Algebra

Michał Dobranowski

semestr zimowy 2022
v1.1

Poniższy skrypt zawiera materiał obejmujący wykłady z Algebry prowadzone przez dr. hab. Jakuba Przybyło na pierwszym roku Informatyki na AGH. Jest uzupełniony o dowody niemal wszystkich twierdzeń oraz przykłady i uwagi, które mają zwykle za zadanie umożliwienie głębszego zrozumienia tematu.

Spis treści

§1 Liczy zespolone

Definicja 1.1. Liczba zespolona z to uporządkowana para liczb rzeczywistych. Pierwszy element tej pary to *część rzeczywista*, oznaczana symbolem $\mathrm{Re}(z)$, a drugi to *część urojona*, oznaczana symbolem $\mathrm{Im}(z)$. Zbiór liczb zespolonych oznaczamy przez $\mathbb{C}$.

Liczby zespolone można reprezentować w kilku postaciach, jedna z nich to *postać algebraiczna*. Używając jej, liczba $z = (x, y)$ jest zapisywana jako

$$z = x + iy,$$

gdzie i nazywamy *jednostką urojoną*, która spełnia

$$i^2 = -1.$$

Niech $z_1 = x_1 + iy_1$ oraz $z_2 = x_2 + iy_2$. Określamy:

- dodawanie $z_1 + z_2 = x_1 + x_2 + i(y_1 + y_2)$,
- mnożenie $z_1 z_2 = x_1 x_2 + i x_1 y_2 + i x_2 y_1 + i^2 y_1 y_2$
  $= x_1 x_2 - y_1 y_2 + i(x_1 y_2 + x_2 y_1)$.

Wniosek 1.2

Dodawanie i mnożenie liczb zespolonych jest przemienne i łączne. Mnożenie jest rozdzielne względem dodawania.

Definicja 1.3. Sprzężenie liczby zespolonej $z = x + iy$ to liczba $\overline{z} = x - iy$.

Definicja 1.4. Moduł liczby zespolonej $z = x + iy$ to liczba $|z| = \sqrt{x^2 + y^2}$.

Zachodzi pewna własność, wynikająca ze wzoru skróconego mnożenia:

$$z\overline{z} = (x + iy)(x - iy) = x^2 - i^2 y^2 = x^2 + y^2$$

$$z\overline{z} = |z|^2 \tag{1}$$

Powyższa liczba jest liczbą rzeczywistą, więc znaleźliśmy prosty sposób na dzielenie liczb zespolonych przez siebie, mnożąc licznik i mianownik przez sprzężenie mianownika. Na przykład:

$$\frac{1 + 2i}{-1 - i} = \frac{(1 + 2i)(-1 + i)}{(-1 - i)(-1 + i)} = \frac{-3 - i}{2} = \frac{-3}{2} - \frac{i}{2}.$$

Lemat 1.5

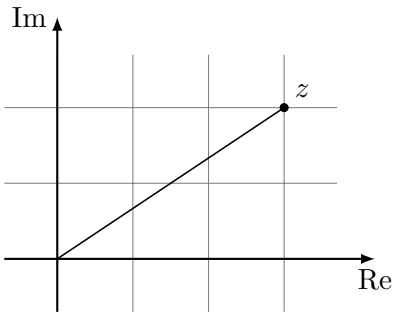
Oprócz $z\overline{z} = |z|^2$, zachodzą również równości:

- $|\overline{z}| = |z|$
- $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$
- $\overline{z_1 z_2} = \overline{z_1} \cdot \overline{z_2}$
- $|z_1 z_2| = |z_1||z_2|$

Ich dowody można w łatwy sposób przeprowadzić z definicji poszczególnych działań.

§1.1 Interpretacja geometryczna liczb zespolonych

Liczby zespolone można interpretować jako punkty na *płaszczyźnie zespolonej*. Dla przykładu liczba $z = 3 + 2i$.



Fakt 1.6. Moduł liczby zespolonej z to długość wektora wodzącego tej liczby na płaszczyźnie zespolonej.

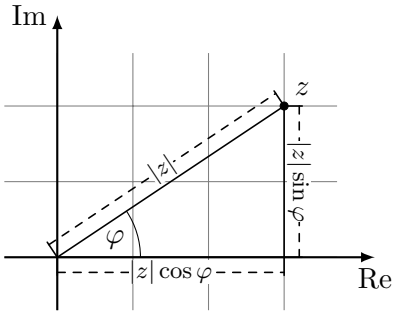
Dowód. Wynika to z twierdzenia Pitagorasa oraz definicji modułu (1.4). □

Możemy wyprowadzić *postać trygonometryczną* liczby zespolonej, która będzie operować na długości wektora wodzącego oraz kącie skierowanym. Mamy więc

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

gdzie φ to miara kąta skierowanego między wektorem wodzącym liczby zespolonej z a osią liczb rzeczywistych. Ten kąt nazywany jest *argumentem* i oznaczany przez $\mathrm{Arg}(z)$. Argument nie jest określony jednoznacznie – dowolne dwa argumenty jednej liczby różnią się o wielokrotność 2π. Jeśli argument jest w przedziale $[0, 2\pi)$, to mówimy, że jest to *argument główny* liczby z i oznaczamy $\arg(z)$.

Za pomocą podstawowej trygonometrii możemy łatwo zamieniać postać algebraiczną i trygonometryczną między sobą.



$$\mathrm{Re}\, z = |z|\cos\varphi, \qquad \mathrm{Im}\, z = |z|\sin\varphi \tag{2}$$

Na potrzeby dalszych rozważań przyjmujemy, że $\arg(0) = 0$.

Fakt 1.7. Odległość między liczbami z_1 i z_2 na płaszczyźnie zespolonej wynosi $|z_1 - z_2|$.

Lemat 1.8

Zachodzą następujące nierówności:

- $|z_1 + z_2| \leq |z_1| + |z_2|$
- $||z_1| - |z_2|| \leq |z_1 - z_2|$

Możemy łatwo mnożyć dwie liczby zespolone w postaci trygonometrycznej przez siebie za pomocą poniższego wzoru.

$$\begin{aligned} z_1 \cdot z_2 &= |z_1|(\cos\varphi_1 + i\sin\varphi_1)|z_2|(\cos\varphi_2 + i\sin\varphi_2) \\ &= |z_1||z_2|(\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2 + i(\cos\varphi_1\sin\varphi_2 + \sin\varphi_1\cos\varphi_2)) \\ &= |z_1||z_2|(\cos(\varphi_1+\varphi_2) + i\sin(\varphi_1+\varphi_2)) \end{aligned} \tag{3}$$

Stosując wzór 3 n razy otrzymujemy dowód następującego twierdzenia.

Twierdzenie 1.9 (wzór de Moivre'a)

Dla $z = |z|(\cos\varphi + i\sin\varphi)$ oraz $n \in \mathbb{Z}$ zachodzi równość

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

Wzór de Moivre'a zapewnia prosty sposób na potęgowanie liczb zespolonych. Dlatego, mając za zadanie obliczyć

$$(-2\sqrt{3} - 2i)^{16}$$

najłatwiej będzie zmienić postać liczby do postaci trygonometrycznej, a następnie skorzystać ze wzoru de Moivre'a.

Definicja 1.10 (pierwiastek liczby zespolonej)**.** Jeśli z jest liczbą zespoloną, to $\sqrt[n]{z}$ jest zbiorem wszystkich takich $w \in \mathbb{C}$, że $w^n = z$.

Korzystając ze wzoru de Moivre'a (twierdzenie 1.9), łatwo wyprowadzić wzór

$$\sqrt[n]{z} = \sqrt[n]{|z|}\left(\cos\frac{\varphi + 2k\pi}{n} + i\sin\frac{\varphi + 2k\pi}{n}\right), k \in \mathbb{Z} \tag{4}$$

Fakt 1.11. Pierwiastków n-tego stopnia z $z \neq 0$ jest dokładnie n i leżą one w równych odstępach na okręgu o środku w 0 i promieniu $\sqrt[n]{|z|}$.

Dowód. Dla $k \in \{0, 1, \ldots, n-1\}$ liczba z równości 4 będzie przyjmować różne wartości (wynika to z okresowości funkcji trygonometrycznych). Liczby te będą na wspomnianym okręgu (to wynika wprost z postaci trygonometrycznej), a ich argumenty główne różnić będzie wielokrotność $\frac{2\pi}{n}$. □

§1.2 Postać wykładnicza

Postać $z = |z|e^{i\varphi}$ liczby zespolonej będziemy nazywać *postacią wykładniczą* tej liczby.

Twierdzenie 1.12 (wzór Eulera)

Dla każdego $\varphi \in \mathbb{R}$ zachodzi

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

Dowód. Standardowo dowodzi się wzoru Eulera za pomocą szeregów Taylora. Pokażemy inny, mniej oczywisty, ale bardziej elementarny dowód.

Niech $f(\varphi) = e^{-i\varphi}(\cos\varphi + i\sin\varphi)$. Zróżniczkujmy:

$$f'(\varphi) = -ie^{-i\varphi}(\cos\varphi + i\sin\varphi) + e^{-i\varphi}(-\sin\varphi + i\cos\varphi) =$$
$$= e^{-i\varphi}(-i\cos\varphi + i\cos\varphi + \sin\varphi - \sin\varphi) = 0.$$

Z tego wynika, że funkcja f jest stała, więc

$$f(\varphi) \equiv f(0) = 1,$$

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

□

§2 Relacje

Definicja 2.1. Relacja to trójka $\mathcal{R} = (X, \operatorname{gr}\mathcal{R}, Y)$, gdzie X i Y są zbiorami, a $\operatorname{gr}\mathcal{R} \subset X \times Y$.

Zbiór X nazywamy *naddziedziną*, Y *zapasem*, $\operatorname{gr}\mathcal{R}$ to *wykres* relacji. Piszemy, że $x\mathcal{R}y$, jeśli $(x, y) \in \operatorname{gr}\mathcal{R}$. *Dziedzina* relacji $\mathcal{R}$ to zbiór

$$D_{\mathcal{R}} = \{x \in X : \exists y \in Y : x\mathcal{R}y\},$$

a jej *przeciwdziedzina* to zbiór

$$\mathcal{C}_{\mathcal{R}} = \{y \in Y : \exists x \in X : x\mathcal{R}y\}.$$

Definicja 2.2. Relacja odwrotna do relacji $\mathcal{R} = (X, \operatorname{gr}\mathcal{R}, Y)$ to taka relacja $\mathcal{R}^{-1} = (Y, \operatorname{gr}\mathcal{R}^{-1}, X)$, że

$$\operatorname{gr}\mathcal{R}^{-1} = \{(y, x) \in Y \times X : (x, y) \in \operatorname{gr}\mathcal{R}\}.$$

Definicja 2.3. Złożenie relacji $\mathcal{R} = (X, \operatorname{gr}\mathcal{R}, Y)$ z relacją $\mathcal{S} = (Y, \operatorname{gr}\mathcal{S}, Z)$ to relacja

$$\mathcal{S} \circ \mathcal{R} = (X, \operatorname{gr}(\mathcal{S} \circ \mathcal{R}), Z),$$

gdzie

$$\operatorname{gr}(\mathcal{S} \circ \mathcal{R}) = \{(x, z) \in X \times Z : \exists y \in Y : x\mathcal{R}y \wedge y\mathcal{S}z\}.$$

Definicja 2.4 (rodzaje relacji)**.** Relacja $\mathcal{R} = (X, \operatorname{gr}\mathcal{R}, X)$ jest:

- *zwrotna* $\Leftrightarrow \forall x \in X : x\mathcal{R}x$,
- *symetryczna* $\Leftrightarrow \forall x, y \in X : x\mathcal{R}y \Rightarrow y\mathcal{R}x$,
- *antysymetryczna* $\Leftrightarrow \forall x, y \in X : x\mathcal{R}y \wedge y\mathcal{R}x \Rightarrow x = y$,
- *asymetryczna* $\Leftrightarrow \forall x, y \in X : x\mathcal{R}y \Rightarrow \neg y\mathcal{R}x$,
- *przechodnia* $\Leftrightarrow \forall x, y, z \in X : x\mathcal{R}y \wedge y\mathcal{R}z \Rightarrow x\mathcal{R}z$,
- *spójna* $\Leftrightarrow \forall x, y \in X : x\mathcal{R}y \vee y\mathcal{R}x \vee x = y$.

Definicja 2.5. Relacja równoważności to relacja $\mathcal{R} = (X, \operatorname{gr}\mathcal{R}, X)$, która jest zwrotna, przechodnia i symetryczna.

Definicja 2.6. Jeżeli $(X, \mathcal{R})$ zbiorem z relacją równoważności, to dla każdego $x \in X$ klasą abstrakcji (klasą równoważności) tego elementu nazywamy zbiór

$$[x] = \{y \in X : x\mathcal{R}y\}.$$

Definicja 2.7. Zbiór ilorazowy relacji $\mathcal{R}$ to zbiór klas abstrakcji tej relacji; przyjmujemy oznaczenie

$$X/\mathcal{R} = \{[x] : x \in X\}.$$

Twierdzenie 2.8

Niech $(X, \mathcal{R})$ będzie zbiorem z relacją równoważności. Wtedy

$$\forall x, y \in X : [x] \neq [y] \Leftrightarrow [x] \cap [y] = \emptyset.$$

Dowód wystarczalności. Załóżmy przez sprzeczność, że $[x] \cap [y] \neq \emptyset$, a więc $\exists z \in X : x\mathcal{R}z \wedge y\mathcal{R}z$. Teraz weźmy dowolny element $a \in [x]$. Mamy więc $x\mathcal{R}a$. Korzystając z symetryczności i przechodniości relacji $\mathcal{R}$, mamy

$$a\mathcal{R}x \wedge x\mathcal{R}z \wedge z\mathcal{R}y,$$

$$\therefore y\mathcal{R}a.$$

Z tego wynika, że $[x] \subset [y]$. Analogicznie (przyjmując na początku $a \in [y]$) dostaniemy, że $[y] \subset [x]$, wiec $[x] = [y]$, co jest sprzeczne z założeniem.

Dowód konieczności. Załóżmy przez sprzeczność, że $[x] = [y]$. Wtedy $[x] \cap [y] = [x] \cap [x] = [x]$ nie może być zbiorem pustym, ponieważ ze zwrotności relacji $\mathcal{R}$ wynika, że $x\mathcal{R}x$, więc $[x]$ to zbiór przynajmniej jednoelementowy. □

Z powyższego twierdzenie wynika, że relacja równoważności w danym zbiorze X dzieli ten zbiór na niepuste i rozłączne podzbiory, których suma daje cały zbiór X.

§2.1 Porządki

Definicja 2.9. Porządek (częściowy) to relacja $\mathcal{R} = (X, \operatorname{gr}\mathcal{R}, X)$, która jest zwrotna, przechodnia i antysymetryczna. Zbiór X nazywamy zbiorem (częściowo) uporządkowanym.

Definicja 2.10. Porządek liniowy (totalny) to porządek, który jest spójny.

Niech $(X, \preceq)$ będzie zbiorem z porządkiem częściowym. Wtedy *element największy* $\overline{M} \in X$ zbioru X to taki element, że

$$\forall x \in X : x \preceq \overline{M},$$

a *element maksymalny* $M_{\max} \in X$ to taki element, że

$$\forall x \in X : (M_{\max} \preceq x) \Rightarrow (M_{\max} = x).$$

Analogicznie definiujemy *element najmniejszy* $\overline{m} \in X$:

$$\forall x \in X : \overline{m} \preceq x$$

oraz *element minimalny* $m_{\min} \in X$:

$$\forall x \in X : x \preceq m_{\min} \Rightarrow (x = m_{\min})$$

Twierdzenie 2.11

Niech $(X, \preceq)$ będzie zbiorem z porządkiem częściowym. Jeśli w zbiorze X istnieje element największy, to jest on jedyny.

Dowód. Załóżmy przeciwnie, że istnieją dwa elementy największe M_1, M_2. Z definicji zachodzi

$$M_1 \preceq M_2$$

oraz

$$M_2 \preceq M_1,$$

co jest sprzeczne z antysymetrycznością porządków. □

Twierdzenie 2.12

Niech $(X, \preceq)$ będzie zbiorem z porządkiem częściowym. Jeśli $M \in X$ jest elementem największym zbioru X, to jest on jedynym elementem maksymalnym tego zbioru.

Dowód. Skoro M jest elementem największym, to poprzednik implikacji w definicji elementu maksymalnego będzie prawdziwy tylko dla $x = M$, więc sama implikacja zawsze będzie prawdziwa. □

Oczywiście dwa powyższe twierdzenia są prawdziwe również odpowiednio dla elementów najmniejszych/minimalnych.

Fakt 2.13. W zbiorach z porządkiem totalnym pojęcia elementu największego i maksymalnego oraz najmniejszego i minimalnego są tożsame ze sobą. Wynika to ze spójności porządków totalnych.

Niech $(X, \preceq)$ będzie zbiorem uporządkowanym, a zbiór $A \subset X$ jego podzbiorem. Element $M \in X$ jest *majorantą* (ograniczeniem górnym) zbioru A jeśli

$$\forall x \in A : x \preceq M.$$

Kresem górnym (supremum) zbioru A (w zbiorze X) jest element najmniejszy zbioru majorant. Oznaczamy go symbolem

$$\sup A.$$

Analogicznie definiujemy *minorantę* (ograniczenie dolne) $m \in X$ zbioru $A \subset X$:

$$\forall x \in A : m \preceq x$$

oraz *kres dolny* (infimum) tego zbioru (jest nim element największy zbioru minorant), który oznaczamy symbolem

$$\inf A.$$

Twierdzenie 2.14

Niech $(X, \preceq)$ będzie zbiorem z porządkiem częściowym oraz $A \subset X$. Jeśli A ma element największy, to jest on również supremum tego zbioru.

Dowód. Z definicji majoranty wynika, że element największy zbioru A jest również jego majorantą. Każda majoranta $M \in X$ zbioru A oczywiście jest „większa" niż dowolny element zbioru A (w tym również jego element największy $\overline{M}$), to znaczy

$$\forall M : \overline{M} \preceq M,$$

z czego wynika, że $\overline{M}$ jest elementem najmniejszym zbioru majorant zbioru A, a więc supremum tego zbioru. □

Wniosek 2.15

Jeśli zbiór częściowo uporządkowany X ma supremum, które nie należy do tego zbioru, to zbiór X nie ma elementu największego.

Dowód. Ponieważ dowolny zbiór (na mocy twierdzenia 2.11) ma co najwyżej jedno supremum, to gdyby zbiór X miał element największy, to na mocy twierdzenia 2.14 byłoby ono również supremum, które należy do zbioru X. □

Oczywiście dwa poprzednia twierdzenia są również prawdziwe odpowiednio dla elementów najmniejszych/infimów.

Przykład 2.16

Weźmy zbiór liniowo uporządkowany $(\mathbb{R}, \leq)$ oraz jego podzbiór $A = [0, 1) \subset \mathbb{R}$. Zbiór majorant zbioru A to przedział $[1, \infty)$, a jego najmniejszy element (a zarazem supremum zbioru A) to liczba 1. Mamy więc

$$\sup A = 1.$$

Liczba 1 nie należy jednak do zbioru A, więc, na mocy wniosku 2.15, element największy (a z faktu 2.13 również maksymalny) nie istnieje.

Przykład 2.17

Weźmy zbiór częściowo uporządkowany $(\mathbb{C}, \preceq)$, gdzie zdefiniujemy

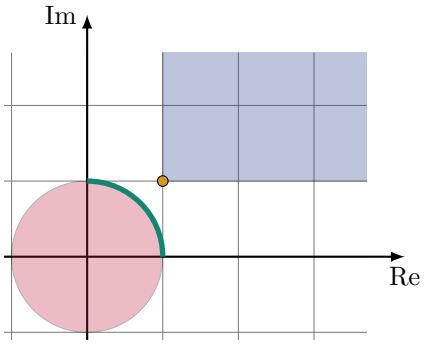
$$x \preceq y \Leftrightarrow \operatorname{Re} x \leq \operatorname{Re} y \wedge \operatorname{Im} x \leq \operatorname{Im} y$$

oraz podzbiór $A \subset \mathbb{C}$ taki, że

$$A = \{z : |z| \leq 1\}.$$

Znajdź zbiór majorant i elementów maksymalnych. Stwierdź, czy istnieje supremum i element największy.

Rozwiązanie. Na rysunku zaznaczono zbiór A (■), zbiór majorant M zbioru A (■), supremum zbioru A (■) oraz zbiór elementów maksymalnych (■). Na mocy wniosku 2.15 element największy nie istnieje.



□

Definicja 2.18. Łańcuch to taki podzbiór $C \subset X$, że $(X, \preceq)$ jest zbiorem z porządkiem częściowym, a $(C, \preceq)$ jest zbiorem z porządkiem liniowym.

Definicja 2.19. Silny porządek to relacja, która jest przechodnia i asymetryczna. Silnie uporządkowany zbiór X oznaczamy przez $(X, \prec)$.

§3 Struktury algebraiczne

Działaniem (wewnętrznym) w zbiorze A nazwiemy każde odwzorowanie h takie, że

$$h : A \times A \to A.$$

Działaniem zewnętrznym w zbiorze A jest odwzorowanie

$$h : F \times A \to A.$$

Jeśli zamiast h weźmiemy jakiś symbol, na przykład $\circ$, to, zamiast $h(a, b)$ będziemy pisać $a \circ b$.

Definicja 3.1 (rodzaje działań)**.** W zbiorze z działaniem $(A, \circ)$ działanie $\circ$ jest:

- *łączne* $\Leftrightarrow \forall x, y, z \in A : (x \circ y) \circ z = x \circ (y \circ z)$,
- *przemienne* $\Leftrightarrow \forall x, y, \in A : x \circ y = y \circ x$.

Jeśli dla pewnego elementu $e \in A$ zachodzi

$$\forall x \in A : x \circ e = e \circ x = x,$$

to e jest *elementem neutralnym*.

Fakt 3.2. Jeżeli w zbiorze A z działaniem $\circ$ istnieje element neutralny, to jest on jedyny.

Dowód. Jeśli mielibyśmy dwa elementy neutralne e_1, e_2 to mamy

$$e_1 \circ e_2 = e_1 = e_2.$$

□

Jeżeli istnieje element neutralny $e \in A$ działania $\circ$, to *elementem symetrycznym* do $x \in A$ jest taki element $x' \in A$, że

$$x \circ x' = e = x' \circ x.$$

Lemat 3.3

Jeśli działanie $\circ$ jest łączne w zbiorze A i istnieje element neutralny $e \in A$, to jeśli dany element $x \in A$ ma element symetryczny, to jest on jedyny oraz zachodzi $(x')' = x$.

Dowód. Jeśli mielibyśmy dwa elementy symetryczne x_1', x_2', to mamy

$$x_1' = x_1' \circ e = x_1' \circ (x \circ x_2') = (x_1' \circ x) \circ x_2' = e \circ x_2' = x_2'.$$

Ponadto z definicji elementu symetrycznego mamy

$$x' \circ x = e$$

oraz

$$x' \circ (x')' = e,$$

a więc x jest elementem symetrycznym x', ergo $(x')' = x$. □

§3.1 Grupy

Definicja 3.4. Grupa to para $(A, \circ)$, gdzie A jest zbiorem, a działanie $\circ$ jest:
1. wewnętrzne,
2. łączne,
3. ma element neutralny,
4. a każdy element $x \in A$ ma element symetryczny.

Definicja 3.5. Grupa abelowa (przemienna) to grupa, w której działanie $\circ$ jest przemienne.

Przykład 3.6

Przykłady grup:

1. $(\mathbb{Z}, +)$ – grupa abelowa,
2. $(\mathbb{Z}_n, +_n)$ – grupa abelowa[a],
3. $(\mathbb{Q}_+, \cdot)$ – grupa abelowa,
4. grupą nieabelową jest grupa obrotów danego obiektu o $90°$ względem dowolnej z trzech osi.

[a] Symbol $\mathbb{Z}_n$ oznacza zbiór $\{0, 1, \ldots, n-1\}$, a $+_n$ operację dodawania modulo n.

Twierdzenie 3.7

$(\mathbb{Z}_n \setminus \{0\}, \cdot_n)$ jest grupą wtedy i tylko wtedy, gdy $n \geq 2$ jest liczbą pierwszą.

Łatwo sprawdzić, że mnożenie modulo n w zbiorze $\mathbb{Z}_n \setminus \{0\}$ jest wewnętrzne i łączne. Ma również element neutralny 1. Będziemy więc dowodzić jedynie istnienia elementu symetrycznego dla każdego elementu.

Dowód wystarczalności. Załóżmy przeciwnie, że istnieje $k \in \mathbb{Z}_n \setminus \{0, 1\}$ takie, że $k \mid n$. Skoro $(\mathbb{Z}_n \setminus \{0\}, \cdot_n)$ jest grupą, to k ma element symetryczny k^{-1}. Zachodzi więc

$$kk^{-1} \equiv 1 \pmod{n},$$

czyli inaczej

$$\exists m \in \mathbb{Z} : kk^{-1} - 1 = mn.$$

Co jednak prowadzi do sprzeczności, ponieważ

$$kk^{-1} - 1 \not\equiv mn \pmod{k}$$

$$-1 \not\equiv 0 \pmod{k}.$$

Dowód dostateczności. Skoro n jest liczbą pierwszą, to z małego twierdzenia Fermata mamy

$$a^{n-1} \equiv 1 \pmod{n}$$

dla każdego $a \in \mathbb{Z}_n \setminus \{0\}$. Z tego wynika, że dla dowolnego elementu a jego elementem symetrycznym będzie a^{n-2}. □

§3.2 Pierścienie i ciała

Definicja 3.8. Pierścień to trójka $(P, \circ, *)$, gdzie P jest zbiorem, $\circ, *$ to działania wewnętrzne oraz

1. $(P, \circ)$ jest grupą abelową
2. działanie $*$ jest łączne
3. działanie $*$ jest rozdzielne względem $\circ$, czyli

$$\forall x, y, z \in P : \begin{array}{l} (x \circ y) * z = (x * z) \circ (y * z), \\ x * (y \circ z) = (x * y) \circ (x * z). \end{array}$$

Definicja 3.9. Pierścień przemienny to pierścień $(P, \circ, *)$, w którym $*$ jest działaniem przemiennym[1].

Pierwsze działanie w pierścieniu nazywamy *działaniem addytywnym* i oznaczamy symbolem $+$. Element neutralny tego działania nazywamy zerem ($\mathbf{0}$), a element symetryczny do elementu x nazywamy elementem przeciwnym i oznaczamy $-x$.

Drugie działanie nazywamy *działaniem multiplikatywnym* i oznaczamy przez $\cdot$. Jeśli w P dodatkowo istnieje element neutralny tego działania, to ten element nazywamy jedynką ($\mathbf{1}$), a pierścień nazywamy *pierścieniem z jedynką*. Element symetryczny do elementu x nazywamy elementem odwrotnym i oznaczamy x^{-1}.

Definicja 3.10. Dzielnikiem zera jest taki element pierścienia $a \neq \mathbf{0}$, że istnieje niezerowy element b, dla którego zachodzi $a \cdot b = \mathbf{0}$.

Definicja 3.11. Pierścień całkowity to pierścień przemienny z jedynką, w którym nie ma dzielników zera.

Lemat 3.12

W pierścieniach całkowitych zachodzi *własność skracania*, to znaczy, że dla elementów pierścienia a, b, c przy $c \neq \mathbf{0}$ zachodzi

$$ac = bc \Rightarrow a = b.$$

Dowód. Jeśli $ac = bc$, to $ac - bc = \mathbf{0}$. Z rozdzielności dostajemy

$$(a - b)c = \mathbf{0}.$$

W pierścieniu całkowitym nie ma jednak dzielników zera, więc $a - b = \mathbf{0}$, co dowodzi tezy. □

Definicja 3.13. Ciało to pierścień z jedynką, w którym dla każdego elementu $x \neq \mathbf{0}$ istnieje element odwrotny x^{-1}.

Ciałem przemiennym będzie takie ciało, w którym działanie multiplikatywne $\cdot$ jest przemienne[2].

Można zauważyć, że struktura $(K, +, \cdot)$ jest ciałem (przemiennym), jeżeli:

[1] Wtedy też rozdzielność prawo- i lewostronna stają się tożsame.

[2] Większość autorów już w definicji ciała wymaga przemienności (wtedy ciało nazywamy pierścieniem z dzieleniem). Przyjęło się tak zwłaszcza w literaturze angielskiej (ciało przemienne to *field*, a ciało to *division ring* lub *skew field*) i niemieckiej (odpowiednio *Körper* i *Schiefkörper*). Odwrotnie — czyli zgodnie z naszą konwencją — jest w literaturze francuskiej (odpowiednio *corps commutatif* i *corps*) oraz rosyjskiej (*поле* [pole] i *тело* [tieło]).

1. $(K, +)$ jest grupą abelową,
2. $(K \setminus \{\mathbf{0}\}, \cdot)$ jest grupą (przemienną),
3. zachodzi warunek rozdzielności $\cdot$ względem $+$.

Lemat 3.14

Dla każdego elementu ciała a zachodzi $a \cdot \mathbf{0} = \mathbf{0}$.

Dowód.

$$\begin{aligned} a \cdot \mathbf{0} &= a \cdot (\mathbf{0} + \mathbf{0}) \\ a \cdot \mathbf{0} &= a \cdot \mathbf{0} + a \cdot \mathbf{0} \\ a \cdot \mathbf{0} + -a \cdot \mathbf{0} &= a \cdot \mathbf{0} + a \cdot \mathbf{0} + -a \cdot \mathbf{0} \\ \mathbf{0} &= a \cdot \mathbf{0} + \mathbf{0} \\ \mathbf{0} &= a \cdot \mathbf{0} \end{aligned}$$

□

Twierdzenie 3.15

Każde ciało przemienne jest pierścieniem całkowitym.

Dowód. Załóżmy przeciwnie, że istnieją dzielniki zera, czyli takie dwa elementy ciała x, y, że $x, y \neq \mathbf{0}$ oraz $x \cdot y = \mathbf{0}$. Mamy

$$\begin{aligned} x \cdot y &= \mathbf{0} \\ x^{-1} \cdot x \cdot y &= x^{-1} \cdot \mathbf{0} \\ y &= x^{-1} \cdot \mathbf{0}, \end{aligned}$$

co, na mocy lematu 3.14, jest sprzecznością z założeniem. □

Twierdzenie 3.16

Każdy skończony pierścień całkowity jest ciałem przemiennym.

Dowód. Załóżmy przeciwnie, że istnieje element pierścienia $a \neq \mathbf{0}$, który nie ma elementu odwrotnego. Rozważmy iloczyny $aa_1, aa_2, aa_3, \ldots$ elementu a ze wszystkimi innymi elementami pierścienia (w tym z $\mathbf{1}$). Z założenia nie ma wśród nich jedynki, więc, skoro $\cdot$ jest działaniem wewnętrznym, to z zasady szufladkowej istnieją takie $a_k \neq a_l$, że $aa_k = aa_l$. To stwierdzenie jest jednak sprzecznością na mocy lematu 3.12, ponieważ rozważamy pierścienie całkowite, w których nie ma dzielników zera. □

Przykład 3.17

Przykłady pierścieni i ciał:

- $(\mathbb{Z}, +, \cdot)$ – pierścień całkowity, który nie jest ciałem (nie ma dzielników zera, ale często elementy odwrotne nie zawierają się w zbiorze $\mathbb{Z}$),
- $(\mathbb{Q}, +, \cdot)$ – ciało przemienne liczb wymiernych,
- $(\mathbb{R}, +, \cdot)$ – ciało przemienne liczb rzeczywistych,
- $(\mathbb{C}, +, \cdot)$ – ciało przemienne liczb zespolonych,
- $(\mathbb{Z}_n, +_n, \cdot_n)$ – pierścień przemienny z jedynką.

Wniosek 3.18 (z twierdzenia 3.7)

Pierścień $(\mathbb{Z}_n, +_n, \cdot_n)$ jest ciałem wtedy i tylko wtedy, gdy n jest liczbą pierwszą.

§3.3 Morfizmy

Definicja 3.19. Homomorfizmem grupy $(A_1, +)$ w grupę $(A_2, \oplus)$ jest takie odwzorowanie $h : A_1 \to A_2$, że

$$\forall x, y \in A_1 : h(x + y) = h(x) \oplus h(y).$$

Fakt 3.20. Jeśli $h : A_1 \to A_2$ jest homomorfizmem grupy $(A_1, +)$ w $(A_2, \oplus)$, to

1. $e \in A_1$ jest elementem neutralnym w $(A_1, +) \implies h(e) \in A_2$ jest elementem neutralnym w $(A_2, \oplus)$,
2. $\forall x \in A_1 : h(x') = h(x)'$.

Definicja 3.21. Izomorfizm między grupami $(A_1, +), (A_2, \oplus)$ jest homomorfizmem bijektywnym. Jeśli taki izomorfizm istnieje, to dwie grupy nazywamy izomorficznymi.

Definicja 3.22. Automorfizm to izomorfizm struktury na samą siebie.

Analogicznie definiujemy morfizmy między pierścieniami i ciałami (wtedy równość z definicji 3.19 musi zachodzić dla obydwu działań).

Przykład 3.23

Przykłady morfizmów:

- $h(x) = x^2$ jest homomorfizmem grupy $(\mathbb{R} \setminus \{0\}, \cdot)$ w $(\mathbb{R}_+, \cdot)$,
- $h(x) = e^x$ jest izomorfizmem grupy $(\mathbb{R}, +)$ w $(\mathbb{R}_+, \cdot)$, ponieważ

$$h(x + y) = e^{x+y} = e^x \cdot e^y = h(x) \cdot g(y),$$

- $h(z) = \overline{z}$ jest automorfizmem grupy $(\mathbb{C}, +)$.

Na podobnej zasadzie jak w przykładzie drugim, można pokazać izomorfizm grupy $(\mathbb{Z}_n, +_n)$ z grupą pierwiastków n-tego stopnia z jedności względem mnożenia $(\mu_n(\mathbb{C}), \cdot)$. Biorąc funkcję $h(x) = \cos(\frac{2\pi}{n}x) + i\sin(\frac{2\pi}{n}x)$, mamy

$$\begin{aligned} h(x + y) &= \cos(\tfrac{2\pi}{n}(x + y)) + i\sin(\tfrac{2\pi}{n}(x + y)) \\ &= \big(\cos(\tfrac{2\pi}{n}x) + i\sin(\tfrac{2\pi}{n}x)\big) \cdot \big(\cos(\tfrac{2\pi}{n}y) + i\sin(\tfrac{2\pi}{n}y)\big) = h(x) \cdot h(y) \end{aligned}$$

§3.4 Przestrzenie wektorowe

Definicja 3.24. Przestrzeń wektorowa (inaczej liniowa) nad ciałem $(K, \oplus, \otimes)$ to struktura $(V, K, +, \cdot)$, gdzie

1. $(V, +)$ jest grupą abelową,
2. działanie $\cdot : K \times V \to V$ jest zewnętrzne
3. działanie $\cdot$ jest rozdzielne względem działania $+$, to znaczy
$$\bigvee_{u,v \in V} \bigvee_{\alpha \in K} \alpha \cdot (u + v) = (\alpha \cdot u) + (\alpha \cdot v),$$
4. zachodzi „rozdzielność" działania $\cdot$ względem $+$ i $\oplus$, to znaczy
$$\bigvee_{v \in V} \bigvee_{\alpha, \beta \in K} (\alpha \oplus \beta) \cdot v = (\alpha \cdot v) + (\beta \cdot v),$$
5. zachodzi „łączność" działań $\cdot$ i $\otimes$, to znaczy
$$\bigvee_{v \in V} \bigvee_{\alpha, \beta \in K} (\alpha \otimes \beta) \cdot v = \alpha \cdot (\beta \cdot v),$$
6. jedynka z ciała $(K, \oplus, \otimes)$ jest elementem neutralnym również dla działania $\cdot$, to znaczy
$$\bigvee_{v \in V} \mathbf{1} \cdot v = v.$$

Elementy zbioru V nazywamy *wektorami*, a zbioru K – *skalarami*. Często zamiast przestrzeni $(V, K, +, \cdot)$ piszemy o przestrzeni V, a zamiast symboli $\oplus, \otimes$ piszemy po prostu $+, \cdot$. Element neutralny dodawania wektorów to wektor zerowy $\overline{0}$.

Przykład 3.25

Przestrzenią wektorową nad ciałem liczb rzeczywistych jest struktura $(\mathbb{R}^n, \mathbb{R}, +, \cdot)$, często oznaczana jako $\mathbb{R}^n(\mathbb{R})$, gdzie

- $(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$,
- $\alpha \cdot (x_1, x_2, \ldots, x_n) = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n)$.

Przykład 3.26

Jeśli przez $\mathbb{R}[x]_n$ oznaczymy zbiór wielomianów rzeczywistych o stopniu równym co najwyżej n, to struktura
$$(\mathbb{R}[x]_n, \mathbb{R}, +, \cdot)$$
będzie przestrzenią liniową.

Twierdzenie 3.27

W przestrzeni liniowej $(V, K, +, \cdot)$ dla każdych $u, v \in V$ oraz $\alpha, \beta \in K$ zachodzą następujące własności:

1. $\mathbf{0} \cdot v = \overline{0}$,
2. $\alpha \cdot \overline{0} = \overline{0}$,
3. $(-\alpha) \cdot v = -(\alpha \cdot v)$,
4. $\alpha \cdot (-v) = -(\alpha \cdot v)$,
5. $\alpha \cdot v = \overline{0} \Leftrightarrow (\alpha = \mathbf{0} \vee v = \overline{0})$,
6. $\alpha \cdot u = \alpha \cdot v \Rightarrow u = v$, dla $\alpha \neq \mathbf{0}$,
7. $\alpha \cdot v = \beta \cdot v \Rightarrow \alpha = \beta$, dla $v \neq \overline{0}$.

Dowód. W dowodach wszystkich własności posługujemy się wyłącznie definicją przestrzeni wektorowej (3.24), wektora zerowego oraz poprzednimi w kolejności udowadnianymi własnościami.

1. $v + \mathbf{0} \cdot v = \mathbf{1} \cdot v + \mathbf{0} \cdot v = (\mathbf{1} + \mathbf{0}) \cdot v = v = v + \overline{0}$
   $\therefore \mathbf{0} \cdot v = \overline{0}$
2. $\alpha \cdot \overline{0} = \alpha \cdot (\overline{0} + \overline{0}) = \alpha \cdot \overline{0} + \alpha \cdot \overline{0}$
   $\therefore \overline{0} = \alpha \cdot \overline{0}$
3. $\overline{0} = \alpha \cdot v - (\alpha \cdot v)$ oraz $\overline{0} = \mathbf{0} \cdot v = (\alpha - \alpha) \cdot v = \alpha \cdot v + (-\alpha) \cdot v$
   $\therefore -(\alpha \cdot v) = (-\alpha) \cdot v$
4. $\overline{0} = \alpha \cdot v - (\alpha \cdot v)$ oraz $\overline{0} = \alpha \cdot \overline{0} = \alpha \cdot (v - v) = \alpha \cdot v + \alpha \cdot (-v)$
   $\therefore -(\alpha \cdot v) = \alpha \cdot (-v)$
5. implikacja $\Leftarrow$ (konieczność) trywialna; implikacja $\Rightarrow$ (dostateczność) wynika z tego, że jeśli założymy, że $\alpha \neq \mathbf{0}, v \neq \overline{0}$, to mamy
   $$\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v = \alpha \cdot u.$$
   Mnożąc przez a^{-1} (które istnieje, bo $(K, +, \cdot)$ jest ciałem) otrzymujemy
   $$u + v = u,$$
   a dodając obustronnie $-u$ (które istnieje z definicji 3.24) dochodzimy do sprzeczności z założeniem
   $$v = \overline{0}.$$
6. dowód analogiczny do dowodu lematu 3.12,
7. dowód analogiczny do dowodu lematu 3.12.

□

Definicja 3.28. Podprzestrzeń liniowa $(U, K, +, \cdot)$ to taka struktura, że

1. $(V, K, +, \cdot)$ jest przestrzenią liniową oraz $U \subset V, U \neq \emptyset$,
2. $\underset{u,v \in U}{\forall} (u + v) \in U$,
3. $\underset{\alpha \in K}{\forall} \underset{u \in U}{\forall} (\alpha \cdot u) \in U$.

Fakt 3.29 (równoważna charakterystyka podprzestrzeni)**.** Dwa ostatnie warunki z powyższej definicji są równoważne warunkowi:

$$\underset{\alpha,\beta \in K}{\forall} \underset{u,v \in V}{\forall} \alpha \cdot u + \beta \cdot v \in U.$$

Dowód. Implikacja w jedną stroną jest trywialna, w drugą stronę można ją udowodnić przez stwierdzenie, że każdy wektor ma wektor przeciwny (bo z definicji 3.24 $(V, +)$ jest grupą abelową) oraz że pod α, β można podstawić $\mathbf{1}$ (i znowu użyć definicji 3.24). □

Definicja 3.30. Kombinacja liniowa wektorów $v_1, v_2, \ldots, v_n$ to wektor

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n,$$

gdzie skalary $\alpha_1, \alpha_2, \ldots, \alpha_n$ nazywamy współczynnikami tej kombinacji.

Definicja 3.31. Wektory $v_1, v_2, \ldots, v_n$ są liniowo niezależne, jeśli dla każdego ciągu współczynników α zachodzi implikacja

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = \overline{0} \Rightarrow \alpha_1, \alpha_2, \ldots, \alpha_n = 0.$$

Mówimy również, że wektory są liniowo zależne, jeśli nie są liniowo niezależne.

Przykład 3.32

W przestrzeni wektorowej $\mathbb{R}^3(\mathbb{R})$ weźmy wektory

$$u = (3, 2, -1), v = (1, -2, 1), w = (1, 1, 1).$$

Rozwiązujemy układ równań $\alpha u + \beta v + \gamma w = \overline{0} \Rightarrow$

$$\begin{cases} 3\alpha + \beta + \gamma = 0 \\ 2\alpha - 2\beta + \gamma = 0 \\ -\alpha + \beta + \gamma = 0 \end{cases} \Rightarrow \begin{cases} 4\alpha = 0 \\ 2\alpha - 2\beta + \gamma = 0 \\ -\alpha + \beta + \gamma = 0 \end{cases} \Rightarrow \begin{cases} \alpha = 0 \\ -2\beta + \gamma = 0 \\ \beta + \gamma = 0 \end{cases} \Rightarrow \begin{cases} \alpha = 0 \\ \beta = 0 \\ \gamma = 0 \end{cases}$$

pokazując, że wektory u, v, w są liniowo niezależne.

Twierdzenie 3.33

Wektory $v_1, \ldots, v_n$ są liniowo zależne wtedy i tylko wtedy, gdy przynajmniej jeden jest kombinacją liniową pozostałych.

Dowód. Jeśli istnieje taki ciąg $\alpha_1, \alpha_2, \ldots, \alpha_n$, że $\{\alpha_1, \ldots, \alpha_n\} \neq \{0\}$ oraz

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = \overline{0},$$

to bez straty ogólności możemy przyjąć, że $\alpha_n \neq 0$. Równoważnie przekształcamy równość do postaci

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_{n-1} v_{n-1} = -\alpha_n v_n$$
$$\frac{-\alpha_1}{\alpha_n} v_1 + \frac{-\alpha_2}{\alpha_n} v_2 + \ldots + \frac{-\alpha_{n-1}}{\alpha_n} v_{n-1} = v_n,$$

więc otrzymujemy równoważność między założeniem i stwierdzeniem, że v_n jest kombinacją liniową wektorów $\alpha_1, \ldots, \alpha_{n-1}$. □

Twierdzenie 3.34

Jeśli wektory $v_1, v_2, \ldots, v_n$ są liniowo niezależne oraz wektor u jest kombinacją liniową tych wektorów, to współczynniki tej kombinacji są wyznaczone jednoznacznie.

Dowód. Weźmy takie ciągi (α_n) i (β_n), że

$$u = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$$
$$u = \beta_1 v_1 + \beta_2 v_2 + \ldots + \beta_n v_n$$

Mamy

$$u - u = \overline{0} = (\alpha_1 - \beta_1) v_1 + (\alpha_2 - \beta_2) v_2 + \ldots + (\alpha_n - \beta_n) v_n,$$

co, skoro $v_1, v_2, \ldots, v_n$ są liniowo niezależne, dowodzi, że dla każdego i zachodzi $\alpha_i - \beta_i = 0$, więc ciągi (α_n) i (β_n) są równe. □

Definicja 3.35. Powłoka liniowa zbioru $A \subset V, A \neq \emptyset$, gdzie V jest przestrzenią wektorową nad ciałem K to zbiór

$$\operatorname{Lin} A = \{v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_k v_k : \alpha_i \in K, v_i \in A\}$$

$\operatorname{Lin} A$ jest podprzestrzenią przestrzeni A nazywaną podprzestrzenią generowaną przez zbiór A. Dla danego zbioru A mówimy, że *rozpina* on przestrzeń wektorową V, jeśli $\operatorname{Lin} A = V$.

Definicja 3.36. Baza B przestrzeni wektorowej V to taki zbiór, że $\operatorname{Lin} B = V$ oraz wszystkie wektory w B są liniowo niezależne.

B jest bazą danej przestrzeni liniowej wtedy i tylko wtedy, gdy B jest maksymalnym (w sensie inkluzji) zbiorem wektorów liniowo niezależnych oraz wtedy i tylko wtedy, gdy B jest minimalnym (w sensie inkluzji) zbiorem wektorów rozpinających. Przestrzeń $\{\overline{0}\}$ nie ma bazy.

Twierdzenie 3.37

Każde dwie bazy danej przestrzeni wektorowej są równoliczne.

Dowód. Weźmy dwie bazy A, B przestrzeni liniowej V oraz niech $|A| = k$. Załóżmy przeciwnie, że $|B| > k, B = \{b_1, b_2, \ldots, b_k, \ldots\}$. Skoro A jest bazą przestrzeni V, to każdy wektor ze zbioru B jest kombinacją liniową wektorów ze zbioru A, czyli

$$b_1 = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k.$$

Bez straty ogólności możemy założyć, że $\alpha_1 \neq 0$ (ponieważ wszystkie nie mogą być zerowe). Wtedy

$$a_1 = \frac{1}{\alpha_1} b_1 + \frac{-\alpha_2}{\alpha_1} a_2 + \ldots + \frac{-\alpha_k}{\alpha_1} a_k,$$

a więc a_1 jest kombinacją liniową wektorów ze zbioru $A \cup \{b_1\} \setminus \{a_1\}$. Wektory z tego zbioru oczywiście rozpinają całą przestrzeń liniową V oraz są liniowo niezależne (wszystkie $a_2, a_3, \ldots, a_k$ są liniowo niezależne, a b_1 jest liniowo niezależny od nich, ponieważ założyliśmy, że $\alpha_1 \neq 0$). Z tego powodu zbiór $A \cup \{b_1\} \setminus \{a_1\}$ jest bazą. Kontynuujemy rozumowanie, pokazując, że zbiór

$$A \cup \{b_1, b_2, \ldots, b_k\} \setminus \{a_1, a_2, \ldots a_k\} = \{b_1, b_2, \ldots, b_k\}$$

jest bazą. Z tego powodu każdy wektor $b_{k+1}, b_{k+2}, \ldots$ jest liniowo zależny od $\{b_1, \ldots, b_k\}$, więc dochodzimy do sprzeczności z założeniem, że B jest bazą. □

Definicja 3.38. Wymiar $\dim V$ przestrzeni wektorowej V to liczność bazy tej przestrzeni. Jeśli $V = \{\overline{0}\}$, to $\dim V = 0$.

Przykład 3.39

Przestrzeń $(\mathbb{R}^n, \mathbb{R}, +, \cdot)$ jest przestrzenią skończenie wymiarową z

$$\dim \mathbb{R}^n = n,$$

natomiast $(\mathcal{F}(\mathbb{R}, \mathbb{R}), \mathbb{R}, +, \cdot)$ jest przestrzenią nieskończenie wymiarową, więc

$$\dim(\mathcal{F}(\mathbb{R}, \mathbb{R})) = \infty.$$

Definicja 3.40. Reper bazowy (lub po prostu baza) to baza, w której ustaliliśmy kolejność wektorów.

Jeśli $B = (e_1, e_2, \ldots, e_n)$ jest reperem bazowym przestrzeni wektorowej V, to dla dowolnego wektora

$$v = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n$$

skalary α_i nazwiemy *współrzędnymi* wektora v w bazie B i zapiszemy

$$v = [\alpha_1, \alpha_2, \ldots, \alpha_n]_B.$$

Definicja 3.41. Baza kanoniczna to reper bazowy przestrzeni $\mathbb{R}^n(\mathbb{R})$, w którym

$$B_k = ((1,0,0,\ldots,0),(0,1,0,\ldots,0),(0,0,0,\ldots,1)).$$

Łatwo uzasadnić, że jeśli $\dim V = n$, to każdy zbiór $n+1$ wektorów jest liniowo zależny, a każdy zbiór n wektorów jest liniowo niezależny wtedy i tylko wtedy, gdy generuje przestrzeń V.

Twierdzenie 3.42

Niech V będzie przestrzenią skończenie wymiarową, a U jej podprzestrzenią. Wówczas

$$\dim U = \dim V \quad \Leftrightarrow \quad U = V.$$

Implikacja w lewą stronę jest trywialna, pokażemy implikację w prawo.

Dowód. Jeśli weźmiemy pewną bazę B przestrzeni U i zachodzi warunek $\dim U = \dim V$, to zgodnie z tym, co powiedzieliśmy wcześniej, jest ona również bazą przestrzeni V, ponieważ $U \subset V$, z czego wynika teza. □

Przykład 3.43

Przykłady przestrzeni wektorowych wraz z wymiarami:

- dla $(\mathbb{K}^n, \mathbb{K}, +, \cdot)$ przy $\mathbb{K} = \mathbb{R}, \mathbb{C}, \ldots$ mamy $\dim \mathbb{K}^n = n$,
- dla $(\mathbb{C}^n, \mathbb{R}, +, \cdot)$ mamy $\dim \mathbb{C}^n = 2n$.

Definicja 3.44. Suma podprzestrzeni V_1, V_2 przestrzeni V to zbiór

$$V_1 + V_2 = \{v = v_1 + v_2 : v_1 \in V_1, v_2 \in V_2\}.$$

Fakt 3.45. Jeśli V_1, V_2 są podprzestrzeniami przestrzeni V, to $V_1 \cap V_2$ jest podprzestrzenią przestrzeni V.

Uwaga 3.46. O ile $V_1 \cap V_2$ oraz $V_1 + V_2$ (z definicji) są przestrzeniami, o tyle już $V_1 \cup V_2$ na ogół nią nie jest, więc nie będziemy raczej używać tego zapisu.

Definicja 3.47. Suma prosta $V_1 \oplus V_2$ dwóch podprzestrzeni przestrzeni V to taka suma $V_1 + V_2$, że zachodzi warunek

$$\forall_{v \in V_1+V_2} \; \exists!_{v_1 \in V_1} \; \exists!_{v_2 \in V_2} \; v = v_1 + v_2.$$

Twierdzenie 3.48

Suma dwóch podprzestrzeni jest sumą prostą wtedy i tylko wtedy, gdy ich częścią wspólną jest zbiór $\{\overline{0}\}$.

Dowód. Jeśli część wspólna dwóch podprzestrzeni jest równa $\{\overline{0}\}$, to ich bazy są rozłączne, a więc teza wynika z twierdzenia 3.34. □

Definicja 3.49. Przestrzeń uzupełniająca V_2 podprzestrzeni V_1 przestrzeni V to taka przestrzeń, że

$$V_1 \oplus V_2 = V.$$

Fakt 3.50. Dla każdej podprzestrzeni dowolnej przestrzeni istnieje przestrzeń uzupełniająca.

Twierdzenie 3.51 (Grassmana)

Dla skończenie wymiarowych podprzestrzeni V_1, V_2 przestrzeni wektorowej V zachodzi

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2),$$

a w szczególności

$$V = V_1 \oplus V_2 \Rightarrow \dim V = \dim V_1 + \dim V_2.$$

Dowód. Możemy wziąć bazę B przestrzeni V oraz bazy B_1, B_2 odpowiednio podprzestrzeni V_1, V_2 takie, że $B_1, B_2 \subset B$. Oczywiste jest, że

$$|B_1 \cup B_2| = |B_1| + |B_2| - |B_1 \cap B_2|,$$

więc z definicji sumy podprzestrzeni (3.44) i twierdzenia 3.37 wynika teza. □

§4 Macierze

Definicja 4.1. Macierz o wymiarach $m \times n$ i elementach ze zbioru K to odwzorowanie

$$\{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\} \ni (i, j) \to a_{ij} \in K,$$

które reprezentujemy w następujący sposób:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Definicja 4.2. Macierz transponowana do macierzy $A = [a_{ij}]_{m\times n}$ to macierz

$$A^T = [a_{ji}]_{n\times m}.$$

Jeśli $A = A^T$, to macierz jest *symetryczna*.

Macierz zerowa $\mathbf{0}_{m\times n}$ to taka macierz, że wszystkie jej elementy są zerowe. *Macierz kwadratowa* to macierz o wymiarach $n \times n$. *Przekątną główną* macierzy kwadratowej tworzą elementy a_{ii}.

Definicja 4.3. Macierz diagonalna to macierz kwadratowa, w której wszystkie elementy poza jej główną przekątną są zerowe.

Definicja 4.4. Macierz jednostkowa to macierz kwadratowa, w której wszystkie elementy na głównej przekątnej są jedynkami. Oznaczamy ją często I_n, gdzie $n \times n$ to wymiary tej macierzy.

Definicja 4.5. Macierz jest trójkątna górna/dolna, jeśli wszystkie elementy poniżej/powyżej głównej przekątnej są równe 0.

§4.1 Działania na macierzach

Zdefiniowane są pewne działania na macierzach:

Suma macierzy dla macierzy $A = [a_{ij}]_{m\times n}$ i $B = [b_{ij}]_{m\times n}$ tych samych wymiarach:

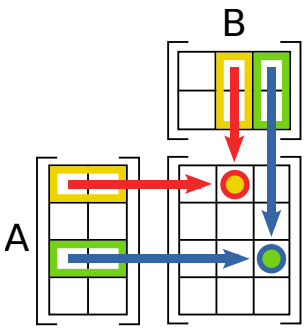
$$A + B = [a_{ij} + b_{ij}]_{m\times n},$$

Mnożenie przez skalar dla macierzy $A = [a_{ij}]_{m\times n}$:

$$\alpha A = [\alpha a_{ij}]_{m\times n},$$

Mnożenie macierzy jeśli liczba kolumn macierzy $A = [a_{ij}]_{m\times p}$ jest równa liczbie wierszy macierzy $B = [a_{ij}]_{p\times n}$, to

$$A \cdot B = [c_{ij}]_{m\times n}, \qquad c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}.$$



Rysunek 1: Mnożenie macierzy, źródło: Wikipedia.

> **Uwaga 4.6.** Mnożenie macierzy nie jest przemienne, jest za to łączne i obustronnie rozdzielne względem dodawania.

Fakt 4.7. Zbiór $\mathcal{M}_{m\times n}(\mathbb{K})$ macierzy o wymiarach $m \times n$ i elementach z ciała przemiennego $\mathbb{K}$, $|\mathbb{K}| \geq 2$ tworzy przestrzeń wektorową nad ciałem $\mathbb{K}$.

Fakt 4.8. Elementem neutralnym mnożenia macierzy kwadratowych jest macierz jednostkowa[3].

Fakt 4.9. Zachodzi równość

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

[3] Jeśli macierz $A_{m\times n}$ nie jest kwadratowa, to również zachodzi $I'M = M$ oraz $M = I''M$ dla pewnych macierzy jednostkowych I', I'', lecz $I' \neq I''$ (są różnych wymiarów.).

§4.2 Wyznacznik macierzy

Definicja 4.10. Inwersja w permutacji $\sigma \in S_n$ to taka para $\sigma(i), \sigma(j)$, że

$$i < j, \qquad \sigma(i) > \sigma(j).$$

Definicja 4.11. Znak permutacji σ to

$$\varepsilon(\sigma) = (-1)^{(\text{liczba inwersji w } \sigma)}.$$

Funkcję ε nazywamy symbolem Leviego-Civity.

Jeśli $\varepsilon(\sigma) = 1$, to permutacja σ jest *parzysta*, a jeśli $\varepsilon(\sigma) = -1$, to jest *nieparzysta*.

Fakt 4.12. Każda transpozycja (zamiana miejscami) dwóch różnych elementów permutacji zmienia jej znak.

Dowód. Weźmy permutację wraz z poniższymi oznaczeniami:

$$\sigma = (\underbrace{\sigma_1, \sigma_2, \ldots, \sigma_{i-1}}_{A}, \sigma_i, \underbrace{\sigma_{i+1}, \ldots, \sigma_{j-1}}_{B}, \sigma_j, \underbrace{\sigma_{j+1}, \ldots, \sigma_{n-1}, \sigma_n}_{A}).$$

Zamieniając σ_i oraz σ_j, nie zmienia się liczba inwersji zawierających dowolny element $\sigma_k \in A$. Nie zmieni się również liczba inwersji zawierających dowolny element $\sigma_k \in B$, który jest większy lub mniejszy jednocześnie od σ_i i σ_j.

Dla pozostałych elementów $\sigma_k \in B$, jeśli istnieje inwersja (σ_i, σ_k), to istnieje również (σ_k, σ_j), a jeśli istnieje inwersja (σ_j, σ_k), to istnieje również (σ_k, σ_i). Tak więc jedyną inwersją, która zmienia parzystość ogólnej liczby inwersji — i tym samym $\varepsilon(\sigma)$ — jest inwersja (σ_i, σ_j), która istnieje przed transpozycją, albo po niej. □

Definicja 4.13. Wyznacznik macierzy kwadratowej A to taki element ciała, że

$$\det A = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Oznaczamy $\det \begin{bmatrix} \cdots \\ \cdots \\ \cdots \end{bmatrix} = \begin{vmatrix} \cdots \\ \cdots \\ \cdots \end{vmatrix}$.

Twierdzenie 4.14 (własności wyznaczników)
Dla macierzy kwadratowej $A = [a_{ij}] \in \mathcal{M}_{n\times n}(\mathbb{K})$ zachodzi:

1. $\det A = \det A^T$,
2. $\det I_n = 1$,
3. jeśli istnieje zerowy wiersz (lub kolumna) to $\det A = 0$,
4. jeśli pomnożymy jeden wiersz (lub kolumnę) przez skalar α, to wyznacznik również będzie α razy większy,
5. $\det \alpha A = (\det A)^\alpha$,
6. jeśli $A = [k_1, \ldots, k'_j + k''_j, \ldots, k_n]$, gdzie k_i są wierszami (lub kolumnami), to
$$\det A = \det[k_1, \ldots, k'_j, \ldots, k_n] + \det[k_1, \ldots, k''_j, \ldots, k_n],$$
7. przestawienie dwóch wierszy macierzy zmienia znak wyznacznika na przeciwny,
8. jeśli macierz ma dwa jednakowe wiersze (lub kolumny) to $\det A = 0$,
9. wyznacznik nie zmieni się, jeśli do wiersza (albo kolumny) dodamy kombinację liniową pozostałych wierszy (kolumn).

Dowód.

1. Wszystkich par elementów w permutacji $\sigma \in S_n$ jest $n(n-1)$. Jeśli (σ_i, σ_j) jest inwersją w σ, to w σ^{-1} nią nie jest, a skoro $2 \mid n(n-1)$, to $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$. Pozostałe czynniki sumowanych iloczynów zostaną takie same (jedynie w innej kolejności).
2. Dla permutacji identycznościowej dany w definicji iloczyn jest równy 1, dla każdej innej permutacji jest równy 0.
3. W każdym z sumowanych iloczynów występuje 0 jako czynnik.
4. W każdym z sumowanych iloczynów występuje jeden dodatkowy skalar α.
5. Wniosek z poprzedniego.
6. Dowód podobny do poprzednich dwóch.
7. Wynika z faktu 4.12.
8. Wniosek z poprzedniego, ponieważ $d = -d \Rightarrow d = 0$.
9. Stwierdzenie jest prawdziwe, jeśli wyznacznik nie zmieni się, gdy do jednego wiersza (lub kolumny) dodamy inny (uzasadnienie przez indukcję). Ten fakt można udowodnić, łącząc punkty 6 i 8 (jeden z wyznaczników będzie zerowy).

□

Twierdzenie 4.15

Wyznacznik macierzy 2×2 jest równy

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

Dowód. Prosty, z definicji. □

Twierdzenie 4.16 (reguła Sarrusa)

Wyznacznik macierzy 3×3 jest równy

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

Dowód. Prosty, z definicji. □

Regułę Sarrusa bardzo łatwo zapamiętać: wystarczy przepisać na koniec macierzy dwie pierwsze kolumny i liczyć podobnie jak macierz 2×2.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad \left[\begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}\right]$$

Twierdzenie 4.17 (Cauchy'ego)

Dla dowolnych macierzy $A, B \in \mathcal{M}_{n\times n}(\mathbb{K})$ zachodzi

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

Definicja 4.18. Minor stopnia k macierzy $A_{m\times n}$ to wyznacznik podmacierzy kwadratowej $k \times k$ powstałej przez wykreślenie $n - k$ kolumn oraz $m - k$ wierszy.

Jeśli $A_{n\times n}$ jest macierzą kwadratową, to wyznacznik macierzy powstałej przez wykreślenie i-tego wiersza i j-tej kolumny nazywamy *minorem odpowiadającym* elementowi a_{ij} macierzy A i oznaczamy M_{ij}.

Definicja 4.19. Dopełnienie algebraiczne elementu a_{ij} macierzy kwadratowej $A = [a_{ij}]_{m\times n}$ to skalar

$$A_{ij} = (-1)^{i+j} M_{ij}.$$

Twierdzenie 4.20 (rozwinięcie Laplace'a)

Niech $A = [a_{ij}]_{n\times n}$ będzie macierzą kwadratową. Wtedy dla każdego ustalonego $i \in \{1, 2 \ldots, n\}$

$$\det A = \sum_{j=1}^{n} a_{ij} A_{ij}$$

i dla każdego ustalonego $j \in \{1, 2 \ldots, n\}$

$$\det A = \sum_{i=1}^{n} a_{ij} A_{ij}.$$

Dowód. Dla wygody w oznaczeniach udowodnimy rozwinięcie wzdłuż pierwszego wiersza, $i = 1$. Weźmy taką permutację τ^j, że permutuje ona zbiór $\{1, \ldots, n\} \setminus \{j\}$. Oznaczmy

$$\sigma^* = (j, \tau_1^j, \tau_2^j, \ldots, \tau_{j-1}^j, \tau_{j+1}^j, \ldots, \tau_n^j).$$

Zauważmy, że $\varepsilon(\sigma^*) = (-1)^{j-1}\varepsilon(\tau^j)$. Teraz wystarczy ustalić j; z definicji wyznacznika (4.13) mamy

$$\begin{aligned}
\det A &= \sum_{\sigma\in S_n} \left(\varepsilon(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} \right) = \sum_{\sigma\in S_n} \left(a_{1\sigma(1)} \cdot \varepsilon(\sigma) \prod_{i=2}^{n} a_{i\sigma(i)} \right) = \\
&= \sum_{j=1}^{n} \left(a_{1j} \sum_{\sigma^*} \left(\varepsilon(\sigma^*) \prod_{i=2}^{n} a_{i\sigma^*(i)} \right) \right) = \\
&= \sum_{j=1}^{n} \left(a_{1j} \sum_{\tau^j} \left((-1)^{j-1} \varepsilon(\tau^j) \prod_{\substack{i=1 \\ i\neq j}}^{n} a_{i\tau^j(i)} \right) \right) = \\
&= \sum_{j=1}^{n} \left(a_{1j} \left((-1)^{j-1} M_{1j} \right) \right) = \\
&= \sum_{j=1}^{n} a_{1j} A_{1j}.
\end{aligned}$$

Wiemy, że $\det A = \det A^T$ (z 4.14), więc możemy rozwijać również wzdłuż kolumn. □

Wniosek 4.21

Wyznacznik macierzy trójkątnej jest równy iloczynowi elementów na jej przekątnej.

Dowód. Wystarczy rozwijać wyznacznik wzdłuż pierwszej kolumny. □

§4.3 Rząd macierzy

Definicja 4.22. Rząd macierzy to maksymalna liczba liniowo niezależnych wektorów kolumnowych tej macierzy. Rząd macierzy A oznaczamy przez $\operatorname{rank}(A)$.

Twierdzenie 4.23

Dla każdej macierzy $A \in \mathcal{M}_{m\times n}$ zachodzi

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

Dowód. Niech $r = \operatorname{rank} A$. Weźmy macierz $L \in \mathcal{M}_{m\times r}$, która zawiera r liniowo niezależnych wektorów kolumnowych macierzy A. Każda kolumna macierzy A jest oczywiście kombinacją liniową kolumn macierzy L, czyli $A = LZ$ dla pewnej macierzy współczynników $Z \in \mathcal{M}_{r\times n}$.

Z faktu 4.9 mamy $A^T = Z^T L^T$, więc $\operatorname{rank}(A^T)$ jest ograniczony z góry przez $\operatorname{rank}(Z^T)$, który jest ograniczony z góry przez r (liczbę kolumn macierzy Z^T), więc

$$\operatorname{rank}(A^T) \leq \operatorname{rank}(A).$$

Powtarzamy argument dla macierzy A^T i, wykorzystując fakt $(A^T)^T = A$, otrzymujemy

$$\operatorname{rank}(A) \leq \operatorname{rank}(A^T),$$

więc

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

□

Możemy teraz zmodyfikować definicję rzędu (4.22): jest to maksymalna liczba liniowo niezależnych wektorów kolumnowych lub wektorów wierszowych. Oczywiście zachodzi nierówność $\operatorname{rank}(A_{m\times n}) \leq n, m$.

Definicja 4.24. Macierz schodkowa to macierz, której pierwsze niezerowe elementy (schodki) kolejnych niezerowych wierszy znajdują się w coraz dalszych kolumnach, a wiersze zerowe umieszczone są najniżej.

Fakt 4.25. Rząd macierzy schodkowej jest równy liczbie jej schodków.

Definicja 4.26. Operacje elementarne na macierzach to:

- zamiana miejscami wierszy (kolumn) macierzy,
- dodanie do wiersza (kolumny) kombinacji liniowej pozostałych wierszy (kolumn),
- pomnożenie wiersza przez niezerowy skalar.

Jeśli macierz B można otrzymać z macierzy A za pomocą operacji elementarnych, to będziemy oznaczać $A \sim B$.

Fakt 4.27. Rząd macierzy nie zmienia się pod wpływem operacji elementarnych.

Dowód. Wynika z twierdzenia 3.33. □

Każdą macierz można łatwo doprowadzić do postaci schodkowej, za pomocą metody *eliminacji Gaussa*, która polega na stosowaniu operacji elementarnych na wierszach, „pozbywając się" niezerowych elementów z dolnego trójkąta. W ten sposób można odczytać jej rząd oraz, z pomocą wniosku 4.21, obliczyć jej wyznacznik (jeśli jest kwadratowa)[4].

[4]Warto jednak zwrócić uwagę, że przy obliczaniu wyznacznika lepiej nie mnożyć wierszy i nie zamieniać ich miejscami, bo te operacje wpływają na wyznacznik.

Przykład 4.28

Obliczyć wyznacznik macierzy

$$\begin{bmatrix} 1 & 3 & 1 \\ 1 & 1 & -1 \\ 3 & 11 & 6 \end{bmatrix}.$$

Rozwiązanie.

$$\begin{vmatrix} 1 & 3 & 1 \\ 1 & 1 & -1 \\ 3 & 11 & 6 \end{vmatrix} \overset{\substack{r_2-r_1 \\ r_3-3r_1}}{=} \begin{vmatrix} 1 & 3 & 1 \\ 0 & -2 & -2 \\ 0 & 2 & 3 \end{vmatrix} \overset{r_3+r_2}{=} \begin{vmatrix} 1 & 3 & 1 \\ 0 & -2 & -2 \\ 0 & 0 & 1 \end{vmatrix} = -2$$

□

Twierdzenie 4.29

Rząd macierzy A jest równy największemu ze stopni niezerowych minorów tej macierzy.

Dowód. Jeśli pewna podmacierz kwadratowa $k \times k$ jest nieosobliwa (czyli minor jest niezerowy), to jej kolumny są liniowo niezależne, czyli $\operatorname{rank} A \geq k$. Natomiast jeśli każda podmacierz wymiaru $(k+1) \times (k+1)$ jest osobliwa (czyli minory są zerowe), to żaden podzbiór $k+1$ wektorów kolumnowych nie jest liniowo niezależny, więc $\operatorname{rank} A < k+1$. Z tego wynika, że $\operatorname{rank} A = k$. □

Przykład 4.30

Obliczyć rząd macierzy

$$A = \begin{bmatrix} 1 & 3 & 4 \\ 3 & 4 & 1 \\ 1 & 2 & 7 \\ 3 & 5 & -1 \end{bmatrix}.$$

Rozwiązanie. Obliczmy minor

$$\begin{vmatrix} 1 & 3 & 4 \\ 3 & 4 & 1 \\ 3 & 5 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 4 \\ 3 & 4 & 1 \\ 0 & 1 & -2 \end{vmatrix} = (-8) + 0 + 12 - 0 - 1 - (-18) = 21 \neq 0.$$

Ten minor jest niezerowy i jednocześnie ma największy stopień (bo nie wykreśliliśmy żadnej kolumny), więc na mocy twierdzenia 4.29 $\operatorname{rank} A = 3$.

Dla pewności można pokazać również inną metodę — eliminację Gaussa:

$$\begin{bmatrix} 1 & 3 & 4 \\ 3 & 4 & 1 \\ 1 & 2 & 7 \\ 3 & 5 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 \\ 0 & -5 & -11 \\ 0 & -1 & 3 \\ 0 & -4 & -13 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 \\ 0 & 0 & -26 \\ 0 & -1 & 3 \\ 0 & 0 & -25 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 \\ 0 & -1 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\therefore \operatorname{rank} A = \operatorname{rank} \begin{bmatrix} 1 & 3 & 4 \\ 0 & -1 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = 3.$$

□

§4.4 Macierz odwrotna

Definicja 4.31. Macierz odwrotna A^{-1} do macierzy kwadratowej $A_{n\times n}$ to taka macierz, że

$$A \cdot A^{-1} = A^{-1} \cdot A = I_n.$$

Jeśli taka macierz istnieje, to mówimy, że A jest macierzą *odwracalną*.

Twierdzenie 4.32

Jeśli macierz $A_{n\times n}$ jest odwracalna, to

1. $\det A \neq 0$ oraz $\det(A^{-1}) = (\det A)^{-1}$,
2. $A^{-1} = \frac{1}{\det A}(A^D)^T$, gdzie A^D jest macierzą dopełnień algebraicznych A_{ij} macierzy A.

Dowód. 1. Z definicji macierzy odwrotnej (4.31)

$$A \cdot A^{-1} = I$$
$$\det(A \cdot A^{-1}) = \det I = 1.$$

Na mocy twierdzenia Cauchy'ego (4.17) otrzymujemy

$$\det A \cdot \det(A^{-1}) = 1,$$

z czego wynika teza.

2. Weźmy macierz B taką, że

$$B = \frac{1}{\det A}(A^D)^T,$$

to znaczy, że dla każdego i, j zachodzi

$$b_{ij} = \det A^{-1} \cdot A_{ji},$$

Obliczmy teraz macierz $C = AB$:

$$\begin{aligned} c_{ij} &= \sum_{k=1}^{n} a_{ik} b_{kj} \\ &= \sum_{k=1}^{n} a_{ik} \cdot \det A^{-1} \cdot A_{jk} \\ &= \det A^{-1} \sum_{k=1}^{n} a_{ik} \cdot A_{jk}. \end{aligned}$$

Jeśli $i = j$, to z rozwinięcia Laplace'a (4.20) wzdłuż i-tego rzędu otrzymujemy

$$c_{ij} = \det A^{-1} \cdot \det A = 1,$$

w przeciwnym wypadku rozwijamy macierz, która ma dwa takie same rzędy (i-ty oraz j-ty), więc jej wyznacznik jest zerowy, stąd

$$c_{ij} = 0.$$

Z tego wynika, że C jest macierzą jednostkową, więc $B = A^{-1}$.

□

Definicja 4.33. Macierz osobliwa to macierz A, której wyznacznik jest zerowy. W innym wypadku A jest macierzą nieosobliwą.

Na podstawie twierdzenia 4.32 łatwo zauważyć, że pojęcie macierzy nieosobliwej jest równoznaczne macierzy odwracalnej, a macierzy osobliwej — nieodwracalnej. Ponadto, jeśli macierz $A_{n\times n}$ jest nieosobliwa, to $\operatorname{rank} A = n$, a $A^{-1}, A^T, \alpha A, A^n$ również są macierzami nieosobliwymi.

Aby znaleźć macierz odwrotną, można oczywiście wykorzystać wzór

$$A^{-1} = \frac{1}{\det A}(A^D)^T$$

z twierdzenia 4.32, ale zwykle szybszą[5] metodą będzie eliminacja Gaussa, którą możemy wykorzystać wraz z poniższym faktem.

Fakt 4.34. Jeśli macierz kwadratowa A jest odwracalna, to

$$\left[A \mid I\right] \sim \left[I \mid B\right] \quad \Rightarrow \quad B = A^{-1}.$$

Dowód. Aby zrozumieć poniższy dowód trzeba się zapoznać z treścią sekcji Odwzorowania liniowe.

Łatwo udowodnić, że każda operacja elementarna może być rozumiana jako pewne odwzorowanie liniowe, a tym samym — jako operacja pomnożenia przez pewną macierz (odwzorowania liniowego). Drugi fakt, z którego należy sobie zdać sprawę, to równość

$$X\left[A \mid I\right] = \left[XA \mid XI\right],$$

która wynika bezpośrednio z definicji mnożenia macierzy. Z tych dwóch stwierdzeń wynika

$$\left[A \mid I\right] \sim \left[X_1A \mid X_1I\right] \sim \left[X_2X_1A \mid X_2X_1I\right] \sim \ldots \sim \left[A^{-1}A \mid A^{-1}I\right] = \left[I \mid A^{-1}\right],$$

gdzie A^{-1} jest niejako ciągiem odwzorowań liniowych X_i, które wykonujemy podczas eliminacji Gaussa. □

Przykład 4.35

Znaleźć macierz odwrotną do macierzy

$$A = \begin{bmatrix} 1 & 4 & 6 \\ 2 & 5 & 3 \\ 0 & 1 & 4 \end{bmatrix}.$$

Rozwiązanie.

$$\left[\begin{array}{ccc|ccc} 1 & 4 & 6 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 0 & 1 & 4 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 4 & 6 & 1 & 0 & 0 \\ 0 & -3 & -9 & -2 & 1 & 0 \\ 0 & 1 & 4 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 4 & 6 & 1 & 0 & 0 \\ 0 & 1 & 3 & \frac{2}{3} & \frac{-1}{3} & 0 \\ 0 & 1 & 4 & 0 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 4 & 6 & 1 & 0 & 0 \\ 0 & 1 & 3 & \frac{2}{3} & \frac{-1}{3} & 0 \\ 0 & 0 & 1 & \frac{-2}{3} & \frac{1}{3} & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & -6 & \frac{-5}{3} & \frac{4}{3} & 0 \\ 0 & 1 & 3 & \frac{2}{3} & \frac{-1}{3} & 0 \\ 0 & 0 & 1 & \frac{-2}{3} & \frac{1}{3} & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & -6 & \frac{-5}{3} & \frac{4}{3} & 0 \\ 0 & 1 & 0 & \frac{8}{3} & \frac{-4}{3} & -3 \\ 0 & 0 & 1 & \frac{-2}{3} & \frac{1}{3} & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{-17}{3} & \frac{10}{3} & 6 \\ 0 & 1 & 0 & \frac{8}{3} & \frac{-4}{3} & -3 \\ 0 & 0 & 1 & \frac{-2}{3} & \frac{1}{3} & 1 \end{array}\right],$$

[5] Na pewno w sensie złożoności obliczeniowej, w zadaniach to kwestia preferencji.

a więc

$$A^{-1} = \frac{1}{3}\begin{bmatrix} -17 & 10 & 18 \\ 8 & -4 & -9 \\ -2 & 1 & 3 \end{bmatrix}.$$

Możemy zweryfikować swoje obliczenia, znajdując macierz odwrotną metodą macierzy dopełnień algebraicznych.

$$\begin{aligned} A^{-1} &= \frac{1}{5\cdot 4 + 6\cdot 2 - 4\cdot 2\cdot 4 - 3}\begin{bmatrix} 5\cdot 4 - 3 & -(2\cdot 4) & 2 \\ -(4\cdot 4 - 6) & 4 & -(1) \\ 4\cdot 3 - 6\cdot 5 & -(3 - 6\cdot 2) & 5 - 4\cdot 2 \end{bmatrix}^T \\ &= \frac{1}{-3}\begin{bmatrix} 17 & -8 & 2 \\ -10 & 4 & -1 \\ -18 & 9 & -3 \end{bmatrix}^T = \frac{1}{3}\begin{bmatrix} -17 & 10 & 18 \\ 8 & 4 & 9 \\ -2 & 1 & 3 \end{bmatrix}. \end{aligned}$$

□

Macierz odwrotną można znaleźć również rozwiązując układ równań liniowych

$$A \cdot X = B,$$

wtedy $A^{-1} \cdot B = X$ — o czym więcej w następnej sekcji.

§5 Układy równań liniowych

Układ m równań liniowych z n niewiadomymi $x_1, \ldots, x_n$ w postaci

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = b_2 \\ \ldots & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = b_m \end{cases} \tag{5}$$

możemy reprezentować jako równanie macierzy. Macierz

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

nazywamy *macierzą główną* (macierzą współczynników) układu 5, a macierze

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \qquad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

nazywamy odpowiednio *kolumną wyrazów wolnych* oraz *kolumną niewiadomych*. Połączenie macierzy A i B

$$\left[A \mid B\right] = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right]$$

jest *macierzą uzupełnioną* tego układu. Wtedy układ 5 zapisujemy macierzowo jako

$$A \cdot X = B.$$

Definicja 5.1. Układ jednorodny to taki układ równań liniowych, że kolumna wyrazów wolnych jest zerowa.

Definicja 5.2. Układ jest:

- *oznaczony*, jeśli ma jedno rozwiązanie,
- *nieoznaczony*, jeśli ma więcej niż jedno rozwiązanie,
- *sprzeczny*, jeśli nie ma rozwiązań.

Definicja 5.3. Układ kwadratowy to układ równań liniowych, w którym liczba niewiadomych jest równa liczbie równań (czyli macierz główna jest kwadratowa).

Definicja 5.4. Układ Cramera to układ kwadratowy, w którym wyznacznik macierzy głównej jest niezerowy, $\det A \neq 0$.

Twierdzenie 5.5 (Cramera)
Jeśli dany układ jest układem Cramera, to jest oznaczony oraz

$$x_j = \frac{D_{x_j}}{\det A},$$

gdzie D_{x_j} jest wyznacznikiem macierzy powstałej przez zastąpienie j-tej kolumny macierzy głównej A kolumną wyrazów wolnych B.

Dowód. Pierwsze stwierdzenie jest prawdziwe, ponieważ jeśli $\det A \neq 0$, to kolumny tworzą bazę pewnej przestrzeni liniowej, do której należy kolumna B, więc na mocy twierdzenia 3.34 jest ona jednoznacznie wyznaczona przez kombinację liniową kolumn z A (a współczynniki tej kombinacji liniowej są właśnie kolumną niewiadomych X).

Oznaczając j-tą kolumną A jako $\mathbf{a_j}$ oraz $B = \mathbf{b}$, na mocy poprzedniego akapitu równanie

$$x_1\mathbf{a_1} + x_2\mathbf{a_2} + \ldots + x_n\mathbf{a_n} = \mathbf{b}$$

spełnia dokładnie jeden wektor $\mathbf{x}$. Zatem

$$D_{x_j} = \det(\mathbf{a_1}, \ldots, \mathbf{b}, \ldots, \mathbf{a_n}) = \det(\mathbf{a_1}, \ldots, \sum_{i=1}^{n} x_i\mathbf{a_i}, \ldots, \mathbf{a_n}).$$

Z własności wyznaczników (4.14) wynika, że wyznacznika nie zmieni odjęcie od pewnej kolumny innej kolumny przemnożonej przez skalar (nawet zerowy), więc

$$D_{x_j} = \det(\mathbf{a_1}, \ldots, x_j\mathbf{a_j}, \ldots, \mathbf{a_n}) = x_j \det(\mathbf{a_1}, \ldots, \mathbf{a_j}, \ldots, \mathbf{a_n}) = x_j \det A,$$

$$\therefore x_j = \frac{D_{x_j}}{\det A}.$$

□

Twierdzenie 5.6 (Kroneckera-Capellego)
Układ $AX = B$ ma co najmniej jedno rozwiązanie wtedy i tylko wtedy, gdy

$$\operatorname{rank}(A) = \operatorname{rank}(\left[A \mid B\right]).$$

Dowód. Oznaczając j-tą kolumną A jako $\mathbf{a_j}$ oraz $B = \mathbf{b}$, mamy

$$x_1\mathbf{a_1} + x_2\mathbf{a_2} + \ldots + x_n\mathbf{a_n} = \mathbf{b},$$

a więc X istnieje wtedy i tylko wtedy, gdy kolumna B jest kombinacją liniową kolumn z A (a więc nie jest liniowo niezależna, ergo $\operatorname{rank}(A) = \operatorname{rank}(\left[A \mid B\right])$). □

Twierdzenie 5.7

Układ $AX = B$ ma dokładnie jedno rozwiązanie wtedy i tylko wtedy, gdy

$$\operatorname{rank}(A) = \operatorname{rank}(\left[A \mid B\right]) = n,$$

gdzie n jest liczbą niewiadomych.

Dowód. Jak poprzednio, lecz z wykorzystaniem twierdzenia 3.34. □

Prosty wniosek z tego twierdzenia jest taki, że jeśli $\operatorname{rank}(A) = \operatorname{rank}(\left[A \mid B\right])$, ale $\operatorname{rank}(A) \neq n$, to układ jest nieoznaczony, a jego rozwiązania zależą od $n - \operatorname{rank}(A)$ parametrów[6].

Układy równań liniowych można łatwo rozwiązać eliminacją Gaussa w podobny sposób, jak robiliśmy to, szukając macierzy odwrotnej w przykładzie 4.35.

Przykład 5.8

Rozwiązać układ równań

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

Rozwiązanie.

$$\left[\begin{array}{ccc|c} 1 & 3 & -1 & 2 \\ 2 & 0 & -3 & -5 \\ 3 & 2 & -3 & -1 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 3 & -1 & 2 \\ 0 & -6 & -1 & -9 \\ 0 & -7 & 0 & -7 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 3 & -1 & 2 \\ 0 & 1 & -1 & -2 \\ 0 & 1 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 3 & -1 & 2 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Rząd macierzy jest równy liczbie zmiennych, więc (na mocy twierdzenia 5.7) układ jest oznaczony. Teraz możemy kontynuować przekształcenia, aby otrzymać macierz $\left[I \mid X\right]$, ale w praktyce łatwiej będzie teraz wrócić do układu równań. Mamy więc

$$z = 3,$$

$$y - z = -2 \Rightarrow y = -2 + 3 = 1,$$

$$x + 3x - z = 2 \Rightarrow x = 2 - 3 + 3 = 2.$$

□

§6 Geometria analityczna

W tej sekcji skupimy się na przestrzeni $\mathbb{R}^3(\mathbb{R})$, w której wektory będziemy interpretować często jako punkty lub wektory zaczepione w środku układu współrzędnych. Przez $\mathbb{R}^n$

[6] Jeśli układ jest określony nad ciałem $\mathbb{R}$ lub $\mathbb{C}$, to układ nieoznaczony ma nieskończenie wiele rozwiązań.

oznaczymy zbiór punktów, a przez $\overrightarrow{\mathbb{R}^n}$ zbiór wektorów. W przestrzeni $\mathbb{R}^3$ osie *prawoskrętnego*[7] układu współrzędnych (x, y, z) będą rozpięte przez *wersory* (wektory o jednostkowej długości):

$$\hat{\mathbf{i}} = (1, 0, 0), \quad \hat{\mathbf{j}} = (0, 1, 0), \quad \hat{\mathbf{k}} = (0, 0, 1).$$

Definicja 6.1. Metryka euklidesowa w $\mathbb{R}^n$ to funkcja $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, która dla punktów $P = (x_1, x_2, \ldots, x_n), Q = (y_1, y_2, \ldots, y_n)$ jest zdefiniowana jako

$$d(P, Q) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2}.$$

Wartość tej funkcji dla punktów X, Y to *odległość euklidesowa* tych punktów.

Definicja 6.2. Norma euklidesowa w $\mathbb{R}^n$ to funkcja $\|\cdot\| : \overrightarrow{\mathbb{R}^n} \to \mathbb{R}_{\geq 0}$, która dla wektora $v = [v_1, v_2, \ldots, v_n]$ jest zdefiniowana jako

$$\|v\| = \sqrt{\sum_{i=1}^{n} v_i^2}.$$

Wartość normy wektora v to *długość* tego wektora.

Łatwo zauważyć korelację między tymi dwoma wzorami: dla dwóch punktów P, Q, wektor $\overrightarrow{PQ}$ jest równy

$$\overrightarrow{PQ} = [y_1 - x_1, \ldots, y_n - x_n],$$

więc

$$d(P, Q) = \|\overrightarrow{PQ}\|.$$

Definicja 6.3. Iloczyn skalarny wektorów $u = [u_1, \ldots, u_n]$ i $v = [v_1, \ldots, v_n]$ w przestrzeni $\mathbb{R}^n$ to liczba

$$u \circ v = \sum_{i=1}^{n} u_i v_i.$$

Fakt 6.4. Jeśli U^T jest jednokolumnową macierzą powstałą z wektora u, a V to jednowierszową macierz powstała z wektora v, to

$$u \circ v = U^T \cdot V.$$

Fakt 6.5. Dla każdego wektora $v \in \mathbb{R}^n$ zachodzi

$$\sqrt{v \circ v} = \|v\|.$$

Jeśli dla przestrzeni wektorowej $\mathbb{R}^n$ określimy iloczyn skalarny wektorów, to taka przestrzeń jest *przestrzenią euklidesową*, którą oznaczamy przez E_n. Warto zauważyć, że taki iloczyn skalarny jest łączny, przemienny, zgodny z mnożeniem przez skalar oraz rozdzielny względem dodawania.

[7] To znaczy zgodnego z regułą prawej ręki — wnętrze obracającej się dłoni zakreśla łuk od osi OX do OY, przy czym kciuk ma zwrot zgodny z osią OZ.

Twierdzenie 6.6 (Cauchy'ego-Schwarza)
Dla dowolnych wektorów $\mathbf{u}, \mathbf{v} \in E_n$ zachodzi nierówność

$$|\mathbf{u} \circ \mathbf{v}| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|,$$

przy czym równość zachodzi wtedy i tylko wtedy, gdy wektory są liniowo zależne.

Dowód. Twierdzenie jest trywialne, jeśli któryś z wektorów jest zerowy, dlatego przyjmijmy $\mathbf{u}, \mathbf{v} \neq \overline{0}$. Dla dowolnego $\alpha \in \mathbb{R}$ mamy

$$0 \leq \|\mathbf{u} - \alpha\mathbf{v}\|^2 = (\mathbf{u} - \alpha\mathbf{v}) \circ (\mathbf{u} - \alpha\mathbf{v}) = \mathbf{u} \circ \mathbf{u} - 2\alpha(\mathbf{u} \circ \mathbf{v}) + \alpha^2(\mathbf{v} \circ \mathbf{v}).$$

Podstawiając $\alpha = (\mathbf{u} \circ \mathbf{v})(\mathbf{v} \circ \mathbf{v})^{-1}$ otrzymamy

$$\begin{aligned}
0 &\leq (\mathbf{u} \circ \mathbf{u}) - (\mathbf{v} \circ \mathbf{v})^{-1}(\mathbf{u} \circ \mathbf{v})^2 \\
(\mathbf{v} \circ \mathbf{v})^{-1}(\mathbf{u} \circ \mathbf{v})^2 &\leq (\mathbf{u} \circ \mathbf{u}) \\
(\mathbf{u} \circ \mathbf{v})^2 &\leq (\mathbf{u} \circ \mathbf{u})(\mathbf{v} \circ \mathbf{v}) \\
(\mathbf{u} \circ \mathbf{v})^2 &\leq \|\mathbf{u}\|^2 \cdot \|\mathbf{v}\|^2 \\
|\mathbf{u} \circ \mathbf{v}| &\leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|.
\end{aligned}$$

Równość zachodzi tylko w przypadku, gdy $\alpha = 0$, czyli gdy $\mathbf{u}, \mathbf{v}$ są liniowo zależne. □

Wniosek 6.7 (nierówność trójkąta)
Dla dowolnych wektorów $\mathbf{u}, \mathbf{v} \in E_n$ zachodzi nierówność

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|,$$

Dowód. Z nierówności Cauchy'ego-Schwarza wynika, że

$$\begin{aligned}
\mathbf{u} \circ \mathbf{v} &\leq \sqrt{\mathbf{u} \circ \mathbf{u}} \cdot \sqrt{\mathbf{v} \circ \mathbf{v}} \\
\mathbf{u} \circ \mathbf{u} + 2 \cdot \mathbf{u} \circ \mathbf{v} + \mathbf{v} \circ \mathbf{v} &\leq \mathbf{u} \circ \mathbf{u} + 2 \cdot \sqrt{\mathbf{u} \circ \mathbf{u}} \cdot \sqrt{\mathbf{v} \circ \mathbf{v}} + \mathbf{v} \circ \mathbf{v} \\
(\mathbf{u} + \mathbf{v}) \circ (\mathbf{u} + \mathbf{v}) &\leq (\sqrt{\mathbf{u} \circ \mathbf{u}} + \sqrt{\mathbf{v} \circ \mathbf{v}})^2 \\
\|\mathbf{u} + \mathbf{v}\|^2 &\leq (\|\mathbf{u}\| + \|\mathbf{v}\|)^2 \\
\|\mathbf{u} + \mathbf{v}\| &\leq \|\mathbf{u}\| + \|\mathbf{v}\|.
\end{aligned}$$

□

Definicja 6.8. Kąt między niezerowymi wektorami $\mathbf{u}, \mathbf{v} \in E_n$ to taka liczba $\sphericalangle(\mathbf{u}, \mathbf{v}) = \varphi \in [0, \pi]$, że

$$\cos\varphi = \frac{\mathbf{u} \circ \mathbf{v}}{\|\mathbf{u}\| \cdot \|\mathbf{v}\|}.$$

Jeśli $\sphericalangle(\mathbf{u}, \mathbf{v}) = \frac{\pi}{2}$, to wektory są *prostopadłe* $\mathbf{u} \perp \mathbf{v}$, a jeśli $\sphericalangle(\mathbf{u}, \mathbf{v}) = 0$ lub π, to są *równoległe* $\mathbf{u} \parallel \mathbf{v}$. Przyjmujemy, że wektor zerowy jest prostopadły i równoległy do wszystkich innych wektorów.

Fakt 6.9. Dla dowolnych wektorów $\mathbf{u}, \mathbf{v} \in E_n$

$$\mathbf{u} \perp \mathbf{v} \quad \Leftrightarrow \quad \mathbf{u} \circ \mathbf{v} = 0$$

Dowód. Wynika z definicji. □

Oczywiście $\mathbf{u} \parallel \mathbf{v}$ wtedy i tylko wtedy, gdy wektory $\mathbf{u}, \mathbf{v}$ są liniowo zależne.

§6.1 Przestrzeń trójwymiarowa

Fakt 6.10. Trójka liniowo niezależnych wektorów $\mathbf{u}, \mathbf{v}, \mathbf{w} \in E_3$ tworzy układ prawoskrętny, jeśli

$$\begin{vmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 \end{vmatrix} > 0.$$

Uzasadnienie. Jest to intuicyjnie prawdziwe — jeśli zamienimy wiersze ze sobą, to układ wektorów stanie się lewoskrętny i jednocześnie (z własności wyznaczników 4.14) wyznacznik macierzy będzie przeciwny. Podobnie, jeśli zmienimy zwrot danego wektora, to układ stanie się lewoskrętny, a wyznacznik macierzy będzie przeciwny.

Dociekliwy Czytelnik może zaznajomić się z dowodem zawartym w artykule „*A Simple Proof of the Right-Hand Rule*" autorstwa Fuchang Gao. □

Definicja 6.11. Iloczyn wektorowy to takie działanie $\times : (\overrightarrow{E_3})^2 \to \overrightarrow{E_3}$, że:

1. jeśli $\mathbf{u} \parallel \mathbf{v}$, to $\mathbf{u} \times \mathbf{v} = \overline{0}$,
2. w przeciwnym wypadku $\mathbf{u} \times \mathbf{v} = \mathbf{w}$, gdzie
   - $\|\mathbf{w}\| = \|\mathbf{u}\| \cdot \|\mathbf{v}\| \cdot \sin \sphericalangle(\mathbf{u}, \mathbf{v})$,
   - $\mathbf{w} \perp \mathbf{u}$ oraz $\mathbf{w} \perp \mathbf{v}$,
   - wektory $\mathbf{u}, \mathbf{v}, \mathbf{w}$ tworzą układ prawoskrętny.

Twierdzenie 6.12

Dla dowolnych wektorów $\mathbf{u}, \mathbf{v} \in E_3$

$$\mathbf{u} \times \mathbf{v} = \left[\begin{vmatrix} \mathbf{u}_2 & \mathbf{u}_3 \\ \mathbf{v}_2 & \mathbf{v}_3 \end{vmatrix}, \begin{vmatrix} \mathbf{u}_3 & \mathbf{u}_1 \\ \mathbf{v}_3 & \mathbf{v}_1 \end{vmatrix}, \begin{vmatrix} \mathbf{u}_1 & \mathbf{u}_2 \\ \mathbf{v}_1 & \mathbf{v}_2 \end{vmatrix} \right]$$

Dowód. Żmudny, ale prosty; z definicji. □

W praktyce łatwiej stosować (zapamiętać) „wzór"

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \widehat{\mathbf{i}} & \widehat{\mathbf{j}} & \widehat{\mathbf{k}} \\ \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{vmatrix} = \widehat{\mathbf{i}} \begin{vmatrix} \mathbf{u}_2 & \mathbf{u}_3 \\ \mathbf{v}_2 & \mathbf{v}_3 \end{vmatrix} + \widehat{\mathbf{j}} \begin{vmatrix} \mathbf{u}_3 & \mathbf{u}_1 \\ \mathbf{v}_3 & \mathbf{v}_1 \end{vmatrix} + \widehat{\mathbf{k}} \begin{vmatrix} \mathbf{u}_1 & \mathbf{u}_2 \\ \mathbf{v}_1 & \mathbf{v}_2 \end{vmatrix} \tag{6}$$

Warto zauważyć, że iloczyn wektorowy jest antyprzemienny ($\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$), zgodny z mnożeniem przez skalar oraz rozdzielny względem dodawania.

Fakt 6.13. Dla dowolnych wektorów $\mathbf{u}, \mathbf{v} \in E_3$

$$\mathbf{u} \parallel \mathbf{v} \quad \Leftrightarrow \quad \mathbf{u} \times \mathbf{v} = \overline{0}$$

Dowód. Wynika z definicji. □

Twierdzenie 6.14

Dla dowolnych wektorów $\mathbf{u}, \mathbf{v} \in E_3$ liczba $\|\mathbf{u} \times \mathbf{v}\|$ jest (skierowanym) polem równoległoboku rozpiętego przez wektory $\mathbf{u}, \mathbf{v}$.

Dowód. Z definicji iloczynu wektorowego (6.11) mamy

$$\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \cdot \|\mathbf{v}\| \cdot \sin \sphericalangle(\mathbf{u}, \mathbf{v}),$$

czyli iloczyn długości obu boków oraz sinusa kąta między nimi, który istotnie jest równy polu równoległoboku. □

Prosty wniosek z tego twierdzenia jest taki, że pole trójkąta rozpiętego przez wektory $\mathbf{u}, \mathbf{v}$ jest równe $\frac{1}{2}\|\mathbf{u} \times \mathbf{v}\|$.

Działanie $(\mathbf{u} \times \mathbf{v}) \circ \mathbf{w}$ nazywamy *iloczynem mieszanym*.

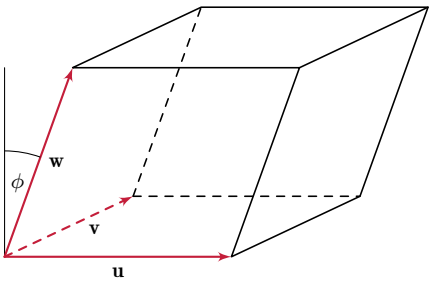
Twierdzenie 6.15

Dla dowolnych wektorów $\mathbf{u}, \mathbf{v}, \mathbf{w} \in E_3$ liczba $((\mathbf{u} \times \mathbf{v}) \circ \mathbf{w})$ jest (skierowaną) objętością równoległościanu rozpiętego przez wektory $\mathbf{u}, \mathbf{v}, \mathbf{w}$.

Dowód. Z definicji kąta między wektorami (6.8) mamy

$$(\mathbf{u} \times \mathbf{v}) \circ \mathbf{w} = \|\mathbf{u} \times \mathbf{v}\| \cdot \|\mathbf{w}\| \cdot \cos \phi,$$

gdzie $\|\mathbf{u} \times \mathbf{v}\|$ to pole równoległoboku rozpiętego przez wektory $\mathbf{u}, \mathbf{v}$, jak na rysunku poniżej.



□

Prosty wniosek z tego twierdzenia jest taki, że objętość czworościanu rozpiętego przez wektory $\mathbf{u}, \mathbf{v}, \mathbf{w}$ jest równa $\frac{1}{6}|(\mathbf{u} \times \mathbf{v}) \circ \mathbf{w}|$.

Fakt 6.16. Dla dowolnych wektorów $\mathbf{u}, \mathbf{v}, \mathbf{w} \in E_3$

$$(\mathbf{u} \times \mathbf{v}) \circ \mathbf{w} = \begin{vmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 \end{vmatrix}.$$

Jest to prostszy sposób na liczenie objętości równoległościanu.

Dowód. Łatwo zauważyć zależność między rozwinięciem Laplace'a (4.20) oraz wzorem 6. □

§6.1.1 Równanie płaszczyzny w przestrzeni

Płaszczyznę jednoznacznie wyznaczają trzy niewspółliniowe punkty (lub wyznaczone przez nie dwa wektory). Płaszczyznę jednoznacznie wyznaczyć może również jeden niezerowy wektor, zwany *wektorem normalnym*; jest on prostopadły do wyznaczanej płaszczyzny.

Równanie parametryczne płaszczyzny Jeśli $P_0 = (x_0, y_0, z_0) \in \pi$ oraz $\mathbf{u} = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]$, $\mathbf{v} = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3] \in E_3$ są liniowo niezależne i równoległe do płaszczyzny π, to

$$\pi : \begin{cases} x = x_0 + s\mathbf{u}_1 + t\mathbf{v}_1 \\ y = y_0 + s\mathbf{u}_2 + t\mathbf{v}_2 \\ z = z_0 + s\mathbf{u}_3 + t\mathbf{v}_3 \end{cases} \qquad s, t \in \mathbb{R} \tag{7}$$

jest równaniem parametrycznym płaszczyzny. Wtedy każdy punkt płaszczyzny jest po prostu punktem P_0, który został przesunięty o pewien wektor równoległy do płaszczyzny.

Równanie normalne płaszczyzny Jeśli $n = [A, B, C]$ jest wektorem normalnym płaszczyzny π oraz $P_0 = (x_0, y_0, z_0) \in \pi$, to

$$\pi : [x - x_0, y - y_0, z - z_0] \circ [A, B, C] = 0$$

czyli

$$\pi : A(x - x_0) + B(y - y_0) + C(z - z_0) = 0 \tag{8}$$

nazywamy równaniem normalnym płaszczyzny π. Wtedy każdy punkt $P_1 \in \pi$ jest taki, że wektor $\overrightarrow{P_0P_1}$ jest prostopadły do wektora normalnego, czyli równoległy do płaszczyzny.

Równanie ogólne płaszczyzny Równanie normalne można wymnożyć do równania ogólnego

$$\pi : Ax + By + Cz + D = 0. \tag{9}$$

Równanie odcinkowe płaszczyzny Jeśli $a, b, c \in \mathbb{R}$ są niezerowe, to

$$\pi : \frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1 \tag{10}$$

jest równaniem odcinkowym płaszczyzny. Taka płaszczyzna przecina się z osiami układu współrzędnych w punktach $(a, 0, 0), (0, b, 0), (0, 0, c)$; tak więc nie każda płaszczyzna ma równianie odcinkowe.

§6.1.2 Równanie prostej w przestrzeni

Prostą jednoznacznie wyznaczają dwa punkty (lub jeden wyznaczony przez nie wektor). Prosta jest również jednoznacznie wyznaczona przez przecięcie dwóch nierównoległych płaszczyzn.

Równanie parametryczne prostej Jeśli $P_0 = (x_0, y_0, z_0) \in l$ oraz $\mathbf{v} = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3] \in E_3$ jest niezerowym wektorem równoległym do prostej l, to

$$l : \begin{cases} x = x_0 + t\mathbf{v}_1 \\ y = y_0 + t\mathbf{v}_2 \\ z = z_0 + t\mathbf{v}_3 \end{cases} \qquad t \in \mathbb{R} \tag{11}$$

jest równaniem parametrycznym prostej. Wektor $\mathbf{v}$ nazywamy wektorem *kierunkowym* (lub tworzącym, rozpinającym) prostej l.

Równanie kierunkowe prostej Jeśli $P_0 = (x_0, y_0, z_0) \in l$ oraz $\mathbf{v} = [a, b, c] \in E_3$ jest równoległy do prostej l i $a, b, c \in \mathbb{R}$ są niezerowe, to

$$l : \frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c} \tag{12}$$

jest równaniem kierunkowym prostej.

Równanie krawędziowe prostej Niech

$$\pi_1 : A_1x + B_1y + C_1z + D_1 = 0, \pi_2 : A_2x + B_2y + C_2z + D_2 = 0.$$

Jeśli π_1 nie jest równoległa z π_2, to równanie krawędziowe prostej ma postać

$$l : \begin{cases} A_1x + B_1y + C_1z + D_1 = 0 \\ A_2x + B_2y + C_2z + D_2 = 0 \end{cases}. \tag{13}$$

Uwaga. Jeśli chcemy łatwo przejść z równania krawędziowego do parametrycznego, to wystarczy zauważyć, że $n_1 = [A_1, B_1, C_1], n_2 = [A_2, B_2, C_2]$ są wektorami normalnymi płaszczyzn. Chcemy znaleźć więc wektor (kierunkowy), który leży na obu tych płaszczyznach, a więc jest prostopadły do obu wektorów normalnych. Tę własność ma wektor $\mathbf{v} = n_1 \times n_2$.

§6.2 Odległości

Zdefiniowaliśmy już odległość między dwoma punktami w definicji 6.1. Bez formalnego wyprowadzenia będziemy używać pojęcia odległości również w kontekście odległości między punktem a płaszczyzną, punktem a prostą, prostą a płaszczyzną czy między płaszczyznami lub prostymi. Taka odległość będzie najmniejszą odległością między pewnym punktem jednej figury oraz pewnym punktem drugiej figury.

$$d(\Phi, \Psi) = \min_{A \in \Phi, B \in \Psi} d(A, B)$$

Z twierdzenia Pitagorasa wynika, że wektor między tymi dwoma punktami będzie prostopadły do powierzchni obu danych figur.

Twierdzenie 6.17 (odległość punktu od płaszczyzny)
Odległość punktu $Q = (x_1, y_1, z_1)$ od płaszczyzny $\pi : Ax + Bx + Cx + D = 0$ jest równa

$$d(Q, \pi) = \frac{|Ax_1 + By_1 + Cz_1 + D|}{\sqrt{A^2 + B^2 + C^2}}.$$

Dowód. Niech prosta l będzie prostopadła do płaszczyzny π oraz niech $Q \in l$. Taka prosta jest równoległa do wektora normalnego $n = [A, B, C]$, więc

$$l : \begin{cases} x = x_1 + At \\ y = y_1 + Bt \\ z = z_1 + Ct \end{cases} \qquad t \in \mathbb{R}.$$

Niech $Q' = \pi \cap l$ będzie punktem przecięcia prostej l i płaszczyzny π. Podstawiamy równanie prostej do równania płaszczyzny:

$$Q' \in l : A(x_1 + tA) + B(y_1 + tB) + C(z_1 + tC) + D = 0$$
$$Q' \in l : t = \frac{Ax_1 + By_1 + Cz_1 + D}{-(A^2 + B^2 + C^2)}$$

Mamy więc

$$d(Q,Q') = \sqrt{(tA)^2+(tB)^2+(tC)^2} = |t|\cdot\sqrt{A^2+B^2+C^2} = \frac{|Ax_1+By_1+Cz_1+D|}{\sqrt{A^2+B^2+C^2}}.$$

□

Jeśli szukamy odległości między dwoma płaszczyznami, to wystarczy sprawdzić, czy są one równoległe (to znaczy, czy ich wektory normalne są równoległe). Jeśli nie, to odległość jest oczywiście zerowa; w przeciwnym wypadku wystarczy wziąć dowolny punkt z jednej płaszczyzny i znaleźć odległość między tym punktem a drugą płaszczyzną.

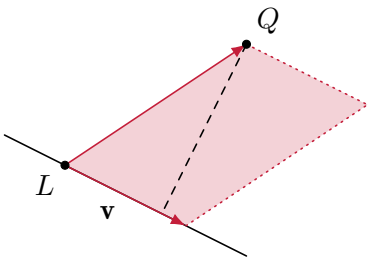
Twierdzenie 6.18 (odległość punktu od prostej)
Odległość punktu Q od prostej l jest równa

$$d(Q,l) = \frac{\|\mathbf{v}\times\overrightarrow{LQ}\|}{\|\mathbf{v}\|},$$

gdzie $\mathbf{v}$ jest wektorem kierunkowym prostej l, a $L\in l$.

Dowód. Mamy $l:(x,y,z) = L+t\mathbf{v}$. Na podstawie twierdzenia 6.14, pole równoległoboku rozpiętego przez wektory $\mathbf{v}, \overrightarrow{LQ}$ jest równe

$$\|\mathbf{v}\times\overrightarrow{LQ}\|.$$



Ze standardowego wzoru na pole równoległoboku jest ono również równe

$$\|\mathbf{v}\|\cdot d(Q,l),$$

z czego wynika teza. □

Jeśli dwie proste są równoległe, to odległość między nimi można obliczyć przez wzięcie dowolnego punktu z jednej prostej i obliczenie jego odległości od drugiej prostej. Jeśli proste się przecinają, to odległość między nimi jest zerowa. Sytuacja się komplikuje, jeśli dane dwie proste nie są równoległe lub przecinające się (to znaczy są *skośne*).

Twierdzenie 6.19 (odległość prostych skośnych)
Niech l_1, l_2 są prostymi skośnymi. Odległość między nimi jest równa

$$d(l_1,l_2) = \frac{|(\mathbf{v}\times\mathbf{u})\circ\overrightarrow{L_1L_2}|}{\|\mathbf{v}\times\mathbf{u}\|},$$

gdzie $\mathbf{v},\mathbf{u}$ są odpowiednio wektorami kierunkowym prostych l_1, l_2, a $L_1\in l_1$ i $L_2\in l_2$.

Dowód. Mamy $l_1 : (x, y, z) = L_1 + t\mathbf{v}$ oraz $l_2 : (x, y, z) = L_2 + t\mathbf{u}$. Na podstawie twierdzenia 6.15 objętość równoległościanu rozpiętego przez wektory $\mathbf{v}, \mathbf{u}, \overrightarrow{L_1L_2}$ jest równa

$$|(\mathbf{v} \times \mathbf{u}) \circ \overrightarrow{L_1L_2}|.$$

Ze standardowego wzoru na objętość (tzn. pole podstawy · wysokość) jest ona również równa

$$\|\mathbf{v} \times \mathbf{u}\| \cdot d(l_1, l_2),$$

z czego wynika teza. □

§6.3 Przykłady

Przykład 6.20 (współliniowość punktów)
Sprawdź, czy punkty $A = (1, 0, 2), B = (3, 1, -1), C = (-1, -1, 5)$ są współliniowe.

Rozwiązanie. Punkty A, B, C są współliniowe wtedy i tylko wtedy, gdy wektory $\overrightarrow{AB}, \overrightarrow{AC}$ są współliniowe, czyli rozpinają równoległobok o zerowym polu. Na podstawie twierdzenia 6.14 wystarczy obliczyć

$$\left\|\overrightarrow{AB} \times \overrightarrow{AC}\right\| = \|[2, 1, -3] \times [-2, -1, 3]\| = \|[3 - 3, 6 - 6, -2 + 2]\| = 0,$$

z czego wynika, że punkty A, B, C są współliniowe. □

Przykład 6.21 (współpłaszczyznowość punktów)
Sprawdź, czy punkty $A = (0, 2, 2), B = (2, 1, 0), C = (3, -1, 2), D = (1, -2, 3)$ są współpłaszczyznowe.

Rozwiązanie. Punkty A, B, C, D są współpłaszczyznowe wtedy i tylko wtedy, gdy wektory $\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}$ rozpinają równoległościan o zerowej objętości. Na podstawie twierdzenia 6.15 i faktu 6.16 wystarczy obliczyć

$$\begin{vmatrix} \overrightarrow{AB} \\ \overrightarrow{AC} \\ \overrightarrow{AD} \end{vmatrix} = \begin{vmatrix} 2 & -1 & -2 \\ 3 & -3 & 0 \\ 1 & -4 & 1 \end{vmatrix} = -6 + 24 - 6 + 3 = 15 \neq 0,$$

z czego wynika, że punkty A, B, C, D nie są współpłaszczyznowe. □

Przykład 6.22 (wzajemne położenie prostych, odległość)
Zbadaj wzajemne położenie prostych

$$l_1 : \begin{cases} x = 1 - 3t \\ y = 2 - 6t \\ z = 3 - 5t \end{cases} \quad t \in \mathbb{R}, \qquad l_2 : \begin{cases} x = 2 + t \\ y = -1 + 2t \\ z = 4 + 2t \end{cases} \quad t \in \mathbb{R}$$

oraz oblicz odległość między nimi.

Rozwiązanie. Zaczniemy od sprawdzenia, czy proste są równoległe. Weźmy wektory kierunkowe tych prostych i obliczmy ich iloczyn wektorowy. Mamy

$$[-3, -6, -5] \times [1, 2, 2] = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ -3 & -6 & -5 \\ 1 & 2 & 2 \end{vmatrix} = [-12 + 10, -5 + 6, -6 + 6] = [-2, 1, 0] \neq \overline{0},$$

więc proste nie są równoległe. Odległość między nimi, zgodnie z twierdzeniem 6.19, jest równa

$$d(l_1, l_2) = \frac{|[-2,1,0] \circ [2-1, -1-2, 4-3]|}{\|[-2,1,0]\|} = \frac{|-2-3|}{\sqrt{5}} = \sqrt{5}.$$

Z tego wynika również, że proste są skośne; nie są przecinające się, bo $d(l_1, l_2) \neq 0$. □

Przykład 6.23 (wzajemne położenie prostych, wspólna płaszczyzna)
Zbadaj wzajemne położenie prostych

$$l_1 : \begin{cases} x = t \\ y = -2t \\ z = 3t \end{cases} \quad t \in \mathbb{R}, \qquad l_2 : \begin{cases} x = -1 + t \\ y = 2 - t \\ z = -3 + 4t \end{cases} \quad t \in \mathbb{R}$$

oraz wyznacz ich wspólną płaszczyznę (jeśli istnieje).

Rozwiązanie. Obliczmy najpierw iloczyn wektory wektorów kierunkowych danych prostych.

$$[1,-2,3] \times [1,-1,4] = \begin{vmatrix} \widehat{\mathbf{i}} & \widehat{\mathbf{j}} & \widehat{\mathbf{k}} \\ 1 & -2 & 3 \\ 1 & -1 & 4 \end{vmatrix} = [-8+3, 3-4, -1+2] = [-5, -1, 1]$$

Widzimy więc, że proste nie są równoległe. Zamiast liczyć odległość między nimi (i tak stwierdzić, czy się przecinają), możemy spróbować znaleźć punkt przecięcia.

$$\begin{cases} t = -1 + s \\ -2t = 2 - s \\ 3t = -3 + 4s \end{cases} \Rightarrow s = 0, t = -1.$$

Z tego wynika, że punkt przecięcia istnieje, jest nim $P = (-1, 2, -3)$, więc proste nie są skośne, a więc tworzą płaszczyznę.

Do tej płaszczyzny należą wektory kierunkowe prostych l_1, l_2, więc jej wektorem normalnym będzie ich iloczyn wektorowy. Podstawiając do równania normalnego płaszczyzny mamy

$$\begin{aligned} \pi : &-5(x-(-1)) - 1(y-2) + 1(z-(-3)) = 0 \\ &-5x - 5 - y + 2 + z + 3 = 0 \\ &-5x - y + z = 0 \end{aligned}$$

co jest równaniem ogólnym płaszczyzny π. □

Uwaga. Jeśli chcemy znaleźć rzut prostokątny punktu P na prostą l, to najłatwiej będzie znaleźć płaszczyznę π zawierającą punkt P i prostopadłą do prostej l (wektor normalny szukanej płaszczyzny będzie wektorem kierunkowym prostej l). Następnie wystarczy znaleźć punkt przecięcia $\pi \cap l$.

Podobnie, jeśli chcemy znaleźć rzut prostokątny punktu P na płaszczyznę π, to należy znaleźć prostą l taką, że $l \perp \pi$ oraz $P \in l$ (jak poprzednio, wektor kierunkowy szukanej prostej będzie wektorem normalnym płaszczyzny π).

Zastosowanie powyższej uwagi niech będzie ćwiczeniem dla Czytelnika[8].

[8]Który to Czytelnik z pewnością zauważył już, że wykorzystaliśmy ją w dowodzie twierdzenia 6.17.

§7 Odwzorowania liniowe

Definicja 7.1. Odwzorowanie

$$f : V \to W,$$

gdzie V, W są przestrzeniami liniowymi nad tym samym ciałem $\mathbb{K}$, jest *liniowe*, jeśli

$$\bigvee_{u,v \in V} f(u+v) = f(u) + f(v)$$

oraz

$$\bigvee_{v \in V, \alpha \in \mathbb{K}} f(\alpha v) = \alpha f(v),$$

to znaczy, kiedy jest *addytywne* oraz *jednorodne*.

Podobnie jak w przypadku homomorfizmu grup (definicja 3.19), elementy przeciwne oraz neutralne są zachowywane. Analogicznie do równoważnej charakterystyki podprzestrzeni (fakt 3.29) warunki z powyższej definicji są równoważne warunkowi

$$\bigvee_{\alpha,\beta \in \mathbb{K}} \bigvee_{u,v \in V} f(\alpha u + \beta v) = \alpha f(u) + \beta f(v). \tag{14}$$

Wniosek 7.2 (z równania 14)

Odwzorowanie liniowe f jest jednoznacznie określone przez wartości f na wektorach bazowych dziedziny.

Dowód. Wzór 14 można rozszerzyć do większej liczby składników, używając $(\alpha, \beta, \gamma, \ldots)$ oraz $(u, v, w, \ldots)$. Taka postać również będzie równoważna definicji 7.1. Możemy policzyć wartość f dla każdego wektora, znając jego współrzędne w pewnej bazie B oraz wartości f dla wszystkich wektorów z tej bazy. □

Definicja 7.3. Jądro odwzorowania liniowego $f : V \to W$ to zbiór

$$\operatorname{Ker} f = f^{-1}(\{\overline{0}\}) = \{v \in V \mid f(v) = \overline{0}\}.$$

Definicja 7.4. Obraz odwzorowania liniowego $f : V \to W$ to zbiór

$$\operatorname{Im} f = f(V) = \{w \in W \mid \exists v \in V : w = f(v)\}.$$

Jeśli B jest bazą przestrzeni V, to

$$\operatorname{Im} f = \operatorname{Lin} f(B).$$

Fakt 7.5. Dla każdego odwzorowania liniowego $f : V \to W$, jądro f jest podprzestrzenią V, a obraz f jest podprzestrzenią W.

Wymiar jądra pewnego odwzorowania f nazywamy *zerowością* i oznaczamy $\operatorname{null} f$, a wymiar jego obrazu nazywamy *rzędem* i oznaczamy $\operatorname{rank} f$.

Przykład 7.6

Weźmy odwzorowanie $f : \mathbb{R}^3 \to \mathbb{R}^2$ takie, że

$$f(x, y, z) = (x, y).$$

Znajdźmy jądro i obraz tego odwzorowania.

$$\operatorname{Ker} f = \{(0, 0, z) : z \in \mathbb{R}\}, \quad \dim = 1.$$

$$\operatorname{Im} f = \mathbb{R}^2, \quad \dim = 2.$$

Teraz weźmy $g : \mathbb{R}^2 \to \mathbb{R}^3$ takie, że

$$g(x, y) = (x, y, x + y).$$

Znajdźmy jądro i obraz tego odwzorowania.

$$\operatorname{Ker} g = \{\overline{0}\}, \quad \dim = 0.$$

$$\operatorname{Im} g = \{(x, y, x + y) : x, y \in \mathbb{R}\}, \quad \dim = 2.$$

Twierdzenie 7.7 (o rzędzie)

Jeśli V, W są skończenie wymiarowymi przestrzeniami wektorowymi nad ciałem $\mathbb{K}$ oraz $f : V \to W$ jest odwzorowaniem liniowym, to

$$\operatorname{null} f + \operatorname{rank} f = \dim V,$$

$$\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim V.$$

Dowód. Niech $n = \dim V$ oraz $k = \dim \operatorname{Ker} f$. Skoro $\operatorname{Ker} f$ jest podprzestrzenią przestrzeni V, to jeśli $n = k$, to dla każdego $\mathbf{v} \in V$ zachodzi $f(\mathbf{v}) = \overline{0}$, więc $\operatorname{Im} f = \{\overline{0}\}$, ergo teza jest spełniona. Dalej załóżmy więc, że $n > k$. Istnieje taka baza przestrzeni V, która ma postać

$$\{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\},$$

gdzie $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ jest bazą $\operatorname{Ker} f$. Weźmy dowolny wektor $\mathbf{v} \in V$,

$$\mathbf{v} = t_1\mathbf{v}_1 + \ldots + t_k\mathbf{v}_k + t_{k+1}\mathbf{u}_{k+1} + \ldots + t_n\mathbf{u}_n.$$

Wtedy

$$\begin{aligned}
\mathbf{w} = f(\mathbf{v}) &= f(t_1\mathbf{v}_1 + \ldots + t_k\mathbf{v}_k + t_{k+1}\mathbf{u}_{k+1} + \ldots + t_n\mathbf{u}_n) \\
&= f(t_1\mathbf{v}_1) + \ldots + f(t_k\mathbf{v}_k) + f(t_{k+1}\mathbf{u}_{k+1}) + \ldots + f(t_n\mathbf{u}_n) \\
&= t_1 f(\mathbf{v}_1) + \ldots + t_k f(\mathbf{v}_k) + t_{k+1} f(\mathbf{u}_{k+1}) + \ldots + t_n f(\mathbf{u}_n) \\
&= t_1(0) + \ldots + t_k(0) + t_{k+1} f(\mathbf{u}_{k+1}) + \ldots + t_n f(\mathbf{u}_n) \\
&= t_{k+1} f(\mathbf{u}_{k+1}) + \ldots + t_n f(\mathbf{u}_n),
\end{aligned}$$

więc

$$\operatorname{Im} f = \operatorname{Lin}\{f(\mathbf{u}_{k+1}), \ldots, f(\mathbf{u}_n)\}.$$

Wystarczy już tylko udowodnić, że wektory $f(\mathbf{u}_{k+1}), \ldots, f(\mathbf{u}_n)$ są liniowo niezależne.

Weźmy ciąg skalarów (s_i) taki, że

$$\begin{aligned} s_{k+1}f(\mathbf{u}_{k+1}) + \ldots + s_n f(\mathbf{u}_n) &= \overline{0} \\ f(s_{k+1}\mathbf{u}_{k+1}) + \ldots + f(s_n\mathbf{u}_n) &= \overline{0} \\ f(s_{k+1}\mathbf{u}_{k+1} + \ldots + s_n\mathbf{u}_n) &= \overline{0}. \end{aligned}$$

Z tego wynika, że

$$(s_{k+1}\mathbf{u}_{k+1} + \ldots + s_n\mathbf{u}_n) \in \operatorname{Ker} f,$$

więc $(s_{k+1}\mathbf{u}_{k+1} + \ldots + s_n\mathbf{u}_n) \in V$ możemy zapisać jako kombinację liniową wektorów $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Zakładając, że ciąg (s_i) jest niezerowy, mamy dwa sposoby zapisu jednego wektora z V, co stoi w sprzeczności z twierdzeniem 3.34. Z tego wynika, że $\forall i : s_i = 0$, więc wektory $f(\mathbf{u}_{k+1}), \ldots, f(\mathbf{u}_n)$ są liniowo niezależne, ergo

$$\dim \operatorname{Im} f = n - k,$$

$$\operatorname{rank} f = \dim V - \operatorname{null} f.$$

□

Definicja 7.8. Przy danych przestrzeniach wektorowych V, W nad ciałem $\mathbb{K}$, odwzorowanie liniowe $f : V \to W$ to:

- *monomorfizm*, jeśli jest injekcją,
- *epimorfizm*, jeśli jest surjekcją,
- *izomorfizm*, jeśli jest bijekcją,
- *endomorfizm*, jeśli $V = W$,
- *automorfizm*, jeśli jest endomorfizmem i izomorfizmem,
- *forma liniowa*, jeśli $W = \mathbb{K}$.

Twierdzenie 7.9

Odwzorowanie liniowe $f : V \to W$ jest epimorfizmem wtedy i tylko wtedy, gdy $\dim \operatorname{Im} f = \dim W$.

Dowód. $\operatorname{Im} f$ jest podprzestrzenią W, więc wynika z twierdzenia 3.42. □

Twierdzenie 7.10

Odwzorowanie liniowe $f : V \to W$ jest monomorfizmem wtedy i tylko wtedy, gdy $\operatorname{Ker} f = \{\overline{0}\}$.

Dowód. Implikacja prawostronna jest trywialna, dlatego udowodnimy tylko lewostronną. Załóżmy przeciwnie, że istnieje takie $\mathbf{v}_1 \neq \mathbf{v}_2$, że

$$f(\mathbf{v}_1) = f(\mathbf{v}_2).$$

Wtedy

$$\begin{aligned} f(\mathbf{v}_1) - f(\mathbf{v}_2) = f(\mathbf{v}_2) - f(\mathbf{v}_1) &= \overline{0} \\ f(\mathbf{v}_1 - \mathbf{v}_2) = f(\mathbf{v}_2 - \mathbf{v}_1) &= \overline{0}, \end{aligned}$$

co, skoro $\mathbf{v}_1 - \mathbf{v}_2 \neq \overline{0}$, przeczy założeniu $\operatorname{Ker} f = \{\overline{0}\}$. □

Twierdzenie 7.11

Jeśli V, W są skończenie wymiarowymi przestrzeniami wektorowymi nad ciałem $\mathbb{K}$, to ich izomorficzność jest równoważna równości ich wymiarów

$$V \sim W \quad \Leftrightarrow \quad \dim V = \dim W.$$

Dowód. Wynika z wniosku 7.2. □

Twierdzenie 7.12

Niech V, W będą pewnymi przestrzeniami nad ciałem $\mathbb{K}$, a $\mathcal{L}(V, W)$ zbiorem wszystkich odwzorowań liniowych między nimi. Struktura $(\mathcal{L}(V, W), \mathbb{K}, +, \cdot)$ jest przestrzenią wektorową.

Dowód. Z definicji. □

§7.1 Macierze odwzorowań liniowych

Definicja 7.13. Niech $B_V = (e_1, e_2, \ldots, e_n), B_W = (l_1, l_2, \ldots, l_m)$ będą pewnymi bazami odpowiednio przestrzeni V, W nad ciałem $\mathbb{K}$. Niech $f : V \to W$ będzie odwzorowaniem liniowym takim, że

$$\begin{aligned} f(e_1) &= a_{11}l_1 + a_{21}l_2 + \ldots + a_{m1}l_m, \\ f(e_2) &= a_{12}l_1 + a_{22}l_2 + \ldots + a_{m2}l_m, \\ &\vdots \\ f(e_n) &= a_{1n}l_1 + a_{2n}l_2 + \ldots + a_{mn}l_m. \end{aligned}$$

Macierz odwzorowania liniowego f w bazach B_V, B_W to macierz

$$M_f(B_V, B_W) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Fakt 7.14. Równanie $y = f(x)$ można zapisać w postaci macierzowej jako

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

gdzie $f : B_V \to B_W$ jest odwzorowaniem liniowym, $[a_{ij}]_{m\times n} = M_f(B_V, B_W)$ oraz

$$x = [x_1, x_2, \ldots, x_n]^T_{B_V}, \quad y = [y_1, y_2, \ldots, y_m]^T_{B_W}.$$

Dowód. Skoro $x = [x_1, x_2, \ldots, x_n]^T_{B_V}$ oraz $y = [y_1, y_2, \ldots, y_m]^T_{B_W}$, to

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n, \quad y = y_1l_1 + y_2l_2 + \ldots + y_ml_m,$$

więc

$$\begin{aligned} f(x) &= f(x_1e_1 + x_2e_2 + \ldots + x_ne_n) \\ f(x) &= x_1f(e_1) + x_2f(e_2) + \ldots + x_nf(e_n) \\ f(x) &= x_1\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ \vdots \\ a_{m1} \end{bmatrix} + x_2\begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots + x_n\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ \vdots \\ a_{mn} \end{bmatrix} \end{aligned} \tag{15}$$

Mnożąc macierze jak w udowadnianej tezie, otrzymalibyśmy

$$y = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{bmatrix},$$

czyli to samo co wyżej. □

Jeśli odwzorowanie f jest endomorfizmem, to macierz tego odwzorowania w bazie B oznaczamy jako $M_f(B)$.

Twierdzenie 7.15

Jeśli $f : V \to W$ będzie odwzorowaniem liniowym, a B_V, B_W to bazy odpowiednio przestrzeni V, W, to

$$\operatorname{rank} f = \operatorname{rank} M_f(B_V, B_W).$$

Dowód. Przypomnijmy, że

$$\operatorname{Im} f = \{y \in W \mid \exists x \in V : y = f(x)\}.$$

Korzystając z faktu 7.14, możemy przekształcić powyższe do

$$\operatorname{Im} f = \{M_f(B_V, B_W) \cdot x \mid x \in V\},$$

z czego wynika, że obraz przekształcenia f jest tożsamy z przestrzenią rozpinaną przez kolumny macierzy $M_f(B_V, B_W)$ (zobacz równanie 15). Z tego powodu również ich wymiary są równe. □

Twierdzenie 7.16

Niech $f, g : V \to W$ będą odwzorowaniami liniowymi, B_V, B_W to bazy odpowiednio przestrzeni V, W nad ciałem $\mathbb{K}$, a $\alpha \in \mathbb{K}$, to

1. $M_{f+g}(B_V, B_W) = M_f(B_V, B_W) + M_g(B_V, B_W)$,
2. $M_{\alpha f}(B_V, B_W) = \alpha M_f(B_V, B_W)$,

Dowód. Z definicji. □

Twierdzenie 7.17

Niech $f : V \to W, g : W \to U$ będą odwzorowaniami liniowymi, a B_V, B_W, B_U to bazy odpowiednio przestrzeni V, W, U, to

$$M_{g\circ f}(B_V, B_U) = M_g(B_W, B_U) \cdot M_f(B_V, B_W).$$

Dowód. Z definicji oraz faktu 7.14. □

Wniosek 7.18 (z faktu 7.14)

Endomorfizm $f : V \to V$ jest automorfizmem (bijekcją) wtedy i tylko wtedy, gdy macierz przekształcenia liniowego jest odwracalna (nieosobliwa). Łatwo również znaleźć przekształcenie odwrotne f^{-1}:

$$M_{f^{-1}}(B_2, B_1) = (M_f(B_1, B_2))^{-1}.$$

Definicja 7.19. Macierz przejścia $P_{B\to B'}$ od bazy $B = (e_1, e_2, \ldots, e_n)$ do bazy $B' = (e'_1, e'_2, \ldots, e'_n)$ przestrzeni V to macierz odwzorowania identycznościowego danej przestrzeni w bazach B', B,

$$P_{B\to B'} = M_{\mathrm{Id}_V}(B', B).$$

Wtedy, jeśli X, X' są wektorami kolumnowymi z V względem baz B, B', to

$$X = P_{B\to B'}X'.$$

Uwaga. W powyżej opisanej macierzy j-tą kolumnę tworzą współrzędne j-tego wektora bazy B' względem bazy B. Najłatwiej jest więc wtedy, gdy baza B jest bazą kanoniczną.

Twierdzenie 7.20 (o zmianie macierzy odwzorowania przy zmianie baz)

Niech $f : V \to W$ będzie odwzorowaniem liniowym, a B_V, B'_V, B_W, B'_W pewnymi bazami odpowiednich przestrzeni liniowych. Wtedy

$$M_f(B'_V, B'_W) = (P_{B_W\to B'_W})^{-1} \cdot M_f(B_V, B_W) \cdot P_{B_V\to B'_V}.$$

Dowód. Mamy

$$P_{B'_W\to B_W} \cdot M_f(B_V, B_W) \cdot P_{B_V\to B'_V} = M_{\mathrm{Id}_W}(B_W, B'_W) \cdot M_f(B_V, B_W) \cdot M_{\mathrm{Id}_V}(B'_V, B_V)$$

□

Przykład 7.21

Dane jest odwzorowanie liniowe $f : \mathbb{R}^3 : \mathbb{R}^2$ takie, że

$$M_f(B_1', B_2') = \begin{bmatrix} 2 & 1 & 2 \\ -1 & 0 & 0 \end{bmatrix},$$

gdzie

$$B_1' = ((1,1,0),(0,1,1),(1,1,1)),$$
$$B_2' = ((1,1),(-1,1)).$$

Znajdź macierz odwzorowania f w bazach

$$B_1 = ((1,0,0),(0,1,0),(0,0,1)),$$
$$B_2 = ((1,0),(0,1)).$$

Rozwiązanie. Z twierdzenia 7.20 mamy

$$M_f(B_1, B_2) = P_{B_2 \to B_2'} \cdot M_f(B_1', B_2') \cdot P_{B_1' \to B_1}.$$

Macierz przejścia $B_2 \to B_2'$ jest bardzo łatwo wyznaczyć, ponieważ B_2 to baza kanoniczna. W przypadku macierzy przejścia $B_1' \to B_1$ najłatwiej będzie wyznaczyć jej odwrotność, ponieważ B_1 jest bazą kanoniczną. Mamy więc

$$M_f(B_1, B_2) = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 & 2 \\ -1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}^{-1} = \cdots = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \end{bmatrix}$$

□

Definicja 7.22. Macierze $A, B \in M_{m\times n}$ są równoważne, jeśli istnieją takie nieosobliwe macierze P, Q, że

$$B = Q^{-1} \cdot A \cdot P.$$

Definicja 7.23. Macierze $A, B \in M_{m\times n}$ są podobne, jeśli istnieje taka nieosobliwa macierz P, że

$$B = P^{-1} \cdot A \cdot P.$$

Z definicji wynika, że dwie macierze pewnego odwzorowania liniowego są równoważne, a dwie macierze pewnego endomorfizmu są podobne. Ponadto, łatwo wykazać, że macierze równoważne mają równe rzędy, a podobne — rzędy i wyznaczniki.

§7.2 Wartości własne i wektory własne

Definicja 7.24. Niech $f : V \to V$ będzie endomorfizmem. Skalar $\lambda \in \mathbb{K}$ jest *wartością własną* tego endomorfizmu, jeśli istnieje niezerowy wektor $\mathbf{v}$ taki, że

$$f(\mathbf{v}) = \lambda\mathbf{v}.$$

Każdy niezerowy wektor $\mathbf{v}$, który spełnia powyższą równość, jest *wektorem własnym* odpowiadającym wartości własnej λ.

Definicja 7.25. Widmo endomorfizmu to zbiór jego wartości własnych.

Definicja 7.26. Podprzestrzeń własną V_λ endomorfizmu $f : V \to V$ tworzy zbiór

$$V_\lambda = \{\mathbf{v} \in V : f(\mathbf{v}) = \lambda\mathbf{v}\},$$

czyli zbiór wektorów własnych z wektorem zerowym.

Uzasadnienie. Fakt, że zbiór V_λ tworzy podprzestrzeń liniową przestrzeni V, wynika z addytywności i jednorodności endomorfizmów. □

Twierdzenie 7.27

Każdy wektor własny odpowiada dokładnie jednej wartości własnej.

Dowód. Załóżmy przeciwnie, że są dwie wartości własne λ_1, λ_2 dla których

$$f(\mathbf{v}) = \lambda_1\mathbf{v} = \lambda_2\mathbf{v}.$$

Wtedy

$$\overline{0} = f(\mathbf{v}) - f(\mathbf{v}) = \lambda_1\mathbf{v} - \lambda_2\mathbf{v} = \mathbf{v}(\lambda_1 - \lambda_2).$$

Skoro $\mathbf{v} \neq \overline{0}$, to $\lambda_1 = \lambda_2$. □

Twierdzenie 7.28

Niech $f : V \to V$ będzie endomorfizmem oraz $A = M_f(B)$ dla bazy $B = (e_1, \ldots, e_n)$. Wtedy $\lambda \in \mathbb{K}$ jest wartością własną wtedy i tylko wtedy, gdy

$$\det(A - \lambda I) = 0,$$

a niezerowy wektor $\mathbf{v} \in V$ jest wektorem własnym odpowiadającym λ wtedy i tylko wtedy, gdy

$$(A - \lambda I)X = \overline{0},$$

gdzie X jest macierzą kolumnową współrzędnych wektora $\mathbf{v}$ w bazie B.

Dowód. Mamy

$$f(\mathbf{v}) = \lambda\mathbf{v} \Leftrightarrow AX = \lambda X,$$

co możemy przekształcić do

$$AX - \lambda X = \overline{0}$$
$$(A - \lambda I)X = \overline{0}.$$

Z twierdzenia Kroneckera-Capellego (5.6) wynika, że powyższy układ równań liniowych ma co najmniej jedno rozwiązanie. A ponieważ jest to układ jednorodny, jednym z rozwiązań na pewno jest $X = \overline{0}$. A więc stwierdzenie, że λ jest wartością własną f (to znaczy, że istnieje inne, niezerowe rozwiązanie X), jest (na mocy twierdzenia 5.7) równoważne temu, że $\operatorname{rank}(A - \lambda I) < n$, a więc $\dim(A - \lambda I) = 0$. □

Definicja 7.29. Wielomian charakterystyczny endomorfizmu $f : V \to V$ to wielomian

$$\Delta(f) = \det(A - \lambda I) = (-\lambda)^n + \ldots + \det A,$$

gdzie A jest macierzą endomorfizmu f w dowolnej bazie.

Fakt 7.30. Wartości własne endomorfizmu f są wyznaczane przez pierwiastki jego wielomianu charakterystycznego $\Delta(f)$.

Dowód. Wynika z twierdzenia 7.28 oraz twierdzenia Bézouta. □

Lemat 7.31

Wektory własne odpowiadające różnym wartościom własnym danego endomorfizmu są liniowo niezależne.

Dowód. Przeprowadzimy dowód indukcyjny. Jeden niezerowy wektor jest oczywiście liniowo niezależny, więc załóżmy, że zbiór $n-1$ wektorów własnych odpowiadających różnym wartościom własnym jest liniowo niezależny i pokażmy ten fakt również dla n wektorów.

Oznaczmy wektory własne $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Niech odpowiadają one różnym wartościom własnym $\lambda_1, \ldots, \lambda_n$. Weźmy takie skalary $\alpha_1, \ldots, \alpha_n$, że

$$\alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n = \overline{0}, \tag{16}$$

$$f(\alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n) = \overline{0},$$

$$\alpha_1\lambda_1\mathbf{v}_1 + \ldots + \alpha_n\lambda_n\mathbf{v}_n = \overline{0}. \tag{17}$$

Mnożąc równanie 16 obustronnie przez λ_n i odejmując od 17 otrzymamy

$$\alpha_1\mathbf{v}_1(\lambda_1 - \lambda_n) + \ldots + \alpha_{n-1}\mathbf{v}_1(\lambda_{n-1} - \lambda_n) = \overline{0}.$$

Wektory $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ są, na mocy założenia indukcyjnego, liniowo niezależne, więc dla każdego $1 \leq i \leq n-1$ zachodzi

$$\alpha_i(\lambda_i - \lambda_n) = 0,$$

a więc, skoro rozważamy różne wartości własne,

$$\alpha_i = 0.$$

Podstawiając to do równości 16 otrzymujemy

$$\alpha_n\mathbf{v}_n = \overline{0}.$$

Wektory własne są niezerowe, więc $\alpha_n = 0$, co, na mocy zasady indukcji matematycznej, dowodzi tezy. □

Definicja 7.32. Endomorfizm $f : V \to V$ jest *diagonalizowalny*, jeśli istnieje baza przestrzeni V, w której macierz endomorfizmu f jest diagonalna.

Twierdzenie 7.33

Endomorfizm $f : V \to V$, $\dim V < \infty$, jest diagonalizowalny wtedy i tylko wtedy, gdy istnieje baza przestrzeni V utworzona z wektorów własnych endomorfizmu V.

Dowód. Weźmy taką bazę $B = (e_1, \ldots, e_n)$, że macierz

$$M_f(B) = \begin{bmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_2 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \lambda_n \end{bmatrix}$$

jest diagonalna. Wtedy dla każdego $1 \leq i \leq n$ zachodzi

$$M_f(B) \cdot e_i = f(e_i) = \lambda_i e_i,$$

więc każdy wektor e_i jest wektorem własnym endomorfizmu f.

Aby uzyskać dowód konieczności, wystarczy powtórzyć powyższy wywód w odwrotnej kolejności — wszystkie implikacje były równoważnościami. □

Wniosek 7.34 (z twierdzenia 7.33)
Jeśli endomorfizm $f : V \to V$, ma $n = \dim V$ różnych wartości własnych, to jest diagonalizowalny.

Dowód. Na mocy lematu 7.31, wektory odpowiadające różnym wartościom własnym są liniowo niezależne, więc stanowią bazę, ergo endomorfizm f jest diagonalizowalny. Warto zauważyć, że jeśli endomorfizm nie ma n różnych wartości własnych, to nie znaczy, że nie jest diagonalizowalny. □

Krotność algebraiczna k_a wartości własnej λ to jej krotność jako pierwiastka wielomianu charakterystycznego, a jej *krotność geometryczna* k_g to wymiar przestrzeni własnej V_λ, czyli $k_g = \dim V_\lambda$.

Lemat 7.35
Jeżeli λ jest wartością własną endomorfizmu f o krotności algebraicznej k_a i geometrycznej k_g, to

$$1 \leq k_g \leq k_a,$$

czyli krotność algebraiczna jest większa lub równa geometrycznej.

Dowód. Pierwsza nierówność wynika wprost z definicji obu krotności.

Niech $A \in \mathcal{M}_{n\times n}$ będzie macierzą endomorfizmu f, który ma pewną wartość własną λ_0 o krotności geometrycznej k_g. Z twierdzenia 7.28 mamy, że

$$k_g = \dim \operatorname{Ker}(A - \lambda_0 I) = \operatorname{null}(A - \lambda_0 I),$$

a z twierdzenia o rzędzie (7.7) wynika, że $\operatorname{rank}(A - \lambda_0 I) = n - k_g$. To z kolei znaczy, że możemy wykonać taki ciąg operacji elementarnych, że uzyskamy macierz schodkową $B \sim (A - \lambda_0 I)$, której ostatnie k_g wiersze są zerowe.

Weźmy pewną zmienną t i wykonajmy identyczny ciąg operacji elementarnych jak poprzednio, otrzymując macierz $C \sim (A - tI)$. Oczywiście, dla $t = \lambda_0$ ostatnie k_g wierszy macierzy C się zeruje, więc $\det C$ jest podzielny przez $(t-\lambda_0)^{k_g}$, ergo $\det(A-\lambda_0 I)$ również (różnią ewentualnie o stałą), stąd $k_g \leq k_a$. □

Uwaga (mnemotechnika)**.** Krotność algebraiczna (ang. *algebraic multiplicity*) i krotność geometryczna (ang. *geometric multiplicity*) to w skrócie odpowiednio AM i GM. Podobnie średnia arytmetyczna (ang. *arithmetic mean*) i średnia geometryczna (ang. *geometric mean*) to w skrócie odpowiednio AM i GM. W obu przypadkach zachodzi zależność

$$\text{AM} \geq \text{GM}.$$

Jeśli nierówność między średnimi nie jest dla Czytelnika oczywista, warto zobaczyć artykuł na wikipedii.

Twierdzenie 7.36

Endomorfizm $f : V \to V$ jest diagonalizowalny wtedy i tylko wtedy, gdy wielomian charakterystyczny jest rozkładalny do postaci

$$\Delta(\lambda) = (\lambda_1 - \lambda)^{k_{a1}} \cdots (\lambda_p - \lambda)^{k_{ap}}$$

oraz dla każdego i zachodzi

$$k_{gi} = k_{ai},$$

czyli gdy krotności geometryczne i algebraiczne poszczególnych wartości własnych są sobie równe.

Uwaga 7.37. Jeśli istnieje baza B przestrzeni V, w której macierz endomorfizmu $f : V \to V$ jest diagonalna, to V jest sumą (prostą, co wynika z lematu 7.31) przestrzeni własnych, więc B jest sumą mnogościową baz tych przestrzeni.

Przykład 7.38

Sprawdź, czy endomorfizm $f : \mathbb{R}^3 \to \mathbb{R}^3$ dany wzorem

$$f(x, y, z) = (-x - 3z, x + 2y + 3z, -3x - z)$$

jest diagonalizowalny. Jeśli tak, wyznacz bazę B taką, że $D = M_f(B)$ jest diagonalna, podaj D oraz macierz P taką, że $M_f(B_k) = PDP^{-1}$.

Rozwiązanie. Najpierw wyznaczymy macierz $M_f(B_k)$. Mamy

$$f(1,0,0) = [-1,3,-3]_{B_k}, \quad f(0,1,0) = [0,2,0]_{B_k}, \quad f(0,0,1) = [-3,3,-1]_{B_k},$$

tak więc

$$M_f(B_k) = \begin{bmatrix} -1 & 0 & -3 \\ 3 & 2 & 3 \\ -3 & 0 & -1 \end{bmatrix}.$$

Aby wyznaczyć wartości własne, skorzystamy z faktu 7.30.

$$\begin{aligned} \Delta(\lambda) &= \begin{vmatrix} -1-\lambda & 0 & -3 \\ 3 & 2-\lambda & 3 \\ -3 & 0 & -1-\lambda \end{vmatrix} = (2-\lambda) \begin{vmatrix} -1-\lambda & -3 \\ -3 & -1-\lambda \end{vmatrix} \\ &= (2-\lambda)((1+\lambda)^2 - 3^2) = (2-\lambda)(\lambda+4)(\lambda-2) \\ &= -(2-\lambda)^2(4+\lambda) \end{aligned}$$

Tak więc mamy wartości własne $\lambda_1 = 2, \lambda_2 = -4$ z krotnościami odpowiednio $k_{a1} = 2, k_{a2} = 1$. Nie mamy trzech różnych wartości własnych, więc, aby sprawdzić, czy f jest diagonalizowalny, nie możemy zastosować wniosku 7.34; będziemy musieli skorzystać z pełnego twierdzenia 7.36.

Wektor własny $[a, b, c]$ odpowiada wartości własnej λ_1 wtw, gdy

$$\begin{bmatrix} -3 & 0 & -3 \\ 3 & 0 & 3 \\ -3 & 0 & -3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Leftrightarrow a = -c.$$

W takim razie

$$k_{g1} = \dim V_{\lambda_1} = \dim \operatorname{Lin}\{(1,0,-1),(0,1,0)\} = 2.$$

Podobnie dla wartości własnej λ_2 mamy

$$\begin{bmatrix} 3 & 0 & -3 \\ 3 & 6 & 3 \\ -3 & 0 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Leftrightarrow a = c, a + 2b + c = 0 \Leftrightarrow a = c, a = -b.$$

A więc

$$k_{g2} = \dim V_{\lambda_2} = \dim \operatorname{Lin}\{(1, -1, 1)\} = 1,$$

z czego wynika, że f jest diagonalizowalny.

Zgodnie z uwagą 7.37 mamy

$$B = ((1, 0, -1), (0, 1, 0), (1, -1, 1))$$

oraz

$$M_f(B) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -4 \end{bmatrix}.$$

Musimy wiec jeszcze znaleźć macierz $P = P_{B_k \to B}$:

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

□

Możemy również wykorzystać własności macierzy diagonalnych bez potrzeby odwoływania się do endomorfizmów. Na podstawie definicji 7.22 i 7.23 będziemy mówili, że macierz jest diagonalizowalna, jeśli jest podobna do pewnej macierzy diagonalnej; że ma wartości własne oraz wektory własne (gdy spełnia równania z twierdzenia 7.28) i tym podobne. Wszystkie wyprowadzone dla endomorfizmów twierdzenia będą również prawdziwe dla macierzy.

Przykład 7.39

Niech $A = \begin{bmatrix} 1 & -3 \\ 1 & 5 \end{bmatrix}$. Oblicz A^n.

Rozwiązanie.

$$\Delta(\lambda) = \begin{vmatrix} 1 - \lambda & -3 \\ 1 & 5 - \lambda \end{vmatrix} = (1 - \lambda)(5 - \lambda) - (-3) = 8 - 6\lambda + \lambda^2 = (2 - \lambda)(4 - \lambda),$$

a więc $\lambda_1 = 2, \lambda_2 = 4$. Korzystając z wniosku 7.34, dostajemy, że macierz A jest diagonalizowalna, więc możemy przedstawić ją jako

$$A^n = (PDP^{-1})^n = PD^nP^{-1}$$

przy

$$D^n = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}^n = \begin{bmatrix} 2^n & 0 \\ 0 & 4^n \end{bmatrix}.$$

Powyższa równość wynika z faktu, że macierz diagonalną potęguje się bardzo łatwo — wystarczy podnieść wszystkie jej elementy do potęgi — co nietrudno jest udowodnić.

Aby znaleźć macierz P, najłatwiej będzie potraktować A oraz D jako macierze tego samego endomorfizmu w różnych bazach. Przyjmiemy, że to A jest w bazie kanonicznej i znajdziemy bazę, w której jest D. Najpierw znajdźmy bazy przestrzeni własnych:

$$\begin{bmatrix} -1 & -3 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Leftrightarrow a = -3b,$$

więc $V_{\lambda_1} = \mathrm{Lin}\{(3,-1)\}$ oraz

$$\begin{bmatrix} -3 & -3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\Leftrightarrow a = -b,$$

więc $V_{\lambda_2} = \mathrm{Lin}\{(1,-1)\}$.

Mamy już $B = ((3,-1),(1,-1))$ oraz

$$P = P_{B_k \to B} = \begin{bmatrix} 3 & 1 \\ -1 & -1 \end{bmatrix}.$$

Obliczmy jeszcze macierz P^{-1}:

$$\left[\begin{array}{cc|cc} 3 & 1 & 1 & 0 \\ -1 & -1 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & -1 & 1 & 2 \\ -1 & -1 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & -1 & 1 & 2 \\ 0 & -2 & 1 & 3 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} & -\frac{3}{2} \end{array}\right],$$

tak więc

$$P^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & -3 \end{bmatrix}.$$

Mamy już więc ostateczny wynik

$$\begin{aligned} A^n &= \frac{1}{2} \begin{bmatrix} 3 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 2^n & 0 \\ 0 & 4^n \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & -3 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 3 \cdot 2^n & 4^n \\ -2^n & -4^n \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & -3 \end{bmatrix} \\ &= 2^{n-1} \begin{bmatrix} 3 - 2^n & 3 - 3 \cdot 2^n \\ 2^n - 1 & 3 \cdot 2^n - 1 \end{bmatrix}. \end{aligned}$$

□