + \ldots + x_n \cdot v_n = 0$ dla pewnych skalarów $x_1, \ldots, x_n \in \mathbb{F}$. Ponieważ

$$0 \cdot x_1 + \ldots + 0 \cdot x_n = 0 = x_1 \cdot v_1 + \ldots + x_n \cdot v_n,$$

więc z założonej jednoznaczności skalarów wynika, że $x_1 = \ldots = x_n = 0$. $\square$

**Lemat 4.1.** *Niech $v_1, \ldots, v_n \in V$. Jeśli $v \in \operatorname{span}\{v_1, \ldots, v_n\}$, to*

$$\operatorname{span}\{v_1, \ldots, v_n\} = \operatorname{span}\{v_1, \ldots, v_n, v\}.$$

*Dowód.* Oczywiście $\operatorname{span}\{v_1, \ldots, v_n\} \subset \operatorname{span}\{v_1, \ldots, v_n, v\}$. Niech $w \in \operatorname{span}\{v_1, \ldots, v_n, v\}$. Wtedy

$$w = x_1 \cdot v_1 + \ldots + x_n \cdot v_n + x \cdot v,$$

dla pewnych skalarów $x_1, \ldots, x_n, x \in \mathbb{F}$. Ponieważ $v \in \operatorname{span}\{v_1, \ldots, v_n\}$, więc $v = y_1 \cdot v_1 + \ldots + y_n \cdot v_n$ dla pewnych skalarów $y_1, \ldots, y_n \in \mathbb{F}$. Stąd

$$w = (x_1 + xy_1) \cdot v_1 + \ldots + (x_n + xy_n) \cdot v_n \in \operatorname{span}\{v_1, \ldots, v_n\}. \qquad \square$$

**Lemat 4.2.** *Załóżmy, że wektory $v_1, \ldots, v_k \in V$ są liniowo niezależne. Jeśli $v_{k+1} \notin \operatorname{span}\{v_1, \ldots, v_k\}$, to*

$$v_1, \ldots, v_k,\ v_{k+1}$$

*są liniowo niezależne.*

*Dowód.* Załóżmy, że

$$x_1 \cdot v_1 + \ldots + x_k \cdot v_k + x_{k+1} \cdot v_{k+1} = 0$$

dla pewnych skalarów $x_i \in \mathbb{F}$. Zauważmy, że $x_{k+1} = 0$, bo w przeciwnym razie

$$v_{k+1} = -\frac{1}{x_{k+1}}(x_1 \cdot v_1 + \ldots + x_k \cdot v_k) \in \operatorname{span}(v_1, \ldots, v_k).$$

W konsekwencji, również $x_1 = \ldots = x_k = 0$, bo $v_1, \ldots, v_k \in V$ są liniowo niezależne. $\square$

**Twierdzenie 4.1.** *Załóżmy, że* $\operatorname{span}\{v_1, \ldots, v_k\} = V \neq 0$. *Wtedy istnieje*

$$B \subset \{v_1, \ldots, v_k\}$$

*baza dla* $V$.

*Dowód.* Bez straty ogólności, na podstawie lematu 4.1, możemy założyć, że wektory $v_i$ są niezerowe. Jeżeli $\operatorname{span}\{v_1\} = V$, to dowód jest zakończony. Jeśli $\operatorname{span}\{v_1\} \neq V$, to któryś z wektorów $v_2, \ldots, v_k$ nie należy do $\operatorname{span}\{v_1\}$. Powiedzmy, że jest to $v_2$. Wtedy $v_1, v_2$ są liniowo niezależne z lematu 4.2. Możemy ten proces kontynuować, rozważając teraz $\operatorname{span}\{v_1, v_2\}$. W skończonej liczbie kroków wybierzemy bazę $B$ ze zbioru $\{v_1, \ldots, v_k\}$. □

**Twierdzenie 4.2** (Steinitz)**.** *Załóżmy, że* $V = \operatorname{span}\{v_1, \ldots, v_n\}$ *dla pewnych wektorów* $v_1, \ldots, v_n \in V$ $(n \geqslant 1)$. *Jeśli wektory* $w_1, \ldots, w_m \in V$ *są liniowo niezależne, to*

- $m \leqslant n$
- *po ewentualnym przenumerowaniu wektorów* $v_1, \ldots, v_n$ *mamy*

$$\operatorname{span}\{w_1, \ldots, w_m, v_{m+1}, \ldots, v_n\} = V.$$

*Dowód.* Zastosujemy indukcję względem $k \in \{1, \ldots, m\}$. Niech $k = 1$. Oczywiście $n \geqslant 1$. Rozważmy wektor $w_1$. Ponieważ $\operatorname{span}\{v_1, \ldots, v_n\} = V$, więc

$$w_1 = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$$

dla pewnych skalarów $x_1, \ldots, x_n \in \mathbb{F}$. Ponieważ z założenia $w_1 \neq 0$, więc któryś ze skalarów $x_i$ jest różny od zera. Przenumerowując wektory $v_i$ i skalary $x_i$ możemy założyć, że $x_1 \neq 0$. Uzasadnimy, że $\operatorname{span}\{w_1, v_2, \ldots, v_n\} = V$. Ponieważ $w_1 \in V = \operatorname{span}\{v_1, \ldots, v_n\}$, więc z lematu 4.1 wynika, że

$$\operatorname{span}\{v_1, \ldots, v_n\} = \operatorname{span}\{w_1, v_1, \ldots, v_n\}.$$

Z drugiej strony $x_1 \neq 0$, więc $v_1 = \frac{1}{x_1}(w_1 - x_2 \cdot v_2 - \ldots - x_n \cdot v_n) \in \operatorname{span}\{w_1, v_2, \ldots, v_n\}$, czyli

$$\operatorname{span}\{w_1, v_2, \ldots, v_n\} = \operatorname{span}\{w_1, v_1, v_2, \ldots, v_n\} = \operatorname{span}\{v_1, v_2, \ldots, v_n\} = V.$$

Zakładamy teraz, że teza zachodzi dla pewnego $1 \leqslant k < m$. Pokażemy, że zachodzi dla $k+1$. Z założenia indukcyjnego $k \leqslant n$ oraz

$$\operatorname{span}\{w_1, \ldots, w_k, v_{k+1}, \ldots, v_n\} = V.$$

Istnieją więc takie skalary $c_1, \ldots, c_n \in \mathbb{F}$, że

$$w_{k+1} = c_1 \cdot w_1 + \ldots + c_k \cdot w_k + c_{k+1} \cdot v_{k+1} + \ldots + c_n \cdot v_n.$$

Zauważmy, że istnieje takie $k < i_0 \leqslant n$, że $c_{i_0} \neq 0$, bo w przeciwnym razie $w_{k+1} = c_1 \cdot w_1 + \ldots + c_k \cdot w_k$ co jest sprzeczne z ich liniową zależnością. Po ewentualnym przenumerowaniu możemy założyć, że $i_0 = k+1$. W szczególności $k+1 \leqslant n$. Mamy

$$v_{k+1} = \frac{1}{c_{k+1}}(w_{k+1} - \sum_{i=1}^{k} c_i \cdot w_i - \sum_{i=k+2} c_i \cdot v_i).$$

Ponieważ $v_{k+1} \in \operatorname{span}\{w_1, \ldots, w_{k+1}, v_{k+2}, \ldots, v_n\}$, więc

$$\operatorname{span}\{w_1, \ldots, w_{k+1}, v_{k+2}, \ldots, v_n\} = \operatorname{span}\{w_1, \ldots, w_{k+1}, v_{k+1} v_{k+2}, \ldots, v_n\} =$$

$$\operatorname{span}\{w_1, \ldots, w_k, v_{k+1} v_{k+2}, \ldots, v_n\} = V.$$ □

**Twierdzenie 4.3.** *Załóżmy, że wektory* $v_1, \ldots, v_n$ *stanowią bazę przestrzeni wektorowej* $V$. *Jeśli wektory* $w_1, \ldots, w_m \in V$ *są liniowo niezależne, to* $m \leqslant n$.

*Dowód.* Przypuśćmy, że $m > n$. Pokażemy, że wektory $w_1, \ldots, w_n$ są bazą $V$. Prowadzi to do sprzeczności, bo wtedy $w_{n+1}$ jest kombinacją liniową wektorów $w_1, \ldots, w_n$, a z założenia wektory $w_1, \ldots, w_n, w_{n+1}$ są liniowo niezależne.

Rozważmy wektor $w_1$. Ponieważ $v_1, \ldots, v_n$ są bazą $V$, więc

$$w_1 = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$$

dla jednoznacznie wyznaczonych skalarów $x_1, \ldots, x_n \in \mathbb{F}$. Ponieważ z założenia $w_1 \neq 0$, więc któryś ze skalarów $x_i$ jest różny od zera. Przenumerowując wektory $v_i$ i skalary $x_i$ możemy założyć, że $x_1 \neq 0$. Uzasadnimy, że $w_1, v_2, \ldots, v_n$ są bazą dla $V$. Przypuśćmy, że

$$\lambda_1 \cdot w_1 + \lambda_2 \cdot v_2 + \ldots + \lambda_n \cdot v_n = 0,$$

dla pewnych skalarów $\lambda_i$. Jeśli $\lambda_1 = 0$, to również $\lambda_2 = \ldots = \lambda_n = 0$ bo wektory $v_2, \ldots, v_n$ są liniowo niezależne. Jeśli $\lambda_1 \neq 0$, to

$$w_1 = -\frac{1}{\lambda_1}(\lambda_2 \cdot v_2 + \ldots + \lambda_n \cdot v_n),$$

co jest sprzeczne z jednoznacznością zapisu $w_1$ w bazie $v_1, \ldots, v_n$, bo $x_1 \neq 0$.

Pokażemy teraz, że $w_1, v_2, \ldots, v_n$ generują $V$. Ponieważ $w_1 \in V = \operatorname{span}\{v_1, \ldots, v_n\}$, więc z lematu 4.1 wynika, że

$$\operatorname{span}\{v_1, \ldots, v_n\} = \operatorname{span}\{w_1, v_1, \ldots, v_n\}.$$

Z drugiej strony $v_1 \in \operatorname{span}\{w_1, v_2, \ldots, v_n\}$, więc

$$\operatorname{span}\{w_1, v_2, \ldots, v_n\} = \operatorname{span}\{w_1, v_1, v_2, \ldots, v_n\}.$$

Uzasadnimy, że możemy ten proces iterować. Przypuśćmy, że dla pewnego $1 \leqslant k < n$ wektory

$$w_1, \ldots, w_k,\ v_{k+1}, \ldots, v_n$$

tworzą bazę $V$. Istnieją więc jedyne takie skalary $c_1, \ldots, c_n \in \mathbb{F}$, że

$$w_{k+1} = c_1 \cdot w_1 + \ldots + c_k \cdot w_k + c_{k+1} \cdot v_{k+1} + \ldots + c_n \cdot v_n.$$

Zauważmy, że istnieje takie $k < i_0 \leqslant n$, że $c_{i_0} \neq 0$, bo w przeciwnym razie $w_{k+1} = c_1 \cdot w_1 + \ldots + c_k \cdot w_k$ co jest sprzeczne z liniową zależnością. Z dokładnością do permutacji możemy założyć, że $c_{k+1} \neq 0$ i możemy podmienić wektor $v_{k+1}$ na $w_{k+1}$, otrzymując bazę

$$w_1, \ldots, w_k, w_{k+1},\ v_{k+2}, \ldots, v_n.$$ □

**Wniosek 4.2.** *Załóżmy, że $v_1, \ldots, v_n$ jest bazą $V$. Wtedy każda baza w $V$ ma $n$ elementów.*

**Definicja 4.4** (Przestrzeń skończenie wymiarowa)**.**

Niech $V$ będzie przestrzenią wektorową nad ciałem $\mathbb{F}$. Powiemy, że przestrzeń $V$ jest *skończenie wymiarowa*, jeśli posiada bazę skończoną. Wymiarem przestrzeni skończenie wymiarowej $V$ nazywamy liczbę elementów jej dowolnej bazy i oznaczamy przez $\dim_{\mathbb{F}} V$ lub $\dim V$. Przyjmujemy z definicji, że $\dim V = 0$, gdy $V = \{0\}$. Jeśli przestrzeń $V$ nie jest skończenie wymiarowa, to mówimy, że jest nieskończenie wymiarowa.

U

Istnienie bazy skończonej $v_1, \ldots, v_n$ w przestrzeni wektorowej $V$ bardzo ułatwia analizę jej własności, bo pozwala ograniczyć się do rozważania kombinacji liniowych $x_1 \cdot v_1 + \ldots + x_n \cdot v_n$ skończonej liczby wektorów. Z drugiej strony musimy być świadomi w jaki sposób, jeśli w ogóle, rozważane przez nas pojęcia lub przeprowadzane rozumowania zależą od wyboru bazy w przestrzeni $V$. Przykładowo definiując wymiar przestrzeni wektorowej musieliśmy uzasadnić, że wszystkie bazy mają tyle samo elementów.

**Wniosek 4.3.** $\dim_{\mathbb{F}} \mathbb{F}^n = n$ *dla* $n \in \mathbb{N}$.

*Dowód.* Wektory $e_1, \ldots, e_n$ stanowią bazę dla $\mathbb{F}^n$ złożoną z $n$-wektorów. $\square$

**Przykład 4.5.** W przestrzeni wektorowej $V = \mathbb{Z}_2$ nad $\mathbb{Z}_2$ istnieje tylko jedna baza, bo w $\mathbb{Z}_2 = \{0, 1\}$ mamy tylko jeden niezerowy wektor.

**Przykład 4.6** (Przestrzeń $\mathbb{C}^n$ nad ciałami $\mathbb{C}$ i $\mathbb{R}$)**.** Rozważmy przestrzenie wektorowe $(\mathbb{C}^n, +, \cdot_{\mathbb{C}})$ oraz $(\mathbb{C}^n, +, \cdot_{\mathbb{R}})$. Pierwsza jest przestrzenią zespoloną oraz

$$\dim(\mathbb{C}^n, +, \cdot_{\mathbb{C}}) = n.$$

Bazą dla $(\mathbb{C}^2, +, \cdot_{\mathbb{C}})$ są przykładowo wektory $e_1 = [1, 0]^T, e_2 = [0, 1]^T \in \mathbb{C}^2$. Druga przestrzeń, $(\mathbb{C}^n, +, \cdot_{\mathbb{R}})$ jest przestrzenią rzeczywistą oraz

$$\dim(\mathbb{C}^n, +, \cdot_{\mathbb{R}}) = 2n.$$

Bazą dla $(\mathbb{C}^2, +, \cdot_{\mathbb{R}})$ są na przykład wektory:

$$e_1, \quad e_2, \quad [i, 0]^T, \quad [0, i]^T.$$

Rzeczywiście, dla $[a + ib, c + id]^T$ i $a, b, c, d \in \mathbb{R}$ mamy

$$[a + ib, c + id]^T = a \cdot [1, 0]^T + c \cdot [0, 1]^T + b \cdot [i, 0]^T + d \cdot [0, i]^T,$$

i to przedstawienie jest jednoznaczne.

**Przykład 4.7** (Przestrzeń wektorowa macierzy $M_{k\times n}(\mathbb{F})$)**.** Wprowadzimy strukturę przestrzeni wektorowej w zbiorze $M_{k\times n}(\mathbb{F})$. W tym celu określimy działania

$$+ : M_{k\times n}(\mathbb{F}) \times M_{k\times n}(\mathbb{F}) \longrightarrow M_{k\times n}(\mathbb{F})$$

oraz

$$\cdot : \mathbb{F} \times M_{k\times n}(\mathbb{F}) \longrightarrow M_{k\times n}(\mathbb{F}).$$

Dla $A, B \in M_{k\times n}(\mathbb{F})$ i $t \in \mathbb{F}$ definiujemy

$$\underbrace{\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{k1} & \ldots & a_{kn} \end{bmatrix}}_{A} + \underbrace{\begin{bmatrix} b_{11} & \ldots & b_{1n} \\ \vdots & & \vdots \\ b_{k1} & \ldots & b_{kn} \end{bmatrix}}_{B} = \underbrace{\begin{bmatrix} a_{11} + b_{11} & \ldots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{k1} + b_{k1} & \ldots & a_{kn} + b_{kn} \end{bmatrix}}_{A+B},$$

$$t \cdot \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & & & \vdots \\ a_{k1} & a_{k2} & \ldots & a_{kn} \end{bmatrix} = \begin{bmatrix} t \cdot a_{11} & t \cdot a_{12} & \ldots & t \cdot a_{1n} \\ \vdots & & & \vdots \\ t \cdot a_{k1} & t \cdot a_{k2} & \ldots & t \cdot a_{kn} \end{bmatrix}.$$

Sprawdzamy łatwo, że trójka $(M_{k\times n}(\mathbb{F}), +, \cdot)$ jest przestrzenią wektorową nad ciałem $\mathbb{F}$. Znajdziemy bazę dla przestrzeni $(M_{2\times 2}(\mathbb{F}), +, \cdot)$. Zauważmy, że dla dowolnej macierzy $A \in M_{2\times 2}(\mathbb{F})$ mamy

$$\begin{aligned} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} &= \begin{bmatrix} a_{11} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & a_{12} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ a_{21} & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & a_{22} \end{bmatrix} \\ &= a_{11} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a_{12} \cdot \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + a_{21} \cdot \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + a_{22} \cdot \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

Ponieważ współczynniki $a_{ij}$ są w tym przedstawieniu wyznaczone jednoznacznie przez macierz $A$, więc macierze

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

są bazą $M_{2\times 2}(\mathbb{F})$, więc $\dim_{\mathbb{F}} M_{2\times 2}(\mathbb{F}) = 4$.

Z powyższego wynika, że $\dim_{\mathbb{C}} M_{2\times 2}(\mathbb{C}) = 4$. Przestrzeń $M_{2\times 2}(\mathbb{C})$ możemy również traktować jako przestrzeń wektorową nad $\mathbb{R}$. Zastanówmy się jaki jest wymiar $\dim_{\mathbb{R}} M_{2\times 2}(\mathbb{C})$. Każdy ze współczynników macierzy $A$ jest liczbą zespoloną. Macierz $A$ możemy teraz mnożyć tylko przez liczby rzeczywiste. Mamy

$$\begin{aligned} A &= \begin{bmatrix} a_{11}+ib_{11} & a_{12}+ib_{12} \\ a_{21}+ib_{21} & a_{22}+ib_{22} \end{bmatrix} = \begin{bmatrix} a_{11}+ib_{11} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & a_{12}+ib_{12} \\ 0 & 0 \end{bmatrix} \\ &+ \begin{bmatrix} 0 & 0 \\ a_{21}+ib_{21} & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & a_{22}+ib_{22} \end{bmatrix} \\ &= a_{11}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a_{12}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + a_{21}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + a_{22}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &+ b_{11}\begin{bmatrix} i & 0 \\ 0 & 0 \end{bmatrix} + b_{12}\begin{bmatrix} 0 & i \\ 0 & 0 \end{bmatrix} + b_{21}\begin{bmatrix} 0 & 0 \\ i & 0 \end{bmatrix} + b_{22}\begin{bmatrix} 0 & 0 \\ 0 & i \end{bmatrix}. \end{aligned}$$

Wynika stąd, że $\dim_{\mathbb{R}} M_{2\times 2}(\mathbb{C}) = 8$.

**Przykład 4.8.** Pokażemy, że rozważanie przestrzeni wektorowych nad $\mathbb{Z}_2$ może być użyteczne w rozwiązywaniu problemów kombinatorycznych.

W pewnym mieście mieszka $n$ osób i działa $m$ klubów dyskusyjnych. Każda z nich ma nieparzystą liczbę członków oraz każde dwa kluby mają parzystą liczbę wspólnych członków. Pokażemy, że $m \leqslant n$.

Dla $i = 1,\ldots,m$, przez $v^i = [v^i_1,\ldots,v^i_n]^T \in \mathbb{Z}_2^n$ oznaczamy wektor przynależności mieszkańców do klubu $i$. Oznacza to, że $v^i_j = 1$, gdy $j$-ty mieszkaniec należy do klubu $i$. W przeciwnym razie $v^i_j = 0$. Wtedy $(v^i|v^i) = \sum_{j=1}(v^i_j)^2$ jest liczbą członków klubu $i$. Wiemy, że jest to liczba nieparzysta. W $\mathbb{Z}_2$ mamy więc, że $(v^i|v^i) = 1$. Analogicznie dla $i \neq j$, $(v^i|v^j)$ jest liczbą osób będących w klubach $i$ oraz $j$. Z założenia wynika, że $(v^i|v^j) = 0$ dla $i \neq j$. Wystarczy pokazać, że $v_1,\ldots,v_m$ są liniowo niezależne w $\mathbb{Z}_2^n$. Przypuśćmy, że dla pewnych $c_1,\ldots,c_m \in \mathbb{Z}_2$ mamy

$$c_1 \cdot v_1 + \ldots + c_m \cdot v_m = 0.$$

Ponieważ $(v_i|v_j) = \delta_{ij}$, więc $c_1 = \ldots = c_m = 0$.

**Lemat 4.3.** *Niech $V$ będzie przestrzenią wektorową. Następujące warunki są równoważne*

*(i) $V$ jest nieskończenie wymiarowa,*

*(ii) istnieje taki nieskończony ciąg wektorów $v_1, v_2, \ldots$, że dla każdego $n \in \mathbb{N}$ wektory $v_1,\ldots,v_n$ są liniowo niezależne.*

*Dowód.* Załóżmy, że $V$ jest nieskończenie wymiarowa. Niech $v_1 \in V$ będzie niezerowym wektorem. Ponieważ $\operatorname{span}\{v_1\} \neq V$, więc istnieje wektor $v_2 \in V \setminus \operatorname{span}\{v_1\}$. Wtedy $v_1, v_2$ są liniowo niezależne. Analogicznie, $\operatorname{span}\{v_1, v_2\} \neq V$, więc istnieje wektor $v_3 \in V \setminus \operatorname{span}\{v_1, v_2\}$. Wtedy $v_1, v_2, v_3$ są liniowo niezależne. Możemy ten proces kontynuować dla każdego $n \in \mathbb{N}$, co kończy dowód punktu (ii).

Załóżmy, że zachodzi warunek (ii). Gdyby $V$ miała wymiar skończony $n \geqslant 1$, to wektory $v_1,\ldots,v_{n+1}$ byłyby liniowo zależne. □

U

Pojęcie bazy w przestrzeni nieskończenie wymiarowej $V$ możemy zdefiniować następująco. Powiemy, że podzbiór $A \subset V$ jest liniowo niezależny, jeżeli dla dla każdego $n \in \mathbb{N}$ dowolne różne wektory $v_1,\ldots,v_n \in A$ są liniowo niezależne. Podzbiór $A$ *generuje* $V$, gdy dla każdego wektora $v \in V$ istnieją takie wektory $w_1,\ldots,w_n \in A$ oraz skalary $x_1,\ldots,x_n \in \mathbb{F}$, że $v = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$. Podzbiór $B \subset V$ jest *bazą*, jeśli $B$ jest liniowo niezależny i generuje $V$. W oparciu o lemat Kuratowskiego–Zorna można udowodnić, że w każdej przestrzeni wektorowej istnieje baza. Dokładniej, jeśli $A \subset V$ jest liniowo niezależny, $C \subset V$ generuje $V$ oraz $A \subset C$, to istnieje taka baza $B$, że $A \subset B \subset C$.

**Przykład 4.9** (Przestrzeń wektorowa ciągów rzeczywistych)**.** Rozważmy przestrzeń wektorową

$$\mathbb{R}^\infty = \Big\{[x_1, x_2, \ldots]^T : x_i \in \mathbb{R}\Big\}$$

wszystkich ciągów nieskończonych o wyrazach rzeczywistych z naturalnie określonymi działaniami. Sprawdzamy łatwo, że ciąg wektorów

$$e_1 = [1, 0, 0 \ldots]^T, \quad e_2 = [0, 1, 0, \ldots]^T, \ldots$$

tworzy zbiór liniowo niezależny $A$, więc $\mathbb{R}^\infty$ jest nieskończenie wymiarowa. $A$ nie generuje przestrzeni $\mathbb{R}^\infty$, bo na przykład ciąg

$$1^\infty = [1, 1, 1, \ldots]^T$$

nie jest skończoną kombinacją liniową wektorów $e_1, e_2, \ldots$.

**Przykład 4.10** (Przestrzeń wektorowa wielomianów)**.** Niech $(P, +, \cdot)$ będzie przestrzenią wektorową wszystkich wielomianów o współczynnikach rzeczywistych. Nie jest to przestrzeń skończonego wymiaru, bo wielomiany

$$1, x, x^2, x^3, \ldots$$

są liniowo niezależne.

Niech $P_n \subset P$ będzie zbiorem wielomianów stopnia **co najwyżej** $n$. Wtedy, $P_n$ jest podprzestrzenią wektorową oraz

$$\dim P_n = n + 1.$$

Bazą dla $P_n$ są wielomiany $1, x, x^2, \ldots, x^n$.

Jeśli $P(n) \subset P$ jest zbiorem wielomianów stopnia $n \geqslant 1$, to $P$ nie jest podprzestrzenią wektorową, bo wielomian zerowy nie należy do $P_n$.

**Twierdzenie 4.4.** *Niech $V \in \mathrm{Vekt}_{\mathbb{F}}$ i $\dim_{\mathbb{F}} V = n \in \mathbb{N}$. Dla wektorów $v_1, \ldots, v_n \in V$ następujące warunki są równoważne:*

*(1) $v_1, \ldots, v_n$ są liniowo niezależne;*

*(2) $\operatorname{span}\{v_1, \ldots, v_n\} = V$;*

*(3) $v_1, \ldots, v_n$ stanowią bazę $V$.*

*Dowód.* (1) $\Rightarrow$ (2). Zakładamy, że $v_1, \ldots v_n$ są liniowo niezależne. Przypuśćmy, że

$$\operatorname{span}\{v_1, \ldots, v_n\} \neq V.$$

Istnieje wtedy wektor $v \notin V \setminus \operatorname{span}\{v_1, \ldots, v_n\}$. Z lematu 4.2 wynika, że wektory $v_1, \ldots v_n, v$ są liniowo niezależne. Jest to sprzeczne z twierdzeniem 4.3.

(2) $\Rightarrow$ (3) Wynika z lematu 4.1 i wniosku 31.2.

(3) $\Rightarrow$ (1) Jest konsekwencją definicji bazy. □

**Wniosek 4.4.** *Wektory $a_1, a_2 \in \mathbb{F}^2$ tworzą bazę $\mathbb{F}^2$ wtedy i tylko wtedy, gdy*

$$\det \Big[a_1 \mid a_2\Big] \neq 0.$$

**Wniosek 4.5.** *Niech $A = \Big[a_1 \mid \ldots \mid a_n\Big] \in M_{n\times n}(\mathbb{F})$. Następujące warunki są równoważne*

*(1) układ $Ax = 0$ ma jednoznaczne rozwiązanie zerowe,*

*(2) układ $Ax = b$ ma rozwiązanie dla każdego $b \in \mathbb{F}^n$,*

*(3) kolumny $a_1, \ldots, a_n$ tworzą bazę $\mathbb{F}^n$.*

*Dowód.* Punkt (1) oznacza, że wektory $a_1, \ldots, a_n$ są liniowo niezależne, a punkt (2) mówi, że $\operatorname{span}\{a_1, \ldots, a_n\} = \mathbb{F}^n$. □

**Wniosek 4.6.** *Załóżmy, że* $\dim V = n$, $1 \leqslant k < n$ *i wektory* $v_1, \ldots, v_k$ *są liniowo niezależne. Wtedy istnieją takie wektory* $v_{k+1}, \ldots, v_n$, *że wektory* $v_1, \ldots, v_n$ *tworzą bazę* $V$.

*Dowód.* Ponieważ $k < n$, więc $v_1, \ldots, v_k$ nie jest bazą. Istnieje więc wektor

$$v_{k+1} \notin \operatorname{span}\{v_1, \ldots, v_k\}.$$

Z lematu 4.2 wektory $v_1, \ldots, v_k, v_{k+1}$ są liniowo niezależne. Iterujemy tę procedurę i zakończy się ona po skończonej liczbie kroków. □

## 4.2 Podprzestrzenie wektorowe

**Definicja 4.5** (Podprzestrzeń wektorowa)**.**

Niech $(V, +, \cdot)$ będzie przestrzenią wektorową nad ciałem $\mathbb{F}$ i niech $S \subset V$. Powiemy, że $S$ jest *podprzestrzenią wektorową przestrzeni* $V$, jeśli zachodzą warunki

(P1) $0 \in S$ tzn. wektor zerowy należy do $S$;

(P2) $v + w \in S$ dla dowolnych $v, w \in S$ tzn. dodawanie wektorów nie wyprowadza nas ze zbioru $S$;

(P3) $t \cdot v \in S$ dla dowolnych $t \in \mathbb{F}$ i $v \in V$ tzn. mnożenie przez skalar nie wyprowadza nas ze zbioru $S$.

U Czasami (P1) zastępuje się warunkiem, że $S \neq \emptyset$. Te dwie definicje są równoważne. Rzeczywiście, załóżmy, że $S \neq \emptyset$ i zachodzą warunki (P2) i (P3). Niech $v \in S$. Wtedy $-v \in S$ z warunku (P3) i $0 = v + (-v) \in S$ z warunku (P2).

**Wniosek 4.7.** *Jeśli* $v_1, \ldots, v_k \in V$, *to* $\operatorname{span}\{v_1, \ldots, v_k\}$ *jest podprzestrzenią przestrzeni wektorowej* $V$.

**Wniosek 4.8.** *Zbiór rozwiązań równania jednorodnego* $Ax = 0$ *jest podprzestrzenią wektorową przestrzeni* $\mathbb{F}^n$ *dla* $A \in M_{k \times n}(\mathbb{F})$.

**Przykład 4.11.** Rozważmy płaszczyznę $S = \left\{[x_1, x_2, x_3]^T \in \mathbb{R}^3 : 2x_1 - 3x_2 + x_3 = 0\right\}$. Zauważmy, że wtedy $x_3 = 3x_2 - 2x_1$, więc

$$\begin{aligned}
[x_1, x_2, x_3]^T &= [x_1, x_2, 3x_2 - 2x_1]^T \\
&= [x_1, 0, -2x_1]^T + [0, x_2, 3x_2]^T \\
&= x_1 \cdot [1, 0, -2]^T + x_2 \cdot [0, 1, 3]^T.
\end{aligned}$$

Stąd $S = \operatorname{span}\left\{[1, 0, -2]^T, [0, 1, 3]^T\right\}$, czyli jest to podprzestrzeń wektorowa w $\mathbb{R}^3$.

**Przykład 4.12.** Pokażemy, że zbiór $S = \left\{[a + b, a - b + 2c, b, c]^T : a, b, c \in \mathbb{R}\right\} \subset \mathbb{R}^4$ jest podprzestrzenią wektorową w $\mathbb{R}^4$. Zauważmy, że

$$\begin{aligned}
[a + b, a - b + 2c, b, c]^T &= [a, a, 0, 0]^T + [b, -b, b, 0]^T + [0, 2c, 0, c]^T \\
&= a \cdot [1, 1, 0, 0]^T + b \cdot [1, -1, 1, 0]^T + c \cdot [0, 2, 0, 1]^T.
\end{aligned}$$

Oznacza to, że $S = \operatorname{span}\left\{[1, 1, 0, 0]^T, [1, -1, 1, 0]^T, [0, 2, 0, 1]^T\right\}$, więc jest to podprzestrzeń wektorowa.

**Przykład 4.13.** Okrąg $\mathbb{S}^1 \subset \mathbb{R}^2$ nie spełnia ani jednego z warunków (P1)–(P3). Nie jest więc podprzestrzenią wektorową.

**Przykład 4.14.** Zbiór $S = \left\{[x_1, x_2]^T \in \mathbb{R}^2 : x_1 x_2 = 0\right\}$ spełnia warunki (P1) i (P3), ale nie spełnia (P2). Nie jest więc podprzestrzenią wektorową.

**Przykład 4.15.** Zbiór $S = \Big\{[x_1, x_2]^T \in \mathbb{R}^2 : x_2 \geqslant 0\Big\}$ spełnia warunki (P1) i (P2), ale nie spełnia (P3). Nie jest więc podprzestrzenią wektorową.

**Przykład 4.16.** Zbiór pusty $\emptyset \subset \mathbb{R}^2$ spełnia warunki (P2) i (P3), ale nie spełnia (P1). Nie jest więc podprzestrzenią wektorową.

**Przykład 4.17.** Rozważmy podprzestrzeń $V \subset \mathbb{Z}_2^3$ daną przez

$$V = \Big\{[x_1, x_2, x_3]^T \in \mathbb{Z}_2^3 : x_1 + x_2 + x_3 = 0\Big\}.$$

W ciele $\mathbb{Z}_2$ równość $x_1 + x_2 + x_3 = 0$ oznacza, że $x_1 = x_2 + x_3$, czyli wektory z $V$ mają postać

$$[x_2 + x_3, x_2, x_3]^T = x_2 \cdot [1, 1, 0]^T + x_3 \cdot [1, 0, 1], \quad x_2, x_3 \in \mathbb{Z}_2.$$

Wektory $[1, 1, 0]^T, [1, 0, 1]^T \in \mathbb{Z}_2^3$ są liniowo niezależne, więc $\dim V = 2$. Sprawdzamy łatwo rozważając wszystkie możliwości wyboru $x_2$ i $x_3$, że

$$V = \Big\{[0, 0, 0]^T, [1, 0, 1]^T, [1, 1, 0]^T, [0, 1, 1]^T\Big\}.$$

**Przykład 4.18.** Zastanówmy się ile 2-wymiarowych podprzestrzeni ma przestrzeń wektorowa $\mathbb{Z}_2^4$? Niech $v \in \mathbb{Z}_2^4$ będzie niezerowym wektorem. Wtedy

$$\operatorname{span}\{v\} = \{0, v\},$$

więc jeśli $v, u$ są różnymi niezerowymi wektorami, to są one liniowo niezależne w $\mathbb{Z}_2^4$. Ponadto,

$$\operatorname{span}\{u, v\} = \{0, u, v, u + v\}$$

składa się z 4 różnych wektorów. W przestrzeni $\mathbb{Z}_2^4$ mamy $2^4$ wektorów i $2^4 - 1$ wektorów niezerowych. Par liniowo niezależnych wektorów $u, v$ mamy więc $\dfrac{(2^4-1)(2^4-2)}{2}$. Ponieważ

$$\operatorname{span}\{u, v\} = \operatorname{span}\{u, u + v\} = \operatorname{span}\{v, u + v\},$$

więc różnych podprzestrzeni 2 wymiarowych jest $\dfrac{(2^4-1)(2^4-2)}{6}$.

TEST $\longrightarrow$ Oceń prawdziwość zdań:

- Jeśli $v_1, v_2, v_3$ tworzą bazę $\mathbb{R}^3$ i $b \in \mathbb{R}^3$ jest niezerowym wektorem, to $b + v_1, v_2, v_3$ również tworzy bazę $\mathbb{R}^3$.
- Istnieją takie 2-wymiarowe podprzestrzenie wektorowe $U$ i $W$ w $\mathbb{R}^3$, że $U \cap W = \{0\}$.
- Istnieją takie 2-wymiarowe podprzestrzenie wektorowe $U$ i $W$ w $\mathbb{R}^4$, że $U \cap W = \{0\}$.
- Jeśli $U, W$ są podprzestrzeniami wektorowymi $\mathbb{R}^2$, to $U \cup W$ jest również podprzestrzenią wektorową.
- Jeśli wektory $v_1, \ldots, v_n$ generują $\mathbb{R}^n$, to są liniowo niezależne.
- Jeśli wektory $v_1, \ldots, v_n$ generują przestrzeń wektorową $V$, to są liniowo niezależne.
- Jeśli wektory $v_1, \ldots, v_n$ są liniowo zależne, to któryś z nich jest wektorem zerowym.
- Jeśli wektory $v_1, \ldots, v_k$ są liniowo zależne w $\mathbb{R}^n$, to $k > n$.
- Jeśli $\operatorname{span}\{v_1, \ldots, v_k\} = \mathbb{R}^n$, to $k = n$.
- Jeśli $\operatorname{span}\{u, v\} = \mathbb{R}^2$, to $\operatorname{span}\{u, v + w\} = \mathbb{R}^2$ dla dowolnego $w \in \mathbb{R}^2$.

- Przestrzenie $M_{n\times n}(\mathbb{C})$ i $M_{n\times n}(\mathbb{R})$ mają ten sam wymiar nad ciałem liczb rzeczywistych $\mathbb{R}$.
- Jeśli $\operatorname{span}\{u,v,w\}=\mathbb{R}^2$, to któreś dwa wektory spośród $u,v,w$ są liniowo niezależne.
- Dla dowolnej macierzy $A\in M_{2\times 2}(\mathbb{R})$ macierze $I,A,A^2,A^3,A^4$ są liniowo zależne w przestrzeni wektorowej $M_{2\times 2}(\mathbb{R})$.
- Dla dowolnej macierzy niezerowej $A\in M_{2\times 2}(\mathbb{R})$ macierze
$$I,A,A^2,A^3,A^4$$
generują przestrzeń wektorową $M_{2\times 2}(\mathbb{R})$.
- Zbiór wszystkich takich funkcji $f:\mathbb{R}\to\mathbb{R}$, że $f(0)=3f(1)$ jest przestrzenią wektorową z naturalnymi działaniami.
- Funkcje $x^2$ i $\sin x$ są liniowo niezależne w przestrzeni funkcji $f:\mathbb{R}\to\mathbb{R}$.
- Funkcje $x^2$ i $\sin x$ generują przestrzeń funkcji $f:\mathbb{R}\to\mathbb{R}$.
- Jeśli $v\in\mathbb{R}^3$ jest dowolnym wektorem, to
$$v^\perp=\{w\in\mathbb{R}^3:(v|w)=0\}$$
jest podprzestrzenią wektorową $\mathbb{R}^3$.
- Jeśli $u,v\in\mathbb{R}^n$ są takie, że $(u|v)=0$, to $u,v$ są liniowo niezależne.

$\longleftarrow$

**TEST** $\longrightarrow$ Oceń które z poniższych podzbiorów przestrzeni wektorowej $M_{2\times 2}(\mathbb{C})$ są jej podprzestrzeniami wektorowymi:

- $\{A\in M_{2\times 2}(\mathbb{C}):\det(A)=0\}$,
- $\{A\in M_{2\times 2}(\mathbb{C}):\operatorname{tr}(A)=0\}$,
- $\{A\in M_{2\times 2}(\mathbb{C}):\det(A)\neq 0\}$,
- $\{A\in M_{2\times 2}(\mathbb{C}):\operatorname{tr}(A)\neq 0\}$,
- $\{A\in M_{2\times 2}(\mathbb{C}):A=A^*\}$,
- $\left\{A\in M_{2\times 2}(\mathbb{C}):A=\overline{A}\right\}$,
- $\left\{A\in M_{2\times 2}(\mathbb{C}):A=A^T\right\}$,
- $\{A\in M_{2\times 2}(\mathbb{C}):A^2=0\}$,
- $\{A\in M_{2\times 2}(\mathbb{C}):a_{11}=0\}$,
- $\{A\in M_{2\times 2}(\mathbb{C}):a_{12}=a_{21}=0\}$,
- $\left\{A\in M_{2\times 2}(\mathbb{C}):\operatorname{tr}(A)=\operatorname{tr}(\overline{A})\right\}$,
- $\{A\in M_{2\times 2}(\mathbb{C}):AJ=JA\}$, gdzie $J=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$,
- $\{A\in M_{2\times 2}(\mathbb{C}):AB=0\}$, gdzie $B\in M_{2\times 2}(\mathbb{C})$ jest ustaloną macierzą.

$\longleftarrow$

**TEST** $\longrightarrow$ Oceń które z poniższych podzbiorów przestrzeni wektorowej $F(\mathbb{R},\mathbb{R})$ funkcji $F:\mathbb{R}\longrightarrow\mathbb{R}$ są jej podprzestrzeniami wektorowymi:

- $C(\mathbb{R},\mathbb{R})=\{f\in F(\mathbb{R},\mathbb{R}):f$ jest ciągła$\}$,
- $B(\mathbb{R},\mathbb{R})=\{f\in F(\mathbb{R},\mathbb{R}):f$ jest ograniczona$\}$,

- $\{f \in F(\mathbb{R}, \mathbb{R}) : |f(x)| < 1$ dla każdego $x \in \mathbb{R}\}$,
- $CB(\mathbb{R}, \mathbb{R}) = \{f \in F(\mathbb{R}, \mathbb{R}) : f$ jest ciągła i ograniczona $\}$,
- $\{f \in F(\mathbb{R}, \mathbb{R}) : f$ jest injekcją$\}$,
- $\{f \in F(\mathbb{R}, \mathbb{R}) : f$ jest surjekcją$\}$,
- $\{f \in F(\mathbb{R}, \mathbb{R}) : f$ jest bijekcją$\}$,
- $\{f \in F(\mathbb{R}, \mathbb{R}) : f$ jest różniczkowalna$\}$,
- $C^1(\mathbb{R}, \mathbb{R}) = \{f \in F(\mathbb{R}, \mathbb{R}) : f$ ma ciągłą pochodną$\}$,
- $C^2(\mathbb{R}, \mathbb{R}) = \{f \in F(\mathbb{R}, \mathbb{R}) : f$ ma ciągłą pochodną rzędu $2\}$,
- $C^k(\mathbb{R}, \mathbb{R}) = \{f \in F(\mathbb{R}, \mathbb{R}) : f$ ma ciągłą pochodną rzędu $k \in \mathbb{N}\}$,
- $C^\infty(\mathbb{R}, \mathbb{R}) = \{f \in F(\mathbb{R}, \mathbb{R}) : f$ ma ciągłą pochodną dowolnego rzędu $k \in \mathbb{N}\}$,
- $\{f \in F(\mathbb{R}, \mathbb{R}) : f(0) = 0\}$,
- $\{f \in F(\mathbb{R}, \mathbb{R}) : f(0) = 1\}$,
- $F_o(\mathbb{R}, \mathbb{R}) = \{f \in F(\mathbb{R}, \mathbb{R}) : f$ jest nieparzysta$\}$,
- $F_e(\mathbb{R}, \mathbb{R}) = \{f \in F(\mathbb{R}, \mathbb{R}) : f$ jest parzysta$\}$.

$\longleftarrow$

Nie za dużo, nie za mało, tylko w sam raz

Niech $n = \dim V$.

- jeśli $w_1, \ldots, w_k$ są liniowo niezależne, to $k \leqslant n$; wtedy $w_1, \ldots, w_k$ da się uzupełnić do bazy,
- jeśli $\operatorname{span}\{w_1, \ldots, w_k\} = V$, to $k \geqslant n$; wtedy ze zbioru $\{w_1, \ldots, w_k\}$ da się wybrać bazę.

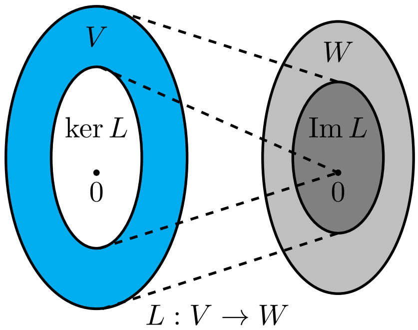
Poniższe warunki są równoważne

- $v_1, \ldots, v_n$ są bazą $V$,
- $v_1, \ldots, v_n$ są liniowo niezależne,
- $\operatorname{span}\{v_1, \ldots, v_n\} = V$.

# Rozdział 5

# Odwzorowania liniowe

Słowa kluczowe rozdziału
odwzorowanie liniowe ♢ macierz jako odwzorowanie liniowe ♢ rzut prostopadły na prostą ♢ symetria względem prostej ♢ obrót ♢ przestrzeń wektorowa odwzorowań liniowych ♢ jądro i obraz ♢ mono-epi-izo ♢ formuła wymiaru




Dwie podstawowe podprzestrzenie: jądro i obraz.

**Definicja 5.1** (Odwzorowanie liniowe)**.**
Niech $V$, $W$ będą przestrzeniami wektorowymi nad tym samym ciałem $\mathbb{F}$. Odwzorowanie $L : V \longrightarrow W$ nazywamy *odwzorowaniem liniowym*, jeśli zachodzą warunki

(L1) (addytywność)

$$L(v_1 + v_2) = L(v_1) + L(v_2), \quad v_1, v_2 \in V,$$

(L2) (jednorodność)

$$L(t \cdot v) = t \cdot L(v), \quad t \in \mathbb{F},\ v \in V.$$

**Wniosek 5.1.** *Jeśli $L : V \longrightarrow W$ jest odwzorowaniem liniowym, to $L(0) = 0$.*

*Dowód.* Teza zachodzi, bo $L(0) = L(0 + 0) = L(0) + L(0)$, więc $L(0) = 0$. □

**Przykład 5.1.** Odwzorowanie $L : \mathbb{R} \ni x \longmapsto x + 1 \in \mathbb{R}$ nie jest odwzorowaniem liniowym, bo $L(0) = 1 \neq 0$.

**Przykład 5.2.** Odwzorowanie $L : \mathbb{R} \ni x \longmapsto x^2 \in \mathbb{R}$ nie jest odwzorowaniem liniowym pomimo, że $L(0) = 0$. Nie spełnia ono żadnego z warunków (L1)–(L2). Rzeczywiście, $L(x+y) = (x+y)^2$ oraz $L(x) + L(y) = x^2 + y^2$ i równość $(x + y)^2 = x^2 + y^2$ nie zachodzi dla wszystkich liczb rzeczywistych. Podobnie, $L(tx) = t^2x^2$, a $tL(x) = tx^2$.

**Przykład 5.3.** Pokażemy, że jeśli $L : \mathbb{R} \longrightarrow \mathbb{R}$ jest liniowe, to istnieje taka liczba rzeczywista $a \in \mathbb{R}$, że $L(x) = ax$. Rzeczywiście,

$$L(x) = L(x \cdot 1) = xL(1),$$

więc wystarczy przyjąć, że $a = L(1)$.

Odwzorowanie liniowe zadane macierzą

Rozważmy macierz $A = \begin{bmatrix} a_1 \mid \ldots \mid a_n \end{bmatrix} \in M_{k\times n}(\mathbb{F})$ i wektor $x = [x_1, \ldots, x_n]^T \in \mathbb{F}^n$. Definiujemy

$$\begin{aligned} Ax &= \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{k1} & \ldots & a_{kn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &:= \begin{bmatrix} x_1 a_{11} + \ldots + x_n a_{1n} \\ \vdots \\ x_1 a_{k1} + \ldots + x_n a_{kn} \end{bmatrix} \\ &= x_1 \cdot a_1 + \ldots + x_n \cdot a_n \in \mathbb{F}^k. \end{aligned}$$

Możemy określić odwzorowanie $L_A : \mathbb{F}^n \longrightarrow \mathbb{F}^k$ przez

$$L_A(x) := Ax.$$

Jest to odwzorowanie liniowe, bo jak łatwo sprawdzamy

$$A(x+y) = Ax + Ay, \quad A(t \cdot x) = t \cdot Ax.$$

**Przykład 5.4.** Odwzorowanie

$$L : \mathbb{R}^2 \ni [x_1, x_2]^T \longrightarrow [3x_1 - 2x_2, x_1 + 4x_2, x_1 + x_2]^T \in \mathbb{R}^3$$

jest liniowe, bo $L = L_A$ (jest ono zadane macierzą $A$) dla

$$A = \begin{bmatrix} 3 & -2 \\ 1 & 4 \\ 1 & 1 \end{bmatrix}.$$

Zauważmy, że $A = \begin{bmatrix} L(e_1) \mid L(e_2) \end{bmatrix}$.

**Przykład 5.5** (Rzut prostopadły na prostą w $\mathbb{R}^2$)**.** Niech $u = [u_1, u_2]^T \in \mathbb{R}^2$ będzie niezerowym wektorem. Rozważmy odwzorowanie $P : \mathbb{R}^2 \to \mathbb{R}^2$ dane wzorem

$$P(v) = \frac{(v|u)}{(u|u)} u.$$

Jest to rzut prostopadły na prostą generowaną przez wektor $u$. Zauważmy, że

$$P(e_1) = \frac{1}{u_1^2 + u_2^2} \begin{bmatrix} u_1^2 \\ u_1 u_2 \end{bmatrix}, \quad P(e_2) = \frac{1}{u_1^2 + u_2^2} \begin{bmatrix} u_1 u_2 \\ u_2^2 \end{bmatrix}.$$

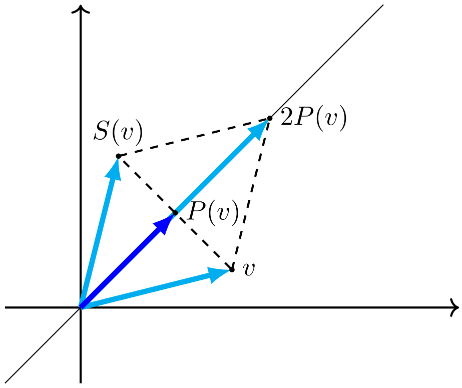
Rozważmy macierz

$$A = \frac{1}{u_1^2 + u_2^2} \begin{bmatrix} u_1^2 & u_1 u_2 \\ u_1 u_2 & u_2^2 \end{bmatrix}.$$

Wtedy $P(v) = L_A v$. Zauważmy, że dla wektora jednostkowego $u = [\cos\theta, \sin\theta]^T$ mamy

$$A = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix} = \underbrace{\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \cdot \begin{bmatrix} \cos\theta & \sin\theta \end{bmatrix}}_{\text{iloczyn macierzy}}.$$

**Przykład 5.6** (Symetria względem prostej na płaszczyźnie)**.** Załóżmy, że prosta $l$ tworzy z osią $x$ kąt $\theta$. Znajdziemy macierz $M_S$ symetrii $S$ względem prostej $l$. Wektor $u = [\cos\theta, \sin\theta]^T$ jest




Rysunek 5.1: Symetria $S$ względem prostej $l$. Jeśli $P$ jest rzutem prostopadłym na prostą $l$, to $S(v)+v = 2P(v)$.

wektorem jednostkowym na prostej $l$. Niech $M_P$ będzie macierzą rzutu prostopadłego $P$ na prostą $l$. Sprawdzamy łatwo, że

$$S + I = 2P,$$

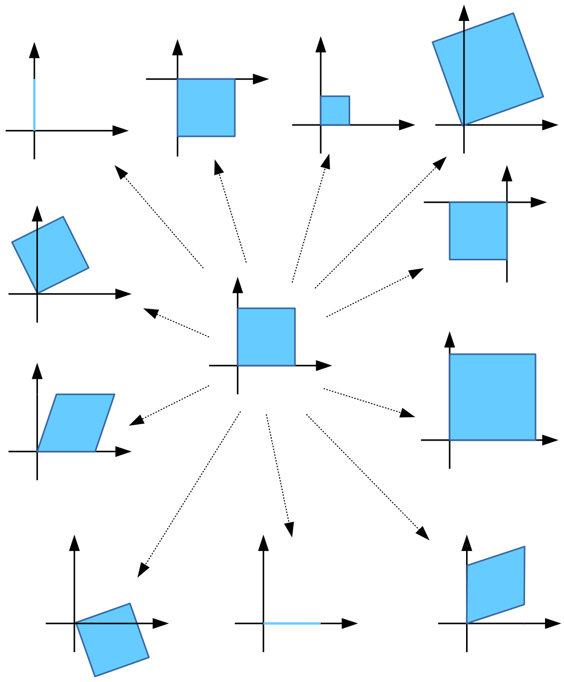
czyli $S = 2P - I$. Stąd

$$\begin{aligned} M_S &= 2M_P - I \\ &= \begin{bmatrix} 2\cos^2\theta - 1 & 2\sin\theta\cos\theta \\ 2\sin\theta\cos\theta & 2\sin^2\theta - 1 \end{bmatrix} \\ &= \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}. \end{aligned}$$

TEST

Rysunek 5.2 przedstawia obrazy kwadratu jednostkowego rozpiętego na wektorach bazowych $e_1, e_2$ pod wpływem działania macierzy $A \in M_{2\times 2}(\mathbb{R})$. Dopasować poniższe macierze z odpowiednim obrazkiem:

- $A = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$ z $0 < a < 1$,
- $A = \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix}$ z $0 < a < 1$,
- jednokładność: $A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$ z $0 < a < 1$,
- jednokładność: $A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$ z $1 < a$,
- symetria względem osi $x$: $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$,
- obrót: $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$
- symetria względem punktu 0: $A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$,
- symetria względem prostej nachylonej do osi $x$ pod kątem $\theta$: $A = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$,
- $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ z $a, b > 0$,



Rysunek 5.2: Obrazy kwadratu jednostkowego pod wpływem macierzy $A$.

- rzutowanie na oś $x$: $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$,
- rzutowanie na oś $x$: $A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$,

Przeanalizować jak zmienia się pole obrazu kwadratu jednostkowego przez $A$ w zależności od wyznacznika macierzy $A$.

$\longleftarrow$

## 5.1 Przestrzeń wektorowa odwzorowań liniowych

**Definicja 5.2** (Przestrzeń wektorowa odwzorowań liniowych)**.**

Niech $V, W \in \mathrm{Vekt}_{\mathbb{F}}$, czyli $V, W$ będą przestrzeniami wektorowymi nad ciałem $\mathbb{F}$. Definiujemy zbiór

$$\mathcal{L}(V, W) = \left\{ L : V \longrightarrow W : \ L \text{ jest liniowe} \right\}.$$

Wprowadzimy w zbiorze $\mathcal{L}(V, W)$ strukturę przestrzeni wektorowej nad ciałem $\mathbb{F}$. Dla $L, G \in \mathcal{L}(V, W)$ oraz $t \in \mathbb{F}$ definiujemy:

$$(L + G)(v) := L(v) + G(v), \quad (t \cdot L)(v) := t \cdot L(v), \quad v \in V$$

Tak określona trójka $(\mathcal{L}(V, W), +, \cdot)$ jest przestrzenią wektorową nad ciałem $\mathbb{F}$.

U Załóżmy, że $v_1, \ldots, v_n$ jest bazą przestrzeni wektorowej $V$ i $L : V \longrightarrow W$ jest odwzorowaniem liniowym. Dla dowolnego $v \in V$ istnieją jednoznacznie wyznaczone takie skalary $x_1, \ldots, x_n$, że

$v = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$. Wtedy z liniowości $L$ otrzymujemy, że

$$\begin{aligned} L(v) &= L(x_1 \cdot v_1 + \ldots + x_n \cdot v_n) \\ &= L(x_1 \cdot v_1) + \ldots + L(x_n \cdot v_n) \\ &= x_1 \cdot L(v_1) + \ldots + x_n \cdot L(v_n), \end{aligned}$$

czyli

$$L(v) = x_1 \cdot L(v_1) + \ldots + x_n \cdot L(v_n). \tag{5.1}$$

Równość (5.1) oznacza, że odwzorowanie $L$ jest jednoznacznie wyznaczone przez swoje wartości $L(v1), \ldots, L(v_n)$ na dowolnej bazie przestrzeni $V$. Jest to bardzo przyjemna własność odwzorowań liniowych: wystarczy znać skończoną liczbę wartości, aby znać całe odwzorowanie.

**Wniosek 5.2.** *Niech $v_1, \ldots, v_n$ będzie bazą przestrzeni wektorowej $V$. Dla dowolnych wektorów $w_1, \ldots, w_n \in W$ istnieje dokładnie jedno takie odwzorowanie liniowe $L : V \longrightarrow W$, że*

$$L(v_1) = w_1, \ \ldots, \ L(v_n) = w_n.$$

*W szczególności, jeśli $L, G \in \mathcal{L}(V, W)$ są takie, że*

$$L(v_1) = G(v_1), \ \ldots, \ L(v_n) = G(v_n),$$

*to $L = G$.*

*Dowód.* Pokażemy najpierw, że takie odwzorowanie liniowe $L$ istnieje. Wektor $v \in V$ zapisujemy jednoznacznie w bazie $v_1, \ldots, v_n$ jako kombinację liniową $v = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$. Definiujemy odwzorowanie $L : V \to W$ przez

$$L(v) = x_1 \cdot L(v_1) + \ldots + x_n \cdot L(v_n) := x_1 \cdot w_1 + \ldots + x_n \cdot w_n.$$

Tak zdefiniowane $L$ jest oczywiście liniowe. Jeśli, $L' : V \to W$ jest innym odwzorowaniem liniowym spełniającym tezę, to

$$\begin{aligned} L'(v) &= L'(x_1 \cdot v_1 + \ldots + x_n \cdot v_n) \\ &= L'(x_1 \cdot v_1) + \ldots + L'(x_n \cdot v_n) \\ &= x_1 \cdot L'(v_1) + \ldots + x_n \cdot L'(v_n) \\ &= x_1 \cdot w_1 + \ldots + x_n \cdot w_n = L(v). \end{aligned}$$

□

**Lemat 5.1.** *Jeśli $L : V \longrightarrow W$ i $G : W \longrightarrow Z$ są odwzorowaniami liniowymi, to $G \circ L : V \longrightarrow Z$ jest również liniowe.*

*Dowód.* Warunek (L1) zachodzi, bo dla $v_1, v_2 \in V$ mamy

$$\begin{aligned} (G \circ L)(v_1 + v_2) &= G(L(v_1 + v_2)) \\ &= G(L(v_1) + L(v_2)) \\ &= G(L(v_1)) + G(L(v_2)) \\ &= (G \circ L)(v_1) + (G \circ L)(v_2). \end{aligned}$$

Podobnie,

$$\begin{aligned} (G \circ L)(t \cdot v) &= G(L(t \cdot v)) \\ &= G(t \cdot L(v)) \\ &= t \cdot G(L(v)) \\ &= t \cdot (G \circ L)(v). \end{aligned}$$

□

## 5.2 Jądro i obraz odwzorowania liniowego

**Definicja 5.3** (Jądro odwzorowania liniowego)**.**
*Jądro* odwzorowania liniowego $L : V \longrightarrow W$ definiujemy jako

$$\ker L := \{v \in V : L(v) = 0\} = L^{-1}(\{0\}).$$

**Wniosek 5.3.** *Jeśli* $L : V \longrightarrow W$ *jest odwzorowaniem liniowym, to* $\ker L$ *jest podprzestrzenią przestrzeni wektorowej* $V$*.*

*Dowód.* Ponieważ $L(0) = 0$, więc $0 \in \ker L$, czyli zachodzi warunek (P1). Jeśli $v, w \in \ker L$, to

$$\begin{aligned} L(v + w) &= L(v) + L(w) \\ &= 0 + 0 = 0, \end{aligned}$$

więc $v + w \in \ker L$, czyli spełniony jest warunek (P2). Dla $t \in \mathbb{F}$ i $v \in \ker L$, to

$$\begin{aligned} L(t \cdot v) &= t \cdot L(v) \\ &= t \cdot 0 = 0, \end{aligned}$$

czyli zachodzi warunek (P3). □

**Przykład 5.7.** Rozważmy płaszczyznę $P$ o równaniu $3x_1 - x_2 + 2x_3 = 0$ w $\mathbb{R}^3$. Ponieważ odwzorowanie

$$L : \mathbb{R}^3 \ni [x_1, x_2, x_3]^T \longmapsto 3x_1 - x_2 + 2x_3 \in \mathbb{R}$$

jest liniowe oraz $P = \ker L$, więc $P$ jest podprzestrzenią wektorową w $\mathbb{R}^3$.

**Definicja 5.4** (Obraz odwzorowania liniowego)**.**
*Obraz* odwzorowania liniowego $L : V \longrightarrow W$ to zbiór

$$\operatorname{Im} L := L(V) = \{L(v) : v \in V\} \subset W.$$

**Wniosek 5.4.** *Jeśli* $L : V \longrightarrow W$ *jest odwzorowaniem liniowym, to jego obraz* $L(V)$ *jest podprzestrzenią wektorową w* $W$*.*

*Dowód.* Ponieważ $L(0) = 0$, więc $0 \in L(V)$. Niech $w_1, w_2 \in L(W)$. Wtedy istnieją takie $v_1, v_2 \in V$, że

$$L(v_1) = w_2, \quad L(v_2) = w_2.$$

Stąd

$$\begin{aligned} w_1 + w_2 &= L(v_1) + L(v_2) \\ &= L(v_1 + v_2), \end{aligned}$$

czyli $w_1 + w_2 \in L(V)$. Jeśli $t \in \mathbb{F}$ i $w \in L(V)$, to istnieje taki wektor $v \in V$, że $L(v) = w$. Stąd

$$\begin{aligned} t \cdot w &= t \cdot L(v) \\ &= L(t \cdot v), \end{aligned}$$

czyli $t \cdot w \in L(V)$. □

**Przykład 5.8.** Zbiór

$$S = \Big\{[x_1 + x_2, 2x_1 - 3x_2, -x_1 + 4x_2, x_2]^T \in \mathbb{R}^4 : x_1, x_2 \in \mathbb{R}\Big\}$$

jest podprzestrzenią wektorową w $\mathbb{R}^4$, bo jest on obrazem odwzorowania liniowego

$$L : \mathbb{R}^2 \ni [x_1, x_2]^T \longmapsto [x_1 + x_2, 2x_1 - 3x_2, -x_1 + 4x_2, x_2]^T \in \mathbb{R}^4.$$

## 5.3 Mono-Epi-Izo. Formuła wymiaru

**Definicja 5.5** (Monomorfizm-epimorfizm-izomorfizm)**.**
Niech $L : V \longrightarrow W$ będzie odwzorowaniem liniowy. Mówimy, że

(i) $L$ jest *monomorfizmem*, gdy $L$ jest różnowartościowe (injekcją),

(ii) $L$ jest *epimorfizmem*, gdy $W = L(V)$, czyli $L$ jest surjekcją,

(iii) $L$ jest *izomorfizmem*, gdy $L$ jest bijekcją, czyli jest zarówno monomorfizmem jak i epimorfizmem.

U

Możecie zapytać dlaczego wprowadzamy nowe nazwy monomorfizm, epimorfizm, izomorfizm na znane nam z teorii mnogości pojęcia injekcja, surjekcja, bijekcja. Odpowiedź jest prosta. Robimy to dla wygody. Przykładowo, stwierdzenie $L : V \longrightarrow W$ jest monomorfizmem niesie w sobie więcej treści niże samo stwierdzenie, że $L$ jest odwzorowaniem różnowartościowym. Zamiast krótkiego $L : V \longrightarrow W$ jest monomorfizmem powinniśmy powiedzieć: $L : V \longrightarrow W$ jest odwzorowaniem liniowym przestrzeni wektorowych i jest różnowartościowe. Monomorfizm oznacza więc różnowartościowość, ale jednocześnie liniowość odwzorowania.

Wiele problemów matematycznych da się sprowadzić do zagadnienia istnienia i jednoznaczności rozwiązań równania
$$L(x) = y,$$
gdzie $L : V \longrightarrow W$ jest odwzorowaniem i $y \in W$. Różnowartościowość $L$ oznacza, że przy ustalonym $y$ istnieje co najwyżej jedno rozwiązanie $x$. Z drugiej strony, surjektywność gwarantuje, że dla dowolnego $y \in W$ istnieje pewne rozwiązanie $x \in V$. Połączenie tych warunków, czyli bijektywność $L$ oznacza, że dla dowolnego $y \in W$ istnieje dokładnie jeden taki $x \in V$, że $L(x) = y$.

**Lemat 5.2.** *$L : V \longrightarrow W$ jest monomorfizmem wtedy i tylko wtedy, gdy* $\ker L = \{0\}$*.*

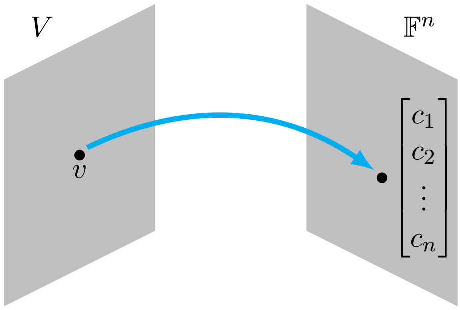
*Dowód.* Niech $L$ będzie monomorfizmem. Jeśli $v \in \ker L$, to $L(0) = 0 = L(v)$, czyli $v = 0$. Oznacza to, że $\ker L = \{0\}$. Załóżmy teraz, że $\ker L = \{0\}$. Jeśli $v, w \in V$ oraz $L(v) = L(w)$, to z liniowości $L$ mamy
$$0 = L(v) - L(w) = L(v - w),$$
czyli $v - w \in \ker L$, więc $v - w = 0$. □

**Lemat 5.3.** *Jeśli $L : V \longrightarrow W$ jest monomorfizmem i $v_1, \dots, v_n \in V$ są liniowo niezależne, to wektory*
$$L(v_1), \dots, L(v_n)$$
*są liniowo niezależne.*

*Dowód.* Załóżmy, że $x_1 \cdot L(v_1) + \ldots + x_n \cdot L(v_n) = 0$ dla pewnych skalarów $x_i \in \mathbb{F}$. Wtedy
$$0 = L(x_1 \cdot v_1 + \ldots + x_n \cdot v_n),$$
czyli
$$x_1 \cdot v_1 + \ldots + x_n \cdot v_n \in \ker L = \{0\},$$
więc
$$x_1 \cdot v_1 + \ldots + x_n \cdot v_n = 0.$$
Ponieważ $v_1, \dots, v_n \in V$ są liniowo niezależne, więc $x_1 = \ldots = x_n = 0$. □

**Lemat 5.4.** *Jeśli $L : V \longrightarrow W$ jest epimorfizmem i $V = \mathrm{span}\{v_1, \dots, v_n\}$, to $W = \mathrm{span}\,\{L(v_1), \dots, L(v_n)\}$.*

*Dowód.* Niech $w \in W$. Ponieważ $L$ jest epimorfizmem, więc istnieje taki wektor $v \in V$, że $L(v) = w$. Z założenia $V = \mathrm{span}\{v_1, \dots, v_n\}$, więc $v = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$ dla pewnych $x_i \in \mathbb{F}$. Wtedy
$$\begin{aligned} w &= L(v) \\ &= L(x_1 \cdot v_1 + \ldots + x_n \cdot v_n) \\ &= x_1 \cdot L(v_1) + \ldots + x_n \cdot L(v_n), \end{aligned}$$
czyli $w \in \mathrm{span}\{L(v_1), \dots, L(v_n)\}$. □




Rysunek 5.3: Wybór bazy $v_1, \ldots, v_n$ w $V$ zadaje izomorfizm z $\mathbb{F}^n$. Wektorowi $v = c_1 \cdot v_1 + \ldots + c_n \cdot v_n$ przypisujemy wektor $[c_1, \ldots, c_n]^T$ jego współrzędnych.

**Wniosek 5.5.** *Jeśli $L : V \longrightarrow W$ jest izomorfizmem liniowym i $v_1, \ldots, v_n \in V$ jest bazą $V$, to $L(v_1), \ldots, L(v_n)$ jest bazą $W$.*

**Wniosek 5.6.** *Niech $L : V \longrightarrow W$ będzie odwzorowaniem liniowym i $\dim V = n$. Następujące warunki są równoważne:*

*(i)* *$L : V \longrightarrow W$ jest izomorfizmem,*

*(ii)* *dla każdej bazy $v_1, \ldots, v_n$ w $V$, wektory $L(v_1), \ldots, L(v_n)$ tworzą bazę w $W$,*

*(iii)* *istnieje taka baza $v_1, \ldots, v_n$ w $V$, że wektory $L(v_1), \ldots, L(v_n)$ tworzą bazę w $W$.*

**Wniosek 5.7.** *Jeśli odwzorowanie $L : V \longrightarrow W$ jest liniowe i $\dim V < \infty$, to $\dim L(V) \leqslant \dim V < \infty$.*

*Dowód.* Zauważmy, że odwzorowanie liniowe

$$L : V \longrightarrow L(V)$$

jest epimorfizmem, więc jeśli $v_1, \ldots, v_n$ jest bazą dla $V$, to

$$L(V) = \operatorname{span}\{L(v_1), \ldots, L(v_n)\}. \qquad \square$$

U Jeśli $L : V \longrightarrow W$ jest izomorfizmem liniowym, to odwzorowanie odwrotne $L^{-1} : W \to V$ jest również liniowe. Rzeczywiście, niech $w_1, w_2 \in W$. Musimy uzasadnić, że $L^{-1}(w_1 + w_2) = L^{-1}(w_1) + L^{-1}(w_2)$. Niech $v_1 = L^{-1}(w_1)$ i $v_2 = L^{-1}(w_2)$. Z liniowości $L$ mamy

$$w_1 + w_2 = L(v_1) + L(v_2) = L(v_1 + v_2),$$

czyli $v_1 + v_2 = L^{-1}(w_1 + w_2)$.
Analogicznie uzasadniamy, że $L^{-1}(t \cdot w) = t \cdot L^{-1}(w)$. Niech $v = L^{-1}(w)$. Wtedy

$$L(t \cdot v) = t \cdot L(v) = t \cdot w,$$

czyli $t \cdot L^{-1}(w) = t \cdot v = L^{-1}(t \cdot w)$.

**Wniosek 5.8.** *Niech $L : V \longrightarrow V$ będzie odwzorowaniem liniowym i $\dim V < \infty$. Wtedy następujące warunki są równoważne*

*(1)* *$L$ jest monomorfizmem,*

*(2)* *$L$ jest epimorfizmem,*

*(3)* *$L$ jest izomorfizmem.*

**Twierdzenie 5.1.** *Niech $V \in \mathrm{Vekt}_{\mathbb{F}}$. Jeśli $\dim_{\mathbb{F}} V = n < \infty$, to $V$ jest izomorficzna z $\mathbb{F}^n$.*

*Dowód.* Niech $v_1, \ldots, v_n$ będzie bazą dla $V$, a $e_1, \ldots, e_n$ bazą standardową w $\mathbb{F}^n$. Definiujemy odwzorowanie liniowe $L : V \longrightarrow \mathbb{F}^n$, zadając jego wartości na bazie:

$$L(v_1) = e_1, \ldots, L(v_n) = e_n.$$

□

**Wniosek 5.9.** *Jeśli* $V, W \in \mathrm{Vekt}_{\mathbb{F}}$ *są skończenie wymiarowe, to* $V$ *jest izomorficzna z* $W$ *wtedy i tylko wtedy, gdy* $\dim V = \dim W$.

*Dowód.* Jeśli $\dim V = \dim W = n$, to $V$ i $W$ są izomorficzne, gdyż obydwie są izomorficzne z $\mathbb{F}^n$. Z drugiej strony, jeśli $V$ i $W$ są izomorficzne, to mają taki sam wymiar, bo izomorfizm przekształca bijektywnie bazę na bazę. □

**Twierdzenie 5.2** (Formuła wymiaru)**.** *Załóżmy, że odwzorowanie* $L : V \longrightarrow W$ *jest liniowe i* $\dim V < \infty$. *Wtedy*

$$\dim V = \dim \ker L + \dim \operatorname{Im} L.$$

*Dowód.* Wybieramy bazę $v_1, \ldots, v_p$ dla $\ker L \subset V$ i rozszerzamy ją do bazy

$$v_1, \ldots, v_p,\ u_1, \ldots, u_q$$

dla $V$. Wtedy

$$p = \dim \ker L, \quad p + q = \dim V.$$

Mamy udowodnić, że $q = \dim L(V)$. Wystarczy pokazać, że wektory $L(u_1), \ldots, L(u_q)$ tworzą bazę $L(V)$. Pokażemy najpierw, że wektory $L(u_1), \ldots, L(u_q)$ generują $L(V)$. Niech $w \in L(V)$. Istnieje więc taki wektor

$$v = c_1 \cdot v_1 + \ldots + c_p \cdot v_p + x_1 \cdot u_1 + \ldots + x_q \cdot u_q,$$

że $L(v) = w$. Stąd

$$\begin{aligned} w &= L(c_1 \cdot v_1 + \ldots + c_p \cdot v_p + x_1 \cdot u_1 + \ldots + x_q \cdot u_q) \\ &= c_1 \cdot L(v_1) + \ldots + c_p \cdot L(v_p) + x_1 \cdot L(u_1) + \ldots + x_q \cdot L(u_q) \\ &= x_1 \cdot L(u_1) + \ldots + x_q \cdot L(u_q), \end{aligned}$$

czyli $w \in \operatorname{span}\{L(u_1), \ldots, L(u_q)\}$.

Pokażemy teraz, że wektory $L(u_1), \ldots, L(u_q)$ są liniowo niezależne. Przypuśćmy, że

$$0 = x_1 \cdot L(u_1) + \ldots + x_q \cdot L(u_q)$$

dla pewnych skalarów $x_i \in \mathbb{R}$. Musimy pokazać, że

$$x_1 = \ldots = x_q = 0.$$

Z liniowości $L$ otrzymujemy, że

$$0 = L(x_1 \cdot u_1 + \ldots + x_q \cdot u_q),$$

czyli $x_1 \cdot u_1 + \ldots + x_q \cdot u_q \in \ker L$. Stąd

$$x_1 \cdot u_1 + \ldots + x_q \cdot u_q$$

jest kombinacją liniową wektorów $v_1, \ldots, v_p$, czyli

$$x_1 \cdot u_1 + \ldots + x_q \cdot u_q = c_1 \cdot v_1 + \ldots + c_p \cdot v_p$$

dla pewnych skalarów $c_i$. Wtedy

$$x_1 \cdot u_1 + \ldots + x_q \cdot u_q - c_1 \cdot v_1 - \ldots - c_p \cdot v_p = 0,$$

więc $x_1 = \ldots = x_q = c_1 = \ldots = c_p = 0$, bo $v_1, \ldots, v_p, u_1, \ldots, u_q$ tworzą bazę $V$. □

**Przykład 5.9.** Rozważmy odwzorowanie liniowe $L : \mathbb{Z}_2^3 \to \mathbb{Z}_2^3$ dane przez

$$L([x_1, x_2, x_3]^T) = [x_1 + x_2, x_3 + x_1, x_2 + x_3]^T.$$

Przestrzeń $\mathbb{Z}_2^3$ składa się tylko z 8 wektorów, więc możemy wypisać wszystkie wartości $L$:

$$L([0,0,0]^T) = L([1,1,1]^T) = [0,0,0]^T,$$

$$L([1,1,0]^T) = L([0,0,1]^T) = [0,1,1]^T,$$
$$L([1,0,1]^T = L([0,1,0]^T) = [1,0,1]^T,$$
$$L([0,1,1]^T) = L([1,0,0]^T) = [1,1,0]^T.$$

Stąd,

$$\ker L = \operatorname{span}\left\{[1,1,1]^T\right\} = \left\{[0,0,0]^T, [1,1,1]^T\right\},$$

$$\operatorname{Im} L = \operatorname{span}\left\{[1,1,0]^T, [1,0,1]^T\right\} = \left\{[1,1,0]^T, [1,0,1]^T, [1,1,0]^T\right\}.$$

# Rozdział 6

# Działania na macierzach

SŁOWA KLUCZOWE ROZDZIAŁU
transponowanie ◊ macierz symetryczna i antysymetryczna ◊ macierz identycznościowa ◊ macierz nieosobliwa ◊ odwrotność iloczynu ◊ transpozycja iloczynu ◊ jądro i obraz ◊ rząd macierzy ◊ formuła wymiaru ◊ rząd macierzy transponowanej ◊ macierz diagonalna ◊ macierz trójkątna ◊ ślad macierzy ◊ grafy ◊ rekurencje liniowe ◊ macierze elementarne ◊ macierze wierszowo równoważne ◊ rozkład $LU$ ◊ zmiana bazy ◊ macierz przejścia

## 6.1 Transponowanie

W zbiorze macierzy $M_{k\times n}(\mathbb{F})$ zdefiniowaliśmy dodawanie

$$+ : M_{k\times n}(\mathbb{F}) \times M_{k\times n}(\mathbb{F}) \longrightarrow M_{k\times n}(\mathbb{F})$$

oraz mnożenie macierzy przez skalar

$$\cdot : \mathbb{F} \times M_{k\times n}(\mathbb{F}) \longrightarrow M_{k\times n}(\mathbb{F}).$$

Otrzymaliśmy w ten sposób przestrzeń wektorową $(M_{k\times n}(\mathbb{F}), +, \cdot)$. Przypomnijmy, że przyjęliśmy konwencję

$$[x_1, \dots, x_n]^T = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Wektor $[x_1, \dots, x_n]^T \in \mathbb{F}^n$ możemy interpretować jako macierz o $n$-wierszach i jednej kolumnie, czyli element $M_{n\times 1}(\mathbb{F})$. Z drugiej strony $[x_1, \dots, x_n]$ jest w naturalny sposób macierzą o jednym wierszu i $n$-kolumnach, czyli elementem $M_{1\times n}(\mathbb{F})$.

**Definicja 6.1** (Macierz transponowana).
Niech $A \in M_{n\times k}(\mathbb{F})$ będzie macierzą o wierszach $a(1), \dots, a(n) \in M_{1\times k}(\mathbb{F})$. *Macierz transponowana* $A^T \in M_{k\times n}(\mathbb{F})$ jest zdefiniowana jako

$$A^T = \left[ a(1)^T \mid \dots \mid a(n)^T \right].$$

Oznacza to, że wiersze macierzy $A$ stają się kolumnami macierzy transponowanej $A^T$. Wynika stąd, że jeśli $A = [a_{ij}]$ oraz $A^T = [a_{ij}^T]$, to

$$a_{ij}^T = a_{ji}.$$

Przykładowo,

$$\begin{bmatrix} 3 & -2 \\ 2 & 4 \\ 1 & -3 \end{bmatrix}^T = \begin{bmatrix} 3 & 2 & 1 \\ -2 & 4 & -3 \end{bmatrix}.$$

Macierz transponowana

$$A = \begin{bmatrix} - & a(1) & - \\ - & a(2) & - \\ & \vdots & \\ - & a(k) & - \end{bmatrix} \in M_{k\times n}(\mathbb{F}), \qquad A^T = \left[a(1)^T \mid \ldots \mid a(k)^T\right] \in M_{n\times k}(\mathbb{F})$$

gdzie

$$a(i) = [a_{i1}, \ldots, a_{in}] \in M_{1\times n}(\mathbb{F}), \qquad a(i)^T = \begin{bmatrix} a_{i1} \\ \vdots \\ a_{in} \end{bmatrix} \in M_{n\times 1}(\mathbb{F})$$

**Wniosek 6.1.** *Dla macierzy $A, B \in M_{k\times n}(\mathbb{F})$ i $t \in \mathbb{F}$ zachodzą warunki*

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

*(ii)* $(tA)^T = tA^T$,

*(iii)* $(A+B)^T = A^T + B^T$.

U Warunki (ii) i (iii) oznaczają, że transponowanie macierzy

$$L : M_{k\times n}(\mathbb{F}) \ni A \longmapsto A^T \in M_{n\times k}(\mathbb{F})$$

jest odwzorowaniem liniowym i jest to izomorfizm. Warunek (i) gwarantuje dla $k = n$, że $L \circ L = \mathrm{id}$.

**Definicja 6.2** (Macierz symetryczna i antysymetryczna)**.**
Macierz kwadratową $A \in M_{n\times n}(\mathbb{F})$ nazywamy *symetryczną*, gdy

$$A^T = A.$$

Powiemy, że $A$ jest *skośnie symetryczna* (*antysymetryczna*), gdy

$$A^T = -A.$$

**Wniosek 6.2.** *Niech $\mathbb{F} = \mathbb{C}$. Dla dowolnej macierzy $A \in M_{n\times n}(\mathbb{C})$ macierz $A + A^T$ jest symetryczna, a macierz $A - A^T$ jest skośnie symetryczna. W szczególności, każda macierz $A \in M_{n\times n}$ jest sumą macierzy symetrycznej i skośnie symetrycznej:*

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T).$$

Podamy teraz geometryczną interpretację rzeczywistej macierzy transponowanej.

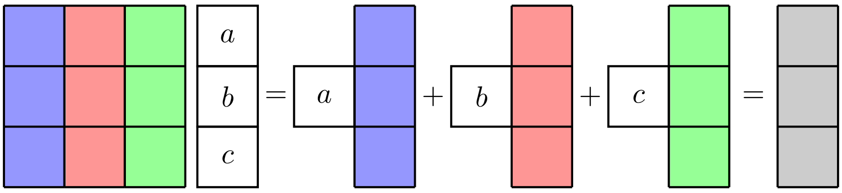
**Lemat 6.1.** *Niech $A \in M_{n\times k}(\mathbb{R})$ oraz $B \in M_{k\times n}(\mathbb{R})$. Następujące warunki są równoważne*

*(i)* $(Ax|y) = (x|By)$ *dla wszystkich* $x \in \mathbb{R}^k$, $y \in \mathbb{R}^n$,

*(ii)* $B = A^T$.

*Dowód.* Dla $x = [x_1, \ldots, x_k]^T \in \mathbb{R}^k$, $y = [y_1, \ldots, y_n]^T \in \mathbb{R}^n$ mamy

$$\begin{aligned}(Ax|y) &= \sum_{i=1}^{k}\sum_{j=1}^{n} x_i y_j (Ae_i|e_j) \\ &= \sum_{i=1}^{k}\sum_{j=1}^{n} x_i y_j (a_i|e_j) \\ &= \sum_{i=1}^{k}\sum_{j=1}^{n} x_i y_j a_{ji}\end{aligned}$$




Rysunek 6.1: Macierz mnoży wektor – wersja kolumnowa.

oraz

$$\begin{aligned}(x|By) &= \sum_{i=1}^{k}\sum_{j=1}^{n} x_i y_j (e_i|Be_j)\\ &= \sum_{i=1}^{k}\sum_{j=1}^{n} x_i y_j (e_i|b_j)\\ &= \sum_{i=1}^{k}\sum_{j=1}^{n} x_i y_j b_{ij}.\end{aligned}$$

Wynika stąd, że $(Ax|y) = (x|By)$ dla wszystkich $x \in \mathbb{R}^k$ i $y \in \mathbb{R}^n$ wtedy i tylko wtedy, gdy $b_{ij} = a_{ji}$ dla wszystkich $i = 1, \ldots, k$ oraz $j = 1, \ldots n$, czyli gdy $B = A^T$. $\square$

**Wniosek 6.3.** *Dla dowolnej macierzy $A = [a_{ij}] \in M_{n\times n}(\mathbb{R})$ mamy*

$$(Ax|y) = (x|A^T y), \quad x, y \in \mathbb{R}^n.$$

*Ponadto, $A \in M_{n\times n}(\mathbb{R})$ jest symetryczna wtedy i tylko wtedy, gdy*

$$(Ax|y) = (x|Ay), \quad x, y \in \mathbb{R}^n.$$

## 6.2 Iloczyn macierzy

Rozważmy macierze

$$a = \begin{bmatrix} a_{11} & \ldots & a_{1k}\end{bmatrix} \in M_{1\times k}(\mathbb{F}), \quad b = \begin{bmatrix} b_{11}\\ \vdots \\ b_{k1}\end{bmatrix} \in M_{k\times 1}(\mathbb{F}).$$

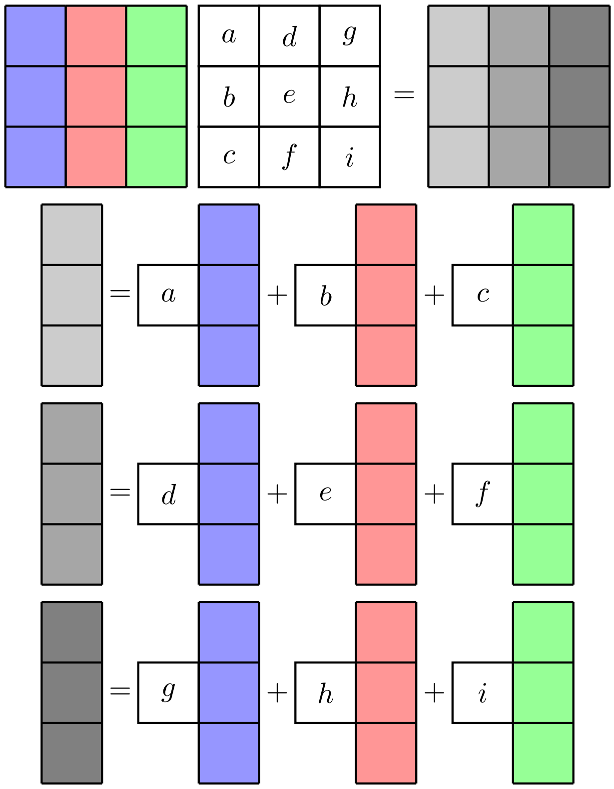
Definiujemy *iloczyn*

$$(a^T|b) = \begin{bmatrix} a_{11} & \ldots & a_{1k}\end{bmatrix}\begin{bmatrix} b_{11}\\ \vdots \\ b_{k1}\end{bmatrix} = a_{11}b_{11} + a_{12}b_{21} + \ldots + a_{1k}b_{k1}.$$

Dla macierzy $A = \begin{bmatrix} a_1 \mid \ldots \mid a_k\end{bmatrix} \in M_{n\times k}(\mathbb{F})$ o wierszach $a(1), \ldots, a(n) \in M_{1\times k}(F)$ oraz macierzy (wektora) $b = [b_1, \ldots, b_k]^T \in M_{k\times 1}(\mathbb{R})$ przyjmujemy, że

$$Ab = \begin{bmatrix} (a(1)^T|b)\\ \vdots \\ (a(n)^T|b)\end{bmatrix} = b_1 \cdot a_1 + \ldots + b_k \cdot a_k \in \mathbb{R}^n.$$

**Definicja 6.3** (Iloczyn macierzy)**.**
Niech $A \in M_{k\times n}(\mathbb{F})$ i $B = \begin{bmatrix} b_1 \mid \ldots \mid b_m\end{bmatrix} \in M_{n\times m}(\mathbb{F})$. Definiujemy *iloczyn macierzy $AB \in M_{k\times m}(\mathbb{F})$* wzorem




Rysunek 6.2: Macierz mnoży macierz w wersji kolumnowej.

$$AB := \Big[ Ab_1 \mid \ldots \mid Ab_m \Big] \in M_{k\times m}(\mathbb{F}).$$

Z definicji wynika, że jeśli $A = [a_{ij}] \in M_{k\times n}(\mathbb{F})$, $B = [b_{jl}] \in M_{n\times m}(\mathbb{F})$ i $AB = [c_{il}] \in M_{k\times m}(\mathbb{F})$, to

$$\begin{aligned} c_{il} &= (a(i)^T | b_l) \\ &= a_{i1}b_{1l} + \ldots + a_{in}b_{nl} \end{aligned}$$

$$\begin{aligned} AB &= \begin{bmatrix} - & w_1 & - \\ - & w_2 & - \\ & \vdots & \\ - & w_k & - \end{bmatrix} \begin{bmatrix} | & & | \\ a_1 & \ldots & a_n \\ | & & | \end{bmatrix} \\ &= \begin{bmatrix} (w_1^T|a_1) & (w_1^T|a_2) & \ldots & (w_1^T|a_n) \\ (w_2^T|a_1) & (w_2^T|a_2) & \ldots & (w_2^T|a_n) \\ \vdots & & \ddots & \vdots \\ (w_k^T|a_1) & (w_k^T|a_2) & \ldots & (w_k^T|a_n) \end{bmatrix} \end{aligned}$$

Wyraz o numerze $ij$ macierzy $AB$ jest iloczynem skalarnym $i$-tego wiersza $A$ oraz $j$-tej kolumny macierzy $B$.

**Przykład 6.1.** Dla macierzy

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

mamy

$$
\begin{aligned}
AB &= \begin{bmatrix} 3 & -2 \\ 2 & 4 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} -2 & 1 & 3 \\ 4 & 1 & 6 \end{bmatrix} \\
&= \begin{bmatrix} 3\cdot(-2) - 2\cdot 4 & 3\cdot 1 - 2\cdot 1 & 3\cdot 3 - 2\cdot 6 \\ 2\cdot(-2) + 4\cdot 4 & 2\cdot 1 + 4\cdot 1 & 2\cdot 3 + 4\cdot 6 \\ 1\cdot(-2) - 3\cdot 4 & 1\cdot 1 - 3\cdot 1 & 1\cdot 3 - 3\cdot 6 \end{bmatrix}.
\end{aligned}
$$

Czasami stosuje się poniższą konwencję:

$$
\begin{array}{cc}
 & \overbrace{\begin{bmatrix} -2 & 1 & 3 \\ 4 & 1 & 6 \end{bmatrix}}^{B} \\
\underbrace{\begin{bmatrix} 3 & -2 \\ 2 & 4 \\ 1 & -3 \end{bmatrix}}_{A} & \underbrace{\begin{bmatrix} -16 & 1 & -3 \\ 12 & 6 & 30 \\ -14 & -2 & -15 \end{bmatrix}}_{AB}
\end{array}
$$

**Przykład 6.2.** Uzupełnić pozostałe miejsca $\boxed{?}$.

$$
\begin{bmatrix} 1 & 2 & 4 \\ \mathbf{2} & \mathbf{6} & \mathbf{0} \end{bmatrix} \begin{bmatrix} 4 & 1 & \mathbf{4} & 3 \\ 0 & -1 & \mathbf{3} & 1 \\ 2 & 7 & \mathbf{5} & 2 \end{bmatrix} = \begin{bmatrix} \boxed{?} & \boxed{?} & \boxed{?} & \boxed{?} \\ \boxed{?} & \boxed{?} & \boxed{26} & \boxed{?} \end{bmatrix}
$$

Dla macierzy kwadratowych $A, B \in M_{n\times n}(\mathbb{R})$ zdefiniowane są obydwa iloczyny $AB$ i $BA$, ale nie muszą one być równe. Mnożenie macierzy w zbiorze $M_{n\times n}(\mathbb{R})$ nie jest przemienne dla $n \geqslant 2$. Przykładowo, dla

$$
A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix},
$$

mamy

$$
\begin{aligned}
AB &= \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 0 & 0 \end{bmatrix}, \\
BA &= \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}.
\end{aligned}
$$

**Lemat 6.2** (Mnożenie macierzy jest łączne)**.** *Załóżmy, że* $A \in M_{m\times n}(\mathbb{F})$*,* $B \in M_{n\times r}(\mathbb{F})$ *i* $C \in M_{r\times s}(\mathbb{F})$*. Wtedy*

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

*Dowód.* Z definicji iloczynu macierzy mamy

$$
\begin{aligned}
(AB)C &= \Big[ ABc_1 \mid \ldots \mid ABc_s \Big] \\
&= A \Big[ Bc_1 \mid \ldots \mid Bc_s \Big] \\
&= A(BC).
\end{aligned}
$$

□

**Lemat 6.3.** *Niech* $A \in M_{k\times n}(\mathbb{F})$ *i* $B \in M_{n\times m}(\mathbb{F})$*. Wtedy*

$$
L_{AB} = L_A \circ L_B
$$

*Dowód.* Wystarczy pokazać, że odwzorowania liniowe $L_{AB}$ oraz $L_A \circ L_B : \mathbb{F}^m \to \mathbb{F}^k$ przyjmują takie same wartości na bazie $e_1, \ldots, e_m$. Mamy

$$
\begin{aligned}
L_{AB}(e_i) &= ABe_i = Ab_i \\
&= L_A(b_i) = L_A(L_B(e_i)) \\
&= (L_A \circ L_B)e_i.
\end{aligned}
$$

□

**Definicja 6.4** (Macierz identycznościowa)**.**
Macierz identycznościowa $I = I_n = [\delta_{ij}] \in M_{n\times n}(\mathbb{F})$ to macierz o współczynnikach zdefiniowanych *symbolem Kroneckera*:

$$\delta_{ij} := \begin{cases} 1, & \text{gdy } i = j, \\ 0, & \text{gdy } i \neq j. \end{cases}$$

Innymi słowy, $I_n = \left[e_1 \mid \ldots \mid e_n\right]$. Przykładowo,

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

**Wniosek 6.4.** *Dla dowolnej macierzy $A \in M_{k\times n}(\mathbb{F})$ mamy*

$$I_k A = A = A I_n.$$

**Definicja 6.5** (Macierz nieosobliwa)**.**
Macierz $A \in M_{n\times n}(\mathbb{F})$ jest *nieosobliwa* (*odwracalna*) wtedy i tylko wtedy, gdy istnieje taka macierz $B \in M_{n\times n}(\mathbb{F})$, że

$$BA = AB = I_n$$

Macierz $B$ nazywamy wtedy *macierzą odwrotną* do macierzy $A$ i oznaczamy przez $A^{-1}$.

U Jeżeli $A$ jest nieosobliwa, to macierz odwrotna do $A$ jest wyznaczona jednoznacznie. Rzeczywiście, jeśli $B$ i $C$ są odwrotne do $A$, to

$$B = BI = B(AC) = (BA)C = IC = C.$$

**Wniosek 6.5** (Odwrotność iloczynu)**.** *Jeśli $A, B \in M_{n\times n}(\mathbb{F})$ są nieosobliwe, to macierz $AB$ jest nieosobliwa oraz*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

*Dowód.* Sprawdzamy, że

$$\begin{aligned}(B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B \\ &= B^{-1}IB \\ &= B^{-1}B \\ &= I.\end{aligned}$$

Analogicznie pokazujemy, że $(AB)(B^{-1}A^{-1}) = I$. □

**Lemat 6.4.** *Dla macierzy kwadratowej $A \in M_{n\times n}(\mathbb{F})$ następujące warunki są równoważne:*

*(i) $A$ jest nieosobliwa,*

*(ii) $L_A : \mathbb{R}^n \ni x \longmapsto Ax \in \mathbb{R}^n$ jest izomorfizmem.*

*Dowód.* Jeśli $A$ jest nieosobliwa, to $AA^{-1} = A^{-1}A = I$, więc

$$I = L_I = L_{AA^{-1}} = L_A \circ L_{A^{-1}}, \quad I = L_I = L_{A^{-1}A} = L_{A^{-1}} \circ L_A,$$

czyli $L_A$ jest izomorfizmem liniowym o odwrotnym $L_{A^{-1}}$.

Załóżmy, że $L_A : \mathbb{R}^n \to \mathbb{R}^n$ jest izomorfizmem liniowym. Istnieje więc odwzorowanie odwrotne $L^{-1}$. Definiujemy macierz

$$B = \left[L^{-1}e_1 \mid \ldots \mid L^{-1}e_n\right].$$

Wtedy $L^{-1} = L_B$, bo $L^{-1}e_i = Be_i = L_B(e_i)$. Ponadto,

$$I = L^{-1} \circ L = L_B \circ L_A = L_{BA}, \quad I = L \circ L^{-1} = L_A \circ L_B = L_{AB},$$

więc $BA = I = AB$, czyli $A$ jest nieosobliwa i $B = A^{-1}$. □

**Lemat 6.5** (Transpozycja iloczynu)**.** *Dla macierzy* $A, B \in M_{k\times n}(\mathbb{F})$ *i* $B \in M_{n\times k}(\mathbb{F})$ *mamy*

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

*Dowód.* Porównamy elementy na pozycji $il$ po obu stronach równości. Dla macierzy $(AB)^T$ na pozycji $il$ stoi element $\sum_{j=1}^{n} a_{lj}b_{ji}$. Dla macierz $B^T A^T$ jest to element $\sum_{j=1}^{n} b_{ji}a_{lj}$. □

Wprowadzimy jeszcze kilka pojęć związanych z macierzami kwadratowymi.

**Definicja 6.6** (Macierz diagonalna)**.**
Macierz $A \in M_{n\times n}(\mathbb{F})$ nazywamy *diagonalną*, jeśli $a_{ij} = 0$ dla $i \neq j$. Będziemy wtedy czasem pisać $A = \mathrm{diag}(a_{11}, \ldots, a_{nn})$.

**Definicja 6.7** (Macierz górnie/dolnie trójkątna)**.**
Macierz $A \in M_{n\times n}(\mathbb{F})$ nazywamy *górnie trójkątną*, gdy $a_{ij} = 0$ dla $i > j$. Analogicznie, $A$ jest *dolnie trójkątna*, jeśli $a_{ij} = 0$ dla $i < j$.

**Przykład 6.3.** Macierz diagonalna, górnie trójkątna i dolnie trójkątna:

$$\begin{bmatrix} \mathbf{a_{11}} & 0 & 0 \\ 0 & \mathbf{a_{22}} & 0 \\ 0 & 0 & \mathbf{a_{33}} \end{bmatrix} \qquad \begin{bmatrix} \mathbf{a_{11}} & \mathbf{a_{12}} & \mathbf{a_{13}} \\ 0 & \mathbf{a_{22}} & \mathbf{a_{23}} \\ 0 & 0 & \mathbf{a_{33}} \end{bmatrix} \qquad \begin{bmatrix} \mathbf{a_{11}} & 0 & 0 \\ \mathbf{a_{21}} & \mathbf{a_{22}} & 0 \\ \mathbf{a_{31}} & \mathbf{a_{32}} & \mathbf{a_{33}} \end{bmatrix}.$$

U

Macierze górnie trójkątne mają użyteczną charakteryzację geometryczną. Niech $e_1, \ldots, e_n$ będzie bazą standardową $\mathbb{F}^n$. Dla $i = 1, \ldots, n$ rozważamy podprzestrzeń

$$V_i = \mathrm{span}\{e_1, \ldots, e_i\}$$

generowaną przez pierwszych $i$ wektorów bazowych. Bezpośrednio z definicji, wynika, że macierz $A$ jest górnie trójkątna wtedy i tylko wtedy, gdy dla każdego $i = 1, \ldots, n$ podprzestrzeń $V_i$ jest niezmiennicza dla $A$ tzn. jeśli $v \in V_i$, to $Av \in V_i$. Wynika, to z faktu, że $Ae_1, \ldots, Ae_i \in V_i$ dla każdego $i$ dokładnie wtedy, gdy $A$ jest górnie trójkątna.

**Wniosek 6.6.** *Macierz* $A \in M_{n\times n}(\mathbb{F})$ *jest diagonalna wtedy i tylko wtedy, gdy* $A$ *jest dolnie i górnie trójkątna.*

**Definicja 6.8** (Ślad macierzy)**.**
*Ślad* $\mathrm{tr}(A)$ *macierzy* $A \in M_{n\times n}(\mathbb{F})$ definiujemy jako liczbę

$$\mathrm{tr}(A) = a_{11} + \ldots + a_{nn}$$

## 6.3 Rząd i jądro macierzy

Rozważmy macierz $A \in M_{k\times n}(\mathbb{F})$ i skojarzone odwzorowanie liniowe $L_A(v) = Av$ dla $v \in \mathbb{F}^n$. Przypomnijmy, że jeśli $v = [x_1, \ldots, x_n]^T = x_1 \cdot e_1 + \ldots + x_n \cdot e_n$, to

$$\begin{aligned} L_A(v) &= Av \\ &= \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & & & \vdots \\ a_{k1} & a_{k2} & \ldots & a_{kn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &= \begin{bmatrix} x_1 a_{11} + \ldots + x_n a_{1n} \\ \vdots \\ x_1 a_{k1} + \ldots + x_n a_{kn} \end{bmatrix} \\ &= x_1 \cdot a_1 + \ldots + x_n \cdot a_n. \end{aligned}$$

Wynika stąd, że obraz $L_A(\mathbb{F}^n)$ jest generowany przez kolumny macierzy $A$, czyli

$$L_A(\mathbb{F}^n) = \mathrm{span}\,\{a_1, \ldots, a_n\}.$$

**Definicja 6.9** (Rząd macierzy).
*Rząd* $\operatorname{rank} A$ *macierzy* $A$ definiujemy jako wymiar obrazu

$$L_A(\mathbb{F}^n) = \operatorname{span}\{a_1, \dots, a_n\},$$

czyli $\operatorname{rank} A$ jest (maksymalną) liczbą liniowo niezależnych kolumn macierzy $A$.

Rząd macierzy został wprowadzony w 1878 przez Georga Frobeniusa (1849–1917).

**Definicja 6.10** (Jądro macierzy).
*Jądro* $\ker A$ *macierzy* $A$ definiujemy jako jądro odwzorowania $L_A$. Jest to więc zbiór takich wektorów $v \in \mathbb{F}^n$, że $Av = 0$.

Jądro macierzy zostało wprowadzone w 1884 przez Jamesa Josepha Sylvestera (1814–1897).

**Wniosek 6.7** (Formuła wymiaru dla macierzy). *Dla* $A \in M_{k\times n}(\mathbb{F})$ *mamy*

$$n = \dim\ker A + \operatorname{rank} A.$$

**Wniosek 6.8.** *Dla macierzy kwadratowej* $A \in M_{n\times n}(\mathbb{F})$ *następujące warunki są równoważne*

*(1)* $A$ *jest nieosobliwa,*

*(2)* $\ker A = \{0\}$,

*(3)* $\operatorname{rank} A = n$.

*W szczególności, macierz* $A \in M_{n\times n}(\mathbb{F})$ *jest odwracalna wtedy i tylko wtedy, gdy jej kolumny są liniowo niezależne.*

**Lemat 6.6.** *Dla dowolnej macierzy* $A \in M_{k\times n}(\mathbb{F})$ *maksymalna liczba liniowo niezależnych kolumn macierzy* $A$ *jest większa bądź równa od maksymalnej liczby liniowo niezależnych wierszy macierzy* $A$. *W szczególności,*

$$\operatorname{rank} A \geqslant \operatorname{rank} A^T.$$

*Dowód.* Niech $1 \leqslant r \leqslant k$ będzie maksymalną liczbą liniowo niezależnych wierszy macierzy $A$. Niech $v_1, \dots, v_r$ będą takimi wierszami $A$, że wektory $v_1^T, \dots, v_r^T$ są liniowo niezależne. Pokażemy, że $Av_1^T, \dots, Av_r^T$ są liniowo niezależne. Przypuśćmy, że

$$x_1 \cdot Av_1^T + \dots + x_r \cdot Av_r^T = 0,$$

dla pewnych skalarów $x_i \in \mathbb{F}$. Wtedy dla $v = x_1 \cdot v_1^T + \dots + x_r v_r^T$ mamy

$$\begin{aligned} Av &= A(x_1 \cdot v_1^T + \dots + x_r \cdot v_r^T) \\ &= x_1 \cdot Av_1^T + \dots + x_r \cdot Av_r^T \\ &= 0. \end{aligned}$$

Z definicji iloczynu $Av$ oznacza to, że $(v_i^T|v) = 0$ dla $i = 1, \dots, r$. Ponieważ $v \in \operatorname{span}\{v_1^T, \dots, v_r^T\}$, więc $(v|v) = 0$. Stąd $v = 0$, czyli $x_1 = \dots = x_r = 0$. Oznacza to, że $Av_1^T, \dots, Av_r^T$ są liniowo niezależne. Ponieważ należą one do $L_A(\mathbb{F}^n)$, więc z definicji rzędu wynika teza. □

**Twierdzenie 6.1** (Rząd macierzy transponowanej). *Dla dowolnej macierzy* $A \in M_{k\times n}(\mathbb{F})$ *zachodzi równość*

$$\operatorname{rank} A = \operatorname{rank} A^T.$$

*Dowód.* Stosujemy lemat 6.6 do macierzy $A$ i $A^T$. □

**Wniosek 6.9.** *Dla dowolnej macierzy* $A \in M_{k\times n}(\mathbb{F})$ *maksymalna liczba liniowo niezależnych kolumn jest równa maksymalnej liczbie liniowo niezależnych wierszy.*

**Przykład 6.4** (Macierze elementarne)**.** Wprowadzimy teraz pojęcie macierzy *elementarnych*. Będą to macierze nieosobliwe $M$, które zastosowane do układu (6.1) skutkują operacjami elementarnymi na wierszach macierzy $A$. Zrobimy to przykładowo dla $k = n = 3$.

Startujemy z macierzy identycznościowej

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Macierze elementarne powstają z macierzy $I$ przez zastosowanie do $I$ operacji elementarnych na wierszach macierzy. Będziemy więc mieć do czynienia z trzema typami macierzy elementarnych.

**Typ I.** Macierze elementarne odpowiadające zamianie miejscami wierszy w macierzy $I$. Rozważmy przykładowo macierz

$$E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

powstałą z $I$ przez zamianę pierwszego i drugiego wiersza. Zauważmy, że dla $A \in M_{3\times 3}(\mathbb{R})$ mamy

$$E_1A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{\mathbf{11}} & a_{\mathbf{12}} & a_{\mathbf{13}} \\ a_{\mathbf{21}} & a_{\mathbf{22}} & a_{\mathbf{23}} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{\mathbf{21}} & a_{\mathbf{22}} & a_{\mathbf{23}} \\ a_{\mathbf{11}} & a_{\mathbf{12}} & a_{\mathbf{13}} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

$$AE_1 = \begin{bmatrix} a_{\mathbf{11}} & a_{\mathbf{12}} & a_{13} \\ a_{\mathbf{21}} & a_{\mathbf{22}} & a_{\mathbf{23}} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{\mathbf{12}} & a_{\mathbf{11}} & a_{13} \\ a_{\mathbf{22}} & a_{\mathbf{21}} & a_{23} \\ a_{\mathbf{32}} & a_{\mathbf{31}} & a_{33} \end{bmatrix},$$

czyli iloczyn $E_1A$ odpowiada wykonanej operacji na wierszach macierzy $I$. Iloczyn $AE_1$ permutuje pierwszą i drugą kolumnę.

**Typ II.** Są to macierze elementarne otrzymane z $I$ przez przemnożenie jej wiersza przez niezerową liczbę rzeczywistą. Przykładowo,

$$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Wtedy

$$E_2A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \mathbf{3} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \mathbf{3}a_{31} & \mathbf{3}a_{32} & \mathbf{3}a_{33} \end{bmatrix}.$$

$$AE_2 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \mathbf{3} \end{bmatrix} = \begin{bmatrix} a_{12} & a_{11} & \mathbf{3}a_{13} \\ a_{22} & a_{21} & \mathbf{3}a_{23} \\ a_{32} & a_{31} & \mathbf{3}a_{33} \end{bmatrix}.$$

**Typ III.** Macierze elementarne powstałe z $I$ przez dodanie do pewnego jej wiersza innego wiersza pomnożonego przez liczbę rzeczywistą. Dla przykładu,

$$E_3 = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Mamy

$$E_3A = \begin{bmatrix} 1 & 0 & \mathbf{3} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} + \mathbf{3}a_{31} & a_{12} + \mathbf{3}a_{32} & a_{13} + \mathbf{3}a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

$$AE_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & \mathbf{3} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{12} & a_{11} & \mathbf{3}a_{11} + a_{13} \\ a_{22} & a_{21} & \mathbf{3}a_{21} + a_{23} \\ a_{32} & a_{31} & \mathbf{3}a_{31} + a_{33} \end{bmatrix}.$$

**Wniosek 6.10.** *Jeśli $E$ jest macierzą elementarną, to $E$ jest nieosobliwa i $E^{-1}$ jest macierzą elementarną tego samego typu.*

**Definicja 6.11** (Macierze wierszowo równoważne)**.**

Niech $A$, $B \in M_{k\times n}(\mathbb{F})$. Powiemy, że $A$ jest *wierszowo równoważna* z $B$, jeśli istnieje taki ciąg macierzy elementarnych $E_1, E_2, \ldots, E_k$, że

$$B = E_k \ldots E_1 A$$

Jest to relacja równoważności w zbiorze $M_{k\times n}(\mathbb{F})$.

**Wniosek 6.11.** *Jeśli $A$, $B \in M_{k\times n}(\mathbb{F})$ są wierszowo równoważne, to* $\operatorname{rank} A = \operatorname{rank} B$.

**Wniosek 6.12.** *Macierz $A \in M_{k\times n}(\mathbb{F})$ jest wierszowo równoważna ze swoją postacią schodkową i zredukowaną postacią schodkową otrzymanymi metodą eliminacji Gaussa.*

**Wniosek 6.13.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą kwadratową. Następujące warunki są równoważne:*

*(1) $A$ jest nieosobliwa,*

*(2) $A$ jest wierszowo równoważna z macierzą identycznościową.*

*Dowód.* Załóżmy, że zachodzi warunek (1). Wtedy układ $Ax = 0$ ma tylko zerowe rozwiązanie. Stosując metodę eliminacji Gaussa możemy sprowadzić macierz $A$ do postaci schodkowej $U$. Oczywiście, $U$ jest wierszowo równoważna z $A$. Jeśli któryś z elementów diagonalnych $U$ jest zerem, to ostatni wiersz macierzy $U$ jest zerowy. Wtedy $\ker U \neq \{0\}$, więc $Ux = 0$ ma niezerowe rozwiązania. Otrzymujemy sprzeczność, bo $Ax = 0$ i $Ux = 0$ mają takie same zbiory rozwiązań. Macierz $U$ jest więc macierzą górnie trójkątną ze wszystkimi elementami na przekątnej równymi 1. Taka macierz jest oczywiście wierszowo równoważna z macierzą $I$, bo $I$ jest jej zredukowaną postacią schodkową.

Jeśli zachodzi warunek (2), to istnieje taki ciąg macierzy elementarnych $E_1, E_2, \ldots, E_k$, że

$$I = E_k \ldots E_1 A.$$

Stąd $A^{-1} = E_k \ldots E_1$. $\square$

**U** W dowodzie powyższym skorzystaliśmy z następującego faktu: jeśli $A, B \in M_{n\times n}(\mathbb{R})$ są macierzami kwadratowymi i $BA = I$, to $A$ jest nieosobliwa i $A^{-1} = B$. Musimy tu być trochę ostrożni. Jak wiemy mnożenie macierzy nie jest przemienne. Potencjalnie mogłoby się więc zdarzyć, że $AB \neq I$. Pokażemy, że jednak $AB = I$. Uzasadnimy najpierw, że $A$ jest nieosobliwa. Wystarczy pokazać, że $\ker A = \{0\}$. Jeśli $v \in \ker A$, to

$$v = Iv = BAv = B0 = 0,$$

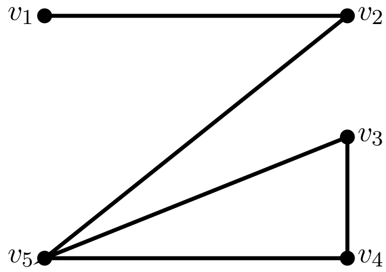
więc $v = 0$. Istnieje więc macierz odwrotna $A^{-1}$. Wtedy

$$BA = I \Rightarrow (BA)A^{-1} = A^{-1} \Rightarrow B = B(AA^{-1}) = A^{-1}.$$

**U** Jak wiemy, macierz $A \in M_{n\times n}(\mathbb{R})$ jest nieosobliwa wtedy i tylko wtedy, gdy $A$ jest wierszowo równoważna z $I$. Ponadto, $A^{-1} = E_k \ldots E_1$ dla pewnych macierzy elementarnych. Wynika stąd, że macierz $\left[A \mid I\right]$ możemy z pomocą operacji elementarnych sprowadzić do postaci $\left[I \mid A^{-1}\right]$. Pozwala to na wyznaczenie macierzy odwrotnej $A^{-1}$.

**Przykład 6.5.** Znajdziemy macierz odwrotną do macierzy

$$\begin{bmatrix} 1 & 4 & 3 \\ -1 & -2 & 0 \\ 2 & 2 & 3 \end{bmatrix}$$




Rysunek 6.3: Przykładowy graf o pięciu wierzchołkach.

wykorzystując elementarne operacje na jej wierszach. Tworzymy macierz $\left[A \mid I\right]$. Jeżeli przekształcimy ją w macierz $\left[I \mid B\right]$, to wtedy $B = A^{-1}$. W naszym przypadku

$$\left[\begin{array}{ccc|ccc} 1 & 4 & 3 & 1 & 0 & 0 \\ -1 & -2 & 0 & 0 & 1 & 0 \\ 2 & 2 & 3 & 0 & 0 & 1 \end{array}\right] \dashrightarrow \left[\begin{array}{ccc|ccc} 1 & 4 & 3 & 1 & 0 & 0 \\ 0 & 2 & 3 & 1 & 1 & 0 \\ 0 & -6 & -3 & -2 & 0 & 1 \end{array}\right]$$

$$\dashrightarrow \left[\begin{array}{ccc|ccc} 1 & 4 & 3 & 1 & 0 & 0 \\ 0 & 2 & 3 & 1 & 1 & 0 \\ 0 & 0 & 6 & 1 & 3 & 1 \end{array}\right]$$

$$\dashrightarrow \left[\begin{array}{ccc|ccc} 1 & 4 & 0 & \frac{1}{2} & -\frac{3}{2} & -\frac{1}{2} \\ 0 & 2 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 6 & 1 & 3 & 1 \end{array}\right]$$

$$\dashrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 2 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 6 & 1 & 3 & 1 \end{array}\right]$$

$$\dashrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{array}\right]$$

więc

$$A^{-1} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{bmatrix}.$$

**Przykład 6.6** (Macierze i grafy)**.** Teoria grafów pełni bardzo ważną rolę dla zastosowań matematyki. *Graf $G$* jest zdefiniowany jako skończony zbiór *wierzchołków*, czyli pewien $n$-elementowy zbiór skończony $\{v_1, \ldots, v_n\}$ oraz zbiór *krawędzi*, czyli pewien zbiór par wierzchołków. Przykładowo, na rysunku 6.3 przedstawiono graf o pięciu wierzchołkach $v_1, v_2, v_3, v_4, v_5$ oraz krawędziach

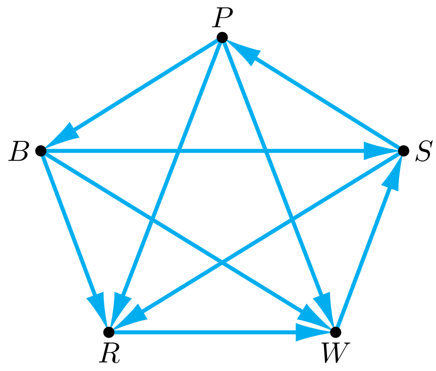
$$\{v_1, v_2\}, \quad \{v_2, v_5\}, \quad \{v_3, v_4\}, \quad \{v_3, v_5\}, \quad \{v_4, v_5\}.$$

Z grafem $G$ o $n$ wierzchołkach możemy skojarzyć pewną macierz $A = [a_{ij}] \in M_{n\times n}(\mathbb{R})$, zwaną *macierzą połączeń* w grafie $G$. Jest ona zdefiniowana regułą:

$$a_{ij} = \begin{cases} 1, & \text{gdy } \{v_i, v_j\} \text{ jest krawędzią,} \\ 0, & \text{w przeciwnym razie.} \end{cases}$$

Dla grafu z rysunku 6.3 mamy

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix}.$$




Rysunek 6.4: Graf skierowany opisujący zwycięstwa drużyn w turnieju.

Możemy myśleć o *ścieżce* w grafie $G$ jako o skończonym ciągu krawędzi od jednego wierzchołka do drugiego. Przykładowo, krawędzie $\{v_1, v_2\}$, $\{v_2, v_5\}$ są pewną ścieżką od wierzchołka $v_1$ do $v_5$. Prostym sposobem określenia ścieżki jest jest opisanie ruchu używając wierzchołków. Powyższa ścieżka ma opis

$$v_1 \to v_2 \to v_5$$

i ma długość 2. Podobnie,

$$v_5 \to v_3 \to v_5 \to v_3$$

jest ścieżką długości 3 z $v_5$ do $v_3$.

Niech $A^k = [a_{ij}^{(k)}] \in M_{n\times n}$. Uzasadnimy, że $a_{ij}^{(k)}$ jest liczbą ścieżek długości $k$ z wierzchołka $v_i$ do wierzchołka $v_j$. Zastosujemy indukcję względem $k$. Dla $k = 1$ teza zachodzi z definicji macierzy $A$. Załóżmy, że teza zachodzi dla pewnej liczby naturalnej $m$, czyli $a_{il}^{(m)}$ jest liczbą ścieżek długości $m$ z wierzchołka $v_i$ do wierzchołka $v_l$. Jeśli istnieje krawędź $\{v_l, v_j\}$, to $a_{il}^{(m)} a_{lj} = a_{il}^{(m)}$ jest liczbą ścieżek długości $m + 1$ z $v_i$ do $v_j$ postaci

$$v_1 \to \ldots \to v_l \to v_j.$$

Ogólna liczba ścieżek długości $m + 1$ z $v_i$ do $v_j$ jest równa,

$$a_{i1}^{(m)} a_{1j} + \ldots + a_{in}^{(m)} a_{nj},$$

ale ta liczba to $a_{ij}^{(m+1)}$ z definicji iloczynu macierzy.

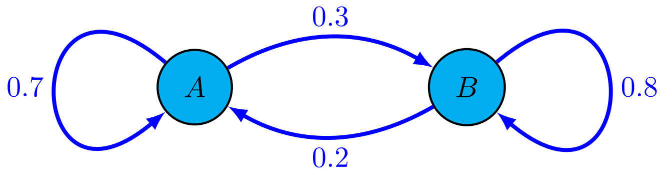
W przypadku grafu z rysunku 6.3 mamy

$$A^3 = \begin{bmatrix} 0 & 2 & 1 & 1 & 0 \\ 2 & 0 & 1 & 1 & 4 \\ 1 & 1 & 2 & 3 & 4 \\ 1 & 1 & 3 & 2 & 4 \\ 0 & 4 & 4 & 4 & 2 \end{bmatrix}.$$

Przykładowo, liczba ścieżek długości 3 z $v_3$ do $v_5$ jest więc równa $a_{35}^{(3)} = 4$. Zauważmy, że macierz $A^3$ (podobnie jak $A^k$) jest symetryczna. Odzwierciedla to fakt, że liczba ścieżek długości 3 z $v_i$ do $v_j$ jest równa liczbie ścieżek długości 3 z $v_j$ do $v_i$.

Powyższe rozważania pozostają prawdziwe dla grafów skierowanych w których krawędzie mają kierunek. Oznacza to, że może istnieć krawędź z wierzchołka $v_i$ do $v_j$, ale niekoniecznie również z $v_j$ do $v_i$. Macierz $A$ nie musi być wtedy symetryczna.

**Przykład 6.7.** Pięć reprezentacji siatkarskich: Polska, Brazylia, Stany Zjednoczone, Rosja i Włochy rozgrywały turniej, grając mecze każdy z każdym. Chcemy ustalić ranking drużyn, w którym liczą się tylko zwycięstwa – nieważne jakim stosunkiem setów. Tworzymy macierz $A$, w której wierszach kodujemy wyniki kolejnych reprezentacji: piszemy 1, jeśli reprezentacja




Rysunek 6.5: Łańcuch Markowa.

wygrała mecz, i 0, jeśli nie wygrała (na przekątnej są zera, bo reprezentacje nie grają ze sobą). Załóżmy, że $A$ ma postać

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Jest to macierz skojarzona z grafem skierowanym, którego wierzchołkami jest pięć reprezentacji, a krawędzie są zadane zwycięstwami. Przykładowo, z wierzchołka Polska wychodzą krawędzie do wierzchołków Brazylia, Rosja i Włochy. Z kolei, Polska jest końcem krawędzi o początku w wierzchołku Stany Zjednoczone. Przykładowo, pierwszy wiersz oznacza, że Polska wygrała z Brazylią, Rosją i Włochami, a przegrała ze Stanami Zjednoczonymi. Rosja (czwarty wiersz) wygrała tylko z Włochami. Liczbę zwycięstw poszczególnych drużyn otrzymujemy, obliczając

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 2 \\ 1 \\ 1 \end{bmatrix}.$$

Oznacza to, że najlepsze w takim rankingu są Polska i Brazylia z trzema zwycięstwami. Następne są Stany Zjednoczone, a ostanie są Rosja i Włochy z jednym zwycięstwem.

Zauważmy, że Polska może argumentować, że jest najlepszą drużyną bo wygrała z Brazylią. Rosja wygrała z Włochami, ale Włochy mogą powiedzieć, że pokonały Stany Zjednoczone, które wygrały z Polską i Rosją. Włochy mają dwa „niebezpośrednie” zwycięstwa. Spróbujmy więc policzyć zwycięstwa reprezentacji łącznie z ich „niebezpośrednimi” zwycięstwami. Odpowiada to sumie

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}^2 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 2 & 3 \\ 1 & 0 & 2 & 2 & 2 \\ 1 & 1 & 0 & 2 & 2 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \\ 6 \\ 2 \\ 3 \end{bmatrix}.$$

**Przykład 6.8** (Łańcuchy Markowa)**.** W badaniach preferencji rynkowych dotyczących dwóch marek pasty do zębów $A$ i $B$ bierze udział 200 osób. Z badań wynika, że każdego miesiąca 70 procent użytkowników marki $A$ używa jej w następnym miesiącu, a 30 procent dokonuje zmiany na markę $B$. Wśród użytkowników marki $B$ te proporcje są odpowiednio równe 80 i 20 procent. Załóżmy, że na początku było 120 użytkowników marki $A$ i 80 marki $B$. Zbadamy jak wiele osób będzie używać poszczególnych marek w kolejnych miesiącach. Po miesiącu marki $A$ będzie używać

$$0.7 \cdot 120 + 0.2 \cdot 80 = 100$$

osób, a marki $B$

$$0.3 \cdot 120 + 0.8 \cdot 80 = 100$$

osób. Rozważmy macierz

$$A = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}.$$

Wtedy

$$\begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix} \begin{bmatrix} 120 \\ 80 \end{bmatrix} = \begin{bmatrix} 100 \\ 100 \end{bmatrix}.$$

Po $n$ miesiącach liczbę osób otrzymujemy jako wektor

$$\begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}^n \begin{bmatrix} 120 \\ 80 \end{bmatrix}.$$

Czasami wygodniej zamiast posługiwania się liczbą użytkowników marek 120 i 80 wygodniej jest użyć ich procentowego udziału, czyli zamiast wektora $[120, 80]^T$ używamy wtedy wektora $[0.6, 0.4]^T$. Wtedy

$$\begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix} \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix},$$

$$\begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}^2 \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix} = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.45 \\ 0.55 \end{bmatrix}.$$

**Przykład 6.9** (Macierze Lesliego)**.** Pewien gatunek żuka żyje co najwyżej 3 lata. Podzielimy populację jego samic na trzy grupy:

- *młode*: wiek $0 - 1$,
- *dojrzałe*: wiek $1 - 2$,
- *dorosłe*: wiek $2 - 3$.

Młode z prawdopodobieństwem $1/2$ staną się dojrzałe, dojrzałe z prawdopodobieństwem $1/4$ staną się dorosłymi. Młode nie znoszą jajek, dojrzałe produkują średnio 4 samice, a dorosłe średnio 3 samice. Przypuśćmy, że wśród populacji 100 samic jest 40 młodych, 40 dojrzałych, 20 dorosłych. Spróbujemy przewidzieć populację po 3 latach. Zobaczmy, jak populacja będzie wyglądała po roku. Młodych będzie

$$40 \cdot 4 + 20 \cdot 3 = 220.$$

Dojrzałych będzie tyle ile młodych, które przetrwały, czyli

$$40 \cdot 0.5 = 20,$$

a dorosłych tyle ile dojrzałych, które przetrwały, czyli

$$40 \cdot 0.25 = 10.$$

Rozważmy macierz

$$L = \begin{bmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix}$$

Zauważmy, że

$$\begin{bmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix} \begin{bmatrix} 40 \\ 40 \\ 20 \end{bmatrix} = \begin{bmatrix} 220 \\ 20 \\ 10 \end{bmatrix}$$

Możemy ten proces iterować otrzymując po dwóch latach

$$\begin{bmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix} \begin{bmatrix} 220 \\ 20 \\ 10 \end{bmatrix} = \begin{bmatrix} 110 \\ 110 \\ 5 \end{bmatrix},$$

po trzech latach

$$\begin{bmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix} \begin{bmatrix} 110 \\ 110 \\ 5 \end{bmatrix} = \begin{bmatrix} 455 \\ 55 \\ 27.5 \end{bmatrix}.$$

TEST $\longrightarrow$ Oceń prawdziwość zadań

- Jeśli macierze $A, B \in M_{2\times 2}(\mathbb{R})$ są nieosobliwe, to macierz $A+B$ jest nieosobliwa i $(A+B)^{-1} = A^{-1} + B^{-1}$.
- Dla dowolnych $A, B \in M_{2\times 2}(\mathbb{R})$ zachodzi równość $(A-B)^2 = A^2 - 2AB + B^2$.
- Jeśli $D_1, D_2 \in M_{n\times n}(\mathbb{R})$ są macierzami diagonalnymi, to $D_1D_2 = D_2D_1$.
- Jeśli $A \in M_{n\times n}(\mathbb{R})$ i $B = 3A^4 - 5A^3 + A^2 + 7A + I$, to $AB = BA$.
- Jeśli $A, B \in M_{n\times n}(\mathbb{R})$ są symetryczne, to $AB = BA$ wtedy i tylko wtedy, gdy $AB$ jest symetryczna.
- Jeśli $A = [a_{ij}] \in M_{n\times n}(\mathbb{R})$ jest skośnie symetryczna, to $a_{11} = \ldots = a_{nn} = 0$.
- Iloczyn macierzy górnie trójkątnych jest macierzą górnie trójkątną.
- Macierz górnie trójkątna jest nieosobliwa wtedy i tylko wtedy, gdy wszystkie jej wyrazy diagonalne $a_{ii}$ są niezerowe.
- Jeśli $A, B, C$ są takimi $2\times 2$-macierzami, że $AB = AC$, to $B = C$.

$\longleftarrow$

## 6.4 Układy równań liniowych raz jeszcze

**Wniosek 6.14.** *Niech* $A = \left[a_1 \mid \ldots \mid a_n\right] \in M_{k\times n}(\mathbb{F})$ *i* $b \in \mathbb{F}^k$. *Następujące warunki są równoważne:*

*(i) zbiór rozwiązań układu* $Ax = b$ *jest niepusty,*

*(ii)* $b \in \operatorname{span}\{a_1, \ldots, a_n\}$,

*(iii)* $\operatorname{rank} A = \operatorname{rank}\left[A \mid b\right]$.

**Wniosek 6.15.** *Niech* $A \in M_{k\times n}(\mathbb{R})$. *Wektor zerowy* $0 \in \mathbb{F}^n$ *jest jedynym rozwiązaniem układu jednorodnego* $Ax = 0 \in \mathbb{F}^k$ *wtedy i tylko wtedy, gdy wektory* $a_1, \ldots, a_n$ *są liniowo niezależne. Wtedy* $\operatorname{rank}(A) = n$ *oraz* $k \geqslant n$.

**Twierdzenie 6.2.** *Niech* $A = \left[a_1 \mid \ldots \mid a_n\right] \in M_{k\times n}(\mathbb{F})$. *Następujące warunki są równoważne:*

*(i) Dla każdego wektora* $b \in \mathbb{F}^k$ *układ* $Ax = b$ *ma dokładnie jedno rozwiązanie* $x \in \mathbb{F}^n$,

*(ii)* $k = n$ *i macierz* $A$ *jest nieosobliwa.*

*Dowód.* Załóżmy najpierw, że zachodzi warunek (ii). Ustalmy wektor $b \in \mathbb{F}^k = \mathbb{F}^n$. Mamy pokazać, że istnieje dokładnie jeden taki wektor $x \in \mathbb{F}^n$, że $Ax = b$. Z warunku (ii) istnieje $A^{-1} \in M_{n\times n}(\mathbb{F})$. Wtedy dla $x = A^{-1}b$ mamy

$$Ax = AA^{-1}b = Ib = b,$$

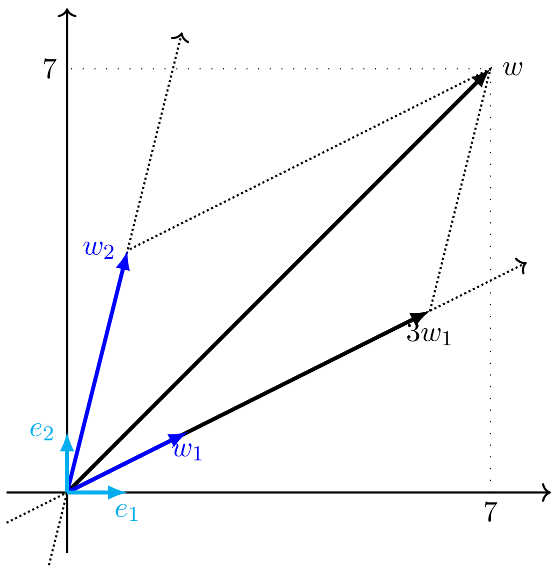
czyli wektor $x = A^{-1}b$ jest rozwiązaniem układu $Ax = b$. Takie rozwiązanie jest jedyne, bo jeśli $Av = b$, to

$$A^{-1}Av = A^{-1}b \Rightarrow v = A^{-1}b.$$

Niech teraz zachodzi warunek (i). Wystarczy pokazać, że wektory $a_1, \ldots, a_n$ są bazą $\mathbb{F}^k$. Zauważmy, że warunek (i) implkuje, że

$$\operatorname{span}\{a_1, \ldots, a_n\} = \mathbb{F}^k.$$

Z drugiej strony, stosując (i) do wektora $b = 0 \in \mathbb{F}^k$, otrzymujemy, że wektory $a_1, \ldots, a_n$ są liniowo niezależne. Stąd $a_1, \ldots, a_n$ tworzą bazę $\mathbb{F}^k$, czyli $k = n$. $\square$




Rysunek 6.6: Zmiana bazy. Baza $w_1 = [2,1]^T$, $w_2 = [1,4]^T$ zadaje nowy układ współrzędnych na płaszczyźnie. Wektor $x = [7,7]^T$ ma w bazie standardowej współrzędne $x_1 = 7$, $x_2 = 7$, bo $w = 7 \cdot e_1 + 7 \cdot e_2$. W bazie $w_1$, $w_2$ wektor $w$ ma współrzędne $c_1 = 3$, $c_2 = 1$, bo $w = 3 \cdot w_1 + 1 \cdot w_2$.

U Dla $A = \big[a_1 \mid \ldots \mid a_n\big] \in M_{k\times n}(\mathbb{F})$ i $b \in \mathbb{F}^k$ rozważmy ponownie układ

$$Ax = b. \tag{6.1}$$

Dla macierzy nieosobliwej $M \in M_{k\times k}(\mathbb{F})$ rozważmy układ

$$MAx = Mb. \tag{6.2}$$

Sprawdzamy łatwo, że układy (6.1) i (6.2) są równoważne tzn. mają równe zbiory rozwiązań. Zamiast więc rozwiązywać układ (6.1), możemy rozwiązać układ (6.2), który przy odpowiednim wyborze macierzy nieosobliwej $M$ może okazać się układem prostszym do analizy.

## 6.5 Zmiana bazy

### 6.5.1 Zmiana bazy w $\mathbb{F}^2$

Dla dowolnego wektora $x = [x_1, x_2]^T \in \mathbb{F}^2$ zachodzi równość

$$x = x_1 \cdot e_1 + x_2 \cdot e_2,$$

gdzie $e_1, e_2$ jest bazą standardową. Skalary $x_1$ i $x_2$ nazywamy *współrzędnymi wektora $x$ w bazie standardowej*. Możemy to pojęcie uogólnić. Załóżmy, że $w_1$, $w_2$ jest ustaloną bazą $\mathbb{F}^2$. Dla dowolnego wektora $w \in \mathbb{F}^2$ istnieją jednoznacznie wyznaczone takie skalary $c_1, c_2 \in \mathbb{F}$, że

$$w = c_1 \cdot w_1 + c_2 \cdot w_2.$$

Wektor $c = [c_1, c_2]^T \in \mathbb{F}^2$ będziemy nazywali *współrzędnymi wektora $w$* w *bazie $w_1$, $w_2$*.

U Zwróćcie uwagę, że mówiąc o bazie $w_1$, $w_2$ mamy tu na myśli bazę uporządkowaną tzn. ciąg wektorów $(w_1, w_2)$. Podając wektor współrzędnych w bazie $c = [c_1, c_2]^T$ istotne jest który wektor bazowy jest pierwszy, a który drugi.

**Przykład 6.10.** Rozważmy bazę $w_1 = [2,1]^T$, $w_2 = [1,4]^T$ w $\mathbb{R}^2$. Łatwo sprawdzamy, że dla wektora $w = [7,7]^T$ mamy

$$w = 3 \cdot w_1 + 1 \cdot w_2.$$

Współrzędne wektora $w = [7,7]^T$ w bazie $w_1, w_2$ są dane wektorem $c = [3,1]^T$.

Niech $w_1$, $w_2$ będzie dowolną bazą $\mathbb{F}^2$. Zajmiemy się teraz następującymi problemami:

(I) Dla danego wektora $x = [x_1, x_2]^T = x_1 \cdot e_1 + x_2 \cdot e_2$ (znamy jego współrzędne w bazie standardowej $e_1$, $e_2$) znajdziemy jego współrzędne $c = [c_1, c_2]^T$ w bazie $w_1$, $w_2$.

(II) Dla danego wektora $c_1 \cdot w_1 + c_2 \cdot w_2$ (znamy jego współrzędne w bazie $w_1$, $w_2$) znajdziemy jego współrzędne $x = [x_1, x_2]^T$ w bazie $e_1$, $e_2$.

Zaczniemy od problemu (II), bo jest on łatwiejszy. Załóżmy, że

$$w_1 = [a_{11}, a_{21}]^T = a_{11} \cdot e_1 + a_{21} \cdot e_2, \quad w_2 = [a_{12}, a_{22}]^T = a_{12} \cdot e_1 + a_{22} \cdot e_2.$$

Znamy więc współrzędne wektorów bazowych $w_1$, $w_2$ w bazie $e_1$, $e_2$. Zauważmy, że

$$\begin{aligned} c_1 \cdot w_1 + c_2 \cdot w_2 &= (a_{11}c_1 \cdot e_1 + a_{21}c_1 \cdot e_2) + (a_{12}c_2 \cdot e_1 + a_{22}c_2 \cdot e_2) \\ &= (a_{11}c_1 + a_{12}c_2) \cdot e_1 + (a_{21}c_1 + a_{22}c_2) \cdot e_2. \end{aligned}$$

Wynika stąd, że współrzędne wektora $c_1 \cdot w_1 + c_2 \cdot w_2$ w bazie standardowej $e_1$, $e_2$ są dane wektorem

$$\begin{aligned} x &= [a_{11}c_1 + a_{12}c_2, a_{21}c_1 + a_{22}c_2]^T \\ &= \begin{bmatrix} a_{11}c_1 + a_{12}c_2 \\ a_{21}c_1 + a_{22}c_2 \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}. \end{aligned}$$

Prowadzi nas to do następującej konkluzji. Jeśli $c = [c_1, c_2]^T$ są współrzędnymi wektora $w$ w bazie $w_1$, $w_2$ oraz $x = [x_1, x_2]^T$ są współrzędnymi wektora $w$ w bazie $e_1$, $e_2$, to

$$x = Wc, \qquad \text{gdzie } W = \begin{bmatrix} w_1 \mid w_2 \end{bmatrix} \in M_{2\times 2}(\mathbb{F}).$$

**Definicja 6.12** (Macierz przejścia do bazy standardowej)**.**
Macierz $W = \begin{bmatrix} w_1 \mid w_2 \end{bmatrix}$ nazywamy *macierzą przejścia* od bazy $w_1, w_2$ do bazy standardowej $e_1, e_2$.

Możemy teraz łatwo rozwiązać problem (I). Ponieważ macierz $W = \begin{bmatrix} w_1 \mid w_2 \end{bmatrix}$ jest nieosobliwa, więc

$$c = W^{-1}x.$$

Przypuśćmy teraz, że mamy dwie dowolne bazy $w_1, w_2$ oraz $u_1, u_2$ w $\mathbb{F}^2$. Rozważmy wektor

$$w = c_1 \cdot w_1 + c_2 \cdot w_2 = c_1^* \cdot u_1 + c_2^* \cdot u_2.$$

Znajdziemy związek pomiędzy współrzędnymi $c = [c_1, c_2]^T$ wektora $w$ w bazie $w_1, w_2$, a jego współrzędnymi $c^* = [c_1^*, c_2^*]^T$ w bazie $u_1, u_2$. Niech $[x_1, x_2]^T$ będą współrzędnymi wektora $w$ w bazie standardowej, czyli $w = x_1 \cdot e_1 + x_2 \cdot e_2$. Jak wiemy dla macierzy $W = \begin{bmatrix} w_1 \mid w_2 \end{bmatrix}$ oraz $U = \begin{bmatrix} u_1 \mid u_2 \end{bmatrix}$ mamy

$$Wc = [x_1, x_2]^T = Uc^*.$$

Stąd

$$c = W^{-1}Uc^*, \quad c^* = U^{-1}Wc.$$

Macierz $S := U^{-1}W$ możemy zinterpretować jeszcze inaczej. Niech $S = \begin{bmatrix} s_1 \mid s_2 \end{bmatrix}$. Zobaczymy jak wyglądają kolumny $s_1$ i $s_2$ macierzy $S$. Oczywiście

$$s_1 = Se_1, \quad s_2 = Se_2.$$

Dla wektora $w = w_1$ mamy $c = e_1$, bo $w_1 = 1 \cdot w_1 + 0 \cdot w_2$, czyli $s_1 = c^*$ są współrzędnymi wektora $w_1$ w bazie $u_1, u_2$, więc

$$w_1 = s_{11} \cdot u_1 + s_{21} \cdot u_2.$$

Analogicznie, $w_2 = s_{12} \cdot u_1 + s_{22} \cdot u_2$, czyli

$$S = U^{-1}W = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}.$$

Nazywamy ją *macierzą przejścia* od bazy $w_1, w_2$ do bazy $u_1, u_2$.

### 6.5.2 Zmiana bazy w $\mathbb{F}^n$

**Definicja 6.13** (Współrzędne wektora w bazie)**.**
Niech $V$ będzie przestrzenią wektorową wymiaru $n$ (nad ciałem $\mathbb{F}$) i $w_1, \ldots, w_n$ pewną bazą w $V$. Dowolny wektor $w \in V$ możemy jednoznacznie zapisać jako kombinację liniową

$$w = c_1 \cdot w_1 + \ldots + c_n \cdot w_n.$$

Wektor $c = [c_1, \ldots, c_n]^T \in \mathbb{F}^n$ nazywamy *współrzędnymi* wektora $w$ w bazie (uporządkowanej) $w_1, \ldots, w_n$.

**Definicja 6.14** (Macierz przejścia – zmiana bazy)**.**
Niech $w_1, \ldots, w_n$ oraz $u_1, \ldots, u_n$ będą bazami $V$. Istnieją jednoznacznie wyznaczone takie skalary $s_{ij}$, że

$$\begin{aligned} w_1 &= s_{11} \cdot u_1 + s_{21} \cdot u_2 + \ldots + s_{n1} \cdot u_n, \\ &\vdots \\ w_n &= s_{1n} \cdot u_1 + s_{2n} \cdot u_2 + \ldots + s_{nn} \cdot u_n. \end{aligned}$$

Macierz $S \in M_{n\times n}(\mathbb{F})$ daną przez

$$S = \Big[ s_1 \mid \ldots \mid s_n \Big] = \begin{bmatrix} s_{11} & s_{12} & \ldots & s_{1n} \\ s_{21} & s_{22} & \ldots & s_{2n} \\ \vdots & \vdots & & \vdots \\ s_{n1} & s_{n2} & \ldots & s_{nn} \end{bmatrix}$$

nazywamy *macierzą przejścia* od bazy $w_1, \ldots, w_n$ do bazy $u_1, \ldots, u_n$. Oznacza to, że $i$-ta kolumna $s_i$ macierzy $S$ jest wektorem współrzędnych wektora $w_i$ w bazie $u_1, \ldots, u_n$.

**Lemat 6.7.** *Niech $c = [c_1, \ldots, c_n]^T, c^* = [c_1^*, \ldots, c_n^*]^T \in \mathbb{F}^n$. Równość*

$$c_1 \cdot w_1 + \ldots + c_n \cdot w_n = c_1^* \cdot u_1 + \ldots + c_n^* \cdot u_n$$

*zachodzi wtedy i tylko wtedy, gdy $c^* = Sc$.*

*Dowód.* Zauważmy, że

$$c_1 \cdot w_1 + \ldots + c_n \cdot w_n = \left( \sum_{j=1}^{n} s_{1j}c_j \right) \cdot u_1 + \ldots + \left( \sum_{j=1}^{n} s_{nj}c_j \right) \cdot u_n,$$

więc teza zachodzi. □

**Wniosek 6.16.** *Macierz przejścia $S \in M_{n\times n}(\mathbb{F})$ jest nieosobliwa.*

*Dowód.* Wystarczy pokazać, że jedynym rozwiązaniem $Sc = 0$ jest $c = 0$. Jeśli $Sc = 0$, to $c^* = 0$, czyli

$$c_1 \cdot w_1 + \ldots + c_n \cdot w_n = 0 \cdot u_1 + \ldots + 0 \cdot u_n = 0,$$

więc $c_1 = \ldots = c_n = 0$, bo $w_1, \ldots, w_n$ są liniowo niezależne. □

**Przykład 6.11.** Znajdziemy macierz przejścia od bazy 1, $x$, $x^2$ dla $P_2$ do bazy 1, $2x$, $4x^2-2$. Wygodnie jest znaleźć najpierw macierz przejścia $S$ od bazy 1, $2x$, $4x^2-2$ do bazy 1, $x$, $x^2$. Ponieważ

$$\begin{aligned} 1 &= 1\cdot 1 + 0\cdot x + 0\cdot x^2 \\ 2\cdot x &= 0\cdot 1 + 2\cdot x + 0\cdot x^2 \\ 4\cdot x^2 &= -2\cdot 1 + 0\cdot x + 4\cdot x^2 \end{aligned}$$

więc

$$S = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}.$$

Macierz przejścia od bazy 1, $x$, $x^2$ dla $P_2$ do bazy 1, $2\cdot x$, $4\cdot x^2-2$ jest równa

$$S^{-1} = \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/4 \end{bmatrix}.$$

Zmiana bazy: od dowolnej do standardowej

Niech $v_1,\dots,v_n$ będzie bazą $\mathbb{F}^n$.
$S = \left[v_1 \mid \dots \mid v_n\right]$ jest nieosobliwa. Jeśli

$$v = x_1\cdot e_1 + \ldots + x_n\cdot e_n = c_1\cdot v_1 + \ldots + c_n\cdot v_n,$$

to

$$S[c_1,\dots,c_n]^T = [x_1,\dots,x_n]^T, \quad [c_1,\dots,c_n]^T = S^{-1}[x_1,\dots,x_n]^T.$$

Zmiana bazy. Podsumowanie

Niech $v_1,\dots,v_n$ oraz $w_1,\dots,w_n$ będą bazami $\mathbb{F}^n$.

$$v_i = s_{1i}\cdot w_1 + \ldots + s_{ni}\cdot w_n, \quad i=1,\dots,n.$$

Macierz

$$S = \left[s_1 \mid \dots \mid s_n\right], \quad s_i = [s_{1i},\dots,s_{ni}]^T$$

jest nieosobliwa.
Jeśli

$$v = x_1\cdot w_1 + \ldots + x_n\cdot w_n = c_1\cdot v_1 + \ldots + c_n\cdot v_n,$$

to

$$S[c_1,\dots,c_n]^T = [x_1,\dots,x_n]^T.$$

TEST → Oceń które ze zdań jest prawdziwe:

- Transpozycja macierzy górnie trójkątnej jest macierzą górnie trójkątną.
- Odwrotność nieosobliwej macierzy górnie trójkątnej jest macierzą dolnie trójkątną.
- Macierz górnie trójkątna i symetryczna jest diagonalna.
- Jeśli $A+B$ jest górnie trójkątna, to $A$ i $B$ są górnie trójkątne.
- Jeśli $A^2$ jest symetryczna, to $A$ jest symetryczna.
- Jeśli $A$ jest nieosobliwa, to $A^TA$ i $AA^T$ są nieosobliwe.
- Jeśli $A$, $B \in M_{n\times n}(\mathbb{F})$ są nieosobliwe, to $AB$ jest nieosobliwa.
- Jeśli $A$, $B \in M_{n\times n}(\mathbb{F})$ są nieosobliwe, to $A+B$ jest nieosobliwa.
- Dla dowolnych $n\times n$ macierzy kwadratowych zachodzi $A^2-B^2=(A-B)(A+B)$.

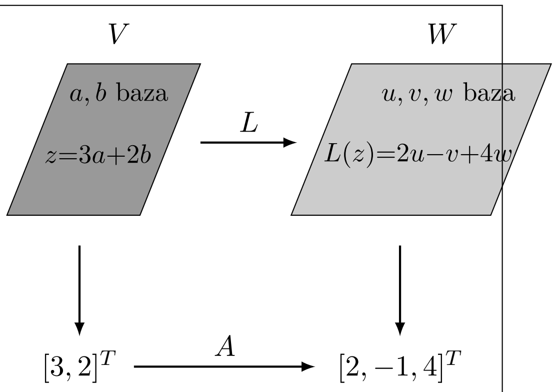
- Dla dowolnej macierzy kwadratowej $I - A^2 = (I - A)(I + A)$.
- Istnieją takie macierze nieosobliwe, że $(AB)^{-1} = A^{-1}B^{-1}$.
- Istnieje taka macierz $A \in M_{2\times 2}(\mathbb{R})$, że $\ker A = \operatorname{im} A$.
- Istnieje taka macierz $A \in M_{3\times 3}(\mathbb{R})$, że $\ker A = \operatorname{im} A$.

$\longleftarrow$

# Rozdział 7

# Reprezentacja macierzowa

Słowa kluczowe rozdziału
macierz odwzorowania liniowego ◊ macierze podobne ◊ ślad iloczynu ◊ ślad macierzy podobnych




Reprezentacja odwzorowania liniowego.

## 7.1 Macierz odwzorowania liniowego

**Lemat 7.1.** *Niech $L : \mathbb{F}^n \longrightarrow \mathbb{F}^m$ będzie odwzorowaniem liniowym. Istnieje taka macierz $A \in M_{m\times n}(\mathbb{F})$, że*

$$L(v) = Av, \quad v \in \mathbb{F}^n.$$

*Ponadto,*

$$A = \Big[L(e_1) \mid \ldots \mid L(e_n)\Big].$$

*Dowód.* Definiujemy wektory

$$a_i = L(e_i), \quad i = 1, \ldots, n$$

oraz macierz $A := \Big[a_1 \mid \ldots \mid a_n\Big]$. Niech $v = x_1 \cdot e_1 + \ldots + x_n \cdot e_n$ będzie dowolnym wektorem w $\mathbb{F}^n$. Wtedy

$$\begin{aligned} L(v) &= x_1 \cdot L(e_1) + \ldots + x_n \cdot L(e_n) \\ &= x_1 \cdot a_1 + \ldots + x_n \cdot a_n \\ &= Av. \end{aligned}$$

□

**Definicja 7.1** (Macierz standardowa odwzorowania liniowego)**.**
Niech $L : \mathbb{F}^n \longrightarrow \mathbb{F}^m$ będzie odwzorowaniem liniowym. Macierz

$$A = \Big[L(e_1) \mid \ldots \mid L(e_n)\Big]$$

nazywamy *macierzą odwzorowania liniowego* $L$ w bazach standardowych lub *macierzą standardową* $L$.

**Przykład 7.1.** Niech $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ będzie odwzorowaniem liniowym danym przez

$$L([x_1, x_2, x_3]^T) = [x_1 - x_2 + x_3, 2x_2 + 3x_3]^T.$$

Wtedy
$$L([1,0,0]^T) = [1,0]^T, \quad L([0,1,0]^T) = [-1,2]^T, \quad L([0,0,1]^T) = [1,3]^T,$$
czyli macierz standardowa odwzorowania $L$ jest równa
$$A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & 3 \end{bmatrix}.$$

U Załóżmy, że $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $G : \mathbb{R}^m \longrightarrow \mathbb{R}^k$. Niech $A \in M_{m\times n}(\mathbb{R})$ będzie standardową reprezentacją macierzową dla $L$, a $B \in M_{k\times m}(\mathbb{R})$ dla $G$. Wtedy $BA \in M_{k\times n}(\mathbb{R})$ jest reprezentacją macierzową dla $G \circ L : \mathbb{R}^n \longrightarrow \mathbb{R}^k$. Rzeczywiście niech $C \in M_{k\times n}(\mathbb{R})$ będzie macierzową reprezentacją odwzorowania $G \circ L : \mathbb{R}^n \longrightarrow \mathbb{R}^k$. Z definicji, $C = \begin{bmatrix} c_1 \mid \ldots \mid c_n \end{bmatrix}$, gdzie $c_i = (G \circ L)(e_i)$ dla $i = 1, \ldots, n$. Zauważmy, że
$$\begin{aligned} c_i &= (G \circ L)(e_i) \\ &= G(L(e_i)) = G(a_i) \\ &= G(a_{1i} \cdot e_1 + \ldots + a_{mi} \cdot e_m) \\ &= a_{1i} \cdot G(e_1) + \ldots + a_{mi} \cdot G(e_m) \\ &= a_{1i} \cdot b_1 + \ldots + a_{mi} \cdot b_m = Ba_i, \end{aligned}$$
czyli $c_i$ jest $i$-tą kolumną macierzy $BA$.

Uogólnimy powyższą konstrukcję macierzy skojarzonej z odwzorowaniem liniowym. Niech $V$ będzie przestrzenią wektorową wymiaru $n$ i niech $W$ będzie przestrzenią wektorową wymiaru $m$. Ustalmy bazę
$$v_1, \ldots, v_n$$
dla przestrzeni $V$ oraz bazę
$$w_1, \ldots, w_m$$
dla przestrzeni $W$. Dla wektora
$$v = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$$
wektor
$$x = [x_1, \ldots, x_n]^T \in \mathbb{F}^n$$
zadaje współrzędne wektora $v$ w bazie $v_1, \ldots, v_n$. Pokażemy, że istnieje (jedyna) taka macierz $A \in M_{m\times n}(\mathbb{F})$, że
$$Ax = y$$
wtedy i tylko wtedy, gdy
$$L(v) = y_1 \cdot w_1 + y_2 \cdot w_2 + \ldots + y_n \cdot w_m$$
dla każdego $v \in V$, czyli $y$ jest wektorem współrzędnych $L(v)$ w bazie $w_1, \ldots, w_m$.

Istnieją wyznaczone jednoznacznie takie skalary $a_{ij}$, że
$$L(v_j) = a_{1j} \cdot w_1 + a_{2j} \cdot w_2 + \ldots + a_{mj} \cdot w_m, \quad 1 \leqslant j \leqslant n.$$
Wektor
$$a_j = [a_{1j}, a_{2j}, \ldots, a_{mj}]^T \in \mathbb{F}^m, \quad j = 1, \ldots, n$$
zadaje współrzędne wektora $L(v_j)$ w bazie $w_1, \ldots, w_m$ dla $W$.

**Definicja 7.2** (Macierz odwzorowania liniowego)**.**

Macierz $A = \begin{bmatrix} a_1 \mid \ldots \mid a_n \end{bmatrix} \in M_{m\times n}(\mathbb{F})$ nazywamy *macierzą odwzorowania liniowego* $L : V \longrightarrow W$ w bazach $v_1, \ldots, v_n$ dla $V$ oraz $w_1, \ldots, w_m$ dla $W$, jeśli
$$a_j = [a_{1j}, a_{2j}, \ldots, a_{mj}]^T \in \mathbb{F}^m,$$
gdzie
$$L(v_j) = a_{1j} \cdot w_1 + a_{2j} \cdot w_2 + \ldots + a_{mj} \cdot w_m, \quad 1 \leqslant j \leqslant n.$$

**Wniosek 7.1.** *Niech $x = [x_1, \ldots, x_n]^T \in \mathbb{F}^n$, $y = [y_1, \ldots, y_m]^T \in \mathbb{F}^m$. Następujące warunki są równoważne:*

*(1)* $v = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$ *i* $L(v) = y_1 \cdot w_1 + \ldots + y_m \cdot w_m$*,*

*(2)* $Ax = y$*.*

*Dowód.* Wystarczy zauważyć, że

$$\begin{aligned} L(v) &= L(x_1 \cdot v_1 + \ldots + x_n \cdot v_n) \\ &= \sum_{j=1}^{n} x_j L(v_j) \\ &= \sum_{j=1}^{n} x_j \left( \sum_{i=1}^{m} a_{ij} w_i \right) \\ &= \sum_{i=1}^{m} \left( \sum_{j=1}^{n} a_{ij} x_j \right) w_i. \end{aligned}$$

□

**Wniosek 7.2.** *Załóżmy, że $v_1, \ldots, v_n$ jest bazą $V$ i $L : V \to V$ jest odwzorowaniem liniowym. Niech $A \in M_{n \times n}(\mathbb{F})$ będzie macierzą odwzorowania $L$ w bazie $v_1, \ldots, v_n$. Następujące warunki są równoważne*

*(1)* $v = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$ *i* $L(v) = y_1 \cdot v_1 + \ldots + x_n \cdot v_n$*,*

*(2)* $Ax = y$*.*

*Oznacza to, że macierz $A$ przekształca współrzędne wektora $v$ w bazie $v_1, \ldots, v_n$ na współrzędne wektora $L(v)$ w tej bazie.*

U Zauważmy, że jeżeli $v_1, \ldots, v_n$ oraz $w_1, \ldots w_n$ są bazami przestrzeni $V$, to macierz przejścia $S$ od bazy $v_1, \ldots, v_n$ do bazy $w_1, \ldots w_n$ jest macierzą odwzorowania identycznościowego na $V$ w tych bazach.

U Popatrzmy na konstrukcję macierzy odwzorowania liniowego jeszcze trochę inaczej. Niech $v_1, \ldots, v_n$ będzie bazą dla $V$ oraz $w_1, \ldots, w_m$ bazą dla $W$. Rozważmy izomorfizmy

$$R : V \ni x_1 \cdot v_1 + \ldots + x_n \cdot v_n \longmapsto [x_1, \ldots, x_n]^T \in \mathbb{R}^n,$$

$$S : W \ni y_1 \cdot v_1 + \ldots + y_m \cdot v_m \longmapsto [y_1, \ldots, y_m]^T \in \mathbb{R}^m,$$

czyli przykładowo $R(v)$ jest wektorem współrzędnych $v$ w bazie $v_1, \ldots, v_n$.
Dla odwzorowania liniowego $L : V \to W$ możemy rozważyć odwzorowanie liniowe

$$S \circ L \circ R^{-1} : \mathbb{R}^n \to \mathbb{R}^m.$$

Zobaczmy jak wygląda jego macierz w bazach standardowych. Mamy

$$\begin{aligned} (S \circ L \circ R^{-1})(e_i) &= (S \circ L)(v_i) \\ &= S(L(v_i)), \end{aligned}$$

czyli $i$-ta kolumna tej macierzy to współrzędne wektora $L(v_i)$ w bazie $w_1, \ldots, w_n$.

**Przykład 7.2.** Niech $P_1$ będzie przestrzenią wielomianów stopnia co najwyżej 1. Rozważmy odwzorowanie liniowe

$$L : P_1 \ni a + bx \longmapsto (a + b)x \in P_1.$$

Znajdziemy jego macierz $A$ w bazie $1 - x, 1 + x$. Ponieważ

$$L(1 - x) = 0 = 0 \cdot (1 - x) + 0 \cdot (1 + x),$$

$$L(1 + x) = 2x = (-1) \cdot (1 - x) + 1 \cdot (1 + x),$$

więc

$$A = \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}.$$

**Lemat 7.2.** *Niech $L : \mathbb{F}^n \longrightarrow \mathbb{F}^m$ będzie odwzorowaniem liniowym. Jeśli*

$$A = \left[a_1 \mid \ldots \mid a_n\right]$$

*jest macierzą $L$ w bazach $u_1, \ldots, u_n$ dla $\mathbb{F}^n$ i $b_1, \ldots, b_m$ dla $\mathbb{F}^m$, to*

$$a_j = B^{-1}L(u_j), \quad j = 1, \ldots, n$$

*gdzie $B = \left[b_1 \mid \ldots \mid b_m\right]$.*

*Dowód.* Z definicji

$$L(u_j) = a_{1j} \cdot b_1 + \ldots + a_{mj} \cdot b_m = Ba_j.$$

□

**Wniosek 7.3.** *Niech $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ będzie odwzorowaniem liniowym. Jeśli*

$$A = \left[a_1 \mid \ldots \mid a_n\right]$$

*jest macierzą $L$ w bazach $u_1, \ldots, u_n$ dla $\mathbb{R}^n$ i $b_1, \ldots, b_m$ dla $\mathbb{R}^m$ oraz $B = \left[b_1 \mid \ldots \mid b_m\right]$, to macierze $\left[B \mid L(u_1) \mid \ldots \mid L(u_n)\right]$ i $\left[I \mid A\right]$ są wierszowo równoważne.*

*Dowód.* Macierze

$$\left[B \mid L(u_1) \mid \ldots \mid L(u_n)\right], \quad B^{-1}\left[B \mid L(u_1) \mid \ldots \mid L(u_n)\right]$$

są wierszowo równoważne. Teza zachodzi, bo

$$\begin{aligned} B^{-1}\left[B \mid L(u_1) \mid \ldots \mid L(u_n)\right] &= \left[I \mid B^{-1}L(u_1) \mid \ldots \mid B^{-1}L(u_n)\right] \\ &= \left[I \mid A\right]. \end{aligned}$$

□

**Przykład 7.3.** Rozważmy odwzorowanie liniowe $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ dane wzorem

$$L[x_1, x_2]^T = [x_2, x_1 + x_2, x_1 - x_2]^T.$$

Znajdziemy jego reprezentację macierzową $A$ w bazach $u_1 = [1, 2]^T$, $u_2 = [3, 1]^T$ oraz

$$b_1 = [1, 0, 0]^T, \quad b_2 = [1, 1, 0]^T, \quad b_3 = [1, 1, 1]^T.$$

Ponieważ $L(u_1) = [2, 3, -1]^T$ i $L(u_2) = [1, 4, 2]^T$, więc rozważamy macierz

$$\left[\begin{array}{ccc|cc} 1 & 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 3 & 4 \\ 0 & 0 & 1 & -1 & 2 \end{array}\right].$$

Sprowadzamy ją łatwo (poprzez dozwolone operacje na wierszach) do postaci

$$\left[\begin{array}{ccc|cc} 1 & 0 & 0 & -1 & -3 \\ 0 & 1 & 0 & 4 & 2 \\ 0 & 0 & 1 & -1 & 2 \end{array}\right].$$

czyli

$$A = \begin{bmatrix} -1 & -3 \\ 4 & 2 \\ -1 & 2 \end{bmatrix}.$$

## 7.2 Macierze podobne

Rozważmy odwzorowanie liniowe $L : \mathbb{F}^2 \longrightarrow \mathbb{F}^2$. Reprezentacja macierzowa $A$ dla $L$ w bazie standardowej $e_1, e_2$ (w dziedzinie i przeciwdziedzinie) ma postać

$$A = \left[a_1 \mid a_2\right], \quad a_1 = L(e_1), \quad a_2 = L(e_2).$$

Rozważmy teraz inną bazę $u_1, u_2$ w $\mathbb{R}^2$. Znajdziemy reprezentację $B$ odwzorowania $L$ w tej bazie. W tym celu musimy znaleźć współrzędne wektorów $L(u_1)$ i $L(u_2)$ w bazie $u_1$, $u_2$, bo z definicji $b_{ij}$ są zadane przez równości

$$L(u_1) = b_{11} \cdot u_1 + b_{21} \cdot u_2, \quad L(u_2) = b_{12} \cdot u_1 + b_{22} \cdot u_2.$$

Współrzędne wektora $L(u_1)$ w bazie standardowej $e_1, e_2$ są równe

$$L(u_1) = Au_1.$$

Jak już wiemy jego współrzędne w bazie $u_1, u_2$ są dane przez

$$U^{-1}Au_1, \quad U = \left[u_1 \mid u_2\right].$$

Analogicznie współrzędne wektora $L(u_2) = Au_2$ w bazie $u_1, u_2$ są równe $U^{-1}Au_2$. Z definicji reprezentacji macierzowej mamy

$$B = \left[U^{-1}Au_1 \mid U^{-1}Au_2\right].$$

Z kolei $\left[U^{-1}Au_1 \mid U^{-1}Au_2\right] = U^{-1}AU$ z definicji mnożenia macierzy, więc

$$B = U^{-1}AU.$$

**Twierdzenie 7.1** (Reprezentacja macierzowa w różnych bazach)**.** *Załóżmy, że $V$ jest przestrzenią wektorową wymiaru $n$ oraz $v_1, \ldots, v_n$ i $w_1, \ldots, w_n$ są bazami $V$. Niech*

- *$L : V \longrightarrow V$ będzie odwzorowaniem liniowym,*
- *$S$ będzie macierzą przejścia od bazy $w_1, \ldots, w_n$ do bazy $v_1, \ldots, v_n$,*
- *$A \in M_{n\times n}(\mathbb{F})$ będzie reprezentacją $L$ w bazie $v_1, \ldots, v_n$,*
- *$B \in M_{n\times n}(\mathbb{F})$ będzie reprezentacją $L$ w bazie $w_1, \ldots, w_n$.*

*Wtedy*

$$B = S^{-1}AS.$$

*Dowód.* Dla wektora $x = [x_1, \ldots, x_n]^T \in \mathbb{F}^n$ rozważmy wektor

$$v = x_1 \cdot w_1 + \ldots + x_n \cdot w_n \in V.$$

Z definicji macierzy przejścia $S$ dla wektora $y = Sx \in \mathbb{R}^n$ mamy

$$v = y_1 \cdot v_1 + \ldots + y_n \cdot v_n.$$

Ponadto, $Ay$ zadaje współrzędne $L(v)$ w bazie $v_1, \ldots, v_n$ oraz $Bx$ zadaje współrzędne $L(v)$ w bazie $w_1, \ldots, w_n$. Stąd

$$S^{-1}Ay = Bx,$$

czyli $S^{-1}ASx = Bx$ dla każdego $x \in \mathbb{F}^n$, więc $S^{-1}AS = B$. □

**Definicja 7.3** (Macierze podobne)**.**

Niech $A, B \in M_{n\times n}(\mathbb{F})$. Mówimy, że $B$ jest *podobna* do $A$, jeśli istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{F})$, że

$$B = S^{-1}AS.$$

Podobieństwo macierzy jest relacją równoważności w $M_{n\times n}(\mathbb{F})$.

**Wniosek 7.4** (Ślad iloczynu macierzy)**.** *Dla dowolnych macierzy* $A \in M_{n\times k}(\mathbb{F})$ *i* $B \in M_{k\times n}(\mathbb{F})$ *mamy*

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

*W szczególności,*

$$\operatorname{tr}(A) = \operatorname{tr}(S^{-1}AS)$$

*dla dowolnej macierzy nieosobliwej* $S \in M_{n\times n}(\mathbb{F})$ *i* $A \in M_{n\times n}(\mathbb{F})$.

*Dowód.* Ślad macierzy kwadratowej jest sumą wyrazów na głównej przekątnej. Z definicji iloczynu macierzy mamy

$$\begin{aligned}
\operatorname{tr}(AB) &= \sum_{j=1}^{n}(AB)_{jj} = \sum_{j=1}^{n}(a(j)^T|b_j) \\
&= \sum_{j=1}^{n}\left(\sum_{i=1}^{k} a_{ji}b_{ij}\right) \\
&= \sum_{j=1}^{n}\sum_{i=1}^{k} a_{ji}b_{ij} = \sum_{i=1}^{k}\left(\sum_{j=1}^{n} b_{ij}a_{ji}\right) \\
&= \sum_{i=1}^{k}(b(i)^T|a_i) = \sum_{i=1}^{k}(BA)_{ii} \\
&= \operatorname{tr}(BA).
\end{aligned}$$

Jeśli $B = S^{-1}AS$, to

$$\begin{aligned}
\operatorname{tr} B &= \operatorname{tr}(S^{-1}A)S \\
&= \operatorname{tr} S(S^{-1}A) \\
&= \operatorname{tr} A.
\end{aligned}$$

□

Pozwala to nam zdefiniować ślad odwzorowania liniowego $L : V \longrightarrow V$ przestrzeni skończenie wymiarowej $V$ jako ślad macierzy $A$ odwzorowania $L$ w dowolnej bazie dla $V$.

U Przypuśćmy, że $L : M_{n\times n}(\mathbb{F}) \longrightarrow \mathbb{F}$ jest takim odwzorowaniem liniowym, że

$$L(AB) = L(BA), \quad A, B \in \mathbb{F}.$$

Pokażemy, że istnieje taka $\lambda \in \mathbb{F}$, że

$$L(A) = \lambda \operatorname{tr} A, \quad A \in M_{n\times n}(\mathbb{F}).$$

Niech $E_{ji}$ ($i, j \in \{1, \dots, n\}$) będzie bazą standardową $M_{n\times n}(\mathbb{F})$, czyli na przecięciu $j$-tego wiersza oraz $i$-tej kolumny jest 1, a pozostałe wyrazy są równe 0. Zauważmy, że

$$E_{lk}E_{ji} = \delta_{kj}E_{li}, \quad i, j, k, l \in \{1, \dots, n\}.$$

Stąd dla $i \neq j$ mamy

$$\begin{aligned}
L(E_{ji}) &= L(E_{jj}E_{ji}) \\
&= L(E_{ji}E_{jj}) \\
&= L(0) \\
&= 0
\end{aligned}$$

oraz

$$\begin{aligned}
L(E_{ii}) &= L(E_{i1}E_{1i}) \\
&= L(E_{1i}E_{i1}) \\
&= L(E_{11}).
\end{aligned}$$

Dla $\lambda = L(E_{11})$ mamy więc

$$L(E_{ji}) = \delta_{ji}\lambda, \quad i, j \in \{1, \dots, n\}.$$

Ponieważ dla $A = [a_{ji}]$ mamy $A = \sum_{i,j=1}^{n} a_{ji}E_{ji}$, więc

$$\begin{aligned} L(A) &= \sum_{i,j=1}^{n} a_{ji}L(E_{ji}) \\ &= \sum_{i,j=1}^{n} a_{ji}\delta_{ji}\lambda \\ &= \sum_{i=1}^{n} a_{ii}\lambda \\ &= \lambda \operatorname{tr} A. \end{aligned}$$

Dowód poniższego twierdzenia jest bardzo podobny do dowodu twierdzenia 7.1.

**Twierdzenie 7.2.** *Załóżmy, że*

- $u_1, \dots, u_p$ *i* $u'_1, \dots, u'_p$ *są bazami przestrzeni wektorowej* $U$*,*
- $v_1, \dots, v_q$ *i* $v'_1, \dots, v'_q$ *są bazami przestrzeni wektorowej* $V$*,*
- $L : U \longrightarrow V$ *jest odwzorowaniem liniowym,*
- $A$ *jest macierzą* $L$ *w bazach* $u_1, \dots, u_p$ *i* $v_1, \dots, v_q$*,*
- $A'$ *jest macierzą* $L$ *w bazach* $u'_1, \dots, u'_p$ *i* $v'_1, \dots, v'_q$*,*
- $P$ *jest macierzą przejścia od bazy* $u_1, \dots, u_p$ *do bazy* $u'_1, \dots, u'_p$*,*
- $Q$ *jest macierzą przejścia od bazy* $v_1, \dots, v_q$ *do bazy* $v'_1, \dots, v'_q$*.*

*Wtedy*

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

TEST $\longrightarrow$ Oceń prawdziwość zdań:

- Jeśli $A$ i $B$ są reprezentacjami odwzorowań liniowych $L, G : \mathbb{R}^n \to \mathbb{R}^n$ w bazie standardowej, to istnieje taka macierz nieosobliwa $S$, że $S^{-1}AS = B$.
- Reprezentacją macierzową odwzorowania identycznościowego jest w dowolnej bazie macierz identycznościowa.
- Odwzorowanie $L : \mathbb{R}^2 \ni [x, y]^T \mapsto [y, x + y]^T \in \mathbb{R}^2$, to $L$ ma w pewnej bazie reprezentację
$$\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}.$$
- Jeśli $L : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ jest izomorfizmem i $A$ jest jego reprezentacją w pewnej bazie, to $A$ jest nieosobliwa.

$\longleftarrow$

## 7.3 Pouczający przykład

Zobrazujemy wprowadzone pojęcia na przykładzie przestrzeni $\mathbb{R}^\infty$ i pewnej jej podprzestrzeni. Przypomnijmy, że

$$\mathbb{R}^\infty = \Big\{(x_1, x_2, x_3, \ldots) : x_i \in \mathbb{R}\Big\},$$

jest przestrzenią wektorową wszystkich ciągów o wyrazach rzeczywistych z działaniami

$$(x_1, x_2, x_3, \ldots) + (y_1, y_2, y_3, \ldots) = (x_1 + y_1, x_2 + y_2, x_3 + y_3, \ldots),$$

$$t \cdot (x_1, x_2, x_3, \ldots) = (tx_1, tx_2, tx_3, \ldots).$$

Dla liczby naturalnej $n \geqslant 1$ definiujemy

$$\mathbb{R}^\infty_n = \Big\{(x_1, x_2, x_3, \ldots) \in \mathbb{R}^\infty : \forall i > n \; x_i = 0\Big\}.$$

Są to podprzestrzenie wektorowe przestrzeni $\mathbb{R}^\infty$ oraz

$$\mathbb{R}^\infty_1 \subset \mathbb{R}^\infty_2 \subset \mathbb{R}^\infty_3 \subset \ldots$$

Odwzorowanie liniowe

$$L_n : \mathbb{R}^\infty_n \ni (x_1, \ldots, x_n, 0, 0, \ldots) \longmapsto [x_1, \ldots, x_n]^T \in \mathbb{R}^n$$

jest izomorfizmem. W szczególności, $\dim \mathbb{R}^\infty_n = n$, czyli przestrzeń $\mathbb{R}^\infty$ zawiera podprzestrzeń wymiaru $n$ dla dowolnej liczby naturalnej.

Zbiór

$$\mathbb{R}^\infty_0 := \bigcup_{n=1}^{\infty} \mathbb{R}^\infty_n$$

jest również podprzestrzenią wektorową przestrzeni $\mathbb{R}^\infty$. Składa się ona z wszystkich ciągów, które mają tylko skończenie wiele niezerowych wyrazów. Naturalnie zdefiniowany nieskończony zbiór wektorów

$$e_1, e_2, e_3, \ldots, e_n, \ldots \in \mathbb{R}^\infty_0$$

składa się z wektorów liniowo niezależnych. Jest on bazą dla $\mathbb{R}^\infty_0$, ale nie jest bazą dla $\mathbb{R}^\infty$. Przestrzeń $\mathbb{R}^\infty$ nie ma przeliczalnej bazy.

Dla niezerowej liczby rzeczywistej $a \in \mathbb{R} \setminus \{0\}$ definiujemy ciąg $x^a \in \mathbb{R}^\infty$ przez

$$x^a = (a, a^2, a^3, \ldots).$$

Uzasadnimy, że każdy skończony podzbiór zbioru (równolicznego z $\mathbb{R}$)

$$\Big\{x^a : a \in \mathbb{R} \setminus \{0\}\Big\}$$

jest zbiorem wektorów liniowo niezależnych.

Rozważmy odwzorowanie liniowe (przesunięcie) $\sigma : \mathbb{R}^\infty \to \mathbb{R}^\infty$ zdefiniowane wzorem

$$\sigma((x_1, x_2, x_3, \ldots)) = (x_2, x_3, x_4, \ldots).$$

Zauważmy, że

$$\sigma(x^a) = a \cdot x^a.$$

Niech $a_1, \ldots, a_n \in \mathbb{R} \setminus \{0\}$ będą różnymi liczbami rzeczywistymi. Pokażemy, że wektory $x^{a_1}, \ldots, x^{a_n} \in \mathbb{R}^\infty$ są liniowo niezależne. Przypuśćmy, że maksymalna liczba wektorów liniowo niezależnych spośród nich jest równa $1 \leqslant r < n$. Możemy założyć, że $x^{a_1}, \ldots, x^{a_r}$ są liniowo niezależne. Z określenia $r$ wynika, że $x^{a_1}, \ldots, x^{a_r}, x^{a_{r+1}}$ są liniowo zależne, czyli

$$x^{a_{r+1}} = t_1 \cdot x^{a_1} + \ldots + t_r \cdot x^{a_r}$$

dla pewnych $t_1, \ldots, t_r \in \mathbb{R}$. Stąd

$$a_{r+1} \cdot x^{a_{r+1}} = \sigma(x^{a_{r+1}}) = t_1 a_1 \cdot x^{a_1} + \ldots + t_r a_r \cdot x^{a_r},$$

więc

$$0 = t_1(a_1 - a_{r+1}) \cdot x^{a_1} + \ldots + t_r(a_r - a_{r+1})x^{a_r}.$$

Ponieważ $x^{a_1}, \ldots, x^{a_r}$ są liniowo niezależne oraz $a_i \neq a_{r+1}$, więc $t_1 = \ldots = t_r = 0$. Prowadzi to do sprzeczności, bo $x^{a_{r+1}} \neq 0$.

Przyglądnijmy się jeszcze odwzorowaniu $\sigma$.

Rozważmy odwzorowanie liniowe $\sigma_0 : \mathbb{R}^\infty \longrightarrow \mathbb{R}^\infty$ zdefiniowane wzorem

$$\sigma_0((x_1, x_2, x_3, \ldots)) = (0, x_1, x_2, x_3, x_4, \ldots).$$

Zauważmy, że

$$\sigma \circ \sigma_0((x_1, x_2, x_3, \ldots)) = ((x_1, x_2, x_3, \ldots)),$$

czyli

$$\sigma \circ \sigma_0 = \mathrm{id}_{\mathbb{R}^\infty}.$$

W szczególności, $\sigma$ jest epimorfizmem i $\sigma_0$ jest monomorfizmem. Żadne z nich nie jest izomorfizmem. Jądro $\ker \sigma$ jest 1-wymiarowe, bo

$$\ker \sigma = \mathbb{R}^\infty_1.$$

Z drugiej strony, $\sigma_0$ nie jest epimorfizmem, bo

$$\mathbb{R}^\infty_1 \not\subset \mathrm{im}\, \sigma_0.$$

Zbadamy teraz jeszcze jedną podprzestrzeń wektorową przestrzeni $\mathbb{R}^\infty$. Powiemy, że ciąg

$$x = (x_1, x_2, x_3, \ldots) \in \mathbb{R}^\infty$$

jest *typu Fibonacciego*, gdy

$$x_{n+1} = x_{n-1} + x_n, \quad n \geqslant 2.$$

Przypomnijmy, że klasyczny ciąg Fibonacciego

$$F = (F_1, F_2, F_3, \ldots) = (0, 1, 1, 2, 3, 5, 8, 13, \ldots)$$

otrzymujemy przyjmując $x_1 = 0$ i $x_2 = 1$.

Niech $\mathcal{F} \subset \mathbb{R}^\infty$ będzie zbiorem wszystkich ciągów typu Fibonacciego. Jest to podprzestrzeń wektorowa przestrzeni $\mathbb{R}^\infty$. Pokażemy, że jest ona izomorficzna z $\mathbb{R}^2$. Szukanym izomorfizmem jest odwzorowanie $L : \mathcal{F} \longrightarrow \mathbb{R}^2$ dane przez

$$L((x_1, x_2, x_3, \ldots)) = [x_1, x_2]^T \in \mathbb{R}^2.$$

Odwzorowanie odwrotne przypisuje wektorowi $[x_1, x_2]^T \in \mathbb{R}^2$ ciąg Fibonacciego o dwóch pierwszych wyrazach $x_1, x_2$.

Bazie standardowej $e_1 = [1, 0]^T$, $e_2 = [0, 1]^T \in \mathbb{R}^2$ odpowiada przez izomorfizm $L$ baza $\mathcal{F}$ złożona z ciągów

$$F' = (1, 0, 1, 1, 2, 3, \ldots), \quad F = (0, 1, 1, 2, 3, \ldots).$$

Ponadto, dla ciągu $(x_1, x_2, x_3, \ldots) \in \mathcal{F}$ mamy

$$(x_1, x_2, x_3, \ldots) = x_1 \cdot F' + x_2 \cdot F.$$

Zauważmy, że

$$\sigma(F') = F, \quad F' = \sigma(F) - F,$$

czyli

$$(x_1, x_2, x_3, \ldots) = x_1 \cdot (\sigma(F) - F) + x_2 \cdot F$$

dla dowolnego ciągu $x = (x_1, x_2, x_3, \ldots) \in \mathcal{F}$. Rozważmy ciąg $x^k = \sigma^k(F)$ dla $k \geqslant 1$. Z definicji odwzorowania przesunięcia wynika, że $n$-ty wyraz ciągu $x^k$ jest równy

$$x_n^k = (\sigma^k(F))_n = F_{n+k}.$$

Ponieważ

$$x_1^k = F_{k+1}, \quad x_2^k = F_{k+2},$$

więc

$$\sigma^k(F) = F_{k+1} \cdot (\sigma(F) - F) + F_{k+2} \cdot F,$$

czyli otrzymujemy równość

$$F_{n+k} = F_{k+1}(F_{n+1} - F_n) + F_{k+2} F_n.$$

Zastanówmy się teraz, czy w przestrzeni $\mathcal{F}$ leży jakiś ciąg geometryczny $(1, q, q^2, q^3, \ldots)$? Sprawdzamy, że musi zachodzić warunek $q^2 = q + 1$, czyli jest tak dla

$$q_1 = \frac{1+\sqrt{5}}{2}, \quad q_2 = \frac{1-\sqrt{5}}{2}.$$

Ponieważ wektory $[1, q_1]^T$, $[1, q_2]^T \in \mathbb{R}^2$ są liniowo niezależne, więc ciągi

$$x = (1, q_1, q_1^2, \ldots), \quad y = (1, q_2, q_2^2, \ldots)$$

tworzą bazę dla $\mathcal{F}$. Zapiszmy ciąg Fibonacciego w tej bazie. Szukamy takich liczb rzeczywistych $a, b \in \mathbb{R}$, że

$$F = a \cdot x + b \cdot y.$$

Liczby $a, b$ spełniają układ równań

$$a + b = 1, \quad aq_1 + bq_2 = 1.$$

Jego jedynym rozwiązaniem są

$$a = \frac{1}{\sqrt{5}}, \quad b = -\frac{1}{\sqrt{5}}.$$

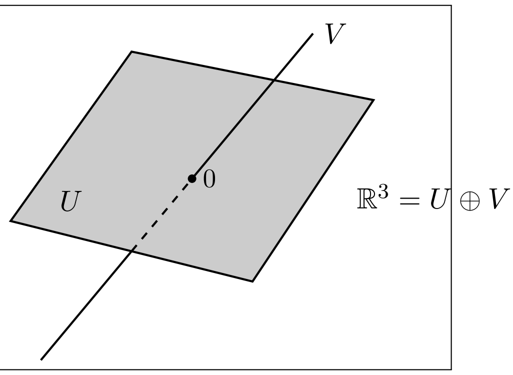
Ostatecznie

$$F_n = \frac{q_1^n - q_2^n}{\sqrt{5}}.$$

# Rozdział 8

# Suma prosta

Słowa kluczowe rozdziału
suma algebraiczna podprzestrzeni ◊ suma prosta ◊ macierze blokowe ◊ podprzestrzenie niezmiennicze




**Definicja 8.1** (Suma algebraiczna i suma prosta)**.**
Załóżmy, że $V$ jest przestrzenią wektorową nad ciałem $\mathbb{F}$. Niech $U_1, U_2 \subset V$ będą podprzestrzeniami wektorowymi $V$. Definiujemy *sumę algebraiczną* $U_1, U_2$ jako

$$U_1 + U_2 = \Big\{u_1 + u_2 : u_i \in U_i,\ i = 1, 2\Big\}$$

Wtedy $U_1 + U_2$ jest podprzestrzenią wektorową $V$. Mówimy, że $V$ jest sumą prostą $U_1$ oraz $U_2$ , jeśli zachodzą warunki

$$V = U_1 + U_2, \quad U_1 \cap U_2 = \{0\}.$$

Piszemy wtedy, że $V = U_1 \oplus U_2$.

**Wniosek 8.1.** *Dla podprzestrzeni $U_1, U_2 \subset V$ przestrzeni wektorowej $V$ następujące warunki są równoważne*

*(i)* $V = U_1 \oplus U_2$*,*

*(ii) dowolny wektor $v \in V$ można jednoznacznie przedstawić w postaci*

$$v = v_1 + v_2,$$

*gdzie $v_i \in U_i$ dla $i = 1, 2$.*

*Dowód.* Załóżmy, że $V = U_1 \oplus U_2$ i niech $v \in V$. Ponieważ $V = U_1 + U_2$, więc istnieją takie wektory $v_i \in U_i$ $(i = 1, 2)$, że $v = v_1 + v_2$. Pokażemy, że takie przedstawienie jest jednoznaczne. Przypuśćmy, że

$$v = v_1 + v_2 = u_1 + u_2,$$

dla pewnych $u_i \in U_i$. Wtedy

$$\underbrace{v_1 - u_1}_{\in U_1} = \underbrace{u_2 - v_2}_{\in U_2} \in U_1 \cap U_2 = \{0\},$$

czyli $v_1 = u_1$ oraz $v_2 = u_2$.

Załóżmy teraz, że zachodzi warunek (ii). Wtedy oczywiście $V = U_1 + U_2$. Pokażemy, że $U_1 \cap U_2 = \{0\}$. Niech $v \in U_1 \cap U_2$. Wtedy

$$v = \underbrace{0}_{\in U_1} + \underbrace{v}_{\in U_2} = \underbrace{v}_{\in U_1} + \underbrace{0}_{\in U_2},$$

czyli z jednoznaczności w warunku (ii) wynika, że $v = 0$. $\square$

**Przykład 8.1.** Rozważmy przestrzeń wektorową $M_{2\times 2}(\mathbb{R})$ i jej dwie podprzestrzenie wektorowe

$$\mathrm{Sym}_2(\mathbb{R}) = \{A \in M_{2\times 2}(\mathbb{R}) : A = A^T\}, \quad \mathrm{Asym}_2(\mathbb{R}) = \{A \in M_{2\times 2}(\mathbb{R}) : A = -A^T\}.$$

Wtedy

$$\mathrm{Sym}_2(\mathbb{R}) \cap \mathrm{Asym}_2(\mathbb{R}) = \{0\}, \quad \mathrm{Sym}_2(\mathbb{R}) + \mathrm{Asym}_2(\mathbb{R}) = M_{2\times 2}(\mathbb{R}),$$

czyli

$$M_{2\times 2}(\mathbb{R}) = \mathrm{Sym}_2(\mathbb{R}) \oplus \mathrm{Asym}_2(\mathbb{R})$$

**Przykład 8.2.** Niech $F(\mathbb{R},\mathbb{R})$ będzie przestrzenią wektorową funkcji $f : \mathbb{R} \longrightarrow \mathbb{R}$. Rozważmy podprzestrzeń $F_o(\mathbb{R},\mathbb{R})$ funkcji nieparzystych i podprzestrzeń $F_e(\mathbb{R},\mathbb{R})$ funkcji parzystych. Sprawdzimy, że

$$F(\mathbb{R},\mathbb{R}) = F_e(\mathbb{R},\mathbb{R}) \oplus F_o(\mathbb{R},\mathbb{R}).$$

Jeśli $f \in F(\mathbb{R},\mathbb{R})$, to

$$f(x) = \underbrace{\frac{1}{2}(f(x) + f(-x))}_{\in F_e(\mathbb{R},\mathbb{R})} + \underbrace{\frac{1}{2}(f(x) - f(-x))}_{\in F_o(\mathbb{R},\mathbb{R})},$$

czyli

$$F(\mathbb{R},\mathbb{R}) = F_e(\mathbb{R},\mathbb{R}) + F_o(\mathbb{R},\mathbb{R}).$$

Jeśli $f \in F_e(\mathbb{R},\mathbb{R}) \cap F_o(\mathbb{R},\mathbb{R})$, to dla $x \in \mathbb{R}$ mamy

$$f(x) = f(-x) = -f(x),$$

więc $f(x) = 0$. W konsekwencji,

$$F_e(\mathbb{R},\mathbb{R}) \cap F_o(\mathbb{R},\mathbb{R}) = \{0\}.$$

Możemy to pojęcie uogólnić na większą liczbę podprzestrzeni $U_1, \ldots, U_k$.

**Definicja 8.2.** Niech $U_1, \ldots, U_k \subset V$ będą podprzestrzeniami wektorowymi $V$. Mówimy, że $V$ jest sumą prostą podprzestrzeni $U_1, \ldots, U_n$, tzn.

$$V = U_1 \oplus \ldots \oplus V_k$$

jeśli każdy wektor $v \in V$ można jednoznacznie przedstawić w postaci sumy

$$v = u_1 + \ldots + u_k,$$

gdzie $u_i \in U_i$ dla $i = 1, \ldots, k$.

**Wniosek 8.2.** *Niech $U_1, U_2 \subset V$ będą podprzestrzeniami przestrzeni wektorowej $V$. Załóżmy, że $v_1, \ldots, v_k$ jest bazą $U_1$ oraz $v_{k+1}, \ldots, v_n$ jest bazą $U_2$. Następujące warunki są równoważne*

*(1)* $V = U_1 \oplus U_2$,

*(2)* $v_1, \ldots, v_n$ *jest bazą* $V$.

*W szczególności, jeśli* $\dim V < \infty$*, to* $V = U_1 \oplus U_2$ *wtedy i tylko wtedy, gdy*

$$\dim V = \dim U_1 + \dim U_2, \quad U_1 \cap U_2 = \{0\}.$$

*Dowód.* Załóżmy, że $V = U_1 \oplus U_2$. Wtedy $\operatorname{span}\{v_1, \ldots, v_n\} = V$, bo $V = U_1 + U_2$. Wektory $v_1, \ldots, v_n$ są liniowo niezależne, bo jeśli

$$x_1 \cdot v_1 + \ldots + x_k \cdot v_k + x_{k+1} \cdot v_{k+1} + \ldots x_n \cdot v_n = 0,$$

to

$$\underbrace{x_1 \cdot v_1 + \ldots + x_k \cdot v_k}_{\in U_1} = \underbrace{-(x_{k+1} \cdot v_{k+1} + \ldots x_n \cdot v_n)}_{\in U_2} \in U_1 \cap U_2 = \{0\},$$

czyli z założenia o bazach dla $U_1$ i $U_2$ mamy, że $x_1 = \ldots = x_k = 0$ oraz $x_{k+1} = \ldots = x_n = 0$.

Jeśli $v_1, \ldots, v_n$ jest bazą dla $V$, to każdy wektor $v \in V$ da się jednoznacznie zapisać w postaci $v = u_1 + u_2$, gdzie $u_1 \in U_1$ i $u_2 \in U_2$, czyli $V = U_1 \oplus U_2$. □

**U** $\boxed{\text{MACIERZE BLOKOWE}}$

Niech $\mathbb{F}^n = U_1 \oplus U_2$ i niech $v_1, \ldots, v_n$ będzie taką bazą $V$, że $v_1, \ldots, v_k \in U_1$ tworzą bazę $U_1$ oraz $v_{k+1}, \ldots, v_n \in U_2$ tworzą bazę $U_2$. Odwzorowanie liniowe $L : \mathbb{F}^n \to \mathbb{F}^n$ ma w tej bazie macierz postać blokową

$$A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right].$$

Przykładowo, jeśli $U_1 = \operatorname{span}\{e_1, e_2, e_3\}$ oraz $U_2 = \operatorname{span}\{e_4, e_5\}$ oraz $A$ jest macierzą odwzorowania liniowego $L : \mathbb{F}^5 \longrightarrow \mathbb{F}^5$ w bazie $e_1, e_2, e_3, e_4, e_5$, to

$$A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] = \left[\begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ \hline a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{array}\right].$$

Zauważmy, że $A_{12} = 0$ wtedy i tylko wtedy, gdy $A(U_2) \subset U_2$, czyli $U_2$ jest niezmiennicza dla $A$. Analogicznie, $A_{21} = 0$ wtedy i tylko wtedy, gdy $A(U_1) \subset U_1$. Jeśli $U_1$ i $U_2$ są niezmiennicze dla $A$, to $A$ ma postać

$$A = \left[\begin{array}{c|c} A_{11} & 0 \\ \hline 0 & A_{22} \end{array}\right] = \left[\begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 \\ \hline 0 & 0 & 0 & a_{44} & a_{45} \\ 0 & 0 & 0 & a_{54} & a_{55} \end{array}\right].$$

**U** Rozważmy macierz $A \in M_{n\times n}(\mathbb{F})$. Jak wiemy

$$n = \dim\ker A + \operatorname{rank} A = \dim\ker A + \dim\operatorname{im} A.$$

Zazwyczaj jednak nie jest prawdą, że $\mathbb{F}^n$ jest sumą prostą podprzestrzeni $\ker A$ oraz $\operatorname{im} A$. Może się nawet zdarzyć, że $\ker A = \operatorname{im} A$. Przykładowo, rozważmy niezerowy wektor $a \in \mathbb{R}^2$ i macierz $A = \big[a \mid a\big]$. Wtedy $\ker A = \operatorname{span}\{[1, -1]^T\}$ niezależenie do wyboru wektora $a$. Przyjmując, że $a = [1, -1]^T$ otrzymujemy, że

$$\ker A = \operatorname{im} A.$$

Z drugiej strony, jeśli $a$ jest liniowo niezależny z $[1, -1]^T$, to

$$\mathbb{R}^2 = \ker A \oplus \operatorname{im} A.$$

Ponadto, gdy $a = [-1, 1]^T$, to proste $\ker A$ i $\operatorname{im} A$ są prostopadłe.

# Rozdział 9

# Wyznacznik

| |
|---|
| Słowa kluczowe rozdziału |
| wyznacznik jako funkcja kolumn ◊ permutacje ◊ wyznacznik iloczynu ◊ wyznacznik macierzy odwrotnej ◊ wyznacznik macierzy transponowanej ◊ wyznacznik a liniowa niezależność ◊ dopełnienie algebraiczne ◊ rozwinięcie Laplace'a ◊ macierz odwrotna ◊ wzory Cramera ◊ rozkład Schura ◊ orientacja ◊ iloczyn wektorowy ◊ iloczyn mieszany ◊ objętość równoległościanu |

## 9.1 Grupa permutacji

**Definicja 9.1** (Permutacja)**.** Rozważmy zbiór skończony $\{1, \ldots, n\}$ dla $n \in \mathbb{N}$. *Permutacją* zbioru $\{1, \ldots, n\}$ nazywamy dowolną bijekcję

$$\sigma : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}.$$

Przez $S_n$ oznaczamy zbiór wszystkich permutacji zbioru $\{1, \ldots, n\}$ tzn.

$$S_n = \Big\{\sigma | \sigma : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\} \quad \text{jest bijekcją}\Big\}$$

Ponieważ złożenie bijekcji jest bijekcją, więc składanie odwzorowań $\circ$ określa strukturę grupy w zbiorze $S_n$. Jej elementem neutralnym jest odwzorowanie identycznościowe id na zbiorze $\{1, \ldots, n\}$. Elementem odwrotnym do permutacji $\sigma$ jest odwzorowanie odwrotne do $\sigma$ (istnieje ono, bo $\sigma$ jest bijekcją).

**Przykład 9.1.** Permutację $\sigma : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$ zapisujemy często w postaci tabeli

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \ldots & \sigma(n) \end{pmatrix}$$

Przykładowo,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 4 & 1 & 5 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 3 & 1 \end{pmatrix}$$

są pewnymi permutacjami zbioru 5-elementowego $\{1, 2, 3, 4, 5\}$. Ich iloczyn jest równy

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

Istnieje bardziej zwarty sposób zapisu permutacji $\sigma \in S_n$. Pochodzi on od Cauchy'ego. Niech $x_1, \ldots, x_r$ będą różnymi elementami zbioru $\{1, \ldots, n\}$. W szczególności, $2 \leqslant r \leqslant n$. Przez $(x_1, \ldots, x_r)$ oznaczamy taką permutację z $S_n$, że

$$x_1 \to x_2 \to x_3 \to \ldots \to x_r \to x_1$$

oraz

$$x \to x, \quad \text{dla} \quad x \in \{1, \ldots, n\} \setminus \{x_1, \ldots, x_r\}.$$

Permutację $(x_1, \ldots, x_r) \in S_n$ nazywamy $r$*-cyklem.* Jeśli $r = 2$, to 2-cykl nazywamy *transpozycją.* Dwa cykle z $S_n$ nazywamy *rozłącznymi*, jeśli nie mają wspólnych wyrazów.

**Przykład 9.2.** Dla permutacji

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 5 & 4 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 1 & 2 & 3 \end{pmatrix}$$

mamy (opuszczamy w zapisie 1-cykle):

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

**Przykład 9.3.** Niech $\sigma = (1,2,3) \in S_n$. Wtedy $\sigma^{-1} = (3,2,1)$, bo

$$(1,2,3)(3,2,1) = \mathrm{id}\,.$$

Ponadto, $(1,2,3)^3 = (1,2,3)(1,2,3)(1,2,3) = \mathrm{id}$ oraz 3 jest najmniejszą taką liczbą naturalną $n$, że $(1,2,3)^n = \mathrm{id}$.

U Jeśli $n > 2$, to grupa $(S_n, \circ)$ nie jest abelowa. Przykładowo,

$$(2,3)(1,2) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = (1,3,2)$$

oraz

$$(1,2)(2,3) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (1,2,3).$$

Dla $\sigma \in S_n$ definiujemy zbiór

$$\sigma^* = \{i \in \{1,\ldots,n\} : \sigma(i) \neq i\}.$$

Zauważmy, że jeśli $i \in \sigma^*$, to $\sigma(i) \in \sigma^*$. Jeśli $\sigma = (x_1,\ldots,x_r) \in S_n$ jest cyklem, to $\sigma^* = \{x_1,\ldots,x_r\}$. Cykle $\sigma$ i $\tau$ są *rozłączne*, gdy $\sigma^* \cap \tau^* = \emptyset$.

**Lemat 9.1.** *Jeśli $\sigma, \tau \in S_n$ i $\sigma^* \cap \tau^* = \emptyset$, to $\sigma\tau = \tau\sigma$.*

*Dowód.* Niech $i \in \{1,\ldots,n\}$ będzie ustalone. Sprawdzimy, że $\sigma\tau(i) = \tau\sigma(i)$. Ponieważ cykle $\tau$ i $\sigma$ są rozłączne, więc mamy trzy możliwości:

(1) $\tau(i) = i = \sigma(i)$,

(2) $\tau(i) \neq i$ oraz $\sigma(i) = i$,

(3) $\sigma(i) \neq i$ oraz $\tau(i) = i$.

Przypadek (1) jest oczywisty, a (2) i (3) są symetryczne, więc można założyć, że zachodzi (2). Ponieważ cykle są rozłączne, więc również $\sigma(\tau(i)) = \tau(i)$, czyli teza zachodzi. □

**Lemat 9.2.** *Każda permutacja $\sigma \in S_n \setminus \{\mathrm{id}\}$ jest iloczynem rozłącznych cykli. Taki rozkład jest jednoznaczny z dokładnością do kolejności czynników.*

*Dowód.* Jeśli $\sigma^* = \emptyset$, to $\sigma = \mathrm{id} = (1)(2)\ldots(n)$, więc teza zachodzi. Niech $\sigma^* \neq \emptyset$ i $s \in \sigma^*$. Rozważmy zbiór

$$\{\sigma^k(s) : k \in \mathbb{N}\}.$$

Ponieważ jest on skończony, więc $\sigma^i(s) = \sigma^j(s)$ dla pewnych liczb naturalnych $i < j$. Wtedy $\sigma^{j-i}(s) = s$. Istnieje więc najmniejsza liczba naturalna $l \geqslant 2$ taka, że $\sigma^l(s) = s$. Wtedy liczby $s, \ldots, \sigma^{l-1}(s)$ są różne z definicji $l$ oraz $\tau = (s, \sigma(s), \ldots, \sigma^{l-1}(s))$ jest $l$-cyklem. Dla $\rho = \sigma\tau^{-1}$ mamy $\sigma = \rho\tau$ oraz $\rho$ nie rusza punktów $s, \sigma(s), \ldots, \sigma^{l-1}(s)$. Stosujemy teraz powyższą procedurę do $\rho$. Zauważmy, że $\rho^*$ jest istotnym podzbiorem $\sigma^*$, więc procedura skończy się po skończonej liczbie kroków, bo $\sigma^*$ jest zbiorem skończonym. Dowodzi to rozkładu $\sigma$ na rozłączne cykle.

Przypuśćmy, że

$$\sigma = \sigma_1 \ldots \sigma_k = \tau_1 \ldots \tau_l$$

są dwoma rozkładami $\sigma$ na rozłączne cykle. Przypuśćmy, że $\sigma_1(i) = j \neq i$. Wtedy również $\tau_p(i) \neq i$ dla pewnego $p$. Z rozłączności cykli mamy $\tau_p(i) = j = \sigma_1(i)$. Rozważmy teraz $\sigma_1(j)$. Z tych samych względów mamy $\tau_p(j) = \sigma_1(j)$. Kontynuując, dostajemy, że $\sigma_1 = \tau_p$ i teza wynika poprzez indukcję. □

**Lemat 9.3.** *Dowolną permutację $\sigma \in S_n$ można przedstawić jako złożenie skończonej liczby transpozycji. Ponadto, jeśli $\sigma$ jest złożeniem parzystej liczby pewnych transpozycji, to każde jej przedstawienie jako złożenie transpozycji (nie jest ono jednoznacznie wyznaczone) składa się z parzystej liczby transpozycji.*

*Dowód.* Wiemy, że $\sigma$ możemy rozłożyć na iloczyn rozłącznych cykli. Każdy cykl $(a_1, \ldots, a_k)$ jest iloczynem transpozycji, bo

$$(a_1, \ldots, a_k) = (a_1, a_k) \ldots (a_1, a_2).$$

Przypuśćmy, że mamy dwa rozkłady $\sigma$ na iloczyny transpozycji:

$$\sigma = \tau_1 \ldots \tau_r = \rho_1 \ldots \rho_p.$$

Pokażemy, że $r \equiv p \pmod 2$. Ponieważ

$$\mathrm{id} = \tau_1 \ldots \tau_r \rho_p^{-1} \ldots \rho_1^{-1},$$

więc wystarczy pokazać, że id może być iloczynem tylko parzystej liczby transpozycji. Przypuśćmy, że

$$\mathrm{id} = (a_1 b_1)(a_2 b_2) \ldots (a_k b_k) \tag{9.1}$$

gdzie $k \geqslant 1$ oraz $a_i \neq b_i$ dla $i = 1, \ldots, k$. Pokażemy, że $k$ jest parzyste. Zauważmy, że nie może być $k = 1$, bo $\mathrm{id} \neq (a_1, b_1)$. Dla $k = 2$ teza zachodzi. Załóżmy, że $k \geqslant 3$ i teza zachodzi dla wszystkich iloczynów mniejszej liczby transpozycji. Jedna z transpozycji $(a_i b_i)$ dla $i \in \{2, \ldots, k\}$ musi ruszać $a_1$, bo inaczej (9.1) nie może zachodzić. Stąd $a_1$ musi być jednym z $a_i$ dla pewnego $i \geqslant 2$ (zmieniając ewentualnie rolami $a_i$ z $b_i$). Zauważmy, że jeśli różne litery oznaczają różne liczby to zachodzą równości

$$(cd)(ab) = (ab)(cd), \quad (bc)(ab) = (ab)(bc),$$

więc możemy założyć, że $a_2 = a_1$.

Jeśli $b_2 = b_1$, to iloczyn $(a_1 b_1)(a_1 b_2)$ jest identycznością i możemy go pominąć otrzymując id jako iloczyn $k - 2$ transpozycji. Z założenia indukcyjnego $k - 2$ jest parzyste, więc $k$ jest parzyste.

Jeśli $b_2 \neq b_1$, to iloczyn $(a_1 b_1)(a_1 b_2)$ jest równy $(a_1 b_2)(b_1 b_2)$, czyli (9.1) przyjmuje postać

$$\mathrm{id} = (a_1 b_2)(b_1 b_2) \ldots (a_k b_k). \tag{9.2}$$

Zauważmy, że w formule (9.2) jest mniej transpozycji, które ruszają $a_1$ niż w (9.1). Rozumujemy teraz analogicznie. Pewna transpozycja w iloczynie (9.2) inna niż $(a_1 b_1)$ musi ruszać $a_1$. Mamy znowu dwie możliwości albo możemy zredukować liczbę transpozycji o 2 i zastosować indukcję albo możemy zmniejszyć o 1 liczbę transpozycji ruszających $a_1$ w formule (9.2). Ponieważ id nie może być iloczynem transpozycji, w którym tylko jedna rusza $a_1$, więc w skończonej liczbie kroków musi zajść sytuacja, że iloczyn pierwszych dwóch transpozycji będzie identycznością, więc teza wynika poprzez indukcję. □

Lemat 9.3 uzasadnia poprawność następującej definicji.

**Definicja 9.2** (Permutacja parzysta i nieparzysta)**.**

Permutacje, które są złożeniem parzystej liczby transpozycji nazywamy *parzystymi.* W przeciwnym razie permutacja jest *nieparzysta.*

**Przykład 9.4.** Rozważmy permutację

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 5 & 6 & 4 & 3 & 2 \end{pmatrix} \in S_6.$$

Powyższy napis oznacza, że

$$\sigma(1) = 1, \quad \sigma(2) = 5, \quad \sigma(3) = 6, \quad \sigma(4) = 4, \quad \sigma(5) = 3, \quad \sigma(6) = 2.$$

Sprawdzamy łatwo, że

$$\begin{aligned}\sigma &= (2,6)\circ(2,3)\circ(2,5)\\ &= (5,2)\circ(3,5)\circ(6,3)\\ &= (1,3)\circ(3,1)\circ(2,6)\circ(2,3)\circ(2,5).\end{aligned}$$

Widzimy, że przedstawienie permutacji $\sigma$ jako złożenie transpozycji nie jest jednoznaczne, ale parzystość liczby transpozycji w przedstawieniu nie zmienia się. Permutacja $\sigma$ jest nieparzysta.

**Definicja 9.3** (Znak permutacji)**.**
Znak permutacji $\sigma \in S_n$ definiujemy jako liczbę:

$$\mathrm{sgn}(\sigma) = \begin{cases} 1, & \text{gdy } \sigma \text{ jest parzysta,}\\ -1, & \text{gdy } \sigma \text{ jest nieparzysta.}\end{cases}$$

Sprawdzamy łatwo, że

$$\mathrm{sgn}(\sigma\tau) = \mathrm{sgn}(\sigma)\,\mathrm{sgn}(\tau), \quad \mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma).$$

## 9.2 Definicja wyznacznika i jego własności

Pojęcie wyznacznika pojawiło się pod koniec XVII wieku w pracach Gottfrieda Wilhelma Leibniza (1646–1716) oraz japońskiego matematyka Seki Kōwa znanego również jako Takakazu (1642–1708). Systematyczna teoria wyznacznika pochodzi od Cauchy'ego (1789–1857) oraz Jacobiego (1804–1851).

Dla macierzy kwadratowej $A \in M_{2\times 2}(\mathbb{F})$ zdefiniowaliśmy wyznacznik wzorem

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}.$$

**Wniosek 9.1.** *Dla dowolnych wektorów $u, v, w \in \mathbb{F}^2$ i liczb rzeczywistych $a, b \in \mathbb{F}$ zachodzą warunki:*

*(1) dwuliniowość wyznacznika jako funkcji kolumn macierzy*

$$(i) \quad \det\left[a\cdot u + b\cdot v \mid w\right] = a\det\left[u \mid w\right] + b\det\left[v \mid w\right],$$

$$(ii) \quad \det\left[u \mid a\cdot v + b\cdot w\right] = a\det\left[u \mid v\right] + b\det\left[u \mid w\right];$$

*(2)*

$$\det\left[u \mid u\right] = 0,$$

*(3)* $\det I_2 = 1$.

*Dowód.* Wystarczy przeprowadzić bezpośredni rachunek. Uzasadnimy dla przykładu warunek (1). Niech $u = [u_1, u_2]^T$, $v = [v_1, v_2]^t$ oraz $w = [w_1, w_2]^T$. Wtedy

$$\left[a\cdot u + b\cdot v \mid w\right] = \begin{bmatrix} au_1 + bv_1 & w_1\\ au_2 + bv_2 & w_2\end{bmatrix},$$

więc

$$\begin{aligned}\det\left[a\cdot u + b\cdot v \mid w\right] &= (au_1 + bv_1)w_2 - (au_2 + bv_2)w_1\\ &= a(u_1w_2 - u_2w_1) + b(v_1w_2 - v_2w_1)\\ &= a\det\left[u \mid w\right] + b\det\left[v \mid w\right].\end{aligned}$$

□

U Pokażemy, że każde odwzorowanie $F : M_{2\times2}(\mathbb{F}) \longrightarrow \mathbb{F}$ spełniające warunki (1)-(3) musi być dane wzorem:

$$F(A) = a_{11}a_{22} - a_{12}a_{21}.$$

Oznacza to, że wyznacznik jest jednoznacznie wyznaczony przez warunki (1)-(3) i musi być zadany powyższym wzorem.

**Lemat 9.4.** *Jeśli $F : M_{2\times2}(\mathbb{F}) \longrightarrow \mathbb{F}$ spełnia warunki (1)-(2), to dla dowolnych macierzy $A, B \in M_{2\times2}(\mathbb{F})$ zachodzi wzór*

$$F(AB) = F(A)(b_{11}b_{22} - b_{12}b_{21}). \tag{9.3}$$

*Dowód.* Niech

$$A = \begin{bmatrix} a_1 \mid a_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}.$$

Wtedy

$$\begin{aligned} AB &= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} \\ &= \begin{bmatrix} b_{11} \cdot a_1 + b_{21} \cdot a_2 \mid b_{12} \cdot a_1 + b_{22} \cdot a_2 \end{bmatrix}. \end{aligned}$$

Z warunku (1) zastosowanego dwukrotnie mamy

$$\begin{aligned} F(AB) &= F(b_{11} \cdot a_1 + b_{21} \cdot a_2, b_{12} \cdot a_1 + b_{22} \cdot a_2) \\ &= b_{11}F(a_1, b_{12} \cdot a_1 + b_{22} \cdot a_2) + b_{21}F(a_2, b_{12} \cdot a_1 + b_{22} \cdot a_2) \\ &= b_{11}b_{12}F(a_1, a_1) + b_{11}b_{22}F(a_1, a_2) \\ &+ b_{21}b_{12}F(a_2, a_1) + b_{21}b_{22}F(a_2, a_2). \end{aligned}$$

Z warunku (2) wynika, że $F(v, v) = 0$ dla dowolnego wektora $v \in \mathbb{F}^2$ oraz

$$F(a_2, a_1) = -F(a_1, a_2).$$

Stąd

$$\begin{aligned} F(AB) &= F(a_1, a_2)(b_{11}b_{22} - b_{12}b_{21}) \\ &= F(A)(b_{11}b_{22} - b_{12}b_{21}). \end{aligned}$$

□

**Twierdzenie 9.1.** *Istnieje dokładnie jedno odwzorowanie $F : M_{2\times2}(\mathbb{F}) \longrightarrow \mathbb{F}$ spełniające warunki (1)-(3). Ponadto,*

$$F(B) = b_{11}b_{22} - b_{12}b_{21}, \quad B \in M_{2\times2}(\mathbb{F}).$$

*Dowód.* Przyjmując $A = I$ w Lemacie 9.4 i korzystając z warunku (3) otrzymujemy, że dla dowolnej macierzy $B \in M_{2\times2}(\mathbb{F})$ mamy

$$\begin{aligned} F(B) &= F(IB) \\ &= F(I)(b_{11}b_{22} - b_{12}b_{21}) \\ &= b_{11}b_{22} - b_{12}b_{21}, \end{aligned}$$

co kończy dowód. □

Uogólnimy ten rezultat na dowolny wymiar i zdefiniujemy wyznacznik dowolnej macierzy kwadratowej $A \in M_{n\times n}(\mathbb{F})$. Będziemy w tym celu potrzebowali pewnych faktów dotyczących permutacji zbiorów skończonych.

Będziemy potrzebowali jeszcze jednego pojęcia. Odwzorowanie

$$F : M_{n\times n}(\mathbb{F}) \longrightarrow \mathbb{F}$$

będziemy traktowali jako funkcję zależną od $n$ kolumn macierzy $A = \begin{bmatrix} a_1 \mid \dots \mid a_n \end{bmatrix} \in M_{n\times n}(\mathbb{F})$, czyli $F = F(a_1, \dots, a_n)$ jest funkcją zależną od $n$ wektorów.

**Definicja 9.4** (Odwzorowanie $n$-liniowe)**.**

Odwzorowanie $F : M_{n\times n}(\mathbb{F}) \longrightarrow \mathbb{F}$ nazywamy $n$-liniowym, jeśli jest ono liniowe ze względu na każdą zmienną. Oznacza to, że dla dowolnie ustalonego $i \in \{1, \dots, n\}$ oraz wektorów $a_i, b_i \in \mathbb{F}^n$ i skalarów $\alpha, \beta \in \mathbb{F}$ zachodzi równość

$$F(a_1, \dots, \underbrace{\alpha \cdot a_i + \beta \cdot b_i}_{i}, \dots a_n) = \alpha \cdot F(a_1, \dots, \underbrace{a_i}_{i}, \dots a_n) + \beta \cdot F(a_1, \dots, \underbrace{b_i}_{i}, \dots a_n).$$

**Przykład 9.5.** Iloczyn skalarny

$$\mathbb{R}^n \times \mathbb{R}^n \ni (x,y) \mapsto (x|y) \in \mathbb{F}$$

jest odwzorowaniem dwuliniowym.

**Twierdzenie 9.2** (Definicja wyznacznika)**.** *Istnieje dokładnie jedno odwzorowanie*

$$F : M_{n\times n}(\mathbb{F}) \longrightarrow \mathbb{F}$$

*spełniające warunki*

*(1) $F$ jest $n$-liniowe (jako funkcja kolumn macierzy $A \in M_{n\times n}(\mathbb{F})$),*

*(2) jeśli $A$ ma dwie identyczne kolumny tzn., $a_i = a_j$ dla pewnych $i \neq j$, to $F(A) = 0$,*

*(3) $F(I) = 1$.*

*Ponadto, dla dowolnej macierzy $B = [b_{ij}] \in M_{n\times n}(\mathbb{F})$ zachodzi*

$$F(B) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\, b_{\sigma(1)1} \dots b_{\sigma(n)n}. \tag{9.4}$$

*Nazywamy ją* wyznacznikiem *i oznaczamy przez* det. *Czasami będziemy również stosować oznaczenie $|B| := \det B$.*

U Z warunków (1) i (2) wynika, że jeśli $c \in \mathbb{F}$ oraz $i \neq j$, to

$$F(a_1,\dots,a_i + c\cdot a_j,\dots,a_n) = F(a_1,\dots,a_i,\dots,a_n).$$

**Wniosek 9.2.** *Jeśli $F$ spełnia warunki (1)-(2), to*

*(2') $F$ jest antysymetryczne tzn. jeśli $i \neq j$, to*

$$F(a_1,\dots,\underbrace{a_i}_{i},\dots,\underbrace{a_j}_{j},\dots a_n) = -F(a_1,\dots,\underbrace{a_j}_{i},\dots,\underbrace{a_i}_{j},\dots a_n).$$

*Dowód.* Niech $i \neq j$. Wtedy z warunków (1)-(2) mamy

$$\begin{aligned}
0 &= F(a_1,\dots,\underbrace{a_i + a_j}_{i},\dots,\underbrace{a_i + a_j}_{j},\dots a_n)\\
&= F(a_1,\dots,\underbrace{a_i}_{i},\dots,\underbrace{a_i}_{j},\dots a_n) + F(a_1,\dots,\underbrace{a_i}_{i},\dots,\underbrace{a_j}_{j},\dots a_n)\\
&+ F(a_1,\dots,\underbrace{a_j}_{i},\dots,\underbrace{a_i}_{j},\dots a_n) + F(a_1,\dots,\underbrace{a_j}_{i},\dots,\underbrace{a_j}_{j},\dots a_n)\\
&= F(a_1,\dots,\underbrace{a_i}_{i},\dots,\underbrace{a_j}_{j},\dots a_n) + F(a_1,\dots,\underbrace{a_j}_{i},\dots,\underbrace{a_i}_{j},\dots a_n).
\end{aligned}$$

□

U Jeśli $F$ spełnia (2'), to

$$F(a_1,\dots,\underbrace{a_i}_{i},\dots,\underbrace{a_i}_{j},a_n) = -F(a_1,\dots,\underbrace{a_i}_{i},\dots,\underbrace{a_i}_{j},a_n),$$

więc jeśli $\mathbb{F}$ jest takim ciałem, że $1 + 1 \neq 0$, to warunek (2') implikuje (2).

U Jeśli $F$ spełnia (1) oraz $F(A) = 0$ dla macierzy które mają takie same dwie sąsiadujące kolumny, to $F$ spełnia (2). Rzeczywiście, z dowodu wniosku 9.2 wynika wtedy, że $F$ zmienia znak, gdy zmienimy miejscami dwie sąsiadujące kolumny.

**Przykład 9.6.** Załóżmy, że $n = 2$. Wtedy $S_2$ składa się z dwóch permutacji: id (parzysta) oraz transpozycji (12) (nieparzysta). Wynika stąd, że dla $B \in M_{2\times 2}(\mathbb{F})$ mamy

$$\det(B) = F(B) = b_{11}b_{22} - b_{12}b_{21}.$$

**Przykład 9.7.** Zobaczmy jakie są konsekwencje warunków (1)-(2) dla $n = 3$. Rozważmy macierz $A = \begin{bmatrix} a_1 \mid a_2 \mid a_3 \end{bmatrix} \in M_{3\times 3}(\mathbb{F})$ oraz

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}.$$

Wtedy

$$AB = \begin{bmatrix} b_{11}a_1 + b_{21}a_2 + b_{31}a_3 \mid b_{12}a_1 + b_{22}a_2 + b_{32}a_3 \mid b_{13}a_1 + b_{23}a_2 + b_{33}a_3 \end{bmatrix}.$$

Z warunku (1) tzn. 3-liniowości $F$ i warunku (2) otrzymujemy, że

$$\begin{aligned} F(AB) &= b_{11}b_{12}b_{13}\underbrace{F(a_1,a_1,a_1)}_{=0} + b_{11}b_{12}b_{23}\underbrace{F(a_1,a_1,a_2)}_{=0} + b_{11}b_{12}b_{33}\underbrace{F(a_1,a_1,a_3)}_{=0} \\ &+ b_{11}b_{22}b_{13}\underbrace{F(a_1,a_2,a_1)}_{=0} + b_{11}b_{22}b_{23}\underbrace{F(a_1,a_2,a_2)}_{=0} + b_{11}b_{22}b_{33}F(a_1,a_2,a_3) \\ &+ b_{11}b_{32}b_{13}\underbrace{F(a_1,a_3,a_1)}_{=0} + b_{11}b_{32}b_{23}F(a_1,a_3,a_2) + b_{11}b_{32}b_{33}\underbrace{F(a_1,a_3,a_3)}_{=0} \\ &+ b_{21}b_{12}b_{13}\underbrace{F(a_2,a_1,a_1)}_{=0} + b_{21}b_{12}b_{23}\underbrace{F(a_2,a_1,a_2)}_{=0} + b_{21}b_{12}b_{33}F(a_2,a_1,a_3) \\ &+ b_{21}b_{22}b_{13}\underbrace{F(a_2,a_2,a_1)}_{=0} + b_{21}b_{22}b_{23}\underbrace{F(a_2,a_2,a_2)}_{=0} + b_{21}b_{22}b_{33}\underbrace{F(a_2,a_2,a_3)}_{=0} \\ &+ b_{21}b_{32}b_{13}F(a_2,a_3,a_1) + b_{21}b_{32}b_{23}\underbrace{F(a_2,a_3,a_2)}_{=0} + b_{21}b_{32}b_{33}\underbrace{F(a_2,a_3,a_3)}_{=0} \\ &+ b_{31}b_{12}b_{13}\underbrace{F(a_3,a_1,a_1)}_{=0} + b_{31}b_{12}b_{23}F(a_3,a_1,a_2) + b_{31}b_{12}b_{33}\underbrace{F(a_3,a_1,a_3)}_{=0} \\ &+ b_{31}b_{22}b_{13}F(a_3,a_2,a_1) + b_{31}b_{22}b_{23}\underbrace{F(a_3,a_2,a_2)}_{=0} + b_{31}b_{22}b_{33}\underbrace{F(a_3,a_2,a_3)}_{=0} \\ &+ b_{31}b_{32}b_{13}\underbrace{F(a_3,a_3,a_1)}_{=0} + b_{31}b_{32}b_{23}\underbrace{F(a_3,a_3,a_2)}_{=0} + b_{31}b_{32}b_{33}\underbrace{F(a_3,a_3,a_3)}_{=0} \end{aligned}$$

W efekcie

$$\begin{aligned} F(AB) &= b_{11}b_{22}b_{33}F(a_1,a_2,a_3) + b_{11}b_{32}b_{23}F(a_1,a_3,a_2) \\ &+ b_{21}b_{12}b_{33}F(a_2,a_1,a_3) + b_{21}b_{32}b_{13}F(a_2,a_3,a_1) \\ &+ b_{31}b_{12}b_{23}F(a_3,a_1,a_2) + b_{31}b_{22}b_{13}F(a_3,a_2,a_1). \end{aligned}$$

Z warunku (2) tzn. antysymetryczności $F$ otrzymujemy, że

$$\begin{aligned} F(AB) &= b_{11}b_{22}b_{33}F(a_1,a_2,a_3) - b_{11}b_{32}b_{23}F(a_1,a_2,a_3) \\ &- b_{21}b_{12}b_{33}F(a_1,a_2,a_3) + b_{21}b_{32}b_{13}F(a_1,a_2,a_3) \\ &+ b_{31}b_{12}b_{23}F(a_1,a_2,a_3) - b_{31}b_{22}b_{13}F(a_1,a_2,a_3). \end{aligned}$$

Wzór na $F(AB)$ możemy zapisać w bardziej zwartej formie korzystając z grupy permutacji $S_3$. Składa się ona z $3! = 6$ permutacji:

$$S_3 = \{\mathrm{id}, (12), (13), (23), (123), (132)\}.$$

Transpozycje są nieparzyste, a permutacje id $=(12)(21)$, $(123)=(13)(12)$ i $(132)=(12)(13)$ są parzyste. Stąd

$$\begin{aligned} F(AB) &= \operatorname{sgn}\mathrm{id}\, b_{11}b_{22}b_{33}F(a_1,a_2,a_3) + \operatorname{sgn}(23)\, b_{11}b_{32}b_{23}F(a_1,a_2,a_3) \\ &\quad + \operatorname{sgn}(12)\, b_{21}b_{12}b_{33}F(a_1,a_2,a_3) + \operatorname{sgn}(123)\, b_{21}b_{32}b_{13}F(a_1,a_2,a_3) \\ &\quad + \operatorname{sgn}(132)\, b_{31}b_{12}b_{23}F(a_1,a_2,a_3) + \operatorname{sgn}(13)\, b_{31}b_{22}b_{13}F(a_1,a_2,a_3) \\ &= \left(\sum_{\sigma\in S_3} \operatorname{sgn}(\sigma) b_{\sigma(1)1}b_{\sigma(2)2}b_{\sigma(3)3}\right) F(A). \end{aligned}$$

Z warunku (3) dla $A=I$ mamy

$$F(B) = \sum_{\sigma\in S_3} \operatorname{sgn}(\sigma) b_{\sigma(1)1}b_{\sigma(2)2}b_{\sigma(3)3}.$$

**Lemat 9.5.** *Załóżmy, że $F : M_{n\times n}(\mathbb{F}) \longrightarrow \mathbb{F}$ spełnia warunki (1)–(2). Dla dowolnych $A, B \in M_{n\times n}(\mathbb{F})$ mamy*

$$F(AB) = F(A)\left(\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) b_{\sigma(1)1}\dots b_{\sigma(n)n}\right).$$

*Dowód.* Niech $A = \left[a_1 \mid \dots \mid a_n\right]$ i $B=[b_{ij}]$. Dla $AB = \left[c_1 \mid \dots \mid c_n\right]$, z definicji iloczynu macierzy wynika, że

$$c_k = \sum_{i=1}^{n} b_{ik}a_i = b_{1k}a_1 + \ldots + b_{nk}a_n.$$

Z warunku (1) wynika, że

$$\begin{aligned} F(AB) &= \sum_{i_1,\dots,i_n=1}^{n} F(b_{i_1 1}a_{i_1},\dots,b_{i_n n}a_{i_n}) \\ &= \sum_{i_1,\dots,i_n=1}^{n} b_{i_1 1}\dots b_{i_n n}\, F(a_{i_1},\dots,a_{i_n}). \end{aligned}$$

Użyjemy teraz warunku (2). Wynika z niego, że jeśli $i_k = i_j$ dla pewnych $k\neq j$, to $F(a_{i_1},\dots,a_{i_n}) = 0$. Stąd zamiast sumy

$$\sum_{i_1,\dots,i_n=1}^{n} b_{i_1 1}\dots b_{i_n n}\, F(a_{i_1},\dots,a_{i_n})$$

możemy rozważać sumę

$$\sum_{\sigma\in S_n} b_{\sigma(1)1}\dots b_{\sigma(n)n}F(a_{\sigma(1)},\dots,a_{\sigma(n)}).$$

Teraz wystarczy tylko zauważyć, że warunek (2) gwarantuje, że

$$F(a_{\sigma(1)},\dots,a_{\sigma(n)}) = \operatorname{sgn}(\sigma)F(a_1,\dots,a_n). \qquad \square$$

*Dowód twierdzenia 9.2.* Stosujemy lemat 9.5 dla $A=I$ i warunek (3), otrzymując, że dla dowolnej macierzy $B\in M_{n\times n}(\mathbb{F})$ zachodzi

$$F(B) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) b_{\sigma(1)1}\dots b_{\sigma(n)n}$$

co dowodzi jednoznaczności funkcji $F$ spełniającej (1)-(3). Dla dowodu istnienia definiujemy odwzorowanie $F$ wzorem (9.4). Sprawdzamy łatwo, że tak określone odwzorowanie $F$ spełnia warunki (1) i (3). Uzasadnimy, że zachodzi również warunek (2). Załóżmy, że dwie kolumny macierzy $B$ są równe, tzn. $b_j = b_k$ dla pewnych $j\neq k$. Wtedy

$$F(B) = \sum_{\sigma\in A_n} b_{\sigma(1)1}\dots b_{\sigma(n)n} - \sum_{\sigma\in A_n} b_{\sigma\tau(1)1}\dots b_{\sigma\tau(n)n},$$

więc wystarczy pokazać, że dla $\sigma \in A_n$ mamy

$$b_{\sigma(1)1} \dots b_{\sigma(n)n} = b_{\sigma\tau(1)1} \dots b_{\sigma\tau(n)n}.$$

Jeśli $i \notin \{j, k\}$, to $\sigma\tau(i) = \sigma(i)$, więc $b_{\sigma(i)i} = b_{\sigma\tau(i)i}$. Ponadto, ponieważ $b_j = b_k$, więc

$$b_{\sigma\tau(j)j} = b_{\sigma(k)j} = b_{\sigma(k)k},$$

oraz

$$b_{\sigma\tau(k)k} = b_{\sigma(j)k} = b_{\sigma(j)j},$$

co kończy dowód. □

**Wniosek 9.3** (Wyznacznik iloczynu macierzy)**.** *Dla dowolnych macierzy $A, B \in M_{n\times n}(\mathbb{F})$ zachodzi wzór Cauchy'ego*

$$\det(AB) = \det(A) \cdot \det(B).$$

*W szczególności, jeśli $A$ jest nieosobliwa, to*

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

**Lemat 9.6.** *Niech $A \in M_{n\times n}(\mathbb{F})$. Kolumny macierzy $A$ są liniowo zależne wtedy i tylko wtedy, gdy* $\det(A) = 0$.

*Dowód.* Jeśli kolumny $A$ są liniowo zależne, to z warunków (1) i (2) wynika, że $\det(A)$ jest równy wyznacznikowi macierzy z pewną kolumną zerową, więc $\det(A) = 0$. Załóżmy, że $\det(A) = 0$ i przypuśćmy, że kolumny macierzy $A$ są liniowo niezależne. Wtedy $\ker A = \{0\}$, czyli $A$ jest nieosobliwa. Istnieje więc taka macierz $B \in M_{n\times n}(\mathbb{F})$, że $I = AB$. Wtedy

$$1 = \det(I) = \det(A) \cdot \det(B) = 0,$$

co prowadzi do sprzeczności. □

**Przykład 9.8.** Uzasadnimy, że

$$\begin{vmatrix} \sin\alpha & \cos\alpha & \sin(\alpha+\delta) \\ \sin\beta & \cos\beta & \sin(\beta+\delta) \\ \sin\gamma & \cos\gamma & \sin(\gamma+\delta) \end{vmatrix} = 0.$$

Ponieważ

$$\sin(x+\delta) = \sin x \cos\delta + \cos x \sin\delta,$$

więc trzecia kolumna jest kombinacją liniową dwóch pierwszych kolumn.

**Wniosek 9.4.** *Niech $A \in M_{n\times n}(\mathbb{F})$. Kolumny macierzy $A$ są liniowo niezależne wtedy i tylko wtedy, gdy* $\det(A) \neq 0$. *W szczególności, $A$ jest nieosobliwa wtedy i tylko wtedy, gdy* $\det A \neq 0$.

**Twierdzenie 9.3** (Wyznacznik transpozycji macierzy)**.** *Dla dowolnej macierzy $A \in M_{n\times n}(\mathbb{F})$ mamy*

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

*Dowód.* Z definicji wyznacznika i macierzy transponowanej wystarczy zaobserwować, że

$$\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n} = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)}.$$

Niech $\sigma^{-1}$ będzie permutacją odwrotną do $\sigma$. Wtedy $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$ oraz

$$\begin{aligned} a_{\sigma(1)1} \dots a_{\sigma(n)n} &= a_{\sigma(1)\sigma^{-1}(\sigma(1))} \dots a_{\sigma(n)\sigma^{-1}(\sigma(n))} \\ &= a_{1\sigma^{-1}(1)} \dots a_{n\sigma^{-1}(n)}, \end{aligned}$$

co kończy dowód. □

**Wniosek 9.5.** *Kolumny (wiersze) macierzy kwadratowej $A$ są liniowo zależne wtedy i tylko wtedy, gdy* $\det(A) = 0$.

**Definicja 9.5** (Dopełnienie algebraiczne wyrazu macierzy)**.**
Dla macierzy $A \in M_{n\times n}(\mathbb{F})$ oraz $i, j \in \{1, \dots, n\}$ definiujemy macierz

$$A(i,j) \in M_{(n-1)\times(n-1)}(\mathbb{F})$$

otrzymaną z macierzy $A$ przez wykreślenie $i$-tego wiersza oraz $j$-tej kolumny. *Dopełnieniem algebraicznym wyrazu* $a_{ij}$ nazywamy liczbę

$$A_{ij} = (-1)^{i+j} \det A(i,j).$$

**Przykład 9.9.** Rozważmy macierz

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \in M_{3\times 3}(\mathbb{R}).$$

Przykładowo, dla $i = 2$ oraz $j = 3$ w macierzy

$$\begin{bmatrix} a_{11} & a_{12} & \mathbf{a_{13}} \\ \mathbf{a_{21}} & \mathbf{a_{22}} & \mathbf{a_{23}} \\ a_{31} & a_{32} & \mathbf{a_{33}} \end{bmatrix}$$

wykreślamy drugi wiersz i trzecią kolumnę otrzymując macierz

$$A(2,3) = \begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix}.$$

Znaki $(-1)^{i+j}$ w zależności od pozycji wyrazu $a_{ij}$ mają następujący rozkład

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}.$$

**Twierdzenie 9.4** (Rozwinięcie względem pierwszego wiersza)**.** *Załóżmy, że* $A \in M_{3\times 3}(\mathbb{F})$*. Wtedy*

$$\begin{aligned} \det(A) &= \sum_{i=1}^{3} a_{1i}A_{1i} = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} \\ &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \end{aligned}$$

*Dowód.* Wystarczy sprawdzić, że odwzorowanie $F : M_{3\times 3}(\mathbb{F}) \longrightarrow \mathbb{F}$ zdefiniowane dla macierzy $A = \begin{bmatrix} a_1 \mid a_2 \mid a_3 \end{bmatrix}$ wzorem

$$F(A) = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

spełnia warunki (1)–(3) z Twierdzenia 9.2. Są one łatwe do weryfikacji bezpośrednim rachunkiem. W warunku (1) sprawdzimy przykładowo liniowość $F$ względem pierwszej zmiennej. Dla $b_1 = [b_{11}, b_{21}, b_{31}]^T$ mamy

$$\begin{aligned} F(a_1 + b_1, a_2, a_3) &= (a_{11} + b_{11})\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} + b_{21} & a_{23} \\ a_{31} + b_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} + b_{21} & a_{22} \\ a_{31} + b_{31} & a_{32} \end{vmatrix} \\ &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &+ b_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} b_{21} & a_{23} \\ b_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} b_{21} & a_{22} \\ b_{31} & a_{32} \end{vmatrix} \\ &= F(a_1, a_2, a_3) + F(b_1, a_2, a_3). \end{aligned}$$

Dla $c \in \mathbb{F}$ mamy

$$\begin{aligned} F(c \cdot a_1, a_2, a_3) &= ca_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} ca_{21} & a_{23} \\ ca_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} ca_{21} & a_{22} \\ ca_{31} & a_{32} \end{vmatrix} \\ &= ca_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - ca_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + ca_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= cF(a_1, a_2, a_3). \end{aligned}$$

Dla sprawdzenia warunku (2) zobaczmy przykładowo co się stanie, gdy z pierwsza i trzecia kolumna są identyczne. Mamy

$$F(a_1, a_2, a_1) = a_{11} \begin{vmatrix} a_{22} & a_{21} \\ a_{32} & a_{31} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{21} \\ a_{31} & a_{31} \end{vmatrix} + a_{11} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} = 0.$$

Warunek (3) zachodzi, bo

$$F(I) = 1 \cdot \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} - 0 \cdot \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix} + 0 \cdot \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix} = 1.$$

□

U Analogiczny jak w twierdzeniu 9.4 wzór na wyznacznik zachodzi dla rozwinięcia względem dowolnego wiersza i dowolnej kolumny. Dowód jest ideologicznie taki sam jak dowód twierdzenia 9.4. Przykładowo, dla rozwinięcia względem drugiej kolumny przyjmuje on postać

$$\det(A) = -a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{22} \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{32} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}.$$

Analogiczne rozumowanie prowadzi nas do wzoru na wyznacznik dla macierzy $n \times n$ dla dowolnego $n \in \mathbb{N}$. Pozwala on zredukować wyznaczenie wyznacznika macierzy $A \in M_{n\times n}(\mathbb{F})$ do obliczenia $n$ wyznaczników macierzy o rozmiarze $(n-1) \times (n-1)$.

**Wniosek 9.6** (Rozwinięcie Laplace'a)**.** *Niech* $A \in M_{n\times n}(\mathbb{F})$ *z* $n \geqslant 2$*. Dla dowolnych* $i = 1, \ldots, n$ *i* $j = 1, \ldots, n$ *zachodzą równości*

$$\det(A) = a_{i1}\, A_{i1} + \ldots + a_{in}\, A_{in} = a_{1j}\, A_{1j} + \ldots + a_{nj}\, A_{nj}. \tag{9.5}$$

*Dowód.* Ponieważ $\det(A) = \det(A^T)$, więc wystarczy udowodnić, że

$$\det(A) = a_{i1}A_{i1} + \ldots + a_{in}A_{in},$$

czyli zachodzi wzór na rozwinięcie Laplace'a względem $i$-tego wiersza. Wystarczy sprawdzić, że funkcja $F(A) = a_{i1}A_{i1} + \ldots + a_{in}A_{in}$ spełnia warunki (1)-(3). Warunki (1) i (3) są łatwe do sprawdzenia. Dla dowodu (2) wystarczy zauważyć, że $F(A) = 0$, gdy $A$ ma dwie sąsiadujące kolumny identyczne. □

U Stosując wzór Laplace'a, mamy swobodę wyboru wiersza lub kolumny, względem której zastosujemy rozwinięcie. Oczywiście korzystnie jest wybrać kolumnę (wiersz) z największą liczbą zer. Zanim zastosujemy wzór Laplace'a, możemy najpierw użyć operacji na kolumnach (wierszach) nie zmieniających wyznacznika, aby wprowadzić możliwie dużo wyrazów zerowych w danej kolumnie (wierszu). Możemy więc, do ustalonej kolumny (wiersza) dodać kombinację liniową innych kolumn (wierszy).

**Przykład 9.10.** Odejmując pierwszy wiersz od pozostałych wierszy i stosując rozwinięcie względem pierwszej kolumny, mamy:

$$\begin{aligned} \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} &= \begin{vmatrix} 1 & a & a^2 \\ 0 & b-a & b^2-a^2 \\ 0 & c-a & c^2-a^2 \end{vmatrix} \\ &= \begin{vmatrix} b-a & b^2-a^2 \\ c-a & c^2-a^2 \end{vmatrix} \\ &= (b-a)(c^2-a^2) - (c-a)(b^2-a^2) \\ &= (b-a)(c-a)(c-b) \end{aligned}$$

**Wniosek 9.7** (Własności wyznacznika). *(i) Jeśli $A$ ma dwa identyczne wiersze (kolumny), to* $\det(A) = 0$.

*(ii) Jeśli $A$ ma zerowy wiersz (kolumnę), to* $\det(A) = 0$.

*(iii) $A$ jest nieosobliwa wtedy i tylko wtedy, gdy* $\det(A) \neq 0$ *wtedy i tylko wtedy, gdy kolumny (wiersze) $A$ są liniowo niezależne.*

*(iv) Jeśli $A$ jest górnie (dolnie) trójkątna, to*

$$\det(A) = a_{11}a_{22}\dots a_{nn}.$$

*(v) Jeśli $\sigma \in S_n$, to*

$$\det\left[a_{\sigma(1)} \mid \dots \mid a_{\sigma(n)}\right] = \operatorname{sgn}(\sigma)\det\left[a_1 \mid \dots \mid a_n\right].$$

*(vi) Jeśli $S$ jest nieosobliwa, to*

$$\det(A) = \det(S^{-1}AS).$$

**Przykład 9.11.** Rozważmy macierz Vandermonde'a

$$A = \begin{bmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{bmatrix}.$$

Uzasadnimy formułę rekurencyjną

$$\begin{vmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{vmatrix} = (b-a)(c-a)(d-a)\begin{vmatrix} 1 & b & b^2 \\ 1 & c & c^2 \\ 1 & d & d^2 \end{vmatrix}$$

łączącą wyznacznik macierzy Vandermonde'a rozmiaru 4 z wyznacznikiem macierzy Vandermonde'a rozmiaru 3. W tym celu stosujemy najpierw operacje nie zmieniające wyznacznika, aby wprowadzić zerowe wyrazy w pierwszym wierszu macierzy $A$:

$$\begin{aligned}
\begin{bmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{bmatrix} &\rightsquigarrow \begin{bmatrix} 1 & a & a^2 & 0 \\ 1 & b & b^2 & b^2(b-a) \\ 1 & c & c^2 & c^2(c-a) \\ 1 & d & d^2 & d^2(d-a) \end{bmatrix} \quad \text{operacja} \quad k_4 - a\cdot k_3 \\
&\rightsquigarrow \begin{bmatrix} 1 & a & 0 & 0 \\ 1 & b & b(b-a) & b^2(b-a) \\ 1 & c & c(c-a) & c^2(c-a) \\ 1 & d & d(d-a) & d^2(d-a) \end{bmatrix} \quad \text{operacja} \quad k_3 - a\cdot k_2 \\
&\rightsquigarrow \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & b-a & b(b-a) & b^2(b-a) \\ 1 & c-a & c(c-a) & c^2(c-a) \\ 1 & d-a & d(d-a) & d^2(d-a) \end{bmatrix} \quad \text{operacja} \quad k_2 - a\cdot k_1.
\end{aligned}$$

Z rozwinięcia Lapalce'a względem pierwszego wiersza mamy

$$\begin{aligned}
\begin{vmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{vmatrix} &= \begin{vmatrix} b-a & b(b-a) & b^2(b-a) \\ c-a & c(c-a) & c^2(c-a) \\ d-a & d(d-a) & d^2(d-a) \end{vmatrix} \\
&= (b-a)(c-a)(d-a)\begin{vmatrix} 1 & b & b^2 \\ 1 & c & c^2 \\ 1 & d & d^2 \end{vmatrix}.
\end{aligned}$$

**Przykład 9.12.** Niech

$$\Delta_n = \begin{vmatrix} a & 0 & 0 & \dots & 0 & 0 & b \\ 0 & a & 0 & \dots & 0 & b & 0 \\ 0 & 0 & a & \dots & b & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & b & \dots & a & 0 & 0 \\ 0 & b & 0 & \dots & 0 & a & 0 \\ b & 0 & 0 & \dots & 0 & 0 & a \end{vmatrix}$$

będzie wyznacznikiem powyższej macierzy rozmiaru $2n$. Przykładowo,

$$\Delta_1 = \begin{vmatrix} a & b \\ b & a \end{vmatrix} = a^2 - b^2,$$

$$\begin{aligned} \Delta_2 &= \begin{vmatrix} a & 0 & 0 & b \\ 0 & a & b & 0 \\ 0 & b & a & 0 \\ b & 0 & 0 & a \end{vmatrix} \\ &= a \begin{vmatrix} a & b & 0 \\ b & a & 0 \\ 0 & 0 & a \end{vmatrix} - b \begin{vmatrix} 0 & 0 & b \\ a & b & 0 \\ b & a & 0 \end{vmatrix} \\ &= a^2\,\Delta_1 - b^2\,\Delta_1 = (a^2 - b^2)\,\Delta_1 \\ &= (a^2 - b^2)^2. \end{aligned}$$

Stosując rozwinięcie względem pierwszej kolumny i dokonując dalszej redukcji, otrzymujemy, że

$$\begin{aligned} \Delta_n &= a \begin{vmatrix} a & 0 & \dots & 0 & b & 0 \\ 0 & a & \dots & b & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & b & \dots & a & 0 & 0 \\ b & 0 & \dots & 0 & a & 0 \\ 0 & 0 & \dots & 0 & 0 & a \end{vmatrix} - b \begin{vmatrix} 0 & 0 & \dots & 0 & 0 & b \\ a & 0 & \dots & 0 & b & 0 \\ 0 & a & \dots & b & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & b & \dots & a & 0 & 0 \\ b & 0 & \dots & 0 & a & 0 \end{vmatrix} \\ &= a^2\,\Delta_{n-1} - b^2\,\Delta_{n-1} = (a^2 - b^2)\,\Delta_{n-1}. \end{aligned}$$

Indukcyjnie dostajemy, że $\Delta_n = (a^2 - b^2)^n$.

**Przykład 9.13.** Odejmując ostatni wiersz od pozostałych wierszy, otrzymujemy, że

$$\begin{vmatrix} 1-n & 1 & 1 & \dots & 1 & 1 \\ 1 & 1-n & 1 & \dots & 1 & 1 \\ 1 & 1 & 1-n & \dots & 1 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & 1 & 1 & \dots & 1-n & 1 \\ 1 & 1 & 1 & \dots & 1 & 1-n \end{vmatrix} = \begin{vmatrix} -n & 0 & 0 & \dots & 0 & n \\ 0 & -n & 0 & \dots & 0 & n \\ 0 & 0 & -n & \dots & 0 & n \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -n & n \\ 1 & 1 & 1 & \dots & 1 & 1-n \end{vmatrix}$$

$$= n^{n-1} \begin{vmatrix} -1 & 0 & 0 & \dots & 0 & 1 \\ 0 & -1 & 0 & \dots & 0 & 1 \\ 0 & 0 & -1 & \dots & 0 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 \\ 1 & 1 & 1 & \dots & 1 & 1-n \end{vmatrix}$$

$$= n^{n-1} \begin{vmatrix} -1 & 0 & 0 & \dots & 0 & 1 \\ 0 & -1 & 0 & \dots & 0 & 1 \\ 0 & 0 & -1 & \dots & 0 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{vmatrix} = 0.$$

**Przykład 9.14.** Pokażemy, że

$$\begin{vmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 3 & 3 & 4 & 5 & 6 \\ 1 & 2 & 5 & 4 & 5 & 6 \\ 1 & 2 & 3 & 7 & 5 & 6 \\ 1 & 2 & 3 & 4 & 9 & 6 \\ 1 & 2 & 3 & 4 & 5 & 11 \end{vmatrix} = 5! = 120.$$

Dla $j = 2, 3, 4, 5, 6$ od $j$-tej kolumny odejmujemy pierwszą kolumnę pomnożoną przez $j$, otrzymując macierz dolnie trójkątną:

$$\begin{vmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 3 & 3 & 4 & 5 & 6 \\ 1 & 2 & 5 & 4 & 5 & 6 \\ 1 & 2 & 3 & 7 & 5 & 6 \\ 1 & 2 & 3 & 4 & 9 & 6 \\ 1 & 2 & 3 & 4 & 5 & 11 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 & 4 & 0 \\ 1 & 0 & 0 & 0 & 0 & 5 \end{vmatrix} = 5!.$$

**Wniosek 9.8.** *Dla dowolnej macierzy $A \in M_{n\times n}(\mathbb{F})$ zachodzi równość*

$$a_{i1}\,A_{j1} + \ldots + a_{in}\,A_{jn} = \begin{cases} \det(A), & \text{gdy } i = j, \\ 0, & \text{gdy } i \neq j. \end{cases} \tag{9.6}$$

*Dowód.* Dla $i = j$ jest to równość (9.5). Załóżmy, że $i \neq j$. Niech $A^*$ będzie macierzą powstałą z macierzy $A$ poprzez wstawienie w miejsce $j$-tego wiersza $i$-tego wiersza macierzy $A$. Wtedy $\det(A^*) = 0$, bo $A^*$ ma dwa takie same wiersze. Ze wzoru (9.5) wynika, że

$$0 = \det(A^*) = a_{i1}A_{j1} + \ldots + a_{in}A_{jn}.$$

□

**Definicja 9.6.** Dla macierzy $A \in M_{n\times n}(\mathbb{F})$ definiujemy *macierz dołączoną* $\operatorname{adj} A \in M_{n\times n}(\mathbb{F})$ jako macierz

$$\operatorname{adj} A = \begin{bmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & & & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{bmatrix}.$$

**Wniosek 9.9** (Macierz odwrotna)**.** *Jeśli macierz $A \in M_{n\times n}(\mathbb{F})$ jest nieosobliwa, to*

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj} A. \tag{9.7}$$

*Dowód.* Ze wzoru (9.6) wynika, że

$$\begin{aligned} A(\operatorname{adj} A) &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & & & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{bmatrix} \\ &= \det(A)I. \end{aligned}$$

□

**Przykład 9.15.** Zobaczmy jak wzór na macierz odwrotną wygląda w niskich wymiarach. Dla $n = 2$ i $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ mamy

$$A_{11} = a_{22}, \quad A_{21} = -a_{12}, \quad A_{12} = -a_{21}, \quad A_{22} = a_{22},$$

czyli otrzymujemy znany nam wzór

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

Niech teraz $n = 3$ i $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. Wtedy

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix},$$

gdzie

$$A_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \quad A_{21} = -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}, \quad A_{31} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix},$$

$$A_{12} = -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}, \quad A_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}, \quad A_{32} = -\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix},$$

$$A_{13} = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}, \quad A_{23} = -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}, \quad A_{33} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

Przykładowo dla $A = \begin{bmatrix} 2 & 4 & 2 \\ 3 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$ otrzymujemy, że

$$A^{-1} = -\frac{1}{8} \begin{bmatrix} 1 & -4 & 2 \\ -2 & 0 & 4 \\ -1 & 4 & -10 \end{bmatrix}.$$

**Twierdzenie 9.5** (Wzory Cramera)**.** *Załóżmy, że macierz $A \in M_{n\times n}(\mathbb{F})$ jest nieosobliwa. Niech $A_i$ oznacza macierz otrzymaną z $A$ przez zastąpienie $i$-tej kolumny w $A$ przez wektor $b \in \mathbb{R}^n$. Jedyne rozwiązanie $x = [x_1, \dots, x_n]^T \in \mathbb{F}^n$ równania $Ax = b$ jest dane przez*

$$x_i = \frac{\det(A_i)}{\det(A)}, \quad \textit{dla} \quad i = 1, \dots, n.$$

*Dowód.* Ponieważ $x = A^{-1}b$, więc

$$x_i = \frac{b_1 A_{1i} + \ldots + b_n A_{ni}}{\det(A)} = \frac{\det(A_i)}{\det(A)}.$$

$\square$

**Przykład 9.16.** Rozwiążemy układ równań

$$\begin{bmatrix} 5 & 2 & 1 \\ 3 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 8 \\ 5 \\ 3 \end{bmatrix}.$$

Macierz układu $A = \begin{bmatrix} 5 & 2 & 1 \\ 3 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$ jest nieosobliwa ($\det(A) = 6$), więc

$$x = \frac{\begin{vmatrix} 8 & 2 & 1 \\ 5 & 2 & 0 \\ 3 & 0 & 2 \end{vmatrix}}{\det(A)} = \frac{6}{6} = 1, \quad y = \frac{\begin{vmatrix} 5 & 8 & 1 \\ 3 & 5 & 0 \\ 1 & 3 & 2 \end{vmatrix}}{\det(A)} = \frac{6}{6} = 1, \quad z = \frac{\begin{vmatrix} 5 & 2 & 8 \\ 3 & 2 & 5 \\ 1 & 0 & 3 \end{vmatrix}}{\det(A)} = \frac{6}{6} = 1.$$

**Przykład 9.17** (Rozkład Schura)**.** Przy obliczaniu wyznacznika często stosuje się metodę rozkładu pochodzącą od Schura. Rozważmy macierz blokową

$$P = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right],$$

gdzie $A$ i $D$ są macierzami kwadratowymi. Spróbujmy znaleźć takie macierze $X$ i $Y$, że

$$\begin{aligned} P = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] &= \left[\begin{array}{c|c} A & 0 \\ \hline C & I \end{array}\right] \left[\begin{array}{c|c} I & Y \\ \hline 0 & X \end{array}\right] \\ &= \left[\begin{array}{c|c} A & AY \\ \hline C & CY + X \end{array}\right]. \end{aligned}$$

Jeśli nam się to uda, to $\det P = \det A \cdot \det X$. Taki rozkład jest zawsze możliwy, gdy macierz $A$ jest nieosobliwa. Możemy wtedy przyjąć, że

$$Y = A^{-1}B, \quad X = D - CA^{-1}B.$$

Zauważmy, że

$$\left[\begin{array}{c|c} A & AY \\ \hline C & CY + X \end{array}\right] = \left[\begin{array}{c|c} A & 0 \\ \hline C & X \end{array}\right] \left[\begin{array}{c|c} I & Y \\ \hline 0 & I \end{array}\right],$$

więc możemy również użyć rozkładu

$$\begin{aligned} P &= \left[\begin{array}{c|c} A & 0 \\ \hline C & X \end{array}\right] \left[\begin{array}{c|c} I & Y \\ \hline 0 & I \end{array}\right] \\ &= \left[\begin{array}{c|c} I & 0 \\ \hline CA^{-1} & I \end{array}\right] \left[\begin{array}{c|c} A & 0 \\ \hline 0 & X \end{array}\right] \left[\begin{array}{c|c} I & Y \\ \hline 0 & I \end{array}\right]. \end{aligned}$$

Jeśli macierz $D$ jest nieosobliwa, to mamy analogiczny rozkład

$$P = \left[\begin{array}{c|c} I & BD^{-1} \\ \hline 0 & I \end{array}\right] \left[\begin{array}{c|c} A - BD^{-1}C & 0 \\ \hline 0 & D \end{array}\right] \left[\begin{array}{c|c} I & 0 \\ \hline D^{-1}C & I \end{array}\right].$$

Ponieważ

$$\left[\begin{array}{c|c} I & Y \\ \hline 0 & I \end{array}\right]^{-1} = \left[\begin{array}{c|c} I & -Y \\ \hline 0 & I \end{array}\right],$$

więc powyższe rozkłady pozwalają łatwo wyznaczyć $P^{-1}$.

TEST $\longrightarrow$ Oceń prawdziwość zdań

- Dla dowolnych macierzy $A, B \in M_{2\times 2}(\mathbb{R})$ zachodzi równość $\det(A+B) = \det(A) + \det(B)$.
- Dla dowolnych macierzy $A, B \in M_{2\times 2}(\mathbb{R})$ zachodzi równość $\det(AB) = \det(BA)$.
- Dla dowolnych macierzy $A, B \in M_{2\times 2}(\mathbb{R})$ równość $\det(A) = \det(B)$ implikuje, że $A = B$.
- Niech $A \in M_{2\times 2}(\mathbb{R})$. Jeśli $Ax = 0$ ma niezerowe rozwiązanie, to $\det(A) = 0$.
- Jeśli $A$ jest macierzą kwadratową i $A^k = 0$ dla pewnego $k \in \mathbb{N}$, to $\det(A) = 0$.
- Jeśli $A \in M_{n\times n}(\mathbb{R})$ jest skośnie symetryczna (tzn. $A^T = -A$), to $\det(A) = 0$.
- Jeśli $A \in M_{n\times n}(\mathbb{R})$ jest skośnie symetryczna i $n$ jest nieparzyste, to $\det(A) = 0$.
- Dla dowolnej macierzy kwadratowej $\det A^2 = (\det A)^2$.
- Jeśli $A$ jest nieosobliwa i górnie trójkątna, to

$$\det A^{-1} = \frac{1}{a_{11}} \cdots \frac{1}{a_{nn}}.$$

- Jeśli $A$, $B \in M_{n\times n}(\mathbb{F})$ i $\det A = \det B$, to $\det(A + B) = 2\det A$.
- Dla macierzy kwadratowej $A$ mamy $\det A^T A = \det AA^T$.

$\longleftarrow$

**Definicja 9.7** (Wyznacznik odwzorowania liniowego)**.**
Niech $V$ będzie przestrzenią wektorową wymiaru $n$. *Wyznacznik* $\det L$ odwzorowania liniowego $L : V \longrightarrow V$ definiujemy jako $\det(A)$, gdzie $A$ jest macierzową reprezentacją $L$ w pewnej bazie $v_1, \ldots, v_n$ dla $V$.

U Definicja wyznacznika $\det L$ nie zależy od wyboru reprezentacji macierzowej $A$ dla $L$. Rzeczywiście, jeśli $B$ jest inną reprezentacją macierzową $L$, to $B$ jest podobna do $A$, więc $\det(A) = \det(B)$.

**Wniosek 9.10.** *Niech $L : V \longrightarrow V$ będzie odwzorowaniem liniowym i $\dim(V) < \infty$. Wyznacznik $\det L \neq 0$ wtedy i tylko wtedy, gdy $L$ jest izomorfizmem liniowym.*

**Wniosek 9.11.** *Jeśli $L, G : V \longrightarrow V$ są odwzorowaniami liniowymi, to $\det G \circ L = \det G \cdot \det L$.*

## 9.3 Geometryczna interpretacja wyznacznika

**Definicja 9.8** (Iloczyn wektorowy w $\mathbb{R}^3$)**.**
Iloczynem wektorowym wektorów $u = [u_1, u_2, u_3]^T, v = [v_1, v_2, v_3]^T \in \mathbb{R}^3$ nazywamy wektor

$$u \times v := \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} \cdot e_1 - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} \cdot e_2 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \cdot e_3 \in \mathbb{R}^3.$$

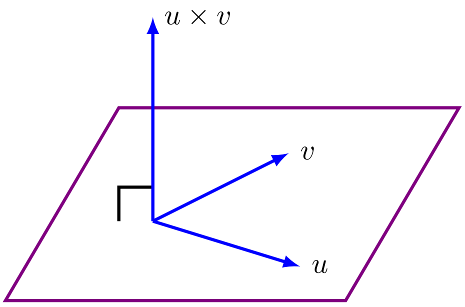
Przez analogię do wzoru Laplace'a będziemy stosować następujące niezbyt formalne oznaczenie:

$$u \times v = \begin{vmatrix} \mathbf{e_1} & \mathbf{e_2} & \mathbf{e_3} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}.$$

**Wniosek 9.12** (Własności iloczynu wektorowego)**.** *Iloczyn wektorowy*

$$\times : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

*ma dla dowolnych $u, v, w \in \mathbb{R}^3$ i $a, b \in \mathbb{R}$ następujące własności:*




Rysunek 9.1: Iloczyn wektorowy.

*(1) (dwuliniowość)*

$$(a \cdot u + b \cdot v) \times w = a \cdot (u \times w) + b \cdot (v \times w)$$
$$u \times (a \cdot v + b \cdot w) = a \cdot (u \times v) + b \cdot (u \times w)$$

*(2) (antysymetryczność)*

$$u \times v = -v \times u$$

*(3) (reguła śruby prawoskrętnej)*

$$e_1 \times e_2 = e_3, \quad e_2 \times e_3 = e_1, \quad e_3 \times e_1 = e_2$$

*W szczególności, $u \times u = 0$ dla dowolnego wektora $u \in \mathbb{R}^3$.*

**Wniosek 9.13.** *Wektor $u \times v$ jest prostopadły do wektorów $u$ i $v$.*

*Dowód.* Sprawdzamy łatwo, że

$$(u \times v | u) = \det \left[ u \mid u \mid v \right] = 0. \qquad \square$$

(U) Iloczyn wektorowy możemy zapisać w konwencji macierzowej. Niech $u = [a, b, c]^T$, $v = [x, y, z]^T \in \mathbb{R}^3$. Bezpośredni rachunek pokazuje, że

$$u \times v = \underbrace{\begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix}}_{:= A_u} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = A_u(v).$$

Stąd

$$u \times (u \times v) = A_u(u \times v) = A_u^2(v).$$

Z drugiej strony, sprawdzamy łatwo, że

$$u \times (u \times v) = (u|v) \cdot u - (u|u) \cdot v.$$

Jeśli $u$ jest wektorem jednostkowym, to otrzymujemy, że

$$v - (v|u) \cdot u = -A_u^2(v) = \begin{bmatrix} -b^2 - c^2 & ab & ac \\ ab & -a^2 - c^2 & bc \\ ac & ba & -a^2 - b^2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Odwzorowanie $v \mapsto v - (v|u) \cdot u$ jest rzutem ortogonalnym na płaszczyznę ortogonalną do wektora jednostkowego $u$.

**Lemat 9.7.** *Dla dowolnych wektorów $0 \neq u, v \in \mathbb{R}^3$ zachodzi równość*

$$\|u \times v\| = \|u\| \|v\| \sin \angle(u, v).$$

*Dowód.* Bezpośredni rachunek pokazuje, że zachodzi *tożsamość Lagrange'a*:

$$\|u \times v\|^2 = \|u\|^2\|v\|^2 - ((u|v))^2.$$

Rzeczywiście, dla $u = [u_1, u_2, u_3]^T$ i $v = [v_1, v_2, v_3]^T$ mamy

$$\begin{aligned}
\|u \times v\|^2 &= \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}^2 + \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix}^2 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}^2 \\
&= (u_2v_3 - v_2u_3)^2 + (u_1v_3 - u_3v_1)^2 + (u_1v_2 - v_1u_2)^2 \\
&= (u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) - (u_1v_1 + u_2v_2 + u_3v_3)^2 \\
&= \|u\|^2\|v\|^2 - ((u|v))^2.
\end{aligned}$$

Stąd

$$\|u \times v\|^2 = \|u\|^2\|v\|^2(1 - \cos^2 \angle(u, v)) = \|u\|^2\|v\|^2 \sin^2 \angle(u, v) \qquad \square$$

**Wniosek 9.14** (Pole równoległoboku)**.** *Niech $u, v \in \mathbb{R}^3$. Wtedy $\|u \times v\|$ jest polem równoległoboku rozpiętego na wektorach $u$ i $v$. W szczególności, jeśli $u$ i $v$ są liniowo niezależne, to $u \times v \neq 0$.*

**Wniosek 9.15.** *Niech $u = [u_1, u_2]^T, v = [v_1, v_2]^T \in \mathbb{R}^2$. Wtedy*

$$\left|\det \left[u \mid v\right]\right|$$

*jest polem równoległoboku rozpiętego na wektorach $u$ i $v$.*

*Dowód.* Rozważmy wektory $u^* = [u_1, u_2, 0]^T, v^* = [v_1, v_2, 0]^T \in \mathbb{R}^3$. Zauważmy, że

$$u^* \times v^* = \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} e_3,$$

więc

$$\|u^* \times v^*\| = \sqrt{\begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}^2} = \left|\det \left[u \mid v\right]\right|. \qquad \square$$

**Definicja 9.9** (Iloczyn mieszany w $\mathbb{R}^3$)**.**
Dla wektorów $u, v, w \in \mathbb{R}^3$ definiujemy ich *iloczyn mieszany* jako liczbę

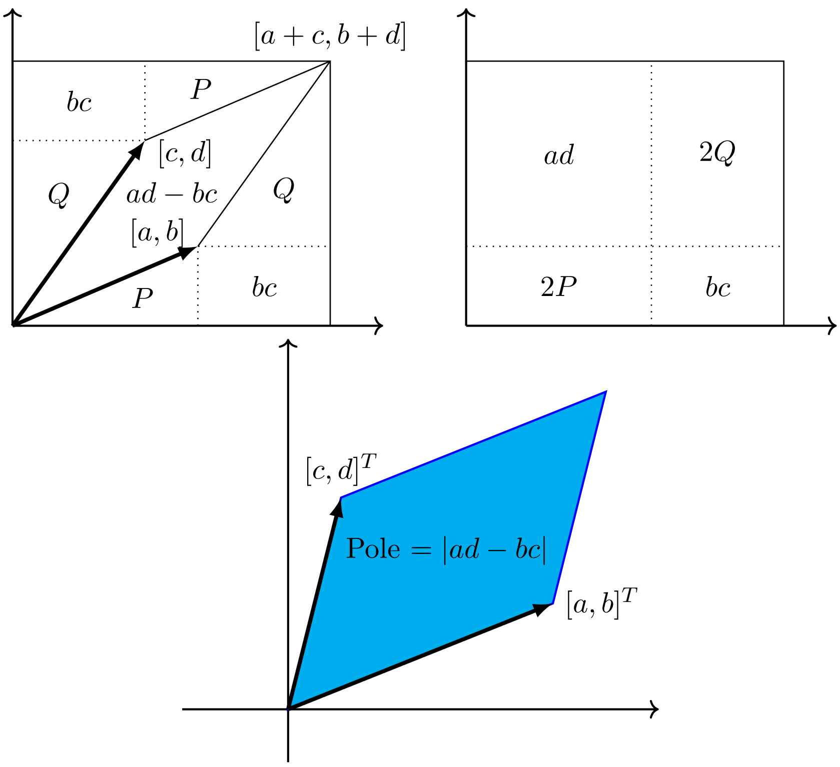
$$(u \times v|w).$$

Zauważmy, że

$$\begin{aligned}
\det \left[w \mid u \mid v\right] &= \det \left[w \mid u \mid v\right]^T \\
&= \begin{vmatrix} w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \\
&= \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} w_1 - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} w_2 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} w_3 \\
&= (u \times v|w).
\end{aligned}$$

**Wniosek 9.16** (Objętość równoległościanu)**.** *Objętość równoległościanu rozpiętego na wektorach $u, v, w$ jest równa*

$$\left|(u \times v|w)\right|.$$




Rysunek 9.2: Wyznacznik i pole

*Dowód.* Objętość równoległościanu rozpiętego na wektorach $u, v, w$ jest równa iloczynowi pola równoległoboku rozpiętego na wektorach $u$, $v$ (jest ono równe $\|u \times v\|$) i odległości $h$ punktu $w$ od płaszczyzny $P$ rozpiętej na wektorach $u$, $v$. Wysokość $h$ jest długościom rzutu prostopadłego wektora $w$ na wektor $u \times v$ prostopadły do $P$. Wtedy $h = \|\frac{(w|u\times v)}{\|u\times v\|^2}(u \times v)\|$, czyli

$$h\|u \times v\| = \left\|\frac{(w|u \times v)}{\|u \times v\|^2}(u \times v)\right\|\|u \times v\| = \big|(u \times v|w)\big|. \qquad \square$$

**Wniosek 9.17.** *Niech $u, v, w \in \mathbb{R}^3$. Objętość równoległościanu rozpiętego na wektorach $u, v, w$ jest równa*

$$\left|\det\left[u \mid v \mid w\right]\right|.$$

**Wniosek 9.18** (Iloczyn wektorowy i mieszany – podsumowanie)**.** *Niech $u, v, w, q \in \mathbb{R}^3$. Zachodzą równości*

*(1)* $u \times v = -v \times u$.

*(3)* $(u|v \times w) = (u \times v|w)$.

*(4) Zachodzi tożsamość Jacobiego*

$$u \times (v \times w) + v \times (w \times u) + w \times (u \times v) = 0.$$

*(5) Zachodzi tośamość Lagrange'a*

$$\|u\|^2\|v\|^2 = \|u \times v\|^2 + ((u|v))^2.$$

*(6)*

$$(u \times v|w \times q) = \begin{vmatrix} (u|w) & (v|w) \\ (u|q) & (v|q) \end{vmatrix}.$$

*(8)* $\det\left[u \mid v \mid w\right] = (u \times v|w)$.

*(9)* $\left|\det\left[u \mid v \mid w\right]\right| = |(u \times v|w)|$ *jest objętością równoległościanu rozpiętego na wektorach* $u, v, w$.

*(10) Iloczyn wektorowy nie jest przemienny ani łączny.*

**Lemat 9.8.** *Jeśli* $A \in M_{3\times 3}(\mathbb{R})$ *i* $u, v \in \mathbb{R}^3$*, to*

$$A^T((Au) \times (Av)) = \det(A) \cdot (u \times v).$$

*W szczególności, jeśli* $A^T A = I$*, to*

$$(Au) \times (Av) = \det(A) A(u \times v).$$

*Jeśli* $A \in SO(3)$ *tzn.* $AA^T = I$ *i* $\det(A) = 1$*, to*

$$(Au) \times (Av) = A(u \times v).$$

*Dowód.* Dla dowolnego $w \in \mathbb{R}^3$ zachodzą równości

$$\begin{aligned}((Au) \times (Av)|Aw) &= \det\left[Au \mid Av \mid Aw\right] \\ &= \det\left(A\left[u \mid v \mid w\right]\right) \\ &= \det(A) \cdot \det\left(\left[u \mid v \mid w\right]\right) \\ &= (\det(A) \cdot (u \times v)|w),\end{aligned}$$

oraz

$$((Au) \times (Av)|Aw) = (A^T((Au) \times (Av))|w).$$

Stąd

$$A^T((Au) \times (Av)) = \det(A) \cdot (u \times v).$$

□

## 9.4 Orientacja $\mathbb{R}^n$

Niech $V$ będzie przestrzenią wektorową nad ciałem $\mathbb{R}$ o wymiarze skończonym $n \geqslant 1$. Dla dwóch uporządkowanych baz $B = b_1, \ldots, b_n$ oraz $B' = b'_1, \ldots, b'_n$ przez $S_{B',B} = [s_{ij}]$ oznaczamy macierz przejścia od bazy $B$ do $B'$, czyli

$$b_i = \sum_{j=1}^{n} s_{ji} b'_j, \quad i = 1, \ldots, n.$$

Niech $\mathcal{B}$ oznacza zbiór wszystkich baz przestrzeni $V$. W zbiorze $\mathcal{B}$ definiujemy relację równoważności $\mathcal{R}$ przez: $B\mathcal{R}B'$ wtedy i tylko wtedy, gdy

$$\det S_{B',B} > 0.$$

**Definicja 9.10** (Orientacja)**.**
Powiemy, że bazy $B$ i $B'$ zadają taką samą *orientację* $V$, jeśli $B\mathcal{R}B'$. Przez wybór orientacji w $V$ rozumiemy wybór dowolnej uporządkowanej bazy (jej klasy abstrakcji w relacji $\mathcal{R}$) $B = b_1, b_2, \ldots, b_n$.

**Przykład 9.18** (Standardowa orientacja $\mathbb{R}^n$)**.** Dla $n \geqslant 1$ standardowa orientacja $\mathbb{R}^n$ jest zadana przez bazę standardową $e_1, \ldots, e_n$.

(1) Przypomnijmy, że macierz przejścia od bazy $B = b_1, \ldots, b_n$ do bazy standardowej jest równa $S = \left[b_1 \mid \ldots \mid b_n\right]$. W szczególności, $B$ zadaje standardową orientację $\mathbb{R}^n$, gdy $\det\left[b_1 \mid \ldots \mid b_n\right] > 0$.

(2) Jeśli $b_1, b_2 \in \mathbb{R}^3$ są liniowo niezależne, to $b_1, b_2, b_1 \times b_2$ zadaje standardową orientację $\mathbb{R}^3$. Rzeczywiście,

$$\det \Big[ b_1 \mid b_2 \mid b_1 \times b_2 \Big] = (b_1 \times b_2 | b_1 \times b_2) > 0.$$

**Wniosek 9.19.** *Niech $L : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ będzie izomorfizmem liniowym. Następujące warunki są równoważne*

*(1) baza $L(e_1), \ldots, L(e_n)$ zadaje standardową orientację $\mathbb{R}^n$,*

*(2)* $\det L > 0$,

*(3) jeśli $b_1, \ldots, b_n$ zadaje standardową orientację $\mathbb{R}^n$, to $L(b_1), \ldots, L(b_n)$ zadaje standardową orientację $\mathbb{R}^n$.*

# Rozdział 10

# Do czego zmierzamy?

Słowa kluczowe rozdziału
wektor własny ♢ wartość własna ♢ diagonalizacja ♢ baza wektorów własnych ♢ diagonalizacja ortogonalna ♢ macierz stochastyczna ♢ dominująca wartość własna ♢ wielomian charakterystyczny ♢ zespolona postać Jordana ♢ rzeczywista postać Jordana ♢ krotność geometryczna i algebraiczna ♢ twierdzenie Cayleya–Hamiltona ♢ diagonalizacja ortogonalna macierzy symetrycznej

Przedstawimy teraz motywację dla naszych dalszych rozważań. Jak już pewnie zauważyliście szczególnie łatwe do analizy są macierze diagonalne

$$D = \operatorname{diag}(\lambda_1, \dots, \lambda_n), \quad \lambda_1, \dots, \lambda_n \in \mathbb{R}.$$

Przykładowo,

$$D^k = \operatorname{diag}(\lambda_1^k, \dots, \lambda_n^k), \quad \det D = \lambda_1 \cdot \ldots \cdot \lambda_n.$$

Jeśli $D$ jest nieosobliwa, to $D^{-1} = \operatorname{diag}(1/\lambda_1, \dots, 1/\lambda_n)$.

Załóżmy, że macierz $A \in M_{n\times n}(\mathbb{R})$ jest podobna do macierzy diagonalnej $D$, czyli istnieje taka macierz nieosobliwa

$$S = \left[v_1 \mid \ldots \mid v_n\right] \in M_{n\times n}(\mathbb{R}),$$

że

$$S^{-1}AS = D.$$

Wtedy $A = SDS^{-1}$, więc

$$A^k = SD^kS^{-1}, \quad \det A = \det D,$$

czyli $A$ jest równie prosta do analizy jak macierz diagonalna $D$.

Problem 1
Scharakteryzować macierze $A \in M_{n\times n}(\mathbb{R})$, które są diagonalizowalne tzn. istnieje taka macierz nieosobliwa $S = \left[v_1 \mid \ldots \mid v_n\right] \in M_{n\times n}(\mathbb{R})$, że

$$S^{-1}AS = D = \operatorname{diag}(\lambda_1, \dots, \lambda_n) = \left[\lambda_1 \cdot e_1 \mid \ldots \mid \lambda_n \cdot e_n\right].$$

Załóżmy, że $S^{-1}AS = D$. Wtedy $AS = SD$, czyli

$$\begin{aligned}
\left[Av_1 \mid \ldots \mid Av_n\right] &= AS = SD \\
&= S\left[\lambda_1 \cdot e_1 \mid \ldots \mid \lambda_n \cdot e_n\right] \\
&= \left[\lambda_1 \cdot Se_1 \mid \ldots \mid \lambda_n \cdot Se_n\right] \\
&= \left[\lambda_1 \cdot v_1 \mid \ldots \mid \lambda_n \cdot v_n\right]
\end{aligned}$$

Stąd

$$Av_1 = \lambda_1 \cdot v_1, \quad \ldots, \quad Av_n = \lambda_n \cdot v_n.$$

Z powyższych rozważań wynika następujące:

**Twierdzenie 10.1** (Diagonalizacja). *Dla macierzy $A \in M_{n\times n}(\mathbb{R})$ poniższe warunki są równoważne*

*(i) istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{R})$, że $S^{-1}AS$ jest diagonalna,*

*(ii) istnieje baza $v_1, \dots, v_n$ dla $\mathbb{R}^n$ oraz takie skalary $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, że*

$$Av_i = \lambda_i \cdot v_i, \quad i = 1, \dots, n.$$

Warto jakoś nazwać własność jaką spełniają wektory $v_i$ w punkcie (ii). Będą one pełniły podstawową rolę w naszych dalszych rozważaniach.

**Definicja 10.1** (Wektor własny i wartość własna).
Powiemy, że **niezerowy** wektor $0 \neq v \in \mathbb{R}^n$ jest *wektorem własnym* macierzy rzeczywistej $A \in M_{n\times n}(\mathbb{R})$, jeśli istnieje taka liczba rzeczywista $\lambda \in \mathbb{R}$, że

$$Av = \lambda \cdot v.$$

Liczbę $\lambda \in \mathbb{R}$ nazywamy *wartością własną* macierzy $A \in M_{n\times n}(\mathbb{R})$ dla wektora własnego $v \in \mathbb{R}^n$.

Twierdzenie 11.1 możemy sformułować następująco:

**Twierdzenie 10.2** (Diagonalizacja = baza wektorów własnych). *Dla macierzy $A \in M_{n\times n}(\mathbb{R})$ poniższe warunki są równoważne*

*(i) istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{R})$, że $S^{-1}AS$ jest diagonalna,*

*(ii) istnieje baza $v_1, \dots, v_n$ dla $\mathbb{R}^n$ złożona z wektorów własnych $A$.*

**Definicja 10.2** (Podprzestrzeń niezmiennicza).
Podprzestrzeń $V \subset \mathbb{R}^n$ nazywamy *niezmienniczą* dla macierzy $A \in M_{n\times n}(\mathbb{R})$, jeśli

$$AV \subset V,$$

albo, równoważnie, $Av \in V$ dla dowolnego wektora $v \in V$.

Podamy geometryczną interpretację wektora własnego.

**Wniosek 10.1.** *Dla niezerowego wektora $v \in \mathbb{R}^n$ i macierzy $A \in M_{n\times n}(\mathbb{R})$ poniższe warunki są równoważne*

*(1) $v$ jest wektorem własnym $A$,*

*(2) prosta generowana przez $v$ jest* niezmiennicza *dla $A$ tzn.*

$$A(\operatorname{span}\{v\}) \subset \operatorname{span}\{v\}.$$

U

Nie każda macierz rzeczywista ma jakiś wektor własny. Wystarczy znaleźć macierz rzeczywistą $A$, która nie ma prostych niezmienniczych. Przykładowo, macierz obrotu o kąt $\frac{\pi}{2}$, czyli

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

nie ma żadnej prostej niezmienniczej.

Możemy na to spojrzeć jeszcze trochę inaczej. Wektor $0 \neq v \in \mathbb{R}^2$ jest wektorem własnym $A$, gdy dla pewnej liczby rzeczywistej $\lambda \in \mathbb{R}$ zachodzi warunek

$$(A - \lambda I)v = 0,$$

czyli $\ker(A - \lambda I) \neq 0$, więc $\det(A - \lambda I) = 0$ dla pewnej liczby rzeczywistej $\lambda \in \mathbb{R}$. Zauważmy, że w naszym przykładzie

$$\det(A - \lambda I) = \lambda^2 + 1 \neq 0,$$

dla $\lambda \in \mathbb{R}$.

Macierz rzeczywistą $A$ możemy traktować w naturalny sposób jako macierz zespoloną. Możemy, więc spytać o istnienie zespolonych wektorów własnych $v \in \mathbb{C}^2$. Zauważmy, że $\det(A - \lambda I) = 0$ dla $\lambda = \pm i$. Rozważmy macierz

$$A - iI = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix}$$

i znajdźmy taki niezerowy wektor $v = [a + ib, c + id] \in \mathbb{C}^2$, że $(A - iI)v = 0$. Mamy

$$(A - iI)v = [-ia + b - c - id, a + ib - ic + d]^T = [0, 0]^T,$$

czyli

$$a + d = 0, \quad b - c = 0.$$

Wektor $v$ ma więc postać

$$v = [a + ib, b - ia]^T = [a + ib, -i(a + ib)]^T = (a + ib)[1, -i]^T.$$

Wynika stąd, że

$$A[1, -i]^T = i[1, -i]^T,$$

czyli $[1, -i]^T$ jest zespolonym wektorem własnym $A$ z wartością własną $i$. Analogicznie sprawdzamy, że $[1, i]^T$ jest zespolonym wektorem własnym $A$ dla wartości własnej $-i$. Macierz $A$ jest diagonalizowalna jako macierz zespolona. Dla

$$S = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$$

mamy

$$S^{-1}AS = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

U

Pojęcie wektora własnego możemy też zdefiniować w bardziej ogólnym kontekście. Niech $L : V \longrightarrow V$ będzie odwzorowaniem liniowym. Powiemy, że niezerowy wektor $v \in V$ jest *wektorem własnym* $L$ jeśli

$$L(v) = \lambda \cdot v.$$

Przestrzeń $V$ nie musi być skończenie wymiarowa. Przykładowo, niech $C^\infty(\mathbb{R}, \mathbb{R})$ będzie przestrzenią wszystkich funkcji $f : \mathbb{R} \longrightarrow \mathbb{R}$, mających ciągłą pochodną dowolnego rzędu. Wtedy odwzorowanie (operator różniczkowania)

$$L : C^\infty(\mathbb{R}, \mathbb{R}) \ni f \longmapsto f' \in C^\infty(\mathbb{R}, \mathbb{R})$$

jest liniowe. Ponadto, dla $f = e^{\lambda x} \in C^\infty(\mathbb{R}, \mathbb{R})$ mamy

$$\begin{aligned} L(f) &= (e^{\lambda x})' \\ &= \lambda e^{\lambda x} \\ &= \lambda \cdot f, \end{aligned}$$

czyli $f = e^{\lambda x}$ jest wektorem własnym $L$ dla wartości własnej $\lambda$.

Uzasadnimy teraz, że jeśli $L(f) = \lambda \cdot f$, to $f = Ce^{\lambda x}$ dla pewnego $C \in \mathbb{R}$. Rzeczywiście, równość $f' = \lambda f$ oznacza, że $f' - \lambda f = 0$, czyli również

$$\underbrace{f' e^{-\lambda x} - \lambda e^{-\lambda x} f}_{=(fe^{-\lambda x})'} = 0,$$

czyli $fe^{-\lambda x}$ jest funkcją stałą, więc $fe^{-\lambda x} = C$.

Wynika stąd, że jądro $\ker(L - \lambda I)$ ma wymiar 1 oraz

$$\ker(L - \lambda I) = \operatorname{span}\{e^{\lambda x}\}.$$

U

Niech $L : V \longrightarrow V$ będzie odwzorowaniem liniowym i niech $v_1, \ldots, v_n \in V$ będzie bazą $V$. Rozważmy macierz $A \in M_{n\times n}(\mathbb{F})$ odwzorowania $L$ w tej bazie. Wektor $v = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$ jest wektorem własnym $L$ dla wartości własnej $\lambda \in \mathbb{F}$ wtedy i tylko wtedy, gdy wektor współrzędnych $[x_1, \ldots, x_n]^T \in \mathbb{F}^n$ jest wektorem własnym macierzy $A$ dla wartości własnej $\lambda$.

Możemy pójść jeszcze krok dalej. Wektory własne macierzy diagonalnej $D$ to wektory bazy standardowej $e_1, \ldots, e_n$. Ma ona własność

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

Takie bazy $\mathbb{R}^n$ będziemy nazywać *ortonormalnymi.* Wektory $e_i$ oraz $e_j$ są ortogonalne i mają normę 1.

> PROBLEM 2
> Scharakteryzować macierze $A \in M_{n\times n}(\mathbb{R})$, które są ortogonalnie diagonalizowalne tzn. istnieje taka macierz nieosobliwa $S = \left[v_1 \mid \ldots \mid v_n\right] \in M_{n\times n}(\mathbb{R})$, że
>
> $$S^{-1}AS = D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) = \left[\lambda_1 \cdot e_1 \mid \ldots \mid \lambda_n \cdot e_n\right]$$
>
> oraz
>
> $$(v_i|v_j) = \delta_{ij}.$$

Zaczniemy od przyglądnięcia się takim macierzom

$$S = \left[v_1 \mid \ldots \mid v_n\right] \in M_{n\times n}(\mathbb{R}),$$

że $(v_i|v_j) = \delta_{ij}$. Bezpośrednio z definicji iloczynu macierzy i definicji macierzy transponowanej wynika, że wtedy

$$S^T S = I, \quad \text{czyli} \quad S^{-1} = S^T.$$

Takie macierze będziemy nazywać *ortogonalnymi.* W problemie 2 pytamy więc o istnienie takiej macierzy ortogonalnej $S \in M_{n\times n}(\mathbb{R})$, że

$$S^{-1}AS = S^T AS = D$$

jest diagonalna. Zauważmy, że wtedy $A = SDS^T$ oraz

$$\begin{aligned} A^T &= (SDS^T)^T \\ &= (S^T)^T D^T S^T \\ &= SDS^T \\ &= A, \end{aligned}$$

czyli $A$ musi być symetryczna. Jest to warunek konieczny, aby zaszła sytuacja opisana w problemie 2. W następnych rozdziałach pokażemy, że macierze symetryczne to jedyne macierze rzeczywiste rozwiązujące problem 2.

## 10.1 Rozkład Jordana w wymiarze 2

Spróbujmy nabrać intuicji w przypadku wymiaru 2. Zakładamy, że ciało $\mathbf{F}$ jest równe $\mathbb{R}$ lub $\mathbb{C}$. Rozważmy macierz $A \in M_{2\times 2}(\mathbb{F})$.

**Definicja 10.3** (Wektor własny i wartość własna)**.**
Liczbę $\lambda \in \mathbb{F}$ nazywamy *wartością własną* macierzy $A \in M_{2\times 2}(\mathbb{F})$, jeśli istnieje taki **niezerowy** wektor $v \in \mathbb{F}^2$, że

$$Av = \lambda \cdot v.$$

Każdy taki wektor $v$ nazywamy *wektorem własnym* $A$ odpowiadającym wartości własnej $\lambda$.

U Niech $\mathbb{F} = \mathbb{R}$. Interpretacja geometryczna jest następująca. Jeśli $v \in \mathbb{R}^2$ jest wektorem własnym dla wartości własnej $\lambda \in \mathbb{R}$, to prosta $l = \operatorname{span}\{v\}$ jest niezmiennicza dla $A$ tzn. jeśli $w \in l$, to $Aw \in l$.

Jeśli $v \in \mathbb{F}^2$ jest wektorem własnym dla wartości własnej $\lambda \in \mathbb{F}$, to $(A - \lambda I)v = 0$. Oznacza to, że $\ker(A - \lambda I) \neq \{0\}$, czyli $\det(A - \lambda I) = 0$. Odwrotnie, jeśli $\det(A - \lambda I) = 0$ dla pewnego skalara $\lambda \in \mathbb{F}$, to macierz $A - \lambda I$ jest osobliwa. W szczególności, $\ker(A - \lambda I) \neq \{0\}$, czyli istnieje wektor własny $v \in \ker(A - \lambda I)$.

**Wniosek 10.2.** *Dla $A \in M_{2\times 2}(\mathbb{F})$ i $\lambda \in \mathbb{F}$ następujące warunki są równoważne*

*(1) $\lambda$ jest wartością własną $A$;*

*(2) $\ker(A - \lambda I) \neq \{0\}$;*

*(3) $A - \lambda I$ jest osobliwa;*

*(4) $\det(A - \lambda I) = 0$.*

Warunek (4) oznacza, że wartość własna $\lambda \in \mathbb{F}$ jest pierwiastkiem wielomianu

$$p_A(x) = \det(A - xI)$$

Wielomian $p_A$ nazywamy *wielomianem charakterystycznym* macierzy $A$. Bezpośredni rachunek pokazuje, że

$$\begin{aligned} p_A(x) &= \det \begin{bmatrix} a_{11} - x & a_{12} \\ a_{21} & a_{22} - x \end{bmatrix} \\ &= x^2 - (a_{11} + a_{22})x + (a_{11}a_{22} - a_{12}a_{22}) \\ &= x^2 - \operatorname{tr}(A)x + \det(A). \end{aligned}$$

Macierze podobne mają takie same wielomiany charakterystyczne. Rzeczywiście, jeśli $B = S^{-1}AS$, to

$$\begin{aligned} \det(B - xI) &= \det(S^{-1}AS - xI) \\ &= \det(S^{-1}(A - xI)S) \\ &= \det(A - xI). \end{aligned}$$

Jeśli macierze $A$ i $B$ mają takie same wielomiany charakterystyczne, to nie muszą być podobne. Przykładowo, $A = 0$ oraz $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ mają taki sam wielomian charakterystyczny $p(x) = x^2$, ale nie są podobne, bo macierz zerowa jest podobna tylko do siebie samej.

Dla $\mathbb{F} = \mathbb{C}$ zasadnicze twierdzenie algebry gwarantuje, że wielomian $p_A$ ma dwa (niekoniecznie różne) pierwiastki zespolone $\lambda_1, \lambda_2 \in \mathbb{C}$. Są one wtedy wartościami własnymi macierzy $A \in M_{2\times 2}(\mathbb{C})$. Macierz zespolona ma więc zawsze dwie wartości własne. Nie oznacza to, że ma ona dwa liniowo niezależne wektory własne. Przykładowo wartościami własnymi macierzy

$$A = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \quad \lambda \in \mathbb{C},$$

są $\lambda_1 = \lambda_2 = \lambda$. Mamy więc dwukrotną wartość własną $\lambda$. Zobaczmy jak wyglądają wektory własne. Niezerowy wektor $v = [v_1, v_2]^T \in \mathbb{C}^2$ jest wektorem własnym dla $\lambda$ jeśli $(A - \lambda I)v = 0$, czyli

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_2 \\ 0 \end{bmatrix}.$$

Stąd $v_2 = 0$, czyli wektor własny $v$ ma postać

$$\begin{bmatrix} v_1 \\ 0 \end{bmatrix} = v_1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = v_1 \cdot e_1.$$

W przypadku $\mathbb{F} = \mathbb{R}$, z algebraicznego punktu widzenia, dla wielomianu charakterystycznego mamy trzy możliwości:

(i) $p_A$ ma dwa różne pierwiastki rzeczywiste;

(ii) $p_A$ ma dwukrotny pierwiastek rzeczywisty;

(iii) $p_A$ ma dwa sprzężone (nierzeczywiste) pierwiastki zespolone.

W przypadku (iii) macierz rzeczywista $A$ nie ma rzeczywistych wartości własnych.

**Wniosek 10.3.** *Niech $A \in M_{2\times 2}(\mathbb{C})$. Jeśli $\lambda_1, \lambda_2 \in \mathbb{C}$ są pierwiastkami równania charakterystycznego*

$$x^2 - \operatorname{tr}(A)x + \det(A) = 0,$$

*to zachodzą wzory Viète'a*

$$\operatorname{tr}(A) = \lambda_1 + \lambda_2, \quad \det(A) = \lambda_1 \cdot \lambda_2.$$

*Dowód.* Z założenia mamy

$$(x - \lambda_1)(x - \lambda_2) = p_A(x) = x^2 - \operatorname{tr}(A)x + \det(A),$$

wystarczy więc porównać współczynniki wielomianów po obu stronach równości. □

U Dla dowolnych macierzy $X, Y \in M_{2\times 2}(\mathbb{C})$ macierze $XY$ i $YX$ mają takie same wielomiany charakterystyczne. Wynika, to z faktu, że

$$\operatorname{tr} XY = \operatorname{tr} YX, \quad \det XY = \det YX.$$

**Lemat 10.1.** *Jeśli $A \in M_{2\times 2}(\mathbb{F})$ ma dwie różne rzeczywiste wartości własne $\lambda_1 \neq \lambda_2$ oraz $v_i \in \mathbb{F}^2$ są wektorami własnymi odpowiadającymi $\lambda_i \in \mathbb{F}$, to $v_1$, $v_2$ są liniowo niezależne. W szczególności, $v_1, v_2$ tworzą bazę $\mathbb{F}^2$.*

*Dowód.* W przeciwnym razie $v_2 = p \cdot v_1$ dla pewnego skalara $p \in \mathbb{F}$. Wtedy

$$\begin{aligned} \lambda_2 \cdot v_2 &= A(v_2) \\ &= A(p \cdot v_1) \\ &= p \cdot A(v_1) \\ &= p\lambda_1 \cdot v_1 \\ &= \lambda_1 \cdot v_2, \end{aligned}$$

czyli $\lambda_1 = \lambda_2$, co prowadzi do sprzeczności. Skorzystaliśmy tu z faktu, że $v_2 \neq 0$. □

**Twierdzenie 10.3.** *Jeśli $A \in M_{2\times 2}(\mathbb{F})$ ma dwa liniowo niezależne wektory własne $v_1, v_2 \in \mathbb{F}^2$ odpowiadające wartościom własnym $\lambda_1$, $\lambda_2 \in \mathbb{F}$ (niekoniecznie różnym), to dla macierzy $B = \begin{bmatrix} v_1 \mid v_2 \end{bmatrix}$ mamy*

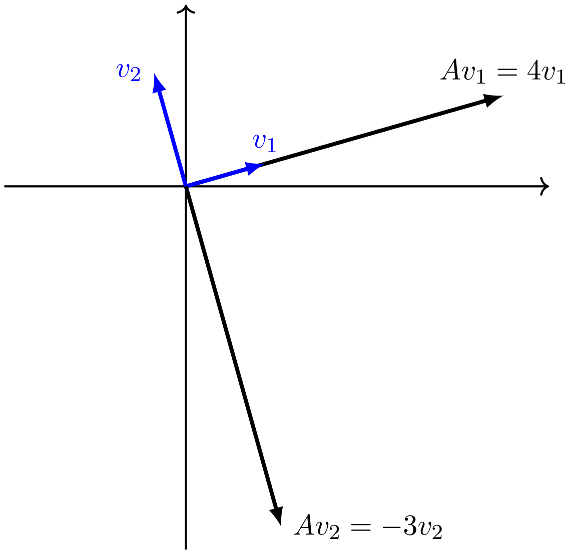
$$\Lambda := B^{-1}AB = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

*Dowód.* Z definicji $B$ mamy, że $B(e_i) = v_i$ dla $i = 1, 2$. Oczywiście $B$ jest odwracalna oraz dla $i = 1, 2$ mamy

$$\begin{aligned} B^{-1}ABe_i &= B^{-1}Av_i \\ &= B^{-1}(\lambda_i \cdot v_i) \\ &= \lambda_i \cdot e_i. \end{aligned}$$

□

U Twierdzenie 11.3 orzeka, że jeśli $\mathbb{F}^2$ ma bazę złożoną z wektorów własnych, to $A \in M_{2\times 2}(\mathbb{F})$ ma w tej bazie postać diagonalną $\Lambda$. Macierz $\Lambda$ nazywamy *postacią Jordana* macierzy $A$.




Rysunek 10.1: Działanie macierzy $A$ na wektorach własnych $v_1$, $v_2$ o wartościach własnych $\lambda_1 = 4$, $\lambda_2 = -3$.

**Przykład 10.1.** Znajdziemy postać Jordana macierzy

$$A = \begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix}.$$

Szukamy najpierw wartości własnych, czyli pierwiastków równania charakterystycznego

$$p_A(x) = x^2 - x - 12 = 0.$$

Są nimi liczby $\lambda_1 = 4$ oraz $\lambda_2 = -3$. Aby znaleźć wektor własny dla $\lambda_1 = 4$, znajdujemy jądro macierzy

$$A - 4I = \begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix}.$$

Rozwiązując równanie $(A - 4I)x = 0$, otrzymujemy, że $x = [2x_2, x_2]^T$. Jednym z wektorów własnych jest $v_1 = [2, 1]^T$. Analogicznie $v_2 = [-1, 3]^T$ jest wektorem własnym dla $\lambda_2$. W bazie $v_1$, $v_2$ macierz $A$ ma postać diagonalną

$$\Lambda = \begin{bmatrix} 4 & 0 \\ 0 & -3 \end{bmatrix}.$$

**!** Macierz zespolona $A \in M_{2\times 2}(\mathbb{C})$ może mieć rzeczywistą wartość własną, ale może nie mieć odpowiadającego jej rzeczywistego wektora własnego w $\mathbb{R}^2$. Dla przykładu rozważmy macierz

$$A = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}.$$

Wielomian charakterystyczny ma postać

$$p_A(x) = x^2 - 1,$$

więc $A$ ma wartości własne $\lambda_1 = 1$ oraz $\lambda_2 = -1$. Wektor własny $[x, y]^T \in \mathbb{C}^2$ dla $\lambda_1 = 1$ jest niezerowym rozwiązaniem równania

$$\begin{bmatrix} -1 & -i \\ i & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

czyli

$$-x - iy = 0, \quad ix - y = 0.$$

Stąd, $y = ix$ oraz każdy wektor własny jest postaci

$$c \cdot [1, i]^T, \quad c \in \mathbb{C} \setminus \{0\}.$$

Nie może on być wektorem z $\mathbb{R}^2$.

Zajmiemy się teraz przypadkiem, gdy $\lambda \in \mathbb{F}$ jest *dwukrotnym pierwiastkiem* wielomianu charakterystycznego $p_A$. Mogą zachodzić dwie możliwości:

- $\dim \ker(A - \lambda I) = 2$. Mówimy wtedy, że *krotność geometryczna* wartości własnej $\lambda$ jest równa jej *krotności algebraicznej*. Wówczas $A$ ma dwa liniowo niezależne wektory własne odpowiadające $\lambda$ i jesteśmy w sytuacji opisanej twierdzeniem 11.3. Wtedy postać Jordana $\Lambda$ macierzy $A$ jest diagonalna:

$$\Lambda := B^{-1}AB = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \lambda I.$$

  Z powyższej równości wynika, że wtedy $A = \lambda I$.

- $\dim \ker(A - \lambda I) = 1$. Istnieje wtedy tylko jeden liniowo niezależny wektor własny $v \in \mathbb{F}^2$ dla $\lambda \in \mathbb{F}$.

**Wniosek 10.4.** *Jeśli $\lambda \in \mathbb{F}$ jest dwukrotnym pierwiastkiem charakterystycznym macierzy $A \in M_{2\times 2}(\mathbb{F})$ i $A \neq \lambda I$, to* $\dim \ker(A - \lambda I) = 1$ *i $A$ ma tylko jeden liniowo niezależny wektor własny.*

**Lemat 10.2** (Twierdzenie Cayleya–Hamiltona)**.** *Dla macierzy $A \in M_{2\times 2}(\mathbb{F})$ zachodzi równość*

$$p_A(A) = A^2 - \operatorname{tr}(A)A + \det(A)I = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

*Dowód.* Zastosujemy brutalną siłę, czyli bezpośredni rachunek. Niech $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Ponieważ $p_A(x) = (x-a)(x-d) - bc$, więc

$$\begin{aligned} p_A(A) &= (A - aI)(A - dI) - bcI \\ &= \begin{bmatrix} 0 & b \\ c & d-a \end{bmatrix} \begin{bmatrix} a-d & b \\ c & 0 \end{bmatrix} - \begin{bmatrix} bc & 0 \\ 0 & bc \end{bmatrix} \\ &= \begin{bmatrix} bc & 0 \\ c(a-d)+(d-a)c & bc \end{bmatrix} - \begin{bmatrix} bc & 0 \\ 0 & bc \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

□

**Przykład 10.2.** Niech $A \in M_{2\times 2}(\mathbb{C})$. Twierdzenie Cayleya–Hamiltona pozwala łatwo wyznaczyć $A^n$ dla $n \in \mathbb{N}$. Niech $\lambda_1, \lambda_2 \in \mathbb{C}$ będą wartościami własnymi $A$. Wtedy

$$A^2 - (\lambda_1 + \lambda_2)A + \lambda_1\lambda_2 I = 0,$$

czyli

$$(A - \lambda_1 I)(A - \lambda_2 I) = 0.$$

Jeśli $\lambda_1 \neq \lambda_2$, to rozważamy macierze

$$X = \frac{A - \lambda_2 I}{\lambda_1 - \lambda_2}, \quad Y = \frac{A - \lambda_1 I}{\lambda_2 - \lambda_1}.$$

Wtedy

$$X^2 = X, \quad XY = YX = 0, \quad Y^2 = Y$$

Środkowa równość wynika bezpośrednio z twierdzenia Cayleya–Hamiltona. Sprawdzimy równość $X^2 = X$. Mamy

$$\begin{aligned}X^2 &= \frac{(A-\lambda_2 I)^2}{(\lambda_1-\lambda_2)^2}\\ &= \frac{(A-\lambda_2 I)(A-\lambda_1 I + (\lambda_1-\lambda_2)I)}{(\lambda_1-\lambda_2)^2}\\ &= \frac{(A-\lambda_2 I)(A-\lambda_1 I)}{(\lambda_1-\lambda_2)^2} + \frac{(A-\lambda_2)(\lambda_1-\lambda_2)I}{(\lambda_1-\lambda_2)^2}\\ &= \frac{A-\lambda_2 I}{(\lambda_1-\lambda_2)}\\ &= X.\end{aligned}$$

Stąd dla $k \geqslant 2$ mamy

$$X^k = X, \quad Y^k = Y, \quad \lambda_1 \neq \lambda_2.$$

Dla $\lambda_1 = \lambda_2$ rozważamy

$$Z = A - \lambda_1 I.$$

Wtedy $Z^k = 0$ dla $k \geqslant 2$. Mamy więc

$$A^n = \begin{cases}(\lambda_1 X + \lambda_2 Y)^n = \lambda_1^n X + \lambda_2^n Y, & \text{gdy } \lambda_1 \neq \lambda_2\\ (\lambda_1 I + Z)^n = \lambda_1^n I + n\lambda_1^{n-1} Z, & \text{gdy } \lambda_1 = \lambda_2\end{cases}$$

**Wniosek 10.5.** *Jeśli $A \in M_{2\times 2}(\mathbb{C})$ jest nieosobliwa, to*

$$A^{-1} = \frac{1}{\det A}(\operatorname{tr} A \cdot I - A), \quad \operatorname{tr} A^{-1} = \frac{\operatorname{tr} A}{\det A}.$$

*Ponadto,*

$$\det A = \frac{1}{2}((\operatorname{tr} A)^2 - \operatorname{tr} A^2).$$

*Dowód.* Z twierdzenia Cayleya–Hamiltona mamy

$$A^2 - \operatorname{tr}(A)A + \det(A)I = \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}.$$

Mnożąc powyższą równość przez $A^{-1}$ otrzymujemy, że

$$A^{-1} = \frac{1}{\det A}(\operatorname{tr} A \cdot I - A).$$

Biorąc ślady obu stron w tej równości, dostajemy, że $\operatorname{tr} A^{-1} = \frac{\operatorname{tr} A}{\det A}$. Wzór na $\det A$ otrzymujemy, obliczając ślady macierzy po obu stronach w równości Cayleya–Hamiltona. □

**Wniosek 10.6.** *Niech $A \in M_{2\times 2}(\mathbb{C})$. Jeśli $A \notin \{\alpha I : \alpha \in \mathbb{C}\}$ oraz $A^2 - aA + bI =$ dla pewnych $a, b \in \mathbb{C}$, to*

$$a = \operatorname{tr} A, \quad b = \det A.$$

*Dowód.* Z założenia i twierdzenia Cayleya–Hamiltona wynika, że

$$(a - \operatorname{tr} A)A = (b - \det A)I.$$

Zauważmy, że $a - \operatorname{tr} A = 0$, bo inaczej $A = \frac{b-\det A}{a-\operatorname{tr} A} I$, co jest sprzeczne z założeniem. Stąd również $b - \det A = 0$. □

U Macierz $A \in M_{2\times 2}(\mathbb{C})$ jest nilpotentna, jeśli $A^n = 0$ dla pewnego $n \in \mathbb{N}$. Pokażemy, że wtedy 0 jest jedyną wartością własną $A$ oraz $A^2 = 0$. Zauważmy, że jeśli $\lambda \in \mathbb{C}$ jest wartością własną dla wektora własnego $v \neq 0$, to $0 = A^n v = \lambda^n \cdot v$, czyli $\lambda = 0$. W szczególności, $\operatorname{tr} A = 0$ i $\det A = 0$, czyli $p_A(x) = x^2$. Z twierdzenia Cayleya–Hamiltona wynika, że $A^2 = 0$.

**Lemat 10.3.** *Jeśli $\lambda \in \mathbb{F}$ jest dwukrotnym pierwiastkiem wielomianu charakterystycznego $p_A$, to*

$$(A - \lambda I)^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

*Dowód.* Ze wzorów Viète'a mamy $\operatorname{tr}(A) = 2\lambda$ i $\det(A) = \lambda^2$. Z twierdzenia Cayleya–Hamiltona wynika, że

$$A^2 - 2\lambda A + \lambda^2 I = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Ponieważ macierze $A$ i $\lambda I$ komutują ze sobą (tzn. $A(\lambda I) = (\lambda I)A$), więc

$$(A - \lambda I)^2 = A^2 - 2\lambda A + \lambda^2 I. \qquad \square$$

**Twierdzenie 10.4.** *Jeśli $\lambda \in \mathbb{F}$ jest dwukrotną wartością własną macierzy $A \in M_{2\times 2}(\mathbb{F})$ i $A \neq \lambda I$, to istnieje taka baza $v, w$ dla $\mathbb{F}^2$, że*

$$Av = \lambda \cdot v, \quad Aw = v + \lambda \cdot w$$

*W szczególności, dla macierzy $S = \begin{bmatrix} v \mid w \end{bmatrix}$ mamy następującą postać Jordana dla $A$:*

$$\Lambda = S^{-1}AS = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$

*Dowód.* Niech $0 \neq v \in \ker(A - \lambda I)$ będzie dowolnym wektorem własnym dla $\lambda$. Każdy wektor własny $A$ jest postaci $p \cdot v$, gdzie $v$ jest pewnym wektorem własnym $A$ dla $\lambda$ oraz $0 \neq p \in \mathbb{F}$. Niech $0 \neq u \in \mathbb{F}^2$ będzie dowolnym wektorem liniowo niezależnym z $v$. W szczególności, $u$ nie jest wektorem własnym odpowiadającym $\lambda$, więc $(A - \lambda I)u \neq 0$. Ponieważ

$$0 = (A - \lambda I)^2 u = (A - \lambda I)(A - \lambda I)u,$$

więc $(A - \lambda I)u \in \ker(A - \lambda I)$ jest wektorem własnym odpowiadającym $\lambda$. Stąd

$$(A - \lambda I)u = p \cdot v,$$

dla pewnego $0 \neq p \in \mathbb{R}$. Dla wektora $w = \frac{1}{p} \cdot u$ otrzymujemy, że $(A - \lambda I)w = v$, czyli $Aw = v + \lambda \cdot w$. $\square$

| |
|---|
| Postać jordana macierzy zespolonej<br>Macierz zespolona $A \in M_{2\times 2}(\mathbb{C})$ ma jedną z dwóch postaci Jordana<br>$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \qquad \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}.$$ |

**Wniosek 10.7.** *(i) Jeśli $0 \neq A \in M_{2\times 2}(\mathbb{C})$ jest macierzą nilpotentną, to*

$$A = S \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} S^{-1}$$

*dla pewnej macierzy nieosobliwej $S$.*

*(ii) Jeśli $A \in M_{2\times 2}(\mathbb{C})$ jest taka, że $A^2 = A$, to $A = 0$ lub $A = I$ lub*

$$A = S \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} S^{-1}$$

*dla pewnej macierzy nieosobliwej $S$.*

*(iii) Jeśli $A \in M_{2\times 2}(\mathbb{C})$ jest taka, że $A^2 = I$, to $A = \pm I$ lub*

$$A = S \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} S^{-1}$$

*dla pewnej macierzy nieosobliwej $S$.*

*(iv) Jeśli $A \in M_{2\times 2}(\mathbb{C})$ jest taka, że $A^2 = -I$, to $A = \pm i \cdot I$ lub*

$$A = S \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} S^{-1}$$

*dla pewnej macierzy nieosobliwej $S$.*

**Wniosek 10.8.** *Dla dowolnej macierzy $A \in M_{2\times 2}(\mathbb{C})$ mamy*

*(i) $A$ jest podobna do $A^T$.*

*(ii) $A$ jest podobna do macierzy symetrycznej.*

*(iii) $A$ jest iloczynem macierzy symetrycznych.*

*Dowód.* Punkt (i). Możemy założyć, że $A$ jest w postaci Jordana. Jeśli jest ona diagonalna, to nie ma czego dowodzić. Załóżmy więc, że

$$A = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}.$$

Wtedy dla $S = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ mamy $S^{-1}AS = A^T$.

Punkt (ii). Ponownie możemy założyć, że $A = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$. Szukamy takiej macierzy nieosobliwej $S = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{C})$, że $S^{-1}AS$ jest symetryczna. Dla $\Delta = \det S \neq 0$ mamy

$$S^{-1}AS = \begin{bmatrix} cd + \lambda\Delta & d^2 \\ -c^2 & -cd + \lambda\Delta \end{bmatrix}.$$

Ponieważ szukamy macierzy symetrycznej, więc musi zachodzić warunek $c^2 + d^2 = 0$. Wystarczy więc przyjąć takie $a, b, c, d$, że $ad - bc \neq 0$ i $c^2 + d^2 = 0$. Zwróćmy uwagę, że $S$ nie może być rzeczywista. Przykładowo, dla $d = i = a$, $c = 1$ i $b = 0$ mamy

$$S = \begin{bmatrix} i & 0 \\ 1 & i \end{bmatrix}, \quad S^{-1}AS = \begin{bmatrix} \lambda - i & 1 \\ 1 & \lambda + i \end{bmatrix}.$$

Punkt (iii). Zauważmy najpierw, że możemy założyć, że $A$ jest w postaci Jordana $\Lambda$, bo jeśli $\Lambda = BC$ jest iloczynem macierzy symetrycznych, to

$$A = S\Lambda S^{-1} = (SBS^T)((S^{-1})^T CS^{-1})$$

jest również iloczynem macierzy symetrycznych. Dla $\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ wystarczy przyjąć

$$B = \begin{bmatrix} \lambda_1 & 0 \\ 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

Jeśli $\Lambda = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$, to przyjmujemy

$$B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & \lambda \\ \lambda & 1 - \lambda \end{bmatrix}. \qquad \square$$

U Rozważmy macierz $A$ w postaci bloku Jordana

$$A = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}.$$

Dla macierzy

$$S = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

mamy $S^{-1} = S$ oraz

$$\begin{aligned} S^{-1}AS &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \lambda & 0 \\ 1 & \lambda \end{bmatrix} \\ &= A^T. \end{aligned}$$

Naturalne jest pytanie o postać macierzy $A \in M_{2\times 2}(\mathbb{R})$, gdy jej wielomian charakterystyczny $p_A$ ma dwie zespolone wartości własne $\lambda_1 = \alpha + i\beta$, $\lambda_2 = \alpha - i\beta$ oraz $\beta \neq 0$. Macierz $A \in M_{2\times 2}(\mathbb{R})$ możemy traktować jako element przestrzeni $M_{2\times 2}(\mathbb{C})$. Istnieje więc wektor własny $v \in \mathbb{C}^2$ odpowiadający wartości własnej $\lambda_1 = \alpha + i\beta$. Wektor $v \in \mathbb{C}^2$ wygodnie jest zapisać w postaci

$$v = \begin{bmatrix} a_1 + ia_2 \\ b_1 + ib_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} + i \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} := v_1 + iv_2, \quad v_1, v_2 \in \mathbb{R}^2.$$

Wówczas

$$Av = Av_1 + iAv_2.$$

Ponadto

$$A(v_1 - iv_2) = (\alpha - i\beta) \cdot (v_1 - iv_2),$$

czyli wektor $v_1 - iv_2$ jest wektorem własnym odpowiadającym wartości własnej $\lambda_2 = \alpha - i\beta$.

Pokażemy, że wektory $v_1, v_2 \in \mathbb{R}^2$ są liniowo niezależne. Rzeczywiście, przypuśćmy, że $v_1 = p \cdot v_2$ dla pewnej liczby rzeczywistej $p \in \mathbb{R}$. Wtedy

$$A(v_1 + iv_2) = (\alpha + i\beta) \cdot (v_1 + iv_2) = (\alpha + i\beta)(p + i) \cdot v_2$$

oraz

$$A(v_1 + iv_2) = A((p + i)v_2) = (p + i) \cdot A(v_2),$$

więc $A(v_2) = (\alpha + i\beta) \cdot v_2$, co prowadzi do sprzeczności z warunkiem $\beta \neq 0$.

Zauważmy, że

$$A(v_1) + iA(v_2) = A(v_1 + iv_2) = (\alpha + i\beta) \cdot (v_1 + iv_2),$$

więc

$$A(v_1) = \alpha \cdot v_1 - \beta \cdot v_2, \quad A(v_2) = \beta \cdot v_1 + \alpha \cdot v_2.$$

Dla macierzy $S = \begin{bmatrix} v_1 \mid v_2 \end{bmatrix}$ otrzymujemy równość

$$\Lambda := S^{-1}AS = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}.$$

Dla $S_1 = \begin{bmatrix} v_2 \mid v_1 \end{bmatrix}$ mamy

$$S_1^{-1}AS_1 = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}.$$

**Wniosek 10.9** (Rzeczywista postać Jordana). *Jeśli $\lambda_1 = \alpha + i\beta$, $\lambda_2 = \alpha - i\beta$ z $\beta \neq 0$ są pierwiastkami wielomianu charakterystycznego $p_A$ dla macierzy $A \in M_{2\times 2}(\mathbb{R})$, to istnieje taka nieosobliwa macierz $S \in M_{2\times 2}(\mathbb{R})$, że*

$$\Lambda = S^{-1}AS = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}.$$

*Ponadto,* $S = \begin{bmatrix} v_1 \mid v_2 \end{bmatrix}$ *dla dowolnego takiego niezerowego wektora* $v_1 + iv_2 \in \mathbb{C}$*, że*

$$A(v_1 + iv_2) = (\alpha + i\beta) \cdot (v_1 + iv_2).$$

**Przykład 10.3** (Ciąg Fibonacciego)**.** Ciąg Fibonacciego $u_n$ jest zdefiniowany rekurencyjnie wzorem

$$u_{n+1} = u_n + u_{n-1}, \quad u_1 = 1, u_0 = 0.$$

Dla macierzy

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

mamy więc

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_n \\ u_{n-1} \end{bmatrix} = \begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix}.$$

Wynika stąd, że

$$\begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix} = A^n \begin{bmatrix} u_1 \\ u_0 \end{bmatrix} = A^n e_1.$$

Wielomian charakterystyczny $A$ ma postać $p_A(x) = x^2 - x - 1$, więc wartościami własnymi $A$ są

$$\lambda_+ = \frac{\sqrt{5}+1}{2}, \quad \lambda_- = \frac{1-\sqrt{5}}{2}.$$

Wektory własne mają postać

$$v_+ = [(1+\sqrt{5})/2, 1]^T, \quad v_- = [(1-\sqrt{5})/2, 1]^T.$$

Ponieważ

$$e_1 = \frac{1}{\sqrt{5}} \cdot (v_+ - v_-),$$

więc

$$\begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix} = \frac{1}{\sqrt{5}}(\lambda_+^n \cdot v_+ - \lambda_-^n \cdot v_-),$$

czyli otrzymujemy *formułę Bineta*

$$u_n = \frac{1}{\sqrt{5}} \left( \frac{1+\sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left( \frac{1-\sqrt{5}}{2} \right)^n.$$

**Przykład 10.4.** Rozważmy macierz

$$A = \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix}.$$

Jej wielomian charakterystyczny ma postać

$$p_A(x) = x^2 - 4x + 4 = (x-2)^2.$$

Stąd $\lambda = 2$ jest dwukrotną wartością własną macierzy $A$. Ponieważ

$$A - 2I = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix},$$

więc $[x, y]^T \in \ker(A - 2I)$ jeśli $x + y = 0$, czyli

$$\ker(A - 2I) = \operatorname{span}\left\{ [1, -1]^T \right\}.$$

Wektor $v = [1, -1]^T$ jest wektorem własnym i każdy inny wektor własny jest liniowo zależny z $v$.

Szukamy teraz takiego wektora $w = [x, y]^T \in \mathbb{R}^2$, że

$$(A - 2I)w = v = [1, -1]^T.$$

Musimy więc rozwiązać układ równań

$$x + y = 1, \quad -x - y = -1.$$

Rozwiązaniami są wektory postaci $[x, 1 - x]^T = x[1, -1]^T + [0, 1]^T$. Możemy więc przyjąć, że

$$w = [1, 0]^T = e_1.$$

Dla macierzy

$$S = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix},$$

mamy

$$S^{-1}AS = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

**Przykład 10.5.** Macierz

$$A = \begin{bmatrix} 0 & -4 \\ 1 & 0 \end{bmatrix}$$

ma zespolone wartości własne $\lambda_1 = 2i$ oraz $\lambda_2 = -2i$. Znajdziemy wektor własny $v_1 = [a + bi, c + di]^T \in \mathbb{C}^2$ dla $\lambda_1 = 2i$. Szukamy niezerowych rozwiązań równania

$$\begin{aligned} 0 = (A - 2iI)v &= \begin{bmatrix} -2i & -4 \\ 1 & -2i \end{bmatrix} \begin{bmatrix} a + bi \\ c + di \end{bmatrix} \\ &= \begin{bmatrix} -2ia + 2b - 4c - 4di \\ a + bi - 2ic + 2d \end{bmatrix} \\ &= \begin{bmatrix} -2ia + 2(b - 2c) - 2i(a + 2d) \\ a + 2d + i(b - 2c) \end{bmatrix}. \end{aligned}$$

Stąd $a + 2d = 0$ oraz $b - 2c = 0$. Możemy więc przyjąć, że

$$v_1 = [2, -i]^T = [2, 0]^T + i[0, -1]^T.$$

Dla wartości własnej $\lambda_2 = -2i$ wektorem własnym jest więc $v_2 = [2, i]^T$. Dla macierzy

$$S = \begin{bmatrix} 2 & 2 \\ -i & i \end{bmatrix}$$

mamy

$$S^{-1}AS = \frac{1}{4i} \begin{bmatrix} i & -2 \\ i & 2 \end{bmatrix} \begin{bmatrix} 0 & -4 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -i & i \end{bmatrix} = \begin{bmatrix} 2i & 0 \\ 0 & -2i \end{bmatrix}$$

Aby znaleźć rzeczywistą postać Jordana, rozważamy macierz

$$P = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$$

której kolumnami są część rzeczywista i urojona wektora $v_1 = [2, -i]^T = [2, 0]^T + i[0, -1]^T$. Wtedy

$$P^{-1}AP = \begin{bmatrix} 1/2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & -4 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}.$$

## 10.2 Rozkład Jordana w wymiarze 3

Podamy dowód Twierdzenia Jordana dla macierzy zespolonej $A \in M_{3\times 3}(\mathbb{C})$. Zastosujemy podejście, które uogólnia się na wyższe wymiary.

**Twierdzenie 10.5** (Rozkład Jordana macierzy zespolonej)**.** *Niech $A \in M_{3\times 3}(\mathbb{C})$. Istnieje taka macierz nieosobliwa $S \in M_{3\times 3}(\mathbb{C})$, że $S^{-1}AS$ jest jednej z poniższych postaci:*

$$\begin{bmatrix} \boxed{\lambda_1} & 0 & 0 \\ 0 & \boxed{\lambda_2} & 0 \\ 0 & 0 & \boxed{\lambda_3} \end{bmatrix}, \quad \left[\begin{array}{c|cc} \lambda_1 & 0 & 0 \\ \hline 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{array}\right], \quad \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{bmatrix}.$$

*Ponadto, $\lambda_i \in \mathbb{C}$ są wartościami własnymi macierzy $A$.*

**Lemat 10.4.** *Załóżmy, że wartości własne $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$ macierzy $A \in M_{3\times 3}(\mathbb{C})$ są różne. Wtedy odpowiadające im wektory własne $v_1, v_2, v_3$ są liniowo niezależne oraz dla $S = \begin{bmatrix} v_1 \mid v_2 \mid v_3 \end{bmatrix}$ mamy*

$$S^{-1}AS = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3).$$

*Dowód.* Zauważmy najpierw, że $v_1$ i $v_2$ są liniowo niezależne. Przypuśćmy, że $v_2 = x \cdot v_1$ dla pewnego $x \in \mathbb{C}$. Wtedy $x \neq 0$, bo $v_2 \neq 0$. Mamy

$$\begin{aligned} \lambda_2 \cdot v_2 &= Av_2 \\ &= A(x \cdot v_1) \\ &= x\lambda_1 \cdot v_1 \\ &= \lambda_1 \cdot v_2, \end{aligned}$$

czyli $\lambda_1 = \lambda_2$, sprzeczność.

Przypuśćmy, że $v_3 \in \operatorname{span}\{v_1, v_2\}$. Wtedy $v_3 = x_1 \cdot v_1 + x_2 \cdot v_2$ dla pewnych $x_1, x_2 \in \mathbb{C}$. Przynajmniej jedna z liczb $x_1$, $x_2$ jest różna od zera, bo $v_3 \neq 0$. Otrzymujemy, że

$$\begin{aligned} \lambda_3 \cdot v_3 &= Av_3 \\ &= A(x_1 \cdot v_1 + x_2 \cdot v_2) \\ &= x_1\lambda_1 \cdot v_1 + x_2\lambda_2 \cdot v_2, \end{aligned}$$

czyli

$$\lambda_3(x_1 \cdot v_1 + x_2 \cdot v_2) = x_1\lambda_1 \cdot v_1 + x_2\lambda_2 \cdot v_2.$$

W efekcie

$$x_1(\lambda_3 - \lambda_1) \cdot v_1 + x_2(\lambda_2 - \lambda_1) \cdot v_2 = 0,$$

co jest sprzeczne z liniową niezależnością $v_1$ i $v_2$. □

**Lemat 10.5.** *Niech $A \in M_{3\times 3}(\mathbb{C})$. Istnieje taka macierz nieosobliwa $S \in M_{3\times 3}(\mathbb{C})$, że macierz $S^{-1}AS$ jest górnie trójkątna.*

*Dowód.* Niech $\lambda \in \mathbb{C}$ będzie wartością własną $A$ i niech $v \in \mathbb{C}^3$ będzie odpowiadającym jej wektorem własnym. Uzupełniamy $v$ do bazy $v$, $u$, $w$ dla $\mathbb{C}^3$. Rozważmy macierz $W = \begin{bmatrix} v \mid u \mid w \end{bmatrix}$. Macierz $W^{-1}AW$ ma postać blokową

$$W^{-1}AW = \left[\begin{array}{c|cc} \lambda & * & * \\ \hline 0 & a & b \\ 0 & c & d \end{array}\right].$$

Z rozkładu Jordana w wymiarze 2 istnieje taka macierz nieosobliwa $V \in M_{2\times 2}(\mathbb{C})$, że

$$V^{-1} \begin{bmatrix} a & b \\ c & d \end{bmatrix} V = T$$

jest górnie trójkątna. Definiujemy macierz nieosobliwą

$$U = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & V \end{array}\right] \in M_{3\times 3}(\mathbb{C}).$$

Wtedy

$$\begin{aligned} U^{-1}W^{-1}AWU &= U^{-1}\left[\begin{array}{c|cc} \lambda & * & * \\ \hline 0 & a & b \\ 0 & c & d \end{array}\right] U \\ &= \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & V^{-1} \end{array}\right] \left[\begin{array}{c|cc} \lambda & * & * \\ \hline 0 & a & b \\ 0 & c & d \end{array}\right] \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & V \end{array}\right] \\ &= \left[\begin{array}{c|c} \lambda & * \\ \hline 0 & T \end{array}\right], \end{aligned}$$

jest górnie trójkątna i teza zachodzi dla $S = WU$. □

**Lemat 10.6.** *Rozważmy macierz górnie trójkątną z* $0$ *jako jedyną wartością własną*

$$A = \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix}.$$

*Wtedy zachodzi jeden z warunków*

*(i)* $A = 0$*,*

*(ii)* $\dim\ker A = 2$ *i* $A$ *jest podobna do macierzy*

$$\left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right],$$

*(iii)* $\dim\ker A = 1$ *oraz* $A$ *jest podobna do macierzy*

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

*Dowód.* Sprawdzamy, że

$$A^2 = \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & ac \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$A^3 = \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & ac \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 0.$$

Mamy trzy możliwości

1. $A = 0$,
2. $A \neq 0$, $A^2 = 0$,
3. $A^2 \neq 0$, $A^3 = 0$.

W przypadku (ii) $a = 0$ lub $c = 0$, bo $A^2 = 0$ i macierz $A$ jest jednej z postaci

$$A = \begin{bmatrix} 0 & 0 & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix} \quad c \neq 0, \quad A = \begin{bmatrix} 0 & a & b \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad a \neq 0, \quad A = \begin{bmatrix} 0 & 0 & b \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad b \neq 0.$$

We wszystkich przypadkach $\operatorname{rank} A = \dim \operatorname{im} A = 1$ oraz

$$\operatorname{im} A \subset \ker A, \quad \dim \ker A = 2.$$

Inkluzja $\operatorname{im} A \subset \ker A$ wynika stąd, że $A^2 = 0$. Niech $v_1 \in \operatorname{im} A$ będzie niezerowym wektorem. Istnieje więc taki (niezerowy) wektor $v_2 \in \mathbb{C}^3$, że $Av_2 = v_1$. Wybieramy teraz wektor $v_3 \in \ker A$ liniowo niezależny z $v_1$. Rozważmy macierz

$$S = \begin{bmatrix} v_1 \mid v_2 \mid v_3 \end{bmatrix}.$$

Ponieważ $v_2 \notin \ker A$, więc $v_1, v_2, v_3$ są liniowo niezależne, czyli $S$ jest nieosobliwa. Ponadto,

$$S^{-1}AS = \left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right]$$

W przypadku (iii) mamy $a \neq 0$ i $c \neq 0$, więc

$$\operatorname{rank} A = \dim \operatorname{im} A = 2.$$

Ponadto,

$$\operatorname{im} A^2 \subset \ker A, \quad \dim \ker A = 1, \quad \text{więc} \quad \operatorname{im} A^2 = \ker A.$$

Niech $v \in \ker A$ będzie wektorem własnym dla wartości własnej 0. Ponieważ

$$\operatorname{im} A^2 = \ker A,$$

więc istnieje taki wektor $u \in \mathbb{C}^3$, że $A^2 u = v$. Wektory

$$v, \quad Au, \quad u$$

są liniowo niezależne. Rzeczywiście, przypuśćmy, że $x \cdot v + y \cdot Au + z \cdot u = 0$ dla pewnych $x, y, z \in \mathbb{C}$. Wtedy

$$0 = A^2(x \cdot v + y \cdot Au + z \cdot u) = z \cdot v,$$

czyli $z = 0$, bo $v \neq 0$. W konsekwencji,

$$0 = A(x \cdot v + y \cdot Au) = y \cdot v,$$

czyli $y = 0$. Z równości $x \cdot v = 0$ również $x = 0$. Dla

$$S = \begin{bmatrix} v \mid Au \mid u \end{bmatrix}$$

mamy

$$S^{-1}AS = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad \square$$

**Wniosek 10.10.** *Załóżmy, że* $\lambda \in \mathbb{C}$ *jest* 3*-krotną wartością własną macierzy* $A \in M_{3\times 3}(\mathbb{C})$. *Wtedy*

*(i) jeśli* $\dim \ker(A - \lambda I) = 3$, *to* $A = \lambda I$,

*(ii) jeśli* $\dim\ker(A-\lambda I)=2$*, to* $A$ *jest podobna do macierzy*

$$\left[\begin{array}{cc|c} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ \hline 0 & 0 & \lambda \end{array}\right].$$

*(iii) jeśli* $\dim\ker(A-\lambda I)=1$*, to* $A$ *jest podobna do macierzy*

$$\begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}.$$

*Dowód.* Z lematu 11.5 możemy założyć, że $A$ jest górnie trójkątna. Stosujemy lemat 11.6 do macierzy $A-\lambda I$. □

**Lemat 10.7.** *Załóżmy, że* $A \in M_{3\times 3}(\mathbb{C})$ *jest macierzą górnie trójkątną postaci*

$$A = \begin{bmatrix} \lambda & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix}, \quad \lambda \neq 0.$$

*Wtedy*

*(i)* $\dim\ker A < 3$*,*

*(ii) jeśli* $\dim\ker A = 2$*, to* $S^{-1}AS = \begin{bmatrix} \boxed{\lambda} & 0 & 0 \\ 0 & \boxed{0} & 0 \\ 0 & 0 & \boxed{0} \end{bmatrix}$ *dla pewnej macierzy nieosobliwej* $S \in M_{3\times 3}(\mathbb{C})$*,*

*(iii) jeśli* $\dim\ker A = 1$*, to* $S^{-1}AS = \left[\begin{array}{c|cc} \lambda & 0 & 0 \\ \hline 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right]$ *dla pewnej macierzy nieosobliwej* $S \in M_{3\times 3}(\mathbb{C})$.

*Dowód.* Zauważmy, że

$$\dim\operatorname{im} A = \begin{cases} 1, & \text{gdy } c = 0, \\ 2, & \text{gdy } c \neq 0, \end{cases}$$

czyli

$$\dim\ker A = \begin{cases} 2, & \text{gdy } c = 0, \\ 1, & \text{gdy } c \neq 0, \end{cases}$$

Załóżmy, że $\dim\ker A = 2$. Niech $v_2, v_3 \in \ker A$ będą bazą dla jądra. Jeśli $v_1$ jest wektorem własnym dla $\lambda$, to $v_1, v_2, v_3$ są liniowo niezależne. Przypuśćmy bowiem, że $v_1 = x_2 \cdot v_2 + x_3 \cdot v_3$. Wtedy

$$\lambda \cdot v_1 = Av_1 = x_2 \cdot Av_2 + x_3 \cdot Av_3 = 0,$$

sprzeczność. Dla macierzy nieosobliwej $S = \begin{bmatrix} v_1 \mid v_2 \mid v_3 \end{bmatrix}$ mamy

$$S^{-1}AS = \begin{bmatrix} \boxed{\lambda} & 0 & 0 \\ 0 & \boxed{0} & 0 \\ 0 & 0 & \boxed{0} \end{bmatrix}.$$

Niech teraz $\dim\ker A = 1$, czyli $c \neq 0$. Zauważmy, że teza zajdzie, gdy znajdziemy taką bazę $v_1, v_2, v_3$, że

$$Av_1 = \lambda \cdot v_1, \quad Av_2 = 0, \quad Av_3 = v_2.$$

Z wyborem $v_1$ i $v_2$ nie mamy problemu: bierzemy wektory własne dla $\lambda$ i 0. Musimy dobrać wektor $v_3$, aby $Av_3 = v_2$. W tym celu wystarczy zauważyć, że

$$\ker A \subset \operatorname{im} A.$$

Rzeczywiście, $A[x,y,z]^T = [\lambda x + ay + bz, cz, 0]^T$, czyli $[x,y,z]^T \in \ker A$, gdy $z = 0$ i $x = \frac{a}{\lambda}y$, czyli jądro jest generowane przez wektor $[a/\lambda, 1, 0]^T$. Sprawdzamy łatwo, że jest on w obrazie $A$ generowanym przez wektory $[\lambda, 0, 0]^T$ i $[b, c, 0]^T$.

Pozostaje sprawdzić, że $v_1, v_2, v_3$ są liniowo niezależne. Wiemy, że $v_1$ i $v_2$ są liniowo niezależne. Przypuśćmy, że $v_3 = x_1 \cdot v_1 + x_2 \cdot v_2$. Wtedy

$$\begin{aligned} v_2 &= Av_3 \\ &= x_1 \cdot Av_1 + x_2 \cdot Av_2 \\ &= x_1\lambda \cdot v_1, \end{aligned}$$

co jest sprzeczne z liniową niezależnością $v_1$ i $v_2$. □

**Wniosek 10.11.** *Załóżmy, że $\lambda, \lambda_1 \in \mathbb{C}$ są różnymi wartościami własnymi macierzy $A \in M_{3\times 3}(\mathbb{C})$ oraz $\lambda_1$ ma krotność 2. Wtedy*

*(i)* $\dim\ker(A - \lambda_1 I) < 3$,

*(ii) jeśli* $\dim\ker(A - \lambda_1 I) = 2$, *to $A$ jest podobna do macierzy*

$$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{bmatrix}.$$

*(iii) jeśli* $\dim\ker(A - \lambda_1 I) = 1$, *to $A$ jest podobna do macierzy*

$$\left[\begin{array}{c|cc} \lambda & 0 & 0 \\ \hline 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{array}\right].$$

*Dowód.* Z lematu 11.5 możemy założyć, że $A$ jest górnie trójkątna. Stosujemy rozumowanie jak w dowodzie lematu 11.7. □

Zajmiemy się teraz macierzami rzeczywistymi $A \in M_{3\times 3}(\mathbb{R})$. Wtedy wielomian charakterystyczny $p_A$ jest stopnia 3 i ma współczynniki rzeczywiste. Jeśli liczba zespolona $\lambda \in \mathbb{C}$ jest pierwiastkiem $p_A$, to jest nim również $\overline{\lambda}$. Wynika stąd, że istnieje co najmniej jeden pierwiastek rzeczywisty $p_A$. Jest on oczywiście wartością własną $p_A$. Jeśli $A$ ma trzy (liczone z krotnościami) rzeczywiste wartości własne, to $A$ ma taką samą postać Jordana jak w przypadku zespolonym.

Pozostaje nam rozważyć przypadek, gdy $p_A$ ma rzeczywisty pierwiastek $\lambda_1 \in \mathbb{R}$ i parę pierwiastków $\lambda_2 = \alpha + i\beta$ oraz $\overline{\lambda_2} = \alpha - i\beta$ z $\beta \neq 0$.

**Wniosek 10.12** (Rzeczywista postać Jordana)**.** *Jeśli dla $A \in M_{3\times 3}(\mathbb{R})$ ma rzeczywisty pierwiastek $\lambda_1 = \lambda \in \mathbb{R}$ i parę pierwiastków sprzężonych $\lambda_2 = \alpha + i\beta$ oraz $\overline{\lambda_2} = \alpha - i\beta$ z $\beta \neq 0$, to istnieje taka nieosobliwa macierz $S \in M_{3\times 3}(\mathbb{R})$, że*

$$S^{-1}AS = \left[\begin{array}{c|cc} \lambda & 0 & 0 \\ \hline 0 & \alpha & \beta \\ 0 & -\beta & \alpha \end{array}\right].$$

*Dowód.* Niech $v \in \mathbb{R}^3$ będzie wektorem własnym dla $\lambda$, a $w = w_1 + iw_2$ zespolonym wektorem własnym dla $\alpha + i\beta$. Wtedy $\overline{w} = w_1 - iw_2$ jest wektorem własnym dla $\alpha - i\beta$. Z naszych rozważań w wymiarze 2 wystarczy pokazać, że $v, w_1, w_2$ są liniowo niezależne nad $\mathbb{R}$, bo wtedy teza zachodzi dla $S = \left[\begin{array}{c|c|c} v & w_1 & w_2 \end{array}\right]$. Wiemy już, że $w_1$ i $w_2$ są liniowo niezależne. Przypuśćmy, że

$$v = x_1 \cdot w_1 + x_2 \cdot w_2,$$

dla pewnych $x_1, x_2 \in \mathbb{R}$. Ponieważ

$$w_1 = \frac{1}{2} \cdot (w + \overline{w}), \quad w_2 = \frac{1}{2i} \cdot (w - \overline{w}),$$

więc

$$v = \frac{x_1 - ix_2}{2} \cdot w + \frac{x_1 + ix_2}{2} \cdot \overline{w},$$

czyli $v, w, \overline{w}$ są liniowo zależne nad $\mathbb{C}$. Prowadzi to do sprzeczności, bo wektory te odpowiadają różnym wartościom własnym. □

**Przykład 10.6.** Macierz

$$A = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 3 & -2 \\ 0 & 1 & 1 \end{bmatrix}$$

ma wartości własne $\lambda_1 = -3$, $\lambda_2 = 2 + i$, $\lambda_3 = \overline{\lambda_2} = 2 - i$. Odpowiadającymi im wektorami własnymi są przykładowo

$$v_1 = [1, 0, 0]^T = e_1, \quad v_2 = [0, 1 + i, 1]^T, \quad v_3 = [0, 1 - i, 1]^T.$$

Dla

$$S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1+i & 1-i \\ 0 & 1 & 1 \end{bmatrix}$$

mamy

$$S^{-1}AS = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 2+i & 0 \\ 0 & & 2-i \end{bmatrix}.$$

Aby otrzymać postać rzeczywistą, rozważamy macierz

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

której druga i trzecia kolumna, to odpowiednio, część rzeczywista i urojona wektora własnego $v_2 = [0, 1, 1]^T + i[0, 1, 0]^T$. Otrzymujemy, że

$$P^{-1}AP = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 2 \end{bmatrix}.$$

**Przykład 10.7.** Wartości własne macierzy

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 1 & 1 & 2 \end{bmatrix}$$

to $\lambda_1 = 1$ i $\lambda_2 = \lambda_3 = 2$. Sprawdzamy łatwo, że odpowiadające im wektory własne, to

$$v_1 = [1, 1, -2]^T, \quad v_2 = [0, 0, 1]^T.$$

Krotność algebraiczna wartości własnej $\lambda_2 = 2$ jest równa 2, ale jej krotność geometryczna jest równa 1. Macierz $A$ nie ma więc bazy wektorów własnych, czyli nie jest diagonalizowalna. Aby wyznaczyć postać Jordana $A$, szukamy takiego wektora $w$, że

$$(A - 2I)w = v_2.$$

Ponieważ

$$A - 2I = \begin{bmatrix} -1 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix},$$

więc możemy przyjąć, że $w = [0, 1, 0]^T$. Dla macierzy

$$S = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix}$$

mamy

$$S^{-1}AS = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

**Przykład 10.8.** Macierz

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 2 \end{bmatrix}$$

ma tylko jedną wartość własną $\lambda = 2$ o krotności algebraicznej 3. Zauważmy, że

$$A - 2I = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix},$$

więc

$$\ker(A - 2I) = \operatorname{span}\{e_1, e_3\}.$$

Ponadto,

$$\operatorname{im}(A - 2I) = \operatorname{span}\{[1, 0, -1]^T\} \subset \ker(A - 2I).$$

Szukamy teraz takiego wektora $w$, że

$$(A - 2I)w = [1, 0, -1]^T.$$

Łatwo sprawdzamy, że możemy przyjąć $w = [0, 1, 0]^T = e_2$. Dla bazy

$$v_1 = [1, 0, -1]^T, \quad v_2 = [0, 1, 0]^T, \quad v_3 = [1, 0, 0]^T$$

Dla

$$S = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

mamy

$$S^{-1}AS = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

**Przykład 10.9.** Przyglądniemy się teraz macierzom antysymetrycznym $A \in M_{3\times 3}(\mathbb{R})$. Ponieważ $A^T = -A$, więc $A$ ma postać

$$A = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}, \quad a, b, c \in \mathbb{R}.$$

Zakładamy, że $A$ jest niezerowa, czyli $a^2+b^2+c^2 \neq 0$. Ponieważ $A$ jest antysymetryczna i wymiar $n = 3$ jest nieparzysty, więc $\det A = 0$. W szczególności, 0 jest zawsze wartością własną. Ma ona krotność 1, bo jak łatwo sprawdzić pozostałe wartości własne, to $\pm i\sqrt{a^2+b^2+c^2}$. Wektor własny dla wartości własnej 0 jest równy

$$v_A = [-c, b, -a]^T.$$

Ponadto jeśli $A, B \in M_{3\times 3}(\mathbb{R})$ są antysymetryczne, to

$$\operatorname{tr}(AB) = -2(v_A|v_B).$$

Co więcej, macierz

$$[A, B] := AB - BA$$

jest antysymetryczna oraz

$$v_{[A,B]} = v_A \times v_B.$$

## 10.3 Diagonalizacja w wymiarach 2 i 3

Niech

$$A = \begin{bmatrix} a & c \\ c & b \end{bmatrix}$$

będzie rzeczywistą macierzą symetryczną. Równanie charakterystyczne dla $A$ ma postać

$$p_A(x) = x^2 - (a+b)x + ab - c^2 = 0.$$

Ponieważ

$$\Delta := (a+b)^2 - 4(ab - c^2) = (a-b)^2 + 4c^2 \geqslant 0,$$

więc wartości własne $\lambda_1$, $\lambda_2$ macierzy $A$ są rzeczywiste oraz

$$\lambda_1 = \frac{a+b+\sqrt{\Delta}}{2}, \quad \lambda_2 = \frac{a+b-\sqrt{\Delta}}{2}.$$

**Twierdzenie 10.6.** *Dla macierzy rzeczywistej $A \in M_{2\times 2}(\mathbb{R})$ następujące warunki są równoważne*

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

*(ii)* $\mathbb{R}^2$ *ma bazę ortonormalną złożoną z wektorów własnych $A$,*

*(iii) Istnieje taka macierz ortogonalna $Q \in M_{2\times 2}(\mathbb{R})$, że*

$$Q^T AQ = D,$$

*gdzie $D = \operatorname{diag}(\lambda_1, \lambda_2)$ jest diagonalna z wartościami własnymi $A$ na przekątnej.*

*Dowód.* $(i) \Rightarrow (ii)$

Załóżmy, że $A$ jest symetryczna. Jeśli $c = 0$, to $e_1$, $e_2$ jest bazą ortonormalną złożoną z wektorów własnych macierzy $A = \operatorname{diag}(a, b)$.

Jeśli $c \neq 0$, to $\Delta > 0$ i $A$ ma dwie różne rzeczywiste wartości własne $\lambda_1$, $\lambda_2$. Niech $v_1$, $v_2$ będą odpowiadającymi im wektorami własnymi o normie 1. Pokażemy, że wektory $v_1$, $v_2$ są ortogonalne. Mamy

$$\begin{aligned} \lambda_1(v_1|v_2) &= (\lambda_1 v_1|v_2) \\ &= (Av_1|v_2) \\ &= (v_1|Av_2) \\ &= (v_1|\lambda_2 v_2) \\ &= \lambda_2(v_1|v_2), \end{aligned}$$

więc $(v_1|v_2) = 0$, bo $\lambda_1 \neq \lambda_2$.

*(ii)* $\Rightarrow$ *(iii)* Wystarczy przyjąć $Q = \left[\begin{array}{c|c} v_1 & v_2 \end{array}\right]$ dla bazy ortonormalnej $v_1$, $v_2$ złożonej z wektorów własnych.

*(iii)* $\Rightarrow$ *(i)*

Wiemy, że $A = QDQ^T$. Wtedy

$$\begin{aligned} A^T &= (Q^T)^T D^T Q^T \\ &= QDQ^T \\ &= A. \end{aligned}$$

□

Rozważmy teraz macierz symetryczną $A \in M_{3\times3}(\mathbb{R})$. Istnieje $\lambda_1 \in \mathbb{R}$ rzeczywista wartość własna dla $A$. Niech $v_1$ będzie odpowiadającym jej wektorem własnym o normie 1. Niech $V \subset \mathbb{R}^3$ będzie płaszczyzną prostopadłą do wektora $v_1$. Wybierzmy bazę ortonormalną $v_2$, $v_3 \in V$ dla $V$. Wtedy $v_1$, $v_2$, $v_3$ jest bazą ortonormalną dla $\mathbb{R}^3$.

Pokażemy, że $Av_2$, $Av_3 \in V$, czyli $A(V) \subset V$. Ze względu na symetrię wystarczy pokazać, że $Av_2 \in V$. Ponieważ $A$ jest symetryczna, więc

$$\begin{aligned} (v_1|Av_2) &= (Av_1|v_2) \\ &= (\lambda_1 v_1|v_2) \\ &= \lambda_1 (v_1|v_2) \\ &= 0. \end{aligned}$$

Rozważmy macierz ortogonalną $W = \left[v_1 \mid v_2 \mid v_3\right]$. Wtedy

$$W^{-1}AW = W^TAW = \left[\begin{array}{c|cc} \lambda_1 & 0 & 0 \\ \hline 0 & a & c \\ 0 & d & b \end{array}\right].$$

Ponieważ $W^TAW$ jest symetryczna, więc $d = c$, czyli

$$W^TAW = \left[\begin{array}{c|cc} \lambda_1 & 0 & 0 \\ \hline 0 & a & c \\ 0 & c & b \end{array}\right]$$

Wiemy już, że istnieje taka macierz ortogonalna $B \in M_{2\times2}(\mathbb{R})$, że

$$B^T \begin{bmatrix} a & c \\ c & b \end{bmatrix} B = \operatorname{diag}(\lambda_2, \lambda_3).$$

Dla macierzy ortogonalnej

$$U = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & B \end{array}\right] \in M_{3\times3}(\mathbb{R})$$

mamy

$$U^TW^TAWU = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3).$$

Macierz $Q = WU$ jest ortogonalna i jej kolumny są bazą ortonormalną złożoną z wektorów własnych.

**Wniosek 10.13.** *Dla macierzy rzeczywistej $A \in M_{3\times3}(\mathbb{R})$ następujące warunki są równoważne*

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

*(ii)* $\mathbb{R}^3$ *ma bazę ortonormalną złożoną z wektorów własnych* $A$,

*(iii) Istnieje taka macierz ortogonalna* $Q \in M_{3\times3}(\mathbb{R})$, *że*

$$Q^TAQ = D,$$

*gdzie* $D = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$ *jest diagonalna z wartościami własnymi* $A$ *na przekątnej.*

## 10.4 Twierdzenie Cayleya–Hamiltona w wymiarze 3

**Twierdzenie 10.7** (Twierdzenie Cayleya–Hamiltona)**.** *Niech $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$ będą wartościami własnymi macierzy $A \in M_{3\times 3}(\mathbb{C})$. Wtedy*

$$(A - \lambda_1 I)(A - \lambda_2 I)(A - \lambda_3 I) = 0.$$

*Dowód.* Z lematu 11.5 istnieje taka macierz nieosobliwa $S \in M_{3\times 3}(\mathbb{C})$, że $S^{-1}AS$ jest górnie trójkątna. Zauważmy, że

$$\begin{aligned}(S^{-1}AS - \lambda_1 I)(S^{-1}AS - \lambda_2 I)(S^{-1}AS - \lambda_3 I) = \\ &(S^{-1}AS - \lambda_1 S^{-1}S)(S^{-1}AS - \lambda_2 S^{-1}S)(S^{-1}AS - \lambda_3 S^{-1}S) = \\ &S^{-1}(A - \lambda_1 I)SS^{-1}(A - \lambda_2 I)SS^{-1}(A - \lambda_3 I)S = \\ &S^{-1}(A - \lambda_1 I)(A - \lambda_2 I)(A - \lambda_3 I)S.\end{aligned}$$

Wynika stąd, że teza zachodzi dla $A$ wtedy i tylko wtedy, gdy zachodzi dla $S^{-1}AS$. Możemy więc założyć, że $A$ jest górnie trójkątna postaci

$$A = \begin{bmatrix} \lambda_1 & a & b \\ 0 & \lambda_2 & c \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

Niech $v = [x, y, z]^T \in \mathbb{C}^3$ będzie dowolnym wektorem. Zauważmy, że

$$(A - \lambda_3 I)v = [x_1, y_1, 0]^T,$$

dla pewnych $x_1, y_1 \in \mathbb{R}$. Następnie,

$$(A - \lambda_2 I)[x_1, y_1, 0]^T = [x_2, 0, 0]^T,$$

dla pewnego $x_2 \in \mathbb{C}$. Ostatecznie,

$$(A - \lambda_1 I)[x_2, 0, 0]^T = [0, 0, 0]^T,$$

czyli

$$(A - \lambda_1 I)(A - \lambda_2 I)(A - \lambda_3 I)v = 0.$$ □

U

Z pewnością zauważyliście, że dokładnie te same argumenty pokazują, że dla macierzy górnie trójkątnej $A \in M_{n\times n}(\mathbb{C})$ zachodzi

$$p_A(A) = 0.$$

Dowód opiera się na obserwacji, że dla macierzy górnie trójkątnej $A$ podprzestrzenie $\operatorname{span}\{e_1, \ldots, e_k\}$ są niezmiennicze dla $A$. Pokażemy później, że każda macierz $A \in M_{n\times n}(\mathbb{C})$ jest podobna do macierzy górnie trójkątnej, więc twierdzenie Cayleya–Hamiltona zachodzi dla dowolnej macierzy $A \in M_{n\times n}(\mathbb{C})$.

TEST → Niech $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{R})$. Oceń prawdziwość poniższych zdań:

- Jeśli $a + b = c + d = 1$, to $[1, 1]^T$ jest wektorem własnym $A$.
- Jeśli $A$ ma prostą niezmienniczą, to $A$ ma rzeczywiste wartości własne.
- $A$ ma rzeczywisty wektor własny.
- Jeśli $A$ jest górnie trójkątna tzn. $c = 0$, to $A$ jest diagonalizowalna.
- Jeśli $\det A < 0$, to $A$ jest diagonalizowalna.
- Jeśli $\det A < 0$, to $A$ jest ortogonalnie diagonalizowalna.
- Jeśli $A^3 = 0$, to $A^2 = 0$.

- Jeśli $\mathbb{R}^2$ ma bazę złożoną z wektorów własnych $A$, to $A$ jest nieosobliwa.
- 0 jest jedyną wartością własną $A$ wtedy i tylko wtedy, gdy $A^2 = 0$.
- Jeśli $\operatorname{tr} A = 0$, to $A$ jest diagonalizowalna lub $A^2 = 0$.
- Jeśli $\det A = 0$, to $A$ ma prostą niezmienniczą.
- $A$ i $A^2$ mają takie same wartości własne.
- $A$ i $A^2$ mają takie same wektory własne.
- Jeśli $A$ jest nieosobliwa, to $A$ i $A^{-1}$ mają takie same wartości własne.
- Jeśli $A$ jest nieosobliwa, to $A$ i $A^{-1}$ mają takie same wektory własne.
- $A$ i $A^T$ mają takie same wartości własne.
- $A$ i $A^T$ mają takie same wektory własne.
- Jeśli $S \in M_{2\times 2}(\mathbb{R})$ jest nieosobliwa, to $A$ i $S^{-1}AS$ mają takie same wartości własne.
- Jeśli $S \in M_{2\times 2}(\mathbb{R})$ jest nieosobliwa, to $A$ i $S^{-1}AS$ mają takie same wektory własne.
- Jeśli $A$ jest nieosobliwa i diagonalizowalna, to $A^{-1}$ jest diagonalizowalna.
- Jeśli macierz kwadratowa $A$ jest diagonalizowalna, to $A^T$ jest również diagonalizowalna.
- Jeśli macierz kwadratowa ma dwa identyczne wiersze (kolumny), to 0 jest jej wartością własną.

$\longleftarrow$

# Część II

# Algebra liniowa z geometrią 2

# Rozdział 11

# Do czego zmierzamy?

Słowa kluczowe rozdziału
wektor własny ♢ wartość własna ♢ diagonalizacja ♢ baza wektorów własnych ♢ diagonalizacja ortogonalna ♢ macierz stochastyczna ♢ dominująca wartość własna ♢ wielomian charakterystyczny ♢ zespolona postać Jordana ♢ rzeczywista postać Jordana ♢ krotność geometryczna i algebraiczna ♢ twierdzenie Cayleya–Hamiltona ♢ diagonalizacja ortogonalna macierzy symetrycznej

Przedstawimy teraz motywację dla naszych dalszych rozważań. Jak już pewnie zauważyliście, szczególnie łatwe do analizy są macierze diagonalne

$$D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n), \quad \lambda_1, \ldots, \lambda_n \in \mathbb{R}.$$

Przykładowo

$$D^k = \operatorname{diag}(\lambda_1^k, \ldots, \lambda_n^k), \quad \det D = \lambda_1 \cdot \ldots \cdot \lambda_n.$$

Jeśli $D$ jest nieosobliwa, to $D^{-1} = \operatorname{diag}(1/\lambda_1, \ldots, 1/\lambda_n)$.

Załóżmy, że macierz $A \in M_{n\times n}(\mathbb{R})$ jest podobna do macierzy diagonalnej $D$, czyli istnieje taka macierz nieosobliwa

$$S = \left[\begin{array}{c|c|c} v_1 & \ldots & v_n \end{array}\right] \in M_{n\times n}(\mathbb{R}),$$

że

$$S^{-1}AS = D.$$

Wtedy $A = SDS^{-1}$, więc

$$A^k = SD^kS^{-1}, \quad \det A = \det D,$$

czyli $A$ jest równie prosta do analizy jak macierz diagonalna $D$.

Problem 1
Scharakteryzować macierze $A \in M_{n\times n}(\mathbb{R})$, które są diagonalizowalne, tzn. istnieje taka macierz nieosobliwa $S = \left[\begin{array}{c|c|c} v_1 & \ldots & v_n \end{array}\right] \in M_{n\times n}(\mathbb{R})$, że

$$S^{-1}AS = D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) = \left[\begin{array}{c|c|c} \lambda_1 \cdot e_1 & \ldots & \lambda_n \cdot e_n \end{array}\right].$$

Załóżmy, że $S^{-1}AS = D$. Wtedy $AS = SD$, czyli

$$\begin{aligned} \left[\begin{array}{c|c|c} Av_1 & \ldots & Av_n \end{array}\right] &= AS = SD \\ &= S\left[\begin{array}{c|c|c} \lambda_1 \cdot e_1 & \ldots & \lambda_n \cdot e_n \end{array}\right] \\ &= \left[\begin{array}{c|c|c} \lambda_1 \cdot Se_1 & \ldots & \lambda_n \cdot Se_n \end{array}\right] \\ &= \left[\begin{array}{c|c|c} \lambda_1 \cdot v_1 & \ldots & \lambda_n \cdot v_n \end{array}\right]. \end{aligned}$$

Stąd

$$Av_1 = \lambda_1 \cdot v_1, \quad \ldots, \quad Av_n = \lambda_n \cdot v_n.$$

Z powyższych rozważań wynika następujące:

**Twierdzenie 11.1** (Diagonalizacja)**.** *Dla macierzy $A \in M_{n\times n}(\mathbb{R})$ poniższe warunki są równoważne*

*(i) istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{R})$, że $S^{-1}AS$ jest diagonalna,*

*(ii) istnieje baza $v_1, \ldots, v_n$ dla $\mathbb{R}^n$ oraz takie skalary $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, że*

$$Av_i = \lambda_i \cdot v_i, \quad i = 1, \ldots, n.$$

Warto jakoś nazwać własność, którą spełniają wektory $v_i$ w punkcie (ii). Będą one pełniły podstawową rolę w naszych dalszych rozważaniach.

**Definicja 11.1.** **Wektor własny i wartość własna**
Powiemy, że **niezerowy** wektor $0 \neq v \in \mathbb{R}^n$ jest *wektorem własnym* macierzy rzeczywistej $A \in M_{n\times n}(\mathbb{R})$, jeśli istnieje taka liczba rzeczywista $\lambda \in \mathbb{R}$, że

$$Av = \lambda \cdot v.$$

Liczbę $\lambda \in \mathbb{R}$ nazywamy *wartością własną* macierzy $A \in M_{n\times n}(\mathbb{R})$ dla wektora własnego $v \in \mathbb{R}^n$.

Twierdzenie 11.1 możemy sformułować następująco:

**Twierdzenie 11.2** (Diagonalizacja = baza wektorów własnych)**.** *Dla macierzy $A \in M_{n\times n}(\mathbb{R})$ poniższe warunki są równoważne*

*(i) istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{R})$, że $S^{-1}AS$ jest diagonalna,*

*(ii) istnieje baza $v_1, \ldots, v_n$ dla $\mathbb{R}^n$ złożona z wektorów własnych $A$.*

Podamy geometryczną interpretację wektora własnego.

**Wniosek 11.1.** *Dla niezerowego wektora $v \in \mathbb{R}^n$ i macierzy $A \in M_{n\times n}(\mathbb{R})$ poniższe warunki są równoważne*

*(1) $v$ jest wektorem własnym $A$,*

*(2) prosta generowana przez $v$ jest niezmiennicza dla $A$ tzn.*

$$A(\mathrm{span}\{v\}) \subset \mathrm{span}\{v\}.$$

U

Nie każda macierz rzeczywista ma jakiś wektor własny. Wystarczy znaleźć macierz rzeczywistą $A$, która nie ma prostych niezmienniczych. Przykładowo, macierz obrotu o kąt $\frac{\pi}{2}$, czyli

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

nie ma żadnej prostej niezmienniczej.
Możemy na to spojrzeć jeszcze trochę inaczej. Wektor $0 \neq v \in \mathbb{R}^2$ jest wektorem własnym $A$, gdy dla pewnej liczby rzeczywistej $\lambda \in \mathbb{R}$ zachodzi warunek

$$(A - \lambda I)v = 0,$$

czyli $\ker(A - \lambda I) \neq 0$, więc $\det(A - \lambda I) = 0$ dla pewnej liczby rzeczywistej $\lambda \in \mathbb{R}$. Zauważmy, że w naszym przykładzie

$$\det(A - \lambda I) = \lambda^2 + 1 \neq 0,$$

dla $\lambda \in \mathbb{R}$.
Macierz rzeczywistą $A$ możemy traktować w naturalny sposób jako macierz zespoloną. Możemy więc spytać o istnienie zespolonych wektorów własnych $v \in \mathbb{C}^2$. Zauważmy, że $\det(A - \lambda I) = 0$ dla $\lambda = \pm i$. Rozważmy macierz

$$A - iI = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix}$$

i znajdźmy taki niezerowy wektor $v = [a + ib, c + id] \in \mathbb{C}^2$, że $(A - iI)v = 0$. Mamy

$$(A - iI)v = [-ia + b - c - id, a + ib - ic + d]^T = [0, 0]^T,$$

czyli

$$a + d = 0, \quad b - c = 0.$$

Wektor $v$ ma więc postać

$$v = [a + ib, b - ia]^T = [a + ib, -i(a + ib)]^T = (a + ib)[1, -i]^T.$$

Wynika stąd, że

$$A[1, -i]^T = i[1, -i]^T,$$

czyli $[1, -i]^T$ jest zespolonym wektorem własnym $A$ z wartością własną $i$. Analogicznie sprawdzamy, że $[1, i]^T$ jest zespolonym wektorem własnym $A$ dla wartości własnej $-i$. Macierz $A$ jest diagonalizowalna jako macierz zespolona. Dla

$$S = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$$

mamy

$$S^{-1}AS = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

U

Pojęcie wektora własnego możemy też zdefiniować w bardziej ogólnym kontekście. Niech $L : V \longrightarrow V$ będzie odwzorowaniem liniowym. Powiemy, że niezerowy wektor $v \in V$ jest wektorem własnym $L$ jeśli

$$L(v) = \lambda \cdot v.$$

Przestrzeń $V$ nie musi być skończenie wymiarowa. Przykładowo, niech $C^\infty(\mathbb{R}, \mathbb{R})$ będzie przestrzenią wszystkich funkcji $f : \mathbb{R} \longrightarrow \mathbb{R}$ mających ciągłą pochodną dowolnego rzędu. Wtedy odwzorowanie (operator różniczkowania)

$$L : C^\infty(\mathbb{R}, \mathbb{R}) \ni f \longmapsto f' \in C^\infty(\mathbb{R}, \mathbb{R})$$

jest liniowe. Ponadto dla $f = e^{\lambda x} \in C^\infty(\mathbb{R}, \mathbb{R})$ mamy

$$\begin{aligned} L(f) &= (e^{\lambda x})' \\ &= \lambda e^{\lambda x} \\ &= \lambda \cdot f, \end{aligned}$$

czyli $f = e^{\lambda x}$ jest wektorem własnym $L$ dla wartości własnej $\lambda$.
Uzasadnimy teraz, że jeśli $L(f) = \lambda \cdot f$, to $f = Ce^{\lambda x}$ dla pewnego $C \in \mathbb{R}$. Rzeczywiście, równość $f' = \lambda f$ oznacza, że $f' - \lambda f = 0$, czyli również

$$\underbrace{f'e^{-\lambda x} - \lambda e^{-\lambda x} f}_{=(fe^{-\lambda x})'} = 0,$$

czyli $fe^{-\lambda x}$ jest funkcją stałą, więc $fe^{-\lambda x} = C$.
Wynika stąd, że jądro $\ker(L - \lambda I)$ ma wymiar 1 oraz

$$\ker(L - \lambda I) = \operatorname{span}\{e^{\lambda x}\}.$$

U

Niech $L : V \longrightarrow V$ będzie odwzorowaniem liniowym i niech $v_1, \ldots, v_n \in V$ będzie bazą $V$. Rozważmy macierz $A \in M_{n\times n}(\mathbb{F})$ odwzorowania $L$ w tej bazie. Wektor $v = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$ jest wektorem własnym $L$ dla wartości własnej $\lambda \in \mathbb{F}$ wtedy i tylko wtedy, gdy wektor współrzędnych $[x_1, \ldots, x_n]^T \in \mathbb{F}^n$ jest wektorem własnym macierzy $A$ dla wartości własnej $\lambda$.

Możemy pójść jeszcze krok dalej. Wektory własne macierzy diagonalnej $D$ to wektory bazy standardowej $e_1, \ldots, e_n$. Ma ona własność

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

Takie bazy dla $\mathbb{R}^n$ będziemy nazywać *ortonormalnymi*. Wektory $e_i$ oraz $e_j$ są ortogonalne oraz mają normę 1.

PROBLEM 2
Scharakteryzować macierze $A \in M_{n\times n}(\mathbb{R})$, które są ortogonalnie diagonalizowalne tzn. istnieje taka macierz nieosobliwa $S = \left[\begin{array}{c|c|c} v_1 & \ldots & v_n \end{array}\right] \in M_{n\times n}(\mathbb{R})$, że

$$S^{-1}AS = D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = \left[\begin{array}{c|c|c} \lambda_1 \cdot e_1 & \ldots & \lambda_n \cdot e_n \end{array}\right]$$

oraz

$$(v_i|v_j) = \delta_{ij}.$$

Zaczniemy od przyglądnięcia się takim macierzom

$$S = \left[\begin{array}{c|c|c} v_1 & \ldots & v_n \end{array}\right] \in M_{n\times n}(\mathbb{R}),$$

że $(v_i|v_j) = \delta_{ij}$. Bezpośrednio z definicji iloczynu macierzy i definicji macierzy transponowanej wynika, że wtedy

$$S^T S = I, \quad \text{czyli} \quad S^{-1} = S^T.$$

Takie macierze będziemy nazywać *ortogonalnymi*. W problemie 2 pytamy więc o istnienie takiej macierzy ortogonalnej $S \in M_{n\times n}(\mathbb{R})$, że

$$S^{-1}AS = S^TAS = D$$

jest diagonalna. Zauważmy, że wtedy $A = SDS^T$ oraz

$$\begin{aligned} A^T &= (SDS^T)^T \\ &= (S^T)^T D^T S^T \\ &= SDS^T \\ &= A, \end{aligned}$$

czyli $A$ musi być symetryczna. Jest to warunek konieczny, aby zaszła sytuacja zaprezentowana w problemie 2. W następnych rozdziałach pokażemy, że macierze symetryczne to jedyne macierze rzeczywiste rozwiązujące problem 2.

## 11.1 Rozkład Jordana w wymiarze 2

Spróbujmy nabrać intuicji w przypadku wymiaru 2. Zakładamy, że ciało $\mathbb{F}$ jest równe $\mathbb{R}$ lub $\mathbb{C}$. Rozważmy macierz $A \in M_{2\times 2}(\mathbb{F})$.

**Definicja 11.2.** **Wektor własny i wartość własna**

Liczbę $\lambda \in \mathbb{F}$ nazywamy wartością własną macierzy $A \in M_{2\times 2}(\mathbb{F})$, jeśli istnieje taki **niezerowy** wektor $v \in \mathbb{F}^2$, że

$$Av = \lambda \cdot v.$$

Każdy taki wektor $v$ nazywamy wektorem własnym $A$ odpowiadającym wartości własnej $\lambda$.

U Niech $\mathbb{F} = \mathbb{R}$. Interpretacja geometryczna jest następująca. Jeśli $v \in \mathbb{R}^2$ jest wektorem własnym odpowiadającym wartości własnej $\lambda \in \mathbb{R}$, to prosta $l = \mathrm{span}\{v\}$ jest niezmiennicza dla $A$ tzn. jeśli $w \in l$, to $Aw \in l$.

Jeśli $v \in \mathbb{F}^2$ jest wektorem własnym odpowiadającym wartości własnej $\lambda \in \mathbb{F}$, to $(A-\lambda I)v = 0$. Oznacza to, że $\ker(A-\lambda I) \neq \{0\}$, czyli $\det(A-\lambda I) = 0$. Odwrotnie, jeśli $\det(A-\lambda I) = 0$ dla pewnego skalara $\lambda \in \mathbb{F}$, to macierz $A - \lambda I$ jest osobliwa. W szczególności, $\ker(A - \lambda I) \neq \{0\}$, czyli istnieje wektor własny $v \in \ker(A - \lambda I)$.

**Wniosek 11.2.** *Dla $A \in M_{2\times 2}(\mathbb{F})$ i $\lambda \in \mathbb{F}$ następujące warunki są równoważne*

*(1) $\lambda$ jest wartością własną $A$;*

*(2)* $\ker(A - \lambda I) \neq \{0\}$*;*

*(3)* $A - \lambda I$ *jest osobliwa;*

*(4)* $\det(A - \lambda I) = 0$*.*

Warunek (4) oznacza, że wartość własna $\lambda \in \mathbb{F}$ jest pierwiastkiem wielomianu

$$p_A(x) = \det(A - xI)$$

Wielomian $p_A$ nazywamy *wielomianem charakterystycznym* macierzy $A$. Bezpośredni rachunek pokazuje, że

$$\begin{aligned} p_A(x) &= \det \begin{bmatrix} a_{11} - x & a_{12} \\ a_{21} & a_{22} - x \end{bmatrix} \\ &= x^2 - (a_{11} + a_{22})x + (a_{11}a_{22} - a_{12}a_{22}) \\ &= x^2 - \operatorname{tr}(A)x + \det(A). \end{aligned}$$

U Macierze podobne mają takie same wielomiany charakterystyczne. Rzeczywiście, jeśli $B = S^{-1}AS$, to

$$\begin{aligned} \det(B - xI) &= \det(S^{-1}AS - xI) \\ &= \det(S^{-1}(A - xI)S) \\ &= \det(A - xI). \end{aligned}$$

Jeśli macierze $A$ i $B$ mają takie same wielomiany charakterystyczne, to nie muszą być podobne. Przykładowo, $A = 0$ oraz $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ mają taki sam wielomian charakterystyczny $p(x) = x^2$, ale nie są podobne, bo macierz zerowa jest podobna tylko do siebie samej.

Dla $\mathbb{F} = \mathbb{C}$ zasadnicze twierdzenie algebry gwarantuje, że wielomian $p_A$ ma dwa (niekoniecznie różne) pierwiastki zespolone $\lambda_1, \lambda_2 \in \mathbb{C}$. Są one wtedy wartościami własnymi macierzy $A \in M_{2\times 2}(\mathbb{C})$. Macierz zespolona ma więc zawsze dwie wartości własne. Nie oznacza to, że ma ona dwa liniowo niezależne wektory własne. Przykładowo, wartościami własnymi macierzy

$$A = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \quad \lambda \in \mathbb{C},$$

są $\lambda_1 = \lambda_2 = \lambda$. Mamy więc dwukrotną wartość własną $\lambda$. Zobaczmy jak wyglądają wektory własne. Niezerowy wektor $v = [v_1, v_2]^T \in \mathbb{C}^2$ jest wektorem własnym dla $\lambda$, jeśli $(A - \lambda I)v = 0$, czyli

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_2 \\ 0 \end{bmatrix}.$$

Stąd $v_2 = 0$, czyli wektor własny $v$ ma postać

$$\begin{bmatrix} v_1 \\ 0 \end{bmatrix} = v_1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = v_1 \cdot e_1.$$

W przypadku $\mathbb{F} = \mathbb{R}$, z algebraicznego punktu widzenia dla wielomianu charakterystycznego mamy trzy możliwości:

(i) $p_A$ ma dwa różne pierwiastki rzeczywiste;

(ii) $p_A$ ma dwukrotny pierwiastek rzeczywisty;

(iii) $p_A$ ma dwa sprzężone (nierzeczywiste) pierwiastki zespolone.

W przypadku (iii) macierz rzeczywista $A$ nie ma rzeczywistych wartości własnych.

**Wniosek 11.3.** *Niech $A \in M_{2\times 2}(\mathbb{C})$. Jeśli $\lambda_1, \lambda_2 \in \mathbb{C}$ są pierwiastkami równania charakterystycznego*

$$x^2 - \operatorname{tr}(A)x + \det(A) = 0,$$

*to zachodzą wzory Viéte'a*

$$\operatorname{tr}(A) = \lambda_1 + \lambda_2, \quad \det(A) = \lambda_1 \cdot \lambda_2.$$

*Dowód.* Z założenia mamy

$$(x - \lambda_1)(x - \lambda_2) = p_A(x) = x^2 - \operatorname{tr}(A)x + \det(A),$$

wystarczy więc porównać współczynniki wielomianów po obu stronach równości. □

U Dla dowolnych macierzy $X, Y \in M_{2\times 2}(\mathbb{C})$ macierze $XY$ i $YX$ mają takie same wielomiany charakterystyczne. Wynika to z faktu, że

$$\operatorname{tr} XY = \operatorname{tr} YX, \quad \det XY = \det YX.$$

**Lemat 11.1.** *Jeśli $A \in M_{2\times 2}(\mathbb{F})$ ma dwie różne rzeczywiste wartości własne $\lambda_1 \neq \lambda_2$ oraz $v_i \in \mathbb{F}^2$ są wektorami własnymi odpowiadającymi $\lambda_i \in \mathbb{F}$, to $v_1$, $v_2$ są liniowo niezależne. W szczególności $v_1, v_2$ tworzą bazę $\mathbb{F}^2$.*

*Dowód.* W przeciwnym razie $v_2 = p \cdot v_1$ dla pewnego skalara $p \in \mathbb{F}$. Wtedy

$$\begin{aligned} \lambda_2 \cdot v_2 &= A(v_2) \\ &= A(p \cdot v_1) \\ &= p \cdot A(v_1) \\ &= p\lambda_1 \cdot v_1 \\ &= \lambda_1 \cdot v_2, \end{aligned}$$

czyli $\lambda_1 = \lambda_2$, co prowadzi do sprzeczności. Skorzystaliśmy tu z faktu, że $v_2 \neq 0$. □

**Twierdzenie 11.3.** *Jeśli $A \in M_{2\times 2}(\mathbb{F})$ ma dwa liniowo niezależne wektory własne $v_1, v_2 \in \mathbb{F}^2$ odpowiadające wartościom własnym $\lambda_1$, $\lambda_2 \in \mathbb{F}$ (niekoniecznie różnym), to dla macierzy $B = \left[\begin{array}{c|c} v_1 & v_2 \end{array}\right]$ mamy*

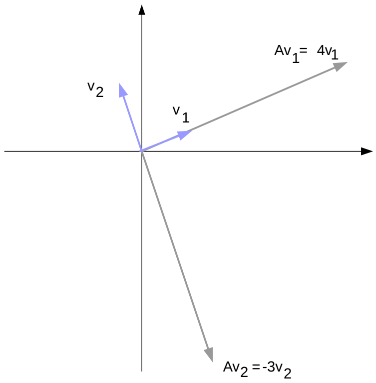
$$\Lambda := B^{-1}AB = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

*Dowód.* Z definicji $B$ mamy, że $B(e_i) = v_i$ dla $i = 1, 2$. Oczywiście $B$ jest odwracalna oraz dla $i = 1, 2$ mamy

$$\begin{aligned} B^{-1}ABe_i &= B^{-1}Av_i \\ &= B^{-1}(\lambda_i \cdot v_i) \\ &= \lambda_i \cdot e_i. \end{aligned}$$

□

U Twierdzenie 11.3 orzeka, że jeśli $\mathbb{F}^2$ ma bazę złożoną z wektorów własnych, to $A \in M_{2\times 2}(\mathbb{F})$ ma w tej bazie postać diagonalną $\Lambda$. Macierz $\Lambda$ nazywamy *postacią Jordana* macierzy $A$.




Rysunek 11.1: Działanie macierzy $A$ na wektorach własnych $v_1$, $v_2$ o wartościach własnych $\lambda_1 = 4$, $\lambda_2 = -3$.

**Przykład 11.1.** Znajdziemy postać Jordana macierzy

$$A = \begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix}.$$

Szukamy najpierw wartości własnych czyli pierwiastów równania charakterystycznego

$$p_A(x) = x^2 - x - 12 = 0.$$

Są nimi liczby $\lambda_1 = 4$ oraz $\lambda_2 = -3$. Aby znaleźć wektor własny dla $\lambda_1 = 4$ znajdujemy jądro macierzy

$$A - 4I = \begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix}.$$

Rozwiązując równanie $(A - 4I)x = 0$ otrzymujemy, że $x = [2x_2, x_2]^T$. Jednym z wektořow własnych jest $v_1 = [2, 1]^T$. Analogicznie, $v_2 = [-1, 3]^T$ jest wektorem własnym dla $\lambda_2$. W bazie $v_1$, $v_2$ macierz $A$ ma postać diagonalną

$$\Lambda = \begin{bmatrix} 4 & 0 \\ 0 & -3 \end{bmatrix}.$$

! Macierz zespolona $A \in M_{2\times 2}(\mathbb{C})$ może mieć rzeczywistą wartość własną, ale może nie mieć odpowiadającego jej rzeczywistego wektora własnego w $\mathbb{R}^2$. Na przykład rozważmy macierz

$$A = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}.$$

Wielomian charakterystyczny ma postać

$$p_A(x) = x^2 - 1,$$

więc $A$ ma wartości własne $\lambda_1 = 1$ oraz $\lambda_2 = -1$. Wektor własny $[x, y]^T \in \mathbb{C}^2$ dla $\lambda_1 = 1$ jest niezerowym rozwiązaniem równania

$$\begin{bmatrix} -1 & -i \\ i & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

czyli

$$-x - iy = 0, \quad ix - y = 0.$$

Stąd $y = ix$ oraz każdy wektor własny jest postaci

$$c \cdot [1, i]^T, \quad c \in \mathbb{C} \setminus \{0\}.$$

Nie może on być wektorem z $\mathbb{R}^2$.

Zajmiemy się teraz przypadkiem, gdy $\lambda \in \mathbb{F}$ jest *dwukrotnym pierwiastkiem* wielomianu charakterystycznego $p_A$. Mogą zachodzić dwie możliwości:

- $\dim \ker(A - \lambda I) = 2$. Mówimy wtedy, że *krotność geometryczna* wartości własnej $\lambda$ jest równa jej *krotności algebraicznej*. Wówczas $A$ ma dwa liniowo niezależne wektory własne odpowiadające $\lambda$ i jesteśmy w sytuacji opisanej twierdzeniem 11.3. Wtedy postać Jordana $\Lambda$ macierzy $A$ jest diagonalna:

$$\Lambda := B^{-1}AB = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \lambda I.$$

  Z powyższej równości wynika, że wtedy $A = \lambda I$.

- $\dim \ker(A - \lambda I) = 1$. Istnieje wtedy tylko jeden liniowo niezależny wektor własny $v \in \mathbb{F}^2$ dla $\lambda \in \mathbb{F}$.

**Wniosek 11.4.** *Jeśli $\lambda \in \mathbb{F}$ jest dwukrotnym pierwiastkiem charakterystycznym macierzy $A \in M_{2\times 2}(\mathbb{F})$ i $A \neq \lambda I$, to* $\dim\ker(A-\lambda I) = 1$ *i $A$ ma tylko jeden liniowo niezależny wektor własny.*

**Lemat 11.2** (twierdzenie Cayleya–Hamiltona)**.** *Dla macierzy $A \in M_{2\times 2}(\mathbb{F})$ zachodzi równość*

$$p_A(A) = A^2 - \operatorname{tr}(A)A + \det(A)I = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

*Dowód.* Zastosujemy brutalną siłę, czyli bezpośredni rachunek. Niech $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Ponieważ $p_A(x) = (x-a)(x-d) - bc$, więc

$$\begin{aligned} p_A(A) &= (A - aI)(A - dI) - bcI \\ &= \begin{bmatrix} 0 & b \\ c & d-a \end{bmatrix} \begin{bmatrix} a-d & b \\ c & 0 \end{bmatrix} - \begin{bmatrix} bc & 0 \\ 0 & bc \end{bmatrix} \\ &= \begin{bmatrix} bc & 0 \\ c(a-d) + (d-a)c & bc \end{bmatrix} - \begin{bmatrix} bc & 0 \\ 0 & bc \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

□

**Przykład 11.2.** Niech $A \in M_{2\times 2}(\mathbb{C})$. Twierdzenie Cayleya–Hamiltona pozwala łatwo wyznaczyć $A^n$ dla $n \in \mathbb{N}$. Niech $\lambda_1, \lambda_2 \in \mathbb{C}$ będą wartościami własnymi $A$. Wtedy

$$A^2 - (\lambda_1 + \lambda_2)A + \lambda_1\lambda_2 I = 0,$$

czyli

$$(A - \lambda_1 I)(A - \lambda_2 I) = 0.$$

Jeśli $\lambda_1 \neq \lambda_2$, to rozważamy macierze

$$X = \frac{A - \lambda_2 I}{\lambda_1 - \lambda_2}, \quad Y = \frac{A - \lambda_1 I}{\lambda_2 - \lambda_1}.$$

Wtedy

$$X^2 = X, \quad XY = YX = 0, \quad Y^2 = Y$$

środkowa równość wynika bezpośrednio z twierdzenia Cayleya–Hamiltona. Sprawdzimy równość $X^2 = X$. Mamy

$$
\begin{aligned}
X^2 &= \frac{(A-\lambda_2 I)^2}{(\lambda_1-\lambda_2)^2}\\
&= \frac{(A-\lambda_2 I)(A-\lambda_1 I+(\lambda_1-\lambda_2)I)}{(\lambda_1-\lambda_2)^2}\\
&= \frac{(A-\lambda_2 I)(A-\lambda_1 I)}{(\lambda_1-\lambda_2)^2}+\frac{(A-\lambda_2)(\lambda_1-\lambda_2)I}{(\lambda_1-\lambda_2)^2}\\
&= \frac{A-\lambda_2 I}{(\lambda_1-\lambda_2)}\\
&= X.
\end{aligned}
$$

Stąd dla $k \geqslant 2$ mamy

$$X^k = X, \quad Y^k = Y, \quad \lambda_1 \neq \lambda_2.$$

Dla $\lambda_1 = \lambda_2$ rozważamy

$$Z = A - \lambda_1 I.$$

Wtedy $Z^k = 0$ dla $k \geqslant 2$. Mamy więc

$$A^n = \begin{cases} (\lambda_1 X + \lambda_2 Y)^n = \lambda_1^n X + \lambda_2^n Y, & \text{gdy } \lambda_1 \neq \lambda_2 \\ (\lambda_1 I + Z)^n = \lambda_1^n I + n\lambda_1^{n-1} Z, & \text{gdy } \lambda_1 = \lambda_2 \end{cases}$$

**Wniosek 11.5.** *Jeśli $A \in M_{2\times 2}(\mathbb{C})$ jest nieosobliwa, to*

$$A^{-1} = \frac{1}{\det A}(\operatorname{tr} A \cdot I - A), \quad \operatorname{tr} A^{-1} = \frac{\operatorname{tr} A}{\det A}.$$

*Ponadto*

$$\det A = \frac{1}{2}((\operatorname{tr} A)^2 - \operatorname{tr} A^2).$$

*Dowód.* Z twierdzenia Cayleya–Hamiltona mamy

$$A^2 - \operatorname{tr}(A)A + \det(A)I = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Mnożąc powyższą równość przez $A^{-1}$ otrzymujemy, że

$$A^{-1} = \frac{1}{\det A}(\operatorname{tr} A \cdot I - A).$$

Biorąc ślady obu stron w tej równości dostajemy, że $\operatorname{tr} A^{-1} = \frac{\operatorname{tr} A}{\det A}$. Wzór na $\det A$ otrzymujemy, obliczając ślady macierzy po obu stronach w równości Cayleya–Hamiltona. □

**Wniosek 11.6.** *Niech $A \in M_{2\times 2}(\mathbb{C})$. Jeśli $A \notin \{\alpha I : \alpha \in \mathbb{C}\}$ oraz $A^2 - aA + bI =$ dla pewnych $a, b \in \mathbb{C}$, to*

$$a = \operatorname{tr} A, \quad b = \det A.$$

*Dowód.* Z twierdzenia Cayleya–Hamiltona wynika, że

$$(a - \operatorname{tr} A)A = (b - \det A)I.$$

Zauważmy, że $a - \operatorname{tr} A = 0$, bo inaczej $A = \frac{b-\det A}{a-\operatorname{tr} A} I$, co jest sprzeczne z założeniem. Stąd również $b - \det A = 0$. □

U

Macierz $A \in M_{2\times 2}(\mathbb{C})$ jest nilpotentna, jeśli $A^n = 0$ dla pewnego $n \in \mathbb{N}$. Pokażemy, że wtedy 0 jest jedyną wartością własną $A$ oraz $A^2 = 0$. Zauważmy, że jeśli $\lambda \in \mathbb{C}$ jest wartością własną dla wektora własnego $v \neq 0$, to $0 = A^n v = \lambda^n \cdot v$, czyli $\lambda = 0$.

W szczególności, $\operatorname{tr} A = 0$ i $\det A = 0$, czyli $p_A(x) = x^2$. Z twierdzenia Cayleya–Hamiltona wynika, że $A^2 = 0$.

**Lemat 11.3.** *Jeśli $\lambda \in \mathbb{F}$ jest dwukrotnym pierwiastkiem wielomianu charakterystycznego $p_A$, to*

$$(A - \lambda I)^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

*Dowód.* Ze wzorów Viéte'a mamy $\mathrm{tr}(A) = 2\lambda$ i $\det(A) = \lambda^2$. Z lematu Cayleya–Hamiltona wynika, że

$$A^2 - 2\lambda A + \lambda^2 I = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Ponieważ macierze $A$ i $\lambda I$ komutują ze sobą, więc

$$(A - \lambda I)^2 = A^2 - 2\lambda A + \lambda^2 I.$$

□

**Twierdzenie 11.4.** *Jeśli $\lambda \in \mathbb{F}$ jest dwukrotną wartością własną macierzy $A \in M_{2\times 2}(\mathbb{F})$ oraz $A \neq \lambda I$, to istnieje taka baza $v, w$ dla $\mathbb{F}^2$, że*

$$Av = \lambda \cdot v, \quad Aw = v + \lambda \cdot w$$

*W szczególności, dla macierzy $S = \left[\begin{array}{c|c} v & w \end{array}\right]$ mamy następującą postać Jordana dla $A$:*

$$\Lambda = S^{-1}AS = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$

*Dowód.* Niech $0 \neq v \in \ker(A - \lambda I)$ będzie dowolnym wektorem własnym dla $\lambda$. Każdy wektor własny $A$ jest postaci $p \cdot v$, gdzie $v$ jest pewnym wektorem własnym $A$ dla $\lambda$ oraz $0 \neq p \in \mathbb{F}$. Niech $0 \neq u \in \mathbb{F}^2$ będzie dowolnym wektorem liniowo niezależnym z $v$. Stąd $u$ nie jest wektorem własnym odpowiadającym $\lambda$, więc $(A - \lambda I)u \neq 0$. Ponieważ,

$$0 = (A - \lambda I)^2 u = (A - \lambda I)(A - \lambda I)u,$$

więc $(A - \lambda I)u \in \ker(A - \lambda I)$ jest wektorem własnym odpowiadającym $\lambda$. Stąd

$$(A - \lambda I)u = p \cdot v,$$

dla pewnego $0 \neq p \in \mathbb{R}$. Dla wektora $w = \frac{1}{p} \cdot u$ otrzymujemy, że $(A - \lambda I)w = v$, czyli $Aw = v + \lambda \cdot w$. □

> POSTAĆ JORDANA MACIERZY ZESPOLONEJ
> Macierz zespolona $A \in M_{2\times 2}(\mathbb{C})$ ma jedną z dwóch postaci Jordana
>
> $$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \qquad \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}.$$

**Wniosek 11.7.** *(i) Jeśli $0 \neq A \in M_{2\times 2}(\mathbb{C})$ jest macierzą nilpotentną, to*

$$A = S \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} S^{-1}$$

*dla pewnej macierzy nieosobliwej $S$.*

*(ii) Jeśli $A \in M_{2\times 2}(\mathbb{C})$ jest taka, że $A^2 = A$, to $A = 0$ lub $A = I$ lub*

$$A = S \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} S^{-1}$$

*dla pewnej macierzy nieosobliwej $S$.*

*(iii) Jeśli $A \in M_{2\times 2}(\mathbb{C})$ jest taka, że $A^2 = I$, to $A = \pm I$ lub*

$$A = S \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} S^{-1}$$

*dla pewnej macierzy nieosobliwej $S$.*

*(iv) Jeśli $A \in M_{2\times 2}(\mathbb{C})$ jest taka, że $A^2 = -I$, to $A = \pm i \cdot I$ lub*

$$A = S \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} S^{-1}$$

*dla pewnej macierzy nieosobliwej $S$.*

**Wniosek 11.8.** *Dla dowolnej macierzy $A \in M_{2\times 2}(\mathbb{C})$ mamy*

*(i) $A$ jest podobna do $A^T$.*

*(ii) $A$ jest podobna do macierzy symetrycznej.*

*(iii) $A$ jest iloczynem macierzy symetrycznych.*

*Dowód.* Punkt (i). Możemy założyć, że $A$ jest w postaci Jordana. Jeśli jest ona diagonalna, to nie ma czego dowodzić. Załóżmy więc, że

$$A = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}.$$

Wtedy dla $S = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ mamy $S^{-1}AS = A^T$.

Punkt (ii). Ponownie możemy założyć, że $A = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$. Szukamy takiej macierzy nieosobliwej $S = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{C})$, że $S^{-1}AS$ jest symetryczna. Dla $\Delta = \det S \neq 0$ mamy

$$S^{-1}AS = \begin{bmatrix} cd + \lambda\Delta & d^2 \\ -c^2 & -cd + \lambda\Delta \end{bmatrix}.$$

Ponieważ szukamy macierzy symetrycznej, więc musi zachodzić warunek $c^2+d^2 = 0$. Wystarczy więc przyjąć takie $a, b, c, d$, że $ad - bc \neq 0$ i $c^2 + d^2 = 0$. Zwróćmy uwagę, że $S$ nie może być rzeczywista. Przykładowo, dla $d = i = a$, $c = 1$ i $b = 0$ mamy

$$S = \begin{bmatrix} i & 0 \\ 1 & i \end{bmatrix}, \quad S^{-1}AS = \begin{bmatrix} \lambda - i & 1 \\ 1 & \lambda + i \end{bmatrix}.$$

Punkt (iii). Zauważmy najpierw, że możemy założyć, że $A$ jest w postaci Jordana $\Lambda$, bo jeśli $\Lambda = BC$ jest iloczynem macierzy symetrycznych, to

$$A = S\Lambda S^{-1} = (SBS^T)((S^{-1})^T CS^{-1})$$

jest również iloczynem macierzy symetrycznych. Dla $\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ wystarczy przyjąć

$$B = \begin{bmatrix} \lambda_1 & 0 \\ 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

Jeśli $\Lambda = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$, to przyjmujemy

$$B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & \lambda \\ \lambda & 1 - \lambda \end{bmatrix}.$$

□

U Rozważmy macierz $A$ w postaci bloku Jordana

$$A = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}.$$

Dla macierzy

$$S = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

mamy $S^{-1} = S$ oraz

$$\begin{aligned} S^{-1}AS &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \lambda & 0 \\ 1 & \lambda \end{bmatrix} \\ &= A^T. \end{aligned}$$

Naturalne jest pytanie o postać macierzy $A \in M_{2\times2}(\mathbb{R})$, gdy jej wielomian charakterystyczny $p_A$ ma dwie zespolone wartości własne $\lambda_1 = \alpha + i\beta$, $\lambda_2 = \alpha - i\beta$ oraz $\beta \neq 0$. Macierz $A \in M_{2\times2}(\mathbb{R})$ możemy traktować jako element przestrzeni $M_{2\times2}(\mathbb{C})$. Istnieje więc wektor własny $v \in \mathbb{C}^2$ odpowiadający wartości własnej $\lambda_1 = \alpha + i\beta$. Wektor $v \in \mathbb{C}^2$ wygodnie jest zapisać w postaci

$$v = \begin{bmatrix} a_1 + ia_2 \\ b_1 + ib_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} + i \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} := v_1 + iv_2, \quad v_1, v_2 \in \mathbb{R}^2.$$

Wówczas

$$Av = Av_1 + iAv_2.$$

Ponadto

$$A(v_1 - iv_2) = (\alpha - i\beta) \cdot (v_1 - iv_2),$$

czyli wektor $v_1 - iv_2$ jest wektorem własnym odpowiadającym wartości własnej $\lambda_2 = \alpha - i\beta$.

Pokażemy, że wektory $v_1, v_2 \in \mathbb{R}^2$ są liniowo niezależne. Rzeczywiście, przypuśćmy, że $v_1 = p \cdot v_2$ dla pewnej liczby rzeczywistej $p \in \mathbb{R}$. Wtedy

$$A(v_1 + iv_2) = (\alpha + i\beta) \cdot (v_1 + iv_2) = (\alpha + i\beta)(p + i) \cdot v_2$$

oraz

$$A(v_1 + iv_2) = A((p + i)v_2) = (p + i) \cdot A(v_2),$$

więc $A(v_2) = (\alpha + i\beta) \cdot v_2$, co prowadzi do sprzeczności z warunkiem $\beta \neq 0$.

Zauważmy, że

$$A(v_1) + iA(v_2) = A(v_1 + iv_2) = (\alpha + i\beta) \cdot (v_1 + iv_2),$$

więc

$$A(v_1) = \alpha \cdot v_1 - \beta \cdot v_2, \quad A(v_2) = \beta \cdot v_1 + \alpha \cdot v_2.$$

Dla macierzy $S = \left[\begin{array}{c|c} v_1 & v_2 \end{array}\right]$ otrzymujemy równość

$$\Lambda := S^{-1}AS = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}.$$

Dla $S_1 = \left[\begin{array}{c|c} v_2 & v_1 \end{array}\right]$ mamy

$$S_1^{-1}AS_1 = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}.$$

**Wniosek 11.9** (Rzeczywista postać Jordana)**.** *Jeśli $\lambda_1 = \alpha + i\beta$, $\lambda_2 = \alpha - i\beta$ z $\beta \neq 0$ są pierwiastkami wielomianu charakterystycznego $p_A$ dla macierzy $A \in M_{2\times2}(\mathbb{R})$, to istnieje taka nieosobliwa macierz $S \in M_{2\times2}(\mathbb{R})$, że*

$$\Lambda = S^{-1}AS = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}.$$

*Ponadto* $S = \left[\begin{array}{c|c} v_1 & v_2 \end{array}\right]$ *dla dowolnego takiego niezerowego wektora* $v_1 + iv_2 \in \mathbb{C}$*, że*

$$A(v_1 + iv_2) = (\alpha + i\beta) \cdot (v_1 + iv_2).$$

**Przykład 11.3** (Ciąg Fibonacciego)**.** Ciąg Fibonacciego $u_n$ jest zdefiniowany rekurencyjnie wzorem

$$u_{n+1} = u_n + u_{n-1}, \quad u_1 = 1, u_0 = 0.$$

Dla macierzy

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

mamy więc

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_n \\ u_{n-1} \end{bmatrix} = \begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix}.$$

Wynika stąd, że

$$\begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix} = A^n \begin{bmatrix} u_1 \\ u_0 \end{bmatrix} = A^n e_1.$$

Wielomian charakterystyczny $A$ ma postać $p_A(x) = x^2 - x - 1$, więc wartościami własnymi $A$ są

$$\lambda_+ = \frac{\sqrt{5}+1}{2}, \quad \lambda_- = \frac{1-\sqrt{5}}{2}.$$

Wektory własne mają postać

$$v_+ = [(1+\sqrt{5})/2, 1]^T, \quad v_- = [(1-\sqrt{5})/2, 1]^T.$$

Ponieważ

$$e_1 = \frac{1}{\sqrt{5}} \cdot (v_+ - v_-),$$

więc

$$\begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix} = \frac{1}{\sqrt{5}}(\lambda_+^n \cdot v_+ - \lambda_-^n \cdot v_-),$$

czyli otrzymujemy *formułę Bineta*

$$u_n = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2}\right)^n.$$

**Przykład 11.4.** Rozważmy macierz

$$A = \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix}.$$

Jej wielomian charakterystyczny ma postać

$$p_A(x) = x^2 - 4x + 4 = (x-2)^2.$$

Stąd $\lambda = 2$ jest dwukrotną wartością własną macierzy $A$. Ponieważ

$$A - 2I = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix},$$

więc $[x, y]^T \in \ker(A - 2I)$ jeśli $x + y = 0$, czyli

$$\ker(A - 2I) = \operatorname{span}\left\{[1, -1]^T\right\}.$$

Wektor $v = [1, -1]^T$ jest wektorem własnym i każdy inny wektor własny jest liniowo zależny z $v$.

Szukamy teraz takiego wektora $w = [x, y]^T \in \mathbb{R}^2$, że

$$(A - 2I)w = v = [1, -1]^T.$$

Musimy więc rozwiązać układ równań

$$x + y = 1, \quad -x - y = -1.$$

Rozwiązaniami są wektory postaci $[x, 1 - x]^T = x[1, -1]^T + [0, 1]^T$. Możemy więc przyjąć, że

$$w = [1, 0]^T = e_1.$$

Dla macierzy

$$S = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix},$$

mamy

$$S^{-1}AS = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

**Przykład 11.5.** Macierz

$$A = \begin{bmatrix} 0 & -4 \\ 1 & 0 \end{bmatrix}$$

ma zespolone wartości własne $\lambda_1 = 2i$ oraz $\lambda_2 = -2i$. Znajdziemy wektor własny $v_1 = [a + bi, c + di]^T \in \mathbb{C}^2$ dla $\lambda_1 = 2i$. Szukamy niezerowych rozwiązań równania

$$\begin{aligned} 0 = (A - 2iI)v &= \begin{bmatrix} -2i & -4 \\ 1 & -2i \end{bmatrix} \begin{bmatrix} a + bi \\ c + di \end{bmatrix} \\ &= \begin{bmatrix} -2ia + 2b - 4c - 4di \\ a + bi - 2ic + 2d \end{bmatrix} \\ &= \begin{bmatrix} -2ia + 2(b - 2c) - 2i(a + 2d) \\ a + 2d + i(b - 2c) \end{bmatrix}. \end{aligned}$$

Stąd $a + 2d = 0$ oraz $b - 2c = 0$. Możemy więc przyjąć, że

$$v_1 = [2, -i]^T = [2, 0]^T + i[0, -1]^T.$$

Dla wartości własnej $\lambda_2 = -2i$ wektorem własnym jest więc $v_2 = [2, i]^T$. Dla macierzy

$$S = \begin{bmatrix} 2 & 2 \\ -i & i \end{bmatrix}$$

mamy

$$S^{-1}AS = \frac{1}{4i} \begin{bmatrix} i & -2 \\ i & 2 \end{bmatrix} \begin{bmatrix} 0 & -4 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -i & i \end{bmatrix} = \begin{bmatrix} 2i & 0 \\ 0 & -2i \end{bmatrix}$$

Aby znaleźć rzeczywistą postać Jordana rozważamy macierz

$$P = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$$

której kolumnami są część rzeczywista i urojona wektora $v_1 = [2, -i]^T = [2, 0]^T + i[0, -1]^T$. Wtedy

$$P^{-1}AP = \begin{bmatrix} 1/2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & -4 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}.$$

**Twierdzenie 11.5.** *Niech $A \in M_{2\times 2}(\mathbb{R})$ będzie nieosobliwa. Wtedy istnieje odwzorowanie*

$$\alpha : [0,1] \ni t \longmapsto A_t \in M_{2\times 2}(\mathbb{R})$$

*o własnościach:*

*(1)* $A_0 = A$,

*(2)*

$$A_1 = \begin{cases} I, & \text{gdy } \det(A) > 0, \\ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, & \text{gdy } \det(A) < 0. \end{cases}$$

*(3)* $A_t$ *jest nieosobliwa dla każdego* $t \in [0,1]$ *oraz* $\det A_t = \det A$,

*(4)* $\alpha$ *jest ciągła.*

*Dowód.* Zauważmy najpierw, że wystarczy rozważyć postać Jordana macierzy $A$, bo jeśli $\alpha$ jest szukanym odwzorowaniem dla $A$, to

$$\tilde{\alpha}(t) = S^{-1}\alpha(t)S = S^{-1}A_tS,$$

jest nim dla $S^{-1}AS$.

Zajmiemy się najpierw przypadkiem, gdy $\det A > 0$. Będziemy stosować oznaczenie $A_0 \simeq A_1$ jeśli istnieje takie szukane odwzorowanie, że $\alpha(0) = A_0$ i $\alpha(1) = A_1$. Zwróćmy uwagę, że jeśli $A_0 \simeq A_1$ i $A_1 \simeq A_2$, to $A_0 \simeq A_2$.

**Krok. 1** Rozważmy macierz

$$A = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}, \quad b \neq 0.$$

Wtedy przyjmując, że

$$A_t = \begin{bmatrix} t & b \\ -b & t \end{bmatrix}, \quad B_t = \begin{bmatrix} 1 & (1-t)b \\ -(1-t)b & 1 \end{bmatrix},$$

otrzymujemy, że

$$A_0 = A \simeq A_1 = B \simeq B_1 = I.$$

Podkreślmy, że $\det A_t > 0$ i $\det B_t > 0$.

**Krok. 2** Zakładamy, że $A$ ma postać Jordana

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}, \quad b \neq 0,$$

i definiujemy

$$A_t = \begin{bmatrix} ta & b \\ -b & ta \end{bmatrix}, \quad b \neq 0,$$

co sprowadza nas do kroku 1.

**Krok. 3** Jeśli postać Jordana jest diagonalna

$$A = \begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix}, \quad ac > 0,$$

to rozważamy

$$A_t = \begin{bmatrix} (1-t)a & t \\ -t & (1-t)c \end{bmatrix},$$

i kończymy w kroku 1.

**Krok. 4** Jeśli $A$ ma trójkątną postać Jordana

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

to przyjmując

$$A_t = \begin{bmatrix} a & t \\ 0 & a \end{bmatrix}$$

lądujemy ponownie w kroku 3.

Pozostaje nam rozważyć przypadek $\det A < 0$. Wtedy wartości własne $A$ są rzeczywiste i $A$ ma postać Jordana

$$A = \begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix}, \quad c < 0 < a.$$

Definiujemy wówczas

$$A_t = \begin{bmatrix} ta+1 & 0 \\ 0 & tc-1 \end{bmatrix}.$$

Podkreślmy, że $\det A_t < 0$. □

## 11.2 Rozkład Jordana w wymiarze 3

Podamy dowód twierdzenia Jordana dla macierzy zespolonej $A \in M_{3\times 3}(\mathbb{C})$. Zastosujemy podejście, które uogólnia się na wyższe wymiary.

**Twierdzenie 11.6** (Rozkład Jordana macierzy zespolonej)**.** *Niech $A \in M_{3\times 3}(\mathbb{C})$. Istnieje taka macierz nieosobliwa $S \in M_{3\times 3}(\mathbb{C})$, że $S^{-1}AS$ jest jednej z poniższych postaci:*

$$\begin{bmatrix} \boxed{\lambda_1} & 0 & 0 \\ 0 & \boxed{\lambda_2} & 0 \\ 0 & 0 & \boxed{\lambda_3} \end{bmatrix}, \quad \left[\begin{array}{c|cc} \lambda_1 & 0 & 0 \\ \hline 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{array}\right], \quad \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{bmatrix}.$$

*Ponadto $\lambda_i \in \mathbb{C}$ są wartościami własnymi macierzy $A$.*

**Lemat 11.4.** *Załóżmy, że wartości własne $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$ macierzy $A \in M_{3\times 3}(\mathbb{C})$ są różne. Wtedy odpowiadające im wektory własne $v_1, v_2, v_3$ są liniowo niezależne oraz dla $S = \left[\begin{array}{c|c|c} v_1 & v_2 & v_3 \end{array}\right]$ mamy*

$$S^{-1}AS = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3).$$

*Dowód.* Zauważmy najpierw, że $v_1$ i $v_2$ są liniowo niezależne. Przypuśćmy, że $v_2 = x \cdot v_1$ dla pewnego $x \in \mathbb{C}$. Wtedy $x \neq 0$, bo $v_2 \neq 0$. Mamy

$$\begin{aligned} \lambda_2 \cdot v_2 &= Av_2 \\ &= A(x \cdot v_1) \\ &= x\lambda_1 \cdot v_1 \\ &= \lambda_1 \cdot v_2, \end{aligned}$$

czyli $\lambda_1 = \lambda_2$, sprzeczność.

Przypuśćmy, że $v_3 \in \operatorname{span}\{v_1, v_2\}$. Wtedy $v_3 = x_1 \cdot v_1 + x_2 \cdot v_2$ dla pewnych $x_1, x_2 \in \mathbb{C}$. Przynajmniej jedna z liczb $x_1$, $x_2$ jest różna od zera, bo $v_3 \neq 0$. Otrzymujemy, że

$$\begin{aligned} \lambda_3 \cdot v_3 &= Av_3 \\ &= A(x_1 \cdot v_1 + x_2 \cdot v_2) \\ &= x_1\lambda_1 \cdot v_1 + x_2\lambda_2 \cdot v_2, \end{aligned}$$

czyli

$$\lambda_3(x_1 \cdot v_1 + x_2 \cdot v_2) = x_1\lambda_1 \cdot v_1 + x_2\lambda_2 \cdot v_2.$$

W efekcie

$$x_1(\lambda_3 - \lambda_1) \cdot v_1 + x_2(\lambda_2 - \lambda_1) \cdot v_2 = 0,$$

co jest sprzeczne z liniową niezależnością $v_1$ i $v_2$. □

**Lemat 11.5.** *Niech $A \in M_{3\times 3}(\mathbb{C})$. Istnieje taka macierz nieosobliwa $S \in M_{3\times 3}(\mathbb{C})$, że macierz $S^{-1}AS$ jest górnie trójkątna.*

*Dowód.* Niech $\lambda \in \mathbb{C}$ będzie wartością własną $A$ i niech $v \in \mathbb{C}^3$ będzie odpowiadającym jej wektorem własnym. Uzupełniamy $v$ do bazy $v$, $u$, $w$ dla $\mathbb{C}^3$. Rozważmy macierz $W = \left[\begin{array}{c|c|c} v & u & w \end{array}\right]$. Macierz $W^{-1}AW$ ma postać blokową

$$W^{-1}AW = \left[\begin{array}{c|cc} \lambda & * & * \\ \hline 0 & a & b \\ 0 & c & d \end{array}\right].$$

Z rozkładu Jordana w wymiarze 2 istnieje taka macierz nieosobliwa $V \in M_{2\times 2}(\mathbb{C})$, że

$$V^{-1}\begin{bmatrix} a & b \\ c & d \end{bmatrix} V = T$$

jest górnie trójkątna. Definiujemy macierz nieosobliwą

$$U = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & V \end{array}\right] \in M_{3\times 3}(\mathbb{C}).$$

Wtedy

$$\begin{aligned} U^{-1}W^{-1}AWU &= U^{-1}\left[\begin{array}{c|cc} \lambda & * & * \\ \hline 0 & a & b \\ 0 & c & d \end{array}\right] U \\ &= \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & V^{-1} \end{array}\right] \left[\begin{array}{c|cc} \lambda & * & * \\ \hline 0 & a & b \\ 0 & c & d \end{array}\right] \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & V \end{array}\right] \\ &= \left[\begin{array}{c|c} \lambda & 0 \\ \hline 0 & T \end{array}\right], \end{aligned}$$

jest górnie trójkątna i teza zachodzi dla $S = WU$. □

**Lemat 11.6.** *Rozważmy macierz górnie trójkątną z $0$ jako jedyną wartością własną*

$$A = \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix}.$$

*Wtedy zachodzi jeden z warunków*

*(i)* $A = 0$,

*(ii)* $\dim \ker A = 2$ *i $A$ jest podobna do macierzy*

$$\left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right],$$

*(iii)* $\dim \ker A = 1$ *oraz $A$ jest podobna do macierzy*

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

*Dowód.* Sprawdzamy, że

$$A^2 = \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & ac \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$A^3 = \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & ac \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 0.$$

Mamy trzy możliwości

(i) $A = 0$,

(ii) $A \neq 0$, $A^2 = 0$,

(iii) $A^2 \neq 0$, $A^3 = 0$.

W przypadku (ii) $a = 0$ lub $c = 0$, bo $A^2 = 0$ i macierz $A$ jest jednej z postaci

$$A = \begin{bmatrix} 0 & 0 & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix} \quad c \neq 0, \quad A = \begin{bmatrix} 0 & a & b \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad a \neq 0, \quad A = \begin{bmatrix} 0 & 0 & b \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad b \neq 0.$$

We wszystkich przypadkach $\operatorname{rank} A = \dim \operatorname{im} A = 1$ oraz

$$\operatorname{im} A \subset \ker A, \quad \dim \ker A = 2.$$

Inkluzja $\operatorname{im} A \subset \ker A$ wynika stąd, że $A^2 = 0$. Niech $v_1 \in \operatorname{im} A$ będzie niezerowym wektorem. Istnieje wię taki (niezerowy) wektor $v_2 \in \mathbb{C}^3$, że $Av_2 = v_1$. Wybieramy teraz wektor $v_3 \in \ker A$ liniowo niezależny z $v_1$. Rozważmy macierz

$$S = \left[ \begin{array}{c|c|c} v_1 & v_2 & v_3 \end{array} \right].$$

Ponieważ $v_2 \notin \ker A$, więc $v_1, v_2, v_3$ są liniowo niezależne, czyli $S$ jest nieosobliwa. Ponadto

$$S^{-1}AS = \left[ \begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array} \right]$$

W przypadku (iii) mamy $a \neq 0$ i $c \neq 0$, więc

$$\operatorname{rank} A = \dim \mathrm{imA} \; = 2.$$

Ponadto

$$\ker A \subset \operatorname{im} A, \quad \dim \ker A = 1, \quad \operatorname{im} A^2 = \ker A.$$

Niech $v \in \ker A$ będzie wektorem własnym dla wartości własnej 0. Ponieważ

$$\operatorname{im} A^2 = \ker A,$$

więc istnieje taki wektor $u \in \mathbb{C}^3$, że $A^2 u = v$. Wektory

$$v, \quad Au, \quad u$$

są liniowo niezależne. Rzeczywiście, przypuśćmy, że $x \cdot v + y \cdot Au + z \cdot u = 0$ dla pewnych $x, y, z \in \mathbb{C}$. Wtedy

$$0 = A^2(x \cdot v + y \cdot Au + z \cdot u) = z \cdot v,$$

czyli $z = 0$, bo $v \neq 0$. W konsekwencji

$$0 = A(x \cdot v + y \cdot Au) = y \cdot v,$$

czyli $y = 0$. Z równości $x \cdot v = 0$ również $x = 0$. Dla

$$S = \left[\begin{array}{c|c|c} v & Au & u \end{array}\right]$$

mamy

$$S^{-1}AS = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

□

**Wniosek 11.10.** *Załóżmy, że $\lambda \in \mathbb{C}$ jest 3-krotną wartością własną macierzy $A \in M_{3\times 3}(\mathbb{C})$. Wtedy*

*(i) jeśli* $\dim\ker(A - \lambda I) = 3$, *to* $A = \lambda I$,

*(ii) jeśli* $\dim\ker(A - \lambda I) = 2$, *to $A$ jest podobna do macierzy*

$$\left[\begin{array}{cc|c} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ \hline 0 & 0 & \lambda \end{array}\right].$$

*(iii) jeśli* $\dim\ker(A - \lambda I) = 1$, *to $A$ jest podobna do macierzy*

$$\begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}.$$

*Dowód.* Z lematu 11.5 możemy założyć, że $A$ jest górnie trójkątna. Stosujemy lemat 11.6 do macierzy $A - \lambda I$. □

**Lemat 11.7.** *Załóżmy, że $A \in M_{3\times 3}(\mathbb{C})$ jest macierzą górnie trójkątną postaci*

$$A = \begin{bmatrix} \lambda & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix}, \quad \lambda \neq 0.$$

*Wtedy*

*(i)* $\dim\ker A < 3$,

*(ii) jeśli* $\dim\ker A = 2$, *to* $S^{-1}AS = \begin{bmatrix} \boxed{\lambda} & 0 & 0 \\ 0 & \boxed{0} & 0 \\ 0 & 0 & \boxed{0} \end{bmatrix}$ *dla pewnej macierzy nieosobliwej $S \in M_{3\times 3}(\mathbb{C})$,*

*(iii) jeśli* $\dim\ker A = 1$, *to* $S^{-1}AS = \left[\begin{array}{c|cc} \lambda & 0 & 0 \\ \hline 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right]$ *dla pewnej macierzy nieosobliwej $S \in M_{3\times 3}(\mathbb{C})$.*

*Dowód.* Zauważmy, że

$$\dim \operatorname{im} A = \begin{cases} 1, & \text{gdy } c = 0, \\ 2, & \text{gdy } c \neq 0, \end{cases}$$

czyli

$$\dim\ker A = \begin{cases} 2, & \text{gdy } c = 0, \\ 1, & \text{gdy } c \neq 0, \end{cases}$$

Załóżmy, że $\dim \ker A = 2$. Niech $v_2, v_3 \in \ker A$ będą bazą dla jądra. Jeśli $v_1$ jest wektorem własnym dla $\lambda$, to $v_1, v_2, v_3$ są liniowo niezależne. Przypuśćmy bowiem, że $v_1 = x_2 \cdot v_2 + x_3 \cdot v_3$. Wtedy

$$\lambda \cdot v_1 = Av_1 = x_2 \cdot Av_2 + x_3 \cdot Av_3 = 0,$$

sprzeczność. Dla macierzy nieosobliwej $S = \left[\begin{array}{c|c|c} v_1 & v_2 & v_3 \end{array}\right]$ mamy

$$S^{-1}AS = \begin{bmatrix} \boxed{\lambda} & 0 & 0 \\ 0 & \boxed{0} & 0 \\ 0 & 0 & \boxed{0} \end{bmatrix}.$$

Niech teraz $\dim \ker A = 1$, czyli $c \neq 0$. Zauważmy, że teza zajdzie, gdy znajdziemy taką bazę $v_1, v_2, v_3$, że

$$Av_1 = \lambda \cdot v_1, \quad Av_2 = 0, \quad Av_3 = v_2.$$

Z wyborem $v_1$ i $v_2$ nie mamy problemu: bierzemy wektory własne dla $\lambda$ i $0$. Musimy dobrać wektor $v_3$, aby $Av_3 = v_2$. W tym celu wystarczy zauważyć, że

$$\ker A \subset \operatorname{im} A.$$

Rzeczywiście, $A[x, y, z]^T = [\lambda x + ay + bz, cz, 0]^T$, czyli $[x, y, z]^T \in \ker A$, gdy $z = 0$ i $x = \frac{a}{\lambda} y$, czyli jądro jest generowane przez wektor $[a/\lambda, 1, 0]^T$. Sprawdzamy łatwo, że jest on w obrazie $A$ generowanym przez wektory $[\lambda, 0, 0]^T$ i $[b, c, 0]^T$.

Pozostaje sprawdzić, że $v_1, v_2, v_3$ są liniowo niezależne. Wiemy, że $v_1$ i $v_2$ są liniowo niezależne. Przypuśćmy, że $v_3 = x_1 \cdot v_1 + x_2 \cdot v_2$. Wtedy

$$\begin{aligned} v_2 &= Av_3 \\ &= x_1 \cdot Av_1 + x_2 \cdot Av_2 \\ &= x_1 \lambda \cdot v_1, \end{aligned}$$

co jest sprzeczne z liniową niezależnością $v_1$ i $v_2$. □

**Wniosek 11.11.** *Załóżmy, że $\lambda, \lambda_1 \in \mathbb{C}$ są różnymi wartościami własnymi macierzy $A \in M_{3\times 3}(\mathbb{C})$ oraz $\lambda_1$ ma krotność 2. Wtedy*

*(i)* $\dim \ker(A - \lambda_1 I) < 3$,

*(ii) jeśli* $\dim \ker(A - \lambda_1 I) = 2$, *to $A$ jest podobna do macierzy*

$$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{bmatrix}.$$

*(iii) jeśli* $\dim \ker(A - \lambda_1 I) = 1$, *to $A$ jest podobna do macierzy*

$$\left[\begin{array}{c|cc} \lambda & 0 & 0 \\ \hline 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{array}\right].$$

*Dowód.* Z lematu 11.5 możemy założyć, że $A$ jest górnie trójkątna. Stosujemy lemat 11.7 do macierzy $A - \lambda_1 I$. □

Zajmiemy się teraz macierzami rzeczywistymi $A \in M_{3\times 3}(\mathbb{R})$. Wtedy wielomian charakterystyczny $p_A$ jest stopnia 3 i ma współczynniki rzeczywiste. Jeśli liczba zespolona $\lambda \in \mathbb{C}$ jest pierwiastkiem $p_A$, to jest nim również $\overline{\lambda}$. Wynika stąd, że istnieje co najmniej jeden pierwiastek rzeczywisty $p_A$. Jest on oczywiście wartością własną $p_A$. Jeśli $A$ ma trzy (liczone z krotnościami) rzeczywiste wartości własne, to $A$ ma taką samą postać Jordana jak w przypadku zespolonym.

Pozostaje nam rozważyć przypadek, gdy $p_A$ ma rzeczywisty pierwiastek $\lambda_1 \in \mathbb{R}$ i parę pierwiastków $\lambda_2 = \alpha + i\beta$ oraz $\overline{\lambda_2} = \alpha - i\beta$ z $\beta \neq 0$.

**Wniosek 11.12** (Rzeczywista postać Jordana)**.** *Jeśli dla* $A \in M_{3\times 3}(\mathbb{R})$ *ma rzeczywisty pierwiastek* $\lambda_1 = \lambda \in \mathbb{R}$ *i parę pierwiastków sprzężonych* $\lambda_2 = \alpha + i\beta$ *oraz* $\overline{\lambda_2} = \alpha - i\beta$ *z* $\beta \neq 0$*, to istnieje taka nieosobliwa macierz* $S \in M_{3\times 3}(\mathbb{R})$*, że*

$$S^{-1}AS = \left[\begin{array}{c|cc} \lambda & 0 & 0 \\ \hline 0 & \alpha & \beta \\ 0 & -\beta & \alpha \end{array}\right].$$

*Dowód.* Niech $v \in \mathbb{R}^3$ będzie wektorem własnym dla $\lambda$, a $w = w_1 + iw_2$ zespolonym wektorem własnym dla $\alpha + i\beta$. Wtedy $\overline{w} = w_1 - iw_2$ jest wektorem własnym dla $\alpha - i\beta$. Z naszych rozważań w wymiarze 2 wystarczy pokazać, że $v, w_1, w_2$ są liniowo niezależne nad $\mathbb{R}$, bo wtedy teza zachodzi dla $S = \left[\begin{array}{c|c|c} v & w_1 & w_2 \end{array}\right]$. Wiemy już, że $w_1$ i $w_2$ są liniowo niezależne. Przypuśćmy, że

$$v = x_1 \cdot w_1 + x_2 \cdot w_2,$$

dla pewnych $x_1, x_2 \in \mathbb{R}$. Ponieważ

$$w_1 = \frac{1}{2} \cdot (w + \overline{w}), \quad w_2 = \frac{1}{2i} \cdot (w - \overline{w}),$$

więc

$$v = \frac{x_1 - ix_2}{2} \cdot w + \frac{x_1 + ix_2}{2} \cdot \overline{w},$$

czyli $v, w, \overline{w}$ są liniowo zależne nad $\mathbb{C}$. Prowadzi to do sprzeczności, bo wektory te odpowiadają różnym wartościom własnym. □

**Przykład 11.6.** Macierz

$$A = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 3 & -2 \\ 0 & 1 & 1 \end{bmatrix}$$

ma wartości własne $\lambda_1 = -3$, $\lambda_2 = 2 + i$, $\lambda_3 = \overline{\lambda_2} = 2 - i$. Odpowiadającymi im wektorami własnymi są przykładowo

$$v_1 = [1, 0, 0]^T = e_1, \quad v_2 = [0, 1 + i, 1]^T, \quad v_3 = [0, 1 - i, 1]^T.$$

Dla

$$S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1+i & 1-i \\ 0 & 1 & 1 \end{bmatrix}$$

mamy

$$S^{-1}AS = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 2+i & 0 \\ 0 & & 2-i \end{bmatrix}.$$

Aby otrzymać postać rzeczywistą rozważamy macierz

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

której druga i trzecia kolumna, to odpowiednio, część rzeczywista i urojona wektora własnego $v_2 = [0, 1, 1]^T + i[0, 1, 0]^T$. Otrzymujemy, że

$$P^{-1}AP = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 2 \end{bmatrix}.$$

**Przykład 11.7.** Wartości własne macierzy

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 1 & 1 & 2 \end{bmatrix}$$

to $\lambda_1 = 1$ i $\lambda_2 = \lambda_3 = 2$. Sprawdzamy łatwo, że odpowiadające im wektory własne, to

$$v_1 = [1, 1, -2]^T, \quad v_2 = [0, 0, 1]^T.$$

Krotność algebraiczna wartości własnej $\lambda_2 = 2$ jest równa 2, ale jej krotność geometryczna jest równa 1. Macierz $A$ nie ma więc bazy wektorów własnych, czyli nie jest diagonalizowalna. Aby wyznaczyć postać Jordana $A$, szukamy takiego wektora $w$, że

$$(A - 2I)w = v_2.$$

Ponieważ

$$A - 2I = \begin{bmatrix} -1 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix},$$

więc możemy przyjąć, że $w = [0, 1, 0]^T$. Dla macierzy

$$S = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix}$$

mamy

$$S^{-1}AS = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

**Przykład 11.8.** Macierz

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 2 \end{bmatrix}$$

ma tylko jedną wartość własną $\lambda = 2$ o krotności algebraicznej 3. Zauważmy, że

$$A - 2I = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix},$$

więc

$$\ker(A - 2I) = \operatorname{span}\{e_1, e_3\}.$$

Ponadto

$$\operatorname{im}(A - 2I) = \operatorname{span}\{[1, 0, -1]^T\} \subset \ker(A - 2I).$$

Szukamy teraz takiego wektora $w$, że

$$(A - 2I)w = [1, 0, -1]^T.$$

łatwo sprawdzamy, że możemy przyjąć $w = [0, 1, 0]^T = e_2$. Dla bazy

$$v_1 = [1, 0, -1]^T, \quad v_2 = [0, 1, 0]^T, \quad v_3 = [1, 0, 0]^T$$

Dla

$$S = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

mamy

$$S^{-1}AS = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

**Przykład 11.9.** Przyglądniemy się teraz macierzom antysymetrycznym $A \in M_{3\times 3}(\mathbb{R})$. Ponieważ $A^T = -A$, więc $A$ ma postać

$$A = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}, \quad a, b, c \in \mathbb{R}.$$

Zakładamy, że $A$ jest niezerowa, czyli $a^2+b^2+c^2 \neq 0$. Ponieważ $A$ jest antysymetryczna i wymiar $n = 3$ jest nieparzysty, więc $\det A = 0$. W szczególności, 0 jest zawsze wartością własną. Ma ona krotność 1, bo jak łatwo sprawdzić pozostałe wartości własne, to $\pm i\sqrt{a^2 + b^2 + c^2}$. Wektor własny dla wartości własnej 0 jest równy

$$v_A = [-c, b, -a]^T.$$

Ponadto jeśli $A, B \in M_{3\times 3}(\mathbb{R})$ są antysymetryczne, to

$$\operatorname{tr}(AB) = -2(v_A|v_B).$$

Co więcej, macierz

$$[A, B] := AB - BA$$

jest antysymetryczna oraz

$$v_{[A,B]} = v_A \times v_B.$$

## 11.3 Diagonalizacja w wymiarach 2 i 3

Niech

$$A = \begin{bmatrix} a & c \\ c & b \end{bmatrix}$$

będzie rzeczywistą macierzą symetryczną. Równanie charakterystyczne dla $A$ ma postać

$$p_A(x) = x^2 - (a + b)x + ab - c^2 = 0.$$

Ponieważ

$$\Delta := (a + b)^2 - 4(ab - c^2) = (a - b)^2 + 4c^2 \geqslant 0,$$

więc wartości własne $\lambda_1$, $\lambda_2$ macierzy $A$ są rzeczywiste oraz

$$\lambda_1 = \frac{a + b + \sqrt{\Delta}}{2}, \quad \lambda_2 = \frac{a + b - \sqrt{\Delta}}{2}.$$

**Twierdzenie 11.7.** *Dla macierzy rzeczywistej $A \in M_{2\times 2}(\mathbb{R})$ następujące warunki są równoważne*

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

*(ii)* $\mathbb{R}^2$ *ma bazę ortonormalną złożoną z wektorów własnych* $A$,

*(iii) Istnieje taka macierz ortogonalna* $Q \in M_{2\times 2}(\mathbb{R})$, *że*

$$Q^T AQ = D,$$

*gdzie* $D = \operatorname{diag}(\lambda_1, \lambda_2)$ *jest diagonalna z wartościami własnymi* $A$ *na przekątnej.*

*Dowód.* $(i) \Rightarrow (ii)$

Załóżmy, że $A$ jest symetryczna. Jeśli $c = 0$, to $e_1$, $e_2$ jest bazą ortonormalną złożoną z wektorów własnych macierzy $A = \operatorname{diag}(a, b)$.

Jeśli $c \neq 0$, to $\Delta > 0$ i $A$ ma dwie różne rzeczywiste wartości własne $\lambda_1$, $\lambda_2$. Niech $v_1$, $v_2$ będą odpowiadającymi im wektorami własnymi o normie 1. Pokażemy, że wektory $v_1$, $v_2$ są ortogonalne. Mamy

$$\begin{aligned}\lambda_1(v_1|v_2) &= (\lambda_1 v_1|v_2)\\ &= (Av_1|v_2)\\ &= (v_1|Av_2)\\ &= (v_1|\lambda_2 v_2)\\ &= \lambda_2(v_1|v_2),\end{aligned}$$

więc $(v_1|v_2) = 0$, bo $\lambda_1 \neq \lambda_2$.

*(ii)* $\Rightarrow$ *(iii)* Wystarczy przyjąć $Q = \left[\begin{array}{c|c} v_1 & v_2 \end{array}\right]$ dla bazy ortonormalnej $v_1$, $v_2$ złożonej z wektorów własnych.

*(iii)* $\Rightarrow$ *(i)*

Wiemy, że $A = QDQ^T$. Wtedy

$$\begin{aligned}A^T &= (Q^T)^T D^T Q^T\\ &= QDQ^T\\ &= A.\end{aligned}$$

□

Rozważmy teraz macierz symetryczną $A \in M_{3\times 3}(\mathbb{R})$. Istnieje $\lambda_1 \in \mathbb{R}$ rzeczywista wartość własna dla $A$. Niech $v_1$ będzie odpowiadającym jej wektorem własnym o normie 1. Niech $V \subset \mathbb{R}^3$ będzie płaszczyzną prostopadłą do wektora $v_1$. Wybierzmy bazę ortonormalną $v_2$, $v_3 \in V$ dla $V$. Wtedy $v_1$, $v_2$, $v_3$ jest bazą ortonormalną dla $\mathbb{R}^3$.

Pokażemy, że $Av_2$, $Av_3 \in V$, czyli $A(V) \subset V$. Ze względu na symetrię wystarczy pokazać, że $Av_2 \in V$. Ponieważ $A$ jest symetryczna, więc

$$\begin{aligned}(v_1|Av_2) &= (Av_1|v_2)\\ &= (\lambda_1 v_1|v_2)\\ &= \lambda_1(v_1|v_2)\\ &= 0.\end{aligned}$$

Rozważmy macierz ortogonalną $W = \left[\begin{array}{c|c|c} v_1 & v_2 & v_3 \end{array}\right]$. Wtedy

$$W^{-1}AW = W^T AW = \left[\begin{array}{c|cc} \lambda_1 & 0 & 0 \\ \hline 0 & a & c \\ 0 & d & b \end{array}\right].$$

Ponieważ $W^T AW$ jest symetryczna, więc $d = c$, czyli

$$W^T AW = \left[\begin{array}{c|cc} \lambda_1 & 0 & 0 \\ \hline 0 & a & c \\ 0 & c & b \end{array}\right]$$

Wiemy już, że istnieje taka macierz ortogonalna $B \in M_{2\times 2}(\mathbb{R})$, że

$$B^T \begin{bmatrix} a & c \\ c & b \end{bmatrix} B = \operatorname{diag}(\lambda_2, \lambda_3).$$

Dla macierzy ortonogonalnej

$$U = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & B \end{array}\right] \in M_{3\times 3}(\mathbb{R})$$

mamy

$$U^T W^T AWU = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3).$$

Macierz $Q = WU$ jest ortogonalna i jej kolumny są bazą ortonormalną złożoną z wektorów własnych.

**Wniosek 11.13.** *Dla macierzy rzeczywistej $A \in M_{3\times 3}(\mathbb{R})$ następujące warunki są równoważne*

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

*(ii)* $\mathbb{R}^3$ *ma bazę ortonormalną złożoną z wektorów własnych* $A$,

*(iii)* *Istnieje taka macierz ortogonalna* $Q \in M_{3\times 3}(\mathbb{R})$, *że*

$$Q^T A Q = D,$$

*gdzie* $D = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$ *jest diagonalna z wartościami własnymi* $A$ *na przekątnej.*

## 11.4 Twierdzenie Cayleya–Hamiltona w wymiarze 3

**Twierdzenie 11.8** (twierdzenie Cayleya–Hamiltona)**.** *Niech* $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$ *będą wartościami własnymi macierzy* $A \in M_{3\times 3}(\mathbb{C})$. *Wtedy*

$$(A - \lambda_1 I)(A - \lambda_2 I)(A - \lambda_3 I) = 0.$$

*Dowód.* Z lematu 11.5 istnieje taka macierz nieosobliwa $S \in M_{3\times 3}(\mathbb{C})$, że $S^{-1}AS$ jest górnie trójkątna. Zauważmy, że

$$\begin{aligned}(S^{-1}AS - \lambda_1 I)(S^{-1}AS - \lambda_2 I)(S^{-1}AS - \lambda_3 I) = &\\ &(S^{-1}AS - \lambda_1 S^{-1}S)(S^{-1}AS - \lambda_2 S^{-1}S)(S^{-1}AS - \lambda_3 S^{-1}S) =\\ &S^{-1}(A - \lambda_1 I)SS^{-1}(A - \lambda_2 I)SS^{-1}(A - \lambda_3 I)S =\\ &S^{-1}(A - \lambda_1 I)(A - \lambda_2 I)(A - \lambda_3 I)S.\end{aligned}$$

Wynika stąd, że teza zachodzi dla $A$ wtedy i tylko wtedy, gdy zachodzi dla $S^{-1}AS$. Możemy więc założyć, że $A$ jest górnie trójkątna postaci

$$A = \begin{bmatrix} \lambda_1 & a & b \\ 0 & \lambda_2 & c \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

Niech $v = [x, y, x]^T \in \mathbb{C}^3$ będzie dowolnym wektorem. Zauważmy, że

$$(A - \lambda_3 I)v = [x_1, y_1, 0]^T,$$

dla pewnych $x_1, y_1 \in \mathbb{R}$. Następnie,

$$(A - \lambda_2 I)[x_1, y_1, 0]^T = [x_2, 0, 0]^T,$$

dla pewnego $x_2 \in \mathbb{C}$. Ostatecznie,

$$(A - \lambda_1 I)[x_2, 0, 0]^T = [0, 0, 0]^T,$$

czyli

$$(A - \lambda_1 I)(A - \lambda_2 I)(A - \lambda_3 I)v = 0.$$

□

U Z pewnością zauważyliscie, że dokładnie te same argumenty pokazują, że dla macierzy górnie trójkątnej $A \in M_{n\times n}(\mathbb{C})$ zachodzi

$$p_A(A) = 0.$$

Dowód opiera się na obserwacji, że dla macierzy górnie trójkątnej $A$ podprzestrzenie $\operatorname{span}\{e_1, \ldots, e_k\}$ są niezmiennicze dla $A$. Pokażemy później, że każda macierz $A \in M_{n\times n}(\mathbb{C})$ jest podobna do macierzy górnie trójkątnej, więc twierdzenie Cayleya–Hamiltona zachodzi dla dowolnej macierzy $A \in M_{n\times n}(\mathbb{C})$.

TEST $\longrightarrow$ Niech $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{R})$. Oceń prawdziwość poniższych zdań:

- Jeśli $a + b = c + d = 1$, to $[1, 1]^T$ jest wektorem własnym $A$.
- Jeśli $A$ ma prostą niezmienniczą, to $A$ ma rzeczywiste wartości własne.
- $A$ ma rzeczywisty wektor własny.
- Jeśli $A$ jest górnie trójkątna tzn. $c = 0$, to $A$ jest diagonalizowalna.
- Jeśli $\det A < 0$, to $A$ jest diagonalizowalna.
- Jeśli $\det A < 0$, to $A$ jest ortogonalnie diagonalizowalna.
- Jeśli $A^3 = 0$, to $A^2 = 0$.
- Jeśli $\mathbb{R}^2$ ma bazę złożoną z wektorów własnych $A$, to $A$ jest nieosobliwa.
- 0 jest jedyną wartością własną $A$ wtedy i tylko wtedy, gdy $A^2 = 0$.
- Jeśli $\operatorname{tr} A = 0$, to $A$ jest diagonalizowalna lub $A^2 = 0$.
- Jeśli $\det A = 0$, to $A$ ma prostą niezmienniczą.
- $A$ i $A^2$ mają takie same wartości własne.
- $A$ i $A^2$ mają takie same wektore wartości własne.
- Jeśli $A$ jest nieosobliwa, to $A$ i $A^{-1}$ mają takie same wartości własne.
- Jeśli $A$ jest nieosobliwa, to $A$ i $A^{-1}$ mają takie same wektory własne.
- $A$ i $A^T$ mają takie same wartości własne.
- $A$ i $A^T$ mają takie same wektory własne.
- Jeśli $S \in M_{2\times 2}(\mathbb{R})$ jest nieosobliwa, to $A$ i $S^{-1}AS$ mają takie same wartości własne.
- Jeśli $S \in M_{2\times 2}(\mathbb{R})$ jest nieosobliwa, to $A$ i $S^{-1}AS$ mają takie same wektory własne.
- Jeśli $A$ jest nieosobliwa i diagonalizowalna, to $A^{-1}$ jest diagonalizowalna.
- Jeśli macierz kwadratowa $A$ jest diagonalizowalna, to $A^T$ jest również diagonalizowalna.
- Jeśli macierz kwadratowa ma dwa identyczne wiersze (kolumny), to 0 jest jej wartością własną.

$\longleftarrow$

# Rozdział 12

# Wektory i wartości własne

Słowa kluczowe rozdziału
wartości własne macierzy podobnych $\lozenge$ wektory własne różnych wartości własnych $\lozenge$ baza wektorów własnych $\lozenge$ diagonalizacja $\lozenge$ postać Jordana macierzy nilpotentnej $\lozenge$ zespolona postać Jordana $\lozenge$ rzeczywista postać Jordana $\lozenge$ spektrum $\lozenge$ promień spektralny $\lozenge$ formuła Gelfanda $\lozenge$ twierdzenie Cayleya–Hamiltona $\lozenge$ postać Frobeniusa $\lozenge$ exponenta macierzy

## 12.1 Co jest możliwe w wyższych wymiarach?

Rozważmy macierz $A \in M_{n\times n}(\mathbb{F})$, gdzie $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Wektory i wartości własne macierzy $A$ są zdefiniowane analogicznie jak w przypadku $n = 2$. Skalar $\lambda \in \mathbb{F}$ jest wartością własną $A$ wtedy i tylko wtedy, gdy istnieje taki niezerowy wektor $v \in \mathbb{F}^n$, że $Av = \lambda \cdot v$. Jest to równoważne z faktem, że $\ker(A - \lambda I) \neq \{0\}$, więc $\det(A - \lambda I) = 0$. Wartość własna $\lambda \in \mathbb{F}$ jest pierwiastkiem wielomianu charakterystycznego (o współczynnikach w $\mathbb{F}$)

$$p_A(x) = \det(A - xI) = 0.$$

Z zasadniczego twierdzenia algebry wielomian $p_A$ ma w $\mathbb{C}$ dokładnie $n$ (liczonych z krotnościami) pierwiastków zespolonych $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$. Dla $\mathbb{F} = \mathbb{C}$ są one wartościami własnymi macierzy $A$. Jeśli $A \in M_{n\times n}(\mathbb{R})$, to $p_A(\lambda) = 0 \Leftrightarrow p_A(\overline{\lambda}) = 0$. Macierz rzeczywista $A \in M_{n\times n}(\mathbb{R})$ może nie mieć rzeczywistych wartości własnych.

**Wniosek 12.1** (ślad, wyznacznik i wartości własne)**.** *Jeśli* $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ *są pierwiastkami wielomianu charakterystycznego* $p_A$*, to*

$$\lambda_1 + \ldots + \lambda_n = \mathrm{tr}(A), \quad \lambda_1 \cdot \ldots \cdot \lambda_n = \det(A).$$

*Dowód.* Z zasadniczego twierdzenia algebry wynika, że

$$\det(A - xI) = (-1)^n (x - \lambda_1) \ldots (x - \lambda_n).$$

Teza wynika z porównania współczynników wielomianów po obu stronach powyższej równości. $\square$

**Wniosek 12.2.** $0$ *jest wartością własną macierzy* $A$ *wtedy i tylko wtedy, gdy* $\det(A) = 0$.

**Wniosek 12.3.** *Jeśli macierze* $A, B \in M_{n\times n}(\mathbb{F})$ *są podobne, to* $p_A = p_B$*. W szczególności,* $A$ *i* $B$ *mają te same wartości własne.*

*Dowód.* Z założenia istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{F})$, że $B = S^{-1}AS$. Stąd

$$\begin{aligned} p_B(x) &= \det(B - xI) \\ &= \det(S^{-1}AS - xI) \\ &= \det(S^{-1}(A - xI)S) \\ &= \det(A - xI) = p_A(x). \end{aligned}$$

$\square$

**Przykład 12.1.** Jeśli $A = [a_{ij}] \in M_{n\times n}(\mathbb{F})$ jest macierzą (górnie/dolnie) trójkątną, to

$$p_A(x) = \det(A - xI) = (-1)^n (x - a_{11}) \dots (x - a_{nn}),$$

czyli wartości własne są elementami na przekątnej macierzy $A$.

**Wniosek 12.4.** *Jeśli $A \in M_{n\times n}(\mathbb{F})$ jest nieosobliwa oraz $\lambda \in \mathbb{F}$ jest wartością własną $A$, to $\frac{1}{\lambda}$ jest wartością własną $A^{-1}$. Ponadto $A$ i $A^{-1}$ mają takie same wektory własne.*

*Dowód.* Jeśli $\lambda$ jest wartością własną $A$, to $\lambda \neq 0$, bo $A$ jest nieosobliwa. Mamy

$$\begin{aligned} 0 &= \det(A - \lambda I) \\ &= \det(\lambda A(1/\lambda I - A^{-1})) = \\ &= (-1)^n \det(\lambda A) \det(A^{-1} - 1/\lambda I), \end{aligned}$$

czyli $\det(A^{-1} - 1/\lambda I) = 0$.

Niech $v$ będzie wektorem własnym dla $\lambda$, czyli $Av = \lambda \cdot v$. Ponieważ $A$ jest nieosobliwa, więc $\frac{1}{\lambda} \cdot v = A^{-1} v$. □

**Wniosek 12.5.** *Jeśli $\lambda$ jest wartością własną $A$ i $v$ jest wektorem własnym, to $\lambda^m$ jest wartością własną $A^m$ i $v$ jest jej wektorem własnym.*

*Dowód.* Rzeczywiście,

$$\begin{aligned} A^m v &= A^{m-1} A v \\ &= A^{m-1}(\lambda \cdot v) \\ &= \lambda \cdot A^{m-1} v \\ &\vdots \\ &= \lambda^m \cdot v. \end{aligned}$$

□

**Wniosek 12.6.** *Jeśli $A$ jest nilpotentna, tzn. $A^m = 0$ dla pewnego $m \in \mathbb{N}$, to $0$ jest jedyną wartością własną $A$.*

**Przykład 12.2.** Jeśli $A^2 = A$ i $\lambda$ jest wartością własną, to $\lambda \in \{0, 1\}$. Rzeczywiście, jeśli $\lambda$ jest wartością własną oraz $v$ jest odpowiadającym wektorem własnym, to

$$\begin{aligned} \lambda \cdot v &= Av \\ &= A^2 v \\ &= \lambda^2 \cdot v. \end{aligned}$$

Ponieważ $v \neq 0$, więc $\lambda = \lambda^2$.

Zajmiemy się najpierw pewnymi specjalnymi przypadkami.

**Twierdzenie 12.1** (Różne wartości własne)**.** *Jeśli $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ są różnymi wartościami własnymi macierzy $A \in M_{n\times n}(\mathbb{F})$ z odpowiadającymi wektorami własnymi $v_1, \dots, v_k$, to $v_1, \dots, v_k$ są liniowo niezależne.*

*Dowód.* Niech

$$r = \dim \operatorname{span}\{v_1, \dots, v_k\}.$$

Przypuśćmy, że $r < k$. Możemy założyć, że $v_1, \dots, v_r$ są liniowo niezależne. Ponieważ $v_1, \dots, v_r, v_{r+1}$ są liniowo zależne, więc istnieją takie, nie wszyskie równe zero, skalary $x_i \in \mathbb{F}$, że

$$x_1 \cdot v_1 + \ldots + x_r \cdot v_r + x_{r+1} \cdot v_{r+1} = 0.$$

Zauważmy, że $x_{r+1} \neq 0$, bo w przeciwnym razie $v_1, \ldots, v_r$ byłyby liniowo zależne. Wtedy

$$\begin{aligned} 0 &= x_1 \cdot Av_1 + \ldots + x_r \cdot Av_r + x_{r+1} \cdot Av_{r+1} \\ &= x_1\lambda_1 \cdot v_1 + \ldots + x_r\lambda_r \cdot v_r + x_{r+1}\lambda_{r+1} \cdot v_{r+1}. \end{aligned}$$

Z powyższych równości wynika, że

$$x_1(\lambda_1 - \lambda_{r+1}) \cdot v_1 + \ldots + x_r(\lambda_r - \lambda_{r+1}) \cdot v_r = 0,$$

co jest sprzeczne z liniową niezależnością $v_1, \ldots, v_r$. □

**Definicja 12.1.** **Macierz diagonalizowalna**
Macierz $A \in M_{n\times n}(\mathbb{F})$ nazywamy *diagonalizowalną*, jeśli istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{F})$, że macierz $D = S^{-1}AS$ jest diagonalna. Mówimy wtedy, że $S$ *diagonalizuje* $A$.

**Twierdzenie 12.2** (Diagonalizacja=baza wektorów własnych)**.** *Macierz kwadratowa $A \in M_{n\times n}(\mathbb{F})$ jest diagonalizowalna wtedy i tylko wtedy, gdy $A$ ma $n$ liniowo niezależnych wektorów własnych.*

*Dowód.* Przypuśćmy, że $v_1, \ldots, v_n$ są liniowo niezależnymi wektorami własnymi z wartościami własnymi $\lambda_i$. Niech $S = \left[\begin{array}{c|c|c} v_1 & \ldots & v_n \end{array}\right]$. Wtedy

$$\begin{aligned} AS &= \left[\begin{array}{c|c|c} Av_1 & \ldots & Av_n \end{array}\right] \\ &= \left[\begin{array}{c|c|c} \lambda_1 \cdot v_1 & \ldots & \lambda_n \cdot v_n \end{array}\right] \\ &= S\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n). \end{aligned}$$

Ponieważ $S$ jest nieosobliwa, więc $S^{-1}AS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

Przypuśćmy teraz, że $A$ jest diagonalizowalna, czyli

$$AS = S\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$$

dla pewnej macierzy nieosobliwej $S = \left[\begin{array}{c|c|c} v_1 & \ldots & v_n \end{array}\right]$. Ponieważ $Se_i = v_i$, więc

$$Av_i = \lambda_i \cdot v_i, \quad i = 1, \ldots, n,$$

co kończy dowód. □

U

Diagonalizowalność macierzy rzeczywistej zależy od tego czy traktujemy ją jako macierz rzeczywistą, czy zespoloną. Przykładowo, macierz

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

nie jest diagonalizowalna jako macierz rzeczywista, bo nie ma rzeczywistych wartości i wektorów własnych. Jest ona diagonalizowalna jako macierz zespolona, bo ma dwie różne wartości własne $i$ oraz $-i$, więc dopowiadające im wektory własne są liniowo niezależne. Jej postać diagonalna jako macierzy zespolonej, to

$$\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

## 12.2 Idea dowodu twierdzenia Jordana

### 12.2.1 Pewne ogólne obserwacje o iteracjach macierzy

Niech $A \in M_{n\times n}(\mathbb{F})$. Rozważmy ciąg podprzestrzeni

$$\ker A \subset \ker A^2 \subset \ker A^3 \subset \ldots$$

Ponieważ $\dim \ker A^k \leqslant n$, więc istnieje najmniejsze takie $N \geqslant 1$, że

$$\ker A^N = \ker A^{N+1}.$$

**Lemat 12.1.** *Dla dowolnego $l \geqslant 1$ mamy*

$$\ker A^{N+l} = \ker A^{N+l-1}.$$

*Dowód.* Zastosujemy indukcję względem $l$. Dla $l = 1$ teza zachodzi z definicji $N$. Załóżmy, że $\ker A^{N+l} = \ker A^{N+l-1}$. Pokażemy, że $\ker A^{N+l+1} = \ker A^{N+l}$. Ponieważ $\ker A^{N+l} \subset \ker A^{N+l+1}$, więc wystarczy pokazać, że $\ker A^{N+l+1} \subset \ker A^{N+l}$. Niech $x \in \ker A^{N+l+1}$. Ponieważ

$$0 = A^{N+l+1}x = A^{N+l}(Ax),$$

więc $Ax \in \ker A^{N+l} = \ker A^{N+l-1}$. Stąd,

$$0 = A^{N+l-1}(Ax) = A^{N+l}x,$$

czyli $x \in \ker A^{N+l}$. □

**Twierdzenie 12.3.** *Jeśli $N$ jest najmniejszą liczbą naturalną, dla której $\ker A^k = \ker A^{k+1}$, to*

$$\mathbb{F}^n = \ker A^N \oplus \operatorname{im} A^N.$$

*Dowód.* Z formuły wymiaru wynika, że

$$n = \dim \ker A^N + \dim \operatorname{im} A^N,$$

więc wystarczy sprawdzić, że

$$\ker A^N \cap \operatorname{im} A^N = \{0\}.$$

Przypuśćmy, że $0 \neq x \in \ker A^N \cap \operatorname{im} A^N$. Wtedy $A^N x = 0$ oraz $x = A^N w$ dla pewnego $w \in \mathbb{F}$. Stąd $0 = A^N(x) = A^{2N}w$, czyli $w \in \ker A^{2N} = \ker A^N$. W efekcie, $x = A^N w = 0$. □

**Wniosek 12.7.** *Dla macierzy $A \in M_{n\times n}(\mathbb{F})$ następujące warunki są równoważne*

*(i)* $\ker A = \ker A^2$,

*(ii)* $\mathbb{F}^n = \ker A \oplus \operatorname{im} A$.

*Dowód.* Z twierdzenia 12.7 wynika, że (i) implikuje (ii). Załóżmy, że zachodzi warunek (ii). Przypuśćmy, że nie zachodzi warunek (i). Wtedy istnieje taki $v \in \ker A^2$, że $v \notin \ker A$. Wówczas $Av \neq 0$ i $A^2 v = 0$, czyli

$$0 \neq Av \in \ker A \oplus \operatorname{im} A = \{0\},$$

co prowadzi do sprzeczności. □

**Lemat 12.2.** *Podprzestrzenie $\ker A^N$ oraz $\operatorname{im} A^N$ są niezmiennicze dla $A$.*

*Dowód.* Teza wynika z równości $AA^N = A^N A$. Jeśli $x \in \ker A^N$, to

$$\begin{aligned} A^N Ax &= AA^N x \\ &= A0 \\ &= 0, \end{aligned}$$

czyli $Ax \in \ker A^N$.

Z drugiej strony, $AA^N x = A^N Ax \in \operatorname{im} A^N$. □

Przypomnijmy, że macierz $B \in M_{k\times k}(\mathbb{F})$ jest *nilpotentna*, gdy $B^m = 0$ dla pewnego $m \in \mathbb{N}$.

**Wniosek 12.8.** *Załóżmy, że $A \in M_{n\times n}(\mathbb{C})$. Niech $v_1, \ldots, v_k$ będzie bazą dla $\ker A^N$ oraz $v_{k+1}, \ldots, v_n$ bazą dla $\operatorname{im} A^N$. Macierz $A$ ma w tej bazie postać blokową*

$$\left[\begin{array}{c|c} B & 0 \\ \hline 0 & C \end{array}\right],$$

*gdzie $B$ jest macierzą restrykcji $A$ do $\ker A^N$ w bazie $v_1, \ldots, v_k$, a $C$ macierzą restrykcji $A$ do $\operatorname{im} A^N$ w bazie $v_{k+1}, \ldots, v_n$. Ponadto* *(i)* *macierz $B$ jest nilpotentna,*

*(ii)* $0$ *jest jedyną wartością własną $B$ oraz* $0$ *nie jest wartością własną $C$,*

*(iii)* $k = \dim\ker A^N$ *jest algebraiczną krotnością $n_0$ wartości własnej* $0$*, tzn. jako pierwiastka wielomianu charakterystycznego $p_A$.*

*Dowód.* Ponieważ $A^N$ zeruje się na $\ker A^N$, więc $B^N = 0$, czyli $B$ jest nilpotentna. Jedyną wartością własną macierzy nilpotentnej jest 0, bo jeśli $\lambda$ jest wartością własną $B$, to $\lambda^N$ jest wartością własną $B^N = 0$.

Pokażemy, że 0 nie jest wartością własną $C$. Przypuśćmy, że $Cx = 0$ dla pewnego $0 \neq x \in \operatorname{im} A^N$. Wtedy $x = A^N w$ dla pewnego $w \in \mathbb{C}^n$. Stąd

$$\begin{aligned} 0 = Cx &= CA^N w \\ &= AA^N w \\ &= A^N A w \\ &= A^{N+1} w, \end{aligned}$$

czyli $w \in \ker A^{N+1} = \ker A^N$. Stąd $x = A^N w = 0$, sprzeczność.

Punkt $(iii)$ wynika z równości $p_A = p_B \cdot p_C$ i punktu $(ii)$. □

U

Jeśli $B \in M_{k\times k}(\mathbb{C})$ jest nilpotentna, to

$$B^k = 0.$$

Rzeczywiście, ponieważ 0 jest jedyną wartością własną $B$, więc $p_B(x) = (-1)^k x^k$, więc z twierdzenia Cayleya–Hamiltona wynika, że $p_B(B) = (-1)^k B^k$ jest macierzą zerową.

### 12.2.2 Postać Jordana macierzy nilpotentnej

Zacznijmy od przykładu modelowego macierzy nilpotentnej.

**Przykład 12.3.** Rozważmy macierz

$$B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Mamy

$$B^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B^4 = 0.$$

Zauważmy, że

$$Be_4 = e_3, \quad Be_3 = e_2, \quad Be_2 = e_1, \quad Be_1 = 0.$$

Oznacza to, że

$$e_1 = B^3 e_4, \quad e_2 = B^2 e_4, \quad e_3 = Be_4, \quad e_4 = B^0 e_4.$$

**Definicja 12.2.** **Nilpotentny blok Jordana**

Macierz

$$N = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 1 \\ 0 & \dots & 0 & 0 & 0 \end{bmatrix} \in M_{m\times m}(\mathbb{R})$$

nazywamy *nilpotentnym blokiem Jordana* rozmiaru $m \geqslant 1$. Rozmiar $m$ jest najmniejszą taką liczbą, że $N^m = 0$. Macierz blokową

$$\operatorname{diag}(N_1, \ldots, N_s)$$

nazywamy *nilpotentną macierzą Jordana*, jeśli $N_i$ są nilpotentnymi blokami Jordana.

U Załóżmy, że macierz $B \in M_{k\times k}(\mathbb{C})$ ma w pewnej bazie postać nilpotentnej macierzy Jordana $B = \operatorname{diag}(N_1, \ldots, N_s) \in M_{k\times k}(\mathbb{R})$ oraz $n_i$ jest rozmiarem $N_i$, to

(1) 0 jest jedyną wartością własną $B$,

(2) istnieją takie wektory $v_1, \ldots, v_s$, że

$$B^{n_1-1}v_1, \ldots, Bv_1, v_1, \quad \ldots \quad B^{n_s-1}v_s, \ldots, Bv_s, v_s$$

tworzą bazę $\mathbb{C}^n$,

(3) bazę dla $\ker B$ tworzą wektory

$$B^{n_1-1}v_1, \ldots, B^{n_s-1}v_s,$$

czyli $\dim\ker B = s$ jest liczbą nilpotentnych bloków Jordana.

(4) $\max\{n_1, \ldots, n_s\}$ jest najmniejszą liczbą naturalną $n$ o własności $B^n = 0$,

(5) krotność algebraiczna wartości własnej 0 jest równa $k = \dim\ker B^{\max\{n_1,\ldots,n_s\}}$.

**Lemat 12.3.** *Dla macierzy $A, B \in M_{k\times k}(\mathbb{C})$ zachodzi równość*

$$\operatorname{rank}(AB) = \operatorname{rank}(B) - \dim \ker(A) \cap \operatorname{im}(B).$$

*Dowód.* Niech $l = \operatorname{rank}(AB)$. Można założyć, że $l \geqslant 1$, bo w przeciwnym razie $AB = 0$, czyli $\operatorname{im}(B) \subset \ker(B)$ i teza zachodzi.

Niech $u_1, \ldots, u_l$ będzie bazą dla $\operatorname{im}(AB)$. Istnieją takie wektory $v_i$, że $ABv_i = u_i$ dla $i = 1, \ldots, l$. Wektory $Bv_1, \ldots, Bv_l$ są liniowo niezależne. Rzeczywiście, jeśli $x_1 \cdot Bv_1 + \ldots + x_l \cdot Bv_l = 0$, to również $x_1 \cdot ABv_1 + \ldots + x_l \cdot ABv_l = 0$, więc $x_1 = \ldots = x_l = 0$. Niech $v_{l+1}, \ldots, v_{l+s}$ będą bazą dla takie, że $Bv_1, \ldots, Bv_{l+s}$ są bazą dla $\ker(A) \cap \operatorname{im}(B)$. Wystarczy pokazać, że $Bv_{l+1}, \ldots, Bv_{l+s}$ tworzą bazę dla $\operatorname{im}(B)$. Niech $Bv \in \operatorname{im}(B)$. Wtedy $ABv = x_1 \cdot ABv_1 + \ldots + x_l \cdot ABv_l$ dla pewnych skalarów $x_i$, więc

$$A(Bv - x_1 \cdot Bv_1 - \ldots - x_l \cdot Bv_l) = 0,$$

czyli

$$Bv - x_1 \cdot Bv_1 - \ldots - x_l \cdot Bv_l \in \ker(A) \cap \operatorname{im}(B).$$

Stąd $Bv \in \operatorname{span}\{Bv_1, \ldots, Bv_{l+s}\}$, więc

$$\operatorname{im}(B) = \operatorname{span}\{Bv_1, \ldots, Bv_{l+s}\}.$$

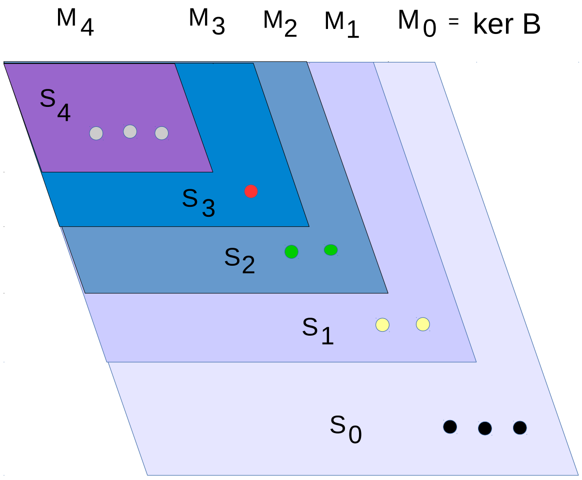
Wystarczy pokazać, że $Bv_1, \ldots, Bv_{l+s}$ są liniowo niezależne. Przypuśćmy, że $x_1 \cdot Bv_1 + \ldots x_{l+s}Bv_{l+s} = 0$. Wtedy

$$0 = x_1 \cdot ABv_1 + \ldots x_{l+s}ABv_{l+s} = x_1 \cdot ABv_1 + \ldots x_l ABv_l,$$

bo $Bv_{s+1}, \ldots, Bv_{l+s} \in \ker(A)$. W efekcie $x_1 = \ldots = x_l = 0$, bo $ABv_1, \ldots, ABv_l$ są liniowo niezależne. W konsekwencji również $x_{l+1} = \ldots = x_{l+s} = 0$. □

**Lemat 12.4.** *Niech $B \in M_{k\times k}(\mathbb{C})$ będzie macierzą nilpotentną i niech $1 \leqslant N$ będzie najmniejszą taką liczbą naturalną, że $B^N = 0$. Istnieją takie wektory $u_1, \ldots, u_s$ oraz takie liczby naturalne $1 \leqslant a_1, \ldots, a_s \leqslant N$, że*

*(i)* $B^{a_i}(u_i) = 0$ *dla* $i = 1, \ldots, s$,




Rysunek 12.1: $S_0 \cup S_1 \cup S_2 \cup S_3 \cup S_4$ tworzy bazę dla $\ker B$ pewnej macierzy nilpotentnej, dla której $B^5 = 0$.

*(ii)* *wektory*

$$\underbrace{B^{a_1-1}u_1}_{\in \ker B} \ldots,\ Bu_1,\ u_1, \quad \ldots, \quad \underbrace{B^{a_s-1}u_s}_{\in \ker B} \ldots,\ Bu_s,\ u_s.$$

*tworzą bazę* $\mathbb{C}^k$*,*

*(iii)* $s = \dim \ker B$*,*

*(iv)* $N = \max\{a_1, \ldots, a_s\} \leqslant k$*.*

*Dowód.* Dla ustalenia uwagi załóżmy, że $N = 5$. Dowód przypadku ogólnego z dokładnością do szczegółów technicznych jest identyczny.

Zaczniemy od konstrukcji pewnej specjalnej bazy dla $\ker B$. W tym celu rozważamy zstępujący ciąg podprzestrzeni

$$M_i = \ker B \cap \operatorname{im} B^i, \quad i = 0, 1, 2, 3, 4, 5,$$

czyli

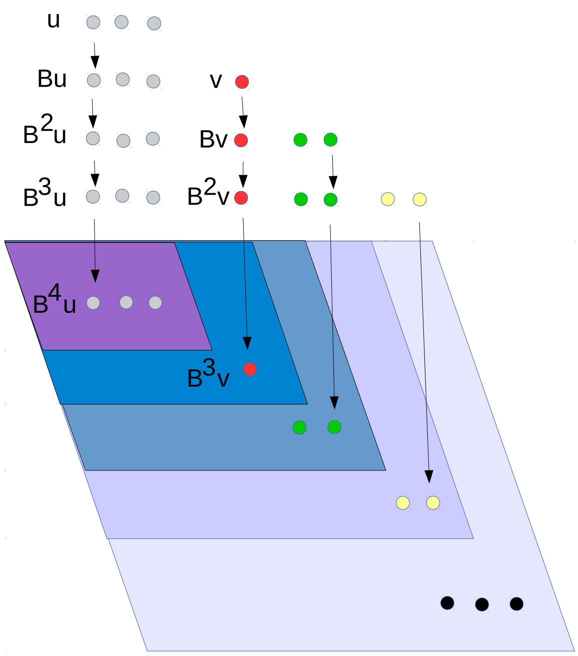
$$\{0\} = M_5 \subset M_4 \subset M_3 \subset M_2 \subset M_1 \subset M_0 = \ker B.$$

Zajmiemy się przykładową sytuacją przedstawioną na rys. 12.1 Zaczynamy od wyboru bazy $S_4$ dla $M_4 = \ker B \cap \operatorname{im} B^i$. Na rys. 12.1 $S_4$ składa się z 3 wektorów zaznaczonych kolorem szarym. Następnie dobieramy wektory $S_3 \subset M_3$, aby suma $S_4 \cup S_3$ tworzyła bazę dla $M_3$.

W zbiorze $S_3$ są takie liniowo niezależne wektory z $\ker B \cap \operatorname{im} B^3$, które nie są w $M_4$, więc nie są one w obrazie $B^4$. Mamy jeden taki wektor zaznaczony kolorem czerwonym. Kontynuujemy ten proces i otrzymujemy bazę $S_0 \cup S_1 \cup S_2 \cup S_3 \cup S_4$ dla $\ker B$.

Składa się ona z 11 wektorów:

- 3 wektory czarne: nie należą one do $\operatorname{im} B$,
- 2 wektory żółte: należą do $\operatorname{im} B$, ale nie należą do $\operatorname{im} B^2$,
- 2 wektory zielone: należą do $\operatorname{im} B^2$, ale nie należą do $\operatorname{im} B^3$,
- 1 wektor czerwony: należy on do $\operatorname{im} B^3$, ale nie należy do $\operatorname{im} B^4$,
- 3 wektory szary: należą do $\operatorname{im} B^4$.




Rysunek 12.2: Rozszerzamy bazę dla $\ker B$ do bazy dla $\mathbb{C}^k$.

Następnie rozszerzamy $S_0 \cup S_1 \cup S_2 \cup S_3 \cup S_4$ do bazy $\mathbb{C}^k$ tak jak to pokazano na Rys. 12.2.

Przykładowo, każdy wektor szary z jądra jest obrazem $B^4 u \in \ker B$ pewnego wektora $u \in \mathbb{C}^k$. Dostarcza on więc 5 wektorów

$$B^4 u,\ B^3 u,\ B^2 u,\ Bu,\ u.$$

Mamy 3 wektory szare w jądrze $B^4 u_1,\ B^4 u_2,\ B^4 u_3 \in \ker B$ dające 15 wektorów

$$B^4 u_1,\ B^3 u_1,\ B^2 u_1,\ Bu_1,\ u_1, \quad B^4 u_2,\ B^3 u_2,\ B^2 u_2,\ Bu_2,\ u_2, \quad B^4 u_3,\ B^3 u_3,\ B^2 u_3,\ Bu_3,\ u_3.$$

Analogicznie:

- wektor czerwony $B^3 u_4 \in \ker B$ daje 4 wektory
$$B^3 u_4,\ B^2 u_4,\ Bu_4,\ u_4,$$
- 2 wektory zielone $B^2 u_5,\ B^2 u_6 \in \ker B$ dają 6 wektorów
$$B^2 u_5,\ Bu_5,\ u_5, \quad B^2 u_6,\ Bu_6,\ u_6,$$
- 2 wektory żółte $Bu_7,\ Bu_8 \in \ker B$ dają 4 wektory
$$Bu_7,\ u_7, \quad Bu_8,\ u_8,$$
- mamy jeszcze 3 wektory czarne $u_9,\ u_{10},\ u_{11} \in \ker B$.

W efekcie otrzymaliśmy $15 + 4 + 6 + 4 + 3 = 32$ wektory. Pokażemy, że jest ich dokładnie $k$, gdy $B \in M_{k\times k}(\mathbb{C})$. Oznaczmy

$$d_i = \dim M_i = \dim\left(\ker B \cap \operatorname{im} B^i\right), \quad i = 0, 1, 2, 3, 4, 5,$$

oraz

$$r_i = \operatorname{rank} B^i = \dim \operatorname{im} B^i, \quad i = 0, 1, 2, 3, 4, 5.$$

Z lematu 12.3 wynika, że

$$d_i = r_i - r_{i+1}, \quad d_5 = 0 = r_5.$$

Oznacza to, że w zbiorze $S_i$ mamy

$$\nu_i = d_i - d_{i+1} = r_i - 2r_{i+1} + r_{i+2}$$

wektorów i każdy z nich dostarcza $i+1$ wektorów. Wszystkich wektorów jest więc

$$\begin{aligned}\sum_{i=0}^{4}(i+1)\nu_i = \sum_{i=0}^{4}(i+1)(d_i - d_{i+1})\\
&= d_0 - d_1 + 2(d_1 - d_2) + 3(d_2 - d_3) + 4(d_3 - d_4)\\
&= d_0 + d_1 + d_2 + d_3 + d_4\\
&= (r_0 - r_1) + (r_1 - r_2) + (r_2 - r_3) + (r_3 - r_4) + (r_4 - r_5)\\
&= r_0\\
&= k.\end{aligned}$$

Wystarczy więc pokazać, że skonstruowane wektory są liniowo niezależne. Uporządkujmy wektory następująco: $X_i$ dla $i = 0, 1, 2, 3, 4$ jest macierzą której kolumnami są wektory na poziomie $i$, czyli przykładowo kolumnami $X_0$ są wektory z $\ker B$, a kolumnami $X_4$, są wektory $u_1, u_2, u_3$. Kolumnami $X_3$ są $Bu_1, Bu_2, Bu_3, u_4$ itd. Definiujemy macierze

$$\begin{aligned}
Q_0 &= \left[\begin{array}{c|c|c|c|c} X_0 & BX_1 & B^2X_2 & B^3X_3 & B^4X_4 \end{array}\right] && \text{wektory na poziomie 0}\\
Q_1 &= \left[\begin{array}{c|c|c|c} X_1 & BX_2 & B^2X_3 & B^3X_4 \end{array}\right] && \text{wektory na poziomie 1}\\
Q_2 &= \left[\begin{array}{c|c|c} X_2 & BX_3 & B^2X_4 \end{array}\right] && \text{wektory na poziomie 2}\\
Q_3 &= \left[\begin{array}{c|c} X_3 & BX_4 \end{array}\right] && \text{wektory na poziomie 3}\\
Q_4 &= \left[\begin{array}{c} X_4 \end{array}\right] && \text{wektory na poziomie 4}
\end{aligned}$$

.

Pokażemy, że macierz

$$Q = \left[\begin{array}{c|c|c|c|c} Q_0 & Q_1 & Q_2 & Q_3 & Q_4 \end{array}\right]$$

jest nieosobliwa. Zauważmy, że kolumny macierzy $B^jX_j$ są na poziomie 0 i są częścią bazy dla $\ker B$, więc są liniowo niezależne. Stąd kolumny macierzy $B^jQ_j$ są też częścią bazy dla $\ker B$, więc

$$\ker B^jQ_j = \{0\}.$$

Pokażemy, że $\ker Q = \{0\}$. Niech $z = [z_0, z_1, z_2, z_3, z_4] \in \mathbb{C}^k$ będzie taki, że $Qz = 0$ oraz $z_j$ ma tyle współrzędnych ile kolumn ma $Q_j$. Ponieważ $Qz = 0$, więc

$$\begin{aligned}0 &= B^4Qz\\
&= \left[\begin{array}{c|c|c|c|c} B^4Q_0 & b^4Q_1 & B^4Q_2 & B^4Q_3 & B^4Q_4 \end{array}\right] z,\end{aligned}$$

czyli $z_4 \in \ker B^4Q_4$, czyli $z_4 = 0$. Stosując analogiczne rozumowanie dla $0 = B^3Qz$ otrzymujemy, że $z_3 = 0$. Kontynuując, otrzymujemy, że $z = 0$.

□

Jako wniosek otrzymujemy:

**Twierdzenie 12.4** (Postać Jordana macierzy nilpotentenj)**.** *Dla macierzy nilpotentnej $B \in M_{k\times k}(\mathbb{C})$ istnieje taka macierz nieosobliwa $S \in M_{k\times k}(\mathbb{C})$, że $S^{-1}AS$ jest nilpotentną macierzą Jordana.*

**Przykład 12.4.** Przeanalizujemy lemat 12.4 dla $k = 4$. Z punktu (ii) wynika, że $k = a_1 + \ldots + a_s, \quad a_i \geqslant 1$. Mamy następujące możliwości

(1) $s = 4$ i $a_1 = a_2 = a_3 = a_4 = 1$,

(2) $s = 3$ i $a_1 = a_2 = 1$, $a_3 = 2$,

(3) $s = 2$ i $a_1 = a_2 = 2$,

(4) $s = 2$ i $a_1 = 1$, $a_2 = 3$,

(5) $s = 1$ i $a_1 = 4$.

W przypadku (1) wektory $u_1$, $u_2$, $u_3$, $u_4$ tworzą bazę $\ker B$, czyli $B = 0$.

W przypadku (2) mamy bazę $u_1$, $u_2$, $Bu_3$, $u_3$, przy czym $u_1$, $u_2$, $Bu_3$ są bazą $\ker B$. Macierz $B$ ma w tej bazie postać

$$\left[\begin{array}{cc|cc} \boxed{0} & 0 & 0 & 0 \\ 0 & \boxed{0} & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right], \quad B^2 = 0.$$

W przypadku (3) mamy bazę $Bu_1$, $u_1$, $Bu_2$, $u_2$, przy czym $Bu_1$, $Bu_2$ tworzą bazę $\ker B$. W tej bazie $B$ ma postać

$$\left[\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right], \quad B^2 = 0.$$

W przypadku (4) mamy bazę $u_1$, $B^2u_2$, $Bu_2$, $u_2$, przy czym $u_1$, $B^2u_2$ tworzą bazę $\ker B$. Macierz $B$ ma w niej postać

$$\left[\begin{array}{c|ccc} 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right], \quad B^3 = 0.$$

W przypadku (5) mamy bazę $B^3u$, $B^2u$, $Bu, u$, pzy czym bazą dla $\ker B$ jest $B^3u$. W tej bazie $B$ ma postać

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B^4 = 0.$$

## 12.3 Zespolona i rzeczywista postać Jordana

**Twierdzenie 12.5** (Zespolona postać Jordana)**.** *Dla macierzy $A \in M_{n\times n}(\mathbb{C})$ istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{C})$, że*

$$\Lambda = S^{-1}AS = \begin{bmatrix} \boxed{\Lambda_1} & & & \\ & \boxed{\Lambda_2} & & \\ & & \ddots & \\ & & & \boxed{\Lambda_k} \end{bmatrix},$$

*gdzie każdy blok $\Lambda_i$ $(i = 1, \dots, k)$ jest macierzą postaci*

$$\Lambda_i = \begin{bmatrix} \boxed{\lambda} & \mathbf{1} & 0 & \dots & 0 \\ 0 & \boxed{\lambda} & \mathbf{1} & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \boxed{\lambda} & \mathbf{1} \\ 0 & \dots & 0 & 0 & \boxed{\lambda} \end{bmatrix}$$

*z wartością własną $\lambda \in \mathbb{C}$ na przekątnej. Macierz $\Lambda$ nazywamy zespoloną postacią Jordana macierzy $A$, a macierze $\Lambda_i \in M_{i\times i}(\mathbb{C})$ blokami Jordana $A$.*

Przedstawimy ideę dowodu twierdzenia 12.5. Pierwszy krok polega na rozkładzie $\mathbb{C}^n$ na takie podprzestrzenie niezmiennicze dla $A \in M_{n\times n}(\mathbb{C})$, w których $A$ ma tylko jedną wartość własną. Następnie w każdej z nich stsoujemy lemat 12.4 do $A - \lambda I$.

U Dla macierzy $A \in M_{n\times n}(\mathbb{C})$ i wartości własnej $\lambda \in \mathbb{C}$ definiujemy ciąg podprzestrzeni

$$\ker(A - \lambda I)^k, \quad k = 1, 2, 3, \ldots$$

Oczywiście,

$$\ker(A - \lambda I)^1 \subset \ker(A - \lambda I)^2 \subset \ldots \subset \ker(A - \lambda I)^m \subset \ldots$$

**Wniosek 12.9.** *Istnieje najmniejsze takie $N_\lambda \geqslant 1$, że*

$$\ker(A - \lambda I)^{N_\lambda} = \ker(A - \lambda I)^{N_\lambda + 1} = \ker(A - \lambda I)^{N_\lambda + 2} = \ldots$$

**Wniosek 12.10.** *Zachodzi równość*

$$\ker(A - \lambda I)^{N_\lambda} \cap \operatorname{im}(A - \lambda I)^{N_\lambda} = \{0\}.$$

*W szczególności,*

$$\mathbb{C}^n = \ker(A - \lambda I)^{N_\lambda} \oplus \operatorname{im}(A - \lambda I)^{N_\lambda}.$$

*Ponadto*

- *Podprzestrzenie* $\ker(A - \lambda I)^{N_\lambda}$ *oraz* $\operatorname{im}(A - \lambda I)^{N_\lambda}$ *są niezmiennicze dla $A$.*
- *Jedyną wartością własną $A$ w restrykcji do* $\ker(A - \lambda I)^{N_\lambda}$ *jest $\lambda$. Ponadto $\lambda$ nie jest wartością własną $A$ w restrykcji do* $\operatorname{im}(A - \lambda I)^{N_\lambda}$.
- *Niech $n_\lambda$ oznacza krotność wartości własnej $\lambda$ jako pierwiastka wielomianu charakterystycznego $p_A$ macierzy $A \in M_{n\times n}(\mathbb{C})$. Wtedy*

$$n_\lambda = \dim \ker(A - \lambda I)^{N_\lambda}.$$

**Lemat 12.5.** *Przypuśćmy, że $\lambda_1, \ldots, \lambda_k$ są różnymi wartościami własnymi macierz $A \in M_{n\times n}(\mathbb{C})$. Wtedy*

$$\mathbb{C}^n = \ker(A - \lambda_1 I)^{N_{\lambda_1}} \oplus \ldots \oplus \ker(A - \lambda_k I)^{N_{\lambda_k}}.$$

*Dowód.* Zastosujemy indukcję względem $n = \dim V$. Jeśli $n = 1$, to teza zachodzi, bo $\mathbb{C} = \ker(A - \lambda_1 I)$. Dla $n > 1$ stosujemy rozkład

$$\mathbb{C}^n = \ker(A - \lambda_1 I)^{N_{\lambda_1}} \oplus \operatorname{im}(A - \lambda_1 I)^{N_{\lambda_1}}$$

i założenie indukcyjne do $A$ w restrykcji do $\operatorname{im}(A - \lambda_1 I)^{N_{\lambda_1}}$. □

*Dowód.* Dowód twierdzenia 12.5. Z lematu 12.5 wynika, że $A$ ma postać blokową

$$A = \begin{bmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_k} \end{bmatrix},$$

gdzie $A_i$ jest macierzą $A$ w restrykcji do $\ker(A - \lambda_1 I)^{N_{\lambda_1}}$. Następnie stosujemy lemat 12.4 do macierzy nilpotentnej $A_i - \lambda_i I$. □

**Twierdzenie 12.6** (Rzeczywista postać Jordana). *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą rzeczywistą. Istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{R})$, że*

$$\Lambda = S^{-1}AS = \begin{bmatrix} \boxed{\Lambda_1} & & & \\ & \boxed{\Lambda_2} & & \\ & & \ddots & \\ & & & \boxed{\Lambda_k} \end{bmatrix},$$

*gdzie każdy blok $\Lambda_i$ jest jednej z dwóch postaci*

$$\begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda & 1 \\ 0 & \dots & 0 & 0 & \lambda \end{bmatrix}, \qquad \begin{bmatrix} B & I_2 & 0 & \dots & 0 \\ 0 & B & I_2 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & B & I_2 \\ 0 & \dots & 0 & 0 & B \end{bmatrix},$$

*gdzie $B = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$ dla pewnej rzeczywistej wartości własnej $\lambda$ i zespolonej wartości własnej $\alpha \pm i\beta$ $(\beta \neq 0)$ macierzy $A$.*

Macierz $\Lambda$ nazywamy rzeczywistą postacią Jordana macierzy $A$.

**Twierdzenie 12.7.** *Załóżmy, że dla macierzy rzeczywistej $A \in M_{2n\times 2n}(\mathbb{R})$ wielomian charakterystyczny $p_A$ ma $n$ różnych pierwiastków $\lambda_j = \alpha_j + i\beta_j$ oraz $\overline{\lambda_j}$ z $\beta_j \neq 0$ dla $j = 1, \dots, n$. Wtedy istnieje taka macierz nieosobliwa $S \in M_{2n\times 2n}(\mathbb{R})$, że*

$$S^{-1}AS = \operatorname{diag}(D_1, \dots, D_n),$$

*gdzie*

$$D_j = \begin{bmatrix} \alpha_j & \beta_j \\ -\beta_j & \alpha_j \end{bmatrix}.$$

*Dowód.* Niech $v_j + iw_j \in \mathbb{C}^n$ będą wektorami własnymi dla $\lambda_j$. Wtedy $v_j + iw_j \in \mathbb{C}^n$ są wektorami własnymi dla $\overline{\lambda_j}$. Pokażemy, że wektory $v_1, w_1, \dots, v_n, w_n$ są liniowo niezależne, więc są bazą $\mathbb{R}^{2n}$. Przypuśćmy, że istnieją takie liczby rzeczywiste $c_j, d_j$ $(j = 1, \dots, n)$, że

$$\sum_{j=1}^{n}(c_j \cdot v_j + d_j \cdot w_j) = 0,$$

ale nie wszystkie liczby $c_j$, $d_j$ są równe zero. Wtedy

$$\sum_{j=1}^{n}((c_j - id_j)\cdot(v_j + iw_j) + (c_j + id_j)\cdot(v_j - iw_j)) = 0.$$

Ponieważ wektory własne $v_j \pm iw_j$ są liniowo niezależne, więc $c_j = d_j = 0$ dla wszystkich $j$. Otrzymana sprzeczność kończy dowód liniowej niezleżności.

Rozważmy macierz nieosobliwą $S = \left[\ v_1 \mid w_1 \mid \dots \mid v_n \mid w_n\ \right]$. Wtedy

$$\begin{aligned} S^{-1}ASe_{2j-1} &= S^{-1}Av_j \\ &= S^{-1}(\alpha_j \cdot v_j - \beta_j \cdot w_j) \\ &= \alpha \cdot e_{2j-1} - \beta_j \cdot e_{2j}. \end{aligned}$$

Podobnie,

$$S^{-1}ASe_{2j} = \beta_j \cdot e_{2j-1} + \alpha_j \cdot e_{2j},$$

skąd otrzymujemy, że

$$S^{-1}AS = \operatorname{diag}(D_1, \dots, D_n).$$

□

**Wniosek 12.11.** *Załóżmy, że wielomian charakterystyczny $p_A$ macierzy rzeczywistej $A \in M_{n\times n}(\mathbb{R})$ ma $n$ różnych pierwiastków w $\mathbb{C}$. Wtedy istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{R})$, że*

$$S^{-1}AS = \begin{bmatrix} \lambda_j & & & & & \\ & \ddots & & & & \\ & & \lambda_k & & & \\ & & & D_1 & & \\ & & & & \ddots & \\ & & & & & D_l \end{bmatrix},$$

*gdzie*

$$D_j = \begin{bmatrix} \alpha_j & \beta_j \\ -\beta_j & \alpha_j \end{bmatrix}.$$

*Ponadto pierwiastkami wielomianu charakterystycznego $p_A$ są $\lambda_s$ oraz $\alpha_j \pm i\beta_j$.*

*Dowód.* Wybieramy wektory własne $u_1, \dots, u_k \in \mathbb{R}^n$ dla rzeczywistych wartości własnych $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ oraz wektory $v_1, w_1, \dots, v_l, w_l$ odpowiadające zespolonym wartościom własnym opisane w twierdzeniu 12.7. Oczywiście, $n = k + 2l$. Wektory

$$u_1, \dots, u_k, v_1, w_1, \dots, v_l, w_l$$

są liniowo niezależne, bo

$$\begin{aligned} 0 &= \sum_m a_m u_m + \sum_j (c_j v_j + d_j w_j) \\ &= \sum_m a_m u_m + \sum_j \Big( \frac{c_j - id_j}{2}(v_j + iw_j) + \frac{c_j + id_j}{2}(v_j - iw_j) \Big), \end{aligned}$$

więc $a_m = c_j = d_j = 0$, bo powyższe wektory własne są liniowo niezależne nad $\mathbb{C}$.

Wystarczy przyjąć, że $S = \left[\, u_1 \mid \dots \mid u_k \mid v_1 \mid w_1 \mid \dots \mid v_l \mid w_l \,\right]$. □

Przypadek macierzy $A \in M_{n\times n}(\mathbb{R})$ opisanych we wniosku 12.11 jest szczególnie ważny z punktu widzenia zastosowań. Okazuje się, że możemy dowolnie mało zmienić współczynniki dowolnej macierzy $B \in M_{n\times n}(\mathbb{R})$, aby otrzymać macierz $A \in M_{n\times n}(\mathbb{R})$ spełniającą założenia wniosku 12.11. Ponadto jeśli macierz $A \in M_{n\times n}(\mathbb{R})$ spełnia założenia wniosku 12.11 i $B$ powstaje z $A$ przez dostatecznie małą zmianę współczynników macierzy $A$, to $B$ też spełnia założenia wniosku 12.11. Topolodzy mówią wtedy, że macierze $A$ spełniające założenia wniosku 12.11 stanowią zbiór otwarty i gęsty w przestrzeni $M_{n\times n}(\mathbb{R})$. Mówimy wtedy, że są one *generyczne*. Bardzo nieściśle oznacza to, że jeśli wybierzecie losowo macierz $A \in M_{n\times n}(\mathbb{R})$, to z bardzo dużym prawdopodobieństwem spełnia ona założenia wniosku 12.11.

**Przykład 12.5.** Zastanówmy się, jak może wyglądać zespolona postać Jordana macierzy $A \in M_{4\times 4}(\mathbb{C})$ w zależności od liczby liniowo niezależnych wektorów własnych macierzy $A$.

(i) $A$ ma 4 liniowo niezależne wektory własne $z_1, z_2, z_3, z_4 \in \mathbb{C}^4$. Jest to równoważne z faktem, że macierz $A$ jest diagonalizowalna. Dokładniej, jeśli $\lambda_i \in \mathbb{C}$ jest wartością własną wektora $z_i$, czyli $Az_i = \lambda_i \cdot z_i$, to dla macierzy $S = \left[\, z_1 \mid z_2 \mid z_3 \mid z_4 \,\right]$ mamy

$$S^{-1}AS = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3, \lambda_4).$$

Sytuacja ta ma miejsce przykładowo, gdy $A$ ma 4 różne wartości własne.

(ii) $A$ ma tylko 3 liniowo niezależne wektory własne $z_1, z_2, z_3 \in \mathbb{C}^4$. W konsekwencji $A$ ma co najwyżej 3 różne wartości własne. Rozważmy podprzypadki:
(a) $A$ ma 3 różne wartości własne $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$ odpowiadające wektorom własnym $z_1, z_2, z_3$. Wtedy któraś z wartości własnych, powiedzmy $\lambda_1$, ma krotność algebraiczną

(jako pierwiastek wielomianu charakterystycznego) równą 2. Istnieje taki wektor $z_1^* \in \mathbb{C}^4$, że $z_1, z_1^*, z_2, z_3$ jest bazą $\mathbb{C}^4$, $(A-\lambda_1 I)z_1^* = z_1$ i $A$ ma w tej bazie postać

$$\left[\begin{array}{cc|cc} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ \hline 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{array}\right].$$

(b) $A$ ma 2 różne wartości własne $\lambda_1, \lambda_2$. Wtedy którejś wartości własnej, powiedzmy $\lambda_1$, odpowiadają 2 liniowo niezależne wektory własne. Macierz $A$ ma w pewnej bazie postać

$$\left[\begin{array}{cc|cc} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ \hline 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{array}\right].$$

(c) $A$ ma jedną wartosść własną $\lambda$ krotności 4. Wtedy $A$ ma w pewnej bazie postać

$$\left[\begin{array}{cc|cc} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ \hline 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{array}\right].$$

(iii) $A$ ma dwa liniowo niezależne wektory własne. Stąd $A$ ma co najwyżej 2 różne wartości własne. Mamy następujące możliwości
(a) $A$ ma 2 róże wartości własne $\lambda_1, \lambda_2$. Wtedy albo obydwie mają krotność algebraiczną 2 albo na przykład $\lambda_1$ ma krotność algebraiczną 3, a $\lambda_2$ ma krotność algebraiczną 1. Macierz $A$ ma wtedy w pewnej bazie jedną z postaci:

$$\left[\begin{array}{cc|cc} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ \hline 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & \lambda_2 \end{array}\right] \quad \left[\begin{array}{ccc|c} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ \hline 0 & 0 & 0 & \lambda_2 \end{array}\right].$$

(b) $A$ ma jedną wartość własną $\lambda$ krotności algebraicznej 4. Analogicznie jak w przypadku (a) macierz $A$ ma jedną z postaci

$$\left[\begin{array}{cc|cc} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ \hline 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{array}\right] \quad \left[\begin{array}{ccc|c} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ \hline 0 & 0 & 0 & \lambda \end{array}\right].$$

(iv) $A$ ma tylko jeden liniowo niezależny wektor własny. Wtedy $A$ ma tylko jedną wartość własną $\lambda$ (dlaczego?) i $A$ ma w pewnej bazie postać

$$\begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}.$$

**Przykład 12.6.** Zajmiemy się teraz rzeczywistą postacią Jordana macierzy $A \in M_{4\times 4}(\mathbb{R})$. Jeśli wszystkie własne $A$ są rzeczywiste, to jesteśmy w sytuacji opisanej przez zespoloną postać Jordana. Przypuśćmy, że $A$ ma wartość własną $\lambda_1 = \alpha_1 + i\beta_1$ z $\beta_1 \neq 0$. Niech $z_1 = u_1 + iv_1 \in \mathbb{C}^4$ będzie wektorem własnym dla $\lambda_1$. Zacznijmy od kilku obserwacji:

(1) Ponieważ $A$ jest macierzą rzeczywistą, więc $\overline{\lambda_1} = \alpha_1 - i\beta_1$ jest również wartością własną $A$, bo wielomian charakterystyczny $p_A(z) = \det(A - zI)$ macierzy $A$ ma współczynniki rzeczywiste. Wtedy $\overline{z_1} = u_1 - iv_1 \in \mathbb{C}^4$ jest wektorem własnym dla $\overline{\lambda_1}$.

(2) Krotności algebraiczne wartości własnych $\lambda_1$ i $\overline{\lambda_1}$ są równe. Zauważmy, że

$$\underbrace{p_A(z)}_{\in\mathbb{R}[z]} = \underbrace{(z-\lambda_1)(z-\overline{\lambda_1})}_{=z^2-2\mathrm{Re}\,(\lambda_1)z+|\lambda_1|\in\mathbb{R}[z]} \quad p(z),$$

więc wielomian $p(z)$ ma również współczynniki rzeczywiste. Jeżeli $\lambda_1$ ma krotność 2, to $p(\lambda_1) = 0$, więc również $p(\overline{\lambda_1}) = 0$.

(3) Załóżmy, że $\lambda_1$ ma krotność algebraiczną 2. Wartość własna $\lambda_1$ ma dwa liniowo niezleżne wektory własne nad $\mathbb{C}$ wtedy i tylko wtedy, gdy $\overline{\lambda_1}$ ma 2 liniowo niezależne wektory własne.

Z powyższych uwag wynika, że mamy następujące możliwości ze względu na zespoloną postać Jordana macierzy $A$:
(a) $A$ ma w bazie $z_1, \overline{z_1}, w_2, w_3$ postać

$$\left[\begin{array}{cc|cc} \lambda_1 & 0 & 0 & 0 \\ 0 & \overline{\lambda_1} & 0 & 0 \\ \hline 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{array}\right], \quad \lambda_2, \lambda_3 \in \mathbb{R}.$$

Wektory $w_2$, $w_3 \in \mathbb{R}^4$ są wektorami własnymi dla rzeczywistych wartości własnych $\lambda_2, \lambda_3$. Pokazażemy, że wektory $u_1, v_1, w_2, w_3$ są liniowo niezależne nad $\mathbb{R}$. Przypuśćmy, że

$$a \cdot u_1 + b \cdot v_1 + c \cdot w_2 + d \cdot w_3 = 0,$$

dla pewnych $a, b, c, d \in \mathbb{R}$. Ponieważ

$$u_1 = \frac{z_1 + \overline{z_1}}{2}, \quad v_1 = \frac{z_1 - \overline{z_1}}{2i} = -i\frac{z_1 - \overline{z_1}}{2},$$

więc

$$\frac{a-ib}{2} \cdot z_1 + \frac{a+ib}{2} \cdot \overline{z_1} + c \cdot w_2 + d \cdot w_3 = 0.$$

W konsekwencji $a = b = c = d = 0$, bo wektory $z_1, \overline{z_1}, w_2, w_3$ są liniowo niezależne nad $\mathbb{C}$. Ponieważ $A(u_1 + iv_1) = (\alpha_1 + i\beta_1) \cdot (u_1 + iv_1)$, więc

$$A(u_1) = \alpha_1 \cdot u_1 - \beta_1 \cdot v_1, \quad A(v_1) = \beta_1 \cdot u_1 + \alpha_1 \cdot v_1.$$

W konsekwencji $A$ ma w bazie $u_1, v_1, w_2, w_3$ rzeczywistą postać Jordana:

$$\left[\begin{array}{cc|cc} \alpha_1 & \beta_1 & 0 & 0 \\ -\beta_1 & \alpha_1 & 0 & 0 \\ \hline 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{array}\right].$$

(b) $A$ ma w bazie $z_1, \overline{z_1}, w_2, w_3$ postać

$$\left[\begin{array}{cc|cc} \lambda_1 & 0 & 0 & 0 \\ 0 & \overline{\lambda_1} & 0 & 0 \\ \hline 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & \lambda_2 \end{array}\right], \quad \lambda_2 \in \mathbb{R}.$$

Wtedy możemy założyć, że $w_2, w_3 \in \mathbb{R}^4$. Z rozumowania w punkcie (a) wynika, że $u_1, v_1, w_2, w_3$ tworzą bazę $\mathbb{R}^4$ i $A$ ma w tej bazie rzeczywistą postać Jordana:

$$\left[\begin{array}{cc|cc} \alpha_1 & \beta_1 & 0 & 0 \\ -\beta_1 & \alpha_1 & 0 & 0 \\ \hline 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & \lambda_2 \end{array}\right].$$

(c) $A$ ma w bazie $z_1, \overline{z_1}, z_2 = u_2 + iv_2, \overline{z_2}$ postać

$$\left[\begin{array}{cc|cc} \lambda_1 & 0 & 0 & 0 \\ 0 & \overline{\lambda_1} & 0 & 0 \\ \hline 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \overline{\lambda_2} \end{array}\right], \quad \lambda_2 = \alpha_2 + i\beta_2, \quad \beta_2 \neq 0.$$

Pokażemy, że wektory $u_1, v_1, u_2, v_2 \in \mathbb{R}^4$ tworzą bazę $\mathbb{R}^4$. Wystarczy pokazać, że są one liniowo niezależne. Przypuśćmy, że

$$a \cdot u_1 + b \cdot v_1 + c \cdot u_2 + d \cdot v_2 = 0,$$

dla pewnych $a, b, c, d \in \mathbb{R}$. Wtedy

$$\frac{a - ib}{2} \cdot z_1 + \frac{a + ib}{2} \cdot \overline{z_1} + \frac{c - id}{2} \cdot z_2 + \frac{c + id}{2} \cdot \overline{z_2} = 0,$$

więc wszystkie współczynniki są zerowe, czyli $a = b = c = d = 0$. Macierz $A$ ma w bazie $u_1, v_1, u_2, v_2 \in \mathbb{R}^4$ rzeczywistą postać Jordana:

$$\left[\begin{array}{cc|cc} \alpha_1 & \beta_1 & 0 & 0 \\ -\beta_1 & \alpha_1 & 0 & 0 \\ \hline 0 & 0 & \alpha_2 & \beta_2 \\ 0 & 0 & -\beta_2 & \alpha_2 \end{array}\right].$$

(d) $A$ ma w bazie $z_1, z_2 = u_2 + iv_2, \overline{z_1}, \overline{z_2}$ postać

$$\left[\begin{array}{cc|cc} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ \hline 0 & 0 & \overline{\lambda_1} & 1 \\ 0 & 0 & 0 & \overline{\lambda_1} \end{array}\right].$$

Analogicznie jak w przypadku (c) pokazujemy, że wektory $u_1, v_1, u_2, v_2$ tworzą bazę $\mathbb{R}^4$. Ponadto

$$A(u_1 + iv_1) = (\alpha_1 + i\beta_1) \cdot (u_1 + iv_1)$$

oraz

$$A(u_2 + iv_2) = u_1 + iv_1 + (\alpha_1 + i\beta_1) \cdot (u_2 + iv_2).$$

Stąd

$$A(u_1) = \alpha_1 \cdot u_1 - \beta_1 \cdot v_1, \quad A(v_1) = \beta_1 \cdot u_1 + \alpha_1 \cdot v_1$$

oraz

$$A(u_2) = u_1 + \alpha_1 \cdot u_2 - \beta_1 \cdot v_2, \quad A(v_2) = v_1 + \beta_1 \cdot u_2 + \alpha_1 \cdot v_2,$$

więc $A$ ma w bazie $u_1, v_1, u_2, v_2$ rzeczywistą postać Jordana:

$$\left[\begin{array}{cc|cc} \alpha_1 & \beta_1 & 1 & 0 \\ -\beta_1 & \alpha_1 & 0 & 1 \\ \hline 0 & 0 & \alpha_1 & \beta_1 \\ 0 & 0 & -\beta_1 & \alpha_1 \end{array}\right].$$

**Przykład 12.7.** Znajdziemy postać Jordana macierzy

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 11 & 6 & -4 & -4 \\ 22 & 15 & -8 & -9 \\ -3 & -2 & 1 & 2 \end{bmatrix}.$$

Bezpośredni rachunek pokazuje, że wielomian charakterystyczny $p_A(z) = \det(A - zI)$ macierzy $A$ jest dany przez

$$p_A(z) = z^4 - 2z^2 + 1 = (z^2 - 1)^2 = (z - 1)^2(z + 1)^2.$$

Wartościami własnymi macierzy $A$ są $-1$ oraz $1$ i mają one krotność algebraiczną 2. Dla wartości własnej $-1$ mamy

$$A + I = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 11 & 7 & -4 & -4 \\ 22 & 15 & -7 & -9 \\ -3 & -2 & 1 & 3 \end{bmatrix},$$

a dla 1 mamy

$$A - I = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 11 & 5 & -4 & -4 \\ 22 & 15 & -9 & -9 \\ -3 & -2 & 1 & 1 \end{bmatrix}.$$

Ponieważ

$$\operatorname{rank}(A - I) = \operatorname{rank}(A + I) = 3,$$

więc jądra są 1-wymiarowe. Każdej z wartości własnych $\pm 1$ odpowiada więc tylko jeden liniowo niezależny wektor własny. Klatki Jordana odpowiadające $-1$ i $1$ są więc (odpowiednio) postaci:

$$\begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Sprawdzamy łatwo, że

$$\ker(A + I) = \operatorname{span}\{[-1, 1, -1, 0]^T\}, \quad \ker(A - I) = \operatorname{span}\{[0, 0, 1, -1]^T\}.$$

Szukamy teraz takich wektorów $v, w$, że

$$(A + I)v = [-1, 1, -1, 0]^T, \quad (A - I)w = [0, 0, 1, -1]^T.$$

Rozwiązaniami są

$$v = [0, -1, -2, 0]^T, \quad w = [1, 1, 0, 4]^T.$$

W bazie $(A + I)v, v, (A - I)w, w$ macierz $A$ ma postać

$$\left[\begin{array}{cc|cc} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

## 12.4 Postać Frobeniusa

**Definicja 12.3.** Blok Frobeniusa

*Blokiem Frobeniusa* nazywamy macierz $B \in M_{n\times n}(\mathbb{C})$ postaci

$$B = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & 0 & \dots & 0 & -a_2 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & \dots & 0 & 1 & 0 & -a_{n-2} \\ 0 & \dots & 0 & 0 & 1 & -a_{n-1} \end{bmatrix}.$$

U Sprawdzamy łatwo, że wielomian charakterystyczny bloku Frobeniusa $B$ jest równy

$$p_B(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0.$$

**Przykład 12.8.** Załóżmy, że macierz $A \in M_{3\times3}(\mathbb{C})$ ma w bazie $v_1$, $v_2$, $v_3$ postać bloku Jordana

$$\Lambda = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}.$$

Oznacza to, że

$$Av_1 = \lambda \cdot v_1, \quad Av_2 = v_1 + \lambda \cdot v_2, \quad Av_3 = v_2 + \lambda \cdot v_3,$$

czyli

$$(A - \lambda I)v_1 = 0, \quad (A - \lambda I)v_2 = v_1, \quad (A - \lambda I)v_3 = v_2,$$

więc

$$(A - \lambda I)^3 v_3 = 0.$$

Rozważmy wektory $v_3$, $Av_3$, $A^2v_3$. Zauważmy, że

$$\begin{aligned} A^2v_3 &= A(v_2 + \lambda \cdot v_3) \\ &= v_1 + \lambda \cdot v_2 + \lambda \cdot v_3. \end{aligned}$$

Wynika stąd, że

$$v_3, \quad Av_3 = v_2 + \lambda \cdot v_3, \quad A^2v_3 = v_1 + \lambda \cdot v_2 + \lambda \cdot v_3$$

tworzą bazę $\mathbb{C}^3$. Ponieważ $(A - \lambda I)^3 v_3 = 0$, więc

$$\begin{aligned} A(A^2v_3) &= A^3v_3 \\ &= 3\lambda A^2 v_3 - 3\lambda^2 Av_3 + \lambda^3 v_3. \end{aligned}$$

Wynika stąd, że $A$ ma w bazie $v_3$, $Av_3$, $A^2v_3$ postać blokową Frobeniusa

$$B = \begin{bmatrix} 0 & 0 & \lambda^3 \\ 1 & 0 & -3\lambda^2 \\ 0 & 1 & 3\lambda \end{bmatrix}.$$

Zwróćmy uwagę, że

$$a_0 = -\lambda^3 = -\det \Lambda, \quad a_1 = 3\lambda^2, \quad a_2 = -3\lambda = -\operatorname{tr} \Lambda$$

są współczynnikami wielomianu charakterystycznego $p_A = p_B = \det(A - xI) = (\lambda - x)^3$.

Stosując analogiczne rozumowanie do bloków Jordana dowolnej macierzy $A \in M_{n\times n}(\mathbb{C})$ otrzymujemy następujące:

**Twierdzenie 12.8** (Postać Frobeniusa)**.** *Dla macierzy $A \in M_{n\times n}(\mathbb{C})$ istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{C})$, że*

$$S^{-1}AS = \operatorname{diag}(B_1, \ldots, B_k),$$

*gdzie $B_i$ są blokami Frobeniusa.*

## 12.5 Eksponenta macierzy $e^A$

Niech $A \in M_{k\times k}(\mathbb{C})$. Rozważmy macierz

$$S_n := I + A + \frac{1}{2}A^2 + \ldots + \frac{1}{n!}A^n.$$

Pokażemy, że ciąg macierzy $S_n$ jest zbieżny do pewnej macierzy którą będziemy oznaczać $e^A$ i nazywać *eksponentą* macierzy $A$. Piszemy wtedy, że

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}, \quad A^0 := I.$$

**Lemat 12.6.** *Niech*
$$\mu = \max\{|a_{ij}| : 1 \leqslant i, j \leqslant k\}.$$
*Jeśli* $A^n = \left[a_{ij}^{(n)}\right]$, *to*
$$|a_{ij}^{(n)}| \leqslant (k\mu)^n, \quad 1 \leqslant i, j \leqslant k.$$
*W szczególności, każdy z* $k^2$ *szeregów*
$$\sum_{n=0}^{\infty} \frac{a_{ij}^{(n)}}{n!}$$
*jest bezwzględnie zbieżny, więc macierz*
$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}$$
*jest dobrze określona.*

*Dowód.* Zastosujemy indukcję względem $n$. Dla $n = 0$, mamy $A^0 = I$, $(k\mu)^0 = 1$, więc teza zachodzi. Załóżmy, że
$$|a_{ij}^{(n)}| \leqslant (k\mu)^n \quad \text{dla} \quad 1 \leqslant i, j \leqslant k.$$
Stąd
$$\begin{aligned} |a_{ij}^{(n+1)}| &= \left|\sum_{l=1}^{k} a_{il}^{(n)} a_{lj}\right| \\ &\leqslant \sum_{l=1}^{k} |a_{il}^{(n)}||a_{lj}| \\ &\leqslant \mu \sum_{l=1}^{k} |a_{il}^{(n)}| \\ &\leqslant k\mu(k\mu)^n \\ &= (k\mu)^{n+1}. \end{aligned}$$
□

**Przykład 12.9.** Niech $A$ będzie macierzą diagonalną
$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R}.$$
Wtedy
$$D^n = \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix},$$
czyli
$$S_n = \begin{bmatrix} \sum_{k=0}^{n} \frac{1}{k!}\lambda_1^k & 0 \\ 0 & \sum_{k=0}^{n} \frac{1}{k!}\lambda_2^k \end{bmatrix} \underset{n\to\infty}{\longrightarrow} \begin{bmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{bmatrix} = e^D.$$
**Przykład 12.10.** Niech
$$C = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}, \quad b \neq 0.$$
Wtedy
$$C = bJ, \quad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$
Zauważmy, że
$$J^2 = -I, \quad J^3 = -J. \quad J^4 = I,$$

czyli ciąg macierzy $J^n$ jest okresowy o okresie 4. Stąd

$$S_n \underset{n\to\infty}{\longrightarrow} e^C = \begin{bmatrix} \cos b & \sin b \\ -\sin b & \cos b \end{bmatrix},$$

gdyż dla $x \in \mathbb{R}$ mamy

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots, \qquad \cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots .$$

U Wiemy, że dla liczby zespolonej $\lambda \in \mathbb{C}$ szereg

$$\sum_{n=0}^{\infty} \frac{\lambda^n}{n!}$$

jest bezwzględnie zbieżny. Możemy więc zdefiniować

$$e^\lambda := \sum_{n=0}^{\infty} \frac{\lambda^n}{n!}.$$

Sprawdzamy łatwo, że

$$e^{ib} = \cos b + i \sin b, \quad e^{a+ib} = e^a(\cos b + i \sin b).$$

**Twierdzenie 12.9.** *Jeśli* $A, B \in M_{k\times k}(\mathbb{C})$ *oraz* $AB = BA$, *to*

$$e^{A+B} = e^A e^B.$$

*Dowód.* Ponieważ $AB = BA$, więc zachodzi wzór dwumianowy Newtona

$$(A+B)^n = \sum_{l=0}^{n} \binom{n}{l} A^l B^{n-l},$$

więc

$$\frac{1}{n!}(A+B)^n = \sum_{l=0}^{n} \frac{A^l B^{n-l}}{l!(n-l)!}.$$

Dla dowolnej liczby naturalnej $N \geqslant 0$ mamy

$$\begin{aligned}
\sum_{n=0}^{2N} \frac{1}{n!}(A+B)^n &= \underbrace{\sum_{n=0}^{2N}\sum_{l=0}^{n} \frac{A^l B^{n-l}}{l!(n-l)!}}_{(2N+1)(N+1) \quad \text{składników}} \\
&= \underbrace{\left(\sum_{n=0}^{N} \frac{A^n}{n!}\right)\left(\sum_{n=0}^{N} \frac{B^n}{n!}\right)}_{(N+1)^2 \quad \text{składników}} + \sum_{\max(p,l)>N,\ p+l\leqslant 2N} \frac{A^p}{p!}\frac{B^l}{l!}.
\end{aligned}$$

Zauważmy, że w drugim składniku sumy mamy $N(N+1)$ par $(p, l)$. Dla

$$\mu_A = \max\{|a_{ij}| : 1 \leqslant i,j \leqslant k\}, \quad \mu_B = \max\{|b_{ij}| : 1 \leqslant i,j \leqslant k\}$$

oraz $\mu = \max(\mu_A, \mu_B)$, dla wyrazów $c_{ij}$ macierzy $\frac{A^p}{p!}\frac{B^l}{l!}$ mamy

$$\begin{aligned}
|c_{ij}| &\leqslant k\frac{(k\mu)^p}{p!}\frac{(k\mu)^l}{l!} \\
&\leqslant \frac{(k^2\mu)^{2N}}{N!}.
\end{aligned}$$

Wtedy wartość bezwzględna każdego wyrazu macierzy

$$\sum_{\max(p,l)>N,\ p+l\leqslant 2N} \frac{A^p}{p!}\frac{B^l}{l!}$$

jest ograniczona przez

$$N(N+1)\frac{(k^2\mu)^{2N}}{N!} \underset{N\to\infty}{\longrightarrow} 0.$$

Wynika stąd, że

$$e^{A+B} = e^A e^B.$$

□

**Wniosek 12.12.** *Dla dowolnej macierzy $A \in M_{n\times n}(\mathbb{C})$ eksponenta $e^A$ jest macierzą nieosobliwą o macierzy odwrotnej $e^{-A}$.*

**Przykład 12.11.** Ponieważ

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} = aI + bJ$$

oraz macierze $aI$ i $bJ$ komutują, więc

$$e^A = e^{aI}e^{bJ} = e^a \begin{bmatrix} \cos b & \sin b \\ -\sin b & \cos b \end{bmatrix}.$$

**Przykład 12.12.** Rozważmy blok Jordana

$$\Lambda = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda & 1 \\ 0 & \dots & 0 & 0 & \lambda \end{bmatrix} \in M_{k\times k}(\mathbb{C}), \quad \lambda \in \mathbb{C}.$$

Zauważmy, że

$$\Lambda = \lambda I + N,$$

gdzie

$$N = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda & 1 \\ 0 & \dots & 0 & 0 & \lambda \end{bmatrix}, \quad N^k = 0.$$

Ponieważ $\lambda I$ i $N$ komutują, więc

$$\begin{aligned} e^\Lambda &= e^{\lambda I}e^N \\ &= e^{\lambda I}\left(I + N + \frac{N^2}{2!} + \ldots + \frac{N^{k-1}}{(k-1)!}\right) \\ &= e^{\lambda}I\left(I + N + \frac{N^2}{2!} + \ldots + \frac{N^{k-1}}{(k-1)!}\right) \\ &= e^{\lambda}\left(I + N + \frac{N^2}{2!} + \ldots + \frac{N^{k-1}}{(k-1)!}\right). \end{aligned}$$

**Wniosek 12.13.** *Jeśli $S \in M_{n\times n}(\mathbb{C})$ jest nieosobliwa, to*

$$e^{S^{-1}AS} = S^{-1}e^AS.$$

*Dowód.* Zauważmy, że

$$\begin{aligned}(S^{-1}AS)^n &= (S^{-1}A\underbrace{S)(S^{-1}}_{=I}AS)\dots(S^{-1}AS)(S^{-1}AS)\\ &= S^{-1}A^nS.\end{aligned}$$

Stąd

$$\begin{aligned}\sum_{n=0}^{N}\frac{(S^{-1}AS)^n}{n!} &= \sum_{n=0}^{N}\frac{S^{-1}A^nS}{n!}\\ &= S^{-1}\left(\sum_{n=0}^{N}\frac{A^n}{n!}\right)S \quad \text{korzystamy z liniowości} \quad S^{-1}.\end{aligned}$$

□

U Jeśli $A$ ma postać blokową

$$A = \operatorname{diag}(\Lambda_1, \dots, \Lambda_s),$$

to $e^A$ ma postać blokową

$$e^A = \operatorname{diag}(e^{\Lambda_1}, \dots, e^{\Lambda_s})$$

**Lemat 12.7.** *Jeśli $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ są wartościami własnymi macierzy $A \in M_{n\times n}(\mathbb{C})$, to $e^{\lambda_1}, \dots, e^{\lambda_n}$ są wartościami własnymi $e^A$. Ponadto jeśli $u_i$ jest wektorem własnym $A$ odpowiadającym $\lambda_i$, to $u_i$ jest wektorem własnym $e^{\lambda_i}$.*

*Dowód.* Z twierdzenia Jordana istnieje taka macierz nieosobliwa $S$, że

$$S^{-1}AS = \operatorname{diag}(\Lambda_1, \dots, \Lambda_s)$$

oraz wyrazy diagonalne bloków Jordana $\Lambda_i$ są wartości własne $\lambda_i$ macierzy $A$. Wtedy

$$e^{S^{-1}AS} = \operatorname{diag}(e^{\Lambda_1}, \dots, e^{\Lambda_s})$$

jest macierzą górnie trójkątną z wyrazami diagonalnymi $e^{\lambda_i}$. Ponieważ $e^A$ i $e^{S^{-1}AS}$ są podobne, więc mają takie same wartości własne.

Niech $u_i$ będzie wektorem własnym $A$ odpowiadającym $\lambda_i$. Wtedy

$$\begin{aligned}e^Au_i &= \left(\lim_{n\to+\infty}\sum_{k=0}^{n}\frac{A^k}{k!}\right)u_i\\ &= \lim_{n\to+\infty}\left(\sum_{k=0}^{n}\frac{A^ku_i}{k!}\right)\\ &= \lim_{n\to+\infty}\left(\sum_{k=0}^{n}\frac{\lambda_i^ku_i}{k!}\right)\\ &= \left(\lim_{n\to+\infty}\sum_{k=0}^{n}\frac{\lambda_i^k}{k!}\right)u_i\\ &= e^{\lambda_i}u_i,\end{aligned}$$

więc teza zachodzi. □

**Wniosek 12.14.** *Dla macierzy $A \in M_{n\times n}(\mathbb{C})$ mamy*

$$\det e^A = e^{\operatorname{tr} A}.$$

*W szczególności, jeśli $A \in M_{n\times n}(\mathbb{R})$, to $\operatorname{tr} A \in \mathbb{R}$, więc*

$$\det e^A > 0.$$

*Dowód.* Mamy

$$\begin{aligned} \det e^A &= e^{\lambda_1} \dots e^{\lambda_n} \\ &= e^{\lambda_1 + \dots + \lambda_n} \\ &= e^{\operatorname{tr} A}. \end{aligned}$$

□

**Lemat 12.8.** *Dla dowolnej macierzy $A \in M_{n\times n}(\mathbb{R})$ mamy $(e^A)^T = e^{A^T}$ oraz jeśli $A^T = -A$, to $e^A \in SO(n)$.*

*Dowód.* Równość $(e^A)^T = e^{A^T}$ jest konsekwencją tożsamości

$$\left(I + A + \frac{A^2}{2!} + \dots + \frac{A^n}{n!}\right)^T = I + A^T + \frac{(A^T)^2}{2!} + \dots + \frac{(A^T)^n}{n!}$$

Jeśli $A^T = -A$, to

$$(e^A)^T = e^{A^T} = e^{-A},$$

czyli

$$(e^A)^T e^A = e^{-A} e^A = e^0 = I.$$

Ponadto $\operatorname{tr} A = 0$, więc

$$\det e^A = e^{\operatorname{tr} A} = 1.$$

□

TEST → Dla macierzy symetrycznej $A = [a_{ij}] \in M_{3\times 3}(\mathbb{R})$ oceń prawdziwość zdań:

- Jeśli $A$ ma dodatnie wartości własne, to $a_{ii} > 0$ dla $i = 1, 2, 3$.
- Jeśli $a_{ii} > 0$ dla $i = 1, 2, 3$, to $A$ ma dodatnią wartość własną.
- Jeśli $a_{ii} > 0$ dla $i = 1, 2, 3$, to $A$ ma dodatnie wartości własne.
- Jeśli $\det A > 0$, to $A$ ma dodatnią wartość własną.
- Jeśli $\operatorname{tr} A > 0$, to $A$ ma dodatnią wartość własną.
- Jeśli $\operatorname{tr} A > 0$ i $\det A > 0$, to wszystkie wartości własne $A$ są dodatnie.

⟵

TEST → Dla macierzy $A, B \in M_{n\times n}(\mathbb{C})$ oceń które z poniższych zdań jest prawdziwe:

- Jeśli $A^k = 0$ dla pewnego $k \geqslant 2$, to $A = 0$.
- Jeśli $A^k = 0$ dla pewnego $k \geqslant 1$, to $\operatorname{tr} A = 0$.
- Jeśli $A^k = 0$ dla pewnego $k \geqslant 1$, to $\det A = 0$.
- Jeśli $A^k = 0$ dla pewnego $k \geqslant 2$, to $\operatorname{rank} A = 0$.
- Jeśli $\operatorname{tr} A = 0$, to $\det A = 0$.
- Jeśli $\operatorname{rank} A = r$, to $A$ ma $r$ niezerowych wartości własnych.
- Jeśli $A$ i $B$ są podobne, to mają takie same wartości własne.
- Jeśli $A$ i $B$ mają takie same wartości własne, to są podobne.
- $A$ i $B$ mają takie same wartości własne wtedy i tylko wtedy, gdy mają identyczne wielomiany charakterystyczne.

- $A$ nie może być podobna do $A + I$.
- Jeśli $AB - BA = A$, to $A$ jest osobliwa.
- Jeśli $\operatorname{tr} A^k = \operatorname{tr} B^k$ dla każdego $k \geqslant 1$, to $A = B$.
- Jeśli $A$ ma wartości własne $\lambda_1, \ldots, \lambda_n$, to $A$ jest podobna do macierzy $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.

- Macierze $\operatorname{diag}(1, 2, \ldots, n)$ i $\operatorname{diag}(n, \ldots, 2, 1)$ są podobne.
- Jeśli $A$ ma wielokrotne wartości własne, to nie jest diagonalizowalna.

$\longleftarrow$

# Rozdział 13

# Przestrzenie euklidesowe i ortogonalność

| |
|---|
| SŁOWA KLUCZOWE ROZDZIAŁU |
| rzut ortogonalny wektora na wektor $\Diamond$ podprzestrzenie ortogonalne $\Diamond$ dopełnienie ortogonalne $\Diamond$ ortogonalna suma prosta $\Diamond$ dopełnienie ortogonalne jądra i obrazu $\Diamond$ baza ortonormalna $\Diamond$ macierz ortogonalna $\Diamond$ rzut ortogonalny na podprzestrzeń $\Diamond$ macierz idempotentna $\Diamond$ ortogonalizacja Grama–Schmidta $\Diamond$ rozkład $QR$ $\Diamond$ macierz Grama $A^TA$ $\Diamond$ grupa ortogonalna $O(n)$ $\Diamond$ wartości własne macierzy symetrycznej $\Diamond$ twierdzenie Schura dla macierzy symetrycznej $\Diamond$ ortogonalna diagonalizacja macierzy symetrycznych $\Diamond$ grupy macierzowe $\Diamond$ izometrie $\Diamond$ metoda najmniejszych kwadratów |

Przypomnijmy, że euklidesowy iloczyn skalarny w $\mathbb{R}^n$ jest określony wzorem

$$(x|y) = \sum_{i=1}^{n} x_i y_i, \quad x, y \in \mathbb{R}^n.$$

Zadaje on normę euklidesową wzorem

$$\|x\| = \sqrt{(x|x)}, \quad x \in \mathbb{R}^n.$$

Ponadto wektory $x, y$ są *ortogonalne (prostopadłe)*, gdy $(x|y) = 0$.

## 13.1 Macierze ortogonalne

**Definicja 13.1.** **Zbiór wektorów ortogonalnych**
Niech $v_1, \ldots, v_k \in \mathbb{R}^n$. Powiemy, że $\{v_1, \ldots, v_k\}$ jest *zbiorem ortogonalnym*, jeśli $(v_i|v_j) = 0$ dla $i \neq j$.

**Wniosek 13.1.** *Jeśli $\{v_1, \ldots, v_k\} \subset \mathbb{R}^n$ jest zbiorem ortogonalnym i wektory $v_i$ są niezerowe, to wektory $v_1, \ldots, v_k$ są liniowo niezależne.*

*Dowód.* Przypuśćmy, że

$$c_1 \cdot v_1 + \ldots c_k \cdot v_k = 0, \quad c_i \in \mathbb{R}.$$

Wtedy dla każdego $j = 1, \ldots, k$ mamy

$$\begin{aligned} 0 &= (0|v_j) \\ &= (c_1 \cdot v_1 + \ldots c_k \cdot v_k | v_j) \\ &= c_1(v_1|v_j) + \ldots + c_k(v_k|v_j) \\ &= c_j(v_j|v_j). \end{aligned}$$

Stąd $c_j(v_j|v_j) = 0$, czyli $c_j = 0$. □

**Definicja 13.2.** Zbiór ortonormalny wektorów

Zbiór wektorów ortogonalnych $\{v_1, \ldots, v_k\} \subset \mathbb{R}^n$ nazywamy *zbiorem ortonormalnym*, jeśli dodatkowo $\|v_j\| = 1$ dla $j = 1, \ldots, k$.

**Definicja 13.3.** Rzut ortogonalny wektora na wektor

Rozważmy wektory $x, y \in \mathbb{R}^n$ i $y \neq 0$. Rzutem ortogonalnym wektora $x$ na $y$ nazywamy taki wektor $p \in \operatorname{span}\{y\}$, że $x - p$ jest ortogonalny do $y$.

**Wniosek 13.2.** *Niech wektory $x, y \in \mathbb{R}^n$ i $y$ jest niezerowy. Wtedy rzut ortogonalny $p$ wektora $x$ na wektor $y$ jest dany przez*

$$p = \frac{(x|y)}{\|y\|^2} \cdot y = (x|u) \cdot u, \quad u = \frac{y}{\|y\|}.$$

*Dowód.* Ponieważ $p \in \operatorname{span}\{y\}$, więc $p = a \cdot y$ dla pewnego $a \in \mathbb{R}$. Musimy tak dobrać $a$, żeby $(x - p|y) = 0$. Mamy

$$\begin{aligned} 0 &= (x - p|y) \\ &= (x - a \cdot y|y) \\ &= (x|y) - a(y|y), \end{aligned}$$

czyli

$$a = \frac{(x|y)}{(y|y)} = \frac{(x|y)}{\|y\|^2}.$$

□

Jeśli wektor $y$ jest jednostkowy ($\|y\| = 1$), to rzut ortogonalny wektora $x$ na prostą generowaną przez $y$ jest równy

$$p = (v|y)y.$$

**Lemat 13.1** (Rzut ortogonalny na podprzestrzeń)**.** *Niech $V \subset \mathbb{R}^n$ będzie podprzestrzenią wektorową. Załóżmy, że wektory $u_1, \ldots, u_k$ tworzą bazę ortonormalną dla $V$. Dla wektora $w \in \mathbb{R}^n$ rozważmy wektor*

$$Pw := (w|u_1) \cdot u_1 + \ldots + (w|u_k) \cdot u_k \in V.$$

*Wtedy*

*(i) Odwzorowanie $P : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ jest liniowe.*

*(ii) $(w - Pw|v) = 0$ dla dowolnych $v \in V$ oraz $w \in \mathbb{R}^n$.*

*(iii) Dla macierzy $U = \left[\begin{array}{c|c|c} u_1 & \ldots & u_k \end{array}\right] \in M_{n \times k}(\mathbb{R})$ mamy*

$$P = UU^T.$$

*Dowód.* Sprawdzamy łatwo, że $P$ jest liniowe i nazywamy je *rzutem ortogonalnym na podprzestrzeń $V$*. Pokażemy, że wektor $w - Pw$ jest ortogonalny do $V$ tzn. $(w - Pw|v) = 0$ dla dowolnego wektora $v \in V$. Wystarczy pokazać, że $(w - Pw|u_i) = 0$ dla $i = 1, \ldots, k$. Mamy

$$\begin{aligned} (w - Pw|u_i) &= (w|u_i) - ((w|u_1) \cdot u_1 + \ldots + (w|u_k) \cdot u_k|u_i) \\ &= (w|u_i) - (w|u_i) = 0. \end{aligned}$$

Rozważmy macierz $U = \left[\begin{array}{c|c|c} u_1 & \ldots & u_k \end{array}\right] \in M_{n \times k}(\mathbb{R})$. Pokażemy, że macierz odwzorowania $P$ w bazie standardowej $e_1, \ldots, e_n$ jest równa $UU^T \in M_{n \times n}(\mathbb{R})$. Dla $i = 1, \ldots, n$ mamy

$$\begin{aligned} Pe_i &= (e_i|u_1) \cdot u_1 + \ldots + (e_i|u_k) \cdot u_k \\ &= U[(u_1|e_i), \ldots, (u_k|e_i)]^T \\ &= UU^T e_i. \end{aligned}$$

□

**Wniosek 13.3.** *Jeśli* $\{u_1, \ldots, u_n\} \subset \mathbb{R}^n$ *jest zbiorem ortonormalnym, to wektory* $u_1, \ldots, u_n$ *tworzą bazę* $\mathbb{R}^n$*. Ponadto dla dowolnego wektora* $v \in \mathbb{R}^n$ *mamy*

$$v = (v|u_1) \cdot u_1 + \ldots + (v|u_n) \cdot u_n.$$

*Skalary* $(v|v_i)$ *nazywamy współczynnikami Fouriera wektora* $v$ *względem bazy ortonormalnej* $u_1, \ldots, u_n$*.*

*Dowód.* Ponieważ wektory $u_i$ są liniowo niezależne z wniosku 13.1, więc $u_1, \ldots, u_n$ tworzą bazę $\mathbb{R}^n$. Wektor $v$ ma jednoznaczne przedstawienie w tej bazie:

$$v = c_1 \cdot u_1 + \ldots + c_n \cdot u_n, \quad c_i \in \mathbb{R}.$$

Stąd dla $j = 1, \ldots, n$ mamy

$$(v|u_j) = c_1(u_1|u_j) + \ldots + c_n(u_n|u_j) = c_j(u_j|u_j) = c_j.$$

□

**Przykład 13.1.** Rozważmy wektor jednostkowy $u = [x, y, z]^T \in \mathbb{R}^3$ tzn. $\|u\| = 1$. Rzut ortogonalny wektora $v \in \mathbb{R}^3$ na prostą $\operatorname{span}\{u\}$ jest dany przez

$$P(v) = (v|u) \cdot u.$$

Z lematu 13.1 wynika, że

$$P = uu^T = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \begin{bmatrix} x & y & z \end{bmatrix} = \begin{bmatrix} x^2 & xy & xz \\ yx & y^2 & yz \\ zx & zy & z^2 \end{bmatrix} \in M_{3\times 3}(\mathbb{R}).$$

Oznaczmy przez $u^\perp$ zawierającą 0 płaszczyznę w $\mathbb{R}^3$, ortogonalną do prostej $\operatorname{span}\{u\}$. Wtedy rzut ortogonalny $R$ na płaszczyznę $u^\perp$ jest dany macierzą

$$R = I - uu^T.$$

Przyglądniemy się teraz symetrii $Q$ względem przechodzącej przez 0 płaszczyzny $u^\perp$ prostopadłej do wektora $u$. Zauważmy, że dla wektora $v \in \mathbb{R}^3$ mamy równość

$$v + Q(v) = 2R(v),$$

czyli

$$Q = I - 2uu^T.$$

Oczywiście, $Q^T = Q$ oraz $Q^2 = I$, więc $Q$ jest symetryczna i ortogonalna tzn. $Q^T Q = I$. Powyższe rozważania możemy powtórzyć w dowolnym wymiarze. Powyższy wzór nazywamy *transformacją A. S. Householdera (1904-1993).*

**Przykład 13.2.** Znajdziemy macierz $Q$ symetrii względem płaszczyzny $x+y-z=0$. Jest ona prostopadła do wektora $v = [1, 1, -1]^T$. Zaczynamy od normalizacji wektora $v$, czyli definiujemy wektor

$$u = \frac{v}{\|v\|} = \frac{1}{\sqrt{3}} \cdot v = \frac{1}{\sqrt{3}}[1, 1, -1]^T.$$

Wtedy

$$uu^T = \frac{1}{3} \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix},$$

więc

$$Q = I - 2uu^T = \frac{1}{3} \begin{bmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}.$$

**Definicja 13.4.** Macierz ortogonalna

Macierz $Q = \left[\begin{array}{c|c|c} q_1 & \ldots & q_n \end{array}\right] \in M_{n\times n}(\mathbb{R})$ nazywamy *ortogonalną*, jeśli jej kolumny $q_i$ tworzą zbiór ortonormalny.

**Twierdzenie 13.1** (Charakteryzacja macierzy ortogonalnych)**.** *Dla macierzy* $Q = \left[\begin{array}{c|c|c} q_1 & \ldots & q_n \end{array}\right] \in M_{n\times n}(\mathbb{R})$ *następujące warunki są równoważne*

*(i)* $Q$ *jest ortogonalna,*

*(ii)* $Q^TQ = I$,

*(iii)* $Q$ *jest nieosobliwa oraz* $Q^T = Q^{-1}$,

*(iv)* $(Qx|Qy) = (x|y)$ *dla dowolnych* $x, y \in \mathbb{R}^n$,

*(v)* $\|Qx\| = \|x\|$ *dla każdego* $x \in \mathbb{R}^n$.

*Dowód.* $(i) \Leftrightarrow (ii)$ Obydwa warunki (i) i (ii) są równoważne z równościami

$$(q_i|q_j) = \delta_{ij}, \quad i, j = 1, \ldots, n.$$

$(ii) \Leftrightarrow (iii)$ Wynika wprost z definicji.

$(iv) \Leftrightarrow (v)$ Oczywiście (iv) implikuje (v). Załóżmy, że zachodzi warunek (v). Dla dowolnych wektorów $x, y \in \mathbb{R}^n$ mamy

$$\begin{aligned}(Q(x+y)|Q(x+y)) &= \|Q(x+y)\|^2\\ &= \|x+y\|^2\\ &= (x+y|x+y).\end{aligned}$$

Ponieważ

$$\begin{aligned}(Q(x+y)|Q(x+y)) &= \|Qx\|^2 + \|Qy\|^2 + 2(Qx|Qy),\\ (x+y|x+y) &= \|x\|^2 + \|y\|^2 + 2(x|y),\end{aligned}$$

więc $(Qx|Qy) = (x|y)$, co dowodzi warunku (iv).

$(i) \Leftrightarrow (iv)$ Zauważmy, że z warunku (iv) wynika, że

$$(q_i|q_j) = (Qe_i|Qe_j) = (e_i|e_j) = \delta_{ij},$$

czyli (iv) implikuje (i).

Z warunku (i) mamy $Q^TQ = I$, więc dla $x, y \in \mathbb{R}^n$ zachodzi równość

$$(Qx|Qy) = (x|Q^TQy) = (x|y).$$

□

## 13.2 Ortogonalizacja Grama–Schmidta. Rozkład $QR$.

Załóżmy, że $v_1, v_2, v_3$ tworzą bazę $\mathbb{R}^3$. Znajdziemy taką bazę ortonormalną $u_1$, $u_2$, $u_3$, że

$$\begin{aligned}\operatorname{span}\{u_1\} &= \operatorname{span}\{v_1\}\\ \operatorname{span}\{u_1, u_2\} &= \operatorname{span}\{v_1, v_2\}\\ \operatorname{span}\{u_1, u_2, u_3\} &= \operatorname{span}\{v_1, v_2, v_3\} = \mathbb{R}^3.\end{aligned}$$

Zastosujemy procedurę zaproponowanaą przez J. P. Grama (1850–1916) i E. Schmidta (1876–1959). Zaczynamy proces, definiując

$$u_1 = \frac{v_1}{\|v_1\|}.$$

Rozważmy $p_1$ rzut wektora $v_2$ na $\text{span}\{u_1\} = \text{span}\{v_1\}$, czyli

$$p_1 = (v_2|u_1) \cdot u_1.$$

Wtedy $v_2 - p_1 \perp u_1$. Zauważmy, że $v_2 - p_1 \neq 0$, bo

$$v_2 - p_1 = \frac{-(v_2|u_2)}{\|v_1\|} \cdot v_1 + v_2,$$

a wektory $v_1$, $v_2$ są liniowo niezależne. Możemy więc zdefiniować

$$u_2 = \frac{v_2 - p_1}{\|v_2 - p_1\|}.$$

Wektor $u_2$ jest prostopadły do $u_1$ i ma normę jeden. Ponadto

$$\text{span}\{u_1, u_2\} = \text{span}\{v_1, v_2\}.$$

Dla konstrukcji $u_3$ rozważamy projekcję $p_2$ wektora $v_3$ na podprzestrzeń $\text{span}\{u_1, u_2\} = \text{span}\{v_1, v_2\}$, czyli

$$p_2 = (v_3|u_1) \cdot u_1 + (v_3|u_2) \cdot u_2.$$

Wektor $v_3 - p_2$ jest niezerowy, bo w przeciwnym razie

$$v_3 = p_2 \in \text{span}\{u_1, u_2\} = \text{span}\{v_1, v_2\},$$

co jest niemożliwe, bo $v_1$, $v_2$, $v_3$ są liniowo niezależne. Możemy więc zdefiniować

$$u_3 = \frac{v_3 - p_2}{\|v_3 - p_2\|}.$$

Wtedy $u_3 \in (\text{span}\{u_1, u_2\})^\perp$ oraz

$$\text{span}\{u_1, u_2, u_3\} = \text{span}\{v_1, v_2, v_3\} = \mathbb{R}^3.$$

Algorytm Grama–Schmidta
Niech $v_1, \ldots, v_n$ będzie bazą $\mathbb{R}^n$.

---

**Krok 1.** Definiujemy

$$u_1 = \frac{v_1}{\|v_1\|}$$

**Krok 2.** Definiujemy rekurencyjnie wektory $u_2, \ldots, u_n$ jako

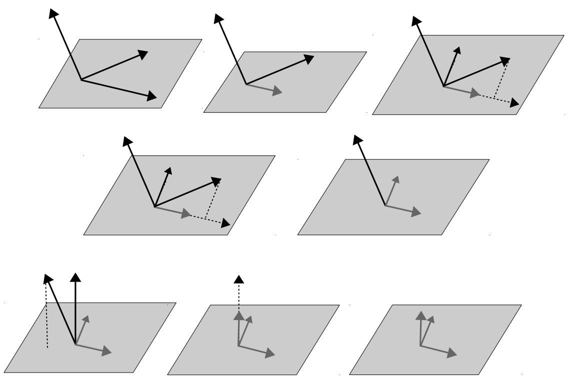
$$u_{k+1} = \frac{v_{k+1} - p_k}{\|v_{k+1} - p_k\|} \qquad p_k = (v_{k+1}|u_1) \cdot u_1 + \ldots + (v_{k+1}|u_k) \cdot u_k$$

tzn. $p_k$ jest rzutem ortogonalnym $v_{k+1}$ na podprzestrzeń $\text{span}\{u_1, \ldots, u_k\}$.

Wtedy

(1) $\text{span}\{u_1, \ldots, u_k\} = \text{span}\{v_1, \ldots, v_k\}$,

(2) $u_1, \ldots, u_n$ jest bazą ortonormalną dla $\mathbb{R}^n$.



Rysunek 13.1: Komiks o procesie Grama–Schmidta. Postarajcie się skojarzyć kolejne obrazki z krokami dokonywanymi w algorytmie Grama–Schmidta.

---

GRAM–SCHMIDT W ZAPISIE MACIERZOWYM
Niech $v_1, \ldots, v_n \in \mathbb{R}^m$ będą liniowo niezależne.

---

Procedura Grama–Schmidta może być wyrażona formułami

$$u_k = \frac{(I - U_k U_k^T)v_k}{\|(I - U_k U_k^T)v_k\|}, \quad k = 1, \ldots, n,$$

gdzie $U_1 = 0 \in M_{m\times 1}(\mathbb{R})$ oraz

$$U_k = \left[\begin{array}{c|c|c} u_1 & \ldots & u_{k-1} \end{array}\right] \in M_{m\times(k-1)}(\mathbb{R}), \quad k > 1.$$

---

**Przykład 13.3.** Załóżmy, że niezerowe wektory $v_1, \ldots, v_k \in \mathbb{R}^n$ są ortogonalne. Definiując $u_i = \frac{v_i}{\|v_i\|}$, otrzymujemy bazę ortonormalną dla podprzestrzeni $V = \operatorname{span}\{v_1, \ldots, v_k\} = \operatorname{span}\{u_1, \ldots, u_k\}$. Rzut ortogonalny wektora $v \in \mathbb{R}^n$ na podprzestrzeń $V$ jest dany przez

$$\begin{aligned} Pv &= (v|u_1)u_1 + \ldots + (v|u_n)u_n \\ &= \frac{(v|v_1)}{\|v_1\|^2}v_1 + \ldots + \frac{(v|v_n)}{\|v_n\|^2}v_n \\ &= \frac{(v|v_1)}{(v_1|v_1)}v_1 + \ldots + \frac{(v|v_n)}{(v_n|v_n)}v_n \end{aligned}$$

Zauważmy, że $\|v - Pv\|$ jest odległością wektora $v$ od podprzestrzeni $V$.

U Niech $v_1, \ldots, v_n$ będzie bazą $\mathbb{R}^n$. Stosując algorytm Grama–Schmidta otrzymujemy bazę ortogonalną $u_1, \ldots, u_n$, daną przez

$$\begin{aligned} u_1 &= \frac{v_1}{\|v_1\|} \\ u_2 &= \frac{v_2 - (v_2|u_1)u_1}{\|v_2 - (v_2|u_1)u_1\|} \\ u_3 &= \frac{v_3 - (v_3|u_1)u_1 - (v_3|u_2)u_2}{\|v_3 - (v_3|u_1)u_1 - (v_3|u_2)u_2\|} \\ &\cdots\cdots\cdots \\ u_n &= \frac{v_n - (v_n|u_1)u_1 - \ldots - (v_n|u_{n-1})u_{n-1}}{\|v_n - (v_n|u_1)u_1 - \ldots - (v_n|u_{n-1})u_{n-1}\|}. \end{aligned}$$

Jeżeli chcemy otrzymać tylko bazę ortogonalną $w_1, \ldots, w_n$, a nie ortonormalną, to modyfikujemy nieco powyższą procedurę pomijając ortonormalizację. Wtedy

$$\begin{aligned}
w_1 &= v_1 \\
w_2 &= v_2 - \frac{(v_2|w_1)}{(w_1|w_1)} w_1 \\
w_3 &= v_3 - \frac{(v_3|w_1)}{(w_1|w_1)} w_1 - \frac{(v_3|w_2)}{(w_2|w_2)} w_2 \\
&\cdots\cdots\cdots \\
w_n &= v_n - \frac{(v_n|w_1)}{(w_1|w_1)} w_1 - \ldots - \frac{(v_n|w_{n-1})}{(w_{n-1}|w_{n-1})} w_{n-1}.
\end{aligned}$$

Wtedy $\|w_k\|$ jest odległością wektora $v_k$ od podprzestrzeni $\operatorname{span}\{v_1, \ldots, v_{k-1}\}$. Zauważmy, że procedura ortogonalizacji nie wymaga liniowej niezależności wektorów $v_1, \ldots, v_n$. Ważne jest tylko aby były one niezerowe.

**Przykład 13.4.** Rozważmy płaszczyznę $V$ w $\mathbb{R}^3$ rozpiętą na wektorach $v_1 = [1, 2, 2]^T$ i $v_2 = [-1, 0, 2]^T$. Przypuśćmy, że chcemy znaleźć bazę ortogonalną $V$ i rozszerzyć ją do bazy ortogonalnej $\mathbb{R}^3$. Możemy postąpić następująco. Najpierw rozszerzamy wektory $v_1$, $v_2$ do bazy $\mathbb{R}^3$. Możemy to zrobić przykładowo dokładając wektor $v_3 = [0, 0, 1]^T = e_3$. Następnie stosujemy proces ortogonalizacji Grama–Schmidta do tej bazy. Mamy

$$\begin{aligned}
w_1 &= v_1 = [1, 2, 2]^T \\
w_2 &= v_2 - \frac{(v_2|w_1)}{(w_1|w_1)} w_1 \\
&= [-1, 0, 2]^T - \frac{3}{9}[1, 2, 2]^T \\
&= [-4/3, -2/3, 4/3]^T \\
w_3 &= v_3 - \frac{(v_3|w_1)}{(w_1|w_1)} w_1 - \frac{(v_3|w_2)}{(w_2|w_2)} w_2 \\
&= [0, 0, 1]^T - \frac{2}{9}[1, 2, 2]^T - \frac{4/3}{4}[-4/3, -2/3, 4/3]^T \\
&= [2/9, -2/9, 1/9]^T.
\end{aligned}$$

Jeśli interesuje nas baza ortonormalna $u_1, u_2, u_3$, to wystarczy teraz tylko znormalizować wektory $w_1, w_2, w_3$, przyjmując $u_i = \frac{w_i}{\|w_i\|}$. Ponieważ

$$\|w_1\| = 3, \quad \|w_2\| = 2, \quad \|w_3\| = 1/3,$$

więc

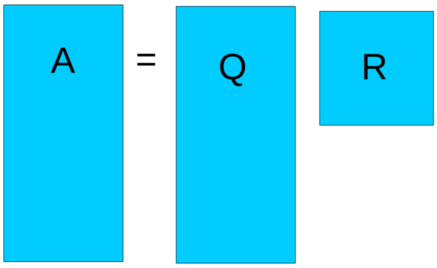
$$u_1 = \frac{1}{3}[1, 2, 2]^T, \quad u_2 = \frac{1}{3}[-2, -1, 2]^T, \quad u_3 = \frac{1}{3}[2, -2, 1]^T.$$

**Przykład 13.5.** Znajdziemy odległość $d$ wektora $v = [0, 0, 0, 1]^T$ od podprzestrzeni $V \subset \mathbb{R}^4$, rozpiętej na wektorach

$$v_1 = [1, -1, 1, -1]^T, \quad v_2 = [1, 1, 3, -1]^T, \quad v_3 = [-3, 7, 1, 3]^T.$$

W tym celu zastosujemy procedurę ortogonalizacji do wektorów $v_1, v_2, v_3, v$, prowadzącą do układu ortogonalnego $w_1, w_2, w_3, w_4$. Wtedy $d = \|w_4\|$. Otrzymujemy, że

$$\begin{aligned}
w_1 &= v_1 = [1, -1, 1, -1]^T \\
w_2 &= v_2 - \frac{(v_2|w_1)}{(w_1|w_1)} w_1 \\
&= [1, 1, 3, -1]^T - \frac{4}{4}[1, -1, 1, -1]^T = [0, 2, 2, 0]^T \\
w_3 &= v_3 - \frac{(v_3|w_1)}{(w_1|w_1)} w_1 - \frac{(v_3|w_2)}{(w_2|w_2)} w_2 \\
&= [-3, 7, 1, 3]^T - \frac{-12}{4}[1, -1, 1, -1]^T - \frac{16}{8}[0, 2, 2, 0]^T \\
&= [0, 0, 0, 0]^T.
\end{aligned}$$




Rysunek 13.2: Rozkład $QR$ macierzy $A$ o liniowo niezależnych kolumnach.

Oznacza to, że wektor $v_3 \in \operatorname{span}\{v_1, v_2\}$, więc $\operatorname{span}\{v_1, v_2, v_3\} = \operatorname{span}\{v_1, v_2\}$ jest płaszczyzną. Powinniśmy zatem dokonać ortogonalizacji wektorów $v_1, v_2, v$. Otrzymujemy, że

$$\begin{aligned}\tilde{w}_3 &= v - \frac{(v|w_1)}{(w_1|w_1)}w_1 - \frac{(v|w_2)}{(w_2|w_2)}w_2 \\ &= [0,0,0,1]^T - \frac{-1}{4}[1,-1,1,-1]^T - \frac{0}{8}[0,2,2,0]^T \\ &= [1/4,-1/4,1/4,3/4]^T,\end{aligned}$$

więc

$$d = \|\tilde{w}_3\| = \frac{1}{4}\|[1,-1,1,3]^T| = \frac{\sqrt{12}}{4} = \frac{\sqrt{3}}{2}.$$

U Algorytmu Grama-Schmidta nie musimy ograniczać tylko do ortonormalizacji bazy. Jeśli $v_1, \ldots, v_k \in \mathbb{R}^n$ są liniowo niezależne, to stosując algorytm otrzymujemy układ wektorów ortonormalnych $u_1, \ldots, u_k$. Pozwala to na rozkład dowolnej macierzy $A = \left[\begin{array}{c|c|c} a_1 & \ldots & a_k \end{array}\right] \in M_{n\times k}(\mathbb{R})$ o liniowo niezależnych kolumnach przedstawić w postaci iloczynu macierzy ortogonalnej $Q = \left[\begin{array}{c|c|c} u_1 & \ldots & u_k \end{array}\right] \in M_{n\times k}(\mathbb{R})$ i macierzy górnie trójkątnej $R \in M_{k\times k}(\mathbb{R})$. Opiszemy ten proces na przykładzie macierzy o trzech liniowo niezależnych kolumnach $A = \left[\begin{array}{c|c|c} v_1 & v_2 & v_3 \end{array}\right] \in M_{n\times 3}(\mathbb{R})$. Z ortogonalizacji Grama–Schmidta mamy

$$\begin{aligned} v_1 &= \|v_1\| \cdot u_1 \\ v_2 &= \underbrace{(v_2|u_1)\cdot u_1}_{=p_1} + \|v_2 - p_1\| \cdot u_2 \\ v_3 &= \underbrace{(v_3|u_1)\cdot u_1 + (v_3|u_2)\cdot u_2}_{=p_2} + \|v_3 - p_2\| \cdot u_3 \end{aligned}$$

Z definicji iloczynu macierzy mamy

$$A = \underbrace{\left[\begin{array}{c|c|c} u_1 & u_2 & u_3 \end{array}\right]}_{Q} \underbrace{\begin{bmatrix} \|v_1\| & (v_2|u_1) & (v_3|u_1) \\ 0 & \|v_2 - p_1\| & (v_3|u_2) \\ 0 & 0 & \|v_3 - p_2\| \end{bmatrix}}_{R}.$$

Analogicznie, dla $A = \left[\begin{array}{c|c} v_1 & v_2 \end{array}\right] \in M_{n\times 2}(\mathbb{R})$ o liniowo niezależnych kolumnach mamy

$$A = \left[\begin{array}{c|c} u_1 & u_2 \end{array}\right] \begin{bmatrix} \|v_1\| & (v_2|u_1) \\ 0 & \|v_2 - p_1\| \end{bmatrix}.$$

Zauważmy, że

$$\det \begin{bmatrix} \|v_1\| & (v_2|u_1) \\ 0 & \|v_2 - p_1\| \end{bmatrix} = \|v_1\|\|v_2 - p_1\|$$

jest polem równoległoboku rozpiętego na wektorach $v_1, v_2$.
Dla macierzy $A = \left[\begin{array}{c|c|c} v_1 & v_2 & v_3 \end{array}\right] \in M_{n\times 3}(\mathbb{R})$ o liniowo niezależnych kolumnach

$$\det \begin{bmatrix} \|v_1\| & (v_2|u_1) & (v_3|u_1) \\ 0 & \|v_2 - p_1\| & (v_3|u_2) \\ 0 & 0 & \|v_3 - p_2\| \end{bmatrix} = \|v_1\|\|v_2 - p_1\|\|v_3 - p_2\|$$

jest objętością równoległościanu rozpiętego na wektorach $v_1, v_2, v_3$.

Rozkład $QR$ macierzy $A$ o liniowo niezależnych kolumnach

$$\underbrace{\begin{bmatrix} | & | & | \\ | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \\ | & | & | \end{bmatrix}}_{\mathbf{A}} = \underbrace{\begin{bmatrix} | & | & | \\ | & | & | \\ u_1 & u_2 & u_3 \\ | & | & | \\ | & | & | \end{bmatrix}}_{\mathbf{Q}} \underbrace{\begin{bmatrix} \|v_1\| & (v_2|u_1) & (v_3|u_1) \\ 0 & \|v_2 - p_1\| & (v_3|u_2) \\ 0 & 0 & \|v_3 - p_2\| \end{bmatrix}}_{\mathbf{R}}.$$

Kolumny $v_1$, $v_2$, $v_3$ macierzy $A$ są liniowo niezależne. Wektory $u_1$, $u_2$, $u_3$ stanowią układ ortonormalny otrzymany z $v_1$, $v_2$, $v_3$ w wyniku procedury Grama–Schmidta.

$$\begin{aligned} v_1 &= \|v_1\| \cdot u_1 \\ v_2 &= \underbrace{(v_2|u_1) \cdot u_1}_{=p_1} + \|v_2 - p_1\| \cdot u_2 \\ v_3 &= \underbrace{(v_3|u_1) \cdot u_1 + (v_3|u_2) \cdot u_2}_{=p_2} + \|v_3 - p_2\| \cdot u_3 \end{aligned}$$

**Wniosek 13.4.** *Jeśli* $A = \begin{bmatrix} v_1 \mid v_2 \end{bmatrix} \in M_{n\times 2}(\mathbb{R})$ *ma liniowo niezależne kolumny, to pole* $P$ *równoległoboku rozpiętego na wektorach* $v_1, v_2$ *jest równe*

$$P = \sqrt{\det A^T A}.$$

*W szczególności, dla* $n = 2$ *otrzymujemy, że* $P = |\det A|$.

*Analogicznie, jeśli* $A = \begin{bmatrix} v_1 \mid v_2 \mid v_3 \end{bmatrix} \in M_{n\times 3}(\mathbb{R})$ *ma liniowo niezależne kolumny, to objętość* $O$ *równoległościanu rozpiętego na wektorach* $v_1, v_2, v_3$ *jest równa*

$$O = \sqrt{\det A^T A}.$$

*Dla* $n = 3$ *mamy* $O = |\det A|$.

*Dowód.* Przeprowadzimy dowód dla $A = \begin{bmatrix} v_1 \mid v_2 \mid v_3 \end{bmatrix} \in M_{n\times 3}(\mathbb{R})$. Rozważmy rozkład $A = QR$, gdzie $Q = \begin{bmatrix} u_1 \mid u_2 \mid u_3 \end{bmatrix}$. Zauważmy, że

$$Q^T Q = \begin{bmatrix} (u_1|u_1) & (u_1|u_2) & (u_1|u_3) \\ (u_2|u_1) & (u_2|u_2) & (u_2|u_3) \\ (u_3|u_1) & (u_3|u_2) & (u_3|u_3) \end{bmatrix} = I,$$

bo wektory $u_1, u_2, u_3$ są ortonormalne. Stąd

$$\begin{aligned} A^T A &= (QR)^T QR \\ &= R^T Q^T QR \\ &= R^T R, \end{aligned}$$

więc

$$\begin{aligned} \det A^T A &= \det R^T R \\ &= (\det R)^2 \\ &= O^2. \end{aligned}$$

Jeśli $n = 3$, to $\det A^T A = (\det A)^2$, co kończy dowód. □

**Wniosek 13.5** (Nierówność Hadamarda)**.** *Niech* $A = \left[\begin{array}{c|c|c} v_1 & v_2 & v_3 \end{array}\right] \in M_{3\times3}(\mathbb{R})$*. Wtedy*

$$|\det A| \leqslant \|v_1\|\|v_2\|\|v_3\|.$$

*Dowód.* Dla rozkładu $A = QR$ mamy

$$\begin{aligned}(\det A)^2 &= \det A^T A \\ &= (\det R)^2 \\ &= \|v_1\|^2\|v_2 - p_1\|^2\|v_3 - p_2\|^2 \\ &\leqslant \|v_1\|^2\|v_2\|^2\|v_3\|^2.\end{aligned}$$

□

U Niech $A \in M_{m\times n}(\mathbb{R})$. Zbiór rozwiązań równania $Ax = 0$ jest równy $\ker A \subset \mathbb{R}^n$. Zauważmy, że $x \in \ker A$ wtedy i tylko wtedy, gdy

$$a_{i1} \cdot x_1 + \ldots + a_{in} \cdot x_n = 0.$$

Oznacza to, że $x$ jest prostopadły do każdej kolumny macierzy $A^T$. Wynika stąd, że $x$ jest prostopadły do każdego wektora $y \in \operatorname{im} A^T \subset \mathbb{R}^n$. Powiemy wtedy, że podprzestrzenie $\ker(A)$ oraz $\operatorname{im} A^T$ są ortogonalne.

**Definicja 13.5. Podprzestrzenie ortogonalne**

Niech $V, W \subset \mathbb{R}^n$ będą podprzestrzeniami wektorowymi. Powiemy, że $V$, $W$ są *ortogonalne*, jeśli $(v|w) = 0$ dla dowolnych $v \in V$, $w \in W$. Piszemy wtedy, że $V \perp W$.

**Wniosek 13.6.** *Jeśli* $V \perp W$*, to* $V \cap W = \{0\}$*.*

*Dowód.* Niech $x \in V \cap W$. Wtedy $(x|x) = 0$, czyli $x = 0$. □

! Musimy być trochę ostrożni. Pojęcie ortogonalności podprzestrzeni $V, W$ nie zawsze jest zgodne z naszą intuicją. W przestrzeni $\mathbb{R}^3$ płaszczyzna $xy$ i $xz$ nie są ortogonalne, ale kąt między nimi jest kątem prostym. Przestrzenią ortogonalną do płaszczyzny $xy$ jest natomiast prosta $z$.

**Definicja 13.6. Dopełnienie ortogonalne podprzestrzeni**

Niech $Y \subset \mathbb{R}^n$ będzie podprzestrzenią wektorową. Dopełnienie ortogonalne $Y^\perp$ podprzestrzeni $Y$ definiujemy jako

$$Y^\perp = \Big\{x \in \mathbb{R}^n : (x|y) = 0 \quad \text{dla} \quad y \in Y\Big\}.$$

**Wniosek 13.7.** $Y^\perp$ *jest podprzestrzenią wektorową w* $\mathbb{R}^n$ *oraz* $Y \perp Y^\perp$*.*

**Twierdzenie 13.2** (Dopełnienie ortogonalne jądra i obrazu)**.** *Niech* $A \in M_{m\times n}(\mathbb{R})$*. Wtedy*

$$\ker A = (\operatorname{im} A^T)^\perp, \qquad \ker A^T = (\operatorname{im} A)^\perp.$$

*Dowód.* Pokażemy najpierw, że $\ker A \perp (\operatorname{im} A^T)$. Niech $v \in \ker A$ i $w = A^T u \in \operatorname{im} A^T$. Wtedy

$$(v|w) = (v|A^T u) = (Av|u) = (0, |u) = 0,$$

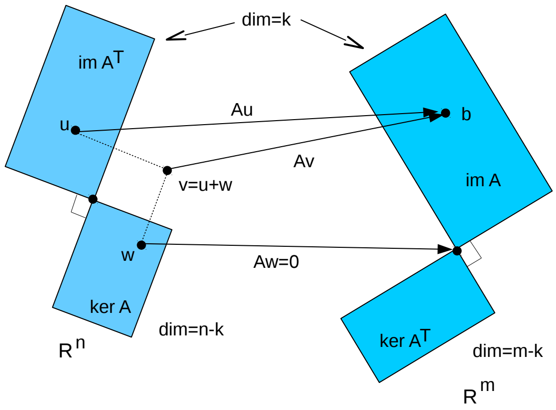
czyli $\ker A \perp (\operatorname{im} A^T)$, więc

$$\ker A \subset (\operatorname{im} A^T)^\perp.$$

Niech $x \in (\operatorname{im} A^T)^\perp$. Wtedy $x$ jest prostopadły do każdej kolumny $A^T$, więc $Ax = 0$, czyli $x \in \ker A$. □

**Twierdzenie 13.3** (Rozkład na ortogonalną sumę prostą)**.** *Jeśli* $S \subset \mathbb{R}^n$ *jest podprzestrzenią, to*

$$\mathbb{R}^n = S \oplus S^\perp$$




Rysunek 13.3: Macierz $A$ i rozkłady na ortogonalne sumy proste: $\mathbb{R}^n = \ker A \oplus \operatorname{im} A^T$ oraz $\mathbb{R}^m = \ker A^T \oplus \operatorname{im} A$.

*W szczególności,*

$$n = \dim S + \dim S^\perp.$$

*Dowód.* Ponieważ $S \perp S^\perp$, więc $S \cap S^\perp = \{0\}$. Wystarczy pokazać, że $\mathbb{R}^n = S + S^\perp$. Niech $v_1, \ldots, v_k$ będzie bazą ortonormalną dla $S$. Dla wektora $v \in \mathbb{R}^n$ rozważamy wektor $s = (v|v_1) \cdot v_1 + \ldots + (v|v_k) \cdot v_k \in S$. Wystarczy pokazać, że $v - s \in S^\perp$. Jest tak, bo dla $i = 1, \ldots, k$ mamy

$$(v - s|v_i) = (v|v_i) - (s|v_i) = (v|v_i) - (v|v_i) = 0.$$

□

**Wniosek 13.8** (Dopełnienie dopełnienia)**.** *Jeśli $S \subset \mathbb{R}^n$ jest podprzestrzenią, to*

$$(S^\perp)^\perp = S$$

*Dowód.* Jeśli $x \in S$, to $(x|y) = 0$ dla każdego $y \in S^\perp$. Stąd $x \in (S^\perp)^\perp$, czyli $S \subset (S^\perp)^\perp$. Ponieważ

$$\mathbb{R}^n = S \oplus S^T = S^T \oplus (S^T)^T,$$

więc $\dim S = \dim (S^T)^T$, czyli $S = (S^T)^T$. □

**Wniosek 13.9.** *Dla macierzy $A \in M_{m\times n}(\mathbb{R})$ mamy ortogonalne sumy proste*

$$\mathbb{R}^n = \ker A \oplus \operatorname{im} A^T, \quad \mathbb{R}^m = \ker A^T \oplus \operatorname{im} A.$$

## 13.3 Macierze symetryczne–diagonalizacja

**Twierdzenie 13.4** (Wartości własne rzeczywistej macierzy symetrycznej)**.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie symetryczna. Wtedy wartości własne $A$ są liczbami rzeczywistymi. Ponadto wektory własne odpowiadające różnym wartościom własnym są ortogonalne.*

*Dowód.* Niech $\lambda = a + ib \in \mathbb{C}$ będzie zespoloną wartością własną macierzy $A$ jako macierzy zespolonej, a $v + iw \in \mathbb{C}^n$ jej wektorem własnym dla $v, w \in \mathbb{R}^n$. Przypuśćmy, że $b \neq 0$. Ponieważ $A(v + iw) = \lambda(v + iw)$, więc

$$Av = a \cdot v - b \cdot w, \quad Aw = b \cdot v + a \cdot w.$$

Zwróćmy uwagę, że wtedy $v \neq 0$ i $w \neq 0$. Stąd

$$\begin{aligned}(Av, w) &= (a \cdot v - b \cdot w, w)\\ &= a(v|w) - b\|w\|^2,\end{aligned}$$

$$
\begin{aligned}
(v|Aw) &= (v, bv + aw) \\
&= b\|v\|^2 + a(v|w).
\end{aligned}
$$

Ponieważ $A$ jest symetryczna, więc $(Av|w) = (v|Aw)$, czyli

$$0 < b\|v\|^2 = -b\|w\|^2 < 0.$$

Otrzymana sprzeczność pokazuje, że $b = 0$, czyli $\lambda \in \mathbb{R}$.

Niech $\lambda_1, \lambda_2 \in \mathbb{R}$ będą różnymi wartościami własnymi macierzy $A$, a $v_1$, $v_2 \in \mathbb{R}^n$ odpowiadającymi im wektorami własnymi. Wtedy

$$
\begin{aligned}
\lambda_1(v_1|v_2) &= (\lambda_1 \cdot v_1|v_2) \\
&= (Av_1|v_2) \\
&= (v_1|Av_2) \\
&= (v_1|\lambda_2 \cdot v_2) \\
&= \lambda_2(v_1|v_2).
\end{aligned}
$$

Ponieważ $\lambda_1 \neq \lambda_2$, więc $(v_1|v_2) = 0$. □

**Wniosek 13.10.** *Jeśli macierz symetryczna $A \in M_{n\times n}(\mathbb{R})$ ma różne wartości własne, to istnieje taka macierz ortogonalna $Q \in M_{n\times n}(\mathbb{C})$, że $Q^T AQ$ jest macierzą diagonalną.*

*Dowód.* Niech $x_i$ będzie wektorem własnym odpowiadającym wartości własnej $\lambda_i$. Wtedy dla $u_i = \frac{x_i}{\|x_i\|}$, wektory $u_1, \ldots, u_n$ tworzą układ ortonormalny. W szczególności, macierz $Q = \left[\begin{array}{c|c|c} u_1 & \ldots & u_n \end{array}\right]$ jest ortogonalna i $Q$ diagonalizuje $A$. □

Pokażemy, że każdą macierz symetryczną $A \in M_{n\times n}(\mathbb{R})$ da się zdiagonalizować przy pomocy macierzy ortogonalnej.

**Twierdzenie 13.5** (Symetryczna ortogonalnie podobna do diagonalnej)**.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie rzeczywistą macierzą symetryczną. Istnieje taka macierz ortogonalna $Q \in O(n)$, że $Q^T AQ$ jest diagonalna.*

*Dowód.* Zastosujemy indukcję względem $n$. Dla $n = 1$ teza zachodzi. Załóżmy, że teza zachodzi dla macierzy rozmiaru $k \times k$. Pokażemy, że zachodzi też dla macierzy rozmiaru $(k+1)\times(k+1)$. Niech $\lambda_1 \in \mathbb{R}$ będzie wartością własną $A$, a $w_1 \in \mathbb{R}^n$ odpowiadającym jej wektorem własnym o normie jeden. Wybieramy bazę ortonormalną $w_2, \ldots, w_{k+1}$ dla $(\mathrm{span}\{w_1\})^\perp$. Wtedy $V = \mathrm{span}\{w_2, \ldots, w_{k+1}\} = (\mathrm{span}\{w_1\})^\perp$ jest niezmiennicza dla $A$. Rzeczywiście, jeśli $v \in V$, to

$$(w_1|Av) = (Aw_1|v) = \lambda(w_1|v) = 0,$$

czyli $Av \in V$. Ponadto w restrykcji do $V$ mamy

$$(Ax|y) = (x|Ay), \quad x, y \in V.$$

Niech $W = \left[\begin{array}{c|c|c} w_1 & \ldots & w_{k+1} \end{array}\right] \in M_{(k+1)\times(k+1)}(\mathbb{R})$. Pierwszą kolumną macierzy $W^T AW$ jest wektor $W^T AWe_1$. Mamy

$$
\begin{aligned}
W^T AWe_1 &= W^T Aw_1 \\
&= W^T(\lambda_1 \cdot w_1) \\
&= \lambda_1 \cdot W^T w_1 \\
&= \lambda_1 \cdot e_1.
\end{aligned}
$$

Stąd, $W^T AW$ ma postać

$$
\left[\begin{array}{c|cccc}
\lambda_1 & 0 & 0 & \ldots & 0 \\
\hline
0 & & & & \\
\vdots & & & M & \\
0 & & & &
\end{array}\right],
$$

gdzie $M$ ma rozmiar $k \times k$ i jest symetryczna. Z założenia indukcyjnego istnieje taka macierz ortogonalna $V_1 \in M_{k\times k}(\mathbb{R})$, ze $V_1^T M V_1 = D_1$ jest diagonalna. Definiujemy

$$V = \left[\begin{array}{c|cccc} 1 & 0 & 0 & \dots & 0 \\ \hline 0 & & & & \\ \vdots & & & V_1 & \\ 0 & & & & \end{array}\right].$$

Wtedy $V$ jest ortogonalna oraz

$$\begin{aligned} V^T W^T A W V &= \left[\begin{array}{c|cccc} \lambda_1 & 0 & 0 & \dots & 0 \\ \hline 0 & & & & \\ \vdots & & & V_1^T M V_1 & \\ 0 & & & & \end{array}\right] \\ &= \left[\begin{array}{c|cccc} \lambda_1 & 0 & 0 & \dots & 0 \\ \hline 0 & & & & \\ \vdots & & & D_1 & \\ 0 & & & & \end{array}\right] \end{aligned}$$

jest macierzą diagonalną. Definiujemy $Q = WV$. Wtedy $Q$ jest ortogonalna, bo

$$\begin{aligned} Q^T Q &= (WV)^T WV \\ &= V^T W^T W V \\ &= I. \end{aligned}$$

□

**Twierdzenie 13.6** (Spektralne dla rzeczywistych macierzy symetrycznych). *$A \in M_{n\times n}(\mathbb{R})$ jest symetryczna wtedy i tylko wtedy, gdy istnieje taka macierz ortogonalna $Q \in O(n)$, że $Q^T A Q$ jest diagonalna.*

*Dowód.* Jeśli $A$ jest symetryczna, to z twierdzenia 13.5 jest ortogonalnie diagonalizowalna.

Załóżmy teraz, że $Q^T A Q = D$ jest diagonalna dla pewnej macierzy ortogonalnej $Q$. Wtedy $A = QDQ^T$ jest symetryczna, bo

$$\begin{aligned} A^T &= (QDQ^T)^T \\ &= (Q^T)^T D^T Q^T \\ &= QDQ^T \\ &= A. \end{aligned}$$

□

**Wniosek 13.11.** *Jeśli $A \in M_{n\times n}(\mathbb{R})$ jest symetryczna, to istnieje baza* **ortonormalna** *$u_1, \dots, u_n$ dla $\mathbb{R}^n$ złożona z wektorów własnych $A$. Jeśli $\lambda_i$ są wartościami własnymi dla $u_i$, to*

$$A = \lambda_1 u_1 u_1^T + \ldots + \lambda_n u_n u_n^T.$$

*Dowód.* Niech macierz ortogonalna $Q = \left[\begin{array}{c|c|c} u_1 & \dots & u_n \end{array}\right] \in O(n)$ diagonalizuje $A$. Wystarczy pokazać, że $u_i$ są wektorami własnymi $A$. Ponieważ $Q^T A Q = D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ oraz $Q^T = Q^{-1}$, więc

$$Q^T A u_i = Q^T A Q e_i = \lambda_i \cdot e_i,$$

więc $A u_i = \lambda_i \cdot u_i$, czyli $u_i$ jest wektorem własnym dla $\lambda_i$. □

U Załóżmy, że macierz symetryczna $A \in M_{n\times n}(\mathbb{R})$ ma dodatnie wartości własne $\lambda_1, \ldots, \lambda_n$. Wiemy, że istnieje taka macierz ortogonalna $Q \in O(n)$, że

$$Q^T A Q = D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

Rozważmy macierz

$$\sqrt{D} = \operatorname{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}).$$

Wtedy

$$\begin{aligned} A &= QDQ^T \\ &= Q\sqrt{D}\sqrt{D}Q^T \\ &= \underbrace{Q\sqrt{D}Q^T}_{B} Q\sqrt{D}Q^T \\ &= B^2. \end{aligned}$$

Zauważmy, że macierz $B$ jest symetryczna i ma dodatnie wartości własne, bo jest ortogonalnie podobna do $\sqrt{D}$. Można pokazać, że istnieje dokładnie jedna taka macierz symetryczna $B$ o dodatnich wartościach własnych, że $B^2 = A$.
Zwróćmy uwagę, że dla macierzy

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

istnieje nieskończenie wiele takich macierzy symetrycznych $B \in M_{2\times 2}(\mathbb{R})$, że $B^2 = I$. Oprócz macierzy

$$\begin{bmatrix} 1 & 0 \\ 0 & \pm 1 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 \\ 0 & \pm 1 \end{bmatrix}$$

są nimi również macierze

$$\frac{1}{\sqrt{a^2+b^2}}\begin{bmatrix} a & b \\ b & -a \end{bmatrix}, \quad \frac{1}{\sqrt{a^2+b^2}}\begin{bmatrix} a & -b \\ -b & -a \end{bmatrix},$$
$$\frac{1}{\sqrt{a^2+b^2}}\begin{bmatrix} -a & b \\ b & a \end{bmatrix}, \quad \frac{1}{\sqrt{a^2+b^2}}\begin{bmatrix} -a & -b \\ -b & a \end{bmatrix},$$

gdzie $a, b > 0$. Jeśli $(a, b, c)$ są *trójkami pitagorejskimi*, czyli takimi liczbami naturalnymi, że $a^2+b^2 = c^2$, to powyższe macierze mają wyrazy wymierne. Trójek pitagorejskich jest nieskończenie wiele, bo jeśli $(a, b, c)$ jest taką trójką, to jest nią również trójka $(ad, bd, cd)$ dla dowolnej liczby naturalnej $d$.

## 13.4 Macierz Grama $A^T A$

**Lemat 13.2.** *Niech $A \in M_{m\times n}(\mathbb{R})$. Wtedy*

*(i)* $\operatorname{im} A^T A = \operatorname{im} A^T$ *oraz* $\ker A^T A = \ker A$,

*(ii)* $\operatorname{rank} A = \operatorname{rank} A^T A$,

*(iii) wartości własne $A^T A$ są nieujemne.*

*Dowód.* Z wniosku 13.9 dla dowolnej macierzy $A$ mamy rozkład na ortogonalną sumę prostą

$$\mathbb{R}^n = \ker A \oplus \operatorname{im} A^T, \quad \ker A = (\operatorname{im} A^T)^\perp. \tag{13.1}$$

W szczególności,

$$\mathbb{R}^n = \ker A^T A \oplus \operatorname{im} A^T A \tag{13.2}$$

bo $A^T A$ jest symetryczna. Zauważmy, że bezpośrednio z definicji wynika, że

$$\ker A \subset \ker A^T A, \quad \operatorname{im} A^T A \subset \operatorname{im} A^T.$$

Pokażemy, że $\ker A^T A \subset \ker A$. Jeśli $x \in \ker A^T A$, to

$$\begin{aligned} Ax &\in \ker A^T \cap \operatorname{im} A \\ &= (\operatorname{im} A)^\perp \cap \operatorname{im} A \\ &= \{0\}, \end{aligned}$$

czyli $x \in \ker A$.

Z udowodnionej równości $\ker A = \ker A^TA$ i rozkładów (13.1), (13.2) wynika, że $\dim \operatorname{im} A^T = \dim \operatorname{im} A^TA$. Ponieważ $\operatorname{im} A^TA \subset \operatorname{im} A^T$, więc są one równe.

Ponieważ $A^TA$ jest symetryczna, więc ma rzeczywiste wartości własne (twierdzenie 13.4). Pokażemy, że są one nieujemne. Jeśli $\lambda \in \mathbb{R}$ jest wartością własną $A^TA$ i $x \in \mathbb{R}^n$ jest jej wektorem własnym, to

$$\begin{aligned}\|Ax\|^2 &= (Ax|Ax)\\ &= (A^TAx|x)\\ &= \lambda(x|x)\\ &= \lambda\|x\|^2,\end{aligned}$$

więc $\lambda = \frac{\|Ax\|^2}{\|x\|^2} \geqslant 0$. □

<table><tr><td>

MACIERZ GRAMA I JEJ KOLUMNY

Niech $A = \left[\begin{array}{c|c|c} a_1 & \dots & a_k \end{array}\right] \in M_{n\times k}(\mathbb{R})$ dla pewnych wektorów $a_1, \dots, a_k \in \mathbb{R}^n$. Wtedy

$$A^TA = \begin{bmatrix} (a_1|a_1) & \dots & (a_1|a_k)\\ \vdots & \ddots & \vdots\\ (a_k|a_1) & \dots & (a_k|a_k)\end{bmatrix} \in M_{k\times k}(\mathbb{R}).$$

</td></tr></table>

**Lemat 13.3** (Macierz Grama a liniowa niezależność)**.** *Wektory $a_1, \dots, a_k$ są liniowo niezależne wtedy i tylko wtedy, gdy macierz Grama $A^TA$ dla $A = \left[\begin{array}{c|c|c} a_1 & \dots & a_k \end{array}\right]$ jest nieosobliwa.*

*Dowód.* Z lematu 13.2 wiemy, że $\ker A = \ker A^TA$. Macierz $A^TA$ jest więc nieosobliwa wtedy i tylko wtedy, gdy $\ker A = \{0\}$. □

**Definicja 13.7.** **Wielościan rozpięty na wektorach**

Dla wektorów $a_1, \dots, a_k \in \mathbb{R}^k$ definiujemy wielościan na nich rozpięty przez

$$P(a_1, \dots, a_k) = \{x_1 \cdot a_1 + \dots + x_k \cdot a_k : 0 \leqslant x_i \leqslant 1,\ i = 1, \dots, k\}$$

Definiujemy jego objętość rekurencyjnie:

$$\operatorname{vol} P(a) = \sqrt{(a|a)} = \|a\|,$$

$$\operatorname{vol} P(a_1, \dots, a_k, a_{k+1}) = \operatorname{vol} P(a_1, \dots, a_k)\|a_{k+1} - p_k\|,$$

gdzie $p_k$ jest rzutem ortogonalnym $a_{k+1}$ na $\operatorname{span}\{a_1, \dots, a_k\}$.

**Lemat 13.4** (Macierz Grama i objętość)**.** *Dla $A = \left[\begin{array}{c|c|c} a_1 & \dots & a_k \end{array}\right]$ mamy*

$$\operatorname{vol} P(a_1, \dots, a_k) = \sqrt{\det A^TA}.$$

*Dowód.* Możemy założyć, że kolumny macierzy $A$ są liniowo niezależne, bo w przeciwnym razie obie strony są równe 0. Stosując algorytm Grama–Schmidta otrzymujemy bazę ortonormalną $u_1, \dots, u_k$ dla $\operatorname{im} A = \operatorname{span}\{a_1, \dots, a_k\}$ daną przez

$$\begin{aligned}u_1 &= \frac{a_1}{\|a_1\|}\\ u_2 &= \frac{a_2 - (a_2|u_1)u_1}{\|a_2 - (a_2|u_1)u_1\|}\\ u_3 &= \frac{a_3 - (a_3|u_1)u_1 - (a_3|u_2)u_2}{\|a_3 - (a_3|u_1)u_1 - (a_3|u_2)u_2\|}\\ &\cdots\cdots\cdots\\ u_k &= \frac{a_k - (a_k|u_1)u_1 - \dots - (a_k|u_{k-1})u_{k-1}}{\|a_k - (a_k|u_1)u_1 - \dots - (a_k|u_{k-1})u_{k-1}\|}.\end{aligned}$$

Rozważmy rozkład $A = QR$ otrzymany w wyniku procedury Grama–Schmidta, czyli $Q = \left[\begin{array}{c|c|c} u_1 & \ldots & u_k \end{array}\right]$ oraz

$$R = \begin{bmatrix} \|a_1\| & (a_2|u_1) & \ldots & (a_{k-1}|u_1) & (a_k|u_1) \\ 0 & \|a_2 - p_1\| & & & \vdots \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \|a_{k-1} - p_{k-2}\| & (a_k|u_{k-1}) \\ 0 & 0 & \ldots & 0 & \|a_k - p_{k-1}\| \end{bmatrix} \in M_{k\times k}(\mathbb{R}).$$

Zauważmy, że

$$\operatorname{vol} P(a_1, \ldots, a_k) = \det R.$$

Ponadto $A^T A = R^T R$, bo $Q^T Q = I$, gdyż kolumny macierzy $Q$ są ortonormalne. Stąd

$$\det A^T A = (\det R)^2 = (\operatorname{vol} P(a_1, \ldots, a_k))^2.$$

□

**Lemat 13.5.** *Niech $A \in M_{n\times n}(\mathbb{R})$ oraz $A^T A = AA^T$. Wtedy*

*(1) $\|Av\| = \|A^T v\|$ dla każdego $v \in \mathbb{R}^n$,*

*(2) $(I - A)^T (I - A) = (I - A)(I - A)^T$.*

*Dowód.* Dla $v \in \mathbb{R}^n$ mamy

$$\begin{aligned} \|Av\|^2 &= (Av|Av) \\ &= (v|A^T Av) \\ &= (v|AA^T v) \\ &= (A^T v|A^T v) \\ &= \|A^T v\|^2. \end{aligned}$$

Ponadto

$$\begin{aligned} (I - A)^T (I - A) &= (I - A^T)(I - A) \\ &= I - A - A^T + A^T A \\ &= I - A - A^T + AA^T \\ &= (I - A)(I - A)^T. \end{aligned}$$

□

## 13.5 Rzutowania

Niech $V \subset \mathbb{R}^n$ będzie podprzestrzenią wektorową. Wtedy z wniosku 13.9 mamy rozkład na ortogonalną sumę prostą

$$\mathbb{R}^n = V^\perp \oplus V, \quad V^\perp \perp V.$$

Jeśli $u_1, \ldots, u_k \in V$ jest bazą ortonormalną dla $V$, to rzutowanie ortogonalne $P : \mathbb{R}^n \to \mathbb{R}^n$ na podprzestrzeń $V$ jest dane przez

$$Pw = (w|u_1) \cdot u_1 + \ldots + (w|u_k) \cdot u_k, \quad w \in \mathbb{R}^n.$$

Z lematu 13.1 wiemy, że

$$P = UU^T,$$

gdzie $U = [u_1|\dots|u_k] \in M_{n\times k}(\mathbb{R})$ jest macierzą której kolumnami są wektory z bazy ortonormalnej $u_1,\dots,u_k$. Bezpośrednio z definicji macierzy $U$ wynika, że

$$\operatorname{im} U = \operatorname{span}\{u_1,\dots,u_k\} = V.$$

Ponadto z twierdzenia 13.2 wynika, że

$$\ker U^T = (\operatorname{im} U)^\perp = V^\perp.$$

Z lematu 13.2 otrzymujemy, że

$$\ker P = \ker UU^T = \ker U^T = V^\perp, \quad \operatorname{im} P = \operatorname{im} UU^T = \operatorname{im} U = V.$$

Zauważmy, że $P$ ma własności:

(1) $P^2 = P$,
bo $U^TU = I \in M_{k\times k}(\mathbb{R})$, więc

$$P^2 = U\underbrace{U^TU}_{=I}U^T = UU^T = P.$$

(2) $(P(v)|w) = (v|P(w))$ dla $v,w \in \mathbb{R}^n$,
bo $UU^T$ jest symetryczna.

(3) $v - P(v) \in V^\perp$.

(4) $\ker P = V^\perp$ oraz $\operatorname{im} P = V$,

(5) $\mathbb{R}^n = \ker P \oplus \operatorname{im} P$ jest ortogonalną sumą prostą,

(6) jeśli $w = u + v \in \mathbb{R}^n = V^\perp \oplus V$ jest jednoznacznym rozkładem $w$ w sumie prostej, to

$$P(w) = v.$$

(7) Dla dowolnego $v \in \mathbb{R}^n$ mamy

$$\|w - v\| > \|P(v) - v\|, \quad w \in V \setminus \{P(v)\},$$

czyli rzut prostopadły $p = P(v)$ jest najbliższym wektora $v$ wektorem z podprzestrzeni $V$. Rzeczywiście,

$$\begin{aligned}\|w - v\|^2 &= \|(w - P(v)) + (P(v) - v)\|^2\\ &= \|w - P(v)\|^2 + \|P(v) - v\|^2 + 2(\underbrace{w - P(v)}_{\in V}\,|\,\underbrace{v - P(v)}_{\in V^\perp})\\ &= \underbrace{\|w - P(v)\|^2}_{>0} + \|P(v) - v\|^2\\ &> \|P(v) - v\|^2.\end{aligned}$$

Przyglądniemy się bliżej warunkom (1) i (6). Zwróćmy uwagę, że warunek (1) ma sens dla dowolnej macierzy kwadratowej. Warto więc przyglądnąć się takim macierzom $P \in M_{n\times n}(\mathbb{R})$, że $P^2 = P$. Z drugiej strony w warunku (6) sumę prostą $\mathbb{R}^n = V^\perp \oplus V$ możemy zastąpić przez dowolną sumę prostą $\mathbb{R}^n = U \oplus V$.

**Definicja 13.8.** Powiemy, że odwzorowanie $P : \mathbb{R}^n \to \mathbb{R}^n$ jest *rzutowaniem* w sumie prostej $\mathbb{R}^n = U \oplus V$ ($U, V \subset \mathbb{R}^n$ podprzestrzenie wektorowe) na $V$, jeżeli

$$P(w) = v,$$

gdzie $w = u + v$ jest jednoznacznym rozkładem wektora $w \in \mathbb{R}^n$ w sumie prostej $\mathbb{R}^n = U \oplus V$. Wtedy $I - P$ jest rzutowaniem w sumie prostej $\mathbb{R}^n = U \oplus V$ na $U$.

Rzutowanie $P$ nazywamy *ortogonalnym*, gdy $U = V^\perp$.

**Wniosek 13.12.** *Jeśli $P : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ jest rzutowaniem w sumie prostej $\mathbb{R}^n = U \oplus V$ na $V$, to $P$ jest odwzorowaniem liniowym.*

*Dowód.* Niech $w = u + v$ i $w_1 = u_1 + v_1$. Wtedy

$$w + w_1 = \underbrace{(u + u_1)}_{\in U} + \underbrace{(v + v_1)}_{\in V},$$

więc

$$\begin{aligned} P(w + w_1) &= P((u + u_1) + (v + v_1)) \\ &= v + v_1 \\ &= P(w) + P(w_1). \end{aligned}$$

Dla $a \in \mathbb{R}$ mamy $a \cdot w = a \cdot u + a \cdot v$, więc

$$\begin{aligned} P(a \cdot w) &= P(a \cdot u + a \cdot v) \\ &= a \cdot v \\ &= a \cdot P(w). \end{aligned}$$

□

**Wniosek 13.13.** *Jeśli $P : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ jest rzutowaniem w sumie prostej $\mathbb{R}^n = U \oplus V$ na $V$ oraz $A \in M_{n\times n}(\mathbb{R})$ jest macierzą $P$ w bazie standardowej, to*

*(i)* $U = \ker A$ *oraz* $V = \operatorname{im} A$, *czyli*

$$\mathbb{R}^n = \ker A \oplus \operatorname{im} A,$$

*(ii)* $A(v) = v$ *dla* $v \in \operatorname{im} A$,

*(iii)* $A^2 = A$.

*Dowód.* Z definicji mamy $Aw = A(u+v) = v$, gdzie $w = u+v \in \mathbb{R}^n = U\oplus V$ jest jednoznacznym rozkładem $w$ w sumie prostej. Stąd $Aw = 0$ wtedy i tylko wtedy, gdy $v = 0$ czyli $w = u \in U$. Oznacza to, że $\ker A = U$. Ponadto $Aw = v \in V$, czyli $\operatorname{im} A \subset V$. Dla $v \in V$ mamy

$$\operatorname{im} A \ni A(v) = A(\underbrace{0}_{\in U} + v) = v.$$

W szczególności, $V \subset \operatorname{im} A$. W efekcie $V = \operatorname{im} A$.

Pokażemy, że $A^2 = A$. Dla $w \in \mathbb{R}^n$ mamy $A(w) \in \operatorname{im} A$, więc $A(A(w)) = A(w)$. □

**Definicja 13.9.** Macierz idempotentna

Powiemy, że macierz $A \in M_{n\times n}(\mathbb{F})$ jest *idempotentna*, jeśli

$$A^2 = A.$$

**Lemat 13.6.** *Niech $A \in M_{n\times n}(\mathbb{R})$. Następujące warunki są równoważne*

*(1)* $\mathbb{R}^n = \ker A \oplus \operatorname{im} A$ *oraz $A$ jest rzutowaniem w tej sumie prostej na* $\operatorname{im} A$,

*(2)* $A^2 = A$.

*Dowód.* Z wniosku 13.13 wynika, że (1) implikuje (2).

Pokażemy że (2) implikuje (1). Uzasadnimy najpierw, że

$$\mathbb{R}^n = \ker A \oplus \operatorname{im} A$$

Niech $w \in \mathbb{R}^n$. Wtedy $Aw \in \operatorname{im} A$. Ponieważ

$$w = (w - Aw) + Aw,$$

więc wystarczy pokazać, że $w - Aw \in \ker A$. Z warunku $A^2 = A$ mamy

$$\begin{aligned} A(w - Aw) &= Aw - A^2 w \\ &= Aw - Aw \\ &= 0 \end{aligned}$$

Niech $w \in \ker A \cap \operatorname{im} A$. Wtedy $Aw = 0$ i $w = Az$ dla pewnego $z \in \mathbb{R}^n$. Stąd $0 = Aw = AAz = A^2 z = Az = w$.

Pokażemy, że $A$ jest rzutowaniem w sumie prostej $\mathbb{R}^n = \ker A \oplus \operatorname{im} A$. Niech

$$w = u + v \in \mathbb{R}^n, \quad u \in \ker A, \quad v = Az \in \operatorname{im} A.$$

Wystarczy uzasadnić, że $Aw = v$. Mamy

$$\begin{aligned} A(w) &= A(u) + A(v) = A(v) \\ &= A^2 z = Az \\ &= v. \end{aligned}$$

□

**Lemat 13.7** (Własności macierzy idempotentnych)**.** *Załóżmy, że macierz $A \in M_{n\times n}(\mathbb{R})$ jest idempotentna. Wtedy*

*(1) jeśli $A \in M_{n\times n}(\mathbb{R})$ jest nieosobliwa, to $A = I$,*

*(2) $I - A$ jest idempotentna,*

*(3) jeśli 1 nie jest wartością własną $A$, to $A = 0$,*

*(4) jeśli $\lambda$ jest wartością własną $A$, to $\lambda = 0$ lub $\lambda = 1$,*

*(5) $\mathbb{R}^n = \ker A \oplus \operatorname{im} A$,*

*(6) $\operatorname{im} A = \ker(A - I)$, więc*

$$\mathbb{R}^n = \ker(A - I) \oplus \ker A,$$

*(7) $A$ w pewnej bazie postać diagonalną*

$$\operatorname{diag}(1, \ldots, 1, 0, \ldots, 0) = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right], \quad r = \operatorname{rank} A = \dim \operatorname{im} A.$$

*Dowód.* (1) Załóżmy, że $A$ jest nieosobliwa i $A^2 = A$. Wtedy

$$\begin{aligned} I &= A^{-1} A \\ &= A^{-1} A^2 \\ &= A. \end{aligned}$$

(2) Jeśli $A^2 = A$, to

$$\begin{aligned} (I - A)^2 &= I - 2A + A^2 \\ &= I - 2A + A \\ &= I - A. \end{aligned}$$

(3) Jeśli 1 nie jest wartością własną $A$, to $I - A$ jest nieosobliwa, więc

$$I - A = I,$$

czyli $A = 0$.

(4) Niech $v \neq 0$ będzie wektorem własnym dla $\lambda$. Wtedy

$$\begin{aligned}\lambda v = Av &= A^2 v \\ &= \lambda^2 v,\end{aligned}$$

czyli $\lambda^2 = \lambda$.

(5) Wynika z lematu 13.6.

(6) Pokażemy, że $\operatorname{im} A = \ker(A - I)$.

Ponieważ $A^2 - A = 0$, więc $(A - I)A = 0$, czyli $\operatorname{im} A \subset \ker(A - I)$. Jeśli $v \in \ker(A - I)$, to $Av = v$, czyli $v \in \operatorname{im} A$. Stąd $\operatorname{im} A = \ker(A - I)$.

(7) Z punktów (5) i (6) wynika, że mamy rozkład na podprzestrzenie niezmiennicze

$$\mathbb{R}^n = \ker(A - I) \oplus \ker A,$$

czyli $A$ ma bazę wektorów własnych dla wartości własnych 1 i 0. Wynika stąd teza o postaci diagonalnej $A$. □

**Definicja 13.10.** **Inwolucja**

Macierz $A \in M_{n\times n}(\mathbb{R})$ nazywamy *inwolucją*, gdy $A^2 = I$.

**Wniosek 13.14.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie inwolucją. Wtedy $A$ ma w pewnej bazie postać diagonalną*

$$\operatorname{diag}(1, \ldots, 1, -1, \ldots, -1).$$

*Dowód.* Zauważmy, że macierz $P \in M_{n\times n}(\mathbb{R})$ jest idempotentna wtedy i tylko wtedy, gdy $A = 2P - I$ jest inwolucją. Rzeczywiście,

$$\begin{aligned}A^2 &= (2P - I)^2 \\ &= 4P^2 - 4P + I \\ &= 4(P^2 - P) + I\end{aligned}$$

więc równość $I = A^2$ zachodzi wtedy i tylko wtedy, gdy $P^2 = P$. Stąd,

$$P = \frac{A + I}{2},$$

jest rzutowaniem. Z lematu 13.7 wiemy, że rzutowanie $P$ ma w pewnej bazie macierz $\operatorname{diag}(1, \ldots, 1, 0, \ldots, 0)$. W tej samej bazie $A = 2P - I$ ma macierz

$$\operatorname{diag}(1, \ldots, 1, -1, \ldots, -1).$$

□

U

Wniosek 13.14 możemy również udowodnić bardziej bezpośrednio, korzystając z rozkładu Jordana. Niech $\lambda \in \mathbb{C}$ będzie wartością inwolucji $A$ z wektorem własnym $v \in \mathbb{C}^n$. Wtedy

$$v = A^2 v = \lambda^2 v,$$

czyli $\lambda^2 = 1$, więc $\lambda = \pm 1$. Z teorii Jordana wynika, że istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{R})$, że

$$S^{-1}AS = \Lambda = \operatorname{diag}(\Lambda_1, \ldots, \Lambda_k),$$

oraz klatki Jordana $\Lambda_i$ są postaci

$$\begin{bmatrix} -1 & 1 & 0 & \ldots & 0 \\ 0 & -1 & 1 & \ldots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & -1 & 1 \\ 0 & \ldots & 0 & 0 & -1 \end{bmatrix} \quad \text{lub} \quad \begin{bmatrix} 1 & 1 & 0 & \ldots & 0 \\ 0 & 1 & 1 & \ldots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & 1 & 1 \\ 0 & \ldots & 0 & 0 & 1 \end{bmatrix},$$

czyli

$$\Lambda_i = \pm I + N,$$

gdzie

$$N = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 1 \\ 0 & \dots & 0 & 0 & 0 \end{bmatrix}.$$

Wystarczy pokazać, że wszystkie bloki Jordana są jednowymiarowe, czyli $\Lambda_i = [\pm 1]$. Ponieważ

$$\begin{aligned} \Lambda^2 &= (S^{-1}AS)(S^{-1}AS) \\ &= S^{-1}AS \\ &= I. \end{aligned}$$

Z postaci $\Lambda$ wynika, że również $\Lambda_i^2 = I$ dla każdej klatki Jordana. Przypuśćmy, że pewna klatka Jordana $\Lambda_i \in M_{k\times k}$ z $k > 1$. Wtedy $\Lambda_i = \pm I + N$ oraz $N \neq 0$. Wtedy

$$\begin{aligned} I = \Lambda_i^2 &= (\pm I + N)^2 \\ &= I \pm 2N + N^2, \end{aligned}$$

czyli $2N + N^2 = 0$, co jak łatwo sprawdzić prowadzi do sprzeczności.

**Lemat 13.8.** *Niech $A \in M_{n\times n}(\mathbb{R})$ oraz $A^2 = -I$. Wtedy $n$ jest parzyste oraz w pewnej bazie macierz $A$ ma postać*

$$\operatorname{diag}(J, \dots, J),$$

*gdzie $J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ jest standardową macierzą symplektyczną.*

*Dowód.* Ponieważ $A^2 = -I$, więc $(\det A)^2 = (-1)^n$, więc $n$ jest parzyste. Niech $\lambda \in \mathbb{C}$ będzie wartością własną $A$ z wektorem własnym $v \in \mathbb{C}^n$. Wtedy

$$-v = A^2 v = \lambda^2 v,$$

więc $\lambda^2 = -1$. Stąd, wartościami własnymi $A$ są $i$, $-i$ oraz mają taką samą krotność algebraiczną, więc $n$ jest parzyste. Z twierdzenia 12.6 wynika, że

$$\Lambda = S^{-1}AS = \begin{bmatrix} \boxed{\Lambda_1} & & & \\ & \boxed{\Lambda_2} & & \\ & & \ddots & \\ & & & \boxed{\Lambda_k} \end{bmatrix},$$

gdzie każdy blok $\Lambda_i$ jest postaci

$$\Lambda_i = \begin{bmatrix} J & I_2 & 0 & \dots & 0 \\ 0 & J & I_2 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & J & I_2 \\ 0 & \dots & 0 & 0 & J \end{bmatrix}.$$

Wystarczy pokazać, że $\Lambda_i = J$ dla $i = 1, \dots, k$. Zauważmy, że

$$\begin{aligned} \Lambda^2 &= (S^{-1}AS)(S^{-1}AS) \\ &= S^{-1}A^2S \\ &= -I. \end{aligned}$$

Z postaci $\Lambda$ wynika, że również $\Lambda_i^2 = -I$. Przypuśćmy, że $\Lambda_i \neq J$ dla pewnego $i$. Dla ustalenia uwagi zakładamy, że

$$\Lambda_i = \begin{bmatrix} J & I_2 \\ 0 & J \end{bmatrix} = \operatorname{diag}(J, J) + \underbrace{\begin{bmatrix} 0 & I_2 \\ 0 & 0 \end{bmatrix}}_{=N}.$$

Dowód w przypadku ogólnym jest analogiczny. Ponieważ $\operatorname{diag}(J, J)$ i $N$ komutują oraz $N^2 = 0$ $J^2 = -I_2$, więc

$$\begin{aligned} -I = \Lambda_i^2 &= (\operatorname{diag}(J, J) + N)^2 \\ &= \operatorname{diag}(J^2, J^2) + 2\operatorname{diag}(J, J)N + N^2 \\ &= \operatorname{diag}(-I_2, -I_2) + 2 \begin{bmatrix} 0 & J \\ 0 & 0 \end{bmatrix} \\ &= -I + 2 \begin{bmatrix} 0 & J \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

Stąd $2 \begin{bmatrix} 0 & J \\ 0 & 0 \end{bmatrix}$ jest macierzą zerową, co prowadzi do sprzeczności. □

**Przykład 13.6.** Rozważmy macierz

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}.$$

Wtedy $A^2 = A$, czyli $A$ jest rzutowaniem. Nie jest to rzut ortogonalny, bo

$$\operatorname{im} A = \operatorname{span}\{e_2\}, \quad \ker A = \operatorname{span}\{[1, -1]^T\}.$$

**U** Niech $A \in M_{n\times n}(\mathbb{R})$. Jeśli dla dowolnych $x, y \in \mathbb{R}^n$ zachodzi równość

$$(Ax|y) = 0,$$

to $A = 0$. Rzeczywiście, dla $A = [a_1, \dots, a_n] = [a_{ij}]$ mamy

$$(Ae_i|e_j) = (a_i|e_j) = a_{ij}.$$

**Wniosek 13.15** (Kiedy rzut jest ortogonalny?)**.** *Załóżmy, że $A^2 = A \in M_{n\times n}(\mathbb{R})$. Następujące warunki są równoważne*

*(1) $A$ jest rzutem ortogonalnym tzn.* $\ker A = (\operatorname{im} A)^\perp$,

*(2)* $A = A^T A$,

*(3)* $A = A^T$,

*(4)* $A^T A = AA^T$.

*Dowód.* (1) implikuje (2). Zakładamy, że $A$ jest rzutem ortogonalnym. Wystarczy pokazać, że

$$((A - A^T A)x|y) = 0, \quad \text{dla} \quad x, y \in \mathbb{R}.$$

Mamy

$$\begin{aligned} ((A - A^T A)x|y) &= ((I - A)^T Ax|y) \\ &= (Ax|(I - A)y) = 0, \end{aligned}$$

bo $Ax \in \operatorname{im} A$ oraz $(I - A)y \in \ker A = (\operatorname{im} A)^\perp$.

Oczywiście warunek (2) implikuje (3) oraz (3) implikuje (4). Pokażemy, że z warunku (4) wynika (1). Jeśli $A^T A = AA^T$, to z lematu 13.2 i twierdzenia 13.2 otrzymujemy, że

$$\begin{aligned} \ker A &= \ker A^T A \\ &= \ker AA^T \\ &= \ker A^T \\ &= (\operatorname{im} A)^{\perp}. \end{aligned}$$

□

**Twierdzenie 13.7** (Rzut ortogonalny na obraz macierzy)**.** *Rozważmy macierz $B \in M_{m\times n}(\mathbb{R})$. Jeśli* $\operatorname{rank}(B) = n$ *(równoważnie* $\ker B = \{0\}$*), to rzut ortogonalny $P$ na* $\operatorname{im} B \subset \mathbb{R}^m$ *jest dany przez*

$$P = B(B^T B)^{-1} B^T.$$

*Dowód.* Z definicji rzutu ortogonalnego mamy

$$v - P(v) \in (\operatorname{im} B)^{\perp} = \ker B^T,$$

czyli

$$0 = B^T(v - P(v)) \Leftrightarrow B^T v = B^T P(v).$$

Ponieważ $P(v) \in \operatorname{im} B$ i $\ker B = \{0\}$, więc istnieje jednoznacznie wyznaczony taki wektor $x \in \mathbb{R}^n$, że $P(v) = Bx$. Stąd

$$B^T Bx = B^T v.$$

Z lematu 13.3 wynika, że $B^T B$ jest nieosobliwa, więc

$$x = (B^T B)^{-1} B^T v,$$

czyli

$$P(v) = Bx = B(B^T B)^{-1} B^T v.$$

□

**Przykład 13.7.** Rozważmy macierz

$$B = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}.$$

Obrazem $\operatorname{im} B$ jest prosta w $\mathbb{R}^4$ generowana przez wektor $[1, 2, 0, 1]^T$. Ponieważ $B^T B = [6]$, więc

$$B(B^T B)^{-1} B^T = \frac{1}{6} \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & 1 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 1 & 2 & 0 & 1 \\ 2 & 4 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix}$$

**Przykład 13.8.** Rozważmy macierz $A \in M_{m\times n}(\mathbb{R})$ i załóżmy, że $\operatorname{rank} A = r$. Niech $B \in M_{m\times r}(\mathbb{R})$ będzie macierzą, której kolumny są bazą dla $\operatorname{im} A$. Wtedy

$$B(B^T B)^{-1} B^T$$

jest rzutem ortogonalnym na $\operatorname{im} A$, a

$$I - B(B^T B)^{-1} B^T$$

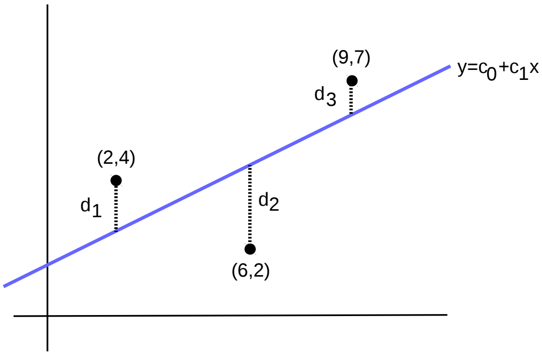
jest rzutem ortogonalnym na $\ker A^T = (\operatorname{im} A)^{\perp}$. Analogicznie, jeśli $N \in M_{n\times(n-r)}(\mathbb{R})$ jest macierzą, której kolumny tworzą bazę dla $\ker A$, to

$$N(N^T N)^{-1} N^T$$

jest rzutem ortogonalnym na $\ker A$, a

$$I - N(N^T N)^{-1} N^T$$

jest rzutem ortogonalnym na $\operatorname{im} A^T = (\ker A)^{\perp}$.




Rysunek 13.4: Metoda najmniejszych kwadratów. Szukamy takiej prostej $y = c_0 + c_1x$, aby suma $d_1^2 + d_2^2 + d_3^2$ była jak najmniejsza.

## 13.6 Metoda najmniejszych kwadratów

Przedstawimy najpierw pewną motywację.

**Przykład 13.9.** Przypuśćmy, że przeprowadzamy pewne doświadczenie (obserwacje) i w efekcie otrzymaliśmy tabelę wyników.

| | | | | |
|---|---|---|---|---|
| $x$ | $x_1$ | $x_2$ | $\dots$ | $x_m$ |
| $y$ | $y_1$ | $y_2$ | $\dots$ | $y_m$ |

Jeśli wyniki składają się z $n+1$ punktów, to zawsze można znaleźć wielomian stopnia co najwyżej $n$ przechodzący przez wszystkie punkty.

Przypuśćmy, że chcemy znaleźć prostą o równaniu

$$y = c_0 + c_1x$$

przechodzącą przez wszystkie punkty $(x_1, y_1), \dots, (x_m, y_m)$, tzn.

$$y_i = c_0 + c_1x_i, \quad i = 1, \dots, m.$$

Otrzymaliśmy układ równań

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad Ac = b.$$

Jeśli $m > 2$ (punktów jest za dużo), to otrzymany układ będzie zazwyczaj sprzeczny. Trudno oczekiwać, aby punkty $(x_1, y_1), \dots, (x_m, y_m)$ leżały na jednej prostej. Możemy się zapytać dla jakiego wektora $c = [c_0, c_1]^T$ kwadrat odległości

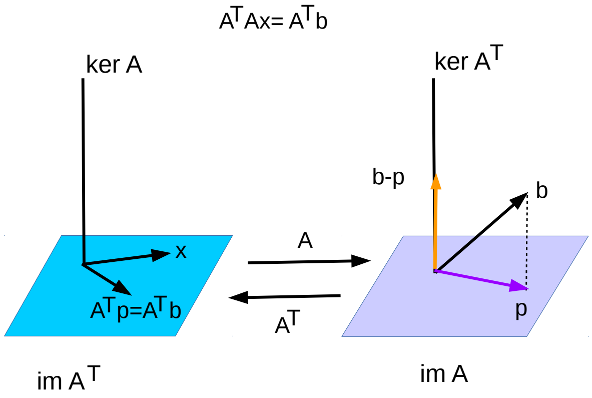
$$\|b - Ac\|^2$$

jest najmniejszy. W naszym przypadku mamy

$$\begin{aligned} \|b - Ac\|^2 &= (y_1 - (c_0 + x_1c_1))^2 + \dots + ((y_m - (c_0 + c_1x_m))^2 \\ &= d_1^2 + \dots + d_m^2. \end{aligned}$$

Rozważmy równanie liniowe $Ax = b$, gdzie $A \in M_{m\times n}(\mathbb{R})$ i $b \in \mathbb{R}^m$. Jeśli $m > n$ tzn. mamy więcej równań niż niewiadomych, to układ $Ax = b$ jest często sprzeczny. Oznacza to, że $b \notin \operatorname{im}(A)$, czyli $b$ nie leży w podprzestrzeni $\mathbb{R}^m$ rozpiętej przez kolumny macierzy $A$. Możemy zapytać wtedy, dla jakiego wektora $x \in \mathbb{R}^n$ wektor $Ax$ jest najbliższy $b$. Naturalne jest rozważenie różnicy

$$r(x) = b - Ax$$




Rysunek 13.5: Metoda najmniejszych kwadratów

Szukamy więc wektora $x \in \mathbb{R}^n$, dla którego norma

$$\|r(x)\| = \|b - Ax\|$$

jest najmniejsza. Zamiast $\|r(x)\|$ wygodnie jest minimalizować $\|r(x)\|^2$.

W metodzie najmniejszych kwadratów szukamy takiego wektora $\hat{x} \in \mathbb{R}^n$, że $A\hat{x}$ jest najbliższym w $S = \operatorname{im} A$ wektorem do wektora $b$. Będzie tak, gdy

$$A\hat{x} = p,$$

gdzie $p$ jest rzutem ortogonalnym $b$ na $S = \operatorname{im} A$. Oznacza to, że musi zachodzić warunek

$$b - A\hat{x} = b - p \in (\operatorname{im} A)^{\perp}.$$

Wektor $\hat{x}$ jest więc rozwiązaniem $Ax = b$ *w sensie najmniejszych kwadratów* wtedy i tylko wtedy, gdy

$$b - A\hat{x} \in (\operatorname{im} A)^{\perp} = \ker(A^T).$$

Równoważnie,

$$0 = A^T(b - A\hat{x}).$$

Musimy więc rozwiązać równanie

$$A^T Ax = A^T b.$$

**U** Niech $A \in M_{m\times n}(\mathbb{R})$. Przypomnijmy, że (wniosek 13.2)

$$\operatorname{im} A^T A = \operatorname{im} A^T,$$

więc równanie

$$A^T Ax = A^T b,$$

ma zawsze rozwiązanie nawet jeśli zbiór rozwiązań równania $Ax = b$ jest pusty.

**Wniosek 13.16.** *Niech $A \in M_{m\times n}(\mathbb{R})$ i $b \in \mathbb{R}^m$. Jeśli* $\operatorname{rank}(A) = n$, *to równanie*

$$A^T Ax = A^T b$$

*ma jedyne rozwiązanie $\hat{x}$ dane przez*

$$\hat{x} = (A^T A)^{-1} A^T b.$$

*W szczególności,*

$$p = A\hat{x} = A(A^T A)^{-1} A^T b$$

*Dowód.* Wystarczy pokazać, że macierz $A^T A \in M_{n\times n}(\mathbb{R})$ jest nieosobliwa. Z formuły wymiaru wynika, że $\ker A = \{0\}$, więc również $\ker A^T A = \{0\}$. $\square$

**Przykład 13.10.** Rozważmy wyniki eksperymentu dane w tabelce

| $x$ | 0 | 3 | 6 |
|---|---|---|---|
| $y$ | 1 | 4 | 5 |

Mamy tu

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 3 \\ 1 & 6 \end{bmatrix}, \quad c = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 4 \\ 5 \end{bmatrix}.$$

Interesuje nas rozwiązanie równania

$$A^T A c = A^T b,$$

czyli

$$\begin{bmatrix} 3 & 9 \\ 9 & 45 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 10 \\ 42 \end{bmatrix}.$$

Rozwiązaniem jest $[4/3, 2/3]^T$. Stąd najlepszą (w sensie najmniejszych kwadratów) przybliżeniem postaci $y = c_0 + c_1 x$ jest

$$y = \frac{4}{3} + \frac{2}{3}x.$$

TEST $\longrightarrow$ Oceń, które ze zdań jest prawdziwe:

- Wektory ortonormalne $v_1, \ldots, v_k \in \mathbb{R}^n$ są liniowo niezależne.
- Każda niezerowa podprzestrzeń $V \subset \mathbb{R}^n$ ma bazę ortonormalną.
- Macierz ortogonalna jest nieosobliwa.
- Jeśli $\det A = 1$, to $A$ jest ortogonalna.
- Jeśli $A$ jest symetryczna i ortogonalna, to $A^2 = I$.
- Każda macierz ortogonalnie diagonalizowalna jest nieosobliwa.
- Jeśli $\mathbb{R}^n = \ker A \oplus \operatorname{im} A$, to $A \in M_{n\times n}(\mathbb{R})$ jest idempotentna.
- Jeśli $A \neq I$ jest idempotentna, to $\det A = 0$.
- Jeśli $A$ jest idempotentna, to $\operatorname{tr} A = \operatorname{rank} A$.
- Suma macierzy idempotentnych jest macierzą idempotentną.
- Iloczyn macierzy idempotentnych jest macierzą idempotentną.
- Jeśli $Q \in O(n)$ i $0 \neq t \in \mathbb{R}$, to $t \cdot Q \in O(n)$.
- Suma izometrii $\mathbb{R}^n$ jest izometrią $\mathbb{R}^n$.
- Jeśli $Q \in O(3)$, to $Q$ ma prostą niezmienniczą.
- Jeśli $Q \in O(3)$, to istnieje taki $v \in \mathbb{R}^3$, że $Av = v$ lub istnieje taki $v \in \mathbb{R}^3$, że $Av = -v$.

$\longleftarrow$

## 13.7 Izometrie przestrzeni euklidesowych $\mathbb{R}^n$

**Definicja 13.11.** Izometria $\mathbb{R}^n$

Odwzorowanie $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ nazywamy *izometrią*, jeśli dla dowolnych $x, y \in \mathbb{R}^n$ zachodzi równość

$$\|f(x) - f(y)\| = \|x - y\|,$$

czyli $f$ zachowuje odległość euklidesową.

**Wniosek 13.17.** *Jeśli $Q \in O(n)$, to $Q : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ jest izometrią.*

*Dowód.* Z twierdzenia 13.1 i liniowości $Q$ mamy

$$\|Q(x) - Q(y)\| = \|Q(x - y)\| = \|x - y\|, \quad x, y \in \mathbb{R}^n.$$

□

**Twierdzenie 13.8.** *Jeśli $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ jest taką izometrią, że $f(0) = 0$, to istnieje taka macierz ortogonalna $Q \in O(n)$, że $f = Q$. W szczególności $f$ jest izomorfizmem liniowym.*

*Dowód.* Niech $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ będzie izometrią oraz $f(0) = 0$. Dla każdego $x \in \mathbb{R}^n$ mamy

$$\begin{aligned}\|f(x)\| &= \|f(x) - f(0)\| \\ &= \|x - 0\| \\ &= \|x\|,\end{aligned}$$

czyli $f$ zachowuje normę. Dla dowolnych $x, y \in \mathbb{R}^n$ mamy

$$\begin{aligned}\|f(x) - f(y)\|^2 &= \|f(x)\|^2 - 2(f(x)|f(y)) + \|f(y)\|^2 \\ &= \|x\|^2 - 2(f(x)|f(y)) + \|y\|^2,\end{aligned}$$

oraz

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

Ponieważ $\|f(x) - f(y)\| = \|x - y\|$, więc

$$(f(x)|f(y)) = (x|y),$$

czyli $f$ zachowuje iloczyn skalarny.

Niech $Q = \left[\begin{array}{c|c|c} f(e_1) & \ldots & f(e_n) \end{array}\right]$. Ponieważ $f$ zachowuje iloczyn skalarny, więc

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

czyli $Q \in O(n)$ jest izometrią (wniosek 13.17). Pokażemy, że $f = Q$. Rozważmy odwzorowanie

$$g := Q^{-1} \circ f : \mathbb{R}^n \longrightarrow \mathbb{R}^n.$$

Zakończymy dowód, pokazując, że $g(x) = x$ dla każdego $x \in \mathbb{R}^n$. Ponieważ złożenie izometrii jest izometrię, więc $g$ jest izometrią oraz $g(0) = 0$. Ponadto

$$g(e_i) = e_i, \quad i = 1, \ldots, n.$$

Rzeczywiście, $g(e_i) = Q^{-1}f(e_i)$, czyli $Q(g(e_i)) = f(e_i)$. Z definicji $Q$ mamy $Q(e_i) = f(e_i)$, czyli $g(e_i) = e_i$, bo $Q$ jest izomorfizmem.

Ustalmy $x = x_1 \cdot e_1 + \ldots + x_n \cdot e_n \in \mathbb{R}^n$ i niech

$$g(x) = b_1 \cdot e_1 + \ldots + b_n \cdot e_n.$$

Ponieważ $g$ zachowuje iloczyn skalarny, więc

$$\begin{aligned}b_i &= (g(x)|e_i) \\ &= (g(x)|g(e_i)) \\ &= (x|e_i) \\ &= x_i,\end{aligned}$$

czyli $g(x) = x$. □

**Wniosek 13.18.** *Jeśli* $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *jes izometrią, to*

$$f(x) = f(0) + Qx, \quad x \in \mathbb{R}^n$$

*dla pewnej macierzy ortogonalnej* $Q \in O(n)$.

*Dowód.* Odwzorowanie $F(x) = f(x) - f(0)$ jest izometrią i $F(0) = 0$, więc teza wynika z twierdzenia 13.8. □

U Pojęcie izometrii można wprowadzić w dowolnej przestrzeni metrycznej $(X, d)$. Odwzorowanie $f : X \longrightarrow X$ jest izometrią, gdy

$$d(f(x), f(y)) = d(x, y), \quad x, y \in X.$$

Izometria jest odwzorowaniem różnowartościowym, ale nie musi być bijekcją. Przykładowo, izometria zbioru $\mathbb{N}$

$$f : \mathbb{N} \ni n \longmapsto n + 1 \in \mathbb{N}$$

z metryką $d(n, m) = |n - m|$ nie jest surjekcją, bo $1 \notin f(\mathbb{N})$.

## 13.8 Grupa ortogonalna $O(n)$ dla $n = 2, 3$

Przypomnijmy, że

$$O(n) = \{Q \in M_{n\times n}(\mathbb{R}) : Q^TQ = QQ^T = I\}, \quad SO(n) = \{Q \in O(n) : \det Q = 1\}.$$

**Wniosek 13.19.** $O(n)$ *i* $SO(n)$ *z działaniem mnożenia macierzy tworzą grupy.*

*Dowód.* Zauważmy, że jeśli $A, B \in O(n)$ $(SO(n))$, to

- $AB \in O(n)$ $(SO(n))$,
- $A^{-1} \in O(n)$ $(SO(n))$,
- $I \in O(n)$ $(SO(n))$.

Uzasadnimy przykładowo pierwszy warunek. Jeśli $A, B \in O(n)$, to

$$(AB)^TAB = B^T \underbrace{A^TA}_{=I} B = B^TB = I,$$

więc $AB \in O(n)$. Jeśli dodatkowo, $A, B \in SO(n)$, to

$$\det(AB) = \underbrace{\det(A)}_{=1} \cdot \underbrace{\det(B)}_{=1} = 1$$

□

**Przykład 13.11.** Przyglądniemy się bliżej grupom $O(2)$ oraz $SO(2)$. Jeśli

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in O(2),$$

to z ortonormalności kolumn macierz $A$ mamy

$$a^2 + c^2 = b^2 + d^2 = 1, \quad ab + cd = 0.$$

Ponieważ $a^2 + c^2 = 1$, więc istnieje takie $\theta \in [0, 2\pi)$, że

$$a = \cos\theta, \quad c = \sin\theta.$$

Na okręgu jednostkowym $\mathbb{S}^1$ są dokładnie dwa wektory prostopadłe do wektora

$$[a, b]^T = [\cos\theta, \sin\theta]^T.$$

Są nimi

$$[c, d]^T = [-\sin\theta, \cos\theta]^T \quad \text{lub} \quad [c, d]^T = [\sin\theta, -\cos\theta]^T.$$

Macierz $A$ ma jedną z postaci:

(1) $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$;

(2) $A = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$.

Dla $\theta = 0$ i $\theta = \pi$ macierze są diagonalne i mamy cztery możliwości

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Odpowiadają one odwzorowaniu identycznościowemu, symetrii względem osi $x$, symetrii względem osi $y$ i symetrii środkowej.

W przypadku (1) mamy do czynienia z obrotami.

W przypadku (2), mamy macierz

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

odpowiadającą złożeniu obrotu z symetrią osiową. Jest to symetria względem prostej nachylonej do osi $x$ pod kątem $\theta/2$.

**Wniosek 13.20.** *Każda izometria płaszczyzny zachowująca* $0$ *jest albo obrotem, albo złożeniem symetrii względem osi* $x$ *z obrotem. Obrót jest złożeniem co najwyżej dwóch symetrii osiowych. Każda izometria płaszczyzny zachowująca* $0$ *jest złożeniem co najwyżej* $2$ *symetrii osiowych.*

**Przykład 13.12.** W przypadku grupy $SO(2)$ mamy dodatkową restrykcję, że $\det A = 1$ dla $A \in O(2)$. W szczególności musimy wykluczyć te macierze z $O(2)$, które zmieniają orientację. Prowadzi nas to do konkluzji, że $SO(2)$ składa się z macierzy obrotów postaci

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Jako zbiór mamy więc, że

$$SO(2) = \left\{ \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad \theta \in [0, 2\pi) \right\}.$$

Możemy więc zdefiniować bijekcję

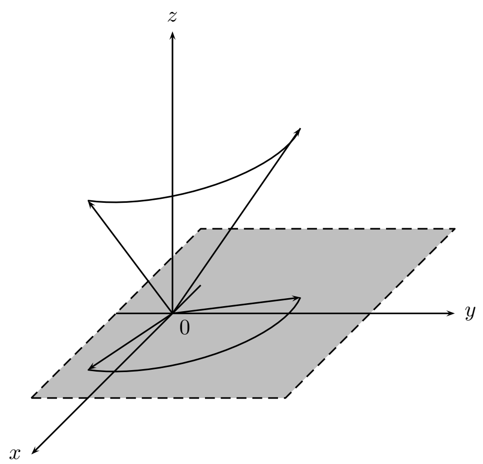
$$f : SO(2) \in \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \longmapsto (\cos\theta, \sin\theta) \in \mathbb{S}^1.$$

Jako zbiory (czyli z punktu widzenia teorii mnogości) $SO(2)$ oraz $\mathbb{S}^1$ są nierozróżnialne. My jednak mamy w tych zbiorach zdefiniowane dodatkowe struktury algebraiczne: $SO(2)$ z działaniem mnożenia macierzy jest grupą, a $\mathbb{S}^1$ jest grupą z mnożeniem liczb zespolonych. Czy możemy również te dwie struktury algebraiczne uznać za równoważne? Algebraicy chcą, aby wtedy bijekcja $f$ była zgodna ze strukturą algebraiczną w rozważanych zbiorach. W naszym przypadku żądamy więc, aby

$$f(AB) = f(A)f(B).$$

Mówimy wtedy, że $f$ jest *homomorfizmem* grup $SO(2)$ i $\mathbb{S}^1$ tzn. odwzorowanie $f$ jest zgodne z działaniami w grupach $SO(2)$ oraz $\mathbb{S}^1$. łatwo sprawdzamy bezpośrenim rachunkiem, że jest tak w naszym przypadku. Oznacza to, że struktury algebraiczne zadane w $SO(2)$ oraz $\mathbb{S}^1$ są identyczne z dokładnościa do przyjętej konwencji. Mówimy wtedy, że grupy $SO(2)$ (z mnożeniem macierzy) oraz $\mathbb{S}^1$ z mnożeniem liczb zespolonych są *izomorficzne*, tzn. nierozróżnialne z algebraicznego punktu widzenia.

**Lemat 13.9.** *Jeśli* $Q \in O(n)$ *i* $\lambda \in \mathbb{R}$ *jest rzeczywistą wartością własną, to* $\lambda = 1$ *lub* $\lambda = -1$.




Rysunek 13.6: Obrót w $\mathbb{R}^3$ wokół osi $z$. Jest on również obrotem w płaszczyźnie niezmienniczej $xy$, ortogonalnej do osi $z$.

*Dowód.* Oczywiście, $\lambda \neq 0$. Niech $v$ będzie wektorem własnym dla $\lambda$. Wtedy

$$\begin{aligned}\lambda(v|v) &= (\lambda \cdot v|v)\\ &= (Qv|v)\\ &= (v, Q^T v)\\ &= (v|Q^{-1}v)\\ &= (v|\lambda^{-1} \cdot v)\\ &= \lambda^{-1}(v|v),\end{aligned}$$

czyli $\lambda = \lambda^{-1}$, więc $\lambda^2 = 1$. $\square$

**Lemat 13.10** (Obroty w $\mathbb{R}$ dają $SO(3)$)**.** *Niech $Q \in SO(3)$. Istnieje taka baza ortonormalna $v_1, v_2, v_3$ w $\mathbb{R}^3$, że $A$ ma w tej bazie macierz*

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}.$$

*Dowód.* Rozważmy wielomian charakterystyczny macierzy $Q$

$$p_Q(z) = \det(Q - xI) = -\det(xI - Q).$$

Jeśli $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$ są wartościami własnymi $A$, to

$$p_Q(x) = -(x - \lambda_1)(x - \lambda_2)(x - \lambda_2).$$

Ponieważ $p_Q(0) = \lambda_1\lambda_2\lambda_3 = \det(Q) = 1$ oraz $\lim_{x\to+\infty} P_Q(x) = -\infty$, więc $p_Q$ ma pierwiastek w $(0, \infty)$. Z wniosku 13.10 wynika, że 1 jest wartością własną $Q$. Niech $v_1$ będzie jednostkowym wektorem własnym dla wartości własnej 1. Wtedy dla $w \in l := \operatorname{span}\{v_1\}$ mamy $Qw = w$, czyli na prostej $l$ każdy wektor jest punktem stałym $Q$. Niech $H = l^\perp$ będzie płaszczyną prostopadłą do prostej $l$, czyli

$$H = \{v : (v|v_1) = 0\}.$$

Dla $v \in H$ mamy

$$\begin{aligned}(Qv|v_1) &= (Qv|Qv_1)\\ &= (v|v_1)\\ &= 0,\end{aligned}$$

więc $Qv \in H$. Stąd, płaszczyna $H$ jest niezmiennicza dla $Q$ oraz $Q$ zacieśniona do $H$ jest izometrią płaszczyny $H$ o wyznaczniku równym 1. Oznacza to, że jeśli $v_2, v_3$ jest bazą ortonormalną dla $H$, to $Q$ ma w bazie $v_1, v_2, v_3$ macierz

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}.$$

□

**Wniosek 13.21** (Grupa $O(3)$)**.** *Jeśli $Q \in O(3)$, to w pewnej bazie ortonormalnej macierz $Q$ ma jedną z postaci*

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix}.$$

*Dowód.* Jeśli $\det Q = 1$, to $Q$ ma w pewnej bazie postać

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix}.$$

Załóżmy, że $\det Q = -1$. Wtedy $-1$ jest wartością własną $Q$. Niech $v$ będzie wektorem własnym odpowiadającym $-1$ o normie 1. Rozważmy $S = \operatorname{span}\{v\}$ prostą generowaną przez $v$. Wtedy

$$\mathbb{R}^3 = S \oplus S^{\perp},$$

gdzie $S^{\perp}$ jest płaszczyzną prostopadłą do prostej $S$. Zauważmy, że dla $w \in S^{\perp}$ mamy

$$\begin{aligned}(Qw|v) &= (w|Qv) \\ &= (w| \pm v_{\pm}) \\ &= \pm(w|v_{\pm}) \\ &= 0,\end{aligned}$$

czyli $Qw \in S^{\perp}$. Możemy więc traktować $Q$ jako izometrię płaszczyzny $S^{\perp}$. Niech $u_1, u_2$ będzie bazą ortonormalną $S^T$. Wtedy $Q$ ma w bazie $v, u_1, u_2$ macierz postaci

$$\begin{bmatrix} -1 & 0 \\ 0 & B \end{bmatrix},$$

gdzie $B$ jest macierzą izometrii $Q$ w restrykcji do $S^T$ w bazie $u_1, u_2$. Ponadto $\det B > 0$. Teza wynika z postaci grupy $SO(2)$. □

**Definicja 13.12. Grupa izometrii**

Zbiór izometrii

$$\operatorname{Isom}(\mathbb{R}^n) := \{f : \mathbb{R}^n \longrightarrow \mathbb{R}^n : f \text{ jest izometrią}\}$$

jest grupą z działaniem składania odwzorowań. Nazywamy ją grupą izometrii przestrzeni euklidesowej $\mathbb{R}^n$.

**Wniosek 13.22.** *$f \in \operatorname{Isom}(\mathbb{R}^n)$ wtedy i tylko wtedy, gdy $f = Q + v$ dla pewnej macierzy ortogonalnej $Q \in O(n)$ i wektora $v \in \mathbb{R}^n$.*

Pokażemy teraz pewien trik, pozwalający interpretować izometrię $f$ jako pewną macierz nawet gdy $f(0) \neq 0$, czyli dla $v \neq 0$. Zrobimy to dla $n = 3$. Niech $Q = [q_{ij}] \in O(3)$ oraz $v = [v_1, v_2, v_3]^T \in \mathbb{R}^3$. Definiujemy macierz

$$F := \begin{bmatrix} Q & 0 \\ v^T & 1 \end{bmatrix} = \left[\begin{array}{ccc|c} q_{11} & q_{12} & q_{13} & 0 \\ q_{21} & q_{22} & q_{23} & 0 \\ q_{31} & q_{32} & q_{33} & 0 \\ \hline v_1 & v_2 & v_3 & 1 \end{array}\right] \in GL_4(\mathbb{R}).$$

Niech $[x_1, x_2, x_3]^T \in \mathbb{R}^3$. Wtedy

$$[x_1, x_2, x_3, 1] \begin{bmatrix} Q & 0 \\ v^T & 1 \end{bmatrix} = \begin{bmatrix} x_1q_{11} + x_2q_{21} + x_3q_{31} + v_1 \\ x_1q_{12} + x_2q_{22} + x_3q_{32} + v_2 \\ x_1q_{13} + x_2q_{23} + x_3q_{33} + v_3 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} Qx + v \\ 1 \end{bmatrix} \in \mathbb{R}^4.$$

## 13.9 Symetrie w $\mathbb{R}^n$

Niech $u \in \mathbb{R}^n$ będzie wektorem jednostkowym. Wtedy odwzorowanie $L : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ dane wzorem

$$L(x) = x - 2(x|u) \cdot u$$

jest symetrią względem hiperpłaszczyzny (odbiciem)

$$u^\perp = \{x \in \mathbb{R}^n : (x|u) = 0\}.$$

**Lemat 13.11.** *Następujące warunki są równoważne:*

*(1) $L$ jest symetrią względem $u^\perp$ dla pewnego wektora jednostkowego $u$,*

*(2) $L$ ma w pewnej bazie ortonormalnej $u_1, \ldots, u_n$ postać diagonalną*

$$D = \operatorname{diag}(-1, 1, \ldots, 1).$$

*Dowód.* Jeśli zachodzi warunek (1), to definiujemy $u_1 = u$ i rozszerzamy go do bazy ortonormalnej $u_1, u_2, \ldots, u_n$ i wtedy zachodzi (2).

Załóżmy, że jest spełniony warunek (2). Niech $u = u_1$. Wtedy $L$ oraz

$$L_1(x) = x - 2(x|u) \cdot u,$$

przyjmują te same wartości na bazie $u_1, \ldots, u_n$, czyli $L = L_1$. □

**Wniosek 13.23.** *Jeśli $L : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ jest symetrią względem $u^\perp$ dla pewnego wektora jednostkowego $u$, to*

- $L^2 = I$,
- $L \in O(n)$,
- $\det L = -1$.

**Lemat 13.12.** *Załóżmy, że $a, b \in \mathbb{R}^n$ są różne oraz $\|a\| = \|b\|$. Wtedy dla $u = \frac{a-b}{\|a-b\|}$ oraz $L(x) = x - 2(x|u) \cdot u$ mamy $L(a) = b$. Ponadto jeśli $w$ jest ortogonalny do $a$ i $b$, to $L(w) = w$.*

*Dowód.* Ponieważ $\|a\| = \|b\|$, więc

$$\begin{aligned} 2(a|a-b) &= 2(a|a) - 2(a|b) \\ &= (a|a) + (b|b) - 2(a|b) \\ &= (a-b|a-b). \end{aligned}$$

Stąd

$$\begin{aligned} L(a) &= a - 2(a|u) \cdot u \\ &= a - \frac{2(a|a-b)}{\|a-b\|^2} \cdot (a-b) \\ &= a - (a-b) \\ &= b. \end{aligned}$$

□

**Lemat 13.13.** *Niech $I \neq Q \in O(n)$. Istnieją taki wektor jednostkowy $u_1$ oraz taka symetria $L$ względem $u^\perp$ dla pewnego wektora jednostkowego $u$, że*

$$Q(u_1) \neq u_1, \quad LQ(u_1) = u_1.$$

*Dowód.* Ponieważ $Q \neq I$, więc istnieje taki wektor $x \in \mathbb{R}^n$, że $Qx \neq x$. Oczywiście $x \neq 0$. Definiujemy $u_1 = x/\|x\|$. Wtedy $Qu_1 \neq u_1$ oraz

$$\|Qu_1\| = \frac{\|Qx\|}{\|x\|} = \frac{\|x\|}{\|x\|} = 1.$$

Z lematu 13.12 zastosowanego do $a = Qx$, $b = x$ wynika, że teza zachodzi dla

$$u = \frac{Qx - x}{\|Qx - x\|}.$$

□

**Lemat 13.14.** *Załóżmy, że $Q \in O(n)$. Niech $u_1, \ldots, u_k$ będzie bazą ortonormalną przestrzeni $\ker(Q-I)$. Wtedy albo $k = n$ (czyli $Q = I$), albo istnieją taki wektor jednostkowy $u_{k+1}$ oraz taka symetria $L$ względem $u^\perp$ pewnego wektora jednostkowego $u$, że $Qu_{k+1} \neq u_{k+1}$ oraz $LQ(u_i) = u_i$ dla $i = 1, \ldots, k+1$. Ponadto wektory $u_1, \ldots, u_{k+1}$ są ortonormalne.*

*Dowód.* Załóżmy, że $Q \neq I$, czyli $k < n$. Istnieje takie $x \in \mathbb{R}^n$, że $Qx \neq x$. Definiujemy

$$v = x - \sum_{j=1}^{k} (x|u_i) \cdot u_i \in (\ker(Q - I))^\perp.$$

Zauważmy, że $v \neq 0$, bo w przeciwnym razie $Qx = x$. Niech

$$u_{k+1} = \frac{v}{\|v\|}.$$

Wtedy $Qu_{k+1} \neq u_{k+1}$, $\|Qu_{k+1}\| = 1$ oraz $u_{k+1}, Qu_{k+1}$ są ortogonalne do $u_1, \ldots, u_k$. Z lematu 13.12 istnieje takie odbicie $L$, że

$$LQ(u_{k+1}) = u_{k+1}, \quad L(u_j) = u_j, \quad j = 1, \ldots, k.$$

□

**Twierdzenie 13.9.** *Jeśli $Q \in O(n)$, to istnieją takie inwolucje $L_1, \ldots, L_k$ ($i = 1, \ldots, k$), że*

$$Q = L_1 \ldots L_k, \quad k \leqslant n.$$

*Dowód.* Ponieważ w $\mathbb{R}^n$ nie może być więcej niż $n$ wektorów ortonoramlnych, więc z lematu 13.14 istnieją takie odbicia $L_1, \ldots, L_k$ dla $0 \leqslant k \leqslant n$, że

$$L_k \ldots L_1 Q = I,$$

co kończy dowód, bo $L_i^{-1} = L_i$. □

# Rozdział 14

# Macierze dodatnio określone

| |
|---|
| SŁOWA KLUCZOWE ROZDZIAŁU<br>macierz dodatnio określona ◊ minor główny ◊ rozkład polarny ◊ forma kwadratowa ◊ krzywe stopnia drugiego ◊ postać kanoniczna formy kwadratowej ◊ powierzchnie stopnia drugiego |

## 14.1 Macierze dodatnio określone

**Definicja 14.1.** **Macierz symetryczna dodatnio określona**
Macierz symetryczną $A \in M_{n\times n}(\mathbb{R})$ nazywamy

(i) *dodatnio określoną*, jeśli $(Av|v) = v^T Av > 0$ dla każdego niezerowego wektora $v \in \mathbb{R}^n$,

(ii) *ujemnie określoną*, jeśli $(Av|v) = v^T Av < 0$ dla każdego niezerowego wektora $v \in \mathbb{R}^n$,

(iii) *dodatnio półokreśloną*, jeśli $(Av|v) = v^T Av \geqslant 0$ dla każdego wektora $v \in \mathbb{R}^n$,

(iv) *ujemnie półokreśloną*, jeśli $(Av|v) = v^T Av \leqslant 0$ dla każdego wektora $v \in \mathbb{R}^n$.

Analogicznej terminologii używamy w stosunku do skojarzonych form kwadratowych $Q_A$.

**Przykład 14.1.** Rozważmy macierz diagonalną $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \in M_{n\times n}(\mathbb{R})$. Ponieważ dla $v = [v_1, \ldots, v_n]^T \in \mathbb{R}^n$ mamy

$$(Dv|v) = \lambda_1 v_1^2 + \ldots + \lambda_n v_n^2,$$

więc $D$ jest dodatnio półokreślona (dodatnio określona) wtedy i tylko wtedy, gdy wszystkie wartości własne $\lambda_1, \ldots, \lambda_n$ są nieujemny (dodatnie).

**Lemat 14.1.** *Niech $Q \in O(n)$ będzie ortogonalna. Wtedy macierz symetryczna $A \in M_{n\times n}(\mathbb{R})$ jest dodatnio określona (półokreślona) wtedy i tylko wtedy, gdy $Q^T AQ$ jest dodatnio określona (półokreślona).*

*Dowód.* Ze względu na symetrią wystarczy pokazać, że jeśli $A$ jest dodatnio określona (półokreślona) wtedy i tylko wtedy, gdy $Q^T AQ$ jest dodatnio określona (półokreślona). Mamy

$$(Q^T AQv|v) = (AQv|Qv)$$

oraz $Q$ jest nieosobliwa, więc $A$ jest dodatnio określona wtedy i tylko wtedy, gdy $Q^T AQ$ jest dodatnio określona. □

**Twierdzenie 14.1** (Dodatnie określona = dodatnie wartości własne)**.** *Macierz symetryczna $A \in M_{n\times n}(\mathbb{R})$ jest dodatnio określona (półokreślona) wtedy i tylko wtedy, gdy wszystkie wartości własne macierzy $A$ są dodatnie (nieujemne).*

*Dowód.* Wynika z przykładu 14.1 i lematu 14.1. □

**Definicja 14.2.** Pierwiastek kwadratowy z macierzy

Niech $A \in M_{n\times n}(\mathbb{C})$. Macierz $B \in M_{n\times n}(\mathbb{C})$ nazywamy *pierwiastkiem kwadratowym* z macierzy $A$, gdy $B^2 = A$.

**Przykład 14.2.** Macierz nilpotentna

$$N = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

nie posiada pierwiastka kwadratowego. Rzeczywście, przypuśćmy, że $B^2 = N$ dla pewnej macierzy $B \in M_{2\times 2}(\mathbb{C})$. Wtedy $B$ jest również nilpotentna, więc z twierdzenia Cayleya–Hamiltona wynika, że $B^2 = 0$. Stąd, $N = 0$ co prowadzi do sprzeczności.

**Przykład 14.3.** Macierz identycznościowa

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

ma nieskończenie wiele symetrycznych pierwiastków kwadratowych $B \in M_{2\times 2}(\mathbb{Q})$:

$$\frac{1}{c}\begin{bmatrix} a & b \\ b & -a \end{bmatrix}, \quad \frac{1}{c}\begin{bmatrix} a & -b \\ -b & -a \end{bmatrix}, \quad \frac{1}{c}\begin{bmatrix} -a & -b \\ -b & a \end{bmatrix}, \quad \frac{1}{c}\begin{bmatrix} -a & b \\ b & a \end{bmatrix},$$

gdzie $a, b, c \in \mathbb{N}$ są takie, że $a^2 + b^2 = c^2$.

**Twierdzenie 14.2** (Pierwiastek z macierzy dodatnio określonej)**.** *Załóżmy, że macierz symetryczna $A \in \mathbb{R}$ jest dodatnio półokreślona. Istnieje dokładnie jedna taka dodatnio półokreślona macierz symetryczna $B \in M_{n\times n}(\mathbb{R})$, że $B^2 = A$. Ponadto macierz dodatnio określona $A$ ma dokładnie jeden dodatnio określony pierwiastek kwadratowy.*

*Dowód.* Pokażemy najpierw, że taka macierz $B$ istnieje. Wartości własne $\lambda_1, \dots, \lambda_n$ macierzy $A$ są rzeczywiste i nieujemne. Ponadto istnieje taka macierz ortogonalna $Q \in O(n)$, że

$$Q^T A Q = \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

Dla

$$B = Q \operatorname{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}) Q^T,$$

otrzymujemy, że

$$\begin{aligned} B^2 &= \big(Q \operatorname{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}) Q^T\big)\big(Q \operatorname{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}) Q^T\big) \\ &Q \operatorname{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}) \operatorname{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}) Q^T \\ &= Q \operatorname{diag}(\lambda_1, \dots, \lambda_n) Q^T \\ &= A. \end{aligned}$$

Uzasadnimy, że $B$ jest wyznaczona jednoznacznie. Przypuśćmy, że macierze symetryczne $B, C$ są dodatnio półokreślone, $B \neq C$ oraz

$$B^2 = A = C^2.$$

Wtedy $B - C \neq 0$ jest macierzą symetryczną, czyli ma niezerową wartość własną $\lambda \in \mathbb{R}$. Jeśli $0 \neq v \in \mathbb{R}^n$ jest jej wektorem własnym, to

$$\begin{aligned} 0 &= (v|(\underbrace{B^2 - C^2}_{=0})v) \\ &= (v|B(B-C)v) + (v|C(B-C)v) \\ &= \lambda\big((v|Bv) + (v|Cv)\big). \end{aligned}$$

Ponieważ $B$ i $C$ są dodatnio półokreślone, więc

$$(v|Bv) = (v|Cv) = 0.$$

Stąd

$$0 = (v|Bv) - (v|Cv) = (v|(B - C)v) = \lambda(v|v),$$

czyli $\lambda = 0$. Otrzymana sprzeczność dowodzi, że $B = C$. $\square$

**Twierdzenie 14.3.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie symetryczna. Następujące warunki są równoważne*

*(i) $A$ ma nieujemne wartości własne,*

*(ii) $A = B^T B$ dla pewnej macierzy $B \in M_{n\times n}(\mathbb{R})$.*

*Ponadto $A$ ma dodatnie wartości własne wtedy i tylko wtedy, gdy $A = B^T B$ i $B$ jest nieosobliwa.*

*Dowód.* Załóżmy, że wartości własne $\lambda_1, \ldots, \lambda_n$ są nieujemne. Dla pewnej macierzy ortogonalnej $Q$ mamy

$$Q^T AQ = D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

Ponieważ wartości własne są nieujemne, więc możemy zdefiniować

$$\sqrt{D} = \mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$$

oraz

$$B = \sqrt{D}Q^T.$$

Wtedy

$$\begin{aligned} A &= QDQ^T \\ &= Q\sqrt{D}\sqrt{D}Q^T \\ &= B^T B. \end{aligned}$$

Oczywiście, wartości własne są dodatnie wtedy i tylko wtedy, gdy $B$ jest nieosobliwa.

Załóżmy teraz, że zachodzi warunek (ii). Wartości własne macierzy Grama $A = B^T B$ są nieujemne (lemat 13.2) oraz są dodatnie wtedy i tylko wtedy, gdy $B$ jest nieosobliwa. $\square$

**Wniosek 14.1.** *Jeśli macierz symetryczna $A$ jest dodatnio określona, to* $\det(A) > 0$.

*Dowód.* Jeśli $\lambda_1, \ldots, \lambda_n > 0$ są wartościami własnymi $A$, to $\det(A) = \lambda_1 \ldots \lambda_n > 0$. $\square$

**Definicja 14.3.** **Minor wiodący**

Niech $A \in M_{n\times n}(\mathbb{R})$. *Wiodący minor* $A_r \in M_{r\times r}(\mathbb{R})$ $(r = 1, \ldots, n)$ macierzy $A$ jest macierzą powstałą z $A$ przez wykreślenie ostatnich $n - r$ wierszy i kolumn.

**Przykład 14.4.** Dla macierzy

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

mamy wiodące minory

$$A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad A_1 = \begin{bmatrix} a_{11} \end{bmatrix}.$$

**Lemat 14.2.** *Jeśli macierz symetryczna $A$ jest dodatnio określona, to wiodące minory $A_1, \ldots, A_n$ są dodatnio określone. W szczególności,* $\det A_r > 0$.

*Dowód.* Niech $0 \neq v(r) = [v_1, \ldots, v_r]^T \in \mathbb{R}^r$. Rozważmy wektor $v = [v_1, \ldots, v_r, 0, \ldots, 0]^T \in \mathbb{R}^n$. Wtedy

$$(A_r v(r)|v(r)) = (Av|v) > 0.$$

□

Zajmiemy się teraz bardziej szczegółowo macierzami dodatnio określonymi $A$ w wymiarze $n = 3$. Wszystko co dalej powiemy uogólnia się na wyższe wymiary. Zaczniemy od pewnych obserwacji dotyczących macierzy dolnie i górnie trójkątnych.

Definiujemy

$$\mathcal{L}_1 = \Big\{A \in M_{3\times 3}(\mathbb{R}) : \ A \quad \text{jest dolnie trójkątna z jedynkami na przekątnej}\Big\},$$

$$\mathcal{U}_1 = \Big\{A \in M_{3\times 3}(\mathbb{R}) : \ A \quad \text{jest górnie trójkątna z jedynkami na przekątnej}\Big\},$$

$$\mathcal{L}_+ = \Big\{A \in M_{3\times 3}(\mathbb{R}) : \ A \quad \text{jest dolnie trójkątna z dodatnimi wyrazami na przekątnej}\Big\},$$

$$\mathcal{U}_+ = \Big\{A \in M_{3\times 3}(\mathbb{R}) : \ A \quad \text{jest górnie trójkątna z dodatnimi wyrazami na przekątnej}\Big\}.$$

Sprawdzamy łatwo, że z działaniem mnożenia macierzy powyższe zbiory są grupami.

**Lemat 14.3.** *Niech $D$ będzie nieosobliwą macierzą diagonalną (tzn. z niezerowymi elementami na przekątnej). Jeśli*

$$LD = DU, \quad L \in \mathcal{L}_1, \quad U \in \mathcal{U}_1,$$

*to $L = D = I$.*

*Dowód.* Niech

$$L = \begin{bmatrix} 1 & 0 & 0 \\ b_1 & 1 & 0 \\ b_2 & b_3 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}, \quad U = \begin{bmatrix} 1 & a_1 & a_2 \\ 0 & 1 & a_3 \\ 0 & 0 & 1 \end{bmatrix}, \quad a, b, c \neq 0.$$

Wtedy

$$LD = \begin{bmatrix} 1 & 0 & 0 \\ b_1 & 1 & 0 \\ b_2 & b_3 & 1 \end{bmatrix} \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} = \begin{bmatrix} a & 0 & 0 \\ b_1 a & b & 0 \\ b_2 a & b_3 b & c \end{bmatrix}.$$

Ponieważ $LD = DU$, więc $LD$ jest górnie trójkątna. Stąd $b_1 = b_2 = b_3 = 0$, bo $a, b, c$ są różne od zera. Analogicznie pokazujemy, że $U = I$. □

**Wniosek 14.2.** *Jeśli $D$ jest nieosobliwą macierzą diagonalną oraz*

$$LDU = L_1 D U_1, \quad L, L_1 \in \mathcal{L}_1, \quad U, U_1 \in \mathcal{U}_1,$$

*to $L = L_1$ oraz $U = U_1$.*

*Dowód.* Mamy

$$(L_1^{-1} L) D = D(U_1 U^{-1}),$$

więc z lematu 14.3 wynika, że

$$L_1^{-1} L = I = U_1 U^{-1}.$$

□

**Lemat 14.4.** *Niech $A \in M_{3\times 3}(\mathbb{R})$ będzie macierzą symetryczną z minorami wiodącymi o dodatnich wyznacznikach. Używając wyłącznie (III) dozwolonej operacji na wierszach, możemy macierz $A$ sprowadzić do postaci górnie trójkątnej z dodatnimi wyrazami na przekątnej. W szczególności,*

$$A = LU, \quad L \in \mathcal{L}_1, \quad U \in \mathcal{U}_+.$$

*Dowód.* Zauważmy najpierw, że

$$a_{11} = \det A_1 > 0.$$

**Krok 1.** Używając pierwszego wiersza i operacji (III) możemy macierz $A$ sprowadzić do postaci

$$A = \left[\begin{array}{cc|c} a_{11} & * & * \\ * & a_{22} & * \\ \hline * & * & a_{33} \end{array}\right] \dashrightarrow \left[\begin{array}{cc|c} a_{11} & * & * \\ 0 & a_{22}^{(1)} & * \\ \hline * & * & a_{33} \end{array}\right].$$

Wykonaną tu operację możemy zapisać przy pomocy mnożenia z lewej strony macierzy $A$ przez macierz elementarną

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ b_1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \in \mathcal{L}_1,$$

gdzie $b_1 = -\frac{a_{21}}{a_{11}}$. Oznacza to, że

$$E_1 A = \left[\begin{array}{cc|c} a_{11} & * & * \\ 0 & a_{22}^{(1)} & * \\ \hline a_{31} & a_{32} & a_{33} \end{array}\right].$$

**Krok 2.** Następnie macierz $E_1A$ sprowadzamy używając operacji (III) do postaci

$$E_1 A = \left[\begin{array}{cc|c} a_{11} & * & * \\ 0 & a_{22}^{(1)} & * \\ \hline a_{31} & a_{32} & a_{33} \end{array}\right] \dashrightarrow \left[\begin{array}{cc|c} a_{11} & * & * \\ 0 & a_{22}^{(1)} & * \\ \hline 0 & * & a_{33}^{(1)} \end{array}\right] := A^{(1)}.$$

W krokach 1 i 2 minor

$$A_2 = \begin{bmatrix} a_{11} & * \\ * & a_{22} \end{bmatrix}$$

został przekształcony na macierz

$$\begin{bmatrix} a_{11} & * \\ 0 & a_{22}^{(1)} \end{bmatrix}$$

o takim samym wyznaczniku, więc $a_{22}^{(1)} = \frac{\det A_2}{\det A_1} > 0$.

**Krok 3.** Możemy więc stosując (III) operację sprowadzić macierz $A^{(1)}$ do postaci

$$A^{(1)} = \begin{bmatrix} a_{11} & * & * \\ 0 & a_{22}^{(1)} & * \\ 0 & * & a_{33}^{(1)} \end{bmatrix} \dashrightarrow \begin{bmatrix} a_{11} & * & * \\ 0 & a_{22}^{(1)} & * \\ 0 & 0 & a_{33}^{(2)} \end{bmatrix}.$$

Po dwóch krokach minor $A_3$ został przekształcony na macierz

$$U = \begin{bmatrix} a_{11} & * & * \\ 0 & a_{22}^{(1)} & * \\ 0 & 0 & a_{33}^{(2)} \end{bmatrix}$$

o takim samym wyznaczniku, czyli

$$0 < \det A_3 = a_{11} a_{22}^{(1)} a_{33}^{(2)},$$

więc

$$a_{33}^{(2)} = \frac{\det A_3}{\det A_2} > 0.$$

W tym kroku dla

$$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & b_3 & 1 \end{bmatrix},$$

mamy

$$E_3E_2E_1A = U.$$

Zauważmy, że

$$E_3E_2E_1 = \begin{bmatrix} 1 & 0 & 0 \\ b_1 & 1 & 0 \\ b_2 & b_3 & 1 \end{bmatrix} \in \mathcal{L}_1.$$

Ponadto

$$\begin{aligned} L &= (E_3E_2E_1)^{-1} \\ &= E_1^{-1}E_2^{-1}E_3^{-1} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ -b_1 & 1 & 0 \\ -b_2 & -b_3 & 1 \end{bmatrix} \in \mathcal{L}_1, \end{aligned}$$

więc

$$A = LU, \quad L \in \mathcal{L}_1, \quad U \in \mathcal{U}_+.$$

□

**Wniosek 14.3.** *Niech $A = LU \in M_{3\times 3}(\mathbb{R})$ dla pewnych $L \in \mathcal{L}_1$ oraz $U \in \mathcal{U}_+$. Istnieje macierz diagonalna $D$ z dodatnimi wyrazami na przekątnej oraz takie macierze $L \in \mathcal{L}_1$, $U_1 \in \mathcal{U}_1$, że*

$$A = LDU_1.$$

*Ponadto jeśli $A$ jest symetryczna, to*

*(a)* $A = LDL^T$,

*(b)* $A = L_1L_1^T$ *dla pewnej macierzy* $L_1 \in \mathcal{L}_+$.

*Dowód.* Zauważmy, że

$$\begin{aligned} U &= \begin{bmatrix} a_{11} & a & b \\ 0 & a_{22} & c \\ 0 & 0 & a_{33} \end{bmatrix} \\ &= \underbrace{\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} 1 & a/a_{11} & b/a_{11} \\ 0 & 1 & c/a_{22} \\ 0 & 0 & 1 \end{bmatrix}}_{U_1} \\ &= DU_1. \end{aligned}$$

Pokażemy, że zachodzi warunek (a). Rozważmy rozkład $A = LDU_1$. Wtedy

$$\begin{aligned} LDU_1 &= A \\ &= A^T \\ &= (U_1)^T DL^T. \end{aligned}$$

Z wniosku 14.2 wynika, że $L^T = U_1$, czyli $A = LDL^T$.

Dla dowodu punktu (b) rozważmy macierze $\sqrt{D} := \operatorname{diag}(\sqrt{d_1}, \sqrt{d_2}, \sqrt{d_3})$ oraz $L_1 = L\sqrt{D}$. Wtedy

$$\begin{aligned} A &= LDL^T \\ &= L\sqrt{D}\sqrt{D}^T L^T \\ &= L_1L_1^T. \end{aligned}$$

□

**Twierdzenie 14.4** (Sylvester)**.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą symetryczną. Następujące warunki są równoważne*

*(i)* *$A$ jest dodatnio określona,*

*(ii)* *$A$ ma dodatnie wartości własne,*

*(iii)* *wiodące minory $A_1, \dots, A_n$ mają dodatnie wyznaczniki,*

*(iv)* *$A = LU$, gdzie $L \in \mathcal{L}_1$ oraz $U \in \mathcal{U}_+$,*

*(v)* *$A = LDL^T$, gdzie $L \in \mathcal{L}_1$ oraz $D$ jest diagonalna z dodatnimi wyrazami digonalnymi,*

*(vi)* *$A = LL^T$, gdzie $L \in \mathcal{L}_+$,*

*(vii)* *$A = B^T B$ dla pewnej macierzy nieosobliwej $B$.*

Rozkład $A = LL^T$ nazywany jest rozkładem Choleskiego (1875–1918).

*Dowód.* Z naszych poprzednich rozważań wynika, że

$$(i) \xRightarrow{\text{tw. 14.1}} (ii) \xRightarrow{\text{lem. 14.2}} (iii) \xRightarrow{\text{lem. 14.4}} (iv) \xRightarrow{\text{wn. 14.3 (a)}} (v) \xRightarrow{\text{wn. 14.3 (b)}} (vi) \Rightarrow (vii) \xRightarrow{\text{tw. 14.3}} (i).$$

□

## 14.2 Rozkład polarny (biegunowy)

**Lemat 14.5.** *Jeśli $A \in M_{n\times n}(\mathbb{R})$ jest nieosobliwa, to macierz Grama $A^T A$ ma dodatnie wartości własne.*

*Dowód.* Z lematu 13.2 wiemy, że wartości własne macierzy $A^T A$ są nieujemne. Ponieważ $A^T A$ jest nieosobliwa jako iloczyn macierzy nieosobliwych, więc 0 nie jest wartością własną $A^T A$. □

**Twierdzenie 14.5** (Rozkład polarny macierzy nieosobliwej)**.** *Jeśli $A \in M_{n\times n}(\mathbb{R})$ jest nieosobliwa, to istnieją jednoznacznie wyznaczone macierz ortogonalna $Q \in O(n)$ oraz dodatnio określona macierz symetryczna $P \in M_{n\times n}(\mathbb{R})$ takie, że*

$$A = PQ.$$

*Dowód.* Macierz $AA^T$ jest symetryczna i ma dodatnie wartości własne, więc jest dodatnio określona. Z twierdzenia 14.2 istnieje dokładnie jedna taka dodatnio określona macierz symetryczna $P \in M_{n\times n}(\mathbb{R})$, że $P^2 = AA^T$. Definiujemy

$$Q := P^{-1}A.$$

Wtedy

$$\begin{aligned} QQ^T &= P^{-1}AA^T(P^{-1})^T \\ &= P^{-1}\underbrace{A^T A}_{=P^2}(\underbrace{P^T}_{=P})^{-1} \\ &= P^{-1}P^2P^{-1} \\ &= I, \end{aligned}$$

czyli $Q$ jest ortogonalna.

Pokażemy, że taki rozkład jest jednoznaczny. Przypuśćmy, że $A = P_1Q_1$, gdzie $Q_1 \in O(n)$ oraz $P_1$ jest symetryczna i dodatnio określona. Wtedy

$$\begin{aligned} AA^T &= P_1Q_1Q_1^TP_1^T \\ &= P_1P_1^T \\ &= P_1^2. \end{aligned}$$

Z twierdzenia 14.2 wynika, że $P = P_1$. Stąd

$$Q_1 = P^{-1}A = Q.$$

□

U Nazwa rozkładu polarnego (biegunowego) macierzy nieosobliwej $A = PQ$ pochodzi od *współrzędnych biegunowych* (postaci trygonometrycznej) liczby zespolonej. Niech $z = a + ib \in \mathbb{C} \setminus \{0\}$. W macierzowej interpretacji ciała liczb zespolonych $\mathbb{C}$ odpowiada jej macierz

$$a + ib \longmapsto \begin{bmatrix} a & b \\ -b & a \end{bmatrix}.$$

W postaci trygonometrycznej $z = |z|(\cos\theta + i\sin\theta)$ powyższą odpowiedniość możemy zapisać w postaci

$$|z|(\cos\theta + i\sin\theta) \longmapsto \underbrace{\begin{bmatrix} |z| & 0 \\ 0 & |z| \end{bmatrix}}_{\text{P dodatnio określona}} \underbrace{\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}}_{\text{Q ortogonalna}}.$$

## 14.3 Formy kwadratowe w wymiarze 2

**Definicja 14.4.** Forma kwadratowa

Funkcję $Q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ postaci

$$Q(v) = ax^2 + 2bxy + cy^2, \quad a, b, c \in \mathbb{R}, \quad v = [x, y]^T \in \mathbb{R}^2$$

nazywamy *formą kwadratową dwóch zmiennych*. Zakładamy, że $a^2 + b^2 + c^2 > 0$.

Zauważmy, że dla macierzy symetrycznej

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

oraz wektora $v = [x, y]^T$ zachodzą równości

$$\begin{aligned} Q(v) &= ax^2 + 2bxy + cy^2 \\ &= ([ax + by, bx + cy]^T | [x, y]^T) \\ &= (Av|v) \\ &= v^TAv \\ &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \end{aligned}$$

Jak wiemy dla macierzy symetrycznej $A$ istnieją wektory własne $u_1, u_2$ macierzy $A$, tworzące bazę ortonormalną. Wtedy $P = \left[\begin{array}{c|c} u_1 & u_2 \end{array}\right]$ jest macierzą ortogonalną, więc $P^{-1} = P^T$. Ponadto macierz $D = P^{-1}AP = P^TAP$ jest diagonalna z wartościami własnymi $\lambda$, $\mu$ na przekątnej. Niech $[x', y']^T = w = P^{-1}(v)$. Wtedy

$$\begin{aligned} Q_A(v) &= (Av|v) \\ &= (APw|Pw) \\ &= (P^TAPw|w) \\ &= Q_D(w) \\ &= (Dw|w) \\ &= \lambda(x')^2 + \mu(y')^2. \end{aligned}$$

**Definicja 14.5.** Krzywe drugiego stopnia
*Krzywymi drugiego stopnia* nazywamy krzywe zadane równaniem postaci

$$ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0, \quad a^2 + b^2 + c^2 > 0.$$

Równanie krzywej stopnia drugiego

$$ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0, \quad a^2 + b^2 + c^2 > 0 \tag{14.1}$$

możemy zapisać w postaci

$$(Av|v) + 2(v|u) + f = 0,$$

gdzie $v = [x, y]^T$, $u = [d, f]^T$. Załóżmy, że $\det A \neq 0$.

Znajdziemy teraz tzw. *postać kanoniczną* równania krzywej stopnia drugiego. Pokażemy najpierw, że istnieje takie $z = [z_1, z_2]^T = \in \mathbb{R}^2$, że dla $w = v + z$ przyjmuje ono postać

$$(Aw|w) + \tilde{f} = 0,$$

dla pewnego $\tilde{f} \in \mathbb{R}$. Zauważmy, że

$$\begin{aligned} 0 &= (Av|v) + 2(v|u) + f = (A(w - z)|w - z) + 2(w - z|u) + f \\ &= (Aw|w) - (Aw|z) - (Az|w) + (Az|z) + 2(w|u) - 2(z|u) + f \\ &= (Aw|w) + \tilde{f} - 2(Aw|z) + 2(w|u), \end{aligned}$$

dla $\tilde{f} = (Az|z) - 2(z|u) + f$. Szukamy więc takiego $z$, aby

$$0 = -2(w|Az) + 2(w|u) = 2(w|u - Az)$$

dla każdego $w \in \mathbb{R}^2$. Będzie tak tylko wtedy, gdy $Az = u$, czyli $z = A^{-1}u$. We współrzędnych

$$x' = x + z_1, \quad y' = y + z_2$$

równanie (14.1) ma postać

$$a(x')^2 + 2bx'y' + c(y')^2 + \tilde{f} = 0.$$

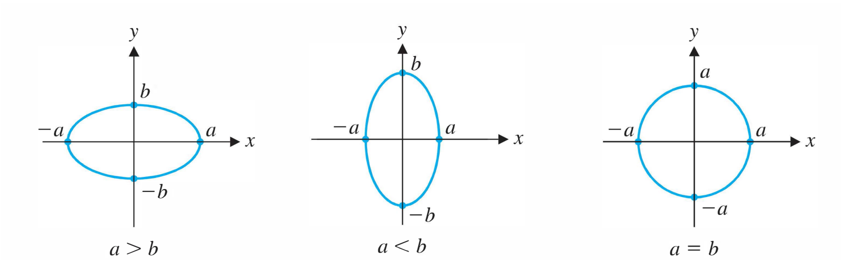
Z powyższych rozważań wynika, że po liniowej zmianie zmiennych przez macierz ortogonalną $P = \left[\begin{array}{c|c} u_1 & u_2 \end{array}\right]$ przyjmuje ono postać kanoniczną

$$\lambda(x'')^2 + \mu(y'')^2 + \tilde{f} = 0,$$

gdzie $\lambda$, $\mu \in \mathbb{R}$ są niezerowymi wartościami własnymi macierzy symetrycznej $A$.

**Przykład 14.5.** Okrąg o środku w punkcie $[x_0, y_0]^T$ i promieniu $r > 0$ jest krzywą drugiego stopnia o równaniu

$$(x - x_0)^2 + (y - y_0)^2 = r^2.$$




Rysunek 14.1: Elipsa i okrąg.

**Przykład 14.6.** Elipsa o środku w punkcie $(x_0, y_0)^T$ jest krzywą drugiego stopnia o równaniu

$$\frac{(x-x_0)^2}{a^2} + \frac{(y-y_0)^2}{b^2} = 1, \quad a, b \neq 0.$$

Elipsa

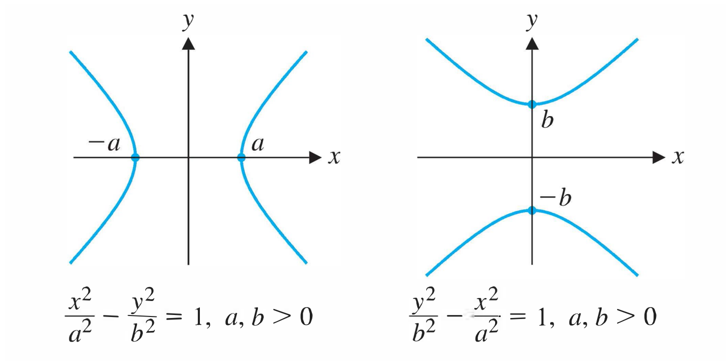
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

jest obrazem okręgu $x^2 + y^2 = 1$ przez izomorfizm liniowy zadany macierzą $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$. Rzeczywiście, jeśli

$$[x', y']^T = A[x, y]^T = [ax, by]^T,$$

to $[x, y]^T = [x'/a, y'/b]^T$, czyli $x^2 + y^2 = 1$ oznacza, że

$$\frac{(x')^2}{a^2} + \frac{(y')^2}{b^2} = 1.$$



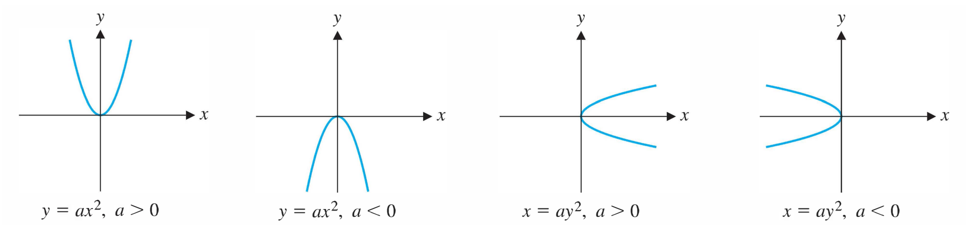

Rysunek 14.2: Hiperbola.

**Przykład 14.7.** Hiperbola o środku w punkcie $[x_0, y_0]^T$ jest krzywą drugiego stopnia o równaniu

$$\frac{(x-x_0)^2}{a^2} - \frac{(y-y_0)^2}{b^2} = 1, \quad a, b \neq 0.$$

**Przykład 14.8.** Parabola o wierzchołku $[x_0, y_0]^T$ jest krzywą drugiego stopnia o równaniu

$$a(x-x_0)^2 - (y-y_0) = 0, \quad a \neq 0.$$



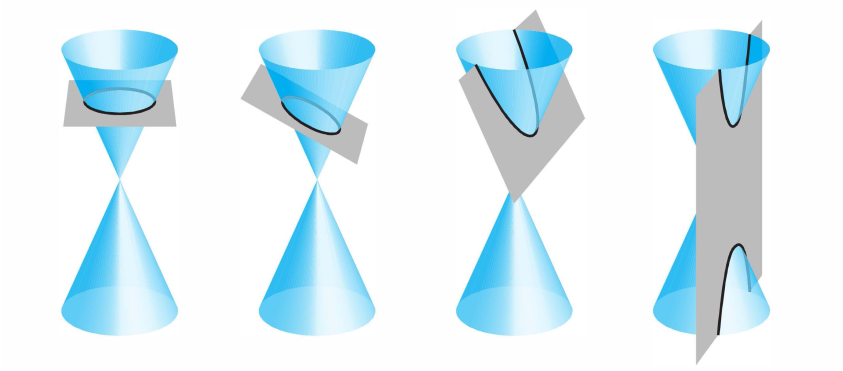

Rysunek 14.3: Parabola.

**Przykład 14.9.** Rozważmy równanie

$$3x^2 + 2xy + 3y^2 = 8.$$

Dla macierzy symetrycznej

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \quad \text{oraz} \quad z = [x, y]^T$$



Rysunek 14.4: Krzywe stopnia drugiego (stożkowe): okrąg, elipsa, parabola i hiperbola.

możemy je zapisać w postaci

$$(Az|z) = z^T Az = [x\ y] \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 8.$$

Macierz ma ma wartości własne $\lambda_1 = 2$ i $\lambda_2 = 4$. Odpowiadającymi im jednostkowymi wektorami własnymi są na przykład

$$u_1 = [1/\sqrt{2}, -1/\sqrt{2}]^T, \quad u_2 = [1/\sqrt{2}, 1/\sqrt{2}]^T.$$

Dla macierzy ortogonalnej (obrót)

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \cos\pi/4 & \sin\pi/4 \\ -\sin\pi/4 & \cos\pi/4 \end{bmatrix}$$

mamy

$$Q^T AQ = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} := D.$$

Stąd, dla $z' = [x', y']^T := Q^T z$ (czyli $z = Qz'$) mamy

$$(Az|z) = (AQz'|Qz') = (Q^T AQz'|z') = (Dz'|z'),$$

więc równanie w nowym układzie współrzędnych przyjmuje postać

$$2(x')^2 + 4(y')^2 = 8,$$

czyli

$$\frac{(x')^2}{4} + \frac{(y')^2}{2} = 1.$$

Jest to więc elipsa.

**Przykład 14.10.** Rozważmy równanie

$$x^2 + \sqrt{3}xy - 3 = 0.$$

Wtedy

$$A = \begin{bmatrix} 1 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & 0 \end{bmatrix}$$

ma wartości własne $-\frac{1}{2}$ i $\frac{3}{2}$ o odpowiadających jednostkowych wektorach własnych $u_1 = [1/2, -\sqrt{3}/2]^T$, $u_2 = [\sqrt{3}/2, 1/2]^T$. Równanie tej krzywej możemy zapisać w postaci kanonicznej

$$-\frac{1}{2}(x')^2 + \frac{3}{2}(y')^2 - 3 = 0,$$

czyli

$$\frac{(x')^2}{(\sqrt{6})^2} - \frac{(y')^2}{(\sqrt{2})^2} = -1,$$

więc jest ona hiperbolą.

## 14.4 Formy kwadratowe w wyższych wymiarach

Nie ma powodu abyśmy ograniczali się wyłącznie do wymiaru 2. Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą symetryczną.

**Definicja 14.6.** Forma kwadratowa
*Formą kwadratową* nazywamy funkcję $Q_A : \mathbb{R}^n \longrightarrow \mathbb{R}$, daną przez

$$\begin{aligned} Q_A(v) &= (Av|v) \\ &= v^T A v \\ &= \sum_{i,j=1}^{n} a_{ij} x_i x_j, \quad v = [x_1, \ldots, x_n]^T \in \mathbb{R}^n. \end{aligned}$$

**Przykład 14.11.** W wymiarze 3 dla macierzy

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$$

mamy

$$Q_A[x, y, z]^T = a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz.$$

**Twierdzenie 14.6** (Postać kanoniczna formy kwadratowej)**.** *Dla macierzy symetrycznej $A \in M_{n\times n}(\mathbb{R})$ istnieje taka macierz ortogonalna $P \in O(n)$, że $P^T AP = D$ jest diagonalna oraz dla $u = P^T v$ zachodzi równość*

$$v^T Av = u^T Du$$

*czyli $Q_A(v) = Q_{P^T AP}(u) = Q_D(u)$.*

U

Niech $Q_A(x) = (Ax|x)$ będzie formą kwadratową na $\mathbb{R}^n$ dla pewnej macierzy symetrycznej $A \in M_{n\times n}(\mathbb{R})$. Jeśli $S \in M_{n\times n}(\mathbb{R})$ jest nieosobliwa (nie musi być ortogonalna) oraz $x = Sy$, to

$$\begin{aligned} Q_A(x) &= (Ax|x) \\ &= (ASy|Sy) \\ &= (S^T ASy|y) \\ &= Q_{S^T AS}(y). \end{aligned}$$

Zmiana bazy $S$ zmienia macierz formy $A$ na macierz $S^T AS$. Macierze $A$ i $S^T AS$ nie muszą być podobne.

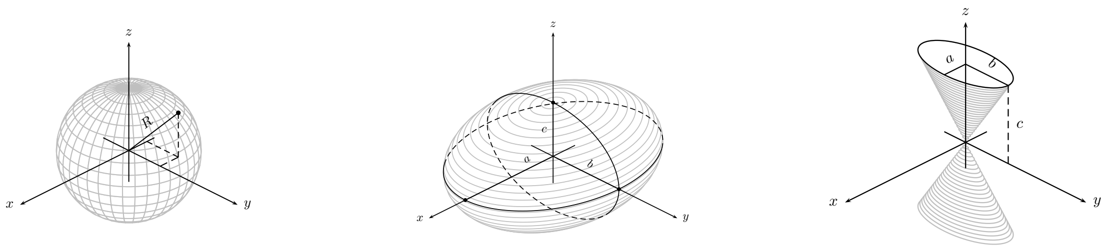
POWIERZCHNIE STOPNIA DRUGIEGO
Załóżmy, że $A = [a_{ij}] \in M_{3\times 3}(\mathbb{R})$ jest niezerową macierzą symetryczną, $b, c \in \mathbb{R}^3$. *Powierzchnią stopnia drugiego* nazywamy powierzchnię zadaną równaniem

$$(Av|v) + (b|v) + c = 0 \quad v = [x, y, z]^T$$

Jeśli $A$ jest nieosobliwa, to dokładnie tak samo jak w przypadku $n = 2$, powyższe równanie w pewnym ortonormalnym układzie współrzędnych $[x', y', z']^T$ ma postać kanoniczną

$$\lambda_1(x')^2 + \lambda_2(y')^2 + \lambda_3(z')^2 + c' = 0$$

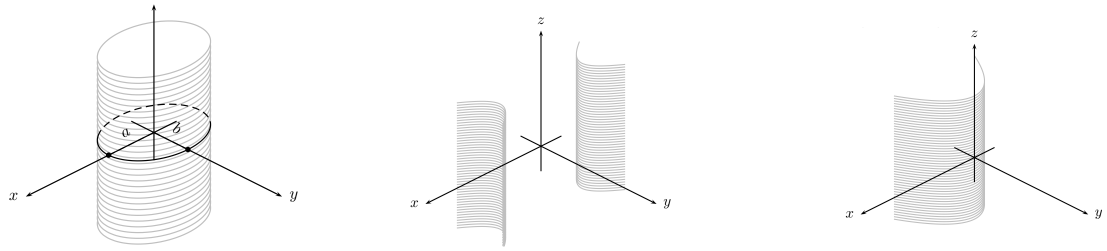
dla pewnego $c' \in \mathbb{R}^3$. Liczby $\lambda_1, \lambda_2, \lambda_3$ są wartościami własnymi macierzy $A$.




Rysunek 14.5: Po lewej sfera $x^2 + y^2 + z^2 = R^2$. W środku elipsoida $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$. Po prawej stożek eliptyczny $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$.


Rysunek 14.6: Po lewej hiperboloida jednopowłokowa $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$. Po prawej hiperboloida dwupowłokowa $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$.


Rysunek 14.7: Po lewej paraboloida eliptyczna $\frac{x^2}{a^2} + \frac{y^2}{b^2} - z = 0$. Po prawej paraboloida hiperboliczna $\frac{x^2}{a^2} - \frac{y^2}{b^2} - z = 0$.




Rysunek 14.8: Walce. W kolejności:walec eliptyczny $\frac{x^2}{a^2}+\frac{y^2}{b^2} = 1$, walec hiperboliczny $\frac{x^2}{a^2}-\frac{y^2}{b^2} = 1$, walec paraboliczny $ax^2 - y = 0$.

# Rozdział 15

# Rozkład singularny

| Słowa kluczowe rozdziału |
|---|
| liniowe twierdzenie o rzędzie $\Diamond$ wartości singularne $\Diamond$ macierz quasi diagonalna $\Diamond$ rozkład singularny |

## 15.1 Liniowe twierdzenie o rzędzie

Zajmiemy się teraz postacią macierzy $A \in M_{n\times m}(\mathbb{R})$ o rzędzie $r = \operatorname{rank} A$. Z definicji rzędu wynika, że $r \leqslant m, n$.

Dla $r \leqslant m, n$ definiujemy macierz

$$U_r = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] \in M_{n\times m}.$$

Oznacza to, że

$$U_r[x_1, \ldots, x_m]^T = [x_1, \ldots, x_r, 0, \ldots, 0]^T.$$

**Twierdzenie 15.1** (Liniowe twierdzenie o rzędzie)**.** *Dla macierzy $A \in M_{n\times m}(\mathbb{R})$ mającej rząd $r$ istnieją takie macierze nieosobliwe $B \in M_{m\times m}(\mathbb{R})$, $C \in M_{n\times n}(\mathbb{R})$, że*

$$C^{-1}AB = U_r.$$

*Dowód.* Z formuły wymiaru wynika, że $\dim\ker A = m - r$. Niech $v_1, \ldots, v_{m-r}$ będzie bazą $\ker A$. Dopełnieniem ortogonalnym dla $\ker A$ jest $\operatorname{im} A^T$. Wybieramy $u_1, \ldots, u_r$ bazę $\operatorname{im} A^T$. Rozważmy macierz nieosobliwą

$$B = \left[\begin{array}{c|c|c|c|c|c} u_1 & \ldots & u_r & v_1 & \ldots & v_{m-r} \end{array}\right].$$

Niech $w_i = Au_i$ dla $i = 1, \ldots, r$. Wtedy

$$\begin{aligned} AB &= \left[\begin{array}{c|c|c|c|c|c} Au_1 & \ldots & Au_r & Av_1 & \ldots & Av_{m-r} \end{array}\right] \\ &= \left[\begin{array}{c|c|c|c|c|c} w_1 & \ldots & w_r & 0 & \ldots & 0 \end{array}\right]. \end{aligned}$$

Wektory $w_1, \ldots, w_r$ są liniowo niezależne i leżą w $\operatorname{im} A$. Uzupełniamy je wektorami $w_{r+1}, \ldots, w_n$ do bazy $\mathbb{R}^n$. Definiujemy $C = \left[\begin{array}{c|c|c} w_1 & \ldots & w_n \end{array}\right]$. Ponieważ $C^{-1}C = I$, więc

$$C^{-1}\left[\begin{array}{c|c|c|c|c|c} w_1 & \ldots & w_r & 0 & \ldots & 0 \end{array}\right] = \left[\begin{array}{c} I_r \\ \hline 0_{(n-r)\times r} \end{array}\right],$$

czyli $C^{-1}AB = U_r$. $\square$

## 15.2 Twierdzenie o rozkładzie singularnym

Niech $A \in M_{n\times m}(\mathbb{R})$. Przypomnijmy, że wtedy

(i) $\ker A = (\operatorname{im} A^T)^\perp$ oraz $\ker A^T = (\operatorname{im} A)^\perp$. W szczególności, mamy rozkłady na ortogonalne sumy proste

$$\mathbb{R}^n = \operatorname{im} A^T \oplus \ker A, \quad \mathbb{R}^m = \ker A^T \oplus \operatorname{im} A.$$

(ii) $\operatorname{im} A^T A = \operatorname{im} A^T$ oraz $\ker A^T A = \ker A$,

(iii) $\operatorname{rank} A = \operatorname{rank} A^T A$,

(iv) wartości własne $A^T A$ są nieujemne.

Zaczniemy od pewnego szczególnego przypadku. Rozważmy macierz nieosobliwą $A \in M_{n\times n}(\mathbb{R})$. Jeśli $A$ nie ma bazy wektorów własnych, to $A$ nie jest diagonalizowalna. Sytuacja zmienia się, gdy dopuścimy wybór innej bazy w dziedzinie niż przeciwdziedzinie. Okazuje się, że wtedy istnieją takie macierze ortogonalne $U = [u_1|\ldots|u_n], V = [v_1|\ldots|v_n] \in O(n)$, że

$$V^T AU = \operatorname{diag}(\sigma_1, \ldots, \sigma_n) =: D, \quad \sigma_1 \geqslant \ldots \geqslant \sigma_1 > 0.$$

Zastanówmy się, jak znaleźć takie $U, V, D$. Równość

$$AU = VD,$$

oznacza, że $A = VDU^T$, więc

$$A^T A = (UD^T V^T)(VDU^T) = UD^T DU^T.$$

Stąd, $A^T A$ i $D^T D$ mają te same wartości własne. Ponieważ

$$D^T D = \operatorname{diag}(\sigma_1^2, \ldots, \sigma_n^2),$$

więc $\lambda_1 = \sigma_1^2 \geqslant \ldots \geqslant \lambda_n = \sigma_n^2$ są wartościami własnymi macierzy Grama $A^T A$. Kolumny macierzy $U$ tworzą bazę ortonormalną, złożoną z wektorów własnych macierzy Grama $A^T A$.

Ponieważ $\sigma_i > 0$ oraz

$$Au_i = \sigma_i \cdot v_i, \quad i = 1, \ldots, n,$$

więc $v_i = \frac{1}{\sigma_i} Au_i$. Sprawdzimy, że wektory $v_1, \ldots, v_n$ tworzą bazę ortonormalną. Mamy

$$\begin{aligned}(v_i|v_j) &= \frac{1}{\sigma_i\sigma_j}(Au_i|Au_j)\\ &= \frac{1}{\sigma_i\sigma_j}(A^T Au_i|u_j)\\ &= \frac{\sigma_i^2}{\sigma_i\sigma_j}(u_i|u_j)\\ &= \delta_{ij}.\end{aligned}$$

Zauważmy, że

$$AA^T = VDD^T V^T,$$

więc kolumny $V$ tworzą ortonormalne wektory własne macierzy $AA^T$.

**Definicja 15.1.** **Wartości singularne**

Niech $\lambda_1 \geqslant \ldots \geqslant \lambda_m \geqslant 0$ będą wartościami własnymi macierzy Grama $A^T A$. Liczby

$$\sigma_i = \sqrt{\lambda_i}, \quad i = 1 \ldots, m$$

nazywamy *wartościami singularnymi* macierzy $A$.

**Przykład 15.1.** Ponieważ rank $A = \operatorname{rank} A^T A$, więc rząd macierzy $A$ jest równy liczbie niezerowych wartości singularnych macierzy $A$. Zwróćmy uwagę, że nie jest to prawda dla wartości własnych. Przykładowo macierz

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

ma rząd 1, a nie ma niezerowych wartości własnych.

**Lemat 15.1.** *Jeśli $A \in M_{n\times n}(\mathbb{R})$ i $Q, U \in O(n)$, to macierze $A$ i $B = UAQ$ mają te same wartości singularne.*

*Dowód.* Wystarczy sprawdzić, że macierze $A^T A$ i $B^T B$ są podobne. Mamy

$$\begin{aligned} B^T B &= (Q^T A^T U^T)(UAQ) \\ &= Q^T (A^T A) Q. \end{aligned}$$

□

U Niech $\lambda_i \geqslant 0$ będzie wartością własną macierzy Grama $A^T A$ z jednostkowym wektorem własnym $w_i$. Wtedy

$$\begin{aligned} \lambda_i &= \lambda_i (w_i | w_i) \\ &= (\lambda_i \cdot w_i | w_i) \\ &= (A^T A w_i | w_i) \\ &= (A w_i | A w_i) \\ &= \|A w_i\|^2, \end{aligned}$$

czyli

$$\sigma_i = \sqrt{\lambda_i} = \|A w_i\|.$$

**Przykład 15.2.** Znajdziemy wartości singularne macierzy

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Macierz Grama

$$A^T A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

ma wartości własne $\lambda_1 = 3$ i $\lambda_2 = 1$, czyli

$$\sigma_1 = \sqrt{3}, \quad \sigma_2 = 1.$$

**Definicja 15.2.** **Macierz quasi-diagonalna**

Macierz $A \in M_{n\times m}(\mathbb{R})$ nazywamy *quasi-diagonalną*, jeśli $a_{ij} = 0$ dla $i \neq j$, czyli jej niediagonalne wyrazy są zerowe.

**Twierdzenie 15.2** (Rozkład singularny)**.** *Jeśli macierz $A \in M_{n\times m}(\mathbb{R})$ ma rząd $r$, to*

$$A = VDU^T,$$

*gdzie $U \in O(m)$, $V \in O(n)$ oraz $D \in M_{n\times m}(\mathbb{R})$ jest quasi-diagonalna postaci*

$$D = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & \sigma_r & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & & \vdots & & & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

*Wyrazy diagonalne macierzy $D$ są wartościami singularnymi $A$.*

*Dowód.* Macierz $A^TA$ jest symetryczna, więc jest ortogonalnie diagonalizowalna przy pomocy ortogonalnej bazy wektorów własnych $u_1, \ldots, u_m$. Wartości własne $\lambda_i$ są nieujemne,

$$(A^TA)u_i = \lambda_i \cdot u_i, \quad i = 1, \ldots, m.$$

Niech $r = \operatorname{rank} A$. Jeśli $r < m$, to

$$\dim \ker A = \dim \ker A^TA = m - r,$$

czyli $m - r$ ostatnich wartości własnych jest równe 0. Dla wyróżnienia wektory własne z bazy $u_1, \ldots, u_m$ które są w jądrze $A^TA$ oznaczamy przez $k_1, \ldots, k_{m-r}$. W tej konwencji bazą ortonormalną w $\mathbb{R}^m$ jest

$$u_1, \ldots, u_r, \; k_1, \ldots, k_{m-r} \in \mathbb{R}^m.$$

Wektory $u_1, \ldots, u_r$ odpowiadają niezerowym wartościom własnym.

Skonstruujemy teraz odpowiednią bazę w $\mathbb{R}^n$. Ponieważ $\operatorname{rank} A^T = \operatorname{rank} A = r$, więc

$$\dim \ker A^T = n - r.$$

Niech $k_1^*, \ldots, k_{n-r}^*$ będzie bazą ortonormalną dla $\ker A^T$. Dobieramy $v_1, \ldots, v_r$ bazę ortonormalną dla $\operatorname{im} A$ kładąc

$$v_i = \frac{1}{\sqrt{\lambda_i}} \cdot Au_i = \frac{1}{\sigma_i} \cdot Au_i, \quad i = 1, \ldots, r.$$

Jest to baza ortonormalna, bo

$$\begin{aligned}
(v_i|v_j) &= \frac{1}{\sigma_i\sigma_j}(Au_i|Au_j) \\
&= \frac{1}{\sigma_i\sigma_j}(A^TAu_i|u_j) \\
&= \frac{\sigma_i}{\sigma_j}(u_i|u_j) \\
&= \delta_{ij}.
\end{aligned}$$

Sprawdzimy, że teza zachodzi dla

$$U = \left[\; u_1 \;\middle|\; \ldots \;\middle|\; u_r \;\middle|\; k_1 \;\middle|\; \ldots \;\middle|\; k_{m-r} \;\right], \quad V = \left[\; v_1 \;\middle|\; \ldots \;\middle|\; v_r \;\middle|\; k_1^* \;\middle|\; \ldots \;\middle|\; k_{n-r}^* \;\right].$$

Z konstrukcji wynika, że

$$A^TAu_i = \lambda_i \cdot u_i, \quad Ak_i = 0, \quad Au_i = \sigma_i \cdot v_i, \quad A^Tk_j^* = 0.$$

Stąd

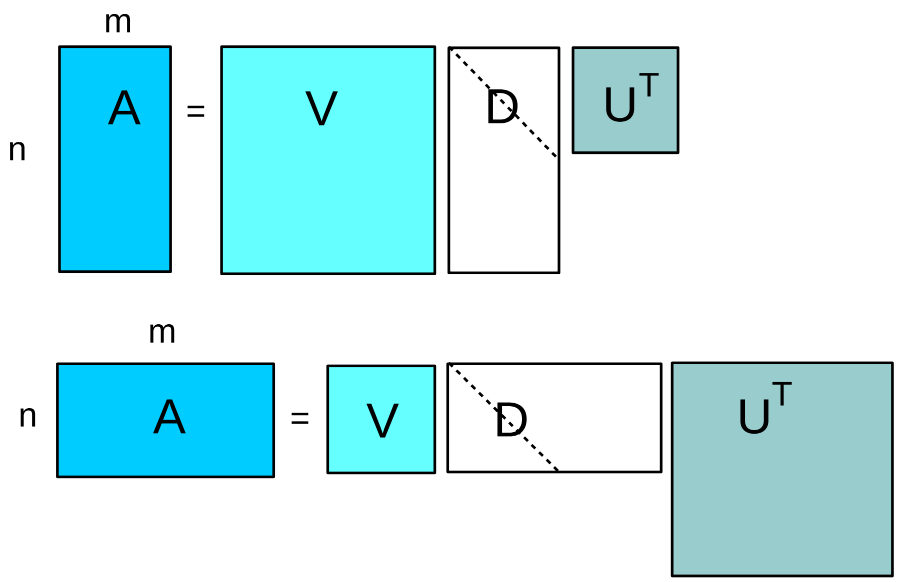
$$AU = \left[\; \sigma_i v_1 \;\middle|\; \ldots \;\middle|\; \sigma_r v_r \;\middle|\; 0 \;\middle|\; \ldots \;\middle|\; 0 \;\right] = VD,$$

dla

$$D = \begin{bmatrix}
\sigma_1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\
0 & \sigma_2 & & 0 & 0 & \ldots & 0 \\
\vdots & & \ddots & \vdots & \vdots & & \vdots \\
0 & \ldots & 0 & \sigma_r & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\
\vdots & & & \vdots & & & \vdots \\
0 & \ldots & 0 & 0 & 0 & \ldots & 0
\end{bmatrix}.$$

$\square$

Rysunek 15.1: Rozkład singularny.




U Podamy interpretację geometryczną twierdzenia 15.2. Każdą macierz $A$ możemy przedstawić w następujący sposób. Najpierw izometria $U^T$ przekształca kolumny $U$ na bazę standardową. Następnie $D$ przeskalowuje pierwszych $r$ osi w $\mathbb{R}^m$ i posyła je do $\mathbb{R}^n$. Na końcu, $V$ przekształca izometrycznie bazę standardową na kolumny $V$.

**Wniosek 15.1.** *Dla dowolnej macierzy $A \in M_{n\times m}(\mathbb{R})$ istnieje taka baza ortonormalna $u_1, \dots, u_m$ w $\mathbb{R}^m$, że*

$$(Au_i|Au_j) = 0, \quad i \neq j.$$

*Dowód.* Rozważmy rozkład singularny

$$A = VDU^T.$$

z $U = \left[\ u_1 \mid \dots \mid u_m\ \right]$. Stąd $AU = VD$, czyli

$$\begin{aligned} Au_j &= (AU)e_j \\ &= (VD)e_j \\ &= V(De_j) \\ &= V(\sigma_j e_j) \\ &= \sigma_j V(e_j). \end{aligned}$$

Ponieważ $V$ jest ortogonalna, więc jej kolumny są ortonormalne. Stąd

$$(Au_i|Au_j) = \sigma_i\sigma_j(V(e_i)|V(e_j)) = 0, \quad i \neq j.$$

$\square$

**Wniosek 15.2.** *Załóżmy, że $A \in M_{n\times n}(\mathbb{R})$ jest kwadratowa. Wtedy istnieją takie macierze ortogonalne, $U = \left[\ u_1 \mid \dots \mid u_n\ \right], V = \left[\ v_1 \mid \dots \mid v_n\ \right] \in O(n)$, że*

$$A = U\,\mathrm{diag}(\sigma_1, \dots, \sigma_n)V^T, \quad VAU^T = \mathrm{diag}(\sigma_1, \dots, \sigma_n).$$

*Oznacza to, że $A$ w bazach ortonormalnych $v_1, \dots, v_n$ i $u_1, \dots, u_n$ ma postać diagonalną.*

**Wniosek 15.3.** *Jeśli $A \in M_{n\times m}(\mathbb{R})$ $(n \geqslant m)$, to istnieją takie macierze $U \in O(n)$, $V \in O(m)$ oraz $D \in M_{n\times m}(\mathbb{R})$ postaci*

$$D = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \sigma_n \\ 0 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & 0 & 0 \end{bmatrix},$$

*że*

$$A = UDV^T.$$

U

(1) Wartości singularne $\sigma_1, \dots, \sigma_m$ macierzy $A \in M_{n\times m}(\mathbb{R})$ są wyznaczone jednoznacznie, ale macierze ortogonalne $U$ i $V$ już nie.

(2) Z dowodu twierdzenia 15.2 wynika, że $V = \left[\ v_1 \mid \dots \mid v_m\ \right]$ diagonalizuje $AA^T$, więc $v_i$ są wektorami własnymi $AA^T$.

(3) $U = \left[\ u_1 \mid \dots \mid u_n\ \right]$ diagonalizuje $A^TA$ oraz $u_j$ są wektorami własnymi $A^TA$.

(4) Z dowodu twierdzenia 15.2 wynika, że jeśli $\operatorname{rank} A = r$, to

**(a)** $v_1, \dots, v_r$ tworzą bazę ortonormalną $\operatorname{im} A$,

**(b)** $v_{r+1}, \dots, v_n$ tworzą bazę ortonormalną $\ker A^T$,

**(c)** $u_1, \dots, u_r$ tworzą bazę ortonormalną $\operatorname{im} A^T$,

**(d)** $u_{r+1}, \dots, u_m$ tworzą bazę ortonormalną $\ker A$.

**Wniosek 15.4.** *Przy powyższych oznaczeniach mamy*

$$A = \sigma_1 \cdot v_1u_1^T + \dots + \sigma_r \cdot v_ru_r^T.$$

*Dowód.* Przy oznaczeniach twierdzenia 15.2 mamy $A = VDU^T$, więc

$$\begin{aligned} A &= VDU^T \\ &= \left[\ v_1 \mid \dots \mid v_r \mid v_{r+1} \mid \dots \mid v_m\ \right] \left[\begin{array}{c|c} \operatorname{diag}(\sigma_1, \dots, \sigma_r) & 0 \\ \hline 0 & 0 \end{array}\right] \left[\ u_1^T \mid \dots \mid u_r^T \mid u_{r+1}^T \mid \dots \mid u_m^T\ \right] \\ &= \left[\ v_1 \mid \dots \mid v_r\ \right] \operatorname{diag}(\sigma_1, \dots, \sigma_r) \left[\ u_1^T \mid \dots \mid u_r^T\ \right] \\ &= \sigma_1 \cdot v_1u_1^T + \dots + \sigma_r \cdot v_ru_r^T. \end{aligned}$$

□

**Przykład 15.3.** Znajdziemy rozkład singularny macierzy

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Macierz Grama

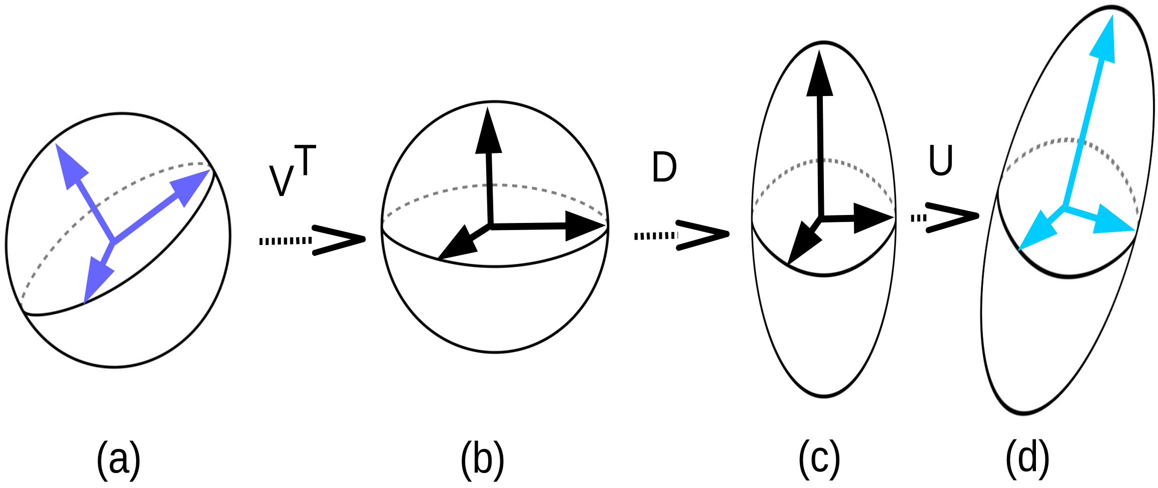
$$A^TA = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

ma wartości własne $\lambda_1 = 2$, $\lambda_2 = 1$ oraz $\lambda_3 = 0$ z odpowiadającymi jednostkowymi wektorami własnymi

$$u_1 = [1/\sqrt{2}, 1/\sqrt{2}, 0]^T, \quad u_2 = [0, 0, 1]^T, \quad u_3 = [-1/\sqrt{2}, 1/\sqrt{2}, 0]^T.$$

Wartości singularne są dane przez

$$\sigma_1 = \sqrt{2}, \quad \sigma_2 = 1, \quad \sigma_3 = 0.$$




Rysunek 15.2: Rozkład singularny $A = UDV^T$ macierzy nieosobliwej $A \in M_{3\times3}(\mathbb{R})$. W punkcie (a) baza ortonormalna $v_1, v_2, v_3$ (kolumny $V$). Poprzez $V^T$ przechodzą one na bazę standardową $e_1, e_2, e_3$ w punkcie (b). Macierz diagonalna $D$ przekształca bazę standardową na bazę $\sigma_1 \cdot e_1, \sigma_2 \cdot e_2, \sigma_3 \cdot e_3$ w punkcie (c), która poprzez $U$ przechodzi na bazę $\sigma_1 \cdot u_1, \sigma_2 \cdot u_2, \sigma_3 \cdot u_3$ w punkcie (d). Obrazem sfery jednostkowej jest elipsoida.

Mamy

$$U = \left[\begin{array}{c|c|c} u_1 & u_2 & u_3 \end{array}\right] = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 \end{bmatrix}$$

oraz

$$D = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Macierz $V = \left[\begin{array}{c|c} v_1 & v_2 \end{array}\right]$, gdzie

$$v_1 = \frac{1}{\sigma_1} A u_1 = [1, 0]^T, \quad v_2 = \frac{1}{\sigma_2} A v_2 = [0, 1]^T,$$

czyli $V = I_2$. W efekcie,

$$A = VDU^T.$$

**Przykład 15.4.** Przyglądnijmy się rozkładowi singularnemu macierzy nieosobliwej $A \in M_{3\times3}(\mathbb{R})$. Wtedy

$$\sigma_1 \geqslant \sigma_2 \geqslant \sigma_3 > 0$$

oraz istnieją takie bazy ortonormalne $v_1, v_2, v_3$ oraz $u_1, u_2, u_3$, że

$$A = \underbrace{\left[\begin{array}{c|c|c} u_1 & u_2 & u_3 \end{array}\right]}_{U} \underbrace{\operatorname{diag}(\sigma_1, \sigma_2, \sigma_3)}_{D} \underbrace{\left[\begin{array}{c|c|c} v_1 & v_2 & v_3 \end{array}\right]^T}_{V^T}.$$

Ponieważ $V^T = V^{-1}$ oraz $V(e_i) = v_i$, więc

$$\begin{aligned} A(v_i) &= UDV^T(v_i) \\ &= UDe_i \\ &= U(\sigma_i \cdot e_i) \\ &= \sigma_i \cdot u_i. \end{aligned}$$

Rozważmy teraz sferę

$$\mathbb{S}^2 = \{[x_1, x_2, x_3]^T : x_1^2 + x_2^2 + x_3^2 = 1\}.$$

Zobaczmy jak wygląda jej obraz poprzez $U^T A$. Niech $y = Ax$ dla pewnego $x \in \mathbb{S}^2$. Ponieważ $x = A^{-1}y$ oraz $U$ i $V$ są ortogonalne, więc dla $w = U^T y$ mamy

$$\begin{aligned} 1 = \|x\| &= \|A^{-1}Ax\| \\ &= \|A^{-1}y\| \\ &= \|VD^{-1}U^T y\| \\ &= \|D^{-1}w\| \\ &= \frac{w_1^2}{\sigma_1^2} + \frac{w_2^2}{\sigma_2^2} + \frac{w_3^2}{\sigma_3^2}. \end{aligned}$$

Wynika stąd, że $U^T A(\mathbb{S}^2)$ jest elipsoidą, której półosie mają długości $\sigma_1, \sigma_2, \sigma_3$ oraz są one skierowane wzdłuż kierunków bazy standardowej $e_1, e_2, e_3$. Ponieważ $U$ przekształca $U^T A(\mathbb{S}^2)$ na $A(\mathbb{S}^2)$ oraz $Ue_i = u_i$, więc osie elipsoidy $A(\mathbb{S}^2)$ są skierowane wzdłuż bazy $u_1, u_2, u_3$ i są równe $\sigma_i \cdot u_i = A(v_i)$. Zauważmy, że

$$\begin{aligned} \max_{x\in\mathbb{S}^2} \|Ax\| &= \max_{x\in\mathbb{S}^2} \|UDV^T x\| \\ &= \max_{y\in\mathbb{S}^2} \|Dy\| \\ &= \sigma_1. \end{aligned}$$

Analogicznie,

$$\min_{x\in\mathbb{S}^2} \|Ax\| = \sigma_3.$$

**Przykład 15.5.** Niech $A \in M_{m\times n}(\mathbb{R})$ ma rząd $r = \operatorname{rank} A$. Zobaczymy jak wygląda obraz sfery jednostkowej $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$ pod wpływem działania $A$. Rozważmy rozkład singularny

$$A = VDU^T,$$

gdzie

$$U = \left[\begin{array}{c|c|c} u_1 & \dots & u_n \end{array}\right] \in O(n), \quad V = \left[\begin{array}{c|c|c} v_1 & \dots & v_m \end{array}\right] \in O(m)$$

z wartościami singularnymi

$$\sigma_1 \geqslant \ldots \geqslant \sigma_r > 0, \quad \sigma_{r+1} = \ldots = \sigma_n = 0.$$

Niech $x = [x_1, \ldots, x_n]^T$ będzie jednostkowym wektorem w $\mathbb{R}^n$. Ponieważ $U \in O(n)$, więc $\|U^T x\| = \|x\| = 1$, czyli

$$(u_1^T x)^2 + \ldots + (u_n^T x)^2 = 1.$$

Ponieważ

$$A = \sigma_1 v_1 u_1^T + \ldots + \sigma_r v_r u_r^T,$$

więc

$$\begin{aligned} Ax &= \sigma_1 v_1 u_1^T x + \ldots + \sigma_r v_r u_r^T x \\ &= \underbrace{(\sigma_1 u_1^T x)}_{=y_1} \cdot v_1 + \ldots + \underbrace{(\sigma_r u_r^T x)}_{=y_r} \cdot v_r \\ &= y_1 \cdot v_1 + \ldots + y_r \cdot v_r. \end{aligned}$$

Jeśli $r = n$, to $n \leqslant m$ oraz

$$Ax = y_1 \cdot v_1 + \ldots + y_n \cdot v_n = Vy, \quad y = [y_1, \ldots, y_n]^T.$$

Stąd,

$$\|Ax\| = \|Vy\| = \|y\|,$$

oraz

$$\left(\frac{y_1}{\sigma_1}\right)^2 + \ldots + \left(\frac{y_n}{\sigma_n}\right)^2 = 1.$$

Jeśli $r < n$, to podobne rozumowanie pokazuje, że

$$\left(\frac{y_1}{\sigma_1}\right)^2 + \ldots + \left(\frac{y_r}{\sigma_r}\right)^2 \leqslant 1.$$

## 15.3 Inaczej o rozkładzie singularnym

Podamy teraz alternatywny dowód twierdzenia 15.2. Niech $A \in M_{n\times m}(\mathbb{R})$ ma rząd $r \geqslant 1$. Ponieważ sfera

$$\mathbb{S}^{m-1} = \{u \in \mathbb{R}^m : \|u\| = 1\}$$

jest zbiorem zwartym i odwzorowanie $\mathbb{R}^m \ni u \longmapsto Au \in \mathbb{R}^n$ jest ciągłe, więc istnieje taki wektor $u_1 \in \mathbb{S}^{m-1}$, że

$$\|Au_1\| = \max_{u\in\mathbb{S}^{m-1}} \|Au\|.$$

Ponieważ $r \geqslant 1$, więc $\sigma_1 := \|Au_1\| > 0$. Definiujemy

$$v_1 := \frac{Au_1}{\|Au_1\|} = \frac{Au_1}{\sigma_1}.$$

Niech

$$\mathbb{S}_{m-1} := \mathbb{S}^{m-1} \cap u_1^{\perp}$$

będzie sferą jednostkową w dopełnieniu ortogonalnym prostej span$\{u_\}$. Istnieje taki wektor $u_2 \in \mathbb{S}_{m-1}$, że

$$\|Au_2\| = \max_{u\in\mathbb{S}_{m-1}} \|Au\|.$$

Jeśli $\|Au_2\| > 0$ (czyli $r \geqslant 2$), to definiujemy

$$\sigma_2 := \|Au_2\|, \quad v_2 := \frac{Au_2}{\|Au_2\|} = \frac{Au_2}{\sigma_2}.$$

Wtedy $\sigma_1 \geqslant \sigma_2 > 0$. Z konstrukcji wynika, że $u_1, u_2$ są ortonormalne. Nie jest oczywiste, że $v_1, v_2$ są ortonormalne.

**Lemat 15.2.** *Jeśli* $\|Au_1\| = \max_{u\in\mathbb{S}^{m-1}} \|Au\| > 0$ *i* $v_1 = \frac{Au_1}{\|Au_1\|}$, *to*

$$A(\mathbb{S}_{m-1}) \subset v_1^{\perp}.$$

*W szczególności,* $v_1, v_2$ *są ortonormalne.*

*Dowód.* Niech $u \in \mathbb{S}_{m-1}$. W szczególności, $\|u\| = 1$ oraz $(u|u_1) = 0$. Dla $t \in \mathbb{R}$ definiujemy

$$u(t) = (\sin t) \cdot u + (\cos t) \cdot u_1.$$

Zauważmy, że $u(0) = u_1$ oraz

$$\begin{aligned} \|u(t)\|^2 &= (u(t)|u(t)) \\ &= \sin^2 t \underbrace{(u|u)}_{=1} + 2\sin t\cos t \underbrace{(u|u_1)}_{=0} + \cos^2 t \underbrace{(u_1|u_1)}_{=1} \\ &= 1. \end{aligned}$$

Rozważmy funkcję klasy $C^\infty$ daną przez

$$f(t) = \|Au(t)\|^2 = (Au(t)|Au(t)), \quad t \in \mathbb{R}.$$

Funkcja $f$ osiąga maksimum dla $t = 0$, bo $f(0) = \|Au_1\|$. Stąd, $f'(0) = 0$. Z drugiej strony,

$$\begin{aligned} f'(t) &= 2(Au'(t)|Au(t)) \\ &= \sin 2t(\|Au\|^2 - \|Au_1\|^2) + 2\cos t(Au_1|Au). \end{aligned}$$

Ponieważ $f'(0) = 0$, więc $(Au_1|Au) = 0$, czyli $(v_1|Au_1) = 0$. $\square$

U Zauważmy, że okrąg jednostkowy w płaszczyźnie $\operatorname{span}\{u_1, u_2\}$ jest równy zbiorowi

$$\{\cos t \cdot u_1 + \sin t \cdot u_2 : t \in \mathbb{R}\}.$$

Ponieważ

$$A(\cos t \cdot u_1 + \sin t \cdot u_2) = \sigma_1 \cos t \cdot v_1 + \sigma_2 \sin t \cdot v_2,$$

więc obrazem tego okręgu jest elipsa w płaszczyźnie $\operatorname{span}\{v_1, v_2\}$ o półosiach długości $\sigma_1, \sigma_2$.

Jak dotąd znaleźliśmy takie wektory ortonormalne $u_1, u_2 \in \mathbb{R}^m$ i $v_1, v_2 \in \mathbb{R}^n$, że

$$\begin{aligned} A[u_1|u_2] &= [Au_1|Au_2] \\ &= [\sigma_1 \cdot v_1|\sigma_2 \cdot v_2] \\ &= [v_1|v_2] \operatorname{diag}(\sigma_1, \sigma_2), \end{aligned}$$

czyli

$$AU = V \operatorname{diag}(\sigma_1, \sigma_2).$$

Rozszerzymy teraz indukcyjnie macierze $U = [u_1|u_2] \in M_{m\times 2}(\mathbb{R})$ i $V = [v_1|v_2] \in M_{n\times 2}(\mathbb{R})$ do macierzy kwadratowych. Załóżmy, że

$$AU = V \operatorname{diag}(\sigma_1, \dots, \sigma_k), \quad k < r$$

gdzie $\sigma_1 \geqslant \dots \geqslant \sigma_k > 0$. Ponieważ $k < r = \operatorname{rank} A$, więc $(\operatorname{span}\{u_1, \dots, u_k\})^\perp \neq \{0\}$. Niech

$$\mathbb{S}_{m-k} = \mathbb{S}^{m-1} \cap (\operatorname{span}\{u_1, \dots, u_k\})^\perp$$

będzie sferą jednostkową w dopełnieniu ortogonalnym do obrazu $U$. Istnieje taki wektor $u_{k+1} \in \mathbb{S}_{m-k}$, że

$$\|Au_{k+1}\| = \max_{u \in \mathbb{S}_{m-k}} \|Au\|.$$

Ponieważ $k < r$, więc $\|Au_{k+1}\| > 0$. Przyjmujemy, że

$$\sigma_{k+1} = \|Au_{k+1}\|, \quad v_{k+1} = \frac{Au_{k+1}}{\|A_{k+1}\|} = \frac{Au_{k+1}}{\sigma_{k+1}}.$$

Wtedy

$$A[u_1|\dots|u_k|u_{k+1}] = [v_1|\dots|v_k|v_{k+1}] \operatorname{diag}(\sigma_1, \dots, \sigma_k, \sigma_{k+1}).$$

Wektory $u_1, \dots, u_{k+1}$ są ortonormalne. Pozostaje sprawdzić, że $v_1, \dots, v_{k+1}$ są ortonormalne. Dla ustalonego $j = 1, \dots, k$ stosujemy rozumowanie z dowodu lematu 15.2 do okręgu jednostkowego w płaszczyźnie $\operatorname{span}\{u_j, u\}$ dla $u \in \mathbb{S}_{m-k}$ pokazujemy, że $Au_{k+1} \perp v_j$, czyli $v_{k+1} \perp v_j$.

Możemy tę procedurę kontynuować, aż otrzymamy macierze $U_1 = [u_1|\dots|u_r$ i $V_1 = [v_1, \dots, v_r]$ o $r$ ortonormalnych kolumnach:

$$AU_1 = V_1 \underbrace{\operatorname{diag}(\sigma_1, \dots, \sigma_r)}_{=\Sigma}, \quad \sigma_1 \geqslant \dots \geqslant \sigma_r > 0.$$

Niech $U_2$ będzie macierzą o $m-r$ kolumnach będących bazą ortonormalną dla $(\operatorname{span}\{u_1, \dots, u_r\})^\perp$, a $V_2$ macierzą o $n-r$ ortonormalnych kolumnach, będących bazą ortonormalną dla $(\operatorname{span}\{v_1, \dots, v_r\})^\perp$. Ponieważ $\sigma_{r+1} = 0$, więc $AV_2 = 0$, czyli

$$A[U_1|U_2] = [V_1|V_2] \left[\begin{array}{c|c} \Sigma & 0 \\ \hline 0 & 0 \end{array}\right]$$

Z dowodu wynika, że

$$\begin{aligned} \sigma_1 &= \max_{\|x\|=1} \|Ax\| \\ \sigma_2 &= \max_{\substack{\|x\|=1 \\ x \in (\operatorname{span}\{u_1\})^\perp}} \|Ax\| \\ \sigma_3 &= \max_{\substack{\|x\|=1 \\ x \in (\operatorname{span}\{u_1, u_2\})^\perp}} \|Ax\| \\ &\vdots \\ \sigma_r &= \max_{\substack{\|x\|=1 \\ x \in (\operatorname{span}\{u_1, \dots, u_r\})^\perp}} \|Ax\|. \end{aligned}$$

**Twierdzenie 15.3** (Minimax dla wartości singularnych)**.** *Niech $A \in M_{n\times m}(\mathbb{R})$. Jeśli $S_{k-1}$ jest zbiorem podprzestrzeni wektorowych w $\mathbb{R}^m$ wymiaru $m-k+1$ dla $k = 2, \ldots, r$, to*

$$\sigma_k = \min_{V\in S_{k-1}} \max_{\substack{\|x\|=1\\ x\in V}} \|Ax\|.$$

*Dowód.* Niech $V \in S_{k-1}$ będzie ustalona. Wtedy

$$V \cap \operatorname{span}\{u_1, \ldots, u_k\} \neq \{0\},$$

bo w przeciwnym razie

$$\dim(V + \operatorname{span}\{u_1, \ldots, u_k\}) = m - k + 1 + k = m + 1.$$

Wybieramy $x \in V \cap \operatorname{span}\{u_1, \ldots, k\}$ o normie jeden. Wtedy

$$x = (x|u_1)\cdot u_1 + \ldots + (x|u_k)\cdot u_k, \quad \sum_{i=1}^{k}(x|u_i)^2 = 1.$$

Stąd

$$\begin{aligned}\|Ax\|^2 &= (Ax|Ax)\\ &= \sum_{i=1}^{k}(x|u_i)^2\sigma_i^2\\ &\geqslant \left(\sum_{i=1}^{k}(x|u_i)^2\right)\sigma_k^2\\ &= \sigma_k^2,\end{aligned}$$

czyli $\max_{\|x\|=1,\, x\in V}\|Ax\| \geqslant \sigma_k$. Z dowolności wyboru $V$ wynika, że

$$\sigma_k \leqslant \min_{V\in S_{k-1}} \max_{\substack{\|x\|=1\\ x\in V}} \|Ax\|.$$

Minimum jest osiągane dla

$$V = \operatorname{span}\{u_k, \ldots, u_m\} \quad \text{oraz} \quad x = u_k.$$

□

U Z równości

$$V^T A U = D$$

otrzymujemy, że

$$\operatorname{diag}(\sigma_1^2, \ldots, \sigma_r^2, \underbrace{0, \ldots, 0}_{m-r}) = D^T D = U^T A^T A U,$$

więc wartości singularne są wartościami własnymi macierzy Grama $A^T A$.

# Rozdział 16

# Macierze hermitowskie i unitarne.

SŁOWA KLUCZOWE ROZDZIAŁU
iloczyn hermitowski ◊ sprzężenie hermitowskie $A^*$ ◊ macierz hermitowska ◊ wartości i wektory własne macierzy hermitowskiej ◊ macierz unitarna ◊ twierdzenie Schura ◊ diagonalizacja unitarna macierzy hermitowskiej ◊ macierz normalna ◊ diagonalizacja unitarna macierzy normalnej ◊ rzeczywiste macierze normalne

**Definicja 16.1.** **Iloczyn hermitowski**
Iloczynem hermitowskim nazywamy odwzorowanie

$$(\cdot|\cdot)_* : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}$$

dane dla wektorów $u = [u_1, \ldots, u_n]^T, v = [v_1, \ldots, v_n]^T \in \mathbb{C}^n$ wzorem

$$(u|v)_* = u_1\overline{v_1} + \ldots + u_n\overline{v_n}$$

Powiemy, że wektory $u$ i $v$ są *ortogonalne* (względem iloczynu hermitowskiego), gdy

$$(u|v)_* = 0.$$

Zauważmy, że dla wektorów rzeczywistych $u, v \in \mathbb{R}^n$ mamy

$$\begin{aligned}(u|v)_* &= u_1\overline{v_1} + \ldots + u_n\overline{v_n}\\ &= u_1v_1 + \ldots + u_nv_n\\ &= (u|v).\end{aligned}$$

*Norma wektora* $z = [z_1, \ldots, z_n]^T \in \mathbb{C}^n$ jest określona przez

$$\begin{aligned}\|z\|_*^2 &:= (z|z)_*\\ &= z_1\overline{z_1} + \ldots + z_n\overline{z_n}\\ &= |z_1|^2 + \ldots + |z_n|^2 \geqslant 0.\end{aligned}$$

**Wniosek 16.1.** *Dla dowolnych wektorów $u, v, w \in \mathbb{C}^n$ i skalarów $a, b \in \mathbb{C}$ spełnione są warunki:*

*(1) (hermitowska dwuliniowość)*

$$(a \cdot u + b \cdot v|w)_* = a(u|w)_* + b(v|w)_*$$

$$(u|a \cdot v + b \cdot w)_* = \overline{a}(u|v)_* + \overline{b}(u|w)_*$$

*(2) (hermitowska symetryczność)*

$$(u|v)_* = \overline{(v|u)_*}$$

*(3) (dodatnia określoność)*

$$(u|u)_* \geqslant 0$$

*przy czym* $(u|u)_* = 0$ *wtedy i tylko wtedy, gdy* $u = 0$.

**Lemat 16.1.** *Dla wektorów* $u, v \in \mathbb{C}^n$ *zachodzą:*

*(i) (norma sumy)*

$$\|u+v\|_*^2 = \|u\|_*^2 + 2\mathrm{Re}\,(u|v)_* + \|v\|_*^2.$$

*(ii) (nierówność Cauchy'ego–Schwarza)*

$$|(u|v)_*| \leqslant \|u\|_* \|v\|_*.$$

*(iii) (nierówność trójkąta)*

$$\|u+v\|_* \leqslant \|u\|_* + \|v\|_*.$$

*(iv) (wzory polaryzacyjne)*

$$4(u|v)_* = \|u+v\|_*^2 - \|u-v\|_*^2 + i\|u+iv\|_*^2 - i\|u-iv\|_*^2$$

$$2(u|v)_* = (1+i)(\|u\|_*^2 + \|v\|_*^2) - \|u-v\|_*^2 - i\|u-iv\|_*^2$$

*Dowód.* Dla dowodu punktu (i) zauważmy, że

$$\begin{aligned}
\|u+v\|_*^2 &= (u+v|u+v)_* \\
&= (u|u)_* + (u|v)_* + (v|u)_* + (v|v)_* \\
&= \|u\|_*^2 + (u|v)_* + \overline{(u|v)_*} + \|v\|_*^2,
\end{aligned}$$

co kończy dowód, bo $(u|v)_* + \overline{(u|v)_*} = 2\mathrm{Re}\,(u|v)_*$.

Niech $u, v \in \mathbb{C}^n$. Możemy założyć, że $u \neq 0$, bo dla $u = 0$ nierówność jest prawdziwa. Dla dowolnego $t \in \mathbb{C}$ mamy

$$\begin{aligned}
0 \leqslant \|tu+v\|_*^2 &= \|tu\|_*^2 + 2\mathrm{Re}\,(tu|v)_* + \|v\|_*^2 \\
&= |t|^2\|u\|_*^2 + 2\mathrm{Re}\,t(u|v)_* + \|v\|_*^2.
\end{aligned}$$

Wstawiając $t = \frac{-\overline{(u|v)_*}}{\|u\|_*^2}$ dostajemy, że

$$0 \leqslant \frac{|(u|v)_*|^2}{\|u\|_*^2} - 2\frac{|(u|v)_*|^2}{\|u\|_*^2} + \|v\|_*^2,$$

co dowodzi, że $|(u|v)_*| \leqslant \|u\|_* \|v\|_*$.

Punkt (iii) wynika z (ii), bo

$$\begin{aligned}
\|u+v\|_*^2 &= (u+v|u+v)_* \\
&= \|u\|_*^2 + (u|v)_* + (v|u)_* + \|v\|_*^2 \\
&\leqslant \|u\|_*^2 + 2\|u\|_*\|v\|_* + \|v\|_*^2 \\
&= (\|u\|_* + \|v\|_*)^2.
\end{aligned}$$

Punkt (iv) wynika z bezpośrednich rachunków i jest pozostawiony jako ćwiczenie. □

**U** Niech $v_1, \ldots, v_n \in \mathbb{C}^n$ będzie dowolną bazą. Stosując odpowiednik procedury Grama–Schmidta z przypadku rzeczywistego (zastępując euklidesowy iloczyn skalarny iloczynem hermitowskim), otrzymujemy taką bazę ortonormalną $u_1, \ldots, u_n$, że

$$\mathrm{span}\{v_1, \ldots, v_k\} = \mathrm{span}\{u_1, \ldots, u_k\}, \quad 1 \leqslant k \leqslant n.$$

**Lemat 16.2** (Nierówność Bessela)**.** *Niech $u_1, \ldots, u_m \in \mathbb{C}^n$ będzie układem wektorów ortonormalnych. Dla $v \in \mathbb{C}^n$ mamy*

$$\sum_{i=1}^{m} \left|(v|u_i)_*\right|^2 \leqslant \|v\|_*^2.$$

*Ponadto równość zachodzi wtedy i tylko wtedy, gdy $v \in \operatorname{span}\{u_1, \ldots, u_m\}$.*

*Dowód.* Niech $v \in \mathbb{C}^n$. Dla

$$w = v - \sum_{i=1}^{m} (v|u_i)_* \cdot u_i$$

mamy

$$\begin{aligned}
(w|w)_* &= \Big(v - \sum_{i=1}^{m} (v|u_i)_* \cdot u_i \Big| v - \sum_{i=1}^{m} (v|u_i)_* \cdot u_i\Big)_* \\
&= (v|v)_* - \Big(\sum_{i=1}^{m} (v|u_i)_* \cdot u_i \Big| v\Big)_* - \Big(v\Big|\sum_{i=1}^{m} (v|u_i)_* \cdot u_i\Big)_* + \Big(\sum_{i=1}^{m} (v|u_i)_* \cdot u_i \Big| \sum_{i=1}^{m} (v|u_i)_* \cdot u_i\Big)_* \\
&= \|v\|_*^2 - \sum_{i=1}^{m} (v|u_i)_* (u_i|v)_* - \sum_{i=1}^{m} \overline{(v|u_i)_*} (v|u_i)_* + \sum_{i,j=1}^{m} (v|u_i)_* \overline{(v|u_j)_*} (u_i|u_j)_* \\
&= \|v\|_*^2 - \sum_{i=1}^{m} |(v|u_i)_*|^2.
\end{aligned}$$

Ponieważ $(w|w)_* \geqslant 0$, więc nierówność zachodzi.

Przypuśćmy, że zachodzi równość. Wtedy $(w|w)_* = 0$, czyli $w = 0$, więc $v \in \operatorname{span}\{u_1, \ldots, u_m\}$. Z drugiej strony, jeśli $v = \sum_{i=1}^{m} x_i \cdot u_i \in \operatorname{span}\{u_1, \ldots, u_n\}$, to $x_i = (v|u_i)_*$. W efekcie

$$(v|v)_* = \sum_{i=1}^{m} |(v|u_i)_*|^2.$$

□

Rozważmy macierz $M = [m_{ij}] \in M_{m \times n}(\mathbb{C})$, gdzie $m_{ij} = a_{ij} + i b_{ij}$. Dla $A = [a_{ij}], B = [b_{ij}] \in M_{m \times n}(\mathbb{R})$ mamy

$$M = A + iB.$$

Definiujemy

$$\overline{M} = A - iB.$$

**Definicja 16.2.** **Macierz hermitowska**

Dla macierzy $M = [m_{ij}] \in M_{m \times n}(\mathbb{C})$ definiujemy jej *sprzężenie hermitowskie* przez

$$M^* = \overline{M^T} = \overline{M}^T.$$

Mówimy, że $M \in M_{n \times n}(\mathbb{C})$ jest *hermitowska* (*samosprzężona*) , jeśli $M^* = M$.

**Przykład 16.1.** Jeśli $A \in M_{n \times n}(\mathbb{R})$ jest symetryczna, to $A$ jest hermitowska.

**Przykład 16.2.** Macierz $A = [a_{ij}] \in M_{2 \times 2}(\mathbb{C})$ jest hermitowska, gdy

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} \\ \overline{a_{12}} & \overline{a_{22}} \end{bmatrix} = A^*.$$

Jeśli $A = [a_{ij}] \in M_{2 \times 2}(\mathbb{C})$ jest hermitowska, to $a_{11}, a_{22} \in \mathbb{R}$ oraz $a_{12} = \overline{a_{21}}$. Macierz hermitowska $A \in M_{2 \times 2}(\mathbb{C})$ ma więc postać

$$A = \begin{bmatrix} a & z \\ \overline{z} & b \end{bmatrix}, \quad a, b \in \mathbb{R}, \quad z \in \mathbb{C}.$$

**Wniosek 16.2.** *Jeśli $M \in M_{n\times n}(\mathbb{C})$ jest hermitowska, to elementy diagonalne $m_{11}, \dots, m_{nn}$ są liczbami rzeczywistymi.*

**Wniosek 16.3.** *Zachodzą własności*

(I) $(A^*)^* = A$,

(II) $(\alpha A + \beta B)^* = \overline{\alpha} A^* + \overline{\beta} B^*$,

(III) $(AC)^* = C^* A^*$.

naprawdę przecinek?

**U** Niech $u, v \in \mathbb{C}^n$. Wektory $u, v$ możemy traktować jako macierze należące do $M_{n\times 1}(\mathbb{C})$. Wtedy $u^T \in M_{1\times n}(\mathbb{C})$. Przy tych oznaczeniach mamy

$$(u|v)_* = u^T \overline{v}.$$

**Lemat 16.3.** *Niech $A \in M_{m\times n}(\mathbb{C})$. Wtedy $A^* \in M_{n\times m}(\mathbb{C})$ jest jedyną taką macierzą z $M_{n\times m}(\mathbb{C})$, że*

$$(Av|w)_* = (v|A^*w)_*, \quad v \in \mathbb{C}^n, w \in \mathbb{C}^m.$$

*Dowód.* Mamy

$$\begin{aligned}(Av|w)_* &= (Av)^T \overline{w} \\ &= v^T A^T \overline{w},\end{aligned}$$

oraz

$$\begin{aligned}(v|A^*w)_* &= v^T \overline{A^* w} \\ &= v^T \overline{\overline{A^T}w} \\ &= v^T A^T \overline{w}.\end{aligned}$$

Przypuśćmy, że $B \in M_{n\times m}(\mathbb{C})$ jest taka, że

$$(Av|w)_* = (v|Bw)_*, \quad v \in \mathbb{C}^n, \quad w \in \mathbb{C}^m.$$

Niech $w \in \mathbb{C}^w$ będzie dowolnie ustalone. Wtedy $(v|A^*w)_* = (v|Bw)_*$ dla wszystkich $v \in \mathbb{C}^n$, czyli

$$(v|A^*w - Bw)_* = 0 \quad \text{dla wszystkich} \quad v \in \mathbb{C}^n.$$

Dla $v = A^*w - Bw$ mamy $(A^*w - Bw|A^*w - Bw)_* = 0$, czyli $A^*w = Bw$. □

**Lemat 16.4.** *Niech $A \in M_{n\times n}(\mathbb{C})$. Jeśli $\lambda \in \mathbb{C}$ jest wartością własną $A$, to $\overline{\lambda}$ jest wartością własną $A^*$.*

*Dowód.* Zauważmy, że

$$\begin{aligned}0 &= \det(A - \lambda I) \\ &= \det(A^T - \lambda I) \\ &= \overline{\det(A^* - \overline{\lambda} I)},\end{aligned}$$

więc $\det(A^* - \overline{\lambda} I)$. □

**U** Niech

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

Wektorem własnym dla wartości własnej $\lambda = 1$ jest wektor $[1,1]^T$. Dla macierzy $A^* = A^T$ i wartości własnej $\overline{\lambda} = 1$ wektorem własnym jest $[0,1]^T$.

**Wniosek 16.4** (Geometryczna charakteryzacja macierzy hermitowskich)**.** *Macierz $A \in M_{n\times n}(\mathbb{C})$ jest hermitowska wtedy i tylko wtedy, gdy*

$$(Av|w)_* = (v|Aw)_*. \tag{16.1}$$

*dla dowolnych $v, w \in \mathbb{C}^n$*

*Dowód.* Jeśli $A = A^*$, to z lematu 16.3 spełniony jest warunek (16.1). Załóżmy teraz, że zachodzi warunek (16.1). Wtedy $A = A^*$ z jedyności w lemacie 16.3. □

**Twierdzenie 16.1** (Wartości własne macierzy hermitowskich)**.** *Wartości własne $A \in M_{n\times n}(\mathbb{C})$ macierzy hermitowskiej są liczbami rzeczywistymi. Ponadto wektory własne odpowiadające różnym wartościom własnym są ortogonalne.*

*Dowód.* Niech $\lambda \in \mathbb{C}$ będzie wartością własną $A$, a $x$ wektorem własnym, odpowiadającym tej wartości własnej.

Wtedy

$$\begin{aligned}\overline{\lambda}\|x\|_*^2 &= (x|\lambda x)_* \\ &= (x|Ax)_* \\ &= (Ax|x)_* \\ &= (\lambda x|x)_* \\ &= \lambda(x, x)_* \\ &= \lambda\|x\|_*^2,\end{aligned}$$

czyli $\lambda = \overline{\lambda}$.

Niech $\lambda_1$, $\lambda_2 \in \mathbb{R}$ będą różnymi wartościami własnymi, a $v_1$, $v_2$ odpowiadającymi im wektorami własnymi. Wtedy

$$\begin{aligned}(v_2|Av_1)_* &= (v_2|\lambda_1 v_1)_* \\ &= \lambda_1(v_2|v_1)_*.\end{aligned}$$

Z drugiej strony

$$\begin{aligned}(v_2|Av_1)_* &= (Av_2|v_1)_* \\ &= (\lambda_2 v_2|v_1)_* \\ &= \lambda_2(v_2|v_1)_*.\end{aligned}$$

Ponieważ $\lambda_1 \neq \lambda_2$, więc $(v_2|v_1)_* = 0$. □

**Definicja 16.3.** **Macierz unitarna**

Macierz $U = \left[\begin{array}{c|c|c} u_1 & \dots & u_n \end{array}\right] \in M_{n\times n}(\mathbb{C})$ jest unitarna, jeśli jej kolumny tworzą układ ortonormalny tzn.

$$u_i^* u_j = (u_i|u_j)_* = \delta_{ij}.$$

**Wniosek 16.5.** *Macierz $U \in M_{n\times n}(\mathbb{C})$ jest unitarna wtedy i tylko wtedy, gdy $U^*U = I$. W szczególności, $U^{-1} = U^*$. Rzeczywiste macierze unitarne, to macierze ortogonalne.*

**Wniosek 16.6** (Wartości własne macierzy unitarnej)**.** *Niech $U \in M_{n\times n}(\mathbb{C})$ będzie unitarna. Jeśli $\lambda \in \mathbb{C}$ jest wartością własną $U$, to $|\lambda| = 1$.*

*Dowód.* Niech $0 \neq v \in \mathbb{C}^n$ będzie wektorem własnym $U$ dla $\lambda$. Oczywiście, $\lambda \neq 0$, bo $U$ jest odwracalna. Wtedy

$$\begin{aligned}\lambda(v|v)_* &= (\lambda v|v)_* \\ &= (Uv|v)_* \\ &= (v|U^*v)_* \\ &= (v|U^{-1}v)_* \\ &= (v|(1/\overline{\lambda})v)_* \\ &= \frac{1}{\overline{\lambda}}(v|v)_*.\end{aligned}$$

Stąd, $\lambda = \frac{1}{\bar{\lambda}}$, czyli $|\lambda| = 1$. □

| Różnice w terminologii | |
|---|---|
| $\mathbb{R}^n$ | $\mathbb{C}^n$ |
| euklidesowy iloczyn skalarny $(\cdot\|\cdot)$ | hermitowski iloczyn skalarny $(\cdot\|\cdot)_*$ |
| macierz symetryczna $A = A^T$ | macierz hermitowska $A = A^*$ |
| macierz ortogonalna $Q^TQ = I$ | macierz unitarna $U^*U = I$ |
| grupa ortogonalna $O(n)$ | grupa unitarna $U(n)$ |

U Dla podprzestrzeni wektorowej $U \subset \mathbb{C}^n$ definiujemy podprzestrzeń ortogonalną

$$U^\perp = \{v \in \mathbb{C}^n : (v|u)_* = 0 \quad \text{dla każdego} \quad u \in U\}.$$

Podobnie jak w przypadku rzeczywistym pokazujemy, że mamy rozkład na ortogonalną sumę prostą

$$\mathbb{C}^n = U \oplus U^\perp.$$

**Twierdzenie 16.2.** *Niech $A \in M_{m\times n}(\mathbb{C})$. Wtedy*

$$\ker A^* = (\operatorname{im} A)^\perp, \quad \operatorname{im} A^* = (\ker A)^\perp.$$

*Dowód.* Przypomnijmy, że jeśli $v \in \mathbb{C}^n$, to $(u|v)_* = 0$ dla każdego $u \in \mathbb{C}^n$ wtedy i tylko wtedy, gdy $v = 0$. Stąd

$$\begin{aligned}
\ker A^* &= \{v \in \mathbb{C}^m : A^*v = 0\} \\
&= \{v \in \mathbb{C}^m : (u|A^*v)_* = 0 \quad \text{dla każdego} \quad u \in \mathbb{C}^n\} \\
&= \{v \in \mathbb{C}^m : (Au|v)_* = 0 \quad \text{dla każdego} \quad u \in \mathbb{C}^n\} \\
&= \{v \in \mathbb{C}^m : (w|v)_* = 0 \quad \text{dla każdego} \quad w \in \operatorname{im} A\} \\
&= (\operatorname{im} A)^\perp.
\end{aligned}$$

Uzasadnimy teraz drugą równość z tezy. Jeśli $v = A^*w \in \operatorname{im} A^*$ oraz $u \in \ker A$, to

$$\begin{aligned}
(u|v)_* &= (u|A^*w)_* \\
&= (Au|w)_* \\
&= (0, w)_* \\
&= 0,
\end{aligned}$$

czyli $\operatorname{im} A^*$ jest podprzestrzenią $(\ker A)^\perp$. Wystarczy pokazać, że mają one ten sam wymiar. Mamy

$$\begin{aligned}
\dim \operatorname{im} A^* &= m - \dim \ker A^* \\
&= m - \dim(\operatorname{im} A)^\perp \\
&= \dim\ \operatorname{im} A \\
&= n - \dim \ker A \\
&= \dim(\ker A)^\perp.
\end{aligned}$$

□

**Lemat 16.5.** *Załóżmy, że $A \in M_{n\times n}(\mathbb{C})$ jest hermitowska. Jeśli $U \subset \mathbb{C}^n$ jest podprzestrzenią niezmienniczą dla $A$ (tzn. $A(U) \subset U$), to $U^\perp$ jest również podprzestrzenią niezmienniczą dla $A$.*

*Dowód.* Niech $v \in U^\perp$. Pokażemy, że $Av \in U^\perp$. Dla $u \in U$ mamy

$$(Av|u)_* = (\underbrace{v}_{\in U^\perp} | \underbrace{Au}_{\in U})_* = 0.$$

□

**Przykład 16.3.** Jeśli $A \in M_{n\times m}(\mathbb{C})$ jest dowolną macierzą, to macierze $A^*A$ i $AA^*$ są hermitowskie. Podobnie jak w przypadku rzeczywistym (lemat 13.2) pokazujemy, że

$$\ker A^*A = \ker A, \quad \operatorname{im} A^*A = \operatorname{im} A^*.$$

Pokażemy, że każdą macierz hermitowską da się zdiagonalizować przy pomocy macierzy unitarnej. Udowodnimy najpierw:

**Twierdzenie 16.3** (Schur)**.** *Niech $A \in M_{n\times n}(\mathbb{C})$ będzie* **dowolną** *macierzą. Istnieje taka macierz unitarna, że $U^*AU$ jest górnie trójkątna.*

*Dowód.* Zastosujemy indukcję względem $n$. Dla $n = 1$ teza zachodzi. Załóżmy, że teza zachodzi dla macierzy rozmiaru $k\times k$. Pokażemy, że zachodzi też dla macierzy rozmiarów $(k+1)\times(k+1)$. Niech $\lambda_1$ będzie wartością własną $A$, a $w_1$ odpowiadającym jej wektorem własnym o normie jeden. Stosując procedurę Grama-Schmidta znajdujemy bazę ortonormalną $w_1, w_2, \ldots, w_{k+1}$ dla $\mathbb{C}^{k+1}$. Niech

$$W = \left[\begin{array}{c|c|c} w_1 & \ldots & w_{k+1} \end{array}\right] \in M_{(k+1)\times(k+1)}(\mathbb{C}).$$

Wtedy pierwszą kolumną macierzy $W^*AW$ jest wektor $W^*Aw_1$. Mamy

$$\begin{aligned} W^*Aw_1 &= W^*\lambda_1 \cdot w_1 \\ &= \lambda_1 \cdot W^*w_1 \\ &= \lambda_1 \cdot e_1. \end{aligned}$$

Stąd, $W^*AW$ ma postać

$$\left[\begin{array}{c|cccc} \lambda_1 & * & * & \ldots & * \\ \hline 0 & & & & \\ \vdots & & & M & \\ 0 & & & & \end{array}\right],$$

gdzie $M$ ma rozmiar $k\times k$. Z założenia indukcyjnego istnieje taka macierz unitarna $V_1 \in M_{k\times k}(\mathbb{C})$, ze $V_1^*MV_1 = T_1$ jest górnie trójkątna. Definiujemy

$$V = \left[\begin{array}{c|cccc} 1 & 0 & 0 & \ldots & 0 \\ \hline 0 & & & & \\ \vdots & & & V_1 & \\ 0 & & & & \end{array}\right].$$

Wtedy $V$ jest unitarna oraz

$$\begin{aligned} V^*W^*AWV &= \left[\begin{array}{c|cccc} \lambda_1 & * & * & \ldots & * \\ \hline 0 & & & & \\ \vdots & & & V_1^*MV_1 & \\ 0 & & & & \end{array}\right] \\ &= \left[\begin{array}{c|cccc} \lambda_1 & * & * & \ldots & * \\ \hline 0 & & & & \\ \vdots & & & T_1 & \\ 0 & & & & \end{array}\right] \end{aligned}$$

jest macierzą górnie trójkątną. Definiujemy $U = WV$. Wtedy, $U$ jest unitarna, bo

$$\begin{aligned} U^*U &= (WV)^*WV \\ &= V^*W^*WV \\ &= I. \end{aligned}$$

□

U Z dowodu twierdzenia Schura wynika, że jeśli $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ są wartościami własnymi macierzy $A \in M_{n\times n}(\mathbb{C})$, to istnieje taka macierz unitarna $U$, że

$$U^*AU = \left[\begin{array}{c|cccc} \lambda_1 & * & * & \ldots & * \\ \hline 0 & & & & \\ \vdots & & & T_1 & \\ 0 & & & & \end{array}\right],$$

macierz $T_1 \in M_{(n-1)\times(n-1)}(\mathbb{C})$ jest górnie trójkątna oraz $\lambda_2, \ldots, \lambda_n$ są wartościami własnymi $T_1$.

**Twierdzenie 16.4** (A. Brauer)**.** *Załóżmy, że $\lambda, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$ są wartościami własnymi macierzy $A \in M_{n\times n}(\mathbb{C})$. Niech $x \in \mathbb{C}^n = M_{n\times 1}(\mathbb{C})$ będzie wektorem własnym dla $\lambda$. Wtedy dla każdego wektora $v \in \mathbb{C}^n = M_{n\times 1}(\mathbb{C})$ macierz $A + xv^*$ ma wartości własne $\lambda + v^*x, \lambda_2, \ldots, \lambda_n$.*

*Dowód.* Niech $\xi = x/\|x\|_*$. Istnieje taka macierz unitarna

$$U = \left[\begin{array}{c|c|c|c} \xi & u_2 & \ldots & u_n \end{array}\right],$$

że

$$U^*AU = \left[\begin{array}{c|cccc} \lambda & * & * & \ldots & * \\ \hline 0 & & & & \\ \vdots & & & T_1 & \\ 0 & & & & \end{array}\right],$$

oraz $\lambda_2, \ldots, \lambda_n$ są wartościami własnymi $T_1 \in M_{(n-1)\times(n-1)}(\mathbb{C})$. Zwróćmy uwagę, że

$$\begin{aligned} U^*xv^*U &= \begin{bmatrix} \xi^*x \\ u_2^*x \\ \vdots \\ u_n^*x \end{bmatrix} v^*U \\ &= \begin{bmatrix} \|x\|_* \\ 0 \\ \vdots \\ 0 \end{bmatrix} \left[\begin{array}{c|c|c|c} v^*\xi & v^*u_2 & \ldots & v^*u_n \end{array}\right] \\ &= \left[\begin{array}{c|cccc} v^*x & * & * & \ldots & * \\ \hline 0 & & & & \\ \vdots & & & 0 & \\ 0 & & & & \end{array}\right]. \end{aligned}$$

Stąd

$$U^*(A + xv^*)U = \left[\begin{array}{c|cccc} \lambda + v^*x & * & * & \ldots & * \\ \hline 0 & & & & \\ \vdots & & & T_1 & \\ 0 & & & & \end{array}\right]$$

ma wartości własne $\lambda + v^*x, \lambda_2, \ldots, \lambda_n$. □

**Wniosek 16.7.** *Dla dowolnych wektorów $v, x \in \mathbb{C}^n$ mamy*

$$\det(I + xv^*) = 1 + v^*x.$$

**Twierdzenie 16.5** (Spektralne dla macierzy hermitowskich)**.** *Jeśli $A \in M_{n\times n}(\mathbb{C})$ jest hermitowska, to istnieje taka macierz unitarna $U \in M_{n\times n}(\mathbb{C})$, że $U^*AU$ jest diagonalna.*

*Dowód.* Wiemy, że istnieje taka macierz unitarna $U$, że $T = U^*AU$ jest górnie trójkątna. Wtedy

$$\begin{aligned} T^* &= (U^*AU)^* \\ &= U^*A^*U \\ &= U^*AU \\ &= T, \end{aligned}$$

czyli $T$ jest hermitowska. Stąd $T$ musi być diagonalna. □

**Wniosek 16.8.** *Jeśli $A \in M_{n\times n}(\mathbb{R})$ jest rzeczywistą macierzą symetryczną, to istnieje taka macierz ortogonalna $Q \in O(n)$, że $Q^T AQ$ jest diagonalna.*

Twierdzenie 16.5 wynika również z następującego rozumowania. Załóżmy, że $A \in M_{n\times n}(\mathbb{C})$ jest hermitowska oraz $\lambda_1, \ldots, \lambda_k$ są różnymi wartościami własnymi $A$. Definiujemy

$$V_i = \{v \in \mathbb{C}^n : Av = \lambda_i \cdot v\}, \quad i = 1, \ldots, k.$$

Podprzestrzeń

$$U = V_1 + \ldots + V_k$$

jest rozpięta na wektorach własnych $A$. Wystarczy pokazać, że $U = \mathbb{C}^n$. Równoważnie, pokażemy, że $U^\perp = \{0\}$. Przypuśćmy, że $U^\perp \neq \{0\}$. Z lematu 16.5 podprzestrzeń $U^\perp$ jest niezmiennicza dla $A$. Stąd $U^\perp$ musi zawierać wektor własny $A$, co prowadzi do sprzeczności, bo wszystkie wektory własne należą do $U$.

**Wniosek 16.9.** *Niech $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ będą wartościami własnymi macierzy $A = [a_{ij}] \in M_{n\times n}(\mathbb{C})$. Wtedy*

$$\sum_{i=1}^{n} |\lambda_i|^2 \leqslant \sum_{i,j=1}^{n} |a_{ij}|^2.$$

*Ponadto równość zachodzi wtedy i tylko wtedy, gdy $A$ jest unitarnie diagonalizowalna.*

*Dowód.* Z twierdzenia Schura istnieje taka macierz unitarna $U$, że $T = U^*AU$ jest górnie trójkątna z wartościami własnymi $A$ na przekątnej. Zauważmy, że

$$\operatorname{tr} A^*A = \sum_{i,j=1}^{n} |a_{ij}|^2.$$

Ponadto $T^*T = U^*A^*AU$, więc $\operatorname{tr} T^*T = \operatorname{tr} A^*A$. Sprawdzamy, że

$$\begin{aligned}\sum_{i,j=1}^{n} |a_{ij}|^2 &= \operatorname{tr} A^*A \\ &= \operatorname{tr} T^*T \\ &\geqslant |t_{11}|^2 + \ldots + |t_{nn}|^2 \\ &= \sum_{i=1}^{n} |\lambda_i|^2.\end{aligned}$$

□

**Lemat 16.6.** *Niech $P \in M_{n\times n}(\mathbb{C})$ będzie idempotentna (tzn. $P^2 = P$). Wtedy $P$ jest hermitowska wtedy i tylko wtedy, gdy $(\ker P)^\perp = \operatorname{im} P$.*

*Dowód.* Ponieważ $P$ jest idempotentna, więc

$$\mathbb{C}^n = \ker P \oplus \operatorname{im} P$$

oraz $P$ jest rzutowaniem w tej sumie prostej na $\operatorname{im} P$.

Jeśli $P$ jest hermitowska oraz $u \in \ker P$ i $w \in \operatorname{im} P$, to

$$\begin{aligned}(u|w)_* &= (u|Pw)_* \\ &= (Pu|w)_* \\ &= (0|w)_* \\ &= 0.\end{aligned}$$

Stąd $\operatorname{im} P \subset (\ker P)^\perp$. Ponieważ mają one ten sam wymiar, więc są równe.

Załóżmy, że $(\ker P)^\perp = \operatorname{im} P$. Niech $v_i = u_i + w_i$, gdzie $u_i \in \ker P$ i $w_i \in \operatorname{im} P$ dla $i = 1, 2$. Wtedy

$$\begin{aligned}(Pv_1|v_2)_* &= (w_1|u_2 + w_2)_* \\ &= (w_1|w_2)_* \\ &= (u_1 + w_1|w_2)_* \\ &= (v_1|Pv_2)_*.\end{aligned}$$

□

**U** Niech $A \in M_{n\times n}(\mathbb{C})$ będzie hermitowska i niech $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ będą różnymi wartościami własnymi $A$. Dla podprzestrzeni własnych

$$V_i = \{v \in \mathbb{C}^n : Av = \lambda_i v\}, \quad i = 1, \ldots, k$$

mamy rozkład na ortogonalną sumę prostą

$$\mathbb{C}^n = V_1 \oplus \ldots \oplus V_k.$$

Niech $P_i$ będzie rzutowaniem na $V_i$ w tej sumie prostej. Z lematu 16.6 wynika, że rzutowania $P_i$ są hermitowskie oraz $P_iP_j = 0$ dla $i \neq j$. Ponadto

$$I = P_1 + \ldots + P_k, \quad A = \lambda_1 \cdot P_1 + \ldots + \lambda_k \cdot P_k.$$

**Lemat 16.7.** *Niech $A \in M_{n\times m}(\mathbb{C})$. Wtedy*

*(i) wartości własne macierzy $A^*A$ są nieujemne,*

*(ii) jeśli $\lambda > 0$ jest wartością własną $A^*A$ z wektorem własnym $v$, to $\lambda > 0$ jest wartością własną $AA^*$ z wektorem własnym $Av$. Ponadto krotności algebraiczne $\lambda$ jako wartości własnych $A^*A$ i $AA^*$ są równe.*

*Dowód.* Ponieważ $A^*A$ jest hermitowska, więc ma rzeczywiste wartości własne. Niech $\lambda \in \mathbb{R}$ będzie wartością własną z wektorem własnym $v$. Mamy

$$\begin{aligned}\lambda(v|v)_* &= (\lambda \cdot v|v)_* \\ &= (A^*Av|v)_* \\ &= (Av|Av)_*,\end{aligned}$$

czyli

$$\lambda = \frac{\|A\|_*^2}{\|v\|_*^2} \geqslant 0.$$

Niech $\lambda > 0$ będzie wartością własną $A^*A$ z wektorem własnym $v$. Wtedy $A^*Av = \lambda \cdot v \neq 0$, więc $Av \neq 0$ oraz

$$AA^*Av = A(\lambda \cdot v) = \lambda \cdot Av,$$

czyli $Av$ jest wektorem własnym $AA^*$ z wartością własną $\lambda$.

Niech $n_{A^*A}$ i $n_{AA^*}$ oznaczają algebraiczne krotności $\lambda > 0$ jako wartości własnych $A^*A$ i $AA*$. Ponieważ $A^*A$ i $AA^*$ są diagonalizowalne, więc krotności są równe krotnościom geometrycznym, czyli liczbie liniowo niezależnych wektorów własnych dla $\lambda$. Ze względu na symetrię wystarczy pokazać, że $n_{A^*A} \leqslant n_{AA^*}$. Niech $v_1, \ldots, v_{n_{A^*A}}$ będzie bazą podprzestrzeni własnej dla $\lambda$. Wystarczy pokazać, że $Av_1, \ldots, Av_{n_{A^*A}}$ są liniowo niezależne. Jeśli $a_1 \cdot Av_1 + \ldots a_{n_{A^*A}} \cdot Av_{n_{A^*A}} = 0$ dla pewnych skalarów $a_i$, to również

$$0 = A^*(0) = a_1 \cdot A^*Av_1 + \ldots a_{n_{A^*A}} \cdot A^*Av_{n_{A^*A}},$$

czyli

$$(\lambda a_1) \cdot v_1 + \ldots (\lambda a_{n_{A^*A}}) \cdot v_{n_{A^*A}} = 0.$$

Z liniowej niezależności wynika, że $a_1 = \ldots = a_{n_{A^*A}} = 0$, bo $\lambda > 0$. □

**U** Lemat 16.7 można czasami wykorzystać do łatwego wyznaczenia dodatnich wartości własnych macierzy $A^*A$. Przykładowo, niech $A \in M_{100\times 2}(\mathbb{C})$. Wtedy $A^*A \in M_{100\times 100}(\mathbb{C})$ ma duży rozmiar. Macierz $AA^* \in M_{2\times 2}(\mathbb{C})$ ma takie same dodatnie wartości własne liczone z krotnościami.

## 16.1 Macierze normalne. Ortonormalna baza wektorów własnych.

Przypuśćmy, że dla $A \in M_{n\times n}(\mathbb{C})$ przestrzeń $\mathbb{C}^n$ ma bazę ortonormalną $u_1, \ldots, u_n$ złożoną z wektorów własnych $A$. Dla macierzy unitarnej $U = \left[\begin{array}{c|c|c} u_1 & \ldots & u_n \end{array}\right]$ mamy wtedy, że $U^*AU = D$ jest diagonalna oraz

$$DD^* = U^*AUU^*A^*U = U^*AA^*U,$$
$$D^*D = U^*A^*UU^*AU = U^*A^*AU.$$

Ponieważ $DD^* = D^*D$ dla macierzy diagonalnej $D$, więc

$$AA^* = A^*A.$$

Jest to więc warunek konieczny na istnienie bazy ortonormalnej $u_1, \ldots, u_n$ złożonej z wektorów własnych $A$.

**Definicja 16.4.** Macierz normalna

Macierz $A \in M_{n\times n}(\mathbb{C})$ nazywamy *normalną*, gdy

$$AA^* = A^*A.$$

**Przykład 16.4.** Macierz rzeczywista jest normalna, gdy $A^TA = AA^T$.

**Przykład 16.5.** Macierze symetryczne i hermitowskie są normalne. Podobnie maciecierze ortogonalne i unitarne.

**Przykład 16.6.** Macierz

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \in O(2)$$

jest normalna. Jeśli $\sin\theta \neq 0$, to $A$ nie jest ortogonalnie diagonalizowalna (nad $\mathbb{R}$). Jest ona jednak unitanie diagonalizowalna (nad $\mathbb{C}$).

**Lemat 16.8.** *Niech $A \in M_{n\times n}(\mathbb{C})$ będzie macierzą normalną. Wtedy*

*(i)* $\|Av\|_* = \|A^*v\|_*$ *dla każdego* $v \in \mathbb{C}^n$,

*(ii)* $\ker A = \ker A^*$,

*(iii) macierz $A - \lambda I$ jest normalna dla każdego $\lambda \in \mathbb{C}$,*

*(iv) wektory własne odpowiadające różnym wartościom własnym $A$ są ortogonalne,*

*(v) jeśli $v$ jest wektorem własnym macierzy $A$ dla wartości własnej $\lambda$, to $v$ jest wektorem własnym macierzy $A^*$ dla wartości własnej $\overline{\lambda}$,*

*(vi) $B = U^*AU$ jest normalna dla dowolnej macierzy unitarnej $U$.*

*Dowód.* Mamy

$$\begin{aligned} \|Av\|_*^2 &= (Av|Av)_* \\ &= (v|A^*Av)_* \\ &= (v|AA^*v)_* \\ &= (A^*v|A^*v)_* \\ &= \|A^*v\|_*^2. \end{aligned}$$

Punkt (ii) jest konsekwencją (i). Punkt (iii) wynika z bezpośredniego rachunku.

Udowodnimy punkt (iv). Niech $U = A - \lambda I$. Wtedy $U^* = A^* - \overline{\lambda}I$. Ponieważ $U$ jest normalna, więc z punktu (ii) $U(v) = 0$ wtedy i tylko wtedy, gdy $U^*v = 0$.

Niech $\lambda_1, \lambda_2$ będą różnymi wartościami własnymi z wektorami własnymi $v_1, v_2$. Wtedy

$$\begin{aligned}\lambda_1(v_1|v_2)_* &= (\lambda_1 \cdot v_1|v_2)_* \\ &= (Av_1|v_2)_* \\ &= (v_1|A^*v_2)_* \\ &= (v_1|\overline{\lambda_2} \cdot v_2)_* \\ &= \lambda_2(v_1|v_2)_*.\end{aligned}$$

Ponieważ $\lambda_1 \neq \lambda_2$, więc $(v_1|v_2)_* = 0$.

Dla dowodu (vi) zauważmy, że

$$\begin{aligned}B^*B &= (U^*AU)^*(U^*AU) \\ &= U^*A^*UU^*AU \\ &= U^*A^*AU \\ &= U^*AA^*U \\ &= U^*AUU^*A^*U \\ &= BB^*.\end{aligned}$$

□

**Lemat 16.9.** *Jeśli macierz górnie trójkątna $T \in M_{n\times n}(\mathbb{C})$ jest normalna, to jest diagonalna.*

*Dowód.* Przyglądnijmy się, jak to wygląda dla $n = 2$. Niech

$$T = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}.$$

Wtedy

$$T^*T = \begin{bmatrix} |t_{11}|^2 & * \\ * & |t_{12}|^2 + |t_{22}|^2 \end{bmatrix}, \quad TT^* = \begin{bmatrix} |t_{11}|^2 + |t_{12}|^2 & * \\ * & |t_{22}|^2 \end{bmatrix}.$$

Ponieważ $T^*T = TT^*$, więc (porównując wyrazy diagonalne tych macierzy) otrzymujemy, że $t_{12} = 0$, więc $T$ jest diagonalna.

Ogólnie, jeśli $T = [t_{ij}]$ jest górnie trójkątna, to (porównując elementy diagonalne macierzy $TT^*$ oraz $T^*T$) otrzymujemy, że

$$\begin{aligned}|t_{11}|^2 + |t_{12}|^2 + |t_{13}|^2 + \ldots + |t_{1n}|^2 &= |t_{11}|^2, \\ |t_{22}|^2 + |t_{23}|^2 + \ldots + |t_{2n}|^2 &= |t_{12}|^2 + |t_{22}|^2 \\ &\vdots \\ |t_{nn}|^2 &= |t_{1n}|^2 + |t_{2n}|^2 + \ldots + |t_{nn}|^2.\end{aligned}$$

Wynika stąd, że $t_{ij} = 0$ dla $i \neq j$, więc $T$ jest diagonalna. □

**Twierdzenie 16.6** (Unitarna diagonalizacja macierzy normalnej)**.** *Macierz $A \in M_{n\times n}(\mathbb{C})$ jest normalna wtedy i tylko wtedy, gdy $\mathbb{C}^n$ ma bazę ortonormalną $u_1, \ldots, u_n$ złożoną z wektorów własnych $A$ co oznacza, że istnieje taka macierz unitarna $U$, że $U^*AU$ jest diagonalna.*

*Dowód.* Załóżmy, że $A$ jest normalna. Wiemy, że istnieje taka macierz unitarna $U$, że

$$T = U^*AU$$

jest macierzą górnie trójkątną. Wystarczy sprawdzić, że $T$ jest diagonalna. Wtedy kolumny $U$ są wektorami własnymi $A$. Pokażemy najpierw, że $T$ jest normalna. Mamy

$$T^*T = U^*A^*AU, \quad TT^* = U^*AA^*U,$$

czyli $T^*T = TT^*$. Z lematu 16.9 wynika, że $T$ jest diagonalna. □

**Wniosek 16.10.** *Jeśli macierz normalna $A \in M_{n\times n}(\mathbb{C})$ ma rzeczywiste wartości własne, to $A$ jest hermitowska.*

**Wniosek 16.11.** *Niech $A \in M_{n\times n}(\mathbb{C})$. Następujące warunki są równoważne*

*(i) $A$ jest unitarna,*

*(ii) istnieje taka macierz unitarna $U \in M_{n\times n}(\mathbb{C})$, że*

$$U^*AU = \operatorname{diag}(e^{i\theta_1}, \dots, e^{i\theta_n}).$$

*Dowód.* Warunek (ii) implikuje (i), bo $\operatorname{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})$ jest unitarna.

Jeśli zachodzi (i), to $A$ jest normalna, więc jest też unitarnie diagonalizowalna. Teza zachodzi, bo wartości własne $A$ jako macierzy unitarnej leżą na okręgu jednostkowym $\mathbb{S}^1$. □

Przypuśćmy, że macierz $A \in M_{n\times n}(\mathbb{C})$ jest unitarna. Wtedy

$$A = U \operatorname{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})U^*$$

dla pewnej macierzy unitarnej. Rozważmy odwzorowanie

$$H : [0,1] \times \mathbb{C}^n \ni (t,z) \longmapsto U \operatorname{diag}(e^{it\theta_1}, \dots, e^{it\theta_n})U^* z \in \mathbb{C}^n.$$

Ma ono następujące własności:

- $H(0,z) = z, \quad H(1,z) = Az,$
- $H(t,\cdot) = U \operatorname{diag}(e^{it\theta_1}, \dots, e^{it\theta_n})U^* \in M_{n\times n}(\mathbb{C})$ jest unitarna dla każdego $t \in [0,1]$,
- $H$ jest odwzorowaniem ciągłym.

Topolodzy mówią wtedy, że odwzorowanie $A$ jest homotopijne z odwzorowaniem identycznościowym $I$ i odwzorowania pośrednie $H(t,\cdot)$ są również unitarne.

**Lemat 16.10.** *Jeśli wektory zespolone $a \pm bi = (a_1 \pm ib_1, \dots, a_n \pm ib_n) \in \mathbb{C}^n$ z $b \neq 0$ są ortonormalne, to $(a|b) = 0$ oraz $\|a\| = \|b\| = 1/2$.*

*Dowód.* Mamy

$$\begin{aligned} 0 &= (a+bi|a-bi)_* \\ &= (a|a)_* + i^2(b|b)_* + i(b|a)_* + i(a|b)_* \\ &= (a|a) - (b|b) + 2i(a|b), \end{aligned}$$

więc

$$0 = \|a\|^2 - \|b\|^2 + i2(a|b).$$

Stąd $\|a\| = \|b\|$ oraz $(a|b) = 0$. Z równości $1 = \|a+ib\|_*^2 = \|a\|^2 + \|b\|^2$ otrzymujemy, że

$$(a|b) = 0, \quad \|a\|^2 = \|b\|^2 = \frac{1}{2}.$$

□

**Lemat 16.11.** *Jeśli*

$$0 = (a+ib|u+iw)_* = (a+ib|u-iw)_*,$$

*to*

$$(a|u) = (a|w) = (b|u) = (b|w) = 0.$$

*Dowód.* Wynika z równości

$$(a+ib|u+iw)_* = (a|u) + (b|w) + i((b|u) - (a|w)),$$

$$(a+ib|u-iw)_* = (a|u) - (b|w) + i((b|u) + (a|w)).$$

□

U

Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą rzeczywistą. Załóżmy, że

$$u_1 = v_1 + iw_1, \dots, u_n = v_n + iw_n$$

jest bazą ortonormalną $\mathbb{C}^n$ złożoną z wektorów własnych macierzy $A$. Zmodyfikujemy bazę ortonormalną $u_1, \dots, u_n$ dla $\mathbb{C}^n$ do pewnej bazy ortonormalnej $x_1, \dots, x_n$ dla $\mathbb{R}^n$. Jeśli wartość własna $\lambda_i$ dla $u_i$ jest rzeczywista, to $u_i \in \mathbb{R}^n$ i kładziemy $x_i = u_i$.

Załóżmy, że $\lambda_j = \alpha_j + i\beta_j$ jest zespolona z $\beta_j \neq 0$. Ponieważ $A$ jest rzeczywista, więc $\overline{\lambda_j} = \alpha_j - i\beta_j$ jest również wartością własną. Możemy założyć, że odpowiada ona wektorowi własnemu $u_{i+1} = \overline{u_i}$. Wtedy w bazie $u_1, \dots, u_n$ wektory $u_j$ oraz $u_{j+1}$ zastępujemy przez wektory $2v_j$ oraz $2w_j$ tzn. podwojoną część rzeczywistą i urojoną wektora $u_j$. Otrzymujemy wektory $x_1, \dots, x_n$. Z lematów 16.10 i 16.11 wynika, że $x_1, \dots, x_n$ jest bazą ortonormalną $\mathbb{R}^n$. Jeśli $u_j = v_j + iw_j$, to

$$Ax_j = \alpha_j x_j - \beta_j x_{j+1}, \quad Ax_{j+1} = \alpha_j x_{j+1} + \beta_j x_j.$$

Wynika stąd następujące:

**Twierdzenie 16.7** (Rzeczywista macierz normalna)**.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą rzeczywistą. Następujące warunki są równoważne*

*(1) $A$ jest normalna tzn. $AA^T = A^T A$,*

*(2) istnieje taka macierz ortogonalna $Q \in M_{n\times n}(\mathbb{R})$, że*

$$Q^T A Q = \operatorname{diag}(A_1, A_2, \dots, A_k),$$

*gdzie $A_i \in M_{1\times 1}(\mathbb{R})$ lub $A_j \in M_{2\times 2}(\mathbb{R})$ i ma postać*

$$A_j = \begin{bmatrix} \alpha_j & \beta_j \\ -\beta_j & \alpha_j \end{bmatrix}, \quad \beta_j \neq 0.$$

*Ponadto jeśli $A$ jest symetryczna, to $A_i$ są jednowymiarowe, a gdy $A$ jest ortogonalna, to jednowymiarowe klatki $A_i$ są równe $\pm 1$.*

| Własności spektralne | |
|---|---|
| macierz | wartości własne |
| hermitowska/symetryczna | rzeczywiste |
| unitarna/ortogonalna | $\lvert\lambda\rvert = 1$ |
| Grama $A^*A/A^TA$ | nieujemne |
| nieosobliwa | niezerowe |

# Rozdział 17

# Przestrzeń dualna

| |
|---|
| Słowa kluczowe rozdziału |
| przestrzeń dualna ♢ funkcjonał liniowy ♢ baza dualna ♢ twierdzenie Riesza |

Niech $V$ będzie przestrzenią wektorową nad ciałem $\mathbb{F}$. Szczególną rolę pełnią odwzorowania liniowe $L : V \longrightarrow \mathbb{F}$. Warto je wyróżnić.

**Definicja 17.1.** Przestrzeń dualna

Niech $V \in \mathrm{Vekt}_{\mathbb{F}}$. Przestrzeń $\mathcal{L}(V, \mathbb{F})$ nazywamy przestrzenią dualną do $V$ i oznaczamy przez

$$V' = \mathcal{L}(V, \mathbb{F}) = \{f | \ f : V \longrightarrow \mathbb{F} \quad \text{odwzorowanie liniowe}\}.$$

U Przestrzeń dualna $V'$ jest często oznaczana przez $V^*$.

**Przykład 17.1.** Niech $V = C([a, b], \mathbb{R})$ będzie przestrzenią wektorową funkcji ciągłych na przedziale $[a, b]$ z $a < b$. Wtedy całka Riemanna

$$C([a, b], \mathbb{R}) \ni f \longmapsto \int_a^b f(x)dx \in \mathbb{R}$$

jest elementem $C([a, b], \mathbb{R})'$.

**Przykład 17.2.** Niech $v \in M_{1\times n}(\mathbb{F})$. Odwzorowanie

$$M_{n\times 1}(\mathbb{F}) \ni w \longmapsto vw \in M_{1\times 1}(\mathbb{F}) = \mathbb{F}$$

jest elementem przestrzeni dualnej do $M_{n\times 1}(\mathbb{F})$. Zwróćmy uwagę, że dla $\mathbb{F} = \mathbb{R}$ mamy

$$vw = (v^T|w).$$

Opiszemy teraz przestrzeń dualną $V'$ do przestrzeni skończenie wymiarowej $V$. Niech $v_1, \dots, v_n$ będzie bazą $V$. Rozważmy $f \in V'$, czyli odwzorowanie liniowe $f : V \longrightarrow \mathbb{F}$. Takie odwzorowania (tzn. gdy prowadzą do $\mathbb{F}$) nazywa się *funkcjonałami liniowymi* albo *kowektorami* albo 1-*formami.* Zdefiniujemy pewną bazę dla $V'$ skojarzoną z bazą $v_1, \dots, v_n$ w $V$. Dla wektora bazowego $v_i$ zdefiniujemy pewien kowektor

$$v_i' : V \longrightarrow \mathbb{F}.$$

Jest on jednoznacznie wyznaczony przez wartości na bazie $v_1, \dots, v_n$. Definiujemy

$$v_i'(v_j) := \delta_{ij}.$$

Wtedy dla wektora $v = x_1 \cdot v_1 + \dots + x_n \cdot v_n \in V$ $(x_i \in \mathbb{F})$ mamy

$$v_i'(v) = x_1 v_i'(v_1) + \dots + x_n v_i'(v_n) = x_i.$$

W szczególności,

$$v = v_1'(v) \cdot v_1 + \dots + v_n'(v) \cdot v_n.$$

**Przykład 17.3.** Jeśli $e_1, \ldots, e_n$ jest bazą standardową $\mathbb{F}^n$, to

$$e_i'([x_1, \ldots, x_n]^T) = x_i,$$

czyli $e_i' : \mathbb{R}^n \longrightarrow \mathbb{R}$ jest rzutowaniem na $i$-tą współrzędną.

**Twierdzenie 17.1** (Baza dualna)**.** *Niech $V$ będzie przestrzenią wektorową skończonego wymiaru. Dla dowolnej bazy $v_1, \ldots, v_n$ dla $V$, kowektory $v_1', \ldots, v_n'$ tworzą bazę dla $V'$. Nazywamy ją bazą dualną do bazy $v_1, \ldots, v_n$ dla $V$. W szczególności,*

$$\dim V = \dim V'.$$

*Dowód.* Pokażmy najpierw, że $v_1', \ldots, v_n'$ są liniowo niezależne. Przypuśćmy, że

$$x_1 \cdot v_1' + \ldots + x_n \cdot v_n' = 0$$

dla pewnych $x_i \in \mathbb{F}$. Wtedy

$$0 = (x_1 \cdot v_1' + \ldots + x_n \cdot v_n')(v_i) = x_i$$

dla $i = 1, \ldots, n$, więc $v_1', \ldots, v_n'$ są liniowo niezależne.

Uzasadnimy teraz, że

$$\operatorname{span}\{v_1', \ldots, v_n'\} = V'.$$

Niech $f \in V'$. Dla $x_i = f(v_i)$ otrzymujemy, że

$$f = x_1 \cdot v_1' + \ldots + x_n \cdot v_n',$$

bo dla $i = 1, \ldots, n$ mamy

$$f(v_i) = x_i = (x_1 \cdot v_1' + \ldots + x_n \cdot v_n')(v_i).$$

□

**Wniosek 17.1.** *Niech $v_1', \ldots, v_n' \in V'$ będzie bazą dualną do bazy $v_1, \ldots, v_n$ dla $V$. Wtedy dla $v \in V$ i $f \in V'$ zachodzą równości*

$$v = v_1'(v) \cdot v_1 + \ldots v_n'(v) \cdot v_n, \quad f = f(v_1) \cdot v_1' + \ldots + f(v_n) \cdot v_n'.$$

**Przykład 17.4.** Rozważmy bazę

$$v_1 = [2, 1]^T, \quad v_2 = [3, 1]^T$$

dla $\mathbb{R}^2$. Znajdziemy jawne formuły na bazę dualną $v_1'$, $v_2'$ dla $(\mathbb{R}^2)'$. Ponieważ $v_1' : \mathbb{R}^2 \longrightarrow \mathbb{R}$ jest odwzorowaniem liniowym, więc ma ono postać

$$v_1'([x, y]^T) = ax + by$$

dla pewnych $a, b \in \mathbb{R}$. Z definicji $v_1'$ wynika, że

$$1 = v_1'(v_1) = 2a + b, \quad 0 = v_1'(v_2) = 3a + b.$$

Stąd $a = -1$, $b = 3$, więc

$$v_1'([x, y]^T) = -x + 3y.$$

Analogicznie pokazujemy, że

$$v_2'([x, y]^T) = x - 2y.$$

U

Niech $W \subset V$ będzie $n - 1$ wymiarową podprzestrzenią $n$-wymiarowej przestrzeni $V$. Niech $w_1, \ldots, w_{n-1}, v_n$ będzie taką bazą $V$, że $w_1, \ldots, w_{n-1}$ jest bazą $W$. Definiujemy $f \in V'$ przyjmując, że

$$f(w_i) = 0, \quad i = 1, \ldots, n-1, \quad f(v_n) = 1.$$

Wtedy $\ker f = W$.

**Lemat 17.1.** *Niech $V$ będzie skończenie wymiarowa i niech funkcjonał $f \in V'$ będzie niezerowy. Wtedy istnieje taka baza $u_1, \ldots, u_n$ dla $V$, że*

$$f(u_1) = 1, \quad f(u_2) = \ldots = f(u_n) = 0.$$

*Dowód.* Ponieważ $f \neq 0$, więc istnieje taki wektor $u \in V$, że $f(u) \neq 0$. Definiujemy

$$u_1 = \frac{u}{f(u)}.$$

Wtedy $f(u_1) = 1$. Uzupełniamy wektor $u_1$ do bazy $u_1, w_2, \ldots, w_n$ dla $V$. Wystarczy teraz przyjąć, że

$$u_i = w_i - f(w_i) \cdot u_1, \quad 2 \leqslant i \leqslant n.$$

Sprawdzamy łatwo, że wektory $u_1, u_2, \ldots, u_n \in V$ są liniowo niezależne oraz $f(u_i) = 0$ dla $i = 2, \ldots, n$. □

Niech $V$ będzie przestrzenią wektorową. Wtedy możemy rozważyć przestrzeń dualną $(V')'$ do przestrzeni dualnej $V'$. Nazywamy ją przestrzenią *bidualną* i oznaczamy $V''$.

**Lemat 17.2.** *Niech $V$ będzie przestrzenią wektorową. Odwzorowanie $j : V \longrightarrow V''$ zadane dla $v \in V$ oraz $f \in V'$ formułą*

$$(j(v))(f) = f(v)$$

*jest monomorfizmem. W szczególności, jeśli $\dim V < \infty$, to $j : V \longrightarrow V''$ jest izomorfizmem.*

*Dowód.* Zauważmy, że $j(v) \in V''$, więc $j$ jest dobrze określone. Ponadto $j$ jest liniowe, bo dla $a, b \in \mathbb{R}$, $v, w \in V$ oraz $f \in V'$ mamy

$$\begin{aligned}(j(a \cdot v + b \cdot w))(f) &= f(a \cdot v + b \cdot w) \\ &= a \cdot f(v) + b \cdot f(w) \\ &= a \cdot (j(v))(f) + b \cdot (j(w))(f) \\ &= (a \cdot j(v) + b \cdot j(w))(f),\end{aligned}$$

czyli

$$j(a \cdot v + b \cdot w) = a \cdot j(v) + b \cdot j(w).$$

Przypuśćmy, że $j(v_1) = 0$ i $v_1 \neq 0$. Oznacza to, że dla dowolnego funkcjonału $f \in V'$

$$0 = (j(v_1))(f) = f(v_1).$$

Rozszerzymy wektor $v_1$ do bazy $v_1, v_2, \ldots, v_n$ i rozważmy bazę dualną $v_1', v_2', \ldots, v_n'$ dla $V'$. Dla $f = v_1'$ mamy

$$0 = (j(v_1))(v_1') = v_1'(v_1) = 1,$$

co prowadzi do sprzeczności. □

U Jeśli $V$ jest skończonego wymiaru z bazą $v_1, \ldots, v_n$, to izomorfizm $j : V \longrightarrow V''$ pozwala utożsamić bazę dualną do bazy dualnej $v_1', \ldots, v_n'$ z bazą $v_1, \ldots, v_n$ dla $V$.

**Wniosek 17.2.** *Przestrzeń wektorowa $V$ jest skończonego wymiaru wtedy i tylko wtedy, gdy $V'$ jest skończonego wymiaru.*

*Dowód.* Jeśli $V'$ jest skończonego wymiaru, to $V''$ jest skończonego wymiaru. Ponieważ $j : V \longrightarrow V''$ jest monomorfizmem, więc $V$ jest skończonego wymiaru. □

**Definicja 17.2.** **Odwzorowanie dualne**

Niech $L : V \longrightarrow W$ będzie odwzorowaniem liniowym. *Odwzorowanie dualne* $L' : W' \longrightarrow V'$ jest zdefiniowane przez

$$L'(f) = f \circ L, \quad f \in W'.$$

**Wniosek 17.3.** *Niech $L : V \longrightarrow W$ będzie odwzorowaniem liniowym. Wtedy odwzorowanie dualne $L' : W' \longrightarrow V'$ jest liniowe.*

*Dowód.* Dla $f, g \in W'$ i $t \in \mathbb{F}$ mamy

$$\begin{aligned} L'(f+g) &= (f+g) \circ L \\ &= f \circ l + g \circ L \\ &= L'(f) + L'(g) \end{aligned}$$

oraz

$$L'(t \cdot f) = (t \cdot f) \circ L = t \cdot (f \circ L) = t \cdot L'(f).$$

□

**Przykład 17.5.** Załóżmy, że odwzorowanie liniowe $L : V \longrightarrow W$ ma w bazach $v_1, v_2$ dla $V$ oraz $w_1, w_2, w_3$ dla $W$ macierz

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix},$$

czyli

$$L(v_1) = a_{11} \cdot w_1 + a_{21} \cdot w_2 + a_{31} \cdot w_3, \quad L(v_2) = a_{12} \cdot w_1 + a_{22} \cdot w_2 + a_{32} \cdot w_3$$

Zobaczymy, jak wygląda macierz odwzorowania dualnego $L' : W' \longrightarrow V'$ w bazach dualnych $w_1', w_2', w_3'$ oraz $v_1', v_2'$. Pokażemy, że

$$L'(w_1') = a_{11} \cdot v_1' + a_{12} \cdot v_2' \in V^*.$$

Wystarczy sprawdzić, że obydwie strony zgadzają się na bazie $v_1, v_2$ dla $V$. Mamy

$$\begin{aligned} L'(w_1')(v_1) &= (w_1' \circ L)(v_1) \\ &= w_1'(L(v_1)) \\ &= w_1'(a_{11} \cdot w_1 + a_{21} \cdot w_2 + a_{31} \cdot w_3) \\ &= a_{11} \\ &= (a_{11} \cdot v_1' + a_{12} \cdot v_2')(v_1). \end{aligned}$$

Podobnie,

$$\begin{aligned} L'(w_1')(v_2) &= (w_1' \circ L)(v_2) \\ &= w_1'(L(v_2)) \\ &= w_1'(a_{12} \cdot w_1 + a_{22} \cdot w_2 + a_{32} \cdot w_3) \\ &= a_{12} \\ &= (a_{11} \cdot v_1' + a_{12} \cdot v_2')(v_2). \end{aligned}$$

Analogicznie pokazujemy, że

$$L'(w_2') = a_{21} \cdot v_1' + a_{22} \cdot v_2' \in V',$$

oraz

$$L'(w_3') = a_{31} \cdot v_1' + a_{32} \cdot v_2' \in V'.$$

Wynika stąd, że odwzorowanie dualne $L^*$ ma w bazach dualnych macierz

$$A^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \end{bmatrix}$$

**Wniosek 17.4.** *Niech $V$, $W$ będą przestrzeniami wektorowymi nad ciałem $\mathbb{F}$. Wtedy odwzorowanie*

$$\mathcal{L}(V, W) \ni L \longmapsto L' \in \mathcal{L}(W', V')$$

*jest liniowe.*

*Dowód.* Niech $L_1, L_2 \in \mathcal{L}(V, W)$ i $a_1, a_2 \in \mathbb{F}$. Dla $f \in W'$ mamy

$$\begin{aligned}(a_1 \cdot L_1 + a_2 \cdot L_2)'(f) &= f \circ (a_1 \cdot L_1 + a_2 \cdot L_2) \\ &= a_1 \cdot f \circ L_1 + a_2 \cdot f \circ L_2 \quad \text{liniowość} \quad f \\ &= a_1 \cdot L_1'(f) + a_2 \cdot L_2'(f) \\ &= (a_1 \cdot L_1' + a_2 \cdot L_2')(f).\end{aligned}$$

□

**Wniosek 17.5.** *Jeśli $L \in \mathcal{L}(V, W)$, $G \in \mathcal{L}(W, U)$, to*

$$(G \circ L)' = L' \circ G'.$$

*Ponadto*

$$(\mathrm{id}_V)' = \mathrm{id}_{V'}.$$

*Dowód.* Ponieważ $G \circ L \in \mathcal{L}(V, U)$, więc $(G \circ L)' \in \mathcal{L}(U', V')$. Niech $f \in U'$. Wtedy

$$\begin{aligned}(G \circ L)'(f) &= f \circ G \circ L \\ &= L'(f \circ G) \\ &= L'(G'(f)) \\ &= (L' \circ G')(f).\end{aligned}$$

□

**Lemat 17.3.** *Niech $L : V \longrightarrow W$ będzie odwzorowaniem liniowym. Jeśli $L$ ma w bazach $v_1, \ldots v_n$ dla $V$ i $w_1, \ldots, w_m$ dla $W$ macierz $A = [a_{ij}] \in M_{m \times n}(\mathbb{F})$, to odwzorowanie dualne $L' : W' \longrightarrow V'$ ma w bazach dualnych macierz $A^T$.*

*Dowód.* Niech $B = [b_{ij}]$ będzie macierzą $L'$ w bazach dualnych. Sprawdzimy, że $B = A^T$. Z definicji $L'$ i wniosku 17.1 mamy

$$\begin{aligned}L'(w_j') &= w_j' \circ L \\ &= (w_j' \circ L)(v_1) \cdot v_1' + \ldots + (w_j' \circ L)(v_n) \cdot v_n',\end{aligned}$$

czyli

$$b_{ij} = (w_j' \circ L)(v_i).$$

Z drugiej strony,

$$\begin{aligned}(w_j' \circ L)(v_i) &= w_j'(L(v_i)) \\ &= w_j'(a_{1i} \cdot w_1 + \ldots + a_{mi} \cdot w_m) \\ &= a_{ji},\end{aligned}$$

czyli $b_{ij} = a_{ji}$. □

**Wniosek 17.6.** *Niech $L : V \longrightarrow W$ będzie odwzorowaniem liniowym między przestrzeniami skończonego wymiaru. Wtedy*

$$\operatorname{rank} L = \operatorname{rank} L'.$$

**Definicja 17.3** (Anihilator podprzestrzeni)**.** Dla podprzestrzeni wektorowej $M \subset V$ definiujemy jej *anihilator* $M^0$ jako

$$\begin{aligned}M^0 &= \{f \in V' : f(v) = 0 \quad \text{dla każdego} \quad v \in M\} \\ &= \{f \in V' : M \subset \ker f\}.\end{aligned}$$

**Wniosek 17.7.** *Niech $M \subset V$ będzie podprzestrzenią wektorową. Wtedy $M^0$ jest podprzestrzenią wektorową przestrzeni dualnej $V'$.*

*Dowód.* Oczywiście $f = 0 \in M^0$. Jeśli $f_1, f_2 \in M^0$ oraz $a_1, a_2 \in \mathbb{F}$, to dla każdego $v \in M$ mamy

$$(a_1 \cdot f_1 + a_2 \cdot f_2)(v) = a_1 \underbrace{f_1(v)}_{=0} + a_2 \underbrace{f_2(v)}_{=0} = 0$$

□

**Lemat 17.4.** *Załóżmy, że $M \subset V$ jest podprzestrzenią przestrzeni skończonego wymiaru $V$. Niech $v_1, \ldots, v_n$ będzie taką bazą $V$, że*

$$M = \operatorname{span}\{v_1, \ldots, v_m\}.$$

*Wtedy*

$$M^0 = \operatorname{span}\{v'_{m+1}, \ldots, v'_n\}.$$

*W szczególności,*

$$\dim M + \dim M^0 = \dim V = \dim V'.$$

*Dowód.* Z definicji bazy dualnej i anihilatora wynika, że

$$\operatorname{span}\{v'_{m+1}, \ldots, v'_n\} \subset M^0.$$

Niech $f \in M^0$ oraz $f = a_1 \cdot v'_1 + \ldots + a_n \cdot v'_n$. Dla $i = 1, \ldots, m$ mamy

$$0 = f(v_i) = a_i,$$

czyli $f \in \operatorname{span}\{v'_{m+1}, \ldots, v'_n\}$. □

U Jeśli $V$ jest skończonego wymiaru i $N \subset V'$ jest podprzestrzenią, to izomorfizm $j : V \longrightarrow V''$ pozwala utożsamić anihilator $N^0$ z podprzestrzenią przestrzeni $V$ daną przez

$$N^0 = \{v \in V : f(v) = 0 \quad \text{dla każdego} \quad f \in N\}.$$

Przy tej identyfikacji z lematu 17.4 wynika, że dla podprzestrzeni $M \subset V$ i $N \subset V'$ mamy

$$M^{00} = M, \quad N^{00} = N.$$

**Lemat 17.5.** *Załóżmy, że $V$ ma skończony wymiar oraz $V = M \oplus N$. Wtedy*

$$V' = M^0 \oplus N^0$$

*oraz odwzorowania zacieśnień*

$$V' \ni f \longmapsto f|_M \in M', \quad V' \ni f \longmapsto f|_N \in N'$$

*indukują izomorfizmy*

$$M^0 \ni f \longmapsto f|_N \in N', \quad N^0 \ni f \longmapsto f|_M \in M'.$$

*Dowód.* Wybieramy taką bazę $v_1, \ldots, v_n$ dla $V$, że

$$M = \operatorname{span}\{v_1, \ldots, v_m\}, \quad N = \operatorname{span}\{v_{m+1}, \ldots, v_n\}.$$

Z lematu 17.4 wynika, że

$$M^0 = \operatorname{span}\{v'_{m+1}, \ldots, v'_n\}, \quad N^0 = \operatorname{span}\{v'_1, \ldots, v'_m\}.$$

Dowodzi to, że $V' = M^0 \oplus N^0$. Zauważmy, że restrukcje $v'_{m+1}|_N, \ldots, v'_n|_N$ tworzą bazę dualną do bazy $v_{m+1}, \ldots, v_n$ dla $N$, co dowodzi drugiej części tezy. □

Zajmiemy się teraz przestrzenią $(\mathbb{R}^n)'$ dualną do przestrzeni euklidesowej $\mathbb{R}^n$. Zauważmy, że wektor $v \in \mathbb{R}^n$ definiuje funkcjonał $f_v \in (\mathbb{R}^n)'$ dany przez

$$f_v(u) = (v|u),$$

czyli $f_v$ jest mnożeniem skalarnym przez wektor $v$.

**Twierdzenie 17.2.** *Niech $L : V \longrightarrow W$ będzie odwzorowaniem liniowym przestrzeni skończonego wymiaru. Wtedy*

$$\ker L' = (\operatorname{im} L)^0 \quad \textit{oraz} \quad \operatorname{im} L' = (\ker L)^0.$$

*Dowód.* Udowodnimy równość $\ker L' = (\operatorname{im} L)^0$. Nie zależy ona od skończoności wymiarów - jest prawdziwa dla dowolnych przestrzeni wektorowych. Mamy

$$\begin{aligned}
\ker L' &= \{f \in W' : L'(f) = 0\} \\
&= \{f \in W' : f \circ L = 0\} \\
&= \{f \in W' : f(\operatorname{im} L) = \{0\}\} \\
&= \{f \in W' : \operatorname{im} L \subset \ker f\} \\
&= (\operatorname{im} L)^0.
\end{aligned}$$

Udowodnimy teraz drugą równość z tezy. Pokażemy najpierw, że $\operatorname{im} L'$ jest podprzestrzenią $(\ker L)^0$. Niech $f = L'(g) \in \operatorname{im} L'$ dla pewnego $g \in W'$. Dla $v \in \ker L$ mamy

$$\begin{aligned}
f(v) &= L'(g)(v) \\
&= g(L(v)) \\
&= g(0) = 0.
\end{aligned}$$

Dowód, że $\operatorname{im} L' \subset (\ker L)^0$ jest podprzestrzenią, nie zależał od skończoności wymiarów. Użyjemy tego założenia dla dowodu przeciwnej inkluzji. Wystarczy pokazać, że powyższe przestrzenie mają ten sam wymiar. Mamy

$$\begin{aligned}
\dim \operatorname{im} L' &= \dim, W' - \dim L' \quad \text{(formuła wymiaru)} \\
&= \dim W - \dim(\operatorname{im} L)^0 \quad \text{(twierdzenie 17.1 i pierwsza równość z tezy)} \\
&= \dim \operatorname{im} L \quad \text{(lemat 17.4)} \\
&= \dim V - \dim \ker L \quad \text{(formuła wymiaru)} \\
&= \dim(\ker L)^0 \quad \text{(lemat 17.4)}.
\end{aligned}$$

□

**Twierdzenie 17.3** (Riesza o postaci funkcjonału)**.** *Odwzorowanie*

$$R : \mathbb{R}^n \ni v \longmapsto f_v \in (\mathbb{R}^n)'$$

*jest izomorfizmem liniowym.*

*Dowód.* Odwzorowanie $R$ jest liniowe, bo jak łatwo sprawdzić

$$f_{v+w} = f_v + f_w, \quad f_{tv} = t \cdot f_v, \quad t \in \mathbb{F}, \quad v, w \in V.$$

Wystarczy pokazać, że $R$ jest monomorfizmem. Jeśli $R(v) = f_v = 0$, to

$$(v|u) = 0,$$

dla każdego $u \in V$. W szczególności, $(v|v) = 0$, czyli $v = 0$. □

**Przykład 17.6.** Wskażemy izomorfizm między przestrzenią dualną $\Big(M_{n\times m}(\mathbb{F})\Big)'$ a przestrzenią $M_{m\times n}(\mathbb{F})$. Dla macierzy $A \in M_{m\times n}(\mathbb{F})$ definiujemy $f_A \in \Big(M_{n\times m}(\mathbb{F})\Big)'$ wzorem

$$f_A(B) := \operatorname{tr}(AB) = \operatorname{tr}(BA), \quad B \in M_{n\times m}(\mathbb{F}).$$

Odwzorowanie liniowe

$$L : M_{m\times n}(\mathbb{F}) \ni A \longmapsto f_A \in \Big(M_{n\times m}(\mathbb{F})\Big)'$$

jest izomorfizmem. Rzeczywiście, dziedzina i przeciwdziedzina mają ten sam wymiar $nm$, więc wystarczy sprawdzić, że $L$ jest monomorfizmem. Przypuśćmy, że $f_A = 0$ dla pewnej macierz $A = [a_1|\dots|a_n]$. Wtedy $f_A(B) = 0$ dla każdej macierzy $B \in M_{n\times m}(\mathbb{F})$. W szczególności dla $B = A^T$ mamy

$$\begin{aligned} 0 &= f_A(A^T) \\ &= \operatorname{tr} A^T A \\ &= (a_1|a_1) + \dots + (a_n|a_n). \end{aligned}$$

Stąd $a_1 = \dots = a_n = 0$, czyli $A = 0$.

# Rozdział 18

# W telegraficznym skrócie

Układy równań

Twierdzenie Kroneckera-Capellego

Niech

$$A = \left[\begin{array}{c|c|c} a_1 & \ldots & a_n \end{array}\right] \in M_{m\times n}(\mathbb{F}), \quad b \in \mathbb{F}.$$

Następujące warunki są równoważne

(i) $Ax = b$ ma rozwiązanie,

(ii) $b \in \operatorname{span}\{a_1, \ldots, a_n\}$,

(iii) $\operatorname{rank}\left[\begin{array}{c|c} A & b \end{array}\right] = \operatorname{rank} A$.

Następujące warunki są równoważne

- $Ax = 0$ ma jednoznaczne rozwiązanie,
- $\ker A = \{0\}$.

Jeśli $A \in M_{n\times n}(\mathbb{F})$, to następujące warunki są równoważne

- $Ax = b$ ma jednoznaczne rozwiązanie dla każdego $b \in \mathbb{F}^n$,
- $A$ jest nieosobliwa.

Wtedy $x = A^{-1}b$. Współrzędne rozwiązania $x$ są też dane wzorami Cramera

$$x_i = \frac{\det A_i}{\det A},$$

gdzie $A_i$ jest macierzą powstałą z macierzy $A$ przez zastąpienie $i$-tej kolumny $A$ przez wektor $b$.

Wyznacznik i kilka ważnych wzorów

Niech $A \in M_{n\times n}(\mathbb{F})$.

- $\det(A) = \sum_{\sigma\in S_n} a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n}$,
- wyznacznik jest $n$-liniowym, antysymetrycznym odwzorowaniem kolumn $A$,
- rozwinięcie Laplace'a

$$\det A = a_{i1}A_{i1} + \ldots + a_{in}A_{in} = a_{1j}A_{1j} + \ldots + a_{nj}A_{nj},$$

  gdzie $A_{ij} = (-1)^{i+j} \det A(i,j)$ oraz $A(i,j) \in M_{(n-1)\times(n-1)}(\mathbb{F})$ jest macierzą otrzymaną z macierzy $A$ przez wykreślenie $i$-tego wiersza oraz $j$-tej kolumny.
- $A$ jest nieosobliwa wtedy i tylko wtedy, gdy $\det A \neq 0$.
- $A^{-1} = \frac{1}{\det(A)} \operatorname{adj} A$, gdzie

$$\operatorname{adj} A = \begin{bmatrix} A_{11} & A_{21} & \ldots & A_{n1} \\ A_{12} & A_{22} & \ldots & A_{n2} \\ \vdots & & & \vdots \\ A_{1n} & A_{2n} & \ldots & A_{nn} \end{bmatrix}.$$

- $|\det A|$ jest objętością równoległościanu rozpiętego na kolumnach macierzy $A$.
- $(AB)^T = B^T A^T$, o ile $AB$ ma sens,
- $\operatorname{rank} A = \operatorname{rank} A^T$, dla $A \in M_{n\times m}(\mathbb{R})$,
- $\det(AB) = \det(A) \cdot \det(B)$,
- $\det(A^{-1}) = \frac{1}{\det(A)}$,
- $\det A^T = \det A$,
- $(AB)^{-1} = B^{-1}A^{-1}$,
- $\operatorname{tr}(AB) = \operatorname{tr}(BA)$,
- $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.
- Dla wartości własnych $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ macierzy $A$ mamy

$$\operatorname{tr} A = \lambda_1 + \ldots + \lambda_n, \quad \det(A) = \lambda_1 \cdot \ldots \cdot \lambda_n.$$

- Zazwyczaj

$$AB \neq BA, \quad \det(A+B) \neq \det A + \det B.$$

- Dla macierzy $A \in M_{2\times 2}(\mathbb{C})$ oraz $\lambda \in \mathbb{C}$ mamy

$$\det(A - \lambda I) = \lambda^2 - \operatorname{tr}(A)\lambda + \det(A).$$

Macierze nieosobliwe

Dla macierzy $A = \left[\begin{array}{c|c|c} u_1 & \dots & u_n \end{array}\right] \in M_{n\times n}(\mathbb{R})$ następujące warunki są równoważne:

- $A$ jest nieosobliwa,
- $Ax = 0$ ma tylko zerowe rozwiązanie,
- $Ax = b$ ma rozwiązanie dla dowolnego $b \in \mathbb{R}^n$,
- $Ax = b$ ma dokładnie jedno rozwiązanie dla dowolnego $b \in \mathbb{R}^n$,
- $\det A \neq 0$,
- kolumny $a_1, \dots, a_n$ macierzy $A$ są liniowo niezależne,
- $\operatorname{span}\{a_1, \dots, a_n\} = \mathbb{R}^n$,
- kolumny $a_1, \dots, a_n$ tworzą bazę $\mathbb{R}^n$,
- $\ker A = \{0\}$,
- $\operatorname{rank} A = n$,
- $A^T$ jest nieosobliwa,
- $\lambda = 0$ nie jest wartością własną $A$,
- $(\ker A)^\perp = \mathbb{R}^n$,
- $(\operatorname{im} A)^\perp = \{0\}$,
- $A$ jest iloczynem macierzy elementarnych,
- zredukowana postać schodkowa $A$ jest równa $I_n$.

Podprzestrzenie fundamentalne macierzy $A$

Niech $A \in M_{n\times k}(\mathbb{R})$. Podprzestrzenie fundamentalne, to

$$\ker A, \quad \operatorname{im} A^T \subset \mathbb{R}^k,$$

$$\ker A^T, \quad \operatorname{im} A \subset \mathbb{R}^n.$$

Mamy ortogonalne sumy proste

$$\mathbb{R}^k = \ker A \oplus \operatorname{im} A^T, \qquad \mathbb{R}^n = \ker A^T \oplus \operatorname{im} A,$$

czyli

$$\ker A = (\operatorname{im} A^T)^\perp, \quad \ker A^T = (\text{imA})^\perp.$$

Niezmienniki podobieństwa macierzy

Macierze podobne $A$ i $S^{-1}AS$ mają takie same:

- wyznaczniki,
- rzędy,
- wielomiany charakterystyczne,
- wartości własne,
- ślady,
- wymiary podprzestrzeni własnych dla wartości własnej $\lambda$.

Rozkład Jordana macierzy zespolonej $A$

Niech $A \in M_{n\times n}(\mathbb{C})$.

- Wielomian charakterystyczna $p_A(x) = \det(A - xI)$ ma $n$ pierwiastków zespolonych $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ liczonych z krotnościami.
- Istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{C})$, że

  $$\Lambda = S^{-1}AS = \mathrm{diag}(\Lambda_1, \ldots, \Lambda_k),$$

  gdzie każdy blok Jordana $\Lambda_i$ $(i = 1, \ldots, k)$ jest macierzą postaci

  $$\Lambda_i = \begin{bmatrix} \boxed{\lambda} & \mathbf{1} & 0 & \ldots & 0 \\ 0 & \boxed{\lambda} & \mathbf{1} & \ldots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & \boxed{\lambda} & \mathbf{1} \\ 0 & \ldots & 0 & 0 & \boxed{\lambda} \end{bmatrix}$$

  z wartością własną $\lambda \in \mathbb{C}$ na przekątnej.
- Wymiar $\dim\ker(A - \lambda I)$ podprzestrzeni własnej dla wartości własnej $\lambda$ jest równy liczbie bloków Jordana $\Lambda_i$ z $\lambda$ na przekątnej.

Rzeczywista postać Jordana

Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą rzeczywistą. Istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{R})$, że

$$\Lambda = S^{-1}AS = \mathrm{diag}(\Lambda_1, \ldots, \Lambda_k),$$

gdzie każdy blok $\Lambda_i$ jest jednej z dwóch postaci

$$\begin{bmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ldots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda & 1 \\ 0 & \ldots & 0 & 0 & \lambda \end{bmatrix}, \qquad \begin{bmatrix} B & I_2 & 0 & \ldots & 0 \\ 0 & B & I_2 & \ldots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & B & I_2 \\ 0 & \ldots & 0 & 0 & B \end{bmatrix},$$

dla $B = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$ przy pewnej rzeczywistej wartości własnej $\lambda$ i zespolonej wartości własnej $\alpha \pm i\beta$ macierzy $A$.

Diagonalizacja

Niech $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Dla macierzy $A \in M_{n\times n}(\mathbb{F})$ następujące warunki są równoważne

- istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{F})$, że $S^{-1}AS$ jest diagonalna,
- istnieje baza $v_1, \ldots, v_n \in \mathbb{F}^n$ złożona z wektorów własnych $A$. $Av_i = \lambda_i \cdot v_i$

W szczególności, jest tak gdy $A$ ma $n$ różnych wartości własnych $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$.

Dla macierzy $A \in M_{n\times n}(\mathbb{R})$ następujące warunki są równoważne

- $A = A^T$,
- istnieje taka macierz ortogonalna $Q \in M_{n\times n}(\mathbb{R})$, że $Q^T AQ$ jest diagonalna, $Q^{-1} = Q^T$
- istnieje baza ortonormalna $v_1, \ldots, v_n \in \mathbb{R}^n$ złożona z wektorów własnych $A$. $(u_i|v_j) = \delta_{ij}$

Dla macierzy $A \in M_{n\times n}(\mathbb{C})$ następujące warunki są równoważne

- $A^*A = AA^*$, (macierz normalna)
- istnieje taka macierz unitarna $U \in M_{n\times n}(\mathbb{C})$, że $U^*AU$ jest diagonalna, $U^{-1} = U^*$
- istnieje baza hermitowsko ortonormalna $v_1, \ldots, v_n \in \mathbb{C}^n$ złożona z wektorów własnych $A$, czyli
$$(u_i|v_j)_* = \delta_{ij}$$

Macierz ortogonalna
$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$
nie ma rzeczywistych wektorów własnych, więc nie jest diagonalizowalna jako macierz rzeczywista. Jej wartościami własnymi są $\pm i$. Wektorem własnym dla $i$ jest wektor $[1, i]^T$, bo
$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} i \\ -1 \end{bmatrix} = i \cdot \begin{bmatrix} 1 \\ i \end{bmatrix}.$$
Dla macierzy unitarnej
$$U = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$$
mamy
$$U^{-1}AU = U^*AU = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

Algorytm Grama–Schmidta
Niech $v_1, \ldots, v_l \in \mathbb{R}^n$ będą liniowo niezależne.

---

**Krok 1.** Definiujemy
$$u_1 = \frac{v_1}{\|v_1\|}$$

**Krok 2.** Definiujemy rekurencyjnie wektory $u_2, \ldots, u_l$ jako
$$u_{k+1} = \frac{v_{k+1} - p_k}{\|v_{k+1} - p_k\|} \qquad p_k = (v_{k+1}|u_1) \cdot u_1 + \ldots + (v_{k+1}|u_k) \cdot u_k$$

tzn. $p_k$ jest rzutem ortogonalnym $v_{k+1}$ na podprzestrzeń $\operatorname{span}\{u_1, \ldots, u_k\}$.

Wtedy

(1) $\operatorname{span}\{u_1, \ldots, u_k\} = \operatorname{span}\{v_1, \ldots, v_k\}$,

(2) $u_1, \ldots, u_l$ jest układem ortonormalnym.

Gram–Schmidt w zapisie macierzowym
Niech $v_1, \ldots, v_n \in \mathbb{R}^m$ będą liniowo niezależne.

---

Procedura Grama–Schmidta może być wyrażona formułami
$$u_k = \frac{(I - U_k U_k^T) v_k}{\|(I - U_k U_k^T) v_k\|}, \quad k = 1, \ldots, n,$$
gdzie $U_1 = 0 \in M_{m\times 1}(\mathbb{R})$ oraz
$$U_k = \left[\, u_1 \,\middle|\, \ldots \,\middle|\, u_{k-1} \,\right] \in M_{m\times(k-1)}(\mathbb{R}), \quad k > 1.$$

Rozkład $QR$ macierzy $A$ o liniowo niezależnych kolumnach

$$\underbrace{\begin{bmatrix} | & | & | \\ | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \\ | & | & | \end{bmatrix}}_{\mathbf{A}} = \underbrace{\begin{bmatrix} | & | & | \\ | & | & | \\ u_1 & u_2 & u_3 \\ | & | & | \\ | & | & | \end{bmatrix}}_{\mathbf{Q}} \underbrace{\begin{bmatrix} \|v_1\| & (v_2|u_1) & (v_3|u_1) \\ 0 & \|v_2 - p_1\| & (v_3|u_2) \\ 0 & 0 & \|v_3 - p_2\| \end{bmatrix}}_{\mathbf{R}}.$$

Kolumny $v_1$, $v_2$, $v_3$ macierzy $A$ są liniowo niezależne. Wektory $u_1$, $u_2$, $u_3$ stanowią układ ortonormalny otrzymany z $v_1$, $v_2$, $v_3$ w wyniku procedury Grama–Schmidta.

$$\begin{aligned} v_1 &= \|v_1\| \cdot u_1 \\ v_2 &= \underbrace{(v_2|u_1) \cdot u_1}_{=p_1} + \|v_2 - p_1\| \cdot u_2 \\ v_3 &= \underbrace{(v_3|u_1) \cdot u_1 + (v_3|u_2) \cdot u_2}_{=p_2} + \|v_3 - p_2\| \cdot u_3 \end{aligned}$$

Tak samo jest dla większej liczby wektorów liniowo niezależnych.

Macierz Grama $A^TA$

Niech $A = \begin{bmatrix} a_1 \mid \ldots \mid a_k \end{bmatrix} \in M_{n\times k}(\mathbb{R})$. Macierz Grama jest macierzą symetryczną

$$A^TA = \begin{bmatrix} (a_1|a_1) & \ldots & (a_1|a_k) \\ \vdots & \ddots & \vdots \\ (a_k|a_1) & \ldots & (a_k|a_k) \end{bmatrix} \in M_{k\times k}(\mathbb{R}).$$

- wartości własne $A^TA$ są nieujemne.
- $\det A^TA \geqslant 0$ oraz $\operatorname{tr} A^TA \geqslant 0$.
- $\ker A = \ker A^TA$ oraz $\operatorname{im} A^TA = \operatorname{im} A^T$.
- $\operatorname{rank} A^TA = \operatorname{rank} A = \operatorname{rank} A^T = \operatorname{rank} AA^T$.
- mamy rozkład na ortogonalną sumę prostą
$$\mathbb{R}^k = \ker A^TA \oplus \operatorname{im} A^TA.$$
- kolumny macierzy $A$ są liniowo niezależne wtedy i tylko wtedy, gdy $\det A^TA > 0$.
- $\sqrt{\det A^TA}$ jest objętością wielościanu rozpiętego na kolumnach macierzy $A$.

Rozkład polarny macierzy nieosobliwej

Niech $A \in M_{n\times n}(\mathbb{R})$ będzie nieosobliwa.

- Macierz Grama $A^TA$ ma dodatnie wartości własne $\lambda_1, \ldots, \lambda_n$.
- $A = B \operatorname{diag}(\lambda_1, \ldots, \lambda_n)B^T$ dla pewnej macierzy $B \in O(n)$.
- $P = B \operatorname{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})B^T$ jest dodatnio określona.
- Macierz $Q = AP^{-1}$ jest ortogonalna.
- $A = QP$.

Macierze dodatnio określone

Dla macierzy symetrycznej $A \in M_{n\times n}(\mathbb{R})$ poniższe warunki są równoważne

(i) $A$ jest dodatnio określona,

(ii) wiodące minory macierzy $A$ mają dodatnie wyznaczniki,

(iii) $A = LU$, gdzie $L$ jest dolnie trójkątna z 1 na przekątnej oraz $U$ jest górnie trójkątna z dodatnimi wyrazami diagonalnymi,

(iv) $A = LL^T$, gdzie $L$ jest dolnie trójkątna z dodatnimi wyrazami na przekątnej,

(v) $A = B^TB$ dla pewnej macierzy nieosobliwej $B$.

Rozkład singularny

Niech $A \in M_{n\times m}(\mathbb{R})$ ma rząd $r = \operatorname{rank} A$.

Mamy rozkłady na ortogonalne sumy proste

$$\mathbb{R}^m = \underbrace{\ker A}_{=\ker A^TA} \oplus \underbrace{\operatorname{im} A^T}_{=\operatorname{im} A^TA}, \qquad \mathbb{R}^n = \underbrace{\ker A^T}_{=\ker AA^T} \oplus \underbrace{\operatorname{im} A}_{=\operatorname{im} AA^T}.$$

Macierz $A^TA$ ma bazę ortonormalną wektorów własnych

$$\underbrace{u_1, \ldots, u_r,}_{\in \operatorname{im} A^TA} \quad \underbrace{u_{r+1}, \ldots, u_m}_{\in \ker A^TA}.$$

Wartości własne $A^TA$ są nieujemne i $\lambda_1, \ldots, \lambda_r$ odpowiadające wektorom własnym $u_1, \ldots, u_r$ są dodatnie.

Wartościami singularnymi macierzy $A$ są

$$\sigma_1 = \sqrt{\lambda_1} \geqslant \ldots \geqslant \sigma_r = \sqrt{\lambda_r} > 0.$$

Macierz $AA^T$ ma bazę ortonormalną wektorów własnych postaci

$$\underbrace{v_1 = \frac{1}{\sigma_1} \cdot Au_1, \ldots, v_r = \frac{1}{\sigma_r} \cdot Au_r,}_{\in \operatorname{im} AA^T} \quad , \underbrace{v_{r+1}, \ldots, v_n}_{\in \ker AA^T}.$$

Dla macierzy ortogonalnych

$$U = \left[\begin{array}{c|c|c} u_1 & \ldots & u_m \end{array}\right] \in O(m), \quad V = \left[\begin{array}{c|c|c} v_1 & \ldots & v_n \end{array}\right] \in O(n),$$

oraz quasi-diagonalnej

$$D = \left[\begin{array}{c|c} \operatorname{diag}(\sigma_1, \ldots, \sigma_r) & 0 \\ \hline 0 & 0 \end{array}\right] \in M_{n\times m}(\mathbb{R}),$$

mamy

$$A = VDU^T.$$

# Część III

# Algebra liniowa z geometrią 3

# Rozdział 19

# Grupa unitarna i kwaterniony

SŁOWA KLUCZOWE ROZDZIAŁU
grupa unitarna ♢ specjalna grupa unitarna SU(2) ♢ kwaterniony macierzowe ♢ kwaterniony ♢ grupa $SO(3)$

**Definicja 19.1.** **Grupa unitarna** $\mathrm{U}(n)$

*Grupa unitarna* jest zdefiniowana jako zbiór

$$\mathrm{U}(n) = \Big\{A \in M_{n\times n}(\mathbb{C}) : AA^* = I\Big\},$$

z działaniem mnożenia macierzy.

U Definicja jest poprawna, bo

- jeśli $A, B \in \mathrm{U}(n)$, to $AB \in \mathrm{U}(n)$,
- $I \in \mathrm{U}(n)$,
- jeśli $A \in \mathrm{U}(n)$, to $A^{-1} = A^* \in \mathrm{U}(n)$.

Ponadto dla $A \in \mathrm{U}(n)$ mamy

$$\begin{aligned} 1 &= \det AA^* \\ &= \det A \cdot \det A^* \\ &= \det A \cdot \overline{\det A} \\ &= |\det A|, \end{aligned}$$

czyli $\det A \in \mathbb{S}^1$.

*Specjalną grupę unitarną* definiujemy jako

$$\mathrm{SU}(n) = \Big\{A \in \mathrm{U}(n) : \det A = 1\Big\}.$$

**Przykład 19.1.** Jeśli $A \in M_{1\times 1}(\mathbb{C})$, to $A = [z]$ dla pewnego $z \in \mathbb{C}$ oraz $|z| = |\det A| = 1$, czyli jako zbiór $\mathrm{U}(1) = \mathbb{S}^1$. Odwzorowanie identycznościowe

$$\mathrm{U}(1) \ni [z] \longmapsto z \in \mathbb{S}^1$$

oczywiście jest zgodne z mnożeniem w tych grupach, więc jest izomorfizmem grup. Stąd

$$\mathrm{U}(1) \cong \mathbb{S}^1, \quad \mathrm{SU}(1) = \{1\}.$$

Odwzorowanie

$$\mathbb{S}^1 \ni \cos\theta + i\sin\theta \longmapsto \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \in SO(2)$$

jest izomorfizmem grup, czyli

$$\mathrm{U}(1) \cong \mathbb{S}^1 \cong SO(2).$$

Przyglądniemy się teraz bliżej grupie SU(2).

**Lemat 19.1.** *Jeśli* $A \in \mathrm{SU}(2)$*, to*

$$A = \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix},$$

*dla pewnych* $z, w \in \mathbb{C}$ *spełniających warunek* $|z|^2 + |w|^2 = 1$*.*

*Dowód.* Niech

$$A = \begin{bmatrix} z_1 & w_1 \\ w_2 & z_2 \end{bmatrix} \in \mathrm{SU}(2).$$

Ponieważ $\det A = 1$, więc

$$A^{-1} = \begin{bmatrix} z_2 & -w_1 \\ -w_2 & z_1 \end{bmatrix}.$$

Stąd,

$$\underbrace{\begin{bmatrix} z_2 & -w_1 \\ -w_2 & z_1 \end{bmatrix}}_{A^{-1}} = \underbrace{\begin{bmatrix} \overline{z_1} & \overline{w_2} \\ \overline{w_1} & \overline{z_2} \end{bmatrix}}_{A^*},$$

czyli

$$z_2 = \overline{z_1}, \quad w_2 = -\overline{w_1}.$$

□

**Wniosek 19.1.**

$$\mathrm{SU}(2) = \left\{ \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} : z, w \in \mathbb{C} \textit{ oraz } |z|^2 + |w|^2 = 1 \right\}.$$

Rozważmy sferę

$$\begin{aligned} \mathbb{S}^3 &= \left\{ [x_1, x_2, x_3, x_4]^T \in \mathbb{R}^4 : x_1^2 + x_2^2 + x_3^2 + x_4^2 \right\} \\ &= \left\{ [z, w]^T \in \mathbb{C}^2 : z = x_1 + ix_2,\ w = x_3 + ix_4,\ |z|^2 + |w|^2 = 1 \right\}. \end{aligned}$$

Z wniosku 19.1 wynika, że odwzorowanie

$$h : \mathrm{SU}(2) \ni \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} \longmapsto [z, w]^T \in \mathbb{S}^3$$

jest bijekcją, więc jako zbiory SU(2) i $\mathbb{S}^3$ są identyczne. Zbiór SU(2) ma dodatkową strukturę algebraiczną – jest grupą z mnożeniem macierzy. Możemy użyć bijekcji $h$, aby zdefiniować mnożenie w zbiorze $\mathbb{S}^3$. Przyjmujemy, że

$$h(A) \cdot h(B) := h(AB).$$

Zobaczmy jak to mnożenie wygląda. Niech

$$A = \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix}, \quad B = \begin{bmatrix} z_1 & w_1 \\ -\overline{w_1} & \overline{z_1} \end{bmatrix}.$$

Wtedy

$$AB = \begin{bmatrix} zz_1 - w\overline{w_1} & zw_1 + w\overline{z_1} \\ -(\overline{zw_1 + w\overline{z_1}}) & \overline{zz_1 - w\overline{w_1}} \end{bmatrix}.$$

Stąd

$$\begin{aligned} [z, w]^T \cdot [z_1, w_1]^T &= h(A) \cdot h(B) \\ &= h(AB) \\ &= [zz_1 - w\overline{w_1}, zw_1 + w\overline{z_1}]^T. \end{aligned}$$

Z tak określonym działaniem mnożenia $\mathbb{S}^3$ jest grupą, zwaną *grupą kwaternionów jednostkowych.*

Zauważmy, że powyższa formuła pozwala nam zdefiniować mnożenie nie tylko w $\mathbb{S}^3$, ale również w $\mathbb{C}^2$. Możemy po prostu przyjąć, że

$$[z, w]^T \cdot [z_1, w_1]^T := [zz_1 - w\overline{w_1}, zw_1 + w\overline{z_1}]^T.$$

Zwróćcie uwagę na analogię do konstrukcji ciała liczb zespolonych $\mathbb{C}$. W zbiorze $\mathbb{R}^2$ zdefiniowaliśmy mnożenie, korzystając z działań w $\mathbb{R}$. Teraz robimy to samo dla $\mathbb{C}^2$. Oczywiście w zbiorze $\mathbb{C}^2$ mamy też naturalnie zdefiniowane dodawanie. Naturalne jest więc pytanie, czy trójka $(\mathbb{C}^2, +, \cdot)$ jest ciałem? Okazuje się, że nie, bo mnożenie w $\mathbb{C}^2$ nie jest przemienne. Trójkę $(\mathbb{C}^2, +, \cdot)$ moglibyśmy nazwać ciałem nieprzemiennym. Są to tzw. *kwaterniony*, oznaczane przez $\mathbb{H}$. Struktura ta została odkryta przez Hamiltona w 1843 roku.

Ponieważ $\mathbb{C}^2$ możemy utożsamić z $\mathbb{R}^4$ poprzez

$$[x_1 + ix_2, x_3 + ix_4]^T \longmapsto [x_1, x_2, x_3, x_4]^T,$$

więc kwaterniony możemy traktować jako czwórki liczb rzeczywistych. W tej konwencji działania są zdefiniowane przez

$$[a_1, b_1, c_1, d_1]^T + [a_2, b_2, c_2, d_2]^T = [a_1 + a_2, b_1 + b_2, c_1 + c_2, d_1 + d_2]^T,$$

$$\begin{aligned}
&[a_1, b_1, c_1, d_1]^T \cdot [a_2, b_2, c_2, d_2]^T = \\
&[a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2, a_1b_2 + b_1a_2 + c_1d_2 - d_1c_2, \\
&a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2, a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2]^T.
\end{aligned}$$

Przez analogię do liczb zespolonych kwaterniowy możemy zapisać w postaci algebriacznej

$$a_1 \cdot 1 + a_2 \cdot i + a_3 \cdot j + a_4 \cdot k,$$

gdzie

$$1 = [1, 0, 0, 0]^T, \quad i = [0, 1, 0, 0]^T, \quad j = [0, 0, 1, 0]^T, \quad k = [0, 0, 0, 1]^T.$$

W konwencji macierzowej, wyróżnionym elementom odpowiadają macierze

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

Łatwo sprawdzamy, że zachodzą równości

$$i \cdot 1 = 1 \cdot i = i, \quad j \cdot 1 = 1 \cdot j = j, \quad k \cdot 1 = 1 \cdot k = k.$$

Kwadraty pozostałych symboli są równe $-1$:

$$i^2 = j^2 = k^2 = -1.$$

Iloczyn dwóch symboli spośród $i, j, k$ jest równy plus/minus trzeciemu:

$$i \cdot j = k, \quad j \cdot k = i, \quad k \cdot i = j,$$

$$j \cdot i = -k, \quad k \cdot j = -i, \quad i \cdot k = -j.$$

Reguła mnożenia symboli $1, i, j, k$ rozszerza się na $\mathbb{R}^4$, dając formułę

$$\begin{aligned}
&(a + bi + cj + dk) \cdot (x + yi + zj + wk) = \\
&(ax - by - cz - dw) + (ay + bx + cw - dz)i + \\
&(az + cx + dy - bw)j + (aw + dx + bz - cy)k.
\end{aligned}$$

Mnożenie kwaternionów bez składowej $k$ i $j$ zgadza się z mnożeniem liczb zespolonych. Mamy więc inkluzje

$$\mathbb{R} \subset \mathbb{C} \subset \mathbb{H}.$$

Liczbę zespoloną przedstawiamy w postaci $z = a + bi$. Podobnie, każdy kwaternion $q = a + bi + cj + dk$ można zapisać w postaci $q = z + wj$ dla pewnych $z, w \in \mathbb{C}$, bo

$$(a + bi) + (c + di)j = a + bi + cj + dk = q.$$

Zajmiemy się teraz związkiem pomiędzy SU(2) (czyli kwaternionami jednostkowymi), a grupą obrotów $SO(3)$. W tym celu wygodnie jest wprowadzić jeszcze jedną konwencję zapisu kwaternionu.

Dla kwaternionu $q = [q_0, q_1, q_2, q_3]^T$ stosujemy notację

$$q = [q_0, \mathbf{q}]^T, \quad \mathbf{q} = [q_1, q_2, q_3]^T \in \mathbb{R}^3.$$

Bezpośredni rachunek pokazuje, że w tej konwencji iloczyn kwaternionów możemy zapisać przy pomocy formuły

$$\begin{aligned} p \cdot q &= [p_0, p_1, p_2, p_3]^T \cdot [q_0, q_1, q_2, q_3]^T \\ &= [p_0 q_0 - (\mathbf{p}|\mathbf{q}), p_0\mathbf{q} + q_0\mathbf{p} + \mathbf{p} \times \mathbf{q}]^T. \end{aligned}$$

Iloczyn skalarny kwaternionów jest dany przez

$$(p|q) = p_0q_0 + p_1q_1 + p_2q_2 + p_3q_3 = p_0q_0 + (\mathbf{p}|\mathbf{q}).$$

Definiujemy

$$\overline{q} = [q_0, -q_1, -q_2, -q_3]^T = [q_0, -\mathbf{q}]^T, \quad |q| = \sqrt{(q|q)} = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2}.$$

Bezpośredni rachunek pokazuje, że

$$q \cdot \overline{q} = \overline{q} \cdot q = (|q|^2, \mathbf{0}).$$

Wynika stąd, że dla $q \neq 0$ mamy

$$q^{-1} = \frac{\overline{q}}{|q|^2}.$$

Przyglądniemy się teraz grupie obrotów $SO(3)$. Wyróżnimy trzy obroty względem osi standardowego układu współrzędnych

$$\begin{aligned} R_x(\theta) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}, \\ R_y(\theta) &= \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}, \\ R_z(\theta) &= \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Jak wiemy 1 jest wartością własną obrotu $R \in SO(3)$. Niech $n = [n_1, n_2, n_3]^T \in \mathbb{R}^3$ będzie jednostkowym wektorem własnym $R$. Wtedy obrót $R$ jest obrotem wokół prostej generowanej przez wektor $n$ o kąt $\theta$. Znajdziemy jawną formułę na $R$ w zależności od $n$ i $\theta$. Ponieważ $n$ jest wektorem jednostkowym, więc możemy go zapisać (we współrzędnych sferycznych) w postaci

$$n = [\cos\alpha \sin\beta, \sin\alpha \sin\beta, \cos\beta]^T,$$

dla pewnych $0 \leqslant \alpha < 2\pi$ oraz $0 \leqslant \beta \leqslant \pi$. Sprawdzamy łatwo, że

$$n = R_z(\alpha) R_y(\beta) e_3.$$

Obrót $R$ możemy teraz wyrazić przez

$$R = R_z(\alpha)R_y(\beta)R_z(\theta)R_y^T(\beta)R_z^T(\alpha),$$

czyli obracamy najpierw układ współrzędnych, aby oś obrotu $n$ przeszła na $e_3$, dokonujemy obrotu o $\theta$ wokół osi $z$ i obracamy ponownie $e_3$ na $n$. Wstawiając formuły na $R_x$, $R_y$, $R_z$ oraz stosując oznaczenie $c = \cos\theta$, $s = \sin\theta$, otrzymujemy, że $R = R(\theta, n)$ ma postać

$$R = \begin{bmatrix} c + (n_1)^2(1-c) & n_1n_2(1-c) - sn_3 & n_1n_3(1-c) + sn_2 \\ n_2n_1(1-c) + sn_3 & c + (n_2)^2(1-c) & n_2n_3(1-c) - sn_1 \\ n_3n_1(1-c) - sn_2 & n_3n_2(1-c) + sn_1 & c + (n_3)^2(1-c) \end{bmatrix}.$$

Możemy ten wzór zapisać jeszcze trochę inaczej, podstawiając $a = \cos\theta/2$ i $b = \sin\theta/2$:

$$R = \begin{bmatrix} a^2 - b^2 + (n_1)^2(2b^2) & 2b^2n_1n_2 - 2abn_3 & 2b^2n_1n_3 + 2abn_2 \\ 2b^2n_2n_1 + 2abn_3 & a^2 - b^2 + 2b^2(n_2)^2 & 2b^2n_2n_3 - 2abn_1 \\ 2b^2n_3n_1 - 2abn_2 & 2b^2n_3n_2 + 2abn_1 & a^2 - b^2 + 2b^2(n_3)^2 \end{bmatrix}.$$

Ponieważ $n_1^2 + n_2^2 + n_3^2 = 1$, więc otrzymujemy, że

$$R = \begin{bmatrix} a^2 + b^2(n_1^2 - n_2^2 - n_3^2) & 2b^2n_1n_2 - 2abn_3 & 2b^2n_1n_3 + 2abn_2 \\ 2b^2n_2n_1 + 2abn_3 & a^2 + b^2(n_2^2 - n_3^2 - n_1^2) & 2b^2n_2n_3 - 2abn_1 \\ 2b^2n_3n_1 - 2abn_2 & 2b^2n_3n_2 + 2abn_1 & a^2 + b^2(n_3^2 - n_1^2 - n_2^2) \end{bmatrix}.$$

Rozważmy teraz kwaternion jednostkowy $q = [q_0, q_1, q_2, q_3]^T \in \mathbb{S}^3$. Sferę $\mathbb{S}^3$ możemy sparametryzować za pomocą $0 \leqslant \theta \leqslant 4\pi$ i $n \in \mathbb{S}^2$, przyjmując

$$\begin{aligned} q_0 &= a = \cos\theta/2, \\ q_1 &= n_1 b = n_1 \sin\theta/2, \\ q_2 &= n_2 b = n_2 \sin\theta/2, \\ q_3 &= n_3 b = n_3 \sin\theta/2. \end{aligned}$$

Wtedy

$$R(q) = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2q_1q_2 - 2q_0q_3 & 2q_1q_3 + 2q_0q_2 \\ 2q_1q_2 + 2q_0q_3 & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2q_2q_3 - 2q_0q_1 \\ 2q_1q_3 - 2q_0q_2 & 2q_2q_3 + 2q_0q_1 & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix}.$$

Obrót $R(q) \in SO(3)$ możemy wyrazić za pomocą mnożenia kwaternionów. Dla $q \in \mathbb{S}^3$ oraz $x = [x_1, x_2, x_3]^T \in \mathbb{R}^3$ bezpośredni rachunek pokazuje, że

$$R(q)x = q \cdot [0, x_1, x_2, x_3]^T \cdot \overline{q}.$$

**Twierdzenie 19.1.** *Odwzorowanie*

$$f : \mathbb{S}^3 \ni q \longmapsto R(q) \in SO(3)$$

*jest surjektywnym homomorfizmem grupy kwaternionów jednostkowych i grupy obrotów. Ponadto $R(q) = R(-q)$.*

*Dowód.* Niech $q_1, q_2 \in \mathbb{S}^3$. Mamy pokazać, że

$$f(q_1 \cdot q_2) = f(q_1)f(q_2).$$

Dla $x \in \mathbb{R}^3$ mamy

$$\begin{aligned} f(q_1 \cdot q_2)x &= (q_1 \cdot q_2) \cdot [0, x_1, x_2, x_3]^T \cdot \overline{(q_1 \cdot q_2)} \\ &= (q_1 \cdot q_2) \cdot [0, x_1, x_2, x_3]^T \cdot (q_1 \cdot q_2)^{-1} \\ &= q_1 \cdot \underbrace{q_2 \cdot [0, x_1, x_2, x_3]^T \cdot q_2^{-1}}_{f(q_2)x} \cdot q_1^{-1} \\ &= (f(q_1)f(q_2))x. \end{aligned}$$

$\square$

# Rozdział 20

# Iloraz Rayleigha

| | |
|---|---|
| Słowa kluczowe rozdziału<br>iloraz Rayleigha ◊ twierdzenie Couranta–Fischera ◊ nierówność Weyla ◊ twierdzenie Cauchy'ego | $\lambda_1 = \max_{x \neq 0} \frac{(Ax\|x)}{(x\|x)}$<br>$\lambda_n = \min_{x \neq 0} \frac{(Ax\|x)}{(x\|x)}$ |

**Twierdzenie 20.1** (Rayleigh)**.** *Niech $Q(x) = (Ax|x)$ będzie formą kwadratową skojarzoną z macierzą symetryczną $A \in M_{n\times n}(\mathbb{R})$. Niech $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$ będą wartościami własnymi $A$. Wtedy*

*(i)* $\lambda_n \leqslant Q(x) \leqslant \lambda_1$ *dla* $\|x\| = 1$,

*(ii) Jeśli $v_1$ jest jednostkowym wektorem własnym $A$ odpowiadającym wartości własnej $\lambda_1$, to*

$$Q(v_1) = \lambda_1 = \max_{\|x\|=1} Q(x),$$

*(iii) Jeśli $v_n$ jest jednostkowym wektorem własnym $A$ odpowiadającym wartości własnej $\lambda_n$, to*

$$Q(v_n) = \lambda_n = \min_{\|x\|=1} Q(x).$$

Obserwacja to pochodzi od fizyka brytyjskiego Lorda Rayleigha (1842–1919).

*Dowód.* Niech $Q = \left[\begin{array}{c|c|c} u_1 & \ldots & u_n \end{array}\right]$ będzie taką macierzą ortogonalną, że

$$Q^T A Q = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

Wektory $u_1, \ldots, u_n$ tworzą bazę ortonormalną złożoną z wektorów własnych dla $\lambda_1, \ldots, \lambda_n$. Niech $x$ będzie wektorem jednostkowym. Istnieje dokładnie jeden taki wektor $y$, że $x = Q(y)$. Ponieważ $Q$ jest ortogonalna, więc $\|x\| = \|Q(y)\| = \|y\|$. Stąd,

$$\begin{aligned} Q(x) = (Ax|x) &= (AQ(y)|Q(y)) \\ &= (Q^T AQ(y)|y) \\ &= (Dy|y) \\ &= \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2 \\ &\leqslant \lambda_1(y_1^2 + \ldots + y_n^2) \\ &= \lambda_1. \end{aligned}$$

Ten sam argument pokazuje, że $Q(x) \geqslant \lambda_n$. □

U Równości

$$\lambda_1 = \max_{\|x\|=1} (Ax|x), \quad \lambda_n = \min_{\|x\|=1} (Ax|x),$$

są często zapisywane w równoważnej formie

$$\lambda_1 = \max_{x\neq 0} \frac{(Ax|x)}{(x|x)}, \quad \lambda_n = \min_{x\neq 0} \frac{(Ax|x)}{(x|x)}.$$

Wyrażenie $\frac{(Ax|x)}{(x|x)}$ jest nazywane *ilorazem Rayleigha.*

**Twierdzenie 20.2** (Courant–Fischer)**.** *Niech*

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$$

*będą wartościami własnymi macierzy symetrycznej $A \in M_{n\times n}(\mathbb{R})$. Wtedy*

$$\lambda_i = \max_{\dim V=i} \quad \min_{\substack{x\in V\\ \|x\|=1}} (Ax|x),$$

$$= \min_{\dim V=n-i+1} \quad \max_{\substack{x\in V\\ \|x\|=1}} (Ax|x).$$

*Dowód.* Udowodnimy tylko wersję min-max. Dowód wersji max-min jest analogiczny. Ponieważ dla $x = Q(y)$ i $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ mamy $\|x\| = \|y\|$ oraz

$$(Ax|x) = (Dy|y),$$

więc wystarczy udowodnić, że

$$\lambda_i = \min_{\dim V=n-i+1} \max_{\substack{y\in V\\ \|y\|=1}} (Dy|y).$$

Dla podprzestrzeni $V$ wymiaru $n-i+1$ rozważamy

$$S_V = \{y \in V : \|y\| = 1\}, \quad S'_V = \{y \in V \cap F : \|y\| = 1\},$$

gdzie $F = \operatorname{span}\{e_1, e_2, \ldots, e_i\}$. Zauważmy, że $V \cap F \neq \{0\}$, bo w przeciwnym razie

$$\dim(V + F) = \dim V + \dim F = n + 1,$$

co jest niemożliwe. Zauważmy, że $S'_V$ składa się z wektorów w $S_V$, które są postaci

$$y = [y_1, \ldots, y_i, 0, \ldots, 0]^T, \quad y_1^2 + \ldots + y_i^2 = 1.$$

Stąd dla $y \in S'_V$ mamy

$$\begin{aligned}(Dy|y) &= \lambda_1 y_1^2 + \ldots + \lambda_i y_i^2\\ &\geqslant \lambda_i(y_1^2 + \ldots + y_i^2)\\ &= \lambda_i.\end{aligned}$$

Ponieważ $S'_V \subset S_V$, więc

$$\max_{S_V}(Dy|y) \geqslant \max_{S'_V}(Dy|y) \geqslant \lambda_i,$$

czyli

$$\min_V \max_{S_V}(Dy|y) \geqslant \lambda_i.$$

Dla $V = F$ i $y = e_1$ mamy $(De_i|e_i) = \lambda_i$, więc

$$\min_V \max_{S_V}(Dy|y) = \lambda_i.$$

□

U Twierdzenie Couranta–Fischera ma następujące alternatywne („ortogonalne”) sformułowanie: dla $1 < i < n$ mamy

$$\lambda_i = \max_{v_1,\ldots,v_{n-i}} \min_{\substack{x \perp v_1,\ldots,v_{n-i} \\ \|x\|=1}} (Ax|x),$$

$$\lambda_i = \min_{v_i,\ldots,v_{i-1}} \max_{\substack{x \perp v_1,\ldots,v_{i-1} \\ \|x\|=1}} (Ax|x).$$

**Wniosek 20.1.** *Niech* $A \in M_{m\times n}(\mathbb{R})$ *oraz* $n \leqslant m$. *Jeśli*

$$\sigma_1 \geqslant \ldots \geqslant \sigma_n$$

*są wartościami singularnymi* $A$*, to*

$$\begin{aligned}
\sigma_i &= \max_{v_1,\ldots,v_{n-i}} \min_{\substack{x \perp v_1,\ldots,v_{n-i} \\ \|x\|=1}} \underbrace{\sqrt{(Ax|Ax)}}_{=\|Ax\|} \\
&= \min_{v_i,\ldots,v_{i-1}} \max_{\substack{x \perp v_1,\ldots,v_{i-1} \\ \|x\|=1}} \|Ax\| \\
&= \max_{\dim V = i} \min_{\substack{x \in V \\ \|x\|=1}} \|Ax\| \\
&= \min_{\dim V = n-i+1} \max_{\substack{x \in V \\ \|x\|=1}} \|Ax\|.
\end{aligned}$$

*W szczególności,*

$$\sigma_1 = \max_{\|x\|=1} \|Ax\| \quad \textit{oraz} \quad \sigma_n = \min_{\|x\|=1} \|Ax\|.$$

*Dowód.* Ponieważ $\sigma_i = \sqrt{\lambda_i}$, gdzie $\lambda_i$ są wartościami własnymi macierzy Grama $A^T A$ oraz $(A^T Ax|x) = (Ax|Ax)$, więc wystarczy zastosować twierdzenie 20.2 do macierzy $A^T A$. $\square$

**Twierdzenie 20.3** (Nierówność Weyla)**.** *Niech*

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$$

*będą wartościami własnymi macierzy symetrycznej* $A \in M_{n\times n}(\mathbb{R})$. *Przypuśćmy, że* $B = A + E$ *jest perturbacją macierzy* $A$ *przez macierz symetryczną* $E$ *o wartościach własnych* $\epsilon_1 \geqslant \ldots \geqslant \epsilon_n$. *Jeśli*

$$\beta_1 \geqslant \ldots \geqslant \beta_n$$

*są wartościami własnymi macierzy* $B$*, to*

$$\lambda_i + \epsilon_1 \geqslant \beta_i \geqslant \lambda_i + \epsilon_n, \quad i = 1, \ldots, n.$$

*Dowód.* Jeśli $Q$ jest taką macierzą ortogonalną, że

$$Q^T AQ = D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n),$$

to $\tilde{B} = Q^T BQ$ oraz $\tilde{E} = Q^T EQ$ mają te same wartości własne, co $B$ i $E$ oraz $\tilde{B} = D + \tilde{E}$. Jeśli $x \in F = \operatorname{span}\{e_1, \ldots, e_i\}$ ma normę 1, to $x = [x_1, \ldots, x_i, 0, \ldots, 0]^T$ oraz

$$(Dx|x) \geqslant \lambda_i,$$

więc, stosując twierdzenie Couranta–Fischera do macierzy $\tilde{B}$, otrzymujemy, że

$$\begin{aligned}
\beta_i &= \max_{\dim V = i} \min_{\substack{x\in V\\ \|x\|=1}} (\tilde{B}x|x) \\
&\geqslant \min_{\substack{x\in F\\ \|x\|=1}} (\tilde{B}x|x) \\
&= \min_{\substack{x\in F\\ \|x\|=1}} \Big((Dx|x) + (\tilde{E}x|x)\Big) \\
&\geqslant \min_{\substack{x\in F\\ \|x\|=1}} (Dx|x) + \min_{\substack{x\in F\\ \|x\|=1}} (\tilde{E}x|x) \\
&\geqslant \lambda_i + \min_{\substack{x\in \mathbb{R}^n\\ \|x\|=1}} (\tilde{E}x|x) \\
&= \lambda_i + \epsilon_n.
\end{aligned}$$

Analogicznie, dla $x \in T = \operatorname{span}\{e_i, \ldots, e_n\}$ o normie 1 mamy $(Dx|x) \leqslant \lambda_i$ oraz

$$\begin{aligned}
\beta_i &= \min_{\dim V = n-i+1} \max_{\substack{x\in V\\ \|x\|=1}} (\tilde{B}x|x) \\
&\leqslant \max_{\substack{x\in T\\ \|x\|=1}} (\tilde{B}x|x) \\
&= \max_{\substack{x\in T\\ \|x\|=1}} \Big((Dx|x) + (\tilde{E}x|x)\Big) \\
&\leqslant \max_{\substack{x\in T\\ \|x\|=1}} (Dx|x) + \min_{\substack{x\in F\\ \|x\|=1}} (\tilde{E}x|x) \\
&\leqslant \lambda_i + \max_{\substack{x\in \mathbb{R}^n\\ \|x\|=1}} (\tilde{E}x|x) \\
&= \lambda_i + \epsilon_1.
\end{aligned}$$

□

**Twierdzenie 20.4** (Cauchy)**.** *Załóżmy, że macierz symetryczna $A \in M_{n\times n}(A)$ ma wartości własne*

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n.$$

*Rozważmy symetryczną macierz blokową*

$$B = \left[\begin{array}{c|c} A & c \\ \hline c^T & \alpha \end{array}\right] \in M_{(n+1)\times(n+1)}(\mathbb{R}),$$

*dla pewnego $c \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$. Jeśli*

$$\beta_1 \geqslant \ldots \geqslant \beta_{n+1}$$

*są wartościami własnymi $B$, to*

$$\beta_1 \geqslant \lambda_1 \geqslant \beta_2 \geqslant \lambda_2 \geqslant \ldots \geqslant \beta_n \geqslant \lambda_n \geqslant \beta_{n+1}.$$

*Dowód.* Mamy pokazać, że

$$\beta_i \geqslant \lambda_i \geqslant \beta_{i+1}, \quad i = 1, \ldots, n.$$

Niech $Q$ będzie taką macierzą ortogonalną, że

$$Q^T A Q = D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

Ponieważ

$$P = \left[\begin{array}{c|c} Q & 0 \\ \hline 0 & 1 \end{array}\right]$$

jest również ortogonalna, więc $B$ ma te same wartości własne jak macierz

$$\tilde{B} = P^T B P = \left[\begin{array}{c|c} D & y \\ \hline y^T & \alpha \end{array}\right], \quad y = Q^T c.$$

Dla $x \in F = \operatorname{span}\{e_1, \ldots, e_i\} \subset \mathbb{R}^{n+1}$ o normie jeden mamy

$$(\tilde{B}x|x) = \sum_{j=1}^{i} \lambda_j x_j^2 \geqslant \lambda_i,$$

czyli

$$\beta_i = \max_{\dim V = i} \min_{\substack{x \in V \\ \|x\|=1}} (\tilde{B}x|x) \geqslant \lambda_i.$$

Z drugiej strony, dla $x \in T = \operatorname{span}\{e_{i-1}, \ldots, e_n\} \subset \mathbb{R}^{n+1}$ o normie jeden mamy

$$(\tilde{B}x|x) = \sum_{j=i-1}^{n} \lambda_j x_j^2 \leqslant \lambda_{i-1}.$$

Stąd,

$$\beta_i = \min_{\dim V = n-i+2} \max_{\substack{x \in V \\ \|x\|=1}} (\tilde{B}x|x) \leqslant \lambda_{i-1}.$$

□

U

Niech $A \in M_{n\times n}(\mathbb{R})$. Załóżmy, że wartości własne $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ macierzy $A$ spełaniają nierówności

$$|\lambda_1| \geqslant \ldots \geqslant |\lambda_n|.$$

Jeśli

$$\sigma_1 \geqslant \ldots \geqslant \sigma_n$$

są wartościami singularnymi macierzy $A$, to zachodzą *nierówności Weyla*

$$\sigma_1 \sigma_2 \ldots \sigma_k \geqslant |\lambda_1 \lambda_2 \ldots \lambda_k|, \quad k = 1, \ldots, n.$$

Można również udowodnić, że jeśli $\sigma_k(A^m)$ jest $k$-tą wartością singularną macierzy $A^m$, to

$$\lim_{m\to\infty} \sqrt[m]{\sigma_k(A^m)} = |\lambda_k|, \quad k = 1, \ldots, n.$$

**Lemat 20.1.** *Niech $A_r$ będzie macierzą otrzymaną z macierzy $A \in M_{m\times n}(\mathbb{R})$ przez wykreślenie $r$ kolumn (lub wierszy). Wtedy (uporządkowane malejąco) wartości singularne spełniają nierówności*

$$\sigma_k(A) \geqslant \sigma_k(A_r) \geqslant \sigma_{k+r}(A), \quad k = 1, \ldots, \min\{m, n\},$$

*gdzie dla $B \in M_{p\times q}(\mathbb{R})$ stosujemy konwencję: $\sigma_j(B) = 0$, gdy $j > \min\{p, q\}$.*

*Dowód.* Wystarczy rozważyć przypadek $r = 1$. Niech $A_1$ powstaje z $A$ przez wykreślenie $s$-tej kolumny. Dla wektora $x \in \mathbb{R}^n$ przez $x^s$ oznaczamy wektor otrzymany z $x$ przez usunięcie $s$-tej współrzędnej. Mamy

$$\begin{aligned}
\sigma_i(A) &= \min_{v_i, \ldots, v_{i-1} \in \mathbb{R}^n} \max_{\substack{x \perp v_1, \ldots, v_{i-1} \\ \|x\|=1}} \|Ax\| \\
&\geqslant \min_{v_1, \ldots, v_{i-1} \in \mathbb{R}^n} \max_{\substack{x \perp v_1, \ldots, v_{i-1}, e_s \\ \|x\|=1}} \|Ax\| \\
&= \min_{v_1, \ldots, v_{i-1} \in \mathbb{R}^{n-1}} \max_{\substack{x^s \perp v_1, \ldots, v_{i-1} \\ \|x^s\|=1}} \|Ax\| \\
&= \sigma_i(A_1).
\end{aligned}$$

Z drugiej strony

$$\begin{aligned}
\sigma_{i+1}(A) &= \max_{v_i,\dots,v_{n-i-1}\in\mathbb{R}^n} \min_{\substack{x\perp v_1,\dots,v_{n-i-1}\\ \|x\|=1}} \|Ax\| \\
&\leqslant \max_{v_1,\dots,v_{n-i-1}\in\mathbb{R}^n} \min_{\substack{x\perp v_1,\dots,v_{n-i-1},e_s\\ \|x\|=1}} \|Ax\| \\
&= \max_{v_1,\dots,v_{n-i-1}\in\mathbb{R}^{n-1}} \min_{\substack{x^s\perp v_1,\dots,v_{n-i-1}\\ \|x^s\|=1}} \|Ax\| \\
&= \sigma_i(A_1).
\end{aligned}$$

$\square$

**Lemat 20.2.** *Niech* $B = \frac{1}{2}(A + A^T)$ *będzie symetryzacją macierzy kwadratowej* $A \in M_{n\times n}(\mathbb{R})$. *Jeśli*

$$\sigma_1(A) \geqslant \ldots \geqslant \sigma_n(A),$$

*są wartościami singularnymi* $A$ *oraz*

$$\lambda_1(B) \geqslant \ldots \geqslant \lambda_n(B)$$

*są wartościami własnymi* $B$, *to*

$$\sigma_k(A) \geqslant \lambda_k(B), \quad k = 1, \ldots, n.$$

*Dowód.* Dla wektora jednostkowego $x \in \mathbb{R}^n$ mamy

$$\begin{aligned}
(Bx|x) &= \frac{1}{2}(Ax + A^T x|x) \\
&= (Ax|x) \\
&\leqslant \|Ax\|\|x\| \\
&= \|Ax\|.
\end{aligned}$$

Stąd

$$\begin{aligned}
\lambda_k(B) &= \min_{v_1,\dots,v_{k-1}} \max_{\substack{x\perp v_1,\dots,v_{k-1}\\ \|x\|=1}} (Bx|x) \\
&\leqslant \min_{v_1,\dots,v_{k-1}} \max_{\substack{x\perp v_1,\dots,v_{k-1}\\ \|x\|=1}} \|Ax\| \\
&= \sigma_k(A).
\end{aligned}$$

$\square$

U

Rezultaty tego rozdziału pozostają prawdziwe dla macierzy zespolonych, jeśli macierze symetryczne zastąpimy przez hermitowskie, a ortogonalne przez unitarne. Wartości singularne są wtedy pierwiastkami z wartości własnych macierzy $A^*A$.

Jeśli $A \in M_{n\times n}(\mathbb{C})$ jest kwadratowa, to

$$\big|\det(A)\big|^2 = \det(A^*A) = \sigma_1^2(A)\ldots\sigma_n^2(A),$$

więc

$$\big|\det(A)\big| = \sigma_1(A)\ldots\sigma_n(A).$$

**Lemat 20.3.** *Niech* $A \in M_{n\times n}(\mathbb{C})$ *i niech* $U \in \mathrm{U}(n)$ *będzie macierzą unitarną. Jeśli* $U_k \in M_{n\times k}(\mathbb{C})$ *jest macierzą utworzoną z pierwszych* $k$ *kolumn macierzy* $U$, *to*

*(i)* $\sigma_i(U_k^* A U_k) \leqslant \sigma_i(A)$ *dla* $i = 1, \ldots, k$,

*(ii)* $\big|\det U_k^* A U_k\big| \leqslant \sigma_1(A)\ldots\sigma_k(A)$.

*Dowód.* Ponieważ macierz $U_k^*AU_k$ powstaje z macierzy $U^*AU$ przez wykreślenie ostanich $n-k$ kolumn i wierszy, więc z lematu 20.1 i unitarnej niezmienniczości wartości singularnych mamy

$$\sigma_i(U_k^*AU_k) \leqslant \sigma_i(U^*AU) = \sigma_i(A).$$

Stąd

$$\big|\det U_k^*AU_k\big| = \sigma_1(U_k^*AU_k)\ldots\sigma_k(U_k^*AU_k) \leqslant \sigma_1(A)\ldots\sigma_k(A).$$

□

**Twierdzenie 20.5** (Nierówności Weyla)**.** *Niech $A \in M_{n\times n}(\mathbb{C})$. Jeśli wartości singularne i wartości własne są uporządkowane malejąco:*

$$\sigma_1(A) \geqslant \ldots \geqslant \sigma_n(A) \geqslant 0$$

*oraz*

$$|\lambda_1(A)| \geqslant \ldots \geqslant |\lambda_n(A)|,$$

*to*

$$|\lambda_1(A)\ldots\lambda_k(A)| \leqslant \sigma_1(A)\ldots\sigma_k(A), \quad k = 1,\ldots,n.$$

*Ponadto dla $k = n$ zachodzi równość.*

*Dowód.* Z twierdzenia Schura istnieje taka macierz unitarna $U \in \mathrm{U}(n)$, że macierz $U^*AU = \Delta$ jest górnie trójkątna z wartościami własnymi $A$ na przekątnej. Jeśli $U_k$ składa się z pierwszych $k$ kolumn macierzy $U$, to

$$\begin{aligned} U^*AU &= [U_k|\star]^*A[U_k|\star] \\ &= \left[\begin{array}{c|c} U_k^*AU_k & \star \\ \hline \star & \star \end{array}\right] \\ &= \Delta. \end{aligned}$$

Stąd, $U_k^*AU_k$ jest górnie trójkątna o wyrazach diagonalnych $\lambda_1(A),\ldots,\lambda_k(A)$. Z lematu 20.3 mamy

$$\begin{aligned} \big|\lambda_1(A)\ldots\lambda_k(A)\big| &= \big|\det U_k^*AU\big| \\ &\leqslant \sigma_1(A)\ldots\sigma_k(A). \end{aligned}$$

□

# Rozdział 21

# Twierdzenie Gerszgorina

| |
|---|
| SŁOWA KLUCZOWE ROZDZIAŁU |
| dyski Gerszgorina $\lozenge$ twierdzenie Gerszgorina |

Rozważmy macierz $A \in M_{n\times n}(\mathbb{C})$. Pokażemy warunek pozwalający na lokalizację wartości własnych $A$ na płaszczyźnie zespolonej $\mathbb{C}$.

**Definicja 21.1.** **Dyski Gerszgorina**
Dla $A \in M_{n\times n}(\mathbb{C})$ oraz $i = 1, \ldots, n$ definiujemy

$$r_i = \sum_{j\neq i} |a_{ij}| \quad \text{oraz} \quad D_i = \Big\{z \in \mathbb{C} : |z - a_{ii}| \leqslant r_i\Big\}.$$

Zbiór $D_i$ nazywamy *dyskiem Gerszgorina*

**Przykład 21.1.** Jeśli $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, to $D_i = \{\lambda_i\}$ dla $i = 1, \ldots, n$.

**Twierdzenie 21.1** (Twierdzenie Gerszgorina)**.** *Wartości własne macierzy $A \in M_{n\times n}(\mathbb{C})$ są zawarte w zbiorze*

$$D_1 \cup \ldots \cup D_n.$$

*Dowód.* Niech $\lambda \in \mathbb{C}$ będzie wartością własną macierzy $A$ dla wektora własnego $x = [x_1, \ldots, x_n]^T$. Niech $0 \neq x_i$ będzie współrzędną $x$ o największym module. Ponieważ $Ax = \lambda x$, więc

$$\sum_{j=1}^{n} a_{ij}x_j = \lambda x_j,$$

więc

$$(\lambda - a_{ii})x_i = \sum_{j\neq i} a_{ij}x_j.$$

Stąd

$$\begin{aligned} |\lambda - a_{ii}| &= \frac{1}{|x_i|}\Big|\sum_{j\neq i} a_{ij}x_j\Big| \\ &\leqslant \frac{1}{|x_i|}\sum_{j\neq i} |a_{ij}||x_j| \\ &\leqslant \sum_{j\neq i} |a_{ij}||x_j| = r_i. \end{aligned}$$

$\square$

**Przykład 21.2.** Wartość własna $\lambda \in \mathbb{C}$ macierzy

$$A = \begin{bmatrix} c & 1 & 0 & 0 & 0 \\ 1 & c & 1 & 0 & 0 \\ 0 & 1 & c & 0 & 0 \\ 0 & 0 & 1 & c & 1 \\ 0 & 0 & 0 & 1 & c \end{bmatrix}.$$

spełnia warunek $|c-\lambda| \leqslant 2$. Rzeczywiście, sprawdzamy łatwo, że

$$r_1 = r_5 = 1, \quad r_2 = r_3 = r_4 = 2.$$

Wszystkie dyski Gerszgorina mają środek w punkcie $c \in \mathbb{C}$, więc ich suma jest równa

$$\{z \in \mathbb{C} : |z-c| \leqslant 2\}.$$

**Przykład 21.3.** Rozważmy macierz

$$A = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$$

o wartościach własnych $\pm\sqrt{2}$. Dyski Gerszgorina $D_1$ i $D_2$ mają środki w punkcie 0 i promienie

$$r_1 = 1, \quad r_2.$$

W dysku $D_1$ nie ma żadnej wartości własnej. Obydwie zawierają się w $D_2$.

U Ponieważ $A$ i $A^T$ mają takie same wartości własne, to są one również zawarte w sumie dysków Gerszgorina

$$\tilde{D}_1 \cup \ldots \cup \tilde{D}_n$$

dla $A^T$. W konsekwencji, wartości własne są zawarte w przecięciu

$$(D_1 \cup \ldots \cup D_n) \cap (\tilde{D}_1 \cup \ldots \cup \tilde{D}_n).$$

**Przykład 21.4.** Rozważmy macierz

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0.5 & 6 & 0.5 \\ 2 & 0 & 8 \end{bmatrix}.$$

Dyski Gerszgorina mają środki w punktach

$$z_1 = [2,0]^T, \quad z_2 = [6,0]^T, \quad z_3 = [8,0]^T$$

i promienie

$$r_1 = 1, \quad r_2 = 1, \quad r_3 = 2.$$

Zauważmy, że 0 nie należy do ich sumy, więc nie jest wartością własną, czyli macierz $A$ jest nieosobliwa.

**Wniosek 21.1.** *Jeśli suma $k$ dysków Gerszgorina macierzy $A \in M_{n\times n}(\mathbb{C})$ jest rozłączna z sumą pozostałych $n-k$ dysków, to pierwsza z nich zawiera dokładnie $k$ wartości własnych (liczonych z krotnościami), a druga suma zawiera ich $n-k$. W szczególności, jeśli pewien dysk Gerszgorina jest rozłączny z sumą pozostałych dysków, to zawiera dokładnie jedną wartość własną macierzy $A$.*

*Dowód.* Niech $A = [a_{ij}]$ oraz $D = \operatorname{diag}(a_{11}, \ldots, a_{nn})$. Załóżmy, że $U_1$ jest sumą $k$ dysków, a $U_2$ sumą pozostałych $n-k$ dysków. Zbiory $U_1$, $U_2$ są zwarte i rozłączne, więc odległość $d = d(U_1, U_2) > 0$ jest dodatnia. Definiujemy macierze

$$A_t = (1-t)D + tA, \quad t \in [0,1].$$

Macierz $A_t$ ma te same wyrazy diagonalne co macierz $A$ dla każdego $t \in [0,1]$. Stąd dyski Gerszgorina macierzy $A_t$ mają te same środki co dyski macierzy $A$. Natomiast ich promienie są równe $t$-razy promień dysków dla $A$. Wynika stąd, że odpowiadające sumy dysków Gerszgorina $U_1(t)$, $U_2(t)$ dla $A_t$ są rozłączne oraz

$$d(t) = d(U_1(t), U_2(t)) \geqslant d, \quad t \in [0,1].$$

Zauważmy, że teza zachodzi dla $A_0 = D$. Niech $\lambda(0) \in U_1(0)$ będzie wartością własną $D$. Funkcja

$$t \longmapsto \lambda(t)$$

jest ciągła, więc również funkcja

$$d(t) := d(\lambda(t), U_2(t))$$

jest ciągła. Przypuśćmy, że $\lambda(1) \in U_2 = U_2(1)$. Wtedy $d(1) = 0$ oraz $d(0) = d > 0$. Z własności Darboux istnieje takie $t_0$, że

$$0 < d(t_0) < d.$$

Oznacza to, że

$$\lambda(t_0) \notin U_1(t_0) \cup U_2(t_0),$$

co jest sprzeczne z twierdzeniem 21.1. □

**Wniosek 21.2.** *Załóżmy, że dla pewnego dysku Gerszgorina $D_i = \{a_{ii}\}$. Wtedy $a_{ii}$ jest wartością własną $A$.*

*Dowód.* Z założenia wynika, że

$$r_i = \sum_{j \neq i} |a_{ij}| = 0,$$

czyli $a_{ij} = 0$ dla $j \neq i$. Oznacza to, że $i$-ty wiersz macierzy $A$ ma postać

$$a(i) = [0, \ldots, a_{ii}, \ldots, 0].$$

Wtedy $i$-ta kolumna macierzy $A^T$ ma postać $[0, \ldots, a_{ii}, \ldots, 0]$. Stąd, $A^T e_i = a_{ii} e_i$, czyli $a_{ii}$ jest wartością własną $A^T$, więc również $A$. □

# Rozdział 22

# Uogólniony problem własny

| Słowa kluczowe rozdziału | |
|---|---|
| uogólniony problem własny ♢<br>uogólniona wartość własna ♢<br>uogólniony wektor własny ♢<br>symultaniczna diagonalizacja | $$Av = \lambda \cdot Bv$$ |

Niech $A, B \in M_{n\times n}(\mathbb{C})$. Powiemy, że niezerowy wektor $v \in \mathbb{C}^n$ jest *uogólnionym wektorem własnym* pary $(A, B)$, jeśli

$$Av = \lambda \cdot Bv$$

dla pewnego $\lambda \in \mathbb{C}$. Skalar $\lambda$ nazywamy wtedy *uogólnioną wartością własną* dla $v$.

Podobnie jak w przypadku $B = I$, czyli klasycznego problemu własnego, $\lambda \in \mathbb{C}$ jest uogólnioną wartością własną, jeśli $A - \lambda B$ jest osobliwa, czyli $\lambda$ jest pierwiastkiem wielomianu charakterystycznego

$$\pi(x) = \det(A - xB).$$

(!) Jeśli $B \neq I$, to $\pi(x)$ może mieć 0 lub nieskończenie wiele wartości własnych. Dla przykładu rozważmy macierze

$$A = \begin{bmatrix} 1 & 0 \\ 0 & a \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & b \end{bmatrix}.$$

Wtedy

$$\det(A - xB) = (1 - x)(a - bx).$$

Mamy następujące przypadki:

(i) Jeśli $a \neq 0$, $b \neq 0$, to mamy dwie uogólnione wartości własne 1 i $a/b$.

(ii) Jeśli $a = 0$ i $b \neq 0$, to mamy dwie uogólnione wartości własne 1 i 0.

(iii) Jeśli $a \neq 0$ i $b = 0$, to mamy tylko jedną uogólnioną jednokrotną wartość własną 1.

(iv) Jeśli $a = b = 0$, to każda liczba zespolona $\lambda$ jest uogólnioną wartością własną.

**Twierdzenie 22.1.** *Niech $A, B \in M_{n\times n}(\mathbb{R})$ będą macierzami symetrycznymi i niech $B$ będzie dodatnio określona. Wtedy uogólniony problem własny*

$$Av = \lambda \cdot Bv$$

*ma $n$ rzeczywistych wartości własnych i odpowiadające im uogólnione wektory własne mogą być tak wybrane, że są ortogonalne względem iloczynu skalarnego*

$$(u|w)_B := (Bw|u).$$

*Ponadto jeśli $A$ jest dodatnio określona, to uogólnione wartości własne są dodatnie.*

*Dowód.* Ponieważ $B$ jest dodatnio określona, więc $B = LL^T$ dla pewnej macierzy nieosobliwej $L \in M_{n\times n}(\mathbb{R})$. Możemy więc problem własny

$$Av = \lambda \cdot Bv = \lambda \cdot LL^T v$$

zapisać w postaci

$$\underbrace{(L^{-1}A(L^T)^{-1})}_{C}(\underbrace{L^T v}_{z}) = \lambda \cdot \underbrace{L^T v}_{z}.$$

Jest to więc klasyczny problem własny

$$Cz = \lambda \cdot z.$$

Ponieważ $C = C^T$, więc mamy $n$ rzeczywistych wartości własnych, z odpowiadającymi takimi wektorami własnymi $z_1, \ldots, z_n$, że

$$(z_i|z_j) = \delta_{ij}.$$

Wtedy $v_i = (L^T)^{-1}(z_i)$ są uogólnionymi wektorami własnymi dla uogólnionego własnego oraz

$$\begin{aligned}(Bv_j|v_i) &= v_i^T Bv_j \\ &= z_i^T L^{-1}(LL^T)(L^T)^{-1}z_j \\ &= \delta_{ij}.\end{aligned}$$

Jeśli dodatkowo $A$ jest dodatnio określona, to $C$ jest również dodatnio określona, więc wartości własne są dodatnie. □

**Twierdzenie 22.2** (Symultaniczna diagonalizacja)**.** *Niech $A, B \in M_{n\times n}(\mathbb{R})$ będą macierzami symetrycznymi i niech $B$ będzie dodatnio określona. Wtedy istnieje taka macierz nieosobliwa $Q \in M_{n\times n}(\mathbb{R})$, że*

$$Q^T AQ = D, \quad Q^T BQ = I,$$

*gdzie $D$ jest diagonalna. Ponadto $D = Q^{-1}(B^{-1}A)Q$, więc wyrazy diagonalne $D$ są wartościami własnymi $B^{-1}A$.*

*Dowód.* Ponownie, niech $B = LL^T$ dla pewnej macierzy nieosobliwej i niech $C = L^{-1}A(L^T)^{-1}$. Ponieważ $C$ jest symetryczna, więc istnieje taka macierz ortogonalna $P \in O(n)$, że $P^T CP = D$, gdzie $D$ jest diagonalna. Niech $Q = (L^T)^{-1}P$. Podkreślmy, że $Q$ nie musi być ortogonalna. Mamy

$$\begin{aligned}Q^T AQ &= P^T L^{-1}A(L^T)^{-1}P \\ &= P^T CP \\ &= D\end{aligned}$$

oraz

$$\begin{aligned}Q^T BQ &= P^T L^{-1}(LL^T)(L^T)^{-1}P \\ &= P^T P \\ &= I.\end{aligned}$$

Ponadto,

$$\begin{aligned}QDQ^{-1} &= QQ^T AQQ^{-1} \\ &= (L^T)^{-1}PP^T L^{-1}A \\ &= (L^T)^{-1}L^{-1}A \\ &= B^{-1}A.\end{aligned}$$

□

# Rozdział 23

# Alternatywny dowód twierdzenia Cayleya–Hamiltona

| |
|---|
| SŁOWA KLUCZOWE ROZDZIAŁU<br>twierdzenie Cayleya–Hamiltona |

Rozważmy macierz $A \in M_{n\times n}(\mathbb{C})$ i wielomian

$$f(x) = a_0 + a_1 x + \ldots a_n x^n \in \mathbb{C}[x].$$

Definiujemy macierz

$$f(A) = a_0 \cdot I + a_1 \cdot A + \ldots + a_n \cdot A^n \in M_{n\times n}(\mathbb{C}).$$

**Wniosek 23.1.** *Jeśli* $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$*, to* $f(D) = \operatorname{diag}(f(\lambda_1), \ldots, f(\lambda_n))$.

**Twierdzenie 23.1** (Cayley–Hamilton)**.** *Niech* $A \in M_{n\times n}(\mathbb{C})$*. Jeśli* $p_A(x) = (-1)^n \det(xI - A)$ *jest wielomianem charakterystycznym macierzy* $A$*, to*

$$p_A(A) = 0,$$

*tzn.* $p(A)$ *jest macierzą zerową.*

(!) Moglibyście powiedzieć, że twierdzenie Cayleya–Hamiltona jest trywialne, bo przecież

$$p_A(A) = (-1)^n \det(AI - A) = (-1)^n \det 0 = 0.$$

Powyższe rozumowanie nie ma wiele wspólnego z dowodem twierdzenia Cayleya–Hamiltona, bo $p_A(A)$ jest macierzą, a nie liczbą.

(U) Niech $A \in M_{n\times n}(\mathbb{C})$ będzie dowolną macierzą. Ponieważ $M_{n\times n}(\mathbb{C})$ ma wymiar $n^2$, więc macierze

$$I, A, \ldots, A^{n^2}$$

są liniowo zależne. W szczególności, istnieje taki wielomian $f$ stopnia co najwyżej $n^2$, że $f(A) = 0$.

**Lemat 23.1.** *Twierdzenie Cayleya–Hamiltona zachodzi dla macierzy diagonalnych.*

*Dowód.* Niech $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Wtedy $\lambda_i$ są wartościami własnymi $D$, więc $p_D(\lambda_i) = 0$. Stąd

$$p_D(D) = \operatorname{diag}(p_D(\lambda_1), \ldots, p_D(\lambda_n)) = \operatorname{diag}(0, \ldots, 0) = 0.$$

□

**Lemat 23.2.** *Załóżmy, że $B$ jest nieosobliwa. Wtedy*

$$p_A = p_{B^{-1}AB}$$

*oraz dla dowolnego wielomianu $f$ mamy*

$$f(B^{-1}AB) = B^{-1}f(A)B.$$

*Dowód.* Zauważmy, że $(B^{-1}AB)^k = B^{-1}A^kB$. □

**Lemat 23.3.** *Twierdzenie Cayleya–Hamiltona zachodzi dla macierzy diagonalizowalnych.*

*Dowód.* Jeśli $D = B^{-1}AB$ jest diagonalna, to $p_A(D) = p_D(D) = 0$. Ale

$$0 = p_A(D) = p_A(B^{-1}AB) = B^{-1}p_A(A)B,$$

więc $p_A(A) = 0$. □

**Lemat 23.4.** *Twierdzenie Cayleya–Hamiltona zachodzi dla macierzy górnie trójkątnych.*

*Dowód.* Niech $A$ będzie macierzą górnie trójkątną. Wtedy $a_{11}, \dots, a_{nn}$ są wartościami własnymi $A$. Możemy dowolnie mało zmienić wyrazy diagonalne, aby były one różne. Otrzymamy macierz diagonalizowalną $\tilde{A}$. Wynika stąd, że istnieje taki ciąg górnie trójkątnych macierzy diagonalizowalnych $A_n$, że $A_n \to A$. Wtedy $p_{A_n} \to p_A$ oraz

$$p_A(A) = \lim_{n\to\infty} p_{A_n}(A_n) = 0.$$

□

**Lemat 23.5.** *Twierdzenie Cayleya–Hamiltona zachodzi dla dowolnych macierzy.*

*Dowód.* Dla dowolnej macierzy $A$ istnieje taka macierz unitarna $U$ (tzn. $U^{-1} = U^*$), że $UAU^*$ jest górnie trójkątna. □

**Przykład 23.1.** Pokażemy, że twierdzenie Cayleya–Hamiltona można użyć do oblicznia macierzy odwrotnej do macierzy nieosobliwej. Rozważmy przykładowo macierz

$$A = \begin{bmatrix} 5 & 7 & -5 \\ 0 & 4 & -1 \\ 2 & 8 & -3 \end{bmatrix}.$$

Bezpośredni rachunek pokazuje, że

$$p_A(x) = -x^3 + 6x^2 - 11x + 6.$$

Ponieważ wyraz wolny $p_A$ jest niezerowy, więc $A$ jest nieosobliwa. Z twierdzenia Cayleya–Hamiltona wiemy, że

$$-A^3 + 6A^2 - 11A + 6I = 0.$$

Stąd

$$-A^2 + 6A - 11I + 6A^{-1} = 0,$$

czyli

$$A^{-1} = \frac{1}{6}(A^2 - 6A + 11I) = \frac{1}{6}\begin{bmatrix} -4 & -19 & 13 \\ -2 & -5 & 5 \\ -8 & -26 & 20 \end{bmatrix}.$$

# Rozdział 24

# Równanie Sylvestera i równanie Lapunowa

| |
|---|
| Słowa kluczowe rozdziału<br>równanie Sylvestera ◊ symultaniczne sprowadzanie do postaci trójkątnej ◊ istnienie i jednoznaczność rozwiązań równania Sylvestera ◊ zasada Rotha ◊ równanie Lapunowa |

Niech $A, B, C \in M_{n\times n}(\mathbb{C})$. Naszym celem jest zbadanie zbioru rozwiązań równania macierzowego

$$AX - XB = C$$

względem $X \in M_{n\times n}(\mathbb{C})$. Jest ono nazywane *równaniem Sylvestera*. Równanie to pełni ważną rolę w teorii stabilności w sensie Lapunowa punktów stacjonarnych równań różniczkowych.

**Lemat 24.1.** *Niech $A, B \in M_{n\times n}(\mathbb{C})$ będą takie, że $AB = BA$. Wtedy, istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{C})$, że*

$$S^{-1}AS, \quad S^{-1}BS$$

*są górnie trójkątne.*

*Dowód.* Stosujemy indukcję względem $n$. Dla $n = 1$ teza zachodzi. Niech $n > 1$. Jeśli $A = \lambda I$ i $B = \mu I$, to teza zachodzi. Możemy więc założyć, że przykładowo $A \neq \lambda I$ dla każdego $\lambda \in \mathbb{C}$. Niech $\lambda \in \mathbb{C}$ będzie wartością własną $A$ i niech $U = \ker(A - \lambda I)$ będzie odpowiadającą podprzestrzenią własną. Wtedy $U$ jest niezmiennicza dla $B$, bo z $AB = BA$ wynika, że

$$B(A - \lambda I) = (A - \lambda I)B.$$

Niech $1 \leqslant m = \dim U < n$. Wybieramy bazę dla $U$ i rozszerzamy ją do bazy $\mathcal{B}$ dla $\mathbb{C}^n$. Niech $P$ będzie macierzą przejścia od bazy $\mathcal{B}$ do bazy standardowej. Wtedy

$$P^{-1}AP = \begin{bmatrix} B_1 & C_1 \\ 0 & D_1 \end{bmatrix}, \quad P^{-1}BP = \begin{bmatrix} B_2 & C_2 \\ 0 & D_2 \end{bmatrix}.$$

Ponieważ $P^{-1}AP$ komutuje z $P^{-1}BP$, więc również $B_1$ komutuje z $B_2$ oraz $D_1$ komutuje z $D_2$. Z założenia indukcyjnego mogą one być symultanicznie sprowadzone do postaci górnie trójkątnej przez macierze nieosobliwe $R \in M_{m\times m}(\mathbb{C})$ oraz $S \in M_{(n-m)\times(n-m)}(\mathbb{C})$. Wtedy

$$Q = \begin{bmatrix} R & 0 \\ 0 & S \end{bmatrix}$$

sprowadza $P^{-1}AP$ oraz $P^{-1}BP$ do postaci górnie trójkątnej. Wystarczy więc przyjąć, że $C = PQ$. □

**Twierdzenie 24.1** (Sylvester)**.** *Niech $A, B \in M_{n\times n}(\mathbb{C})$. Jeśli $A$ i $B$ nie mają wspólnych wartości własnych, to dla dowolnej macierzy $C \in M_{n\times n}(\mathbb{C})$ równanie*

$$AX - XB = C$$

*ma jednoznaczne rozwiązanie $X \in M_{n\times n}(\mathbb{C})$.*

*Dowód.* Wystarczy pokazać, że odwzorowanie

$$L : M_{n\times n}(\mathbb{C}) \ni X \longmapsto AX - XB \in M_{n\times n}(\mathbb{C})$$

jest izomorfizmem. Zauważmy, że

$$L = L_A - L_B, \quad \text{gdzie} \quad L_A(X) = AX, \quad L_B(X) = XB.$$

Ponadto, $L_A \circ L_B = L_B \circ L_A$, więc $L_A$ i $L_B$ można symultanicznie sprowadzić do postaci górnie trójkątnej. Wynika stąd, że wartości własne $L$ są różnicami wartości własnych $L_A$ i $L_B$. Ponieważ wartości własne $L_A$, to wartości własne $A$ i wartości własne $L_B$, są wartościami własnymi $B$, więc wartości własne $L$ są niezerowe. W szczególności, $L$ jest izomorfizmem. □

**Wniosek 24.1** (Zasada Rotha)**.** *Niech $A, B, C \in M_{n\times n}(\mathbb{C})$. Jeśli równanie $AX - XB = C$ ma rozwiązanie, to macierze*

$$\left[\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right], \quad \left[\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array}\right]$$

*są podobne. Implikacja przeciwna jest również prawdziwa.*

*Dowód.* Wystarczy zauważyć, że

$$\left[\begin{array}{c|c} I & X \\ \hline 0 & I \end{array}\right] \left[\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right] \left[\begin{array}{c|c} I & -X \\ \hline 0 & I \end{array}\right] = \left[\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array}\right].$$

□

Szczególnym przypadkiem równania Sylvestera jest następujące *równanie Lapunowa*

$$A^T C + CA = -I, \quad A \in M_{n\times n}(\mathbb{R}). \tag{24.1}$$

Wiemy, że jeśli $A$ i $-A^T$ mają różne wartości własne, to ma ono jednoznaczne rozwiązanie $C = C_1 + iC_2 \in M_{n\times n}(\mathbb{C})$, gdzie $C_i \in M_{n\times n}(\mathbb{R})$. Wtedy $C_1$ jest również rozwiązaniem równania (24.1), więc $C = C_1$ jest macierzą rzeczywistą. Jest tak w szczególności, gdy wszystkie wartości własne macierzy $A$ mają ujemne części rzeczywiste. Jeśli $C$ jest rozwiązaniem równania (24.1), to

$$\begin{aligned} -I &= (A^T C + CA)^T \\ &= C^T A + A^T C^T, \end{aligned}$$

czyli $C^T$ jest również rozwiązaniem równania (24.1). Z jednoznaczności rozwiązań wynika, że $C = C^T$ jest macierzą symetryczną. Ponadto można udowodnić, że w tym przypadku jedyne rozwiązanie $C$ jest macierzą dodatnio określoną.

Przy wyznaczaniu rozwiązania $C$ równania Lapunowa często stosuje się następującą technikę. Przypuśćmy, że $C$ jest szukanym rozwiązaniem. Definiujemy macierz

$$S := CA - A^T C. \tag{24.2}$$

Jest ona antysymetryczna, bo

$$\begin{aligned} S^T &= (CA - A^T C)^T \\ &= A^T C^T - C^T A \\ &= A^T C - CA \\ &= -S. \end{aligned}$$

Z równości (24.1) i (24.2) otrzymujemy, że $2CA = S - I$, czyli

$$C = \frac{1}{2}(S - I)A^{-1}.$$

Wystarczy więc znaleźć macierz antysymetryczną $S$. Mamy

$$\begin{aligned} A^T S + SA &= A^T CA - A^T A^T C + CAA - A^T CA \\ &= -A^T A^T C - (I + A^T C)A \\ &= -A^T (A^T C + CA) - A \\ &= A^T - A. \end{aligned}$$

Równanie

$$A^T S + SA = A^T - A, \tag{24.3}$$

jest prostsze do rozwiązania od równania (24.1). Przykładowo, w wymiarze 2 macierze antysymetryczna $S$ ma postać

$$S = \begin{bmatrix} 0 & s \\ -s & 0 \end{bmatrix},$$

więc równanie (24.3) jest równaniem z jedną niewiadomą $s \in \mathbb{R}$. Dla

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

przyjmuje ono postać

$$\begin{bmatrix} 0 & s(a+d) \\ -s(a+d) & 0 \end{bmatrix} = \begin{bmatrix} 0 & c-b \\ b-c & 0 \end{bmatrix}.$$

Z założenia o wartościach własnych macierzy $A$ wynika, że $a + d = \operatorname{tr} A < 0$, więc

$$s = \frac{c-b}{a+b}.$$

W efekcie

$$\begin{aligned} C &= \frac{1}{2\det A} \begin{bmatrix} -1 & \frac{c-b}{a+d} \\ -\frac{c-b}{a+d} & -1 \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \\ &= \frac{1}{2\det A} \begin{bmatrix} -d - \frac{c(c-b)}{a+d} & b + \frac{a(c-b)}{a+d} \\ c - \frac{d(c-b)}{a+d} & -a + \frac{b(c-b)}{a+d} \end{bmatrix}. \end{aligned}$$

Macierz $C$ jest symetryczna, bo $b + \frac{a(c-b)}{a+d} = c - \frac{d(c-b)}{a+d}$. Ponieważ $\det A > 0$ i $\operatorname{tr} A < 0$, więc

$$\operatorname{tr} C = \frac{1}{2\det A}\left(-\operatorname{tr} A - \frac{(c-b)^2}{\operatorname{tr} A}\right) > 0,$$

$$\det C = \frac{1}{4\det A}\left(1 + \frac{(c-b)^2}{(a+d)^2}\right) > 0.$$

Stąd $C$ ma dodatnie wartości własne, czyli jest dodatnio określona.

U Jeśli $x$ jest rozwiązaniem liniowego równania różniczkowego $\dot{x} = Ax$, to

$$\begin{aligned} \frac{d}{dt}(Cx|x) &= (C\dot{x}|x) + (Cx|\dot{x}) \\ &= (CAx|x) + (Cx|Ax) \\ &= (CAx|x) + (A^T Cx|x) \\ &= ((A^T C + CA)x|x) \\ &= (-x|x) \\ &= -\|x\|^2. \end{aligned}$$

# Rozdział 25

# Liniowe układy dynamiczne

| |
|---|
| SŁOWA KLUCZOWE ROZDZIAŁU |
| dyskretne liniowe układy dynamiczne ◊ punkt stały i okresowy ◊ orbita ◊ punkt siodłowy $\qquad A^n v$ |

Rozważmy macierz

$$A = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix} = 0.5I.$$

Zauważmy, że

$$A^n = A = \begin{bmatrix} (0.5)^n & 0 \\ 0 & (0.5)^n \end{bmatrix} = (0.5)^n I.$$

Wynika stąd, że dla dowolnego wektora $v \in \mathbb{R}^2$ mamy

$$\lim_{n\to+\infty} A^n v = \lim_{n\to+\infty} (0.5)^n v = 0.$$

Rozważmy macierz $A \in M_{k\times k}(\mathbb{R})$ i wektor $v \in \mathbb{R}^k$. Naturalne jest pytanie: co możemy powiedzieć o zachowaniu ciągu

$$A^n v.$$

Czasami odpowiedź jest bardzo prosta. Przykładowo, jeśli $v = 0$, to oczywiście $A^n 0 = 0$ dla każdego $n \geqslant 1$. Oznacza to, że punkt 0 nie porusza się w tym *liniowym układzie dynamicznym.* Takie punkty nazywamy *punktami stałymi.*

**Definicja 25.1.** Punkt stały

Wektor $v \in \mathbb{R}^n$ jest *punktem stałym macierzy* $A \in M_{k\times k}(\mathbb{R})$, gdy $Av = v$.

**Wniosek 25.1.** *Niezerowy wektor* $0 \neq v$ *jest punktem stałym* $A$ *wtedy i tylko wtedy, gdy* $v$ *jest wektorem własnym* $A$ *dla wartości własnej* 1.

Wynika stąd, że $A$ ma dokładnie jeden punkt stały wtedy i tylko wtedy, gdy macierz $A - I$ jest nieosobliwa.

Zauważmy, że jeśli $Av = v$ to dla każdego $n \geqslant 1$ mamy $A^n v = v$, czyli punkt $v$ nie porusza się, gdy zmieniamy $n$.

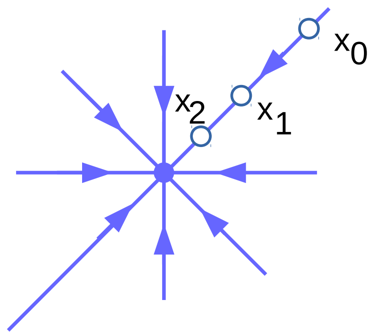
Bardziej ogólnie, jeśli $\lambda \in \mathbb{R}$ jest wartością własną, to dla odpowiadającego jej wektora własnego $v \in \mathbb{R}^k$ mamy

$$A^n v = \lambda^n v.$$

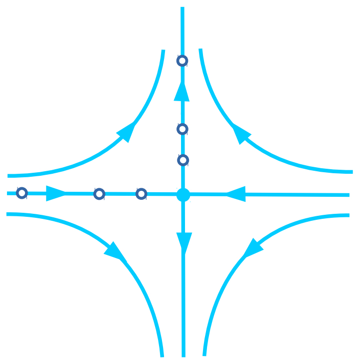
Wynika stąd, że jeśli $w \in \operatorname{span}\{v\}$, to $Aw \in \operatorname{span}\{v\}$. Mówimy wtedy, że prosta $\operatorname{span}\{v\}$ jest *dodatnio niezmiennicza* dla $A$.

**Definicja 25.2.** Podzbiór niezmienniczy

Podzbiór $S \subset \mathbb{R}^n$ jest *zbiorem dodatnio niezmienniczym dla macierzy* $A \in M_{k\times k}(\mathbb{R})$, gdy $Aw \in S$ dla $w \in S$.




Rysunek 25.1: Dynamika zadana macierzą diagonalną $A = 0.5I$. Ponieważ $A^n v = (0.5)^n v$, więc $A^n v$ zbiega do wektora zerowego 0. Dla wektora $x_0 = [1, 1]^T$ mamy: $x_1 = Ax_0 = [0.5, 0.5]^T$, $x_2 = Ax_1 = A^2 x_0 = [(0.5)^2, (0.5)^2]^T$. Ogólnie $x_n = A^n x_0 = [(0.5)^n, (0.5)^n]^T$.



Rysunek 25.2: Dynamika zadana taką macierzą diagonalną, że $Ae_1 = 0.5e_1$ i $Ae_2 = 2e_2$. Układ ma dwie proste niezmiennicze odpowiadające wektorom własnym dla wartości własnych $\lambda_1 = 0.5$ i $\lambda_2 = 2$. Dla każdego wektora $v$ na osi $x$ mamy $A^n v \to 0$, a dla wektora $w$ na osi $y$ zachodzi $\|A^n w\| \to \infty$.

**Definicja 25.3.** Orbita

Dla wektora $v_0 = v$ ciąg $v_n = A^n v_{n-1}$, czyli

$$v,\ Av,\ A^2 v,\ A^3 v,\ \dots$$

nazywamy *dodatnią orbitą* wektora $v$. Jeśli $A$ jest nieosobliwa, to możemy również rozważać *orbitę ujemną*

$$\dots\ A^{-3} v,\ A^{-2} v,\ A^{-1} v,\ v$$

oraz *orbitę*

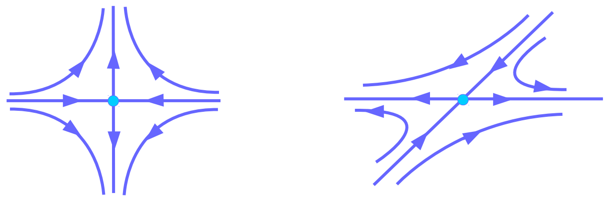
$$\dots\ A^{-3} v,\ A^{-2} v,\ A^{-1} v,\ v\ Av,\ A^2 v,\ A^3 v,\ \dots$$

**Przykład 25.1.** Niech

$$A = \begin{bmatrix} 0.5 & 0 \\ 0 & 2 \end{bmatrix}.$$

Wektorami własnymi macierzy $A$ są $e_1$ i $e_2$. Zauważmy, że

$$A^n e_1 = (0.5)^n e_1, \quad A^n e_2 = 2^n e_n.$$



Rysunek 25.3: Liniowa zmiana zmiennych w układzie dynamicznym.

Proste generowane przez $e_1$ i $e_2$ są niezmiennicze dla tego układu. Dla niezerowego wektora $[x, 0]^T$ na osi $x$ mamy

$$\lim_{n\to+\infty} A^n[x,0]^T = \lim_{n\to+\infty}[(0.5)^n x, 0]^T = 0,$$

$$\lim_{n\to-\infty} \|A^n[x,0]^T\| = \lim_{n\to-\infty}\|[(0.5)^n x, 0]^T\| = +\infty.$$

Natomiast dla niezerowego wektora $[0, y]^T$ na osi $y$ mamy

$$\lim_{n\to-\infty} A^n[0,y]^T = \lim_{n\to-\infty}[0, 2^n y]^T = 0,$$

$$\lim_{n\to+\infty} \|A^n[0,y]^T\| = \lim_{n\to+\infty}\|[0, 2^n y]^T\| = +\infty.$$

Rozważmy teraz taki wektor $[x, y]^T$, że $x \neq 0$ i $y \neq 0$. Wtedy

$$A^n[x,y]^T = [(0.5)^n x, 2^n y]^T.$$

Zauważmy, że wtedy

$$(0.5)^n x \cdot 2^n y = xy = \text{const},$$

czyli orbita leży na hiperboli. Osie układu współrzędnych są jedynymi prostymi niezmienniczymi dla układu. Oś $x$ jest nazywana *podprzestrzenią stabilną*, bo wszyskie punkty na niej zbiegają do początku układu 0 pod wpływem działania macierzy $A$. Punkt 0 nazywamy w tym przypadku *punktem siodłowym*.

Skupmy się na przypadku macierzy $A \in M_{2\times 2}(\mathbb{R})$. Zauważmy, że wystarczy przeanalizować macierze $A$ w postaci Jordana, bo dla macierzy nieosobliwej $S \in M_{2\times 2}(\mathbb{R})$ mamy

$$(S^{-1}AS)^n = S^{-1}A^n S.$$

**Przykład 25.2.** Dla układu generowanego przez macierz

$$A = \begin{bmatrix} 0.5 & 1 \\ 0 & 0.5 \end{bmatrix}$$

oś $x$ odpowiadająca wektorowi własnemu $e_1$ jest jedyną podprzestrzenią niezmienniczą. Jest ona stabilna (rys. **??** (a)). Macierz $A$ możemy przedstawić jako sumę macierzy diagonalnej $D$ i nilpotentnej $N$

$$A = \begin{bmatrix} 0.5 & 1 \\ 0 & 0.5 \end{bmatrix} = \underbrace{\begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}}_{D} + \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}_{N}.$$

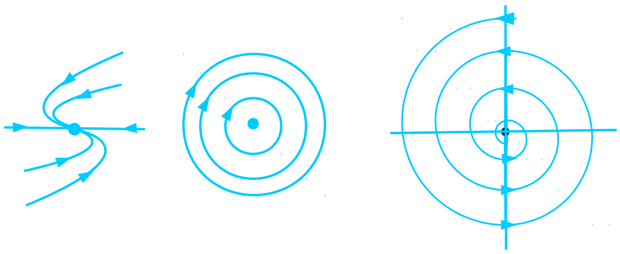
Ponieważ $DN = ND$ oraz $N^2 = 0$, więc

$$A^n = (D+N)^n = \sum_{k=0}^{n} \binom{n}{k} D^k N^{n-k} = D^n + nD^{n-1}N.$$

Zauważmy, że dla dowolonego wektora $v \in \mathbb{R}^2$ mamy

$$\lim_{n\to+\infty} D^n v = 0, \qquad \lim_{n\to+\infty} nD^{n-1}v = 0,$$

więc $A^n v \to 0$.




Rysunek 25.4: Układy dynamiczne na płaszczyźnie.

**Przykład 25.3.** Rozważmy teraz macierz $A \in M_{2\times 2}(\mathbb{R})$ z zespolonymi wartościami własnymi

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}, \quad b \neq 0.$$

Jeśli $a^2 + b^2 = 1$, to $A$ jest obrotem i ma postać

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

dla pewnego $\theta \in [0, 2\pi]$. Dynamika generowana przez macierz $A$ jest pokazana na rys. **??** (b). Okręgi o środku w punkcie 0 są zbiorami niezmienniczymi, bo $A$ jest izometrią.

Jeśli $a^2 + b^2 \neq 1$, to $A$ jest złożeniem jednokładności z obrotem, bo

$$A = \begin{bmatrix} \sqrt{a^2+b^2} & 0 \\ 0 & \sqrt{a^2+b^2} \end{bmatrix} \begin{bmatrix} \frac{a}{\sqrt{a^2+b^2}} & \frac{b}{\sqrt{a^2+b^2}} \\ \frac{-b}{\sqrt{a^2+b^2}} & \frac{a}{\sqrt{a^2+b^2}} \end{bmatrix}.$$

Dynamika układu dla $a^2 + b^2 > 1$ jest pokazana na rys. **??** (c).

**Wniosek 25.2.** *Niech $A \in M_{2\times 2}(\mathbb{R})$. Jeśli $|\lambda_1|, |\lambda_2| < 1$ dla wartości własnych $A$, to*

$$\lim_{n\to\infty} A^n v = 0, \quad v \in \mathbb{R}^2.$$

Wniosek 25.2 uogólnia się na wyższe wymiary.

**Twierdzenie 25.1.** *Jeśli $A \in M_{n\times n}(\mathbb{C})$ oraz $|\lambda| < 1$ dla każdej wartości własnej macierzy $A$, to*

$$\lim_{n\to\infty} A^n v = 0, \quad v \in \mathbb{C}^n$$

**Lemat 25.1** (Układ dynamiczny-potęgujemy macierz)**.** *Załóżmy, że*

$$S = [s_1, \ldots, s_n] \in M_{n\times n}(\mathbb{R})$$

*jest nieosobliwa oraz $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \in M_{n\times n}(\mathbb{R})$. Rozważmy macierz $A = SDS^{-1}$. Wtedy*

*(i) $As_i = \lambda_i s_i$ dla $i = 1, \ldots, n$;*

*(ii) $A^k = SD^kS^{-1}$ dla $k \in \mathbb{N}$;*

*(iii) Dla $v = x_1 \cdot s_1 + \ldots + x_n \cdot s_n$ i $k \in \mathbb{N}$ mamy*

$$A^k v = x_1\lambda_1^k \cdot s_1 + \ldots + x_n\lambda_n^k \cdot s_n.$$

*W szczególności, jeśli $|\lambda_i| < 1$ dla każdego $i = 1, \ldots, n$, to*

$$\lim_{k\to\infty} A^k v = 0.$$

*Dowód.* Dla dowodu punktu (i) zauważmy, że $s_i = Se_i$, więc

$$\begin{aligned} As_i &= SDS^{-1}s_i \\ &= SDS^{-1}Se_i \\ &= SDe_i \\ &= S(\lambda_i e_i) \\ &= \lambda_i S(e_i) \\ &= \lambda_i s_i. \end{aligned}$$

Punkt (ii) wynika z faktu, że

$$\begin{aligned} A^k &= (SDS^{-1})^k \\ &= \underbrace{(SDS^{-1})(SDS^{-1})\dots(SDS^{-1})}_{k} \\ &= SD^kS^{-1}. \end{aligned}$$

Punkt (iii) jest konsekwencją punktu (i), bo $A^k s_i = \lambda_i^k s_i$. $\square$

# Rozdział 26

# Przestrzenie unormowane i unitarne

| |
|---|
| Słowa kluczowe rozdziału |
| norma ◊ metryka ◊ iloczyn skalarny ◊ normy równoważne $\qquad \|v\| = \sqrt{(v\|v)}$ |

Skupimy się na przestrzeniach wektorowych $V$ nad ciałem $\mathbb{R}$.

**Definicja 26.1.** Norma
Normą w przestrzeni wektorowej $V$ nazywamy dowolne odwzorowanie

$$\|\cdot\| : V \ni v \longmapsto \|v\| \in \mathbb{R}$$

spełniające dla dowolnych $v, w \in V$ oraz $c \in \mathbb{R}$ warunki

(1) $\|v\| \geqslant 0$, oraz $\|v\| = 0$ wtedy i tylko wtedy, gdy $v = 0$,

(2) $\|c \cdot v\| = |c|\|v\|$,

(3) $\|v + w\| \leqslant \|v\| + \|w\|$.

Przestrzeń wektorową z ustaloną normą nazywamy *przestrzenią unormowaną*.

**Przykład 26.1.** Oto kilka popularnych norm w przestrzeni $\mathbb{R}^n$:

- norma euklidesowa
$$\|x\|_2 = \sqrt{x_1^2 + \ldots + x_n^2},$$
- $\|x\|_1 = |x_1| + \ldots + |x_n|$,
- $\|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$,
- dla liczby rzeczywistej $p \geqslant 1$ mamy normę
$$\|x\|_p = \sqrt[p]{|x_1|^p + \ldots + |x_n|^p}.$$

Stąd $\|\cdot\|_E = \|\cdot\|_2$ oraz $\|\cdot\|_s = \|\cdot\|_1$.

Nierówność trójkąta dla normy $\|\cdot\|_p$, czyli

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{1/p} \leqslant \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p}$$

jest nazywana *nierównością Minkowskiego*. Jest ona związana z inną znaną nierównością. Dla $p > 1$ rozważamy takie $q > 1$, że

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Wtedy zachodzi *nierówność Höldera*

$$\sum_{i=1}^{n} |x_i y_i| \leqslant \|x\|_p \|y\|_q.$$

Dla $p = 2$ jest to nierówność Cauchy'ego-Schwarza.

**Przykład 26.2.** Rozważmy przestrzeń wektorową $C([0,1],\mathbb{R})$ wszystkich funkcji ciągłych $f : [0,1] \longrightarrow \mathbb{R}$. Możemy w niej rozważyć normę

$$\|f\|_M = \max_{x\in[0,1]} |f(x)|.$$

Inne normy w tej przestrzeni są zadane wzorami

$$\|f\|_2 = \sqrt{\int_0^1 (f(x))^2 dx}, \quad \|f\|_1 = \int_0^1 |f(x)| dx.$$

**Przykład 26.3.** Dla niepustego zbioru $X$ definiujemy przestrzeń wektorową $F(X,\mathbb{R})$ złożoną z wszystkich funkcji $f : X \longrightarrow \mathbb{R}$, które są ograniczone, tzn. istnieje takie $M \geqslant 0$, że $|f(x)| \leqslant M$ dla wszyskich $x \in X$. Możemy w niej zdefiniować normę wzorem

$$\|f\|_\infty := \sup_{x\in X} |f(x)|.$$

Każda norma $\|\cdot\|$ w przestrzeni wektorowej $V$ zadaje w niej *metrykę* wzorem

$$d(v,w) := \|v - w\|, \quad v, w \in V.$$

Jest ona funkcją

$$d : V \times V \longrightarrow \mathbb{R}$$

o własnościach

(i) $d(v,w) \geqslant 0$, oraz $d(v,w) = 0$ wtedy i tylko wtedy, gdy $v = w$,

(ii) $d(v,w) = d(w,v)$,

(iii) $d(v,w) \leqslant d(v,u) + d(u,w)$,

dla dowolnych $v, w, u \in V$.

**Definicja 26.2.** **Iloczyn skalarny**

Iloczynem skalarnym w przestrzeni $V$ nad ciałem $\mathbb{R}$ nazywamy dowolne odwzorowanie dwuliniowe

$$(\cdot|\cdot) : V \times V \longrightarrow \mathbb{R},$$

spełniające dla $v, w \in V$ warunki

(a) $(v|v) \geqslant 0$, oraz $(v|v) = 0$ tylko, gdy $v = 0$,

(b) $(v|w) = (w|v)$.

Przestrzeń $V$ z ustalonym iloczynem skalarnym nazywamy *przestrzenią unitarną*.

U

Iloczyn skalarny pozwala zdefiniować ortogonalność wektorów i inne pojęcia rozważane wcześniej w kontekście euklidesowego iloczynu skalarnego w $\mathbb{R}^n$. W każdej przestrzeni unitarnej zachodzi nierówność Cauchy'ego-Schwarza

$$|(u|v)| \leqslant \sqrt{(u|u)}\sqrt{(v|v)}, \quad u, v \in V.$$

Iloczyn skalarny zadaje normę wzorem

$$\|v\| = \sqrt{(v|v)}.$$

Taka norma spełnia tożsamość równoległoboku

$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2), \quad u, v \in V. \tag{26.1}$$

Nie każda norma w przestrzeni wektorowej spełnia tożsamość (26.1). Możecie się zastanowić, które spośród norm $\|\cdot\|_p$ w $\mathbb{R}^n$ dla $p \geqslant 1$ ją spełniają.

Okazuje się, że norma $\|\cdot\|$ spełnia tożsamość (26.1) wtedy i tylko wtedy, gdy pochodzi ona od pewnego iloczynu skalarnego, tzn. istnieje taki iloczyn skalarny $(\cdot|\cdot)$, że

$$\|v\| = \sqrt{(v|v)}.$$

Jeśli norma spełnia tożsamość (26.1), to można sprawdzić, że wzór polaryzacyjny

$$(u|v) := \frac{1}{4}(\|u+v\|^2 - \|u-v\|^2)$$

definiuje iloczyn skalarny.

**Przykład 26.4.** Każdą przestrzeń unitarną możemy traktować jako przestrzeń unormowaną z normą zadaną przez iloczyn skalarny wzorem

$$\|v\| := \sqrt{(v|v)}, \quad v \in V.$$

**Przykład 26.5.** Wzór

$$(f|g) = \int_0^1 f(x)g(x)dx$$

zadaje iloczyn skalarny w $C([0,1], \mathbb{R})$.

**Definicja 26.3.** **Normy równoważne**

Normy $\|\cdot\|$ oraz $\|\cdot\|_*$ na przestrzeni wektorowej $V$ są *równoważne*, jeśli istnieją takie stałe dodatnie $C_1, C_2 > 0$, że

$$C_1\|v\| \leqslant \|v\|_* \leqslant C_2\|v\|, \quad v \in V.$$

**Twierdzenie 26.1** (równoważność norm w skończonym wymiarze)**.** *Wszystkie normy w $\mathbb{R}^n$ są równoważne.*

*Dowód.* Pokażemy, że dowolna norma $\|\cdot\|$ w $\mathbb{R}^n$ jest równoważna z normą

$$\|v\|_1 = \max_{i=1,\ldots,n} |v_i|.$$

Niech $e_1, \ldots, e_n$ będzie bazą standardową $\mathbb{R}^n$. Wtedy

$$\begin{aligned}
\|x-y\| &= \|(x_1-y_1)e_1 + \ldots + (x_n-y_n)e_n\| \\
&\leqslant |x_1-y_1|\|e_1\| + \ldots + |x_n-y_n|\|e_n\| \\
&\leqslant \max_i\{\|e_i\|\}\left(|x_1-y_1| + \ldots + |x_n-y_n|\right),
\end{aligned}$$

czyli dla $C_2 = \max_i\{\|e_i\|\}$ mamy

$$\|x-y\| \leqslant C_2\|x-y\|_1.$$

Z nierówności tej wynika, że funkcja $\|\cdot\| : \mathbb{R}^n \longrightarrow \mathbb{R}$ jest ciągła. Z twierdzenia Weierstrassa osiąga ona minimum $C_1$ na zbiorze zwartym $S = \{x \in \mathbb{R}^n : \|x\|_1 = 1\}$. Stąd

$$\|x\| \geqslant C_1, \quad \|x\|_1 = 1.$$

Dla $x \neq 0$ mamy

$$\|x/\|x\|_1\| \geqslant C_1,$$

czyli $\|x\| \geqslant C_1\|x\|_1$.

Pozostaje pokazać, że $C_1 > 0$. Oczywiście $C_1 \geqslant 0$. Przypuśćmy, że $C_1 = 0$. Istnieje więc taki $x$, że $\|x\|_1 = 1$ oraz $\|x\| = 0$. Ale wtedy $x = 0$, co prowadzi do sprzeczności. □

Niech $A \in M_{2\times 2}(\mathbb{R})$. Istnieje taka baza $v_1, v_2 \in \mathbb{R}^2$, że zachodzi jeden z warunków

(1) dla wartości własnych $\lambda_1, \lambda_2 \in \mathbb{R}$ mamy

$$A(v_1) = \lambda_1 v_1, \quad A(v_2) = \lambda_2 v_2,$$

(2) dla wartości własnej $\lambda = \lambda_{1,2} \in \mathbb{R}$ mamy

$$A(v_1) = \lambda v_1, \quad A(v_2) = v_1 + \lambda v_2,$$

(3) dla zespolonych wartości własnych $\lambda_{1,2} = a \pm ib$ mamy

$$A(v_1) = av_1 - bv_2, \quad A(v_2) = bv_1 + av_2.$$

Poniższe twierdzenie pełni ważną rolę w jakościowej teorii równań różniczkowych. Jest ono prawdziwe w dowolnym wymiarze, ale przykładowo udowodnimy je tylko w wymiarze 2.

**Twierdzenie 26.2.** *Niech $\lambda_1, \lambda_2 \in \mathbb{C}$ będą wartościami własnymi macierzy $A \in M_{2\times 2}(\mathbb{R})$ (liczonymi z krotnościami), spełniającymi warunek*

$$m < \operatorname{Re}\lambda_1, \operatorname{Re}\lambda_2 < M$$

*dla pewnych stałych $m, N \in \mathbb{R}$. Istnieje taki iloczyn skalarny $(\,\cdot\,|\,\cdot\,)$ w $\mathbb{R}^2$ i odpowiadająca mu norma $\|z\| := \sqrt{(z|z)}$, że*

$$m\|z\| \leqslant (Az|z) \leqslant M\|z\|, \quad z \in \mathbb{R}^2.$$

*Dowód.* Jeśli $A$ spełnia warunek (1) lub (3), to definiujemy iloczyn skalarny na bazie $v_1$, $v_2$ przez

$$(v_1|v_2) = \delta_{ij},$$

czyli jest to baza ortonormalna względem tego iloczynu skalarnego. Niech $z = x_1v_1 + x_2v_2$. Wtedy $\|z\| = \sqrt{x_1^2 + x_2^2}$. W przypadku (1) mamy

$$\begin{aligned}(Az|z) &= (x_1\lambda_1 v_1 + x_2\lambda v_2 | x_1v_1 + x_2v_2)\\ &= x_1^2\lambda_1 + x_2^2\lambda_2,\end{aligned}$$

więc teza zachodzi. W przypadku (3) mamy

$$\begin{aligned}(Az|z) &= (x_1(av_1 - bv_2) + x_2(bv_1 + av_2)|x_1v_1 + x_2v_2)\\ &= x_1^2 a + x_2^2 a\\ &= a\|z\|^2,\end{aligned}$$

czyli teza zachodzi.

Jeśli powtórzymy to rozumowanie w przypadku (2), czyli przyjmiemy, że $v_1, v_2$ jest bazą ortonormalną względem iloczynu skalarnego $(\cdot|\cdot)$, to otrzymamy, że

$$\begin{aligned}(Az|z) &= (x_1\lambda v_1 + x_2(v_1 + \lambda v_2)|x_1v_1 + x_2v_2)\\ &= x_1^2\lambda + x_1x_2 + x_2^2\lambda\\ &= \lambda\|z\|^2 + x_1x_2.\end{aligned}$$

Teza nie musi zachodzić. Musimy więc zmodyfikować nasze podejście. Ustalmy $\epsilon > 0$ i rozważmy bazę

$$u_1 = v_1, \quad u_2 = \frac{1}{\epsilon}v_2.$$

Wtedy

$$Au_1 = \lambda u_1,$$

oraz

$$\begin{aligned} Au_2 &= \frac{1}{\epsilon} Av_1 \\ &= \frac{1}{\epsilon}(v_1 + \lambda v_2) \\ &= \epsilon u_1 + \lambda u_2. \end{aligned}$$

Rozważmy iloczyn skalarny $(\,\cdot\,|\,\cdot\,)$ w którym baza $u_1, u_2$ jest ortonormalna. Wtedy dla $z = x_1u_1 + x_2u_2$ mamy

$$\begin{aligned} (Az|z) &= (x_1\lambda u_1 + x_2\epsilon u_1 + \lambda u_2 | x_1u_1 + x_2u_2) \\ &= \lambda\|z\|^2 + \epsilon x_1x_2. \end{aligned}$$

Ponieważ

$$|x_1x_2| \leqslant \frac{1}{2}(x_1^2 + x_2^2),$$

więc

$$(\lambda - \epsilon/2)\|z\|^2 \leqslant (Az|z) \leqslant (\lambda + \epsilon/2)\|z\|^2,$$

więc teza zachodzi dla dostatecznie małego $\epsilon > 0$. □

**U** Z dowodu twierdzenia 26.2 wynika, że w przypadku gdy $A$ spełnia warunek (1) lub (3), to wystarczy założyć, że

$$m \leqslant \operatorname{Re}\lambda_1,\ \operatorname{Re}\lambda_2 \leqslant M.$$

# Rozdział 27

# Normy macierzowe

| SŁOWA KLUCZOWE ROZDZIAŁU |
|---|
| norma macierzowa ♢ norma kompatybilna ♢ norma Frobeniusa ♢ norma operatorowa ♢ promień spektralny ♢ formuła Gelfanda |

Zajmiemy się teraz przykładami norm w przestrzeni $M_{n\times n}(\mathbb{F})$, gdzie $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Powiemy, że norma $\|\cdot\|$ w $M_{n\times n}(\mathbb{F})$ jest *normą macierzową*, jeśli spełnia dodatkowo warunek

$$\|AB\| \leqslant \|A\|\|B\|, \quad A, B \in M_{n\times n}(\mathbb{F}).$$

U Jeśli $\|\cdot\|$ jest normą macierzową, to

$$\|A^n\| \leqslant \|A\|^n.$$

W szczególności, jeśli $\|A\| < 1$, to

$$\lim_{n\to\infty} \|A^n\| = 0.$$

Jeżeli w $\mathbb{F}^n$ jest ustalona norma, to mówimy, że norma macierzowa jest z nią *kompatybilna*, gdy zachodzi warunek

$$\|Ax\| \leqslant \|A\|\|x\|, \quad x \in \mathbb{F}^n.$$

**Przykład 27.1.** Przykładem normy macierzowej w $M_{n\times n}(\mathbb{R})$ kompatybilnej z normą euklidesową jest *norma Frobeniusa*

$$\|A\|_F = \sqrt{\sum_{i,j=1}^{n} a_{ij}^2} = \sqrt{\operatorname{tr} A^T A}.$$

Norma Frobeniusa ma następujące własności:

(1) jeśli $a_1, \ldots, a_n \in \mathbb{R}^n$ są kolumnami macierzy $A$, to

$$\|A\|_F = \sqrt{\sum_{i=1}^{n} \|a_i\|_2^2},$$

(2) jeśli $a(1), \ldots a(n) \in M_{1\times n}(\mathbb{R})$ są wierszami macierzy $A$, to

$$\|A\|_F = \sqrt{\sum_{i=1}^{n} \|a(i)^T\|_2^2},$$

(3) dla $x \in \mathbb{R}^n$ mamy $Ax = [(a(1)^T|x), \ldots, (a(n)^T|x)]^T$, więc

$$\begin{aligned}
\|Ax\|_2 &= \sqrt{\sum_{i=1}^{n} (a(i)^T|x)^2} \\
&\leqslant \sqrt{\sum_{i=1}^{n} \|a(i)^T\|_2^2 \|x\|_2^2}, \quad \text{(nierówność Cauchy'ego-Schwarza)} \\
&= \|A\|_F \|x\|_2.
\end{aligned}$$

(4) dla $A, B \in M_{n\times n}(\mathbb{R})$ mamy

$$\begin{aligned}\|AB\|_F &= \|[Ab_1|\dots|Ab_n]\|_F \\ &= \sqrt{\sum_{i=1}^n \|Ab_i\|_2^2} \\ &\leqslant \|A\|_F \sqrt{\sum_{i=1}^n \|b_i\|_2^2} \\ &= \|A\|_F \|B\|_F.\end{aligned}$$

(5) jeśli $Q \in O(n)$, to

$$\begin{aligned}\|QA\|_F &= \sqrt{\operatorname{tr}(QA)^T QA} \\ &= \sqrt{\operatorname{tr} A^T Q^T QA} \\ &= \sqrt{\operatorname{tr} A^T A} \\ &= \|A\|_F.\end{aligned}$$

(6) jeśli $\sigma_1, \dots, \sigma_r$ są dodatnimi wartościami singularnymi $A$, to

$$\|A\|_F = \sqrt{\sigma_1^2 + \dots + \sigma_r^2}.$$

**Przykład 27.2.** Dla macierz $A \in M_{n\times n}(\mathbb{F})$ definiujemy jej *normę operatorową* przez

$$\|A\| = \max_{\|x\|=1} \|Ax\|,$$

gdzie $\|x\|$ jest ustaloną normą w $\mathbb{F}^n$. Z twierdzenia Weierstrassa wynika, że $\|A\| < +\infty$. Ponadto dla dowolnego $x \in \mathbb{F}^n$ mamy

$$\|Ax\| \leqslant \|A\|\|x\|.$$

Istotnie, nierówność zachodzi dla $x = 0$. Jeśli $x \neq 0$, to

$$\begin{aligned}\|Ax\| &= A(\|x\|x/\|x\|)\| \\ &= \|x\|A(x/\|x\|)\| \\ &\leqslant \|x\|\|A\|.\end{aligned}$$

Przyglądniemy się normom operatorowym w $M_{n\times n}(\mathbb{R})$, zdefiniowanym następująco

$$\|A\|_1 = \max_{\|x\|_1=1} \|Ax\|_1, \quad \|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2, \quad \|A\|_\infty = \max_{\|x\|_\infty=1} \|Ax\|_\infty$$

**Lemat 27.1.** *Dla macierzy $A \in M_{n\times n}(\mathbb{R})$ mamy*

$$\|A\|_1 = \max_{j=1,\dots,n} \left( \sum_{i=1}^n |a_{ij}| \right) = \max\{\|a_1\|_1, \dots, \|a_n\|_1\},$$

$$\|A\|_\infty = \max_{i=1,\dots,n} \left( \sum_{j=1}^n |a_{ij}| \right) = \max\{\|a(1)^T\|_1, \dots, \|a(n)^T\|_1\}.$$

*Dowód.* Pokażemy, że

$$\|A\|_1 = \max\{\|a_1\|_1, \dots, \|a_n\|_1\}.$$

Niech

$$\alpha := \max\{\|a_1\|_1, \dots, \|a_n\|_1\} = \|a_k\|_1,$$

dla pewnego $k \in \{1, \dots, n\}$. Dla $x = [x_1, \dots, x_n]^T \in \mathbb{R}^n$ mamy

$$Ax = [(a(1)^T|x), \dots, (a(n)^T|x)]^T,$$

więc

$$\begin{aligned}
\|Ax\|_1 &= \sum_{i=1}^{n} |(a(i)^T|x)|_1 \\
&\leqslant \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}x_j| \\
&= \sum_{j=1}^{n} \left( |x_j| \sum_{i=1}^{n} |a_{ij}| \right) \\
&\leqslant \alpha \sum_{j=1}^{n} |x_j| \\
&= \alpha \|x\|_1.
\end{aligned}$$

Stąd, $\|Ax\|_1 \leqslant \alpha$, gdy $\|x\|_1 = 1$. W efekcie,

$$\|A\|_1 \leqslant \alpha.$$

Z drugiej strony,

$$\|Ae_k\|_1 = \|a_k\|_1 = \alpha,$$

czyli $\|A\|_1 = \alpha$.

Pokażemy teraz, że

$$\|A\|_\infty = \max\{\|a(1)^T\|_1, \dots, \|a(n)^T\|_1\}.$$

Niech

$$\beta := \max\{\|a(1)^T\|_1, \dots, \|a(n)^T\|_1\} = \|a(k)^T|_1,$$

dla pewnego $k \in \{1, \dots, n\}$. Dla $x = [x_1, \dots, x_n]^T \in \mathbb{R}^n$ z $\|x\|_\infty = \max\{|x_1|, \dots, |x_n|\} = 1$ mamy

$$\begin{aligned}
\|Ax\|_\infty &= \max\{|(a(1)^T|x)|, \dots, |(a(n)|x)|\} \\
&\leqslant \max\Big\{ \sum_{j=1}^{n} |a_{1j}x_j|, \dots, \sum_{j=1}^{n} |a_{nj}x_j| \Big\} \\
&\leqslant \max\Big\{ \sum_{j=1}^{n} |a_{1j}|, \dots, \sum_{j=1}^{n} |a_{nj}| \Big\} \\
&= \beta.
\end{aligned}$$

W efekcie,

$$\|A\|_\infty \leqslant \beta.$$

Z drugiej strony,

$$\|A[\operatorname{sgn} a_{k1}, \dots, \operatorname{sgn} s_{kn}]^T\|_\infty = \|a(k)\|_1 = \beta,$$

czyli $\|A\|_\infty = \beta$. $\square$

Niech $A \in M_{n\times n}(\mathbb{R})$. Jeśli

$$\lambda_1 \geqslant \ldots \geqslant \lambda_n \geqslant 0,$$

są wartościami własnymi macierzy Grama $A^TA$, a $\sigma_i = \sqrt{\lambda_i}$ odpowiadającymi wartościami singularnymi, to istnieją takie macierze ortogonalne $V, U \in O(n)$, że

$$A = VDU^T, \quad D = \operatorname{diag}(\sigma_1, \dots, \sigma_n).$$

**Twierdzenie 27.1.** *Dla $A \in M_{n\times n}(\mathbb{R})$ mamy*

$$\|A\|_2 = \sigma_1, \quad \textit{największa wartość singularna.}$$

*Ponadto jeśli $A$ jest nieosobliwa, to*

$$\|A^{-1}\|_2 = \frac{1}{\sigma_n}.$$

*Dowód.* Ponieważ $V, U$ są ortogonalne, więc

$$\|A\|_2 = \|VDU^T\|_2 = \|D\|_2.$$

Sprawdzamy łatwo, że $\|D\|_2 = \sigma_1$. □

## 27.1 Promień spektralny. Formuła Gelfanda

**Definicja 27.1.** **Promień spektralny i spektrum**

Przez $\sigma(A)$ oznaczamy zbiór wartości własnych macierzy $A \in M_{n\times n}(\mathbb{C})$ i nazywamy go *spektrum macierzy A*. *Promień spektralny* macierzy $A$ definiujemy jako nieujemną liczbę rzeczywistą

$$\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\} \geqslant 0.$$

**Lemat 27.2.** *Dla kompatybilnej normy macierzowej i każdego $k \geqslant 1$ mamy*

$$\rho(A) \leqslant \sqrt[k]{\|A^k\|} \leqslant \|A\|$$

*Dowód.* Jeśli $\lambda \in \sigma(A)$ i $v$ jest odpowiadającym wektorem własnym, to

$$\begin{aligned} |\lambda|^k\|v\| &= \|\lambda^k v\| \\ &= \|A^k v\| \\ &\leqslant \|A^k\|\|v\|, \end{aligned}$$

więc

$$|\lambda|^k \leqslant \|A^k\|.$$

□

**Lemat 27.3.** *Niech*

$$B = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda & 1 \\ 0 & \dots & 0 & 0 & \lambda \end{bmatrix}$$

*będzie blokiem Jordana odpowiadającym wartości własnej $\lambda$. Wtedy*

$$(B^n)_{ij} = \binom{n}{j-i}\lambda^{n-j+i}, \quad j \geqslant i.$$

*Dowód.* Wynika to z bezpośredniego rachunku. Zauważmy, że

$$B = \begin{bmatrix} \lambda & 0 & 0 & \dots & 0 \\ 0 & \lambda & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda & 0 \\ 0 & \dots & 0 & 0 & \lambda \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 1 \\ 0 & \dots & 0 & 0 & 0 \end{bmatrix} = \lambda I + N,$$

gdzie

$$N = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 1 \\ 0 & \dots & 0 & 0 & 0 \end{bmatrix}$$

jest macierzą nilpotentną. Wystarczy porównać współczynniki po obu stronach równości

$$B^n = (\lambda I + N)^n = \sum_{k=0}^{n} \binom{n}{k} \lambda^k N^{n-k}.$$

□

**Twierdzenie 27.2** (Formuła Gelfanda)**.** *Zachodzi równość*

$$\rho(A) = \lim_{n\to+\infty} \sqrt[n]{\|A^n\|_\infty}$$

*Dowód.* Z twierdzenia Jordana istnieje taka macierz nieosobliwa $S$, że $S^{-1}AS$ ma postać Jordana

$$\Lambda = S^{-1}AS = \operatorname{diag}(B_1, \dots, B_s).$$

Wtedy

$$\begin{aligned} \rho(A) &\leqslant \sqrt[n]{\|A^n\|_\infty} \\ &= \sqrt[n]{\|S\Lambda^n S^{-1}\|_\infty} \\ &\leqslant \sqrt[n]{\|S\|_\infty \|S^{-1}\|_\infty} \sqrt[n]{\|\Lambda^n\|_\infty}. \end{aligned}$$

Ponieważ

$$\lim_{n\to+\infty} \sqrt[n]{\|S\|_\infty \|S^{-1}\|_\infty} = 1,$$

więc z twierdzenia o trzech ciągach wystarczy pokazać, że

$$\lim_{n\to\infty} \sqrt[n]{\|\Lambda^n\|_\infty} = \rho(A).$$

Ponieważ

$$\sqrt[n]{\|\Lambda^n\|_\infty} = \max_{k=1,\dots,s} \left\{ \sqrt[n]{\|B_k^n\|_\infty} \right\},$$

więc wystarczy przyglądnąć się blokom Jordana. Z jawnej postaci $B_k^n$ dla $\lambda_k \neq 0$ mamy

$$\begin{aligned} \|B_k^n\|_\infty &= \sum_{j=1}^{d_k} |(B_k^n)_{1j}| \\ &= \sum_{j=1}^{d_k} \binom{n}{j-1} |\lambda_k|^{n-j+1} \\ &= |\lambda_k|^n \left( |\lambda_k|^{1-d_k} \sum_{j=1}^{d_k} d_k \binom{n}{j-1} |\lambda_k|^{d_k-j} \right). \end{aligned}$$

Dla

$$M_k = |\lambda_k|^{1-d_k} \sum_{j=1}^{d_k} |\lambda_k|^{d_k-j}$$

z nierówności $1 \leqslant \binom{n}{j-1} \leqslant n^{d_k}$ mamy

$$M_k |\lambda_k|^n \leqslant \|B_k^n\|_\infty \leqslant M_k n^{d_k} |\lambda_k|^n.$$

Stąd
$$\lim_{n\to\infty} \sqrt[n]{\|B_k^n\|_\infty} = |\lambda_k|.$$
Jeśli $\lambda_k = 0$, to $B_k$ jest nilpotentna, więc powyższa równość jest również spełniona.

Wynika stąd, że
$$\lim_{n\to\infty} \sqrt[n]{\|\Lambda^n\|_\infty} = \max_k\{|\lambda_k|\} = \rho(A).$$
□

**Wniosek 27.1.** *Formuła Gelfanda zachodzi dla dowolnej normy w $M_{n\times n}(\mathbb{R})$.*

*Dowód.* Wynika to z równoważności norm i twierdzenia o trzech ciągach. □

**Wniosek 27.2.** *Niech $A \in M_{n\times n}(\mathbb{C})$. Wtedy, $\rho(A) < 1$ wtedy i tylko wtedy, gdy*
$$\lim_{k\to+\infty} A^k = 0.$$
*Ponadto, jeśli $\rho(A) > 1$, to*
$$\lim_{k\to+\infty} \|A^k\| = \infty.$$

**Lemat 27.4.** *Niech $A \in M_{n\times n}(\mathbb{C})$ i $\epsilon > 0$ będą ustalone. Wtedy, istnieje taka norma $\|\cdot\|_*$ w $\mathbb{C}^n$, że dla indukowanej normy operatorowej mamy*
$$\rho(A) \leqslant \|A\|_* \leqslant \rho(A) + \epsilon.$$

*Dowód.* Z twierdzenia Schura $A = U^*TU$ dla pewnej macierzy unitarnej $U \in M_{n\times n}(\mathbb{C})$ i macierzy górnie trójkątnej $T$ postaci
$$T = \begin{bmatrix} \lambda_1 & t_{12} & \dots & t_{1n} \\ 0 & \lambda_2 & \dots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix},$$
gdzie $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ są wartościami własnymi $A$. Dla ustalonego $\delta > 0$ definiujemy macierz diagonalną
$$D_\delta := \operatorname{diag}(1, \delta, \delta^2, \dots, \delta^{n-1})$$
oraz normę $\|\cdot\|_*$ w $\mathbb{R}^n$ przez
$$\|x\|_* := \|D_\delta^{-1}Ux\|_\infty.$$
Dla indukowanej normy macierzowej mamy
$$\begin{aligned} \|A\|_* &= \max_{x\neq 0} \frac{\|Ax\|_*}{\|x\|_*} \\ &= \max_{x\neq 0} \frac{\|D_\delta^{-1}UAx\|_\infty}{\|D_\delta^{-1}Ux\|_\infty} \\ &= \max_{x\neq 0} \frac{\|D_\delta^{-1}UU^*TUx\|_\infty}{\|D_\delta^{-1}Ux\|_\infty}. \end{aligned}$$
Ponieważ $D_\delta^{-1}U$ jest izomorfizmem liniowym na $\mathbb{C}^n$, więc dla $y = D_\delta^{-1}Ux$ mamy
$$\|A\|_* = \max_{y\neq 0} \frac{\|D_\delta^{-1}TD_\delta y\|_\infty}{\|y\|_\infty} = \|D_\delta^{-1}TD_\delta\|_\infty.$$
Sprawdzamy bezpośrednim rachunkiem, że
$$D_\delta^{-1}TD_\delta = \begin{bmatrix} \lambda_1 & \delta t_{12} & \dots & & \delta^{n-1}t_{1n} \\ 0 & \lambda_2 & \dots & & \delta^{n-2}t_{2n} \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & \dots & & \lambda_{n-1} & \delta t_{(n-1)n} \\ 0 & 0 & \dots & 0 & \lambda_n \end{bmatrix},$$

więc jeśli $\delta > 0$ jest dostatecznie mała, to

$$\|D_\delta^{-1} T D_\delta\|_\infty \leqslant \rho(A) + \epsilon.$$

□

**Wniosek 27.3.** *Dla dowolnej macierzy $A \in M_{n\times n}(\mathbb{C})$ mamy*

$$\rho(A) = \inf\{\|A\| : \|\cdot\| \quad \textit{jest kompatybilną normą macierzową}\}.$$

## 27.2 Wartości singularne a wartości własne

Niech $A \in M_{n\times n}(\mathbb{C})$ oraz niech

$$\sigma_1(A) \geqslant \ldots \geqslant \sigma_n(A) \geqslant 0$$

będą jej wartościami singularnymi, a $\lambda_1(A), \ldots, \lambda_n(A) \in \mathbb{C}$ jej wartościami własnymi uporządkowanymi przez

$$\rho(A) = |\lambda_1(A)| \geqslant \ldots \geqslant |\lambda_n(A)|.$$

**Wniosek 27.4.** *Dla $A \in M_{n\times n}(\mathbb{C})$ mamy*

*(a)* $\rho(A) \leqslant \|A\|_2 = \sigma_1(A)$,

*(b)*

$$\rho(A) = \lim_{m\to\infty} \sqrt[m]{\|A^m\|_2} = \lim_{m\to\infty} \sqrt[m]{\sigma_1(A^m)},$$

*(c)*

$$|\lambda_1(A)\ldots\lambda_k(A)| \leqslant \sigma_1(A)\ldots\sigma_k(A), \quad k = 1, \ldots, n.$$

*(d) jeśli $A_p$ powstaje z $A$ przez wykreślenie $p$ kolumn (lub $p$ wierszy), to*

$$\sigma_i(A) \geqslant \sigma_i(A_p) \geqslant \sigma_{i+p}(A), \quad i = 1, \ldots, n-p.$$

**Wniosek 27.5.** *Dla $A \in M_{n\times n}(\mathbb{C})$ mamy*

$$\lim_{m\to\infty} \sqrt[m]{\sigma_n(A^m)} = |\lambda_n(A)|.$$

*Dowód.* Jeśli $A$ jest singularna, to $A^m$ jest singularna dla każdego $m = 1, 2, \ldots$, więc

$$\sigma_n(A^m) = \lambda_n(A^m) = 0,$$

czyli teza zachodzi. Jeśli $A$ jest nieosobliwa, to stosujemy punkt (b) wniosku 27.4 do macierzy $A^{-1}$ korzystając z faktu, że

$$\sigma_1(A^{-1}) = \frac{1}{\sigma_n(A)}, \quad \lambda_1(A^{-1}) = \frac{1}{\lambda_n(A)}.$$

□

**Twierdzenie 27.3.** *Dla $A \in M_{n\times n}(\mathbb{C})$ oraz $k = 1, \ldots, n$ mamy*

$$\lim_{m\to\infty} \sqrt[m]{\sigma_k(A^m)} = |\lambda_k(A)|.$$

*Dowód.* Wiemy, że teza zachodzi dla $k = 1$ oraz $k = n$. Z twierdzenia Schura istnieje taka macierz unitarna $U \in \mathrm{U}(n)$, że $U^*AU = T$ jest górnie trójkątna oraz $t_{ii} = \lambda_i(A) = \lambda_i(T)$. Dla $k = 2, \dots, n-1$ przez $T_{[k]} \in M_{k\times k}(\mathbb{C})$ oznaczamy macierz powstałą z $T$ przez wykreślenie $k$ ostatnich wierszy i kolumn, a przez $T_{<k>} \in M_{(n-k+1)\times(n-k+1)}(\mathbb{C})$ macierz powstałą z $T$ przez wykreślenie pierwszych $k-1$ wierszy i kolumn. Zauważmy, że dla macierzy górnie trójkątnej dla $m \geqslant 1$ mamy

$$(T_{[k]})^m = (T^m)_{[k]}, \quad (T_{<k>})^m = (T^m)_{<k>}.$$

Z punktu (d) we wniosku 27.4 oraz unitarnej niezmienniczości wartości singularnych mamy

$$\begin{aligned}\sigma_k(A^m) &= \sigma_k(T^m)\\ &\geqslant \sigma_k((T^m)_{[k]})\\ &= \sigma_k((T_{[k]})^m).\end{aligned}$$

Z drugiej strony, dla $i = 1$ oraz $p = k-1$ w (d) mamy

$$\begin{aligned}\sigma_k(A^m) &= \sigma_k(T^m)\\ &= \sigma_{1+(k-1)}(T^m)\\ &\leqslant \sigma_1((T^m)_{<k>})\\ &= \sigma_1((T_{<k>})^m).\end{aligned}$$

Otrzymujemy, że

$$\sqrt[m]{\sigma_k((T_{[k]})^m)} \leqslant \sqrt[m]{\sigma_k(A^m)} \leqslant \sqrt[m]{\sigma_1((T_{<k>})^m)}.$$

Z wniosku 27.5 wynika, że

$$\lim_{m\to\infty} \sqrt[m]{\sigma_k((T_{[k]})^m)} = \left|\lambda_k(T_{[k]})\right|,$$

a z punktu (b) we wniosku 27.4 mamy

$$\lim_{m\to\infty} \sqrt[m]{\sigma_1((T_{<k>})^m)} = \left|\lambda_1(T_{<k>})\right|.$$

Zauważmy, że

$$\lambda_k(T_{[k]}) = \lambda_k(T) = \lambda_k(A),$$

oraz

$$\lambda_1(T_{<k>}) = \lambda_k(T) = \lambda_k(A),$$

czyli teza zachodzi. $\square$

# Rozdział 28

# Pseudoodwrotność Moore'a-Penrose'a

| | |
|---|---|
| SŁOWA KLUCZOWE ROZDZIAŁU<br>pseudoodwrotność Moore'a–Penrose'a ♢ warunki Penrose'a ♢ metoda najmniejszych kwadratów | $AA^+A = A$<br>$A^+AA^+ = A^+$<br>$(AA^+)^T = AA^+$<br>$(A^+A)^T = A^+A$ |

Jeśli macierz kwadratowa $A \in M_{n\times n}(\mathbb{R})$ jest nieosobliwa, to równanie $Ax = b$ ma dokładnie jedno rozwiązanie

$$x = A^{-1}b.$$

Jeśli $A \in M_{m\times n}$ i $m > n$, to zazwyczaj równanie $Ax = b$ nie ma rozwiązań. Jeśli kolumny macierzy $A$ są liniowo niezależne, to przybliżone rozwiązanie w sensie najmniejszych kwadratów jest dane przez

$$\tilde{x} = (A^TA)^{-1}A^Tb.$$

Macierz $(A^TA)^{-1}A^T$ pełni tu w pewnym sensie rolę „odwrotności" macierzy $A$.

**Definicja 28.1.** Pseudoodwrotność macierzy

Niech $A \in M_{m\times n}$ będzie macierzą o liniowo niezależnych kolumnach. Macierz

$$A^+ = (A^TA)^{-1}A^T \in M_{n\times m}(\mathbb{R})$$

nazywamy *pseudoodwrotnością* macierzy $A$.

U Jeśli $m = n$ i $A$ ma liniowo niezależne kolumny, to $A^+ = A^{-1}$.

**Lemat 28.1** (Warunki Penrose'a)**.** *Jeśli $A \in M_{m\times n}(\mathbb{R})$ ma liniowo niezależne kolumny, to jej pseudoodwrotność $A^+$ spełnia warunki*

*(a)* $AA^+A = A$,

*(b)* $A^+AA^+ = A^+$,

*(c) macierze $AA^+$ i $A^+A$ są symetryczne.*

*Dowód.* Udowodnimy warunek (a). Dowód warunku (b) jest analogiczny. Mamy

$$\begin{aligned} AA^+A &= A\big((A^TA)^{-1}A^T\big)A \\ &= A(A^TA)^{-1}(A^TA) \\ &= A. \end{aligned}$$

Dla dowodu (c) pokażemy, że $AA^+$ jest symetryczna. Macierz Grama $A^TA$ jest symetryczna i odwracalna, więc jej odwrotność $(A^TA)^{-1}$ jest również symetryczna. Mamy

$$\begin{aligned} (AA^+)^T &= \big(A(A^TA)^{-1}A^T\big)^T \\ &= (A^T)^T\big((A^TA)^{-1}\big)^TA^T \\ &= A(A^TA)^{-1}A^T \\ &= AA^+. \end{aligned}$$

□

Wykorzystamy teraz rozkład singularny do rozszerzenie pojęcia pseudoodwrotności macierzy $A \in M_{m\times n}(\mathbb{R})$ dla dowolnej macierzy $A$, czyli o niekoniecznie liniowo niezależnych kolumnach. Konstrukcja ta pochodzi od E. H. Moore'a (1862-1932) i Rogera Penrose'a (ur. 1931).

**Definicja 28.2.** **Pseudoodwrotność macierzy**

Niech $A \in M_{m\times n}$. Załóżmy, że

$$A = V \underbrace{\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}}_{=\Sigma\in M_{m\times n}(\mathbb{R})} U^T, \quad D = \operatorname{diag}(\sigma_1, \ldots, \sigma_r),$$

gdzie $\sigma_1 \geqslant \ldots \geqslant \sigma_1 > 0$ są niezerowymi wartościami singularnymi macierzy $A$. *Pseudoodwrotność Moore'a–Penrose'a* macierzy $A$ jest zdefiniowana jako

$$A^+ := U\Sigma^+V^T, \quad \Sigma^+ = \begin{bmatrix} D^{-1} & 0 \\ 0 & 0 \end{bmatrix} \in M_{n\times m}(\mathbb{R}).$$

**Lemat 28.2.** *Dla macierzy $A \in M_{m\times n}(\mathbb{R})$ o liniowo niezależnych kolumnach mamy $A^+A = I \in M_{m\times m}(\mathbb{R})$, czyli $A^+ \in M_{n\times m}(\mathbb{R})$ jest lewą odwrotnością $A$.*

*Dowód.* Ponieważ $r = n$, więc

$$\Sigma = \left[\begin{array}{c} D \\ \hline 0 \end{array}\right] \in M_{m\times n}(\mathbb{R}), \quad \Sigma^+ = \left[\begin{array}{c|c} D^{-1} & 0 \end{array}\right] \in M_{n\times m}(\mathbb{R}).$$

Mamy

$$\begin{aligned} A^+A &= (U\Sigma^+V^T)(V\Sigma U^T) \\ &= U\Sigma^+\Sigma U^T \\ &= UIU^T \\ &= I. \end{aligned}$$

□

**Lemat 28.3.** *Jeśli $A \in M_{m\times n}(\mathbb{R})$ ma liniowo niezależne kolumny, to*

$$A^+ = (A^TA)^{-1}A^T.$$

*Dowód.* Mamy pokazać, że

$$(A^T A)A^+ = A^T.$$

Z lematu 28.2 wynika, że

$$\begin{aligned}(A^T A)A^+ A &= A^T AI\\ &= A^T A,\end{aligned}$$

czyli $(A^T A)A^+$ i $A^T$ pokrywają się na obrazie $\operatorname{im} A$. Ponieważ

$$\mathbb{R}^m = \ker A^T \oplus \operatorname{im} A,$$

więc wystarczy pokazać, że

$$\ker A^+ = \ker A^T.$$

Zauważmy, że

$$A^T = U\Sigma^T V^T = U \left[\begin{array}{c|c} D & 0 \end{array}\right] V^T$$

oraz

$$A^+ = U \left[\begin{array}{c|c} D^{-1} & 0 \end{array}\right] V^T,$$

więc $A^T$ i $A^+$ mają takie same jądra. □

**Lemat 28.4.** *Dla dowolnej macierzy $A \in M_{m\times n}(\mathbb{R})$ psedodowrotność $A^+$ spełnia warunki Penrose'a*

*(a) $AA^+A = A$,*

*(b) $A^+AA^+ = A^+$,*

*(c) macierze $AA^+$ i $A^+A$ są symetryczne oraz idempotentne, więc są rzutami ortogonalnymi.*

*Dowód.* Przy oznaczeniach definicji 28.2 sprawdzamy łatwo, że

$$\Sigma\Sigma^+\Sigma = \Sigma, \quad \Sigma^+\Sigma\Sigma^+ = \Sigma^+$$

oraz macierze

$$\Sigma\Sigma^+, \quad \Sigma^+\Sigma$$

są symetryczne. Udowodnimy punkt (a). dowód (b) jest analogiczny. Mamy

$$\begin{aligned}AA^+A &= (V\Sigma U^T)(U\Sigma^+V^T)(V\Sigma U^T)\\ &= V(\Sigma\Sigma^+\Sigma)U^T\\ &= V\Sigma U^T\\ &= A.\end{aligned}$$

Dla dowodu punktu (c) sprawdzimy, że $AA^+$ jest symetryczna. Rzeczywiście,

$$\begin{aligned}AA^+ &= (V\Sigma U^T)(U\Sigma^+V^T)\\ &= V(\Sigma\Sigma^+)V^T,\end{aligned}$$

więc $AA^+$ jest symetryczna, bo $\Sigma\Sigma^+$ jest symetryczna.

Macierz $AA^+$ jest idempotentna, bo $AA^+A = A$, więc

$$A^+AA^+A = A^+A.$$

□

**Lemat 28.5.** *Dla macierzy $A \in M_{m\times n}(\mathbb{R})$ pseudoodwrotność $A^+ \in M_{n\times m}(\mathbb{R})$ jest jedyną macierzą spełniającą warunki Penrose'a (a)-(c).*

*Dowód.* Przypuśćmy, że $A'$ spełnia warunki (a)-(c). Pokażemy, że $A^+ = A'$. Mamy

$$\begin{aligned} AA^+ &= AA'AA^+ \\ &= (AA')^T(AA^+)^T \\ &= (A')^T(AA^+A)^T \\ &= (A')^TA^T \\ &= (AA')^T \\ &= AA'. \end{aligned}$$

Analogicznie pokazujemy, że $A^+A = A'A$. Stąd

$$\begin{aligned} A^+ &= A^+AA^+ \\ &= A^+AA' \\ &= A'AA' \\ &= A'. \end{aligned}$$

□

**Wniosek 28.1.** *Dla dowolnej macierzy $A \in M_{m\times n}(\mathbb{R})$ mamy $(A^+)^+ = A$.*

*Dowód.* Z lematu 28.5 wystarczy zauważyć, że $A$ spełnia warunki Penrose'a dla $A^+$. □

**Wniosek 28.2.** *Jeśli $A \in M_{n\times n}(\mathbb{R})$ jest nieosobliwa, to $A^+ = A^{-1}$.*

**Wniosek 28.3.** *Jeśli $A \in M_{n\times n}(\mathbb{R})$ jest symetryczna i idempotentna (jest rzutem ortogonalnym), to $A^+ = A$.*

*Dowód.* Wynika z Lematu 28.5, bo $A$ spełnia warunki Penrosa dla $A$. □

Podamy teraz zastosowanie macierzy pseudodwrotnej do metody najmniejszych kwadratów. Zaczniemy od prostego przykładu.

**Przykład 28.1.** Rozważmy

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad b = [0,1]^T.$$

Układ równań $Ax = b$ jest sprzeczny, więc pozostaje nam szukać rozwiązań w sensie najmniejszych kwadratów. Są one rozwiązaniami układu

$$A^TA[x,y]^T = A^T[0,1]^T.$$

Ponieważ kolumny $A$ są liniowo zależne, więc ma on nieskończenie wiele rozwiązań. Spełniają one

$$\underbrace{\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}}_{=A^TA} \begin{bmatrix} x \\ y \end{bmatrix} = \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{=A^Tb},$$

czyli rozwiązania spełniają warunek

$$x + y = \frac{1}{2}.$$

Zastanówmy się które z nich ma najmniejszą normę. Chcemy, więc znaleźć minimum funkcji

$$f(x) = x^2 + (\frac{1}{2} - x)^2 = \frac{1}{2}(x - 1/4)^2.$$

Jest ono osiągane dla $x = 1/4$, czyli rozwiązaniem (w sensie najmniejszych kwadratów) o najmniejszej normie jest

$$[x,y]^T = [1/4, 1/4]^T.$$

**Twierdzenie 28.1.** *Niech $A \in M_{m\times n}(\mathbb{R})$ i $b \in \mathbb{R}^m$. Wtedy*

$$\min_{x\in\mathbb{R}^n} \|b - Ax\|^2$$

*jest osiągane dla*

$$\tilde{x} = A^+ b.$$

*Ponadto, $\tilde{x}$ ma najmniejszą normę spośód wektorów minimalizujących.*

*Dowód.* Niech $r = \operatorname{rank} A$. Rozważmy rozkład singularny

$$A = V\Sigma U^T,$$

czyli $A^+ = U\Sigma^+ V^T$. Niech

$$y = U^T x = [y_1, y_2]^T, \quad c = V^T b = [c_1, c_2]^T.$$

Ponieważ $V^T$ jest ortogonalne, więc

$$\begin{aligned}
\|b - Ax\|^2 &= \|V^T(b - Ax)\|^2 \\
&= \|V^T b - \Sigma U^T x\|^2 \\
&= \|c - \Sigma y\|^2 \\
&= \|c_1 - Dy_1, c_2\|^2.
\end{aligned}$$

Minimum będzie osiągnięte, gdy $c_1 = Dy_1$, czyli dla $y_1 = D^{-1}c_1$. Interesują nas więc wektory

$$x = Uy = U[D^{-1}c_1, y_2]^T. \tag{28.1}$$

Pokażemy, że

$$\tilde{x} = U\underbrace{[D^{-1}c_1, 0]^T}_{=\tilde{y}}$$

ma najmniejszą normę spośród wektorów spełniających (28.1). Rzeczywiście, jeśli $y_2 \neq 0$, to

$$\begin{aligned}
\|\tilde{x}\| &= \|U\tilde{y}\| = \|\tilde{y}\| \\
&< \|y\| = \|Uy\| = \|x\|.
\end{aligned}$$

Pozostaje pokazać, że $\tilde{x} = A^+ b$. Mamy

$$\begin{aligned}
\tilde{x} &= U\tilde{y} \\
&= U[D^{-1}c_1, 0]^T \\
&= U\Sigma^+[c_1, c_2] \\
&= U\Sigma^+ V^T b \\
&= A^+ b.
\end{aligned}$$

□

**Przykład 28.2.** Rozważmy

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad b = [0, 1]^T.$$

z przykładu 28.1. Rozkład singularny macierzy $A$ jest dany przez

$$A = \underbrace{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}}_{=V} \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}}_{=\Sigma} \underbrace{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}}_{=U^T}.$$

Stąd

$$A^+ = \underbrace{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}}_{=U} \underbrace{\begin{bmatrix} 1/2 & 0 \\ 0 & 0 \end{bmatrix}}_{=\Sigma^+} \underbrace{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}}_{=V^T} = \begin{bmatrix} 1/4 & 1/4 \\ 1/4 & 1/4 \end{bmatrix},$$

więc

$$\tilde{x} = A^+ b = \begin{bmatrix} 1/4 & 1/4 \\ 1/4 & 1/4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/4 \\ 1/4 \end{bmatrix}.$$

Otrzymane rozwiązanie pokrywa się z tym otrzymanym analitycznie w przykładzie 28.1.

**Przykład 28.3.** Pseudoodwrotność Moore'a–Penrose'a $A^+$ nie jest funkcją ciągłą współczynników macierzy $A$. Dla $\epsilon \neq 0$ rozważmy przykładowo macierz

$$A_\epsilon = \begin{bmatrix} 1 & 1 \\ 1 & 1+\epsilon \end{bmatrix}.$$

Jest to macierz odwracalna, więc

$$A_\epsilon^+ = A_\epsilon^{-1} = \begin{bmatrix} 1+\frac{1}{\epsilon} & -\frac{1}{\epsilon} \\ -\frac{1}{\epsilon} & \frac{1}{\epsilon} \end{bmatrix}.$$

Z drugiej strony dla $\epsilon = 0$, mamy

$$A_0^+ = \begin{bmatrix} 1/4 & 1/4 \\ 1/4 & 1/4 \end{bmatrix},$$

więc $A_\epsilon \not\to A_0$, gdy $\epsilon \to 0$.

## 28.1 Fundamentalne twierdzenia algebry liniowej

Rozważmy macierz

$$A = \left[\, a_1 \mid \ldots \mid a_n \,\right] = \begin{bmatrix} a(1) \\ \vdots \\ a(m) \end{bmatrix} \in M_{m\times n}(\mathbb{R})$$

o kolumnach $a_1, \ldots, a_n \in \mathbb{R}^m$ i wierszach $a(1)^T, \ldots, a(m)^T \in \mathbb{R}^n$. W wielu zagadnieniach podstawową rolę pełnią podprzestrzenie wektorowe

$$\ker A,\ \operatorname{im} A^T \subset \mathbb{R}^n, \quad \operatorname{im} A,\ \ker A^T \subset \mathbb{R}^m$$

oraz ich wzajemne zależności. Podsumujmy nasze dotychczasowe rozważania.

(i) **Podprzestrzenie rozpięte na kolumnach i wierszach:**

$$\operatorname{im} A = \operatorname{span}\{a_1, \ldots, a_n\}, \quad \operatorname{im} A^T = \operatorname{span}\{a(1)^T, \ldots, a(m)^T\}$$

oraz

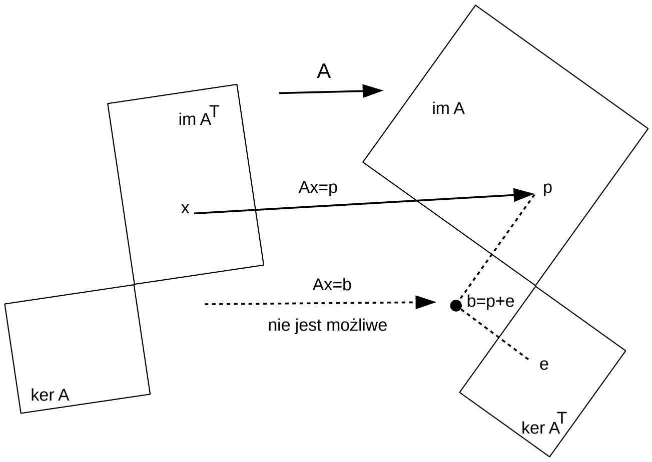
$$\operatorname{rank} A := \dim \operatorname{im} A = \dim \operatorname{im} A^T =: \operatorname{rank} A.$$

(ii) **Ortogonalne sumy proste:**

$$(\ker A)^\perp = \operatorname{im} A^T, \quad (\ker A^T)^\perp = \operatorname{im} A.$$

W szczególności, mamy ortogonalne sumy proste

$$\mathbb{R}^n = \ker A \oplus \operatorname{im} A^T, \quad \mathbb{R}^m = \ker A^T \oplus \operatorname{im} A.$$




Rysunek 28.1: Metoda najmniejszych kwadratów.

(iii) **Macierz Grama $A^TA$:**

$$\operatorname{im} A^TA = \operatorname{im} A^T, \quad \ker A^TA = \ker A.$$

Macierz $A^TA \in M_{n\times n}(\mathbb{R})$ jest kwadratowa oraz $A^TA$ jest nieosobliwa, gdy $\operatorname{rank} A = n$ (kolumny macierzy $A$ są liniowo niezależne lub równoważnie $\ker A = \{0\}$).

(iv) **Metoda najmniejszych kwadratów:**

Równanie $Ax = b$ dla ustalonego wektora $b \in \mathbb{R}^m$ ma rozwiązanie wtedy i tylko wtedy, gdy $b \in \operatorname{im} A$ czyli jeśli $b$ jest kombinacją liniową kolumn macierzy $A$.

Jeśli $b \notin \operatorname{span}\{a_1, \ldots, a_n\}$, to zbiór rozwiązań $Ax = b$ jest pusty. Możemy wtedy poszukać takiego wektora $x \in \mathbb{R}^n$ aby norma $\|b - Ax\|$ była jak najmniejsza. Będzie tak wtedy, gdy $Ax$ będzie rzutem ortogonalnym $b$ na obraz $\operatorname{im} A$. Wtedy

$$b - Ax \in (\operatorname{im} A)^\perp = \ker A^T,$$

czyli $x$ spełnia równanie

$$A^TAx = A^Tb.$$

Jak wiemy, macierz Grama $A^TA$ jest odwracalna, gdy kolumny macierzy $A$ są liniowo niezależne. Wtedy

$$x = (A^TA)^{-1}A^Tb.$$

(v) **Rozkład singularny:**

Macierz Grama $A^TA$ jest symetryczna i ma nieujemne wartości własne $\lambda_1 \geqslant \ldots \geqslant \lambda_n$. Z twierdzenia o rozkładzie singularnym, dla wartości singularnych

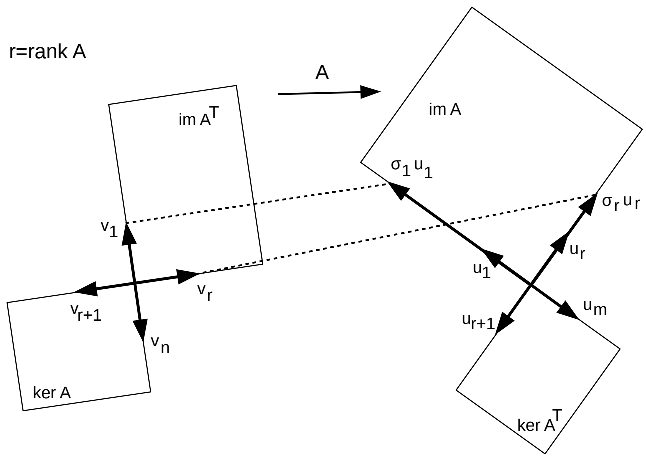
$$\sigma_1 = \sqrt{\lambda_1} \geqslant \ldots \geqslant \sigma_n = \sqrt{\lambda_n} \geqslant 0$$

mamy

$$A^TAv_i = \sigma_i^2 v_i, \quad i = 1, \ldots, n$$

dla pewnej bazy ortonormalnej $v_1, \ldots, v_n \in \mathbb{R}^n$. Wtedy

$$\begin{aligned}\|Av_i\|^2 &= (Av_i|Av_i)\\ &= (A^TAv_i|v_i)\\ &= \sigma_i^2(v_i|v_i)\\ &= \sigma_i^2,\end{aligned}$$




Rysunek 28.2: Rozkład singularny macierzy $A \in M_{m\times n}(\mathbb{R})$ rzędu $r$.

więc

$$\|Av_i\| = \sigma_i.$$

Niech $r = \operatorname{rank} A^T A = \operatorname{rank} A$. Ponieważ

$$AA^T Av_i = \sigma_i^2 Av_i, \quad i = 1, \ldots, r$$

więc

$$u_i = \frac{1}{\sigma_i} Av_i, \quad i = 1, \ldots, r$$

jest jednostkowym wektorem własnym macierzy $AA^T$. Wektory $u_1, \ldots, u_r$ tworzą układ ortonormalny. Ponieważ

$$Av_i = \sigma_i u_i,$$

więc

$$AV = U\Sigma,$$

gdzie

**(1)** $U = \left[\begin{array}{c|c|c} u_1 & \ldots & u_m \end{array}\right] \in O(m)$, gdzie $u_1, \ldots, u_r$ jest bazą ortononormalną dla $\operatorname{im} A$ i $u_{r+1}, \ldots, u_m$ jest bazą ortonormalną dla $\ker A^T$,

**(2)** $V = \left[\begin{array}{c|c|c} v_1 & \ldots & v_n \end{array}\right] \in O(n)$, gdzie $v_1, \ldots, v_r$ jest bazą ortononormalną dla $\operatorname{im} A^T$ i $v_{r+1}, \ldots, v_n$ jest bazą ortonormalną dla $\ker A$,

**(3)** $\Sigma \in M_{m\times n}(\mathbb{R})$ jest pseudodiagonalna z niezerowymi wyrazami diagonalnymi

$$\sigma_1 > \ldots > \sigma_r > 0.$$

**Przykład 28.4.** Niech

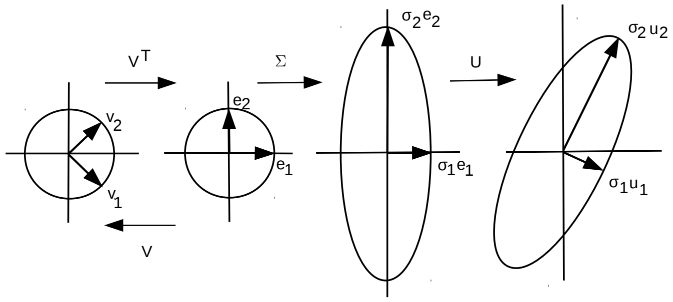
$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}.$$

Wszystkie cztery podprzestrzenie fundamentalne macierzy $A$ są jednowymiarowe. Mamy

$$\operatorname{im} A^T = \operatorname{span}\{[1,2]^T\}, \quad \ker A = \{[-2,1]^T\},$$

więc

$$V = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}.$$




Rysunek 28.3: Diagonalizacja przez dobrze dobrane bazy ortonormalne.

Z drugiej strony

$$\operatorname{im} A = \operatorname{span}\{[1,3]^T\}, \quad \ker A^T = \{[-3,1]^T\},$$

czyli

$$U = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}.$$

Macierz Grama

$$A^T A = \begin{bmatrix} 10 & 20 \\ 20 & 40 \end{bmatrix}$$

ma wartości własne $\lambda_1 = 50$ i $\lambda_2 = 0$, więc wartości singularne macierzy $A$ są równe $\sigma_1 = \sqrt{50}$ i $\sigma_2 = 0$. W efekcie

$$A = \frac{\begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}}{\sqrt{10}} \begin{bmatrix} \sqrt{50} & 0 \\ 0 & 0 \end{bmatrix} \frac{\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}}{\sqrt{5}} = U\Sigma V^T.$$

Macierz $A$ możemy zapisać jako kombinację $r = \operatorname{rank} A$ macierzy o rzędzie 1:

$$A = U\Sigma V^T = u_1 \sigma_1 v_1^T + \ldots + u_r \sigma_r v_r^T.$$

W rozważanym przypadku

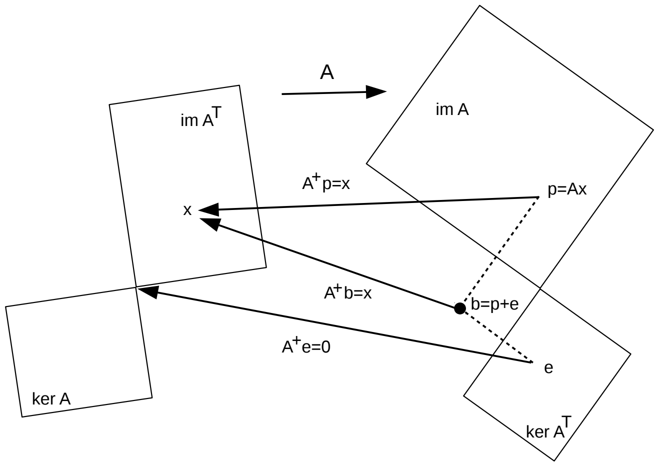
$$A = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix}.$$

Porównajmy rozkład Jordana macierzy kwadratowej $A \in M_{n\times n}(\mathbb{R})$ z jej rozkładem singularnym. W rozkładzie Jordana $SAS^{-i} = J$, więc wybieramy tę samą bazę w dziedzinie i przeciwdziedzinie. Mamy bardzo związane ręce. W szczególności, $J$ nie musi być nawet rzeczywista. Jeśli $J$ jest rzeczywista, to nie musi być diagonalna. Nawet jeśli, $J$ jest diagonalna, to $S$ nie musi być ortogonalna. Dopuszczając różne bazy w rozkładzie singularnym $U^T AV = \Sigma$, każda macierz działa tak, jak macierz symetryczna tzn. $U$ i $V$ są ortogonalne i $\Sigma$ jest diagonalna. Z punktu widzenia metody najmniejszych kwadratów rozkład singularny jest najlepszym wyborem. Z drugiej strony, do obliczania potęg $A^n$ lepszy jest rozkład Jordana, bo chcemy aby $S$ i $S^{-1}$ się „skróciły".

(vi) **pseudoodwrotność Moore'a–Penrose'a**

Rozkład singularny $A = U\Sigma V^T$ macierzy $A \in M_{m\times n}(\mathbb{R})$ prowadzi bezpośrednio do pseudoodwrotności Moore'a–Penrose'a $A^+$. Ma ona pełnić rolę macierzy odwrotnej do $A$, gdy $A$ nie jest odwracalna. Dla dodatniej wartości singularnej $\sigma_i > 0$ mamy $Av_i = \sigma_i \cdot u_i$, więc musi zachodzić

$$A^+ u_i = \frac{1}{\sigma_i} \cdot v_i, \quad i = 1, \ldots, r = \operatorname{rank} A.$$




Rysunek 28.4: Pseudoodwrotność Moore'a–Penrose'a $A^+$.

Pseudoodwrotność $\Sigma^+$ macierzy $\Sigma$ ma niezerowe wyrazy $1/\sigma_i$ dla dodatnich wartości singularnych $\sigma_i$. Macierze ortogonalne $U$ i $V^T$ są odwracane przez $U^T$ i $V$. Pseudoodwrotność macierzy $A = U\Sigma V^T$ jest równa

$$A^+ = V\Sigma^+U^T.$$

Podsumujmy, że dla $i = 1, \ldots, r$ mamy

$$Av_i = \sigma_i \cdot u_i, \quad A^T u_i = \sigma_i \cdot v_i, \quad A^+ u_i = \frac{1}{\sigma_i} \cdot v_i.$$

**Przykład 28.5.** Dla macierzy

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$$

mamy

$$A^+ = \frac{\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}}{\sqrt{5}} \begin{bmatrix} 1/\sqrt{50} & 0 \\ 0 & 0 \end{bmatrix} \frac{\begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix}}{\sqrt{10}} = \frac{1}{50} \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}.$$

Macierz $A^+A$ jest zawsze identycznością na podprzestrzeni $\operatorname{im} A^T$ i zeruje się na $\ker A$. W rozważanym przypadku

$$A^+A = \frac{1}{50} \begin{bmatrix} 10 & 20 \\ 20 & 40 \end{bmatrix}$$

jest rzutem ortogonalnym na span $\{[1, 2]^T\}$.

Z drugiej strony, $AA^+$ jest identycznościowa na podprzestrzeni $\operatorname{im} A$ i zeruje się na $\ker A^T$. U nas

$$AA^+ = \frac{1}{50} \begin{bmatrix} 5 & 15 \\ 15 & 45 \end{bmatrix}$$

jest rzutem ortogonalnym na span $\{[1, 3]^T\}$.

# Rozdział 29

# Macierze nieujemne

| |
|---|
| SŁOWA KLUCZOWE ROZDZIAŁU |
| macierz dodatnia ◊ macierz nieujemna ◊ twierdzenie Frobeniusa–Perrona ◊ dominująca wartość własna |

**Definicja 29.1.** Macierz dodatnia i nieujemna

Macierz $A \in M_{n\times n}(\mathbb{R})$ nazywamy *nieujemną*, jeśli $a_{ij} \geqslant 0$ dla $i, j = 1, \dots n$. Jeśli $a_{ij} > 0$ dla $i, j = 1, \dots n$, to $A$ nazywamy *dodatnią*.

Podobnie, wektor $x = [x_1, \dots, x_n]^T$ jest *nieujemny*, gdy $x_i \geqslant 0$ oraz jest dodatni, gdy $x_i > 0$ dla $i = 1, \dots, n$.

Powiemy, że $A \geqslant B$ $(A > B)$, jeśli $a_{ij} \geqslant b_{ij}$ $(a_{ij} > b_{ij})$ dla wszystkich $i, j$.

**Definicja 29.2.** Promień spektralny

*Promień spektralny macierzy* $A \in M_{n\times n}(\mathbb{C})$ definiujemy jako

$$\rho(A) = \max\{|\lambda| : \lambda \quad \text{wartość własna} \quad A\}.$$

**Twierdzenie 29.1** (Frobenius–Perron)**.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą dodatnią. Wtedy*

*(i) promień spektralny $\rho(A) > 0$ jest wartością własną $A$,*

*(ii) $\rho(A)$ ma algebraiczną krotność $1$,*

*(iii) $\rho(A)$ ma dodatni wektor własny $v$,*

*(iv) $|\lambda| < \rho(A)$ dla każdej innej wartości własnej $\lambda$ macierzy $A$.*

## 29.1 Twierdzenie Frobeniusa–Perrona w wymiarze 2

Rozważmy macierz nieujemną

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad a, b, c, d \geqslant 0.$$

Wielomian charakterystyczny $A$ ma postać

$$p_A(x) = x^2 - (a + d)x + (ad - bc).$$

Ponieważ

$$\Delta := (a + d)^2 - 4(ad - bc) = (a - d)^2 + 4bc \geqslant 0,$$

więc wartości własne $A$ są rzeczywiste oraz

$$\lambda = \frac{a + d + \sqrt{\Delta}}{2}, \quad \lambda' = \frac{a + d - \sqrt{\Delta}}{2}$$

Zauważmy, że $\lambda \geqslant 0$ oraz $\lambda \geqslant \lambda' \geqslant 0$.

Załóżmy, że $A$ jest dodatnia. Wtedy $\lambda > 0$ oraz $\lambda > |\lambda'|$. Pokażemy, że istnieje dodatni wektor własny dla $\lambda$. Niech $x = [x_1, x_2]^T$ będzie wektorem własnym dla $\lambda$. Ponieważ $-x$ jest również wektorem własnym, więc wystarczy pokazać, że $x_1 x_2 > 0$, czyli $x_1$ i $x_2$ są tego samego znaku. Z założenia wynika, że $x_1 \neq 0$ lub $x_2 \neq 0$. Załóżmy przykładowo, że $x_1 \neq 0$. Ponieważ $Ax = \lambda x$, więc

$$ax_1 + bx_2 = \lambda x_1, \quad cx_1 + dx_2 = \lambda x_2.$$

Stąd

$$a + b\frac{x_2}{x_1} = \lambda, \quad \frac{x_2}{x_1}(\lambda - d) = c > 0.$$

Chcemy pokazać, że $\frac{x_2}{x_1} > 0$. W tym celu wystarczy sprawdzić, że $\lambda > a$ lub $\lambda > d$. Tak jest bo $\lambda > \frac{a+d}{2}$.

Przypuśćmy, że $v$ jest dodatnim wektorem własnym dla wartości własnej $\lambda_v$. Pokażemy, że $\lambda_v = \lambda$. Mamy

$$a + b\frac{v_2}{v_1} = \lambda_v, \quad c + d\frac{v_2}{v_1} = \lambda_v \frac{v_2}{v_1}.$$

Stąd

$$\frac{v_2}{v_1} b = \lambda_v - a, \quad \frac{v_2}{v_1}(\lambda_v - d) = c \geqslant 0,$$

więc $\lambda_v \geqslant a$ i $\lambda_v \geqslant d$. Stąd $\lambda_v \geqslant \frac{a+d}{2}$, czyli $\lambda_v = \lambda$.

U

Naturalne jest pytanie czy twierdzenie Frobeniusa–Perrona zachodzi dla macierzy nieujemnych $A \in M_{2\times 2}(\mathbb{R})$. Odpowiedź jest negatywna. Przykładowo, macierz nilpotentna

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

ma promień spektralny 0, a macierz

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

ma dwukrotną wartość własną $\lambda = 1$ i nie ma dodatniego wektora własnego. Z drugiej strony teza twierdzenia zachodzi dla macierzy

$$B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

Geniusz Frobeniusa polegał na dostrzeżeniu różnicy między macierzami $A$ i $B$ w języku tzw. redukowalności. Macierz $A \in M_{n\times n}(\mathbb{R})$ jest *redukowalna*, jeśli dla pewnej macierzy permutacji $P$ macierz $P^T AP$ ma postać blokową

$$P^T AP = \left[\begin{array}{c|c} X & Y \\ \hline 0 & Z \end{array}\right].$$

W wymiarze 2 mamy tylko dwie permutacje: $I$ oraz

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Wynika stąd, że

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

jest *nieredukowalna* wtedy i tylko wtedy, gdy $b \neq 0$ i $c \neq 0$.

## 29.2 Dowód twierdzenia Frobeniusa–Perrona

**Twierdzenie 29.2** (Twierdzenie Frobeniusa–Perrona dla macierzy dodatnich)**.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą dodatnią. Wtedy*

*(1) $\rho(A) > 0$ jest wartością własną $A$ i ma ona dodatni wektor własny,*

*(2) $\rho(A)$ jest jedyną wartością własną na okręgu o promieniu $\rho(A)$,*

*(3) $\rho(A)$ ma geometryczną krotność 1,*

*(4) $\rho(A)$ ma algebraiczną krotność 1,*

*(5) $\rho(A)$ jest jedyną wartością własną mającą nieujemny wektor własny.*

**Lemat 29.1.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą dodatnią. Jeśli $v > w$, to $Av > Aw$. W szczególności, istnieje taki $\epsilon > 0$, że*

$$Av > (1+\epsilon)Aw.$$

*Dowód.* Niech $(A(v-w))_i$ będzie $i$-tą współrzędną wektora $A(v-w)$, czyli

$$(A(v-w))_i = \sum_{j=1}^{n} a_{ij}(v_j - w_j).$$

Wtedy dla $i = 1, \ldots, n$ mamy

$$\begin{aligned}(A(v-w))_i &= \sum_{j=1}^{n} a_{ij}(v_j - w_j)\\ &\geqslant \sum_{j=1}^{n} \min_j\{a_{ij}\}(v_j - w_j)\\ &= \min_j\{a_{ij}\}\sum_{j=1}^{n}(v_j - w_j) > 0,\end{aligned}$$

więc $Av > Aw$. □

**Lemat 29.2.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą dodatnią. Wtedy $\rho(A) > 0$ jest wartością własną $A$ i ma ona dodatni wektor własny.*

*Dowód.* Zauważmy najpierw, że $\operatorname{tr} A > 0$, bo $A$ jest dodatnia. Stąd $\rho(A) > 0$. Istnieje taka wartość własna $\lambda \in \mathbb{C}$, że $|\lambda| = \rho(A)$. Niech $v = [v_1, \ldots, v_n]^T \in \mathbb{C}^n$ będzie odpowiadającym jej wektorem własnym. Rozważmy wektor

$$w = [|v_1|, \ldots, |v_n|]^T \in \mathbb{R}^n.$$

Pokażemy, że $Aw = \rho(A)w$, czyli $w$ jest wektorem własnym dla $\lambda$. Mamy

$$\begin{aligned}(Aw)_i &= \sum_{j=1}^{n} a_{ij}|v_j|\\ &\geqslant |\sum_{j=1}^{n} a_{ij}v_j|\\ &= |(Av)_i|\\ &= |\lambda v_i|\\ &= \rho(A)w_i,\end{aligned}$$

czyli

$$Aw \geqslant \rho(A)w.$$

Przypuśćmy, że $Aw > \rho(A)w$. Wtedy z lematu 29.2 wynika, że istnieje taki $\epsilon > 0$, że

$$A^2w \geqslant (1+\epsilon)\rho(A)Aw.$$

Ponieważ $A$ jest nieujemna, więc również każda jej potęga jest nieujemna, czyli

$$A^{n+1}w \geqslant (1+\epsilon)\rho(A)A^n w.$$

Stąd

$$\begin{aligned} A^n A w = A^{n+1}w & \\ &\geqslant (1+\epsilon)\rho(A)A^n w \\ &\geqslant ((1+\epsilon)\rho(A))^2 A^{n-1}w \\ &\geqslant ((1+\epsilon)\rho(A))^n A w, \end{aligned}$$

więc

$$\|A^n\| \geqslant ((1+\epsilon)\rho(A))^n.$$

Z formuły Gelfanda mamy

$$\rho(A) = \lim_{n\to\infty} \|A^n\|^{1/n} \geqslant (1+\epsilon)\rho(A),$$

co oznacza sprzeczność. Mamy więc

$$Aw = \rho(A)w,$$

czyli $w$ jest nieujemny wektorem własnym odpowiadającym $\rho(A)$. Z dodatniości $A$ wynika, że każda współrzędna wektora $Aw$ jest dodatnia, więc $w$ jest dodatni. □

**Lemat 29.3** (Dominująca wartość własna)**.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą dodatnią. Załóżmy, że $v = [v_1, \ldots, v_n]$ jest dodatnim wektorem własnym odpowiadającym $\lambda_v = \rho(A)$. Wtedy dla dowolnej innej wartości własnej $\lambda \in \mathbb{C}$ macierzy $A$ mamy*

$$|\lambda| < \lambda_v.$$

*Ponadto, jeśli $\lambda \neq \rho(A)$ jest wartością własną $A$, to nie ma ona nieujemnego wektora własnego.*

*Dowód.* Zauważmy, że

$$v = \begin{bmatrix} v_1 \\ \vdots \\ v_2 \end{bmatrix} = \begin{bmatrix} v_1 & 0 & \ldots & 0 \\ 0 & v_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & v_n \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Oznaczmy

$$V = \begin{bmatrix} v_1 & 0 & \ldots & 0 \\ 0 & v_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & v_n \end{bmatrix}, \quad \mathbf{1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Wtedy

$$\lambda_v V\mathbf{1} = \lambda_v v = Av = AV\mathbf{1},$$

czyli

$$V^{-1}AV\mathbf{1} = \lambda_v \mathbf{1}.$$

Zauważmy, że $B = V^{-1}AV$ jest nieujemna i ma te same wartości własne co $A$. Możemy więc założyć, że $v = \mathbf{1}$.

Dla $z = [z_1, \ldots, z_n]^T \in \mathbb{C}^n$ będziemy używać normy

$$\|z\| = \max_{i=1,\ldots,n} |z_i|.$$

Dla dowolnego wektora $z \in \mathbb{C}^n$, $i$-ta współrzędna wektora $Az$ jest równa

$$a_{i1}z_1 + \ldots + a_{in}z_n.$$

Mamy

$$\begin{aligned}|(Az)_i| &= |a_{i1}z_1 + \ldots + a_{in}z_n| \\ &\leqslant a_{i1}|z_1| + \ldots a_{in}|z_n| \\ &\leqslant \|z\|(a_{i1} + \ldots + a_{in}) \\ &= \lambda_v\|z\|.\end{aligned}$$

Stąd,

$$\|Az\| \leqslant \lambda_v\|z\|$$

oraz $\|Az\| = \lambda_v\|z\|$ tylko wtedy, gdy $z_1 = \ldots = z_n$, czyli gdy $z$ jest liniowo zależny z $v$.

Jeśli $v'$ jest wektorem własnym dla $\lambda' \neq \lambda_v$, to $v'$ i $v$ są liniowo niezależne, więc

$$\begin{aligned}\lambda_v\|z'\| &> \|Az'\| \\ &= \|\lambda' z'\| \\ &= |\lambda'|\|z'\|,\end{aligned}$$

czyli $\lambda_v > |\lambda'|$.

Przypuśćmy, że wartość własna $\lambda$ ma nieujemny wektor własny $y$. Niech $v$ będzie dodatnim wektorem własnym macierzy $A^T$ dla wartości własnej $\rho(A)$. Wtedy oczywiście $(y|v) > 0$ oraz

$$\begin{aligned}\rho(A)(y|v) &= (y|\rho(A)v) \\ &= (y|A^Tv) \\ &= (Ay|v) \\ &= \lambda(y|v),\end{aligned}$$

czyli $\lambda = \rho(A)$. □

**Lemat 29.4.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą dodatnią. Wtedy wartość własna $\rho(A)$ ma geometryczną krotność* 1*.*

*Dowód.* Niech $v$ będzie dodatnim wektorem własnym dla $\lambda_v = \rho(A)$. Przypuśćmy, że $v'$ jest liniowo niezależnym z $v$ wektorem własnym dla $\lambda_v$. Możemy założyć, że $v'$ jest rzeczywisty. Wybierzmy $c > 0$ tak, aby $v - c \cdot v' \geqslant 0$ i przynajmniej jedna jego współrzędna była zerowa. Wektor $v - c\cdot v'$ jest niezerowy, bo $v, v'$ są liniowo niezależne. Wtedy

$$A(v - c\cdot v') > 0,$$

więc

$$v - c\cdot v' = \frac{A(v - c\cdot v')}{\rho(A)} > 0,$$

co prowadzi do sprzeczności, bo przynajmniej jedna współrzędna $v - cv'$ jest zerowa. □

**Lemat 29.5.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą dodatnią. Wtedy wartość własna $\rho(A)$ ma algebraiczną krotność* 1*.*

*Dowód.* Niech $v$ będzie dodatnim wektorem własnym $A$ dla $\lambda = \rho(A)$, a $w$ dodatnim wektorem własnym $A^T$ dla $\lambda$. Pokażemy, że

$$w^\perp = \{x \in \mathbb{R}^n : (x|w) = 0\}$$

jest poprzestrzenią niezmienniczą dla $A$. Załóżmy, że $(x|w) = 0$. Wtedy

$$\begin{aligned}(Ax|w) &= (x|A^Tw) \\ &= (x|\lambda w) \\ &= \lambda(x|w) \\ &= 0,\end{aligned}$$

czyli $Ax \in w^{\perp}$.

Zauważmy, że $\dim w^{\perp} = n-1$, oraz $v \notin w^{\perp}$, bo

$$(v|w) = \sum_{i=1}^{n} v_i w_i > 0.$$

Stąd

$$\mathbb{R}^n = \operatorname{span}\{v\} \oplus w^{\perp}.$$

Niech $v_2, \ldots, v_n$ będzie bazą $w^{\perp}$. Rozważmy macierz nieosobliwą $B = \left[ \, v \mid v_2 \mid \ldots \mid v_n \, \right]$. Wtedy podprzestrzenie $B^{-1} \operatorname{span}\{v\} = \operatorname{span}\{e_1\}$ i $B^{-1}w^{\perp} = \operatorname{span}\{e_2, \ldots, e_n\}$ są niezmiennicze dla $B^{-1}AB$. Macierz $B^{-1}AB$ ma postać blokową

$$B^{-1}AB = \left[ \begin{array}{c|c} \rho(A) & 0 \\ \hline 0 & Y \end{array} \right].$$

Przypuśćmy, że $\lambda = \rho(A)$ ma krotność algebraiczną większą od 1. Wtedy $\lambda$ jest również wartością własną $Y$. Istnieje więc wektor własny $v^*$ dla $Y$ odpowiadający $\lambda$ w $B^{-1}w^{\perp} = \operatorname{span}\{e_2, \ldots, e_n\}$. Wtedy $Bv^*$ jest liniowo niezależnym z $v$ wektorem własnym $A$ dla $\lambda$, sprzeczność. $\square$

**Wniosek 29.1** (Formuła Collatza-Wielandta)**.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą dodatnią. Wtedy*

$$\rho(A) = \max_{x\in N} f(x),$$

*gdzie*

$$f(x) = \min_{1\leqslant i\leqslant n:\, x_i\neq 0} \frac{(Ax)_i}{x_i} \quad \textit{oraz} \quad N = \{x \neq 0 : x \geqslant 0\}.$$

*Dowód.* Zauważmy najpierw, że $f$ jest dobrze określona. Jeśli $p > 0$ jest dodatnim wektorem własnym dla $\rho(A)$, czyli $Ap = \rho(A)\cdot p$, to $f(p) = \rho(A)$. Wystarczy pokazać, że $f(x) \leqslant \rho(A)$ dla $x \in N$. Niech $q > 0$ bedzie wektorem własnym dla wartości własnej $\rho(A)$ i macierzy $A^T$. Tranponując równość $A^T q = \rho(A)\cdot q$ otrzymujemy, że

$$q^T A = \rho(A)\cdot q^T.$$

Dla $x \in N$ mamy

$$0 \leqslant f(x)\cdot x \leqslant Ax \quad \text{oraz} \quad q^T x > 0.$$

Stąd $f(x)\cdot x \leqslant Ax$ implikuje, że

$$f(x) q^T x \leqslant q^T A x = \rho(A) q^T x,$$

więc $f(x) \leqslant \rho(A)$ dla $x \in N$. $\square$

**Twierdzenie 29.3** (Nieujemna para własna)**.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą nieujemną. Wtedy*

*(1) $\rho(A) \in \sigma(A)$,*

*(2) wartość własna $\rho(A)$ ma nieujemny wektor własny.*

*Dowód.* Niech $E \in M_{n\times n}(\mathbb{R})$ będzie macierzą, której wszystkie wyrazy są równe 1. Rozważmy ciąg macierzy dodatnich

$$A_k = A + (1/k)\cdot E.$$

Z twierdzenia 29.2 wynika, że $r_k = \rho(A_k) > 0$ jest dodatnią wartością własną $A_k$, mającą dodatni jednostkowy wektor własny $p_k > 0$. Ponieważ $\|p_k\| = 1$, więc przechodząc do podciągu możemy założyć, że $p_k \to z$. Wtedy $z \geqslant 0$ i $z \neq 0$, bo $\|z\| = 1$. Ponieważ

$$A_1 > A_2 > \ldots > A,$$

więc z formuły Gelfanda (twierdzenie 27.2) wynika, że

$$r_1 \geqslant r_2 \geqslant \ldots \geqslant r.$$

Niech $r^* = \lim_{k\to\infty} r_k$. Wtedy $r^* \geqslant r$. Pokażemy, że $r^* = r$. Wystarczy pokazać, że $r^* \in \sigma(A)$, bo wtedy $r^* \leqslant r = \rho(A)$. Pokażemy, że $Az = r^* z$, czyli $r^*$ jest wartością własną z wektorem własnym $z$. Ponieważ $\lim_{k\to\infty} A_k = A$, więc

$$Az = \lim_{k\to\infty} A_k p_k = \lim_{k\to\infty} r_k \cdot p_k = r^* z.$$

□

**Twierdzenie 29.4** (Perrona–Frobeniusa dla macierzy nieujemnych)**.** *Niech $A \in M_{n\times n}(\mathbb{R})$ będzie taką macierzą nieujemną, że $B := (I + A)^k$ jest dodatnia dla pewnego $k \in \mathbb{N}$. Wtedy*

*(i)* $r = \rho(A) \in \sigma(A)$ *i* $r > 0$,

*(ii) istnieje taki dodatnie wektor własny* $x > 0$, *że* $Ax = r \cdot x$,

*(iii)* $r$ *ma algebraiczną krotność* 1,

*(iv) jeśli* $p > 0$ *jest jedynym dodatnim jednostkowym wektorem własnym dla* $r$, *to nieujemne wektory własne* $A$ *są postaci* $t \cdot p$ *dla* $t > 0$.

*Dowód.* Z twierdzenia 29.3 $r = \rho(A) \in \sigma(A)$. Pokażemy, że $r$ ma algebraiczną krotność 1. Zauważmy, że $\lambda \in \sigma(A)$ wtedy i tylko wtedy, gdy $(1 + \lambda)^k \in \sigma(B)$ i krotność algebraiczna $\lambda$ jako wartości własnej $A$ jest równa krotności algebraicznej $(1 + \lambda)^k$ jako wartości własnej $B$. Ponieważ $B$ jest dodatnia, więc z twierdzenia 29.2 wartość własna $\rho(B)$ macierzy $B$ ma krotność algebraiczną 1. Wystarczy więc pokazać, że $\rho(B) = (1 + r)^k$. Wynika to z równości

$$\begin{aligned}\rho(B) &= \max_{\lambda\in\sigma(A)} |(1+\lambda)|^k \\ &= \Big(\max_{\lambda\in\sigma(A)} |1+\lambda|\Big)^k \\ &= (1+r)^k.\end{aligned}$$

Pokażemy, że wartość własna $r$ ma dodatni wektor własny. Z twierdzenia 29.3 wynika, że $A$ ma nieujemny wektor własny $x \geqslant 0$ dla $r$. Wtedy $x$ jest wektorem własnym $B$ dla $(1 + r)^k$. Z twierdzenia 29.2 $\frac{x}{\|x\|}$ jest dodatnie, więc $x$ jest dodatni.

Zakończymy dowód pokazując, że $r > 0$. W przeciwnym razie, $Ax = 0$ co jest niemożliwe, bo $A \geqslant 0$ i $x > 0$. □

U

W kontekście twierdzenia 29.4 interesujące jest pytanie, dla jakich macierzy nieujemnych $A$ macierz $(I + A)^k$ jest dodatnia dla pewnego $k \in \mathbb{N}$. Oczywiście będzie tak, jeśli $A^k > 0$ jest dodatnia dla pewnego $k$. Takie macierze nieujemne nazywamy *regularnymi.*

Również jeśli $A^l = [a_{ij}^{(l)}]$ i dla każdej pary $(i, j)$ istnieje takie $l \geqslant 1$, że $a_{ij}^{(n)} > 0$, to $(I + A)^k > 0$ dla pewnego $k$. Macierze spełniające ten warunek nazywa się macierzami *nieredukowalnymi.* Można pokazać, że macierz nieujemna jest nieredukowalna wtedy i tylko wtedy, gdy nie istnieje taka macierz permutacji $P \in M_{n\times n}(\mathbb{R})$, że

$$P^T AP = \begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix},$$

gdzie $X$ i $Z$ są macierzami kwadratowymi. W macierzy $P^T AP$ wiersze i kolumny macierzy $A$ są permutowane w taki sam sposób.

# Rozdział 30

# Macierze stochastyczne

| |
|---|
| Słowa kluczowe rozdziału |
| macierz stochastyczna ♢ macierz dodatnia i regularna ♢ wiodąca wartość własna 1 |

**Definicja 30.1.** Macierz stochastyczna

Mówimy, że macierz $A = [a_{ij}] \in M_{n\times n}(\mathbb{R})$ o wyrazach nieujemnych $a_{ij} \geqslant 0$ jest *stochastyczna*, gdy

$$a_{1i} + \ldots + a_{ni} = 1, \quad i = 1, \ldots, n,$$

czyli suma wyrazów w każdej kolumnie jest równa 1.

**Lemat 30.1.** *Jeśli $A \in M_{n\times n}(\mathbb{R})$ jest stochastyczna, to 1 jest wartością własną $A$.*

*Dowód.* Zauważmy, że dla wektora $\mathbf{1} = [1, \ldots, 1]^T$ mamy

$$\begin{aligned} A^T\mathbf{1} &= [(a_1^T|\mathbf{1}), \ldots, (a_n^T|\mathbf{1})]^T \\ &= [a_{11} + \ldots + a_{n1}, \ldots, a_{1n} + \ldots + a_{nn}]^T \\ &= \mathbf{1}, \end{aligned}$$

czyli 1 jest wartością własną $A^T$, więc również $A$. □

**Definicja 30.2.** Macierz dodatnia

Macierz $A = [a_{ij}] \in M_{n\times n}(\mathbb{R})$ nazywamy *dodatnią*, gdy $a_{ij} > 0$. Powiemy, że $A$ jest *regularna*, gdy pewna jej potęga jest dodatnia.

**Lemat 30.2.** *Załóżmy, że liczby zespolone $z_1, \ldots, z_n$ mają moduł 1 oraz dodatnie liczby rzeczywiste $x_i > 0$ są takie, że $x_1 + \ldots + x_n = 1$. Jeśli $|x_1z_1 + \ldots + x_nz_n| = 1$, to $z_1 = \ldots = z_n$.*

**Twierdzenie 30.1.** *Niech $A$ będzie regularną macierzą stochastyczną. Jeśli $\lambda \neq 1$ jest wartością własną $A$, to $|\lambda| < 1$. W szczególności, $\rho(A) = 1$.*

*Dowód.* Załóżmy, że $z = [z_1, \ldots, z_n]^T \in \mathbb{C}^n$ jest wektorem własnym $A$ odpowiadającym wartości własnej $\lambda \in \mathbb{C}$. Niech $z_k$ będzie współrzędną wektora $z$ o największym module $m = |z_k| > 0$. Ponieważ $Az = \lambda \cdot z$, czyli

$$\lambda z_k = a_{k1}z_1 + \ldots + a_{kn}z_n,$$

więc

$$\begin{aligned} |\lambda|m = |\lambda z_k| &= |a_{k1}z_1 + \ldots + a_{kn}z_n| \\ &\leqslant a_{k1}|z_1| + \ldots + a_{kn}|z_n| \\ &\leqslant \underbrace{(a_{k1} + \ldots + a_{kn})}_{=1} m \\ &= m. \end{aligned}$$

W efekcie $|\lambda|m \leqslant m$, czyli $|\lambda| \leqslant 1$.

Pokażemy teraz, że jeśli $|\lambda| = 1$, to $\lambda = 1$. Załóżmy, najpierw, że $A$ jest dodatnia. Z powyższych nierówności wynika, że dla $|\lambda| = 1$ mamy

$$m \leqslant a_{k1}|z_1| + \ldots + a_{kn}|z_n| \leqslant (a_{k1} + \ldots + a_{kn})m = m,$$

czyli

$$a_{k1}|z_1| + \ldots + a_{kn}|z_n| = (a_{k1} + \ldots + a_{kn})m,$$

więc

$$a_{k1}(m - |z_1|) + \ldots + a_{kn}(m - |z_n|) = 0.$$

Ponieważ $A$ jest dodatnia oraz $(m - |z_i|) \geqslant 0$ dla $i = 1, \ldots, n$, więc $m = |z_i|$ dla $i = 1, \ldots, n$, czyli $z_i$ mają równe moduły. Ponadto z lematu 30.2 wynika, że równość

$$|a_{k1}z_1 + \ldots + a_{kn}z_n| = a_{k1}|z_1| + \ldots + a_{kn}|z_n|$$

zachodzi tylko wtedy, gdy liczby zespolone $z_1, \ldots, z_n$ mają ten sam argument, więc $z_1 = \ldots = z_n$. Wynika stąd, że $z \in \operatorname{span}\{[1, \ldots, 1]^T\}$, więc $\lambda = 1$. Załóżmy teraz, że $A$ jest regularna i niech $A^k$ będzie dodatnia. Wtedy $A^{k+1}$ jest również dodatnia. Ponieważ $\lambda^k$ i $\lambda^{k+1}$ są wartościami własnymi $A^k$ oraz $A^{k+1}$ i $|\lambda^k| = |\lambda^{k+1}| = 1$, więc $\lambda^k = \lambda^{k+1} = 1$, czyli $\lambda^k(\lambda - 1) = 0$, co implikuje, że $\lambda = 1$, bo $|\lambda| = 1$. □

**Twierdzenie 30.2.** *Załóżmy, że $A \in M_{n\times n}(\mathbb{R})$ jest diagonalizowalna oraz wartości własne $A$ spełniają warunek*

$$1 = \lambda_1 > |\lambda_2| \geqslant \ldots \geqslant |\lambda_n|.$$

*Niech $v_1, \ldots, v_n$ będzie odpowiadającą bazą wektorów własnych. Jeśli $v = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$ jest takim wektorem, że $x_1 \neq 0$, to*

$$\lim_{k\to\infty} A^k v = x_1 \cdot v_1.$$

*Dowód.* Zauważmy, że

$$A^k v = x_1 \cdot v_1 + x_2\lambda_2^k \cdot v_2 + \ldots + x_n\lambda_n^k \cdot v_n \to x_1 \cdot v_1.$$

□

**Przykład 30.1.** Rozważmy Rynek Główny w Krakowie z czterema narożnikami $v_1$, $v_2$, $v_3$, $v_4$. Będziemy spacerować po rynku od narożnika do narożnika poruszając się po bokach rynku. Nie wolno nam iść przez Sukiennice. Mamy więc krawędzie

$$v_1 \longleftrightarrow v_2, \quad v_1 \longleftrightarrow v_4, \quad v_2 \longleftrightarrow v_3, \quad v_3 \longleftrightarrow v_4.$$

Będąc w narożniku, za każdym razem rzucamy monetą, aby zdecydować, do którego z sąsiadujących narożników się udamy. Definiujemy macierz $A$ przez

$$a_{ij} = \frac{1}{2},$$

jeśli istnieje krawędź łącząca $i$ z $j$ oraz

$$a_{ij} = 0,$$

jeśli takiej krawędzi w grafie nie ma. Otrzymujemy macierz stochastyczną

$$\begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{bmatrix}.$$

**Lemat 30.3.** *Iloczyn $AB$ macierzy stochastycznych $A$ i $B$ jest macierzą stochastyczną. W szczególności, $A^n$ jest stochastyczna dla każdego $n \geqslant 1$.*

*Dowód.* Załóżmy, że macierze $A, B \in M_{n\times n}(\mathbb{R})$ są stochastyczne. Wtedy

$$\begin{aligned}(AB)_{i1} + \ldots + (AB)_{in} &= \sum_{j=1}^{n}\sum_{k=1}^{n} a_{ik}b_{kj} \\ &= a_{i1}(b_{11} + \ldots + b_{1n}) + \ldots + a_{in}(b_{n1} + \ldots + b_{nn}) \\ &= a_{i1} + \ldots + a_{in} = 1.\end{aligned}$$

$\square$

U

- Macierz stochastyczna nie musi być nieosobliwa. Przykładowo, macierze

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$$

  są stochastyczne.
- Macierz stochastyczna może mieć ujemne wartości własne i wyznacznik. Przykładowo,

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

- Jeśli $A \in M_{2\times 2}$ jest dodatnią macierzą stochastyczną, to 1 jest jednokrotną wartością własną i istnieje wartość własna mniejsza od 1. Rzeczywiście, $A$ ma postać

$$A = \begin{bmatrix} a & 1-a \\ 1-b & b \end{bmatrix}, \quad 0 < a, b < 1.$$

  Wtedy suma wartości własnych jest równa $a + b < 2$. Ponieważ jedną z wartości własnych jest 1, więc druga wartość własna jest mniejsza od 1.
- Macierz stochastyczna może mieć zespolone wartości własne i może być ortogonalna. Przykładowo,

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

- Rozważmy macierz

$$A = \begin{bmatrix} 5/12 & 1/4 & 1/3 \\ 5/12 & 1/4 & 1/3 \\ 1/6 & 1/2 & 1/3 \end{bmatrix}.$$

  Jest ona „podwójnie stochastyczna" tzn. $A^T$ jest również stochastyczna. Jej wielomianem charakterystycznym jest $p_A(x) = x^2 - x^3$, więc 0 jest dwukrotną wartością własną. Wymiar jądra $A$ jest równy 1. Macierz $A$ nie jest więc digonalizowalna.

Przyglądnijmy się macierzom stochastycznym w wymiarze 2. Macierz stochastyczna $A \in M_{2\times 2}(\mathbb{R})$ ma postać

$$A = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}, \quad a, b \in [0, 1].$$

Wiemy, że 1 jest wartością własną $A$ z wektorem własnym $[1, 1]^T$. Ponieważ $\operatorname{tr} A = 2 - a - b$, więc druga wartość własna jest równa $1 - a - b$. Załóżmy, że $A$ jest dodatnia, czyli

$$0 < a, b < 1.$$

Sprawdzamy łatwo, że $[a, -b]^T$ jest jej wektorem własnym. Wtedy dla

$$S = \begin{bmatrix} 1 & a \\ 1 & -b \end{bmatrix}$$

mamy

$$S^{-1}AS = \begin{bmatrix} 1 & 0 \\ 0 & 1-a-b \end{bmatrix}.$$

Wiemy już, że wtedy $|1-a-b|<1$. Wynika stąd, że

$$\begin{aligned} A^n &= S \begin{bmatrix} 1 & 0 \\ 0 & 1-a-b \end{bmatrix}^n S^{-1} \\ &= S \begin{bmatrix} 1 & 0 \\ 0 & (1-a-b)^n \end{bmatrix} S^{-1} \\ &\to S \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} S^{-1}. \end{aligned}$$

Ponieważ

$$S^{-1} = \frac{1}{a+b} \begin{bmatrix} b & a \\ 1 & -1 \end{bmatrix},$$

więc

$$A^n \to S \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} S^{-1} = \frac{1}{a+b} \begin{bmatrix} b & a \\ b & a \end{bmatrix}.$$

U Zauważmy, że dla macierzy stochastycznej

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

mamy

$$A^{2k} = I, \quad A^{2k+1} = A,$$

więc nie każda macierz stochastyczna ma powyższą własność zbieżności ciągu $A^n$.

U Można łatwo udowodnić indukcyjnie, że dla dodatniej macierzy stochastycznej

$$A = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$$

zachodzi wzór

$$A^n = \frac{1}{a+b} \begin{bmatrix} b & a \\ b & a \end{bmatrix} + \frac{(1-a-b)^n}{a+b} \begin{bmatrix} a & -a \\ -b & b \end{bmatrix}.$$

## 30.1 Macierze bistochastyczne

**Definicja 30.3.** Macierz bistochastyczna

Macierz nieujemną $A \in M_{n\times n}(\mathbb{R})$ nazywamy *bistochastyczną*, jeśli suma wyrazów w każdym jej wierszu i w każdej kolumnie jest równa 1. Przez $\Omega_n \subset M_{n\times n}(\mathbb{R})$ oznaczamy zbiór wszystkich macierzy bistochastycznych.

U Z definicji wynika, że macierz nieujemna $A \in M_{n\times n}(\mathbb{R})$ jest bistochastyczna wtedy i tylko wtedy, gdy

$$A[1,\ldots,1]^T = [1,\ldots,1]^T, \quad A^T[1,\ldots,1]^T = [1,\ldots,1]^T,$$

czyli wektor $\mathbf{1} := [1,\ldots,1]^T$ jest wektorem własnym $A$ i $A^T$ z wartością własną $\lambda = 1$.

**Przykład 30.2.** Dwuwymiarowa macierz bistochastyczna $A \in \Omega_2$ ma postać

$$\begin{bmatrix} a & 1-a \\ 1-a & a \end{bmatrix}, \quad 0 \leqslant a \leqslant 1.$$

**Lemat 30.4.** *Jeśli $A, B \in \Omega_n$ i $\lambda \in [0,1]$, to $\lambda \cdot A + (1-\lambda)\cdot B \in \Omega_n$.*

*Dowód.* Oczywiście macierz $\lambda \cdot A + (1-\lambda) \cdot B \in \Omega_n$ jest nieujemna. Ponadto

$$\begin{aligned}(\lambda \cdot A + (1-\lambda) \cdot B)\mathbf{1} &= \lambda \cdot A\mathbf{1} + (1-\lambda) \cdot B\mathbf{1} \\ &= \lambda \cdot \mathbf{1} + (1-\lambda) \cdot \mathbf{1} \\ &= \mathbf{1}.\end{aligned}$$

Analogicznie sprawdzamy, że

$$(\lambda \cdot A^T + (1-\lambda) \cdot B^T)\mathbf{1} = \mathbf{1},$$

czyli $\lambda \cdot A + (1-\lambda) \cdot B \in \Omega_n$. □

**Wniosek 30.1.** *Zbiór $\Omega_n$ traktowany jako podzbiór $\mathbb{R}^{n^2}$ jest ograniczony, domknięty i wypukły.*

Punkt $x_0$ należący do zbioru wypukłego $S$ jest nazywany *punktem ekstremalnym* zbioru $S$, gdy $S \setminus \{x_0\}$ jest również wypukły.

**Lemat 30.5.** *Punkt $x_0 \in S$ nie jest punktem ekstremalnym zbioru wypukłego $S$ wtedy i tylko wtedy, gdy istnieją takie różne punkty $u, w \in S$, że $x_0 = \frac{1}{2}u + \frac{1}{2}w$.*

*Dowód.* Załóżmy, że $x_0$ nie jest punktem ekstremalnym, czyli $S \setminus \{x_0\}$ nie jest wypukły. Ponieważ zbiór pusty i jednoelementowy są wypukłe, więc istnieją takie dwa różne punkty $u, w \in S \setminus \{x_0\}$, że $\lambda u + (1-\lambda)w \notin S \setminus \{x_0\}$ dla pewnego $\lambda \in (0,1)$. Wtedy

$$x_0 = \lambda u + (1-\lambda)w.$$

Ze względu na symetrię możemy założyć, że $1/2 < \lambda < 1$. Wtedy $\lambda = \frac{\alpha+1}{2}$ dla pewnego $0 < \alpha < 1$ oraz

$$x_0 = \frac{1}{2}u + \frac{1}{2}(\alpha u + (1-\alpha)w).$$

Z drugiej strony, jeśli $x_0 = \frac{1}{2}u + \frac{1}{2}w$ dla różnych punktów $u, w \in S$, to $S \setminus \{x_0\}$ nie jest wypukły, czyli $x_0$ nie jest punktem ekstremalnym. □

**Lemat 30.6.** *Niech $P \in \Omega_n$ będzie macierzą permutacji. Wtedy $P$ jest punktem ekstremalnym zbioru $\Omega_n$.*

*Dowód.* Przypuśćmy, że macierz permutacji $P = [p_{ij}]$ nie jest punktem ekstremalnym $\Omega_n$. Z lematu 30.5

$$P = \frac{A+B}{2}, \quad A \neq B, \quad A, B \in \Omega_n.$$

Ponieważ $p_{ij} \in \{0,1\}$ oraz $a_{ij}, b_{ij} \in [0,1]$, więc z równości $p_{ij} = \frac{a_{ij}+b_{ij}}{2}$ wynika, że

$$a_{ij} = b_{ij} = \begin{cases} 0, & \text{gdy } p_{ij} = 0 \\ 1, & \text{gdy } p_{ij} = 1. \end{cases}$$

Stąd $A = B$, co prowadzi do sprzeczności. □

**Definicja 30.4.** **Permanent macierzy**

*Permanent* $\operatorname{per}(A)$ macierzy kwadratowej $A \in M_{n \times n}(\mathbb{R})$ jest zdefiniowany wzorem

$$\operatorname{per}(A) = \sum_{\sigma \in S_n} \underbrace{\prod_{i=1}^{n} a_{i\sigma(i)}}_{\text{iloczyn diagonalny}}.$$

U [Rozwinięcie Laplace'a] Podobnie jak w przypadku wyznacznika, zachodzi wzór na rozwinięcie Laplace'a względem $i$-tego wiersza:

$$\operatorname{per}(A) = \sum_{j=1}^{n} a_{ij} \operatorname{per} A(i,j).$$

*Podmacierzą* macierzy $A$ nazywamy dowolną macierz, powstałą z $A$ przez wykreślenie pewnej liczby wierszy i pewnej liczby kolumn.

**Twierdzenie 30.3** (Frobenius-König). *Dla macierzy nieujemnej $A \in M_{n\times n}(\mathbb{R})$ następujące warunki są równoważne:*

*(i)* $\operatorname{per} A = 0$,

*(ii) $A$ ma zerową podmacierz rozmiaru $r \times s$, gdzie liczby $r$, $s$ spełniają warunek $r+s=n+1$.*

*Dowód.* Zakładamy, że zachodzi warunek (i). Stosujemy indukcję względem $n$. Dla $n=1$ teza zachodzi. Niech $A \in M_{n\times n}(\mathbb{R})$. Możemy założyć, że $A \neq 0$. Istnieje więc niezerowy wyraz $a_{ij} \neq 0$. Z rozwinięcia Laplace'a wynika, że $\operatorname{per} A(i,j) = 0$. Z założenia indukcyjnego, macierz $A(i,j)$, więc również $A$ ma zerową podmacierz rozmiaru $u \times v$, gdzie liczby $u$, $v$ spełniają równość $u+v=n$. Istnieją takie macierze permutacji $P_1, P_2 \in M_{n\times n}(\mathbb{R})$, że

$$P_1 A P_2 = \left[\begin{array}{c|c} B & 0 \\ \hline C & D \end{array}\right],$$

gdzie $B \in M_{u\times u}(\mathbb{R})$ i $D \in M_{(n-u)\times(n-u)}(\mathbb{R})$. Ponieważ $\operatorname{per}(A) = 0$, więc

$$\operatorname{per}(B) = 0 \quad \text{lub} \quad \operatorname{per}(D) = 0.$$

Niech przykładowo $\operatorname{per}(B) = 0$. Z założenia indukcyjnego $B$ ma zerową podmacierz rozmiaru $p \times q$ gdzie $p+q = u+1$. Bez straty ogólności możemy założyć, że składa się ona z pierwszych $p$ wierszy i ostatnich $q$ kolumn macierzy $B$. Wtedy $A$ ma zerową podmacierz rozmiaru $p \times (q+n-u)$ oraz

$$p+q+n-u = u+1+n-u = n+1.$$

Załóżmy, że zachodzi warunek (ii). Możemy przyjąć, że pierwszych $r$ wierszy i pierwszych $s$ kolumn tworzy podmacierz zerową z $r+s=n+1$. Dla ustalonej permutacji $\sigma \in S_n$ mamy pokazać, że $a_{i\sigma(i)} = 0$ dla pewnego $i$. Przypuśćmy, że $a_{i\sigma(i)} = 0$ dla każdego $i$. Wtedy

$$\{\sigma(1), \dots, \sigma(r)\} \cap \{1, \dots, s\} = \emptyset,$$

co prowadzi do sprzeczności, gdyż $r+s=n+1$. □

**Lemat 30.7.** *Jeśli $A \in \Omega_n$ jest macierzą bistochastyczną, to* $\operatorname{per}(A) > 0$.

*Dowód.* Przypuśćmy, że $\operatorname{per}(A) = 0$. Z twierdzenia 30.3 wynika, że $A$ ma zerową podmacierz rozmiaru $r \times s$ z $r+s=n+1$. Możemy założyć, że

$$\left[\begin{array}{c|c} B & 0 \\ \hline C & D \end{array}\right], \quad 0 \in M_{r\times s}(\mathbb{R}), \quad B \in M_{r\times(n-s)}(\mathbb{R}), \quad D \in M_{(n-r)\times s}(\mathbb{R}).$$

Ponieważ $A$ jest bistochastyczna, więc suma wyrazów w wierszach macierzy $B$ jest równa 1. Stąd suma wszystkich wyrazów macierzy $B$ jest równa $r$. Z drugiej strony suma wszystkich wyrazów macierzy $B$ i $C$ jest równa liczbie kolumn, czyli $n-s$. Stąd $r \leqslant n-s$, co prowadzi do sprzeczności. □

**Twierdzenie 30.4** (Birkhoff-von Neumann). *Jeśli $A \in \Omega_n$ jest macierzą bistochastyczną, to*

$$A = \lambda_1 \cdot P_1 + \ldots + \lambda_m \cdot P_m,$$

*gdzie $P_i \in M_{n\times n}(\mathbb{R})$ są macierzami permutacji, $\lambda_i \in [0,1]$ oraz $\lambda_1 + \ldots + \lambda_m = 1$.*

*Dowód.* Jeśli $A$ jest macierzą permutacji, to nie ma czego dowodzić. Z lematu 30.7 wynika, że $\operatorname{per}(A) > 0$. Istnieje taka permutacja $\sigma \in S_n$, że

$$\prod_{i=1}^{n} a_{i\sigma(i)} > 0.$$

Niech $P_1$ będzie macierzą permutacji, odpowiadającą permutacji $\sigma$. Dokładniej,

$$(P_1)_{ij} = \begin{cases} 1, & \text{gdy } (i,j) \in \{(1,\sigma(1)), \dots, (n,\sigma(n))\} \\ 0, & \text{w przeciwnym razie .} \end{cases}$$

Niech

$$\lambda_1 = a_{i\sigma(i)} := \min\{a_{1\sigma(1)}, \dots, a_{n\sigma(n)}\} > 0.$$

Oczywiście, $\lambda_1 < 1$, bo $P_1$ nie jest macierzą permutacji. Wtedy

$$A = \lambda_1 \cdot P_1 + (1-\lambda_1)\Big(\frac{1}{1-\lambda_1}(A - \lambda_1 P_1)\Big).$$

Macierz

$$B = \frac{1}{1-\lambda_1}(A - \lambda_1 P_1)$$

jest bistochastyczna i ma przynajmniej o 1 zerowy wyraz więcej niż macierz $A$. Jeśli $B$ jest macierzą permutacji, to

$$A = \lambda_1 \cdot P_1 + (1-\lambda_1) \cdot B,$$

więc teza zachodzi. Jeśli $B$ nie jest macierzą permutacji, to

$$B = \mu_2 \cdot P_2 + (1-\mu_2) \cdot C,$$

gdzie $C$ jest bistochastyczna i ma więcej wyrazów zerowych niż $B$. Jeśli $C$ jest macierzą permutacji, to

$$A = \lambda_1 \cdot P_1 + \lambda_2 \cdot P_2 + \lambda_3 \cdot P_3, \quad C = P_3$$

dla

$$\lambda_2 = (1-\lambda_1)\mu_2, \quad \lambda_3 = (1-\lambda_1)(1-\mu_2).$$

Wtedy dowód jest zakończony. Jeśli $C$ nie jest macierzą permutacji, to możemy ten proces kontynuować i skończymy po co najwyżej $n^2 - n$ krokach. Dowodzi to, że $A$ jest kombinacją wypukłą co najwyżej $n^2 - n + 1$ permutacji. $\square$

# Rozdział 31

# Odwzorowania dwuliniowe

| |
|---|
| Słowa kluczowe rozdziału<br>odwzorowanie dwuliniowe symetryczne i antysymetryczne ♢ niezdegenerowanie ♢ forma kwadratowa ♢ forma symplektyczna ♢ macierz symplektyczna ♢ macierze hamiltonowskie ♢ twierdzenie Darboux |

W tym rozdziale zakładamy, że $V$ jest przestrzenią wektorową wymiaru $n \in \mathbb{N}$ nad ciałem $\mathbb{R}$.

**Definicja 31.1.** **Odwzorowanie dwuliniowe**
Odwzorowanie $g : V \times V \longrightarrow \mathbb{R}$ nazywamy dwuliniowym, jeśli dla $a, b \in \mathbb{R}$ i $v, u, w \in V$ spełnione są warunki

- $g(a \cdot v + b \cdot u, w) = ag(v, w) + bg(u, w)$,
- $g(w, a \cdot v + b \cdot u) = ag(w, v) + bg(w, u)$.

**Definicja 31.2.** **Odwzorowanie symetryczne i antysymetryczne**
Odwzorowanie dwuliniowe $g : V \times V \longrightarrow \mathbb{R}$ nazywamy

- *symetrycznym*, gdy $g(u, v) = g(v, u)$ dla $u, v \in V$,
- *antysymetrycznym*, gdy $g(u, v) = -g(v, u)$ dla $u, v \in V$.

Dwuliniowe odwzorowania antysymetryczne nazywamy czasami *2-formami*. Zauważmy, że wówczas $g(u, u) = 0$.

**Przykład 31.1.** Niech $A \in M_{n\times n}(\mathbb{R})$. Wtedy odwzorowanie

$$g : \mathbb{R}^n \times \mathbb{R}^n \ni (v, u) \longmapsto (Av|u) \in \mathbb{R}$$

jest dwuliniowe.

**Przykład 31.2.** Iloczyn skalarny $(\cdot|\cdot) : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ jest odwzorowaniem dwuliniowym symetrycznym. Bardziej ogólnie, jeśli macierz $A \in M_{n\times n}(\mathbb{R})$ jest symetryczna, to odwzorowanie

$$g : \mathbb{R}^n \times \mathbb{R}^n \ni (v, u) \longmapsto (Av|u) \in \mathbb{R}$$

jest dwuliniowym odwzorowaniem symetrycznym. Jest ono symetryczne, bo

$$\begin{aligned} g(v, u) &= (Av|u) \\ &= (v|A^T u) \\ &= (v|Au) \\ &= (Au|v) \\ &= g(u, v). \end{aligned}$$

**Przykład 31.3.** Odwzorowanie

$$g : \mathbb{R}^2 \times \mathbb{R}^2 \ni (u, v) \longmapsto \det \left[ \begin{array}{c|c} u & v \end{array} \right] \in \mathbb{R}$$

jest dwuliniowe i antysymetryczne. Zauważmy, że jeśli

$$u = [u_1, u_2]^T, \quad v = [v_1, v_2]^T,$$

to

$$g(u, v) = u_1 v_2 - v_1 u_2.$$

Odwzorowanie $g$ możemy zapisać w postaci

$$g(u, v) = (Ju|v),$$

gdzie $J$ jest standardową macierzą symplektyczną:

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad Ju = [-u_2, u_1]^T.$$

Zauważmy, że $J^T = -J$. Możemy taką konstrukcję uogólnić na dowolny wymiar parzysty. Dla

$$J_{2n} = \operatorname{diag} \underbrace{(J, \ldots, J)}_{n}$$

odwzorowanie

$$g : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \ni (u, v) \longmapsto (J_{2n} u|v) \in \mathbb{R}$$

jest dwuliniowe antysymetryczne.

Definiujemy przestrzeń wektorową

$$\mathcal{L}_2(V) = \Big\{ g : V \times V \longrightarrow \mathbb{R} : g \quad \text{jest dwuliniowe} \Big\}$$

i jej dwie podprzestrzenie wektorowe

$$\mathcal{L}_2^s(V) = \Big\{ g \in \mathcal{L}_2(V) : g \quad \text{jest symetryczne} \Big\},$$

$$\mathcal{L}_2^a(V) = \Big\{ g \in \mathcal{L}_2(V) : g \quad \text{jest antysymetryczne} \Big\}.$$

Rozważmy $g \in \mathcal{L}_2(V)$. Ustalmy bazę $v_1, \ldots, v_n$ dla $V$. Jeśli $v = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$ i $u = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$, to

$$\begin{aligned} g(v, u) &= g\left( \sum_{i=1}^n x_i \cdot v_i, \sum_{j=1}^n y_j \cdot v_j \right) \\ &= \sum_{i,j=1}^n x_i y_j g(v_i, v_j). \end{aligned}$$

Definiujemy macierz

$$A_g = A = \begin{bmatrix} g(v_1, v_1) & \ldots & g(v_1, v_n) \\ g(v_2, v_1) & \ldots & g(v_2, v_n) \\ \vdots & \ddots & \vdots \\ g(v_n, v_1) & \ldots & g(v_n, v_n) \end{bmatrix}.$$

Wtedy

$$g(v, u) = (A[x_1, \ldots, x_n]^T, [y_1, \ldots, y_n]). \tag{31.1}$$

Macierz $A$ nazywamy *macierzą odwzorowania dwuliniowego* $g$ w bazie $v_1, \ldots, v_n$ dla $V$.

**Wniosek 31.1.** *Niech $v_1, \ldots, v_n$ będzie bazą $V$. Odwzorowanie*

$$T : \mathcal{L}_2(V) \ni g \longmapsto A_g \in M_{n\times n}(\mathbb{R})$$

*jest izomorfizmem. W szczególności,* $\dim \mathcal{L}_2(V) = n^2$.

*Dowód.* Sprawdzamy łatwo, że $T$ jest odwzorowaniem liniowym. Jest ono monomorfizmem, bo jeśli $T(g) = A_g = 0$, to $g = 0$. Dla dowodu epimorficzności zauważmy, że dla macierzy $A \in M_{n\times n}(\mathbb{R})$ wzór (31.1) definiuje takie $g \in \mathcal{L}_2(V)$, że $T(g) = A$. □

U

Jeśli $\mathrm{Sym}_n$ i $\mathrm{ASym}_n$ są podprzestrzeniami $M_{n\times n}(\mathbb{R})$ macierzy symetrycznych i antysymetrycznych, to

$$T(\mathcal{L}_2^s(V)) = \mathrm{Sym}_n, \quad T(\mathcal{L}_2^a(V)) = \mathrm{ASym}_n\,.$$

Wynika stąd, że

$$\dim \mathcal{L}_2^s(V) = \frac{n(n+1)}{2}, \quad \dim \mathcal{L}_2^a(V) = \frac{n(n-1)}{2}.$$

Ponadto, ponieważ $A = \frac{A+A^T}{2} + \frac{A-A^T}{2}$ oraz $\mathrm{Sym}_n \cap \mathrm{ASym}_n = \{0\}$, czyli

$$M_{n\times n}(\mathbb{R}) = \mathrm{Sym}_n \oplus \mathrm{ASym}_n,$$

więc również

$$\mathcal{L}_2(V) = \mathcal{L}_2^s(V) \oplus \mathcal{L}_2^a(V).$$

**Definicja 31.3.** **Niezdegenerowanie**

Powiemy, że odwzorowanie dwuliniowe $g \in \mathcal{L}_2(V)$ jest *niezdegenerowane*, jeśli dla dowolnego niezerowego wektora $v \in V$ istnieje taki wektor $u \in V$, że $g(u, v) \neq 0$.

**Przykład 31.4.** Iloczyn skalarny jest niezdegenerowany, bo $(v|v) \neq 0$, gdy $v \neq 0$.

**Lemat 31.1.** *Dla odwzorowania $g \in \mathcal{L}_2(V)$ następujące warunki są równoważne*

*(i)* *$g$ jest niezdegenerowane,*

*(ii)* *odwzorowanie liniowe*

$$L_g : V \ni v \longmapsto g_v \in V^*, \quad g_v(u) := g(u|v)$$

*jest izomorfizmem,*

*(iii)* $\det A_g \neq 0$, *gdzie $A_g$ jest macierzą $g$ w pewnej bazie $v_1, \ldots, v_n$.*

*Dowód.* Zauważmy, że $L_g$ jest monomorfizmem wtedy i tylko wtedy, gdy zachodzi implikacja

$$(\forall u \in V \ g(u, v) = 0) \Rightarrow v = 0.$$

Wynika stąd równoważność warunków (i) oraz (ii).

Dla dowodu równoważności warunków (ii) i (iii) pokażemy, że macierz odwzorowania $L_g$ w bazach $v_1, \ldots, v_n$ i $v_1^*, \ldots, v_n^*$ jest równa $A_g$. Rzeczywiście, jeśli

$$L_g(v_i) = a_{1i} \cdot v_1^* + \ldots + a_{ni} \cdot v_n^*,$$

to

$$\begin{aligned} g(v_j, v_i) &= (L_g(v_i))(v_j) \\ &= (a_{1i} \cdot v_1^* + \ldots + a_{ni} \cdot v_n^*)(v_j) \\ &= a_{ji}. \end{aligned}$$

□

**Definicja 31.4.** Forma kwadratowa

Odwzorowanie $q : V \longrightarrow \mathbb{R}$ postaci

$$q(v) = g(v,v), \quad v \in V$$

dla pewnego $g \in \mathcal{L}_2(V)$ nazywamy *formą kwadratową*. Przez $Q(V)$ oznaczamy przestrzeń wektorową form kwadratowych na $V$.

Każdej formie kwadratowej $q \in Q(V)$ możemy przypisać pewne odwzorowanie dwuliniowe i symetryczne $\Psi(q) \in \mathcal{L}_2^s(V)$, kładąc

$$\Psi(q)(u,v) = \frac{1}{2}(q(u+v) - q(u) - q(v)) = \frac{g(u,v)+g(v,u)}{2},$$

czyli $\Psi(q)$ jest symetryzacją odwzorowania dwuliniowego $g$, zadającego formę kwadratową $q$. Odwzorowanie

$$\Psi : Q(V) \longrightarrow \mathcal{L}_2^s(V)$$

jest liniowym izomorfizmem.

Niech $q \in Q(V)$ będzie formą kwadratową zadaną przez $g \in \mathcal{L}_2^s(V)$, czyli

$$q(v) = g(v,v).$$

Jeśli $A_g \in \mathrm{Sym}_n$ jest macierzą $g$ w pewnej bazie $v_1, \ldots, v_n$, to

$$q(v) = g(v,v) = (A_g x | x), \quad x = [x_1, \ldots, x_n]^T,$$

gdzie $v = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$. Ponieważ $A_g$ jest symetryczna, więc w pewnej bazie ortonormalnej $A_g$ jest diagonalna, czyli istnieje taka macierz ortonormalna $Q \in O(n)$, że

$$Q^T A_g Q = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

Liczby rzeczywiste $\lambda_1, \ldots, \lambda_n$ są wartościami własnymi macierzy $A_g$. Wynika stąd, że forma $q$ ma w pewnej bazie ortonormalnej postać

$$q(v) = \lambda_1 x_1^2 + \ldots + \lambda_n x_n^2.$$

Dla macierzy symetrycznej $A$ definiujemy trójkę $(\rho, \nu, \xi)$, gdzie $\rho$ jest liczbą dodatnich wartości własnych, $\nu$ liczbą ujemnych wartości własnych i $\xi$ krotnością wartości własnej 0.

**Twierdzenie 31.1** (Sylvestera o bezwładności)**.** *Jeśli $S \in M_{n\times n}(\mathbb{R})$ jest nieosobliwa, to macierze symetryczne $A$ i $S^T AS$ mają takie same trójki $(\rho, \nu, \xi)$.*

*Dowód.* Niech macierz symetryczna $A \in M_{n\times n}(\mathbb{R})$ ma trójkę $(p, j, s)$. Pokażemy najpierw, że istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{R})$, że

$$S^T AS = \underbrace{\begin{bmatrix} I_p & & \\ & -I_j & \\ & & 0_s \end{bmatrix}}_{=E}.$$

Wiemy, że jeśli

$$\lambda_1, \ldots, \lambda_p,\ -\lambda_{p+1}, \ldots, -\lambda_{p+j},\ \underbrace{0, \ldots, 0}_{s}$$

są wartościami własnymi $A$ z $\lambda_j > 0$, to istnieje taka macierz ortogonalna $P \in O(n)$, że

$$P^T AP = \mathrm{diag}(\lambda_1, \ldots, \lambda_p, -\lambda_{p+1}, \ldots, -\lambda_{p+j}, 0, \ldots, 0).$$

Definiujemy macierz nieosobliwą $S = PD$ dla

$$D = \mathrm{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_{p+j}}}, 1, \ldots, 1\right).$$

Wtedy

$$\begin{aligned} S^T AS &= (PD)^T APD \\ &= D^T P^T APD \\ &= D^T \operatorname{diag}(\lambda_1, \ldots, \lambda_p, -\lambda_{p+1}, \ldots, -\lambda_{p+j}, 0, \ldots, 0)D \\ &= E. \end{aligned}$$

Wystarczy pokazać, że jeśli

$$F = \begin{bmatrix} I_q & & \\ & -I_k & \\ & & 0_t \end{bmatrix},$$

czyli $F$ ma trójkę $(q, k, t)$ oraz

$$F = C^T EC,$$

dla pewnej macierzy nieosobliwej $C$, to $p = q$, $j = k$ i $s = t$. Z równości $F = C^T EC$ wynika, że $\operatorname{rank} E = \operatorname{rank} F$, czyli $s = t$.

Ponieważ $p + j + s = n = q + k + t$, więc wystarczy pokazać, że $p = q$. Przypuśćmy, że $p > q$. Zapiszmy macierz $C$ w postaci blokowej

$$C = \left[\begin{array}{c|c} X & Y \end{array}\right], \quad X \in M_{n\times q}(\mathbb{R}), \quad Y \in M_{n\times(n-q)}(\mathbb{R}).$$

Rozważmy podprzestrzenie

$$V = \operatorname{im} Y \subset \mathbb{R}^n, \quad W = \operatorname{span}\{e_1, \ldots, e_p\} \subset \mathbb{R}^n.$$

Ponieważ

$$\begin{aligned} \dim(V \cap W) &= \dim V + \dim W - \dim(V + W) \\ &= (n - q) + p - \dim(V + W) \\ &> 0, \end{aligned}$$

więc istnieje niezerowy wektor $x \in V \cap W$. Wtedy $x = Yy = C[0, y]^T$ dla pewnego $y \in \mathbb{R}^{n-q}$. Mamy

$$(Ex|x) = (F[0, y]^T|[0, y]^T) \leqslant 0$$

oraz $x \in W$, więc $x = [x_1, \ldots, x_p, 0, \ldots, 0]^T$, czyli

$$(Ex|x) > 0,$$

co prowadzi do sprzeczności. Analogicznie pokazujemy, że nie jest możliwe, aby $q > p$, więc $p = q$. □

## 31.1 Forma symplektyczna.

**Definicja 31.5.** Forma symplektyczna

Niech $V$ będzie skończenie wymiarową rzeczywistą przestrzenią wektorową. *Formą symplektyczną* na $V$ nazywamy dowolne niezdegenerowane odwzorowanie $\omega \in \mathcal{L}_2^a(V)$.

**Wniosek 31.2.** *Jeśli na przestrzeni $V$ istnieje forma symplektyczna $\omega$, to $\dim V$ jest parzysty.*

*Dowód.* Niech $A_\omega$ będzie macierzą formy $\omega$ w pewnej bazie $v_1, \ldots, v_n$. Ponieważ $\omega$ jest niezdegenerowana, więc $\det A_\omega \neq 0$. Z drugiej strony macierz $A_\omega$ jest antysymetryczna, więc

$$\begin{aligned} \det A_\omega &= \det A_\omega^T \\ &= \det(-A_\omega) \\ &= (-1)^n \det A_\omega, \end{aligned}$$

czyli $1 = (-1)^n$. □

**Przykład 31.5** (Standardowa forma symplektyczna). *Standardowa forma symplektyczna* $\omega_0$ *w* $\mathbb{R}^{2n}$ jest określona dla $u = [u_1, u_2]^T, v = [v_1, v_2]^T \in \mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$ przez

$$\omega_0(u, v) := (v_1|u_2) - (v_2|u_1) = (Ju|v),$$

gdzie

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \in M_{2n\times 2n}(\mathbb{R}).$$

Dla $u, v \in \mathbb{R}^{2n}$ zachodzi tożsamość

$$\omega_0(u, Jv) = (Ju|Jv) = (u|v).$$

Możemy powyższą konstrukcję uogólnić na przestrzeń wektorową $V$ wymiaru $n$ wyposażoną w iloczyn skalarny. Jeżeli $(\cdot|\cdot) : V \times V \longrightarrow \mathbb{R}$ jest iloczynem skalarnym, to w przestrzeni wektorowej $V \times V$ można wprowadzić formę $\omega$ daną dla $u = (u_1, u_2),\ v = (v_1, v_2) \in V \times V$ przez

$$\omega(u, v) = \omega((u_1, u_2), (v_1, v_2)) := (v_1|u_2) - (v_2|u_1).$$

Jeżeli $e_1, \ldots, e_n$ jest bazą ortonormalną dla $V$ względem iloczynu skalarnego $(\cdot|\cdot)$, to macierz formy $\omega$ w bazie

$$\overline{e}_i = (e_i, 0), \quad \overline{e}_{n+i} = (0, e_i), \quad i = 1, \ldots, n$$

jest dana przez $2n \times 2n$-macierz $J$. Wynika stąd, że $\omega$ jest niezdegenerowana, czyli jest formą symplektyczną. Zauważmy, że

$$\det J = 1, \quad J^T = J^{-1} = -J.$$

U *Strukturą zespoloną na rzeczywistej przestrzeni wektorowej* $V$ nazywamy takie odwzorowanie liniowe, że

$$J : V \longrightarrow V, \quad J^2 = -I.$$

Pozwala ono zadać na $V$ mnożenie przez skalar zespolony wzorem

$$(a + ib) \cdot u := a \cdot u + b \cdot Ju, \quad a, b \in \mathbb{R},\ u \in V.$$

W przypadku $V = \mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n \cong \mathbb{C}^n$ standardowa macierz symplektyczna $J$ jest mnożeniem przez $-i$, bo dla $z = x + iy$

$$Jz = y - ix.$$

U Jeżeli $(\cdot|\cdot)_* : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}$ jest hermitowskim iloczynem skalarnym, czyli

$$(z|w)_* = \sum_{j=1}^{n} z_j \overline{w_j}, \quad z = [z_1, \ldots, z_n]^T, w = [w_1, \ldots, w_n]^T \in \mathbb{C}^n,$$

to

$$(z|w)_* = \sum_{j=1}^{n}(x_j u_j + y_j v_j) + i\sum_{j=1}^{n}(u_j y_j - v_j x_j),$$

gdzie $x_j = x_j + iy_j$, $w_j = u_j + iv_j$. Wynika stąd, że

$$(z|w)_* = (z|w) + i\, \omega_0(z, w),$$

więc jego część rzeczywista jest euklidesowym iloczynem skalarnym w $\mathbb{R}^{2n}$, a część urojona jest standardową formą symplektyczną $\omega_0$.

Niech $(V, \omega)$ będzie liniową przestrzenią symplektyczną tzn. $\omega$ jest formą symplektyczną na $V$. Dla $u, v \in V$ definiujemy $\omega$-*ortogonalność* przez

$$u \perp_\omega v \Leftrightarrow \omega(u, v) = 0.$$

Ponieważ dla $u, v \in V$ mamy

$$u \perp_\omega v \Leftrightarrow v \perp_\omega u, \quad u \perp_\omega u,$$

więc każdy wektor jest do siebie $\omega$-ortogonalny. Geometria przestrzeni symplektycznych różni się więc od geometrii przestrzeni euklidesowych zadanej iloczynem skalarnym.

Dla podprzestrzeni liniowej $E \subset V$ przestrzeni symplektycznej $V$ definiujemy jej *dopełnienie $\omega$-ortogonalne* przez

$$E^{\perp_\omega} = \Big\{ u \in V : \forall v \in E \quad \omega(u, v) = 0 \Big\}.$$

Sprawdzamy łatwo, że $E^{\perp_\omega}$ jest podprzestrzenią liniową przestrzeni $V$. Zauważmy, że jeżeli $\dim E = 1$, to $E \subset E^{\perp_\omega}$.

**Lemat 31.2.** *Dla dowolnej podprzestrzeni $E \subset V$ zachodzi*

$$\dim V = \dim E + \dim E^{\perp_\omega}. \tag{31.2}$$

*Dowód.* Jeżeli $e_1, \dots e_d$ jest bazą $E$, to funkcjonały

$$\omega(e_1, \cdot), \dots, \omega(e_d, \cdot) \in V^*$$

są liniowo niezależne w $V^*$, bo $\omega$ jest niezdegenerowana. Stąd

$$F : V \ni v \longmapsto [\omega(e_1, v), \dots, \omega(e_d, v)]^T \in \mathbb{R}^d$$

jest epimorfizmem oraz $\ker F = E^{\perp_\omega}$, co kończy dowód. □

**Wniosek 31.3.**

$$(E^{\perp_\omega})^{\perp_\omega} = E$$

*Dowód.* Oczywiście $E \subset (E^{\perp_\omega})^{\perp_\omega}$ i z (31.2) mają one ten sam wymiar. □

Podprzestrzenie $E$ i $E^\perp$ nie muszą dawać w sumie prostej przestrzeni symplektycznej $V$. Na przykład, jeżeli $E$ jest jednowymiarowa, to

$$E \subset E^\perp,$$

czyli wtedy

$$E + E^\perp = E^\perp \neq V,$$

bo z (31.2) wynika, że

$$\dim E^\perp = \dim V - 1.$$

Jeżeli $E \subset V$ jest podprzestrzenią wektorową, to zawężenie $\omega|_{E\times E}$ jest dwuliniowe, antysymetryczne, ale nie musi być niezdegenerowane. Przykładowo, jeżeli $E$ jest jednowymiarowa, to $\omega|_{E\times E} = 0$.

Mówimy, że podprzestrzeń $E \subset V$ jest *symplektyczna*, jeżeli forma

$$\omega|_{E\times E} : E \times E \longrightarrow \mathbb{R}$$

jest niezdegenerowana, czyli $\omega|_{E\times E}$ jest formą symplektyczną w $E$.

**Lemat 31.3.** *Jeżeli $E$ jest podprzestrzenią przestrzeni symplektycznej $(V, \omega)$, to*

**(1)** *$E$ jest symplektyczna $\Leftrightarrow E \cap E^\perp = \{0\}$,*

**(2)** *$E$ jest symplektyczna $\Leftrightarrow E^\perp$ jest symplektyczna,*

**(3)** *jeżeli $E$ jest symplektyczna, to $V = E \oplus E^\perp$,*

**(4)** *jeżeli $V = E \oplus F$ i $\omega(E, F) = 0$, to $E$, $F$ są symplektyczne.*

*Dowód.* Udowodnimy punkt (4). Przypuśćmy, że $\omega$ jest zdegenerowana na $E$, więc istnieje taki wektor $0 \neq e \in E$, że $\omega(e, E) = 0$. Ponieważ $V = E \oplus F$ i $\omega(E, F) = 0$, więc $\omega(e, V) = 0$, czyli $\omega$ jest zdegenerowana na $V$, sprzeczność. □

Poniższe twierdzenie pokazuje, że każda liniowa przestrzeń symplektyczna $(V, \omega)$ wygląda jak przestrzeń $(\mathbb{R}^{2n}, \omega_0)$, to znaczy może być sprowadzona do tej samej postaci normalnej, co różni formy symplektyczne od niezdegenerowanych, symetrycznych form dwuliniowych.

**Twierdzenie 31.2** (Darboux)**.** *Jeżeli $(V, \omega)$ jest niezerową przestrzenią symplektyczną, to*

**(1)** $\dim V = 2n$ *jest parzysty,*

**(2)** *istnieje taka baza $e_1, \ldots, e_n, f_1, \ldots, f_n$ dla $V$, że*

$$\omega(e_i, e_j) = \omega(f_i, f_j) = 0, \quad \omega(f_i, e_j) = \begin{cases} 1, & i = j; \\ 0, & i \neq j. \end{cases}$$

*dla $i, j = 1, \ldots, n$,*

**(3)** *jeżeli $u, v \in V$ mają w tej bazie reprezentacje*

$$u = \sum_{i=1}^{n} (x_i e_i + x_{n+i} f_i), \quad v = \sum_{i=1}^{n} (y_i e_i + y_{n+i} f_i),$$

*to forma $\omega$ jest dana przez*

$$\omega(u, v) = (Jx|y) = \omega_0(x, y),$$

*gdzie $x = [x_1, \ldots, x_{2n}]^T, y = [y_1, \ldots, y_{2n}]^T \in \mathbb{R}^{2n}$.*

**(4)** *Podprzestrzenie $V_i = \operatorname{span}\{e_i, f_i\}$ $(i = 1, \ldots, n)$ są symplektyczne i $\omega$-ortogonalne, gdy $i \neq j$ (to znaczy $\omega(V_i, V_j) = 0$), czyli mamy rozkład*

$$V = V_1 \oplus \ldots \oplus V_n, \quad \dim V_i = 2.$$

*Dowód.* Wybieramy dowolnie $0 \neq e_1 \in V$. Ponieważ $\omega$ jest niezdegenerowana, więc istnieje $u \in V$ taki, że $\omega(u, e_1) \neq 0$. Normalizujemy $u$ tak, że $f_1 = \alpha u$ spełnia $\omega(f_1, e_1) = 1$. Oczywiście $e_1$, $f_1$ są liniowo niezeleżne, więc $E = \operatorname{span}\{e_1, f_1\}$ jest 2-wymiarowa i $E$ jest symplektyczna. Jeżeli $\dim V = 2$, to dowód jest zakończony. Jeżeli $\dim V > 2$, to ponieważ $E$ jest symplektyczna, więc mamy rozkład $V = E \oplus E^{\perp_\omega}$ i $E^{\perp_\omega}$ jest też symplektyczna. Możemy teraz zastosować poprzedni krok do $E^{\perp_\omega}$ i proces zakończy się w skończonej liczbie kroków. □

## 31.2 Liniowe odwzorowania symplektyczne

Niech $(V, \omega)$ będzie przestrzenią symplektyczną. Dla odwzorowania liniowego $A : V \longrightarrow V$ definiujemy 2-formę $A^*\omega$ („*cofnięcie formy*”) przez

$$A^*\omega(u, v) = \omega(Au, Av), \quad u, v \in V.$$

Mówimy, że odwzorowanie liniowe $A : V \longrightarrow V$ jest *symplektyczne*, jeżeli

$$A^*\omega = \omega. \tag{31.3}$$

**Lemat 31.4.** *Niech $V = \mathbb{R}^{2n}$ i $\omega_0(u, v) = (Ju|v)$. Następujące warunki są równoważne*

**(a)** *macierz $A \in M_{2n \times 2n}(\mathbb{R})$ jest symplektyczna,*

**(b)** $A^T J A = J$.

*Dowód.* Dla dowolnych $u, v \in \mathbb{R}^{2n}$ mamy

$$\begin{aligned} A^*\omega_0(u,v) &= \omega_0(u,v) \\ &\iff \omega_0(Au, Av) = \omega_0(u,v) \\ &\iff (JAu|Av) = (Ju|v) \\ &\iff (A^T JAu|v) = (Ju|v) \\ &\iff A^T JA = J. \end{aligned}$$

□

**Twierdzenie 31.3** (Wyznacznik macierzy symplektycznej)**.** *Jeśli* $A \in M_{2n\times 2n}(\mathbb{R})$ *jest symplektyczna, to* $\det A = 1$.

*Dowód.* Ponieważ $A^T JA = J$, więc $\det A = \pm 1$. Wystarczy więc pokazać, że $\det A > 0$. Rozważmy macierz $A^T A + I$. Ponieważ macierz Grama $A^T A$ ma dodatnie wartości własne (0 nie jest wartością własną, bo $A$ jest nieosobliwa), więc wartości własne $A^T A + I$ są większe od 1, czyli

$$\det(A^T A + I) > 1.$$

Ponieważ $A$ jest nieosobliwa, więc

$$\begin{aligned} A^T A + I &= A^T(A + (A^T)^{-1}) \\ &= A^T(A + JAJ^{-1}). \end{aligned}$$

Stąd

$$0 < 1 < \det(A^T A + I) = \det A \cdot \det(A + JAJ^{-1}),$$

czyli zakończymy dowód pokazując, że

$$\det(A + JAJ^{-1}) > 0.$$

Zauważmy, że $\det(A + JAJ^{-1}) \neq 0$. Zapiszmy $A$ w postaci blokowej

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad A_{ij} \in M_{n\times n}(\mathbb{R}).$$

Wtedy

$$\begin{aligned} A + JAJ^{-1} &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} + \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix} \\ &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} + \begin{bmatrix} A_{22} & -A_{21} \\ -A_{12} & A_{11} \end{bmatrix} \\ &= \begin{bmatrix} A_{11} + A_{22} & A_{12} - A_{21} \\ -A_{12} + A_{21} & A_{11} + A_{22} \end{bmatrix}. \end{aligned}$$

Oznaczmy $C = A_{11} + A_{22}$ i $D = A_{12} - A_{21}$. Wtedy

$$\begin{aligned} A + JAJ^{-1} &= \begin{bmatrix} C & D \\ -D & C \end{bmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} I & I \\ iI & -iI \end{bmatrix} \begin{bmatrix} C + iD & 0 \\ 0 & C - iD \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} I & -iI \\ I & iI \end{bmatrix}. \end{aligned}$$

Stąd

$$\begin{aligned} \det(A + JAJ^{-1}) &= \det(C + iD) \cdot \det(C - iD) \\ &= \det(C + iD) \cdot \det(\overline{C + iD}) \\ &= \det(C + iD) \cdot \overline{\det(C + iD)} \\ &= |\det(C + iD)|^2 > 0 \end{aligned}$$

□

Macierze symplektyczne pełnią ważną rolę w dziale teorii równań różniczkowych, dotyczących równań hamiltonowskich. Obejmują one równania ruchu Newtona z mechaniki klasycznej.

Definiujemy

$$\operatorname{Symp}(2) := \left\{A \in M_{2\times 2}(\mathbb{R}) : A^T J_2 A = J_2\right\}, \quad J_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Bezpośredni rachunek pokazuje, że

$$A^T J_2 A = \begin{bmatrix} 0 & -\det A \\ \det A & 0 \end{bmatrix}.$$

Wynika stąd, że $A \in \operatorname{Symp}(2)$ wtedy i tylko wtedy, gdy $\det(A) = 1$. Oznacza to, że

$$\operatorname{Symp}(2) = SL_2(\mathbb{R})$$

**Przykład 31.6.**

$$\operatorname{Symp}(2n) := \left\{A \in M_{2n\times 2n}(\mathbb{R}) : A^T J_{2n} A = J_{2n}\right\}, \quad J_{2n} = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}.$$

Oczywiście $I_{2n}, J_{2n} \in \operatorname{Symp}(2n)$. W dalszym ciągu stosujemy oznacznia $I = I_{2n}$, $J = J_{2n}$.

**Lemat 31.5.** $\operatorname{Symp}(2n)$ *jest grupą z działaniem mnożenia macierzy.*

*Dowód.* Jeśli $A^T JA = J$, to

$$\begin{aligned} J &= (A^T)^{-1} A^T JAA^{-1} \\ &= (A^T)^{-1} JA^{-1} \\ &= (A^{-1})^T JA^{-1}, \end{aligned}$$

czyli $J = (A^{-1})^T JA^{-1}$, zatem $A^{-1} \in \operatorname{Symp}(2n)$. Ponieważ

$$J^{-1} = A^{-1} J^{-1} (A^T)^{-1},$$

więc $AJ^{-1}A^T = J^{-1}$ oraz $J^{-1} = -J$, czyli

$$(A^T)^T JA^T = J,$$

zatem $A^T \in \operatorname{Symp}(2n)$. Ponadto,

$$\begin{aligned} (AB)^T JAB &= B^T A^T JAB \\ &= B^T JB \\ &= J, \end{aligned}$$

zatem $AB \in \operatorname{Symp}(2n)$. □

U Zauważmy, że jeśli $A \in \operatorname{Symp}(2n)$, to

$$A^{-1} = -JA^T J,$$

co pozwala łatwo wyznaczyć $A^{-1}$.

U Z definicji wynika, że jeśli $A \in \operatorname{Symp}(2n)$, to $\det(A) = \pm 1$. Z twierdzenia 31.3 $\det(A) = 1$ dla $A \in \operatorname{Symp}(2n)$. Dla $n > 1$ warunek $\det(A) = 1$ nie gwarantuje, że $A \in \operatorname{Symp}(2n)$. Na przykład macierz

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

nie jest symplektyczna, a ma wyznacznik równy 1.

**Lemat 31.6** (Wartości własne macierzy symplektycznej)**.** *Wartości własne $A \in \mathrm{Symp}(2n)$ mają następujące własności*

*(1)* $p_A(\lambda) = \lambda^{2n} p_A(\lambda^{-1})$*,*

*(2)* $\lambda \in \mathbb{C}$ *jest wartością własną $A$ wtedy i tylko wtedy, gdy $\lambda^{-1}$ jest wartością własną $A$. Ponadto, $\lambda$ i $\lambda^{-1}$ mają taką samą krotność,*

*(3) jeżeli $\pm 1$ jest wartością własną $A$, to ma parzystą krotność.*

*Dowód.* W dowodzie przyjmujemy, że $p := p_A$. Otrzymujemy, że

$$\begin{aligned}
p(\lambda) &= \det(A - \lambda I) \\
&= \det(A^T - \lambda I) \\
&= \det(-JA^{-1}J + \lambda JJ) \\
&= \det(J)\det(-A^{-1} + \lambda I)\det(J) \\
&= \det(-A^{-1} + \lambda I) \\
&= \det(A^{-1})\det(-I + \lambda A) \\
&= \det(-I + \lambda A) \\
&= \lambda^{2n}\det(A - \lambda^{-1} I) \\
&= \lambda^{2n} p\left(\frac{1}{\lambda}\right).
\end{aligned}$$

Ponieważ $A$ jest izomorfizmem, więc 0 nie jest wartością własną $A$, czyli punkt (2) wynika z (1). Jeżeli $-1$ jest wartością własną $A$, to ma ona krotność parzystą, bo $\det A = 1$ i z punktu (1) inne ujemne wartości własne pojawiają się „parami". Podobnie, jeżeli 1 jest wartością własną, to 1 ma krotność parzystą, gdyż z (1) iloczyn wszystkich innych wartości własnych ma „krotność parzystą". □

**Przykład 31.7** (Macierze hamiltonowskie)**.** Definiujemy

$$H_{2n} = \{A \in M_{2n\times 2n}(\mathbb{R}) : JA^T + AJ = 0\}.$$

Jest to grupa z dodawaniem macierzy. Jej elementy nazywamy *macierzami hamiltonowskimi.*

Pokażemy, że jeśli $A \in H_{2n}$ i $S \in \mathrm{Symp}(2n)$, to $S^{-1}AS \in H_{2n}$, czyli symplektyczna zmiana zmiennych nie zmienia hamiltonowskiego charakteru macierzy. Mamy

$$\begin{aligned}
JS^{-1}AS + S^TA^T(S^{-1})^TJ =& JS^{-1}AS + S^TA^T(S^T)^{-1}J \\
&= S^T(\underbrace{(S^T)^{-1}JS^{-1}}_{J}AS + A^T(S^T)^{-1}J) \\
&= S^T(JAS + A^T(S^T)^{-1}J) \\
&= S^T(JA + A^T\underbrace{(S^T)^{-1}JS^{-1}}_{J})S \\
&= S^T(JA + A^TJ)S \\
&= 0.
\end{aligned}$$

Jeśli $A \in H_{2n}$, to $\operatorname{tr} A = 0$. Rzeczywiście, $JA^T = -AJ$, więc $A^T = -J^{-1}AJ$, czyli $A^T$ jest podobna do $-A$. Stąd, $\operatorname{tr} A^T = -\operatorname{tr} A$. Z drugiej strony, $\operatorname{tr} A^T = \operatorname{tr} A$, więc $\operatorname{tr} A = -\operatorname{tr} A$.

**Przykład 31.8** (Grupa Heisenberga)**.** Zbiór wszyskich macierzy rzeczywistych postaci

$$A = \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}, \quad a, b, c \in \mathbb{R}$$

jest zamknięty ze względu na mnożenie macierzy i tworzy z tym działaniem grupę, zwaną *grupą Heisenberga*. Macierz odwrotna do macierzy $A$ jest dana wzorem

$$A^{-1} = \begin{bmatrix} 1 & -a & ac-b \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{bmatrix}.$$

Grupa Heisenberga jest podgrupą grupy $GL_3(\mathbb{R})$.

# Część IV

# Zadania

# Rozdział 32

# Grupy i ciała.

**Zadanie 32.1.** Uzasadnić, że $\sqrt{2}$ jest liczbą niewymierną.

**Zadanie 32.2.** Dodawanie i mnożenie jest dobrze zdefiniowane w zbiorach liczba naturalnych $\mathbb{N}$ i całkowitych $\mathbb{Z}$. Sprawdzić które z warunków (i)-(ix) definiujących ciało zachodzą dla trójek $(\mathbb{N}, +, \cdot)$ i $(\mathbb{Z}, +, \cdot)$.

**Zadanie 32.3.** Sprawdzić, że trójki $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ oraz $(\mathbb{C}, +, \cdot)$ z naturalnie określonymi działaniami są ciałami.

**Zadanie 32.4.** Zaznaczyć na płaszczyźnie liczby $1 + i$, $1 - i\sqrt{3}$ i liczby do nich sprzężone. Wyznaczyć liczby do nich odwrotne. Podać interpretację geometryczną sprzężenia liczby zespolonej.

**Zadanie 32.5.** Dla jakich liczb zespolonych zachodzi równość $z = \overline{z}$?

**Zadanie 32.6.** Uzasadnić, że dla dowolnych liczb zespolonych $z_1$ oraz $z_2$ zachodzą równości

$$\overline{z_1 + z_2} = \overline{z}_1 + \overline{z_2}, \quad \overline{z_1 \cdot z_2} = \overline{z}_1 \cdot \overline{z_2}$$

**Zadanie 32.7.** Niech $z_1$, $z_2 \in \mathbb{C}$. Pokazać, że

$$|z_1 \cdot z_2| = |z_1||z_2|,$$

$$|z_1 + z_2| \leqslant |z_1| + |z_2|.$$

**Zadanie 32.8.** Obliczyć

$$(1 + i)(2 + 3i) - 1 + i, \quad (1 - i)^3, \quad (3 + 4i)(5 - 6i) + (1 + 2i)(1 - i).$$

Dla $z = x + iy \neq 0 + i0$ stosujemy oznaczenie

$$\frac{1}{z} := z^{-1} = \frac{\overline{z}}{|z|^2} = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2}.$$

**Zadanie 32.9.** Liczby

$$\frac{2 + i}{3 + i}, \quad \frac{i}{1 + i}, \quad \frac{1 - 5i}{1 - i}$$

zapisać w postaci $x + iy$.

**Zadanie 32.10.** Znaleźć takie liczby rzeczywiste $x, y \in \mathbb{R}$, że

$$(1 + 2i)x + (3 - 5i)y = 1 - 3i.$$

**Zadanie 32.11.** Rozwiązać równania

$$|z|^2 - z = 1 - 2i, \quad z\overline{z} + z - \overline{z} = 3 + 2i.$$

**Zadanie 32.12.** Znaleźć liczby zespolone $z, w \in \mathbb{C}$ spełniające układ równań

$$\begin{cases} (2+i)z + (2-i)w = 6 \\ (3+2i)z + (3-2i)w = 8 \end{cases}$$

**Zadanie 32.13.** Korzystając ze wzoru de Moivre'a wyprowadzić wzory na $\cos 3\phi$ oraz $\sin 3\phi$.

**Zadanie 32.14.** Obliczyć $(1-\sqrt{3}i)^5$ oraz $(1+i)^{2017}$.

**Zadanie 32.15.** Wyznaczyć pierwiastki stopnia $n$ z liczb zespolonych $1$, $-1$, $i$, $1+i$, $1+i\sqrt{3}$.

**Zadanie 32.16.** Wyznaczyć pierwiastki stopnia 1, 2, 3, 4, 5 z liczby 1. Podaj ich interpretację geometryczną.

**Zadanie 32.17.** Rozważmy równanie kwadratowe o współczynnikach rzeczywistych ($a, b, c \in \mathbb{R}$)

$$az^2 + bz + c = 0, \quad a \neq 0$$

oraz $\Delta = b^2 - 4ac < 0$. Uzasadnić, że ma ono dwa rozwiązania zespolone

$$z_1 = \frac{-b - i\sqrt{|\Delta|}}{2a}, \quad z_2 = \frac{-b + i\sqrt{|\Delta|}}{2a}.$$

Zauważmy, że $\overline{z_2} = z_1$.

**Zadanie 32.18.** Rozważmy równanie kwadratowe o współczynnikach zespolonych ($a, b, c \in \mathbb{C}$)

$$az^2 + bz + c = 0, \quad a \neq 0.$$

Niech $\Delta = b^2 - 4ac$ i $w \in \mathbb{C}$ będzie takie, że $w^2 = \Delta$, czyli $w$ jest jednym z dwóch pierwistków stopnia 2 z liczby zespolonej $\Delta$ (drugim jest $-w$). Uzasadnić, że rozwiązaniami równania $az^2 + bz + c = 0$ są

$$z_1 = \frac{-b - w}{2a}, \quad z_2 = \frac{-b + w}{2a}.$$

**Zadanie 32.19.** Rozwiązać równania w zbiorze liczb zespolonych

- $z^2 + 4 = 0$,
- $z^2 - z + 3 = 0$,
- $(1-i)z^2 + (-4+2i)z + 7 - 9i = 0$,
- $z^2 + 2iz - 5 = 0$,
- $z^2 + 2i = 0$.

**Zadanie 32.20.** Uzasadnić, że jeśli $p \in \mathbb{N}$ jest liczbą pierwszą, to $(\mathbb{Z}_p, +, \cdot)$ jest ciałem.

**Zadanie 32.21.** Scharakteryzować macierze $A \in M_{2\times 2}(\mathbb{R})$ spełniające warunki:

(i) $A^2 = -I$,

(ii) $A^T A = -I$.

**Zadanie 32.22.** Niech $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in M_{2\times 2}(\mathbb{C})$ oraz $a_{11} \neq 0$ i $\alpha = a_{21}/a_{11}$. Pokazać, że $A$ można zapisać w postaci iloczynu postaci

$$A = \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ 0 & b \end{bmatrix}.$$

Ile jest równe $b$ w tym przedstawieniu?

**Zadanie 32.23.** Sprawdzić, że dla $A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$ zachodzi $A^2 = 0$. Czy istnieje taka niezerowa macierz $B \in M_{2\times 2}(\mathbb{R})$, że $B^T = B$ i $B^2 = 0$?

**Zadanie 32.24.** Niech $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Oblicz $A^2$, $A^3$. Zgadnij wzór na $A^n$ i go uzasadnij.

**Zadanie 32.25.** Sprawdzić czy dla dowolnych macierzy $A$, $B \in M_{2\times 2}(\mathbb{R})$ zachodzi równość $(A+B)^2 = A^2 + 2AB + B^2$.

**Zadanie 32.26.** Podać przykład takich macierzy niezerowych $A$, $B \in M_{2\times 2}(\mathbb{R})$, że $AB = 0$ i $A \neq B$.

**Zadanie 32.27.** Podać przykład takich niezerowych macierzy $A$, $B$, $C \in M_{2\times 2}(\mathbb{R})$, że $A \neq B$ oraz $AC = BC$.

**Zadanie 32.28.** Pokazać, że dla $A = \begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix}$ zachodzi wzór

$$A^n = \begin{bmatrix} 1+6n & 4n \\ -9n & 1-6n \end{bmatrix}, \quad n \geqslant 1.$$

**Zadanie 32.29.** Dla macierzy $A \in M_{2\times 2}(\mathbb{F})$ zachodzą wzory

- $$\det(A) = \frac{1}{2} \det \begin{bmatrix} \operatorname{tr}(A) & 1 \\ \operatorname{tr}(A^2) & \operatorname{tr}(A) \end{bmatrix}.$$
- $\operatorname{tr} A^2 = (\operatorname{tr} A)^2 - 2\det A$,
- $\operatorname{tr} A^3 = (\operatorname{tr} A)^3 - 3(\det A)(\operatorname{tr} A)$.

**Zadanie 32.30.** Niech $A$, $B \in M_{2\times 2}(\mathbb{R})$ będą takimi macierzami, że $\det A = 7$ i $\det B = 3$. Obliczyć $\det(2AB)$ oraz $\det(A^{-1}B)$.

**Zadanie 32.31.** Niech $A \in M_{2\times 2}(\mathbb{R})$. Pokazać, że

$$\det(A^2 + I) = |\det(A + iI)|^2 \geqslant 0.$$

**Zadanie 32.32.** Niech $A, B \in M_{2\times 2}(\mathbb{C})$. Pokazać, że

$$\det(A+B) + \det(A-B) = 2\det A + 2\det B.$$

**Zadanie 32.33.** Niech $A \in M_{2\times 2}(\mathbb{R})$. Obliczyć $\det A$ jeśli $\det(A+A^T) = 8$ i $\det(A+2A^T) = 27$.

**Zadanie 32.34.** Niech $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{R})$ będzie taka, że $A^2 = -I$. Sprawdzić, że

- $a + d = 0$, $b \neq 0$ i $c \neq 0$,
- $A = \begin{bmatrix} a & b \\ -\frac{1+a^2}{b} & -a \end{bmatrix}$.
- dla $P = \begin{bmatrix} ab & -b \\ -(1+a^2) & 0 \end{bmatrix}$ mamy

$$P^{-1}AP = J.$$

**Zadanie 32.35.** Niech $A = \begin{bmatrix} 2 & i \\ i & 0 \end{bmatrix}$. Pokazać, że $A^n = \begin{bmatrix} n+1 & ni \\ ni & 1-n \end{bmatrix}$ dla $n \geqslant 1$.

**Zadanie 32.36.** Pokazać, że $A \in M_{2\times 2}(\mathbb{C})$ komutuje z każdą macierzą symetryczną wtedy i tylko wtedy, gdy $A = \alpha I$ dla pewnego $\alpha \in \mathbb{C}$.

**Zadanie 32.37.** Niech $A \in M_{2\times 2}(\mathbb{R})$. Pokazać, że

- $AA^T = A^T A$ wtedy i tylko wtedy, gdy
$$A = c\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$
dla pewnych $c, \theta \in \mathbb{R}$.
- $A^2 = 0$ wtedy i tylko wtedy, gdy
$$A = c\begin{bmatrix} \cos\theta & 1+\sin\theta \\ -1+\sin\theta & \cos\theta \end{bmatrix}$$
dla pewnych $c, \theta \in \mathbb{R}$.

**Zadanie 32.38.** Dla $A \in M_{2\times 2}(\mathbb{C})$ definiujemy
$$C(A) = \{X \in M_{2\times 2}(\mathbb{C}) : AX = XA\}.$$
Pokazać, że

- jeśli $A = kI$ dla pewnego $k \in \mathbb{C}$, to $C(A) = M_{2\times 2}(\mathbb{C})$,
- jeśli $A \neq kI$, to
$$C(A) = \{\alpha A + \beta I : \alpha, \beta \in \mathbb{C}\}.$$
- Jeśli $AB \neq BA$, to
$$C(A) \cap C(B) = \{\alpha I : \alpha \in \mathbb{C}\}.$$

**Zadanie 32.39.** Niech $A, B, C \in M_{2\times 2}(\mathbb{C})$ będą takie, że
$$A^2 = BC, \quad B^2 = CA, \quad C^2 = AB.$$
Pokazać, że $A^3 = B^3 = C^3$.

**Zadanie 32.40.** Niech $A \in M_{2\times 2}(\mathbb{C})$. Pokazać, że $A^2 = A$ wtedy i tylko wtedy, gdy $(2A-I)^2 = I$. Uzasadnić, ze jeśli $A^2 = A$, to $I + A$ jest odwracalna oraz
$$(I + A)^{-1} = I - \frac{1}{2}A.$$

**Zadanie 32.41.** Sprawdzić, czy
$$G = \left\{ \begin{bmatrix} 2-a & a-1 \\ 2(1-a) & 2a-1 \end{bmatrix} : a \in \mathbb{R} \setminus \{0\} \right\}$$
jest grupą abelową z działaniem mnożenia macierzy?

**Zadanie 32.42.** Sprawdzić, czy
$$G = \left\{ \begin{bmatrix} x & 2y \\ y & x \end{bmatrix} : x, y \in \mathbb{Q} \setminus \{0\} \right\}$$
jest grupą abelową z działaniem mnożenia macierzy?

**Zadanie 32.43.** Wykazać, że

$$SO(2) := \left\{ R_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} : \theta \in \mathbb{R} \right\}$$

jest grupą abelową z działaniem mnożenia macierzy. Ponadto,

$$f : \mathbb{S}^1 \ni \cos\theta + i\sin\theta \mapsto R_\theta \in SO(2)$$

jest izomorfizmem grup. Dla dowolnego $n \geqslant 1$ zbiór

$$\mathcal{R}_n = \{R_{\frac{2k\pi}{n}} : k = 0, 1, \ldots, n-1\} \subset SO(2)$$

jest grupą abelową z mnożeniem macierzy oraz

$$f : G_n = \{z \in \mathbb{C} : z^n = 1\} \to \mathcal{R}_n$$

jest izomorfizmem grup.

**Zadanie 32.44.** Sprawdzić. czy

$$G = \left\{ \begin{bmatrix} \cos\theta & 3\sin\theta \\ -\frac{1}{3}\sin\theta & \cos\theta \end{bmatrix} : \theta \in \mathbb{R} \right\}$$

jest grupą abelową z działaniem mnożenia macierzy?

**Zadanie 32.45.** Pokazać, że

$$G = \left\{ A \in M_{2\times 2}(\mathbb{Z}) : \det A = \pm 1 \right\}$$

jest grupą z mnożeniem macierzy.

**Zadanie 32.46.**

$$G = \left\{ \begin{bmatrix} x & 2y \\ \frac{5}{2}y & x \end{bmatrix} : x, y \in \mathbb{Q}\ x^2 - 5y^2 = 1 \right\}$$

jest grupą abelową z działaniem mnożenia macierzy.

**Zadanie 32.47** (Grupa czwórkowa Kleina)**.** Niech

$$K_4 = \{I, S_x, S_y, S_0\},$$

gdzie

$$S_x = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad S_y = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad S_0 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -I.$$

$K_4$ jest grupą abelową z mnożeniem macierzy. Sprawdzić, że

| $\cdot$ | I | $S_x$ | $S_y$ | $S_z$ |
|---|---|---|---|---|
| I | I | $S_x$ | $S_y$ | $S_0$ |
| $S_x$ | $S_x$ | I | $S_0$ | $S_y$ |
| $S_y$ | $S_y$ | $S_0$ | I | $S_x$ |
| $S_z$ | $S_0$ | $S_y$ | $S_x$ | I |

**Zadanie 32.48.** Dla $a, b, \alpha, \beta \in \mathbb{C}$ definiujemy macierze

$$U(a,b) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \quad V(\alpha,\beta) = \begin{bmatrix} \alpha & \beta \\ \beta & -\alpha \end{bmatrix}.$$

Pokazać, że

- $U(a,b)U(a',b') = U(aa' - bb', a'b + ab')$,

- jeśli $a^2+b^2 \neq 0$, to $U(a,b)$ jest odwracalna oraz

$$U^{-1}(a,b) = U\left(\frac{a}{a^2+b^2}, -\frac{b}{a^2+b^2}\right),$$

- $U(a,b)V(\alpha,\beta) = V(a\alpha - b\beta, a\beta + b\alpha)$,
- $V(\alpha,\beta)U(a,b) = V(\alpha a + \beta b, \beta a - \alpha b)$,
- $U(\alpha,\beta)V(1,0) = V(\alpha,\beta)$,
- $V(\alpha,\beta)V(\alpha',\beta') = U(\alpha\alpha' + \beta\beta', \alpha'\beta - \alpha\beta')$,
- jeśli $\alpha^2+\beta^2 \neq 0$, to $V(\alpha,\beta)$ jest odwracalna oraz

$$V^{-1}(\alpha,\beta) = V\left(\frac{\alpha}{\alpha^2+\beta^2}, \frac{\beta}{\alpha^2+\beta^2}\right),$$

- jeśli $a^2+b^2 \neq 0$, to

$$U(a,b) = \sqrt{a^2+b^2}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

  gdzie $\cos\theta = \frac{a}{\sqrt{a^2+b^2}}$, $\sin\theta = \frac{b}{\sqrt{a^2+b^2}}$ dla pewnego $\theta \in [0, 2\pi)$,

- jeśli $\alpha^2+\beta^2 \neq 0$, to

$$V(\alpha,\beta) = \sqrt{\alpha^2+\beta^2}\begin{bmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{bmatrix},$$

  gdzie $\cos t = \frac{\alpha}{\sqrt{\alpha^2+\beta^2}}$, $\sin t = \frac{\beta}{\sqrt{\alpha^2+\beta^2}}$ dla pewnego $\theta \in [0, 2\pi)$.

**Zadanie 32.49** (Grupa dihedralna)**.** Dla $n \geqslant 3$ definiuejmy zbiór

$$D_{2n} = \left\{ \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad \begin{bmatrix} -\cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \text{ gdzie } \theta = 0, \frac{2\pi}{n}, \ldots, \frac{2(n-1)\pi}{n} \right\}.$$

Wykazać, że $D_{2n}$ jest grupą z mnożeniem macierzy.

**Zadanie 32.50.** Sprawdzić czy dla $d \in \mathbb{R}$ zbiór

$$M_d = \left\{ \begin{bmatrix} a & db \\ b & a \end{bmatrix} : a, b \in \mathbb{R},\ a^2 - db^2 \neq 0 \right\}$$

jest grupą z mnożeniem macierzy?

**Zadanie 32.51.** Rozważmy macierze, należące do $\mathrm{SL}_2(\mathbb{R})$:

$$U = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad V = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

$$W = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad P = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}.$$

Sprawdzić, że

- $Q := P^2 = \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}$,
- $U^k = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$ dla $k \in \mathbb{Z}$,
- $V^2 = -I$, $V^{-1} = -V$, $V^4 = I$,

- $W^k = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$ dla $k \in \mathbb{Z}$ oraz $W = UVU$,
- $P = VU = V^2Q^2$ oraz $P^3 = -I$,
- $Q = VUVU = VW$, $Q^2 = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}$, $Q^3 = I$,
- $U = V^{-1}P = -VP = VP^4 = VQ^2$.

**Zadanie 32.52.** Ustalmy $n \in \mathbb{N}$ i taką macierz $A \in M_{2\times 2}(\mathbb{C})$, że $A \neq 0, I$. Rozważmy zbiór

$$\mathcal{M}_n = \{X \in M_{2\times 2}(\mathbb{C}) : X^n = A\}.$$

Pokazać, że następujące warunki są równoważne

(1) $(\mathcal{M}_n, \cdot)$ jest grupą,

(2) $A^2 = A$,

(3) $\mathcal{M}_n = \{zA : z \in \mathbb{C}, z^n = 1\}$.

**Zadanie 32.53** (kwateriony macierzowe)**.** Rozważmy macierze

$$\mathcal{I} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad K = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

Rozważmy zbiór macierzy

$$\mathcal{H} = \{M = aI + b\mathcal{I} + cJ + dK : a, b, c, d \in \mathbb{R}\}.$$

Pokazać, że

- $\mathcal{H}$ z działaniami dodawania i mnożenia macierzy spełnia wszystkie warunki definiujące ciało z wyjątkiem przemienności mnożenia.
- Obliczyć $M\tilde{M}$ i $\tilde{M}M$, gdzie $\tilde{M} = aI - b\mathcal{I} - cJ - dK$.
- Znaleźć wszystkie takie macierze $X \in \mathcal{H}$, że $X^2 + I = 0$.

# Rozdział 33

# Przestrzeń wektorowa $\mathbb{F}^n$

**Zadanie 33.1.** Niech $V$ będzie przestrzenią wektorową nad $\mathbb{F} \in \{\mathbb{C}, \mathbb{R}, \mathbb{Q}\}$. Pokazać, że albo $V$ jest jednoelementowa albo $V$ jest zbiorem nieskończonym. Czy powyższe jest prawdą dla przestrzeni wektorowej nad dowolnym ciałem?

**Zadanie 33.2.** Dla jakich wartości parametru $x \in \mathbb{R}$ wektory $[-3x, 2x]^T$ i $[4, x]^T$ są prostopadłe?

**Zadanie 33.3.** Dla jakich wartości parametru $k \in \mathbb{R}$ wektory

$$u = [1, -1, 2]^T, \quad v = [k^2, k, -3]^T,$$

są prostopadłe.

**Zadanie 33.4.** Znaleźć wszystkie wektory prostopadłe do wektora $[1, 2]^T$, które mają normę jeden.

**Zadanie 33.5.** Znajdź kąt $\angle(u, v)$ pomiędzy wektorami:

- $u = [3, 4]^T$, $v = [5, 12]^T$,
- $u = [1, 2, 3]^T$, $v = [2, 1, 2]^T$.

**Zadanie 33.6.** Znaleźć rzut wektora $v$ na wektor $u$ dla

- $v = [-1, 3]^T$, $u = [2, 1]^T$,
- $v = [1, 2, 3]^T$, $u = [0, 0, 1]^T$,
- $v = [1, 2, 3]^T$, $u = [1/2, 1/2, 1/\sqrt{2}]^T$.

**Zadanie 33.7.** Udowodnić następujące *twierdzenie cosinusów* (wzór Carnota). W dowolnym trójkącie o bokach długości $a, b, c$ i przeciwległych kątach $\alpha$, $\beta$, $\gamma$ zachodzi równość

$$a^2 = b^2 + c^2 - 2bc \cos \alpha.$$

Wskazówka: Bez straty ogólności można założyć, że jednym z wierzchołków trójkąta jest punkt 0. Kładąc $b = \|u\|$, $c = \|v\|$ oraz $a = \|u - v\|$ obliczyć $a^2 = \|u - v\|^2$.

**Zadanie 33.8.** Pokazać, że odległość punktu $p = [p_1, p_2, p_3]^T$ od płaszczyzny prostopadłej do niezerowego wektora $n = [A, B, C]^T$ i przechodzącej przez 0 jest równa

$$d = \frac{|(p|n)|}{\|n\|} = \frac{|Ap_1 + Bp_2 + Cp_3|}{\sqrt{A^2 + B^2 + C^2}}.$$

Znaleźć odległość punktu $[1, 1, 1]^T$ od płaszczyzny $2x + 2y + z = 0$.

**Zadanie 33.9.** Uzasadnić, że $\|u - v\| \geqslant |\|u\| - \|v\||$.

**Zadanie 33.10.** Uzasadnić, że $\|u + v\| = \|u - v\|$ wtedy i tylko wtedy, gdy $u$ i $v$ są ortogonalne.

**Zadanie 33.11.** Niech $u, v, w \in \mathbb{R}^3$ będą parami ortogonalne. Pokazać, że

$$\|u\|^2 + \|v\|^2 + \|w\|^2 = \|u + v + w\|^2.$$

**Zadanie 33.12.** Rozważmy wektory w $\mathbb{R}^3$:

$$n_1 = [1, 0, 0]^T, \quad n_2 = [0, 1, 0]^T, \quad n_3 = [2^{-1/2}, 2^{-1/2}, 0]^T$$

oraz liczby $p_1 = p_2 = 0$, $p_3 = 2^{1/2}$. Sprawdzić, że każde dwie spośród płaszczyzn

$$\pi_i = \{x \in \mathbb{R}^3 : (n_i|x) = p_i, \quad i = 1, 2, 3$$

przecinają się wzdłuż prostej, ale wszystkie trzy nie mają punktów wspólnych.

**Zadanie 33.13.** Pokazać, że dla niezerowych wektorów $x, y \in \mathbb{R}^n$ zachodzi równość

$$\left\| \frac{x}{\|x\|^2} - \frac{y}{\|y\|^2} \right\| = \frac{1}{\|x\|\|y\|} \|x - y\|.$$

**Zadanie 33.14.** Pokazać, że dla $x, y, z \in \mathbb{R}^n$ mamy

$$\|z\|\|x - y\| \leqslant \|y\|\|z - x\| + \|x\|\|y - z\|.$$

Ponadto, dla niezerowych wektorów zachodzi równość wtedy i tylko wtedy, gdy $\|x\|^{-2}x$, $\|y\|^{-2}y$, $\|z\|^{-2}z$ leżą na jednej prostej.

**Zadanie 33.15.** W oparciu o nierówność Cauchy'ego-Schwarza, pokazać, że

$$x^2 + y^2 + z^2 \geqslant yz + zx + xy,$$

dla $x, y, z \in \mathbb{R}$.

**Zadanie 33.16.** Niech $a \in \mathbb{R}^n$ będzie ustalony. Załóżmy, że $x, y \in \mathbb{R}^n$ spełniają równość

$$x + (x|y) \cdot y = a.$$

- Pokazać, że $(x|y)^2 = \frac{\|a\|^2 - \|x\|^2}{2 + \|y\|^2}$.
- Uzasadnić, że $\|x\|(1 + \|y\|^2) \geqslant \|a\| \geqslant \|x\|$. Wyjaśnić, przy jakich jakich założeniach zachodzą równości.

**Zadanie 33.17.** Obliczyć

- $[2, 1, 2]^T + [2, 0, 1]^T$ w $\mathbb{Z}_3^3$,
- $2 \cdot [2, 2, 1]^T$ w $\mathbb{Z}_3^3$,
- $2 \cdot ([3, 1, 1, 2]^T + [3, 3, 2, 1]^T)$ w $\mathbb{Z}_5^4$.

**Zadanie 33.18.** Korzystając z nierówności Cauchy'ego-Schwarza, pokazać, że jeśli $a_1, \ldots, a_n > 0$, to

$$\left(\sum_{i=1}^{n} a_i\right)\left(\sum_{i=1}^{n} \frac{1}{a_i}\right) \geqslant n^2.$$

**Zadanie 33.19.** Sprawdzić, czy $\mathbb{R}^2$ z określonymi poniżej działaniami jest przestrzenią wektorową

- $[x_1, y_1]^T + [x_2, y_2]^T = [x_1 + x_2, y_1 + y_2]^T$ i $t \cdot [x, y]^T = [t \cdot x, y]^T$;
- $[x_1, y_1]^T + [x_2, y_2]^T = [x_1, 0]^T$ i $t \cdot [x, y]^T = [t \cdot x, t \cdot y]^T$

**Zadanie 33.20.** Niech $V$ będzie przestrzenią wektorową nad ciałem $\mathbb{F}$. Pokazać, że dla dowolnego $v \in V$ zachodzi równość:

$$(-1) \cdot v = -v.$$

**Zadanie 33.21.** Niech $V$ będzie przestrzenią wektorową nad ciałem $\mathbb{F}$. Pokazać, że

- $(a-b)\cdot u = a\cdot u - b\cdot u$ dla $u\in V$ i $a,b\in\mathbb{F}$,
- $a\cdot(u-v) = a\cdot u - a\cdot v$ dla $u,v\in V$ i $a\in\mathbb{F}$,
- jeśli $a\in\mathbb{F}\setminus\{0\}$ i $a\cdot u = a\cdot v$ dla pewnych $u,v\in V$, to $u=v$,
- jeśli $u\in\mathbb{F}\setminus\{0\}$ i $a\cdot u = b\cdot u$ dla pewnych $a,b\in\mathbb{F}$, to $a=b$.

**Zadanie 33.22.** Znaleźć równanie ogólne płaszczyzny zawierającej wektory

$$[1,2,3]^T,\quad [2,3,4]^T,\quad [1,3,5]^T.$$

**Zadanie 33.23.** Niech $V=(0,+\infty)$ będzie zbiorem dodatnich liczb rzeczywistych. Definiujemy działanie dodawania w $V$ i mnożenia przez skalar rzeczywisty wzorami

$$x+y := xy,\quad x,y\in V,$$

$$c\cdot x := x^c,\quad x\in V,\ c\in\mathbb{R}.$$

Sprawdzić, czy trójka $(V,+,\cdot)$ jest przestrzenią wektorową.

**Zadanie 33.24.** Niech $f:\mathbb{R}\to\mathbb{R}$ będzie injekcją. W zbiorze $V:=f(\mathbb{R})$ definiujemy działanie dodawania i mnożenia przez skalar rzeczywisty wzorami

$$x+y := f(f^{-1}(x)+f^{-1}(y)),\quad x,y\in V,$$

$$c\cdot x := f(c\cdot f^{-1}(x)),\quad x\in V,\ c\in\mathbb{R}.$$

Sprawdzić, że trójka $(V,+,\cdot)$ jest przestrzenią wektorową. Jako szczególny przypadek rozważyć funkcję $f:\mathbb{R}\ni x\mapsto e^x\in\mathbb{R}$.

**Zadanie 33.25.** W zbiorze $V=\mathbb{R}$ definiujemy działanie dodawania i mnożenia przez skalar rzeczywisty wzorami

$$x\oplus y := x+y+7,\quad x,y\in\mathbb{R},$$

$$c\circ x := cx+7(c-1).$$

Sprawdzić, czy trójka $(\mathbb{R},\oplus,\circ)$ jest przestrzenią wektorową.

**Zadanie 33.26.** Niech $V$ będzie zbiorem macierzy zespolonych postaci

$$\begin{bmatrix} a & b \\ -\overline{b} & \overline{a} \end{bmatrix}.$$

Sprawdzić, że z naturalnie określonymi działaniami $V$ jest przestrzenią wektorową nad $\mathbb{R}$. Czy jest również przestrzenią wektorową nad $\mathbb{C}$?

# Rozdział 34

# Układy równań liniowych

**Zadanie 34.1.** Rozwiązać układ równań

$$\begin{cases} x_1 + 2x_2 + x_3 = 3 \\ 3x_1 - x_2 - 3x_3 = -1, \\ 2x_1 + 3x_2 + x_3 = 4 \end{cases}$$

sprowadzając go metodą eliminacji Gaussa do postaci schodkowej.

**Zadanie 34.2.** Rozwiązać poniższe układy równań metodą eliminacji Gaussa

$$\begin{cases} 2x_1 + 3x_2 + x_3 = 1 \\ x_1 + x_2 + x_3 = 3, \\ 3x_1 + 4x_2 + 2x_3 = 4 \end{cases},$$

$$\begin{cases} x_1 - x_2 + 2x_3 = 4 \\ 2x_1 + 3x_2 - x_3 = 1 \\ 7x_1 + 3x_2 + 4x_3 = 7 \end{cases},$$

$$\begin{cases} -x_1 + 2x_2 - x_3 = 2 \\ -2x_1 + 2x_2 + x_3 = 4 \\ 3x_1 + 2x_2 + 2x_3 = 5 \\ -3x_1 + 8x_2 + 5x_3 = 17 \end{cases}$$

**Zadanie 34.3.** Rozwiązać układy równań skojarzone z macierzami

$$\left[\begin{array}{ccc|c} 2 & 1 & 4 & -1 \\ 4 & -2 & 3 & 4 \\ 5 & 2 & 6 & -1 \end{array}\right], \quad \left[\begin{array}{ccc|c} 5 & -2 & 1 & 3 \\ 2 & 3 & -4 & 0 \end{array}\right]$$

**Zadanie 34.4.** Dla jakich wartości parametru $a \in \mathbb{R}$ układ równań

$$\begin{cases} x + 2y + z = 1 \\ -x + 4y + 3z = 2, \\ 2x - 2y + az = 3 \end{cases}$$

ma dokładnie jedno rozwiązanie?

**Zadanie 34.5.** Zbadać liczbę rozwiązań poniższego układu w zależności od parametrów $a, b \in \mathbb{R}$.

$$\begin{cases} x + y + 3z = 2 \\ x + 2y + 4z = 3, \\ x + 3y + az = b \end{cases}$$

**Zadanie 34.6.** Sprawdzić dla jakich wartości $p \in \mathbb{R}$ układ równań

$$\begin{cases} 6p^2x - 3y = 3p \\ 2x - y = 7 \end{cases}$$

ma dokładnie jedno rozwiązanie i je wyznaczyć.

**Zadanie 34.7.** Rozwiązać układy równań nad $\mathbb{Z}_3$

$$\begin{cases} x+2y=1 \\ x+y=2, \end{cases} \qquad \begin{cases} x+y=1 \\ y+z=0, \\ x+z=1 \end{cases}$$

oraz układ nad $\mathbb{Z}_5$:

$$\begin{cases} 3x+2y=1 \\ x+4y=1, \end{cases}$$

**Zadanie 34.8.** Znaleźć wielomian czwartego stopnia przechodzący przez punkty $(-2,3)$, $(-1,5)$, $(0,1)$, $(1,4)$ i $(2,10)$.

# Rozdział 35

# Baza

**Zadanie 35.1.** Sprawdzić, które z poniższych zbiorów z naturalnymi działaniami są przestrzeniami wektorowymi. Dla tych które są określić ich wymiar.

- $\mathbb{C}$ nad $\mathbb{C}$.
- $\mathbb{C}$ nad $\mathbb{R}$.
- $\mathbb{R}$ nad $\mathbb{C}$.
- $\mathbb{Q}$ nad $\mathbb{R}$.
- $\mathbb{R}$ nad $\mathbb{Q}$.
- $\{a + b\sqrt{3} : a, b \in \mathbb{Q}\}$ nad $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.

**Zadanie 35.2.** Niech $u = [2,1]^T$, $v = [4,3]^T$, $w = [7,-3]^T \in \mathbb{R}^2$. Sprawdzić, że $u, v$ są bazą dla $\mathbb{R}^2$. Pokazać, że $u, v, w$ są liniowo zależne.

**Zadanie 35.3.** Czy wektory $u = [3,-2,4]^T, v = [-3,2,-4]^T, w = [-6,4,8]^T$ tworzą bazę dla $\mathbb{R}^3$?

**Zadanie 35.4.** Czy wektory $v_1 = [1+i, 2+3i]^T$, $v_2 = [2+4i, 5-i]^T$ są bazą dla $\mathbb{C}^2$ nad ciałem $\mathbb{C}$? Czy są one bazą $\mathbb{C}^2$ na ciałem $\mathbb{R}$?

**Zadanie 35.5.** Sprawdzić, czy liczby 1, $\sqrt{2}$, $\sqrt{3}$ są liniowo niezależne w $\mathbb{R}$ nad ciałem $\mathbb{Q}$.

**Zadanie 35.6.** Sprawdzić, że funkcje $\sin x$ i $\cos x$ są liniowo niezależne w przestrzeni funkcji $f : \mathbb{R} \to \mathbb{R}$.

**Zadanie 35.7.** Niech $v_1, v_2, v_3, v_4 \in \mathbb{R}^5$ będą liniowo niezależne. Które z poniższych zdań jest prawdziwe:

- $v_1 + v_2$, $v_2 + v_3$, $v_3 + v_4$, $v_4 + v_1$ są liniowo niezależne.
- $v_1 - v_2$, $v_2 - v_3$, $v_3 - v_4$, $v_4 - v_1$ są liniowo niezależne.
- $v_1 + v_2$, $v_2 + v_3$, $v_3 + v_4$, $v_4 - v_1$ są liniowo niezależne.
- $v_1 + v_2$, $v_2 + v_3$, $v_3 - v_4$, $v_4 - v_1$ są liniowo niezależne.

**Zadanie 35.8.** Niech $v_1, v_2, v_3 \in \mathbb{R}^n$ będą liniowo niezależne. Dla jakich wartości $k \in \mathbb{R}$ wektory $v_2 - v_1$, $kv_3 - v_2$, $v_1 - v_3$ są liniowo niezależne.

**Zadanie 35.9.** Niech $v_1, v_2, v_3 \in \mathbb{R}^n$ będą liniowo zależne i $v_2, v_3, v_4$ są liniowo niezależne. Pokazać, że $v_1 \in \operatorname{span}\{v_2, v_3\}$ i $v_4 \notin \operatorname{span}\{v_1, v_2, v_3\}$.

**Zadanie 35.10.** Niech $v_1, v_2, \dots, v_n$ będzie bazą $\mathbb{R}^n$. Pokazać, że

$$v_1, v_1 + v_2, \dots, v_1 + \dots + v_n$$

jest też bazą. Czy

$$v_1 + v_2, v_2 + v_3, \dots, v_{n-1} + v_n, v_n + v_1$$

również tworzą bazę?

**Zadanie 35.11.** Jeśli $V$, $W$ są przestrzeniami wektorowymi nad $\mathbb{F}$, to jest nią również iloczyn kartezjański $V \times W$ z działaniami

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2), \quad t \cdot (v, w) = (t \cdot v, t \cdot w).$$

Pokazać, że jeśli $V, W$ są skończenie wymiarowe, to $\dim(V \times W) = \dim V + \dim W$.

**Zadanie 35.12.** Wskazać dowolną bazę $\mathbb{R}^3$ zawierającą wektory $[1,5,2]^T$, $[1,4,1]^T$ i zawartą w zbiorze

$$\{[1,5,2]^T, [1,4,1]^T, [2,9,3]^T, [1,7,5]^T\}.$$

**Zadanie 35.13.** Obliczyć $A + 3B$ oraz $2A - 3B$ dla macierzy

$$A = \begin{bmatrix} 3 & 1 & 4 \\ -2 & 0 & 1 \\ 1 & 2 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 2 \\ -3 & 1 & 1 \\ 2 & -4 & 1 \end{bmatrix}.$$

**Zadanie 35.14.** Znaleźć bazę dla przestrzeni wektorowej $(M_{k\times n}(\mathbb{R}), +, \cdot)$ i uzasadnić, że $\dim M_{k\times n}(\mathbb{R}) = nk$.

**Zadanie 35.15.** Pokaż, że wektory $u = [1,1,1]^T, v = [3,-1,4]^T \in \mathbb{R}^3$ są liniowo niezależne. Znajdź taki wektor $w \in \mathbb{R}^3$, że $u, v, w$ są bazą dla $\mathbb{R}^3$.

**Zadanie 35.16.** Sprawdzić czy wektory $[1,1,0]^T$, $[1,0,1]^T$, $[0,1,1]^T$ tworzą bazę $\mathbb{Z}_2^3$.

**Zadanie 35.17.** Sprawdzić liniową niezależność wektorów w $\mathbb{R}^2$:

- $u = [2,1]^T$, $v = [3,2]^T$
- $u = [2,3]^T$, $v = [4,6]^T$
- $u = [-2,1]^T$, $v = [1,3]^T$, $w = [2,4]^T$
- $u = [1,2]^T$, $v = [-1,1]^T$
- $u = [-1,2]^T$, $v = [1,-2]^T$, $w = [2,-4]^T$.

**Zadanie 35.18.** Sprawdzić liniową niezależność wektorów w $\mathbb{R}^3$:

- $u = [1,0,0]^T$, $v = [0,1,1]^T$, $w = [1,0,1]^T$
- $u = [1,0,0]^T$, $v = [0,1,1]^T$, $w = [1,0,1]^T$, $p = [1,2,3]^T$
- $u = [2,1,-2]^T$, $v = [3,2,-2]^T$, $w = [2,2,0]^T$
- $u = [2,1,-2]^T$, $v = [-2,-1,2]^T$, $w = [4,2,-4]^T$
- $u = [1,1,3]^T$, $v = [0,2,1]^T$.

**Zadanie 35.19.** Niech $u = [-1,2,3]^T$, $v = [3,4,2]^T$, $w = [2,6,6]^T$, $p = [-9,-2,5]^T \in \mathbb{R}^3$. Sprawdzić czy

- $w \in \operatorname{span}\{u, v\}$,
- $p \in \operatorname{span}\{u, v\}$.

**Zadanie 35.20.** Znaleźć podprzestrzenie $\mathrm{span}\{v_1, \ldots, v_k\}$ dla poniższych zbiorów wektorów $\{v_1, \ldots, v_k\}$ w podanych przestrzeniach wektorowych:

- $\{\sqrt{2}\}$ dla przestrzeni $\mathbb{R}$ nad ciałem $\mathbb{Q}$,
- $\{\sqrt{2}\}$ dla przestrzeni $\mathbb{R}$ nad ciałem $\mathbb{R}$,
- $\{\sqrt{2} + i\sqrt{2}\}$ dla przestrzeni $\mathbb{C}$ nad ciałem $\mathbb{C}$,
- $\{\sqrt{2} + i\sqrt{2}\}$ dla przestrzeni $\mathbb{C}$ nad ciałem $\mathbb{R}$,

**Zadanie 35.21.** Które z poniższych zbiorów są podprzestrzeniami wektorowymi

- $S = \{[x_1, x_2, x_3]^T \in \mathbb{R}^3 : x_1 = x_2\}$,
- $S = \{[x, 1]^T \in \mathbb{R}^2 : x \in \mathbb{R}\}$,
- $S = \{[x, y]^T \in \mathbb{R}^2 : y = 0\}$,
- $S = \{[x, y]^T \in \mathbb{R}^2 : xy = 0\}$,
- $S = \{[x, y]^T \in \mathbb{R}^2 : |x| = |y|\}$,
- $S = \{[x, y, z]^T \in \mathbb{R}^3 : x + z = 1\}$,
- $S = \{[x, y, z]^T \in \mathbb{R}^3 : z = x + y\}$,
- $S = \{[x, y, z]^T \in \mathbb{R}^3 : z = x \ \text{ lub } \ z = y\}$.

**Zadanie 35.22.** Niech $A \in M_{2\times 2}(\mathbb{F})$. Pokazać, że

$$V = \{B \in M_{2\times 2}(\mathbb{F}) : AB = BA\}$$

jest podprzestrzenią $M_{2\times 2}(\mathbb{F})$. Dla macierzy $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ i $\mathbb{F} = \mathbb{R}$ znaleźć wymiar $V$.

**Zadanie 35.23.** Niech $v_1, v_2, v_3, v_4, v_5 \in \mathbb{R}^3$. Pokazać, że istnieją takie skalary $a_1, a_2, a_3, a_4, a_5 \in \mathbb{R}$, że zachodzą warunki

- nie wszystkie $a_i$ są równe zero,
- $a_1 + a_2 + a_3 + a_4 + a_5 = 0$,
- $a_1v_1 + a_2v_2a_3v_3 + a_4v_4 + a_5v_5 = 0$.

**Zadanie 35.24.** Niech $v_1, \ldots, v_{n+1} \in \mathbb{R}^n$ będą takimi wektorami, że $(v_i|v_j) < 0$ dla $i \neq j$. Pokazać, że każde $n$ wektorów spośród nich tworzy bazę $\mathbb{R}^n$.

**Zadanie 35.25.** Uzasadnić, że jeżeli $v_1, \ldots, v_k \in V$ są liniowo niezależne, to $v_i \neq 0$ dla każdego $i = 1, \ldots, k$.

**Zadanie 35.26.** Niech $v_1, \ldots, v_m \in \mathbb{R}^n$ będą takimi wektorami, że $(v_i|v_j) < 0$ dla $i \neq j$. Pokazać, że $m \leqslant n + 1$.

**Zadanie 35.27.** Niech $U, V \subset \mathbb{R}^n$ będą podprzestrzeniami wektorowaymi. Które ze zbiorów $U \cap V$, $U \cup V$, $U + V = \{u + v : u \in U, v \in V\}$ są podprzestrzeniami wektorowymi.

**Zadanie 35.28.** Rozważmy przestrzeń wektorową $P_4$ wielomianów stopnia co najwyżej 4. Które z poniższych podzbiorów $P_4$ są podprzestrzeniami wektorowymi (bądźcie ostrożni!):

- $S = \{w \in P_4 : \deg w \text{ jest liczbą parzystą}\}$;
- $S = \{w \in P_4 : \deg w = 3\}$;
- $S = \{p \in P_4 : p(0) = 0\}$.

**Zadanie 35.29.** Niech $P_2$ będzie przestrzenią wielomianów stopnia co najwyżej 2. Który z poniższych podzbiorów $P_2$ generuje $P_2$?

- $\{1, x^2, x^2 - 2\}$,
- $\{2, x^2, x, 2x + 3\}$,
- $\{x + 2, x + 1, x^2 - 1\}$,
- $\{x + 2, x^2 - 1\}$.

**Zadanie 35.30.** Niech $W_1, W_2 \subset \mathbb{R}^n$ będą podprzestrzeniami wektorowymi oraz $\{0\} \neq W_i \neq \mathbb{R}^n$ dla $i = 1, 2$. Pokazać, że istnieje taki wektor $v \in \mathbb{R}^n$, że $v \notin W_1$ i $v \notin W_2$.

**Zadanie 35.31.** Niech

$$W = \{[z_1, z_2, z_3, z_4]^T \in \mathbb{C}^4 : z_3 = z_1 + z_2,\ z_4 = z_1 - z_2\}.$$

- Pokazać, że $W$ jest podprzestrzenią wektorową $\mathbb{C}^4$.
- Znaleźć bazę i wymiar $W$.
- Czy $\{t \cdot [1, 0, 1, 1]^T : t \in \mathbb{C}\}$ jest podprzestrzenią $W$?

**Zadanie 35.32.** Sprawdzić, czy podane wektory są liniowo niezależne w $P_2$:

- $1, x^2, x^2 - 2$;
- $2, x^2, x, 2x + 3$;
- $x + 2, x + 1, x^2 - 1$;
- $x + 2, x^2 - 1$.

**Zadanie 35.33.** Wyznacz wymiar podprzestrzeni przestrzeni $P_2$ generowanej przez wektory

- $x, x - 1, x^2 + 1$;
- $x, x - 1, x^2 + 1, x^2 - 1$;
- $x^2, x^2 - x - 1, x + 1$;
- $2x, x - 2$.

**Zadanie 35.34.** Niech $S$ będzie podprzestrzenią wektorową przestrzeni $P_2$, złożoną z takich wielomianów $p$, że $p(0) = 0$ i niech $T$ będzie podprzestrzenią wektorową $P_2$, złożoną z takich wielomianów $q$, że $q(1) = 0$ (sprawdzić, że są to faktycznie podprzestrzenie). Znaleźć bazy dla $S$, $T$ i $S \cap T$.

**Zadanie 35.35.** Niech $u, v, w \in \mathbb{R}^n$ będą liniowo niezależne. Pokazać, że $u+v, u+w, v+w$ są liniowo niezależne.

**Zadanie 35.36.** Rozważmy zbiór $V \subset M_{3\times 3}(\mathbb{R})$ złożony z takich macierzy $A$, że

$$\sum_{r=1}^{3} a_{rj} = k = \sum_{r=1}^{3} a_{ir}, \quad i, j = 1, 2, 3$$

dla pewnej stałej $k \in \mathbb{R}$. Pokazać, że $V$ jest podprzestrzenią wektorową. Znaleźć wymiar i jakąś „naturalną bazę” dla $V$. Rozwiązać analogiczny problem jeśli macierze z $V$ spełniają dodatkowo warunek

$$\sum_{j=1}^{3} a_{jj} = K = \sum_{i=1}^{3} a_{i,3-i}.$$

**Zadanie 35.37.** Niech $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \in M_{2n\times 2n}(\mathbb{R})$. Niech $V \in M_{2n\times 2n}(\mathbb{R})$ będzie zbiorem takich macierzy $A$, że $A^T J + JA = 0$. Pokazać, że jest to przestrzeń wektorowa i znaleźć jej wymiar.

**Zadanie 35.38.** Pokazać, że dla podprzestrzeni $V, W \subset \mathbb{R}^n$ zachodzi równość

$$\dim(V + W) = \dim V + \dim W - \dim(V \cap W).$$

Ponadto, jeśli $V \not\subset W$ i $W \not\subset V$, to

$$\dim(V + W) \geqslant 2 + \dim(V \cap W).$$

**Zadanie 35.39.** Podać przykład takich trzech podprzestrzeni wektorowych w przestrzeni $V$, że

$$W_1 \cap (W_2 + W_3) \neq (W_1 \cap W_2) + (W_1 \cap W_3).$$

**Zadanie 35.40.** Niech $V \subset \mathbb{R}^n$ będzie niezerową podprzestrzenią. Pokazać, że jeśli istnieje dokładnie jedna taka podprzestrzeń $W \subset \mathbb{R}^n$, że $\mathbb{R}^n = V \oplus W$, to $V = \mathbb{R}^n$.

**Zadanie 35.41.** Niech $V = \{A \in M_{2\times 2}(\mathbb{R}) : \operatorname{tr} A = 0\}$. Znaleźć $\dim V$. Wskazać bazę $V$.

**Zadanie 35.42.** Sprawdzić, czy poniższe funkcje są liniowo niezależne w przestrzeni funkcji $f : \mathbb{R} \to \mathbb{R}$:

- $\sin x$ oraz $\cos x$,
- $\sin^2 x$ oraz $\cos^2 x$.

**Zadanie 35.43.** Niech

$$V = \Big\{ \begin{bmatrix} u & -u-x \\ 0 & x \end{bmatrix} : u, x \in \mathbb{R} \Big\},$$

$$W = \Big\{ \begin{bmatrix} v & 0 \\ w & -v \end{bmatrix} : v, w \in \mathbb{R} \Big\}.$$

Znaleźć bazy dla $V + W$ oraz $V \cap W$.

**Zadanie 35.44.** Uzasadnić, że przestrzeń wektorowa $\mathbb{R}$ nad $\mathbb{F} = \mathbb{Q}$ jest nieskończenie wymiarowa.

**Zadanie 35.45.** Rozważmy przestrzeń wektorową $\mathcal{F}(\mathbb{R}, \mathbb{R})$ wszystkich funkcji $\mathbb{R} \to \mathbb{R}$ (z naturalnie zdefiniowanymi działaniami). Które z poniższych zbiorów są jej podprzestrzeniami

- Zbiór $C(\mathbb{R}) \subset \mathcal{F}(\mathbb{R}, \mathbb{R})$ wszystkich funkcji ciągłych,
- Zbiór $L^\infty(\mathbb{R}) \subset \mathcal{F}(\mathbb{R}, \mathbb{R})$ funkcji ograniczonych,
- $C(\mathbb{R}) \cap L^\infty(\mathbb{R})$,
- Zbiór takich funkcji $f \in \mathcal{F}(\mathbb{R}, \mathbb{R})$, że $f(0) = 1$,
- Zbiór takich funkcji $f \in \mathcal{F}(\mathbb{R}, \mathbb{R})$, że $f(1) = 0$,
- Zbiór takich funkcji $f \in \mathcal{F}(\mathbb{R}, \mathbb{R})$, że $f(0) = f(1)$,
- Zbiór takich funkcji $f \in \mathcal{F}(\mathbb{R}, \mathbb{R})$, że $f(\mathbb{R}) \subset \mathbb{Q}$,
- Zbiór $C^1(\mathbb{R}) \subset \mathcal{F}(\mathbb{R}, \mathbb{R})$ wszystkich funkcji klasy $C^1$.

**Zadanie 35.46.** Niech $V$ będzie skończenie wymiarową przestrzenią wektorową nad $\mathbb{R}$ i niech $S \subset V$ będzie jej podprzestrzenią. Pokazać, że

- $\dim S \leqslant \dim V$,

- $\dim S = \dim V$ wtedy i tylko wtedy, gdy $S = V$,
- Każda baza dla $S$ jest zawarta w bazie dla $V$,
- Baza dla $V$ nie musi zawierać bazy dla $S$.

**Zadanie 35.47.** Sprawdzić, które z poniższych rzeczywistych przestrzeni wektorowych mają skończony wymiar:

- przestrzeń wszystkich ciągów o wyrazach rzeczywistych,
- przestrzeń ciągów rzeczywistych zbieżnych do 0,
- przestrzeń takich ciągów rzeczywistych $(a_n)$, że $a_n = 0$ dla $n > 2017$.

**Zadanie 35.48.** Niech $V$ będzie przestrzenią wektorową wszystkich nieskończonych ciągów $[x_1, x_2, x_3, \ldots]^T$ o wyrazach z ciała $\mathbb{F}$ z naturalnie zdefiniowanymi działaniami. Rozważmy ciąg wektorów

$$e_1 = [1, 0, 0, \ldots]^T, \quad e_2 = [0, 1, 0, \ldots]^T, \quad \ldots$$

Pokazać, że

- każdy skończony podzbiór zbioru $\{e_1, e_2, \ldots\}$ składa się z wektorów liniowo niezależnych,
- uzasadnić, że $e_1, e_2, \ldots$ nie jest bazą $V$,
- opisać zbiór wszystkich kombinacji liniowych wektorów $e_1, e_2, \ldots$.

**Zadanie 35.49.** Uzasadnić, że $\dim_{\mathbb{F}} \; Sym_n(\mathbb{F}) = \frac{n(n+1)}{2}$. Znaleźć wymiar przestrzeni $n \times n$ macierzy antysymetrycznych.

# Rozdział 36

# Odwzorowania liniowe

**Zadanie 36.1.** Niech $L : \mathbb{R}^n \to \mathbb{R}^2$ będzie odwzorowaniem liniowym, $v_1, v_2 \in \mathbb{R}^n$ oraz $L(v_1) = [1,1]^T$, $L(v_2) = [-1,2]^T$. Obliczyć $L(v)$ dla wektora $v = 3v_1 - 4v_2$.

**Zadanie 36.2.** Niech $L : \mathbb{R}^3 \to \mathbb{R}^2$ będzie odwzorowaniem liniowym oraz $L[2,3,4]^T = [2,1]^T$ i $L[1,2,3]^T = [4,5]^T$. Obliczyć $L([7,10,13]^T)$.

**Zadanie 36.3.** Niech $L$ będzie odwzorowaniem liniowym $L : \mathbb{R}^3 \to \mathbb{R}^3$ zadanym wzorem

$$L([x,y,z]^T) = [x+2y+z, x+3y+2z, 2x+3y-z]^T.$$

Uzasadnić, że

$$\dim\ker(L) = 1, \quad \dim\operatorname{im}(L) = 2.$$

Oznacza to, że $\ker(L)$ jest prostą w $\mathbb{R}^3$. Znaleźć jej równanie parametryczne. Wyznaczyć równanie ogólne płaszczyzny będącej obrazem $L$.

**Zadanie 36.4.** Niech $L : \mathbb{R}^3 \to \mathbb{R}^3$ będzie takim odwzorowaniem liniowym, że

$$L([1,6,9]^T) = [2,4,7]^T, \quad L([1,2,4]^T) = [9,-3,1]^T,$$
$$L([5,22,34]^T) = [16,14,1]^T.$$

Znaleźć wzór ogolny na $L([x,y,z]^T)$. Wyznaczyć jądro i obraz $L$.

**Zadanie 36.5.** Załóżmy, że $V, W$ są rzeczywistymi przestrzeniami wektorowymi. Niech $\dim V = 2$ oraz $\dim W = 3$. Uzasadnić, że przestrzeń odwzorowań liniowych $\mathcal{L}(V,W)$ jest izomorficzna z przestrzenią macierzy $M_{3\times 2}(\mathbb{R})$.

**Zadanie 36.6.** Niech $V \subset \mathbb{R}^n$ będzie podprzestrzenią wektorową. Pokazać, że istnieje takie odwzorowanie liniowe $L : \mathbb{R}^n \to \mathbb{R}^n$, że $\ker L = V$.

**Zadanie 36.7.** Niech $V \subset \mathbb{R}^n$ będzie podprzestrzenią wektorową. Pokazać, że istnieje takie odwzorowanie liniowe $L : \mathbb{R}^n \to \mathbb{R}^n$, że $\operatorname{im} L = V$.

**Zadanie 36.8.** Niech $f : \mathbb{R}^n \to \mathbb{R}^m$ jest takim odwzorowaniem niezerowym, że $\|f(x)\| \leqslant 1$ dla $x \in \mathbb{R}^n$. Pokazać, że $f$ nie jest odwzorowaniem liniowym.

**Zadanie 36.9.** Czy odwzorowanie $L([x,y]^T) = [x+y, xy]^T$ ($[x,y]^T \in \mathbb{R}^2$) jest liniowe?

**Zadanie 36.10.** Dla ustalonych $a, b \in \mathbb{R}$, rozważmy odwzorowanie

$$L : \mathbb{C} \ni \alpha + i\beta \mapsto a\alpha + ib\beta \in \mathbb{C}$$

przestrzeni wektorowej $\mathbb{C}$ nad $\mathbb{C}$. Dla jakich $a, b \in \mathbb{R}$ jest ono liniowe.

**Zadanie 36.11.** Które z poniższych rzeczywistych przestrzeni wektorowych są izomorficzne (mają ten sam wymiar)?

- $\mathbb{R}^2$ oraz $\{(x_1,x_2,x_3,x_4) \in \mathbb{R}^4 : x_1 = x_2 = x_3 = x_4\}$,

- $\mathbb{R}^2$ oraz $\{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1 = x_2 = x_3\}$,
- $\mathbb{R}^2$ oraz $\{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1 = x_2 i x_3 = x_4\}$.

**Zadanie 36.12.** Niech $L : V \to W$ będzie odwzorowaniem liniowym. Pokazać, że jeśli $S \subset W$ jest podprzestrzenią wektorową $W$, to przeciwobraz $L^{-1}(S) = \{v \in V : L(v) \in S\}$ jest podprzestrzenią wektorową $V$.

**Zadanie 36.13.** Niech $L : \mathbb{R}^n \to \mathbb{R}^n$ będzie odwzorowaniem liniowym oraz $v_1, \ldots, v_k \in \mathbb{R}^k$. Jeśli któreś z poniższych zdań jest prawdziwe, to je uzasadnić.

- Jeśli $v_1, \ldots, v_k$ są liniowo niezależne, to $L(v_1), \ldots, L(v_k)$ są liniowo niezależne.
- Jeśli $L(v_1), \ldots, L(v_k)$ są liniowo niezależne, to $v_1, \ldots, v_k$ są liniowo niezależne.

**Zadanie 36.14.** Dla odwzorowań liniowych

$$L : \mathbb{R}^2 \ni [x, y]^T \mapsto [2x + 6y, x + 3y]^T \in \mathbb{R}^2,$$

$$L' : \mathbb{R}^2 \ni [x, y]^T \mapsto [x + 2y, 2x + 4y]^T \in \mathbb{R}^2,$$

znaleźć takie, izomorfizmy $G, H : \mathbb{R}^2 \to \mathbb{R}^2$, że $L' \circ G = H \circ L$.

**Zadanie 36.15.** Niech $m < n$ będą liczbami naturalnymi. Pokazać, że jądro dowolnego epimorfizmu $L : \mathbb{F}^n \to \mathbb{F}^m$ jest izomorficzne z $\mathbb{F}^{n-m}$. Podać konkretny przykład takiego epimorfizmu oraz wskazać izomorfizm pomiędzy jego jądrem a $\mathbb{F}^{n-m}$.

**Zadanie 36.16.** Sprawdzić, że odwzorowanie

$$L : M_{n\times n}(\mathbb{R}) \ni A \mapsto \operatorname{tr} A \in \mathbb{R}$$

jest liniowe.

**Zadanie 36.17.** Uzasadnić, że przestrzenie $\mathbb{R}^{nm}$ oraz $M_{n\times m}(\mathbb{R})$ są izomorficzne.

**Zadanie 36.18.** Wykres odwzorowania $f : \mathbb{R}^n \to \mathbb{R}^m$ definiujemy jako zbiór

$$G_f = \{(x, f(x)) : x \in \mathbb{R}^n\} \subset \mathbb{R}^{n+m}.$$

Uzasadnić, że $f$ jest liniowe wtedy i tylko wtedy, gdy $G_f$ jest podprzestrzenią wektorową $\mathbb{R}^{n+m}$.

**Zadanie 36.19.** Niech $L : \mathbb{R}^n \to \mathbb{R}^n$. Pokazać, że obrazem prostej przez $L$ jest albo prosta albo punkt.

**Zadanie 36.20.** Niech $L : V \to W$ będzie odwzorowaniem liniowym. Pokazać, że jeśli $S \subset W$ jest podprzestrzenią wektorową $W$, to przeciwobraz $L^{-1}(S) = \{v \in V : L(v) \in S\}$ jest podprzestrzenią wektorową $V$.

**Zadanie 36.21.** Niech $L : \mathbb{R}^n \to \mathbb{R}^m$ będzie odwzorowaniem liniowym. Pokazać, że jeśli $V \subset \mathbb{R}^n$, $W \subset \mathbb{R}^m$ są podprzestrzeniami, to

- $L^{-1}(L(V)) = V + \ker f$,
- $L(L^{-1}(W)) = W \cap \operatorname{im} L$.

**Zadanie 36.22.** Podać przykład przestrzeni wektorowej $V$ nieskończonego wymiaru i takich odwzorowań liniowych $L, G : V \to V$, że

- $L$ jest monomorfizmem, ale nie jest epimorfizmem,
- $G$ jest epimorfizmem, ale nie jest monomorfizmem.

**Zadanie 36.23.** Niech $L : V \to W$ będzie odwzorowaniem liniowym. Załóżmy, że $v_1, \ldots, v_n \in V$ są takie, że

- $\ker L \subset \operatorname{span}\{v_1, \ldots, v_n\}$,
- $\operatorname{span}\{L(v_1), \ldots, L(v_n)\} = W$.

Pokazać, że $\operatorname{span}\{v_1, \ldots, v_n\} = V$.

**Zadanie 36.24.** Niech

$$L_1 : V_1 \to V_2, \quad L_2 : V_2 \to V_3, \quad L_3 : V_3 \to V_4$$

będą odwzorowaniami liniowymi oraz

$$\dim V_1 = 8, \quad \dim V_2 = 5, \quad \dim V_3 = 7, \quad \dim V_4 = 6.$$

Pokazać, że $L_3 \circ L_2 \circ L_1 : V_1 \to V_4$ nie jest epimorfizmem.

**Zadanie 36.25.** Niech $L : V \to V$ będzie odwzorowaniem liniowym. Pokazać, że $L^2 = 0$ wtedy i tylko wtedy, gdy $\operatorname{im} L \subset \ker L$.

**Zadanie 36.26.** Załóżmy, że $L, S : V \to V$ są odwzorowaniami liniowymi ora $S$ jest izomorfizmem. Pokazać, że

$$\operatorname{im}(SLS^{-1}) = S(\operatorname{im} L), \quad \ker(SLS^{-1}) = S(\ker L)$$

**Zadanie 36.27.** Niech $L : \mathbb{R}^n \to \mathbb{R}^n$ będzie odwzorowaniem liniowym. Pokazać, że

$$\dim V + \dim L^2(V) \geqslant 2 \dim L(V).$$

# Rozdział 37

# Macierze

**Zadanie 37.1.** Dla podanych macierzy $A$ i $B$ oblicz iloczyny $AB$ i $BA$, o ile są one wykonalne:

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

**Zadanie 37.2.** Niech $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Oblicz $A^2$, $A^3$. Zgadnij wzór na $A^n$ i go uzasadnij.

**Zadanie 37.3.** Podać przykład takich macierzy niezerowych $A$, $B \in M_{2\times 2}(\mathbb{R})$, że $AB = 0$.

**Zadanie 37.4.** Niech $A \in M_{2\times 2}(\mathbb{C})$. Pokazać, że $A^2 = 0$ wtedy i tylko wtedy, gdy $A$ ma jedną z postaci

$$\begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ a & 0 \end{bmatrix}, \quad \begin{bmatrix} a & ab \\ -ab^{-1} & -a \end{bmatrix},$$

dla pewnych $a, b \in \mathbb{C}$ z $b \neq 0$.

**Zadanie 37.5.** Podać przykład takich niezerowych macierzy $A$, $B$, $C \in M_{2\times 2}(\mathbb{R})$, że $A \neq B$ oraz $AC = BC$.

**Zadanie 37.6.** Niech

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Sprawdzić, że $A^n = 0$ dla $n \geqslant 4$.

**Zadanie 37.7.** Uzasadnić na podstawie definicji, że macierz $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ nie posiada macierzy odwrotnej.

**Zadanie 37.8.** Pokazać, że jeśli $A \in M_{n\times n}(\mathbb{F})$ jest nieosobliwa, to $A^m$ jest nieosobliwa dla dowolnego $m \in \mathbb{N}$ oraz $(A^m)^{-1} = (A^{-1})^m$.

**Zadanie 37.9.** Niech macierz $B \in M_{n\times n}(\mathbb{R})$ będzie nieosobliwa. Pokazać, że odwzorowanie

$$L : M_{n\times n}(\mathbb{R}) \ni A \mapsto AB \in M_{n\times n}(\mathbb{R})$$

jest izomorfizmem.

**Zadanie 37.10.** Sprawdzić, że dla $A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$ zachodzi $A^2 = 0$. Czy istnieje taka niezerowa macierz symetryczna $B \in M_{2\times 2}(\mathbb{R})$, że $B^2 = 0$?

**Zadanie 37.11.** Niech $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ oraz $a_{11} \neq 0$ i $\alpha = a_{21}/a_{11}$. Pokazać, że $A$ można zapisać w postaci iloczynu postaci

$$A = \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ 0 & b \end{bmatrix}.$$

Ile jest równe $b$ w tym przedstawieniu?

**Zadanie 37.12.** Pokazać, że dla $A = \begin{bmatrix} 7 & 4 \\ -9 & -5 \end{bmatrix}$ zachodzi wzór

$$A^n = \begin{bmatrix} 1+6n & 4n \\ -9n & 1-6n \end{bmatrix}, \quad n \geqslant 1.$$

**Zadanie 37.13.** Załóżmy, że $A \in M_{2\times 2}(\mathbb{Z})$ jest macierzą o wyrazach całkowitych. Pokazać, że jeśli $p \in \mathbb{N}$ jest liczbą pierwszą, to liczba $\operatorname{tr} A^p - \operatorname{tr} A$ dzieli się przez $p$.

**Zadanie 37.14.** Dla macierzy $A \in M_{2\times 2}(\mathbb{F})$ zachodzą wzory

- $$\det(A) = \frac{1}{2} \det \begin{bmatrix} \operatorname{tr}(A) & 1 \\ \operatorname{tr}(A^2) & \operatorname{tr}(A) \end{bmatrix}.$$
- $\operatorname{tr} A^2 = (\operatorname{tr} A)^2 - 2 \det A$,
- $\operatorname{tr} A^3 = (\operatorname{tr} A)^3 - 3(\det A)(\operatorname{tr} A)$.

**Zadanie 37.15.** Niech $A, B, C \in M_{n\times n}(\mathbb{R})$. Pokazać, że

- $\operatorname{tr}(ABC) = \operatorname{tr}(BCA) = \operatorname{tr}(CAB)$,
- $\operatorname{tr}(ABC)$ może być różne od $\operatorname{tr}(BAC)$,
- $\operatorname{tr}(A^T B) = \operatorname{tr}(AB^T)$.

**Zadanie 37.16.** Niech $A \in M_{n\times n}(\mathbb{R})$. Pokazać, że $\operatorname{tr} A^T A = 0$ wtedy i tylko wtedy, gdy $A = 0$.

**Zadanie 37.17.** Niech $A \in M_{n\times n}(\mathbb{R})$ będzie skośnie symetryczna. Pokazać, że

- $I - A$ jest nieosobliwa,
- jeśli $B = (I + A)(I - A)^{-1}$, to $B^{-1} = B^T$.

**Zadanie 37.18.** Niech $A, B, A + B \in M_{n\times n}(\mathbb{R})$ będą nieosobliwe. Pokazać, że

$$A(A + B)^{-1}B = B(A + B)^{-1}A = (A^{-1} + B^{-1})^{-1}.$$

**Zadanie 37.19.** Niech $A \in M_{m\times n}(\mathbb{R})$ i $B \in M_{n\times m}(\mathbb{R})$ będą takie, że

$$AB = I_m, \quad BA = I_n.$$

Pokazać, że $n = m$.

**Zadanie 37.20.** Znaleźć macierz odwrotną do macierzy $A = \begin{bmatrix} 2 & 1 & 2 \\ 3 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}$.

**Zadanie 37.21.** Pokazać, że jeśli $A, B \in M_{n\times n}(\mathbb{F})$ oraz macierze $A$, $B$ i $A+B$ są nieosobliwe, to

$$A(A^{-1}+B^{-1})B(A+B)^{-1} = I.$$

Co można powiedzieć o odwracalności macierzy $A^{-1}+B^{-1}$?

**Zadanie 37.22.** Jeśli $A$ jest macierzą kwadratową oraz $A^k = 0$ dla $k > 1$, to $I - A$ jest nieosobliwa oraz

$$(I-A)^{-1} = I + A + \ldots + A^{k-1}.$$

**Zadanie 37.23.** Uzasadnić, że jeśli $A^T A = A$, to $A$ jest symetryczna oraz $A = A^2$.

**Zadanie 37.24.** Niech $A \in M_{n\times n}(\mathbb{R})$ oraz $A^2 = 0$. Pokazać, że $\operatorname{rank} A \leqslant n/2$.

**Zadanie 37.25.** Pokazać, że jeśli $A$ jest antysymetryczna ($A^T = -A$), to $A+I$ jest nieosobliwa.

**Zadanie 37.26.** Pokazać, że

$$\operatorname{rank}(AB) \leqslant \min\{\operatorname{rank} A, \operatorname{rank} B\},$$

dla $A, B \in M_{n\times n}(\mathbb{R})$.

**Zadanie 37.27.** Pokazać, że

$$\operatorname{rank}(A+B) \leqslant \operatorname{rank} A + \operatorname{rank} B,$$

dla $A, B \in M_{m\times n}(\mathbb{R})$.

**Zadanie 37.28.** Podać przykłady takich macierzy $A, B \in M_{2\times 2}(\mathbb{R})$, że

- $\operatorname{rank}(A+B) < \operatorname{rank} A$ i $\operatorname{rank}(A+B) < \operatorname{rank} B$,
- $\operatorname{rank}(A+B) = \operatorname{rank} A$ i $\operatorname{rank}(A+B) = \operatorname{rank} B$,
- $\operatorname{rank}(A+B) > \operatorname{rank} A$ i $\operatorname{rank}(A+B) > \operatorname{rank} B$.

**Zadanie 37.29.** Niech $A, B \in M_{n\times n}(\mathbb{R})$. Pokazać, że jeśli $B^T A = 0$, to

$$\operatorname{rank}(A+B) = \operatorname{rank} A + \operatorname{rank} B.$$

**Zadanie 37.30.** Niech $A, B \in M_{n\times n}(\mathbb{R})$ i $n$ jest nieparzyste. Pokazać, że jeśli $AB = 0$, to przynajmniej jedna z macierzy $A + A^T$ i $B + B^T$ jest osobliwa.

**Zadanie 37.31.** Niech $A, B \in M_{n\times n}(\mathbb{R})$ będą takie, że $AB = 0$. Pokazać, że

$$\operatorname{rank} A + \operatorname{rank} B \leqslant n.$$

**Zadanie 37.32.** Uzasadnić, że jeśli $A \in M_{m\times n}(\mathbb{R})$ i $B \in M_{n\times p}(\mathbb{R})$, to

$$\operatorname{rank}(AB) = \operatorname{rank}(B) - \dim(\ker A \cap \operatorname{im} B).$$

**Zadanie 37.33.** Dla macierzy kwadratowej $X \in M_{n\times n}(\mathbb{R})$ oznaczamy $\nu(X) = \dim\ker X$. Uzasadnić prawo Sylwestera

$$\max\{\nu(A), \nu(B)\} \leqslant \nu(AB) \leqslant \nu(A) + \nu(B).$$

**Zadanie 37.34.** Pokazać, że dla macierzy $A \in M_{n\times n}(\mathbb{R})$ następujące warunki są równoważne

(a) $\ker A = \ker A^2$,

(b) $\operatorname{im} A = \operatorname{im} A^2$,

(c) $\ker A \cap \operatorname{im} A = \{0\}$.

**Zadanie 37.35.** Niech $A \in M_{n\times n}(\mathbb{R})$ będzie taką macierzą, że dla dowolnej macierzy $B \in M_{n\times n}(\mathbb{R})$ mamy

$$AB = BA.$$

Pokazać, że $A = \lambda I$ dla pewnego $\lambda \in \mathbb{R}$.

**Zadanie 37.36.** Znaleźć rozkład $LU$ macierzy

$$\begin{bmatrix} 2 & -4 \\ 3 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 8 & 7 & 9 \end{bmatrix}, \quad \begin{bmatrix} 2 & 2 & -1 \\ 4 & 0 & 4 \\ 3 & 4 & 4 \end{bmatrix}.$$

**Zadanie 37.37.** Rozważmy przestrzeń wektorową $M_{2\times 2}(\mathbb{R})$. Które z poniższych podzbiorów $S \subset M_{2\times 2}(\mathbb{R})$ jest podprzestrzenią wektorową:

- $S$-zbiór macierzy diagonalnych;
- $S$-zbiór macierzy górnie trójkątnych;
- $S = \{A = [a_{ij}] \in M_{2\times 2}(\mathbb{R}) : a_{12} = 1\}$;
- $S = \{A = [a_{ij}] \in M_{2\times 2}(\mathbb{R}) : a_{11} = 0\}$;
- $S = \{A \in M_{2\times 2}(\mathbb{R}) : A = A^T\}$;
- $S$ zbiór macierzy nieosobliwych;

Jeśli $S$ jest podprzestrzenią, to znajdź jakąś bazę dla $S$ i wyznacz jej wymiar.

**Zadanie 37.38.** Niech $A \in M_{2\times 2}(\mathbb{R})$ będzie ustaloną macierzą. Które z podzbiorów $S \subset M_{2\times 2}(\mathbb{R})$ jest podprzestrzenią wektorową:

- $S = \{B \in M_{2\times 2}(\mathbb{R}) : AB = BA\}$;
- $S = \{B \in M_{2\times 2}(\mathbb{R}) : AB \neq BA\}$;
- $S = \{B \in M_{2\times 2}(\mathbb{R}) : BA = 0\}$.

**Zadanie 37.39.** Znaleźć współrzędne wektora $[7,4]^T$ w bazie $v_1 = [3,2]^T$, $v_2 = [1,1]^T$.

**Zadanie 37.40.** Znaleźć macierz przejścia od bazy $v_1 = [5,2]^T$, $v_2 = [7,3]^T$ do bazy $u_1 = [3,2]^T$, $u_2 = [1,1]^T$.

**Zadanie 37.41.** Rozważmy przestrzeń wektorową $\mathbb{C}$ nad ciałem $\mathbb{R}$.

- Pokazać, że $1$, $i$ oraz $1+i$, $1+2i$ są bazami tej przestrzeni wektorowej.
- Znaleźć macierz przejścia od bazy $1$, $i$ do bazy $1+i$, $1+2i$.

**Zadanie 37.42.** Znaleźć macierz $A$ odwzorowania liniowego

$$L([x_1,x_2,x_3]^T) = [2x_1 - x_2 - x_3, 2x_2 - x_1 - x_3, 2x_3 - x_1 - x_2]^T$$

w bazach standardowych w $\mathbb{R}^3$.

**Zadanie 37.43.** Niech $A \in M_{m\times n}(\mathbb{R})$. Pokazać, że jeśli $\operatorname{tr}(AX) = 0$ dla dowolnej macierzy $X \in M_{n\times m}(\mathbb{R})$, to $A = 0$.

**Zadanie 37.44.** Macierz $A \in M_{n\times n}(\mathbb{R})$ spełnia równanie

$$A^3 - 4A^2 + 3A - 5I = 0.$$

Uzasadnić, że $A$ jest nieosobliwa.

**Zadanie 37.45.** Niech $A = \begin{bmatrix} 1 & 1/3 \\ c & d \end{bmatrix}$. Dla jakich wartości parametrów $c, d \in \mathbb{R}$ zachodzi $A^2 = 0$.

**Zadanie 37.46.** Niech $A, B \in M_{n\times n}(\mathbb{R})$ będą symetryczne. Pokazać, że $AB$ jest symetryczna wtedy i tylko wtedy, gdy $AB = BA$.

**Zadanie 37.47.** Czy iloczyn macierzy symetrycznych jest macierzą symetryczną? Co można powiedzieć o ilocznie macierzy antysymetrycznych?

**Zadanie 37.48.** Uzasadnić, że jeśli $A \in M_{n\times n}(\mathbb{F})$ jest nieosobliwa i symetryczna, to $A^{-1}$ jest symetryczna.

**Zadanie 37.49.** Niech $A \in M_{k\times n}(\mathbb{F})$ będzie dowolną macierzą. Uzasadnić, że iloczyny $AA^T$ i $A^TA$ są określone i są macierzami symetrycznymi.

**Zadanie 37.50.** Niech $A \in M_{n\times n}(\mathbb{R})$. Pokazać, że $A$ jest nieosobliwa wtedy i tylko wtedy, gdy $A^2$ jest nieosobliwa.

**Zadanie 37.51.** Niech $A \in M_{2\times 2}(\mathbb{R})$ będzie taka, że $\operatorname{tr} A = 0$. Pokazać, że $A^2 = aI$ dla pewnego $a \in \mathbb{R}$.

**Zadanie 37.52.** Dla macierzy $A, B \in M_{n\times n}(\mathbb{R})$ definiujemy ich *komutator* (*nawias Poissona*) przez

$$[A, B] = AB - BA \in M_{n\times n}(\mathbb{R}).$$

Pokazać, że

- $[A + B, C] = [A, C] + [B, C]$,
- $[A, A] = 0$,
- $[A, B] = -[B, A]$,
- $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$,
- $[A, BC] = [A, B]C + B[A, C]$,
- $\operatorname{tr}[A, B] = 0$,
- $[A, B] \neq I$.

**Zadanie 37.53.** Znaleźć rząd i jądro macierzy

$$\begin{bmatrix} 1 & 1 & 1 & i \\ 1 & i & 1 & i \\ 0 & 1+i & 0 & 0 \\ 2 & 0 & 2 & 2i \end{bmatrix}$$

**Zadanie 37.54.** Niech $C \in M_{2\times 2}(\mathbb{R})$. Pokazać, że poniższe warunki są równoważne

- $\operatorname{tr} C = 0$,
- $C = [A, B]$ dla pewnych $A, B \in M_{2\times 2}(\mathbb{R})$.

Spróbować udowodnić analogiczną równoważność w dowolnym wymiarze.

**Zadanie 37.55.** Niech $A, B \in M_{m\times n}(\mathbb{R})$ będą takie, że

$$(Ax|y) = (Bx|y), \quad x \in \mathbb{R}^n,\ y \in \mathbb{R}^m.$$

Pokazać, że $A = B$.

**Zadanie 37.56.** Rozważmy zbiór

$$G_{2n} = \{A \in M_{2n\times 2n}(\mathbb{R}) : (JA)^T = JA\},$$

gdzie $J = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}$. Uzasadnić, że

(i) $A \in G_{2n}$ wtedy i tylko wtedy, gdy $JA + A^T J = 0$,

(ii) $G_{2n}$ jest grupą abelową z działaniem dodawania macierzy,

(iii) Dla $A \in G_2$ zachodzi równość

$$JA + A^T J = \begin{bmatrix} 0 & -\operatorname{tr} A \\ \operatorname{tr} A & 0 \end{bmatrix}.$$

W szczególności, $A \in G_2$ wtedy i tylko wtedy, gdy $\operatorname{tr} A = 0$.

Czy (iii) jest prawdą dla $n > 2$?

**Zadanie 37.57.** Pokazać, że zbiór $\mathrm{GL}_n(\mathbb{R})$ wszystkich macierzy nieosobliwych $A \in M_{n\times n}(\mathbb{R})$ jest grupą z działaniem mnożenia macierzy. Uzasadnić, że dla ustalonego wektora $v \in \mathbb{R}^n$ zbiór $\{A \in \mathrm{GL}_n(\mathbb{R}) : Av = v\}$ jest jej podgrupą.

## 37.1 Reprezentacja macierzowa

**Zadanie 37.58.** Załóżmy, że $B$ jest podobna do $A$, czyli $B = S^{-1}AS$. Sprawdzić, że dla każdego $k \in \mathbb{N}$ mamy $B^k = S^{-1}A^kS$.

**Zadanie 37.59.** Uzasadnić, że jeśli macierze $A, B \in M_{n\times n}(\mathbb{R})$ są podobne, to $\operatorname{rank} A = \operatorname{rank} B$. Pokazać na przykładzie, że implikacja odwrotna nie jest prawdziwa.

**Zadanie 37.60.** Niech

$$L : \mathbb{R}^3 \ni [x, y, z]^T \mapsto [x + y + z, 2z - x, 2y - z]^T \in \mathbb{R}^3.$$

Znaleźć reprezentację odwzorowania liniowego $L$ w bazach

$$[2, 01, ]^T, \quad [0, 2, 1]^T, \quad [1, 2, 1]^T$$

oraz

$$[1, 1, 1]^T, \quad [1, 1, 0]^T, \quad [0, 1, 1]^T.$$

**Zadanie 37.61.** Odwzorowanie liniowe $L : \mathbb{R}^2 \to \mathbb{R}^3$ ma w bazach $u_1, u_2$ oraz $v_1, v_2, v_3$ ma macierz

$$\begin{bmatrix} 3 & 2 \\ 4 & 1 \\ 2 & 1 \end{bmatrix}.$$

Obliczyć $L(u)$ jeśli wektor $u \in \mathbb{R}^2$ ma w bazie $u_1, u_2$ współrzędne $[1, 3]^T$.

**Zadanie 37.62.** Obliczyć macierz odwzorowania liniowego

$$L : M_{2\times 2}(\mathbb{R}) \ni A \mapsto \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} A \in M_{2\times 2}(\mathbb{R})$$

w bazie standardowej dla $M_{2\times 2}(\mathbb{R})$.

**Zadanie 37.63.** Rozważmy odwzorowanie liniowe $L : \mathbb{R}^3 \to \mathbb{R}^3$ zadane na bazie standardowej przez

$$L(e_1) = [a, b, b]^T, \quad L(e_2) = [b, a, b]^T, \quad L(e_3) = [b, b, a]^T.$$

Dla jakich wartości $a, b \in \mathbb{R}$ odwzorowanie $L$ ma odpowiednio rząd $0, 1, 2, 3$.

**Zadanie 37.64.** Uzasadnić, że jeśli $B$ jest podobna do $A$, to $\det(B) = \det(A)$.

**Zadanie 37.65.** Pokazać, że jeśli $B$ jest podobna do $A$ i $\lambda \in \mathbb{R}$, to $B - \lambda I$ jest podobna do $A - \lambda I$. W szczególności, $\det(A - \lambda I) = \det(B - \lambda I)$.

**Zadanie 37.66.** Przypuśćmy, że $S, T \in M_{n\times n}(\mathbb{R})$ i $S$ jest nieosobliwa. Pokazać, że $TS$ jest podobna do $ST$.

**Zadanie 37.67.** Niech $A, B \in M_{n\times n}(\mathbb{R})$. Uzasadnić, że jeśli $A$ jest podobna do $B$, to istnieją takie macierze $S, T \in M_{n\times n}(\mathbb{R})$, że $S$ jest nieosobliwa oraz $A = ST$ i $B = TS$.

**Zadanie 37.68.** Niech $V$, $W$ będą przestrzeniami wektorowymi nad ciałem $\mathbb{F}$ wymiaru $n$ i $m$ odpowiednio. Pokazać, że przestrzeń odwzorowań liniowych $\mathcal{L}(V, W)$ jest izomorficzna z przestrzenią macierzy $M_{m\times n}(\mathbb{F})$.

**Zadanie 37.69.** Rozważmy odwzorowanie liniowe

$$L : M_{2\times 2}(\mathbb{F}) \ni A \mapsto A^T \in M_{2\times 2}(\mathbb{F}).$$

Znaleźć macierz $L$ w bazie standardowej $M_{2\times 2}(\mathbb{F})$.

**Zadanie 37.70.** Pokazać, że dla dowolnych macierzy $A, B, C \in M_{n\times n}(\mathbb{F})$ zachodzą warunki

(1) $A$ jest podobna do $A$;

(2) jeśli $B$ jest podobna do $A$, to $A$ jest podobna do $B$;

(3) jeśli $A$ jest podobna do $B$ i $B$ jest podobna do $C$, to $A$ jest podobna do $C$.

**Zadanie 37.71.** Pokazać, że dla $A \in M_{n\times n}(\mathbb{C})$ istnieje taka macierz nieosobliwa $S \in M_{n\times n}(\mathbb{C})$, że $A^T = S^{-1}AS$.

# Rozdział 38

# Wyznacznik

**Zadanie 38.1.** Dla jakich wartości parametru $\lambda \in \mathbb{R}$ macierz $\begin{bmatrix} 2-\lambda & 4 \\ 3 & 3-\lambda \end{bmatrix}$ jest odwracalna?

**Zadanie 38.2.** Niech $A$, $B \in M_{n\times n}(\mathbb{R})$ będą takimi macierzami, że $\det A = 7$ i $\det B = 3$. Obliczyć $\det(2AB)$ oraz $\det(A^{-1}B)$.

**Zadanie 38.3.** Pokazać, że wyznacznik macierzy Vandermonde'a

$$V(x_1, x_2, x_3) = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{bmatrix}$$

jest równy $(x_2 - x_1)(x_3 - x_1)(x_3 - x_2)$.

**Zadanie 38.4.** Niech

$$A = \begin{bmatrix} 1+\lambda_1 & 1 & 1 \\ 1 & 1+\lambda_2 & 1 \\ 1 & 1 & 1+\lambda_3 \end{bmatrix}, \quad \lambda_i \neq 0.$$

Sprawdzić, że

$$\det A = \lambda_1\lambda_2\lambda_3 \left(1 + \frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3}\right).$$

**Zadanie 38.5.** Uzasadnić, korzystając z własności wyznacznika i nie dokonując obliczeń, że wyznacznik macierzy

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \\ b_1 & b_2 & b_3 & b_4 & b_5 \\ c_1 & c_2 & 0 & 0 & 0 \\ d_1 & d_2 & 0 & 0 & 0 \\ e_1 & e_2 & 0 & 0 & 0 \end{bmatrix} \in M_{5\times 5}(\mathbb{R})$$

jest równy zero.

**Zadanie 38.6.** Niech $A \in M_{n\times n}(\mathbb{R})$ będzie taką macierzą, że $a_{ij} \in \{-1, 1\}$. Pokazać, że $\det A$ jest podzielny przez $2^{n-1}$.

**Zadanie 38.7.** Niech $A \in M_{n\times n}(\mathbb{R})$ będzie taką macierzą, że układ $Ax = b$ dla dowolnego wektora $b = [b_1, \ldots, b_n]^T$ o współrzędnych całkowitych ma rozwiązanie $x = [x_1, \ldots, x_n]^T$ o współczynnikach całkowitych. Pokazać, że $\det A = \pm 1$.

**Zadanie 38.8.** Niech $A \in M_{n\times n}(\mathbb{R})$. Pokazać, że istnieje taka niezerowa macierz $B \in M_{n\times n}(\mathbb{R})$, że $AB = 0$ wtedy i tylko wtedy, gdy $\det A = 0$.

**Zadanie 38.9.** Uzasadnić, że jeśli $A \in M_{n\times n}(\mathbb{R})$ oraz $A^2 + I = 0$, to $n$ jest parzyste. Czy jest tak również, gdy $A$ jest zespolona?

**Zadanie 38.10.** Niech $A \in M_{n\times n}(\mathbb{R})$ będzie taką macierzą, że $A^3 = 2I$. Pokazać, że $A^2 - 2A + 2I$ jest nieosobliwa.

**Zadanie 38.11.** Niech $A \in M_{n\times n}(\mathbb{Z})$ będzie nieosobliwa. Pokazać, że $A^{-1} \in M_{n\times n}(\mathbb{Z})$ wtedy i tylko wtedy, gdy $\det A = \pm 1$.

**Zadanie 38.12.** Pokazać, że równanie prostej przechodzącej przez dwa różne punkty $(a_1, b_1)$ i $(a_2, b_2)$ można zapisać w postaci

$$\det \begin{bmatrix} x & y & 1 \\ a_1 & b_1 & 1 \\ a_2 & b_2 & 1 \end{bmatrix} = 0.$$

**Zadanie 38.13.** Niech $A \in M_{n\times n}(\mathbb{R})$ będzie skośnie symetryczna tzn. $A^T = -A$. Pokazać, że jeśli $n$ jest nieparzyste, to $\det A = 0$. Pokazać, że gdy $n$ jest parzyste, to dodanie do każdego wyrazu macierzy $A$ tej samej liczby nie zmienia wyznacznika. Uzasadnić, że dla $n$ parzystego oraz $A \neq 0$ mamy $\det A \neq 0$.

**Zadanie 38.14.** Uzasadnić, że dla macierzy $A \in M_{n\times n}(\mathbb{R})$ górnie (dolnie) trójkątnej zachodzi wzór

$$\det A = a_{11} \dots a_{nn}.$$

**Zadanie 38.15.** Obliczyć wyznacznik macierzy

$$A = \begin{bmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}.$$

**Zadanie 38.16.** Oblicz wyznacznik macierzy $\begin{bmatrix} 1 & a & b \\ 1 & a+x & b \\ 1 & a & b+y \end{bmatrix}$.

**Zadanie 38.17.** Sprawdzić, że $\det \begin{bmatrix} 1+x_1 & x_2 & x_3 \\ x_1 & 1+x_2 & x_3 \\ x_1 & x_2 & 1+x_3 \end{bmatrix} = 1 + x_1 + x_2 + x_3$.

**Zadanie 38.18.** Niech $A = [a_{ij}]$, $B = [b_{ij}] \in M_{2\times 2}(\mathbb{R})$ oraz

$$C = \begin{bmatrix} a_{11} & a_{12} \\ b_{21} & b_{22} \end{bmatrix}, \quad D = \begin{bmatrix} b_{11} & b_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad E = \begin{bmatrix} 0 & \alpha \\ \beta & 0 \end{bmatrix}.$$

Pokazać, że

(a) $\det(A+B) = \det(A) + \det(B) + \det(C) + \det(D)$.

(b) jeśli $B = EA$, to $\det(A+B) = \det(A) + \det(B)$.

**Zadanie 38.19.** Dla ustalonego $n \in \mathbb{N}$ definiujemy macierz $A = [a_{ij}] \in M_{n\times n}(\mathbb{R})$, biorąc $a_{ij} = \min\{i, j\}$. Obliczyć wyznacznik macierzy $A$.

**Zadanie 38.20.** Sprawdzić, że dla macierzy Frobeniusa

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a_0 & a_1 & a_2 \end{bmatrix}$$

zachodzi równość

$$\det(\lambda I - A) = \lambda^3 - a_2\lambda^2 - a_1\lambda - a_0.$$

**Zadanie 38.21.** Uzasadnić, że

$$\det\begin{bmatrix} b_0 & b_2 & b_1 \\ b_1 & b_0 & b_2 \\ b_2 & b_1 & b_0 \end{bmatrix} = f(1)f(\epsilon_1)f(\epsilon_2),$$

gdzie 1, $\epsilon_1$, $\epsilon_2$ są zespolonymi pierwiastkami stopnia trzeciego z jedynki oraz

$$f(x) = b_0 + b_1 x + b_2 x^2.$$

**Zadanie 38.22.** Rozważmy macierz

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Pokazać, że jeśli

$$|a_{11}| > |a_{12}| + |a_{13}|,$$
$$|a_{22}| > |a_{21}| + |a_{23}|,$$
$$|a_{33}| > |a_{31}| + |a_{32}|,$$

to macierz $A$ jest nieosobliwa.

**Zadanie 38.23.** Niech $A \in M_{m\times n}(\mathbb{R})$ i $B \in M_{n\times m}(\mathbb{R})$. Pokazać, że

$$\det\left[\begin{array}{c|c} 0 & A \\ \hline -B & I \end{array}\right] = \det(AB).$$

Wskazówka: Rozważyć iloczyn $\left[\begin{array}{c|c} 0 & A \\ \hline -B & I \end{array}\right]\left[\begin{array}{c|c} I & 0 \\ \hline B & I \end{array}\right]$

**Zadanie 38.24.** Pokazać, że dla macierzy $A, B \in M_{n\times n}(\mathbb{R})$ zachodzi równość

$$\det\left[\begin{array}{c|c} A & B \\ \hline B & A \end{array}\right] = \det(A+B)\det(A-B).$$

Wskazówka: Rozważyć iloczyn $\left[\begin{array}{c|c} I & B \\ \hline 0 & A-B \end{array}\right]\left[\begin{array}{c|c} A+B & 0 \\ \hline 0 & I \end{array}\right]$.

**Zadanie 38.25.** Pokazać, że dla dowolnej macierzy $A \in M_{n\times n}(\mathbb{R})$ istnieje taka macierz diagonalna $J = \operatorname{diag}(a_{11}, \ldots, a_{nn})$, że $a_{ii} \in \{-1, 1\}$ oraz $\det(A+J) \neq 0$.

**Zadanie 38.26.** Macierz $A = [a_{ij}] \in M_{n\times n}(\mathbb{R})$ nazywamy *binarną*, gdy $a_{ij} \in \{0, 1\}$. Ile jest macierzy binernych dla $n = 3$? Która z nich ma największy wyznacznik?

**Zadanie 38.27.** Niech $A, B \in M_{n\times n}(\mathbb{F})$ i $A$ jest nieosobliwa. Pokazać, że

$$\det(I + AB) = \det(I + BA).$$

Uzasadnić, że założenie o odwracalności $A$ nie jest istotne.

**Zadanie 38.28.** Uzasadnić, że jeśli $A \in M_{n\times m}(\mathbb{F})$ i $B \in M_{m\times n}(\mathbb{F})$, to

$$\det(I_n + AB) = \det(I_m + BA).$$

Wskazówka: Sprawdzić, że

$$\left[\begin{array}{c|c} I_n & 0 \\ \hline -B & I_m \end{array}\right]\left[\begin{array}{c|c} I_n & 0 \\ \hline 0 & I_m + BA \end{array}\right]\left[\begin{array}{c|c} I_n & A \\ \hline 0 & I_m \end{array}\right] = \left[\begin{array}{c|c} I_n & A \\ \hline -B & I_m \end{array}\right]$$
$$= \left[\begin{array}{c|c} I_n & A \\ \hline 0 & I_m \end{array}\right]\left[\begin{array}{c|c} I_n + AB & 0 \\ \hline 0 & I_m \end{array}\right]\left[\begin{array}{c|c} I_n & 0 \\ \hline -B & I_m \end{array}\right]$$

**Zadanie 38.29.** Niech $u, v \in \mathbb{R}^n$. Pokazać, że $\det(I + uv^T) = 1 + (u|v)$.

**Zadanie 38.30.** Niech

$$K = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad L = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad M = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Niech

$$D : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}$$

będzie taką funkcją, że

(i) $D(AB) = D(A)D(B)$ dla dowolnych $A, B \in M_{2\times 2}(\mathbb{R})$,

(ii) $D(I) \neq D(K)$.

Pokazać, że

- $D(0) = D(L) = D(M) = 0$, $D(I) = 1$, $D(K) = -1$.
- jeśli $B$ powstaje z $A$ poprzez zamianę miejscami wierszy lub kolumn, to $D(A) = -D(B)$.
- Jeśli jakiś wiersz lub kolumna $A$ jest zerowy, to $D(A) = 0$.
- $D(A) = 0$ wtedy i tylko wtedy, gdy $A$ jest osobliwa.
- Podać przykład takiej funkcji $D$, która nie jest wyznacznikiem.

# Rozdział 39

# Wektory i wartości własne

**Zadanie 39.1.** Znaleźć wartości i wektory własne macierzy rzeczywistych

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}, \quad \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}, \quad \begin{bmatrix} 1 & 3 \\ \sqrt{2} & 3\sqrt{2} \end{bmatrix}$$

**Zadanie 39.2.** Znaleźć wartości i wektory własne macierzy zespolonych

$$\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 2 & -3 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1+i \\ 1-i & 1 \end{bmatrix}$$

**Zadanie 39.3.** Znaleźć wartości i wektory własne poniższych macierzy nad $\mathbb{Z}_3$

$$\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

**Zadanie 39.4.** Co można powiedzieć o wartościach własnych macierzy rzeczywistej

$$A = \begin{bmatrix} a & b \\ c & a \end{bmatrix},$$

gdy $bc > 0$?

**Zadanie 39.5.** Obliczyć $\begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}^{10} \begin{bmatrix} 5 \\ 1 \end{bmatrix}$.

**Zadanie 39.6.** Niech $\lambda$, $\mu \in \mathbb{C}$ będą wartościami własnymi macierzy $A \in M_{2\times 2}(\mathbb{C})$. Uzasadnić, że każda niezerowa kolumna macierzy $A - \lambda I$ jest wektorem własnym dla wartości własnej $\mu$.

**Zadanie 39.7.** Niech $A \in M_{n\times n}(\mathbb{C})$ i $\lambda \in \mathbb{C}$ jest wartością własną. Pokazać, że dla wielomianu zespolonego $p(x) = a_n x^n + \ldots + a_1 x + a_0$ liczba $p(\lambda)$ jest wartością własną macierzy

$$p(A) = a_n A^n + \ldots + a_1 A + a_0 I.$$

**Zadanie 39.8.** Wyznaczyć wszystkie możliwe rzeczywiste postaci Jordana dla macierzy

$$A = \begin{bmatrix} 2a & b \\ b & 0 \end{bmatrix}$$

w zależności od parametrów $a, b \in \mathbb{R}$.

**Zadanie 39.9.** Niech $A \in M_{2\times 2}(\mathbb{R})$. Uzasadnić, że jeśli $\operatorname{tr} A \neq 0$ oraz $\det A = 0$, to $A$ jest diagonalizowalna.

**Zadanie 39.10.** Niech $A \in M_{n\times n}(\mathbb{C})$ będzie taką macierzą, że $A^m = I$ dla pewnego $m \geqslant 1$. Pokazać, że $\det(I - A) \neq 0$ wtedy i tylko wtedy, gdy

$$I + A + A^2 + \ldots + A^{m-1} = 0.$$

**Zadanie 39.11.** Niech $A \in M_{2\times 2}(\mathbb{R})$ ma dwukrotną wartość własną $\lambda \in \mathbb{R}$. Uzasadnić, że dla dowolnego wektora $v \in \mathbb{R}^2$, $v$ jest wektorem własnym $A$ lub $(A - \lambda I)v$ jest wektorem własnym $A$.

**Zadanie 39.12.** Znaleźć postać Jordana macierzy

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & -2 \\ 0 & 2 & -2 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & 1 \\ -1 & -1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 0 \\ -1 & 3 & 0 \\ -1 & 1 & 2 \end{bmatrix}.$$

**Zadanie 39.13.** Dla jakich wartości parametrów $a, b, c \in \mathbb{R}$ macierz

$$\begin{bmatrix} 2 & 0 & 0 \\ a & 2 & 0 \\ b & c & 2 \end{bmatrix}$$

jest diagonalizowalna?

**Zadanie 39.14.** Dla wektorów $u = [1, 2, 3]^T$ i $v = [1, 1/2, 1/3]^T$ definiujemy macierz $A = u^T v \in M_{3\times 3}(\mathbb{R})$. Wyznaczyć $A^n$ dla $n \geqslant 1$.

**Zadanie 39.15.** Niech $B \in M_{3\times 3}(\mathbb{C})$. Pokazać, że

$$\det(xI - B) = x^3 - \operatorname{tr}(B)x^2 + \operatorname{tr}(\operatorname{adj}(B))x - \det B.$$

**Zadanie 39.16.** Wyznaczyć możliwe rzeczywiste postaci Jordana macierzy

$$A = \begin{bmatrix} 0 & 0 & a \\ 0 & b & 0 \\ c & 0 & 0 \end{bmatrix}$$

w zależności od parametrów $a, b, c \in \mathbb{R}$.

**Zadanie 39.17.** Niech $v$ będzie wektorem własnym macierzy $A$ dla wartości własnej $\lambda$. Pokazać, że dla skalara $c$ wektor $v$ jest wektorem własnym macierzy $A - cI$ odpowiadającym wartości własnej $\lambda - c$.

**Zadanie 39.18.** Rozważmy macierz

$$A = \begin{bmatrix} a_0 & a_1 & a_2 \\ a_2 & a_0 & a_1 \\ a_1 & a_2 & a_0 \end{bmatrix} \in M_{3\times 3}(\mathbb{C}).$$

Pokazać, że $A$ ma wektor własny postaci $[1, \eta, \eta^2]^T \in \mathbb{C}$.

**Zadanie 39.19.** Podać przykład takich macierzy $A$, $B \in M_{n\times n}(\mathbb{R})$, że $\lambda$ jest wartością własną $A$, $\mu$ jest wartością własną $B$, ale $\lambda + \mu$ nie jest wartością własną $A + B$.

**Zadanie 39.20.** Rozważmy macierze $A$ oraz $B = A - \alpha I$ dla pewnego $\alpha \in \mathbb{F}$. Jaki jest związek pomiędzy wartościami własnymi macierzy $A$ i $B$?

**Zadanie 39.21.** Uzasadnić, że macierze $A$ i $A^T$ mają te same wartości własne. Czy wektory własne też są takie same?

**Zadanie 39.22.** Niech $A, B \in M_{n\times n}(\mathbb{F})$. Pokazać, że

(1) jeśli $AB$ jest nieosobliwa i $\lambda$ jest wartością własną $AB$, to $\lambda$ jest wartością własną $BA$.

(2) jeśli 0 jest wartością własną $AB$, to 0 jest wartością własną $BA$.

(3) czy macierze $AB$ i $BA$ mają te same wartości własne?

**Zadanie 39.23.** Uzasadnić, że nie istnieją takie macierze $A, B \in M_{n\times n}(\mathbb{F})$, że $AB - BA = I$.

**Zadanie 39.24.** Podać przykład takich macierzy $A$, $B \in M_{n\times n}(\mathbb{R})$, że $\lambda$ jest wartością własną $A$, $\mu$ jest wartością własną $B$, ale $\lambda \cdot \mu$ nie jest wartością własną $AB$.

**Zadanie 39.25.** Uzasadnić, że macierze

$$A = \begin{bmatrix} 2 & 1 \\ -4 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & -1 \\ -5 & 7 \end{bmatrix}$$

nie są podobne.

**Zadanie 39.26.** Co można powiedzieć o wartościach własnych takiej macierzy kwadratowej $A$, że $A^3 = A$.

**Zadanie 39.27.** Niech $A \in M_{n\times n}(\mathbb{R})$ oraz $A^3 v = Av$ dla pewnego niezerowego wektora $v \in \mathbb{R}^n$. Pokazać, że co najmniej jedna z liczb 0, 1, $-1$ jest wartością własną macierzy $A$. Podać przykład takiej macierzy $A$ spełniającej powyższy warunek, że 0 jest wartością własną, a $-1$ i 1 nie są.

**Zadanie 39.28.** Niech $A \in M_{n\times n}(\mathbb{R})$ ma $n$ różnych wartości własnych $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Pokazać, że jeśli $v_1, \ldots, v_n$ są odpowiadającymi wektorami własnymi, to dla $v = v_1 + \ldots + v_n$ wektory $v, Av, \ldots, A^{n-1}v$ tworzą bazę $\mathbb{R}^n$.

**Zadanie 39.29.** Dla jakich wartości parametru $k \in \mathbb{R}$ poniższe macierze są diagonalizowalne jako macierze rzeczywiste?

$$\begin{bmatrix} 1 & 1 \\ 0 & k \end{bmatrix}, \quad \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} k & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

**Zadanie 39.30.** Niech $A \in M_{n\times n}(\mathbb{C})$. Uzasadnić, że $A$ jest diagonalizowalna wtedy i tylko wtedy, gdy krotność algebraiczna każdej wartości własnej jest równa jej krotności geometrycznej.

**Zadanie 39.31.** Rozważmy *macierz Toeplitza*

$$A = \begin{bmatrix} b & a \\ c & b \end{bmatrix} \in M_{2\times 2}(\mathbb{C}), \quad a \neq 0 \neq c.$$

Znaleźć jej wartości i wektory własne. Czy jest ona diagonalizowalna? Co można, w tym kontekście, powiedzieć o macierzy Toeplitza

$$A = \begin{bmatrix} b & a & 0 \\ c & b & a \\ 0 & c & b \end{bmatrix}, \quad a \neq 0 \neq c?$$

**Zadanie 39.32.** Niech $A, B \in M_{n\times n}(\mathbb{R})$ mają $n$ różnych wartości własnych. Uzasadnić, że $A$ i $B$ mają takie same wektory własne wtedy i tylko wtedy, gdy $AB = BA$.

**Zadanie 39.33.** Rozważmy rekurencję liniową

$$x_n = a x_{n-1} + b x_{n-2}, \quad a, b \in \mathbb{R}.$$

Sprawdzić, że

- $\begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix} = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_{n-1} \\ x_{n-2} \end{bmatrix}$,
- Jeśli $A = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix}$ ma dwie różne rzeczywiste wartości własne $\lambda_1 \neq \lambda_2$, to $x_n = c_1 \lambda_1^n + c_2 \lambda_2^n$ dla pewnych $c_1, c_2 \in \mathbb{R}$,
- Jeśli $A$ ma dwukrotną wartość własną $\lambda$, to $x_n = c_1 \lambda^n + c_2 n \lambda^n$ dla pewnych $c_1, c_2 \in \mathbb{R}$.

**Zadanie 39.34.** Niech $A \in M_{n\times n}(\mathbb{R})$ będzie macierzą antysymetryczną. Pokazać, że

- $A^2$ jest symetryczna,
- $x^T A^2 x \leqslant 0$ dla dowolnego $x \in \mathbb{R}^n$,
- niezerowe wartości własne macierzy antysymetrycznej $A$ są urojone.

Wskazówka: Zauważcie, że $x^T A^2 x = -xA^T Ax = (Ax)^T Ax$.

**Zadanie 39.35.** Niech $\lambda_k = x_k + iy_k$ dla $k = 1, \ldots, n$ będą wartościami własnymi macierzy rzeczywistej $A \in M_{n\times n}(\mathbb{R})$. Pokazać, że

(1) $y_1 + \ldots + y_n = 0$,

(2) $x_1y_1 + \ldots + x_ny_n = 0$,

(3) $\operatorname{tr} A^2 = (x_1^2 + \ldots + x_n^2) - (y_1^2 + \ldots + y_n^2)$.

**Zadanie 39.36.** Niech $\lambda_1$, $\lambda_2$ bądą różnymi wartościami własnymi macierzy $A \in M_{n\times n}(\mathbb{C})$ i niech $v_1$, $v_2$ będą ich wektorami własnymi. Pokazać, że $v_1 + v_2$ nie jest wektorem własnym $A$.

**Zadanie 39.37.** Niech $A \in M_{2\times 2}(\mathbb{R})$. Pokazać, że istnieje taka macierz nieosobliwa $S \in M_{2\times 2}(\mathbb{C})$, że wyrazami diagonalnymi macierzy $S^{-1}AS$ są 0 i $\operatorname{tr} A$.

**Zadanie 39.38.** Niech $A \in M_{n\times n}(\mathbb{C})$ będzie taka, że $\operatorname{rank} A = 1$. Pokazać, że $A^2 = \operatorname{tr} A \cdot A$.

**Zadanie 39.39.** Niech $A \in M_{5\times 5}(\mathbb{C})$ będzie taką macierzą, że $A^4 = A^2 \neq A$. Jakie są możliwe zespolone postacie Jordana macierzy $A$?

**Zadanie 39.40.** Narysować dyski Gerszgorina dla macierzy

$$\begin{bmatrix} 1 & 1 & 0 \\ 0.5 & 4 & 0.5 \\ 1 & 0 & 5 \end{bmatrix}, \quad \begin{bmatrix} 2 & -i & 0 \\ 1 & 2i & 1+i \\ 0 & 1 & -2i \end{bmatrix}$$

**Zadanie 39.41.** Podać przykład takich macierzy $A, B \in M_{2\times 2}(\mathbb{R})$, że $\lambda$ jest wartością własną $A$, $\mu$ jest wartością własną $B$, ale $\lambda + \mu$ nie jest wartością własną $A + B$. Podać analogiczny przykład dla $AB$.

**Zadanie 39.42.** Niech $A \in M_{n\times n}(\mathbb{R})$. Pokazać, że zbiory

$$V = \{x \in \mathbb{R}^n : \lim_{n\to+\infty} \|A^n x\| = 0\},$$

$$W = \{x \in \mathbb{R}^n : \sup_{n\geqslant 1} \|A^n x\| < \infty\},$$

są podprzestrzeniami wektorowymi $\mathbb{R}^n$. Podać przykład takiej macierzy $A$, że

$$E \neq \{0\}, \quad F \neq E, \quad F \neq \mathbb{R}^n.$$

# Rozdział 40

# Ortogonalność

**Zadanie 40.1.** Dla jakich wartości $a, b, c \in \mathbb{R}$ macierz $\begin{bmatrix} 1/2 & a \\ b & c \end{bmatrix}$ jest ortogonalna?

**Zadanie 40.2.** Zastosować procedurę Grama-Schmidta do bazy w $\mathbb{R}^3$:

$$v_1 = [1, -1, -1]^T, \quad v_2 = [0, 3, 3]^T, \quad v_3 = [3, 2, 4]^T.$$

**Zadanie 40.3.** Uzupełnić miejsca oznaczone $*$ tak, aby macierz

$$Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} & * \\ 0 & 1/\sqrt{3} & * \\ -1/\sqrt{2} & 1/\sqrt{3} & * \end{bmatrix}$$

była ortogonalna.

**Zadanie 40.4.** Znaleźć bazę ortogonalną $\mathbb{R}^3$, która zawiera wektor $[3, 1, 5]^T$.

**Zadanie 40.5.** Dla macierzy

$$A = \begin{bmatrix} 2 & 8 & 2 \\ 1 & 7 & -1 \\ -2 & -2 & 1 \end{bmatrix}$$

macierz ortogonalna

$$Q = \begin{bmatrix} 2/3 & 1/3 & 2/3 \\ 1/3 & 2/3 & -2/3 \\ -2/3 & 2/3 & 1/3 \end{bmatrix}$$

jest otrzymana w wyniku procedury Grama-Schmidta zastosowanej do kolumn macierzy $A$. Znaleźć rozkład $QR$ dla $A$.

**Zadanie 40.6.** Dla poniższych macierzy symetrycznych $A$:

$$\begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}, \quad \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix}, \quad \begin{bmatrix} 2 & 3 & 0 \\ 3 & 2 & 4 \\ 0 & 4 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$$

znaleźć takie macierze ortogonalne $Q$ i diagonalne $D$, że $Q^T AQ = D$.

**Zadanie 40.7.** Niech $A \in M_{n\times n}(\mathbb{R})$ będzie taką macierzą symetryczną, że $A^m = I$ dla pewnego $m \geqslant 1$. Pokazać, że $A^2 = I$.

**Zadanie 40.8.** Niech $A \in M_{3\times 3}(\mathbb{R})$ będzie macierzą symetryczną o wyznaczniku równym 6. Załóżmy, że $v_1 = [1, 2, 3]^T$, $v_2 = [0, 3, -2]^T$ są wektorami własnymi odpowiednio dla wartości własnych $\lambda_1 = 1$ i $\lambda_2 = 2$. Znaleźć taki wektor własny $v_3$ postaci $[1, x, y]$ $(x, y \in \mathbb{R})$, że $v_1, v_2, v_3$ tworzą bazę $\mathbb{R}^3$. Jaka jest wartość własna odpowiadająca $v_3$.

**Zadanie 40.9.** Uzasadnić, że jeśli macierz $A \in M_{n\times n}(\mathbb{R})$ jest symetryczna oraz $\operatorname{tr} A^2 = 0$, to $A = 0$.

**Zadanie 40.10.** Znaleźć dopełnienie ortogonalne $W^\perp$ podprzestrzeni $W$ w przypadku, gdy

- $W$ jest prostą w $\mathbb{R}^3$ o równaniu parametrycznym

$$x = t, \quad y = 2t, \quad z = -t.$$

- $W = \operatorname{span}\{[1,1,1,1]^T, [1,-1,1,1]^T\} \subset \mathbb{R}^4$.

**Zadanie 40.11.** Znaleźć bazę ortogonalną przestrzeni $\mathbb{R}^4$ zawierającą wektory

$$[1,0,2,2]^T, \quad [0,1,1,-1]^T.$$

**Zadanie 40.12.** Zastosować algorytm Grama-Schmidta do znalezienia bazy ortonormalnej przestrzeni $\operatorname{span}\{x_1, x_2, x_3\}$, gdzie

$$x_1 = [1,1,1,1]^T, \quad x_2 = [1,1,1,0]^T, \quad x_3 = [0,1,1,1]^T.$$

Następnie znaleźć rozkład $QR$ macierzy

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

**Zadanie 40.13.** Znaleźć bazę ortonormalną podprzestrzeni $\mathbb{R}^4$ opisanej równaniem $x_1 + x_2 + x_3 + x_4 = 0$.

**Zadanie 40.14.** Znaleźć macierz symetryczną $A \in M_{3\times 3}$ o wartościach własnych $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = -2$ i takich, że

$$\ker(A - \lambda_1 I) = \operatorname{span}\{[1,1,0]^T, [1,1,1]^T\},$$
$$\ker(A - \lambda_2 I) = \operatorname{span}\{[1,-1,0]^T\}.$$

**Zadanie 40.15.** Uzasadnić, że jeśli $u_1, \ldots, u_n \in \mathbb{R}^n$ tworzą bazę ortonormalną, to

$$(x|y) = \sum_{i=1}^{n} (x|u_i)(u_i|y),$$

dla dowolnych $x, y \in \mathbb{R}^n$.

**Zadanie 40.16.** Pokazać, że jeśli wektory $v_1, \ldots, v_k \in \mathbb{R}^n$ są ortonormalne, to dla dowolnego $v \in \mathbb{R}^n$ zachodzi nierówność:

$$\|v\|^2 \geqslant |(v|v_1)|^2 + \ldots + |(v|v_k)|^2.$$

**Zadanie 40.17.** Niech $V \subset \mathbb{R}^n$ będzie podprzestrzenią wektorową wymiaru $n-1$ oraz $a, b \in \mathbb{R}^n \setminus V$. Pokazać, że istnieje dokładnie jeden taki wektor $c \in V$, że

$$\|c - a\| + \|c - b\| \leqslant \|u - a\| + \|u - b\|,$$

dla każdego $u \in V$.

**Zadanie 40.18.** W przestrzeni $M_{n\times n}(\mathbb{R})$ definiujemy iloczyn skalarny wzorem

$$(A|B) = \operatorname{tr} A^T B.$$

Uzasadnić, że $(A|A) \geqslant 0$ oraz $(A|A) = 0$ tylko wtedy, gdy $A = 0$. Dla macierzy

$$A = \begin{bmatrix} 1 & 4 \\ -3 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} a^2 & a-1 \\ a+1 & -1 \end{bmatrix}$$

znaleźć wartości parametru $a \in \mathbb{R}$ dla których $(A|B) = 0$.

**Zadanie 40.19.** Niech przestrzeń $M_{n\times n}(\mathbb{R})$ będzie wyposażona w iloczyn skalarny i normę

$$(A|B) = \operatorname{tr} A^T B, \quad \|A\| = \sqrt{(A|A)}.$$

Rozważmy podprzestrzeń $V$ macierzy symetrycznych i podprzestrzeń $W$ macierzy antysymetrycznych. Pokazać, że

- $V^\perp = W$,
- rzut ortogonalny $P$ na $V$ jest dany przez

$$P(A) = \frac{A + A^T}{2},$$

- $\|A - P(A)\| = \|(A - A^T)/2\| = \sqrt{\frac{\operatorname{tr} A^T A - \operatorname{tr} A^2}{2}}$.

**Zadanie 40.20.** Niech $V, W \subset \mathbb{R}^n$ będą podprzestrzeniami. Pokazać, że

$$(V + W)^\perp = V^\perp \cap W^\perp.$$

**Zadanie 40.21.** Niech $A \in M_{n\times n}(\mathbb{R})$. Pokazać, że $A^T A = 0$ wtedy i tylko wtedy, gdy $A = 0$. Uzasadnić, że jeśli $A^T A = AA^T$, to $\ker A = \ker A^T$ oraz $\operatorname{im} A = \operatorname{im} A^T$. W szczególności,

$$\mathbb{R}^n = \ker A \oplus \operatorname{im} A.$$

**Zadanie 40.22.** Podać przykład takiej macierzy $A \in M_{n\times n}(\mathbb{R})$, że

$$\ker A \cap \operatorname{im} A = \{0\},$$

ale podprzestrzenie $\ker A$ i $\operatorname{im} A$ nie są ortogonalne.

**Zadanie 40.23.** Niech $V \subset \mathbb{R}^n$. Pokazać, że $V$ jest podprzestrzenią wektorową wtedy i tylko wtedy, gdy $V = \ker A$ dla pewnej macierzy $A \in M_{n\times n}(\mathbb{R})$.

**Zadanie 40.24.** Znaleźć macierz rzutu ortogonalnego na podprzestrzeń $\operatorname{span}\{[1, 2, 1, 1]^T, [-1, 1, 0, 1]^T\}$ przestrzeni $\mathbb{R}^4$.

**Zadanie 40.25.** Niech $A \in M_{2\times 2}(\mathbb{R})$. Pokazać, że istnieje taka macierz ortogonalna $S$, że wyrazy diagonalne macierzy $SAS^{-1}$ są równe.

**Zadanie 40.26.** Dla jakich wartości $b \in \mathbb{R}$ macierz

$$\begin{bmatrix} 2 & 2 & 4 \\ 2 & b & 8 \\ 4 & 8 & 7 \end{bmatrix}$$

jest dodatnio określona?

**Zadanie 40.27.** Niech $A \in M_{n\times n}(\mathbb{R})$ będzie idempotentna ($A^2 = A$). Pokazać, że $\operatorname{rank} A = \operatorname{tr} A$.

**Zadanie 40.28.** Niech $A, B \in M_{n\times n}(\mathbb{R})$ będą idempotentne oraz $\ker A = \ker B$ i $\operatorname{im} A = \operatorname{im} B$. Pokazać, że $A = B$.

**Zadanie 40.29.** Niech $A \in M_{n\times n}(\mathbb{R})$ będzie idempotentna. Pokazać, że istnieje dokładnie jedna taka macierz idempotentna $B \in M_{n\times n}(\mathbb{R})$, że $\ker A = \operatorname{im} B$ i $\operatorname{im} A = \ker B$.

**Zadanie 40.30.** Niech $A, B \in M_{n\times n}(\mathbb{R})$ będą idempotentne. Pokazać, że $A + B$ jest idempotentna wtedy i tylko wtedy, gdy zachodzą warunki

$$\operatorname{im} A \subset \ker B, \quad \operatorname{im} B \subset \ker A.$$

**Zadanie 40.31.** Załóżmy, że $\mathbb{R}^n = V \oplus W$ dla pewnych podprzestrzeni $V, W \subset \mathbb{R}^n$. Pokazać, że istnieje dokładnie jedna taka macierz idempotentna $A \in M_{n\times n}(\mathbb{R})$, że $\operatorname{im} A = V$ oraz $\ker A = W$.

**Zadanie 40.32.** Niech $A, B \in M_{n\times n}(\mathbb{R})$ będą takimi macierzami idempotentnymi, że $A + B$ jest również idempotentna. Uzasadnić, że

- $\operatorname{im}(A + B) = \operatorname{im} A \oplus \operatorname{im} B$,
- $\ker(A + B) = \ker A \cap \ker B$,
- $\mathbb{R}^n = \operatorname{im} A \oplus \operatorname{im} B \oplus (\ker A \cap \ker B)$.

**Zadanie 40.33.** Uzasadnić, że liniowa izometria przestrzeni $\mathbb{R}^3$ jest złożeniem obrotu wokół prostej i symetrii względem płaszczyzny.

**Zadanie 40.34.** Sprawdzić, że

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

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

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

**Zadanie 40.35.** Macierz nieujemną $A = [a_{ij}] \in M_{n\times n}(\mathbb{R})$ nazywamy *podwójnie stochastyczną*, jeśli

$$\sum_{i=1}^{n} a_{ik} = 1 = \sum_{j=1}^{n} a_{kj}, \quad k = 1, \ldots, n.$$

Pokazać, że iloczyn macierzy podwójnie stochastycznych jest macierzą podwójnie stochastyczną.

**Zadanie 40.36.** Pokazać, że jeśli $A = [a_{ij}] \in M_{n\times n}(\mathbb{R})$ jest ortogonalna, to macierz $B = [a_{ij}^2]$ jest podwójnie stochastyczna.

**Zadanie 40.37.** Niech $A \in M_{n\times n}(\mathbb{R})$ i niech $A = QR$ będzie jej rozkładem $QR$ otrzymany z procedury Grama-Schmidta. Sprawdzić, że

$$\|Ax - b\| = \|Rx - c\|,$$

gdzie $x, b \in \mathbb{R}^n$ oraz $c = Q^T b$.

**Zadanie 40.38.** Pokazać, że

(1) Jeśli $A \in O(n)$ jest taka, że $AB = BA$ dla każdej macierzy $B \in O(n)$, to $A = \pm I$.

(2) Jeśli $A \in SO(n)$ jest taka, że $AB = BA$ dla każdej macierzy $B \in SO(n)$, to $A = \pm I$, gdy $n$ jest parzyste oraz $= I$, gdy $n$ jest nieparzyste.

**Zadanie 40.39.** Niech $P \in O(n)$. Pokazać, że $P^T AP$ jest symetryczna wtedy i tylko wtedy, gdy $A \in M_{n\times n}(\mathbb{R})$ jest symetryczna.

# Rozdział 41

# Macierze hermitowskie

**Zadanie 41.1.** Które z poniższych macierzy są hermitowskie

$$\begin{bmatrix} 1 & 3-i \\ 3+i & i \end{bmatrix}, \quad \begin{bmatrix} 0 & 3-2i \\ 3-2i & 4 \end{bmatrix},$$

$$\begin{bmatrix} 2 & 2-i & -4i \\ 2+i & 0 & 1 \\ 4i & 1 & 5 \end{bmatrix}, \quad \begin{bmatrix} -2 & 2 & 4 \\ 2 & 0 & 8 \\ 4 & 8 & 7 \end{bmatrix}?$$

**Zadanie 41.2.** Znaleźć wartości i wektory własne macierzy hermitowskiej

$$A = \begin{bmatrix} 3 & 2-i & 3i \\ 2+i & 0 & -1+i \\ -3i & -1-i & 0 \end{bmatrix}$$

**Zadanie 41.3.** Które z poniższych macierzy są unitarne

$$\begin{bmatrix} 1+i & 1+i \\ 1-i & 1-i \end{bmatrix}, \quad \begin{bmatrix} 1+i & 1-i \\ 1-i & 1+i \end{bmatrix},$$

$$\begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}, \quad \frac{1}{1\sqrt{2}} \begin{bmatrix} \sqrt{3}-i & 1+\sqrt{3}i \\ \sqrt{3}+i & 1-\sqrt{3}i \end{bmatrix}?$$

**Zadanie 41.4.** Niech $z \in \mathbb{S}^1$. Pokazać, że macierz

$$\frac{1}{\sqrt{2}} \begin{bmatrix} z & \overline{z} \\ iz & -i\overline{z} \end{bmatrix}$$

jest unitarna.

**Zadanie 41.5.** Pokazać, że dla $A \in M_{n\times n}(\mathbb{C})$ mamy

$$\det A^* = \overline{\det A}.$$

**Zadanie 41.6.** Podaj przykład macierzy normalnej $A \in M_{2\times 2}(\mathbb{C})$, która nie jest ani hermitowska, ani unitarna.

**Zadanie 41.7.** Niech $p = [p_0, p_1, p_2, p_3]^T$, $q = [q_0, q_1, q_2, q_3]^T \in \mathbb{R}^4$ będą kwaternionami. Pokazać, że

- iloczyn kwaternionów $p \cdot q$ można zapisać macierzowo

$$\begin{aligned} p \cdot q &= Pq \\ &= \underbrace{\begin{bmatrix} p_0 & -p_1 & -p_2 & -p_3 \\ p_1 & p_0 & -p_3 & p_2 \\ p_2 & p_3 & p_0 & -p_1 \\ p_3 & -p_2 & p_1 & p_0 \end{bmatrix}}_{P} \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix}. \end{aligned}$$

- jeśli $|p| = 1$, to $P \in O(4)$,
- $|p \cdot q|^2 = |p|^2 |q|^2$

**Zadanie 41.8.** Rozważmy *macierze Pauliego*

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Pokazać, że

- $I = \sigma_1^2 = \sigma_2^2 = \sigma_3^2$,
- $\sigma_1 \sigma_2 = i\sigma_3$,
- $\sigma_2 \sigma_3 = i\sigma_1$,
- $\sigma_3 \sigma_1 = i\sigma_2$.

Sprawdzić, że jeśli

$$q = \begin{bmatrix} q_0 - iq_3 & -iq_1 - q_2 \\ -iq_1 + q_2 & q_0 + iq_3 \end{bmatrix}$$

jest kwaternionem traktowanym jako element SU(2), to

$$q = q_0 I - iq_1\sigma_1 - iq_2\sigma_2 - iq_3\sigma_3.$$

# Rozdział 42

# Różne

**Zadanie 42.1.** Które z poniższych macierzy są dodatnio określone

$$\begin{bmatrix} 1 & -1 & -1 \\ -1 & 5 & 1 \\ -1 & 1 & 5 \end{bmatrix}, \quad \begin{bmatrix} 20 & 6 & 8 \\ 6 & 3 & 0 \\ 8 & 0 & 8 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 & 2 \\ 0 & 6 & 2 \\ 2 & 2 & 4 \end{bmatrix}.$$

**Zadanie 42.2.** Obliczyć $e^A$ dla

$$A = \begin{bmatrix} 6 & 2 & 8 \\ -2 & 2 & -2 \\ 0 & 0 & 2 \end{bmatrix}.$$

**Zadanie 42.3.** Podać przykład takich macierzy $A, B \in M_{2\times 2}(\mathbb{R})$, że $e^{A+B} \neq e^A e^B$.

**Zadanie 42.4.** Niech $A \in M_{2\times 2}(\mathbb{C})$ będzie nieosobliwa. Pokazać, że istnieje taka macierz $B \in M_{2\times 2}(\mathbb{C})$, że $B^2 = A$. Podać warunek konieczny i wystarczający na istnienie takiej macierzy $B$ w przypadku, gdy $A$ jest osobliwa.

**Zadanie 42.5.** Niech $A \in M_{2\times 2}(\mathbb{C})$. Pokazać, że następujące warunki są równoważne

- $\lim_{n\to\infty} A^n = 0$,
- $\rho(A) < 1$, gdzie $\rho(A)$ jest promieniem spektralnym $A$,
- *szereg Neumanna* $I + A + A^2 + \ldots$ jest zbieżny.

Jeśli zachodzi którykolwiek z powyższych warunków, to $I - A$ jest nieosobliwa oraz

$$(I - A)^{-1} = \sum_{n=0}^{\infty} A^n.$$

**Zadanie 42.6.** Zweryfikować twierdzenie Frobeniusa–Perrona, obliczając wartości i wektory własne macierzy

$$A = \begin{bmatrix} 7 & 2 & 3 \\ 1 & 8 & 3 \\ 1 & 2 & 9 \end{bmatrix}.$$

**Zadanie 42.7.** Znaleźć promień spektralny $\rho(A)$ oraz dodatni wektor własny Perrona dla

$$A = \begin{bmatrix} 1-a & b \\ a & 1-b \end{bmatrix},$$

gdzie $a + b = 1$ oraz $a, b > 0$.

**Zadanie 42.8.** Uzasadnić, że jeśli $A \in M_{n\times n}(\mathbb{R})$ jest dodatnia oraz jej kolumny (wiersze) sumują się do liczby $\rho$, to $\rho(A) = \rho$.

**Zadanie 42.9.** Pokazać, że jeśli $A \in M_{n\times n}(\mathbb{R})$ jest dodatnia, to

$$\min_i \sum_{j=1}^n a_{ij} \leqslant \rho(A) \leqslant \max_i \sum_{j=1}^n a_{ij}.$$

**Zadanie 42.10.** Podać przykłady takich macierzy nieujemnych $A \in M_{n\times n}(\mathbb{R})$, które nie są dodatnie i $\rho(A)$ jest wartością własną $A$ oraz

- $\rho(A) = 0$,
- algebraiczna krotność $\rho(A)$ jest większa od 1,
- $\rho(A)$ nie ma dodatniego wektora własnego,
- istnieje taka $\lambda \neq \rho(A)$ wartość własna, że $|\lambda| = \rho(A)$.

**Zadanie 42.11.** Znaleźć formułę na odwzorowanie liniowe $L : \mathbb{R}^2 \to \mathbb{R}^2$, będące rzutem na podprzestrzeń $\{[x, y]^T \in \mathbb{R}^2 : x+2y = 0\}$ równolegle do podprzestrzeni $\{[x, y]^T \in \mathbb{R}^2 : 2x+3y = 0\}$.

**Zadanie 42.12.** Niech $A \in M_{n\times n}(\mathbb{R})$ będzie taka, że $A^2 - 5A + 6I = 0$. Uzasadnić, że

$$\mathbb{R}^n = \ker(A - 2I) \oplus \ker(A - 3I).$$

**Zadanie 42.13.** Niech $A \in M_{n\times n}(\mathbb{R})$ i $\lambda \in \mathbb{R}$ będą dowolne. Pokazać, że podprzestrzeń $\ker(A - \lambda I)$ jest niezmiennicza dla $A$.

**Zadanie 42.14.** Niech $V_1, V_2$ i $W_1, W_2$ będą przestrzeniami wektorowymi nad ciałem $\mathbb{F}$. Pokazać, że jeśli $L_1 : V_1 \to W_1$ i $L_2 : V_2 \to W_2$, to

$$L_1 \times L_2 : V_1 \times V_2 \ni (v_1, v_2) \mapsto (L_1(v_1), L_2(v_2)) \in W_1 \times W_2$$

jest liniowe wtedy i tylko wtedy, gdy $L_1$ i $L_2$ są liniowe. Wtedy

$$\ker(L_1 \times L_2) = \ker(L_1) \times \ker(L_2),$$

$$\mathrm{im}(L_1 \times L_2) = \mathrm{im}(L_1) \times \mathrm{im}(L_2).$$

**Zadanie 42.15.** Niech $V, V_1, V_2$ będą przestrzeniami wektorowymi nad ciałem $\mathbb{F}$. Pokazać, że jeśli $L_1 : V \to V_1$ i $L_2 : V \to V_2$, to

$$(L_1, L_2) : V \ni v \mapsto (L_1(v), L_2(v)) \in W_1 \times W_2$$

jest liniowe wtedy i tylko wtedy, gdy $L_1$ i $L_2$ są liniowe. Wtedy

$$\ker(L_1, L_2) = \ker(L_1) \cap \ker(L_2).$$

Sprawdzić, czy

$$\mathrm{im}(L_1, L_2) = \mathrm{im}(L_1) \times \mathrm{im}(L_2)?$$

**Zadanie 42.16.** Znaleźć bazę dla $(\mathbb{R}^2)^*$ dualną do bazy $[1, 2]^T$, $[2.5]^T \in \mathbb{R}^2$.

**Zadanie 42.17.** Niech

$$L : \mathbb{R}^2 \ni [x, y]^T \mapsto [x - 2y, x + y, 3x + y]^T \in \mathbb{R}^3.$$

Obliczyć $L^*(\beta)$, gdzie

$$\beta : \mathbb{R}^3 \ni [x, y, z]^T \mapsto 2x - 4y + 3z \in \mathbb{R}.$$

**Zadanie 42.18.** Znaleźć dowolne takie odwzorowanie liniowe $L : \mathbb{R}^3 \to \mathbb{R}^3$, że

$$\ker L = \mathrm{span}\{[1, 1, 1]^T\},$$

$$\mathrm{im}\, L = \{[x, y, z]^T \in \mathbb{R}^3 : 2x - y + z = 0\}.$$

**Zadanie 42.19.** Znaleźć dowolne bazy podprzestrzeni $V, W, V \cap W$ i $V + W$ dla

$$V = \{[x, y, z]^T \in \mathbb{R}^3 : x + y - z = 0\},$$

$$W = \{[x, y, z]^T \in \mathbb{R}^3 : x - 3y + z = 0\}.$$

**Zadanie 42.20.** Sprawdzić, czy

$$(\mathbb{R}^2)^* = V \oplus W,$$

jeśli

$$V = \{\alpha \in (\mathbb{R}^2)^* : \alpha([2, 3]^T) = 0\},$$
$$W = \{\alpha \in (\mathbb{R}^2)^* : \alpha([1, 2]^T) = 0\}.$$

**Zadanie 42.21.** Rozwiązać równanie macierzowe

$$\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} X = \begin{bmatrix} 3 & 2 & 4 \\ 1 & 2 & 3 \end{bmatrix}.$$

**Zadanie 42.22.** Niech $L : \mathbb{R}^2 \ni [x, y]^T \mapsto [2x+y, x+y]^T \in \mathbb{R}^2$. Znaleźć macierz odwzorowania dualnego $L^*$ w bazie dla $(\mathbb{R}^2)^*$ danej przez odwzorowania

$$\alpha : \mathbb{R}^2 \ni [x, y]^T \mapsto 2x + y \in \mathbb{R},$$

$$\beta : \mathbb{R}^2 \ni [x, y]^T \mapsto 3x + y \in \mathbb{R}$$

**Zadanie 42.23.** Niech $V \subset \mathbb{R}^3$ i $W \subset \mathbb{R}^4$ będą ustalonymi podprzestrzeniami wymiaru 2.

- Pokazać, że
$$X = \{A \in M_{4\times3}(\mathbb{R}) : A(V) \subset W\}$$
jest podprzestrzenią wektorową w $A \in M_{4\times3}(\mathbb{R})$.

- Jaki jest wymiar $X$?

**Zadanie 42.24.** Sprawdzić, że macierz $\begin{bmatrix} 1 & 4 & 3 \\ 2 & 7 & 5 \\ 2 & 5 & 3 \end{bmatrix}$ ma rząd 2. Znaleźć takie macierze nieosobliwe $P, Q \in M_{3\times3}(\mathbb{R})$, że

$$Q^{-1} \begin{bmatrix} 1 & 4 & 3 \\ 2 & 7 & 5 \\ 2 & 5 & 3 \end{bmatrix} P = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

**Zadanie 42.25.** Obliczyć ślad macierzy $A \in M_{2\times2}(\mathbb{R})$, wiedząc, że

$$A[1, 2]^T = [2, 3]^T, \quad A[1, 3]^T = [3, 8]^T.$$

**Zadanie 42.26.** Wyznaczyć takie odwzorowanie dwuliniowe antysymetryczne $f : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, że

$$f([1, 0, 2]^T, [1, 1, 1]^T) = 1,$$
$$f([1, 0, 2]^T, [2, 1, 4]^T) = 2,$$
$$f([1, 1, 1]^T, [2, 1, 4]^T) = 0.$$

**Zadanie 42.27.** Znaleźć postacie kanoniczne form kwadratowych

- $Q([x, y]^T) = x^2 - 2xy + 5y^2$,
- $Q([x, y]^T) = 4x^2 - 8xy + 3y^2$,
- $Q([x, y]^T) = x^2 - 6xy + 9y^2$,

- $Q([x,y,z]^T) = x^2 - 2xy - 2xz + 5y^2 - 6yz + 9z^2$,
- $Q([x,y,z]^T) = 2xy + 6xz + y^2 + 2yz - 7z^2$,
- $Q([x,y,z]^T) = x^2 + 2xy - 6xz - 4yz + 8z^2$.

**Zadanie 42.28.** Sprawdzić, że odwzorowanie

$$\mathbb{R}^2 \times \mathbb{R}^2 \ni ([a,b]^T, [a,d]^T) \mapsto 8ac - 2ad - 2bc + bd \in \mathbb{R}$$

jest iloczynem skalarnym.

**Zadanie 42.29.** Znaleźć rozkład polarny macierzy $\begin{bmatrix} 14 & -2 \\ 22 & 4 \end{bmatrix}$.

**Zadanie 42.30.** Niech $a, b, c \in \mathbb{R}^3$. Pokazać, że

- $(a \times b) \times c = (a|c) \cdot b - (b|c) \cdot a$,
- iloczyn wektorowy nie jest łączny,
- $a \times (b \times c) + b \times (c \times a) + c \times (a \times b) = 0$,
- $(a \times b|a \times b) + (a|b)^2 = \|a\|^2\|b\|^2$,
- $(a \times b) \times (a \times c) = (a|b \times c)a$.

**Zadanie 42.31.** Dla ustalonych wektorów $a, b \in \mathbb{R}^n$ definiujemy odwzorowanie liniowe

$$L_{a,b} : \mathbb{R}^n \ni x \mapsto (x|b) \cdot a \in \mathbb{R}^n.$$

- Znaleźć macierz $A$ odwzorowania $L$ w bazie standardowej.
- Pokazać, że $A^T$ jest macierzą odwzorowania $L_{b,a}$.
- Pokazać, że $\operatorname{tr} A = (a|b)$.

Podać warunek konieczny i wystarczający, aby $A$ była symetryczna.

**Zadanie 42.32.** Niech $V = M_{2\times 2}(\mathbb{R})$. Dla macierzy $B \in V$ definiujemy

$$\tau_B : V \ni A \mapsto \operatorname{tr} AB \in \mathbb{R}.$$

Sprawdzić, czy odwzorowanie

$$V \ni B \mapsto \tau_B \in V^*$$

jest izomorfizmem?

**Zadanie 42.33.** Niech $A \in M_{n\times n}(\mathbb{R})$ będzie taką macierzą, że $a_{ij} \in \{0,1\}$ oraz wartości własne $\lambda_1, \ldots, \lambda_n$ macierzy $A$ są dodatnimi liczbami rzeczywistymi.

- Korzystając z nierówności pomiędzy średnią geometryczną i arytmetyczną, uzasadnić, że
$$n \geqslant \operatorname{tr} A \geqslant n(\det A)^{1/n} \geqslant n.$$
- Wywnioskować, że $\lambda_1 = \ldots = \lambda_n = 1$ oraz $a_{11} = \ldots = a_{nn} = 1$. W szczególności, $\det A = 1$.

**Zadanie 42.34.** Niech $A \in M_{m\times m}(\mathbb{R})$ będzie dodatnia. Pokazać, że

$$\rho(A) = \lim_{n\to\infty} \sqrt[n]{\operatorname{tr} A^n}.$$

**Zadanie 42.35.** Niech $A \in M_{n\times n}(\mathbb{C})$. Pokazać, że jeśli $A^3 = A$, to $\operatorname{rank} A = \operatorname{tr} A^2$.

**Zadanie 42.36.** Dla macierzy $A \in M_{n\times n}(\mathbb{C})$ następujące warunki są równoważne

(1) $A$ jest diagonalizowalna,

(2) $\operatorname{rank}(A - \lambda I) = \operatorname{rank}(A - \lambda I)^2$ dla każdej wartości własnej $\lambda$.

(3) $\ker(A - \lambda I) = \ker(A - \lambda I)^2$ dla każdej wartości własnej $\lambda$.

**Zadanie 42.37.** Dla macierzy $A \in M_{n\times n}(\mathbb{C})$ następujące warunki są równoważne

(i) $A^2 = BA$ dla pewnej macierzy nieosobliwej $B$,

(ii) $\operatorname{rank} A = \operatorname{rank} A^2$,

(iii) $\ker A = \ker A^2$,

(iv) $\ker A \cap \operatorname{im} A = \{0\}$,

(v) Istnieją macierze nieosobliwe $P \in M_{n\times n}(\mathbb{C})$ i $C \in M_{\operatorname{rank} A\times \operatorname{rank} A}(\mathbb{C})$, że

$$A = P \operatorname{diag}(C, 0)P^{-1}.$$

**Zadanie 42.38.** Rozważmy odwzorowania liniowe $\mathbb{R}^3 \to \mathbb{R}^3$ dane wzorami

$$L_1([a, b, c]^T) = [a + b, b + c, c + a]^T, \quad L_2([a, b, c]^T) = [a - b, b - c, 0]^T,$$

$$L_3([a, b, c]^T) = [-b, a, c]^T, \quad L_4([a, b, c]^T) = [a, b, b]^T.$$

- Sprawdzić, czy $\mathbb{R}^3 = \ker L_i \oplus \operatorname{im} L_i$ dla $i = 1, 2, 3, 4$?
- Czy $\operatorname{im} L_2$ i $\ker L_2$ są niezmiennicze dla $L_3$?
- Wyznaczyć $L_3 \circ L_4$ i $L_4 \circ L_3$ oraz ich jądra i obrazy.

**Zadanie 42.39.** Niech $V$ będzie przestrzenią wymiaru 3 nad ciałem $\mathbb{F}$. Odwzorowanie liniowe $L : V \to V$ ma w pewnej bazie macierz

$$\begin{bmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{bmatrix}.$$

Znaleźć wymiary $\dim\ker L$ i $\dim\operatorname{im} L$ jeśli

(1) $\mathbb{F} = \mathbb{R}$,

(2) $\mathbb{F} = \mathbb{Z}_2$,

(3) $\mathbb{F} = \mathbb{Z}_3$.

Czy $V = \ker L \oplus \operatorname{im} L$ w którymś z powyższych przypadków?

**Zadanie 42.40.** Niech $A \in M_{n\times n}(\mathbb{C})$ będzie taką macierzą, że $A^2 = 0$. Pokazać, że

$$2 \operatorname{rank} A \leqslant n.$$

**Zadanie 42.41.** Niech $A \in M_{n\times n}(\mathbb{F})$ będzie taka, że $A^2 = 0$. Przypuśćmy, że $\mathbb{F}^n = \ker A \oplus W$. Pokazać, że

- $\dim W = \operatorname{rank} A$,
- Jeśli $w_1, \ldots, w_r$ jest bazą $W$, to $Aw_1, \ldots, Aw_r \in \ker A$ są liniowo niezależne.
- Istnieją takie $x_1, \ldots, x_{n-2r} \in \ker A$, że

  $$w_1, \ldots, w_r, Aw_1, \ldots, Aw_r, x_1, \ldots, x_{n-2r}$$

  tworzą bazę $\mathbb{F}^n$.

**Zadanie 42.42.** Rozważmy ciąg

$$\frac{1}{1}, \frac{3}{2}, \frac{7}{5}, \ldots, \frac{a_n}{b_n}, \ldots$$

gdzie

$$a_{n+1} = a_n + 2b_n, \quad b_{n+1} = a_n + b_n.$$

- Wskazać macierz $A$ dla której $[a_{n+1}, b_{n+1}]^T = A[a_n, b_n]^T$.
- Znaleźć jawne formuły na $a_n$ i $b_n$, diagonalizując macierz $A$.
- Wynioskować, że $\frac{a_n}{b_n} \to \sqrt{2}$.

**Zadanie 42.43.** Niech

$$A = \begin{bmatrix} 2 & -2 & 0 \\ 1 & -1 & 0 \\ -1 & 3 & 1 \end{bmatrix}.$$

Uzasadnić, że $A$ nie jest diagonalizowalna. Znaleźć bazę dla $\mathbb{R}^3$, w której $A$ jest górnie trójkątna.

**Zadanie 42.44.** Pokazać, że jeśli któraś z macierzy $A, B \in M_{n\times n}(\mathbb{C})$ jest nieosobliwa, to macierze $AB$ i $BA$ są podobne. Uzasadnić, że twierdzenie odwrotne nie jest prawdziwe.

**Zadanie 42.45.** Niech $A, B \in M_{n\times n}(\mathbb{R})$. Zakładamy, że $A$ jest symetryczna, $B$ jest antysymetryczna. Jeśli $AB = BA$ i $\det(A - B) \neq 0$, to

$$(A + B)(A - B)^{-1}$$

jest ortogonalna.

**Zadanie 42.46.** Załóżmy, że $A, B \in O(n)$ oraz $\det A = -\det B$. Pokazać, że $A + B$ jest osobliwa.

**Zadanie 42.47.** Niech $A, B \in M_{2\times 2}(\mathbb{C})$. Pokazać, że

$$\det(A + B) + \det(A - B) = 2\det A + 2\det B.$$

**Zadanie 42.48.** Niech $A \in M_{2\times 2}(\mathbb{R})$. Obliczyć $\det A$ jeśli $\det(A+A^T) = 8$ i $\det(A+2A^T) = 27$.

**Zadanie 42.49.** Niech $A \in M_{2\times 2}(\mathbb{R})$. Pokazać, że

- $\det(A^2 + I) = |\det(A + iI)|^2 \geqslant 0$,
- jeśli $a, b, c \in \mathbb{R}$ oraz $b^2 - 4ac \leqslant 0$, to $\det(aA^2 + bA + cI) \geqslant 0$.

**Zadanie 42.50.** Dla $A \in M_{2\times 2}(\mathbb{C})$ definiujemy

$$C(A) = \{X \in M_{2\times 2}(\mathbb{C}) : AX = XA\}.$$

Pokazać, że

- jeśli $A = kI$ dla pewnego $k \in \mathbb{C}$, to $C(A) = M_{2\times 2}(\mathbb{C})$,
- jeśli $A \neq kI$, to
  $$C(A) = \{\alpha A + \beta I : \alpha, \beta \in \mathbb{C}\}.$$
  Wywnioskować, że jeśli $p$ jest wielomianem zespolonym, to $p(A) = \alpha A + \beta I$ dla pewnych $\alpha, \beta \in \mathbb{C}$.
- Jeśli $AB \neq BA$, to
  $$C(A) \cap C(B) = \{\alpha I : \alpha \in \mathbb{C}\}.$$

**Zadanie 42.51.** Niech $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{R})$ będzie taka, że $A^2 = -I$. Sprawdzić, że

- $a + d = 0$, $b \neq 0$ i $c \neq 0$,
- $A = \begin{bmatrix} a & b \\ -\frac{1+a^2}{b} & -a \end{bmatrix}$.
- dla $P = \begin{bmatrix} ab & -b \\ -(1+a^2) & 0 \end{bmatrix}$ mamy

$$P^{-1}AP = J,$$

czyli $A$ jest podobna do $J$.

**Zadanie 42.52.** Niech $A = \begin{bmatrix} 2 & i \\ i & 0 \end{bmatrix}$. Pokazać, że $A^n = \begin{bmatrix} n+1 & ni \\ ni & 1-n \end{bmatrix}$ dla $n \geqslant 1$.

**Zadanie 42.53.** Pokazać, że $A \in M_{2\times 2}(\mathbb{C})$ komutuje z każdą macierzą symetryczną wtedy i tylko wtedy, gdy $A = \alpha I$ dla pewnego $\alpha \in \mathbb{C}$.

**Zadanie 42.54.** Niech $A \in M_{2\times 2}(\mathbb{R})$. Pokazać, że

- $AA^T = A^T A$ wtedy i tylko wtedy, gdy
$$A = c \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$
dla pewnych $c, \theta \in \mathbb{R}$.
- $A^2 = 0$ wtedy i tylko wtedy, gdy
$$A = c \begin{bmatrix} \cos\theta & 1+\sin\theta \\ -1+\sin\theta & \cos\theta \end{bmatrix}$$
dla pewnych $c, \theta \in \mathbb{R}$.

**Zadanie 42.55.** Wyznaczyć wszystkie takie macierze $A \in M_{2\times 2}(\mathbb{R})$, że $(I + iA)^{-1} = I - iA$.

**Zadanie 42.56.** Niech $A \in M_{2\times 2}(\mathbb{C})$. Pokazać, że $A^2 = A$ wtedy i tylko wtedy, gdy $(2A - I)^2 = I$. Uzasadnić, ze jeśli $A^2 = A$, to $I + A$ jest nieosobliwa oraz

$$(I + A)^{-1} = I - \frac{1}{2}A.$$

**Zadanie 42.57.** Niech $A \in M_{2\times 2}(\mathbb{C})$. Pokazać, że

- jeśli $A^2 = A$ i $\alpha \neq -1$, to $(I + \alpha A)^{-1} = I - \frac{\alpha}{\alpha+1}A$,
- jeśli $A^2 = -A$ i $\alpha \neq 1$, to $(I + \alpha A)^{-1} = I + \frac{\alpha}{\alpha-1}A$,
- jeśli $A^2 = 0$ i $\alpha \in \mathbb{C}$, to $(I + \alpha A)^{-1} = I - \alpha A$.

**Zadanie 42.58.** Niech $A, B, C \in M_{2\times 2}(\mathbb{C})$ będą takie, że

$$A^2 = BC, \quad B^2 = CA, \quad C^2 = AB.$$

Pokazać, że $A^3 = B^3 = C^3$.

**Zadanie 42.59.** Pokazać, że jeśli $A, B \in M_{n\times n}(\mathbb{F})$ są takie, że $AB = A + B$, to $AB = BA$.

**Zadanie 42.60.** Uzasadnić, że jeśli $A \in M_{n\times n}(\mathbb{R})$ i $n$ jest nieparzyste, to $A^2 \neq -I$.

**Zadanie 42.61.** Niech $A, B \in M_{n\times n}(\mathbb{F})$. Pokazać, że jeśli $AB = 0$, to $\operatorname{rank} A + \operatorname{rank} B \leqslant n$.

**Zadanie 42.62.** Niech $A \in M_{n\times n}(\mathbb{C})$. Pokazać, że jeśli $v_1$, $v_2$ są wektorami własnymi dla różnych wartości własnych $\lambda_1, \lambda_2 \in \mathbb{C}$, to $v_1 + v_2$ nie jest wektorem własnym $A$.

**Zadanie 42.63.** Niech $A \in M_{m\times n}(\mathbb{R})$ i $B \in M_{n\times m}(\mathbb{R})$. Pokazać, że jeśli $m > n$, to $\det AB = 0$.

**Zadanie 42.64.** Niech $A \in M_{3\times 2}(\mathbb{R})$ i $B \in M_{2\times 3}$ oraz

$$AB = \begin{bmatrix} 8 & 2 & -2 \\ 2 & 5 & 4 \\ -2 & 4 & 5 \end{bmatrix}$$

Obliczyć $BA$.

Wskazówka:

- Sprawdzić, że $\operatorname{rank} AB = 2$. Wynioskować, że $B$ jest epimorfizmem, a $A$ monomorfizmem.
- Sprawdzić, że $(AB)^2 = 9AB$.
- Ponieważ $A(BA - 9I)B = 0$, więc $BA = 9I$.

**Zadanie 42.65.** Niech $m > n$. Załóżmy, że $A \in M_{m\times n}(\mathbb{R})$, $B \in M_{n\times n}(\mathbb{R})$ i $C \in M_{n\times m}(\mathbb{R})$. Pokazać, że jeśli $\operatorname{rank} AC = n$ oraz $ABC = 0$, to

- $\ker A = \{0\}$,
- $\operatorname{rank} C = n$,
- $B = 0$.

**Zadanie 42.66.** Niech $\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} \in M_{2\times 2}(\mathbb{C})$. Wyznaczyć wszystkie liczby zespolone $z \in \mathbb{C}$, dla których $(zA)^2 = A$.

**Zadanie 42.67.** Niech $A \in M_{2\times 2}(\mathbb{R})$. Pokazać, że jeśli $\operatorname{tr} A = \operatorname{tr} A^2 = 0$, to $A^2 = 0$.

**Zadanie 42.68.** Niech $A \in M_{2\times 2}(\mathbb{R})$ będzie nieosobliwa. Definiujemy odwzorowanie

$$L : M_{2\times 2}(\mathbb{R}) \ni B \to AB^T A^{-1} \in M_{2\times 2}(\mathbb{R}).$$

Podać warunek konieczny i wystarczający (na $A$), aby dla dowolnych $B, C \in M_{2\times 2}(\mathbb{R})$ zachodziły warunki

- $L(L(B)) = B$,
- $L(BC) = L(C)L(B)$.

**Zadanie 42.69.** Niech $A \in M_{n\times n}(\mathbb{C})$ będzie taka, że $A^k = A^{k+1}$ dla pewnego $k > 0$. Pokazać, że

$$\operatorname{tr} A^n = \operatorname{tr} A$$

dla każdego $n > 1$.

**Zadanie 42.70.** Niech $A \in M_{2\times 2}(\mathbb{R})$ będzie taka, że $\det A = 1$ oraz $|\operatorname{tr} A| < 2$. Pokazać, że $|\operatorname{tr} A^n| < 2$ dla każdego $n \geqslant 1$.

**Zadanie 42.71.** Niech $A \in M_{2\times 2}(\mathbb{C})$ będzie taka, że $\operatorname{tr} A^n = \operatorname{tr} A^{n+1}$ dla pewnego $n \in \mathbb{N}$. Pokazać, że $A^2 = 0$.

**Zadanie 42.72.** Pokazać, że jeśli $A \in M_{2\times 2}(\mathbb{C})$ oraz $\operatorname{tr} A = 0$, to $\operatorname{tr} A^{2n+1} = 0$ dla $n \in \mathbb{N}$.

**Zadanie 42.73.** Pokazać, że macierze $A, B \in M_{2\times 2}(\mathbb{C})$ mają takie same wielomiany charakterystyczne wtedy i tylko wtedy, gdy $\operatorname{tr} A^k = \operatorname{tr} B^k$ dla $k \in \mathbb{N}$. Wywnioskować, że $A$ jest nilpotentna wtedy i tylko wtedy, gdy $\operatorname{tr} A^k = 0$ dla $k \in \mathbb{N}$.

# Bibliografia

[1] S. Axler, *Linear Algebra Done Right*, Springer, 1997

[2] J. Komorowski, *Od liczb zespolonych do tensorów, spinorów, algebr Liego i kwadryk*, PWN, 1978

[3] S. J. Leon, *Linear Algebra with Applications*, Pearson 2015

[4] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM

[5] D. Poole, *Linear Algebra: A Modern Introduction*, Cengage Learning, 2015

[6] B. Solomon, *Linear Algebra, Geometry and Transformations*, CRC Press, 2015

[7] K. Tapp, *Matrix Groups for Undergraduates*, AMS, 1971.

# Indeks