Algebra liniowa

Jerzy Topp

Politechnika Gdańska 2005

Spis treści

PRZEDMOWA 5

Rozdział 1. PODSTAWOWE STRUKTURY ALGEBRAICZNE 7

1.1. Działania i ich własności 7
1.2. Grupa i jej podgrupy 10
1.3. Pierścień 14
1.4. Ciało 16
1.5. Ćwiczenia 17

Rozdział 2. LICZBY ZESPOLONE 19

2.1. Liczby zespolone i działania na liczbach zespolonych 19
2.2. Sprzężenie i moduł liczby zespolonej 23
2.3. Postać trygonometryczna liczby zespolonej 25
2.4. Pierwiastkowanie liczb zespolonych 29
2.5. Wzory Eulera 33
2.6. Postać wykładnicza liczby zespolonej 35
2.7. Ćwiczenia 36

Rozdział 3. WIELOMIANY 39

3.1. Pierścień wielomianów 39
3.2. Podzielność wielomianów 42
3.3. Schemat Hornera 44
3.4. Pierwiastki wielomianów 46
3.5. Wielomiany względnie pierwsze 53
3.6. Funkcje wymierne i ułamki proste 54
3.7. Ćwiczenia 62

Rozdział 4. MACIERZE 64

4.1. Podstawowe definicje 64
4.2. Działania na macierzach 66
4.3. Macierz odwrotna 74
4.4. Ślad macierzy kwadratowej 76
4.5. Ćwiczenia 77

Rozdział 5. UKŁADY RÓWNAŃ LINIOWYCH 80

5.1. Podstawowe definicje i fakty 80
5.2. Równania macierzowe 92
5.3. Kolejne własności macierzy odwracalnej 94
5.4. Wyznaczanie macierzy odwrotnej 95
5.5. Struktura rozwiązań układu równań liniowych 97
5.6. Ćwiczenia 99

Rozdział 6. WYZNACZNIKI 102

6.1. Definicja i pierwsze własności wyznacznika 102
6.2. Wyznacznik iloczynu macierzy 111
6.3. Macierze odwracalne i nieosobliwe 112
6.4. Wyznacznik macierzy podobnych 114
6.5. Układy równań i wzory Cramera 115
6.6. Ćwiczenia 117

Rozdział 7. PRZESTRZEŃ WEKTOROWA 120

7.1. Przestrzeń wektorowa i jej podprzestrzenie 120
7.2. Kombinacje liniowe wektorów 125
7.3. Przestrzeń kolumnowa macierzy 128
7.4. Liniowa zależność i liniowa niezależność wektorów 131
7.5. Baza przestrzeni wektorowej 135
7.6. Współrzędne wektora 141
7.7. Rząd macierzy 149

7.8. Suma i suma prosta podprzestrzeni 152
7.9. Ćwiczenia 155

Rozdział 8. PRZEKSZTAŁCENIA LINIOWE 159
8.1. Definicja przekształcenia liniowego 159
8.2. Jądro i obraz przekształcenia liniowego 164
8.3. Mono- i epimorficzność przekształcenia liniowego 169
8.4. Suma i złożenie przekształceń liniowych 171
8.5. Macierz przekształcenia liniowego 173
8.6. Odwracalność odwzorowania liniowego 180
8.7. Podobieństwo macierzy 183
8.8. Ćwiczenia 186

Rozdział 9. ILOCZYN SKALARNY I ORTOGONALNOŚĆ WEKTORÓW 191
9.1. Definicja i przykłady iloczynów skalarnych 191
9.2. Kąt pomiędzy wektorami 195
9.3. Ortogonalizacja bazy 198
9.4. Dopełnienie ortogonalne 200
9.5. Rzut ortogonalny 201
9.6. Macierz rzutu ortogonalnego 204
9.7. Metoda najmniejszych kwadratów 206
9.8. Najlepsze rozwiązanie układu równań 207
9.9. Dopasowanie prostej 208
9.10. Macierz i przekształcenie ortogonalne 210
9.11. Ćwiczenia 213

Rozdział 10. WARTOŚCI WŁASNE I WEKTORY WŁASNE 216
10.1. Wartości własne i wektory własne macierzy i operatora 216
10.2. Diagonalizowalność macierzy i operatora liniowego 221
10.3. Diagonalizacja macierzy symetrycznej 228
10.4. Potęga macierzy diagonalizowalnej 232
10.5. Granica ciągu macierzy 233
10.6. Podprzestrzenie niezmiennicze 236
10.7. Twierdzenie Cayleya-Hamiltona 238
10.8. Zależności rekurencyjne 242
10.9. Ćwiczenia 245

Rozdział 11. FORMY KWADRATOWE 250
11.1. Rzeczywista forma kwadratowa 250
11.2. Postać kanoniczna formy kwadratowej 252
11.3. Określoność macierzy i formy kwadratowej 259
11.4. Ćwiczenia 264

Rozdział 12. ELEMENTY GEOMETRII ANALITYCZNEJ 265
12.1. Iloczyn wektorowy wektorów 265
12.2. Iloczyn mieszany wektorów 267
12.3. Prosta i płaszczyzna 269
12.4. Ćwiczenia 281

Skorowidz 284

Bibliografia 288

PRZEDMOWA

Niniejszy skrypt jest pierwszą wersją zbioru notatek do wykładów algebry liniowej prowadzonych dla studentów pierwszego roku na Wydziale Elektroniki, Telekomunikacji i Informatyki Politechniki Gdańskiej.

Treści skryptu podzielono na 12 rozdziałów i obejmują one następujące tematy: podstawowe struktury algebraiczne, liczby zespolone, wielomiany, macierze, układy równań liniowych, wyznaczniki, przestrzenie wektorowe, przekształcenia liniowe, wartości własne, formy kwadratowe i elementy geometrii analitycznej. Ze względów praktycznych niektóre z tych rozdziałów rozbudowano. Dotyczy to m.in. rozdziału zawierającego przestrzenie wektorowe z iloczynem skalarnym oraz rozdziału poświęconego wartościom własnym i wektorom własnym. Pewne części tych rozdziałów pozostawiamy Czytelnikowi do samodzielnej lektury.

Teorię przedstawiono w skrypcie w sposób ścisły, dowodząc prawie wszystkich twierdzeń. Język i notację dobrano w taki sposób, aby całość była bardzo czytelna. Skrypt zawiera wielką liczbę rozwiązanych przykładów. Ilustrują one ważniejsze pojęcia i twierdzenia. Tam gdzie było to możliwe, pojęcia i zależności pomiędzy rozważanymi pojęciami zilustrowano rysunkami. Powinno to ułatwić czytanie i zrozumienie przedstawionego tekstu.

Każdy rozdział kończy się dużą liczbą stosunkowo prostych zadań, których rozwiązanie powinno doprowadzić Czytelnika do pełniejszego zrozumienia wcześniejszych definicji i twierdzeń oraz do osiągnięcia niezbędnej biegłości myślowej i rachunkowej.

Czytelnikowi pozostawia się wybór sposobu korzystania z tego skryptu i wybór sposobu uczenia się języka algebry liniowej. Warto jedynie przypomnieć, że nauka języka algebry liniowej (tak jak i nauka każdego języka obcego) wymaga umiejętności słuchania, mówienia, czytania i pisania. Pełne opanowanie materiału przedstawionego w skrypcie wymaga starannego nauczenia się definicji, poznania dokładnych sformułowań twierdzeń i zrozumienia ich dowodów oraz wyćwiczenia w sobie umiejętności rozwiązywania zadań.

Autor jest świadom faktu, że w tym skrypcie występować mogą niedoskonałości i usterki. Wszelkie uwagi o skrypcie i informacje o zauważonych usterkach prosimy przesłać na adres `topp@mif.pg.gda.pl`. Pełna informacja o poprawionych fragmentach dostępna będzie na stronie internetowej `www.mif.pg.gda.pl/topp`. Tam też znajdą się wskazówki i/lub odpowiedzi do zadań przedstawionych w tym skrypcie.

Gdynia, lipiec 2005 Jerzy Topp

Rozdział 1

PODSTAWOWE STRUKTURY ALGEBRAICZNE

1.1. Działania i ich własności

Definicja 1.1.1. *Działaniem dwuargumentowym* (krótko, *działaniem*) w niepustym zbiorze X nazywamy każdą funkcję

$$f\colon X \times X \to X.$$

Działanie w zbiorze X

Działanie f przyporządkowuje każdej uporządkowanej parze (x, y) elementów zbioru X jednoznacznie wyznaczony element $f(x, y)$ zbioru X, nazywany *wynikiem działania* f na uporządkowanej parze (x, y) elementów zbioru X. Zwykle zamiast $f(x, y)$ pisze się xfy, $x * y$, $x \bullet y$, $x + y$ (lub używa się jeszcze innych symboli na oznaczenie działania f i jego wyniku $f(x, y)$).

Wynik działania

Przykład 1. Działaniem w zbiorze liczb rzeczywistych R jest funkcja $\star : R \times R \to R$ określona za pomocą zwykłego dodawania, zwykłego odejmowania i zwykłego mnożenia liczb rzeczywistych i taka, że

$$x \star y = x + y - 2xy \tag{1.1}$$

dla każdej pary $(x, y) \in R \times R$.

Definicja 1.1.2. Mówimy, że działanie $*$ w zbiorze X jest *przemienne*, gdy

$$x * y = y * x$$

Przemienność działania

dla każdych dwóch elementów x i y zbioru X. Jeśli natomiast dla każdych trzech elementów x, y, $z \in X$ mamy

$$x * (y * z) = (x * y) * z,$$

Łączność działania

to mówimy, że działanie $*$ jest *łączne* w zbiorze X.

Przykład 2. Zwykłe dodawanie liczb rzeczywistych jest działaniem przemiennym i łącznym. Podobnie, zwykłe mnożenie liczb rzeczywistych jest przemienne i łączne. Natomiast odejmowanie liczb rzeczywistych nie jest ani przemienne, ani łączne. Z przemienności zwykłego dodawania i mnożenia liczb rzeczywistych wynika, że działanie $\star : R \times R \to R$ określone wzorem (1.1) jest przemienne, bo dla każdych liczb x, $y \in R$ mamy

$$x \star y = x + y - 2xy = y + x - 2yx = y \star x.$$

Działanie $\star$ jest także łączne, bo dla każdych liczb x, y, $z \in R$ mamy

$$\begin{aligned}
x \star (y \star z) &= x \star (y + z - 2yz) \\
&= x + (y + z - 2yz) - 2x(y + z - 2yz) \\
&= x + y + z - 2yz - 2xy - 2xz + 4xyz \\
&= (x + y - 2xy) + z - 2(x + y - 2xy)z \\
&= (x + y - 2xy) \star z \\
&= (x \star y) \star z.
\end{aligned}$$

Element neutralny

Definicja 1.1.3. Niech $*$ będzie działaniem w zbiorze X. Element e należący do zbioru X nazywamy *elementem neutralnym* działania $*$, gdy dla każdego $x \in X$ jest

$$e * x = x * e = x.$$

Przykład 3. Liczba 1 jest elementem neutralnym zwykłego mnożenia liczb rzeczywistych, bo $x \cdot 1 = 1 \cdot x = x$. Liczba 0 jest elementem neutralnym zwykłego dodawania liczb rzeczywistych, bo $x + 0 = 0 + x = x$.

Niech teraz $\circ$ będzie działaniem w zbiorze R, gdzie

$$x \circ y = x + y + 3. \tag{1.2}$$

Liczba -3 jest elementem neutralnym działania $\circ$, bo dla każdego $x \in R$ mamy

$$x \circ (-3) = x + (-3) + 3 = x \quad \text{i} \quad (-3) \circ x = (-3) + x + 3 = x.$$

Twierdzenie 1.1.1. *Dla dowolnego działania $*$ w zbiorze X istnieje co najwyżej jeden element neutralny działania $*$.*

Dowód. Niech e i e' będą elementami neutralnymi działania $*$. Z definicji 1.1.3 mamy wtedy

$$e * e' = e' \quad \text{i} \quad e * e' = e.$$

Stąd $e' = e$ i to dowodzi, że istnieje co najwyżej jeden element neutralny działania $*$. □

Definicja 1.1.4. Załóżmy, że e jest elementem neutralnym działania $*$ w zbiorze X. Mówimy, że element x zbioru X jest *odwracalny* (względem działania $*$), gdy istnieje element $y \in X$ taki, że

Odwracalność elementu

$$x * y = y * x = e. \tag{1.3}$$

Element odwrotny

Element przeciwny

Element y o powyższych własnościach nazywamy *elementem odwrotnym* (albo *przeciwnym*) do elementu x (względem działania $*$) i zwykle oznaczamy go przez x^{-1} (albo przez $-x$). Z równości (1.3), czyli z równości $x * x^{-1} = x^{-1} * x = e$, wynika, że element odwrotny x^{-1} także jest odwracalny i dla niego dodatkowo mamy

$$(x^{-1})^{-1} = x \quad \text{(i odpowiednio} \quad -(-x) = x). \tag{1.4}$$

Twierdzenie 1.1.2. *Niech e będzie elementem neutralnym działania $*$ w zbiorze X. Jeśli działanie $*$ jest łączne, to dowolny element x zbioru X ma co najwyżej jeden element odwrotny.*

Dowód. Niech y i y' będą elementami odwrotnymi do elementu x względem działania $*$. Ponieważ $y * x = e = x * y'$, więc mamy

$$y = y * e = y * (x * y') = (y * x) * y' = e * y' = y'$$

i to dowodzi, że x ma co najwyżej jeden element odwrotny. □

Przykład 4. Elementem odwrotnym do elementu $x \in R - \{0\}$ względem zwykłego mnożenia liczb rzeczywistych jest liczba $1/x$.

Niech teraz x_0 będzie ustaloną liczbą rzeczywistą i niech $\circ$ będzie takim działaniem w zbiorze R, że dla każdych x, $y \in R$ jest

$$x \circ y = x + y - x_0.$$

Łatwo zauważyć, że liczba x_0 jest elementem neutralnym działania $\circ$ (zob. (1.2) dla $x_0 = -3$). Każdy element $x \in R$ jest odwracalny (względem działania $\circ$)

i elementem odwrotnym do x jest $-x + 2x_0$, bo $x \circ (-x + 2x_0) = x + (-x + 2x_0) - x_0 = x_0$ i podobnie $(-x + 2x_0) \circ x = x_0$.

Definicja 1.1.5. Niech $*$ i $\circ$ będą działaniami w zbiorze X. Mówimy, że działanie $\circ$ jest *lewostronnie rozdzielne* względem działania $*$, gdy dla dowolnych x, y, $z \in X$ jest

$$x \circ (y * z) = (x \circ y) * (x \circ z). \tag{1.5}$$

Lewostronna rozdzielność

Działanie $\circ$ jest *prawostronnie rozdzielne* względem działania $*$, gdy dla dowolnych x, y, $z \in X$ jest

$$(y * z) \circ x = (y \circ x) * (z \circ x). \tag{1.6}$$

Prawostronna rozdzielność

Mówimy, że działanie $\circ$ jest *rozdzielne względem działania* $*$, gdy jest ono jednocześnie lewo- i prawostronnie rozdzielne względem działania $*$. Z (1.5) i (1.6) łatwo wynika, że jeśli działanie $\circ$ jest przemienne, to jest ono rozdzielne względem działania $*$ pod warunkiem, że jest ono lewostronnie (lub prawostronnie) rozdzielne względem działania $*$.

Przykład 5. W zbiorze R dane są działania $*$ i $\circ$ takie, że

$$x * y = x + y + 2 \quad \text{i} \quad x \circ y = \frac{x+y}{2}.$$

Zbadać: (a) rozdzielność działania $*$ względem działania $\circ$; (b) rozdzielność działania $\circ$ względem działania $*$.

Ponieważ działania $*$ i $\circ$ są przemienne, wystarczy zbadać ich lewostronne rozdzielności.

(a) Dla każdych trzech liczb x, y, $z \in R$ mamy

$$x * (y \circ z) = x * \left(\frac{y}{2} + \frac{z}{2}\right) = x + \frac{y}{2} + \frac{z}{2} + 2$$

i

$$(x * y) \circ (x * z) = (x + y + 2) \circ (x + z + 2) = x + \frac{y}{2} + \frac{z}{2} + 2.$$

Stąd i z definicji 1.1.5 wynika rozdzielność działania $*$ względem działania $\circ$.

(b) Dla każdych trzech liczb x, y, $z \in R$ jest

$$x \circ (y * z) = x \circ (y + z + 2) = \frac{x}{2} + \frac{y}{2} + \frac{z}{2} + 1,$$

ale

$$(x \circ y) * (x \circ z) = \left(\frac{x}{2} + \frac{y}{2}\right) * \left(\frac{x}{2} + \frac{z}{2}\right) = x + \frac{y}{2} + \frac{z}{2} + 2$$

i dlatego działanie $\circ$ nie jest rozdzielne względem działania $*$.

Definicja 1.1.6. Niech $*$ będzie działaniem dwuargumentowym w zbiorze X i niech Y będzie niepustym podzbiorem zbioru X. Mówimy, że zbiór Y jest *zamknięty ze względu na działanie* $*$, gdy

$$x * y \in Y \quad \text{dla każdych} \quad x,\, y \in Y.$$

Przykład 6. Niech $*$ będzie działaniem dwuargumentowym na zbiorze R, gdzie

$$x * y = |x| - y \quad \text{dla} \quad (x, y) \in R \times R.$$

Zbiór liczb naturalnych $N \subseteq R$ nie jest zamknięty ze względu na działanie $*$, bo przykładowo $x = 4 \in N$ i $y = 5 \in N$, ale

$$x * y = 4 * 5 = |4| - 5 = -1 \notin N.$$

1.2. Grupa i jej podgrupy

Grupa

Definicja 1.2.1. Niech $\circ$ będzie działaniem w niepustym zbiorze G. Parę $(G, \circ)$, czyli system algebraiczny $(G, \circ)$, nazywamy *grupą*, gdy:

(G_1) działanie $\circ$ jest łączne w zbiorze G, czyli dla każdych elementów x, y i z ze zbioru G jest
$$x \circ (y \circ z) = (x \circ y) \circ z;$$

(G_2) w zbiorze G istnieje element neutralny e działania $\circ$, czyli taki element, że dla każdego x ze zbioru G jest
$$x \circ e = e \circ x = x;$$

(G_3) każdy element zbioru G jest odwracalny względem działania $\circ$, czyli dla każdego $x \in G$ istnieje element $x^{-1} \in G$ taki, że
$$x \circ x^{-1} = x^{-1} \circ x = e.$$

Grupa przemienna

Jeśli $(G, \circ)$ jest grupą i działanie $\circ$ jest przemienne w zbiorze G, to mówimy, że grupa $(G, \circ)$ jest *przemienna* lub *abelowa*.

W naszych rozważaniach grupę $(G, \circ)$ i zbiór jej elementów G oznaczać będziemy tym samym symbolem (zwykle literą G). Ufamy, że nie doprowadzi to do nieporozumień. Analogicznie będziemy czynić w przypadku innych systemów algebraicznych.

Przykład 7. Zbiór liczb rzeczywistych R ze zwykłym dodawaniem $+$ tworzy grupę przemienną $(R, +)$. Także para $(Z, +)$, gdzie Z jest zbiorem liczb całkowitych, jest grupą przemienną. Struktura $(Z, \cdot)$, gdzie $\cdot$ jest zwykłym mnożeniem, nie jest grupą (bo prawie żaden element zbioru Z nie jest odwracalny ze względu na mnożenie $\cdot$). Zbiór liczb wymiernych Q ze zwykłym mnożeniem, czyli para $(Q, \cdot)$, także nie jest grupą, bo liczba 0 nie jest elementem odwracalnym. Natomiast para $(Q - \{0\}, \cdot)$ (jak i para $(R - \{0\}, \cdot)$) jest już grupą i jest to grupa przemienna.

Przykład 8. Niech X będzie niepustym zbiorem i niech $\mathcal{F} = \mathcal{F}(X)$ będzie zbiorem wszystkich bijekcji na zbiorze X, czyli zbiorem wszystkich odwzorowań wzajemnie jednoznacznych zbioru X na siebie. Jeśli $f\colon X \to X$ i $g\colon X \to X$ są bijekcjami, to ich złożenie $g \circ f$ (będące funkcją $g \circ f\colon X \to X$ określoną wzorem $(g \circ f)(x) = g(f(x))$) też jest bijekcją. Składanie odwzorowań jest łączne: dla każdych trzech f, g, $h \in \mathcal{F}$ jest $f \circ (g \circ h) = (f \circ g) \circ h$, bo dla każdego $x \in X$ mamy
$$\begin{aligned}\big(f \circ (g \circ h)\big)(x) &= f\big((g \circ h)(x)\big) = f\big(g(h(x))\big)\\ &= (f \circ g)\big(h(x)\big) = \big((f \circ g) \circ h\big)(x).\end{aligned}$$

Odwzorowanie tożsamościowe I_X zbioru X (czyli funkcja $I_X\colon X \to X$ taka, że $I_X(x) = x$ dla każdego $x \in X$) jest bijekcją na zbiorze X. Ponieważ dla każdego $f \in \mathcal{F}$ i każdego $x \in X$ mamy
$$(f \circ I_X)(x) = f\big(I_X(x)\big) = f(x) \quad \text{i} \quad (I_X \circ f)(x) = I_X\big(f(x)\big) = f(x),$$

więc $f \circ I_X = f = I_X \circ f$ i to dowodzi, że odwzorowanie tożsamościowe I_X jest elementem neutralnym złożenia $\circ$. Dla każdej bijekcji $f \in \mathcal{F}$ istnieje odwzorowanie odwrotne $f^{-1}\colon X \to X$ (czyli takie, że $f \circ f^{-1} = I_X = f^{-1} \circ f$), które też jest bijekcją na zbiorze X. Stąd wynika, że para $(\mathcal{F}, \circ)$ jest grupą, jest to tzw. grupa symetryczna zbioru X. Bez problemu można zauważyć, że grupa $(\mathcal{F}, \circ)$ jest nieprzemienna, gdy zbiór X ma co najmniej trzy elementy (zob. zadanie 4).

Twierdzenie 1.2.1 (o skracaniu w grupie). *W grupie $(G, \circ)$ dla dowolnych elementów a, b, $c \in G$ jest:*

(a) *jeśli $a \circ b = a \circ c$, to $b = c$;*

(b) *jeśli $a \circ b = c \circ b$, to $a = c$.*

Dowód. Załóżmy, że $a \circ b = a \circ c$. Wtedy także $a^{-1} \circ (a \circ b) = a^{-1} \circ (a \circ c)$. Jednocześnie z łączności działania $\circ$ oraz z własności elementu odwrotnego a^{-1} i elementu neutralnego e jest

$$a^{-1} \circ (a \circ b) = (a^{-1} \circ a) \circ b = e \circ b = b$$

oraz

$$a^{-1} \circ (a \circ c) = (a^{-1} \circ a) \circ c = e \circ c = c$$

i stąd wynika implikacja (a). Analogicznie dowodzi się implikację (b). □

Twierdzenie 1.2.2. *W grupie $(G, \circ)$ dla dowolnych elementów a, $b \in G$ jest*

$$(a \circ b)^{-1} = b^{-1} \circ a^{-1}.$$

Dowód. Z łączności działania $\circ$ oraz z własności elementu odwrotnego i elementu neutralnego jest

$$(a \circ b) \circ (b^{-1} \circ a^{-1}) = a \circ (b \circ b^{-1}) \circ a^{-1} = a \circ e \circ a^{-1} = e$$

oraz

$$(b^{-1} \circ a^{-1}) \circ (a \circ b) = b^{-1} \circ (a^{-1} \circ a) \circ b = b^{-1} \circ e \circ b = e.$$

Stąd i z definicji 1.1.4 wynika teza twierdzenia. □

Definicja 1.2.2. W grupie $(G, \circ)$ definiujemy całkowitą potęgę elementu $x \in G$, przyjmując, że

$$x^0 = e \quad \text{i} \quad x^{n+1} = x^n \circ x \quad \text{oraz} \quad x^{-n} = (x^n)^{-1}$$

dla każdego naturalnego n.

Analogicznie definiujemy całkowitą krotność elementu x grupy G z działaniem "dodawania" $\oplus$:

$$0x = 0 \quad \text{i} \quad (n+1)x = nx \oplus x \quad \text{oraz} \quad (-n)x = -(nx)$$

dla każdego naturalnego n.

Korzystając z tej definicji oraz z twierdzenia 1.2.2, można udowodnić następujące twierdzenie.

Twierdzenie 1.2.3. *W grupie $(G, \circ)$ dla każdego $x \in G$ i każdych liczb całkowitych m i n jest*

$$x^m \circ x^n = x^{m+n} \quad \textit{oraz} \quad (x^m)^n = x^{mn}.$$

Dodawanie i mnożenie modulo n

Niech $n > 1$ będzie liczbą naturalną. Jeśli x jest liczbą całkowitą, to przez $[x]_n$ oznaczamy resztę z dzielenia x przez n. Dla liczb całkowitych x i y niech l_x, l_y, r_x i r_y będą liczbami całkowitymi takimi, że

$$x = nl_x + r_x, \quad y = nl_y + r_y \quad \text{i} \quad 0 \leqslant r_x,\, r_y < n.$$

Wtedy

$$[x+y]_n = \left[\, n(l_x + l_y) + r_x + r_y \,\right]_n = [r_x + r_y]_n$$

oraz

$$\left[\, x + [y]_n \,\right]_n = [\, nl_x + r_x + r_y \,]_n = [r_x + r_y]_n$$

i to dowodzi, że dla każdych x, $y \in Z$ mamy

$$\big[\, x + [y]_n \,\big]_n = [x + y]_n. \tag{1.7}$$

Analogicznie uzasadnia się, że dla każdych liczb całkowitych x i y mamy

$$\big[\, x \cdot [y]_n \,\big]_n = [x \cdot y]_n. \tag{1.8}$$

W zbiorze

$$Z_n = \{0, 1, 2, \ldots, n-1\},$$

który jest zbiorem wszystkich reszt z dzielenia liczb całkowitych przez liczbę n, określamy *dodawanie* $\oplus_n$ i *mnożenie* $\otimes_n$ *modulo* n, przyjmując, że dla liczb x i y ze zbioru Z_n mamy

Dodawanie modulo n

$$x \oplus_n y = [x + y]_n \tag{1.9}$$

i

Mnożenie modulo n

$$x \otimes_n y = [x\, y]_n. \tag{1.10}$$

Wyniki działań $\oplus_4$ i $\otimes_4$ w zbiorze Z_4 przedstawiają odpowiednio następujące dwie tabelki, w których $x \oplus_4 y$ (i $x \otimes_4 y$) umieszczono na przecięciu wiersza oznaczonego przez x i kolumny oznaczonej przez y:

$\oplus_4$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

i

$\otimes_4$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

$(Z_n, \oplus_n)$ – grupa reszt modulo n

Twierdzenie 1.2.4. $(Z_n, \oplus_n)$ *jest grupą przemienną.*

Dowód. Wobec (1.7) dla każdych x, y, $z \in Z_n$ jest

$$\big[x + [y+z]_n\big]_n = [x + (y+z)]_n \quad \text{i} \quad [(x+y)+z]_n = \big[[x+y]_n + z\big]_n$$

i stąd wynika, że działanie $\oplus_n$ jest łączne, bo mamy

$$\begin{aligned} x \oplus_n (y \oplus_n z) &= x \oplus_n [y+z]_n = \big[x + [y+z]_n\big]_n \\ &= [x + (y+z)]_n = [(x+y)+z]_n \\ &= \big[[x+y]_n + z\big]_n = [x+y]_n \oplus_n z = (x \oplus_n y) \oplus_n z. \end{aligned}$$

Elementem neutralnym działania $\oplus_n$ jest 0, bo dla każdego $x \in Z_n$ mamy

$$x \oplus_n 0 = [x+0]_n = [x]_n = x.$$

Łatwo zaobserwować, że elementem przeciwnym do $x \in Z_n$ jest

$$-x = \begin{cases} 0, & \text{gdy } x = 0, \\ n - x, & \text{gdy } x \neq 0. \end{cases}$$

Działanie $\oplus_n$ jest przemienne, bo dla każdych liczb x, $y \in Z_n$ mamy

$$x \oplus_n y = [x+y]_n = [y+x]_n = y \oplus_n x. \ \square$$

Ponieważ dla każdego $x \in Z_n$ mamy $1 \otimes_n x = [1 \cdot x]_n = [x]_n = x$, więc liczba 1 jest elementem neutralnym mnożenia $\otimes_n$ w Z_n. Zauważmy także, że mnożenie $\otimes_n$ jest przemienne, bo dla każdych x, $y \in Z_n$ mamy

$$x \otimes_n y = [x \cdot y]_n = [y \cdot x]_n = y \otimes_n x. \tag{1.11}$$

Korzystając z własności (1.8), bez problemu pokazuje się, że mnożenie $\otimes_n$ jest łączne w zbiorze Z_n (zob. zadanie 5). Jednakże struktura $(Z_n, \otimes_n)$ nie jest grupą, bo w Z_n co najmniej 0 nie jest elementem odwracalnym ze względu na działanie $\otimes_n$. Pełną charakteryzację elementów zbioru Z_n odwracalnych ze względu na działanie $\otimes_n$ przedstawiamy w następnym twierdzeniu.

Definicja 1.2.3. Mówimy, że liczby całkowite a i b są względnie pierwsze, gdy liczba 1 jest największym wspólnym dzielnikiem liczb a i b. Można udowodnić, że tak jest wtedy i tylko wtedy, gdy

$$ax + by = 1$$

dla pewnych liczb całkowitych x i y.

Twierdzenie 1.2.5. *Element x zbioru Z_n jest odwracalny (ze względu na działanie $\otimes_n$) wtedy i tylko wtedy, gdy liczby x i n są względnie pierwsze.*

Dowód. Załóżmy, że y jest elementem odwrotnym dla $x \in Z_n$ względem działania $\otimes_n$. Wtedy $x \otimes_n y = [x \cdot y]_n = 1$ i dlatego $x\,y = nk + 1$ dla pewnego $k \in Z$. Stąd

$$xy + n(-k) = 1$$

i to dowodzi, że liczby x i n są względnie pierwsze.

Załóżmy teraz, że liczby x i n są względnie pierwsze. Wtedy dla pewnych liczb całkowitych a i b jest

$$xa + nb = 1. \tag{1.12}$$

Zauważmy teraz, że $[a]_n = a - nk$ (dla pewnego $k \in Z$) i wobec (1.12) jest

$$x \otimes_n [a]_n = [x \cdot (a - nk)]_n = [1 - nb - nkx]_n = 1.$$

To dowodzi, że $[a]_n$ jest odwrotnością elementu x. □

Wniosek 1.2.1. *Każdy niezerowy element x zbioru Z_n jest odwracalny (ze względu na działanie $\otimes_n$) wtedy i tylko wtedy, gdy n jest liczbą pierwszą.*

Wniosek 1.2.2. *Struktura $(Z_n - \{0\}, \otimes_n)$ jest grupą wtedy i tylko wtedy, gdy n jest liczbą pierwszą.*

Definicja 1.2.4. Niech $(G, *)$ będzie grupą i niech H będzie podzbiorem zbioru G. Mówimy, że H jest *podgrupą* grupy G, jeśli $(H, *)$ jest grupą. Podgrupa

Jeśli e jest elementem neutralnym grupy G, to zbiór $H = \{e\}$ jest podgrupą grupy G. Podobnie $H = G$ jest podgrupą grupy G. Obie te podgrupy są tzw. trywialnymi podgrupami grupy G. Każdą inną podgrupę grupy G (jeśli taka istnieje) nazywamy jej podgrupą nietrywialną.

Z faktu, że $(G, *)$ jest grupą wynika, że działanie $*$ jest łączne na każdym podzbiorze H zbioru G zamkniętym ze względu na działanie $*$. Stąd zaś wynika, że podzbiór H grupy G jest jej podgrupą pod warunkiem, że:

(1) H jest zamknięty ze względu na działanie $*$;

(2) H zawiera element neutralny;

(3) H zawiera odwrotność każdego swojego elementu.

Następujące twierdzenie przedstawia prostszy warunek konieczny i dostateczny na to, aby podzbiór H grupy G był jej podgrupą. (W zadaniu 23 przedstawiamy jeszcze inny warunek konieczny i dostateczny na to, aby skończony podzbiór grupy był jej podgrupą.)

Twierdzenie 1.2.6. *Podzbiór H grupy G jest jej podgrupą wtedy i tylko wtedy, gdy spełnione są następujące dwa warunki:*

(1) *H jest niepusty;*

(2) *$a * b^{-1} \in H$ dla każdych elementów a i b zbioru H.*

Dowód. Załóżmy najpierw, że zbiór H jest podgrupą grupy G. Ponieważ element neutralny należy do H, więc $H \neq \emptyset$. Niech teraz a i b będą dowolnymi elementami zbioru H. Wtedy także $a, b^{-1} \in H$ (bo grupa H zawiera też odwrotności wszystkich swoich elementów) i $a * b^{-1} \in H$ (bo H jest zamknięty ze względu na działanie $*$).

Załóżmy teraz, że H jest niepustym podzbiorem zbioru G i $a*b^{-1} \in H$ dla każdych a, $b \in H$. Ponieważ $H \neq \emptyset$, więc istnieje co najmniej jeden element a w zbiorze H i dlatego $e = a*a^{-1} \in H$. Stąd zaś wynika, że dla każdego $a \in H$ jest $a^{-1} = e*a^{-1} \in H$ i to dowodzi, że H zawiera odwrotności wszystkich swoich elementów.

W końcu zauważmy, że jeżeli a i b są elementami zbioru H, to wobec powyższego także a i b^{-1} są elementami zbioru H i wtedy też $a * b = a * (b^{-1})^{-1} \in H$. Zatem zbiór H jest zamknięty ze względu na działanie $*$ i to kończy dowód faktu, że H jest podgrupą grupy G. □

$\otimes_7$	1	2	4
1	1	2	4
2	2	4	1
4	4	1	2

Przykład 9. Wobec wniosku 1.2.2 zbiór $Z_7 - \{0\} = \{1, 2, \ldots, 6\}$ jest grupą ze względu na mnożenie $\otimes_7$. Z przedstawionej obok tabelki wynika zaś, że niepusty zbiór

$$H = \{1,\ 2,\ 4\}$$

jest podgrupą grupy $(Z_7 - \{0\}, \otimes_7)$, bo zbiór H zawiera odwrotność każdego swojego elementu ($1^{-1} = 1$, $2^{-1} = 4$ i $4^{-1} = 2$) i jest zamknięty ze względu na mnożenie $\otimes_7$.

Przykład 10. Zbiór liczb wymiernych Q jest grupą ze względu na zwykłe dodawanie. Jego nietrywialny podzbiór $2Z$, czyli zbiór wszystkich parzystych liczb całkowitych, jest nietrywialną podgrupą grupy Q, bo różnica dowolnych dwóch liczb parzystych jest liczbą parzystą, tj. $a - b \in 2Z$ dla każdych a, $b \in 2Z$.

Przykład 11. Jeśli a jest ustalonym elementem grupy G, to wobec twierdzeń 1.2.3 i 1.2.6 zbiór wszystkich całkowitych potęg elementu a, czyli zbiór

$$H = \{a^n \colon\ n \in Z\},$$

jest podgrupą grupy G. Podgrupa ta zwykle jest oznaczana symbolem $\langle a \rangle$ i nazywana *podgrupą cykliczną* grupy G generowaną przez element a. Samą grupę G nazywamy *grupą cykliczną*, gdy $G = \langle a \rangle$ dla pewnego $a \in G$. W takim przypadku mówimy też, że element a jest generatorem grupy G.

Grupa cykliczna

Przykładowo, grupa liczb całkowitych Z jest grupą cykliczną ze względu na dodawanie i jej jedynymi generatorami są liczby 1 i -1, czyli $Z = \langle 1 \rangle$ oraz $Z = \langle -1 \rangle$.

W grupie $Z_7 - \{0\}$ z mnożeniem modulo 7 podgrupami cyklicznymi są: $\langle 1 \rangle = \{1\}$, $\langle 2 \rangle = \langle 4 \rangle = \{1,\ 2,\ 4\}$, $\langle 3 \rangle = \langle 5 \rangle = \{1, \ldots, 6\} = Z_7 - \{0\}$ i $\langle 6 \rangle = \{1,\ 6\}$.

1.3. Pierścień

Definicja 1.3.1. Niech $+$ i $\circ$ będą działaniami w niepustym zbiorze P; działania te będziemy nazywać odpowiednio dodawaniem i mnożeniem w zbiorze P. Mówimy, że system algebraiczny $(P, +, \circ)$ (albo, dla krótkości, że zbiór P) jest *pierścieniem*, gdy:

Pierścień

(P_1) P jest grupą przemienną ze względu na dodawanie $+$;

(P_2) mnożenie $\circ$ jest łączne w zbiorze P;

(P_3) mnożenie $\circ$ jest lewo- i prawostronnie rozdzielne względem dodawania, czyli dla każdych elementów x, y, $z \in P$ jest

$$x \circ (y + z) = x \circ y + x \circ z \quad \text{i} \quad (x + y) \circ z = x \circ z + y \circ z.$$

O pierścieniu P mówimy, że jest *pierścieniem przemiennym*, gdy

(P_4)
$$x \circ y = y \circ x$$

Pierścień przemienny

dla każdych elementów x, $y \in P$.

Z warunku (P_1) powyższej definicji oraz z definicji grupy (zob. def. 1.2.1) i z twierdzenia 1.1.1 wynika, że pierścień P ma dokładnie jeden element neutralny ze względu na działanie dodawania. Element ten zwykle oznaczamy symbolem 0 i nazywamy *zerem pierścienia* P. Z podobnych powodów dla każdego $x \in P$ istnieje dokładnie jeden element $\tilde{x} \in P$ taki, że $x+\tilde{x} = 0$. Element ten oznaczamy przez $-x$ i nazywamy *elementem przeciwnym* do x.

0 – zero pierścienia

$-x$ – element przeciwny do x w pierścieniu

Proste własności zera i elementów przeciwnych przedstawiamy w następnym twierdzeniu.

Twierdzenie 1.3.1. *Jeśli x i y są elementami pierścienia P, to:*

(1) $-(-x) = x$;

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

(3) $x \circ 0 = 0 \circ x = 0$;

(4) $(-x) \circ y = -(x \circ y) = x \circ (-y)$;

(5) $(-x) \circ (-y) = x \circ y$.

Dowód. Własność (1) jest natychmiastową konsekwencją definicji elementu przeciwnego (zob. def. 1.1.4). Równość (2) jest treścią twierdzenia 1.2.2 dla elementów grupy przemiennej $(P, +)$. Dla dowodu (3) zauważmy, że mamy $0 + x \circ 0 = x \circ 0 = x \circ (0 + 0) = x \circ 0 + x \circ 0$ i stąd wobec twierdzenia 1.2.1 (w grupie $(P, +)$) jest $x \circ 0 = 0$. Podobnie pokazuje się, że $0 \circ x = 0$. Z rozdzielności mnożenia względem dodawania i z (3) mamy $(-x) \circ y + x \circ y = ((-x) + x) \circ y = 0 \circ y = 0$ i, podobnie, $x \circ y + (-x) \circ y = 0$. Stąd zaś wynika, że $(-x) \circ y = -(x \circ y)$. Analogicznie dowodzi się, że $x \circ (-y) = -(x \circ y)$ i to kończy dowód własności (4). W końcu własność (5) wynika z (4) i (1), bo $(-x) \circ (-y) = -(x \circ (-y)) = -(-(x \circ y)) = x \circ y$. □

Przykład 12. Każdy ze zbiorów Z, $2Z$, Q i R jest pierścieniem przemiennym ze zwykłym dodawaniem i zwykłym mnożeniem liczb.

Definicja 1.3.2. *Jedynką pierścienia* P nazywamy element $e \in P$ taki, że dla każdego $x \in P$ jest

Jedynka pierścienia

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

Z twierdzenia 1.1.1 wynika, że każdy pierścień ma co najwyżej jedną jedynkę. Jedynkę pierścienia zwykle oznacza się symbolem 1. Każdy z pierścieni Z, Q i R z poprzedniego przykładu jest pierścieniem z jedynką. Jednakże pierścień $2Z$ jest pierścieniem bez jedynki.

Definicja 1.3.3. Element x pierścienia P nazywamy *dzielnikiem zera*, jeśli $x \neq 0$ i istnieje element $y \in P - \{0\}$ taki, że

Dzielnik zera

$$x \circ y = 0 \quad \text{lub} \quad y \circ x = 0.$$

Przykład 13. Żaden z pierścieni Z, $2Z$, Q i R z przykładu 12 nie ma dzielników zera.

Przykład 14. Wiemy już (zob. twierdzenie 1.2.4), że zbiór Z_n $(n > 1)$ jest grupą przemienną ze względu na dodawanie modulo n. Bezpośrednio po dowodzie twierdzenia 1.2.4 wspomnieliśmy także, że mnożenie modulo n jest łączne i przemienne w zbiorze Z_n. Mnożenie to jest także lewostronnie rozdzielne względem dodawania modulo n, bo dla każdych x, y, $z \in Z_n$ mamy

$$\begin{aligned}
x \otimes_n (y \oplus_n z) &= x \otimes_n [y + z]_n = \big[x \cdot [y + z]_n\big]_n && \text{(z definicji działań } \oplus_n \text{ i } \otimes_n) \\
&= [x \cdot (y + z)]_n && \text{(z własności (1.8))} \\
&= [(x \cdot y) + (x \cdot z)]_n && \text{(z rozdzielności zwykłego mnożenia liczb} \\
& && \text{całkowitych względem zwykłego dodawania liczb całkowitych)} \\
&= [x \cdot y]_n \oplus_n [x \cdot z]_n && \text{(z definicji działania } \oplus_n) \\
&= (x \otimes_n y) \oplus_n (x \otimes_n z). && \text{(z definicji działania } \otimes_n)
\end{aligned}$$

Stąd i z przemienności mnożenia $\otimes_n$ wynika także prawostronna rozdzielność mnożenia $\otimes_n$ względem dodawania $\oplus_n$. Ze wszystkich tych obserwacji wynika, że system $(Z_n, \oplus_n, \otimes_n)$ jest pierścieniem przemiennym. Jest oczywiste, że liczba 1 jest jedynką tego pierścienia.

Element odwracalny
Element odwrotny

Definicja 1.3.4. Niech P będzie pierścieniem z jedynką. Element $x \in P$ nazywamy *odwracalnym w pierścieniu* P, gdy istnieje element $x' \in P$ (zwany elementem odwrotnym elementu x) taki, że

$$x \circ x' = x' \circ x = 1.$$

Z twierdzenia 1.1.2 wynika, że dla każdego elementu x pierścienia istnieje co najwyżej jeden element odwrotny do elementu x. Element ten zwykle oznaczamy przez x^{-1}.

1.4. Ciało

Ciało

Definicja 1.4.1. Niech K będzie pierścieniem przemiennym z jedynką. Mówimy, że K jest *ciałem*, gdy każdy niezerowy element x pierścienia K jest odwracalny. Równoważnie, system algebraiczny $(K, \oplus, \otimes)$, w którym K jest zbiorem mającym co najmniej dwa elementy, a $\oplus$ i $\otimes$ są działaniami w zbiorze K (zwanymi odpowiednio dodawaniem i mnożeniem), nazywamy ciałem, gdy:

(C_1) $(K, \oplus)$ jest grupą przemienną:

(1) $\forall_{x,\, y \in K}\ x \oplus y = y \oplus x$; (przemienność dodawania)

(2) $\forall_{x,\, y,\, z \in K}\ x \oplus (y \oplus z) = (x \oplus y) \oplus z$; (łączność dodawania)

(3) $\exists_{0 \in K} \forall_{x \in K}\ x \oplus 0 = x$; (element neutralny dodawania)

(4) $\forall_{x \in K} \exists_{-x \in K}\ x \oplus (-x) = 0$; (element przeciwny względem dodawania)

(C_2) $(K - \{0\}, \otimes)$ jest grupą przemienną:

(5) $\forall_{x,\, y \in K - \{0\}}\ x \otimes y = y \otimes x$; (przemienność mnożenia)

(6) $\forall_{x,\, y,\, z \in K - \{0\}}\ x \otimes (y \otimes z) = (x \otimes y) \otimes z$; (łączność mnożenia)

(7) $\exists_{1 \in K - \{0\}} \forall_{x \in K - \{0\}}\ x \otimes 1 = x$; (element neutralny mnożenia)

(8) $\forall_{x \in K - \{0\}} \exists_{x^{-1} \in K - \{0\}}\ x \otimes x^{-1} = 1$; (element odwrotny względem mnożenia)

(C_3) mnożenie $\otimes$ jest rozdzielne względem dodawania $\oplus$:

(9) $\forall_{x,\, y,\, z \in K}\ x \otimes (y \oplus z) = (x \otimes y) \oplus (x \otimes z)$.

Przykład 15. Działania zwykłego dodawania i zwykłego mnożenia liczb rzeczywistych mają własności (1)—(9) powyższej definicji, więc zbiór liczb rzeczywistych R ze zwykłym dodawaniem i zwykłym mnożeniem liczb rzeczywistych jest ciałem. Podobnie, zbiór liczb wymiernych Q (ze zwykłym dodawaniem i zwykłym mnożeniem liczb wymiernych) jest ciałem. Zbiór liczb całkowitych Z (ze zwykłym dodawaniem i zwykłym mnożeniem liczb całkowitych) jest pierścieniem przemiennym z jedynką, ale nie jest on ciałem, bo niektóre jego niezerowe elementy nie są odwracalne (jedynymi odwracalnymi elementami pierścienia Z są 1 i -1).

Przykład 16. Z faktu, że $(Z_n, \oplus_n, \otimes_n)$ jest pierścieniem przemiennym z jedynką (zob. przykład 14) wynika, że system ten jest ciałem wtedy i tylko wtedy, gdy każdy element zbioru $Z_n - \{0\}$ jest odwracalny. Wobec wniosku 1.2.2 tak jest wtedy i tylko wtedy, gdy n jest liczbą pierwszą. Zatem zbiór Z_n (z dodawaniem i mnożeniem modulo n) jest ciałem wtedy i tylko wtedy, gdy n jest liczbą pierwszą.

Inne ważne przykłady grup, pierścieni i ciał przedstawiamy w kolejnych rozdziałach.

1.5. Ćwiczenia

1. W zbiorze liczb całkowitych Z dane jest działanie $*$, gdzie $x * y = x + |y|$ dla $x, y \in Z$. (a) Zbadać przemienność i łączność działania $*$. (b) Czy w zbiorze Z istnieje element neutralny działania $*$?

2. W zbiorze liczb rzeczywistych R określone jest zwykłe dodawanie $+$ i zwykłe mnożenie $\cdot$ liczb rzeczywistych oraz działanie $*$, gdzie $a*b = a+b-a\cdot b$. Sprawdzić, czy działanie $*$ jest: (a) łączne; (b) rozdzielne względem dodawania $+$; (c) rozdzielne względem mnożenia $\cdot$.

3. W zbiorze $R \times R$ dane jest działanie $*$ takie, że
$$(x_1, y_1) * (x_2, y_2) = (x_1x_2, y_1 + y_2).$$
(a) Wskazać element neutralny działania $*$. (b) Które elementy (x, y) zbioru $R\times R$ są odwracalne ze względu na działanie $*$? (c) Czy działanie $*$ jest łączne?

4. Niech X będzie zbiorem mającym co najmniej 3 elementy. Wskazać dwa wzajemnie jednoznaczne odwzorowania f i g zbioru X na siebie takie, że $f\circ g \neq g\circ f$.

5. Udowodnić, że dla każdych x, y, $z \in Z_n$ jest
$$x \otimes_n (y \otimes_n z) = (x \otimes_n y) \otimes_n z.$$

6. W zbiorze $R_+ = \{x \in R\colon x > 0\}$ dane jest działanie $\circ$, gdzie dla każdych x, $y \in R_+$ jest
$$x \circ y = x^{\ln y}.$$
(a) Czy działanie $\circ$ jest przemienne? (b) Czy para $(R_+, \circ)$ jest grupą?

7. W zbiorze $R \times (R - \{0\})$ określone jest działanie $\otimes$, gdzie
$$(x, y) \otimes (x', y') = (x + x'y, yy').$$
(a) Czy działanie $\otimes$ jest przemienne? (b) Wykazać, że $(R \times (R - \{0\}), \otimes)$ jest grupą. (c) Czy zbiór $S = \{(0, y) : y \in R - \{0\}\}$ z działaniem $\otimes$ jest podgrupą grupy $(R \times (R - \{0\}), \otimes)$?

8. Wykazać, że podzbiór $H = \{2, 4, 6, 8\}$ zbioru Z_{10} jest grupą przemienną ze względu na mnożenie modulo 10.

9. Czy zbiór $H = \{1, 4, 7, 13\}$ z mnożeniem modulo 15 tworzy grupę?

10. (a) W grupie Z_{96} z dodawaniem modulo 96 wskazać podgrupę mającą cztery elementy. (b) Czy grupa $(Z_{96}, \oplus_{96})$ ma podgrupę trzyelementową?

11. Niech a będzie ustalonym elementem grupy G. Pokazać, że zbiór
$$H = \{x \in G : ax = xa\}$$
jest podgrupą grupy G.

12. Niech G będzie grupą. Udowodnić, że zbiór
$$H = \{x \in G\colon xg = gx \quad \text{dla każdego} \quad g \in G\},$$
zwany centrum grupy G, jest podgrupą grupy G.

13. Niech H będzie podgrupą grupy G i niech a będzie ustalonym elementem grupy G. Udowodnić, że zbiór
$$F = \{aha^{-1}\colon\ h \in H\}$$
jest podgrupą grupy G.

14. Niech G będzie grupą przemienną. Udowodnić, że zbiór
$$H = \{x \in G\colon\ x^{-1} = x\}$$
jest podgrupą grupy G.

15. Udowodnić, że jeśli H_1 i H_2 są podgrupami grupy G, to także ich część wspólna $H_1 \cap H_2$ jest podgrupą grupy G.

16. Ile elementów ma podgrupa $H = \{4^n\colon\ n \in Z\}$ grupy $Z_{13} - \{0\}$ z mnożeniem modulo 13.

17. Pokazać, że zbiór $G = \{1,\ 3,\ 5,\ 9,\ 11,\ 13\}$ z mnożeniem modulo 14 jest grupą cykliczną.

18.* Udowodnić, że każda podgrupa grupy cyklicznej jest cykliczna.

19. Rzędem elementu x w skończonej grupie nazywamy najmniejszą liczbę naturalną k, dla której $x^k = e$. Udowodnić, że w skończonej grupie rząd elementu x jest identyczny z rzędem elementu odwrotnego x^{-1}.

20. Udowodnić, że dla każdych elementów a i b grupy G (z działaniem $\circ$) równania $a \circ x = b$ i $y \circ a = b$ mają jednoznaczne rozwiązania w grupie G, czyli istnieją jednoznacznie wyznaczone elementy x, $y \in G$ takie, że $a \circ x = b$ i $y \circ a = b$.

21.* Niech $\circ$ będzie działaniem dwuargumentowym w niepustym zbiorze G. Udowodnić, że $(G, \circ)$ jest grupą wtedy i tylko wtedy, gdy jednocześnie spełnione są warunki: (1) $a \circ (b \circ c) = (a \circ b) \circ c$ dla każdych a, b, $c \in G$; (2) równania $a \circ x = b$ i $y \circ a = b$ mają rozwiązania w zbiorze G dla każdych a, $b \in G$.

22.* Pokazać, że jeśli $\circ$ jest działaniem łącznym w skończonym zbiorze G, to para $(G, \circ)$ jest grupą wtedy i tylko wtedy, gdy dla każdych trzech elementów a, b, c ze zbioru G spełnione są warunki: (a) jeśli $a \circ b = a \circ c$, to $b = c$; (b) jeśli $a \circ b = c \circ b$, to $a = c$.

23.* Udowodnić, że skończony i niepusty podzbiór H grupy G jest jej podgrupą wtedy i tylko wtedy, gdy $ab \in H$ dla każdych dwóch elementów a i b ze zbioru H.

24. Udowodnić, że jeśli x jest dzielnikiem zera w pierścieniu przemiennym P i $y \in P$, to $xy = 0$ lub xy jest dzielnikiem zera.

25. Udowodnić, że jeśli niezerowy element x pierścienia P nie jest dzielnikiem zera i dla elementów y, $z \in P$ jest $xy = xz$, to $y = z$.

26. Działania $\oplus$ i $\otimes$ w zbiorze liczb rzeczywistych R są określone za pomocą zwykłego dodawania i zwykłego mnożenia liczb rzeczywistych i dla każdych x, $y \in R$ jest $x \oplus y = x + y + 1$ oraz $x \otimes y = xy + x + y$. Wykazać, że system $(R, \oplus, \otimes)$ jest ciałem.

27. Udowodnić, że ciało nie ma dzielników zera.

28. Niech P będzie skończonym i przemiennym pierścieniem z jedynką. Udowodnić, że albo P ma dzielnik zera, albo P jest ciałem.

29. W zbiorze L wszystkich nieskończonych ciągów rzeczywistych określamy dodawanie $\oplus$ i mnożenie $\otimes$ w następujący sposób:
$$\begin{aligned}(x_1, x_2, \ldots) \oplus (x_1, y_2, \ldots) &= (x_1 + y_1, x_2 + y_2, \ldots),\\ (x_1, x_2, \ldots) \otimes (x_1, y_2, \ldots) &= (x_1 y_1, x_2 y_2, \ldots).\end{aligned}$$
Wykazać, że $(L, \oplus, \otimes)$ jest pierścieniem i nie jest ciałem.

Rozdział 2

LICZBY ZESPOLONE

2.1. Liczby zespolone i działania na liczbach zespolonych

Niech $\mathcal{C}$ będzie zbiorem wszystkich uporządkowanych par (a, b) liczb rzeczywistych a i b,

$$\mathcal{C} = \{(a, b) : a, b \in R\}.$$

Za pomocą równości, zwykłego dodawania (i odejmowania) oraz zwykłego mnożenia liczb rzeczywistych definiujemy równość, dodawanie $\oplus$ oraz mnożenie $\otimes$ w zbiorze $\mathcal{C}$. Jeśli pary (a, b) i (c, d) są elementami zbioru $\mathcal{C}$, to przyjmujemy, że:

$$(a, b) = (c, d) \;\Leftrightarrow\; a = c \;\text{ i }\; b = d, \tag{2.1}$$

Równość liczb zespolonych

$$(a, b) \oplus (c, d) \;=\; (a + c, b + d), \tag{2.2}$$

Suma liczb zespolonych

$$(a, b) \otimes (c, d) \;=\; (ac - bd, ad + bc). \tag{2.3}$$

Iloczyn liczb zespolonych

Definicja 2.1.1. Elementy zbioru $\mathcal{C}$ (z równością (2.1) oraz działaniami dodawania i mnożenia określonymi wzorami (2.2) i (2.3)) nazywamy *liczbami zespolonymi.*

Liczby zespolone

Jeśli $z = (a, b)$ jest liczbą zespoloną, to liczby rzeczywiste a i b nazywamy odpowiednio *częścią rzeczywistą* i *częścią urojoną* liczby z i piszemy

$$\mathrm{Re}\,(z) = a \quad \text{ i } \quad \mathrm{Im}\,(z) = b.$$

Pokażemy teraz, jak z własności zwykłego dodawania i zwykłego mnożenia liczb rzeczywistych (zob. przykład 15 i definicję 1.4.1) wynika, że zbiór liczb zespolonych z wyżej określonym dodawaniem i mnożeniem liczb zespolonych jest ciałem.

Twierdzenie 2.1.1. *Zbiór* $\mathcal{C}$ *z działaniami dodawania i mnożenia określonymi wzorami* (2.2) *i* (2.3) *jest ciałem, więc działania te mają następujące własności:*

$\mathcal{C}$ – ciało liczb zespolonych

(a) $\forall_{z,w\in\mathcal{C}}\;\; z \oplus w = w \oplus z,$ (przemienność dodawania)

(b) $\forall_{z,w,t\in\mathcal{C}}\;\; z \oplus (w \oplus t) = (z \oplus w) \oplus t,$ (łączność dodawania)

(c) $\exists_{z_0\in\mathcal{C}}\; \forall_{z\in\mathcal{C}}\;\; z \oplus z_0 = z,$ ($z_0 = (0, 0)$ – zero zespolone)

(d) $\forall_{z\in\mathcal{C}}\; \exists_{-z\in\mathcal{C}}\;\; z \oplus (-z) = z_0,$ ($-z = (-a, -b)$ – liczba przeciwna do $z = (a, b)$)

(e) $\forall_{z,w\in\mathcal{C}}\;\; z \otimes w = w \otimes z,$ (przemienność mnożenia)

(f) $\forall_{z,w,t\in\mathcal{C}}\;\; z \otimes (w \otimes t) = (z \otimes w) \otimes t,$ (łączność mnożenia)

(g) $\exists_{z_1\in\mathcal{C}}\; \forall_{z\in\mathcal{C}}\;\; z \otimes z_1 = z,$ ($z_1 = (1, 0)$ – jedynka zespolona)

(h) $\forall_{z\in\mathcal{C}-\{z_0\}}\; \exists_{z^{-1}\in\mathcal{C}}\;\; z \otimes z^{-1} = z_1,$ ($z^{-1} = \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right)$, gdy $z = (a, b) \neq z_0$)

(i) $\forall_{z,w,t\in\mathcal{C}}\;\; z \otimes (w \oplus t) = (z \otimes w) \oplus (z \otimes t).$ (rozdzielność działnia $\otimes$ względem $\oplus$)

Dowód. (a) i (b). Z przemienności i łączności zwykłego dodawania liczb rzeczywistych wynika przemienność i łączność dodawania $\oplus$ określonego wzorem (2.2). Istotnie, jeśli (a,b), (c,d), $(e,f) \in \mathcal{C}$, to mamy

$$\begin{aligned} (a,b) \oplus (c,d) &= (a+c, b+d) && \text{(definicja działania } \oplus) \\ &= (c+a, d+b) && \text{(przemienność działania } +) \\ &= (c,d) \oplus (a,b) && \text{(definicja działania } \oplus) \end{aligned}$$

oraz

$$\begin{aligned} (a,b) \oplus \big((c,d) \oplus (e,f)\big) &= (a,b) \oplus (c+e, d+f) && \text{(definicja działania } \oplus) \\ &= \big(a+(c+e), b+(d+f)\big) && \text{(definicja działania } \oplus) \\ &= \big((a+c)+e, (b+d)+f\big) && \text{(łączność działania } +) \\ &= (a+c, b+d) \oplus (e,f) && \text{(definicja działania } \oplus) \\ &= \big((a,b) \oplus (c,d)\big) \oplus (e,f). && \text{(definicja działania } \oplus) \end{aligned}$$

(c) Liczba $(0,0)$ jest elementem neutralnym działania $\oplus$, bo dla każdej liczby zespolonej (a,b) mamy

$$(a,b) \oplus (0,0) = (a+0, b+0) = (a,b).$$

(d) Liczba $(-a,-b)$ jest liczbą przeciwną do liczby (a,b), bo

$$(a,b) \oplus (-a,-b) = \big(a+(-a), b+(-b)\big) = (0,0).$$

(e) i (f). Mnożenie $\otimes$ jest przemienne i łączne w zbiorze $\mathcal{C}$, bo dla każdych liczb zespolonych (a,b), (c,d), (e,f) jest

$$\begin{aligned} (a,b) \otimes (c,d) &= (ac-bd, ad+bc) && \text{(definicja działania } \otimes) \\ &= (ca-db, cb+da) && \text{(przemienność mnożenia} \\ & && \text{i dodawania liczb rzeczywistych)} \\ &= (c,d) \otimes (a,b) && \text{(definicja działania } \otimes) \end{aligned}$$

oraz

$$\begin{aligned} (a,b) \otimes \big((c,d) \otimes (e,f)\big) &= (a,b) \otimes (ce-df, cf+de) \\ &= \big(a(ce-df) - b(cf+de), a(cf+de) + b(ce-df)\big) \\ &= \big((ac-bd)e - (ad+bc)f, (ac-bd)f + (ad+bc)e\big) \\ &= (ac-bd, ad+bc) \otimes (e,f) \\ &= \big((a,b) \otimes (c,d)\big) \otimes (e,f). \end{aligned}$$

(g) Liczba $(1,0)$ jest elementem neutralnym działania $\otimes$, bo dla każdej liczby zespolonej (a,b) jest

$$(a,b) \oplus (1,0) = (a \cdot 1 - b \cdot 0, a \cdot 0 + b \cdot 1) = (a,b).$$

(h) Niech teraz (a,b) będzie dowolną liczbę ze zbioru $\mathcal{C} - \{(0,0)\}$. Wtedy $a^2+b^2 \neq 0$ i liczba $\big(a/(a^2+b^2), -b/(a^2+b^2)\big)$ istnieje i jest to liczba odwrotna do liczby (a,b),

$$(a,b) \otimes \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right) = \left(\frac{a^2}{a^2+b^2} - \frac{-b^2}{a^2+b^2}, \frac{-ab}{a^2+b^2} + \frac{ab}{a^2+b^2}\right) = (1,0).$$

(i) W końcu dla każdych liczb zespolonych (a,b), (c,d) i (e,f) mamy

$$\begin{aligned} (a,b) \otimes \big((c,d) \oplus (e,f)\big) &= (a,b) \otimes (c+e, d+f) \\ &= \big(a(c+e) - b(d+f), a(d+f) + b(c+e)\big) \\ &= \big((ac-bd) + (ae-bf), (ad+bc) + (af+be)\big) \\ &= (ac-bd, ad+bc) \oplus (ae-bf, af+be) \\ &= \big((a,b) \otimes (c,d)\big) \oplus \big((a,b) \otimes (e,f)\big). \end{aligned}$$

To dowodzi, że mnożenie $\otimes$ jest rozdzielne względem dodawania $\oplus$ i to jednocześnie kończy dowód twierdzenia. $\square$

Począwszy od tego miejsca działania na liczbach zespolonych oznaczać będziemy za pomocą symboli używanych dla liczb rzeczywistych, tzn. będziemy pisać $(a, b) + (c, d)$ zamiast $(a, b) \oplus (c, d)$ oraz $(a, b) \cdot (c, d)$ (lub nawet $(a, b)(c, d)$) zamiast $(a, b) \otimes (c, d)$. Nie spowoduje to większej niejednoznaczności. Dodatkowo, sądzimy, że dzięki takim zabiegom, Czytelnik szybciej dojdzie do przekonania, że działania na liczbach zespolonych określone wzorami (2.2) i (2.3) są "naturalnymi" uogólnieniami dodawania i mnożenia liczb rzeczywistych.

Definicja 2.1.2. *Różnicą* i *ilorazem* liczby zespolonej $z = (a, b)$ i liczby zespolonej $w = (c, d)$ nazywamy odpowiednio liczby $z - w$ i z/w, gdzie

$$z - w = (a - c, b - d), \tag{2.4}$$

Różnica liczb zespolonych

$$\frac{z}{w} = \left(\frac{ac + bd}{c^2 + d^2}, \frac{bc - ad}{c^2 + d^2}\right), \quad \text{gdy } w \neq (0, 0). \tag{2.5}$$

Iloraz liczb zespolonych
Tego wzoru nie warto pamiętać!

Definicja 2.1.3. Jeśli n jest liczbą naturalną, to n-tą potęgę liczby zespolonej z definiuje się indukcyjnie (zob. def. 1.2.2), przyjmując, że

$$z^1 = z \quad \text{i} \quad z^n = z \cdot z^{n-1} \tag{2.6}$$

Potęga liczby zespolonej

dla $n = 2, 3, \ldots$. Dla $z \neq 0$ przyjmujemy także, że

$$z^0 = 1 \quad \text{i} \quad z^{-n} = \frac{1}{z^n} \tag{2.7}$$

dla naturalnego n.

Postać kanoniczna liczby zespolonej

Niech $\mathcal{C}_R$ będzie zbiorem liczb zespolonych, których część urojona jest zerem. Każda liczba ze zbioru $\mathcal{C}_R$ jest postaci $(a, 0)$ i jest ona jednoznacznie wyznaczona przez swoją część rzeczywistą a, więc liczbę zespoloną $(a, 0)$ utożsamiamy z liczbą rzeczywistą a i wprost piszemy

$$(a, 0) = a. \tag{2.8}$$

Ze względu na powyższe utożsamianie (które formalnie jest izomorfizmem ciała liczb rzeczywistych R z podciałem $\mathcal{C}_R$ ciała liczb zespolonych $\mathcal{C}$, zob. zad. 39 i 40) możemy powiedzieć, że zbiór liczb zespolonych $\mathcal{C}$, w którym sumę i iloczyn określono wzorami (2.2) i (2.3), jest rozszerzeniem zbioru liczb rzeczywistych R z tradycyjnie rozumianą sumą i iloczynem liczb rzeczywistych.

Do zbioru $\mathcal{C}$ należy liczba $(0, 1)$, którą nazywamy *jednostką urojoną* i oznaczamy symbolem j, czyli

$$j = (0, 1). \tag{2.9}$$

Zauważmy, że wobec (2.3) i (2.8) mamy

$$j \cdot j = (0, 1)(0, 1) = (-1, 0) = -1,$$

więc

$$j^2 = -1. \tag{2.10}$$

Równie łatwo zauważamy, że $j^3 = -j$ i $j^4 = 1$. Stąd zaś można wywnioskować, że każda naturalna potęga liczby j jest jedną z liczb j, -1, $-j$ i 1. W szczególności mamy

$$j^5 = j, \quad j^6 = -1, \quad j^7 = -j \quad \text{i} \quad j^8 = 1.$$

Twierdzenie 2.1.2. *Każdą liczbę zespoloną $z = (a, b)$ można przedstawić w postaci*

$$z = a + bj. \tag{2.11}$$

Dowód. Łatwo zauważyć, że dla liczby zespolonej $z = (a, b)$ wobec (2.2), (2.3), (2.8) i (2.9) kolejno mamy

$$z = (a, b) = (a, 0) + (0, b) = (a, 0) + (b, 0)(0, 1) = a + bj. \ \square$$

$z = a + jb$ – postać kanoniczna liczby $z = (a, b)$

Postać $z = a + bj$ liczby zespolonej $z = (a, b)$ nazywamy jej *postacią kanoniczną* (*algebraiczną* lub *kartezjańską*). Wobec (2.2)–(2.5) suma, różnica, iloczyn i iloraz liczb zespolonych w postaci kanonicznej określone są odpowiednio wzorami:

Działania na liczbach w postaci kanonicznej

$$z + w = (a + bj) + (c + dj) = (a + c) + (b + d)j; \tag{2.12}$$

$$z - w = (a + bj) - (c + dj) = (a - c) + (b - d)j; \tag{2.13}$$

$$zw = (a + bj)(c + dj) = (ac - bd) + (ad + bc)j; \tag{2.14}$$

$$\frac{z}{w} = \frac{a + bj}{c + dj} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2}j. \tag{2.15}$$

Przykład 17. Obliczyć $z + w$, $z - w$ i zw, gdy $z = 1 + 5j$ i $w = 3 + 2j$.

Łatwo zauważyć, że

$$z + w = (1 + 5j) + (3 + 2j) = (1 + 3) + (5j + 2j) = 4 + 7j,$$

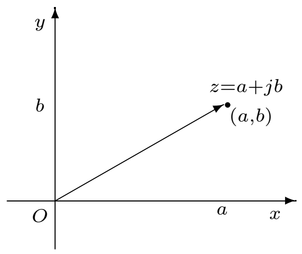
$$z - w = (1 + 5j) - (3 + 2j) = (1 - 3) + (5j - 2j) = -2 + 3j.$$

Przy wyznaczaniu iloczynu liczb zespolonych w postaci kanonicznej nie trzeba pamiętać wzoru (2.14). Wystarczy skorzystać z rozdzielności mnożenia względem dodawania, zastąpić j^2 przez -1 i zgrupować "podobne" czynniki. Dlatego dla liczb $z = 1+5j$ i $w = 3 + 2j$ mamy

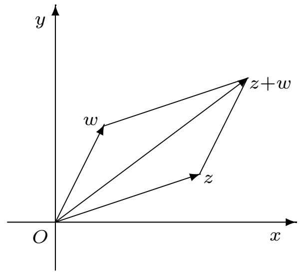
$$\begin{aligned} zw &= (1 + 5j)(3 + 2j) = 1(3 + 2j) + 5j(3 + 2j) \\ &= 3 + 2j + 15j + 10j^2 = 3 + 17j - 10 = -7 + 17j. \end{aligned}$$

Interpretacja geometryczna liczby zespolonej

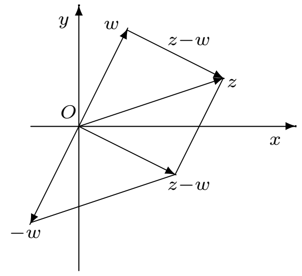
Każdą liczbę zespoloną $z = a + jb$ można utożsamiać z punktem (a, b), którego współrzędne są odpowiednio równe części rzeczywistej i części urojonej liczby $z = a+jb$ (zob. rys. 2.1). Płaszczyznę, w której każdy punkt utożsamiamy z liczbą zespoloną, nazywamy *płaszczyzną zespoloną* lub *płaszczyzną Gaussa* (albo *płaszczyzną Arganda*). W tym przypadku oś odciętych i oś rzędnych nazywamy odpowiednio osią rzeczywistych i osią urojonych. Liczbę zespoloną $z = a + jb$ można również utożsamiać z wektorem wodzącym punktu (a, b), tj. z wektorem $\overrightarrow{Oz}$, gdzie O jest początkiem układu. W tej interpretacji suma liczb zespolonych z i w jest czwartym wierzchołkiem równoległoboku zbudowanego na wektorach $\overrightarrow{Oz}$ i $\overrightarrow{Ow}$, zob. rys. 2.2. Natomiast różnica $z - w$ jest czwartym wierzchołkiem równoległoboku zbudowanego na wektorach $\overrightarrow{Oz}$ i $\overrightarrow{O(-w)}$; różnicę $z - w$ można także utożsamiać z wektorem $\overrightarrow{wz}$ (zob. rys. 2.3).



Rys. 2.1



Rys. 2.2



Rys. 2.3

2.2. Sprzężenie i moduł liczby zespolonej

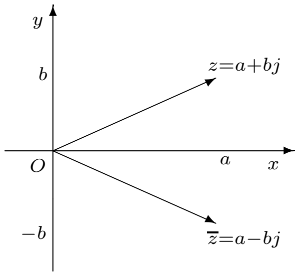
Okazuje się, że także przy wyznaczaniu ilorazu dwóch liczb zespolonych w postaci kanonicznej nie trzeba pamiętać wzoru (2.15). W zamian warto skorzystać z własności sprzężeń liczb zespolonych.

Sprzężenie liczby zespolonej

Definicja 2.2.1. *Sprzężeniem* liczby zespolonej $z = a+bj$ (gdzie a i b są liczbami rzeczywistymi) nazywamy liczbę

$$\overline{z} = a - bj.$$

Przykładowo, jeśli $z = 2 + 5j$, to jej sprzężeniem jest $\overline{z} = 2 - 5j$. Geometrycznie $\overline{z}$ jest symetrycznym odbiciem z względem osi rzeczywistej (zob. rys. 2.4). Inne własności sprzężenia podajemy w następnym twierdzeniu (zob. także tw. 2.2.2 (a)).



Rys. 2.4

Twierdzenie 2.2.1. *Jeśli z i w są liczbami zespolonymi, to:*

(a) $\overline{z+w} = \overline{z} + \overline{w}$, $\overline{z-w} = \overline{z} - \overline{w}$;

(b) $\overline{zw} = \overline{z}\,\overline{w}$, $\overline{z/w} = \overline{z}/\overline{w}$ *(gdy* $w \neq 0$*);*

(c) *dla każdej liczby całkowitej* n *jest* $\overline{(z^n)} = \left(\overline{z}\right)^n$ $(z \neq 0$, *gdy* $n \leqslant 0)$;

(d) z *jest liczbą rzeczywistą wtedy i tylko wtedy, gdy* $\overline{z} = z$.

Dodatkowo, jeśli $z = a + bj$ *(gdzie* $a, b \in R$*), to*

(e) $z + \overline{z} = 2a$, $z - \overline{z} = 2bj$ *i* $z\overline{z} = a^2 + b^2$.

Dowód. Udowodnimy pierwszą część własności (b). Dowody pozostałych własności są równie łatwe. Załóżmy, że $z = a + bj$ i $w = c + dj$, gdzie a, b, c i d są liczbami rzeczywistymi. Wtedy

$$\begin{aligned} \overline{z}\,\overline{w} &= \overline{(a+bj)}\;\overline{(c+dj)} = (a-bj)(c-dj) = (ac-bd) - (ad+bc)j \\ &= \overline{(ac-bd) + (ad+bc)j} = \overline{(a+bj)(c+dj)} = \overline{zw}. \ \square \end{aligned}$$

Możemy teraz wyznaczyć postać kanoniczną ilorazu dwóch liczb zespolonych bez odwoływania się do wzoru (2.15). W tym celu wystarczy skorzystać z tożsamości

Sposób wyznaczania ilorazu liczb zespolonych

$$\frac{z}{w} = \frac{z}{w}\frac{\overline{w}}{\overline{w}} = \frac{z\overline{w}}{w\overline{w}} \tag{2.16}$$

oraz z umiejętności mnożenia liczb zespolonych i dzielenia liczby zespolonej przez liczbę rzeczywistą.

Przykład 18. Znaleźć postać kanoniczną liczby $(1 + 5j)/(3 + 2j)$.

Wobec poprzednich uwag mamy

$$\frac{1+5j}{3+2j} = \frac{(1+5j)(3-2j)}{(3+2j)(3-2j)} = \frac{3-2j+15j-10j^2}{9-4j^2} = \frac{13+13j}{13} = 1 + j.$$

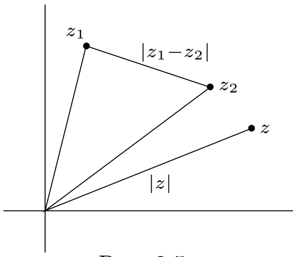
Geometryczną relację między liczbami z i w oraz ich iloczynem zw (oraz ilorazem z/w) można opisać w terminach modułu i argumentu liczby zespolonej.

Definicja 2.2.2. *Modułem* (lub *wielkością*) liczby zespolonej $z = a + bj$ (gdzie a, $b \in R$) nazywamy nieujemną liczbę rzeczywistą

Moduł liczby zespolonej

$$|z| = \sqrt{a^2 + b^2}. \tag{2.17}$$

Moduł liczby $z = a + bj$ jest odległością punktu (a, b) od początku układu współrzędnych. Ogólniej, dla liczb zespolonych $z_1 = a_1 + jb_1$ i $z_2 = a_2 + jb_2$ (gdzie a_1, a_2, b_1, $b_2 \in R$) moduł $|z_1 - z_2|$ jest odległością pomiędzy z_1 i z_2 (zob. rys. 2.5), bo $|z_1 - z_2| = |(a_1 + jb_1) - (a_2 + jb_2)| = |(a_1 - a_2) + j(b_1 - b_2)| = \sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2}$. Jeśli z jest liczbą rzeczywistą, to $z = a + 0j$ (dla pewnego $a \in R$) i $|z| = \sqrt{a^2}$, co jest wartością bezwzględną z liczby $a = z$. Pojęcie modułu liczby zespolonej jest więc uogólnieniem pojęcia wartości bezwzględnej. Podstawowe własności modułu przedstawiono w następującym twierdzeniu.



Rys. 2.5

Twierdzenie 2.2.2. *Jeśli $z, w \in \mathcal{C}$, to:*

(*a*) $|z| = \sqrt{z\overline{z}}, \quad |z| = |\overline{z}| = |-z|$;

(*b*) $|zw| = |z||w|, \quad \left|\frac{z}{w}\right| = \frac{|z|}{|w|}$ $(w \neq 0)$;

(*c*) $|z| \geqslant |\mathrm{Re}\,(z)| \geqslant \mathrm{Re}\,(z), \quad |z| \geqslant |\mathrm{Im}\,(z)| \geqslant \mathrm{Im}\,(z)$;

(*d*) $\big||z| - |w|\big| \leqslant |z + w| \leqslant |z| + |w|$.

Dowód. Udowodnimy tylko część (*d*). Z części (*a*)$-$(*c*) oraz z twierdzenia 2.2.1 kolejno mamy

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

i z tego wynika, że $|z + w| \leqslant |z| + |w|$. Dodatkowo wobec (a) mamy

$$|z| = |(z + w) + (-w)| \leqslant |z + w| + |-w| = |z + w| + |w|,$$

czyli $|z| - |w| \leqslant |z + w|$. Z tych samych powodów jest $|w| - |z| \leqslant |z + w|$. Z tych nierówności wynika, że $\big||z| - |w|\big| \leqslant |z + w|$ i to kończy dowód (*d*). □

Wniosek 2.2.1. *Dla każdej liczby naturalnej n i każdych liczb zespolonych $z, z_1, \ldots, z_n$ jest:*

(1) $|z^n| = |z|^n$ *i* $|z^{-n}| = |z|^{-n}$ $(z \neq 0)$;

(2) $|z_1 + z_2 + \ldots + z_n| \leqslant |z_1| + |z_2| + \ldots + |z_n|$.

Przykład 19. Znaleźć moduł liczby $z = \dfrac{(2+j)^2}{(1+6j)(1-7j)}$.

Wobec definicji 2.2.2 i twierdzenia 2.2.2 mamy

$$|z| = \frac{|2+j|^2}{|1+6j||1-7j|} = \frac{(\sqrt{2^2+1^2})^2}{\sqrt{1^2+6^2}\sqrt{1^2+(-7)^2}} = \frac{5}{\sqrt{37}\sqrt{50}} = \frac{1}{\sqrt{74}}.$$

Przykład 20. Wyznaczyć zbiór punktów z spełniających równanie

$$|z + 2j| = 4|z - 2j|.$$

Korzystając z postaci kanonicznej $x + jy$ liczby zespolonej z, równanie $|z + 2j| = 4|z - 2j|$ można zapisać w postaci $|x + jy + 2j| = 4|x + jy - 2j|$ lub $\sqrt{x^2 + (y+2)^2} = 4\sqrt{x^2 + (y-2)^2}$. Po podniesieniu obu stron do kwadratu i redukcji otrzymujemy $15x^2 + 15y^2 - 68y + 60 = 0$, czyli

$$x^2 + \left(y - \frac{34}{15}\right)^2 = \left(\frac{16}{15}\right)^2,$$

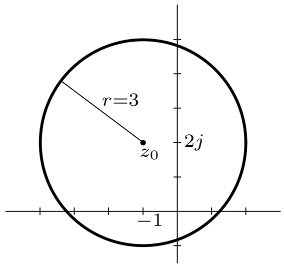
co jest równaniem okręgu o środku w punkcie $\left(0, \frac{34}{15}\right)$ i promieniu długości $\frac{16}{15}$.

Przykład 21. Na płaszczyźnie zespolonej zaznaczyć zbiór tych liczb zespolonych z, które spełniają podane warunki:

(a) $|z+1-2j|=3$;

(b) $2 \leqslant |z-1-3j| \leqslant 4$;

(c) $|z-1| \leqslant \mathrm{Im}\,(z)+2$.



Rys. 2.6

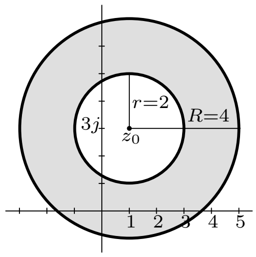
(a) Ponieważ

$$|z+1-2j|=3 \Leftrightarrow |z-(-1+2j)|=3,$$

więc rozważany zbiór jest zbiorem wszystkich punktów z położonych w odległości $r=3$ od punktu $z_0=-1+2j$. Zatem jest to okrąg o środku w punkcie $z_0=-1+2j$ i promieniu $r=3$, zob. rys. 2.6.

(b) Rozważany zbiór składa się z tych i tylko tych z, dla których

$$2 \leqslant |z-(1+3j)| \leqslant 4,$$

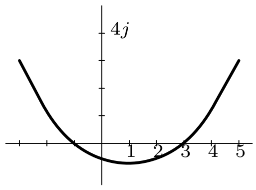


Rys. 2.7

czyli jest to zbiór tych z, których odległość od punktu $z_0=1+3j$ jest liczbą z przedziału $\langle 2;4\rangle$. Zatem jest to pierścień kołowy o środku w punkcie $z_0=1+3j$ i promieniu wewnętrznym $r=2$ oraz promieniu zewnętrznym $R=4$, zob. rys. 2.7.

(c) Dla liczby $z=x+jy$ (gdzie $x,\, y \in R$) jest

$$\begin{aligned} |z-1| \leqslant \mathrm{Im}\,(z)+2 \quad &\Leftrightarrow \quad |(x-1)+jy| \leqslant \mathrm{Im}\,(x+jy)+2 \\ &\Leftrightarrow \quad \sqrt{(x-1)^2+y^2} \leqslant y+2 \\ &\Leftrightarrow \quad y \geqslant \tfrac{(x-1)^2-4}{4} \end{aligned}$$



Rys. 2.8

i stąd wynika, że rozważany zbiór jest zbiorem wszystkich punktów $z=x+jy$, które leżą na lub nad parabolą $y=\frac{(x-1)^2-4}{4}$, zob. rys. 2.8.

2.3. Postać trygonometryczna liczby zespolonej

Każdą różną od zera liczbę zespoloną $z=a+bj$ można przedstawić w postaci

$$z=|z|\left(\frac{a}{|z|}+j\frac{b}{|z|}\right), \tag{2.18}$$

a ponieważ

$$\left(\frac{a}{|z|}\right)^2+\left(\frac{b}{|z|}\right)^2=\frac{a^2}{a^2+b^2}+\frac{b^2}{a^2+b^2}=1,$$

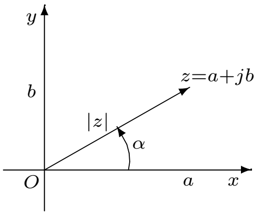
więc jedna z liczb $a/|z|$ i $b/|z|$ jest sinusem, a druga cosinusem tej samej liczby rzeczywistej α i dlatego możemy przedstawić następującą definicję.

Definicja 2.3.1. *Argumentem* liczby zespolonej $z=a+bj\neq 0$ nazywamy każdą liczbę rzeczywistą α, oznaczamy ją także symbolem $\arg(z)$, dla której

Argument liczby zespolonej

$$\frac{a}{|z|}=\cos\alpha \quad \text{i} \quad \frac{b}{|z|}=\sin\alpha. \tag{2.19}$$

Geometrycznie argument liczby z jest miarą kąta skierowanego jaki wektor $\overrightarrow{Oz}$ tworzy z dodatnim kierunkiem osi Ox (rys. 2.9). Z okresowości sinusa i cosinusa wynika, że każda liczba zespolona $z\neq 0$ ma nieskończenie wiele argumentów α i każde dwa z nich różnią się o całkowitą krotność liczby 2π. Spośród argumentów α liczby z dokładnie jeden spełnia nierówności $-\pi<\alpha\leqslant\pi$; nazywamy



Rys. 2.9

go *argumentem głównym* liczby z i oznaczamy przez $\mathrm{Arg}\,(z)$.[1] Argument każdej liczby $z = a + bj$ ($\neq 0$) można wyznaczyć z równości (2.19). W szczególności argumentem (głównym) liczby urojonej $z = bj$ ($b \in R - \{0\}$) jest $\frac{\pi}{2}$ lub $-\frac{\pi}{2}$ zależnie od tego, czy b jest liczbą dodatnią, czy ujemną. Argument liczby różnej od liczby urojonej można także wyznaczyć ze wzoru

$$\mathrm{Arg}\,(a+bj) = \begin{cases} \arctan(b/a) & \text{jeśli } a > 0, \\ \arctan(b/a) + \pi & \text{jeśli } a < 0 \text{ i } b \geqslant 0, \\ \arctan(b/a) - \pi & \text{jeśli } a < 0 \text{ i } b < 0. \end{cases} \tag{2.20}$$

Argument liczby $z = 0$ jest nieokreślony (ale możemy także przyjąć, że argumentem zera jest dowolna liczba rzeczywista).

Przykład 22. Dla liczb $-1 + j$ oraz $-\sqrt{3} - j$ wobec (2.20) mamy

$$\mathrm{Arg}\,(-1+j) = \arctan(-1/1) + \pi = -\frac{\pi}{4} + \pi = \frac{3}{4}\pi$$

oraz

$$\mathrm{Arg}\,(-\sqrt{3}-j) = \arctan(1/\sqrt{3}) - \pi = \frac{\pi}{6} - \pi = -\frac{5}{6}\pi.$$

Postać trygonometryczna liczby zespolonej

Z (2.18) i (2.19) wynika, że liczbę zespoloną z można przedstawić w postaci

$$z = |z|(\cos\alpha + j\sin\alpha), \tag{2.21}$$

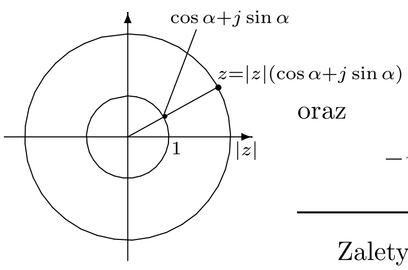
zwanej *postacią trygonometryczną* (lub *biegunową*) liczby z, w której α jest argumentem liczby z, czyli $\alpha = \arg\,(z)$.

Przykład 23. Dla liczb z poprzedniego przykładu mamy

$$\begin{aligned} -1+j &= |-1+j|\big(\cos\arg\,(-1+j) + j\sin\arg\,(-1+j)\big) \\ &= \sqrt{2}(\cos\tfrac{3}{4}\pi + j\sin\tfrac{3}{4}\pi) \end{aligned}$$

oraz

$$-\sqrt{3} - j = 2\left(\cos\left(-\tfrac{5}{6}\pi\right) + j\sin\left(-\tfrac{5}{6}\pi\right)\right) = 2\left(\cos\tfrac{5}{6}\pi - j\sin\tfrac{5}{6}\pi\right).$$



Rys. 2.10

Zalety postaci trygonometrycznej liczby zespolonej uwidaczniają się przy mnożeniu, dzieleniu, potęgowaniu i pierwiastkowaniu liczb zespolonych. Załóżmy, że znamy postać trygonometryczną liczb z i w, powiedzmy $z = |z|(\cos\alpha + j\sin\alpha)$ i $w = |w|(\cos\beta + j\sin\beta)$. Łatwo zauważyć, że liczby te są równe wtedy i tylko wtedy, gdy mają one równe moduły i gdy ich argumenty różnią się o całkowitą krotność liczby 2π. Dla iloczynu liczb z i w mamy

$$\begin{aligned} zw &= |z||w|(\cos\alpha + j\sin\alpha)(\cos\beta + j\sin\beta) \\ &= |z||w|\big((\cos\alpha\cos\beta - \sin\alpha\sin\beta) + j(\sin\alpha\cos\beta + \cos\alpha\sin\beta)\big) \\ &= |z||w|\big(\cos(\alpha+\beta) + j\sin(\alpha+\beta)\big) \\ &= |zw|\big(\cos(\alpha+\beta) + j\sin(\alpha+\beta)\big), \end{aligned}$$

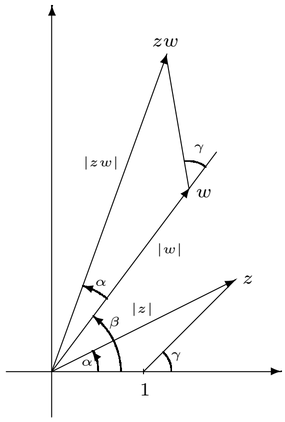
co jest postacią trygonometryczną liczby zw i skąd ponownie wynika, że "moduł iloczynu jest równy iloczynowi modułów", $|zw| = |z||w|$, i dodatkowo, że suma argumentów liczb z i w jest argumentem iloczynu zw,

$$\arg\,(z) + \arg\,(w) = \arg\,(zw). \tag{2.22}$$

[1] W niektórych podręcznikach argument główny liczby z oznacza się przez $\arg\,(z)$, a symbolu $\mathrm{Arg}\,(z)$ używa się na oznaczenie zbioru wszystkich argumentów liczby z.

Obie te obserwacje pozwalają uzasadnić poprawność naszkicowanej na rys. 2.11 geometrycznej konstrukcji iloczynu zw liczb zespolonych z i w. Łatwo także zaobserwować, że dla argumentu odwrotności liczby i ilorazu liczb zespolonych mamy

$$-\arg(z) = \arg\left(\frac{1}{z}\right) \quad \text{i} \quad \arg(z) - \arg(w) = \arg\left(\frac{z}{w}\right). \tag{2.23}$$



Rys. 2.11

Zatem mamy następujące twierdzenie.

Twierdzenie 2.3.1. *Jeśli $z = |z|(\cos\alpha + j\sin\alpha)$ i $w = |w|(\cos\beta + j\sin\beta)$, to*

$$zw = |z||w|\big(\cos(\alpha+\beta) + j\sin(\alpha+\beta)\big) \tag{2.24}$$

i

$$\frac{z}{w} = \frac{|z|}{|w|}\big(\cos(\alpha-\beta) + j\sin(\alpha-\beta)\big), \tag{2.25}$$

gdy $w \neq 0$. □

Przykład 24. Liczby $z = \sqrt{3} + j$ oraz $w = 1 + j$ przedstawić w postaci trygonometrycznej. Następnie znaleźć postać trygonometryczną każdej z liczb zw i z/w.

Ponieważ $|z| = 2$, $\cos\alpha = \frac{\sqrt{3}}{2}$ i $\sin\alpha = \frac{1}{2}$, więc $\alpha = \frac{\pi}{6}$ i $z = 2\big(\cos\frac{\pi}{6} + j\sin\frac{\pi}{6}\big)$. Podobnie $|w| = \sqrt{2}$, $\cos\beta = \frac{1}{\sqrt{2}}$ i $\sin\beta = \frac{1}{\sqrt{2}}$, więc $\beta = \frac{\pi}{4}$ i $w = \sqrt{2}\big(\cos\frac{\pi}{4} + j\sin\frac{\pi}{4}\big)$. Zatem

$$zw = 2\sqrt{2}\left(\cos\left(\frac{\pi}{6}+\frac{\pi}{4}\right) + j\sin\left(\frac{\pi}{6}+\frac{\pi}{4}\right)\right) = 2\sqrt{2}\left(\cos\frac{5\pi}{12} + j\sin\frac{5\pi}{12}\right)$$

i

$$\frac{z}{w} = \frac{2}{\sqrt{2}}\left(\cos\left(\frac{\pi}{6}-\frac{\pi}{4}\right) + j\sin\left(\frac{\pi}{6}-\frac{\pi}{4}\right)\right) = \sqrt{2}\left(\cos\left(\frac{-\pi}{12}\right) + j\sin\left(\frac{-\pi}{12}\right)\right).$$

Przy obliczaniu potęg liczb zespolonych można posłużyć się tzw. wzorem de Moivre'a. Wzór ten jest prostą konsekwencją twierdzenia 2.3.1.

Wniosek 2.3.1. *Jeśli $z = |z|(\cos\alpha + j\sin\alpha)$ i n jest liczbą całkowitą, to*

$$z^n = |z|^n(\cos n\alpha + j\sin n\alpha), \tag{2.26}$$

Wzór de Moivre'a

gdzie $z \neq 0$ dla $n \leqslant 0$. □

Wniosek 2.3.2. *Dla każdej liczby rzeczywistej α i każdej liczby całkowitej n jest*

$$(\cos\alpha + j\sin\alpha)^n = \cos n\alpha + j\sin n\alpha. \; \square \tag{2.27}$$

Przykład 25. Obliczyć

$$z_1 = \left(-\frac{1}{2} + j\frac{\sqrt{3}}{2}\right)^{17} \quad \text{i} \quad z_2 = \left(\frac{1+j\sqrt{3}}{1-j}\right)^{12}.$$

Mamy $z_1 = \left(-\frac{1}{2} + j\frac{\sqrt{3}}{2}\right)^{17} = \left(\cos\frac{2}{3}\pi + j\sin\frac{2}{3}\pi\right)^{17}$ i z wniosku 2.3.2

$$\begin{aligned} z_1 &= \cos\tfrac{34}{3}\pi + j\sin\tfrac{34}{3}\pi = \cos\left(10\pi + \tfrac{4}{3}\pi\right) + j\sin\left(10\pi + \tfrac{4}{3}\pi\right) \\ &= \cos\tfrac{4}{3}\pi + j\sin\tfrac{4}{3}\pi = -\tfrac{1}{2} - j\tfrac{\sqrt{3}}{2}. \end{aligned}$$

Ponieważ $1 + j\sqrt{3} = 2\big(\cos\frac{\pi}{3} + j\sin\frac{\pi}{3}\big)$ i $1 - j = \sqrt{2}\big(\cos(-\frac{\pi}{4}) + j\sin(-\frac{\pi}{4})\big)$, więc z twierdzenia 2.3.1 otrzymujemy

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

Stąd i z wniosku 2.3.1 mamy

$$z_2 = \left(\sqrt{2}\left(\cos\frac{7\pi}{12} + j\sin\frac{7\pi}{12}\right)\right)^{12} = 2^6\left(\cos 7\pi + j\sin 7\pi\right) = -64.$$

Przykład 26. Liczbę $z = \sin\alpha - j\cos\alpha$ zapisać w postaci trygonometrycznej.

Liczbę z chcemy zapisać w postaci $|z|(\cos\beta + j\sin\beta) = |z|\left(\frac{a}{|z|} + j\frac{b}{|z|}\right)$. Można to zrobić na dwa sposoby. Ponieważ $|z| = 1$ i $\cos\beta = \frac{a}{|z|} = \sin\alpha = \cos(\alpha - \pi/2)$ oraz $\sin\beta = \frac{b}{|z|} = -\cos\alpha = \sin(\alpha - \pi/2)$, więc

$$z = \cos(\alpha - \pi/2) + j\sin(\alpha - \pi/2).$$

Równoważnie, $z = -j(\cos\alpha + j\sin\alpha)$ i ponieważ $-j = \cos(-\pi/2) + j\sin(-\pi/2)$, więc także jest

$$\begin{aligned} z &= \big(\cos(-\pi/2) + j\sin(-\pi/2)\big)(\cos\alpha + j\sin\alpha) \\ &= \cos(\alpha - \pi/2) + j\sin(\alpha - \pi/2). \end{aligned}$$

Przykład 27. Korzystając ze wzorów de Moivre'a i Newtona, wyrazić $\cos 3x$ i $\sin 3x$ za pomocą $\cos x$ i $\sin x$.

Wzór Newtona:

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

Ze wzoru de Moivre'a mamy $(\cos x + j\sin x)^3 = \cos 3x + j\sin 3x$. Jednocześnie ze wzoru Newtona

$$\begin{aligned} (\cos x + j\sin x)^3 &= \cos^3 x + 3j\cos^2 x\sin x - 3\cos x\sin^2 x - j\sin^3 x \\ &= (\cos^3 x - 3\cos x\sin^2 x) + j(3\cos^2 x\sin x - \sin^3 x). \end{aligned}$$

Stąd i z warunku równości liczb zespolonych otrzymujemy

$$\cos 3x = \cos^3 x - 3\cos x\sin^2 x$$

oraz

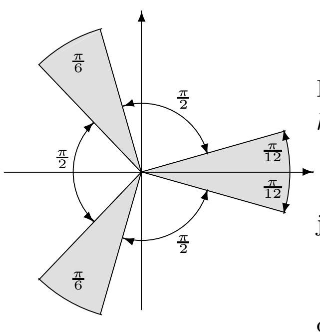
$$\sin 3x = 3\cos^2 x\sin x - \sin^3 x.$$

Jeśli uwzględni się, że $\sin^2 x + \cos^2 x = 1$, to prawe strony ostatnich tożsamości można wyrazić za pomocą potęg tylko $\cos x$ lub tylko $\sin x$,

$$\cos 3x = 4\cos^3 x - 3\cos x \quad \text{i} \quad \sin 3x = 3\sin x - 4\sin^3 x.$$

Przykład 28. Wyznaczyć zbiór liczb zespolonych z spełniających nierówności

$$\frac{\pi}{2} \leqslant \operatorname{Arg}\left((-1+j)z^3\right) \leqslant \pi.$$



Rys. 2.12

Ponieważ $\operatorname{Arg}\left((-1+j)z^3\right) = \operatorname{Arg}(-1+j) + 3\operatorname{Arg}(z) + 2k\pi$ dla pewnej liczby całkowitej k i $\operatorname{Arg}(-1+j) = \frac{3\pi}{4}$, więc nierówność

$$\frac{\pi}{2} \leqslant \operatorname{Arg}\left((-1+j)z^3\right) \leqslant \pi$$

jest równoważna nierówności

$$\frac{\pi}{2} \leqslant \operatorname{Arg}(-1+j) + 3\operatorname{Arg}(z) + 2k\pi \leqslant \pi$$

dla pewnego $k \in Z$. Stąd

$$-\frac{\pi}{12} - \frac{2k\pi}{3} \leqslant \operatorname{Arg}(z) \leqslant \frac{\pi}{12} - \frac{2k\pi}{3}$$

dla pewnego $k \in Z$. Jednocześnie $-\pi < \operatorname{Arg}(z) \leqslant \pi$, więc $k = 0$, $k = 1$ lub $k = -1$ i $-\frac{\pi}{12} \leqslant \operatorname{Arg}(z) \leqslant \frac{\pi}{12}$, $-\frac{9\pi}{12} \leqslant \operatorname{Arg}(z) \leqslant -\frac{7\pi}{12}$ lub $\frac{7\pi}{12} \leqslant \operatorname{Arg}(z) \leqslant \frac{9\pi}{12}$. Szukany zbiór składa się z trzech części przedstawionych na rys. 2.12.

2.4. Pierwiastkowanie liczb zespolonych

Definicja 2.4.1. Liczbę w nazywamy *pierwiastkiem n-tego stopnia* z liczby zespolonej z ($n \in N$), gdy

$$w^n = z.$$

Pierwiastki n-tego stopnia

O pierwiastkach stopnia naturalnego z liczby zespolonej mówi następujące twierdzenie.

Twierdzenie 2.4.1. *Każda liczba zespolona $z = |z|(\cos\alpha + j\sin\alpha)$ różna od zera ma dokładnie n różnych pierwiastków n-tego stopnia i wszystkie one określone są wzorem*

$$w_k = \sqrt[n]{|z|}\left(\cos\frac{\alpha+2k\pi}{n} + j\sin\frac{\alpha+2k\pi}{n}\right), \tag{2.28}$$

Pierwiastki n-tego stopnia z liczby zespolonej z

gdzie $k = 0, 1, \ldots, n-1$, a $\sqrt[n]{|z|}$ jest pierwiastkiem arytmetycznym.

Dowód. Liczba $w = |w|(\cos\varphi + j\sin\varphi)$ jest pierwiastkiem n-tego stopnia z liczby z wtedy i tylko wtedy, gdy zachodzą następujące równoważności

$$\begin{aligned}
& w^n = z && \text{(z def. pierwiastka)}\\
\Leftrightarrow\ & |w|^n(\cos\varphi + j\sin\varphi)^n = |z|(\cos\alpha + j\sin\alpha) &&\\
\Leftrightarrow\ & |w|^n(\cos n\varphi + j\sin n\varphi) = |z|(\cos\alpha + j\sin\alpha) && \text{(ze wzoru de Moivre'a)}\\
\Leftrightarrow\ & |w|^n = |z| \quad \text{i} \quad n\varphi = \alpha + 2k\pi \quad \text{dla} \quad k \in Z && \text{(z równości liczb)}\\
\Leftrightarrow\ & |w| = \sqrt[n]{|z|} \quad \text{i} \quad \varphi = \frac{\alpha+2k\pi}{n} \quad \text{dla} \quad k \in Z &&\\
\Leftrightarrow\ & w \in \left\{w_k = \sqrt[n]{|z|}\left(\cos\tfrac{\alpha+2k\pi}{n} + j\sin\tfrac{\alpha+2k\pi}{n}\right) : k \in Z\right\} = \mathcal{P}. &&
\end{aligned}$$

Dla zakończenia dowodu wystarczy teraz pokazać, że liczby $w_0, w_1, \ldots, w_{n-1}$ są różne i $\mathcal{P} \subseteq \{w_0, w_1, \ldots, w_{n-1}\}$. W tym celu zauważmy najpierw, że dwie liczby w_k i w_l ze zbioru $\mathcal{P}$ są równe wtedy i tylko wtedy, gdy

$$\begin{aligned}
\tfrac{\alpha+2k\pi}{n} &= \tfrac{\alpha+2l\pi}{n} + 2m\pi \quad \text{dla pewnej liczby } m \in Z\\
&\Leftrightarrow\ k = l + mn \text{ dla pewnej liczby } m \in Z\\
&\Leftrightarrow\ k \text{ i } l \text{ różnią się o całkowitą krotność liczby } n.
\end{aligned}$$

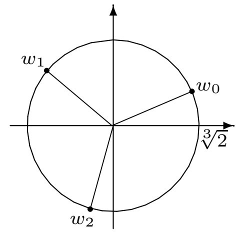
Ponieważ żadne dwie liczby ze zbioru $\{0, 1, \ldots, n-1\}$ nie różnią się o całkowitą krotność liczby n, więc liczby $w_0, w_1, \ldots, w_{n-1}$ są różne. Niech teraz w_l będzie dowolnym elementem ze zbioru $\mathcal{P}$. Ponieważ liczba l różni się o całkowitą krotność liczby n od dokładnie jednej liczby k ze zbioru $\{0, 1, \ldots, n-1\}$, więc $w_l = w_k$. Stąd $\mathcal{P} \subseteq \{w_0, w_1, \ldots, w_{n-1}\}$ i to kończy dowód twierdzenia. □

Moduł każdej z liczb określonych wzorem (2.28) jest równy $\sqrt[n]{|z|}$, więc wszystkie one leżą na okręgu o promieniu $\sqrt[n]{|z|}$ i środku w początku układu współrzędnych. Dodatkowo, dzielą one ten okrąg na n równych części, bo $\arg(w_k) - \arg(w_{k-1}) = 2\pi/n$ dla $k = 1, \ldots, n-1$. Równoważnie, pierwiastki $w_0, w_1, \ldots, w_{n-1}$ stopnia $n \geqslant 3$ z liczby zespolonej z są wierzchołkami n-kąta foremnego wpisanego w okrąg o promieniu $\sqrt[n]{|z|}$ i środku w początku układu współrzędnych.

Przykład 29. Obliczyć pierwiastki stopnia trzeciego z liczby

$$z = 1 + j\sqrt{3}.$$

Ponieważ $|z| = |1 + j\sqrt{3}| = 2$ i $\cos\alpha = \frac{1}{2}$ oraz $\sin\alpha = \frac{\sqrt{3}}{2}$, więc można przyjąć, że $\arg(z) = \frac{\pi}{3}$. Wtedy (na podstawie (2.28)) pierwiastkami stopnia trzeciego z liczby



Rys. 2.13

$z = 1 + j\sqrt{3}$ są liczby

$$\begin{aligned} w_0 &= \sqrt[3]{2}\big(\cos\tfrac{\pi}{9} + j\sin\tfrac{\pi}{9}\big),\\ w_1 &= \sqrt[3]{2}\big(\cos\tfrac{7}{9}\pi + j\sin\tfrac{7}{9}\pi\big),\\ w_2 &= \sqrt[3]{2}\big(\cos\tfrac{13}{9}\pi + j\sin\tfrac{13}{9}\pi\big) \end{aligned}$$

przedstawione na rys. 2.13.

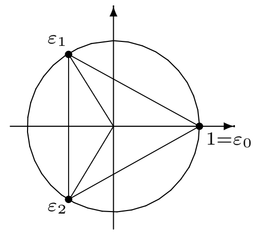
Z twierdzenia 2.4.1 dla jedności, tj. dla $z = 1 = 1(\cos 0 + j\sin 0)$, mamy następujący wniosek.

Pierwiastki z jedności

Wniosek 2.4.1. *Pierwiastki n-tego stopnia z jedności określone są wzorem*

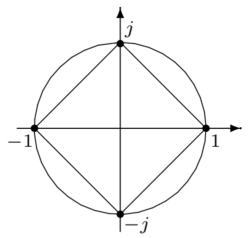
$$\varepsilon_k = \cos\frac{2k\pi}{n} + j\sin\frac{2k\pi}{n} \tag{2.29}$$

dla $k = 0, 1, \ldots, n-1$. $\square$



Rys. 2.14

Warto zauważyć, że jeśli $\varepsilon = \varepsilon_1$, to (ponieważ $\varepsilon^k = \varepsilon_1^k = \varepsilon_k$ dla $k = 0, 1, \ldots, n-1$) liczby $1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{n-1}$ są wszystkimi pierwiastkami n-tego stopnia z jedności i – jak to już wcześniej powiedzieliśmy – liczby te są wierzchołkami n-kąta foremnego wpisanego w okrąg o środku w początku układu współrzędnych i promieniu długości jeden, zob. rys. 2.14 dla $n = 3$, rys. 2.15 dla $n = 4$ i rys. 2.16 dla $n = 6$. Łatwo także zauważyć następującą własność pierwiastków n-tego stopnia z jedności, własność która może być przydatna przy wyznaczaniu pierwiastków n-tego stopnia z innych liczb zespolonych.



Rys. 2.15

Wniosek 2.4.2. *Jeśli w jest jakimkolwiek pierwiastkiem n-tego stopnia z liczby $z \neq 0$ i $\varepsilon = \cos\frac{2\pi}{n} + j\sin\frac{2\pi}{n}$, to liczby*

$$w,\ w\varepsilon,\ w\varepsilon^2,\ \ldots,\ w\varepsilon^{n-1} \tag{2.30}$$

są wszystkimi pierwiastkami n-tego stopnia z liczby z. $\square$

Przykład 30. Wyznaczyć pierwiastki stopnia trzeciego z liczby $(1+j)^6$.

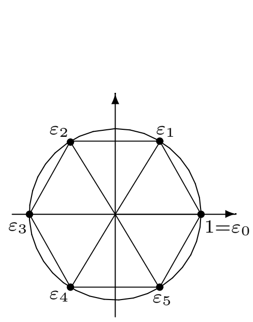
Ponieważ $w = (1+j)^2 = 2j$ jest pierwiastkiem stopnia trzeciego z liczby $(1+j)^6$ i $\varepsilon = \cos\frac{2\pi}{3} + j\sin\frac{2\pi}{3} = -\frac{1}{2} + j\frac{\sqrt{3}}{2}$ jest pierwiastkiem stopnia trzeciego z jedności, więc wobec wniosku 2.4.2 także liczby

$$w\varepsilon = 2j\left(-\frac{1}{2} + j\frac{\sqrt{3}}{2}\right) = -\sqrt{3} - j$$

i

$$w\varepsilon^2 = 2j\left(-\frac{1}{2} + j\frac{\sqrt{3}}{2}\right)^2 = 2j\left(-\frac{1}{2} - j\frac{\sqrt{3}}{2}\right) = \sqrt{3} - j$$

są pierwiastkami stopnia trzeciego z liczby $(1+j)^6$.



Rys. 2.16

Przykład 31. Znaleźć wszystkie rozwiązania równania

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

Zauważmy, że $x = 2$ nie jest rozwiązaniem równania $(2x+1)^4 = (x-2)^4$. Natomiast dla $x \neq 2$ mamy równoważności

$$(2x+1)^4 = (x-2)^4 \Leftrightarrow \left(\frac{2x+1}{x-2}\right)^4 = 1 \Leftrightarrow \frac{2x+1}{x-2} = \varepsilon_k \Leftrightarrow x = \frac{1+2\varepsilon_k}{\varepsilon_k - 2},$$

gdzie ε_k jest pierwiastkiem stopnia czwartego z jedności, $\varepsilon_k = \cos\frac{2k\pi}{4} + j\sin\frac{2k\pi}{4}$ $(k = 0, 1, 2, 3)$. Stąd już łatwo można zaobserwować, że x jest rozwiązaniem równania $(2x+1)^4 = (x-2)^4$ wtedy i tylko wtedy, gdy x jest elementem zbioru $\{-3,\ -j,\ 1/3,\ j\}$.

Pierwiastki stopnia drugiego z liczby zespolonej

Ponieważ pierwiastki stopnia drugiego z liczby zespolonej będą często pojawiały się w naszych rozważaniach, poświęcamy im więcej uwagi. Przede wszystkim, jeśli znana jest postać trygonometryczna liczby z, powiedzmy $z = |z|(\cos\varphi + j\sin\varphi)$, to wobec twierdzenia 2.4.1 pierwiastkami stopnia drugiego z liczby z są różniące się znakiem liczby w_0 i w_1, gdzie

w_0, w_1 – pierwiastki stopnia drugiego z liczby zespolonej z

$$w_0 = \sqrt{|z|}\big(\cos(\varphi/2) + j\sin(\varphi/2)\big)$$

i

$$w_1 = \sqrt{|z|}\big(\cos(\varphi/2 + \pi) + j\sin(\varphi/2 + \pi)\big) = -w_0.$$

Przykład 32. Pierwiastkami stopnia drugiego z liczby $-25 = 25(\cos\pi + j\sin\pi)$ są liczby

$$w_0 = \sqrt{25}\left(\cos\frac{\pi}{2} + j\sin\frac{\pi}{2}\right) = 5j \quad \text{i} \quad w_1 = \sqrt{25}\left(\cos\frac{3}{2}\pi + j\sin\frac{3}{2}\pi\right) = -5j.$$

Podobnie pierwiastkami stopnia drugiego z liczby $4j = 4(\cos\frac{\pi}{2} + j\sin\frac{\pi}{2})$ są liczby

$$\pm 2\left(\cos\frac{\pi}{4} + j\sin\frac{\pi}{4}\right) = \pm(\sqrt{2} + j\sqrt{2}).$$

Przy rozwiązywaniu równań kwadratowych zwykle wyznacza się pierwiastki stopnia drugiego z liczby zespolonej zapisanej w postaci kanonicznej. Zauważmy, że liczba $w = x + jy$ jest pierwiastkiem stopnia drugiego z liczby zespolonej $z = a + jb$ (gdzie a, b, x, $y \in R$) wtedy i tylko wtedy, gdy $w^2 = z$, tj. wtedy i tylko wtedy, gdy

$$(x + jy)^2 = x^2 - y^2 + 2xyj = a + jb. \tag{2.31}$$

Równanie to jest równoważne układowi równań

$$\begin{cases} x^2 - y^2 &= a, \\ 2xy &= b, \end{cases} \tag{2.32}$$

z którego łatwo można wyznaczyć liczby x i y (i dlatego także otrzymać liczbę $w = x + jy$).

Przykład 33. Obliczyć pierwiastki stopnia drugiego z liczby $9 - 40j$.

Niech $x + jy$ (gdzie x, $y \in R$) będzie pierwiastkiem stopnia drugiego z liczby $9 - 40j$. Wtedy

$$(x + jy)^2 = x^2 - y^2 + 2xyj = 9 - 40j$$

i dlatego $x^2 - y^2 = 9$, $xy = -20$ i $x^2y^2 = 400$. Otrzymane z pierwszego z tych równań $x^2 = y^2 + 9$ podstawiamy do trzeciego równania i otrzymujemy

$$y^4 + 9y^2 - 400 = 0.$$

Stąd zaś $y^2 = 16$ i dlatego

$$y = \pm 4 \quad \text{i} \quad x = -\tfrac{20}{y} = \mp 5.$$

Zatem pierwiastkami stopnia drugiego z liczby $9 - 40j$ są liczby $5 - 4j$ oraz $-5 + 4j$.

$$\operatorname{sign}(x) = \begin{cases} 1 & \text{dla } x \geqslant 0 \\ -1 & \text{dla } x < 0 \end{cases}$$

W kolejnym twierdzeniu przedstawiamy jeszcze jeden sposób wyznaczania pierwiastków stopnia drugiego z liczby zespolonej. W twierdzeniu tym $\operatorname{sign}(x)$ jest "znakiem" liczby rzeczywistej x, tj. $\operatorname{sign}(x) = 1$ dla $x \geqslant 0$ oraz $\operatorname{sign}(x) = -1$ dla $x < 0$.

Twierdzenie 2.4.2. *Jeśli a i b są liczbami rzeczywistymi, to pierwiastkami stopnia drugiego z liczby $z = a + jb$ są liczby*

$$w = \pm\left(\sqrt{\frac{|z| + a}{2}} + j\operatorname{sign}(b)\sqrt{\frac{|z| - a}{2}}\right). \tag{2.33}$$

Dowód. Wystarczy zauważyć, że istotnie

$$\left(\pm\left(\sqrt{(|z| + a)/2} + j\operatorname{sign}(b)\sqrt{(|z| - a)/2}\right)\right)^2 = a + jb. \ \square$$

Przykład 34. Korzystając z ostatniego twierdzenia, wyznaczyć pierwiastki stopnia drugiego z liczby $z = 3 + 4j$.

Ponieważ $z = a + jb = 3 + 4j$ i $|z| = 5$, więc wobec (2.33) pierwiastkami stopnia drugiego z liczby $3 + 4j$ są

$$w = \pm\left(\sqrt{(5 + 3)/2} + j\operatorname{sign}(4)\sqrt{(5 - 3)/2}\right) = \pm(2 + j).$$

Umiejętność wyznaczania pierwiastków stopnia drugiego z liczby zespolonej jest ważna przy wyznaczaniu pierwiastków równania stopnia drugiego.

Twierdzenie 2.4.3. *Pierwiastkami równania kwadratowego $ax^2 + bx + c = 0$, w którym $a, b, c \in \mathcal{C}$ i $a \neq 0$, są liczby*

$$x = \frac{-b - \delta}{2a} \quad i \quad x = \frac{-b + \delta}{2a}, \tag{2.34}$$

gdzie δ jest jednym z dwóch pierwiastków stopnia drugiego z liczby $\Delta = b^2 - 4ac$.

Dowód. Jeśli δ jest pierwiastkiem stopnia drugiego z liczby $\Delta = b^2 - 4ac$, to mamy $b^2 - 4ac = \delta^2$ i

$$\begin{aligned} ax^2 + bx + c &= a\left(x^2 + \tfrac{b}{a}x + \tfrac{c}{a}\right) = a\left(\left(x + \tfrac{b}{2a}\right)^2 + \tfrac{c}{a} - \tfrac{b^2}{4a^2}\right) \\ &= a\left(\left(x + \tfrac{b}{2a}\right)^2 - \tfrac{b^2 - 4ac}{4a^2}\right) = a\left(\left(x + \tfrac{b}{2a}\right)^2 - \tfrac{\delta^2}{4a^2}\right) \\ &= a\left(\left(x + \tfrac{b}{2a}\right) + \tfrac{\delta}{2a}\right)\left(\left(x + \tfrac{b}{2a}\right) - \tfrac{\delta}{2a}\right) \\ &= a\left(x - \tfrac{-b-\delta}{2a}\right)\left(x - \tfrac{-b+\delta}{2a}\right). \end{aligned}$$

Stąd wynika, że liczby $x = \frac{-b \pm \delta}{2a}$ są pierwiastkami równania $ax^2 + bx + c = 0$. $\square$

Przykład 35. Rozwiązać równanie $x^2 - 2x + 2 = 0$.

Ponieważ $\Delta = b^2 - 4ac = -4$ i pierwiastkiem stopnia drugiego z liczby $\Delta = -4$ jest $\delta = \pm 2j$, więc wobec (2.34) rozwiązaniem równania są liczby

$$x = \frac{-b \pm \delta}{2a} = 1 \pm j.$$

Przykład 36. Rozwiązać równanie $x^2 - (2+j)x + (-1+7j) = 0$.

W tym przypadku jest $\Delta = (2+j)^2 - 4(-1+7j) = 7 - 24j$, $|\Delta| = 25$ i wobec (2.33) jednym pierwiastkiem stopnia drugiego z liczby $\Delta = 7 - 24j$ jest

$$\delta = \sqrt{(25+7)/2} - j\sqrt{(25-7)/2} = 4 - 3j.$$

Stąd i z (2.34) wynika, że rozwiązaniem równania są liczby

$$x_1 = 3 - j \quad \text{i} \quad x_2 = -1 + 2j.$$

2.5. Wzory Eulera

Funkcją wykładniczą e^z i funkcjami trygonometrycznymi $\cos z$ i $\sin z$ zmiennej zespolonej z nazywa się funkcje określone szeregami potęgowymi: $e \approx 2,7182\ldots$

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \ldots + \frac{z^n}{n!} + \ldots, \tag{2.35}$$

$$\cos z = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!} = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \ldots, \tag{2.36}$$

$$\sin z = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!} = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \ldots. \tag{2.37}$$

Funkcje określone powyższymi równościami są naturalnymi uogólnieniami rzeczywistej funkcji wykładniczej e^x i rzeczywistych funkcji trygonometrycznych $\cos x$ i $\sin x$.[2] Z (2.35) po podstawieniu jz zamiast z otrzymujemy

$$\begin{aligned} e^{jz} &= 1 + jz + \frac{(jz)^2}{2!} + \frac{(jz)^3}{3!} + \frac{(jz)^4}{4!} + \ldots \\ &= 1 + jz - \frac{z^2}{2!} - \frac{jz^3}{3!} + \frac{z^4}{4!} + \frac{jz^5}{5!} + \ldots. \end{aligned}$$

Stąd zaś, po oddzieleniu składników zawierających z w potęgach parzystych od tych, w których z występuje w potęgach nieparzystych, otrzymujemy

$$e^{jz} = \left(1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \ldots\right) + j\left(z - \frac{z^3}{3!} + \frac{z^5}{5!} - \ldots\right). \tag{2.38}$$

Z (2.38), (2.36) i (2.37) jest oczywiste, że dla każdego $z \in \mathcal{C}$ mamy

$$e^{jz} = \cos z + j \sin z. \tag{2.39}$$

Z tych samych powodów

$$e^{-jz} = \cos z - j \sin z. \tag{2.40}$$

[2] W analizie matematycznej dowodzi się, że funkcje $E(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$, $C(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}$ i $S(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}$ są określone dla każdej liczby zespolonej $z \in \mathcal{C}$. Tamże dowodzi się, że dla każdego rzeczywistego x jest $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$, $\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$ i $\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$. Zatem funkcje zespolone $E(z)$, $C(z)$ i $S(z)$ są uogólnieniami funkcji rzeczywistych e^x, $\cos x$ i $\sin x$. Z tego też względu są one oznaczane przez e^z, $\cos z$ i $\sin z$ i nazywane funkcją wykładniczą, cosinusem i sinusem zmiennej zespolonej z.

Dodając lub odejmując stronami równości (2.39) i (2.40), otrzymujemy

$$e^{jz} + e^{-jz} = 2\cos z \quad \text{i} \quad e^{jz} - e^{-jz} = 2j\sin z.$$

Zatem między funkcją wykładniczą e^z oraz funkcjami trygonometrycznymi $\cos z$ i $\sin z$ zmiennej zespolonej z zachodzą następujące związki, zwane wzorami Eulera.

Wzory Eulera

Twierdzenie 2.5.1. *Dla każdej liczby zespolonej z jest*

$$e^{jz} = \cos z + j\sin z, \tag{2.41}$$

oraz

$$\cos z = \frac{e^{jz} + e^{-jz}}{2} \quad i \quad \sin z = \frac{e^{jz} - e^{-jz}}{2j}. \ \square \tag{2.42}$$

Przykład 37. Z (2.41) mamy $e^{j\pi} = \cos\pi + j\sin\pi = -1 + j\cdot 0 = -1$, więc także

$$e^{j\pi} + 1 = 0.$$

Ostatnia równość jest związkiem pomiędzy podstawowymi stałymi matematycznymi (0, 1, π, e oraz j). Przez studentów amerykańskich zależność ta została uznana za najpiękniejszy wzór matematyki.

Z powyższych definicji i ze wzoru Eulera można otrzymać następujące własności funkcji wykładniczej e^z.

$e^{x+jy} = e^x(\cos y + j\sin y)$

Twierdzenie 2.5.2. *Dla liczby zespolonej $z = x + jy$ (gdzie $x, y \in R$) jest:*

(*a*) $e^z = e^{x+jy} = e^x e^{jy} = e^x(\cos y + j\sin y)$;

(*b*) $|e^{jy}| = 1$, $|e^{x+jy}| = e^x$ *i* $\arg\left(e^{x+jy}\right) = y$. $\square$

Dowodząc (*a*) trzeba pokazać, że funkcja $e^x \cdot e^{jy}$, czyli iloczyn szeregów $\left(\sum_{n=0}^{\infty} \frac{x^n}{n!}\right)$ i $\left(\sum_{n=0}^{\infty} \frac{(jx)^n}{n!}\right)$, jest identyczna z funkcją e^{x+jy}, czyli z szeregiem $\sum_{n=0}^{\infty} \frac{(x+jy)^n}{n!}$. Formalny dowód tych zależności można znaleźć w podręcznikach analizy matematycznej (przykładowo zob. [3]).[3] Stwierdzenie (*b*) jest już prostą konsekwencją (*a*) (oraz definicji modułu i argumentu liczby zespolonej). Z drugiej części (*b*) łatwo wynika, że funkcja wykładnicza e^z nie przyjmuje wartości 0.

Wniosek 2.5.1. *Dla każdej liczby zespolonej z jest $e^z \neq 0$.* $\square$

Dla rzeczywistych liczb x_1, x_2, x oraz dla liczby naturalnej n jest $e^{x_1}e^{x_2} = e^{x_1+x_2}$, $e^{x_1}/e^{x_2} = e^{x_1-x_2}$ i $(e^x)^n = e^{nx}$. Teraz udowodnimy, że te same własności ma funkcja wykładnicza e^z zmiennej zespolonej z.

Wniosek 2.5.2. *Dla każdych liczb zespolonych z_1, z_2 i z oraz każdej liczby naturalnej n jest*

$$e^{z_1}e^{z_2} = e^{z_1+z_2}, \quad \frac{e^{z_1}}{e^{z_2}} = e^{z_1-z_2} \quad i \quad (e^z)^n = e^{nz}.$$

[3] W wielu podręcznikach przyjmuje się, że iloczyn $e^x(\cos y + j\sin y)$ jest definicją funkcji wykładniczej e^{x+jy} (dla $x, y \in R$).

Dowód. Udowodnimy tylko pierwszą równość. Dwie następne wynikają z pierwszej. Zauważmy, że jeżeli $z_1 = x_1 + jy_1$ i $z_2 = x_2 + jy_2$ $(x_i, y_i \in R)$, to wobec twierdzenia 2.5.2 i 2.3.1 mamy

$$\begin{aligned} e^{z_1}e^{z_2} &= e^{x_1+jy_1}\cdot e^{x_2+jy_2} \\ &= e^{x_1}(\cos y_1 + j\sin y_1)\cdot e^{x_2}(\cos y_2 + j\sin y_2) \\ &= e^{x_1+x_2}\big(\cos(y_1+y_2) + j\sin(y_1+y_2)\big) \\ &= e^{x_1+x_2+j(y_1+y_2)} = e^{(x_1+jy_1)+(x_2+jy_2)} = e^{z_1+z_2}. \ \square \end{aligned}$$

Wniosek 2.5.3. *$e^z = 1$ wtedy i tylko wtedy, gdy $z = 2k\pi j$, gdzie k jest liczbą całkowitą.*

Dowód. Zauważmy najpierw, że jeśli k jest liczbą całkowitą, to $e^{2k\pi j} = \cos 2k\pi + j\sin 2k\pi = 1 + j0 = 1$.

Niech teraz $z = x + jy$ $(x, y \in R)$ będzie takie, że $e^z = e^x(\cos y + j\sin y) = 1$. Wtedy $e^x \sin y = 0$ i dlatego $y = l\pi$ (dla pewnej liczby całkowitej l). Jednocześnie $e^x \cos y = e^x \cos l\pi = e^x(-1)^l = 1$ wtedy i tylko wtedy, gdy $x = 0$ i l jest liczbą parzystą, $l = 2k$. Stąd $z = x + jy = 0 + j2k\pi = 2k\pi j$ dla pewnej liczby całkowitej k. $\square$

Wniosek 2.5.4. *Dla liczb zespolonych z_1 i z_2 jest $e^{z_1} = e^{z_2}$ wtedy i tylko wtedy, gdy $z_1 - z_2 = 2k\pi j$ dla pewnej liczby całkowitej k.* $\square$

Z definicji funkcji trygonometrycznych $\cos z$ i $\sin z$, ze wzorów Eulera oraz z wniosków 2.5.1–2.5.4 łatwo wyprowadza się (i to bez odwoływania się do geometrii) wszystkie wzory redukcyjne i wszystkie tożsamości trygonometryczne znane dla rzeczywistych funkcji trygonometrycznych. Tu przykładowo wyprowadzimy wzór na sumę sinusów.

Przykład 38. Dla wszystkich zespolonych x i y jest

$$\sin x + \sin y = 2\sin\frac{x+y}{2}\cos\frac{x-y}{2}.$$

Ze wzorów Eulera ((2.41) i (2.42)) oraz z wniosku 2.5.2 mamy

$$\begin{aligned} 2\sin\frac{x+y}{2}\cos\frac{x-y}{2} &= 2\cdot\frac{e^{j\frac{x+y}{2}} - e^{-j\frac{x+y}{2}}}{2j}\cdot\frac{e^{j\frac{x-y}{2}} + e^{-j\frac{x-y}{2}}}{2} \\ &= \frac{1}{2j}\left(e^{jx} + e^{jy} - e^{-jy} - e^{-jx}\right) \\ &= \frac{e^{jx} - e^{-jx}}{2j} + \frac{e^{jy} - e^{-jy}}{2j} = \sin x + \sin y. \end{aligned}$$

2.6. Postać wykładnicza liczby zespolonej

Ponieważ każdą różną od zera liczbę zespoloną z można zapisać w postaci trygonometrycznej

$$z = |z|(\cos\varphi + j\sin\varphi),$$

gdzie $\varphi = \arg(z)$, więc wobec wzoru Eulera można ją także zapisać w postaci

$$z = |z|e^{j\varphi}, \tag{2.43}$$

Postać wykładnicza liczby zespolonej

zwanej *postacią wykładniczą* liczby zespolonej z.

Przykładowo mamy

$$1 + j = \sqrt{2}e^{j\frac{\pi}{4}} \quad \text{i} \quad -\sqrt{3} - j = 2e^{-j\frac{5\pi}{6}} = 2e^{j\frac{7\pi}{6}}.$$

Z własności liczb zespolonych w postaci trygonometrycznej natychmiast wynikają następujące własności liczb zespolonych w postaci wykładniczej.

Wniosek 2.6.1. *Dla liczb zespolonych* $z = |z|e^{j\varphi}$ *i* $w = |w|e^{j\psi}$ *oraz liczby naturalnej* n *mamy:*

(1) $z = w$ *wtedy i tylko wtedy, gdy* $|z| = |w| = 0$ *albo* $|z| = |w| > 0$ *i* $\varphi = \psi + 2k\pi$ *dla pewnej liczby całkowitej* k*;*

(2) $zw = |z||w|e^{j(\varphi+\psi)}$*;*

(3) $z^n = |z|^n e^{jn\varphi}$ *(oraz* $(e^{j\varphi})^n = e^{jn\varphi}$*, co jest wzorem de Moivre'a);*

(4) $\overline{z} = |z|e^{-j\varphi}$*;*

(5) $\frac{1}{z} = \frac{1}{|z|}e^{-j\varphi}$ *(gdy* $z \neq 0$*);*

(6) $\frac{w}{z} = \frac{|w|}{|z|}e^{j(\psi-\varphi)}$ *(gdy* $z \neq 0$*);*

(7) *pierwiastkami* n*-tego stopnia z liczby* $z = |z|e^{j\varphi} \neq 0$ *są liczby*

$$z_k = \sqrt[n]{|z|}e^{j\frac{\varphi+2k\pi}{n}} \quad dla \quad k = 0, 1, \ldots, n-1,$$

a pierwiastkami n*-tego stopnia z jedności są liczby*

$$\varepsilon_k = e^{j\frac{2k\pi}{n}} \quad dla \quad k = 0, 1, \ldots, n-1. \ \square$$

Przykład 39. Korzystając z postaci wykładniczej liczby zespolonej, rozwiązać równanie $(\overline{z})^4 = -9|z^2|$.

Załóżmy, że $z = re^{j\varphi}$. Ponieważ $\overline{z} = re^{-j\varphi}$, $|z| = r$ i $-9 = 9e^{j\pi}$, więc mamy równoważności

$$\begin{aligned} (\overline{z})^4 = -9|z^2| &\iff r^4e^{-4j\varphi} = 9r^2e^{j\pi} \\ &\iff r = 0 \quad \text{albo} \quad r = 3 \ \text{ i } \ -4\varphi = \pi + 2k\pi \ \ (k \in Z) \\ &\iff r = 0 \quad \text{albo} \quad r = 3 \ \text{ i } \ \varphi = \tfrac{\pi}{4} + \tfrac{l\pi}{2} \ \ (l = 0, 1, 2, 3) \\ &\iff z = 0 \ \text{ albo } \ z = 3e^{j\left(\frac{\pi}{4}+\frac{l\pi}{2}\right)} \ \ (l = 0, 1, 2, 3). \end{aligned}$$

Zatem rozwiązaniem równania są liczby $z_1 = 0$, $z_2 = \frac{3}{\sqrt{2}}(1 + j)$, $z_3 = \frac{3}{\sqrt{2}}(-1 + j)$, $z_4 = \frac{3}{\sqrt{2}}(-1 - j)$ i $z_5 = \frac{3}{\sqrt{2}}(1 - j)$.

2.7. Ćwiczenia

1. Znaleźć część rzeczywistą i urojoną następujących liczb zespolonych:
   (a) $(1 + 2j)^6$; (b) $\dfrac{4 - 3j}{1 + j}$; (c) $\dfrac{1 + j\operatorname{tg}\alpha}{1 - j\operatorname{tg}\alpha}$;
2. Rozwiązać równania, w których niewiadome x i y są liczbami rzeczywistymi:
   (a) $(7 + 2j)x - (5 - 4j)y = -1 - j$;
   (b) $(1 + 2j)x + (3 - 5j)y = 1 - 3j$;
   (c) $(5 - 8j)x + (7 + 3j)y = 2 - j$;
   (d) $(2 + 3j)x^2 + (2 + j)x + (4 - 3j)y = 8 + 17j$.
3. Rozwiązać następujące równania, w których niewiadoma z jest liczbą zespoloną:
   (a) $(1 + j)z + 2j\overline{z} = 1 + 5j$;
   (b) $z\overline{z} + 2z = 19 + 4j$;
   (c) $|z| - z = 1 + 2j$;
   (d) $|z| + (1 + j)z = 4 + 7j$;
   (e) $|(2 + j)z| - (3 - j)z = -5j$;
   (f) $z\overline{z} + 2(\overline{z} - z) = 25 - 12j$.
4. Wyznaczyć i zaznaczyć w płaszczyźnie zbiór wszystkich z spełniających warunek:
   (a) $|z| + |z - 2j| = 2$;
   (b) $|z + j| + |z - j| = 4$.
5. Następujące liczby zapisać w postaci kartezjańskiej:
   (a) $\overline{(2 + 3j)(1 + j)}(7 - 3j)\overline{(7 - 3j)}$;
   (b) $30\left(\cos\pi + j\sin\pi\right)\left(\cos\frac{3\pi}{4} + j\sin\frac{3\pi}{4}\right)$;
   (c) $6\left(\cos\frac{\pi}{6} + j\sin\frac{\pi}{6}\right)\left(\cos\frac{5}{6}\pi + j\sin\frac{5}{6}\pi\right)$;
   (d) $3\left(\cos 42^\circ + j\sin 42^\circ\right)\left(\cos 168^\circ + j\sin 168^\circ\right)$;
   (e) $\frac{\sqrt{8}(\cos 147^\circ + j\sin 147^\circ)}{\sqrt{2}(\cos 57^\circ + j\sin 57^\circ)}j$;
   (f) $\left(\cos\frac{\pi}{6} + j\sin\frac{\pi}{6}\right)^{12}\left(\cos\frac{\pi}{12} + j\sin\frac{\pi}{12}\right)^{20}$.
6. Wyznaczyć postać trygonometryczną następujących liczb:
   (a) $(5 + 5j) \cdot \frac{3-j}{2+j}$; (b) $\frac{1+j\operatorname{tg}\alpha}{\cos\alpha+j\sin\alpha}$;
   (c) $\operatorname{tg}\alpha + j$; (d) $1 - \cos x + j\sin x$;
   (e) $\left(\frac{1+j\sqrt{3}}{1+j}\right)^n$; (f) $\frac{(-1+j\sqrt{3})^{4n}}{(1-j)^{8n}}$;

(g) $\left(\frac{1}{4} - j\frac{\sqrt{3}}{4}\right)(-\cos\alpha + j\sin\alpha)$;
(h)* $\sqrt{6} + \sqrt{2} + j(\sqrt{6} - \sqrt{2})$.

7. Obliczyć: (a) $\sum_{k=1}^{100} j^k$; (b) $\sum_{k=-15}^{27} j^k$; (c) $\sum_{k=0}^{11}(1+j)^k$.
8. W tym zadaniu przez $\mathrm{Arg}\,(z)$ oznaczamy ten argument liczby z, który należy do przedziału $(-\pi;\pi\rangle$ (albo do innego przedziału długości 2π). W płaszczyźnie $\mathcal{C}$ lub R^2 zaznaczyć zbiór tych liczb z, dla których:
(a) $\overline{z} = z^2$; (b) $|z-1+2j| = |z-4|$;
(c) $|z| < 1 - \mathrm{Re}\,(z)$; (d) $\mathrm{Arg}\,(z-2+j) = \frac{\pi}{4}$;
(e) $-\frac{\pi}{2} \leqslant \mathrm{Arg}\,(z^4) \leqslant \frac{3}{4}\pi$;
(f) $|z-4+5j| \geqslant 2$ i $-\frac{\pi}{6} \leqslant \mathrm{Arg}\,(z-1-3j) \leqslant \frac{\pi}{2}$;
(g) $-\frac{\pi}{4} \leqslant \mathrm{Arg}\,\big((1+j)(z-1+2j)\big) \leqslant \frac{3\pi}{4}$
i $|(3+4j)z-25| \leqslant 25$;
(h) $\mathrm{Arg}\,(z+1-j) = \mathrm{Arg}\,(z-3-2j)$;
(i) $\mathrm{Arg}\,(z+1-2j) = \mathrm{Arg}\,(z-3-j)+\pi$;
(j) $|\mathrm{Arg}\,(z-1) - \mathrm{Arg}\,(z+1)| = \frac{\pi}{2}$.
9. Wykazać, że dla liczby zespolonej z jest
$$\mathrm{Arg}\,(z-1) = \mathrm{Arg}\,(z+1) + \frac{\pi}{4}$$
wtedy i tylko wtedy, gdy $|z-j| = \sqrt{2}$ i $\mathrm{Im}\,(z) > 0$.
10. Wyznaczyć wszystkie rozwiązania równania:
(a) $z^8 = 1$; (b) $z^3 = j$;
(c) $z^3 = -2+2j$; (d) $z^4 = 8+8\sqrt{3}j$;
(e) $z^4 = -j$; (f) $z^4 = (1-j)^8$;
(g) $z^6 = (1+2j)^6$; (h) $z^3 = 4\sqrt{2}(1+j)$.
11. Wyznaczyć wszystkie zespolone rozwiązania następujących równań:
(a) $(x+j)^n + (x-j)^n = 0$;
(b) $(x+2)^n - (x-2)^n = 0$;
(c) $(x+3j)^n + j(x-3j)^n = 0$;
(d) $\left(\frac{1+xj}{1-xj}\right)^n = \frac{1+j\,\mathrm{tg}\,\alpha}{1-j\,\mathrm{tg}\,\alpha}$;
(e) $x^5+x^4+x^3+x^2+x+1 = 0$.
12. Wyznaczyć pierwiastki stopnia drugiego z następujących liczb zespolonych:
(a) $(1+2j)^2$; (b) $-5+12j$; (c) $-24-10j$.
13. Rozwiązać następujące równania:
(a) $x^2-3x+2,5 = 0$;
(b) $x^2+2jx-5 = 0$;
(c) $x^2+(2+2j)x+1+2j = 0$;
(d) $x^2-(3+7j)x-10+11j = 0$;
(e) $(3+j)x^2+(1-j)x-6j = 0$;
(f) $x^2+(8-5j)x-19+43j = 0$.
14. Znajdź cześć rzeczywistą i urojoną następujących liczb:
(a) e^{2+3j}; (c) $e^{\frac{\pi j}{6}}e^{\frac{3\pi j}{4}}$; (e) $\sin 2j$;
(b) $e^{\frac{5\pi j}{4}}$; (d) $2e^{\frac{\pi j}{4}}+3e^{\frac{\pi j}{6}}$; (f) $\cos(1+j)$.
15. Dana jest liczba $z = e^{a+bj}$, gdzie a i b są rzeczywiste. Wyznaczyć:
(a) $|z|$; (b) $\overline{z}$; (c) z^{-1};
(d) $\mathrm{Re}(z)$; (e) $\mathrm{Im}(z)$; (f) $\arg\,(z)$.
16. Korzystając z postaci wykładniczej liczby zespolonej, rozwiązać następujące równania:
(a) $z^3 = j\frac{|z^5|}{z\overline{z}}$; (b) $8z|\overline{z}| = (\overline{z})^5$;
(c) $z^6 = (\overline{z})^6$; (d) $z^4 = (1+j\sqrt{3})|\overline{z}|^2$.
17. (a) Oblicz $(5+4j)^3$ i stąd wyznaczyć postać kanoniczną liczby $(5-4j)^3$; (b) Zauważ, że $41 = 5^2+4^2 = (5+4j)(5-4j)$, a następnie przedstaw liczbę 41^3 jako sumę dwóch kwadratów.
18. Za pomocą sprzężenia wykazać, że $|z-w|^2+|z+w|^2 = 2|z|^2+2|w|^2$. Wywnioskować stąd, że suma kwadratów wszystkich boków równoległoboku jest równa sumie kwadratów jego przekątnych.
19. Biorąc pod uwagę potęgi liczby $2+j$, pokaż, że wektor $[3,4]$ dzieli kąt między wektorami $[2,1]$ i $[2,11]$ na dwie równe części.
20. Wykazać, że trzy różne punkty z_1, z_2 i z_3 są wierzchołkami trójkąta równobocznego, jeśli wiadomo, że $|z_1| = |z_2| = |z_3| = r > 0$ i $z_1+z_2+z_3 = 0$.
21. Niech $\omega_1 \neq 0$ będzie wierzchołkiem n-kąta foremnego o środku w punkcie $s_0 = 0$. Znaleźć pozostałe wierzchołki tego n-kąta.
22. Wyznaczyć część rzeczywistą i urojoną liczby $\frac{-1+j}{1+\sqrt{3}j}$. Przedstawić liczby $-1+j$ oraz $1+\sqrt{3}j$ w postaci trygonometrycznej, a następnie pokazać, że $\cos\frac{5\pi}{12} = \frac{\sqrt{3}-1}{2\sqrt{2}}$ i wyznaczyć $\sin\frac{5\pi}{12}$.
23. Pokazać, że jeśli $|z| = 1$, to $\overline{z} = \dfrac{1}{z}$.
24. Pokazać, że jeśli $z \neq -1$, to $\frac{z-1}{z+1}$ jest liczbą urojoną wtedy i tylko wtedy, gdy $|z| = 1$ i $\mathrm{Im}(z) \neq 0$.
25. Pokazać, że jeśli $z \neq -1$, to $\frac{z-1}{z+1}$ jest liczbą rzeczywistą wtedy i tylko wtedy, gdy $\mathrm{Im}(z) = 0$.
26. Korzystając ze wzoru de Moivre'a, przedstawić $\sin 4x$ i $\cos 4x$ za pomocą $\sin x$ i $\cos x$.
27. Korzystając ze wzoru de Moivre'a, uzasadnić, że
$$\cos 2nx = \sum_{k=0}^{n}\binom{2n}{2k}(-1)^k\cos^{2(n-k)}x\sin^{2k}x$$
i
$$\sin 2nx = \sum_{k=0}^{n-1}\binom{2n}{2k+1}(-1)^k\cos^{2(n-k)-1}x\sin^{2k+1}x.$$
28. Znaleźć wzory na sumy $\cos\alpha+\cos 2\alpha+\ldots+\cos n\alpha$ oraz $\sin\alpha+\sin 2\alpha+\ldots+\sin n\alpha$.
29. Wykazać, że jeśli z jest pierwiastkiem n-tego stopnia z jedynki, to także $\overline{z}$ jest pierwiastkiem n-tego stopnia z jedynki.
30. Niech z będzie pierwiastkiem n-tego stopnia z jedynki. Obliczyć wartości następujących sum:
(a) $1+z+z^2+\ldots+z^{n-1}$;
(b) $1+2z+3z^2+\ldots+nz^{n-1}$;
(c) $1+4z+9z^2+\ldots+n^2z^{n-1}$.
31. Obliczyć iloczyn wszystkich pierwiastków n-tego stopnia z jedynki.
32. Obliczyć sumę wszystkich pierwiastków n-tego stopnia z jedynki.
33. Obliczyć sumę k-tych potęg wszystkich pierwiastków n-tego stopnia z jedynki, gdzie k jest liczbą całkowitą.
34. Pokazać, że zbiór $H = \{z \in \mathcal{C} : |z| = 1\}$ jest grupą ze względu na mnożenie.
35. Wykazać, że zbiór $H = \{1, -1, j, -j\}$ z mnożeniem jest grupą.

36. Pokazać, że zbiór $\mathcal{P}_n$ wszystkich pierwiastków n-tego stopnia z jedynki jest grupą przemienną ze względu na mnożenie liczb zespolonych.

37. Udowodnić, że zbiór $\mathcal{P} = \bigcup_{n=1}^{\infty} \mathcal{P}_n$ jest multiplikatywną grupą abelową, gdzie $\mathcal{P}_n$ jest zbiorem wszystkich pierwiastków n-tego stopnia z jedynki dla $n \in N$.

38. Pokazać, że jeżeli liczba zespolona z_0 jest pierwiastkiem wielomianu $P(z)$ o współczynnikach rzeczywistych, to także liczba $\overline{z}_0$ jest pierwiastkiem wielomianu $P(z)$.

39. Wykazać, że podzbiór $\mathcal{C}_R = \{(a, 0)\colon a \in R\}$ zbioru liczb zespolonych $\mathcal{C}$ ze zwykłym dodawaniem $\oplus$ i zwykłym mnożeniem $\otimes$ liczb zespolonych (zob. (2.2) i (2.3)) jest ciałem.

40. Niech $(F_1, \oplus_1, \otimes_1)$ i $(F_2, \oplus_2, \otimes_2)$ będą ciałami. Mówimy, że ciało $(F_1, \oplus_1, \otimes_1)$ jest izomorficzne z ciałem $(F_2, \oplus_2, \otimes_2)$, gdy istnieje funkcja różnowartościowa $\varphi\colon F_1 \to F_2$ odwzorowująca zbiór F_1 na zbiór F_2 i taka, że $\varphi(x \oplus_1 y) = \varphi(x) \oplus_2 \varphi(y)$ oraz $\varphi(x \otimes_1 y) = \varphi(x) \otimes_2 \varphi(y)$ dla każdych elementów x, $y \in F_1$. Wykazać, że ciało liczb rzeczywistych $(R, +, \cdot)$, gdzie $+$ i $\cdot$ jest odpowiednio zwykłym dodawaniem i zwykłym mnożeniem liczb rzeczywistych, jest izomorficzne z ciałem $(\mathcal{C}_R, \oplus, \otimes)$ z poprzedniego zadania.

41. Wpisując TAK albo NIE, stwierdzić prawdziwość każdego z następujących zdań:

$\boxed{1}$ e^z jest liczbą dodatnią dla każdej liczby zespolonej z.

$\boxed{2}$ Nierówność $|\sin z| \leqslant 1$ jest prawdziwa dla każdej liczby zespolonej z.

$\boxed{3}$ Dla każdej liczby rzeczywistej φ i każdej liczby całkowitej n jest $(\cos\varphi - j\sin\varphi)^n = \cos n\varphi - j\sin n\varphi$.

$\boxed{4}$ Kwadrat każdej liczby zespolonej jest nieujemną liczbą rzeczywistą.

$\boxed{5}$ Każda niezerowa liczba zespolona ma dwa różne pierwiastki stopnia drugiego.

$\boxed{6}$ Jeśli $z + \overline{z} = 2z$, to z jest liczbą rzeczywistą.

$\boxed{7}$ Dla liczb rzeczywistych a i b jest $\sqrt{a^2 + b^2} = 2$, gdy $(a + bj)^3 = 8$.

$\boxed{8}$ Dla każdej liczby zespolonej z jest $e^{\overline{z}} = \overline{e^z}$.

$\boxed{9}$ Jeśli z jest liczbą zespoloną, a n jest liczbą naturalną, to częścią rzeczywistą liczby $z^n - (\overline{z})^n$ jest zero.

$\boxed{10}$ Jeśli z i w są liczbami zespolonymi, to $z\overline{w} + \overline{z}w$ jest liczbą rzeczywistą.

$\boxed{11}$ Jeśli z i w są liczbami zespolonymi, to $z\overline{w} - \overline{z}w$ jest liczbą czysto urojoną.

Rozdział 3

WIELOMIANY

3.1. Pierścień wielomianów

Definicja 3.1.1. Nieskończony ciąg

$$V = (a_0, a_1, a_2, \dots) \tag{3.1}$$

Wielomian

elementów a_i ciała K nazywamy *wielomianem* nad ciałem K (albo wielomianem o współczynnikach z ciała K), jeśli $a_i = 0$ dla prawie wszystkich $i \in N$.

Element a_i nazywamy i-tym *współczynnikiem* wielomianu (3.1). Czasami wygodnie jest i-ty współczynnik wielomianu V oznaczać przez $(V)_i$. Zbiór wszystkich wielomianów nad ciałem K oznaczamy przez $K[x]$. Wielomiany nad ciałem liczb rzeczywistych, czyli elementy zbioru $R[x]$, nazywamy *wielomianami rzeczywistymi*. Podobnie, elementy zbioru $\mathcal{C}[x]$ nazywamy *wielomianami zespolonymi*. Wielomian $(0, 0, \dots) \in K[x]$, w którym wszystkie współczynniki są równe zeru, nazywamy *wielomianem zerowym.*

Stopień wielomianu

Jeśli wielomian (3.1) nie jest zerowy, to jego *stopniem* nazywamy największą liczbę naturalną k taką, że $a_k \neq 0$. Przyjmujemy, że stopniem wielomianu zerowego jest $-\infty$. Stopień wielomianu V oznaczamy przez $\deg V$.

Równość wielomianów

Mówimy, że wielomiany $V = (a_0, a_1, \dots)$ i $W = (b_0, b_1, \dots)$ ze zbioru $K[x]$ są *równe*, piszemy $V = W$, wtedy i tylko wtedy, gdy $a_i = b_i$ dla $i = 0, 1, 2, \dots$. (Równoważnie, dla wielomianów V i W mamy $V = W$ wtedy i tylko wtedy, gdy $(V)_i = (W)_i$ dla każdej liczby naturalnej i.) W zbiorze $K[x]$ określamy dwa działania dwuargumentowe – dodawanie i mnożenie wielomianów. (Wyniki tych działań nazywamy odpowiednio *sumą* i *iloczynem* wielomianów.)

Definicja 3.1.2. Jeśli V i W są wielomianami nad ciałem K, powiedzmy $V = (a_0, a_1, \dots)$ i $W = (b_0, b_1, \dots)$, to ich *sumą* jest ciąg

$$V + W = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \dots). \tag{3.2}$$

Suma wielomianów

Ciąg (3.2) jest wielomianem nad ciałem K, bo $a_i + b_i \in K$ dla każdego indeksu $i \in N$ i jednocześnie $a_i + b_i = 0$ dla prawie wszystkich i, a na pewno $a_i + b_i = 0$ dla $i > \max\{\deg V, \deg W\}$. Stąd też wynika, że

$$\deg(W + V) \leqslant \max\{\deg W, \deg V\}.$$

Definicja 3.1.3. Jeśli $V = (a_0, a_1, \dots)$ i $W = (b_0, b_1, \dots)$ są wielomianami nad ciałem K, to ich *iloczynem* nazywamy ciąg

$$VW = (c_0, c_1, c_2, \dots), \tag{3.3}$$

Iloczyn wielomianów

w którym

$$c_0 = a_0 b_0, \;\; c_1 = a_0 b_1 + a_1 b_0, \;\; c_2 = a_0 b_2 + a_1 b_1 + a_2 b_0$$

i ogólnie

$$c_i = \sum_{j=0}^{i} a_j b_{i-j} = \sum_{k+l=i} a_k b_l, \tag{3.4}$$

gdzie ostatnia suma rozciąga się na wszystkie indeksy naturalne k i l takie, że $k + l = i$ dla $i \in N$.

Jest oczywiste, że $c_i = \sum_{j=0}^{i} a_j b_{i-j} \in K$ dla każdego $i \in N$. Zauważmy teraz, że jeśli V i W są odpowiednio wielomianami stopnia n i m, $\deg V = n$ i $\deg W = m$, to wtedy $a_n \neq 0$ i $a_j = 0$ dla $j > n$ oraz $b_m \neq 0$ i $b_j = 0$ dla $j > m$. Stąd łatwo wynika, że

$$\begin{aligned} c_{n+m} &= (a_0 b_{n+m} + \ldots + a_{n-1} b_{m+1}) + a_n b_m + (a_{n+1} b_{m-1} + \ldots + a_{n+m} b_0) \\ &= (a_0\, 0 + \ldots + a_{n-1}\, 0) + a_n b_m + (0\, b_{m-1} + \ldots + 0\, b_0) = a_n b_m \neq 0, \end{aligned} \tag{3.5}$$

a dla każdej liczby naturalnej $l > 0$ jest

$$\begin{aligned} c_{n+m+l} &= (a_0 b_{n+m+l} + \ldots + a_n b_{m+l}) + (a_{n+1} b_{m+l-1} + \ldots + a_{n+m+l} b_0) \\ &= (a_0\, 0 + \ldots + a_n\, 0) + (0\, b_{m+l-1} + \ldots + 0\, b_0) = 0. \end{aligned} \tag{3.6}$$

To dowodzi, że ciąg (3.3), którego współczynniki są określone wzorem (3.4), jest wielomianem nad ciałem K. Z (3.5) i (3.6) wynika jednocześnie, że jeśli V i W są niezerowymi wielomianami, to

$$\deg(VW) = \deg V + \deg W. \tag{3.7}$$

(Jeśli przyjmiemy, że dla każdej liczby naturalnej n jest $(-\infty) + n = -\infty = n + (-\infty)$ i jednocześnie $(-\infty) + (-\infty) = -\infty$, to równość (3.7) zachodzi także i wtedy, gdy V lub W jest wielomianem zerowym.)

Przykład 40. Jeśli $V = (2, 3, 0, 2, 0, \ldots)$ i $W = (4, 2, 3, 0, \ldots)$ są wielomianami nad ciałem Z_5, to ich sumą jest wielomian

$$\begin{aligned} V + W &= (2, 3, 0, 2, 0, \ldots) + (4, 2, 3, 0, \ldots) \\ &= (2 \oplus_5 4, 3 \oplus_5 2, 0 \oplus_5 3, 2 \oplus_5 0, 0 \oplus_5 0, \ldots) \\ &= (1, 0, 3, 2, 0, \ldots) \end{aligned}$$

i dla tego wielomianu jest $\deg(W + V) = 3 = \max\{\deg V, \deg W\}$. Iloczynem wielomianów V i W także jest wielomian i mamy

$$\begin{aligned} VW &= (2, 3, 0, 2, 0, \ldots)(4, 2, 3, 0, \ldots) \\ &= (c_0, c_1, c_2, \ldots) = (3, 1, 2, 2, 4, 1, 0, \ldots), \end{aligned}$$

bo wobec (3.4) jest

$$\begin{aligned} c_0 &= \sum_{j=0}^{0} a_j b_{0-j} = 2 \otimes_5 4 = 3, \\ c_1 &= \sum_{j=0}^{1} a_j b_{1-j} = (2 \otimes_5 2) \oplus_5 (3 \otimes_5 4) = 1, \\ c_2 &= \sum_{j=0}^{2} a_j b_{2-j} = (2 \otimes_5 3) \oplus_5 (3 \otimes_5 2) \oplus_5 (0 \otimes_5 4) = 2, \\ c_3 &= \sum_{j=0}^{3} a_j b_{3-j} = (2 \otimes_5 0) \oplus_5 (3 \otimes_5 3) \oplus_5 (0 \otimes_5 2) \oplus_5 (2 \otimes_5 4) = 2, \\ c_4 &= \sum_{j=0}^{4} a_j b_{4-j} = (2 \otimes_5 0) \oplus_5 (3 \otimes_5 0) \oplus_5 (0 \otimes_5 3) \oplus_5 (2 \otimes_5 2) \ominus_5 (0 \otimes_5 4) = 4, \\ c_5 &= \sum_{j=0}^{5} a_j b_{5-j} = (2 \cdot_5 0) \oplus_5 (3 \otimes_5 0) \oplus_5 (0 \otimes_5 0) \oplus_5 (2 \otimes_5 3) \oplus_5 (0 \otimes_5 2) \oplus_5 (0 \otimes_5 4) = 1, \end{aligned}$$

oraz

$$c_i = 0 \quad \text{dla } i > 5 = 3 + 2 = \deg V + \deg W = \deg(VW).$$

Twierdzenie 3.1.1. *Jeśli K jest ciałem, to zbiór wielomianów $K[x]$ z działaniami określonymi wzorami (3.2) – (3.4) jest pierścieniem całkowitym, czyli pierścieniem przemiennym z jednością i bez dzielników zera.*

Uwaga. Teza twierdzenia pozostaje prawdziwa, gdy założymy, że K jest pierścieniem całkowitym.

Dowód. Z przemienności i łączności dodawania w ciele K wynika przemienność oraz łączność dodawania w zbiorze $K[x]$, bo dla każdych wielomianów $V, W, U \in K[x]$ i każdej liczby naturalnej i jest

$$(V+W)_i = (V)_i + (W)_i = (W)_i + (V)_i = (W+V)_i$$

oraz

$$\begin{aligned}\big(V+(W+U)\big)_i &= (V)_i + (W+U)_i = (V)_i + \big((W)_i + (U)_i\big)\\ &= \big((V)_i + (W)_i\big) + (U)_i = (V+W)_i + (U)_i\\ &= \big((V+W)+U\big)_i.\end{aligned}$$

Łatwo zauważyć, że elementem neutralnym dodawania jest wielomian zerowy. Elementem przeciwnym do $V = (a_0, a_1, \ldots)$ jest $-V = (-a_0, -a_1, \ldots)$.

Z przemienności dodawania i mnożenia w ciele K wynika także przemienność mnożenia w zbiorze $K[x]$, bo dla każdych wielomianów V i W oraz liczby naturalnej i mamy

$$(VW)_i = \sum_{k+l=i} (V)_k (W)_l = \sum_{l+k=i} (W)_l (V)_k = (WV)_i.$$

Zauważmy, że dla każdych wielomianów $V, W, U \in K[x]$ oraz liczby naturalnej i mamy

$$\begin{aligned}\big(V(W+U)\big)_i &= \sum_{k+l=i} (V)_k (W+U)_l = \sum_{k+l=i} (V)_k \big((W)_l + (U)_l\big)\\ &= \sum_{k+l=i} (V)_k (W)_l + \sum_{k+l=i} (V)_k (U)_l\\ &= (VW)_i + (VU)_i = (VW+VU)_i.\end{aligned}$$

To dowodzi, że w $K[x]$ mnożenie jest rozdzielne względem dodawania, więc zachodzi równość $V(W+U) = VW + VU$. Mnożenie w zbiorze $K[x]$ jest także łączne, bo dla każdych wielomianów V, W i U oraz liczby naturalnej i jest

$$\begin{aligned}\big(V(WU)\big)_i &= \sum_{k+l=i} (V)_k (WU)_l = \sum_{k+l=i} (V)_k \sum_{s+t=l} (W)_s (U)_t\\ &= \sum_{k+s+t=i} (V)_k (W)_s (U)_t = \sum_{r+t=i} (VW)_r (U)_t = \big((VW)U\big)_i,\end{aligned}$$

co dowodzi, że $V(WU) = (VW)U$.

Wielomian $E = (1, 0, 0, \ldots)$ jest elementem neutralnym mnożenia, bo dla każdego wielomianu $V = (a_0, a_1, a_2, \ldots)$ wobec (3.3) i (3.4) jest

$$EV = (1, 0, 0, \ldots)(a_0, a_1, a_2, \ldots) = (a_0, a_1, a_2, \ldots) = V.$$

Z powyższych obserwacji wynika, że $K[x]$ jest pierścieniem przemiennym z jednością. Niech teraz V i W będą niezerowymi wielomianami i niech $\deg V = n$ oraz $\deg W = m$. Wtedy $(VW)_{n+m} = (V)_n (W)_m \neq 0$ (zob. (3.5)) i dlatego iloczyn VW jest niezerowy. Zatem $K[x]$ jest pierścieniem bez dzielników zera i to kończy dowód twierdzenia. □

Wniosek 3.1.1. *Jeśli V i W są wielomianami nad ciałem K i jeśli ich iloczyn VW jest wielomianem zerowym, to przynajmniej jeden z wielomianów V i W jest zerowy.*

Niech X i Y będą niepustymi zbiorami. Oznaczmy przez $\mathcal{F}(X,Y)$ zbiór wszystkich funkcji $f : X \to Y$. Weźmy teraz pod uwagę przekształcenie

$$T : K[x] \to \mathcal{F}(K,K),$$

które każdemu wielomianowi $V = (a_0, a_1, a_2, \ldots) \in K[x]$ przyporządkowuje funkcję $T_V \in \mathcal{F}(K, K)$ taką, że dla każdego $x \in K$ jest

$$T_V(x) = a_0 + a_1 x + a_2 x^2 + \ldots.$$

Funkcja wielomianowa

Funkcję $T_V(x)$ nazywamy *funkcją wielomianową* odpowiadającą wielomianowi V. Okazuje się, że to przyporządkowanie nie jest różnowartościowe. Przykładowo, wielomiany $V = (0, 1, 0, 0, 0, 3, 0, 0, \ldots)$ i $W = (0, 4, 0, 0, 0, \ldots)$ z pierścienia $Z_5[x]$ są różne, ale odpowiadające im funkcje wielomianowe $T_V(x) = x + 3x^5$ i $T_W(x) = 4x$ są równe, bo dla każdego x z ciała Z_5 jest $T_V(x) = x + 3x^5 = 4x = T_W(x)$. Okazuje się, że jeśli ciało K jest nieskończone, to wyżej wspomniane przekształcenie T jest już wzajemnie jednoznaczne. Dodatkowo, przekształcenie T jest izomorfizmem, bo dla każdych dwóch wielomianów $V = (a_0, a_1, a_2, \ldots)$ i $W = (b_0, b_1, b_2, \ldots)$ z pierścienia $K[x]$ mamy

$$\begin{aligned} T(V + W) &= T(a_0 + b_0, a_1 + b_1, \ldots) = (a_0 + b_0) + (a_1 + b_1)x + \ldots \\ &= (a_0 + a_1 x + \ldots) + (b_0 + b_1 x + \ldots) = T(V) + T(W). \end{aligned}$$

Dla iloczynu wielomianów $V = (a_0, a_1, a_2, \ldots)$ i $W = (b_0, b_1, b_2, \ldots)$, czyli dla wielomianu $VW = (c_0, c_1, \ldots)$, gdzie $c_i = \sum_{j=0}^{i} a_j b_{i-j}$, także mamy

$$\begin{aligned} T(VW) &= T(c_0, c_1, \ldots) = c_0 + c_1 x + c_2 x^2 + \ldots \\ &= (a_0 b_0) + (a_0 b_1 + a_1 b_0)x + (a_0 b_2 + a_1 b_1 + a_2 b_0)x^2 + \ldots \\ &= (a_0 + a_1 x + a_2 x^2 + \ldots)(b_0 + b_1 x + b_2 x^2 + \ldots) \\ &= T(V)T(W). \end{aligned}$$

Stąd i z twierdzenia 3.1.1 wynika, że zbiór funkcji wielomianowych

$$\{a_0 + a_1 x + a_2 x^2 + \ldots : (a_0, a_1, a_2, \ldots) \in K[x]\} \tag{3.8}$$

jest pierścieniem izomorficznym z pierścieniem $K[x]$. Z tego też względu także zbiór (3.8) oznaczamy symbolem $K[x]$, a elementy zbioru (3.8) nazywamy wielomianami.

3.2. Podzielność wielomianów

Podzielność wielomianów

$V(x)$ – dzielna

$W(x)$ – dzielnik

$Q(x)$ – iloraz

Definicja 3.2.1. Niech $V(x)$ i $W(x)$ będą wielomianami z pierścienia $K[x]$. Mówimy, że wielomian $V(x)$ jest *podzielny* przez wielomian $W(x)$, gdy istnieje wielomian $Q(x) \in K[x]$ taki, że

$$V(x) = W(x)Q(x). \tag{3.9}$$

Wielomian nierozkładalny

Jeżeli dla wielomianów $V(x)$, $W(x)$ i $Q(x)$ zachodzi równość (3.9), to mówimy, że $Q(x)$ jest *ilorazem* wielomianu $V(x)$ przez wielomian $W(x)$. W takim przypadku mówimy też, że wielomian $W(x)$ (jak i $Q(x)$) jest *dzielnikiem* albo *czynnikiem* wielomianu $V(x)$. Mówimy, że wielomian $V(x)$ jest *nierozkładalny* w pierścieniu $K[x]$, gdy nie istnieją wielomiany dodatniego stopnia $W(x)$, $Q(x) \in K[x]$ takie, że $V(x) = W(x)Q(x)$.

Przykład 41. Wielomian $W(x) = x + 4$ dzieli wielomian $V(x) = x^3 + 2x^2 - 4x + 16$ w pierścieniu $R[x]$, bo

$$W(x)Q(x) = (x + 4)(x^2 - 2x + 4) = x^3 + 2x^2 - 4x + 16 = V(x).$$

Twierdzenie 3.2.1 (Twierdzenie o dzieleniu wielomianów). *Jeżeli $V(x)$ i $W(x)$ są wielomianami z pierścienia $K[x]$ i $W(x) \neq 0$, to w pierścieniu $K[x]$ istnieją jednoznacznie wyznaczone wielomiany $Q(x)$ i $R(x)$ takie, że*

$$V(x) = W(x)Q(x) + R(x) \quad i \quad \deg R(x) < \deg W(x). \tag{3.10}$$

Uwagi. 1. Wielomiany $Q(x)$ i $R(x)$ spełniające warunek (3.10) nazywamy odpowiednio *ilorazem* i *resztą* z dzielenia wielomianu $V(x)$ przez $W(x)$. **2**. Ponieważ przyjęliśmy, że stopniem wielomianu zerowego jest $-\infty$, więc nierówność $\deg R(x) < \deg W(x)$ w (3.10) jest równoważna temu, że albo $R(x)$ jest wielomianem zerowym, albo wielomian $R(x)$ jest niezerowy i wtedy $0 \leqslant \deg R(x) < \deg W(x)$.

Dowód. Wykażemy najpierw istnienie wielomianów $Q(x)$ i $R(x)$. Niech $V(x)$ i $W(x)$ będą odpowiednio wielomianami stopnia n i m. Jeśli $n < m$, to wielomiany $Q(x) \equiv 0$ i $R(x) \equiv V(x)$ mają własność (3.10). Załóżmy teraz, że $0 \leqslant m \leqslant n$. W tym przypadku tezy dowodzimy indukcyjnie ze względu na n. Jeśli $n = 0$, to także $m = 0$ i $V(x)$ oraz $W(x)$ są niezerowymi elementami ciała K. Wtedy $Q(x) = V(x)/W(x)$ i $R(x) \equiv 0$ mają własność (3.10).

Niech teraz $V(x)$ będzie wielomianem stopnia $n > 0$ i załóżmy, że dowodzone stwierdzenie jest prawdziwe dla wszystkich wielomianów mniejszego stopnia. Jeśli $V(x) = a_nx^n + \ldots + a_1x + a_0$ i $W(x) = b_mx^m + \ldots + b_1x + b_0$, to bierzemy pod uwagę wielomian

$$\begin{aligned} U(x) &= V(x) - a_nb_m^{-1}x^{n-m}W(x) \\ &= (a_nx^n + \ldots + a_0) - a_nb_m^{-1}x^{n-m}(b_mx^m + b_{m-1}x^{m-1} \ldots + b_0) \\ &= 0x^n + (a_{n-1} - a_nb_m^{-1}b_{m-1})x^{n-1} + \ldots, \end{aligned} \tag{3.11}$$

którego stopień jest mniejszy od n. Z założenia indukcyjnego istnieją więc wielomiany $Q_1(x)$ i $R(x)$ takie, że

$$U(x) = W(x)Q_1(x) + R(x) \quad \text{i} \quad \deg R(x) < \deg W(x). \tag{3.12}$$

Z (3.11) i (3.12) otrzymujemy

$$\begin{aligned} V(x) &= a_nb_m^{-1}x^{n-m}W(x) + U(x) = a_nb_m^{-1}x^{n-m}W(x) + W(x)Q_1(x) + R(x) \\ &= W(x)\big(a_nb_m^{-1}x^{n-m} + Q_1(x)\big) + R(x) = W(x)Q(x) + R(x), \end{aligned}$$

gdzie $Q(x) = a_nb_m^{-1}x^{n-m} + Q_1(x)$ i $\deg R(x) < \deg W(x)$. To kończy dowód pierwszej części twierdzenia.

Dla dowodu jednoznaczności przypuśćmy, że istnieją wielomiany $Q_1(x)$, $Q_2(x)$, $R_1(x)$ i $R_2(x)$ takie, że

$$V(x) = W(x)Q_1(x) + R_1(x) \quad \text{i} \quad V(x) = W(x)Q_2(x) + R_2(x),$$

gdzie $\deg R_1(x) < \deg W(x)$ i $\deg R_2(x) < \deg W(x)$. Wtedy

$$W(x)\big(Q_1(x) - Q_2(x)\big) = R_2(x) - R_1(x) \tag{3.13}$$

i ponieważ $\deg\big(R_2(x) - R_1(x)\big) < \deg W(x)$, więc $Q_1(x) - Q_2(x)$ musi być wielomianem zerowym. Stąd $Q_1(x) = Q_2(x)$ i wtedy też wobec (3.13) jest $R_1(x) = R_2(x)$. To dowodzi, że iloraz $Q(x)$ i reszta $R(x)$, spełniające warunek (3.10), są wyznaczone jednoznacznie. $\square$

W praktyce wyznaczając iloraz $Q(x)$ i resztę $R(x)$ z dzielenia wielomianu $V(x)$ przez wielomian $W(x)$, gdzie $V(x) = a_nx^n + \ldots + a_1x + a_0$, $W(x) = b_mx^m + \ldots + b_1x + b_0$ i $\deg V(x) = n \geqslant m = \deg W(x)$, posługujemy się (dobrze znanym) algorytmem, który jest praktyczną realizacją dowodu poprzedniego twierdzenia. W algorytmie tym wielokrotnie tworzymy różnicę (3.11) i powtarzamy ten proces do chwili otrzymania różnicy, której stopień jest mniejszy od $m = \deg W(x)$. Pierwszy krok wspomnianego algorytmu przedstawia następujący schemat:

$$\begin{array}{rcl} \underline{a_nb_m^{-1}x^{n-m}\phantom{+ a_{n-1}x^{n-1} + \ldots + a_1x + a_0}} & & \\ a_nx^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0 & : & b_mx^m + \ldots + b_0 \\ \underline{a_nb_m^{-1}x^{n-m} \cdot W(x)} & & \\ V(x) - a_nb_m^{-1}x^{n-m} \cdot W(x) & & \end{array}$$

Przykład 42. Obliczyć iloraz $Q(x)$ i resztę $R(x)$ z dzielenia wielomianu $V(x) = 4x^3 + x^2 - x + 2$ przez wielomian $W(x) = x^2 - x - 2$ w pierścieniu $R[x]$.

Zgodnie z powyższym schematem, dzieląc wielomian $V(x)$ przez $W(x)$, otrzymujemy

$$
\begin{array}{rrrrrrrl}
\multicolumn{7}{l}{4x + 5} & \\
\hline
4x^3 & + x^2 & - & x & + & 2 & & : x^2 - x - 2 \\
4x^3 & - 4x^2 & - & 8x & & & & \\
\hline
& 5x^2 & + & 7x & + & 2 & & \\
& 5x^2 & - & 5x & - & 10 & & \\
\cline{2-6}
& & & 12x & + & 12. & &
\end{array}
$$

Stąd $Q(x) = 4x + 5$ i $R(x) = 12x + 12$.

Przykład 43. Wyznaczyć iloraz $Q(x)$ i resztę $R(x)$ z dzielenia wielomianu $V(x) = 4x^4 + 2x^3 + 6x^2 + 4x + 3$ przez wielomian $W(x) = 3x^2 + 2$ w pierścieniu $Z_7[x]$.

Ponieważ pierwszymi współczynnikami wielomianów $V(x)$ i $W(x)$ są $a_n = a_4 = 4$ i $b_m = b_2 = 3$, więc pierwszym składnikiem ilorazu tych wielomianów jest $a_n b_m^{-1} x^{n-m} = 4 \cdot 3^{-1} x^2 = 6x^2$ (bo w Z_7 jest $3^{-1} = 5$ i $4 \cdot 5 = 6$) i cały proces dzielenia wygląda następująco:

$$
\begin{array}{rrrrrl}
\multicolumn{5}{l}{6x^2 + 3x + 5} & \\
\hline
4x^4 & + 2x^3 & + 6x^2 & + 4x & + 3 & : 3x^2 + 2 \\
4x^4 & & + 5x^2 & & & \\
\hline
& 2x^3 & + x^2 & + 4x & + 3 & \\
& 2x^3 & & + 6x & & \\
\cline{2-5}
& & x^2 & + 5x & + 3 & \\
& & x^2 & & + 3 & \\
\cline{3-5}
& & & 5x. & &
\end{array}
$$

Zatem $Q(x) = 6x^2 + 3x + 5$ i $R(x) = 5x$.

Wniosek 3.2.1 (Twierdzenie o reszcie). *Jeśli $V(x) \in K[x]$ i $x_0 \in K$, to resztą z dzielenia wielomianu $V(x)$ przez $x - x_0$ jest $V(x_0)$, czyli istnieje wielomian $Q(x) \in K[x]$ taki, że*

$$V(x) = (x - x_0)Q(x) + V(x_0). \tag{3.14}$$

Dowód. Wobec twierdzenia 3.2.1 istnieją wielomiany $Q(x)$ i $R(x)$ takie, że

$$V(x) = (x - x_0)Q(x) + R(x) \quad \text{i} \quad \deg R(x) < \deg(x - x_0). \tag{3.15}$$

Ponieważ $\deg(x - x_0) = 1$, więc $R(x)$ jest elementem ciała K i dlatego w szczególności jest $R(x) = R(x_0)$. Z równości (3.15) dla $x = x_0$ mamy $V(x_0) = (x_0 - x_0)Q(x_0) + R(x_0) = R(x_0)$, więc także $R(x) = V(x_0)$ i dlatego mamy (3.14). □

3.3. Schemat Hornera

$(3x + 2)x + 7$ – postać zagnieżdżona wielomianu $3x^2 + 2x + 7$

Wielomian $V(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$ możemy zapisać w tzw. postaci zagnieżdżonej,

$$V(x) = \big(\ldots\big((a_n x + a_{n-1})x + a_{n-2}\big)x + \ldots + a_1\big)x + a_0.$$

Takie ustawienie nawiasów sugeruje następującą kolejność obliczania wartości wielomianu $V(x)$ dla $x = x_0$:

— przyjmujemy, że

$$y_0 = a_n \tag{3.16}$$

jest początkową wartością wielomianu;

— następnie, aż do wyczerpania wszystkich współczynników wielomianu, bieżącą wartość y_i wielomianu mnożymy przez x_0 i powiększamy o współczynnik a_{n-i-1}, otrzymując

$$y_{i+1} = x_0 y_i + a_{n-i-1}. \tag{3.17}$$

Ostatnia tak wyliczona liczba $y_n = x_0 y_{n-1} + a_0$ jest wartością $V(x_0)$ wielomianu $V(x)$. Ten sposób obliczania wartości wielomianu zwykle nazywa się schematem (lub algorytmem) Hornera i realizuje za pomocą dwuwierszowej tabeli

Schemat Hornera

	a_n	a_{n-1}	$\cdots$	a_1	a_0
x_0	y_0	y_1	$\cdots$	y_{n-1}	y_n

W górnym wierszu tej tabeli stoją kolejne współczynniki wielomianu $V(x)$. Natomiast w dolnym wierszu mamy x_0, $y_0 = a_n$ oraz kolejno wyznaczone liczby $y_1, \ldots, y_n$. Te ostatnie wyznaczamy ze wzoru (3.17). Warto zauważyć, że w ogólnym przypadku obliczenie wartości wielomianu stopnia n za pomocą schematu Hornera wymaga wykonania n mnożeń i n dodawań. A. Borodin udowodnił w 1971 roku, że jest to najszybszy sposób obliczania wartości wielomianu.

Przykład 44. Za pomocą schematu Hornera obliczyć wartość wielomianu $V(x) = 5x^3 - 3x^2 + 10x + 7$ dla $x_0 = 2$.

Tworzymy tabelę

	5	-3	10	7
2	5			

W górnym wierszu wpisaliśmy kolejne współczynniki wielomianu $V(x)$. W dolnym wpisaliśmy $x_0 = 2$ i $y_0 = a_3 = 5$. W wolne miejsca kolejno wpisujemy liczby $y_1 = x_0 y_0 + a_2 = 2 \cdot 5 + (-3) = 7$, $y_2 = x_0 y_1 + a_1 = 2 \cdot 7 + 10 = 24$ i $y_3 = x_0 y_2 + a_0 = 2 \cdot 24 + 7 = 55$ otrzymując tabelę

	5	-3	10	7
2	5	7	24	55

z której wnioskujemy, że $V(2) = 55$.

Schemat Hornera jest niezwykle przydatny przy wyznaczaniu ilorazu $Q(x)$ i reszty $R(x)$ z dzielenia wielomianu $V(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ przez dwumian $x - x_0$. To ostatnie jest oczywiste, bo wobec wniosku 3.2.1 jest $R(x) = V(x_0)$, a wartość $V(x_0)$ można wyznaczyć za pomocą schematu Hornera. Dla uzasadnienia pierwszej części przyjmijmy, że ilorazem z dzielenia wielomianu $V(x)$ przez dwumian $x - x_0$ jest wielomian $Q(x) = b_0 x^{n-1} + b_1 x^{n-2} + \ldots + b_{n-2}x + b_{n-1}$. Wtedy

$$\begin{aligned} V(x) &= (x - x_0)Q(x) + V(x_0) \\ &= (x - x_0)(b_0 x^{n-1} + b_1 x^{n-2} + \ldots + b_{n-2}x + b_{n-1}) + V(x_0) \\ &= b_0 x^n + (b_1 - x_0 b_0)x^{n-1} + \ldots + (b_{n-1} - x_0 b_{n-2})x - x_0 b_{n-1} + V(x_0) \end{aligned}$$

i dlatego $a_n = b_0$, $a_{n-1} = b_1 - x_0 b_0$, $\ldots$, $a_1 = b_{n-1} - x_0 b_{n-2}$ oraz $a_0 = V(x_0) - x_0 b_{n-1}$. Stąd zaś wynika, że

$$b_0 = a_n,\ b_1 = x_0 b_0 + a_{n-1},\ \ldots,\ b_{n-1} = x_0 b_{n-2} + a_1 \text{ oraz } V(x_0) = x_0 b_{n-1} + a_0,$$

więc (wobec (3.16) i (3.17)) współczynniki wielomianu $Q(x)$ (oraz reszta $R(x)$) są dokładnie tymi, które otrzymujemy w schemacie Hornera dla wielomianu $V(x)$ i $x = x_0$.

Przykład 45. Wyznaczyć iloraz i resztę z dzielenia wielomianu $V(x) = x^5 + 2x^4 - x + 7$ przez $x + 2$.

Stosując schemat Hornera dla wielomianu $V(x)$ i $x_0 = -2$ otrzymujemy tabelę

	1	2	0	0	-1	7
-2	1	0	0	0	-1	9

i z niej wynika, że ilorazem oraz resztą z dzielenia wielomianu $V(x)$ przez $x + 2$ są odpowiednio $Q(x) = x^4 - 1$ i $R(x) = 9 = V(-2)$.

3.4. Pierwiastki wielomianów

Pierwiastek wielomianu

Definicja 3.4.1. Niech $V(x)$ będzie wielomianem nad ciałem K. Element x_0 ciała K nazywamy *pierwiastkiem* (albo *zerem*) wielomianu $V(x)$, gdy $V(x_0) = 0$.

Przykład 46. Wyznaczyć pierwiastki wielomianu $V(x) = 2x^2 + 3$ w ciele Z_5.

Ponieważ dla elementów ciała $Z_5 = \{0, 1, 2, 3, 4\}$ jest

$$\begin{aligned} &V(0) = 2 \cdot 0^2 + 3 = 3, \quad V(1) = 2 \cdot 1^2 + 3 = 0, \quad V(2) = 2 \cdot 2^2 + 3 = 1, \\ &V(3) = 2 \cdot 3^2 + 3 = 1, \quad V(4) = 2 \cdot 4^2 + 3 = 0, \end{aligned}$$

więc w ciele Z_5 pierwiastkami wielomianu $V(x)$ są $x_1 = 1$ i $x_2 = 4$.

Z definicji 3.2.1 i 3.4.1 oraz z wniosku 3.2.1 natychmiast otrzymujemy następującą ważną własność pierwiastka wielomianu.

Twierdzenie Bézout

Twierdzenie 3.4.1 (Bézout). *Element x_0 ciała K jest pierwiastkiem wielomianu $V(x) \in K[x]$ wtedy i tylko wtedy, gdy wielomian $V(x)$ jest podzielny przez dwumian $x - x_0$.*

Dowód. Wobec wniosku 3.2.1 istnieje jednoznacznie wyznaczony wielomian $Q(x)$ taki, że

$$V(x) = (x - x_0)Q(x) + V(x_0).$$

Stąd widać, że $V(x)$ jest podzielny przez $x - x_0$ wtedy i tylko wtedy, gdy $V(x_0) = 0$, czyli wtedy i tylko wtedy, gdy x_0 jest pierwiastkiem wielomianu $V(x)$. □

Definicja 3.4.2. Element x_0 ciała K jest *k-krotnym pierwiastkiem* wielomianu $V(x) \in K[x]$, gdy istnieje wielomian $Q(x) \in K[x]$ taki, że

Pierwiastek k-krotny

$$V(x) = (x - x_0)^k Q(x) \quad \text{i} \quad Q(x_0) \neq 0.$$

Pierwiastek wielokrotny

Mówimy też, że x_0 jest pierwiastkiem wielokrotnym wielomianu $V(x)$, gdy x_0 jest k-krotnym pierwiastkiem wielomianu $V(x)$ dla pewnej liczby naturalnej $k \geqslant 2$.

Przykład 47. Liczba $x_0 = 2$ jest 2-krotnym pierwiastkiem wielomianu $V(x) = x^3 - 3x^2 + 4$, bo mamy $V(x) = (x - 2)^2(x + 1)$ i dla wielomianu $Q(x) = x + 1$ jest $Q(2) = 3 \neq 0$.

Krotność pierwiastka wielomianu o współczynnikach zespolonych możemy wyznaczyć korzystając z następującego twierdzenia.

Twierdzenie 3.4.2. *Liczba zespolona x_0 jest k-krotnym pierwiastkiem wielomianu $V(x) \in \mathcal{C}[x]$ wtedy i tylko wtedy, gdy*

$$V(x_0) = V'(x_0) = \ldots = V^{(k-1)}(x_0) = 0 \quad \textit{oraz} \quad V^{(k)}(x_0) \neq 0.$$

Dowód. Załóżmy, że x_0 jest k-krotnym pierwiastkiem wielomianu $V(x)$. Wtedy $V(x) = (x-x_0)^k Q(x)$ dla pewnego wielomianu $Q(x) \in \mathcal{C}[x]$ i $Q(x_0) \neq 0$. Ze wzoru Leibniza[1] na pochodną iloczynu funkcji mamy

$$\begin{aligned} V^{(m)}(x) &= \sum_{j=0}^{m} \binom{m}{j} \left((x-x_0)^k\right)^{(j)} Q^{(m-j)}(x) \\ &= \sum_{j=0}^{m} \binom{m}{j} k(k-1)\dots(k-j+1)(x-x_0)^{k-j} Q^{(m-j)}(x). \end{aligned}$$

Łatwo teraz zauważyć, że dla $m < k$ jest $V^{(m)}(x_0) = 0$. Natomiast dla $m = k$ mamy

$$V^{(k)}(x_0) = k!\, Q^{(0)}(x_0) = k!\, Q(x_0) \neq 0.$$

Załóżmy teraz, że $V(x)$ jest wielomianem stopnia n i dla pewnego k ($k \leqslant n$) jest $V(x_0) = V'(x_0) = \ldots = V^{(k-1)}(x_0) = 0$ i $V^{(k)}(x_0) \neq 0$. Wtedy ze wzoru Taylora[2] mamy

$$\begin{aligned} V(x) &= \sum_{i=0}^{n} \frac{V^{(i)}(x_0)}{i!}(x-x_0)^i = \sum_{i=k}^{n} \frac{V^{(i)}(x_0)}{i!}(x-x_0)^i \\ &= (x-x_0)^k \sum_{i=k}^{n} \frac{V^{(i)}(x_0)}{i!}(x-x_0)^{i-k} = (x-x_0)^k Q(x), \end{aligned}$$

gdzie

$$Q(x) = \frac{V^{(k)}(x_0)}{k!} + \frac{V^{(k+1)}(x_0)}{(k+1)!}(x-x_0) + \ldots + \frac{V^{(n)}(x_0)}{n!}(x-x_0)^{n-k}$$

i widać stąd, że $Q(x_0) = \frac{V^{(k)}(x_0)}{k!} \neq 0$. To dowodzi, że x_0 jest k-krotnym pierwiastkiem wielomianu $V(x)$. □

Przykład 48. Wyznaczyć krotność pierwiastka $x_0 = 2j$ wielomianu

$$V(x) = x^3 - 2jx^2 + 4x - 8j \in \mathcal{C}[x].$$

Ponieważ

$$\begin{array}{lcll} V(x) &=& x^3 - 2jx^2 + 4x - 8j & \text{i} \quad V(2j) = -8j + 8j + 8j - 8j = 0, \\ V'(x) &=& 3x^2 - 4jx + 4 & \text{i} \quad V'(2j) = -12 + 8 + 4 = 0, \\ V''(x) &=& 6x - 4j & \text{i} \quad V''(2j) = 12j - 4j \neq 0, \end{array}$$

więc (wobec twierdzenia 3.4.2) liczba $x_0 = 2j$ jest 2-krotnym pierwiastkiem wielomianu $V(x)$.

Przykład 49. Wykazać, że wielomian rzeczywisty

$$V(x) = x^5 + x^4 + 1$$

nie ma pierwiastka wielokrotnego.

Jeśliby wielomian $V(x)$ miał wielokrotny pierwiastek, to wobec twierdzenia 3.4.2 byłby to jednocześnie pierwiastek wielomianu $V'(x)$. Ponieważ pierwiastkami wielomianu $V'(x) = 5x^4 + 4x^3$ są tylko liczby $x_0 = 0$ i $x_0' = -4/5$ i ponieważ żadna z tych liczb nie

[1] Wzór Leibniza: $(f(x)g(x))^{(m)} = \sum_{j=0}^{m} \binom{m}{j} f^{(j)}(x) g^{(m-j)}(x)$

[2] Wzór Taylora: Jeśli $\deg V(x) = n$, to $V(x) = \sum_{i=0}^{n} \frac{V^{(i)}(x_0)}{i!}(x-x_0)^i$

jest pierwiastkiem wielomianu $V(x)$ (bo $V(x_0) = V(0) = 1 \neq 0$ i $V(x_0') = V(-4/5) = (-4/5)^5 + (-4/5)^4 + 1 \neq 0$), więc żadna z nich nie jest pierwiastkiem wielokrotnym wielomianu $V(x)$.

Twierdzenie 3.4.3. *Wielomian stopnia $n \geqslant 0$ nad ciałem K ma co najwyżej n pierwiastków w ciele K.*

Dowód indukcyjny ze względu na n. Jeśli $V(x) \in K[x]$ i $\deg V(x) = 0$, to $V(x) \equiv a_0 \neq 0$ i $V(x)$ nie ma pierwiastków. Załóżmy prawdziwość tezy dla wielomianów stopnia $n-1$, $n \geqslant 1$. Niech teraz $V(x)$ będzie wielomianem stopnia n. Jeśli $V(x)$ nie ma pierwiastków w ciele K, to oczywiście ma on co najwyżej n pierwiastków w ciele K. Jeśli $x_0 \in K$ jest pierwiastkiem wielomianu $V(x)$, to $V(x) = (x - x_0)Q(x)$ dla pewnego $Q(x) \in K[x]$ i $\deg Q(x) = n - 1$. Z założenia indukcyjnego wielomian $Q(x)$ ma co najwyżej $n - 1$ pierwiastków w ciele K, więc $V(x)$ ma co najwyżej n pierwiastków w ciele K – są nimi x_0 i pierwiastki wielomianu $Q(x)$. $\square$

Wniosek 3.4.1. *Wielomiany $V(x)$ i $W(x)$ nad nieskończonym ciałem K są równe wtedy i tylko wtedy, gdy mają one identyczne współczynniki, czyli gdy $(V)_i = (W)_i$ dla każdego $i \in N$.*

Dowód. Jest oczywiste, że jeśli $(V)_i = (W)_i$ dla $i \in N$, to dla każdego $x \in K$ jest

$$V(x) = \sum_{i=0}^{\infty} (V)_i x^i = \sum_{i=0}^{\infty} (W)_i x^i = W(x).$$

Załóżmy teraz, że $V(x) = W(x)$ dla każdego $x \in K$. Wtedy każdy element x_0 nieskończonego ciała K jest pierwiastkiem wielomianu $T(x) = V(x) - W(x)$. To zaś jest możliwe pod warunkiem, że $T(x)$ jest wielomianem zerowym (bo wobec twierdzenia 3.4.3 każdy niezerowy wielomian ma skończoną liczbę pierwiastków). Stąd $(V)_i - (W)_i = (T)_i = 0$ i dlatego $(V)_i = (W)_i$ dla każdego $i \in N$. $\square$

Wniosek 3.4.2. *Załóżmy, że $V(x)$ i $W(x)$ są wielomianami stopnia co najwyżej n i $V(x), W(x) \in K[x]$. Wielomiany $V(x)$ i $W(x)$ są równe wtedy i tylko wtedy, gdy $V(x_i) = W(x_i)$ dla $n+1$ różnych elementów $x_1, x_2, \ldots, x_{n+1}$ z ciała K.*

Dowód. Załóżmy, że $V(x_i) = W(x_i)$ dla różnych elementów $x_1, x_2, \ldots, x_{n+1}$ z ciała K. Wtedy elementy $x_1, x_2, \ldots, x_{n+1}$ są różnymi pierwiastkami wielomianu $T(x) = V(x) - W(x)$. Ponieważ $T(x)$ jest wielomianem stopnia co najwyżej n i ponieważ ma on więcej niż n pierwiastków, więc wobec twierdzenia 3.4.3 wielomian $T(x)$ musi być wielomianem zerowym. Stąd zaś wynika równość wielomianów $V(x)$ i $W(x)$.

Przeciwna implikacja jest oczywista. $\square$

Następujące twierdzenie – zwane zasadniczym twierdzeniem algebry – powiada, że każdy wielomian (dodatniego stopnia) nad ciałem liczb zespolonych ma pierwiastek zespolony.

Zasadnicze tw. algebry

Twierdzenie 3.4.4 (Zasadnicze twierdzenie algebry). *Jeśli $V(x)$ jest wielomianem nad ciałem liczb zespolonych i $\deg V(x) > 0$, to $V(x)$ ma pierwiastek w ciele $\mathcal{C}$.* $\square$

Znanych jest wiele dowodów tego twierdzenia. Pierwszy podał C. F. Gauss w swojej pracy doktorskiej z roku 1799 i noszącej tytuł "Dowód twierdzenia, że każdą całkowitą funkcję wymierną jednej zmiennej można rozłożyć na rzeczywiste czynniki pierwszego lub drugiego stopnia". (Gauss podał też trzy inne dowody tego twierdzenia.) Ponieważ w każdym z nich korzysta się z dość zaawansowanych metod analizy matematycznej lub algebry, nie przedstawimy tu dowodu tego twierdzenia. Dowody te można znaleźć m.in. w [5], [12], [13] i [14]. Następujący wniosek o rozkładzie wielomianu na iloczyn dwumianów jest prostą konsekwencją twierdzenia 3.4.4.

Wniosek 3.4.3. *Jeśli $V(x)$ jest wielomianem nad ciałem liczb zespolonych i $\deg V(x) = n > 0$, to $V(x)$ ma dokładnie n (niekoniecznie różnych) pierwiastków $x_1, x_2, \ldots, x_n$ w ciele $\mathcal{C}$ i może on być przedstawiony w postaci iloczynu*

Rozkład wielomianu na czynniki liniowe – iloczyn dwumianów

$$V(x) = a(x - x_1)(x - x_2)\ldots(x - x_n), \tag{3.18}$$

gdzie a jest liczbą różną od zera.

Dowód indukcyjny ze względu na n. Teza jest oczywista dla $n = 1$. Załóżmy teraz prawdziwość tezy dla wielomianów stopnia $n - 1$, gdzie $n \geqslant 2$, i weźmy pod uwagę wielomian $V(x) \in \mathcal{C}[x]$ stopnia n. Wobec twierdzenia 3.4.4 istnieje liczba $x_1 \in \mathcal{C}$ taka, że $V(x_1) = 0$. Zatem, wobec twierdzenia 3.4.1, istnieje wielomian $Q(x) \in \mathcal{C}[x]$ stopnia $n - 1$ taki, że

$$V(x) = (x - x_1)Q(x).$$

Z założenia indukcyjnego dla wielomianu $Q(x)$ istnieją liczby zespolone $x_2, x_3, \ldots x_n$ i a takie, że

$$Q(x) = a(x - x_2)(x - x_3)\ldots(x - x_n).$$

Stąd i z twierdzenia 3.4.3 wynika, że $x_1, x_2, \ldots, x_n$ są wszystkimi pierwiastkami wielomianu $V(x)$ i mamy

$$V(x) = (x - x_1)Q(x) = a(x - x_1)(x - x_2)\ldots(x - x_n). \quad \square$$

Wniosek 3.4.4. *Niech $V(x)$ będzie wielomianem stopnia n nad ciałem liczb zespolonych. Jeśli liczby zespolone $x_1, x_2, \ldots, x_m$ są pierwiastkami wielomianu $V(x)$ o krotnościach odpowiednio $k_1, k_2, \ldots, k_m$ i $k_1 + k_2 + \ldots + k_m = n$, to*

$$V(x) = a(x - x_1)^{k_1}(x - x_2)^{k_2}\ldots(x - x_m)^{k_m} \tag{3.19}$$

dla pewnej niezerowej liczby a. $\square$

Teoretycznie każdy wielomian o współczynnikach zespolonych można przedstawić w postaci (3.18) i (3.19). Praktycznie są z tym problemy, bo dla wielomianów stopnia co najmniej piątego nie znamy (i – jak to wynika z teorii Abela – nigdy nie poznamy) możliwości przedstawienia pierwiastków wielomianu za pomocą pierwiastkowania i działań arytmetycznych na współczynnikach wielomianu. W praktyce, przedstawiając wielomian w postaci iloczynu dwumianów, korzystamy (tam gdzie jest to możliwe i/lub celowe) ze wzorów na pierwiastki wielomianów niskiego stopnia, ze wzorów na pierwiastki z liczb zespolonych oraz z przekształceń algebraicznych.

Przykład 50. Podane wielomiany zespolone przedstawić w postaci iloczynu dwumianów:

(a) $V(x) = x^3 + 8j$;

(b) $V(x) = x^3 - (3 + j)x^2 + (1 + 8j)x + 1 - 7j$;

(c) $V(x) = x^4 + jx^2 + 6$.

(a) Pierwiastkami wielomianu $V(x) = x^3 + 8j$ są pierwiastki trzeciego stopnia z liczby $-8j$, czyli liczby $2j$, $-\sqrt{3} - j$ i $\sqrt{3} - j$. Zatem

$$V(x) = x^3 + 8j = (x - 2j)(x + \sqrt{3} + j)(x - \sqrt{3} + j).$$

(b) Łatwo zauważyć, że $x = 1$ jest pierwiastkiem wielomianu $V(x)$. Ponieważ ilorazem z dzielenia wielomianu $V(x)$ przez dwumian $x - 1$ jest trójmian $Q(x) = x^2 - (2+j)x - 1 + 7j$ i ponieważ pierwiastkami tego trójmianu są liczby $3 - j$ i $-1 + 2j$ (zob. przykład 36), więc

$$V(x) = (x - 1)(x - 3 + j)(x + 1 - 2j).$$

(c) Tym razem wielomian $V(x)$ przedstawimy w postaci iloczynowej bez uprzedniego wyznaczania jego pierwiastków. Za pomocą prostych przekształceń otrzymujemy

$$\begin{aligned} V(x) &= x^4 + jx^2 + 6 = \left(x^4 + jx^2 + \tfrac{j^2}{4}\right) - \tfrac{j^2}{4} + 6 \\ &= \left(x^2 + \tfrac{j}{2}\right)^2 + \tfrac{25}{4} = \left(x^2 + \tfrac{j}{2}\right)^2 - \left(\tfrac{5}{2}j\right)^2 \\ &= \left(x^2 + \tfrac{j}{2} - \tfrac{5}{2}j\right)\left(x^2 + \tfrac{j}{2} + \tfrac{5}{2}j\right) = (x^2 - 2j)(x^2 + 3j) \\ &= (x - 1 - j)(x + 1 + j)\left(x + \tfrac{\sqrt{6}}{2} - \tfrac{\sqrt{6}}{2}j\right)\left(x - \tfrac{\sqrt{6}}{2} + \tfrac{\sqrt{6}}{2}j\right). \end{aligned}$$

Każdy wielomian stopnia n o współczynnikach rzeczywistych ma wobec wniosku 3.4.3 dokładnie n pierwiastków zespolonych (uwzględniając pierwiastki wielokrotne) i oczywiście nie muszą to być liczby rzeczywiste. Przykładowo, wielomian $x^2 + 1 \in R[x]$ ma dwa pierwiastki zespolone, j i $-j$, ale nie ma on żadnego pierwiastka rzeczywistego. Wielomian $x^3 + 1 \in R[x]$ ma trzy pierwiastki zespolone – są nimi liczby -1, $\frac{1}{2} + j\frac{\sqrt{3}}{2}$ i $\frac{1}{2} - j\frac{\sqrt{3}}{2}$ (które są pierwiastkami stopnia 3 z liczby -1) i tylko jeden z tych pierwiastków jest liczbą rzeczywistą. Warto zwrócić uwagę na fakt, że w obu przykładach liczby sprzężone z_0 i $\overline{z_0}$ jednocześnie są pierwiastkami tego samego wielomianu. Następne twierdzenie pokazuje, że to "chodzenie parami" pierwiastków sprzężonych jest własnością charakterystyczną pierwiastków wielomianów o współczynnikach rzeczywistych (zob. także zadanie 23). W dowodzie tego twierdzenia skorzystamy z następującego oznaczenia: jeśli $V(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \in \mathcal{C}[x]$, to przez $\overline{V}(x)$ oznaczamy wielomian sprzężony z wielomianem $V(x)$, wielomian, którego współczynniki są liczbami sprzężonymi do odpowiednich współczynników wielomianu $V(x)$, czyli

$$\overline{V}(x) = \overline{a_n}x^n + \overline{a_{n-1}}x^{n-1} + \ldots + \overline{a_1}x + \overline{a_0}.$$

Twierdzenie 3.4.5. *Niech $V(x)$ będzie wielomianem o współczynnikach rzeczywistych. Jeśli liczba zespolona z_0 jest k-krotnym pierwiastkiem wielomianu $V(x)$, to także jej sprzężenie $\overline{z_0}$ jest k-krotnym pierwiastkiem wielomianu $V(x)$.*

Dowód. Niech z_0 będzie k-krotnym pierwiastkiem wielomianu $V(x)$. Wtedy wobec definicji 3.4.2 istnieje wielomian $Q(x) \in \mathcal{C}[x]$ taki, że

$$V(x) = (x - z_0)^k Q(x) \text{ i } Q(z_0) \neq 0.$$

Stąd i z własności sprzężenia liczb zespolonych (zob. twierdzenie 2.2.1) mamy

$$\overline{V}(x) = (x - \overline{z_0})^k \overline{Q}(x) \text{ i } \overline{Q(z_0)} \neq 0.$$

Ponieważ $V(x)$ jest wielomianem o współczynnikach rzeczywistych, więc mamy $\overline{V}(x) = V(x)$ i dlatego także

$$V(x) = (x - \overline{z_0})^k \overline{Q}(x) \text{ i } \overline{Q}(\overline{z_0}) \neq 0 \text{ (bo } \overline{Q}(\overline{z_0}) = \overline{Q(z_0)} \neq 0).$$

To oznacza, że liczba $\overline{z_0}$ jest k-krotnym pierwiastkiem wielomianu $V(x)$. □

Przykład 51. Liczba $1 - j$ jest pierwiastkiem wielomianu

$$V(x) = x^4 + x^3 - 14x^2 + 26x - 20.$$

Wyznaczyć wszystkie pozostałe pierwiastki wielomianu $V(x)$.

Ponieważ $V(x)$ jest wielomianem o współczynnikach rzeczywistych i liczba $1 - j$ jest jego pierwiastkiem, więc wobec twierdzenia 3.4.5 także liczba $1 + j$ jest pierwiastkiem wielomianu $V(x)$. Zatem wielomian $V(x)$ jest podzielny przez dwumiany $x - (1 - j)$ i $x - (1 + j)$, więc

$$V(x) = \big(x - (1 - j)\big)\big(x - (1 + j)\big)Q(x) = (x^2 - 2x + 2)Q(x).$$

Dzieląc $V(x)$ przez x^2-2x+2 znajdujemy $Q(x)=x^2+3x-10$. Stąd

$$\begin{aligned} V(x) &= \big(x-(1-j)\big)\big(x-(1+j)\big)(x^2+3x-10) \\ &= \big(x-(1-j)\big)\big(x-(1+j)\big)(x-2)(x+5) \end{aligned}$$

i dlatego liczby $1-j$, $1+j$, 2 oraz -5 są pierwiastkami wielomianu $V(x)$.

Z wniosku 3.4.4 i twierdzenia 3.4.5 wynika, że każdy wielomian rzeczywisty można przedstawić w postaci iloczynu rzeczywistych nierozkładalnych wielomianów stopnia co najwyżej drugiego. Dokładniej mamy następujący wniosek.

Wniosek 3.4.5. *Niech $V(x)$ będzie rzeczywistym wielomianem stopnia $n>0$. Niech $x_1,\ldots,x_r$ będą jego rzeczywistymi pierwiastkami o krotnościach odpowiednio $k_1,\ldots,k_r$ i niech z_1, $\overline{z_1}$, $\ldots$, z_s, $\overline{z_s}$ (gdzie $\operatorname{Im}(z_j)\neq 0$ dla $j=1,\ldots,s$) będą jego zespolonymi pierwiastkami o krotnościach odpowiednio $l_1,\ldots,l_s$. Jeśli $k_1+k_2+\ldots+k_r+2(l_1+\ldots+l_s)=n$, to*

$$V(x)=a\prod_{i=1}^{r}(x-x_i)^{k_i}\prod_{j=1}^{s}(x^2+p_jx+q_j)^{l_j}, \tag{3.20}$$

gdzie $a\in R-\{0\}$ oraz $\Delta_j=p_j^2-4q_j<0$ dla $j=1,\ldots,s$.

Dowód. Z twierdzenia 3.4.5 i wniosku 3.4.4 wynika, że istnieje liczba $a\in\mathcal{C}-\{0\}$ taka, że

$$\begin{aligned} V(x) &= a\prod_{i=1}^{r}(x-x_i)^{k_i}\prod_{j=1}^{s}\big((x-z_j)(x-\overline{z_j})\big)^{l_j} \\ &= a\prod_{i=1}^{r}(x-x_i)^{k_i}\prod_{j=1}^{s}\big(x^2-(z_j+\overline{z_j})x+z_j\overline{z_j}\big)^{l_j} \\ &= a\prod_{i=1}^{r}(x-x_i)^{k_i}\prod_{j=1}^{s}(x^2+p_jx+q_j)^{l_j}, \end{aligned}$$

gdzie $p_j=-(z_j+\overline{z_j})=-2\operatorname{Re}(z_j)$, $q_j=z_j\overline{z_j}=\operatorname{Re}^2(z_j)+\operatorname{Im}^2(z_j)$ oraz $\Delta_j=p_j^2-4q_j=-4\operatorname{Im}^2(z_j)<0$. Z faktu, że współczynniki $V(x)$ oraz liczby p_j i q_j $(j=1,\ldots,s)$ są liczbami rzeczywistymi wynika, że także liczba a jest rzeczywista. □

Przykład 52. Wielomian $V(x)=x^4+1$ przedstawić w postaci iloczynu rzeczywistych wielomianów stopnia co najwyżej drugiego.

Ponieważ pierwiastkami wielomianu $V(x)=x^4+1$ są pierwiastki czwartego stopnia z liczby -1, więc mamy

$$V(x)=\left(x-\frac{\sqrt2}{2}-\frac{\sqrt2}{2}j\right)\left(x+\frac{\sqrt2}{2}-\frac{\sqrt2}{2}j\right)\left(x+\frac{\sqrt2}{2}+\frac{\sqrt2}{2}j\right)\left(x-\frac{\sqrt2}{2}+\frac{\sqrt2}{2}j\right).$$

Teraz, tak jak w dowodzie wniosku 3.4.5, grupujemy i wymnażamy czynniki odpowiadające sprzężonym pierwiastkom i otrzymujemy

$$\begin{aligned} V(x) &= \left[\left(x-\tfrac{\sqrt2}{2}-\tfrac{\sqrt2}{2}j\right)\left(x-\tfrac{\sqrt2}{2}+\tfrac{\sqrt2}{2}j\right)\right]\left[\left(x+\tfrac{\sqrt2}{2}-\tfrac{\sqrt2}{2}j\right)\left(x+\tfrac{\sqrt2}{2}+\tfrac{\sqrt2}{2}j\right)\right] \\ &= \left[\left(x-\tfrac{\sqrt2}{2}\right)^2+\tfrac12\right]\left[\left(x+\tfrac{\sqrt2}{2}\right)^2+\tfrac12\right]=(x^2-\sqrt2x+1)(x^2+\sqrt2x+1). \end{aligned}$$

Ten sam rozkład można otrzymać przez proste przekształcenia, bo mamy

$$\begin{aligned} V(x) &= x^4+1=(x^4+2x^2+1)-2x^2=(x^2+1)^2-(\sqrt2x)^2 \\ &= (x^2+1-\sqrt2x)(x^2+1+\sqrt2x)=(x^2-\sqrt2x+1)(x^2+\sqrt2x+1). \end{aligned}$$

Skupimy teraz naszą uwagę na pierwiastkach wymiernych wielomianów z wymiernymi współczynnikami. Ponieważ pierwiastki wielomianu nie zmieniają się przy mnożeniu wielomianu przez niezerową liczbę, więc w następnym twierdzeniu możemy ograniczyć się do wielomianów o współczynnikach całkowitych.

Twierdzenie 3.4.6. *Niech $V(x) = a_nx^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0$ będzie wielomianem dodatniego stopnia, którego współczynniki są liczbami całkowitymi. Niech p i q będą liczbami całkowitymi względnie pierwszymi. Jeśli ułamek p/q jest pierwiastkiem wielomianu $V(x)$, to p dzieli wyraz wolny a_0, a q dzieli współczynnik wiodący a_n wielomianu $V(x)$.*

Dowód. Załóżmy, że p/q jest pierwiastkiem wielomianu $V(x)$. Wtedy

$$V\left(\frac{p}{q}\right) = a_n\left(\frac{p}{q}\right)^n + a_{n-1}\left(\frac{p}{q}\right)^{n-1} + \ldots + a_1\left(\frac{p}{q}\right) + a_0 = 0. \tag{3.21}$$

Z (3.21), po pomnożeniu przez q^n, otrzymujemy

$$a_np^n + a_{n-1}p^{n-1}q + \ldots + a_1pq^{n-1} + a_0q^n = 0, \tag{3.22}$$

więc także

$$p(a_np^{n-1} + a_{n-1}p^{n-2}q + \ldots + a_1q^{n-1}) = -a_0q^n.$$

Stąd wynika, że p dzieli liczbę a_0q^n i dlatego p dzieli a_0, bo p i q są względnie pierwsze. Podobnie z równości (3.22) mamy

$$q(a_{n-1}p^{n-1} + \ldots + a_1pq^{n-2} + a_0q^{n-1}) = -a_np^n$$

i stąd także wynika, że q jest dzielnikiem liczby a_n. $\square$

Przykład 53. Wyznaczyć wszystkie pierwiastki wielomianu $V(x) = 2x^3 - 9x^2 + 14x - 5$ znajdując najpierw wymierne pierwiastki tego wielomianu.

Jeśli liczba wymierna p/q (zapisana w postaci nieskracalnej) jest pierwiastkiem wielomianu $V(x) = 2x^3 - 9x^2 + 14x - 5$, to wobec twierdzenia 3.4.6 jej licznik p jest dzielnikiem wyrazu wolnego $a_0 = -5$, a mianownik q jest dzielnikiem współczynnika wiodącego $a_3 = 2$. Dlatego

$$p \in \{\pm 1, \pm 5\}, \quad q \in \{\pm 1, \pm 2\} \quad \text{i} \quad \frac{p}{q} \in \left\{\pm 1, \pm\frac{1}{2}, \pm 5, \pm\frac{5}{2}\right\}.$$

Ponieważ żadna rzeczywista liczba ujemna nie może być pierwiastkiem wielomianu $V(x)$ (bo dla każdej liczby $x_0 < 0$ jest $V(x_0) < 0$), więc spośród liczb wymiernych jedynie liczba $\frac{p}{q} \in \left\{1, \frac{1}{2}, 5, \frac{5}{2}\right\}$ może być pierwiastkiem wielomianu $V(x)$. Obliczając wartość wielomianu $V(x)$ dla możliwych ułamków p/q, znajdujemy

$$V(1) = 2, \quad V\left(\frac{1}{2}\right) = 0, \quad V(5) = 90 \quad \text{i} \quad V\left(\frac{5}{2}\right) = -\frac{25}{2}.$$

Dlatego jedynym pierwiastkiem wymiernym wielomianu $V(x)$ jest $\frac{1}{2}$. Dzieląc teraz $V(x)$ przez $x - \frac{1}{2}$, otrzymujemy

$$V(x) = \left(x - \frac{1}{2}\right)(2x^2 - 8x + 10) = (2x - 1)(x^2 - 4x + 5).$$

Pozostałe pierwiastki wielomianu $V(x)$ są pierwiastkami ilorazu

$$\begin{aligned} Q(x) &= x^2 - 4x + 5 = (x^2 - 4x + 4) + 1 \\ &= (x - 2)^2 - j^2 = (x - 2 - j)(x - 2 + j). \end{aligned}$$

Ostatecznie pierwiastkami wielomianu $V(x)$ są liczby $1/2$, $2 + j$ i $2 - j$.

3.5. Wielomiany względnie pierwsze

Definicja 3.5.1. Niech $V(x)$ i $W(x)$ będą wielomianami z pierścienia $K[x]$. Mówimy, że wielomiany $V(x)$ i $W(x)$ są *względnie pierwsze,* gdy nie są one jednocześnie wielomianami zerowymi i nie są jednocześnie podzielne przez żaden wielomian dodatniego stopnia.

Wielomiany względnie pierwsze

Przykład 54. Wielomiany

$$V(x) = (x-1)(x+2) \quad \text{i} \quad W(x) = (x+2)(x+3)$$

nie są względnie pierwsze, bo każdy z nich jest podzielny przez $x+2$, wielomian dodatniego stopnia. Natomiast wielomiany

$$V(x) = (x-1)(x+2) \quad \text{i} \quad U(x) = (x-2)(x+3)$$

są względnie pierwsze, bo nie mają one wspólnego dzielnika dodatniego stopnia.

Twierdzenie 3.5.1. *Jeśli wielomiany $V(x)$ i $W(x)$ z pierścienia $K[x]$ są niezerowe i względnie pierwsze, to w pierścieniu $K[x]$ istnieją wielomiany $P(x)$ i $Q(x)$ takie, że*

$$P(x)V(x) + Q(x)W(x) = 1. \tag{3.23}$$

Dowód. Przedstawiamy dowód indukcyjny ze względu na liczbę $n = \min\{\deg V(x), \deg W(x)\}$. Jeśli $n = 0$, to co najmniej jeden z wielomianów $V(x)$ i $W(x)$ jest niezerową stałą i, oczywiście, dla nich istnieją takie wielomiany $P(x)$ i $Q(x)$, że zachodzi (3.23). (Przykładowo, jeśli $V(x) \equiv V$ i $V \neq 0$, to dla wielomianów $P(x) \equiv 1/V$ i $Q(x) \equiv 0$ mamy (3.23).)

Załóżmy z kolei, że dowodzone stwierdzenie jest prawdziwe dla każdych niezerowych i względnie pierwszych wielomianów $V_1(x)$ i $W_1(x)$ takich, że $\min\{\deg V_1(x), \deg W_1(x)\} < n$, gdzie n jest pewną dodatnią liczbą naturalną. Niech teraz $V(x)$ i $W(x)$ będą względnie pierwszymi wielomianami takimi, że $\min\{\deg V(x), \deg W(x)\} = n$. Bez zmniejszenia ogólności możemy założyć, że $\deg V(x) \geqslant \deg W(x) = n$.

Wobec twierdzenia 3.2.1 w pierścieniu $K[x]$ istnieją wielomiany $\overline{Q}(x)$ i $R(x)$ takie, że

$$V(x) = \overline{Q}(x)W(x) + R(x) \quad \text{i} \quad \deg R(x) < n = \deg W(x). \tag{3.24}$$

Z faktu, że $W(x)$ i $V(x)$ są względnie pierwsze łatwo wynika, że $R(x)$ nie jest wielomianem zerowym. (Inaczej byłoby $V(x) = \overline{Q}(x)W(x)$ i wbrew założeniu wielomiany $V(x)$ i $W(x)$ byłyby podzielne przez wielomian dodatniego stopnia, np. byłyby one podzielne przez sam wielomian $W(x)$.) Dodatkowo, wielomiany $W(x)$ i $R(x)$ są względnie pierwsze. (Gdyby było inaczej, to istniałby wielomian dodatniego stopnia, powiedzmy $S(x)$, dzielący jednocześnie $W(x)$ i $R(x)$. Wtedy, wobec (3.24), wielomian $S(x)$ byłby jednoczesnym dzielnikiem wielomianów $V(x)$ i $W(x)$, co przeczyłoby założeniu, że są one względnie pierwsze.) Ponieważ $\deg R(x) < n$, więc $\min\{\deg V(x), \deg R(x)\} < n$ i wobec założenia indukcyjnego istnieją wielomiany $\widetilde{Q}(x)$ i $P(x)$ takie, że

$$\widetilde{Q}(x)W(x) + P(x)R(x) = 1. \tag{3.25}$$

A wtedy z (3.24) i (3.25) otrzymujemy (3.23), bo mamy

$$\begin{aligned} 1 &= \widetilde{Q}(x)W(x) + P(x)\big(V(x) - \overline{Q}(x)W(x)\big) \\ &= P(x)V(x) + \big(\widetilde{Q}(x) - P(x)\overline{Q}(x)\big)W(x) \\ &= P(x)V(x) + Q(x)W(x), \end{aligned}$$

gdzie $Q(x) = \widetilde{Q}(x) - P(x)\overline{Q}(x)$. □

Dowód twierdzenia 3.5.1 sugeruje, w jaki sposób dla względnie pierwszych wielomianów $V(x)$ i $W(x)$ można wyznaczyć wielomiany $P(x)$ i $Q(x)$ spełniajace zależność (3.23).

Przykład 55. Dla względnie pierwszych wielomianów $V(x)$ i $W(x)$ wyznaczyć wielomiany $P(x)$ i $Q(x)$ takie, że $P(x)V(x) + Q(x)W(x) \equiv 1$, gdy:
(a) $V(x) = x^2 + x - 1$ i $W(x) = 4$;
(b) $V(x) = x^2 + x - 1$ i $W(x) = x - 2$;
(c) $V(x) = x^2 + x - 1$ i $W(x) = x^2 + 2$.

(a) Zauważmy, że wielomian $W(x)$ jest niezerową stałą, $W(x) \equiv 4$, więc dla wielomianów $P(x) \equiv 0$ i $Q(x) = \frac{1}{W(x)} = \frac{1}{4}$ mamy

$$P(x)V(x) + Q(x)W(x) = 0 \cdot V(x) + \frac{1}{4} \cdot 4 = 1.$$

(b) Tym razem żaden z wielomianów $V(x) = x^2 + x - 1$ i $W(x) = x - 2$ nie jest stały, ale niezerową stałą jest reszta z dzielenia $V(x)$ przez $W(x)$,

$$x^2 + x - 1 = V(x) = \overline{Q}(x)W(x) + R(x) = (x+3)(x-2) + 5, \tag{3.26}$$

gdzie $\overline{Q}(x) = x + 3$ i $R(x) = 5$ są odpowiednio ilorazem i resztą z dzielenia $V(x)$ przez $W(x)$. Wielomiany $R(x) \equiv 5 = R$ i $W(x) = x - 2$ są względnie pierwsze i dla nich (tak jak w części (a) było dla wielomianów $W(x)$ i $V(x)$) mamy

$$1 = \frac{1}{R} \cdot R(x) + 0 \cdot W(x),$$

a ponieważ wobec (3.26) jest $R(x) = V(x) - \overline{Q}(x)W(x)$, więc

$$1 = \frac{1}{R} \cdot \big(V(x) - \overline{Q}(x)W(x)\big) = P(x)V(x) + Q(x)W(x),$$

gdzie $P(x) = \frac{1}{R} = \frac{1}{5}$ i $Q(x) = -\frac{1}{R} \cdot \overline{Q}(x) = -\frac{1}{5}(x+3)$.

(c) Dla wielomianów $V(x)$ i $W(x)$ tym razem mamy

$$x^2 + x - 1 = V(x) = \overline{Q}(x)W(x) + R(x) = 1 \cdot (x^2+2) + (x-3), \tag{3.27}$$

i

$$x^2 + 2 = W(x) = \widetilde{Q}(x)R(x) + \overline{R}(x) = (x+3)(x-3) + 11, \tag{3.28}$$

gdzie $\overline{Q}(x) = 1$ i $R(x) = x - 3$ są odpowiednio ilorazem i resztą z dzielenia $V(x)$ przez $W(x)$, a $\widetilde{Q}(x) = x + 3$ i $\overline{R}(x) = 11$ są ilorazem i resztą z dzielenia $W(x)$ przez $R(x)$. Ponieważ $R(x)$ i $\overline{R}(x)$ są względnie pierwsze i wielomian $\overline{R}(x)$ jest stały, więc wobec (3.28) i (3.27) otrzymujemy

$$\begin{aligned}
1 &= \frac{1}{\overline{R}(x)} \cdot \overline{R}(x) + 0 \cdot R(x) \overset{(3.28)}{=} \frac{1}{\overline{R}(x)} \cdot \big(W(x) - \widetilde{Q}(x)R(x)\big) \\
&= \frac{1}{\overline{R}(x)} \cdot W(x) - \frac{\widetilde{Q}(x)}{\overline{R}(x)} \cdot R(x) \overset{(3.27)}{=} \frac{1}{\overline{R}(x)} \cdot W(x) - \frac{\widetilde{Q}(x)}{\overline{R}(x)} \cdot \big(V(x) - \overline{Q}(x)W(x)\big) \\
&= -\frac{\widetilde{Q}(x)}{\overline{R}(x)} \cdot V(x) + \left(\frac{1}{\overline{R}(x)} + \frac{\overline{Q}(x)\widetilde{Q}(x)}{\overline{R}(x)}\right) W(x) = P(x)V(x) + Q(x)W(x),
\end{aligned}$$

gdzie $P(x) = -\frac{\widetilde{Q}(x)}{\overline{R}(x)} = -\frac{1}{11}(x+3)$ i $Q(x) = \frac{1}{\overline{R}(x)} + \frac{\overline{Q}(x)\widetilde{Q}(x)}{\overline{R}(x)} = \frac{1}{11}(x+4)$.

3.6. Funkcje wymierne i ułamki proste

Definicja 3.6.1. Niech $P(x)$ i $Q(x)$ będą wielomianami nad ciałem K, powiedzmy $P(x) = a_m x^m + a_{m-1}x^{m-1} + \ldots + a_1 x + a_0$ i $Q(x) = b_n x^n + b_{n-1}x^{n-1} + \ldots + b_1 x + b_0$. Wtedy funkcję postaci

Funkcja wymierna

$$\frac{P(x)}{Q(x)} = \frac{a_m x^m + a_{m-1}x^{m-1} + \ldots + a_1 x + a_0}{b_n x^n + b_{n-1}x^{n-1} + \ldots + b_1 x + b_0} \tag{3.29}$$

nazywamy *funkcją wymierną* nad ciałem K. Funkcję wymierną (3.29) nazywamy *właściwą funkcją wymierną*, gdy $\deg P(x) < \deg Q(x)$.

Przykład 56. Funkcje

$$f(x) = \frac{x^5 + 2x - 1}{x^4 + 1}, \quad g(x) = \frac{1}{x} \quad \text{i} \quad h(x) = \frac{x - 7}{(x + 1)^3}$$

są funkcjami wymiernymi nad ciałem liczb rzeczywistych, a dwie ostatnie są także właściwymi funkcjami wymiernymi.

Definicja 3.6.2. *Ułamkiem prostym* (nad ciałem K) nazywamy funkcję postaci

Ułamek prosty

$$\frac{P(x)}{\big(Q(x)\big)^k},$$

gdzie $P(x)$, $Q(x) \in K[x]$, $\deg P(x) < \deg Q(x)$, wielomian $Q(x)$ jest nierozkładalny w pierścieniu $K[x]$ i k jest dodatnią liczbą naturalną.

Z definicji ułamka prostego i z wniosku 3.4.3 wynika, że każdy ułamek prosty nad ciałem liczb zespolonych jest postaci $A/(x - x_0)^k$, gdzie A i x_0 są liczbami zespolonymi, a k jest dodatnią liczbą naturalną. Natomiast wobec wniosku 3.4.5 ułamkami prostymi nad ciałem liczb rzeczywistych są tylko i wyłącznie funkcje postaci

Ułamki proste nad ciałem liczb rzeczywistych

$$\frac{A}{(x - x_0)^k} \quad \text{i} \quad \frac{Bx + C}{(x^2 + px + q)^k},$$

gdzie A, B, C, x_0, p i q są liczbami rzeczywistymi, $p^2 - 4q < 0$, a k jest dodatnią liczbą naturalną.

Znajdując iloraz $I(x)$ i resztę $R(x)$ z dzielenia wielomianu $P(x)$ przez $Q(x)$, otrzymujemy $P(x) = I(x)Q(x) + R(x)$, czyli

$$\frac{P(x)}{Q(x)} = I(x) + \frac{R(x)}{Q(x)},$$

gdzie $\deg R(x) < \deg Q(x)$. Oznacza to, że każdą funkcję wymierną można przedstawić w postaci sumy wielomianu i funkcji wymiernej właściwej. Okazuje się, i to jest ważne z praktycznego punktu widzenia, że funkcję wymierną właściwą można przedstawić w postaci sumy skończonej liczby ułamków prostych.

Twierdzenie 3.6.1. *Niech P i Q będą wielomianami o współczynnikach z ciała K i $0 \leqslant \deg P < \deg Q$. Załóżmy, że wielomian Q można przedstawić w postaci iloczynu*

$$Q = Q_1^{n_1} \cdot Q_2^{n_2} \cdot \ldots \cdot Q_k^{n_k}, \tag{3.30}$$

gdzie $Q_1, Q_2, \ldots, Q_k$ są nierozkładalnymi wielomianami w pierścieniu $K[x]$, a każde dwa spośród nich są względnie pierwsze i $n_1, n_2, \ldots, n_k$ są dodatnimi liczbami naturalnymi. Wtedy funkcję wymierną P/Q można jednoznacznie przedstawić w postaci sumy ułamków prostych,

$$\frac{P}{Q} = \sum_{i=1}^{k} \left(\frac{R_{i,n_i-1}}{Q_i} + \frac{R_{i,n_i-2}}{Q_i^2} + \ldots + \frac{R_{i,1}}{Q_i^{n_i-1}} + \frac{R_{i,0}}{Q_i^{n_i}} \right), \tag{3.31}$$

gdzie $R_{i,j}$ $(i = 1, \ldots, k,\ j = 1, \ldots, n_i - 1)$ są wielomianami o współczynnikach z ciała K i $\deg R_{i,j} < \deg Q_i$ $(j = 0, \ldots, n_i - 1)$.

Uwaga. Ze względów praktycznych warto pamiętać, że każdemu czynnikowi $Q_i^{n_i}$ iloczynu (3.30) w sumie (3.31) odpowiada suma n_i ułamków prostych

$$\frac{R_{i,n_i-1}}{Q_i} + \frac{R_{i,n_i-2}}{Q_i^2} + \ldots + \frac{R_{i,1}}{Q_i^{n_i-1}} + \frac{R_{i,0}}{Q_i^{n_i}},$$

których mianownikami są kolejne potęgi wielomianu Q_i.

Dowód. Udowodnimy tylko istnienie rozkładu (3.31). Z założenia, że każde dwa spośród wielomianów $Q_1, Q_2, \ldots, Q_k$ są względnie pierwsze wynika, że także każde dwie spośród ich potęg $Q_1^{n_1}, Q_2^{n_2}, \ldots, Q_k^{n_k}$ są względnie pierwsze. Stąd zaś łatwo wynika, że jedynka jest największym wspólnym podzielnikiem wielomianów $\frac{Q}{Q_1^{n_1}}, \frac{Q}{Q_2^{n_2}}, \ldots, \frac{Q}{Q_k^{n_k}}$. Dlatego istnieją wielomiany $S_1, S_2, \ldots, S_k \in K[x]$ takie, że

$$\frac{Q}{Q_1^{n_1}} S_1 + \frac{Q}{Q_2^{n_2}} S_2 + \ldots + \frac{Q}{Q_k^{n_k}} S_k = 1. \tag{3.32}$$

Zatem

$$\begin{aligned} \frac{P}{Q} = \frac{P}{Q} \cdot 1 &= \frac{P}{Q} \cdot \left(\frac{Q}{Q_1^{n_1}} S_1 + \frac{Q}{Q_2^{n_2}} S_2 + \ldots + \frac{Q}{Q_k^{n_k}} S_k \right) \\ &= \frac{PS_1}{Q_1^{n_1}} + \frac{PS_2}{Q_2^{n_2}} + \ldots + \frac{PS_k}{Q_k^{n_k}}. \end{aligned} \tag{3.33}$$

Udowodnimy obecnie, że każda funkcja $PS_i/Q_i^{n_i}$ jest sumą wielomianu i skończonej liczby ułamków prostych. Przez $I_{i,k}$ oraz $R_{i,k}$ $(0 \leqslant k < n_i)$ oznaczmy wielomiany określone rekurencyjnie w następujący sposób: niech $I_{i,0}$ oraz $R_{i,0}$ będą odpowiednio ilorazem i resztą z dzielenia wielomianu $P S_i$ przez Q_i, a dla $k = 1, \ldots, n_i - 1$ niech $I_{i,k}$ oraz $R_{i,k}$ będą odpowiednio ilorazem i resztą z dzielenia $I_{i,k-1}$ przez Q_i. Dla tak określonych wielomianów kolejno mamy

$$\begin{aligned} P S_i &= I_{i,0}Q_i + R_{i,0} \\ &= (I_{i,1}Q_i + R_{i,1})Q_i + R_{i,0} \\ &= ((I_{i,2}Q_i + R_{i,2})Q_i + R_{i,1})Q_i + R_{i,0} \\ &\vdots \\ &= ((\ldots(I_{i,n_i-1}Q_i + R_{i,n_i-1})Q_i + \ldots + R_{i,2})Q_i + R_{i,1})Q_i + R_{i,0} \\ &= I_{i,n_i-1}Q_i^{n_i} + R_{i,n_i-1}Q_i^{n_i-1} + \ldots + R_{i,2}Q_i^2 + R_{i,1}Q_i + R_{i,0} \end{aligned} \tag{3.34}$$

i jednocześnie $\deg R_{i,j} < \deg Q_i$ $(j = 0, \ldots, n_i - 1)$, bo $R_{i,j}$ jest resztą z dzielenia (pewnego wielomianu) przez wielomian Q_i. Z (3.34) wynika, że iloraz $PS_i/Q_i^{n_i}$ jest sumą wielomianu i ułamków prostych,

$$\begin{aligned} \frac{PS_i}{Q_i^{n_i}} &= \frac{I_{i,n_i-1}Q_i^{n_i} + R_{i,n_i-1}Q_i^{n_i-1} + \ldots + R_{i,2}Q_i^2 + R_{i,1}Q_i + R_{i,0}}{Q_i^{n_i}} \\ &= I_{i,n_i-1} + \frac{R_{i,n_i-1}}{Q_i} + \frac{R_{i,n_i-2}}{Q_i^2} + \ldots + \frac{R_{i,1}}{Q_i^{n_i-1}} + \frac{R_{i,0}}{Q_i^{n_i}}. \end{aligned} \tag{3.35}$$

Jeśli teraz przez J oznaczymy sumę wielomianów I_{i,n_i-1} $(i = 1, \ldots, k)$, to z (3.32) i (3.35) otrzymujemy rozkład funkcji P/Q na sumę wielomianu i ułamków prostych,

$$\frac{P}{Q} = J + \sum_{i=1}^{k} \left(\sum_{j=1}^{n_i} \frac{R_{i,n_i-j}}{Q_i^j} \right). \tag{3.36}$$

Dla dowodu (3.31) pozostaje wykazać, że J jest wielomianem zerowym. Przypuśćmy, że jest inaczej. Wtedy $\deg J \geqslant 0$ i z (3.36), po przemnożeniu przez Q, otrzymujemy

$$P = JQ + \sum_{i=1}^{k} \left(\sum_{j=1}^{n_i} \frac{R_{i,n_i-j}}{Q_i^j} Q \right). \tag{3.37}$$

Z nierówności $\deg R_{i,n_i-j} < \deg Q_i$ łatwo wynika, że

$$\deg \frac{R_{i,n_i-j}}{Q_i^j} Q < \deg Q \leqslant \deg JQ$$

dla $i = 1, \ldots, k$ oraz $j = 1, \ldots, n_i$. To zaś implikuje, że stopień całego wielomianu $\sum_{i=1}^{k} \left(\sum_{j=1}^{n_i} \frac{R_{i,n_i-j}}{Q_i^j} Q \right)$ jest mniejszy od stopnia wielomianu JQ. Z tej obserwacji i z (3.37) wynika, że wielomiany P i JQ mają identyczne stopnie. Jednakże wtedy $\deg P = \deg JQ \geqslant \deg Q$ i to jest sprzeczne z założeniem, że $\deg P < \deg Q$. Otrzymana sprzeczność kończy dowód twierdzenia. □

Dowód twierdzenia 3.6.1 podpowiada ogólną metodę przedstawiania funkcji wymiernej w postaci sumy ułamków prostych. Uzyskamy rozkład funkcji wymiernej właściwej P/Q na sumę ułamków prostych, gdy: (a) jej mianownik Q zapisze się w postaci (3.30), czyli w postaci iloczynu nierozkładalnych czynników; (b) wyznaczy się wielomiany S_i takie, że spełniony jest warunek (3.32) (można to zrobić za pomocą uogólnionego algorytmu Euklidesa) ; (c) ułamek P/Q, tak jak w (3.33), zapisze się w postaci sumy ułamków $PS_i/Q_i^{n_i}$ i (d) każdy z tych ułamków $PS_i/Q_i^{n_i}$ (za pomocą n_i-krotnego dzielenia z resztą przez Q_i) przedstawi się w postaci (3.35), tj. w postaci sumy wielomianu i n_i ułamków prostych.

Przykład 57. Przedstawić funkcję wymierną

$$\frac{P(x)}{Q(x)} = \frac{9x+9}{x^3+3x^2-4}$$

w postaci sumy ułamków prostych nad ciałem liczb rzeczywistych.

Rozkładu funkcji P/Q dokonamy w sposób (do pewnego stopnia) podobny do tego z dowodu twierdzenia 3.6.1. Mianownik rozważanej funkcji przedstawiamy w postaci iloczynu nierozkładalnych i względnie pierwszych czynników,

$$Q = x^3 + 3x^2 - 4 = (x-1)(x+2)^2 = Q_1 Q_2^2,$$

gdzie $Q_1 = x - 1$, $n_1 = 1$, $Q_2 = x + 2$ i $n_2 = 2$. Wyznaczamy teraz wielomiany S_1 i S_2, dla których spełniony jest warunek (3.32). Łatwo zauważyć, że jeśli $S_1 = 1/9$ i $S_2 = -(x+5)/9$, to istotnie mamy

$$\frac{Q}{Q_1} S_1 + \frac{Q}{Q_2^2} S_2 = \frac{(x+2)^2}{9} - \frac{(x-1)(x+5)}{9} = 1.$$

Stąd

$$\begin{aligned}\frac{P}{Q} = \frac{P}{Q} \cdot 1 &= \frac{9(x+1)}{(x-1)(x+2)^2} \cdot \left(\frac{(x+2)^2}{9} - \frac{(x-1)(x+5)}{9} \right) \\ &= \frac{x+1}{x-1} - \frac{x^2+6x+5}{(x+2)^2}\end{aligned}$$

i każdy z ostatnich ułamków łatwo przedstawia się w postaci sumy wielomianu i ułamków prostych i mamy

$$\begin{aligned}&= \frac{(x-1)+2}{x-1} - \frac{(x+2)^2+2(x+2)-3}{(x+2)^2} \\ &= \left(1 + \frac{2}{x-1}\right) - \left(1 + \frac{2}{x+2} - \frac{3}{(x+2)^2}\right) \\ &= \frac{2}{x-1} - \frac{2}{x+2} + \frac{3}{(x+2)^2}.\end{aligned}$$

Obecnie dokładniej opiszemy praktyczne sposoby przedstawiania rzeczywistej funkcji wymiernej właściwej w postaci sumy ułamków prostych (nad ciałem liczb rzeczywistych). Niech $P(x)/Q(x)$ będzie funkcją wymierną, gdzie $P(x)$ i $Q(x)$ są wielomianami o współczynnikach rzeczywistych i $0 \leqslant \deg P(x) < \deg Q(x) = n$. Wobec wniosku 3.4.5 wielomian $Q(x)$ można przedstawić w postaci iloczynu

$$Q(x) = a \prod_{i=1}^{r} (x - x_i)^{m_i} \prod_{j=1}^{s} (x^2 + p_j x + q_j)^{n_j}, \tag{3.38}$$

gdzie a jest pewną liczbą rzeczywistą, $x_1, \ldots, x_r$ są pierwiastkami rzeczywistymi wielomianu $Q(x)$, liczby rzeczywiste p_j i q_j są takie, że trójmian $x^2+p_jx+q_j$ jest nierozkładalny w pierścieniu $R[x]$ (dla $j = 1, \dots, s$), a $m_1, \dots, m_r$ i $n_1, \dots, n_s$ są dodatnimi liczbami naturalnymi takimi, że $m_1 + \ldots + m_r + 2(n_1 + \ldots + n_s) = n$. Wtedy, wobec twierdzenia 3.6.1 i naszych uwag o ułamkach prostych nad ciałem liczb rzeczywistych, funkcję $P(x)/Q(x)$ można przedstawić w postaci sumy

$$\frac{P(x)}{Q(x)} = \sum_{i=1}^{r} \left(\sum_{t=1}^{m_i} \frac{A_{i,t}}{(x - x_i)^t} \right) + \sum_{j=1}^{s} \left(\sum_{u=1}^{n_j} \frac{B_{j,u}x + C_{j,u}}{(x^2 + p_j x + q_j)^u} \right), \tag{3.39}$$

gdzie $A_{i,t}$ $(i = 1, \dots, r,\ t = 1, \dots, m_i)$ oraz $B_{j,u}$ i $C_{j,u}$ $(j = 1, \dots, s,\ u = 1, \dots, n_j)$ są liczbami rzeczywistymi.

W celu wyznaczenia współczynników $A_{i,t}$, $B_{j,u}$ i $C_{j,u}$ mnożymy obie strony równości (3.39) przez $Q(x)$ otrzymując równość dwóch wielomianów. Z lewej strony nowej równości wystąpi wielomian $P(x)$, powiedzmy $P(x) = b_m x^m + \ldots + b_1 x + b_0$ $(m \leqslant n - 1)$, a z prawej strony pojawi się pewien wielomian $W(x)$ stopnia $n-1$, powiedzmy $W(x) = c_{n-1}x^{n-1} + \ldots + c_1 x + c_0$, którego współczynniki są kombinacjami liniowymi współczynników $A_{i,t}$, $B_{j,u}$ i $C_{j,u}$. Równość

$$P(x) = W(x), \tag{3.40}$$

wobec wniosku 3.4.1, jest równoważna identyczności stopni i współczynników wielomianów $P(x)$ i $W(x)$. (Z drugiej strony, ponieważ $\deg P(x) \leqslant \deg W(x) = n - 1$, więc wobec wniosku 3.4.2 równość (3.40) jest równoważna temu, że dla pewnych n różnych liczb $\overline{x}_1, \dots, \overline{x}_n$ jest $P(\overline{x}_1) = W(\overline{x}_1)$, ..., $P(\overline{x}_n) = W(\overline{x}_n)$.)

Metoda porównywania współczynników

Przyrównując współczynniki stojące przy jednakowych potęgach zmiennej x w obu wielomianach, otrzymujemy układ n równań liniowych

$$c_0 = b_0,\ c_1 = b_1,\ \ldots,\ c_{n-1} = b_{n-1} \tag{3.41}$$

o n niewiadomych $A_{i,t}$, $B_{j,u}$ i $C_{j,u}$. Z twierdzenia 3.6.1 wynika, że układ (3.41) zawsze ma rozwiązanie. (Praktyczne sposoby rozwiązywania takich układów równań omawiamy w następnych rozdziałach.)

Przykład 58. Rozłożyć na sumę ułamków prostych nad ciałem R funkcję wymierną

$$\frac{P(x)}{Q(x)} = \frac{-15x + 6}{x^3 - 3x^2 - 6x + 8}. \tag{3.42}$$

Ponieważ $Q(x) = x^3 - 3x^2 - 6x + 8 = (x - 1)(x + 2)(x - 4)$, więc wobec (3.39) dla pewnych stałych A, B, C jest

$$\frac{P(x)}{Q(x)} = \frac{-15x + 6}{(x - 1)(x + 2)(x - 4)} = \frac{A}{x - 1} + \frac{B}{x + 2} + \frac{C}{x - 4}. \tag{3.43}$$

Dla wyznaczenia stałych A, B i C obie strony równości (3.43) mnożymy przez mianownik $Q(x)$ otrzymując równość

$$-15x + 6 = A(x + 2)(x - 4) + B(x - 1)(x - 4) + C(x - 1)(x + 2). \tag{3.44}$$

Wykonujemy działania po prawej stronie równości (3.44), grupujemy składniki według malejących potęg zmiennej x i otrzymujemy

$$-15x + 6 = (A + B + C)x^2 + (-2A - 5B + C)x + (-8A + 4B - 2C). \tag{3.45}$$

Przyrównujemy teraz współczynniki stojące przy jednakowych potęgach zmiennej x po obu stronach równości (3.45) otrzymując, układ równań

$$\begin{cases} A + B + C = 0, \\ -2A - 5B + C = -15, \\ -8A + 4B - 2C = 6. \end{cases}$$

Stąd $A = 1$, $B = 2$ oraz $C = -3$, więc szukanym rozkładem funkcji (3.42) na sumę ułamków prostych jest

$$\frac{-15x+6}{x^3-3x^2-6x+8} = \frac{1}{x-1} + \frac{2}{x+2} - \frac{3}{x-4}. \tag{3.46}$$

Przykład 59. Przedstawić w postaci sumy ułamków prostych (nad ciałem liczb rzeczywistych) funkcję

$$\frac{P(x)}{Q(x)} = \frac{12x^2+8x+26}{x^4+x^3-4x^2+2x-12}.$$

Ponieważ $Q(x) = x^4+x^3-4x^2+2x-12 = (x-2)(x+3)(x^2+2)$, więc dla pewnych stałych A, B, C i D jest

$$\frac{P(x)}{Q(x)} = \frac{A}{x-2} + \frac{B}{x+3} + \frac{Cx+D}{x^2+2}.$$

Mnożąc obie strony tej równości przez Q, otrzymujemy

$$12x^2+8x+26 = A(x+3)(x^2+2) + B(x-2)(x^2+2) + (Cx+D)(x-2)(x+3). \tag{3.47}$$

Stałe A, B, C i D, podobnie jak to zrobiliśmy w poprzednim przykładzie, można wyznaczyć z układu równań otrzymanego w wyniku przyrównania współczynników stojących przy jednakowych potęgach zmiennej x po obu stronach równości (3.47). Tym razem postąpimy inaczej. Stałe te wyznaczymy biorąc pod uwagę wartości lewej i prawej strony równości (3.47) dla czterech różnych konkretnych wartości x. Do równości (3.47), która jest prawdziwa dla każdego x, wygodnie jest wstawiać takie wartości x, dla których jeden z czynników iloczynu $Q(x) = (x-2)(x+3)(x^2+2)$ jest zerowy i/lub takie, dla których obliczanie wartości obu stron równości (3.47) jest proste.

Metoda porównywania wartości wielomianów

Z (3.47) dla $x = 2$ otrzymujemy $A = 3$, bo

$$90 = A(5)(6) + B(0) + (C \cdot 2 + D)(0).$$

Podobnie dla $x = -3$ mamy

$$110 = A(0) + B(-5)(11) + (C(-3) + D)(0),$$

więc $B = -2$. Dla $x = 0$ otrzymujemy

$$26 = A(6) + B(-4) + D(-6) = 18 + 8 - 6D$$

i dlatego $D = 0$. W końcu dla $x = 1$ jest

$$46 = A(4)(3) + B(-1)(3) + (C+D)(-1)(4) = 42 - 4C$$

i stąd $C = -1$. Zatem mamy rozkład

$$\frac{P(x)}{Q(x)} = \frac{3}{x-2} - \frac{2}{x+3} - \frac{x}{x^2+2}.$$

Przedstawiamy teraz kolejny dość wygodny sposób wyznaczania współczynników $A_{i,t}$ rozkładu (3.39) funkcji $P(x)/Q(x)$, sposób przez niektórych nazywany "metodą zakrywania". Tak jak w (3.38), niech x_i będzie m_i-krotnym pierwiastkiem wielomianu $Q(x)$. Przez $Z_{x_i}(x)$ oznaczamy funkcją powstałą z ułamka $(x-x_i)^{m_i}P(x)/Q(x)$ w wyniku podzielenia jego licznika i mianownika przez $(x-x_i)^{m_i}$, albo – co na to samo wychodzi – funkcją powstałą z funkcji $P(x)/Q(x)$ przez "zakrycie" czynnika $(x-x_i)^{m_i}$ w mianowniku $Q(x)$,

Metoda zakrywania

$$\begin{aligned} Z_{x_i}(x) &= \tfrac{(x-x_i)^{m_i}P(x)}{Q(x)} \\ &= \tfrac{P(x)}{(x-x_1)^{m_1}\ldots\cancel{(x-x_i)^{m_i}}\ldots(x-x_r)^{m_r}\prod_{j=1}^{s}(x^2+p_jx+q_j)^{n_j}}, \end{aligned} \tag{3.48}$$

gdzie kreskami ||||||| zakryliśmy czynnik $(x - x_i)^{m_i}$.

Każdemu pierwiastkowi x_i wielomianu $Q(x)$ w rozkładzie (3.39) odpowiada suma $\sum_{t=1}^{m_i} \frac{A_{i,t}}{(x-x_i)^t}$. Chcąc wyznaczyć jej współczynniki, zapiszmy rozkład (3.39) w postaci

$$\frac{P(x)}{Q(x)} = \frac{A_{i,m_i}}{(x-x_i)^{m_i}} + \frac{A_{i,m_i-1}}{(x-x_i)^{m_i-1}} + \ldots + \frac{A_{i,2}}{(x-x_i)^2} + \frac{A_{i,1}}{x-x_i} + S(x), \tag{3.49}$$

gdzie $S(x)$ jest sumą tych składników prawej strony rozkładu (3.39), które nie zawierają czynnika $x - x_i$, $S(x) = \frac{P(x)}{Q(x)} - \sum_{t=1}^{m_i} \frac{A_{i,t}}{(x-x_i)^t}$. Mnożąc obie strony równości (3.49) przez $(x - x_i)^{m_i}$ i uwzględniając (3.48), otrzymujemy

$$\begin{aligned}
Z_{x_i}(x) &= \sum_{t=0}^{m_i-1} A_{i,m_i-t}(x-x_i)^t + (x-x_i)^{m_i} S(x),\\
Z'_{x_i}(x) &= \sum_{t=1}^{m_i-1} tA_{i,m_i-t}(x-x_i)^{t-1} + \Big((x-x_i)^{m_i} S(x)\Big)',\\
Z''_{x_i}(x) &= \sum_{t=2}^{m_i-1} t(t-1)A_{i,m_i-t}(x-x_i)^{t-2} + \Big((x-x_i)^{m_i} S(x)\Big)'',\\
&\ \vdots\\
Z^{(m_i-1)}_{x_i}(x) &= (m_i-1)!A_{i,1} + \Big((x-x_i)^{m_i} S(x)\Big)^{(m_i-1)}.
\end{aligned}$$

Zauważmy, że każdy składnik pochodnej $((x-x_i)^{m_i} S(x))^{(k)}$ dla $k = 0, 1, \ldots, m_i - 1$ zawiera czynnik $x - x_i$ w dodatniej potędze, więc jego wartość dla $x = x_i$ jest równa zeru i dlatego z powyższych równości dla $x = x_i$ otrzymujemy

$$Z^{(k)}_{x_i}(x_i) = k!\, A_{i,m_i-k},$$

a stąd mamy wygodne wzory na współczynniki $A_{i,t}$ rozkładu (3.39) i (3.49),

$$A_{i,m_i-k} = \frac{Z^{(k)}_{x_i}(x_i)}{k!}. \tag{3.50}$$

Z (3.49) i (3.50) wynika, że wszystkie składniki rozkładu (3.39) odpowiadające czynnikowi $x - x_i$ można uzyskać metodą zakrywania i mamy

$$\begin{aligned}
\frac{P(x)}{Q(x)} = {} & \frac{Z_{x_i}(x_i)}{0!(x-x_i)^{m_i}} + \frac{Z'_{x_i}(x_i)}{1!(x-x_i)^{m_i-1}}\\
& + \frac{Z''_{x_i}(x_i)}{2!(x-x_i)^{m_i-2}} + \ldots + \frac{Z^{(m_i-1)}_{x_i}(x_i)}{(m_i-1)!(x-x_i)} + S(x).
\end{aligned} \tag{3.51}$$

Przykład 60. Przedstawić w postaci sumy ułamków prostych funkcję wymierną

$$\frac{P(x)}{Q(x)} = \frac{x+18}{x^2+x-6}.$$

Ponieważ mianownik $Q(x) = x^2+x-6 = (x-2)(x+3)$ ma tylko jednokrotne pierwiastki, więc metodą zakrywania (zob. (3.51)) otrzymujemy szukany rozkład

$$\begin{aligned}
\frac{P(x)}{Q(x)} &= \frac{x+18}{(x-2)(x+3)} = \frac{Z_2(2)}{x-2} + \frac{Z_{-3}(-3)}{x+3}\\
&= \frac{\left(\frac{x+18}{\cancel{(x-2)}\,(x+3)}\right)_{|x=2}}{x-2} + \frac{\left(\frac{x+18}{(x-2)\cancel{(x+3)}}\right)_{|x=-3}}{x+3} = \frac{4}{x-2} + \frac{-3}{x+3}.
\end{aligned}$$

Przykład 61. Funkcję $\dfrac{9x-27}{(x+1)(x-2)^2}$ przedstawić w postaci sumy ułamków prostych.

Wobec (3.51) mamy

$$\begin{aligned}\frac{9x-27}{(x+1)(x-2)^2} &= \frac{Z_{-1}(-1)}{x+1}+\frac{Z_2(2)}{(x-2)^2}+\frac{Z_2'(2)}{x-2}\\ &= \frac{\left(\frac{9x-27}{\cancel{(x+1)}(x-2)^2}\right)_{|x=-1}}{x+1}+\frac{\left(\frac{9x-27}{(x+1)\cancel{(x-2)^2}}\right)_{|x=2}}{(x-2)^2}+\frac{\left(\frac{9x-27}{(x+1)\cancel{(x-2)^2}}\right)'_{|x=2}}{x-2}\\ &= \frac{-4}{x+1}+\frac{-3}{(x-2)^2}+\frac{4}{x-2}.\end{aligned}$$

Uzasadniliśmy, że w rozkładzie właściwej funkcji wymiernej P/Q na sumę ułamków prostych współczynniki odpowiadające pierwiastkom wielomianu Q można wyznaczyć metodą zakrywania (zob. (3.51)). Ponieważ w rozkładzie funkcji P/Q na sumę ułamków prostych nad ciałem liczb zespolonych każdy składnik odpowiada jakiemuś pierwiastkowi wielomianu Q, więc wszystkie współczynniki tego rozkładu można wyznaczyć metodą zakrywania. Stąd zaś wynika, że metodą zakrywania można także otrzymać wszystkie współczynniki rozkładu funkcji P/Q na sumę ułamków prostych nad ciałem liczb rzeczywistych. W tym celu funkcję P/Q rozkładamy najpierw na sumę ułamków prostych nad ciałem $\mathcal{C}$ (wyznaczając współczynniki tego rozkładu metodą zakrywania). Z tego rozkładu po dodaniu do siebie każdych dwóch ułamków postaci $A/(x-x_0)^k$ i $B(x-\overline{x_0})^k$ (odpowiadających tej samej potędze k i sprzężonym pierwiastkom x_0 i $\overline{x_0}$ wielomianu Q) otrzymujemy rozkład funkcji P/Q na sumę rzeczywistych ułamków prostych. W następnym przykładzie w ten sposób otrzymujemy rozkład funkcji wymiernej na sumę ułamków prostych nad ciałem liczb rzeczywistych.

Przykład 62. Funkcję $\dfrac{P(x)}{Q(x)} = \dfrac{2x-8}{(x-1)(x^2+2)}$ rozłożyć na sumę ułamków prostych: (a) nad ciałem $\mathcal{C}$; (b) nad ciałem R.

W pierścieniu $\mathcal{C}[x]$ jest $Q(x)=(x-1)(x-\sqrt{2}j)(x+\sqrt{2}j)$ i pierwiastki wielomianu $Q(x)$ są jednokrotne, więc wobec (3.51) mamy

$$\frac{P(x)}{Q(x)}=\frac{2x-8}{(x-1)(x-\sqrt{2}j)(x+\sqrt{2}j)}=\frac{Z_1(1)}{x-1}+\frac{Z_{\sqrt{2}j}(\sqrt{2}j)}{x-\sqrt{2}j}+\frac{Z_{-\sqrt{2}j}(-\sqrt{2}j)}{x+\sqrt{2}j}.$$

Zauważmy, że

$$\begin{aligned}Z_1(1) &= \left(\frac{2x-8}{\cancel{(x-1)}(x-\sqrt{2}j)(x+\sqrt{2}j)}\right)_{|x=1} = -2\\ Z_{\sqrt{2}j}(\sqrt{2}j) &= \left(\frac{2x-8}{(x-1)\cancel{(x-\sqrt{2}j)}(x+\sqrt{2}j)}\right)_{|x=\sqrt{2}j} = 1-\sqrt{2}j\\ Z_{-\sqrt{2}j}(-\sqrt{2}j) &= \left(\frac{2x-8}{(x-1)(x-\sqrt{2}j)\cancel{(x+\sqrt{2}j)}}\right)_{|x=-\sqrt{2}j} = 1+\sqrt{2}j,\end{aligned}$$

więc

$$\frac{P(x)}{Q(x)}=\frac{-2}{x-1}+\frac{1-\sqrt{2}j}{x-\sqrt{2}j}+\frac{1+\sqrt{2}j}{x+\sqrt{2}j}$$

jest rozkładem funkcji $P(x)/Q(x)$ na sumę zespolonych ułamków prostych.

Z rozkładu tego, po dodaniu do siebie dwóch ostatnich ułamków (i wyeliminowaniu "zespoloności"), otrzymujemy rozkład funkcji $P(x)/Q(x)$ na sumę rzeczywistych ułamków prostych,

$$\frac{P(x)}{Q(x)}=\frac{-2}{x-1}+\frac{2x+4}{x^2+2}.$$

3.7. Ćwiczenia

1. W pierścieniu $Z_5[x]$ wyznaczyć $V(x)+W(x)$, $V(x)-2W(x)$, $V(x)W(x)$ i $V^2(x)+W^3(x)$, gdy $V(x)=x^3+2x^2+3x+4$ i $W(x)=x^3+3x^2+2$.
2. Niech $Q(x)=x^3+5x-1$ i $R(x)=-23x+5$ będą ilorazem i resztą z dzielenia wielomianu $V(x)$ przez wielomian $W(x)=x^2+2x+5$ w pierścieniu $R[x]$. Wyznaczyć wielomian $V(x)$.
3. Zbadać podzielność wielomianu $V(x)=x^3+x^2+x+1$ przez wielomian $W(x)=x^2+3x+2$ w pierścieniu: (a) $R[x]$; (b) $Z_5[x]$; (c) $Z_7[x]$.
4. Korzystając ze schematu Hornera, wyznaczyć iloraz $Q(x)$ i resztę $R(x)$ z dzielenia wielomianu $V(x)$ przez dwumian $W(x)$, gdy:
   (a) $V(x)=x^3-4x^2+x-3$, $W(x)=x-2$;
   (b) $V(x)=x^4-4x^3-10x^2-4x+4$, $W(x)=x+1$;
   (c) $V(x)=-4x^7+12x^6-15x^5+21x^4-8x^3+7x^2-4x+6$, $W(x)=x-2$.
5. Obliczyć iloraz $Q(x)$ i resztę $R(x)$ z dzielenia wielomianu $V(x)$ przez wielomian $W(x)$, gdy:
   (a) $V(x)=8x^4+3x^2+6$, $W(x)=x+2$;
   (b) $V(x)=x^3+8$, $W(x)=x^2-2x+4$;
   (c) $V(x)=jx^2+x-j$, $W(x)=x-j$ w $\mathcal{C}[x]$;
   (d) $V(x)=x^2+2x+2$, $W(x)=2x+2$ w $Z_3[x]$;
   (e) $V(x)=2x^3+3x^2+4x+1$, $W(x)=3x+1$ w $Z_5[x]$;
   (f) $V(x)=x^5+2x^4+3x^3+4x^2+3x+1$, $W(x)=x+1$ w $Z_5[x]$.
6. Wyznaczyć resztę $R(x)$ z dzielenia wielomianu $V(x)$ przez wielomian $W(x)$, gdy:
   (a) $V(x)=x^4-1$, $W(x)=x-2$ w $Z_5[x]$;
   (b) $V(x)=2x^{2000}+1999x+2$, $W(x)=x^2-1$ w $R[x]$;
   (c) $V(x)=6x^{13}+x$, $W(x)=x^2+1$ w $\mathcal{C}[x]$;
   (d) $V(x)=x^{110}-2x^{55}+1$, $W(x)=x^2+1$ w $\mathcal{C}[x]$.
7. Wyznaczyć wszystkie pierwiastki wielomianu $V(x)$, jeśli x_1 jest jednym z pierwiastków wielomianu $V(x)$:
   (a) $V(x)=x^3-x^2-7x+15$, $x_1=2+j$;
   (b) $V(x)=x^3-6x^2+21x-26$, $x_1=2+3j$;
   (c) $V(x)=x^4-2x^3+9x^2-8x+20$, $x_1=2j$;
   (d) $V(x)=x^4+x^3+2x^2+x+1$, $x_1=-j$;
   (e) $V(x)=x^4-6x^3+11x^2+12x-26$, $x_1=3+2j$;
   (f) $V(x)=x^4+3x^3+2x^2-x+5$, $x_1=-2+j$.
8. Wyznaczyć krotność pierwiastka $x_0=1$ wielomianu $V(x)$, gdy:
   (a) $V(x)=x^4-x^3-3x^2+5x-2$;
   (b) $V(x)=x^5-2x^4-2x^3+8x^2-7x+2$.
9. Znaleźć pierwiastki wielomianu $V(x)=x^3-15x^2+76x-140$, jeśli jednym z jego pierwiastków jest liczba całkowita.
10. Wyznaczyć pierwiastki wielomianu $V(x)=x^4-8x^3+20x^2-72x+99$, jeśli jednym z nich jest liczba czysto urojona.
11. Wyznaczyć pierwiastki wielomianu $V(x)$, gdy:
    (a) $V(x)=x^2+3x+3-j$;
    (b) $V(x)=x^2+(2j-1)x+1+5j$;
    (c) $V(x)=x^2-(5+j)x+8+j$;
    (d) $V(x)=12x^3-4x^2-3x+1$;
    (e) $V(x)=x^4-x^3-2x^2+6x-4$.
12. Rozwiązać następujące równania:
    (a) $x^2-(3+7j)x-10+11j=0$;
    (b) $(3+j)x^2+(1-j)x-6j=0$;
    (c) $x^4+2jx^2+8=0$;
    (d) $x^4-(3+6j)x^2-8+6j$.
13. Wielomian $V(x)$ przedstawić w postaci iloczynu wielomianów stopnia pierwszego, gdy:
    (a) $V(x)=x^3-6x^2+11x-6$; (b) $V(x)=x^4+16$.
14. Wielomian $V(x)$ przedstawić w postaci iloczynu rzeczywistych wielomianów stopnia co najwyżej drugiego, gdy:
    (a) $V(x)=x^3+x^2-x+2$;
    (b) $V(x)=x^4+4x^3+4x^2-1$;
    (c) $V(x)=3x^4-5x^3+3x^2+4x-2$;
    (d) $V(x)=x^6+27$.
15. Znaleźć największy wspólny dzielnik i najmniejszą wspólną wielokrotność wielomianów $V(x)$ i $W(x)$, gdy:
    (a) $V(x)=6(x-1)^3(x+2)^2(x-3)(x^2+4)^2$ i $W(x)=4(x-1)^2(x+2)^3(x+5)(x^2+4)$;
    (b) $V(x)=x^5+x^4-x^3-2x-1$ i $W(x)=3x^4+2x^3+x^2+2x-2$.
16. Wyznaczyć wielomiany $P(x)$ i $Q(x)$ takie, że $V(x)P(x)+W(x)Q(x)$ jest największym wspólnym dzielnikiem wielomianów $V(x)=x^4+2x^3-x^2-4x-2$ i $W(x)=x^4+x^3-x^2-2x-2$.
17. Podane funkcje wymierne przedstawić w postaci sumy ułamków prostych nad ciałem liczb rzeczywistych:
    (a) $\dfrac{4}{(x-1)(x+3)}$; (b) $\dfrac{2}{(x-1)(x^2+1)}$;
    (c) $\dfrac{x}{(x-1)(x-2)^2}$; (d) $\dfrac{x^2-1}{x^2(2x+1)}$;
    (e) $\dfrac{x^4+x^2+1}{x(x^2+1)^2}$; (f) $\dfrac{1}{(x^3-1)^2}$; (g) $\dfrac{4x^2}{x^4-1}$;
    (h) $\dfrac{x^2-6x+4}{x^4-3x^3+2x^2}$; (i) $\dfrac{3x^2-2x-1}{(x-3)(x^2+1)}$.
18. Podane funkcje wymierne rozłożyć na sumę ułamków prostych nad ciałem liczb zespolonych:
    (a) $\dfrac{2xj+8}{x^2+4}$; (b) $\dfrac{16j}{x^4+4}$; (c) $\dfrac{x^2+x+1}{x^4+x^2}$;
    (d) $\dfrac{x^2+2x}{(x^2+2x+2)^2}$; (e) $\dfrac{24x-8j}{(x^2+1)^2(x-j)}$.
19. Udowodnić, że wielomian rzeczywisty $V(x)$ jest podzielny przez dwumian $x-1$ wtedy i tylko wtedy, gdy suma współczynników wielomianu $V(x)$ jest równa zeru.

20. Udowodnić, że jeśli liczba naturalna m jest dzielnikiem liczby naturalnej n, to wielomian $x^m - a^m$ jest dzielnikiem wielomianu $x^n - a^n$ $(a \in R)$.
21. Niech x_0 będzie jednokrotnym pierwiastkiem wielomianu $Q(x)$ i niech $\frac{A}{x-x_0} + S(x)$ będzie rozkładem funkcji wymiernej właściwej $P(x)/Q(x)$ na sumę ułamków prostych, gdzie A jest liczbą i $S(x)$ jest funkcją ciągłą w punkcie x_0. Udowodnić, że $A = \frac{P(x_0)}{Q'(x_0)}$.
22. Niech $V(x) = a_n x^n + \ldots + a_1 x + a_0$ będzie wielomianem, którego współczynniki $a_n, \ldots, a_1, a_0$ są liczbami rzeczywistymi. Pokazać (bez odwoływania się do twierdzenia 3.4.5), że jeśli liczba zespolona z_0 jest pierwiastkiem wielomianu $V(x)$, to także liczba sprzężona $\overline{z_0}$ jest pierwiastkiem wielomianu $V(x)$.
23. Niech $x_1, x_2, \ldots, x_n$ będą pierwiastkami wielomianu $V(x) = x^n + a_1 x^{n-1} + \ldots + a_n$. Udowodnić, że prawdziwe są następujące związki

$$\begin{cases} a_1 = -(x_1 + x_2 + \ldots + x_n), \\ a_2 = x_1x_2 + x_1x_3 + \ldots + x_{n-1}x_n, \\ a_3 = -(x_1x_2x_3 + x_1x_2x_4 + \ldots + x_{n-2}x_{n-1}x_n), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_n = (-1)^n x_1 x_2 \cdot \ldots \cdot x_n, \end{cases}$$

nazywane wzorami Viète'a.

24. Wpisując TAK albo NIE, stwierdzić prawdziwość każdego z następujących zdań:

$\boxed{1\quad}$ Jeśli $V(x) = 3x^3 + 5x + 1$ i $W(x) = 2x + 4$ są wielomianami z pierścienia $Z_6[x]$, to $\deg(V(x)W(x)) = \deg V(x) + \deg W(x)$.

$\boxed{2\quad}$ Jeśli $V(x) = 3x^3 + 5x + 1$ i $W(x) = 2x + 4$ są wielomianami z pierścienia $Z_7[x]$, to $\deg(V(x)W(x)) = \deg V(x) + \deg W(x)$.

$\boxed{3\quad}$ Największy wspólny dzielnik dwóch niezerowych wielomianów z pierścienia $R[x]$ wyznaczony jest w sposób jednoznaczny.

$\boxed{4\quad}$ Wielomiany $2x + 6$ oraz $4x^2 + 12$ są względnie pierwsze w pierścieniu $R[x]$.

$\boxed{5\quad}$ Jeśli $V(x)$ i $W(x)$ są niezerowymi wielomianami z pierścienia $R[x]$ takimi, że $V(x)$ jest dzielnikiem $W(x)$ i $W(x)$ jest dzielnikiem $V(x)$, to $V(x) = W(x)$.

$\boxed{6\quad}$ Jeśli wielomian $V(x)$ jest dzielnikiem wielomianów $W(x)$ i $U(x)$ w pierścieniu $R[x]$, to $V(x)$ jest dzielnikiem wielomianu $W(x)S(x)+U(x)T(x)$ dla każdych wielomianów $S(x)$ i $T(x)$ z pierścienia $R[x]$.

$\boxed{7\quad}$ Zbiór wszystkich funkcji wymiernych (nad ustalonym ciałem K) jest ciałem ze względu na dodawanie i mnożenie funkcji.

Rozdział 4

MACIERZE

4.1. Podstawowe definicje

Definicja 4.1.1. *Macierzą* (dokładniej – macierzą o m wierszach i n kolumnach albo macierzą wymiaru $m \times n$) nazywamy prostokątną tablicę

Macierz wymiaru $m \times n$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \cdots & a_{mn} \end{bmatrix} \tag{4.1}$$

utworzoną z elementów a_{ij} ($1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$) ustalonego zbioru lub struktury algebraicznej K. W naszych dalszych rozważaniach K będzie ciałem liczb rzeczywistych lub ciałem liczb zespolonych. (W pewnych przypadkach będziemy także rozważać macierze, których elementy są funkcjami, wektorami lub symbolami pewnych obiektów.) Elementy a_{ij} macierzy $\mathbf{A}$ nazywamy jej *współczynnikami*. Na oznaczenie macierzy (4.1) będziemy także używać symbolu $[a_{ij}]_{m\times n}$ lub $[a_{ij}]$ (gdy znany będzie jej wymiar). Symbolem $K_{m\times n}$ oznaczać będziemy zbiór wszystkich macierzy wymiaru $m \times n$, których współczynniki należą do ciała K. Zatem $R_{m\times n}$ jest zbiorem wszystkich rzeczywistych macierzy wymiaru $m \times n$, a $\mathcal{C}_{m\times n}$ jest zbiorem wszystkich zespolonych macierzy wymiaru $m \times n$.

a_{ij} – współczynniki macierzy $\mathbf{A} = [a_{ij}]$

$K_{m\times n}$ – zbiór macierzy wymiaru $m \times n$

Macierze

Wiersz i kolumna macierzy

$$\mathbf{a}_{i*} = \begin{bmatrix} a_{i1} & a_{i2} & a_{i3} & \dots & a_{in} \end{bmatrix} \in K_{1\times n} \quad \text{oraz} \quad \mathbf{a}_{*j} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \in K_{m\times 1}$$

nazywamy odpowiednio i-tym *wierszem* oraz j-tą *kolumną* macierzy (4.1). Warto zauważyć, że współczynnik a_{ij} jest elementem macierzy (4.1) znajdującym się jednocześnie w i-tym wierszu oraz w j-tej kolumnie. (W niektórych dowodach element macierzy $\mathbf{X}$ znajdujący się w jej i-tym wierszu oraz w j-tej kolumnie oznaczać będziemy przez $(\mathbf{X})_{ij}$.)

$$\begin{array}{cl} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \cdots & a_{mn} \end{bmatrix} & \leftarrow i\text{-ty wiersz} \\ \uparrow j\text{-ta kolumna} & \end{array}$$

Przykład 63. Macierz

$$\mathbf{A} = [a_{ij}] = \begin{bmatrix} 2 & 3 & -1 \\ 0 & 1 & 4 \end{bmatrix}$$

ma dwa wiersze i trzy kolumny, a jej elementami są $a_{11} = 2$, $a_{12} = 3$, $a_{13} = -1$, $a_{21} = 0$, $a_{22} = 1$ i $a_{23} = 4$.

W pewnych przypadkach wygodnie będzie macierz (4.1) utożsamiać z macierzą jednowierszową lub jednokolumnową

$$\begin{bmatrix} \mathbf{a}_{*1} & \mathbf{a}_{*2} & \dots & \mathbf{a}_{*n} \end{bmatrix} \quad \text{i} \quad \begin{bmatrix} \mathbf{a}_{1*} \\ \mathbf{a}_{2*} \\ \vdots \\ \mathbf{a}_{m*} \end{bmatrix},$$

której elementami są odpowiednio kolejne kolumny i kolejne wiersze macierzy (4.1). Będziemy także pisać

$$\begin{bmatrix} | & | & & | \\ \mathbf{a}_{*1} & \mathbf{a}_{*2} & \dots & \mathbf{a}_{*n} \\ | & | & & | \end{bmatrix} \quad \text{zamiast} \quad \begin{bmatrix} \mathbf{a}_{*1} & \mathbf{a}_{*2} & \dots & \mathbf{a}_{*n} \end{bmatrix}$$

i

$$\begin{bmatrix} - & \mathbf{a}_{1*} & - \\ - & \mathbf{a}_{2*} & - \\ & \vdots & \\ - & \mathbf{a}_{m*} & - \end{bmatrix} \quad \text{zamiast} \quad \begin{bmatrix} \mathbf{a}_{1*} \\ \mathbf{a}_{2*} \\ \vdots \\ \mathbf{a}_{m*} \end{bmatrix}$$

dla podkreślenia, że chodzi o kolumny lub wiersze macierzy.

Przykład 64. Macierz $\mathbf{A}$ z poprzedniego przykładu możemy utożsamiać z macierzą jednowierszową $\mathbf{B}$ i z macierzą jednokolumnową $\mathbf{C}$, gdzie

$$\mathbf{B} = \begin{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} & \begin{bmatrix} 3 \\ 1 \end{bmatrix} & \begin{bmatrix} -1 \\ 4 \end{bmatrix} \end{bmatrix} \quad \text{i} \quad \mathbf{C} = \begin{bmatrix} \begin{bmatrix} 2 & 3 & -1 \end{bmatrix} \\ \begin{bmatrix} 0 & 1 & 4 \end{bmatrix} \end{bmatrix}.$$

Definicja 4.1.2. Macierz $\mathbf{O} = [o_{ij}] \in K_{m \times n}$ nazywamy *macierzą zerową* wymiaru $m \times n$, gdy każdy jej współczynnik o_{ij} jest równy zeru,

$$\mathbf{O} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Macierz zerowa

Definicja 4.1.3. *Macierzą kwadratową* stopnia n nazywamy każdą macierz $\mathbf{A} = [a_{ij}]$ należącą do zbioru $K_{n \times n}$, czyli macierz mającą tyle samo wierszy co kolumn,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}. \tag{4.2}$$

Macierz kwadratowa

Definicja 4.1.4. Ciąg $(a_{11},\, a_{22}, \ldots, a_{nn})$ elementów macierzy kwadratowej (4.2) nazywa się *główną przekątną* macierzy (4.2). O macierzy kwadratowej (4.2) mówimy, że jest *macierzą diagonalną*, jeśli wszystkie jej elementy znajdujące się poza główną przekątną są równe zeru, czyli jeśli

$$a_{ij} = 0 \ \text{dla} \ i \neq j.$$

Taką macierz diagonalną oznaczamy przez diag $(a_{11}, a_{22}, \ldots, a_{nn})$. Zatem mamy

Macierz diagonalna

$$\operatorname{diag}(a_{11}, a_{22}, \ldots, a_{nn}) = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}.$$

Macierzą jednostkową stopnia n, oznaczamy ją przez $\mathbf{I}_n$ lub $\mathbf{I}$, nazywamy macierz diagonalną diag $(a_{11}, a_{22}, \ldots, a_{nn})$, w której $a_{11} = a_{22} = \ldots = a_{nn} = 1$. Mamy zatem

$\mathbf{I}_n$ – macierz jednostkowa

$$\mathbf{I}_n = \operatorname{diag}(\underbrace{1, 1, \ldots, 1}_{n}) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Przykład 65. Macierze

$$\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

są diagonalne i macierz $\mathbf{C}$ jest macierzą jednostkową stopnia 4, $\mathbf{C} = \mathbf{I}_4$.

4.2. Działania na macierzach

Równość macierzy

Definicja 4.2.1. Macierze $\mathbf{A} = [a_{ij}] \in K_{m\times n}$ i $\mathbf{B} = [b_{ij}] \in K_{s\times t}$ nazywamy *równymi* i piszemy $\mathbf{A} = \mathbf{B}$, gdy

$$m = s, \quad n = t \quad \text{i} \quad a_{ij} = b_{ij}$$

dla $i = 1 \ldots, m$ oraz $j = 1, \ldots, n$.

Definicja 4.2.2. *Sumą macierzy* $\mathbf{A} = [a_{ij}]$ i $\mathbf{B} = [b_{ij}]$ należących do zbioru $K_{m\times n}$ (gdzie K jest ciałem, a m i n są liczbami naturalnymi) nazywamy macierz $\mathbf{A} + \mathbf{B} = [c_{ij}] \in K_{m\times n}$, której elementy określone są wzorami

Suma macierzy:
$[a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}]$

$$c_{ij} = a_{ij} + b_{ij}$$

dla $i = 1, \ldots, m$ oraz $j = 1, \ldots, n$. Dlatego mamy

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

Przykład 66. Dla macierzy rzeczywistych

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & -1 \\ 1 & 3 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 0 & -2 \end{bmatrix} \quad \text{i} \quad \mathbf{C} = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}$$

suma $\mathbf{A} + \mathbf{B}$ istnieje i

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} 0 & 2 & -1 \\ 1 & 3 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 3 & 5 \\ 2 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 5 & 4 \\ 3 & 3 & 0 \end{bmatrix},$$

ale suma macierzy $\mathbf{A}$ i $\mathbf{C}$ (oraz $\mathbf{B}$ i $\mathbf{C}$) nie istnieje, bo macierze $\mathbf{A}$ i $\mathbf{C}$ (oraz $\mathbf{B}$ i $\mathbf{C}$) mają różne wymiary.

Definicja 4.2.3. *Iloczynem macierzy* $\mathbf{A} = [a_{ij}] \in K_{m \times n}$ *przez skalar* r z ciała K nazywamy macierz $r\mathbf{A} = [b_{ij}] \in K_{m \times n}$, w której $b_{ij} = ra_{ij}$ dla $i = 1, \ldots, m$ oraz $j = 1, \ldots, n$, czyli

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

Iloczyn macierzy przez skalar: $r[a_{ij}] = [ra_{ij}]$

Przykład 67. Mamy

$$2 \begin{bmatrix} -1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} -2 & 4 \\ 6 & 8 \end{bmatrix}.$$

Definicja 4.2.4. Dla macierzy $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k \in K_{m \times n}$ i dla skalarów $\alpha_1, \alpha_2, \ldots, \alpha_k$ z ciała K, macierz

$$\alpha_1 \mathbf{A}_1 + \alpha_2 \mathbf{A}_2 + \ldots + \alpha_k \mathbf{A}_k$$

Kombinacja macierzy

nazywamy *kombinacją liniową* macierzy $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ ze współczynnikami $\alpha_1, \ldots, \alpha_k$.

Definicja 4.2.5. *Różnicą macierzy* $\mathbf{A}$ i $\mathbf{B}$ tego samego wymiaru nazywamy macierz $\mathbf{A} - \mathbf{B}$, gdzie

$$\mathbf{A} - \mathbf{B} = \mathbf{A} + (-1)\mathbf{B}.$$

Różnica macierzy: $[a_{ij}] - [b_{ij}] = [a_{ij} - b_{ij}]$

Przykład 68. Jeśli

$$\mathbf{A} = \begin{bmatrix} -2 & 1 \\ 3 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 3 & -2 \\ 1 & 4 \end{bmatrix} \quad \text{i} \quad \mathbf{O} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

to macierz

$$2\mathbf{A} - \mathbf{B} + 3\mathbf{O} = 2 \begin{bmatrix} -2 & 1 \\ 3 & 2 \end{bmatrix} - \begin{bmatrix} 3 & -2 \\ 1 & 4 \end{bmatrix} + 3 \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -7 & 4 \\ 5 & 0 \end{bmatrix}$$

jest kombinacją liniową macierzy $\mathbf{A}$, $\mathbf{B}$ i $\mathbf{O}$ ze współczynnikami 2, -1 i 3.

Z własności ciała K oraz z definicji 4.2.1, 4.2.2 i 4.2.3 natychmiast wynikają następujące własności dodawania macierzy i mnożenia macierzy przez skalary.

Twierdzenie 4.2.1. *Dla skalarów α, $\beta \in K$ i macierzy* $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C} \in K_{m\times n}$ *oraz macierzy zerowej* $\mathbf{O}$ *wymiaru* $m \times n$ *jest:*

(*a*) $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$;

(*b*) $\mathbf{A}+(\mathbf{B}+\mathbf{C})=(\mathbf{A}+\mathbf{B})+\mathbf{C}$;

(*c*) $\mathbf{A}+\mathbf{O}=\mathbf{O}+\mathbf{A}=\mathbf{A}$;

(*d*) $\mathbf{A}+(-1)\mathbf{A}=\mathbf{O}$;

(*e*) $\alpha(\mathbf{A}+\mathbf{B})=\alpha\mathbf{A}+\alpha\mathbf{B}$;

(*f*) $(\alpha+\beta)\mathbf{A}=\alpha\mathbf{A}+\beta\mathbf{A}$;

(*g*) $(\alpha\beta)\mathbf{A}=\alpha(\beta\mathbf{A})$;

(*h*) $1\cdot\mathbf{A}=\mathbf{A}$.

Łatwe dowody powyższych własności pozostawiamy Czytelnikowi. Tu zauważmy, że bezpośrednią konsekwencją własności (*a*) – (*d*) jest następujący wniosek.

$(K_{m\times n},+)$ – grupa przemienna

Wniosek 4.2.1. *System algebraiczny* $(K_{m\times n},+)$, *gdzie* $+$ *jest działaniem dodawania macierzy, jest grupą przemienną.* □

Definicja 4.2.6. Niech $\mathbf{A}=[a_{ij}]$ i $\mathbf{B}=[b_{ij}]$ będą odpowiednio macierzami należącymi do zbiorów $K_{m\times n}$ i $K_{n\times p}$. *Iloczynem macierzy* $\mathbf{A}$ i $\mathbf{B}$ nazywamy macierz $\mathbf{AB}=[c_{ij}]\in K_{m\times p}$, w której

Iloczyn macierzy

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\ldots+a_{in}b_{nj}=\sum_{k=1}^{n}a_{ik}b_{kj} \tag{4.3}$$

dla $i=1,\ldots,m$ oraz $j=1,\ldots,p$.

Zauważmy, że iloczyn $\mathbf{AB}$ jest określony, gdy ilość kolumn macierzy $\mathbf{A}$ jest równa ilości wierszy macierzy $\mathbf{B}$ i wtedy też macierz $\mathbf{AB}$ ma tyle wierszy co macierz $\mathbf{A}$, a kolumn tyle co macierz $\mathbf{B}$. Warto także zauważyć, że element c_{ij} iloczynu $\mathbf{AB}$ jest standardowym iloczynem skalarnym i-tego wiersza $\mathbf{a}_i$ macierzy $\mathbf{A}$ oraz j-tej kolumny $\mathbf{b}_j$ macierzy $\mathbf{B}$,

$$c_{ij}=\mathbf{a}_i\mathbf{b}_j,$$

czyli mamy

$$\mathbf{AB}=\begin{bmatrix} - & \mathbf{a}_1 & - \\ - & \mathbf{a}_2 & - \\ & \vdots & \\ - & \mathbf{a}_m & - \end{bmatrix}\begin{bmatrix} | & | & & | \\ \mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_p \\ | & | & & | \end{bmatrix}=\begin{bmatrix} \mathbf{a}_1\mathbf{b}_1 & \mathbf{a}_1\mathbf{b}_2 & \cdots & \mathbf{a}_1\mathbf{b}_p \\ \mathbf{a}_2\mathbf{b}_1 & \mathbf{a}_2\mathbf{b}_2 & \cdots & \mathbf{a}_2\mathbf{b}_p \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_m\mathbf{b}_1 & \mathbf{a}_m\mathbf{b}_2 & \cdots & \mathbf{a}_m\mathbf{b}_p \end{bmatrix}. \tag{4.4}$$

Przykład 69. Niech

$$\mathbf{A}=[a_{ij}]=\begin{bmatrix} 1 & 2 \\ -2 & 3 \\ 0 & 4 \end{bmatrix} \quad \text{i} \quad \mathbf{B}=[b_{ij}]=\begin{bmatrix} 2 & 3 & 0 \\ -1 & 1 & -2 \end{bmatrix}.$$

Ponieważ $\mathbf{A}\in R_{3\times 2}$ i $\mathbf{B}\in R_{2\times 3}$ i liczba kolumn macierzy $\mathbf{A}$ jest równa liczbie wierszy macierzy $\mathbf{B}$, więc iloczyn $\mathbf{AB}$ istnieje i jest macierzą wymiaru 3×3,

$$\mathbf{AB}=\begin{bmatrix} 1 & 2 \\ -2 & 3 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 2 & 3 & 0 \\ -1 & 1 & -2 \end{bmatrix}=\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix},$$

gdzie

$$c_{ij}=\mathbf{a}_i\mathbf{b}_j=\sum_{k=1}^{2}a_{ik}b_{kj}=a_{i1}b_{1j}+a_{i2}b_{2j} \quad \text{dla} \quad 1\leqslant i,\, j\leqslant 3.$$

Łatwo znajdujemy

$c_{11} = \mathbf{a}_1\mathbf{b}_1 = 1 \cdot 2 + 2 \cdot (-1) = 0,$

$c_{12} = \mathbf{a}_1\mathbf{b}_2 = 1 \cdot 3 + 2 \cdot 1 = 5,$

$c_{13} = \mathbf{a}_1\mathbf{b}_3 = 1 \cdot 0 + 2 \cdot (-2) = -4,$

$c_{21} = \mathbf{a}_2\mathbf{b}_1 = (-2) \cdot 2 + 3 \cdot (-1) = -7,$

$c_{22} = \mathbf{a}_2\mathbf{b}_2 = (-2) \cdot 3 + 3 \cdot 1 = -3,$

$c_{23} = \mathbf{a}_2\mathbf{b}_3 = (-2) \cdot 0 + 3 \cdot (-2) = -6,$

$c_{31} = \mathbf{a}_3\mathbf{b}_1 = 0 \cdot 2 + 4 \cdot (-1) = -4,$

$c_{32} = \mathbf{a}_3\mathbf{b}_2 = 0 \cdot 3 + 4 \cdot 1 = 4,$

$c_{33} = \mathbf{a}_3\mathbf{b}_3 = 0 \cdot 0 + 4 \cdot (-2) = -8,$

więc

$$\mathbf{AB} = \begin{bmatrix} 0 & 5 & -4 \\ -7 & -3 & -6 \\ -4 & 4 & -8 \end{bmatrix}.$$

Iloczyn **BA** także istnieje, ale jest macierzą wymiaru 2×2,

$$\mathbf{BA} = \begin{bmatrix} 2 & 3 & 0 \\ -1 & 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & 3 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} -4 & 13 \\ -3 & -7 \end{bmatrix} \neq \mathbf{AB}.$$

Przykład ten ilustruje, że **mnożenie macierzy nie jest przemienne**. Niżej mamy podobne przykłady:

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

$$\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \begin{bmatrix} 2 & 5 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 5 & 1 \\ 0 & 0 & 0 \\ 4 & 10 & 2 \end{bmatrix} \neq \begin{bmatrix} 4 \end{bmatrix} = \begin{bmatrix} 2 & 5 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 2 \\ -2 & -4 \end{bmatrix} \begin{bmatrix} 2 & -4 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 10 & 20 \\ -5 & -10 \end{bmatrix} = \begin{bmatrix} 2 & -4 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & -4 \end{bmatrix}.$$

Warto zauważyć, że jeśli macierz $\mathbf{A} = [a_{ij}] \in K_{m \times n}$ pomnoży się przez macierz jednokolumnową $\mathbf{x} = [\, x_1\, x_2 \ldots x_n \,]^T \in K_{n \times 1}$, to wobec (4.4) mamy

$$\begin{aligned} \mathbf{Ax} &= \begin{bmatrix} - & \mathbf{a}_{1*} & - \\ - & \mathbf{a}_{2*} & - \\ & \vdots & \\ - & \mathbf{a}_{m*} & - \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{a}_{1*}\mathbf{x} \\ \mathbf{a}_{2*}\mathbf{x} \\ \vdots \\ \mathbf{a}_{m*}\mathbf{x} \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{bmatrix} \\ &= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \\ &= x_1 \begin{bmatrix} | \\ \mathbf{a}_{*1} \\ | \end{bmatrix} + x_2 \begin{bmatrix} | \\ \mathbf{a}_{*2} \\ | \end{bmatrix} + \ldots + x_n \begin{bmatrix} | \\ \mathbf{a}_{*n} \\ | \end{bmatrix}, \end{aligned}$$

więc także

$$\mathbf{Ax} = x_1\mathbf{a}_{*1} + x_2\mathbf{a}_{*2} + \ldots + x_n\mathbf{a}_{*n}. \tag{4.5}$$

Oznacza to, że iloczyn **Ax** jest kombinacją liniową kolumn macierzy **A** ze współczynnikami, które są kolejnymi elementami kolumny **x**.

Przykład 70. Z (4.5) wynika, że mamy

$$\begin{bmatrix} 1 & 3 & 0 & 1 \\ 2 & -4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 3 \\ -2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 3 \\ -4 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 5 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ 6 \end{bmatrix} = \begin{bmatrix} -3 \\ 11 \end{bmatrix}.$$

Podobnie, jeśli macierz $\mathbf{B} = [b_{ij}] \in K_{n\times p}$ przemnażamy przez macierz jednowierszową $\mathbf{y} = \begin{bmatrix} y_1 \ y_2 \ \dots \ y_n \end{bmatrix} \in K_{1\times n}$, to mamy

$$\begin{aligned} \mathbf{yB} &= \begin{bmatrix} \mathbf{yb}_{*1} \ \mathbf{yb}_{*2} \ \dots \ \mathbf{yb}_{*p} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^{n} y_i b_{i1} \ \sum_{i=1}^{n} y_i b_{i2} \ \dots \ \sum_{i=1}^{n} y_i b_{ip} \end{bmatrix} \\ &= \sum_{i=1}^{n} y_i \begin{bmatrix} b_{i1} \ b_{i2} \ \dots \ b_{ip} \end{bmatrix}, \end{aligned}$$

czyli

$$\mathbf{yB} = \begin{bmatrix} y_1 \ y_2 \ \dots \ y_n \end{bmatrix} \begin{bmatrix} - \ \mathbf{b}_{1*} \ - \\ - \ \mathbf{b}_{2*} \ - \\ \vdots \\ - \ \mathbf{b}_{n*} \ - \end{bmatrix} = \sum_{i=1}^{n} y_i \mathbf{b}_{i*}. \tag{4.6}$$

Zatem iloczyn $\mathbf{yB}$ jest kombinacją liniową wierszy macierzy $\mathbf{B}$ ze współczynnikami będącymi kolejnymi elementami macierzy $\mathbf{y}$. Teraz z (4.4) i (4.5) mamy

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} \mathbf{a}_{1*}\mathbf{b}_{*1} & \mathbf{a}_{1*}\mathbf{b}_{*2} & \cdots & \mathbf{a}_{1*}\mathbf{b}_{*p} \\ \mathbf{a}_{2*}\mathbf{b}_{*1} & \mathbf{a}_{2*}\mathbf{b}_{*2} & \cdots & \mathbf{a}_{2*}\mathbf{b}_{*p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_{m*}\mathbf{b}_{*1} & \mathbf{a}_{m*}\mathbf{b}_{*2} & \cdots & \mathbf{a}_{m*}\mathbf{b}_{*p} \end{bmatrix} \\ &= \begin{bmatrix} \begin{bmatrix} \mathbf{a}_{1*}\mathbf{b}_{*1} \\ \mathbf{a}_{2*}\mathbf{b}_{*1} \\ \vdots \\ \mathbf{a}_{m*}\mathbf{b}_{*1} \end{bmatrix} \begin{bmatrix} \mathbf{a}_{1*}\mathbf{b}_{*2} \\ \mathbf{a}_{2*}\mathbf{b}_{*2} \\ \vdots \\ \mathbf{a}_{m*}\mathbf{b}_{*2} \end{bmatrix} \cdots \begin{bmatrix} \mathbf{a}_{1*}\mathbf{b}_{*p} \\ \mathbf{a}_{2*}\mathbf{b}_{*p} \\ \vdots \\ \mathbf{a}_{m*}\mathbf{b}_{*p} \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} | & | & & | \\ \mathbf{Ab}_{*1} & \mathbf{Ab}_{*2} & \dots & \mathbf{Ab}_{*p} \\ | & | & & | \end{bmatrix}. \end{aligned}$$

Analogicznie z (4.4) i (4.6) otrzymujemy

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} \mathbf{a}_{1*}\mathbf{b}_{*1} & \mathbf{a}_{1*}\mathbf{b}_{*2} & \cdots & \mathbf{a}_{1*}\mathbf{b}_{*p} \\ \mathbf{a}_{2*}\mathbf{b}_{*1} & \mathbf{a}_{2*}\mathbf{b}_{*2} & \cdots & \mathbf{a}_{2*}\mathbf{b}_{*p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_{m*}\mathbf{b}_{*1} & \mathbf{a}_{m*}\mathbf{b}_{*2} & \cdots & \mathbf{a}_{m*}\mathbf{b}_{*p} \end{bmatrix} \\ &= \begin{bmatrix} [\, \mathbf{a}_{1*}\mathbf{b}_{*1} \ \mathbf{a}_{1*}\mathbf{b}_{*2} \ \cdots \ \mathbf{a}_{1*}\mathbf{b}_{*p} \,] \\ [\, \mathbf{a}_{2*}\mathbf{b}_{*1} \ \mathbf{a}_{2*}\mathbf{b}_{*2} \ \cdots \ \mathbf{a}_{2*}\mathbf{b}_{*p} \,] \\ \vdots \\ [\mathbf{a}_{m*}\mathbf{b}_{*1} \ \mathbf{a}_{m*}\mathbf{b}_{*2} \ \cdots \ \mathbf{a}_{m*}\mathbf{b}_{*p}] \end{bmatrix} = \begin{bmatrix} - \ \mathbf{a}_{1*}\mathbf{B} \ - \\ - \ \mathbf{a}_{2*}\mathbf{B} \ - \\ \vdots \\ - \ \mathbf{a}_{m*}\mathbf{B} \ - \end{bmatrix}. \end{aligned}$$

Dlatego mamy następujące własności iloczynu macierzy.

Twierdzenie 4.2.2. *Jeśli* $\mathbf{a}_1, \ldots, \mathbf{a}_m$ *są kolejnymi wierszami macierzy* $\mathbf{A} \in K_{m\times n}$*, a* $\mathbf{b}_1, \ldots, \mathbf{b}_p$ *są kolejnymi kolumnami macierzy* $\mathbf{B} \in K_{n\times p}$*, to iloczyn* $\mathbf{AB}$ *jest macierzą, której kolejnymi kolumnami oraz wierszami są odpowiednio* $\mathbf{Ab}_1, \ldots, \mathbf{Ab}_p$ *oraz* $\mathbf{a}_1\mathbf{B}, \ldots, \mathbf{a}_m\mathbf{B}$*, czyli*

$$\mathbf{AB} = \mathbf{A}\begin{bmatrix} | & | & & | \\ \mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \mathbf{Ab}_1 & \mathbf{Ab}_2 & \ldots & \mathbf{Ab}_p \\ | & | & & | \end{bmatrix} \tag{4.7}$$

i

$$\mathbf{AB} = \begin{bmatrix} - & \mathbf{a}_1 & - \\ - & \mathbf{a}_2 & - \\ & \vdots & \\ - & \mathbf{a}_m & - \end{bmatrix} \cdot \mathbf{B} = \begin{bmatrix} - & \mathbf{a}_1\mathbf{B} & - \\ - & \mathbf{a}_2\mathbf{B} & - \\ & \vdots & \\ - & \mathbf{a}_m\mathbf{B} & - \end{bmatrix}. \tag{4.8}$$

Dodatkowo, i*-ta kolumna macierzy* $\mathbf{AB}$ *jest kombinacją liniową kolumn macierzy* $\mathbf{A}$ *o współczynnikach z* i*-tej kolumny macierzy* $\mathbf{B}$ $(1 \leqslant i \leqslant p)$*, a* j*-ty wiersz macierzy* $\mathbf{AB}$ *jest kombinacją liniową wierszy macierzy* $\mathbf{B}$ *o współczynnikach z* j*-tego wiersza macierzy* $\mathbf{A}$ $(1 \leqslant j \leqslant m)$*.* □

Przykład 71. Dla macierzy $\mathbf{A}$ i $\mathbf{B}$ z przykładu 69 mamy

$$\begin{aligned}
\mathbf{AB} &= \begin{bmatrix} 1 & 2 \\ -2 & 3 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 2 & 3 & 0 \\ -1 & 1 & -2 \end{bmatrix} = \begin{bmatrix} | & | & | \\ \mathbf{Ab}_1 & \mathbf{Ab}_2 & \mathbf{Ab}_3 \\ | & | & | \end{bmatrix} \\
&= \left[2\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \;\; 3\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \;\; 0\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} - 2\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \right] = \begin{bmatrix} 0 & 5 & -4 \\ -7 & -3 & -6 \\ -4 & 4 & -8 \end{bmatrix}
\end{aligned}$$

i jednocześnie

$$\begin{aligned}
\mathbf{AB} &= \begin{bmatrix} 1 & 2 \\ -2 & 3 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 2 & 3 & 0 \\ -1 & 1 & -2 \end{bmatrix} = \begin{bmatrix} - & \mathbf{a}_1\mathbf{B} & - \\ - & \mathbf{a}_2\mathbf{B} & - \\ - & \mathbf{a}_3\mathbf{B} & - \end{bmatrix} \\
&= \begin{bmatrix} 1\,[2\;3\;0] + 2\,[-1\;1-2] \\ -2\,[2\;3\;0] + 3\,[-1\;1-2] \\ 0\,[2\;3\;0] + 4\,[-1\;1-2] \end{bmatrix} = \begin{bmatrix} 0 & 5 & -4 \\ -7 & -3 & -6 \\ -4 & 4 & -8 \end{bmatrix}.
\end{aligned}$$

Następne twierdzenie daje nam jeszcze jedną możliwość rozumienia (i wyznaczania) iloczynu macierzy.

Twierdzenie 4.2.3. *Jeśli* $\mathbf{a}_1, \ldots, \mathbf{a}_n$ *są kolejnymi kolumnami macierzy* $\mathbf{A} = [a_{ij}] \in K_{m\times n}$*, a* $\mathbf{b}_1, \ldots, \mathbf{b}_n$ *kolejnymi wierszami macierzy* $\mathbf{B} = [b_{ij}] \in K_{n\times p}$*, to*

$$\mathbf{AB} = \begin{bmatrix} | & & | \\ \mathbf{a}_1 & \ldots & \mathbf{a}_n \\ | & & | \end{bmatrix}\begin{bmatrix} - \mathbf{b}_1 - \\ \vdots \\ - \mathbf{b}_n - \end{bmatrix} = \mathbf{a}_1\mathbf{b}_1 + \ldots + \mathbf{a}_n\mathbf{b}_n. \tag{4.9}$$

Dowód. Równość (4.9) jest konsekwencją tego, że dla liczb naturalnych $i = 1, \ldots, m$ oraz $j = 1, \ldots, p$ mamy

$$\begin{aligned}
\left(\sum_{k=1}^{n} \mathbf{a}_k\mathbf{b}_k\right)_{ij} &= \sum_{k=1}^{n} (\mathbf{a}_k\mathbf{b}_k)_{ij} = \sum_{k=1}^{n}\left(\begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix}\begin{bmatrix} b_{k1} & b_{k2} & \ldots & b_{kp} \end{bmatrix}\right)_{ij} \\
&= \sum_{k=1}^{n} a_{ik}b_{kj} = (\mathbf{AB})_{ij},
\end{aligned}$$

gdzie przez $(\mathbf{X})_{ij}$ oznaczyliśmy element znajdujący w i-tym wierszu oraz j-tej kolumnie macierzy $\mathbf{X}$. □

Przykład 72. Korzystając z ostatniego twierdzenia, obliczyć iloczyn $\mathbf{AB}$, gdy

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 0 \\ 2 & 1 & -1 \end{bmatrix} \quad \text{i} \quad \mathbf{B} = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 3 & 4 \\ 5 & 0 & -1 \end{bmatrix}.$$

Mamy

$$\begin{aligned}
\mathbf{AB} &= \begin{bmatrix} | & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} -\,\mathbf{b}_1 - \\ -\,\mathbf{b}_2 - \\ -\,\mathbf{b}_3 \end{bmatrix} = \mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \mathbf{a}_3\mathbf{b}_3 \\
&= \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 3 & 4 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix} \begin{bmatrix} 5 & 0 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 1 & 2 \\ 0 & 2 & 4 \end{bmatrix} + \begin{bmatrix} -3 & 9 & 12 \\ -1 & 3 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ -5 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -3 & 10 & 14 \\ -6 & 5 & 9 \end{bmatrix}.
\end{aligned}$$

W następującym twierdzeniu przedstawiamy pewne własności mnożenia macierzy.

Twierdzenie 4.2.4. *Dla skalara α i macierzy $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ oraz macierzy jednostkowej $\mathbf{I}$ zachodzą następujące równości (pod warunkiem, że występujące w nich działania są wykonalne):*

(*a*) $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$;

(*b*) $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$;

(*c*) $\mathbf{IA} = \mathbf{A}$ *i* $\mathbf{BI} = \mathbf{B}$;

(*d*) $(\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B}) = \alpha(\mathbf{AB})$;

(*e*) $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$.

Dowód. Przykładowo udowodnimy (*a*) i (*e*). Pozostałe równości dowodzi się podobnie. Załóżmy, że macierze $\mathbf{A}$, $\mathbf{B}$ i $\mathbf{C}$ są odpowiednio macierzami wymiaru $m \times n$, $n \times p$ oraz $n \times p$ i o współczynnikach z ciała K. Ponieważ $\mathbf{A}(\mathbf{B}+\mathbf{C})$ i $\mathbf{AB}+\mathbf{AC}$ są macierzami wymiaru $m \times p$, więc dla dowodu równości (*a*) wystarczy udowodnić, że $\big(\mathbf{A}(\mathbf{B}+\mathbf{C})\big)_{ij} = \big(\mathbf{AB}+\mathbf{AC}\big)_{ij}$ dla $i = 1, \ldots, m$ oraz $j = 1, \ldots, p$. Tak jest istotnie, bo z definicji iloczynu i sumy macierzy (zob. definicje (4.2.6) i (4.2.2)) oraz z rozdzielności mnożenia względem dodawania w ciele K mamy

$$\begin{aligned}
\big(\mathbf{A}(\mathbf{B}+\mathbf{C})\big)_{ij} &= \sum_{k=1}^{n} (\mathbf{A})_{ik}(\mathbf{B}+\mathbf{C})_{kj} = \sum_{k=1}^{n} (\mathbf{A})_{ik}\big((\mathbf{B})_{kj} + (\mathbf{C})_{kj}\big) \\
&= \sum_{k=1}^{n} (\mathbf{A})_{ik}(\mathbf{B})_{kj} + \sum_{k=1}^{n} (\mathbf{A})_{ik}(\mathbf{C})_{kj} \\
&= (\mathbf{AB})_{ij} + (\mathbf{AC})_{ij} = \big(\mathbf{AB}+\mathbf{AC}\big)_{ij}.
\end{aligned}$$

Dla dowodu równości (*e*) zakładamy, że macierze $\mathbf{A}$, $\mathbf{B}$ i $\mathbf{C}$ są odpowiednio wymiaru $m \times n$, $n \times p$ oraz $p \times s$. Tym razem wystarczy wykazać, że $\big(\mathbf{A}(\mathbf{BC})\big)_{ij} = \big((\mathbf{AB})\mathbf{C}\big)_{ij}$ dla $i = 1, \ldots, m$ oraz $j = 1, \ldots, s$. To zaś wynika z definicji iloczynu macierzy, bo mamy

$$\begin{aligned}
\big(\mathbf{A}(\mathbf{BC})\big)_{ij} &= \sum_{k=1}^{n} (\mathbf{A})_{ik}(\mathbf{BC})_{kj} = \sum_{k=1}^{n} (\mathbf{A})_{ik}\Big(\sum_{l=1}^{p} (\mathbf{B})_{kl}(\mathbf{C})_{lj}\Big) \\
&= \sum_{l=1}^{p} \Big(\sum_{k=1}^{n} (\mathbf{A})_{ik}(\mathbf{B})_{kl}\Big)(\mathbf{C})_{lj} = \sum_{l=1}^{p} (\mathbf{AB})_{il}(\mathbf{C})_{lj} = \big((\mathbf{AB})\mathbf{C}\big)_{ij}.
\end{aligned}$$

Równość (*e*) (tak samo jak i pozostałe równości) można także udowodnić w inny sposób. Przyjmijmy, że $\mathbf{A} \in K_{m\times n}$, $\mathbf{B} = [\,\mathbf{b}_1 \ldots \mathbf{b}_p\,] \in K_{n\times p}$ i $\mathbf{C} = [\,\mathbf{c}_1 \ldots \mathbf{c}_s\,] = [c_{ij}] \in K_{p\times s}$. Zauważmy najpierw, że dla iloczynu macierzy $\mathbf{A}$, $\mathbf{B}$ i macierzy jednokolumnowej $\mathbf{c}_i$ (która jest i-tą kolumną macierzy $\mathbf{C}$) wobec (4.5), (*a*) oraz (*d*) mamy

$$\begin{aligned}\mathbf{A}(\mathbf{B}\mathbf{c}_i) &= \mathbf{A}\left(\left[\ \mathbf{b}_1 \dots \mathbf{b}_p\ \right]\begin{bmatrix} c_{1i} \\ \vdots \\ c_{pi}\end{bmatrix}\right) = \mathbf{A}(c_{1i}\mathbf{b}_1 + \dots + c_{pi}\mathbf{b}_p) \\ &= c_{1i}(\mathbf{A}\mathbf{b}_1) + \dots + c_{pi}(\mathbf{A}\mathbf{b}_p) = [\ \mathbf{A}\mathbf{b}_1 \dots \mathbf{A}\mathbf{b}_p\ ]\mathbf{c}_i = (\mathbf{AB})\mathbf{c}_i.\end{aligned}$$

Zatem ogólnie mamy

$$\begin{aligned}\mathbf{A}(\mathbf{BC}) &= \mathbf{A}\big(\mathbf{B}[\ \mathbf{c}_1 \dots \mathbf{c}_s\ ]\big) = \mathbf{A}\big[\ \mathbf{B}\mathbf{c}_1 \dots \mathbf{B}\mathbf{c}_s\ \big] = \big[\ \mathbf{A}(\mathbf{B}\mathbf{c}_1) \dots \mathbf{A}(\mathbf{B}\mathbf{c}_s)\ \big] \\ &= \big[\ (\mathbf{AB})\mathbf{c}_1 \dots (\mathbf{AB})\mathbf{c}_s\ \big] = (\mathbf{AB})[\ \mathbf{c}_1 \dots \mathbf{c}_s\ ] = (\mathbf{AB})\mathbf{C}. \ \square\end{aligned}$$

Definicja 4.2.7. *Macierzą transponowaną* macierzy $\mathbf{A} = [a_{ij}] \in K_{m\times n}$ nazywamy macierz $\mathbf{A}^T = [b_{ij}] \in K_{n\times m}$, w której

$\mathbf{A}^T$ – transpozycja macierzy $\mathbf{A}$
$(\mathbf{A}^T)_{ij} = (\mathbf{A})_{ji}$

$$b_{ij} = a_{ji}$$

dla $i = 1, \dots, n$ oraz $j = 1, \dots, m$.

Obrazowo, $\mathbf{A}^T$ powstaje z macierzy $\mathbf{A}$ przez zamianę wierszy na kolumny i kolumn na wiersze; kolejne wiersze macierzy $\mathbf{A}$ stają się kolejnymi kolumnami macierzy $\mathbf{A}^T$.

Przykład 73. Macierzami transponowanymi macierzy

$$\mathbf{A} = \begin{bmatrix} 2 & -1 \\ 3 & 4 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 2 & -3 \\ -4 & 0 & 8 \end{bmatrix} \quad \text{i} \quad \mathbf{C} = \begin{bmatrix} 1 & -2 & 0 & 4 \end{bmatrix}$$

są odpowiednio

$$\mathbf{A}^T = \begin{bmatrix} 2 & 3 \\ -1 & 4 \end{bmatrix}, \quad \mathbf{B}^T = \begin{bmatrix} 1 & -4 \\ 2 & 0 \\ -3 & 8 \end{bmatrix} \quad \text{i} \quad \mathbf{C}^T = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 4 \end{bmatrix}.$$

Definicja 4.2.8. Macierz kwadratową $\mathbf{A}$ nazywamy *macierzą symetryczną*, gdy

Macierz symetryczna

$$\mathbf{A}^T = \mathbf{A}, \tag{4.10}$$

a jest ona *skośnie symetryczna*, gdy

$$\mathbf{A}^T = -\mathbf{A}. \tag{4.11}$$

Macierz skośnie symetryczna

Równoważnie, macierz kwadratowa $\mathbf{A} = [a_{ij}]$ stopnia n jest symetryczna, gdy

$$a_{ij} = a_{ji}, \tag{4.12}$$

a jest ona skośnie symetryczna, gdy

$$a_{ij} = -a_{ji} \tag{4.13}$$

dla każdych indeksów i, $j = 1, \dots, n$.

Elementy głównej przekątnej macierzy skośnie symetrycznej $\mathbf{A} = [a_{ij}]$ są równe zeru, $a_{11} = a_{22} = \dots = a_{nn} = 0$, bo wobec (4.13) jest $a_{ii} = -a_{ii}$ dla $i = 1, \dots, n$.

Przykład 74. Spośród macierzy

$$\begin{bmatrix} 0 & 1 & 2 \\ 1 & 3 & -1 \\ 2 & -1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix} \quad \text{i} \quad \begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & -1 \\ -2 & 1 & 2 \end{bmatrix},$$

pierwsza jest symetryczna, druga – skośnie symetryczna, a trzecia nie jest ani symetryczna, ani skośnie symetryczna.

Podstawowe własności transpozycji macierzy przedstawiamy w kolejnym twierdzeniu.

Twierdzenie 4.2.5. *Dla każdych macierzy* $\mathbf{A}, \mathbf{B} \in K_{m\times n}$ *i* $\mathbf{C} \in K_{n\times p}$ *oraz skalara* $\alpha \in K$ *jest:*

(*a*) $(\mathbf{A}^T)^T = \mathbf{A}$;

(*b*) $(\alpha\mathbf{A})^T = \alpha\mathbf{A}^T$;

(*c*) $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$;

(*d*) $(\mathbf{A}\mathbf{C})^T = \mathbf{C}^T\mathbf{A}^T$.

Dowód. Równości (*a*) – (*c*) są oczywiste. Dla dowodu równości (*d*) wystarczy wykazać, że $\left((\mathbf{AC})^T\right)_{ij} = (\mathbf{C}^T\mathbf{A}^T)_{ij}$ dla $i = 1,\ldots,p$ oraz $j = 1,\ldots,m$. Z definicji transpozycji i iloczynu macierzy mamy

$$\begin{aligned} \left((\mathbf{AC})^T\right)_{ij} &= (\mathbf{AC})_{ji} = \sum_{k=1}^{n} (\mathbf{A})_{jk}(\mathbf{C})_{ki} \\ &= \sum_{k=1}^{n} (\mathbf{A}^T)_{kj}(\mathbf{C}^T)_{ik} = \sum_{k=1}^{n} (\mathbf{C}^T)_{ik}(\mathbf{A}^T)_{kj} = (\mathbf{C}^T\mathbf{A}^T)_{ij} \end{aligned}$$

i to kończy dowód twierdzenia. $\square$

Przykład 75. Jeśli $\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 5 & 4 \end{bmatrix}$ i $\mathbf{B} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$, to mamy

$$\begin{aligned} (\mathbf{AB})^T &= \left(\begin{bmatrix} 1 & 2 & 1 \\ 3 & 5 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \right)^T = \begin{bmatrix} 7 \\ 23 \end{bmatrix}^T \\ &= \begin{bmatrix} 7 & 23 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 5 \\ 1 & 4 \end{bmatrix} = \mathbf{B}^T\mathbf{A}^T. \end{aligned}$$

4.3. Macierz odwrotna

Definicja 4.3.1. Mówimy, że macierz kwadratowa $\mathbf{A} \in K_{n\times n}$ jest *odwracalna*, jeśli istnieje macierz $\mathbf{B} \in K_{n\times n}$ taka, że

Macierz odwracalna

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n. \tag{4.14}$$

Twierdzenie 4.3.1. *Jeśli* $\mathbf{A}$, $\mathbf{B}$ *i* $\mathbf{C}$ *są macierzami kwadratowymi stopnia* n *takimi, że* $\mathbf{AB} = \mathbf{CA} = \mathbf{I}_n$, *to* $\mathbf{B} = \mathbf{C}$.

Dowód. Jeśli $\mathbf{AB} = \mathbf{I}_n$ i $\mathbf{CA} = \mathbf{I}_n$, to z własności macierzy jednostkowej i z łączności mnożenia macierzy otrzymujemy

$$\mathbf{B} = \mathbf{I}_n\mathbf{B} = (\mathbf{CA})\mathbf{B} = \mathbf{C}(\mathbf{AB}) = \mathbf{CI}_n = \mathbf{C}. \ \square$$

Z definicji 4.3.1 i twierdzenia 4.3.1 wynika, że jeśli macierz $\mathbf{A} \in K_{n\times n}$ jest odwracalna, to istnieje dokładnie jedna macierz $\mathbf{B} \in K_{n\times n}$ taka, że $\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n$. W takim przypadku mówimy, że $\mathbf{B}$ jest *macierzą odwrotną* do macierzy $\mathbf{A}$ i oznaczamy ją symbolem $\mathbf{A}^{-1}$.

$\mathbf{A}^{-1}$ – macierz odwrotna

Przykład 76. Jeśli $\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix}$ i $\mathbf{B} = \begin{bmatrix} 5 & -2 \\ -7 & 3 \end{bmatrix}$, to mamy

$$\mathbf{AB} = \begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix}\begin{bmatrix} 5 & -2 \\ -7 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

i

$$\mathbf{BA} = \begin{bmatrix} 5 & -2 \\ -7 & 3 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Dlatego macierz $\mathbf{A}$ jest odwracalna i jej macierzą odwrotną jest $\mathbf{B}$, czyli mamy $\mathbf{A}^{-1} = \mathbf{B}$. Z tych samych powodów macierz $\mathbf{B}$ jest odwracalna i $\mathbf{B}^{-1} = \mathbf{A}$.

Przykład 77. Macierz $\mathbf{C} = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}$ nie jest odwracalna, bo dla każdej macierzy $\mathbf{D} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ jest

$$\mathbf{DC} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} a+2b & 0 \\ c+2d & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}_2.$$

Warunki konieczne i dostateczne odwracalności macierzy oraz sposoby wyznaczania macierzy odwrotnej poznamy w kolejnych rozdziałach. Teraz przedstawiamy trzy podstawowe własności macierzy odwracalnych i ich macierzy odwrotnych.

Twierdzenie 4.3.2. *Jeśli* $\mathbf{A}$ *i* $\mathbf{B}$ *są macierzami odwracalnymi ze zbioru* $K_{n\times n}$*, to:*

(a) *macierz* $\mathbf{A}^{-1}$ *jest odwracalna i* $\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$*;*

(b) *macierz* $\mathbf{A}^{T}$ *jest odwracalna i* $\left(\mathbf{A}^{T}\right)^{-1} = \left(\mathbf{A}^{-1}\right)^{T}$*;*

(c) *macierz* $\mathbf{AB}$ *jest odwracalna i* $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

Dowód. Dla dowodu (a) musimy wskazać macierz $\mathbf{X} \in K_{n\times n}$, dla której zachodzą równości $\mathbf{A}^{-1}\mathbf{X} = \mathbf{X}\mathbf{A}^{-1} = \mathbf{I}_n$. Ponieważ wiemy już, że te równości zachodzą dla $\mathbf{X} = \mathbf{A}$, więc macierz $\mathbf{A}^{-1}$ jest odwracalna i $\mathbf{A}$ jest jej macierzą odwrotną.

Z odwracalności macierzy $\mathbf{A}$ i z własności transpozycji macierzy (zob. tw. 4.2.5 (d)) mamy

$$\mathbf{A}^{T}\left(\mathbf{A}^{-1}\right)^{T} = \left(\mathbf{A}^{-1}\mathbf{A}\right)^{T} = \mathbf{I}_n^{T} = \mathbf{I}_n$$

oraz

$$\left(\mathbf{A}^{-1}\right)^{T}\mathbf{A}^{T} = \left(\mathbf{A}\mathbf{A}^{-1}\right)^{T} = \mathbf{I}_n^{T} = \mathbf{I}_n$$

i stąd wynika (b).

Z odwracalności macierzy $\mathbf{A}$ i $\mathbf{B}$ oraz z łączności iloczynu macierzy (tw. 4.2.4 (e)) mamy

$$(\mathbf{AB})\left(\mathbf{B}^{-1}\mathbf{A}^{-1}\right) = \mathbf{A}\left(\mathbf{B}\mathbf{B}^{-1}\right)\mathbf{A}^{-1} = \mathbf{A}\mathbf{I}_n\mathbf{A}^{-1} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n$$

i

$$\left(\mathbf{B}^{-1}\mathbf{A}^{-1}\right)(\mathbf{AB}) = \mathbf{B}^{-1}\left(\mathbf{A}^{-1}\mathbf{A}\right)\mathbf{B} = \mathbf{B}^{-1}\mathbf{I}_n\mathbf{B} = \mathbf{B}^{-1}\mathbf{A} = \mathbf{I}_n,$$

a to dowodzi (c). □

Potęga macierzy

Definicja 4.3.2. Potęgę macierzy kwadratowej $\mathbf{A} \in K_{n\times n}$ definiujemy przyjmując, że

$$\mathbf{A}^0 = \mathbf{I}_n \quad \text{i} \quad \mathbf{A}^{k+1} = \mathbf{A}^k\mathbf{A}$$

dla $k \geqslant 0$. Przykładowo, $\mathbf{A}^1 = \mathbf{A}^0\mathbf{A} = \mathbf{I}_n\mathbf{A} = \mathbf{A}$ i $\mathbf{A}^2 = \mathbf{A}^1\mathbf{A} = \mathbf{A}\mathbf{A}$.

Przykład 78. Dana jest macierz $\mathbf{A} = \begin{bmatrix} 4 & -4 \\ 1 & 0 \end{bmatrix}$. (a) Wykazać, że $\mathbf{A}^2 - 4\mathbf{A} + 4\mathbf{I}_2 = \mathbf{0}$. Wywnioskować stąd, że macierz $\mathbf{A}$ jest odwracalna i wyznaczyć $\mathbf{A}^{-1}$. (b) Indukcyjnie wykazać, że dla każdej liczby naturalnej n jest

$$\mathbf{A}^n = \begin{bmatrix} (n+1)2^n & -n2^{n+1} \\ n2^{n-1} & (1-n)2^n \end{bmatrix}.$$

(a) Łatwo sprawdza się, że $\mathbf{A}^2 - 4\mathbf{A} + 4\mathbf{I}_2 = \mathbf{0}$. Stąd zaś wyliczając $\mathbf{I}_2$, otrzymujemy

$$\mathbf{A}\left(-\frac{1}{4}\mathbf{A} + \mathbf{I}_2\right) = \left(-\frac{1}{4}\mathbf{A} + \mathbf{I}_2\right)\mathbf{A} = \mathbf{I}_2.$$

To dowodzi, że macierz $\mathbf{A}$ jest odwracalna i $\mathbf{A}^{-1} = -\frac{1}{4}\mathbf{A} + \mathbf{I}_2 = \begin{bmatrix} 0 & 1 \\ -\frac{1}{4} & 1 \end{bmatrix}$.

(b) Dowodzona równość jest prawdziwa dla $n = 1$, bo mamy

$$\mathbf{A}^1 = \mathbf{A} = \begin{bmatrix} 4 & -4 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} (1+1)2^1 & -1\cdot 2^{1+1} \\ 1\cdot 2^{1-1} & (1-1)2^1 \end{bmatrix}.$$

Załóżmy teraz, że dla liczby naturalnej $n \geqslant 1$ jest $\mathbf{A}^n = \begin{bmatrix} (n+1)2^n & -n2^{n+1} \\ n2^{n-1} & (1-n)2^n \end{bmatrix}$. Wtedy dla macierzy $\mathbf{A}^{n+1} = \mathbf{A}^n\mathbf{A}$ mamy

$$\begin{aligned} \mathbf{A}^{n+1} &= \begin{bmatrix} (n+1)2^n & -n2^{n+1} \\ n2^{n-1} & (1-n)2^n \end{bmatrix}\begin{bmatrix} 4 & -4 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 4(n+1)2^n - n2^{n+1} & -4(n+1)2^n \\ 4n2^{n-1} + (1-n)2^n & -4n2^{n-1} \end{bmatrix} \\ &= \begin{bmatrix} (n+2)2^{n+1} & -(n+1)2^{n+2} \\ (n+1)2^n & (1-(n+1))2^{n+1} \end{bmatrix}. \end{aligned}$$

Stąd i z twierdzenia o indukcji wynika, że $\mathbf{A}^n = \begin{bmatrix} (n+1)2^n & -n2^{n+1} \\ n2^{n-1} & (1-n)2^n \end{bmatrix}$ dla każdej liczby naturalnej n.

4.4. Ślad macierzy kwadratowej

Definicja 4.4.1. Jeśli $\mathbf{A}$ jest macierzą kwadratową, $\mathbf{A} = [a_{ij}] \in K_{n\times n}$, to sumę elementów należących do jej głównej przekątnej nazywamy *śladem macierzy* $\mathbf{A}$ i oznaczamy przez $\operatorname{tr}(\mathbf{A})$, czyli

Ślad macierzy

$$\operatorname{tr}(\mathbf{A}) = a_{11} + a_{22} + \ldots + a_{nn}.$$

Przykład 79. Jeśli

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 3 \\ -1 & 0 & 5 \end{bmatrix} \quad \text{i} \quad \mathbf{B} = \begin{bmatrix} 2 & 2 \\ -1 & 2 \\ 1 & 3 \end{bmatrix},$$

to macierze **AB** i **BA** są kwadratowe,

$$\mathbf{AB} = \begin{bmatrix} 6 & 15 \\ 3 & 13 \end{bmatrix} \quad \text{i} \quad \mathbf{BA} = \begin{bmatrix} 2 & 2 & 16 \\ -4 & -1 & 7 \\ -1 & 1 & 18 \end{bmatrix},$$

i mają one równe ślady, $\mathrm{tr}\,(\mathbf{AB}) = 19 = \mathrm{tr}\,(\mathbf{BA})$. Równość ta nie jest przypadkowa. Jest ona ilustracją bardzo ważnej własności śladu iloczynu macierzy i udowodnimy teraz, że zawsze $\mathrm{tr}\,(\mathbf{AB}) = \mathrm{tr}\,(\mathbf{BA})$, jeśli tylko macierze **AB** i **BA** są określone (i ich współczynniki należą do pierścienia przemiennego).

Twierdzenie 4.4.1. *Jeśli* $\mathbf{A} \in K_{m\times n}$ *i* $\mathbf{B} \in K_{n\times m}$, *to ślady macierzy* **AB** *i* **BA** *są sobie równe, czyli*

$$\mathrm{tr}\,(\mathbf{AB}) = \mathrm{tr}\,(\mathbf{BA}). \tag{4.15}$$

Dowód. Załóżmy, że $\mathbf{A} = [a_{ij}]$, $\mathbf{B} = [b_{ij}]$, $\mathbf{AB} = [c_{ij}]$ i $\mathbf{BA} = [d_{ij}]$. Wtedy z definicji śladu i definicji iloczynu macierzy mamy

$$\mathrm{tr}\,(\mathbf{AB}) = \sum_{i=1}^{m} c_{ii} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{m} b_{ji}a_{ij} = \sum_{j=1}^{n} d_{jj} = \mathrm{tr}\,(\mathbf{BA}). \quad \square$$

Twierdzenie 4.4.2. *Jeśli* $\mathbf{A}, \mathbf{B} \in K_{n\times n}$, *to*

$$\mathrm{tr}\,(\mathbf{A}+\mathbf{B}) = \mathrm{tr}\,(\mathbf{A}) + \mathrm{tr}\,(\mathbf{B}) \quad \textit{i} \quad \mathrm{tr}\,(\mathbf{AB}) = \mathrm{tr}\,(\mathbf{BA}).$$

Dowód. Pierwsza część tezy jest oczywista, a druga wynika z poprzedniego twierdzenia. $\square$

Definicja 4.4.2. Macierze $\mathbf{A}, \mathbf{B} \in K_{n\times n}$ nazywamy *podobnymi*, gdy istnieje macierz odwracalna $\mathbf{C} \in K_{n\times n}$ taka, że

Macierze podobne

$$\mathbf{B} = \mathbf{C}^{-1}\mathbf{AC}.$$

Twierdzenie 4.4.3. *Jeśli macierze kwadratowe* **A** *i* **B** *są podobne, to*

$$\mathrm{tr}\,(\mathbf{A}) = \mathrm{tr}\,(\mathbf{B}).$$

Dowód. Niech **C** będzie macierzą taką, że $\mathbf{B} = \mathbf{C}^{-1}\mathbf{AC}$. Wtedy wobec twierdzenia 4.4.1 mamy

$$\begin{aligned} \mathrm{tr}\,(\mathbf{B}) &= \mathrm{tr}\,(\mathbf{C}^{-1}\mathbf{AC}) = \mathrm{tr}\,(\mathbf{C}^{-1}(\mathbf{AC})) \\ &= \mathrm{tr}\,((\mathbf{AC})\mathbf{C}^{-1}) = \mathrm{tr}\,(\mathbf{A}(\mathbf{CC}^{-1})) = \mathrm{tr}\,(\mathbf{AI}) = \mathrm{tr}\,(\mathbf{A}). \quad \square \end{aligned}$$

4.5. Ćwiczenia

1. Dane są macierze **A**, **B**, **C**, **x** i **y**, gdzie
$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 3 \\ 1 & 1 & 5 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 3 & 1 \end{bmatrix},$$
$$\mathbf{C} = \begin{bmatrix} 3 & 2 \\ 1 & 0 \\ 1 & 2 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$
Jeśli to możliwe, obliczyć następujące wielkości:

(a) $\mathbf{A}+2\mathbf{B}$;	(f) $(\mathbf{A}+\mathbf{C}^T)\mathbf{B}^T$;	(k) $\mathbf{y}^T\mathbf{Ax}$;
(b) $2\mathbf{A}-\mathbf{C}$;	(g) $\mathbf{A}^T(\mathbf{B}+\mathbf{C})$;	(l) $\mathbf{Axy}^T$;
(c) $\mathbf{AC}$;	(h) $\mathbf{Ax}-5\mathbf{y}$;	(m) $\mathbf{AA}^T$;
(d) $\mathbf{CA}$;	(i) $\mathbf{y}^T\mathbf{A}$;	(n) $\mathbf{xx}^T$;
(e) $\mathbf{AB}^T$;	(j) $\mathbf{x}^T\mathbf{C}$;	(o) $\mathbf{x}^T\mathbf{x}$.

2. Oblicz A^2 i A^5, jeśli $A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

3. Wyznaczyć liczbę $\left(\mathbf{AB}^T\right)_{21}$, gdy
$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 1 \end{bmatrix} \quad \text{i} \quad \mathbf{B} = \begin{bmatrix} 1 & 3 & 5 \\ 4 & 3 & 2 \\ 1 & 1 & 1 \end{bmatrix}.$$

4. Na przykładzie macierzy
$$\mathbf{A} = \begin{bmatrix} 1 & -2 \\ 2 & -4 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 3 & -1 \\ 2 & -2 \end{bmatrix} \text{ i } \mathbf{C} = \begin{bmatrix} -1 & 5 \\ 0 & 1 \end{bmatrix}$$
sprawdzić, czy z równości $\mathbf{AB} = \mathbf{AC}$ wynika równość $\mathbf{B} = \mathbf{C}$.

5. Czy macierze $\mathbf{AB}$ i $\mathbf{BA}$ są identyczne, gdy
$$\mathbf{A}=\begin{bmatrix}1&2&3\\3&2&0\\-1&-1&-1\end{bmatrix}\quad\text{i}\quad\mathbf{B}=\begin{bmatrix}-2&-1&-6\\3&2&9\\-1&-1&-4\end{bmatrix}.$$
6. Podać przykład niezerowych macierzy $\mathbf{A}$ i $\mathbf{B}$ wymiaru 2×2 i takich, że $\mathbf{AB}=\mathbf{0}$.
7. Podać przykład macierzy $\mathbf{A}$ i $\mathbf{B}$ takich, że dokładnie jeden z iloczynów $\mathbf{AB}$ i $\mathbf{BA}$ jest macierzą jednostkową.
8. Wyznaczyć macierze $\mathbf{B}$ i $\mathbf{C}$ takie, że
$$\mathbf{AB}=\begin{bmatrix}1&2\\-1&3\\2&0\end{bmatrix}\quad\text{i}\quad\mathbf{ACA}=\begin{bmatrix}2&-1&0\\0&-2&2\\3&1&4\end{bmatrix},\ \text{gdy}$$
$$\mathbf{A}^{-1}=\begin{bmatrix}1&0&1\\0&2&0\\1&0&3\end{bmatrix}.$$
9. Wyznaczyć, jeśli to możliwe, macierz $\mathbf{X}$ taką, że:
(a) $\mathbf{AX}=\mathbf{A}^3+2\mathbf{A}$, gdy $\mathbf{A}=\begin{bmatrix}4&-1&1\\-1&4&-1\\1&-1&4\end{bmatrix}$;
(b) $\begin{bmatrix}4&0&0\\0&4&0\\0&0&4\end{bmatrix}\mathbf{X}-\mathbf{X}\begin{bmatrix}2&0&0\\0&2&0\\0&0&2\end{bmatrix}=\begin{bmatrix}2&0&-2\\0&8&0\\4&0&-6\end{bmatrix}$.
10. Pokazać, że jeśli macierz $\mathbf{A}\in R_{m\times n}$ ma zerowy wiersz i $\mathbf{B}\in R_{n\times p}$, to także macierz $\mathbf{AB}$ ma zerowy wiersz.
11. Pokazać, że jeśli $\mathbf{A}\in R_{m\times n}$ i macierz $\mathbf{B}\in R_{n\times p}$ ma dwie identyczne kolumny, to także macierz $\mathbf{AB}$ ma dwie identyczne kolumny.
12. (a) Wykazać, że iloczyn dwóch macierzy trójkątnych górnych (zob. definicję 6.1.2) jest macierzą trójkątną górną. (b) Czy analogiczną własność mają macierze trójkątne dolne?
13. Korzystając z twierdzenia 4.2.5 udowodnić, że jeśli $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ $(k\geqslant 2)$ są macierzami takimi, że istnieje iloczyn $\mathbf{A}_1\cdot\mathbf{A}_2\cdot\ldots\cdot\mathbf{A}_k$, to
$$(\mathbf{A}_1\cdot\mathbf{A}_2\cdot\ldots\cdot\mathbf{A}_k)^T=\mathbf{A}_k^T\cdot\ldots\cdot\mathbf{A}_2^T\cdot\mathbf{A}_1^T.$$
14. Pokazać, że jeśli dla macierzy kwadratowych $\mathbf{A}$ i $\mathbf{B}$ jest $\mathbf{AB}=\mathbf{BA}$, to także $\mathbf{A}^n\mathbf{B}=\mathbf{BA}^n$ dla każdej liczby naturalnej n.
15. Uzasadnić, że jeśli $\mathbf{A}$ jest rzeczywistą macierzą kwadratową, to macierz $\frac{1}{2}(\mathbf{A}+\mathbf{A}^T)$ jest symetryczna, a macierz $\frac{1}{2}(\mathbf{A}-\mathbf{A}^T)$ jest skośnie-symetryczna. Wywnioskować stąd, że każda rzeczywista macierz kwadratowa jest sumą macierzy symetrycznej i skośnie-symetrycznej.
16. Pokazać, że jeśli $\mathbf{A}$ jest macierzą, to każda z macierzy $\mathbf{AA}^T$ i $\mathbf{A}^T\mathbf{A}$ jest symetryczna.
17. Niech $\mathbf{A}$ będzie macierzą symetryczną wymiaru $n\times n$ i niech $\mathbf{B}$ bedzie macierzą wymiaru $n\times m$. Wykazać, że $\mathbf{B}^T\mathbf{AB}$ jest macierzą symetryczną.
18. Niech $K^s_{n\times n}$ będzie zbiorem wszystkich macierzy symetrycznych wymiaru $n\times n$. Udowodnić, że $K^s_{n\times n}$ jest podgrupą grupy $K_{n\times n}$ (z działaniem dodawania macierzy).
19. Niech $K^o_{n\times n}$ będzie zbiorem wszystkich macierzy odwracalnych należących do zbioru $K_{n\times n}$. Udowodnić, że $K^o_{n\times n}$ jest grupą ze względu na mnożenie macierzy.
20. Pokazać, że zbiór macierzy postaci $\begin{bmatrix}1&x\\0&1\end{bmatrix}$, gdzie $x\in R$, jest grupą ze względu na mnożenie macierzy.
21. Macierz $\mathbf{A}$ nazywa się macierzą okresową, gdy $\mathbf{A}^{k+1}=\mathbf{A}$ dla pewnej dodatniej liczby naturalnej k. Wykazać, że macierz
$$\mathbf{A}=\begin{bmatrix}1&-2&-6\\-3&2&9\\2&0&-3\end{bmatrix}$$
jest okresowa. Dodatkowo wyznaczyć okres macierzy $\mathbf{A}$, czyli wyznaczyć najmniejszą dodatnią liczbę naturalną k, dla której $\mathbf{A}^{k+1}=\mathbf{A}$.
22. Macierz $\mathbf{A}$ nazywamy macierzą idempotentną, gdy $\mathbf{A}^2=\mathbf{A}$. (a) Wykazać, że macierz
$$\mathbf{A}=\begin{bmatrix}2&-2&-4\\-1&3&4\\1&-2&-3\end{bmatrix}$$
jest macierzą idempotentną. (b) Udowodnić, że macierze $\mathbf{A}$ i $\mathbf{B}$ są idempotentne, gdy $\mathbf{AB}=\mathbf{A}$ i $\mathbf{BA}=\mathbf{B}$.
23. O macierzy $\mathbf{X}$ mówi się, że jest macierzą nilpotentną rzędu k, gdy $\mathbf{X}\neq\mathbf{0}$ i k jest najmniejszą liczbą naturalną taką, że $\mathbf{X}^k=\mathbf{0}$. Sprawdzić nilpotentność macierzy
$$\mathbf{A}=\begin{bmatrix}1&-3&-4\\-1&3&4\\1&-3&-4\end{bmatrix}\quad\text{i}\quad\mathbf{B}=\begin{bmatrix}1&1&3\\5&2&6\\-2&-1&-3\end{bmatrix}.$$
24. Udowodnić, że dla macierzy kwadratowych $\mathbf{A}$ i $\mathbf{B}$ jest $(\mathbf{A}+\mathbf{B})(\mathbf{A}-\mathbf{B})=\mathbf{A}^2-\mathbf{B}^2$ i $(\mathbf{A}+\mathbf{B})^2=\mathbf{A}^2+2\mathbf{AB}+\mathbf{B}^2$ wtedy i tylko wtedy, gdy $\mathbf{AB}=\mathbf{BA}$.
25. O macierzy $\mathbf{X}$ mówi się, że jest macierzą inwolującą, gdy $\mathbf{X}^2=\mathbf{I}$. (a) Wykazać, że macierz $\mathbf{X}$ jest macierzą inwolującą wtedy i tylko wtedy, gdy $(\mathbf{I}-\mathbf{X})(\mathbf{I}+\mathbf{X})=\mathbf{0}$. (b) Sprawdzić inwolucyjność macierzy
$$\mathbf{A}=\begin{bmatrix}0&1&-1\\4&-3&4\\3&-3&4\end{bmatrix}\quad\text{i}\quad\mathbf{B}=\begin{bmatrix}4&3&3\\-1&0&-1\\-4&-4&-3\end{bmatrix}.$$
26. Przeprowadzić dowód równości (b) z twierdzenia 4.2.4.
27. Udowodnić, że jeśli $\mathbf{A}$, $\mathbf{B}$ i $\mathbf{C}$ są macierzami kwadratowymi takimi, że $\mathbf{AC}=\mathbf{BC}$ i macierz $\mathbf{C}$ jest odwracalna, to $\mathbf{A}=\mathbf{B}$.
28. Wykazać, że jeśli $\mathbf{A}$ i $\mathbf{B}$ są macierzami takimi, że $\mathbf{AB}=\mathbf{0}$ i macierz $\mathbf{A}$ jest odwracalna, to $\mathbf{B}=\mathbf{0}$.
29. Uzasadnić, że dla macierzy $\mathbf{A}=\begin{bmatrix}2&-3&3\\4&-5&3\\4&-4&2\end{bmatrix}$ jest

$\mathbf{A}^3 + \mathbf{A}^2 - 4\mathbf{A} - 4\mathbf{I}_3 = \mathbf{0}$. Wywnioskować stąd, że macierz $\mathbf{A}$ jest odwracalna i wyznaczyć macierz $\mathbf{A}^{-1}$.

30. Dana jest macierz $\mathbf{A} = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 1 \\ 2 & 1 & 2 \end{bmatrix}$. Wykazać, że $\mathbf{A}^3 = 3\mathbf{A}^2 - 3\mathbf{A} + \mathbf{I}_3$. Stąd wywnioskować, że $\mathbf{A}$ jest odwracalna i wyznaczyć $\mathbf{A}^{-1}$. Macierz $\mathbf{A}^4$ wyrazić poprzez macierze $\mathbf{A}^2$, $\mathbf{A}$ i $\mathbf{I}_3$ i znaleźć jej bezpośrednią postać.
31. Macierzą Heisenberga nazywamy każdą macierz postaci $\begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix}$, gdzie $x,\, y \in R$ i $x > 0$. (a) Udowodnić, że iloczyn macierzy Heisenberga jest macierzą Heisenberga. (b) Wykazać, że każda macierz Heisenberga jest odwracalna i uzasadnić, że jej macierzą odwrotną jest macierz Heisenberga. (c) Znaleźć macierz Heisenberga $\mathbf{X}$ taką, że $\begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}\mathbf{X} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$.
32. Niech $\mathbf{A}$ i $\mathbf{B}$ będą macierzami kwadratowymi i niech $\mathbf{A}$ będzie macierzą odwracalną. Udowodnić, że $\mathbf{AB} = \mathbf{BA}$ wtedy i tylko wtedy, gdy $\mathbf{A}^{-1}\mathbf{B} = \mathbf{BA}^{-1}$.
33. Udowodnić, że jeśli $\mathbf{A}$ jest macierzą wymiaru $n \times n$ i $\mathbf{X}$ jest jedyną macierzą taką, że $\mathbf{AX} = \mathbf{I}_n$, to $\mathbf{XA} = \mathbf{I}_n$ i dlatego $\mathbf{X} = \mathbf{A}^{-1}$.
34. Dana jest macierz $\mathbf{A} = \begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix}$. ($a$) Udowodnić, że $\mathbf{A}^2 = 4\mathbf{A} - 3\mathbf{I}_2$. ($b$) Indukcyjnie wykazać, że
$$\mathbf{A}^n = \frac{3^n - 1}{2}\mathbf{A} + \frac{3 - 3^n}{2}\mathbf{I}_2 \quad \text{dla} \quad n \geqslant 1.$$
35. Wykazać, że jeśli dla macierzy kwadratowych $\mathbf{A}$, $\mathbf{B}$ i $\mathbf{P}$ jest $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{B}$, to $\mathbf{P}^{-1}\mathbf{A}^n\mathbf{P} = \mathbf{B}^n$ dla każdej liczby naturalnej n.
36. Udowodnić, że jeśli $\mathbf{A} = \begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix}$ i $\mathbf{P} = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$, to $\mathbf{P}^{-1}\mathbf{AP} = \operatorname{diag}(3, 1)$. Korzystając z tej równości wykazać, że
$$\mathbf{A}^n = \frac{3^n - 1}{2}\mathbf{A} + \frac{3 - 3^n}{2}\mathbf{I}_2 \quad \text{dla} \quad n \geqslant 1.$$
37. Wykazać, że jeśli macierz $\mathbf{A}$ jest odwracalna, to $(\mathbf{A}^k)^{-1} = (\mathbf{A}^{-1})^k$ dla każdej liczby naturalnej k.
38. Ciągiem Fibonacciego (zob. (10.13)) nazywamy ciąg (F_n), w którym $F_0 = 0$, $F_1 = 1$ i $F_n = F_{n-1} + F_{n-2}$ dla $n \geqslant 2$. Korzystając z tego ciągu podać i udowodnić wzór na n-tą potęgę macierzy $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.
39. Podać i udowodnić wzór na n-tą potęgę macierzy $\mathbf{A}$ i następnie obliczyć $\lim_{n\to\infty} \mathbf{A}^n$, gdy:
(a) $\mathbf{A} = \begin{bmatrix} 1/2 & \alpha \\ 0 & 1/2 \end{bmatrix}$; ($b$) $\mathbf{A} = \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix}$.
40. Wykazać, że $\mathbf{A}^2 - (a + d)\mathbf{A} + (ad - bc)\mathbf{I}_2 = \mathbf{0}$, gdy $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.
41. Rzeczywista macierz kwadratowa o nieujemnych współczynnikach jest macierzą Markowa, jeśli suma współczynników każdej kolumny jest równa 1. (a) Czy w macierzy Markowa suma współczynników każdego wiersza także jest równa 1? (b) Czy iloczyn macierzy Markowa jest macierzą Markowa?
42. Udowodnić, że dla macierzy $\mathbf{A} \in R_{m\times n}$ jest $\operatorname{tr}(\mathbf{A}^T\mathbf{A}) = 0$ wtedy i tylko wtedy, gdy $\mathbf{A} = \mathbf{0}$.
43. Wpisując TAK albo NIE, stwierdzić prawdziwość każdego z następujących zdań:

$\boxed{1}$ Jeśli macierze $\mathbf{A}$ i $\mathbf{B}$ są równe, to dla każdej macierzy $\mathbf{C}$ jest $\mathbf{AC} = \mathbf{BC}$.

$\boxed{2}$ Jeśli dla macierzy $\mathbf{A}$, $\mathbf{B}$ i $\mathbf{C}$ jest $\mathbf{AC} = \mathbf{BC}$, to także jest $\mathbf{A} = \mathbf{B}$.

$\boxed{3}$ Jeśli $\mathbf{A}$ i $\mathbf{B}$ są macierzami kwadratowymi takimi, że $\mathbf{AB} = \mathbf{0}$, to także $\mathbf{BA} = \mathbf{0}$.

$\boxed{4}$ Jeśli $\mathbf{A}$ i $\mathbf{B}$ są macierzami takimi, że $\mathbf{AB} = \mathbf{0}$, to koniecznie $\mathbf{A} = \mathbf{0}$ lub $\mathbf{B} = \mathbf{0}$.

$\boxed{5}$ Jeśli $\mathbf{A}$ jest macierzą i $\mathbf{A}^2 = \mathbf{I}$, to $\mathbf{A} = \mathbf{I}$ lub $\mathbf{A} = -\mathbf{I}$.

$\boxed{6}$ Jeśli $\mathbf{A}$ jest macierzą i $\mathbf{A}^2 = \mathbf{I}$, to $\mathbf{A}^k = \mathbf{I}$ dla każdej liczby naturalnej $k \geqslant 2$.

$\boxed{7}$ Jeśli $\mathbf{A}$ jest macierzą i $\mathbf{A}^2 = \mathbf{I}$, to $\mathbf{A}^k = \mathbf{I}$ dla każdej parzystej liczby naturalnej $k \geqslant 2$.

$\boxed{8}$ Jeśli $\mathbf{A}$ i $\mathbf{B}$ są macierzami, to z równości $\mathbf{AB} = \mathbf{I}$ wynika odwracalność macierzy $\mathbf{A}$ i $\mathbf{B}$.

$\boxed{9}$ Jeśli macierze $\mathbf{A}$ i $\mathbf{B}$ są takie, że $\mathbf{AB}$ jest macierzą jednostkową i $\mathbf{BA}$ jest macierzą diagonalną, to $\mathbf{A}$ i $\mathbf{B}$ są wzajemnie odwrotne.

$\boxed{10}$ Niech $\mathbf{A}$, $\mathbf{B}$ i $\mathbf{C}$ będą macierzami. Czy z równości $\mathbf{A} + \mathbf{C} = \mathbf{B} + \mathbf{C}$ wynika równość $\mathbf{A} = \mathbf{B}$?

$\boxed{11}$ Jeśli $\mathbf{A}$ jest macierzą symetryczną wymiaru $n \times n$ i $\mathbf{A}^2 = [b_{ij}]$, to $b_{ii} \geqslant 0$ dla $i = 1, \ldots, n$.

$\boxed{12}$ Jeśli dla macierzy $\mathbf{A}$, $\mathbf{B}$ i $\mathbf{C}$ jest $\mathbf{AB} = \mathbf{C}$ i dwie spośród nich są kwadratowe, to i trzecia macierz jest kwadratowa.

$\boxed{13}$ Dla każdej macierzy kwadratowej $\mathbf{A}$ i liczby α jest $\operatorname{tr}(\alpha\mathbf{A}) = \alpha\operatorname{tr}(\mathbf{A})$.

$\boxed{14}$ Dla każdej macierzy kwadratowej $\mathbf{A}$ jest $\operatorname{tr}(\mathbf{A}^T) = \operatorname{tr}(\mathbf{A})$.

$\boxed{15}$ $\begin{bmatrix} -1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^2 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix}^3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \end{bmatrix}^4$.

Rozdział 5

UKŁADY RÓWNAŃ LINIOWYCH

5.1. Podstawowe definicje i fakty

Definicja 5.1.1. Układem m równań liniowych o n niewiadomych $x_1, x_2, \ldots, x_n$ nazywamy układ postaci

Układ równań liniowych

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \quad\vdots \qquad\qquad\qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases} \tag{5.1}$$

w którym *współczynniki* a_{ij} oraz b_i (te ostatnie nazywa się *wyrazami wolnymi* układu) są elementami ciała K. (W naszych rozważaniach tym ciałem będzie ciało liczb rzeczywistych R lub ciało liczb zespolonych $\mathcal{C}$.)

Ze współczynników, niewiadomych i wyrazów wolnych układu (5.1) można utworzyć macierze

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

i

$$[\mathbf{A}|\mathbf{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right]$$

nazywane odpowiednio *macierzą główną*, *macierzą niewiadomych*, *macierzą wyrazów wolnych* i *macierzą rozszerzoną* układu (5.1).

Układ (5.1) jest równoważny jednemu równaniu macierzowemu

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

które można zapisać w postaci

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \tag{5.2}$$

(symbolicznie $\mathbf{Ax} = \mathbf{b}$), zwanej *postacią macierzową* układu (5.1), oraz w postaci

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \tag{5.3}$$

(symbolicznie $x_1\mathbf{a}_{*1}+x_2\mathbf{a}_{*2}+\ldots+x_n\mathbf{a}_{*n} = \mathbf{b}$), zwanej *postacią wektorową* układu (5.1).

Przykład 80. Postacią macierzową i wektorową układu równań

$$\begin{cases} x_1 - 2x_2 + x_3 = 0, \\ 2x_2 - 8x_3 = 8, \\ -4x_1 + 5x_2 + 9x_3 = -9 \end{cases}$$

jest odpowiednio

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ -9 \end{bmatrix}$$

i

$$x_1 \begin{bmatrix} 1 \\ 0 \\ -4 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 2 \\ 5 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ -8 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ -9 \end{bmatrix}.$$

Definicja 5.1.2. *Rozwiązaniem układu* (5.1) nazywamy macierz jednokolumnową $\mathbf{s} \in K_{n\times 1}$, dla której

$$\mathbf{As} = \mathbf{b}.$$

Rozwiązanie układu

Równoważnie, przez rozwiązanie układu (5.1) możemy rozumieć taki ciąg $s_1, s_2, \ldots, s_n$ elementów ciała K, dla którego jest $a_{i1}s_1 + a_{i2}s_2 + \ldots + a_{in}s_n = b_i$ dla $i = 1, \ldots, m$. Mówimy, że układ (5.1) jest *niesprzeczny*, jeśli ma on co najmniej jedno rozwiązanie. Z drugiej strony układ (5.1) nazywamy układem *sprzecznym*, jeśli nie ma on żadnego rozwiązania.

Definicja 5.1.3. Dwa układy równań $\mathbf{Ax} = \mathbf{b}$ i $\mathbf{Cx} = \mathbf{d}$, gdzie $\mathbf{A}, \mathbf{C} \in K_{m\times n}$ i $\mathbf{b}, \mathbf{d} \in K_{m\times 1}$, nazywamy *równoważnymi*, gdy mają one identyczne zbiory rozwiązań.

Układy równoważne

Operacje elementarne

Przedstawimy teraz tzw. operacje elementarne na równaniach układu (5.1), operacje, za pomocą których układ sprowadza się do postaci, z której łatwo odczytuje się rozwiązania układu lub stwierdza się brak rozwiązań tego układu. Mamy trzy typy operacji elementarnych na równaniach układu:

Operacje elementarne na równaniach układu

- zamiana miejscami dwóch równań układu (5.1) (symbolem $r_i \leftrightarrow r_j$ będziemy oznaczać przestawienie miejscami równań r_i oraz r_j, gdzie $i \neq j$); $r_i \leftrightarrow r_j$
- pomnożenie równania układu (5.1) przez liczbę różną od zera (symbolu tr_i użyjemy na oznaczenie faktu pomnożenia obu stron równania r_i przez liczbę $t \neq 0$); tr_i
- dodanie do jednego równania układu (5.1) innego równania układu (5.1) pomnożonego przez liczbę różną od zera (symbolu $r_i + tr_j$ użyjemy, gdy do równania r_i dodamy równanie r_j ($j \neq i$) pomnożone przez liczbę t). $r_i + tr_j$

Przykład 81. Za pomocą operacji elementarnych można dany układ równań przekształcić w prostszy. Przykładowo mamy

$$\begin{cases} x_1 - x_2 + x_3 = 4 \\ x_1 + x_2 - x_3 = 10 \\ 3x_1 - x_2 + 2x_3 = 20 \end{cases} \xrightarrow[r_3 - 3r_1]{r_2 - r_1} \begin{cases} x_1 - x_2 + x_3 = 4 \\ 2x_2 - 2x_3 = 6 \\ 2x_2 - x_3 = 8 \end{cases} \xrightarrow[\frac{1}{2}r_2]{r_3 - r_2}$$

$$\begin{cases} x_1 - x_2 + x_3 = 4 \\ x_2 - x_3 = 3 \\ x_3 = 2. \end{cases}$$

Ostatni układ równań jest prostszy od wyjściowego układu i można go rozwiązać "metodą cofania". W tym celu otrzymane z ostatniego równania $x_3 = 2$ podstawiamy do drugiego równania ($x_2 - 2 = 3$) i z niego wyliczamy $x_2 = 5$. W końcu $x_2 = 5$ i $x_3 = 2$ podstawiamy do pierwszego równania ($x_1 - 5 + 2 = 4$) i otrzymujemy $x_1 = 7$. Zatem rozwiązaniem ostatniego (i wyjściowego) układu równań jest

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 5 \\ 2 \end{bmatrix}.$$

Ponieważ wszystkie informacje o układzie (5.1) zawarte są w jego macierzy rozszerzonej, zamiast wykonywania operacji elementarnych na równaniach układu wygodniej jest wykonywać odpowiednie operacje na wierszach macierzy rozszerzonej. Wyżej przedstawionym operacjom elementarnym na równaniach wzajemnie jednoznacznie odpowiadają operacje elementarne na wierszach macierzy. Niech $w_1, w_2, \ldots, w_m$ będą wierszami macierzy $\mathbf{A} \in K_{m\times n}$. Operacjami elementarnymi na wierszach macierzy $\mathbf{A}$ są:

Operacje elementarne na wierszach macierzy

- zamiana miejscami dwóch wierszy macierzy $\mathbf{A}$ (symbolicznie $w_i \leftrightarrow w_j$ dla $i \neq j$); $w_i \leftrightarrow w_j$
- pomnożenie wiersza macierzy $\mathbf{A}$ przez liczbę różną od zera (symbolicznie tw_i dla $t \neq 0$); tw_i
- dodanie do jednego wiersza macierzy $\mathbf{A}$ innego wiersza macierzy $\mathbf{A}$ pomnożonego przez liczbę różną od zera (symbolicznie $w_i + tw_j$). $w_i + tw_j$

Za pomocą operacji elementarnych na wierszach macierzy można daną macierz przekształcić w macierz mającą określone własności. Następujący przykład ilustruje przekształcenie macierzy kwadratowej w macierz trójkątną górną.

Przykład 82.

$$\begin{bmatrix} 0 & 3 & 5 & 5 \\ 2 & 4 & 2 & 6 \\ 0 & 3 & 4 & 5 \\ 0 & 6 & 8 & 7 \end{bmatrix} \xrightarrow{w_1 \leftrightarrow w_2} \begin{bmatrix} 2 & 4 & 2 & 6 \\ 0 & 3 & 5 & 5 \\ 0 & 3 & 4 & 5 \\ 0 & 6 & 8 & 7 \end{bmatrix} \xrightarrow{w_4 - 2w_3}$$

$$\begin{bmatrix} 2 & 4 & 2 & 6 \\ 0 & 3 & 5 & 5 \\ 0 & 3 & 4 & 5 \\ 0 & 0 & 0 & -3 \end{bmatrix} \xrightarrow{w_3 - w_2} \begin{bmatrix} 2 & 4 & 2 & 6 \\ 0 & 3 & 5 & 5 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}.$$

Macierze elementarne

Macierz elementarna

Definicja 5.1.4. *Macierzą elementarną* nazywamy macierz otrzymaną z macierzy jednostkowej w wyniku jednej operacji elementarnej na jej wierszach.

Niech $\mathbf{E}_{ij}$, $\mathbf{E}_i(t)$ i $\mathbf{E}_{ij}(t)$ będą odpowiednio macierzami elementarnymi otrzymanymi w wyniku operacji elementarnych $w_i \leftrightarrow w_j$, tw_i i $w_i + tw_j$ na wierszach macierzy jednostkowej $\mathbf{I}_m$. Warto zauważyć, że jeśli $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m$ są kolejnymi wierszami macierzy $\mathbf{I}_m$, to

$$
\mathbf{E}_{ij} = \left[\begin{matrix} - & \mathbf{e}_1 & - \\ & \vdots & \\ - & \mathbf{e}_{i-1} & - \\ - & \mathbf{e}_j & - \\ - & \mathbf{e}_{i+1} & - \\ & \vdots & \\ - & \mathbf{e}_{j-1} & - \\ - & \mathbf{e}_i & - \\ - & \mathbf{e}_{j+1} & - \\ & \vdots & \\ - & \mathbf{e}_m & - \end{matrix}\right]\begin{matrix} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \leftarrow j \\ \\ \\ \\ \end{matrix} = \left[\begin{matrix} 1 &&&&&&&&&& \\ & \ddots &&&&&&&&& \\ && 1 &&&&&&&& \\ &&& 0 & 0 & \cdots && 1 &&& \\ &&& 0 & 1 &&& 0 &&& \\ &&& \vdots && \ddots && \vdots &&& \\ &&& 0 &&& 1 & 0 &&& \\ &&& 1 && \cdots & 0 & 0 &&& \\ &&&&&&&& 1 && \\ &&&&&&&&& \ddots & \\ &&&&&&&&&& 1 \end{matrix}\right]\begin{matrix} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \leftarrow j \\ \\ \\ \\ \end{matrix},
$$

$$
\mathbf{E}_i(t) = \left[\begin{matrix} - & \mathbf{e}_1 & - \\ & \vdots & \\ - & \mathbf{e}_{i-1} & - \\ - & t\mathbf{e}_i & - \\ - & \mathbf{e}_{i+1} & - \\ & \vdots & \\ - & \mathbf{e}_m & - \end{matrix}\right]\begin{matrix} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \end{matrix} = \left[\begin{matrix} 1 & \cdots & 0 &&&& \\ \vdots & \ddots & \vdots &&&& \\ 0 & \cdots & 1 &&&& \\ &&& t &&& \\ &&&& 1 & \cdots & 0 \\ &&&& \vdots & \ddots & \vdots \\ &&&& 0 & \cdots & 1 \end{matrix}\right]\begin{matrix} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \end{matrix},
$$

$$
\mathbf{E}_{ij}(t) = \left[\begin{matrix} - & \mathbf{e}_1 & - \\ & \vdots & \\ - & \mathbf{e}_{i-1} & - \\ - & \mathbf{e}_i + t\mathbf{e}_j & - \\ - & \mathbf{e}_{i+1} & - \\ & \vdots & \\ - & \mathbf{e}_j & - \\ & \vdots & \\ - & \mathbf{e}_m & - \end{matrix}\right]\begin{matrix} \\ \\ \\ \leftarrow i \\ \\ \\ \leftarrow j \\ \\ \\ \end{matrix} = \left[\begin{matrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 & \cdots & t && \\ && \vdots & \ddots & \vdots && \\ && 0 & \cdots & 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{matrix}\right]\begin{matrix} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{matrix}.
$$

Przykład 83. Przykładowymi macierzami elementarnymi stopnia 4 są

$$
\mathbf{E}_{34} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \mathbf{E}_3(5) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ i } \mathbf{E}_{24}(5) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
$$

Niech teraz $\mathbf{A}$ będzie macierzą wymiaru $m \times n$, której kolejnymi wierszami są $\mathbf{a}_1$, $\mathbf{a}_2, \ldots, \mathbf{a}_m$. Niech $\mathbf{B}$, $\mathbf{C}$ i $\mathbf{D}$ będą odpowiednio macierzami otrzymanymi z macierzy $\mathbf{A}$ w wyniku przestawienia i-tego wiersza z j-tym, przemnożenia i-tego wiersza przez liczbę t oraz dodania j-tego wiersza pomnożonego przez t do i-tego wiersza. Z faktu, że iloczyn $\mathbf{e}_k\mathbf{A}$ jest kombinacją wierszy macierzy $\mathbf{A}$ ze współczynnikami z jednowierszowej macierzy $\mathbf{e}_k = [\underbrace{0\ldots0}_{k-1}\,1\,0\ldots0]$ (zob. (4.6)) wynika, że

$$\mathbf{e}_k\mathbf{A} = \mathbf{a}_k$$

dla $k = 1, \ldots, m$. Stąd zaś wynika, że mamy

$$\mathbf{B} = \begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & \mathbf{a}_j & - \\ & \vdots & \\ - & \mathbf{a}_i & - \\ & \vdots & \\ - & \mathbf{a}_m & - \end{bmatrix} \begin{array}{l} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{array} = \begin{bmatrix} - & \mathbf{e}_1\mathbf{A} & - \\ & \vdots & \\ - & \mathbf{e}_j\mathbf{A} & - \\ & \vdots & \\ - & \mathbf{e}_i\mathbf{A} & - \\ & \vdots & \\ - & \mathbf{e}_m\mathbf{A} & - \end{bmatrix} = \begin{bmatrix} - & \mathbf{e}_1 & - \\ & \vdots & \\ - & \mathbf{e}_j & - \\ & \vdots & \\ - & \mathbf{e}_i & - \\ & \vdots & \\ - & \mathbf{e}_m & - \end{bmatrix} \cdot \mathbf{A} = \mathbf{E}_{ij}\mathbf{A},$$

$$\mathbf{C} = \begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & t\mathbf{a}_i & - \\ & \vdots & \\ - & \mathbf{a}_m & - \end{bmatrix} \leftarrow i = \begin{bmatrix} - & \mathbf{e}_1\mathbf{A} & - \\ & \vdots & \\ - & t\mathbf{e}_i\mathbf{A} & - \\ & \vdots & \\ - & \mathbf{e}_m\mathbf{A} & - \end{bmatrix} = \begin{bmatrix} - & \mathbf{e}_1 & - \\ & \vdots & \\ - & t\mathbf{e}_i & - \\ & \vdots & \\ - & \mathbf{e}_m & - \end{bmatrix} \cdot \mathbf{A} = \mathbf{E}_i(t)\mathbf{A},$$

$$\mathbf{D} = \begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & \mathbf{a}_i + t\mathbf{a}_j & - \\ & \vdots & \\ - & \mathbf{a}_m & - \end{bmatrix} \leftarrow i = \begin{bmatrix} - & \mathbf{e}_1\mathbf{A} & - \\ & \vdots & \\ - & (\mathbf{e}_i + t\mathbf{e}_j)\mathbf{A} & - \\ & \vdots & \\ - & \mathbf{e}_m\mathbf{A} & - \end{bmatrix} = \begin{bmatrix} - & \mathbf{e}_1 & - \\ & \vdots & \\ - & \mathbf{e}_i + t\mathbf{e}_j & - \\ & \vdots & \\ - & \mathbf{e}_m & - \end{bmatrix} \cdot \mathbf{A} = \mathbf{E}_{ij}(t)\mathbf{A}.$$

Oznacza to, że efekt wykonania jakiejkolwiek operacj elementarnej na wierszach macierzy $\mathbf{A}$ jest tożsamy z przemnożeniem macierzy $\mathbf{A}$ przez macierz elementarną odpowiadającą tej operacji elementarnej na wierszach. Można także rozważać operacje elementarne na kolumnach macierzy $\mathbf{A}$ i zaobserwować, że w wyniku każdej takiej operacji uzyskujemy dokładnie to samo, co w wyniku pomnożenia macierzy $\mathbf{A}$ przez macierz elementarną odpowiadającą tej operacji elementarnej.

Przykład 84.

$$\begin{aligned} \mathbf{E}_{23}(4)\mathbf{A} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} + 4a_{31} & a_{22} + 4a_{32} & a_{23} + 4a_{33} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}. \end{aligned}$$

Inną ważną własność macierzy elementarnych przedstawia następujące twierdzenie.

Twierdzenie 5.1.1. *Każda macierz elementarna jest odwracalna i macierz odwrotna macierzy elementarnej jest macierzą elementarną.*

Dowód. Łatwo zauważyć, że dla macierzy elementarnych mamy

$$\mathbf{E}_{ij}\mathbf{E}_{ij} = \mathbf{I}_m, \quad \mathbf{E}_i(t)\mathbf{E}_i(1/t) = \mathbf{I}_m \quad \text{i} \quad \mathbf{E}_{ij}(t)\mathbf{E}_{ij}(-t) = \mathbf{I}_m.$$

Zatem

$$(\mathbf{E}_{ij})^{-1} = \mathbf{E}_{ij}, \quad (\mathbf{E}_i(t))^{-1} = \mathbf{E}_i(1/t) \quad \text{oraz} \quad (\mathbf{E}_{ij}(t))^{-1} = \mathbf{E}_{ij}(-t)$$

i stąd wynika teza. □

Przykład 85. Znaleźć macierze odwrotne następujących macierzy elementarnych

$$\mathbf{E} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{i} \quad \mathbf{G} = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Z dowodu poprzedniego twierdzenia wynika, że

$$\mathbf{E}^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{F}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{i} \quad \mathbf{G}^{-1} = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Wierszowa równoważność macierzy

Definicja 5.1.5. Mówimy, że macierz $\mathbf{B}$ jest *wierszowo równoważna* macierzy $\mathbf{A}$, piszemy $\mathbf{A} \sim \mathbf{B}$, gdy macierz $\mathbf{B}$ można uzyskać z macierzy $\mathbf{A}$ za pomocą skończonego ciągu operacji elementarnych na wierszach.

Z odpowiedniości pomiędzy operacjami elementarnymi na wierszach i przemnażaniem macierzy przez macierze elementarne wynika, że macierz $\mathbf{B}$ jest wierszowo równoważna macierzy $\mathbf{A}$ (czyli $\mathbf{A} \sim \mathbf{B}$) wtedy i tylko wtedy, gdy istnieją macierze elementarne $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_k$ takie, że $\mathbf{B} = \mathbf{E}_k\mathbf{E}_{k-1}\ldots\mathbf{E}_1\mathbf{A}$. Ponieważ macierze $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_k$ są odwracalne, więc także macierz $\mathbf{P} = \mathbf{E}_k\mathbf{E}_{k-1}\ldots\mathbf{E}_1$ jest odwracalna i dlatego mamy: jeśli dla macierzy $\mathbf{A}$ i $\mathbf{B}$ jest $\mathbf{A} \sim \mathbf{B}$, to istnieje taka macierz odwracalna $\mathbf{P}$, że $\mathbf{B} = \mathbf{PA}$. (Udowodnimy, że stwierdzenie odwrotne też jest prawdziwe, tj. udowodnimy, że jeśli $\mathbf{B} = \mathbf{PA}$ i macierz $\mathbf{P}$ jest odwracalna, to $\mathbf{A} \sim \mathbf{B}$.) Warto zauważyć, że dla macierzy $\mathbf{P}$ jest $\mathbf{P} = \mathbf{E}_k\mathbf{E}_{k-1}\ldots\mathbf{E}_1\mathbf{I}$, a to oznacza, że $\mathbf{P}$ powstaje z macierzy jednostkowej $\mathbf{I}$ za pomocą operacji na wierszach odpowiadających kolejno macierzom $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_k$.

Przykład 86. Pokazać, że macierze $\mathbf{A}$ i $\mathbf{B}$ są wierszowo równoważne i znaleźć taką macierz $\mathbf{P}$, że $\mathbf{B} = \mathbf{PA}$, gdy

$$\mathbf{A} = \begin{bmatrix} 2 & 8 & 7 \\ 0 & 2 & 2 \\ 1 & 4 & 3 \end{bmatrix} \quad \text{i} \quad \mathbf{B} = \begin{bmatrix} 1 & 4 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Dla dowodu wierszowej równoważności macierzy $\mathbf{A}$ i $\mathbf{B}$ wystarczy wskazać ciąg operacji elementarnych przekształcających macierz $\mathbf{A}$ w macierz $\mathbf{B}$ (zob. dwie pierwsze kolumny następującej tablicy). Każdej takiej operacji elementarnej odpowiada macierz elementarna, czyli macierz uzyskana w wyniku wykonania tej samej operacji elementarnej na wierszach macierzy $\mathbf{I}_3$. Iloczyn uzyskanych macierzy elementarnych tworzy macierz $\mathbf{P}$.

Równoważność macierzy $\mathbf{A}$ i $\mathbf{B}$	Operacje elementarne	Macierze elementarne	Równoważność macierzy $\mathbf{I}_3$ i $\mathbf{P}$
$\mathbf{A} = \begin{bmatrix} 2 & 8 & 7 \\ 0 & 2 & 2 \\ 1 & 4 & 3 \end{bmatrix}$	$w_1 - 2w_3$	$E_{13}(-2) = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
$\sim \begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 2 \\ 1 & 4 & 3 \end{bmatrix}$	$\frac{1}{2}w_2$	$E_2(\frac{1}{2}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
$\sim \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 4 & 3 \end{bmatrix}$	$w_2 + w_1$	$E_{21}(1) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$
$\sim \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 4 & 3 \end{bmatrix}$	$w_1 \leftrightarrow w_3$	$E_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$	$\sim \begin{bmatrix} 1 & 0 & -2 \\ 1 & \frac{1}{2} & -2 \\ 0 & 0 & 1 \end{bmatrix}$
$\sim \begin{bmatrix} 1 & 4 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$			$\sim \begin{bmatrix} 0 & 0 & 1 \\ 1 & \frac{1}{2} & -2 \\ 1 & 0 & -2 \end{bmatrix}$
$= \mathbf{B}$			$= \mathbf{P}$

Ponieważ $\mathbf{B} = \mathbf{E}_{13}\mathbf{E}_{21}(1)\mathbf{E}_2(\frac{1}{2})\mathbf{E}_{13}(-2)\mathbf{A}$, więc mamy

$$\begin{aligned}\mathbf{P} &= \mathbf{E}_{13}\mathbf{E}_{21}(1)\mathbf{E}_2(\tfrac{1}{2})\mathbf{E}_{13}(-2)\\ &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & \frac{1}{2} & -2 \\ 1 & 0 & -2 \end{bmatrix}.\end{aligned}$$

Dla wyznaczenia macierzy $\mathbf{P}$ nie trzeba wymnażać macierzy $\mathbf{E}_{13}$, $\mathbf{E}_{21}(1)$, $\mathbf{E}_2(\frac{1}{2})$ i $\mathbf{E}_{13}(-2)$. Wystarczy zauważyć, że $\mathbf{P}$ jest macierzą uzyskaną z macierzy $\mathbf{I}_3$ kolejno za pomocą operacji $w_1 - 2w_3$, $\frac{1}{2}w_2$, $w_2 + w_1$ i $w_1 \leftrightarrow w_3$. Zatem $\mathbf{P}$ jest macierzą wygenerowaną w ostatniej kolumnie powyższej tabeli.

Rozwiązywanie układu równań liniowych

Omówimy teraz praktyczną metodę rozwiązywania układów równań. W metodzie tej macierz rozszerzoną układu przekształca się w wierszowo równoważną macierz rozszerzoną prostszego i równoważnego układu równań.

Twierdzenie 5.1.2. *Jeśli macierze $[\mathbf{A}|\mathbf{b}]$ i $[\mathbf{C}|\mathbf{d}]$ są wierszowo równoważne, to układy równań liniowych $\mathbf{Ax} = \mathbf{b}$ i $\mathbf{Cx} = \mathbf{d}$ są równoważne.*

Dowód. Załóżmy, że macierze $[\mathbf{A}|\mathbf{b}]$ i $[\mathbf{C}|\mathbf{d}]$ są wierszowo równoważne. Wtedy istnieje macierz odwracalna $\mathbf{P}$ taka, że $[\mathbf{C}|\mathbf{d}] = \mathbf{P}[\mathbf{A}|\mathbf{b}] = [\mathbf{PA}|\mathbf{Pb}]$ i dlatego także $\mathbf{C} = \mathbf{PA}$ oraz $\mathbf{d} = \mathbf{Pb}$. Ponieważ macierz $\mathbf{P}$ jest odwracalna, więc mamy równoważności

$$\mathbf{Ax} = \mathbf{b} \Leftrightarrow \mathbf{PAx} = \mathbf{Pb} \Leftrightarrow \mathbf{Cx} = \mathbf{d},$$

a to oznacza, że układy $\mathbf{Ax} = \mathbf{b}$ i $\mathbf{Cx} = \mathbf{d}$ mają takie same rozwiązania i dlatego są one równoważne. $\square$

Praktycznie rozwiązując układ równań liniowych $\mathbf{Ax} = \mathbf{b}$, przechodzić będziemy od jego macierzy rozszerzonej $[\mathbf{A}|\mathbf{b}]$ do wierszowo równoważnej macierzy mającej tzw. postać schodkową.

Macierz schodkowa

Definicja 5.1.6. Mówimy, że macierz $\mathbf{M}$ ma *postać schodkową* (albo jest *macierzą schodkową*), jeśli ma ona następujące własności:

- każdy wiersz składający się z samych zer (jeśli taki wiersz w macierzy istnieje) występuje w niej po wszystkich niezerowych wierszach;
- w każdym niezerowym wierszu pierwszym niezerowym elementem jest jedynka – nazywamy ją *wiodącą jedynką* tego wiersza – i występuje ona na prawo od wiodącej jedynki każdego wcześniejszego wiersza.

Wiodąca jedynka

Normalna macierz schodkowa

Powiemy także, że macierz $\mathbf{M}$ ma *normalną postać schodkową* (albo jest *normalną macierzą schodkową*), jeśli ma ona postać schodkową i dodatkowo

- każda wiodąca jedynka jest jedynym niezerowym elementem kolumny, w której ona występuje.

Przykład 87. Dane są macierze

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 1 & 3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & 1 & 3 \\ 1 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{i} \quad \mathbf{D} = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}.$$

Macierz $\mathbf{A}$ nie ma postaci schodkowej, bo wiodąca jedynka trzeciego wiersza nie występuje na prawo od wiodącej jedynki drugiego wiersza. Z podobnych powodów macierz $\mathbf{C}$ nie ma postaci schodkowej. Macierze $\mathbf{B}$ i $\mathbf{D}$ są w postaci schodkowej. Dodatkowo, macierz $\mathbf{D}$ ma normalną postać schodkową. “Schodkowość” następnych dwóch macierzy $\mathbf{E}$ i $\mathbf{F}$ podkreślona została schodkowymi łamanymi oddzielającymi zerowe części ich wierszy od tych, które zaczynają się od wiodących jedynek. Macierz **F** oczywiście ma normalną postać schodkową. W obu macierzach "gwiazdkowe" współczynniki mogą mieć dowolną wartość.

$$\mathbf{E} = \begin{bmatrix} 0 & 1 & * & * & * & * & * & * \\ 0 & 0 & 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{i} \quad \mathbf{F} = \begin{bmatrix} 0 & 1 & * & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 1 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Twierdzenie 5.1.3. *Każda macierz* $\mathbf{A} = [a_{ij}] \in K_{m\times n}$ *jest wierszowo równoważna macierzy schodkowej* (*normalnej macierzy schodkowej*).

Uwaga. Można udowodnić, że każda macierz jest wierszowo równoważna dokładnie jednej macierzy w normalnej postaci schodkowej.

Dowód. Przedstawiamy tu dowód indukcyjny ze względu na $s = mn$. Twierdzenie jest oczywiste, gdy **A** jest macierzą zerową i łatwo wynika z założenia indukcyjnego, jeśli **A** ma zerowy wiersz lub zerową kolumnę. W pozostałych przypadkach przestawiając miejscami dwa wiersze, można macierz **A** sprowadzić do wierszowo równoważnej macierzy $\mathbf{B} = [b_{ij}]$, w której $b_{11} \neq 0$. Mnożąc teraz pierwszy wiersz macierzy **B** przez $1/b_{11}$ otrzymamy wierszowo równoważną macierz $\mathbf{C} = [c_{ij}]$ z $c_{11} = 1$. Mnożąc teraz pierwszy wiersz macierzy **C** przez $-c_{k1}$ i dodając go do k-tego wiersza (dla $k = 2,\dots,m$), otrzymamy wierszowo równoważną macierz

$$\mathbf{D} = \begin{bmatrix} 1 & c_{12} & \cdots & c_{1n} \\ 0 & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & d_{m2} & \cdots & d_{mn} \end{bmatrix}.$$

Wobec założenia indukcyjnego macierz $\mathbf{D}'$, którą otrzymujemy z **D** przez wykreślenie pierwszego wiersza, jest wierszowo równoważna pewnej macierzy schodkowej $\mathbf{D}''$. Wstawiając $\mathbf{D}''$ na miejsce $\mathbf{D}'$ uzyskujemy macierz schodkową $\mathbf{E} = [e_{ij}]$ wierszowo równoważną macierzy **D** (i **A**).

W celu otrzymania normalnej macierzy schodkowej wystarczy kolejno (zaczynając od ostatniego niezerowego wiersza) każdy i-ty wiersz ($i \geqslant 2$) macierzy **E** zawierający wiodącą jedynkę, powiedzmy w j-tej kolumnie, pomnożyć przez $-e_{lj}$ i dodać go do l-tego wiersza dla $l = 1,\dots,i-1$. □

Niezerową macierz możemy sprowadzić do wierszowo równoważnej macierzy schodkowej, posługując się następującym algorytmem.

Algorytm redukcji macierzy do postaci schodkowej

Dla redukcji niezerowej macierzy **A** do wierszowo równoważnej macierzy schodkowej wykonujemy następujące kroki:

1. Zaczynając od lewej strony, odnajdujemy pierwszą niezerową kolumnę macierzy **A** i niezerowy współczynnik a tej kolumny. Jeśli trzeba, przestawiamy wiersze tak, aby niezerowy współczynnik a znalazł się w pierwszym wierszu.

2. Pierwszy wiersz mnożymy przez $1/a$ dla otrzymania wiodącej jedynki w pierwszym wierszu.

3. Za pomocą operacji zastępowania wierszy (tj. poprzez dodawanie do jednego wiersza macierzy innego wiersza pomnożonego przez liczbę) tworzymy zera wszędzie poniżej wiodącej jedynki.

4. Przykrywamy (lub ignorujemy) wiersz zawierający wiodącą jedynkę i przykrywamy wszystkie wiersze powyżej niego, jeśli takie są. Dla powstałej podmacierzy powtarzamy kroki 1–3. Proces ten powtarzamy do wyczerpania niezerowych wierszy.

Normalną macierz schodkową otrzymamy, wykonując jeszcze jeden krok.

5. Kolejno zaczynając od ostatniej wiodącej jedynki, za pomocą operacji zastępowania wierszy tworzymy zera powyżej każdej wiodącej jedynki.

Przykład 88. Za pomocą powyższego algorytmu macierz $\mathbf{A}$ sprowadzić do wierszowo równoważnej macierzy mającej normalną postać schodkową, gdy

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 1 \\ 2 & 6 & 2 & 4 & 2 & 0 \\ 3 & 9 & 5 & 6 & 4 & 4 \end{bmatrix}.$$

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 1 \\ 2 & 6 & 2 & 4 & 2 & 0 \\ 3 & 9 & 5 & 6 & 4 & 4 \end{bmatrix}$$

Krok 1. Pierwsza kolumna jest nizerowa i $a = 2$ znajdujące się w drugim wierszu może być jej niezerowym elementem. Przestawiamy miejscami dwa pierwsze wiersze.

$$\sim \begin{bmatrix} 2 & 6 & 2 & 4 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 3 & 9 & 5 & 6 & 4 & 4 \end{bmatrix}$$

Krok 2. Pierwszy wiersz mnożymy przez $1/2$; uzyskujemy wiodącą jedynkę w pierwszym wierszu.

$$\sim \begin{bmatrix} 1 & 3 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 3 & 9 & 5 & 6 & 4 & 4 \end{bmatrix}$$

Krok 3. Do trzeciego wiersza dodajemy pierwszy wiersz pomnożony przez -3; uzyskujemy zera pod wiodącą jedynką.

$$\sim \begin{bmatrix} 1 & 3 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 & 1 & 4 \end{bmatrix}$$

Krok 4. Przykrywamy pierwszy wiersz, wiersz zawierający wiodącą jedynką.

$$\sim \left[\begin{array}{cccccc} 1 & 3 & 1 & 2 & 1 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 & 1 & 4 \end{array}\right]$$

Kroki 1-3. Trzecia kolumna jest pierwszą niezerową kolumną podmacierzy i $a = 1$ jest niezerowym współczynnikiem tej kolumny znajdującym sie w “pierwszym” wierszu (kroki 1 i 2). Do “drugiego” wiersza dodajemy “pierwszy” pomnożony przez -2 (krok 3).

$$\sim \left[\begin{array}{cccccc} 1 & 3 & 1 & 2 & 1 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right]$$

Krok 4. Przykrywamy “pierwszy” wiersz i wiersz poprzedzający; otrzymujemy macierz mającą tylko jeden wiersz.

$$\sim \left[\begin{array}{cccccc} 1 & 3 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right]$$

Kroki 1-3 niczego nie zmieniają w tej macierzy jednowierszowej. **Krokiem 4** wyczerpujemy wszystkie wiersze macierzy. Otrzymaliśmy macierz schodkową.

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

Krok 5. Od pierwszego wiersza odejmujemy trzeci i następnie od pierwszego odejmujemy drugi.

$$\sim \begin{bmatrix} 1 & 3 & 0 & 2 & 0 & -3 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

Otrzymaliśmy macierz mającą normalną postać schodkową.

Wiodąca kolumna macierzy

Definicja 5.1.7. *Kolumną wiodącą* macierzy $\mathbf{A}$ nazywamy każdą kolumnę, która w macierzy $\mathbf{A}$ lub w macierzy wierszowo równoważnej ma wiodącą jedynkę.

Z pojęcia tego skorzystamy w kolejnym twierdzeniu, które mówi o istnieniu i ilości rozwiązań układu równań liniowych.

Przykład 89. Kolumną wiodącą w macierzy

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 1 \\ 2 & 6 & 2 & 4 & 2 & 0 \\ 3 & 9 & 5 & 6 & 4 & 4 \end{bmatrix}$$

(z poprzedniego przykładu) jest odpowiednio pierwsza, trzecia i piąta kolumna.

Twierdzenie 5.1.4. *Niech* $\mathbf{Ax} = \mathbf{b}$ *będzie układem równań liniowych i niech macierz* $[\mathbf{C}|\mathbf{d}] \in K_{m\times(n+1)}$ *mająca normalną postać schodkową i dokładnie* r *wiodących kolumn będzie wierszowo równoważna macierzy* $[\mathbf{A}|\mathbf{b}]$. *Układ* $\mathbf{Ax} = \mathbf{b}$ *ma rozwiązanie wtedy i tylko wtedy, gdy* $\mathbf{d}$ *nie jest wiodącą kolumną macierzy* $[\mathbf{C}|\mathbf{d}]$. *Dodatkowo, jeśli* $\mathbf{d}$ *nie jest wiodącą kolumną macierzy* $[\mathbf{C}|\mathbf{d}]$, *to układ* $\mathbf{Ax} = \mathbf{b}$ *ma dokładnie jedno rozwiązanie, gdy* $r = n$, *a ma on co najmniej dwa rozwiązania, gdy* $r < n$ (*nieskończenie wiele rozwiązań, gdy* $r < n$ *i ciało* K *jest nieskończone*).

Dowód. Załóżmy, że $\mathbf{d}$ jest wiodącą kolumną macierzy $[\mathbf{C}|\mathbf{d}]$. Wtedy $[0 \ldots 0\,|\,1]$ jest jej r-tym wierszem i dlatego $0x_1+0x_2+\ldots+0x_n = 1$ jest r-tym równaniem układu $\mathbf{Cx} = \mathbf{d}$. Równanie to nie ma rozwiązania, więc także układ $\mathbf{Cx} = \mathbf{d}$ nie ma rozwiązania. Stąd i z twierdzenia 5.1.2 wynika, że także układ $\mathbf{Ax} = \mathbf{b}$ nie ma rozwiązania.

Załóżmy teraz, że $\mathbf{d}$ nie jest wiodącą kolumną macierzy $[\mathbf{C}|\mathbf{d}]$. Jeśli $r = n$, to wszystkie kolumny macierzy $[\mathbf{C}|\mathbf{d}]$ poza ostatnią są wiodące i dlatego mamy

$$[\mathbf{C}|\mathbf{d}] = \left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & d_1 \\ 0 & 1 & \cdots & 0 & d_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & d_n \\ \hline 0 & 0 & \cdots & 0 & 0 \end{array}\right],$$

przy czym ostatni zerowy wiersz może nie pojawić się wcale albo takich zerowych wierszy może być więcej. Macierzy tej odpowiada układ równań

$$\begin{cases} x_1 & & & & = & d_1, \\ & x_2 & & & = & d_2, \\ & & \ddots & & & \vdots \\ & & & x_n & = & d_n, \\ & & & 0 & = & 0, \end{cases}$$

którego jedynym rozwiązaniem jest $x_1 = d_1, \ldots, x_n = d_n$. Wobec twierdzenia 5.1.2 jest to także jedyne rozwiązanie układu $\mathbf{Ax} = \mathbf{b}$.

Jeśli $r < n$, to macierzą $[\mathbf{C}|\mathbf{d}]$ (po ewentualnej permutacji pierwszych n kolumn) jest

$$\left[\begin{array}{cccc|ccc|c} 1 & 0 & \cdots & 0 & c_{1\,r+1} & \cdots & c_{1\,n} & d_1 \\ 0 & 1 & \cdots & 0 & c_{2\,r+1} & \cdots & c_{2\,n} & d_2 \\ & & \ddots & & & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & c_{r\,r+1} & \cdots & c_{r\,n} & d_r \\ \hline 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{array}\right].$$

Macierzy tej odpowiada układ równań

$$\begin{cases} x_1 & & & & + & c_{1\,r+1}x_{r+1} & + & \cdots & + & c_{1\,n}x_n & = & d_1, \\ & x_2 & & & + & c_{2\,r+1}x_{r+1} & + & \cdots & + & c_{2\,n}x_n & = & d_2, \\ & & \ddots & & & & & & & & & \vdots \\ & & & x_r & + & c_{r\,r+1}x_{r+1} & + & \cdots & + & c_{r\,n}x_n & = & d_r, \\ & & & & & & & & & 0 & = & 0, \end{cases}$$

którego rozwiązaniem jest

$$\mathbf{x} = \begin{bmatrix} d_1 - c_{1\,r+1}x_{r+1} - \cdots - c_{1\,n}x_n \\ d_2 - c_{2\,r+1}x_{r+1} - \cdots - c_{2\,n}x_n \\ \vdots \\ d_r - c_{r\,r+1}x_{r+1} - \cdots - c_{r\,n}x_n \\ x_{r+1} \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_r \\ 0 \\ \vdots \\ 0 \end{bmatrix} - x_{r+1} \begin{bmatrix} c_{1\,r+1} \\ c_{2\,r+1} \\ \vdots \\ c_{r\,r+1} \\ 1 \\ \vdots \\ 0 \end{bmatrix} - \ldots - x_n \begin{bmatrix} c_{1\,n} \\ c_{2\,n} \\ \vdots \\ c_{r\,n} \\ 0 \\ \vdots \\ 1 \end{bmatrix},$$

gdzie niewiadomym $x_{r+1}, \ldots, x_n$ (które nazywa się *parametrami* lub *niewiadomymi wolnymi*) nadano dowolne wartości z ciała K. Zgodnie z twierdzeniem 5.1.2, tak wyznaczone $\mathbf{x}$ (dla każdych parametrów $x_{r+1}, \ldots, x_n \in K$) jest także rozwiązaniem wyjściowego układu rówań liniowych $\mathbf{Ax} = \mathbf{b}$. Stąd wynika ostatnia część tezy. □

Z ostatniego twierdzenia i z jego dowodu otrzymujemy prosty algorytm rozwiązywania układów równań liniowych. Algorytm ten zwykle nazywa się algorytmem Gaussa albo algorytmem Gaussa-Jordana.

Algorytm Gaussa i Gaussa-Jordana rozwiązywania układów równań liniowych

Dla rozwiązania układu równań liniowych $\mathbf{Ax} = \mathbf{b}$ wykonujemy następujące kroki:

1. Tworzymy macierz rozszerzoną $[\mathbf{A}|\mathbf{b}]$ układu $\mathbf{Ax} = \mathbf{b}$.

2. Macierz $[\mathbf{A}|\mathbf{b}]$ redukujemy do wierszowo równoważnej macierzy schodkowej (normalnej macierzy schodkowej) $[\mathbf{C}|\mathbf{d}]$. Jeśli $\mathbf{d}$ jest wiodącą kolumną macierzy $[\mathbf{C}|\mathbf{d}]$, układ $\mathbf{Ax} = \mathbf{b}$ jest sprzeczny. W przeciwnym przypadku przechodzimy do następnego kroku.

3. Wypisujemy układ równań liniowych $\mathbf{Cx} = \mathbf{d}$.

4. Z układu $\mathbf{Cx} = \mathbf{d}$ metodą cofania (wprost, gdy $[\mathbf{C}|\mathbf{d}]$ jest normalną macierzą schodkową) wyznaczamy niewiadome odpowiadające wiodącym kolumnom macierzy $[\mathbf{C}|\mathbf{d}]$.

Metoda Gaussa i metoda Gaussa-Jordana

Sposób rozwiązywania układów równań liniowych zgodnie z powyższym algorytmem nazywamy odpowiednio *metodą Gaussa* lub *metodą Gaussa-Jordana* rozwiązywania układów równań liniowych w zależności od tego, czy macierz rozszerzoną $[\mathbf{A}|\mathbf{b}]$ rozwiązywanego układu równań $\mathbf{Ax} = \mathbf{b}$ sprowadzamy do wierszowo równoważnej macierzy mającej postać schodkową, czy schodkową normalną. Metody te ilustrujemy trzema przykładami.

Przykład 90. Metodą Gaussa rozwiązać układ równań liniowych

$$\begin{cases} \phantom{-4x_1 +{}} 2x_2 - 8x_3 = 8, \\ x_1 - 2x_2 + x_3 = 0, \\ -4x_1 + 5x_2 + 10x_3 = -6. \end{cases}$$

Macierz rozszerzoną powyższego układu za pomocą operacji elementarnych przekształcamy w macierz schodkową. Tu mamy

$$[\mathbf{A}|\mathbf{b}] = \left[\begin{array}{rrr|r} 0 & 2 & -8 & 8 \\ 1 & -2 & 1 & 0 \\ -4 & 5 & 10 & -6 \end{array}\right]$$ Przestawiamy pierwszy wiersz z drugim dla uzyskania wiodącej jedynki w pierwszym wierszu ($w_1 \leftrightarrow w_2$).

$$\sim \left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 10 & -6 \end{array}\right]$$ Pierwszy wiersz pomnożony przez 4 dodajemy do trzeciego dla otrzymania zer pod wiodącą jedynką ($w_3 + 4w_1$).

$$\sim \left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 14 & -6 \end{array}\right]$$ Drugi wiersz mnożymy przez $\frac{1}{2}$ dla uzyskania wiodącej jedynki w drugim wierszu ($\frac{1}{2}w_2$).

$$\sim \left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 14 & -6 \end{array}\right]$$ Drugi wiersz pomnożony przez 3 dodajemy do trzeciego dla otrzymania zera pod kolejną wiodącą jedynką ($w_3 + 3w_2$).

$$\sim \left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 2 & 6 \end{array}\right]$$ Trzeci wiersz mnożymy przez $\frac{1}{2}$ dla uzyskania w nim wiodącej jedynki ($\frac{1}{2}w_3$).

$$\sim \left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right].$$

Ostatnia macierz ma postać schodkową i odpowiada jej układ równań

$$\begin{cases} x_1 - 2x_2 + x_3 = 0, \\ \phantom{x_1 - {}} x_2 - 4x_3 = 4, \\ \phantom{x_1 - 2x_2 + {}} x_3 = 3, \end{cases}$$

którego rozwiązaniem (otrzymanym metodą cofania) jest $x_3 = 3$, $x_2 = 16$ i $x_1 = 29$. Wobec twierdzenia 5.1.2 macierz

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 29 \\ 16 \\ 3 \end{bmatrix}$$

jest jedynym rozwiązaniem wyjściowego układu równań.

Przykład 91. Rozwiązać układ równań liniowych

$$\begin{cases} x_1 + 2x_2 + x_4 = 1, \\ x_1 + 3x_2 + 3x_3 + 3x_4 = 4, \\ x_2 + 4x_3 + 6x_4 = 8, \\ 2x_1 + 5x_2 + 3x_3 + 4x_4 = 6. \end{cases}$$

Dla macierzy rozszerzonej tego układu mamy następujące równoważności

$$\left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 1 \\ 1 & 3 & 3 & 3 & 4 \\ 0 & 1 & 4 & 6 & 8 \\ 2 & 5 & 3 & 4 & 6 \end{array}\right] \underset{\sim}{\substack{w_2 - w_1 \\ w_4 - 2w_1}} \left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 1 \\ 0 & 1 & 3 & 2 & 3 \\ 0 & 1 & 4 & 6 & 8 \\ 0 & 1 & 3 & 2 & 4 \end{array}\right] \underset{\sim}{\substack{w_3 - w_2 \\ w_4 - w_2}} \left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 1 \\ 0 & 1 & 3 & 2 & 3 \\ 0 & 0 & 1 & 4 & 5 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

Ostatnia kolumna ostatniej macierzy jest jej kolumną wiodącą, więc z twierdzenia 5.1.4 wynika, że rozważany układ równań liniowych nie ma rozwiązania.

Zaprezentujemy teraz przykład układu mającego nieskończenie wiele rozwiązań.

Przykład 92. Metodą Gaussa-Jordana rozwiązać układ równań liniowych $\mathbf{A}\mathbf{x} = \mathbf{b}$, gdzie

$$\begin{cases} x_1 + 3x_3 + 2x_4 = 10, \\ 2x_1 + x_2 + x_3 + x_4 = 7, \\ 2x_1 + 2x_3 + 4x_4 = 8, \\ x_1 + 2x_2 + 3x_3 - 4x_4 = 14. \end{cases}$$

Macierz rozszerzoną $[\mathbf{A}|\mathbf{b}]$ tego układu sprowadzamy do macierzy $[\mathbf{C}|\mathbf{d}]$ mającej normalną postać schodkową,

$$[\mathbf{A}|\mathbf{b}] = \left[\begin{array}{cccc|c} 1 & 0 & 3 & 2 & 10 \\ 2 & 1 & 1 & 1 & 7 \\ 2 & 0 & 2 & 4 & 8 \\ 1 & 2 & 3 & -4 & 14 \end{array}\right] \overset{\substack{w_2 - 2w_1 \\ w_3 - 2w_1 \\ w_4 - w_1}}{\sim} \left[\begin{array}{cccc|c} 1 & 0 & 3 & 2 & 10 \\ 0 & 1 & -5 & -3 & -13 \\ 0 & 0 & -4 & 0 & -12 \\ 0 & 2 & 0 & -6 & 4 \end{array}\right] \begin{array}{c} w_4 - 2w_2 \\ \left(-\frac{1}{4}\right)w_3 \end{array}$$

$$\sim \left[\begin{array}{cccc|c} 1 & 0 & 3 & 2 & 10 \\ 0 & 1 & -5 & -3 & -13 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 10 & 0 & 30 \end{array}\right] \overset{\substack{w_1 - 3w_3 \\ w_2 + 5w_3 \\ w_4 - 10w_3}}{\sim} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & -3 & 2 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] = [\mathbf{C}|\mathbf{d}].$$

Ostatniej macierzy odpowiada układ równań $\mathbf{C}\mathbf{x} = \mathbf{d}$,

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

z którego wyznaczamy niewiadome odpowiadające wiodącym jedynkom macierzy $[\mathbf{C}|\mathbf{d}]$: $x_1 = 1 - 2x_4$, $x_2 = 2 + 3x_4$, $x_3 = 3$. Zatem rozwiązaniem układu $\mathbf{A}\mathbf{x} = \mathbf{b}$ (i układu $\mathbf{C}\mathbf{x} = \mathbf{d}$) jest

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 - 2x_4 \\ 2 + 3x_4 \\ 3 \\ x_4 \end{bmatrix}$$

dla każdego $x_4 \in R$.

Z twierdzenia 5.1.4 i jego dowodu wynika następujący wniosek.

Wniosek 5.1.1. *Niech* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *będzie układem równań liniowych, gdzie* $[\mathbf{A}|\mathbf{b}] \in K_{m\times(n+1)}$.

(1) *Jeśli* $m < n$, *to układ* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *jest sprzeczny albo ma on co najmniej dwa rozwiązania.*

(2) *Jeśli* $m = n$, *to układ* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *ma dokładnie jedno rozwiązanie wtedy i tylko wtedy, gdy macierz* $\mathbf{A}$ *jest wierszowo równoważna macierzy jednostkowej* $\mathbf{I}_n$. $\square$

5.2. Równania macierzowe

Weźmy pod uwagę równanie macierzowe

$$\mathbf{A}\mathbf{X} = \mathbf{B},$$

w którym $\mathbf{A}$ i $\mathbf{B}$ są macierzami mającymi tyle samo wierszy, powiedzmy $\mathbf{A} \in K_{m\times n}$ i $\mathbf{B} = [\mathbf{b}_1\, \mathbf{b}_2 \ldots \mathbf{b}_p] \in K_{m\times p}$, a wtedy macierz niewiadoma $\mathbf{X}$ jest macierzą wymiaru $n \times p$,

$$\mathbf{X} = \begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \ldots & x_{1p} \\ x_{21} & x_{22} & & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \ldots & x_{np} \end{bmatrix}.$$

Ponieważ

$$\mathbf{AX} = \mathbf{A}\begin{bmatrix} | & & | \\ \mathbf{x}_1 & \ldots & \mathbf{x}_p \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ \mathbf{Ax}_1 & \ldots & \mathbf{Ax}_p \\ | & & | \end{bmatrix} \quad \text{i} \quad \mathbf{B} = \begin{bmatrix} | & & | \\ \mathbf{b}_1 & \ldots & \mathbf{b}_p \\ | & & | \end{bmatrix},$$

więc równanie $\mathbf{AX} = \mathbf{B}$ jest równoważne następującemu układowi p równań

$$\mathbf{Ax}_1 = \mathbf{b}_1,\ \mathbf{Ax}_2 = \mathbf{b}_2,\ \ldots,\ \mathbf{Ax}_p = \mathbf{b}_p.$$

Każdy z tych układów jest układem m równań liniowych o n niewiadomych i każdy z nich można z osobna rozwiązać metodą Gaussa-Jordana. W tym przypadku każdą z macierzy $[\mathbf{A}|\mathbf{b}_1]$, $[\mathbf{A}|\mathbf{b}_2]$, $\ldots$, $[\mathbf{A}|\mathbf{b}_p]$ sprowadza się do wierszowo równoważnej normalnej macierzy schodkowej. Wygodniej (i oszczędniej) jest utworzyć macierz rozszerzoną $[\mathbf{A}|\mathbf{B}]$ i sprowadzić ją do wierszowo równoważnej normalnej macierzy schodkowej $[\mathbf{C}|\mathbf{D}] = [\mathbf{C}|\mathbf{d}_1 \ldots \mathbf{d}_p]$. Z tej ostatniej macierzy można utworzyć macierze $[\mathbf{C}|\mathbf{d}_1]$, $\ldots$, $[\mathbf{C}|\mathbf{d}_p]$ (wierszowo równoważne macierzom $[\mathbf{A}|\mathbf{b}_1]$, $\ldots$, $[\mathbf{A}|\mathbf{b}_p]$) i z nich lub wprost z macierzy $[\mathbf{C}|\mathbf{D}]$ można odczytać rozwiązania układów $\mathbf{Ax}_1 = \mathbf{b}_1$, $\ldots$, $\mathbf{Ax}_p = \mathbf{b}_p$, więc także rozwiązanie równania $\mathbf{AX} = \mathbf{B}$ (albo stwierdzić brak rozwiązania równania $\mathbf{AX} = \mathbf{B}$). Z twierdzenia 5.1.4 wynika, że równanie $\mathbf{AX} = \mathbf{B}$ ma rozwiązanie wtedy i tylko wtedy, gdy żadna kolumna macierzy $\mathbf{D}$ nie jest wiodącą kolumną macierzy $[\mathbf{C}|\mathbf{D}]$. Wyżej przedstawiony sposób rozwiązywania równań macierzowych postaci $\mathbf{AX} = \mathbf{B}$ jest prostym uogólnieniem metody Gaussa-Jordana rozwiązywania układów równań liniowych i dalej będzie nosił miano metody Gaussa-Jordana.

Przykład 93. Rozwiązać układy równań liniowych $\mathbf{Ax} = \mathbf{b}$ i $\mathbf{Ay} = \mathbf{b}'$, gdzie

$$\begin{cases} x_1 - 2x_2 + x_3 = -4 \\ 2x_1 - x_2 + 3x_3 = -7 \\ 4x_1 + 2x_2 + x_3 = 0 \end{cases} \quad \text{i} \quad \begin{cases} y_1 - 2y_2 + y_3 = 1 \\ 2y_1 - y_2 + 3y_3 = 10 \\ 4y_1 + 2y_2 + y_3 = 18 \end{cases}.$$

Układ powyższych dwóch układów równań liniowych jest równoważny równaniu macierzowemu

$$\begin{bmatrix} 1 & -2 & 1 \\ 2 & -1 & 3 \\ 4 & 2 & 1 \end{bmatrix}\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{bmatrix} = \begin{bmatrix} -4 & 1 \\ -7 & 10 \\ 0 & 18 \end{bmatrix}.$$

Łatwo sprawdzić, że dla macierzy rozszerzonej tego równania mamy

$$\left[\begin{array}{ccc|cc} 1 & -2 & 1 & -4 & 1 \\ 2 & -1 & 3 & -7 & 10 \\ 4 & 2 & 1 & 0 & 18 \end{array}\right] \sim \ldots \sim \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & -2 & 2 \end{array}\right].$$

Stąd widać, że rozwiązaniami układów są odpowiednio liczby $x_1 = 0$, $x_2 = 1$, $x_3 = -2$ oraz $y_1 = 3$, $y_2 = 2$, $y_3 = 2$.

Z faktu, że umiemy rozwiązać równanie macierzowe $\mathbf{AX} = \mathbf{B}$ wynika, że umiemy także rozwiązać równanie postaci $\mathbf{XC} = \mathbf{D}$, bo jest ono równoważne równaniu $\mathbf{C}^T\mathbf{X}^T = \mathbf{D}^T$.

Przykład 94. Rozwiązać równanie macierzowe $\mathbf{XA} + \mathbf{X} = \mathbf{B}$, gdzie

$$\mathbf{A} = \begin{bmatrix} 0 & -1 & 2 \\ 1 & 1 & 0 \\ 0 & -3 & 2 \end{bmatrix} \quad \text{i} \quad \mathbf{B} = \begin{bmatrix} 0 & -6 & 5 \\ 2 & -5 & 7 \end{bmatrix}.$$

Mamy równoważności

$$\mathbf{XA} + \mathbf{X} = \mathbf{B} \Leftrightarrow \mathbf{X}(\mathbf{A} + \mathbf{I}) = \mathbf{B} \Leftrightarrow (\mathbf{A} + \mathbf{I})^T\mathbf{X}^T = \mathbf{B}^T$$

i z ostatniego równania metodą Gaussa-Jordana (lub inną) możemy wyznaczyć $\mathbf{X}^T$ i dlatego także $\mathbf{X}$. Mamy

$$[\,(\mathbf{A}+\mathbf{I})^T\,|\,\mathbf{B}^T\,] = \left[\begin{array}{rrr|rr} 1 & 1 & 0 & 0 & 2 \\ -1 & 2 & -3 & -6 & -5 \\ 2 & 0 & 3 & 5 & 7 \end{array}\right] \sim \left[\begin{array}{rrr|rr} 1 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right] = \left[\,\mathbf{I}\,|\,\mathbf{X}^T\,\right],$$

więc macierz $\mathbf{X} = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \end{bmatrix}$ jest rozwiązaniem równania $\mathbf{XA}+\mathbf{X}=\mathbf{B}$.

Przykład 95. Rozwiązać równania macierzowe $\mathbf{AX}_1 = \mathbf{B}_1$ i $\mathbf{AX}_2 = \mathbf{B}_2$, gdzie

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & -2 \\ 0 & 1 & 1 \\ 2 & -5 & -5 \end{bmatrix}, \quad \mathbf{B}_1 = \begin{bmatrix} 1 & -4 & -3 \\ 0 & 2 & 1 \\ 2 & -10 & -7 \end{bmatrix} \quad \text{i} \quad \mathbf{B}_2 = \begin{bmatrix} 2 & -6 \\ 0 & 3 \\ 4 & -14 \end{bmatrix}.$$

Oba równania $\mathbf{AX}_1 = \mathbf{B}_1$ i $\mathbf{AX}_2 = \mathbf{B}_2$ można rozwiązywać jednocześnie. W tym celu macierz $[\mathbf{A}|\mathbf{B}_1|\mathbf{B}_2]$ sprowadzamy do wierszowo równoważnej normalnej macierzy schodkowej. Mamy

$$\begin{aligned} [\,\mathbf{A}\,|\,\mathbf{B}_1\,|\,\mathbf{B}_2\,] &= \left[\begin{array}{rrr|rrr|rr} 1 & -2 & -2 & 1 & -4 & -3 & 2 & -6 \\ 0 & 1 & 1 & 0 & 2 & 1 & 0 & 3 \\ 2 & -5 & -5 & 2 & -10 & -7 & 4 & -14 \end{array}\right] \\ &\sim \left[\begin{array}{rrr|rrr|rr} 1 & 0 & 0 & 1 & 0 & -1 & 2 & 0 \\ 0 & 1 & 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \end{aligned}$$

i stąd wynika, że rozwiązaniem równania $\mathbf{AX}_1 = \mathbf{B}_1$ jest macierz

$$\mathbf{X}_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ -\alpha & -\beta & -\gamma \\ \alpha & \beta & \gamma \end{bmatrix}$$

dla każdych $\alpha,\ \beta,\ \gamma \in R$. Równanie $\mathbf{AX}_2 = \mathbf{B}_2$ nie ma rozwiązania, bo macierz $[\,\mathbf{A}\,|\,\mathbf{B}_2\,]$ jest wierszowo równoważna macierzy

$$\left[\begin{array}{rrr|rr} 1 & 0 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right],$$

w której ostatnia kolumna jest kolumną wiodącą.

5.3. Kolejne własności macierzy odwracalnej

Wiemy już, że macierze elementarne i ich iloczyny są macierzami odwracalnymi. Teraz przedstawiamy kolejne związki macierzy odwracalnej z macierzami elementarnymi oraz odwracalności macierzy z istnieniem rozwiązań pewnych równań macierzowych i układów równań liniowych.

Twierdzenie 5.3.1. *Jeśli* $\mathbf{A}$ *jest macierzą kwadratową stopnia* n*, to następujące stwierdzenia są równoważne:*

- (i) $\mathbf{A}$ *jest odwracalna;*
- (ii) *Równanie* $\mathbf{AX} = \mathbf{I}_n$ *ma rozwiązanie;*
- (iii) *Równanie* $\mathbf{Ax} = \mathbf{b}$ *ma rozwiązanie dla każdego* $\mathbf{b} \in K_{n\times 1}$*;*
- (iv) $\mathbf{A}$ *jest wierszowo równoważna macierzy jednostkowej* $\mathbf{I}_n$*;*
- (v) $\mathbf{A}$ *jest iloczynem macierzy elementarnych.*

Dowód. $(i) \Rightarrow (ii)$. Załóżmy, że macierz $\mathbf{A}$ jest odwracalna. Wtedy macierz $\mathbf{A}^{-1}$ istnieje i z równości $\mathbf{AX} = \mathbf{I}_n$ po przemnożeniu obu jej stron przez $\mathbf{A}^{-1}$ otrzymujemy rozwiązanie $\mathbf{X} = \mathbf{A}^{-1}$ równania $\mathbf{AX} = \mathbf{I}_n$.

$(ii) \Rightarrow (iii)$. Załóżmy, że dla macierzy $\mathbf{C} \in K_{n\times n}$ jest $\mathbf{AC} = \mathbf{I}_n$. Jeśli $\mathbf{b} \in K_{n\times 1}$, to $\mathbf{Cb}$ jest rozwiązaniem równania $\mathbf{Ax} = \mathbf{b}$, bo $\mathbf{A}(\mathbf{Cb}) = (\mathbf{AC})\mathbf{b} = \mathbf{I}_n\mathbf{b} = \mathbf{b}$.

$(iii) \Rightarrow (iv)$. Załóżmy, że równanie $\mathbf{Ax} = \mathbf{b}$ ma rozwiązanie dla każdego $\mathbf{b} \in K_{n\times 1}$. Niech $\mathbf{B}$ będzie macierzą wierszowo równoważną macierzy $\mathbf{A}$ i mającą normalną postać schodkową. Wtedy istnieje macierz odwracalna $\mathbf{P}$ (będąca iloczynem macierzy elementernych) taka, że $\mathbf{B} = \mathbf{PA}$. Twierdzimy, że $\mathbf{B} = \mathbf{I}_n$. Gdyby było inaczej, to (co najmniej) ostatni wiersz macierzy $\mathbf{B}$ byłby zerowy i układ $\mathbf{Bx} = \mathbf{e}_n$ byłby sprzeczny. Wtedy także równoważny układ $\mathbf{Ax} = \mathbf{P}^{-1}\mathbf{e}_n$ byłby sprzeczny, co zaprzeczałoby naszemu założeniu.

$(iv) \Rightarrow (v)$. Jeśli macierze $\mathbf{A}$ i $\mathbf{I}_n$ są wierszowo równoważne, to istnieją macierze elementarne $\mathbf{E}_1, \ldots, \mathbf{E}_k$ takie, że $\mathbf{I}_n = \mathbf{E}_k \ldots \mathbf{E}_2\mathbf{E}_1\mathbf{A}$. Wtedy $\mathbf{A} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1} \ldots \mathbf{E}_k^{-1}$ i także macierze $\mathbf{E}_1^{-1}, \ldots, \mathbf{E}_k^{-1}$ są elementarne (zob. tw. 5.1.1).

$(v) \Rightarrow (i)$. Załóżmy teraz, że $\mathbf{A} = \mathbf{E}_1\mathbf{E}_2 \ldots \mathbf{E}_k$, gdzie $\mathbf{E}_1, \ldots, \mathbf{E}_k$ są macierzami elementarnymi. Ponieważ $\mathbf{E}_1, \ldots, \mathbf{E}_k$ są odwracalne, to także ich iloczyn jest odwracalny. To dowodzi, że macierz $\mathbf{A}$ jest odwracalna i $\mathbf{A}^{-1} = (\mathbf{E}_1 \ldots \mathbf{E}_k)^{-1} = \mathbf{E}_k^{-1} \ldots \mathbf{E}_1^{-1}$. □

Przykład 96. Macierz $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ -2 & -2 \end{bmatrix}$ przedstawić w postaci iloczynu macierzy elementarnych. Następnie wyznaczyć macierz $\mathbf{A}^{-1}$.

Szukany iloczyn można uzyskać tak jak w dowodzie implikacji $(iv) \Rightarrow (v)$ poprzedniego twierdzenia. Macierze $\mathbf{A}$ i $\mathbf{I}_2$ są wierszowo równoważne i mamy

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ -2 & -2 \end{bmatrix} \overset{w_2+2w_1}{\longrightarrow} \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \overset{w_1-w_2}{\longrightarrow} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \overset{(\frac{1}{2})w_2}{\longrightarrow} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}_2.$$

Stąd $\mathbf{E}_2\left(\frac{1}{2}\right)\mathbf{E}_{1\,2}(-1)\mathbf{E}_{2\,1}(2)\mathbf{A} = \mathbf{I}_2$, więc

$$\begin{aligned} \mathbf{A} &= \big(\mathbf{E}_{2\,1}(2)\big)^{-1}\big(\mathbf{E}_{1\,2}(-1)\big)^{-1}\big(\mathbf{E}_2(\tfrac{1}{2})\big)^{-1} \\ &= \mathbf{E}_{2\,1}(-2)\mathbf{E}_{1\,2}(1)\mathbf{E}_2(2) \\ &= \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \end{aligned}$$

i

$$\mathbf{A}^{-1} = \mathbf{E}_2(\tfrac{1}{2})\mathbf{E}_{1\,2}(-1)\mathbf{E}_{2\,1}(2) = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & \frac{1}{2} \end{bmatrix}.$$

5.4. Wyznaczanie macierzy odwrotnej

Metoda Gaussa-Jordana rozwiązywania równań macierzowych postaci $\mathbf{AX} = \mathbf{B}$ jest przydatna przy wyznaczaniu macierzy odwrotnej. Następujący wniosek wprost powiada, że rozwiązaniem równania $\mathbf{AX} = \mathbf{I}_n$ (jeśli takie istnieje) jest macierz $\mathbf{A}^{-1}$.

Wniosek 5.4.1. *Jeśli* $\mathbf{A}$ *i* $\mathbf{B}$ *są macierzami kwadratowymi stopnia* n*, to* $\mathbf{B}$ *jest macierzą odwrotną macierzy* $\mathbf{A}$ *wtedy i tylko wtedy, gdy* $\mathbf{B}$ *jest rozwiązaniem równania* $\mathbf{AX} = \mathbf{I}_n$*.*

Dowód. Jeśli $\mathbf{B} = \mathbf{A}^{-1}$, to oczywiście mamy $\mathbf{AB} = \mathbf{AA}^{-1} = \mathbf{I}_n$. Załóżmy teraz, że $\mathbf{AB} = \mathbf{I}_n$. Oznacza to, że równanie $\mathbf{AX} = \mathbf{I}_n$ ma rozwiazanie. Stąd i z twierdzenia 5.3.1 wynika, że macierz $\mathbf{A}$ jest odwracalna. Zatem $\mathbf{A}^{-1}$ istnieje i z równości $\mathbf{AB} = \mathbf{I}_n$, po jej przemnożeniu przez $\mathbf{A}^{-1}$, otrzymujemy $\mathbf{B} = \mathbf{A}^{-1}$. □

Przykład 97. Wyznaczyć macierz odwrotną macierzy

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 2 & 1 \\ 2 & 5 & 4 & 2 \\ 3 & 9 & 5 & 3 \\ 4 & 12 & 8 & 3 \end{bmatrix}.$$

Dla wyznaczenia macierzy $\mathbf{A}^{-1}$ wystarczy rozwiązać równanie $\mathbf{AX} = \mathbf{I}_4$. Możemy to zrobić metodą Gaussa-Jordana. W tym celu macierz $[\mathbf{A}|\mathbf{I}_4]$ sprowadzamy do wierszowo równoważnej normalnej macierzy schodkowej,

$$\left[\begin{array}{cccc|cccc} 1 & 3 & 2 & 1 & 1 & 0 & 0 & 0 \\ 2 & 5 & 4 & 2 & 0 & 1 & 0 & 0 \\ 3 & 9 & 5 & 3 & 0 & 0 & 1 & 0 \\ 4 & 12 & 8 & 3 & 0 & 0 & 0 & 1 \end{array}\right] \underset{\substack{w_3-3w_1 \\ w_4-4w_1}}{\overset{w_2-2w_1}{\sim}} \left[\begin{array}{cccc|cccc} 1 & 3 & 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & -2 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & -4 & 0 & 0 & 1 \end{array}\right] \begin{array}{c} w_1+3w_2+2w_3+w_4 \\ (-1)w_2 \\ (-1)w_3 \\ (-1)w_4 \end{array}$$

$$\sim \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -15 & 3 & 2 & 1 \\ 0 & 1 & 0 & 0 & 2 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 3 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 4 & 0 & 0 & -1 \end{array}\right].$$

Zatem

$$\mathbf{A}^{-1} = \begin{bmatrix} -15 & 3 & 2 & 1 \\ 2 & -1 & 0 & 0 \\ 3 & 0 & -1 & 0 \\ 4 & 0 & 0 & -1 \end{bmatrix}.$$

Problemu odwracalności macierzy $\mathbf{A}$ i sposobu wyznaczania macierzy $\mathbf{A}^{-1}$ nie musimy kojarzyć z rozwiązalnością i sposobem wyznaczania rozwiązania równania $\mathbf{AX} = \mathbf{I}_n$. Wobec twierdzenia 5.3.1 odwracalność macierzy $\mathbf{A}$ jest tożsama wierszowej równoważności macierzy $\mathbf{A}$ i $\mathbf{I}_n$. Sposób wyznaczania macierzy $\mathbf{A}^{-1}$ jest także praktyczną konsekwencją dowodu tej równoważności.

Wniosek 5.4.2. *Jeśli* $\mathbf{A}$ *jest macierzą kwadratową stopnia* n*, to dla pewnej macierzy* $\mathbf{P}$ *jest* $\mathbf{PA} = \mathbf{I}_n$ *wtedy i tylko wtedy, gdy* $\mathbf{P}[\mathbf{A}|\mathbf{I}_n] = [\mathbf{I}_n|\mathbf{A}^{-1}]$.

Dowód. Załóżmy, że dla pewnej macierzy $\mathbf{P}$ jest $\mathbf{PA} = \mathbf{I}_n$. Wtedy $\mathbf{A} = \mathbf{P}^{-1}$ (zob. wniosek 5.4.1) i $\mathbf{A}^{-1} = \mathbf{P} = \mathbf{PI}_n$. Zatem $\mathbf{P}[\mathbf{A}|\mathbf{I}_n] = [\mathbf{PA}|\mathbf{PI}_n] = [\mathbf{I}_n|\mathbf{A}^{-1}]$. Implikacja odwrotna jest oczywista. □

Ponieważ macierz odwracalna jest iloczynem macierzy elementarnych (zob. twierdzenie 5.3.1), a każdej macierzy elementarnej jednoznacznie odpowiada operacja elementarna na wierszach macierzy, więc tezę ostatniego twierdzenia możemy przedstawić następująco: ciąg operacji elementarnych na wierszach przekształca macierz $\mathbf{A}$ w macierz jednostkową $\mathbf{I}_n$ wtedy i tylko wtedy, gdy ten sam ciąg przekształca macierz $\mathbf{I}_n$ w macierz $\mathbf{A}^{-1}$. Zatem mamy wygodną metodę wyznaczania macierzy odwrotnej (tożsamą z praktyczną realizacją metody Gaussa-Jordana rozwiązywania równania macierzowego $\mathbf{AX} = \mathbf{I}_n$).

Algorytm wyznaczania macierzy odwrotnej

Dla danej macierzy $\mathbf{A}$ wymiaru $n \times n$ wykonujemy następujące czynności:

1° Tworzymy macierz rozszerzoną $[\mathbf{A}|\mathbf{I}_n]$.

2° Macierz $[\mathbf{A}|\mathbf{I}_n]$ za pomocą operacji elementarnych na wierszach sprowadzamy do normalnej postaci schodkowej. Jeśli $[\mathbf{A}|\mathbf{I}_n]$ jest wierszowo równoważna macierzy postaci $[\mathbf{I}_n|\mathbf{P}]$, to macierz $\mathbf{A}$ jest odwracalna i $\mathbf{A}^{-1} = \mathbf{P}$. W innym przypadku macierz $\mathbf{A}$ nie jest odwracalna.

Przykład 98. Znaleźć (jeśli to możliwe) macierze odwrotne macierzy

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & -4 \\ 6 & 0 & 2 \end{bmatrix} \quad \text{i} \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 6 \\ 2 & 2 & 8 \end{bmatrix}.$$

Dla macierzy $\mathbf{A}$ mamy następujące równoważności

$$\begin{aligned} [\mathbf{A}|\mathbf{I}_3] &= \left[\begin{array}{ccc|ccc} 2 & 0 & 2 & 1 & 0 & 0 \\ 0 & 2 & -4 & 0 & 1 & 0 \\ 6 & 0 & 2 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 2 & 0 & 2 & 1 & 0 & 0 \\ 0 & 2 & -4 & 0 & 1 & 0 \\ 0 & 0 & -4 & -3 & 0 & 1 \end{array}\right] \\ &\sim \left[\begin{array}{ccc|ccc} 2 & 0 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 2 & 0 & 3 & 1 & -1 \\ 0 & 0 & -4 & -3 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 4 & 0 & 0 & -1 & 0 & 1 \\ 0 & 4 & 0 & 6 & 2 & -2 \\ 0 & 0 & 4 & 3 & 0 & -1 \end{array}\right], \end{aligned}$$

z których wynika, że $\mathbf{A}$ jest odwracalna i

$$\mathbf{A}^{-1} = \frac{1}{4} \begin{bmatrix} -1 & 0 & 1 \\ 6 & 2 & -2 \\ 3 & 0 & -1 \end{bmatrix}.$$

Natomiast dla macierzy $\mathbf{B}$ mamy

$$[\mathbf{B}|\mathbf{I}_3] = \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & 1 & 6 & 0 & 1 & 0 \\ 2 & 2 & 8 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 1 & 2 & 0 & -1 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 0 & 0 & 2 & -2 & 1 \end{array}\right],$$

a stąd widać, że macierz $[\mathbf{B}|\mathbf{I}_3]$ nie jest wierszowo równoważna macierzy $[\mathbf{I}_3|\mathbf{X}]$ (dla żadnej macierzy $\mathbf{X} \in R_{3\times 3}$). Zatem macierz $\mathbf{B}$ nie jest odwracalna.

5.5. Struktura rozwiązań układu równań liniowych

Definicja 5.5.1. Układ równań

$$\mathbf{Ax} = \mathbf{b}, \tag{5.4}$$

gdzie $\mathbf{A} \in K_{m\times n}$ i $\mathbf{b} \in K_{m\times 1}$, nazywamy *układem jednorodnym*, gdy $\mathbf{b} = \mathbf{0}$. W przeciwnym przypadku mówimy, że układ (5.4) jest *niejednorodny*.

$\mathbf{Ax} = \mathbf{0}$ – jednorodny
$\mathbf{Ax} = \mathbf{b}$ – niejednorodny, gdy $\mathbf{b} \neq \mathbf{0}$

Zauważmy, że zbiór rozwiązań układu jednorodnego $\mathbf{Ax} = \mathbf{0}$ jest niepusty, bo macierz zerowa $\mathbf{x}_0 = \mathbf{0} \in K_{n\times 1}$ zawsze jest rozwiązaniem układu $\mathbf{Ax} = \mathbf{0}$. Rozwiązanie $\mathbf{x}_0 = \mathbf{0}$ nazywamy *rozwiązaniem zerowym* (albo *rozwiązaniem trywialnym*) układu $\mathbf{Ax} = \mathbf{0}$. W wielu miejscach będziemy zainteresowani istnieniem niezerowych rozwiązań układu jednorodnego. Następujące obserwacje są natychmiastowymi konsekwencjami wniosku 5.1.1.

Wniosek 5.5.1. *Dla macierzy* $\mathbf{A} \in K_{m\times n}$ *mamy:*

(1) *układ jednorodny* $\mathbf{Ax} = \mathbf{0}$ *ma niezerowe rozwiązanie, gdy* $m < n$*;*

(2) *jeśli* $m = n$*, to układ jednorodny* $\mathbf{Ax} = \mathbf{0}$ *ma niezerowe rozwiązanie wtedy i tylko wtedy, gdy macierz* $\mathbf{A}$ *nie jest wierszowo równoważna macierzy jednostkowej* $\mathbf{I}_n$. □

Przykład 99. Jednorodny układ równań

$$\begin{cases} x_1 - x_2 + x_3 + x_4 = 0 \\ x_1 - 2x_2 + x_3 + 5x_4 = 0 \end{cases}$$

ma nieskończenie wiele rozwiązań, bo ma on mniej równań niż niewiadomych ($m = 2 < 4 = n$).

Przykład 100. Zbadać istnienie niezerowych rozwiązań jednorodnego układu równań liniowych

$$\begin{cases} \phantom{x_1 + {}} x_2 + x_3 + x_4 = 0 \\ x_1 \phantom{{}+x_2} - x_3 - x_4 = 0 \\ x_1 + x_2 \phantom{{}+x_3} - x_4 = 0 \\ x_1 + x_2 + x_3 \phantom{{}-x_4} = 0 \end{cases}.$$

Ponieważ macierz główna tego układu jest wierszowo równoważna macierzy jednostkowej,

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & -1 & -1 \\ 1 & 1 & 0 & -1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \overset{\substack{w_4 - w_3 \\ w_3 - w_2}}{\sim} \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & -1 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \overset{\substack{w_1 - w_4 \\ w_2 + w_4}}{\sim} \ldots \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \overset{w_4 - w_3}{\sim} \mathbf{I}_4,$$

więc układ ten ma tylko zerowe rozwiązanie.

Dla macierzy $\mathbf{A} \in K_{m\times n}$ i $\mathbf{b} \in K_{m\times 1}$ oznaczmy przez $\mathcal{R}_\mathbf{0}$ zbiór wszystkich rozwiązań jednorodnego układu $\mathbf{Ax} = \mathbf{0}$, a przez $\mathcal{R}_\mathbf{b}$ zbiór rozwiązań niejednorodnego układu równań $\mathbf{Ax} = \mathbf{b}$, czyli

$$\mathcal{R}_\mathbf{0} = \{\mathbf{x} \in K_{n\times 1} \colon \mathbf{Ax} = \mathbf{0}\}$$

i

$$\mathcal{R}_\mathbf{b} = \{\mathbf{x} \in K_{n\times 1} \colon \mathbf{Ax} = \mathbf{b}.\}$$

Wiemy już, że jednorodny układ równań $\mathbf{Ax} = \mathbf{0}$ zawsze ma rozwiązanie, więc zbiór $\mathcal{R}_\mathbf{0}$ jest niepusty. Teraz zauważmy, że jego elementy mają dodatkową własność.

$\mathcal{R}_\mathbf{0} \neq \emptyset$

Twierdzenie 5.5.1. *Jeśli* $\mathbf{x}_0$, $\mathbf{y}_0 \in \mathcal{R}_\mathbf{0}$, *to*

$$\alpha\mathbf{x}_0 + \beta\mathbf{y}_0 \in \mathcal{R}_\mathbf{0}$$

dla każdych α, $\beta \in K$.

Dowód. Jeśli $\mathbf{x}_0$, $\mathbf{y}_0 \in \mathcal{R}_\mathbf{0}$, to $\mathbf{Ax}_0 = \mathbf{0}$ i $\mathbf{Ay}_0 = \mathbf{0}$, więc z własności iloczynu macierzy dla każdych α, $\beta \in K$ mamy

$$\begin{aligned} \mathbf{A}(\alpha\mathbf{x}_0 + \beta\mathbf{y}_0) &= \mathbf{A}(\alpha\mathbf{x}_0) + \mathbf{A}(\beta\mathbf{y}_0) \\ &= \alpha(\mathbf{Ax}_0) + \beta(\mathbf{Ay}_0) \\ &= \alpha\mathbf{0} + \beta\mathbf{0} = \mathbf{0}, \end{aligned}$$

a to oznacza, że $\alpha\mathbf{x}_0 + \beta\mathbf{y}_0 \in \mathcal{R}_\mathbf{0}$. □

Obecnie udowodnimy, że zbiór $\mathcal{R}_\mathbf{b}$ (zbiór rozwiązań układu niejednorodnego $\mathbf{Ax} = \mathbf{b}$) jest jednoznacznie wyznaczony przez zbiór $\mathcal{R}_\mathbf{0}$ (zbiór rozwiązań układu jednorodnego $\mathbf{Ax} = \mathbf{0}$) i przez dowolne ustalone (czyli szczególne) rozwiązanie $\overline{\mathbf{x}}$ układu niejednorodnego $\mathbf{Ax} = \mathbf{b}$.

$\overline{\mathbf{x}}$ – szczególne rozwiązanie układu $\mathbf{Ax} = \mathbf{b}$

Twierdzenie 5.5.2. *Jeśli* $\overline{\mathbf{x}} \in \mathcal{R}_\mathbf{b}$, *to*

$$\mathcal{R}_\mathbf{b} = \mathcal{R}_\mathbf{0} + \{\overline{\mathbf{x}}\}.$$

Dowód. Uzasadnimy, że $\mathcal{R}_\mathbf{b} \subseteq \mathcal{R}_\mathbf{0} + \{\overline{\mathbf{x}}\}$ i $\mathcal{R}_\mathbf{0} + \{\overline{\mathbf{x}}\} \subseteq \mathcal{R}_\mathbf{b}$.

Załóżmy najpierw, że $\tilde{\mathbf{x}} \in \mathcal{R}_\mathbf{b}$. Wtedy $\mathbf{x}' = \tilde{\mathbf{x}} - \overline{\mathbf{x}} \in \mathcal{R}_\mathbf{0}$, bo

$$\mathbf{A}\mathbf{x}' = \mathbf{A}(\tilde{\mathbf{x}} - \overline{\mathbf{x}}) = \mathbf{A}\tilde{\mathbf{x}} - \mathbf{A}\overline{\mathbf{x}} = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

Stąd $\tilde{\mathbf{x}} = \mathbf{x}' + \overline{\mathbf{x}} \in \mathcal{R}_\mathbf{0} + \{\overline{\mathbf{x}}\}$ i to dowodzi, że $\mathcal{R}_\mathbf{b} \subseteq \mathcal{R}_\mathbf{0} + \{\overline{\mathbf{x}}\}$.

Z drugiej strony, jeśli $\mathbf{x}_0 + \overline{\mathbf{x}} \in \mathcal{R}_\mathbf{0} + \{\overline{\mathbf{x}}\}$, to $\mathbf{x}_0 \in \mathcal{R}_\mathbf{0}$ i dlatego

$$\mathbf{A}(\mathbf{x}_0 + \overline{\mathbf{x}}) = \mathbf{A}\mathbf{x}_0 + \mathbf{A}\overline{\mathbf{x}} = \mathbf{0} + \mathbf{b} = \mathbf{b}.$$

Zatem $\mathbf{x}_0 + \overline{\mathbf{x}} \in \mathcal{R}_\mathbf{b}$ i to dowodzi, że $\mathcal{R}_\mathbf{0} + \{\overline{\mathbf{x}}\} \subseteq \mathcal{R}_\mathbf{b}$. $\square$

Przykład 101. Rozwiązanie układu $\mathbf{A}\mathbf{x} = \mathbf{b}$, gdzie

$$\begin{cases} x_1 + 2x_2 + 2x_3 + 3x_4 = 7, \\ 2x_1 + x_2 + x_3 = 5, \\ 3x_1 + x_2 + x_3 - x_4 = 6, \\ x_1 - x_4 = 1, \end{cases}$$

przedstawić jako sumę rozwiązania szczególnego układu niejednorodnego $\mathbf{A}\mathbf{x} = \mathbf{b}$ i rozwiązania układu jednorodnego $\mathbf{A}\mathbf{x} = \mathbf{0}$.

Mamy

$$[\mathbf{A}|\mathbf{b}] = \left[\begin{array}{cccc|c} 1 & 2 & 2 & 3 & 7 \\ 2 & 1 & 1 & 0 & 5 \\ 3 & 1 & 1 & -1 & 6 \\ 1 & 0 & 0 & -1 & 1 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 1 \\ 0 & 1 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

i stąd wynika, że rozwiązaniem układu $\mathbf{A}\mathbf{x} = \mathbf{b}$ jest

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 + x_4 \\ 3 - x_3 - 2x_4 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} x_4 \\ -x_3 - 2x_4 \\ x_3 \\ x_4 \end{bmatrix}$$

dla każdych $x_3, x_4 \in R$. Rozwiązanie to jest sumą rozwiązania szczególnego $\overline{\mathbf{x}} = [1\,3\,0\,0]^T$ układu $\mathbf{A}\mathbf{x} = \mathbf{b}$ i rozwiązania $\mathbf{x}_0 = [x_4\ (-x_3 - 2x_4)\ x_3\ x_4]^T$ (gdzie $x_3, x_4 \in R$) układu jednorodnego $\mathbf{A}\mathbf{x} = \mathbf{0}$.

5.6. Ćwiczenia

1. Wskazać macierz elementarną $\mathbf{E}$ taką, że:

   (a) $\mathbf{E}\begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 3 & 3 \\ -2 & 4 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 3 & 3 \\ 0 & 8 & -2 & 9 \end{bmatrix}$;

   (b) $\mathbf{E}\begin{bmatrix} 2 & -3 & 1 & 3 \\ -2 & 7 & -4 & 2 \\ 2 & 1 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 2 & -3 & 1 & 3 \\ 2 & 1 & -2 & 8 \\ 2 & 1 & 0 & 3 \end{bmatrix}$.

2. Za pomocą macierzy elementarnych wyznaczyć macierz $\mathbf{A}$ taką, że:

   (a) $\mathbf{A}\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & -2 \\ 0 & -4 \end{bmatrix}$;

   (b) $\mathbf{A}\begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$.

3. Macierz $\mathbf{A} \in R_{3\times 3}$ jest iloczynem trzech macierzy elementarnych, $\mathbf{A} = \mathbf{E}_{23}(-1)\mathbf{E}_{21}\mathbf{E}_2(3)$. Wyznaczyć macierze $\mathbf{A}$, $\mathbf{A}^{-1}$ oraz $\mathbf{A} + 2\mathbf{A}^{-1}$.

4. Macierz $\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}$ przedstawić w postaci iloczynu macierzy elementarnych.

5. Rozwiązać następujące układy równań:

   (a) $\begin{cases} -3x_1 + x_2 + 4x_3 = -5, \\ x_1 + x_2 + x_3 = 2, \\ -2x_1 + x_3 = -3, \\ x_1 + x_2 - 2x_3 = 5; \end{cases}$

   (b) $\begin{cases} 2x_1 - 4x_2 + 2x_3 = 16, \\ 2x_1 - 4x_2 + 3x_3 = 19, \\ x_1 + 2x_2 + 3x_3 = 6, \\ - 2x_2 + x_3 = 7; \end{cases}$

   (c) $\begin{cases} x_1 + x_2 + x_3 + x_4 = 1, \\ x_1 + x_2 + 3x_3 + 3x_4 = 3, \\ x_1 + x_2 + 2x_3 + 3x_4 = 3, \\ x_1 + 3x_2 + 3x_3 + 3x_4 = 4; \end{cases}$

$$(d)\ \begin{cases} 2x_1 - x_2 + x_3 + 3x_4 = 7, \\ x_1 + 2x_2 + x_3 + 3x_4 = 12, \\ 2x_1 + 3x_2 + x_3 + 2x_4 = 19, \\ x_1 + x_2 + 2x_3 + 3x_4 = 11; \end{cases}$$

$$(e)\ \begin{cases} x_1 + x_2 + 3x_3 - 2x_4 = 0, \\ 2x_1 + 3x_2 + 7x_3 - 2x_4 = 9, \\ 3x_1 + 3x_2 + 13x_3 - 9x_4 = 1, \\ -2x_1 + x_2 - x_4 = 0; \end{cases}$$

$$(f)\ \begin{cases} -x_1 + 2x_2 + x_4 = 1, \\ -4x_1 + 8x_2 + 2x_3 + 10x_4 = 4, \\ -2x_1 + 4x_2 + 2x_3 + 8x_4 = 2; \end{cases}$$

$$(g)\ \begin{cases} 8x_1 + 6x_2 + 5x_3 + 2x_4 = 21, \\ 3x_1 + 3x_2 + 2x_3 + x_4 = 10, \\ 4x_1 + 2x_2 + 3x_3 + x_4 = 8, \\ 3x_1 + 5x_2 + x_3 + x_4 = 15, \\ 7x_1 + 4x_2 + 5x_3 + 2x_4 = 18; \end{cases}$$

$$(h)\ \begin{cases} 3x_1 + 3x_2 - 3x_3 - 3x_4 = 7, \\ -7x_1 + 3x_2 - x_3 + 2x_4 = 0, \\ -x_1 + 6x_2 + 2x_3 - x_4 = 3, \\ 2x_1 + 2x_2 - 2x_3 - 2x_4 = 5; \end{cases}$$

$$(i)\ \begin{cases} x_1 + x_2 - x_3 + 2x_4 = 3, \\ 2x_1 + 2x_2 - 2x_3 + 4x_4 = 6, \\ -3x_1 - 3x_2 + 3x_3 - 6x_4 = -9, \\ -2x_1 - 2x_2 + 2x_3 - 4x_4 = -6; \end{cases}$$

$$(j)\ \begin{cases} 2x_1 + x_2 - x_3 - x_4 + x_5 = 1, \\ x_1 - x_2 + x_3 + x_4 - 2x_5 = 0, \\ 3x_1 + 3x_2 - 3x_3 - 3x_4 + 4x_5 = 2, \\ 4x_1 + 5x_2 - 5x_3 - 5x_4 + 7x_5 = 3; \end{cases}$$

$$(k)\ \begin{cases} 2x_1 - x_2 + x_4 - 5x_5 = -1, \\ x_2 + 3x_4 + x_5 = 1, \\ 4x_1 - x_2 + x_3 + 6x_4 - 6x_5 = 1, \\ 6x_1 - 2x_2 + x_3 + 7x_4 - 11x_5 = 0; \end{cases}$$

$$(l)\ \begin{cases} x_1 + x_2 - x_3 + x_4 + x_5 = 0, \\ 2x_1 + x_2 + 2x_3 - x_4 - 2x_5 = 0, \\ 4x_1 + 3x_2 + x_4 = 0, \\ 5x_1 + 3x_2 + 3x_3 - x_4 - 3x_5 = 0, \\ x_1 - x_2 + 7x_3 - 5x_4 - 7x_5 = 0; \end{cases}$$

$$(m)\ \begin{cases} (1-n)x_1 + x_2 + \ldots + x_n = 0, \\ x_1 + (1-n)x_2 + \ldots + x_n = 0, \\ \vdots \qquad \ddots \qquad \vdots \\ x_1 + x_2 + \ldots + (1-n)x_n = 0. \end{cases}$$

6. Dla jakich wartości parametru k następujący układ ma niezerowe rozwiązanie:

$$(a)\ \begin{cases} kx_1 - x_2 + 3x_3 = 0, \\ x_1 + 2x_2 - x_3 = 0, \\ 2x_1 - x_2 - 2x_3 = 0; \end{cases}$$

$$(b)\ \begin{cases} x_1 + 2x_2 + 3x_3 = 0, \\ 2x_1 + kx_2 + 6x_3 = 0, \\ -x_1 + x_2 - kx_3 = 0. \end{cases}$$

7. Czy równanie

$$\begin{bmatrix} 2 & 4 & -2 \\ 3 & -1 & 11 \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

ma rozwiązanie dla każdych b_1, b_2, b_3?

8. Dla których b_1, b_2 i b_3 równanie

$$\begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & -1 \\ 2 & -1 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

ma rozwiązanie?

9. Dla jakich a, b i c układ równań

$$\begin{cases} 2x_1 - x_2 + 3x_3 = a \\ 3x_1 + x_2 - 5x_3 = b \\ -5x_1 - 5x_2 + 21x_3 = c \end{cases}$$

jest sprzeczny?

10. Dany jest układ równań

$$\begin{cases} x_1 + x_2 - x_3 = 2 \\ x_1 + 2x_2 + x_3 = 3 \\ x_1 + x_2 + (a^2-5)x_3 = a \end{cases}.$$

Wyznaczyć wszystkie wartości a, dla których układ ten: (a) nie ma rozwiązań; (b) ma dokładnie jedno rozwiązanie; (c) ma nieskończenie wiele rozwiązań.

11. Dane są macierze

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ 2 & -3 & 2 \\ -1 & 12 & -7 \end{bmatrix} \quad \text{i} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

(a) Rozwiązać równanie $\mathbf{Ax} = \mathbf{0}$. (b) Wszystkie rozwiązania równania $\mathbf{Ax} = \mathbf{b}$ zapisać w postaci $\mathbf{s} + \mathbf{n}$, gdzie $\mathbf{s}$ jest jednym konkretnym rozwiązaniem równania $\mathbf{Ax} = \mathbf{b}$. (c) Wskazać taki wektor $\mathbf{c}$, że równanie $\mathbf{Ax} = \mathbf{c}$ nie ma rozwiązania. Uzasadnić swój wybór.

12. Wyznaczyć liczby x_1, x_2 i x_3 takie, że

$$\begin{bmatrix} -6 \\ 37 \\ -59 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ -3 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ -8 \\ 12 \end{bmatrix}.$$

13. Rozwiązać następujące równania macierzowe:

$$(a)\ \begin{bmatrix} 7 & -2 & 1 \\ -4 & 5 & -3 \\ 5 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{bmatrix} = \begin{bmatrix} 11 & 1 \\ 0 & -4 \\ 7 & 5 \end{bmatrix};$$

$$(b)\ \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 0 & 1 \end{bmatrix} \mathbf{X} = \begin{bmatrix} 2 & 3 & 0 & 4 \\ 1 & 0 & -3 & -1 \\ 1 & 2 & 1 & 3 \end{bmatrix};$$

$$(c)\ \begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & 0 \\ 2 & 1 & -1 \end{bmatrix} \mathbf{X} - \mathbf{X} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{bmatrix};$$

$$(d)\ \mathbf{X} \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 5 & 7 \end{bmatrix};$$

$$(e)\ \mathbf{X} \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 5 \\ 3 & 0 & 9 \end{bmatrix};$$

$$(f)\ \mathbf{X} \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 6 \\ -2 & 2 & 12 \end{bmatrix}.$$

14. Metodą Gaussa-Jordana wyznaczyć macierz odwrotną $\mathbf{A}_i^{-1}$ macierzy $\mathbf{A}_i$ ($i = 1, 2, 3, 4$), gdy:

$$\mathbf{A}_1 = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & -1 & 2 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 2 & 0 & 5 & 2 \\ 1 & 1 & 1 & -1 \\ -1 & 0 & 3 & 0 \\ 1 & 0 & 5 & 2 \end{bmatrix};$$

$$\mathbf{A}_3 = \begin{bmatrix} 1 & 4 & 2 \\ 1 & 5 & 2 \\ 2 & 6 & 2 \end{bmatrix}, \quad \mathbf{A}_4 = \begin{bmatrix} 2 & 6 & 0 & 5 \\ 6 & 21 & 8 & 17 \\ 4 & 12 & -4 & 13 \\ 0 & -3 & -12 & 2 \end{bmatrix}.$$

15. Wyznaczyć macierze odwrotne dla następujących czterech macierzy wymiaru $n \times n$:

$$(a) \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}; (c) \begin{bmatrix} 1 & 2 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 3 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 4 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix};$$

$$(b) \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 2 & 0 & \cdots & 0 \\ 1 & 2 & 3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & n \end{bmatrix}; (d) \begin{bmatrix} 1 & 2 & 3 & \cdots & n \\ 0 & 1 & 2 & \cdots & n-1 \\ 0 & 0 & 1 & \cdots & n-2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

16. Rozwiązać następujące układy równań macierzowych:

$$(a) \begin{cases} \mathbf{X} + 2\mathbf{Y} = \begin{bmatrix} 2 & 4 \\ -2 & 6 \end{bmatrix}, \\ 3\mathbf{X} + 4\mathbf{Y} = \begin{bmatrix} 0 & 4 \\ 8 & 16 \end{bmatrix}; \end{cases}$$

$$(b) \begin{cases} \mathbf{X} + \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \mathbf{Y} = \begin{bmatrix} 0 & -1 \\ 2 & 1 \end{bmatrix}, \\ \begin{bmatrix} 3 & 2 \\ 0 & -1 \end{bmatrix} \mathbf{X} + \mathbf{Y} = \begin{bmatrix} 0 & 1 \\ 6 & 3 \end{bmatrix}. \end{cases}$$

17. Rozwiązać następujące układy równań liniowych:

$$(a) \begin{cases} z + jw = 1 \\ jz + w = 1+j \end{cases};$$

$$(b) \begin{cases} (1+j)z - jw = 3+j \\ (2+j)z + (2-j)w = 2j \end{cases};$$

$$(c) \begin{cases} jz + (1+j)w = 3+j \\ (1+j)\overline{z} - (6+j)\overline{w} = 4 \end{cases};$$

$$(d) \begin{cases} (1+j)z + (1-3j)w = -2+14j \\ 3z + (3-9j)w = 9+36j \end{cases}.$$

18. Wykazać, że jeśli macierz $\mathbf{A}$ jest wierszowo równoważna macierzy odwracalnej $\mathbf{B}$, to także macierz $\mathbf{A}$ jest odwracalna.

19. Znaleźć wszystkie macierze przemienne z macierzą $\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$, czyli wszystkie macierze $\mathbf{X}$, dla których $\mathbf{AX} = \mathbf{XA}$.

20. Niech $\mathbf{A}$ będzie macierzą odwracalną i niech $\mathbf{B}$ będzie macierzą powstałą z $\mathbf{A}$ w wyniku przestawienia miejscami dwóch wierszy. (a) Pokazać, że macierz $\mathbf{B}$ jest odwracalna. (b) Jaki jest związek macierzy $\mathbf{B}^{-1}$ z macierzą $\mathbf{A}^{-1}$?

21. Wpisując TAK albo NIE, stwierdzić prawdziwość każdego z następujących zdań:

$\boxed{\scriptstyle 1 \quad}$ Każdy układ równań liniowych, w którym liczba równań jest równa liczbie niewiadomych ma tylko jedno rozwiązanie.

$\boxed{\scriptstyle 2 \quad}$ Każdy układ równań liniowych, w którym liczba równań jest równa liczbie niewiadomych ma co najmniej jedno rozwiązanie.

$\boxed{\scriptstyle 3 \quad}$ Układ równań liniowych, w którym jest więcej równań niż niewiadomych ma nieskończenie wiele rozwiązań.

$\boxed{\scriptstyle 4 \quad}$ Układ równań liniowych, w którym jest mniej równań niż niewiadomych może nie mieć rozwiązania.

$\boxed{\scriptstyle 5 \quad}$ Układ równań liniowych, w którym jest mniej równań niż niewiadomych ma nieskończenie wiele rozwiązań.

$\boxed{\scriptstyle 6 \quad}$ Każda macierz jest wierszowo równoważna macierzy schodkowej.

$\boxed{\scriptstyle 7 \quad}$ Iloczyn dwóch macierzy elementarnych jest macierzą elementarną.

$\boxed{\scriptstyle 8 \quad}$ Jeśli macierze $[\mathbf{A}|\mathbf{b}]$ i $[\mathbf{B}|\mathbf{c}]$ są wierszowo równoważne, to układy równań liniowych $\mathbf{Ax} = \mathbf{b}$ i $\mathbf{Bx} = \mathbf{c}$ mają identyczne zbiory rozwiązań.

$\boxed{\scriptstyle 9 \quad}$ Układ równań liniowych $\mathbf{Ax} = \mathbf{b}$, w którym $\mathbf{A}$ jest macierzą kwadratową, ma dokładnie jedno rozwiązanie wtedy i tylko wtedy, gdy $\mathbf{A}$ jest wierszowo równoważna macierzy jednostkowej.

$\boxed{\scriptstyle 10 \quad}$ Układ równań liniowych $\mathbf{Ax} = \mathbf{b}$ ma nieskończenie wiele rozwiązań wtedy i tylko wtedy, gdy $\mathbf{A}$ jest wierszowo równoważna macierzy schodkowej, w której pewna kolumna nie zawiera wiodącej jedynki.

$\boxed{\scriptstyle 11 \quad}$ Jeśli układ równań $\mathbf{Ax} = \mathbf{b}$, gdzie $[\mathbf{A}|\mathbf{b}] \in R_{n\times(n+1)}$, ma rozwiązanie dla każdego $\mathbf{b} \in R_{n\times 1}$, to układ $\mathbf{Ax} = \mathbf{b}$ ma dokładnie jedno rozwiązanie dla każdego $\mathbf{b}$.

$\boxed{\scriptstyle 12 \quad}$ Jeśli układ równań $\mathbf{Ax} = \mathbf{0}$ ma tylko zerowe rozwiązanie i $\mathbf{A} \in R_{n\times n}$, to układ $\mathbf{Ax} = \mathbf{b}$ ma dokładnie jedno rozwiązanie dla każdego $\mathbf{b} \in R_{n\times 1}$.

$\boxed{\scriptstyle 13 \quad}$ Jeśli $\mathbf{x}_0$ i $\mathbf{y}_0$ są rozwiązaniami układu $\mathbf{Ax} = \mathbf{b}$, to także $\mathbf{x}_0 + \mathbf{y}_0$ jest rozwiązaniem układu $\mathbf{Ax} = \mathbf{b}$.

Rozdział 6

WYZNACZNIKI

6.1. Definicja i pierwsze własności wyznacznika

W rozdziale tym każdej macierzy kwadratowej

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

o elementach z ciała K przyporządkowujemy element z ciała K nazywany wyznacznikiem (lub determinantem) macierzy $\mathbf{A}$ i oznaczany symbolem

det $\mathbf{A} = |\mathbf{A}|$

$$|\mathbf{A}|, \ \det \mathbf{A}, \ \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad \text{lub} \quad \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

W literaturze matematycznej można spotkać różne sposoby definiowania wyznacznika. Tu przedstawiamy indukcyjną definicję wyznacznika, definicję wyznacznika macierzy kwadratowej stopnia n za pomocą wyznaczników z n macierzy kwadratowych stopnia $n-1$. Definicja ta może nie jest matematycznie najładniejsza, ale jest ona dostatecznie praktyczna. W tej definicji korzystać będziemy z następującego oznaczenia: dla macierzy kwadratowej $\mathbf{A}$ stopnia $n > 1$ i dla liczb naturalnych i, j $(1 \leqslant i, j \leqslant n)$ przez $\mathbf{A}_{ij}$ oznaczamy macierz kwadratową stopnia $n-1$ powstałą z macierzy $\mathbf{A}$ przez wykreślenie z niej i-tego wiersza oraz j-tej kolumny. Przykładowo, jeśli

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 4 & 0 \\ 2 & 3 & 5 & 1 \\ 0 & 7 & 2 & 4 \\ 1 & 3 & 2 & 1 \end{bmatrix},$$

to macierzą powstałą w wyniku wykreślenia drugiego wiersza i trzeciej kolumny jest

$$\mathbf{A}_{2\,3} = \begin{bmatrix} 1 & 2 & \cancel{4} & 0 \\ \cancel{2} & \cancel{3} & \cancel{5} & \cancel{1} \\ 0 & 7 & \cancel{2} & 4 \\ 1 & 3 & \cancel{2} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 7 & 4 \\ 1 & 3 & 1 \end{bmatrix}.$$

Definicja 6.1.1. *Wyznacznikiem* macierzy kwadratowej $\mathbf{A}$ nazywamy liczbę $\det \mathbf{A}$, gdzie:

Definicja wyznacznika

1° jeśli $\mathbf{A} \in K_{1\times 1}$, $\mathbf{A} = [a_{11}]$, to $\det \mathbf{A} = a_{11}$;

2° jeśli $\mathbf{A} = [a_{ij}] \in K_{n\times n}$ i $n > 1$, to

$$\det \mathbf{A} = \sum_{j=1}^{n} a_{1j} (-1)^{1+j} \det \mathbf{A}_{1j}, \tag{6.1}$$

czyli

$$|\mathbf{A}| = a_{11}(-1)^{1+1}|\mathbf{A}_{11}| + a_{12}(-1)^{1+2}|\mathbf{A}_{12}| + \ldots + a_{1n}(-1)^{1+n}|\mathbf{A}_{1n}|.$$

Prawą stronę równości (6.1) nazywamy *rozwinięciem Laplace'a wyznacznika* $\det \mathbf{A}$ względem elementów pierwszego wiersza macierzy $\mathbf{A}$. Zgodnie z tą definicją wyznacznik macierzy kwadratowej $\mathbf{A}$ stopnia n obliczamy za pomocą n wyznaczników macierzy kwadratowych $\mathbf{A}_{1j}$ stopnia $n-1$. Dla macierzy kwadratowych małego stopnia można podać proste wzory wyznaczania wartości wyznacznika.

Przykład 102. Wobec (6.1) dla wyznacznika macierzy kwadratowej stopnia 2 mamy

$$\begin{aligned}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} &= a_{11}(-1)^{1+1}|\mathbf{A}_{11}| + a_{12}(-1)^{1+2}|\mathbf{A}_{12}| \\ &= a_{11}|[a_{22}]| - a_{12}|[a_{21}]| = a_{11}a_{22} - a_{12}a_{21}.\end{aligned}$$

Stąd

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & \\ & a_{22} \end{vmatrix} - \begin{vmatrix} & a_{12} \\ a_{21} & \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}, \tag{6.2}$$

co oznacza, że wyznacznik macierzy kwadratowej stopnia dwa jest równy różnicy iloczynu elementów stojących na głównej przekątnej i iloczynu elementów stojących na "drugiej" przekątnej tej macierzy. Przykładowo, jeśli

$$\mathbf{A} = \begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix}, \quad \text{to} \quad \det \mathbf{A} = 2 \cdot 5 - (-1) \cdot 3 = 13.$$

Przykład 103. Wobec (6.1) i (6.2) dla wyznacznika macierzy kwadratowej stopnia trzy mamy

$$\begin{aligned}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}(-1)^{1+1}|\mathbf{A}_{11}| + a_{12}(-1)^{1+2}|\mathbf{A}_{12}| + a_{13}(-1)^{1+3}|\mathbf{A}_{13}| \\ &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.\end{aligned}$$

Zatem otrzymaliśmy następujący wzór na obliczanie wyznacznika macierzy kwadratowej stopnia trzy:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{aligned} &a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &- a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \end{aligned} \tag{6.3}$$

Warto zaobserwować, że powyższy wzór na wyznacznik macierzy kwadratowej stopnia trzy można uzyskać za pomocą tzw. schematu Sarrusa. W tym celu na prawo od wyznacznika kopiujemy pierwszą i drugą kolumnę macierzy. Łatwo teraz zauważyć, że wyznacznik jest sumą iloczynów elementów stojących na prostych p_1, p_2 i p_3 oraz opatrzonych znakiem minus iloczynów elementów stojących na prostych p_4, p_5 i p_6,

$$\begin{array}{ccc|cc} p_1 & p_2 & p_3 & p_4 & p_5 \;\; p_6 \\ \hline a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \\ - & - & - & + & + \;\; + \end{array} = \begin{aligned} &a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &-a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \end{aligned}$$

Schemat Sarrusa

Zgodnie z (6.3) (i schematem Sarrusa) mamy

$$\begin{vmatrix} 1 & 2 & 3 \\ -2 & 3 & 1 \\ 3 & 5 & -1 \end{vmatrix} = 1{\cdot}3{\cdot}(-1) + 2{\cdot}1{\cdot}3 + 3{\cdot}(-2){\cdot}5 - 3{\cdot}3{\cdot}3 - 1{\cdot}5{\cdot}1 - (-1){\cdot}2(-2) = -63.$$

Przykład 104. Korzystajc z równości (6.1) i ze schematu Sarrusa szybko można obliczyć wyznacznik macierzy kwadratowej stopnia cztery. Przykładowo mamy

$$\begin{aligned} \begin{vmatrix} 0 & 3 & 2 & 0 \\ 3 & 3 & 0 & 0 \\ -4 & 2 & 1 & 2 \\ 3 & -1 & -2 & 2 \end{vmatrix} &= a_{11}|\mathbf{A}_{11}| - a_{12}|\mathbf{A}_{12}| + a_{13}|\mathbf{A}_{13}| - a_{14}|\mathbf{A}_{14}| \\ &= 0|\mathbf{A}_{11}| - 3|\mathbf{A}_{12}| + 2|\mathbf{A}_{13}| - 0|\mathbf{A}_{14}| \\ &= -3\begin{vmatrix} 3 & 0 & 0 \\ -4 & 1 & 2 \\ 3 & -2 & 2 \end{vmatrix} + 2\begin{vmatrix} 3 & 3 & 0 \\ -4 & 2 & 2 \\ 3 & -1 & 2 \end{vmatrix} \\ &= -3 \cdot 18 + 2 \cdot 60 = 66. \end{aligned}$$

Nasza definicja wyznacznika wyróżnia pierwszy wiersz macierzy. Okazuje się (ale tego tu nie udowodnimy), że to wyróżnienie nie jest niczym usprawiedliwione, bo mamy następujące twierdzenie.

Twierdzenie 6.1.1 (Laplace). *Niech* $\mathbf{A} = [a_{ij}]$ *będzie macierzą wymiaru* $n \times n$ *i niech* i *oraz* j *będą liczbami ze zbioru* $\{1, \ldots, n\}$. *Wtedy*

$$\det \mathbf{A} = a_{i1}(-1)^{i+1}|\mathbf{A}_{i1}| + a_{i2}(-1)^{i+2}|\mathbf{A}_{i2}| + \ldots + a_{in}(-1)^{i+n}|\mathbf{A}_{in}| \tag{6.4}$$

Rozwinięcia Laplace'a

oraz

$$\det \mathbf{A} = a_{1j}(-1)^{1+j}|\mathbf{A}_{1j}| + a_{2j}(-1)^{2+j}|\mathbf{A}_{2j}| + \ldots + a_{nj}(-1)^{n+j}|\mathbf{A}_{nj}|. \tag{6.5}$$

Prawe strony równości (6.4) i (6.5) nazywa się odpowiednio *rozwinięciem Laplace'a wyznacznika* $\det \mathbf{A}$ względem i-tego wiersza oraz j-tej kolumny macierzy $\mathbf{A}$. Zgodnie z tym twierdzeniem, obliczając wyznacznik macierzy $\mathbf{A}$, możemy posłużyć się rozwinięciem względem dowolnego wiersza lub dowolnej kolumny. W praktyce, dla skrócenia czasu obliczeń, warto posługiwać się rozwinięciem względem wiersza lub kolumny z dużą ilością zer. W skrajnym przypadku, gdy macierz ma zerowy wiersz lub zerową kolumnę, to z (6.4) lub (6.5) natychmiast otrzymujemy następującą ważną obserwację.

Twierdzenie 6.1.2. *Jeśli w macierzy kwadratowej* $\mathbf{A}$ *jest zerowy wiersz lub zerowa kolumna, to* $\det \mathbf{A} = 0$.

Dowód. Niech $\mathbf{A} = [a_{ij}]$ będzie macierzą kwadratową stopnia n, w której i-ty wiersz jest zerowy. Wtedy $a_{i1} = a_{i2} = \ldots = a_{in} = 0$ i dlatego wobec (6.4) mamy

$$\begin{aligned} \det \mathbf{A} &= a_{i1}(-1)^{i+1}|\mathbf{A}_{i1}| + a_{i2}(-1)^{i+2}|\mathbf{A}_{i2}| + \ldots + a_{in}(-1)^{i+n}|\mathbf{A}_{in}| \\ &= 0(-1)^{i+1}|\mathbf{A}_{i1}| + 0(-1)^{i+2}|\mathbf{A}_{i2}| + \ldots + (-1)^{i+n}|\mathbf{A}_{in}| = 0. \end{aligned}$$

Analogicznie dowodzi się, że $\det \mathbf{A} = 0$, gdy macierz $\mathbf{A}$ ma zerową kolumnę. □

Przykład 105. Obliczyć wyznaczniki macierzy

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 & 4 \\ 7 & 3 & 0 & 3 \\ 1 & 2 & 0 & 5 \\ 1 & 4 & 0 & 0 \end{bmatrix} \quad \text{i} \quad \mathbf{B} = \begin{bmatrix} 7 & 3 & 4 & 2 & 1 \\ 6 & 0 & 3 & 0 & 2 \\ 2 & 1 & 2 & 2 & 4 \\ 4 & 0 & 5 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \end{bmatrix}.$$

Ponieważ macierz $\mathbf{A}$ ma zerową kolumnę, więc natychmiast $\det \mathbf{A} = 0$. Licząc wyznacznik macierzy $\mathbf{B}$ trzykrotnie możemy skorzystać z zalet rozwinięcia wyznacznika względem wiersza (u nas obramowanego) z największą liczbą zer. Mamy

$$
\begin{aligned}
|\mathbf{B}| &= \begin{vmatrix} 7 & 3 & 4 & 2 & 1 \\ 6 & 0 & 3 & 0 & 2 \\ 2 & 1 & 2 & 2 & 4 \\ 4 & 0 & 5 & 0 & 0 \\ \boxed{0 \;\; 0 \;\; 3 \;\; 0 \;\; 0} \end{vmatrix} = 3 \cdot (-1)^{5+3} \begin{vmatrix} 7 & 3 & 2 & 1 \\ 6 & 0 & 0 & 2 \\ 2 & 1 & 2 & 4 \\ \boxed{4 \;\; 0 \;\; 0 \;\; 0} \end{vmatrix} \\
&= 3 \cdot 4 \cdot (-1)^{4+1} \begin{vmatrix} 3 & 2 & 1 \\ \boxed{0 \;\; 0 \;\; 2} \\ 1 & 2 & 4 \end{vmatrix} = -12 \cdot 2 \cdot (-1)^{2+3} \begin{vmatrix} 3 & 2 \\ 1 & 2 \end{vmatrix} = 24(6-2) = 96.
\end{aligned}
$$

Definicja 6.1.2. Macierz kwadratową $\mathbf{A} = [a_{ij}]$ stopnia n nazywa się *macierzą trójkątną* dolną (górną), gdy

$$a_{ij} = 0 \quad \text{dla} \quad 1 \leqslant i < j \leqslant n \ \ (1 \leqslant j < i \leqslant n).$$

Macierz trójkątna dolna

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

Korzystając z definicji wyznacznika lub z twierdzenia 6.1.1, indukcyjnie ze względu na stopień macierzy, łatwo wykazuje się prawdziwość następującego twierdzenia o wyznaczniku macierzy trójkątnej.

Twierdzenie 6.1.3. *Wyznacznik macierzy trójkątnej jest równy iloczynowi elementów stojących na jej głównej przekątnej,*

$$\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ & & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22}\ldots a_{nn} \quad i \quad \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ & & \ddots & \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22}\ldots a_{nn}.$$

W szczególności dla macierzy diagonalnej, więc także dla macierzy jednostkowej, mamy

$$\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22}\ldots a_{nn} \quad i \quad |\mathbf{I}_n| = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & 1 \end{vmatrix} = 1. \ \square$$

Przykład 106. Wobec poprzedniego twierdzenia mamy

$$\begin{vmatrix} 5 & 0 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ 1 & 3 & 3 & 0 \\ 2 & 3 & 1 & 2 \end{vmatrix} = 5 \cdot 4 \cdot 3 \cdot 2,$$

bo jest to wyznacznik macierzy trójkątnej.

Ustalimy teraz związek między wyznacznikiem macierzy kwadratowej $\mathbf{A}$ i wyznacznikiem jej transpozycji $\mathbf{A}^T$. Zaczynamy od prostego przykładu.

Przykład 107. Obliczyć wyznaczniki macierzy $\mathbf{A}$ i $\mathbf{A}^T$, gdy

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

Wyznacznik macierzy $\mathbf{A}$ możemy obliczyć za pomocą rozwinięcia względem drugiego wiersza,

$$|\mathbf{A}| = \begin{vmatrix} 2 & 3 & 5 & 1 \\ \boxed{2 \quad 0 \quad 3 \quad 0} \\ 1 & -1 & 2 & 4 \\ 3 & 2 & 2 & 1 \end{vmatrix} = 2(-1)^{2+1}\begin{vmatrix} 3 & 5 & 1 \\ -1 & 2 & 4 \\ 2 & 2 & 1 \end{vmatrix} + 3(-1)^{2+3}\begin{vmatrix} 2 & 3 & 1 \\ 1 & -1 & 4 \\ 3 & 2 & 1 \end{vmatrix}$$
$$= (-2)\cdot 21 + (-3)\cdot 20 = -102.$$

Natomiast wyznacznik macierzy $\mathbf{A}^T$ liczymy poprzez rozwinięcie względem drugiej kolumny,

$$|\mathbf{A}^T| = \begin{vmatrix} 2 & \boxed{2} & 1 & 3 \\ 3 & \boxed{0} & -1 & 2 \\ 5 & \boxed{3} & 2 & 2 \\ 1 & \boxed{0} & 4 & 1 \end{vmatrix} = 2(-1)^{1+2}\begin{vmatrix} 3 & -1 & 2 \\ 5 & 2 & 2 \\ 1 & 4 & 1 \end{vmatrix} + 3(-1)^{3+2}\begin{vmatrix} 2 & 1 & 3 \\ 3 & -1 & 2 \\ 1 & 4 & 1 \end{vmatrix}$$
$$= (-2)\cdot 21 + (-3)\cdot 20 = -102.$$

Nie jest sprawą przypadku, że dla macierzy $\mathbf{A}$ z ostatniego przykładu mamy $|\mathbf{A}| = |\mathbf{A}^T|$. Pokażemy teraz, że macierz kwadratowa i jej transpozycja zawsze mają identyczne wyznaczniki.

Twierdzenie 6.1.4. *Dla każdej macierzy kwadratowej* $\mathbf{A}$ *jest*

$|\mathbf{A}^T| = |\mathbf{A}|$

$$\det \mathbf{A}^T = \det \mathbf{A}. \tag{6.6}$$

Dowód. Twierdzenie jest trywialne dla macierzy kwadratowych stopnia 1. Załóżmy jego prawdziwość dla macierzy kwadratowych stopnia $n-1$, gdzie $n \geqslant 2$ jest ustaloną liczbą naturalną. Niech teraz $\mathbf{A} = [a_{ij}]$ będzie macierzą kwadratową stopnia n i niech $\mathbf{B} = [b_{ij}]$ będzie jej transpozycją. Rozwijając wyznaczniki macierzy $\mathbf{A}$ i $\mathbf{B}$ odpowiednio względem pierwszego wiersza i pierwszej kolumny, otrzymujemy

$$\det \mathbf{A} = \sum_{i=1}^{n} a_{1i}(-1)^{1+i}|\mathbf{A}_{1i}| \quad \text{i} \quad \det \mathbf{B} = \sum_{i=1}^{n} b_{i1}(-1)^{i+1}|\mathbf{B}_{i1}|. \tag{6.7}$$

Ponieważ $\mathbf{B} = \mathbf{A}^T$, więc dla $i, j \in \{1, \ldots, n\}$ jest $b_{ij} = a_{ji}$ oraz

$$\mathbf{B}_{i1} = \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ \vdots & \vdots & & \vdots \\ b_{i1} & b_{i2} & \ldots & b_{in} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \ldots & b_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ \vdots & \vdots & & \vdots \\ a_{1i} & a_{2i} & \ldots & a_{ni} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{bmatrix}$$
$$= \begin{bmatrix} a_{11} & \ldots & a_{1i} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2i} & \ldots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \ldots & a_{ni} & \ldots & a_{nn} \end{bmatrix}^T = (\mathbf{A}_{1i})^T.$$

Dodatkowo, ponieważ $\mathbf{A}_{1i}$ jest macierzą kwadratową stopnia $n-1$, więc z założenia indukcyjnego $|\mathbf{A}_{1i}| = |\left(\mathbf{A}_{1i}\right)^T| = |\mathbf{B}_{i1}|$. Stąd i z (6.7) wynika (6.6), bo mamy

$$\det \mathbf{A}^T = \det \mathbf{B} = \sum_{i=1}^{n} b_{i1}(-1)^{i+1}|\mathbf{B}_{i1}| = \sum_{i=1}^{n} a_{1i}(-1)^{1+i}|\mathbf{A}_{1i}| = \det \mathbf{A}. \ \square$$

Konsekwencją twierdzenia 6.1.4 jest następujące metatwierdzenie o wyznacznikach: każde twierdzenie o wyznacznikach, w którym występuje słowo "kolumna" pozostaje prawdziwe po zamianie słowa "kolumna" na słowo "wiersz". Z tego też powodu w dowodach następnych twierdzeń zajmujemy się tylko ich wersjami kolumnowymi.

Twierdzenie 6.1.5. *Jeśli macierz* $\mathbf{B}$ *powstaje z macierzy kwadratowej* A *w wyniku przestawienia miejscami dwóch kolumn (albo dwóch wierszy), to*

$$\det \mathbf{B} = -\det \mathbf{A}.$$

Dowód. Łatwo sprawdza się prawdziwość tezy dla macierzy kwadratowych stopnia 2, bo wobec (6.2) mamy

$$\begin{vmatrix} b & a \\ d & c \end{vmatrix} = bc - ad = -(ad - bc) = -\begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Załóżmy teraz prawdziwość tezy twierdzenia dla macierzy kwadratowych stopnia $n-1$, gdzie $n > 2$ jest ustaloną liczbą naturalną. Niech $\mathbf{A}$ będzie dowolną macierzą kwadratową stopnia n i niech $\mathbf{B}$ będzie macierzą powstałą z $\mathbf{A}$ w wyniku przestawienia miejscami k-tej i l-tej kolumny, $k \neq l$, powiedzmy

$$\mathbf{A} = [a_{ij}] = [\,\mathbf{a}_1 \ldots \mathbf{a}_k \ldots \mathbf{a}_l \ldots \mathbf{a}_n\,]$$

i

$$\mathbf{B} = [b_{ij}] = [\,\mathbf{a}_1 \ldots \mathbf{a}_l \ldots \mathbf{a}_k \ldots \mathbf{a}_n\,].$$

Weźmy pod uwagę dowolną liczbę $s \in \{1, \ldots, n\} - \{k, l\}$ i rozważmy rozwinięcia wyznaczników macierzy $\mathbf{B}$ i $\mathbf{A}$ względem ich s-tych kolumn,

$$\det \mathbf{B} = \sum_{i=1}^{n} b_{is}(-1)^{i+s}|\mathbf{B}_{is}| \quad \text{i} \quad \det \mathbf{A} = \sum_{i=1}^{n} a_{is}(-1)^{i+s}|\mathbf{A}_{is}|. \tag{6.8}$$

Ponieważ $b_{is} = a_{is}$, a macierze $\mathbf{B}_{is}$ oraz $\mathbf{A}_{is}$ są wymiaru $(n-1)\times(n-1)$ i różnią się tylko kolejnością dwóch kolumn,

$$\mathbf{A}_{is} = \begin{bmatrix} a_{11} & \ldots & a_{1k} & \ldots & a_{1s} & \ldots & a_{1l} & \ldots & a_{1n} \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{i1} & \ldots & a_{ik} & \ldots & a_{is} & \ldots & a_{il} & \ldots & a_{in} \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{n1} & \ldots & a_{nk} & \ldots & a_{ns} & \ldots & a_{nl} & \ldots & a_{nn} \end{bmatrix}$$

i

$$\mathbf{B}_{is} = \begin{bmatrix} a_{11} & \ldots & a_{1l} & \ldots & a_{1s} & \ldots & a_{1k} & \ldots & a_{1n} \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{i1} & \ldots & a_{il} & \ldots & a_{is} & \ldots & a_{ik} & \ldots & a_{in} \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{n1} & \ldots & a_{nl} & \ldots & a_{ns} & \ldots & a_{nk} & \ldots & a_{nn} \end{bmatrix},$$

więc wobec założenia indukcyjnego jest $|\mathbf{B}_{is}| = -|\mathbf{A}_{is}|$ i dlatego wobec (6.8) mamy

$$\det \mathbf{B} = \sum_{i=1}^{n} b_{is}(-1)^{i+s}|\mathbf{B}_{is}| = \sum_{i=1}^{n} a_{is}(-1)^{i+s}(-|\mathbf{A}_{is}|) = -\det \mathbf{A}. \ \square$$

Przykład 108. Obliczyć wyznacznik macierzy $\mathbf{A} = \begin{bmatrix} 2 & 4 & 0 & 2 \\ 2 & 3 & 0 & 0 \\ 4 & 5 & 3 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$.

Przestawiając pierwszy wiersz z czwartym, przekształcamy macierz $\mathbf{A}$ w macierz trójkątną i wobec twierdzenia 6.1.5 oraz twierdzenia 6.1.3 mamy

$$|\mathbf{A}| = -\begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 4 & 5 & 3 & 0 \\ 2 & 4 & 0 & 2 \end{vmatrix} = -18.$$

Twierdzenie 6.1.6. *Jeśli macierz kwadratowa* $\mathbf{A}$ *ma dwie równe kolumny* (*dwa równe wiersze*), *to*

$$\det \mathbf{A} = 0.$$

Dowód. Załóżmy, że macierz $\mathbf{A} = [\,\mathbf{a}_1 \dots \mathbf{a}_l \dots \mathbf{a}_k \dots \mathbf{a}_n\,]$ ma dwie równe kolumny, $\mathbf{a}_l = \mathbf{a}_k$ i $l \neq k$. Niech $\mathbf{B}$ będzie macierzą powstałą z $\mathbf{A}$ w wyniku przestawienia miejscami kolumn $\mathbf{a}_l$ i $\mathbf{a}_k$. Wtedy $\det \mathbf{B} = -\det \mathbf{A}$ wobec twierdzenia 6.1.5. Z drugiej strony $\mathbf{B} = \mathbf{A}$, więc $\det \mathbf{A} = \det \mathbf{B} = -\det \mathbf{A}$ i dlatego $\det \mathbf{A} = 0$. □

Twierdzenie 6.1.7. *Niech* $\mathbf{A} = [a_{ij}]$ *będzie macierzą kwadratową stopnia* n. *Jeśli* i *oraz* j *są różnymi liczbami ze zbioru* $\{1, \dots, n\}$, *to*

$$a_{i1}(-1)^{j+1}|\mathbf{A}_{j1}| + a_{i2}(-1)^{j+2}|\mathbf{A}_{j2}| + \ldots + a_{in}(-1)^{j+n}|\mathbf{A}_{jn}| = 0 \tag{6.9}$$

oraz

$$a_{1i}(-1)^{1+j}|\mathbf{A}_{1j}| + a_{2i}(-1)^{2+j}|\mathbf{A}_{2j}| + \ldots + a_{ni}(-1)^{n+j}|\mathbf{A}_{nj}| = 0, \tag{6.10}$$

czyli suma iloczynów kolejnych elementów jednego wiersza (*jednej kolumny*) *i dopełnień algebraicznych*[1] *kolejnych elementów innego wiersza* (*innej kolumny*) *jest równa zeru.*

Dowód. Przyjmijmy, że $\mathbf{A} = [a_{ij}] = [\,\mathbf{a}_1 \dots \mathbf{a}_i \dots \mathbf{a}_j \dots \mathbf{a}_n\,]$ $(i \neq j)$ i weźmy pod uwagę macierz $\mathbf{B} = [b_{ij}] = [\,\mathbf{a}_1 \dots \mathbf{a}_i \dots \mathbf{a}_i \dots \mathbf{a}_n\,]$ powstałą z $\mathbf{A}$ przez zastąpienie w niej j-tej kolumny przez i-tą ($b_{kj} = a_{ki}$ dla $k = 1, \dots, n$). Macierz $\mathbf{B}$ ma dwie równe kolumny, więc wobec twierdzenia 6.1.6 jest $\det \mathbf{B} = 0$. Z drugiej strony ponieważ macierze $\mathbf{B}$ i $\mathbf{A}$ różnią się tylko j-tą kolumną, więc

$$\begin{aligned}
\mathbf{A}_{kj} &= \begin{bmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & & & \vdots & & \vdots \\ a_{k1} & \dots & a_{ki} & \dots & a_{kj} & \dots & a_{kn} \\ \vdots & & & & \vdots & & \vdots \\ a_{n1} & \dots & a_{ni} & \dots & a_{nj} & \dots & a_{nn} \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1i} & \dots & a_{1n} \\ \vdots & & & & \vdots & & \vdots \\ a_{k1} & \dots & a_{ki} & \dots & a_{ki} & \dots & a_{kn} \\ \vdots & & & & \vdots & & \vdots \\ a_{n1} & \dots & a_{ni} & \dots & a_{ni} & \dots & a_{nn} \end{bmatrix} = \mathbf{B}_{kj}
\end{aligned}$$

(dla $k = 1, \dots, n$) i dlatego rozwijając wyznacznik macierzy $\mathbf{B}$ względem j-tej kolumny otrzymamy

$$\det \mathbf{B} = \sum_{k=1}^{n} b_{kj}(-1)^{k+j}|\mathbf{B}_{kj}| = \sum_{k=1}^{n} a_{ki}(-1)^{k+j}|\mathbf{A}_{kj}|$$

i stąd wynika (6.10). Analogicznie dowodzi się (6.9). □

Następujący wniosek jest prostą konsekwencją twierdzeń 6.1.1 i 6.1.7.

Wniosek 6.1.1. *Niech* $\mathbf{A} = [a_{ij}]$ *będzie macierzą kwadratową stopnia* n. *Jeśli* i *oraz* j *są liczbami ze zbioru* $\{1, \dots, n\}$, *to*

$$\sum_{k=1}^{n} a_{ik}(-1)^{j+k}|\mathbf{A}_{jk}| = \begin{cases} |\mathbf{A}|, & \text{gdy } j = i, \\ 0, & \text{gdy } j \neq i, \end{cases} \tag{6.11}$$

i

$$\sum_{k=1}^{n} a_{ik}(-1)^{k+j}|\mathbf{A}_{kj}| = \begin{cases} |\mathbf{A}|, & \text{gdy } j = i, \\ 0, & \text{gdy } j \neq i. \end{cases} \tag{6.12}$$

[1] Dopełnieniem algebraicznym elementu a_{ij} macierzy kwadratowej $\mathbf{A} = [a_{ij}]$ nazywa się liczbę $(-1)^{i+j}|\mathbf{A}_{ij}|$.

Twierdzenie 6.1.8. *Jeśli macierz* $\mathbf{B}$ *powstaje z macierzy kwadratowej* $\mathbf{A}$ *w wyniku pomnożenia jednej kolumny* (*jednego wiersza*) *przez liczbę* r, *to*

$$\det \mathbf{B} = r \cdot \det \mathbf{A}.$$

Dowód. Niech $\mathbf{A} = [a_{ij}] = [\,\mathbf{a}_1 \dots \mathbf{a}_k \dots \mathbf{a}_n\,]$ będzie macierzą kwadratową stopnia n i niech $\mathbf{B}$ będzie macierzą powstałą z $\mathbf{A}$ w wyniku pomnożenia k- tej kolumny przez liczbę r. Wtedy $\mathbf{B} = [b_{ij}] = [\,\mathbf{a}_1 \dots r\,\mathbf{a}_k \dots \mathbf{a}_n\,]$, $b_{ik} = ra_{ik}$ oraz

$$\mathbf{B}_{ik} = \begin{bmatrix} a_{11} & \dots & ra_{1k} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & ra_{ik} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & ra_{nk} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1k} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ik} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nk} & \dots & a_{nn} \end{bmatrix} = \mathbf{A}_{ik}$$

dla $i = 1, \dots, n$. Zatem rozwijając wyznaczniki macierzy $\mathbf{B}$ i $\mathbf{A}$ względem k-tych kolumn, zauważamy, że

$$\det \mathbf{B} = \sum_{i=1}^{n} b_{ik}(-1)^{i+k}|\mathbf{B}_{ik}| = \sum_{i=1}^{n} r\, a_{ik}(-1)^{i+k}|\mathbf{A}_{ik}| = r \cdot \det \mathbf{A}. \ \square$$

Wniosek 6.1.2. *Jeśli* $\mathbf{A}$ *jest macierzą kwadratową stopnia* n, *to dla każdej liczby* r *jest*

$$\det(r\mathbf{A}) = r^n \cdot \det \mathbf{A}. \quad \square$$

Przykład 109. Z twierdzenia 6.1.8 i wniosku 6.1.2 wynika, że mamy

$$\begin{vmatrix} 6 & 2 & 3 \\ 2r & 3r & 6r \\ 0 & 3 & 2 \end{vmatrix} = r \begin{vmatrix} 6 & 2 & 3 \\ 2 & 3 & 6 \\ 0 & 3 & 2 \end{vmatrix} \qquad \begin{vmatrix} 6 & 2 & 3p \\ 2 & 3 & 6p \\ 0 & 3 & 2p \end{vmatrix} = p \begin{vmatrix} 6 & 2 & 3 \\ 2 & 3 & 6 \\ 0 & 3 & 2 \end{vmatrix}$$

i

$$\begin{vmatrix} 10 & 20 & 50 \\ 30 & 10 & 40 \\ -10 & 20 & 20 \end{vmatrix} = 10^3 \begin{vmatrix} 1 & 2 & 5 \\ 3 & 1 & 4 \\ -1 & 2 & 2 \end{vmatrix} = 1000 \cdot 9 = 9000.$$

Twierdzenie 6.1.9. *Jeśli macierz* $\mathbf{B}$ *powstaje z macierzy kwadratowej* $\mathbf{A}$ *przez dodanie do jednej kolumny innej kolumny pomnożonej przez dowolną liczbę* (*albo przez dodanie do jednego wiersza innego wiersza pomnożonego przez dowolną liczbę*), *to*

$$\det \mathbf{B} = \det \mathbf{A}.$$

Dowód. Niech $\mathbf{A} = [a_{ij}] = [\,\mathbf{a}_1 \dots \mathbf{a}_l \dots \mathbf{a}_k \dots \mathbf{a}_n\,]$ będzie macierzą kwadratową stopnia n i niech $\mathbf{B}$ będzie macierzą powstałą z $\mathbf{A}$ przez dodanie iloczynu k-tej kolumny i liczby r do l-tej kolumny, $k \neq l$. Wtedy $\mathbf{B} = [b_{ij}] = [\,\mathbf{a}_1 \dots \mathbf{a}_l + r\,\mathbf{a}_k \dots \mathbf{a}_k \dots \mathbf{a}_n\,]$, $b_{il} = a_{il} + ra_{ik}$ oraz

$$\begin{aligned} \mathbf{B}_{il} &= \begin{bmatrix} a_{11} & \dots & a_{1l} + ra_{1k} & \dots & a_{1k} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{il} + ra_{ik} & \dots & a_{ik} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nl} + ra_{nk} & \dots & a_{nk} & \dots & a_{nn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & \dots & a_{1l} & \dots & a_{1k} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{il} & \dots & a_{ik} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nl} & \dots & a_{nk} & \dots & a_{nn} \end{bmatrix} = \mathbf{A}_{il} \end{aligned}$$

dla $i = 1, \ldots, n$. Ponieważ liczby k i l są różne, więc wobec twierdzenia 6.1.7 jest $\sum_{i=1}^{n} a_{ik}(-1)^{i+l}|\mathbf{A}_{il}| = 0$. Zatem rozwijając wyznacznik macierzy $\mathbf{B}$ względem l-tej kolumny otrzymujemy

$$\begin{aligned} \det \mathbf{B} &= \sum_{i=1}^{n} b_{il}(-1)^{i+l}|\mathbf{B}_{il}| = \sum_{i=1}^{n}(a_{il} + ra_{ik})(-1)^{i+l}|\mathbf{A}_{il}| \\ &= \sum_{i=1}^{n} a_{il}(-1)^{i+l}|\mathbf{A}_{il}| + r\sum_{i=1}^{n} a_{ik}(-1)^{i+l}|\mathbf{A}_{il}| \\ &= \det \mathbf{A} + r \cdot 0 = \det \mathbf{A}. \ \square \end{aligned}$$

Przykład 110. Obliczyć wyznacznik macierzy $\begin{bmatrix} 3 & 2 & -3 & 2 \\ 0 & 2 & 3 & -1 \\ -9 & -4 & 12 & -7 \\ 2 & 3 & 2 & 1 \end{bmatrix}$.

Dodając pierwszy wiersz pomnożony przez 3 do trzeciego, nie zmieniamy wyznacznika i otrzymujemy

$$\begin{vmatrix} 3 & 2 & -3 & 2 \\ 0 & 2 & 3 & -1 \\ -9 & -4 & 12 & -7 \\ 2 & 3 & 2 & 1 \end{vmatrix} = \begin{vmatrix} 3 & 2 & -3 & 2 \\ 0 & 2 & 3 & -1 \\ 0 & 2 & 3 & -1 \\ 2 & 3 & 2 & 1 \end{vmatrix} = 0,$$

bo jest to wyznacznik z macierzy mającej dwa identyczne wiersze.

Przykład 111. Obliczyć wyznacznik macierzy $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 3 & 4 & 3 & 0 \\ 5 & 4 & 6 & 6 \\ 4 & 2 & 4 & 3 \end{bmatrix}$.

Ponieważ w ostatniej kolumnie macierzy $\mathbf{A}$ mamy już dwa zera, więc dobrym pomysłem jest eliminacja kolejnego niezerowego wyrazu z tej kolumny. W tym celu możemy do trzeciego wiersza dodać wiersz czwarty pomnożony przez -2 i rozwinąć otrzymany wyznacznik względem ostatniej kolumny,

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

Ostatni wyznacznik można już policzyć za pomocą schematu Sarrusa, ale można także pierwszy wiersz pomnożony przez -2 dodać do drugiego i obliczyć wartość wyznacznika za pomocą rozwinięcia względem elementów drugiej kolumny,

$$\det \mathbf{A} = 3 \begin{vmatrix} 1 & 2 & 3 \\ 1 & 0 & -3 \\ -3 & 0 & -2 \end{vmatrix} = 3(-2) \begin{vmatrix} 1 & -3 \\ -3 & -2 \end{vmatrix} = -6(-2-9) = 66.$$

Przykład 112. Obliczyć wyznacznik macierzy $\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 1 & 3 & 3 & 3 \\ 1 & 1 & 1 & 4 & 4 \\ 1 & 1 & 1 & 1 & 5 \end{bmatrix}$.

Odejmując kolejno pierwszy wiersz od drugiego, trzeciego, czwartego i piątego wiersza, otrzymujemy

$$\det \mathbf{A} = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 & 4 \end{vmatrix} = 24,$$

bo jest to wyznacznik macierzy trójkątnej i jest on równy iloczynowi elementów stojących na głównej przekątnej.

6.2. Wyznacznik iloczynu macierzy

Ważną rolę w naszych rozważaniach będzie pełniło twierdzenie o wyznaczniku iloczynu macierzy. Udowodnimy tzw. twierdzenie Cauchy'ego: jeśli $\mathbf{A}$ i $\mathbf{B}$ są macierzami kwadratowymi tego samego stopnia, to wyznacznik iloczynu $\mathbf{AB}$ jest iloczynem wyznacznika macierzy $\mathbf{A}$ i wyznacznika macierzy $\mathbf{B}$, czyli udowodnimy równość

$$\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B}).$$

Zaczynamy od twierdzenia, które jest prostą konsekwencją twierdzeń 6.1.5, 6.1.8 i 6.1.9.

Twierdzenie 6.2.1. *Jeśli* $\mathbf{A}$ *i* $\mathbf{E}$ *są macierzami wymiaru* $n \times n$ *i macierz* $\mathbf{E}$ *jest elementarna, to*

$$\det(\mathbf{EA}) = (\det \mathbf{E})(\det \mathbf{A}) \ \ i \ \ \det(\mathbf{AE}) = (\det \mathbf{A})(\det \mathbf{E}).$$

Dowód. Niech $\mathbf{E}$ będzie jedną z macierzy elementarnych $\mathbf{E}_{ij}$, $\mathbf{E}_i(t)$ i $\mathbf{E}_{ij}(t)$ wymiaru $n \times n$. Na początek przyjmijmy, że $\mathbf{E} = \mathbf{E}_{ij}$ dla $i, j \in \{1, \ldots, n\}$, $i \neq j$. Ponieważ macierz $\mathbf{EA} = \mathbf{E}_{ij}\mathbf{A}$ powstaje z macierzy $\mathbf{A}$ w wyniku przestawienia miejscami dwóch wierszy, to wobec twierdzenia 6.1.5 jest $|\mathbf{E}_{ij}\mathbf{A}| = -|\mathbf{A}|$. Dodatkowo mamy $|\mathbf{E}_{ij}| = -|\mathbf{I}_n| = -1$, bo $|\mathbf{I}_n| = 1$ i macierz $\mathbf{E}_{ij}$ także powstaje z macierzy $\mathbf{I}_n$ w wyniku przestawienia miejscami dwóch wierszy. Stąd wynika, że $|\mathbf{E}_{ij}\mathbf{A}| = -|\mathbf{A}| = |\mathbf{E}_{ij}||\mathbf{A}|$. Jednocześnie, ponieważ $\mathbf{E}_{ij}^T = \mathbf{E}_{ij}$, więc z własności transpozycji macierzy (twierdzenia 4.2.5 i 6.1.4) i z już udowodnionej równości kolejno mamy $|\mathbf{AE}_{ij}| = |(\mathbf{E}_{ij}^T\mathbf{A}^T)^T| = |\mathbf{E}_{ij}^T\mathbf{A}^T| = |\mathbf{E}_{ij}\mathbf{A}^T| = |\mathbf{E}_{ij}||\mathbf{A}^T| = |\mathbf{E}_{ij}||\mathbf{A}| = |\mathbf{A}||\mathbf{E}_{ij}|$.

Równości $|\mathbf{EA}| = |\mathbf{E}||\mathbf{A}| = |\mathbf{AE}|$, gdy $\mathbf{E} = \mathbf{E}_i(t)$ lub $\mathbf{E} = \mathbf{E}_{ij}(t)$, dowodzi się analogicznie. □

Możemy już teraz udowodnić zapowiedziane twierdzenie Cauchy'ego o wyznaczniku iloczynu macierzy kwadratowych.

Twierdzenie 6.2.2 (Cauchy). *Jeśli* $\mathbf{A}$ *i* $\mathbf{B}$ *są macierzami kwadratowymi tego samego stopnia, to* Twierdzenie Cauchy'ego

$$\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B}). \tag{6.13}$$

Dowód. Jeśli macierz $\mathbf{A}$ jest odwracalna, to wobec twierdzenia 5.3.1 istnieją macierze elementarne $\mathbf{E}_1, \ldots, \mathbf{E}_k$ takie, że $\mathbf{A} = \mathbf{E}_1 \cdot \ldots \cdot \mathbf{E}_k$. Wtedy $\mathbf{AB} = \mathbf{E}_1 \cdot \ldots \cdot \mathbf{E}_k\mathbf{B}$ i wobec twierdzenia 6.2.1 mamy

$$\begin{aligned}|\mathbf{AB}| &= |\mathbf{E}_1 \cdot \ldots \cdot \mathbf{E}_k\mathbf{B}| = |\mathbf{E}_1||\mathbf{E}_2 \cdot \ldots \cdot \mathbf{E}_k\mathbf{B}| = \ldots = |\mathbf{E}_1| \cdot \ldots \cdot |\mathbf{E}_k||\mathbf{B}| \\ &= |\mathbf{E}_1\mathbf{E}_2| \cdot \ldots \cdot |\mathbf{E}_k||\mathbf{B}| = \ldots = |\mathbf{E}_1 \cdot \ldots \cdot \mathbf{E}_k||\mathbf{B}| = |\mathbf{A}||\mathbf{B}|.\end{aligned}$$

Załóżmy teraz, że macierz $\mathbf{A}$ nie jest odwracalna. Niech wtedy $\mathbf{C}$ będzie macierzą wierszowo równoważną macierzy $\mathbf{A}$ i mającą postać schodkową normalną. Z faktu, że $\mathbf{A}$ nie jest odwracalna i z twierdzenia 5.3.1 wynika, że macierz $\mathbf{C}$ ma zerowy wiersz. Wtedy także macierz $\mathbf{CB}$ ma zerowy wiersz i wobec twierdzenia 6.1.2 jest

$$|\mathbf{C}| = 0 = |\mathbf{CB}|. \tag{6.14}$$

Z wierszowej równoważności macierzy $\mathbf{A}$ i $\mathbf{C}$ wynika istnienie macierzy elementarnych $\mathbf{E}_1, \ldots, \mathbf{E}_k$ takich, że $\mathbf{A} = \mathbf{E}_1 \cdot \ldots \cdot \mathbf{E}_k\mathbf{C}$. Teraz zauważmy, że wobec (6.14) i twierdzenia 6.2.1 mamy

$$|\mathbf{A}| = |\mathbf{E}_1 \cdot \ldots \cdot \mathbf{E}_k\mathbf{C}| = |\mathbf{E}_1| \cdot \ldots \cdot |\mathbf{E}_k||\mathbf{C}| = 0$$

oraz

$$|\mathbf{AB}| = |\mathbf{E}_1 \cdot \ldots \cdot \mathbf{E}_k\mathbf{CB}| = |\mathbf{E}_1| \cdot \ldots \cdot |\mathbf{E}_k||\mathbf{CB}| = 0 = 0|\mathbf{B}| = |\mathbf{A}||\mathbf{B}|$$

i to kończy dowód twierdzenia. □

Przykład 113. Obliczyć $\det \mathbf{A}$, gdy $\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 3 & 1 & 0 \\ 3 & 3 & 3 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & -1 \\ 0 & -2 & 2 \\ 4 & 3 & 3 \end{bmatrix}$.

Wobec twierdzenia 6.2.2 mamy

$$\det \mathbf{A} = \begin{vmatrix} 2 & 0 & 0 \\ 3 & 1 & 0 \\ 3 & 3 & 3 \end{vmatrix} \cdot \begin{vmatrix} 0 & 0 & -1 \\ 0 & -2 & 2 \\ 4 & 3 & 3 \end{vmatrix} = 6(-8) = -48.$$

Przykład 114. Wyznaczyć $\det(4\mathbf{A}^5)$, gdy $\mathbf{A}$ jest macierzą wymiaru 3×3 i $\det \mathbf{A} = 2$.

Z twierdzenia 6.1.8 (lub z wniosku 6.1.2) oraz z twierdzenia 6.2.2 mamy

$$\det(4\mathbf{A}^5) = 4^3 \det(\mathbf{A}^5) = 4^3(\det \mathbf{A})^5 = 4^3 \cdot 2^5 = 2048.$$

6.3. Macierze odwracalne i nieosobliwe

Macierz $\mathbf{A}$ jest nieosobliwa, gdy $|\mathbf{A}| \neq 0$

Definicja 6.3.1. Macierz kwadratową $\mathbf{A}$ nazywamy *macierzą nieosobliwą*, gdy

$$\det \mathbf{A} \neq 0.$$

Macierz $\mathbf{A}$ jest odwracalna, gdy $\mathbf{A}^{-1}$ istnieje

Udowodnimy teraz jedno z piękniejszych i ważniejszych twierdzeń całej matematyki. Udowodnimy, że macierz jest nieosobliwa wtedy i tylko wtedy, gdy jest ona odwracalna. Przedstawimy także jeszcze jeden sposób wyznaczania macierzy odwrotnej. Zauważmy najpierw, że każda odwracalna macierz jest nieosobliwa.

Twierdzenie 6.3.1. *Jeśli macierz kwadratowa* $\mathbf{A}$ *jest odwracalna, to jest ona nieosobliwa. Wtedy także macierz odwrotna* $\mathbf{A}^{-1}$ *jest nieosobliwa i dodatkowo*

$|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1}$

$$\det(\mathbf{A}^{-1}) = \frac{1}{\det \mathbf{A}}.$$

Dowód. Jeśli macierz $\mathbf{A}$ jest odwracalna, to macierz $\mathbf{A}^{-1}$ istnieje i $\mathbf{A}\cdot\mathbf{A}^{-1} = \mathbf{A}^{-1}\cdot\mathbf{A} = I$. Zatem, wobec twierdzenia 6.2.2,

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

Stąd zaś wynika nieosobliwość macierzy $\mathbf{A}$ i $\mathbf{A}^{-1}$ oraz zależność $\det(\mathbf{A}^{-1}) = (\det \mathbf{A})^{-1}$ pomiędzy ich wyznacznikami. □

Udowodnimy teraz, że każda macierz nieosobliwa jest odwracalna. (Inny dowód tego samego faktu zaproponowaliśmy w ćwiczeniach.) Niech $\mathbf{A} = [a_{ij}]$ będzie macierzą kwadratową stopnia n. Z dopełnień algebraicznych elementów macierzy $\mathbf{A}$, czyli z liczb $D_{ij} = (-1)^{i+j}|\mathbf{A}_{ij}|$, tworzymy macierz

Macierz dołączona

$$\mathbf{A}^D = \begin{bmatrix} D_{11} & D_{21} & \dots & D_{n1} \\ D_{12} & D_{22} & \dots & D_{n2} \\ & & \ddots & \\ D_{1n} & D_{2n} & \dots & D_{nn} \end{bmatrix} \quad \left(= \begin{bmatrix} D_{11} & D_{12} & \dots & D_{1n} \\ D_{21} & D_{22} & \dots & D_{2n} \\ & & \ddots & \\ D_{n1} & D_{n2} & \dots & D_{nn} \end{bmatrix}^T \right),$$

którą nazywamy *macierzą dołączoną* macierzy $\mathbf{A}$. Zauważmy, że jeśli

$$\mathbf{A}\mathbf{A}^D = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} D_{11} & D_{21} & \dots & D_{n1} \\ D_{12} & D_{22} & \dots & D_{n2} \\ & & \ddots & \\ D_{1n} & D_{2n} & \dots & D_{nn} \end{bmatrix} = [c_{ij}],$$

to z definicji iloczynu macierzy oraz z twierdzeń 6.1.1 i 6.1.7 mamy

$$c_{ij} = \sum_{k=1}^{n} a_{ik} D_{jk} = \sum_{k=1}^{n} a_{ik}(-1)^{j+k}|\mathbf{A}_{jk}| = \begin{cases} |\mathbf{A}|, & \text{gdy } j = i, \\ 0, & \text{gdy } j \neq i. \end{cases}$$

To oznacza, że

$$\mathbf{A} \cdot \mathbf{A}^D = [c_{ij}] = |\mathbf{A}| \cdot \mathbf{I}_n. \tag{6.15}$$

Podobnie można udowodnić, że

$$\mathbf{A}^D \cdot \mathbf{A} = |\mathbf{A}| \cdot \mathbf{I}_n. \tag{6.16}$$

Dlatego, jeśli $|\mathbf{A}| \neq 0$, to z (6.15) i (6.16) mamy

$$\mathbf{A} \cdot \frac{\mathbf{A}^D}{|\mathbf{A}|} = \frac{\mathbf{A}^D}{|\mathbf{A}|} \cdot \mathbf{A} = \mathbf{I}_n$$

i $\frac{\mathbf{A}^D}{|\mathbf{A}|}$ jest macierzą odwrotną macierzy $\mathbf{A}$. Zatem udowodniliśmy następujące twierdzenie o odwracalności macierzy nieosobliwej.

Twierdzenie 6.3.2. *Każda macierz nieosobliwa* $\mathbf{A}$ *jest odwracalna i jej macierz odwrotna jest określona wzorem*

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \cdot \mathbf{A}^D. \ \square \tag{6.17}$$

Macierz odwrotna

Z ostatnich dwóch twierdzeń wynika bardzo ważny wniosek.

Wniosek 6.3.1. *Macierz kwadratowa* $\mathbf{A}$ *jest odwracalna wtedy i tylko wtedy, gdy jest ona nieosobliwa, tj. wtedy i tylko wtedy, gdy* $\det \mathbf{A} \neq 0$. (*Równoważnie, macierz kwadratowa* $\mathbf{A}$ *jest odwracalna wtedy i tylko wtedy, gdy liczba* $\det \mathbf{A}$ *jest odwracalna.*) $\square$

$\mathbf{A}^{-1}$ istnieje $\Leftrightarrow$ $|\mathbf{A}| \neq 0$

$\mathbf{A}^{-1}$ istnieje $\Leftrightarrow$ $|\mathbf{A}|^{-1}$ istnieje

---

Przykład 115. Jeśli macierz $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ jest nieosobliwa, tj. gdy $ad - bc \neq 0$, to wobec (6.17) jej macierz odwrotna określona jest wzorem

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

Przykładowo mamy

$$\begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}^{-1} = \frac{1}{10} \begin{bmatrix} 4 & -1 \\ -2 & 3 \end{bmatrix}.$$

Przykład 116. Wyznaczyć macierz odwrotną macierzy

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

Ponieważ $|\mathbf{A}| = 7 \neq 0$, więc macierz $\mathbf{A}$ jest odwracalna i wobec (6.17) mamy

$$\mathbf{A}^{-1} = \frac{\mathbf{A}^D}{|\mathbf{A}|} = \frac{1}{|\mathbf{A}|} \begin{bmatrix} D_{11} & D_{21} & D_{31} \\ D_{12} & D_{22} & D_{32} \\ D_{13} & D_{23} & D_{33} \end{bmatrix}.$$

Dla rozważanej macierzy $\mathbf{A}$ jest

$$\mathbf{A}^{-1} = \frac{1}{7}\begin{bmatrix} +\begin{vmatrix} 1 & 1 \\ -1 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ -1 & 3 \end{vmatrix} & +\begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} \\ -\begin{vmatrix} -1 & 1 \\ 2 & 3 \end{vmatrix} & +\begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} \\ +\begin{vmatrix} -1 & 1 \\ 2 & -1 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 2 & -1 \end{vmatrix} & +\begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 4 & -5 & -1 \\ 5 & -1 & -3 \\ -1 & 3 & 2 \end{bmatrix}.$$

Przykład 117. Za pomocą macierzy odwrotnej rozwiązać równanie $\mathbf{XA} = \mathbf{B}$, w którym

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \text{ i } \mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & -1 & 3 \end{bmatrix}.$$

Ponieważ macierz $\mathbf{A}$ jest odwracalna, więc pomnażając obie strony równania $\mathbf{XA} = \mathbf{B}$ przez $\mathbf{A}^{-1}$, otrzymujemy $\mathbf{X} = \mathbf{BA}^{-1}$, i dlatego

$$\begin{aligned} \mathbf{X} &= \begin{bmatrix} 1 & 2 & 3 \\ 1 & -1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 2 & 3 \\ 1 & -1 & 3 \end{bmatrix}\begin{bmatrix} 1/2 & 0 & 1/2 \\ -1 & 1 & -1 \\ -3/2 & 1 & -1/2 \end{bmatrix} = \begin{bmatrix} -6 & 5 & -3 \\ -3 & 2 & 0 \end{bmatrix}. \end{aligned}$$

Przykład 118. Rozwiązać równanie macierzowe $\mathbf{AX} + 2\mathbf{X} = \mathbf{B}$, w którym

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} \quad \text{i} \quad \mathbf{B} = \begin{bmatrix} 5 & -2 & 5 \\ -1 & 3 & 12 \end{bmatrix}.$$

Zauważmy najpierw, że mamy równoważności

$$\mathbf{AX} + 2\mathbf{X} = \mathbf{B} \Leftrightarrow \mathbf{AX} + 2\mathbf{I}\,\mathbf{X} = \mathbf{B} \Leftrightarrow (\mathbf{A} + 2\mathbf{I})\mathbf{X} = \mathbf{B} \Leftrightarrow \mathbf{X} = (\mathbf{A} + 2\mathbf{I})^{-1}\mathbf{B},$$

a ponieważ

$$(\mathbf{A} + 2\mathbf{I})^{-1} = \left(\begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\right)^{-1} = \begin{bmatrix} 3 & 1 \\ 2 & 5 \end{bmatrix}^{-1} = \frac{1}{13}\begin{bmatrix} 5 & -1 \\ -2 & 3 \end{bmatrix},$$

więc

$$\mathbf{X} = \frac{1}{13}\begin{bmatrix} 5 & -1 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 5 & -2 & 5 \\ -1 & 3 & 12 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 1 & 2 \end{bmatrix}.$$

6.4. Wyznacznik macierzy podobnych

Tu ograniczamy swoje zainteresowania do najprostszych własności wyznaczników macierzy podobnych. Przypomnijmy, że macierze kwadratowe $\mathbf{A}$ i $\mathbf{B}$ nazywamy podobnymi, gdy istnieje nieosobliwa macierz $\mathbf{P}$ taka, że $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$ (zob. def. 4.4.2).

Twierdzenie 6.4.1. *Jeśli macierze* $\mathbf{A}$ *i* $\mathbf{B}$ *wymiaru* $n \times n$ *są podobne, to:*

(a) *macierze* $\mathbf{A} - \lambda\mathbf{I}_n$ *i* $\mathbf{B} - \lambda\mathbf{I}_n$ *są podobne dla każdego* $\lambda \in R$;

(b) $\det(\mathbf{A} - \lambda\mathbf{I}_n) = \det(\mathbf{B} - \lambda\mathbf{I}_n)$ *dla każdego* $\lambda \in R$;

(c) $\det \mathbf{A} = \det \mathbf{B}$.

Dowód. Niech $\mathbf{P}$ będzie nieosobliwą macierzą taką, że $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$. Wtedy

$$\mathbf{B} - \lambda\mathbf{I}_n = \mathbf{P}^{-1}\mathbf{AP} - \lambda\mathbf{P}^{-1}\mathbf{P} = \mathbf{P}^{-1}(\mathbf{AP} - \lambda\mathbf{P}) = \mathbf{P}^{-1}(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{P}$$

i to dowodzi własność (a). Ponieważ $(\det \mathbf{P}^{-1})(\det \mathbf{P}) = 1$ (zob. tw. 6.3.1), więc z powyższej równości i z twierdzenia Cauchy'ego (tw. 6.2.2) mamy własność (b), bo

$$\begin{aligned} \det(\mathbf{B} - \lambda\mathbf{I}_n) &= \det\big(\mathbf{P}^{-1}(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{P}\big) \\ &= \det(\mathbf{P}^{-1}) \cdot \det(\mathbf{A} - \lambda\mathbf{I}_n) \cdot \det(P) \\ &= \det(\mathbf{A} - \lambda\mathbf{I}_n). \end{aligned}$$

W końcu, (c) wynika z (b) dla $\lambda = 0$. □

6.5. Układy równań i wzory Cramera

Wyznacznik macierzy może być przydatny przy stwierdzaniu istnienia, wyznaczaniu liczby rozwiązań i przy samym wyznaczaniu rozwiązań układu równań liniowych $\mathbf{Ax} = \mathbf{b}$, gdzie $\mathbf{A} \in K_{n\times n}$ i $\mathbf{b} \in K_{n\times 1}$. Dla takiego układu mamy równoważności:

$$\begin{aligned} \mathbf{Ax} = \mathbf{b} \quad & \text{ma dokładnie jedno rozwiązanie} \\ \Leftrightarrow \quad & \text{macierz } \mathbf{A} \text{ jest wierszowo równoważna macierzy } \mathbf{I} \text{ (wn. 5.1.1)} \\ \Leftrightarrow \quad & \text{macierz } \mathbf{A} \text{ jest odwracalna (tw. 5.3.1)} \\ \Leftrightarrow \quad & \det \mathbf{A} \neq 0 \text{ (wn. 6.3.1).} \end{aligned}$$

Zatem mamy następujące wnioski o rozwiązaniach układów równań linowych.

Wniosek 6.5.1. *Jeśli* $\mathbf{A} \in K_{n\times n}$ *i* $\mathbf{b} \in K_{n\times 1}$*, to układ równań liniowych* $\mathbf{Ax} = \mathbf{b}$ *ma dokładnie jedno rozwiązanie wtedy i tylko wtedy, gdy* $\det \mathbf{A} \neq 0$. □

Wniosek 6.5.2. *Jeśli* $\mathbf{A}$ *jest macierzą kwadratową, to jednorodny układ równań liniowych* $\mathbf{Ax} = \mathbf{0}$ *ma niezerowe rozwiązanie wtedy i tylko wtedy, gdy* $\det \mathbf{A} = 0$. □

Przykład 119. Znaleźć te wartości parametru a, dla których jednorodny układ równań

$$\begin{cases} x_1 + 2x_2 - 3x_3 = 0 \\ 2x_1 + ax_2 + x_3 = 0 \\ 2x_1 + x_2 + ax_3 = 0 \end{cases}$$

ma niezerowe rozwiązanie. Rozwiązać układ dla otrzymanych wartości parametru a.

Wyznacznikiem macierzy głównej tego układu jest

$$\begin{aligned} |\mathbf{A}| &= \begin{vmatrix} 1 & 2 & -3 \\ 2 & a & 1 \\ 2 & 1 & a \end{vmatrix} = \begin{vmatrix} 1 & 2 & -3 \\ 0 & a-4 & 7 \\ 0 & 1-a & a-1 \end{vmatrix} \\ &= \begin{vmatrix} a-4 & 7 \\ 1-a & a-1 \end{vmatrix} = \begin{vmatrix} a+3 & 7 \\ 0 & a-1 \end{vmatrix} = (a+3)(a-1). \end{aligned}$$

Zatem $|\mathbf{A}| = 0$ wtedy i tylko wtedy, gdy $a = -3$ lub $a = 1$ i wobec wniosku 6.5.1 są to jedyne wartości a, dla których jednorodny układ równań ma niezerowe rozwiązanie.

Jeśli $a = -3$, to dla macierzy rozszerzonej tego układu mamy

$$[\mathbf{A}|\mathbf{0}] = \left[\begin{array}{ccc|c} 1 & 2 & -3 & 0 \\ 2 & -3 & 1 & 0 \\ 2 & 1 & -3 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

i dlatego $\mathbf{x} = [\,t\ t\ t\,]^T$ jest rozwiązaniem układu dla każdego $t \in R$.

Dla $a = 1$ mamy

$$[\mathbf{A}|\mathbf{0}] = \left[\begin{array}{rrr|r} 1 & 2 & -3 & 0 \\ 2 & 1 & 1 & 0 \\ 2 & 1 & 1 & 0 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & 0 & 5/3 & 0 \\ 0 & 1 & -7/3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

i rozwiązaniem układu jest $\mathbf{x} = [-5t\ 7t\ 3t]^T$ dla każdego $t \in R$.

Przedstawimy teraz zastosowania wyznaczników do rozwiązywania układów n równań liniowych o n niewiadomych. Niech $\mathbf{A}$ będzie macierzą kwadratową stopnia n i niech $\mathbf{x}$ oraz $\mathbf{b}$ będą macierzami wymiaru $n \times 1$,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

których współczynniki są elementami ciała K.

Definicja 6.5.1. Układ równań

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases}$$

Układ Cramera

(symbolicznie $\mathbf{Ax} = \mathbf{b}$) nazywamy *układem Cramera*, gdy jego macierz główna $\mathbf{A}$ jest macierzą nieosobliwą, tj. gdy $\det \mathbf{A} \neq 0$.

Wobec wniosku 6.5.1 układ Cramera ma dokładnie jedno rozwiązanie. Teraz poznamy nową metodę wyznaczania tego rozwiązania. Zaczynamy od przydatnej notacji. Dla macierzy $\mathbf{A} = [\,\mathbf{a}_1 \ldots \mathbf{a}_{i-1}\, \mathbf{a}_i\, \mathbf{a}_{i+1} \ldots\ \mathbf{a}_n\,] \in K_{n\times n}$ i $\mathbf{x} \in K_{n\times 1}$ symbolem $\mathbf{A}_i(\mathbf{x})$ oznaczamy macierz powstałą z $\mathbf{A}$ przez zastąpienie w niej i-tej kolumny kolumną $\mathbf{x}$, czyli

$$\mathbf{A}_i(\mathbf{x}) = \begin{bmatrix} | & & | & | & | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_{i-1} & \mathbf{x} & \mathbf{a}_{i+1} & \cdots & \mathbf{a}_n \\ | & & | & | & | & & | \end{bmatrix}.$$

Twierdzenie 6.5.1. *Jedynym rozwiązaniem układ Cramera* $\mathbf{Ax} = \mathbf{b}$*, gdzie* $\mathbf{A} \in K_{n\times n}$ *i* $\mathbf{b} \in K_{n\times 1}$*, jest macierz* $\mathbf{x} = [\,x_1 \ldots x_n\,]^T$*, w której*

Wzór Cramera

$$x_i = \frac{\det \mathbf{A}_i(\mathbf{b})}{\det \mathbf{A}} \quad dla \quad i = 1, \ldots, n. \tag{6.18}$$

Dowód. Załóżmy, że macierz $\mathbf{A} = [\,\mathbf{a}_1 \ldots \mathbf{a}_n\,]$ jest nieosobliwa i niech $\mathbf{x} = [\,x_1 \ldots x_n\,]^T$ będzie jedynym rozwiązaniem układu Cramera $\mathbf{Ax} = \mathbf{b}$. Wtedy $x_1\mathbf{a}_1 + \ldots + x_i\mathbf{a}_i + \ldots + x_n\mathbf{a}_n = \mathbf{b}$ i wobec twierdzeń 6.1.9 i 6.1.8 mamy

$$\begin{aligned} \det \mathbf{A}_i(\mathbf{b}) &= \det \begin{bmatrix} | & & | & | & | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_{i-1} & \mathbf{b} & \mathbf{a}_{i+1} & \cdots & \mathbf{a}_n \\ | & & | & | & | & & | \end{bmatrix} \\ &= \det \begin{bmatrix} | & & | & | & | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_{i-1} & x_1\mathbf{a}_1 + \ldots + x_i\mathbf{a}_i + \ldots + x_n\mathbf{a}_n & \mathbf{a}_{i+1} & \cdots & \mathbf{a}_n \\ | & & | & | & | & & | \end{bmatrix} \\ &= \det \begin{bmatrix} | & & | & | & | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_{i-1} & x_i\mathbf{a}_i & \mathbf{a}_{i+1} & \cdots & \mathbf{a}_n \\ | & & | & | & | & & | \end{bmatrix} \\ &= x_i \det \mathbf{A}. \end{aligned}$$

Stąd otrzymujemy wzór (6.18), bo $\det \mathbf{A} \neq 0$. $\square$

Wzory (6.18), nazywane *wzorami Cramera*, mają wielorakie zastosowania w rozważaniach teoretycznych.

Przykład 120. Za pomocą wzorów Cramera rozwiązać układ równań

$$\begin{cases} x_1 - x_2 + x_3 = 6, \\ 3x_1 + 2x_2 + 2x_3 = 10, \\ x_1 + 3x_2 + 3x_3 = 8. \end{cases}$$

Dla tego układu mamy

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 1 \\ 3 & 2 & 2 \\ 1 & 3 & 3 \end{bmatrix}, \quad \mathbf{A}_1(\mathbf{b}) = \begin{bmatrix} 6 & -1 & 1 \\ 10 & 2 & 2 \\ 8 & 3 & 3 \end{bmatrix},$$

$$\mathbf{A}_2(\mathbf{b}) = \begin{bmatrix} 1 & 6 & 1 \\ 3 & 10 & 2 \\ 1 & 8 & 3 \end{bmatrix}, \quad \mathbf{A}_3(\mathbf{b}) = \begin{bmatrix} 1 & -1 & 6 \\ 3 & 2 & 10 \\ 1 & 3 & 8 \end{bmatrix}.$$

Ponieważ $|\mathbf{A}| = 14$, $|\mathbf{A}_1(\mathbf{b})| = 28$, $|\mathbf{A}_2(\mathbf{b})| = -14$ i $|\mathbf{A}_3(\mathbf{b})| = 42$, więc wobec (6.18) jedynym rozwiązaniem układu są liczby

$$x_1 = \frac{|\mathbf{A}_1(\mathbf{b})|}{|\mathbf{A}|} = 2, \ x_2 = \frac{|\mathbf{A}_2(\mathbf{b})|}{|\mathbf{A}|} = -1, \ x_3 = \frac{|\mathbf{A}_3(\mathbf{b})|}{|\mathbf{A}|} = 3.$$

6.6. Ćwiczenia

1. Posługując się rozwinięciem wyznacznika względem wiersza lub kolumny z największą liczbą zer, obliczyć wyznaczniki następujących macierzy:

$$(a) \begin{bmatrix} 2 & 3 & 0 \\ 7 & 10 & 3 \\ 1 & 5 & 0 \end{bmatrix}; \quad (c) \begin{bmatrix} 5 & 4 & 0 & 7 \\ 0 & 3 & 0 & 2 \\ 2 & 3 & 5 & 4 \\ 0 & 7 & 0 & 2 \end{bmatrix};$$

$$(b) \begin{bmatrix} 2 & 0 & 0 & 2 \\ 0 & 3 & 1 & 3 \\ 2 & 1 & 0 & 4 \\ 4 & 3 & 0 & 1 \end{bmatrix}; \quad (d) \begin{bmatrix} 1 & 3 & 7 & 4 \\ -2 & 0 & 3 & 1 \\ -1 & 2 & 4 & 0 \\ 0 & 3 & 0 & 1 \end{bmatrix}.$$

2. Obliczyć następujące wyznaczniki:

$$(a) \begin{vmatrix} 116 & 98 & 145 \\ 100 & 23 & 100 \\ 13 & 75 & 42 \end{vmatrix}; \quad (d) \begin{vmatrix} 1 & 2 & 3 & 4 \\ -2 & 1 & -4 & 3 \\ 3 & -4 & -1 & 2 \\ 4 & 3 & -2 & -1 \end{vmatrix};$$

$$(b) \begin{vmatrix} 1 & 1 & 2 & 2 \\ 3 & 1 & 4 & 8 \\ 7 & 6 & 1 & 3 \\ 1 & 1 & 3 & 4 \end{vmatrix}; \quad (e) \begin{vmatrix} 1001 & 2001 & 3001 & 4001 \\ 1002 & 2002 & 3002 & 4002 \\ 1003 & 2003 & 3003 & 4003 \\ 1004 & 2004 & 3004 & 4004 \end{vmatrix};$$

$$(c) \begin{vmatrix} 1 & 1 & 1 & 1 \\ x & 1 & 1 & 1 \\ x & x & 1 & 1 \\ x & x & x & 1 \end{vmatrix}; \quad (f) \begin{vmatrix} 1 & 2 & 11 & 2 & 6 \\ 4 & 3 & 9 & 7 & 6 \\ 2 & 3 & 4 & 0 & 0 \\ 4 & 5 & 6 & 5 & 0 \\ 1 & 3 & 5 & 7 & 6 \end{vmatrix};$$

$$(g) \begin{vmatrix} 1 & 2 & 11 & 2 & 0 \\ 4 & 3 & 9 & 0 & 6 \\ 2 & 3 & 0 & 0 & 0 \\ 4 & 0 & 6 & 5 & 0 \\ 0 & 3 & 5 & 7 & 6 \end{vmatrix}; \quad (h) \begin{vmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{vmatrix}.$$

3. Wyznacznikiem Vandermonde'a stopnia n nazywamy wyznacznik

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

Udowodnić, że $V_3 = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2)$. (Udowodnić ogólnie, że $V_n = \prod_{n \geqslant i > j \geqslant 1} (x_i - x_j)$ dla $n \geqslant 2$).

4. Obliczyć wyznaczniki następujących macierzy:

$$(a)\ \mathbf{A}_n = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 1 & \dots & 1 \\ 1 & 1 & 3 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \dots & n \end{bmatrix};$$

$$(b)\ \mathbf{A}_n = \begin{bmatrix} 1 & n & n & \dots & n \\ n & 2 & n & \dots & n \\ n & n & 3 & \dots & n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n & n & n & \dots & n \end{bmatrix};$$

(c) $\mathbf{A}_n = \begin{bmatrix} 1 & 2 & 3 & \dots & n \\ -1 & 0 & 3 & \dots & n \\ -1 & -2 & 0 & \dots & n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -2 & -3 & \dots & 0 \end{bmatrix}$;

(d) $\mathbf{A}_n = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & x & 1 & \dots & 1 \\ 1 & 1 & x-1 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \dots & x-n \end{bmatrix}$;

(e) $\mathbf{A}_n = \begin{bmatrix} 1 & 2 & 3 & \dots & n-2 & n-1 & n \\ 2 & 3 & 4 & \dots & n-1 & n & n \\ 3 & 4 & 5 & \dots & n & n & n \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ n & n & n & \dots & n & n & n \end{bmatrix}$;

(f) $\mathbf{A}_n = \begin{bmatrix} 1 & 2 & 3 & 4 & \dots & n \\ 1 & 1 & 2 & 3 & \dots & n-1 \\ 1 & 1 & 1 & 2 & \dots & n-2 \\ 1 & 1 & 1 & 1 & \dots & n-3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & 1 & \dots & 1 \end{bmatrix}$.

5. Dany jest ciąg macierzy $(\mathbf{A}_n)$, gdzie $\mathbf{A}_n$ jest macierzą wymiaru $n \times n$ i $\mathbf{A}_1 = [2]$, $\mathbf{A}_2 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ oraz

$$\mathbf{A}_n = \begin{bmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & -1 & 2 & -1 \\ 0 & \dots & 0 & -1 & 2 \end{bmatrix} \text{ dla } n > 2.$$

(a) Zauważyć, że $|\mathbf{A}_1| = 2$, $|\mathbf{A}_2| = 3$ i następnie wykazać, że $|\mathbf{A}_n| = 2|\mathbf{A}_{n-1}| - |\mathbf{A}_{n-2}|$ dla $n > 2$. (b) Udowodnić, że $|\mathbf{A}_n| = n + 1$.

6. Dany jest ciąg macierzy $(\mathbf{A}_n)$, gdzie $\mathbf{A}_1 = [12]$, $\mathbf{A}_2 = \begin{bmatrix} 12 & 9 \\ 4 & 12 \end{bmatrix}$ oraz

$$\mathbf{A}_n = \begin{bmatrix} 12 & 9 & 0 & 0 & \dots & 0 \\ 4 & 12 & 9 & 0 & \dots & 0 \\ 0 & 4 & 12 & 9 & \dots & 0 \\ 0 & 0 & 4 & 12 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 12 \end{bmatrix} \text{ dla } n > 2.$$

(a) Zauważyć, że $|\mathbf{A}_1| = 12$, $|\mathbf{A}_2| = 108$ i następnie wykazać, że $|\mathbf{A}_n| = 12|\mathbf{A}_{n-1}| - 36|\mathbf{A}_{n-2}|$ dla $n > 2$. (b) Udowodnić, że $|\mathbf{A}_n| = 6^n(n + 1)$.

7. Rozwiązać następujące równania:

(a) $\begin{vmatrix} x & 2 & 3 \\ 4 & 1 & x \\ 1 & 2+x & 2 \end{vmatrix} = 0$; (b) $\begin{vmatrix} x^2 & 1 & 2 \\ 4 & 2-2x & x \\ 1 & 5-4x & 2x \end{vmatrix} = 0$.

8. Pokazać, że macierz $\mathbf{A}$ jest nieosobliwa dla każdych liczb rzeczywistych a, b i c, gdy

$$\mathbf{A} = \begin{bmatrix} 1 & a & b \\ -a & 1 & c \\ -b & -c & 1 \end{bmatrix}.$$

9. Za pomocą wyznacznika wyznaczyć te wartości parametru x, dla których następujące macierze są odwracalne:

(a) $\begin{bmatrix} 1 & 0 & 0 & 1 \\ 2 & x & 0 & 3 \\ 3 & 0 & 1 & 0 \\ 6 & 1 & 2 & 5 \end{bmatrix}$; (b) $\begin{bmatrix} 8 & x+3 & x+2 & 2x \\ 5 & x & x+2 & x \\ 3 & 1 & x+2 & 1 \\ 2 & 3 & 4 & x \end{bmatrix}$.

10. Za pomocą wzoru (6.17) wyznaczyć macierze odwrotne następujących macierzy:

(a) $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}$; (b) $\begin{bmatrix} 1 & -2 & 3 \\ 1 & -3 & 4 \\ 2 & -5 & 8 \end{bmatrix}$;

(c) $\begin{bmatrix} 15 & 0 & 4 \\ 17 & 1 & 19 \\ 27 & 1 & 22 \end{bmatrix}$; (d) $\begin{bmatrix} 0 & 0 & -2 & 3 \\ 1 & 0 & 1 & 2 \\ -1 & 1 & 2 & 1 \\ 0 & 2 & -3 & 0 \end{bmatrix}$.

11. Wyznaczyć pierwszą kolumnę macierzy $\mathbf{A}^{-1}$, gdy

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

12. Za pomocą wzorów Cramera wyznaczyć niewiadomą x_3 z każdego z następujących układów równań:

(a) $\begin{cases} 2x_1 + 3x_2 - x_3 = 16, \\ \quad - x_2 + 3x_3 = 34, \\ -x_1 + 2x_2 + x_3 = 16; \end{cases}$

(b) $\begin{cases} x_1 - 2x_2 + 3x_3 + 2x_4 = 0, \\ -2x_1 + 3x_2 + x_3 = 2, \\ 3x_1 + x_2 + 2x_3 + x_4 = 1, \\ 2x_1 + x_3 = 0; \end{cases}$

(c) $\begin{cases} 2x_1 - 4x_2 - 2x_3 + 6x_4 = 4, \\ x_1 + x_2 - 2x_4 = -3, \\ 7x_1 - x_2 - 6x_4 = 3, \\ 3x_1 - 3x_2 - 2x_3 = 1; \end{cases}$

(d) $x_1 \begin{bmatrix} 2 \\ 0 \\ 2 \\ 4 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 2 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 1 \\ 2 \\ 5 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \\ 7 \\ 6 \end{bmatrix}$;

(e) $\begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 2 & 3 & 4 \\ 0 & 3 & 2 & 1 \\ 2 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2 \\ 16 \\ 14 \\ 1 \end{bmatrix}$.

13. Dany jest układ równań

$$\begin{cases} 2x_1 + x_2 - 2x_3 = 2, \\ -3x_1 + 3x_2 + 5x_3 = 3, \\ -x_1 + 2x_2 + 2x_3 = 4. \end{cases}$$

(a) Wyznaczyć macierz odwrotną macierzy głównej $\mathbf{A}$ tego układu. (b) Sprawdzić poprzednie rachunki obliczając $\mathbf{A}\mathbf{A}^{-1}$. (c) Za pomocą macierzy $\mathbf{A}^{-1}$ wyznaczyć rozwiązanie tego układu.

14. Rozwiązać następujące równania macierzowe:
(a) $\left(\begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} + 4\mathbf{X}\right)^{-1} = \begin{bmatrix} 3 & 1 \\ 4 & 8 \end{bmatrix}$;
(b) $\mathbf{X} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 9 & 8 \\ 0 & 1 & 6 \end{bmatrix}$;
(c) $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 2 & 3 & 4 & 5 \\ 0 & 0 & 3 & 4 & 5 \\ 0 & 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} \cdot \mathbf{X} = \begin{bmatrix} -1 & 0 & 1 & 0 & -1 \\ 0 & 2 & 0 & -2 & 0 \\ 0 & 0 & 6 & 0 & -6 \\ 0 & 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 0 & 20 \end{bmatrix}$.

15. Wyznaczyć macierz $\mathbf{A}$, jeśli $\mathbf{A}^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ i $\det(\mathbf{A}^{-1}) = 3$.

16. Niech $\mathbf{A}$ będzie macierzą o całkowitych współczynnikach i niech $\det \mathbf{A} = \pm 1$. Pokazać, że macierz $\mathbf{A}^{-1}$ ma te same własności.

17. Niech $\mathbf{A}$ i $\mathbf{B}$ będą macierzami kwadratowymi takimi, że $\mathbf{AB} + \mathbf{B} + \mathbf{I} = \mathbf{0}$. Pokazać, że $\mathbf{B}$ jest macierzą nieosobliwą i znaleźć macierz $\mathbf{B}^{-1}$.

18. Niech $\mathbf{A}$ będzie nieosobliwą macierzą wymiaru $n \times n$ i niech $\mathbf{A}^D$ będzie jej macierzą dołączoną. Wykazać, że ma ona następujące własności:
(a) $|\mathbf{A}^D| = |\mathbf{A}|^{n-1}$;
(b) $\left(\mathbf{A}^D\right)^{-1} = \frac{1}{|\mathbf{A}|}\mathbf{A}$;
(c) $(\alpha\mathbf{A})^D = \alpha^{n-1}\mathbf{A}^D$.

19. Udowodnić lemat 6.2.1 dla macierzy elementarnej $\mathbf{E} = \mathbf{E}_{ij}(t)$.

20. Podać bezpośredni i elementarny dowód twierdzenia Cauchy'ego (tw. 6.2.2) dla macierzy wymiaru 2×2.

21. Korzystając ze wzorów Cramera, wykazać, że jeśli $\mathbf{A}$ jest nieosobliwą macierzą wymiaru $n \times n$, to dla $i, j \in \{1, \ldots, n\}$ jest $(\mathbf{A}^{-1})_{ij} = |\mathbf{A}_i(\mathbf{e}_j)|/|\mathbf{A}|$, gdzie $\mathbf{e}_j$ jest j-tym wektorem jednostkowym. Następnie obliczyć $(\mathbf{A}^{-1})_{11}$ i $(\mathbf{A}^{-1})_{23}$, gdy

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 0 & 6 \\ -3 & 9 & 1 \end{bmatrix}.$$

22. Wykazać, że jeśli macierz $\mathbf{A}$ jest wierszowo równoważna macierzy odwracalnej $\mathbf{B}$, to także macierz $\mathbf{A}$ jest odwracalna.

23. Udowodnić, że jeśli $\mathbf{A}$ i $\mathbf{B}$ są macierzami kwadratowymi stopnia n, to $\mathbf{AB} = \mathbf{I}_n$ wtedy i tylko wtedy, gdy $\mathbf{BA} = \mathbf{I}_n$.

24. Niech $\mathbf{A}$ będzie macierzą nieosobliwą i niech macierz $\mathbf{B}$, mająca normalną postać schodkową, będzie wierszowo równoważna macierzy $\mathbf{A}$. (1) Wykazać, że macierz $\mathbf{B}$ jest nieosobliwa. (2) Z (1) wywnioskować, że macierz $\mathbf{B}$ nie ma ani zerowego wiersza, ani zerowej kolumny. (3) Z tego ostatniego wywnioskować, że $\mathbf{B}$ jest macierzą jednostkową. (4) W końcu z (3) wywnioskować, że macierz $\mathbf{A}$ jest odwracalna.

25. Dana jest macierz $\mathbf{A}$ taka, że $\mathbf{A}^2 = \mathbf{0}$. Wyjaśnić dlaczego macierz $\mathbf{A}$ nie jest odwracalna.

26. Niech S będzie zbiorem nieosobliwych macierzy postaci $\begin{bmatrix} x & y \\ y & x \end{bmatrix}$, gdzie $x, y \in R$. Sprawdzić, czy zbiór S ze zwykłym mnożeniem macierzy jest grupą.

27. Wpisując TAK albo NIE, stwierdzić prawdziwość każdego z następujących zdań:

$\boxed{1\quad}$ Wyznacznik macierzy kwadratowej stopnia n można wyznaczyć za pomocą wyznaczników macierzy kwadratowych stopnia $n-1$.

$\boxed{2\quad}$ Dopełnieniem algebraicznym elementu a_{ij} macierzy $\mathbf{A}$ jest macierz $\mathbf{A}_{ij}$ otrzymana z $\mathbf{A}$ przez wykreślenie z niej i-tego wiersza oraz j-tej kolumny.

$\boxed{3\quad}$ Rozwinięcie wyznacznika $\det \mathbf{A}$ względem kolumny różni się znakiem od rozwinięcia wyznacznika względem wiersza.

$\boxed{4\quad}$ Wyznacznik macierzy trójkątnej $\mathbf{A}$ jest równy sumie elementów stojących na głównej przekątnej macierzy $\mathbf{A}$.

$\boxed{5\quad}$ Jeśli $\mathbf{A}$ i $\mathbf{B}$ są macierzami kwadratowymi stopnia n, to $\det(\mathbf{A} + \mathbf{B}) = \det \mathbf{A} + \det \mathbf{B}$.

$\boxed{6\quad}$ Jeśli macierze kwadratowe $\mathbf{A}$ i $\mathbf{B}$ różnią się tylko pierwszymi kolumnami i $\det \mathbf{A} = 2$ oraz $\det \mathbf{B} = 3$, to $\det(\mathbf{A} + \mathbf{B}) = 5$.

$\boxed{7\quad}$ Jeśli $\mathbf{A}$ jest macierzą prostokątną, to macierze $\mathbf{AA}^T$ i $\mathbf{A}^T\mathbf{A}$ są kwadratowe i $\det(\mathbf{AA}^T) = \det(\mathbf{A}^T\mathbf{A})$.

$\boxed{8\quad}$ $\det(\mathbf{AA}^T) = (\det \mathbf{A})^2$, jeśli $\mathbf{A}$ jest macierzą kwadratową.

$\boxed{9\quad}$ $\det \mathbf{A}^T = -\det \mathbf{A}$, jeśli $\mathbf{A}$ jest macierzą kwadratową.

$\boxed{10\quad}$ Jeśli $\mathbf{A}$ jest macierzą kwadratową stopnia n i r jest liczbą rzeczywistą, to $\det(r\mathbf{A}) = r^{n^2} \det \mathbf{A}$.

$\boxed{11\quad}$ Jeśli $\mathbf{A}$ jest macierzą kwadratową stopnia n i r jest liczbą rzeczywistą, to $\det(r\mathbf{A}) = r^n \det \mathbf{A}$ i $\operatorname{tr}(r\mathbf{A}) = r\operatorname{tr}(\mathbf{A})$.

$\boxed{12\quad}$ Każda macierz elementarna jest nieosobliwa.

$\boxed{13\quad}$ Macierz kwadratowa $\mathbf{A}$ jest odwracalna wtedy i tylko wtedy, gdy $\det \mathbf{A} > 0$.

$\boxed{14\quad}$ Jeśli $\mathbf{A}$ jest macierzą kwadratową, to macierz $\mathbf{A}^2$ jest odwracalna wtedy i tylko wtedy, gdy macierz $\mathbf{A}^3$ jest odwracalna.

$\boxed{15\quad}$ Jeśli $\mathbf{A}$ i $\mathbf{B}$ są macierzami kwadratowymi stopnia n, to $\det(\mathbf{AB}) = \det(\mathbf{BA})$.

$\boxed{16\quad}$ Jeśli $\mathbf{A}$ jest macierzą wymiaru $m \times n$ i $\mathbf{B}$ jest macierzą wymiar $n \times m$, to $\det(\mathbf{AB}) = \det(\mathbf{BA})$.

$\boxed{17\quad}$ Jeśli $\mathbf{A}$ jest macierzą wymiaru $m \times n$ i $m < n$, to $\det(\mathbf{A}^T\mathbf{A}) = 0$.

$\boxed{18\quad}$ Jeśli $\mathbf{A}$ i $\mathbf{B}$ są macierzami kwadratowymi i $\mathbf{AB} = \mathbf{A}^{-1}$, to $\det \mathbf{B} = 1$.

$\boxed{19\quad}$ Jeśli $\mathbf{A}$ jest rzeczywistą macierzą kwadratową i $\mathbf{A}^2 = 8\mathbf{A}^{-1}$, to $\det \mathbf{A} = 0$.

$\boxed{20\quad}$ Jeśli $\mathbf{A}$ jest macierzą kwadratową stopnia n i $\mathbf{A}^T = 4\mathbf{A}^{-1}$, to $\det \mathbf{A} = 2^n$ lub $\det \mathbf{A} = -2^n$.

Rozdział 7

PRZESTRZEŃ WEKTOROWA

7.1. Przestrzeń wektorowa i jej podprzestrzenie

Przypomnijmy, że jeśli V jest niepustym zbiorem, to przez dwuargumentowe działanie w zbiorze V (zob. def. 1.1.1) rozumiemy każdą funkcję $\psi\colon V \times V \to V$. W przestrzeniach wektorowych, którymi teraz się zajmiemy, ważna będzie funkcja przez niektórych nazywana działaniem zewnętrznym.

Działanie zewnętrzne

Definicja 7.1.1. Jeśli K i V są niepustymi zbiorami, to każdą funkcję

$$\varphi\colon K \times V \to V$$

nazywamy *działaniem zewnętrznym* w zbiorze V nad zbiorem K.

$\varphi(a, \mathbf{x}) = a\mathbf{x}$

Wartość działania zewnętrznego $\varphi : K \times V \to V$ dla $a \in K$ i $\mathbf{x} \in V$, czyli element $\varphi(a, \mathbf{x})$ zbioru V, oznaczać będziemy przez $a \cdot \mathbf{x}$ lub $a\mathbf{x}$.

Przykład 121. Funkcja $\varphi\colon R \times R_{m\times n} \to R_{m\times n}$ taka, że $\varphi(a, \mathbf{A})$ jest iloczynem macierzy $\mathbf{A} \in R_{m\times n}$ przez liczbę $a \in R$,

$$\varphi(a, \mathbf{A}) = a\mathbf{A},$$

jest działaniem zewnętrznym w zbiorze $R_{m\times n}$ nad zbiorem R.

Przestrzeń wektorowa

Definicja 7.1.2. System algebraiczny $(V, K, +, \cdot)$, w którym V jest niepustym zbiorem, K jest ciałem, $+$ jest działaniem dwuargumentowym w zbiorze V, a $\cdot$ jest działaniem zewnętrznym w zbiorze V nad zbiorem K nazywamy *przestrzenią wektorową* nad ciałem K, gdy spełnione są następujące warunki:

1° $\forall_{\mathbf{x},\,\mathbf{y}\in V}\ \ \mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$;

2° $\forall_{\mathbf{x},\,\mathbf{y},\,\mathbf{z}\in V}\ \ \mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{y} + \mathbf{x}) + \mathbf{z}$;

3° $\exists_{\mathbf{0}\in V}\forall_{\mathbf{x}\in V}\ \ \mathbf{x} + \mathbf{0} = \mathbf{x}$;

4° $\forall_{\mathbf{x}\in V}\exists_{-\mathbf{x}\in V}\ \ \mathbf{x} + (-\mathbf{x}) = \mathbf{0}$;

5° $\forall_{\alpha\in K}\forall_{\mathbf{x},\mathbf{y}\in V}\ \ \alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$;

6° $\forall_{\alpha,\beta\in K}\forall_{\mathbf{x}\in V}\ \ (\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$;

7° $\forall_{\alpha,\beta\in K}\forall_{\mathbf{x}\in V}\ \ \alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$;

8° $\forall_{\mathbf{x}\in V}\ \ 1\mathbf{x} = \mathbf{x}$, gdzie 1 jest jedynką ciała K.

V – zbiór wektorów
K – ciało skalarów

$\mathbf{x} + \mathbf{y}$ – suma wektorów
$\alpha\mathbf{x}$ – iloczyn wektora $\mathbf{x}$ przez skalar α

Jeśli $(V, K, +, \cdot)$ jest *przestrzenią wektorową* nad ciałem K, to elementy zbioru V nazywamy *wektorami*, a elementy zbioru K – *skalarami*, działanie $+$ nazywamy *dodawaniem* wektorów, działanie $\cdot$ – *mnożeniem* wektorów przez skalary. (Tych samych symboli $+$ i $\cdot$ używamy na oznaczenie działań w ciele K.) Jeśli $\mathbf{x}$, $\mathbf{y} \in V$ i $\alpha \in K$, to wektor $\mathbf{x} + \mathbf{y}$ nazywamy sumą wektorów $\mathbf{x}$ i $\mathbf{y}$, a $\alpha\mathbf{x}$ – iloczynem wektora $\mathbf{x}$ przez skalar α. Wektor $\mathbf{0}$, element neutralny działania $+$, nazywamy wektorem zerowym; wektor $-\mathbf{x}$ nazywamy wektorem przeciwnym do wektora $\mathbf{x}$. Tam gdzie nie będzie to prowadziło do nieporozumień, przestrzeń wektorową $(V, K, +, \cdot)$ nad ciałem K będziemy, dla krótkości, oznaczać symbolem V lub $V(K)$. Warto pamiętać o tym, że jeśli $(V, K, +, \cdot)$ jest przestrzenią wektorową, to system algebraiczny $(V, +)$ jest grupą przemienną.

Przykład 122. Z własności zwykłego dodawania macierzy i zwykłego mnożenia macierzy przez skalary (zob. tw. 4.2.1) wynika, że zbiór $K_{m\times n}$ macierzy wymiaru $m \times n$ o współczynnikach z ciała K jest przestrzenią wektorową nad ciałem K. W szczególności zbiór macierzy jednokolumnowych $K_{n\times 1}$ jest przestrzenią wektorową z działaniami określonymi wzorami

$K_{m\times n}$ – przestrzeń macierzy

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix} \quad \text{i} \quad r \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} rx_1 \\ \vdots \\ rx_n \end{bmatrix}.$$

Z tych samych powodów $K_{1\times n}$ jest przestrzenią wektorową. Niech teraz K^n będzie zbiorem n-elementowych ciągów o wyrazach z ciała K,

$$K^n = \{(x_1, x_2, \dots, x_n) : x_1, x_2, \dots, x_n \in K\}.$$

Łatwo sprawdza się, że zbiór K^n jest przestrzenią wektorową nad ciałem K, jeśli dodawanie ciągów i mnożenie ciągów przez skalary są działaniami określonymi wzorami

K^n – przestrzeń n-elementowych ciągów
R^2, R^3 są przestrzeniami wektorowymi

$$\begin{aligned} (x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) &= (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n), \\ r(x_1, x_2, \dots, x_n) &= (rx_1, rx_2, \dots, rx_n). \end{aligned}$$

Ponieważ przestrzenie $K_{n\times 1}$, $K_{1\times n}$ i K^n są do siebie podobne, w wielu miejscach (a przynajmniej tam, gdzie nie jest istotne, czy wektor zapisujemy w postaci kolumny, wiersza lub ciągu) użyjemy wspólnego symbolu K^n na oznaczenie tych trzech przestrzeni i będziemy pisać $(x_1, x_2, \dots, x_n)$ zamiast

$$\begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \quad \text{lub} \quad \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Przykład 123. Niech X będzie dowolnym niepustym zbiorem, a K – ustalonym ciałem. Oznaczmy przez $\mathcal{F}(X, K)$ zbiór wszystkich funkcji $f : X \to K$. Dla $f, g \in \mathcal{F}(X, K)$ i $\alpha \in K$ określamy sumę $f + g$ i iloczyn αf jako zwykłe dodawanie funkcji i mnożenie funkcji przez skalar, tj. przyjmujemy, że dla każdego $x \in X$ jest

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

i

$$(\alpha f)(x) = \alpha f(x).$$

Z własności dodawania i mnożenia w ciele K łatwo wynika, że zbiór $\mathcal{F}(X, K)$ z wyżej określonymi działaniami jest przestrzenią wektorową.

Udowodnimy teraz pewne proste i ważne własności działań na wektorach.

Twierdzenie 7.1.1. *Dla wektorów* $\mathbf{x}$, $\mathbf{y}$ *z przestrzeni* $V(K)$ *i skalarów* α, β *z ciała* K *jest:*

(1) $\alpha\mathbf{x} = \mathbf{0}$ *wtedy i tylko wtedy, gdy* $\alpha = 0$ *lub* $\mathbf{x} = \mathbf{0}$*;*

(2) *jeśli* $\mathbf{x} \neq \mathbf{0}$ *i* $\alpha\mathbf{x} = \beta\mathbf{x}$*, to* $\alpha = \beta$*;*

(3) *jeśli* $\alpha \neq 0$ *i* $\alpha\mathbf{x} = \alpha\mathbf{y}$*, to* $\mathbf{x} = \mathbf{y}$*;*

(4) $(-\alpha)\mathbf{x} = \alpha(-\mathbf{x}) = -(\alpha\mathbf{x})$.

Dowód. (1) Z faktu, że w grupie $(V(K), +)$ jest $0\mathbf{x} + 0\mathbf{x} = (0+0)\mathbf{x} = 0\mathbf{x} = 0\mathbf{x} + \mathbf{0}$ i z twierdzenia 1.2.1 (o skracaniu w grupie) wynika, że $0\mathbf{x} = \mathbf{0}$. Podobnie z równości $\alpha\mathbf{0} + \alpha\mathbf{0} = \alpha(\mathbf{0} + \mathbf{0}) = \alpha\mathbf{0} = \alpha\mathbf{0} + \mathbf{0}$ wynika, że $\alpha\mathbf{0} = \mathbf{0}$. Załóżmy teraz, że $\alpha\mathbf{x} = \mathbf{0}$ i $\alpha \neq 0$. Wtedy α^{-1} istnieje i $\mathbf{x} = 1\mathbf{x} = (\alpha^{-1}\alpha)\mathbf{x} = \alpha^{-1}(\alpha\mathbf{x}) = \alpha^{-1}\mathbf{0} = \mathbf{0}$ i to kończy dowód pierwszej części twierdzenia.

(2) Załóżmy teraz, że $\mathbf{x} \neq \mathbf{0}$ i $\alpha\mathbf{x} = \beta\mathbf{x}$. Wtedy $(\alpha - \beta)\mathbf{x} = \mathbf{0}$ i z (1) otrzymujemy $\alpha - \beta = 0$, więc także $\alpha = \beta$.

(3) Jeśli $\alpha \neq 0$ i $\alpha\mathbf{x} = \alpha\mathbf{y}$, to $\alpha(\mathbf{x} - \mathbf{y}) = \mathbf{0}$ i wobec (1) jest $\mathbf{x} - \mathbf{y} = \mathbf{0}$.

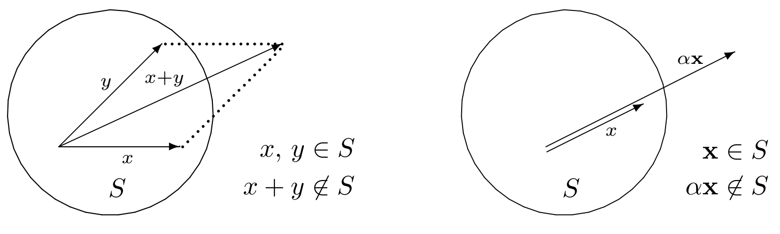
(4) Ponieważ $(-\alpha)\mathbf{x} + \alpha\mathbf{x} = \big((-\alpha) + \alpha\big)\mathbf{x} = 0\mathbf{x} = \mathbf{0}$, więc $(-\alpha)\mathbf{x} = -(\alpha\mathbf{x})$. Analogicznie z równości $\alpha(-\mathbf{x}) + \alpha\mathbf{x} = \alpha\big((-\mathbf{x}) + \mathbf{x}\big) = \alpha\mathbf{0} = \mathbf{0}$ wynika, że $\alpha(-\mathbf{x}) = -(\alpha\mathbf{x})$. $\square$

Definicja 7.1.3. Niech $V = V(K)$ będzie przestrzenią wektorową i niech S będzie podzbiorem zbioru V. Mówimy, że zbiór S jest *zamknięty ze względu na dodawanie*, jeśli

$$\forall_{\mathbf{x},\,\mathbf{y}\in S}\ \ \mathbf{x} + \mathbf{y} \in S.$$

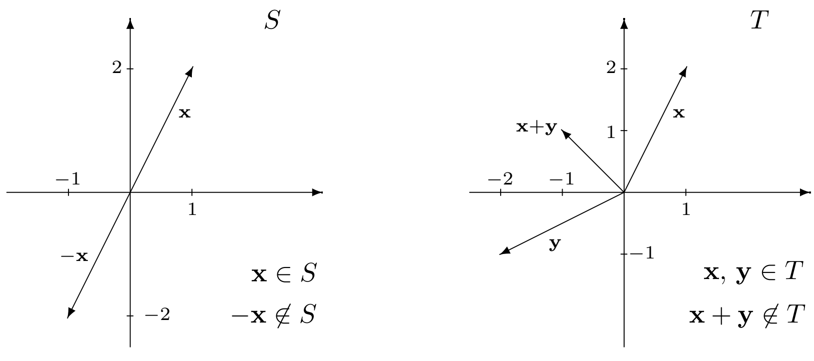
Podobnie zbiór S jest *zamknięty ze względu na mnożenie przez skalary*, jeśli

$$\forall_{\mathbf{x}\in S}\forall_{\alpha\in K}\ \ \alpha\mathbf{x} \in S.$$



Rys. 7.1. Zbiór S nie jest zamknięty ze względu na dodawanie wektorów, ani ze względu na mnożenie wektorów przez skalary

Przykład 124. W przestrzeni wektorowej R^2 zbiór $S = \{(x, y) \in R^2 : x,\, y \geqslant 0\}$ jest zamknięty ze względu na dodawanie, ale nie jest on zamknięty ze względu na mnożenie skalarne, bo przykładowo $(1, 2) \in S$ i $-1 \in R$, ale $-1(1, 2) \notin S$. Zbiór $T = \{(x, y) \in R^2 : xy \geqslant 0\}$ jest zamknięty ze względu na mnożenie skalarne, ale nie jest on zamknięty ze względu na dodawanie: wektory $(1, 2)$ i $(-2, -1)$ należą do zbioru T, ale ich suma $(1, 2) + (-2, -1) = (-1, 1)$ już nie należy do zbioru T, zob. rys. 7.2.



Rys. 7.2

Definicja 7.1.4. Niech $(V, K, +, \cdot)$ będzie przestrzenią wektorową i niech W będzie podzbiorem zbioru V. Mówimy, że W jest *podprzestrzenią* (przestrzeni wektorowej V), gdy $(W, K, +, \cdot)$ jest przestrzenią wektorową (gdzie $+$ i $\cdot$ są działaniami z przestrzeni V obciętymi do zbioru W).

Podprzestrzeń przestrzeni wektorowej

Procedura sprawdzania, czy dany niepusty podzbiór wektorów jest podprzestrzenią, jest prosta i – jak to wynika z następującego twierdzenia – ogranicza się do weryfikacji zamkniętości tego podzbioru ze względu na dodawanie i mnożenie przez skalary (rys. 7.3).

Twierdzenie 7.1.2. *Podzbiór S zbioru wektorów przestrzeni V jest podprzestrzenią wtedy i tylko wtedy, gdy spełnione są warunki:*

(a) $S \neq \emptyset$*;*

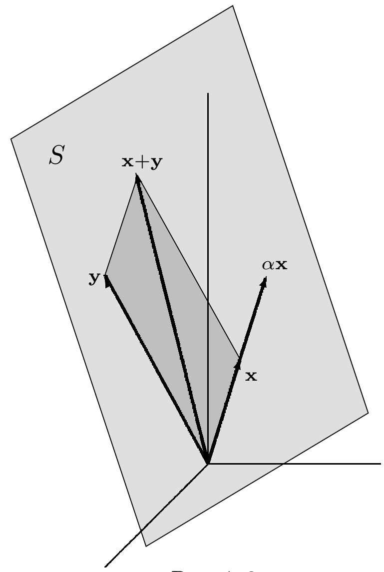
(b) $\forall_{\mathbf{x},\mathbf{y}\in S}\ \mathbf{x}+\mathbf{y} \in S$*;*

(c) $\forall_{\mathbf{x}\in S}\forall_{\alpha\in K}\ \alpha\mathbf{x} \in S$.

Dowód. Jeśli S jest podprzestrzenią przestrzeni V, to oczywiście S ma własności (a), (b) i (c).

Załóżmy teraz, że podzbiór S zbioru V ma własności (a), (b) i (c). Musimy pokazać, że S spełnia wszystkie warunki definicji przestrzeni. Większość z nich wynika z faktu, że S jest podzbiorem zbioru V. Dla przykładu, jeśli $\mathbf{x}, \mathbf{y} \in S$, to $\mathbf{x}+\mathbf{y},\ \mathbf{y}+\mathbf{x} \in S$ (wobec (b)), więc także $\mathbf{x}+\mathbf{y},\ \mathbf{y}+\mathbf{x} \in V$. Ponieważ $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$ w przestrzeni V, więc także $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$ w zbiorze S. To oznacza, że dodawanie w zbiorze S jest przemienne. Analogicznie uzasadnia się łączność dodawania i wszystkie własności mnożenia skalarnego w zbiorze S. Zatem pozostaje pokazać, że w S jest wektor zerowy i każdy wektor z S ma wektor przeciwny w S.

Na mocy (a) istnieje co najmniej jeden wektor $\mathbf{x} \in S$. Wobec (c) iloczyn $0\mathbf{x}$ należy do S, ale $0\mathbf{x} = \mathbf{0}$ (zob. tw. 7.1.1), więc wektor zerowy $\mathbf{0}$ należy do S. Weźmy teraz dowolny wektor $\mathbf{x} \in S$. Ponieważ $(-1)\mathbf{x} \in S$ (wobec (c)) i $(-1)\mathbf{x} = -\mathbf{x}$ (zob. tw. 7.1.1), więc $-\mathbf{x} \in S$. Zatem S jest przestrzenią wektorową i dlatego S jest podprzestrzenią przestrzeni V. $\square$



Rys. 7.3

Łatwo zauważyć, że twierdzenie 7.1.2 jest równoważne następującemu twierdzeniu.

Twierdzenie 7.1.3. *Niepusty zbiór S wektorów przestrzeni V jest podprzestrzenią wtedy i tylko wtedy, gdy*

$$\forall_{\mathbf{x},\mathbf{y}\in S}\forall_{\alpha,\beta\in K}\ \alpha\mathbf{x}+\beta\mathbf{y} \in S. \ \square$$

Jeśli V jest przestrzenią wektorową, to cały zbiór V spełnia warunki twierdzenia 7.1.2 (i 7.1.3). Zatem przestrzeń V jest swoją własną podprzestrzenią. Inne podprzestrzenie przestrzeni V (jeśli takie istnieją) są *właściwymi podprzestrzeniami* w V. Zbiór $\{\mathbf{0}\}$, zawierający tylko wektor zerowy przestrzeni V, też jest podprzestrzenią przestrzeni V. Nazywamy ją *podprzestrzenią zerową* (lub *trywialną*) przestrzeni V.

Przykład 125. Zbiór wielomianów rzeczywistych $R[x]$ jest zawarty w przestrzeni wektorowej $\mathcal{F}(R, R)$ (z przykładu 123). Ponieważ $f(x) = x \in R[x]$, więc $R[x]$ jest niepustym podzbiorem zbioru $\mathcal{F}(R, R)$. Dodatkowo, ponieważ suma wielomianów jest wielomianem i iloczyn wielomianu przez liczbę rzeczywistą jest wielomianem, więc wobec twierdzenia 7.1.2 zbiór wielomianów $R[x]$ jest podprzestrzenią przestrzeni wektorowej $\mathcal{F}(R, R)$.

$R[x]$ – przestrzeń

Niech teraz n będzie nieujemną liczbą całkowitą i niech $R_n[x]$ będzie zbiorem wielomianów rzeczywistych stopnia co najwyżej n. Ponieważ $f(x) = 1 \in R_n[x]$, więc $R_n[x]$ jest niepustym podzbiorem zbioru $R[x]$. Ponieważ suma wielomianów

$R_n[x]$ – przestrzeń

stopnia co najwyżej n jest wielomianem stopnia co najwyżej n i iloczyn wielomianu stopnia co najwyżej n przez liczbę rzeczywistą jest wielomianem stopnia co najwyżej n, więc z twierdzenia 7.1.2 wynika, że $R_n[x]$ jest podprzestrzenią przestrzeni $R[x]$ (oraz przestrzeni $\mathcal{F}(R,R)$).

Przykład 126. Zbiór $\mathcal{A} = \{f \in \mathcal{F}(R,R) : f(1) = 0\}$ jest podprzestrzenią przestrzeni $\mathcal{F}(R,R)$.

Zbiór $\mathcal{A}$ jest niepusty, bo np. funkcja $f(x) = x - 1$ należy do zbioru $\mathcal{A}$. Weźmy teraz pod uwagę dowolne funkcje f i g ze zbioru $\mathcal{A}$ oraz dowolne skalary $\alpha, \beta \in R$. Wtedy $\alpha f + \beta g \in \mathcal{F}(R,R)$. Jednocześnie $f(1) = g(1) = 0$, więc także $(\alpha f + \beta g)(1) = (\alpha f)(1) + (\beta g)(1) = \alpha(f(1)) + \beta(g(1)) = \alpha 0 + \beta 0 = 0$ i dlatego $\alpha f + \beta g \in \mathcal{A}$. Stąd i z twierdzenia 7.1.3 wynika, że $\mathcal{A}$ jest podprzestrzenią przestrzeni $\mathcal{F}(R,R)$.

Przykład 127. Zbiór $S = \{(x,y,z) \in R^3 : x + 2y + 3z = 0\}$ jest właściwą podprzestrzenią przestrzeni R^3.

Ponieważ $(0,0,0) \in S$ i $(1,1,1) \in R^3 - S$, więc S jest właściwym podzbiorem zbioru R^3. Weźmy teraz dowolne wektory $\mathbf{x} = (x,y,z)$ i $\mathbf{y} = (x',y',z')$ ze zbioru S i dowolne skalary α, β z ciała R. Wtedy $x + 2y + 3z = 0$ oraz $x' + 2y' + 3z' = 0$ i dlatego dla wektora

$$\alpha\mathbf{x} + \beta\mathbf{y} = \alpha(x,y,z) + \beta(x',y',z') = (\alpha x + \beta x', \alpha y + \beta y', \alpha z + \beta z')$$

jest

$$(\alpha x + \beta x') + 2(\alpha y + \beta y') + 3(\alpha z + \beta z') = \alpha(x + 2y + 3z) + \beta(x' + 2y' + 3z') = 0.$$

To dowodzi, że $\alpha\mathbf{x} + \beta\mathbf{y} \in S$. Stąd i z twierdzenia 7.1.3 wynika, że S jest właściwą podprzestrzenią w R^3.

$N_{\mathbf{A}}$ – przestrzeń zerowa macierzy

Definicja 7.1.5. Jeśli $\mathbf{A}$ jest macierzą, powiedzmy $\mathbf{A} \in K_{m\times n}$, to zbiór

$$N_{\mathbf{A}} = \{\mathbf{x} \in K_{n\times 1} \colon \mathbf{A}\mathbf{x} = \mathbf{0}\},$$

czyli zbiór rozwiązań równania jednorodnego $\mathbf{A}\mathbf{x} = \mathbf{0}$, nazywamy *przestrzenią zerową* (lub *przestrzenią rozwiązań*) macierzy $\mathbf{A}$.

Z twierdzeń 5.3.1 i 7.1.3 wynika, że $N_{\mathbf{A}}$ jest podprzestrzenią przestrzeni $K_{n\times 1}$ (gdy $\mathbf{A} \in K_{m\times n}$).

Przykład 128. Wyznaczyć przestrzeń zerową $N_{\mathbf{A}}$ rzeczywistej macierzy

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & -1 \\ 6 & 1 & -5 \end{bmatrix}.$$

Przestrzenią zerową macierzy $\mathbf{A}$ jest zbiór wszystkich rozwiązań równania $\mathbf{A}\mathbf{x} = \mathbf{0}$, czyli równania

$$\begin{bmatrix} 3 & -1 & -1 \\ 6 & 1 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Ponieważ

$$[\mathbf{A}|\mathbf{0}] = \left[\begin{array}{rrr|r} 3 & -1 & -1 & 0 \\ 6 & 1 & -5 & 0 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & 0 & -2/3 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right],$$

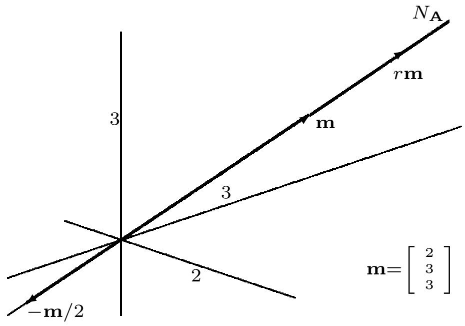
więc rozwiązaniem równania $\mathbf{Ax} = \mathbf{0}$ jest każda macierz

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2r \\ 3r \\ 3r \end{bmatrix} \quad \text{dla} \quad r \in R.$$

Stąd

$$N_{\mathbf{A}} = \left\{ \begin{bmatrix} 2r \\ 3r \\ 3r \end{bmatrix} : r \in R \right\} = \left\{ r \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix} : r \in R \right\}$$

i jest to prosta w przestrzeni R^3 (zob. rys. 7.4).



Rys. 7.4

7.2. Kombinacje liniowe wektorów

Definicja 7.2.1. Jeśli $\mathbf{v}_1, \ldots, \mathbf{v}_n$ są wektorami z przestrzeni $V(K)$, a α_1, α_2, $\ldots, \alpha_n$ są skalarami z ciała K, to wektor

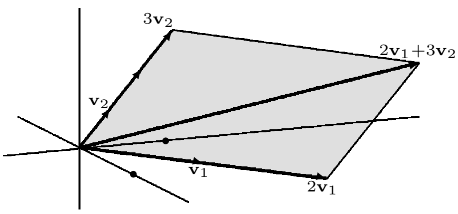
$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n = \sum_{i=1}^{n} \alpha_i\mathbf{v}_i$$

Kombinacja liniowa wektorów

nazywamy *kombinacją liniową* wektorów $\mathbf{v}_1, \ldots, \mathbf{v}_n$ o współczynnikach α_1, α_2, $\ldots, \alpha_n$.

Przykład 129. Wektor $\mathbf{v} = (4, 7, 3)$ (zob. rys. 7.5) jest kombinacją liniową wektorów $\mathbf{v}_1 = (2, 2, 0)$ i $\mathbf{v}_2 = (0, 1, 1)$, bo

$$2\mathbf{v}_1 + 3\mathbf{v}_2 = 2(2, 2, 0) + 3(0, 1, 1) = (4, 7, 3) = \mathbf{v}.$$



Rys. 7.5

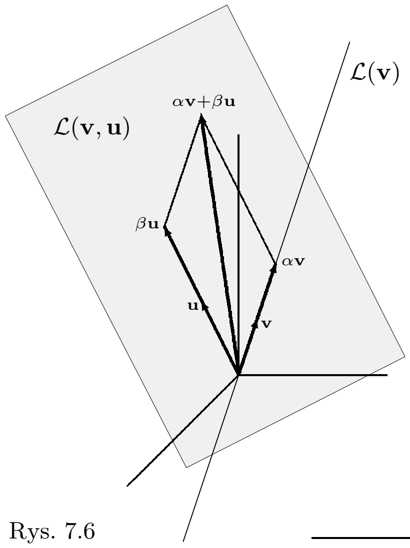
Jeśli S jest podzbiorem zbioru $V(K)$, to przez $\mathcal{L}(S)$ oznaczamy zbiór wszystkich kombinacji liniowych skończonej liczby wektorów ze zbioru S. Zatem

$$\mathcal{L}(S) = \left\{ \sum_{i=1}^{n} \alpha_i\mathbf{v}_i \colon \mathbf{v}_i \in S,\ \alpha_i \in K \text{ dla } i = 1, \ldots, n, \text{ gdzie } n \in N \right\}.$$

$\mathcal{L}(S)$

Przy tym przyjmujemy, że $\mathcal{L}(\emptyset) = \{\mathbf{0}\}$. Jeśli S jest niepustym zbiorem skończonym, powiedzmy $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, to piszemy $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ zamiast $\mathcal{L}(\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\})$. W tym przypadku

$$\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = \{\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n \colon \alpha_1, \alpha_2, \ldots, \alpha_n \in K\}.$$



Rys. 7.6

W szczególności, zbiór $\mathcal{L}(\mathbf{v})$ jest zbiorem wszystkich wektorów postaci $\alpha\mathbf{v}$. Jeśli wektor $\mathbf{v}$ jest niezerowy, to zbiór $\mathcal{L}(\mathbf{v})$ zwykle utożsamiamy z prostą przechodzącą przez punkt $\mathbf{0}$ i równoległą do wektora $\mathbf{v}$, zob. rys. 7.6. Natomiast $\mathcal{L}(\mathbf{v},\mathbf{u})$ jest zbiorem wszystkich wektorów postaci $\alpha\mathbf{v}+\beta\mathbf{u}$ i zbiór ten możemy utożsamiać z płaszczyzną (także zob. rys. 7.6), gdy żaden z wektorów $\mathbf{v}$ i $\mathbf{u}$ nie jest krotnością drugiego wektora. Zauważmy, że wektor $\mathbf{b}$ należy do zbioru $\mathcal{L}(\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n)$ wtedy i tylko wtedy, gdy istnieją skalary $\alpha_1,\alpha_2,\ldots,\alpha_n\in K$ takie, że

$$\alpha_1\mathbf{v}_1+\alpha_2\mathbf{v}_2+\ldots+\alpha_n\mathbf{v}_n=\mathbf{b},$$

tj. wtedy i tylko wtedy, gdy równanie wektorowe

$$x_1\mathbf{v}_1+x_2\mathbf{v}_2+\ldots+x_n\mathbf{v}_n=\mathbf{b}$$

ma co najmniej jedno rozwiązanie.

Przykład 130. Dane są wektory

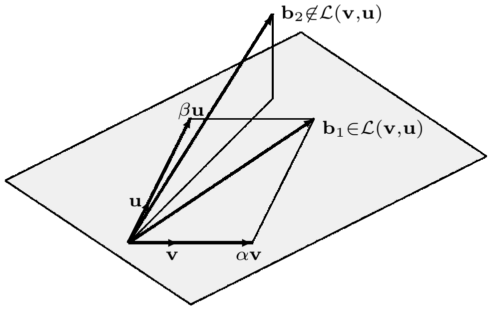
$$\mathbf{v}=\begin{bmatrix}0\\1\\-1\end{bmatrix},\quad \mathbf{u}=\begin{bmatrix}2\\1\\3\end{bmatrix},\quad \mathbf{b}_1=\begin{bmatrix}6\\5\\7\end{bmatrix},\quad \mathbf{b}_2=\begin{bmatrix}2\\1\\0\end{bmatrix}.$$

Sprawdzić, który z wektorów $\mathbf{b}_1$, $\mathbf{b}_2$ należy do zbioru $\mathcal{L}(\mathbf{v},\mathbf{u})$.

Wektor $\mathbf{b}_i$ należy do zbioru $\mathcal{L}(\mathbf{v},\mathbf{u})$ wtedy i tylko wtedy, gdy równanie $x_1\mathbf{v}+x_2\mathbf{u}=\mathbf{b}_i$ ma rozwiązanie. Tak jest wtedy i tylko wtedy, gdy $\mathbf{b}_i$ nie jest wiodącą kolumną macierzy $[\,\mathbf{v}\,\mathbf{u}\,|\,\mathbf{b}_i\,]$. Ponieważ

$$[\,\mathbf{v}\,\mathbf{u}\,|\,\mathbf{b}_1\,\mathbf{b}_2\,]=\left[\begin{array}{rr|rr}0&2&6&2\\1&1&5&1\\-1&3&7&0\end{array}\right]\sim\ldots\sim\left[\begin{array}{rr|rr}1&1&5&1\\0&1&3&1\\0&0&0&1\end{array}\right],$$

więc $\mathbf{b}_i$ nie jest kolumną wiodącą macierzy $[\,\mathbf{v}\,\mathbf{u}\,|\,\mathbf{b}_i\,]$ tylko dla $i=1$. Zatem $\mathbf{b}_1\in\mathcal{L}(\mathbf{v},\mathbf{u})$ i $\mathbf{b}_2\notin\mathcal{L}(\mathbf{v},\mathbf{u})$, zob. rys. 7.7.



Rys. 7.7

Przykład 131. Uzasadnimy, że w przestrzeni wektorowej $R[x]$ jest

$$\mathcal{L}(1,x,x^2,\ldots)=R[x].$$

Jeśli $\varphi(x)\in R[x]$, powiedzmy $\varphi(x)=a_0+a_1x+\ldots+a_kx^k$, to oczywiście $\varphi(x)$ jest kombinacją liniową wektorów $1,x,x^2,\ldots,x^k$ i dlatego $\varphi(x)\in\mathcal{L}(1,x,x^2,\ldots)$. Zatem $R[x]\subseteq\mathcal{L}(1,x,x^2,\ldots)$.

Z drugiej strony zbiór $\{1,x,x^2,\ldots\}$ jest nieskończony, ale na podstawie definicji każdy wektor $\psi(x)$ ze zbioru $\mathcal{L}(1,x,x^2,\ldots)$ jest kombinacją liniową skończonej ilości wektorów ze zbioru $\{1,x,x^2,\ldots\}$,

$$\psi(x)=\alpha_{k_1}x^{k_1}+\ldots+\alpha_{k_m}x^{k_m}$$

(dla pewnych liczb $k_1, \ldots, k_m \in \{0, 1, 2, \ldots\}$ i pewnych liczb $\alpha_{k_1}, \ldots, \alpha_{k_m} \in R$), więc $\psi(x)$ jest wielomianem. Stąd wynika, że $\mathcal{L}(1, x, x^2, \ldots) \subseteq R[x]$ i dlatego też $\mathcal{L}(1, x, x^2, \ldots) = R[x]$.

Zbiór $\mathcal{L}(S)$ kombinacji liniowych wektorów ze zbioru $S \subseteq V(K)$ nie tylko jest podzbiorem zbioru $V(K)$, ale i jest on podprzestrzenią przestrzeni $V(K)$.

Twierdzenie 7.2.1. *Jeśli S jest zbiorem wektorów przestrzeni $V(K)$, to $\mathcal{L}(S)$ jest podprzestrzenią przestrzeni $V(K)$.*

$\mathcal{L}(S)$ – przestrzeń wektorowa

Dowód. Jeśli $S = \emptyset$, to $\mathcal{L}(S) = \{\mathbf{0}\}$ i teza twierdzenia jest oczywista. Zatem załóżmy, że zbiór S jest niepusty. Ponieważ $\mathbf{v} = 1\mathbf{v}$ i $1\mathbf{v} \in \mathcal{L}(S)$ dla każdego $\mathbf{v} \in S$, więc $S \subseteq \mathcal{L}(S)$ i zbiór $\mathcal{L}(S)$ jest niepusty. Wobec twierdzenia 7.1.3 wystarczy teraz pokazać, że kombinacja liniowa wektorów ze zbioru $\mathcal{L}(S)$ należy do zbioru $\mathcal{L}(S)$.

Weźmy dowolne dwa wektory $\mathbf{v}, \mathbf{u} \in \mathcal{L}(S)$ i skalary $\alpha, \beta \in K$. Z definicji zbioru $\mathcal{L}(S)$ wynika, że istnieją wektory $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{u}_1, \ldots, \mathbf{u}_m \in S$ i skalary $\alpha_1, \ldots, \alpha_n$, $\beta_1, \ldots, \beta_m \in K$, dla których $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$ i $\mathbf{u} = \sum_{j=1}^{m} \beta_j \mathbf{u}_j$. Wtedy

$$\alpha\mathbf{v} + \beta\mathbf{u} = (\alpha\alpha_1)\mathbf{v}_1 + \ldots + (\alpha\alpha_n)\mathbf{v}_n + (\beta\beta_1)\mathbf{u}_1 + \ldots + (\beta\beta_m)\mathbf{u}_m$$

i stąd widać, że wektor $\alpha\mathbf{v} + \beta\mathbf{u}$ jest kombinacją liniową wektorów $\mathbf{v}_1, \ldots, \mathbf{v}_n$, $\mathbf{u}_1, \ldots,$ $\mathbf{u}_m$ ze zbioru S (o współczynnikach $\alpha\alpha_i$ i $\beta\beta_j$ z ciała K), więc $\alpha\mathbf{x} + \beta\mathbf{y} \in \mathcal{L}(S)$. $\square$

Definicja 7.2.2. Jeśli S jest zbiorem wektorów przestrzeni $V(K)$, to wobec twierdzenia 7.2.1 zbiór $\mathcal{L}(S)$ jest podprzestrzenią przestrzeni $V(K)$. Podprzestrzeń $\mathcal{L}(S)$ nazywa się *podprzestrzenią generowaną* przez zbiór S, zbiór S – *zbiorem generującym* przestrzeń $\mathcal{L}(S)$, a elementy zbioru S – *generatorami przestrzeni* $\mathcal{L}(S)$. Mówimy także, że zbiór S *generuje* (pod)przestrzeń $\mathcal{L}(S)$. Czasami mówi się, że $\mathcal{L}(S)$ jest powłoką liniową zbioru S.

$\mathcal{L}(S)$ – podprzestrzeń generowana przez zbiór S

Każda przestrzeń wektorowa V jest generowana przez pewien zbiór wektorów, np. zawsze mamy $V = \mathcal{L}(V)$. Jeśli przestrzeń V ma skończony zbiór generujący, to mówimy, że jest ona *skończenie generowana.* Zatem przestrzeń V jest skończenie generowana, gdy $V = \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ dla pewnej liczby naturalnej n i pewnych wektorów $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$.

Przestrzeń skończenie generowana

Przykład 132. Przestrzeń $R_n[x]$ jest skończenie generowana, bo

$$R_n[x] = \mathcal{L}(1, x, x^2, \ldots, x^n).$$

Ponieważ do przestrzeni $R[x]$ należą wielomiany dowolnie wysokiego stopnia, więc żaden skończony zbiór wielomianów nie generuje przestrzeni $R[x]$. Dlatego przestrzeń $R[x]$ nie jest skończenie generowana.

Przykład 133. Wyznaczyć generatory przestrzeni rozwiązań jednorodnego układu równań

$$\begin{cases} x_1 + \ \ x_2 + 2x_3 - 2x_4 = 0, \\ x_1 + 2x_2 + 3x_3 + 4x_4 = 0. \end{cases}$$

Niech $\mathbf{A}$ będzie macierzą główną powyższego układu. Układ ten rozwiązujemy metodą Gaussa-Jordana sprowadzając jego macierz rozszerzoną $[\mathbf{A}|\mathbf{0}]$ do postaci schodkowej. Ponieważ

$$[\mathbf{A}|\mathbf{0}] = \left[\begin{array}{cccc|c} 1 & 1 & 2 & -2 & 0 \\ 1 & 2 & 3 & 4 & 0 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 1 & 2 & -2 & 0 \\ 0 & 1 & 1 & 6 & 0 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 0 & 1 & -8 & 0 \\ 0 & 1 & 1 & 6 & 0 \end{array}\right],$$

więc każde rozwiązanie powyższego układu jest postaci

$$\mathbf{x} = \begin{bmatrix} -x_3 + 8x_4 \\ -x_3 - 6x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 8 \\ -6 \\ 0 \\ 1 \end{bmatrix},$$

gdzie x_3 i x_4 są dowolnymi liczbami rzeczywistymi. Zatem przestrzeń rozwiązań tego układu (i przestrzeń zerowa jego macierzy głównej) jest generowana przez wektory $\mathbf{v}_1 = \begin{bmatrix} -1 & -1 & 1 & 0 \end{bmatrix}^T$ i $\mathbf{v}_2 = \begin{bmatrix} 8 & -6 & 0 & 1 \end{bmatrix}^T$.

7.3. Przestrzeń kolumnowa macierzy

$C_{\mathbf{A}}$ – przestrzeń kolumnowa macierzy $\mathbf{A}$

Definicja 7.3.1. *Przestrzenią kolumnową* macierzy $\mathbf{A} \in K_{m\times n}$, oznaczamy ją przez $C_{\mathbf{A}}$, nazywamy przestrzeń generowaną przez kolumny macierzy $\mathbf{A}$. Zatem, jeśli $\mathbf{A} = [\mathbf{a}_1\, \mathbf{a}_2\, \ldots\, \mathbf{a}_n]$, to

$$C_{\mathbf{A}} = \mathcal{L}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n).$$

Przykład 134. Wyznaczyć przestrzeń $C_{\mathbf{A}}$ macierzy $\mathbf{A} = \begin{bmatrix} -2 & 1 \\ 0 & 2 \\ 1 & 0 \end{bmatrix}$.

Mamy

$$\begin{aligned} C_{\mathbf{A}} &= \left\{ x_1 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} :\ x_1,\, x_2 \in R \right\} \\ &= \left\{ \begin{bmatrix} -2 & 1 \\ 0 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} :\ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in R^2 \right\} \\ &= \left\{ \mathbf{Ax} :\ \mathbf{x} \in R^2 \right\}. \end{aligned}$$

Korzystając z pojęcia przestrzeni kolumnowej macierzy, przedstawimy teraz kolejny warunek konieczny i dostateczny istnienia rozwiązania układu $\mathbf{Ax} = \mathbf{b}$. Niech $\mathbf{a}_1, \mathbf{a}_2, \ldots\, \mathbf{a}_n$ będą kolejnymi kolumnami macierzy $\mathbf{A}$. Ponieważ dla przestrzeni kolumnowej macierzy $\mathbf{A} = \begin{bmatrix}\mathbf{a}_1\, \mathbf{a}_2\, \ldots\, \mathbf{a}_n\end{bmatrix}$ mamy

$$\begin{aligned} C_{\mathbf{A}} = \mathcal{L}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) &= \{x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n :\ x_1, \ldots, x_n \in K\} \\ &= \left\{ [\mathbf{a}_1\, \mathbf{a}_2\, \ldots\, \mathbf{a}_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} : \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in K^n \right\} \\ &= \{\mathbf{Ax} :\ \mathbf{x} \in K^n\} \\ &= \{\mathbf{b} : \mathbf{b} = \mathbf{Ax} \text{ dla pewnego } \mathbf{x} \in K^n\}, \end{aligned}$$

więc układ równań $\mathbf{Ax} = \mathbf{b}$ ma rozwiązanie wtedy i tylko wtedy, gdy wektor $\mathbf{b}$ należy do przestrzeni $C_{\mathbf{A}}$. Stąd natychmiast otrzymujemy następujące twierdzenie.

Twierdzenie 7.3.1. *Układ równań liniowych* $\mathbf{Ax} = \mathbf{b}$, *gdzie* $\mathbf{A} \in K_{m\times n}$ *i* $\mathbf{b} \in K_{m\times 1}$, *ma rozwiązanie wtedy i tylko wtedy, gdy wektor* $\mathbf{b}$ *należy do przestrzeni kolumnowej macierzy* $\mathbf{A}$. *Równoważnie, układ* $\mathbf{Ax} = \mathbf{b}$ *ma rozwiązanie wtedy i tylko wtedy, gdy*

$$C_{\mathbf{A}} = C_{[\mathbf{A}|\mathbf{b}]}, \tag{7.1}$$

tj. wtedy i tylko wtedy, gdy przestrzeń kolumnowa macierzy $\mathbf{A}$ *jest równa przestrzeni kolumnowej macierzy rozszerzonej* $[\mathbf{A}|\mathbf{b}]$ *układu* $\mathbf{Ax} = \mathbf{b}$. $\square$

Przykład 135. Dany jest układ równań $\mathbf{Ax} = \mathbf{b}$, gdzie

$$\mathbf{A} = [\,\mathbf{a}_1,\, \mathbf{a}_2,\, \mathbf{a}_3\,] = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix} \quad \text{i} \quad \mathbf{b} = \begin{bmatrix} a \\ 2a \\ a+b \\ 2a+b \end{bmatrix}.$$

Dla jakich a i b układ ten ma rozwiązanie?

Wobec twierdzenia 7.3.1 układ $\mathbf{Ax} = \mathbf{b}$ ma rozwiązanie wtedy i tylko wtedy, gdy $\mathbf{b} \in \mathcal{L}(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$. Ponieważ $\mathcal{L}(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) = \mathcal{L}(\mathbf{a}_1, \mathbf{a}_2)$ i

$$[\,\mathbf{a}_1,\, \mathbf{a}_2\,|\,\mathbf{b}\,] = \left[\begin{array}{cc|c} 1 & 1 & a \\ 2 & 1 & 2a \\ 3 & 1 & a+b \\ 4 & 1 & 2a+b \end{array}\right] \sim \ldots \sim \left[\begin{array}{cc|c} 1 & 0 & a \\ 0 & 1 & 0 \\ 0 & 0 & b-2a \\ 0 & 0 & 0 \end{array}\right],$$

więc układ $\mathbf{Ax} = \mathbf{b}$ ma rozwiązanie wtedy i tylko wtedy, gdy $b = 2a$ i a jest dowolną liczbą rzeczywistą.

Przykład 136. Sprawdzić przynależność wektora $\mathbf{b} = (5, 0, 11)$ do podprzestrzeni generowanej przez wektory $\mathbf{v}_1 = (1, -2, 2)$, $\mathbf{v}_2 = (1, 0, 4)$ i $\mathbf{v}_3 = (2, -4, 1)$.

Niech $\mathbf{A}$ będzie macierzą, której kolumnami są $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$. Ponieważ mamy $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = C_{\mathbf{A}}$, więc wobec twierdzenia 7.3.1 wektor $\mathbf{b}$ należy do podprzestrzeni $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ wtedy i tylko wtedy, gdy układ równań $\mathbf{Ax} = \mathbf{b}$ ma rozwiązanie. Dla macierzy rozszerzonej tego układu mamy równoważności

$$\left[\begin{array}{rrr|r} 1 & 1 & 2 & 5 \\ -2 & 0 & -4 & 0 \\ 2 & 4 & 1 & 11 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & 1 & 2 & 5 \\ 0 & 2 & 0 & 10 \\ 0 & 2 & -3 & 1 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & 1 & 2 & 5 \\ 0 & 2 & 0 & 10 \\ 0 & 0 & -3 & -9 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & 1 & 2 & 5 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

i z ostatniej macierzy widać, że układ $\mathbf{Ax} = \mathbf{b}$ ma rozwiązanie. Zatem wektor $\mathbf{b}$ należy do podprzestrzeni $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$.

Podamy teraz prosty sposób sprawdzania, czy danych n wektorów z przestrzeni K^n generuje całą przestrzeń K^n. Niech $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ będą wektorami z przestrzeni K^n i niech $\mathbf{A}$ będzie macierzą wymiaru $n \times n$, której kolumnami są $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, czyli $\mathbf{A} = [\mathbf{v}_1\, \mathbf{v}_2 \ldots \mathbf{v}_n]$. Ponieważ $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) \subseteq K^n$, więc mamy następujący ciąg równoważności:

$$\begin{aligned} \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = K^n &\Leftrightarrow K^n \subseteq \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) \\ &\Leftrightarrow \text{każdy wektor } \mathbf{b} \text{ z przestrzeni } K^n \text{ jest kombinacją liniową wektorów } \mathbf{v}_1,\, \mathbf{v}_2, \ldots, \mathbf{v}_n \\ &\Leftrightarrow \text{układ } \mathbf{Ax} = \mathbf{b} \text{ ma rozwiązanie dla każdego wektora } \mathbf{b} \in K^n \text{ (twierdzenie 7.3.1)} \\ &\Leftrightarrow \text{macierz } \mathbf{A} \text{ jest odwracalna (twierdzenie 5.3.1)} \\ &\Leftrightarrow \text{macierz } \mathbf{A} \text{ jest nieosobliwa (wniosek 6.3.1).} \end{aligned}$$

Zatem za pomocą wyznacznika możemy badać czy danych n wektorów z przestrzeni K^n generuje całą przestrzeń K^n i mamy następujące twierdzenie.

Twierdzenie 7.3.2. *Wektory $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ z przestrzeni K^n generują całą przestrzeń K^n wtedy i tylko wtedy, gdy macierz $\mathbf{A} = [\mathbf{v}_1\, \mathbf{v}_2 \ldots \mathbf{v}_n]$ jest nieosobliwa.* $\square$

Przykład 137. Czy wektory

$$\mathbf{v}_1 = (1, 2, -2), \quad \mathbf{v}_2 = (3, 2, 1) \quad \text{i} \quad \mathbf{v}_3 = (-1, 2, -5)$$

generują całą przestrzeń R^3?

Wobec twierdzenia 7.3.2 wektory $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ z przestrzeni R^3 generowałyby całą przestrzeń R^3 wtedy i tylko wtedy, gdyby macierz $\mathbf{A} = [\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3]$ była nieosobliwa. W tym przypadku

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

więc macierz $\mathbf{A}$ jest osobliwa i dlatego wektory $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ nie generują całej przestrzeni R^3.

Wniosek 7.3.1. *Żadnych k wektorów z przestrzeni K^n nie generuje całej przestrzeni K^n, gdy $k < n$.*

Dowód. Przypuśćmy, że wektory $\mathbf{v}_1, \ldots, \mathbf{v}_k$ generują przestrzeń K^n i $k < n$. Wtedy także wektory $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1} = \mathbf{0}, \ldots, \mathbf{v}_n = \mathbf{0}$ generują przestrzeń K^n i wobec twierdzenia 7.3.2 macierz $[\mathbf{v}_1 \ldots \mathbf{v}_k\, \mathbf{0} \ldots \mathbf{0}]$ (mająca $n - k > 0$ zerowych kolumn) jest nieosobliwa, co jest niemożliwe. □

Niech teraz k będzie liczbą naturalną większą od n. Dla sprawdzenia czy wektory $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ z przestrzeni K^n generują przestrzeń K^n tworzymy macierz $\mathbf{A} = [\mathbf{v}_1\, \mathbf{v}_2\, \ldots\, \mathbf{v}_k]$ wymiaru $n \times k$ i za pomocą operacji elementarnych na wierszach przekształcamy ją w wierszowo równoważną macierz $\mathbf{B}$ mającą postać schodkową. Jeśli $\mathbf{B}$ ma n wiodących jedynek, kolumny macierzy $\mathbf{A}$ odpowiadające wiodącym jedynkom z $\mathbf{B}$ tworzą macierz odwracalną i wobec twierdzenia 7.3.2 generują one całą przestrzeń K^n. Jeśli macierz $\mathbf{B}$ ma mniej niż n wiodących jedynek, to (tak jak w dowodzie twierdzenia 5.3.1) ostatni wiersz macierzy $\mathbf{B}$ jest zerowy i układ $\mathbf{Bx} = \mathbf{e}_n = [0 \ldots 0\, 1]^T$ jest sprzeczny. Ponieważ macierz rozszerzona $[\mathbf{B}|\mathbf{e}_n]$ tego układu jest wierszowo równoważna macierzy $[\mathbf{A}|\mathbf{b}]$ (dla pewnego wektora $\mathbf{b} \in K^n$), układ $\mathbf{Ax} = \mathbf{b}$ jest sprzeczny, co oznacza, że $\mathbf{b}$ nie należy do przestrzeni kolumnowej macierzy $\mathbf{A}$. Zatem kolumny $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ macierzy $\mathbf{A}$ nie generują przestrzeni K^n. Z tych rozważań wynika następujący wniosek.

Wniosek 7.3.2. *Wektory $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ z przestrzeni K^n ($k > n$) generują przestrzeń K^n wtedy i tylko wtedy, gdy macierz $\mathbf{A} = [\mathbf{v}_1\, \mathbf{v}_2\, \ldots\, \mathbf{v}_k]$ jest wierszowo równoważna macierzy schodkowej mającej dokładnie n wiodących jedynek.* □

Przykład 138. Czy przestrzeń R^3 jest generowana przez wektory

$$\mathbf{v}_1 = (1, -2, 2), \quad \mathbf{v}_2 = (0, 1, 1), \quad \mathbf{v}_3 = (1, 0, 4) \quad \text{i} \quad \mathbf{v}_4 = (2, 3, 6)?$$

Tworzymy macierz $\mathbf{A} = [\mathbf{v}_1\, \ldots\, \mathbf{v}_4]$ i przekształcamy ją w wierszowo równoważną macierz mającą postać schodkową,

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 & 2 \\ -2 & 1 & 0 & 3 \\ 2 & 1 & 4 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 2 & 7 \\ 0 & 2 & 4 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Ostatnia macierz ma trzy wiodące jedynki i wektory $\mathbf{v}_1$, $\mathbf{v}_2$ i $\mathbf{v}_4$ (odpowiadające kolumnom zawierającym wiodące jedynki) generują przestrzeń R^3.

7.4. Liniowa zależność i liniowa niezależność wektorów

W poniższym tekście przez *układ wektorów* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ rozumiemy ciąg wektorów $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$.

Definicja 7.4.1. Układ wektorów $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ z przestrzeni wektorowej $V = V(K)$ nazywamy *liniowo niezależnym*, jeśli równanie wektorowe

Liniowo niezależny układ wektorów

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_k\mathbf{v}_k = \mathbf{0}$$

ma tylko zerowe rozwiązanie. Z drugiej strony, układ wektorów nazywamy *liniowo zależnym*, jeśli nie jest on liniowo niezależny. Zatem układ wektorów $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ jest liniowo zależny, jeśli istnieją skalary $\alpha_1, \ldots, \alpha_k \in K$, nie wszystkie równe zeru i takie, że

Liniowo zależny układ wektorów

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_k\mathbf{v}_k = \mathbf{0}. \tag{7.2}$$

Równość (7.2), w której nie wszystkie skalary $\alpha_1, \ldots, \alpha_k$ są zerowe, nazywamy *relacją liniowej zależności* pomiędzy wektorami $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

Tam gdzie nie będzie to prowadziło do nieporozumień, będziemy mówić o liniowej zależności (albo liniowej niezależności) wektorów $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ lub zbioru wektorów $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$, rozumiejąc przez to liniową zależność (albo liniową niezależność) układu $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$.

Przykład 139. Wektory

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ -2 \\ -2 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ -6 \\ -2 \end{bmatrix}$$

są liniowo zależne w R^3 i równość $2\mathbf{v}_1 + 3\mathbf{v}_2 + (-1)\mathbf{v}_3 = \mathbf{0}$ jest relacją liniowej zależności pomiędzy nimi.

Przykład 140. Zbadać liniową niezależność wektorów $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3 \in K_{2\times 2}$, gdy

$$\mathbf{v}_1 = \begin{bmatrix} 1 & 3 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 4 & 5 \\ 2 & -3 \end{bmatrix} \quad \text{i} \quad \mathbf{v}_3 = \begin{bmatrix} 10 & 9 \\ 6 & -7 \end{bmatrix}.$$

Dla zbadania liniowej niezależności wektorów $\mathbf{v}_1$, $\mathbf{v}_2$ i $\mathbf{v}_3$ wystarczy zbadać istnienie niezerowych rozwiązań równania $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{0}$. Równanie to, czyli równanie

$$\begin{aligned} &x_1\begin{bmatrix} 1 & 3 \\ 0 & -1 \end{bmatrix} + x_2\begin{bmatrix} 4 & 5 \\ 2 & -3 \end{bmatrix} + x_3\begin{bmatrix} 10 & 9 \\ 6 & -7 \end{bmatrix} \\ &\quad = \begin{bmatrix} x_1 + 4x_2 + 10x_3 & 3x_1 + 5x_2 + 9x_3 \\ 2x_2 + 6x_3 & -x_1 - 3x_2 - 7x_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \end{aligned}$$

jest równoważne jednorodnemu układowi równań liniowych

$$\begin{cases} x_1 + 4x_2 + 10x_3 = 0, \\ 3x_1 + 5x_2 + 9x_3 = 0, \\ 2x_2 + 6x_3 = 0, \\ -x_1 - 3x_2 - 7x_3 = 0. \end{cases}$$

Dla jego macierzy rozszerzonej mamy równoważności

$$\left[\begin{array}{ccc|c} 1 & 4 & 10 & 0 \\ 3 & 5 & 9 & 0 \\ 0 & 2 & 6 & 0 \\ -1 & -3 & -7 & 0 \end{array}\right] \sim \ldots \sim \left[\begin{array}{ccc|c} 1 & 0 & -2 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

i stąd widać, że równanie $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{0}$ ma niezerowe rozwiązanie, np. $2\mathbf{v}_1 - 3\mathbf{v}_2 + \mathbf{v}_3 = \mathbf{0}$. Zatem wektory $\mathbf{v}_1$, $\mathbf{v}_2$ i $\mathbf{v}_3$ są liniowo zależne.

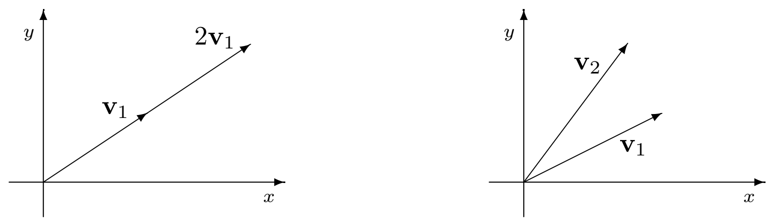
Przykład 141. Funkcje $\mathbf{v}_1 = \sin$ i $\mathbf{v}_2 = \cos$ są liniowo niezależne w przestrzeni $\mathcal{F}(\langle 0;\pi\rangle, R)$.

Załóżmy, że w przestrzeni $\mathcal{F}(\langle 0;\pi\rangle, R)$ zachodzi równość $x_1 \sin + x_2 \cos = \mathbf{0}$ dla pewnych skalarów $x_1, x_2 \in R$. Wtedy

$$x_1 \sin t + x_2 \cos t = \mathbf{0}(t) = 0$$

dla każdego $t \in \langle 0;\pi\rangle$. Z równości tej dla $t = \pi/2$ i $t = 0$ otrzymujemy $x_1 = 0$ i $x_2 = 0$. To dowodzi, że funkcje $\mathbf{v}_1 = \sin$ i $\mathbf{v}_2 = \cos$ są liniowo niezależne.

Z definicji 7.4.1 (i z twierdzenia 7.1.1) wynika, że układ $(\mathbf{v})$ składający się z jednego wektora jest liniowo zależny wtedy i tylko wtedy, gdy $\mathbf{v}$ jest wektorem zerowym. Istotnie, jeśli $\mathbf{v} \neq \mathbf{0}$ i $\alpha\mathbf{v} = \mathbf{0}$, to $\alpha = 0$ i układ $(\mathbf{v})$ jest liniowo niezależny. Z drugiej strony, układ $(\mathbf{0})$ jest liniowo zależny, bo $1{\cdot}\mathbf{0} = \mathbf{0}$ jest kombinacją spełniającą warunki definicji liniowej zależności. Łatwo także zauważyć, że układ dwóch wektorów $(\mathbf{v}_1, \mathbf{v}_2)$ jest liniowo zależne wtedy i tylko wtedy, gdy jeden z wektorów $\mathbf{v}_1$ i $\mathbf{v}_2$ jest równy iloczynowi drugiego wektora przez skalar, np. $\mathbf{v}_2 = \alpha\mathbf{v}_1$ (zob. rys. 7.9). Uogólnieniem tej obserwacji są następujące dwa twierdzenia i wynikające z nich wnioski o liniowej zależności (i niezależności) wektorów.



Wektory liniowo zależne   Wektory liniowo niezależne

Rys. 7.9

Twierdzenie 7.4.1. *Układ wektorów $B = (\mathbf{v}_1, \ldots, \mathbf{v}_k)$, $k \geqslant 2$, z przestrzeni wektorowej V jest liniowo zależny wtedy i tylko wtedy, gdy jeden z jego wektorów jest kombinacją liniową pozostałych wektorów z tego układu, tj. gdy $\mathbf{v}_i \in \mathcal{L}(B - \{\mathbf{v}_i\})$ dla pewnego $\mathbf{v}_i \in B$.*

Dowód. Jeśli wektor $\mathbf{v}_i$ jest kombinacją liniową wektorów $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k$, powiedzmy

$$\mathbf{v}_i = \alpha_1\mathbf{v}_1 + \ldots + \alpha_{i-1}\mathbf{v}_{i-1} + \alpha_{i+1}\mathbf{v}_{i+1} + \ldots + \alpha_k\mathbf{v}_k,$$

to

$-1 \neq 0$

$$\alpha_1\mathbf{v}_1 + \ldots + \alpha_{i-1}\mathbf{v}_{i-1} + (-1)\mathbf{v}_i + \alpha_{i+1}\mathbf{v}_{i+1} + \ldots + \alpha_k\mathbf{v}_k = \mathbf{0}$$

jest relacją liniowej zależności pomiędzy wektorami $\mathbf{v}_1, \ldots, \mathbf{v}_k$ i wektory te są liniowo zależne.

Załóżmy teraz, że układ wektorów $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ jest liniowo zależny. Wtedy istnieją skalary $\alpha_1, \ldots, \alpha_k$, nie wszystkie równe zeru i takie, że

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_k\mathbf{v}_k = \mathbf{0}. \tag{7.3}$$

Jeśli $\alpha_i \neq 0$, to z (7.3) otrzymujemy

$$\alpha_i\mathbf{v}_i = -\alpha_1\mathbf{v}_1 - \ldots - \alpha_{i-1}\mathbf{v}_{i-1} - \alpha_{i+1}\mathbf{v}_{i+1} - \ldots - \alpha_k\mathbf{v}_k$$

i stąd

$$\mathbf{v}_i = (-\alpha_1\alpha_i^{-1})\mathbf{v}_1 + \ldots + (-\alpha_{i-1}\alpha_i^{-1})\mathbf{v}_{i-1} + (-\alpha_{i+1}\alpha_i^{-1})\mathbf{v}_{i+1} + \ldots + (-\alpha_k\alpha_i^{-1})\mathbf{v}_k.$$

To oznacza, że wektor $\mathbf{v}_i$ jest kombinacją liniową wektorów $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k$. □

Wniosek 7.4.1. *Układ wektorów $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$, $k \geqslant 2$, z przestrzeni wektorowej V jest liniowo niezależny wtedy i tylko wtedy, gdy żaden z wektorów $\mathbf{v}_j$ nie jest kombinacją liniową pozostałych wektorów, tj. gdy $\mathbf{v}_j \notin \mathcal{L}(B - \{\mathbf{v}_j\})$ dla każdego $\mathbf{v}_j \in B$.* $\Box$

Przykład 142. Wektory $\mathbf{v}_1$, $\mathbf{v}_2$ i $\mathbf{v}_3$ z przykładu 139 są liniowo zależne i $\mathbf{v}_3$ jest kombinacją liniową wektorów $\mathbf{v}_1$ i $\mathbf{v}_2$, $\mathbf{v}_3 = 2\mathbf{v}_1 + 3\mathbf{v}_2$.

Twierdzenie 7.4.2. *Układ $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ niezerowych wektorów przestrzeni wektorowej V jest liniowo zależny wtedy i tylko wtedy, gdy pewien wektor $\mathbf{v}_j$ $(2 \leqslant j \leqslant k)$ jest kombinacją liniową swoich poprzedników, tj. gdy $\mathbf{v}_j \in \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$ dla pewnego j, $2 \leqslant j \leqslant k$.*

Dowód. Załóżmy, że układ $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ jest liniowo zależny. Wtedy, ponieważ układ $(\mathbf{v}_1)$ jest liniowo niezależny (bo $\mathbf{v}_1 \neq \mathbf{0}$), istnieje j $(2 \leqslant j \leqslant k)$ takie, że układ $(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$ jest liniowo niezależny, a układ $(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}, \mathbf{v}_j)$ jest już liniowo zależny. Zatem istnieją skalary $\alpha_1, \ldots, \alpha_j$ nie wszystkie równe zeru i takie, że

$$\alpha_1\mathbf{v}_1 + \ldots + \alpha_{j-1}\mathbf{v}_{j-1} + \alpha_j\mathbf{v}_j = \mathbf{0}.$$

Zauważmy, że $\alpha_j \neq 0$ (bo inaczej byłoby $\alpha_1\mathbf{v}_1 + \ldots + \alpha_{j-1}\mathbf{v}_{j-1} = \mathbf{0}$ i nie wszystkie współczynniki tej kombinacji byłyby zerowe i układ $(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$ byłby liniowo zależny), więc

$$\mathbf{v}_j = -\alpha_j^{-1}(\alpha_1\mathbf{v}_1 + \ldots + \alpha_{j-1}\mathbf{v}_{j-1}) \in \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}).$$

Załóżmy teraz, że $\mathbf{v}_j \in \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$ dla pewnego j, $2 \leqslant j \leqslant k$. Wtedy także $\mathbf{v}_j \in \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_k)$ i wobec twierdzenia 7.4.1 wektory $\mathbf{v}_1, \ldots, \mathbf{v}_k$ są liniowo zależne. $\Box$

Wniosek 7.4.2. *Układ $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ niezerowych wektorów przestrzeni V jest liniowo niezależny wtedy i tylko wtedy, gdy żaden wektor $\mathbf{v}_j$ $(2 \leqslant j \leqslant k)$ nie jest kombinacją liniową swoich poprzedników, tj. gdy $\mathbf{v}_j \notin \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$ dla $j = 2, \ldots, k$.* $\Box$

Przestrzeń wektorowa K^n jest najważniejsza w naszych rozważaniach, więc z osobna zajmiemy się liniową niezależnością wektorów w tej przestrzeni. Dla wektorów $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ z przestrzeni K^n równanie wektorowe $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_k\mathbf{v}_k = \mathbf{0}$ jest równoważne równaniu macierzowemu $\mathbf{Ax} = \mathbf{0}$, w którym $\mathbf{A} = [\mathbf{v}_1\, \mathbf{v}_2 \ldots \mathbf{v}_k]$.

Ponieważ istnieje odpowiedniość wzajemnie jednoznaczna pomiędzy relacjami liniowej zależności wektorów $\mathbf{v}_1$, $\mathbf{v}_2, \ldots, \mathbf{v}_k$ i niezerowymi rozwiązaniami równania $\mathbf{Ax} = \mathbf{0}$, więc mamy następujące twierdzenie o związku pomiędzy liniową niezależnością wektorów przestrzeni K^n i istnieniem niezerowych rozwiązań jednorodnego układu równań liniowych.

Twierdzenie 7.4.3. *Kolumny macierzy $\mathbf{A} = [\mathbf{v}_1\, \mathbf{v}_2 \ldots \mathbf{v}_k] \in K_{n\times k}$ są wektorami liniowo niezależnymi wtedy i tylko wtedy, gdy wektor zerowy jest jedynym rozwiązaniem układu równań liniowych $\mathbf{Ax} = \mathbf{0}$.* $\Box$

Przykład 143. Zbadać liniową niezależność wektorów $\mathbf{v}_1 = (1, 2, 3, -1)$, $\mathbf{v}_2 = (2, 1, 2, 1)$ i $\mathbf{v}_3 = (1, -4, -5, 5)$ w przestrzeni R^4.

Ponieważ wektory $\mathbf{v}_1$, $\mathbf{v}_2$ i $\mathbf{v}_3$ są liniowo zależne wtedy i tylko wtedy, gdy równanie $\mathbf{Ax} = \mathbf{0}$ (gdzie $\mathbf{A} = [\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3]$) ma niezerowe rozwiązanie, więc zbadamy istnienie niezerowych rozwiązań równania $\mathbf{Ax} = \mathbf{0}$. Dla macierzy rozszerzonej tego równania istnieją równoważności

$$[\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3 \,|\, \mathbf{0}] = \left[\begin{array}{rrr|r} 1 & 2 & 1 & 0 \\ 2 & 1 & -4 & 0 \\ 3 & 2 & -5 & 0 \\ -1 & 1 & 5 & 0 \end{array}\right] \sim \ldots \sim \left[\begin{array}{rrr|r} 1 & 0 & -3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

i z ostatniej macierzy widać, że równanie $x_1\mathbf{v}_1+x_2\mathbf{v}_2+x_3\mathbf{v}_3 = \mathbf{0}$ ma niezerowe rozwiązanie. Dla przykładu, jeśli $x_3 = 1$, to $x_1 = 3$, $x_2 = -2$ i dlatego mamy $3\mathbf{v}_1-2\mathbf{v}_2+\mathbf{v}_3 = \mathbf{0}$. Zatem wektory $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ są liniowo zależne.

Jeśli wektory $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ należą do przestrzeni K^n i $k > n$, to w jednorodnym układzie równań $[\mathbf{v}_1\, \mathbf{v}_2 \ldots \mathbf{v}_k]\mathbf{x} = \mathbf{0}$ liczba niewiadomych jest większa od liczby równań, więc układ ten ma niezerowe rozwiązanie (zob. wniosek 5.5.1) i z twierdzenia 7.4.3 mamy następujący wniosek.

Wniosek 7.4.3. *Wektory* $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ *z przestrzeni* K^n *są liniowo zależne, gdy* $k > n$. □

Przykład 144. Każde cztery wektory w przestrzeni R^3 są liniowo zależne. Zatem wektory $\mathbf{v}_1 = (2, 1, 1)$, $\mathbf{v}_2 = (1, 2, 1)$, $\mathbf{v}_3 = (1, 1, 2)$ i $\mathbf{v}_4 = (3, 4, 5)$ są liniowo zależne w R^3.

Z wniosków 6.3.1 i 6.5.1 oraz z twierdzeń 7.3.2 i 7.4.3 otrzymujemy następujący wniosek o liniowej niezależności n wektorów z przestrzeni K^n.

Wniosek 7.4.4. *Dla wektorów* $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *z przestrzeni* K^n *następujące stwierdzenia są równoważne*:

(1) *Układ wektorów* $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ *jest liniowo niezależny;*

(2) *Równanie* $[\mathbf{v}_1 \ldots \mathbf{v}_n]\mathbf{x} = \mathbf{0}$ *ma tylko zerowe rozwiązanie;* (tw. 7.4.3)

(3) *Macierz* $[\mathbf{v}_1 \ldots \mathbf{v}_n]$ *jest nieosobliwa;* (wn. 6.3.1)

(4) *Macierz* $[\mathbf{v}_1 \ldots \mathbf{v}_n]$ *jest odwracalna;* (wn. 6.5.1)

(5) *Wektory* $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *generują całą przestrzeń* K^n. □ (tw. 7.3.2)

Przykład 145. Zbadać liniową zależność wektorów $\mathbf{v}_1 = (8, 6, 5, 4)$, $\mathbf{v}_2 = (10, 8, 5, 4)$, $\mathbf{v}_3 = (12, 10, 8, 4)$ i $\mathbf{v}_4 = (7, 6, 5, 4)$ w przestrzeni R^4.

Ponieważ macierz $\left[\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3\, \mathbf{v}_4\right]$ jest nieosobliwa,

$$\det \begin{bmatrix} | & | & | & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 \\ | & | & | & | \end{bmatrix} = \begin{vmatrix} 8 & 10 & 12 & 7 \\ 6 & 8 & 10 & 6 \\ 5 & 5 & 8 & 5 \\ 4 & 4 & 4 & 4 \end{vmatrix} \overset{\substack{\mathbf{v}_1 - \mathbf{v}_4 \\ \mathbf{v}_2 - \mathbf{v}_4 \\ \mathbf{v}_3 - \mathbf{v}_4}}{=} \begin{vmatrix} 1 & 3 & 5 & 7 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & 4 \end{vmatrix} = 24,$$

więc wobec wniosku 7.4.4 wektory $\mathbf{v}_1, \ldots, \mathbf{v}_4$ są liniowo niezależne (i generują całą przestrzeń R^4).

Na zakończenie zauważmy, że definicję liniowej niezależności (i zależności) można tak rozszerzyć, aby obejmowała ona także zbiory nieskończone.

Definicja 7.4.2. Mówimy, że zbiór S wektorów przestrzeni V jest liniowo niezależny, jeżeli każdy skończony podzbiór zbioru S jest liniowo niezależny. W przeciwnym przypadku mówimy, że zbiór S jest liniowo zależny.

Przykład 146. Zbiór $S = \{1, x, x^2, \ldots, x^n, \ldots\}$ jednomianów jest liniowo niezależny w przestrzeni wielomianów rzeczywistych $R[x]$.

Tak jest, bo każda liniowa kombinacja skończonej ilości różnych wektorów ze zbioru S jest wielomianem, $\varphi(x) = \alpha_1 x^{k_1} + \ldots + \alpha_m x^{k_m}$. Taki wielomian jest wielomianem zerowym, gdy wszystkie jego współczynniki są zerowe.

7.5. Baza przestrzeni wektorowej

Definicja 7.5.1. Układ $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ wektorów z przestrzeni V nazywamy *bazą przestrzeni* V, gdy ma on następujące dwie własności:

Baza przestrzeni

(1) B jest liniowo niezależny;

(2) B generuje przestrzeń V, tj. $\mathcal{L}(B) = V$.

Zacznijmy od kilku przykładów baz przestrzeni wektorowych.

Przykład 147. Sprawdzić, czy układ wektorów $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ jest bazą przestrzeni R^3, gdy $\mathbf{v}_1 = (2, 1, 3)$, $\mathbf{v}_2 = (1, 0, 1)$ i $\mathbf{v}_3 = (1, -1, 1)$.

Wobec wniosku 7.4.4 wystarczy sprawdzić nieosobliwość macierzy $[\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3]$. Ponieważ

$$\det[\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3] = \det \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & -1 \\ 3 & 1 & 1 \end{bmatrix} = -1 \neq 0,$$

więc macierz ta jest nieosobliwa i układ wektorów $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ jest bazą przestrzeni R^3.

Przykład 148. Podobnie jak w poprzednim przykładzie można uzasadnić, że układ wektorów $B = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$, w którym

Baza standardowa

$$\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$$

(jedynka na i-tym miejscu) dla $i = 1, \ldots, n$, jest bazą przestrzeni K^n. Nazywamy ją *bazą standardową* (lub *kanoniczną*) przestrzeni K^n.

Przykład 149. Wyznaczyć bazę podprzestrzeni W przestrzeni R^4, gdy

$$W = \{(x_1, x_2, x_3, x_4) \in R^4 \colon\ x_1 = 3x_2,\ x_3 = x_2 + 4x_4\}.$$

Mamy

$$\begin{aligned} W &= \{(3x_2, x_2, x_2 + 4x_4, x_4) \colon\ x_2,\ x_4 \in R\} \\ &= \{(3x_2, x_2, x_2, 0) + (0, 0, 4x_4, x_4) \colon\ x_2, x_4 \in R\} \\ &= \{x_2(3, 1, 1, 0) + x_4(0, 0, 4, 1) \colon\ x_2, x_4 \in R\}, \end{aligned}$$

co oznacza, że wektory $\mathbf{v}_1 = (3, 1, 1, 0)$ i $\mathbf{v}_2 = (0, 0, 4, 1)$ generują podprzestrzeń W. Ponieważ wektory $\mathbf{v}_1$ i $\mathbf{v}_2$ są liniowo niezależne (bo żaden z nich nie jest krotnością drugiego), więc uporządkowany zbiór $(\mathbf{v}_1, \mathbf{v}_2)$ (i zbiór $(\mathbf{v}_2, \mathbf{v}_1)$) jest bazą podprzestrzeni W.

Przykład 150. Wyznaczyć bazę przestrzeni zerowej $N_{\mathbf{A}}$ macierzy

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

Przestrzeń zerowa $N_{\mathbf{A}}$ macierzy $\mathbf{A}$ jest zbiorem wszystkich rozwiązań jednorodnego układu równań $\mathbf{Ax} = \mathbf{0}$. Ponieważ dla macierzy rozszerzonej tego układu mamy

$$[\mathbf{A}|\mathbf{0}] = \left[\begin{array}{ccc|c} 1 & 3 & 2 & 0 \\ 2 & 6 & 4 & 0 \\ 2 & 1 & -1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 3 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -5 & -5 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

więc

$$N_{\mathbf{A}} = \left\{ \begin{bmatrix} x_3 \\ -x_3 \\ x_3 \end{bmatrix} : x_3 \in R \right\} = \left\{ x_3 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} : x_3 \in R \right\} = \mathcal{L}\left(\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right)$$

i bazę przestrzeni $N_{\mathbf{A}}$ tworzy pojedynczy wektor $(1, -1, 1)$.

Przykład 151. Z naszych wcześniejszych rozważań wynika, że uporządkowany nieskończony zbiór jednomianów $(1, x, x^2, \ldots, x^n, \ldots)$ jest bazą przestrzeni $R[x]$.

Korzystając z tzw. lematu Kuratowskiego-Zorna, można udowodnić, że każdy uporządkowany zbiór niezależnych wektorów przestrzeni V jest zawarty w pewnej bazie przestrzeni V (zob. tw. 6.2 w [2]). Stąd w szczególności wynika, że każda przestrzeń wektorowa ma bazę. Twierdzenia te mają charakter egzystencjalny i w ogólnym przypadku nie dostarczają one efektywnych metod wyznaczania bazy przestrzeni wektorowej. W przestrzeniach skończenie generowanych, a te są głównym obiektem naszych zainteresowań, istnieją – jak to wynika z następnych twierdzeń – efektywne metody wyznaczania baz.

Następujące twierdzenie podaje podstawowe własności bazy skończenie generowanej przestrzeni wektorowej.

Twierdzenie 7.5.1. *Niech $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ będzie układem wektorów z przestrzeni V. Wówczas następujące warunki są równoważne:*

(a) *B jest bazą przestrzeni V;*

(b) *Każdy wektor $\mathbf{v} \in V$ można, i to tylko na jeden sposób, przedstawić jako kombinację liniową wektorów układu B;*

(c) *B jest minimalnym układem generującym przestrzeń V;*

(d) *B jest maksymalnym liniowo niezależnym układem w przestrzeni V.*

Dowód. $(a) \Rightarrow (b)$. Załóżmy, że B jest bazą przestrzeni V. Wtedy $V = \mathcal{L}(B)$, więc każdy wektor $\mathbf{v} \in V$ jest kombinacją liniową wektorów układu B, czyli

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n$$

dla pewnych skalarów $\alpha_1, \ldots, \alpha_n \in K$. Twierdzimy, że współczynniki tej kombinacji są określone jednoznacznie (przez wektory $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ i $\mathbf{v}$). Przypuśćmy bowiem, że także mamy

$$\mathbf{v} = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \ldots + \beta_n\mathbf{v}_n$$

dla pewnych skalarów $\beta_1, \ldots, \beta_n \in K$. Wtedy

$$\mathbf{0} = \mathbf{v} - \mathbf{v} = (\beta_1 - \alpha_1)\mathbf{v}_1 + (\beta_2 - \alpha_2)\mathbf{v}_2 + \ldots + (\beta_n - \alpha_n)\mathbf{v}_n$$

i z liniowej niezależności wektorów $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ wynika, że wszystkie współczynniki tej kombinacji są zerami. Stąd $\beta_1 = \alpha_1, \ldots, \beta_n = \alpha_n$ i to dowodzi, że współczynniki kombinacji $\mathbf{v} = \alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n$ są określone jednoznacznie.

$(b) \Rightarrow (c)$. Z (b) wynika, że B jest układem generującym przestrzeń V, tj. $V = \mathcal{L}(B)$. Dla dowodu, że B jest minimalnym układem generującym przestrzeń V wystarczy zauważyć, że gdyby układ $B - \{\mathbf{v}_i\}$ (powstały z B przez odrzucenie zeń wektora $\mathbf{v}_i$) był układem generującym przestrzeń V, to mielibyśmy $\mathcal{L}(B - \{\mathbf{v}_i\}) = V$ i dla wektora $\mathbf{v}_i$, który należy do $V = \mathcal{L}(B - \{\mathbf{v}_i\})$, istniałyby skalary $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n$ takie, że

$$\mathbf{v}_i = \alpha_1\mathbf{v}_1 + \ldots + \alpha_{i-1}\mathbf{v}_{i-1} + \alpha_{i+1}\mathbf{v}_{i+1} \ldots + \alpha_n\mathbf{v}_n,$$

a ponieważ jednocześnie jest

$$\mathbf{v}_i = 0\mathbf{v}_1 + \ldots + 0\mathbf{v}_{i-1} + 1\mathbf{v}_i + 0\mathbf{v}_{i+1} + \ldots + 0\mathbf{v}_n,$$

więc istniałyby dwie różne możliwości przedstawienia wektora $\mathbf{v}_i$ w postaci kombinacji liniowej wektorów $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ i to przeczyłoby założeniu (b).

$(c) \Rightarrow (d)$. Załóżmy teraz, że B jest minimalnym układem generującym przestrzeń V. Wtedy $\mathcal{L}(B) = V$ i $V - \mathcal{L}(B - \{\mathbf{v}_i\}) \neq \emptyset$ dla każdego wektora $\mathbf{v}_i$ z układu B. Stąd wynika, że $\mathbf{v}_i \notin \mathcal{L}(B - \{\mathbf{v}_i\})$ dla każdego $\mathbf{v}_i \in B$. Wobec wniosku 7.4.1 dowodzi to, że układ B jest liniowo niezależny. Jednocześnie dla każdego $\mathbf{u} \in V - B$ jest $\mathbf{u} \in \mathcal{L}(B)$, więc układ $B \cup \{\mathbf{u}\}$ (czyli uklad powstały z B przez dołączenie doń wektora $\mathbf{u}$) jest liniowo zależny. Zatem B jest maksymalnym liniowo niezależnym układem.

$(d) \Rightarrow (a)$. Załóżmy, że B jest maksymalnym układem liniowo niezależnym i niech $\mathbf{u}$ będzie dowolnym wektorem ze zbioru $V - B$. Wtedy $B \cup \{\mathbf{u}\}$ jest liniowo zależny i z twierdzenia 7.4.1 wynika, że $\mathbf{u} \in \mathcal{L}(B)$. To dowodzi, że układ B generuje przestrzeń V. Zatem B jest bazą przestrzeni V. □

Wniosek 7.5.1. *Każda skończenie generowana przestrzeń wektorowa ma bazę.*

Dowód. Teza jest oczywista dla przestrzeni zerowej, jej bazą jest zbiór pusty. Zatem niech V będzie niezerową przestrzenią wektorową generowaną przez zbiór $G = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subset V$ i niech B będzie maksymalnym liniowo niezależnym układem utworzonym z elementów zbioru G (lub minimalnym układem generującym przestrzeń V i utworzonym z elementów zbioru G). Zgodnie z twierdzeniem 7.5.1 układ B jest bazą przestrzeni V. □

Niech $G = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ będzie zbiorem generującym przestrzeń wektorową V. Ponieważ $V = \mathcal{L}(G) = \mathcal{L}(G - \{\mathbf{0}\})$, możemy założyć, że wektory $\mathbf{v}_1, \ldots, \mathbf{v}_n$ są niezerowe. Bazę B przestrzeni V, w praktyce (uporządkowany) maksymalny liniowo niezależny podzbiór zbioru G, o którym mowa w dowodzie poprzedniego wniosku, możemy efektywnie wyznaczyć za pomocą *procedury odrzucania.* W tym celu bierzemy pod uwagę układ $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$. Zaczynając od $\mathbf{v}_2$, z układu tego odrzucamy każdy wektor, który jest kombinacją liniową jego poprzedników pozostałych po wcześniejszych odrzuceniach. Po przetestowaniu wszystkich wektorów $\mathbf{v}_i$ $(2 \leqslant i \leqslant n)$, wektory, które nie zostały odrzucone są liniowo niezależne (zob. wniosek 7.4.2) i – jak łatwo zauważyć – generują one przestrzeń V, więc tworzą bazę przestrzeni V.

Procedura odrzucania

Przykład 152. Za pomocą procedury odrzucania wyznaczyć bazę przestrzeni $V = \mathcal{L}(x^2,\, x^2+1,\, 2,\, 3x-2,\, x)$ w $R[x]$.

Weźmy pod uwagę ciąg $(\mathbf{v}_1, \ldots, \mathbf{v}_5) = (x^2,\, x^2+1,\, 2,\, 3x-2,\, x)$. Ponieważ wektory $\mathbf{v}_1 = x^2$ i $\mathbf{v}_2 = x^2+1$ są liniowo niezależne, więc zgodnie z procedurą odrzucania zachowujemy wektor $\mathbf{v}_2$ i badamy zależność wektora $\mathbf{v}_3 = 2$ od jego poprzedników. Wektor $\mathbf{v}_3$ jest liniową kombinacją swoich poprzedników, $\mathbf{v}_3 = 2 = -2x^2 + 2(x^2+1)$ $= -2\mathbf{v}_1 + 2\mathbf{v}_2$, więc odrzucamy $\mathbf{v}_3$ i badamy zależność wektora $\mathbf{v}_4$ od jego poprzedników w ciągu $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4, \mathbf{v}_5)$. Wektor $\mathbf{v}_4 = 3x - 2$ nie jest kombinacją wektorów $\mathbf{v}_1$ i $\mathbf{v}_2$, bo x w pierwszej potędze nie występuje ani w $\mathbf{v}_1$, ani w $\mathbf{v}_2$. Zatem $\mathbf{v}_4$ pozostaje w ciągu i analizujemy zależność wektora $\mathbf{v}_5$ od jego poprzedników w ciągu $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4, \mathbf{v}_5)$.

W tym celu sprawdzamy czy istnieją skalary a, b i c takie, że $\mathbf{v}_5 = a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_4$. Łatwo zauważyć, że $\mathbf{v}_5 = x = -\frac{2}{3}x^2 + \frac{2}{3}(x^2+1) + \frac{1}{3}(3x-2) = -\frac{2}{3}\mathbf{v}_1 + \frac{2}{3}\mathbf{v}_2 + \frac{1}{3}\mathbf{v}_4$, więc z ciągu $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4, \mathbf{v}_5)$ odrzucamy wektor $\mathbf{v}_5$. Ostatecznie wektory $\mathbf{v}_1$, $\mathbf{v}_2$ i $\mathbf{v}_4$ tworzą bazę przestrzeni V.

W przypadku wektorów przestrzeni K^n procedurę odrzucania wektorów, które są liniowymi kombinacjami swoich poprzedników w ciągu $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ można wykonać w następujący sposób: Tworzymy macierz $\mathbf{A} = [\mathbf{v}_1\, \mathbf{v}_2 \ldots \mathbf{v}_k]$ i sprowadzamy ją do wierszowo równoważnej macierzy $\mathbf{H}$ mającej normalną postać schodkową. Z ciągu $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ odrzucamy wszystkie te wektory, które odpowiadają kolumnom bez wiodących jedynek w macierzy $\mathbf{H}$. Pozostałe wektory tworzą bazę podprzestrzeni generowanej przez wektory $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, więc także bazę przestrzeni kolumnowej macierzy $\mathbf{A}$.

Tak jest istotnie, bo wobec definicji 7.4.1 każdej relacji liniowej zależności między wektorami $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ odpowiada niezerowe rozwiązanie jednorodnego układu $\mathbf{Ax} = \mathbf{0}$. A ponieważ zbiór rozwiązań układu $\mathbf{Ax} = \mathbf{0}$ jest taki sam jak zbiór rozwiązań układu $\mathbf{Hx} = \mathbf{0}$, więc każda relacja liniowej zależności między kolumnami macierzy $\mathbf{A}$ jest relacją liniowej zależności między kolumnami macierzy $\mathbf{H}$ (i odwrotnie). Kolumny macierzy $\mathbf{H}$ zawierające wiodące jedynki są różnymi wektorami standardowej bazy przestrzeni K^n, więc są one niezależne. Natomiast każda kolumna bez wiodącej jedynki macierzy $\mathbf{H}$ jest liniową kombinacją poprzedzających ją kolumn z wiodącymi jedynkami. Zatem kolumnami macierzy $\mathbf{A}$, które są kombinacjami swoich poprzedników (i które odrzucamy) są dokładnie te kolumny, które odpowiadają kolumnom bez wiodących jedynek w macierzy $\mathbf{H}$.

Nasze rozważania ilustruje następujący przykład.

Przykład 153. Za pomocą procedury odrzucania wyznaczyć bazę podprzestrzeni V przestrzeni R^5 generowanej przez wektory $\mathbf{v}_1 = (1, 2, 3, 1, 1)$, $\mathbf{v}_2 = (-2, -3, -5, -1, -2)$, $\mathbf{v}_3 = (0, 1, 1, 1, 0)$, $\mathbf{v}_4 = (1, -1, 0, -2, 1)$ i $\mathbf{v}_5 = (2, 6, 9, 4, 3)$.

Tworzymy macierz $\mathbf{A} = [\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3\, \mathbf{v}_4\, \mathbf{v}_5]$ i sprowadzamy ją do wierszowo równoważnej normalnej macierzy schodkowej:

$$\begin{bmatrix} 1 & -2 & 0 & 1 & 2 \\ 2 & -3 & 1 & -1 & 6 \\ 3 & -5 & 1 & 0 & 9 \\ 1 & -1 & 1 & -2 & 4 \\ 1 & -2 & 0 & 1 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 0 & 1 & 2 \\ 0 & 1 & 1 & -3 & 2 \\ 0 & 1 & 1 & -3 & 3 \\ 0 & 1 & 1 & -3 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 & -5 & 0 \\ 0 & 1 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Ponieważ wiodące jedynki są tylko w pierwszej, drugiej i piątej kolumnie normalnej macierzy schodkowej, więc wektory $\mathbf{v}_1$, $\mathbf{v}_2$ i $\mathbf{v}_5$ są liniowo niezależne. Wektory $\mathbf{v}_3$ i $\mathbf{v}_4$ są kombinacjami wektorów $\mathbf{v}_1$ i $\mathbf{v}_2$ ($\mathbf{v}_3 = 2\mathbf{v}_1 + \mathbf{v}_2$, $\mathbf{v}_4 = -5\mathbf{v}_1 - 3\mathbf{v}_2$) i odrzucamy je ze zbioru generatorów przestrzeni V. Zatem układ $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_5)$ jest bazą przestrzeni V (i bazą przestrzeni kolumnowej macirzy $\mathbf{A}$).

Uwaga. Przy wyznaczaniu bazy przestrzeni $\mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_k) \subseteq K^n$ wystarczy sprowadzić macierz $\mathbf{A} = [\mathbf{v}_1\, \mathbf{v}_2 \ldots \mathbf{v}_k]$ do wierszowo równoważnej macierzy schodkowej, nie musi to być normalna macierz schodkowa.

Niech V będzie skończenie generowaną przestrzenią wektorową. Chociaż przestrzeń ta może mieć wiele różnych baz, pokażemy, że każde dwie bazy przestrzeni V mają taką samą liczbę elementów. Fakt ten jest konsekwencją następującego twierdzenia Steinitza.

Twierdzenie 7.5.2 (Steinitz). *Jeśli wektory* $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *generują przestrzeń* V*, a wektory* $\mathbf{u}_1, \ldots, \mathbf{u}_m$ *są liniowo niezależne w przestrzeni* V*, to* $m \leqslant n$*.*

Dowód. Przypuśćmy dla sprzeczności, że jest $m > n$. Ponieważ wektory $\mathbf{v}_1, \ldots, \mathbf{v}_n$ generują przestrzeń V, więc istnieją skalary a_{ij} takie, że

$$\mathbf{u}_i = a_{i1}\mathbf{v}_1 + a_{i2}\mathbf{v}_2 + \ldots + a_{in}\mathbf{v}_n = \sum_{j=1}^{n} a_{ij}\mathbf{v}_j \tag{7.4}$$

dla $i = 1, \ldots, m$. Weźmy teraz pod uwagę jednorodny układ równań

$$\begin{cases} a_{11}x_1 + a_{21}x_2 + \ldots + a_{m1}x_m = 0, \\ a_{12}x_1 + a_{22}x_2 + \ldots + a_{m2}x_m = 0, \\ \qquad\vdots \\ a_{1n}x_1 + a_{2n}x_2 + \ldots + a_{mn}x_m = 0. \end{cases} \tag{7.5}$$

W układzie (7.5) jest więcej niewiadomych niż równań ($m > n$), więc wobec wniosku 5.5.1 układ ten ma niezerowe rozwiązanie $\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_m$. Dla tego niezerowego rozwiązania $\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_m$ z (7.4) i (7.5) kolejno otrzymujemy

$$\sum_{i=1}^{m} \overline{x}_i\mathbf{u}_i = \sum_{i=1}^{m} \overline{x}_i\Big(\sum_{j=1}^{n} a_{ij}\mathbf{v}_j\Big) = \sum_{j=1}^{n}\Big(\sum_{i=1}^{m} a_{ij}\overline{x}_i\Big)\mathbf{v}_j = \sum_{j=1}^{n} 0\mathbf{v}_j = \mathbf{0},$$

co oznacza, że wektory $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ są liniowo zależne i co jest sprzeczne z założeniem. To dowodzi, że $m \leqslant n$. □

Wniosek 7.5.2. *Każde dwie bazy skończenie generowanej przestrzeni wektorowej mają tyle samo elementów.*

Dowód. Niech $B = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m)$ i $B' = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ będą bazami przestrzeni V. Ponieważ wektory $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ są liniowo niezależne w przestrzeni generowanej przez wektory $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, to wobec twierdzenia 7.5.2 jest $m \leqslant n$. Jednocześnie wektory $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ są liniowo niezależne w przestrzeni generowanej przez wektory $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$, więc także $n \leqslant m$. Stąd wynika, że $m = n$. □

Wobec wniosku 7.5.2 liczba wektorów w dowolnej bazie skończenie generowanej przestrzeni wektorowej jest niezmiennikiem tej przestrzeni, więc możemy przyjąć następującą definicję.

Definicja 7.5.2. Liczbę elementów dowolnej bazy skończenie generowanej przestrzeni wektorowej V nazywamy *wymiarem przestrzeni* V i oznaczamy symbolem $\dim V$. Jeśli $\dim V = n$, to mówimy, że V jest *przestrzenią wymiaru* n (lub *przestrzenią* n*-wymiarową*). W tym przypadku mówimy także, że V jest *przestrzenią skończenie wymiarową*. Jeśli przestrzeń V nie jest skończenie wymiarowa, to mówimy, że jest ona *przestrzenią nieskończenie wymiarową* (lub że jej *wymiar jest nieskończony*) i piszemy $\dim V = \infty$.

$\dim V$ – wymiar przestrzeni V

Przestrzeń K^n jest skończenie wymiarowa i wobec przykładu 121 mamy $\dim K^n = n$. Wielomiany $1, x, x^2, \ldots, x^n$ tworzą bazę przestrzeni $R_n[x]$ wielomianów stopnia co najwyżej n. Zatem $\dim R_n[x] = n + 1$. Natomiast przestrzeń wszystkich wielomianów $R[x]$ (i przestrzeń funkcji $\mathcal{F}(R, R)$) jest przestrzenią nieskończenie wymiarową. Bazą przestrzeni zerowej $V = \{\mathbf{0}\}$ jest zbiór pusty i dlatego mówimy, że jest to przestrzeń 0-wymiarowa, $\dim\{\mathbf{0}\} = 0$.

$\dim K^n = n$

$\dim R_n[x] = n + 1$

$\dim\{\mathbf{0}\} = 0$

Twierdzenie 7.5.3. *Każdy uporządkowany zbiór* $I_k = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ *liniowo niezależnych wektorów skończenie wymiarowej przestrzeni* V *można uzupełnić do bazy przestrzeni* V*.*

Dowód. Niech $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$ będzie jakąkolwiek bazą przestrzeni V. Ponieważ wektory $\mathbf{v}_1, \ldots, \mathbf{v}_k$, $\mathbf{u}_1, \ldots, \mathbf{u}_n$ generują przestrzeń V, więc po odrzuceniu z ciągu $(\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{u}_1, \ldots, \mathbf{u}_n)$ wektorów będących kombinacjami liniowymi swoich poprzedników otrzymamy bazę B przestrzeni V. Baza ta zawiera wektory $\mathbf{v}_1, \ldots, \mathbf{v}_k$, bo żaden z nich nie jest kombinacją swoich poprzedników w ciągu $(\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{u}_1, \ldots, \mathbf{u}_n)$. $\square$

Inne uzasadnienie. Jeśli $\mathcal{L}(I_k) = V$, to I_k jest żądaną bazą przestrzeni V. Jeśli $\mathcal{L}(I_k) \neq V$, to wybieramy dowolny wektor $\mathbf{v}_{k+1}$ ze zbioru $V - \mathcal{L}(I_k)$ i tworzymy uporządkowany zbiór $I_{k+1} = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1})$. Wobec wniosku 7.4.2 zbiór I_{k+1} jest liniowo niezależny. Jeśli tym razem $\mathcal{L}(I_{k+1}) = V$, to I_{k+1} jest stosowną bazą przestrzeni V. Jeśli $\mathcal{L}(I_{k+1}) \neq V$, to wybieramy dowolny wektor $\mathbf{v}_{k+2}$ ze zbioru $V - \mathcal{L}(I_{k+1})$ i tworzymy uporządkowany zbiór $I_{k+2} = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \mathbf{v}_{k+2})$. Zbiór I_{k+2} jest liniowo niezależny. Zatem, jeśli $\mathcal{L}(I_{k+2}) = V$, to I_{k+2} jest bazą przestrzeni V. W przypadku przeciwnym zbiór $V - \mathcal{L}(I_{k+2})$ jest niepusty, więc wybieramy z niego dowolny wektor $\mathbf{v}_{k+3}$, itd. Proces ten kończy się po pewnej liczbie kroków, bo przestrzeń V jest skończenie wymiarowa i nie może zawierać nieskończonego zbioru niezależnych wektorów. Dlatego po skończonej liczbie kroków otrzymamy zbiór $I_n = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n)$ liniowo niezależnych wektorów generujących przestrzeń V, czyli bazę przestrzeni V. $\square$

Niech V będzie przestrzenią skończenie wymiarową. Na to aby układ wektorów $B \subset V$ był bazą przestrzeni V potrzeba i wystarcza, aby miał on dwie własności: (a) B generuje przestrzeń V i (b) B jest liniowo niezależny. Okazuje się, że jeśli znamy wymiar przestrzeni V i liczba elementów zbioru B jest równa wymiarowi przestrzeni V, to własności (a) i (b) są sobie równoważne i mamy następujący odpowiednik wniosku 7.4.4.

Twierdzenie 7.5.4. *Jeśli $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ jest układem n wektorów n-wymiarowej przestrzeni V, to następujące warunki są równoważne:*

(1) *B jest bazą przestrzeni V;*

(2) *B jest liniowo niezależny;*

(3) *B generuje przestrzeń V.*

Dowód. Jeśli zbiór B jest niezależny, to wobec twierdzenia 7.5.3 można go uzupełnić do bazy B' przestrzeni V. Ponieważ $B \subseteq B'$ i $n = |B| \leqslant |B'| = \dim V = n$, więc $B = B'$ i B jest bazą przestrzeni V.

Z drugiej strony, jeśli B generuje przestrzeń V, to można B zredukować (np. za pomocą procedury odrzucania) do bazy B'' przestrzeni V. Ponieważ $B'' \subseteq B$ i $n = \dim V = |B''| \leqslant |B| = n$, więc $B = B''$ i B jest bazą przestrzeni V.

Stąd i z definicji bazy wynikają równoważności warunków (1), (2) i (3). $\square$

Przykład 154. Pokazać, że układ $B = (1,\ x-1,\ (x-1)^2,\ \ldots,\ (x-1)^{10})$ jest bazą przestrzeni $R_{10}[x]$.

Ponieważ $\dim R_{10}[x] = 11$ i układ B ma 11 elementów, więc wobec twierdzenia 7.5.4 wystarczy pokazać, że B generuje przestrzeń $R_{10}[x]$.

Niech $\varphi(x)$ będzie dowolnym wielomianem z przestrzeni $R_{10}[x]$. Wielomian $\varphi(x)$ ma pochodną dowolnego rzędu i $\varphi^{(k)}(x) = 0$ dla $k \geqslant 11$, więc ze wzoru Taylora (znanego z kursu analizy matematycznej) mamy

$$\varphi(x) = \varphi(1) + \frac{\varphi'(1)}{1!}(x-1) + \frac{\varphi''(1)}{2!}(x-1)^2 + \ldots + \frac{\varphi^{(10)}(1)}{10!}(x-1)^{10},$$

co oznacza, że $\varphi(x)$ jest kombinacją liniową wektorów z układu B. Stąd wynika, że B generuje przestrzeń $R_{10}[x]$ i dlatego jest on bazą przestrzeni $R_{10}[x]$.

Przykład 155. Liniowo niezależny układ wektorów

$$B = \left(\begin{bmatrix} 3 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right)$$

uzupełnić do bazy przestrzeni $R_{2\times 2}$.

Ponieważ $\dim R_{2\times 2} = 4$ i wektory należące do układu B są liniowo niezależne, więc wobec twierdzenia 7.5.4 wystarczy układ B uzupełnić o wektor z przestrzeni $R_{2\times 2}$, który nie należy do podprzestrzeni $\mathcal{L}(B)$. Można to zrobić "metodą prób i błędów" lub – jak to tu zrobimy – za pomocą opisu podprzestrzeni $\mathcal{L}(B)$.

Zauważmy, że macierz $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ należy do $\mathcal{L}(B)$ pod warunkiem, że dla pewnych liczb x, y i z jest

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = x\begin{bmatrix} 3 & 1 \\ 0 & 0 \end{bmatrix} + y\begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix} + z\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3x+2y+z & x \\ y & z \end{bmatrix}.$$

To jest możliwe tylko wtedy, gdy $a = 3b+2c+d$. Zatem $\mathcal{L}(B)$ jest zbiorem tych macierzy $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in R_{2\times 2}$, dla których $a = 3b+2c+d$. W szczególności $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \notin \mathcal{L}(B)$ i zbiór $B' = \left(\begin{bmatrix} 3 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right)$ jest uzupełnieniem zbioru B do bazy przestrzeni $R_{2\times 2}$.

Ze względów praktycznych warto pamiętać o następujących werbalnych konsekwencjach wcześniejszych twierdzeń i definicji.

Wniosek 7.5.3. (1) *Jeśli V jest n-wymiarową przestrzenią wektorową, to każdy podzbiór zbioru V mający więcej niż n wektorów jest liniowo zależny. Jednocześnie żaden podzbiór zbioru V mający mniej niż n wektorów nie generuje przestrzeni V.*

(2) *Bazą skończenie wymiarowej przestrzeni V jest każdy największy zbiór liniowo niezależny w przestrzeni V oraz każdy najmniejszy podzbiór zbioru V generujący przestrzeń V.* $\square$

W naszych dalszych rozważaniach skorzystamy z następujących relacji pomiędzy wymiarem przestrzeni i jej podprzestrzeni.

Twierdzenie 7.5.5. *Niech W będzie podprzestrzenią skończenie wymiarowej przestrzeni V. Wtedy $\dim W \leqslant \dim V$ i, dodatkowo, $\dim W = \dim V$ wtedy i tylko wtedy, gdy $W = V$.*

Dowód. Teza twierdzenia jest oczywista, gdy $W = \{\mathbf{0}\}$. Załóżmy zatem, że $W \neq \{\mathbf{0}\}$. Niech $\dim V = n$ i niech $(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k)$ będzie bazą przestrzeni W. Ponieważ wektory $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k$ są liniowo niezależne, więc wobec wniosku 7.5.3 jest $\dim W = k \leqslant n = \dim V$. Dla dowodu drugiej części twierdzenia odnotujmy najpierw, że z równości $W = V$ w oczywisty sposób wynika równość $\dim W = \dim V$. Załóżmy teraz, że $\dim W = \dim V$. Wtedy $k = n$ i wobec twierdzenia 7.5.4 zbiór $(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k)$ jest bazą przestrzeni V i dlatego $W = \mathcal{L}(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k) = V$. $\square$

7.6. Współrzędne wektora

W naszych dotychchczasowych rozważaniach najwięcej uwagi poświęciliśmy przestrzeni K^n i nie było to przypadkowe, bo – jak teraz pokażemy – każdy wektor z przestrzeni V (nad ciałem K) mającej bazę składającą się z n wektorów można utożsamiać z wektorem z przestrzeni K^n i w konsekwencji także działania na wektorach z przestrzeni V można utożsamiać z działaniami na wektorach z przestrzeni K^n.

Definicja 7.6.1. Niech układ $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ będzie bazą przestrzeni wektorowej V nad ciałem K i niech $\mathbf{v}$ będzie wektorem z przestrzeni V. Wobec twierdzenia 7.5.1 istnieją jednoznacznie wyznaczone skalary $r_1, r_2, \ldots, r_n \in K$ takie, że

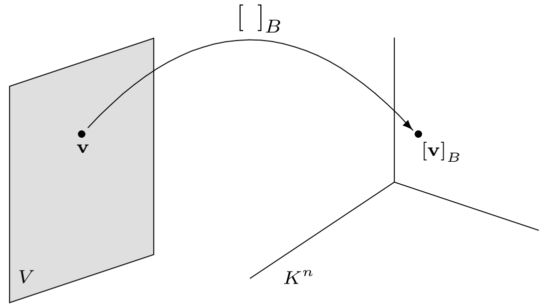
$$\mathbf{v} = r_1\mathbf{b}_1 + r_2\mathbf{b}_2 + \ldots + r_n\mathbf{b}_n.$$

Wektor

$$[\mathbf{v}]_B = (r_1, r_2, \dots, r_n) \quad \text{lub częściej} \quad [\mathbf{v}]_B = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}$$

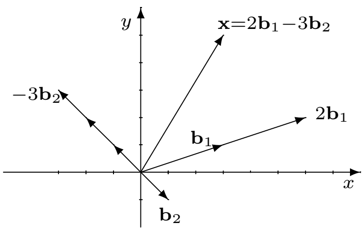
$[\mathbf{v}]_B$ – wektor współrzędnych

nazywamy *wektorem współrzędnych* wektora $\mathbf{v}$ względem bazy B przestrzeni V, a skalary $r_1, r_2, \dots, r_n$ – *współrzędnymi* wektora $\mathbf{v}$ względem bazy B (lub B-*współrzędnymi* wektora $\mathbf{v}$).



Rys. 7.10. Wektorowi $\mathbf{v} \in V$ odpowiada wektor $[\mathbf{v}]_B \in K^n$

Przykład 156. Dana jest baza $B = (\mathbf{b}_1, \mathbf{b}_2) = \left(\begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$ przestrzeni R^2. Wyznaczyć wektor współrzęnych $[\mathbf{v}]_B$ wektora $\mathbf{v} = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$. Wyznaczyć także wektor $\mathbf{x} \in R^2$, którego wektorem współrzędnych jest $[\mathbf{x}]_B = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$.



Rys. 7.11. Wektor $\mathbf{x} = 2\mathbf{b}_1 - 3\mathbf{b}_2$

Ponieważ

$$\mathbf{v} = \begin{bmatrix} 5 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 2 \\ -2 \end{bmatrix} = 1\mathbf{b}_1 + 2\mathbf{b}_2,$$

więc

$$[\mathbf{v}]_B = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Z definicji wektora współrzędnych mamy

$$\mathbf{x} = r_1\mathbf{b}_1 + r_2\mathbf{b}_2 = 2\begin{bmatrix} 3 \\ 1 \end{bmatrix} + (-3)\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}.$$

Przykład 157. Znaleźć wektory współrzędnych wektorów $\mathbf{v} = (9, 8)$, $\mathbf{u} = (-1, 3)$, $\mathbf{v} + \mathbf{u} = (8, 11)$ i $3\mathbf{u} = (-3, 9)$ względem bazy $B = (\mathbf{b}_1, \mathbf{b}_2) = \big((1, 2), (3, 1)\big)$ przestrzeni R^2.

Szukane wektory współrzędnych $[\mathbf{v}]_B$, $[\mathbf{u}]_B$, $[\mathbf{v} + \mathbf{v}]_B$ i $[3\mathbf{u}]_B$ są odpowiednio rozwiązaniami równań

$$[\,\mathbf{b}_1\,\mathbf{b}_2\,]\,\mathbf{x} = \mathbf{v}, \quad [\,\mathbf{b}_1\,\mathbf{b}_2\,]\,\mathbf{x} = \mathbf{u}, \quad [\,\mathbf{b}_1\,\mathbf{b}_2\,]\,\mathbf{x} = \mathbf{v} + \mathbf{u} \quad \text{i} \quad [\,\mathbf{b}_1\,\mathbf{b}_2\,]\,\mathbf{x} = 3\mathbf{u}.$$

Dla ich wyznaczenia macierz $[\,\mathbf{b}_1\,\mathbf{b}_2\,|\,\mathbf{v}\ \mathbf{u}\ \mathbf{v}{+}\mathbf{u}\ 3\mathbf{u}]$ (odpowiadającą powyższym czterem równaniom) sprowadzamy do wierszowo równoważnej normalnej macierzy schodkowej[1],

$$[\,\mathbf{b}_1\,\mathbf{b}_2\,|\,\mathbf{v}\ \mathbf{u}\ \mathbf{v}+\mathbf{u}\ 3\mathbf{u}] = \left[\begin{array}{cc|cccc} 1 & 3 & 9 & -1 & 8 & -3 \\ 2 & 1 & 8 & 3 & 11 & 9 \end{array}\right] \sim \dots \sim \left[\begin{array}{cc|cccc} 1 & 0 & 3 & 2 & 5 & 6 \\ 0 & 1 & 2 & -1 & 1 & -3 \end{array}\right].$$

[1] W ogólnym przypadku wektor współrzędnych $[\mathbf{v}]_B$ wektora $\mathbf{v}$ z przestrzeni K^n względem bazy $B = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ jest rozwiązaniem równania $[\mathbf{b}_1\,\mathbf{b}_2 \dots \mathbf{b}_n]\mathbf{x} = \mathbf{v}$. Rozwiązanie to możemy wyznaczyć metodą Gaussa-Jordana, czyli sprowadzając macierz $[\mathbf{b}_1\,\mathbf{b}_2 \dots \mathbf{b}_n|\,\mathbf{v}]$ do wierszowo równoważnej normalnej macierzy schodkowej $[I_n|\,[\mathbf{v}]_B]$.

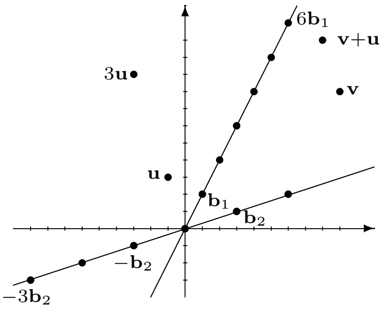
Z ostatniej macierzy otrzymujemy $[\mathbf{v}]_B = (3,2)$, $[\mathbf{u}]_B = (2,-1)$, $[\mathbf{v}+\mathbf{u}]_B = (5,1)$ i $[3\mathbf{u}]_B = (6,-3)$, zobacz rysunek 7.12.

Warto zauważyć, że dla wektorów z ostatniego przykładu mamy:

$$[\mathbf{v}+\mathbf{u}]_B = (5,1) = (3,2) + (2,-1) = [\mathbf{v}]_B + [\mathbf{u}]_B$$

i

$$[3\mathbf{u}]_B = (6,-3) = 3(2,-1) = 3[\mathbf{u}]_B.$$



Rys. 7.12

Powyższe własności są szczególnymi przypadkami dowodzonych w następnym twierdzeniu ogólniejszych własności: w skończenie wymiarowej przestrzeni z ustaloną bazą wektor współrzędnych sumy wektorów jest równy sumie wektorów współrzędnych tych wektorów i wektor współrzędnych iloczynu wektora przez skalar jest równy iloczynowi wektora współrzędnych przez ten sam skalar.

Twierdzenie 7.6.1. *Niech B będzie bazą n-wymiarowej przestrzeni V nad ciałem K. Jeśli $\mathbf{v}, \mathbf{u} \in V$ i $t \in K$, to*

$$[\mathbf{v}+\mathbf{u}]_B = [\mathbf{v}]_B + [\mathbf{u}]_B \quad i \quad [t\,\mathbf{v}]_B = t\,[\mathbf{v}]_B.$$

Dowód. Niech $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ będzie bazą przestrzeni V i niech $(r_1, r_2, \ldots, r_n)$ oraz $(s_1, s_2, \ldots, s_n)$ będą odpowiednio wektorami współrzędnych wektorów $\mathbf{v}$ i $\mathbf{u}$ względem bazy B. Wtedy $\mathbf{v} = r_1\mathbf{b}_1 + r_2\mathbf{b}_2 + \ldots + r_n\mathbf{b}_n$, $\mathbf{u} = s_1\mathbf{b}_1 + s_2\mathbf{b}_2 + \ldots + s_n\mathbf{b}_n$ i $\mathbf{v}+\mathbf{u} = (r_1+s_1)\mathbf{b}_1 + (r_2+s_2)\mathbf{b}_2 + \ldots + (r_n+s_n)\mathbf{b}_n$. Zatem dla wektora współrzędnych wektora $\mathbf{v}+\mathbf{u}$ mamy

$$\begin{aligned}[\mathbf{v}+\mathbf{u}]_B &= (r_1+s_1, r_2+s_2, \ldots, r_n+s_n)\\ &= (r_1, r_2, \ldots, r_n) + (s_1, s_2, \ldots, s_n)\\ &= [\mathbf{v}]_B + [\mathbf{u}]_B.\end{aligned}$$

Podobnie mamy

$$t\,\mathbf{v} = t\,(r_1\mathbf{b}_1 + r_2\mathbf{b}_2 + \ldots + r_n\mathbf{b}_n) = (tr_1)\mathbf{b}_1 + (tr_2)\mathbf{b}_2 + \ldots + (tr_n)\mathbf{b}_n,$$

więc także

$$[t\,\mathbf{v}]_B = (t\,r_1, t\,r_2, \ldots, t\,r_n) = t\,(r_1, r_2, \ldots, r_n) = t\,[\mathbf{v}]_B. \ \square$$

Definicja 7.6.2. Niech V i W będą przestrzeniami wektorowymi nad tym samym ciałem K. Przekształcenie $\varphi : V \to W$ nazywamy *izomorfizmem* przestrzeni V na przestrzeń W, gdy ma ono następujące własności:

Izomorfizm przestrzeni

(1) φ jest różnowartościowe i $\varphi(V) = W$;

(2) $\varphi(\mathbf{x}+\mathbf{y}) = \varphi(\mathbf{x}) + \varphi(\mathbf{y})$ dla każdych $\mathbf{x}, \mathbf{y} \in V$, tj. φ zachowuje działanie dodawania;

(3) $\varphi(\alpha\mathbf{x}) = \alpha\varphi(\mathbf{x})$ dla każdych $\mathbf{x} \in V$ i $\alpha \in K$, tj. φ zachowuje mnożenie wektorów przez skalary.

Izomorficzność przestrzeni

Dwie przestrzenie wektorowe nazywamy *izomorficznymi*, gdy istnieje izomorfizm odwzorowujący jedną z nich na drugą. Łatwo zauważyć, że: (1) przekształcenie tożsamościowe przestrzeni V na siebie jest izomorfizmem; (2) przekształcenie odwrotne do izomorfizmu przestrzeni V na przestrzeń W jest izomorfizmem przestrzeni W na przestrzeń V; (3) złożenie izomorfizmu przestrzeni V na przestrzeń W i izomorfizmu przestrzeni W na przestrzeń U jest izomorfizmem przestrzeni V na przestrzeń U. Zatem relacja izomorfizmu przestrzeni wektorowych jest relacją równoważności w zbiorze wszystkich przestrzeni wektorowych (nad tym samym ciałem) i jest podstawą algebraicznej identyfikacji przestrzeni izomorficznych.

W konsekwencji, jeżeli w przestrzeni wektorowej V zostało udowodnione jakiekolwiek twierdzenie sformułowane w terminach dodawania wektorów i mnożenia wektorów przez skalary, to dokładnie to samo twierdzenie jest prawdziwe w każdej przestrzeni izomorficznej z przestrzenią V.

$\dim V(K) = n$ / $V(K) \approx K^n$

Twierdzenie 7.6.2. *Każda n-wymiarowa przestrzeń wektorowa V nad ciałem K jest izomorficzna z przestrzenią K^n.*

Dowód. Niech $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ będzie bazą przestrzeni V. Odwzorowanie $\varphi : V \to K^n$, które każdemu wektorowi $\mathbf{v}$ z przestrzeni V przyporządkowuje jego wektor współrzędnych względem bazy B, tj. $\varphi(\mathbf{v}) = [\mathbf{v}]_B$, zachowuje działania przestrzeni V (zob. twierdzenie 7.6.1), jest różnowartościowe (bo jeśli $\varphi(\mathbf{v}) = (r_1, r_2, \ldots, r_n) = \varphi(\mathbf{u})$, to $\mathbf{v} = r_1\mathbf{b}_1 + r_2\mathbf{b}_2 + \ldots + r_n\mathbf{b}_n = \mathbf{u}$) i odwzorowuje zbiór V na cały zbiór K^n (bo dla każdego wektora $(s_1, s_2, \ldots, s_n) \in K^n$ mamy $(s_1, s_2, \ldots, s_n) = \varphi(\mathbf{v})$, gdy $\mathbf{v} = s_1\mathbf{b}_1 + s_2\mathbf{b}_2 + \ldots + s_n\mathbf{b}_n$), więc jest izomorfizmem przestrzeni V i K^n. □

Jeśli V i W są n-wymiarowymi przestrzeniami wektorowymi nad ciałem K, to wobec twierdzenia 7.6.2 każda z nich jest izomorficzna z przestrzenią K^n i dlatego przestrzenie V i W są wzajemnie izomorficzne.

Wniosek 7.6.1. *Każde dwie n-wymiarowe przestrzenie wektorowe (nad tym samym ciałem) są izomorficzne.* □

Ponieważ n-wymiarowa przestrzeń V nad ciałem K jest izomorficzna z przestrzenią K^n, więc przestrzeń K^n, wraz z jej technikami macierzowymi, można stosować do badania przestrzeni V. Tego rodzaju możliwości ilustrujemy w kolejnym wniosku i przykładzie.

Wniosek 7.6.2. *Wektory $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ z n-wymiarowej przestrzeni wektorowej $V(K)$ z bazą B są liniowo niezależne wtedy i tylko wtedy, gdy wektory $[\mathbf{v}_1]_B$, $[\mathbf{v}_2]_B$, ..., $[\mathbf{v}_k]_B$ są liniowo niezależne w przestrzeni K^n.*

Dowód. Ponieważ odwzorowanie $\varphi : V(K) \to K^n$, gdzie $\varphi(\mathbf{x}) = [\mathbf{x}]_B$ dla $\mathbf{x} \in V(K)$, jest izomorfizmem (zob. dowód twierdzenia 7.6.2), więc dla skalarów $\alpha_1, \ldots, \alpha_k \in K$ mamy równoważności

$$\begin{aligned} \alpha_1\mathbf{v}_1 + \ldots + \alpha_k\mathbf{v}_k = \mathbf{0} \quad &\Leftrightarrow \quad [\alpha_1\mathbf{v}_1 + \ldots + \alpha_k\mathbf{v}_k]_B = [\mathbf{0}]_B = \mathbf{0} \\ &\Leftrightarrow \quad \alpha_1[\mathbf{v}_1]_B + \ldots + \alpha_k[\mathbf{v}_k]_B = \mathbf{0}. \end{aligned}$$

To oznacza, że pomiędzy wektorami $\mathbf{v}_1, \ldots, \mathbf{v}_k$ istnieją dokładnie takie same relacje liniowej zależności jak pomiędzy wektorami $[\mathbf{v}_1]_B, \ldots, [\mathbf{v}_k]_B$. Stąd wynika teza wniosku. □

Przykład 158. Pokazać, że uporządkowany zbiór

$$C = (\mathbf{a}, \mathbf{b}, \mathbf{c}) = (x^2 + 2x - 3, 2x^2 + 2x + 1, -3x^2 + x + 1)$$

jest bazą przestrzeni $R_2[x]$. Następnie znaleźć współrzędne wektora $\mathbf{v} = -7x^2 - x - 8$ względem bazy C i przedstawić wektor $\mathbf{v}$ jako kombinację liniową wektorów bazy C.

Weźmy pod uwagę bazę $B = (x^2, x, 1)$ przestrzeni $R_2[x]$. Wektorami współrzędnych wektorów $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ i $\mathbf{v}$ względem bazy B są $[\mathbf{a}]_B = (1, 2, -3)$, $[\mathbf{b}]_B = (2, 2, 1)$, $[\mathbf{c}]_B = (-3, 1, 1)$ i $[\mathbf{v}]_B = (-7, -1, -8)$.

Dla dowodu, że wektory $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ tworzą bazę przestrzeni $R_2[x]$ (izomorficznej z R^3) wystarczy pokazać, że wektory $[\mathbf{a}]_B$, $[\mathbf{b}]_B$, $[\mathbf{c}]_B$ tworzą bazę przestrzeni R^3. Podobnie wyznaczając współrzędne wektora $\mathbf{v}$ względem bazy C wystarczy znaleźć współrzędne wektora $[\mathbf{v}]_B$ względem bazy $([\mathbf{a}]_B, [\mathbf{b}]_B, [\mathbf{c}]_B)$ (bo wobec twierdzenia 7.6.1 jest $\mathbf{v} = r_1\mathbf{a} + r_2\mathbf{b} + r_3\mathbf{c}$ wtedy i tylko wtedy, gdy $[\mathbf{v}]_B = [r_1\mathbf{a} + r_2\mathbf{b} + r_3\mathbf{c}]_B = r_1[\mathbf{a}]_B + r_2[\mathbf{b}]_B + r_3[\mathbf{c}]_B$, co oznacza, że współrzędne wektora $\mathbf{v}$ względem bazy $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ są równe współrzędnym wektora $[\mathbf{v}]_B$ względem bazy $([\mathbf{a}]_B, [\mathbf{b}]_B, [\mathbf{c}]_B)$). Dla jednoczesnego osiągnięcia obu wspomnianych celów tworzymy macierz $\Big[[\mathbf{a}]_B \, [\mathbf{b}]_B \, [\mathbf{c}]_B \,|\, [\mathbf{v}]_B \Big]$ i sprowadzamy ją do wierszowo równoważnej macierzy mającej normalną postać schodkową:

$$\left[\begin{array}{rrr|r} 1 & 2 & -3 & -7 \\ 2 & 2 & 1 & -1 \\ -3 & 1 & 1 & -8 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & 2 & -3 & -7 \\ 0 & -2 & 7 & 13 \\ 0 & 7 & -8 & -29 \end{array}\right] \sim \ldots \sim \left[\begin{array}{rrr|r} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 1 \end{array}\right].$$

Ponieważ macierz $\Big[[\mathbf{a}]_B \, [\mathbf{b}]_B \, [\mathbf{c}]_B \Big]$ jest wierszowo równoważna macierzy jednostkowej I_3, więc jest ona nieosobliwa i wobec wniosku 7.4.4 układ $\big([\mathbf{a}]_B, [\mathbf{b}]_B, [\mathbf{c}]_B\big)$ jest bazą przestrzeni R^3. Zatem układ $C = (\mathbf{a}, \mathbf{b}, \mathbf{c})$ jest bazą przestrzeni $R_2[x]$. Dodatkowo mamy $[\mathbf{v}]_C = (2, -3, 1)$ i dlatego

$$\begin{aligned} \mathbf{v} &= 2\mathbf{a} - 3\mathbf{b} + \mathbf{c} \\ &= 2(x^2 + 2x - 3) - 3(2x^2 + 2x + 1) + (-3x^2 + x + 1) \\ &= -7x^2 - x - 8. \end{aligned}$$

Macierz przejścia od bazy do bazy

Niech $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ i $C = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n)$ będą dwiema bazami przestrzeni wektorowej $V(K)$ i niech $\mathbf{x}$ będzie dowolnym wektorem z przestrzeni $V(K)$. Zbadamy teraz jaki jest związek pomiędzy wektorami współrzędnych

$$[\mathbf{x}]_B = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} \quad \text{i} \quad [\mathbf{x}]_C = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}$$

wektora $\mathbf{x}$ względem baz B i C. Opiszemy sposób przedstawiania współrzędnych wektora $[\mathbf{x}]_B$ poprzez współrzędne wektora $[\mathbf{x}]_C$.

Z definicji wektora współrzędnych jest $\mathbf{x} = r_1\mathbf{b}_1 + r_2\mathbf{b}_2 + \ldots + r_n\mathbf{b}_n$. Zatem wobec twierdzenia 7.6.1 mamy

$$\begin{aligned} [\mathbf{x}]_C &= [r_1\mathbf{b}_1 + r_2\mathbf{b}_2 + \ldots + r_n\mathbf{b}_n]_C \\ &= r_1[\mathbf{b}_1]_C + r_2[\mathbf{b}_2]_C + \ldots + r_n[\mathbf{b}_n]_C \\ &= \begin{bmatrix} | & | & & | \\ [\mathbf{b}_1]_C & [\mathbf{b}_2]_C & \cdots & [\mathbf{b}_n]_C \\ | & | & & | \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} \\ &= \begin{bmatrix} | & | & & | \\ [\mathbf{b}_1]_C & [\mathbf{b}_2]_C & \cdots & [\mathbf{b}_n]_C \\ | & | & & | \end{bmatrix} [\mathbf{x}]_B. \end{aligned}$$

Stąd zaś wynika, że mamy następujące twierdzenie o zależnościach pomiędzy B-współrzędnymi i C-współrzędnymi wektora $\mathbf{x}$.

Twierdzenie 7.6.3. *Jeśli $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ i $C = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n)$ są bazami przestrzeni wektorowej V, to dla każdego wektora $\mathbf{x}$ z przestrzeni V jest*

$$[\mathbf{x}]_C = \mathbf{P}_C^B \, [\mathbf{x}]_B, \tag{7.6}$$

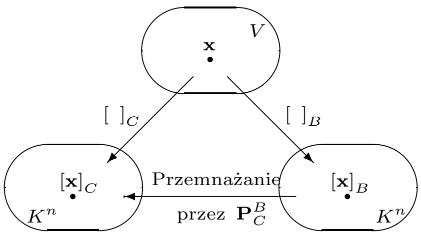
gdzie

$$\mathbf{P}_C^B = \begin{bmatrix} | & | & & | \\ [\mathbf{b}_1]_C & [\mathbf{b}_2]_C & \cdots & [\mathbf{b}_n]_C \\ | & | & & | \end{bmatrix}. \tag{7.7}$$

Macierz przejścia

Macierz $\mathbf{P}_C^B$ określoną równością (7.7) nazywa się *macierzą przejścia* od bazy B do bazy C (lub *macierzą zamiany współrzędnych* przy przejściu od bazy B do bazy C). W macierzy tej i-tą kolumną jest wektor $[\mathbf{b}_i]_C$, czyli wektor C-współrzędnych wektora $\mathbf{b}_i$, i jest on jedynym rozwiązaniem równania

$$x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \ldots + x_n\mathbf{c}_n = \mathbf{b}_i.$$



Rys. 7.13

Równość $[\mathbf{x}]_C = \mathbf{P}_C^B\,[\mathbf{x}]_B$ jest związkiem pomiędzy B-współrzędnymi i C-współrzędnymi wektora $\mathbf{x}$, zob. rys. 7.13. Ponieważ wektory $\mathbf{b}_1, \ldots, \mathbf{b}_n$ bazy B są liniowo niezależne, więc także kolumny $[\mathbf{b}_1]_C, \ldots, [\mathbf{b}_n]_C$ macierzy $\mathbf{P}_C^B$ są liniowo niezależne (zob. wniosek 7.6.2). Stąd i z wniosku 7.4.4 wynika, że macierz $\mathbf{P}_C^B$ jest odwracalna. Zatem z równości $[\mathbf{x}]_C = \mathbf{P}_C^B\,[\mathbf{x}]_B$ mamy także równość $[\mathbf{x}]_B = \left(\mathbf{P}_C^B\right)^{-1}[\mathbf{x}]_C$. To zaś oznacza, że macierz odwrotna $\left(\mathbf{P}_C^B\right)^{-1}$ przekształca C-współrzędne wektora $\mathbf{x}$ w jego B-współrzędne. Dlatego macierz $\left(\mathbf{P}_C^B\right)^{-1}$ jest macierzą przejścia od bazy C do bazy B i mamy

$$\left(\mathbf{P}_C^B\right)^{-1} = \mathbf{P}_B^C. \tag{7.8}$$

Przykład 159. Układy wektorów $B = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4)$ i $C = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4)$ są bazami przestrzeni R^4, gdy $\mathbf{b}_1 = (1,2,3,0)$, $\mathbf{b}_2 = (1,-1,1,1)$, $\mathbf{b}_3 = (3,1,6,3)$, $\mathbf{b}_4 = (5,3,11,6)$ i $\mathbf{c}_1 = (1,2,3,0)$, $\mathbf{c}_2 = (2,1,4,1)$, $\mathbf{c}_3 = (0,1,1,1)$, $\mathbf{c}_4 = (0,0,0,1)$. Ponieważ

$$\begin{cases} \mathbf{b}_1 = \mathbf{c}_1, \\ \mathbf{b}_2 = -\mathbf{c}_1 + \mathbf{c}_2, \\ \mathbf{b}_3 = -\mathbf{c}_1 + 2\mathbf{c}_2 + \mathbf{c}_3, \\ \mathbf{b}_4 = -\mathbf{c}_1 + 3\mathbf{c}_2 + 2\mathbf{c}_3 + \mathbf{c}_4, \end{cases}$$

więc macierzą przejścia od bazy B do bazy C jest

$$\mathbf{P}_C^B = \begin{bmatrix} | & & | \\ [\mathbf{b}_1]_C & \cdots & [\mathbf{b}_4]_C \\ | & & | \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 & -1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Łatwo zauważyć, że jednocześnie mamy

$$\begin{cases} \mathbf{c}_1 = \mathbf{b}_1, \\ \mathbf{c}_2 = \mathbf{b}_1 + \mathbf{b}_2, \\ \mathbf{c}_3 = -\mathbf{b}_1 - 2\mathbf{b}_2 + \mathbf{b}_3, \\ \mathbf{c}_4 = \mathbf{b}_2 - 2\mathbf{b}_3 + \mathbf{b}_4, \end{cases}$$

więc także

$$\mathbf{P}_B^C = \begin{bmatrix} | & & | \\ [\mathbf{c}_1]_B & \cdots & [\mathbf{c}_4]_B \\ | & & | \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \left(\mathbf{P}_C^B\right)^{-1}.$$

Za pomocą tak otrzymanych macierzy $\mathbf{P}_C^B$ i $\mathbf{P}_B^C$ można wyrazić związek pomiędzy B-współrzędnymi i C-współrzędnymi każdego wektora z przestrzeni R^4. Przykładowo dla wektora $\mathbf{x} = (3,2,6,-1) \in R^4$ jest $\mathbf{x} = 3\mathbf{b}_1 + 2\mathbf{b}_2 + \mathbf{b}_3 - \mathbf{b}_4$ i dlatego

$$[\mathbf{x}]_B = \begin{bmatrix} 3 \\ 2 \\ 1 \\ -1 \end{bmatrix} \quad \text{oraz } [\mathbf{x}]_C = \mathbf{P}_C^B \cdot [\mathbf{x}]_B = \begin{bmatrix} 1 & -1 & -1 & -1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}.$$

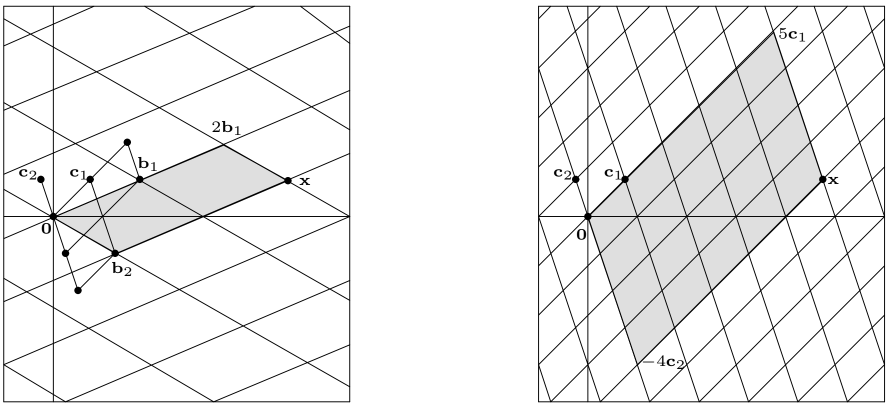
Przykład 160. Dane są bazy $B = (\mathbf{b}_1, \mathbf{b}_2)$ i $C = (\mathbf{c}_1, \mathbf{c}_2)$ przestrzeni V, gdzie $\mathbf{b}_1 = 2\mathbf{c}_1 - \mathbf{c}_2$ i $\mathbf{b}_2 = \mathbf{c}_1 - 2\mathbf{c}_2$, i dany jest wektor $\mathbf{x} = 2\mathbf{b}_1 + \mathbf{b}_2$. Wyznaczyć wektor $[\mathbf{x}]_C$.

Dla wektora $\mathbf{x}$ (zob. rys. 7.14) podobnie jak w wyprowadzeniu twierdzenia 7.6.3 mamy następujące zależności między jego B- i C-współrzędnymi:

$$\begin{aligned}
[\mathbf{x}]_C &= [2\mathbf{b}_1 + \mathbf{b}_2]_C = 2[\mathbf{b}_1]_C + [\mathbf{b}_2]_C \\
&= \left[\ [\mathbf{b}_1]_C \quad [\mathbf{b}_2]_C\ \right]\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \left[\ [2\mathbf{c}_1 - \mathbf{c}_2]_C \quad [\mathbf{c}_1 - 2\mathbf{c}_2]_C\ \right]\begin{bmatrix} 2 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} 2 & 1 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ -4 \end{bmatrix}.
\end{aligned}$$

Drugi sposób. Ponieważ wektor $\mathbf{x}$ jest kombinacją liniową wektorów $\mathbf{b}_1$ i $\mathbf{b}_2$, a te ostatnie są kombinacjami wektorów $\mathbf{c}_1$ i $\mathbf{c}_2$, więc $\mathbf{x}$ jest kombinacją liniową wektorów $\mathbf{c}_1$ oraz $\mathbf{c}_2$ i współczynniki tej kombinacji tworzą wektor $[\mathbf{x}]_C$,

$$[\mathbf{x}]_C = [\,2\mathbf{b}_1 + \mathbf{b}_2\,]_C = [\,2(2\mathbf{c}_1 - \mathbf{c}_2) + (\mathbf{c}_1 - 2\mathbf{c}_2)\,]_C = [\,5\mathbf{c}_1 - 4\mathbf{c}_2\,]_C = \begin{bmatrix} 5 \\ -4 \end{bmatrix}.$$



Rys. 7.14

Zmiana bazy w przestrzeni $\mathbf{K^n}$

Niech $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ i $C = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n)$ będą bazami przestrzeni wektorowej K^n i niech $\mathbf{B}$ i $\mathbf{C}$ będą macierzami utworzonymi z kolejnych wektorów baz B i C, $\mathbf{B} = [\,\mathbf{b}_1\,\mathbf{b}_2\,\ldots\,\mathbf{b}_n\,]$ i $\mathbf{C} = [\,\mathbf{c}_1\,\mathbf{c}_2\,\ldots\,\mathbf{c}_n\,]$. Pokażemy teraz związek macierzy $\mathbf{B}$ i $\mathbf{C}$ z macierzą przejścia od bazy B do bazy C, czyli z macierzą $\mathbf{P}_C^B$.

Zauważmy, że jeśli $E = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ jest bazą standardową przestrzeni K^n, to dla każdego wektora $\mathbf{x} \in K^n$ jest $[\mathbf{x}]_E = \mathbf{x}$ i dlatego wobec twierdzenia 7.6.3 dla macierzy przejścia od baz B i C do bazy standardowej E mamy

$$\mathbf{P}_E^B = [\,\mathbf{b}_1\,\mathbf{b}_2\,\ldots\,\mathbf{b}_n\,] = \mathbf{B} \quad \text{i} \quad \mathbf{P}_E^C = [\,\mathbf{c}_1\,\mathbf{c}_2\,\ldots\,\mathbf{c}_n\,] = \mathbf{C}.$$

Dodatkowo wobec (7.6) dla każdego wektora $\mathbf{x} \in K^n$ jest

$$\mathbf{C}\,[\mathbf{x}]_C = \mathbf{P}_E^C[\mathbf{x}]_C = [\mathbf{x}]_E = \mathbf{P}_E^B[\mathbf{x}]_B = \mathbf{B}\,[\mathbf{x}]_B.$$

Stąd i z odwracalności macierzy $\mathbf{C}$ wynika, że $[\mathbf{x}]_C = \mathbf{C}^{-1}\mathbf{B}\,[\mathbf{x}]_B$. Jednocześnie dla każdego $\mathbf{x} \in K^n$ jest $[\mathbf{x}]_C = \mathbf{P}_C^B[\mathbf{x}]_B$. Zatem mamy następujące zależności pomiędzy macierzą przejścia $\mathbf{P}_C^B$ i macierzami $\mathbf{B}$ oraz $\mathbf{C}$:

$$\mathbf{P}_C^B = \mathbf{C}^{-1}\mathbf{B}. \tag{7.9}$$

Macierz $\mathbf{C}^{-1}\mathbf{B} = \mathbf{P}_C^B$ jest jednocześnie rozwiązaniem równania macierzowego

$$\mathbf{CX} = \mathbf{B}$$

(reprezentującego równania $[\,\mathbf{c}_1\,\mathbf{c}_2\,\ldots\,\mathbf{c}_n\,]\mathbf{x} = \mathbf{b}_1, \ldots, [\,\mathbf{c}_1\,\mathbf{c}_2\,\ldots\,\mathbf{c}_n\,]\mathbf{x} = \mathbf{b}_n$, których rozwiązaniami są kolumny $[\mathbf{b}_1]_C, \ldots, [\mathbf{b}_n]_C$ macierzy $\mathbf{P}_C^B$) i dlatego można ją wyznaczyć metodą Gaussa-Jordana. Mamy następujący wygodny sposób wyznaczania macierzy przejścia.

Algorytm wyznaczania macierzy przejścia w $\mathbf{K}^\mathbf{n}$

Jeśli $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ i $C = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n)$ są bazami przestrzeni wektorowej K^n, to macierz przejścia od bazy B do bazy C jest macierzą $\mathbf{P}_C^B = \mathbf{C}^{-1}\mathbf{B}$ otrzymaną w wyniku wierszowej redukcji macierzy $[\,\mathbf{C}|\mathbf{B}\,]$ do normalnej macierzy schodkowej,

$$[\,\mathbf{C}|\mathbf{B}\,] = \left[\begin{array}{cccc|cccc} | & | & & | & | & | & & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n & \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \\ | & | & & | & | & | & & | \end{array}\right] \sim \left[\,\mathbf{I}_n|\mathbf{P}_C^B\,\right]. \tag{7.10}$$

Nowa baza Stara baza

Przykład 161. Wyznaczyć macierz przejścia $\mathbf{P}_C^B$ od bazy B do bazy C przestrzeni R^3, gdy $B = \big((1,2,3),(-2,-6,-10),(4,16,16)\big)$ i $C = \big((1,2,3),(1,1,1),(1,3,1)\big)$. Za pomocą tej macierzy wyznaczyć wektor C-współrzędnych wektora $\mathbf{v}$, dla którego $[\mathbf{v}]_B = (1,1,1)$.

Dla otrzymania macierzy $\mathbf{P}_C^B$ tworzymy macierz $[\mathbf{C}|\mathbf{B}]$ i sprowadzamy ją do wierszowo równoważnej normalnej macierzy schodkowej (zob. (7.10)). Ponieważ

$$[\mathbf{C}|\mathbf{B}] = \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & -2 & 4 \\ 2 & 1 & 3 & 2 & -6 & 16 \\ 3 & 1 & 1 & 3 & -10 & 16 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -4 & 6 \\ 0 & 1 & 0 & 0 & 2 & -5 \\ 0 & 0 & 1 & 0 & 0 & 3 \end{array}\right],$$

więc

$$\mathbf{P}_C^B = \begin{bmatrix} 1 & -4 & 6 \\ 0 & 2 & -5 \\ 0 & 0 & 3 \end{bmatrix}$$

i

$$[\mathbf{v}]_C = \mathbf{P}_C^B\,[\mathbf{v}]_B = \begin{bmatrix} 1 & -4 & 6 \\ 0 & 2 & -5 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \\ 3 \end{bmatrix}.$$

Przykład 162. Znaleźć macierz przejścia od bazy $B = \big(4x^2 - 6x, 2x^2 - 2, 4x\big)$ do bazy $C = \big(x+1, 2x^2, x-1\big)$ przestrzeni $R_2[x]$. Wyznaczyć także współrzędne wektora $\mathbf{v} = 10x^2 - 8x - 2 = 2(4x^2 - 6x) + (2x^2 - 2) + 4x$ względem bazy C.

Utożsamiając każdy wektor $a_2x^2 + a_1x + a_0$ z przestrzeni $R_2[x]$ z wektorem (a_2, a_1, a_0) z przestrzeni R^3, możemy utożsamiać bazy B i C odpowiednio z bazami $B' = \big((4,-6,0), (2,0,-2),(0,4,0)\big)$ i $C' = \big((0,1,1),(2,0,0),(0,1,-1)\big)$ przestrzeni R^3. Wtedy szukana macierz przejścia od bazy B do C jest macierzą przejścia od bazy B' do C'. Ponieważ

$$[\mathbf{C}'|\mathbf{B}'] = \left[\begin{array}{ccc|ccc} 0 & 2 & 0 & 4 & 2 & 0 \\ 1 & 0 & 1 & -6 & 0 & 4 \\ 1 & 0 & -1 & 0 & -2 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -3 & -1 & 2 \\ 0 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -3 & 1 & 2 \end{array}\right],$$

więc

$$\mathbf{P}_C^B = \begin{bmatrix} -3 & -1 & 2 \\ 2 & 1 & 0 \\ -3 & 1 & 2 \end{bmatrix}.$$

Zatem dla wektora $[\mathbf{v}]_C$ mamy

$$[\mathbf{v}]_C = \mathbf{P}_C^B[\mathbf{v}]_B = \begin{bmatrix} -3 & -1 & 2 \\ 2 & 1 & 0 \\ -3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ 5 \\ -3 \end{bmatrix}.$$

Zauważmy, że istotnie mamy $-5(x+1)+5(2x^2)-3(x-1) = 10x^2-8x-2 = \mathbf{v}$.

7.7. Rząd macierzy

Niech $\mathbf{A}$ będzie macierzą wymiaru $m \times n$ o współczynnikach z ciała K,

$$\mathbf{A} = \begin{bmatrix} | & | & & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} - & \mathbf{w}_1 & - \\ - & \mathbf{w}_2 & - \\ & \vdots & \\ - & \mathbf{w}_m & - \end{bmatrix},$$

gdzie $\mathbf{a}_1, \ldots, \mathbf{a}_n$ są kolumnami, a $\mathbf{w}_1, \ldots, \mathbf{w}_m$ wierszami macierzy $\mathbf{A}$.

Definicja 7.7.1. Liczbę $\dim \mathcal{L}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$, czyli wymiar przestrzeni kolumnowej $C_\mathbf{A}$ macierzy $\mathbf{A}$, nazywa się *rzędem kolumnowym* macierzy $\mathbf{A}$. Podobnie liczbę $\dim \mathcal{L}(\mathbf{w}_1, \ldots, \mathbf{w}_m)$, czyli wymiar przestrzeni wierszowej $R_\mathbf{A} = \mathcal{L}(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m)$ macierzy $\mathbf{A}$, nazywamy *rzędem wierszowym* macierzy $\mathbf{A}$. (Te dwa rzędy są odpowiednio największą liczbą liniowo niezależnych kolumn i największą liczbą liniowo niezależnych wierszy macierzy $\mathbf{A}$.)

Rząd kolumnowy

Rząd wierszowy

Przykład 163. Dane są wierszowo równoważne macierze $\mathbf{A}$ i $\mathbf{B}$, gdzie

$$\mathbf{A} = [\,\mathbf{a}_1 \ldots \mathbf{a}_5\,] = \begin{bmatrix} 3 & 4 & -2 & 7 & 6 \\ 2 & 3 & -1 & 6 & 7 \\ 2 & 4 & 0 & 7 & 7 \\ 1 & 1 & -1 & 2 & 2 \end{bmatrix} \sim \ldots \sim \begin{bmatrix} 1 & 0 & -2 & 0 & -1 \\ 0 & 1 & 1 & 0 & -3 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = [\,\mathbf{b}_1 \ldots \mathbf{b}_5\,] = \mathbf{B}.$$

Macierz $\mathbf{B}$ ma normalną postać schodkową i jej wiodące kolumny $\mathbf{b}_1$, $\mathbf{b}_2$ i $\mathbf{b}_4$ tworzą bazę jej przestrzeni kolumnowej, bo są one liniowo niezależne (dlaczego?) i pozostałe kolumny są ich kombinacjami liniowymi (tu oczywiście $\mathbf{b}_3 = -2\mathbf{b}_1 + \mathbf{b}_2$ i $\mathbf{b}_5 = -\mathbf{b}_1 - 3\mathbf{b}_2 + 3\mathbf{b}_4$). Zatem rząd kolumnowy macierzy $\mathbf{B}$ jest równy 3.

Trzy niezerowe wiersze macierzy $\mathbf{B}$ także są liniowo niezależne (bo żaden z nich – ze względu na położenie wiodących jedynek – nie jest kombinacją liniową pozostałych wierszy) i dlatego tworzą one bazę przestrzeni generowanej przez wiersze macierzy $\mathbf{B}$. Zatem rząd wierszowy macierzy $\mathbf{B}$ także jest równy 3.

Macierz $\mathbf{A}$ jest wierszowo równoważna macierzy $\mathbf{B}$ i – jak to będzie wynikało z następnego twierdzenia – jej przestrzeń wierszowa jest równa przestrzeni wierszowej macierzy $\mathbf{B}$. Z kolejnych twierdzeń będzie wynikało, że między kolumnami macierzy $\mathbf{A}$ są dokładnie takie same relacje liniowej zależności, jak między kolumnami macierzy $\mathbf{B}$. Z tego także będzie wynikała równość rzędów kolumnowych macierzy $\mathbf{A}$ i $\mathbf{B}$ (oraz równość rzędu wierszowego i kolumnowego macierzy $\mathbf{A}$).

Twierdzenie 7.7.1. *Jeśli macierze* $\mathbf{A}$ *i* $\mathbf{B}$ *są wierszowo równoważne, to ich przestrzenie wierszowe są równe. Jeśli macierz* $\mathbf{B}$ *jest w postaci schodkowej, to niezerowe wiersze macierzy* $\mathbf{B}$ *tworzą bazę przestrzeni wierszowej macierzy* $\mathbf{A}$ *i* $\mathbf{B}$.

Dowód. Niech $\mathbf{A}$ i $\mathbf{B}$ będą macierzami wierszowo równoważnymi. Ponieważ macierz $\mathbf{B}$ można otrzymać z macierzy $\mathbf{A}$ za pomocą operacji elementarnych na wierszach, więc wiersze macierzy $\mathbf{B}$ są kombinacjami liniowymi wierszy macierzy $\mathbf{A}$ i dlatego każda kombinacja liniowa wierszy macierzy $\mathbf{B}$ jest kombinacją liniową wierszy macierzy $\mathbf{A}$. Zatem przestrzeń wierszowa macierzy $\mathbf{B}$ jest zawarta w przestrzeni wierszowej macierzy $\mathbf{A}$. Z drugiej strony operacje elementarne są odwracalne, więc macierz $\mathbf{A}$ można otrzymać z macierzy $\mathbf{B}$ za pomocą operacji elementarnych na wierszach i dlatego przestrzeń wierszowa macierzy $\mathbf{A}$ jest zawarta w przestrzeni wierszowej macierzy $\mathbf{B}$. To dowodzi, że przestrzenie wierszowe macierzy $\mathbf{A}$ i $\mathbf{B}$ są równe.

$$\mathbf{B} = \begin{bmatrix} 0 & 1 & * & * & * & * & * & * \\ 0 & 0 & 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Rys. 7.15

Jeśli macierz $\mathbf{B}$ jest w postaci schodkowej (zob. rys. 7.15), to jej niezerowe wiersze są liniowo niezależne, bo żaden z nich nie jest kombinacją liniową wierszy leżących poniżej niego. Zatem niezerowe wiersze macierzy $\mathbf{B}$ tworzą bazę przestrzeni wierszowej macierzy $\mathbf{B}$ i dlatego także tworzą one bazę przestrzeni wierszowej macierzy $\mathbf{A}$. □

Twierdzenie 7.7.2. *Operacje elementarne na wierszach macierzy nie zmieniają relacji liniowej zależności pomiędzy kolumnami tej macierzy.*

Dowód. Jeśli macierze $\mathbf{A}$ i $\mathbf{B}$ są wierszowo równoważne, to także macierze $[\mathbf{A}|\mathbf{0}]$ i $[\mathbf{B}|\mathbf{0}]$ są wierszowo równoważne i wobec twierdzenia 5.1.2 układy równań liniowych $\mathbf{Ax} = \mathbf{0}$ i $\mathbf{Bx} = \mathbf{0}$ mają dokładnie te same rozwiązania. To zaś oznacza, że pomiędzy kolumnami macierzy $\mathbf{A}$ są dokładnie takie same liniowe zależności, jak pomiędzy kolumnami macierzy $\mathbf{B}$. □

Przykład 164. Weźmy pod uwagę wierszowo równoważne macierze $\mathbf{A}$ i $\mathbf{B}$ z poprzedniego przykładu. Wobec twierdzenia 7.7.2 pomiędzy kolumnami macierzy $\mathbf{A}$ są dokładnie takie same relacje liniowej zależności, jak pomiędzy kolumnami macierzy $\mathbf{B}$. Przykładowo mamy $\mathbf{a}_5 = -\mathbf{a}_1 - 3\mathbf{a}_2 + 3\mathbf{a}_4$, bo było $\mathbf{b}_5 = -\mathbf{b}_1 - 3\mathbf{b}_2 + 3\mathbf{b}_4$. Podobnie kolumny $\mathbf{a}_1$, $\mathbf{a}_2$ i $\mathbf{a}_4$ tworzą bazę przestrzeni kolumnowej macierzy $\mathbf{A}$, bo kolumny $\mathbf{b}_1$, $\mathbf{b}_2$ i $\mathbf{b}_4$ tworzą bazę przestrzeni kolumnowej macierzy $\mathbf{B}$.

Twierdzenie 7.7.3. *Wiodące kolumny macierzy* $\mathbf{A}$ *tworzą bazę przestrzeni kolumnowej macierzy* $\mathbf{A}$.

Dowód. Załóżmy, że macierz $\mathbf{A}$ jest wierszowo równoważna macierzy schodkowej $\mathbf{B}$. Wiodące kolumny macierzy $\mathbf{B}$ są liniowo niezależne, bo żadna z nich nie jest liniową kombinacją swoich poprzedniczek. Z tego i z twierdzenia 7.7.2 wynika, że wiodące kolumny macierzy $\mathbf{A}$ są liniowo niezależne. Z tych samych powodów każda inna kolumna macierzy $\mathbf{A}$ jest liniową kombinacją wiodących kolumn tej macierzy. Stąd wynika, że wiodące kolumny macierzy $\mathbf{A}$ tworzą bazę przestrzeni kolumnowej macierzy $\mathbf{A}$. □

Uwaga. Przestrzenie kolumnowe wierszowo równoważnych macierzy $\mathbf{A}$ i $\mathbf{B}$ mają równe wymiary, ale przestrzenie te nie muszą być równe. Dla przykładu, przestrzenie kolumnowe macierzy $\mathbf{A}$ i $\mathbf{B}$ z ostatniego przykładu są generowane odpowiednio przez kolumny $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_4$ oraz przez kolumny $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_4$ i przestrzenie te są różne, bo czwarta współrzędna wektora z przestrzeni kolumnowej macierzy $\mathbf{A}$ może być dowolną liczbą rzeczywistą, ale czwarta współrzędna każdego wektora z przestrzeni kolumnowej macierzy $\mathbf{B}$ jest równa zeru.

Twierdzenie 7.7.4. *Rząd kolumnowy macierzy* $\mathbf{A}$ *jest równy jej rzędowi wierszowemu.*

Dowód. Niech macierz $\mathbf{A}$ będzie wierszowo równoważna macierzy schodkowej $\mathbf{B}$. Wobec twierdzenia 7.7.1 rząd wierszowy macierzy $\mathbf{A}$ jest równy rzędowi wierszowemu macierzy $\mathbf{B}$, a ten jest równy liczbie niezerowych wierszy macierzy $\mathbf{B}$, więc także liczbie wiodących jedynek macierzy $\mathbf{B}$. Z drugiej strony wobec twierdzenia 7.7.3 rząd kolumnowy macierzy $\mathbf{A}$ jest równy liczbie wiodących kolumn macierzy $\mathbf{A}$, a ta liczba jest równa liczbie wiodących jedynek macierzy $\mathbf{B}$. Stąd wynika teza twierdzenia. □

Definicja 7.7.2. Wspólną wartość rzędu kolumnowego i wierszowego macierzy $\mathbf{A} \in K_{m\times n}$ nazywamy *rzędem macierzy* $\mathbf{A}$ i oznaczamy ją przez $r(\mathbf{A})$, czyli mamy

$r(\mathbf{A})$ – rząd macierzy

$$r(\mathbf{A}) = \dim C_{\mathbf{A}} = \dim R_{\mathbf{A}}. \tag{7.11}$$

Przykład 165. Wyznaczyć rząd macierzy

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & -4 & 0 & 8 \\ 1 & 1 & -3 & 0 & 6 \\ 3 & 3 & -9 & 1 & 20 \\ 2 & 2 & -6 & 2 & 16 \end{bmatrix}.$$

Łatwo zauważyć, że macierz $\mathbf{A}$ jest wierszowo równoważna normalnej macierzy schodkowej

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & -2 & 0 & 4 \\ 0 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Ponieważ ta ostatnia ma trzy niezerowe wiersze, więc jej rząd, jak i rząd macierzy $\mathbf{A}$ jest równy trzy.

Z faktu, że rząd macierzy $\mathbf{A}$ jednocześnie jest największą liczbą liniowo niezależnych wierszy i największą liczbą liniowo niezależnych kolumn macierzy $\mathbf{A}$ wynikają następujące dwa wnioski.

Wniosek 7.7.1. *Dla każdej macierzy* $\mathbf{A}$ *jest* $r(\mathbf{A}) = r(\mathbf{A}^T)$. $\square$

Wniosek 7.7.2. *Jeśli* $\mathbf{A}$ *jest macierzą wymiaru* $m \times n$, *to*

$$r(\mathbf{A}) \leqslant \min\{m, n\}. \quad \square$$

Nasze rozważania o rzędzie macierzy kończymy obserwacjami dotyczącymi związku rzędu macierzy kwadratowej $\mathbf{A}$ z jej odwracalnością oraz wyrażonego w terminach rzędu macierzy warunku koniecznego i dostatecznego istnienia rozwiązania układu równań liniowych $\mathbf{Ax} = \mathbf{b}$. Inne twierdzenie o związku rzędu macierzy $\mathbf{A}$ z rzędem macierzy $\mathbf{A}^T\mathbf{A}$ przedstawimy w następnym rozdziale (zob. twierdzenie 8.2.3).

Twierdzenie 7.7.5. *Macierz kwadratowa* $\mathbf{A}$ *stopnia* n *jest odwracalna wtedy i tylko wtedy, gdy* $r(\mathbf{A}) = n$.

Dowód. Teza twierdzenia jest natychmiastową konsekwencją wniosku 7.4.4 i definicji 7.7.2. $\square$

Twierdzenie 7.7.6 (Kroneckera-Capellego). *Układ równań liniowych* $\mathbf{Ax} = \mathbf{b}$ *ma rozwiązanie wtedy i tylko wtedy, gdy rząd macierzy* $\mathbf{A}$ *jest równy rzędowi macierzy rozszerzonej* $[\mathbf{A}|\mathbf{b}]$ *tego układu,*

$$r(\mathbf{A}) = r([\mathbf{A}|\mathbf{b}]). \tag{7.12}$$

Dowód. Wobec twierdzenia 7.3.1 układ $\mathbf{Ax} = \mathbf{b}$ ma rozwiązanie wtedy i tylko wtedy, gdy przestrzeń $C_{[\mathbf{A}|\mathbf{b}]}$ jest równa swojej podprzestrzeni $C_{\mathbf{A}}$. To zaś wobec twierdzenia 7.5.5 zachodzi wtedy i tylko wtedy, gdy $\dim C_{\mathbf{A}} = \dim C_{[\mathbf{A}|\mathbf{b}]}$, tj. wtedy i tylko wtedy, gdy $r(\mathbf{A}) = r([\mathbf{A}|\mathbf{b}])$ (bo $r(\mathbf{A}) = \dim C_{\mathbf{A}}$ i $r([\mathbf{A}|\mathbf{b}]) = \dim C_{[\mathbf{A}|\mathbf{b}]}$). $\square$

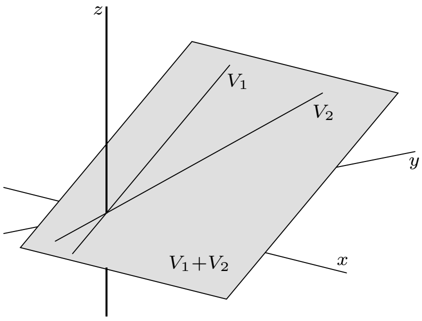
7.8. Suma i suma prosta podprzestrzeni

Suma podprzestrzeni

Definicja 7.8.1. Jeżli V_1 i V_2 są podprzestrzeniami przestrzeni wektorowej V, to zbiór

$$V_1 + V_2 = \{\mathbf{v}_1 + \mathbf{v}_2 : \mathbf{v}_1 \in V_1 \quad \text{i} \quad \mathbf{v}_2 \in V_2\}$$

nazywamy *sumą podprzestrzeni* V_1 i V_2.



Rys. 7.16

Przykład 166. Jeżli $V_1 = \mathcal{L}\big((0,1,1)\big)$ i $V_2 = \mathcal{L}\big((1,1,1)\big)$, to

$$\begin{aligned} V_1 + V_2 &= \{t(0,1,1) + s(1,1,1) \colon t,\, s \in R\} \\ &= \mathcal{L}\big((0,1,1),(1,1,1)\big). \end{aligned}$$

Geometrycznie V_1 i V_2 są różnymi prostymi w przestrzeni R^3 przechodzącymi przez początek układu współrzędnych, a $V_1 + V_2$ jest płaszczyzną zawierającą te dwie proste,

$$V_1 + V_2 = \{(x, y, z) \in R^3 : y - z = 0\},$$

zobacz rys. 7.16.

Twierdzenie 7.8.1. *Jeśli V_1 i V_2 są podprzestrzeniami przestrzeni wektorowej V, to $V_1 + V_2$ jest podprzestrzenią przestrzeni V.*

Dowód. Z faktu, że $\mathbf{0} \in V_i$ $(i = 1, 2)$ wynika, że $\mathbf{0} = \mathbf{0} + \mathbf{0} \in V_1 + V_2$ i dlatego $V_1 + V_2 \neq \emptyset$. Weźmy teraz dowolne dwa wektory $\mathbf{v}$, $\mathbf{u} \in V_1 + V_2$ i dwa skalary s, $t \in K$. Wobec twierdzenia 7.1.2 wystarczy pokazać, że $s\mathbf{v} + t\mathbf{u} \in V_1 + V_2$. Ponieważ $\mathbf{v}$, $\mathbf{u} \in V_1 + V_2$, więc istnieją $\mathbf{v}_1$, $\mathbf{u}_1 \in V_1$ i $\mathbf{v}_2$, $\mathbf{u}_2 \in V_2$ takie, że $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$ i $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$. Jednocześnie $s\mathbf{v}_1 + t\mathbf{u}_1 \in V_1$ i $s\mathbf{v}_2 + t\mathbf{u}_2 \in V_2$, bo V_1 i V_2 są podprzestrzeniami. Zatem $(s\mathbf{v}_1 + t\mathbf{u}_1) + (s\mathbf{v}_2 + t\mathbf{u}_2) \in V_1 + V_2$ wobec definicji sumy podprzestrzeni. Stąd wynika, że $s\mathbf{v} + t\mathbf{u} \in V_1 + V_2$, bo $s\mathbf{v} + t\mathbf{u} = (s\mathbf{v}_1 + t\mathbf{u}_1) + (s\mathbf{v}_2 + t\mathbf{u}_2)$. □

Część wspólna dwóch podprzestrzeni nigdy nie jest pusta (bo zawsze zawiera ona wektor zerowy) i – podobnie jak w poprzednim twierdzeniu – łatwo dowodzi się, że jest ona podprzestrzenią.

Twierdzenie 7.8.2. *Jeśli V_1 i V_2 są podprzestrzeniami przestrzeni wektorowej V, to $V_1 \cap V_2$ jest podprzestrzenią przestrzeni V.* □

Twierdzenie 7.8.3. *Jeśli V_1 i V_2 są skończenie wymiarowymi podprzestrzeniami przestrzeni wektorowej V, to*

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

Dowód. Niech $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ będzie bazą przestrzeni $V_1 \cap V_2$. Wobec twierdzenia 7.5.3 bazę B można uzupełnić do bazy przestrzeni V_1, jak i do bazy przestrzeni V_2. Niech $(\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b}_1, \ldots, \mathbf{b}_n)$ będzie uzupełnieniem bazy B do bazy przestrzeni V_1, a $(\mathbf{b}_1, \ldots, \mathbf{b}_n, \mathbf{c}_1, \ldots, \mathbf{c}_k)$ – uzupełnieniem bazy B do bazy przestrzeni V_2. Dla dowodu twierdzenia wystarczy teraz pokazać, że $(\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b}_1, \ldots, \mathbf{b}_n, \mathbf{c}_1, \ldots, \mathbf{c}_n)$ jest bazą przestrzeni $V_1 + V_2$, bo wtedy będzie $\dim (V_1 + V_2) = m + n + k = (m + n) + (n + k) - n$ $= \dim V_1 + \dim V_2 - \dim (V_1 \cap V_2)$.

Ponieważ $V_1 = \mathcal{L}(\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b}_1, \ldots, \mathbf{b}_n)$ i $V_2 = \mathcal{L}(\mathbf{b}_1, \ldots, \mathbf{b}_n,\, \mathbf{c}_1, \ldots,\, \mathbf{c}_k)$, więc $V_1 + V_2 = \mathcal{L}(\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b}_1, \ldots, \mathbf{b}_n, \mathbf{c}_1, \ldots, \mathbf{c}_k)$, co oznacza, że wektory $\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b}_1, \ldots, \mathbf{b}_n$, $\mathbf{c}_1, \ldots, \mathbf{c}_k$ generują przestrzeń $V_1 + V_2$. Pozostaje nam uzasadnić, że wektory $\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b}_1, \ldots,\, \mathbf{b}_n, \mathbf{c}_1, \ldots, \mathbf{c}_k$ są liniowo niezależne. Załóżmy zatem, że dla pewnych skalarów s_i, t_i i r_i jest

$$s_1\mathbf{a}_1 + \ldots + s_m\mathbf{a}_m + t_1\mathbf{b}_1 + \ldots + t_n\mathbf{b}_n + r_1\mathbf{c}_1 + \ldots + r_k\mathbf{c}_k = \mathbf{0}.$$

Wtedy wektor $\mathbf{v} = s_1\mathbf{a}_1 + \ldots + s_m\mathbf{a}_m + t_1\mathbf{b}_1 + \ldots + t_n\mathbf{b}_n = -(r_1\mathbf{c}_1 + \ldots + r_k\mathbf{c}_k)$ musi być wektorem zerowym, bo inaczej wektor $\mathbf{v} = s_1\mathbf{a}_1 + \ldots + s_m\mathbf{a}_m + t_1\mathbf{b}_1 + \ldots + t_n\mathbf{b}_n$ należałby do V_1, a wektor $\mathbf{v} = -(r_1\mathbf{c}_1 + \ldots + r_k\mathbf{c}_k)$ do $V_2 - V_1$, co jest niemożliwe. Stąd

$$s_1\mathbf{a}_1 + \ldots + s_m\mathbf{a}_m + t_1\mathbf{b}_1 + \ldots + t_n\mathbf{b}_n = \mathbf{0} \quad \text{i} \quad r_1\mathbf{c}_1 + \ldots + r_k\mathbf{c}_k = \mathbf{0}.$$

Z równości tych i z niezależności wektorów tworzących układy $(\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b}_1, \ldots, \mathbf{b}_n)$ oraz $(\mathbf{c}_1, \ldots, \mathbf{c}_k)$ wynika, że $s_1 = \ldots = s_m = t_1 = \ldots = t_n = 0$ i $r_1 = \ldots = r_k = 0$. To dowodzi, że wektory tworzące układ $(\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b}_1, \ldots, \mathbf{b}_n, \mathbf{c}_1, \ldots, \mathbf{c}_k)$ są liniowo niezależne i to kończy dowód twierdzenia. □

Definicja 7.8.2. Jeśli V_1 i V_2 są podprzestrzeniami przestrzeni wektorowej V takimi, że $V = V_1 + V_2$ i $V_1 \cap V_2 = \{\mathbf{0}\}$, to mówimy, że przestrzeń V jest *sumą prostą* podprzestrzeni V_1 oraz V_2 i piszemy $V = V_1 \oplus V_2$.

$V_1 \oplus V_2$ – suma prosta

Twierdzenie 7.8.4. *Jeśli V_1 i V_2 są skończenie wymiarowymi podprzestrzeniami przestrzeni wektorowej V, to następujące warunki są równoważne:*

(a) *$V = V_1 \oplus V_2$, tj. $V = V_1 + V_2$ i $V_1 \cap V_2 = \{\mathbf{0}\}$;*

(b) *$V = V_1 + V_2$ i $\dim V = \dim V_1 + \dim V_2$;*

(c) *każdy wektor $\mathbf{v} \in V$ można w jeden i tylko w jeden sposób przedstawić w postaci sumy $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$, gdzie $\mathbf{v}_1 \in V_1$ i $\mathbf{v}_2 \in V_2$.*

Dowód. Równoważność warunków (a) i (b) jest oczywista, bo wobec twierdzenia 7.8.3 jest $\dim V = \dim V_1 + \dim V_2$ wtedy i tylko wtedy, gdy $\dim(V_1 \cap V_2) = 0$, tj. wtedy i tylko wtedy, gdy $V_1 \cap V_2 = \{\mathbf{0}\}$.

Wykażemy teraz równoważność warunków (a) i (c). Załóżmy najpierw, że $V = V_1 \oplus V_2$. Wówczas $V_1 \cap V_2 = \{\mathbf{0}\}$. Weźmy teraz dowolny wektor $\mathbf{v}$ z przestrzeni V. Ponieważ $V = V_1 + V_2$, więc istnieją wektory $\mathbf{v}_1 \in V_1$ i $\mathbf{v}_2 \in V_2$ takie, że $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$. Przypuśćmy, że jednocześnie jest $\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2$ dla pewnych $\mathbf{u}_1 \in V_1$ i $\mathbf{u}_2 \in V_2$. Wtedy $\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{u}_1 + \mathbf{u}_2$ i dlatego $\mathbf{v}_1 - \mathbf{u}_1 = \mathbf{u}_2 - \mathbf{v}_2$. Ponieważ $\mathbf{v}_1 - \mathbf{u}_1 \in V_1$ i $\mathbf{u}_2 - \mathbf{v}_2 \in V_2$, więc $\mathbf{v}_1 - \mathbf{u}_1 = \mathbf{u}_2 - \mathbf{v}_2 \in V_1 \cap V_2 = \{\mathbf{0}\}$. Stąd $\mathbf{u}_1 = \mathbf{v}_1$ i $\mathbf{u}_2 = \mathbf{v}_2$, co dowodzi, że rozkład $\mathbf{v}_1 + \mathbf{v}_2$ wektora $\mathbf{v}$ jest jednoznaczny.

Załóżmy teraz, że dla każdego wektora $\mathbf{v} \in V$ istnieją jednoznacznie wyznaczone wektory $\mathbf{v}_1 \in V_1$ i $\mathbf{v}_2 \in V_2$ takie, że $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$. Wtedy $V = V_1 + V_2$ i pozostaje pokazać, że $V_1 \cap V_2 = \{\mathbf{0}\}$. W tym celu weźmy dowolny wektor $\mathbf{u}$ ze zbioru $V_1 \cap V_2$. Ponieważ $\mathbf{u} = \mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} \in V_1 + V_2$, więc wobec jednoznaczności rozkładu mamy $\mathbf{u} = \mathbf{0}$ i to dowodzi, że $V_1 \cap V_2 = \{\mathbf{0}\}$. □

Przykład 167. Weźmy pod uwagę podprzestrzenie $V_1 = \{(x, y, z) \in R^3 : x + y + z = 0\}$ i $V_2 = \mathcal{L}\big((1, 2, 3)\big)$ przestrzeni R^3. Ponieważ $V_1 = \{(x, y, -x-y) \in R^3 : x, y \in R\} = \mathcal{L}\big((1, 0, -1), (0, 1, -1)\big)$ i wektory $(1, 0, -1)$, $(0, 1, -1)$ oraz $(1, 2, 3)$ są liniowo niezależne, więc $V_1 + V_2 = \mathcal{L}\big((1, 0, -1), (0, 1, -1), (1, 2, 3)\big)$ jest całą przestrzenią R^3. Dodatkowo, $V_1 \cap V_2 = \{\mathbf{0}\}$, bo $\dim(V_1 \cap V_2) = \dim V_1 + \dim V_2 - \dim(V_1 + V_2) = 2 + 1 - 3 = 0$. Zatem przestrzeń R^3 jest sumą prostą podprzestrzeni V_1 i V_2, $R^3 = V_1 \oplus V_2$. Wobec twierdzenia 7.8.4 (c) każdy wektor $\mathbf{v} \in R^3$ można w jeden i tylko w jeden sposób przedstawić w postaci sumy wektorów $\mathbf{v}_1 \in V_1$ i $\mathbf{v}_2 \in V_2$. Przykładowo, aby znaleźć stosowny rozkład wektora $\mathbf{v} = (7, 3, 2)$ warto poznać jego współrzędne względem bazy $B = \big((1, 0, -1), (0, 1, -1), (1, 2, 3)\big)$. Rozwiązując równanie $x(1, 0, -1) + y(0, 1, -1) + z(1, 2, 3) = (7, 3, 2)$, stwierdzamy, że $[\mathbf{v}]_B = (5, -1, 2)$. Zatem $\mathbf{v} = \big(5(1, 0, -1) - (0, 1, -1)\big) + 2(1, 2, 3) = (5, -1, -4) + (2, 4, 6)$ i $\mathbf{v}_1 = (5, -1, -4) \in V_1$, $\mathbf{v}_2 = (2, 4, 6) \in V_2$.

Geometrycznie $\mathbf{v}_1$ jest wektorem wodzącym punktu przecięcia się płaszczyzny $x + y + z = 0$ z prostą przechodzącą przez punkt $(7, 3, 2)$ i równoległą do wektora $(1, 2, 3)$, tj. z prostą $(x, y, z) = (7, 3, 2) + t(1, 2, 3)$. Zaś $\mathbf{v}_2$ jest wektorem wodzącym punktu przecięcia się prostej $\mathcal{L}\big((1, 2, 3)\big)$ z płaszczyzną $x + y + z - 12 = 0$, tj. z płaszczyzną przechodzącą przez punkt $(7, 3, 2)$ i równoległą do płaszczyzny $x + y = z = 0$ (zob. rys. 7.17).

Przykład 168. Sumą podprzestrzeni

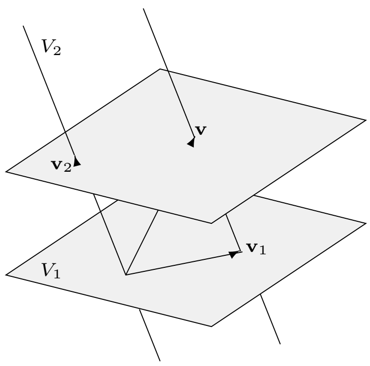
$$V_1 = \mathcal{L}\big((1,-1,0),(1,0,-1)\big) \quad \text{i} \quad V_2 = \mathcal{L}\big((2,1,0),(3,0,1)\big)$$

jest przestrzeń

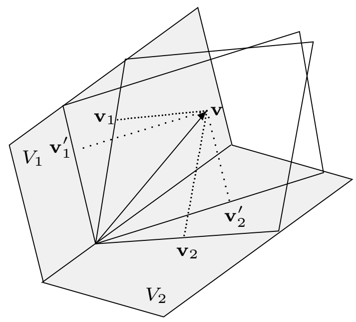
$$V_1 + V_2 = \mathcal{L}\big((1,-1,0),(1,0,-1),(2,1,0),(3,0,1)\big).$$

Ponieważ (dowolne) trzy spośród wektorów $(1,-1,0)$, $(1,0,-1)$, $(2,1,0)$, $(3,\ 0,1)$ są liniowo niezależne w przestrzeni R^3, więc suma $V_1 + V_2$ jest całą przestrzenią R^3. Jednakże, ponieważ $V_1 \cap V_2 \neq \{\mathbf{0}\}$ (bo $\dim(V_1 \cap V_2) = \dim V_1 + \dim V_2 - \dim(V_1 + V_2) = 1$), więc przestrzeń R^3 nie jest sumą prostą podprzestrzeni V_1 i V_2.

Geometrycznie V_1 i V_2 są płaszczyznami przechodzącymi przez początek układu współrzędnych, a $V_1 + V_2$ jest podprzestrzenią zawierającą obie te płaszczyzny, więc $V_1 + V_2 = R^3$. Podprzestrzeń $V_1 \cap V_2$ jest prostą wzdłuż której przecinają się płaszczyzny V_1 i V_2 (zob. rys. 7.18) i – jak łatwo zauważyć – mamy $V_1 \cap V_2 = \mathcal{L}\big((1,-4,3)\big)$.



Rys. 7.17



Rys. 7.18. $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}_1' + \mathbf{v}_2'$

Pojęcie sumy i sumy prostej dwóch podprzestrzeni łatwo uogólnia się na dowolną (skończoną) liczbę podprzestrzeni: jeśli $V_1, V_2, \ldots, V_n$ są podprzestrzeniami przestrzeni V, to zbiór

$$V_1 + V_2 + \ldots + V_n = \{\mathbf{v}_1 + \mathbf{v}_2 + \ldots + \mathbf{v}_n : \mathbf{v}_i \in V_i \ \text{ dla } \ i = 1, \ldots, n\}$$

nazywamy *sumą podprzestrzeni* $V_1, V_2, \ldots, V_n$. Dodatkowo, jeśli $V = V_1 + V_2 + \ldots + V_n$ i $V_i \cap V_j = \{\mathbf{0}\}$, gdy $i, j = 1, \ldots, n$, $i \neq j$, to mówimy, że przestrzeń V jest *sumą prostą* podprzestrzeni $V_1, V_2, \ldots, V_n$ i piszemy $V = V_1 \oplus V_2 \oplus \ldots \oplus V_n$.

Następujące twierdzenie jest łatwym uogólnieniem twierdzenia 7.8.4 dla sumy n podprzestrzeni.

Twierdzenie 7.8.5. *Jeśli $V_1, V_2, \ldots, V_n$ $(n \geqslant 2)$ są skończenie wymiarowymi podprzestrzeniami przestrzeni wektorowej V, to następujące warunki są równoważne:*

(*a*) $V = V_1 \oplus V_2 \oplus \ldots \oplus V_n$, *tj.* $V = V_1 + \ldots + V_n$ *i* $(V_1 + \ldots + V_j) \cap V_{j+1} = \{\mathbf{0}\}$ *dla* $j = 1, \ldots, n-1$*;*

(*b*) $V = V_1 + \ldots + V_n$ *i* $\dim V = \dim V_1 + \ldots + \dim V_n$*;*

(*c*) *każdy wektor* $\mathbf{v} \in V$ *można w jeden i tylko w jeden sposób przedstawić w postaci sumy* $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 + \ldots + \mathbf{v}_n$*, gdzie* $\mathbf{v}_i \in V_i$ *dla* $i = 1, \ldots, n$. $\square$

7.9. Ćwiczenia

1. Dla wektorów $\mathbf{x} = (x_1, \ldots, x_n)$ i $\mathbf{y} = (y_1, \ldots, y_n)$ ze zbioru R^n i skalara $\alpha \in R$, sumę wektorów i iloczyn wektora przez skalar definiujemy wzorami:
$$\begin{aligned}\mathbf{x} + \mathbf{y} &= (x_1 + y_1, \ldots, x_n + y_n);\\ \alpha\mathbf{x} &= (\alpha x_1, \ldots, \alpha x_n).\end{aligned}$$
Pokazać, że R^n z tak określonymi działaniami jest przestrzenią wektorową.
2. Pokazać, że zbiór $K[x]$ wielomianów nad ciałem K z dodawaniem wielomianów i mnożeniem wielomianów przez skalary jest przestrzenią wektorową.
3. Niech R^R będzie zbiorem wszystkich funkcji odwzorowujących zbiór R w zbiór R. Dla $f, g \in R^R$ i $\alpha \in R$, suma $f + g$ oraz iloczyn αf są funkcjami takimi, że
$$(f + g)(x) = f(x) + g(x) \text{ oraz } (\alpha f)(x) = \alpha f(x)$$
dla każdego $x \in R$. Pokazać, że R^R z tak określonymi działaniami jest przestrzenią wektorową.
4. Sprawdzić, czy niżej podane zbiory ze wskazanymi działaniami są przestrzeniami wektorowymi nad ciałem R:
(a) R^2 ze zwykłym mnożeniem wektorów przez skalary i z dodawaniem określonym wzorem
$$(x, y) \oplus (x', y') = (x + x', y + 2y');$$
(b) R^2 ze zwykłym mnożeniem wektorów przez skalary i z dodawaniem określonym wzorem
$$(x, y) \oplus (x', y') = (x - x', y - y');$$
(c) R^2 ze zwykłym mnożeniem wektorów przez skalary i z dodawaniem określonym wzorem
$$(x, y) \oplus (x', y') = (0, 0);$$
(d) R^2 ze zwykłym mnożeniem wektorów przez skalary i z dodawaniem określonym wzorem
$$(x, y) \oplus (\overline{x}, \overline{y}) = \left((x^3 + \overline{x}^3)^{1/3}, (y^3 + \overline{y}^3)^{1/3}\right);$$
(e) R^2 ze zwykłym dodawaniem, ale z mnożeniem przez skalary określonym wzorem
$$r \odot (x, y) = (ry, rx);$$
(f) R^2 ze zwykłym dodawaniem, ale z mnożeniem przez skalary określonym wzorem
$$r \odot (x, y) = (rx, r^2 y);$$
(g) $R_+ = \{x \in R : x > 0\}$ z dodawaniem $\oplus$ i mnożeniem $\odot$, gdzie
$$x \oplus y = xy \quad \text{i} \quad r \odot x = x^r$$
dla $x, y \in R_+$ i $r \in R$;
(h) zbiór macierzy $R_{2\times 2}$ ze zwykłym mnożeniem macierzy przez skalary, ale z dodawaniem określonym wzorem
$$A \oplus B = 0$$
dla każdych $A, B \in R_{2\times 2}$;
(i) zbiór funkcji R^R ze zwykłym mnożeniem funkcji przez skalary, ale z dodawaniem określonym wzorem
$$(f \oplus g)(x) = \max\{f(x), g(x)\}.$$
5. W zbiorze $R_{3\times 1}$ określono dodawanie wektorów i mnożenie wekorów przez skalary w taki sposób, że
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x + x' - 1 \\ y + y' \\ z + z' \end{bmatrix}$$
i
$$r \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} rx - r + 1 \\ ry \\ rz \end{bmatrix}.$$
(a) Pokazać, że $R_{3\times 1}$ z tak określonymi działaniami jest przestrzenią wektorową nad ciałem R.
(b) Sprawdzić, czy zbiór
$$S = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in R_{3\times 1} : x + y + z = 1 \right\}$$
jest podprzestrzenią przestrzeni $R_{3\times 1}$ z wyżej określonymi działaniami.
6. Niech $(V, R, +, \cdot)$ będzie przestrzenią wektorową nad ciałem R i niech $\mathbf{p}$ będzie ustalonym wektorem z przestrzeni V. Określamy nowe działanie dodawania $\oplus$ i nowe mnożenie $\otimes$ wektorów przez skalary przyjmując, że
$$\mathbf{x} \oplus \mathbf{y} = \mathbf{x} + \mathbf{y} + \mathbf{p} \quad \text{i} \quad \alpha \otimes \mathbf{x} = \alpha\mathbf{x} + (\alpha - 1)\mathbf{p}$$
dla każdych wektorów $\mathbf{x}, \mathbf{y} \in V$ i każdego skalara $\alpha \in R$. Pokazać, że system algebraiczny $(V, R, \oplus, \otimes)$ jest przestrzenią wektorową.
7. Sprawdzić, który ze zbiorów jest podprzestrzenią danej przestrzeni wektorowej:
(a) $\{(x, -x)\colon x \in R\}$ w R^2;
(b) $\{(x, x - 1)\colon x \in R\}$ w R^2;
(c) $\{(x, y) \in R^2\colon x, y \geqslant 0\}$ w R^2;
(d) $\{(x, y) \in R^2\colon xy \geqslant 0\}$ w R^2;
(e) $\{(x, y, z) \in R^3 : x + y + z = 0\}$ w R^3;
(f) $\{(x, y, z) \in R^3\colon y = x,\ y = 1\}$ w R^3;
(g) $\{(x, y, z, t) \in R^4\colon xy = 0\}$ w R^4;
(h) $\{\mathbf{x} \in R^n\colon \mathbf{xa} = 0\}$ w R^n, gdzie $\mathbf{a} \in R^n$;
(i) $\{(x_1, \ldots, x_n) \in R^n\colon |x_1| = \ldots = |x_n|\}$ w R^n;
(j) $\{(x_n) \in R^N\colon$ ciąg (x_n) jest zbieżny$\}$ w R^N;
(k) $\{(x_n) \in R^N\colon$ ciąg (x_n) jest rosnący$\}$ w R^N;
(l) $\{A \in R_{n\times n}\colon A$ jest odwracalna$\}$ w $R_{n\times n}$;
(m) $\{f \in R^R\colon f$ jest funkcją nieparzystą$\}$ w R^R;
(n) $\{f \in R^R\colon f(0) = 1\}$ w R^R;
(o) $\{f \in R^R\colon f$ jest różniczkowalna na R$\}$ w R^R;
(p) $\{f \in R^R\colon f'' + f = 0\}$ w R^R.

8. Sprawdzić, czy wektor $\mathbf{x}$ jest kombinacją liniową wektorów $\mathbf{x}_i$ w danej przestrzeni wektorowej:

   (*a*) $\mathbf{x} = (1,2)$, $\mathbf{x}_1 = (3,1)$, $\mathbf{x}_2 = (1,3)$ w R^2;

   (*b*) $\mathbf{x} = (1,2,1)$, $\mathbf{x}_1 = (1,1,1)$, $\mathbf{x}_2 = (1,1,2)$, $\mathbf{x}_3 = (0,1,1)$ w R^3;

   (*c*) $\mathbf{x} = t^2+3t$, $\mathbf{x}_1 = t^3+3t$, $\mathbf{x}_2 = t^2+t$ i $\mathbf{x}_3 = t^2$ w $R_3[t]$;

   (*d*) $\mathbf{x}=\mathbf{I}_2$, $\mathbf{x}_1 = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$ w $R_{2\times 2}$.

9. Dane są funkcje $f(x) = 0$, $g(x) = 1$, $h(x) = \sin x$, $k(x) = 1 + x$ i $l(x) = \cos 2x$ z przestrzeni R^R. Która z nich należy do podprzestrzeni $\mathcal{L}(\cos^2 x, \sin^2 x)$?

10. Pokazać, że jeśli $\mathbf{b} \in \mathcal{L}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ i $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \subseteq \mathcal{L}(\mathbf{x}_1, \ldots, \mathbf{x}_k)$, to $\mathbf{b} \in \mathcal{L}(\mathbf{x}_1, \ldots, \mathbf{x}_k)$.

11. Pokazać, że $\mathcal{L}(\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{y}) = \mathcal{L}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ wtedy i tylko wtedy, gdy $\mathbf{y} \in \mathcal{L}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$.

12. Wykazać, że $S \subseteq \mathcal{L}(S)$ i $\mathcal{L}(S) = \mathcal{L}\big(\mathcal{L}(S)\big)$ dla każdego zbioru wektorów S z przestrzeni wektorowej.

13. Sprawdzić następujące równości:

   (*a*) $\mathcal{L}(\mathbf{x}, \mathbf{y}) = \mathcal{L}(\mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y})$;

   (*b*) $\mathcal{L}(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \mathcal{L}(\mathbf{x}, \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}+\mathbf{z})$;

   (*c*) $\mathcal{L}(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \mathcal{L}(\mathbf{x}+\mathbf{y}, \mathbf{y}+\mathbf{z}, \mathbf{x}+\mathbf{z})$;

   (*d*) $\mathcal{L}\big((1,2,1),(4,1,2)\big) = \mathcal{L}\big((2,3,0),(3,1,1)\big)$;

   (*e*) $\mathcal{L}(t^2+1, t^2-1, t^2+t) = \mathcal{L}(t, t^2+1)$;

   (*f*) $\{\mathbf{x} \in R^4\colon\ \mathbf{x}\mathbf{a} = 0,\ \mathbf{x}\mathbf{b} = 0\} = \mathcal{L}\big((1,0,1,0), (2,1,0,1)\big)$, gdzie $\mathbf{a} = (1,1,-2,-3)$ i $\mathbf{b} = (3,-2,0,-4)$.

14. Zbadać liniową niezależność podanych wektorów $\mathbf{x}_i$ w danej przestrzeni wektorowej:

   (*a*) $\mathbf{x}_1 = (1,3)$, $\mathbf{x}_2 = (-2,-6)$ w R^2;

   (*b*) $\mathbf{x}_1 = (1,3)$, $\mathbf{x}_2 = (2,-6)$ w R^2;

   (*c*) $\mathbf{x}_1 = (-1,2,1)$, $\mathbf{x}_2 = (2,-4,3)$ w R^3;

   (*d*) $\mathbf{x}_1 = (1,3,2)$, $\mathbf{x}_2 = (2,5,3)$, $\mathbf{x}_3 = (4,0,1)$ w R^3;

   (*e*) $\mathbf{x}_1 = (-2,3,1)$, $\mathbf{x}_2 = (3,-1,2)$, $\mathbf{x}_3 = (1,\ 2,3)$, $\mathbf{x}_4 = (1,1,1)$ w R^3;

   (*f*) $\mathbf{x}_1 = x^2-1$, $\mathbf{x}_2 = x^2+1$, $\mathbf{x}_3 = x$, $\mathbf{x}_4 = 2x+1$ w $R[x]$;

   (*g*) $\mathbf{x}_1 = 1$, $\mathbf{x}_2 = \sin^2 x$, $\mathbf{x}_3 = \cos^2 x$, $\mathbf{x}_4 = \cos 2x$ w R^R.

15. W przestrzeni R^R zbadać liniową niezależność następujących układów funkcji rzeczywistych: (*a*) $(1, 2\sin^2 x,\ 3\cos^2 x)$; (*b*) $(1,\ \sin x,\ \sin 2x)$; (*c*) $(1,\ x^2+2x,\ (x+1)^2)$.

16. Niech A i B będą niepustymi zbiorami wektorów przestrzeni wektorowej V. Udowodnić, że: (*a*) jeśli $\mathbf{0} \in A$, to A jest liniowo zależny; (*b*) jeśli A jest liniowo zależny i $A \subseteq B$, to B też jest liniowo zależny; (*c*) jeśli B jest liniowo niezależny i $A \subseteq B$, to A też jest liniowo niezależny.

17. Niech A będzie liniowo niezależnym zbiorem wektorów przestrzeni wektorowej V i niech $\mathbf{v} \in V$. Pokazać, że zbiór $A \cup \{\mathbf{v}\}$ jest liniowo zależny wtedy i tylko wtedy, gdy $\mathbf{v} \in \mathcal{L}(A)$.

18. Pokazać, że jeśli $\mathbf{x}_1, \ldots, \mathbf{x}_n$ są wektorami w przestrzeni V, to wektory $\mathbf{x}_1-\mathbf{x}_2, \mathbf{x}_2-\mathbf{x}_3, \ldots, \mathbf{x}_{n-1}-\mathbf{x}_n, \mathbf{x}_n-\mathbf{x}_1$ są liniowo zależne.

19. Wykazać, że wektory $\mathbf{x}_1, \ldots, \mathbf{x}_n$ są liniowo niezależne w przestrzeni V wtedy i tylko wtedy, gdy wektory $\mathbf{x}_1, \mathbf{x}_1+\mathbf{x}_2, \ldots, \mathbf{x}_1+\mathbf{x}_2+\ldots+\mathbf{x}_n$ są liniowo niezależne w V.

20. Uzasadnić, że wektory $(1,\alpha,\alpha^2)$, $(1,\beta,\beta^2)$, $(1,\gamma,\gamma^2)$ są liniowo niezależne, gdy α, β i γ są różnymi liczbami rzeczywistymi.

21. Dobrać liczbę m tak, aby wektory $(1,2,3,1)$, $(0,3,1,2)$, $(1,0,3,4)$ i $(2,5,0,m)$ były liniowo zależne w R^4.

22. Wskazać te parametry t, dla których wektory $(1,-1,0,0)$, $(1,1,1,0)$, $(1,0,0,1)$ i $(t,1,1,2t)$ są liniowo niezależne.

23. Czy istnieje taka liczba m, że wektory $(1,0,1)$, $(2,m,3)$ i $(1,-m,0)$ są liniowo niezależne?

24. Sprawdzić, czy dany układ wektorów jest bazą przestrzeni:

   (*a*) $\big((1,1),(2,3)\big)$ w R^2;

   (*b*) $\big((j,j),(0,j)\big)$ w $\mathcal{C}^2$;

   (*c*) $\big((2,3,0),(3,2,1)\big)$ w $L\big((6,4,2),(-2,2,-2)\big)$;

   (*d*) $\big(x+1, x^2-1, (x+1)^2\big)$ w $R_2[x]$.

25. Znaleźć takie wektory $\mathbf{v}$ i $\mathbf{u}$, aby układ $(\mathbf{a}, \mathbf{b}, \mathbf{v}, \mathbf{u})$ był bazą przestrzeni R^4, gdy $\mathbf{a} = (1,1,0,0)$ i $\mathbf{b} = (1,0,1,0)$.

26. Znaleźć bazę i wymiar przestrzeni generowanej przez wektory:

   (*a*) $\mathbf{a}_1 = (1,3,2)$, $\mathbf{a}_2 = (2,2,-1)$, $\mathbf{a}_3 = (1,7,7)$;

   (*b*) $\mathbf{b}_1 = (1,0,1,-1)$, $\mathbf{b}_2 = (2,1,1,0)$, $\mathbf{b}_3 = (1,1,1,1)$, $\mathbf{b}_4 = (1,2,3,4)$, $\mathbf{b}_5 = (0,1,2,3)$.

27. Niech W_1 i W_2 będą podprzestrzeniami przestrzeni wektorowej V. Wykazać, że $W_1 \cap W_2$ i $W_1 + W_2 = \{\mathbf{x}+\mathbf{y} : \mathbf{x} \in W_1 \text{ i } \mathbf{y} \in W_2\}$ są podprzestrzeniami przestrzeni V. Pokazać, że $W_1 \cup W_2$ jest podprzestrzenia przestrzeni V wtedy i tylko wtedy, gdy $W_1 \subseteq W_2$ lub $W_2 \subseteq W_1$.

28. Niech $\mathbf{x}_1, \ldots, \mathbf{x}_m$ oraz $\mathbf{y}_1, \ldots, \mathbf{y}_n$ będą wektorami z przestrzeni wektorowej V. Wykazać, że $\mathcal{L}(\mathbf{x}_1, \ldots, \mathbf{x}_m) + \mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_n) = \mathcal{L}(\mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{y}_1, \ldots, \mathbf{y}_n)$.

29. Znaleźć współrzędne wektora $\mathbf{v} = (1,1,1)$ względem baz $B = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ i $B' = (\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3)$ przestrzeni R^3, gdy $\mathbf{b}_1 = (1,2,1), \mathbf{b}_2 = (2,3,3), \mathbf{b}_3 = (3,7,1)$ i $\mathbf{b}'_1 = (3,1,4), \mathbf{b}'_2 = (5,2,1), \mathbf{b}'_3 = (1,1,-6)$.

30. W przestrzeni $R_n[x]$ znaleźć współrzędne wektora $f(x) = a_0 + a_1 x + \ldots + a_n x^n$ względem bazy $B = \big(1, x-1, (x-1)^2, \ldots, (x-1)^n\big)$.

31. Znaleźć macierz przejścia $\mathbf{P}$ od bazy B do bazy B', gdy:

   (*a*) $B = \big((1,2,0),(3,4,2),(2,2,1)\big)$ i $B' = \big((1,-1,2),\ (3,1,-1),(4,0,2)\big)$ w R^3;

   (*b*) $B = (x^3, x^2, x, 1)$ i $B' = (x^3-x^2, x^2-x, x-1, x^3+1)$ w $R_3[x]$.

32. (*a*) Wskazać przykład macierzy kwadratowej $\mathbf{A}$ takiej, że $\mathbf{A} \neq \mathbf{0}$ i $\mathbf{A}^2 = \mathbf{0}$. (*b*) Następnie wykazać, że dla każdej takiej macierzy $\mathbf{A}$ jest $C_{\mathbf{A}} \subseteq N_{\mathbf{A}}$.

33. Zbadać, czy wektor $\mathbf{b}$ należy do podprzestrzeni wierszowej $R_{\mathbf{A}}$ macierzy $\mathbf{A}$, gdy

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & -2 & 1 \\ 1 & 0 & 2 & 4 & 2 \\ 1 & 0 & 2 & 5 & 3 \\ 1 & 0 & 2 & 6 & 4 \\ 1 & 0 & 2 & 6 & 4 \end{bmatrix} \quad \text{i} \quad \mathbf{b} = \begin{bmatrix} -2 \\ 4 \\ 5 \\ 6 \\ 6 \end{bmatrix}.$$

34. Wyznaczyć bazy przestrzeni kolumnowej $C_{\mathbf{A}}$, wierszowej $R_{\mathbf{A}}$ i zerowej $N_{\mathbf{A}}$ macierzy

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & 1 & 5 \\ -2 & -4 & 0 & 4 & -2 \\ 1 & 2 & 2 & 4 & 9 \end{bmatrix}.$$

35. Wyznaczyć rząd macierzy

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 3 & 1 & 2 \\ 2 & 3 & 7 & 4 & 8 \\ 2 & 2 & 6 & 1 & 2 \\ 2 & 3 & 7 & 5 & 10 \end{bmatrix}.$$

36. Czy jest możliwe, że dla rzędu macierzy $\mathbf{A}+\mathbf{B}$ jest $\mathrm{r}(\mathbf{A}+\mathbf{B}) = \mathrm{r}(\mathbf{A}) + \mathrm{r}(\mathbf{B})$?
37. Niech $\mathbf{A}$ będzie macierzą ze zbioru $K_{n\times n}$. Wykazać, że $\mathrm{r}(\mathbf{A}) \geqslant \mathrm{r}(\mathbf{A}^2)$. Wskazać przykład macierzy $\mathbf{A} \in R_{2\times 2}$ takiej, że $\mathrm{r}(\mathbf{A}) > \mathrm{r}(\mathbf{A}^2)$.
38. Pokazać, że zbiór macierzy symetrycznych (diagonalnych) wymiaru $n \times n$ jest podprzestrzenią przestrzeni $R_{n\times n}$.
39. Niech $\mathbf{A}$ będzie macierzą ze zbioru $R_{n\times n}$ i niech V będzie zbiorem macierzy przemiennych z macierzą $\mathbf{A}$, czyli $V = \{\mathbf{X} \in R_{n\times n}\colon \mathbf{AX} = \mathbf{XA}\}$. (*a*) Udowodnić, że V jest podprzestrzenią w $R_{n\times n}$. (*b*) Wykazać, że zbiór $U = \{a_0\mathbf{I}_n + a_1\mathbf{A} + \ldots + a_m\mathbf{A}^m\colon m \in N,\ a_0, \ldots, a_m \in R\}$ jest podprzestrzenią w $R_{n\times n}$. (*c*) Wykazać, że $U \subseteq V$. (*d*) Wyznaczyć bazy przestrzeni V i U, gdy $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. (Uwzględnić, że $\mathbf{A}^2 = 2\mathbf{A} - \mathbf{I}$).
40. Znaleźć bazę przestrzeni macierzy przemiennych z macierzą $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$.
41. Niech S i T będą niepustymi zbiorami wektorów z przestrzeni wektorowej V. Udowodnić, że

$$\mathcal{L}(S \cup T) = \mathcal{L}(S) + \mathcal{L}(T).$$

42. Pokazać, że przestrzeń R^2 jest sumą prostą podprzestrzeni $\mathcal{L}\big((1,2)\big)$ i $\mathcal{L}\big((1,3)\big)$.
43. Pokazać, że przestrzeń $\mathcal{L}\big((1,2,3)\big) + \mathcal{L}\big((1,2,0)\big)$ jest sumą prostą podprzestrzeni $\mathcal{L}\big((1,2,3)\big)$ i $\mathcal{L}\big((1,2,0)\big)$. Czy przestrzeń R^3 jest sumą prostą podprzestrzeni $\mathcal{L}\big((1,2,3)\big)$ i $\mathcal{L}\big((1,2,0)\big)$?
44. Niech V_1 i V_2 będą podprzestrzeniami w R^3, gdzie $V_1 = \{(x,y,z) \in R^3 : x+y+z = 0\}$ i $V_2 = \{(x,y,z) \in R^3 : x+y-z = 0\}$. Wykazać, że $V_1 + V_2 = R^3$.
45. Wyznaczyć część wspólną (przekrój) $V_1 \cap V_2$ i sumę $V_1 + V_2$ podprzestrzeni $V_1 = \mathcal{L}\big((1,2,-1), (1,0,2), (1,-4,8)\big)$ i $V_2 = \mathcal{L}\big((1,0,0), (0,0,1)\big)$ przestrzeni R^3. Czy przestrzeń R^3 jest sumą prostą podprzestrzeni V_1 i V_2?
46. Niech V_1 i V_2 będą podprzestrzeniami przestrzeni R^4, gdzie $V_1 = \mathcal{L}\big((1,\ 1,1,1), (1,2,-1,0)\big)$ i $V_2 = \mathcal{L}\big((-1,1,-1,1),\ (2,1,1,2)\big)$. Wykazać, że $R^4 = V_1 \oplus V_2$.
47. Znaleźć bazę i wymiar przestrzeni $V_1 \cap V_2$ oraz przestrzeni $V_1 + V_2$, gdy:
    (*a*) $V_1 = \mathcal{L}((1,2,0,1), (1,1,1,0))$ i $V_2 = \mathcal{L}((1,0,1,0), (1,3,1,3))$;
    (*b*) $V_1 = \mathcal{L}((1,1,1,1), (1,-1,1,-1), (1,3,1,3))$ i $V_2 = \mathcal{L}((1,2,0,2), (1,2,1,\ 2), (3,1,3,1))$.
48. Niech $\mathbf{a}_1, \ldots, \mathbf{a}_6$ będą kolejnymi kolumnami macierzy

$$\begin{bmatrix} 3 & -1 & 1 & 1 & 8 & 12 \\ 1 & -1 & -1 & 1 & 4 & 6 \\ 1 & 0 & 1 & 1 & 5 & 8 \\ 1 & 1 & 3 & 1 & 6 & 10 \end{bmatrix}.$$

Wyznaczyć bazy przestrzeni V, U, $V+U$ i $V \cap U$, gdy V i U są podprzestrzeniami przestrzeni R^4 takimi, że $V = \mathcal{L}(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ i $U = \mathcal{L}(\mathbf{a}_4, \mathbf{a}_5, \mathbf{a}_6)$.
49. Wykazać, że zbiory $V = \left\{ \begin{bmatrix} x & y \\ -y & x \end{bmatrix} : x, y \in R \right\}$ i $U = \left\{ \begin{bmatrix} x & y \\ y & z \end{bmatrix} : x, y, z \in R \right\}$ są podprzestrzeniami przestrzeni $R_{2\times 2}$. Wskazać także bazy i wymiary przestrzeni V, U, $V \cap U$ oraz $V + U = \{\mathbf{x}+\mathbf{y}\colon \mathbf{x} \in V \ \text{ i } \ \mathbf{y} \in U\}$.
50. Niech V_1 i V_2 będą podprzestrzeniami przestrzeni R^n, gdzie $V_1 = \{(x_1, \ldots, x_n) \in R^n : x_1 + \ldots + x_n = 0\}$ i $V_2 = \{(x_1, \ldots, x_n) \in R^n : x_1 = \ldots = x_n\}$. Wykazać, że $R^n = V_1 \oplus V_2$.
51. Pokazać, że przestrzeń wektorowa $R_n[x]$ ($n \geqslant 1$) jest sumą prostą podprzestrzeni $V_1 = \{\varphi(x) \in R_n[x] : \varphi(-x) = -\varphi(x)$ dla każdego $x \in R\}$ i $V_2 = \{\varphi(x) \in R_n[x] : \varphi(-x) = \varphi(x)$ dla każdego $x \in R\}$.
52. Wykazać, że dla każdej podprzestrzeni V_1 skończenie wymiarowej przestrzeni V istnieje podprzestrzeń V_2 taka, że $V = V_1 \oplus V_2$.
53. Pokazać, że n-wymiarowa przestrzeń wektorowa V jest sumą prostą n podprzestrzeni, gdy $n \geqslant 2$.
54. Niech $(\mathbf{a}_1, \ldots, \mathbf{a}_m)$ i $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ będą odpowiednio bazami podprzestrzeni V_1 i V_2 przestrzeni wektorowej V. Udowodnić, że przestrzeń $V_1 + V_2$ jest sumą prostą podprzestrzeni V_1 i V_2 wtedy i tylko wtedy, gdy $(\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{b}_1, \ldots, \mathbf{b}_n)$ jest bazą przestrzeni $V_1 + V_2$.
55. Udowodnić twierdzenie 7.8.5.
56. Wpisując TAK albo NIE, stwierdzić prawdziwość każdego z następujących zdań:

$\boxed{1}$ Zbiór składający się z wektora zerowego przestrzeni V jest podprzestrzenią przestrzeni wektorowej V.

$\boxed{2}$ Każda przestrzeń wektorowa ma co najmniej dwie różne podprzestrzenie.

$\boxed{3}$ Każda przestrzeń wektorowa, w której jest niezerowy wektor, ma co najmniej dwie różne podprzestrzenie.

4 Jeśli S i T są zbiorami wektorów z przestrzeni wektorowej V, to zawsze $\mathcal{L}(S \cap T) \subseteq \mathcal{L}(S) \cap \mathcal{L}(T)$.

5 Jeśli $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ jest podzbiorem przestrzeni wektorowej, to $\mathbf{v}_i \in \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ dla każdego $i \in \{1, \ldots, n\}$.

6 Jeśli $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ jest podzbiorem przestrzeni wektorowej, to $\mathbf{v}_i + \mathbf{v}_j \in \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ dla każdych $i, j \in \{1, \ldots, n\}$.

7 Jeśli $\mathbf{v} + \mathbf{u}$ należy do podprzestrzeni W przestrzeni V, to $\mathbf{v}$ i $\mathbf{u}$ należą do W.

8 Część wspólna dwóch podprzestrzeni przestrzeni wektorowej V może być zbiorem pustym.

9 Każda prosta w R^2 jest podprzestrzenią przestrzeni R^2 generowaną przez jeden wektor.

10 Każda prosta przechodząca przez początek układu w R^2 jest podprzestrzenią przestrzeni R^2 generowaną przez jeden wektor.

11 Każdy zbiór generujący przestrzeń wektorową R^2 zawiera co najwyżej dwa wektory.

12 Przestrzeń R^2 jest podprzestrzenią przestrzeni R^3.

13 Jeśli wiersze macierzy kwadratowej $\mathbf{A}$ są liniowo zależne, to $\det \mathbf{A} = 0$.

14 Czy układ wektorów $B = ((0, 1, 0, 1), (1, 1, 1, 1))$ jest bazą przestrzeni $W = \{(x, y, z, t) \in R^4 : x = z \text{ i } y = t\}$?

15 Jeśli wektory $\mathbf{a}$, $\mathbf{b}$ i $\mathbf{c}$ są liniowo niezależne w przestrzeni wektorowej V, to podprzestrzenie $\mathcal{L}(\mathbf{a}, \mathbf{b})$ i $\mathcal{L}(\mathbf{b}, \mathbf{c})$ są izomorficzne.

16 Jeśli $\mathbf{A}$ jest macierzą kwadratową i $\mathbf{A}^4 = \mathbf{0}$, to $\mathbf{A}$ jest macierzą nieosobliwą.

17 Jeśli $\mathbf{A}$ jest rzeczywistą macierzą kwadratową i $\mathbf{A}^2 = 8\mathbf{A}^{-1}$, to $\det \mathbf{A} = 2$.

18 Jeśli $\mathbf{A}$ jest macierzą kwadratową i $\mathbf{A}^3 - \mathbf{A} = \mathbf{0}$, to $\det \mathbf{A} = 0$.

19 Jeśli $\mathbf{A}$ jest macierzą kwadratową stopnia n i $\mathbf{A}^T = 4\mathbf{A}^{-1}$, to $\det \mathbf{A} = 2^n$ lub $\det \mathbf{A} = -2^n$.

Rozdział 8

PRZEKSZTAŁCENIA LINIOWE

8.1. Definicja przekształcenia liniowego

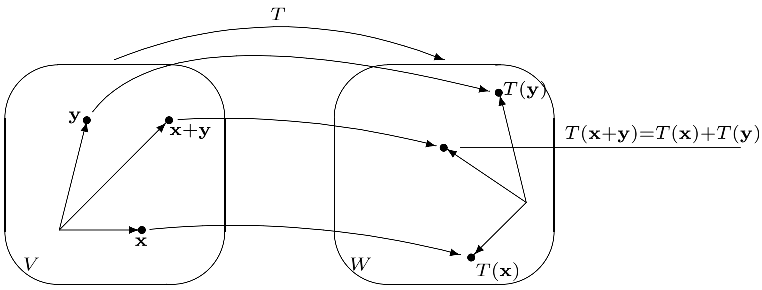
Definicja 8.1.1. Niech V i W będą przestrzeniami wektorowymi nad tym samym ciałem K. Funkcję $T\colon V \to W$ nazywamy *przekształceniem* (*odwzorowaniem* lub *homomorfizmem*) *liniowym* przestrzeni V w przestrzeń W, jeżeli dla każdych wektorów $\mathbf{x}, \mathbf{y} \in V$ i każdego skalara $\alpha \in K$ spełnione są warunki: Przekształcenie liniowe

(a) $T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$, (addytywność przekształcenia)

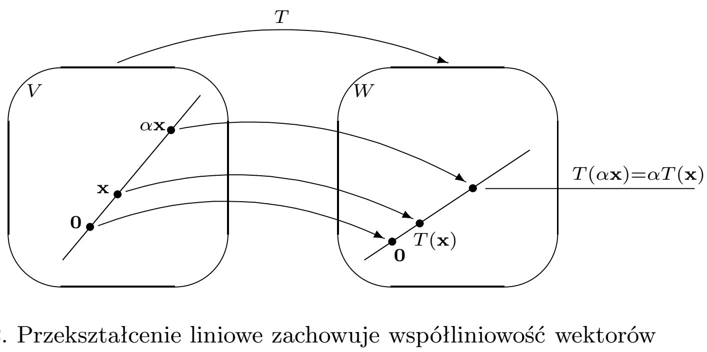
(b) $T(\alpha\mathbf{x}) = \alpha T(\mathbf{x})$. (jednorodność przekształcenia)

Przekształcenie liniowe $T\colon V \to V$ (przestrzeni V w przestrzeń V) nazywa się *operatorem liniowym* na przestrzeni V lub *endomorfizmem* przestrzeni V. Operator liniowy
Zbiór wszystkich przekształceń liniowych przestrzeni V w przestrzeń W oznaczać będziemy symbolem $L(V,W)$, a zbiór wszystkich operatorów na przestrzeni V – symbolem $L(V)$. $L(V,W)$ $L(V) = L(V,V)$

Własność (a) przekształcenia liniowego T powiada, że wektor $T(\mathbf{x}+\mathbf{y})$, który otrzymujemy dodając $\mathbf{x}$ i $\mathbf{y}$ w przestrzeni V i następnie wyznaczając obraz sumy $\mathbf{x}+\mathbf{y}$ poprzez przekształcenie T, jest identyczny z wektorem, jaki otrzymamy wyznaczając najpierw obrazy wektorów $\mathbf{x}$ i $\mathbf{y}$ poprzez przekształcenie T i następnie dodając $T(\mathbf{x})$ i $T(\mathbf{y})$ w przestrzeni W. Własność (b) gwarantuje, że wektor $T(\alpha\mathbf{x})$, obraz iloczynu wektora $\mathbf{x}$ przez skalar α, jest taki sam jak iloczyn $\alpha T(\mathbf{x})$ obrazu $T(\mathbf{x})$ (wektora $\mathbf{x}$) i skalara α, zob. rys. 8.1 i 8.2.



Rys. 8.1. Przekształcenie liniowe zachowuje dodawanie wektorów



Rys. 8.2. Przekształcenie liniowe zachowuje współliniowość wektorów

Z własności (a) i (b) przekształcenia liniowego łatwo wynika, że jeśli V i W są przestrzeniami wektorowymi nad ciałem K, to funkcja $T\colon V \to W$ jest przekształceniem liniowym wtedy i tylko wtedy, gdy spełniony jest jeden z następujących trzech warunków:

(c) Dla każdych wektorów $\mathbf{x}$, $\mathbf{y} \in V$ i każdego skalara $\alpha \in K$ jest

$$T(\mathbf{x} + \alpha\mathbf{y}) = T(\mathbf{x}) + \alpha T(\mathbf{y}). \tag{8.1}$$

Przekształcenie liniowe zachowuje kombinacje liniowe dwóch wektorów

(d) Dla każdych wektorów $\mathbf{x}$, $\mathbf{y} \in V$ i każdych skalarów α, $\beta \in K$ jest

$$T(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha T(\mathbf{x}) + \beta T(\mathbf{y}); \tag{8.2}$$

Przekształcenie liniowe zachowuje kombinacje liniowe skończonej liczby wektorów

(e) Dla każdych wektorów $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$ i skalarów $\alpha_1, \dots, \alpha_n \in K$ jest

$$T\Big(\sum_{i=1}^{n} \alpha_i \mathbf{x}_i\Big) = \sum_{i=1}^{n} \alpha_i T(\mathbf{x}_i). \tag{8.3}$$

Z powyższych własności oraz z twierdzenia 7.1.1 wynikają następujące ważne i naturalne własności przekształcenia liniowego.

Przekształcenie liniowe przekształca wektor zerowy w wektor zerowy

(f) Jeśli $T\colon V \to W$ jest przekształceniem liniowym, to

$$T(\mathbf{0}) = \mathbf{0} \tag{8.4}$$

i

$$T(\mathbf{x} - \mathbf{y}) = T(\mathbf{x}) - T(\mathbf{y}) \tag{8.5}$$

dla każdych dwóch wektorów $\mathbf{x}$, $\mathbf{y} \in V$.[1]

Przykład 169. Jeśli $\mathbf{A}$ jest macierzą wymiaru $m \times n$ i jej elementy należą do ciała K, to symbolem $T_{\mathbf{A}}$ oznaczać będziemy odwzorowanie $T_{\mathbf{A}}\colon K^n \to K^m$ takie, że

$$T_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\cdot\mathbf{x} \tag{8.6}$$

dla każdego $\mathbf{x} \in K^n$. Twierdzimy, że $T_{\mathbf{A}}$ jest przekształceniem liniowym (przestrzeni wektorowej K^n w przestrzeń wektorową K^m).

Z własności iloczynu macierzy mamy

$$T_{\mathbf{A}}(\mathbf{x} + \mathbf{y}) = \mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y} = T_{\mathbf{A}}(\mathbf{x}) + T_{\mathbf{A}}(\mathbf{y})$$

i

$$T_{\mathbf{A}}(\alpha\mathbf{x}) = \mathbf{A}(\alpha\mathbf{x}) = \alpha(\mathbf{A}\mathbf{x}) = \alpha T_{\mathbf{A}}(\mathbf{x})$$

dla każdych $\mathbf{x}$, $\mathbf{y} \in K^n$ i $\alpha \in K$. To dowodzi, że funkcja $T_{\mathbf{A}}$ określona wzorem (8.6) jest przekształceniem liniowym.

Przykład 170. Dana jest funkcja $T\colon R^2 \to R^2$, gdzie

$$T(x_1, x_2) = (x_1 + 2x_2, 3x_1 - x_2)$$

dla $(x_1, x_2) \in R^2$. Niech P będzie prostokątem o wierzchołkach w punktach $A(0,0)$, $B(1,0)$, $C(1,2)$ i $D(0,2)$. Wykazać liniowość przekształcenia T i wyznaczyć obraz $T(P)$ prostokąta P poprzez przekształcenie T.

[1] Własność (8.4) wynika z (b) i z tw. 7.1.1, bo mamy $T(\mathbf{0}) = T(0\mathbf{x}) = 0T(\mathbf{x}) = \mathbf{0}$. Własność (8.5) wynika z (8.1) i tw. 7.1.1, bo $T(\mathbf{x} - \mathbf{y}) = T(\mathbf{x} + (-1)\mathbf{y}) = T(\mathbf{x}) + (-1)T(\mathbf{y}) = T(\mathbf{x}) - T(\mathbf{y})$.

Dla wektorów $\mathbf{x} = (x_1, x_2)$ i $\mathbf{y} = (y_1, y_2)$ z przestrzeni R^2 i skalarów α, $\beta \in R$ mamy

$$\begin{aligned}
T(\alpha\mathbf{x} + \beta\mathbf{y}) &= T(\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2) \\
&= \big((\alpha x_1 + \beta y_1) + 2(\alpha x_2 + \beta y_2), 3(\alpha x_1 + \beta y_1) - (\alpha x_2 + \beta y_2)\big) \\
&= \big(\alpha(x_1 + 2x_2) + \beta(y_1 + 2y_2), \alpha(3x_1 - x_2) + \beta(3y_1 - y_2)\big) \\
&= \big(\alpha(x_1 + 2x_2), \alpha(3x_1 - x_2)\big) + \big(\beta(y_1 + 2y_2), \beta(3y_1 - y_2)\big) \\
&= \alpha(x_1 + 2x_2, 3x_1 - x_2) + \beta(y_1 + 2y_2, 3y_1 - y_2) \\
&= \alpha T(\mathbf{x}) + \beta T(\mathbf{y}),
\end{aligned}$$

więc T jest przekształceniem liniowym. Do tego samego stwierdzenia dochodzimy korzystając z poprzedniego przykładu i obserwując, że $T(\mathbf{x})$ jest iloczynem pewnej macierzy $\mathbf{A} \in R_{2\times 2}$ i wektora $\mathbf{x} = (x_1, x_2)$. Istotnie, mamy

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

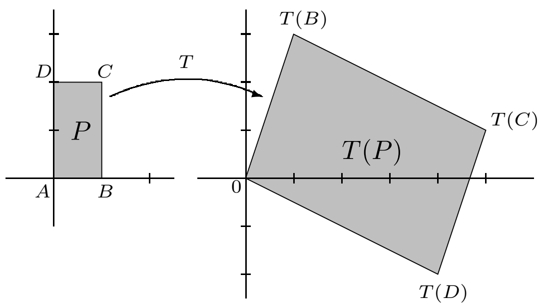
czyli $T(\mathbf{x}) = T_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x}$, gdzie $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$.

Zauważmy, że czworokąt P jest równoległobokiem zbudowanym na wektorach $\overrightarrow{AB} = (1,0)$ i $\overrightarrow{AD} = (0,2)$, czyli $P = \{\alpha(1,0) + \beta(0,2)\colon\ 0 \leqslant \alpha,\ \beta \leqslant 1\}$, a jego obrazem jest zbiór

$$\begin{aligned}
T(P) &= \{T(\mathbf{x})\colon\ \mathbf{x} \in P\} = \{T\big(\alpha(1,0) + \beta(0,2)\big)\colon\ 0 \leqslant \alpha,\ \beta \leqslant 1\} \\
&= \{\alpha T(1,0) + \beta T(0,2)\colon\ 0 \leqslant \alpha,\ \beta \leqslant 1\} \\
&= \{\alpha(1,3) + \beta(4,-2)\colon\ 0 \leqslant \alpha,\ \beta \leqslant 1\}
\end{aligned}$$

i jest to równoległobok zbudowany na wektorach $(1,3) = T(1,0)$ i $(4,-2) = T(0,2)$, zob. rys. 8.3. Warto zauważyć, że pole równoległoboku $T(P)$ jest iloczynem pola równoległoboku P i wartości bezwzględnej wyznacznika macierzy $\mathbf{A}$,

$$|T(P)| = 14 = 7 \cdot 2 = |\det \mathbf{A}| \cdot |P|.$$



Rys. 8.3

Przykład 171. Sprawdzić, czy przekształcenie $T_i\colon\ R^3 \to R^3$ jest liniowe, gdy:

(*a*) $T_1(x_1, x_2, x_3) = (x_1, 2x_2 - x_3, x_1 - 3x_2 + 1)$;

(*b*) $T_2(x_1, x_2, x_3) = (x_2 + x_3, x_1x_2, 6x_3)$;

(*c*) $T_3(x_1, x_2, x_3) = (-x_1, x_1 + x_2^2, x_1 - x_3)$.

(*a*) Zauważmy, że $T_1(\mathbf{0}) = T_1(0,0,0) = (0,0,1) \neq (0,0,0) = \mathbf{0}$, więc z (8.4) wynika, że T_1 nie jest przekształceniem liniowym.

(b) Jeśli $\mathbf{x} = (x_1, x_2, x_3)$ i $\mathbf{y} = (y_1, y_2, y_3)$ są wektorami z przestrzeni R^3, to mamy

$$\begin{aligned}
T_2(\mathbf{x}+\mathbf{y}) &= T_2(x_1+y_1, x_2+y_2, x_3+y_3)\\
&= \big((x_2+y_2)+(x_3+y_3), (x_1+y_1)(x_2+y_2), 6(x_3+y_3)\big)\\
&= (x_2+x_3, x_1x_2, 6x_3) + (y_2+y_3, y_1y_2, 6y_3) + (0, x_1y_2+x_2y_1, 0)\\
&= T_2(\mathbf{x}) + T_2(\mathbf{y}) + (0, x_1y_2+x_2y_1, 0)\\
&\neq T_2(\mathbf{x}) + T_2(\mathbf{y}) \ \ (\text{gdy } x_1y_2+x_2y_1 \neq 0)
\end{aligned}$$

i stąd wynika, że T_2 nie jest przekształceniem liniowym.

(c) Jeśli $\mathbf{x} = (x_1, x_2, x_3)$ jest wektorem z przestrzeni R^3 i α jest liczbą rzeczywistą, to

$$\begin{aligned}
T_3(\alpha\mathbf{x}) &= T_3(\alpha x_1, \alpha x_2, \alpha x_3) = (-\alpha x_1, \alpha x_1 + \alpha^2 x_2^2, \alpha x_1 - \alpha x_3)\\
&= \alpha(-x_1, x_1 + \alpha x_2^2, x_1 - x_3)\\
&\neq \alpha(-x_1, x_1 + x_2^2, x_1 - x_3) \ (\text{gdy } \alpha \neq 1 \text{ i } x_2 \neq 0)\\
&= \alpha\, T_3(\mathbf{x}),
\end{aligned}$$

więc także T_3 nie jest przekształceniem liniowym.

Przykład 172. Niech $C(R)$ i $C'(R)$ będą odpowiednio przestrzenią ciągłych funkcji rzeczywistych i przestrzenią funkcji mających ciągłą pochodną na całym zbiorze R. Odwzorowanie $T\colon\ C'(R) \to C(R)$ takie, że $T(f) = f'$, gdzie f' jest pochodną funkcji f, jest przekształceniem liniowym przestrzeni $C'(R)$ w przestrzeń $C(R)$.

Istotnie, z elementarnych własności pochodnej mamy

$$T(f+g) = (f+g)' = f' + g' = T(f) + T(g)$$

i

$$T(\alpha f) = (\alpha f)' = \alpha f' = \alpha T(f)$$

dla każdych funkcji f, $g \in C'(R)$ i każdej liczby $\alpha \in R$. To dowodzi, że tu rozważana funkcja T jest przekształceniem liniowym.

Ważnymi przykładami przekształceń są przekształcenie zerowe i przekształcenie tożsamościowe.

Definicja 8.1.2. Jeśli V i W są przestrzeniami wektorowymi (nad tym samym ciałem), to *przekształceniem zerowym* przestrzeni V w przestrzeń W nazywamy funkcję $0\colon\ V \to W$ taką, że

Przekształcenie zerowe

$$0(\mathbf{x}) = \mathbf{0} \tag{8.7}$$

dla każdego $\mathbf{x} \in V$. Natomiast *przekształceniem tożsamościowym* (lub *przekształceniem identycznościowym*) przestrzeni V nazywamy funkcję $I_V\colon\ V \to V$ taką, że

Przekształcenie tożsamościowe

$$I_V(\mathbf{x}) = \mathbf{x} \tag{8.8}$$

dla każdego $\mathbf{x} \in V$.

Łatwo pokazuje się, że przekształcenie zerowe oraz przekształcenie tożsamościowe są przekształceniami liniowymi.

W pierwszym twierdzeniu tego rozdziału pokazujemy, że każde przekształcenie liniowe $T\colon\ V \to W$ jest jednoznacznie wyznaczone przez swoje wartości na wektorach bazy przestrzeni V. Praktycznie oznacza to, że jeśli $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ jest bazą przestrzeni V i jeśli znamy wektory $T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)$, to możemy wyznaczyć wartość $T(\mathbf{x})$ przekształcenia T dla każdego wektora $\mathbf{x}$ z przestrzeni V.

Twierdzenie 8.1.1. *Niech V i W będą przestrzeniami wektorowymi nad ciałem K. Jeśli $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ jest bazą przestrzeni V, a $\mathbf{c}_1, \ldots, \mathbf{c}_n$ są dowolnymi wektorami przestrzeni W, to istnieje dokładnie jedno takie przekształcenie liniowe $T\colon V \to W$, że $T(\mathbf{b}_i) = \mathbf{c}_i$ dla $i = 1, \ldots, n$.*

Dowód. Niech $\mathbf{x}$ będzie dowolnym wektorem z przestrzeni V. Ponieważ $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ jest bazą przestrzeni V, więc istnieją jednoznacznie wyznaczone skalary $x_1, \ldots, x_n \in K$ takie, że $\mathbf{x} = x_1\mathbf{b}_1 + \ldots + x_n\mathbf{b}_n$. Definiujemy przekształcenie $T\colon V \to W$, przyjmując, że

$$T(\mathbf{x}) = T(x_1\mathbf{b}_1 + \ldots + x_n\mathbf{b}_n) = x_1\mathbf{c}_1 + \ldots + x_n\mathbf{c}_n. \tag{8.9}$$

Tak określone przekształcenie T odwzorowuje przestrzeń V w przestrzeń W i obrazem każdego wektora $\mathbf{b}_i$ poprzez T jest wektor $\mathbf{c}_i$, bo wobec (8.9) mamy

$$T(\mathbf{b}_i) = T(0\mathbf{b}_1 + \ldots + 1\mathbf{b}_i + \ldots + 0\mathbf{b}_n) = 0\mathbf{c}_1 + \ldots + 1\mathbf{c}_i + \ldots + 0\mathbf{c}_n = \mathbf{c}_i.$$

Dodatkowo, jest to przekształcenie liniowe, bo znowu z (8.9) dla każdych wektorów $\mathbf{x} = x_1\mathbf{b}_1 + \ldots + x_n\mathbf{b}_n$ i $\mathbf{y} = y_1\mathbf{b}_1 + \ldots + y_n\mathbf{b}_n$ z przestrzeni V oraz skalarów α, β z ciała K mamy

$$\begin{aligned}
T(\alpha\mathbf{x} + \beta\mathbf{y}) &= T\big((\alpha x_1 + \beta y_1)\mathbf{b}_1 + \ldots + (\alpha x_n + \beta y_n)\mathbf{b}_n\big) \\
&= (\alpha x_1 + \beta y_1)\mathbf{c}_1 + \ldots + (\alpha x_n + \beta y_n)\mathbf{c}_n \\
&= \alpha(x_1\mathbf{c}_1 + \ldots + x_n\mathbf{c}_n) + \beta(y_1\mathbf{c}_1 + \ldots + y_n\mathbf{c}_n) \\
&= \alpha T(\mathbf{x}) + \beta T(\mathbf{y}).
\end{aligned}$$

W końcu, dla dowodu jedyności T, przypuśćmy, że $U\colon V \to W$ jest przekształceniem liniowym takim, że $U(\mathbf{b}_i) = \mathbf{c}_i$ dla $i = 1, \ldots, n$. Wystarczy teraz udowodnić, że $U = T$. Ponieważ dla każdego wektora $\mathbf{x} = x_1\mathbf{b}_1 + \ldots + x_n\mathbf{b}_n$ z przestrzeni V mamy

$$\begin{aligned}
U(\mathbf{x}) &= U(x_1\mathbf{b}_1 + \ldots + x_n\mathbf{b}_n) \\
&= x_1U(\mathbf{b}_1) + \ldots + x_nU(\mathbf{b}_n) && (\text{z liniowości } U) \\
&= x_1T(\mathbf{b}_1) + \ldots + x_nT(\mathbf{b}_n) && (\text{bo } U(\mathbf{b}_i) = \mathbf{c}_i = T(\mathbf{b}_i)) \\
&= T(x_1\mathbf{b}_1 + \ldots + x_n\mathbf{b}_n) = T(\mathbf{x}), && (\text{z liniowości } T)
\end{aligned}$$

więc $U = T$ i to kończy dowód twierdzenia. □

Przykład 173. Wskazać przekształcenie liniowe $T\colon R^3 \to R^2$, dla którego

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad \text{i} \quad T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Ponieważ wektory

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \mathbf{e}_1, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \mathbf{e}_2 \quad \text{i} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \mathbf{e}_3$$

tworzą bazę przestrzeni R^3 i dla każdego wektora

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in R^3 \quad \text{jest} \quad \mathbf{x} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3,$$

więc wobec żądanej liniowości przekształcenia T musi być

$$\begin{aligned}
T(\mathbf{x}) &= T(x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3) = xT(\mathbf{e}_1) + yT(\mathbf{e}_2) + zT(\mathbf{e}_3) \\
&= x\begin{bmatrix} 1 \\ -1 \end{bmatrix} + y\begin{bmatrix} 2 \\ 3 \end{bmatrix} + z\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 3 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}.
\end{aligned}$$

Przykład 174. Wyznaczyć obraz wektora $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ poprzez przekształcenie liniowe $T\colon R^3 \to R^3$ takie, że $T(\mathbf{b}_i) = \mathbf{c}_i$ $(i = 1, 2, 3)$, gdzie

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{b}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \ \mathbf{b}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \ \mathbf{c}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \ \mathbf{c}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \ \mathbf{c}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Wyznaczając $\mathbf{s}$ z równania $\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 \end{bmatrix} \mathbf{s} = \mathbf{x}$, stwierdzamy, że wektor $\mathbf{x}$ jest kombinacją liniową wektorów $\mathbf{b}_1$, $\mathbf{b}_2$ oraz $\mathbf{b}_3$ i jednocześnie zauważamy, że

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = (x - z)\mathbf{b}_1 + (-x + y + 2z)\mathbf{b}_2 + (x - y - z)\mathbf{b}_3.$$

Stąd, z liniowości przekształcenia T oraz z równości $T(\mathbf{b}_i) = \mathbf{c}_i$ otrzymujemy obraz $T(\mathbf{x})$ wektora $\mathbf{x}$,

$$\begin{aligned}
T(\mathbf{x}) &= T\big((x - z)\mathbf{b}_1 + (-x + y + 2z)\mathbf{b}_2 + (x - y - z)\mathbf{b}_3\big) \\
&= (x - z)T(\mathbf{b}_1) + (-x + y + 2z)T(\mathbf{b}_2) + (x - y - z)T(\mathbf{b}_3) \\
&= (x - z)\mathbf{c}_1 + (-x + y + 2z)\mathbf{c}_2 + (x - y - z)\mathbf{c}_3 \\
&= (x - z)\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + (-x + y + 2z)\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + (x - y - z)\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} y + z \\ z \\ -y - z \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & -1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.
\end{aligned}$$

Wniosek 8.1.1. *Niech V i W będą przestrzeniami wektorowymi i niech $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ będzie bazą przestrzeni V. Jeśli $U, T\colon V \to W$ są przekształceniami liniowymi i $U(\mathbf{b}_i) = T(\mathbf{b}_i)$ $(i = 1, \dots, n)$, to $U = T$.* $\Box$

Przykład 175. Niech $T\colon R^2 \to R^3$ będzie funkcją taką, że $T(x, y) = (x - y, 0, 3x)$. Niech $U\colon R^2 \to R^3$ będzie przekształceniem liniowym takim, że

$$U(2, 1) = (1, 0, 6) \quad \text{i} \quad U(0, 1) = (-1, 0, 0).$$

Ponieważ układ wektorów $B = \big((2, 1), (0, 1)\big)$ jest bazą przestrzeni R^2 i przekształcenia liniowe T oraz U przyjmują identyczne wartości na wektorach bazy B,

$$U(2, 1) = (1, 0, 6) = T(2, 1) \quad \text{i} \quad U(0, 1) = (-1, 0, 0) = T(0, 1),$$

więc wobec ostatniego wniosku przekształcenia liniowe T oraz U są identyczne, czyli $U = T$.

8.2. Jądro i obraz przekształcenia liniowego

Zajmiemy się teraz dwiema ważnymi podprzestrzeniami związanymi z przekształceniem liniowym — jądrem i obrazem przekształcenia liniowego. Zaczynamy od pokazania, że obraz (i przeciwobraz) podprzestrzeni poprzez przekształcenie liniowe jest podprzestrzenią.

Twierdzenie 8.2.1. *Niech* $T\colon V \to W$ *będzie przekształceniem liniowym i niech* V' *oraz* W' *będą odpowiednio podprzestrzeniami przestrzeni wektorowych* V *i* W*. Wtedy:*

(*a*) $T(V') = \{\, T(\mathbf{x})\colon\ \mathbf{x} \in V'\,\}$ *jest podprzestrzenią przestrzeni* W*;*

(*b*) $T^{-1}(W') = \{\, \mathbf{x} \in V\colon\ T(\mathbf{x}) \in W'\,\}$ *jest podprzestrzenią przestrzeni* V*.*

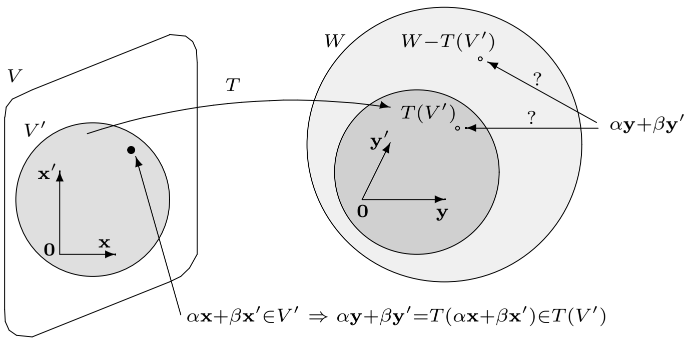
V'−podprzestrzeń w V / $T(V')$− podprzestrzeń w W

W'−podprzestrzeń w W / $T^{-1}(W')$−podprzestrzeń w V

Dowód. Zbiór $T(V')$ jest niepusty, bo $\mathbf{0} \in V'$ i $\mathbf{0} = T(\mathbf{0}) \in T(V')$. Weźmy teraz dowolne wektory $\mathbf{y}$, $\mathbf{y}' \in T(V')$ i skalary α, $\beta \in K$. Wobec twierdzenia 7.1.3 dla dowodu części (*a*) wystarczy pokazać, że $\alpha\mathbf{y} + \beta\mathbf{y}' \in T(V')$, zob. rys. 8.4. Niech $\mathbf{x}$, $\mathbf{x}' \in V'$ będą takie, że $T(\mathbf{x}) = \mathbf{y}$ i $T(\mathbf{x}') = \mathbf{y}'$. Wtedy $\alpha\mathbf{x} + \beta\mathbf{x}' \in V'$, bo V' jest podprzestrzenią. Stąd i z liniowości przekształcenia T wynika, że

$$\alpha\mathbf{y} + \beta\mathbf{y}' = \alpha T(\mathbf{x}) + \beta T(\mathbf{x}') = T(\alpha\mathbf{x} + \beta\mathbf{x}') \in T(V').$$

Równie łatwo dowodzi się drugą część twierdzenia. Ponieważ $T(\mathbf{0}) = \mathbf{0} \in W'$, więc $\mathbf{0} \in T^{-1}(W')$ i dlatego zbiór $T^{-1}(W')$ jest niepusty. Niech $\mathbf{x}$, $\mathbf{x}' \in T^{-1}(W')$ i α, $\beta \in K$. Dla dowodu (*b*) wystarczy pokazać, że $\alpha\mathbf{x} + \beta\mathbf{x}' \in T^{-1}(V')$. Ponieważ $\mathbf{x}$, $\mathbf{x}' \in T^{-1}(W')$, więc $T(\mathbf{x})$, $T(\mathbf{x}') \in W'$ i dlatego $\alpha T(\mathbf{x}) + \beta T(\mathbf{x}') \in W'$ (bo W' jest podprzestrzenią). Stąd i z liniowości przekształcenia T wynika, że $T(\alpha\mathbf{x} + \beta\mathbf{x}') \in W'$ i dlatego $\alpha\mathbf{x} + \beta\mathbf{x}' \in T^{-1}(W')$. $\square$



Rys. 8.4

Definicja 8.2.1. Niech V i W będą przestrzeniami wektorowymi (nad ciałem K) i niech $T\colon V \to W$ będzie przekształceniem liniowym. Wtedy zbiór

$$\operatorname{Im} T = T(V) = \{\, T(\mathbf{x})\colon\ \mathbf{x} \in V\} \tag{8.10}$$

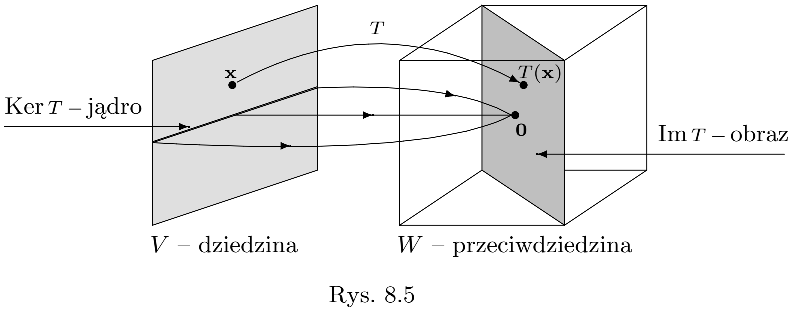
Obraz przekształcenia

(będący obrazem zbioru V poprzez przekształcenie T) nazywamy *obrazem przekształcenia* T*.* Natomiast zbiór

$$\operatorname{Ker} T = T^{-1}(\{\mathbf{0}\}) = \{\, \mathbf{x} \in V\colon\ T(\mathbf{x}) = \mathbf{0}\}, \tag{8.11}$$

Jądro przekształcenia

czyli przeciwobraz wektora zerowego, nazywamy *jądrem* (lub *przestrzenią zerową*) *przekształcenia* T*.*



Rys. 8.5

Wniosek 8.2.1. *Jeśli* $T\colon V\to W$ *jest przekształceniem liniowym, to:*

(*a*) $\operatorname{Im} T$ *jest podprzestrzenią przestrzeni* W*;*

(*b*) $\operatorname{Ker} T$ *jest podprzestrzenią przestrzeni* V*.*

Dowód. Oba stwierdzenia wynikają z twierdzenia 8.2.1 odpowiednio dla $V'=V$ i $W' = \{\mathbf{0}\}$. $\square$

Wniosek 8.2.2. *Niech* $B=(\mathbf{b}_1,\ldots,\mathbf{b}_n)$ *będzie bazą przestrzeni wektorowej* V *i niech* $T\colon V\to W$ *będzie przekształceniem liniowym. Wtedy*

$T\big(\mathcal{L}(B)\big)=\mathcal{L}\big(T(B)\big)$

$$\operatorname{Im} T=\mathcal{L}\big(T(\mathbf{b}_1),\ldots,T(\mathbf{b}_n)\big). \tag{8.12}$$

Dowód. Mamy $B\subset V$, więc także $T(B)\subset T(V)$ i $\mathcal{L}\big(T(B)\big)\subseteq\mathcal{L}\big(T(V)\big)$. Jednocześnie $\mathcal{L}\big(T(V)\big)=T(V)$ (bo $T(V)$ wobec wniosku 8.2.1 jest podprzestrzenią) i dlatego $\mathcal{L}\big(T(B)\big)\subseteq T(V)$.

Weźmy teraz $\mathbf{y}\in T(V)$. Niech wektor $\mathbf{x}\in V$ będzie taki, że $\mathbf{y}=T(\mathbf{x})$. Ponieważ B jest bazą przestrzeni V, więc $\mathbf{x}=\sum_{i=1}^n\alpha_i\mathbf{b}_i$ (dla pewnych skalarów $\alpha_1,\ldots,\alpha_n$) i wtedy też z liniowości przekształcenia T mamy

$$\mathbf{y}=T(\mathbf{x})=T\Big(\sum_{i=1}^n\alpha_i\mathbf{b}_i\Big)=\sum_{i=1}^n\alpha_iT(\mathbf{b}_i)\in\mathcal{L}\big(T(B)\big).$$

Stąd $T(V)\subseteq\mathcal{L}\big(T(B)\big)$ i to kończy dowód równości $T(V)=\mathcal{L}\big(T(B)\big)$. $\square$

Przykład 176. Dane jest przekształcenie liniowe $T\colon R_{2\times2}\to R_2[x]$ takie, że

$$T\left(\begin{bmatrix} a & b\\ c & d\end{bmatrix}\right)=a+(b+c)x+dx^2.$$

Wyznaczyć obraz i jądro przekształcenia T.

Bazą przestrzeni $R_{2\times2}$ jest układ $\left(\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}\right)$, więc wobec wniosku 8.2.2 obrazem przekształcenia T jest

$$\begin{aligned}\operatorname{Im} T &= \mathcal{L}\left(T\left(\begin{bmatrix}1&0\\0&0\end{bmatrix}\right),T\left(\begin{bmatrix}0&1\\0&0\end{bmatrix}\right),T\left(\begin{bmatrix}0&0\\1&0\end{bmatrix}\right),T\left(\begin{bmatrix}0&0\\0&1\end{bmatrix}\right)\right)\\ &= \mathcal{L}(1,x,x,x^2)=\mathcal{L}(1,x,x^2)=R_2[x].\end{aligned}$$

Zauważmy teraz, że

$$\begin{aligned}\begin{bmatrix}a&b\\c&d\end{bmatrix}\in\operatorname{Ker}T \quad &\Leftrightarrow\quad T\left(\begin{bmatrix}a&b\\c&d\end{bmatrix}\right)=a+(b+c)x+dx^2\equiv0\\ &\Leftrightarrow\quad a=0,\quad b+c=0\quad\text{i}\quad d=0.\end{aligned}$$

Stąd otrzymujemy

$$\operatorname{Ker}T=\left\{\begin{bmatrix}0&b\\-b&0\end{bmatrix}:\ b\in R\right\}=\mathcal{L}\left(\begin{bmatrix}0&1\\-1&0\end{bmatrix}\right).$$

Przykład 177. Wyznaczyć jądro i obraz przekształcenia $T_{\mathbf{A}}\colon R^5\to R^4$, dla którego

$$T_{\mathbf{A}}(\mathbf{x})=\mathbf{A}\mathbf{x}\quad\text{i}\quad\mathbf{A}=\begin{bmatrix}1&1&1&1&-1\\2&1&1&0&1\\1&0&0&1&0\\3&2&2&-1&2\end{bmatrix}.$$

Z definicji jądra mamy

$$\operatorname{Ker} T_{\mathbf{A}} = \{\, \mathbf{x} \in R^5 \colon\ T_{\mathbf{A}}(\mathbf{x}) = \mathbf{0}\,\} = \{\, \mathbf{x} \in R^5 \colon\ \mathbf{A}\mathbf{x} = \mathbf{0}\,\}.$$

To oznacza, że $\operatorname{Ker} T_{\mathbf{A}}$ jest zbiorem rozwiązań jednorodnego układu równań $\mathbf{A}\mathbf{x} = \mathbf{0}$ (i przestrzenią zerową $N_{\mathbf{A}}$ macierzy $\mathbf{A}$). Dla macierzy rozszerzonej układu $\mathbf{A}\mathbf{x} = \mathbf{0}$ mamy wierszowe równoważności

$\operatorname{Ker} T_{\mathbf{A}} = N_{\mathbf{A}}$

$$[\mathbf{A}|\mathbf{0}] = \left[\begin{array}{rrrrr|r} 1 & 1 & 1 & 1 & -1 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 3 & 2 & 2 & -1 & 2 & 0 \end{array}\right] \sim \ldots \sim \left[\begin{array}{rrrrr|r} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Stąd zaś wynika, że

$$\operatorname{Ker} T_{\mathbf{A}} = \left\{ \begin{bmatrix} -x_5 \\ -x_3 + x_5 \\ x_3 \\ x_5 \\ x_5 \end{bmatrix} \colon\ x_3,\, x_5 \in R \right\} = \mathcal{L}\left(\begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right).$$

Obrazem przekształcenia liniowego $T_{\mathbf{A}}$ jest

$$\operatorname{Im} T_{\mathbf{A}} = T_{\mathbf{A}}(R^5) = \{\, T_{\mathbf{A}}(\mathbf{x}) \colon\ \mathbf{x} \in R^5\,\} = \{\, \mathbf{A}\mathbf{x} \colon\ \mathbf{x} \in R^5\,\}$$

$\operatorname{Im} T_{\mathbf{A}} = C_{\mathbf{A}}$

i jest to przestrzeń kolumnowa macierzy $\mathbf{A}$, czyli przestrzeń $C_{\mathbf{A}}$. Korzystając z powyższej wierszowej równoważności, dochodzimy do wniosku, że jest ona generowana przez pierwszą, drugą i czwartą kolumnę macierzy $\mathbf{A}$,

$$\operatorname{Im} T_{\mathbf{A}} = C_{\mathbf{A}} = \mathcal{L}\left(\begin{bmatrix} 1 \\ 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix} \right).$$

Następujące twierdzenie, zwane *twierdzeniem wymiarowym, zasadniczym twierdzeniem algebry liniowej* lub *twierdzeniem Sylvestera,* podaje związek pomiędzy wymiarem jądra i wymiarem obrazu (zob. rys. 8.6) przekształcenia liniowego.[2]

Twierdzenie 8.2.2 (Twierdzenie wymiarowe). *Jeśli $T\colon\ V \to W$ jest przekształceniem liniowym i V jest przestrzenią skończonego wymiaru, to*

Twierdzenie wymiarowe

$$\dim\ \operatorname{Ker} T + \dim \operatorname{Im} T = \dim V. \tag{8.13}$$

Dowód. Niech $(\mathbf{a}_1, \ldots, \mathbf{a}_r)$ i $(\mathbf{b}_1, \ldots, \mathbf{b}_p)$ będą odpowiednio bazą przestrzeni $\operatorname{Ker} T$ i $\operatorname{Im} T$. Niech $\mathbf{b}'_i \in V$ będzie wektorem takim, że $T(\mathbf{b}'_i) = \mathbf{b}_i$ dla $i \in \{1, \ldots, p\}$. Dla dowodu równości (8.13) wystarczy udowodnić, że układ $B = (\mathbf{a}_1, \ldots, \mathbf{a}_r, \mathbf{b}'_1, \ldots, \mathbf{b}'_p)$ jest bazą przestrzeni V. Pokażemy najpierw, że układ B jest liniowo niezależny. W tym celu bierzemy pod uwagę kombinację liniową

$$\sum_{i=1}^{r} x_i \mathbf{a}_i + \sum_{j=1}^{p} y_j \mathbf{b}'_j = \mathbf{0}. \tag{8.14}$$

Wystarczy pokazać, że $x_1 = \ldots = x_r = y_1 \ldots = y_p = 0$. Zauważmy najpierw, że z liniowości T i z faktu, że $\mathbf{a}_i \in \operatorname{Ker} T$ mamy

$$\mathbf{0} = T(\mathbf{0}) = T\Big(\sum_{i=1}^{r} x_i \mathbf{a}_i + \sum_{j=1}^{p} y_j \mathbf{b}'_j \Big) = \sum_{i=1}^{r} x_i T(\mathbf{a}_i) + \sum_{j=1}^{p} y_j T(\mathbf{b}'_j) = \sum_{j=1}^{p} y_j \mathbf{b}_j$$

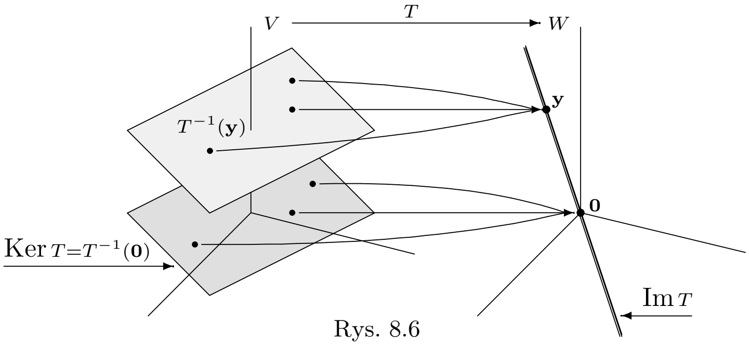
[2] Jeśli $T\colon\ V \to W$ jest przkształceniem liniowym skończenie wymiarowej przestrzeni wektorowej V, to liczby $\dim \operatorname{Ker} T$ i $\dim \operatorname{Im} T$ nazywa się odpowiednio *zerowością* i *rzędem* przekształcenia T.

i dlatego

$$\sum_{j=1}^{p} y_j \mathbf{b}_j = \mathbf{0}. \tag{8.15}$$

Ponieważ wektory $\mathbf{b}_1, \ldots, \mathbf{b}_p$ są liniowo niezależne, więc z równości (8.15) otrzymujemy $y_1 = \ldots = y_p = 0$. Stąd i z równości (8.14) mamy $\sum_{i=1}^{r} x_i \mathbf{a}_i = \mathbf{0}$. Z tej zaś równości i z liniowej niezależności wektorów $\mathbf{a}_1, \ldots, \mathbf{a}_r$ wynika, że $x_1 = \ldots = x_r = 0$.

Pokażemy teraz, że każdy wektor $\mathbf{v} \in V$ jest kombinacją liniową wektorów układu B. Przede wszystkim, ponieważ wektory $\mathbf{b}_1, \ldots, \mathbf{b}_p$ generują przestrzeń $T(V)$ i $T(\mathbf{v}) \in T(V)$, więc istnieją skalary $y_1, \ldots, y_p$ takie, że $T(\mathbf{v}) = \sum_{j=1}^{p} y_j \mathbf{b}_j$. Zauważmy teraz, że wektor $\mathbf{v} - \sum_{j=1}^{p} y_j \mathbf{b}'_j \in \operatorname{Ker} T$ (bo $T(\mathbf{v} - \sum_{j=1}^{p} y_j \mathbf{b}'_j) = T(\mathbf{v}) - \sum_{j=1}^{p} y_j T(\mathbf{b}'_j) = T(\mathbf{v}) - \sum_{j=1}^{p} y_j \mathbf{b}_j = \mathbf{0}$), więc istnieją skalary $x_1, \ldots, x_r$ takie, że $\mathbf{v} - \sum_{j=1}^{p} y_j \mathbf{b}'_j = \sum_{i=1}^{r} x_i \mathbf{a}_i$ i dlatego $\mathbf{v} = \sum_{i=1}^{r} x_i \mathbf{a}_i + \sum_{j=1}^{p} y_j \mathbf{b}'_j$. Z powyższego wynika, że zbiór B jest bazą przestrzeni V i dlatego $\dim V = |B| = r + p = \dim \operatorname{Ker} T + \dim \operatorname{Im} T$. $\square$



Rys. 8.6

Przykład 178. Wyznaczyć wymiar jądra i wymiar obrazu przekształcenia liniowego $T\colon\ R^4 \to R^3$, gdzie

$$T(x, y, z, t) = (x + y + z, y + 2z, x - z).$$

Łatwo zauważyć, że mamy

$$\operatorname{Ker} T = \left\{ \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} \in R^4 \colon \begin{cases} x + y + z = 0 \\ \quad\; y + 2z = 0 \\ x \qquad - z = 0 \end{cases} \right\} = \mathcal{L}\left(\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right).$$

Stąd $\dim \operatorname{Ker} T = 2$ i wobec twierdzenia 8.2.2 jest

$$\dim \operatorname{Im} T = \dim R^4 - \dim \operatorname{Ker} T = 2.$$

Przykład 179. Korzystając z twierdzenia wymiarowego (tw. 8.2.2), uzasadnić, że nie istnieje przekształcenie liniowe $T\colon\ R^3 \to R^4$ takie, że wektory $\mathbf{c}_1 = (1, 1, 2, 1)$, $\mathbf{c}_2 = (1, 2, 3, 1)$, $\mathbf{c}_3 = (3, 2, 7, 5)$ i $\mathbf{c}_4 = (1, 0, 2, 5)$ należą do przestrzeni $\operatorname{Im} T$.

Łatwo sprawdzić, że wektory $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{c}_3$ i $\mathbf{c}_4$ są liniowo niezależne. Stąd wynika, że $\dim \mathcal{L}(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4) = 4$. Zauważmy teraz, że jeśli istniałoby przekształcenie liniowe $T\colon\ R^3 \to R^4$ takie, że $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{c}_3$, $\mathbf{c}_4 \in \operatorname{Im} T$, to wtedy przestrzeń $\mathcal{L}(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4)$ zawierałaby się w przestrzeni $\operatorname{Im} T$ i dlatego byłoby $\dim \operatorname{Im} T \geqslant \dim \mathcal{L}(\mathbf{c}_1, \ldots, \mathbf{c}_4) = 4$. To zaś byłoby niemożliwe, bo z twierdzenia 8.2.2 mamy

$$\dim \operatorname{Im} T \leqslant \dim \operatorname{Im} T + \dim \operatorname{Ker} T = \dim R^3 = 3.$$

Za pomocą twierdzenia 8.2.2 udowodnimy teraz, że rząd macierzy $\mathbf{A}$ wymiaru $m \times n$ i o współczynnikach z ciała K jest identyczny z rzędem macierzy $\mathbf{A}^T\mathbf{A}$. W tym celu kolejny raz warto uświadomić sobie, że przestrzeń kolumnowa macierzy $\mathbf{A}$ jest identyczna z obrazem przekształcenia liniowego $T_{\mathbf{A}}\colon K_{n\times 1} \to K_{m\times 1}$, czyli jest $C_{\mathbf{A}} = \operatorname{Im} T_{\mathbf{A}}$. Natomiast jądro przekształcenia liniowego $T_{\mathbf{A}}$ pokrywa się z przestrzenią zerową macierzy $\mathbf{A}$, czyli $\operatorname{Ker} T_{\mathbf{A}} = N_{\mathbf{A}}$. Wymiar tej ostatniej przestrzeni, czyli liczbę $\dim N_{\mathbf{A}}$, nazywa się *zerowością macierzy* $\mathbf{A}$. Z twierdzenia 8.2.2 dla przekształcenia $T_{\mathbf{A}}\colon K_{n\times 1} \to K_{m\times 1}$, czyli z równości $\dim \operatorname{Ker} T_{\mathbf{A}} + \dim \operatorname{Im} T_{\mathbf{A}} = n$, mamy następującą zależność pomiędzy rzędem (zob. definicję 7.7.2) i zerowością macierzy $\mathbf{A}$:

Zerowość macierzy

$$r(\mathbf{A}) = \dim C_{\mathbf{A}} = \dim \operatorname{Im} T_{\mathbf{A}} = n - \dim \operatorname{Ker} T_{\mathbf{A}} = n - \dim N_{\mathbf{A}}$$

i stąd

$$r(\mathbf{A}) = n - \dim N_{\mathbf{A}}. \tag{8.16}$$

Twierdzenie 8.2.3. *Macierze rzeczywiste* $\mathbf{A}$ *i* $\mathbf{A}^T\mathbf{A}$ *mają identyczne rzędy, czyli*

$$r\,(\mathbf{A}) = r\,(\mathbf{A}^T\mathbf{A}). \tag{8.17}$$

Dowód. Załóżmy, że $\mathbf{A} \in R_{m\times n}$. Wtedy $\mathbf{A}^T\mathbf{A} \in R_{n\times n}$ i wobec (8.16) jest $r\,(\mathbf{A}) = n - \dim\, N_{\mathbf{A}}$ oraz $r\,(\mathbf{A}^T\mathbf{A}) = n - \dim\, N_{\mathbf{A}^T\mathbf{A}}$. Zatem dla dowodu równości (8.17) wystarczy wykazać równość przestrzeni zerowych $N_{\mathbf{A}}$ i $N_{\mathbf{A}^T\mathbf{A}}$.

W tym celu zauważmy najpierw, że jeśli $\mathbf{x} \in N_{\mathbf{A}}$, to $\mathbf{A}\mathbf{x} = \mathbf{0}$ i wtedy $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T(\mathbf{A}\mathbf{x}) = \mathbf{A}^T\mathbf{0} = \mathbf{0}$, więc także $\mathbf{x} \in N_{\mathbf{A}^T\mathbf{A}}$. To dowodzi, że $N_{\mathbf{A}} \subseteq N_{\mathbf{A}^T\mathbf{A}}$.

Weźmy teraz dowolny wektor $\mathbf{y}$ ze zbioru $N_{\mathbf{A}^T\mathbf{A}}$. Dla takiego $\mathbf{y}$ jest $\mathbf{A}^T\mathbf{A}\mathbf{y} = \mathbf{0}$. Zatem $(\mathbf{A}\mathbf{y})^T(\mathbf{A}\mathbf{y}) = \mathbf{y}^T\mathbf{A}^T\mathbf{A}\mathbf{y} = \mathbf{y}^T\mathbf{0} = 0$ i dlatego $\mathbf{A}\mathbf{y} = \mathbf{0}$, co dowodzi, że $\mathbf{y} \in N_{\mathbf{A}}$. Stąd $N_{\mathbf{A}^T\mathbf{A}} \subseteq N_{\mathbf{A}}$.

To kończy dowód równości $N_{\mathbf{A}} = N_{\mathbf{A}^T\mathbf{A}}$. $\square$

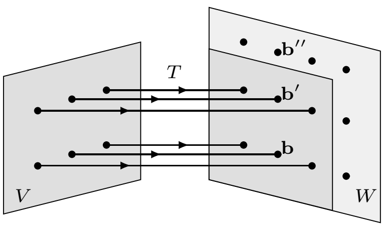
8.3. Mono- i epimorficzność przekształcenia liniowego

Definicja 8.3.1. Przekształcenie liniowe $T\colon V \to W$ nazywamy *monomorfizmem*, gdy każdy wektor $\mathbf{b}$ z przestrzeni W jest obrazem co najwyżej jednego wektora $\mathbf{x}$ z przestrzeni V, zob. rys. 8.7. Przekształcenie liniowe $T\colon V \to W$ jest *epimorfizmem*, gdy każdy wektor $\mathbf{b}$ z przestrzeni W jest obrazem co najmniej jednego wektora $\mathbf{x}$ z przestrzeni V, zob. rys. 8.8.

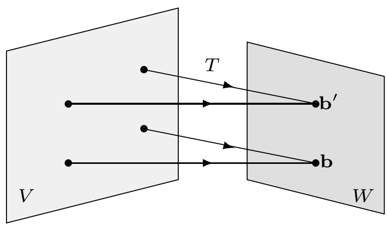
Monomorfizm

Epimorfizm

Równoważnie, przekształcenie liniowe $T\colon V \to W$ jest monomorfizmem (epimorfizmem), gdy dla każdego wektora $\mathbf{b} \in W$ równanie $T(\mathbf{x}) = \mathbf{b}$ ma co najwyżej (co najmniej) jedno rozwiązanie. Jeszcze inaczej: Przekształcenie liniowe $T\colon V \to W$ jest monomorfizmem, gdy T jest <u>różnowartościowym</u> odwzorowaniem przestrzeni wektorowej w przestrzeń W. Przekształcenie liniowe $T\colon V \to W$ jest epimorfizmem, gdy T odwzorowuje przestrzeń wektorową V <u>na przestrzeń</u> W, tj. gdy $\operatorname{Im} T = W$.[3]



Rys. 8.7



Rys. 8.8

[3] Odwzorowanie $T\colon V \to W$ jest <u>różnowartościowe</u>, gdy dla dowolnych $\mathbf{x}, \mathbf{x}' \in V$ z równości $T(\mathbf{x}) = T(\mathbf{x}')$ wynika równość $\mathbf{x} = \mathbf{x}'$. Odwzorowanie $T\colon V \to W$ przekształca zbiór V <u>na</u> zbiór W, gdy $T(V) = W$, czyli gdy dla każdego $\mathbf{y} \in W$ istnieje $\mathbf{x} \in V$, dla którego $\mathbf{y} = T(\mathbf{x})$.

Następne dwa twierdzenia pokazują bezpośrednie związki monomorfizmu i epimorfizmu z jądrem i obrazem przekształcenia liniowego.

Twierdzenie 8.3.1. *Dla przekształcenia liniowego $T\colon\ V \to W$ przestrzeni wektorowej V w przestrzeń wektorową W następujące warunki są równoważne:*

(*a*) *T jest monomorfizmem;*

(*b*) $\operatorname{Ker} T = \{\mathbf{0}\}$*;*

(*c*) $\dim \operatorname{Im} T = \dim V$.

Dowód. (*a*) $\Leftrightarrow$ (*b*) Załóżmy, że T jest monomorfizmem i niech $\mathbf{x}$ będzie dowolnym wektorem ze zbioru $\operatorname{Ker} T$. Wtedy $T(\mathbf{x}) = \mathbf{0} = T(\mathbf{0})$ i z różnowartościowości T mamy $\mathbf{x} = \mathbf{0}$. To dowodzi, że $\operatorname{Ker} T = \{\mathbf{0}\}$.

Załóżmy teraz, że $\operatorname{Ker} T = \{\mathbf{0}\}$ i przypuśćmy, że dla pewnych wektorów $\mathbf{x}$ i $\mathbf{y}$ jest $T(\mathbf{x}) = T(\mathbf{y})$. Wtedy $\mathbf{x} - \mathbf{y} \in \operatorname{Ker} T = \{\mathbf{0}\}$ (bo $T(\mathbf{x} - \mathbf{y}) = T(\mathbf{x}) - T(\mathbf{y}) = \mathbf{0}$). Stąd $\mathbf{x} - \mathbf{y} = \mathbf{0}$, czyli $\mathbf{x} = \mathbf{y}$. To dowodzi, że T jest monomorfizmem.

(*b*) $\Leftrightarrow$ (*c*) Wobec twierdzenia 8.2.2 mamy $\dim V = \dim \operatorname{Ker} T + \dim \operatorname{Im} T$. Stąd $\dim V = \dim \operatorname{Im} T$ wtedy i tylko wtedy, gdy $\dim \operatorname{Ker} T = 0$, czyli wtedy i tylko wtedy, gdy $\operatorname{Ker} T = \{\mathbf{0}\}$. $\square$

Okazuje się, że monomorficzność przekształcenia liniowego jest równoważna jego epimorficzności w jednym bardzo ważnym przypadku.

Twierdzenie 8.3.2. *Jeśli V i W są przestrzeniami wektorowymi tego samego skończonego wymiaru* ($\dim V = \dim W < \infty$) *i $T\colon\ V \to W$ jest przekształceniem liniowym, to T jest monomorfizmem wtedy i tylko wtedy, gdy T jest epimorfizmem.*

Dowód. Wobec twierdzenia 8.3.1 odwzorowanie T jest monomorfizmem wtedy i tylko wtedy, gdy $\dim \operatorname{Im} T = \dim V$, czyli wtedy i tylko wtedy, gdy $\dim \operatorname{Im} T = \dim W$ (bo $\dim W = \dim V$). Równość $\dim \operatorname{Im} T = \dim W$ wobec twierdzenia 7.5.5 jest równoważna równości $\operatorname{Im} T = W$, a ta z kolei jest równoważna epimorficzności odwzorowania liniowego T. $\square$

Następujące przykłady ilustrują przydatność powyższych twierdzeń przy badaniu i rozumieniu natury monomorficzności i/lub epimorficzności przekształcenia liniowego.

Przykład 180. Przekształcenie liniowe $T\colon\ R^4 \to R^3$, gdzie

$$T(\mathbf{x}) = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 1 & 2 & 1 \\ 2 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} \quad \text{dla} \quad \mathbf{x} = \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix},$$

nie jest monomorfizmem. (Z tych samych powodów, jeśli V i W są przestrzeniami wektorowymi (nad tym samym ciałem) i $\dim V > \dim W$, to żadne przekształcenie liniowe $T\colon\ V \to W$ nie jest monomorfizmem, tak jak nie było nim przekształcenie T z przykładu 173, 176, 177 lub 178.)

Ponieważ $\operatorname{Im} T \subseteq R^3$, więc $\dim \operatorname{Im} T \leqslant 3$ i wobec twierdzenia 8.2.2 mamy

$$\dim \operatorname{Ker} T = \dim R^4 - \dim \operatorname{Im} T \geqslant 1.$$

To zaś oznacza, że $\operatorname{Ker} T \neq \{\mathbf{0}\}$ i (wobec twierdzenia 8.3.1) odwzorowanie T nie jest monomorfizmem.

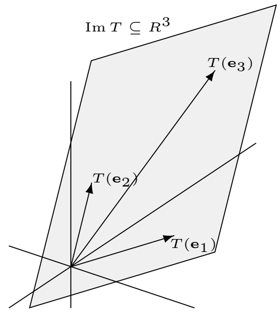
Przykład 181. Dane jest przekształcenie liniowe $T\colon R^3 \to R^3$ takie, że $T(\mathbf{e}_1) = (2,3,0)$, $T(\mathbf{e}_2) = (-1,1,3)$ i $T(\mathbf{e}_3) = (0,5,6)$, gdzie $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ jest bazą przestrzeni $V = R^3$. Ponieważ

$$T(\mathbf{e}_3) = (0,5,6) = (2,3,0) + 2(-1,1,3) = T(\mathbf{e}_1) + 2T(\mathbf{e}_2)$$

(zob. rys. 8.9), więc

$$\operatorname{Im} T = \mathcal{L}\big(T(\mathbf{e}_1), T(\mathbf{e}_2), T(\mathbf{e}_3)\big) = \mathcal{L}\big(T(\mathbf{e}_1), T(\mathbf{e}_2)\big)$$

i $\dim \operatorname{Im} T \leqslant 2 < 3 = \dim V$. Z twierdzenia 8.3.1 wynika, że przekształcenie T nie jest monomorfizmem. Wobec twierdzenia 8.3.2 przekształcenie T nie jest także epimorfizmem.



Rys. 8.9

Poprzedni przykład pokazuje, że obraz poprzez przekształcenie liniowe liniowo niezależnego zbioru wektorów może być zbiorem liniowo zależnym. Na zakończenie tej części udowodnimy, że tak nie może być, gdy przekształcenie jest monomorfizmem.

Twierdzenie 8.3.3. *Jeżeli przekształcenie liniowe $T\colon V \to W$ jest monomorfizmem, to wektory $\mathbf{b}_1, \ldots, \mathbf{b}_n$ są liniowo niezależne w przestrzeni V wtedy i tylko wtedy, gdy ich obrazy $T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)$ są liniowo niezależne w przestrzeni W.*

Dowód. Niech $T\colon V \to W$ będzie monomorfizmem (więc $\operatorname{Ker} T = \{\mathbf{0}\}$) i niech wektory $\mathbf{b}_1, \ldots, \mathbf{b}_n$ będą liniowo niezależne w przestrzeni V. Weźmy skalary $x_1, \ldots, x_n$ takie, że $\sum_{i=1}^n x_i T(\mathbf{b}_i) = \mathbf{0}$. Wtedy $\sum_{i=1}^n x_i T(\mathbf{b}_i) = T\big(\sum_{i=1}^n x_i \mathbf{b}_i\big) = \mathbf{0}$, więc $\sum_{i=1}^n x_i \mathbf{b}_i \in \operatorname{Ker} T = \{\mathbf{0}\}$ i dlatego $\sum_{i=1}^n x_i \mathbf{b}_i = \mathbf{0}$. Stąd i z niezależności wektorów $\mathbf{b}_1, \ldots, \mathbf{b}_n$ wynika, że $x_1 = \ldots = x_n = 0$. To dowodzi liniową niezależność wektorów $T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)$.

Załóżmy teraz, że wektory $T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)$ są liniowo niezależne i przypuśćmy, że $\sum_{i=1}^n x_i \mathbf{b}_i = \mathbf{0}$. Wtedy $\mathbf{0} = T(\mathbf{0}) = T(\sum_{i=1}^n x_i \mathbf{b}_i) = \sum_{i=1}^n x_i T(\mathbf{b}_i)$ i dlatego $x_1 = \ldots = x_n = 0$, bo wektory $T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)$ są liniowo niezależne. To dowodzi że wektory $\mathbf{b}_1, \ldots, \mathbf{b}_n$ są liniowo niezależne. (Zauważmy, że w dowodzie drugiej części twierdzenia nie korzystaliśmy z monomorficzności przekształcenia T.) $\square$

8.4. Suma i złożenie przekształceń liniowych

Definicja 8.4.1. Niech $T, U : V \to W$ będą funkcjami, gdzie V i W są przestrzeniami wektorowymi nad ciałem K i niech $\alpha \in K$. Wtedy $T + U : V \to W$ i $\alpha T\colon V \to W$ są funkcjami takimi, że dla każdego $\mathbf{x} \in V$ jest

$$(T+U)(\mathbf{x}) = T(\mathbf{x}) + U(\mathbf{x}) \quad \text{i} \quad (\alpha T)(\mathbf{x}) = \alpha T(\mathbf{x}). \tag{8.18}$$

$T+U$ – suma funkcji T i U

αT – iloczyn funkcji T i skalara α

Okazuje się, że suma przekształceń liniowych jest przekształceniem liniowym. Podobnie iloczyn przekształcenia liniowego i skalara jest przekształceniem liniowym. Można udowodnić, że zbiór $L(V, W)$, czyli zbiór wszystkich przekształceń liniowych przestrzeni wektorowej V w przestrzeń wektorową W, jest przestrzenią wektorową (z wyżej określonym dodawaniem przekształceń i mnożeniem przekształceń przez skalary).

Twierdzenie 8.4.1. *Niech V i W będą przestrzeniami wektorowymi nad tym samym ciałem K. Wtedy:*

(*a*) *$T + U \in L(V, W)$, gdy $T, U \in L(V, W)$;*

(*b*) *$\alpha T \in L(V, W)$, gdy $T \in L(V, W)$ i $\alpha \in K$;*

(*c*) *$L(V, W)$ jest przestrzenią wektorową nad ciałem K.*

Dowód. (a) Niech $\mathbf{x}$, $\mathbf{y} \in V$ i a, $b \in K$. Wtedy

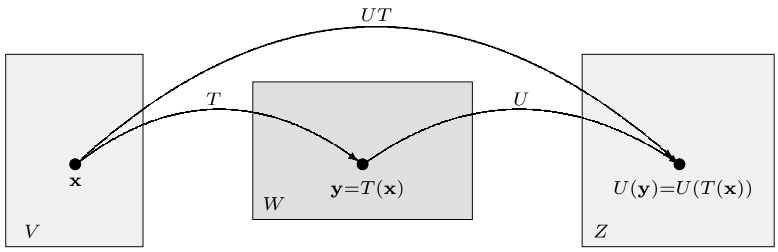
$$\begin{aligned}(T+U)(a\mathbf{x}+b\mathbf{y}) &= T(a\mathbf{x}+b\mathbf{y})+U(a\mathbf{x}+b\mathbf{y}) && \text{(z definicji sumy } T+U)\\ &= \big(aT(\mathbf{x})+bT(\mathbf{y})\big)+\big(aU(\mathbf{x})+bU(\mathbf{y})\big) && \text{(z liniowości } T \text{ i } U)\\ &= a\big(T(\mathbf{x})+U(\mathbf{x})\big)+b\big(T(\mathbf{y})+U(\mathbf{y})\big)\\ &= a\big(T+U\big)(\mathbf{x})+b\big(T+U\big)(\mathbf{y}) && \text{(z definicji sumy } T+U)\end{aligned}$$

i to oznacza, że $T+U$ jest przekształceniem liniowym. Podobnie dowodzi się stwierdzenie (b). Łatwy dowód stwierdzenia (c), sprawdzenie wszystkich warunków przestrzeni wektorowej, pozostawiamy czytelnikowi. □

Definicja 8.4.2. Jeśli V, W i Z są zbiorami, a $T\colon V \to W$ i $U\colon W \to Z$ są funkcjami, to ich *złożeniem* (lub *superpozycją*) nazywamy funkcję $UT\colon V \to Z$ taką, że dla każdego $\mathbf{x} \in V$ mamy

UT – złożenie funkcji T i U

$$(UT)(\mathbf{x}) = U(T(\mathbf{x})). \tag{8.19}$$



Rys. 8.10. Złożenie przekształceń T i U

W następnym twierdzeniu dowodzimy, że złożenie przekształceń liniowych jest przekształceniem liniowym. Inne własności złożenia operatorów liniowych przedstawiamy w ćwiczeniach.

Twierdzenie 8.4.2. *Niech V, W i Z będą przestrzeniami wektorowymi nad tym samym ciałem K. Jeśli $T\colon V \to W$ i $U\colon W \to Z$ są przekształceniami liniowymi, to także ich złożenie $UT\colon V \to Z$ jest przekształceniem liniowym.*

Dowód. Dla wektorów $\mathbf{x}$, $\mathbf{y} \in V$ i skalarów a, $b \in K$ mamy

$$\begin{aligned}\big(UT\big)(a\mathbf{x}+b\mathbf{y}) &= U\big(T(a\mathbf{x}+b\mathbf{y})\big) && \text{(z definicji złożenia } UT)\\ &= U\big(aT(\mathbf{x})+bT(\mathbf{y})\big) && \text{(z liniowości } T)\\ &= aU\big(T(\mathbf{x})\big)+bU\big(T(\mathbf{y})\big) && \text{(z liniowości } U)\\ &= a\big(UT\big)(\mathbf{x})+b\big(UT\big)(\mathbf{y}). && \text{(z definicji złożenia } UT)\end{aligned}$$

To dowodzi, że złożenie UT jest przekształceniem liniowym. □

Przykład 182. Jeśli $T\colon R^2 \to R^3$ i $U\colon R^3 \to R^2$ są przekształceniami liniowymi takimi, że

$$T(x,y) = (x, x+y, -3x-y) \quad \text{i} \quad U(v,u,w) = (v, v+u-w),$$

to ich złożeniem jest przekształcenie liniowe $UT\colon R^2 \to R^2$, dla którego

$$(UT)(x,y) = U(T(x,y)) = U(x, x+y, -3x-y) = (x, 5x+2y).$$

Dla tych samych przekształceń T i U istnieje także złóżenie $TU\colon R^3 \to R^3$ i dla każdego wektora $(v,u,w) \in R^3$ mamy

$$\begin{aligned}(TU)(v,u,w) &= T(U(v,u,w)) = T(v, v+u-w)\\ &= (v, 2v+u-w, -4v-u+w).\end{aligned}$$

8.5. Macierz przekształcenia liniowego

Definicja 8.5.1. Niech V i W będą przestrzeniami wektorowymi (nad tym samym ciałem K) i niech odpowiednio $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ oraz $C = (\mathbf{c}_1, \ldots, \mathbf{c}_m)$ będą ich bazami. Jeśli $T\colon V \to W$ jest przekształceniem liniowym, to macierz

$$[T]_C^B = \begin{bmatrix} | & & | \\ [T(\mathbf{b}_1)]_C & \cdots & [T(\mathbf{b}_n)]_C \\ | & & | \end{bmatrix}, \tag{8.20}$$

Macierz przekształcenia liniowego

której kolejnymi kolumnami są wektory C-współrzędnych wektorów $T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)$ jest nazywana *macierzą przekształcenia* liniowego $T\colon V \to W$ względem baz B i C (przestrzeni wektorowych V i W). Macierz operatora liniowego $T\colon V \to V$ względem baz B i B, czyli macierz $[T]_B^B$, nazywamy macierzą operatora T względem bazy B i oznaczamy symbolem $[T]_B$.

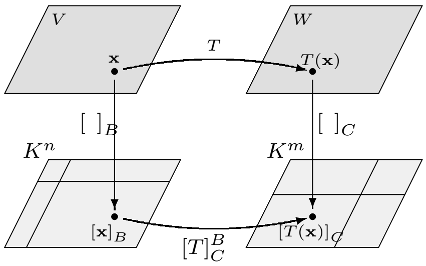
$[T]_B$ – macierz operatora T względem bazy B

Niech teraz $\mathbf{x}$ będzie dowolnym wektorem z przestrzeni V, powiedzmy $\mathbf{x} = x_1\mathbf{b}_1 + \ldots + x_n\mathbf{b}_n$ (dla pewnych skalarów $x_1, \ldots, x_n \in K$). Wtedy

$$T(\mathbf{x}) = T(x_1\mathbf{b}_1 + \ldots + x_n\mathbf{b}_n) = x_1T(\mathbf{b}_1) + \ldots + x_nT(\mathbf{b}_n)$$

i mamy następujący związek pomiędzy wektorem $[\mathbf{x}]_B$, czyli wektorem współrzędnych wektora $\mathbf{x}$ względem bazy B, i wektorem $[T(\mathbf{x})]_C$, wektorem współrzędnych wektora $T(\mathbf{x})$ względem bazy C,

$$\begin{aligned} [T(\mathbf{x})]_C &= \big[x_1T(\mathbf{b}_1) + \ldots + x_nT(\mathbf{b}_n)\big]_C \\ &= x_1[T(\mathbf{b}_1)]_C + \ldots + x_n[T(\mathbf{b}_n)]_C \\ &= \begin{bmatrix} | & & | \\ [T(\mathbf{b}_1)]_C & \cdots & [T(\mathbf{b}_n)]_C \\ | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &= [T]_C^B \cdot [\mathbf{x}]_B, \end{aligned}$$



Rys. 8.11

zob. rys. 8.11. Zatem udowodniliśmy następujące twierdzenie.

Twierdzenie 8.5.1. *Jeśli B i C są odpowiednio bazami skończenie wymiarowych przestrzeni wektorowych V i W, a $T\colon V \to W$ jest przekształceniem liniowym, to dla każdego wektora $\mathbf{x} \in V$ jest*

$$[T(\mathbf{x})]_C = [T]_C^B \cdot [\mathbf{x}]_B. \quad \square \tag{8.21}$$

Przykład 183. Niech $T\colon R^3 \to R^2$ będzie "rzutem" na płaszczyznę Oxy, czyli przekształceniem określonym wzorem $T(x, y, z) = (x, y)$, zob. rys. 8.12. Wyznaczyć macierz $[T]_C^B$ przekształcenia T względem baz $B = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ oraz $C = (\mathbf{c}_1, \mathbf{c}_2)$, gdzie $\mathbf{b}_1 = (1, 0, 0)$, $\mathbf{b}_2 = (-1, 1, 0)$, $\mathbf{b}_3 = (1, -1, 1)$, $\mathbf{c}_1 = (1, 2)$ i $\mathbf{c}_2 = (1, 3)$. Następnie za pomocą macierzy $[T]_C^B$ wyznaczyć $[T(\mathbf{x})]_C$ i $T(\mathbf{x})$, gdy $\mathbf{x} = (2, 3, 7)$.

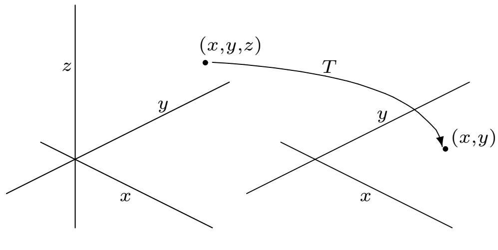
Łatwo sprawdzić, że

$$T(\mathbf{b}_1) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} - 2\begin{bmatrix} 1 \\ 3 \end{bmatrix} = 3\mathbf{c}_1 - 2\mathbf{c}_2,$$

$$T(\mathbf{b}_2) = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -4\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 3 \end{bmatrix} = -4\mathbf{c}_1 + 3\mathbf{c}_2$$

i

$$T(\mathbf{b}_3) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 2 \end{bmatrix} - 3\begin{bmatrix} 1 \\ 3 \end{bmatrix} = 4\mathbf{c}_1 - 3\mathbf{c}_2.$$



Rys. 8.12

Zatem

$$[T(\mathbf{b}_1)]_C = \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \quad [T(\mathbf{b}_2)]_C = \begin{bmatrix} -4 \\ 3 \end{bmatrix} \quad \text{oraz} \quad [T(\mathbf{b}_3)]_C = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

i wobec (8.20) mamy

$$[T]_C^B = \begin{bmatrix} | & | & | \\ [T(\mathbf{b}_1)]_C & [T(\mathbf{b}_2)]_C & [T(\mathbf{b}_2)]_C \\ | & | & | \end{bmatrix} = \begin{bmatrix} 3 & -4 & 4 \\ -2 & 3 & -3 \end{bmatrix}.$$

Z definicji przekształcenia T dla wektora $\mathbf{x} = (2, 3, 7)$ jest $T(\mathbf{x}) = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. Z drugiej strony

$$\mathbf{x} = \begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix} = 5\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 10\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + 7\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = 5\mathbf{b}_1 + 10\mathbf{b}_2 + 7\mathbf{b}_3,$$

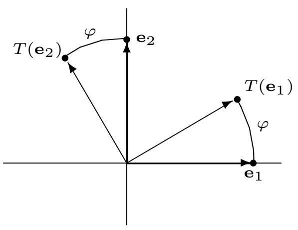
więc $[\mathbf{x}]_B = \begin{bmatrix} 5 \\ 10 \\ 7 \end{bmatrix}$ i wobec twierdzenia 8.5.1 (zob. (8.21)) mamy

$$[T(\mathbf{x})]_C = [T]_C^B\,[\mathbf{x}]_B = \begin{bmatrix} 3 & -4 & 4 \\ -2 & 3 & -3 \end{bmatrix}\begin{bmatrix} 5 \\ 10 \\ 7 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

i dlatego znowu

$$T(\mathbf{x}) = 3\mathbf{c}_1 + (-1)\mathbf{c}_2 = 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

W ostatnim przykładzie obraz wektora $\mathbf{x}$, czyli wektor $T(\mathbf{x})$, wyznaczyliśmy na dwa sposoby — wprost z definicji przekształcenia T oraz za pomocą macierzy przekształcenia T. Ten drugi sposób był dłuższy. Jednakże, jak to pokazuje następny przykład, i ten dłuższy sposób może mieć swoje zalety.



Rys. 8.13

Macierz obrotu płaszczyzny o kąt φ

Przykład 184. Niech φ będzie ustaloną liczbą rzeczywistą i niech T będzie obrotem płaszczyzny R^2 o kąt φ dookoła początku układu współrzędnych (w kierunku przeciwnym do kierunku ruchu wskazówek zegara), zob. rys. 8.13. Ponieważ jest to przekształcenie, które wektorom $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ i $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ bazy kanonicznej $E = (\mathbf{e}_1, \mathbf{e}_2)$ przestrzeni R^2 przyporządkowuje wektory

$$T(\mathbf{e}_1) = \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix} \quad \text{i} \quad T(\mathbf{e}_2) = \begin{bmatrix} -\sin\varphi \\ \cos\varphi \end{bmatrix},$$

więc jego macierzą względem bazy E jest

$$[T]_E = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}. \tag{8.22}$$

Znając obrazy wektorów bazowych, czyli znając macierz $[T]_E$, łatwo wyznacza się obraz $T(\mathbf{x})$ każdego innego wektora $\mathbf{x}$. Istotnie, ponieważ E jest bazą kanoniczną przestrzeni R^2, więc dla każdego $\mathbf{x} = (x, y) \in R^2$ jest $[\mathbf{x}]_E = \mathbf{x}$ i wobec (8.21) oraz (8.22) mamy

$$T(\mathbf{x}) = [T(\mathbf{x})]_E = [T]_E\,[\mathbf{x}]_E = [T]_E\,\mathbf{x} = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}.$$

Przedstawiona w twierdzeniu 8.5.1 równość

$$[T(\mathbf{x})]_C = [T]_C^B \cdot [\mathbf{x}]_B$$

pokazuje, że wektor C-współrzędnych wektora $T(\mathbf{x})$ można otrzymać pomnażając macierz $[T]_C^B$ (przekształcenia $T\colon V \to W$ względem baz B i C) przez wektor B-współrzędnych wektora $\mathbf{x}$. (Zob. także rys. 8.11 i/lub rys. 8.14.) Warto także pamiętać, że macierz $[T]_C^B$ ma m wierszy i n kolumn, gdzie $m = \dim W$ i $n \dim V$. Zauważmy, że i-ta kolumna $[T(\mathbf{b}_i)]_C$ macierzy $[T]_C^B$ jest wektorem C-współrzędnych wektora $T(\mathbf{b}_i)$ i jest to jedyne rozwiązanie równania wektorowego

$$x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \ldots + x_m\mathbf{c}_m = T(\mathbf{b}_i).$$

$$\begin{tikzcd} \mathbf{x} \arrow[r, "T"] \arrow[d, "{[\ ]_B}"'] & T(\mathbf{x}) \arrow[d, "{[\ ]_C}"] \\ {[\mathbf{x}]_B} \arrow[r, "{[T]_C^B}"'] & {[T(\mathbf{x})]_C} \end{tikzcd}$$

Rys. 8.14

Stąd w szczególności wynika, że jeśli przestrzeniami V i W są odpowiednio K^n i K^m i jeśli $\mathbf{C}$ oraz $T(\mathbf{B})$ są macierzami takimi, że

$$\mathbf{C} = \begin{bmatrix} | & | & & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_m \\ | & | & & | \end{bmatrix} \quad \text{i} \quad T(\mathbf{B}) = \begin{bmatrix} | & | & & | \\ T(\mathbf{b}_1) & T(\mathbf{b}_2) & \cdots & T(\mathbf{b}_n) \\ | & | & & | \end{bmatrix},$$

to jedynymi rozwiązaniami równań macierzowych

$$\mathbf{C}\,\mathbf{x}_1 = T(\mathbf{b}_1),\ \mathbf{C}\ \mathbf{x}_2 = T(\mathbf{b}_2), \ldots,\ \mathbf{C}\,\mathbf{x}_n = T(\mathbf{b}_n) \tag{8.23}$$

są kolumny macierzy $[T]_C^B$, czyli

$$\mathbf{x}_1 = [T(\mathbf{b}_1)]_C,\ \ \mathbf{x}_2 = [T(\mathbf{b}_2)]_C, \ldots, \mathbf{x}_n = [T(\mathbf{b}_n)]_C.$$

Równania (8.23) wyznaczają (i są wyznaczone przez) równanie macierzowe

$$\mathbf{C}\,\mathbf{X} = T(\mathbf{B}). \tag{8.24}$$

Jego jedynym rozwiązaniem jest

$$\mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ [T(\mathbf{b}_1)]_C & \cdots & [T(\mathbf{b}_n)]_C \\ | & & | \end{bmatrix} = [T]_C^B,$$

czyli macierz przekształcenia T względem baz B i C. Ponieważ macierz $\mathbf{C}$ jest kwadratowa i jej kolumny są liniowo niezależne, więc jest ona nieosobliwa i jedynym rozwiązaniem równania (8.24) jest $\mathbf{X} = \mathbf{C}^{-1}\,T(\mathbf{B})$. Stąd mamy

$$[T]_C^B = \mathbf{C}^{-1}\,T(\mathbf{B}),$$

więc macierz $[T]_C^B$ jest iloczynem macierzy $\mathbf{C}^{-1}$ i $T(\mathbf{B})$. Praktycznie, macierz $[T]_C^B$, jako rozwiązanie równania (8.24), możemy wyznaczyć metodą Gaussa-Jordana. W tym celu korzystamy z wierszowej równoważności

Sposób wyznaczania macierzy przekształcenia liniowego względem baz B i C

$$[\,\mathbf{C}\,|\,T(\mathbf{B})\,] \sim \left[\mathbf{I}_m\,|\,[T]_C^B\right]. \tag{8.25}$$

Przykład 185. Przekształcenie liniowe $T\colon R^3 \to R^3$ określone jest wzorem

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

Wyznaczyć: (a) macierz $[T]_C^B$, gdy $B = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ i $C = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3)$, gdzie $\mathbf{b}_1 = (1, 0, 0)$, $\mathbf{b}_2 = (1, 1, 1)$, $\mathbf{b}_3 = (1, -1, 1)$, $\mathbf{c}_1 = (1, 1, 2)$, $\mathbf{c}_2 = (1, 2, 1)$ i $\mathbf{c}_3 = (2, 1, 1)$; (b) macierz $[T]_D^B$, gdzie $D = (T(\mathbf{b}_1), T(\mathbf{b}_2),\ T(\mathbf{b}_3))$, a $\mathbf{b}_1$, $\mathbf{b}_2$

oraz $\mathbf{b}_3$ są takie jak w części (a); (c) macierz $[T]_E = [T]_E^E$, gdzie $E = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ jest bazą standardową przestrzeni R^3.

(a) Mamy

$$T(\mathbf{b}_1) = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \quad T(\mathbf{b}_2) = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \quad \text{i} \quad T(\mathbf{b}_3) = \begin{bmatrix} 5 \\ 0 \\ 3 \end{bmatrix}.$$

Musimy teraz poznać C-współrzędne każdego z wektorów $T(\mathbf{b}_i)$. W tym celu tworzymy macierz rozszerzoną $[\,\mathbf{C}\,|\,T(\mathbf{B})\,]$ i sprowadzamy ją do wierszowo równoważnej macierzy schodkowej normalnej $\left[\,\mathbf{I}_m\,|\,[T]_C^B\,\right]$ (zob. (8.25)). Mamy

$$\begin{aligned}
[\,\mathbf{C}\,|\,T(\mathbf{B})\,] \quad &= \left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 2 & 3 & 5 \\ 1 & 2 & 1 & 1 & 2 & 0 \\ 2 & 1 & 1 & 3 & 1 & 3 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 2 & 3 & 5 \\ 0 & 1 & -1 & -1 & -1 & -5 \\ 0 & -1 & -3 & -1 & -5 & -7 \end{array}\right] \\
&\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 3 & 4 & 10 \\ 0 & 1 & -1 & -1 & -1 & -5 \\ 0 & 0 & -4 & -2 & -6 & -12 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{2} & -\frac{1}{2} & 1 \\ 0 & 1 & 0 & -\frac{1}{2} & \frac{1}{2} & -2 \\ 0 & 0 & 1 & \frac{1}{2} & \frac{3}{2} & 3 \end{array}\right].
\end{aligned}$$

Stąd wynika, że macierzą przekształcenia T jest $[T]_C^B = \frac{1}{2}\begin{bmatrix} 3 & -1 & 2 \\ -1 & 1 & -4 \\ 1 & 3 & 6 \end{bmatrix}$.

(b) Przede wszystkim zauważmy, że układ $D = \big(T(\mathbf{b}_1), T(\mathbf{b}_2), T(\mathbf{b}_3)\big)$ jest liniowo niezależny w 3-wymiarowej przestrzeni R^3 (bo macierz $\left[\,T(\mathbf{b}_1)\,T(\mathbf{b}_2)\;T(\mathbf{b}_3)\,\right]$ $= \begin{bmatrix} 2 & 3 & 5 \\ 1 & 2 & 0 \\ 3 & 1 & 3 \end{bmatrix}$ jest nieosobliwa) i dlatego jest on bazą przestrzeni R^3. Łatwo teraz zauważyć, że

$$[T]_D^B = \begin{bmatrix} | & | & | \\ [T(\mathbf{b}_1)]_D & [T(\mathbf{b}_2)]_D & [T(\mathbf{b}_3)]_D \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ | & | & | \end{bmatrix} = \mathbf{I}_3.$$

(c) Ponieważ dla każdego $\mathbf{x} \in R^3$ jest $[\mathbf{x}]_E = \mathbf{x}$, więc mamy

$$[T]_E = \begin{bmatrix} | & | & | \\ [T(\mathbf{e}_1)]_E & [T(\mathbf{e}_2)]_E & [T(\mathbf{e}_3)]_E \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & T(\mathbf{e}_3) \\ | & | & | \end{bmatrix} = \begin{bmatrix} 2 & -1 & 2 \\ 1 & 1 & 0 \\ 3 & -1 & -1 \end{bmatrix}.$$

Ostatni przykład pokazuje, że w ogólnym przypadku macierz przekształcenia liniowego $T\colon V \to W$ zależy od wyboru baz przestrzeni V i W.

Przykład 186. Wyznaczyć $T(\mathbf{x})$, gdy $\mathbf{x} = (14, 7, 3)$ i macierzą przekształcenia liniowego $T\colon R^3 \to R^3$ względem baz $B = ((1,1,1), (2,1,0), (3,0,0))$ i $C = ((1,0,1), (2,1,0), (0,3,2))$ jest

$$[T]_C^B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 2 & 0 & 1 \end{bmatrix}.$$

Ponieważ znamy macierz $[T]_C^B$ przekształcenia T względem baz B i C, więc możemy kolejno wyznaczyć $[\mathbf{x}]_B$, $[T(\mathbf{x})]_C$ i $T(\mathbf{x})$.

Bez trudu stwierdzamy, że $[\mathbf{x}]_B = (3, 4, 1)$. Wtedy wobec (8.21) mamy

$$[T(\mathbf{x})]_C = [T]_C^B\,[\mathbf{x}]_B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 2 & 0 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 14 \\ 9 \\ 7 \end{bmatrix}$$

i dlatego

$$T(\mathbf{x}) = 14\begin{bmatrix}1\\0\\1\end{bmatrix} + 9\begin{bmatrix}2\\1\\0\end{bmatrix} + 7\begin{bmatrix}0\\3\\2\end{bmatrix} = \begin{bmatrix}32\\30\\28\end{bmatrix}.$$

Przykład 187. Dana jest macierz $\mathbf{A} = \begin{bmatrix}2 & 1 & 0\\2 & 4 & 1\end{bmatrix} \in R_{2\times 3}$. Wyznaczyć macierz przekształcenia $T_\mathbf{A}\colon R^3 \to R^2$, gdzie $T_\mathbf{A}(\mathbf{x}) = \mathbf{A}\mathbf{x}$, względem baz kanonicznych $E_3 = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ i $E_2 = (\mathbf{e}'_1, \mathbf{e}'_2)$ przestrzeni R^3 i R^2.

Niech $\mathbf{a}_1$, $\mathbf{a}_2$ i $\mathbf{a}_3$ będą kolejnymi kolumnami macierzy $\mathbf{A}$. Ponieważ $[\mathbf{x}]_{E_2} = \mathbf{x}$ dla $\mathbf{x} \in R^2$ i $[T_\mathbf{A}(\mathbf{e}_i)]_{E_2} = T_\mathbf{A}(\mathbf{e}_i) = \mathbf{A}\mathbf{e}_i = \mathbf{a}_i$, więc mamy

$$[T_\mathbf{A}]_{E_2}^{E_3} = \begin{bmatrix} | & | & | \\ [T(\mathbf{e}_1)]_{E_2} & [T(\mathbf{e}_2)]_{E_2} & [T(\mathbf{e}_3)]_{E_2} \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ | & | & | \end{bmatrix} = \mathbf{A}.$$

W ten sam sposób pokazuje się, że macierz $\mathbf{A} \in K_{m\times n}$ jest równa macierzy przekształcenia liniowego $T_\mathbf{A}\colon K^n \to K^m$ względem baz standardowych przestrzeni K^n i K^m.

$[T_\mathbf{A}]_{E_m}^{E_n} = \mathbf{A}$

W naszych dalszych rozważaniach wykorzystamy fakt, że macierz przekształcenia tożsamościowego względem różnych baz przestrzeni jest znaną nam (z poprzedniego rozdziału) macierzą przejścia od jednej bazy przestrzeni wektorowej do innej bazy tej samej przestrzeni.

Twierdzenie 8.5.2. *Niech $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ i $B' = (\mathbf{b}'_1, \ldots, \mathbf{b}'_n)$ będą bazami przestrzeni wektorowej V. Wtedy macierzą przekształcenia tożsamościowego $I_V\colon V \to V$ względem baz B i B' jest macierz przejścia od bazy B do bazy B',*

$$[I_V]_{B'}^B = \mathbf{P}_{B'}^B. \tag{8.26}$$

$I_V(\mathbf{x}) = \mathbf{x}$ dla $\mathbf{x} \in V$

$\mathbf{P}_{B'}^B = [I_V]_{B'}^B$ – macierz przejścia

Dowód. Z definicji macierzy przekształcenia liniowego (zob. def. 8.5.1) oraz z definicji macierzy przejścia (zob. tw. 7.6.3) mamy

$$[I_V]_{B'}^B = \begin{bmatrix} | & & | \\ [I_V(\mathbf{b}_1)]_{B'} & \cdots & [I_V(\mathbf{b}_n)]_{B'} \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ [\mathbf{b}_1]_{B'} & \cdots & [\mathbf{b}_n]_{B'} \\ | & & | \end{bmatrix} = \mathbf{P}_{B'}^B. \quad \square$$

$$\begin{tikzcd} V \arrow[r, "I_V"] \arrow[d, "{[\ ]_B}"'] & V \arrow[d, "{[\ ]_{B'}}"] \\ K^n \arrow[r, "\mathbf{P}_{B'}^B = [I_V]_{B'}^B"'] & K^n \end{tikzcd}$$

Rys. 8.15

Kolejne dwa twierdzenia pokazują zależności pomiędzy działaniami na przekształceniach liniowych i działaniami na odpowiadających im macierzach. Dokładniej, pokazują one, że macierz kombinacji liniowej przekształceń jest kombinacją liniową macierzy przekształceń, a macierz złożenia przekształceń jest iloczynem macierzy przekształceń.

Twierdzenie 8.5.3. *Niech V i W będą skończenie wymiarowymi przestrzeniami wektorowymi, których bazami są odpowiednio B i C, i niech $T, U\colon V \to W$ będą przekształceniami liniowymi. Wtedy:*

(a) $[T + U]_C^B = [T]_C^B + [U]_C^B$;

(b) $[aT]_C^B = a[T]_C^B$ *dla każdego skalara a.*

Dowód. (*a*) Jeśli $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ jest bazą przestrzeni V, to stwierdzenie (*a*) jest konsekwencją następującego ciągu równości:

$$\begin{aligned}
[T+U]_C^B &= \Big[[(T+U)(\mathbf{b}_1)]_C \ \ldots \ [(T+U)(\mathbf{b}_n)]_C \Big] && \text{(def. macierzy przekształcenia)}\\
&= \Big[[T(\mathbf{b}_1)+U(\mathbf{b}_1)]_C \ \ldots \ [T(\mathbf{b}_n)+U(\mathbf{b}_n)]_C \Big] && \text{(def. sumy } T+U\text{)}\\
&= \Big[[T(\mathbf{b}_1)]_C + [U(\mathbf{b}_1)]_C \ \ldots \ [T(\mathbf{b}_n)]_C + [U(\mathbf{b}_n)]_C \Big] && \text{(wobec tw. 7.6.1)}\\
&= \Big[[T(\mathbf{b}_1)]_C \ \ldots \ [T(\mathbf{b}_n)]_C \Big] + \Big[[U(\mathbf{b}_1)]_C \ \ldots \ [U(\mathbf{b}_n)]_C \Big] \\
& && \text{(def. sumy macierzy)}\\
&= [T]_C^B + [U]_C^B. && \text{(def. macierzy przekształcenia)}
\end{aligned}$$

Dowód stwierdzenia (*b*) jest podobny do tego z części (*a*). □

Twierdzenie 8.5.4. *Niech A, B i C będą odpowiednio bazami skończenie wymiarowych przestrzeni wektorowych V, W i Z. Jeśli $T\colon V \to W$ i $U\colon W \to Z$ są przekształceniami liniowymi, to*

$$[UT]_C^A = [U]_C^B \cdot [T]_B^A. \tag{8.27}$$

Dowód. Załóżmy, że bazami przestrzeni V, W i Z są odpowiednio $A = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$, $B = (\mathbf{b}_1, \ldots, \mathbf{b}_m)$ oraz $C = (\mathbf{c}_1, \ldots, \mathbf{c}_p)$. Niech a_{ij}, b_{ij} i c_{ij} będą odpowiednio współczynnikami macierzy $[T]_B^A$, $[U]_C^B$ i $[UT]_C^A$. Dokładniej, zakładamy, że

$$\begin{aligned}
[T]_B^A &= \Big[[T(\mathbf{a}_1)]_B \ [T(\mathbf{a}_2)]_B \ \ldots \ [T(\mathbf{a}_n)]_B \Big] = [a_{ij}] \in K_{m\times n},\\
[U]_C^B &= \Big[[U(\mathbf{b}_1)]_C \ [U(\mathbf{b}_2)]_C \ \ldots \ [U(\mathbf{b}_m)]_C \Big] = [b_{ij}] \in K_{p\times m},\\
[UT]_C^A &= \Big[[(UT)(\mathbf{a}_1)]_C \ [(UT)(\mathbf{a}_2)]_C \ \ldots \ [(UT)(\mathbf{a}_n)]_C \Big] = [c_{ij}] \in K_{p\times n}.
\end{aligned}$$

Dla dowodu równości (8.27) wystarczy wykazać, że $c_{ij} = \sum_{k=1}^m b_{ik}a_{kj}$ $(i = 1, \ldots, p,$ $j = 1, \ldots, n)$.

Ponieważ

$$\begin{bmatrix} c_{1j} \\ \vdots \\ c_{pj} \end{bmatrix} = \big[(UT)(\mathbf{a}_j)\big]_C, \quad \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} = \big[T(\mathbf{a}_j)\big]_B \ \text{ i } \ \begin{bmatrix} b_{1k} \\ \vdots \\ b_{pk} \end{bmatrix} = \big[U(\mathbf{b}_k)\big]_C,$$

więc kolejno mamy

$$\begin{aligned}
\sum_{i=1}^p c_{ij}\mathbf{c}_i &= \big(UT\big)(\mathbf{a}_j) = U\big(T(\mathbf{a}_j)\big) = U\left(\sum_{k=1}^m a_{kj}\mathbf{b}_k\right) = \sum_{k=1}^m a_{kj}U(\mathbf{b}_k)\\
&= \sum_{k=1}^m a_{kj}\left(\sum_{i=1}^p b_{ik}\mathbf{c}_i\right) = \sum_{i=1}^p \left(\sum_{k=1}^m b_{ik}a_{kj}\right)\mathbf{c}_i.
\end{aligned}$$

Z tych równości oraz z liniowej niezależności wektorów $\mathbf{c}_1, \ldots, \mathbf{c}_p$ wynika, że mamy $c_{ij} = \sum_{k=1}^m b_{ik}a_{kj}$. □

Wniosek 8.5.1. *Jeśli B jest bazą skończenie wymiarowej przestrzeni wektorowej V i jeśli T oraz U są operatorami liniowymi na przestrzeni V, to*

$$[UT]_B = [U]_B \cdot [T]_B. \quad \square \tag{8.28}$$

Wniosek 8.5.2. *Jeśli A, B i C są bazami skończenie wymiarowej przestrzeni wektorowej V, to macierz przejścia $[1_V]_C^A$ jest iloczynem macierzy $[1_V]_B^A$ i $[1_V]_C^B$,*

$$[1_V]_C^A = [1_V]_C^B \cdot [1_V]_B^A. \quad \square \tag{8.29}$$

Przykład 188. Niech przekształcenie $T\colon R^2 \to R^2$ będzie złożeniem $T_3T_2T_1$ trzech przekształceń T_1, T_2 i T_3 płaszczyzny R^2 w siebie, gdzie T_1 jest symetrią względem prostej $y = x$, T_2 jest obrotem o kąt $\pi/3$ (dookoła początku układu współrzędnych i w kierunku przeciwnym do kierunku ruch wskazówek zegara) i T_3 jest jednokładnością (o środku w początku układu współrzędnych) i skali 2. Wyznaczyć macierze przekształceń T_1, T_2, T_3 i T względem bazy standardowej $E = (\mathbf{e}_1, \mathbf{e}_2)$ przestrzeni R^2.

$T_1(x, y) = (y, x)$

$T_2(x, y) = \begin{bmatrix} \cos\frac{\pi}{3} & -\sin\frac{\pi}{3} \\ \sin\frac{\pi}{3} & \cos\frac{\pi}{3} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

$T_3(x, y) = 2(x, y)$

Ponieważ E jest bazą standardową przestrzeni R^2, więc $[\mathbf{x}]_E = \mathbf{x}$ dla $\mathbf{x} \in R^2$ i dlatego mamy

$$[T_i]_E = \Big[[T_i(\mathbf{e}_1)]_E \ [T_i(\mathbf{e}_2)]_E \Big] = \big[T_i(\mathbf{e}_1) \ T_i(\mathbf{e}_2) \big].$$

Stąd i z definicji przekształceń T_1, T_2 i T_3 (oraz z przykładu 184) mamy

$$[T_1]_E = \big[T_1(\mathbf{e}_1) \ T_1(\mathbf{e}_2) \big] = [\, \mathbf{e}_2 \ \mathbf{e}_1 \,] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

$$[T_2]_E = \begin{bmatrix} \cos\frac{\pi}{3} & -\sin\frac{\pi}{3} \\ \sin\frac{\pi}{3} & \cos\frac{\pi}{3} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{bmatrix}$$

i

$$[T_3]_E = \big[T_3(\mathbf{e}_1) \ T_3(\mathbf{e}_2) \big] = [\, 2\mathbf{e}_1 \, 2\mathbf{e}_2 \,] = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Zatem wobec wniosku 8.5.1 mamy

$$[T]_E = [T_3]_E \, [T_2]_E \, [T_1]_E = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \left(\frac{1}{2} \begin{bmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{bmatrix} \right) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -\sqrt{3} & 1 \\ 1 & \sqrt{3} \end{bmatrix}.$$

Na zakończenie tej części rozdziału udowodnimy twierdzenie o izomorficzności przestrzeni $L(V, W)$, czyli przestrzeni przekształceń liniowych, z przestrzenią macierzy. Przypomnijmy (zob. definicję 7.6.2), że dwie przestrzenie wektorowe X i Y (nad tym samym ciałem) są izomorficzne, gdy istnieje przekształcenie liniowe $\varphi\colon X \to Y$ przestrzeni X w przestrzeń Y, które jest mono- i epimorfizmem. Takie przekształcenie φ nazywa się *izomorfizmem przestrzeni wektorowych* X i Y.

Izomorfizm przestrzeni wektorowych

Twierdzenie 8.5.5. *Niech V i W będą przestrzeniami wektorowymi nad ciałem K. Jeżeli $\dim V = n$ i $\dim W = m$, to przestrzeń wektorowa $L(V, W)$ jest izomorficzna z przestrzenią $K_{m\times n}$.*

$L(V, W) \approx K_{m\times n}$

Dowód. Niech $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ i $C = (\mathbf{c}_1, \ldots, \mathbf{c}_m)$ będą bazami przestrzeni V i W. Udowodnimy, że odwzorowanie $\varphi\colon L(V, W) \to K_{m\times n}$, gdzie

$$\varphi(T) = [T]_C^B \quad \text{dla} \quad T \in L(V, W),$$

jest izomorfizmem.

Liniowość odwzorowania φ wynika z twierdzenia 8.5.3. Dla dowodu mono- i epimorficzności φ wystarczy pokazać, że dla każdej macierzy $\mathbf{A} \in K_{m\times n}$ istnieje dokładnie jedno przekształcenie liniowe $T\colon V \to W$, dla którego $\varphi(T) = \mathbf{A}$.

Weźmy dowolną macierz $\mathbf{A} = [a_{ij}] \in K_{m\times n}$. Wobec twierdzenia 8.1.1 istnieje dokładnie jedno przekształcenie liniowe $T\colon V \to W$ takie, że

$$T(\mathbf{b}_j) = \sum_{i=1}^{m} a_{ij}\mathbf{c}_i \quad \text{dla} \quad j = 1, \ldots, n.$$

Dla tego przekształcenia T wobec (8.20) mamy $[T]_C^B = \mathbf{A}$, więc $\varphi(T) = \mathbf{A}$, i to kończy dowód twierdzenia. $\square$

8.6. Odwracalność odwzorowania liniowego

Funkcja odwracalna

Definicja 8.6.1. Funkcję $T\colon V \to W$ nazywamy *odwracalną*, jeśli istnieje funkcja $U\colon W \to V$ taka, że

$$UT = I_V \quad \text{i} \quad TU = I_W. \tag{8.30}$$

Funkcję U mającą powyższe własności nazywa się *funkcją odwrotną* funkcji T. Łatwo dowodzi się, że funkcja odwracalna $T\colon V \to W$ ma dokładnie jedną funkcję odwrotną.[4] Tę jedyną funkcję odwrotną odwracalnej funkcji T oznacza się przez T^{-1}.

T^{-1} – funkcja odwrotna

Przykład 189. (*a*) Funkcja $T\colon R^2 \to R^2$ taka, że

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

jest odwracalna i jej funkcją odwrotną jest

$$U(x, y) = (2x - 3y, x - y).$$

Istotnie tak jest, bo mamy

$$(UT)(x, y) = U(-x + 3y, -x + 2y) = (x, y) = I_{R^2}(x, y)$$

i

$$(TU)(x, y) = T(2x - 3y, x - y) = (x, y) = I_{R^2}(x, y).$$

(*b*) Niech teraz $T\colon R^3 \to R^2$ będzie funkcją taką, że $T(x, y, z) = (x, y)$. Dla funkcji T i dla funkcji $U\colon R^2 \to R^3$ określonej wzorem $U(x, y) = (x, y, 0)$ jest

$$(TU)(x, y) = T(x, y, 0) = (x, y) = I_{R^2}(x, y).$$

Jednak dla tej funkcji U (i dla każdej innej funkcji $U\colon R^2 \to R^3$) jest

$$(UT)(x, y, z) = U(x, y) \neq (x, y, z) = I_{R^3}(x, y, z),$$

więc tym razem funkcja T nie jest odwracalna.

Mamy następującą charakteryzację funkcji odwracalnych.

Twierdzenie 8.6.1. *Funkcja $T\colon V \to W$ jest odwracalna wtedy i tylko wtedy, gdy jest ona różnowartościowa i odwzorowuje zbiór V na cały zbiór W.*

Dowód. Niech $U\colon W \to V$ będzie funkcją odwrotną funkcji $T\colon V \to W$. Wtedy wobec definicji 8.6.1 jest $UT = I_V$ i $TU = I_W$. Z różnowartościowości przekształcenia tożsamościowego $I_V = UT$ łatwo wynika różnowartościowość przekształcenia T. Równie łatwo z faktu, że przekształcenie tożsamościowe $I_W = TU$ jest odwzorowaniem na cały zbiór W wynika, że T jest odwzorowaniem na cały zbiór W.

Załóżmy teraz, że $T\colon V \to W$ jest odwzorowaniem różnowartościowym zbioru V na zbiór W. Ponieważ $T(V) = W$, więc dla każdego $\mathbf{y} \in W$ istnieje $\mathbf{x} \in V$ taki, że $\mathbf{y} = T(\mathbf{x})$. Stąd i z różnowartościowości T wynika, że dla każdego $\mathbf{y} \in W$ istnieje dokładnie jeden $\mathbf{x} \in V$ taki, że $\mathbf{y} = T(\mathbf{x})$. Zatem istnieje funkcja $U\colon W \to V$ taka, że dla każdego $\mathbf{y} \in W$ i każdego $\mathbf{x} \in V$ jest

$$U(\mathbf{y}) = \mathbf{x} \quad \text{wtedy i tylko wtedy, gdy} \quad T(\mathbf{x}) = \mathbf{y}.$$

[4] Jeśli funkcje $U_1, U_2\colon W \to V$ są funkcjami odwrotnymi funkcji $T\colon V \to W$, to z równości $U_iT = I_V$ oraz $TU_i = I_W$ mamy $U_1 = U_1I_W = U_1(TU_2) = (U_1T)U_2 = I_VU_2 = U_2$ i to dowodzi, że funkcja odwracalna ma tylko jedną funkcję odwrotną.

Ponieważ dla $\mathbf{x}$ i $\mathbf{y}$ spełniających powyższe relacje mamy

$$(UT)(\mathbf{x}) = U(T(\mathbf{x})) = U(\mathbf{y}) = \mathbf{x}$$

i

$$(TU)(\mathbf{y}) = T(U(\mathbf{y})) = T(\mathbf{x}) = \mathbf{y},$$

więc $UT = I_V$ oraz $TU = I_W$ i funkcja T jest odwracalna. □

Z twierdzeń 8.2.2 – 8.3.2 i 8.6.1 otrzymujemy następujące warunki konieczne i dostateczne odwracalności przekształcenia liniowego.

Wniosek 8.6.1. *Przekształcenie liniowe $T\colon V \to W$ skończenie wymiarowej przestrzeni wektorowej V w przestrzeń wektorową W jest odwracalne wtedy i tylko wtedy, gdy* $\operatorname{Ker} T = \{\mathbf{0}\}$ *i* $\dim V = \dim W$.

Dowód. Jeśli przekształcenie liniowe $T\colon V \to W$ jest odwracalne, to wobec twierdzenia 8.6.1 jest ono mono- i epimorfizmem. Zatem $\operatorname{Ker} T = \{\mathbf{0}\}$ (wobec tw. 8.3.1) i $\operatorname{Im} T = W$. Stąd i z twierdzenia 8.2.2 wynikają także równości

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

Załóżmy teraz, że $\dim V = \dim W$ i $T\colon V \to W$ jest przekształceniem liniowym takim, że $\operatorname{Ker} T = \{\mathbf{0}\}$. Z tych założeń oraz z twierdzeń 8.3.1 i 8.3.2 jest oczywiste, że T jest jednocześnie mono- i epimorfizmem. Stąd i z twierdzenia 8.6.1 wynika odwracalność odwzorowania T. □

Wniosek 8.6.2. *Operator liniowy $T\colon V \to V$ na skończenie wymiarowej przestrzeni wektorowej V jest odwracalny wtedy i tylko wtedy, gdy* $\operatorname{Ker} T = \{\mathbf{0}\}$. □

Przykład 190. Dany jest operator liniowy

$$T(a + bx + cx^2) = a + b + c + (2a + b + c)x + (a + b)x^2$$

na przestrzeni $R_2[x]$. Uzasadnić odwracalność operatora T.

Wobec wniosku 8.6.2 wystarczy pokazać, że wielomian zerowy jest jedynym elementem jądra odwzorowania T. Ponieważ mamy

$$\begin{aligned} a + bx + cx^2 \in \operatorname{Ker} T \quad &\Leftrightarrow \quad T(a + bx + cx^2) \equiv 0 \\ &\Leftrightarrow \quad a + b + c + (2a + b + c)x + (a + b)x^2 \equiv 0 \\ &\Leftrightarrow \quad \begin{cases} a + b + c = 0 \\ 2a + b + c = 0 \\ a + b = 0 \end{cases} \\ &\Leftrightarrow \quad a = b = c = 0 \\ &\Leftrightarrow \quad a + bx + cx^2 \equiv 0, \end{aligned}$$

więc $\operatorname{Ker} T = \{0\}$ i dlatego operator T jest odwracalny.

Udowodnimy teraz, że przekształcenie odwrotne przekształcenia liniowego jest przekształceniem liniowym.

Twierdzenie 8.6.2. *Jeżeli V i W są przestrzeniami wektorowymi i $T\colon V \to W$ jest odwracalnym przekształceniem liniowym, to także przekształcenie odwrotne $T^{-1}\colon W \to V$ jest przekształceniem liniowym.*

Dowód. Z definicji przekształcenia odwrotnego T^{-1} i z liniowości przekształcenia T wynika, że dla każdych wektorów $\mathbf{y}_1, \mathbf{y}_2 \in W$ i skalarów a_1 i a_2 jest

$$\begin{aligned} T^{-1}(a_1\mathbf{y}_1 + a_2\mathbf{y}_2) &= T^{-1}\big(a_1T(T^{-1}(\mathbf{y}_1)) + a_2T(T^{-1}(\mathbf{y}_2))\big) && (TT^{-1} = I_W) \\ &= T^{-1}\big(T\big(a_1T^{-1}(\mathbf{y}_1) + a_2T^{-1}(\mathbf{y}_2)\big)\big) && (\text{z liniowości } T) \\ &= a_1T^{-1}(\mathbf{y}_1) + a_2T^{-1}(\mathbf{y}_2). && (T^{-1}T = I_V) \end{aligned}$$

To kończy dowód liniowości przekształcenia T^{-1}. □

Z wniosku 8.6.1 wynika, że jeśli V i W są skończenie wymiarowymi przestrzeniami wektorowymi i $\dim V \neq \dim W$, to nie istnieje odwracalne przekształcenie liniowe przestrzeni V w przestrzeń W. Z tego względu w następnym twierdzeniu, w którym dowodzimy, że przekształcenie liniowe jest odwracalne wtedy i tylko wtedy, gdy jego macierz jest odwracalna, rozważane przestrzenie wektorowe są tego samego skończonego wymiaru.

Twierdzenie 8.6.3. *Załóżmy, że V i W są n-wymiarowymi przestrzeniami wektorowymi oraz $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ i $C = (\mathbf{c}_1, \ldots, \mathbf{c}_n)$ są odpowiednio ich bazami. Jeżeli $T\colon V \to W$ jest przekształceniem liniowym, to T jest odwzorowaniem odwracalnym wtedy i tylko wtedy, gdy jego macierz $[T]_C^B$ jest odwracalna. Dodatkowo,*

$$\left([T]_C^B\right)^{-1} = [T^{-1}]_B^C. \tag{8.31}$$

Dowód. Załóżmy najpierw, że odwzorowanie $T\colon V \to W$ jest odwracalne. Wtedy odwzorowanie odwrotne $T^{-1}\colon W \to V$ istnieje i dla niego jest $TT^{-1} = I_W$ oraz $T^{-1}T = I_V$. Stąd i z twierdzenia 8.5.3 wynika, że dla macierzy odwzorowań T, T^{-1}, TT^{-1} i $T^{-1}T$ mamy

$$\mathbf{I}_n = [I_W]_C = [TT^{-1}]_C = [T]_C^B \cdot [T^{-1}]_B^C$$

oraz

$$\mathbf{I}_n = [I_V]_B = [T^{-1}T]_B = [T^{-1}]_B^C \cdot [T]_C^B,$$

a to oznacza odwracalność macierzy $[T]_C^B$ i równość (8.31).

Załóżmy teraz, że macierz $\mathbf{A} = [T]_C^B$ jest odwracalna i niech $\mathbf{B} = [b_{ij}]$ będzie jej macierzą odwrotną. Wtedy $\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n$. Niech $U\colon W \to V$ będzie przekształceniem liniowym takim, że

$$U(\mathbf{c}_j) = \sum_{i=1}^{n} b_{ij}\mathbf{b}_i \tag{8.32}$$

dla $j = 1, \ldots, n$. (Istnienie (i jednoznaczność) takiego przekształcenia U wynika z twierdzenia 8.1.1.) Dla dowodu odwracalności przekształcenia T (i równości $T^{-1} = U$) wystarczy wobec definicji 8.6.1 pokazać, że złożenia UT i TU są przekształceniami tożsamościowymi, $UT = I_V$ i $TU = I_W$.

Wobec (8.32)) jest oczywiste, że mamy $[U]_B^C = \Big[\, [U(\mathbf{c}_1)]_B \ \ldots \ [U(\mathbf{c}_n)]_B \,\Big] = \mathbf{B}$. Stąd zaś i z twierdzenia 8.5.3 otrzymujemy

$$[UT]_B = [U]_B^C \cdot [T]_C^B = \mathbf{BA} = \mathbf{I}_n = [I_V]_B.$$

W końcu z równości $[UT]_B = [I_V]_B$ wynika równość $UT = I_V$. (Jeśli byłoby $UT \neq I_V$, to wobec wniosku 8.1.1 byłoby $\big(UT\big)(\mathbf{b}_i) \neq I_V(\mathbf{b}_i)$ dla pewnego $i \in \{1, \ldots, n\}$ i wtedy i-ta kolumna macierzy $[UT]_B$ byłaby różna od i-tej kolumny macierzy $[I_V]_B$ $\Big(\big[(UT)(\mathbf{b}_i)\big]_B \neq [I_V(\mathbf{b}_i)]_B\Big)$, co przeczyłoby równości $[UT]_B = [I_V]_B$.) Podobnie dowodzi się, że $TU = I_W$. □

Z powyższego twierdzenia wynikają dwa proste wnioski.

Wniosek 8.6.3. *Niech $T\colon V \to V$ będzie operatorem liniowym na skończenie wymiarowej przestrzeni wektorowej V z bazą B. Operator T jest odwracalny wtedy i tylko wtedy, gdy jego macierz $[T]_B$ jest odwracalna. Dla odwracalnego operatora T mamy też*

$\big([T]_B\big)^{-1} = \big[T^{-1}\big]_B$

$$([T]_B)^{-1} = \big[T^{-1}\big]_B. \quad \square \tag{8.33}$$

Wniosek 8.6.4. *Jeżeli $\mathbf{A}$ jest macierzą wymiaru $n \times n$ i o współczynnikach z ciała K, to macierz $\mathbf{A}$ jest odwracalna wtedy i tylko wtedy, gdy operator $T_{\mathbf{A}}\colon K^n \to K^n$ jest odwracalny. Dla odwracalnej macierzy $\mathbf{A}$ mamy też*

$(T_{\mathbf{A}})^{-1} = T_{\mathbf{A}^{-1}}$

$$(T_{\mathbf{A}})^{-1} = T_{\mathbf{A}^{-1}}. \quad \square \tag{8.34}$$

Przykład 191. Dane jest przekształcenie liniowe $T\colon R^2 \to R^2$ takie, że

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

(a) Wyznaczyć macierz $[T]_E$ przekształcenia T względem bazy kanonicznej E przestrzeni R^2. (b) Wykazać odwracalność odwzorowania T oraz wyznaczyć $[T^{-1}]_E$ i $T^{-1}(x, y)$. (c) Wyznaczyć $T^{-1}(\mathbf{b}_1)$ i $T^{-1}(\mathbf{b}_2)$ oraz $[T^{-1}]_B$, gdzie $B = (\mathbf{b}_1, \mathbf{b}_2) = ((4, -1), (1, -1))$ jest bazą przestrzeni R^2.

Mamy $[T]_E = \begin{bmatrix} 3 & 4 \\ -1 & -2 \end{bmatrix}$ i macierz ta jest odwracalna. Zatem wobec wniosku 8.6.3 także i przekształcenie T jest odwracalne oraz

$$[T^{-1}]_E = \big([T]_E\big)^{-1} = \frac{1}{2}\begin{bmatrix} 2 & 4 \\ -1 & -3 \end{bmatrix}.$$

Dlatego

$$T^{-1}(x, y) = \frac{1}{2}\begin{bmatrix} 2 & 4 \\ -1 & -3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{2}(2x + 4y, -x - 3y).$$

Stąd $T^{-1}(\mathbf{b}_1) = \frac{1}{2}(4, -1) = \frac{1}{2}\mathbf{b}_1$, $T^{-1}(\mathbf{b}_2) = \frac{1}{2}(-2, 2) = -\mathbf{b}_2$ i w końcu mamy

$$[T^{-1}]_B = \big[\ [T^{-1}(\mathbf{b}_1)]_B\ [T^{-1}(\mathbf{b}_2)]_B\ \big] = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & -1 \end{bmatrix}.$$

8.7. Podobieństwo macierzy

Definicja 8.7.1. Dwie macierze $\mathbf{A}, \mathbf{B} \in K_{m\times n}$ nazywamy *równoważnymi*, jeśli istnieją nieosobliwe macierze $\mathbf{P}$ i $\mathbf{Q}$ takie, że

Równoważność macierzy

$$\mathbf{B} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{P}. \tag{8.35}$$

Podobieństwo macierzy, o którym mówiliśmy w rozdziale trzecim, jest szczególnym przypadkiem równoważności macierzy. Przypomnijmy, że macierze kwadratowe $\mathbf{A}$ i $\mathbf{B}$ nazywamy *podobnymi*, jeśli istnieje macierz nieosobliwa $\mathbf{P}$ (zwana *macierzą podobieństwa* macierzy $\mathbf{A}$ do macierzy $\mathbf{B}$) taka, że

Podobieństwo macierzy

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}. \tag{8.36}$$

Wspomnieliśmy już, że macierz przekształcenia liniowego $T\colon V \to W$ przestrzeni wektorowej V w przestrzeń wektorową W zależy jednocześnie od wyboru bazy przestrzeni V i od wyboru bazy przestrzeni W. Badaniu tych zależności poświęcamy pozostałą część tego rozdziału. Pokazujemy, że macierze przekształcenia T względem różnych baz przestrzeni V i/lub W są równoważne. Dowodzimy też, że każde dwie równoważne macierze mogą być utożsamiane z macierzami tego samego przekształcenia liniowego względem różnych wyborów baz.

Twierdzenie 8.7.1. *Jeśli B i B' są bazami przestrzeni wektorowej V, a C i C' są bazami przestrzeni wektorowej W oraz $T\colon V \to W$ jest przekształceniem liniowym, to macierze $[T]_C^B$ i $[T]_{C'}^{B'}$ są równoważne,*

Wzór na zmianę macierzy przekształcenia liniowego przy zmianie baz

$$[T]_{C'}^{B'} = \mathbf{Q}^{-1}\,[T]_C^B\,\mathbf{P}, \tag{8.37}$$

gdzie $\mathbf{P} = [I_V]_B^{B'}$ jest macierzą przejścia od bazy B' do bazy B, a $\mathbf{Q} = [I_W]_C^{C'}$ jest macierzą przejścia od bazy C' do bazy C.

Dowód. Niech I_V oraz I_W będą przekształceniami tożsamościowymi odpowiednio na V oraz W. Wtedy $I_W T = T I_V$ (co oznacza przemienność diagramu z rys. 8.16) i wobec twierdzenia 8.5.4 mamy

$$\mathbf{Q}\,[T]_{C'}^{B'} = [I_W]_C^{C'}\,[T]_{C'}^{B'} = [I_W T]_C^{B'} = [T I_V]_C^{B'} = [T]_C^B\,[I_V]_B^{B'} = [T]_C^B\,\mathbf{P}.$$

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "I_V"'] & W \arrow[d, "I_W"] \\ V \arrow[r, "T"'] & W \end{tikzcd}$$

Rys. 8.16

Stąd otrzymujemy $[T]_{C'}^{B'} = \mathbf{Q}^{-1}\,[T]_C^B\,\mathbf{P}$. $\square$

Wniosek 8.7.1. *Jeśli $T\colon V \to V$ jest operatorem liniowym oraz B i B' są bazami przestrzeni V, to macierze $[T]_B$ i $[T]_{B'}$ są podobne i*

$$[T]_{B'} = \mathbf{P}^{-1}[T]_B\mathbf{P}, \tag{8.38}$$

gdzie $\mathbf{P} = \mathbf{P}_B^{B'}$. $\square$

Przykład 192. Dane jest przekształcenie liniowe $T\colon\ R^3 \to R^2$, gdzie $T(x,y,z) = (x-y, 2x+3z)$. Wyznaczyć macierze $[T]_C^B$ i $[T]_{C'}^{B'}$ przekształcenia T oraz macierze przejścia $\mathbf{P} = [I_{R^3}]_B^{B'}$ i $\mathbf{Q}^{-1} = [I_{R^2}]_{C'}^{C}$, gdy $B = \big((1,-1,1),(1,0,1),(0,1,1)\big)$ i $B' = \big((3,1,1),(2,2,1),(1,2,1)\big)$ są bazami przestrzeni R^3, a $C = \big((0,1),(-1,3)\big)$ i $C' = \big((2,1),(7,3)\big)$ – bazami przestrzeni R^2. Sprawdzić, czy dla otrzymanych macierzy zachodzi równość $[I_{R^2}]_{C'}^{C}\,[T]_C^B\,[I_{R^3}]_B^{B'} = [T]_{C'}^{B'}$.

Macierze $[T]_C^B$ i $[T]_{C'}^{B'}$ (tak jak w przykładzie 185) wyznaczamy z równoważności

$$[\,\mathbf{C}\,|\,T(\mathbf{B})\,] \sim \big[\,\mathbf{I}\,|\,[T]_C^B\,\big] \quad \text{i} \quad [\,\mathbf{C}'\,|\,T(\mathbf{B}')\,] \sim \big[\,\mathbf{I}\,|\,[T]_{C'}^{B'}\,\big].$$

Ponieważ mamy

$$\big[\,\mathbf{C}\,|\,T(\mathbf{B})\,\big] = \left[\begin{array}{cc|ccc} 0 & -1 & 2 & 1 & -1 \\ 1 & 3 & 5 & 5 & 3 \end{array}\right] \sim \left[\begin{array}{cc|ccc} 1 & 0 & 11 & 8 & 0 \\ 0 & 1 & -2 & -1 & 1 \end{array}\right] = \big[\,\mathbf{I}\,|\,[T]_C^B\,\big]$$

i

$$\big[\,\mathbf{C}'\,|\,T(\mathbf{B}')\,\big] = \left[\begin{array}{cc|ccc} 2 & 7 & 2 & 0 & -1 \\ 1 & 3 & 9 & 7 & 5 \end{array}\right] \sim \left[\begin{array}{cc|ccc} 1 & 0 & 57 & 49 & 38 \\ 0 & 1 & -16 & -14 & -11 \end{array}\right] = \big[\,\mathbf{I}\,|\,[T]_{C'}^{B'}\,\big],$$

więc

$$[T]_C^B = \begin{bmatrix} 11 & 8 & 0 \\ -2 & -1 & 1 \end{bmatrix} \quad \text{i} \quad [T]_{C'}^{B'} = \begin{bmatrix} 57 & 49 & 38 \\ -16 & -14 & -11 \end{bmatrix}.$$

W podobny sposób wyznaczamy macierze przejścia $\mathbf{P} = [I_{R^3}]_B^{B'}$ i $\mathbf{Q}^{-1} = [I_{R^2}]_{C'}^{C}$. Tym razem korzystamy z równoważności $\big[\,\mathbf{B}\,|\,I_{R^3}(\mathbf{B}')\,\big] = \big[\,\mathbf{B}\,|\,\mathbf{B}'\,\big] \sim \big[\,\mathbf{I}\,|\,\mathbf{P}\,\big]$ i $\big[\,\mathbf{C}'\,|\,I_{R^2}(\mathbf{C})\big] = \big[\,\mathbf{C}'\,|\,\mathbf{C}\,\big] \sim \big[\,\mathbf{I}\,|\,\mathbf{Q}^{-1}\,\big]$, z których otrzymujemy

$$\mathbf{P} = [I_{R^3}]_B^{B'} = \begin{bmatrix} -3 & -3 & -2 \\ 6 & 5 & 3 \\ -2 & -1 & 0 \end{bmatrix} \quad \text{i} \quad \mathbf{Q}^{-1} = [I_{R^2}]_{C'}^{C} = \begin{bmatrix} 7 & 24 \\ -2 & -7 \end{bmatrix}.$$

Łatwo teraz zauważyć, że istotnie mamy

$$\begin{aligned} \mathbf{Q}^{-1}\,[T]_C^B\,\mathbf{P} &= \begin{bmatrix} 7 & 24 \\ -2 & -7 \end{bmatrix}\begin{bmatrix} 11 & 8 & 0 \\ -2 & -1 & 1 \end{bmatrix}\begin{bmatrix} -3 & -3 & -2 \\ 6 & 5 & 3 \\ -2 & -1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 29 & 32 & 24 \\ -8 & -9 & -7 \end{bmatrix}\begin{bmatrix} -3 & -3 & -2 \\ 6 & 5 & 3 \\ -2 & -1 & 0 \end{bmatrix} = \begin{bmatrix} -57 & 49 & 38 \\ -16 & -14 & -11 \end{bmatrix} = [T]_{C'}^{B'}. \end{aligned}$$

Przykład 193. Dana jest pewna przestrzeń wektorowa V z bazami $B = (\mathbf{b}_1, \mathbf{b}_2)$ i $B' = (\mathbf{b}_1+\mathbf{b}_2, 2\mathbf{b}_1+\mathbf{b}_2)$. Niech $T\colon\ V \to V$ będzie operatorem liniowym, którego macierzą względem bazy B jest $\mathbf{A} = \begin{bmatrix} 0 & 2 \\ -1 & 3 \end{bmatrix}$. Wyznaczyć:

(a) $[T(\mathbf{b}_1)]_B$ i $T(\mathbf{b}_1)$;

(b) $T(\mathbf{b}_1+\mathbf{b}_2)$, $[T(\mathbf{b}_1+\mathbf{b}_2)]_{B'}$, $T(2\mathbf{b}_1+\mathbf{b}_2)$ i $[T(2\mathbf{b}_1+\mathbf{b}_2)]_{B'}$;

(c) $[T]_{B'}$;

(d) macierz podobieństwa $\mathbf{P}$ macierzy $\mathbf{B}=\begin{bmatrix}2&0\\0&1\end{bmatrix}$ do macierzy $\mathbf{A}$.

(a) Ponieważ macierz $\mathbf{A}$ jest macierzą przekształcenia T względem bazy B, więc oczywiście $\mathbf{A}=[T]_B=\Big[[T(\mathbf{b}_1)]_B\,[T(\mathbf{b}_2)]_B\Big]$ i dlatego mamy

$$[T(\mathbf{b}_1)]_B=\begin{bmatrix}0\\-1\end{bmatrix}\quad\text{oraz}\quad T(\mathbf{b}_1)=0\,\mathbf{b}_1+(-1)\mathbf{b}_2=-\mathbf{b}_2.$$

Z tych samych powodów mamy $T(\mathbf{b}_2)=2\mathbf{b}_1+3\mathbf{b}_2$.

(b) Z liniowości przekształcenia T oraz z (a) mamy

$$\begin{aligned}T(\mathbf{b}_1+\mathbf{b}_2)&=T(\mathbf{b}_1)+T(\mathbf{b}_2)=-\mathbf{b}_2+(2\mathbf{b}_1+3\mathbf{b}_2)\\&=2(\mathbf{b}_1+\mathbf{b}_2)=2(\mathbf{b}_1+\mathbf{b}_2)+0(2\mathbf{b}_1+\mathbf{b}_2)\end{aligned}$$

i stąd

$$[T(\mathbf{b}_1+\mathbf{b}_2)]_{B'}=\begin{bmatrix}2\\0\end{bmatrix}.$$

Podobnie znajdujemy

$$T(2\mathbf{b}_1+\mathbf{b}_2)=0(\mathbf{b}_1+\mathbf{b}_2)+1(2\mathbf{b}_1+\mathbf{b}_2)\quad\text{i}\quad[T(2\mathbf{b}_1+\mathbf{b}_2)]_{B'}=\begin{bmatrix}0\\1\end{bmatrix}.$$

(c) Korzystając z (b), otrzymujemy

$$[T]_{B'}=\Big[[T(\mathbf{b}_1+\mathbf{b}_2)]_{B'}\,[T(2\mathbf{b}_1+\mathbf{b}_2)]_{B'}\Big]=\begin{bmatrix}2&0\\0&1\end{bmatrix}.$$

(d) Ponieważ $\mathbf{A}=[T]_B$ i $\mathbf{B}=[T]_{B'}$, więc wobec wniosku 8.7.1 jest $\mathbf{B}=\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ i macierzą podobieństwa $\mathbf{P}$ macierzy $\mathbf{A}$ do macierzy $\mathbf{B}$ jest macierz przejścia od bazy B' do bazy B,

$$\mathbf{P}=\mathbf{P}_B^{B'}=\Big[[\mathbf{b}_1+\mathbf{b}_2]_B\,[2\mathbf{b}_1+\mathbf{b}_2]_B\Big]=\begin{bmatrix}1&2\\1&1\end{bmatrix}.$$

Przykład 194. Macierzą przekształcenia liniowego $T\colon V\to W$ względem bazy $B=(\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3)$ przestrzeni V i bazy $C=(\mathbf{c}_1,\mathbf{c}_2)$ przestrzeni W jest

$$[T]_C^B=\begin{bmatrix}1&3&2\\2&-1&-1\end{bmatrix}.$$

Wyznaczyć macierz $[T]_{C'}^{B'}$, gdy $B'=(\mathbf{b}_1+\mathbf{b}_2,\mathbf{b}_1+\mathbf{b}_3,\mathbf{b}_1+\mathbf{b}_2+\mathbf{b}_3)$ oraz $C'=(2\mathbf{c}_1+3\mathbf{c}_2,3\mathbf{c}_1+5\mathbf{c}_2)$.

Korzystamy ze wzoru na zmianę macierzy przekształcenia przy zmianie baz (zob. (8.37)) i macierz $[T]_{C'}^{B'}$ wyznaczamy z zależności

$$[T]_{C'}^{B'}=[I_W]_{C'}^{C}[T]_C^B[I_V]_B^{B'}=\left([I_W]_C^{C'}\right)^{-1}[T]_C^B[I_V]_B^{B'}.$$

Ponieważ każdy wektor bazy B' jest już wskazaną kombinacją liniową wektorów bazy B, więc mamy $[I_V]_B^{B'}=\begin{bmatrix}1&1&1\\1&0&1\\0&1&1\end{bmatrix}$. Z tych samych powodów $[I_W]_C^{C'}=\begin{bmatrix}2&3\\3&5\end{bmatrix}$ i w końcu mamy

$$[T]_{C'}^{B'}=\begin{bmatrix}2&3\\3&5\end{bmatrix}^{-1}\begin{bmatrix}1&3&2\\2&-1&-1\end{bmatrix}\begin{bmatrix}1&1&1\\1&0&1\\0&1&1\end{bmatrix}=\begin{bmatrix}17&12&30\\-10&-7&-18\end{bmatrix}.$$

Udowodnimy teraz, że każde dwie równoważne macierze są macierzami tego samego przekształcenia liniowego względem różnych baz przestrzeni wektorowych.

Twierdzenie 8.7.2. *Niech V i W będą odpowiednio n- i m-wymiarową przestrzenią wektorową nad ciałem K. Macierze $\mathbf{A}, \mathbf{B} \in K_{m\times n}$ są macierzami pewnego przekształcenia liniowego $T\colon V \to W$ (względem pewnych baz przestrzeni V i W) wtedy i tylko wtedy, gdy są one równoważne, tj. gdy istnieją nieosobliwe macierze $\mathbf{P}$ i $\mathbf{Q}$ takie, że*

$$\mathbf{B} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{P}. \tag{8.39}$$

Dowód. Pierwsza część jest prostą konsekwencją twierdzenia 8.7.1.

Dla dowodu drugiej części załóżmy, że $\mathbf{P}$ i $\mathbf{Q}$ są nieosobliwymi macierzami takimi, że $\mathbf{B} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{P}$. Niech $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ będzie ustaloną bazą przestrzeni V, a $C = (\mathbf{c}_1, \ldots, \mathbf{c}_m)$ – ustaloną bazą przestrzeni W. Pokażemy teraz jak za pomocą macierzy $\mathbf{A} = [a_{ij}]$, $\mathbf{P} = [p_{ij}]$ i $\mathbf{Q} = [q_{ij}]$ (oraz baz B i C) określić przekształcenie liniowe $T\colon V \to W$ oraz bazę B' przestrzeni V i bazę C' przestrzeni W takie, że $\mathbf{B} = [T]_{C'}^{B'} = [I_W]_{C'}^{C}[T]_C^B[I_V]_B^{B'} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{P}$.

Niech $T\colon V \to W$ będzie przekształceniem liniowym takim, że dla $i = 1, \ldots, n$ jest $T(\mathbf{b}_i) = \sum_{k=1}^m a_{ki}\mathbf{c}_k$. (Istnienie takiego przekształcenia wynika z twierdzenia 8.1.1.) Jest oczywiste, że macierzą tak określonego przekształcenia T względem baz B i C jest macierz $\mathbf{A}$, czyli $[T]_C^B = \mathbf{A}$.

Weźmy teraz pod uwagę wektory $\mathbf{b}'_1, \ldots, \mathbf{b}'_n$, gdzie $\mathbf{b}'_i = \sum_{k=1}^n p_{ki}\mathbf{b}_k$ dla $i = 1, \ldots, n$. Macierz $\left[\,[\mathbf{b}'_1]_B \ \ldots \ [\mathbf{b}'_n]_B\,\right]$ jest identyczna z nieosobliwą macierzą $\mathbf{P}$ i z tego faktu wynika, że uporządkowany zbiór $B' = (\mathbf{b}'_1, \ldots, \mathbf{b}'_n)$ jest bazą przestrzeni V. Macierzą przejścia od bazy B' do bazy B jest $[I_V]_B^{B'} = [p_{ij}] = \mathbf{P}$. Z tych samych powodów uporządkowany zbiór $C' = (\mathbf{c}'_1, \ldots, \mathbf{c}'_m)$, w którym $\mathbf{c}'_i = \sum_{k=1}^m q_{ki}\mathbf{c}_k$ (dla $i = 1, \ldots, m$), jest bazą przestrzeni W i $[I_W]_C^{C'} = \mathbf{Q}$. Stąd i z poprzedniego twierdzenia wynika, że dla macierzy przekształcenia T mamy $[T]_{C'}^{B'} = [I_W]_{C'}^{C}\,[T]_C^B\,[I_V]_B^{B'} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{P}$. □

W analogiczny sposób dowodzi się prawdziwość następującego twierdzenia.

Twierdzenie 8.7.3. *Niech V będzie n-wymiarową przestrzenią wektorową nad ciałem K. Dla macierzy $\mathbf{A}, \mathbf{B} \in K_{n\times n}$ istnieje operator liniowy $T\colon V \to V$ oraz bazy B i B' przestrzeni V takie, że*

$$\mathbf{A} = [T]_B \quad i \quad \mathbf{B} = [T]_{B'}$$

wtedy i tylko wtedy, gdy macierze $\mathbf{A}$ i $\mathbf{B}$ są podobne, tj. gdy istnieje macierz nieosobliwa $\mathbf{P}$ taka, że

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}. \tag{8.40}$$

8.8. Ćwiczenia

1. Zbadać liniowość następujących przekształceń:
   (*a*) $T\colon R^2 \to R^2$, $T(x, y) = (x + y, 3x)$;
   (*b*) $T\colon R^2 \to R^2$, $T(x, y) = (x, -x)$;
   (*c*) $T\colon R^2 \to R^2$, $T(x, y) = (x + 1, x + y)$;
   (*d*) $T\colon R^3 \to R^2$, $T(x, y, z) = (z, 2y - x)$;
   (*e*) $T\colon R^3 \to R^3$, $T(x, y, z) = (0, x, x + y + z)$;
   (*f*) $T\colon R^4 \to R^3$, $T(x, y, z, t) = (0, xy, z - t)$;
   (*g*) $T\colon R^3 \to R^2$, $T(x, y, z) = (x, y - 2z + 3)$;
   (*h*) $T\colon R^3 \to R^2$, $T(x, y, z) = (0, \sqrt{3}(x + z))$;
   (*i*) $T\colon R^3 \to R^2$, $T(x, y, z) = (x^2, y + z - x)$;
   (*j*) $T\colon R^3 \to R^3$, $T(x, y, z) = (x, x + y, x + y)$;
   (*k*) $T\colon R^R \to R^R$, $T(f) = -f$;
   (*l*) $T\colon R^R \to R$, $T(f) = \big(f(1)\big)^2$;
   (*m*) $T\colon R_2[x] \to R_3[x]$, $T\big(\varphi(x)\big) = x\varphi(x)$;
   (*n*) $T\colon R[x] \to R[x]$, $T\big(\varphi(x)\big) = \int_0^x \varphi(t)\,dt$;
   (*o*) $T\colon R[x] \to R$, $T(\varphi(x)) = \varphi(0)$.
2. Dana jest macierz $\mathbf{B} \in R_{2\times 2}$. Stwierdzić, czy funkcja $T\colon R_{2\times 2} \to R_{2\times 2}$ jest przekształceniem liniowym, gdy dla każdej macierzy $\mathbf{A} \in R_{2\times 2}$ jest: (*a*) $T(\mathbf{A}) = \mathbf{A}\mathbf{B}$; (*b*) $T(\mathbf{A}) = (\det \mathbf{A})\mathbf{B}$; (*c*) $T(\mathbf{A}) = \mathbf{B}\mathbf{A}$; (*d*) $T(\mathbf{A}) = (\det \mathbf{A})\mathbf{A}$; (*e*) $T(\mathbf{A}) = (\det \mathbf{B})(\mathbf{A} + \mathbf{A}^T)$.
3. Niech $T\colon R^2 \to R^2$ będzie przekształceniem liniowym takim, że $T(0, 1) = (4, 3)$ i $T(1, 3) = (5, 9)$. Wyznaczyć $T(2, 5)$ i $T^3(-1, 4)$. Czy przekształcenie T jest różnowartościowe?
4. Przekształcenie liniowe $T\colon R^2 \to R^2$ jest takie, że $T(2, 3) = (4, 6)$ i $T(4, -1) = (-8, 2)$. Wektor $(-5, 3)$ przedstawić jako kombinację liniową wektorów $(2, 3)$ oraz $(4, -1)$ i obliczyć $T^{50}(-5, 3)$.

5. Niech $T\colon R^3 \to R^3$ będzi przekształceniem liniowym takim, że $T(x, y, z) = (x + y - z, x + 3y - z, -3x - y + z)$. Wyznaczyć bazę jądra i bazę obrazu przekształcenia T.
6. Przekształcenie liniowe $T\colon R^4 \to R^4$ określone jest wzorem $T(x, y, z, t) = (x+3y+2z+3t, y-z+2t, 2x+5y+5z+5t, x+4y+z+4t)$. Wyznaczyć bazy i wymiary przestrzeni $\operatorname{Ker} T$ i $\operatorname{Im} T$.
7. Dane jest przekształcenie liniowe $T\colon R^4 \to R^2$, gdzie $T(\mathbf{x}) = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 2 & 1 & 1 \end{bmatrix}\mathbf{x}$ dla $\mathbf{x} \in R^4$. Wyznaczyć: (a) bazę przestrzeni $\operatorname{Ker} T$ i (b) wymiar przestrzeni $\operatorname{Im} T$.
8. Niech $T\colon R_2[x] \to R_3[x]$ będzie przekształceniem takim, że $T(\varphi(x)) = x\varphi(x)$ dla $\varphi(x) \in R_2[x]$. Znaleźć obraz i jądro przekształcenia T. Wyznaczyć także rząd i zerowość przekształcenia T.
9. Dane jest przekształcenie $T\colon R_2[x] \to R_3[x]$, gdzie $$T(\varphi(x)) = \varphi'(x) + 6\int_0^x \varphi(t)dt.$$ (a) Wyznaczyć $\operatorname{Im} T$ i $\operatorname{Ker} T$. (b) Zbadać mono- i epimorficzność przekształcenia T.
10. Wskazać przekształcenie liniowe $T\colon R^3 \to R^3$, dla którego $\operatorname{Ker} T = \{(x, y, z) \in R^3\colon\ x + 2y - z = 0\}$ oraz $\operatorname{Im} T = \big\{(x, y, z) \in R^3\colon\ x + y - z = 0 \text{ i } x + 3y + z = 0\big\}$. Czy istnieje tylko jedno takie przekształcenie? Uzasadnić swoją odpowiedź.
11. Uzasadnić, że nie istnieje przekształcenie liniowe $T\colon R^3 \to R^3$ takie, że $\operatorname{Ker} T = \{(x, y, z) \in R^3\colon x + y + z = 0\}$ i $\operatorname{Im} T = \{(x, y, z) \in R^3\colon x - y + 2z = 0\}$.
12. Dane jest przekształcenie $T\colon R^3 \to R^3$, gdzie $T(x, y, z) = (x + 2y - 2z, 3x + 3z, 2x - 2y + 5z)$. (a) Pokazać, że $\operatorname{Im} T$ jest płaszczyzną i znaleźć równanie tej płaszczyzny. (b) Pokazać, że T przekształca wszystkie punkty prostej $\frac{x-1}{-2} = \frac{y+1}{3} = \frac{z}{2}$ w jeden punkt i znaleźć ten punkt. (c) Pokazać, że wszystkie punkty, których obrazem jest punkt $(0, 0, 0)$ leżą na pewnej prostej. Wyznaczyć równanie tej prostej.
13. Dane jest przekształcenie $T\colon R^3 \to R^3$, gdzie $$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & -2 \\ 3 & -2 & -7 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$ (a) Pokazać, że $T(R^3)$ jest płaszczyzną o równaniu $x - 2y + z = 0$. (b) Wyznaczyć obrazy prostej $x = -y = z/2$ oraz płaszczyzny $x - y - z = 0$ poprzez przekształcenie T.
14. Niech V i W będą przestrzeniami wektorowymi i niech ich bazami będą odpowiednio $B = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ i $C = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4)$. Znaleźć macierz $[T]_C^B$ przekształcenia liniowego $T\colon V \to W$ względem baz B i C, gdy $T(\mathbf{b}_1) = 2\mathbf{c}_1 + \mathbf{c}_2 + 3\mathbf{c}_3 - \mathbf{c}_4$, $T(\mathbf{b}_2) = \mathbf{c}_1 + 2\mathbf{c}_2 - \mathbf{c}_3 + 2\mathbf{c}_4$ i $T(\mathbf{b}_3) = -\mathbf{c}_1 + \mathbf{c}_2 + \mathbf{c}_4$.
15. Przekształcenie liniowe $T\colon R^4 \to R^4$ jest takie, że $T(\mathbf{e}_1) = (1, 3, 1, 5)$, $T(\mathbf{e}_2) = (2, 5, 1, 13)$, $T(\mathbf{e}_3) = (1, 2,\ 0, 8)$ i $T(\mathbf{e}_4) = (2, 6, 3, 12)$, gdzie $E = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$ jest bazą standardową przestrzeni R^4. Wyznaczyć: (a) $\operatorname{Ker} T$ i bazę tej podprzestrzeni; (b) $\operatorname{Im} T$ i $\dim \operatorname{Im} T$; (c) $[T]_E$ oraz $\det [T]_E$.
16. Dane jest przekształcenie liniowe $T\colon R^3 \to R^3$ takie, że $T(x, y, z) = (z, 0, x)$. Wyznaczyć jego macierz $[T]_C^B$, gdy $B = \big((3, 1, 2),\ (1, 2, 1),\ (2, -1, 0)\big)$ i $C = \big((1, 2, 1),\ (2, 1, -1), (5, 4, 1)\big)$.
17. Niech V i W będą przestrzeniami wektorowymi i niech ich bazami będą odpowiednio $B = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4)$ i $C = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3)$. Znaleźć $T(\mathbf{b}_1 - \mathbf{b}_2 + 3\mathbf{b}_3)$, jeśli macierzą przekształcenia liniowego $T\colon V \to W$ względem baz B i C jest: $$(a)\ \begin{bmatrix} 2 & 4 & 1 & -1 \\ 3 & 0 & 0 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix};\quad (b)\ \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$
18. Wyznaczyć macierz przekształcenia liniowego $T\colon R^n \to R^m$ względem bazy B przestrzeni R^n i bazy C przestrzeni R^m, gdy:
    (a) $T(x, y) = (x - 2y, 2x + 3y)$, $B = \big((1, 1),\ (1, 0)\big)$ i $C = \big((0, 1), (1, 0)\big)$;
    (b) $T(x, y) = (2x + y, x + y, x - 3y)$, $B = \big((1, 0), (0, 1)\big)$ i $C = \big((1, 1, 1), (1, 1, 0), (1, 0, 0)\big)$;
    (c) $T(x, y, z) = (x + 2y + 3z, x, y + z)$, $B = \big((1,\ 0, 1), (1, 1, 0), (0, 1, 1)\big)$ i $C = \big((0,\ 1, 1), (1, 1, 0), (1, 0, 1)\big)$;
    (d) $T(x, y, z) = (x, y, z)$, $B = \big((1, 1, 0), (-1, 1, 0), (0, 1, 2)\big)$ i $C = \big((1, 0, 0),\ (0, 1, 0), (0, 0, 1)\big)$;
    (e) $T(x, y) = (x + y, x - 2y)$, $B = C = \big((1, 1), (1, -2)\big)$.
19. Wyznaczyć macierz przekształcenia liniowego $T\colon R^3 \to R^3$ względem bazy standardowej $E = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ przestrzeni R^3, jeśli $T(\mathbf{a}_i) = \mathbf{b}_i$ dla $i = 1, 2, 3$, gdy:
    (a) $\mathbf{a}_1 = (2, 3, 5)$, $\mathbf{a}_2 = (0, 1, 2)$, $\mathbf{a}_3 = (1, 0, 1)$, $\mathbf{b}_1 = (1, 1, 1)$, $\mathbf{b}_2 = (1, 1, -1)$, $\mathbf{b}_3 = (2, 1, 2)$;
    (b) $\mathbf{a}_1 = (2, 0, 3)$, $\mathbf{a}_2 = (4, 1, 5)$, $\mathbf{a}_3 = (3, 1, 2)$, $\mathbf{b}_1 = (1, 2, -1)$, $\mathbf{b}_2 = (4, 5, -2)$, $\mathbf{b}_3 = (1, -1, 1)$.
20. Niech $T\colon R_4[x] \to R_4[x]$ będzie funkcją taką, że $T(f(x)) = x^2 f''(x) - (2x+2)f'(x) + 2f(x)$ dla $f(x) \in R_4[x]$. (a) Pokazać, że T jest przekształceniem liniowym. (b) Wyznaczyć jądro i obraz przekształcenia T oraz ich wymiary. (c) Wyznaczyć macierz $[T]_B$ przekształcenia T względem bazy $B = (1, x, x^2, x^3, x^4)$ przestrzeni $R_4[x]$. (d) Wyznaczyć rząd macierzy $[T]_B$ i wymiar przestrzeni zerowej macierzy $[T]_B$.
21. Niech $B = \big((1, 1, 0), (0, 1, 0), (1, 0, 2)\big)$, $C = \big((0, 1),\ (1, 1)\big)$ i $D = \big((1, 0, 0, 1), (0, 1, 1, 0), (0, 0, 1, 1), (1, 0, 0, 0)\big)$ będą odpowiednio bazami przestrzeni R^3, R^2 i R^4. Niech $T\colon R^3 \to R^2$ i $U\colon R^2 \to R^4$ będą funkcjami takimi, że $T(x, y, z) = (x + z, y - z)$ i $U(u, v) = (u, u + v, v, v - u)$. (a) Znaleźć $\big(UT\big)(0, 1, 1)$. (b) Pokazać, że T, U i UT są przekształceniami liniowymi. (c) Utworzyć macierze $[T]_C^B$, $[U]_D^C$ i $[UT]_D^B$ przekształceń T, U i UT. (d) Wyznaczyć:
    (1) $\dim \operatorname{Ker} T$, $\dim \operatorname{Im} T$ i bazę przestrzeni $\operatorname{Ker} T$;
    (2) $\dim \operatorname{Ker} U$, $\dim \operatorname{Im} U$ i bazę przestrzeni $\operatorname{Im} U$ oraz
    (3) $\dim \operatorname{Ker} UT$ i $\dim \operatorname{Im} UT$.

22. Niech $T\colon R_3[x] \to R_3[x]$ będzie przekształceniem liniowym takim, że $T(\varphi(x)) = \varphi''(x) - \varphi'(x) + 2\varphi(x)$ dla $\varphi(x) \in R_3[x]$. Znaleźć macierz $[T]_B$ przekształcenia T względem bazy $B = (x, 1+x, x+x^2, x^3)$ przestrzeni $R_3[x]$.
23. W przestrzeni wektorowej $R_n[x]$ wielomianów stopnia co najwyżej n dana jest funkcja $T\colon R_n[x] \to R_n[x]$ taka, że $T(f(x)) = f'(x)$. (a) Pokazać, że T jest przekształceniem liniowym. (b) Wyznaczyć jądro i obraz przekształcenia T. (c) Wyznaczyć macierz przekształcenia T względem bazy $(1, x, \ldots, x^n)$ przestrzeni $R_n[x]$. (d) Czy przekształcenie T jest różnowartościowe?
24. Dane jest przekształcenie $T : R_n[x] \to R_n[x]$, gdzie $T(f(x)) = (x^2 f(x))''$. (a) Pokazać, że T jest przekształceniem liniowym. (b) Znaleźć macierz przekształcenia T względem bazy $(1, x, \ldots, x^n)$ przestrzeni $R_n[x]$. (c) Pokazać, że przekształcenie T jest epimorfizmem.
25. Przekształcenie liniowe $T\colon R_2[x] \to R_{2\times 2}$ jest takie, że $T(\varphi(x)) = \begin{bmatrix} \varphi(0) & 2\varphi'(1) \\ 3\varphi''(2) & 0 \end{bmatrix}$. Wyznaczyć macierz $[T]_C^B$, gdy $B = (1, x, x^2)$ i $C = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right)$.
26. Przekształcenie liniowe $T\colon R_{2\times 2} \to R_{2\times 2}$ jest takie, że $T\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right) = T\left(\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \right) = \begin{bmatrix} 1 & 3 \\ 1 & 0 \end{bmatrix}$ i $T\left(\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \right) = T\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \right) = \begin{bmatrix} 2 & 0 \\ 3 & 0 \end{bmatrix}$. Wyznaczyć macierz przekształcenia T względem bazy standardowej E przestrzeni $R_{2\times 2}$. Dodatkowo wyznaczyć $T\left(\begin{bmatrix} x & y \\ z & t \end{bmatrix} \right)$.
27. Znaleźć macierz przejścia $\mathbf{P} = [I_V]_C^B$ od bazy B do bazy C, gdy: (a) $B = \big((1,2,0), (3,4,2), (2,2,1) \big)$, $C = \big((1,-1,2), (3,1,-1), (4,0,2) \big)$ i $V = R^3$; (b) $B = (x^3, x^2, x, 1)$, $C = (x^3 - x^2, x^2 - x, x - 1, x^3 + 1)$ i $V = R_3[x]$.
28. Macierz $\mathbf{A}$ jest macierzą operatora liniowego T względem bazy standardowej przestrzeni R^n. Znaleźć macierz operatora T względem bazy B, gdy:
(a) $\mathbf{A} = \begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix}$, $B = \left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$;
(b) $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 1 & 3 & 2 \end{bmatrix}$; $B = \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right)$;
(c) $\mathbf{A} = \begin{bmatrix} 7 & 2 \\ 1 & 2 \end{bmatrix}$; $B = \left(\begin{bmatrix} -1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right)$;
(d) $\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 7 & 0 \\ 2 & 3 & 14 \end{bmatrix}$, $B = \left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right)$.
29. Macierzą operatora liniowego $T\colon R^3 \to R^3$ względem bazy $A = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ jest
$$[T]_A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & -1 \\ 2 & 5 & 3 \end{bmatrix}.$$
Wyznaczyć macierze $[T]_B$, $[T]_C$ i $[T]_D$ względem baz $B = (\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_2)$, $C = (\mathbf{b}_1, \mathbf{b}_1 + \mathbf{b}_2, \mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3)$ i $D = (2\mathbf{b}_1 + 3\mathbf{b}_2 + \mathbf{b}_3, 3\mathbf{b}_1 + 4\mathbf{b}_2 + \mathbf{b}_3, \mathbf{b}_1 + 2\mathbf{b}_2 + 2\mathbf{b}_3)$.
30. Dana jest funkcja $T\colon R_2[x] \to R_2[x]$ taka, że $T\big(\varphi(x)\big) = \varphi(x+1) + \varphi(x)$ dla $\varphi(x) \in R_2[x]$. (a) Wyznaczyć macierz $[T]_C^B$ względem baz $B = (x^2, x, 1)$ i $C = (x^2+1, x+1, 2)$. (b) Obliczyć det $[T]_C^B$. (c) Czy przekształcenie T jest odwracalne?
31. Dane jest przekształcenie $T\colon R_3[x] \to R_3[x]$ takie, że $T\big(\varphi(x)\big) = \varphi(x) - x\varphi''(x)$. Macierzowo uzasadnić odwracalność przekształcenia T i obliczyć $T^{-1}(x^2)$.
32. Dana jest przestrzeń wektorowa V z bazą $B = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ i przekształcenie liniowe $T\colon V \to V$ takie, że $T(\mathbf{b}_1) = \mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3$, $T(\mathbf{b}_2) = \mathbf{b}_1 + 2\mathbf{b}_2 - \mathbf{b}_3$ i $T(\mathbf{b}_3) = \mathbf{b}_1 - \mathbf{b}_2 + 3\mathbf{b}_3$. Wyznaczyć macierz $[T]_B$ przekształcenia T względem bazy B. Czy przekształcenie T jest odwracalne? Obliczyć $T^{-1}(\mathbf{b}_1)$.
33. Macierzą przekształcenia liniowego $T\colon V \to V$ względem bazy $B = (\mathbf{b}_1, \mathbf{b}_2)$ przestrzeni V jest $[T]_B = \begin{bmatrix} 1 & 1 \\ 3 & 5 \end{bmatrix}$. Wyznaczyć $T^2(\mathbf{b}_1)$ i $T^{-1}(\mathbf{b}_2)$.
34. Macierzą przekształcenia liniowego $T\colon R^3 \to R^3$ względem bazy $B = \big((1,2,3), (2,3,1), (4,5,2) \big)$ jest macierz
$$[T]_B = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \\ 2 & -1 & 1 \end{bmatrix}.$$
Wyznaczyć macierze $[T]_C$ i $[T^{-1}]_C$, gdy $C = \big((1,2,3), (0,1,2), (0,0,1) \big)$.
35. Przekształcenie $T\colon R^3 \to R^3$ jest symetrią względem płaszczyzny $y = 0$. Podać macierze $[T]_E$ i $[T]_B$ przekształcenia T względem bazy kanonicznej E i względem bazy $B = \big((1,1,1), (0,1,1), (0,0,1) \big)$ przestrzeni R^3.
36. Przekształcenie liniowe $T\colon R^3 \to R^3$ jest symetrią względem prostej określonej równaniami $x = y$ i $z = 0$. Wyznaczyć macierz przekształcenia T względem bazy standardowej przestrzeni R^3 i wyznaczyć $T(x, y, z)$.
37. Przekształcenie $T\colon R^3 \to R^3$ jest obrotem o kąt $\pi/2$ wokół osi Ox, zaś przekształcenie $U\colon R^3 \to R^3$ określone jest wzorem $U(x, y, z) = (x, y - x, z + x - y)$. Wyznaczyć macierze przekształceń T, U, UT i TUT względem bazy standardowej przestrzeni R^3.
38. Niech T będzie obrotem w przestrzeni R^3 o kąt $\frac{3}{2}\pi$ wokół osi Oz, zaś U niech będzie symetrią względem płaszczyzny Oxz. Wyznaczyć macierz przekształcenia UT względem bazy standardowej przestrzeni R^3.
39. Zbadać odwracalność przekształcenia liniowego T i, jeśli to możliwe, wyznaczyć przekształcenie odwrotne, gdy $T\colon R^3 \to R^3$ i $T(x, y, z) = (x + y - z, x - 2y + z, -3x + y + z)$.

40. Zbadać, które z przekształceń liniowych T jest izomorfizmem przestrzeni wektorowych:
(a) $T : R^2 \to R^2$, $T(x,y) = (x+y, x)$;
(b) $T : R^2 \to R^2$, $T(x,y) = (x-y, x-y)$;
(c) $T : R^4 \to R^4$, $T(x,y,z,t) = (x,y,t,t-z)$.
41. Wyznaczyć wymiar przestrzeni wektorowej
$$V = \left\{ \begin{bmatrix} a & a+b \\ 0 & b+c \end{bmatrix} : a,\, b,\, c \in R \right\}.$$
Wskazać izomorfizm przestrzeni V i R^3.
42. Zbadać liniowość i różnowartościowość przekształcenia $T\colon R_{n\times n} \to R_{n\times n}$ takiego, że $T(\mathbf{A}) = \mathbf{A}^T + \mathbf{A}$.
43. Pokazać, że każda podprzestrzeń wektorowa W przestrzeni V jest jądrem pewnego przekształcenia liniowego $T : V \to V$.
44. Pokazać, że przekształcenie odwrotne odwracalnego przekształcenia liniowego odwracalnym przekształceniem liniowym.
45. Niech $T\colon V \to W$ będzie monomorfizmem przestrzeni V i W. Wykazać, że wektory $\mathbf{b}_1, \ldots, \mathbf{b}_n$ są liniowo niezależne w przestrzeni V wtedy i tylko wtedy, gdy ich obrazy $T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)$ są liniowo niezależne w przestrzeni W.
46. Wykazać, że jeśli T_1, T_2 i T_3 są operatorami liniowymi na przestrzeni wektorowej V, to: (a) $T_1(T_2+T_3) = T_1T_2 + T_1T_3$ i $(T_1+T_2)T_3 = T_1T_3 + T_2T_3$; (b) $T_1(T_2T_3) = (T_1T_2)T_3$; (c) $a(T_1T_2) = (aT_1)T_2 = T_1(aT_2)$ dla każdego skalara a; (d) $TI_V = I_VT = T$, gdzie I_V jest przekształceniem tożsamościowym na przestrzeni V.
47. Niech V i W będą przestrzeniami wektorowymi (nad tym samym ciałem), niech S będzie podzbiorem zbioru V i $S^0 = \{T \in L(V,W)\colon T(\mathbf{x}) = \mathbf{0}$ dla każdego $\mathbf{x} \in S\}$. Udowodnić następujące stwierdzenia: (a) S^0 jest podprzestrzenią przestrzeni $L(V,W)$; (b) Jeśli $S_1 \subseteq S_2 \subset V$, to $S_2^0 \subseteq S_1^0$; (c) Jeśli V_1 i V_1 są podprzestrzeniami przestrzeni V, to $(V_1+V_2)^0 = V_1^0 \cap V_2^0$.
48. Niech $T\colon V \to V$ będzie operatorem liniowym na przestrzeni wektorowej V. Udowodnić, że $T^2 = 0$ wtedy i tylko wtedy, gdy $\operatorname{Im} T \subseteq \operatorname{Ker} T$.
49. Podać przykład przekształceń liniowych $T, U\colon R^2 \to R^2$ takich, że $UT = 0$ i $TU \neq 0$. Wskazać także macierze $\mathbf{A}$ i $\mathbf{B}$ takie, że $\mathbf{AB} = \mathbf{0}$ i $\mathbf{BA} \neq \mathbf{0}$. (Uzasadnić swoje stwierdzenia.)
50. Podać przykład przekształcenia liniowego $T\colon R^2 \to R^2$ takiego, że $\operatorname{Ker} T = \operatorname{Im} T$. (Uzasadnić swój wybór.)
51. Podać przykład różnych przekształceń liniowych $T, S : R^2 \to R^2$ takich, że $\operatorname{Ker} T = \operatorname{Ker} S$ i $\operatorname{Im} T = \operatorname{Im} S$. (Uzasadnić swój wybór.)
52. Dane jest przekształcenie liniowe $T\colon R^3 \to R^3$, gdzie $T(x,y,z) = (x+y, y+z, x-z)$. Wskazać podprzestrzenie V, W, Z i X przestrzeni R^3, dla których: (a) $\dim T(V) < \dim V$; (b) $\dim T(W) = \dim W$; (c) $\dim T^{-1}(Z) > \dim Z$; (d) $\dim T^{-1}(X) = \dim X$.
53. Wykazać, że macierze $\mathbf{A}$ i $\mathbf{B}$ są równoważne wtedy i tylko wtedy, gdy macierz $\mathbf{B}$ można otrzymać z macierzy $\mathbf{A}$ za pomocą operacji elementarnych na wierszach i kolumnach macierzy.
54. Wyznaczyć wszystkie macierze podobne do macierzy jednostkowej $\mathbf{I}_n$.
55. Wyznaczyć wszystkie macierze równoważne z macierzą jednostkową $\mathbf{I}_n$.
56. Wykazać, że macierze $\begin{bmatrix} 1 & 0 & 4 \\ 1 & 1 & 3 \\ 2 & 1 & 7 \end{bmatrix}$ i $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 3 & 1 & 2 \end{bmatrix}$ nie są podobne.
57. Wykazać, że jeśli macierze $\mathbf{A}$ i $\mathbf{B}$ są podobne i macierz $\mathbf{A}$ jest odwracalna, to także macierz $\mathbf{B}$ jest odwracalna.
58. Wykazać, że jeśli macierze $\mathbf{A}$ i $\mathbf{B}$ są podobne, to także macierze $\mathbf{A}^2$ i $\mathbf{B}^2$ są podobne.
59. Wykazać, że jeśli macierze $\mathbf{A}$ i $\mathbf{B}$ są podobne, to także macierze $a_n\mathbf{A}^n + a_{n-1}\mathbf{A}^{n-1} + \ldots + a_1\mathbf{A} + a_0\mathbf{I}$ i $a_n\mathbf{B}^n + a_{n-1}\mathbf{B}^{n-1} + \ldots + a_1\mathbf{B} + a_0\mathbf{I}$ są podobne dla każdych skalarów $a_0, \ldots, a_n$.
60. Wykazać, że jeśli macierze $\mathbf{A} - \lambda\mathbf{I}$ i $\mathbf{B}$ są podobne, to także macierze $\mathbf{A}$ i $\mathbf{B} + \lambda\mathbf{I}$ są podobne.
61. Niech $T\colon V \to W$ będzie przekształceniem liniowym przestrzeni wektorowej V w przestrzeń wektorową W, gdzie $\dim V = \dim W < \infty$. Pokazać, że istnieje baza B przestrzeni V i baza C przestrzeni W taka, że macierz $[T]_C^B$ jest diagonalna.
62. Niech $\mathbf{A}$ i $\mathbf{B}$ będą macierzami podobnymi wymiaru $n \times n$. Udowodnić, że istnieje przestrzeń wektorowa V, operator liniowy $T : V \to V$ i bazy B oraz C przestrzeni V takie, że $\mathbf{A} = [T]_B$ i $\mathbf{B} = [T]_C$.
63. Wykazać, że przekształcenie liniowe $T\colon V \to W$ jest monomorfizmem wtedy i tylko wtedy, gdy istnieje taki wektor $\mathbf{y} \in W$, że przeciwobraz $T^{-1}(\{\mathbf{y}\})$ jest zbiorem jednoelementowym.
64. Dane jest przekształcenie liniowe $T\colon V \to W$. (a) Wykazać, że T nie jest monomorfizmem, gdy $\dim V > \dim W$. (b) Wykazać, że T nie jest epimorfizmem, gdy $\dim V < \dim W$.
65. Niech $\mathbf{Q}$ będzie macierzą rzeczywistą wymiaru $n \times n$. Wykazać, że odwzorowanie $T\colon R_{n\times n} \to R_{n\times n}$ jest izomorfizmem, gdy $T(\mathbf{A}) = \mathbf{Q}^{-1}\mathbf{AQ}$ dla $\mathbf{A} \in R_{n\times n}$.
66. Niech V i W będą skończenie wymiarowymi przestrzeniami wektorowymi i $T \in L(V,W)$. Niech $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ będzie bazą przestrzeni V. Udowodnić, że T jest izomorfizmem wtedy i tylko wtedy, gdy $\big(T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)\big)$ jest bazą przestrzeni W.
67. Niech V, W i Z będą przestrzeniami wektorowymi i niech $T \in L(V,W)$ oraz $U \in L(W,Z)$. (a) Pokazać, że jeśli UT jest monomorfizmem, to T jest monomorfizmem. (b) Wykazać, że U jest epimorfizmem, gdy UT jest epimorfizmem. (c) Udowodnić, że UT jest izomorfizmem, gdy T i U są izomorfizmami.
68. Niech $T\colon V \to W$ będzie przekształceniem liniowym i niech V' i W' będą odpowiednio podprzestrzeniami przestrzeni V i W. Wykazać, że: (a) $\dim T(V') \leqslant \dim V'$; (b) $\dim T^{-1}(W') \geqslant \dim W'$.
69. Niech $T\colon V \to W$ będzie izomorfizmem skończenie wymiarowych przestrzeni wektorowych V i W. Wykazać, że jeśli V' jest podprzestrzenią przestrzeni V, to $\dim V' = \dim T(V')$.

70. Niech $T\colon R^2 \to R^2$ będzie przekształceniem liniowym takim, że $T(x,y) = (x+2y, 3x+2y)$ i niech $E = ((1,0),(0,1))$ oraz $B = ((1,3),(2,5))$ będą bazami przestrzeni R^2. (a) Wyznaczyć macierze $[T]_E$ i $[T]_B$. (b) Znaleźć $[\mathbf{v}]_B$ i $[T(\mathbf{v})]_B$, gdy $\mathbf{v} = (1,1)$. (c) Wyznaczyć macierze przejścia $[1_{R^2}]_B^E$ i $[1_{R^2}]_E^B$. (d) Sprawdzić, czy $[T]_B = [1_{R^2}]_B^E [T]_E [1_{R^2}]_E^B$. (e) Znaleźć bazę C przestrzeni R^2 taką, że macierz $[T]_C^C$ jest diagonalna. (f) Wskazać macierze $\mathbf{P} = [1_{R^2}]_E^C$ i $\mathbf{P}^{-1}$. (g) Wskazać formułę dla obliczeń potęgi $\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}^n$, gdzie n jest liczbą naturalną.

71. Niech $T, U\colon V \to W$ będą niezerowymi przekształceniami liniowymi takimi, że $\operatorname{Im} T \cap \operatorname{Im} U = \{\mathbf{0}\}$. Udowodnić,że T i U są liniowo niezależne w przestrzeni $L(V,W)$.

72. Niech V i W będą n-wymiarowymi przestrzeniami wektorowymi i niech B i C będą odpowiednio ich bazami. Udowodnić, że dla przekształcenia liniowego $T\colon V \to W$ następujące stwierdzenia są równoważne: (a) T jest monomorfizmem; (b) $\operatorname{Ker} T = \{\mathbf{0}\}$; (c) $\dim \operatorname{Ker} T = 0$; (d) $\dim \operatorname{Im} T = n$; (e) $\operatorname{Im} T = W$; (f) T jest epimorfizmem; (g) macierz $[T]_C^B$ jest odwracalna; (h) układ wektorów $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ jest bazą przestrzeni V wtedy i tylko wtedy, gdy układ $\big(T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)\big)$ jest bazą przestrzeni W.

73. Wpisując TAK albo NIE, stwierdzić prawdziwość każdego z następujących zdań:

$\boxed{1}$ Przekształcenie $T \in L(V,W)$ jest różnowartościowe wtedy i tylko wtedy, gdy $\operatorname{Ker} T = \{\mathbf{0}\}$.

$\boxed{2}$ Przekształcenie $T\colon R_{n\times n} \to R$ jest liniowe, gdy $T(\mathbf{A}) = \operatorname{tr}(\mathbf{A})$.

$\boxed{3}$ Jeśli $T \in L(V,W)$ i wektory $\mathbf{b}_1, \ldots, \mathbf{b}_n$ są liniowo niezależne w przestrzeni V, to wektory $T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)$ są liniowo niezależne w przestrzeni W.

$\boxed{4}$ Jeśli $T \in L(V,W)$ i $\mathcal{L}(\mathbf{b}_1, \ldots, \mathbf{b}_n) = V$, to $\mathcal{L}\big(T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)\big) = W$.

$\boxed{5}$ Jeśli $T \in L(V,W)$ i wektory $\mathbf{b}_1, \ldots, \mathbf{b}_n$ są liniowo zależne w przestrzeni V, to wektory $T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)$ są liniowo zależne w przestrzeni W.

$\boxed{6}$ Jeśli $T \in L(V,W)$ i $\mathbf{b}_1, \ldots, \mathbf{b}_n \in V$ oraz układ wektorów $\big(T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)\big)$ jest bazą przestrzeni $\operatorname{Im} T$, to układ wektorów $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ jest bazą przestrzeni V.

$\boxed{7}$ Jeśli $T \in L(V,W)$ jest epimorfizmem, to $\dim V \geqslant \dim W$.

$\boxed{8}$ Jeśli V i W są przestrzeniami wektorowymi nad tym samym ciałem, to $L(V,W) = L(W,V)$.

$\boxed{9}$ Jeśli V i W są przestrzeniami wektorowymi nad tym samym ciałem, to $L(V,W) = L(W,V)$ wtedy i tylko wtedy, gdy $\dim V = \dim W$.

$\boxed{10}$ Jeśli $T \in L(V,W)$ oraz B i C są bazami przestrzeni wektorowych V i W, to $\big([T]_C^B\big)^{-1} = [T^{-1}]_C^B$.

$\boxed{11}$ Przestrzenie wektorowe $R_{3\times 2}$ i R^5 są izomorficzne.

$\boxed{12}$ Macierze $\mathbf{A}$ i $\mathbf{B}$ ze zbior $R_{n\times n}$ są podobne, gdy $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ dla pewnej macierzy $\mathbf{P} \in R_{n\times n}$.

$\boxed{13}$ Macierze podobne mają identyczne ślady.

$\boxed{14}$ Każda macierz przejścia jest odwracalna.

$\boxed{15}$ Macierz jednostkowa zawsze reprezentuje przekształcenie tożsamościowe.

Rozdział 9

ILOCZYN SKALARNY I ORTOGONALNOŚĆ WEKTORÓW

9.1. Definicja i przykłady iloczynów skalarnych

Definicja 9.1.1. Niech V będzie przestrzenią wektorową nad ciałem liczb rzeczywistych. Funkcję $(\cdot|\cdot)\colon V \times V \to R$, która każdej parze wektorów $\mathbf{x}$, $\mathbf{y} \in V$ przyporządkowuje liczbę $(\mathbf{x}|\mathbf{y}) \in R$, nazywamy *iloczynem skalarnym* w przestrzeni V, jeśli ma ona następujące własności:

Iloczyn skalarny

(S_1) $\forall_{\mathbf{x},\,\mathbf{y}\in V}$ $(\mathbf{x}|\mathbf{y}) = (\mathbf{y}|\mathbf{x})$; (symetria)

(S_2) $\forall_{\mathbf{x},\,\mathbf{y},\,\mathbf{z}\in V}\forall_{\alpha,\,\beta\in R}$ $(\mathbf{x}|\alpha\mathbf{y} + \beta\mathbf{z}) = \alpha(\mathbf{x}|\mathbf{y}) + \beta(\mathbf{x}|\mathbf{z})$; (liniowość)

(S_3) $\forall_{\mathbf{x}\in V}$ $(\mathbf{x}|\mathbf{x}) \geqslant 0$ i $(\mathbf{x}|\mathbf{x}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$. (dodatnia określoność)

Jeśli $(\cdot|\cdot)$ jest iloczynem skalarnym w przestrzeni V, a $\mathbf{x}$ i $\mathbf{y}$ są wektorami z przestrzeni V, to liczbę $(\mathbf{x}|\mathbf{y})$ nazywamy *iloczynem skalarnym wektorów* $\mathbf{x}$ i $\mathbf{y}$. Z własności (S_1) i (S_2) iloczynu skalarnego wynika, że dla każdych wektorów $\mathbf{x}, \mathbf{x}_1, \ldots, \mathbf{x}_n$, $\mathbf{y}_1, \ldots, \mathbf{y}_m$ z przestrzeni wektorowej V i każdych liczb rzeczywistych $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m$ jest

$$\Big(\mathbf{x}\Big|\sum_{j=1}^{m}\beta_j\mathbf{y}_j\Big) = \sum_{j=1}^{m}\alpha_i\beta_j(\mathbf{x}_i|\mathbf{y}_j) \tag{9.1}$$

oraz

$$\Big(\sum_{i=1}^{n}\alpha_i\mathbf{x}_i\Big|\sum_{j=1}^{m}\beta_j\mathbf{y}_j\Big) = \sum_{i=1}^{n}\sum_{j=1}^{m}\alpha_i\beta_j(\mathbf{x}_i|\mathbf{y}_j). \tag{9.2}$$

Definicja 9.1.2. Skończenie wymiarową przestrzeń wektorową V z iloczynem skalarnym $(\cdot|\cdot)$, czyli parę $\big(V, (\cdot|\cdot)\big)$, nazywamy *przestrzenią Euklidesa.*

Przestrzeń Euklidesa

Przykład 195. Niech V będzie n-wymiarową przestrzenią wektorową nad ciałem liczb rzeczywistych i niech B będzie bazą przestrzeni V. Wykażemy, że iloczynem skalarnym w przestrzeni V jest funkcja $(\cdot|\cdot) : V \times V \to R$, która wektorom $\mathbf{x}$, $\mathbf{y}$ z przestrzeni V przyporządkowuje liczbę

$$(\mathbf{x}|\mathbf{y}) = [\mathbf{x}]_B^T[\mathbf{y}]_B. \tag{9.3}$$

Formalnie iloczyn $[\mathbf{x}]_B^T[\mathbf{y}]_B$ jest macierzą wymiaru 1×1, którą utożsamiamy z jej jedynym elementem. Z własności transpozycji macierzy wynika, że dla każdych wektorów $\mathbf{x}$, $\mathbf{y}$ z przestrzeni V jest

$$(\mathbf{x}|\mathbf{y}) = [\mathbf{x}]_B^T[\mathbf{y}]_B = \big([\mathbf{x}]_B^T[\mathbf{y}]_B\big)^T = [\mathbf{y}]_B^T\left([\mathbf{x}]_B^T\right)^T = [\mathbf{y}]_B^T[\mathbf{x}]_B = (\mathbf{y}|\mathbf{x}).$$

To dowodzi, że funkcja $(\cdot|\cdot)$ ma własność (S_1). Własność (S_2) funkcji $(\cdot|\cdot)$ wynika z twierdzenia 7.6.1 oraz z własności iloczynu macierzy, bo dla każdych wektorów $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z} \in V$ i każdych liczb α, $\beta \in R$ mamy

$$\begin{aligned}(\mathbf{x}|\alpha\mathbf{y} + \beta\mathbf{z}) &= [\mathbf{x}]_B^T[\alpha\mathbf{y} + \beta\mathbf{z}]_B = [\mathbf{x}]_B^T\left(\alpha[\mathbf{y}]_B + \beta[\mathbf{z}]_B\right)\\ &= \alpha[\mathbf{x}]_B^T[\mathbf{y}]_B + \beta[\mathbf{x}]_B[\mathbf{z}]_B = \alpha(\mathbf{x}|\mathbf{y}) + \beta(\mathbf{x}|\mathbf{z}).\end{aligned}$$

W końcu zauważmy, że jeśli $\mathbf{x} \in V$ i $[\mathbf{x}]_B = (x_1, x_2, \ldots, x_n)$, to mamy $(\mathbf{x}|\mathbf{x}) = [\mathbf{x}]_B^T[\mathbf{x}]_B = \sum_{i=1}^n x_i^2 \geqslant 0$ i w ciele liczb rzeczywistych jest $(\mathbf{x}|\mathbf{x}) = \sum_{i=1}^n x_i^2 = 0$ wtedy i tylko wtedy, gdy $x_i = 0$ dla $i = 1, \ldots, n$, tj. wtedy i tylko wtedy, gdy $\mathbf{x} = \mathbf{0}$. Zatem funkcja $(\cdot|\cdot)$ określona wzorem (9.3) ma własność (S_3) i jest ona iloczynem skalarnym w przestrzeni V. Dlatego przestrzeń wektorowa V z tak określonym iloczynem skalarnym jest przestrzenią Euklidesa.

W przestrzeni wektorowej R^n ze standardową bazą $E = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ dla każdego wektora $\mathbf{x}$ jest $[\mathbf{x}]_E = \mathbf{x}$, więc z powyższego wynika, że iloczynem skalarnym w przestrzeni R^n jest funkcja $(\cdot|\cdot) : R^n \times R^n \to R$ przyporządkowująca wektorom $\mathbf{x} = (x_1, \ldots, x_n)$ i $\mathbf{y} = (y_1, \ldots, y_n)$ liczbę

Standardowy iloczyn skalarny

$$(\mathbf{x}|\mathbf{y}) = \mathbf{x}^T\mathbf{y} = x_1y_1 + x_2y_2 + \ldots + x_ny_n = \sum_{i=1}^n x_iy_i. \tag{9.4}$$

Ten iloczyn nazywa się *standardowym iloczynem skalarnym* w przestrzeni R^n.

Przykład 196. Z poprzedniego przykładu wynika, że w przestrzeni $R_3[x]$ z bazą $B = (1, x, x^2, x^3)$ iloczynem skalarnym jest funkcja $(\cdot|\cdot)$, która każdym wielomianom $\varphi(x) = a_0+a_1x+a_2x^2+a_3x^3$ i $\psi(x) = b_0+b_1x+b_2x^2+b_3x^3$ z przestrzeni $R_3[x]$ przyporządkowuje liczbę

$$(\varphi|\psi) = [\varphi(x)]_B^T[\psi(x)]_B = [\,a_0\,a_1\,a_2\,a_3\,]\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = a_0b_0 + a_1b_1 + a_2b_2 + a_3b_3.$$

Zatem przestrzeń $R_3[x]$ z iloczynem skalarnym $(\cdot|\cdot)$ jest przestrzenią Euklidesa i w tej przestrzeni przykładowo mamy

$$(x|1+x^2) = [\,0\,1\,0\,0\,]\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = 0 \text{ i } (-1+4x^2|-3+2x) = [\,-1\,0\,4\,0\,]\begin{bmatrix} -3 \\ 2 \\ 0 \\ 0 \end{bmatrix} = 3.$$

Przykład 197. Niech $V = C\big(\langle a;b\rangle\big)$ będzie przestrzenią rzeczywistych funkcji ciągłych na odcinku $\langle a;b\rangle$. Wykażemy, że iloczynem skalarnym w tej przestrzeni jest funkcja $(\cdot|\cdot) : V\times V \to R$, która funkcjom $f,\, g \in V$ przyporządkowuje liczbę

$$(f|g) = \int_a^b f(x)g(x)\,dx. \tag{9.5}$$

Korzystając z własności całki oznaczonej, z łatwością stwierdzamy, że tak określona funkcja ma własności (S_1) i (S_2). Oczywiście dla każdej funkcji $f \in V$ jest $(f|f) = \int_a^b f^2(x)\,dx \geqslant 0$. Zatem dla dowodu własności (S_3) wystarczy pokazać, że $(f|f) = \int_a^b f^2(x)\,dx > 0$ dla każdej niezerowej funkcji f ze zbioru V. Jeśli f jest niezerową funkcją ze zbioru V, to istnieje $x_0 \in \langle a;b\rangle$ takie, że $f^2(x_0) = p > 0$. Stąd i z ciągłości funkcji f wynika, że istnieje przedział $\langle c;d\rangle \subseteq \langle a;b\rangle$ zawierający x_0 i taki, że $f^2(x) \geqslant p/2$ dla każdego $x \in \langle c;d\rangle$. Dlatego jest

$$\begin{aligned}(f|f) &= \int_a^b f^2(x)\,dx = \int_a^c f^2(x)\,dx + \int_c^d f^2(x)\,dx + \int_d^b f^2(x)\,dx \\ &\geqslant \int_c^d f^2(x)\,dx \geqslant \int_c^d p/2\,dx > 0.\end{aligned}$$

To dowodzi, że funkcja $(\cdot|\cdot)$ określona wzorem (9.5) jest iloczynem skalarnym.

Przykład 198. Pokazać, że funkcja $f\colon R^2 \times R^2 \to R$ jest iloczynem skalarnym, jeśli dla każdych wektorów $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ z R^2 jest

$$f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{A} \mathbf{y} \quad \text{i} \quad \mathbf{A} = \begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix}.$$

Ponieważ macierz $\mathbf{A}$ jest symetryczna ($\mathbf{A}^T = \mathbf{A}$), więc dla każdych wektorów $\mathbf{x}$, $\mathbf{y} \in R^2$ jest

$$f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{A} \mathbf{y} = (\mathbf{x}^T \mathbf{A} \mathbf{y})^T = \mathbf{y}^T \mathbf{A}^T \mathbf{x} = \mathbf{y}^T \mathbf{A} \mathbf{x} = f(\mathbf{y}, \mathbf{x})$$

i funkcja f ma własność (S_1). Funkcja f ma także własność (S_2), bo dla każdych wektorów $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z} \in R^2$ i liczb rzeczywistych α, β jest

$$f(\mathbf{x}, \alpha\mathbf{y} + \beta\mathbf{z}) = \mathbf{x}^T \mathbf{A}(\alpha\mathbf{y} + \beta\mathbf{z}) = \alpha\mathbf{x}^T \mathbf{A}\mathbf{y} + \beta\mathbf{x}^T \mathbf{A}\mathbf{z} = \alpha f(\mathbf{x}, \mathbf{y}) + \beta f(\mathbf{x}, \mathbf{z}).$$

Dla dowodu własności (S_3) zauważmy, że dla każdego wektora $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ jest

$$\begin{aligned} f(\mathbf{x}, \mathbf{x}) &= [\, x_1 \, x_2 \,] \begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= x_1^2 - 2x_1x_2 + 3x_2^2 = (x_1 - x_2)^2 + 2x_2^2 \geqslant 0. \end{aligned}$$

Stąd też widać, że $f(\mathbf{x}, \mathbf{x}) = 0$ wtedy i tylko wtedy, gdy $\mathbf{x} = (x_1, x_2) = (0, 0)$.

Przykład 199. Uzasadnić, że funkcja $f\colon R^2 \times R^2 \to R$ nie jest iloczynem skalarnym w przestrzeni R^2, jeśli dla każdych wektorów $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ z przestrzeni R^2 jest

$$f(\mathbf{x}, \mathbf{y}) = 3x_1y_1 + 2x_1y_2 + 2x_2y_1 - 2x_2y_2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Pokażemy, że funkcja f nie ma własności (S_3) iloczynu skalarnego. Dla wektora $\mathbf{x} = (x_1, x_2) \in R^2$ jest $f(\mathbf{x}, \mathbf{x}) = 3x_1^2 + 4x_1x_2 - 2x_2^2$ i w szczególności dla wektora $\mathbf{x}_0 = (0, 1)$ jest $f(\mathbf{x}_0, \mathbf{x}_0) = -2 < 0$. Stąd wynika, że funkcja f nie jest iloczynem skalarnym w przestrzeni R^2.

Definicja 9.1.3. *Normą* albo *długością* wektora $\mathbf{x}$ w przestrzeni Euklidesa nazywamy liczbę rzeczywistą

$$\|\mathbf{x}\| = \sqrt{(\mathbf{x}|\mathbf{x})}. \tag{9.6}$$

Norma wektora

Wektor $\mathbf{x}$ jest *wektorem jednostkowym*, gdy $\|\mathbf{x}\| = 1$. Łatwo zauważyć, że jeśli $\mathbf{x} \neq 0$, to $(1/\|\mathbf{x}\|)\,\mathbf{x}$ jest wektorem jednostkowym. Przykładowo, normą funkcji $f(x) = x^2$ z przestrzeni $C(\langle 0; 1\rangle)$ z iloczynem skalarnym określonym wzorem (9.5) jest $\|f\| = \sqrt{(f|f)} = \sqrt{\int_0^1 x^4\,dx} = 1/\sqrt{5}$ i $(1/\|f\|)f = \sqrt{5}x^2$.

Wektor jednostkowy

Definicja 9.1.4. *Odległością pomiędzy wektorami* $\mathbf{x}$ i $\mathbf{y}$ nazywamy liczbę

$$\mathrm{d}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|. \tag{9.7}$$

Odległość pomiędzy wektorami

Jeśli W jest niepustym podzbiorem wektorów przestrzeni Euklidesa V i $\mathbf{x}$ jest wektorem z przestrzeni V, to liczbę

$$\mathrm{d}(\mathbf{x}, W) = \inf_{\mathbf{y} \in W} \mathrm{d}(\mathbf{x}, \mathbf{y}) \tag{9.8}$$

nazywamy *odległością pomiędzy wektorem* $\mathbf{x}$ *i zbiorem wektorów* W.

Podstawowe własności normy i odległości pomiędzy wektorami przedstawiamy w następnych dwóch twierdzeniach i w kolejnym wniosku.

Nierówność Schwarza

Twierdzenie 9.1.1. *Dla dowolnych wektorów* $\mathbf{x}$ *i* $\mathbf{y}$ *w przestrzeni Euklidesa spełniona jest nierówność*

$$|(\mathbf{x}|\mathbf{y})| \leqslant \|\mathbf{x}\|\,\|\mathbf{y}\|. \tag{9.9}$$

Dowód. Łatwo zauważyć, że nierówność (9.9) jest prawdziwa, gdy $\mathbf{x} = \mathbf{0}$ lub $\mathbf{y} = \mathbf{0}$. Z własności (S_2) i (S_3) iloczynu skalarnego wynika, że jeśli $\mathbf{x} \neq \mathbf{0}$ i $\mathbf{y} \neq \mathbf{0}$, to dla każdej liczby $t \in R$ mamy

$$0 \leqslant (t\mathbf{x} + \mathbf{y}|t\mathbf{x} + \mathbf{y}) = t^2(\mathbf{x}|\mathbf{x}) + 2t(\mathbf{x}|\mathbf{y}) + (\mathbf{y}|\mathbf{y}) = \|\mathbf{x}\|^2 t^2 + 2(\mathbf{x}|\mathbf{y})t + \|\mathbf{y}\|^2,$$

więc wyróżnik trójmianu kwadratowego $\|\mathbf{x}\|^2 t^2 + 2(\mathbf{x}|\mathbf{y})t + \|\mathbf{y}\|^2$ zmiennej t jest niedodatni. Zatem $\Delta = 4(\mathbf{x}|\mathbf{y})^2 - 4\|\mathbf{x}\|^2\|\mathbf{y}\|^2 \leqslant 0$ i stąd już otrzymujemy $|(\mathbf{x}|\mathbf{y})| \leqslant \|\mathbf{x}\|\,\|\mathbf{y}\|$. □

Przykład 200. Wobec nierówności (9.9) dla każdych dwóch wektorów $\mathbf{x} = (x_1, \ldots, x_n)$ i $\mathbf{y} = (y_1, \ldots, y_n)$ z przestrzeni R^n (ze standardowym iloczynem skalarnym) mamy

$$|x_1y_1 + \ldots + x_ny_n| \leqslant (x_1^2 + \ldots + x_n^2)^{1/2}(y_1^2 + \ldots + y_n^2)^{1/2}.$$

Przykład 201. Za pomocą nierówności Schwarza (9.9) w odpowiednio dobranej przestrzeni Euklidesa uzyskać następujące nierówności:

(a) $(x + y + z)^2 \leqslant 3(x^2 + y^2 + z^2)$ dla każdych $x,\ y,\ z \in R$;

(b) $(xy + yz + xz)^2 \leqslant (x^2 + y^2 + z^2)^2$ dla każdych $x,\ y,\ z \in R$;

(c) $\left(\int_0^1 f(x)\,dx\right)^2 \leqslant \int_0^1 f^2(x)\,dx$ dla każdej funkcji $f \in C(\langle 0; 1\rangle)$.

(a) W przestrzeni R^3 ze standardowym iloczynem skalarnym weźmy pod uwagę wektory $\mathbf{x} = (x, y, z)$ i $\mathbf{y} = (1, 1, 1)$. Z nierówności Schwarza mamy

$$(x + y + z)^2 = \big((x, y, z)|(1, 1, 1)\big)^2 \leqslant \|(x, y, z)\|^2\|(1, 1, 1)\|^2 = 3\,(x^2 + y^2 + z^2).$$

(b) Z nierówności Schwarza dla wektorów $\mathbf{x} = (x, y, z)$ i $\mathbf{y} = (y, z, x)$ z przestrzeni R^3 ze standardowym iloczynem skalarnym mamy

$$(xy + yz + xz)^2 = \big((x, y, z)|(y, z, x)\big)^2 \leqslant \|(x, y, z)\|^2\|(y, z, x)\|^2 = (x^2 + y^2 + z^2)^2.$$

(c) W przestrzeni $C\big(\langle 0; 1\rangle\big)$ z iloczynem skalarnym określonym wzorem (9.5) wobec nierówności Schwarza mamy

$$\begin{aligned}\left(\int_0^1 f(x)\,dx\right)^2 &= \left(\int_0^1 f(x)\cdot 1\,dx\right)^2 = (f|1)^2 \leqslant \|f\|^2\|1\|^2 \\ &= \left(\int_0^1 f^2(x)\,dx\right)\left(\int_0^1 1\,dx\right) = \int_0^1 f^2(x)\,dx.\end{aligned}$$

Własności normy

Twierdzenie 9.1.2. *Norma wektora w przestrzeni Euklidesa V ma następujące własności:*

(N_1) $\forall_{\mathbf{x}\in V}\ \|\mathbf{x}\| \geqslant 0$ *i* $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$*;*

(N_2) $\forall_{\mathbf{x}\in V}\forall_{\alpha\in R}\ \|\alpha\mathbf{x}\| = |\alpha|\,\|\mathbf{x}\|$*;*

Nierówność trójkąta (Cauchy'ego-Minkowskiego)

(N_3) $\forall_{\mathbf{x},\mathbf{y}\in V}\ \|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$.

Dowód. Dwie pierwsze własności łatwo wynikają z definicji normy i własności (S_3) oraz (S_2) iloczynu skalarnego. Nierówność (N_3), zwana także *nierównością Cauchy'ego-Minkowskiego*, wynika z nierówności (9.9). Mamy bowiem

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|^2 = (\mathbf{x}+\mathbf{y}|\mathbf{x}+\mathbf{y}) &= (\mathbf{x}|\mathbf{x}) + 2(\mathbf{x}|\mathbf{y}) + (\mathbf{y}|\mathbf{y}) \\ &= \|\mathbf{x}\|^2 + 2(\mathbf{x}|\mathbf{y}) + \|\mathbf{y}\|^2 \\ &\leqslant \|\mathbf{x}\|^2 + 2|(\mathbf{x}|\mathbf{y})| + \|\mathbf{y}\|^2 \\ &\leqslant \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\,\|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2,\end{aligned}$$

a stąd już otrzymujemy nierówność $\|\mathbf{x}+\mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$. □

Z twierdzenia 9.1.2 natychmiast wynikają następujące własności odległości pomiędzy wektorami.

Wniosek 9.1.1. *Dla każdych wektorów* $\mathbf{x}$, $\mathbf{y}$ *i* $\mathbf{z}$ *z przestrzeni Euklidesa jest:*

(1) $d(\mathbf{x},\mathbf{y}) \geqslant 0$ *oraz* $d(\mathbf{x},\mathbf{y}) = 0$ *wtedy i tylko wtedy, gdy* $\mathbf{x} = \mathbf{y}$*;*

(2) $d(\mathbf{x},\mathbf{y}) = d(\mathbf{y},\mathbf{x})$*;*

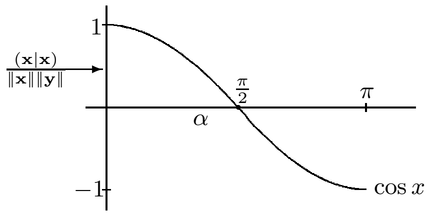
(3) $d(\mathbf{x},\mathbf{y}) \leqslant d(\mathbf{x},\mathbf{z}) + d(\mathbf{z},\mathbf{y})$. □

9.2. Kąt pomiędzy wektorami

Z nierówności Schwarza (zob. (9.9)) wynika, że dla niezerowych wektorów $\mathbf{x}$ i $\mathbf{y}$ jest

$$-1 \leqslant \frac{(\mathbf{x}|\mathbf{y})}{\|\mathbf{x}\|\,\|\mathbf{y}\|} \leqslant 1,$$

więc ułamek $\frac{(\mathbf{x}|\mathbf{y})}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$ jest wartością cosinusa (zob. rys. 9.1) i możemy przyjąć następującą definicję miary kąta pomiędzy wektorami $\mathbf{x}$ i $\mathbf{y}$.



Rys. 9.1

Definicja 9.2.1. *Miarą kąta* pomiędzy niezerowymi wektorami $\mathbf{x}$ i $\mathbf{y}$ w przestrzeni Euklidesa nazywamy liczbę $\varphi \in \langle 0;\pi\rangle$ taką, że

Kąt pomiędzy wektorami

$$\cos\varphi = \frac{(\mathbf{x}|\mathbf{y})}{\|\mathbf{x}\|\,\|\mathbf{y}\|}. \tag{9.10}$$

Definicja 9.2.2. Dwa wektory $\mathbf{x}$ i $\mathbf{y}$ w przestrzeni Euklidesa nazywamy *ortogonalnymi* (*prostopadłymi*), gdy $(\mathbf{x}|\mathbf{y}) = 0$.

Ortogonalność wektorów

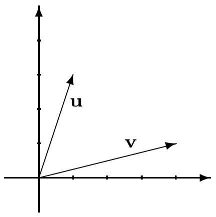
Z definicji tej wynika, że wektory $\mathbf{x}$ i $\mathbf{y}$ są ortogonalne (piszemy $\mathbf{x} \perp \mathbf{y}$), gdy co najmniej jeden z nich jest wektorem zerowym lub gdy miara kąta między tymi wektorami jest równa $\pi/2$.

Przykład 202. W przestrzeni $C(\langle 0;\pi/2\rangle)$ z iloczynem skalarnym określonym wzorem (9.5) dla $a = 0$ i $b = \pi/2$ funkcje $f(x) = \cos x + \sin x$ i $g(x) = \cos x - \sin x$ jest

$$(f|g) = \int_0^{\frac{\pi}{2}} f(x)g(x)\,dx = \int_0^{\frac{\pi}{2}} (\cos^2 x - \sin^2 x)\,dx = 0,$$

więc funkcje te są wektorami ortogonalnymi.

Następny przykład pokazuje, że miara kąta pomiędzy wektorami (i ortogonalność wektorów) zależy od wyboru iloczynu skalarnego w przestrzeni wektorowej.



Rys. 9.2

Przykład 203. Wektory $\mathbf{v} = (4, 1)$ i $\mathbf{u} = (1, 3)$ z rysunku 9.2 nie są ortogonalne wprzestrzeni R^2 ze standardowym iloczynem skalarnym, bo $(\mathbf{v}|\mathbf{u}) = \mathbf{v}^T\mathbf{u} = 7 \neq 0$. Jednakże te same dwa wektory są ortogonalne w tej samej przestrzeni R^2 z loczynem skalarnym określonym wzorem (9.3) dla bazy $B = (\mathbf{v}, \mathbf{u})$, bo mamy

$$(\mathbf{v}|\mathbf{u}) = [\mathbf{v}]_B^T[\mathbf{u}]_B = [\ 1\ 0\ ]\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0.$$

W każdej przestrzeni Euklidesa mamy następującą własność wektorów ortogonalnych.

Twierdzenie Pitagorasa

Twierdzenie 9.2.1 (Pitagorasa). *W przestrzeni Euklidesa wektory* $\mathbf{u}$ *i* $\mathbf{v}$ *są ortogonalne wtedy i tylko wtedy, gdy*

$$\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = \|\mathbf{u} + \mathbf{v}\|^2. \tag{9.11}$$

Dowód. Dla wektorów $\mathbf{u}$ i $\mathbf{v}$ mamy

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

Stąd wynika, że $\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = \|\mathbf{u} + \mathbf{v}\|^2$ wtedy i tylko wtedy, gdy $(\mathbf{u}|\mathbf{v}) = 0$, tj. wtedy i tylko wtedy, gdy wektory $\mathbf{u}$ i $\mathbf{v}$ są ortogonalne. $\square$

Baza ortogonalna

Baza ortonormalna

Definicja 9.2.3. Układ $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ wektorów z przestrzeni Euklidesa nazywamy *ortogonalnym* (*ortonormalnym*), jeśli jego wektory są wzajemnie ortogonalne (i normalne), tj. $(\mathbf{v}_i|\mathbf{v}_j) = 0$ dla $i \neq j$ (i każdy jest długości 1, czyli $\|\mathbf{v}_i\| = 1$ dla $i = 1, \ldots, n$). *Bazą ortogonalną* przestrzeni Euklidesa V nazywamy układ wektorów $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$, który jednocześnie jest bazą przestrzeni V i ortogonalnym układem wektorów w tej przestrzeni. Bazę $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ przestrzeni Euklidesa V nazywamy *bazą ortonormalną*, gdy jest ona ortonormalnym układem wektorów w przestrzeni V.

Przykład 204. Standardowa baza $E = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ przestrzeni R^n (ze standardowym iloczynem skalarnym) jest bazą ortonormalną w R^n. Układy

$$\left(\frac{1}{\sqrt{5}}(2, 1), \frac{1}{\sqrt{5}}(1, -2)\right) \quad \text{i} \quad \left(\frac{1}{\sqrt{14}}(1, 2, 3), \frac{1}{\sqrt{3}}(1, 1, -1), \frac{1}{\sqrt{42}}(5, -4, 1)\right)$$

są innymi przykładami baz ortonormalnych w przestrzeniach R^2 i R^3.

Współczynniki Fouriera

Kombinacja Fouriera

Definicja 9.2.4. Jeśli S jest ortonormalnym układem wektorów w przestrzeni V z iloczynem skalarnym $(\cdot|\cdot)$ i jeśli $\mathbf{x} \in V$, to *współczynnikami Fouriera* wektora $\mathbf{x}$ względem układu S nazywamy liczby $(\mathbf{x}|\mathbf{v}_i)$ dla $\mathbf{v}_i \in S$. Natomiast sumę

$$\sum_{\mathbf{v}_i \in S} (\mathbf{x}|\mathbf{v}_i)\mathbf{v}_i \tag{9.12}$$

nazywamy *kombinacją Fouriera* wektora $\mathbf{x}$ względem układu ortonormalnego S. (Kombinacja ta nie musi być identyczna z wektorem $\mathbf{x}$.)

Przykład 205. W przestrzeni R^3 ze standardowym iloczynem skalarnym wyznaczyć kombinację Fouriera wektora $\mathbf{x} = (-11, 3, 11)$ względem ortonormalnego układu $(\mathbf{v}_1, \mathbf{v}_2)$, gdzie $\mathbf{v}_1 = \frac{1}{\sqrt{35}}(1, -3, 5)$ i $\mathbf{v}_2 = \frac{1}{\sqrt{6}}(1, 2, 1)$. Czy ta kombinacja pokrywa się z wektorem $\mathbf{x}$?

Ponieważ $(\mathbf{x}|\mathbf{v}_1) = \sqrt{35}$ i $(\mathbf{x}|\mathbf{v}_2) = \sqrt{6}$, więc kombinacją Fouriera wektora $\mathbf{x}$ względem układu $(\mathbf{v}_1, \mathbf{v}_2)$ jest

$$(\mathbf{x}|\mathbf{v}_1)\mathbf{v}_1 + (\mathbf{x}|\mathbf{v}_2)\mathbf{v}_2 = \sqrt{35}\frac{1}{\sqrt{35}}(1,-3,5) + \sqrt{6}\frac{1}{\sqrt{6}}(1,2,1) = (2,-1,6)$$

i jest ona różna od wektora $\mathbf{x} = (-11, 3, 11)$.

Pierwsze zalety ortonormalnych układów wektorów i współczynników Fouriera przedstawia następujące twierdzenie i wynikające zeń wnioski.

Twierdzenie 9.2.2. *Niech $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ będzie ortogonalnym układem niezerowych wektorów przestrzeni Euklidesa V i niech $\mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{v}_i$ będzie kombinacją liniową wektorów tego układu. Wtedy $\alpha_i = \frac{(\mathbf{x}|\mathbf{v}_i)}{\|\mathbf{v}_i\|^2}$ dla $i = 1, \ldots, n$, więc*

$$\mathbf{x} = \sum_{i=1}^n \frac{(\mathbf{x}|\mathbf{v}_i)}{\|\mathbf{v}_i\|^2}\mathbf{v}_i \tag{9.13}$$

i jest to kombinacja Fouriera wektora $\mathbf{x}$ względem ortonormalnego układu $\left(\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \ldots, \frac{\mathbf{v}_n}{\|\mathbf{v}_n\|}\right)$.

Dowód. Jeśli $\mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{v}_i$, to dla $i = 1, \ldots, n$ jest

$$\begin{aligned}(\mathbf{x}|\mathbf{v}_i) &= \left(\textstyle\sum_{j=1}^n \alpha_j \mathbf{v}_j|\mathbf{v}_i\right) = \textstyle\sum_{j=1}^n \alpha_j(\mathbf{v}_j|\mathbf{v}_i) \\ &= \alpha_i(\mathbf{v}_i|\mathbf{v}_i) = \alpha_i\|\mathbf{v}_i\|^2\end{aligned}$$

i stąd wynika (9.13), bo $\mathbf{x} = \sum_{i=1}^n \frac{(\mathbf{x}|\mathbf{v}_i)}{\|\mathbf{v}_i\|^2}\mathbf{v}_i = \sum_{i=1}^n \left(\mathbf{x}|\frac{\mathbf{v}_i}{\|\mathbf{v}_i\|}\right)\frac{\mathbf{v}_i}{\|\mathbf{v}_i\|}$. □

Wniosek 9.2.1. *Jeśli $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ jest ortonormalną bazą przestrzeni V, to dla każdego $\mathbf{x} \in V$ jest $\mathbf{x} = \sum_{i=1}^n (\mathbf{x}|\mathbf{v}_i)\mathbf{v}_i$ i dlatego*

$$[\mathbf{x}]_B = \begin{bmatrix} (\mathbf{x}|\mathbf{v}_1) \\ (\mathbf{x}|\mathbf{v}_2) \\ \vdots \\ (\mathbf{x}|\mathbf{v}_n) \end{bmatrix}. \quad \square \tag{9.14}$$

Wniosek 9.2.2. *Układ $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ niezerowych i wzajemnie ortogonalnych wektorów w przestrzeni Euklidesa jest liniowo niezależny.*

Niezerowe wektory ortogonalne są liniowo niezależne

Dowód. Załóżmy, że dla wektorów układu $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ i skalarów $\alpha_1, \ldots, \alpha_n$ jest $\sum_{i=1}^n \alpha_i \mathbf{v}_i = \mathbf{0}$. Wtedy wobec twierdzenia 9.2.2 jest $\alpha_i = (\mathbf{0}|\mathbf{v}_i)/\|\mathbf{v}_i\|^2 = 0$ dla $i = 1, \ldots, n$. To dowodzi, że układ $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ jest liniowo niezależny. □

Przykład 206. Wykazać, że układ

$$B = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \left(\frac{1}{\sqrt{5}}(1,0,-2), \frac{1}{\sqrt{14}}(2,3,1), \frac{1}{\sqrt{70}}(6,-5,3)\right)$$

jest bazą ortonormalną przestrzeni R^3 ze standardowym iloczynem skalarnym. Dodatkowo, wektor $\mathbf{x} = \sqrt{70}(-1,2,-1)$ przedstawić jako kombinację liniową wektorów bazy B.

Ponieważ

$$\begin{aligned}(\mathbf{v}_1|\mathbf{v}_1) = 1,\ \ (\mathbf{v}_1|\mathbf{v}_2) = 0,\ \ (\mathbf{v}_1|\mathbf{v}_3) = 0,\\ (\mathbf{v}_2|\mathbf{v}_2) = 1,\ \ (\mathbf{v}_2|\mathbf{v}_3) = 0,\\ (\mathbf{v}_3|\mathbf{v}_3) = 1,\end{aligned}$$

więc $B = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ jest ortonormalnym układem wektorów w przestrzeni R^3. Stąd i z wniosku 9.2.2 wynika, że układ B jest liniowo niezależny i (wobec twierdzenia 7.5.4) jest on bazą przestrzeni R^3. Dla wektora $\mathbf{x}$ jest

$$(\mathbf{x}|\mathbf{v}_1) = \sqrt{14}, \quad (\mathbf{x}|\mathbf{v}_2) = 3\sqrt{5} \quad \text{i} \quad (\mathbf{x}|\mathbf{v}_3) = -19,$$

więc wobec wniosku 9.2.1 (lub twierdzenia 9.2.2) mamy

$$\mathbf{x} = (\mathbf{x}|\mathbf{v}_1)\mathbf{v}_1 + (\mathbf{x}|\mathbf{v}_2)\mathbf{v}_2 + (\mathbf{x}|\mathbf{v}_3)\mathbf{v}_3 = \sqrt{14}\mathbf{v}_1 + 3\sqrt{5}\mathbf{v}_2 - 19\mathbf{v}_3.$$

9.3. Ortogonalizacja bazy

Pokażemy teraz, że każda skończenie wymiarowa przestrzeń Euklidesa ma bazę ortogonalną i ortonormalną. Zaprezentujemy tu tak zwaną *metodę Grama-Schmidta ortogonalizacji bazy*, czyli metodę przekształcania bazy $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ przestrzeni Euklidesa w pewną bazę ortogonalną $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ i otrzymaną z niej bazę ortonormalną $(\mathbf{y}_1/\|\mathbf{y}_1\|, \mathbf{y}_2/\|\mathbf{y}_2\|, \ldots, \mathbf{y}_n/\|\mathbf{y}_n\|)$ tej samej przestrzeni. W metodzie tej przyjmuje się, że $\mathbf{y}_1 = \mathbf{x}_1$, a każdy następny wektor $\mathbf{y}_k$ $(k = 2, \ldots, n)$ jest kombinacją liniową wektorów $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{k-1}$ i $\mathbf{x}_k$,

$$\mathbf{y}_k = \mathbf{x}_k + a_{k\,1}\mathbf{y}_1 + \ldots + a_{k\,k-1}\mathbf{y}_{k-1},$$

gdzie współczynniki a_{kj} dobiera się tak, aby wektor $\mathbf{y}_k$ był ortogonalny do każdego z wektorów $\mathbf{y}_1, \ldots, \mathbf{y}_{k-1}$. Nasze rozważania zaczynamy od przydatnego lematu.

Lemat 9.3.1. *Niech $(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1})$ będzie układem niezerowych i wzajemnie ortogonalnych wektorów przestrzeni Euklidesa V, niech $\mathbf{b}$ będzie dowolnym wektorem ze zbioru V i niech $\overline{\mathbf{b}}$ będzie jego kombinacją Fouriera względem ortonormalnego układu $\left(\frac{\mathbf{y}_1}{\|\mathbf{y}_1\|}, \ldots, \frac{\mathbf{y}_{k-1}}{\|\mathbf{y}_{k-1}\|}\right)$, czyli $\overline{\mathbf{b}} = \sum_{i=1}^{k-1} \frac{(\mathbf{b}|\mathbf{y}_i)}{\|\mathbf{y}_i\|^2}\,\mathbf{y}_i$. Wtedy wektor $\mathbf{b} - \overline{\mathbf{b}}$ jest ortogonalny do podprzestrzeni $\mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1})$. Dodatkowo, jeśli $\mathbf{b} \in V - \mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1})$, to wektor $\mathbf{b} - \overline{\mathbf{b}}$ jest niezerowy.*

Dowód. Wektor $\mathbf{b} - \overline{\mathbf{b}}$ jest ortogonalny do podprzestrzeni $\mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1})$, bo dla każdego $j = 1, \ldots, k-1$ jest

$$\begin{aligned}(\mathbf{b} - \overline{\mathbf{b}}|\mathbf{y}_j) &= \left(\mathbf{b} - \sum_{i=1}^{k-1} \frac{(\mathbf{b}|\mathbf{y}_i)}{\|\mathbf{y}_i\|^2}\mathbf{y}_i \,\middle|\, \mathbf{y}_j\right)\\ &= (\mathbf{b}|\mathbf{y}_j) - \sum_{i=1}^{k-1} \frac{(\mathbf{b}|\mathbf{y}_i)}{\|\mathbf{y}_i\|^2}(\mathbf{y}_i|\mathbf{y}_j)\\ &= (\mathbf{b}|\mathbf{y}_j) - \frac{(\mathbf{b}|\mathbf{y}_j)}{\|\mathbf{y}_j\|^2}(\mathbf{y}_j|\mathbf{y}_j) = 0.\end{aligned}$$

Dalej, jeśli $\mathbf{b} \in V - \mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1})$, to $\mathbf{b} - \overline{\mathbf{b}} \neq \mathbf{0}$, bo inaczej byłoby $\mathbf{b} = \sum_{i=1}^{k-1} \frac{(\mathbf{b}|\mathbf{y}_i)}{\|\mathbf{y}_i\|^2}\,\mathbf{y}_i$ i wektor $\mathbf{b}$ byłby elementem podprzestrzeni $\mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1})$. □

Twierdzenie 9.3.1. *Jeśli $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ jest bazą przestrzeni Euklidesa V, to układ wektorów $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$, w którym*

Metoda Grama-Schmidta ortogonalizacji bazy

$$\left\{\begin{array}{lcl}\mathbf{y}_1 &=& \mathbf{x}_1\\ \mathbf{y}_2 &=& \mathbf{x}_2 - \frac{(\mathbf{x}_2|\mathbf{y}_1)}{(\mathbf{y}_1|\mathbf{y}_1)}\,\mathbf{y}_1\\ &\vdots& \\ \mathbf{y}_n &=& \mathbf{x}_n - \frac{(\mathbf{x}_n|\mathbf{y}_1)}{(\mathbf{y}_1|\mathbf{y}_1)}\,\mathbf{y}_1 - \ldots - \frac{(\mathbf{x}_n|\mathbf{y}_{n-1})}{(\mathbf{y}_{n-1}|\mathbf{y}_{n-1})}\,\mathbf{y}_{n-1},\end{array}\right. \tag{9.15}$$

jest bazą ortogonalną przestrzeni V, a układ

$$\left(\frac{\mathbf{y}_1}{\|\mathbf{y}_1\|}, \frac{\mathbf{y}_2}{\|\mathbf{y}_2\|}, \ldots, \frac{\mathbf{y}_n}{\|\mathbf{y}_n\|}\right)$$

jest bazą ortonormalną przestrzeni V.

Dowód. Ponieważ drugie stwierdzenie jest oczywistą konsekwencją pierwszego stwierdzenia, udowodnimy tylko to pierwsze. W tym celu z uwagi na wniosek 9.2.2 wystarczy indukcyjnie ze względu na k ($1 \leqslant k \leqslant n$) uzasadnić, że $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k)$ jest ortogonalnym układem niezerowych wektorów i $\mathcal{L}(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k) \subseteq \mathcal{L}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$. Jest to oczywiste dla $k = 1$. Załóżmy teraz, że $(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1})$ jest ortogonalnym układem niezerowych wektorów i $\mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1}) \subseteq \mathcal{L}(\mathbf{x}_1, \ldots, \mathbf{x}_{k-1})$, $2 \leqslant k \leqslant n$. Ponieważ $\mathbf{x}_k \in V - \mathcal{L}(\mathbf{x}_1, \ldots, \mathbf{x}_{k-1}) \subseteq V - \mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1})$, więc z lematu 9.3.1 (dla $\mathbf{b} = \mathbf{x}_k$) wynika, że wektor $\mathbf{y}_k$ jest niezerowy i ortogonalny do podprzestrzeni $\mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1})$. Stąd zaś i z założenia wynika, że także układ $(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1}, \mathbf{y}_k)$ jest ortogonalnym układem niezerowych wektorów.

W końcu, ponieważ $\mathbf{y}_k = \mathbf{x}_k - \sum_{i=1}^{k-1} \frac{(\mathbf{x}_k|\mathbf{y}_i)}{\|\mathbf{y}_i\|^2}\mathbf{y}_i \in \mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1}, \mathbf{x}_k)$, więc

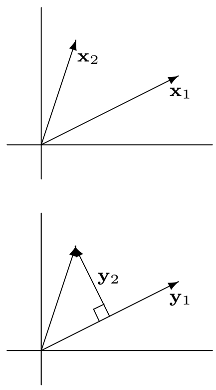
$$\mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1}, \mathbf{y}_k) \subseteq \mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1}, \mathbf{x}_k) \subseteq \mathcal{L}(\mathbf{x}_1, \ldots, \mathbf{x}_{k-1}, \mathbf{x}_k),$$

bo z założenia jest $\mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1}) \subseteq \mathcal{L}(\mathbf{x}_1, \ldots, \mathbf{x}_{k-1})$. $\square$

Przykład 207. Baza $(\mathbf{x}_1, \mathbf{x}_2) = \left(\begin{bmatrix} 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}\right)$ przestrzeni R^2 (ze standardowym iloczynem skalarnym) nie jest ortogonalna, ale wobec twierdzenia 9.3.1 (zob. (9.15)) układ wektorów $(\mathbf{y}_1, \mathbf{y}_2)$ jest już bazą ortogonalną przestrzeni R^2, gdy

$$\mathbf{y}_1 = \mathbf{x}_1 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, \quad \mathbf{y}_2 = \mathbf{x}_2 - \frac{(\mathbf{x}_2|\mathbf{y}_1)}{(\mathbf{y}_1|\mathbf{y}_1)}\mathbf{y}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix},$$

rys. 9.3.



Rys. 9.3

Przykład 208. Metodą Grama-Schmidta utworzyć bazę ortogonalną i ortonormalną podprzestrzeni $W = \mathcal{L}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ przestrzeni R^4 (ze standardowym iloczynem skalarnym), gdzie $\mathbf{x}_1 = (1, 1, 0, 1)$, $\mathbf{x}_2 = (3, 0, 0, 3)$ i $\mathbf{x}_3 = (1, -1, -1, 0)$.

Ponieważ układ wektorów $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ jest bazą przestrzeni W (dlaczego?), więc wobec twierdzenia 9.3.1 bazą ortogonalną przestrzeni W będzie układ $(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3)$, w którym

$$\mathbf{y}_1 = \mathbf{x}_1, \ \ \mathbf{y}_2 = \mathbf{x}_2 - \frac{(\mathbf{x}_2|\mathbf{y}_1)}{(\mathbf{y}_1|\mathbf{y}_1)}\mathbf{y}_1, \ \ \mathbf{y}_3 = \mathbf{x}_3 - \frac{(\mathbf{x}_3|\mathbf{y}_1)}{(\mathbf{y}_1|\mathbf{y}_1)}\mathbf{y}_1 - \frac{(\mathbf{x}_3|\mathbf{y}_2)}{(\mathbf{y}_2|\mathbf{y}_2)}\mathbf{y}_2.$$

Łatwo widać, że mamy $(\mathbf{y}_1|\mathbf{y}_1) = 3$ i $(\mathbf{x}_2|\mathbf{y}_1) = 6$. Dlatego

$$\mathbf{y}_2 = (3, 0, 0, 3) - \frac{6}{3}(1, 1, 0, 1) = (1, -2, 0, 1).$$

Analogicznie, ponieważ $(\mathbf{y}_2|\mathbf{y}_2) = 6$, $(\mathbf{x}_3|\mathbf{y}_1) = 0$ i $(\mathbf{x}_3|\mathbf{y}_2) = 3$, więc

$$\mathbf{y}_3 = (1, -1, -1, 0) - \frac{0}{3}(1, 1, 0, 1) - \frac{3}{6}(1, -2, 0, 1) = \frac{1}{2}(1, 0, -2, -1).$$

Zatem układ

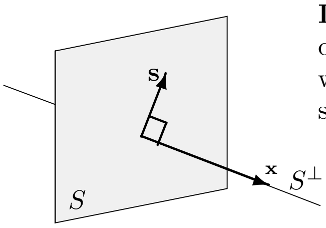
$$(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3) = \left((1, 1, 0, 1), (1, -2, 0, 1), \frac{1}{2}(1, 0, -2, -1)\right)$$

jest bazą ortogonalną przestrzeni W, a układ

$$\left(\frac{\mathbf{y}_1}{\|\mathbf{y}_1\|}, \frac{\mathbf{y}_2}{\|\mathbf{y}_2\|}, \frac{\mathbf{y}_3}{\|\mathbf{y}_3\|}\right) = \left(\frac{1}{\sqrt{3}}(1, 1, 0, 1), \frac{1}{\sqrt{6}}(1, -2, 0, 1), \frac{1}{2\sqrt{6}}(1, 0, -2, -1)\right)$$

jest bazą ortonormalną tej samej przestrzeni.

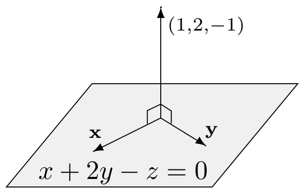
9.4. Dopełnienie ortogonalne



Rys. 9.4

Definicja 9.4.1. Jeśli S jest niepustym zbiorem wektorów w przestrzeni Euklidesa V, to zbiór wszystkich wektorów przestrzeni V ortogonalnych do każdego wektora ze zbioru S nazywamy *ortogonalnym dopełnieniem* zbioru S i oznaczamy symbolem $S^{\perp}$ (zob. rys. 9.4),

$$S^{\perp} = \{\mathbf{x} \in V : (\mathbf{x}|\mathbf{s}) = 0 \quad \text{dla każdego} \quad \mathbf{s} \in S\}. \tag{9.16}$$



Rys. 9.5

Przykład 209. W przestrzeni R^3 ze standardowym iloczynem skalarnym ortogonalnym dopełnieniem zbioru $S = \{(1, 2, -1)\}$ jest

$$\begin{aligned} S^{\perp} = \{(1, 2, -1)\}^{\perp} &= \{(x, y, z) \in R^3 : ((x, y, z)|(1, 2, -1)) = 0\} \\ &= \{(x, y, z) \in R^3 : x + 2y - z = 0\}. \end{aligned}$$

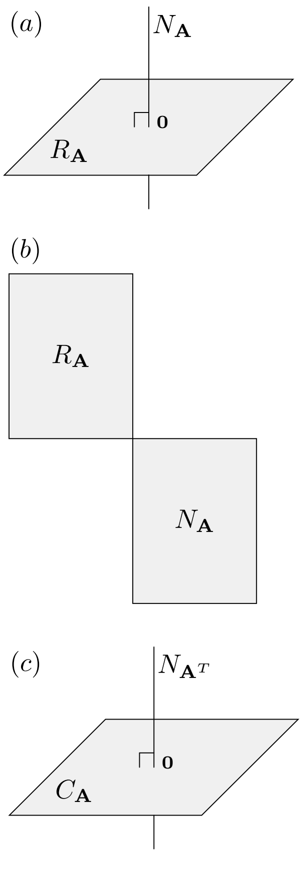
Zatem

$S^{\perp}$ jest zbiorem wektorów odpowiadających punktom płaszczyzny o równaniu $x + 2y - z = 0$, rys. 9.5.

$$\frac{S \subseteq V \quad \text{i} \quad S \neq \emptyset}{S^{\perp} - \text{podprzestrzeń}}$$

$S \cap S^{\perp} \subseteq \{\mathbf{0}\}$

Łatwo zauważyć, że $S^{\perp}$ jest podprzestrzenią przestrzeni V dla każdego niepustego podzbioru S zbioru V. Ponieważ dla wektora $\mathbf{x}$ jest $(\mathbf{x}|\mathbf{x}) = 0$ wtedy i tylko wtedy, gdy $\mathbf{x} = \mathbf{0}$, więc wektor zerowy jest jedynym wektorem, który może jednocześnie należeć do zbioru S i do jego ortogonalnego dopełnienia $S^{\perp}$. Warto także zauważyć, że wektor $\mathbf{x}$ jest ortogonalny do podprzestrzeni $\mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ wtedy i tylko wtedy, gdy $\mathbf{x}$ jest ortogonalny do każdego z wektorów $\mathbf{v}_1, \ldots, \mathbf{v}_m$. Dodatkowo, jeśli $\mathbf{v}_1, \ldots, \mathbf{v}_m$ są wektorami z przestrzeni R^n, to wektor $\mathbf{x}$ jest ortogonalny do podprzestrzeni $\mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ wtedy i tylko wtedy, gdy $\mathbf{Ax} = \mathbf{0}$, gdzie $\mathbf{A}$ jest macierzą, której wierszami są $\mathbf{v}_1, \ldots, \mathbf{v}_m$. Z tych obserwacji skorzystamy przy wyjaśnianiu ortogonalnych zależności pomiędzy podprzestrzeniami $N_{\mathbf{A}}$, $R_{\mathbf{A}}$, $C_{\mathbf{A}}$ i $N_{\mathbf{A}^T}$, czterema podstawowymi podprzestrzeniami odpowiadającymi macierzy $\mathbf{A}$, i przy wyznaczaniu ortogonalnego dopełnienia przestrzeni $\mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$, gdy $\mathbf{v}_1, \ldots, \mathbf{v}_m \in R^n$.



Rys. 9.6

Twierdzenie 9.4.1. *Ortogonalnym dopełnieniem przestrzeni wierszowej macierzy* $\mathbf{A}$ *jest przestrzeń zerowa macierzy* $\mathbf{A}$ *i ortogonalnym dopełnieniem przestrzeni kolumnowej macierzy* $\mathbf{A}$ *jest przestrzeń zerowa macierzy* $\mathbf{A}^T$*,*

$$(R_{\mathbf{A}})^{\perp} = N_{\mathbf{A}} \quad i \quad (C_{\mathbf{A}})^{\perp} = N_{\mathbf{A}^T}. \tag{9.17}$$

Dowód. Niech $\mathbf{v}_1, \ldots, \mathbf{v}_m$ będą wierszami macierzy $\mathbf{A}$ wymiaru $m \times n$. Wtedy

$$\begin{aligned} N_{\mathbf{A}} &= \{\mathbf{x} \in R^n : \mathbf{Ax} = \mathbf{0}\} = \left\{ \mathbf{x} \in R^n : \begin{bmatrix} - & \mathbf{v}_1 & - \\ - & \mathbf{v}_2 & - \\ & \vdots & \\ - & \mathbf{v}_m & - \end{bmatrix} \mathbf{x} = \mathbf{0} \right\} \\ &= \left\{ \mathbf{x} \in R^n : \begin{bmatrix} (\mathbf{v}_1|\mathbf{x}) \\ (\mathbf{v}_2|\mathbf{x}) \\ \vdots \\ (\mathbf{v}_m|\mathbf{x}) \end{bmatrix} = \mathbf{0} \right\} = \{\mathbf{x} \in R^n : \mathbf{x} \perp \mathbf{v}_i \text{ dla } i = 1, \ldots, m\} \\ &= \{\mathbf{x} \in R^n : \mathbf{x} \perp \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_m) = R_{\mathbf{A}}\} = (R_{\mathbf{A}})^{\perp} \end{aligned}$$

i to kończy dowód pierwszej części tezy (zob. rys. 9.6(a) i 9.6(b)). Druga część tezy wynika z pierwszej i z faktu, że $C_{\mathbf{A}} = R_{\mathbf{A}^T}$ (rys. 9.6(c)). $\square$

Przykład 210. W przestrzeni R^4 wyznaczyć bazę ortogonalnego dopełnienia podprzestrzeni

$$S = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \mathcal{L}\big((1,-2,3,-1),(2,-3,4,0),(7,-12,17,-3)\big).$$

Niech $\mathbf{A}$ będzie macierzą, której wierszami są wektory $\mathbf{v}_1$, $\mathbf{v}_2$ i $\mathbf{v}_3$,

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 3 & -1 \\ 2 & -3 & 4 & 0 \\ 7 & -12 & 17 & -3 \end{bmatrix}.$$

Wtedy S jest przestrzenią wierszową macierzy $\mathbf{A}$, $S = R_{\mathbf{A}}$, i wobec twierdzenia 9.4.1 ortogonalnym dopełnieniem przestrzeni S jest przestrzeń zerowa macierzy $\mathbf{A}$. Łatwo zauważyć, że przestrzenią zerową macierzy $\mathbf{A}$ jest

$$N_{\mathbf{A}} = \{s(1,2,1,0) + t(3,2,0,-1):\ s,t \in R\}.$$

Ponieważ wektory $(1,2,1,0)$ i $(3,2,0,-1)$ generujące przestrzeń $N_{\mathbf{A}}$ są liniowo niezależne, więc układ $\big((1,2,1,0),(3,2,0,-1)\big)$ jest bazą przestrzeni $S^{\perp}$.

Następne twierdzenie pokazuje, że przestrzen Euklidesa jest sumą prostą (zob. def. 7.8.2) swojej podprzestrzeni i jej ortogonalnego dopełnienia.

Twierdzenie 9.4.2. *Jeśli W jest podprzestrzenią przestrzeni Euklidesa V, to*

$$V = W \oplus W^{\perp}, \tag{9.18}$$

czyli dla każdego wektora $\mathbf{b} \in V$ istnieje jednoznacznie wyznaczony wektor $\overline{\mathbf{b}} \in W$ taki, że $\mathbf{b} - \overline{\mathbf{b}} \in W^{\perp}$.

Dowód. Niech $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)$ będzie ortonormalną bazą podprzestrzeni W i niech $\mathbf{b}$ będzie wektorem z przestrzeni V. Wtedy wektor $\overline{\mathbf{b}} = \sum_{i=1}^{k}(\mathbf{b}|\mathbf{a}_i)\mathbf{a}_i$ należy do podprzestrzeni W i wobec lematu 9.3.1 wektor $\mathbf{b} - \overline{\mathbf{b}}$ należy do podprzestrzeni $W^{\perp}$. Stąd wynika, że przestrzeń V jest sumą podprzestrzeni W i $W^{\perp}$. Dodatkowo, ponieważ $W \cap W^{\perp} = \{\mathbf{0}\}$, więc (wobec definicji 7.8.2) przestrzeń V jest sumą prostą swoich podprzestrzeni W i $W^{\perp}$, czyli $V = W \oplus W^{\perp}$. □

Kolejne twierdzenie jest inną wersją twierdzenia 9.4.2.

Twierdzenie 9.4.3. *Niech W będzie podprzestrzenią przestrzeni Euklidesa V i niech $\mathbf{b} \in V$. Wtedy istnieją jednoznacznie wyznaczone wektory $\overline{\mathbf{b}} \in W$ i $\widetilde{\mathbf{b}} \in W^{\perp}$ takie, że $\mathbf{b} = \overline{\mathbf{b}} + \widetilde{\mathbf{b}}$. Dodatkwo, jeśli $(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ jest ortonormalną bazą przestrzeni W, to $\overline{\mathbf{b}}$ jest kombinacją Fouriera wektora $\mathbf{b}$ względem ortonormalnego układu $(\mathbf{a}_1, \ldots, \mathbf{a}_k)$,*

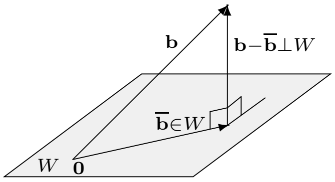
$$\overline{\mathbf{b}} = \sum_{i=1}^{k}(\mathbf{b}|\mathbf{a}_i)\mathbf{a}_i. \tag{9.19}$$

Dowód. Niech $(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ będzie bazą ortonormalną przestrzeni W. Wtedy $\overline{\mathbf{b}} = \sum_{i=1}^{k}(\mathbf{b}|\mathbf{a}_i)\mathbf{a}_i \in W$. Dodatkowo, wobec lematu 9.3.1, wektor $\widetilde{\mathbf{b}} = \mathbf{b} - \overline{\mathbf{b}}$ należy do podprzestrzeni $W^{\perp}$.

Dla dowodu jednoznaczności przypuśćmy, że dodatkowo $\mathbf{b} = \overline{\mathbf{b}}' + \widetilde{\mathbf{b}}'$, gdzie $\overline{\mathbf{b}}' \in W$ i $\widetilde{\mathbf{b}}' \in W^{\perp}$. Wtedy $\overline{\mathbf{b}} + \widetilde{\mathbf{b}} = \mathbf{b} = \overline{\mathbf{b}}' + \widetilde{\mathbf{b}}'$ i dlatego $\overline{\mathbf{b}}' - \overline{\mathbf{b}}' = \widetilde{\mathbf{b}}' - \widetilde{\mathbf{b}} \in W \cap W^{\perp} = \{\mathbf{0}\}$. Stąd zaś wynika, że $\overline{\mathbf{b}}' = \overline{\mathbf{b}}$ i $\widetilde{\mathbf{b}}' = \widetilde{\mathbf{b}}$. □

9.5. Rzut ortogonalny

Definicja 9.5.1. Niech $\mathbf{b}$ będzie wektorem z przestrzeni V i niech W będzie podprzestrzenią przestrzeni V. *Rzutem ortogonalnym* wektora $\mathbf{b}$ na podprzestrzeń W nazywamy wektor $\overline{\mathbf{b}}$ (czasami oznaczany symbolem $\mathrm{proj}_W \mathbf{b}$) taki, że $\overline{\mathbf{b}} \in W$ i $\mathbf{b} - \overline{\mathbf{b}}$ jest ortogonalny do podprzestrzeni W, zob. rys. 9.7. Rzut ortogonalny



Rys. 9.7. Jeśli $\overline{\mathbf{b}} = \operatorname{proj}_W \mathbf{b}$, to $\overline{\mathbf{b}} \in W$ i $\mathbf{b} - \overline{\mathbf{b}} \perp W$

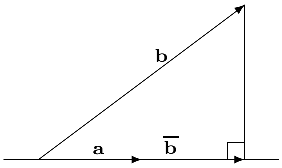
Twierdzenia 9.4.2 i 9.4.3 gwarantują istnienie i jednoznaczność rzutu ortogonalnego każdego wektora $\mathbf{b}$ z przestrzeni V na podprzestrzeń W przestrzeni V. Z twierdzeń tych wynika też, że rzutem ortogonalnym wektora $\mathbf{b}$ na podprzestrzeń W jest kombinacja Fouriera wektora $\mathbf{b}$ względem bazy ortonormalnej podprzestrzeni W.

Wniosek 9.5.1. *Jeśli* $(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ *jest bazą ortogonalną podprzestrzeni* W *przestrzeni Euklidesa* V *i jeśli* $\mathbf{b}$ *jest wektorem z przestrzeni* V, *to*

$$\operatorname{proj}_W \mathbf{b} = \sum_{i=1}^{k} \frac{(\mathbf{b}|\mathbf{a}_i)}{(\mathbf{a}_i|\mathbf{a}_i)} \mathbf{a}_i. \tag{9.20}$$

W szczególności, jeśli $(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ *jest bazą ortonormalną podprzestrzeni* W, *to*

$$\operatorname{proj}_W \mathbf{b} = \sum_{i=1}^{k} (\mathbf{b}|\mathbf{a}_i)\mathbf{a}_i. \quad \square \tag{9.21}$$



$W = \mathcal{L}(\mathbf{a})$ i $\overline{\mathbf{b}} = \frac{(\mathbf{b}|\mathbf{a})}{(\mathbf{a}|\mathbf{a})}\,\mathbf{a}$

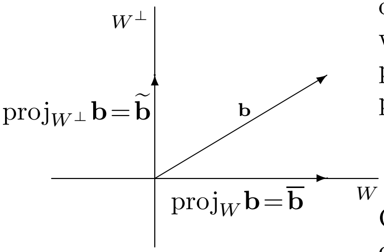
Rys. 9.8

Przykład 211. Z wniosku 9.5.1 jest oczywiste, że rzut ortogonalny wektora $\mathbf{b}$ na 1-wymiarową podprzestrzeń $W = \mathcal{L}(\mathbf{a})$ (rys. 9.8) jest określony wzorem

$$\overline{\mathbf{b}} = \frac{(\mathbf{b}|\mathbf{a})}{(\mathbf{a}|\mathbf{a})}\,\mathbf{a}. \tag{9.22}$$

Przykładowo, w przestrzeni R^3 (ze standardowym iloczynem skalarnym) rzutem ortogonalnym wektora $\mathbf{b} = (2, 4, -3)$ na podprzestrzeń W generowaną przez wektor $\mathbf{a} = (0, 1, -2)$ jest wektor

$$\overline{\mathbf{b}} = \frac{(\mathbf{b}|\mathbf{a})}{(\mathbf{a}|\mathbf{a})}\,\mathbf{a} = \frac{\begin{bmatrix} 2 & 4 & -3 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}}{\begin{bmatrix} 0 & 1 & -2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}} \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix} = \frac{10}{5}\begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ -4 \end{bmatrix}.$$



Rys. 9.9

Z definicji rzutu ortogonalnego łatwo wynika, że wektor $\overline{\mathbf{b}}$ jest rzutem ortogonalnym wektora $\mathbf{b}$ na podprzestrzeń W przestrzeni V wtedy i tylko wtedy, gdy wektor $\widetilde{\mathbf{b}} = \mathbf{b} - \overline{\mathbf{b}}$ jest rzutem ortogonalnym wektora $\mathbf{b}$ na podprzestrzeń $W^{\perp}$, rys. 9.9. Równoważnie, dla każdej podprzestrzeni W przestrzeni V i dla każdego wektora $\mathbf{b} \in V$ jest

$$\mathbf{b} = \operatorname{proj}_W \mathbf{b} + \operatorname{proj}_{W^{\perp}} \mathbf{b}. \tag{9.23}$$

O tej prostej zależności warto pamiętać wtedy, gdy bezpośrednie wyznaczanie rzutu ortogonalnego wektora $\mathbf{b}$ na podprzestrzeń $W^{\perp}$ jest łatwiejsze od wyznaczania rzutu ortogonalnego wektora $\mathbf{b}$ na podprzestrzeń W. Ilustruje to następujący przykład.

Przykład 212. W przestrzeni R^4 ze standardowym iloczynem skalarnym wyznaczyć rzut ortogonalny wektora $\mathbf{b} = (1, 0, -1, 1)$ na podprzestrzeń W generowaną przez wektory $\mathbf{a}_1 = (1, 0, 0, 0)$, $\mathbf{a}_2 = (0, 1, 1, 0)$ i $\mathbf{a}_3 = (0, 0, 1, 1)$.

Ponieważ $W = \mathcal{L}(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ jest 3-wymiarową podprzestrzenią w przestrzeni R^4, więc jej ortogonalne dopełnienie $W^{\perp}$ jest podprzestrzenią 1-wymiarową generowaną przez każdy niezerowy wektor $\mathbf{a}$ ortogonalny do wektorów $\mathbf{a}_1$, $\mathbf{a}_2$ i $\mathbf{a}_3$. W szczególności, wektorem tym może być wektor $\mathbf{a} = (0, -1, 1, -1)$. (Samą przestrzeń $W^{\perp}$ i generujący ją wektor można wyznaczyć – podobnie jak to zrobiliśmy w przykładzie 210 – jako przestrzeń zerową macierzy $\mathbf{A}$, której wierszami są $\mathbf{a}_1$, $\mathbf{a}_2$ i $\mathbf{a}_3$.) Wobec wniosku 9.5.1 rzutem wektora $\mathbf{b}$ na podprzestrzeń $W^{\perp} = \mathcal{L}(\mathbf{a})$ jest wektor

$$\widetilde{\mathbf{b}} = \frac{(\mathbf{b}|\mathbf{a})}{(\mathbf{a}|\mathbf{a})}\,\mathbf{a} = \frac{-2}{3}(0,-1,1,-1).$$

Zatem, wobec własności (9.23), rzutem wektora $\mathbf{b}$ na podprzestrzeń W jest wektor

$$\overline{\mathbf{b}} = \mathbf{b} - \widetilde{\mathbf{b}} = (1,0,-1,1) - \frac{-2}{3}(0,-1,1,-1) = \left(1, -\frac{2}{3}, -\frac{1}{3}, \frac{1}{3}\right).$$

Przedstawimy teraz sposób wyznaczania rzutu ortogonalnego wektora na podprzestrzeń, gdy znamy jakąkolwiek bazę tej podprzestrzeni.

Twierdzenie 9.5.1. *Jeśli układ $(\mathbf{a}_1, \dots, \mathbf{a}_k)$ jest bazą podprzestrzeni W przestrzeni Euklidesa V, to rzutem ortogonalnym wektora $\mathbf{b} \in V$ na podprzestrzeń W jest wektor*

$$\overline{\mathbf{b}} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_k\mathbf{a}_k,$$

gdzie $(x_1, x_2, \dots, x_k)$ jest rozwiązaniem układu równań liniowych

$$\begin{cases} x_1(\mathbf{a}_1|\mathbf{a}_1) + x_2(\mathbf{a}_1|\mathbf{a}_2) + \ldots + x_k(\mathbf{a}_1|\mathbf{a}_k) = (\mathbf{a}_1|\mathbf{b}), \\ x_1(\mathbf{a}_2|\mathbf{a}_1) + x_2(\mathbf{a}_2|\mathbf{a}_2) + \ldots + x_k(\mathbf{a}_2|\mathbf{a}_k) = (\mathbf{a}_2|\mathbf{b}), \\ \qquad\vdots \\ x_1(\mathbf{a}_k|\mathbf{a}_1) + x_2(\mathbf{a}_k|\mathbf{a}_2) + \ldots + x_k(\mathbf{a}_k|\mathbf{a}_k) = (\mathbf{a}_k|\mathbf{b}). \end{cases} \tag{9.24}$$

Dowód. Z faktu, że układ wektorów $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k)$ jest bazą przestrzeni $W = \mathcal{L}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k)$ wynika, że wektor $\overline{\mathbf{b}} = \sum_{j=1}^{k} x_j\mathbf{a}_j$ jest rzutem ortogonalnym wektora $\mathbf{b}$ na podprzestrzeń W wtedy i tylko wtedy, gdy wektor $\mathbf{b} - \overline{\mathbf{b}}$ jest ortogonalny do podprzestrzeni $\mathcal{L}(\mathbf{a}_1, \dots, \mathbf{a}_k)$. Tak jest wtedy i tylko wtedy, gdy

$$(\mathbf{a}_i|\mathbf{b} - \overline{\mathbf{b}}) = (\mathbf{a}_i|\mathbf{b}) - \sum_{j=1}^{k} x_j(\mathbf{a}_i|\mathbf{a}_j) = 0$$

dla $i = 1, \dots, k$, czyli wtedy i tylko wtedy, gdy $(x_1, \dots, x_k)$ jest rozwiązaniem układu (9.24). □

Definicja 9.5.2. Wyznacznik macierzy głównej układu równań liniowych (9.24) nazywa się *wyznacznikiem Grama układu* $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_2, \dots, \mathbf{a}_k)$ i oznacza symbolem $G(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k)$, Wyznacznik Grama

$$G(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k) = \begin{vmatrix} (\mathbf{a}_1|\mathbf{a}_1) & (\mathbf{a}_1|\mathbf{a}_2) & \dots & (\mathbf{a}_1|\mathbf{a}_k) \\ (\mathbf{a}_2|\mathbf{a}_1) & (\mathbf{a}_2|\mathbf{a}_2) & \dots & (\mathbf{a}_2|\mathbf{a}_k) \\ \vdots & \vdots & \ddots & \vdots \\ (\mathbf{a}_k|\mathbf{a}_1) & (\mathbf{a}_k|\mathbf{a}_2) & \dots & (\mathbf{a}_k|\mathbf{a}_k) \end{vmatrix}. \tag{9.25}$$

Łatwo zauważyć, że wyznacznik Grama $G(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k)$ jest niezerowy, jeśli układ wektorów $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k)$ jest liniowo niezależny.

Przykład 213. Znaleźć rzut ortogonalny wektora $\mathbf{b} = (6, 6, 21)$ na podprzestrzeń W generowaną przez wektory $\mathbf{a}_1 = (1, 0, 1)$ i $\mathbf{a}_2 = (0, 1, 1)$.

Wobec twierdzenia 9.5.1 rzutem ortogonalnym wektora $\mathbf{b}$ na podprzestrzeń generowaną przez wektory $\mathbf{a}_1$ i $\mathbf{a}_2$ jest wektor

$$\overline{\mathbf{b}} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2,$$

gdzie (x_1, x_2) jest rozwiązaniem układu równań liniowych

$$\begin{cases} x_1(\mathbf{a}_1|\mathbf{a}_1) + x_2(\mathbf{a}_1|\mathbf{a}_2) = (\mathbf{a}_1|\mathbf{b}), \\ x_1(\mathbf{a}_2|\mathbf{a}_1) + x_2(\mathbf{a}_2|\mathbf{a}_2) = (\mathbf{a}_2|\mathbf{b}). \end{cases}$$

Tu mamy

$$\begin{cases} 2x_1 + \ x_2 = 27, \\ \ x_1 + 2x_2 = 27, \end{cases}$$

więc $x_1 = x_2 = 9$ i dlatego $\overline{\mathbf{b}} = 9\mathbf{a}_1 + 9\mathbf{a}_2 = (9, 9, 18)$.

9.6. Macierz rzutu ortogonalnego

Przedstawimy tu jeszcze jeden sposób wyznaczania rzutu ortogonalnego wektora na podprzestrzeń W przestrzeni R^n. Niech $(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ będzie bazą podprzestrzeni W. Wtedy

$$W = \mathcal{L}(\mathbf{a}_1, \ldots, \mathbf{a}_k) = \{\mathbf{Ax}\colon\ \mathbf{x} \in R^k\} = C_{\mathbf{A}},$$

gdzie $\mathbf{A}$ jest macierzą wymiaru $n \times k$, której kolejnymi kolumnami są $\mathbf{a}_1, \ldots, \mathbf{a}_k$.

Weźmy pod uwagę dowolny wektor $\mathbf{b} \in R^n$ i jego rzut ortogonalny $\overline{\mathbf{b}}$ na podprzestrzeń W. Wtedy $\overline{\mathbf{b}} \in W = C_{\mathbf{A}}$, więc $\overline{\mathbf{b}} = \mathbf{Ax}_0$ (dla pewnego $\mathbf{x}_0 \in R^k$) i wektor $\mathbf{b} - \overline{\mathbf{b}} = \mathbf{b} - \mathbf{Ax}_0$ jest ortogonalny do podprzestrzeni W, czyli do każdego wektora $\mathbf{Ax}$, gdzie $\mathbf{x} \in R^k$. Dlatego

$$\mathbf{x}^T(\mathbf{A}^T\mathbf{b} - \mathbf{A}^T\mathbf{Ax}_0) = (\mathbf{Ax})^T(\mathbf{b} - \mathbf{Ax}_0) = 0$$

dla każdego $\mathbf{x}$. Stąd wynika, że

$$\mathbf{A}^T\mathbf{b} - \mathbf{A}^T\mathbf{Ax}_0 = \mathbf{0}.$$

Ponieważ $\mathbf{A}$ jest macierzą rzędu k (bo jej kolumny są liniowo niezależne), więc wobec twierdzenia 7.5.4 także macierz $\mathbf{A}^T\mathbf{A}$ wymiaru $k \times k$ jest rzędu k. Dlatego $\mathbf{A}^T\mathbf{A}$ jest odwracalna i z równości $\mathbf{A}^T\mathbf{b} - \mathbf{A}^T\mathbf{Ax}_0 = \mathbf{0}$ otrzymujemy $\mathbf{x}_0 = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$. Stąd wynika, że rzutem wektora $\mathbf{b}$ na podprzestrzeń W jest

$$\overline{\mathbf{b}} = \mathbf{Ax}_0 = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}.$$

Zatem udowodniliśmy następujące twierdzenie.

Twierdzenie 9.6.1. *Niech $W = \mathcal{L}(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ będzie k-wymiarową podprzestrzenią przestrzeni R^n i niech $\mathbf{A}$ będzie macierzą, której kolumnami są wektory $\mathbf{a}_1, \ldots, \mathbf{a}_k$. Wtedy rzutem ortogonalnym wektora $\mathbf{b} \in R^n$ na podprzestrzeń W jest*

$$\overline{\mathbf{b}} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}. \quad \square \tag{9.26}$$

Definicja 9.6.1. Jeśli $W = \mathcal{L}(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ jest k-wymiarową podprzestrzenią przestrzeni R^n i $\mathbf{A}$ jest macierzą, której kolumnami są $\mathbf{a}_1, \ldots, \mathbf{a}_k$, to macierz

Macierz rzutu

$$\mathbf{P} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T \tag{9.27}$$

nazywamy *macierzą rzutu* (ortogonalnego) na podprzestrzeń W.

Jeśli $\mathbf{A} = [\mathbf{a}_1 \ \ldots \ \mathbf{a}_k]$ i $\mathbf{B} = [\mathbf{b}_1 \ \ldots \ \mathbf{b}_k]$ są macierzami wymiaru $n \times k$ i układy $(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ i $(\mathbf{b}_1, \ldots, \mathbf{b}_k)$ są bazami tej samej podprzestrzeni W przestrzeni R^n, to $\mathbf{P} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ i $\mathbf{P}' = \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T$ są macierzami rzutu na podprzestrzeń W. Ponieważ dla każdego $\mathbf{x} \in R^n$ jest $\mathbf{Px} = \mathbf{P}'\mathbf{x}$, więc $\mathbf{P} = \mathbf{P}'$ i to oznacza, że macierz rzutu na podprzestrzeń W nie zależy od wyboru bazy tej podprzestrzeni.

Przykład 214. Wyznaczyć macierz $\mathbf{P}$ rzutu ortogonalnego na podprzestrzeń W będącą płaszczyzną o równaniu $x+y-z=0$. Za pomocą tej macierzy znaleźć rzut ortogonalny wektora $\mathbf{b}=(6,6,21)$ na podprzestrzeń W.

Bazę przestrzeni W tworzą dwa nierównoległe wektory $\mathbf{a}_1$ i $\mathbf{a}_2$ leżące w płaszczyźnie $x+y-z=0$, rys. 9.10. Mogą nimi być

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \text{i} \quad \mathbf{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Wtedy dla macierzy

$$\mathbf{A} = \begin{bmatrix} | & | \\ \mathbf{a}_1 & \mathbf{a}_2 \\ | & | \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

mamy

$$(\mathbf{A}^T\mathbf{A})^{-1} = \left(\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \right)^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{-1} = \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

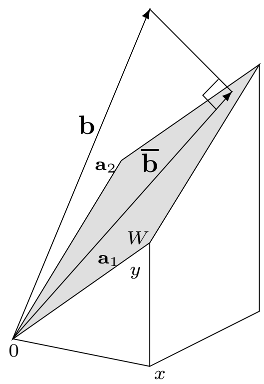
i wobec (9.27) macierzą rzutu na płaszczyznę W jest

$$\begin{aligned} \mathbf{P} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T \quad &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \tfrac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \\ &= \tfrac{1}{3} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \end{bmatrix} = \tfrac{1}{3} \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}. \end{aligned}$$

Zatem, wobec (9.26), rzutem ortogonalnym wektora $\mathbf{b}$ na płaszczyznę W jest wektor

$$\overline{\mathbf{b}} = \mathrm{proj}_W \mathbf{b} = \mathbf{Pb} = \frac{1}{3} \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 6 \\ 21 \end{bmatrix} = \begin{bmatrix} 9 \\ 9 \\ 18 \end{bmatrix}$$

(i oczywiście jest on identyczny z wektorem z przykładu 213).



Rys. 9.10

Jeśli baza $(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ podprzestrzeni W jest ortonormalna, to macierz $\mathbf{A} = [\mathbf{a}_1 \ \ldots \ \mathbf{a}_k]$ ma ortonormalne kolumny ($\mathbf{a}_i^T\mathbf{a}_i = 1$ i $\mathbf{a}_i^T\mathbf{a}_j = 0$, gdy $j \neq i$), więc $\mathbf{A}^T\mathbf{A} = \mathbf{I}$ (i $(\mathbf{A}^T\mathbf{A})^{-1} = \mathbf{I}$) i mamy prostszą niż w (9.27) zależność pomiędzy macierzą $\mathbf{A}$ i macierzą rzutu $\mathbf{P}$ na podprzestrzeń W,

$$\mathbf{P} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \mathbf{AIA}^T = \mathbf{AA}^T. \tag{9.28}$$

Przykładowo dla bazy ortonormalnej

$$(\mathbf{a}_1, \mathbf{a}_2) = \Big((1/\sqrt{2}, 0, 1/\sqrt{2}), \ (-1/\sqrt{6}, \sqrt{2/3}, 1/\sqrt{6}) \Big)$$

przestrzeni W z ostatniego przykładu i dla macierzy

$$\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ 0 & \sqrt{\frac{2}{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

iloczyn

$$\mathbf{AA}^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ 0 & \sqrt{\frac{2}{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{6}} & \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{6}} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

jest znaną z przykładu 214 macierzą rzutu na podprzestrzeń W.

Na koniec zaobserwujmy pewne ogólne własności macierzy rzutu. Przede wszystkim zauważmy, że jeśli $W = \mathcal{L}(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ jest k-wymiarową podprzestrzenią przestrzeni R^n i $\mathbf{A}$ jest macierzą, której kolumnami są $\mathbf{a}_1, \ldots, \mathbf{a}_k$, to macierz rzutu $\mathbf{P} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ jest symetryczna,

$\mathbf{P}^T = \mathbf{P}$

$$\begin{aligned} \mathbf{P}^T &= (\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T)^T = (\mathbf{A}^T)^T\left((\mathbf{A}^T\mathbf{A})^{-1}\right)^T\mathbf{A}^T \\ &= \mathbf{A}\left((\mathbf{A}^T\mathbf{A})^T\right)^{-1}\mathbf{A}^T = \mathbf{A}\left(\mathbf{A}^T(\mathbf{A}^T)^T\right)^{-1}\mathbf{A}^T = \mathbf{P}, \end{aligned}$$

i idempotentna,

$\mathbf{P}^2 = \mathbf{P}$

$$\begin{aligned} \mathbf{P}^2 &= (\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T)(\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T) \\ &= \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \mathbf{P}. \end{aligned}$$

W następnym twierdzeniu udowodnimy, że te dwie własności w pełni charakteryzują macierz rzutu ortogonalnego.

Twierdzenie 9.6.2. *Macierz $\mathbf{P} \in R_{n\times n}$ jest macierzą rzutu ortogonalnego na pewną podprzestrzeń $W \subseteq R^n$ wtedy i tylko wtedy, gdy jest ona symetryczna ($\mathbf{P}^T = \mathbf{P}$) i idempotentna ($\mathbf{P}^2 = \mathbf{P}$).*

Dowód. Symetryczność i idempotentność macierzy rzutu już wyjaśniliśmy. Załóżmy teraz, że macierz $\mathbf{P}$ jest symetryczna i idempotentna. Wykażemy, że $\mathbf{P}$ jest macierzą rzutu ortogonalnego na podprzestrzeń $C_{\mathbf{P}} = \{\mathbf{P}\mathbf{x}\colon\ \mathbf{x} \in R^n\}$, podprzestrzeń kolumnową macierzy $\mathbf{P}$. W tym celu wystarczy pokazać, że jeśli $\mathbf{b} \in R^n$, to $\mathbf{P}\mathbf{b} \in C_{\mathbf{P}}$ (co jest oczywiste) i wektor $\mathbf{b} - \mathbf{P}\mathbf{b}$ jest ortogonalny do każdego wektora $\mathbf{P}\mathbf{x}$ z przestrzeni $C_{\mathbf{P}}$. Ponieważ $\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^T$, więc dla iloczynu skalarnego wektorów $\mathbf{b} - \mathbf{P}\mathbf{b}$ i $\mathbf{P}\mathbf{x}$ mamy

$$\begin{aligned} (\mathbf{b} - \mathbf{P}\mathbf{b})^T\mathbf{P}\mathbf{x} &= ((\mathbf{I} - \mathbf{P})\mathbf{b})^T\,\mathbf{P}\mathbf{x} = \mathbf{b}^T(\mathbf{I} - \mathbf{P})^T\mathbf{P}\mathbf{x} = \mathbf{b}^T(\mathbf{I}^T - \mathbf{P}^T)\mathbf{P}\mathbf{x} \\ &= \mathbf{b}^T(\mathbf{I} - \mathbf{P})\mathbf{P}\mathbf{x} = \mathbf{b}^T(\mathbf{P} - \mathbf{P}^2)\mathbf{x} = \mathbf{b}^T\mathbf{0}\mathbf{x} = \mathbf{0} \end{aligned}$$

i to dowodzi, że wektory $\mathbf{b} - \mathbf{P}\mathbf{b}$ i $\mathbf{P}\mathbf{x}$ są ortogonalne. □

Przykład 215. Czy macierz $\mathbf{P} = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$ jest macierzą rzutu ortogonalnego na pewną podprzestrzeń przestrzeni R^2?

Ponieważ

$$\mathbf{P}^2 = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}^2 = \begin{bmatrix} 20 & 10 \\ 10 & 5 \end{bmatrix} \neq \mathbf{P},$$

macierz $\mathbf{P}$ nie jest idempotentna i nie może być macierzą rzutu ortogonalnego na żadną podprzestrzeń przestrzeni R^2.

9.7. Metoda najmniejszych kwadratów

W tej części zajmujemy się aproksymacją wektora $\mathbf{b}$ z przestrzeni Euklidesa V wektorem z podprzestrzeni $W \subset V$. Poszukujem wektora $\mathbf{x}$ należącego do podprzestrzeni W takiego, że odległość pomiędzy wektorami $\mathbf{b}$ i $\mathbf{x}$, czyli norma $\|\mathbf{b} - \mathbf{x}\|$, jest najmniejsza z możliwych. W przypadku takiej aproksymacji wektor $\mathbf{b} - \mathbf{x}$ nazywa się *wektorem błędu* (jaki popełnia się zastępując wektor $\mathbf{b}$ wektorem $\mathbf{x}$), a jego długość $\|\mathbf{b} - \mathbf{x}\|$ – *wielkością błędu* aproksymacji. Następne twierdzenie zapewnia, że wektor $\mathrm{proj}_W\mathbf{b}$, czyli rzut ortogonalny wektora $\mathbf{b}$ na podprzestrzeń W, jest najlepszą (ze względu na odległość) aproksymacją wektora $\mathbf{b}$ wektorami należącymi do podprzestrzeni W.

Twierdzenie 9.7.1. *Jeśli $\overline{\mathbf{b}}$ jest rzutem ortogonalnym wektora $\mathbf{b}$ na podprzestrzeń W przestrzeni Euklidesa V, to dla każdego wektora $\widetilde{\mathbf{b}} \in W$ różnego od $\overline{\mathbf{b}}$ zachodzi nierówność*

Twierdzenie o najlepszej aproksymacji

$$\|\mathbf{b} - \overline{\mathbf{b}}\| < \|\mathbf{b} - \widetilde{\mathbf{b}}\|, \tag{9.29}$$

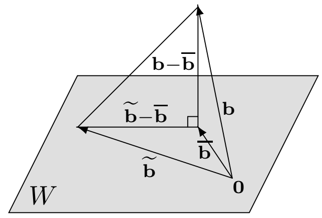
czyli odległość pomiędzy wektorem $\mathbf{b}$ i podprzestrzenią W jest odległością pomiędzy wektorem $\mathbf{b}$ i jego rzutem ortogonalnym $\overline{\mathbf{b}}$ na podprzestrzeń W,

$$\mathrm{d}(\mathbf{b}, W) = \min_{\widetilde{\mathbf{b}} \in W} \|\mathbf{b} - \widetilde{\mathbf{b}}\| = \|\mathbf{b} - \overline{\mathbf{b}}\|.$$

Dowód. Niech $\widetilde{\mathbf{b}}$ będzie dowolnym wektorem ze zbioru $W - \{\overline{\mathbf{b}}\}$. Wtedy $\overline{\mathbf{b}} - \widetilde{\mathbf{b}} \in W - \{\mathbf{0}\}$ i $\|\overline{\mathbf{b}} - \widetilde{\mathbf{b}}\| > 0$. Z definicji rzutu ortogonalnego wektor $\mathbf{b} - \overline{\mathbf{b}}$ jest ortogonalny do podprzestrzeni W, więc w szczególności jest on ortogonalny do wektora $\overline{\mathbf{b}} - \widetilde{\mathbf{b}}$ (zob. rys. 9.11). Stąd i z twierdzenia Pitagorasa (tw. 9.2.1) mamy

$$\|\mathbf{b} - \widetilde{\mathbf{b}}\|^2 = \|(\mathbf{b} - \overline{\mathbf{b}}) + (\overline{\mathbf{b}} - \widetilde{\mathbf{b}})\|^2 = \|\mathbf{b} - \overline{\mathbf{b}}\|^2 + \|\overline{\mathbf{b}} - \widetilde{\mathbf{b}}\|^2 > \|\mathbf{b} - \overline{\mathbf{b}}\|^2$$

i to implikuje nierówność (9.29). □



Rys. 9.11. $\mathrm{proj}_W \mathbf{b}$ jest najlepszą aproksymacją wektora $\mathbf{b}$.

9.8. Najlepsze rozwiązanie układu równań

Dany jest układ równań liniowych

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \quad \vdots \qquad\qquad \ddots \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases}$$

czyli układ $\mathbf{Ax} = \mathbf{b}$, gdzie

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \quad \text{i} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

W wielu konkretnych przypadkach taki układ $\mathbf{Ax} = \mathbf{b}$ nie ma rozwiązania. W takiej sytuacji możemy (a często musimy) wyznaczyć tzw. najlepsze rozwiązanie, czyli taki wektor $\overline{\mathbf{x}}$, że odległość pomiędzy wektorami $\mathbf{A}\overline{\mathbf{x}}$ i $\mathbf{b}$ jest najmniejsza z możliwych.

Definicja 9.8.1. *Najlepszym rozwiązaniem* układu $\mathbf{Ax} = \mathbf{b}$ nazywamy wektor $\overline{\mathbf{x}} \in R^n$ taki, że dla każdego $\mathbf{x} \in R^n$ jest

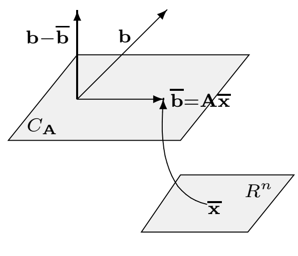
$\overline{\mathbf{x}}$ – najlepsze rozwiązanie układu $\mathbf{Ax} = \mathbf{b}$

$$\|\mathbf{A}\overline{\mathbf{x}} - \mathbf{b}\| \leqslant \|\mathbf{Ax} - \mathbf{b}\|.$$

Ponieważ $\{\mathbf{Ax}\colon\ \mathbf{x} \in R^n\}$ jest przestrzenią kolumnową macierzy $\mathbf{A}$,

$$\{\mathbf{Ax}\colon\ \mathbf{x} \in R^n\} = C_{\mathbf{A}},$$

więc wektor $\overline{\mathbf{x}}$ jest najlepszym rozwiązaniem układu $\mathbf{Ax} = \mathbf{b}$ wtedy i tylko wtedy, gdy $\mathbf{A}\overline{\mathbf{x}}$ jest tym wektorem z podprzestrzeni $C_{\mathbf{A}}$, który leży najbliżej wektora $\mathbf{b}$. Z twierdzenia 9.7.1 wynika, że tak jest wtedy i tylko wtedy, gdy $\mathbf{A}\overline{\mathbf{x}}$ jest rzutem ortogonalnym wektora $\mathbf{b}$ na podprzestrzeń $C_{\mathbf{A}}$, tj. wtedy i tylko wtedy, gdy $\mathbf{A}\overline{\mathbf{x}} = \overline{\mathbf{b}}$, gdzie $\overline{\mathbf{b}} = \mathrm{proj}_{C_{\mathbf{A}}}\mathbf{b}$, zob. rys. 9.12. Ostatni warunek jest równoważny ortogonalności wektora $\mathbf{b} - \overline{\mathbf{b}} = \mathbf{b} - \mathbf{A}\overline{\mathbf{x}}$ i podprzestrzeni $C_{\mathbf{A}}$, czyli ortogonalności wektora $\mathbf{b} - \mathbf{A}\overline{\mathbf{x}}$ do każdej kolumny macierzy $\mathbf{A}$. Na to potrzeba i wystarcza, aby zachodziła równość $\mathbf{A}^T(\mathbf{b} - \mathbf{A}\overline{x}) = 0$. Stąd zaś wynika następujące twierdzenie o najlepszych rozwiązaniach układu równań liniowych.



Rys. 9.12

Twierdzenie 9.8.1. *Zbiór najlepszych rozwiązań układu równań liniowych* $\mathbf{Ax} = \mathbf{b}$ *jest identyczny ze zbiorem rozwiązań układu*

Normalny układ równań

$$\mathbf{A}^T\mathbf{A}\,\mathbf{x} = \mathbf{A}^T\,\mathbf{b}. \quad \Box \tag{9.30}$$

Metoda najmniejszych kwadratów

Wyżej przedstawiony sposób wyznaczania najlepszego rozwiązania układu równań liniowych $\mathbf{Ax} = \mathbf{b}$ poprzez rozwiązywanie układu równań $\mathbf{A}^T\mathbf{A}\,\mathbf{x} = \mathbf{A}^T\mathbf{b}$, zwanego *normalnym układem równań* (odpowiadających układowi $\mathbf{Ax} = \mathbf{b}$), nazywa się *metodą najmniejszych kwadratów.*

Przykład 216. Za pomocą normalnego układu równań wyznaczyć najlepsze rozwiązanie układu równań liniowych $\mathbf{Ax} = \mathbf{b}$, gdy

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

Układowi $\mathbf{Ax} = \mathbf{b}$ odpowiada normalny układ równań $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$, w którym

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & -1 & -3 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ 3 & -3 & 1 \end{bmatrix} = \begin{bmatrix} 14 & -10 & 6 \\ -10 & 11 & -3 \\ 6 & -3 & 3 \end{bmatrix}$$

i

$$\mathbf{A}^T\mathbf{b} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & -1 & -3 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 \\ -6 \\ 4 \end{bmatrix}.$$

Ponieważ dla macierzy rozszerzonej układu $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$ mamy

$$[\mathbf{A}^T\mathbf{A}|\mathbf{A}^T\mathbf{b}] = \left[\begin{array}{rrr|r} 14 & -10 & 6 & 9 \\ -10 & 11 & -3 & -6 \\ 6 & -3 & 3 & 4 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 2/3 & 13/18 \\ 0 & 1 & 1/3 & 1/9 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

więc rozwiązaniem układu $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$ i najlepszym rozwiązaniem układu $\mathbf{Ax} = \mathbf{b}$ jest każdy wektor

$$\overline{\mathbf{x}} = \frac{1}{18}\begin{bmatrix} 13 \\ 2 \\ 0 \end{bmatrix} + t\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix},$$

gdzie $t \in R$.

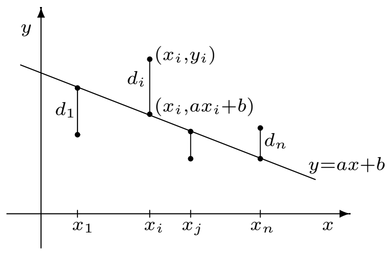
9.9. Dopasowanie prostej

Zajmiemy się teraz prostym, lecz bardzo ważnym i często spotykanym w praktycznych zastosowaniach zagadnieniem najlepszego dopasowania prostej do danego zbioru punktów płaszczyzny. Niech (x_1, y_1), $(x_2, y_2), \ldots, (x_n, y_n)$ będą punktami z płaszczyzny R^2 takimi, że nie wszystkie liczby $x_1, x_2, \ldots, x_n$ są równe. Chcemy wyznaczyć prostą $y = ax+b$, która w sensie metody najmniejszych kwadratów najlepiej pasuje do danych punktów. Jej współczynniki a i b dobieramy w taki sposób aby suma

$$\sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} (ax_i + b - y_i)^2,$$

w której $d_i = |ax_i + b - y_i|$ jest odległością pomiędzy punktami (x_i, y_i) i $(x_i, ax_i + b)$ (zob. rys. 9.13), była najmniejsza z możliwych. Ostatnia suma jest najmniejsza wtedy i tylko wtedy, gdy (a,b) jest najlepszym (w sensie metody najmniejszych kwadratów) rozwiązaniem układu równań liniowych

$$\begin{cases} ax_1 + b = y_1 \\ ax_2 + b = y_2 \\ \quad\vdots \\ ax_n + b = y_n, \end{cases}$$



Rys. 9.13

czyli układu $\mathbf{Ax} = \mathbf{b}$, w którym

$$\mathbf{A} = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{i} \quad \mathbf{b} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Wobec twierdzenia 9.8.1 najlepsze rozwiązanie (a,b) układu $\mathbf{Ax} = \mathbf{b}$ można wyznaczyć z normalnego układu równań $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$, w którym

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} x_1 & \dots & x_n \\ 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} = \begin{bmatrix} x_1^2 + \ldots + x_n^2 & x_1 + \ldots + x_n \\ x_1 + \ldots + x_n & n \end{bmatrix}$$

i

$$\mathbf{A}^T\mathbf{b} = \begin{bmatrix} x_1 & \dots & x_n \\ 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1y_1 + \ldots + x_ny_n \\ y_1 + \ldots + y_n \end{bmatrix}.$$

Bez trudu można zauważyć, że

$$\det(\mathbf{A}^T\mathbf{A}) = \sum_{1 \leqslant i < j \leqslant n} (x_i - x_j)^2$$

i ostatnia suma jest niezerowa, jeśli tylko nie wszystkie liczby $x_1, \ldots, x_n$ są równe. Stąd zaś wynika, że układ $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$ ma dokładnie jedno rozwiązanie (a układ $\mathbf{Ax} = \mathbf{b}$ ma dokładnie jedno najlepsze rozwiązanie), jeśli tylko nie wszystkie liczby $x_1, \ldots, x_n$ są równe.

Przykład 217. Wyznaczyć najlepszą liniową zależność $y = ax + b$ pomiędzy współrzędnymi x_i i y_i punktów $(0,0)$, $(2,2)$, $(4,2)$ i $(6,4)$.

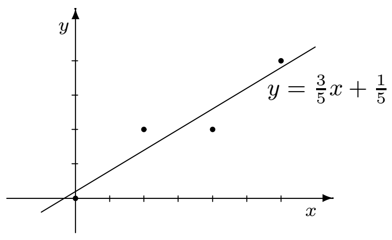
Współczynniki a i b szukanej prostej tworzą najlepsze rozwiązanie układu równań liniowych $\mathbf{Ax} = \mathbf{b}$, gdzie

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 2 & 1 \\ 4 & 1 \\ 6 & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{i} \quad \mathbf{b} = \begin{bmatrix} 0 \\ 2 \\ 2 \\ 4 \end{bmatrix}.$$

Najlepsze rozwiązanie układu $\mathbf{Ax} = \mathbf{b}$ wyznaczamy z normalnego układu równań $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$. Mamy

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 0 & 2 & 4 & 6 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 2 & 1 \\ 4 & 1 \\ 6 & 1 \end{bmatrix} = \begin{bmatrix} 56 & 12 \\ 12 & 4 \end{bmatrix},$$

$$\mathbf{A}^T\mathbf{b} = \begin{bmatrix} 0 & 2 & 4 & 6 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 36 \\ 8 \end{bmatrix}$$



Rys. 9.14

i

$$\mathbf{x}=\begin{bmatrix} a \\ b \end{bmatrix}=\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{b}=\frac{1}{80}\begin{bmatrix} 4 & -12 \\ -12 & 56 \end{bmatrix}\begin{bmatrix} 36 \\ 8 \end{bmatrix}=\begin{bmatrix} 3/5 \\ 1/5 \end{bmatrix}.$$

Zatem $y=\frac{3}{5}x+\frac{1}{5}$ jest najlepszą liniową zależnością pomiędzy współrzędnymi punktów $(0,0)$, $(2,2)$, $(4,2)$ i $(6,4)$, zob. rys. 9.14.

9.10. Macierz i przekształcenie ortogonalne

Macierz ortogonalna

Definicja 9.10.1. Mówimy, że rzeczywista macierz $\mathbf{A}$ wymiaru $n\times n$ jest *macierzą ortogonalną*, jeśli

$$\mathbf{A}^T\mathbf{A}=\mathbf{I}_n. \tag{9.31}$$

Z (9.31) jest oczywiste, że $\mathbf{A}$ jest macierzą ortogonalną wtedy i tylko wtedy, gdy $\mathbf{A}$ jest odwracalna i

$$\mathbf{A}^{-1}=\mathbf{A}^T. \tag{9.32}$$

Dodatkowo, ponieważ $\det\mathbf{A}^T=\det\mathbf{A}$ i $\det\mathbf{A}^T\mathbf{A}=\det\mathbf{A}^T\det\mathbf{A}$, więc z (9.31) otrzymujemy

$$(\det\mathbf{A})^2=\det\mathbf{A}^T\det\mathbf{A}=\det\mathbf{A}^T\mathbf{A}=\det\mathbf{I}_n=1$$

i stąd wynika, że wartość wyznacznika macierzy ortogonalnej jest równa jeden lub minus jeden. Warto także zauważyć, że jeżeli $\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n$ są kolumnami macierzy $\mathbf{A}$, to mamy

$$\mathbf{A}^T\mathbf{A}=\begin{bmatrix} - & \mathbf{a}_1^T & - \\ - & \mathbf{a}_2^T & - \\ & \vdots & \\ - & \mathbf{a}_n^T & - \end{bmatrix}\begin{bmatrix} | & | & & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \\ | & | & & | \end{bmatrix}=\begin{bmatrix} \mathbf{a}_1^T\mathbf{a}_1 & \mathbf{a}_1^T\mathbf{a}_2 & \cdots & \mathbf{a}_1^T\mathbf{a}_n \\ \mathbf{a}_2^T\mathbf{a}_1 & \mathbf{a}_2^T\mathbf{a}_2 & \cdots & \mathbf{a}_2^T\mathbf{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_n^T\mathbf{a}_1 & \mathbf{a}_n^T\mathbf{a}_2 & \cdots & \mathbf{a}_n^T\mathbf{a}_n \end{bmatrix},$$

więc równość (9.31) (i (9.32)) zachodzi wtedy i tylko wtedy, gdy

$$\mathbf{a}_i^T\mathbf{a}_j=\begin{cases} 0, & \text{jeśli} \quad j\neq i, \\ 1, & \text{jeśli} \quad j=i, \end{cases}$$

tj. wtedy i tylko wtedy, gdy kolumny macierzy $\mathbf{A}$ tworzą bazę ortonormalną przestrzeni R^n. Ponieważ równość (9.31) (i (9.32)) zachodzi wtedy i tylko wtedy, gdy $(\mathbf{A}^T)^T\mathbf{A}^T=\mathbf{A}\mathbf{A}^T=I_n$, więc także macierz $\mathbf{A}^T$ jest ortogonalna, a to ma miejsce wtedy i tylko wtedy, gdy jej kolumny (które są wierszami macierzy $\mathbf{A}$) tworzą bazę ortonormalną przestrzeni R^n. Stąd mamy następujące twierdzenie.

Twierdzenie 9.10.1. *Jeśli $\mathbf{A}$ jest rzeczywistą macierzą wymiaru $n\times n$, to następujące stwierdzenia są równoważne:*

(*a*) *wiersze macierzy $\mathbf{A}$ tworzą bazę ortonormalną przestrzeni R^n;*

(*b*) *kolumny macierzy $\mathbf{A}$ tworzą bazę ortonormalną przestrzeni R^n;*

(*c*) *macierz $\mathbf{A}$ jest ortogonalna, tj. $\mathbf{A}$ jest odwracalna i $\mathbf{A}^{-1}=\mathbf{A}^T$.* □

Przykład 218. Pokazać, że macierz

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

jest ortogonalna i wyznaczyć macierz $\mathbf{A}^{-1}$.

Ponieważ mamy

$$\mathbf{A}^T\mathbf{A} = \frac{1}{25}\begin{bmatrix} 3 & -4 & 0 \\ 0 & 0 & 5 \\ 4 & 3 & 0 \end{bmatrix}\begin{bmatrix} 3 & 0 & 4 \\ -4 & 0 & 3 \\ 0 & 5 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

więc macierz $\mathbf{A}$ jest ortogonalna i

$$\mathbf{A}^{-1} = \mathbf{A}^T = \frac{1}{5}\begin{bmatrix} 3 & -4 & 0 \\ 0 & 0 & 1 \\ 4 & 3 & 0 \end{bmatrix}.$$

Następne twierdzenie pokazuje, że przekształcenie liniowe $T_{\mathbf{A}}$ określone przez macierz ortogonalną $\mathbf{A}$, przekształcając wektory, zachowuje ich iloczyn skalarny, więc także długości wektorów i kąty pomiędzy wektorami.

Twierdzenie 9.10.2. *Niech* $\mathbf{A}$ *będzie macierzą ortogonalną stopnia n i niech* $\mathbf{x}$ *oraz* $\mathbf{y}$ *będą wektorami z przestrzeni R^n. Wtedy*

(*a*) $(\mathbf{Ax})^T(\mathbf{Ay}) = \mathbf{x}^T\mathbf{y}$, (zachowanie iloczynu skalarnego)

(*b*) $\|\mathbf{Ax}\| = \|\mathbf{x}\|$, (zachowanie długości wektorów)

(*c*) $\angle(\mathbf{Ax}, \mathbf{Ay}) = \angle(\mathbf{x}, \mathbf{y})$. (zachowanie kąta pomiędzy wektorami)

Dowód. (*a*) Ponieważ $\mathbf{A}^T\mathbf{A} = I_n$, więc mamy $(\mathbf{Ax})^T(\mathbf{Ay}) = \mathbf{x}^T\mathbf{A}^T\mathbf{Ay} = \mathbf{x}^T I_n\mathbf{y} = \mathbf{x}^T\mathbf{y}$. (*b*) Z definicji normy oraz z (*a*) mamy $\|\mathbf{Ax}\| = \sqrt{(\mathbf{Ax})^T(\mathbf{Ax})} = \sqrt{\mathbf{x}^T\mathbf{x}} = \|\mathbf{x}\|$. (*c*) Z definicji miary kąta pomiędzy wektorami (zob. def. 9.2.1) oraz z (*a*) i (*b*) mamy $\angle(\mathbf{Ax}, \mathbf{Ay}) = \arccos\frac{(\mathbf{Ax})^T(\mathbf{Ay})}{\|\mathbf{Ax}\|\|\mathbf{Ay}\|} = \arccos\frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} = \angle(\mathbf{x}, \mathbf{y})$. □

Okazuje się, co dowodzimy w następnym twierdzeniu, że własność (*a*) z twierdzenia 9.10.2 jest także warunkiem dostatecznym ortogonalności macierzy.

Twierdzenie 9.10.3. *Macierz* $\mathbf{A}$ *wymiaru $n \times n$ jest macierzą ortogonalną wtedy i tylko wtedy, gdy* $(\mathbf{Ax})^T(\mathbf{Ay}) = \mathbf{x}^T\mathbf{y}$ *dla każdych dwóch wektorów* $\mathbf{x}$ *i* $\mathbf{y}$ *z przestrzeni R^n.*

Dowód. Załóżmy, że $\mathbf{x}^T\mathbf{A}^T\mathbf{Ay} = (\mathbf{Ax})^T(\mathbf{Ay}) = \mathbf{x}^T\mathbf{y}$, czyli $\mathbf{x}^T(\mathbf{A}^T\mathbf{A} - I_n)\mathbf{y} = \mathbf{0}$ dla każdych wektorów $\mathbf{x}, \mathbf{y} \in R^n$. Wtedy także dla każdego wektora $\mathbf{y} \in R^n$ i dla wektora $\mathbf{x} = (\mathbf{A}^T\mathbf{A} - I_n)\mathbf{y}$ mamy $\big((\mathbf{A}^T\mathbf{A} - I_n)\mathbf{y}\big)^T \cdot \big((\mathbf{A}^T\mathbf{A} - I_n)\mathbf{y}\big) = \mathbf{0}$, co wobec własności iloczynu skalarnego (zob. def. 9.1.1) ma miejsce wtedy i tylko wtedy, gdy $(\mathbf{A}^T\mathbf{A} - I_n)\mathbf{y} = \mathbf{0}$ dla każdego $\mathbf{y} \in R^n$. Stąd $\mathbf{A}^T\mathbf{A} - I_n = 0$ i dlatego macierz $\mathbf{A}$ jest ortogonalna.

Odwrotna implikacja była treścią twierdzenia 9.10.2 (a). □

Pokażemy teraz, że macierz $\mathbf{A}$ jest ortogonalna wtedy i tylko wtedy, gdy przekształcenie $T_{\mathbf{A}} = \mathbf{Ax}$ przekształca bazę ortonormalną w bazę ortonormalną.

Twierdzenie 9.10.4. *Jeśli* $\mathbf{A}$ *jest macierzą wymiaru $n \times n$ i układ wektorów* $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ *jest bazą ortonormalną przestrzeni R^n, to macierz* $\mathbf{A}$ *jest ortogonalna wtedy i tylko wtedy, gdy układ wektorów* $(\mathbf{Aa}_1, \mathbf{Aa}_2, \ldots, \mathbf{Aa}_n)$ *jest bazą ortonormalną przestrzeni R^n.*

Dowód. Konieczność warunku wynika z twierdzenia 9.10.2. Dla dowodu dostateczności zakładamy, że $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ oraz $(\mathbf{b}_1, \ldots, \mathbf{b}_n) = (\mathbf{Aa}_1, \ldots, \mathbf{Aa}_n)$ są bazami ortonormalnymi przestrzeni R^n. Wobec twierdzenia 9.10.3 wystarczy pokazać, że $(\mathbf{Ax})^T(\mathbf{Ay}) = \mathbf{x}^T\mathbf{y}$ dla każdych $\mathbf{x}, \mathbf{y} \in R^n$. Zauważmy, że jeśli $\mathbf{x} = \sum_{i=1}^n x_i\mathbf{a}_i$ i $\mathbf{y} = \sum_{j=1}^n y_j\mathbf{a}_j$, to wobec ortonormalności bazy $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ jest $\mathbf{x}^T\mathbf{y} = \sum_{i=1}^n x_iy_i$. Jednocześnie, ponieważ $\mathbf{Ax} = \mathbf{A}\big(\sum_{i=1}^n x_i\mathbf{a}_i\big) = \sum_{i=1}^n x_i\mathbf{Aa}_i = \sum_{i=1}^n x_i\mathbf{b}_i$ i podobnie $\mathbf{Ay} = \sum_{j=1}^n y_j\mathbf{b}_j$, więc tym razem wobec ortonormalności bazy $(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ mamy $(\mathbf{Ax})^T(\mathbf{Ay}) = \sum_{i=1}^n x_iy_i = \mathbf{x}^T\mathbf{y}$ i to kończy dowód twierdzenia. □

Przykład 219. Z ostatniego twierdzenia łatwo wynika, że macierz obrotu płaszczyzny wokół punktu $(0,0)$ o kąt α, czyli macierz

$$\begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix},$$

jest ortogonalna. (To samo stwierdzenie jest także oczywistą konsekwencją definicji 9.10.1 i/lub twierdzenia 9.10.1.)

Przekształcenie ortogonalne

Definicja 9.10.2. Niech V będzie przestrzenią wektorową z iloczynem skalarnym $(\cdot|\cdot)$. Przekształcenie liniowe $T\colon V \to V$ nazywamy *przekształceniem ortogonalnym*, gdy dla każdych wektorów $\mathbf{x}$, $\mathbf{y} \in V$ jest

$$\big(T(\mathbf{x})|T(\mathbf{y})\big) = (\mathbf{x}|\mathbf{y}). \tag{9.33}$$

Związek przekształcenia ortogonalnego z macierzą ortogonalną przedstawia następujące twierdzenie.

Twierdzenie 9.10.5. *Niech B będzie baza ortonormalną przestrzeni wektorowej V. Przekształcenie liniowe $T\colon V \to V$ jest ortogonalne wtedy i tylko wtedy, gdy jego macierz $[T]_B$ jest ortogonalna.*

Dowód. Niech $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ będzie bazą ortonormalną przestrzeni wektorowej V i niech $[T]_B = \big[\,[T(\mathbf{b}_1)]_B \ldots [T(\mathbf{b}_n)]_B\,\big] = [a_{ij}]$ będzie macierzą przekształcenia T względem bazy B.

Z własności iloczynu skalarnego, liniowości przekształcenia T i z ortonormalności bazy B mamy

$(\mathbf{b}_k|\mathbf{b}_l) = \delta_{kl}$, gdzie

$$\delta_{kl} = \begin{cases} 0, & \text{gdy } k \neq l \\ 1, & \text{gdy } k = l \end{cases}$$

$$\begin{aligned} \big(T(\mathbf{b}_i)|T(\mathbf{b}_j)\big) &= \left(\sum_{k=1}^{n} a_{ki}\mathbf{b}_k \Big| \sum_{l=1}^{n} a_{lj}\mathbf{b}_l\right) = \sum_{k=1}^{n}\sum_{l=1}^{n} a_{ki}a_{lj}(\mathbf{b}_k|\mathbf{b}_l) \\ &= \sum_{k=1}^{n}\sum_{l=1}^{n} a_{ki}a_{lj}\delta_{kl} = \sum_{k=1}^{n} a_{ki}a_{kj} = [T(\mathbf{b}_i)]^T[T(\mathbf{b}_j)]. \end{aligned}$$

Z drugiej strony z ortonormalności bazy B i z ortogonalności przekształcenia T (zob. definicję (9.10.2) mamy

$$\big(T(\mathbf{b}_i)|T(\mathbf{b}_j)\big) = (\mathbf{b}_i|\mathbf{b}_j) = \delta_{ij}.$$

Stąd wynika, że kolumny macierzy $[T]_B$ tworzą bazę ortonormalną przestrzeni R^n, więc macierz $[T]_B$ jest ortogonalna.

Załóżmy teraz, że macierz $[T]_B$ jest ortogonalna. Wtedy

$$\big(T(\mathbf{b}_i)|T(\mathbf{b}_j)\big) = [T(\mathbf{b}_i)]^T[T(\mathbf{b}_j)] = \delta_{ij} = (\mathbf{b}_i|\mathbf{b}_j).$$

Zatem dla dowolnych wektorów $\mathbf{x} = \sum_{i=1}^{n} x_i\mathbf{b}_i$ oraz $\mathbf{y} = \sum_{j=1}^{n} y_j\mathbf{b}_j$ mamy (9.33), bo

$$\begin{aligned} \big(T(\mathbf{x})|T(\mathbf{y})\big) &= \left(\sum_{i=1}^{n} x_iT(\mathbf{b}_i) \Big| \sum_{j=1}^{n} y_jT(\mathbf{b}_j)\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_j\big(T(\mathbf{b}_i)|T(\mathbf{b}_j)\big) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_j(\mathbf{b}_i|\mathbf{b}_j) = \left(\sum_{i=1}^{n} x_i\mathbf{b}_i \Big| \sum_{j=1}^{n} y_j\mathbf{b}_j\right) = (\mathbf{x}|\mathbf{y}), \end{aligned}$$

i dlatego T jest przekształceniem ortogonalnym. □

Przykład 220. Zbadać ortogonalność przekształcenia $T\colon R^3 \to R^3$, gdzie

$$T(x,y,z) = \frac{1}{\sqrt{2}}(x+z, y, x-z).$$

Macierzą przekształcenia T względem bazy kanonicznej E przestrzeni R^3 jest

$$[T]_E = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & -1 \end{bmatrix}.$$

Ponieważ kolumny macierzy $[T]_E$ tworzą ortonormalną bazę przestrzeni R^3, więc macierz $[T]_E$ jest ortogonalna, a przekształcenie T jest ortogonalne.

9.11. Ćwiczenia

1. W przestrzeni R^3 ze standardowym iloczynem skalarnym dane są wektory $\mathbf{x} = (6,-2,3)$ i $\mathbf{y} = (1,2,-3)$. Obliczyć następujące wielkości: $(\mathbf{x}|\mathbf{y})$, $(\mathbf{x}|\mathbf{x})$, $\|\mathbf{x}\|$, $\frac{\mathbf{x}}{\|\mathbf{x}\|}$, $d(\mathbf{x},\mathbf{y})$ i $\frac{(\mathbf{x}|\mathbf{y})}{(\mathbf{y}|\mathbf{y})}\mathbf{y}$.
2. W przestrzeni R^3 z bazą $B = \big((1,-2,3), (0,1,2), (0,0,1)\big)$ iloczyn skalarny określony jest wzorem $(\mathbf{x}|\mathbf{y}) = [\mathbf{x}]_B^T[\mathbf{y}]_B$. Obliczyć $(\mathbf{x}|\mathbf{y})$, $\|\mathbf{x}\|$, $\|\mathbf{y}\|$ i $\|\mathbf{x}+\mathbf{y}\|$, gdy $\mathbf{x} = (1,-4,-1)$ i $\mathbf{y} = (3,-6,10)$.
3. W przestrzeni $C(\langle 0;1\rangle)$ iloczyn skalarny określony jest wzorem $(\mathbf{f}|\mathbf{g}) = \int_0^1 f(x)g(x)\,dx$. Obliczyć $(\mathbf{f}|\mathbf{g})$, $\|\mathbf{f}\|$, $\|\mathbf{g}\|$ i $\|\mathbf{f}-\mathbf{g}\|$, gdy $\mathbf{f}(x) = x$ i $\mathbf{g}(x) = e^x$.
4. Niech $\mathbf{x}$, $\mathbf{y}$ i $\mathbf{z}$ będą wektorami z przestrzeni Euklidesa takimi, że $(\mathbf{x}|\mathbf{y}) = 2$, $(\mathbf{x}|\mathbf{z}) = -3$, $(\mathbf{y}|\mathbf{z}) = 2$, $\|\mathbf{x}\| = 1$, $\|\mathbf{y}\| = 2$ i $\|\mathbf{z}\| = 3$. Obliczyć: (a) $(\mathbf{x}+\mathbf{y}|\mathbf{y}+\mathbf{z})$; (b) $(\mathbf{x}-\mathbf{y}+3\mathbf{z}|2\mathbf{x}+\mathbf{y})$; (c) $\|\mathbf{x}+\mathbf{y}\|$; (d) $\|\mathbf{x}-2\mathbf{y}+4\mathbf{z}\|$.
5. Dana jest macierz $\mathbf{A} \in R_{2\times 2}$ i funkcja $f : R^2 \times R^2 \to R$, która wektorom $\mathbf{x}$ i $\mathbf{y}$ z przestrzeni R^2 przyporządkowuje liczbę $f(\mathbf{x},\mathbf{y}) = \mathbf{x}^T\mathbf{A}\mathbf{y}$. Sprawdzić, czy funkcja ta jest iloczynem skalarnym, gdy:
   (a) $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$; (b) $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$;
   (c) $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$; (d) $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.
6. Niech $w_1, \ldots, w_n$ będą dodatnimi liczbami rzeczywistymi. Wykazać, że funkcja $(\cdot|\cdot)\colon R^n \times R^n \to R$ jest iloczynem skalarnym, gdy
   $$(\mathbf{x}|\mathbf{y}) = w_1x_1y_1 + \ldots + w_nx_ny_n$$
   dla każdych wektorów $\mathbf{x} = (x_1, \ldots, x_n)$ i $\mathbf{y} = (y_1, \ldots, y_n)$ z przestrzeni R^n.
7. Opisać, jak wyznacza się współrzędne wektora względem bazy ortogonalnej. Następnie wektor $(1,2,3)$ przedstawić jako kombinację liniową wektorów ortogonalnej bazy $\big((1,-2,1), (2,1,0), (-1,2,5)\big)$ przestrzeni R^3.
8. Pokazać, że układ $B = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ jest bazą ortogonalną przestrzeni R^3 (ze standardowym iloczynem skalarnym) i wektor $\mathbf{x}$ zapisać jako kombinację Fouriera względem układu $(\mathbf{x}_1/\|\mathbf{x}_1\|, \mathbf{x}_2/\|\mathbf{x}_2\|, \mathbf{x}_3/\|\mathbf{x}_3\|)$, gdy:
   (a) $B = \left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right)$ i $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$;
   (b) $B = \left(\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}\right)$ i $\mathbf{x} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$;
   (c) $B = \left(\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}\right)$ i $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$.
9. Metodą Grama-Schmidta wyznaczyć bazę ortogonalną podprzestrzeni W przestrzeni R^4 (ze standardowym iloczynem skalarnym), gdy:
   (a) $W = \mathcal{L}\big((1,0,1,0), (1,1,0,1), (0,1,1,1)\big)$;
   (b) $W = \mathcal{L}\big((1,2,2,-1), (1,1,-5,3), (3,2,8,-7)\big)$.
10. Metodą Grama-Schmidta w przestrzeni R^n (ze standardowym iloczynem skalarnym) utworzyć bazę ortogonalną z następującej bazy:
   (a) $\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \end{bmatrix}\right)$; (b) $\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right)$;
   (c) $\left(\begin{bmatrix} 1 \\ 2 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 6 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 13 \\ 7 \\ 1 \\ 0 \end{bmatrix}\right)$.
11. Metodą Grama-Schmidta w przestrzeni V z iloczynem skalarnym $(\cdot|\cdot)$ z bazy C utworzyć bazę ortonormalną C' i wyznaczyć współczynniki Fouriera wektora $\mathbf{x}_0$ względem bazy C', gdy:
   (a) $V = R^2$, $(\mathbf{x}|\mathbf{y}) = \mathbf{x}^T\mathbf{y}$, $C = \big((1,1), (1,2)\big)$ i $\mathbf{x}_0 = (3,4)$; (b) $V = R^3$, $(\mathbf{x}|\mathbf{y}) = [\mathbf{x}]_C^T[\mathbf{y}]_C$, gdzie $C = \big((1,0,1), (0,1,1), (1,1,1)\big)$ i $\mathbf{x}_0 = (1,-1,0)$; (c) $V = R_2[x]$, $(f|g) = \int_0^1 f(t)g(t)\,dt$, $C = (1, x, x^2)$ i $f_0(x) = 1 + x^2$.
12. Wskazać ortogonalną bazę przestrzeni R^4 zawierającą wektory $(1,2,-1,0)$ i $(2,-1,0,1)$.
13. W przestrzeni R^3 ze standardowym iloczynem skalarnym wyznaczyć ortogonalne dopełnienia zbiorów $S_1 = \{(1,1,1)\}$, $S_2 = \{(1,1,1), (1,2,-1)\}$ i $S_3 = \{(x,y,z) \in R^3\colon\ x = 0 \text{ i } y + z = 0\}$.

14. W przestrzeni R^4 wyznaczyć bazę ortogonalnego dopełnienia podprzestrzeni generowanej przez wektory: (a) $(2,0,1,2)$, $(1,0,0,1)$ i $(3,0,1,3)$; (b) $(1,0,2,1)$, $(2,1,2,3)$ i $(0,1,-2,1)$.
15. W przestrzeni R^4 wyznaczyć ortogonalną bazę ortogonalnego dopełnienia podprzestrzeni S generowanej przez wektory $(1,2,3,1)$ i $(0,0,1,2)$.
16. Znaleźć bazę, wymiar i samo ortogonalne dopełnienie $L^\perp$ podprzestrzeni $L \subseteq R^4$, która jest zbiorem rozwiązań następującego jednorodnego układu równań
$$\begin{cases} 2x_1 + x_2 + 3x_3 - x_4 = 0, \\ 3x_1 + 2x_2 - 2x_4 = 0, \\ 3x_1 + x_2 + 9x_3 - x_4 = 0. \end{cases}$$
17. Wskazać ortonormalną bazę jądra przekształcenia liniowego $T\colon R^3 \to R$, gdzie $T(x,y,z) = 2x + y + 3z$. Następnie wyznaczyć rzut ortogonalny wektora $(1,1,0)$ na podprzestrzeń $\operatorname{Ker} T$. Dodatkowo, obliczyć kąt pomiędzy wektorem $(1,1,0)$ i płaszczyzną $\operatorname{Ker} T$.
18. Znaleźć rzut ortogonalny wektora $\mathbf{b} = (1,2,2,7)$ na płaszczyznę $x + y + z + u = 0$.
19. Wyznaczyć rzut ortogonalny wektora $\mathbf{b}$ na podprzestrzeń W przestrzeni R^n (ze standardowym iloczynem skalarnym), gdy:
(a) $\mathbf{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $W = \mathcal{L}\left(\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right)$;
(b) $\mathbf{b} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$, $W = \mathcal{L}\left(\begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}\right)$;
(c) $\mathbf{b} = \begin{bmatrix} 4 \\ 0 \\ 1 \\ 2 \end{bmatrix}$, $W = \mathcal{L}\left(\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ -1 \\ 2 \end{bmatrix}\right)$;
(d) $\mathbf{b} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$, $W = \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \\ 2 \end{bmatrix}\right)$.
20. W przestrzeni V z podanym iloczynem skalarnym wyznaczyć (1) rzut ortogonalny wektora $\mathbf{b}$ na podprzestrzeń W i (2) odległość d pomiędzy wektorem $\mathbf{b}$ i podprzestrzenią W, gdy: (a) $V = R^2$ ze standardowym iloczynem skalarnym, $\mathbf{b} = (4,7)$, $W = \{(x,y)\colon x = 2y\}$; ($b$) $V = R^3$ ze standardowym iloczynem skalarnym, $\mathbf{b} = (14,3,2)$, $W = \{(x,y,z)\colon x + 2y - 2z = 0\}$; ($c$) $V = R_2[x]$ z iloczynem skalarnym $(f|g) = \int_0^1 f(t)g(t)\,dt$, $\mathbf{b} = 1 + 3x + 2x^2$, $W = R_1[x]$.
21. (a) Wyznaczyć rzut ortogonalny wektora $\mathbf{b} = (1,2,3)$ na płaszczyznę π zawierającą wektory $\mathbf{a}_1 = (2,2,-1)$ i $\mathbf{a}_2 = (2,-1,2)$. (b) Metodą Grama-Schmidta otrzymać ortonormalny układ wektorów $(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)$ z układu $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{b})$. ($c$) Znaleźć macierz $\mathbf{P}$ rzutu na płaszczyznę π.
22. Wyznaczyć: (1) najlepszą aproksymację wektora $\mathbf{b}$ wektorami z podprzestrzeni W (czyli wyznaczyć wektor $\mathbf{b}_0 \in W$ taki, że $\|\mathbf{b} - \mathbf{b}_0\| = \min\{\|\mathbf{b} - \mathbf{x}\|\colon \mathbf{x} \in W\}$); (2) wektor błędu tej aproksymacji; (3) wielkość błędu tej aproksymacji oraz (4) odległość pomiędzy wektorem $\mathbf{b}$ i podprzestrzenią W, gdy:
(a) $\mathbf{b} = \begin{bmatrix} 1 \\ 29 \\ 1 \end{bmatrix}$ i $W = \mathcal{L}\left(\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix}\right)$;
(b) $\mathbf{b} = \begin{bmatrix} 5 \\ -1 \\ 1 \\ 1 \end{bmatrix}$ i $W = \mathcal{L}\left(\begin{bmatrix} 3 \\ 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 1 \\ -2 \end{bmatrix}\right)$.
23. Znaleźć najlepsze rozwiązanie sprzecznego układu równań:
(a) $\begin{cases} x + 2y = 4, \\ x + y = 5, \\ 3x + 5y = 12; \end{cases}$ (b) $\begin{cases} 2x = 1, \\ y = 2, \\ 2x + 2y = 3. \end{cases}$
24. W przestrzeni R^4 dane są wektory
$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 8 \\ -6 \\ 2 \\ 0 \end{bmatrix} \quad \text{i} \quad \mathbf{b} = \begin{bmatrix} 7 \\ -6 \\ 4 \\ -5 \end{bmatrix}.$$
(a) Algorytmem Grama-Schmidta dokonać ortogonalizacji układu $(\mathbf{v}_1, \mathbf{v}_2)$. ($b$) Wyznaczyć rzut ortogonalny wektora $\mathbf{b}$ na podprzestrzeń $V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2)$. ($c$) Wskazać macierz $\mathbf{A}$, której przestrzenią zerową jest przestrzeń $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2)$.
25. Szukamy prostej $y = Cx + D$ najbliższej punktom $(0,-1)$, $(1,2)$ i $(2,-1)$. (a) Metodą najmniejszych kwadratów wyznaczyć współczynniki C i D. (b) Wyjaśnić jak ma się wektor $\mathbf{b} = (-1,2,-1)$ do płaszczyzny, na którą rzutowano? (c) Wyznaczyć długość wektora błędu $\mathbf{e}$ (= odległość od płaszczyzny $= \|\mathbf{b} - \mathbf{A}\overline{\mathbf{x}}\|$).
26. Wyznaczyć najlepszą (w sensie metody najmniejszych kwadratów) zależność postaci $y = ax + b$ między współrzędnymi x_i oraz y_i punktów $(1,2)$, $(2,3)$, $(3,5)$ i $(4,7)$.
27. Rozwiązaniem sprzecznego układu równań $\mathbf{Ax} = \mathbf{b}$, gdzie $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$ i $\mathbf{b} = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}$, jest macierz $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 11/3 \\ -1 \end{bmatrix}$. ($a$) Wskazać rzut ortogonalny $\overline{\mathbf{b}}$ wektora $\mathbf{b}$ na przestrzeń kolumnową macierzy $\mathbf{A}$. (b) Wyznaczyć bazę ortonormalną $(\mathbf{q}_1, \mathbf{q}_2)$ przestrzeni kolumnowej macierzy $\mathbf{A}$. (c) Wyznaczyć macierz $\mathbf{P}$ rzutu ortogonalnego na przestrzeń kolumnową macierzy $\mathbf{A}$.
28. Wyznaczyć rozwiązanie układu
$$\begin{cases} x_1 + x_2 + x_3 = 0 \\ 2x_1 + 2x_2 + 2x_3 + 3x_4 = 0 \end{cases}$$
najbliższe wektorowi $\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^T$.
29. Dane są macierze $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{bmatrix}$ i $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 7 \end{bmatrix}$.
(a) Wskazać trzy ortonormalne wektory $\mathbf{x}_1$, $\mathbf{x}_2$ i $\mathbf{x}_3$ takie, że układ $(\mathbf{x}_1, \mathbf{x}_2)$ jest bazę przestrzeni kolumnowej $C_\mathbf{A}$ macierzy $\mathbf{A}$. (b) Która z podprzestrzeni $N_\mathbf{A}$, $N_{\mathbf{A}^T}$, $C_\mathbf{A}$ i $C_{\mathbf{A}^T}$ zawiera wektor $\mathbf{x}_3$? (c) Wyznaczyć macierz $\mathbf{P}$ rzutu na podprzestrzeń $N_{\mathbf{A}^T}$. (d) Wyznaczyć najlepsze rozwiązanie równania $\mathbf{Ax} = \mathbf{b}$.
30. (a) Czy macierz $\mathbf{A} = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 1 & -3 \\ 2 & 2 & -4 \end{bmatrix}$ jest odwracalna? (b) Wskazać bazę ortonormalną przestrzeni kolumnowej macierzy $\mathbf{A}$. (c) Dlaczego macierz $\mathbf{P} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ nie jest macierzą rzutu na przestrzeń kolumnową macierzy $\mathbf{A}$?

31. Pokazać, że w przestrzeni wektorowej V z iloczynem skalarnym $(\cdot|\cdot)$ dla każdych wektorów $\mathbf{x}$ i $\mathbf{y}$ jest
$$(\mathbf{x}|\mathbf{y}) = \tfrac{1}{4}\left(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\right).$$
32. Niech $f(\cdot,\cdot)$ i $g(\cdot,\cdot)$ będą iloczynami skalarnymi w przestrzeni wektorowej V. Wykazać, że funkcja
$$h(\cdot,\cdot) = f(\cdot,\cdot) + g(\cdot,\cdot)$$
także jest iloczynem skalarnym w przestrzeni V.
33. Wykazać, że funkcja $f : R_{n\times n} \times R_{n\times n} \to R$ jest iloczynem skalarnym, gdy $f(\mathbf{A},\mathbf{B}) = \operatorname{tr}(\mathbf{A}\mathbf{B}^T)$.
34. Indukcyjnie udowodnić, że jeśli wektory $\mathbf{v}_1,\ldots,\mathbf{v}_n$ są wzajemnie ortogonalne, to
$$\|\mathbf{v}_1\|^2 + \|\mathbf{v}_2\|^2 + \ldots + \|\mathbf{v}_n\|^2 = \|\mathbf{v}_1 + \mathbf{v}_2 + \ldots + \mathbf{v}_n\|^2.$$
35. Niech $(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ będzie ortonormalnym układem wektorów w przestrzeni Euklidesa. Udowodnić, że
$$\|\mathbf{v}_1 + \ldots + \mathbf{v}_n\| = \sqrt{n}.$$
36. Za pomocą nierówności Schwarza udowodnić, że dla liczb rzeczywistych a, b i φ jest
$$(a\cos\varphi + b\sin\varphi)^2 \leqslant a^2 + b^2.$$
37. Za pomocą nierówności Schwarza udowodnić, że dla dodatnich liczb $a_1, a_2, \ldots, a_n$ jest
$$(a_1 + a_2 + \ldots + a_n)\left(\frac{1}{a_1} + \frac{1}{a_2} + \ldots + \frac{1}{a_n}\right) \geqslant n^2.$$
38. Niech $(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ będzie bazą ortonormalną przestrzeni V. Wykazać, że dla każdych wektorów $\mathbf{x}, \mathbf{y} \in V$ spełniona jest tzw. równość Parsevala $(\mathbf{x}|\mathbf{y}) = \sum_{i=1}^n (\mathbf{x}|\mathbf{v}_i)(\mathbf{y}|\mathbf{v}_i)$. Następnie wykazać, że dla każdego wektora $\mathbf{x} \in V$ jest $\|\mathbf{x}\| = \sum_{i=1}^n (\mathbf{x}|\mathbf{v}_i)^2$.
39. Niech $(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ będzie układem ortonormalnych wektorów w przestrzeni Euklidesa V. Wykazać, że dla każdego $\mathbf{x} \in V$ zachodzi nierówność $\|\mathbf{x}\|^2 \geqslant \sum_{i=1}^n (\mathbf{x}|\mathbf{v}_i)^2$. Jest to tzw. nierówność Bessela i wyraża ona fakt, że norma rzutu ortogonalnego wektora $\mathbf{x}$ na podprzestrzeń nie jest większa od normy samego wektora $\mathbf{x}$.
40. Wykazać, że $W \cap W^\perp = \{\mathbf{0}\}$ dla każdej podprzestrzeni W przestrzeni Euklidesa.
41. Wykazać, że jeśli S i T są niepustymi zbiorami wektorów w przestrzeni Euklidesa V, a W jest podprzestrzenią przestrzeni V, to prawdziwe są następujące stwierdzenia: (*a*) $S^\perp$ jest podprzestrzenią przestrzeni V; (*b*) jeśli $S \subseteq T$, to $T^\perp \subseteq S^\perp$; (*c*) $S \subseteq (S^\perp)^\perp$; (*d*) $W = (W^\perp)^\perp$; (*e*) $V = W \oplus W^\perp$.
42. Niech U_1 i U_2 będą podprzestrzeniami przestrzeni Euklidesa. Udowodnić, że $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$ i $(U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp$.
43. Udowodnić, że 0 lub 1 jest wyznacznikiem macierzy rzutu.
44. Wykazać, że jeśli macierze $\mathbf{A}$ i $\mathbf{B}$ z przestrzeni $R_{n\times n}$ są ortogonalne, to także macierze $\mathbf{AB}$ i $\mathbf{A}^2$ są ortogonalne.
45. Wykazać ortogonalność macierzy
$$\mathbf{A} = \tfrac{1}{1+2a^2}\begin{bmatrix} 1 & -2a & 2a^2 \\ 2a & 1-2a^2 & -2a \\ 2a^2 & 2a & 1 \end{bmatrix}$$
dla każdej liczby $a \in R$.
46. Wykazać, że jeśli macierz $\mathbf{A}$ jest macierzą ortogonalną wymiaru $n \times n$, to $\|\mathbf{A}\mathbf{x}\| = \|\mathbf{A}^{-1}\mathbf{x}\|$ dla każdego wektora $\mathbf{x} \in R^n$.
47. Zbadać ortogonalność następujących przekształceń liniowych: (*a*) $T(x,y) = (y,x)$; (*b*) $T(x,y) = (x,-y)$; (*c*) $T(x,y) = \frac{1}{2}(x - \sqrt{3}y, \sqrt{3}x + y)$; (*d*) $T(x,y,z) = (3/3{+}2y/3{+}2z/3, -2x/3{-}y/3{+}2z/3, -2x/3{+}2y/3{-}z/3)$.
48. Niech V będzie przestrzenią wektorową z iloczynem skalarnym $(\,\cdot\,|\,\cdot\,)$ i niech $B = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ będzie bazą ortonormalną przestrzeni V. Niech $T\colon V \to V$ będzie takim przekształceniem, że $T(\mathbf{x}) = (\mathbf{x}|\mathbf{a})\mathbf{a}$, gdzie $\mathbf{a} \in V$ i $[\mathbf{a}]_B = (1,-1,3)$. (*a*) Pokazać, że T jest przekształceniem liniowym. (*b*) Znaleźć macierz $[T]_B$ przekształcenia T.
49. Wpisując TAK albo NIE, stwierdzić prawdziwość każdego z następujących zdań:

$\boxed{1}$ Jeśli $\mathbf{x}$, $\mathbf{y}$ i $\mathbf{z}$ są wektorami z przestrzeni Euklidesa i $(\mathbf{x}|\mathbf{y}) = (\mathbf{x}|\mathbf{z})$, to $\mathbf{y} = \mathbf{z}$.

$\boxed{2}$ Jeśli $\mathbf{x}$ jest wektorem z przestrzeni Euklidesa i $(\mathbf{x}|\mathbf{y}) = 0$ dla każdego wektora $\mathbf{y}$, to $\mathbf{x} = \mathbf{0}$.

$\boxed{3}$ W przestrzeni R^n istnieje tylko jeden iloczyn skalarny.

$\boxed{4}$ W przestrzeni R^2 istnieje iloczyn skalarny $(\cdot|\cdot)$ względem którego wektory $\mathbf{x} = (1,1)$ i $\mathbf{y} = (1,2)$ są ortogonalne.

$\boxed{5}$ Funkcja $f : R_{2\times2} \times R_{2\times2} \to R$ jest iloczynem skalarnym, gdy $f(\mathbf{A},\mathbf{B}) = \operatorname{tr}(\mathbf{A}+\mathbf{B})$ dla każdych $\mathbf{A}, \mathbf{B} \in R_{2\times2}$.

$\boxed{6}$ Metoda Grama-Schmidta umożliwia konstrukcję ortogonalnego zbioru wektorów z dowolnego zbioru wektorów.

$\boxed{7}$ Każda przestrzeń Euklidesa ma ortonormalną bazę.

$\boxed{8}$ Ortogonalne dopełnienie każdego zbioru jest podprzestrzenią.

$\boxed{9}$ Jeśli $(\mathbf{b}_1,\ldots,\mathbf{b}_n)$ jest bazą przestrzeni Euklidesa V i $\mathbf{x} \in V$, to liczby $(\mathbf{x}|\mathbf{b}_i)$ są współczynnikami Fouriera wektora $\mathbf{x}$.

$\boxed{10}$ Każdy zbiór ortonormalny jest liniowo niezależny.

$\boxed{11}$ Każdy zbiór liniowo niezależnych wektorów jest ortogonalny.

$\boxed{12}$ Dla każdych wektorów $\mathbf{x}$ i $\mathbf{y}$ z przestrzeni Euklidesa mamy $\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2$.

$\boxed{13}$ Jeśli $\mathbf{U}$ jest macierzą wymiaru $n \times p$ i $\mathbf{b} \in R^n$, to $\mathbf{U}\mathbf{U}^T\mathbf{b} = \overline{\mathbf{b}}$ jest rzutem ortogonalnym wektora $\mathbf{b}$ na przestrzeń kolumnową macierzy $\mathbf{U}$.

$\boxed{14}$ Jeśli $\mathbf{U}$ jest macierzą wymiaru $n \times p$ i jej kolumny są ortogonalne, to $\mathbf{U}\mathbf{U}^T\mathbf{b} = \mathbf{b}$ dla każdego $\mathbf{b} \in R^n$.

$\boxed{15}$ Jeśli W jest podprzestrznią przestrzeni Euklidesa V i $\mathbf{b} \in V$, to $\mathbf{b} - \operatorname{proj}_W \mathbf{b}$ jest najlepszą aproksymacją wektora $\mathbf{b}$ za pomocą wektorów z podprzestrzeni W.

$\boxed{16}$ Jeśli macierz $\mathbf{A}$ jest ortogonalna, to także macierze $\mathbf{A}^T$ i $\mathbf{A}^{-1}$ są ortogonalne.

$\boxed{17}$ Jeśli $\mathbf{A}$ jest symetryczną macierzą ortogonalną, to $\mathbf{A}^2 = \mathbf{I}$.

Rozdział 10

WARTOŚCI WŁASNE I WEKTORY WŁASNE

10.1. Wartości własne i wektory własne macierzy i operatora

λ – wartość własna

Definicja 10.1.1. Niech V będzie przestrzenią wektorową nad ciałem K (gdzie $K = R$ albo $K = \mathcal{C}$) i niech $T\colon V \to V$ będzie operatorem liniowym na przestrzeni V. Skalar $\lambda \in K$ nazywamy *wartością własną* operatora T, jeśli istnieje niezerowy wektor $\mathbf{v}_0 \in V$ taki, że

$$T(\mathbf{v}_0) = \lambda\mathbf{v}_0. \tag{10.1}$$

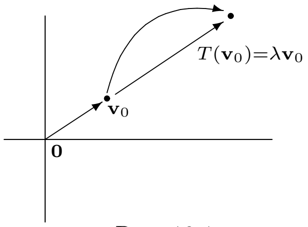
$\mathbf{v}_0$ – wektor własny

W takim przypadku mówimy, że $\mathbf{v}_0$ jest *wektorem własnym* operatora T odpowiadającym wartości własnej λ. Geometryczny efekt działania operatora liniowego $T\colon R^n \to R^n$ na wektor własny $\mathbf{v}_0$ przedstawia rys. 10.1.

Podobnie, skalar $\lambda \in K$ nazywamy *wartością własną* macierzy kwadratowej $\mathbf{A} \in K_{n\times n}$, jeśli istnieje niezerowy wektor $\mathbf{v}_0 \in K^n$, dla którego

$$\mathbf{A}\mathbf{v}_0 = \lambda\mathbf{v}_0. \tag{10.2}$$

Taki wektor $\mathbf{v}_0$ nazywamy *wektorem własnym* macierzy $\mathbf{A}$ odpowiadającym wartości własnej λ.



Rys. 10.1

Przykład 221. Niech V będzie przestrzenią wektorową nad ciałem K i niech λ będzie dowolnym elementem z ciała K. Jeśli T jest operatorem liniowym na przestrzeni V takim, że $T(\mathbf{x}) = \lambda\mathbf{x}$, to każdy niezerowy wektor $\mathbf{x} \in V$ jest wektorem własnym operatora T odpowiadającym wartości własnej λ.

Łatwo sprawdza się, czy dany wektor jest wektorem własnym macierzy (operatora). Równie łatwo bada się, czy dana liczba jest wartością własną macierzy (operatora).

Przykład 222. Sprawdzić, czy wektor $\mathbf{v}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ jest wektorem własnym macierzy $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$.

Z równości

$$\mathbf{A}\mathbf{v}_0 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 3\mathbf{v}_0$$

wynika, że wektor $\mathbf{v}_0$ jest wektorem własnym macierzy $\mathbf{A}$ odpowiadającym wartości własnej $\lambda = 3$.

Przykład 223. Sprawdzić, czy liczba $\lambda = 4$ jest wartością własną macierzy

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

Ponieważ

$$\det(\mathbf{A}-4\mathbf{I}) = \begin{vmatrix} 0 & 0 & 1 \\ -2 & -3 & 0 \\ -2 & 0 & -3 \end{vmatrix} = -6 \neq 0,$$

więc równanie $(\mathbf{A}-4\mathbf{I})\mathbf{x} = \mathbf{0}$ (i równoważne z nim równanie $\mathbf{A}\mathbf{x} = 4\mathbf{x}$) ma tylko zerowe rozwiązanie. Stąd wynika, że liczba 4 nie jest wartością własną macierzy $\mathbf{A}$.

Przykład 224. Niech C_R^∞ będzie przestrzenią funkcji $f\colon R \to R$ mających pochodne każdego rzędu. Niech $T\colon C_R^\infty \to C_R^\infty$ będzie funkcją taką, że $T(f) = f'$ dla $f \in C_R^\infty$.

Zauważmy, że funkcja $f \in C_R^\infty$ jest wektorem własnym operatora T odpowiadającym wartości własnej $\lambda \in R$ wtedy i tylko wtedy, gdy

$$f' = T(f) = \lambda f,$$

tj. wtedy i tylko wtedy, gdy f jest rozwiązaniem równania różniczkowego $y' = \lambda y$. (Rozwiązaniem takiego równania jest funkcja $f(t) = ce^{\lambda t}$ dla każdej stałej rzeczywistej c.)

Po tych kilku przykładach ilustrujących pojęcie wartości własnej i wektora własnego macierzy i operatora liniowego wracamy do ogólnych rozważań. Przede wszystkim zauważmy, że zależność (10.2) można zapisać w postaci równości $(\mathbf{A}-\lambda\mathbf{I}_n)\mathbf{v}_0 = \mathbf{0}$. Stąd wynika, że wektor $\mathbf{v}_0 \in K^n$ jest wektorem własnym macierzy $\mathbf{A}$ odpowiadającym wartości własnej λ wtedy i tylko wtedy, gdy jest on niezerowym rozwiązaniem równania jednorodnego

$$(\mathbf{A}-\lambda\mathbf{I}_n)\mathbf{x} = \mathbf{0}. \tag{10.3}$$

Wartością własną macierzy $\mathbf{A}$ jest więc każda liczba λ, dla której równanie jednorodne (10.3) ma niezerowe rozwiązanie $\mathbf{x}$. Z wniosku 6.5.2 wiadomo, że takie niezerowe rozwiązanie istnieje wtedy i tylko wtedy, gdy macierz $\mathbf{A}-\lambda\mathbf{I}_n$ jest osobliwa, co ma miejsce wtedy i tylko wtedy, gdy

$$\det(\mathbf{A}-\lambda\mathbf{I}_n) = 0. \tag{10.4}$$

Stąd otrzymujemy twierdzenie, z którego dalej korzystamy, wyznaczając wartości własne i wektory własne macierzy.

Twierdzenie 10.1.1. (*a*) *Liczba $\lambda \in K$ jest wartością własną macierzy $\mathbf{A} \in K_{n\times n}$ wtedy i tylko wtedy, gdy λ jest rozwiązaniem równania* (10.4).

(*b*) *Wektor $\mathbf{v} \in K^n$ jest wektorem własnym macierzy $\mathbf{A} \in K_{n\times n}$ odpowiadającym wartości własnej λ wtedy i tylko wtedy, gdy $\mathbf{v}$ jest niezerowym rozwiązaniem równania* (10.3). $\square$

Dla macierzy $\mathbf{A} = [a_{ij}]_{n\times n}$ równanie (10.4) przyjmuje postać

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0$$

i jego lewa strona, $\det(\mathbf{A}-\lambda\mathbf{I}_n)$, jest wielomianem zmiennej λ. Wielomian ten nazywamy *wielomianem charakterystycznym* macierzy $\mathbf{A}$, a równanie

$$\det(\mathbf{A}-\lambda\mathbf{I}_n) = 0$$

Równanie charakterystyczne

– *równaniem charaktrystycznym* macierzy $\mathbf{A}$. Można pokazać, że wielomian charakterystyczny $\det(\mathbf{A} - \lambda\mathbf{I}_n)$ jest wielomianem stopnia n,

$\det(\mathbf{A} - \lambda\mathbf{I}_n)$ – wielomian charakterystyczny

$$\det(\mathbf{A} - \lambda\mathbf{I}_n) = (-1)^n\big(\lambda^n - a_1\lambda^{n-1} + a_2\lambda^{n-1} - \ldots + (-1)^n a_n\big),$$

i każdy jego współczynnik a_k jest sumą wszystkich minorów głównych[1] stopnia k macierzy $\mathbf{A}$. W szczególności mamy

$$a_1 = a_{11} + a_{22} + \ldots + a_{nn} = \operatorname{tr}(\mathbf{A}) \quad \text{i} \quad a_n = \det(\mathbf{A}).$$

Zatem dla macierzy $\mathbf{A}$ wymiaru 2×2, powiedzmy dla macierzy $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, jest

$$\begin{aligned} \det(\mathbf{A} - \lambda\mathbf{I}_2) &= \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = (a-\lambda)(d-\lambda) - bc \\ &= \lambda^2 - \lambda(a+d) + (ad - bc) \\ &= \lambda^2 - \lambda\operatorname{tr}(\mathbf{A}) + |\mathbf{A}|. \end{aligned}$$

Podobnie, jeśli $\mathbf{A}$ jest macierzą wymiaru 3×3, to

$$\det(\mathbf{A} - \lambda\mathbf{I}_3) = -\big(\lambda^3 - \lambda^2\operatorname{tr}(\mathbf{A}) + \lambda\big(|\mathbf{A}_{11}| + |\mathbf{A}_{22}| + |\mathbf{A}_{33}|\big) - |\mathbf{A}|\big).$$

Przykład 225. Wyznaczyć wielomian charakterystyczny $\varphi(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}_3)$ macierzy

$$\mathbf{A} = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}.$$

Mamy

$$\begin{aligned} \varphi(\lambda) &= -\lambda^3 + \lambda^2(4+1+8) - \lambda\left(\begin{vmatrix} 1 & 6 \\ -1 & 8 \end{vmatrix} + \begin{vmatrix} 4 & 6 \\ 2 & 8 \end{vmatrix} + \begin{vmatrix} 4 & -1 \\ 2 & 1 \end{vmatrix}\right) + \begin{vmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{vmatrix} \\ &= -\lambda^3 + 13\lambda^2 - 40\lambda + 36. \end{aligned}$$

Niech $T\colon V \to V$ będzie operatorem liniowym na skończenie wymiarowej przestrzeni wektorowej V i niech $\mathbf{A}$ będzie macierzą operatora T względem jakiejkolwiek bazy B przestrzeni V, czyli $\mathbf{A} = [T]_B$. Twierdzimy, że wielomian $\det(\mathbf{A} - \lambda\mathbf{I})$ nie zależy od wyboru bazy przestrzeni V. Tak jest istotnie, bo jeśli C jest inną bazą przestrzeni V, to macierze $[T]_B$ i $[T]_C$ są podobne (zob. tw. 8.40) i wobec twierdzenia 6.4.1 wyznaczniki $\det([T]_B - \lambda\mathbf{I})$ i $\det([T]_C - \lambda\mathbf{I})$ są identyczne. Wielomian $\varphi(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I})$, gdzie $\mathbf{A}$ jest macierzą operatora T względem dowolnej bazy przestrzeni V, nazywamy *wielomianem charakterystycznym operatora* T.

Wielomian charakterystyczny operatora

Przykład 226. Wyznaczyć wielomian charakterystyczny operatora liniowego $T\colon R^2 \to R^2$, gdzie

$$T(x_1, x_2) = (2x_1 + x_2, 3x_1 - x_2).$$

Ponieważ macierzą operatora T względem bazy standardowej E przestrzeni R^2 jest

$$[T]_E = \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix},$$

[1] Niech $\mathbf{A}$ będzie macierzą kwadratową stopnia n i niech $i_1, i_2, \ldots, i_k$ będą liczbami naturalnymi takimi, że $1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n$. *Minorem głównym stopnia* k macierzy $\mathbf{A}$ nazywamy wyznacznik macierzy kwadratowej stopnia k powstałej z macierzy $\mathbf{A}$ przez wykreślenie z niej wszystkich wierszy i wszystkich kolumn o numerach różnych od $i_1, i_2, \ldots, i_k$.

więc wielomianem charakterystycznym operatora T jest

$$\varphi(\lambda) = \det\left([T]_E - \lambda\mathbf{I}\right) = \begin{vmatrix} 2-\lambda & 1 \\ 3 & -1-\lambda \end{vmatrix} = \lambda^2 - \lambda - 5.$$

Dla macierzy $\mathbf{A} \in \mathcal{C}_{n\times n}$ równanie charakterystyczne $\det(\mathbf{A} - \lambda\mathbf{I}_n) = 0$ jest równaniem wielomianowym stopnia n i wobec zasadniczego twierdzenia algebry ma ono n (niekoniecznie różnych) rozwiązań zespolonych. Stąd wynika, że macierz kwadratowa stopnia n ma n (niekoniecznie różnych) zespolnych wartości własnych. Zbiór wszystkich wartości własnych macierzy $\mathbf{A}$ nazywamy *widmem* tej macierzy.

Widmo macierzy

Przykład 227. Znaleźć wartości własne macierzy $\mathbf{A} = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$.

Równaniem charakterystycznym macierzy $\mathbf{A}$ jest

$$\det(\mathbf{A} - \lambda\mathbf{I}_2) = \begin{vmatrix} 5-\lambda & 4 \\ 1 & 2-\lambda \end{vmatrix} = (5-\lambda)(2-\lambda) - 4 = \lambda^2 - 7\lambda + 6 = 0.$$

Zatem wartościami własnymi macierzy $\mathbf{A}$ są liczby $\lambda_1 = 1$ i $\lambda_2 = 6$.

Przypomnijmy, że wobec twierdzenia 10.1.1 wektory własne macierzy $\mathbf{A}$ odpowiadające wartości własnej λ_0 są niezerowymi rozwiązaniami równania $(\mathbf{A} - \lambda_0\mathbf{I}_n)\mathbf{x} = \mathbf{0}$.

Przykład 228. Wyznaczyć wartości własne i wektory własne macierzy $\mathbf{A}$, gdy

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 4 \\ 0 & 6 & 0 \\ 4 & 0 & 2 \end{bmatrix}.$$

Rozwiązując równanie charakterystyczne $\det(\mathbf{A} - \lambda\mathbf{I}_3) = 0$, znajdujemy najpierw wszystkie wartości własne macierzy $\mathbf{A}$. Ponieważ

$$\begin{aligned} \det(\mathbf{A} - \lambda\mathbf{I}_3) &= \begin{vmatrix} 2-\lambda & 0 & 4 \\ 0 & 6-\lambda & 0 \\ 4 & 0 & 2-\lambda \end{vmatrix} \\ &= (6-\lambda)(2-\lambda)^2 - 16(6-\lambda) = (6-\lambda)(\lambda+2)(\lambda-6), \end{aligned}$$

więc wartościami własnymi macierzy $\mathbf{A}$ są $\lambda_1 = -2$ i $\lambda_2 = \lambda_3 = 6$.

Rozwiązując teraz równanie macierzowe $(\mathbf{A} - \lambda_i\mathbf{I}_3)\mathbf{x} = \mathbf{0}$, znajdujemy wektory własne macierzy $\mathbf{A}$ odpowiadające wartości własnej λ_i.

Dla $\lambda_1 = -2$ wyznaczamy $\mathbf{x} = (x_1, x_2, x_3)$ z równania $(\mathbf{A} - \lambda_1\mathbf{I}_3)\mathbf{x} = \mathbf{0}$. Mamy

$$[\,\mathbf{A} - \lambda_1\mathbf{I}_3 \,|\, \mathbf{0}\,] = [\,\mathbf{A} + 2\mathbf{I}_3 \,|\, \mathbf{0}\,] = \left[\begin{array}{ccc|c} 4 & 0 & 4 & 0 \\ 0 & 8 & 0 & 0 \\ 4 & 0 & 4 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Zatem

$$\mathbf{x} = \begin{bmatrix} -x_3 \\ 0 \\ x_3 \end{bmatrix}, \quad x_3 \in R,$$

jest rozwiązaniem równania $(\mathbf{A} - \lambda_1\mathbf{I}_3)\mathbf{x} = \mathbf{0}$. Dlatego każdy wektor

$$\mathbf{v}_1 = x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad x_3 \in R - \{0\},$$

jest wektorem własnym macierzy $\mathbf{A}$ odpowiadającym wartości własnej $\lambda_1 = -2$.

Dla $\lambda_2 = \lambda_3 = 6$ mamy

$$[\,\mathbf{A} - \lambda_2\mathbf{I}_3 \,|\, \mathbf{0}\,] = [\,\mathbf{A} - 6\mathbf{I}_3 \,|\, 0\,] = \left[\begin{array}{ccc|c} -4 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \\ 4 & 0 & -4 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

i każdy niezerowy wektor

$$\mathbf{x} = \begin{bmatrix} x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

jest wektorem własnym macierzy $\mathbf{A}$ odpowiadającym wartości własnej $\lambda_2 = \lambda_3 = 6$. W szczególności wektory

$$\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{i} \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

są wektorami własnymi macierzy $\mathbf{A}$ odpowiadającymi wartości własnej $\lambda_2 = \lambda_3 = 6$.

Następujące twierdzenie pokazuje, że istnieją ścisłe zależności pomiędzy wartościami i wektorami własnymi operatora liniowego, a wartościami i wektorami własnymi macierzy tego operatora. Zależności te bywają pomocne przy wyznaczaniu wartości własnych i wektorów własnych operatora liniowego.

Twierdzenie 10.1.2. *Niech T będzie operatorem liniowym na skończenie wymiarowej przestrzeni wektorowej V i niech $[T]_B$ będzie macierzą operatora T względem bazy B przestrzeni V.*

(a) Liczba λ jest wartością własną operatora T wtedy i tylko wtedy, gdy λ jest wartością własną macierzy $[T]_B$.

(b) Wektor $\mathbf{v}_0 \in V$ jest wektorem własnym operatora T odpowiadającym wartości własnej λ wtedy i tylko wtedy, gdy wektor $[\mathbf{v}_0]_B$ jest wektorem własnym macierzy $[T]_B$ odpowiadającym wartości własnej λ.

Dowód. Obie części twierdzenia wynikają z następujących równoważności.

Wektor $\mathbf{v}_0 \in V$ jest wektorem własnym operatora T odpowiadającym wartości własnej λ

$\Leftrightarrow T(\mathbf{v}_0) = \lambda\mathbf{v}_0$ i $\mathbf{v}_0 \neq \mathbf{0}$ (definicja 10.1.1)

$\Leftrightarrow [T(\mathbf{v}_0)]_B = [\lambda\mathbf{v}_0]_B$ i $[\mathbf{v}_0]_B \neq \mathbf{0}$ (zob. dowód tw. 7.6.2)

$\Leftrightarrow [T]_B[\mathbf{v}_0]_B = \lambda[\mathbf{v}_0]_B$ i $[\mathbf{v}_0]_B \neq \mathbf{0}$ (zob. tw. 8.5.1 i 7.6.1)

$\Leftrightarrow [\mathbf{v}_0]_B$ jest wektorem własnym macierzy $[T]_B$ odpowiadającym wartości własnej λ. $\square$

Przykład 229. Przekształcenie liniowe $T\colon R_2[x] \to R_2[x]$ określone jest wzorem $T(a_0 + a_1x + a_2x^2) = (-14a_0 + 4a_1 - 14a_2) + (-33a_0 + 9a_1 - 31a_2)x + (11a_0 - 4a_1 + 11a_2)x^2$. Wyznaczyć jego wartości własne i wektory własne.

Macierzą przekształcenia T względem bazy $B = (1, x, x^2)$ jest

$$[T]_B = \begin{bmatrix} -14 & 4 & -14 \\ -33 & 9 & -31 \\ 11 & -4 & 11 \end{bmatrix},$$

a dla jej wielomianu charakterystycznego mamy

$$\begin{aligned} \begin{vmatrix} -14-\lambda & 4 & -14 \\ -33 & 9-\lambda & -31 \\ 11 & -4 & 11-\lambda \end{vmatrix} &= \begin{vmatrix} -\lambda & 4 & -14 \\ -2 & 9-\lambda & -31 \\ \lambda & -4 & 11-\lambda \end{vmatrix} = \begin{vmatrix} -\lambda & 4 & -14 \\ -2 & 9-\lambda & -31 \\ 0 & 0 & -3-\lambda \end{vmatrix} \\ &= -(3+\lambda)\begin{vmatrix} -\lambda & 4 \\ -2 & 9-\lambda \end{vmatrix} = -(\lambda+3)(\lambda^2 - 9\lambda + 8) \\ &= -(\lambda+3)(\lambda-1)(\lambda-8). \end{aligned}$$

Stąd wynika, że liczby $\lambda_1 = -3$, $\lambda_2 = 1$ i $\lambda_3 = 8$ są wartościami własnymi operatora T (i macierzy $[T]_B$). Rozwiązując równania $([T]_B - \lambda_i \mathbf{I}_3)\mathbf{x} = \mathbf{0}$, stwierdzamy, że wektory

$$[\mathbf{v}_1]_B = \begin{bmatrix} 4 \\ 11 \\ 0 \end{bmatrix}, \quad [\mathbf{v}_2]_B = \begin{bmatrix} 4 \\ 1 \\ -4 \end{bmatrix} \quad \text{i} \quad [\mathbf{v}_3]_B = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

są wektorami własnymi macierzy $[T]_B$. Stąd i z ostatniego twierdzenia wynika, że wektory

$$\mathbf{v}_1 = 4 + 11x, \quad \mathbf{v}_2 = 4 + x - 4x^2 \quad \text{i} \quad \mathbf{v}_3 = 1 + 2x - x^2$$

(których wektorami B-współrzędnych są $\mathbf{v}_1$, $\mathbf{v}_2$ i $\mathbf{v}_3$) są wektorami własnymi operatora T odpowiadającymi wartościom własnym $\lambda_1 = -3$, $\lambda_2 = 1$ i $\lambda_3 = 8$.

10.2. Diagonalizowalność macierzy i operatora liniowego

Definicja 10.2.1. Operator liniowy T na skończenie wymiarowej przestrzeni wektorowej V nazywamy *diagonalizowalnym*, jeśli istnieje baza B przestrzeni V taka, że macierz $[T]_B$ jest diagonalna. Macierz kwadratową $\mathbf{A}$ nazywamy *diagonalizowalną*, jeśli jest ona podobna do macierzy diagonalnej. Tak jest wtedy i tylko wtedy, gdy istnieje macierz nieosobliwa $\mathbf{P}$ taka, że macierz $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ jest macierzą diagonalną.

Diagonalizowalność operatora i macierzy

Przykład 230. Operator liniowy $T(x_1, x_2) = (4x_1 + 2x_2, -3x_1 + 11x_2)$ na przestrzeni R^2 jest diagonalizowalny, bo jego macierz $[T]_B$ jest diagonalna dla bazy $B = (\mathbf{b}_1, \mathbf{b}_2)$ przestrzeni R^2, gdzie $\mathbf{b}_1 = (2, 1)$ i$\mathbf{b}_2 = (1, 3)$,

$$\begin{aligned} [T]_B &= \Big[[T(\mathbf{b}_1)]_B \; [T(\mathbf{b}_2)]_B \Big] = \Big[[T(2,1)]_B \, [T(1,3)]_B \Big] \\ &= \Big[[(10,5)]_B \, [(10,30)]_B \Big] = \begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}. \end{aligned}$$

Zauważmy, że także macierz $\mathbf{A} = \begin{bmatrix} 4 & 2 \\ -3 & 11 \end{bmatrix}$ (która jest macierzą operatora T względem bazy standardowej przestrzeni R^2) jest diagonalizowalna i macierz $\mathbf{P} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$ jest macierzą podobieństwa macierzy $\mathbf{A}$ do macierzy diagonalnej, bo mamy

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \frac{1}{5}\begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 4 & 2 \\ -3 & 11 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}.$$

Mamy następujący związek diagonalizowalności operatora z diagonalizowalnością macierzy tego operatora.

Twierdzenie 10.2.1. *Jeśli T jest operatorem liniowym na skończenie wymiarowej przestrzeni wektorowej V i B jest bazą przestrzeni V, to operator T jest diagonalizowalny wtedy i tylko wtedy, gdy macierz $[T]_B$ jest diagonalizowalna.*

T – diagonalizowalny $\Leftrightarrow$ $[T]_B$ – diagonalizowalna

Dowód. Załóżmy, że operator T jest diagonalizowalny. Wtedy istnieje baza B' przestrzeni V taka, że macierz $[T]_{B'}$ jest diagonalna. Stąd i z podobieństwa macierzy $[T]_B$ i $[T]_{B'}$ (zob. tw. 8.7.3) wynika diagonalizowalność macierzy $[T]_B$.

Załóżmy teraz, że macierz $[T]_B$ jest diagonalizowalna. Wtedy jest ona podobna do pewnej macierzy diagonalnej $\mathbf{D}$. Wobec twierdzenia 8.7.3 istnieje baza B' przestrzeni V taka, że $\mathbf{D} = [T]_{B'}$. Stąd wynika diagonalizowalność operatora T. $\square$

Wniosek 10.2.1. *Macierz* $\mathbf{A}$ *jest diagonalizowalna wtedy i tylko wtedy, gdy operator* $T_{\mathbf{A}}$ *jest diagonalizowalny.* $\square$

Ogólny warunek konieczny i dostateczny diagonalizowalności operatora liniowego (i macierzy) przedstawia następujace twierdzenie.

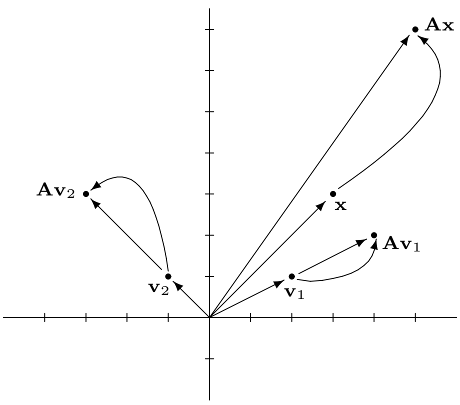
Twierdzenie 10.2.2. *Operator liniowy* T *na* n*-wymiarowej przestrzeni wektorowej* V *jest diagonalizowalny wtedy i tylko wtedy, gdy istnieje baza* $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ *przestrzeni* V *składająca się z wektorów własnych operatora* T*. Dodatkowo, jeśli wektory własne* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ *odpowiadają wartościom własnym* $\lambda_1, \lambda_2, \ldots, \lambda_n$*, to macierz* $[T]_B$ *jest diagonalna i*

$$[T]_B = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}.$$

Dowód. Operator T jest diagonalizowalny wtedy i tylko wtedy, gdy istnieje baza $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ przestrzeni V taka, że macierz $[T]_B = \big[[T(\mathbf{v}_1)]_B \ldots [T(\mathbf{v}_n)]_B\big]$ jest diagonalna. Tak jest wtedy i tylko wtedy, gdy istnieją skalary $\lambda_1, \ldots, \lambda_n$ takie, że

$$[T]_B = \begin{bmatrix} | & | & & | \\ [T(\mathbf{v}_1)]_B & [T(\mathbf{v}_2)]_B & \ldots & [T(\mathbf{v}_n)]_B \\ | & | & & | \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}.$$

Ostatnia równość zachodzi wtedy i tylko wtedy, gdy $T(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$ dla $i = 1, \ldots, n$, tj. wtedy i tylko wtedy, gdy każdy wektor $\mathbf{v}_i$ bazy B jest wektorem własnym operatora T odpowiadającym wartości własnej λ_i. $\square$



Rys. 10.2

Przykład 231. Weźmy pod uwagę macierze

$$\mathbf{A} = \frac{1}{3}\begin{bmatrix} 7 & -2 \\ -1 & 8 \end{bmatrix}, \quad \mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \text{i} \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Ponieważ

$$T_{\mathbf{A}}(\mathbf{v}_1) = \frac{1}{3}\begin{bmatrix} 7 & -2 \\ -1 & 8 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 2\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2\mathbf{v}_1,$$

więc $\mathbf{v}_1$ jest wektorem własnym operatora $T_{\mathbf{A}}$ (i macierzy $\mathbf{A}$) odpowiadającym wartości własnej $\lambda_1 = 2$. Podobnie mamy

$$T_{\mathbf{A}}(\mathbf{v}_2) = \frac{1}{3}\begin{bmatrix} 7 & -2 \\ -1 & 8 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ 3 \end{bmatrix} = 3\begin{bmatrix} -1 \\ 1 \end{bmatrix} = 3\mathbf{v}_2$$

i dlatego $\mathbf{v}_2$ jest wektorem własnym operatora $T_{\mathbf{A}}$ (i macierzy $\mathbf{A}$) odpowiadającym wartości własnej $\lambda_2 = 3$. Ponieważ układ $B = (\mathbf{v}_1, \mathbf{v}_2)$ jest bazą przestrzeni R^2, więc wobec twierdzenia 10.2.2 operator $T_{\mathbf{A}}$ jest diagonalizowalny i mamy

$$[T_{\mathbf{A}}]_B = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}.$$

Także macierz $\mathbf{A}$ jest diagonalizowalna, bo dla macierzy $\mathbf{P} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}$ (która jest macierzą przejścia od bazy B do bazy standardowej E) mamy

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \big[\, T_{\mathbf{A}} \,\big]_B = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}.$$

Efekt oddziaływania przekształcenia $T_{\mathbf{A}}$ (czyli mnożenia przez macierz $\mathbf{A}$) na wektory $\mathbf{v}_1$, $\mathbf{v}_2$ i $\mathbf{x} = (3,3)$ przedstawiono na rys. 10.2.

Następujące twierdzenie przedstawia warunek konieczny i dostateczny diagonalizowalności macierzy. Dla macierzy diagonalizowalnej $\mathbf{A}$, przedstawia ono także macierz podobieństwa $\mathbf{P}$ macierzy $\mathbf{A}$ do macierzy diagonalnej.

Twierdzenie 10.2.3. *Macierz $\mathbf{A} \in K_{n\times n}$ jest diagonalizowalna wtedy i tylko wtedy, gdy istnieje baza $B = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ przestrzeni K^n składająca się z wektorów własnych macierzy $\mathbf{A}$. Dodatkowo, jeśli wektory własne $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ odpowiadają wartościom własnym $\lambda_1, \lambda_2, \ldots, \lambda_n$, to macierz $\mathbf{A}$ jest podobna do macierzy diagonalnej $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ i $\mathbf{P} = \begin{bmatrix} \mathbf{v}_1 \ldots \mathbf{v}_n \end{bmatrix}$ jest macierzą podobienstwa macierzy $\mathbf{A}$ do macierzy $\mathbf{\Lambda}$, tj. $\mathbf{\Lambda} = \mathbf{P}^{-1}\mathbf{AP}$.*

Dowód. Przede wszystkim zauważmy, że dla macierzy $\mathbf{P} = \begin{bmatrix} \mathbf{v}_1 \ldots \mathbf{v}_n \end{bmatrix}$ oraz $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ mamy

$$\mathbf{AP} = \mathbf{A}\begin{bmatrix} \mathbf{v}_1 \ldots \mathbf{v}_n \end{bmatrix} = \begin{bmatrix} \mathbf{A}\mathbf{v}_1 \ldots \mathbf{A}\mathbf{v}_n \end{bmatrix} \tag{10.5}$$

i

$$\mathbf{P\Lambda} = \begin{bmatrix} \mathbf{v}_1 \ldots \mathbf{v}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{v}_1 \ldots \lambda_n\mathbf{v}_n \end{bmatrix}. \tag{10.6}$$

Załóżmy teraz, że macierz $\mathbf{A}$ jest diagonalizowalna i $\Lambda = \mathbf{P}^{-1}\mathbf{AP}$. Wtedy $\mathbf{AP} = \mathbf{P\Lambda}$, więc z (10.5) i (10.6) mamy

$$\begin{bmatrix} \mathbf{A}\mathbf{v}_1 \ldots \mathbf{A}\mathbf{v}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{v}_1 \ldots \lambda_n\mathbf{v}_n \end{bmatrix} \tag{10.7}$$

i dlatego także

$$\mathbf{A}\mathbf{v}_1 = \lambda_1\mathbf{v}_1, \ldots, \mathbf{A}\mathbf{v}_n = \lambda_n\mathbf{v}_n. \tag{10.8}$$

Ponieważ macierz $\mathbf{P}$ jest odwracalna, więc jej kolumny $\mathbf{v}_1, \ldots, \mathbf{v}_n$ są liniowo niezależne i dlatego układ $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ jest bazą przestrzeni K^n. Dodatkowo, ponieważ wektory $\mathbf{v}_1, \ldots, \mathbf{v}_n$ są niezerowe, z (10.8) wynika, że $\mathbf{v}_1, \ldots, \mathbf{v}_n$ są wektorami własnymi macierzy $\mathbf{A}$ odpowiadającymi wartościom własnym $\lambda_1, \ldots, \lambda_n$.

Załóżmy teraz, że $\mathbf{v}_1, \ldots, \mathbf{v}_n$ są wektorami własnymi odpowiadającymi wartościom własnym $\lambda_1, \ldots, \lambda_n$ macierzy $\mathbf{A}$. Wtedy wobec (10.5), (10.6) i (10.7) dla macierzy $\mathbf{P} = \begin{bmatrix} \mathbf{v}_1 \ldots \mathbf{v}_n \end{bmatrix}$ oraz $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ jest $\mathbf{P\Lambda} = \mathbf{AP}$. Przy założeniu liniowej niezależności wektorów $\mathbf{v}_1, \ldots, \mathbf{v}_n$ macierz $\mathbf{P}$ jest odwracalna i równość $\mathbf{P\Lambda} = \mathbf{AP}$ jest równoważna równości $\mathbf{\Lambda} = \mathbf{P}^{-1}\mathbf{AP}$, a ta jest równoważna diagonalizowalności macierzy $\mathbf{A}$. □

Twierdzenie 10.2.4. *Niech T będzie operatorem liniowym na przestrzeni wektorowej V. Jeśli $\mathbf{v}_1, \ldots, \mathbf{v}_k$ są wektorami własnymi operatora T odpowiadającymi jego różnym wartościom własnym $\lambda_1, \ldots, \lambda_k$, to układ wektorów $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ jest liniowo niezależny.*

Dowód. Ponieważ wektory $\mathbf{v}_1, \ldots, \mathbf{v}_k$ są niezerowe, więc wystarczy udowodnić, że żaden z nich nie jest kombinacją liniową swoich poprzedników. Przypuśćmy, że jest inaczej. Niech wtedy j $(2 \leqslant j \leqslant k)$ będzie najmniejszą liczbą naturalną taką, że $\mathbf{v}_j$ jest kombinacją liniową wektorów $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$. Wtedy wektory $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$ są liniowo niezależne i istnieją skalary $a_1, \ldots, a_{j-1}$ takie, że

$$\mathbf{v}_j = a_1\mathbf{v}_1 + \ldots + a_{j-1}\mathbf{v}_{j-1}.$$

Wtedy także

$$\lambda_j \sum_{l=1}^{j-1} a_l\mathbf{v}_l = \lambda_j\mathbf{v}_j = T(\mathbf{v}_j) = T\left(\sum_{l=1}^{j-1} a_l\mathbf{v}_l\right) = \sum_{l=1}^{j-1} a_l T(\mathbf{v}_l) = \sum_{l=1}^{j-1} a_l\lambda_l\mathbf{v}_l$$

i z niezależności wektorów $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$ wynika, że $\lambda_j a_l = \lambda_l a_l$ dla $l = 1, \ldots, j-1$. Stąd i z faktu, że $\lambda_1, \ldots, \lambda_k$ są różne wynika, że $a_l = 0$ dla $l = 1, \ldots, j-1$. Wtedy $\mathbf{v}_j = \sum_{l=1}^{j-1} a_l\mathbf{v}_l = \sum_{l=1}^{j-1} 0\mathbf{v}_l = \mathbf{0}$, co zaprzecza niezerowości wektorów $\mathbf{v}_1, \ldots, \mathbf{v}_k$. □

Wniosek 10.2.2. *Niech T będzie operatorem liniowym na n-wymiarowej przestrzeni wektorowej V. Jeśli operator T ma n różnych wartości własnych, to jest on diagonalizowalny.*

Dowód. Niech $\lambda_1, \ldots, \lambda_n$ będą różnymi wartościami własnymi operatora T. Niech $\mathbf{v}_i$ będzie wektorem własnym operatora T odpowiadającym wartości własnej λ_i ($i = 1, \ldots, n$). Wobec poprzedniego twierdzenia $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ jest układem n liniowo niezależnych wektorów w n-wymiarowej przestrzeni V. Stąd i z twierdzenia 7.5.4 wynika, że B jest bazą przestrzeni V. Ponieważ baza ta składa się z wektorów własnych operatora T, więc wobec twierdzenia 10.2.2 operator T jest diagonalizowalny. □

Wniosek 10.2.3. *Jeśli macierz $\mathbf{A} \in K_{n\times n}$ ma n różnych wartości własnych, to jest ona diagonalizowalna.* □

Przykład 232. Wielomianem charakterystycznym macierzy

$$\mathbf{A} = \begin{bmatrix} 7 & 8 \\ -4 & -5 \end{bmatrix} \in R_{2\times 2}$$

(i operatora $T_\mathbf{A}\colon\ R^2 \to R^2$) jest

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 7-\lambda & 8 \\ -4 & -5-\lambda \end{vmatrix} = (\lambda+1)(\lambda-3),$$

więc wartościami własnymi operatora $T_\mathbf{A}$ (i macierzy $\mathbf{A}$) są dwie różne liczby $\lambda = -1$ i $\lambda = 3$. Stąd i z wniosku 10.2.2 wynika, że operator $T_\mathbf{A}$ i macierz $\mathbf{A}$ są diagonalizowalne.

Warunek z ostatniego wniosku jest tylko warunkiem dostatecznym diagonalizowalności. Przykład operatora tożsamościowego (który jest diagonalizowalny choć ma tylko jedną wartość własną) pokazuje, że nie jest to warunek konieczny diagonalizowalności operatora. Udowodnimy teraz, że wielomian charakterystyczny diagonalizowalnego operatora T na przestrzeni wektorowej V, która jest przestrzenią wektorową nad ciałem K, jest rozkładalny nad ciałem K.

Wielomian charakterystyczny operatora diagonalizowalnego jest rozkładalny

Twierdzenie 10.2.5. *Niech T będzie operatorem liniowym na n-wymiarowej przestrzeni wektorowej V (nad ciałem K) i niech $\varphi(\lambda)$ będzie wielomianem charakterystycznym operatora T. Jeśli operator T jest diagonalizowalny, to istnieją skalary a, $\lambda_1, \ldots, \lambda_n \in K$ takie, że*

$$\varphi(\lambda) = a(\lambda-\lambda_1)(\lambda-\lambda_2)\ldots(\lambda-\lambda_n).$$

Dowód. Niech B będzie taką bazą przestrzeni V, że macierz $[T]_B$ jest diagonalizowalna. Wtedy istnieją skalary $\lambda_1, \ldots, \lambda_n$ takie, że

$$[T]_B = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}.$$

Wtedy także mamy

$$\begin{aligned} \varphi(\lambda) &= \det([T]_B - \lambda\mathbf{I}) = \det \begin{bmatrix} \lambda_1-\lambda & 0 & \ldots & 0 \\ 0 & \lambda_2-\lambda & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n-\lambda \end{bmatrix} \\ &= (\lambda_1-\lambda)(\lambda_2-\lambda)\ldots(\lambda_n-\lambda) = (-1)^n(\lambda-\lambda_1)(\lambda-\lambda_2)\ldots(\lambda-\lambda_n) \end{aligned}$$

i to dowodzi tezę twierdzenia. □

Przykład 233. Dana jest macierz $\mathbf{A} = \begin{bmatrix} 1 & 4 \\ -1 & 1 \end{bmatrix}$ i określony przez nią operator liniowy $T_{\mathbf{A}}\colon\ R^2 \to R^2$. Ponieważ jego wielomian charakterystyczny

$$\varphi(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \det \begin{bmatrix} 1-\lambda & 4 \\ -1 & 1-\lambda \end{bmatrix} = (1-\lambda)^2 + 4$$

nie jest rozkładalny nad ciałem R, więc operator $T_{\mathbf{A}}\colon\ R^2 \to R^2$ nie jest diagonalizowalny. Ta sama macierz $\mathbf{A}$ określa także operator liniowy $U_{\mathbf{A}}\colon\ \mathcal{C}^2 \to \mathcal{C}^2$, gdzie $U_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ dla każdego $\mathbf{x} \in \mathcal{C}^2$. Tym razem jego wielomian charakterystyczny $\varphi(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = (1-\lambda)^2 + 4 = (\lambda - 1 - 2j)(\lambda - 1 + 2j)$ (jest rozkładalny nad ciałem $\mathcal{C}$ i) ma dwie różne wartości własne i dlatego operator $U_{\mathbf{A}}$ jest diagonalizowalny.

W naszych dalszych rozważaniach diagonalizowalności operatorów posłużymy się algebraiczną i geometryczną krotnością wartości własnej.

Definicja 10.2.2. Niech $\varphi(\lambda)$ i λ_0 będą odpowiednio wielomianem charakterystycznym i wartością własną operatora liniowego (lub macierzy kwadratowej). *Algebraiczną krotnością* liczby λ_0 nazywamy największą liczbę naturalną $k = k(\lambda_0)$ taką, że $(\lambda - \lambda_0)^k$ jest dzielnikiem wielomianu $\varphi(\lambda)$.

$k(\lambda_0)$ – algebraiczna krotność wartości własnej

Przykład 234. Wielomianem charakterystycznym macierzy

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

jest $\varphi(\lambda) = -(\lambda-2)^2(\lambda-3)$. Zatem $\lambda = 2$ jest dwukrotną, a $\lambda = 3$ – jednokrotną wartością własną macierzy $\mathbf{A}$ i operatora $T_{\mathbf{A}}$ na przestrzeni R^3.

Definicja 10.2.3. Niech T będzie operatorem liniowym na przestrzeni wektorowej V i niech λ będzie jego wartością własną. Wtedy zbiór

$$V_\lambda = \{\mathbf{x} \in V\colon\ T(\mathbf{x}) = \lambda\mathbf{x}\} = \operatorname{Ker}(T - \lambda I_V)$$

V_λ – przestrzeń własna

który jest podprzestrzenią przestrzeni V, nazywamy *przestrzenią własną* operatora T odpowiadającą wartości własnej λ. Analogicznie definiuje się przestrzeń własną macierzy. Wymiar przestrzeni V_λ, czyli liczbę $\dim(V_\lambda)$, nazywamy *geometryczną krotnością* wartości własnej λ operatora T.

$\dim(V_\lambda)$ – geometryczna krotność wartości własnej

Związek między algebraiczną i geometryczną krotnością wartości własnej operatora przedstawia następujące twierdzenie.

Twierdzenie 10.2.6. *Jeśli λ_0 jest wartością własną operatora liniowego T na skończenie wymiarowej przestrzeni wektorowej V, to*

$$1 \leqslant \dim(V_{\lambda_0}) \leqslant k(\lambda_0).$$

Geometryczna krotność $\leqslant$ algebraiczna krotność

Dowód. Weźmy pod uwagę bazę $(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ przestrzeni V_{λ_0} (gdzie $m = \dim(V_{\lambda_0})$) i jej rozszerzenie $B = (\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_n)$ do bazy całej przestrzeni V. Ponieważ $\mathbf{v}_1, \ldots, \mathbf{v}_m$ są wektorami własnymi operatora T odpowiadającymi wartości własnej λ_0, więc macierzą operatora T względem bazy B jest macierz blokowa

$$[T]_B = \left[\begin{array}{c|c} \lambda_0\mathbf{I}_m & \mathbf{A} \\ \hline \mathbf{0} & \mathbf{B} \end{array}\right],$$

w której $\mathbf{A}$ i $\mathbf{B}$ są macierzami wymiaru odpowiednio $m \times (n-m)$ i $(n-m) \times (n-m)$. Zatem dla wielomianu charakterystycznego operatora T mamy równości

$$\begin{aligned}
\varphi(\lambda) &= \det\Big([T]_B - \lambda \mathbf{I}_n\Big) = \det \left[\begin{array}{c|c} (\lambda_0-\lambda)\mathbf{I}_m & \mathbf{A} \\ \hline \mathbf{0} & \mathbf{B}-\lambda\mathbf{I}_{n-m} \end{array}\right] \\
&= \det\Big((\lambda_0-\lambda)\mathbf{I}_m\Big) \det(\mathbf{B}-\lambda\mathbf{I}_{n-m}) \\
&= (\lambda_0-\lambda)^m \det(\mathbf{B}-\lambda\mathbf{I}_{n-m})
\end{aligned}$$

i z nich wynika, że $(\lambda-\lambda_0)^m$ jest dzielnikiem wielomianu $\varphi(\lambda)$. Stąd i z definicji liczby $k(\lambda_0)$ wynika, $k(\lambda_0) \geqslant m = \dim(V_{\lambda_0})$. □

Z definicji wartości własnej i przestrzeni własnej operatora jest oczywiste, że jeśli λ' i λ'' są różnymi wartościami własnymi operatora liniowego $T\colon V \to V$, to $V_{\lambda'} \cap V_{\lambda''} = \{\mathbf{0}\}$. Stąd zaś wynika, że jeśli $\lambda_1, \ldots, \lambda_m$ są różnymi wartościami własnymi operatora liniowego $T\colon V \to V$, to podprzestrzeń $V_{\lambda_1} + \ldots + V_{\lambda_m}$ jest sumą prostą podprzestrzeni $V_{\lambda_1}, \ldots, V_{\lambda_m}$, czyli $V_{\lambda_1} + V_{\lambda_2} + \ldots + V_{\lambda_m} = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \ldots \oplus V_{\lambda_m}$.

Twierdzenie 10.2.7. *Niech T będzie operatorem liniowym na n-wymiarowej przestrzeni wektorowej V. Jeśli wielomian charakterystyczny operatora T jest rozkładalny i $\lambda_1, \lambda_2, \ldots, \lambda_l$ są wszystkimi różnymi wartościami własnymi operatora T, to T jest diagonalizowalny wtedy i tylko wtedy, gdy algebraiczna krotność każdej wartości własnej λ_i pokrywa się z jej geometryczną krotnością, $k(\lambda_i) = \dim(V_{\lambda_i})$ dla $i = 1, \ldots, l$.*

Dodatkowo, jeśli operator T jest diagonalizowalny i $B_i = (\mathbf{b}_1^{(i)}, \ldots, \mathbf{b}_{k(\lambda_i)}^{(i)})$ jest bazą przestrzeni V_{λ_i} $(i = 1, \ldots, l)$, to

$$B = \Big(\mathbf{b}_1^{(1)}, \ldots, \mathbf{b}_{k(\lambda_1)}^{(1)}, \mathbf{b}_1^{(2)}, \ldots, \mathbf{b}_{k(\lambda_2)}^{(2)}, \ldots, \mathbf{b}_1^{(l)}, \ldots, \mathbf{b}_{k(\lambda_l)}^{(l)}\Big)$$

jest bazą przestrzeni V składającą się z wektorów własnych operatora T.

Dowód. Z założeń wynika, że wielomianem charakterystycznym operatora T jest $\varphi(\lambda) = (-1)^n(\lambda-\lambda_1)^{k(\lambda_1)} \ldots (\lambda-\lambda_l)^{k(\lambda_l)}$ i $\sum_{i=1}^{l} k(\lambda_i) = n$.

Załóżmy teraz, że operator T jest diagonalizowalny. Wtedy istnieje baza B przestrzeni V składająca się z samych wektorów własnych operatora T. Ponieważ $B \cap V_{\lambda_i}$ jest zbiorem liniowo niezależnych wektorów własnych należących do przestrzeni V_{λ_i} wymiaru $\dim(V_{\lambda_i})$ i $\dim(V_{\lambda_i}) \leqslant k(\lambda_i)$ (zob. tw. 10.2.6), więc

$$n = |B| = |\bigcup_{i=1}^{l} (B \cap V_{\lambda_i})| = \sum_{i=1}^{l} |B \cap V_{\lambda_i}| \leqslant \sum_{i=1}^{l} \dim(V_{\lambda_i}) \leqslant \sum_{i=1}^{l} k(\lambda_i) = n$$

i stąd wynika, że $k(\lambda_i) = \dim(V_{\lambda_i})$ dla $i = 1, \ldots, l$.

Dla dowodu przeciwnej implikacji i jednoczesnego dowodu drugiej części twierdzenia załóżmy, że $k(\lambda_i) = \dim(V_{\lambda_i})$ i niech $B_i = (\mathbf{b}_1^{(i)}, \ldots, \mathbf{b}_{k(\lambda_i)}^{(i)})$ będzie bazą przestrzeni V_{λ_i} $(i = 1, \ldots, l)$. Twierdzimy, że układ $B = (\mathbf{b}_1^{(1)}, \ldots, \mathbf{b}_{k(\lambda_1)}^{(1)}, \ldots, \mathbf{b}_1^{(l)}, \ldots, \mathbf{b}_{k(\lambda_l)}^{(l)})$ jest liniowo niezależny. Dla dowodu tego faktu niech $x_j^{(i)}$ będą skalarami takimi, że $\sum_{i=1}^{l} \sum_{j=1}^{k(\lambda_i)} x_j^{(i)} \mathbf{b}_j^{(i)} = \mathbf{0}$. Wtedy dla każdego $i_0 \in \{1, \ldots, l\}$ jest

$$\sum_{j=1}^{k(\lambda_{i_0})} x_j^{(i_0)} \mathbf{b}_j^{(i_0)} = -\sum_{i \neq i_0} \sum_{j=1}^{k(\lambda_i)} x_j^{(i)} \mathbf{b}_j^{(i)}.$$

Ponieważ $\sum_{j=1}^{k(\lambda_{i_0})} x_j^{(i_0)} \mathbf{b}_j^{(i_0)} \in V_{\lambda_0}$ i $-\sum_{i \neq i_0} \sum_{j=1}^{k(\lambda_i)} x_j^{(i)} \mathbf{b}_j^{(i)} \in \sum_{i \neq i_0} V_{\lambda_i}$, a wektor zerowy jest jedynym wspólnym elementem podprzestrzeni V_{λ_0} i $\sum_{i \neq i_0} V_{\lambda_i}$, więc $\sum_{j=1}^{k(\lambda_{i_0})} x_j^{(i_0)} \mathbf{b}_j^{(i_0)} = \mathbf{0}$. Stąd i z liniowej niezależności wektorów układu B_{i_0} wynika, że $x_1^{(i_0)} = x_2^{(i_0)} = \ldots = x_{k(\lambda_{i_0})}^{(i_0)} = 0$ (dla każdego $i_0 \in \{1, \ldots, l\}$). To dowodzi, że układ B jest liniowo niezależny. Dodatkowo, ponieważ układ B zawiera $\sum_{i=1}^{l} k(\lambda_i) = n = \dim(V)$ wektorów, jest on bazą przestrzeni V. Baza ta składa się z samych wektorów własnych operatora T. Zatem wobec twierdzenia 10.2.2 operator T jest diagonalizowalny. □

Przykład 235. Niech T będzie operatorem liniowym na przestrzeni wektorowej R^3 określonym wzorem $T(x,y,z) = (2x-z, 3y, -x+2z)$. Wyznaczyć przestrzenie własne operatora T i zbadać diagonalizowalność operatora T.

Macierzą operatora T względem bazy standardowej E przestrzeni R^3 jest

$$[T]_E = \begin{bmatrix} 2 & 0 & -1 \\ 0 & 3 & 0 \\ -1 & 0 & 2 \end{bmatrix},$$

więc jego wielomianem charakterystycznym jest

$$\det([T]_E - \lambda\mathbf{I}) = \det \begin{bmatrix} 2-\lambda & 0 & -1 \\ 0 & 3-\lambda & 0 \\ -1 & 0 & 2-\lambda \end{bmatrix} = -(\lambda-1)(\lambda-3)^2.$$

Zatem $\lambda_1 = 1$ jest jednokrotną, a $\lambda_2 = 3$ – dwukrotną wartością własną operatora T. Przestrzenią własną operatora T odpowiadającą wartości własnej $\lambda_1 = 1$ jest

$$\begin{aligned} V_{\lambda_1} &= \operatorname{Ker}(T - \lambda_1 1_{R^3}) \\ &= \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in R^3 \colon \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\} = \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right). \end{aligned}$$

Wartości własnej $\lambda_2 = 3$ odpowiada przestrzeń własna

$$\begin{aligned} V_{\lambda_2} = \operatorname{Ker}(T - \lambda_2 1_{R^3}) &= \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in R^3 \colon \begin{bmatrix} -1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\} \\ &= \mathcal{L}\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right). \end{aligned}$$

W tym przypadku algebraiczna krotność każdej wartości własnej λ_i jest równa wymiarowi odpowiadającej jej przestrzeni własnej V_{λ_i}, $k(\lambda_1) = 1 = \dim(V_{\lambda_1})$ i $k(\lambda_2) = 2 = \dim(V_{\lambda_2})$, więc układ

$$B = \left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right),$$

który jest sumą baz przestrzeni V_{λ_1} i V_{λ_2}, jest bazą przestrzeni R^3 składającą się z samych wektorów własnych operatora T. Stąd i z poprzedniego twierdzenia wynika, że operator T jest diagonalizowalny.

Przykład 236. Zbadać diagonalizowalność macierzy

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

Liczba $\lambda_1 = 1$ jest dwukrotną, a liczba $\lambda_2 = 5$ – jednokrotną wartością własną macierzy $\mathbf{A}$ (i operatora $T_\mathbf{A} \colon R^3 \to R^3$), bo $\det(\mathbf{A} - \lambda\mathbf{I}) = (1-\lambda)^2(5-\lambda)$. Przestrzenią własną macierzy $\mathbf{A}$ odpowiadającą wartości własnej $\lambda_1 = 1$ jest

$$\begin{aligned} V_{\lambda_1} &= \left\{ \mathbf{x} \in R^3 \colon (\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x} = \mathbf{0} \right\} \\ &= \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in R^3 \colon \begin{bmatrix} 0 & 0 & 2 \\ 3 & 0 & 4 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\} = \mathcal{L}\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right). \end{aligned}$$

Ponieważ wymiar przestrzeni V_{λ_1} jest mniejszy od algebraicznej krotności liczby λ_1, bo mamy $\dim(V_{\lambda_1}) = 1 < 2 = k(\lambda_1)$, więc macierz $\mathbf{A}$ nie jest diagonalizowalna. (Wymiar przestrzeni V_{λ_1} także można było wyznaczyć za pomocą rzędu macierzy $\mathbf{A} - \lambda_i\mathbf{I}$, bo zawsze jest $\dim(V_{\lambda_i}) = n - \mathrm{r}(\mathbf{A} - \lambda_i\mathbf{I})$, gdzie $n = \dim(V)$.)

10.3. Diagonalizacja macierzy symetrycznej

Problem diagonalizacji macierzy symetrycznych jest ważny zarówno ze względów praktycznych, jak i teoretycznych. Tu udowodnimy, że każda rzeczywista macierz symetryczna jest diagonalizowalna. Wykażemy także, że tej diagonalizacji można dokonać za pomocą macierzy ortogonalnej. Nasze rozważania zaczynamy od dowodu, że wartości własne rzeczywistej macierzy symetrycznej są rzeczywiste.

$\mathbf{A}$ – symetryczna $\Leftrightarrow \mathbf{A}^T = \mathbf{A}$

Twierdzenie 10.3.1. *Każda wartość własna rzeczywistej macierzy symetrycznej jest liczbą rzeczywistą.*

Dowód. Niech $\mathbf{A}$ będzie symetryczną macierzą rzeczywistą wymiaru $n \times n$. Niech λ będzie wartością własną macierzy $\mathbf{A}$ i niech $\mathbf{v} \in \mathcal{C}^n$ będzie wektorem własnym takim, że $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Dla dowodu twierdzenia wystarczy wykazać, że $\overline{\lambda} = \lambda$.

Zwróćmy uwagę, że $\overline{\mathbf{A}} = \overline{[a_{ij}]} = [\overline{a_{ij}}]$

Ponieważ $\mathbf{A} = \overline{\mathbf{A}} = \overline{\mathbf{A}}^T$, więc mamy

$$\begin{aligned}\overline{\lambda}\big(\overline{\mathbf{v}}^T\mathbf{v}\big) &= \big(\overline{\lambda\mathbf{v}}\big)^T\mathbf{v} = \big(\overline{\mathbf{A}\mathbf{v}}\big)^T\mathbf{v} = \big(\overline{\mathbf{v}}^T\overline{\mathbf{A}}^T\big)\mathbf{v} = \big(\overline{\mathbf{v}}^T\mathbf{A}\big)\mathbf{v} \\ &= \overline{\mathbf{v}}^T\big(\mathbf{A}\mathbf{v}\big) = \overline{\mathbf{v}}^T\big(\lambda\mathbf{v}\big) = \lambda\big(\overline{\mathbf{v}}^T\mathbf{v}\big).\end{aligned}$$

Stąd $\overline{\lambda} = \lambda$ i to dowodzi, że $\lambda \in R$. □

Warto zwrócić uwagę na to, że wartościami własnymi niesymetrycznej macierzy rzeczywistej mogą być zarówno liczby rzeczywiste jak i zespolone. Przykładowo, wartościami własnymi niesymetrycznej macierzy rzeczywistej

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

są liczby $\lambda_1 = 1$, $\lambda_2 = j$ i $\lambda_3 = -j$.

Wiemy, że wektory własne odpowiadające różnym wartościom własnym macierzy są liniowo niezależne (zob. tw. 10.2.4). Pokażemy teraz, że w przypadku macierzy symetrycznych mają one dodatkową własność – są wzajemnie ortogonalne.

Twierdzenie 10.3.2. *Wektory własne odpowiadające różnym wartościom własnym rzeczywistej macierzy symetrycznej są ortogonalne.*

Dowód. Niech $\mathbf{v}_1$ i $\mathbf{v}_2$ będą wektorami własnymi odpowiadającymi różnym wartościom własnym λ_1 i λ_2 symetrycznej macierzy $\mathbf{A}$. Dla dowodu twierdzenia wystarczy pokazać, że $(\mathbf{v}_1|\mathbf{v}_2) = 0$.

Ponieważ $\mathbf{A}\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ i $\mathbf{A}\mathbf{v}_2 = \lambda_2\mathbf{v}_2$, więc wobec symetrii macierzy $\mathbf{A}$ oraz własności iloczynu skalarnego mamy

$(\mathbf{x}|\mathbf{y}) = \mathbf{x}^T\mathbf{y}$

$$\begin{aligned}\lambda_1(\mathbf{v}_1|\mathbf{v}_2) &= (\lambda_1\mathbf{v}_1|\mathbf{v}_2) = (\mathbf{A}\mathbf{v}_1|\mathbf{v}_2) = (\mathbf{A}\mathbf{v}_1)^T\mathbf{v}_2 = (\mathbf{v}_1^T\mathbf{A}^T)\mathbf{v}_2 \\ &= (\mathbf{v}_1^T\mathbf{A})\mathbf{v}_2 = \mathbf{v}_1^T(\mathbf{A}\mathbf{v}_2) = \mathbf{v}_1^T(\lambda_2\mathbf{v}_2) = \lambda_2(\mathbf{v}_1^T\mathbf{v}_2) = \lambda_2(\mathbf{v}_1|\mathbf{v}_2).\end{aligned}$$

Stąd $(\lambda_1 - \lambda_2)(\mathbf{v}_1|\mathbf{v}_2) = 0$ i dlatego $(\mathbf{v}_1|\mathbf{v}_2) = 0$, bo $\lambda_1 - \lambda_2 \neq 0$. □

$\lambda_i \neq \lambda_j \Rightarrow V_{\lambda_i} \perp V_{\lambda_j}$

Wniosek 10.3.1. *Jeśli $\mathbf{A}$ jest symetryczną macierzą rzeczywistą i λ_i oraz λ_j są różnymi wartościami własnymi macierzy $\mathbf{A}$, to podprzestrzenie własne V_{λ_i} oraz V_{λ_j} są wzajemnie ortogonalne.* □

Twierdzenie 10.3.3. *Jeśli* $\mathbf{Q}$ *i* $\mathbf{A}$ *są rzeczywistymi macierzami wymiaru* $n \times n$ *i macierz* $\mathbf{Q}$ *jest ortogonalna, to macierz* $\mathbf{A}$ *jest symetryczna wtedy i tylko wtedy, gdy macierz* $\mathbf{Q}^T\mathbf{AQ}$ *jest symetryczna.*

$\mathbf{Q}$ – ortogonalna $\Leftrightarrow \mathbf{Q}^{-1} = \mathbf{Q}^T$

Dowód. Ponieważ $\mathbf{Q}^T = \mathbf{Q}^{-1}$, więc macierze

$$\mathbf{Q}^T\mathbf{AQ} \quad \text{i} \quad (\mathbf{Q}^T\mathbf{AQ})^T = \mathbf{Q}^T\mathbf{A}^T\big(\mathbf{Q}^T\big)^T = \mathbf{Q}^T\mathbf{A}^T\mathbf{Q}$$

są identyczne wtedy i tylko wtedy, gdy macierze $\mathbf{A}$ i $\mathbf{A}^T$ są identyczne. □

Definicja 10.3.1. Mówimy, że macierz $\mathbf{A} \in R_{n\times n}$ jest *ortogonalnie diagonalizowalna*, gdy istnieje macierz ortogonalna $\mathbf{Q}$ taka, że macierz $\mathbf{Q}^T\mathbf{AQ}$ jest diagonalna.

Ortogonalna diagonalizowalność macierzy

Z twierdzenia 10.2.2 jest oczywiste, że macierz $\mathbf{A} \in R_{n\times n}$ jest ortogonalnie diagonalizowalna wtedy i tylko wtedy, gdy istnieje baza $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ przestrzeni R^n składająca się z ortonormalnych wektorów własnych macierzy $\mathbf{A}$. W takim przypadku macierz $\mathbf{Q} = [\,\mathbf{v}_1 \ldots \mathbf{v}_n\,]$ jest ortogonalna i iloczyn $\mathbf{Q}^T\mathbf{AQ}$ jest macierzą diagonalną. Okazuje się, że klasa macierzy ortogonalnie diagonalizowalnych jest łatwo rozpoznawalna i identyczna z klasą macierzy symetrycznych.

Twierdzenie 10.3.4. *Rzeczywista macierz* $\mathbf{A}$ *wymiaru* $n \times n$ *jest ortogonalnie diagonalizowalna wtedy i tylko wtedy, gdy macierz* $\mathbf{A}$ *jest symetryczna.*

Dowód. Jeśli dla macierzy $\mathbf{A}$ istnieje macierz ortogonalna $\mathbf{Q}$ taka, że iloczyn $\mathbf{Q}^T\mathbf{AQ}$ jest macierzą diagonalną (więc także symetryczną), to wobec poprzedniego twierdzenia macierz $\mathbf{A}$ jest symetryczna.

Przeciwną implikację dowodzimy indukcyjnie ze względu na n. Implikacja ta jest oczywista dla $n = 1$. Załóżmy teraz, że każda symetryczna macierz wymiaru $l \times l$, $1 \leqslant l < n$, jest ortogonalnie diagonalizowalna i niech $\mathbf{A}$ będzie symetryczną macierzą wymiaru $n\times n$. Niech $\lambda_1, \ldots, \lambda_k$ będą wszystkimi różnymi wartościami własnymi macierzy $\mathbf{A}$ i niech $\mathbf{v}_1, \ldots, \mathbf{v}_k$ będą odpowiadającymi im jednostkowymi (i wobec twierdzenia 10.2.4 wzajemnie ortogonalnymi) wektorami własnymi.

Jeśli $k = n$, to układ $B = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ jest bazą ortonormalną przestrzeni R^n, macierz $\mathbf{Q} = [\,\mathbf{v}_1 \ldots \mathbf{v}_n\,]$ jest ortogonalna, a macierz $\mathbf{Q}^T\mathbf{AQ} = [T_\mathbf{A}]_B$ (która jest macierzą przekształcenia $T_\mathbf{A}$ względem bazy B) jest diagonalna, $[T_\mathbf{A}]_B = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.

$T_\mathbf{A}(\mathbf{x}) = \mathbf{Ax}$

Załóżmy teraz, że $k < n$. W tym przypadku niech $B = (\mathbf{v}_1, \ldots, \mathbf{v}_k, \ldots, \mathbf{v}_n)$ będzie rozszerzeniem ortonormalnego układu $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ do ortonormalnej bazy przestrzeni R^n. Łatwo zauważyć, że macierzą przekształcenia $T_\mathbf{A}$ względem bazy B jest

$$[T_\mathbf{A}]_B = \left[\begin{array}{c|c} \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_k \end{bmatrix} & \mathbf{B} \\ \hline \mathbf{0} & \mathbf{C} \end{array}\right] = \mathbf{Q}^T\mathbf{AQ},$$

gdzie $\mathbf{Q} = \big[\mathbf{v}_1 \ldots \mathbf{v}_n\big]$ znowu jest macierzą ortogonalną. Z symetrii macierzy $\mathbf{A}$ i z twierdzenia 10.3.3 wynika symetryczność macierzy $\mathbf{Q}^T\mathbf{AQ}$. To zaś wymusza zerowość macierzy $\mathbf{B}$ i symetrię macierzy $\mathbf{C} \in R_{(n-k)\times(n-k)}$. Z tego ostatniego i z założenia indukcyjnego wynika ortogonalna diagonalizowalność macierzy $\mathbf{C}$. Zatem istnieje macierz ortogonalna $\mathbf{P} \in R_{(n-k)\times(n-k)}$ taka, że macierz $\mathbf{P}^T\mathbf{CP}$, oznaczamy ją przez $\mathbf{D}$, jest diagonalna. Łatwo teraz zauważyć, że macierz $\begin{bmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{P} \end{bmatrix}$ jest ortogonalna i jest ona macierzą podobieństwa macierzy $\mathbf{Q}^T\mathbf{AQ}$ do macierzy diagonalnej $\begin{bmatrix} \mathbf{\Lambda} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}$, gdzie $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_k)$. Równoważnie, macierz ortogonalna $\mathbf{Q}\begin{bmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{P} \end{bmatrix}$ jest macierzą podobieństwa macierzy symetrycznej $\mathbf{A}$ do macierzy diagonalnej $\begin{bmatrix} \mathbf{\Lambda} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}$. □

Przykład 237. Wyznaczyć ortogonalną macierz podobieństwa macierzy $\mathbf{A}$ do macierzy diagonalnej, gdy

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

Wartościami własnymi macierzy $\mathbf{A}$ są pierwiastki wielomianu

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

więc są nimi liczby $\lambda_1 = -1$ i $\lambda_2 = 3$. Łatwo sprawdzić, że liczbom tym odpowiadają wektory własne

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{i} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Ponieważ wektory $\mathbf{x}_1/\|\mathbf{x}_1\|$ i $\mathbf{x}_2/\|\mathbf{x}_2\|$ są ortonormalne, więc macierz

$$\mathbf{Q} = \begin{bmatrix} \dfrac{\mathbf{x}_1}{\|\mathbf{x}_1\|} & \dfrac{\mathbf{x}_2}{\|\mathbf{x}_2\|} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

jest ortogonalną macierzą podobieństwa macierzy $\mathbf{A}$ do macierzy diagonalnej,

$$\mathbf{Q}^T\mathbf{A}\mathbf{Q} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}.$$

Przykład 238. Dla macierzy symetrycznej $\mathbf{A} = \begin{bmatrix} 5 & 2 & 2 \\ 2 & 2 & -4 \\ 2 & -4 & 2 \end{bmatrix}$ wyznaczyć macierz ortogonalną $\mathbf{Q}$ taką, że $\mathbf{Q}^T\mathbf{A}\mathbf{Q}$ jest macierzą diagonalną.

Ortogonalną macierzą diagonalizującą macierz $\mathbf{A}$ będzie macierz $\mathbf{Q} = [\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3]$, w której kolumny $\mathbf{v}_1$, $\mathbf{v}_2$ i $\mathbf{v}_3$ są ortonormalnymi wektorami własnymi macierzy $\mathbf{A}$. Dla ich uzyskania wyznaczamy najpierw wartości własne macierzy $\mathbf{A}$. Ponieważ

$$\begin{aligned} \det(\mathbf{A} - \lambda\mathbf{I}) &= -\Big(\lambda^3 - \lambda^2 \operatorname{tr}(\mathbf{A}) + \lambda(|\mathbf{A}_{11}| + |\mathbf{A}_{22}| + |\mathbf{A}_{33}|) - |\mathbf{A}|\Big) \\ &= -(\lambda^3 - 9\lambda^2 - 108) = -(\lambda+3)(\lambda-6)^2, \end{aligned}$$

więc wartościami własnymi macierzy $\mathbf{A}$ są $\lambda_1 = -3$ i $\lambda_2 = \lambda_3 = 6$.

Rozwiązując teraz układ równań $(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{x} = \mathbf{0}$ dla $\lambda_1 = -3$ i $\lambda_2 = 6$, znajdujemy przestrzenie własne V_{-3} i V_6 macierzy $\mathbf{A}$,

$$V_{-3} = \mathcal{L}(\mathbf{x}_1) \quad \text{i} \quad V_6 = \mathcal{L}(\mathbf{x}_2, \mathbf{x}_3),$$

gdzie

$$\mathbf{x}_1 = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \quad \text{i} \quad \mathbf{x}_3 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}.$$

Wystarczy teraz wskazać bazę ortogonalną $(\mathbf{y}_1)$ przestrzeni V_{-3}, bazę ortogonalną $(\mathbf{y}_2, \mathbf{y}_3)$ przestrzeni V_6, z nich zestawić bazę ortonormalną $(\mathbf{y}_1/\|\mathbf{y}_1\|, \mathbf{y}_2/\|\mathbf{y}_2\|, \mathbf{y}_3/\|\mathbf{y}_3\|)$ przestrzeni R^3 i utworzyć macierz ortogonalną

$$\mathbf{Q} = \big[\mathbf{y}_1/\|\mathbf{y}_1\|\ \mathbf{y}_2/\|\mathbf{y}_2\|\ \mathbf{y}_3/\|\mathbf{y}_3\|\big].$$

Możemy przyjąć $\mathbf{y}_1 = \mathbf{x}_1$. Wektory $\mathbf{y}_2$ i $\mathbf{y}_3$ mogą być dowolnymi niezerowymi krotnościami wektorów uzyskanych metodą Grama-Schmidta z bazy $(\mathbf{x}_2, \mathbf{x}_3)$. W szczególności możemy przyjąć

$$\mathbf{y}_2 = \mathbf{x}_2 \quad \text{i} \quad \mathbf{y}_3 = 5\left(\mathbf{x}_3 - \frac{(\mathbf{x}_3|\mathbf{x}_2)}{(\mathbf{x}_2|\mathbf{x}_2)}\mathbf{x}_2\right) = 5\left(\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} - \frac{4}{5}\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ -4 \\ 5 \end{bmatrix}.$$

(Równie dobrze można było przyjąć, że $\mathbf{y}_3$ jest iloczynem wektorowym (definicja 12.1.1) wektorów $\mathbf{y}_1$ i $\mathbf{y}_2$, czyli przyjąć $\mathbf{y}_3 = \mathbf{y}_1 \times \mathbf{y}_2$.) Zatem szukaną macierzą jest

$$\mathbf{Q} = \left[\begin{array}{ccc} \frac{\mathbf{y}_1}{\|\mathbf{y}_1\|} & \frac{\mathbf{y}_2}{\|\mathbf{y}_2\|} & \frac{\mathbf{y}_3}{\|\mathbf{y}_3\|} \end{array} \right] = \begin{bmatrix} -1/3 & 2/\sqrt{5} & 2/\sqrt{45} \\ 2/3 & 1/\sqrt{5} & -4/\sqrt{45} \\ 2/3 & 0 & 5/\sqrt{45} \end{bmatrix}.$$

Twierdzenie spektralne dla macierzy symetrycznej

Konsekwencją poprzednich twierdzeń jest następujący wniosek nazywany twierdzeniem spektralnym dla macierzy symetrycznych.

Wniosek 10.3.2 (Twierdzenie spektralne). *Jeśli $\lambda_1, \ldots, \lambda_m$ są wszystkimi różnymi wartościami własnymi rzeczywistej macierzy symetrycznej $\mathbf{A}$ wymiaru $n \times n$, to:*

(1) *liczby $\lambda_1, \ldots, \lambda_m$ są rzeczywiste;*

(2) *podprzestrzenie własne $V_{\lambda_1}, \ldots, V_{\lambda_m}$ są wzajemnie ortogonalne;*

(3) *macierz $\mathbf{A}$ jest ortogonalnie diagonalizowalna;*

(4) $\dim V_{\lambda_i} = k(\lambda_i)$ *dla* $i = 1, \ldots, m$;

(5) $k(\lambda_1) + \ldots + k(\lambda_m) = n$;

(6) $R^n = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_m}$.

Dowód. Stwierdzenia (1), (2), (3) i (4) wynikają odpowiednio z twierdzenia 10.3.1, wniosku 10.3.1, twierdzenia 10.3.4 i twierdzenia 10.2.7.

Niech teraz $\varphi(\lambda)$ będzie wielomianem charaktrystycznym macierzy $\mathbf{A}$. Ponieważ $\varphi(\lambda)$ ma dokładnie n pierwiastków i liczby $\lambda_1, \ldots, \lambda_m$ są wszystkimi różnymi pierwiastkami wielomianu $\varphi(\lambda)$, więc suma ich krotności algebraicznych musi być równa n, czyli musi być $k(\lambda_1) + \ldots + k(\lambda_m) = n$ i to dowodzi (5).

Z faktu, że podprzestrzenie $V_{\lambda_1}, \ldots, V_{\lambda_m}$ są wzajemnie ortogonalne wynika, że $V_{\lambda_i} \cap V_{\lambda_j} = \{\mathbf{0}\}$, gdy $i \neq j$. Stąd zaś wynika, że podprzestrzeń $V_{\lambda_1} + \ldots + V_{\lambda_m}$ przestrzeni R^n jest sumą prostą podprzestrzeni $V_{\lambda_1}, \ldots, V_{\lambda_m}$, czyli $V_{\lambda_1} + \ldots + V_{\lambda_m} = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_m}$. Ponieważ

$$\begin{aligned} n = \dim R^n &\geqslant \dim (V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_m}) \\ &= \dim V_{\lambda_1} + \ldots + \dim V_{\lambda_m} = k(\lambda_1) + \ldots + k(\lambda_m) = n, \end{aligned}$$

więc $\dim R^n = \dim (V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_m})$ i wobec twierdzenia 7.5.5 przestrzeń R^n jest identyczna ze swoją podprzestrzenią $V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_m}$. To kończy dowód stwierdzenia (6). $\square$

Rozkład spektralny macierzy symetrycznej

Niech macierz ortogonalna $\mathbf{Q} = [\, \mathbf{u}_1 \ldots \mathbf{u}_n \,]$ będzie macierzą podobieństwa macierzy symetrycznej $\mathbf{A}$ do macierzy diagonalnej $\mathbf{\Lambda}$, gdzie $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Wtedy $\mathbf{A} = \mathbf{Q\Lambda Q}^{-1}$ i ponieważ $\mathbf{Q}^{-1} = \mathbf{Q}^T$, więc wobec twierdzenia 4.2.3 mamy

$$\begin{aligned} \mathbf{A} &= \mathbf{Q\Lambda Q}^T \\ &= \begin{bmatrix} | & & | \\ \mathbf{u}_1 & \ldots & \mathbf{u}_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{bmatrix} \begin{bmatrix} - & \mathbf{u}_1^T & - \\ & \vdots & \\ - & \mathbf{u}_n^T & - \end{bmatrix} \\ &= \begin{bmatrix} | & & | \\ \lambda_1\mathbf{u}_1 & \ldots & \lambda_n\mathbf{u}_n \\ | & & | \end{bmatrix} \begin{bmatrix} - & \mathbf{u}_1^T & - \\ & \vdots & \\ - & \mathbf{u}_n^T & - \end{bmatrix} \\ &= \lambda_1\mathbf{u}_1\mathbf{u}_1^T + \ldots + \lambda_n\mathbf{u}_n\mathbf{u}_n^T. \end{aligned}$$

Wniosek 10.3.3 (Rozkład spektralny macierzy). *Jeśli macierz ortogonalna* $\mathbf{Q} = [\,\mathbf{u}_1 \ldots \mathbf{u}_n\,]$ *jest macierzą podobieństwa symetrycznej macierzy* $\mathbf{A} \in R_{n\times n}$ *do macierzy diagonalnej* $\mathbf{\Lambda} = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_n)$*, to*

$$\mathbf{A} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \ldots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T. \tag{10.9}$$

Prawą stronę równości (10.9) nazywamy spektralnym rozkładem macierzy $\mathbf{A}$. Ponieważ każdy iloczyn $\mathbf{u}_i \mathbf{u}_i^T$ jest macierzą rzutu ortogonalnego na podprzestrzeń $\mathcal{L}(\mathbf{u}_i)$ (zob. (9.28) i tw. 9.6.2), więc z (10.9) wynika, że każda rzeczywista macierz symetryczna jest kombinacją liniową macierzy rzutów ortogonalnych (na swoje podprzestrzenie własne).

Przykład 239. Wyznaczyć rozkład spektralny macierzy $\mathbf{A} = \begin{bmatrix} \frac{1}{2} & \frac{3}{2} \\ \frac{3}{2} & \frac{1}{2} \end{bmatrix}$.

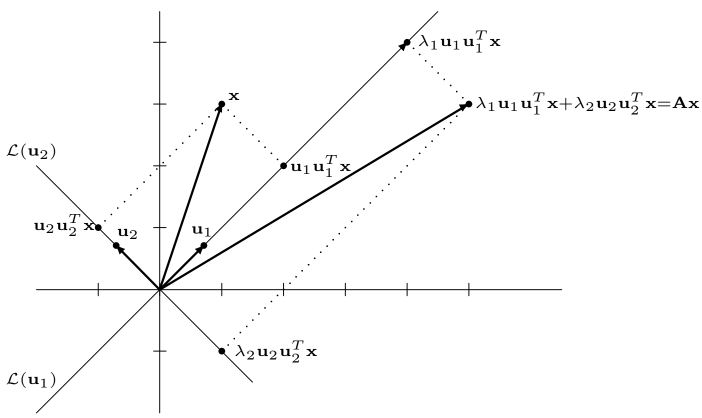
Ponieważ $\det(\mathbf{A} - \lambda\mathbf{I}) = (\lambda - 2)(\lambda + 1)$, więc wartościami własnymi macierzy $\mathbf{A}$ są $\lambda_1 = 2$ i $\lambda_2 = -1$. Łatwo także zauważyć, że

$$\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \quad \text{i} \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

są ortonormalnymi wektorami własnymi macierzy $\mathbf{A}$ odpowiadającymi wartościom własnym $\lambda_1 = 2$ i $\lambda_2 = -1$. Wobec (10.9) dla macierzy $\mathbf{A}$ mamy

$$\begin{aligned} \begin{bmatrix} 1/2 & 3/2 \\ 3/2 & 1/2 \end{bmatrix} &= \mathbf{A} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T \\ &= 2 \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} - \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \\ &= 2 \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} - \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}. \end{aligned}$$

Zauważmy także, że wektor $\mathbf{Ax}$ (dla każdego wektora $\mathbf{x} \in R^2$) jest kombinacją liniową $\lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T$ rzutów $\mathbf{u}_1 \mathbf{u}_1^T \mathbf{x}$ i $\mathbf{u}_2 \mathbf{u}_2^T \mathbf{x}$ wektora $\mathbf{x}$ na podprzestrzenie $\mathcal{L}(\mathbf{u}_1)$ i $\mathcal{L}(\mathbf{u}_2)$, zob. rys. 10.3 dla $\mathbf{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.



Rys. 10.3

10.4. Potęga macierzy diagonalizowalnej

Wyznaczanie naturalnych potęg macierzy kwadratowej jest szczególnie proste dla macierzy diagonalnych. Istotnie, indukcyjnie łatwo wykazuje się, że *jeśli* $\mathbf{\Lambda}$

jest macierzą diagonalną, powiedzmy $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$*, to dla każdej liczby naturalnej* k *jest*

$$\mathbf{\Lambda}^k = \big(\operatorname{diag}(\lambda_1, \ldots, \lambda_n)\big)^k = \operatorname{diag}(\lambda_1^k, \ldots, \lambda_n^k).$$

Potęga macierzy diagonalnej

Przykładowo mamy

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}^5 = \begin{bmatrix} 32 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 243 \end{bmatrix}.$$

Odrobinę trudniej wyznacza się naturalne potęgi macierzy diagonalizowalnych: *jeśli macierz* $\mathbf{A}$ *jest diagonalizowalna i jeśli* $\mathbf{P}$ *jest macierzą podobieństwa macierzy* $\mathbf{A}$ *do macierzy diagonalnej* $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$*, to* $\mathbf{AP} = \mathbf{P\Lambda}$ *i dla każdej liczby naturalnej* k *jest*

$$\begin{aligned} \mathbf{A}^k &= \big(\mathbf{P\Lambda P}^{-1}\big)^k = \mathbf{P\Lambda P}^{-1}\mathbf{P\Lambda P}^{-1}\ldots\mathbf{P\Lambda P}^{-1} = \mathbf{P\Lambda}^k\mathbf{P}^{-1} \\ &= \mathbf{P}\big(\operatorname{diag}(\lambda_1, \ldots, \lambda_n)\big)^k\mathbf{P}^{-1} = \mathbf{P}\operatorname{diag}(\lambda_1^k, \ldots, \lambda_n^k)\,\mathbf{P}^{-1}. \end{aligned}$$

Potęga macierzy diagonalizowalnej

Przykład 240. Wykazać, że macierz $\mathbf{A} = \begin{bmatrix} -4 & 6 \\ -3 & 5 \end{bmatrix}$ jest diagonalizowalna i wskazać macierz podobieństwa $\mathbf{P}$ macierzy $\mathbf{A}$ do macierzy diagonalnej $\mathbf{\Lambda}$. Następnie wyznaczyć $\mathbf{A}^k$ ($k \in N$) i obliczyć $\mathbf{A}^{10}\mathbf{x}$, gdy $\mathbf{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

Wielomianem charakterystycznym macierzy $\mathbf{A}$ jest $\begin{vmatrix} -4-\lambda & 6 \\ -3 & 5-\lambda \end{vmatrix} = (\lambda+1)(\lambda-2)$. Ponieważ ma on różne wartości własne, $\lambda_1 = -1$ i $\lambda_2 = 2$, więc macierz $\mathbf{A}$ jest diagonalizowalna i jest ona podobna do macierzy diagonalnej $\mathbf{\Lambda} = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$. Łatwo sprawdzić, że wektorami własnymi macierzy $\mathbf{A}$ odpowiadającymi wartościom własnym $\lambda_1 = -1$ i $\lambda_2 = 2$ są $\mathbf{x}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ i $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, więc $\mathbf{P} = [\,\mathbf{x}_1\ \mathbf{x}_2\,] = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ jest macierzą podobieństwa macierzy $\mathbf{A}$ do macierzy diagonalnej $\mathbf{\Lambda} = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$. Ponieważ $\mathbf{AP} = \mathbf{P\Lambda}$, więc $\mathbf{A} = \mathbf{P\Lambda P}^{-1}$ i dla każdej liczby naturalnej k mamy

$$\begin{aligned} \mathbf{A}^k &= \mathbf{P\Lambda}^k\mathbf{P}^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}^k\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} (-1)^k & 0 \\ 0 & 2^k \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 2(-1)^k - 2^k & 2^{k+1} + 2(-1)^{k+1} \\ (-1)^k - 2^k & 2^{k+1} + (-1)^{k+1} \end{bmatrix}. \end{aligned}$$

Stąd

$$\mathbf{A}^{10} = \begin{bmatrix} -1022 & 2046 \\ -1023 & 2047 \end{bmatrix} \quad \text{i} \quad \mathbf{A}^{10}\mathbf{x} = \begin{bmatrix} -1022 & 2046 \\ -1023 & 2047 \end{bmatrix}\begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4090 \\ 4093 \end{bmatrix}.$$

10.5. Granica ciągu macierzy

Definicja 10.5.1. Dane są macierze $\mathbf{A}$, $\mathbf{A}_1$, $\mathbf{A}_2, \ldots$ wymiaru $m \times n$. Mówimy, że ciąg macierzy $(\mathbf{A}_k)$ jest *zbieżny* do macierzy $\mathbf{A}$, gdy dla każdych indeksów $i \in \{1, \ldots, m\}$ oraz $j \in \{1, \ldots, n\}$ jest

$$\lim_{k\to\infty} (\mathbf{A}_k)_{ij} = \mathbf{A}_{ij}.$$

Granica ciągu macierzy

W takim przypadku mówimy też, że macierz $\mathbf{A}$ jest *granicą* ciągu macierzy $(\mathbf{A}_k)$ i piszemy
$$\lim_{n\to\infty} \mathbf{A}_k = \mathbf{A}.$$

Przykład 241. Jeśli $\mathbf{A}_k = \begin{bmatrix} 1-\frac{1}{k} & \left(1+\frac{1}{k}\right)^{2k} & 2^{-k} \\ \sqrt[k]{5+k} & 2 & \frac{3k-1}{k+1} \end{bmatrix}$, to mamy

$$\begin{aligned} \lim_{k\to\infty} \mathbf{A}_k &= \lim_{k\to\infty} \begin{bmatrix} 1-\frac{1}{k} & \left(1+\frac{1}{k}\right)^{2k} & 2^{-k} \\ \sqrt[k]{5+k} & 2 & \frac{3k-1}{k+1} \end{bmatrix} \\ &= \begin{bmatrix} \lim\limits_{k\to\infty}\left(1-\dfrac{1}{k}\right) & \lim\limits_{k\to\infty}\left(1+\dfrac{1}{k}\right)^{2k} & \lim\limits_{k\to\infty} 2^{-k} \\ \lim\limits_{k\to\infty} \sqrt[k]{5+k} & \lim\limits_{k\to\infty} 2 & \lim\limits_{k\to\infty} \dfrac{3k-1}{k+1} \end{bmatrix} = \begin{bmatrix} 1 & e^2 & 0 \\ 1 & 2 & 3 \end{bmatrix}. \end{aligned}$$

Dla granicy ciągu macierzy mamy następujące dwie ważne własności.

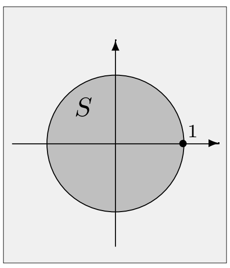
Twierdzenie 10.5.1. *Jeżeli ciąg* $(\mathbf{A}_k)$ *macierzy wymiaru* $m\times n$ *jest zbieżny do macierzy* $\mathbf{A}$*, a ciąg* $(\mathbf{B}_k)$ *macierzy wymiaru* $n\times p$ *jest zbieżny do macierzy* $\mathbf{B}$*, to ciąg* $(\mathbf{A}_k\mathbf{B}_k)$ *jest zbieżny do macierzy* $\mathbf{AB}$*,*
$$\lim_{k\to\infty} \mathbf{A}_k\mathbf{B}_k = \left(\lim_{k\to\infty} \mathbf{A}_k\right)\left(\lim_{k\to\infty} \mathbf{B}_k\right) = \mathbf{AB}.$$

Dowód. Dla każdego $i \in \{1,\ldots,m\}$ oraz $j\in\{1,\ldots,p\}$ jest
$$\begin{aligned} \lim_{k\to\infty} (\mathbf{A}_k\mathbf{B}_k)_{ij} &= \lim_{k\to\infty} \sum_{l=1}^{n} (\mathbf{A}_k)_{il} (\mathbf{B}_k)_{lj} = \sum_{l=1}^{n} \left(\lim_{k\to\infty} (\mathbf{A}_k)_{il}\right)\left(\lim_{k\to\infty} (\mathbf{B}_k)_{lj}\right) \\ &= \textstyle\sum_{l=1}^{n} (\mathbf{A})_{il} (\mathbf{B})_{lj} = (\mathbf{AB})_{ij}. \end{aligned}$$
To zaś oznacza, że $\lim\limits_{k\to\infty} \mathbf{A}_k\mathbf{B}_k = \mathbf{AB}$. $\square$

Wniosek 10.5.1. *Jeśli* $\mathbf{\Lambda}$ *i* $\mathbf{P}$ *są macierzami kwadratowymi tego samego stopnia i* $\mathbf{P}$ *jest macierzą odwracalną, to granica* $\lim\limits_{k\to\infty} \mathbf{\Lambda}^k$ *istnieje wtedy i tylko wtedy, gdy istnieje granica* $\lim\limits_{k\to\infty} \mathbf{P\Lambda}^k\mathbf{P}^{-1}$ *i wtedy też*
$$\lim_{k\to\infty} \mathbf{P\Lambda}^k\mathbf{P}^{-1} = \mathbf{P}\left(\lim_{k\to\infty} \mathbf{\Lambda}^k\right)\mathbf{P}^{-1}.$$

Dowód. Ponieważ $\lim\limits_{k\to\infty} \mathbf{P} = \mathbf{P}$ i $\lim\limits_{k\to\infty} \mathbf{P}^{-1} = \mathbf{P}^{-1}$, więc z istnienia granicy $\lim\limits_{k\to\infty} \mathbf{\Lambda}^k$ i z poprzedniego twierdzenia wynika istnienie granicy $\lim\limits_{k\to\infty} \mathbf{P\Lambda}^k\mathbf{P}^{-1}$, bo
$$\lim_{k\to\infty} \mathbf{P\Lambda}^k\mathbf{P}^{-1} = \left(\lim_{k\to\infty} \mathbf{P}\right)\left(\lim_{k\to\infty} \mathbf{\Lambda}^k\right)\left(\lim_{k\to\infty} \mathbf{P}^{-1}\right) = \mathbf{P}\left(\lim_{k\to\infty} \mathbf{\Lambda}^k\right)\mathbf{P}^{-1}.$$
Odwrotna implikacja także jest oczywista, bo mamy
$$\lim_{k\to\infty} \mathbf{\Lambda}^k = \lim_{k\to\infty} \mathbf{P}^{-1}\big(\mathbf{P\Lambda}^k\mathbf{P}^{-1}\big)\mathbf{P} = \mathbf{P}^{-1}\Big(\lim_{k\to\infty} \mathbf{P\Lambda}^k\mathbf{P}^{-1}\Big)\mathbf{P}. \quad \square$$

W naszych rozważaniach o granicy ciągu macierzy będziemy odwoływać się do zbioru
$$\mathcal{S} = \{\lambda \in \mathcal{C}\colon\ |\lambda| < 1 \quad \text{lub} \quad \lambda = 1\}.$$



Rys. 10.4

Geometrycznie zbiór $\mathcal{S}$ składa się z liczby zespolonej 1 i liczb, które tworzą wnętrze jednostkowego koła o środku w punkcie 0 (zob. rys. 10.4). Warto zauważyć, że zbiór $\mathcal{S}$ składa się z tych i tylko tych liczb zespolonych λ, dla których granica $\lim_{k\to\infty} \lambda^k$ istnieje.

Można udowodnić, że dla macierzy $\mathbf{A} \in \mathcal{C}_{n\times n}$ granica $\lim_{k\to\infty} \mathbf{A}^k$ istnieje wtedy i tylko wtedy, gdy każda wartość własna λ macierzy $\mathbf{A}$ należy do zbioru $\mathcal{S}$ i jeśli liczba 1 jest wartością własną macierzy $\mathbf{A}$, to jej krotność algebraiczna jest identyczna z jej krotnością geometryczną. W naszych zastosowaniach rozważać będziemy tylko granice potęg macierzy diagonalizowalnych. Dla każdej takiej macierzy $\mathbf{A}$ krotność algebraiczna każdej wartości własnej automatycznie jest równa jej krotności geometrycznej i mamy następujący warunek konieczny i dostateczny istnienia granicy $\lim_{k\to\infty} \mathbf{A}^k$.

Twierdzenie 10.5.2. *Jeśli macierz $\mathbf{A} \in \mathcal{C}_{n\times n}$ jest diagonalizowalna, to granica $\lim_{k\to\infty} \mathbf{A}^k$ istnieje wtedy i tylko wtedy, gdy każda wartość własna λ macierzy $\mathbf{A}$ należy do zbioru $\mathcal{S}$. Dodatkowo, jeśli $\mathbf{A} = \mathbf{P\Lambda P}^{-1}$, gdzie $\mathbf{P}$ jest macierzą odwracalną i $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ dla pewnych liczb $\lambda_1, \ldots, \lambda_n \in \mathcal{S}$, to*

$$\lim_{k\to\infty} \mathbf{A}^k = \mathbf{P} \operatorname{diag}\left(\lim_{k\to\infty} \lambda_1^k, \ldots, \lim_{k\to\infty} \lambda_n^k\right) \mathbf{P}^{-1}.$$

Dowód. Ponieważ $\mathbf{A}^k = \left(\mathbf{P\Lambda P}^{-1}\right)^k = \mathbf{P\Lambda}^k\mathbf{P}^{-1} = \mathbf{P} \operatorname{diag}(\lambda_1^k \ldots, \lambda_n^k)\, \mathbf{P}^{-1}$ i granica

$$\lim_{k\to\infty} \Lambda^k = \lim_{k\to\infty} \operatorname{diag}(\lambda_1^k, \ldots, \lambda_n^k) = \operatorname{diag}\left(\lim_{k\to\infty} \lambda_1^k, \ldots, \lim_{k\to\infty} \lambda_n^k\right)$$

istnieje wtedy i tylko wtedy, gdy każda z liczb $\lambda_1, \ldots, \lambda_n$ należy do zbioru $\mathcal{S}$, więc wobec wniosku 10.5.1 granica $\lim\limits_{k\to\infty} \mathbf{A}^k = \lim\limits_{k\to\infty} \mathbf{P\Lambda}^k\mathbf{P}^{-1}$ istnieje i jest równa macierzy $\mathbf{P}\left(\lim\limits_{k\to\infty} \mathbf{\Lambda}^k\right)\mathbf{P}^{-1}$ wtedy i tylko wtedy, gdy $\lambda_1, \ldots, \lambda_n \in \mathcal{S}$. $\square$

Przykład 242. Obliczyć granicę $\lim\limits_{k\to\infty} \mathbf{A}^k$, gdy $\mathbf{A} = \begin{bmatrix} 1/2 & 1/2 & -1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Ponieważ $\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 1/2-\lambda & 1/2 & -1/2 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} = (\frac{1}{2} - \lambda)(1-\lambda)^2$, więc liczby $\lambda_1 = \frac{1}{2}$, $\lambda_2 = \lambda_3 = 1$ są wartościami własnymi macierzy $\mathbf{A}$. Rozwiązując układ równań $(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{x} = \mathbf{0}$ (dla $i = 1, 2, 3$), stwierdzamy, że wektory

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

są wektorami własnymi macierzy $\mathbf{A}$ odpowiadającymi wartościom własnym λ_1, λ_2 i λ_3. Wektory $\mathbf{v}_1$, $\mathbf{v}_2$ i $\mathbf{v}_3$ tworzą bazę przestrzeni R^3, więc macierz $\mathbf{A}$ jest diagonalizowalna i mamy $\mathbf{A} = \mathbf{P\Lambda P}^{-1}$, gdzie $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$ i $\mathbf{P} = [\,\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3\,]$. Dodatkowo, ponieważ wszystkie wartości własne macierzy $\mathbf{A}$ należą do zbioru $\mathcal{S}$, więc wobec ostatniego wniosku granica $\lim_{k\to\infty} \mathbf{A}^k$ istnieje i

$$\begin{aligned}
\lim_{k\to\infty} \mathbf{A}^k &= \lim_{k\to\infty} \mathbf{P\Lambda}^k\mathbf{P}^{-1} = \mathbf{P}\left(\lim_{k\to\infty} \mathbf{\Lambda}^k\right)\mathbf{P}^{-1} \\
&= \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \left(\lim_{k\to\infty} \begin{bmatrix} \left(\frac{1}{2}\right)^k & 0 & 0 \\ 0 & 1^k & 0 \\ 0 & 0 & 1^k \end{bmatrix}\right) \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.
\end{aligned}$$

10.6. Podprzestrzenie niezmiennicze

Podprzestrzeń niezmiennicza

Definicja 10.6.1. Niech T będzie operatorem liniowym na przestrzeni wektorowej V. Podprzestrzeń W przestrzeni V nazywamy *T-niezmienniczą podprzestrzenią*, gdy $T(W) \subseteq W$, czyli gdy $T(\mathbf{x}) \in W$ dla każdego wektora $\mathbf{x} \in W$.

Przykład 243. Jeśli T jest operatorem liniowym na przestrzeni wektorowej V, to T-niezmienniczymi podprzestrzeniami przestrzeni V m.in. są: przestrzeń zerowa $\{\mathbf{0}\}$, cała przestrzeń V, jądro $\operatorname{Ker} T$ przekształcenia T, obraz $\operatorname{Im} T$ przekształcenia T oraz przestrzeń V_λ wektorów własnych operatora T odpowiadających wartości własnej λ.

Przykład 244. Jeśli T jest operatorem liniowym na przestrzeni wektorowej V i $\mathbf{v}$ jest niezerowym wektorem z przestrzeni V, to podprzestrzeń

Podprzestrzeń T-cykliczna

$$W = \mathcal{L}(\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}), \ldots)$$

nazywamy *T-cykliczną podprzestrzenią* przestrzeni V generowaną przez wektor $\mathbf{v}$. Jest oczywiste, że jest to T-niezmiennicza podprzestrzeń. Bez trudu dowodzi się także, że jest to najmniejsza T-niezmiennicza podprzestrzeń przestrzeni V zawierająca wektor $\mathbf{v}$. Z T-cyklicznej podprzestrzeni korzystamy w dowodzie twierdzenia Cayleya-Hamiltona, przy bezwyznacznikowym wyznaczaniu wielomianu charakterystycznego operatora liniowego i przy analizie macierzowych reprezentacji niediagonalizowalnych operatorów liniowych.

Przykład 245. Wyznaczyć T-cykliczną podprzestrzeń przestrzeni R^3 generowaną przez wektor $\mathbf{v} = (1, 1, 1)$, gdy operator T określony jest wzorem

$$T(x, y, z) = (x, x + y, y + z).$$

Ponieważ $T(\mathbf{v}) = T(1, 1, 1) = (1, 2, 2)$, $T^2(\mathbf{v}) = T(1, 2, 2) = (1, 3, 4)$, $T^3(\mathbf{v}) = T(1, 3, 4) = (1, 4, 7)$, więc

$$R^3 \supseteq \mathcal{L}\left(\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}), \ldots\right) \supseteq \mathcal{L}\big((1, 1, 1), (1, 2, 2), (1, 3, 4)\big) = R^3$$

i dlatego mamy $\mathcal{L}\left(\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}), \ldots\right) = R^3$.

Obcięcie operatora

Jeśli T jest operatorem liniowym na przestrzeni wektorowej V i jeśli W jest T-niezmienniczą podprzestrzenią przestrzeni V, to symbolem T_W oznaczamy *obcięcie* operatora T do zbioru W, czyli funkcję $T_W : W \to W$ taką, że $T_W(\mathbf{x}) = T(\mathbf{x})$ dla każdego $\mathbf{x} \in W$. Operator T_W dziedziczy pewne własności po operatorze T. Przede wszystkim łatwo zauważyć, że T_W jest operatorem liniowym na przestrzeni wektorowej W. Kolejną własność, związek pomiędzy wielomianami charakterystycznymi operatorów T i T_W, przynosi następujące twierdzenie.

Twierdzenie 10.6.1. *Niech T będzie operatorem liniowym na skończenie wymiarowej przestrzeni wektorowej V i niech W będzie T-niezmienniczą podprzestrzenią w przestrzeni V. Wtedy wielomian charakterystyczny operatora T_W dzieli wielomian charakterystyczny operatora T.*

Dowód. Niech $B = (\mathbf{b}_1, \ldots, \mathbf{b}_k)$ będzie bazą podprzestrzeni W i niech $C = (\mathbf{b}_1, \ldots, \mathbf{b}_k, \mathbf{b}_{k+1}, \ldots, \mathbf{b}_n)$ będzie rozszerzeniem bazy B do bazy całej przestrzeni V. Dla macierzy $[T_W]_B$ operatora T_W względem bazy B i macierzy $[T]_C$ operatora T względem bazy C mamy

$$[T]_C = \left[\begin{array}{c|c} [T_W]_B & \mathbf{A} \\ \hline \mathbf{0} & \mathbf{B} \end{array}\right],$$

gdzie $\mathbf{0}$ jest macierzą zerową wymiaru $(n-k)\times k$, $\mathbf{A}$ jest pewną macierzą wymiaru $k\times(n-k)$ i $\mathbf{B}$ jest pewną macierzą wymiaru $(n-k)\times(n-k)$. Jeśli $\varphi(\lambda)$ i $\psi(\lambda)$ są wielomianami charakterystycznymi operatorów T i T_W, to mamy

$$\begin{aligned}\varphi(\lambda) &= \det\big([T]_C-\lambda\mathbf{I}_n\big)=\det\left[\begin{array}{c|c}[T_W]_B-\lambda\mathbf{I}_k & \mathbf{A}\\ \hline \mathbf{0} & \mathbf{B}-\lambda\mathbf{I}_{n-k}\end{array}\right]\\ &= \det\big([T_W]_B-\lambda\mathbf{I}_k\big)\cdot\det\left(\mathbf{B}-\lambda\mathbf{I}_{n-k}\right)=\psi(\lambda)\cdot\det\left(\mathbf{B}-\lambda\mathbf{I}_{n-k}\right)\end{aligned}$$

i to dowodzi, że wielomian $\psi(\lambda)$ dzieli wielomian $\varphi(\lambda)$. □

Przykład 246. Dana jest macierz $\mathbf{A}=\begin{bmatrix}1&3&3&-3\\3&1&-3&3\\3&-3&1&3\\-3&3&3&1\end{bmatrix}$ i operator liniowy $T:R^4\to R^4$, gdzie $T(\mathbf{x})=\mathbf{Ax}$ dla $\mathbf{x}\in R^4$. Zauważmy, że podprzestrzeń $W\subset R^4$ generowana przez wektory $\mathbf{a}_1=(1,0,0,-1)$ i $a_2=(0,1,0,1)$ jest T-niezmiennicza, bo $T(\mathbf{a}_1)=4\mathbf{a}_1\in W$ i $T(\mathbf{a}_2)=4\mathbf{a}_2\in W$. Zauważmy także, że $B=(\mathbf{a}_1,\mathbf{a}_2)$ jest bazą przestrzeni W i $C=(\mathbf{a}_1,\mathbf{a}_2,\mathbf{e}_3,\mathbf{e}_4)$ jest rozszerzeniem bazy B do bazy całej przestrzeni R^4. Łatwo teraz sprawdzić, że dla macierzy $[T_W]_B$ operatora T_W względem bazy B i macierzy $[T]_C$ operatora T względem bazy C mamy

$$[T_W]_B=\begin{bmatrix}4&0\\0&4\end{bmatrix}\quad\text{i}\quad [T]_C=\left[\begin{array}{cc|cc}4&0&3&-3\\0&4&-3&3\\ \hline 0&0&1&3\\0&0&9&5\end{array}\right].$$

Jeśli teraz $\varphi(\lambda)$ i $\psi(\lambda)$ są wielomianami charakterystycznymi operatorów T i T_W, to

$$\begin{aligned}\varphi(\lambda) &= \det\left[\begin{array}{cc|cc}4-\lambda&0&3&-3\\0&4-\lambda&-3&3\\ \hline 0&0&1-\lambda&3\\0&0&9&5-\lambda\end{array}\right]\\ &= \det\begin{bmatrix}4-\lambda&0\\0&4-\lambda\end{bmatrix}\cdot\det\begin{bmatrix}1-\lambda&3\\9&5-\lambda\end{bmatrix}=\psi(\lambda)\cdot\det\begin{bmatrix}1-\lambda&3\\9&5-\lambda\end{bmatrix}\\ &= (4-\lambda)^2(\lambda^2+4\lambda-32)=(\lambda-4)^3(\lambda+8).\end{aligned}$$

Ostatnie twierdzenie jest szczególnie wygodne, gdy mamy do czynienia z podprzestrzenią cykliczną operatora, bo (jak zaraz zobaczymy) wielomian charakterystyczny operatora obciętego do podprzestrzeni cyklicznej jest łatwo wyznaczalny. (W dodatku, można ten wielomian wyznaczyć metodą bezwyznacznikową.)

Twierdzenie 10.6.2. *Niech T będzie operatorem liniowym na skończenie wymiarowej przestrzeni wektorowej V i niech W będzie T-cykliczną podprzestrzenią generowaną przez wektor $\mathbf{v}\in V$. Jeśli $\dim W=k$, to wtedy:*

(*a*) *układ $\big(\mathbf{v},T(\mathbf{v}),T^2(\mathbf{v}),\ldots,T^{k-1}(\mathbf{v})\big)$ jest bazą przestrzeni W;*

(*b*) *jeśli $a_0,a_1,\ldots,a_{k-1}$ są skalarami takimi, że*

$$a_0\mathbf{v}+a_1T(\mathbf{v})+\ldots+a_{k-1}T^{k-1}(\mathbf{v})+T^k(\mathbf{v})=\mathbf{0},$$

to

$$\psi(\lambda)=(-1)^k\left(a_0+a_1\lambda+a_2\lambda^2+\ldots+a_{k-1}\lambda^{k-1}+\lambda^k\right)$$

jest wielomianem charakterystycznym operatora T_W.

Dowód. (*a*) Ponieważ wektor $\mathbf{v}$ jest niezerowy i przestrzeń V jest skończenie wymiarowa, więc istnieje największa liczba naturalna l taka, że układ wektorów $B=\big(\mathbf{v},T(\mathbf{v}),\ldots,T^{l-1}(\mathbf{v})\big)$ jest liniowo niezależny. Niech U będzie podprzestrzenią generowaną przez wszystkie wektory układu B. Oczywiście, B jest bazą przestrzeni U. Z wyboru l

jest także oczywiste, że układ $\left(\mathbf{v}, T(\mathbf{v}), \ldots, T^{l-1}(\mathbf{v}), T^{l}(\mathbf{v})\right)$ jest liniowo zależny. Stąd $T^l(\mathbf{v}) \in U$, więc $\mathcal{L}\left(\mathbf{v},\ T(\mathbf{v}), \ldots, T^{l-1}(\mathbf{v}), T^l(\mathbf{v})\right) = \mathcal{L}\left(\mathbf{v},\ T(\mathbf{v}), \ldots, T^{l-1}(\mathbf{v})\right) = U$. Pokażemy teraz, że U jest T-niezmienniczą podprzestrzenią. Weźmy $\mathbf{x} \in U$. Dla takiego $\mathbf{x}$ istnieją skalary $\alpha_0, \alpha_1, \ldots, \alpha_{l-1}$ takie, że $\mathbf{x} = \alpha_0\mathbf{v} + \alpha_1 T(\mathbf{v}) + \ldots + \alpha_{l-1}T^{l-1}(\mathbf{v})$. Wtedy $T(\mathbf{x}) = \alpha_0 T(\mathbf{v}) + \alpha_1 T^2(\mathbf{v}) + \ldots + \alpha_{l-2}T^{l-1}(\mathbf{v}) + \alpha_{l-1}T^l(\mathbf{v})$ i $T(\mathbf{x}) \in U$, bo $\mathcal{L}\left(\mathbf{v}, T(\mathbf{v}), \ldots, T^{l-1}(\mathbf{v}), T^l(\mathbf{v})\right) = \mathcal{L}\left(\mathbf{v}, T(\mathbf{v}), \ldots, T^{l-1}(\mathbf{v})\right) = U$. To dowodzi, że U jest T-niezmienniczą podprzestrzenią zawierającą wektor $\mathbf{v}$. Ponieważ W jest najmniejszą T-niezmienniczą podprzestrzenią zawierającą wektor $\mathbf{v}$, więc $W \subseteq U$. Jednocześnie $U \subseteq W$, więc także mamy $U = W$. To dowodzi, że B jest bazą przestrzeni W i $k = \dim W = |B| = l$.

(*b*) Niech $B = \left(\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}), \ldots, T^{k-1}(\mathbf{v})\right)$ będzie bazą przestrzeni W (zob. (*a*)) i niech $a_0, a_1, \ldots, a_{k-1}$ będą skalarami, dla których $a_0\mathbf{v}+a_1T(\mathbf{v})+\ldots+a_{k-1}T^{k-1}(\mathbf{v})+T^k(\mathbf{v}) = \mathbf{0}$. Wtedy macierzą operatora T_W względem bazy B jest

$$[T_W]_B = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{k-1} \end{bmatrix}$$

i wielomianem charakterystycznym operatora T_W jest

$$\psi(\lambda) = \det\left([T_W]_B - \lambda\mathbf{I}_k\right) = \begin{vmatrix} -\lambda & 0 & 0 & \ldots & 0 & -a_0 \\ 1 & -\lambda & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & -\lambda & \ldots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{k-1}-\lambda \end{vmatrix}.$$

Rozwijając ostatni wyznacznik względem elementów pierwszego wiersza i korzystając z indukcji, łatwo stwierdzamy, że

$$\psi(\lambda) = (-1)^k\left(a_0 + a_1\lambda + a_2\lambda^2 + \ldots + a_{k-1}\lambda^{k-1} + \lambda^k\right). \quad \square$$

Przykład 247. T-cykliczna podprzestrzeń operatora $T(x, y, z) = (x, x + y, y + z)$ generowana przez wektor $\mathbf{v} = (1, 1, 1)$ (zob. przykład 245) jest przestrzenią $W = \mathcal{L}(\mathbf{v}, T(\mathbf{v}), \ldots)$ wymiaru 3 i jej bazą jest układ $\left(\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v})\right)$. Dla skalarów $a_0 = -1$, $a_1 = 3$ i $a_2 = -3$ mamy $a_0\mathbf{v} + a_1T(\mathbf{v}) + a_2T^2(\mathbf{v}) + T^3(\mathbf{v}) = \mathbf{0}$, bo

$$-1\begin{bmatrix}1\\1\\1\end{bmatrix} + 3\begin{bmatrix}1\\2\\2\end{bmatrix} - 3\begin{bmatrix}1\\3\\4\end{bmatrix} + \begin{bmatrix}1\\4\\7\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix},$$

więc z ostatniego twierdzenia wynika, że wielomianem charakterystycznym operatora T_W (i identycznego z nim operatora T) jest

$$\varphi(\lambda) = (-1)^3(-1 + 3\lambda - 3\lambda^2 + \lambda^3) = \det\begin{bmatrix} -\lambda & 0 & 1 \\ 1 & -\lambda & -3 \\ 0 & 1 & 3-\lambda \end{bmatrix}.$$

10.7. Twierdzenie Cayleya-Hamiltona

Niech $\psi(x)$ będzie wielomianem o współczynnikach z ciała K, powiedzmy $\psi(x) = a_0 + a_1x + \ldots + a_kx^k$, i niech $\mathbf{A}$ będzie macierzą wymiaru $n \times n$ o współczynnikach z ciała K. Wtedy

$$\psi(\mathbf{A}) = a_0\mathbf{I}_n + a_1\mathbf{A} + \ldots + a_k\mathbf{A}^k$$

jest macierzą z przestrzeni $K_{n\times n}$. Z faktu, że $K_{n\times n}$ jest przestrzenią wektorową wymiaru n^2 wynika, że (n^2+1)-elementowy zbiór

$$\{\mathbf{I}_n, \mathbf{A}, \ldots, \mathbf{A}^{n^2}\}$$

jest liniowo zależny. Zatem istnieją skalary $a_0, a_1, \ldots, a_{n^2} \in K$ takie, że

$$a_0\mathbf{I}_n + a_1\mathbf{A} + \ldots + a_{n^2}\mathbf{A}^{n^2} = \mathbf{0}.$$

To zaś jest równoważne istnieniu wielomianu $\varphi(x)$ takiego, że $\varphi(\mathbf{A}) = \mathbf{0}$. Podobnie, jeśli V jest n-wymiarową przestrzenią wektorową nad ciałem K i jeśli $T\colon V \to V$ jest operatorem liniowym, to z faktu, że $L(V,V)$ jest n^2-wymiarową przestrzenią wektorową wynika, że przekształcenia liniowe $I_V, T, \ldots, T^{n^2}$ są liniowo zależne. Stąd zaś wynika istnienie wielomianu $\varphi(x)$ takiego, że $\varphi(T)$ jest zerowym przekształceniem liniowym. Udowodnimy teraz, że taką własność ma wielomian charakterystyczny operatora liniowego (macierzy kwadratowej). Innymi słowy udowodnimy, że operator liniowy (macierz kwadratowa) jest "pierwiastkiem" swojego równania charakterystycznego.

Twierdzenie 10.7.1 (Cayleya-Hamiltona). *Niech T będzie operatorem liniowym na skończenie wymiarowej przestrzeni wektorowej V i niech $\varphi(\lambda)$ będzie jego wielomianem charakterystycznym. Wtedy $\varphi(T)$ jest operatorem zerowym na przestrzeni V, czyli $\varphi(T) = \mathbf{0}$.*

Twierdzenie Cayleya-Hamiltona

Dowód. Wystarczy pokazać, że $\varphi(T)(\mathbf{v}) = \mathbf{0}$ dla każdego wektora $\mathbf{v} \in V$. Jest to oczywiste, gdy $\mathbf{v} = \mathbf{0}$ (bo $\varphi(T)$ jest przekształceniem liniowym). Załóżmy teraz, że $\mathbf{v} \neq \mathbf{0}$ i niech W będzie T-cykliczną podprzestrzenią generowaną przez wektor $\mathbf{v}$. Jeśli $\dim W = k$, to wobec twierdzenia 10.6.2 układ $\big(\mathbf{v}, T(\mathbf{v}), \ldots, T^{k-1}(\mathbf{v})\big)$ jest bazą przestrzeni W i dlatego istnieją skalary $a_0, a_1, \ldots, a_{k-1}$ takie, że

$$a_0\mathbf{v} + a_1T(\mathbf{v}) + \ldots + a_{k-1}T^{k-1}(\mathbf{v}) + T^k(\mathbf{v}) = \mathbf{0}.$$

Wtedy także wobec poprzedniego twierdzenia wielomian

$$\psi(\lambda) = (-1)^k \left(a_0 + a_1\lambda + \ldots + a_{k-1}\lambda^{k-1} + \lambda^k\right)$$

jest wielomianem charakterystycznym operatora T_W. Teraz z obu powyższych równości wynika, że

$$\begin{aligned} \psi(T)(\mathbf{v}) &= (-1)^k \left(a_0I + a_1T + \ldots + a_{k-1}T^{k-1} + T^k\right)(\mathbf{v}) \\ &= (-1)^k \left(a_0\mathbf{v} + a_1T(\mathbf{v}) + \ldots + a_{k-1}T^{k-1}(\mathbf{v}) + T^k(\mathbf{v})\right) = \mathbf{0}. \end{aligned}$$

Ponieważ wielomian $\psi(\lambda)$ dzieli wielomian $\varphi(\lambda)$, więc istnieje wielomian $\varrho(\lambda)$ taki, że $\varphi(\lambda) = \varrho(\lambda)\psi(\lambda)$. Wtedy także mamy

$$\varphi(T)(\mathbf{v}) = \big(\varrho(T)\psi(T)\big)(\mathbf{v}) = \varrho(T)\big(\psi(T)(\mathbf{v})\big) = \varrho(T)(\mathbf{0}) = \mathbf{0}$$

i to kończy dowód twierdzenia. □

Wniosek 10.7.1 (Twierdzenie Cayleya-Hamiltona dla macierzy). *Jeśli $\mathbf{A}$ jest macierzą kwadratową i*

$$\varphi(\lambda) = (-1)^n \left(a_0 + a_1\lambda + \ldots + a_{n-1}\lambda^{n-1} + \lambda^n\right)$$

jest jej wielomianem charakterystycznym, to $\varphi(\mathbf{A})$ jest macierzą zerową, więc

$$a_0\mathbf{I} + a_1\mathbf{A} + \ldots + a_{n-1}\mathbf{A}^{n-1} + \mathbf{A}^n = \mathbf{0}. \quad \square$$

Przykład 248. Niech $T\colon R^2 \to R^2$ będzie operatorem liniowym takim, że $T(x,y) = (5x-6y, 3x-4y)$ dla $(x,y) \in R^2$. Macierzą operatora T względem bazy kanonicznej przestrzeni R^2 jest macierz

$$\mathbf{A} = \begin{bmatrix} 5 & -6 \\ 3 & -4 \end{bmatrix},$$

więc jego wielomianem charakterystycznym jest

$$\varphi(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}_2) = \det\begin{bmatrix} 5-\lambda & -6 \\ 3 & -4-\lambda \end{bmatrix} = \lambda^2 - \lambda - 2.$$

Zauważmy teraz, że operator liniowy $\varphi(T) = T^2 - T - 2I$ jest zerowy, bo dla każdego wektora $\mathbf{v} = (x,y) \in R^2$ mamy

$$\begin{aligned} \varphi(T)(\mathbf{v}) &= T\big(T(\mathbf{v})\big) - T(\mathbf{v}) - 2\mathbf{v} = T\big(T(x,y)\big) - T(x,y) - 2(x,y) \\ &= T(5x-6y, 3x-4y) - (5x-6y, 3x-4y) - (2x, 2y) \\ &= (7x-6y, 3x-2y) - (5x-6y, 3x-4y) - (2x,2y) = (0,0). \end{aligned}$$

Podobnie macierz $\varphi(\mathbf{A})$ jest zerowa, bo

$$\begin{aligned} \varphi(\mathbf{A}) &= \mathbf{A}^2 - \mathbf{A} - 2\mathbf{I} = \begin{bmatrix} 5 & -6 \\ 3 & -4 \end{bmatrix}^2 - \begin{bmatrix} 5 & -6 \\ 3 & -4 \end{bmatrix} - 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 7 & -6 \\ 3 & -2 \end{bmatrix} - \begin{bmatrix} 5 & -6 \\ 3 & -4 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

Twierdzenie Cayleya-Hamiltona i potęgi macierzy

Pokażemy teraz w jaki sposób twierdzenie Cayleya-Hamiltona może być pomocne przy wyznaczaniu potęg macierzy. Cała tajemnica tkwi w następującym prostym twierdzeniu.

Twierdzenie 10.7.2. *Jeśli macierz $\mathbf{A}$ o współczynnikach z ciała K jest macierzą wymiaru $n \times n$ i jeśli k jest liczbą naturalną, to macierz $\mathbf{A}^k$ jest kombinacją liniową macierzy $\mathbf{I}_n, \mathbf{A}, \ldots, \mathbf{A}^{n-1}$.*

Dowód. Teza jest oczywista dla $k = 0, 1, \ldots, n-1$. Jeśli $k = n$, to stwierdzenie wynika z ostatniego wniosku. Istotnie, jeśli $\varphi(\lambda) = (-1)^n \left(a_0 + a_1\lambda + \ldots + a_{n-1}\lambda^{n-1} + \lambda^n\right)$ jest wielomianem charakterystycznym macierzy $\mathbf{A}$, to $\varphi(\mathbf{A}) = \mathbf{0}$ i z równości tej mamy

$$\mathbf{A}^n = -a_0\mathbf{I} - a_1\mathbf{A} + \ldots - a_{n-1}\mathbf{A}^{n-1},$$

co oznacza, że macierz $\mathbf{A}^n$ jest kombinacją liniową macierzy $\mathbf{I}, \mathbf{A}, \ldots, \mathbf{A}^{n-1}$. Jeśli jest teraz $k > n$ i jeśli założymy, że macierz $\mathbf{A}^{k-1}$ jest kombinacją liniową macierzy $\mathbf{I}, \mathbf{A}, \ldots, \mathbf{A}^{n-1}$, powiedzmy

$$\mathbf{A}^{k-1} = x_0\mathbf{I} + x_1\mathbf{A} + \ldots + x_{n-1}\mathbf{A}^{n-1}$$

dla pewnych skalarów $x_0, x_1, \ldots, x_{n-1}$, to dla macierzy $\mathbf{A}^k$ mamy

$$\begin{aligned} \mathbf{A}^k &= \mathbf{A}^{k-1}\mathbf{A} = \left(x_0\mathbf{I} + x_1\mathbf{A} + \ldots + x_{n-1}\mathbf{A}^{n-1}\right)\mathbf{A} \\ &= x_0\mathbf{A} + \ldots + x_{n-2}\mathbf{A}^{n-1} + x_{n-1}\mathbf{A}^n \\ &= x_0\mathbf{A} + \ldots + x_{n-2}\mathbf{A}^{n-1} + x_{n-1}\left(-a_0\mathbf{I} - a_1\mathbf{A} + \ldots - a_{n-1}\mathbf{A}^{n-1}\right) \end{aligned}$$

i to dowodzi, że także $\mathbf{A}^k$ jest kombinacją liniową macierzy $\mathbf{I}, \mathbf{A}, \ldots, \mathbf{A}^{n-1}$. $\square$

Przykład 249. Korzystając z twierdzenia Cayleya-Hamiltona, obliczyć $\mathbf{A}^2$, $\mathbf{A}^3$ oraz $\mathbf{A}^4$, gdy $\mathbf{A} = \begin{bmatrix} -2 & 0 \\ 7 & -1 \end{bmatrix}$.

Wielomianem charakterystycznym macierzy $\mathbf{A}$ jest $\varphi(\lambda) = \lambda^2 + 3\lambda + 2$. Stąd i z twierdzenia Cayleya-Hamiltona mamy $\mathbf{A}^2 + 3\mathbf{A} + 2\mathbf{I} = \mathbf{0}$ i dlatego kolejno

$$\begin{aligned} \mathbf{A}^2 &= -3\mathbf{A} - 2\mathbf{I} = \begin{bmatrix} 4 & 0 \\ -21 & 1 \end{bmatrix}, \\ \mathbf{A}^3 &= \mathbf{A}(-3\mathbf{A} - 2\mathbf{I}) = -3(-3\mathbf{A} - 2\mathbf{I}) - 2\mathbf{A} = 7\mathbf{A} + 6\mathbf{I} = \begin{bmatrix} -8 & 0 \\ 49 & -1 \end{bmatrix}, \\ \mathbf{A}^4 &= \mathbf{A}(7\mathbf{A} + 6\mathbf{I}) = 7(-3\mathbf{A} - 2\mathbf{I}) + 6\mathbf{A} = -15\mathbf{A} - 14\mathbf{I} = \begin{bmatrix} 16 & 0 \\ -105 & 1 \end{bmatrix}. \end{aligned}$$

Z ostatniego twierdzenia łatwo wynika, że jeśli macierz $\mathbf{A}$ o współczynnikach z ciała K jest wymiaru $n \times n$ i jeśli $\psi(x)$ jest wielomianem o współczynnikach z ciała K, to macierz $\psi(\mathbf{A})$ jest kombinacją liniową macierzy $\mathbf{I}_n, \mathbf{A} \ldots, \mathbf{A}^{n-1}$, czyli istnieją skalary $x_0, x_1, \ldots, x_{n-1}$ takie, że

$$\psi(\mathbf{A}) = x_0\mathbf{I}_n + x_1\mathbf{A} + \ldots + x_{n-1}\mathbf{A}^{n-1}.$$

Prosty sposób wyznaczania współczynników ostatniej kombinacji jest konsekwencją następującego twierdzenia.

Twierdzenie 10.7.3. *Niech $\varphi(x)$ będzie wielomianem charakterystycznym macierzy $\mathbf{A} \in K_{n\times n}$. Jeśli $\psi(x)$ jest wielomianem o współczynnikach z ciała K i $r(x)$ jest resztą z dzielenia wielomianu $\psi(x)$ przez wielomian $\varphi(x)$, to*

$$\psi(\mathbf{A}) = r(\mathbf{A}).$$

Dowód. Jeśli $q(x)$ jest ilorazem, a $r(x)$ resztą z dzielenia wielomianu $\psi(x)$ przez $\varphi(x)$, to $\psi(x) = q(x)\varphi(x) + r(x)$. Stąd i z równości $\varphi(\mathbf{A}) = \mathbf{0}$ mamy

$$\psi(\mathbf{A}) = q(\mathbf{A})\varphi(\mathbf{A}) + r(\mathbf{A}) = r(\mathbf{A}). \quad \square$$

Przykład 250. Dla macierzy $\mathbf{A}$ z poprzedniego przykładu wyznaczyć $\mathbf{A}^9 - 3\mathbf{A}^7$.

Różnica $\mathbf{A}^9 - 3\mathbf{A}^7$ jest wartością wielomianu $\psi(x) = x^9 - 3x^7$ dla macierzy $\mathbf{A}$, której wielomianem charakterystycznym jest $\varphi(x) = x^2 + 3x + 2$. Ponieważ resztą z dzielenia wielomianu $\psi(x)$ przez $\varphi(x)$ jest $r(x) = -126x - 124$, więc wobec poprzedniego twierdzenia mamy

$$\begin{aligned} \mathbf{A}^9 - 3\mathbf{A}^7 &= \psi(\mathbf{A}) = r(\mathbf{A}) = -126\,\mathbf{A} - 124\,\mathbf{I}_2 \\ &= -126 \begin{bmatrix} -2 & 0 \\ 7 & -1 \end{bmatrix} - 124 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 128 & 0 \\ -882 & 2 \end{bmatrix}. \end{aligned}$$

Twierdzenie Cayleya-Hamiltona i macierz odwrotna

Wyżej uzasadnialiśmy, że jeśli $\mathbf{A}$ jest macierzą wymiaru $n \times n$ i k jest liczbą naturalną, to macierz $\mathbf{A}^k$ jest kombinacją liniową macierzy $\mathbf{I}, \mathbf{A}, \ldots, \mathbf{A}^{n-1}$. Z twierdzenia Cayleya-Hamiltona wynika, że także macierz $\mathbf{A}^{-1}$ (i każda macierz $\mathbf{A}^{-k}$ dla $k \in N$) jest kombinacją liniową macierzy $\mathbf{I}, \mathbf{A}, \ldots, \mathbf{A}^{n-1}$.

Twierdzenie 10.7.4. *Jeśli* $\mathbf{A}$ *jest nieosobliwą macierzą wymiaru* $n \times n$ *i jeśli* $\varphi(\lambda) = (-1)^n\big(a_0 + a_1\lambda + \ldots + a_{n-1}\lambda^{n-1} + \lambda^n\big)$ *jest jej wielomianem charakterystycznym, to*

$$\mathbf{A}^{-1} = -\frac{1}{a_0}\big(a_1\mathbf{I} + a_2\mathbf{A} + \ldots + a_{n-1}\mathbf{A}^{n-2} + \mathbf{A}^{n-1}\big).$$

Dowód. Ponieważ wobec ostatniego wniosku mamy

$$a_0\mathbf{I} + a_1\mathbf{A} + \ldots + a_{n-1}\mathbf{A}^{n-1} + \mathbf{A}^n = \mathbf{0}$$

i ponieważ $a_0\ (= |\mathbf{A}|) \neq 0$ (bo macierz $\mathbf{A}$ jest nieosobliwa), więc także mamy

$$-\frac{1}{a_0}\left(a_1\mathbf{I} + a_2\mathbf{A} + \ldots + a_{n-1}\mathbf{A}^{n-2} + \mathbf{A}^{n-1}\right)\mathbf{A} = \mathbf{I}.$$

Stąd zaś wynika, że macierz $-\frac{1}{a_0}\left(a_1\mathbf{I} + a_2\mathbf{A} + \ldots + a_{n-1}\mathbf{A}^{n-2} + \mathbf{A}^{n-1}\right)$ jest macierzą odwrotną do macierzy $\mathbf{A}$.

Kolejny przykład ilustruje, że powyższy związek może być przydatny przy wyznaczaniu macierzy odwrotnej.

Przykład 251. Za pomocą twierdzenia Cayleya-Hamiltona wyznaczyć macierz odwrotną macierzy $\mathbf{A} = \begin{bmatrix} 0 & 2 & 3 \\ 2 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}$.

Macierz $\mathbf{A}$ jest nieosobliwa ($|\mathbf{A}| = -6$) i jej wielomianem charakterystycznym jest

$$\det\begin{bmatrix} -\lambda & 2 & 3 \\ 2 & -\lambda & 0 \\ 1 & -1 & -\lambda \end{bmatrix} = -\lambda^3 + 7\lambda - 6.$$

Stąd i z twierdzenia Cayleya-Hamiltona jest $-\mathbf{A}^3 + 7\mathbf{A} - 6\,\mathbf{I} = \mathbf{0}$, więc także $\frac{1}{6}(-\mathbf{A}^2 + 7\mathbf{I})\mathbf{A} = \mathbf{I}$ i dlatego mamy

$$\mathbf{A}^{-1} = \frac{1}{6}(-\mathbf{A}^2 + 7\,\mathbf{I}) = \frac{1}{6}\left(-\begin{bmatrix} 0 & 2 & 3 \\ 2 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}^2 + 7\,\mathbf{I}\right) = \frac{1}{6}\begin{bmatrix} 0 & 3 & 0 \\ 0 & 3 & -6 \\ 2 & -2 & 4 \end{bmatrix}.$$

10.8. Zależności rekurencyjne

Ciąg rekurencyjny jest ciągiem (x_n), w którym określone są początkowe wyrazy $x_0, x_1, \ldots, x_p$ (dla pewnej liczby naturalnej p) i dana jest reguła tworzenia wyrazu x_n $(n > p)$ za pomocą wyrazów $x_0, \ldots, x_{n-1}$. Tak rozumiane ciągi rekurencyjne pełnią ważną rolę w wielu działach matematyki i jej zastosowaniach. Tu zajmiemy się ciągiem (x_n), w którym znane są dwa pierwsze wyrazy x_0 i x_1, a pozostałe wyrazy określone są zależnością rekurencyjną

Zależność rekurencyjna

$$x_n = ax_{n-1} + bx_{n-2} \quad (n \geqslant 2), \tag{10.10}$$

gdzie a i b są liczbami rzeczywistymi takimi, że $a^2 + 4b \geqslant 0$. Pokażemy teraz, jak za pomocą wartości własnych i wektorów własnych macierzy można uzyskać jawny wzór na wyrazy ciągu (10.10).[2]

Zauważmy, że wobec (10.10) dla macierzy $\mathbf{A} = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix}$ i dla $n \geqslant 2$ mamy

$$\begin{aligned} \begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} &= \begin{bmatrix} ax_n + bx_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix} \\ &= \mathbf{A} \begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix} = \mathbf{A}^2 \begin{bmatrix} x_{n-1} \\ x_{n-2} \end{bmatrix} = \ldots = \mathbf{A}^n \begin{bmatrix} x_1 \\ x_0 \end{bmatrix}. \end{aligned}$$

Z równości $\begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} = \mathbf{A}^n \begin{bmatrix} x_1 \\ x_0 \end{bmatrix}$ możemy wyznaczyć x_n, ale wcześniej musimy obliczyć wartości iloczynu $\mathbf{A}^n \begin{bmatrix} x_1 \\ x_0 \end{bmatrix}$. Ponieważ a i b są liczbami rzeczywistymi takimi, że $a^2 + 4b \geqslant 0$, więc macierz $\mathbf{A}$ ma rzeczywiste wartości własne i przy wyliczeniach iloczynu $\mathbf{A}^n \begin{bmatrix} x_1 \\ x_0 \end{bmatrix}$ rozróżniamy dwa przypadki.

(a) Jeśli wektor $\begin{bmatrix} x_1 \\ x_0 \end{bmatrix}$ jest kombinacją liniową wektorów własnych $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ i $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ odpowiadających wartościom własnym λ_1 i λ_2 macierzy $\mathbf{A}$, powiedzmy $\begin{bmatrix} x_1 \\ x_0 \end{bmatrix} = \alpha\mathbf{v} + \beta\mathbf{u}$ dla pewnych liczb rzeczywistych α i β, to z równości

$$\begin{aligned} \begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} &= \mathbf{A}^n \begin{bmatrix} x_1 \\ x_0 \end{bmatrix} = \mathbf{A}^n(\alpha\mathbf{v} + \beta\mathbf{u}) = \alpha\mathbf{A}^n\mathbf{v} + \beta\mathbf{A}^n\mathbf{u} \\ &= \alpha\lambda_1^n\mathbf{v} + \beta\lambda_2^n\mathbf{u} = \alpha\lambda_1^n \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \beta\lambda_2^n \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \end{aligned}$$

otrzymujemy jawny wzór na wyrazy ciągu (10.10),

$$x_n = \alpha v_2 \lambda_1^{n-1} + \beta u_2 \lambda_2^{n-1}. \tag{10.11}$$

(b) Załóżmy teraz, że wektor $\begin{bmatrix} x_1 \\ x_0 \end{bmatrix}$ nie jest kombinacją liniową wektorów własnych macierzy $\mathbf{A}$. W tym przypadku macierz $\mathbf{A}$ ma podwójną wartość własną i jej wielomianem charakterystycznym jest $\varphi(x) = (x - \lambda_0)^2$ dla pewnego $\lambda_0 \in R$. Wtedy $x^2 - 2\lambda_0 x + \lambda_0^2 = x^2 - ax - b = \begin{vmatrix} a - x & b \\ 1 & -x \end{vmatrix}$ i dlatego $a = 2\lambda_0$ i $b = -\lambda_0^2$. Łatwo teraz zauważyć, że resztą z dzielenia wielomianu $\psi(x) = x^n$ przez wielomian $\varphi(x)$ jest $r(x) = n\lambda_0^{n-1}x - (n-1)\lambda_0^n$. Stąd i z twierdzenia 10.7.3 wynika, że

$$\begin{aligned} \mathbf{A}^n = \psi(\mathbf{A}) = r(\mathbf{A}) &= n\lambda_0^{n-1}\mathbf{A} - (n-1)\lambda_0^n\mathbf{I} \\ &= \lambda_0^{n-1} \begin{bmatrix} (n+1)\lambda_0 & -n\lambda_0^2 \\ n & -(n-1)\lambda_0 \end{bmatrix}. \end{aligned}$$

[2] W matematyce wypracowano wiele szczegółowych metod rozwiązywania rekurencji, czyli uzyskiwania jawnych wzorów na wyrazy ciągów określonych rekurencyjnie. Przykładowo, elementarnie można uzasadnić, że jeśli λ_1 i λ_2 są różnymi pierwiastkami równania kwadratowego $x^2 - ax - b = 0$, to n-ty wyraz ciągu (10.10) określony jest wzorem

$$x_n = c_1\lambda_1^n + c_2\lambda_2^n \quad (n \geqslant 0),$$

gdzie stałe c_1 i c_2 są takie, że $c_1 + c_2 = x_0$ i $c_1\lambda_1 + c_2\lambda_2 = x_1$.

Jeśli zaś λ_0 jest podwójnym pierwiastkiem równania $x^2 - ax - b = 0$, to

$$x_n = c_1\lambda_0^n + c_2 n\lambda_0^n \quad (n \geqslant 0),$$

gdzie stałe c_1 i c_2 są tak dobrane, że $c_1 = x_0$ i $c_1\lambda_0 + c_2\lambda_2 = x_1$.

Dlatego z równości $\begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} = \mathbf{A}^n \begin{bmatrix} x_1 \\ x_0 \end{bmatrix}$ dla $n \geqslant 0$ mamy

$$x_n = n\lambda_0^{n-1}x_1 - (n-1)\lambda_0^n x_0. \tag{10.12}$$

Nasze rozważania powtarzamy dla dwóch konkretnych ciągów rekurencyjnych.

Ciąg Fibonacciego: $0, 1, 1, 2, 3, 5, 8, 13, \ldots$

Przykład 252. Wyznaczyć jawny wzór na wyrazy ciągu Fibonacciego, czyli ciągu (F_n), w którym $F_0 = 0$, $F_1 = 1$ i dla $n \geqslant 2$ jest

$$F_n = F_{n-1} + F_{n-2}. \tag{10.13}$$

Ciąg (10.13) jest jednoznacznie określony przez macierz $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. Ponieważ macierz ta jest diagonalizowalna, wyznaczymy $\begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = \mathbf{A}^n \begin{bmatrix} F_1 \\ F_0 \end{bmatrix}$, przedstawiając wektor $\begin{bmatrix} F_1 \\ F_0 \end{bmatrix}$ jako kombinację liniową wektorów własnych macierzy $\mathbf{A}$.

Dla macierzy $\mathbf{A}$ mamy

$$|\mathbf{A} - \lambda\mathbf{I}| = \lambda^2 - \lambda - 1 = \left(\lambda - \frac{1+\sqrt{5}}{2}\right)\left(\lambda - \frac{1-\sqrt{5}}{2}\right),$$

więc jej wartościami własnymi są liczby

$$\lambda_1 = \frac{1+\sqrt{5}}{2} \quad \text{i} \quad \lambda_2 = \frac{1-\sqrt{5}}{2}.$$

Ponieważ

$$\left[\mathbf{A} - \lambda_1\mathbf{I} \,|\, \mathbf{0}\right] \sim \left[\begin{array}{cc|c} 1-\sqrt{5} & 2 & 0 \\ 0 & 0 & 0 \end{array}\right] \quad \text{i} \quad \left[\mathbf{A} - \lambda_2\mathbf{I} \,|\, \mathbf{0}\right] \sim \left[\begin{array}{cc|c} 1+\sqrt{5} & 2 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

więc

$$\mathbf{v} = \begin{bmatrix} 2 \\ \sqrt{5}-1 \end{bmatrix} \quad \text{i} \quad \mathbf{u} = \begin{bmatrix} -2 \\ \sqrt{5}+1 \end{bmatrix}$$

są wektorami własnymi odpowiadającymi wartościom własnym λ_1 i λ_2. Zauważmy teraz, że wektor $\begin{bmatrix} F_1 \\ F_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ jest kombinacją liniową wektorów $\mathbf{v}$ i $\mathbf{u}$,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \alpha\mathbf{v} + \beta\mathbf{u}, \quad \text{gdzie} \quad \alpha = \frac{1+\sqrt{5}}{4\sqrt{5}} \quad \text{i} \quad \beta = \frac{1-\sqrt{5}}{4\sqrt{5}}.$$

Zatem

$$\begin{aligned} \begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} &= \mathbf{A}^n \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \mathbf{A}^n(\alpha\mathbf{v} + \beta\mathbf{u}) = \alpha\mathbf{A}^n\mathbf{v} + \beta\mathbf{A}^n\mathbf{u} = \alpha\lambda_1^n\mathbf{v} + \beta\lambda_2^n\mathbf{u} \\ &= \frac{1+\sqrt{5}}{4\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n \begin{bmatrix} 2 \\ \sqrt{5}-1 \end{bmatrix} + \frac{1-\sqrt{5}}{4\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n \begin{bmatrix} -2 \\ \sqrt{5}+1 \end{bmatrix}. \end{aligned}$$

Stąd składowa F_n wektora $\begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix}$, czyli n-ta liczba Fibonacciego, określona jest wzorem

$$F_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n. \tag{10.14}$$

Korzystając ze wzoru (10.14), możemy obliczyć każdą liczbę Fibonacciego F_n bez uprzedniego obliczania liczb $F_2, \ldots, F_{n-1}$. Okazuje się, że sam proces obliczania liczby F_n można jeszcze uprościć. Przede wszystkim można zaobserwować, że dla każdej liczby naturalnej n jest

$$-\frac{1}{2} < \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n < \frac{1}{2}.$$

Stąd zaś i z faktu, że każda liczba Fibonacciego jest liczbą całkowitą wynika, że F_n jest liczbą całkowitą najbliższą liczbie $\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n$, czyli

$$F_n = \left\lfloor \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n \right\rceil. \tag{10.15}$$

W tym miejscu warto zastanowić się, czy łatwiej oblicza się liczbę F_n iteracyjnie (znajdując najpierw $F_2, \ldots, F_{n-1}$), czy posługując się formułą (10.14) lub (10.15). Czytelnikowi proponujemy iteracyjne obliczenie kilku liczb F_n, a następnie obliczenie na kalkulatorze tych samych liczb zz pomocą wzoru (10.14) lub (10.15).

Przykład 253. Wyznaczyć jawny wzór na wyrazy ciągu (x_n), w którym $x_0 = 1$, $x_1 = -3$ i $x_n = 6x_{n-1} - 9x_{n-2}$ dla $n \geqslant 2$.

Ciągowi (x_n) odpowiada macierz $\mathbf{A} = \begin{bmatrix} 6 & -9 \\ 1 & 0 \end{bmatrix}$, której wielomianem charakterystycznym jest $\varphi(x) = x^2 - 6x + 9 = (x-3)^2$. Tym razem macierz $\mathbf{A}$ nie jest diagonalizowalna i wyznaczając $\begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} = \mathbf{A}^n \begin{bmatrix} x_1 \\ x_0 \end{bmatrix}$, wyznaczymy najpierw $\mathbf{A}^n$. Ponieważ $r(x) = 3^{n-1}nx - 3^n(n-1)$ jest resztą z dzielenia wielomianu $\psi(x) = x^n$ przez $\varphi(x)$, więc wobec twierdzenia 10.7.3 mamy

$$\mathbf{A}^n = \psi(\mathbf{A}) = r(\mathbf{A}) = 3^{n-1}n\mathbf{A} - 3^n(n-1)\mathbf{I} = 3^{n-1}\begin{bmatrix} 3n+3 & -9n \\ n & 3-3n \end{bmatrix}.$$

Stąd

$$\begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} = \mathbf{A}^n \begin{bmatrix} x_1 \\ x_0 \end{bmatrix} = 3^{n-1}\begin{bmatrix} 3n+3 & -9n \\ n & 3-3n \end{bmatrix}\begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

i dlatego

$$x_n = 3^n - 2n3^n \quad (n \geqslant 0).$$

10.9. Ćwiczenia

1. Zbadać, czy liczba λ jest wartością własną macierzy $\mathbf{A}$, gdy:

   (a) $\lambda = -2$, $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 18 & 4 \end{bmatrix}$;

   (b) $\lambda = 3$, $\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}$.

2. Zbadać, czy wektor $\mathbf{v}$ jest wektorem własnym macierzy $\mathbf{A}$, gdy:

   (a) $\mathbf{v} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ i $\mathbf{A} = \begin{bmatrix} 7 & 3 \\ 3 & -1 \end{bmatrix}$;

   (b) $\mathbf{v} = \begin{bmatrix} -3 \\ 5 \\ 1 \end{bmatrix}$ i $\mathbf{A} = \begin{bmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$.

3. Wyznaczyć $\operatorname{tr}(\mathbf{A})$, $\det(\mathbf{A})$, $\mathbf{A}^{-1}$, wartości własne i wektory własne macierzy $\mathbf{A}$, gdy:

   (a) $\mathbf{A} = \begin{bmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{bmatrix}$; (b) $\mathbf{A} = \begin{bmatrix} -2 & -1 \\ 5 & 2 \end{bmatrix}$;

   (c) $\mathbf{A} = \begin{bmatrix} -3 & 2 & 4 \\ 2 & -6 & 2 \\ 4 & 2 & -3 \end{bmatrix}$; (d) $\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$.

4. Wyznaczyć wielomian charakterystyczny, wartości własne i wektory własne następujących macierzy:

   $\begin{bmatrix} 3 & 0 \\ 7 & -1 \end{bmatrix}$, $\begin{bmatrix} 2 & 7 \\ 7 & 2 \end{bmatrix}$, $\begin{bmatrix} 3 & -2 \\ 1 & -1 \end{bmatrix}$, $\begin{bmatrix} 5 & -3 \\ -4 & 3 \end{bmatrix}$,

   $\begin{bmatrix} 10 & -9 \\ 4 & -2 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$, $\begin{bmatrix} 2 & 4 \\ 5 & 3 \end{bmatrix}$, $\begin{bmatrix} -1 & 2j \\ -2j & 2 \end{bmatrix}$,

   $\begin{bmatrix} 3 & 4 & 4 \\ -1 & -2 & -4 \\ 1 & 1 & 3 \end{bmatrix}$, $\begin{bmatrix} 2 & -2 & 3 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$,

   $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \\ -1 & 2 & 3 \end{bmatrix}$, $\begin{bmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{bmatrix}$, $\begin{bmatrix} 4 & 9 & 0 \\ 0 & -2 & 8 \\ 0 & 0 & 7 \end{bmatrix}$.

5. Wyznaczyć wielomian charakterystyczny, wartości własne i wektory własne oraz stwierdzić, czy istnieje baza przestrzeni R^3 składająca się z wektorów własnych każdej z następujących macierzy:
$$\begin{bmatrix} 2 & -2 & 3 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{bmatrix}, \begin{bmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ -4 & -4 & 4 \end{bmatrix}, \begin{bmatrix} 2 & -2 & 3 \\ 10 & -4 & 5 \\ 5 & -4 & 6 \end{bmatrix},$$
$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 3 & -1 \\ -3 & 5 & -1 \\ -3 & 3 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 0 & 4 \\ 0 & 6 & 0 \\ 4 & 0 & 2 \end{bmatrix}.$$

6. Wyznaczyć (jeśli to możliwe) macierz $\mathbf{P}$ taką, że macierz $\mathbf{P}^{-1}\mathbf{AP}$ jest diagonalna, gdy $\mathbf{A}$ jest jedną z następujących macierzy:
$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & -3 \\ 4 & -5 \end{bmatrix}, \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}, \begin{bmatrix} 13 & 4 \\ 4 & 7 \end{bmatrix}, \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix},$$
$$\begin{bmatrix} -3 & 4 \\ 2 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & j \\ -j & 0 \end{bmatrix}, \begin{bmatrix} 0 & j \\ j & 0 \end{bmatrix},$$
$$\begin{bmatrix} 2 & -2 & 2 \\ 0 & 1 & 1 \\ -4 & 8 & 3 \end{bmatrix}, \begin{bmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}, \begin{bmatrix} 5 & 2 & -2 \\ 2 & 5 & -2 \\ -2 & -2 & 5 \end{bmatrix},$$
$$\begin{bmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & 6 & 6 \end{bmatrix}, \begin{bmatrix} 1 & -3 & -3 \\ -8 & 6 & -3 \\ 8 & -2 & 7 \end{bmatrix}, \begin{bmatrix} 1 & 1 & -1 \\ -1 & -1 & 2 \\ -5 & -2 & 3 \end{bmatrix}.$$

7. Niech $\mathbf{A}$ będzie macierzą wymiaru 3×3 z wartościami własnymi $\lambda_1 = 0$, $\lambda_2 = 2$ i $\lambda_3 = c \in R$ i wektorami własnymi $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, i $\mathbf{x}_3 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$. (*a*) Dla jakich c macierz $\mathbf{A}$ jest diagonalizowalna? (*b*) Dla jakich c macierz $\mathbf{A}$ jest symetryczna?

8. (*a*) Obliczyć wartości własne i wektory własne macierzy $\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$. (*b*) Czy macierz $\mathbf{A}$ jest diagonalizowalna? (*c*) Wyznaczyć rząd i wyznacznik macierzy $\mathbf{A} + 2\mathbf{I}$.

9. Zbadać diagonalizowalność operatora liniowego $T\colon V \to V$ i, tam gdzie jest to możliwe, wskazać bazę B przestrzeni V taką, że macierz $[T]_B$ jest diagonalna, gdy:

   (*a*) $T\colon R^3 \to R^3$, $T(x, y, z) = (y, x, 3z)$;

   (*b*) $T\colon R_3[x] \to R_3[x]$, $T(\varphi(x)) = \varphi'(x) + \varphi''(x)$;

   (*c*) $T\colon R_2[x] \to R_2[x]$, $T(ax^2 + bx + c) = cx^2 + bx + a$;

   (*d*) $T\colon \mathcal{C}^2 \to \mathcal{C}^2$, $T(z, w) = (jz + w, z + jw)$;

   (*e*) $T\colon R_{2\times 2} \to R_{2\times 2}$, $T(\mathbf{A}) = \mathbf{A}^T$;

   (*f*) $T\colon R_{2\times 2} \to R_{2\times 2}$,
$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 2c & a+c \\ b-2c & d \end{bmatrix}.$$

10. Wyznaczyć macierz $\mathbf{P}$ taką, że macierz $\mathbf{P}^{-1}\mathbf{AP}$ jest diagonalna i, dodatkowo, wyznaczyć $\mathbf{A}^n$, gdy:

    (*a*) $\mathbf{A} = \begin{bmatrix} 3 & -4 \\ 2 & -3 \end{bmatrix}$; (*b*) $\mathbf{A} = \begin{bmatrix} 3 & 4 \\ -1 & -2 \end{bmatrix}$;

    (*c*) $\mathbf{A} = \begin{bmatrix} -3 & 1 \\ -4 & 2 \end{bmatrix}$; (*d*) $\mathbf{A} = \begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix}$.

11. Niech $T\colon R^2 \to R^2$ będzie przekształceniem liniowym takim, że $T(x, y) = (x + 2y, 3x + 2y)$ i niech $E = \big((1, 0), (0, 1)\big)$ oraz $B = \big((1, 3), (2, 5)\big)$ będą bazami przestrzeni R^2. (*a*) Wyznaczyć macierze $[T]_E$ i $[T]_B$. (*b*) Znaleźć $[\mathbf{v}]_B$ i $[T(\mathbf{v})]_B$, gdy $\mathbf{v} = (1, 1)$. (*c*) Wyznaczyć macierze przejścia $[I_{R^2}]_B^E$ i $[I_{R^2}]_E^B$. (*d*) Znaleźć bazę C przestrzeni R^2 taką, że macierz $[T]_C$ jest diagonalna. (*e*) Wskazać macierz $\mathbf{P} = [I_{R^2}]_E^C$ i $\mathbf{P}^{-1}$. (*f*) Obliczyć $\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}^n$ dla $n \in N$.

12. Wyznaczyć wartości własne i wektory własne macierzy $\mathbf{A} = \begin{bmatrix} 25 & 40 \\ -12 & -19 \end{bmatrix}$. Następnie wektor $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ przedstawić jako kombinację liniową wektorów własnych macierzy $\mathbf{A}$ i obliczyć $\mathbf{A}^{10}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

13. Wyznaczyć wartości własne i wektory własne macierzy $\mathbf{A} = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$. Następnie wektor $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ przedstawić jako kombinację liniową wektorów własnych macierzy $\mathbf{A}$ i obliczyć $\mathbf{A}^5\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ oraz $\mathbf{A}^n\begin{bmatrix} 2 \\ -1 \end{bmatrix}$.

14. Wyznaczyć $\mathbf{P}^{-1}\mathbf{AP}$ i $\mathbf{A}^n - \frac{1}{7}\left(\frac{5}{12}\right)^n \begin{bmatrix} 4 & -3 \\ -4 & 3 \end{bmatrix}$, gdy
$$\mathbf{A} = \begin{bmatrix} 2/3 & 1/4 \\ 1/3 & 3/4 \end{bmatrix} \quad \text{i} \quad \mathbf{P} = \begin{bmatrix} 1 & 3 \\ -1 & 4 \end{bmatrix}.$$

15. Wykazać, że $\lim\limits_{n\to\infty} \begin{bmatrix} 3/5 & 4/5 \\ 2/5 & 1/5 \end{bmatrix}^n = \begin{bmatrix} 2/3 & 2/3 \\ 1/3 & 1/3 \end{bmatrix}$.

16. Wyznaczyć (1) wartości własne i wektory własne macierzy $\mathbf{A}$, (2) macierz $\mathbf{P}$ taką, że $\mathbf{P}^{-1}\mathbf{AP}$ jest macierzą diagonalną i (3) $\lim\limits_{n\to\infty} \mathbf{A}^n$, gdy:

    (*a*) $\mathbf{A} = \begin{bmatrix} 1/2 & 3 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; (*b*) $\mathbf{A} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 1/2 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}$.

17. Sprawdzić, czy podprzestrzeń W jest T-niezmienniczą podprzestrzenią przestrzeni V, gdy:

    (*a*) $V = R_4[x]$, $T\big(\varphi(x)\big) = \varphi'(x)$ i $W = R_2[x]$;

    (*b*) $V = R^3$, $T(x, y, z) = (x + 3y + 3z, -3x - 5y - 3z, 3x + 3y + z)$ i $W = \{(t, -t, t)\colon t \in R\}$;

    (*c*) $V = R_{2\times 2}$, $T(\mathbf{A}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\mathbf{A}$ i $W = \{\mathbf{A}\colon \mathbf{A}^T = \mathbf{A}\}$.

18. Wyznaczyć bazę T-cyklicznej podprzestrzeni generowanej przez wektor $\mathbf{v}$, gdy:

    (*a*) $V = R_4[x]$, $T\big(\varphi(x)\big) = \varphi'(x)$, $\mathbf{v} = x^3$;

    (*b*) $V = R^3$, $T(x, y, z) = (x + 3y + 3z, -3x - 5y - 3z, 3x + 3y + z)$, $\mathbf{v} = (1, 0, 0)$;

    (*c*) $V = R_{2\times 2}$, $T(\mathbf{A}) = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}\mathbf{A}^T$, $\mathbf{v} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

19. Niech V będzie przestrzenią wektorową (nad ciałem liczb zespolnych) z iloczynem skalarnym. Niech $T\colon V \to V$ będzie samosprzężonym operatorem linio-

wym na przestrzeni V, czyli takim, że $(T(\mathbf{u})|\mathbf{v}) = (\mathbf{u}|T(\mathbf{v}))$ dla każdych wektorów $\mathbf{u}, \mathbf{v} \in V$. Niech $\mathbf{v}_1$ i $\mathbf{v}_2$ będą wektorami własnymi operatora T odpowiadającymi wartościom własnym λ_1 i λ_2. Udowodnić, że λ_1 i λ_2 są rzeczywiste. Wykazać, że jeśli $\lambda_1 \neq \lambda_2$, to $\mathbf{v}_1$ i $\mathbf{v}_2$ są ortogonalne.

20. Niech V będzie rzeczywistą przestrzenią wektorową z iloczynem skalarnym. Napisać, co to znaczy, że operator liniowy $T : V \to V$ jest ortogonalny. Pokazać, że jeśli $\lambda \in R$ jest wartością własną operatora T, to $\lambda = 1$ albo $\lambda = -1$.

21. Czy macierz $\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$ jest ortogonalna? Wyznaczyć jej wartości własne.

22. Wyznaczyć wartości własne i przestrzeń zerową macierzy $\mathbf{A} = \begin{bmatrix} 0 & a & -b \\ -a & 0 & c \\ b & -c & 0 \end{bmatrix}$. Wyjaśnić dlaczego macierz $\mathbf{A}$ nie jest ortogonalna.

23. Wyznaczyć wartości własne i wektory własne macierzy $\mathbf{A}$ oraz utworzyć macierz ortogonalną $\mathbf{Q}$ taką, że $\mathbf{Q}^T\mathbf{A}\mathbf{Q}$ jest macierzą diagonalną, gdy:

(a) $\mathbf{A} = \begin{bmatrix} 3 & 4 \\ 4 & 3 \end{bmatrix}$; (b) $\mathbf{A} = \begin{bmatrix} 1 & 3 & 3 & -3 \\ 3 & 1 & -3 & 3 \\ 3 & -3 & 1 & 3 \\ -3 & 3 & 3 & 1 \end{bmatrix}$;

(c) $\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$; (d) $\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 3 & 1 \end{bmatrix}$.

24. Znaleźć wszystkie podprzestrzenie niezmiennicze operatora $T\colon R^3 \to R^3$, gdy
$$T(\mathbf{x}) = \begin{bmatrix} 1 & -4 & 1 \\ -4 & 1 & 1 \\ 4 & 4 & 4 \end{bmatrix} \mathbf{x}.$$

25. Napisać twierdzenie Cayleya-Hamiltona. Wskazać macierz wymiaru 6×6, której wielomianem charakterystycznym jest $\lambda^6 - 17\lambda^5 + 2\lambda^3 + \lambda^2 - 5$.

26. Niech $\mathbf{A}$ będzie macierzą wymiaru 4×4 i niech $\lambda_1 = 1$, $\lambda_2 = -1$, $\lambda_3 = \lambda_4 = 2$ będą wartościami własnymi macierzy $\mathbf{A}$. (a) Wyznaczyć $\mathbf{A}^{-1}$ za pomocą nieujemnych potęg macierzy $\mathbf{A}$. (b) Pokazać, że $\mathbf{A}^6 = 36\mathbf{A}^3 - 51\mathbf{A}^2 - 36\mathbf{A} + 52\mathbf{I}$.

27. Wyznaczyć wartości własne i wektory własne $\mathbf{v}_1$, $\mathbf{v}_2$ i $\mathbf{v}_3$ macierzy $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ -2 & 2 & 3 \end{bmatrix}$. Dla macierzy $\mathbf{B} = \begin{bmatrix} \mathbf{v}_1 \, \mathbf{v}_2 \, \mathbf{v}_3 \end{bmatrix}$ znaleźć liczby α, β i γ takie, że $\mathbf{B}^3 + \alpha\mathbf{B}^2 + \beta\mathbf{B} + \gamma\mathbf{I} = \mathbf{0}$.

28. Operator liniowy $T\colon R^3 \to R^3$ określony jest wzorem $T(x, y, z) = (2x+4y+3z, -4x-6y-3z, 3x+3y+z)$. (a) Wyznaczyć wymiar T-cyklicznej podprzestrzeni $W \subseteq R^3$ generowanej przez wektor $\mathbf{v} = (1, 0, 0)$. (b) Wyznaczyć liczby $a_0, \ldots, a_{k-1}$ ($k = \dim W$) takie, że $a_0\mathbf{v} + a_1T(\mathbf{v}) + \ldots + a_{k-1}T^{k-1}(\mathbf{v}) + T^k(\mathbf{v}) = \mathbf{0}$. (c) Korzystając z (b) wyznaczyć wielomian charakterystyczny $\varphi(\lambda)$ operatora T. (d) Korzystając z równości $\varphi(T) = \mathbf{0}$, wyznaczyć T^{-1}. (e) Obliczyć $T^7 + 3T^6 - 3T^4 + 3T^3 + 5T + I_{R^3}$.

29. (a) Wyznaczyć wartości własne i wektory własne macierzy $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ -\frac{1}{6} & \frac{1}{6} \end{bmatrix}$. (b) Dany jest ciąg wektorów $(\mathbf{x}_k)$ taki, że $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k$ i $\mathbf{x}_{100} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Wyznaczyć $\mathbf{x}_0$. (c) Niech $\mathbf{B}$ będzie dowolna macierzą wymiaru 2×2. Wyjaśnić dlaczego macierze $\mathbf{AB}$ i $\mathbf{BA}$ mają identyczne wartości własne.

30. Niech d_n będzie wyznacznikiem macierzy $\mathbf{A}_n = [a_{ij}]$, której współczynniki określone są wzorem
$$a_{ij} = \begin{cases} -2, & i = j-1, \\ 1, & i \in \{j, j+1\}, \\ 0, & i \notin \{j-1, j, j+1\}. \end{cases}$$
Przykładowo: $\mathbf{A}_1 = [1]$, $\mathbf{A}_2 = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}$,
$$\mathbf{A}_3 = \begin{bmatrix} 1 & -2 & 0 \\ 1 & 1 & -2 \\ 0 & 1 & 1 \end{bmatrix}, \mathbf{A}_4 = \begin{bmatrix} 1 & -2 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$
(a) Obliczyć d_4. (b) Wyznaczyć liczby a i b takie, że $d_n = ad_{n-1} + bd_{n-2}$. (c) Wskazać macierz $\mathbf{A}$ taką, że $\begin{bmatrix} d_{n+1} \\ d_n \end{bmatrix} = \mathbf{A} \begin{bmatrix} d_n \\ d_{n-1} \end{bmatrix}$ i obliczyć wartości własne i wektory własne wskazanej macierzy $\mathbf{A}$. (d) Wyznaczyć liczbę λ taką, że $\lim\limits_{n\to\infty} \dfrac{d_n}{\lambda^n} = 0$.

31. Dany jest ciąg (x_n), w którym $x_0 = 0$, $x_1 = 1$ i $x_{n+2} = (x_n + x_{n+1})/2$. (a) Wskazać taką macierz $\mathbf{A}$, dla której jest $\begin{bmatrix} x_{n+2} \\ x_{n+1} \end{bmatrix} = \mathbf{A} \begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix}$. (b) Wyznaczyć wartości własne i wektory własne macierzy $\mathbf{A}$. (c) Wyznaczyć macierz diagonalną $\mathbf{\Lambda}$ i macierz odwracalną $\mathbf{P}$ taką, że $\mathbf{A} = \mathbf{P\Lambda P}^{-1}$. (d) Wyznaczyć $\lim_{n\to\infty} x_n$.

32. (a) Ciąg liczbowy x_0, x_1, $x_2, \ldots$ spełnia zależność rekurencyjną $x_{n+2} = 2x_{n+1} + 3x_n$. Wskazać macierz $\mathbf{A}$ taką, że $\begin{bmatrix} x_{n+2} \\ x_{n+1} \end{bmatrix} = \mathbf{A} \begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix}$. (b) Wyznaczyć wartości własne i wektory własne macierzy $\mathbf{A}$. (c) Dla $x_0 = 2$ wskazać taką wartość x_1, że ciąg (x_n) jest ograniczony.

33. Rozwiązać układ zależności rekurencyjnych
$$\begin{cases} x_{n+1} = 3x_n - y_n, \\ y_{n+1} = -x_n + 3y_n, \end{cases}$$
gdy $x_0 = 1$ i $y_0 = 2$.

34. Rozwiązać układ zależności rekurencyjnych
$$\begin{cases} x_{n+1} = 2x_n - y_n - 1, \\ y_{n+1} = -x_n + 2y_n + 2, \end{cases}$$
gdy $x_0 = 0$ i $y_0 = -1$.

35. Zakładamy, że populacja Polski jest ustalona i stała. Co roku 1/20 liczby ludności wiejskiej przenosi się do miast i 1/10 liczby ludności miejskiej opuszcza miasto. Co po wielu latach stanie się z ludnością miejską?
36. Niech T będzie operatorem liniowym na przestrzeni wektorowej V (nad ciałem K) i niech W będzie T-niezmienniczą podprzestrzenią przestrzeni V. Udowodnić, że W jest $\varphi(T)$-niezmienniczą podprzestrzenią dla każdego wielomianu $\varphi(x) \in K[x]$.
37. Niech $\mathbf{A}$ będzie rzeczywistą macierzą wymiaru $n \times n$ i niech V_λ oznacza zbiór $\{\mathbf{x} \in R^n\colon\ \mathbf{Ax} = \lambda\mathbf{x}\}$ dla liczby rzeczywistej λ. Pokazać, że V_λ jest podprzestrzenią przestrzeni R^n. Wyznaczyć wymiar przestrzeni V_λ, gdy λ nie jest wartością własną macierzy $\mathbf{A}$.
38. Wykazać, że macierz kwadratowa $\mathbf{A}$ i jej transpozycja $\mathbf{A}^T$ mają identyczne wielomiany charakterystyczne.
39. Udowodnić, że macierze podobne mają identyczne wielomiany charakterystyczne, te same wartości własne i to samo widmo.
40. Niech $\mathbf{A}$ będzie macierzą diagonalizowalną. Za pomocą macierzy podobnych pokazać, że ślad macierzy $\mathbf{A}$ jest równy sumie jej wartości własnych.
41. Niech $\mathbf{A}$ będzie macierzą kwadratową stopnia n i niech $\lambda_1, \ldots, \lambda_n$ będą jej wartościami własnymi. Pokazać, że ślad macierzy $\mathbf{A}$ jest równy sumie jej wartości własnych, tj. $\operatorname{tr}(\mathbf{A}) = \lambda_1 + \ldots + \lambda_n$.
42. Niech $\mathbf{A}$ będzie diagonalizowalną macierzą wymiaru $n \times n$ i niech $\lambda_1, \ldots, \lambda_n$ będą wartościami własnymi macierzy $\mathbf{A}$. Udowodnić, że $\det(\mathbf{A}) = \lambda_1\lambda_2 \ldots \lambda_n$.
43. Udowodnić, że jeśli $\mathbf{A}$ jest macierzą kwadratową i sumą elementów każdego jej wiersza jest liczba a, to a jest pierwiastkiem wielomianu charakterystycznego macierzy $\mathbf{A}$.
44. Niech T będzie operatorem liniowym w przestrzeni wektorowej V nad ciałem K i niech $\mathbf{x}_0$ będzie wektorem własnym operatora T odpowiadającym wartości własnej λ. (a) Udowodnić, że $\mathbf{x}_0$ jest wektorem własnym operatora T^n (dla każdej liczby naturalnej n). (b) Wykazać, że $\mathbf{x}_0$ jest wektorem własnym operatora $2T^2 + I_V$ odpowiadającym wartości własnej $2\lambda^2 + 1$. (c) Niech $\psi(t)$ będzie wielomianem o współczynnikach z ciała K. Wykazać, że $\mathbf{x}_0$ jest wektorem własnym operatora $\psi(T)$ odpowiadającym wartości własnej $\psi(\lambda)$.
45. Pokazać, że jeśli λ jest wartością własną macierzy $\mathbf{A}$, to λ^n jest wartością własną macierzy $\mathbf{A}^n$ dla każdej liczby naturalnej n.
46. Liczby $\lambda = -1$ i $\lambda = 1$ są wartościami własnymi macierzy $\mathbf{A}$ wymiaru 3×3 takiej, że $\mathbf{A} + \mathbf{I}$ jest macierzą rzędu jeden. Która wartość własna macierzy $\mathbf{A}$ jest wielokrotna? Dlaczego? Czy macierz $\mathbf{A}$ jest diagonalizowalna?
47. Dana jest macierz $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & x & 4 \\ 3 & 4 & 5 \end{bmatrix}$. Wyznaczyć (jeśli to możliwe) te wartości parametru x, dla których wszystkie wartości własne macierzy $\mathbf{A}$ są dodatnie.
48. Wykazać, że jeśli $\mathbf{A}^2$ jest macierzą zerową, to zero jest jedyną wartością własną macierzy $\mathbf{A}$.
49. Macierz $\mathbf{A}$ jest nilpotentna, gdy $\mathbf{A}^k = \mathbf{0}$ dla pewnego naturalnego k. Wykazać, że zero jest jedyną wartością własną nilpotentnej macierzy $\mathbf{A}$.
50. Wykazać, że macierz $\mathbf{A} \in R_{n\times n}$ jest osobliwa wtedy i tylko wtedy, gdy 0 jest jej wartością własną.
51. Udowodnić, że operator liniowy T na skończenie wymiarowej przestrzeni wektorowej V jest odwracalny wtedy i tylko wtedy, gdy zero nie jest wartością własną operatora T.
52. Niech T będzie odwracalnym operatorem liniowym. Udowodnić, że skalar λ jest wartością własną operatora T wtedy i tylko wtedy, gdy λ^{-1} jest wartością własną operatora T^{-1}.
53. Niech $\varphi(\lambda) = (-1)^n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0$ będzie wielomianem charakterystycznym macierzy $\mathbf{A}$. (a) Udowodnić, że $\varphi(0) = a_0 = \det(\mathbf{A})$. Wywnioskować stąd, że macierz $\mathbf{A}$ jest odwracalna wtedy i tylko wtedy, gdy $a_0 \neq 0$. (b) Udowodnić, że $\operatorname{tr}(\mathbf{A}) = (-1)^{n-1}a_{n-1}$.
54. Wykazać, że jeśli λ jest wartością własną odwracalnej macierzy $\mathbf{A}$, to $1/\lambda$ jest wartością własną macierzy $\mathbf{A}^{-1}$.
55. Wykazać, że jeśli λ jest wartością własną macierzy ortogonalnej $\mathbf{A}$, to także $1/\lambda$ jest wartością własną macierzy $\mathbf{A}$.
56. Niech T będzie operatorem liniowym w n-wymiarowej przestrzeni wektorowej V nad ciałem K. Niech B będzie bazą przestrzeni V. Udowodnić, że wektor $\mathbf{v} \in V$ jest wektorem własnym operatora T odpowiadającym wartości własnej λ wtedy i tylko wtedy, gdy wektor $[\mathbf{v}]_B \in K^n$ jest wektorem własnym macierzy $[T]_B$ odpowiadającym wartości własnej λ.
57. Niech $\mathbf{A}$ będzie rzeczywistą macierzą skośnie-symetryczną (tj. taką, że $\mathbf{A}^T = -\mathbf{A}$). Wykazać, że każda wartość własna λ macierzy $\mathbf{A}$ jest liczbą ściśle urojoną. (W tym celu pokazać, że $\overline{\lambda} = -\lambda$.)
58. Macierz $\mathbf{A}$ jest podobna do macierzy $\mathbf{B}$ wymiaru 3×3, której wartościami własnymi są 1, 1 i 2. Co można powiedzieć o: (a) wartościach własnych macierzy $\mathbf{A}$; (b) diagonalizowalności macierzy $\mathbf{A}$; (c) symetryczności $\mathbf{A}$; (d) wyznaczniku macierzy $\mathbf{A}$?
59. Niech $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ będą wektorami własnymi macierzy $\mathbf{A} \in R_{3\times3}$ odpowiadającymi jej wartościom własnym 0, 1 i 2. (a) Opisać przestrzeń zerową, przestrzeń kolumnową i przestrzeń wierszową macierzy $\mathbf{A}$ za pomocą $\mathbf{u}$, $\mathbf{v}$ i $\mathbf{w}$. (b) Znaleźć wszystkie rozwiązania równania $\mathbf{Ax} = \mathbf{v} - \mathbf{w}$.
60. Wartościami własnymi macierzy $\mathbf{A}$ są 0, 1 i 2. Wyznaczyć: (a) rząd macierzy $\mathbf{A}$; (b) wyznacznik macierzy $\mathbf{A}^T\mathbf{A}$; (c) wyznacznik macierzy $\mathbf{A}+\mathbf{I}$; (d) wartości własne macierzy $(\mathbf{A} + \mathbf{I})^{-1}$.
61. Wpisując TAK albo NIE, stwierdzić prawdziwość każdego z następujących zdań:

$\boxed{1\quad}$ Jeśli $T\colon V \to V$ jest przekształceniem liniowym i jego wielomian charakterystyczny jest iloczynem różnych składników stopnia pierwszego, to T jest diagonalizowalne.

$\boxed{2\quad}$ Każde przekształcenie liniowe $T\colon V \to V$ ma co najmniej jedną wartość własną.

$\boxed{3}$ Macierz $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix}$ jest diagonalizowalna, a macierz $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ nie jest diagonalizowalna.

$\boxed{4}$ Jeśli dla macierzy $\mathbf{A}$, $\mathbf{B}$ i $\mathbf{Q}$ jest $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{B}$, to $\mathbf{A}$ i $\mathbf{B}$ mają te same wartości własne.

$\boxed{5}$ Macierz mająca wielokrotną wartość własną nie może być diagonalizowalna.

$\boxed{6}$ Jeśli macierz $\mathbf{A}$ ma wielokrotne wartości własne, to istnieje baza ortonormalna jej przestrzeni kolumnowej.

$\boxed{7}$ Operator liniowy $T\colon R^n \to R^n$ mający mniej niż n wartości własnych nie może być diagonalizowalny.

$\boxed{8}$ Wektory własne odpowiadające tej samej wartości własnej operatora liniowego nie muszą być liniowo zależne.

$\boxed{9}$ Liniowo niezależne wektory własne $\mathbf{v}_1$ i $\mathbf{v}_2$ macierzy $\mathbf{A}$ odpowiadają jej różnym wartościom własnym.

$\boxed{10}$ Jeśli λ jest wartością własną operatora liniowego $T\colon V \to V$, to każdy element zbioru $V_\lambda - \{\mathbf{0}\}$ jest wektorem własnym operatora T.

$\boxed{11}$ Jeśli $\mathbf{A} \in R_{n\times n}$ i $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ jest bazą przestrzeni R^n składającą się z wektorów własnych macierzy $\mathbf{A}$, to macierz $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ jest diagonalna, gdy $\mathbf{Q} = [\,\mathbf{b}_1 \ldots \mathbf{b}_n\,]$.

$\boxed{12}$ Przestrzeń własna V_λ macierzy kwadratowej $\mathbf{A}$ jest przestrzenią zerową pewnej innej macierzy.

$\boxed{13}$ Jeśli λ jest wartością własną macierzy $\mathbf{A}$ i β jest liczbą zespoloną, to $\lambda-\beta$ jest wartością własną macierzy $\mathbf{A} - \beta\mathbf{I}$.

$\boxed{14}$ Jeśli λ jest wartością własną macierzy $\mathbf{A}$, to także $-\lambda$ jest wartością własną macierzy $\mathbf{A}$.

$\boxed{15}$ Jeśli λ jest wartością własną nieosobliwej macierzy $\mathbf{A}$, to λ^{-1} jest wartością własną macierzy $\mathbf{A}^{-1}$.

$\boxed{16}$ Jeśli wszystkie wartości własne macierzy $\mathbf{A}$ są zerowe, to $\mathbf{A} = \mathbf{0}$.

$\boxed{17}$ Jeśli macierz $\mathbf{A}$ jest diagonalizowalna i wszystkie wartości własne macierzy $\mathbf{A}$ są sobie równe, to $\mathbf{A}$ jest diagonalna.

$\boxed{18}$ Liczba λ_0 jest wartością własną macierzy $\mathbf{A}$ wtedy i tylko wtedy, gdy jest ona wartością własną macierzy $\mathbf{A}^T$.

$\boxed{19}$ Wektor $\mathbf{v}_0$ jest wektorem własnym macierzy $\mathbf{A}$ wtedy i tylko wtedy, gdy jest on wektorem własnym macierzy $\mathbf{A}^T$.

$\boxed{20}$ Jeśli macierz $\mathbf{A}$ jest symetryczna, to każde dwa wektory własne macierzy $\mathbf{A}$ odpowiadające różnym wartościom własnym są ortogonalne.

$\boxed{21}$ Jeśli $\mathbf{A}$ jest macierzą kwadratową stopnia n i $\mathbf{A}$ ma n wzajemnie ortogonalnych wektorów własnych, to macierz $\mathbf{A}$ jest symetryczna.

Rozdział 11

FORMY KWADRATOWE

11.1. Rzeczywista forma kwadratowa

Definicja 11.1.1. Rzeczywistą *formą kwadratową* n zmiennych nazywamy wielomian jednorodny drugiego stopnia postaci

Forma kwadratowa

$$q = \sum_{i=1}^{n} \sum_{j=1}^{n} m_{ij} x_i x_j, \tag{11.1}$$

w którym współczynniki m_{ij} $(i,\ j = 1, \ldots, n)$ są liczbami rzeczywistymi.

Formę kwadratową (11.1) można także zapisać w postaci

$$q = \sum_{i=1}^{n} m_{ii} x_i^2 + \sum_{1 \leqslant i < j \leqslant n} (m_{ij} + m_{ji}) x_i x_j. \tag{11.2}$$

Przykład 254. Wielomian $q = x_1^2 + 4x_3^2 - 2x_2x_3$ jest formą kwadratową zmiennych x_1, x_2 i x_3, ale wielomian $p = x_1^2 + 4x_3^2 - 2x_2x_3 + x_2 - 1$ nie jest już formą kwadratową zmiennych x_1, x_2 i x_3.

Współczynniki m_{ij} i zmienne x_i formy (11.1) tworzą macierze

$$\mathbf{M} = \begin{bmatrix} m_{11} & m_{12} & \ldots & m_{1n} \\ m_{21} & m_{22} & \ldots & m_{2n} \\ \vdots & \vdots & & \vdots \\ m_{n1} & m_{n2} & \ldots & m_{nn} \end{bmatrix} \quad \text{i} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_3 \end{bmatrix},$$

dla których mamy

$$\begin{aligned} q &= \sum_{i=1}^{n} x_i \sum_{j=1}^{n} m_{ij} x_j = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix} \begin{bmatrix} \sum_{j=1}^{n} m_{1j} x_j \\ \vdots \\ \sum_{j=1}^{n} m_{nj} x_j \end{bmatrix} \\ &= \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix} \begin{bmatrix} m_{11} & \ldots & m_{1n} \\ \vdots & & \vdots \\ m_{n1} & \ldots & m_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{x}^T \mathbf{M} \mathbf{x}. \end{aligned}$$

Zatem formę (11.1) można zapisać w postaci

Postać macierzowa formy kwadratowej

$$q = q(\mathbf{x}) = \mathbf{x}^T \mathbf{M} \mathbf{x}, \tag{11.3}$$

nazywanej jej postacią macierzową. Można zauważyć, że każdą formę kwadratową n zmiennych $(n \geqslant 2)$ można zapisać w postaci macierzowej na nieskończenie wiele sposobów.

Przykład 255. Mamy

$$q = 2x_1^2 + 10x_1x_2 + 3x_2^2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 10 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

ale jednocześnie

$$q = 2x_1^2 + 10x_1x_2 + 3x_2^2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 10-a \\ a & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

dla każdej liczby rzeczywistej a i, w szczególności, dla $a = 5$ jest

$$q = 2x_1^2 + 10x_1x_2 + 3x_2^2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 5 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Okazuje się także, że każdą formę kwadratową można zapisać przy użyciu macierzy symetrycznej. Fakt ten będzie miał zasadnicze znaczenie w naszych dalszych rozważaniach, bo z uwagi na twierdzenie 10.3.4 każdą formę kwadratową będzie można zapisać w bardzo wygodnej posatci, w tzw. postaci kanonicznej.

Twierdzenie 11.1.1. *Dla każdej formy kwadratowej* $q = \mathbf{x}^T\mathbf{M}\mathbf{x}$ *istnieje dokładnie jedna macierz symetryczna* $\mathbf{A}$ *taka, że* $q = \mathbf{x}^T\mathbf{A}\mathbf{x}$.

Dowód. Jeśli $\mathbf{M} = [m_{ij}]$ jest macierzą wymiaru $n \times n$, to macierz $\mathbf{A} = \frac{1}{2}(\mathbf{M} + \mathbf{M}^T)$ jest symetryczna oraz $(\mathbf{A})_{ij} + (\mathbf{A})_{ji} = m_{ij} + m_{ji}$ (i w szczególności $(\mathbf{A})_{ii} = m_{ii}$) dla $i, j \in \{1, \ldots, n\}$. Stąd i z (11.2) mamy

$$\begin{aligned} \mathbf{x}^T\mathbf{A}\mathbf{x} &= \sum_{i=1}^{n} (\mathbf{A})_{ii} x_i^2 + \sum_{i \neq j} \big((\mathbf{A})_{ij} + (\mathbf{A})_{ji}\big) x_i x_j \\ &= \sum_{i=1}^{n} m_{ii} x_i^2 + \sum_{i \neq j} (m_{ij} + m_{ji})\, x_i x_j = \mathbf{x}^T\mathbf{M}\mathbf{x} = q. \end{aligned}$$

Dowód jedyności macierzy $\mathbf{A}$ pozostawiamy czytelnikowi. □

Wobec powyższego twierdzenia możemy przyjąć, że *macierzą formy* kwadratowej $q = \mathbf{x}^T\mathbf{M}\mathbf{x}$ (gdzie $\mathbf{M} \in R_{n\times n}$) jest jedyna macierz symetryczna $\mathbf{A}$ taka, że $\mathbf{x}^T\mathbf{M}\mathbf{x} = \mathbf{x}^T\mathbf{A}\mathbf{x}$ dla każdego $\mathbf{x} \in R^n$. Natomiast *rzędem formy* kwadratowej nazywamy rząd jej jedynej macierzy symetrycznej.

Macierz formy kwadratowej

Rząd formy kwadratowej

Przykład 256. Formę kwadratową $q(\mathbf{x}) = 2x_1^2 + 3x_2^2 - 7x_3^2 - 2x_1x_2 + 3x_1x_3$ zapisać w postaci $\mathbf{x}^T\mathbf{A}\mathbf{x}$ z symetryczną macierzą $\mathbf{A}$.

Współczynniki stojące przy x_1^2, x_2^2 i x_3^2 kolejno stawiamy na głównej przekątnej macierzy $\mathbf{A}$, czyli przyjmujemy $(\mathbf{A})_{11} = 2$, $(\mathbf{A})_{22} = 3$ i $(\mathbf{A})_{33} = -7$. Współczynnik stojący przy x_ix_j dla $i \neq j$ w równych częściach rozdzielamy pomiędzy $(\mathbf{A})_{ij}$ i $(\mathbf{A})_{ji}$. Tu mamy $(\mathbf{A})_{12} = (\mathbf{A})_{21} = -1$, $(\mathbf{A})_{13} = (\mathbf{A})_{31} = 3/2$ i $(\mathbf{A})_{23} = (\mathbf{A})_{32} = 0$ (bo współczynnikiem x_2x_3 jest 0). Łatwo sprawdzić, że istotnie mamy

$$q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 2 & -1 & 3/2 \\ -1 & 3 & 0 \\ 3/2 & 0 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Definicja 11.1.2. Niech $q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ i $p(\mathbf{y}) = \mathbf{y}^T\mathbf{B}\mathbf{y}$ będą rzeczywistymi formami kwadratowymi n zmiennych, gdzie $\mathbf{A}$ i $\mathbf{B}$ są macierzami symetrycznymi.

Mówimy, że *forma* $q(\mathbf{x})$ *jest sprowadzalna do formy* $p(\mathbf{y})$, gdy istnieje nieosobliwa macierz $\mathbf{Q} \in R_{n\times n}$ taka, że

$$\forall_{\mathbf{y}\in R^n}\ q(\mathbf{Q}\mathbf{y}) = p(\mathbf{y}). \tag{11.4}$$

W takim przypadku mówimy też, że *podstawienie* $\mathbf{x} = \mathbf{Q}\mathbf{y}$ *sprowadza formę kwadratową* $q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ *do formy kwadratowej* $p(\mathbf{y}) = \mathbf{y}^T\mathbf{B}\mathbf{y}$, czyli

$$q(\mathbf{Q}\mathbf{y}) = (\mathbf{Q}\mathbf{y})^T\mathbf{A}(\mathbf{Q}\mathbf{y}) = \mathbf{y}^T(\mathbf{Q}^T\mathbf{A}\mathbf{Q})\mathbf{y} = \mathbf{y}^T\mathbf{B}\mathbf{y} = p(\mathbf{y}). \tag{11.5}$$

Łatwo zauważyć, że równości (11.5) zachodzą dla każdego $\mathbf{y} \in R^n$ wtedy i tylko wtedy, gdy

$$\mathbf{B} = \mathbf{Q}^T\mathbf{A}\mathbf{Q}. \tag{11.6}$$

To dowodzi, że forma kwadratowa $q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ jest sprowadzalna do formy kwadratowej $p(\mathbf{y}) = \mathbf{y}^T\mathbf{B}\mathbf{y}$ wtedy i tylko wtedy, gdy zachodzi równość (11.6) dla pewnej nieosobliwej macierzy $\mathbf{Q}$.

Przykład 257. Podstawienie

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{x} = \mathbf{Q}\mathbf{y} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1 + y_2 \\ y_2 \end{bmatrix}$$

sprowadza formę kwadratową $q(\mathbf{x}) = x_1^2 + 4x_1x_2 + 2x_2^2$ do formy kwadratowej $p(\mathbf{y}) = y_1^2 + 6y_1y_2 + 7y_2^2$, bo

$$q(\mathbf{Q}\mathbf{y}) = (y_1 + y_2)^2 + 4(y_1 + y_2)y_2 + 2y_2^2 = y_1^2 + 6y_1y_2 + 7y_2^2 = p(\mathbf{y}).$$

Powyższy fakt jest także konsekwencją tego, że dla macierzy

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix} \quad \text{i} \quad \mathbf{B} = \begin{bmatrix} 1 & 3 \\ 3 & 7 \end{bmatrix},$$

które są odpowiednio macierzami symetrycznymi form kwadratowych $q(\mathbf{x})$ i $p(\mathbf{y})$ oraz dla macierzy nieosobliwej $\mathbf{Q} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ spełniony jest warunek (11.6), bo mamy

$$\mathbf{Q}^T\mathbf{A}\mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 3 & 7 \end{bmatrix} = \mathbf{B}.$$

11.2. Postać kanoniczna formy kwadratowej

Definicja 11.2.1. Jeżeli macierz $\mathbf{A}$ jest diagonalna, to mówimy, że forma kwadratowa $q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ ma postać kanoniczną. Zauważmy, że jeśli $\mathbf{A} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ i $\mathbf{x} = \begin{bmatrix} x_1\, x_2 & \ldots & x_n \end{bmatrix}^T$, to forma kwadratowa $q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ jest formą postaci

Forma kwadratowa w postaci kanonicznej

$$q(\mathbf{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2.$$

W tym miejscu nasuwa się pytanie: które formy kwadratowe (i za pomocą jakich podstawień) można sprowadzić do form w postaci kanonicznej? Uzasadnimy, że każdą rzeczywistą formę kwadratową można sprowadzić do postaci kanonicznej za pomocą przekształcenia ortogonalnego oraz tzw. metodą Lagrange'a. Zaczynamy od tego pierwszego sposobu.

Metoda przekształceń ortogonalnych

Niech $\mathbf{A}$ będzie symetryczną macierzą rzeczywistą wymiaru $n \times n$. Wobec twierdzenia 10.3.4 macierz $\mathbf{A}$ jest ortogonalnie diagonalizowalna, czyli istnieje macierz ortogonalna $\mathbf{Q} = [\,\mathbf{q}_1 \ldots \mathbf{q}_n\,]$ taka, że macierz $\mathbf{Q}^T\mathbf{A}\mathbf{Q}$ jest diagonalna,

$$\mathbf{Q}^T\mathbf{A}\mathbf{Q} = \Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n), \tag{11.7}$$

gdzie $\lambda_1, \ldots, \lambda_n$ są wartościami własnymi macierzy $\mathbf{A}$, a $\mathbf{q}_1, \ldots, \mathbf{q}_n$ są znormalizowanymi wektorami własnymi macierzy $\mathbf{A}$ odpowiadającymi wartościom własnym $\lambda_1, \ldots, \lambda_n$. Jednocześnie układ wektorów $Q = (\mathbf{q}_1, \ldots, \mathbf{q}_n)$ jest bazą przestrzeni R^n i macierz $\mathbf{Q}$ jest macierzą przejścia od bazy Q do bazy kanonicznej $E = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ przestrzeni R^n, czyli $\mathbf{Q} = [1_{R^n}]_E^Q$. Natomiast macierz $\mathbf{Q}^{-1} = \mathbf{Q}^T$ jest macierzą przejścia od bazy E od bazy Q, więc $\mathbf{Q}^T = [1_{R^n}]_Q^E$. Niech teraz $\mathbf{x} = [x_1 \ldots x_n]^T$ będzie dowolnym wektorem z przestrzeni R^n i niech $\mathbf{y} = [y_1 \ldots y_n]^T$ będzie wektorem współrzędnych wektora $\mathbf{x}$ względem bazy Q. Wtedy

$$\mathbf{y} = [\mathbf{x}]_Q = [1_{R^n}]_Q^E \cdot [\mathbf{x}]_E = \mathbf{Q}^T\mathbf{x} \quad \text{i} \quad \mathbf{x} = \mathbf{Q}\mathbf{y}.$$

Przekształcenie $\mathbf{y} = \mathbf{Q}^T\mathbf{x}$ nazywa się *przekształceniem do osi głównych* formy kwadratowej $q = \mathbf{x}^T\mathbf{A}\mathbf{x}$. Natomiast przekształcenie $\mathbf{x} = \mathbf{Q}\mathbf{y}$ jest *ortogonalnym podstawieniem* sprowadzającym formę kwadratową $q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ do postaci kanonicznej, bo wobec (11.7) mamy

Przekształceniem do osi głównych

Podstawienie ortogonalne

$$\begin{aligned} q(\mathbf{Q}\mathbf{y}) &= (\mathbf{Q}\mathbf{y})^T\mathbf{A}(\mathbf{Q}\mathbf{y}) = \mathbf{y}^T(\mathbf{Q}^T\mathbf{A}\mathbf{Q})\mathbf{y} \\ &= \mathbf{y}^T\mathbf{\Lambda}\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2. \end{aligned}$$

W ten sposób udowodniliśmy następujące twierdzenie o sprowadzalności formy kwadratowej do postaci kanonicznej za pomocą przekształcenia ortogonalnego.

Twierdzenie 11.2.1. *Formę kwadratową $q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$, gdzie $\mathbf{A}$ jest symetryczną macierzą rzeczywistą wymiaru $n \times n$, za pomocą podstawienia ortogonalnego $\mathbf{x} = \mathbf{Q}\mathbf{y} = [\,\mathbf{q}_1 \ldots \mathbf{q}_n\,]\,\mathbf{y}$ można sprowadzić do postaci*

$$q(\mathbf{Q}\mathbf{y}) = \mathbf{y}^T\mathbf{\Lambda}\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2, \tag{11.8}$$

gdzie $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ i $\lambda_1, \ldots, \lambda_n$ są wszystkimi wartościami własnymi macierzy $\mathbf{A}$, a $\mathbf{q}_1, \ldots, \mathbf{q}_n$ są znormalizowanymi wektorami własnymi macierzy $\mathbf{A}$ odpowiadającymi wartościom własnym $\lambda_1, \ldots, \lambda_n$.

Przykład 258. Wyznaczyć podstawienie ortogonalne $\mathbf{x} = \mathbf{Q}\mathbf{y}$ sprowadzające formę kwadratową $q(\mathbf{x}) = 2x_1x_2 + 4x_1x_3$ do postaci kanonicznej.

Symetryczną macierzą współczynników podanej formy kwadratowej jest

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

Ponieważ

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} -\lambda & 1 & 2 \\ 1 & -\lambda & 0 \\ 2 & 0 & -\lambda \end{vmatrix} = -\lambda(\lambda^2 - 5),$$

więc jej wartościami własnymi są liczby $\lambda_1 = 0$, $\lambda_2 = \sqrt{5}$ i $\lambda_3 = -\sqrt{5}$. Wektory własne macierzy $\mathbf{A}$ odpowiadające wartościom własnym λ_1, λ_2 i λ_3 wybieramy spośród niezerowych rozwiązań układu równań $(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{x} = \mathbf{0}$ dla $i = 1, 2, 3$. Mamy

$$[\,\mathbf{A} - \lambda_1\mathbf{I}\,|\,\mathbf{0}\,] = [\,\mathbf{A}\,|\,\mathbf{0}\,] = \left[\begin{array}{ccc|c} 0 & 1 & 2 & 0 \\ 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

więc $\mathbf{v}_1 = (0, -2, 1)$ jest wektorem własnym odpowiadającym wartości własnej $\lambda_1 = 0$. Podobnie

$$[\,\mathbf{A} - \lambda_2 \mathbf{I} \,|\, \mathbf{0}\,] = \left[\begin{array}{ccc|c} -\sqrt{5} & 1 & 2 & 0 \\ 1 & -\sqrt{5} & 0 & 0 \\ 2 & 0 & -\sqrt{5} & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & -\sqrt{5} & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

i stąd wynika, że $\mathbf{v}_2 = (\sqrt{5}, 1, 2)$ jest wektorem własnym odpowiadającym wartości własnej $\lambda_2 = \sqrt{5}$. Analogicznie pokazuje się, że $\mathbf{v}_3 = (-\sqrt{5}, 1, 2)$ jest wektorem własnym odpowiadającym wartości własnej $\lambda_3 = -\sqrt{5}$. Zatem

$$\mathbf{Q} = \begin{bmatrix} | & | & | \\ \frac{\mathbf{v}_1}{||\mathbf{v}_1||} & \frac{\mathbf{v}_2}{||\mathbf{v}_2||} & \frac{\mathbf{v}_3}{||\mathbf{v}_3||} \\ | & | & | \end{bmatrix} = \begin{bmatrix} 0 & \sqrt{5}/\sqrt{10} & -\sqrt{5}/\sqrt{10} \\ -2/\sqrt{5} & 1/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{5} & 2/\sqrt{10} & 2/\sqrt{10} \end{bmatrix}$$

jest macierzą ortogonalną i wobec twierdzenia 11.2.1 podstawienie

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{x} = \mathbf{Q}\mathbf{y} = \frac{1}{\sqrt{10}} \begin{bmatrix} 0 & \sqrt{5} & -\sqrt{5} \\ -2\sqrt{2} & 1 & 1 \\ \sqrt{2} & 2 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

sprowadza formę kwadratową $q(\mathbf{x}) = 2x_1x_2 + 4x_1x_3$ do postaci kanonicznej $q(\mathbf{Qy}) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2 = \sqrt{5}y_2^2 - \sqrt{5}y_3^2$.

Przykład 259. Formę kwadratową $q = 5x_1^2 + 8x_1x_2 + 5x_2^2$ za pomocą przekształcenia do osi głównych (tj. za pomocą przekształcenia typu $\mathbf{y} = \mathbf{Q}^T\mathbf{x}$) zapisać w postaci kanonicznej.

Mamy

$$q = \mathbf{x}^T \mathbf{A} \mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Łatwe obliczenia pokazują, że liczby $\lambda_1 = 1$ i $\lambda_2 = 9$ są wartościami własnymi macierzy $\mathbf{A} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$, a $\mathbf{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ i $\mathbf{q}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ są ortonormalnymi wektorami własnymi macierzy $\mathbf{A}$ odpowiadającymi wartościom własnym λ_1 i λ_2. Zatem

$$\mathbf{Q} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

i przekształceniem do osi głównych formy $q = 5x_1^2 + 8x_1x_2 + 5x_2^2$ jest

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{y} = \mathbf{Q}^T\mathbf{x} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} x_1 - x_2 \\ x_1 + x_2 \end{bmatrix}.$$

Wobec (11.8) formę $q = 5x_1^2 + 8x_1x_2 + 5x_2^2$ możemy zapisać w postaci

$$\begin{aligned} q &= \mathbf{y}^T \mathbf{\Lambda} \mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 \\ &= 1\left(\tfrac{1}{\sqrt{2}}(x_1 - x_2)\right)^2 + 9\left(\tfrac{1}{\sqrt{2}}(x_1 + x_2)\right)^2 \\ &= \tfrac{1}{2}(x_1 - x_2)^2 + \tfrac{9}{2}(x_1 + x_2)^2. \end{aligned}$$

Metoda Lagrange'a

Tym razem będziemy sprowadzać formę kwadratową do postaci kanonicznej, czyli do sumy kwadratów, poprzez uzupełnianie do pełnych kwadratów wyrażeń zawierających ustaloną zmienną formy. Przykładowo, jeśli w formie $q = x_1^2 +$

$4x_1x_2-7x_2^2$ wyrażenie $x_1^2+4x_1x_2$ zawierające zmienną x_1 uzupełnimy do pełnego kwadratu, $(x_1^2+4x_1x_2)+4x_2^2=(x_1+2x_2)^2$, to formę q możemy zapisać w postaci

$$q=\big((x_1+2x_2)^2-4x_2^2\big)-7x_2^2=(x_1+2x_2)^2-11x_2^2$$

i jest to jej postać kanoniczna. Ten sposób sprowadzania formy kwadratowej do postaci kanonicznej nazywa się *metodą Lagrange'a.* W ogólnym przypadku bierzemy pod uwagę macierz rzeczywistą symetryczną $\mathbf{A}^{(0)}=[a_{ij}^{(0)}]$ wymiaru $n\times n$ i formę kwadratową $q(\mathbf{x})=\mathbf{x}^T\mathbf{A}^{(0)}\mathbf{x}=\sum_{i,\,j=1}^{n}a_{ij}^{(0)}x_ix_j$. Sprowadzając ją do postaci kanonicznej metodą Lagrange'a rozróżniamy dwa przypadki, w zależności od współczynników $a_{ii}^{(0)}$: Metoda Lagrange'a

(a) $a_{ii}^{(0)}\neq 0$ dla pewnego $i\in\{1,\dots,n\}$;

(b) $a_{ii}^{(0)}=0$ dla $i=1,\dots,n$.

Przypadek (a)**.** Załóżmy, że nie wszystkie współczynniki $a_{ii}^{(0)}$ są równe zeru. Bez utraty ogólności możemy założyć, że $a_{11}^{(0)}\neq 0$. W tym przypadku całe wyrażenie $a_{11}^{(0)}x_1^2+2a_{12}^{(0)}x_1x_2+\ldots+2a_{1n}^{(0)}x_1x_n$ zawierające x_1 uzupełniamy do pełnego kwadratu i formę q zapisujemy w postaci

$$q(\mathbf{x})=\frac{1}{a_{11}^{(0)}}(a_{11}^{(0)}x_1+a_{12}^{(0)}x_2+\ldots+a_{1n}^{(0)}x_n)^2+q_1, \tag{11.9}$$

gdzie $q_1=q_1(x_2,\dots,x_n)$ jest pewną formą kwadratową zmiennych $x_2,\ x_3,\dots,x_n$. Z równości tej wynika, że jeśli $t_2y_2^2+\ldots+t_ny_n^2$ będzie postacią kanoniczną formy q_1, to wyrażenie

$$t_1y_1^2+t_2y_2^2+\ldots+t_ny_n^2,$$

w którym $t_1=\frac{1}{a_{11}^{(0)}}$ i $y_1=a_{11}^{(0)}x_1+\ldots+a_{1n}^{(0)}x_n$, będzie postacią kanoniczną formy q.

Przykład 260. Metodą Lagrange'a formę kwadratową

$$q=x_1^2+4x_1x_2-2x_1x_3+3x_2^2+2x_2x_3+2x_3^2 \tag{11.10}$$

sprowadzić do postaci kanonicznej.

Wyrażenie $x_1^2+4x_1x_2-2x_1x_3$, zawierające x_1 w formie q, uzupełniamy do pełnego kwadratu, otrzymując

$$\begin{aligned} q &= (x_1^2+4x_1x_2-2x_1x_3)+3x_2^2+2x_2x_3+2x_3^2\\ &= \big((x_1+2x_2-x_3)^2-4x_2^2+4x_2x_3-x_3^2\big)+3x_2^2+2x_2x_3+2x_3^2\\ &= (x_1+2x_2-x_3)^2+(-x_2^2+6x_2x_3+x_3^2).\end{aligned}$$

Podobnie postępujemy z wyrażeniem $-x_2^2+6x_2x_3$, zawierającym x_2 w formie $q_1=-x_2^2+6x_2x_3+x_3^2$ zmiennych x_2 i x_3. Mamy

$$\begin{aligned} q_1 &= -x_2^2+6x_2x_3+x_3^2=\big(-(-x_2+3x_3)^2+9x_3\big)+x_3^2\\ &= -(-x_2+3x_3)^2+10x_3^2\end{aligned}$$

i to ostatnie wyrażenie jest postacią kanoniczną formy q_1. Stąd otrzymujemy

$$q=(x_1+2x_2-x_3)^2-(-x_2+3x_3)^2+10x_3^2. \tag{11.11}$$

Zatem, jeśli podstawimy $y_1=x_1+2x_2-x_3,\ y_2=-x_2+3x_3$ i $y_3=x_3$ do (11.11) lub $x_1=y_1+2y_2-5y_3,\ x_2=-y_2+3y_3$ i $x_3=y_3$ do (11.10), to otrzymamy postać kanoniczną

$$q=y_1^2-y_2^2+10y_3^2$$

formy (11.10).

Zauważmy jeszcze, że jeśli proces przekształcania formy (11.10) zaczniemy od czynników zawierających zmienną x_2, to otrzymamy

$$q = \frac{1}{3}(2x_1 + 3x_2 + x_3)^2 - \frac{1}{3}(x_1 + 5x_3)^2 + 10x_3^2$$

i jest to inna postać kanoniczna wyjściowej formy q. Oznacza to, że przedstawienie formy kwadratowej w postaci sumy kwadratów nie jest jednoznaczne.

Warto odnotować, że wszystkie współczynniki formy q zapisanej w postaci (11.9) łatwo wyznacza się z macierzy $\mathbf{A}^{(0)}$. W tym celu zauważmy, że dla formy q_1 mamy

$$\begin{aligned} q_1 &= q(x_1, x_2, \ldots, x_n) - \frac{1}{a_{11}^{(0)}}(a_{11}^{(0)}x_1 + \ldots + a_{1n}^{(0)}x_n)^2 \\ &= \sum_{i,\,j=1}^{n} a_{ij}^{(0)}x_ix_j - \sum_{i,\,j=1}^{n} \frac{a_{1i}^{(0)}a_{1j}^{(0)}}{a_{11}^{(0)}}x_ix_j = \sum_{i,\,j=2}^{n} a_{ij}^{(1)}x_ix_j, \end{aligned}$$

gdzie

$$a_{ij}^{(1)} = a_{ij}^{(0)} - \frac{a_{1i}^{(0)}a_{1j}^{(0)}}{a_{11}^{(0)}} \tag{11.12}$$

dla $i,\, j \in \{2, \ldots, n\}$. Wszystkie te współczynniki są elementami macierzy

$$\mathbf{A}^{(1)} = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \ldots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & \ldots & a_{2n}^{(1)} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2}^{(1)} & \ldots & a_{nn}^{(1)} \end{bmatrix}.$$

Wobec (11.12) możemy przyjąć, że macierz $\mathbf{A}^{(1)}$ otrzymano z macierzy $\mathbf{A}^{(0)}$ odejmując kolejno pierwszy wiersz pomnożony przez $a_{1i}^{(0)}/a_{11}^{(0)}$ od i-tego wiersza dla $i = 2, \ldots, n$. Odnotujmy także, że z symetrii macierzy $\mathbf{A}^{(0)}$ i z równości (11.12) wynika, że podmacierz $\mathbf{A}_{11}^{(1)}$ macierzy $\mathbf{A}^{(1)}$ jest symetryczna, bo

$$a_{ij}^{(1)} = a_{ij}^{(0)} - \frac{a_{1i}^{(0)}a_{1j}^{(0)}}{a_{11}^{(0)}} = a_{ji}^{(0)} - \frac{a_{1j}^{(0)}a_{1i}^{(0)}}{a_{11}^{(0)}} = a_{ji}^{(1)}$$

dla $i,\, j \in \{2, \ldots, n\}$. Zatem, jeśli $a_{22}^{(1)} \neq 0$, to także formę $q_1 = q_1(x_2, \ldots, x_n)$ (określoną przez macierz $\mathbf{A}_{11}^{(1)}$) można przedstawić w postaci sumy

$$q_1 = \frac{1}{a_{22}^{(1)}}\left(a_{22}^{(1)}x_2 + \ldots + a_{2n}^{(1)}x_n\right)^2 + q_2. \tag{11.13}$$

Tym razem q_2 jest formą kwadratową $n-2$ zmiennych $x_3, \ldots, x_n$ i

$$\begin{aligned} q_2 &= q_1(x_2, \ldots, x_n) - \frac{1}{a_{22}^{(1)}}(a_{22}^{(1)}x_2 + \ldots + a_{2n}^{(1)}x_n)^2 \\ &= \sum_{i,\,j=2}^{n} a_{ij}^{(1)}x_ix_j - \sum_{i,\,j=2}^{n} \frac{a_{2i}^{(1)}a_{2j}^{(1)}}{a_{22}^{(1)}}x_ix_j \\ &= \sum_{i,\,j=3}^{n} \left(a_{ij}^{(1)} - \frac{a_{2i}^{(1)}a_{2j}^{(1)}}{a_{22}^{(1)}}\right)x_ix_j = \sum_{i,\,j=3}^{n} a_{ij}^{(2)}x_ix_j, \end{aligned}$$

a jej współczynniki $a_{ij}^{(2)} = a_{ij}^{(1)} - \frac{a_{2i}^{(1)} a_{2j}^{(1)}}{a_{22}^{(1)}}$ są elementami macierzy

$$\mathbf{A}^{(2)} = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & a_{13}^{(0)} & \ldots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \ldots & a_{2n}^{(1)} \\ 0 & 0 & a_{33}^{(2)} & \ldots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \ldots & a_{nn}^{(2)} \end{bmatrix},$$

którą otrzymano z macierzy $\mathbf{A}^{(1)}$ odejmując kolejno jej drugi wiersz pomnożony przez $a_{2i}^{(1)}/a_{22}^{(1)}$ od i-tego wiersza (dla $i = 3, \ldots, n$). Jeśli $a_{33}^{(2)} \neq 0$, to procedurę redukcji macierzy (i przedstawiania formy kwadratowej w postaci sumy kwadratów) możemy kontynuować aż do otrzymania macierzy

Macierz zredukowana formy kwadratowej

$$\mathbf{A}^{(k-1)} = \left[\begin{array}{cccc|ccc} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1k}^{(0)} & a_{1\,k+1}^{(0)} & \cdots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & \cdots & a_{2k}^{(1)} & a_{2\,k+1}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{kk}^{(k-1)} & a_{k\,k+1}^{(k-1)} & \cdots & a_{kn}^{(k-1)} \\ \hline 0 & 0 & \cdots & 0 & a_{k+1\,k+1}^{(k-1)} & \cdots & a_{k+1\,n}^{(k-1)} \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{n\,k+1}^{(k-1)} & \cdots & a_{nn}^{(k-1)} \end{array}\right],$$

którą nazywamy *zredukowaną macierzą formy kwadratowej* q, i w której jest $a_{11}^{(0)} \neq 0$, $a_{22}^{(1)} \neq 0 \ldots, a_{kk}^{(k-1)} \neq 0$ i albo $k = n$, albo $k < n$ i wtedy $a_{k+1\,k+1}^{(k-1)} = \ldots = a_{nn}^{(k-1)} = 0$.

W każdym przypadku formę kwadratową q można zapisać za pomocą współczynników macierzy $\mathbf{A}^{(k-1)}$ i mamy

$$q = \sum_{i=1}^{k-1} \frac{1}{a_{ii}^{(i-1)}} \left(\sum_{j=i}^{n} a_{ij}^{(i-1)} x_j \right)^2 + q_k, \tag{11.14}$$

gdzie

$$q_k = \sum_{i,\,j=k+1}^{n} a_{ij}^{(k-1)} x_i x_j. \tag{11.15}$$

Jeśli $k = n$ albo $k < n$ i $a_{ij}^{(k-1)} = 0$ dla $i, j \in \{k+1, \ldots, n\}$, to q_k jest formą zerową i wtedy

$$q = \sum_{i=1}^{k} \frac{1}{a_{ii}^{(i-1)}} \left(\sum_{j=i}^{n} a_{ij}^{(i-1)} x_j \right)^2 \tag{11.16}$$

i jest to postać kanoniczna formy q. Jeśli zaś jest $k < n$ i $a_{ij}^{(k-1)} \neq 0$ dla pewnych indeksów $i, j \in \{k+1, \ldots, n\}$, $i \neq j$, to q_k jest niezerową formą kwadratową mającą własność (b).

Przykład 261. Wyznaczyć macierz zredukowaną formy kwadratowej

$$q = x_1^2 + 5x_2^2 + 4x_3^2 - x_4^2 + 6x_1x_2 - 4x_1x_3 - 12x_2x_3 - 4x_2x_4 \tag{11.17}$$

i następnie zapisać ją w postaci (11.14) (lub, jeśli to będzie możliwe, w postaci (11.16)).

Niżej wskazane operacje elementarne na wierszach sprowadzają macierz symetryczną $\mathbf{A}^{(0)}$ formy q do zredukowanej macierzy $\mathbf{A}^{(2)}$,

$$\begin{bmatrix} 1 & 3 & -2 & 0 \\ 3 & 5 & -6 & -2 \\ -2 & -6 & 4 & 0 \\ 0 & -2 & 0 & -1 \end{bmatrix} \underset{\sim}{\overset{w_2-3w_1}{w_3+2w_1}} \begin{bmatrix} 1 & 3 & -2 & 0 \\ 0 & -4 & 0 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & -1 \end{bmatrix} \overset{w_4+\frac{1}{2}w_2}{\sim} \left[\begin{array}{cc|cc} 1 & 3 & -2 & 0 \\ 0 & -4 & 0 & -2 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] = \mathbf{A}^{(2)}.$$

Ponieważ w macierzy $\mathbf{A}^{(2)}$ jest $a_{ij}^{(2)} = 0$ dla $i,\, j \in \{3,\, 4\}$, więc wobec (11.16) mamy

$$q = (x_1 + 3x_2 - 2x_3)^2 - \frac{1}{4}(-4x_2 - 2x_4)^2.$$

Z ostatniej formy po przyjęciu $y_1 = x_1 + 3x_2 - 2x_3$, $y_2 = 2x_2 + x_4$ i $y_3 = x_3$ oraz $y_4 = x_4$ (albo z formy (11.17) po podstawieniu $x_1 = y_1 - \frac{3}{2}y_2 + 2y_3 + \frac{3}{2}y_4$, $x_2 = \frac{1}{2}y_2 - \frac{1}{2}y_4$, $x_3 = y_4$ oraz $x_4 = y_4$) otrzymujemy

$$q = y_1^2 - y_2^2$$

i to jest postać kanoniczna formy (11.17).

Przypadek (b)**.** Załóżmy teraz, że forma $q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}^{(0)}\mathbf{x}$ nie jest zerowa, ale $a_{11}^{(0)} = \ldots = a_{nn}^{(0)} = 0$. Wtedy $a_{ij}^{(0)} \neq 0$ dla pewnych indeksów $i,\, j \in \{1, \ldots, n\}$, gdzie $i \neq j$. W takim przypadku podstawienie $\mathbf{x} = \mathbf{Q}\mathbf{y}$, gdzie macierz $\mathbf{Q}$ określona jest równościami

$$x_i = y_i - y_j, \quad x_j = y_i + y_j \quad \text{oraz} \quad x_k = y_k \tag{11.18}$$

dla $k \in \{1, \ldots, n\} - \{i,\, j\}$, sprowadza formę kwadratową $q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}^{(0)}\mathbf{x}$ zmiennych $x_1, \ldots, x_n$ do formy kwadratowej $q(\mathbf{Q}\mathbf{y}) = \mathbf{y}^T\mathbf{Q}^T\mathbf{A}^{(0)}\mathbf{Q}\mathbf{y}$ zmiennych $y_1, \ldots, y_n$ i ta ostatnia forma ma już własność (a), bo w niej współczynnik przy y_i^2 (oraz przy y_j^2) jest niezerowy:

$$2a_{ij}^{(0)}x_ix_j = 2a_{ij}^{(0)}(y_i - y_j)(y_i + y_j) = 2a_{ij}^{(0)}y_i^2 - 2a_{ij}^{(0)}y_j^2 \quad \text{i} \quad 2a_{ij}^{(0)} \neq 0.$$

Przykład 262. Metodą Lagrange'a sprowadzić formę kwadratową

$$q = 2x_1x_2 + 4x_2x_3 \tag{11.19}$$

do postaci kanonicznej.

Tym razem forma q ma własność (b), bo w niej $a_{11}^{(0)} = a_{22}^{(0)} = a_{33}^{(0)} = 0$. Ponieważ $a_{12}^{(0)} \neq 0$, więc tak jak w (11.18) podstawienie

$$x_1 = y_1 - y_2, \quad x_2 = y_1 + y_2 \quad \text{i} \quad x_3 = y_3$$

sprowadza ją do formy

$$q = 2(y_1 - y_2)(y_1 + y_2) + 4(y_1 + y_2)y_3 = 2y_1^2 + 4y_1y_3 - 2y_2^2 + 4y_2y_3$$

mającej własność (a). W tym przypadku wyrażenie $2y_1^2 + 4y_1y_3$ zawierające y_1 (oraz wyrażenie $-2y_2^2 + 4y_2y_3$ zawierające y_2) uzupełniamy do pełnego kwadratu, kolejno otrzymując

$$\begin{aligned} q &= 2\big((y_1 + y_3)^2 - y_3^2\big) - 2\big((y_2 - y_3)^2 - y_3^2\big) \\ &= 2(y_1 + y_3)^2 - 2(y_2 - y_3)^2 \\ &= z_1^2 - z_2^2, \end{aligned}$$

gdzie $z_1 = \sqrt{2}(y_1 + y_3) = \frac{\sqrt{2}}{2}(x_1 + x_2 + 2x_3)$, $z_2 = \sqrt{2}(y_2 - y_3) = \frac{\sqrt{2}}{2}(-x_1 + x_2 - 2x_3)$ i $z_3 = x_3$. (Postać $q = z_1^2 - z_2^2$ można też uzyskać wprost z (11.19) za pomocą podstawienia $x_1 = \frac{\sqrt{2}}{2}(z_1 - z_2 - 2\sqrt{2}z_3)$, $x_2 = \frac{\sqrt{2}}{2}(z_1 + z_2)$ i $x_3 = z_3$.)

Przykład 263. Metodą Lagrange'a wyznaczyć postać kanoniczną formy

$$q = 2x_1x_2 + 2x_1x_3 - 2x_1x_4 - 2x_2x_3 + 2x_2x_4 + 2x_3x_4. \tag{11.20}$$

W formie (11.20) jest $a_{11}^{(0)} = a_{22}^{(0)} = a_{33}^{(0)} = a_{44}^{(0)} = 0$ i $a_{12}^{(0)} \neq 0$, więc możemy podstawić $x_1 = y_1 - y_2$, $x_2 = y_1 + y_2$, $x_3 = y_3$ i $x_4 = y_4$, otrzymując

$$q = 2y_1^2 - 2y_2^2 - 4y_2y_3 + 4y_2y_4 + 2y_3y_4.$$

Teraz możemy wyznaczyć zredukowaną macierz tej ostatniej formy. Ponieważ

$$\mathbf{A}^{(0)} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & -2 & -2 & 2 \\ 0 & -2 & 0 & 1 \\ 0 & 2 & 1 & 0 \end{bmatrix} \sim \ldots \sim \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & -2 & -2 & 2 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 3/2 \end{bmatrix} = \mathbf{A}^{(3)},$$

więc wobec (11.16) mamy

$$q = q(y_1, \ldots, y_4) = 2y_1^2 - 2(y_2 + y_3 - y_4)^2 + \frac{1}{2}(2y_3 - y_4)^2 + \frac{3}{2}y_4^2.$$

Stąd zaś wynika, że

$$q = z_1^2 - z_2^2 + z_3^2 + z_4^2, \tag{11.21}$$

gdzie $z_1 = \sqrt{2}y_1 = \frac{\sqrt{2}}{2}(x_1 + x_2)$, $z_2 = \sqrt{2}(y_2 + y_3 - y_4) = \frac{\sqrt{2}}{2}(-x_1 + x_2 + 2x_3 - 2x_4)$, $z_3 = \frac{\sqrt{2}}{2}(2y_3 - y_4) = \frac{\sqrt{2}}{2}(2x_3 - x_4)$ i $z_4 = \sqrt{\frac{3}{2}}y_4 = \sqrt{\frac{3}{2}}x_3$.

Łatwo zauważyć, że postać (11.21) można także otrzymać wprost z (11.20) przez podstawienie $x_1 = \frac{1}{\sqrt{2}}(z_1 - z_2 + z_3 - \frac{1}{\sqrt{3}}z_4)$, $x_2 = \frac{1}{\sqrt{2}}(z_1 + z_2 - z_3 + \frac{1}{\sqrt{3}}z_4)$, $x_3 = \frac{1}{\sqrt{2}}(z_3 + \frac{1}{\sqrt{3}}z_4)$ i $x_4 = \sqrt{\frac{2}{3}}z_4$.

11.3. Określoność macierzy i formy kwadratowej

Definicja 11.3.1. Niech $\mathbf{A}$ będzie rzeczywistą macierzą symetryczną wymiaru $n \times n$. Mówimy, że forma kwadratowa $q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ (oraz definiująca ją macierz $\mathbf{A}$) jest:

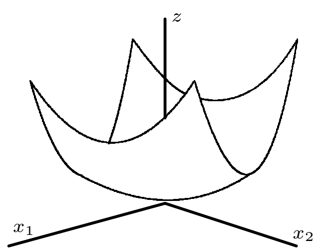
(*a*) *dodatnio określona*, gdy $q(\mathbf{x}) > \mathbf{0}$ dla każdego $\mathbf{x} \in R^n - \{\mathbf{0}\}$; $q(R^n - \{\mathbf{0}\}) \subseteq (0; +\infty)$

(*b*) *dodatnio półokreślona*, gdy $q(\mathbf{x}) \geqslant \mathbf{0}$ dla każdego $\mathbf{x} \in R^n - \{\mathbf{0}\}$ i $q(\mathbf{x}_0) = \mathbf{0}$ dla pewnego $\mathbf{x}_0 \in R^n - \{\mathbf{0}\}$; $0 \in q(R^n - \{\mathbf{0}\}) \subseteq \langle 0; +\infty)$

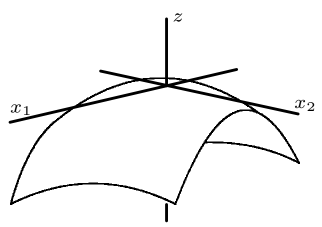
(*c*) *ujemnie określona*, gdy $q(\mathbf{x}) < \mathbf{0}$ dla każdego $\mathbf{x} \in R^n - \{\mathbf{0}\}$; $q(R^n - \{\mathbf{0}\}) \subseteq (-\infty; 0)$

(*d*) *ujemnie półokreślona*, gdy $q(\mathbf{x}) \leqslant \mathbf{0}$ dla każdego $\mathbf{x} \in R^n - \{\mathbf{0}\}$ i $q(\mathbf{x}_0) = \mathbf{0}$ dla pewnego $\mathbf{x}_0 \in R^n - \{\mathbf{0}\}$; $0 \in q(R^n - \{\mathbf{0}\}) \subseteq (-\infty; 0\rangle$

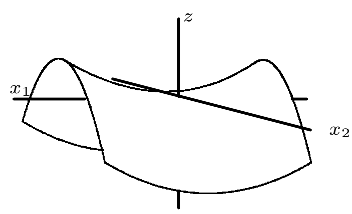
(*e*) *nieokreślona*, gdy $q(\mathbf{x}) > \mathbf{0}$ i $q(\mathbf{y}) < \mathbf{0}$ dla pewnych $\mathbf{x}, \mathbf{y} \in R^n$. $q(R^n - \{\mathbf{0}\}) \cap (-\infty; 0) \neq \emptyset$ i $q(R^n - \{\mathbf{0}\}) \cap (0; +\infty) \neq \emptyset$



Rys. 11.1. Dodatnio określona



Rys. 11.2. Ujemnie określona



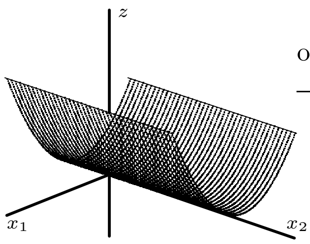
Rys. 11.3. Nieokreślona

Łatwo jest rozpoznać typ określoności formy kwadratowej (i odpowiadającej jej macierzy symetrycznej), gdy forma ta zapisana jest w postaci kanonicznej.

Przykładowo mamy:

$$\begin{aligned} q_1(x_1,x_2) &= x_1^2+3x_2^2 && \text{– forma dodatnio określona;}\\ q_2(x_1,x_2) &= x_1^2 && \text{– forma dodatnio półokreślona;}\\ q_3(x_1,x_2) &= -x_1^2-3x_2^2 && \text{– forma ujemnie określona;}\\ q_4(x_1,x_2) &= -x_1^2 && \text{– forma ujemnie półokreślona;}\\ q_5(x_1,x_2) &= x_1^2-3x_2^2 && \text{– forma nieokreślona.}\end{aligned}$$

Tak samo łatwo wyznacza się typ określoności formy kwadratowej, jeśli jest ona zapisana w postaci sumy lub różnicy kwadratów.



Rys. 11.4. Dodatnio półokreślona

Przykład 264. Formę kwadratową

$$q=-x_1^2+3x_2^2+3x_3^2-4x_1x_2+4x_1x_3$$

przedstawić w postaci sumy (lub różnicy) kwadratów i na tej podstawie wyznaczyć jej typ określoności.

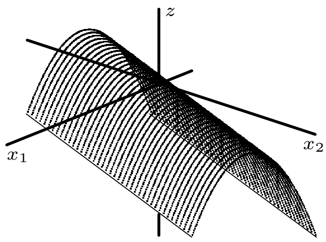
Ponieważ mamy

$$q=-(x_1+2x_2-2x_3)^2+\frac{1}{7}(7x_2-4x_3)^2+\frac{33}{7}x_3^2$$

i ponieważ współczynniki kwadratów tej formy są liczbami różnych znaków, więc można się domyślać, że w jednych punktach forma q przyjmuje wartości ujemne, a w innych dodatnie. Łatwo zauważyć, że

$$q(1,0,0)=-1 \quad \text{i} \quad q(0,1,0)=3$$

i to dowodzi, że forma q jest nieokreślona.



Rys. 11.5. Ujemnie półokreślona

W ogólnym przypadku oczywiste jest następujące twierdzenie.

Twierdzenie 11.3.1. *Jeśli $\lambda_1,\dots,\lambda_n$ są liczbami rzeczywistymi, to macierz diagonalna $\mathbf{A}=\operatorname{diag}(\lambda_1,\dots,\lambda_n)$ oraz forma kwadratowa $q(\mathbf{x})=\mathbf{x}^T\mathbf{A}\mathbf{x}=\lambda_1x_1^2+\lambda_2x_2^2+\ldots+\lambda_nx_n^2$ są:*

(*a*) *dodatnio określone, gdy $\lambda_1>0,\ \dots,\ \lambda_n>0$;*

(*b*) *dodatnio półokreślone, gdy $\lambda_1\geqslant 0,\ \dots,\ \lambda_n\geqslant 0$ i $\lambda_1\cdot\ldots\cdot\lambda_n=0$;*

(*c*) *ujemnie określone, gdy $\lambda_1<0,\ \dots,\ \lambda_n<0$;*

(*d*) *ujemnie półokreślone, gdy $\lambda_1\leqslant 0,\ \dots,\ \lambda_n\leqslant 0$ i $\lambda_1\cdot\ldots\cdot\lambda_n=0$;*

(*e*) *nieokreślone, gdy $\lambda_i\lambda_j<0$ dla pewnych $\lambda_i,\ \lambda_j\in\{\lambda_1,\dots,\lambda_n\}$.* □

Udowodnimy teraz, że typ określoności formy kwadratowej $q=\mathbf{x}^T\mathbf{A}\mathbf{x}$ (i symetrycznej macierzy $\mathbf{A}$) w sposób jednoznaczny zależy od wartości własnych macierzy $\mathbf{A}$. Zaczynamy od dowodu następującego lematu.

Lemat 11.3.1. *Niech $\mathbf{A}$ i $\mathbf{Q}$ będą odpowiednio symetryczną i nieosobliwą macierzą ze zbioru $R_{n\times n}$. Wtedy dla form kwadratowych $q(\mathbf{x})=\mathbf{x}^T\mathbf{A}\mathbf{x}$ i $p(\mathbf{y})=\mathbf{y}^T\mathbf{Q}^T\mathbf{A}\mathbf{Q}\mathbf{y}$ jest*

$$q(R^n-\{\mathbf{0}\})=p(R^n-\{\mathbf{0}\}).$$

Równoważnie, forma kwadratowa $q(\mathbf{x})=\mathbf{x}^T\mathbf{A}\mathbf{x}$ (i macierz $\mathbf{A}$) jest dodatnio określona (dodatnio półokreślona, ujemnie określona, ujemnie półokreślona albo nieokreślona) wtedy i tylko wtedy, gdy forma kwadratowa $p(\mathbf{y})=\mathbf{y}^T\mathbf{Q}^T\mathbf{A}\mathbf{Q}\mathbf{y}$ (i macierz $\mathbf{Q}^T\mathbf{A}\mathbf{Q}$) jest dodatnio określona (dodatnio półokreślona, ujemnie określona, ujemnie półokreślona albo nieokreślona).

Dowód. Z faktu, że macierz $\mathbf{Q}$ jest nieosobliwa wynika, że przekształcenie liniowe $T_{\mathbf{Q}}\colon R^n \to R^n$ (gdzie $T_{\mathbf{Q}}(\mathbf{y}) = \mathbf{Qy}$ dla każdego $\mathbf{y} \in R^n$) wzajemnie jednoznacznie odwzorowuje przestrzeń R^n w siebie. Stąd w szczególności wynika, że

$$T_{\mathbf{Q}}(R^n - \{\mathbf{0}\}) = R^n - \{\mathbf{0}\}. \tag{11.22}$$

Ponieważ dla każdego $\mathbf{y} \in R^n$ jest $q(\mathbf{Qy}) = p(\mathbf{y})$, więc wobec (11.22) mamy

$$\begin{aligned} q(R^n - \{\mathbf{0}\}) &= q(T_{\mathbf{Q}}(R^n - \{\mathbf{0}\})) = q(\{T_{\mathbf{Q}}(\mathbf{y})\colon \mathbf{y} \in R^n - \{\mathbf{0}\}\}) \\ &= q(\{\mathbf{Qy}\colon \mathbf{y} \in R^n - \{\mathbf{0}\}\}) = \{q(\mathbf{Qy})\colon \mathbf{y} \in R^n - \{\mathbf{0}\}\} \\ &= \{p(\mathbf{y})\colon \mathbf{y} \in R^n - \{\mathbf{0}\}\} = p(R^n - \{\mathbf{0}\}). \quad \Box \end{aligned}$$

Wniosek 11.3.1. *Jeśli* $\mathbf{A}$ *jest rzeczywistą macierzą symetryczną wymiaru* $n \times n$ *i* $\lambda_1, \ldots, \lambda_n$ *są jej wszystkimi wartościami własnymi, to macierz* $\mathbf{A}$ *oraz forma kwadratowa* $q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ *są:*

(*a*) *dodatnio określone, gdy* $\lambda_1 > 0, \ldots, \lambda_n > 0$*;*

(*b*) *dodatnio półokreślone, gdy* $\lambda_1 \geqslant 0, \ldots, \lambda_n \geqslant 0$ *i* $\lambda_1 \cdot \ldots \cdot \lambda_n = 0$*;*

(*c*) *ujemnie określone, gdy* $\lambda_1 < 0, \ldots, \lambda_n < 0$*;*

(*d*) *ujemnie półokreślone, gdy* $\lambda_1 \leqslant 0, \ldots, \lambda_n \leqslant 0$ *i* $\lambda_1 \cdot \ldots \cdot \lambda_n = 0$*;*

(*e*) *nieokreślone, gdy* $\lambda_i\lambda_j < 0$ *dla pewnych* $\lambda_i, \lambda_j \in \{\lambda_1, \ldots, \lambda_n\}$*.*

Dowód. Wobec twierdzenia 11.2.1 istnieje przekształcenie ortogonalne $\mathbf{x} = \mathbf{Qy}$ sprowadzające formę kwadratową $q(\mathbf{x})$ do formy kanonicznej $p(\mathbf{y}) = q(\mathbf{Qy}) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2$, w której $\lambda_1, \ldots, \lambda_n$ są wszystkimi wartościami własnymi macierzy $\mathbf{A}$. Wobec lematu 11.3.1 zbiory $q(R^n - \{\mathbf{0}\})$ i $p(R^n - \{\mathbf{0}\})$ są identyczne. To oznacza, że forma $q(\mathbf{x})$ jest dodatnio określona (dodatnio półokreślona, ujemnie określona, ujemnie półokreślona lub nieokreślona) wtedy i tylko wtedy, gdy forma $p(\mathbf{y}) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2$ jest dodatnio określona (dodatnio półokreślona, ujemnie określona, ujemnie półokreślona lub nieokreślona). Stąd i z twierdzenia 11.3.1 wynikają poszczególne części tezy. (Inny dowód wniosku 11.3.1 proponujemy w ćwiczeniach.) $\Box$

Przykład 265. Wyznaczyć wartości własne macierzy symetrycznej formy kwadratowej

$$q = -x_1^2 + 3x_2^2 + 3x_3^2 - 4x_1x_2 + 4x_1x_3$$

i za pomocą wartości własnych wyznaczyć typ określoności formy q.

Wielomianem charakterystycznym macierzy symetrycznej $\mathbf{A}$ formy q jest

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} -1-\lambda & -2 & 2 \\ -2 & 3-\lambda & 0 \\ 2 & 0 & 3-\lambda \end{vmatrix} = (3-\lambda)(\lambda - 1 - 2\sqrt{3})(\lambda - 1 + 2\sqrt{3})$$

i ponieważ wśród wartości własnych $\lambda_1 = 3$, $\lambda_2 = 1 + 2\sqrt{3}$ i $\lambda_3 = 1 - 2\sqrt{3}$ są liczby różnych znaków, więc wobec wniosku 11.3.1 forma q jest nieokreślona.

Wniosek 11.3.2. *Jeśli symetryczna macierz* $\mathbf{A}$ *wymiaru* $n \times n$ *jest dodatnio (ujemnie) określona, to* $\det \mathbf{A} > 0$ $\big((-1)^n \det \mathbf{A} > 0\big)$.

Dowód. Niech $\lambda_1, \ldots, \lambda_n$ będą wszystkimi wartościami własnymi macierzy $\mathbf{A}$. Ponieważ $\det \mathbf{A} = \lambda_1 \cdot \ldots \cdot \lambda_n$ i ponieważ wobec wniosku 11.3.1 liczby $\lambda_1, \ldots, \lambda_n$ (odpowiednio – $-\lambda_1, \ldots, -\lambda_n$) są dodatnie, gdy macierz $\mathbf{A}$ jest dodatnio (odpowiednio – ujemnie) określona, więc $\lambda_1 \cdot \ldots \cdot \lambda_n = \det \mathbf{A} > 0$ $((-\lambda_1) \cdot \ldots \cdot (-\lambda_n) = (-1)^n \det \mathbf{A} > 0)$. $\Box$

Uszczegółowimy teraz poprzedni wniosek i przedstawimy bardziej praktyczne kryterium określoności rzeczywistej macierzy symetrycznej i rzeczywistej formy kwadratowej. Dla potrzeb tego kryterium przyjmujemy następującą definicję.

Wiodąca podmacierz

Wiodący minor główny

Definicja 11.3.2. Niech $\mathbf{A}$ będzie macierzą wymiaru $n \times n$. *Wiodącą podmacierzą główną* stopnia k macierzy $\mathbf{A}$ nazywamy macierz $\mathbf{A}_k$ powstałą z $\mathbf{A}$ przez odrzucenie z niej $n-k$ ostatnich wierszy i $n-k$ ostatnich kolumn. Wyznacznik takiej macierzy $\mathbf{A}_k$ nazywamy *wiodącym minorem głównym* stopnia k macierzy $\mathbf{A}$.

Z definicji tej wynika, że macierz $\mathbf{A} = [a_{ij}]$ wymiaru $n \times n$ ma dokładnie n wiodących podmacierzy głównych i są nimi

$$\mathbf{A}_1 = [\,a_{11}\,],\ \mathbf{A}_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \ldots, \mathbf{A}_n = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}.$$

Twierdzenie Sylvestera

Twierdzenie 11.3.2 (Sylvester). *Symetryczna macierz* $\mathbf{A} = [a_{ij}] \in R_{n\times n}$ (*i forma kwadratowa* $q = \mathbf{x}^T\mathbf{A}\mathbf{x}$) *jest:*

(a) *dodatnio określona* $\Leftrightarrow |\mathbf{A}_k| > 0\,(k = 1, \ldots, n)$;

(b) *ujemnie określona* $\Leftrightarrow (-1)^k|\mathbf{A}_k| > 0\,(k = 1, \ldots, n)$;

(c) *dodatnio półokreślona* $\Leftrightarrow |\mathbf{A}_k| \geqslant 0\,(k = 1, \ldots, n-1),\ |\mathbf{A}_n| = 0$;

(d) *ujemnie półokreślona* $\Leftrightarrow (-1)^k|\mathbf{A}_k| \geqslant 0\,(k{=}1, \ldots, n{-}1),\ |\mathbf{A}_n|{=}0$;

W pozostałych przypadkach macierz $\mathbf{A}$ (*i forma kwadratowa* $q = \mathbf{x}^T\mathbf{A}\mathbf{x}$) *jest nieokreślona.*

Uwaga. Nierówność $(-1)^k|\mathbf{A}_k| > 0$ w (b) jest równoważna stwierdzeniu, że liczby $|\mathbf{A}_k|$ i $(-1)^k$ mają identyczne znaki.

Dowód. (a) Załóżmy najpierw, że macierz $\mathbf{A}$ jest dodatnio określona. Wobec wniosku 11.3.2 jest wtedy $|\mathbf{A}_n| = |\mathbf{A}| > 0$. Udowodnimy teraz, że każda wiodąca podmacierz główna $\mathbf{A}_k$ macierzy $\mathbf{A}$ jest dodatnio określona dla $k = 1, \ldots, n-1$. Weźmy pod uwagę niezerowy wektor $\mathbf{x}_k \in R_{k\times 1}$ i zerowy wektor $\mathbf{0}_{n-k} \in R_{(n-k)\times 1}$. Z dodatniej określoności macierzy $\mathbf{A}$ wynika, że dla niezerowego wektora $\begin{bmatrix} \mathbf{x}_k \\ \mathbf{0}_{n-k} \end{bmatrix}$ jest

$$0 < \begin{bmatrix} \mathbf{x}_k^T & \mathbf{0}_{n-k}^T \end{bmatrix} \mathbf{A} \begin{bmatrix} \mathbf{x}_k \\ \mathbf{0}_{\overline{k}} \end{bmatrix} = \sum_{i,\,j=1}^{k} a_{ij}x_ix_j = \mathbf{x}_k^T\mathbf{A}_k\mathbf{x}_k.$$

To dowodzi, że macierz $\mathbf{A}_k$ jest dodatnio określona. Stąd i z wniosku 11.3.2 znowu wynika, że $|\mathbf{A}_k| > 0$ dla $k = 1, \ldots, n-1$.

Odwrotną implikację udowodnimy indukcyjnie ze względu na n. Implikacja ta jest oczywista dla $n = 1$, bo jeśli $\mathbf{A} = [\,a\,]$ i $|\mathbf{A}_1| = a > 0$, to forma $q = \mathbf{x}^T\mathbf{A}\mathbf{x} = ax_1^2$ (i macierz $\mathbf{A}$) jest dodatnio określona. Niech teraz $n \geqslant 2$ będzie liczbą naturalną i niech $\mathbf{A} = [\,a_{ij}\,]$ będzie symetryczną macierzą wymiaru $n\times n$, której wszystkie wiodące minory główne $|\mathbf{A}_1|, |\mathbf{A}_2|, \ldots, |\mathbf{A}_{n-1}|, |\mathbf{A}_n|$ są dodatnie. Z założenia indukcyjnego wynika, że macierz $\mathbf{A}_{n-1}$ jest dodatnio określona (bo wszystkie jej wiodące minory główne $|\mathbf{A}_1|, \ldots, |\mathbf{A}_{n-1}|$ są dodatnie). Weźmy teraz pod uwagę macierze

$$\mathbf{a} = \begin{bmatrix} a_{1\,n} \\ \vdots \\ a_{n-1\,n} \end{bmatrix},\ \mathbf{A} = \left[\begin{array}{c|c} \mathbf{A}_{n-1} & \mathbf{a} \\ \hline \mathbf{a}^T & a_{nn} \end{array}\right],\ \mathbf{Q} = \left[\begin{array}{c|c} \mathbf{I}_{n-1} & \mathbf{A}_{n-1}^{-1}\mathbf{a} \\ \hline \mathbf{0}_{n-1}^T & 1 \end{array}\right] \text{ i } \mathbf{B} = \left[\begin{array}{c|c} \mathbf{A}_{n-1} & \mathbf{0}_{n-1} \\ \hline \mathbf{0}_{n-1}^T & d \end{array}\right],$$

gdzie $d = a_{nn} - \mathbf{a}^T\mathbf{A}_{n-1}^{-1}\mathbf{a}$ i $\mathbf{0}_{n-1}$ jest macierzą zerową wymiaru $(n-1)\times 1$. Można sprawdzić, że dla tych macierzy mamy

$$\mathbf{Q}^T\mathbf{B}\mathbf{Q} = \left[\begin{array}{c|c} \mathbf{I}_{n-1} & \mathbf{0}_{n-1} \\ \hline (\mathbf{A}_{n-1}^{-1}\mathbf{a})^T & 1 \end{array}\right] \left[\begin{array}{c|c} \mathbf{A}_{n-1} & \mathbf{0}_{n-1} \\ \hline \mathbf{0}_{n-1}^T & d \end{array}\right] \left[\begin{array}{c|c} \mathbf{I}_{n-1} & \mathbf{A}_{n-1}^{-1}\mathbf{a} \\ \hline \mathbf{0}_{n-1}^T & 1 \end{array}\right] = \mathbf{A}.$$

Ponieważ $|\mathbf{Q}| = 1 = |\mathbf{Q}^T|$, więc z powyższej równości i z twierdzenia 6.2.2 mamy

$$|\mathbf{A}| = |\mathbf{Q}^T\mathbf{B}\mathbf{Q}| = |\mathbf{Q}^T||\mathbf{B}||\mathbf{Q}| = |\mathbf{B}| = d\cdot|\mathbf{A}_{n-1}|$$

i stąd wynika, że $d > 0$ (bo liczby $|\mathbf{A}|$ i $|\mathbf{A}_{n-1}|$ są dodatnie). Z faktu, że macierz $\mathbf{A}_{n-1}$ jest dodatnio określona i d jest liczbą dodatnią, wynika, że dla niezerowego wektora

$\mathbf{x} = [\,x_1\ \dots\ x_{n-1}\, x_n\,]^T \in R_{n\times 1}$ i dla wektora $\overline{\mathbf{x}} = [\,x_1\ \dots\ x_{n-1}\,]^T$ (powstałego z $\mathbf{x}$ przez odrzucenie z niego ostatniej współrzędnej) jest

$$\mathbf{x}^T\mathbf{B}\mathbf{x} = \left[\begin{array}{cc} \overline{\mathbf{x}}^T & x_n \end{array}\right] \left[\begin{array}{c|c} \mathbf{A}_{n-1} & \mathbf{0}_{n-1} \\ \hline \mathbf{0}_{n-1}^T & d \end{array}\right] \left[\begin{array}{c} \overline{\mathbf{x}} \\ x_n \end{array}\right] = \overline{\mathbf{x}}^T\mathbf{A}_{n-1}\overline{\mathbf{x}} + d\cdot x_n^2 > 0.$$

To dowodzi, że macierz $\mathbf{B}$ jest dodatnio określona. Stąd zaś i z lematu 11.3.1 wynika dodatnia określoność macierzy $\mathbf{A}$.

(b) Jest oczywiste, że macierz $\mathbf{A}$ jest ujemnie określona wtedy i tylko wtedy, gdy macierz $-\mathbf{A}$ jest dodatnio określona. Wobec już udowodnionego stwierdzenia (a) tak jest wtedy i tylko wtedy, gdy każdy wiodący minor główny $|(-\mathbf{A})_k|$ macierzy $-\mathbf{A}$ jest dodatni, tj. wtedy i tylko wtedy, gdy $(-1)^k|\mathbf{A}_k| = |(-\mathbf{A})_k|$ jest liczbą dodatnią. To kończy dowód stwierdzenia (b).

Analogiczne dowody stwierdzeń (c) i (d) pozostawiamy jako ćwiczenia. □

Podamy teraz dwa przykłady na badanie określoności macierzy i formy kwadratowej z wykorzystaniem twierdzenia Sylvestera.

Przykład 266. Zbadać określoność formy kwadratowej

$$q = -x_1^2 + 3x_2^2 + 3x_3^2 - 4x_1x_2 + 4x_1x_3.$$

Macierzą symetryczną formy q jest

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

której wiodącymi minorami głównymi są

$$|\mathbf{A}_1| = -1,\ \ |\mathbf{A}_2| = \begin{vmatrix} -1 & -2 \\ -2 & 3 \end{vmatrix} = -7 \ \text{ i }\ |\mathbf{A}_3| = \begin{vmatrix} -1 & -2 & 2 \\ -2 & 3 & 0 \\ 2 & 0 & 3 \end{vmatrix} = -33.$$

Ponieważ liczby te nie spełniają żadnego z warunków (a) – (d) twierdzenia 11.3.2, więc rozważana forma kwadratowa q i jej macierz $\mathbf{A}$ są nieokreślone. Do tego samego stwierdzenia w inny sposób doszliśmy w przykładach 264 i 265.

Przykład 267. Korzystając z twierdzenia Sylvestera, zbadać określoność formy kwadratowej

$$q = x_1^2 + 2x_2^2 + 5x_3^2 + 2x_1x_2 + 4x_2x_3.$$

Dla macierzy symetrycznej

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 2 \\ 0 & 2 & 5 \end{bmatrix}$$

formy q jest

$$|\mathbf{A}_1| = 1,\ \ |\mathbf{A}_2| = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1 \ \text{ i }\ |\mathbf{A}_3| = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 2 & 2 \\ 0 & 2 & 5 \end{vmatrix} = 1.$$

Tym razem warunek (a) twierdzenia 11.3.2 jest spełniony, więc forma kwadratowa q i jej macierz $\mathbf{A}$ są dodatnio określone.

11.4. Ćwiczenia

1. Dla formy kwadratowej $q = \mathbf{x}^T\mathbf{A}\mathbf{x}$ wyznaczyć macierz ortogonalną $\mathbf{Q}$ i postać kanoniczną $q = \mathbf{y}^T\mathbf{\Lambda}\mathbf{y}$ uzyskaną za pomocą podstawienia ortogonalnego $\mathbf{x} = \mathbf{Q}\mathbf{y}$, gdy:
   (*a*) $q = 5x_1^2 + 2x_2^2 + 4x_1x_2$;
   (*b*) $q = -2x_1^2 + x_2^2 + 4x_1x_2$;
   (*c*) $q = 3x_1^2 + 3x_2^2 + 3x_3^2 - 2x_1x_3$;
   (*d*) $q = 10x_1^2 + 10x_2^2 - 12x_1x_2$;
   (*e*) $q = 5x_1^2 + 5x_2^2 + 18x_3^2 - 26x_1x_2$;
   (*f*) $q = 3x_2^2 + 3x_3^2 + 4x_1x_2 + 4x_1x_3 - 2x_2x_3$;
   (*g*) $q = 2x_1^2 + 5x_2^2 + 2x_3^2 + 4x_1x_2 + 2x_1x_3 + 4x_2x_3$;
   (*h*) $q = 2x_1^2 + 6x_2^2 + 2x_3^2 + 8x_1x_3$;
   (*i*) $q = x_1^2 + 5x_2^2 + x_3^2 + 2x_1x_2 + 6x_1x_3 + 2x_2x_3$;
   (*j*) $q = 2x_1x_2 + 2x_1x_3 - 2x_1x_4 - 2x_2x_3 + 2x_2x_4 + 2x_3x_4$.
2. Formę $q = q(x_1, x_2)$ przedstawić w postaci sumy kwadratów i następnie zaproponować nowe współrzędne y_1 i y_2 takie, że forma q będzie postaci $\lambda_1 y_1^2 + \lambda_2 y_2^2$, gdy:
   (*a*) $q = x_1^2 + 4x_1x_2 + x_2^2$;
   (*b*) $q = 2x_1^2 + 2x_1x_2 + 2x_2^2$;
   (*c*) $q = x_1^2 - 12x_1x_2 - 4x_2^2$;
   (*d*) $q = 3x_1^2 + 2x_1x_2 + 3x_2^2$.
3. Formę kwadratową $q = q(x_1, \ldots, x_n)$ zapisać w postaci kanonicznej i następnie zbadać jej określoność, gdy:
   (*a*) $q = x_1^2 + 2x_2^2 + x_3^2 - 2x_1x_2 + 4x_1x_3 + 2x_2x_3$;
   (*b*) $q = 5x_1^2 + 5x_2^2 + 2x_3^2 + 8x_1x_2 + 4x_1x_3 + 4x_2x_3$;
   (*c*) $q = x_1^2 + x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3$;
   (*d*) $q = 2x_1^2 + 5x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_1x_3 - 8x_2x_3$;
   (*e*) $q = x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4$.
4. Za pomocą twierdzenia Sylvestera (lub w inny sposób) zbadać określoność macierzy $\mathbf{A}_i$, gdzie:

$$\mathbf{A}_1 = \begin{bmatrix} 2 & 3 \\ 3 & 7 \end{bmatrix}; \quad \mathbf{A}_2 = \begin{bmatrix} 2 & 4 \\ 4 & 7 \end{bmatrix}; \quad \mathbf{A}_3 = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}; \quad \mathbf{A}_4 = \begin{bmatrix} -3 & 4 \\ 4 & -5 \end{bmatrix}; \quad \mathbf{A}_5 = \begin{bmatrix} -3 & 4 \\ 4 & -6 \end{bmatrix};$$

$$\mathbf{A}_6 = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix}; \quad \mathbf{A}_7 = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 5 \\ 0 & 5 & 6 \end{bmatrix}; \quad \mathbf{A}_8 = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix}; \quad \mathbf{A}_9 = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 5 & 1 \\ -1 & 1 & 5 \end{bmatrix};$$

$$\mathbf{A}_{10} = \begin{bmatrix} 20 & 6 & 8 \\ 6 & 3 & 0 \\ 8 & 0 & 8 \end{bmatrix}; \quad \mathbf{A}_{11} = \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 2 & 0 & 5 \\ 3 & 0 & 4 & 0 \\ 0 & 5 & 0 & 6 \end{bmatrix}.$$

5. Udowodnić, że jeśli macierz $\mathbf{A} \in R_{n\times n}$ jest dodatnio określona, to $a_{ii} > 0$ dla $i = 1, \ldots, n$.
6. Dana jest forma kwadratowa $q(x, y, z) = 4x^2 + 3y^2 + 2z^2 + 4xy - 4yz$. (*a*) Wyznaczyć macierz symetryczną $\mathbf{A}$ formy q. (*b*) Wyznaczyć wartości własne i wektory własne macierzy $\mathbf{A}$. (*c*) Zbadać określoność formy q. (*d*) Formę q zapisać w postaci kanonicznej w bazie unormowanych wektorów własnych macierzy $\mathbf{A}$. (*e*) Podać przekształcenie ortogonalne $\mathbf{x} = \mathbf{Q}\mathbf{y}$ sprowadzające formę q do postaci kanonicznej.
7. Bez odwoływania się do wniosku 11.3.1 udowodnić, że jeśli macierz symetryczna $\mathbf{A}$ jest dodatnio (ujemnie) określona, to jest ona nieosobliwa.
8. Udowodnić, że każdy wiodący minor główny macierzy dodatnio półokreślonej jest nieujemny. (Wskazówka. Można wzorować się na dowodzie wniosku 11.3.2 i na dowodzie części (*a*) twierdzenia 11.3.2.)
9. Udowodnić, że każdy minor główny macierzy dodatnio określonej jest dodatni. (Wskazówka. Można wzorować się na dowodzie wniosku 11.3.2 i na dowodzie części (*a*) twierdzenia 11.3.2.)
10. Przedstawić dowody stwierdzeń (*c*) i (*d*) z twierdzenia 11.3.2.
11. (*a*) Udowodnić, że jeśli $\mathbf{x}_0$ jest wektorem własnym macierzy $\mathbf{A}$ odpowiadającym wartości własnej λ_0, to liczby $\mathbf{x}_0^T\mathbf{A}\mathbf{x}_0$ i λ_0 mają zgodne znaki. (*b*) Wyprowadzić nowy dowód wniosku 11.3.1 korzystając z części (*a*) oraz z twierdzenia 10.3.4.
12. Wykazać, że jeśli macierz $\mathbf{A} \in R_{n\times n}$ jest nieosobliwa, to macierz $\mathbf{A}^T\mathbf{A}$ jest dodatnio określona. (Wskazówka. Uwzględnić, że $\mathbf{x}^T\mathbf{I}\mathbf{x} = \mathbf{y}^T\mathbf{A}^T\mathbf{I}\mathbf{A}\mathbf{y}$.)
13. Udowodnić, że dla macierzy $\mathbf{A} \in R_{n\times n}$ następujące stwierdzenia są równoważne: (*a*) $\mathbf{A}$ jest dodatnio określona; (*b*) istnieje macierz nieosobliwa $\mathbf{B}$ taka, że $\mathbf{A} = \mathbf{B}^T\mathbf{B}$; (*c*) istnieje macierz nieosobliwa $\mathbf{P}$ taka, że $\mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{I}_n$.
14. Udowodnić, że jeśli macierz symetryczna $\mathbf{A} \in R_{n\times n}$ jest dodatnio określona, to także macierz $\mathbf{A}^p$ jest dodatnio określona dla każdej liczby całkowitej p.
15. Pokazać, że funkcja $q\colon R_{n\times n} \to R$ jest dodatnio określoną formą kwadratową, gdy $q(\mathbf{X}) = \mathrm{tr}(\mathbf{X}^T\mathbf{X})$ dla $\mathbf{X} \in R_{n\times n}$.
16. Wpisując TAK albo NIE, stwierdzić prawdziwość każdego z następujących zdań:
    $\boxed{1\quad}$ Macierz $\mathbf{A}$ jest dodatnio określona wtedy i tylko wtedy, gdy macierz $\mathbf{A}^{-1}$ jest dodatnio określona.
    $\boxed{2\quad}$ Jeśli macierze $\mathbf{A}$ i $\mathbf{B}$ są dodatnio określone, to także macierz $\mathbf{A} + \mathbf{B}$ jest dodatnio określona.
    $\boxed{3\quad}$ Jeśli macierze $\mathbf{A}$ jest symetryczna, a $\mathbf{P}$ jest macierzą ortogonalną, to podstawienie $\mathbf{x} = \mathbf{P}\mathbf{y}$ przekształca formę kwadratową $q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ w formę kwadratową $q(\mathbf{P}\mathbf{y}) = \mathbf{y}^T(\mathbf{P}^T\mathbf{A}\mathbf{P})\mathbf{y}$ w postaci kanonicznej.
    $\boxed{4\quad}$ Jeśli macierze $\mathbf{A} \in R_{n\times n}$ jest symetryczna i $\det \mathbf{A} < 0$, to forma kwadratowa $q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ jest nieokreślona.

Rozdział 12

ELEMENTY GEOMETRII ANALITYCZNEJ

12.1. Iloczyn wektorowy wektorów

W tej części rozważamy przestrzeń R^3 ze standardową bazą $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ (gdzie $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, $\mathbf{k} = (0, 0, 1)$) i standardowym iloczynem skalarnym.

Definicja 12.1.1. *Iloczynem wektorowym* uporządkowanej pary $(\mathbf{a}, \mathbf{b})$ wektorów $\mathbf{a} = (a_1, a_2, a_3) = a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}$ oraz $\mathbf{b} = (b_1, b_2, b_3) = b_1\mathbf{i} + b_2\mathbf{j}+b_3\mathbf{k}$ nazywamy wektor

$$\mathbf{a}\times\mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \tag{12.1}$$

$$= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}\mathbf{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}\mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\mathbf{k} \tag{12.2}$$

$$= (a_2b_3 - a_3b_2)\mathbf{i} - (a_1b_3 - a_3b_1)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k}$$

$$= (a_2b_3 - a_3b_2, -a_1b_3 + a_3b_1, a_1b_2 - a_2b_1).$$

Przykład 268. Jeśli $\mathbf{a} = (2, -1, 3)$ i $\mathbf{b} = (1, 3, 4)$, to wobec (12.1) mamy

$$\mathbf{a}\times\mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & -1 & 3 \\ 1 & 3 & 4 \end{vmatrix} = \begin{vmatrix} -1 & 3 \\ 3 & 4 \end{vmatrix}\mathbf{i} - \begin{vmatrix} 2 & 3 \\ 1 & 4 \end{vmatrix}\mathbf{j} + \begin{vmatrix} 2 & -1 \\ 1 & 3 \end{vmatrix}\mathbf{k}$$

$$= -13\mathbf{i} - 5\mathbf{j} + 7\mathbf{k} = (-13, -5, 7).$$

W powyższej definicji wykorzystano pojęcie wyznacznika macierzy i możliwość obliczenia jego wartości za pomocą rozwinięcia względem elementów pierwszego wiersza. Uczyniono to w sposób formalny, bez zwracania uwagi na naturę elementów tej macierzy. Dalej, korzystając z formalnych własności wyznacznika, otrzymujemy kolejne własności iloczynu wektorowego.

Twierdzenie 12.1.1. *Dla każdych wektorów* $\mathbf{a}$*,* $\mathbf{b}$ *i* $\mathbf{c}$ *oraz każdej liczby* α *mamy:*

(*a*) $\mathbf{a}\times\mathbf{b} = -(\mathbf{b}\times\mathbf{a})$;

(*b*) $\mathbf{a}\times\mathbf{a} = 0$;

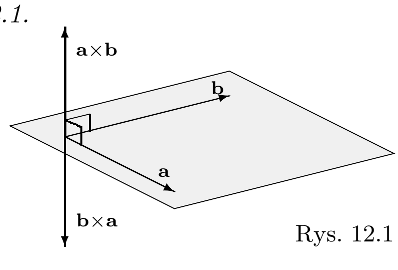
(*c*) $(\alpha\mathbf{a})\times\mathbf{b} = \mathbf{a}\times(\alpha\mathbf{b}) = \alpha(\mathbf{a}\times\mathbf{b})$;

(*d*) $\mathbf{a}\times(\mathbf{b}+\mathbf{c}) = (\mathbf{a}\times\mathbf{b}) + (\mathbf{a}\times\mathbf{c})$ *i* $(\mathbf{a}+\mathbf{b})\times\mathbf{c} = (\mathbf{a}\times\mathbf{c}) + (\mathbf{b}\times\mathbf{c})$;

(*e*) *wektor* $\mathbf{a}\times\mathbf{b}$ *jest ortogonalny do każdego z wektorów* $\mathbf{a}$ *i* $\mathbf{b}$*, rys. 12.1.*

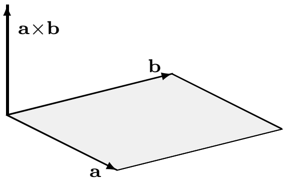
Dowód. Załóżmy, że $\mathbf{a} = (a_1, a_2, a_3)$ i $\mathbf{b} = (b_1, b_2, b_3)$. Ponieważ zamiana miejscami dwóch wierszy macierzy powoduje zmianę znaku wyznacznika (zob. twierdzenie 6.1.5), więc wobec (12.1) mamy

$$\mathbf{a}\times\mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = -\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ b_1 & b_2 & b_3 \\ a_1 & a_2 & a_3 \end{vmatrix} = -(\mathbf{b}\times\mathbf{a})$$



Rys. 12.1

i to dowodzi (a). Podobnie uzasadnia się (b), (c) i (d). Dla dowodu (e) wystarczy pokazać, że $\mathbf{a}\cdot(\mathbf{a}\times\mathbf{b})=0$ i $\mathbf{b}\cdot(\mathbf{a}\times\mathbf{b})=0$. Z definicji iloczynu wektorowego (def. 12.1.1), z definicji wyznacznika (def. 6.1.1) oraz z twierdzenia 6.1.6 mamy



Rys. 12.2. $P_{(\mathbf{a},\mathbf{b})}=\|\mathbf{a}\times\mathbf{b}\|$

$$\begin{aligned}
\mathbf{a}\cdot(\mathbf{a}\times\mathbf{b}) &= (a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k})\cdot\left(\begin{vmatrix}a_2 & a_3\\ b_2 & b_3\end{vmatrix}\mathbf{i}-\begin{vmatrix}a_1 & a_3\\ b_1 & b_3\end{vmatrix}\mathbf{j}+\begin{vmatrix}a_1 & a_2\\ b_1 & b_2\end{vmatrix}\mathbf{k}\right)\\
&= a_1\begin{vmatrix}a_2 & a_3\\ b_2 & b_3\end{vmatrix}-a_2\begin{vmatrix}a_1 & a_3\\ b_1 & b_3\end{vmatrix}+a_3\begin{vmatrix}a_1 & a_2\\ b_1 & b_2\end{vmatrix}\\
&= \begin{vmatrix}a_1 & a_2 & a_3\\ a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\end{vmatrix}=0,
\end{aligned}$$

bo dwa pierwsze wiersze są identyczne. Dowód drugiej części jest podobny. □

Geometryczne własności iloczynu wektorowego przedstawia następujące twierdzenie.

Twierdzenie 12.1.2. *Dla każdych dwóch wektorów* $\mathbf{a}$ *i* $\mathbf{b}$ *mamy:*

(a) *pole* $P_{(\mathbf{a},\mathbf{b})}$ *równoległoboku zbudowanego na wektorach* $\mathbf{a}$ *i* $\mathbf{b}$ *jest równe długości wektora* $\mathbf{a}\times\mathbf{b}$, *tj.* $P_{(\mathbf{a},\mathbf{b})}=\|\mathbf{a}\times\mathbf{b}\|$ *(rys. 12.2);*

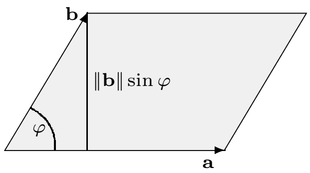
Tożsamość Lagrange'a

(b) $\|\mathbf{a}\times\mathbf{b}\|=\|\mathbf{a}\|\|\mathbf{b}\|\sin\varphi$, *gdzie* φ *jest miarą kąta pomiędzy wektorami* $\mathbf{a}$ *i* $\mathbf{b}$*;*

(c) $\mathbf{a}\times\mathbf{b}=\mathbf{0}$ *wtedy i tylko wtedy, gdy wektory* $\mathbf{a}$ *i* $\mathbf{b}$ *są równoległe;*

(d) $(\mathbf{a}\times\mathbf{b})^2=\mathbf{a}^2\mathbf{b}^2-(\mathbf{a}\cdot\mathbf{b})^2$.

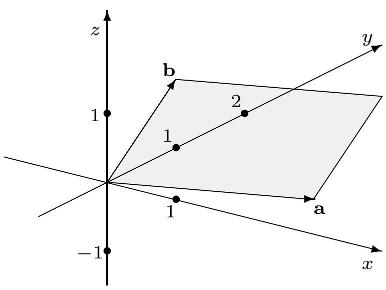
Dowód. Jeżeli φ jest miarą kąta między wektorami $\mathbf{a}$ i $\mathbf{b}$, to pole $P_{(\mathbf{a},\mathbf{b})}$ równoległoboku zbudowanego na wektorach $\mathbf{a}$ i $\mathbf{b}$ (zob. rys. 12.3) jest równe liczbie $\|\mathbf{a}\|\|\mathbf{b}\|\sin\varphi$. Zatem, jeśli $\mathbf{a}=(a_1,a_2,a_3)$ i $\mathbf{b}=(b_1,b_2,b_3)$, to mamy



Rys. 12.3

$$\begin{aligned}
P^2_{(\mathbf{a},\mathbf{b})} &= \|\mathbf{a}\|^2\|\mathbf{b}\|^2\sin^2\varphi\\
&= \|\mathbf{a}\|^2\|\mathbf{b}\|^2(1-\cos^2\varphi)\\
&= \|\mathbf{a}\|^2\|\mathbf{b}\|^2-(\|\mathbf{a}\|\|\mathbf{b}\|\cos\varphi)^2\\
&= \|\mathbf{a}\|^2\|\mathbf{b}\|^2-(\mathbf{a}\cdot\mathbf{b})^2\\
&= (a_1^2+a_2^2+a_3^2)(b_1^2+b_2^2+b_3^2)-(a_1b_1+a_2b_2+a_3b_3)^2\\
&= \begin{vmatrix}a_2 & a_3\\ b_2 & b_3\end{vmatrix}^2+\begin{vmatrix}a_1 & a_3\\ b_1 & b_3\end{vmatrix}^2+\begin{vmatrix}a_1 & a_2\\ b_1 & b_2\end{vmatrix}^2\\
&= \|\mathbf{a}\times\mathbf{b}\|^2
\end{aligned}$$

i z równości tych łatwo wynikają wszystkie cztery dowodzone własności. □



Rys. 12.4

Przykład 269. Wyznaczyć pole równoległoboku zbudowanego na wektorach $\mathbf{a}=(1,2,-1)$ i $\mathbf{b}=(0,1,1)$, zob. rys. 12.4.

Ponieważ

$$\mathbf{a}\times\mathbf{b}=\begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ 1 & 2 & -1\\ 0 & 1 & 1\end{vmatrix}=\begin{vmatrix}2 & -1\\ 1 & 1\end{vmatrix}\mathbf{i}-\begin{vmatrix}1 & -1\\ 0 & 1\end{vmatrix}\mathbf{j}+\begin{vmatrix}1 & 2\\ 0 & 1\end{vmatrix}\mathbf{k}=[3,-1,1],$$

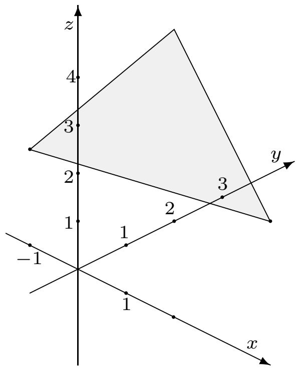
więc z poprzedniego twierdzenia mamy

$$P_{(\mathbf{a},\mathbf{b})}=\|\mathbf{a}\times\mathbf{b}\|=\|(3,-1,1)\|=\sqrt{11}.$$

Przykład 270. Obliczyć pole trójkąta ABC, którego wierzchołkami są punkty $A(1,3,0)$, $B(0,2,5)$ i $C(-1,0,2)$, zob. rys. 12.5.

Ponieważ pole P_{ABC} trójkąta ABC jest równe połowie pola równoległoboku zbudowanego na wektorach $AB = (-1,-1,5)$ i $AC = (-2,-3,2)$, więc z twierdzenia 12.1.2 (a) mamy

$$P_{ABC} = \frac{1}{2}\|AB \times AC\| = \frac{1}{2}\left\| \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -1 & -1 & 5 \\ -2 & -3 & 2 \end{vmatrix} \right\| = \frac{1}{2}\|[13,-8,-5]\| = \sqrt{258}/2.$$



Rys. 12.5

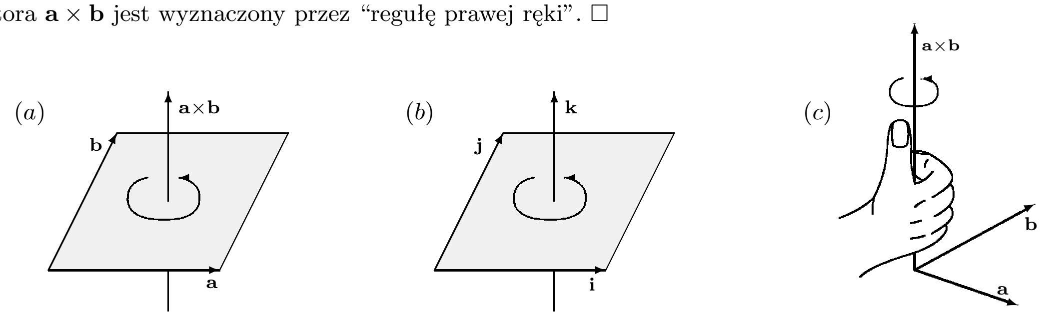
Niech $(\mathbf{a},\mathbf{b},\mathbf{c})$ będzie układem liniowo niezależnych wektorów z przestrzeni R^3. Mówimy, że układ $(\mathbf{a},\mathbf{b},\mathbf{c})$ ma orientację zgodną z orientacją układu $(\mathbf{i},\mathbf{j},\mathbf{k})$, gdy wyznacznik macierzy $[\mathbf{a}\ \mathbf{b}\ \mathbf{c}]^T$ jest dodatni.

Twierdzenie 12.1.3. *Jeśli wektory* $\mathbf{a}$ *i* $\mathbf{b}$ *z przestrzeni* R^3 *są liniowo niezależne, to układ* $(\mathbf{a},\mathbf{b},\mathbf{a}\times\mathbf{b})$ *ma orientację zgodną z orientacją układu* $(\mathbf{i},\mathbf{j},\mathbf{k})$.

Dowód. Z liniowej niezależności wektorów $\mathbf{a} = (a_1,a_2,a_3)$ i $\mathbf{b} = (b_1,b_2,b_3)$ wynika, że co najmniej jeden z wyznaczników $\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$, $\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}$ i $\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}$ jest różny od zera. Wtedy też

$$\det \begin{bmatrix} -\mathbf{a}- \\ -\mathbf{b}- \\ -\mathbf{a}\times\mathbf{b}- \end{bmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ a_2b_3 - a_3b_2 & a_3b_1 - a_1b_3 & a_1b_2 - a_2b_1 \end{vmatrix}$$
$$= \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}^2 + \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}^3 > 0,$$

więc układ $(\mathbf{a},\mathbf{b},\mathbf{a}\times\mathbf{b})$ ma orientację zgodną z orientacją układu $(\mathbf{i},\mathbf{j},\mathbf{k})$ (rys. 12.6), czyli zwrot wektora $\mathbf{a}\times\mathbf{b}$ jest wyznaczony przez "regułę prawej ręki". □



Rys. 12.6

12.2. Iloczyn mieszany wektorów

Definicja 12.2.1. Niech $\mathbf{a}$, $\mathbf{b}$ i $\mathbf{c}$ będą wektorami z przestrzeni R^3. *Iloczynem mieszanym* uporządkowanej trójki $(\mathbf{a},\mathbf{b},\mathbf{c})$ nazywamy liczbę $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})$, czasami oznaczaną przez $\mathbf{abc}$.

Twierdzenie 12.2.1. *Jeśli* $\mathbf{a} = (a_1,a_2,a_3)$, $\mathbf{b} = (b_1,b_2,b_3)$ *i* $\mathbf{c} = (c_1,c_2,c_3)$ *są wektorami z przestrzeni* R^3*, to*

$$\mathbf{abc} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \qquad (12.3)$$

Dowód. Tak jak w dowodzie twierdzenia 12.1.1 (e), mamy

$$\begin{aligned}\mathbf{abc} = \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) &= (a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k})\cdot\left(\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3\end{vmatrix}\mathbf{i} - \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3\end{vmatrix}\mathbf{j} + \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2\end{vmatrix}\mathbf{k}\right)\\ &= a_1\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3\end{vmatrix} - a_2\begin{vmatrix} b_1 & b_3 \\ c_1 & c_3\end{vmatrix} + a_3\begin{vmatrix} b_1 & b_2 \\ c_1 & c_2\end{vmatrix}\\ &= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3\end{vmatrix}. \ \square\end{aligned}$$

Z twierdzenia 12.2.1 i z faktu, że przestawienie miejscami dwóch wierszy macierzy powoduje zmianę znaku jej wyznacznika (zob. wspomniane już twierdzenie 6.1.5), otrzymujemy następną własność iloczynu mieszanego.

Twierdzenie 12.2.2. *Jeśli* $\mathbf{a}$, $\mathbf{b}$ *i* $\mathbf{c}$ *są wektorami z przestrzeni* R^3, *to*

$$\mathbf{abc} = \mathbf{bca} = \mathbf{cab} = -\mathbf{bac} = -\mathbf{acb} = -\mathbf{cba}. \quad \square \tag{12.4}$$

Twierdzenie 12.2.3. *Jeśli* $\mathbf{a}$, $\mathbf{b}$ *i* $\mathbf{c}$ *są wektorami z przestrzeni* R^3, *to*

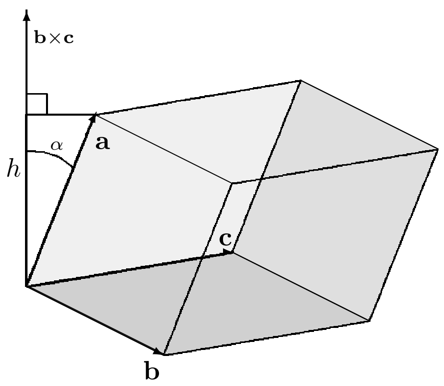
$$\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}. \tag{12.5}$$

Dowód. Z definicji iloczynu mieszanego, z twierdzenia 12.2.2 i z przemienności iloczynu skalarnego wektorów mamy

$$\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \mathbf{abc} = \mathbf{cab} = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b}) = (\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}. \quad \square$$

Twierdzenie 12.2.4. *Objętość* V *równoległościanu zbudowanego na wektorach* $\mathbf{a}$, $\mathbf{b}$ *i* $\mathbf{c}$ *z przestrzeni* R^3 *jest równa wartości bezwzględnej iloczynu mieszanego tych wektorów, tj.*

$$V = |\mathbf{abc}|. \tag{12.6}$$



Rys. 12.7

Dowód. Mamy $\mathbf{abc} = \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \|\mathbf{a}\|\|\mathbf{b}\times\mathbf{c}\|\cos\alpha$, gdzie α jest miarą kąta między wektorami $\mathbf{a}$ i $\mathbf{b}\times\mathbf{c}$. Dla obliczenia objętości równoległościanu możemy przyjąć, że jego podstawą jest równoległobok rozpięty na wektorach $\mathbf{b}$ i $\mathbf{c}$ (zob. rys. 12.7), którego pole jest równe $\|\mathbf{b}\times\mathbf{c}\|$. Wysokość h równoległościanu jest równa długości rzutu wektora $\mathbf{a}$ na prostą prostopadłą do podstawy, mającą kierunek $\mathbf{b}\times\mathbf{c}$, więc h jest wartością bezwzględną liczby $\|\mathbf{a}\|\cos\alpha$. Stąd mamy $V = \|\mathbf{b}\times\mathbf{c}\|\|\mathbf{a}\||\cos\alpha| = |\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})| = |\mathbf{abc}|$. $\square$

Przykład 271. Wyznaczyć objętość równoległościanu zbudowanego na wektorach $\mathbf{a} = (1,0,2)$, $\mathbf{b} = (4,6,2)$ i $\mathbf{c} = (3,3,-6)$.

Wobec twierdzenia 12.2.4 mamy

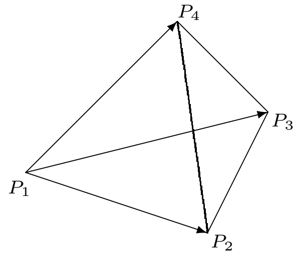
$$V = |\mathbf{abc}| = \left|\begin{vmatrix} 1 & 0 & 2 \\ 4 & 6 & 2 \\ 3 & 3 & -6\end{vmatrix}\right| = 54.$$

Przykład 272. Z poprzedniego twierdzenia i z faktu, że objętość czworościanu rozpiętego na trzech wektorach jest równa szóstej części objętości równoległoboku rozpiętego na tych samych wektorach wynika, że objętość V czworościanu o wierzchołkach w punktach $P_i(x_i, y_i, z_i)$ $(i = 1, 2, 3, 4)$ jest równa szóstej części objętości równoległoboku zbudowanego na wektorach P_1P_4, P_2P_4 i P_3P_4, rys. 12.8. Stąd

$$V = \frac{1}{6}|(P_1P_2)(P_1P_3)(P_1P_4)|, \tag{12.7}$$

czyli

$$V = \frac{1}{6}\left|\begin{vmatrix} x_2-x_1 & y_2-y_1 & z_2-z_1 \\ x_3-x_1 & y_3-y_1 & z_3-z_1 \\ x_4-x_1 & y_4-y_1 & z_4-z_1 \end{vmatrix}\right| = \frac{1}{6}\left|\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix}\right|. \tag{12.8}$$



Rys. 12.8

Z twierdzenia 12.2.1 i z wniosku 7.4.4 natychmiast otrzymujemy następujący związek liniowej niezależności wektorów z wartością iloczynu mieszanego wektorów.

Wniosek 12.2.1. *Trzy wektory* $\mathbf{a}$, $\mathbf{b}$ *i* $\mathbf{c}$ *z przestrzeni* R^3 *są liniowo zależne wtedy i tylko wtedy, gdy* $\mathbf{abc} = 0$. □

12.3. Prosta i płaszczyzna

Definicja 12.3.1. Niech W będzie k-wymiarową podprzestrzenią n-wymiarowej przestrzeni E i niech $\mathbf{r}_0$ będzie ustalonym wektorem z przestrzeni E. Wtedy zbiór

$$\mathbf{r}_0 + W = \{\mathbf{r}_0 + \mathbf{w}\colon\ \mathbf{w} \in W\}$$

nazywamy *k-wymiarową płaszczyzną* w n-wymiarowej przestrzeni E. Wektor $\mathbf{r}_0$ nazywamy *wektorem przesunięcia* płaszczyzny $\mathbf{r}_0 + W$, a podprzestrzeń W jej *kierunkiem.*

Jeśli W jest podprzestrzenią generowaną przez liniowo niezależne wektory $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_k$, to

$$\mathbf{r}_0 + W = \{\mathbf{r}_0 + \sum_{i=1}^{k} t_i\mathbf{m}_i\colon\ t_1, t_2, \ldots, t_k \in R\}$$

i wektory k-wymiarowej płaszczyzny $\mathbf{r}_0 + W$ są określone równaniem

$$\mathbf{r} = \mathbf{r}_0 + \sum_{i=1}^{k} t_i\mathbf{m}_i, \quad t_1, \ldots, t_k \in R. \tag{12.9}$$

Równanie (12.9) nazywamy równaniem k-wymiarowej płaszczyzny $\mathbf{r}_0 + W$ i dalej równanie to utożsamiać będziemy z samą k-wymiarową płaszczyzną. Mówimy także, że k-wymiarowa płaszczyzna (12.9) przechodzi przez punkt P_0 o wektorze wodzącym $\mathbf{r}_0$ i jest równoległa do wektorów $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_k$. Wektory $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_k$ nazywamy *wektorami kierunkowymi* płaszczyzny (12.9), a współczynniki $t_1, \ldots, t_k$ jej parametrami.

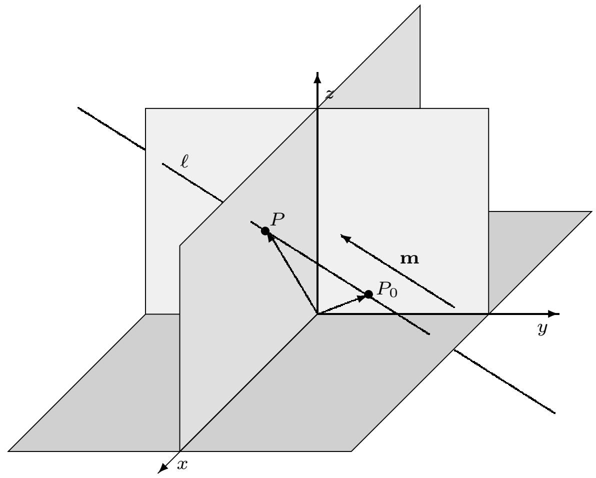
Teraz zajmiemy się przypadkami szczególnymi: $k = 1$ i $k = n - 1$. Każdą 1-wymiarową płaszczyznę zwykle nazywa się *prostą*. Natomiast $(n-1)$-wymiarową płaszczyznę nazywana jest *hiperpłaszczyzną* w n-wymiarowej przestrzeni E lub *płaszczyzną*, gdy $n = 3$.

Prosta przechodząca przez punkt i równoległa do wektora

Niech $\mathbf{r} = (x, y, z)$ i $\mathbf{r}_0 = (x_0, y_0, z_0)$ będą wektorami wodzącymi punktów $P(x, y, z)$ i $P_0(x_0, y_0, z_0)$ i niech $\mathbf{m} = (a, b, c)$ będzie niezerowym wektorem. Wobec (12.9) punkt P leży na prostej ℓ przechodzącej przez punkt P_0 i równoległej do wektora $\mathbf{m}$ (rys. 12.9) wtedy i tylko wtedy, gdy

Równanie wektorowe prostej

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{m}, \quad t \in R. \tag{12.10}$$



Rys. 12.9

Równanie (12.10) nazywamy *równaniem wektorowym* prostej ℓ, a wektor $\mathbf{m}$ jej *wektorem kierunkowym*. Równanie (12.10), czyli równanie

$$(x, y, z) = (x_0, y_0, z_0) + t(a, b, c), \quad t \in R,$$

jest równoważne następującemu układowi trzech równań skalarnych

Równanie parametryczne prostej

$$\begin{cases} x = x_0 + ta, \\ y = y_0 + tb, \quad t \in R. \\ z = z_0 + tc, \end{cases} \tag{12.11}$$

Układ równań (12.11) nazywa się *układem parametrycznych równań prostej* ℓ przechodzącej przez punkt $P_0(x_0, y_0, z_0)$ i równoległej do wektora $\mathbf{m} = (a, b, c)$. Potocznie układ ten nazywa się parametrycznym równaniem prostej ℓ. Rugując z tego układu parametr t, otrzymujemy równoważny układ równości (proporcji)

Równanie kierunkowe prostej

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}. \tag{12.12}$$

Układ ten nazywa się *układem symetrycznych równań* prostej ℓ (albo *równaniem kierunkowym* prostej ℓ) przechodzącej przez punkt $P_0(x_0, y_0, z_0)$ i równoległej do niezerowego wektora $\mathbf{m} = (a, b, c)$. W równościach (12.12) któryś z mianowników (ale nie wszystkie trzy jednocześnie) może być zerowy. W takim przypadku przyjmujemy, że odpowiadający mu licznik także jest zerowy.

Przykład 273. Prosta określona równaniem

$$\frac{x - 4}{5} = \frac{y - 3}{0} = \frac{z + 2}{-1}$$

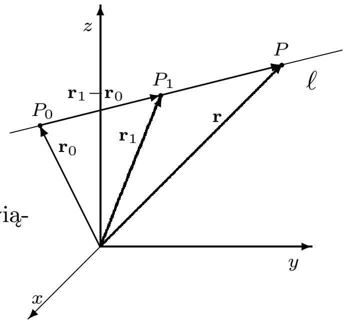
przechodzi przez punkt $(4, 3, -2)$ i jest równoegła do wektora $\mathbf{m} = (5, 0, -1)$. Ta sama prosta jest określona przez równości

$$x = 4 + 5t, \quad y = 3, \quad z = -2 - t \quad \text{dla} \quad t \in R.$$

Przykład 274. Napisać równanie parametryczne prostej ℓ przechodzącej przez punkt $P_0(1,-1,2)$ i równoległej do wektora $\mathbf{m}=(2,1,-1)$. Sprawdzić, czy punkt $P(7,1,1)$ leży na tej prostej.

Wobec (12.11) równaniem parametrycznym prostej ℓ jest

$$\begin{cases} x = 1 + 2t, \\ y = -1 + t, \quad t\in R. \\ z = 2 - t, \end{cases}$$



Rys. 12.10

Dla sprawdzenia czy punkt $P(7,1,1)$ leży na tej prostej, należy zbadać czy ma rozwiązanie następujący układ równań:

$$\begin{cases} 7 = 1 + 2t, \\ 1 = -1 + t, \\ 1 = 2 - t. \end{cases}$$

Rozwiązaniem pierwszego równania, $7=1+2t$, jest $t=2$. Liczba $t=2$ jest także rozwiązaniem drugiego równania, bo $1=-1+2$. Jednakże liczba ta nie jest rozwiązaniem trzeciego równania, $1\neq 2-2$. Zatem punkt $P(7,1,1)$ nie leży na prostej ℓ.

Prosta przechodząca przez dwa punkty

Prosta przechodząca przez dwa różne punkty $P_0(x_0,y_0,z_0)$ i $P_1(x_1,y_1,z_1)$ (o wektorach wodzących $\mathbf{r}_0$ i $\mathbf{r}_1$) jest równoległa do niezerowego wektora $P_0P_1=\mathbf{r}_1-\mathbf{r}_0$, więc jej równanie możemy uzyskać z każdego z równań (12.10)–(12.12) zastępując w nich wektor $\mathbf{m}=(a,b,c)$ lub jego składowe przez wektor $\mathbf{r}_1-\mathbf{r}_0=(x_1-x_0,y_1-y_0,z_1-z_0)$. W szczególności, równaniem wektorowym i równaniem parametrycznym prostej ℓ przechodzącej przez punkty P_0 i P_1 (zob. rys. 12.10) są odpowiednio

$$\mathbf{r}=\mathbf{r}_0+t(\mathbf{r}_1-\mathbf{r}_0), \quad t\in R, \tag{12.13}$$

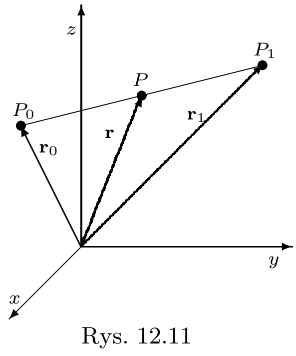
i

$$\begin{cases} x = x_0 + t(x_1-x_0), \\ y = y_0 + t(y_1-y_0), \quad t\in R. \\ z = z_0 + t(z_1-z_0), \end{cases} \tag{12.14}$$

Odcinek

Ograniczając w równaniach (12.13) i (12.14) zakres zmienności parametru t do przedziału $\langle 0;1\rangle$, otrzymujemy równanie *odcinka* o końcach w punktach P_0 i P_1 (rys. 12.11).

Przykład 275. Napisać równanie odcinka o końcach w punktach $P_0(1,2,5)$ i $P_1(-1,3,2)$.



Rys. 12.11

Ponieważ $P_0P_1=(-2,1,-3)$, więc wobec (12.14) odcinek o końcach w punktach P_0 i P_1 określony jest przez równania

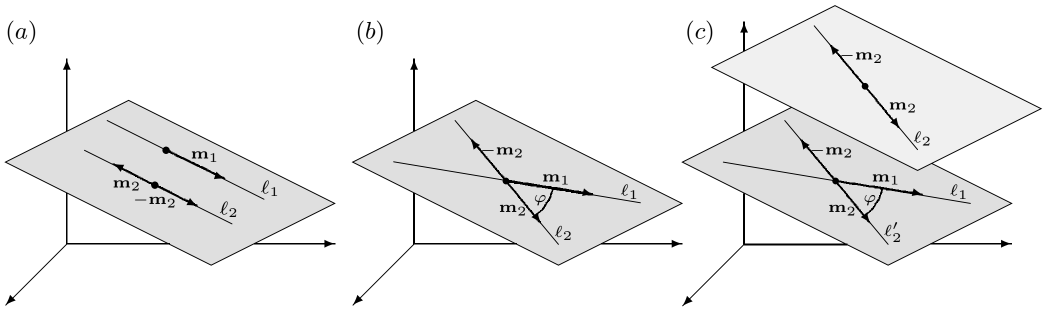
$$\begin{cases} x = 1 - 2t, \\ y = 2 + t, \quad t\in\langle 0;1\rangle. \\ z = 5 - 3t, \end{cases}$$

Kąt nachylenia dwóch prostych

Dane są dwie proste ℓ_1 i ℓ_2 o równaniach

$$\mathbf{r} = \mathbf{r}_1 + t\mathbf{m}_1 \quad \text{i} \quad \mathbf{r} = \mathbf{r}_2 + \lambda\mathbf{m}_2, \ \text{ gdzie } t,\ \lambda \in R. \tag{12.15}$$

Definicja 12.3.2. Miarą kąta nachylenia prostych ℓ_1 i ℓ_2 określonych przez równania (12.15), oznaczamy ją przez $\measuredangle(\ell_1, \ell_2)$, nazywamy miarę $\varphi = \measuredangle(\mathbf{m}_1, \mathbf{m}_2)$ kąta pomiędzy ich wektorami kierunkowymi $\mathbf{m}_1$ i $\mathbf{m}_2$ dobranymi tak, aby φ było liczbą z przedziału $\langle 0; \pi/2\rangle$ (co zawsze można uzyskać przez ewentualne zastąpienie jednego z wektorów $\mathbf{m}_1$ i $\mathbf{m}_2$ przez wektor przeciwny, zob. rys. 12.12 (a)-(c)).



Rys. 12.12

Z ostatniej definicji i ze wzorów na cosinus i sinus kąta pomiędzy wektorami (zob. (9.2.1) i twierdzenie 12.1.2 (b)) otrzymujemy wzory na cosinus i sinus kąta pomiędzy prostymi ℓ_1 i ℓ_2,

$$\cos\measuredangle(\ell_1, \ell_2) = |\cos\measuredangle(\mathbf{m}_1, \mathbf{m}_2)| = \frac{|\mathbf{m}_1 \cdot \mathbf{m}_2|}{\|\mathbf{m}_1\|\|\mathbf{m}_2\|} \tag{12.16}$$

i

$$\sin\measuredangle(\ell_1, \ell_2) = \sin\measuredangle(\mathbf{m}_1, \mathbf{m}_2) = \frac{\|\mathbf{m}_1 \times \mathbf{m}_2\|}{\|\mathbf{m}_1\|\|\mathbf{m}_2\|}. \tag{12.17}$$

Mówimy, że proste ℓ_1 i ℓ_2 określone przez równania (12.15) są prostopadłe (piszemy $\ell_1 \perp \ell_2$) lub równoległe (co oznaczamy przez $\ell_1 \parallel \ell_2$), jeśli ich wektory kierunkowe $\mathbf{m}_1$ i $\mathbf{m}_2$ są odpowiednio prostopadłe lub równoległe. Zatem mamy

Prostopadłość prostych

$$\ell_1 \perp \ell_2 \Leftrightarrow \mathbf{m}_1 \perp \mathbf{m}_2 \Leftrightarrow \mathbf{m}_1 \cdot \mathbf{m}_2 = 0 \tag{12.18}$$

i

Równoległość prostych

$$\ell_1 \parallel \ell_2 \Leftrightarrow \mathbf{m}_1 \parallel \mathbf{m}_2 \Leftrightarrow \mathbf{m}_1 \times \mathbf{m}_2 = \mathbf{0}. \tag{12.19}$$

Wichrowatość prostych

Definicja 12.3.3. Dwie proste nazywamy *skośnymi* (*wichrowymi* lub *niewspółpłaszczyznowymi*), gdy nie są one równoległe i nie przecinają się (rys. 12 (c)).

Jeśli proste (12.15) są nierównoległe (co zachodzi, gdy $\mathbf{m}_1 \times \mathbf{m}_2 \neq 0$) i przecinają się, to dla pewnych liczb rzeczywistych t i λ jest

$$\mathbf{r}_2 + \lambda\mathbf{m}_2 = \mathbf{r}_1 + t\mathbf{m}_1.$$

Wtedy też

$$\mathbf{r}_2 - \mathbf{r}_1 = t\mathbf{m}_1 - \lambda\mathbf{m}_2$$

i

$$(\mathbf{r}_2 - \mathbf{r}_1) \cdot (\mathbf{m}_1 \times \mathbf{m}_2) = t\mathbf{m}_1 \cdot (\mathbf{m}_1 \times \mathbf{m}_2) - \lambda\mathbf{m}_2 \cdot (\mathbf{m}_1 \times \mathbf{m}_2) = 0,$$

bo $\mathbf{m}_1 \cdot (\mathbf{m}_1 \times \mathbf{m}_2) = 0$ i $\mathbf{m}_2 \cdot (\mathbf{m}_1 \times \mathbf{m}_2) = 0$. Stąd zaś wynika, że proste (12.15) są skośne wtedy i tylko wtedy, gdy

$$(\mathbf{m}_1 \times \mathbf{m}_2) \neq 0 \ \text{ i } \ (\mathbf{r}_2 - \mathbf{r}_1) \cdot (\mathbf{m}_1 \times \mathbf{m}_2) \neq 0. \tag{12.20}$$

Przykład 276. Wyznaczyć kąt nachylenia prostych ℓ_1 i ℓ_2 określonych równaniami

$$\frac{x+1}{-2} = y-1 = z+2 \quad \text{i} \quad \begin{cases} x = & t, \\ y = -1 - 2t, \\ z = -1 + t. \end{cases}$$

Wyznaczyć (jeśli jest to możliwe) punkt przecięcia się prostych ℓ_1 i ℓ_2.

Wektorami kierunkowymi prostych ℓ_1 i ℓ_2 są wektory $\mathbf{m}_1 = (-2,1,1)$ i $\mathbf{m}_2 = (1,-2,1)$. Wobec (12.16) mamy

$$\cos\angle(\ell_1,\ell_2) = \frac{|\mathbf{m}_1 \cdot \mathbf{m}_2|}{\|\mathbf{m}_1\|\|\mathbf{m}_2\|} = \frac{|-2-2+1|}{\sqrt{6}\sqrt{6}} = \frac{1}{2}$$

i dlatego $\angle(\ell_1,\ell_2) = \pi/3$.

Proste ℓ_1 i ℓ_2 przetną się wtedy i tylko wtedy, gdy pewien punkt $P(t,-1-2t,-1+t)$ prostej ℓ_2 jest punktem prostej ℓ_1. Podstawiając współrzędne punktu P do równania prostej ℓ_1, otrzymujemy układ równań

$$\frac{t+1}{-2} = -2t-2 = t+1,$$

którego rozwiązaniem jest $t=-1$. Stąd wynika, że punkt $P(-1,1,-2)$ jest punktem przecięcia prostych ℓ_1 i ℓ_2.

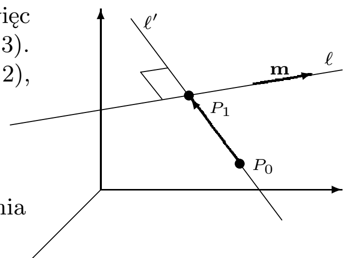
Przykład 277. Prosta ℓ określona jest przez równania $x = 3-t$, $y = 2+t$ i $z = 2t$. Napisać równanie prostej ℓ' przechodzącej przez punkt $P_0(1,0,3)$, przecinającej prostą ℓ i prostopadłej do prostej ℓ.

Prosta ℓ' przechodzi przez pewien punkt $P_1(3-t, 2+t, 2t)$ leżący na prostej ℓ, więc wektor $P_0P_1 = (2-t, 2+t, 2t-3)$ może być jej wektorem kierunkowym (rys. 12.13). Parametr t dobieramy tak, aby wektor P_0P_1 był prostopadły do wektora $\mathbf{m} = (-1,1,2)$, wektora kierunkowego prostej ℓ. Z warunku prostopadłości, czyli z równości

$$\mathbf{m} \cdot P_0P_1 = -(2-t) + (2+t) + 2(2t-3) = 0$$

otrzymujemy $t = 1$. Zatem $P_0P_1 = (1,3,-1)$ i prosta ℓ' określona jest przez równania

$$x = 1+t, \quad y = 3t, \quad z = 3-t.$$



Rys. 12.13

Przykład 278. Zbadać wzajemne położenie prostych ℓ_1 i ℓ_2 określonych odpowiednio przez równania

$$\frac{x-1}{2} = \frac{y+2}{1} = \frac{z-8}{2} \quad \text{i} \quad \frac{x-5}{10} = \frac{y-6}{-4} = \frac{z}{1}.$$

Proste ℓ_1 i ℓ_2 są równoległe do wektorów $\mathbf{m}_1 = (2,1,2)$ oraz $\mathbf{m}_2 = (10,-4,1)$ i przechodzą odpowiednio przez punkty P_1 i P_2 o wektorach wodzących $\mathbf{r}_1 = (1,-2,8)$ i $\mathbf{r}_2 = (5,6,0)$. Ponieważ

$$\mathbf{m}_1 \times \mathbf{m}_2 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 1 & 2 \\ 10 & -4 & 1 \end{vmatrix} = (9,18,-18) \quad \text{i} \quad \mathbf{r}_2 - \mathbf{r}_1 = (4,8,-8)$$

oraz

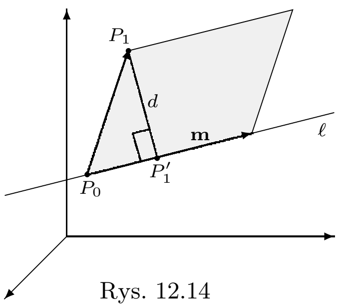
$$(\mathbf{r}_2 - \mathbf{r}_1) \cdot (\mathbf{m}_1 \times \mathbf{m}_2) = (4,8,-8) \cdot (6,18,-18) \neq 0,$$

więc wobec (12.20) proste ℓ_1 i ℓ_2 są skośne.

Odległość punktu od prostej

Niech ℓ będzie prostą przechodzącą przez punkt P_0 i równoległą do niezerowego wektora $\mathbf{m}$. Niech P_1 będzie ustalonym punktem z przestrzeni R^3 i niech P_1' będzie jego rzutem ortogonalnym na prostą ℓ (rys. 12.14). Odległość d punktu P_1 od prostej ℓ jest równa odległości punktu P_1 od punktu P_1'. Dla jej wyznaczenia weźmy pod uwagę pole P równoległoboku rozpiętego na wektorach $\mathbf{m}$ i P_0P_1. Z jednej strony (wobec twierdzenia 12.1.2) mamy $P = \|\mathbf{m} \times P_0P_1\|$. Z drugiej strony pole P jest iloczynem długości podstawy $\mathbf{m}$ i wysokości d tego równoległoboku, $P = \|\mathbf{m}\|\, d$. Zatem $\|\mathbf{m}\|\, d = \|\mathbf{m} \times P_0P_1\|$ i stąd otrzymujemy wzór na odległość d punktu P_1 od prostej ℓ,

$$d = \frac{\|\mathbf{m} \times P_0P_1\|}{\|\mathbf{m}\|}. \tag{12.21}$$



Rys. 12.14

Przykład 279. Obliczyć odległość d punktu $P_1(2,2,3)$ od prostej ℓ określonej przez równania

$$\frac{x-1}{3} = \frac{y-2}{0} = \frac{z+3}{4}.$$

Prosta ℓ przechodzi przez punkt $P_0(1,2,-3)$ i jest równoległa do wektora $\mathbf{m} = (3,0,4)$. Ponieważ $\|\mathbf{m}\| = 5$ i $P_0P_1 = (1,0,6)$, więc wobec (12.21) mamy

$$d = \frac{\|\mathbf{m} \times P_0P_1\|}{\|\mathbf{m}\|} = \frac{1}{5}\left\| \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & 0 & 4 \\ 1 & 0 & 6 \end{vmatrix} \right\| = \frac{1}{5}\|(0,-14,0)\| = \frac{14}{5}.$$

Płaszczyzna w przestrzeni trójwymiarowej

Niech $\mathbf{r}_0 = (x_0, y_0, z_0)$, $\mathbf{r}_1 = (x_1, y_1, z_1)$ i $\mathbf{r}_2 = (x_2, y_2, z_2)$ będą wektorami wodzącymi trzech niewspółliniowych punktów P_0, P_1 i P_2 z przestrzeni R^3. Wtedy wektory

$$\mathbf{m}_1 = \mathbf{r}_1 - \mathbf{r}_0 = (x_1 - x_0, y_1 - y_0, z_1 - z_0) = (a_1, a_2, a_3)$$

i

$$\mathbf{m}_2 = \mathbf{r}_2 - \mathbf{r}_0 = (x_2 - x_0, y_2 - y_0, z_2 - z_0) = (b_1, b_2, b_3)$$

są liniowo niezależne, więc ich iloczyn wektorowy

$$\mathbf{n} = \mathbf{m}_1 \times \mathbf{m}_2 = (A, B, C)$$

jest wektorem niezerowym.

Niech teraz Π będzie płaszczyzną przechodzącą przez punkt P_0 i równoległą do wektorów $\mathbf{m}_1$ i $\mathbf{m}_2$ (rys. 12.15). Wobec (12.9) punkt $P(x,y,z)$ o wektorze wodzącym $\mathbf{r} = (x,y,z)$ leży na płaszczyźnie Π wtedy i tylko wtedy, gdy istnieją liczby t, $\lambda \in R$ takie, że

Równanie wektorowe płaszczyzny

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{m}_1 + \lambda\mathbf{m}_2, \tag{12.22}$$

tzn. gdy wektor $\mathbf{r} - \mathbf{r}_0$ jest kombinacją liniową wektorów $\mathbf{m}_1$ i $\mathbf{m}_2$,

$$\mathbf{r} - \mathbf{r}_0 = t\mathbf{m}_1 + \lambda\mathbf{m}_2. \tag{12.23}$$

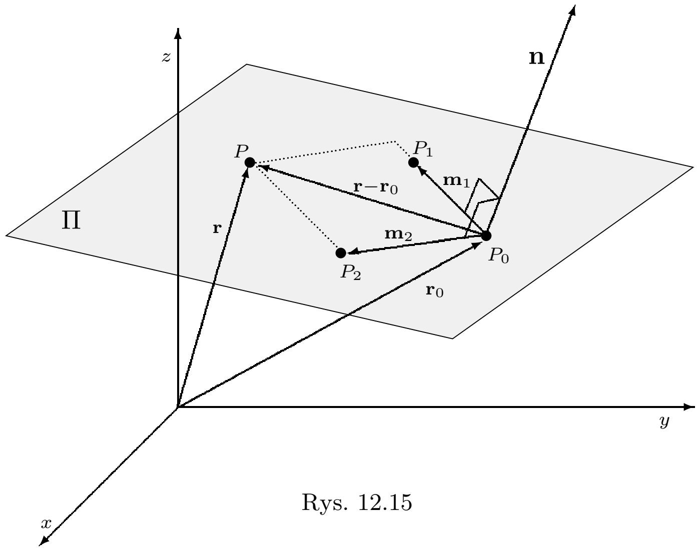
Każde z równań (12.22) i (12.23) nazywamy *równaniem wektorowym płaszczyzny* Π przechodzącej przez punkt P_0 i równoległej do wektorów $\mathbf{m}_1$ i $\mathbf{m}_2$.

Przedstawiając wektory $\mathbf{r}$, $\mathbf{r}_0$, $\mathbf{m}_1$ i $\mathbf{m}_2$ za pomocą ich współrzędnych, łatwo zauważamy, że każde z równań (12.22) i (12.23) jest równoważne układowi równań

$$\begin{cases} x - x_0 &= ta_1 + \lambda b_1, \\ y - y_0 &= ta_2 + \lambda b_2, \\ z - z_0 &= ta_3 + \lambda b_3, \end{cases} \tag{12.24}$$

Układ równań parametrycznych płaszczyzny

który nazywa się *układem parametrycznych równań płaszczyzny* przechodzącej przez punkt $P_0(x_0, y_0, z_0)$ i równoległej do wektorów $\mathbf{m}_1 = (a_1, a_2, a_3)$ i $\mathbf{m}_2 = (b_1, b_2, b_3)$.



Rys. 12.15

Przykład 280. Wyznaczyć równanie płaszczyzny przechodzącej przez punkt $P_0(1, -3, 2)$ i równoległej do prostych ℓ_1 i ℓ_2 określonych równaniami

$$\frac{x-1}{2} = \frac{y+3}{0} = \frac{z-2}{-1} \quad \text{i} \quad \begin{cases} x &= 2 + 3\lambda, \\ y &= 1 - 2\lambda, \\ z &= 1 + \lambda, \end{cases} \quad \lambda \in R.$$

Rozważana płaszczyzna przechodzi przez punkt $P_0(1, -3, 2)$ i jest równoległa do wektorów kierunkowych $\mathbf{m}_1 = (2, 0, -1)$ i $\mathbf{m}_2 = (3, -2, 1)$ prostych ℓ_1 i ℓ_2, więc wobec (12.24) jest ona określona układem parametrycznych równań

$$\begin{cases} x &= 1 + 2t + 3\lambda, \\ y &= -3 - 2\lambda, \\ z &= 2 - t + \lambda, \end{cases}$$

gdzie t, $\lambda \in R$.

Ponieważ $\mathbf{m}_1(\mathbf{m}_1 \times \mathbf{m}_2) = \mathbf{m}_2(\mathbf{m}_1 \times \mathbf{m}_2) = 0$, więc z równania (12.23) wynika, że mamy

$$(\mathbf{r} - \mathbf{r}_0)(\mathbf{m}_1 \times \mathbf{m}_2) = 0 \tag{12.25}$$

i jest to kolejna postać wektorowa równania płaszczyzny przechodzacej przez punkt P_0 i równoległej do wektorów $\mathbf{m}_1$ i $\mathbf{m}_2$ oraz ortogonalnej do wektora $\mathbf{n} = \mathbf{m}_1 \times \mathbf{m}_2$ (rys. 12.16). W tym ostatnim przypadku mowi się, że $\mathbf{n} = \mathbf{m}_1 \times \mathbf{m}_2$ jest *wektorem normalnym* płaszczyzny określonej równaniem (12.25).

Uwzględniając współrzędne wektorów $\mathbf{r} - \mathbf{r}_0 = (x - x_0, y - y_0, z - z_0)$, $\mathbf{m}_1 = (a_1, a_2, a_3)$, $\mathbf{m}_2 = (b_1, b_2, b_3)$ oraz wyznacznikową postać iloczynu mieszanego, równanie (12.25) można zapisać w postaci

Równanie wyznacznikowe płaszczyzny

$$\begin{vmatrix} x-x_0 & y-y_0 & z-z_0 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = 0 \qquad (12.26)$$

zwanej *wyznacznikowym równaniem płaszczyzny* równoległej do wektorów $\mathbf{m}_1$ oraz $\mathbf{m}_2$ i przechodzącej przez punkt $P_0(x_0, y_0, z_0)$.

Jeśli teraz uwzględnimy, że wektory $\mathbf{m}_1$ i $\mathbf{m}_2$ są wyznaczone przez współrzędne punktów P_0, P_1 i P_2, czyli uwzględniając, że $\mathbf{m}_1 = (x_1 - x_0, y_1 - y_0, z_1 - z_0)$ i $\mathbf{m}_2 = (x_2 - x_0, y_2 - y_0, z_2 - z_0)$, to równanie (12.26) można zapisać w postaci

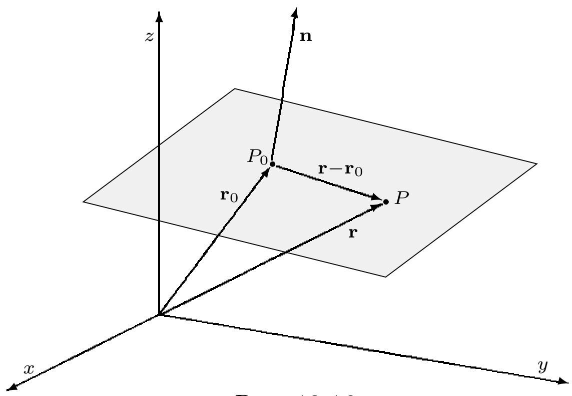
$$\begin{vmatrix} x-x_0 & y-y_0 & z-z_0 \\ x_1-x_0 & y_1-y_0 & z_1-z_0 \\ x_2-x_0 & y_2-y_0 & z_2-z_0 \end{vmatrix} = 0 \qquad (12.27)$$

Wyznacznikowe równania płaszczyzny przechodzącej przez trzy punkty

lub

$$\begin{vmatrix} x & y & z & 1 \\ x_0 & y_0 & z_0 & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \end{vmatrix} = 0. \qquad (12.28)$$

Każde z równań (12.27) i (12.28) jest równaniem płaszczyzny przechodzacej przez trzy niewspółliniowe punkty $P_0(x_0, y_0, z_0)$, $P_1(x_1, y_1, z_1)$ i $P_2(x_2, y_2, z_2)$.



Rys. 12.16

Przykład 281. Napisać równanie płaszczyzny przechodzącej przez punkty $P_0(1, 1, 0)$, $P_1(2, 1, 1)$ i $P_2(0, 2, 1)$

Wobec (12.28) płaszczyzna przechodząca przez punkty P_0, P_1 i P_2 wyznaczona jest przez równanie

$$\begin{vmatrix} x & y & z & 1 \\ 1 & 1 & 0 & 1 \\ 2 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 \end{vmatrix} = 0.$$

Łatwo można zauważyć, że

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

więc rozważana płaszczyzna jest wyznaczona przez równanie $x + 2y - z = 3$.

Z równania (12.25) po podstawieniu (A, B, C) zamiast $\mathbf{m}_1 \times \mathbf{m}_2$ oraz $(x - x_0, y - y_0, z - z_0)$ zamiast $\mathbf{r} - \mathbf{r}_0$ otrzymujemy równania

$$(A, B, C) \cdot (x - x_0, y - y_0, z - z_0) = 0 \tag{12.29}$$

Równania ogólne płaszczyzny

oraz

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0. \tag{12.30}$$

Każde z równań (12.29) i (12.30) jest tzw. *równaniem ogólnym płaszczyzny* przechodzącej przez punkt $P_0(x_0, y_0, z_0)$ i ortogonalnej do wektora $\mathbf{n} = (A, B, C)$. Jeśli przyjmiemy, że $D = -Ax_0 - By_0 - Cz_0$, to każde z powyższych równań płaszczyzny przechodzącej przez punkt $P_0(x_0, y_0, z_0)$ i ortogonalnej do wektora $\mathbf{n} = (A, B, C)$ przyjmuje postać

$$Ax + By + Cz + D = 0 \tag{12.31}$$

i także i to równanie nazywamy równaniem ogólnym płaszczyzny ortogonalnej do wektora (A, B, C).

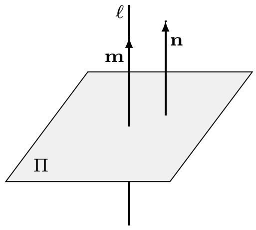
Przykład 282. Płaszczyzna przechodząca przez punkt $P_0(2, 3, -4)$ i ortogonalna do wektora $\mathbf{n} = (-2, 5, 3)$ określona jest przez równanie

$$-2(x - 2) + 5(y - 3) + 3(z + 4) = 0,$$

czyli przez równanie

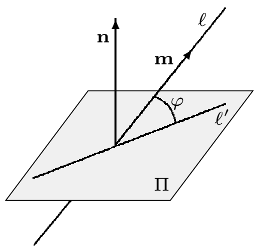
$$-2x + 5y + 3z + 1 = 0.$$

Kąt między prostą i płaszczyzną



Rys. 12.17

Niech $\mathbf{r} = \mathbf{r}_0 + \mathbf{m}t$ i $\mathbf{n}(\mathbf{r} - \mathbf{r}_0') = 0$ będą odpowiednio równaniem prostej ℓ i płaszczyzny Π. Mówimy, że prosta ℓ jest ortogonalna do płaszczyzny Π, jeśli wektory $\mathbf{m}$ i $\mathbf{n}$ są równoległe (rys. 12.17). W takim przypadku mówimy też, że prosta ℓ tworzy kąt $\pi/2$ z płaszczyzną Π. W każdym innym przypadku przez kąt pomiędzy prostą ℓ i płaszczyzną Π rozumiemy kąt φ pomiędzy prostą ℓ a jej rzutem ortogonalnym ℓ' na płaszczyznę Π (rys. 12.18). Zauważmy, że jeśli $\measuredangle(\mathbf{m}, \mathbf{n})$ jest kątem pomiędzy wektorami $\mathbf{m}$ i $\mathbf{n}$, to



Rys. 12.18

$$\varphi = \frac{\pi}{2} - \measuredangle(\mathbf{m}, \mathbf{n}) \quad \text{lub} \quad \varphi = \measuredangle(\mathbf{m}, \mathbf{n}) - \frac{\pi}{2}.$$

W obu przypadkach jest

$$\sin\varphi = |\cos\measuredangle(\mathbf{m}, \mathbf{n})|$$

i dlatego wobec (9.2.1) mamy

$$\sin\varphi = \frac{|\mathbf{m} \cdot \mathbf{n}|}{\|\mathbf{m}\|\|\mathbf{n}\|}.$$

Stąd zaś w szczególności wynika, że prosta ℓ jest równoległa do płaszczyzny Π wtedy i tylko wtedy, gdy wektory $\mathbf{m}$ i $\mathbf{n}$ są ortogonalne. Jeśli prosta ℓ i płaszczyzna Π nie są równoległe, to mają one dokładnie jeden punkt wspólny (czasami nazywany śladem prostej ℓ na płaszczyźnie Π) i punkt ten jest rozwiązaniem układu równań

$$\begin{cases} \mathbf{r} = \mathbf{r}_0 + \mathbf{m}t, \\ \mathbf{n}(\mathbf{r} - \mathbf{r}_0') = 0. \end{cases}$$

Dla jego wyznaczenia wstawiamy $\mathbf{r} = \mathbf{r}_0 + \mathbf{m}t$ do równania $\mathbf{n}(\mathbf{r} - \mathbf{r}_0') = 0$. Wtedy $\mathbf{n}(\mathbf{r}_0 - \mathbf{r}_0') + \mathbf{nm}t = 0$ i stąd $t = -\frac{\mathbf{n}(\mathbf{r}_0 - \mathbf{r}_0')}{\mathbf{nm}}$. Wspólnym punktem prostej ℓ i płaszczyzny Π jest więc

$$\mathbf{r} = \mathbf{r}_0 - \mathbf{m}\frac{\mathbf{n}(\mathbf{r}_0 - \mathbf{r}_0')}{\mathbf{nm}}.$$

Przykład 283. Prosta ℓ i płaszczyzna Π są określone odpowiednio przez równania

$$\frac{x+1}{4} = \frac{y-2}{3} = \frac{z+2}{A} \quad \text{i} \quad Ax + 3y - 5z = 0.$$

Dla jakiej wartości parametru A prosta ℓ będzie równoległa do płaszczyzny Π?

Prosta ℓ jest równoległa do wektora $\mathbf{m} = (4, 3, A)$, a płaszczyzna Π jest ortogonalna do wektora $\mathbf{n} = (A, 3, -5)$. Zatem ℓ będzie równoległa do Π wtedy i tylko wtedy, gdy wektory $\mathbf{m}$ i $\mathbf{n}$ będą ortogonalne. Tak będzie wtedy i tylko wtedy, gdy ich iloczyn skalarny $\mathbf{m} \cdot \mathbf{n} = -A + 9$ będzie równy zeru, tj. wtedy i tylko wtedy, gdy $A = 9$.

Przykład 284. Znaleźć wspólny punkt prostej $(x, y, z) = (1, 2, -12) + (1, 2, 3)t$ i płaszczyzny $x + y + 2z - 6 = 0$.

Dla wyznaczenia wspólnego punktu prostej i płaszczyzny wstawiamy $x = 1 + t$, $y = 2 + 2t$ i $z = -12 + 3t$ do równania płaszczyzny. Wtedy

$$(1 + t) + (2 + 2t) + 2(-12 + 3t) - 6 = 0$$

i stąd $t = 3$. Zatem $(x, y, z) = (1, 2, -12) + (1, 2, 3) \cdot 3 = (4, 8, -3)$ jest wspólnym punktem prostej i płaszczyzny.

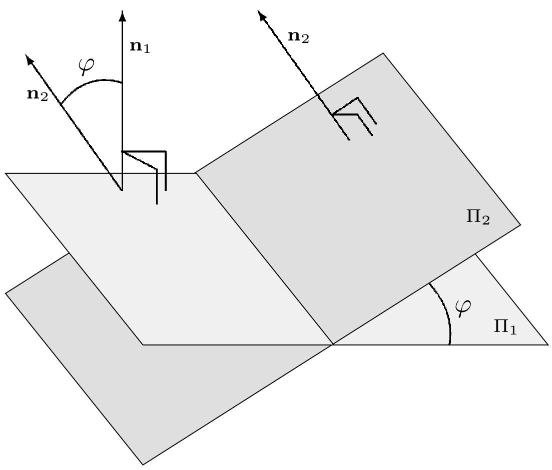
Kąt dwóch płaszczyzn

Niech Π_1 i Π_2 będą płaszczyznami o równaniach

$$\mathbf{n}_1(\mathbf{r} - \mathbf{r}_1) = 0 \quad \text{i} \quad \mathbf{n}_2(\mathbf{r} - \mathbf{r}_2) = 0. \tag{12.32}$$

Przez miarę kąta pomiędzy płaszczyznami Π_1 i Π_2, oznaczamy ją przez $\varphi = \measuredangle(\Pi_1, \Pi_2)$, rozumiemy miarę $\measuredangle(\mathbf{n}_1, \mathbf{n}_2)$ kąta pomiędzy ich wektorami normalnymi $\mathbf{n}_1$ i $\mathbf{n}_2$ dobranymi tak, aby była to liczba z przedziału $\langle 0; \pi/2 \rangle$ (zob. rys. 12.19). Zatem $\varphi = \measuredangle(\mathbf{n}_1, \mathbf{n}_2)$ lub $\varphi = \pi - \measuredangle(\mathbf{n}_1, \mathbf{n}_2)$. W obu przypadkach $\cos\varphi = |\cos\measuredangle(\mathbf{n}_1, \mathbf{n}_2)|$ i ze wzoru na cosinus kąta pomiędzy wektorami otrzymujemy

$$\cos\varphi = \frac{|\mathbf{n}_1 \cdot \mathbf{n}_2|}{\|\mathbf{n}_1\| \|\mathbf{n}_2\|}.$$



Rys. 12.19

Przykład 285. Płaszczyzny o równaniach $x - 12y + 7z + 12 = 0$ oraz $2x + 3y - 4z - 1 = 0$ tworzą kąt

$$\varphi = \arccos \frac{|(1, -12, 7)(2, 3, -4)|}{\|(1, -12, 7)\| \|(2, 3, -4)\|} = \arccos \frac{62}{\sqrt{194}\sqrt{29}} = 34°25'.$$

O płaszczyznach Π_1 i Π_2 określonych równaniami (12.32) mówimy, że są one ortogonalne ($\Pi_1 \perp \Pi_2$) albo równoległe ($\Pi_1 \parallel \Pi_2$), jeśli ich wektory normalne $\mathbf{n}_1$ i $\mathbf{n}_2$ są odpowiednio ortogonalne albo równoległe. Zatem mamy

$$\Pi_1 \perp \Pi_2 \Leftrightarrow \mathbf{n}_1 \perp \mathbf{n}_2 \Leftrightarrow \mathbf{n}_1 \cdot \mathbf{n}_2 = 0 \tag{12.33}$$

i

$$\Pi_1 \parallel \Pi_2 \Leftrightarrow \mathbf{n}_1 \parallel \mathbf{n}_2 \Leftrightarrow \mathbf{n}_1 \times \mathbf{n}_2 = \mathbf{0}. \tag{12.34}$$

Przykład 286. Płaszczyzny Π_1 i Π_2 są określone przez równania

$$x - 2y + z + 7 = 0 \quad \text{oraz} \quad 2x + y - z + 3 = 0.$$

Wyznaczyć równanie płaszczyzny Π przechodzącej przez punkt $A(3, 2, -1)$ i ortogonalnej do płaszczyzn Π_1 i Π_2.

Wektorami normalnymi płaszczyzn Π_1 i Π_2 są wektory $\mathbf{n}_1 = (1, -2, 1)$ i $\mathbf{n}_2 = (2, 1, -1)$. Warunek prostopadłości płaszczyzny Π do płaszczyzn Π_1 i Π_2 jest równoważny równoległości płaszczyzny Π do wektorów $\mathbf{n}_1$ i $\mathbf{n}_2$. Zatem, wobec (12.26), równaniem płaszczyzny Π jest

$$\begin{vmatrix} x-3 & y-2 & z+1 \\ 1 & -2 & 1 \\ 2 & 1 & -1 \end{vmatrix} = 0, \quad \text{czyli} \quad x + 3y + 5z - 4 = 0.$$

Jeśli płaszczyzny Π_1 i Π_2 nie są równoległe, to przecinają się one wzdłuż prostej (rys. 12.19), którą nazywa się *krawędzią przecięcia się płaszczyzn* i której równanie otrzymuje się z układu równań opisujących obie płaszczyzny.

Przykład 287. Płaszczyzny Π_1 i Π_2 o równaniach

$$x - 2y + 2z = 3 \quad \text{i} \quad 2x - 3y + z = 1$$

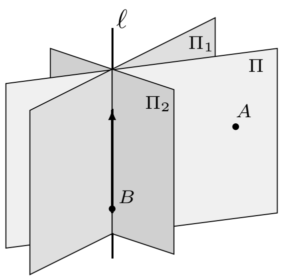
przecinają się wzdłuż prostej ℓ (rys. 12.20). Napisać jej równanie. Wyznaczyć także równanie płaszczyzny Π przechodzącej przez punkt $A(3, 2, 1)$ i zawierającej prostą ℓ.

Prosta ℓ jest zbiorem tych punktów (x, y, z), których współrzędne x, y i z spełniają układ równań

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

Jego rozwiązaniem jest $(x, y, z) = (-7, -5, 0) + (4, 3, 1)t$, $t \in R$, i jest to równanie szukanej prostej ℓ. Prosta ta przechodzi przez punkt $B(-7, -5, 0)$ i jest równoległa do wektora $\mathbf{m} = (4, 3, 1)$. Płaszczyzna Π przechodzi przez punkt $A(3, 2, 1)$ i jest równoległa do wektorów $\mathbf{m} = (4, 3, 1)$ i $AB = (-10, -7, -1)$, więc jest ona określona równaniem

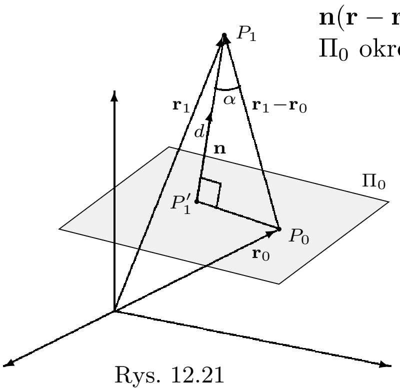
$$\begin{vmatrix} x-3 & y-2 & z-1 \\ 4 & 3 & 1 \\ -10 & -7 & -1 \end{vmatrix} = 0, \quad \text{czyli} \quad 2x - 3y + z = 1.$$



Rys. 12.20

Odległość punktu od płaszczyzny

Niech $\mathbf{r}_0$ oraz $\mathbf{r}_1$ będą wektorami wodzącymi punktów P_0 oraz P_1 i niech $\mathbf{n}$ będzie niezerowym wektorem. Niech Π_0 będzie płaszczyzną określoną równaniem $\mathbf{n}(\mathbf{r}-\mathbf{r}_0)=0$. Pokażemy, że odległość $d=d(P_1,\Pi_0)$ punktu P_1 od płaszczyzny Π_0 określona jest wzorem



Rys. 12.21

$$d(P_1,\Pi_0)=\frac{|\mathbf{n}(\mathbf{r}_1-\mathbf{r}_0)|}{\|\mathbf{n}\|}. \tag{12.35}$$

Dla dowodu powyższego wzoru niech P_1' będzie rzutem ortogonalnym punktu P_1 na płaszczyznę $\mathbf{n}(\mathbf{r}-\mathbf{r}_0)=0$ i niech α będzie miarą kąta między wektorami $\mathbf{n}$ i $\mathbf{r}_1-\mathbf{r}_0$ (rys. 12.21). Z trójkąta prostokątnego $P_0P_1P_1'$ mamy $d=\|\mathbf{r}_1-\mathbf{r}_0\|\,|\cos\alpha|$. Stąd i ze wzoru (9.2.1) mamy

$$|\mathbf{n}(\mathbf{r}_1-\mathbf{r}_0)|=\|\mathbf{n}\|\,\|\mathbf{r}_1-\mathbf{r}_0\|\,|\cos\alpha|=\|\mathbf{n}\|\,d$$

i z tej równości wynika wzór (12.35).

Niech Π_0 i Π_1 będą płaszczyznami równoległymi określonymi przez równania

$$\mathbf{n}(\mathbf{r}-\mathbf{r}_0)=0 \quad \text{i} \quad \mathbf{n}(\mathbf{r}-\mathbf{r}_1)=0, \tag{12.36}$$

gdzie $\mathbf{r}_0$ i $\mathbf{r}_1$ są wektorami wodzącymi punktów P_0 i P_1. Ponieważ odległość $d(\Pi_0,\Pi_1)$ pomiędzy płaszczyznami Π_0 i Π_1 jest równa odległości dowolnego punktu leżącego na jednej płaszczyźnie od drugiej płaszczyzny, więc $d(\Pi_0,\Pi_1)=d(P_1,\Pi_0)$ i (12.35) jest także wzorem na odległość pomiędzy równoległymi płaszczyznami określonymi przez równania (12.36).

Niech teraz Π_0 będzie płaszczyzną określoną równaniem ogólnym (12.31), czyli równaniem

$$Ax+By+Cz+D=0,$$

i niech $P_0(x_0,y_0,z_0)$ będzie dowolnym punktem leżącym na płaszczyźnie Π_0. Dla takiego punktu jest $D=-(Ax_0+By_0+Cz_0)$. Niech teraz $P_1(x_1,y_1,z_1)$ będzie dowolnym punktem z przestrzeni R^3. Jeśli przyjmiemy, że wektorem kierunkowym płaszczyzny π_0 jest wektor $\mathbf{n}=(A,B,C)$ i jeśli $\mathbf{r}_0$ oraz $\mathbf{r}_1$ są wektorami wodzącymi punktów P_0 oraz P_1, to mamy $\|\mathbf{n}\|=\sqrt{A^2+B^2+C^2}$ oraz

$$\begin{aligned} Ax_1+By_1+Cz_1+D &= (A,B,C)(x_1,y_1,z_1)-(A,B,C)(x_0,y_0,z_0)\\ &= (A,B,C)\big((x_1,y_1,z_1)-(x_0,y_0,z_0)\big)=\mathbf{n}(\mathbf{r}_1-\mathbf{r}_0). \end{aligned}$$

Stąd i z (12.35) wynika, że odległość pomiędzy punktem $P_1(x_1,y_1,z_1)$ i płaszczyzną Π_0 określoną równaniem $Ax+By+Cz+D=0$ wyraża się wzorem

$$d(P_1,\Pi_0)=\frac{|Ax_1+By_1+Cz_1+D|}{\sqrt{A^2+B^2+C^2}}. \tag{12.37}$$

Przykład 288. Obliczyć odległość punktu $P_1(1,5,3)$ od: (a) płaszczyzny Π_1 przechodzącej przez punkty $A(1,0,0)$, $B(0,2,0)$ i $C(0,0,3)$; (b) płaszczyzny Π_2 określonej równaniem $2x+3y+6z-7=0$.

Płaszczyzna Π_1 przechodzi przez punkt $P_0=A(1,0,0)$ i jest ortogonalna do wektora

$$\mathbf{n}=AB\times AC=\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k}\\ -1 & 2 & 0\\ -1 & 0 & 3\end{vmatrix}=(6,3,2).$$

Zatem, wobec wzoru (12.35), odległość pomiędzy punktem P_1 i płaszczyzną Π_1 jest równa

$$\begin{aligned} d(P_1,\Pi_1)=\frac{|\mathbf{n}(\mathbf{r}_1-\mathbf{r}_0)|}{\|\mathbf{n}\|} &= \frac{|(6,3,2)\cdot\big((1,5,3)-(1,0,0)\big)|}{\sqrt{6^2+3^2+2^2}}\\ &= \frac{|(6,3,2)\cdot(0,5,3)|}{7}=\frac{21}{7}=3. \end{aligned}$$

Ponieważ płaszczyzna Π_2 jest określona równaniem ogólnym, więc odległość punktu P_1 od płaszczyzny Π_2 możemy wyznaczyć za pomocą wzoru (12.37),

$$d(P_1, \Pi_2) = \frac{|Ax_1 + By_1 + Cz_1 + D|}{\sqrt{A^2 + B^2 + C^2}} = \frac{|2 \cdot 1 + 3 \cdot 5 + 6 \cdot 3 - 7|}{\sqrt{2^2 + 3^2 + 6^2}} = \frac{28}{7} = 4.$$

Odległość dwóch prostych skośnych

Niech ℓ_1 i ℓ_2 będą dwiema prostymi skośnymi określonymi przez równania $\mathbf{r} = \mathbf{r}_1 + t\mathbf{m}_1$ oraz $\mathbf{r} = \mathbf{r}_2 + \lambda\mathbf{m}_2$ i niech P_1 oraz P_2 będą punktami, których wektorami wodzącymi są $\mathbf{r}_1$ i $\mathbf{r}_2$. Odległość $d = d(\ell_1, \ell_2)$ pomiędzy tymi prostymi jest równa odległości dowolnego punktu jednej prostej od płaszczyzny zawierającej drugą prostą i równoległej do obu rozważanych prostych (rys. 12.22). Przykładowo, jest ona równa odległości punktu P_2 od płaszczyzny $(\mathbf{m}_1 \times \mathbf{m}_2)(\mathbf{r} - \mathbf{r}_1) = 0$. Stąd i ze wzoru (12.35) otrzymujemy wzór na odległość pomiędzy prostymi skośnymi,

$$d(\ell_1, \ell_2) = \frac{|(\mathbf{m}_1 \times \mathbf{m}_2)(\mathbf{r}_2 - \mathbf{r}_1)|}{\|\mathbf{m}_1 \times \mathbf{m}_2\|}. \tag{12.38}$$

Przykład 289. Obliczyć odległość $d(\ell_1, \ell_2)$ pomiędzy prostymi skośnymi ℓ_1 i ℓ_2 określonymi przez równania

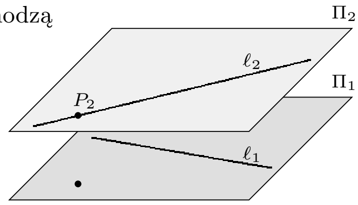
$$\frac{x+7}{3} = \frac{y-4}{4} = \frac{z+3}{-1} \quad \text{i} \quad \frac{x-21}{-6} = \frac{y-5}{4} = \frac{z-2}{1}.$$

Proste ℓ_1 i ℓ_2 są równoległe do wektorów $\mathbf{m}_1 = (3, 4, -1)$ i $\mathbf{m}_2 = (-6, 4, 1)$ i przechodzą przez punkty o wektorach wodzących $\mathbf{r}_1 = (-7, 4, -3)$ i $\mathbf{r}_2 = (21, 5, 2)$. Ponieważ

$$\mathbf{m}_1 \times \mathbf{m}_2 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & 4 & -1 \\ -6 & 4 & 1 \end{vmatrix} = (8, 3, 36) \quad \text{i} \quad \mathbf{r}_2 - \mathbf{r}_1 = (28, 1, 5),$$

więc ze wzoru (12.38) znajdujemy

$$d(\ell_1, \ell_2) = \frac{|(8, 3, 36) \cdot (28, 1, 5)|}{\|(8, 3, 36)\|} = 11.$$



Rys. 12.22

12.4. Ćwiczenia

1. Obliczyć $\mathbf{a} \times \mathbf{b}$, gdy: (a) $\mathbf{a} = (2, 1, 3)$ i $\mathbf{b} = (0, 1, 2)$; (b) $\mathbf{a} = (2, -3, 1)$ i $\mathbf{b} = (-2, 3, -1)$.
2. Obliczyć pole równoległoboku zbudowanego na wektorach $\mathbf{a} = (2, -3, 5)$ i $\mathbf{b} = (4, 1, 0)$.
3. Obliczyć pole równoległoboku, którego trzema wierzchołkami są $A(1, 2, -1)$, $B(2, 1, 4)$ i $C(3, 5, 2)$.
4. Obliczyć pole równoległoboku, którego wierzchołkami są punkty $(-3, 0, 2)$, $(6, 1, 4)$, $(4, 2, 2)$ i $(-5, 1, 0)$.
5. Obliczyć pole trójkąta, którego wierzchołkami są: (a) $A(1, 2, 1)$, $B(2, 1, -3)$ i $C(0, 1, 5)$; (b) $A(1, 4)$, $B(3, 2)$ i $C(-1, 2)$.
6. Obliczyć wysokość h trójkąta o wierzchołkach $A(1, 0, 0)$, $B(0, 2, 0)$, $C(0, 0, 3)$ poprowadzoną z wierzchołka C na bok AB.
7. Obliczyć pole wielokąta o wierzchołkach w punktach: (a) $(1, 1)$, $(4, 2)$, $(3, 4)$ i $(2, 4)$; (b) $(2, 1)$, $(6, 1)$, $(7, 3)$, $(5, 5)$ i $(3, 4)$.
8. Wyznaczyć objętość czworościanu, którego wierzchołkami są: (a) $(0, 0, 0)$, $(1, 1, 2)$, $(-1, 2, -1)$, $(0, -1, 3)$; (b) $(1, 3, -1)$, $(2, 2, 3)$, $(3, 7, 4)$, $(4, 2, -2)$.
9. Wyznaczyć ogólną postać wektora $\mathbf{a} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ takiego, że $\mathbf{a} \times (\mathbf{i} + \mathbf{j} + 2\mathbf{k}) = \mathbf{i} + \mathbf{j} - \mathbf{k}$.
10. Dane są wektory $\mathbf{a} = x\mathbf{i} - 6\mathbf{j} - 3\mathbf{k}$, $\mathbf{b} = 4\mathbf{i} + 3\mathbf{j} - \mathbf{k}$ i $\mathbf{c} = \mathbf{i} - 3\mathbf{j} + 2\mathbf{k}$. (a) Wyznaczyć x takie, że $\mathbf{a}$ i $\mathbf{b}$ są ortogonalne. (b) Wyznaczyć x takie, że $\mathbf{a}$, $\mathbf{b}$ i $\mathbf{c}$ leżą w jednej płaszczyźnie. (c) Obliczyć $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$, gdy $x = 2$.
11. Dane są wektory $\mathbf{a}$, $\mathbf{b}$ i $\mathbf{c}$ z przestrzeni R^3, gdzie $\mathbf{a}$ i $\mathbf{b}$ nie są równoległe. (a) Pokazać, że istnieją liczby x i y takie, że $\mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = x\mathbf{a} + y\mathbf{b}$. (b) Wywnioskować stąd, że $x(\mathbf{ca}) + y(\mathbf{cb}) = 0$. (c) Dla wektorów $\mathbf{c} = c_1\mathbf{i}$, $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j}$ i $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ obliczyć $\mathbf{c} \times (\mathbf{a} \times \mathbf{b})$ oraz $(\mathbf{cb})\mathbf{a} - (\mathbf{ca})\mathbf{b}$ i stąd wywnioskować, że $\mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = (\mathbf{cb})\mathbf{a} - (\mathbf{ca})\mathbf{b}$. (Dla matematyków: czy ostatnia równość jest prawdziwa dla każdych wektorów $\mathbf{a}$, $\mathbf{b}$ i $\mathbf{c}$ z przestrzeni R^3?)

12. Niech $\mathbf{a}$ będzie wektorem niezerowym. Pokazać, że jeśli $\mathbf{b}$ jest wektorem takim, że $\mathbf{a}\times\mathbf{b}=\mathbf{0}$ i $\mathbf{a}\cdot\mathbf{b}=0$, to $\mathbf{b}=\mathbf{0}$.
13. Niech $\mathbf{a}$ i $\mathbf{b}$ będą różnymi niezerowymi wektorami z przestrzeni R^3. Rozwiązać równania: (a) $\mathbf{x}=\mathbf{a}\times\mathbf{x}$; (b) $\mathbf{x}-\mathbf{a}=\ \mathbf{a}\times\mathbf{z}$; (c) $\mathbf{a}\times\mathbf{x}=\mathbf{b}\times\mathbf{x}$; (d) $(\mathbf{a}\times\mathbf{x})\times(\mathbf{b}\times\mathbf{x})=\mathbf{0}$.
14. Napisać równanie prostej przechodzącej przez punkt $P(1,1,3)$ i równoległej do: (a) wektora $\mathbf{n}=(2,-3,5)$; (b) prostej $-(x-1)/3=(y+1)/2=-(z+1)$; (c) osi Ox.
15. Wyznaczyć punkt, w którym prosta przechodząca przez punkty $A(2,1,3)$ i $B(4,-1,5)$ przecina się z płaszczyzną Oxz.
16. Przez punkt $A(0,1,-1)$ poprowadzić prostą prostopadłą do prostych $(x+3)/2=(y-5)/3=-(z+8)$ i $x=4-t$, $y=5+t$, $z=3t$.
17. Obliczyć odległość d punktu $A(2,1,0)$ od prostej $x+1=y-1=-(z-3)/3$.
18. Obliczyć odległość punktu $A(3,1,2)$ od prostej $\frac{x-2}{3}=\frac{y-3}{2}=\frac{z-1}{2}$.
19. Znaleźć rzut ortogonalny A' punktu $A(2,1,-3)$ na prostą $x=2+2t$, $y=8+5t$, $z=2-t$.
20. Znaleźć rzut ortogonalny A' punktu $A(1,1,19)$ na prostą $\frac{x-2}{3}=\frac{y-3}{4}=\frac{z-1}{2}$. Obliczyć odległość między punktami A i A'.
21. Dana jest prosta ℓ przechodząca przez punkty $A(2,4,9)$ i $B(4,6,7)$. Znaleźć odległość d prostej ℓ od punktu $C(0,0,0)$ i wyznaczyć rzut ortogonalny C' punktu C na prostą ℓ.
22. Płaszczyzny $x+y+z=0$ i $x-3y+9z-28=0$ przecinają się wzdłuż prostej. Napisać równanie tej prostej.
23. Napisać równanie parametryczne prostej przecięcia płaszczyzn $4x-5y-2z+3=0$ i $x+4y+3z-8=0$.
24. Na prostej ℓ, wzdłuż której przecinają się płaszczyzny $x+y-2z=1$ i $x+3y-z=4$, wskazać punkt B najbliższy punktowi $A(1,2,4)$.
25. Znaleźć długość d rzutu ortogonalnego odcinka łączącego punkty $A(5,2,3)$ i $B(0,1,-7)$ na prostą przechodzącą przez punkty $C(1,2,0)$ i $D(7,2,-8)$.
26. Proste ℓ_1 i ℓ_2 określone są przez równania $\frac{x-2}{3}=\frac{y+3}{-4}=\frac{z-4}{12}$ i $\frac{x-1}{4}=\frac{y-5}{0}=\frac{z+3}{3}$. Znaleźć cosinus kąta pomiędzy prostymi ℓ_1 i ℓ_2 oraz odległość prostej ℓ_1 od prostej ℓ_2.
27. Obliczyć odległość d pomiędzy prostymi skośnymi $(x-2)/3=(y-8)/4=-(z+6)$ i $-(x-9)/6=(y-3)/4=z-4$.
28. Dana jest prosta przechodząca przez punkty $A(1,1,5)$ i $B(-2,1,2)$ i druga przechodząca przez punkty $C(2,2,1)$ i $D(1,-2,1)$. Obliczyć odległość d między tymi prostymi.
29. Znaleźć punkty A i B leżące odpowiednio na prostych $(x,y,z)=(-1,7,1)+t(-2,5,1)$ i $(x,y,z)=(3,1,4)+s(3,0,1)$, odległość pomiędzy którymi jest równa odległości pomiędzy prostymi.
30. Dane są proste $(x,y,z)=(1,1,1)+t(1,-1,1)$ i $(x,y,z)=(1,-1,1)+s(-1,1,1)$. Znaleźć cosinus kąta nachylenia oraz odległość pomiędzy tymi prostymi.
31. Zbadać wzajemne położenie prostych $-(x+1)=y-1=(z-1)/2$ i $x/2=y=-z$ i obliczyć kąt między tymi prostymi.
32. Przez punkt $P(1,2,3)$ poprowadzić prostą, która pod kątem prostym przecina prostą $x=3+2t$, $y=-4-t$, $z=1+5t$.
33. Wskazać równanie prostej przecinającej proste
$$\frac{x-1}{-1}=\frac{y+2}{2}=\frac{z-3}{1} \quad \text{i} \quad \frac{x+1}{3}=\frac{y-1}{2}=\frac{z+1}{-1}$$
i przechodzącej przez punkt $(0,0,0)$.
34. Wyznaczyć punkt P przecięcia się prostych o równaniach $-x=-(y+5)/7=(z-7)/4$ i $(x+1)/2=(y-2)/0=(z+3)/6$.
35. Wskazać punkt przecięcia się prostych $\frac{x-1}{4}=\frac{y-2}{5}=\frac{z-3}{6}$ i $\frac{x-4}{1}=\frac{y-5}{2}=\frac{z-6}{3}$. Następnie napisać równanie płaszczyzny zawierającej obie proste.
36. Wyznaczyć równanie płaszczyzny równoległej do osi Oy, prostopadłej do płaszczyzny $2x-y+5z=0$ i przechodzącej przez punkt $A(1,2,3)$.
37. Znaleźć równanie płaszczyzny przechodzącej przez punkt $A(2,0,1)$, prostopadłej do płaszczyzny $x+5z-1=0$ i równoległej do prostej $\frac{x-1}{4}=\frac{y}{-1}=\frac{z-2}{2}$.
38. Podać równanie parametryczne rzutu ortogonalnego prostej $\frac{x-1}{1}=\frac{y-1}{0}=\frac{z-2}{1}$ na płaszczyznę $x+2y+3z-1=0$.
39. Napisać równanie prostej przechodzącej przez punkt $(0,0,0)$ i równoległej do płaszczyzn $x+y-2z=1$ i $3x-y+7z=2$.
40. Napisać równanie płaszczyzny przechodzącej przez punkty $(1,0,1)$, $(-1,1,1)$ i $(5,4,-3)$.
41. Pokazać, że punkty $(-5,3,3)$, $(-1,-2,-2)$, $(2,8,3)$ i $(3,4,0)$ leżą w jednej płaszczyźnie. Wyznaczyć równanie tej płaszczyzny.
42. Napisać równanie płaszczyzny przechodzącej przez punkt $P(-1,2,4)$ i prostopadłej do płaszczyzn $6x-2y+3z-\ 12=0$ i $3x+2y-6z+21=0$.
43. Napisać równanie prostej przechodzącej przez punkt $A(6,1,-2)$ i prostopadłej do płaszczyzny $3x-2y+z-4=\ 0$.
44. Jakie jest położenie prostej $(x,y,z)=(1+t,2-3t,1+2t)$ względem płaszczyzny $3x+5y+6z-19=0$?
45. Wyznaczyć odległość punktu $(1,1,1)$ od prostej, wzdłuż której przecinają się płaszczyzny $2x-y+z=3$ i $x-3y+3z=4$.
46. Obliczyć odległość d punktu $P(2,-1,6)$ od płaszczyzny $7x-4y+4z-6=0$.
47. Znaleźć symetryczne odbicie A' punktu $A(4,21,2)$ względem płaszczyzny $2x+3y-4z-5=0$.
48. Znaleźć punkt A' symetryczny do punktu $A(7,5,-3)$ względem płaszczyzny przechodzącej przez punkt $B(3,2,2)$ i równoległej do wektorów $\mathbf{m}=(1,2,2)$ i $\mathbf{n}=(3,1,2)$.
49. Wyznaczyć odległość punktu $P_0(3,-1,2)$ od płaszczyzny $(x,y,z)=(3,1,-2)+t(1,-1,1)+s(1,1,-1)$.
50. Obliczyć odległość d między płaszczyznami $2x-10y+11z-15=0$ i $2x-10y+11z+15=0$.
51. Napisać równania płaszczyzn dwusiecznych kątów dwuściennych między płaszczyznami $x+2y+2z-1=0$ i $4x+4y+7z+1=0$.

52. Wektorami wodzącymi punktów A, B i C są $\mathbf{a} = (2, -4, -3)$, $\mathbf{b} = (6, 0, 4)$ i $\mathbf{c} = (-2+4t, 1+4t, 8+7t)$. ($a$) Obliczyć pole trójkąta ABC. (b) Dlaczego pole trójkąta ABC nie zależy od t? (c) Obliczyć pole trójkąta $A'B'C'$ będącego rzutem prostopadłym trójkąta ABC na płaszczyznę $2x - y + 2z = 0$.
53. Wyznaczyć zbiór tych punktów, które leżą w płaszczyźnie $x - 2y + 3z = 0$ i są równooddalone od punktów $A(1, -3, 4)$ i $B(-1, 1, 0)$.
54. Napisać równanie płaszczyzny przechodzącej przez punkty $A(1, 2, 3)$ i $B(3, -1, 4)$ i prostopadłej do płaszczyzny $x - 3y + 2z + 4 = 0$.
55. Obliczyć kąt nachylenia płaszczyzn: (a) $x-y-2z = 4$ i $2x + y - z = 5$; (b) $2x - 10y + 11z - 1 = 0$ i $4x + 4y - 7z + 2 = 0$.
56. Napisać równanie prostej przechodzącej przez punkt $A(0, 1, 6)$, równoległej do płaszczyzny $2x-y+3z+4 = 0$ i przecinającej prostą $\frac{x-1}{2} = \frac{y+2}{3} = \frac{z}{4}$.
57. Wyznaczyć płaszczyznę przechodzącą przez punkt $(2, -1, -3)$ i przez prostą przecięcia płaszczyzn $3x + 2y - 4z = 7$ i $6x - 3y + 2z = 4$.
58. Napisać równanie parametryczne prostej będącej rzutem ortogonalnym prostej $\frac{x-1}{2} = \frac{y+2}{3} = \frac{z}{4}$ na płaszczyznę $2x - y + 3z + 4 = 0$.
59. Napisać równanie płaszczyzny zawierającej prostą $x-2 = (y + 3)/2 = (z - 4)/3$ i jej ortogonalny rzut na płaszczyznę $3x - y + 2z + 1 = 0$.
60. Wpisując TAK albo NIE, stwierdzić prawdziwość każdego z następujących zdań:

$\boxed{1}$ Załóżmy, że wektory $\mathbf{b}$ i $\mathbf{c}$ są równoległe. Czy wektory $\mathbf{a} \times \mathbf{b}$ i $\mathbf{a} \times \mathbf{c}$ są równoległe?

$\boxed{2}$ Niech wektory $\mathbf{b}$ i $\mathbf{c}$ będą ortogonalne. Czy w takim przypadku wektory $\mathbf{a} \times \mathbf{b}$ i $\mathbf{a} \times \mathbf{c}$ także będą ortogonalne?

$\boxed{3}$ Iloczyn mieszany wektorów jest wektorem.

$\boxed{4}$ Dla każdych $\mathbf{a}, \mathbf{b} \in R^3$ jest $(\mathbf{a} + \mathbf{b}) \times (\mathbf{a} - \mathbf{b}) = 2(\mathbf{b} \times \mathbf{a})$.

$\boxed{5}$ Dla każdych $\mathbf{a}, \mathbf{b} \in R^3$ jest $\|\mathbf{a} \cdot (\mathbf{a} \times \mathbf{b})\| = \|\mathbf{a}\|^2\|\mathbf{b}\| \sin\measuredangle(\mathbf{a}, \mathbf{b})$.

$\boxed{6}$ Dla $\mathbf{a}, \mathbf{b} \in R^3$ jest $\|\mathbf{a}\times\mathbf{b}\|^2 + |\mathbf{a}\cdot\mathbf{b}|^2 = \|\mathbf{a}\|^2\|\mathbf{b}\|^2$.

$\boxed{7}$ Dla każdego $\mathbf{a} \in R^3$ jest $\|\mathbf{a} \times \mathbf{a}\| = \|\mathbf{a}\|^2$.

$\boxed{8}$ Jeśli kąt między wektorami $\mathbf{a}$ i $\mathbf{b}$ z przestrzeni R^3 jest równy $\pi/4$, to $\|\mathbf{a} \times \mathbf{b}\| = |\mathbf{a} \cdot \mathbf{b}|$.

$\boxed{9}$ Dla każdych $\mathbf{a}, \mathbf{b}, \mathbf{c} \in R^3$ jest $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$.

BIBLIOGRAFIA

[1] Banaszak G., Gajda W.: Elementy algebry liniowej, cz. I i II. Warszawa: Wydawnictwa Naukowo-Techniczne 2002.
[2] Białynicki-Birula A.: Algebra liniowa z geometrią. Warszawa: Państwowe Wydawnictwo Naukowe 1979.
[3] Fichtenholz G. M.: Rachunek różniczkowy i całkowy, tom 2. Warszawa: Państwowe Wydawnictwo Naukowe 1966.
[4] Friedberg S. H., Insel A. J., Spence L. E.: Linear Algebra. New Jersey: Prentice Hall 1999.
[5] Gleichgewicht B.: Algebra. Wrocław: Oficyna Wydawnicza GiS 2002.
[6] Jurlewicz T., Skoczylas Z.: Algebra liniowa 1. Definicje, twierdzenia, wzory. Wrocław: Oficyna Wydawnicza GiS 2003.
[7] Jurlewicz T., Skoczylas Z.: Algebra liniowa 2. Definicje, twierdzenia, wzory. Wrocław: Oficyna Wydawnicza GiS 2003.
[8] Jurlewicz T., Skoczylas Z.: Algebra liniowa 1. Przykłady i zadania. Wrocław: Oficyna Wydawnicza GiS 2003.
[9] Jurlewicz T., Skoczylas Z.: Algebra liniowa 2. Przykłady i zadania. Wrocław: Oficyna Wydawnicza GiS 2003.
[10] Kaczorek T.: Wektory i macierze w automatyce i elektrotechnice. Warszawa: Wydawnictwa Naukowo-Techniczne 1998.
[11] Klukowski J., Nabiałek I.: Algebra dla studentów. Warszawa: Wydawnictwa Naukowo-Techniczne 1999.
[12] Kostrykin A.: Wstęp do algebry. Warszawa: Państwowe Wydawnictwo Naukowe 1984.
[13] Mostowski A., Stark M.: Elementy algebry wyższej. Warszawa: Państwowe Wydawnictwo Naukowe 1965.
[14] Rudin W.: Podstawy analizy matematycznej. Warszawa: Państwowe Wydawnictwo Naukowe 1969.
[15] Sołtysiak A.: Algebra liniowa. Poznań: Wydawnictwo Naukowe UAM 1996.

Skorowidz

algorytm
- Gaussa-Jordana, 90
- Grama-Schmidta, 198
- Hornera, 45

argument liczby zespolonej, 25–27

baza
- ortogonalna, 196, 199, 202
- ortonormalna, 196, 197, 199, 201, 202, 211
- standardowa, 196

baza przestrzeni wektorowej, 135–137, 139, 140, 150, 196, 202, 221, 223, 226, 237
Bessela nierówność, 215
bijekcja, 10

Cauchy'ego-Minkowskiego nierówność, 194
centrum grupy, 17
ciąg Fibonacciego, 244
ciąg rekurencyjny, 242
ciało, 16, 18
ciało
- ciało Z_n, 17
- liczb rzeczywistych, 17
- liczb wymiernych, 17
- liczb zespolonych, 19
- skalarów, 120

cosinus kąta pomiędzy prostymi, 272

długość wektora, 193, 211
dodawanie modulo n, 12
dopasowanie prostej, 208–210
dopełnienie algebraiczne, 108
działanie
- łączne, 7
- dwuargumentowe, 7
- modulo n, 11
- przemienne, 7
- zewnętrzne, 120

dzielenie wielomianów z resztą, 42
dzielnik zera, 15, 18

element
- neutralny, 8
- odwracalny, 8, 13
- odwrotny, przeciwny, 8

endomorfizm, 159
epimorfizm, 169, 170

forma kwadratowa, 250–264
forma kwadratowa
- dodatnio (pół)określona, 259, 261, 262
- nieokreślona, 259, 261, 262
- postać kanoniczna, 252–259
- postać macierzowa, 250
- ujemnie (pół)określona, 259, 261, 262

funkcja
- odwracalna, 180
- odwrotna, 180
- wielomianowa, 42
- wymierna, 54–63

generator grupy, 14
generator przestrzeni, 127, 129, 134
granica ciągu macierzy, 234, 235
grupa, 10–14, 18
grupa
- cykliczna, 14, 18
- macierzy, 68
- podgrupa, 13, 18
- podgrupa cykliczna, 14
- podgrupa trywialna, 13
- przemienna, 10, 121
- reszt modulo n, 12
- symetryczna, 10

Hornera schemat, 44

iloczyn
- macierzy, 68–79
- mieszany wektorów, 267
- przekształcenia liniowego przez skalar, 171
- skalarny wektorów, 191, 211
- wektora przez skalar, 120
- wektorowy wektorów, 265–267
- wielomianów, 39

iloczyn macierzy przez skalar, 67
izomorfizm ciał, 38
izomorfizm przestrzeni wektorowych, 143, 144, 179

jądro przekształcenia liniowego, 165, 181
jednokładność, 179
jednostka urojona, 21

kąt
- nachylenia dwóch prostych, 272
- pomiędzy płaszczyznami, 278
- pomiędzy prostą i płaszczyzną, 277

pomiędzy wektorami, 195, 211
kolumna macierzy, 64
kombinacja Fouriera, 196–198, 201, 202
kombinacja liniowa wektorów, 125, 132–133, 136
krawędź przecięcia się płaszczyzn, 279
krotność wartości własnej
algebraiczna, 225, 226
geometryczna, 225, 226

Lagrange'a metoda, 255
liczba zespolona, 19–38
liczba zespolona
część rzeczywista, 19
część urojona, 19
postać kanoniczna, 22
postać trygonometryczna, 26
postać wykładnicza, 35

macierz, 64–79
macierz
diagonalizowalna, 221–224
diagonalna, 66, 105, 222, 223, 232, 233
dołączona, 112
dodatnio (pół)określona, 259, 261, 262
elementarna, 83–86, 94, 111
formy kwadratowej, 251
główna układu, 80
Heisenberga, 79
idempotentna, 78, 206
inwolującą, 78
jednostkowa, 66, 105
kwadratowa, 65
Markowa, 79
nieokreślona, 259, 261, 262
nieosobliwa, 112, 113, 129, 134
nilpotentna, 78, 248
obrotu płaszczyzny, 174
odwracalna, 74, 84, 94, 113, 134, 151, 182, 210
odwrotna, 75, 95–97, 112, 113, 182, 242
okresowa, 78
operatora liniowego, 220
ortogonalna, 210–212, 229, 232
ortogonalnie diagonalizowalna, 229, 231
podobieństwa, 223, 232
przejścia, 145–149, 177, 178
przekształcenia liniowego, 173, 177, 182
rozszerzona układu, 80
rzutu ortogonalnego, 204–206
schodkowa, 86
schodkowa normalna, 86
skośnie symetryczna, 73
symetryczna, 73, 206, 228, 229, 231, 232
trójkątna, 78, 105
transponowana, 73, 106, 210
ujemnie (pół)określona, 259, 261, 262
wierszowo równoważna, 85, 86, 94, 149
złożenia przekształceń liniowych, 178
zerowa, 65
zredukowana formy kwadratowej, 257
metoda
Gaussa i Gaussa-Jordana rozwiązywania układu równań liniowych, 90
Grama-Schmidta, 198, 199
Lagrange'a, 255
najmniejszych kwadratów, 208
zakrywania, 59
minor główny, 218, 262
mnożenie macierzy, 68
mnożenie modulo n, 12
moduł liczby zespolonej, 23
monomorfizm, 169–171

najlepsze rozwiązanie układu równań, 207, 208
nierówność
Bessela, 215
Cauchy'ego-Minkowskiego, 194
Schwarza, 194, 215
trójkąta, 194
norma wektora, 193
normalny układ równań, 208

objętość czworościanu, 269
objętość równoległościanu, 268
obrót płaszczyzny, 174, 179
obraz podprzestrzeni, 165
obraz przekształcenia liniowego, 165
odcinek, 271
odległość
pomiędzy prostymi, 281
pomiędzy wektorami, 193
punktu od płaszczyzny, 280
punktu od prostej, 274
odwracalność
operatora liniowego, 181
przekształcenia liniowego, 181
odwzorowanie tożsamościowe, 10
okres macierzy, 78
operacje elementarne
na równaniach układu, 81
na wierszach macierzy, 82, 150
operator
liniowy odwracalny, 182
operator liniowy, 159, 220, 223
operator liniowy
diagonalizowalny, 221, 222, 224, 226
odwracalny, 182

ortogonalizacja bazy, 198–199
ortogonalne dopełnienie
    (pod)przestrzeni, 200, 201
    zbioru wektorów, 200
ortogonalność wektorów, 195

płaszczyzna, 269
płaszczyzna
    k-wymiarowa, 269
    zespolona Gaussa, 22
Parsevala równość, 215
pierścień, 14
pierścień
    przemienny, 15
    wielomianów, 41
    z jedynką, 15
pierwiastek
    n-tego stopnia z jedności, 30
    n-tego stopnia z liczby zespolonej, 29–33
    wielokrotny wielomianu, 46
    wielomianu, 46–52
podobieństwo macierzy, 77, 183, 184, 186, 223
podprzestrzeń
    cykliczna, 236, 237
    generowana, 127
    niezmiennicza, 236
    przestrzeni wektorowej, 123, 152
    właściwa, 123
    zerowa, 123
podstawienie ortogonalne, 253
podzielność wielomianu, 42
pole równoległoboku, 266
pole trójkąta, 267
potęga elementu, 11
potęga macierzy, 76, 233, 240
prosta, 269
proste skośne, 272
prostopadłość prostych, 272
prostopadłość wektorów, 195
przeciwobraz podprzestrzeni, 165
przekształcenie
    liniowe, 159
    liniowe odwrotne, 181
    na, 169, 180
    ortogonalne, 212
    różnowartościowe, 169, 180
    tożsamościowe, 162, 177
    zerowe, 162
przestrzeń
    n-wymiarowa, 139
    (nie)skończenie wymiarowa, 139
    Euklidesa, 191
    kolumnowa macierzy, 128, 200
    macierzy, 179
    przekształceń liniowych, 171, 179
    skończenie generowana, 127
    własna macierzy, 225, 228, 231
    własna operatora, 225
    wektorowa ciągów n-elementowych, 121
    wektorowa funkcji, 121
    wektorowa macierzy, 121
    wektorowa wielomianów, 123
    wierszowa macierzy, 200
    zerowa macierzy, 124, 200
    zerowa przekształcenia, 165
przestrzeń wektorowa, 120

różnica macierzy, 67
równanie
    charakterystyczne macierzy, 218
    macierzowe, 92
    płaszczyzny
        ogólne, 277
        parametryczne, 275
        wektorowe, 274
        wyznacznikowe, 276
    prostej
        kierunkowe, 270
        parametryczne, 270
        wektorowe, 270
równość macierzy, 66
równość Parsevala, 215
równość wielomianów, 39, 48
równoległość prostych, 272
równoważne układy równań, 81, 86
równoważność macierzy, 183, 186
rozdzielność działania, 9
rozkład spektralny macierzy, 232
rozwiązanie układu równań
    Cramera, 116
    jedyne, 115
    liniowych, 81, 89, 97–99, 128
    najlepsze, 207
    niezerowe, 115
    zerowe, 133, 134
rząd
    formy kwadratowej, 251
    kolumnowy macierzy, 149, 150
    macierzy, 149–151
    przekształcenia liniowego, 167
    wierszowy macierzy, 149, 150
rząd elementu, 18
rzut ortogonalny, 201–204, 207

schemat Hornera, 44, 45
Schwarza nierówność, 194, 215
sinus kąta pomiędzy
    prostymi, 272
    wektorami, 266
skracanie w grupie, 11
sprzężenie liczby zespolonej, 23
standardowy iloczyn skalarny, 192
stopień wielomianu, 39
suma
    macierzy, 66
    podprzestrzeni, 152
    prosta podprzestrzeni, 153, 154, 201

przekształceń liniowych, 171
wektorów, 120
wielomianów, 39
symetria względem prostej, 179

ślad macierzy, 76

twierdzenie
Bézout, 46
Cauchy'ego, 111
Cayleya-Hamiltona, 239, 241
Kroneckera-Capellego, 151
Laplace'a, 104
o dzieleniu wielomianów, 42
o najlepszej aproksymacji, 207
o reszcie, 44
Pitagorasa, 196
spektralne, 231
Steinitza, 139
Sylvestera, 167, 262
wymiarowe, 167

ułamek prosty, 55
układ równań liniowych, 80
układ równań liniowych
(nie)sprzeczny, 81
Cramera, 116
postać macierzowa, 81
postać wektorowa, 81
układ wektorów, 131
układ wektorów
ortogonalnych, 196, 197
ortonormalnych, 196, 198

wartość własna
macierzy, 216, 217, 220, 223, 228, 231
operatora, 216, 220, 222, 223, 226
wektor
błędu, 206
kierunkowy prostej, 270
normalny płaszczyzny, 275
własny
macierzy, 216, 217, 220, 223, 228
operatora, 216, 220, 222, 223
współrzędnych, 142
wektory
liniowo (nie)zależne, 131–135, 144, 197
ortogonalne, 228
ortogonalne (prostopadłe), 195
widmo macierzy, 219
wielkość błędu aproksymacji, 206
wielomian, 39–63
wielomian
charakterystyczny
macierzy, 217, 242
operatora, 218, 224, 226, 236, 237
definicja, 39
nierozkładalny, 42
wielomiany
względnie pierwsze, 53
wiersz macierzy, 64
wiodąca
jedynka macierzy, 86
kolumna macierzy, 88, 150
współczynniki Fouriera, 196
współrzędne wektora, 142
wymiar przestrzeni wektorowej, 139, 141
wyznacznik
definicja, 102
dopełnienie algebraiczne, 108
Grama, 203
iloczynu macierzy, 111
macierzy ortogonalnej, 210
macierzy podobnych, 114
rozwinięcie Laplace'a, 103, 104
schemat Sarrusa, 103
twierdzenie Cauchy'ego, 111
Vandermonde'a, 117
własności, 103–114
wzór de Moivre'a, 27
wzory
Cramera, 116
Eulera, 34
Viète'a, 63

złożenie
przekształceń liniowych, 172
zależność rekurencyjna, 242
zasadnicze twierdzenie algebry, 48
zasadnicze twierdzenie algebry liniowej, 167
zbiór wektorów, 120
zbiór zamknięty
ze względu na dodawanie, 122
ze względu na mnożenie przez skalary, 122
zerowość przekształcenia liniowego, 167

Bibliografia

[1] Banaszak G., Gajda W.: Elementy algebry liniowej, cz. I i II. Warszawa: Wydawnictwa Naukowo-Techniczne 2002.

[2] Białynicki-Birula A.: Algebra liniowa z geometrią. Warszawa: Państwowe Wydawnictwo Naukowe 1979.

[3] Fichtenholz G. M.: Rachunek różniczkowy i całkowy, tom 2. Warszawa: Państwowe Wydawnictwo Naukowe 1966.

[4] Friedberg S. H., Insel A. J., Spence L. E.: Linear Algebra. New Jersey: Prentice Hall 1999.

[5] Gleichgewicht B.: Algebra. Wrocław: Oficyna Wydawnicza GiS 2002.

[6] Jurlewicz T., Skoczylas Z.: Algebra liniowa 1. Definicje, twierdzenia, wzory. Wrocław: Oficyna Wydawnicza GiS 2003.

[7] Jurlewicz T., Skoczylas Z.: Algebra liniowa 2. Definicje, twierdzenia, wzory. Wrocław: Oficyna Wydawnicza GiS 2003.

[8] Jurlewicz T., Skoczylas Z.: Algebra liniowa 1. Przykłady i zadania. Wrocław: Oficyna Wydawnicza GiS 2003.

[9] Jurlewicz T., Skoczylas Z.: Algebra liniowa 2. Przykłady i zadania. Wrocław: Oficyna Wydawnicza GiS 2003.

[10] Kaczorek T.: Wektory i macierze w automatyce i elektrotechnice. Warszawa: Wydawnictwa Naukowo-Techniczne 1998.

[11] Klukowski J., Nabiałek I.: Algebra dla studentów. Warszawa: Wydawnictwa Naukowo-Techniczne 1999.

[12] Kostrykin A.: Wstęp do algebry. Warszawa: Państwowe Wydawnictwo Naukowe 1984.

[13] Mostowski A., Stark M.: Elementy algebry wyższej. Warszawa: Państwowe Wydawnictwo Naukowe 1965.

[14] Rudin W.: Podstawy analizy matematycznej. Warszawa: Państwowe Wydawnictwo Naukowe 1969.

[15] Sołtysiak A.: Algebra liniowa. Poznań: Wydawnictwo Naukowe UAM 1996.