Algebra

Michał Dobranowski

semestr zimowy 2022
v0.8

Poniższy skrypt zawiera materiał obejmujący wykłady z Algebry prowadzone przez dr hab. Jakuba Przybyło na I semestrze Informatyki na AGH oraz tematy, które uznałem za warte uwagi podczas własnych studiów nad tematem.

Spis treści

§1 Liczy zespolone

Definicja 1.1. Liczba zespolona z to uporządkowana para liczb rzeczywistych. Pierwszy element tej pary to **część rzeczywista**, ozaczana 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

$$i^2 = -1.$$

§1.1 Działania na liczbach zespolonych

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 $z_1 z_2 = x_1x_2 + ix_1y_2 + ix_2y_1 + i^2y_1y_2$
  $= x_1x_2 - y_1y_2 + i(x_1y_2 + x_2y_1)$

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:

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

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

Powyższa liczba jest liczbą rzeczywistą, więc znaleźliśmy prosty sposób na dzielenie liczb zespolonych przez siebie, mnożąc licznik i mnianownik 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

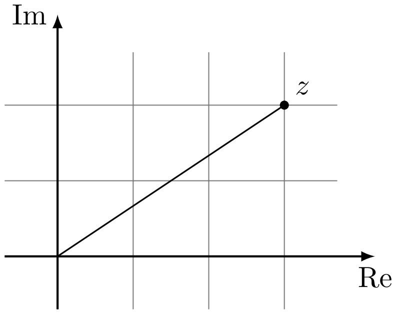
Oprócz $z\overline{z} = |z|^2$, zachodzą również równości:

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

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

§1.2 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.

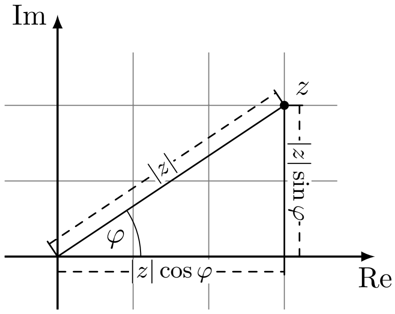
Dowód. Wynika to z twierdzenia Pitagorasa oraz definicji modułu (1.4). □

Możemy wyprowadzić **postać trygonometryczną** liczby zespolonej, która, zamiast dwóch współrzędnych, 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 $\operatorname{Arg}(z)$. 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.



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

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 z twierdzenia 1.9.

§1.3 Pierwiastkowanie liczb zespolonych

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 $\frac{2\pi}{n}$. □

§1.4 Postać wykładnicza

Postać $z = |z|e^{i\varphi}$ liczby zespolonej bedziemy 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. Weźmy $z = \cos\varphi + i\sin\varphi$. Różniczkując po zmiennej φ otrzymujemy

$$\frac{dz}{d\varphi} = -\sin\varphi + i\cos\varphi = iz$$

$$\therefore \frac{dz}{z} = id\varphi.$$

Po obustronnym całkowaniu mamy

$$\int \frac{dz}{z} = \int id\varphi$$
$$\ln z = i\varphi + c$$
$$e^{\ln z} = e^{i\varphi + c}$$
$$z = e^{i\varphi + c}.$$

Podstawiając $\varphi = 0$ otrzymujemy $1 = e^c$, skąd mamy $c = 0$, co kończy dowód. □

§2 Relacje

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

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

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

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

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

Definicja 2.3. Złożeniem relacji $\mathcal{R} = (X, \mathrm{gr}\mathcal{R}, Y)$ z relacją $\mathcal{S} = (Y, \mathrm{gr}\mathcal{S}, Z)$ nazywamy relację

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

gdzie

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

Definicja 2.4 (rodzaje relacji)**.** Relacja $\mathcal{R} = (X, \mathrm{gr}\mathcal{R}, X)$ 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 $\mathcal{R} = (X, \mathrm{gr}\mathcal{R}, X)$, 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. Założmy 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. $\square$

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 $\mathcal{R} = (X, \mathrm{gr}\mathcal{R}, X)$, 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

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

Uwaga 2.11. Analogicznie można zdefiniować **element najmniejszy** $\overline{m}$:

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

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

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

Twierdzenie 2.12

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 2.13

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. □

Fakt 2.14. 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

$$\sup A.$$

Uwaga 2.15. Analogicznie można zdefiniować **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 2.16

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 2.17

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 2.12) ma co najwyżej jedno supremum, to gdyby zbiór X miał element najwiekszy, to na mocy twierdzenia 2.16 byłoby ono również supremum, które należy do zbioru X. □

Przykład 2.18

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 2.17, element największy (a z faktu 2.14 również maksymalny) nie istnieje.

Przykład 2.19

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

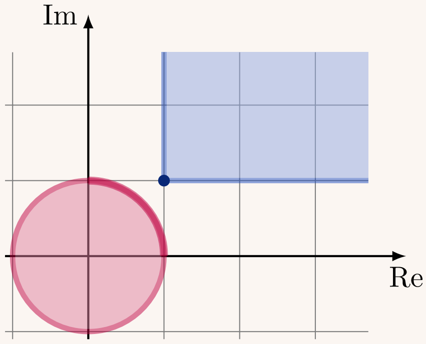
$$x \preceq y \Leftrightarrow \operatorname{Re} x \leq \operatorname{Re} y \wedge \operatorname{Im} x \leq \operatorname{Im} y.$$

Oczywiście niektóre elementy nie będą w tym porządku porównywalne, na przykład 1 oraz i.

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

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

Na rysunku zaznaczono zbiór A, zbiór majorant M zbioru A, supremum zbioru A oraz zbiór elementów maksymalnych (jako ćwierćokrąg). Na mocy wniosku 2.17 element największy nie istnieje.



Definicja 2.20. Łańcuch to taki podziór $C \subset X$, że $(X, \preceq)$ jest zbiorem z porządkiem częściowym, a $(C, \preceq)$ jest zbiorem z porządkiem liniowym.

Definicja 2.21. 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]gdzie $\mathbb{Z}_n$ 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 wewnętrze 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

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

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

Pierwsze działanie w pierścieniu nazywamy **działaniem addytywnym** i oznaczamy przez $+$. 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}.

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

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 ciało, w którym działanie $\cdot$ jest przemienne. Niektórzy autorzy utożsamiają pojęcie ciała z ciałem przemiennym.

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

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 wsród 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 liczb wymiernych,
- $(\mathbb{R}, +, \cdot)$ – ciało liczb rzeczywistych,
- $(\mathbb{C}, +, \cdot)$ – ciało 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 pierwiastów 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 (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

$$\underset{u,v \in V}{\forall} \ \underset{\alpha \in K}{\forall} \ \alpha \cdot (u+v) = (\alpha \cdot u) + (\alpha \cdot v),$$

4. zachodzi „rozdzielność” działania $\cdot$ względem $+$ i $\oplus$, to znaczy

$$\underset{v \in V}{\forall} \ \underset{\alpha,\beta \in K}{\forall} \ (\alpha \oplus \beta) \cdot v = (\alpha \cdot v) + (\beta \cdot v),$$

5. zachodzi „łączność” działań $\cdot$ i $\otimes$, to znaczy

$$\underset{v \in V}{\forall} \ \underset{\alpha,\beta \in K}{\forall} \ (\alpha \otimes \beta) \cdot v = \alpha \cdot (\beta \cdot v),$$

6. jedynka z ciała $(K, \oplus, \otimes)$ jest elementem neutralnym również dla działania $\cdot$, to znaczy

$$\underset{v \in V}{\forall} \ \mathbf{1} \cdot v = v.$$

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 starty 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 $\alpha_1, \ldots, \alpha_{n-1}$. $\square$

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. $\square$

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. $\square$

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 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. $\square$

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

$$\underset{v \in V_1 + V_2}{\forall} \ \underset{v_1 \in V_1}{\exists!} \ \underset{v_2 \in V_2}{\exists!} \ v = v_1 + v_2.$$

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. $\square$

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

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 $B_1, B_2 \subset B$. Oczywistym 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:
$$\begin{bmatrix} a_{11} & a_{12} \cdots\cdots a_{1n} \\ a_{21} & a_{22} \cdots\cdots a_{2n} \\ \vdots & \vdots \quad \ddots \quad \vdots \\ a_{m1} & a_{m2} \cdots\cdots a_{mn} \end{bmatrix}.$$

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:

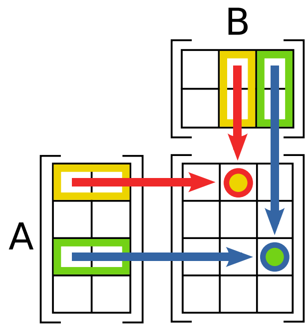
$$A + B = [a_{ij} + b_{ij}]_{m\times n},$$

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 $M_{m\times n}(\mathbb{K})$ 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[2].

Fakt 4.9. Zachodzi równość

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

§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

$$\varepsilon(\sigma) = (-1)^{[\sigma]},$$

gdzie $[\sigma]$ to liczba inwersji w permutacji σ.

[2] 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).

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

$$(\sigma_1, \sigma_2, \ldots, \sigma_i, \ldots, \sigma_j, \ldots, \sigma - n - 1, \sigma_n).$$

Zamieniając σ_i oraz σ_j nie zmieni się liczba inwersji zawierająćh elementy $\sigma_k, k \in [1, i) \cup (j, n]$. Nie zmieni się również liczba inwersji zawierających elementy $\sigma_k, k \in (i, j)$ takie, że σ_k jest większe lub mniejsze jednocześnie od σ_i i σ_j.

Dla pozostałych elementów $\sigma_k, k \in (i, j)$ 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 $[\sigma]$ 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 $A = [a_{ij}] \in M_{n \times n}(\mathbb{K})$ 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 kolumnami lub wierszami, 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 $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$.

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.
9. TODO

□

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. □

Fakt 4.16. Pole równoległoboku rozpiętego przez wektory będące wierszami (lub kolumnami) macierzy $A_{2\times 2}$ jest równe wartości bezwzględnej wyznacznika tej macierzy.

Dowód. Pole takiego równoległoboku to dwukrotność pola trójkąta o współrzędnych $(0, 0), (a_{11}, a_{12}), (a_{21}, a_{22})$. Ze wzoru

$$[\triangle ABC] = \frac{1}{2} |(x_B - x_A)(y_C - y_A) - (y_B - y_A)(x_C - x_A)|$$

otrzymujemy, że szukane pole P równoległoboku jest równe

$$P = 2 \cdot \frac{1}{2} |(a_{11} - 0)(a_{22} - 0) - (a_{12} - 0)(a_{21} - 0)| = |a_{11}a_{22} - a_{12}a_{21}|.$$

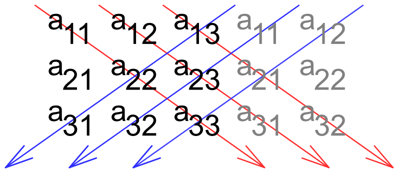
□

Twierdzenie 4.17 (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. □

Fakt 4.18. Objętość równoległościanu rozpiętego przez wektory będące wierszami (lub kolumnami) macierzy $A_{3\times 3}$ jest równa wartości bezwzględnej wyznacznika tej macierzy.



Rysunek 2: Reguła Sarrusa, źródło: Wikipedia.

Twierdzenie 4.19 (Cauchy'ego)
Dla dowolnych macierzy $A, B \in M_{n\times n}(\mathbb{K})$ zachodzi

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

Dowód. TODO □

Definicja 4.20. 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 kadratową, 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 4.21. Dopełnienie algebraiczne elementu a_{ij} macierzy kwadratowej $A = [a_{ij}]_{m\times n}$ to skalar

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

Twierdzenie 4.22 (Laplace'a)
Niech $A = [a_{ij}]_{n\times n}$ będzie macierzą kwadratową. Wtedy dla każdego $i \in \{1, 2 \ldots, n\}$

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

Wniosek 4.23
Wyznacznik macierzy trójkątnej jest równy iloczynowi elementów na jej przekątnej.

§4.3 Rząd macierzy

Twierdzenie 4.24
Maksymalna liczba liniowo niezależnych kolumn (wektorów z $\mathbb{K}^m$) dowolnej macierzy $A \in M_{m\times n}(\mathbb{K})$ jest równa maksymalnej liczbie liniowo niezależnych wierszy (wektorów z $\mathbb{K}^n$).

Dowód. TODO □

Definicja 4.25. Rząd macierzy to maksymalna liczba liniowo niezależnych kolumn (lub wierszy). Rząd macierzy A oznaczamy przez $\operatorname{rank}(A)$.

Z twierdzenia 4.24 i definicji 4.25 wynika, że

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

Definicja 4.26. 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 4.27. Rząd macierzy schodkowej jest równy liczbie jej schodków.

Definicja 4.28. 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 mówimy, że te macierze są **równoważne** i oznaczamy $A \sim B$.

Fakt 4.29. 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, za pomocą wniosku 4.23, obliczyć jej wyznacznik (jeśli jest kwadratowa)[3].

Przykład 4.30

Obliczyć wyznacznik macierzy

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

Rozwiązanie.

$$\begin{vmatrix} 1 & 3 & 1 \\ 1 & 1 & -1 \\ 3 & 11 & 6 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 1 \\ 0 & -2 & -2 \\ 0 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 1 \\ 0 & -2 & -2 \\ 0 & 0 & 1 \end{vmatrix} = -2$$

□

Twierdzenie 4.31

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

Dowód. TODO □

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

Przykład 4.32

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 wykreśliliśmy tylko jeden wiersz), więc na mocy twierdzenia 4.31 rank $A = 3$.

Dla pewności można pokazać również inną metodę — eliminację Gaussa:
$$\begin{bmatrix} 1 & 3 & 4 \\ 3 & 4 & 1 \\ 1 & 2 & 7 \\ 3 & 5 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 \\ 0 & -5 & -11 \\ 0 & -1 & 3 \\ 0 & -4 & -13 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 \\ 0 & 0 & -26 \\ 0 & -1 & 3 \\ 0 & 0 & -25 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 \\ 0 & -1 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$
$$\therefore \operatorname{rank} A = \operatorname{rank} \left[\begin{smallmatrix} 1 & 3 & 4 \\ 0 & -1 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{smallmatrix}\right] = 3.$$

□

§4.4 Macierz odwrotna

Definicja 4.33. 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 4.34

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 macierzy A.

Dowód. 1. Z definicji macierzy odwrotnej (4.33)
$$\begin{aligned} A \cdot A^{-1} &= I \\ \det(A \cdot A^{-1}) &= \det I = 1. \end{aligned}$$

Na mocy twierdzenia Cauchy'ego (4.19) otrzymujemy
$$\det A \cdot \det(A^{-1}) = 1,$$

z czego wynika teza.

2. Weźmy macierz B taką, że

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

to znaczy, że dla każdego i, j zachodzi

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

Obliczmy teraz macierz $C = AB$:

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

Z rozwinięcia Laplace'a (4.22) otrzymujemy, że $c_{ij} = 1$ jeśli $i = j$ oraz $c_{ij} = 0$ w przeciwnym wypadku (TODO). W takim razie $C = I$, więc $B = A^{i-1}$.

□

Definicja 4.35. Macierz osobliwa to macierz A, której wyznacznik jest zerowy. W innym wypadku A jest macierzą nieosobliwą.

Na podstawie twierdzenia 4.34 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 4.34, ale zwykle szybszą[4] metodą będzie eliminacja Gaussa, którą możemy wykorzystać wraz z poniższym faktem.

Fakt 4.36. Jeśli macierz kwadratowa A jest odwracalna, to

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

Dowód. TODO □

Przykład 4.37

Znaleźć macierz odwrotną do macierzy

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

[4] na pewno w sensie złożoności obliczeniowej, w zadaniach to kwestia preferencji

Rozwiązanie.

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

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

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

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.

$$A^{-1}=\frac{1}{5\cdot4+6\cdot2-4\cdot2\cdot4-3}\begin{bmatrix}5\cdot4-3&-(2\cdot4)&2\\-(4\cdot4-6)&4&-(1)\\4\cdot3-6\cdot5&-(3-6\cdot2)&5-4\cdot2\end{bmatrix}^T$$
$$=\frac{1}{-3}\begin{bmatrix}17&-8&2\\-10&4&-1\\-18&9&-3\end{bmatrix}^T=\frac{1}{3}\begin{bmatrix}-17&10&18\\8&4&9\\-2&1&3\end{bmatrix}.$$

□

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 $x_1,\dots,x_n$ 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

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

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

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

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[5].

Układy równań liniowych można łatwo rozwiązać eliminacją Gaussa w podobny sposób, jak robiliśmy to szukając macierzy odwrotnej w przykładzie 4.37.

Przykład 5.8
Rozwiązać układ równań

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

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

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 bedzie teraz wrócic do układu równań. Mamy więc

$$z = 3,$$

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

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

□