Liczby zespolone i wielomiany

Równanie $x^2+1=0$ nie ma rozwiązania w zbiorze liczb rzeczywistych $\mathbb{R}$. Tak więc nie każdy wielomian o współczynnikach należących do $\mathbb{R}$ posiada miejsce zerowe (zwane inaczej pierwiastkiem) w tym zbiorze. Okazuje się jednak, że zbiór liczb rzeczywistych tak naprawdę zawiera się w większym zbiorze, tzw. liczb zespolonych $\mathbb{C}$, który już tej „wady" nie posiada. Początki rachunków na liczbach zespolonych to pierwsza połowa XVI wieku i są zasługą głownie matematyków włoskich – przede wszystkim Girolamo Cardano.[1]

Definicja. Zbiór $\mathbb{C}=\mathbb{R}\times\mathbb{R}$, w którym określono dwa działania dwuargumentowe + oraz · następującymi wzorami

$$(x_1,y_1)+(x_2,y_2):=(x_1+x_2,y_1+y_2), \tag{0.1}$$

$$(x_1,y_1)\cdot(x_2,y_2):=(x_1x_2-y_1y_2,x_1y_2+x_2y_1), \tag{0.2}$$

nazywamy *zbiorem liczb zespolonych*. Elementy tego zbioru nazywamy *liczbami zespolonymi*. Działanie + nazywamy dodawaniem (zespolonym), a działanie · nazywamy mnożeniem (zespolonym). Dla liczby zespolonej $z=(x,y)\in\mathbb{C}$ liczbę rzeczywistą x nazywamy *częścią rzeczywistą*, natomiast y *częścią urojoną* liczby zespolonej z. Część rzeczywistą liczby zespolonej z oznaczamy przez $\operatorname{Re} z$, a część urojoną przez $\operatorname{Im} z$.

Podstawowe własności działań w zbiorze liczb zespolonych

Niech $z,z_1,z_2,z_3\in\mathbb{C}$. Wtedy mamy następujące własności:

1) $z_1+z_2=z_2+z_1$ (*przemienność dodawania*)

2) $(z_1+z_2)+z_3=z_1+(z_2+z_3)$ (*łączność dodawania*)

3) $z+0=z$, gdzie $0=(0,0)$ (*istnienie elementu neutralnego dodawania*)

4) $\forall z\in\mathbb{C}\ \exists w\in\mathbb{C}$ takie, że $z+w=0$ (*istnienie elementu przeciwnego względem dodawania*)

---

[1] Girolamo Cardano (1501-1576) – włoski uczony. Zajmował się wieloma dziedzinami m.in. medycyną, astrologią, filozofia, fizyką i matematyką. Z wykształcenia był lekarzem, ale wydaje się, że największe osiągnięcia miał w matematyce. Opracował m.in. metodę rozwiązywania ogólnego równania trzeciego stopnia (w dzisiejszej symbolice: $ax^3+bx^2+cx+d=0$), badał poprawnie zjawiska losowe (które z sukcesem wykorzystywał w swojej praktyce hazardzisty) jeszcze przed słynną korespondencją pomiędzy Pascalem a Fermatem, którą się uważa z początek nauki o prawdopodobieństwie. W trakcie badań nad równaniem trzeciego stopnia pojawia się w jego pracach pierwiastek kwadratowy z liczb ujemnych (dzieło *Ars Magna sive de regulis algebraicis* zwane krótko *Ars Magna,* czyli *Wielka sztuka*). Na początku rachunki, w których występował symbole takie jak $\sqrt{-5}$ budziły opory, a nawet szyderstwa (stąd określenie na tego typu wyrażenia – *liczby urojone*). Jednak już w połowie XVIII wieku były akceptowane przez wiodących matematyków (L. Euler). Precyzyjne definicje liczb zespolonych pojawiają się pod koniec tego samego wieku (Caspar Wessel, Jean-Robert Argand, Carl Fryderyk Gauss).

5) $z_1 \cdot z_2 = z_2 \cdot z_1$ (*przemienność mnożenia*)

6) $1 \cdot z = z$ gdzie $1 = (1,0)$ (*istnienie elementu neutralnego dla mnożenia*)

7) jeżeli $z \neq 0$, to istnieje $w \in \mathbb{C}$ takie, że $z \cdot w = 1$ (*istnienie elementu przeciwnego względem mnożenia – nazywamy go liczbą odwrotną*)

8) $z_1 \cdot (z_2 + z_3) = z_1 \cdot z_2 + z_1 \cdot z_3$ (*rozdzielność mnożenia względem dodawania*)

DOWÓD. Większość z podanych własności uzasadnia się bardzo prosto w oparciu o definicje działań podane w (0.1) i (0.2). Weźmy dla przykładu własność 1). Niech $z_1 = (x_1, y_1),\ z_2 = (x_2, y_2) \in \mathbb{C}.$ Wtedy

$$z_1 + z_2 = (x_1, y_1) + (x_2, y_2) \overset{def.}{=} (x_1 + x_2, y_1 + y_2) = (x_2 + x_1, y_2 + y_1) = z_2 + z_1,$$

gdzie wykorzystaliśmy definicję dodawania (0.1) oraz przemienność dodawania zbiorze $\mathbb{R}$: $x_1 + x_2 = x_2 + x_1,\ \ y_1 + y_2 = y_2 + y_1.$ Analogicznie „przeliczamy" własność 2) (tym razem będziemy korzystać z łączności w $\mathbb{R}$), a punkt 2) to $z = (x, y) + (0,0) = (x + 0, y + 0) = (x, y) = z.$ Istnienie elementu przeciwnego wprost bazuje na istnieniu liczb przeciwnych w $\mathbb{R}$:

$$z + w = (x, y) + (-x, -y) = (x + (-x), y + (-y)) = (0,0) = 0,$$

tak więc $w = (-x, -y).$ Własność 5) wymaga zastosowania definicji (0.2) oraz przemienności mnożenia w $\mathbb{R}$:

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

Własność 7) wynika wprost z definicji: $1 \cdot z = (1,0) \cdot (x, y) = (1 \cdot x - 0 \cdot y, 1 \cdot y + 0 \cdot x) = (x, y) = z.$

Istnienie elementu odwrotnego w do $z \neq 0$ można wyprowadzić rozważając równość $z \cdot w = 1,$ czyli

$$(x, y) \cdot (w_1, w_2) = (1,0) \ \Rightarrow \ (x w_1 - y w_2, x w_2 + y w_1) = (1,\ 0) \ \Rightarrow$$

$$\begin{cases} x w_1 - y w_2 = 1, \\ x w_2 + y w_1 = 0. \end{cases}$$

Mnożymy pierwsze równanie przez $x,$ a drugie przez y i dodajemy stronami, co daje

$$x^2 w_1 + y^2 w_1 = x \ \Rightarrow \ w_1 = \frac{x}{x^2 + y^2}.$$

Podobnie wyliczamy w_2 (pierwsze mnożymy przez $y,$ drugie przez $x,$ następnie odejmujemy):

$$w_2 = \frac{-y}{x^2 + y^2}.$$

Dzielenie przez x^2+y^2 jest poprawne, gdyż jest to liczba niezerowa, bo $z=(x,y)\neq(0,0)$ z założenia. Możemy więc napisać, że

$$(x,y)^{-1}=\left(\tfrac{x}{x^2+y^2},-\tfrac{y}{x^2+y^2}\right)\in\mathbb{C}.$$

Rozdzielność mnożenia względem dodawania (własność 8)), to znów rachunek. Najlepiej rozpisać lewą i prawą stronę osobno, a następnie porównać. W trakcie rachunków oczywiście korzystamy z odpowiedniej rozdzielności w zbiorze $\mathbb{R}$.

Związek pomiędzy liczbami rzeczywistymi a zespolonymi

Jak się maja liczby rzeczywiste $\mathbb{R}$ do liczb zespolonych $\mathbb{C}$? Odpowiedź jest bardzo prosta: zbiór liczb rzeczywistych zawiera się w zbiorze liczb zespolonych: $\mathbb{R}\subset\mathbb{C}$. Może to budzić pewne wątpliwości, wszak liczby zespolone zostały określone jako pary liczb rzeczywistych (x,y), gdzie $x\in\mathbb{R}$ i $y\in\mathbb{R}$. Zauważmy jednak, że jeżeli rozpatrzymy podzbiór K zbioru $\mathbb{C}$ składający się z wszystkich liczb postaci $(x,0)$, tzn. $K=\{(x,0)\in\mathbb{C}:\ x\in\mathbb{R}\}$, to działania na elementach tego zbioru dają w wyniku elementy tego samego zbioru: dla dowolnych $x,y\in\mathbb{R}$ mamy

$$\begin{aligned}(x,0)+(y,0)&=(x+y,0)\in K,\\(x,0)\cdot(y,0)&=(xy,0)\in K.\end{aligned}\tag{0.3}$$

Oznacza to, że działania na parach postaci $(x,0)\in K\subset\mathbb{C}$ odpowiadają działaniom na samych liczbach $x\in\mathbb{R}$. Ponadto możemy wprowadzić uporządkowanie w zbiorze K tak, aby było zgodne z działaniami[2] dodawania i mnożenia w „zwykłym” zbiorze liczb rzeczywistych. Tak naprawę, gdy się dokładniej przyjrzymy zbiorowi $K\subset\mathbb{C}$, to się okaże że faktycznie nie różni się on własnościami od zbioru $\mathbb{R}$. Dlatego możemy przyjąć, że $\mathbb{R}=K$ i używać zawierania $\mathbb{R}\subset\mathbb{C}$.

---

[2] Liczby rzeczywiste są strukturą, na którą składają zbiór $\mathbb{R}$, działania $+$ i $\cdot$ oraz relacja mniejszości $<$. Dlatego matematycy mówią, że jest to czwórka: $(\mathbb{R},+,\cdot,<)$ przy czym muszą być spełnione pewne własności arytmetyczne (przemienność, łączność, rozdzielność itd.), a relacja mniejszości musi być powiązana z działaniami:

$$\begin{aligned}&(x,y,z\in\mathbb{R}\ \text{ i }\ x<y)\ \Rightarrow x+z<y+z,\\&(x,y,z\in\mathbb{R},\ z>0\ \text{ i }\ x<y)\ \Rightarrow x\cdot z<y\cdot z.\end{aligned}$$

Dodatkowo zbiór $\mathbb{R}$ posiada fundamentalną własność zwaną *aksjomatem ciągłości*: każdy niepusty i ograniczony z góry podzbiór $A\subset\mathbb{R}$ posiada kres górny należący do $\mathbb{R}$. Kres górny nie musi leżeć w zbiorze A czego przykładem może być odcinek $[1,3)$, którego kres górny $3\notin[1,3)$. Aksjomat ciągłości jest tym co różni zbiór liczb rzeczywistych od zbioru liczb wymiernych $\mathbb{Q}$. Dzięki tej własności można np. Udowodnić istnienie pierwiastków $\sqrt{x}$ dla dowlonej liczby rzeczywistej $x>0$.

Szczególna rolę w zbiorze liczb zespolonych pełni liczba $i=(0,1)\in\mathbb{C}$, którą z powodów historycznych nazywamy *jednostka urojoną*. Obliczmy jej kwadrat

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

gdyż jakjuż wiemy liczby postaci $(x,0)\in\mathbb{C}$ możemy utożsamić z liczbami rzeczywistymi $x\in\mathbb{R}$. Oznacza to, że równanie $z^2+1=0$ w zbiorze $\mathbb{C}$ ma rozwiązanie $z=i$ (oczywiście drugim rozwiązaniem jest $z=-i$). Co więcej, łatwo sprawdzić, że są to jedyne rozwiązania, gdyż zachodzi tożsamość

$$z^2+1=(z-i)(z+i),$$

którą sprawdzamy bezpośrednim rachunkiem:

$$(z-i)(z+i)=z^2+zi-iz-i^2=z^2-(-1)=z^2+1.$$

Używając zatem „szkolnego" oznaczenia na pierwiastek kwadratowy możemy napisać: $\sqrt{-1}=i$. Jest to jak najbardziej *realna* równość, tyle że zachodzi w „szerszej" strukturze algebraicznej – w zbiorze zespolonych $\mathbb{C}$.

Postać algebraiczna i sprzężenie liczby zespolonej

Z formalnej definicji liczb zespolonych wiemy, że są to pary $(x,\ y)$ liczb rzeczywistych. W praktyce (zwłaszcza rachunkowej) posługiwanie się takim zapisem nie jest wygodne. Postać algebraiczna takiej pary to wyrażenie $x+iy$ dla $x,y\in\mathbb{R}$. Postać ta wynika z przyjętych działań w $\mathbb{C}$. Mamy bowiem

$$\begin{aligned}z&=(x,y)=(x,0)+(0,y)=(x,0)+(0,1)(y,0)=\\&=x+iy,\end{aligned}$$

gdzie wykorzystaliśmy utożsamienie liczb $(x,0)$, $(y,0)$ z liczbami rzeczywistymi $x,\ y$.

Posługiwanie się taką reprezentacją liczb zespolonych niezwykle ułatwia wykonywanie standardowych rachunków (mnożenie, dzielenie, dodawanie). Należy tylko pamiętać, że $i^2=-1$. Niech $z_1=2+5i,\ z_2=3-2i$. Wtedy mamy

$$\begin{aligned}&z_1+z_2=2+5i+3-2i=5+3i,\\&z_1\cdot z_2=(2+5i)\cdot(3-2i)=6-4i+15i-10i^2=6+11i-10\cdot(-1)=16+11i,\\&\frac{z_1}{z_2}=\frac{2+5i}{3-2i}=\frac{(2+5i)(3+2i)}{(3-2i)(3+2i)}=\frac{6+4i+15i+10i^2}{3^2-(2i)^2}=\frac{6+19i+10(-1)}{9-4(-1)}=\frac{-4+19i}{13}=\frac{-4}{13}+\frac{19}{13}i.\end{aligned}$$

DEFINICJA. Sprzężeniem liczby zespolonej $z=(x,y)=x+iy$ nazywamy liczbę zespoloną $\overline{z}$ określoną wzorem $\overline{z}=(x,-y)=x-iy$.

Geometrycznie liczba sprzężona do z na płaszczyźnie zespolonej $\mathbb{C}=\mathbb{R}^2$ jest jej obrazem w symetrii względem osi *OX*.

FAKT. (własności sprzężenia zespolonego)

Niech $z_1,\, z_2 \in \mathbb{C}$. Wtedy zachodzą równości:

$$\overline{z_1+z_2}=\overline{z}_1+\overline{z}_2, \qquad \overline{z_1-z_2}=\overline{z}_1-\overline{z}_2, \qquad \overline{z_1 z_2}=\overline{z}_1\overline{z}_2,$$

$$\overline{\left(\frac{z_1}{z_2}\right)}=\frac{\overline{z}_1}{\overline{z}_2}, \qquad \overline{\overline{z}}=z, \qquad z+\overline{z}=2\operatorname{Re} z,\ \ z-\overline{z}=2\operatorname{Im} z.$$

Moduł liczby zespolonej

DEFINICJA. *Modułem liczby zespolonej* $z=(x,y)=x+iy$ nazywamy liczbę rzeczywistą $|z|$ określoną wzorem $|z|=\sqrt{x^2+y^2}$.

Z definicji wynika, ze moduł liczby zespolonej jest nieujemny: $|z|\ge 0$. Ponadto $|z|=0 \Leftrightarrow z=0$. Jeżeli $z\neq 0$, to $|z|>0$.

Moduł liczby zespolonej jest uogólnieniem pojęcia wartości bezwzględniej liczy rzeczywistej. Geometrycznie moduł $|z|$ jest równy odległości liczby $z\in\mathbb{C}$ od początku układu współrzędnych na płaszczyźnie.

Dla dwóch liczb zespolonych $z_1,\, z_2 \in \mathbb{C}$ mamy $z_1-z_2=x_1+iy_1-(x_2+iy_2)=(x_1-x_2)+i(y_1-y_2)$, więc

$$|z_1-z_2|=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2},$$

co oznacza, że moduł różnicy, $|z_1-z_2|$, jest równy *odległości pomiędzy punktami* z_1, z_2 na płaszczyźnie zespolonej.

FAKT. (własności modułu liczby zespolonej).

Niech $z,\, z_1,\, z_2 \in \mathbb{C}$. Wtedy mamy

1) $|\overline{z}|=|z|=|-z|$,
2) $z\overline{z}=|z|^2$,
3) $|z_1 z_2|=|z_1||z_2|$,
4) $\left|\dfrac{z_1}{z_2}\right|=\dfrac{|z_1|}{|z_2|}$,
5) $|z_1+z_2|\le|z_1|+|z_2|$,
6) $\big||z_1|-|z_2|\big|\le|z_1-z_2|$,
7) $|\operatorname{Re} z|\le|z|,\ |\operatorname{Im} z|\le|z|$.

DOWÓD.

1) wynika wprost z definicji.

2) Niech $z=x+iy$. Wtedy: $z\overline{z}=(x+iy)(x-iy)=x^2-(iy)^2=x^2-(-1)y^2=x^2+y^2=|z|^2$.

3) Niech $z_1=x_1+iy_1,\ z_2=x_2+iy_2$. Wtedy $z_1z_2=x_1x_2-y_1y_2+i(x_1y_2+x_2y_1)$ więc:

$$
\begin{aligned}
|z_1z_2|^2 &= (x_1x_2-y_1y_2)^2+(x_1y_2+x_2y_1)^2=(x_1x_2)^2-2x_1x_2y_1y_2+(y_1y_2)^2+(x_1y_2)^2+2x_1y_2x_2y_1+(x_2y_1)^2\\
&=(x_1x_2)^2+(y_1y_2)^2+(x_1y_2)^2+(x_2y_1)^2=x_1^2(x_2^2+y_2^2)+y_1^2(x_2^2+y_2^2)\\
&=(x_1^2+y_1^2)(x_2^2+y_2^2)=|z_1|^2|z_2|^2=(|z_1||z_2|)^2,
\end{aligned}
$$

skąd $|z_1z_2|=|z_1||z_2|$, gdyż moduł jest nieujemny.

4) Z punktu 2) mamy $z^{-1}=\bar{z}/|z|^2$, biorąc teraz moduł obu stron tej równości otrzymujemy $|z^{-1}|=|\bar{z}|/|z|^2=|z|/|z|^2=1/|z|$. Korzystając dodatkowo z punktu 3) mamy

$$
\left|\frac{z_1}{z_2}\right|=|z_1z_2^{-1}|=|z_1||z_2^{-1}|=|z_1||z_2|^{-1}=\frac{|z_1|}{|z_2|}.
$$

5) W dowodzie wykorzystamy nierówność (wariant nierówności Cauchy'ego-Schwarza): $x_1x_2+y_1y_2\le\sqrt{x_1^2+y_1^2}\sqrt{x_1^2+y_1^2}$. Mamy więc

$$
\begin{aligned}
&|z_1+z_2|^2=(x_1+x_2)^2+(y_1+y_2)^2=x_1^2+2x_1x_2+x_2^2+y_1^2+2y_1y_2+y_2^2=\\
&=x_1^2+x_2^2+y_1^2+y_2^2=|z_1|^2+|z_2|^2+2(x_1x_2+y_1y_2)\le|z_1|^2+|z_2|^2+2\sqrt{x_1^2+y_1^2}\sqrt{x_1^2+y_1^2}=\\
&=|z_1|^2+|z_2|^2+2|z_1||z_2|=(|z_1|+|z_2|)^2,
\end{aligned}
$$

czyli $|z_1+z_2|^2\le(|z_1|+|z_2|)^2$, skąd $|z_1+z_2|\le|z_1|+|z_2|$.

6) W dowodzie tej nierówności skorzystamy z udowodnionego już punktu 5):

$$
\begin{aligned}
&|z_1|=|z_1-z_2+z_2|\le|z_1-z_2|+|z_2| &&\Rightarrow\quad |z_1|-|z_2|\le|z_1-z_2|\\
&|z_2|=|z_2-z_1+z_1|\le|z_2-z_1|+|z_1|=|z_1-z_2|+|z_1| &&\Rightarrow\quad -|z_1-z_2|\le|z_1|-|z_2|
\end{aligned}
$$

zatem $||z_1|-|z_2||\le|z_1-z_2|$.

Postać trygonometryczna liczby zespolonej

Jak już wiemy liczby zespolone możemy reprezentować jako punkty na płaszczyźnie zespolonej $\mathbb{C}$, która geometrycznie jest tożsama z płaszczyzną $\mathbb{R}^2$. Istotnym dodatkiem do tej struktury jest mnożenie takich punktów (lub wektorów) zdefiniowane wzorem (0.2). Jednak standardowa reprezentacja liczby zespolonej $z=(x,y)=x+iy\in\mathbb{C}$, gdzie $x,y\in\mathbb{R}$ są współrzędnymi kartezjańskim punktu nie zawsze jest wygodna. Czasami wygodna jest reprezentacja odpowiadająca współrzędnym biegunowym: odległość $r\ge 0$ od środka $(0,0)$ oraz kąt $0\le\varphi<2\pi$ pomiędzy osią *OX* a półprostą przechodzącą przez środek i liczbę $z=(x,y)$.

Jeżeli $z=x+iy\neq 0$, to istnieje dokładnie jedna liczba $0\le\varphi<2\pi$ taka, że

$$
\cos\varphi=\frac{x}{|z|},\quad \sin\varphi=\frac{y}{|z|}. \tag{0.4}
$$

Tak określoną liczbę będziemy nazywać *argumentem głównym* liczby zespolonej $z \neq 0.$ Dla wygody i jednoznaczności przyjmujemy dla $z = 0$ argument $\varphi = 0.$ Tak więc mamy $x = |z|\cos\varphi,\quad y = |z|\sin\varphi,$

co daje

$$z = x + iy = r(\cos\varphi + i\sin\varphi), \tag{0.5}$$

gdzie $r = |z| \geq 0.$ Ten sposób przedstawienia liczby zespolonej nazywamy *postacią trygonometryczną*.

Jeżeli dopuścimy w reprezentacji trygonometrycznej (0.5) argumenty dowolne $\varphi \in \mathbb{R}$ (bez ograniczenia się do przedziału $[0, 2\pi)$), to przedstawienie (0.5) nie jest jednoznaczne, ale różne argumenty będą różniły się o całkowitą wielokrotność liczby 2π :

$$z = r(\cos\varphi_1 + i\sin\varphi_1) = r(\cos\varphi_2 + i\sin\varphi_2) \;\Rightarrow\; \varphi_2 - \varphi_2 = 2k\pi \text{ dla pewnego } k \in \mathbb{Z}.$$

Przykład. Podane liczby zespolone zapiszemy w postaci trygonometrycznej.

a) $z = 1$: $z = 1\cdot(\cos 0 + i\sin 0) = \cos 0 + i\sin 0;$

b) $z = -1$: $z = 1\cdot(\cos\pi + i\sin\pi) = \cos\pi + i\sin\pi;$

c) $z = 1 + i : z = |z|(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}) = \sqrt{1^2 + 1^2}(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}) = \sqrt{2}(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}).$

d) $z = 2 + i\frac{2\sqrt{3}}{3} : |z| = \sqrt{2^2 + \left(\frac{2\sqrt{3}}{3}\right)^2} = \sqrt{4 + \frac{4\cdot 3}{9}} = \sqrt{4 + \frac{4}{3}} = \sqrt{\frac{16}{3}} = \frac{4}{\sqrt{3}} = \frac{4\sqrt{3}}{3}.$ Ponadto mamy

$$\cos\varphi = \frac{x}{|z|} = \frac{2}{\frac{4\sqrt{3}}{3}} = \frac{\sqrt{3}}{2},\quad \sin\varphi = \frac{y}{|z|} = \frac{\frac{2\sqrt{3}}{3}}{\frac{4\sqrt{3}}{3}} = \frac{1}{2} \;\Rightarrow\; \varphi = \frac{\pi}{6},$$

zatem $z = \frac{4\sqrt{3}}{3}(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}).$

Mnożenie i dzielenie liczb zespolonych zapisanych w postaci trygonometrycznej: niech $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1),\; z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2).$ Wtedy mamy

$$\begin{aligned} z_1 z_2 &= r_1(\cos\varphi_1 + i\sin\varphi_1) r_2(\cos\varphi_2 + i\sin\varphi_2) = r_1 r_2 \big((\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2) \\ &+ i(\sin\varphi_1\cos\varphi_2 + \cos\varphi_1\sin\varphi_2)\big) = r_1 r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)), \end{aligned}$$

gdzie skorzystaliśmy ze wzorów na kosinus sumy oraz na sinus sumy. Tak więc przy mnożeniu liczb zespolonych moduły się mnożą, a argumenty dodają. Przy dzieleniu argumenty będą się odejmowały

$$\frac{z_1}{z_2}=\frac{r_1(\cos\varphi_1+i\sin\varphi_1)}{r_2(\cos\varphi_2+i\sin\varphi_2)}=\frac{r_1}{r_2}\frac{(\cos\varphi_1+i\sin\varphi_1)(\cos\varphi_2-i\sin\varphi_2)}{(\cos\varphi_2+i\sin\varphi_2)(\cos\varphi_2-i\sin\varphi_2)}$$
$$=\frac{r_1}{r_2}\frac{\cos\varphi_1\cos\varphi_2+\sin\varphi_1\sin\varphi_2+i(\sin\varphi_1\cos\varphi_2-\cos\varphi_1\sin\varphi_2)}{(\cos\varphi_2)^2-(i\sin\varphi_2)^2}$$
$$=\frac{r_1}{r_2}\frac{\cos(\varphi_1-\varphi_2)+i\sin(\varphi_1-\varphi_2)}{(\cos\varphi_2)^2+(\sin\varphi_2)^2}=\frac{r_1}{r_2}\frac{\cos(\varphi_1-\varphi_2)+i\sin(\varphi_1-\varphi_2)}{1}$$
$$=\frac{r_1}{r_2}\big(\cos(\varphi_1-\varphi_2)+i\sin(\varphi_1-\varphi_2)\big).$$

Wzór de Moivre'a

Niech $z=\cos\varphi+i\sin\varphi$. Wtedy $z^2=z\cdot z=\cos(\varphi+\varphi)+i\sin(\varphi+\varphi)$, czyli $z^2=\cos 2\varphi+i\sin 2\varphi$. Podobnie $z^3=\cos 3\varphi+i\sin 3\varphi$. Ogólnie mamy: dla dowolnej liczby rzeczywistej $\varphi\in\mathbb{R}$ oraz liczby całkowitej $n\in\mathbb{Z}$ zachodzi

$$(\cos\varphi+i\sin\varphi)^n=\cos n\varphi+i\sin n\varphi. \tag{0.6}$$

Równość (0.6) nazywamy *wzorem de Moivre'a*.

Przykład. Doprowadzimy do najprostszej postaci liczbę zespoloną $z=\left(\frac{\sqrt{3}}{2}-\frac{1}{2}i\right)^{26}$. Postać trygonometryczna liczby z:

$$\frac{\sqrt{3}}{2}-\frac{1}{2}i=\cos\frac{11\pi}{6}+i\sin\frac{11\pi}{6}.$$

Ze wzoru de Moivre'a mamy:

$$\left(\cos\tfrac{11\pi}{6}+i\sin\tfrac{11\pi}{6}\right)^{26}=\cos(26\cdot\tfrac{11\pi}{6})+i\sin(26\cdot\tfrac{11\pi}{6})=\cos(\tfrac{143\pi}{3})+i\sin(\tfrac{143\pi}{3})=$$
$$\cos(\tfrac{5\pi+138\pi}{3})+i\sin(\tfrac{5\pi+138\pi}{3})=\cos(\tfrac{5\pi}{3}+46\pi)+i\sin(\tfrac{5\pi}{3}+46\pi)=\cos\tfrac{5\pi}{3}+i\sin\tfrac{5\pi}{3}=-\tfrac{\sqrt{3}}{2}+\tfrac{1}{2}i.$$

Pierwiastkowanie liczb zespolonoych

Niech $w\in\mathbb{C}$. Pierwiastkiem stopnia $n\in\mathbb{N}$ z liczby w nazywamy każdą liczbę zespoloną z spełniającą warunek

$$z^n=w. \tag{0.7}$$

Jest to definicja analogiczna do określenia pierwiastka rzeczywistego $x\in\mathbb{R}$ z liczby rzeczywistej $a\in\mathbb{R}$:

$$x^n=a.$$

W przypadku rzeczywistym nie zawsze istnieje pierwiastek (gdy wykładnik n jest parzysty oraz $a<0$). Ponadto gdy $a>0$, to mamy wtedy dwa pierwiastki (różniące się znakiem). W przypadku gdy wykładnik jest nieparzysty, to dla dowolnego $a\in\mathbb{R}$ istnieje dokładnie jeden pierwiastek stopnia n.

Wprzypadku zespolonym sytuacja jest zdecydowanie bardziej klarowna: każda liczba zespolona różna od zero ma dokładnie n różnych pierwiastków w zbiorze $\mathbb{C}$. Spróbujemy teraz znaleźć te pierwiastki.

Zapiszmy obie liczby występujace w równości (0.7) w postaci trygonometrycznej $z = |z|(\cos\psi + i\sin\psi)$, $w = |w|(\sin\varphi + i\sin\varphi)$. Podstawiając do (0.7) otrzymujemy

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

Stąd mamy $|z|^n = |w|$, $n\psi - \varphi = 2k\pi$, gdzie $k \in \mathbb{Z}$. Tak więc

$$|z| = \sqrt[n]{|w|}, \quad \psi_k = \frac{\varphi + 2k\pi}{n}, \ k \in \mathbb{Z}.$$

Jednak w ostnim wyrażeniu mamy tak napprawdę n różnych argumentów ψ_k dla $k = 0, \dots, n-1$, gdyż dla $k = n$ argument ψ_n różni się od ψ_0 o dokładnie 2π, czyli przedstawia tę samą liczbę z. Podobnie jest z pozostałymi wartościami $k > n$ oraz $k < 0$. Ostatecznie mamy następujące różne rozwiązania (pierwiastki):

$$z_k = \sqrt[n]{|w|}(\cos\frac{\varphi + 2k\pi}{n} + i\sin\frac{\varphi + 2k\pi}{n}), \ \text{dla } k = 0, 1, \dots, n-1. \tag{0.8}$$

Przykład. Podać pierwiastki stopnia 3 z liczby 1.

Mamy $w = 1 = 1 \cdot (\cos 0 + i\sin 0)$. Zatem $|w| = 1$, $\varphi = 0$, co po wstwieniu do (0.8) daje

$$z_k = \sqrt[3]{1}(\cos\frac{0 + 2k\pi}{3} + i\sin\frac{0 + 2k\pi}{3}), \ \text{dla } k = 0, 1, 2,$$

czyli

$$\begin{aligned} z_0 &= \cos 0 + i\sin 0 = 1, \\ z_1 &= \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} = -\frac{1}{2} + i\frac{\sqrt{3}}{2}, \\ z_2 &= \cos\frac{3\pi}{3} + i\sin\frac{3\pi}{3} = \cos\pi + i\sin\pi = -i. \end{aligned}$$

Ostatecznie w dziedzinie zespolonej mamy $\sqrt[3]{1} = \{1, \frac{1}{2} + i\frac{\sqrt{3}}{2}, -i\}$.

Wielomiany

Wielomianem o współczynnikach rzeczywistych nazywamy funkcję o następującej postaci:

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0, \tag{0.9}$$

gdzie $a_0, a_1, \dots, a_n \in \mathbb{R}$ są danymi liczbami.

Jeżeli $a_n \neq 0,$ to mówimy, że wielomian ma *stopień* $n.$ Stopień wielomianu p oznaczamy symbolem $\deg(p).$

Oczywiście mozna analogicznie zdefiniować wielomian zmiennej zespolonej o współczynnikach zespolonych:

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0, \tag{0.10}$$

gdzie $a_0, a_1, \ldots, a_n \in \mathbb{C}$ są danymi współczynnikami zespolonymi.

Ważnym problemem w teorii wielomianów jest znajdowanie miejsc zerowych (zwanych też pierwiastkami wielomianu), czyli rozwiązywanie równań postaci

$$a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0 = 0. \tag{0.11}$$

Miejscem zerowym (pierwiastkiem) wielomianu nazywamy taką liczbę $z_0 \in \mathbb{C},$ że:

$$a_n {z_0}^n + a_{n-1} {z_0}^{n-1} + \ldots + a_1 z_0 + a_0 = 0.$$

Jeżeli $a_n \neq 0,$ to równanie (0.11) nazywamy równaniem algebraicznym (czasami wielomianowym) *n*-tego stopnia. Jeżeli rozważamy wielomian rzeczywisty (0.9), to jego pierwiastkiem rzeczywistym będzie liczba $x_0 \in \mathbb{R}$ taka, że $p(x_0) = 0.$ Ponieważ jednak liczby rzeczywiste są podzbiorem liczb zespolonych, to na wielomian rzeczywisty (0.9) możemy też patrzyć jak na wielomian o współczynnikach zespolonych i szukać jego pierwiastków w tym szerszym zbiorze. Na przykład równanie algebraiczne

$$x^2 + 1 = 0,$$

ma współczynniki $a_0 = 1,\ a_1 = 1.$ Nie posiada ono rozwiązań w zbiorze $\mathbb{R},$ ale w zbiorze $\mathbb{C}$ istnieją dwa pieriwaski $z_1 = i$ oraz $z_2 = -i.$ Z kolei równanie

$$x^3 - 1 = 0,$$

posiada w $\mathbb{R}$ tylko jeden pierwiastek $x_1 = 1,$ ale w zbiorze $\mathbb{C}$ ma jeszcze dwa dodatkowe $z_2 = \frac{-1+i\sqrt{3}}{2},$ $z_3 = \frac{-1-i\sqrt{3}}{2}.$

W ogólnym przypadku im wyższy stopień równania, tym trudniej go rozwiązać. Z równaniami 1-ego i 2-ego stopnia spotkamy się w szkole. Są to równanie liniowe i kwadratowe:

- $a_1 x + a_0 = 0,$ gdzie $a_1 \neq 0$. Rozwiązanie takiego równania ma postać: $x = -\dfrac{a_0}{a_1}$

- $a_2 x^2 + a_1 x + a_0 = 0,$ gdzie $a_2 \neq 0.$ W tym przypadku jak wiadomo, liczba rozwiązań (a więc w szczególności to czy w ogóle istnieją) zależy od wartości współczynników $a_2, a_1, a_0 \in \mathbb{R}.$ Prostym sposobem analizy tego równania jest obliczenie tzw. wyróżnika równania kwadratowego (popularnie zwanego deltą – ze względu na powszechnie używane oznaczenie): $\Delta = a_1^{\ 2} - 4 \cdot a_2 \cdot a_0$. W zależności od tego czy wyróżnik jest ujemny, zerowy, czy dodatni mamy zero, jedno lub dwa rozwiązania rzeczywiste. Co więcej są znane proste wzory na te rozwiązania:

Jeżeli $\Delta \geq 0,$ to istnieją rozwiązania rzeczywistwe wyrażajace się wzorami

$$x_1 = \frac{-a_1 + \sqrt{\Delta}}{2a_2}, \quad x_2 = \frac{-a_1 - \sqrt{\Delta}}{2a_2}. \tag{0.12}$$

Jeżeli $\Delta \geq 0,$ to nie ma pierwiastków rzeczywistych, ale w dzidzinie zespolonej nadal wzory (0.12) są poprawne, gdyż w $\mathbb{C}$ istnieją pierwiastki z liczb ujemnych. Możemy wzory (0.12) zapisać wtedy nastepująco

$$z_1 = \frac{-a_1 + i\sqrt{|\Delta|}}{2a_2}, \quad z_2 = \frac{-a_1 - i\sqrt{|\Delta|}}{2a_2}. \tag{0.13}$$

Powstaje naturalne pytanie jak wyglądają rozwiązania równania 3-ego, 4-ego, 5-ego itd. stopnia. Czy można te rozwiązania wyrazić jakimiś „wzorami”, podobnymi do tych powyżej. Podobne – aczkolwiek bardziej skomplikowane – dla równań 3-ego i 4-ego stopnia znaleziono już w średniowieczu. Są to tzw. *wzory Cardano* (dla $n=3$) oraz *wzory Ferrari* (dla $n=4$). Są one raczej mało przydatne do znajdowania pierwiastków (na ogół w zastosowaniach wystarczy przybliżenie numeryczne). Aby zilustrować to przedstawiamy jeden z pierwiastków najprostrzego nietrywialnego równania trzeciego stopnia $x^3 + x + 1 = 0$ otrzymany ze wzorów Cardano:

$$x_1 = -\sqrt[3]{\frac{2}{3(-9+\sqrt{93})}} + \sqrt[3]{\frac{-9+\sqrt{93}}{18}}.$$

Przybliżona wartość numeryczna to $x_1 \approx -0.682328$. Równanie to posiada jeszcze dwa pierwiastki zespolone, których przybliżone wartości to $x_{2,3} \approx 0.341164 \pm 1.6154i$.

Natomiast, mimo dużego wysiłku wielu matematyków, nie udawało się znaleźć ogólnych wzorów na pierwiastki dla równań stopnia 5-ego, 6-ego i wyższych. W XIX wieku okazało się, że nie mogli takich wzorów znaleźć, gdyż one po prostu *nie istnieją*. Oczywiście nie oznacza to, że nie można rozwiązać konkretnego równania, np. 5-ego stopnia. Wiadomo jednak, że nie można podać ogólnego wzoru, który wyrażałby rozwiązania takiego równania poprzez jego współczynniki $a_0, \ldots, a_n$ przy użyciu działań +, ·, –, : oraz operacji $\sqrt[n]{}$. Istnieją jednak metody numeryczne, które pozwalają lokalizować zera wielomianów z dowolną dokładnością.

Zasadnicze twierdzenie algebry

Jak już wiemy zbiór liczb zespolonych pozwala pierwiastkować dowolną liczbę (patrz (0.8)). Okazuje się, że ta własność jest szczgólnym przypadkiem innej fundamentalnej własnosci, która mówi, że w zbiorze $\mathbb{C}$ każdy wielomian ma miejsce zerowe.

Twierdzenie (Zasadnicze twierdzenie algebry). Każdy wielomian zespolony różny od stałej ma przynajmniej jeden pierwiastek zespolony.

Ciekwym wynikiem jest twierdzenie Gaussa-Lucasa, które charakteryzuje geometrycznie miejsca zerowe pochodnej $p'(z)$ wielomianu zespolonego w oparciu o miejsca zerowe j $p(z)$. Pamiętajmy, że zbiór liczb zespolonych $\mathbb{C}$ geometrycznie pokrywa się z płaszczyzną $\mathbb{R}^2$, a na liczby zespolone możemy patrzeć jak punkty na tej płaszczyźnie. W szczególności dla dowolnego zbioru $A \subset \mathbb{C}$ jest określona otoczka wypukła $\mathrm{conv}(A)$, jako najmnejszy zbiór wypukły, który zawiera zbiór A. Zbiór wypukły to taki, który dla każdej pary punktów z tego zbioru zawiera także odcinek łączacy te punkty. W przypadku gdy zbiór A jest skończony, to otoczka wypukła jest wielokątem.

Twierdzenie (Gaussa-Lucasa). Jeżeli $p(z)$ jest wielomianem zmiennej zespolonej $z \in \mathbb{C}$, to wszystkie pierwiastki pochodnej $p'(z)$ tego wielomianu należą do otoczki wypukłej zbioru pierwiastków samego wielomianu $p(z)$.

Dowód. Niech $n \geq 1$ oznacza stopień wielomianu $p(z) = a_n z^n + \ldots + a_1 z + a_0$. Z zasadniczego twierdzenia algebry mamy pierwiastki $z_1, \ldots, z_n$ (niekonieczne różne). Możemy więc wielomian przedstawić w postaci iloczynu

$$p(z) = a_n (z - z_1) \cdot \ldots \cdot (z - z_n).$$

Obliczamy pochodną korzystając wielokrotnie z wzoru na pochodną iloczynu $(fg)' = f'g + fg'$:

$$p'(z) = a_n \big((z - z_2) \cdot \ldots \cdot (z - z_n) + (z - z_1)(z - z_3) \cdot \ldots \cdot (z - z_n) + \ldots + (z - z_1) \cdot \ldots \cdot (z - z_{n-1}) \big).$$

Dzielimy teraz $p'(z)$ przez $p(z)$

$$\frac{p'(z)}{p(z)} = \frac{a_n \big((z - z_2) \cdot \ldots \cdot (z - z_n) + (z - z_1)(z - z_3) \cdot \ldots \cdot (z - z_n) + \ldots + (z - z_1) \cdot \ldots \cdot (z - z_{n-1}) \big)}{a_n (z - z_1) \cdot \ldots \cdot (z - z_n).} =$$
$$= \frac{(z - z_2) \cdot \ldots \cdot (z - z_n) + (z - z_1)(z - z_3) \cdot \ldots \cdot (z - z_n) + \ldots + (z - z_1) \cdot \ldots \cdot (z - z_{n-1})}{(z - z_1) \cdot \ldots \cdot (z - z_n).} =$$
$$= \frac{1}{z - z_1} + \frac{1}{z - z_2} + \ldots + \frac{1}{z - z_n}.$$

Niech w będzie dowolnym miejscem zerowym pochodnej: $p'(w) = 0$. Jeżeli $w \in \{z_1, \ldots, z_n\}$, to oczywiście w należy do otoczki wypukłej zbioru $\{z_1, \ldots, z_n\}$. Niech więc $w \notin \{z_1, \ldots, z_n\}$ tj. $p(w) \neq 0$. Podstawiamy do powyższej zależności

$$\frac{p'(w)}{p(w)} = 0 \ \Rightarrow \ \frac{1}{w - z_1} + \frac{1}{w - z_2} + \ldots + \frac{1}{w - z_n} = 0.$$

Dla dowolnej liczby zespolonej mamy $z\overline{z} = |z|^2$, więc jeżeli $z \neq 0$, to $z^{-1} = \overline{z} / |z|^2$. Stosując tą równość do liczb $(w - z_k)^{-1}$ otrzymamy

$$\frac{\overline{w}-\overline{z}_1}{|w-z_1|^2}+\frac{\overline{w}-\overline{z}_2}{|w-z_2|^2}+\ldots+\frac{\overline{w}-\overline{z}_n}{|w-z_n|^2}=0,$$

co po wzięciu sprzężenia zespolonego obu stron daje

$$\sum_{k=1}^{n}\frac{w-z_k}{|w-z_k|^2}=0.$$

Rozbijamy na dwie sumy

$$\left(\sum_{k=1}^{n}\frac{1}{|w-z_k|^2}\right)w=\sum_{k=1}^{n}\frac{1}{|w-z_k|^2}z_k \quad \text{czyli} \quad w=\sum_{k=1}^{n}t_k z_k \quad \text{gdzie} \quad t_k=\left(\frac{1}{|w-z_k|^2}\right)\Big/\left(\sum_{k=1}^{n}\frac{1}{|w-z_k|^2}\right).$$

Ponieważ $t_k \geq 0$ oraz $\sum_{k=1}^{n} t_k = 1,$ więc w należy do otoczki wypukłej zbioru $\{z_1,\ldots,z_n\}$.

Zad. 1. Rozwiązać poniższe równania liniowe (1-ego stopnia). Niektóre są z parametrem.

a) $4x+15=1-x$

b) $-5x+8-5ax=14-2x$

c) $12+3x+2a=-2ax+5$

d) $(2+a)x-3=5-(-4-2a)x$

Zad. 2. Rozwiązać poniższe równania kwadratowe (2-ego stopnia).

a) $x^2+2x-3=0$

b) $(x-1)(x+2)=10$

c) $2x^2+0{,}3x-0{,}9=0$

d) $-x^2+5x+3=0$

Rozkład wielomianu na czynniki

Problem ten jest ściśle związany z zagadnieniem znajdowania pierwiastków wielomianu. Jeżeli dany wielomian $W(x)$ może być przedstawiony w postaci iloczynu wielomianów $P(x)$ i $Q(x)$, których stopnie są mniejsze od stopnia wielomianu $W(x)$, to jest to przykład rozkładu wielomianu $W(x)$ na czynniki:

$$W(x)=P(x)\cdot Q(x), \quad \text{gdzie} \quad \deg(P)<\deg(W) \ \text{i} \ \deg(Q)<\deg(W).$$

Przykłady

$x^2-1=(x-1)(x+1)$

$x^2+3x-10=(x-2)(x+5)$

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

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

$x^4-1=(x-1)(x+1)(x^2+1)$

Zauważmy, że czynniki x^2+1 I x^2+x+1 z przykładu nie są już dalej rozkładalne. Gdyby tak było to to, że wielomiany te mjałyby rzeczywiste pierwiastki, bo np. $x^2+1=(x+\alpha)(x+\beta)$ oznaczałoby, że $-\alpha$ oraz $-\beta$ są pierwaistkami. Z drugiej stronu wiemy, że równanie $x^2+1=0$ nie ma rzeczywistych pierwiastków. Gdyby jednak rozważać te wielomiany w zbiorze liczb zespolonych, to można je rozłożyć aż do czynników pierwszego stopnia

$$z^3-1=(z-1)(z-\tfrac{-1+i\sqrt{3}}{2})(z-\tfrac{-1-i\sqrt{3}}{2}),$$
$$z^4-1=(z-1)(z+1)(z-i)(z+i).$$

Zad. 3. Poniżej podane są w dwóch kolumnach wielomiany. W lewej są zapisane w postaci nierozłożonej, a w prawej w postaci iloczynu czynników niższego stopnia. Należy je połączyć w pary.

a) x^4-1

b) $x^3-6x^2+12x-8$

c) $2x^3-x^2+2x-1$

d) $x^6-x^4+x^2-1$

a) $(x^2-1)(x^4+1)$

b) $(2x-1)(x^2+1)$

c) $(x-1)(x^3+x^2+x+1)$

d) $(x-2)(x-2)(x-2)$

Jednym z podstawowych problemów w teorii wielomianów jest zagadnienie znalezienia rozkładu danego wielomianu na iloczyn czynników, które nie dadzą się już rozłożyć. Przypomina to nieco zagadnienie rozkładu liczb naturalnych na czynniki pierwsze.[3]

Przykładami wielomianów, które nie dadzą się rozłożyć (w zbiorze liczb rzeczywistych $\mathbb{R}$) są:

$x^2+1, \quad x+2$ (wielomian 2-ego i 1-ego stopnia).

[3] Przypomnijmy, że liczbę $p\in\mathbb{N}$ nazywamy liczbą pierwszą jeżeli $p>1$ oraz jedynymi naturalnymi dzielnikami tej liczby są 1 i p. Przykładami liczb pierwszych są $2, 3, 5, 7, 11, 13, 17$. Już w starożytności matematycy Greccy udowodnili, że liczb pierwszych jest nieskończenie wiele. Okazuje się, że każdą liczbę naturalną $n\in\mathbb{N}$ można rozłożyć na czynniki pierwsze, czyli na iloczyn, w którym występują tylko liczby pierwsze, $n=p_1\cdot p_2\cdot\ldots\cdot p_k$, gdzie p_i są liczbami pierwszymi. Rozkład ten jest jednoznaczny z dokładnością co do kolejności czynników. Na przykład $10=2\cdot 5$, $12=2\cdot 2\cdot 3$, $2835=5\cdot 7\cdot 9\cdot 9$. Ciekawym zagadnieniem w arytmetyce liczb naturalnych jest rozmieszczenie liczb pierwszych w zbiorze $\mathbb{N}$. Okazuje się, że jest ono nierównomierne, a jeden z podstawowych rezultatów mówi, że udział procentowy liczb pierwszych w odcinku $\{1,2,\ldots,n\}$ maleje do zera. Znane są dokładne wyrażenie na ten udział procentowy, a najprostsze z nich ma postać *(ilość liczb pierwszych* $\leq n$ *)* $/n\approx 1/\ln n$. Poza czystą matematyką liczby pierwsze odgrywają bardzo ważną rolę w systemach kryptografii z kluczem publicznym. Na przykład fundamentem bezpieczeństwa szyfrowania metodą **RSA** jest algorytmiczna trudność rozkładu dużych liczb na czynniki pierwsze.

Okazuje się, że każdy wielomian rzeczywisty $W(x)$ można rozłożyć na czynniki liniowe i kwadratowe.

TWIERDZENIE. Każdy wielomian rzeczywisty da się rozłożyć na iloczyn wielomianów co najwyżej drugiego stopnia. Dokładniej, jeżeli $W(x)$ jest dowolnym wielomianem o współczynnikach rzeczywistych, to istnieją wielomiany rzeczywiste $P_1(x),\ldots,P_m(x)$ takie, że

$$W(x)=P_1(x)\cdot\ldots\cdot P_m(x), \tag{0.14}$$

oraz $\deg(P_i)=1$ lub $\deg(P_i)=2.$

DOWÓD. Twierdzenie to jest w gruncie rzeczy wnioskiem z **Zasadniczego Twierdzenia Algebry**. Jeżeli wielomian $W(x)=a_nx^n+a_{n-1}x^{n-1}+\ldots+a_1x+a_0$ ma współczynniki rzeczywiste, to i takmożemy „traktować” go jako wielomian zespolony $\tilde{W}(z)=a_nz^n+a_{n-1}z^{n-1}+\ldots+a_1z+a_0$. Formalnie oznaczo to, żedefiniujemy wielomian $\tilde{W}(z)$ dla $z\in\mathbb{C}$, ale o tych samych współczynnikach $a_0,\ldots,a_n\in\mathbb{R}\subset\mathbb{C}.$ Mamy więc pierwiastki $z_1,\ldots,z_n$ wielomianu $\tilde{W}(z)$, co prowadzi do rozkładu

$$\tilde{W}(z)=a_n(z-z_1)\cdot\ldots\cdot(z-z_n)\ \text{ dla } z\in\mathbb{C}. \tag{0.15}$$

Z drugiej strony wielomian $\tilde{W}(z)$ ma współczynniki rzeczywiste, co oznacza, że jeżeli jakaś liczba $w\in\mathbb{C}$ jest jegopierwiastkiem, to również liczba sprzężona $\overline{w}$ jest pierwiastkiem:

$$\tilde{W}(w)=a_nw^n+a_{n-1}w^{n-1}+\ldots+a_1w+a_0=0 \Rightarrow \overline{a_nw^n+a_{n-1}w^{n-1}+\ldots+a_1w+a_0}=0$$

czyli

$$\begin{aligned}0&=\overline{a_nw^n}+\overline{a_{n-1}w^{n-1}}+\ldots+\overline{a_1w}+\overline{a_0}=\overline{a}_n\overline{w}^n+\overline{a_{n-1}}\overline{w}^{n-1}+\ldots+\overline{a}_1\overline{w}+\overline{a}_0=\\&=a_n\overline{w}^n+a_{n-1}\overline{w}^{n-1}+\ldots+a_1\overline{w}+a_0=0,\end{aligned}$$

gdzie skorzystaliśmy z własności sprzężenia zespolonego: $\overline{z_1+z_2}=\overline{z}_1+\overline{z}_2$, $\overline{z_1z_2}=\overline{z}_1\overline{z}_2$ oraz z faktu, że współczynniki a_k są rzeczywiste, czyli $\overline{a}_k=a_k$. Wszystkie pierwiastki $z_1,\ldots,z_n$ można podzielmy na dwie grupy:

1) pierwiastki, które są rzeczywiste $(z_k\in\mathbb{R})$;
2) pierwiastki, które nie są rzeczywiste $(z_k\in\mathbb{C}\setminus\mathbb{R})$;

Pierwiastki z drugiej grupy występują parami , gdyż jeżeli $z_k\in\mathbb{C}\setminus\mathbb{R}$ jest pierwiastkiem, to również $\overline{z}_k$ jest pierwiastkiem (co pokazaliśmy przed chwilą), a ponadto też $\overline{z}_k\in\mathbb{C}\setminus\mathbb{R}.$

Ustawny pierwiastki w takiej kolejności, aby na początku były rzeczywiste $(\in\mathbb{R})$, a następnie nierzeczywiste $(\in\mathbb{C}\setminus\mathbb{R})$, które będą występowały parami

$$\underbrace{x_1,\ldots,x_r}_{\in\mathbb{R}},\ \underbrace{z_1,\overline{z}_1,\ldots,z_s,\overline{z}_s}_{\in\mathbb{C}\setminus\mathbb{R}}.$$

Możemy teraz rozkład (0.15) zapisać następująco

$$\tilde{W}(z)=a_n(z-x_1)\ldots(z-x_r)(z-z_1)(z-\overline{z}_1)\ldots(z-z_s)(z-\overline{z}_s). \tag{0.16}$$

Zauważmy teraz, że

$$(z-z_k)(z-\overline{z}_k)=z^2-(z_k+\overline{z}_k)z+z_k\overline{z}_k=z^2-2\operatorname{Re}(z_k)z+|z_k|^2=z^2+p_kz+q_k,$$

gdzie $p_k=-2\operatorname{Re}z_k$, $q_k=|z_k|^2$ i oczywiście p_k, $q_k\in\mathbb{R}$. Zatem rozkład (0.16) ma postać

$$\tilde{W}(z)=a_n(z-x_1)\ldots(z-x_r)(z^2+p_1z+q_1)\ldots(z^2+p_sz+q_s).$$

Ostatecznie wystarczy teraz w tej równosci podstawić $z=x\in\mathbb{R}$:

$$W(x)=\tilde{W}(x)=a_n(x-x_1)\ldots(x-x_r)(x^2+p_1x+q_1)\ldots(x^2+p_sx+q_s),$$

co kończy dowód.

Przykłady

$$\begin{aligned}
&2x^3-x^2-x+1=(x+1)(2x^2-x+1),\\
&x^4+x^3-x-1=(x-1)(x+1)(x^2+x+1),\\
&2x^4-4x^3+7x^2-2x+3=(2x^2+1)(x^2-2x+3),\\
&x^4+x^3-3x^2-x+2=(x-1)(x-1)(x+1)(x+2)=(x-1)^2(x+1)(x+2).
\end{aligned}$$

Zauważmy, że rozkład (0.14) jest w równoważny znalezieniu pierwiastków (rzeczywistych) wielomianu $W(x)$. Wynika to stąd, że

$$W(x)=0\Leftrightarrow P_1(x)=0 \text{ lub } P_2(x)=0 \text{ lub } \ldots\ P_k(x)=0.$$

Czynniki $P_i(x)$ są liniowe lub kwadratowe. Jeżeli więc $\deg P_i=1$, to $P_i(x)=a_ix+b_i$, więc pierwiastek $x_i=-b_i/a_i$. Natomiast gdy $\deg P_i=2$, to wielomian jest nierozkładalny, więc nie ma pierwiastków rzeczywistych.

Zad. 4. Znaleźć rozkłady na czynniki nierozkładalne następujących wielomianów:

a) x^2+3x-4

b) x^3+x^2-4x-4

c) $x^4+3x^3-4x^2$

d) x^4-1

e) x^4+x^2+2x-4

f) x^3+3x^2-4x

Zad. 5. Rozwiązać w zbiorze liczb rzeczywistych następujące układy równań liniowych

$$\begin{cases}2x+4y=15\\6x-y=-1\end{cases}\qquad\begin{cases}-3x+0,5y=12\\6x+4y=-1\end{cases}\qquad\begin{cases}x+y=1\\3x-4y=1\end{cases}$$

Zad. 6. Dla jakich wartości parametru $m \in \mathbb{R}$ układ równań

$$\begin{cases} 3x-(m-1)y=3, \\ mx-2y=-2, \end{cases}$$

jest niejednoznaczny?

Potęga o wykładniku całkowitym. Niech $a, b \in \mathbb{R}$ oraz $n \in \mathbb{Z}$. Wtedy definiujemy

$$a^n = 1 \quad \text{dla} \quad n=0,$$
$$a^n = a \quad \text{dla} \quad n=1,$$
$$a^n = \underbrace{a \cdot \ldots \cdot a}_{n\ razy} \quad \text{dla} \quad n>1,$$
$$a^n = \frac{1}{a^{-n}} \quad \text{dla} \quad n<0.$$

Zad. 1. Niech $a, b \in \mathbb{R}$ będą dowolnymi liczbami rzeczywistymi, a $n, m \in \mathbb{Z}$ liczbami całkowitymi. Udowodnić następujące tożsamości:

$$a^n b^n = (ab)^n$$
$$a^n a^m = a^{n+m}$$
$$\frac{a^n}{a^m} = a^{n-m}$$
$$(a^n)^m = a^{nm}$$

Zad. 2. Zapisać następujące prawa arytmetyki liczb rzeczywistych:

- prawo przemienności dodawania i mnożenia
- prawo łączności dodawania i mnożenia
- prawo rozdzielności mnożenia względem dodawania

Zad. 3. Jakie są podstawowe związki łączące operacje dodawania i mnożenia z relacją mniejszości w zborze liczb rzeczywistych?

Zad. 4. Wartość bezwzględna liczby $x \in \mathbb{R}$ jest określona następująco:

$$|x| = \begin{cases} x & \text{dla } x \geq 0, \\ -x & \text{dla } x < 0. \end{cases}$$

Udowodnić następujące związki dla dowolnych liczb rzeczywistych $x, y \in \mathbb{R}.$:

$$|x+y| \leq |x| + |y|,$$
$$|x \cdot y| \leq |x| \cdot |y|.$$

Zad. 5. Rozwiązać poniższe nierówności w zbiorze liczb rzeczywistych (tzn. $x \in \mathbb{R}$):

$|2x+6|<8,$

$|x-1|+|3x+6|>2,$

$||2x+4|-3|x+5||>1+x,$

$|x|3-x|+|x^2-1|+4|<6.$

Symbol sumy: Σ (grecka duża litera sigma)

W matematyce często występują sumy pewnych wartości. Gdy składników sumy jest dużo, to możemy stosować różne pomocne zapisy. Np. sumę liczb naturalnych od *1* do *50* zapiszemy skrótowo: $1+2+3+\ldots+50$. Sumę wyrazów ciągu (a_n) od wyrazu 7-ego do *30*-tego możemy zapisać tak: $a_7+a_8+\ldots+a_{30}$. Okazuje się, że wielu zastosowaniach nawet ten zapis nie jest dostatecznie wygodny, więc wprowadzono jeszcze bardzie zwięzły i wygodniejszy sposób zapisywania sum. Jest kilka wariantów tego zapisu. Jeden z nich wygląda tak:

$$\sum_{k=1}^{n} a_k := a_1 + a_2 + \ldots + a_n.$$

Na dole symbolu sumy, Σ, podajemy wartość (indeks) pierwszego elementu sumy (tutaj: $k=1$), a u góry wartość ostatniego indeksy sumy (nie piszemy już $k=n$ tylko samo *n*). Symbol k, który występuje w tym przykładzie, nie jest tak bardzo istotny. Równie dobrze mogłaby to być inna litera (np. i, j, k – te się najczęściej stosuje).

A zatem:

$$\sum_{k=1}^{n} a_k = \sum_{i=1}^{n} a_i = \sum_{j=1}^{n} a_j = a_1 + a_2 + \ldots + a_n.$$

Oczywiście <u>nie jest konieczne</u>, aby suma była indeksowana od wartości 1. Może zaczynać się np. od liczby 7:

$$\sum_{k=7}^{100} a_k = a_7 + a_8 + \ldots + a_{100}.$$

Zad. 6. Zapisz przy pomocy symbolu sumy następujące wyrażenia:

a) $1+2+\ldots+99,$

b) $1+2+\ldots+n,$

c) $5+6+\ldots+88,$

d) $2+4+6+\ldots+200,$

e) $1+3+5+\ldots+101,$

f) $2^2+3^2+\ldots+n^2$,

g) $\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}$,

h) $x+x^2+x^3+\ldots+x^{100}$,

i) $1+x+x^2+x^3+\ldots+x^n$,

j) $\frac{1}{2}+\frac{2}{3}+\frac{3}{4}+\ldots+\frac{n}{n+1}$,

k) $1-2+3-4+5-6+\ldots+999-1000$,

l) $1-2+3-4+5-6+\ldots+2(n-1)-2n$,

m) $-\frac{1}{1}+\frac{1}{2}-\frac{1}{3}+\frac{1}{4}-\frac{1}{5}++\frac{1}{2n}-\frac{1}{2n+1}$.

n) $-\frac{1}{1}+\frac{1}{2}-\frac{1}{3}+\frac{1}{4}-\frac{1}{5}+\ldots+\frac{1}{2n}-\frac{1}{2(n+1)}$

Zad. 7. Ile składników sumy zawierają podane niżej wyrażenia? (symbol n wszędzie oznacza liczbę naturalną: $n\in\mathbb{N}$)

a) $\sum_{k=0}^{9}a_k$, $\sum_{k=1}^{9}a_k$.

b) $\sum_{i=7}^{21}i^3$, $\sum_{i=24}^{53}i^3$.

c) $\sum_{k=-3}^{5}b_k$, $\sum_{k=-4}^{7}k$.

d) $\sum_{k=1}^{n}a_k$, $\sum_{k=0}^{n}a_k$, $\sum_{k=0}^{n+2}a_k$.

e) $\sum_{k=-2}^{n+1}k^4$, $\sum_{k=-9}^{0}k^4$, $\sum_{k=-n}^{3}k^4$, $\sum_{k=-n}^{n}k^4$, $\sum_{k=-n+2}^{13}k^4$.

f) $\sum_{k=-2n}^{2n+1}a_k$, $\sum_{j=-2n-3}^{3n+2}a_j$.

Zad. 8. Oblicz wartość poniższych wyrażeń:

a) $\sum_{k=1}^{11}k$, $\sum_{k=6}^{13}(-1)^k$, $\sum_{k=1}^{1000}(-1)^k$, $\sum_{k=1}^{n}(-1)^k$, $\sum_{k=0}^{5}2^{-k}$

b) $\sum_{k=1}^{200}\left(\frac{1}{k}-\frac{1}{k+1}\right)$, $\sum_{j=1}^{n}\left(\frac{1}{j}-\frac{1}{j+1}\right)$, $\sum_{i=1}^{n}(-1)^{i+1}$, $\sum_{j=1}^{n}(-1)^{j}(2j+1)$.

Zad. 9. Udowodnić indukcyjnie następujące tożsamości:

- $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$
- $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$
- $\sum_{k=1}^{n} k^3 = \left(\sum_{k=1}^{n} k\right)^2$
- $1+3+5+\ldots+2n-1=n^2$ (zapisać także tą tożsamość przy pomocy symbolu sumy)
- $1^2+3^2+5^2+\ldots+(2n-1)^2=\frac{1}{3}n(4n^2-1)$ (zapisać także tą tożsamość przy pomocy symbolu sumy)
- $\sum_{k=1}^{n} k(k+1)=\frac{1}{3}n(n+1)(n+2)$
- $\sum_{k=0}^{n} a^k = \frac{a^{n+1}-1}{a-1}$ $\quad (n \geq 0,\ a \neq 1)$ (co się dzieje dla $a=1$?)

Zad. 10. Wzór dwumianowy Newtona pozwala zapisać w rozwiniętej formie wyrażenie $(a+b)^n$. Ma on następującą postać

$$(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}.$$

- Udowodnić następującą tożsamość dla współczynników Newtona

$$\binom{n}{k}=\binom{n-1}{k}+\binom{n-1}{k-1}.$$

- Pokazać $\sum_{k=0}^{n}\binom{n}{k}=2^n$. Uzasadnić algebraicznie (ze wzoru dwumianowego) i kombinatorycznie.

Zad. 11. Symbolem $|$ oznaczamy relacje podzielności w zbiorze liczb całkowitych. Tak więc zapis $a|b$ czytamy jako *„a dzieli b”* (lub *„b jest podzielne przez a”*). Udowodnić następujące twierdzenia dotyczące podzielności w zbiorze liczb całkowitych.

- $2|(n^2-n),\ 6|(n^3-n)$
- $30|(n^5-n),\ 42|(n^7-n)$
- $p|(n^p-n)$ gdzie p jest liczbą pierwszą
- $9|(10^n-1),\ 12|(10^n-4),\ 11|(10^n-(-1)^n).$

Przestrzeń wektorowa

Pojęciem *wektora* spotykamy w fizyce. Jest to wielkość fizyczna, która charakteryzuje się pewnymi specyficznymi cechami takimi jak: długość, kierunek i zwrot. Czasami jeszcze dochodzi punkt zaczepienia. Wielkość taką wygodnie wyobrażać sobie w postaci strzałki. Przykładami wielkości fizycznych wektorowych są: prędkość, siła czy natężenie pola elektrycznego.

Przestrzeń wektorowa stanowi matematyczną formalizację pojęcia wektora.

DEFINICJA. *Przestrzenią wektorową nad* K nazywamy czwórkę $(V,\ K,\ +,\ \cdot)$ gdzie V jest dowolnym niepustym zbiorem, K jest ciałem, a $+$ oraz $\cdot$ są działaniami dwuargumentowymi

$$+ : V\times V \to V,$$
$$\cdot\ : K\times V \to V,$$

które spełniają następujące warunki (dla dowolnych $\alpha,\beta\in K,\ u,\ \upsilon,\ w\in V$):

a) Zbiór V z działaniem $+$ (zwanym *dodawaniem*) jest grupą abelową, tzn.

a.1) $(u+\upsilon)+w=u+(\upsilon+w)$ (*łączność dodawania*);

a.2) $\exists\ \Theta\in V\ \forall\ u\in V : u+\Theta=u$ (istnienie elementu neutralnego dla dodawania, $\Theta-$ nazywać go będziemy *wektorem zerowym*);

a.3) $\forall\ u\in V\ \exists -u\in V : u+(-u)=\Theta$ (wektor $-u$ nazywamy wektorem przeciwnym do u);

a.4) $u+\upsilon=\upsilon+u$ (przemienność dodawania);

b) Działanie $\cdot$ (zwane *mnożeniem*) spełnia następujące warunki:

b.1) $\alpha\cdot(u+\upsilon)=\alpha\cdot u+\alpha\cdot\upsilon$

b.2) $(\alpha+\beta)\cdot u=\alpha\cdot u+\beta\cdot u$

b.3) $\alpha\cdot(\beta\cdot u)=(\alpha\beta)\cdot u$

b.4) $1\cdot u=u.$

Uwagi:

1) Elementy zbioru V będziemy nazywać *wektorami*, a elementy zbioru K *skalarami*. Ciało K w naszych zastosowaniach to będzie: albo zbiór liczb rzeczywistych $(K=\mathbb{R})$, albo zbiór liczb zespolonych $(K=\mathbb{C})$, albo zbiór liczb wymiernych $(K=\mathbb{Q})$.

2) Mnożenie wektora przez skalar oznaczamy kropką $(\cdot)$, ale czasami pomijamy tę kropkę (podobnie jak przy zapisie mnożenia liczb); tak więc jeżeli mamy skalar $\alpha \in K$ oraz wektor $\upsilon \in V$, to możemy pisać $\alpha \cdot \upsilon$ lub $\alpha\upsilon$.

3) Jak widzimy definicja przestrzeni wektorowej oznacza, że w zbiorze V jest określone dodawanie wektorów u, υ: $u+\upsilon$ oraz mnożenie przez skalary $\alpha \in K$: $\alpha \cdot \upsilon$ tak, aby zachodziły pewne „naturalne" własności.

4) Wektor zerowy, który w definicji oznaczyliśmy symbolem Θ na ogół będziemy jednak oznaczać zwykłym zerem, 0. Tak samo jest oznaczana liczba zero, $0 \in K$. Nie powinno to prowadzić do nieporozumień, gdyż z kontekstu wiadomo, o które „zero" chodzi. Na przykład w zapisie $0 \cdot u$ lub $0u$ mamy do czynienia ze skalarem 0, ale w zapisie $\alpha \cdot 0$ lub $\alpha 0$ symbol 0 to wektor zerowy.

5) W niektórych opracowaniach wektory wyróżnia się czcionka pogrubioną, np. $\mathbf{u}$, $\mathbf{E}$ lub dopisując strzałkę nad symbolem, np. $\vec{u}$, $\vec{E}$. Konwencje te często stosują fizycy.

Jednym z podstawowych przykładów przestrzeni wektorowej (rzeczywistej) jest przestrzeń $(\mathbb{R}^n, \mathbb{R})$ zwana $n-$wymiarową rzeczywistą przestrzenią kartezjańską. Zbiór wektorów $V=\mathbb{R}^n$ jest określony jako zbiór wszystkich $n-$elementowych ciągów liczb rzeczywistych

$$\mathbb{R}^n = \mathbb{R} \times \ldots \times \mathbb{R} = \{(x_1, \ldots, x_n) : x_1, \ldots, x_n \in \mathbb{R}\}.$$

Aby jednak zbiór $\mathbb{R}^n$ stał się przestrzenią wektorową, musimy jeszcze określić dodawanie elementów $x,\ y \in \mathbb{R}^n$ oraz mnożenie wektora $x \in \mathbb{R}^n$ przez skalar $\alpha \in \mathbb{R}$. Działania te określamy następująco

$$\begin{aligned} &(x_1, \ldots, x_n) + (y_1, \ldots, y_n) \coloneqq (x_1 + y_1, \ldots, x_n + y_n), \\ &\alpha \cdot (x_1, \ldots, x_n) \coloneqq (\alpha x_1, \ldots, \alpha x_n), \end{aligned}$$

dla dowolnych $x=(x_1, \ldots, x_n)$, $y=(y_1, \ldots, y_n) \in \mathbb{R}^n$ oraz $\alpha \in \mathbb{R}$. Mimo, że ściślej rzecz biorąc przestrzeń wektorowa to para $(\mathbb{R}^n, \mathbb{R})$, to w praktyce upraszczamy zapis i posługujemy się zapisem $\mathbb{R}^n$ zamiast pary $(\mathbb{R}^n, \mathbb{R})$.

Zad. 7. Narysować podane wektory w przestrzeni $\mathbb{R}^2$ i $\mathbb{R}^3$.

a) (2, 3), (-2, -3), (3, 5), (0,1), (1, 0), (-1, 0)

b) (0, 1, 1), (1, 1, 1), (-1, -1, -1)

Zad. 8. Niech $v, u, w \in \mathbb{R}^2$: $v = (1, 2), u = (-2, 0), w = (-1, -1)$. Wykonać podane niżej działania i podać ostateczną postać powstałych wektorów.

a) $v + u$

b) $v + u + w$

c) $2v$

d) $-w$

e) $-3w$

f) $2v + w$

g) $w - u$

h) $3v + 4w - u$

i) $(w + u)/2$

j) $(w - 2u)/2$

k) $-3w + u + 3w$

l) $v - 2w + 5w + 2u - 4u + v$

m) $0,5v + 1,2w - 3,4v + 0,9w$

n) $4(v + 2u + 3w) - 2(v - 2,5w) + 6w$

Zad. 9. Niech $v, u, w \in \mathbb{R}^3$: $v = (1, 2, 0), u = (-1, -2, 0), w = (2, 2, 3)$. Wykonać podane niżej działania na tych wektorach.

a) $v + u + w$

b) $2(v + w) + v + 3w$

c) $3(v + 2u + 2w) + 4v - 3w$

d) $(v + w + u)/3$

e) $4(-2v + 3w)/2 + w + v + w$

f) $-3(w + v - 2w + v - 4w + v) + 2(u + 3w)$

Zad. 10. Niech *V* będzie dowolną przestrzenią wektorową. Załóżmy, że mamy dane dwa wektory: a, $b \in V$. Znaleźć wektory *u* i *v* takie, że:

a) $\begin{cases} a = u + v \\ b = u - v \end{cases}$ $\quad$ $\begin{cases} 2u + v = a \\ -4u - 2v = b \end{cases}$ $\quad$ $\begin{cases} a + b = -2u + 4v \\ a - 2b = u + 3v \end{cases}$

Zad. 11. Dane są wektory $u, v, w \in \mathbb{R}^2$. Znaleźć skalary $\alpha, \beta, \gamma \in \mathbb{R}$ takie, że: $\alpha v + \beta u + \gamma w = 0$ dla następujący wektorów:

a) $u = (2, 4)$, $v = (-1, 0)$, $w = (0, 3)$

b) $u = (1, 2)$, $v = (2, 2)$, $w = (1, -1)$

c) $u = (2, 4)$, $v = (-1, 4)$, $w = (0, 0)$

Zad. 12. Dane są pary wektorów $u, v \in \mathbb{R}^2$. Znaleźć $\alpha, \beta \in \mathbb{R}$ takie, że: $\alpha u + \beta v = 0$. Zinterpretować wyniki geometrycznie (w układzie współrzędnych na płaszczyźnie) .

a)

$u = (1, 2)$	$u = (1, -1)$	$u = (1, 0)$
$v = (-1, 3)$	$v = (-2, 2)$	$v = (0, 1)$

Baza przestrzeni wektorowej

Jednym z ważniejszych pojęć w teorii przestrzeni wektorowych jest baza przestrzeni wektorowej. Mówiąc opisowo jest to taki podzbiór wektorów tej przestrzeni, że dowolny wektor przestrzeni można rozłożyć względem wektorów bazowych w postaci tzw. kombinacji liniowej, i to w sposób jednoznaczny. Aby wprowadzić pojęcie bazy wygodnie jest wprowadzić wcześniej dwa inne pojęcia: liniową niezależności oraz układ generatorów.

Mówimy, że zbiór wektorów $u_1,\ldots,u_m$ jest *liniowo niezależny*, gdy zachodzi następująca własność wyrażona w formie implikacji:

$$\lambda_1 u_1 + \ldots + \lambda_m u_m = 0 \;\Rightarrow\; \lambda_1 = 0,\ldots,\lambda_n = 0.$$

Mówimy, że zbiór wektorów $u_1,\ldots,u_m$ jest *układem generatorów* (układem generującym) przestrzeń wektorowa V, gdy

$$\forall\, x \in V\ \exists\ \lambda_1,\ldots,\lambda_m \in K \text{ takie, że } x = \lambda_1 u_1 + \ldots + \lambda_m u_m.$$

Mówimy, że zbiór wektorów $u_1,\ldots,u_m$ jest *bazą przestrzeni* wektorowej, jeżeli jednocześnie jest liniowo niezależny i jest układem generatorów.

Współrzędne wektora względem bazy

Dzięki wybraniu jakiejś bazy w przestrzeni wektorowej, możemy wektory opisywać przy pomocy współrzędnych. Należy wyraźnie podkreślić, że współrzędne zależą od wyboru konkretnej bazy. Oznacza to, że ten sam wektor będzie miał na ogół różne współrzędne w różnych bazach.

Zad. 1. Udowodnić, że współrzędne danego wektora są w danej bazie wyznaczone jednoznacznie. Dokładniej: niech V będzie rzeczywistą (nad ciałem $\mathbb{R}$) lub zespoloną (nad ciałem $\mathbb{C}$) przestrzenią wektorową, $b_1,\ldots,b_n$ pewną bazą w V oraz $x \in V$. Wtedy zachodzi

$$x = \lambda_1 b_1 + \ldots + \lambda_n b_n \text{ oraz } x = \mu_1 b_1 + \ldots \mu_n b_n \;\Rightarrow\; \lambda_1 = \mu_1,\ldots,\lambda_n = \mu_n.$$

Podsumujmy zatem podstawową własność bazy przestrzeni wektorowej

---

Niech $B = \{b_1,\ldots,b_n\}$ będzie bazą przestrzeni wektorowej V. Wtedy każdy wektor $x \in V$ ma jednoznaczne przedstawienie w postaci sumy:

$$x = \lambda_1 b_1 + \ldots + \lambda_n b_n = \sum_{i=1}^{n} \lambda_i b_i.$$

Liczby $\lambda_1,\ldots,\lambda_n$ nazywamy *współrzędnymi* wektora x w bazie B.

Zad. 2. Wyznaczyć współrzędne podanych wektorów w różnych bazach.

1) $v=(2,\ 5)$

- baza: $b_1=(2,\ 3),\ b_2=(1,\ 1)$
- baza: $b_1=(2,\ 4),\ b_2=(-1,\ 2)$

2) $v=(3,\ 2,\ 1)$

- baza: $b_1=(1,\ 0,\ 2),\ \ b_2=(-1,\ 2,\ 3),\ \ b_3=(-2,\ 2,\ 0)$
- baza: $b_1=(2,\ 2,\ 2),\ \ b_2=(0,\ 2,\ 3),\ b_3=(4,\ \ 5,-5)$

Zad. 3. Jak wiemy w przestrzeni $\mathbb{R}^n$ istnieje pewna szczególna, wyróżniona baza składająca się z następujących wektorów: $e_1=(1,0,\ldots,0),\ e_2=(0,1,\ldots,0),\ldots,e_n=(0,0,\ldots,1)$. Jest to tzw. *baza kanoniczna*. Jak wyglądają współrzędne wektora w bazie kanonicznej?

Zad. 4. Wyznaczyć współrzędne wektora $v=(x,y,z)\in\mathbb{R}^3$ w następujących bazach:

a) $(1,\ 0,\ 2),\ (-1,\ 2,\ 3),\ (3,\ 2,\ 1)$

b) $(-1,\ -2,\ -1),\ (0,\ 0,\ 1),\ (3,\ 2,\ 1)$

Podać współrzędne $\lambda_1,\ \lambda_2,\ \lambda_3$ jako funkcje składowych $x,\ y,\ z$ dowolnego wektora $v=(x,y,z)$.

Zad. 5. Wyznaczyć współrzędne podanych wektorów z $\mathbb{R}^n$ w następującej bazie:

$$b_1=(1,0,\ldots,0),\ b_2=(1,1,0\ldots,0),\ b_2=(1,1,1,0\ldots,0)\ldots,b_n=(1,1,\ldots,1).$$

a) $v=(n,n-1,n-2,\ldots,1)$

b) $v=(x_1,x_2,\ldots,x_n)$

Zad. 6. Czy istnieją jakieś wektory, które mają takie same współrzędne w każdej bazie? Jeżeli tak, to podać jakiś przykład.

Zad. 7. Niech V będzie dowolną przestrzenią wektorową i niech $v\in V$. Załóżmy ponadto, że w każdej bazie przestrzeni V współrzędne tego wektora są takie same i równe $\lambda_1,\ldots,\lambda_n$. Udowodnić, że wtedy: $\lambda_1=\ldots=\lambda_n=0$.

Operatory liniowe

W każdej przestrzeni wektorowej możemy wyróżnić specjalną klasę funkcji. Są to tzw. *funkcje liniowe* (zwane też *operatorami liniowymi*). Funkcje te są naturalnym uogólnieniem funkcji liniowej znanej z szkoły $(y=ax)$.

DEFINICJA. Niech (V,K) oraz (U,K) będą dwiema dowolnymi przestrzeniami wektorowymi nad tym samym ciałem liczbowym K (możemy przyjąć, że $K=\mathbb{R}$ lub $K=\mathbb{C}$). Funkcję:

$$F:V\to U,$$

nazywamy *operatorem liniowym*, gdy spełnia następujące warunki:

a) $\forall x,y\in V: F(x+y)=F(x)+F(y),$

b) $\forall x\in V,\ \alpha\in K: F(\alpha x)=\alpha F(x).$

Uwaga!

Oba powyższe warunki można wyrazić jednym:

dla dowolnych $x,y\in V,\ \alpha,\beta\in K$ *zachodzi:*

$$F(\alpha x+\beta y)=\alpha F(x)+\beta F(y).$$

Zad. 8. Niech $F:V\to U$ będzie dowolnym operatorem liniowym. Wtedy

a) $F(0)=0,$

b) $F(-x)=-F(x),$

c) $F(x-y)=F(x)-F(y),$

d) Jeżeli $W\subseteq V$ jest podprzestrzenią wektorową przestrzeni V, to $F(W)$ jest podprzestrzenią wektorowa przestrzeni U.

Zad. 9. Które z podanych funkcji są, a które nie są operatorami liniowymi?

a) $F:\mathbb{R}^2\to\mathbb{R},\ F(x,y)=2x+1$

b) $F:\mathbb{R}^2\to\mathbb{R},\ F(x,y)=2x+3y$

c) $F:\mathbb{R}^2\to\mathbb{R}^2,\ F(x,y)=(2x+3y,x-y)$

d) $F:\mathbb{R}^2\to\mathbb{R},\ F(x,y)=x^3$

e) $F:\mathbb{R}^3\to\mathbb{R}^2,\ F(x,y,z)=(x+y,\ y+z,\ x)$

f) $F:\mathbb{R}^3\to\mathbb{R}^3,\ F(x,y,z)=(z,\ x,\ y)$

g) $F:\mathbb{R}^2\to\mathbb{R}^4,\ F(x,y)=(0,x,\ 0,y)$

h) $F:\mathbb{R}^2\to\mathbb{R},\ F(x,y)=xy$

Zad. 10. Udowodnić, że ogólna postać funkcji liniowej z $\mathbb{R}^n$ do $\mathbb{R}$, czyli $F:\mathbb{R}^n\to\mathbb{R}$ jest następująca:

$$F(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \ldots + a_nx_n = \sum_{k=1}^{n} a_k x_k,$$

gdzie $a_1, \ldots, a_n \in \mathbb{R}$ są pewnymi współczynnikami liczbowymi.

Zad. 11. Można udowodnić, że każde przekształcenie liniowe jest w pełni określone, jeżeli tylko znamy jego wartości na wektorach pewnej ustalonej bazy.

„Odtworzyć" wzór na przekształcenie liniowe na podstawie poniższych danych:

a) $F : \mathbb{R}^1 \to \mathbb{R}^1$ oraz $F(1) = 3$

b) $F : \mathbb{R}^2 \to \mathbb{R}^1$ oraz $F((1,\ 0)) = 2,\ F((0,\ 1)) = 4$

c) $F : \mathbb{R}^2 \to \mathbb{R}^1$ oraz $F((1,\ 1)) = 3,\ F((2,\ 1)) = 5$

d) $F : \mathbb{R}^2 \to \mathbb{R}^2$ oraz $F((1,\ 0)) = (2,\ 1),\ F((0,\ 1)) = (0,\ 4)$

e) $F : \mathbb{R}^2 \to \mathbb{R}^2$ oraz $F((1,\ 4)) = (1,\ 1),\ F((2,\ 1)) = (-1,\ 4)$

f) $F : \mathbb{R}^1 \to \mathbb{R}^2$ oraz $F(3) = (5,\ -1)$

g) $F : \mathbb{R}^n \to \mathbb{R}^1$ oraz $F(e_1) = a_1, \ldots, F(e_n) = a_n,$ gdzie $e_1, \ldots, e_n$ są wektorami bazy kanonicznej, natomiast $a_1, \ldots, a_n \in \mathbb{R}$ są danymi współczynnikami liczbowymi.

h) $F : \mathbb{R}^3 \to V$ oraz $F((1,1,0)) = \upsilon_1,\ F((0,0,1)) = \upsilon_2,\ F((1,0,1)) = \upsilon_3,$ gdzie V jest dowolną przestrzenią wektorową, a $\upsilon_1, \upsilon_2, \upsilon_3 \in V$ to dowolne ustalone wektory z tej przestrzeni.

Zad. 12. Podaj wzór na odwzorowanie liniowe $F : \mathbb{R}^2 \to \mathbb{R}^2,$ które jest:

a) odbiciem względem osi OX

b) odbiciem względem osi OY

c) odbiciem względem prostej o równaniu $y = x$

d) odbiciem względem prostej o równaniu $y = ax$

e) obrotem o kąt φ (w kierunku przeciwnym do ruch wskazów zegara)

f) rzutem na oś OX

g) rzutem na prostą o równaniu $y = ax$

Zad. 13. Podaj wzór na odwzorowanie liniowe $F : \mathbb{R}^3 \to \mathbb{R}^3,$ które jest:

a) rzutem na płaszczyznę OXZ

b) odbiciem względem płaszczyzny o równaniu: $x + y + z = 0.$

Algebra macierzy

Działania na macierzach

Niech n i m będą liczbami naturalnymi. Na potrzeby tego skryptu może nieformalnie określić *macierz* jako prostokątną tablicę liczb rzeczywistych $(\mathbb{R})$ lub zespolonych $(\mathbb{C})$:

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}$$

Liczba wierszy wynosi n, a liczba kolumn wynosi m. Elementy w takiej tablicy indeksujemy dwoma indeksami: a_{ij}. Pierwszy indeks (wskaźnik) oznacza numer wiersza, a drugi numer kolumny. Tak wiec a_{ij} oznacza element z $i-tego$ wiersza oraz $j-tej$ kolumny. Zbiór wszystkich macierzy o m wierszach i n kolumnach oznaczamy o elementach ze zbioru K oznaczamy symbolem $M(n,m;K)$. Czasami, gdy wiadomo o jaki zbiór chodzi piszemy krócej $M(n,m)$. Stosowane są też inne oznaczenia, na przykład $M(n\times m)$, $K^{n\times m}$ lub po prostu $n\times m$. Tak więc zbiór macierzy rzeczywistych możemy też oznaczać przez $\mathbb{R}^{n\times m}$, a zespolonych $\mathbb{C}^{n\times m}$.

Jeżeli $n=m$, to mówimy, że macierz jest *kwadratowa*. Poniżej jest kilka przykładów macierzy

$$\begin{bmatrix} 1 & -2 \\ 4 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 5 & 0 \\ -2 & 1/2 & 6 \\ 0 & -1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 4 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 9 & 1 & 1 \\ 0 & 0 & -2 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 2 \end{bmatrix}.$$

Mamy tu kolejno macierze typu $\mathbb{R}^{2\times2}, \mathbb{R}^{3\times3}, \mathbb{R}^{3\times2}, \mathbb{R}^{2\times3}, \mathbb{R}^{4\times1}, \mathbb{R}^{1\times3}$.
Do oznaczania macierzy używamy na ogół dużych liter A, B, C, X itp. Macierz o elementach a_{ij} jest często oznaczana symbolem (a_{ij}). Tak więc piszemy: macierz $A=(a_{ij})$ lub, gdy chcemy podać jawnie liczbę wierszy i kolumn: $A=(a_{ij})_{n\times m}$ (macierz o wymiarach $n\times m$.

O przydatności macierzy w matematyce i zastosowaniach decydują dopiero działania, które są określone w zbiorze macierzy: *dodawanie* dwóch macierzy, *mnożenie macierzy przez liczbę* oraz *mnożenie dwóch macierzy*.

Jeżeli A, B są dwiema macierzami o tym samym wymiarze $n\times m$

$$A=\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}, \quad B=\begin{bmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nm} \end{bmatrix},$$

to suma tych macierzy, oznaczana przez $A+B,$ jest zdefiniowana następująco

$$A+B=\begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1m}+b_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & \cdots & a_{nm}+b_{nm} \end{bmatrix},$$

a iloczyn macierzy A przez skalar α (liczbę z $\mathbb{R}$ lub $\mathbb{C}$), oznaczany jako $\alpha A,$ jest zdefiniowany następująco

$$\alpha A=\begin{bmatrix} \alpha a_{11} & \cdots & \alpha a_{1m} \\ \vdots & \ddots & \vdots \\ \alpha a_{n1} & \cdots & \alpha a_{nm} \end{bmatrix}.$$

Podkreślmy, że dodawać możemy tylko macierze tego samego rozmiaru. Obie te operacje są jak widzimy bardzo proste: aby dodać dwie macierze, dodajemy odpowiadające sobie elementy a mnożenie przez skalar polega na mnożeniu każdego elementu macierzy przez ten skalar. Jest to zilustrowane na przykładzie poniżej

$$\begin{bmatrix} 2 & 0 & -3 \\ 4 & 2 & 1 \end{bmatrix}+\begin{bmatrix} 3 & 1 & 3 \\ 5 & 0 & 7 \end{bmatrix}=\begin{bmatrix} 5 & 1 & 0 \\ 9 & 2 & 8 \end{bmatrix}, \quad 3\begin{bmatrix} 1 & 0 & -3 \\ 2 & 2 & 1 \end{bmatrix}=\begin{bmatrix} 3 & 0 & -9 \\ 6 & 6 & 3 \end{bmatrix}.$$

Z definicji tych dwóch działań na macierzach oraz z własności działań w zbiorach $\mathbb{R}$ i $\mathbb{C}$ wynikają następujące prawa: jeżeli A,B,C są dowolnymi macierzami rozmiaru $n\times m,$ a α,β dowolnymi skalarami, to

a) $A+B=B+A$

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

c) $\alpha(A+B)=\alpha A+\alpha B$

d) $(\alpha\beta)A=\alpha(\beta A)$

e) $(\alpha+\beta)A=\alpha A+\beta A$

Uzasadnienie jest bardzo proste – wystarczy tylko stosować odpowiednie definicje. Na przykład równość a) możemy uzasadnić tak: niech $A=(a_{ij}),\ B=(b_{ij})$ będą macierzami rozmiaru $n\times m.$ Z definicji $A+B=(a_{ij}+b_{ij}),$ ale dodawanie liczb jest przemienne, więc $a_{ij}+b_{ij}=b_{ji}+a_{ij},$ czyli $A+B=(a_{ij}+b_{ij})=(b_{ij}+a_{ij})=B+A.$ Innym sposobem może być zapisywanie macierzy w postaci tablic. Uzasadnijmy równość e)

$$(\alpha+\beta)A=(\alpha+\beta)\begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix} \overset{def}{=} \begin{bmatrix} (\alpha+\beta)a_{11} & \dots & (\alpha+\beta)a_{1m} \\ \vdots & \ddots & \vdots \\ (\alpha+\beta)a_{n1} & \dots & (\alpha+\beta)a_{nm} \end{bmatrix} =$$

$$= \begin{bmatrix} \alpha a_{11}+\beta a_{11} & \dots & \alpha a_{1m}+\beta a_{1m} \\ \vdots & \ddots & \vdots \\ \alpha a_{n1}+\beta a_{n1} & \dots & \alpha a_{nm}+\beta a_{nm} \end{bmatrix} \overset{def}{=} \begin{bmatrix} \alpha a_{11} & \dots & \alpha a_{1m} \\ \vdots & \ddots & \vdots \\ \alpha a_{n1} & \dots & \alpha a_{nm} \end{bmatrix} + \begin{bmatrix} \beta a_{11} & \dots & \beta a_{1m} \\ \vdots & \ddots & \vdots \\ \beta a_{n1} & \dots & \beta a_{nm} \end{bmatrix} \overset{def}{=}$$

$$= \alpha \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix} + \beta \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix} = \alpha A + \beta A.$$

Jak widać tożsamość $(\alpha+\beta)A=\alpha A+\beta A$ opiera się na tożsamości $(\alpha+\beta)a=\alpha a+\beta a,$ czyli na *prawie rozdzielności mnożenia względem dodawania* w zbiorze $\mathbb{R}$ lub $\mathbb{C}.$

Macierze z dotychczas określonymi działaniami, $A+B,$ $\alpha A,$ nie byłyby zbyt przydatne, gdyby nie kolejne działanie na macierzach, a mianowicie iloczyn dwóch macierzy. Tym razem nie jest to prosta operacja polegająca na mnożeniu odpowiadających sobie elementów macierzy $A=(a_{ij}),$ $B=(b_{ij}).$ Operacja ta wygląda następująco: jeśli A jest macierzą rozmiaru $n\times p,$ a B jest macierzą $p\times m,$ to ich iloczyn $C=A\cdot B$ jest macierzą rozmiaru $n\times m$ o elementach określonych następująco

$$c_{ij}=\sum_{k=1}^{p} a_{ik}b_{kj} \qquad (1\le i\le n,\ 1\le j\le m). \tag{0.17}$$

Jeżeli się dokładniej przyjrzymy temu wzorowi, to zauważymy, że element c_{ij} iloczynu powstaje przez „pomnożenie” $i-$ tego wiersza macierzy $A=(a_{ij})$ oraz $j-$ tej kolumny macierzy $B=(b_{ij}).$ Na przykład mnożąc dwie macierze 2×2 otrzymamy macierz 2×2

$$\begin{bmatrix} 3 & 2 \\ -2 & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 & 5 \\ 2 & 6 \end{bmatrix} = \begin{bmatrix} 3\cdot 1+2\cdot 2 & 3\cdot 5+2\cdot 6 \\ -2\cdot 1+4\cdot 2 & -2\cdot 5+4\cdot 6 \end{bmatrix} = \begin{bmatrix} 7 & 27 \\ 6 & 14 \end{bmatrix},$$

a mnożąc macierz 2×3 przez macierz 3×3 otrzymamy macierz 2×3

$$\begin{bmatrix} 1 & 2 & -3 \\ 0 & 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} 4 & 1 & 0 \\ -1 & 2 & 2 \\ 3 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1\cdot 4+2\cdot(-1)+(-3)\cdot 3 & 1\cdot 1+2\cdot 2+(-3)\cdot 3 & 1\cdot 0+2\cdot 2+(-3)\cdot 3 \\ 0\cdot 4+2\cdot(-1)+(-1)\cdot 3 & 0\cdot 1+2\cdot 2+(-1)\cdot 0 & 0\cdot 0+2\cdot 2+(-1)\cdot 3 \end{bmatrix} =$$

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

Podkreślmy jeszcze raz: aby można było pomnożyć macierz A przez B otrzymując macierz $C=AB$ musi być spełniony warunek

liczba wierszy macierzy A = liczba kolumn macierzy B.

Mnożenie macierzy jest łączne

$$A(BC) = (AB)C. \tag{0.18}$$

Zakładamy oczywiście, że mnożenia można wykonać, czyli $A \in M(n,p)$, $B \in M(p,q)$, $C \in M(q,m)$. Aby udowodnić równość (0.18) najlepiej posłużyć się rachunkiem z wykorzystaniem symbol sumy i zastosować definicję (0.17). Ponadto w odpowiednim momencie musimy skorzystać z łączność mnożenia liczba, $a(bc) = (ab)c$. Mamy mianowicie (symbol $(M)_{ij}$ oznacza odpowiedni element macierzy M : jest te element z $i-$tego wiersza oraz $j-$tej kolumny)

$$\big(A(BC)\big)_{ij} = \sum_{k=1}^{p} a_{ik}(BC)_{kj} = \sum_{k=1}^{p}\left(a_{ik}\sum_{l=1}^{q} b_{kl}c_{lj}\right) = \sum_{k=1}^{p}\sum_{l=1}^{q} a_{ik}(b_{kl}c_{lj}) = \sum_{k=1}^{p}\sum_{l=1}^{q}(a_{ik}b_{kl})c_{lj} =$$
$$= \sum_{l=1}^{q}\sum_{k=1}^{p}(a_{ik}b_{kl})c_{lj} = \sum_{l=1}^{q}(AB)_{il}c_{lj} = \big((AB)C\big)_{ij}.$$

Mnożenie macierzy jest rozdzielne względem dodawania macierzy. Dokładniej, jeżeli $A \in M(n,p)$, $B, C \in M(p,m)$, to zachodzi równość

$$A\cdot(B+C) = A\cdot B + A\cdot C. \tag{0.19}$$

Sposób uzasadnienia jest podobny jak w przypadku łączności – wzór (0.18). Tym razem jednak w odpowiednim miejscu należy skorzystać z rozdzielności mnożenia względem dodawania w zbiorze $\mathbb{R}$ lub $\mathbb{C}$, $a(b+c) = ab + ac$. Mamy bowiem

$$\big(A(B+C)\big)_{ij} = \sum_{k=1}^{p} a_{ik}(B+C)_{kj} = \sum_{k=1}^{p} a_{ik}(b_{kj}+c_{kj}) = \sum_{k=1}^{p} a_{ik}b_{kj} + a_{ik}c_{kj} = \sum_{k=1}^{p} a_{ik}b_{kj} + \sum_{k=1}^{p} a_{ik}c_{kj} =$$
$$= (AB)_{ij} + (AC)_{ij} = (AB+AC)_{ij}.$$

Pokazaliśmy więc, że każdy element $(i,j)-$ty macierzy $A(B+C)$ jest równy elementowi $(i,j)-$temu macierzy $AB + BC$. Stąd wynika równość (0.19).

Kolejna własność mnożenie macierzy: niech $A \in M(n,p)$, $B \in M(p,m)$ oraz α będzie skalarem, wtedy

$$\alpha(A\cdot B) = (\alpha A)\cdot B = A\cdot(\alpha B) \tag{0.20}$$

Uzasadnienie jest całkiem elementarne

$$\big(\alpha(A\cdot B)\big)_{ij} = \alpha(A\cdot B)_{ij} = \alpha\sum_{k=1}^{p} a_{ik}b_{kj} = \sum_{k=1}^{p}(\alpha a_{ik})b_{kj} = \sum_{k=1}^{p}(\alpha A)_{ik}b_{kj} = \big((\alpha A)\cdot B\big)_{ij}.$$

Ponieważ równość zachodzi dla każdej pary indeksów (i,j) więc macierze są równe, $\alpha(A\cdot B) = (\alpha A)\cdot B$. Drugą równość z (0.20) uzasadniamy analogicznie.

Dla danej macierzy A rozmiaru $n\times m$ macierz *transponowana*, oznaczana symbolem A^T, jest zdefiniowana poprzez przestawienie wierszy z kolumnami, czyli jeżeli $A = (a_{ij})$, to $(A^T)_{ij} = a_{ji}$. Oto przykład

$$A = \begin{bmatrix} 1 & 3 & -2 \\ 4 & 0 & 5 \end{bmatrix}, \quad A^T = \begin{bmatrix} 1 & 4 \\ 3 & 0 \\ -2 & 5 \end{bmatrix}.$$

Z definicji transponowania macierzy wynika, że jeżeli $A \in M(n,m)$, to $A \in M(m,n)$. Operacja transponowania ma następujące własności

a) $(A^T)^T = A$

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

c) $(\alpha A)^T = \alpha A^T$

d) $(A \cdot B)^T = B^T \cdot A^T$

Tożsamości a) – c) są oczywiste, jedynie tożsamość d) wymaga bardziej szczegółowego uzasadnienia. Mamy

$$\left((A\cdot B)^T\right)_{ij} = (A\cdot B)_{ji} = \sum_{k=1}^{p} a_{jk} b_{ki} = \sum_{k=1}^{p} b_{ki} a_{jk} = \sum_{k=1}^{p} (B^T)_{ik} (A^T)_{kj} = \left(B^T \cdot A^T\right)_{ij}.$$

Z dowolności i, j mamy więc $(A \cdot B)^T = B^T \cdot A^T$. Zauważmy też, że w tej równości ważna jest kolejność mnożenia, gdyż mnożenie macierzy nie jest przemienne.

Szczególnym przypadkiem macierzy są tzw. *macierze kwadratowe*, czyli takie, w których liczba wierszy jest równa liczbie kolumn $(m = n)$

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}.$$

W przypadku macierzy kwadratowych używamy oznaczeń $M(n,\mathbb{R})$, $M(n,\mathbb{C})$, $M(n)$, lub $\mathbb{R}^{n\times n}$, $\mathbb{C}^{n\times n}$. Mówimy też: macierz kwadratowa stopnia n lub po prostu macierz stopnia n.

Macierz kwadratową stopnia n postaci

$$I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix},$$

nazywamy macierzą jednostkową. Jedynki występują tylko na głównej przekątnej (na diagonalnej) a poza nią występują wszędzie zera. Macierz jednostkowa ma własność

$$A \cdot I = I \cdot A = A, \tag{0.21}$$

Dlatego możemy powiedzieć, że macierz I jest elementem neutralnym mnożenia macierzowego (pełni podobną rolę jak liczba 1 w zbiorze liczb rzeczywistych $\mathbb{R}$). Elementy macierzy jednostkowej zazwyczaj oznaczamy symbolem δ_{ij} (*delta Kroneckera*)

$$\delta_{ij} = \begin{cases} 1 & \text{dla } i = j, \\ 0 & \text{dla } i \neq j. \end{cases} \tag{0.22}$$

Sprawdźmy pierwszą z równości (0.21)

$$(AI)_{ij} = \sum_{k=1}^{n} a_{ik}\delta_{kj} = a_{ij} = (A)_{ij},$$

gdyż w powyższej sumie składniki dla $k \neq j$ równe zero a dla $k = j$ $\delta_{kj} = 1.$

Wspomnieliśmy już, że mnożenie macierzy nie jest przemienne. W przypadku macierzy kwadratowych stopnia n, $A, B \in M(n)$ oba iloczyny AB oraz BA są poprawnie określone, ale macierze te są na ogół różne. Oto przykład

$$AB = \begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 0 & 3 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 3 \\ 2 & 4 \end{bmatrix}, \; BA = \begin{bmatrix} 0 & 3 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 6 & 6 \\ -1 & -2 \end{bmatrix},$$

więc mamy $AB \neq BA.$

Teraz krótko omówimy kolejne pojęcie algebry macierzy ważne w zastosowaniach – macierz odwrotną do danej macierzy kwadratowej.

Definicja. Niech A będzie daną macierzą kwadratową stopnia n. Macierz kwadratową B stopnia n nazywamy macierzą odwrotną względem macierzy, gdy zachodzą równości

$$A \cdot B = B \cdot A = I. \tag{0.23}$$

Zauważmy, że równości definicyjnej (0.23) wynika, że jeżeli macierz A jest odwrotna względem B, to również jest odwrotnie: macierz B jest odwrotna względem $A.$

Na przykład

$$AB = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 3\cdot 1 + 1\cdot(-2) & 3\cdot(-1) + 1\cdot 3 \\ 2\cdot 1 + 1\cdot(-2) & 2\cdot(-1) + 1\cdot 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

oraz

$$BA = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1\cdot 3 + (-1)\cdot 2 & 1\cdot 1 + (-1)\cdot 1 \\ (-2)\cdot 3 + 3\cdot 2 & (-2)\cdot 1 + 3\cdot 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Tak więc macierze

$$A = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix},$$

są wzajemnie odwrotne.

Po wprowadzeniu pojęcia macierzy odwrotnej od razu pojawiają się oczywiste pytania:

a) Jaki warunek musi spełniać macierz kwadratowa, aby istniała do niej odwrotna?
b) Czy macierz odwrotna jest wyznaczona jednoznacznie?
c) Jeżeli istnieje macierz odwrotna, to jak ją wyznaczyć?

Pytanie a)

Łatwo zauważyć, że nie każda macierz kwadratowa posiada odwrotną. Oczywistym przykładem jest macierz zerowa

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

gdyż dla dowolnej macierzy $B \in M(2)$ mamy

$$AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

Ale nawet macierze niezerowe mogą nie posiadać odwrotnej o czym świadczy przykład

$$AB = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} b_{11} + 2b_{21} & b_{12} + 2b_{22} \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

gdyż dla dowolnych elementów macierzy B nie otrzymamy jedynki w prawym dolnym rogu. Widać więc, że kryterium na odwracalność macierzy musi być nieco bardziej subtelne. W kursie algebry liniowej dowodzi się kilku wzajemnie równoważnych postaci takiego kryterium. Przypomnijmy je

TWIERDZENIE. Niech dane będzie macierz kwadratowa A stopnia n. Następujące warunki są równoważne

i) macierz A jest odwracalna
ii) kolumny macierzy A traktowane jak wektory w $\mathbb{R}^n$ $(\mathbb{C}^n)$ są liniowo niezależne
iii) wiersze macierzy A traktowane jak wektory w $\mathbb{R}^n$ $(\mathbb{C}^n)$ są liniowo niezależne
iv) wyznacznik macierzy jest niezerowy $(\det A \neq 0)$

Pojęcie wyznacznika będzie omówione dalej nieco dokładniej, ale z kursu matematyki szkolnej zapewne czytelnicy znają wzór na wyznacznik dla układu dwóch równań liniowych z dwoma niewiadomymi, który dla macierzy współczynników ma postać

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{21}a_{12}. \tag{0.24}$$

Tak więc kryterium na odwracalność macierzy 2×2 jest $a_{11}a_{22}-a_{21}a_{12}\neq 0$. Ponieważ wyznacznik dla dowolnej macierzy kwadratowej może być w zasadzie obliczony (omawiane dalej tzw. *rozwinięcie Laplace'a*), więc mogłoby się wydawać, że najlepszym sposobem na sprawdzenie czy macierz jest nieosobliwa jest obliczenie wg rozwinięcia Laplace'a. Tak jednak nie jest, gdyż analiza tego wzoru pokazuje, że ma on złożoność rzędu $n!$, a więc całkowicie się nie nadaje do większych macierzy. Zdecydowanie lepszy jest sposób oparty o eliminację Gaussa o złożoności rzędu n^3.

<u>Pytanie b)</u>

Przypomnijmy: czy macierz odwrotna jest jedna (jeśli w ogóle istnieje)? Można to sformułować tak: dana jest macierz kwadratowa $A\in M(n)$ oraz dwie macierze kwadratowe $B_1, B_2\in M(n)$. Czy zachodzi implikacja

$$\left.\begin{aligned} AB_1 = B_1A = I \\ AB_2 = B_2A = I \end{aligned}\right\} \;\Rightarrow\; B_1 = B_2.$$

Okazuje się, że odpowiedź jest twierdząca, a dowód jest prosty (opiera się na łączności mnożenia macierzy)

$$B_1 = B_1 I = B_1(AB_2) = (B_1A)B_2 = IB_2 = B_2.$$

Ze względu na jednoznaczność odwrotnej do macierzy A możemy używać jednego oznaczenia, którym jest A^{-1}. Tak więc z definicji mamy

$$A\cdot A^{-1} = A^{-1}\cdot A = I. \tag{0.25}$$

<u>Pytanie c)</u>

Pytanie o sposoby wyznaczania macierzy odwrotnej wkracza już w obszar metod numerycznych. Okazuje się, że zagadnienie to związane jest z rozwiązywaniem układów równań liniowych. Należy podkreślić, że znany z kursu algebry liniowej „wzór" na macierz odwrotną poprzez tzw. dopełnienia algebraiczne (więcej szczegółów dalej)

$$A^{-1} = \frac{1}{\det A}\left[A_{ij}^{D}\right],$$

nie nadaje się do „prawdziwych" obliczeń numerycznych, w których pojawią się macierze 100×100, 1000×1000 lub większe. Powodem jest to, że wzór powyższy ma złożoność obliczeniową rzędu $n\cdot n!$, więc nadaje się w praktyce do małych macierzy. Oczywiście wzór powyższy ma duże znaczenie teoretyczne ale nie numeryczne. Zagadnieniem numerycznego wyznaczania macierzy A^{-1} w oparciu o procedurę eliminacji Gaussa zajmiemy się w rozdziale Układy Liniowe.

Operacja odwracania macierzy ma następujące własności (wszędzie zakładamy, że $A,B\in M(n)$ są nieosobliwe):

a) $(\alpha A)^{-1} = \alpha^{-1}A^{-1}$ jeżeli $\alpha\neq 0$,

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

c) $(A^{-1})^T = (A^T)^{-1}$,

d) $(A^{-1})^{-1} = A$.

Powyższe tożsamości dowodzi się stosunkowo łatwo, jeżeli tylko przypomnimy sobie definicję macierzy odwrotnej (równość (0.23) lub (0.25)).

Ad a) Pytamy czy $\alpha^{-1}A^{-1}$ jest odwrotną dla macierzy αA? Mamy

$$(\alpha A)(\alpha^{-1}A^{-1}) = (\alpha\alpha^{-1})(AA^{-1}) = 1I = I,$$

zatem $(\alpha A)^{-1} = \alpha^{-1}A^{-1}$.

Ad b) Pytamy czy $B^{-1}A^{-1}$ jest odwrotną dla macierzy AB? Mamy

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I,$$

zatem $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.

Ad c) Korzystamy z dwóch faktów: $I^T = I$ oraz $(AB)^T = B^T A^T$. Zachodzi

$$A^{-1}A = I \;\Rightarrow\; (A^{-1}A)^T = I^T \;\Rightarrow\; A^T\left(A^{-1}\right)^T = I \;\Rightarrow\; \left(A^T\right)^{-1} = \left(A^{-1}\right)^T.$$

Ad d) Jest to właściwie ćwiczenie na zrozumienie sensu definicji. Mamy bowiem

$$AA^{-1} = A^{-1}A = I,$$

w szczególności wynika stąd, że odwrotną dla A^{-1} jest właśnie A, więc $\left(A^{-1}\right)^{-1} = A$.

Zad. 1. Czy każde dwie macierze można dodać? Jakie muszą być spełnione warunki aby można było dodać do siebie dwie macierze?

Zad. 2. Jaka jest definicja mnożenia macierzy przez liczbę (skalar)? Niech $\alpha \in \mathbb{R}$. Jaki jest wynik poniższych mnożeń:

$$\alpha\begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix} = \qquad -3\begin{bmatrix} 2 & 4 & 6 \\ 5 & 0 & 7 \\ 0 & 4 & 4 \end{bmatrix} = \qquad x\begin{bmatrix} \frac{1}{x} & x \\ 2 & \frac{1}{x^2} \end{bmatrix} =$$

Przypomnijmy: jeżeli $A \in M(n, p)$, $B \in M(p, m)$, to określone jest mnożenie macierzy. W wyniku uzyskamy nową macierz $C = A \cdot B$, taką że $C \in M(n, m)$. Ważna jest kolejność mnożenia!

Zad. 3. Wykonaj mnożenia następujących par macierzy kwadratowych:

$$\begin{bmatrix} 2 & 5 \\ 4 & 7 \end{bmatrix} \cdot \begin{bmatrix} 3 & 6 \\ 2 & 8 \end{bmatrix} =$$

$$\begin{bmatrix} 2 & 4 & 1 \\ 3 & 5 & 9 \\ 7 & 2 & 6 \end{bmatrix} \begin{bmatrix} 2 & 2 & 10 \\ 5 & 2 & 12 \\ 1 & 3 & -1 \end{bmatrix} =$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 & 10 \\ 1 & 3 & 12 \\ 3 & 3 & -1 \end{bmatrix} =$$

Zad. 4. Jaką postać ma macierz jednostkowa? Jaką własność posiada macierz jednostkowa?

Zad. 5. Wykonaj mnożenia macierzy

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 & 2 \\ 1 & 3 \\ 3 & 3 \end{bmatrix} =$$

$$\begin{bmatrix} 1 & 0 & 5 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 2 & 10 & 0 \\ 1 & 0 & -1 & 2 \\ 3 & 3 & -1 & 3 \end{bmatrix} =$$

$$\begin{bmatrix} 1 & 2 & 3 & -2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \\ 3 \\ 4 \end{bmatrix} =$$

Zad. 6. Podaj przykład dwóch niezerowych macierzy, których iloczyn jest macierzą zerową. Odpowiedni przykład jest już dla macierzy kwadratowych 2×2. Czy podobne zjawisko może się zdarzyć gdy mnożymy dwie liczby rzeczywiste?

Zad. 7. Dane są macierze:

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

Oblicz następujące wyrażenia macierzowe:

a) $(3A+B)(C-2D)$

b) $(A-B)(A+B)$

c) $(A+B)^2$

d) C^2-CD

Zad. 8. Czy w algebrze macierzy prawdziwe są następujące wzory skróconego mnożenia:

a) $(A+B)(A-B)=A^2-B^2$

b) $(A+B)^2=A^2+2AB+B^2$

Jeżeli nie, to z czym to jest związane?

Zad. 9. Oblicz następującą potęgę macierzy:

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{1994} =$$

Zad. 10. Udowodnij następującą tożsamość:

$$\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \cdot \begin{bmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{bmatrix} = \begin{bmatrix} \cos(\varphi+\psi) & -\sin(\varphi+\psi) \\ \sin(\varphi+\psi) & \cos(\varphi+\psi) \end{bmatrix}.$$

Wyznaczniki

Wyznacznik macierzy jest to pewna liczba, którą w sposób jednoznaczny przypisujemy do macierzy, przy czym określamy ją tylko dla macierzy kwadratowych. Jeżeli A jest macierzą kwadratową $n\times n$, to wyznacznik tej macierzy oznaczamy symbolem $\det A$. Dla macierzy rzeczywistych mamy $\det A\in\mathbb{R}$ a dla zespolonych $\det A\in\mathbb{C}$. Czasami stosujemy inne oznaczenie wyznacznika, a mianowicie $|A|$. Wyznacznik ma szereg własności, w szczególności informuje nas o tym czy macierz jest odwracalna: zachodzi to wtedy i tylko wtedy, gdy $\det A\neq 0$. Niestety definicja wyznacznika nie jest elementarna i wprowadzenie jej wymaga udowodnienie szeregu własności. Tutaj podamy podstawową charakteryzację wyznacznika oraz podstawowe sposoby jego obliczania.

Okazuje się, że wyznacznik jako funkcja na zbiorze macierzy kwadratowych

$$\det : M(n)\to\mathbb{R},$$

jest jednoznacznie scharakteryzowany przez następujące warunki:

i) wyznacznik macierzy jednostkowej jest równy 1: $\det I=1$,

ii) wyznacznik jest liniową funkcją względem poszczególnych kolumn:

$$\det[c_1,\ldots,c_{i-1},\alpha c_i+\beta c_i',c_{i+1},\ldots,c_n]=\alpha\det[c_1,\ldots,c_i,\ldots,c_n]+\beta\det[c_1,\ldots,c_i',\ldots,c_n]$$

iii) jeśli dwie kolumny macierzy są równe, to wyznacznik tej macierzy jest równy zeru.

Można udowodnić, że istnieje funkcja o powyższych własnościach i to tylko jedna. Kolumny nie są w jakiś szczególny sposób wyróżnione w tej definicji. Można kolumny zamienić powyżej z wierszami i otrzymamy ten sam wyznacznik. W szczególności wynika stąd, że wyznacznik jest liniowy względem poszczególnych wierszy oraz jeżeli dwa wiersze są takie same, to $\det A = 0.$

Niestety to co dotychczas powiedzieliśmy o wyznaczniku nadal nie pokazuje jak go obliczać. Jednakże dokładniejsza analiza powyższych warunków prowadzi do pewnych „wzorów” na obliczanie wyznaczników.

Wzór I

Jeżeli A jest macierzą stopnia $n,$ to

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \cdot a_{1\sigma_1} \cdot a_{2\sigma_2} \cdot \ldots \cdot a_{n\sigma_n}, \tag{0.26}$$

gdzie S_n jest zbiorem wszystkich permutacji zbioru $\{1, 2, \ldots, n\},$ a $\operatorname{sgn} \sigma$ oznacza tzw. *znak permutacji* $(\operatorname{sgn} \sigma = \pm 1).$[4] Praktyczne obliczanie wyznacznika ze wzoru (0.26) nie bardzo ma sens, gdyż suma w nim występująca ma $n!$ składników. Widać natomiast, że dla $n = 1$ mamy $\det[a_{11}] = a_{11},$ natomiast dla $n = 2,$ gdzie mamy dwie permutacje $S_2 = \{(1, 2), (2, 1)\}$ otrzymamy

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \operatorname{sgn}(1, 2) a_{11} a_{22} + \operatorname{sgn}(2, 1) a_{12} a_{21} = a_{11} a_{22} - a_{12} a_{21},$$

a więc wyrażenie które pojawia się już w szkole średniej. Oczywiście dla $n = 3$ liczba składników będzie $6.$

Wzór II (rozwinięcie Laplace’a)

Ten sposób obliczania wyznacznika ma charakter rekurencyjny: wyznacznik macierzy stopnia n sprowadza się do obliczania n wyznaczników stopnia $n-1.$ Rozwinięcie Laplace’a względem $j-$ tego wiersza dla macierzy $A = (a_{ij})$ stopnia n ma postać

---

[4] Przypomnijmy, że *permutacją* dowolnego skończonego zbioru X nazywamy dowolną wzajemnie jednoznaczną funkcję $f : X \to X.$ Jeżeli elementy zbioru X ponumerujemy $X = \{x_1, x_2, \ldots, x_n\},$ to widać, że każda taka funkcja jest właściwie tym samym co ciąg elementów zbioru $X,$ w którym każdy element występuje dokładnie jeden raz. Na przykład dla $X = \{a, b, c\}$ permutacje są tożsame z ciągami: $abc, acb, cab, cba, bca, bac.$ Jak łatwo zauważyć dla zbioru $n-$ elementowego liczba permutacji wynosi $n!.$ Ponadto można udowodnić, że każda permutacja σ jest złożeniem pewnej liczby *transpozycji*. Transpozycja to taka permutacja, w której tylko dwa wybrane elementy są zamienione miejscami (pozostałe nie są ruszane). Dla parzystej liczby transpozycji przyjmujemy $\operatorname{sgn} \sigma = 1,$ a dla nieparzystej $\operatorname{sgn} \sigma = -1.$

$$\det A = \sum_{k=1}^{n} (-1)^{j+k} a_{jk} \det A_{jk}, \tag{0.27}$$

gdzie macierz A_{jk} jest macierzą stopnia $n-1$ powstałą z macierzy A przez wykreślenie $j-$tego wiersza i $k-$tej kolumny. Dla $n=1$ już nie stosujemy wzoru (0.27) tylko przyjmujemy, że $\det[a]=a$.

Przykład. Obliczyć wyznacznik macierzy kwadratowej stopnia $n=2$ stosując rozwinięcie względem pierwszego, a następnie drugiego wiersza. porównać wyniki.

Rozwiązanie: Mamy daną macierz

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

Zauważmy, że skreślając na przykład pierwszy wiersz i pierwszą kolumnę otrzymamy

$$A_{11} = [a_{22}].$$

Stosujemy teraz rozwinięcie Laplace'a (0.27) dla $j=1$ do naszej macierzy

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = (-1)^{1+1} a_{11} \det A_{11} + (-1)^{1+2} a_{12} \det A_{12} = a_{11} \det[a_{22}] - a_{12} \det[a_{21}] =$$
$$= a_{11}a_{22} - a_{12}a_{21}.$$

Jeżeli teraz zastosujemy rozwinięcie Laplace'a dla $j=2$ (czyli względem drugiego wiersza), to otrzymamy

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = (-1)^{2+1} a_{21} \det A_{21} + (-1)^{2+2} a_{22} \det A_{22} = -a_{21} \det[a_{12}] + a_{22} \det[a_{11}] =$$
$$= -a_{12}a_{21} + a_{11}a_{22} = a_{11}a_{22} - a_{12}a_{21}.$$

Jak widać w obu przypadkach otrzymaliśmy to samo.

Złożoność obliczeniowa rozwinięcia Laplace'a (0.27) jest niestety również $n!$. Oznacza to, że dla dużych n stosowanie tej metody może być w praktyce niewykonalne. Tym niemniej, obliczanie wyznacznika metodą rozwinięcia Laplace'a jest w praktyce stosowane dość często. Dotyczy to przypadków niedużych ($n \le 5$) macierzy lub takich, gdzie jest pewna zależność rekurencyjna pomiędzy wyznacznikiem stopnia n a wyznacznikami mniejszych stopni. Przy obliczaniu wyznacznika pomocne mogą być pewne operacje wykonywane na macierzy, które nie zmieniają wyznacznika macierzy.

Poniżej podano te własności wyznaczników, które mogą być pomocne w obliczeniach. Przyjmujemy, że macierz A jest stopnia n.

1) Jeżeli macierz A posiada dwie jednakowe kolumny (wiersze), to $\det A = 0$.

2) Wyznacznik macierzy zmienia znak, gdy przedstawimy dwie kolumny (dwa wiersze).

3) Jeżeli macierz A zawiera kolumnę (wiersz) złożony z samych zer, to $\det A = 0$.

4) Jeżeli do jakiejś kolumny (wiersza) macierzy A dodamy inną kolumnę (wiersz) pomnożoną przez dowolną liczbę, to wyznacznik się nie zmieni.

5) $\det(\alpha A) = \alpha^n \det A$.

Dzięki wymienionym wyżej własnościom możemy daną macierz tak przekształcić, że obliczanie wyznacznika przez rozwinięcie Laplace'a będzie szybkie, gdyż w macierzy będzie dużo zer.

Przykład. Obliczyć wyznacznik

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 3 & 1 \\ 1 & 1 & 1 & 4 \end{vmatrix}.$$

Rozwiązanie: Odejmujemy pierwszą kolumnę kolejno od drugiej, trzeciej i czwartej

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 3 & 1 \\ 1 & 1 & 1 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 3 & 1 \\ 1 & 0 & 1 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 2 & 1 \\ 1 & 0 & 0 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 3 \end{vmatrix}.$$

Teraz wystarczy dokonać rozwinięcia Laplace'a względem pierwszego wiersza (z wyjątkiem pierwszego składnika występują zera)

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 3 \end{vmatrix} = (-1)^{1+1} 1 \begin{vmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{vmatrix}.$$

Podobnie obliczamy uzyskany powyżej wyznacznika stopnia 3:

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

Tak więc wyznacznik wynosi 6.

Z podanych już rozważań wynika, że macierz trójkątna ma wyznacznik, który jest iloczynem elementów na diagonalnej

$$\begin{vmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ a_{31} & a_{32} & a_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix} = (-1)^{1+1} a_{11} \begin{vmatrix} a_{22} & 0 & \dots & 0 \\ a_{32} & a_{33} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & 0 & \dots & 0 \\ a_{32} & a_{33} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix}.$$

Podobnie

$$\begin{vmatrix} a_{22} & 0 & \dots & 0 \\ a_{32} & a_{33} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix} = (-1)^{1+1} a_{22} \begin{vmatrix} a_{33} & \dots & 0 \\ \vdots & & \vdots \\ a_{n3} & \dots & a_{nn} \end{vmatrix} = a_{22} \begin{vmatrix} a_{33} & \dots & 0 \\ \vdots & & \vdots \\ a_{n3} & \dots & a_{nn} \end{vmatrix}.$$

Postępując tak $n-1$ krotnie oraz wykorzystując $\det[a_{nn}] = a_{nn}$ otrzymujemy

$$\begin{vmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ a_{31} & a_{32} & a_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix} = a_{11} a_{22} \cdot \ldots \cdot a_{nn}.$$

Przykład. Obliczyć wyznacznik następującej macierzy stopnia n

$$A = \begin{bmatrix} x & a & a & \dots & a & a \\ b & x & a & \dots & a & a \\ b & b & x & \dots & a & a \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ b & b & b & \dots & x & a \\ b & b & b & \dots & b & x \end{bmatrix}.$$

Rozwiązanie: Niech $W_n = \det A$, gdzie macierz A jest jak wyżej stopnia n. Jeśli w macierzy tej od ostatniego wiersza odejmiemy wiersz przedostatni, a następnie rozwiniemy względem ostatniego wiersza, to otrzymamy

$$W_n = \begin{vmatrix} x & a & a & \dots & a & a \\ b & x & a & \dots & a & a \\ b & b & x & \dots & a & a \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ b & b & b & \dots & x & a \\ b & b & b & \dots & b & x \end{vmatrix} = \begin{vmatrix} x & a & a & \dots & a & a \\ b & x & a & \dots & a & a \\ b & b & x & \dots & a & a \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ b & b & b & \dots & x & a \\ 0 & 0 & 0 & \dots & b-x & x-a \end{vmatrix} =$$

$$= (-1)^{n-1+n}(b-x)\begin{vmatrix} x & a & a & \dots & a \\ b & x & a & \dots & a \\ b & b & x & & a \\ \vdots & \vdots & & \ddots & \vdots \\ b & b & b & \dots & a \end{vmatrix} + (-1)^{n+n}(x-a)W_{n-1}.$$

Teraz musimy jeszcze obliczyć wyznacznik stopnia $n-1$

$$\begin{vmatrix} x & a & a & \dots & a \\ b & x & a & \dots & a \\ b & b & x & & a \\ \vdots & \vdots & & \ddots & \vdots \\ b & b & b & \dots & a \end{vmatrix}.$$

W tym celu wynosimy z ostatniej kolumny a i odejmujemy ostatni wiersz od wierszy poprzednich

$$\begin{vmatrix} x & a & a & \dots & a \\ b & x & a & \dots & a \\ b & b & x & & a \\ \vdots & \vdots & & \ddots & \vdots \\ b & b & b & \dots & a \end{vmatrix} = a\begin{vmatrix} x & a & a & \dots & 1 \\ b & x & a & \dots & 1 \\ b & b & x & & 1 \\ \vdots & \vdots & & \ddots & \vdots \\ b & b & b & \dots & 1 \end{vmatrix} = a\begin{vmatrix} x-b & a-b & a-b & \dots & 0 \\ 0 & x-b & a-b & \dots & 0 \\ 0 & 0 & x-b & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ b & b & b & \dots & 1 \end{vmatrix},$$

a teraz rozwijamy względem ostatniej kolumny, co daje

$$a\begin{vmatrix} x-b & a-b & a-b & \dots & 0 \\ 0 & x-b & a-b & \dots & 0 \\ 0 & 0 & x-b & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ b & b & b & \dots & 1 \end{vmatrix} = a\begin{vmatrix} x-b & a-b & a-b & \dots & a-b \\ 0 & x-b & a-b & \dots & a-b \\ 0 & 0 & x-b & & a-b \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & x-b \end{vmatrix} = a(x-b)^{n-2},$$

bo ostatni wyznacznik jest dla macierzy trójkątnej, więc jest to iloczyn elementów na przekątnej (pamiętajmy, że ostatni wyznacznik jest stopnia $n-2$). Podstawiając ostatnia wartość do wyrażenia na W_n otrzymujemy

$$W_n = (-1)^{n-1+n}(b-x)a(x-b)^{n-2} + (-1)^{n+n}(x-a)W_{n-1} =$$
$$= (x-a)W_{n-1} + a(x-b)^{n-1}.$$

Otrzymaliśmy w ten sposób zależność rekurencyjną pomiędzy W_n a W_{n-1}. Ponieważ dla $n=1$ macierz A składa się tylko z elementu x, więc $W_1 = \det[x] = x$. Podsumowując mamy

$$\begin{cases} W_n = (x-a)W_{n-1} + a(x-b)^{n-1} & \text{dla} \quad n>1, \\ W_1 = x. \end{cases}$$

Można teraz wykazać, że

$$W_n = \frac{b(x-a)^n - a(x-b)^n}{b-a}.$$

Przykład. (*wyznacznik Vandermonde'a*)

Obliczyć wyznacznik następującej macierzy

$$A = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \dots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{bmatrix},$$

gdzie $x_1, \dots, x_n \in \mathbb{R}$.

Rozwiązanie: Kolejno odejmujemy od ostatniej kolumny kolumnę przedostatnią pomnożoną przez x_n, i tak dalej aż od drugiej kolumny pierwszą pomnożona przez x_n (kolejność jest ważna: od prawej do lewej strony „przebiegamy" macierz). Otrzymujemy

$$V_n = \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \dots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix} = \begin{vmatrix} 1 & x_1 - x_n & x_1^2 - x_n x_1 & \dots & x_1^{n-1} - x_n x_1^{n-2} \\ 1 & x_2 - x_n & x_2^2 - x_n x_2 & \dots & x_2^{n-1} - x_n x_2^{n-2} \\ 1 & x_3 - x_n & x_3^2 - x_n x_3 & \dots & x_3^{n-1} - x_n x_3^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & x_n^2 - x_n x_n & \dots & x_n^{n-1} - x_n x_n^{n-2} \end{vmatrix} =$$

$$= \begin{vmatrix} 1 & x_1 - x_n & x_1^2 - x_n x_1 & \dots & x_1^{n-1} - x_n x_1^{n-2} \\ 1 & x_2 - x_n & x_2^2 - x_n x_2 & \dots & x_2^{n-1} - x_n x_2^{n-2} \\ 1 & x_3 - x_n & x_3^2 - x_n x_3 & \dots & x_3^{n-1} - x_n x_3^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \dots & 0 \end{vmatrix} = (-1)^{n+1} \begin{vmatrix} x_1 - x_n & x_1(x_1 - x_n) & x_1^{n-2}(x_1 - x_n) \\ x_2 - x_n & x_2(x_2 - x_n) & x_2^{n-2}(x_2 - x_n) \\ & & \\ x_{n-1} - x_n & x_{n-1}(x_{n-1} - x_n) & x_{n-1}^{n-2}(x_{n-1} - x_n) \end{vmatrix}.$$

Teraz wyłączamy kolejno z każdego wiersza wspólne czynniki, $(x_1 - x_n)$, $(x_2 - x_n)$, $\dots, (x_{n-1} - x_n)$ co daje

$$V_n = (-1)^{n+1}(x_1 - x_n)\dots(x_{n-1} - x_n)\begin{vmatrix} 1 & x_1 & x_1^{n-2} \\ 1 & x_2 & x_2^{n-2} \\ & & \\ 1 & x_{n-1} & x_{n-1}^{n-2} \end{vmatrix} = (-1)^{n+1}(x_1 - x_n)\dots(x_{n-1} - x_n)V_{n-1},$$

Czyli mamy zależność pomiędzy wyznacznikiem V_n stopnia n dla parametrów $x_1,\dots,x_n$ a analogicznym wyznacznikiem V_{n-1} stopnia $n-1$ dla parametrów $x_1,\dots,x_{n-1}$. Należy tę zależność stosować teraz iteracyjnie aż do wyznacznika $V_1 = 1$:

$$\begin{aligned} V_n &= (-1)^{n+1}(x_1 - x_n)\dots(x_{n-1} - x_n)V_{n-1} = (-1)^2(-1)^{n-1}(x_1 - x_n)\dots(x_{n-1} - x_n)V_{n-1} = \\ &= (x_n - x_1)\dots(x_n - x_{n-1})V_{n-1} = (x_n - x_1)\dots(x_n - x_{n-1})(x_{n-2} - x_1)\dots(x_{n-2} - x_{n-1})V_{n-2} = \\ &= \dots = \prod_{1\le i<j\le n}(x_j - x_i). \end{aligned}$$

Związek pomiędzy wyznacznikiem a operacją mnożenia macierzy

Okazuje się, że zachodzi bardzo fundamentalna zależność pomiędzy operacja mnożenia macierzy a wyznacznikiem macierzy. Jest to *Twierdzenie Cauchy'ego o wyznaczniku iloczynu macierzy*.

Niech A i B będą macierzami kwadratowymi tego samego stopnia. Wtedy

$$\det(AB) = \det A \cdot \det B, \tag{0.28}$$

czyli wyznacznik iloczynu macierzy jest iloczynem wyznaczników.

Jeżeli $A = B,$ to ze wzoru (0.28) mamy natychmiast $\det(A^2) = (\det A)^2$. Dalej $\det(A^3) = \det(A^2 A)$ $= \det(A^2)\det A = (\det A)^2 \det A = (\det A)^3$. Widać, że dla dowolnego $n \in \mathbb{N}$ mamy

$$\det(A^n) = (\det A)^n. \tag{0.29}$$

Zad. 11. Oblicz wyznaczniki:

$$\begin{vmatrix} 2 & 3 \\ 4 & 3 \end{vmatrix} = \quad \begin{vmatrix} a & b \\ b & a \end{vmatrix} = \qquad \begin{vmatrix} a^2+1 & ax+1 \\ ax+1 & x^2+1 \end{vmatrix} =$$

Zad. 12. Oblicz wyznaczniki i zapisz możliwie najprostszej postaci powstałe wyrażenia:

$$\begin{vmatrix} 0 & a & b \\ a & 0 & c \\ b & c & 0 \end{vmatrix} = \qquad \begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix} = \qquad \begin{vmatrix} 1 & a & b \\ 1 & a+x & b \\ 1 & a & b+y \end{vmatrix} = \qquad \begin{vmatrix} 1 & a & b \\ 1 & x & b \\ 1 & a & x \end{vmatrix} =$$

Zad. 13. Oblicz wyznaczniki, stosując w każdym przypadku rozwinięcie Laplace'a. Aby przekonać się, że wybór kolumny lub wiersza nie wpływa na wynik wykonaj w każdym przypadku dwa obliczenie: najpierw względem wybranej kolumny, a następnie względem jakiegoś wiersza.

$$\begin{vmatrix} 1 & 5 & 5 & 3 \\ 2 & 6 & 3 & 3 \\ 3 & 7 & 5 & 2 \\ 4 & 2 & 3 & 3 \end{vmatrix} = \qquad \begin{vmatrix} 0 & 2 & 3 & 3 \\ 1 & 6 & 2 & 3 \\ 2 & 0 & 0 & 1 \\ 3 & 2 & 0 & 2 \end{vmatrix} =$$

Zad. 14. Korzystając z wyznacznika Vandermonde'a obliczyć wyznacznik

$$\begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2^2 & 2^3 & 2^4 \\ 1 & 3 & 3^2 & 3^3 & 3^4 \\ 1 & 4 & 4^2 & 4^3 & 4^4 \\ 1 & 5 & 5^2 & 5^3 & 5^4 \end{vmatrix}.$$

Zad. 15. Udowodnić następującą równość dla wyznacznika stopnia n

$$\begin{vmatrix} 0 & 1 & 1 & \dots & 1 & 1 \\ 1 & x & a & \dots & a & a \\ 1 & a & x & \dots & a & a \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 1 & a & a & \dots & x & a \\ 1 & a & a & \dots & a & x \end{vmatrix} = -(n-1)(x-a)^{n-1}.$$

Zad 16. Dana jest macierz

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

Obliczyć $\det(A^5)$.

Macierz odwzorowania liniowego w bazie

Niech V i U będą dwoma przestrzeniami wektorowymi (nad tym samym zbiorem liczbowym $\mathbb{R}$ lub $\mathbb{C}$) oraz niech $T: U \to V$ będzie operatorem liniowym. Załóżmy, że w obu przestrzeniach zostały wybrane pewne bazy:

$u_1, u_2, \dots, u_n$ w przestrzeni U,

$v_1, v_2, \dots, v_m$ w przestrzeni V.

Jeżeli teraz obliczymy wartości operatora T na kolejnych wektorach bazowych $u_1, u_2, \ldots, u_n$, to możemy wynik $T(u_j) \in V$ rozkładać względem wektorów bazowych $v_1, v_2, \ldots, v_m$. Jeżeli współczynniki rozkładu ustawimy w j – tej kolumnie tablicy, to otrzymamy tzw. *macierz operatora* T w podanych bazach:

$$T(u_j) = a_{1j}v_1 + a_{2j}v_2 + \ldots + a_{mj}v_m \quad \text{dla} \quad j = 1, \ldots, n. \tag{0.30}$$

Zatem

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{0.31}$$

jest macierzą $m \times n$. Widzimy więc, że według przyjętej konwencji (współczynniki rozkładu jako kolejne kolumny) liczba wierszy macierzy A jest równa wymiarowi przeciwdziedziny operatora, a liczba kolumn jest równa wymiarowi dziedziny operatora.

Proszę zwrócić uwagę, że macierz odwzorowania na ogół zależy od wyboru bazy!

Najprościej wyznacza się macierz operatora liniowego w przypadku przestrzeni $U = \mathbb{R}^n$, $V = \mathbb{R}^m$ z wybranymi bazami kanonicznymi $\left(e_i = (0, \ldots 0, 1, 0, \ldots, 0)\right)$. Wynika to stąd, że współrzędne wektora w bazie kanonicznej pokrywają się z jego składowymi, na przykład

$$\mathbb{R}^3 \ni (3, -1, 2) = 3(1,0,0) - 1(0,1,0) + 2(0,0,2) = 3e_1 - e_2 + 2e_3,$$

zatem współrzędne wektora $u = (3, -1, 2)$ względem bazy kanonicznej wynoszą $3, -1, 2$.

Przykład. Dany jest operator liniowy $T : \mathbb{R}^3 \to \mathbb{R}^2$ wzorem

$$T(x, y, z) = (2x - 3y + 4z,\ 5y + z).$$

Jaka jest jego macierz w bazach kanonicznych.

Rozwiązanie: W obu przestrzeniach wybieramy bazy kanoniczne, a więc

$$\begin{aligned} &e_1 = (1,0,0),\ e_2 = (0,1,0),\ e_3 = (0,0,1) \ \text{ w } \ \mathbb{R}^3, \\ &f_1 = (1,0),\ f_2 = (0,1) \ \text{ w } \ \mathbb{R}^2. \end{aligned}$$

Musimy teraz znaleźć współczynniki (współrzędne) rozkładu wektorów $T(e_i)$ względem wektorów f_j i ustawić je kolumnami w szukanej macierzy A. Mamy zatem

$$\begin{aligned} T(e_1) &= T(1,0,0) = (2 \cdot 1 - 3 \cdot 0 + 4 \cdot 0,\ 5 \cdot 0 + 0) = (2,\ 0) = 2f_1 + 0f_2, \\ T(e_2) &= T(0,1,0) = (2 \cdot 0 - 3 \cdot 1 + 4 \cdot 0,\ 5 \cdot 1 + 0) = (-3,\ 5) = -3f_1 + 5f_2, \\ T(e_3) &= T(0,0,1) = (2 \cdot 0 - 3 \cdot 0 + 4 \cdot 1,\ 5 \cdot 0 + 1) = (4,\ 1) = 4f_1 + f_2, \end{aligned}$$

co daje macierz

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

Jeżeli teraz się dokładnie przyjrzymy wzorowi na operator T, to zauważymy, że macierz powyższa powstała ze współczynników $T(x,y,z)=(2x-3y+4z,\, 5y+z)$ kolejnych składowych umieszczanych w wierszach (oczywiście kolejność zmiennych we wzorze musi być $x,y,\ldots$). Na przykład dla operatora liniowego $T:\mathbb{R}^4\to\mathbb{R}^3$ określonego wzorem

$$T(x_1,x_2,x_3,x_4)=(2x_1-4x_2+5x_4,-x_1+x_2+x_3,2x_1+x_2+6x_3+7x_4),$$

macierz w bazach kanonicznych jest następująco

$$A=\begin{bmatrix} 2 & -4 & 0 & 5 \\ -1 & 1 & 1 & 0 \\ 2 & 1 & 6 & 7 \end{bmatrix}.$$

Możemy te obserwacje uogólnić na przypadek ogólny operatora liniowego $T:\mathbb{R}^n\to\mathbb{R}^m$. Jeżeli operator taki jest dany wzorem

$$T(x_1,\ldots,x_n)=(a_{11}x_1+\ldots+a_{1n}x_n,\ldots,a_{m1}x_1+\ldots+a_{mn}x_n), \tag{0.32}$$

to jego macierzą w bazach kanonicznych jest

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}. \tag{0.33}$$

Zauważmy, że zgodnie z definicją mnożenia macierzy możemy wyrażenie (0.32) na operator liniowy uzyskać z macierzy (0.33) następująco

$$T(x_1,\ldots,x_n)=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\cdot\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \tag{0.34}$$

czyli w zwartej postaci

$$T(x)=A\cdot x, \tag{0.35}$$

gdzie wektor x zapisujemy w formie kolumny tak, jak we wzorze (0.34).[5] Dlatego w zastosowaniach często utożsamia się macierz z operatorem liniowym rozumiejąc przez to, że jak mamy daną macierz (0.33), to możemy skojarzyć z nią operator liniowy wg wzoru (0.35). Należy jednak pamiętać, że zależność ta jest prawdziwa jeżeli macierz odwzorowania obliczana jest w bazach kanonicznych przestrzeni kartezjańskich $\mathbb{R}^n$ oraz $\mathbb{R}^m$.

Przykład. Dana jest w bazach kanonicznych macierz

$$A = \begin{bmatrix} 3 & 1 & -2 \\ 1 & 0 & 5 \end{bmatrix},$$

operatora liniowego T. Podaj wzór tego operatora.

Rozwiązanie: Ponieważ macierz operatora liniowego jest obliczona w bazach kanonicznych, więc możemy zastosować wzór (0.34). Zatem

$$T(x) = \begin{bmatrix} 3 & 1 & -2 \\ 1 & 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_2 \end{bmatrix} = (3x_1 + x_2 - 2x_3, x_1 + 0x_2 + 5x_3),$$

czyli $T(x_1, x_2, x_3) = (3x_1 + x_2 - 2x_3, x_1 + 5x_3)$. Widać też, że $T : \mathbb{R}^3 \to \mathbb{R}^2$.

Obrót na płaszczyźnie

Jeżeli punkty płaszczyzny będziemy utożsamiali z dwuwymiarową przestrzenią kartezjańską $\mathbb{R}^2$, to każdemu przekształceniu geometrycznemu na płaszczyźnie odpowiada pewne odwzorowanie $F : \mathbb{R}^2 \to \mathbb{R}^2$. W szczególności obrót o kąt φ wokół środka układu współrzędnych $O = (0,0) \in \mathbb{R}^2$ w kierunku przeciwnym do ruchu wskazówek zegara polega na przekształceniu $(x, y) \to (x', y')$ opisanym wzorami

$$\begin{cases} x' = x\cos\varphi - y\sin\varphi, \\ y' = x\sin\varphi + y\cos\varphi, \end{cases} \tag{0.36}$$

co można wyrazić odwzorowaniem $T_\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ o postaci

$$T_\varphi(x, y) = (x\cos\varphi - y\sin\varphi,\ x\sin\varphi + y\cos\varphi). \tag{0.37}$$

Widać, że jest to operator liniowy, a macierz tego odwzorowania w bazie kanonicznej $e_1 = (1,0),\ e_2 = (0,1)$ ma postać

[5] W zasadzie jest to trochę nieprecyzyjne. W większości książek, gdzie podobnie jak w tym skrypcie wektory z $\mathbb{R}^n$ są zapisywane poziomo (wierszowo): $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, wzór (0.35) jest raczej zapisywany z wykorzystaniem operacji transponowania, czyli $T(x) = A \cdot x^T$, gdyż x^T jest kolumną.

$$A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}. \tag{0.38}$$

Zad. 1. Jaka jest macierz podanych odwzorowań liniowych w bazach kanonicznych przestrzeni $\mathbb{R}^k$?

a) $T(x, y) = (2x+3y,\ 4x-5y)$

b) $T(x, y) = 3x+7y$

c) $T(x, y, z) = (-x+z,\ 2x+6y,\ 4x+7y-2z)$

d) $T(x, y, z) = (3y-z,\ 2x)$

e) $T(x, y, z) = 3x-y+5x$

Zad. 2. Jak jest macierz odwzorowania $T:\mathbb{R}^2 \to \mathbb{R}^2$

$$T(x, y) = (2x+3y,\ 4x-5y),$$

w bazie $b_1 = (1, 1), b_2 = (-2, 1)$. Wybieramy tę samą bazę w dziedzinie i przeciwdziedzinie $(U = V = \mathbb{R}^2)$.

Zad. 4. Czy istnieje operator liniowy, który ma tą samą macierz we wszystkich bazach? Jeżeli tak, to podaj jakieś przykłady.

Zad. 5. Podaj macierze następujących operatorów liniowych w przestrzeni wektorowej $\mathbb{R}^2$ z wybraną bazą kanoniczną:

a) Obrót o kąt φ w kierunku przeciwnym do ruchu wskazówek zegara

b) Odbicie symetryczne względem osi OY

c) Jednokładność o skali λ i środku w początku układu współrzędnych

d) Odbicie symetryczne względem prostej o równaniu $y = ax$

e) Rzut na prostą OX

f) Rzut na prostą równaniu $y = ax$

Zad. 6. Niech $T:\mathbb{R}^2 \to \mathbb{R}^2$ będzie operatorem liniowym, który jest odbiciem symetrycznym względem prostej o równaniu $y = ax$. Udowodnić, że T ma następującą postać:

$$T(x, y) = \frac{1}{1+a^2}((1-a^2)x + 2ay,\ 2ax - (1-a^2)y)$$

Zad. 7. Podaj macierze następujących operatorów liniowych w bazach kanonicznych:

a) $T(x_1, x_2, x_3, x_4, x_5) = (2x_1 + 3x_4, -x_3 + x_5, x_4 + x_5, x_1, 2x_3 + x_4 - x_5, x_3 + x_4, x_1 + x_4)$

b) $T(x_1, \ldots, x_n) = (x_1 + x_2, x_2 + x_3, \ldots, x_{n-1} + x_n, x_n + x_1)$