GEOMETRIA Z ALGEBRĄ LINIOWĄ

Skrypt dla studentów 1. roku informatyki

Wersja 0.6.2

29/01/2018

Paweł Bechler

WMIM, UW

Spis treści

Rozdział 1

Grupy i ciała

1.1 Działania i ich własności

Na początek zajmiemy się zbiorem liczb rzeczywistych $\mathbb{R}$. Na elementach tego zbioru określone są następujące operacje, zwane *działaniami*:

Działania jednoargumentowe. Możemy wyznaczyć *liczbę przeciwną* do danej:

$$x \mapsto -x$$

oraz, dla liczby różnej od zera, jej *odwrotność*:

$$x \mapsto \frac{1}{x} = x^{-1}.$$

Działania dwuargumentowe. Określone jest *dodawanie* oraz *mnożenie* dwóch liczb:

$$(x, y) \mapsto x + y, \qquad (x, y) \mapsto x \cdot y.$$

Celowo nie wymieniamy tutaj działań zwanych *odejmowaniem* i *dzieleniem*, gdyż można je otrzymać jako złożenia operacji brania liczby przeciwnej i dodawania oraz odwrotności i mnożenia.

Jak dobrze wszystkim wiadomo, działania dodawania i mnożenia mają następujące własności:

- Dla dowolnych $a, b \in \mathbb{R}$ zachodzi

  $$a + b = b + a \qquad \text{oraz} \qquad a \cdot b = b \cdot a.$$

  Mówimy, że działania te są *przemienne*.

- Dla dowolnych $a, b, c \in \mathbb{R}$ zachodzi
$$a + (b + c) = (a + b) + c \qquad \text{oraz} \qquad a \cdot (b \cdot c) = (a \cdot b) \cdot c.$$
Mówimy, że działania te są *łączne*.

- Dla dowolnych $a, b, c \in \mathbb{R}$ zachodzi
$$a \cdot (b + c) = a \cdot b + a \cdot c. \tag{1.1}$$
Mówimy, że mnożenie jest *rozdzielne* względem dodawania.

W zbiorze liczb rzeczywistych istnieją również pewne szczególne elementy, które oznaczamy symbolami 0 i 1. Dla dowolnej liczby $a \in \mathbb{R}$:
$$a + 0 = 0 + a = a \qquad \text{oraz} \qquad a \cdot 1 = 1 \cdot a = a.$$
Liczbę 0 nazywamy w związku z tym *elementem neutralnym* dodawania, i, analogicznie, liczba 1 jest zwana *elementem neutralnym* mnożenia. Elementy neutralne mają również następujący związek z operacjami wyznaczania liczby przeciwnej i odwrotności:
$$a + (-a) = 0 \qquad \text{oraz} \qquad a \cdot \frac{1}{a} = 1.$$

1.2 Grupy

Wyobraźmy sobie teraz następujące zbiory liczbowe:

- $\mathbb{N}$ - liczby naturalne (wraz z zerem),
- $\mathbb{Z}$ - liczby całkowite,
- $\mathbb{R}^*$ - liczby rzeczywiste różne od zera,
- $\mathbb{R}^*_+$ - liczby rzeczywiste dodatnie,
- $\mathbb{R}$ - liczby rzeczywiste.

Są to podzbiory zbioru liczb rzeczywistych, co oznacza, że na elementach tych zbiorów także możemy wykonywać cztery wspomniane wcześniej operacje. Nie jest jednak prawdą, że wynik takiej operacji musi należeć do tego samego zbioru, co liczby wejściowe. Jeżeli jednak tak jest, to mówimy, że dany zbiór *jest zamknięty* ze względu na rozważane działanie.

Jeżeli zbiór jest jednocześnie zamknięty ze względu na dane działanie dwuargumentowe (dodawanie lub mnożenie) i odpowiadające mu działanie jednoargumentowe (liczba przeciwna lub odwrotna), i dodatkowo element neutralny danego działania też należy do tego zbioru, to mówimy, że taki zbiór jest *grupą*. Formalna definicja jest następująca:

Tabela 1.1: Zamkniętość wybranych podzbiorów zbioru liczb rzeczywistych ze względu na 4 podstawowe działania

działanie	$\mathbb{N}$	$\mathbb{Z}$	$\mathbb{R}^*$	$\mathbb{R}^*_+$	$\mathbb{R}$
$-a$	NIE	TAK	TAK	NIE	TAK
$a+b$	TAK	TAK	NIE	TAK	TAK
a^{-1}	NIE	NIE	TAK	TAK	NIE
$a \cdot b$	TAK	TAK	TAK	TAK	TAK

Definicja 1.1. Zbiór G wraz z

- elementem wyróżnionym $e \in G$,
- działaniem dwuargumentowym „$\diamond$" o własnościach

$$\forall_{x,y,z \in G} \qquad x \diamond (y \diamond z) = (x \diamond y) \diamond z \tag{G1}$$

$$\forall_{x \in G} \qquad e \diamond x = x \diamond e = x \tag{G2}$$

$$\forall_{x \in G} \ \exists_{y \in G} \qquad x \diamond y = e. \tag{G3}$$

nazywamy *grupą* z *działaniem* „$\diamond$" i *elementem neutralnym* e. Formalnie, grupę możemy zapisać jako trójkę uporządkowaną $(G, \diamond, e)$.

Zgodnie z obserwacjami zawartymi w tabeli 1.1, mamy następujące przykłady grup:

- $(\mathbb{Z}, +, 0)$,
- $(\mathbb{R}, +, 0)$,
- $(\mathbb{R}^*, \cdot, 1)$,
- $(\mathbb{R}^*_+, \cdot, 1)$,

Podana definicja grupy nie wymaga, aby działanie grupowe było przemienne. Jeżeli jednak tak jest, czyli spełniony jest dodatkowy warunek:

$$\forall_{x,y \in G} \qquad x \diamond y = y \diamond x \tag{G4}$$

to taką grupę G nazywamy grupą *abelową* lub *przemienną*.

Wyżej wymienione grupy są przemienne i mają nieskończenie wiele elementów. Dalej podajemy przykłady grup, które nie spełniają co najmniej jednego z tych warunków.

Przykład 1.1 (Grupy addytywne $\mathbb{Z}_n$)**.** Niech $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$. Na zbiorze $\mathbb{Z}_n$ rozważamy działanie „+” jako dodawanie modulo n. Wówczas $(\mathbb{Z}_n, +, 0)$ jest n-elementową grupą abelową.

Przykład 1.2 (Grupy multiplikatywne $\mathbb{Z}_n^*$)**.** Niech $\mathbb{Z}_n^* = \mathbb{Z}_n \setminus \{0\}$. Na zbiorze $\mathbb{Z}_n^*$ rozważamy działanie „$\cdot$” jako mnożenie modulo n. Wówczas, jeżeli n jest liczbą pierwszą, to $(\mathbb{Z}_n^*, \cdot, 1)$ jest grupą abelową.

Zadanie 1.1. Udowodnij powyższe stwierdzenie. Czy jest prawdziwe twierdzenie odwrotne?

Przykład 1.3 (Grupy permutacji)**.** Niech A_n oznacza zbiór n-elementowy, np. $A_n = \{0, 1, \ldots, n-1\}$ i rozważmy zbiór wszystkich permutacji elementów zbioru A_n. Każdą taką permutację p możemy traktować jako bijekcję zbioru A_n w siebie. Przykładowo, permutacji $(2, 1, 0)$ zbioru $A_3 = \{0, 1, 2\}$ odpowiada bijekcja $f : A_3 \to A_3$ taka, że $f(0) = 2, f(1) = 1, f(2) = 0$. Niech S_n oznacza zbiór wszystkich takich bijekcji zbioru A_n. Na zbiorze S_n określamy działanie „$\circ$” polegające na składaniu odwzorowań, czyli

$$(f \circ g)(x) = f(g(x)) \qquad \text{dla } x \in A.$$

Tak określone działanie jest łączne, posiada ono element neutralny, czyli odwzorowanie identycznościowe $\mathrm{id}(x) = x$ i dla każdego elementu zbioru S_n możemy znaleźć odpowiadający mu element odwrotny (odwrócenie danej bijekcji). Zatem zbiór S_n jest grupą.

Zadanie 1.2. Czy działanie w grupie S_n jest przemienne?

Zadanie 1.3. Czy istnieje grupa nieabelowa mająca 2 lub 3 elementy?

Zadanie 1.4. Jaka jest najmniejsza (w sensie liczby elementów) grupa nieprzemienna?

Na zakończenie poświęćmy chwilę uwagi aksjomatowi (G3). Nie mówi on czy, skoro $x \diamond y = e$, to także $y \diamond x = e$ – inaczej mówiąc, czy lewostronna odwrotność elementu x jest także jego prawostronną odwrotnością. W odróżnieniu od przemienności, jest to zawsze prawda. Nie wiemy na razie także, czy dany element może posiadać więcej niż jedną odwrotność – definicja grupy nic o tym nie mówi. Wszelkie wątpliwości rozwiewa jednak następujące

Stwierdzenie 1.1. *Załóżmy, że $(G, \diamond, e)$ jest grupą i $x \in G$. Wówczas istnieje dokładnie jeden element $y \in G$ taki, że $x \diamond y = e$ i wówczas także $y \diamond x = e$.*

Dowód. Istnieją elementy $y, z \in G$ takie, że

$$x \diamond y = e = y \diamond z,$$

co wynika z (G3). Korzystając z (G2) i (G1) otrzymujemy

$$z = e \diamond z = (x \diamond y) \diamond z = x \diamond (y \diamond z) = x \diamond e = x.$$

Czyli $x \diamond y = y \diamond x = e$ i $x' = y$ jest jednocześnie elementem lewo- i prawostronnie odwrotnym do x.

Jeżeli natomiast $x \diamond y = e = x \diamond y'$, to możemy napisać

$$y' = e \diamond y' = (x \diamond y) \diamond y' = (y \diamond x) \diamond y' = y \diamond (x \diamond y') = y \diamond e = y,$$

co oznacza, że element odwrotny jest tylko jeden. □

Zadanie 1.5 (Prawo skracania)**.** Załóżmy, że $(G, \diamond, e)$ jest grupą, $x, y, z \in G$ oraz

$$x \diamond z = y \diamond z.$$

Pokaż, że wówczas $x = y$.

1.3 Ciała

W poprzednim podrozdziale dowiedzieliśmy się, że w zbiorze liczb rzeczywistych $\mathbb{R}$ mamy do czynienia z dwiema grupami, jedną z dodawaniem, drugą z mnożeniem:

$$(\mathbb{R}, +, 0) \qquad \text{oraz} \qquad (\mathbb{R}^*, \cdot, 1).$$

Taki opis struktury zbioru liczb rzeczywistych nie uwzględnia jednak tego, że działanie mnożenia jest rozdzielne względem dodawania, jak to udało nam się wcześniej zauważyć w (1.1). Tutaj na odsiecz przychodzi pojęcie *ciała*.

Definicja 1.2. *Ciałem* nazywamy zbiór X wraz z działaniami „+” i „·”, zwanymi odp. *dodawaniem* i *mnożeniem*, oraz wyróżnionymi elementami 0 i 1 taki, że

$$(X, +, 0) \text{ jest grupą abelową,} \tag{C1}$$

$$(X \backslash \{0\}, \cdot, 1) \text{ jest grupą abelową,} \tag{C2}$$

oraz

$$\text{mnożenie jest rozdzielne względem dodawania,} \tag{C3}$$

Ciało możemy zapisać jako piątkę uporządkowaną $(X, +, \cdot, 0, 1)$.

Uwaga 1.1. W niektórych podręcznikach w warunku (C2) żąda się jedynie, aby $(X, \cdot, 1)$ było grupą. Wówczas można mówić o *ciałach nieprzemiennych* lub *pierścieniach z dzieleniem*. Przykładem takiej struktury algebraicznej są kwaterniony, które mają zastosowanie w grafice trójwymiarowej.

Przykład 1.4 (Ciała Galois)**.** Załóżmy, że p jest liczbą pierwszą. Jak pamiętamy z przykładów 1.1 i 1.2, na zbiorze $\mathbb{Z}_p$ możemy określić działania „+” i „·” modulo p, przy czym tak określone mnożenie jest rozdzielne względem dodawania. Oznacza to, że

$$(\mathbb{Z}_p, +, \cdot, 0, 1)$$

jest ciałem mającym p elementów.

Uwaga 1.2. Ciałem Galois nazywamy każde ciało skończone. Wspomniane powyżej ciała $(\mathbb{Z}_p, +, \cdot, 0, 1)$ nie wyczerpują wszystkich przykładów takich ciał. Ciała skończone mają zastosowania m. in. w kryptografii – zob. protokół Diffiego - Hellmana.

Zadanie 1.6. Niech $(X, +, \cdot, 0, 1)$ będzie ciałem, $a, b \in X$ i $a \cdot b = 0$. Pokaż, że wówczas $a = 0$ lub $b = 0$.

Zadanie 1.7. Niech $(X, +, \cdot, 0, 1)$ będzie ciałem. Pokaż, że dla każdego $a \in X$, $a \cdot 0 = 0$.

Zadanie 1.8. Skonstruuj ciało mające dokładnie 4 elementy.

Kolejny przykład ciała jest na tyle istotny, że poświęcimy mu osobny rozdział.

Rozdział 2

Liczby zespolone

2.1 Ciało liczb zespolonych

Przez wieki jedną z większych bolączek ludzkości, a przynajmniej tej części ludzkości, która zajmowała się matematyką, był fakt, że liczba -1 nie jest kwadratem żadnej innej *liczby rzeczywistej*. Inaczej mówiąc, nie istnieje liczba rzeczywista x taka, że

$$x^2 = -1.$$

Nic jednak nie stoi na przeszkodzie temu, abyśmy zadeklarowali, że rozwiązanie powyższego równania istnieje – nie będzie to jednak liczba rzeczywista. Umawiamy się po prostu, że pewien obiekt, który będziemy nazywali *jednostką urojoną* i oznaczali literką i, ma taką własność, że

$$i \cdot i = -1.$$

Od tej chwili możemy też pisać $i = \sqrt{-1}$. Jesteśmy teraz w stanie rozwiązać dowolne równanie postaci

$$x^2 = a, \qquad \text{gdzie a } \in \mathbb{R},$$

mianowicie, dla $a < 0$ jego rozwiązania to

$$x_1 = i\sqrt{|a|}, \qquad \text{oraz} \qquad x_2 = -i\sqrt{|a|}.$$

Przykład 2.1. Szukamy pierwiastków wielomianu $p(x) = x^2 + 4$. Ponieważ $i^2 = -1$, więc piszemy

$$\begin{aligned} p(x) &= x^2 + 4 \\ &= x^2 - (-1) \cdot 4 \\ &= x^2 - i^2 \cdot 4 \\ &= x^2 - (i \cdot 2)^2 \\ &= (x - i \cdot 2)(x + i \cdot 2). \end{aligned}$$

Jednostkę urojoną i traktujemy przy wykonywaniu działań tak samo jak każdą inną liczbę rzeczywistą, z tym, że zamiast i^2 zawsze możemy napisać -1 i na odwrót, liczbę -1 zawsze możemy przedstawić jako i^2. W powyższym rachunku mogliśmy dzięki temu wykorzystać dobrze znany wzór skróconego mnożenia $b^2 - c^2 = (b-c)(b+c)$.

Skoro ośmieliliśmy się traktować jednostkę urojoną tak samo, jak każdą liczbę rzeczywistą, to nieuchronnie będziemy mieli do czynienia z napisami postaci $a + bi$, $a - bi$, $(a + bi) \cdot (c + di)$, czy

$$\frac{a+bi}{i}, \qquad \text{a nawet} \qquad \frac{a+bi}{c+di},$$

gdzie $a, b, c, d \in \mathbb{R}$. Okazuje się jednak, że każdy z tych napisów można sprowadzić do pierwszego z nich:

$$\begin{aligned}
a - bi &= a + (-b)i, \\
(a+bi)(c+di) &= ac + adi + bic + bidi = ac + i^2bd + adi + bci \\
&= ac - bd + (ad + bc)i, \\
\frac{a+bi}{i} &= \frac{a+bi}{i^2} \cdot i = \frac{ai + bi^2}{-1} = \frac{-b + ai}{-1} \\
&= b - ai, \\
\frac{a+bi}{c+di} &= \frac{(a+bi)(c-di)}{(c+di)(c-di)} = \frac{ac + bd + (bc - ad)i}{c^2 + d^2} \\
&= \frac{ac+bd}{c^2+d^2} + \frac{bc-ad}{c^2+d^2}i.
\end{aligned}$$

Czym jednak jest ten napis $a + bi$?

Definicja 2.1. *Liczba zespolona* jest to para uporządkowana liczb rzeczywistych (a, b), którą tradycyjnie zapisujemy jako $a + bi$. Zbiór wszystkich liczb zespolonych oznaczamy symbolem $\mathbb{C}$.

Jeżeli $(a, b) = a + bi = z$ jest liczbą zespoloną, to

- liczbę rzeczywistą a nazywamy *częścią rzeczywistą* liczby zespolonej z i oznaczamy
$$\mathrm{Re}z = a,$$
- liczbę rzeczywistą b nazywamy *częścią urojoną* liczby zespolonej z i oznaczamy
$$\mathrm{Im}z = b,$$
- *moduł* liczby zespolonej z definiujemy jako
$$|z| = \sqrt{a^2 + b^2},$$

- *sprzężeniem* liczby zespolonej z nazywamy liczbę zespoloną

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

W szczególności, każda liczba rzeczywista a jest także liczbą zespoloną: $a = a + 0i$.

Na zbiorze $\mathbb{C}$ mamy określone działania

- *dodawania:*

$$(a, b) + (c, d) = (a + bi) + (c + di) = (a + c) + (b + d)i = (a + c, b + d)$$

z elementem neutralnym $0 = (0, 0) = 0 + 0i$, oraz

- *mnożenia:*

$$(a, b) \cdot (c, d) = (a + bi)(c + di) = ac - bd + (ad + bc)i$$

z elementem neutralnym $1 = (1, 0) = 1 + 0i$.

Ponadto, jeżeli $z = a + bi$, to możemy określić liczbę *przeciwną*

$$-z = -a - bi,$$

tak, że $z + (-z) = 0$, oraz, gdy $z \neq 0$, liczbę *odwrotną*

$$z^{-1} = \frac{1}{z} = \frac{1}{a + bi} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2} i,$$

tak, że $z \cdot z^{-1} = 1$.

Zadanie 2.1. Pokaż, że działania dodawania i mnożenia liczb zespolonych są przemienne i łączne, oraz mnożenie jest rozdzielne względem dodawania.

Z powyższych rozważań i zadania 2.1 wynika:

Twierdzenie 2.1. *Zbiór liczb zespolonych* $\mathbb{C}$ *z działaniami dodawania i mnożenia określonymi powyżej oraz elementami wyróżnionymi* 0 *i* 1 *jest ciałem.*

2.2 Podstawowe własności

Poniżej podajemy kilka własności liczb zespolonych. Niektóre z nich są oczywiste, inne trochę mniej.

Stwierdzenie 2.2. *Niech $z \in \mathbb{C}$. Wówczas*

(i) $z + \overline{z} = 2\mathrm{Re}z$,

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

(iii) jeżeli $z \neq 0$, *to* $z^{-1} = \dfrac{\overline{z}}{|z|^2}$.

Dowód powyższego stwierdzenia zostawiamy jako bardzo proste ćwiczenie.

Stwierdzenie 2.3. *Niech $z, w \in \mathbb{C}$. Wówczas*

(i) $\overline{z + w} = \overline{z} + \overline{w}$,

(ii) $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$,

(iii) $\overline{z^{-1}} = (\overline{z})^{-1}$,

(iv) $|z + w| \leqslant |z| + |w|$,

(v) $\big||z| - |w|\big| \leqslant |z - w|$,

(vi) $|zw| = |z| \cdot |w|$,

(vii) gdy $w \neq 0$, *to* $\left|\dfrac{z}{w}\right| = \dfrac{|z|}{|w|}$.

Dowód (iv). Niech $z = a + bi$, $w = c + di$. Trzeba pokazać, że

$$(a + c)^2 + (b + d)^2 \leqslant \left(\sqrt{a^2 + b^2} + \sqrt{c^2 + d^2}\right)^2.$$

mamy

$$a^2 + c^2 + 2ac + b^2 + d^2 + 2bd \leqslant a^2 + b^2 + c^2 + d^2 + 2\sqrt{(a^2 + b^2)(c^2 + d^2)}$$

i po uproszczeniu

$$ac + bd \leqslant \sqrt{(a^2 + b^2)(c^2 + d^2)}.$$

Jeżeli lewa strona jest ujemna, to już koniec. W przeciwnym razie podnosimy obie strony do kwadratu:

$$(ac + bd)^2 \leqslant (a^2 + b^2)(c^2 + d^2),$$

skąd, po wymnożeniu i uproszczeniu

$$2abcd \leqslant a^2d^2 + b^2c^2$$

lub równoważnie

$$0 \leqslant (ad - bc)^2.$$

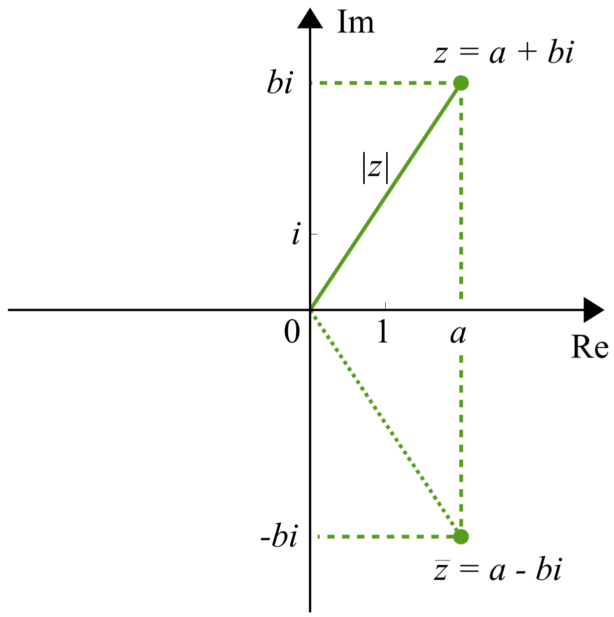
Ostatnia nierówność jest zawsze prawdziwa. □

Zadanie 2.2. Udowodnij pozostałe podpunkty ze stwierdzeń 2.2 i 2.3.

2.3 Interpretacja wektorowa

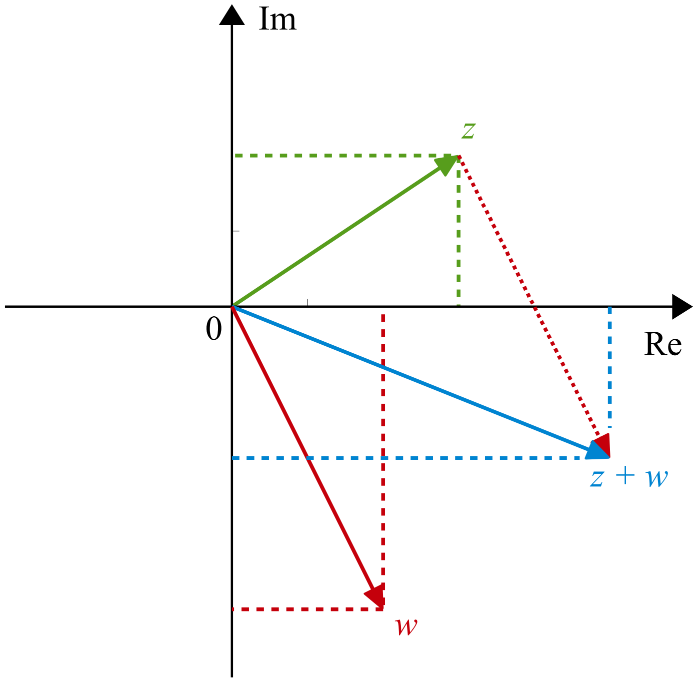
Liczba zespolona została zdefiniowana jako para uporządkowana liczb rzeczywistych. Wobec tego, naturalnym środowiskiem, w którym żyją liczby zespolone jest płaszczyzna (zob. Rysunek 2.1). Określamy ją mianem *płaszczyzny zespolonej* lub *płaszczyzny Gaussa*.

Rysunek 2.1: Interpretacja geometryczna liczby zespolonej.



Oś poziomą, oznaczoną na rysunku 2.1 symbolem Re nazywamy wówczas *osią rzeczywista*, natomiast oś pionowa, oznaczana symbolem Im, to *oś urojona*. Moduł liczby zespolonej z interpretujemy jako długość odcinka łączącego punkt na płaszczyźnie Gaussa odpowiadający z z liczbą 0. Operacja sprzężenia odpowiada odbiciu względem osi rzeczywistej.

Rysunek 2.2: Dodawanie liczb zespolonych.



W zależności od potrzeb, liczbę zespoloną $z = a + bi$ możemy interpretować jako punkt (a, b) na płaszczyźnie, lub jako wektor $\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right]$. Wówczas dodawanie liczb zespolonych to po prostu dodawanie wektorów, co zostało zilustrowane na Rysunku 2.2.

2.4 Postać trygonometryczna

Niech $z = x + iy \in \mathbb{C}$ i $z \neq 0$. Oznaczmy literą ϕ kąt pomiędzy półosią rzeczywistą dodatnią i wektorem odpowiadającym z i ustalmy, że jest to kąt zorientowany przeciwnie do ruchu wskazówek zegara. Wówczas

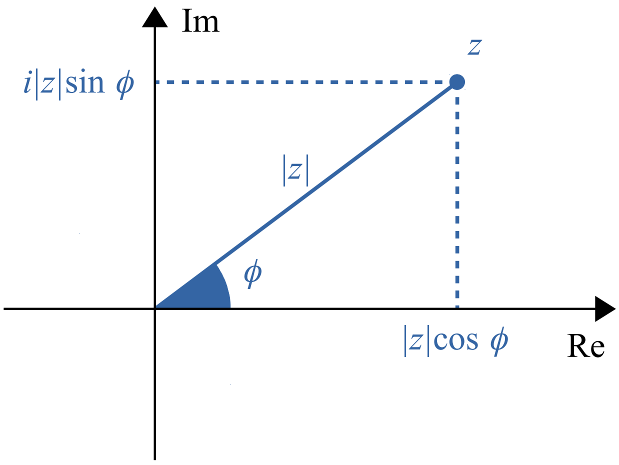
$$\frac{x}{|z|} = \cos\phi, \qquad \frac{y}{|z|} = \sin\phi$$

oraz

$$z = |z|\left(\cos\phi + i\sin\phi\right). \tag{2.1}$$

Zapisaliśmy w ten sposób liczbę zespoloną z w *postaci trygonometrycznej* (zob. Rys. 2.3). Kąt ϕ nazywamy *argumentem* liczby zespolonej z i oznaczamy $\phi = \arg z$. Jeżeli $\phi \in [0, 2\pi)$, to mówimy, że ϕ jest *argumentem głównym*, który oznaczamy $\mathrm{Arg} z$.

Rysunek 2.3: Postać trygonometryczna liczby zespolonej



Przyjrzymy się teraz ponownie iloczynowi dwóch liczb zespolonych. Niech $z, w \in \mathbb{C}$ i zapiszmy te liczby w postaci trygonometrycznej:

$$\begin{aligned} z &= |z|(\cos\phi + i\sin\phi) \\ w &= |w|(\cos\theta + i\sin\theta). \end{aligned}$$

Obliczamy iloczyn:

$$\begin{aligned} zw &= |z|\cdot|w|(\cos\phi\cos\theta - \sin\phi\sin\theta + i(\cos\phi\sin\theta + \sin\phi\cos\theta)) \\ &= |z|\cdot|w|(\cos(\phi+\theta) + i\sin(\phi+\theta)). \end{aligned}$$

Zatem moduł iloczynu zw to iloczyn modułów $|z|$ i $|w|$, natomiast argument iloczynu jest równy sumie argumentów liczb z i w. Geometryczna interpretacja mnożenia liczby zespolonej z przez liczbę w jest następująca: wektor odpowiadający z (zaczepiony w 0) obracamy wokół punktu 0 o kąt równy argumentowi liczby w, natomiast długość wektora z skalujemy o czynnik $|w|$. (zob. Rys. 2.4)

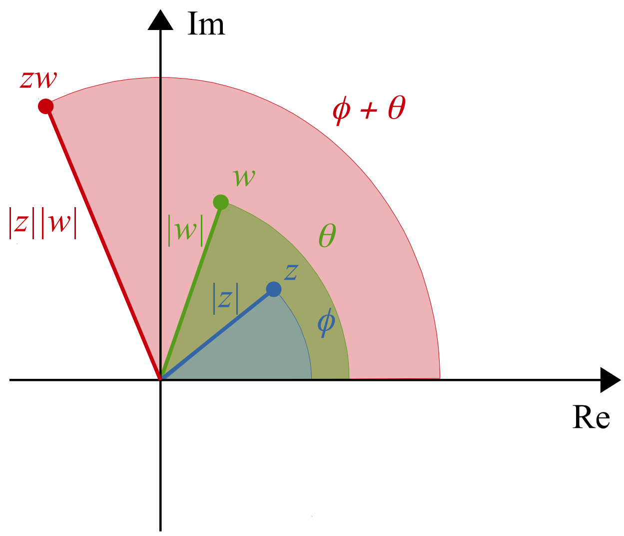
Zadanie 2.3. Pokaż, że dla $z = |z|(\cos\phi + i\sin\phi)$:

$$z^{-1} = |z|^{-1}(\cos\phi - i\sin\phi) = |z|^{-1}(\cos(-\phi) + i\sin(-\phi)).$$

Zapis wykładniczy. Umawiamy się, że będziemy pisać

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

Rysunek 2.4: Iloczyn liczb zespolonych



Taki zapis ma więcej sensu niż się z pozoru może wydawać. Jeżeli bowiem

$$z = |z| \cdot (e^{i\phi} = \cos\phi + i\sin\phi), \qquad w = |w| \cdot (e^{i\theta} = \cos\theta + i\sin\theta),$$

to

$$z \cdot w = |zw| \cdot \big(\cos(\phi+\theta) + i\sin(\phi+\theta)\big) = |zw| \cdot e^{i(\phi+\theta)} = |zw| \cdot e^{i\phi} \cdot e^{i\theta}.$$

Inaczej mówiąc, przy mnożeniu liczb zespolonych argumenty pomnożone przez i zachowują się jak wykładniki.

Przykład 2.2. Niech S^1 oznacza okrąg jednostkowy na płaszczyźnie zespolonej, czyli

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

Zauważmy, że $1 \in S^1$, gdy $z, w \in S^1$ to także $zw \in S^1$ oraz $z^{-1} \in S^1$. Inaczej mówiąc, S^1 jest grupą z działaniem mnożenia liczb zespolonych i 1 jako elementem neutralnym. Każdy element tej grupy możemy zapisać jako $e^{i\phi}$; wówczas działanie grupowe polega na dodawaniu wykładników, a odwrotność to wzięcie wykładnika z przeciwnym znakiem.

Uwaga 2.1. Ktoś może się zapytać, co zapis wykładniczy ma wspólnego z funkcją wykładniczą $x \mapsto e^x$, określoną dla $x \in \mathbb{R}$. Po pierwsze, możemy zdefiniować wartość funkcji wykładniczej dla dowolnej liczby zespolonej z:

$$e^z = e^{\mathrm{Re}z + i\mathrm{Im}z} = e^{\mathrm{Re}z} \cdot e^{i\mathrm{Im}z} = e^{\mathrm{Re}z}(\cos(\mathrm{Im}z) + i\sin(\mathrm{Im}z)).$$

Po drugie, definicja liczby e^x dla $x \in \mathbb{R}$:

$$e^x = \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n \qquad \text{lub} \qquad e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

może również zostać zastosowana dla $x = z$ zespolonego i okazuje się, że otrzymamy wówczas właśnie liczbę $e^{\mathrm{Re}z}(\cos(\mathrm{Im}z) + i\sin(\mathrm{Im}z))$. Zauważmy, że funkcja wykładnicza $z \mapsto e^z$ dla argumentów zespolonych nie jest różnowartościowa (a nawet jest okresowa, jej okres wynosi $2\pi i$), a jej obraz to zbiór $\mathbb{C}\setminus\{0\}$.

Stwierdzenie 2.4 (Wzór Eulera)**.**

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

Uwaga 2.2. Wzór Eulera bywa nazywany najpiękniejszym wzorem matematyki. Pomijając kwestie estetyki, zauważmy że pokazuje on związek pomiędzy pięcioma ważnymi stałymi matematycznymi: elementami wyróżnionymi ciała liczb rzeczywistych (i zespolonych) (0 i 1), jednostką urojoną i oraz liczbami π oraz e.

Stwierdzenie 2.5 (Wzór de Moivre'a)**.** *Jeżeli liczba n jest całkowita, natomiast $z = |z|e^{i\phi} = |z|(\cos\phi + i\sin\phi)$, to*

$$z^n = |z|^n e^{in\phi} = |z|^n(\cos n\phi + i\sin n\phi).$$

Dowód. Jeżeli n jest dodatnie to stosujemy $n-1$ razy wzór na postać trygonometryczną iloczynu liczb zespolonych. Dla n ujemnych najpierw dodatkowo zapisujemy w postaci trygonometrycznej liczbę z^{-1}. □

Przykład 2.3. Obliczymy $(1+i)^{21}$. Mamy $1+i = \sqrt{2}(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4})$. Zatem

$$\begin{aligned}(1+i)^{21} &= (\sqrt{2})^{21}\left(\cos\frac{21\pi}{4} + i\sin\frac{21\pi}{4}\right)\\ &= 2^{10}\sqrt{2}\left(\cos\frac{5\pi}{4} + i\sin\frac{5\pi}{4}\right)\\ &= 2^{10}(-1-i)\end{aligned}$$

2.5 Pierwiastki zespolone

Niech $a \in \mathbb{C}\setminus\{0\}$ będzie ustaloną liczbą zespoloną. Zajmiemy się teraz wyznaczeniem rozwiązań zespolonych równania

$$w^n = a.$$

Każde rozwiązanie tego równania będziemy nazywać *pierwiastkiem zespolonym stopnia n z liczby zespolonej a.*

Rozpatrzmy najpierw przypadek, gdy $|a| = 1$, czyli $a = e^{i\phi}$. Wówczas, jeżeli $w^n = a$, to także $|w| = 1$, czyli w jest postaci $e^{i\theta}$. Skoro $w^n = a$, czyli $e^{in\theta} = e^{i\phi}$, to

$$n\theta = 2k\pi + \phi$$

dla pewnej liczby całkowitej k. Zatem

$$\theta = \frac{2k\pi}{n} + \frac{\phi}{n}$$

Dla liczb k różniących się o n otrzymujemy argumenty θ różniące się o 2π a więc wszystkie różne wartości $e^{i\theta}$ otrzymamy biorąc $k \in \{0, 1, \ldots, n-1\}$. Otrzymaliśmy n różnych pierwiastków zespolonych z liczby zespolonej a:

$$w_k = e^{i(2k\pi+\phi)/n} = \cos\left(\frac{2k\pi + \phi}{n}\right) + i\sin\left(\frac{2k\pi + \phi}{n}\right).$$

Uogólnienie na przypadek, gdy $|a| \neq 1$ jest proste. Wówczas

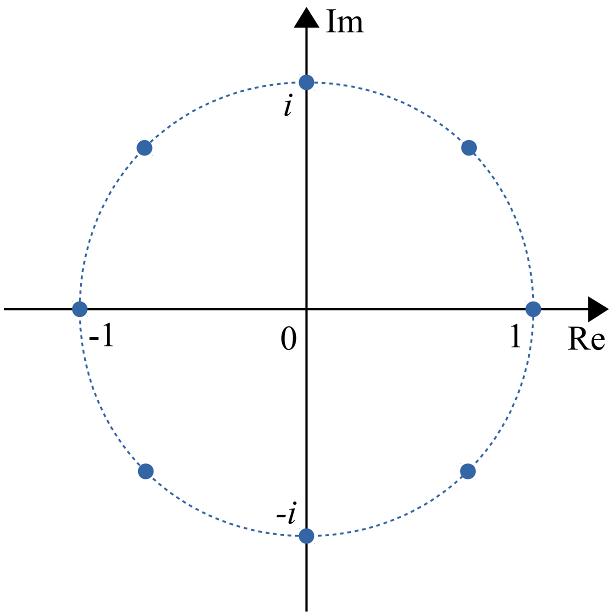
$$w_k = \sqrt[n]{|a|}e^{i(2k\pi+\phi)/n} = \sqrt[n]{|a|}\left(\cos\left(\frac{2k\pi + \phi}{n}\right) + i\sin\left(\frac{2k\pi + \phi}{n}\right)\right).$$

Pierwiastki z jedności stopnia n. Dla $a = 1$ otrzymujemy n różnych zespolonych pierwiastków z jedności stopnia n:

$$u_k = e^{i2k\pi/n} = \cos\left(\frac{2k\pi}{n}\right) + i\sin\left(\frac{2k\pi}{n}\right).$$

Jasne jest, że $u_k \in S^1$ i w szczególności $u_0 = 1$. Zbiór $\{1, u_1, \ldots, u_{n-1}\}$ z działaniem mnożenia liczb zespolonych i 1 jako elementem neutralnym jest grupą. Łatwo też sprawdzić, że $u_j \cdot u_k = u_l$, gdzie l to reszta z dzielenia $j + k$ przez n. Zatem grupa ta ma taką samą strukturę, co grupa addytywna $(\mathbb{Z}_n, +, 0)$.

Rysunek 2.5: Pierwiastki z jedności stopnia 8



Rozdział 3

Wielomiany

Niech $\mathbb{K}$ oznacza ciało $\mathbb{R}$ lub $\mathbb{C}$.

Definicja 3.1. *Wielomianem* zmiennej z *stopnia* n *o współczynnikach z ciała* $\mathbb{K}$ nazywamy funkcję $p : \mathbb{K} \to \mathbb{K}$ postaci

$$p(z) = a_0 + a_1 z + a_2 z^2 + \ldots + a_n z^n = \sum_{k=0}^{n} a_k z^k,$$

gdzie $a_0, a_1, \ldots, a_n \in \mathbb{K}$ oraz $a_n \neq 0$. Stopień wielomianu zapisujemy jako $\deg p$. Jeżeli p jest wielomianem zerowym, czyli $p(z) = 0$ dla każdego z, to przyjmujemy, że $\deg p = -\infty$.

Zbiór wszystkich wielomianów o współczynnikach z ciała $\mathbb{K}$ zmiennej z będziemy oznaczać $\mathbb{K}[z]$. Zbiór wszystkich takich wielomianów stopnia nie większego od n będziemy oznaczać $\mathbb{K}[z]_n$

3.1 Pierwiastki wielomianów

Definicja 3.2. Liczba $z_0 \in \mathbb{K}$ jest *pierwiastkiem* wielomianu $p \in \mathbb{K}[z]$, jeżeli $p(z_0) = 0$.

W poprzednim rozdziale pokazaliśmy, że wielomian o współczynnikach zespolonych postaci $p(z) = z^n - a$, gdzie $a \in \mathbb{C} \setminus \{0\}$ ma n różnych pierwiastków będących liczbami zespolonymi.

Przykład 3.1. Wyznaczymy pierwiastki zespolonego trójmianu kwadratowego $w(z) = z^2 + pz + q$, dla $p, q \in \mathbb{C}$.

$$\begin{aligned} w(z) &= z^2 + pz + q = \\ &= \left(z^2 + pz + \frac{p^2}{4}\right) + q - \frac{p^2}{4} \\ &= \left(z + \frac{p}{2}\right)^2 - \frac{p^2 - 4q}{4}. \end{aligned}$$

Niech $\Delta = p^2 - 4q$. Gdy $\Delta = 0$, to

$$w(z) = \left(z + \frac{p}{2}\right)^2$$

i mówimy, że w ma jeden dwukrotny pierwiastek $z_1 = z_2 = -\frac{p}{2}$.

Jeżeli $\Delta \neq 0$, to niech δ oznacza jeden z pierwiastków zespolonych stopnia 2 z liczby Δ. Wówczas drugi taki pierwiastek to $-\delta$. Przykładowo, jeżeli $\Delta = 4$, to jako δ możemy wybrać 2 lub -2; jeżeli $\Delta = -3$, to $\delta = -i\sqrt{3}$ lub $\delta = i\sqrt{3}$, a jeżeli np. $\Delta = 2i$, to $\delta = 1 + i$ lub $-1 - i$ – możliwe wartości δ różnią się tylko znakiem.

$$\begin{aligned} w(z) &= \left(z + \frac{p}{2}\right)^2 - \frac{\Delta}{4} \\ &= \left(z + \frac{p}{2} - \frac{\delta}{2}\right)\left(z + \frac{p}{2} + \frac{\delta}{2}\right) \end{aligned}$$

Wielomian w ma 2 pierwiastki zespolone:

$$z_1 = -\frac{p}{2} - \frac{\delta}{2}, \qquad z_2 = -\frac{p}{2} + \frac{\delta}{2}.$$

Wielomian w zapisaliśmy jako iloczyn czynników $z - z_1$ i $z - z_2$:

$$w(z) = (z - z_1)(z - z_2).$$

Znaleźliśmy wzory, których możemy użyć do wyznaczenia pierwiastków dowolnego wielomianu zespolonego stopnia 2. Podobne wzory, jednak znacznie bardziej skomplikowane, można wyznaczyć jeszcze dla wielomianów stopnia 3 i 4, ale na tym zabawa się kończy. Twierdzenie Abela-Ruffiniego, udowodnione niezależenie przez E. Galois, mówi, że nie istnieją ogólne wzory na pierwiastki wielomianów stopnia 5 i wyższych.

Stwierdzenie 3.1. *Jeżeli $p \in \mathbb{K}[z]$ jest wielomianem stopnia $n \geqslant 1$ i $u \in \mathbb{K}$, to istnieje wielomian $q \in \mathbb{K}[z]$ stopnia $n - 1$ taki, że dla każdego $z \in \mathbb{C}$*

$$p(z) = q(z)(z - u) + p(u)$$

i, gdy u jest pierwiastkiem p, to $p(z) = q(z)(z - u)$.

Dowód. Niech $p(z) = \sum_{k=0}^{n} a_k z^k$. Wówczas

$$\begin{aligned} p(z) &= p(z) - p(u) + p(u) \\ &= \sum_{k=0}^{n} a_k(z^k - u^k) + p(u) \\ &= \sum_{k=0}^{n} (z-u)q_k(z) + p(u) \\ &= q(z)(z-u) + p(u) \end{aligned}$$

W powyższych przekształceniach q i q_k to pewne wielomiany stopnia odp. $n-1$ i $k-1$ o współczynnikach zależnych od u. □

Obliczając pierwiastki $w_0, \ldots, w_{n-1}$ stopnia n z liczby zespolonej a wyznaczyliśmy jednocześnie pierwiastki wielomianu $z^n - a$. Stosując $n-1$ razy stwierdzenie 3.1 możemy więc napisać:

$$z^n - w = (z - w_0)(z - w_1) \ldots (z - w_{n-1}).$$

Wróćmy teraz do sytuacji, gdy w przykładzie 3.1 mamy $\Delta = 0$. Wówczas $w(z) = (z + \frac{p}{2})^2$ i po podzieleniu wielomianu w przez dwumian $z + \frac{p}{2}$ otrzymujemy wielomian, który też jest podzielny przez $z + \frac{p}{2}$. Mówimy, że $z_1 = -\frac{p}{2}$ jest pierwiastkiem *wielokrotnym*, a konkretnie w tym przypadku – dwukrotnym.

Definicja 3.3. Powiemy, że pierwiastek z_0 wielomianu $p(z)$ stopnia n, ma *krotność* k, jeżeli dla pewnego wielomianu q, $\deg q = n - k$ zachodzi

$$p(z) = q(z)(z - z_0)^k$$

oraz $q(z_0) \neq 0$.

3.2 Zasadnicze twierdzenie algebry

Zarówno zespolone trójmiany kwadratowe, jak i wielomiany $z^n - a$ mają, z uwzględnieniem krotności, tyle pierwiastków zespolonych, ile wynosi ich stopień. Uzasadnione jest więc pytanie, czy jest to prawda dla dowolnego wielomianu zespolonego. Okazuje się, że tak i wynika to z następującego twierdzenia:

Twierdzenie 3.2 (Zasadnicze twierdzenie algebry)**.** *Każdy wielomian zespolony $p(z)$ stopnia co najmniej 1 ma pierwiastek zespolony.*

Twierdzenie to zostało udowodnione na kilkanaście sposobów w 1799 roku przez Gaussa.

Z zasadniczego twierdzenia algebry i stwierdzenia 3.1 wynika natychmiast

Wniosek 3.3. *Każdy wielomian $p \in \mathbb{C}[z]$ stopnia $n > 0$ ma, z uwzględnieniem krotności, dokładnie n pierwiastków zespolonych $z_1, z_2, \ldots, z_n$ i*

$$p(z) = a(z - z_1)(z - z_2)\ldots(z - z_n),$$

gdzie a to współczynnik wielomianu p stojący przy z^n.

Na koniec zajmiemy się pierwiastkami zespolonymi wielomianów rzeczywistych.

Stwierdzenie 3.4. *Niech $p \in \mathbb{R}[z]$ będzie wielomianem zmiennej zespolonej z o współczynnikach rzeczywistych. Wówczas, jeżeli $p(z_0) = 0$ dla pewnej liczby zespolonej z_0, to także $p(\overline{z_0}) = 0$.*

Zadanie 3.1. Udowodnij to stwierdzenie.

Ze stwierdzenia 3.4 wynika, że istnieją liczby rzeczywiste $x_1, \ldots, x_k$ i zespolone o niezerowej części urojonej $z_1, \ldots, z_l$ takie, że $k + 2l = \deg p$ oraz

$$p(z) = (z - x_1)\ldots(z - x_k) \cdot (z - z_1)(z - \overline{z_1})\ldots(z - z_l)(z - \overline{z}_l).$$

Jeżeli $\mathrm{Im} z_j \neq 0$, to $(z - z_j)(z - \overline{z}_j) = (z^2 - 2z\mathrm{Re} z_j + |z_j|^2)$ jest wielomianem rzeczywistym, który nie ma pierwiastków rzeczywistych. Pokazaliśmy więc

Stwierdzenie 3.5. *Każdy wielomian o współczynnikach rzeczywistych stopnia większego lub równego 1 można zapisać jako iloczyn wielomianów stopnia 1 i wielomianów stopnia 2, przy czym żaden z tych wielomianów stopnia 2 nie ma już pierwiastków rzeczywistych.*

Przykład 3.2. Rozłożymy wielomian $p(z) = z^3 + z^2 + z + 1$ na czynniki rzeczywiste i zespolone minimalnego stopnia:

$$\begin{aligned} p(z) &= z^2(z+1) + (z+1) \\ &= (z+1)(z^2+1) \\ &= (z+1)(z+i)(z-i). \end{aligned}$$

3.3 Schemat Hornera

Schemat Hornera jest algorytmem stosowanym do sprawnego dzielenia danego wielomianu przez wielomian stopnia 1 oraz obliczania wartości wielomianu dla danego argumentu $x = x_0$.

Rozważamy wielomian stopnia n

$$p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_{n-2}x^{n-1} + a_{n-1}x^{n-1} + a_nx^n, \tag{3.1}$$

gdzie $a_j \in \mathbb{K}$, który przekształcamy do postaci

$$p(x) = a_0 + x(a_1 + x(a_2 + \ldots + x(a_{n-2} + x(a_{n-1} + a_nx))\ldots)). \tag{3.2}$$

Dla ustalonej wartości $x = x_0$ obliczamy

$$\begin{aligned} b_n &= a_n, \\ b_{n-1} &= a_{n-1} + b_nx_0, \\ b_{n-2} &= a_{n-2} + b_{n-1}x_0, \\ \ldots&, \\ b_0 &= a_0 + b_1x_0. \end{aligned}$$

Liczby b_k to wartości wyrażeń w kolejnych nawiasach we wzorze (3.2) i $b_0 = p(x_0)$. Ponadto

$$p(x) = (b_nx^{n-1} + b_{n-1}x^{n-2} + \ldots + b_2x + b_1)(x - x_0) + b_0.$$

Przykład 3.3. Obliczymy wartość wielomianu $p(x) = -2x^4 + x^2 - 5x + 3$ dla $x = -\frac{3}{2}$. Współczynniki a_n i b_n wygodnie jest zapisywać w tabelce:

a_n	b_n
-2	-2
0	$0 + (-2)(-\frac{3}{2}) = 3$
1	$1 + 3(-\frac{3}{2}) = -\frac{7}{2}$
-5	$-5 + (-\frac{7}{2})(-\frac{3}{2}) = \frac{1}{4}$
3	$3 + \frac{1}{4}(-\frac{3}{2}) = \frac{21}{8}$

Zatem $p(-\frac{3}{2}) = \frac{21}{8}$ oraz $p(x) = (-2x^3 + 3x^2 - \frac{7}{2}x + \frac{1}{4})(x + \frac{3}{2}) + \frac{21}{8}$.

Zadanie 3.2. Ile operacji dodawania i mnożenia należy wykonać, obliczając wartość wielomianu bezpośrednio ze wzoru (3.1) i z wykorzystaniem schematu Hornera?

Rozdział 4

Macierze

4.1 Definicja macierzy

Symbolem $\mathbb{K}$ oznaczamy ciało $\mathbb{R}$ lub $\mathbb{C}$.

Macierzą (o elementach z ciała $\mathbb{K}$) wymiaru $m \times n$ nazywamy tablicę prostokątną

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix}, \tag{4.1}$$

gdzie $a_{i,j} \in \mathbb{K}$ dla $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$ nazywamy elementami macierzy. Można również stosować bardziej zwarty zapis:

$$A = [a_{i,j}]_{\substack{i=1,\dots,m \\ j=1,\dots,n}}.$$

Jeżeli wymiar macierzy jest znany, to możemy również użyć zapisu $A = [a_{i,j}]_{i,j}$. Przyjmiemy, że pierwszy dolny indeks za nawiasem kwadratowym oznacza numer wiersza, a drugi – kolumny.

Zbiór wszystkich macierzy wymiaru $m \times n$ o elementach z ciała $\mathbb{K}$ będziemy oznaczać symbolem $\mathbb{K}^{m,n}$

Macierz wymiaru $n \times 1$ nazywamy *wektorem (n wymiarowym / długości n)*. Zbiór wszystkich takich macierzy oznaczamy po prostu $\mathbb{K}^n$. Przyjmujemy również, że $\mathbb{K}^{1,1} = \mathbb{K}^1 = \mathbb{K}$. Wektory będziemy zapisywać w następujący sposób:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_i]_{i=1}^n = [x_i]_i.$$

Dla $a_{i,j}$ z (4.1) o wektorze $\vec{a}_j = [a_{i,j}]_i$, $j = 1, \ldots, n$, mówimy jako o j-tej kolumnie macierzy A. Możemy również napisać

$$A = [\vec{a}_1, \ldots, \vec{a}_n]. \tag{4.2}$$

Podobnie, jeżeli $W_i = [a_{i,1}, \ldots a_{i,n}]$, $i = 1, \ldots, m$, to możemy napisać

$$A = \begin{bmatrix} W_1 \\ \vdots \\ W_m \end{bmatrix}. \tag{4.3}$$

Macierz W_i określamy wówczas jako i-ty wiersz macierzy A. Równości (4.3) i (4.2) to szczególne przypadki tzw. blokowego zapisu macierzy:

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & \ldots & A_{1,k} \\ A_{2,1} & A_{2,2} & \ldots & A_{2,k} \\ \multicolumn{4}{c}{\dotfill} \\ A_{l,1} & A_{l,2} & \ldots & A_{l,k} \end{bmatrix}. \tag{4.4}$$

Powyżej $A_{i,j}$ oznaczają już podmacierze macierzy A, czyli *bloki*. Bloki z tego samego wiersza powinny wszystkie mieć tyle samo wierszy, a bloki stojące w tej samej kolumnie powinny mieć wszystkie taką samą liczbę kolumn.

Macierz zerowa (wymiaru $m \times n$) jest to macierz, której wszystkie elementy są zerami: $0 = [0]_{i,j} \in \mathbb{K}^{m,n}$.

Macierz *diagonalna* jest to macierz $A = [a_{i,j}]_{i,j} \in \mathbb{K}^{m,n}$, w której $a_{i,j} = 0$ gdy $i \neq j$, czyli, gdy $m \leqslant n$

$$A = \begin{bmatrix} a_{1,1} & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & a_{2,2} & \ldots & 0 & 0 & \ldots & 0 \\ \multicolumn{7}{c}{\dotfill} \\ 0 & 0 & \ldots & a_{m,m} & 0 & \ldots & 0 \end{bmatrix}$$

lub, gdy $m \geqslant n$

$$A = \begin{bmatrix} a_{1,1} & 0 & \ldots & 0 \\ 0 & a_{2,2} & \ldots & 0 \\ \multicolumn{4}{c}{\dotfill} \\ 0 & 0 & \ldots & a_{n,n} \\ 0 & 0 & \ldots & 0 \\ \multicolumn{4}{c}{\dotfill} \\ 0 & 0 & \ldots & 0 \end{bmatrix}.$$

Ciąg pozycji o indeksach (i, i), $i = 1, \ldots, \min(m, n)$ w macierzy $A \in \mathbb{K}^{m,n}$ nazywamy *(główną) przekątną* macierzy A

Jeżeli $A = \mathbb{K}^{n,n}$, to mówimy, że macierz A jest kwadratowa. Macierz kwadratową i diagonalną $n \times n$ o wyrazach $d_1, \ldots, d_n$ możemy zapisać jako:

$$D = \operatorname{diag}(d_1, \ldots, d_n).$$

Macierz

$$I_n = \operatorname{diag}(1, \ldots, 1) \in \mathbb{K}^{n,n}$$

nazywamy *macierzą jednostkową* (rozmiaru $n \times n$).

Jeżeli w macierzy A wszystkie elementy pod przekątną są równe zero, to A nazywamy macierzą *trójkątną górną*. Analogicznie, jeżeli wszystkie elementy macierzy A nad przekątną są równe zero, to mówimy, że A jest *trójkątna dolna*. Zbiór macierzy trójkątnych górnych wymiaru $m \times n$ o elementach z ciała $\mathbb{K}$ oznaczamy $\mathrm{TRIU}^{m,n}(\mathbb{K})$. Zbiór macierzy trójkątnych dolnych wymiaru $m \times n$ oznaczamy $\mathrm{TRIL}^{m,n}(\mathbb{K})$.

Przykład 4.1. Macierze

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

są trójkątne: $A \in \mathrm{TRIL}^{3,3}(\mathbb{R})$, $B \in \mathrm{TRIU}^{2,3}(\mathbb{C})$, $C \in \mathrm{TRIU}^{4,3}(\mathbb{R})$.

4.2 Operacje na macierzach

Niech

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m,1} & a_{m,2} & \ldots & a_{m,n} \end{bmatrix} \in \mathbb{K}^{m,n}.$$

Transpozycją macierzy $A \in \mathbb{K}^{m,n}$ nazywamy macierz

$$A^T = \begin{bmatrix} a_{1,1} & a_{2,1} & \ldots & a_{m,1} \\ a_{1,2} & a_{2,2} & \ldots & a_{m,2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1,n} & a_{2,n} & \ldots & a_{m,n} \end{bmatrix} \in \mathbb{K}^{n,m}.$$

Przykład 4.2.

$$\begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & -2 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 \\ -1 & 3 \\ 2 & -2 \end{bmatrix}.$$

Hermitowskim sprzężeniem macierzy $A \in \mathbb{K}^{m,n}$ nazywamy macierz

$$A^H = \begin{bmatrix} \overline{a}_{1,1} & \overline{a}_{2,1} & \dots & \overline{a}_{m,1} \\ \overline{a}_{1,2} & \overline{a}_{2,2} & \dots & \overline{a}_{m,2} \\ \multicolumn{4}{c}{\dotfill} \\ \overline{a}_{1,n} & \overline{a}_{2,n} & \dots & \overline{a}_{m,n} \end{bmatrix} \in \mathbb{K}^{n,m}.$$

Jasne jest, że w przypadku, gdy $\mathbb{K} = \mathbb{R}$, to $A^T = A^H$. Stosuje się również oznaczenia

$$\overline{A} = [\overline{a}_{i,j}]_{i,j}, \qquad |A| = [|a_{i,j}|]_{i,j}.$$

Jak łatwo zauważyć $A^H = \overline{A}^T = \overline{A^T}$.

Przykład 4.3.

$$\begin{bmatrix} 1+i & 2 & 2-i \\ 2+3i & i & 0 \\ 1 & -3i & 3-2i \end{bmatrix}^H = \begin{bmatrix} 1-i & 2-3i & 1 \\ 2 & -i & 3i \\ 2+i & 0 & 3+2i \end{bmatrix}$$

Definicja 4.1. Macierz $A \in \mathbb{K}^{n,n}$ taką, że $A = A^T$ nazywamy macierzą *symetryczną*, natomiast jeżeli $A = -A^T$, to mówimy, że A jest macierzą *antysymetryczną*. Jeżeli $A = A^H$, to mówimy, że A jest macierzą *hermitowską*.

Zadanie 4.1. Macierz $A = [a_{i,j}]_{i,j}$ jest hermitowska. Pokaż, że $a_{i,i} \in \mathbb{R}$, natomiast dla $i \neq j$ $a_{i,j} = \overline{a}_{j,i}$.

Jeżeli $b \in \mathbb{K}$ jest elementem z ciała, czyli skalarem, oraz $A = [a_{i,j}]_{i,j} \in \mathbb{K}^{m,n}$, to możemy określić iloczyn macierzy A przez skalar b:

$$A \cdot b = b \cdot A = [ba_{i,j}]_{i,j} \in \mathbb{K}^{m,n}.$$

Zauważmy przy okazji, że

$$(bA)^T = bA^T \qquad \text{oraz} \qquad (bA)^H = \overline{b}A^H$$

Mając dwie macierze takiego samego rozmiaru $A = [a_{i,j}]_{i,j} \in \mathbb{K}^{m,n}$, $B = [b_{i,j}]_{i,j} \in \mathbb{K}^{m,n}$, możemy określić ich *sumę*:

$$A + B = \begin{bmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} & \dots & a_{1,n} + b_{1,n} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} & \dots & a_{2,n} + b_{2,n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{m,1} + b_{m,1} & a_{m,2} + b_{m,2} & \dots & a_{m,n} + b_{m,n} \end{bmatrix} \in \mathbb{K}^{m,n}.$$

Tak określone dodawanie macierzy jest oczywiście przemienne. Ponadto, mnożenie macierzy przez skalar jest rozdzielne względem dodawania.

Przy okazji warto zauważyć, że zbiór $\mathbb{K}^{m,n}$ z działaniem dodawania macierzy i macierzą zerową jako elementem neutralnym jest grupą. Macierz przeciwna do macierzy $A = [A_{i,j}]_{i,j}$ to macierz $-A = [-a_{i,j}]_{i,j}$.

Bez większych trudności możemy sprawdzić, że

$$(A+B)^T = A^T + B^T \qquad \text{oraz} \qquad (A+B)^H = A^H + B^H.$$

Dodawanie macierzy w postaci blokowej. Jeżeli dane są dwie macierze w postaci blokowej:

$$A = \begin{bmatrix} A_{1,1} & \dots & A_{1,k} \\ \dots & \dots & \dots \\ A_{l,1} & \dots & A_{l,k} \end{bmatrix}, \qquad B = \begin{bmatrix} B_{1,1} & \dots & B_{1,k} \\ \dots & \dots & \dots \\ B_{l,1} & \dots & B_{l,k} \end{bmatrix}$$

i dla dowolnych i, j bloki $A_{i,j}$ oraz $B_{i,j}$ maja takie same rozmiary, to

$$A + B = \begin{bmatrix} A_{1,1} + B_{1,1} & \dots & A_{1,k} + B_{1,k} \\ \dots & \dots & \dots \\ A_{l,1} + B_{l,1} & \dots & A_{l,k} + B_{l,k} \end{bmatrix}.$$

Znacznie ciekawszym działaniem jest *iloczyn macierzy*. Jeżeli $A = [a_{i,j}]_{ij} \in \mathbb{K}^{m,r}$, $B = [b_{i,j}]_{i,j} \in \mathbb{K}^{r,n}$, to iloczyn tych macierzy jest macierzą

$$A \cdot B = C = [c_{i,j}]_{i,j} \in \mathbb{K}^{m,n},$$

gdzie

$$c_{i,j} = a_{i,1}b_{1,j} + a_{i,2}b_{2,j} + \ldots + a_{i,r}b_{r,j}, \qquad \text{dla } 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n.$$

Przykład 4.4.

$$\begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & -2 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 & 1 & 2 \\ 2 & 0 & 1 & -3 \\ 3 & 1 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 1 & 2 & 13 \\ 0 & -2 & 1 & -17 \end{bmatrix}$$

Stwierdzenie 4.1 (Własności iloczynu macierzy)**.**

(i) Jeżeli $A, B \in \mathbb{K}^{m,r}$ i $C \in \mathbb{K}^{r,n}$, to

$$(A + B) \cdot C = A \cdot C + B \cdot C \in \mathbb{K}^{m,n}.$$

(ii) Jeżeli $A \in \mathbb{K}^{m,r}$ i $B, C \in \mathbb{K}^{r,n}$, to

$$A \cdot (B + C) = A \cdot B + A \cdot C \in \mathbb{K}^{m,n}.$$

(iii) *Jeżeli* $c \in \mathbb{K}$, $A \in \mathbb{K}^{m,r}$, $B \in \mathbb{K}^{r,n}$, *to*

$$c(A \cdot B) = (cA) \cdot B = A \cdot (cB) \in \mathbb{K}^{m,n}.$$

(iv) *Jeżeli* $A \in \mathbb{K}^{m,r}$, $B \in \mathbb{K}^{r,s}$, $C \in \mathbb{K}^{s,nm}$ *to*

$$(A \cdot B) \cdot C = A \cdot (B \cdot C) \in \mathbb{K}^{m,n}.$$

(v) *Jeżeli* $A \in \mathbb{K}^{m,r}$, $B \in \mathbb{K}^{r,n}$, *to*

$$(A \cdot B)^T = B^T \cdot A^T \in \mathbb{K}^{n,m} \quad \textit{oraz} \quad (A \cdot B)^H = B^H \cdot A^H \in \mathbb{K}^{n,m}.$$

(vi) *Jeżeli* $A \in \mathbb{K}^{m,r}$, 0 *to macierz zerowa z* $\mathbb{K}^{r,n}$, *to* $A \cdot 0 = 0 \in \mathbb{K}^{m,n}$ *i, gdy* 0 *to macierz zerowa z* $\mathbb{K}^{m,r}$, $A \in \mathbb{K}_{r,n}$, *to* $0 \cdot A = 0 \in \mathbb{K}^{m,n}$.

(vii) *Jeżeli* $A \in \mathbb{K}^{m,n}$, *to*

$$I_m \cdot A = A = A \cdot I_n.$$

Nie wszystkie własności iloczynu liczb przenoszą się na iloczyn macierzy

Przykład 4.5. Jeżeli $AB = 0$, to wcale nie musi zachodzić $A = 0$ lub $B = 0$:

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

Jeżeli $A \in \mathbb{K}^{m,r}$ i $B \in \mathbb{K}^{r,n}$, to, dla $n \neq m$ iloczyn BA nie jest określony. Oczywiście nawet, gdy $m = n$, to nie musi być prawdą, że $AB = BA$, na przykład, gdy $r \neq m$. Jednak nawet w sytuacji, gdy $r = m = n$ nie możemy powiedzieć, że zawsze $AB = BA$.

Przykład 4.6.

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

ale

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

Zadanie 4.2. Pokaż, że iloczyn macierzy kwadratowych trójkątnych górnych (dolnych) jest macierzą trójkątną górną (dolną).

Iloczyn macierzy w postaci blokowej Jeżeli dane są dwie macierze w postaci blokowej

$$A = \begin{bmatrix} A_{1,1} & \dots & A_{1,l} \\ \dots & \dots & \dots \\ A_{k,1} & \dots & A_{k,l} \end{bmatrix}, \qquad B = \begin{bmatrix} B_{1,1} & \dots & B_{1,r} \\ \dots & \dots & \dots \\ B_{l,1} & \dots & B_{l,r} \end{bmatrix}$$

oraz $A_{i,j} \in \mathbb{K}^{u_i,v_j}$, $B_{i,j} \in \mathbb{K}^{v_i,w_j}$, to dla

$$C_{s,t} = A_{s,1}B_{1,t} + A_{s,2}B_{2,t} + \ldots + A_{s,l}B_{l,t}$$

mamy

$$AB = \begin{bmatrix} C_{1,1} & \dots & C_{1,r} \\ \dots & \dots & \dots \\ C_{k,1} & \dots & C_{k,r} \end{bmatrix}.$$

4.3 Macierz jako odwzorowanie

Dana macierz $A \in \mathbb{K}^{m,n}$ zadaje odwzorowanie ze zbioru $\mathbb{K}^n$ w zbiór $\mathbb{K}^m$:

$$F(\vec{x}) = A\vec{x}, \qquad \text{dla } \vec{x} \in \mathbb{K}^n.$$

Przykład 4.7 (Macierze obrotów płaszczyzny)**.** Jak pamiętamy, mnożenie przez liczbę zespoloną $e^{i\alpha}$ odpowiada obrotowi płaszczyzny Gaussa o kąt α wokół zera. Rozważmy więc macierz

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

i rozważmy przekształcenie $F_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ zadane wzorem

$$F_\alpha\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\alpha - y\sin\alpha \\ x\sin\alpha + y\cos\alpha \end{bmatrix}.$$

Wektor wynikowy to obrót wektora $[x, y]^T$ o kąt α wokół początku układu współrzędnych.

4.4 Macierze nieosobliwe

W zbiorze macierzy kwadratowych $\mathbb{K}^{n,n}$ mnożenie macierzy jest działaniem wewnętrznym i jest to działanie łączne.

Definicja 4.2. Powiemy, że macierz $A \in \mathbb{K}^{n,n}$ jest *nieosobliwa (odwracalna)*, jeżeli istnieje macierz B taka, że $AB = I_n$. Macierz B nazywamy macierzą *odwrotną* do A lub po prostu *odwrotnością* macierzy A i zapisujemy jako A^{-1}.

Macierz, która nie jest odwracalna jest nazywana macierzą *nieodwracalną* lub *osobliwą*.

Przykład 4.8.

- Macierz I_2 jest odwracalna i $I_2^{-1} = I_2$.
- Macierz zerowa $0 \in \mathbb{R}^{2,2}$ jest osobliwa, bowiem $0 \cdot A = 0$ dla każdej macierzy A.
- Macierz $A = \begin{bmatrix} 1 & 1 \\ 3 & 2 \end{bmatrix}$ jest nieosobliwa i $A^{-1} = \begin{bmatrix} -2 & 1 \\ 3 & -1 \end{bmatrix}$.
- Macierz $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ jest osobliwa, bowiem dla dowolnych $a, b, c, d \in \mathbb{R}$
$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a+c & b+d \\ a+c & b+d \end{bmatrix}$$
i nie jest możliwe, żeby macierz po prawej stronie była macierzą jednostkową.
- Jeżeli $d_1, \ldots d_n \neq 0$, to macierz diagonalna $D = \operatorname{diag}(d_1, \ldots, d_n)$ jest nieosobliwa i $D^{-1} = \operatorname{diag}(d_1^{-1}, \ldots, d_n^{-1})$.

Stwierdzenie 4.2. *Jeżeli macierze $A, B \in \mathbb{K}^{n,n}$ są nieosobliwe, to macierz $C = AB$ też jest nieosobliwa i $C^{-1} = B^{-1}A^{-1}$.*

Dowód. Niech $D = B^{-1}A^{-1}$. Wówczas
$$CD = (AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n.$$
□

Przykład 4.9. Macierze skalarne. Rozważmy następujący podzbiór zbioru $\mathbb{R}^{2,2}$:
$$X = \left\{ \begin{bmatrix} x & -y \\ y & x \end{bmatrix} \in \mathbb{R}^{2,2} : x, y \in \mathbb{R} \right\}.$$

Zbiór X jest zamknięty ze względu na dodawanie. Ponadto

$$\begin{bmatrix} x & -y \\ y & x \end{bmatrix} \begin{bmatrix} u & -v \\ v & u \end{bmatrix} = \begin{bmatrix} xu - yv & -xv - yu \\ xv + yu & xu - yv \end{bmatrix},$$

czyli X także jest zamknięty ze względu na iloczyn macierzy, oraz, gdy $x \neq 0$ lub $y \neq 0$

$$\begin{bmatrix} x & -y \\ y & x \end{bmatrix} \cdot \frac{1}{x^2 + y^2} \begin{bmatrix} x & y \\ -y & x \end{bmatrix} = I_2.$$

Rozważmy teraz bijekcję $\Phi : \mathbb{C} \to X$:

$$\Phi(x + iy) = \begin{bmatrix} x & -y \\ y & x \end{bmatrix}.$$

Widzimy, że $\Phi(0) = 0$, $\Phi(1) = I_2$ oraz dla $z, w \in \mathbb{C}$

- $\Phi(zw) = \Phi(z) \cdot \Phi(w)$,
- $\Phi(z + w) = \Phi(z) + \Phi(w)$,
- $\Phi(z^{-1}) = \Phi(z)^{-1}$ dla $z \neq 0$.

Okazuje się, że $(X, +, \cdot, 0, I_2)$ jest ciałem. Na dodatek ciało to jest tak naprawdę ciałem liczb zespolonych $\mathbb{C}$, tyle, że jego elementy zapisujemy w nieco inny sposób. Dodatkowo, określona przez nas bijekcja Φ jest zgodna z działaniami dodawania, mnożenia i brania odwrotności w $\mathbb{C}$ i X.

Ze względu na to, że powyższe macierze mają własności skalarów z ciała liczb zespolonych, są on nazywane *macierzami skalarnymi*.

W trakcie wykładu poznamy wiele warunków równoważnych nieosobliwości macierzy. Pierwszym z nich jest stw. 4.4 w dalszej części tego rozdziału.

Zadanie 4.3. Pokaż, że macierz $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{K}^{2,2}$ jest nieosobliwa wtedy i tylko wtedy, gdy $ad - bc \neq 0$ i wówczas $A^{-1} = \dfrac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

Zadanie 4.4. Macierz $A \in \mathbb{K}^{n,n}$ jest nieosobliwa. Pokaż, że macierze A^T i A^H też są nieosobliwe.

Zadanie 4.5. Pokaż, że macierz trójkątna $A \in \mathbb{K}^{n,n}$ jest nieosobliwa wtedy i tylko wtedy, gdy ma wszystkie wyrazy na przekątnej różne od zera.

4.5 Operacje elementarne

Operacje elementarne na wierszach macierzy. Dla macierzy

$$A = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix} \in \mathbb{K}^{m,n}, \qquad \text{gdzie } w_i \in \mathbb{K}^{1,n}$$

rozważamy następujące przekształcenia:

(I) zamiana miejscami dwóch wierszy: $w_j \leftrightarrow w_k$,

(II) pomnożenie wiersza przez niezerowy skalar: $w_j \to \alpha w_j$, $\alpha \in \mathbb{K}$,

(III) dodanie do pewnego wiersza innego pomnożonego przez skalar: $w_j \to w_j + \beta w_k$.

Przekształcenia te nazywamy operacjami elementarnymi na wierszach macierzy.

Zadanie 4.6. Pokaż, że każdą operację elementarną na wierszach można zapisać w postaci iloczynu MA, gdzie $M \in \mathbb{R}^{m,m}$ jest macierzą nieosobliwą. Wyznacz macierze M dla każdego rodzaju operacji.

Operacje elementarne na kolumnach macierzy. Dla macierzy

$$A = \left[\vec{k}_1 \;\; \vec{k}_2 \dots \vec{k}_n\right] \in \mathbb{K}^{m,n}, \qquad \text{gdzie } \vec{k}_j \in \mathbb{K}^m$$

rozważamy następujące przekształcenia:

(I) zamiana miejscami dwóch kolumn: $\vec{k}_j \leftrightarrow \vec{k}_l$,

(II) pomnożenie kolumny przez niezerowy skalar: $\vec{k}_j \to \alpha \vec{k}_j$, $\alpha \in \mathbb{K}$,

(III) dodanie do pewnej kolumny innej pomnożonej przez skalar: $\vec{k}_j \to \vec{k}_j + \beta \vec{k}_l$.

Przekształcenia te nazywamy operacjami elementarnymi na kolumnach macierzy.

Zadanie 4.7. Pokaż, że każdą operację elementarną na kolumnach można zapisać w postaci iloczynu AM, gdzie $M \in \mathbb{R}^{n,n}$ jest macierzą nieosobliwą. Wyznacz macierze M dla każdego rodzaju operacji.
Wskazówka: Jak wykorzystać rozwiązania zadania 4.6?

4.6 Układy równań liniowych

W tym podrozdziale wprowadzamy wstępne definicje dotyczące układów równań liniowych. Cały temat jest omówiony szczegółowo w rozdziale 7.

Układ m równań liniowych o n niewiadomych jest to układ równości

$$\begin{cases} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \dots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \dots & + & a_{2,n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \dots & + & a_{m,n}x_n & = & b_m \end{cases}$$

Liczby $a_{i,j}$ $i = 1, \dots, m$, $j = 1, \dots, n$ nazywamy *współczynnikami* tego układu, $b_1, \dots, b_m$ nazywamy *wyrazami wolnymi*, natomiast $x_1, \dots, x_n$ nazywamy *niewiadomymi*.

Jeżeli $A = [a_{i,j}] \in \mathbb{K}^{m,n}$, $\vec{b} = [b_i] \in \mathbb{K}^m$, $\vec{x} = [x_j] \in \mathbb{K}^n$, to powyższy układ równań możemy zapisać w postaci macierzowej

$$A\vec{x} = \vec{b}.$$

Każdy wektor $\vec{x} \in \mathbb{K}^n$ taki, że $A\vec{x} = \vec{b}$ nazywamy *rozwiązaniem* danego układu równań.

Układ równań można rozważać w jeszcze bardziej zwartej postaci, zapisując wyłącznie współczynniki $a_{i,j}$ stojące przy niewiadomych i wyrazy wolne b_i, jako macierz o m wierszach i $n + 1$ kolumnach:

$$[A|\vec{b}].$$

Stosując operacje elementarne na wierszach do macierzy $[A|\vec{b}]$ otrzymujemy układy równoważne (czyli mające te same zbiory rozwiązań), jest to więc wygodny sposób przekształcania układu równań w celu wyznaczenia rozwiązania:

Stwierdzenie 4.3. *Załóżmy, że macierz $[A'|\vec{b}']$ otrzymano z macierzy $[A|\vec{b}]$ za pomocą operacji elementarnych na wierszach. Wówczas wektor $\vec{x}$ jest rozwiązaniem układu $A\vec{x} = \vec{b}$ wtedy i tylko wtedy, gdy jest rozwiązaniem układu $A'\vec{x} = \vec{b}'$.*

Dowód. Wystarczy pokazać równoważność z tezy dla pojedynczej operacji elementarnej na wierszach. Dowód dla operacji typu (I) i (II) zostawiamy jako proste ćwiczenie.

Pokażemy, że zbiór rozwiązań układu $A\vec{x} = \vec{b}$ nie zmieni się, jeżeli do macierzy $[A|\vec{b}]$ zastosujemy operację typu (III), czyli do pewnego wiersza macierzy $[A|\vec{b}]$ dodamy inny pomnożony przez skalar z ciała $\mathbb{K}$. Oznaczmy kolejne wiersze macierzy A przez $w_1, \ldots, w_m$. Jeżeli A' otrzymujemy z A przez dodanie do wiersza i-tego wiersza j-tego pomnożone przez skalar λ, to macierz A' ma wiersze $w_1, \ldots, w_{i-1}, w_i + \lambda w_j, w_{i+1}, \ldots, w_m$, natomiast wektor $\vec{b}'$ ma współczynniki $b_1, \ldots, b_{i-1}, b_i + \lambda b_j, b_{i+1}, \ldots, b_m$. W układzie równań zmianie uległo tylko i-te równanie:

$$w_i \vec{x} = b_i \quad \mapsto \quad w_i \vec{x} + \lambda w_j \vec{x} \mapsto b_i + \lambda b_j.$$

Ponieważ $w_j \vec{x} = b_j$, więc każdy wektor $\vec{x}$ taki, że $A\vec{x} = \vec{b}$ spełnia także równanie $A'\vec{x} = \vec{b}'$. Aby udowodnić implikację w drugą stronę wystarczy zauważyć, że macierz $[A|\vec{b}]$ otrzymujemy z macierzy $[A'|\vec{b}']$ za pomocą operacji elementarnej, która do i-tego wiersza dodaje wiersz j-ty pomnożony przez $-\lambda$. $\square$

Istnienie i jednoznaczność rozwiązania układu równań ma związek z nieosobliwością macierzy:

Stwierdzenie 4.4. *Macierz $A \in \mathbb{K}^{n,n}$ jest nieosobliwa wtedy i tylko wtedy, gdy dla każdego wektora $\vec{b}$ układ równań $A\vec{x} = \vec{b}$ ma dokładnie jedno rozwiązanie $\vec{x}$.*

Dowód. Jeżeli macierz A jest nieosobliwa, to wektor $\vec{x} = A^{-1}\vec{b}$ jest rozwiązaniem. Ponadto, jeżeli $\vec{y}$ jest rozwiązaniem układu $A\vec{x} = \vec{b}$, to

$$\vec{y} = A^{-1}(A\vec{y}) = A^{-1}\vec{b} = \vec{x},$$

więc $\vec{x}$ jest jedynym rozwiązaniem.

Załóżmy, że dla każdego $\vec{b} \in \mathbb{K}^n$ układ $A\vec{x} = \vec{b}$ ma rozwiązanie. Niech $\vec{c}_1, \ldots, \vec{c}_n$ to rozwiązania dla $b = \vec{e}_1, \ldots, \vec{e}_n$ (czyli jako $\vec{b}$ bierzemy kolejne kolumny macierzy I_n). Niech $C = [\vec{c}_1, \ldots, \vec{c}_n]$. Wówczas

$$AC = [A\vec{c}_1, \ldots, A\vec{c}_n] = [\vec{e}_1, \ldots, \vec{e}_n] = I_n.$$

Zatem $A^{-1} = C$. $\square$

4.7 Odwracanie macierzy

Załóżmy, że macierz $A \in \mathbb{K}^{n,n}$ jest nieosobliwa. Poniżej pokazany jest sposób wyznaczenia macierzy A^{-1} z wykorzystaniem operacji elementarnych na kolumnach.

Niech $E_1, E_2, \dots, E_k$ to macierze kolejnych operacji elementarnych na kolumnach, dobranych tak, aby macierz A przekształcić w macierz I_n. Inaczej mówiąc

$$A \cdot E_1 \cdot E_2 \cdot \ldots \cdot E_k = I_n,$$

czyli, dla $B = E_1 \cdot E_2 \cdot \ldots \cdot E_k$ mamy $A \cdot B = I_n$. To oznacza, że B jest macierzą odwrotną do A, $B = A^{-1}$. Zauważmy także, że

$$A^{-1} = E_1 \cdot E_2 \cdot \ldots \cdot E_k = I_n \cdot E_1 \cdot E_2 \cdot \ldots \cdot E_k,$$

czyli, stosując operacje elementarne w tej samej kolejności do macierzy jednostkowej I_n otrzymujemy macierz odwrotną do A. Najwygodniej jest po prostu wykonywać te operacje jednocześnie na macierzy $\begin{bmatrix} A \\ I_n \end{bmatrix} \in \mathbb{K}^{2n,n}$:

$$\begin{bmatrix} A \\ I_n \end{bmatrix} \cdot E_1 \cdot E_2 \cdot \ldots \cdot E_k = \begin{bmatrix} I_n \\ A^{-1} \end{bmatrix}.$$

Systematyczny sposób doboru operacji elementarnych $E_1, \dots, E_k$, zwany *metodą Gaussa - Jordana* jest następujący:

- sprowadzamy macierz A do postaci trójkątnej dolnej bez zer na przekątnej za pomocą operacji typu (III) i (I), zerując elementy nad przekątną w kolejnych wierszach poczynając od pierwszego;
- za pomocą operacji typu (II) wszystkie elementy na przekątnej zamieniamy na 1;
- za pomocą operacji typu(III) zerujemy wszystkie elementy pod przekątną.

Rozdział 5

Przestrzenie liniowe

5.1 Definicja przestrzeni liniowej

Jak się dowiedzieliśmy w poprzednim rozdziale, w zbiorze wektorów $\mathbb{K}^n$ lub macierzy $\mathbb{K}^{m,n}$ określone są dwa działania:

- *dodawanie* elementów,
- *mnożenie* elementu *przez skalar z ciała* $\mathbb{K}$.

Zauważyliśmy również, że dodawanie jest przemienne, a mnożenie przez skalar jest rozdzielne względem dodawania elementów.

Podobna sytuacja ma miejsce w przypadku zbiorów wielomianów $\mathbb{K}[x]$ czy $\mathbb{K}[x]_n$: dodając do siebie dwa wielomiany (stopnia nie większego niż n) lub mnożąc taki wielomian przez liczbę, w wyniku otrzymujemy także wielomian (stopnia nie większego niż n). Mówimy, że zbiory te są *przestrzeniami liniowymi*:

Definicja 5.1. Niech $(X, +, 0)$ będzie grupą przemienną, w której element przeciwny do $x \in X$ będziemy zapisywać jako $-x$. Załóżmy dodatkowo, że w zbiorze X określone jest działanie „$\cdot$” mnożenia elementu $x \in X$ przez skalar $a \in \mathbb{K}$, przy czym działanie to ma następujące własności:

- dla dowolnego $x \in X$
$$1 \cdot x = x,$$
gdzie $1 \in \mathbb{K}$ to element neutralny mnożenia w ciele $\mathbb{K}$,
- dla dowolnych $x, y \in X$ i dowolnych $\alpha, \beta \in \mathbb{K}$
$$\begin{aligned}(\alpha + \beta) \cdot x &= \alpha \cdot x + \beta \cdot x,\\ \alpha \cdot (x + y) &= \alpha \cdot x + \alpha \cdot y,\\ (\alpha\beta) \cdot x &= \alpha \cdot (\beta \cdot x).\end{aligned}$$

Wówczas, zbiór X wraz z działaniami „+” i „$\cdot$” nazywamy *przestrzenią liniową nad ciałem* $\mathbb{K}$.

Elementy przestrzeni liniowej nazywamy *wektorami*.

Zapisując mnożenie elementu x przestrzeni liniowej przez skalar α możemy, zgodnie z wieloletnią tradycją, pominąć symbol mnożenia (czyli $\alpha \cdot x = \alpha x$).

Stwierdzenie 5.1 (Podstawowe własności przestrzeni liniowej)**.** *Jeżeli X jest przestrzenią liniową nad ciałem $\mathbb{K}$, to*

(i) dla każdego $x \in X$ zachodzi $0x = 0$ (gdzie 0 po lewej to zero w ciele $\mathbb{K}$, a zero po prawej, to element neutralny dodawania w X),

(ii) dla każdego $x \in X$ zachodzi $(-1) \cdot x = -x$,

(iii) dla każdych $\alpha \in \mathbb{K}$ i $x \in X$

$$\alpha x = 0 \quad \iff \quad \alpha = 0 \text{ lub } x = 0.$$

Dowód. Podpunkt (i):

$$0 + 0x = 0x = (0 + 0)x = 0x + 0x$$

i z prawa skracania wynika, że $0 = 0x$.

Podpunkt (ii): mamy

$$0 = 0x = (1 + (-1))x = x + (-1)x,$$

więc $-x = (-1)x$.

Podpunkt (iii): Załóżmy, że $\alpha = 0$ lub $x = 0$. Jeżeli $\alpha = 0$, to korzystamy z (i), dostając $\alpha \cdot x = 0$. Jeżeli $x = 0$, to

$$\alpha \cdot x = \alpha \cdot 0 = \alpha \cdot (0 + 0) = \alpha \cdot 0 + \alpha \cdot 0,$$

i dodając do obu stron element przeciwny do $\alpha \cdot 0$ (czyli odejmując $\alpha \cdot 0$) dostajemy $\alpha \cdot 0 = 0$

Załóżmy teraz, że $\alpha x = 0$. Gdy $\alpha \neq 0$, to

$$x = 1 \cdot x = (\alpha^{-1}\alpha)x = \alpha^{-1}(\alpha x) = \alpha^{-1} \cdot 0 = 0.$$

□

Znane już przykłady przestrzeni liniowych nad ciałem $\mathbb{K}$ (gdzie, zgodnie z umową, $\mathbb{K} = \mathbb{R}$ lub $\mathbb{K} = \mathbb{C}$), to $\mathbb{K}^n$, $\mathbb{K}^{m,n}$, $\mathbb{K}[x]$ i $\mathbb{K}[x]_n$. Omówimy teraz kilka nowych przykładów przestrzeni liniowych.

Przykład 5.1 (Przestrzenie funkcji)**.** Załóżmy, że Z jest dowolnym zbiorem (np. $Z = \mathbb{N}$, $Z = [0, 1]$, $Z = \mathbb{R}^2$) i rozpatrzmy zbiór $\mathbb{K}^Z$, czyli zbiór wszystkich funkcji $f : Z \to \mathbb{K}$. Funkcje takie możemy mnożyć przez elementy ciała $\mathbb{K}$:

$$(\alpha \cdot f)(x) = \alpha f(x), \qquad \alpha \in \mathbb{K}, \quad x \in Z,$$

gdzie po prawej stronie mamy mnożenie w ciele $\mathbb{K}$, oraz dodawać do siebie:

$$(f + g)(x) = f(x) + g(x), \qquad x \in Z,$$

gdzie po prawej mamy dodawanie w ciele $\mathbb{K}$. Mamy także element neutralny dodawania, a mianowicie funkcję zerową $x \mapsto 0 \in \mathbb{K}$ dla każdego $x \in Z$.

Można łatwo sprawdzić, że spełnione są wszystkie warunki z definicji przestrzeni liniowej, więc $\mathbb{K}^Z$ jest przestrzenią liniową nad ciałem $\mathbb{K}$. Co więcej, przestrzenie $\mathbb{K}^n$ i $\mathbb{K}^{m,n}$ są szczególnymi przypadkami takich przestrzeni. Dla $\mathbb{K}^n$ można wziąć $Z = \{1, 2, \ldots, n\}$, natomiast dla $\mathbb{K}^{m,n}$ za Z bierzemy zbiór wszystkich par (i, j), gdzie $i \in \{1, 2, \ldots, m\}$, $j \in \{1, 2, \ldots, n\}$.

Przykład 5.2 (Ciągi)**.** Jeżeli za Z weźmiemy zbiór liczb naturalnych $\mathbb{N}$, to otrzymamy przestrzeń liniową $\mathbb{K}^{\mathbb{N}} = \mathbb{K}^\infty$, której elementy to ciągi nieskończone o wyrazach z ciała $\mathbb{K}$. Mnożenie ciągu $x = (x_n)_{n=0}^\infty$ przez skalar α jest zdefiniowane jako

$$\alpha x = \alpha \cdot (x_n)_n = (\alpha x_n)_n,$$

a suma ciągów $x = (x_n)_{n=0}^\infty$ i $y = (y_n)_{n=0}^\infty$ to

$$x + y = (x_n)_n + (y_n)_n = (x_n + y_n)_n.$$

Ciąg, którego wszystkie wyrazy są równe zero, jest elementem neutralnym dodawania.

Przykład 5.3. Ciało $\mathbb{C}$ jest przykładem przestrzeni liniowej nad $\mathbb{R}$, określone jest bowiem mnożenie $z \in \mathbb{C}$ przez liczbę rzeczywistą $\alpha \in \mathbb{R}$, jak również dodawanie liczb zespolonych, z liczbą 0 jako elementem neutralnym.

5.2 Podprzestrzenie

Definicja 5.2. Załóżmy, że X jest przestrzenią liniową nad ciałem $\mathbb{K}$ i $Y \subset X$ jest niepustym podzbiorem, który, wraz z odziedziczonymi z X działaniami dodawania i mnożenia przez skalar także jest przestrzenią liniową nad $\mathbb{K}$. Mówimy wówczas, że Y jest *podprzestrzenią liniową* w X.

Przykład 5.4. Zbiór wszystkich wektorów postaci $[a, b, 0]$, gdzie $a, b \in \mathbb{R}$ jest podprzestrzenią liniową w $\mathbb{R}^3$.

Przykład 5.5. Przestrzeń liniowa $\mathbb{K}[x]_m$ jest podprzestrzenią liniową w $\mathbb{K}[x]_n$ dla $n \geqslant m$. $\mathbb{K}[x]_m$ jest także podprzestrzenią liniową w $\mathbb{K}[x]$.

Przykład 5.6. Jeżeli X jest przestrzenią liniową nad $\mathbb{K}$, to X jest także podprzestrzenią liniową w X i zawsze zbiór jednoelementowy $\{0\} \subset X$ jest podprzestrzenią liniową w X.

Jeżeli chcemy wiedzieć, czy dany podzbiór Y jest podprzestrzenią liniową w pewnej przestrzeni liniowej X, to musimy sprawdzić, czy dla Y spełnione są wszystkie warunki z definicji przestrzeni liniowej. Okazuje się jednak, że gdy wiemy już, że nadzbiór X jest przestrzenią liniową, to mamy nieco mniej pracy:

Stwierdzenie 5.2. *Załóżmy, że X jest przestrzenią liniową nad ciałem $\mathbb{K}$. Wówczas $Y \subset X$, $Y \neq \varnothing$ jest podprzestrzenią liniową w X wtedy i tylko wtedy, gdy zachodzą oba z poniższych warunków:*

(i) dla dowolnych $x, y \in Y$ także $x + y \in Y$,

(ii) dla dowolnych $\alpha \in \mathbb{K}$ i $y \in Y$ także $\alpha y \in Y$.

Dowód. Jeżeli Y wraz z działaniami odziedziczonym z X jest podprzestrzenią liniową w X, to oczywiście oba warunki są spełnione. Aby udowodnić implikację w drugą stronę, zauważmy najpierw, że $0 \in Y$: skoro istnieje $y \in Y$, to $0 \cdot y = 0 \in Y$ z (ii). Ponadto, dla $y \in Y$ mamy $-y = (-1) \cdot y \in Y$. Wobec (i), $(Y, +, 0)$ jest grupą przemienną. Skoro działania nie wyprowadzają poza Y, a pozostałe warunki z definicji przestrzeni liniowej są spełnione w X, to także są one spełnione w Y. Oznacza to, że Y jest przestrzenią liniową. $\square$

Omówimy teraz kilka kolejnych przykładów podprzestrzeni liniowych.

Przykład 5.7 (Płaszczyzna w $\mathbb{R}^3$)**.** Załóżmy, że $a, b, c \in \mathbb{R}$ i $a^2 + b^2 + c^2 \neq 0$. Rozważmy podzbiór

$$Y = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : ax + by + cz = 0 \right\}.$$

Krótki rachunek pokazuje, że jeśli $[x, y, z]^T \in Y$, to także $\alpha[x, y, z]^T \in Y$ dla dowolnego $\alpha \in \mathbb{R}$, a gdy także $[x', y', z']^T \in Y$, to również $[x, y, z]^T + [x', y', z']^T \in Y$. Zatem Y jest podprzestrzenią liniową w $\mathbb{R}^3$. Geometrycznie, zbiór Y interpretujemy jako płaszczyznę przechodzącą przez początek układu współrzędnych.

Przykład 5.8. Zbiory $\mathrm{TRIU}^{m,n}(\mathbb{K})$ i $\mathrm{TRIL}^{m,n}(\mathbb{K})$ są podprzestrzeniami liniowymi w $\mathbb{K}^{m,n}$.

Przykład 5.9. Zbiór wszystkich funkcji parzystych $f : \mathbb{R} \to \mathbb{R}$ (czyli takich, że dla wszystkich $x \in \mathbb{R}$ zachodzi $f(x) = f(-x)$) jest podprzestrzenią liniową w $\mathbb{R}^{\mathbb{R}}$. To samo można powiedzieć o zbiorze funkcji nieparzystych (czyli takich, że $f(x) = -f(-x)$ dla $x \in \mathbb{R}$).

Przykład 5.10. W przestrzeni wszystkich ciągów rzeczywistych $\mathbb{R}^\infty$ rozważamy podzbiór

$$X = \left\{ (x_n)_n \in \mathbb{R}^\infty : \text{ istnieje skończona granica } \lim_{n\to\infty} x_n \right\}.$$

Wówczas X jest podprzestrzenią liniową w $\mathbb{R}^\infty$.

5.3 Kombinacje liniowe i zbiory rozpinające

Definicja 5.3. Załóżmy, że X jest przestrzenią liniową nad ciałem $\mathbb{K}$, $x_1, \ldots, x_r \in X$, $\alpha_1, \ldots, \alpha_r \in \mathbb{K}$ i

$$x = \alpha_1 x_1 + \ldots + \alpha_r x_r = \sum_{j=1}^{r} \alpha_j x_j.$$

Mówimy wówczas, że wektor x jest *kombinacją liniową* wektorów $x_1, \ldots, x_r$ (o współczynnikach $\alpha_1, \ldots, \alpha_r$).

Przykład 5.11. Wektor $\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ jest kombinacją liniową wektorów

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

o współczynnikach $1, 2, 3$: $\vec{v} = \vec{e}_1 + 2\vec{e}_2 + 3\vec{e}_3$.

Przykład 5.12. Wielomian $p(x) = -4x^3 + 3x - 5$ jest kombinacją liniową wielomianów

$$q_1(x) = 2x^3 - 2x^2 - 5x + 7, \qquad \text{i} \qquad q_2(x) = x^3 + x^2 + x - 1$$

o współczynnikach -1 i -2.

Jeżeli $B = \{x_j : j \in J\}$ jest pewnym (niekoniecznie skończonym) podzbiorem przestrzeni liniowej X, to kombinacje liniowe wektorów ze zbioru B są to sumy postaci

$$\alpha_1 x_{j_1} + \alpha_2 x_{j_2} + \ldots + \alpha_k x_{j_k},$$

gdzie $j_1, \ldots, j_k \in J$.

Stwierdzenie 5.3. *Jeżeli X jest przestrzenią liniową nad ciałem $\mathbb{K}$, $\{x_j : j \in J\} \subset X$ i*

$$Y = \{\alpha_1 x_{j_1} + \ldots + \alpha_k x_{j_k} \in X : k = 1, 2, \ldots; \alpha_1, \ldots \alpha_k \in \mathbb{K}\},$$

to zbiór Y jest podprzestrzenią liniową w X

Dowód. Jeżeli $y, z \in Y$, to istnieją $j_1, \ldots, j_n \in J$ takie, że

$$y = \sum_{r=1}^{n} \alpha_r x_{j_r}, \qquad z = \sum_{r=1}^{n} \beta_r x_{j_r},$$

gdzie $\alpha_r, \beta_r \in \mathbb{K}$. Zatem także $\gamma y \in Y$ dla $\gamma \in \mathbb{K}$, jak również $y + z \in Y$. □

Definicja 5.4. Zbiór Y w stw. 5.3 nazywamy *podprzestrzenią rozpiętą* przez podzbiór B przestrzeni X. Stosujemy wówczas oznaczenie

$$Y = \operatorname{span} B.$$

Jeżeli $B = \{y_1, \ldots, y_n\}$, to możemy także napisać

$$Y = \operatorname{span}(y_1, \ldots, y_n).$$

Przykład 5.13. W $\mathbb{R}^3$ podprzestrzeń składającą się ze wszystkich wektorów postaci $[a, b, 0]^T$, gdzie $a, b \in \mathbb{R}$ możemy zapisać jako

$$\operatorname{span}(\vec{e}_1, \vec{e}_2),$$

gdzie $\vec{e}_1 = [1, 0, 0]^T$, $\vec{e}_2 = [0, 1, 0]^T$.

Przykład 5.14. W przypadku przestrzeni $\mathbb{K}[x]$ i $\mathbb{K}[x]_n$ mamy

$$\mathbb{K}[x]_n = \operatorname{span}(1, x, x^2, \ldots, x^n),$$
$$\mathbb{K}[x] = \operatorname{span}\{x^j : j \in \mathbb{N}\}.$$

Przykład 5.15. Niech $E_{k,l} = [u_{i,j}]_{i,j} \in \mathbb{K}^{m,n}$ to taka macierz, że $u_{k,l} = 1$ i $u_{i,j} = 0$ gdy $(i, j) \neq (k, l)$. Wówczas

$$\mathbb{K}^{m,n} = \operatorname{span}\{E_{k,l} : k = 1, \ldots, m, l = 1, \ldots, n\}.$$

5.4 Liniowa niezależność

Definicja 5.5. Niech X będzie przestrzenią liniową nad ciałem $\mathbb{K}$. Powiemy, że układ wektorów $x_1, x_2, \ldots, x_k \in X$ jest *liniowo zależny*, jeżeli istnieją skalary $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{K}$, nie wszystkie równe 0, takie, że

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k = 0.$$

Powiemy, że układ $x_1, x_2, \ldots, x_k$ jest *liniowo niezależny*, jeżeli nie jest on liniowo zależny, co oznacza, że spełniony jest warunek

$$\forall_{\alpha_1,\ldots,\alpha_k \in \mathbb{K}} \quad (\alpha_1 x_1 + \ldots + \alpha_k x_k = 0 \implies \alpha_1 = \ldots = \alpha_k = 0.)\,.$$

Przykład 5.16. Wektory w $\mathbb{R}^2$

$$\vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \vec{y} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

są liniowo niezależne: jeżeli $\alpha\vec{x} + \beta\vec{y} = 0$, to

$$\alpha + \beta = 0 \qquad \text{oraz} \qquad \alpha - \beta = 0,$$

skąd wynika, że $\alpha = \beta = 0$.

Przykład 5.17. Wektory w $\mathbb{R}^3$

$$\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

są liniowo zależne, gdyż

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

Stwierdzenie 5.4. *Załóżmy, że X jest przestrzenią liniową nad ciałem $\mathbb{K}$ i $x_1, \ldots, x_k \in X$. Wówczas układ $x_1, \ldots, x_k$ jest liniowo zależny wtedy i tylko wtedy, gdy jeden z wektorów x_j możemy przedstawić jako kombinację liniową pozostałych.*

Dowód. Załóżmy, że układ $x_1, \ldots, x_k$ jest liniowo zależny. Istnieją $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ nie wszystkie równe zero, takie, że

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k = 0.$$

Bez straty ogólności możemy założyć, że $\alpha_1 \neq 0$. Wówczas $\alpha_1 x_1 = -\alpha_2 x_2 - \ldots - \alpha_k x_k$, skąd

$$x_1 = -\frac{\alpha_2}{\alpha_1} x_2 - \ldots - \frac{\alpha_k}{\alpha_1} x_k.$$

Jeżeli natomiast jeden z wektorów x_j, na przykład x_1, jest kombinacją liniową pozostałych, czyli $x_1 = \alpha_2 x_2 + \ldots + \alpha_k x_k$, to dla $\alpha_1 = -1$ mamy

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k = 0.$$

□

Stwierdzenie 5.5. *Załóżmy, że X jest przestrzenią liniową nad ciałem $\mathbb{K}$, $x, x_1, \ldots, x_k \in X$ oraz układ $x_1, \ldots, x_k$ jest liniowo niezależny. Wówczas następujące warunki są równoważne:*

(i) $x \in \operatorname{span}(x_1, \ldots, x_k)$,

(ii) układ $x, x_1, \ldots, x_k$ jest liniowo zależny.

Dowód. Implikacja (i) ⇒ (ii) wynika natychmiast ze stw. 5.4. Implikację (ii) ⇒ (i) dowodzimy w następujący sposób: istnieją skalary $\alpha, \alpha_1, \ldots, \alpha_k$, nie wszystkie równe zero i takie, że

$$\alpha x + \alpha_1 x_1 + \ldots + \alpha_k x_k = 0.$$

gdyby $\alpha = 0$, to układ $x_1, \ldots, x_k$ byłby liniowo zależny, co jest sprzeczne z założeniem. Zatem

$$x = -\frac{\alpha_1}{\alpha} x_1 - \ldots - \frac{\alpha_k}{\alpha} x_k.$$

□

Uwaga 5.1. Definicję liniowej niezależności rozszerzamy na dowolne nieskończone układy wektorów w następujący sposób: powiemy, że układ wektorów $\{x_j : j \in J\}$ z przestrzeni liniowej X jest liniowo niezależny, jeżeli dla dowolnego $n \in \mathbb{N}$ i indeksów $j_1, \ldots j_n \in J$ układ $x_{j_1}, \ldots, x_{j_n}$ jest liniowo niezależny.

Zadanie 5.1. Pokaż, że dla każdego n układ funkcji $f_k : \mathbb{C} \to \mathbb{C}$, $f_k(x) = x^k$ $(k = 0, 1, \ldots, n)$ jest liniowo niezależny w $\mathbb{C}[x]$ (lub $\mathbb{C}^{\mathbb{C}}$). Następnie udowodnij to samo dla przestrzeni $\mathbb{R}[x]$ i $\mathbb{R}^{\mathbb{R}}$.

Wskazówka: Skorzystaj ze wniosku 3.3.

Twierdzenie 5.6 (Twierdzenie Steinitza o wymianie)**.** *Załóżmy, że X jest przestrzenią liniową nad ciałem $\mathbb{K}$, $n, m \geqslant 1$, układ wektorów $x_1, \ldots, x_n \in X$ jest liniowo niezależny, $y_1, \ldots, y_m \in X$ oraz*

$$\operatorname{span}(x_1, \ldots, x_n) \subseteq \operatorname{span}(y_1, \ldots, y_m) = Y.$$

Wówczas

(i) $n \leqslant m$,

(ii) n *wektorów z układu* $y_1, \ldots, y_m$ *można zastąpić wektorami* $x_1, \ldots, x_n$ *w taki sposób, że otrzymany układ nadal będzie rozpinał przestrzeń* Y.

Dowód. Dowód przeprowadzimy przez indukcję po n.

Gdy $n = 1$ mamy $1 \leqslant m$ i $x_1 = \beta_1 y_1 + \ldots + \beta_m y_m$ dla pewnych skalarów $\beta_1, \ldots, \beta_m \in \mathbb{K}$. Z założenia wektory x_i są liniowo niezależne, co w tym przypadku oznacza, że $x_1 \neq 0$. Istnieje zatem indeks j taki, że $\beta_j \neq 0$ i

$$y_j = \frac{1}{\beta_j} x_1 - \sum_{k \neq j} \frac{\beta_k}{\beta_j} y_k.$$

Pokazaliśmy, że

$$\operatorname{span}(y_1, \ldots, y_m) \subseteq \operatorname{span}(x_1, y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_m),$$

skąd wynika, że $Y = \operatorname{span}(x_1, y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_m)$.

Krok indukcyjny: Załóżmy, że teza zachodzi dla $n-1$. Wówczas $n-1 \leqslant m$. Gdyby $n - 1 = m$, to z założenia indukcyjnego wynika, że

$$\operatorname{span}(x_1, \ldots, x_{n-1}) = \operatorname{span}(y_1, \ldots, y_m) = Y$$

(wszystkie wektory y_j wymieniliśmy na x_j). Zatem $x_n \in \operatorname{span}(x_1, \ldots, x_{n-1})$, co przeczy założeniu, że wektory x_j są liniowo niezależne. Inaczej mówiąc, albo wyczerpaliśmy zasób wektorów y_j, które możemy wymienić (i wtedy indukcja zakończyła się na poprzednim kroku), albo $n - 1 < m$ i wtedy $n \leqslant m$.

Zmieniając w razie potrzeby numerację wektorów y_j, możemy napisać, że

$$Y = \operatorname{span}(x_1, \ldots, x_{n-1}, y_n, y_{n+1}, \ldots, y_m).$$

Istnieją więc skalary $\alpha_1, \ldots, \alpha_{n-1}, \beta_n, \ldots, \beta_m \in \mathbb{K}$ takie, że

$$x_n = \sum_{k=1}^{n-1} \alpha_k x_k + \sum_{k=n}^{m} \beta_k y_k.$$

Jeden ze współczynników β_j musi być różny od zera, gdyż w przeciwnym wypadku wektor x_n byłby kombinacją liniową wektorów $x_1, \ldots, x_{n-1}$, a z założenia układ $x_1, \ldots, x_n$ jest liniowo niezależny (zob. stw. 5.4). Bez straty ogólności możemy przyjąć, że $j = n$. Wówczas

$$y_n = \frac{1}{\beta_n} x_n - \sum_{k=1}^{n-1} \frac{\alpha_k}{\beta_n} x_k - \sum_{k=n+1}^{m} \frac{\beta_k}{\beta_n} y_k.$$

Inaczej mówiąc, wektor y_n jest kombinacją liniową wektorów $x_1, \ldots, x_n$, $y_{n+1}, \ldots, y_m$. Wynika z tego, że

$$Y = \operatorname{span}(x_1, \ldots, x_n, y_{n+1}, \ldots, y_m) = \operatorname{span}(y_1, \ldots, y_m).$$

□

Przykład 5.18. Niech

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{e}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

oraz

$$\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Dla dowolnego wektora $\vec{v} = [v_1, v_2, v_3, v_3]^T \in \mathbb{R}^4$ mamy

$$\begin{aligned}
\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} &= v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3 + v_4\vec{e}_4 \\
&= v_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + (v_2 - v_1)\vec{e}_2 + (v_3 - v_1)\vec{e}_3 + (v_4 - v_1)\vec{e}_4 \\
&= v_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + (v_2 - v_1) \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} + (v_3 - v_2)\vec{e}_3 + (v_4 - v_2)\vec{e}_4 \\
&= v_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + (v_2 - v_1) \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} + (v_3 - v_2) \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} + (v_4 - v_3)\vec{e}_4.
\end{aligned}$$

Czyli możemy wymieniać wektory $\vec{e}_1, \vec{e}_2, \vec{e}_3$ kolejno na $\vec{x}_1, \vec{x}_2, \vec{x}_3$ i

$$\begin{aligned}
\mathbb{R}^4 = Y &= \operatorname{span}(\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{e}_4) \\
&= \operatorname{span}(\vec{x}_1, \vec{e}_2, \vec{e}_3, \vec{e}_4) \\
&= \operatorname{span}(\vec{x}_1, \vec{x}_2, \vec{e}_3, \vec{e}_4) \\
&= \operatorname{span}(\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{e}_4).
\end{aligned}$$

Z twierdzenia Steinitza wyprowadzimy kilka wniosków, które będą przydatne w dalszej części wykładu. Poniżej cały czas zakładamy, że X oznacza pewną przestrzeń liniową nad ciałem $\mathbb{K}$.

Wniosek 5.7. *Jeżeli* $X = \operatorname{span}(x_1, \ldots, x_m)$ *i* $Y \subseteq X$ *jest podprzestrzenią liniową, to można wybrać liniowo niezależne wektory* $y_1, \ldots, y_n \in Y$ *takie, że* $n \leqslant m$ *oraz* $Y = \operatorname{span}(y_1, \ldots, y_n)$.

Dowód. Na mocy (i), każdy liniowo niezależny układ wektorów z podprzestrzeni Y ma długość co najwyżej m. Weźmy dowolny taki układ o maksymalnej długości: $y_1, \ldots, y_n$. Wtedy $n \leqslant m$ i $\operatorname{span}(y_1, \ldots, y_n) \subseteq Y$. Pozostaje pokazać, że zachodzi także inkluzja w drugą stronę. Jeżeli $y \in Y$, to układ $y, y_1, \ldots, y_n$ jest już liniowo zależny, gdyż składa się z więcej niż n wektorów. Ze stw. 5.5 wynika, że $y \in \operatorname{span}(y_1, \ldots, y_n)$. Zatem $Y \subseteq \operatorname{span}(y_1, \ldots, y_n)$. □

Wniosek 5.8. *Jeżeli* $x_1, \ldots, x_n \in X$, $y_1, \ldots, y_m \in X$ *są dwoma liniowo niezależnymi układami wektorów oraz*

$$\operatorname{span}(x_1, \ldots, x_n) = \operatorname{span}(y_1, \ldots, y_m),$$

to $n = m$.

Dowód. Z (i) wynika, że $n \leqslant m$ i $m \geqslant n$, więc $n = m$. □

5.5 Bazy i wymiar przestrzeni liniowej

Cały czas X oznacza przestrzeń liniową nad ciałem $\mathbb{K}$.

Definicja 5.6. Układ wektorów $x_1, x_2, \ldots, x_n \in X$ nazywamy *bazą* przestrzeni X, jeżeli

(i) układ ten jest liniowo niezależny,

(ii) $X = \operatorname{span}(x_1, x_2, \ldots, x_n)$.

Przykład 5.19. W przestrzeni $\mathbb{K}^n$ niech

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Wówczas układ $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ jest liniowo niezależny w $\mathbb{K}^n$ i rozpina całą przestrzeń $\mathbb{K}^n$, jest więc bazą.

Liniowa niezależność wynika z warunku $\alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \ldots + \alpha_n\vec{e}_n = \vec{0}$: porównując kolejne współrzędne lewej i prawej strony, otrzymujemy

$$\alpha_1 = 0, \alpha_2 = 0, \ldots, \alpha_n = 0.$$

Układ $\vec{e}_1, \ldots, \vec{e}_n$ będziemy nazywać *bazą standardową* przestrzeni $\mathbb{K}^n$.

Przykład 5.20. W przestrzeni $\mathbb{R}^2$ układ wektorów

$$\vec{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

jest liniowo niezależny: warunek $\alpha\vec{b}_1 + \beta\vec{b}_2 = \vec{0}$ prowadzi do układu równań

$$\alpha - \beta = 0, \qquad \alpha + 2\beta = 0,$$

którego jedyne rozwiązanie to $\alpha = \beta = 0$. Jeżeli natomiast $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$, to $\alpha, \beta \in \mathbb{R}$ takie, że

$$\begin{bmatrix} x \\ y \end{bmatrix} = \alpha\vec{b}_1 + \beta\vec{b}_2$$

wyznaczamy z układu równań

$$\alpha - \beta = x, \qquad \alpha + 2\beta = y.$$

Mamy $\alpha = \frac{1}{3}(2x+y)$, $\beta = \frac{1}{3}(-x+y)$. Zatem układ $\vec{b}_1, \vec{b}_2$ jest liniowo niezależny i rozpina przestrzeń $\mathbb{R}^2$, jest więc bazą tej przestrzeni.

Przykład 5.21. Bazą przestrzeni $\mathbb{K}^{m,n}$ są jednostki macierzowe $E_{i,j}$ dla $i = 1, \ldots, m$, $j = 1, \ldots, n$ (zob. przykład 5.15).

Przykład 5.22. Pokażemy, że układ jednomianów $1, x, x^2, \ldots, x^n$ jest bazą przestrzeni $\mathbb{R}[x]_n$. Z definicji, wielomian stopnia nie większego niż n jest to funkcja

$$p(x) = a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x_2 + \ldots + a_{n-1} \cdot x^{n-1} + a_n \cdot x^n,$$

więc $\mathbb{R}[x]_n = \operatorname{span}(1, x, x^2, \ldots, x^{n-1}, x^n)$. Pozostaje wykazać, że funkcje $f_j(x) = x^j$, $j = 0, 1, \ldots, n$, tworzą układ liniowo niezależny. Załóżmy, że dla pewnych liczb $a_0, a_1, \ldots, a_n$ mamy $a_0 f_0 + a_1 f_1 + \ldots a_n f_n = 0$, czyli, jeśli $p(x) = a_0 + a_1 x + a_2 x^2 + \ldots a_{n-1}x^{n-1} + a_n x^n$, to $p(x) = 0$ dla każdego $x \in \mathbb{R}$. Rozważmy ciąg liczbowy $(c_k)_{k=1}^n$:

$$c_k = \frac{p(k)}{k^n} = \frac{a_0}{k^n} + \frac{a_1}{k^{n-1}} + \ldots + \frac{a_{n-1}}{k} + a_n.$$

Z jednej strony mamy $c_k = 0$ dla każdego $k \in \mathbb{N}$, więc $\lim_{k\to\infty} c_k = 0$. Z drugiej strony $\lim_{k\to\infty} c_k = a_n$, więc $a_n = 0$. Następnie w ten sam sposób pokazujemy, że $a_{n-1} = \ldots = a_1 = 0$ i zostaje nam równość $a_0 = 0$.

Uwaga 5.2. Definicję liniowej niezależności rozszerzamy na dowolne nieskończone układy wektorów w następujący sposób: powiemy, że układ wektorów $\{x_j : j \in J\}$ z przestrzeni liniowej X jest liniowo niezależny, jeżeli dla dowolnego $n \in \mathbb{N}$ i indeksów $j_1, \ldots j_n \in J$ układ $x_{j_1}, \ldots, x_{j_n}$ jest liniowo niezależny.

Najważniejszą cechą baz jest to, że za ich pomocą można *jednoznacznie* przedstawić dowolny wektor z przestrzeni liniowej:

Twierdzenie 5.9. *Niech $x_1, x_2, \ldots, x_n \in X$. Wówczas następujące warunki są równoważne:*

(i) Układ $x_1, \ldots, x_n$ jest bazą przestrzeni X.

(ii) Dla każdego wektora $x \in X$ istnieją jednoznacznie wyznaczone skalary $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ takie, że

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n.$$

Dowód. Najpierw pokażemy implikację (i) $\Rightarrow$ (ii). Załóżmy, że układ $x_1, \ldots, x_n$ jest bazą przestrzeni X. Wtedy $X = \operatorname{span}(x_1, \ldots, x_n)$ i każdy wektor $x \in X$ jest kombinacją liniową wektorów $x_1, \ldots, x_n$. Załóżmy, że mamy dwie takie kombinacje liniowe:

$$x = \sum_{k=0}^{n} \alpha_k x_k = \sum_{k=0}^{n} \beta_k x_k.$$

Wówczas

$$\sum_{k=0}^{n} (\alpha_k - \beta_k) x_k = 0$$

i z liniowej niezależności układu $x_1, \ldots, x_n$ wynika, że

$$\alpha_1 = \beta_1, \quad \alpha_2 = \beta_2, \quad \ldots, \quad \alpha_n = \beta_n.$$

Pokażemy implikację (ii) $\Rightarrow$ (i). Jeżeli zachodzi (ii), to $X = \operatorname{span}(x_1, \ldots x_n)$. Uzasadnimy liniową niezależność. Jeżeli

$$\alpha_1 x_1 + \ldots + \alpha_n x_n = 0 = 0x_1 + \ldots + 0x_n,$$

to z jednoznaczności przedstawienia wektora 0 za pomocą $x_1, \ldots, x_n$ wynika, że $\alpha_1 = \ldots = \alpha_n = 0$. □

Twierdzenie 5.10. *Załóżmy, że $x_1, x_2, \ldots, x_k \in X$. Następujące warunki są równoważne:*

(i) układ $x_1, x_2, \ldots, x_k$ jest bazą przestrzeni X,

(ii) *układ $x_1, x_2, \ldots, x_k$ jest maksymalnym (w sensie inkluzji) układem liniowo niezależnym w przestrzeni X,*

(iii) *układ $x_1, x_2, \ldots, x_k$ jest minimalnym (w sensie inkluzji) układem rozpinającym przestrzeń X.*

Uwaga 5.3. Sformułowanie układ *maksymalny w sensie inkluzji* oznacza, że danego układu nie da się rozszerzyć do większego układu, który nadal spełnia podany warunek (liniową niezależność). Podobnie, układ *minimalny w sensie inkluzji* oznacza, że po usunięciu z niego dowolnego elementu nie będzie spełniony podany warunek (rozpinanie przestrzeni X).

Dowód. Implikację (i) $\Rightarrow$ (ii) udowodnimy nie wprost: Jeżeli układ $x_1, \ldots, x_k$ nie jest maksymalnym układem liniowo niezależnym w X, to znajdziemy wektor $x \in X$ taki, że układ $x, x_1, \ldots, x_k$ też będzie liniowo niezależny. Wówczas $x \notin \text{span}(x_1, \ldots, x_x)$, co jednak przeczy założeniu (i).

Dowód implikacji (ii) $\Rightarrow$ (i): Dla każdego $x \in X$ układ $x, x_1, \ldots, x_k$ jest liniowo zależny. Zgodnie ze stw. 5.5, $x \in \text{span}(x_1, \ldots, x_k)$, więc układ $x_1, \ldots, x_k$ jest bazą.

Dowód implikacji (i) $\Rightarrow$ (iii) również przeprowadzamy nie wprost: Jeżeli układ $x_1, \ldots, x_k$ nie jest minimalny, to zawiera on podukład $x_{j_1}, \ldots, x_{j_r}$ rozpinający przestrzeń X, przy czym $r < k$. Istnieje j takie, że $j \neq j_i$ dla każdego indeksu j_i i wektor x_j jest kombinacją liniową wektorów $x_{j_1}, \ldots, x_{j_r}$. Stąd wynika, że cały układ $x_1, \ldots, x_k$ jest liniowo zależny, nie może więc być bazą.

Dowód implikacji (iii) $\Rightarrow$ (i): Należy pokazać liniową niezależność układu $x_1, \ldots, x_k$. Gdyby ten układ był liniowo zależny, to ze stw. 5.4 wynika, że jeden z wektorów x_j byłby kombinacją liniową pozostałych. Wówczas te pozostałe wektory będą rozpinać całą przestrzeń X. To jednak przeczy założeniu o minimalności układu $x_1, \ldots, x_k$. □

Twierdzenie 5.11. *Jeżeli układy wektorów $x_1, \ldots, x_n$ i $y_1, \ldots, y_m$ są bazami w przestrzeni X, to $n = m$.*

Dowód. Teza jest natychmiastową konsekwencją wniosku 5.8. □

Definicja 5.7. *Wymiarem* przestrzeni liniowej X nazywamy liczbę elementów dowolnej bazy przestrzeni X. Jeżeli wymiar przestrzeni X wynosi n, to piszemy

$$\dim X = n.$$

Z wcześniejszych przykładów wynika, że

- $\dim \mathbb{K}^n = n$,

- $\dim \mathbb{K}^{m,n} = mn$,
- $\dim \mathbb{K}[x]_n = n + 1$,
- wymiar $\mathbb{C}$ jako przestrzeni liniowej nad $\mathbb{R}$ wynosi 2.

Uwaga 5.4. Jeżeli przestrzeń X zawiera nieskończony układ liniowo niezależny, to mówimy, że $\dim X = \infty$.

Przykładem takiej przestrzeni jest $\mathbb{K}[x]$. Ponieważ $\mathbb{K}[x] \subset \mathbb{K}^{\mathbb{K}}$, więc

$$\dim \mathbb{K}[x] = \dim \mathbb{K}^{\mathbb{K}} = \infty.$$

Jednakże o ile w przypadku $\mathbb{K}[x]_n$ jesteśmy w stanie wskazać bazę: jest to układ x^j, $j = 0, 1, 2, \ldots$, to w przypadku przestrzeni $\mathbb{K}^{\mathbb{K}}$ (a także $\mathbb{K}^{\infty}$) nie jesteśmy w stanie podać konstrukcji bazy. Za pomocą lematu Kuratowskiego - Zorna można jedynie pokazać, że takie bazy istnieją.

Twierdzenie 5.12. *Załóżmy, że* $\dim X = n < \infty$. *Wówczas*

(i) Każdy liniowo niezależny układ wektorów z X można uzupełnić do bazy przestrzeni X.

(ii) Z każdego układu wektorów rozpinających przestrzeń X można wybrać bazę przestrzeni X.

(iii) Jeżeli pewne wektory $x_1, \ldots, x_n \in X$ są liniowo niezależne, to są one bazą przestrzeni X.

(iv) Jeżeli pewne wektory $x_1, \ldots, x_n \in X$ rozpinają przestrzeń X, to są one jej bazą.

Dowód. Podpunkt (i): niech $x_1, \ldots, x_k \in X$ będzie układem liniowo niezależnych wektorów. Spośród wszystkich układów $x_1, \ldots, x_k, y_{k+1}, \ldots, y_m$ rozpinających przestrzeń X wybieramy najkrótszy (stosujemy zasadę minimum). Zgodnie z tw. 5.10 otrzymamy bazę przestrzeni X.

(ii): Jeżeli $X = \operatorname{span}(x_1, \ldots, x_r)$, to z układu $x_1, \ldots, x_r$ możemy wybrać najkrótszy podukład rozpinający X. Zgodnie z tw. 5.10 otrzymamy bazę przestrzeni X.

(iii): gdyby układ $x_1, \ldots, x_n$ nie był bazą przestrzeni X, to z (i) moglibyśmy go przedłużyć do bazy. Ale wówczas otrzymalibyśmy bazę składającą się z więcej niż $n = \dim X$ wektorów, a to jest niemożliwe na mocy tw. 5.11.

(iv): Zgodnie z (ii) z układu $x_1, \ldots, x_n$ możemy wybrać bazę. Gdyby baza ta składała się z mniej niż n wektorów, to znowu mielibyśmy sprzeczność z założeniem, że $\dim X = n$. □

Stwierdzenie 5.13. *Załóżmy, że* $\dim X < \infty$ *i* $Y \subset X$ *jest podprzestrzenią liniową. Wówczas*

(i) $\dim Y \leqslant \dim X$,

(ii) $\dim Y = \dim X$ *wtedy i tylko wtedy, gdy* $Y = X$.

Dowód. Niech układ $x_1, \ldots, x_n$ będzie bazą przestrzeni X. Z wniosku 5.7 wynika, że istnieje baza przestrzeni Y złożona z wektorów $y_1, \ldots, y_k \in Y$, przy czym $k \leqslant n$. Zatem $\dim Y \leqslant \dim X$.

Jeżeli $\dim Y = \dim X$ i $y_1, \ldots, y_n$ jest bazą podprzestrzeni Y, to jest to także układ liniowo niezależny w X, składający się z tylu wektorów, ile wynosi $\dim X$. Z tw. 5.12 (ii) wynika, że jest to baza X, więc $X = \operatorname{span}(y_1, \ldots, y_n) = Y$. □

5.6 Przecięcie podprzestrzeni

Często spotykaną konstrukcją podprzestrzeni liniowej jest wzięcie części wspólnej większych podprzestrzeni liniowych.

Stwierdzenie 5.14. *Załóżmy, że* X *jest przestrzenią liniową nad ciałem* $\mathbb{K}$ *i* $Y_j \subset X$ *są podprzestrzeniami liniowymi, gdzie* $j \in J$. *Wówczas zbiór*

$$Y = \bigcap_{j \in J} Y_j$$

też jest podprzestrzenią liniową w X.

Zadanie 5.2. Udowodnij stw. 5.14.

Przykład 5.23. Przecięcie przestrzeni macierzy trójkątnych górnych $n \times n$ i trójkątnych dolnych $n \times n$ daje podprzestrzeń w $\mathbb{K}^{n,n}$, której elementy to macierze diagonalne.

Przykład 5.24. Przecięcie dwóch różnych płaszczyzn w $\mathbb{R}^3$ zawierających 0 jest podprzestrzenią liniową – prostą przechodzącą przez 0.

Zadanie 5.3. Udowodnij, że w $\mathbb{R}^{\mathbb{R}}$ przecięcie podprzestrzeni funkcji parzystych i funkcji nieparzystych daje podprzestrzeń trywialną $\{0\}$, której jedyny element to funkcja zerowa.

5.7 Suma i suma prosta podprzestrzeni

Poznamy teraz inny sposób konstrukcji podprzestrzeni liniowych:

Definicja 5.8. Załóżmy, że X jest przestrzenią liniową nad ciałem $\mathbb{K}$ i $U, V \subset X$ są podprzestrzeniami liniowymi. *Suma podprzestrzeni* U i V jest to zbiór

$$U + V = \{u + v \in X : u \in U, v \in V\}.$$

Uwaga 5.5. Załóżmy, że U i V to dwie podprzestrzenie liniowe w przestrzeni liniowej X. W ogólności nie jest prawdą, że

$$U + V = U \cup V.$$

Łatwo można zobaczyć, że zawsze $U \cup V \subset U + V$. Ponadto, zbiór $U \cup V$ nie musi być podprzestrzenią liniową w X.

Stwierdzenie 5.15. *Załóżmy, że X jest przestrzenią liniową nad ciałem $\mathbb{K}$ i $U, V \subset X$ są podprzestrzeniami liniowymi. Wówczas ich suma $U + V$ też jest podprzestrzenią liniową w X.*

Dowód. Zauważmy, że $U + V \subset \operatorname{span}(U \cup V)$. Z drugiej strony, jeżeli $x \in \operatorname{span}(U \cup V)$, to

$$x = \alpha_1 u_1 + \ldots \alpha_k u_k + \beta_1 v_1 + \ldots + \beta_l v_l$$

dla pewnych $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_l \in \mathbb{K}$, $u_1, \ldots, u_k \in U$, $v_1, \ldots, v_l \in V$. Niech

$$u = \alpha_1 u_1 + \ldots + \alpha_k u_k \in U, \qquad v = \beta_1 v_1 + \ldots + \beta_l v_l \in V.$$

Teraz $x = u + v \in U + V$. Czyli $\operatorname{span}(U \cup V) \subset U + V$. Pokazaliśmy, że $U + V = \operatorname{span}(U \cup V)$, a wiemy, że $\operatorname{span}(U \cup V)$ jest podprzestrzenią liniową. □

Przykład 5.25. W przestrzeni $\mathbb{R}^3$ dane są podprzestrzenie:

$$U = \operatorname{span}(\vec{e}_1, \vec{e}_2) \qquad V = \operatorname{span}(\vec{e}_2, \vec{e}_3).$$

Wówczas $\mathbb{R}^3 = U + V$. Istotnie, dla $\vec{x} \in \mathbb{R}^3$ i dowolnego $\alpha \in \mathbb{R}$ mamy

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \underbrace{x_1\vec{e}_1 + (x_2 - \alpha)\vec{e}_2}_{\vec{u}} + \underbrace{\alpha\vec{e}_2 + x_3\vec{e}_3}_{\vec{v}}$$

i $\vec{u} \in U$, $\vec{v} \in V$. Zatem $\mathbb{R}^3 \subset U + V$. Z drugiej strony jasne jest, że $U + V \subset \mathbb{R}^3$, więc ostatecznie $\mathbb{R}^3 = U + V$.

Zauważmy także, że $U \cap V = \operatorname{span}(\vec{e}_2)$.

Uwaga 5.6. Definicję sumy dwóch podprzestrzeni można indukcyjnie rozszerzyć na sumę dowolnej skończonej liczby podprzestrzeni: jeżeli X jest przestrzenią liniową nad $\mathbb{R}$ i $U_1, U_2, \ldots, U_k \subset U$ są podprzestrzeniami liniowymi, to określamy podprzestrzenie:

$$\begin{aligned} &V_1 = U_1, \\ &V_2 = V_1 + U_2, \\ &\ldots\ldots \\ &V_j = V_{j-1} + U_j, \qquad \text{dla } j = 2, ..., k \end{aligned}$$

i definiujemy

$$U_1 + U_2 + \ldots + U_k = V_k.$$

Równoważnie

$$U_1 + U_2 + \ldots + U_k = \{u_1 + u_2 + \ldots + u_k : u_j \in V_j \text{ dla } j = 1, \ldots, k\}$$

W przykładzie 5.25 zauważyliśmy, że przecięcie podprzestrzeni U i V jest niezerową podprzestrzenią $\mathrm{span}(\vec{e}_2)$. Ponadto, wektory $u \in U$ i $v \in V$ takie, że $x = u + v$ nie były wyznaczone jednoznacznie. Okazuje się, że te dwie obserwacje są ze sobą powiązane.

Stwierdzenie 5.16. *Załóżmy, że U, V są podprzestrzeniami liniowymi w przestrzeni liniowej X nad ciałem $\mathbb{K}$. Wówczas następujące warunki są równoważne:*

(i) $U \cap V = \{0\}$,

(ii) *dla każdego wektora $x \in U + V$ wektory $u \in U$ i $v \in V$ takie, że $x = u + v$ są wyznaczone jednoznacznie.*

Dowód. Pokażemy implikację (i) $\Rightarrow$ (ii): Niech $x \in U + V$ i

$$x = u + v = u' + v', \qquad u, u' \in U, \quad v, v' \in V.$$

Zatem $y = u - u' = v' - v \in U \cap V$, więc $y = 0$, czyli $u = u'$ i $v = v'$.

Teraz pokażemy implikację (ii) $\Rightarrow$ (i): zauważmy, że $U \cap V \subset U + V$. Załóżmy, że $y \in U \cap V$. Istnieją więc jednoznacznie wyznaczone wektory $u \in U$ i $v \in V$ takie, że $y = u + v$. Ale y możemy zapisać jako sumę wektora z U i wektora z V także na następujące sposoby:

$$y = y + 0 = 0 + y.$$

Skoro u i v są wyznaczone jednoznacznie, to $u = v = 0$, więc $y = 0$, czyli $U \cap V = \{0\}$. □

Definicja 5.9. Załóżmy, że X jest przestrzenią liniową nad ciałem $\mathbb{K}$ i $U, V \subset X$ są podprzestrzeniami liniowymi, przy czym $U \cap V = \{0\}$. Wówczas mówimy, że podprzestrzeń $U + V$ jest *sumą prostą* podprzestrzeni U i V i piszemy $U + V = U \oplus V$.

Stwierdzenie 5.17. *Jeżeli X jest przestrzenią liniową nad ciałem $\mathbb{K}$, $U, V \subset X$ to podprzestrzenie liniowe i $X = U \oplus V$, to dla każdego wektora $x \in X$, istnieją jednoznacznie wyznaczone wektory $u \in U$ oraz $v \in V$ takie, że*

$$x = u + v.$$

Dowód. Stwierdzenie to jest bezpośrednim wnioskiem ze stw. 5.16. □

Uwaga 5.7. Jeżeli dane są podprzestrzenie liniowe $U_1, \ldots, U_k \subset X$ takie, że $U_i \cap U_j = \{0\}$ gdy $i \neq j$, to można, podobnie jak w przypadku sumy skończonej liczby podprzestrzeni (Uwaga 5.6) zdefiniować sumę prostą

$$Y = U_1 \oplus U_2 \oplus \ldots \oplus U_k \subset X.$$

Wówczas dla każdego wektora $y \in Y$ istnieją jednoznacznie wyznaczone wektory $u_1 \in U_1, \ldots, u_k \in U_k$ takie, że

$$y = u_1 + u_2 + \ldots + u_k.$$

Przykład 5.26. W $\mathbb{R}^3$ rozważamy podprzestrzenie

$$U = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : x + y + z = 0 \right\}, \qquad V = \operatorname{span}\left(\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right).$$

Pokażemy, że $\mathbb{R}^3 = U \oplus V$. Najpierw uzasadnimy, że $U \cap V = \{0\}$. Niech

$$\vec{w} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in U \cap V.$$

Zatem $\vec{w} \in U$, skąd $x + y + z = 0$, oraz $\vec{w} \in V$, skąd $\vec{w} = [0, \alpha, \alpha]^T$. Wobec tego $x = 0$ i $y = z = \alpha$. Ale jednocześnie $y + z = 0$, więc $y = z = \alpha = 0$. Ostatecznie $\vec{w} = 0$. Pokazaliśmy, że $U \cap V = 0$. Zatem $U + V = U \oplus V$.

Teraz wystarczy pokazać, że każdy wektor $[a, b, c]^T \in \mathbb{R}^3$ możemy przedstawić jako sumę wektorów $\vec{u} \in U$ i $\vec{v} \in V$. Wektor $[a, b, c]^T$ chcemy przedstawić jako

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \alpha \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \qquad \alpha \in \mathbb{R}$$

przy czym $x + y + z = 0$. Mamy zatem równania

$$\begin{aligned} a &= x \\ b &= y + \alpha \\ c &= z + \alpha \\ 0 &= x + y + z \end{aligned}$$

Z 1. i 4. równania dostajemy $y + z = -a$, z 2. i 3. $y - z = b - c$. Dodając i odejmując stronami, oraz dzieląc przez 2, otrzymujemy

$$y = \frac{-a + b - c}{2} \qquad z = \frac{-a - b + c}{2}.$$

Z 2. lub 3. równania wyznaczamy jeszcze

$$\alpha = \frac{a + b + c}{2}.$$

Ostatecznie okazuje się, że wektor $[a, b, c]^T$ możemy zapisać w następujący sposób:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a \\ \frac{-a+b-c}{2} \\ \frac{-a-b+c}{2} \end{bmatrix} + \frac{a+b+c}{2} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Zadanie 5.4. Udowodnij, że przestrzeń $\mathbb{R}^{\mathbb{R}}$ jest sumą prostą przestrzeni funkcji parzystych i przestrzeni funkcji nieparzystych.

Zajmiemy się teraz wymiarem sumy podprzestrzeni liniowych.

Twierdzenie 5.18. *Jeżeli $Y, Z \subset X$ są podprzestrzeniami liniowymi skończonego wymiaru, to*

(i) wymiary podprzestrzeni $Y \cap Z$ oraz $Y + Z$ też są skończone,

(ii) $\dim(Y + Z) = \dim Y + \dim Z - \dim(Y \cap Z)$.

Dowód. Ponieważ $Y \cap Z \subset Y$ i $\dim Y < \infty$, więc ze stw. 5.13 wynika, że $\dim(Y \cap Z) \leqslant \dim Y < \infty$. Niech $m = \dim Y \cap Z$ i przyjmijmy, że $\dim Y = m + k$, $\dim Z = m + l$. Wybierzmy układ $x_1, \ldots, x_m \in (Y \cap Z)$, który jest bazą $Y \cap Z$. Zgodnie z tw. 5.12 (i), układ ten możemy uzupełnić do

- bazy $x_1, \ldots, x_m, y_1, \ldots, y_k$ przestrzeni Y,
- bazy $x_1, \ldots, x_m, z_1, \ldots, z_l$ przestrzeni Z.

Pokażemy, że układ $x_1, \ldots, x_m, y_1, \ldots, y_k, z_1, \ldots, z_l$ jest liniowo niezależny. Załóżmy, że

$$\alpha_1 x_1 + \ldots + \alpha_n x_m + \beta_1 y_1 + \ldots + \beta_k y_k + \gamma_1 z_1 + \ldots + \gamma_l z_l = 0.$$

Niech $x = -\gamma_1 z_1 - \ldots - \gamma_l z_l$. Wówczas

$$\alpha_1 x_1 + \ldots + \alpha_n x_m + \beta_1 y_1 + \ldots + \beta_k y_k = -\gamma_1 z_1 - \ldots - \gamma_l z_l = x.$$

Zauważmy, że $x \in Z \cap Y$, czyli $x = \alpha_1 x_1 + \ldots + \alpha_m x_m$ – korzystamy z tego, że układ $x_1, \ldots, x_m$ jest bazą podprzestrzeni $Y \cap Z$. Z tw 5.9 wynika, że $\beta_1 = \ldots = \beta_k = 0$. Tak samo pokazujemy, że $\alpha_1 = \ldots = \alpha_k = 0$. Zatem

$$\gamma_1 z_1 + \ldots + \gamma_l z_l = 0,$$

i z liniowej niezależności $z_1, \ldots, z_l$ dostajemy, że $\gamma_1 = \ldots = \gamma_l = 0$.

Pozostało nam zauważyć, że

$$Y + Z = \operatorname{span}(x_1, \ldots, x_m, y_1, \ldots, y_k, z_1, \ldots, z_l),$$

więc układ $x_1, \ldots, x_m, y_1, \ldots, y_k, z_1, \ldots, z_l$ jest bazą przestrzeni $Y + Z$ i

$$\dim(Y+Z) = m+k+l = (m+k)+(m+l)-m = \dim Y + \dim Z - \dim(Y \cap Z).$$

Tym samym pokazaliśmy też, że $\dim(Y + Z) < \infty$. $\square$

Wniosek 5.19. *Jeżeli* $Y, Z \subset X$ *i* $Y \cap Z = \{0\}$, *to*

$$\dim(Y \oplus Z) = \dim Y + \dim Z.$$

Rozdział 6

Obraz, jądro i rząd macierzy

6.1 Obraz i jądro macierzy

Dowolną macierz $A \in \mathbb{K}^{m,n}$ możemy interpretować jako funkcję

$$\Phi_A : \mathbb{K}^n \to \mathbb{K}^m \qquad \Phi_A(\vec{x}) = A\vec{x}.$$

(zob. też przykład 4.7).

Przykład 6.1. Niech

$$A = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & -3 \\ 1 & 0 & 2 \end{bmatrix} \in \mathbb{R}^{3,3}.$$

- Wyznaczymy obraz funkcji Φ_A, czyli zbiór

  $$\Phi_A(\mathbb{R}^3) = \left\{\vec{y} \in \mathbb{R}^3 : \vec{y} = A\vec{x} \text{ dla pewnego } x \in \mathbb{R}^3\right\} \subset \mathbb{R}^3.$$

  Weźmy dowolny wektor $\vec{x} = [x_1, x_2, x_3]^T \in \mathbb{R}^3$. Wówczas

  $$\Phi_A(\vec{x}) = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & -3 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ -3 \\ 2 \end{bmatrix}.$$

  Zatem

  $$\Phi_A(\vec{x}) \in \operatorname{span}\left(\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -3 \\ 2 \end{bmatrix}\right).$$

  Jeżeli natomiast

  $$\vec{y} \in \operatorname{span}\left(\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -3 \\ 2 \end{bmatrix}\right),$$

to istnieją $\alpha, \beta, \gamma \in \mathbb{R}$ takie, że

$$\vec{y} = \alpha \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} 0 \\ -3 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & -3 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \in \Phi_A(\mathbb{R}^3).$$

Zatem

$$\Phi_A(\mathbb{R}^3) = \operatorname{span}\left(\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -3 \\ 2 \end{bmatrix} \right).$$

Zbiór $\Phi_A(\mathbb{R}^3)$ okazał się być przestrzenią liniową.

- Wyznaczymy teraz przeciwobraz zbioru $\{0\}$, czyli zbiór

$$\Phi_A^{-1}(\{0\}) = \left\{ \vec{x} \in \mathbb{R}^3 : A\vec{x} = 0 \right\} \subset \mathbb{R}^3.$$

Niech $\vec{x} = [x_1, x_2, x_3]^T$. Wówczas $\vec{x} \in \Phi_A^{-1}(\{0\})$ wtedy i tylko wtedy, gdy liczby x_1, x_2, x_3 są rozwiązaniami układu równań

$$\begin{cases} x_1 + 2x_2 = 0 \\ -x_1 + x_2 - 3x_3 = 0 \\ x_1 + 2x_3 = 0 \end{cases}$$

Z ostatniego równania wyznaczamy $x_1 = -2x_3$, a następnie z pierwszego równania dostajemy $x_2 = -\frac{1}{2}x_1 = x_3$. Dla takich wartości x_1, x_2 środkowe równanie jest spełnione. Zatem

$$\vec{x} = \begin{bmatrix} -2x_3 \\ x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}.$$

Sprawdzamy, że $\alpha[-2, 1, 1]^T \in \Phi_A^{-1}(\{0\})$ dla dowolnego $\alpha \in \mathbb{R}$. Ostatecznie

$$\Phi_A^{-1}(\{0\}) = \operatorname{span}\left(\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \right).$$

Podobnie jak w przypadku obrazu, także zbiór $\Phi_A^{-1}(\{0\})$ jest przestrzenią liniową.

Zauważmy jeszcze, że z równości

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

mamy
$$\begin{bmatrix} 0 \\ -3 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix},$$
czyli wektory rozpinające $\Phi_A(\mathbb{R}^3)$ są liniowo zależne. Możemy więc napisać
$$\Phi_A(\mathbb{R}^3) = \operatorname{span}\left(\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}\right).$$
Te dwa wektory są już liniowo niezależne, jest to więc baza przestrzeni liniowej $\Phi_A(\mathbb{R}^3)$. Pokazaliśmy także, że
$$\dim \Phi_A(\mathbb{R}^3) = 2, \qquad \dim \Phi_A^{-1}(\{0\}) = 1.$$

Definicja 6.1. Rozważamy macierz $A \in \mathbb{K}^{m,n}$.

- *Obrazem* macierzy A nazywamy zbiór
$$\operatorname{im} A = \{A\vec{x} \in \mathbb{K}^m : x \in \mathbb{K}^n\}.$$
- *Jądrem* macierzy A nazywamy zbiór
$$\ker A = \{\vec{x} \in \mathbb{K}^n : A\vec{x} = 0\}.$$

W przykładzie 6.1 wyznaczyliśmy obraz i jądro pewnej konkretnej macierzy $A \in \mathbb{R}^{3,3}$. Okazało się, że oba te zbiory są przestrzeniami liniowymi. Jest tak w ogólnym przypadku.

Stwierdzenie 6.1. *Załóżmy, że* $A \in \mathbb{K}^{m,n}$. *Wówczas*

(i) Zbiór $\operatorname{im} A$ *jest podprzestrzenią liniową w* $\mathbb{K}^m$.

(ii) Zbiór $\ker A$ *jest podprzestrzenią liniową w* $\mathbb{K}^n$.

Dowód. Podpunkt (i) dowodzimy następująco: załóżmy, że $\vec{y}, \vec{z} \in \operatorname{im} A$ i $\alpha, \beta \in \mathbb{K}$. Wystarczy pokazać, że $\alpha\vec{y} + \beta\vec{z} \in \operatorname{im} A$. Z definicji $\operatorname{im} A$, istnieją wektory $\vec{u}, \vec{v} \in \mathbb{K}^n$ takie, że
$$\vec{y} = A\vec{u}, \qquad \vec{z} = A\vec{v}.$$
Wówczas
$$\alpha\vec{y} + \beta\vec{z} = \alpha A\vec{u} + \beta A\vec{v} = A(\alpha\vec{u}) + A(\beta\vec{v}) = A(\alpha\vec{u} + \beta\vec{v}) \in \operatorname{im} A.$$

Podpunkt (ii) dowodzimy następująco: załóżmy, że $\vec{u}, \vec{v} \in \ker A$ i $\alpha, \beta \in \mathbb{K}$. Wystarczy pokazać, że $\alpha\vec{u} + \beta\vec{v} \in \ker A$. Mamy
$$A(\alpha\vec{u} + \beta\vec{v}) = \alpha A\vec{u} + \beta A\vec{v} = \alpha \cdot 0 + \beta \cdot 0 = 0.$$

□

Przykład 6.2. Niech $I_n \in \mathbb{K}^{n,n}$ to macierz jednostkowa. Wówczas

$$\operatorname{im} I_n = \mathbb{K}^n, \qquad \ker I_n = \{0\} \subset \mathbb{K}^n.$$

Uwaga 6.1. Jeżeli macierz $A \in \mathbb{K}^{m,n}$ zapiszemy w postaci blokowo-kolumnowej:

$$A = [\vec{a}_1, \ldots, \vec{a}_n], \qquad \vec{a}_k \in \mathbb{K}^m,$$

to, dla $\vec{x} = [x_1, \ldots, x_n]^T \in \mathbb{K}^n$ mamy

$$A\vec{x} = x_1\vec{a}_1 + \ldots + x_n\vec{a}_n.$$

Podprzestrzeń $\operatorname{im} A$ jest więc rozpięta przez kolumny macierzy A.

Stwierdzenie 6.2. *Załóżmy, że* $A \in \mathbb{K}^{m,n}$, $\dim\ker A = k$ *i układ wektorów* $\vec{u}_1, \ldots, \vec{u}_k \in \mathbb{K}^n$ *jest bazą podprzestrzeni* $\ker A$. *Układ ten uzupełniamy wektorami* $\vec{u}_{k+1}, \ldots, \vec{u}_n$ *do bazy całej przestrzeni* $\mathbb{K}^n$. *Wówczas układ* $n-k$ *wektorów*

$$\vec{y}_{k+1} = A\vec{u}_{k+1}, \quad \ldots, \quad \vec{y}_n = A\vec{u}_n \in \mathbb{K}^m$$

jest bazą podprzestrzeni $\operatorname{im} A$.

Dowód. Pokażemy najpierw, że $\operatorname{im} A = \operatorname{span}(\vec{y}_{k+1}, \ldots, \vec{y}_n)$. Inkluzja „$\supset$” jest oczywista. Załóżmy, że $y \in \operatorname{im} A$. Istnieje $\vec{x} \in \mathbb{K}^n$ takie, że $\vec{y} = A\vec{x}$. Ponieważ układ $\vec{u}_1, \ldots, \vec{u}_n$ jest bazą przestrzeni $\mathbb{K}^n$, istnieją skalary $\alpha_1, \ldots, \alpha_n$, dla których

$$\vec{x} = \alpha_1\vec{u}_1 + \ldots + \alpha_k\vec{u}_k + \alpha_{k+1}\vec{u}_{k+1} + \ldots + \alpha_n\vec{u}_n.$$

Wobec tego

$$\begin{aligned}
\vec{y} &= A(\alpha_1\vec{u}_1 + \ldots + \alpha_k\vec{u}_k + \alpha_{k+1}\vec{u}_{k+1} + \ldots + \alpha_n\vec{u}_n) \\
&= \alpha_1 A\vec{u}_1 + \ldots + \alpha_k A\vec{u}_k + \alpha_{k+1} A\vec{u}_{k+1} + \ldots + \alpha_n A\vec{u}_n \\
&= \alpha_{k+1} A\vec{u}_{k+1} + \ldots + \alpha_n A\vec{u}_n \\
&= \alpha_{k+1}\vec{y}_{k+1} + \ldots + \alpha_n\vec{y}_n \\
&\in \operatorname{span}(\vec{y}_{k+1}, \ldots, \vec{y}_n).
\end{aligned}$$

Pozostaje nam pokazać, że układ $\vec{y}_{k+1}, \ldots, \vec{y}_n$ jest liniowo niezależny. Załóżmy, że $0 = \beta_{k+1}\vec{y}_{k+1} + \ldots + \beta_n\vec{y}_n$ dla pewnych skalarów $\beta_{k+1}, \ldots, \beta_n \in \mathbb{K}$. Mamy

$$\begin{aligned}
0 &= \beta_{k+1}\vec{y}_{k+1} + \ldots + \beta_n\vec{y}_n \\
&= \beta_{k+1} A\vec{u}_{k+1} + \ldots + \beta_n A\vec{u}_n \\
&= A(\beta_{k+1}\vec{u}_{k+1} + \ldots + \beta_n\vec{u}_n).
\end{aligned}$$

Zatem $\vec{v} = \beta_{k+1}\vec{u}_{k+1} + \ldots + \beta_n\vec{u}_n \in \ker A$. Istnieją więc skalary $\beta_1, \ldots, \beta_k$ takie, że $\vec{v} = \beta_1\vec{u}_1 + \ldots + \beta_k\vec{u}_k$ i

$$\beta_1\vec{u}_1 + \ldots + \beta_k\vec{u}_k - \beta_{k+1}\vec{u}_{k+1} - \ldots - \beta_n\vec{u}_n = 0.$$

Ponieważ układ $u_1, \ldots, u_n$ jest liniowo niezależny, jako baza przestrzeni $\mathbb{K}^n$, z ostatniej równości wynika, że $\beta_{k+1} = \ldots = \beta_n = 0$. Tym samym pokazaliśmy liniową niezależność wektorów $y_{k+1}, \ldots, y_n$. $\square$

Wnioskiem z powyższego faktu jest następujące ważne

Twierdzenie 6.3 (Twierdzenie o wymiarze obrazu i jądra)**.** *Niech $A \in \mathbb{K}^{m,n}$. Wówczas*

$$\dim \operatorname{im} A + \dim \ker A = n.$$

Dowód. Na mocy stw. 6.2 wystarczy zauważyć, że baza $\operatorname{im} A$ ma tyle samo elementów, ile potrzeba, aby uzupełnić bazę $\ker A$ do bazy całej przestrzeni $\mathbb{K}^n$. $\square$

6.2 Rząd macierzy

Definicja 6.2. Rozważamy macierz $A \in \mathbb{K}^{m,n}$. *Rząd* macierzy A jest to liczba

$$\operatorname{rank} A = \dim(\operatorname{im} A).$$

Zgodnie z tw 5.12, podpunkt (ii) i tw. 5.10, podpunkt (ii), rząd macierzy A jest równy maksymalnej liczbie liniowo niezależnych kolumn tej macierzy. Pokażemy teraz, że rząd macierzy moglibyśmy zdefiniować równoważnie jako maksymalną liczbę jej liniowo niezależnych wierszy (jako elementów przestrzeni $\mathbb{K}^{1,n}$), lub, inaczej mówiąc, jako maksymalną liczbę liniowo niezależnych kolumn macierzy A^T (lub A^H):

Twierdzenie 6.4. *Dana jest macierz $A \in \mathbb{K}^{m,n}$. Wówczas*

$$\operatorname{rank} A = \operatorname{rank} A^H = \operatorname{rank} A^T.$$

W dowodzie wykorzystamy następujący fakt:

Lemat 6.5. *Załóżmy, że $A \in \mathbb{K}^{m,n}$. Wówczas $\operatorname{im} A^H, \ker A \subset \mathbb{K}^n$ oraz*

$$\ker A \cap \operatorname{im} A^H = \{0\}.$$

Dowód lematu. Załóżmy, że $\vec{y} \in \ker A \cap \operatorname{im} A^H$. Oznacza to, że $A\vec{y} = 0$ oraz istnieje $\vec{x} \in \mathbb{K}^m$ taki, że $A^H\vec{x} = \vec{y}$. Zauważmy też, że gdy $\vec{y} = [y_1, \ldots, y_n]^T$, to $\vec{y}^H\vec{y} = \overline{y}_1 y_1 + \ldots \overline{y}_n y_n = |y_1|^2 + \ldots + |y_n|^2$ i

$$\vec{y} = 0 \quad \text{wtedy i tylko wtedy, gdy} \quad \vec{y}^H\vec{y} = 0 \tag{6.1}$$

Mamy

$$\vec{y}^H\vec{y} = (A^H\vec{x})^H A^H\vec{x} = \vec{x}^H A A^H \vec{x} = \vec{x}^H(A\vec{y}) = \vec{x}^H \cdot 0 = 0.$$

Zatem $\vec{y} = 0$, czyli $\ker A \cap \operatorname{im} A^H = \{0\}$. □

Dowód tw. 6.4. Pokażemy najpierw, że

$$\operatorname{rank} A = \dim \operatorname{im} A \geqslant \dim \operatorname{im} A^H = \operatorname{rank} A^H.$$

Ponieważ taka sama zależność jest prawdziwa, gdy macierz A zastąpimy macierzą A^H, więc wynika stąd, że rzędy obu macierzy są równe.

W przestrzeni $\mathbb{K}^n$ konstruujemy bazę $\vec{u}_1, \ldots, \vec{u}_n$ w następujący sposób: niech $s = \dim \ker A$ i $r - s = \dim \operatorname{im} A^H$. Teraz

- wektory $\vec{u}_1, \ldots, \vec{u}_s$ wybieramy jako bazę podprzestrzeni $\ker A$,
- wektory $\vec{u}_{s+1}, \ldots, \vec{u}_r$ wybieramy jako bazę podprzestrzeni $\operatorname{im} A^H$. Zauważmy, że wobec lematu 6.5 $\ker A \cap \operatorname{im} A^H = \{0\}$, więc układ $\vec{u}_1, \ldots, \vec{u}_s, \vec{u}_{s+1}, \ldots, \vec{u}_r$ jest bazą podprzestrzeni $\ker A \oplus \operatorname{im} A^H \subset \mathbb{R}^n$.
- wektory $\vec{u}_{r+1}, \ldots, \vec{u}_n$ to uzupełnienie układu $\vec{u}_1, \ldots, \vec{u}_r$ do bazy całej przestrzeni $\mathbb{K}^n$.

Z powyższej konstrukcji wynika, że układ $\vec{u}_{s+1}, \ldots, \vec{u}_r, \vec{u}_{r+1}, \ldots, \vec{u}_n$ to uzupełnienie bazy podprzestrzeni $\ker A$ do bazy całej przestrzeni $\mathbb{K}^n$. Wobec stw. 6.2, układ $A\vec{u}_{s+1}, \ldots, A\vec{u}_n$ jest bazą podprzestrzeni $\operatorname{im} A$. Pokazaliśmy więc, że

$$\dim \operatorname{im} A = n - s \geqslant r - s = \dim \operatorname{im} A^H.$$

Na zakończenie pozostaje nam zauważyć, że $\dim \operatorname{im} A^H = \dim \operatorname{im} A^T$. Dowód tego faktu pozostawiamy jako proste ćwiczenie. □

Uwaga 6.2. Lemat 6.5 wykorzystuje fakt, że $\mathbb{K}$ jest ciałem liczb zespolonych lub rzeczywistych – w dowodzie tego lematu korzystamy z tego założenia w warunku (6.1). Tym samym twierdzenie 6.4 pokazaliśmy tylko w przypadku $\mathbb{K} = \mathbb{R}$ lub $\mathbb{K} = \mathbb{C}$ (lub $\mathbb{K} = \mathbb{Q}$). Należy jednak zaznaczyć, że twierdzenie to jest prawdziwe dla macierzy $\mathbb{K}^n$, gdzie $\mathbb{K}$ jest dowolnym ciałem. W ogólnym przypadku dowód jest trochę bardziej skomplikowany.

Zadanie 6.1. Załóżmy, że $A \in \mathbb{K}^{m,n}$. Pokaż, że

$$\mathbb{K}^n = \ker A \oplus \operatorname{im} A^H$$
$$\mathbb{K}^m = \ker A^H \oplus \operatorname{im} A.$$

6.3 Charakteryzacje macierzy nieosobliwych

Twierdzenie 6.6. *Niech $A \in \mathbb{K}^{n,n}$. Następujące warunki są równoważne:*

(i) *Macierz A jest nieosobliwa.*

(ii) $\operatorname{im} A = \mathbb{K}^n$.

(iii) $\operatorname{rank} A = n$.

(iv) $\ker A = \{0\}$.

(v) *Kolumny macierzy A są liniowo niezależne*

(vi) *Kolumny macierzy A rozpinają przestrzeń $\mathbb{K}^n$.*

(vii) *Macierz A^T jest nieosobliwa.*

(viii) *Wiersze macierzy A są liniowo niezależne.*

(ix) *Wiersze macierzy A rozpinają przestrzeń $\mathbb{K}^{1,n}$.*

(x) *Dla każdego wektora $\vec{b} \in \mathbb{K}^n$ układ równań $A\vec{x} = \vec{b}$ ma jednoznaczne rozwiązanie.*

Dowód. (i) $\Rightarrow$ (ii): Zapiszmy macierz A^{-1} w postaci blokowo - kolumnowej

$$A^{-1} = [\vec{b}_1, \dots, \vec{b}_n], \qquad \vec{b}_j \in \mathbb{K}^n.$$

Wówczas

$$[\vec{e}_1, \dots, \vec{e}_n] = I_n = AA^{-1} = A[\vec{b}_1, \dots, \vec{b}_n] = [A\vec{b}_1, \dots, A\vec{b}_n].$$

Zatem $A\vec{b}_j = \vec{e}_j$ dla $j = 1, 2, \dots, n$, co oznacza, że

$$\operatorname{im} A \supset \operatorname{span}(A\vec{b}_1, \dots, A\vec{b}_n) = \operatorname{span}(\vec{e}_1, \dots, \vec{e}_n) = \mathbb{K}^n,$$

a więc $\operatorname{im} A = \mathbb{K}^n$.

(ii) $\Rightarrow$ (iii): skoro $\operatorname{im} A = \mathbb{K}^n$ to $\operatorname{rank} A = \dim \operatorname{im} A = \dim \mathbb{K}^n = n$.

(iii) $\Rightarrow$ (iv): z tw. 6.3 wynika, że $\dim\ker A = n - \operatorname{rank} A = 0$, co oznacza, że $\ker A = \{0\}$.

(iv) $\Rightarrow$ (v): Niech $A = [\vec{a}_1, \ldots, \vec{a}_n]$, gdzie $\vec{a}_j \in \mathbb{K}^n$ i załóżmy, że $\gamma_1\vec{a}_1 + \ldots + \gamma_n\vec{a}_n = 0$ dla pewnych skalarów $\gamma_1, \ldots, \gamma_n$. Niech $\vec{x} = [\gamma_1, \ldots, \gamma_n]^T$. Wówczas $A\vec{x} = 0$, czyli $\vec{x} \in \ker A$. Zatem $\vec{x} = 0$ i $\gamma_1 = \ldots = \gamma_n = 0$. Kolumny macierzy A są więc liniowo niezależne.

(v) $\Rightarrow$ (vi): skoro kolumny macierzy A są liniowo niezależne jako wektory z $\mathbb{K}^n$ i jest ich $n = \dim\mathbb{K}^n$, to stanowią one bazę przestrzeni $\mathbb{K}^n$ i w szczególności ją rozpinają.

(vi) $\Rightarrow$ (i): Niech $A = [\vec{a}_1, \ldots, \vec{a}_n]$, gdzie $\vec{a}_j \in \mathbb{K}^n$. Dla $j = 1, 2, \ldots, n$ istnieją skalary $b_{i,j}$, $i = 1, 2, \ldots, n$, takie, że dla wektora $\vec{b}_j = [b_{1,j}, \ldots, b_{n,j}]^T$

$$\vec{e}_j = b_{1,j}\vec{a}_1 + \ldots + b_{n,j}\vec{a}_n = A\vec{b}_j.$$

Zatem $I_n = [A\vec{b}_1, \ldots, A\vec{b}_n] = A[\vec{b}_1, \ldots, \vec{b}_n]$ i mamy $A^{-1} = [\vec{b}_1, \ldots, \vec{b}_n]$.

(i) $\Leftrightarrow$ (vii): Macierz odwrotna do A^T to $(A^{-1})^T$. Jeżeli $B = (A^T)^{-1}$, to $A^{-1} = B^T$.

(vii) $\Rightarrow$ (viii): stosujemy (i) $\Rightarrow$ (v) do macierzy A^T

(viii) $\Rightarrow$ (ix): stosujemy (v) $\Rightarrow$ (vi) do macierzy A^T

(ix) $\Rightarrow$ (vii): stosujemy (vi) $\Rightarrow$ (i) do macierzy A^T.

□

Stwierdzenie 6.7. *Załóżmy, że $M \in \mathbb{K}^{m,n}$, $A \in \mathbb{K}^{m,m}$, $B \in \mathbb{K}^{n,n}$ oraz macierze A i B są nieosobliwe. Wówczas*

(i) $\ker M = \ker(AM)$,

(ii) $\operatorname{im} M = \operatorname{im}(MB)$.

Dowód. (i): Jeżeli $\vec{x} \in \ker M$, to $M\vec{x} = 0$ i $AM\vec{x} = A0 = 0$. Zatem $\ker M \subset \ker(AM)$. Macierz A jest nieosobliwa, więc $\ker A = \{0\}$. Niech $\vec{x} \in \mathbb{R}^n$ będzie takim wektorem, że $AM\vec{x} = 0$ i oznaczmy $\vec{y} = M\vec{x}$. Wtedy $A\vec{y} = 0$, więc $\vec{y} = 0$, czyli $M\vec{x} = 0$ i $\vec{x} \in \ker M$. Pokazaliśmy inkluzję $\ker(AM) \subset \ker M$.

(ii): Załóżmy, że $\vec{y} \in \operatorname{im}(MB)$. Wtedy $\vec{y} = MB\vec{t}$ dla pewnego wetora $\vec{t} \in \mathbb{K}^n$. Niech $\vec{x} = B\vec{t}$. Wówczas $y = M\vec{x}$, czyli $\vec{y} \in \operatorname{im} M$. Pokazaliśmy, że $\operatorname{im}(MB) \subset \operatorname{im} M$. Załóżmy teraz, że $\vec{y} \in \operatorname{im} M$, czyli istnieje $\vec{x} \in \mathbb{R}^n$ taki, że $\vec{y} = M\vec{x}$. Macierz B jest nieosobliwa, więc $\operatorname{im} B = \mathbb{K}^n$. Istnieje więc $\vec{t} \in \mathbb{K}^n$ taki, że $\vec{x} = B\vec{t}$. Wobec tego $\vec{y} = MB\vec{t} \in \operatorname{im}(MB)$. Zatem także $\operatorname{im} M \subset \operatorname{im}(MB)$. □

Z twierdzenia 6.7 i zadań 4.6, 4.7 otrzymujemy

Wniosek 6.8. *Niech $M \in \mathbb{K}^{m,n}$. Wówczas*

(i) operacje elementarne na wierszach macierzy M nie zmieniają jej jądra,

(ii) operacje elementarne na kolumnach macierzy M nie zmieniają jej obrazu.

Dowód. Wystarczy zauważyć, że operacje na wierszach odpowiadają mnożeniu macierzy M z lewej strony przez pewną macierz nieosobliwą $A \in \mathbb{K}^{m,m}$, natomiast operacje na kolumnach odpowiadają mnożeniu macierzy M z prawej strony przez pewną nieosobliwą macierz $B \in \mathbb{K}^{n,n}$. □

6.4 Ogólna grupa liniowa

Pokażemy teraz, że macierze nieosobliwe z $\mathbb{K}^{n,n}$ tworzą grupę z działaniem mnożenia macierzy.

Stwierdzenie 4.2 mówi, że zbiór macierzy nieosobliwych w $\mathbb{K}^{n,n}$ jest zamknięty ze względu na iloczyn macierzy. Ze stw 4.1 wiemy, że iloczyn macierzy w $\mathbb{K}^{n,n}$ jest działaniem łącznym i macierz identycznościowa I_n jest elementem neutralnym tego działania. Nie udowodniliśmy jednak dotąd, że macierz odwrotna do macierzy nieosobliwej też jest macierzą nieosobliwą. Uczynimy to właśnie teraz.

Stwierdzenie 6.9. *Załóżmy, że $A, B \in \mathbb{K}^{n,n}$, macierz A jest nieosobliwa i $AB = I_n$. Wówczas także macierz B jest nieosobliwa.*

Dowód. Skoro macierz A jest nieosobliwa, to, na mocy tw. 6.6 (iv) $\ker A = \{0\}$. Pokażemy, że także $\ker B = \{0\}$, co jest równoważne z tym, że macierz B jest nieosobliwa. Niech $\vec{y} \in \ker B$. Wówczas

$$\vec{y} = I_n \cdot \vec{y} = (AB) \cdot \vec{y} = A(B \cdot \vec{y}) = A \cdot 0 = 0.$$

Zatem istotnie $\ker B = \{0\}$. □

Powyższe stwierdzenie mówi, że dla każdej macierzy nieosobliwej $A \in \mathbb{K}^{n,n}$ jej macierz odwrotna $B = A^{-1}$ też jest nieosobliwa. Zatem zbiór wszystkich macierzy nieosobliwych $n \times n$ o elementach z ciała $\mathbb{K}$ jest zamknięty ze względu na iloczyn macierzy i z I_n jako elementem neutralnym spełnia wszystkie aksjomaty z definicji grupy (1.1). Grupa macierzy odwracalnych $n \times n$ o elementach z ciała $\mathbb{K}$ jest zapisywana jako $GL(n, \mathbb{K})$; nazywana jest *pełną grupą liniową* lub *ogólną grupą liniową*.

Uwaga 6.3. Teraz ze stw. 1.1 wynika, że jeśli $A \in \mathbb{K}^{n,n}$ jest nieosobliwa, $B \in \mathbb{K}^{n,n}$, to $AB = I_n$ wtedy i tylko wtedy, gdy $BA = I_n$.

Rozdział 7

Układy równań liniowych

Układ m równań liniowych o n niewiadomych jest to układ równości

$$\begin{cases} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \dots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \dots & + & a_{2,n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \dots & + & a_{m,n}x_n & = & b_m \end{cases}$$

Liczby $a_{i,j}$ $i = 1, \dots, m$, $j = 1, \dots, n$ nazywamy *współczynnikami* tego układu, $b_1, \dots, b_m$ nazywamy *wyrazami wolnymi*, natomiast $x_1, \dots, x_n$ nazywamy *niewiadomymi*.

Jeżeli $A = [a_{i,j}] \in \mathbb{K}^{m,n}$, $\vec{b} = [b_i] \in \mathbb{K}^m$, $\vec{x} = [x_j] \in \mathbb{K}^n$, to powyższy układ równań możemy zapisać w postaci macierzowej

$$A\vec{x} = \vec{b}.$$

7.1 Istnienie i jednoznaczność rozwiązań

Definicja 7.1. Powiemy, że układ równań $A\vec{x} = \vec{b}$ jest

(i) *niesprzeczny*, jeżeli istnieje wektor $\vec{x} \in \mathbb{K}^n$ taki, że $A\vec{x} = \vec{b}$;

(ii) *sprzeczny*, jeżeli dla każdego wektora $\vec{x} \in \mathbb{K}^n$ zachodzi $A\vec{x} \neq \vec{b}$;

(iii) *oznaczony*, jeżeli istnieje dokładnie jeden wektor $\vec{x} \in \mathbb{K}^n$ taki, że $A\vec{x} = \vec{b}$.

Przykład 7.1. Rozważmy układ równań

$$\begin{cases} x_1 & + & x_2 & + & x_3 & = & 3 \\ 2x_1 & - & x_2 & + & 4x_3 & = & 5 \\ 3x_1 & & & + & 5x_3 & = & 8 \end{cases}$$

Każda z równości jest spełniona dla $x_1 = x_2 = x_3 = 1$, więc układ ten jest niesprzeczny. Innym rozwiązaniem jest $x_1 = -4$, $x_2 = 3$, $x_3 = 4$, nie jest to więc układ oznaczony.

Przykład 7.2. Rozważmy następującą modyfikację powyższego układu

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_1 - x_2 + 4x_3 = 1 \\ 3x_1 + 5x_3 = 1 \end{cases}$$

Z ostatniego równania możemy wyznaczyć

$$x_3 = \frac{1}{5} - \frac{3}{5}x_1,$$

a następnie z pierwszego równania

$$x_2 = 1 - x_1 - x_3 = \frac{4}{5} - \frac{2}{5}x_1.$$

Wstawiając do drugiego równania otrzymujemy

$$2x_1 - \frac{4}{5} + \frac{2}{5}x_1 + \frac{4}{5} - \frac{12}{5}x_1 = 1,$$

czyli $0 = 1$. Zatem ten układ równań jest sprzeczny.

Twierdzenie 7.1 (Twierdzenie Kroneckera-Capellego)**.** *Układ równań $A\vec{x} = \vec{b}$ jest niesprzeczny wtedy i tylko wtedy, gdy*

$$\operatorname{rank} A = \operatorname{rank}[A \mid \vec{b}].$$

Dowód. Niech $A = [\vec{a}_1, \ldots, \vec{a}_n]$, gdzie $\vec{a}_j \in \mathbb{K}^n$. Zauważmy, że $\operatorname{im} A \subset \operatorname{im}[A \mid \vec{b}]$, więc zawsze $\operatorname{rank} A \leqslant \operatorname{rank}[A \mid \vec{b}]$.

Załóżmy, że $\operatorname{rank} A = \operatorname{rank}[A \mid \vec{b}]$. Wówczas $\operatorname{im} A = \operatorname{im}[A \mid \vec{b}]$, skąd wynika, że $\vec{b} \in \operatorname{im} A = \operatorname{span}(\vec{a}_1, \ldots, \vec{a}_n)$. Istnieją więc $x_1, \ldots, x_n \in \mathbb{K}$ takie, że

$$\vec{b} = x_1\vec{a}_1 + \ldots + x_n\vec{a}_n,$$

co można zapisać jako $A\vec{x} = \vec{b}$.

Jeżeli natomiast istnieje $\vec{x} \in \mathbb{K}^n$ taki, że $A\vec{x} = \vec{b}$, to

$$\vec{b} = x_1\vec{a}_1 + \ldots + x_n\vec{a}_n,$$

czyli $\vec{b} \in \operatorname{span}(\vec{a}_1, \ldots, \vec{a}_n) = \operatorname{im} A$. Tym samym $\operatorname{im} A = \operatorname{im}[A \mid \vec{b}]$ i $\operatorname{rank} A = \operatorname{rank}[A \mid \vec{b}]$. □

Uwaga 7.1. Twierdzenie Kroneckera-Capellego można równoważnie sformułować w następujący sposób: układ $A\vec{x} = \vec{b}$ ma rozwiązanie wtedy i tylko wtedy, gdy $\vec{b} \in \operatorname{im} A$.

Zajmiemy się teraz jednoznacznością rozwiązań układów równań liniowych. (Czyli zbadamy, kiedy dany układ równań jest oznaczony) Załóżmy, że układ równań $A\vec{x} = \vec{b}$ jest niesprzeczny i posiada dwa różne rozwiązania $\vec{x}$ oraz $\vec{u}$. Wówczas

$$0 = \vec{b} - \vec{b} = A\vec{x} - A\vec{u} = A(\vec{x} - \vec{u}).$$

Zatem $\vec{x} - \vec{u} \in \ker A$.

Z drugiej strony, jeżeli $\vec{x}$ jest rozwiązaniem danego układu, natomiast $\vec{t} \in \ker A$, to

$$A(\vec{x} + \vec{t}) = A\vec{x} + A\vec{t} = \vec{b} + 0 = \vec{b},$$

więc wektor $\vec{x} + \vec{t}$ też jest rozwiązaniem układu $A\vec{x} = \vec{b}$.

Pokazaliśmy

Stwierdzenie 7.2. *Niesprzeczny układ równań jest oznaczony wtedy i tylko wtedy, gdy* $\ker A = \{0\}$.

Natomiast dla dowolnego niesprzecznego układu równań prawdziwe jest:

Stwierdzenie 7.3. *Jeżeli wektor $\vec{x}$ jest rozwiązaniem układu równań $A\vec{x} = \vec{b}$, to każde inne rozwiązanie tego układu jest postaci*

$$\vec{x} + \vec{t},$$

gdzie $\vec{t} \in \ker A$.

Przykład 7.3. Wróćmy do układu równań z przykładu 7.1:

$$\begin{cases} x_1 + x_2 + x_3 = 3 \\ 2x_1 - x_2 + 4x_3 = 5 \\ 3x_1 + 5x_3 = 8 \end{cases}$$

lub $A\vec{x} = \vec{b}$ dla

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & 4 \\ 3 & 0 & 5 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} 3 \\ 5 \\ 8 \end{bmatrix}.$$

Znamy 2 rozwiązania tego układu: $\vec{x} = [1, 1, 1]^T$ i $\vec{y} = [-4, 3, 4]^T$. Wobec tego

$$\vec{x} - \vec{y} = \begin{bmatrix} 5 \\ -2 \\ -3 \end{bmatrix} \in \ker A.$$

Ponadto, łatwo sprawdzić, że rząd macierzy A jest większy niż 1, a skoro $\dim(\ker A) \geqslant 1$, to $\operatorname{rank} A = 2$, więc $\dim(\ker A) = 1$ i

$$\ker A = \operatorname{span}\left(\begin{bmatrix} 5 \\ -2 \\ -3 \end{bmatrix}\right).$$

Wobec tego każde rozwiązanie tego układu równań jest postaci

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \alpha \begin{bmatrix} 5 \\ -2 \\ -3 \end{bmatrix}.$$

7.2 Wyznaczanie rozwiązań

Stwierdzenie 7.4. *Dla $A \in \mathbb{K}^{m,n}$, $\vec{b} \in \mathbb{K}^m$ rozważamy układ równań $A\vec{x} = \vec{b}$. Macierz $C \in \mathbb{K}^{m,m}$ jest nieosobliwa. Wówczas układy równań*

$$A\vec{x} = \vec{b} \qquad \textit{oraz} \qquad CA\vec{x} = C\vec{b}$$

mają takie same zbiory rozwiązań.

Dowód. Macierz C jest nieosobliwa, więc $C\vec{y} = 0$ wtedy i tylko wtedy, gdy $\vec{y} = 0$. Zatem

$$\begin{aligned} & 0 = A\vec{x} - \vec{b} \\ \iff \quad & 0 = C(A\vec{x} - \vec{b}) \\ \iff \quad & 0 = CA\vec{x} - C\vec{b} \end{aligned}$$

□

Ze stw. 7.4 wynika, że wykonując operacje elementarne na wierszach macierzy $[A \,|\, \vec{b}]$ otrzymujemy układ mający takie same rozwiązania co układ wyjściowy.

Omówimy teraz procedurę zwaną *eliminacją Gaussa*, za pomocą której można wyznaczyć zbiór rozwiązań dowolnego układu równań liniowych. W przypadku układów sprzecznych przyjmujemy, że zbiór rozwiązań jest zbiorem pustym.

Naszym celem jest sprowadzenie macierzy $[A \,|\, \vec{b}]$ do postaci

$$\left[\begin{array}{c:c|c} R & T & \vec{p} \\ \hdashline 0 & 0 & \vec{q} \end{array}\right], \tag{7.1}$$

gdzie dla $r = \operatorname{rank} A$

- macierz $R \in \mathbb{K}^{r,r}$ jest trójkątna górna i nieosobliwa,
- $T \in \mathbb{K}^{r,n-r}$,
- $\vec{p} \in \mathbb{K}^r$, $\vec{q} \in \mathbb{R}^{m-r}$

Zadanie 7.1. Pokaż, że macierz kwadratowa trójkątna górna jest nieosobliwa wtedy i tylko wtedy, gdy wszystkie wyrazy na jej przekątnej są różne od 0.

Aby doprowadzić macierz $[A \mid \vec{b}]$ do postaci (7.1) będziemy stosować następujące przekształcenia:

(i) zamiana miejscami wierszy macierzy

(ii) dodanie do pewnego wiersza innego pomnożonego przez skalar

(iii) zamiana miejscami kolumn macierzy A.

Na mocy stw. 7.4, operacje typów (i) i (ii) nie zmieniają zbioru rozwiązań układu. Operacje typu (iii) odpowiadają za zmianę kolejności niewiadomych.

Cały proces dzielimy na $r = \operatorname{rank} A$ kroków, przy czym k-ty krok polega na zerowaniu wyrazów pod przekątną w k-tej kolumnie przekształcanej macierzy.

Załóżmy, że po k krokach otrzymaliśmy macierz

$$\left[\begin{array}{c:c|c} R_k & T_k & \vec{p}_k \\ \hdashline 0 & V_k & \vec{q}_k \end{array}\right], \tag{7.2}$$

gdzie $R_k \in \mathbb{K}^{k,k}$ jest macierzą nieosobliwą, trójkątną górną. Wówczas:

- Jeżeli $V_k = 0$ to $r = k$ i zakończyliśmy przekształcenia.
- W przeciwnym przypadku macierz V_k zawiera niezerowy element, który możemy przenieść, przestawiając wiersze $k+1$-szy i l-ty ($l > k+1$) i / lub kolumny $k+1$-szą i j-tą ($j > k+1$), tak, aby znalazł się on na przecięciu $k+1$-go wiersza z $k+1$-szą kolumną.
- Następnie odejmujemy wielokrotności $k+1$-szego wiersza od wierszy $k+2, \dots, m$ tak, aby w $k+1$-szej kolumnie poniżej przekątnej uzyskać zera. Uzyskujemy w ten sposób macierz postaci (7.2) dla $k+1$.

Gdy w wyniku przekształceń otrzymamy macierz w postaci (7.1), możemy przystąpić do wyznaczania rozwiązań. Niech $\vec{y}$ oznacza wektor niewiadomych (z ewentualnym uwzględnieniem zmiany ich kolejności):

$$\vec{y} = \begin{bmatrix} \vec{z} \\ \vec{t} \end{bmatrix}, \qquad \vec{z} \in \mathbb{K}^k, \quad \vec{t} \in \mathbb{K}^{n-k}.$$

Wówczas

- Jeżeli $\vec{q} \neq 0$ to układ jest sprzeczny.
- W przeciwnym wypadku układ równań możemy zapisać w postaci

  $$R\vec{z} + T\vec{t} = \vec{p}.$$

  Kładąc $\vec{t} = 0$ otrzymujemy *rozwiązanie szczególne* układu:

  $$\vec{y_0} = \begin{bmatrix} \vec{z_0} \\ 0 \end{bmatrix},$$

  gdzie $\vec{z_0}$ to jedyne rozwiązanie układu $R\vec{z} = \vec{p}$.
- Następnie, jeżeli $r < n$, wyznaczamy bazę jądra macierzy $[R|T]$. (Po uwzględnieniu przestawień kolumn jest to także jądro macierzy A.) W tym celu rozwiązujemy $n-k$ układów równań

  $$R\vec{u}_j + T\vec{e}_j = 0 \quad \Longleftrightarrow \quad R\vec{u}_j = -T\vec{e}_j,$$

  gdzie $\vec{e}_1, \ldots, \vec{e}_{n-k}$ to baza standardowa w $\mathbb{K}^{n-k}$. Zwróćmy uwagę, że $T\vec{e}_j$ to po prostu j-ta kolumna macierzy T. Powyższe układy też są oznaczone. Wektory

  $$\vec{v}_j = \begin{bmatrix} \vec{u}_j \\ \vec{e}_j \end{bmatrix}, \qquad j = 1, \ldots, n-k$$

  tworzą bazę przestrzeni $\ker[R|T]$.
- Ogólne rozwiązanie układu równań $[R|T]\vec{y} = \vec{p}$ jest postaci

  $$\vec{y} = \vec{y_0} + \alpha_1\vec{v}_1 + \ldots + \alpha_{n-k}\vec{v}_{n-k}.$$

  Aby otrzymać rozwiązania układu $A\vec{x} = b$ należy przestawić współrzędne wektora $\vec{y}$, tak, aby otrzymać wyjściową kolejność niewiadomych.

Rozdział 8

Wyznaczniki

8.1 Permutacje

Niech S_n oznacza grupę permutacji zbioru n-elementowego $\{1, 2, \ldots, n\}$. Każdą funkcję p ze zbioru S_n można zapisać jako różnowartościowy ciąg długości n:

$$p = [a_1, \ldots, a_n], \qquad a_k \in \{1, 2, \ldots, n\} \text{ dla } k = 1, 2, \ldots, n.$$

Zapis ten rozumiemy tak, że $p(k) = a_k$ dla $k = 1, 2, \ldots, n$.

Definicja 8.1. *Cyklem* długości m nazwiemy permutację $p \in S_n$ taką, że istnieje m różnych liczb $b_1, \ldots, b_m$ takich, że

$$p(b_1) = b_2, \quad p(b_2) = b_3, \quad \ldots, \quad p(b_{m-1}) = p(b_m), \quad p(b_m) = b_1$$

oraz $p(k) = k$ dla $k \in \{1, 2, \ldots, n\} \backslash \{b_1, \ldots, b_m\}$. Cykl będziemy zapisywać jako $(b_1, b_2, \ldots, b_m)$. Liczbę m nazywamy *długością* cyklu.

Powiemy, że dwa cykle $(b_1, \ldots, b_m)$ i $(c_1, \ldots, c_l)$ są *rozłączne*, jeżeli zbiory $\{b_1, \ldots, b_m\}$ i $\{c_1, \ldots, c_l\}$ są rozłączne.

Transpozycją nazwiemy permutację $\tau \in S_n$, która jest cyklem długości 2, czyli $\tau = (a, b) = (b, a)$.

Stwierdzenie 8.1. *Załóżmy, że $p \in S_n$. Wówczas*

(i) permutację p można zapisać jako złożenie pewnej liczby parami rozłącznych cykli długości co najmniej 2. Z dokładnością do kolejności cykli taki rozkład jest jednoznaczny.

(ii) permutację p można zapisać jako złożenie pewnej liczby transpozycji.

Dowód. (i): pierwszy cykl otrzymujemy biorąc $b_1 = 1$, $b_2 = p(b_1)$, itd. aż do momentu, gdy $p(b_m) = 1$ – dostajemy cykl $(1, b_2, \ldots, b_{m_1})$ $(m_1 = m)$. Jeżeli $m_1 < n$, wybieramy $b_{m+1} \in \{1, 2, \ldots, n\} \backslash \{b_1, \ldots, b_m\}$ i tak samo konstruujemy drugi cykl $(b_{m+1}, \ldots, b_{m_2})$, itd, aż wyczerpiemy wszystkie elementy zbioru $\{1, 2, \ldots, m\}$.

(ii): wystarczy pokazać, że każdy cykl można zapisać jako złożenie transpozycji:

$$(b_1, b_2, \ldots, b_m) = (b_1, b_m) \circ (b_1, b_{m-1}) \circ \cdots \circ (b_1, b_2).$$

□

Uwaga 8.1. W odróżnieniu od rozkładu na cykle, rozkład permutacji na transpozycje nie jest jednoznaczny. (Dlaczego?)

Stwierdzenie 8.2. *Załóżmy, że permutację $p \in S_n$ zapisano na dwa sposoby jako złożenie transpozycji:*

$$p = \tau_r \circ \ldots \circ \tau_1 = \tau'_s \circ \ldots \circ \tau'_1.$$

Wówczas $r + s$ jest liczbą parzystą.

Dowód. Pokażemy najpierw, że dowolny rozkład permutacji identycznościowej $\mathrm{id} = [1, 2, \ldots, n]$ na transpozycje musi się składać z parzystej liczby takich transpozycji. Załóżmy, że mamy rozkład na transpozycje

$$\mathrm{id} = \tau_r \circ \ldots \circ \tau_2 \circ \tau_1.$$

Patrzymy na pierwsze dwie transpozycje z tego rozkładu, czyli $\tau_2 \circ \tau_1$. Zachodzi jedna z następujących możliwości (a, b, c, d oznaczają różne liczby ze zbioru $\{1, 2, \ldots, n\}$):

- $\tau_2 \circ \tau_1 = (a, b) \circ (a, b)$,
- $\tau_2 \circ \tau_1 = (a, c) \circ (a, b)$,
- $\tau_2 \circ \tau_1 = (b, c) \circ (a, b)$,
- $\tau_2 \circ \tau_1 = (c, d) \circ (a, b)$.

Łatwo sprawdzić, że

- $(a, b) \circ (a, b) = \mathrm{id}$,
- $(a, c) \circ (a, b) = (a, b) \circ (b, c)$,
- $(b, c) \circ (a, b) = (a, c) \circ (b, c)$,

- $(c, d) \circ (a, b) = (a, b) \circ (c, d)$.

Zatem w pierwszym przypadku transpozycje τ_1 i τ_2 możemy po prostu usunąć z naszego rozkładu, skracając go o 2 czynniki. W pozostałych przypadkach możemy złożenie $\tau_2 \circ \tau_1$ zastąpić złożeniem dwóch transpozycji takim, że a nie występuje w pierwszej (czyli prawej) z nich. Wobec tego jesteśmy w stanie przerzucać a do coraz dalszych transpozycji i zawsze w transpozycjach na prawo od tej, do której przerzuciliśmy a, ta liczba nie będzie już występować. Możemy tak robić, dopóki nie trafimy na parę postaci $(a, x) \circ (a, x) = \mathrm{id}$ – wówczas także skracamy łańcuch transpozycji o 2, albo przerzucimy a do ostatniej transpozycji w łańcuchu. Jednak ten drugi przypadek jest niemożliwy, gdyż to oznacza, że a występuje tylko w jednej transpozycji z rozkładu, a wtedy $\mathrm{id}(a) \neq a$.

To oznacza, że łańcuch niezerowej długości zawsze jesteśmy w stanie skrócić o 2. Zatem długość łańcucha musi być liczbą parzystą.

Załóżmy teraz, że mamy permutację p, którą rozłożyliśmy na dwa sposoby na transpozycje:

$$p = \tau_r \circ \ldots \circ \tau_1 = \tau'_s \circ \ldots \circ \tau'_1.$$

Wówczas permutację odwrotną do p możemy zapisać jako

$$p^{-1} = \tau'_1 \circ \ldots \circ \tau'_s$$

i otrzymujemy rozkład permutacji identycznościowej na $r + s$ transpozycji

$$\mathrm{id} = p \circ p^{-1} = \tau_r \circ \ldots \circ \tau_1 \ \circ \ \tau'_1 \circ \ldots \circ \tau'_s.$$

Oznacza to, że $r + s$ jest liczbą parzystą. □

Definicja 8.2. *Znakiem permutacji* $p \in S_n$ nazywamy liczbę $\operatorname{sgn} p = (-1)^r$, gdzie r jest długością dowolnego rozkładu p na transpozycje. Jeżeli $\operatorname{sgn} p = 1$, to mówimy, że p jest permutacją *parzystą*. Jeżeli $\operatorname{sgn} p = -1$, to mówimy, że p jest permutacją *nieparzystą*.

Zgodnie ze stw. 8.2, definicja jest poprawna i $\operatorname{sgn} p$ nie zależy od długości rozkładu p na transpozycje.

Stwierdzenie 8.3. *Jeżeli* $p, q \in S_n$, *to* $\operatorname{sgn}(p \circ q) = \operatorname{sgn} p \cdot \operatorname{sgn} q$ *oraz* $\operatorname{sgn} p^{-1} = \operatorname{sgn} p$.

8.2 Wyznacznik macierzy

Definicja 8.3. Wyznacznikiem macierzy kwadratowej $A \in \mathbb{K}^{n,n}$ nazywamy liczbę

$$\det_n(A) = \sum_{p \in S_n} \operatorname{sgn} p \cdot \prod_{i=1}^{n} a_{i,p(i)}. \tag{8.1}$$

Zauważmy, że w każdym iloczynie $\prod_{i=1}^{n} a_{i,p(i)}$ występuje dokładnie jeden element z każdego wiersza macierzy A i, jednocześnie, dokładnie jeden element z każdej kolumny.

Uwaga 8.2. Jeżeli rozmiar macierzy $A \in \mathbb{K}^{n,n}$ jest znany, to zamiast $\det_n A$ można pisać po prostu $\det A$.

Przykład 8.1. Jeżeli $A = [a] \in \mathbb{K}^{1,1}$, to $\det_1 A = a$.

Przykład 8.2.

$$\det_2 \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Przykład 8.3. Jeżeli macierz $A = [a_{i,j}] \in \mathbb{K}^{n,n}$ jest diagonalna, to jedyna permutacja $p \in S_n$, dla której iloczyn $\prod_{i=1}^{n} a_{i,p(i)}$ nie zawiera elementów zerowych, to permutacja identycznościowa. Zatem

$$\det_n A = a_{11}a_{22}\ldots a_{nn}.$$

W szczególności $\det_n I_n = 1$ oraz $\det_n(-I_n) = (-1)^n$.

Stwierdzenie 8.4. $\det_n A = \det_n(A^T)$.

Dowód. We wzorze definiującym wyznacznik możemy sumować po odwrotnościach wszystkich permutacji:

$$\begin{aligned}
\det_n A &= \sum_{p\in S_n} \operatorname{sgn} p \cdot \prod_{i=1}^{n} a_{i,p(i)} \\
&= \sum_{p\in S_n} \operatorname{sgn} p \cdot \prod_{i=1}^{n} a_{p^{-1}\circ p(i),p(i)} \\
&= \sum_{p\in S_n} \operatorname{sgn} p \cdot \prod_{j=1}^{n} a_{p^{-1}(j),j} && p(i) = j \\
&= \sum_{q\in S_n} \operatorname{sgn} q \cdot \prod_{j=1}^{n} a_{q(j),j} && q = p^{-}1,\ \operatorname{sgn} p = \operatorname{sgn} q \\
&= \det_n(A^T).
\end{aligned}$$

□

Powyższe stwierdzenie oznacza, że wszelkie własności wyznacznika odnoszące się do kolumn macierzy, odnoszą się także do jej wierszy.

Prostym wnioskiem ze wzoru (8.1) jest

Stwierdzenie 8.5. *Jeżeli* $A = [\vec{a}_1, \dots, \vec{a}_n]$ *i* $\vec{a}_j = \beta_1\vec{b}_1 + \dots \beta_r\vec{b}_r$, *dla pewnego* $j \in \{1, \dots, n\}$, *to, dla* $B_k = [\vec{a}_1, \dots, \vec{a}_{j-1}, \vec{b}_k, \vec{a}_{j+1}, \dots, \vec{a}_n]$, $k = 1, 2, \dots, r$*:*

$$\det{}_n A = \sum_{k=1}^{r} \beta_k \det{}_n B_k.$$

Wobec stw. 8.4, analogiczny fakt zachodzi dla wierszy macierzy.

Stwierdzenie 8.6. *Jeżeli w macierzy* A *zamienimy miejscami dwie kolumny (wiersze), to* $\det_n A$ *zmieni znak.*

Dowód. Niech $A = [\vec{a}_1, \dots, \vec{a}_n]$. Zamieniając miejscami kolumny k i l otrzymujemy macierz $\widetilde{A}$. Dla transpozycji $\tau = (k, l)$ mamy $\widetilde{A} = [\vec{a}_{\tau(1)}, \dots, \vec{a}_{\tau(n)}]$ i

$$\begin{aligned}
\det{}_n \widetilde{A} &= \sum_{p \in S_n} \operatorname{sgn} p \cdot \prod_{i=1}^{n} a_{i, p\circ\tau(i)} && q = p \circ \tau \\
&= -\sum_{q \in S_n} \operatorname{sgn} q \cdot \prod_{i=1}^{n} a_{i, q(i)} && \operatorname{sgn}(p \circ \tau) = -\operatorname{sgn} p \\
&= -\det{}_n A.
\end{aligned}$$

Zgodnie ze stw. 8.4, teza zachodzi także dla wierszy. □

Stwierdzenie 8.7. *Niech* $A \in \mathbb{K}^{n,n}$

- *Jeżeli macierz* A *ma dwie identyczne kolumny (wiersze), to* $\det_n A = 0$*,*
- *Jeżeli macierz* A *ma zerową kolumnę (zerowy wiersz), to* $\det_n A = 0$*.*

Dowód. Zamieniając miejscami owe identyczne kolumny nie zmieniamy macierzy, a zmieniamy znak jej wyznacznika, więc w 1. przypadku $\det_n A = 0$. W drugim przypadku $\det_n A = \alpha \det_n A$ dla dowolnego $\alpha \in \mathbb{K}$ (stw. 8.5), więc także $\det_n A = 0$. □

Stwierdzenie 8.8. *Jeżeli*

$$A = [\vec{a}_1, \dots, \vec{a}_n] \quad i \quad B = [\vec{a}_1, \dots, \vec{a}_{j-1}, \vec{a}_j + \alpha\vec{a}_k, \vec{a}_{j+1}, \dots, \vec{a}_n]$$

to $\det_n A = \det_n B$*.*

Dowód. Zgodnie ze stw. 8.5 i wnioskiem 8.7

$$\begin{aligned}
\det{}_n B &= \det{}_n A + \alpha \det{}_n[\dots, \vec{a}_k, \dots, \vec{a}_k, \dots] \\
&= \det{}_n A.
\end{aligned}$$

□

Wobec stw. 8.4, analogiczny fakt zachodzi dla wierszy macierzy.

Zadanie 8.1. Jak operacje elementarne na wierszach / kolumnach macierzy A wpływają na jej wyznacznik?

Dla danej macierzy $A \in \mathbb{K}^{n,n}$ przez $A_{i,j}$ będziemy dalej oznaczać macierz z $\mathbb{K}^{n-1,n-1}$ otrzymaną z A przez usunięcie i-tego wiersza i j-tej kolumny.

Lemat 8.9. *Macierz $A \in \mathbb{K}^{n,n}$ ma postać blokową*

$$A = \begin{bmatrix} B & D \\ 0 & C \end{bmatrix}, \quad B \in \mathbb{K}^{r,r}, C \in \mathbb{K}^{n-r,n-r}, D \in \mathbb{K}^{r,n-r}, 0 \in \mathbb{K}^{n-r,r}.$$

Wówczas $\det_n A = \det_r B \cdot \det_{n-r} C$.

Dowód. Niech $A = [a_{i,j}]$. Stosujemy indukcję po r, przy założeniu, że n jest dowolne. Dla $r = 1$ $B = [a_{1,1}]$, $C = A_{1,1}$ i pierwsza kolumna macierzy A ma postać $[a_{1,1}, 0, \dots, 0]^T$ i

$$\begin{aligned} \det_n A &= \sum_{p \in S_n} \operatorname{sgn} p \cdot \prod_{i=1}^{n} a_{i,p(i)} \\ &= a_{1,1} \sum_{\substack{p \in S_n \\ p(1)=1}} \operatorname{sgn} p \cdot \prod_{i=2}^{n} a_{i,p(i)} \\ &= a_{1,1} \sum_{q \in S_{n-1}} \operatorname{sgn} q \cdot \prod_{i=1}^{n-1} a_{i+1,q(i)+1} \\ &= \det_n B \cdot \det_n C. \end{aligned}$$

Załóżmy teraz, że dowodzony wzór zachodzi dla $r < n$. Pokażemy, że zachodzi dla $r + 1$. Jeżeli w A (i bloku B) pierwsza kolumna jest zerowa, to $\det_n A = \det_r B = 0$ i teza zachodzi.

W przeciwnym wypadku zajmiemy się blokiem $B = [b_{i,j}] = [a_{i,j}]_{i,j=1}^{r} \in \mathbb{K}^{r,r}$. Nie zmieniając wyznacznika macierzy A i B, zawsze możemy zamienić miejscami dwa wiersze w B (i te same dwa wiersze w A), aby $a_{1,1} \neq 0$. Następnie, odejmując pierwszy wiersz od wierszy $2, \dots, r$, możemy wyzerować elementy $a_{i,1}$ $(i = 2, \dots, r)$. Jeżeli $\widetilde{A}$ i $\widetilde{B}$ oznaczają macierze otrzymane z A i B w wyniku tych przekształceń, to

$$\begin{aligned} \det_n A &= \varepsilon \cdot \det_n \widetilde{A} \\ \det_r \widetilde{B} &= \varepsilon \cdot \det_r B \end{aligned}$$

dla pewnego $\varepsilon \in \{-1, 1\}$. Ponadto, z założenia indukcyjnego dla $r = 1$ (czyli tezy pokazanej w 1. kroku indukcji) mamy

$$\det_r \widetilde{B} = \widetilde{a}_{1,1} \cdot \det_{r-1} \widetilde{B}_{1,1} \quad \text{i} \quad \det_n \widetilde{A} = \widetilde{a}_{1,1} \cdot \det_{n-1} \widetilde{A}_{1,1}.$$

Ponadto

$$\widetilde{A}_{1,1} = \begin{bmatrix} \widetilde{B}_{1,1} & \widetilde{D} \\ 0 & C \end{bmatrix}, \qquad \widetilde{D} \in \mathbb{K}^{r-1,n-1}$$

i $\det_{n-1} \widetilde{A}_{1,1} = \det_{r-1} \widetilde{B}_{1,1} \cdot \det_{n-r} C$ z założenia indukcyjnego.

Ostatecznie

$$\begin{aligned} \det_n A &= \varepsilon \cdot \det_n \widetilde{A} \\ &= \varepsilon \cdot \widetilde{a}_{1,1} \cdot \det_{n-1} \widetilde{A}_{1,1} \\ &= \varepsilon \cdot \widetilde{a}_{1,1} \cdot \det_{r-1} \widetilde{B}_{1,1} \cdot \det_{n-r} C \\ &= \varepsilon \cdot \det_r \widetilde{B} \cdot \det_{n-r} C \\ &= \varepsilon \cdot \varepsilon \cdot \det_r B \cdot \det_{n-r} C \\ &= \det_r B \cdot \det_{n-r} C. \end{aligned}$$

□

Wobec stw. 8.4, analogiczny fakt zachodzi dla macierzy postaci

$$\begin{bmatrix} B & 0 \\ D & C \end{bmatrix}.$$

Wniosek 8.10. *Jeżeli* $A = [a_{i,j}] \in \mathbb{K}^{n,n}$ *jest macierzą trójkątną górną (dolną), to* $\det_n A = a_{11} a_{22} \dots a_{nn}$*.*

8.3 Rozwinięcia Laplace'a

Poniższe rekurencyjne wzory na wyznaczniki noszą nazwę *rozwinięć Laplace'a*. Pierwszy wzór przedstawia rozwinięcie wyznacznika macierzy A względem j-tej kolumny, drugi to rozwinięcie względem i-tego wiersza. Wyrażają one wyznacznik stopnia n macierzy A za pomocą wyznaczników stopnia $n-1$ macierzy $A_{i,j}$.

Twierdzenie 8.11 (Rozwinięcia Laplace'a)**.** *Jeżeli* $A = [a_{i,j}] \in \mathbb{K}^{n,n}$*, to*

$$\det_n A = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \cdot \det_{n-1} A_{i,j}, \qquad \textit{dla } j = 1, 2, \dots, n$$

oraz

$$\det_n A = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \cdot \det_{n-1} A_{i,j}, \qquad \textit{dla } i = 1, 2, \ldots, n.$$

Dowód. Wykażemy drugi wzór, czyli rozwinięcie względem i-tego wiersza. Zapiszmy macierz A jako $A = [\vec{a}_1, \ldots, \vec{a}_n]$, gdzie $\vec{a}_j \in \mathbb{K}^n$. Ustalmy i; dla $j = 1, 2, \ldots, n$ niech

$$B_j = [\vec{a}_1, \ldots, \vec{a}_{j-1}, \vec{e}_i, \vec{a}_{j+1}, \ldots, \vec{a}_n]$$

(j-tą kolumnę w A zastępujemy przez e_i). Pokażemy, że

$$\det_n B_j = (-1)^{i+j} \det_{n-1} A_{i,j}. \tag{8.2}$$

Za pomocą $j-1$ przestawień sąsiednich kolumn możemy w macierzy B_j przemieścić j-tą kolumnę w miejsce pierwszej, zachowując kolejność pozostałych kolumn. Podobnie, za pomocą $i-1$ przestawień sąsiednich wierszy możemy przemieścić i-ty wiersz w miejsce pierwszego, nie zmieniając kolejności pozostałych wierszy. Otrzymamy w ten sposób macierz

$$\widetilde{B}_k = \left[\begin{array}{c|c} 1 & w \\ \hline 0 & A_{i,j} \end{array}\right], \qquad w \in \mathbb{K}^{1,r-1}$$

Z lematu 8.9 mamy, że $\det_n \widetilde{B}_j = 1 \cdot \det_{n-1} A_{i,j}$. Jednocześnie

$$\det_n \widetilde{B}_j = (-1)^{i-1+j-1} \det_n B_j = (-1)^{i+j} \det_n B_j.$$

Udowodniliśmy tym samym wzór (8.2).

Zauważmy teraz, że, na mocy stw. 8.5,

$$\det_n A = \sum_{i=1}^{n} a_{i,j} \det_n B_j = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det_{n-1} A_{i,j}.$$

□

8.4 Twierdzenie Cauchy'ego

Udowodnimy teraz wzór na wyznacznik iloczynu macierzy:

Twierdzenie 8.12 (Twierdzenie Cauchy'ego)**.** *Jeżeli* $A, B \in \mathbb{K}^{n,n}$*, to*

$$\det_n(AB) = \det_n A \cdot \det_n B.$$

W dowodzie wykorzystamy następujący fakt:

Lemat 8.13. *Jeżeli* $C \in \mathbb{K}^{2n,2n}$ *i*

$$D = \begin{bmatrix} I_n & B \\ 0 & I_n \end{bmatrix}, \qquad B \in \mathbb{K}^{n,n},$$

to $\det_{2n} C = \det_{2n}(CD)$*.*

Dowód. Niech $B = [b_{i,j}]_{i,j=1}^{n}$. Macierz D możemy uzyskać z macierzy I_{2n} za pomocą n^2 operacji elementarnych na kolumnach:

$$k_{n+j} \mapsto k_{n+j} + b_{i,j} k_i, \qquad i, j = 1, \ldots, n$$

czyli $D = E_1 \cdot \ldots \cdot E_{n^2}$, gdzie $E_k \in \mathbb{K}^{2n,2n}$ to macierze powyższych operacji elementarnych. Zatem $CD = CE_1 \ldots E_{n^2}$, czyli macierz CD otrzymujemy z C za pomocą tych samych operacji elementarnych. Operacje te nie zmieniają wyznacznika macierzy (stw 8.8), więc $\det_{2n} C = \det_{2n}(CD)$. □

Dowód twierdzenia Cauchy'ego. Rozważmy macierze $C, D \in \mathbb{K}^{2n,2n}$:

$$C = \begin{bmatrix} A & 0 \\ -I_n & B \end{bmatrix}, \qquad D = \begin{bmatrix} I_n & B \\ 0 & I_n \end{bmatrix}.$$

Z lematu 8.13 mamy, że $\det_{2n} C = \det_{2n}(CD)$. Z lematu 8.9 wynika, że $\det_{2n} C = \det_n A \cdot \det_n B$. Ponadto

$$CD = \begin{bmatrix} A & AB \\ -I_n & 0 \end{bmatrix}$$

i za pomocą n przestawień kolumn możemy macierz CD sprowadzić do macierzy

$$\begin{bmatrix} AB & A \\ 0 & -I_n \end{bmatrix}.$$

Zatem $(-1)^n \cdot \det_{2n}(CD) = \det_n(AB) \cdot \det_n(-I_n) = \det_n(AB) \cdot (-1)^n$, czyli $\det_n(AB) = \det_{2n}(CD) = \det_{2n} C = \det_n A \cdot \det_n B$. □

Wniosek 8.14. *Jeżeli macierz* $A \in \mathbb{K}^{n,n}$ *jest nieosobliwa, to*

$$\det_n A \neq 0$$

oraz

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

8.5 Macierz dopełnień algebraicznych

Definicja 8.4. Niech $A \in \mathbb{K}^{n,n}$. Macierz dopełnień algebraicznych macierzy A jest to macierz $C_A = [c_{i,j}] \in \mathbb{K}^{n,n}$, gdzie

$$c_{i,j} = (-1)^{i+j} \cdot \det_{n-1} A_{i,j}.$$

Obliczymy teraz iloczyn $M = A \cdot (C_A)^T$. Element w k-tym wierszu i l-tej kolumnie to

$$m_{k,l} = \sum_{j=1}^{n} a_{k,j} c_{l,j} = \sum_{j=1}^{n} a_{k,j} (-1)^{l+j} \cdot \det_{n-1} A_{l,j}$$

Zatem $m_{k,k} = \det_n A$. Gdy $k \neq l$, to $m_{k,l}$ jest wyznacznikiem macierzy powstałej z A przez zastąpienie l-tego wiersza jej k-tym wierszem. Jest to więc macierz o wyznaczniku zerowym, czyli $m_{k,l} = 0$ w tym przypadku. Zatem

$$A \cdot (C_A)^T = \det_n A \cdot I_n.$$

Udowodniliśmy także

Stwierdzenie 8.15. *Jeżeli macierz $A \in \mathbb{K}^{n,n}$ jest nieosobliwa, to*

$$A^{-1} = \frac{1}{\det_n A} \cdot (C_A)^T = \frac{1}{\det_n A} \left[(-1)^{i+j} \cdot \det_{n-1} A_{i,j} \right]_{i,j=1}^{n}.$$

Otrzymaliśmy więc jawny wzór na macierz odwrotną do danej macierzy nieosobliwej. Nie trzeba chyba nikogo przekonywać, że, gdy chcemy policzyć macierz odwrotną, to powyższy wzór nie jest zbyt poręczny.

8.6 Wzory Cramera

Za pomocą wyznaczników można także wyrazić rozwiązanie oznaczonego układu równań liniowych z macierzą kwadratową.

Twierdzenie 8.16 (Wzory Cramera)**.** *Załóżmy że macierz kwadratowa $A = [\vec{a}_1, \ldots, \vec{a}_n] \in \mathbb{K}^{n,n}$ jest nieosobliwa, $\vec{b} \in \mathbb{K}^n$ i $\vec{x} = [x_1, \ldots, x_n]^T \in \mathbb{K}^n$ jest rozwiązaniem układu równań $A\vec{x} = \vec{b}$. Wówczas*

$$x_j = \frac{\det_n [\vec{a}_1, \ldots, \vec{a}_{j-1}, \vec{b}, \vec{a}_{j+1}, \ldots, \vec{a}_n]}{\det_n A}.$$

Dowód. Jedyne rozwiązanie układu można zapisać, korzystając ze stw. 8.15, jako

$$\vec{x} = A^{-1}\vec{b} = \frac{1}{\det_n A} \cdot (C_A)^T \vec{b}.$$

Zatem

$$\begin{aligned} x_j &= \frac{1}{\det_n A} \cdot \sum_{i=1}^{n} (-1)^{i+j} \det_{n-1} A_{i,j} b_i \\ &= \frac{\det_n [\vec{a}_1, \ldots, \vec{a}_{j-1}, \vec{b}, \vec{a}_{j+1}, \ldots, \vec{a}_n]}{\det_n A}. \end{aligned}$$

□

Uwaga 8.3. Wykorzystanie wzorów Cramera do znalezienia rozwiązania oznaczonego, kwadratowego układu równań może być, ze względu na ich pozorną prostotę, bardzo kuszące. Należy jednak zdawać sobie sprawę, że dla większych układów równań liczba operacji arytmetycznych, które trzeba wykonać w ogólnym przypadku, jest ogromna. Obliczenie jednego wyznacznika z rozwinięcia Laplace'a wymaga wykonania $O(n \cdot n!)$ działań. Z drugiej strony, eliminację Gaussa można przeprowadzić wykonując $O(n^3)$ operacji arytmetycznych.

Dodatkowo, obliczenie wyznacznika z definicji lub rozwinięcia Laplace'a na komputerze, w arytmetyce zmiennopozycyjnej, może być obarczone ogromnym błędem (wynikającym nie tylko z zaokrągleń danych i niedokładności działań). Znacznie lepszym i szybszym sposobem obliczenia wyznacznika jest sprowadzenie macierzy do postaci trójkątnej za pomocą eliminacji Gaussa.

Rozdział 9

Przestrzenie z iloczynem skalarnym

W całym rozdziale $\mathbb{K}$ oznacza ciało liczb rzeczywistych lub zespolonych.

Aby w przestrzeni liniowej można było rozważać pojęcia geometryczne, takie jak odległość między wektorami lub kąt jaki tworzą między sobą dwa wektory, konieczne jest wprowadzenie na niej dodatkowej struktury za pomocą iloczynu skalarnego.

9.1 Iloczyny skalarne

Niech X będzie przestrzenią liniową nad ciałem $\mathbb{K}$.

Definicja 9.1. *Iloczyn skalarny* na X jest to funkcja $\phi : X \times X \to \mathbb{K}$, którą będziemy zapisywać jako $\langle x, y\rangle = \phi(x, y)$, o następujących własnościach:

(i) dla dowolnych $x, y_1, y_2 \in X$ i $\alpha_1, \alpha_2 \in \mathbb{K}$:

$$\langle x, \alpha_1 y_1 + \alpha_2 y_2\rangle = \alpha_1\langle x, y_1\rangle + \alpha_2\langle x, y_2\rangle,$$

(ii) dla dowolnych $x, y \in X$

$$\langle x, y\rangle = \overline{\langle y, x\rangle},$$

(iii) dla dowolnego $x \in X\backslash\{0\}$

$$\langle x, x\rangle > 0.$$

Uwaga 9.1. W szczególności, w przypadku zespolonym ($\mathbb{K} = \mathbb{C}$) warunek (iii) oznacza, że iloczyn skalarny niezerowego wektora $x \in X$ z samym sobą jest liczbą rzeczywistą dodatnią.

Przykład 9.1 (Standardowy iloczyn skalarny w $\mathbb{K}^n$)**.** Dla $\vec{x}, \vec{y} \in \mathbb{K}^n$ (gdzie $\mathbb{K} = \mathbb{R}$ lub $\mathbb{K} = \mathbb{C}$) określamy

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^H \vec{y} = \overline{x}_1 y_1 + \ldots + \overline{x}_n y_n$$

(jeżeli $\mathbb{K} = \mathbb{R}$ to oczywiście $\langle \vec{x}, \vec{y} \rangle = \vec{x}^T \vec{y}$).

Sprawdzimy, że istotnie jest to iloczyn skalarny: dla $\vec{x}, \vec{y}, \vec{z} \in \mathbb{K}^n$, $\alpha, \beta \in \mathbb{K}$:

$$\begin{aligned} \langle \vec{x}, \alpha\vec{y} + \beta\vec{z} \rangle &= \vec{x}^H(\alpha\vec{y} + \beta\vec{z}) = \alpha\vec{x}^H\vec{y} + \beta\vec{x}^H\vec{z} \\ &= \alpha\langle \vec{x}, \vec{y} \rangle + \beta\langle \vec{x}, \vec{z} \rangle \end{aligned}$$

oraz

$$\begin{aligned} \langle \vec{x}, \vec{y} \rangle &= \vec{x}^H\vec{y} = \overline{(\vec{y}^H\vec{x})} \\ &= \overline{\langle \vec{y}, \vec{x} \rangle}. \end{aligned}$$

Ponadto, gdy $\vec{x} \neq 0$,

$$\langle \vec{x}, \vec{x} \rangle = \vec{x}^H\vec{x} = \sum_{j=1}^{n} \overline{x_j} x_j = \sum_{j=1}^{n} |x_j|^2 > 0,$$

gdzie $\vec{x} = [x_1, \ldots, x_n]^T$.

Przykład 9.2. Dla $x, y \in \mathbb{R}^2$ niech

$$\langle \vec{x}, \vec{y} \rangle = 2x_1y_1 - x_2y_1 - x_1y_2 + 5x_2y_2.$$

Jest to iloczyn skalarny: warunki (i) i (ii) są oczywiste, warunek (iii) zachodzi, gdyż

$$\begin{aligned} \langle x, x \rangle &= 2x_1^2 - 2x_1x_2 + 5x_2^2 \\ &= (x_1^2 + 2x_1x_2 + x_2^2) + (x_1^2 - 4x_1x_2 + 4x_2^2) \\ &= (x_1 + x_2)^2 + (x_1 - 2x_2)^2 > 0 \qquad \text{gdy } \vec{x} \neq 0 \end{aligned}$$

Zadanie 9.1. Niech $x_0, \ldots x_n \in \mathbb{K}$ będą różnymi punktami. Dla $p, q \in \mathbb{K}[x]_n$ określamy

$$\langle p, q \rangle = \sum_{k=0}^{n} \overline{p(x_k)} q(x_k).$$

Pokaż, że jest to iloczyn skalarny na przestrzeni $\mathbb{K}[x]_n$. Czy będzie to prawda, jeżeli weźmiemy pewne $k+1$ różnych punktów $x_0, x_1 \ldots, x_k$ dla $k \neq n$?

Stwierdzenie 9.1 (Własności iloczynu skalarnego)**.** *Załóżmy, że* $\langle\cdot,\cdot\rangle$ *jest iloczynem skalarnym na przestrzeni liniowej* X *nad ciałem* $\mathbb{K}$*. Wówczas:*

(a) dla dowolnych $x_1, x_2, y \in X$*,* $\alpha_1, \alpha_2 \in \mathbb{K}$

$$\langle \alpha_1 x_1 + \alpha_2 x_2, y\rangle = \overline{\alpha_1}\langle x_1, y\rangle + \overline{\alpha_2}\langle x_2, y\rangle,$$

(b) dla dowolnego $x \in X$

$$\langle x, 0\rangle = \langle 0, x\rangle = 0,$$

(c) dla dowolnego $x \in X$

$$\langle x, x\rangle = 0 \text{ wtedy i tylko wtedy, gdy } x = 0,$$

(d) jeżeli $x \in X$ *i dla każdego* $y \in X$ *zachodzi* $\langle x, y\rangle = 0$*, to* $x = 0$*.*

Dowód. Podpunkt (a) wynika z warunków (i) i (ii) definicji, podpunkt (b) wynika z ciągu równości:

$$\langle x, 0\rangle = \langle x, x - x\rangle = \langle x, x\rangle - \langle x, x\rangle = 0$$

oraz (ii), podpunkt (c) wynika z (b) i (iii), podpunkt (d) wynika z (c). □

Definicja 9.2. Przestrzeń liniową skończenie wymiarową X nad ciałem $\mathbb{R}$ z iloczynem skalarnym $\langle\cdot,\cdot\rangle$ nazywamy *przestrzenią euklidesową*.

Przestrzeń liniową X nad ciałem $\mathbb{C}$ z iloczynem skalarnym $\langle\cdot,\cdot\rangle$ nazywamy *przestrzenią unitarną*.

Przestrzeń liniową X z iloczynem skalarnym $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{K}$ zapisujemy jako parę $(X, \langle\cdot,\cdot\rangle)$.

9.2 Normy

Określenie iloczynu skalarnego na przestrzeni liniowej umożliwia zdefiniowanie pojęcia długości wektora, czyli jego *normy*:

Definicja 9.3. *Norma* na przestrzeni z iloczynem skalarnym $(X, \langle\cdot,\cdot\rangle)$, pochodząca od iloczynu skalarnego, jest to funkcja $\|\cdot\| : X \to [0, +\infty)$ zadana wzorem

$$\|x\| = \sqrt{\langle x, x\rangle}, \qquad x \in X.$$

Przykład 9.3. Na przestrzeni $\mathbb{K}^n$ ze standardowym iloczynem skalarnym $\langle\vec{x}, \vec{y}\rangle = \vec{x}^H\vec{y}$ mamy normę

$$\|x\| = \sqrt{|x_1|^2 + \ldots + |x_n|^2}.$$

Twierdzenie 9.2 (Nierówność Schwarza)**.** *W przestrzeni z iloczynem skalarnym $(X, \langle\cdot,\cdot\rangle)$, dla dowolnych $x, y \in X$ zachodzi nierówność*

$$|\langle x, y\rangle| \leqslant \|x\| \cdot \|y\| .$$

Dowód. Przeprowadzimy dowód w przypadku, gdy $\mathbb{K} = \mathbb{C}$, czyli gdy X jest przestrzenią unitarną.

Niech $x, y \in X$. Weźmy dowolny skalar $t \in \mathbb{K}$ i zapiszmy t i $\langle x, y\rangle$ w postaci trygonometrycznej:

$$t = |t|e^{i\alpha}, \qquad \langle x, y\rangle = |\langle x, y\rangle|e^{i\beta}.$$

Wówczas

$$\begin{aligned} 0 \leqslant \|x+ty\|^2 &= \langle x+ty, x+ty\rangle \\ &= \langle x, x\rangle + \langle ty, x\rangle + \langle x, ty\rangle + \langle ty, ty\rangle = \\ &= \|x\|^2 + \bar{t}\overline{\langle x, y\rangle} + t\langle x, y\rangle + |t|^2\|y\|^2 \\ &= \|x\|^2 + |t|e^{-i\alpha}|\langle x, y\rangle|e^{-i\beta} + |t|e^{i\alpha}|\langle x, y\rangle|e^{i\beta} + |t|^2\|y\|^2 \\ &= \|x\|^2 + |t||\langle x, y\rangle|(e^{-i(\alpha+\beta)} + e^{i(\alpha+\beta)}) + |t|^2\|y\|^2 . \end{aligned}$$

Liczba t jest dowolna, więc jej argument α możemy wybrać tak jak nam pasuje. Dla $\alpha = -\beta$ otrzymujemy

$$0 \leqslant \|x\|^2 + 2|t||\langle x, y\rangle| + |t|^2\|y\|^2 ,$$

natomiast dla $\alpha = \pi - \beta$:

$$0 \leqslant \|x\|^2 - 2|t||\langle x, y\rangle| + |t|^2\|y\|^2 .$$

Zatem, dla dowolnego $s \in \mathbb{R}$:

$$0 \leqslant \|x\|^2 + 2s|\langle x, y\rangle| + s^2\|y\|^2 .$$

Jest to trójmian kwadratowy zmiennej s, przyjmujący wyłącznie wartości nieujemne, więc jego wyróżnik jest niedodatni:

$$0 \geqslant (2|\langle x, y\rangle|)^2 - 4\|x\|^2\|y\|^2 ,$$

lub równoważnie

$$|\langle x, y\rangle| \leqslant \|x\| \cdot \|y\| .$$

□

Zadanie 9.2. Udowodnij nierówność Schwarza w przypadku przestrzeni euklidesowej.

Zadanie 9.3. Pokaż, że nierówność Schwarza staje się równością wtedy i tylko wtedy, gdy wektory x, y są liniowo zależne.

Stwierdzenie 9.3 (Własności normy)**.** *Niech* $(X, \langle\cdot,\cdot\rangle)$ *będzie przestrzenią z iloczynem skalarnym, a* $\|\cdot\|$ *to norma na* X *pochodząca od iloczynu skalarnego. Wówczas*

(a) dla każdego $x \in X$, $\|x\| = 0$ *wtedy i tylko wtedy, gdy* $x = 0$,

(b) dla dowolnych $x \in X$ *i* $\alpha \in \mathbb{K}$

$$\|\alpha x\| = |\alpha| \cdot \|x\|,$$

(c) dla dowolnych $x, y \in X$

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

Dowód. Podpunkty (a) i (b) wynikają z własności iloczynu skalarnego. Podpunkt (c) jest wnioskiem z nierówności Schwarza:

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle \\
&= \langle x,x\rangle + \overline{\langle x,y\rangle} + \langle x,y\rangle + \langle y,y\rangle \\
&= \|x\|^2 + \|y\|^2 + 2\mathrm{Re}\langle x,y\rangle \\
&\leqslant \|x\|^2 + \|y\|^2 + 2|\langle x,y\rangle| && \text{gdyż } \mathrm{Re} z \leqslant |z| \\
&\leqslant \|x\|^2 + \|y\|^2 + 2\,\|x\| \cdot \|y\| && \text{z nier. Schwarza.} \\
&= (\|x\| + \|y\|)^2.
\end{aligned}$$

□

Uwaga 9.2. Norma $\|\cdot\|$ na przestrzeni euklidesowej X zadaje również funkcję odległości wektorów, czyli *metrykę* $d : X \times X \to [0, \infty)$:

$$d(x,y) = \|x - y\|.$$

Z własności normy wynika, że

- $d(x,y) = d(y,x)$
- $d(x,y) = 0$ wtedy i tylko wtedy, gdy $x = y$,

- $d(x, y) \leqslant d(x, z) + d(z, y)$ dla dowolnych $x, y, z \in X$ (nierówność trójkąta)

Metryka (a więc i norma) określona na przestrzeni liniowej pozwala nam mówić o *odległości* między wektorami. Dzięki temu można zdefiniować takie pojęcia, jak granica ciągu wektorów, zbiory otwarte i domknięte. Można wtedy także mówić o *błędzie przybliżenia* jednego wektora innym jako o odległości między tymi wektorami.

Definicja 9.4. Załóżmy, że $\|\cdot\|$ jest normą na przestrzeni z iloczynem skalarnym $(X, \langle\cdot,\cdot\rangle)$.

- *sferą jednostkową* w $(X, \langle\cdot,\cdot\rangle)$ nazywamy zbiór
$$S = S_X = \{x \in X : \|x\| = 1\},$$
- *kulą jednostkową* w $(X, \langle\cdot,\cdot\rangle)$ nazywamy zbiór
$$B = B_X = \{x \in X : \|x\| \leqslant 1\}.$$

Zadanie 9.4. Pokaż, że kula jednostkowa jest *zbiorem wypukłym*, tzn. jeżeli $x, y \in B_X$, $0 \leqslant \alpha, \beta \leqslant 1$ i $\alpha + \beta = 1$, to
$$\alpha x + \beta y \in B_X.$$

Uwaga 9.3. Można także wprowadzić pojęcie normy na przestrzeni liniowej bez odwoływania się do pojęcia iloczynu skalarnego:
Niech X będzie przestrzenią liniową nad ciałem $\mathbb{K}$. Funkcję $f : X \to [0, +\infty)$ nazwiemy *normą* na X, jeżeli $\|x\| = f(x)$ spełnia własności (a), (b) i (c) w stw. 9.3.

Zadanie 9.5. Pokaż, że następujące funkcje są normami (zgodnie z definicją w uwadze 9.3) na przestrzeniach $\mathbb{R}^n$ i $\mathbb{C}^n$:

- $\|\vec{x}\|_1 = |x_1| + \ldots + |x_n|$,
- $\|\vec{x}\|_\infty = \max(|x_1|, \ldots, |x_n|)$.

Zadanie 9.6. Niech $\|\cdot\|$ oznacza normę na $\mathbb{R}^n$ pochodzącą od iloczynu skalarnego $\langle\vec{x}, \vec{y}\rangle = \vec{x}^T\vec{y}$. Pokaż, że
$$\|A\|_M = \sup\{\|A\vec{x}\| : \|\vec{x}\| = 1\}, \qquad \text{dla } A \in \mathbb{R}^{n,n}$$
jest normą (zgodnie z definicją w uwadze 9.3) na przestrzeni $\mathbb{R}^{n,n}$.

9.3 Ortogonalność

W celu uproszczenia zapisu, wprowadźmy oznaczenie (tzw. symbol Kroneckera):

$$\delta_{j,k} = \begin{cases} 1 & \text{gdy } j = k \\ 0 & \text{gdy } j \neq k \end{cases}.$$

Niech $(X, \langle\cdot,\cdot\rangle)$ będzie przestrzenią z iloczynem skalarnym.

Definicja 9.5. Powiemy, że wektory $x, y \in X$ są *ortogonalne* (*prostopadłe*), jeżeli $\langle x, y\rangle = 0$. Piszemy wówczas $x \perp y$.

Przykład 9.4. W przestrzeni $\mathbb{R}^3$ ze standardowym iloczynem skalarnym $\langle x, y\rangle = \vec{x}^T\vec{y}$ dany jest wektor $\vec{x} = [a, b, c]^T$. Jeżeli $\vec{y} = [y_1, y_2, y_3]^T \perp \vec{x}$, to

$$0 = \vec{x}^T\vec{y} = ay_1 + by_2 + cy_3,$$

czyli $\vec{y}$ należy do podprzestrzeni w $\mathbb{R}^3$ opisanej równaniem $ay_1+by_2+cy_3 = 0$.

Definicja 9.6. Powiemy, że układ wektorów $x_1, \ldots, x_k \in X$ jest *układem ortogonalnym*, jeżeli $x_j \neq 0$ dla każdego j oraz $x_i \perp x_j$ dla $i \neq j$.

Ortogonalny układ wektorów $x_1, \ldots, x_n \in X$ nazywamy *układem ortonormalnym*, jeżeli dodatkowo $\|x_j\| = 1$ dla każdego j. Inaczej mówiąc $\langle x_i, x_j\rangle = \delta_{i,j}$.

Przykład 9.5. Wektory $\vec{e}_1, \ldots, \vec{e}_n \in \mathbb{K}^n$ tworzą układ ortonormalny w $\mathbb{K}^n$ ze standardowym iloczynem skalarnym.

Przykład 9.6. W $\mathbb{R}^4$ ze standardowym iloczynem skalarnym $\langle\vec{x}, \vec{y}\rangle = \vec{x}^T\vec{y}$ wektory

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \quad \vec{v}_4 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$$

tworzą układ ortogonalny.

Stwierdzenie 9.4. *Każdy ortogonalny układ wektorów w przestrzeni z iloczynem skalarnym $(X, \langle\cdot,\cdot\rangle)$ jest liniowo niezależny.*

Dowód. Załóżmy, że wektory $x_1, \ldots x_k$ tworzą układ ortogonalny i $\alpha_1 x_1 + \ldots + \alpha_k x_k = 0$ dla $\alpha_j \in \mathbb{K}$. Dla $j = 1, \ldots, k$ możemy pomnożyć stronami ostatnią równość skalarnie przez x_j. Ponieważ $\langle x, 0\rangle = 0$, otrzymujemy

$$0 = \langle x_j, \alpha_1 x_1 + \ldots \alpha_k x_k\rangle = \sum_{i=0}^{k} \alpha_i \langle x_j, x_i\rangle = \alpha_j \langle x_j, x_j\rangle.$$

więc $\alpha_j = 0$. Ostatnia równość wynika z tego, że $\langle x_j, x_i\rangle = 0$ dla $i \neq j$. □

Definicja 9.7. Układ ortogonalny (ortonormalny) w przestrzeni z iloczynem skalarnym $(X, \langle\cdot,\cdot\rangle)$, który jest bazą tej przestrzeni, nazywamy *bazą ortogonalną* (*ortonormalną*).

Przykład 9.7. Układ $\vec{e}_1, \ldots, \vec{e}_n$ jest bazą ortonormalną w $\mathbb{K}^n$ ze standardowym iloczynem skalarnym.

Przykład 9.8. Wektory $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$ z przykładu 9.6 tworzą bazę ortogonalną przestrzeni $\mathbb{R}^4$.

Stwierdzenie 9.5. *Załóżmy, że układ $y_1, \ldots, y_n$ jest bazą ortogonalną przestrzeni z iloczynem skalarnym $(X, \langle\cdot,\cdot\rangle)$. Wówczas każdy wektor $x \in X$ możemy zapisać jako*

$$x = \sum_{k=1}^{n} \frac{\langle y_k, x\rangle}{\langle y_k, y_k\rangle} y_k.$$

Jeżeli natomiast baza $y_1, \ldots, y_n$ jest ortonormalna, to

$$x = \sum_{k=1}^{n} \langle y_k, x\rangle y_k.$$

Dowód. Istnieją skalary $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ takie, że

$$x = \sum_{k=1}^{n} \alpha_k y_k.$$

Biorąc iloczyn skalarny obu stron z wektorem y_j dostajemy

$$\begin{aligned}\langle y_j, x\rangle &= \langle y_j, \sum_{k=1}^{n} \alpha_k y_k\rangle \\ &= \sum_{k=1}^{n} \alpha_k \langle y_j, y_k\rangle \\ &= \alpha_j \langle y_j, y_j\rangle.\end{aligned}$$

W powyższych przekształceniach korzystamy z własności iloczynu skalarnego i tego, że $\langle y_j, y_i\rangle = 0$ dla $i \neq j$. □

Udowodnimy teraz następujący ważny fakt dotyczący baz ortonormalnych:

Twierdzenie 9.6 (Tożsamość Parsevala)**.** *Jeżeli układ* $y_1, \ldots, y_n$ *jest bazą ortonormalną w przestrzeni z iloczynem skalarnym* $(X, \langle\cdot,\cdot\rangle)$ *oraz* $u, v \in X$*, to*

$$\langle u, v\rangle = \sum_{j=1}^{n} \overline{\langle y_j, u\rangle}\langle y_j, v\rangle.$$

W szczególności

$$\|u\|^2 = \sum_{j=1}^{n} |\langle y_j, u\rangle|^2.$$

Dowód. Korzystając z ortonormalności bazy $y_1, \ldots, y_n$ i własności iloczynu skalarnego, obliczamy

$$\begin{aligned}\langle u, v\rangle &= \left\langle \sum_{j=1}^{n}\langle y_j, u\rangle y_j, \sum_{k=1}^{n}\langle y_k, v\rangle y_k \right\rangle \\ &= \sum_{j=1}^{n}\sum_{k=1}^{n} \overline{\langle y_j, u\rangle}\langle y_k, v\rangle\langle y_j, y_k\rangle \\ &= \sum_{j=1}^{n} \overline{\langle y_j, u\rangle}\langle y_j, v\rangle.\end{aligned}$$

□

Uwaga 9.4. Twierdzenie 9.6 mówi, że, jeśli w przestrzeni $(X, \langle\cdot,\cdot\rangle)$ wybierzemy bazę ortonormalną $y_1, \ldots, y_n$, to iloczyn skalarny $\langle\cdot,\cdot\rangle$ wygląda tak samo, jak standardowy iloczyn skalarny na $\mathbb{K}^n$ (zob. przykład 9.1). Jeżeli bowiem $u, v \in X$ i

$$u = \alpha_1 y_1 + \ldots + \alpha_n y_n, \qquad v = \beta_1 y_1 + \ldots + \beta_n y_n,$$

to ze stw. 9.5

$$\alpha_k = \langle y_k, u\rangle, \qquad \beta_k = \langle y_k, v\rangle$$

i

$$\langle u, v\rangle = \sum_{k=1}^{n} \overline{\langle y_k, u\rangle}\langle y_k, v\rangle = \sum_{k=1}^{n} \overline{\alpha}_k \beta_k.$$

Zatem, jeżeli $\vec{\alpha} = [\alpha_1, \ldots, \alpha_n]^T$, $\vec{\beta} = [\beta_1, \ldots, \beta_n]^T$ to wektory współrzędnych wektorów u, v w bazie $y_1, \ldots, y_n$, to

$$\langle u, v\rangle = \vec{\alpha}^H \vec{\beta}.$$

Inaczej mówiąc, przestrzeń euklidesowa lub unitarna wymiaru n, z ustaloną bazą ortonormalną, ma taką samą strukturę geometryczną jak $\mathbb{R}^n$ lub $\mathbb{C}^n$ ze standardowym iloczynem skalarnym i baza standardową.

9.4 Ortogonalizacja Grama - Schmidta

Niech $x_1, \ldots, x_m$ będzie liniowo niezależnym układem wektorów w przestrzeni z iloczynem skalarnym $(X, \langle\cdot,\cdot\rangle)$. Chcemy znaleźć układ ortogonalny $(y_1, \ldots, y_m)$ lub ortonormalny $(\tilde{y}_1, \ldots, \tilde{y}_m)$ taki, że dla $k = 1, \ldots, m$:

$$\operatorname{span}(x_1, \ldots, x_k) = \operatorname{span}(y_1, \ldots, y_k) = \operatorname{span}(\tilde{y}_1, \ldots, \tilde{y}_k). \tag{9.1}$$

Układ taki można skonstruować rekurencyjnie w następujący sposób:

- $y_1 = x_1$ i $\tilde{y}_1 = \frac{1}{\|y_1\|} y_1$,
- jeżeli wyznaczyliśmy już $y_1, \ldots y_k$ i $\tilde{y}_1, \ldots \tilde{y}_k$ oraz $k < m$, to

$$y_{k+1} = x_{k+1} - \sum_{j=1}^{k} \frac{\langle y_j, x_{k+1}\rangle}{\langle y_j, y_j\rangle} y_j = x_{k+1} - \sum_{j=1}^{k} \langle \tilde{y}_j, x_{k+1}\rangle \tilde{y}_j,$$

  oraz $\tilde{y}_{k+1} = \frac{1}{\|y_{k+1}\|} y_{k+1}$

Twierdzenie 9.7. *Otrzymane w powyżej opisany sposób wektory $y_1, \ldots, y_m$ są niezerowe i tworzą układ ortogonalny taki, że spełnione jest (9.1).*

Dowód. Ponieważ dla $k = 1, \ldots, m$ zachodzi $y_k \in \operatorname{span}(x_1, \ldots, x_k)$, więc układ $y_1, \ldots, y_k, x_{k+1}$ jest liniowo niezależny i

$$y_{k+1} = x_{k+1} - \sum_{j=1}^{k} \alpha_j y_j \neq 0.$$

Ortogonalność wektorów y_j jest równoważna ortogonalności wektorów $\tilde{y}_j$. Ortogonalność układu $\tilde{y}_1, \ldots, \tilde{y}_k$ dowodzimy przez indukcję po k. Gdy $k = 1$ mamy układ składający się z jednego wektora, który w oczywisty sposób jest ortogonalny.

Załóżmy, że wektory $\tilde{y}_1, \ldots \tilde{y}_k$ są ortogonalne. Pokażemy, że wektor y_{k+1} jest ortogonalny do każdego z nich:

$$\begin{aligned}
\langle \tilde{y}_j, y_{k+1}\rangle &= \left\langle \tilde{y}_j, x_{k+1} - \sum_{i=1}^{k} \langle \tilde{y}_i, x_{k+1}\rangle \tilde{y}_i \right\rangle \\
&= \langle \tilde{y}_j, x_{k+1}\rangle - \sum_{i=1}^{k} \langle \tilde{y}_i, x_{k+1}\rangle \langle \tilde{y}_j, \tilde{y}_i\rangle \\
&= \langle \tilde{y}_j, x_{k+1}\rangle - \langle \tilde{y}_j, x_{k+1}\rangle \langle \tilde{y}_j, \tilde{y}_j\rangle \\
&= \langle \tilde{y}_j, x_{k+1}\rangle - \langle \tilde{y}_j, x_{k+1}\rangle \\
&= 0.
\end{aligned}$$

Ponadto, $y_1, \ldots, y_k \in \operatorname{span}(x_1, \ldots, x_k)$ są liniowo niezależne jako układ ortogonalny, więc zachodzi (9.1). □

Przykład 9.9. Ortogonalizacja Grama - Schmidta układu

$$\begin{bmatrix}1\\0\\0\end{bmatrix}, \quad \begin{bmatrix}1\\1\\0\end{bmatrix}, \quad \begin{bmatrix}1\\1\\2\end{bmatrix}$$

w $\mathbb{R}^3$ ze standardowym iloczynem skalarnym daje układ $\vec{e}_1, \vec{e}_2, \vec{e}_3$.

Wniosek 9.8. *W każdej przestrzeni z iloczynem skalarnym $(X, \langle\cdot,\cdot\rangle)$ skończonego wymiaru istnieje baza ortogonalna.*

Dowód. Przeprowadzamy ortogonalizację Grama - Schmidta dowolnej bazy przestrzeni X. □

9.5 Rozkład ortogonalny przestrzeni

Definicja 9.8. Załóżmy, że Z jest podzbiorem przestrzeni z iloczynem skalarnym $(X, \langle\cdot,\cdot\rangle)$. *Dopełnienie ortogonalne* zbioru Z jest to zbiór

$$Z^\perp = \{x \in X : x \perp z \text{ dla każdego } z \in Z\}.$$

Stwierdzenie 9.9. *$Z^\perp$ jest podprzestrzenią liniową w X*

Dowód. Jeżeli $x, y \in Z^\perp$, to $\langle z, x\rangle = \langle z, y\rangle = 0$ dla każdego $z \in Z$. Zatem dla dowolnych skalarów $\alpha, \beta \in \mathbb{K}$ $\langle z, \alpha x + \beta y\rangle = \alpha\langle z, x\rangle + \beta\langle z, y\rangle = 0$, więc $\alpha x + \beta y \in Z^\perp$. □

Z własności iloczynu skalarnego wynika następujący fakt:

Stwierdzenie 9.10. *Jeżeli $(X, \langle\cdot,\cdot\rangle)$ jest przestrzenią z iloczynem skalarnym i $Z = \mathrm{span}(z_1, \ldots, z_k) \subset X$, to $x \in Z^\perp$ wtedy i tylko wtedy, gdy $\langle z_1, x\rangle = \ldots = \langle z_k, x\rangle = 0$.*

Przykład 9.10. W $\mathbb{R}^3$ ze standardowym iloczynem skalarnym niech

$$Z = \{\vec{z} \in \mathbb{R}^3 : z_1 - 2z_2 + 4z_3 = 0\}.$$

Wówczas $\vec{x} = [1, -2, 4]^T \in Z^\perp$. Ponadto, niech $W = \mathrm{span}([1, -2, 4]^T)$. Wtedy $W^\perp = Z$:

$$\begin{aligned} W^\perp &= \{\vec{x} \in X : \langle\vec{w}, \vec{x}\rangle = 0 \text{ dla każdego } w \in W\} \\ &= \{\vec{x} \in X : \alpha[1, -2, 4]\vec{x} = 0 \text{ dla każdego } \alpha \in \mathbb{R}\} \\ &= \{\vec{x} \in X : [1, -2, 4]\vec{x} = 0\} \\ &= \{\vec{x} \in X : x_1 - 2x_2 + 4x_3 = 0\} \\ &= Z. \end{aligned}$$

Mamy także $Z = \operatorname{span}([2,1,0]^T, [-4,0,1]^T)$ i $\vec{x} \in Z^\perp$ wtedy i tylko wtedy, gdy $\vec{x} \perp [2,1,0]^T$ oraz $\vec{x} \perp [-4,0,1]^T$, czyli

$$[2,1,0]\vec{x} = [-4,0,1]\vec{x} = 0,$$

co możemy także zapisać jako

$$Z^\perp = \ker \begin{bmatrix} 2 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} = \operatorname{span}([1,-2,4]^T) = W.$$

Ponadto $\mathbb{R}^3 = Z \oplus W = Z \oplus Z^\perp$. Przedstawiliśmy $\mathbb{R}^3$ jako sumę prostą podprzestrzeni Z i jej dopełnienia ortogonalnego.

Stwierdzenie 9.11. *Załóżmy, że $(X, \langle\cdot,\cdot\rangle)$ jest przestrzenią z iloczynem skalarnym i $Z \subset X$ jest podprzestrzenią liniową. Wówczas*

$$Z \cap Z^\perp = \{0\}.$$

Dowód. Jeżeli $x \in Z^\perp$, to dla każdego $z \in Z$ zachodzi $\langle z, x\rangle = 0$. Zatem, dla $x \in Z \cap Z^\perp$ mamy $\langle x, x\rangle = 0$, więc $x = 0$. □

Twierdzenie 9.12 (Rozkład ortogonalny przestrzeni)**.** *Załóżmy, że $(X, \langle\cdot,\cdot\rangle)$ jest przestrzenią z iloczynem skalarnym wymiaru $n < \infty$ i $Z \subset X$ jest podprzestrzenią liniową. Wówczas*

$$X = Z \oplus Z^\perp.$$

Dowód. Wiemy już, że $Z \cap Z^\perp = \{0\}$, wystarczy więc pokazać, że $X = Z + Z^\perp$.

Niech $z_1, \ldots, z_k$ będzie bazą ortonormalną podprzestrzeni Z i niech $x \in X$ będzie dowolnym wektorem. Określmy

$$z = \sum_{j=1}^{k} \langle z_j, x\rangle z_j, \qquad w = x - z.$$

Jasne jest, że $z \in Z$. Pokażemy, że $w \in Z^\perp$. Dla $i = 1, \ldots, k$ mamy:

$$\begin{aligned} \langle z_i, w\rangle &= \left\langle z_i, x - \sum_{j=1}^{k} \langle z_j, x\rangle z_j \right\rangle \\ &= \langle z_i, x\rangle - \sum_{j=1}^{k} \langle z_j, x\rangle \langle z_i, z_j\rangle \\ &= \langle z_i, x\rangle - \sum_{j=1}^{k} \langle z_j, x\rangle \delta_{i,j} \\ &= \langle z_i, x\rangle - \langle z_i, x\rangle \\ &= 0. \end{aligned}$$

Pokazaliśmy, że wektor w jest ortogonalny do każdego wektora z bazy przestrzeni Z, więc w jest ortogonalny do dowolnego wektora z przestrzeni Z. Wobec tego $w \in Z^\perp$. Ponieważ $x = z + w$, więc $X = Z + Z^\perp$. □

Wniosek 9.13. $\dim Z^\perp = \dim X - \dim Z$.

9.6 Rzut ortogonalny

Definicja 9.9. W przestrzeni z iloczynem skalarnym $(X, \langle\cdot,\cdot\rangle)$ dana jest podprzestrzeń liniowa Z. Wówczas, dla każdego $x \in X$ istnieją jednoznacznie wyznaczone wektory $z \in Z$ oraz $w \in Z^\perp$ takie, że $x = z + w$. Wektor z nazywamy *rzutem ortogonalnym* wektora x na podprzestrzeń Z. Będziemy go oznaczać $P_Z(x)$.

Uwaga 9.5. Niech $x_1, \ldots, x_m$ będzie dowolnym układem liniowo niezależnym w X. Proces ortogonalizacji Grama-Schmidta tego układu możemy opisać za pomocą rzutów ortogonalnych:

Niech $X_k = \operatorname{span}(x_1, \ldots, x_k)$ dla $k = 1, \ldots, m$. Wówczas, jeżeli wyznaczyliśmy już wektory ortogonalne $y_1, \ldots, y_k$, to

$$y_{k+1} = x_{k+1} - P_{X_k}(x_{k+1}).$$

Uwaga 9.6. Aby wyznaczyć rzut ortogonalny wektora $x \in X$ na daną podprzestrzeń $Z \subset X$, wystarczy znaleźć bazę ortonormalną $z_1, \ldots z_k$ podprzestrzeni Z i skorzystać ze wzoru

$$P_Z(x) = \sum_{j=1}^{k} \langle z_j, x\rangle z_j.$$

Stwierdzenie 9.14. *Jeżeli $Z \subset X$ jest podprzestrzenią liniową i $x \in X$, to*

$$x = P_Z(x) + P_{Z^\perp}(x).$$

Dowód. Niech bowiem $z_1, \ldots, z_k$ to baza ortonormalna Z, a $z_{k+1}, \ldots, z_n$ to baza ortonormalna $Z^\perp$. Wtedy $z_1, \ldots, z_n$ jest bazą ortonormalną X i

$$x = \sum_{j=1}^{n} \langle z_j, x\rangle z_j,$$

$$P_Z(x) = \sum_{j=1}^{k} \langle z_j, x\rangle z_j,$$

$$P_{Z^\perp}(x) = \sum_{j=k+1}^{n} \langle z_j, x\rangle z_j.$$

□

Przykład 9.11. W przestrzeni $\mathbb{R}^4$ ze standardowym iloczynem skalarnym rozważamy podprzestrzeń

$$Z = \{\vec{z} \in \mathbb{R}^4 : z_1 + z_2 + z_3 + z_4 = 0\} = \operatorname{span}\left(\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}\right).$$

Wówczas

$$\vec{z}_1 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \quad \vec{z}_2 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \quad \vec{z}_3 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$$

jest bazą ortonormalną podprzestrzeni Z i rzut ortogonalny wektora $\vec{x} = [1, 0, 1, -4]^T$ na Z to

$$\begin{aligned} P_Z(\vec{x}) &= (\vec{z}_1^T\vec{x})\vec{z}_1 + (\vec{z}_2^T\vec{x})\vec{z}_2 + (\vec{z}_3^T\vec{x})\vec{z}_3 \\ &= 2 \cdot \vec{z}_1 - 2 \cdot \vec{z}_2 + 3 \cdot \vec{z}_3 \\ &= \tfrac{1}{2}[3, 1, 3, -7]^T. \end{aligned}$$

Zauważmy jednak, że w tym przypadku łatwiej jest obliczyć $P_{Z^\perp}(\vec{x})$. Mamy bowiem $Z^\perp = \operatorname{span}(\vec{z}_4)$, $\vec{z}_4 = \frac{1}{2}[1, 1, 1, 1]^T$. Potem wystarczy skorzystać ze wzoru $P_Z(x) = x - P_{Z^\perp}(x)$.

Rzut ortogonalny wektora na podprzestrzeń można także scharakteryzować jako element tej podprzestrzeni położony najbliżej tego wektora:

Twierdzenie 9.15. *Rzut ortogonalny P_Z na podprzestrzeń liniową Z w przestrzeni z iloczynem skalarnym $(X, \langle\cdot,\cdot\rangle)$ spełnia*

$$\|x - P_Z(x)\| \leqslant \|x - z\|$$

dla każdego wektora $x \in X$ i $z \in Z$. Ponadto równość zachodzi wtedy i tylko wtedy, gdy $z = P_Z(x)$.

Dowód. Niech $z_1, \ldots, z_k$ będzie bazą ortonormalną Z, którą rozszerzamy wektorami $z_{k+1}, \ldots, z_n$ do bazy ortonormalnej całej przestrzeni X, $x \in X$ oraz $z \in Z$. Wówczas

$$x = \sum_{j=1}^{n} \langle z_j, x\rangle z_j \quad \text{oraz} \quad z = \sum_{j=1}^{k} \alpha_j z_j,$$

dla pewnych $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$. Ponadto, na mocy tożsamości Parsevala

$$\|x - P_Z(x)\|^2 = \sum_{j=k+1}^{n} |\langle z_j, x\rangle|^2$$

oraz

$$\|x - z\|^2 = \sum_{j=1}^{k} |\langle z_j, x\rangle - \alpha_j|^2 + \sum_{j=k+1}^{n} |\langle z_j, x\rangle|^2.$$

Zatem dla każdego $z \in Z$ $\|x - P_Z(x)\| \leqslant \|x - z\|$, a równość zachodzi wtedy i tylko wtedy, gdy $\langle z_j, x\rangle - \alpha_j = 0$ dla $j = 1, \ldots, k$, czyli gdy $z = P_Z(x)$. $\square$

Zadanie 9.7. W przestrzeni $\mathbb{K}^n$ ze standardowym iloczynem skalarnym dana jest podprzestrzeń V. Układ wektorów $\vec{v}_1, \ldots, \vec{v}_k$ jest bazą ortonormalną podprzestrzeni V. Niech

$$A = \vec{v}_1\vec{v}_1^H + \ldots + \vec{v}_k\vec{v}_k^H.$$

(a) Pokaż, że dla każdego wektora $\vec{x} \in \mathbb{K}^n$ $P_V(\vec{x}) = A\vec{x}$.

(b) Jaki jest rząd macierzy A? Wyznacz podprzestrzenie im A oraz ker A.

9.7 Kąt między wektorami

Załóżmy, że $(X, \langle\cdot,\cdot\rangle)$ jest przestrzenią **euklidesową**. Dla pary wektorów $x, y \in X\backslash\{0\}$

$$\frac{\langle x, y\rangle}{\|x\| \cdot \|y\|} \in [-1, 1],$$

co wynika z nierówności Schwarza.

Definicja 9.10. *Kosinus kąta* między niezerowymi wektorami x, y definiujemy jako liczbę

$$\cos(\sphericalangle(x, y)) = \frac{\langle x, y\rangle}{\|x\| \cdot \|y\|}.$$

Kosinus kąta daje nam pewną informację o wzajemnym położeniu wektorów x i y.

Stwierdzenie 9.16. *Niech x, y będą dwoma niezerowymi wektorami w przestrzeni euklidesowej $(X, \langle\cdot,\cdot\rangle)$. Wówczas*

(a) $\frac{x}{\|x\|} = \frac{y}{\|y\|}$ *wtedy i tylko wtedy, gdy* $\cos(\sphericalangle(x, y)) = 1$.

(b) $\frac{x}{\|x\|} = -\frac{y}{\|y\|}$ *wtedy i tylko wtedy, gdy* $\cos(\sphericalangle(x,y)) = -1$.

(c) $x \perp y$ *wtedy i tylko wtedy, gdy* $\cos(\sphericalangle(x,y)) = 0$.

Dowód. Udowodnimy (a): Niech

$$u = \frac{x}{\|x\|}, \qquad v = \frac{y}{\|y\|}.$$

Wówczas $\|u\| = \|v\| = 1$ i $\cos(\sphericalangle(x,y)) = \cos(\sphericalangle(u,v))$. Jeżeli $u = v$, to $\cos(\sphericalangle(u,v)) = 1$.

Załóżmy, że $\cos(\sphericalangle(u,v)) = 1$. Wektor v możemy zapisać jako sumę jego składowej równoległej do u i prostopadłej do u: $v = \alpha u + \beta w$, gdzie $w \perp u$ i $\|w\| = 1$. Wówczas $1 = \|v\|^2 = |\alpha|^2 + |\beta|^2$ oraz

$$1 = \langle u, v\rangle = \langle u, \alpha u + \beta w\rangle = \alpha\langle u, u\rangle + \beta\langle u, w\rangle = \alpha,$$

wobec czego $\beta = 0$. Otrzymaliśmy, że $u = v$.

Drugi podpunkt dowodzi się w identyczny sposób. Trzeci podpunkt jest oczywisty. □

Można także zdefiniować kosinus kąta między wektorem i podprzestrzenią:

Definicja 9.11. *Kosinus kąta* między niezerowym wektorem $x \notin Y^\perp$ i niezerową podprzestrzenią liniową Y definiujemy jako liczbę

$$\cos(\sphericalangle(x, Y)) = \cos(\sphericalangle(x, P_Y(x))),$$

gdzie $P_Y(x)$ to rzut ortogonalny wektora x na podprzestrzeń Y. Jeżeli $x \in Y^\perp$, przyjmujemy, że $\cos(\sphericalangle(x, Y)) = 0$.

Zadanie 9.8. Pokaż, że

$$\cos(\sphericalangle(x, Y)) = \inf\{\cos(\sphericalangle(x, y)) : y \in Y\}.$$

9.8 Macierz Grama układu wektorów

Definicja 9.12. W przestrzeni z iloczynem skalarnym $(X, \langle\cdot,\cdot\rangle)$ dany jest układ wektorów $v_1, \ldots, v_k$. Macierz

$$G = G(v_1, \ldots, v_k) = [\langle v_i, v_j\rangle]_{i,j=1}^k$$

nazywamy *macierzą Grama* układu $v_1, \ldots, v_k$.

Twierdzenie 9.17. *Układ wektorów $v_1, \ldots, v_k$ jest liniowo niezależny wtedy i tylko wtedy, gdy macierz $G(v_1, \ldots, v_k)$ jest nieosobliwa.*

Dowód. Załóżmy, że wektory $v_1, \ldots, v_k$ są liniowo niezależne. Pokażemy, że kolumny macierzy G są liniowo niezależne. Rozważmy kombinację liniową kolumn macierzy G o współczynnikach $\alpha_1, \ldots, \alpha_k$, za pomocą której zapisaliśmy wektor zerowy:

$$\alpha_1 G\vec{e}_1 + \ldots + \alpha_k G\vec{e}_k = 0.$$

Inaczej mówiąc, dla $i = 1, \ldots, k$ mamy równości

$$\begin{aligned} 0 &= \alpha_1\langle v_i, v_1\rangle + \ldots + \alpha_k\langle v_i, v_k\rangle \\ &= \langle v_i, \alpha_1 v_1 + \ldots + \alpha_k v_k\rangle. \end{aligned}$$

Oznacza to, że wektor $w = \alpha_1 v_1 + \ldots + \alpha_k v_k$ jest prostopadły do każdego z wektorów v_i, jest więc prostopadły do przestrzeni $\operatorname{span}(v_1, \ldots, v_k)$. To jest możliwe tylko wtedy, gdy $w = 0$. Z liniowej niezależności wektorów $v_1, \ldots, v_k$ wynika, że $\alpha_1 = \ldots = \alpha_k = 0$.

Załóżmy teraz, że macierz $G = [\langle v_i, v_j\rangle]_{i,j}$ jest nieosobliwa. Przypuśćmy, że $\alpha_1 v_1 + \ldots + \alpha_k v_k = 0$. Zatem, dla $i = 1, \ldots, k$

$$\begin{aligned} 0 &= \langle v_i, \alpha_1 v_1 + \ldots + \alpha_k v_k\rangle \\ &= \alpha_1\langle v_i, v_1\rangle + \ldots + \alpha_k\langle v_i, v_k\rangle. \end{aligned}$$

co oznacza, że $\alpha_1 G\vec{e}_1 + \ldots + \alpha_k G\vec{e}_k = 0$. Ponieważ kolumny macierzy G, czyli wektory $G\vec{e}_j$, są liniowo niezależne, więc $\alpha_1 = \ldots = \alpha_k = 0$. □

Zadanie 9.9. Jak wygląda macierz Grama ortogonalnego (ortonormalnego) układu wektorów?

Zadanie 9.10. Pokaż, że macierz Grama jest hermitowska.

Stwierdzenie 9.18. *Załóżmy, że w przestrzeni z iloczynem skalarnym $(X, \langle\cdot,\cdot\rangle)$ dana jest baza ortonormalna $u_1, \ldots, u_n$, oraz wektory $x_1, \ldots, x_k \in X$, przy czym*

$$x_j = \sum_{i=1}^{n} a_{i,j} u_i, \quad a_{i,j} \in \mathbb{K}, \qquad j = 1, \ldots, k$$

Niech $A = [a_{i,j}] \in \mathbb{K}^{n,k}$. Wówczas

$$G(x_1, \ldots, x_k) = A^H A.$$

Dowód. Kolumny macierzy A to wektory współczynników wektorów $x_1, \ldots, x_k$ w bazie $u_1, \ldots, u_n$. Teza wynika z definicji macierzy Grama i tw. 9.6 (tożsamości Parsevala). □

Uwaga 9.7. W szczególności w $\mathbb{R}^n$ lub $\mathbb{C}^n$ ze standardowym iloczynem skalarnym macierz Grama układu wektorów $\vec{a}_1, \ldots, \vec{a}_k$ możemy wyznaczyć jako:

$$G(\vec{a}_1, \ldots, \vec{a}_n) = [\vec{a}_1, \ldots, \vec{a}_n]^H \cdot [\vec{a}_1, \ldots, \vec{a}_n].$$

Definicja 9.13. Wyznacznik macierzy Grama układu wektorów $v_1, \ldots, v_k$ w przestrzeni z z iloczynem skalarnym $(X, \langle\cdot,\cdot\rangle)$ nazywamy *wyznacznikiem Grama* tego układu.

Wobec tw. 9.17, wyznacznik Grama jest różny od zera wtedy i tylko wtedy, gdy układ $v_1, \ldots, v_k$ jest liniowo niezależny.

Stwierdzenie 9.19. *W przestrzeni z iloczynem skalarnym* $(X, \langle\cdot,\cdot\rangle)$ *dane są wektory* $x_1, \ldots, x_k$*,* $k > 1$ *i wektor* z *jest rzutem ortogonalnym wektora* x_k *na podprzestrzeń* $\operatorname{span}(x_1, \ldots, x_{k-1})^\perp$*. Wówczas*

$$\det_k G(x_1, \ldots, x_k) = \det_{k-1} G(x_1, \ldots, x_{k-1}) \cdot \|z\|^2 .$$

Dowód. Niech $V = \operatorname{span}(x_1, \ldots, x_{k-1})$. Wtedy $x_k = y + z$ i $y = P_V(x_k) \in V$, natomiast $z = P_{V^\perp}(x_k) \in V^\perp$. Mamy

$$G(x_1, \ldots, x_k) = \begin{bmatrix} \langle x_1, x_1\rangle & \ldots & \langle x_1, x_{k-1}\rangle & \langle x_1, y+z\rangle \\ \cdots & \cdots & \cdots & \cdots \\ \langle x_{k-1}, x_1\rangle & \ldots & \langle x_{k-1}, x_{k-1}\rangle & \langle x_{k-1}, y+z\rangle \\ \langle y+z, x_1\rangle & \ldots & \langle y+z, x_{k-1}\rangle & \langle y+z, y+z\rangle \end{bmatrix}.$$

Ostatnią kolumnę możemy zapisać jako sumę

$$\begin{bmatrix} \langle x_1, y+z\rangle \\ \vdots \\ \langle x_{k-1}, y+z\rangle \\ \langle y+z, y+z\rangle \end{bmatrix} = \begin{bmatrix} \langle x_1, y\rangle \\ \vdots \\ \langle x_{k-1}, y\rangle \\ \langle y+z, y\rangle \end{bmatrix} + \begin{bmatrix} \langle x_1, z\rangle \\ \vdots \\ \langle x_{k-1}, z\rangle \\ \langle y+z, z\rangle \end{bmatrix},$$

przy czym, wobec $y \in V = \operatorname{span}(x_1, \ldots, x_{k-1})$, pierwszy składnik jest kombinacją liniową pierwszych $k-1$ kolumn macierzy $G(x_1, \ldots, x_k)$. Ponieważ $x_j \perp z$ dla $j = 1, \ldots, k-1$ i $y \perp z$, drugi składnik jest równy

$$\begin{bmatrix} \langle x_1, z\rangle \\ \vdots \\ \langle x_{k-1}, z\rangle \\ \langle y+z, z\rangle \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \langle z, z\rangle \end{bmatrix}.$$

Zatem

$$\begin{aligned}\det{}_k G(x_1,\dots,x_k) &= \det{}_k \begin{bmatrix} \langle x_1,x_1\rangle & \dots & \langle x_1,x_{k-1}\rangle & 0 \\ \multicolumn{4}{c}{\dotfill} \\ \langle x_{k-1},x_1\rangle & \dots & \langle x_{k-1},x_{k-1}\rangle & 0 \\ \langle y+z,x_1\rangle & \dots & \langle y+z,x_{k-1}\rangle & \langle z,z\rangle \end{bmatrix} \\ &= \det{}_{k-1} G(x_1,\dots,x_{k-1})\cdot \|z\|^2 .\end{aligned}$$

Ostatnia równość wynika z rozwinięcia Laplace'a względem ostatniej kolumny macierzy. □

Stwierdzenie 9.20. *Wyznaczniki Grama liniowo niezależnego układu wektorów $x_1,\dots,x_k$ i układu $z_1,\dots,z_k$ otrzymanego w wyniku ortogonalizacji Grama-Schmidta układu $x_1,\dots,x_k$ są równe.*

Dowód. Niech $V_j = \operatorname{span} v_1,\dots,v_j$. Każdy z wektorów x_j możemy zapisać jako

$$x_j = y_j + z_j, \qquad y_j \in V_{j-1}, z_j \in V_{j-1}^{\perp}.$$

Wobec stw. 9.19

$$\begin{aligned}\det{}_k G(x_1,\dots,x_k) &= \det{}_{k-1} G(x_1,\dots,x_{k-1})\cdot \|z_k\|^2 \\ &= \det{}_{k-2} G(x_1,\dots,x_{k-2})\cdot \|z_{k-1}\|^2\cdot \|z_k\|^2 \\ &\dots\dots\dots \\ &= \det{}_1 G(x_1)\cdot \|z_2\|^2\cdot\ldots\cdot \|z_{k-1}\|^2\cdot\|z_k\|^2 \\ &= \|z_1\|^2\cdot \|z_2\|^2\cdot\ldots\cdot \|z_{k-1}\|^2\cdot\|z_k\|^2 \\ &= \det{}_k G(z_1,\dots,z_k).\end{aligned}$$

□

Zadanie 9.11. Pokaż, że wyznacznik Grama jest zawsze liczbą rzeczywistą nieujemną.

Zadanie 9.12. W $\mathbb{R}^2$ ze standardowym iloczynem skalarnym dane są dwa liniowo niezależne wektory $\vec{u},\vec{v}$. Pokaż, że pole powierzchni równoległoboku wyznaczonego przez te wektory (czyli takiego, którego wierzchołki to 0, $\vec{u}$, $\vec{v}$, $\vec{u}+\vec{v}$) jest równe $\sqrt{\det_2 G(u,v)}$

Zadanie 9.13. W $\mathbb{R}^3$ ze standardowym iloczynem skalarnym dane są trzy liniowo niezależne wektory $\vec{u},\vec{v},\vec{w}$. Pokaż, że objętość równoległościanu wyznaczonego przez te wektory (czyli takiego, którego wierzchołki to 0, $\vec{u}$, $\vec{v}$, $\vec{u}+\vec{v}$, $\vec{w}$, $\vec{u}+\vec{w}$, $\vec{v}+\vec{w}$, $\vec{u}+\vec{v}+\vec{w}$) jest równa $\sqrt{\det_3 G(u,v,w)}$.

Rozdział 10

Przekształcenia liniowe

10.1 Definicja i przykłady przekształceń liniowych

W przykładzie 4.7 mówiliśmy o obrocie płaszczyzny $\mathbb{R}^2$ o zadany kąt skierowany θ skierowany wokół punktu $(0,0)$ jako o przekształceniu płaszczyzny, które można wykonać mnożąc dany wektor przez pewną macierz:

$$f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix}.$$

Zgodnie z własnościami działań na macierzach, jeżeli $\alpha, \beta \in \mathbb{R}$ i $\vec{x}, \vec{y} \in \mathbb{R}^2$, to

$$f(\alpha\vec{x} + \beta\vec{y}) = \alpha f(\vec{x}) + \beta f(\vec{y}),$$

czyli wartość przekształcenia f od kombinacji liniowej wektorów jest kombinacją liniową o tych samych współczynnikach wartości f od tych wektorów. Każde takie odwzorowanie F nazywamy *przekształceniem liniowym.*

Definicja 10.1. Załóżmy, że X i Y są przestrzeniami liniowymi nad ciałem $\mathbb{K}$. Powiemy, że odwzorowanie (funkcja) $f : X \to Y$ jest *przekształceniem liniowym z przestrzeni X w przestrzeń Y*, jeżeli dla dowolnych wektorów $x, y \in X$ i dowolnych skalarów $\alpha, \beta \in \mathbb{K}$

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

Przykład 10.1 (Macierz jako przekształcenie liniowe)**.** Rozważmy funkcję $f : \mathbb{R}^3 \to \mathbb{R}^2$, która wektorowi $[x_1, x_2, x_3]^T$ przyporządkowuje wektor $[x_1 + x_2 + x_3, 2x_1 - 3x_2 + 6x_3]^T$. Możemy napisać

$$f(\vec{x}) = \begin{bmatrix} 1 & 1 & 1 \\ 2 & -3 & 6 \end{bmatrix} \vec{x}.$$

Także w tym przypadku $f(\alpha\vec{x}+\beta\vec{y}) = \alpha f(\vec{x})+\beta f(\vec{y})$, czyli f jest przekształceniem liniowym z $\mathbb{R}^3$ w $\mathbb{R}^2$.

Ogólnie, każda macierz $A \in \mathbb{K}^{m,n}$ zadaje przekształcenie liniowe z $\mathbb{K}^n$ w $\mathbb{K}^m$:

$$f_A(\vec{x}) = A\vec{x}.$$

Przykład 10.2. Rozważmy następującą funkcję z przestrzeni $f : \mathbb{R}[x]_2 \to \mathbb{R}$:

$$f(p) = p(0) - 2p(-1).$$

Przeciwdziedzinę $\mathbb{R}$ utożsamiamy z przestrzenią liniową $\mathbb{R}^1$. Tak zadana funkcja f jest przekształceniem liniowym z $\mathbb{R}[x]_2$ w $\mathbb{R}^1 = \mathbb{R}$.

Przykład 10.3. Funkcja $f : \mathbb{K}^n \to \mathbb{K}[t]_{n-1}$ jest zadana wzorem

$$f(\vec{x}) = x_1 + x_2 t + x_3 t^2 + \ldots + x_n t^{n-1}.$$

To jest przekształcenie liniowe z $\mathbb{K}^n$ w $\mathbb{K}[t]_{n-1}$. Jeżeli $\vec{x}, \vec{y} \in \mathbb{K}^n$, $\alpha, \beta \in \mathbb{K}$, to

$$\begin{aligned} f(\alpha\vec{x} + \beta\vec{y}) &= (\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2)t + \ldots + (\alpha x_n + \beta y_n)t^{n-1} \\ &= \alpha(x_1 + x_2 t + \ldots + x_n t^{n-1}) + \beta(y_1 + y_2 t + \ldots + y_n t^{n-1}) \\ &= \alpha f(\vec{x}) + \beta(f\vec{y}), \end{aligned}$$

więc f jest przekształceniem liniowym z $\mathbb{K}^n$ w $\mathbb{K}[x]_{n-1}$

Przykład 10.4 (Pochodna wielomianu)**.** Wielomianowi $p \in \mathbb{K}[x]$,

$$p(x) = \sum_{k=0}^{n} a_k x^k$$

przyporządkowujemy wielomian $D(p) = p' \in \mathbb{K}[x]$ zadany wzorem

$$p'(x) = \sum_{k=1}^{n} k a_k x^{k-1}.$$

D jest przekształceniem liniowym z przestrzeni $\mathbb{K}[x]$ w $\mathbb{K}[x]$.

Przykład 10.5 (Iloczyn skalarny)**.** Jeżeli $(X, \langle\cdot,\cdot\rangle)$ jest przestrzenią z iloczynem skalarnym, $a \in X$ jest ustalonym wektorem, to funkcja $f_a(x) = \langle a, x\rangle$ jest przekształceniem liniowym z X w przestrzeń liniową $\mathbb{K} = \mathbb{K}^1$. ($\mathbb{K} = \mathbb{R}$ lub $\mathbb{K} = \mathbb{C}$)

Przykład 10.6 (Rzut ortogonalny)**.** Niech $(X, \langle\cdot,\cdot\rangle)$ będzie przestrzenią z iloczynem skalarnym, $V \subset X$ jest podprzestrzenią liniową i $v_1, \ldots, v_k$ jest bazą ortonormalną V. Wówczas rzut ortogonalny

$$P_V(x) = \sum_{k=1}^{n} \langle v_k, x\rangle v_k$$

jest przekształceniem linowym z przestrzeni X w siebie, bowiem dla $x, y \in X$, $\alpha, \beta \in \mathbb{K}$

$$\begin{aligned} P_V(\alpha x + \beta y) &= \sum_{k=1}^{n} \langle v_k, \alpha x + \beta y\rangle v_k \\ &= \sum_{k=1}^{n} (\alpha\langle v_k, x\rangle + \beta\langle v_k, y\rangle) v_k \\ &= \alpha \sum_{k=1}^{n} \langle v_k, x\rangle v_k + \beta \sum_{k=1}^{n} \langle v_k, y\rangle v_k \\ &= \alpha P_V(x) + \beta P_V(y). \end{aligned}$$

10.2 Przestrzenie przekształceń liniowych

Załóżmy, że X i Y są przestrzeniami liniowymi nad ciałem $\mathbb{K}$ i $f : X \to Y$ jest przekształceniem liniowym. Dla skalara $\alpha \in \mathbb{K}$ możemy określić przekształcenie liniowe $\alpha f : X \to Y$ w następujący sposób: dla $x \in X$

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

Podobnie, jeżeli $g : X \to Y$ też jest przekształceniem liniowym, to możemy zdefiniować sumę przekształceń f i g: dla $x \in X$

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

Możemy także rozważać *przekształcenie zerowe* z X w Y:

$$0 : X \to Y \qquad x \mapsto 0 \in Y \quad \text{dla} \quad x \in X.$$

Dla przekształcenia liniowego $f : X \to Y$ możemy określić *przekształcenie przeciwne* $-f = (-1) \cdot f$, czyli dla $x \in X$

$$(-f)(x) = -f(x).$$

Definicja 10.2. Zbiór wszystkich przekształceń liniowych z przestrzeni liniowej X w przestrzeń liniową Y będziemy oznaczać $L(X,Y)$. Wraz z działaniami mnożenia przez skalar i dodawania określonymi powyżej oraz przekształceniem zerowym zbiór ten jest przestrzenią liniową nad ciałem $\mathbb{K}$, którą nazywamy *przestrzenią przekształceń liniowych* z X w Y.

Zadanie 10.1. Sprawdź, że zbiór $L(X,Y)$ spełnia wszystkie warunki z definicji przestrzeni liniowej.

Stwierdzenie 10.1 (Składanie przekształceń liniowych)**.** *Jeżeli $f \in L(X,Y)$ i $g \in L(Y,Z)$, gdzie X, Y, Z są przestrzeniami liniowymi nad ciałem $\mathbb{K}$, to odwzorowanie*

$$h = g \circ f, \qquad h(x) = g(f(x)), \qquad x \in X$$

jest przekształceniem liniowym z X w Z, czyli $h \in L(X,Z)$.

Zadanie 10.2. Udowodnij powyższe stwierdzenie.

Odpowiemy teraz na pytanie, jaki jest wymiar przestrzeni $L(X,Y)$ i jak konstruować bazy tej przestrzeni.

Stwierdzenie 10.2. *Załóżmy, że* $\dim X = n < \infty$, *$x_1, \ldots, x_n$ jest bazą przestrzeni X i $z_1, \ldots, z_n \in Y$. Wówczas istnieje dokładnie jedno przekształcenie liniowe $f \in L(X,Y)$ takie, że*

$$f(x_i) = z_i \quad \textit{dla} \quad i = 1, \ldots, n. \tag{10.1}$$

Dowód. Pokażemy, że z warunku (10.1) wynika, że możemy już jednoznacznie określić przekształcenie liniowe $f : X \to Y$. Dla $x_1, \ldots, x_n$ definiujemy $f(x_i) = z_i$. Dla dowolnego $x \in X$ istnieją jednoznacznie wyznaczone skalary $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ takie, że $x = \alpha_1 x_1 + \ldots + \alpha_n x_n$. Określamy funkcję $f : X \to Y$:

$$f(x) = \alpha_1 z_1 + \ldots + \alpha_n z_n = \alpha_1 f(x_1) + \ldots + \alpha_n f(x_n). \tag{10.2}$$

Jeżeli teraz $x, u \in X$,

$$x = \alpha_1 x_1 + \ldots + \alpha_n x_n, \quad u = \beta_1 x_1 + \ldots + \beta_n x_n,$$

i $\psi, \xi \in \mathbb{K}$, to

$$\begin{aligned} f(\psi x + \xi u) &= \psi \alpha_1 f(x_1) + \ldots + \psi \alpha_f(x_n) + \xi \beta_1 f(x_1) + \ldots + \xi \beta_n f(x_n) \\ &= \psi(\alpha_1 f(x_1) + \ldots + \alpha_n f(x_n)) + \xi(\beta_1 f(x_1) + \ldots + \beta_n f(x_n)) \\ &= \psi f(x) + \xi f(u), \end{aligned}$$

czyli f jest przekształceniem liniowym.

Każde przekształcenie liniowe spełniające (10.1) musi spełniać (10.2), skąd wynika jednoznaczność. $\square$

Pokazaliśmy, że przekształcenie liniowe $f \in L(X, Y)$ jest jednoznacznie zadane przez wartości, które przyjmuje na dowolnej bazie przestrzeni X.

Zadanie 10.3. Przekształcenie $f \in L(\mathbb{R}^2, \mathbb{R}^3)$ spełnia

$$f([1, -2]^T) = [1, 0, -3]^T, \qquad f([3, 1]^T) = [2, 1, -5]^T.$$

Znajdź macierz $A \in \mathbb{R}^{3,2}$ taką, że $f(\vec{x}) = A\vec{x}$ dla każdego $x \in \mathbb{R}^2$.

Twierdzenie 10.3. *Jeżeli* $\dim X = n$ *i* $\dim Y = m$*,* $m, n < \infty$*, to*

$$\dim L(X, Y) = mn.$$

Ponadto, jeżeli układ $x_1, \ldots, x_n$ *jest bazą przestrzeni liniowej* X*, natomiast układ* $y_1, \ldots, y_m$ *jest bazą przestrzeni* Y*, to układ przekształceń liniowych* $F_{i,j} : X \to Y$*,*

$$F_{i,j}(x_k) = \begin{cases} y_i & \text{gdy } k = j, \\ 0 & \text{gdy } k \neq j, \end{cases} \quad \text{dla} \quad k = 1, \ldots, n \tag{10.3}$$

$i = 1, \ldots, m$*,* $j = 1, \ldots, n$ *jest bazą przestrzeni* $L(X, Y)$*.*

Dowód. Na mocy stw. 10.2, wzór (10.3) określa przekształcenia liniowe. Pokażemy, że układ F_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$ rozpina przestrzeń $L(X, Y)$. Niech $f \in L(X, Y)$. Dla każdego wektora x_j z bazy przestrzeni X istnieją skalary $\alpha_{i,j}$, $i = 1, \ldots, m$ takie, że

$$f(x_j) = \sum_{i=1}^{m} \alpha_{i,j} y_i.$$

Zatem

$$f(x_j) = \sum_{i=1}^{m} \alpha_{i,j} F_{i,j}(x_j).$$

Zauważmy teraz, że dla $x \in X$, $x = \sum_{k=1}^{n} \gamma_k x_k$ mamy

$$F_{i,j}(x) = \sum_{k=1}^{m} \gamma_k F_{i,j}(x_k) = \gamma_j F_{i,j}(x_j).$$

Wobec tego

$$\begin{aligned} f(x) &= \sum_{j=1}^{n} \gamma_j f(x_j) \\ &= \sum_{j=1}^{n} \gamma_j \sum_{i=1}^{m} \alpha_{i,j} F_{i,j}(x_j) \\ &= \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{i,j} \gamma_j F_{i,j}(x_j) \\ &= \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{i,j} F_{i,j}(x). \end{aligned}$$

Inaczej mówiąc,

$$f = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{i,j} F_{i,j}.$$

Pokazaliśmy, że $L(X, Y) = \operatorname{span}\{F_{i,j} : i = 1, \ldots, m, j = 1, \ldots, n\}$.

Pokażemy teraz, że przekształcenia $F_{i,j}$ są liniowo niezależne. Załóżmy, że

$$0 = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{i,j} F_{i,j}.$$

Oznacza to, że dla każdego $x \in X$

$$0 = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{i,j} F_{i,j}(x).$$

W szczególności, dla $x = x_j$

$$\begin{aligned} 0 &= \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{i,j} F_{i,j}(x_j) \\ &= \sum_{i=1}^{m} \alpha_{i,j} y_i. \end{aligned}$$

Ponieważ układ $y_1, \ldots, y_m$ jest bazą, więc $\alpha_{i,j} = 0$.

Pokazaliśmy, że przestrzeń $L(X, Y)$ ma bazę liczącą mn elementów, więc $\dim L(X, Y) = mn$. □

Uwaga 10.1. Dowolna macierz $\mathbb{K}^{m,n}$ zadaje przekształcenie liniowe z $\mathbb{K}^n$ w $\mathbb{K}^m$. Bazą tej przestrzeni są jednostki macierzowe $E_{i,j}$ (zob. przykład 5.21). Jeżeli w $\mathbb{K}^n$ ustalimy bazę $\vec{x}_j = \vec{e}_j$, a w $\mathbb{K}^m$ bazę $\vec{y}_i = \vec{e}_i$, to otrzymanym w wyżej opisany sposób przekształceniom $F_{i,j}$ będą odpowiadać właśnie macierze $E_{i,j}$.

Przykład 10.7. W $\mathbb{R}^2$ mamy bazę $\vec{e}_1, \vec{e}_2$, w $\mathbb{R}[t]_2$ mamy bazę $1, t, t^2$. Przekształcenia bazowe $F_{i,j} \in L(\mathbb{R}^2, \mathbb{R}[t]_2)$, to, dla $\vec{x} = [x_1, x_2]^t$,

$$\begin{aligned} F_{1,1}(\vec{x}) &= x_1, & F_{1,2}(\vec{x}) &= x_2, \\ F_{2,1}(\vec{x}) &= x_1 \cdot t, & F_{2,2}(\vec{x}) &= x_2 \cdot t, \\ F_{3,1}(\vec{x}) &= x_1 \cdot t^2 & F_{3,2}(\vec{x}) &= x_2 \cdot t^2. \end{aligned}$$

Rozważmy przekształcenie $f \in L(\mathbb{R}^2, \mathbb{R}[t]_2)$,

$$f(\vec{x}) = x_1 - x_2 + (2x_1 - 3x_2)t + (x_1 + 4x_2)t^2.$$

Wówczas

$$f(\vec{e}_1) = 1 + 2t + t^2, \qquad f(\vec{e}_2) = -1 - 3t + 4t^2,$$

więc

$$\begin{array}{ccrcr} f & = & F_{1,1} & - & F_{1,2} \\ & + & 2F_{2,1} & - & 3F_{2,2} \\ & + & F_{3,1} & + & 4F_{3,2} \end{array}$$

10.3 Macierz przekształcenia liniowego

Aby zapisać przekształcenie liniowe $f : X \to Y$ w bazie $F_{i,j}$, zapisujemy obrazy wektorów bazy przestrzeni X w bazie przestrzeni Y:

$$f(x_j) = \sum_{i=1}^{m} \alpha_{i,j} y_i.$$

Wtedy

$$f = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{i,j} F_{i,j}.$$

Definicja 10.3. Jeżeli $f \in L(X, Y)$, układ $x_1, \ldots, x_n$ jest bazą przestrzeni X, układ $y_1, \ldots, y_m$ jest bazą przestrzeni Y, przekształcenia $F_{i,j}$ są bazą $L(X, Y)$ zdefiniowaną wzorami (10.3) i

$$f = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{i,j} F_{i,j},$$

to macierz

$$A = \begin{bmatrix} \alpha_{1,1} & \dots & \alpha_{1,n} \\ \dots & \dots & \dots \\ \alpha_{m,1} & \dots & \alpha_{m,n} \end{bmatrix}.$$

nazywamy *macierzą przekształcenia f w bazach $x_1, \dots, x_n$ i $y_1, \dots, y_m$*. Jeżeli $Y = X$ i $y_j = x_j$ dla $j = 1, \dots, n$, to mówimy o macierzy przekształcenia f w bazie $x_1, \dots, x_n$.

Przykład 10.8. Przekształcenie $f \in L(\mathbb{R}^2, \mathbb{R}[t]_2)$ z przykładu 10.7 ma w bazach $\vec{e}_1, \vec{e}_2$ i $1, t, t^2$ macierz

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

Zadanie 10.4. Znajdź macierz przekształcenia $D \in L(\mathbb{K}[x]_n, \mathbb{K}[x]_n)$,

$$D(p) = p'$$

w bazie $1, x, \dots, x^n$. (Zob. przykład 10.4).

Uwaga 10.2. Wiemy już, że $\dim(\mathbb{K}^n, \mathbb{K}^m) = mn = \dim \mathbb{K}^{m,n}$. Ponadto, każda macierz $A \in \mathbb{K}^{m,n}$ zadaje przekształcenie liniowe $f \in L(\mathbb{K}^n, \mathbb{K}^m)$:

$$f(\vec{x}) = A\vec{x}. \tag{10.4}$$

Macierz A jest wówczas macierzą przekształcenia f w bazach standardowych w $\mathbb{K}^n$ i $\mathbb{K}^m$. Wobec tego każde przekształcenie liniowe $f \in L(\mathbb{K}^n, \mathbb{K}^m)$ możemy zapisać w postaci (10.4).

Uwaga 10.3. Załóżmy, że $f \in L(X, Y)$, układ $x_1, \dots, x_n$ jest bazą przestrzeni X, układ $y_1, \dots, y_m$ – bazą Y i $A \in \mathbb{K}^{m,n}$ jest macierzą przekształcenia f w tych bazach. Jeżeli $x \in X$ i

$$x = \alpha_1 x_1 + \ldots + \alpha_n x_n, \qquad \vec{\alpha} = [\alpha_1, \dots, \alpha_n]^T \in \mathbb{K}^n,$$

to dla wektora $[\beta_1, \dots, \beta_m]^T = \vec{\beta} = A\vec{\alpha} \in \mathbb{K}^m$

$$f(x) = \beta_1 y_1 + \ldots + \beta_m y_m.$$

Stwierdzenie 10.4 (Macierz złożenia przekształceń)**.** *Załóżmy, że $f \in L(X, Y)$, $g \in L(Y, Z)$, gdzie X, Y, Z są przestrzeniami liniowymi nad ciałem $\mathbb{K}$,*

- *$x_1, \dots, x_p$ jest bazą X, $y_1, \dots, y_n$ – bazą Y, $z_1, \dots, z_m$ – bazą Z;*

- *$A \in \mathbb{K}^{n,p}$ jest macierzą f w bazach $x_1, \ldots, x_p$ i $y_1, \ldots, y_n$;*
- *$B \in \mathbb{K}^{m,n}$ jest macierzą g w bazach $y_1, \ldots, y_n$ i $z_1, \ldots, z_m$.*

Wówczas $C = B \cdot A \in \mathbb{K}^{m,p}$ jest macierzą przekształcenia $g \circ f \in L(X, Z)$ w bazach $x_1, \ldots, x_p$ i $z_1, \ldots, z_m$.

Zadanie 10.5. Udowodnij powyższe stwierdzenie.

10.4 Rząd, obraz i jądro przekształcenia liniowego

Zakładamy, że X i Y są przestrzeniami liniowymi nad ciałem $\mathbb{K}$ i $f \in L(X, Y)$ jest przekształceniem liniowym.

Definicja 10.4. Jądrem przekształcenia liniowego $f \in L(X, Y)$ nazywamy zbiór

$$\ker f = \{x \in X : f(x) = 0\} = f^{-1}(\{0\}).$$

Definicja 10.5. Obrazem przekształcenia liniowego $f \in L(X, Y)$ nazywamy zbiór

$$\operatorname{im} f = \{f(x) \in Y : x \in X\} = f(X).$$

Definicja 10.6. Rzędem przekształcenia liniowego $f \in L(X, Y)$ nazywamy liczbę

$$\operatorname{rank} f = \dim \operatorname{im} f.$$

Bardzo łatwo jest udowodnić co następuje:

Stwierdzenie 10.5. *Dla $f \in L(X, Y)$*

(i) Zbiór $\ker f$ jest podprzestrzenią liniową w X.

(ii) Zbiór $\operatorname{im} f$ jest podprzestrzenią liniową w Y.

Przykład 10.9. Wyznaczymy jądro, obraz i rząd przekształcenia $f \in L(\mathbb{R}^2, \mathbb{R}[t]_2)$,

$$f(\vec{x}) = x_1 - x_2 + (2x_1 - 3x_2)t + (x_1 + 4x_2)t^2$$

z przykładu 10.7.

Zauważmy, że $f(\vec{x}) = 0$ wtedy i tylko wtedy, gdy

$$x_1 - x_2 = 0, \quad 2x_1 - 3x_2 = 0, \quad x_1 + 4x_2 = 0.$$

Zachodzi to tylko wtedy, gdy $\vec{x} = 0.$, więc $\ker f = \{0\}$.

Aby wyznaczyć obraz, zauważmy, że

$$f(\vec{x}) = x_1(1 + 2t + t^2) + x_2(-1 - 3t + 4t^2),$$

więc $\operatorname{im} f = \operatorname{span}(1 + 2t + t^2, -1 - 3t + 4t^2)$ i te dwa wielomiany są liniowo niezależne, więc są też bazą $\operatorname{im} f$. Zatem $\operatorname{rank} f = 2$.

Zadanie 10.6. Wyznacz jądro, obraz i rząd przekształcenia D z zadania 10.4.

Odpowiednikiem twierdzenia 6.3, które zachodzi dla macierzy jest:

Twierdzenie 10.6. *Załóżmy, że X i Y są przestrzeniami liniowymi skończonego wymiaru, $f \in L(X, Y)$. Wówczas*

$$\dim \ker f + \dim \operatorname{im} f = \dim X.$$

Dowód tego twierdzenia jest bardzo podobny do dowodu tw. 6.3, dlatego go pominiemy.

Zadanie 10.7. Załóżmy, że $f \in L(X, Y)$ i $A \in \mathbb{K}^{m,n}$ jest macierzą przekształcenia f w bazach $x_1, \ldots, x_n$ i $y_1, \ldots, y_m$. Zastanów się, jak, znając A i obie bazy, wyznaczyć jądro i obraz przekształcenia f.

Zadanie 10.8. Załóżmy, że $f \in L(X, Y)$ i układ $x_1, \ldots, x_n$ jest bazą przestrzeni X. Pokaż, że

$$\operatorname{im} f = \operatorname{span}(f(x_1), \ldots, f(x_n)).$$

10.5 Monomorfizmy i epimorfizmy

Definicja 10.7. Przekształcenie liniowe $f \in L(X, Y)$ nazwiemy

(a) *monomorfizmem*, jeżeli f jest różnowartościowe,

(b) *epimorfizmem*, jeżeli $\operatorname{im} f = Y$.

Przykład 10.10. Przekształcenie liniowe $f \in L(\mathbb{R}^2, \mathbb{R}^3)$ dane wzorem

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_1 - x_2 \\ x_1 + 3x_2 \end{bmatrix}$$

jest monomorfizmem: jeżeli $f(\vec{x}) = \vec{y} \in \mathbb{R}^3$, to

$$\begin{aligned} x_1 + x_2 &= y_1, \\ x_1 - x_2 &= y_2, \\ x_1 + 3x_2 &= y_3 \end{aligned}$$

i z pierwszych dwóch równań wynika, że x_1 i x_2 są jednoznacznie określone przez y_1, y_2:

$$x_1 = \frac{y_1 + y_2}{2}, \qquad x_2 = \frac{y_1 - y_2}{2}.$$

Jeżeli więc

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = f\left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\right) = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix},$$

to

$$x_1 = \frac{y_1 + y_2}{2} = u_1, \qquad x_2 = \frac{y_1 - y_2}{2} = u_2,$$

więc $[u_1, u_2]^T = [x_1, x_2]^T$. Zatem przekształcenie f jest różnowartościowe, jest więc monomorfizmem.

Przykład 10.11. Przekształcenie $D \in L(\mathbb{K}[t]_n, \mathbb{K}[t]_{n-1})$, $D(p) = p'$ jest epimorfizmem, ale nie jest monomorfizmem.

Stwierdzenie 10.7. *Załóżmy, że* $\dim X < \infty$ *i* $f \in L(X, Y)$*. Następujące warunki są równoważne:*

(i) f *jest monomorfizmem,*

(ii) $\ker f = \{0\}$,

(iii) $\dim imf = \dim X$,

(iv) dla pewnej bazy $x_1, \ldots, x_n$ *przestrzeni* X *układ* $f(x_1), \ldots, f(x_n)$ *jest bazą przestrzeni* $\operatorname{im} f$,

(v) dla dowolnej bazy $x_1, \ldots, x_n$ *przestrzeni* X *układ* $f(x_1), \ldots, f(x_n)$ *jest bazą przestrzeni* $\operatorname{im} f$,

(vi) dla każdego liniowo niezależnego układu wektorów $x_1, \ldots, x_k \in X$ *układ* $f(x_1), \ldots, f(x_k)$ *jest liniowo niezależny.*

Dowód tego stwierdzenia zostawiamy jako ćwiczenie.

Zadanie 10.9. Załóżmy, że przekształcenie liniowe $f \in L(X, Y)$ ma w pewnych bazach macierz $A \in \mathbb{K}^{m,n}$. Pokaż, że

(a) f jest monomorfizmem wtedy i tylko wtedy, gdy $\ker A = \{0\}$,

(b) f jest epimorfizmem wtedy i tylko wtedy, gdy $\operatorname{rank} A = m$.

10.6 Izomorfizmy

Definicja 10.8. Przekształcenie liniowe $f \in L(X, Y)$, które jest monomorfizmem i epimorfizmem, nazywamy *izomorfizmem*.

Jeżeli dla pary przestrzeni liniowych X, Y nad ciałem $\mathbb{K}$ istnieje izomorfizm $f : X \to Y$, to mówimy, że przestrzeń X jest *izomorficzna* z przestrzenią Y.

Przykład 10.12. Przekształcenie liniowe $f \in L(\mathbb{K}^n, \mathbb{K}[t]_{n-1})$ zadane wzorem

$$f(\vec{x}) = x_1 + x_2 t + x_3 t^2 + \ldots + x_n t^{n-1}$$

jest izomorfizmem, a przestrzenie $\mathbb{K}^n$ i $\mathbb{K}[t]_{n-1}$ są izomorficzne.

Przykład 10.13. Jeżeli macierz $A \in \mathbb{K}^{n,n}$ jest nieosobliwa, to przekształcenie liniowe $f : \mathbb{K}^n \to \mathbb{K}^n$ zadane wzorem $f(\vec{x}) = A\vec{x}$ jest izomorfizmem: f jest monomorfizmem, gdyż, jeżeli $f(\vec{x}) = f(\vec{u})$, to $A(\vec{x} - \vec{u}) = 0$, więc $\vec{x} - \vec{u} = 0$, gdyż A jest nieosobliwa. Ponadto, dla każdego $\vec{y} \in \mathbb{R}^n$ $f(A^{-1}\vec{y}) = \vec{y}$, więc f jest epimorfizmem.

Przykład 10.14. Przekształcenie liniowe $f \in L(X, X)$ takie, że $f(x) = x$ dla każdego wektora $x \in X$ jest izomorfizmem. Przekształcenie to nazywamy *identycznościowym* i oznaczamy id_X.

Twierdzenie 10.8. *Jeżeli przekształcenie liniowe $f : X \to Y$ jest izomorfizmem, to określona jest funkcja odwrotna do f: $f^{-1} : Y \to X$ taka, że*

$$f \circ f^{-1} = \mathrm{id}_Y, \qquad f^{-1} \circ f = \mathrm{id}_X$$

oraz $f^{-1} \in L(Y, X)$.

Dowód. Istnienie funkcji $f^{-1} : Y \to X$ wynika z tego, że f jest bijekcją z X na Y. Pokażemy, że f^{-1} jest przekształceniem liniowym. Niech $\alpha_1, \alpha_2 \in \mathbb{K}$, $y_1, y_2 \in Y$ i $y = \alpha_1 y_1 + \alpha_2 y_2$. Istnieją jednoznacznie wyznaczone wektory $x_1, x_2, x \in X$ takie, że

$$f(x_1) = y_1, \quad f(x_2) = y_2, \quad f(x) = y.$$

Ponadto,

$$f(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 f(x_1) + \alpha_2 f(x_2) = \alpha_1 y_1 + \alpha_2 y_2 = y,$$

więc $x = \alpha_1 x_1 + \alpha_2 x_2$. Zatem

$$\begin{aligned} f^{-1}(\alpha_1 y_1 + \alpha_2 y_2) &= f^{-1}(\alpha_1 f(x_1) + \alpha_2 f(x_2)) = f^{-1}(f(\alpha_1 x_1 + \alpha_2 x_2)) \\ &= f^{-1}(f(x)) = x \\ &= \alpha_1 x_1 + \alpha_2 x_2 = \alpha_1 f^{-1}(f(x_1)) + \alpha_2 f^{-1}(f(x_2)) \\ &= \alpha_1 f^{-1}(y_1) + \alpha_2 f^{-1}(y_2). \end{aligned}$$

□

Twierdzenie 10.9. *Jeżeli X jest przestrzenią liniową nad ciałem $\mathbb{K}$ i $\dim X = n$, to przestrzeń X jest izomorficzna z przestrzenią $\mathbb{K}^n$.*

Dowód. Wskażemy izomorfizm $f : X \to \mathbb{K}^n$. Niech $x_1, \ldots, x_n$ będzie bazą przestrzeni X. Dla $x \in X$ istnieją jednoznacznie wyznaczone skalary $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ takie, że

$$x = \alpha_1 x_1 + \ldots + \alpha_n x_n.$$

Określamy

$$f(x) = [\alpha_1, \ldots, \alpha_n]^T.$$

Wówczas $\operatorname{im} f = \mathbb{K}^n$ i f jest monomorfizmem, jest to więc izomorfizm. □

Uwaga 10.4. Dla przestrzeni liniowych nad ciałem $\mathbb{K}$ prawdziwe są następujące stwierdzenia:

- każda przestrzeń X jest izomorficzna ze sobą,
- jeżeli przestrzeń X jest izomorficzna z przestrzenią Y, to także przestrzeń Y jest izomorficzna z X,
- jeżeli przestrzeń X jest izomorficzna z przestrzenią Y, a Y jest izomorficzna z przestrzenią Z, to także X jest izomorficzna z Z.

Inaczej mówiąc, izomorfizm przestrzeni liniowych jest relacją równoważności określoną na zbiorze wszystkich przestrzeni liniowych nad danym ciałem $\mathbb{K}$. W przypadku przestrzeni skończonego wymiaru klasy abstrakcji tej relacji równoważności odpowiadają wymiarom przestrzeni, a reprezentantem danej klasy jest przestrzeń $\mathbb{K}^n$.

10.7 Przestrzenie dualne

Ciało $\mathbb{K}$ możemy utożsamić z przestrzenią liniową $\mathbb{K}^1$.

Definicja 10.9. Niech X będzie przestrzenią liniową nad ciałem $\mathbb{K}$. *Przestrzeń dualna (sprzężoną)* do X nazywamy przestrzeń $L(X, \mathbb{K})$. Oznaczamy ją symbolem X^*. Elementy przestrzeni X^* nazywamy *funkcjonałami liniowymi* na X. Funkcjonały liniowe często oznacza się symbolami x^*, u^*, itp.

Z twierdzenia 10.3 wynika natychmiast:

Twierdzenie 10.10. *Jeżeli X jest przestrzenią liniową i $\dim X = n < +\infty$, to także $\dim X^* = n$.*

Przykład 10.15 (Funkcjonały na $\mathbb{K}^n$)**.** Niech $a = [a_1, \dots, a_n] \in \mathbb{K}^{1,n}$. Wówczas a zadaje funkcjonał na $\mathbb{K}^n$:

$$a^*(\vec{x}) = a \cdot \vec{x} = \sum_{k=1}^{n} a_k x_k, \qquad \vec{x} = [x_1, \dots, x_n]^T$$

jest funkcjonałem liniowym na $\mathbb{K}^n$.

Możemy określić $F \in L(\mathbb{K}^{1,n}, (\mathbb{K}^n)^*)$:

$$F(a)(\vec{x}) = a \cdot \vec{x}.$$

Tak zadane przekształcenie jest izomorfizmem. Zatem wzór (10.15) opisuje wszystkie funkcjonały liniowe na $\mathbb{K}^n$ i przestrzeń $(\mathbb{K}^n)^*$ możemy utożsamić z $\mathbb{K}^{1,n}$.

Przykład 10.16. Niech $t_0 \in \mathbb{K}$. Wówczas funkcja $x^* : \mathbb{K}[t] \to \mathbb{K}$ dana wzorem

$$x^*(p) = p(t_0)$$

jest funkcjonałem liniowym na $\mathbb{K}[t]_n$.

Przykład 10.17. *Ślad macierzy* $A = [a_{i,j}]_{i,j} \in \mathbb{K}^{m,n}$ definiujemy następująco:

$$\mathrm{tr} A = \sum_{i=1}^{\min(m,n)} a_{i,i}.$$

Jest to funkcjonał liniowy na przestrzeni $\mathbb{K}^{m,n}$

Przykład 10.18. Jeżeli $(X, \langle \cdot, \cdot \rangle)$ jest przestrzenią euklidesową i $\dim X = n < \infty$, to każdy wektor $a \in X$ określa funkcjonał liniowy na X wzorem

$$a^*(x) = \langle a, x \rangle \tag{10.5}$$

i gdy $a \neq b$, to funkcjonały a^*, b^* są różne: jeżeli $a^*(x) = b^*(x)$ dla każdego $x \in X$, to $\langle a - b, x \rangle = 0$ dla każdego x, więc $a - b = 0$.

Możemy określić monomorfizm $F \in L(X, X^*)$ wzorem

$$F(a)(x) = \langle a, x \rangle.$$

Wobec 10.7, $\dim imF = n$, więc $\operatorname{im} F = X^*$. Zatem F jest izomorfizmem i każdy funkcjonał liniowy na $(X, \langle \cdot, \cdot \rangle)$ jest postaci (10.5).

10.8 Bazy dualne

Rozważamy następujący problem: w przestrzeni liniowej X nad ciałem $\mathbb{K}$ dana jest baza $x_1, \ldots, x_n$. Chcemy znaleźć funkcjonały liniowe $x_1^*, \ldots, x_n^* \in X^*$ takie, że dla dowolnego wektora $x \in X$

$$x = \sum_{j=1}^{n} x_j^*(x)x_j.$$

Inaczej mówiąc, szukamy funkcjonałów, za pomocą których możemy wyznaczyć współrzędne dowolnego wektora z przestrzeni X w zadanej bazie.

Przykład 10.19. W przestrzeni $\mathbb{R}^n$ rozważamy bazę $\vec{e}_1, \ldots, \vec{e}_n$. Wówczas funkcjonały liniowe

$$e_1^*(\vec{x}) = x_1, \quad \ldots, \quad e_n^*(\vec{x}) = x_n$$

spełniają, dla każdego wektora $\vec{x} \in \mathbb{R}^n$

$$\vec{x} = \sum_{j=1}^{n} e_j^*(\vec{x})\vec{e}_j.$$

Za pomocą funkcjonałów e_j^* jesteśmy w stanie wyznaczyć współczynniki każdego wektora $\vec{x}$ w zadanej bazie $\vec{e}_1, \ldots, \vec{e}_n$.

Przykład 10.20. W przestrzeni $\mathbb{R}^2$ dana jest baza

$$\vec{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}.$$

Szukamy funkcjonałów u^*, v^* takich, że dla każdego wektora $\vec{x} \in \mathbb{R}^2$

$$\vec{x} = u^*(\vec{x})\vec{u} + v^*(\vec{x})\vec{v}.$$

Jeżeli $x = [x_1, x_2]^T$, $u^*(\vec{x}) = \alpha$, $v^*(\vec{x}) = \beta$, to

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \alpha \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} 5 \\ 3 \end{bmatrix},$$

skąd

$$\alpha = 3x_1 - 5x_2, \qquad \beta = -x_1 + 2x_2.$$

Zatem

$$u^*(\vec{x}) = 3x_1 - 5x_2, \qquad v^*(\vec{x}) = -x_1 + 2x_2.$$

Przykład 10.21. Jeżeli układ $x_1, \ldots, x_n$ jest bazą ortonormalną w przestrzeni euklidesowej lub unitarnej $(X, \langle\cdot,\cdot\rangle)$, to szukane funkcjonały są postaci $x_j^*(x) = \langle x_j, x\rangle$.

Zauważmy, że szukane funkcjonały są jednoznacznie zadane przez następujący układ warunków:

$$x_i^*(x_j) = \begin{cases} 1 & \text{gdy } i = j, \\ 0 & \text{gdy } i \neq j. \end{cases}.$$

Twierdzenie 10.11. *Załóżmy, że X jest przestrzenią liniową nad ciałem $\mathbb{K}$, $\dim X = n$ i dane są wektory $x_1, \ldots, x_n \in X$ oraz funkcjonały $x_1^*, \ldots, x_n^* \in X^*$ takie, że dla $i, j = 1, \ldots, n$*

$$x_i^*(x_j) = \begin{cases} 1 & \text{gdy } i = j, \\ 0 & \text{gdy } i \neq j. \end{cases} \tag{10.6}$$

Wówczas

(i) *układ $x_1, \ldots, x_n$ jest bazą przestrzeni X,*

(ii) *układ $x_1^*, \ldots, x_n^*$ jest bazą przestrzeni X^*,*

(iii) *dla każdego wektora $x \in X$*

$$x = \sum_{i=1}^{n} x_i^*(x) x_i,$$

(iv) *dla każdego funkcjonału $x^* \in X^*$*

$$x^* = \sum_{i=1}^{n} x^*(x_i) x_i^*.$$

Dowód. (i): $\dim X = n$, wystarczy więc pokazać, że wektory $x_1, \ldots, x_n$ są liniowo niezależne. Załóżmy, że dla pewnych skalarów $\alpha_j \in \mathbb{K}$

$$\alpha_1 x_1 + \ldots + \alpha_n x_n = 0.$$

Wówczas dla $i = 1, \ldots, n$

$$0 = x_i^*(\alpha_1 x_1 + \ldots + \alpha_n x_n) = \alpha_1 x_i^*(x_1) + \ldots + \alpha_n x_i^*(x_n) = \alpha_i x_i^*(x_i) = \alpha_i.$$

Zatem układ $x_1, \ldots, x_n$ jest liniowo niezależny, więc jest bazą X.

(ii): $\dim X^* = n$, wystarczy więc pokazać, że funkcjonały $x_1^*, \ldots, x_n^*$ są liniowo niezależne. Załóżmy, że dla pewnych skalarów $\beta_i \in \mathbb{K}$

$$0 = \beta_1 x_1^* + \ldots + \beta_n x_n^*,$$

czyli, dla każdego $x \in X$

$$0 = \beta_1 x_1^*(x) + \ldots + \beta_n x_n^*(x).$$

Dla $x = x_j$ dostajemy

$$0 = \beta_1 x_1^*(x_j) + \ldots + \beta_n x_n^*(x_j) = \beta_j x_j^*(x_j) = \beta_j,$$

czyli układ funkcjonałów $x_1^*, \ldots, x_n^*$ jest liniowo niezależny, jest więc bazą przestrzeni dualnej do X.

(iii): Niech $x \in X$. Istnieją skalary $\alpha_j \in \mathbb{K}$ takie, że $x = \alpha_1 x_1 + \ldots + \alpha_n x_n$. Dla $i = 1, \ldots, n$ mamy

$$x_i^*(x) = x_i^*(\alpha_1 x_1 + \ldots + \alpha_n x_n) = \alpha_i x_i^*(x_i) = \alpha_i.$$

(iv): Załóżmy, że $x^* \in X^*$. Istnieją $\beta_i \in \mathbb{K}$ takie, że $x^* = \beta_1 x_1^* + \ldots + \beta_n x_n^*$. Zatem, dla $j = 1, \ldots, n$

$$x^*(x_j) = \beta_1 x_1^*(x_j) + \ldots + \beta_n x_n^*(x_j) = \beta_j x_j^*(x_j) = \beta_j.$$

□

Uwaga 10.5. Zauważmy, że w założeniach twierdzenie 10.11 nie zakładamy, że układ $x_1, \ldots, x_n$ jest bazą przestrzeni X. Wynika to jedynie z warunku (10.6).

Definicja 10.10. Załóżmy, że układ $x_1, \ldots, x_n$ jest bazą przestrzeni X. Wówczas z tw. 10.11 wynika istnienie bazy $x_1^*, \ldots, x_n^*$ przestrzeni X^* takiej, że spełnione są warunki

$$x_i^*(x_j) = \delta_{i,j}.$$

Bazę $x_1^*, \ldots, x_n^*$ nazywamy *bazą dualną (sprzężoną)* do bazy $x_1, \ldots, x_n$. Jeżeli zachodzi (10.6), to możemy także mówić, że baza $x_1, \ldots, x_n$ przestrzeni X jest dualna do bazy $x_1^*, \ldots, x_n^*$ przestrzeni X^*.

Uwaga 10.6. Nic nie stoi na przeszkodzie, aby rozważać przestrzeń dualną do X^* (którą możemy oznaczyć X^{**}). Można określić tzw. *kanoniczny monomorfizm* Φ przestrzeni X w X^{**}:

$$\Phi(x)(x^*) = x^*(x), \qquad x \in X, \quad x^* \in X^*$$

W powyższym wzorze $\Phi(x)$ jest funkcjonałem liniowym na X^* i $\Phi(x)(x^*)$ to wartość (skalar), którą ten funkcjonał przyjmuje dla elementu $x^* \in X^*$.

Jeżeli $\dim X < \infty$, to Φ jest „naturalnym" izomorfizmem przestrzeni X i X^{**}.

Przykład 10.22. (Baza Lagrange'a w $\mathbb{K}[t]_n$). Wybierzmy $n+1$ różnych liczb $t_0, t_1, \ldots, t_n$ z ciała $\mathbb{K}$ i rozpatrzmy $n+1$ funkcjonałów liniowych na przestrzeni $\mathbb{K}[t]_n$:

$$l_0^*(p) = p(t_0), \quad l_1^*(p) = p(t_1), \quad \ldots, \quad l_n^*(p) = p(t_n).$$

Określmy też wielomiany:

$$l_j(t) = \prod_{\substack{i=0,\ldots,n \\ i \neq j}} \frac{t - t_i}{t_j - t_i}, \qquad j = 0, 1, \ldots, n.$$

Wówczas $l_i^*(l_j) = \delta_{i,j}$. Układ $l_0, l_1, \ldots, l_n$ jest więc bazą przestrzeni $\mathbb{K}[x]_n$, którą nazywamy *bazą Lagrange'a* dla punktów $t_0, t_1, \ldots, t_n$. Baza dualna to w tym przypadku funkcjonały wyznaczające wartości wielomianu w punktach t_j.

Jeżeli szukamy wielomianu $p \in \mathbb{K}[x]_n$ takiego, że $p(x_j) = y_j$ dla zadanych wartości $y_0, \ldots, y_n$, to wielomian ten możemy zapisać w bazie Lagrange'a jako

$$p = \sum_{j=0}^{n} y_j l_j.$$

Przykład 10.23. Załóżmy, że w przestrzeni $\mathbb{K}^n$ mamy daną bazę $\vec{b}_1, \ldots, \vec{b}_n$. Wówczas bazę dualną możemy utożsamić z układem macierzy $a_i \in \mathbb{K}^{1,n}$ takich, że $a_i \vec{b}_j = \delta_{i,j}$. Rozważmy macierze

$$B = [\vec{b}_1, \ldots, \vec{b}_n] \in \mathbb{K}^{n,n}, \qquad A = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{K}^{n,n}.$$

Wówczas $AB = I_n$, czyli $A = B^{-1}$. Zatem baza dualna do $\vec{b}_1, \ldots, \vec{b}_n$ składa się z funkcjonałów odpowiadających wierszom macierzy odwrotnej do B:

$$b_i^*(\vec{x}) = \vec{e}_i^{\,T} B^{-1} \vec{x}.$$

Rozdział 11

Endomorfizmy

W tym rozdziale zajmiemy się badaniem przestrzeni $L(X, X)$, której elementy to przekształcenia liniowe danej przestrzeni liniowej X w siebie. Wygodnym sposobem zapisu takiego przekształcenia jest jego macierz w ustalonej bazie przestrzeni X. Dlatego najpierw wprowadzimy kilka nowych pojęć dotyczących macierzy kwadratowych, które później wykorzystamy przy omawianiu przekształceń liniowych

11.1 Wielomian charakterystyczny macierzy

Definicja 11.1. Niech $A \in \mathbb{K}^{n,n}$. *Wielomian charakterystyczny* macierzy A jest zdefiniowany jako

$$p_A(\lambda) = \det_n(A - \lambda I_n).$$

Uwaga 11.1. Można pokazać, że p_A jest wielomianem stopnia n o współczynnikach z ciała $\mathbb{K}$.

Przykład 11.1. Jeżeli $A = \begin{bmatrix} -5 & -4 \\ 8 & 7 \end{bmatrix}$, to

$$\begin{aligned} p_A(\lambda) &= \det \begin{bmatrix} -5-\lambda & -4 \\ 8 & 7-\lambda \end{bmatrix} = (-5-\lambda)(7-\lambda) + 32 \\ &= \lambda^2 - 2\lambda - 3 = (\lambda+1)(\lambda-3). \end{aligned}$$

Przykład 11.2. Jeżeli $A = \operatorname{diag}(a_1, a_2, \ldots, a_n)$ (macierz diagonalna), to

$$p_A(\lambda) = (a_1 - \lambda)(a_2 - \lambda)\ldots(a_n - \lambda).$$

11.2 Podobieństwo macierzy

Wprowadźmy następującą definicję:

Definicja 11.2. Powiemy, że macierze $A, B \in \mathbb{K}^{n,n}$ są podobne, jeżeli istnieje macierz nieosobliwa $C \in \mathbb{K}^{n,n}$ taka, że $B = CAC^{-1}$.

Zadanie 11.1. Pokaż, że podobieństwo macierzy jest relacją równoważności na zbiorze $\mathbb{K}^{n,n}$.

Stwierdzenie 11.1 (Niezmienniki podobieństwa)**.** *Jeżeli macierze $A, B \in \mathbb{K}^{n,n}$ są podobne, to*

(i) $\det A = \det B$,

(ii) *macierze A i B mają takie same wielomiany charakterystyczne, czyli $p_A = p_B$.*

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

(iv) $\operatorname{rank} A = \operatorname{rank} B$.

Dowód. Podpunkt (i) jest wnioskiem z twierdzenia Cauchy'ego 8.12. $B = C^{-1}AC$ dla pewnej nieosobliwej macierzy C i dla dowolnej liczby $\lambda \in \mathbb{K}$:

$$C^{-1}(A - \lambda I_n)C = C^{-1}AC - C^{-1}\lambda I_n C = B - \lambda I_n,$$

co oznacza, że macierze $A - \lambda I_n$ i $B - \lambda I_n$ też są podobne, więc ich wyznaczniki są równe, czyli

$$p_A(\lambda) = \det(A - \lambda I_n) = \det(B - \lambda I_n) = p_B(\lambda).$$

Pozostałe podpunkty zostawiamy jako ćwiczenia. □

Zadanie 11.2. Udowodnij podpunkty (iii) i (iv) w stw. 11.1.

Zadanie 11.3. Dla $A \in \mathbb{K}^{n,n}$, czy $\operatorname{im} A$ i $\ker A$ są niezmiennikami podobieństwa macierzy?

Zadanie 11.4. Załóżmy, że macierze $A, B \in \mathbb{K}^{n,n}$ są podobne i $\gamma_0, \gamma_1, \ldots, \gamma_r \in \mathbb{K}$. Pokaż, że macierze

$$C = \sum_{k=0}^{r} \gamma_k A^k, \qquad D = \sum_{k=0}^{r} \gamma_k B^k$$

też są podobne. (Przyjmujemy, że $A^0 = I_n$.)

11.3 Macierz zmiany bazy

Rozważamy następujące zagadnienie: załóżmy, że w przestrzeni liniowej X mamy dane dwie bazy: $x_1, \ldots, x_n$ i $u_1, \ldots, u_n$ i dla danego wektora x znamy jego współrzędne w bazie $x_1, \ldots, x_n$

$$x = \alpha_1 x_1 + \ldots + \alpha_n x_n, \qquad \vec{\alpha} = [\alpha_1, \ldots, \alpha_n]^T \in \mathbb{K}^n.$$

Szukamy macierzy $C \in \mathbb{K}^{n,n}$, niezależnej od x i takiej, że dla wektora

$$\vec{\beta} = [\beta_1, \ldots, \beta_n]^T = C\vec{\alpha}$$

mamy

$$x = \beta_1 u_1 + \ldots + \beta_n u_n.$$

Taką macierz C można zdefiniować następująco:

Definicja 11.3. W przestrzeni liniowej X dane są dwie bazy: $\mathbf{x} = (x_1, \ldots, x_n)$ i $\mathbf{u} = (u_1, \ldots, u_n)$. *Macierz zmiany bazy* z $\mathbf{x}$ na $\mathbf{u}$ jest to macierz przekształcenia identycznościowego $\mathrm{id}_X : X \to X$ w bazach $\mathbf{x}$ i $\mathbf{u}$.

Stwierdzenie 11.2. *Jeżeli w przestrzeni liniowej X dane są dwie bazy: $\mathbf{x} = (x_1, \ldots, x_n)$ i $\mathbf{u} = (u_1, \ldots, u_n)$, oraz baza $\mathbf{u}^* = (u_1^*, \ldots, u_n^*)$ przestrzeni X^* dualna do bazy $\mathbf{u}$, to macierz zmiany bazy z $\mathbf{x}$ na $\mathbf{u}$ jest postaci*

$$C = [u_i^*(x_j)]_{i,j} \in \mathbb{K}^{n,n}.$$

Dowód. W j-tej kolumnie macierzy C muszą się znaleźć współrzędne wektora $x_j = \mathrm{id}_X(x_j)$ w bazie $\mathbf{u}$, czyli liczby $u_1^*(x_j), \ldots, u_n^*(x_j)$. □

Przykład 11.3. W przestrzeni $\mathbb{K}^n$ dane są dwie bazy $\mathbf{a} = (\vec{a}_1, \ldots, \vec{a}_n)$ i $\mathbf{b} = (\vec{b}_1, \ldots, \vec{b}_n)$. Wówczas macierz zmiany bazy z $\mathbf{a}$ na $\mathbf{b}$ ma postać

$$C = [\vec{b}_1, \ldots, \vec{b}_n]^{-1} \cdot [\vec{a}_1, \ldots, \vec{a}_n].$$

Zauważmy, że macierz zmiany bazy zawsze jest nieosobliwa. Odnotujmy jeszcze jedno stwierdzenie:

Stwierdzenie 11.3. *X jest przestrzenią liniową nad ciałem $\mathbb{K}$ wymiaru n*

(i) Jeżeli $A \in \mathbb{K}^{n,n}$ jest macierzą zmiany bazy $\mathbf{a}$ na bazę $\mathbf{b}$, to macierz A^{-1} jest macierzą zmiany bazy $\mathbf{b}$ na bazę $\mathbf{a}$.

(ii) Jeżeli $A \in \mathbb{K}^{n,n}$ jest macierzą zmiany bazy $\mathbf{a}$ na bazę $\mathbf{b}$, $B \in \mathbb{K}^{n,n}$ jest macierzą zmiany bazy $\mathbf{b}$ na bazę $\mathbf{c}$, to $B \cdot A$ jest macierzą zmiany bazy $\mathbf{a}$ na bazę $\mathbf{c}$.

11.4 Definicja endomorfizmu

Definicja 11.4. *Endomorfizmem* przestrzeni liniowej X w siebie nazywamy dowolne przekształcenie liniowe $f \in L(X, X)$. Przestrzeń $L(X, X)$ będziemy zapisywać jako $L(X)$.

Stwierdzenie 11.4. *Załóżmy, że* $\dim X < \infty$ *i* $f \in L(X)$ *jest endomorfizmem. Wówczas następujące warunki są równoważne:*

(i) f *jest monomorfizmem,*

(ii) f *jest epimorfizmem,*

(iii) f *jest izomorfizmem.*

Zadanie 11.5. Udowodnij powyższe stwierdzenie. Podaj przykład na to, że założenie $\dim X < \infty$ jest potrzebne.

Definicja 11.5. Endomorfizm f przestrzeni liniowej X, który jest izomorfizmem, nazywamy *automorfizmem*.

Załóżmy teraz, że w przestrzeni X mamy dane dwie bazy $\mathbf{a} = (a_1, \ldots, a_n)$ oraz $\mathbf{b} = (b_1, \ldots, b_n)$. Niech $f \in L(X)$ i $M_{f,\mathbf{a}} \in \mathbb{K}^{n,n}$ jest macierzą endomorfizmu f w bazie $\mathbf{a}$. Oznacza to, że, dla

$$x = \alpha_1 a_1 + \ldots + \alpha_n a_n \in X, \qquad \vec{\alpha} = [\alpha_1, \ldots, \alpha_n]^T \tag{11.1}$$

mamy

$$f(x) = \theta_1 a_1 + \ldots + \theta_n a_n, \qquad \text{dla } [\theta_1, \ldots, \theta_n]^T = \vec{\theta} = M_{f,\mathbf{a}}\vec{\alpha}.$$

Niech $C \in \mathbb{K}^{n,n}$ będzie macierzą zmiany bazy $\mathbf{a}$ na bazę $\mathbf{b}$. Wówczas macierz $M_{f,\mathbf{b}}$ endomorfizmu f w bazie $\mathbf{b}$ spełnia

$$M_{f,\mathbf{b}} = C \cdot M_{f,\mathbf{a}} \cdot C^{-1}.$$

Istotnie, możemy napisać $f = \mathrm{id}_X \circ f \circ \mathrm{id}_X$, natomiast C to macierz endomorfizmu id_X w bazach $\mathbf{a}$ i $\mathbf{b}$, natomiast C^{-1} to macierz id_X w bazach $\mathbf{b}$ i $\mathbf{a}$. Zatem, jeżeli $x = \beta_1 b_1 + \ldots + \beta_n b_n$, czyli $\vec{\beta} = [\beta_1, \ldots, \beta_n]^T$ jest wektorem współrzędnych wektora x w bazie $\mathbf{b}$, to

- $C^{-1}\vec{\beta}$ jest wektorem współrzędnych wektora x w bazie $\mathbf{a}$,
- $M_{f,\mathbf{a}} \cdot C^{-1}\vec{\beta}$ jest wektorem współrzędnych wektora $f(x)$ w bazie $\mathbf{a}$
- $C \cdot M_{f,\mathbf{a}} \cdot C^{-1}\vec{\beta}$ jest wektorem współrzędnych wektora $f(x)$ w bazie $\mathbf{b}$.

Udowodniliśmy

Stwierdzenie 11.5. *Macierze tego samego endomorfizmu f w różnych bazach są podobne.*

Uwaga 11.2. Jako że macierz C w definicji 11.2 jest nieosobliwa, jest ona macierzą zmiany bazy w przestrzeni $\mathbb{K}^n$. Zatem macierze A i B są podobne wtedy i tylko wtedy, gdy są macierzami pewnego przekształcenie liniowego $F \in L(\mathbb{K}^n)$ w różnych bazach.

Zgodnie ze stw. 11.1, wyznacznik, wielomian charakterystyczny i ślad są niezmiennikami podobieństwa macierzy. Wynika z tego poprawność następującej definicji:

Definicja 11.6. Załóżmy, że X jest przestrzenią liniową nad ciałem $\mathbb{K}$ i $\dim X < \infty$. Niech $f \in L(X)$ i $\mathbf{b}$ będzie dowolną bazą przestrzeni X, $M_{f,\mathbf{b}}$ to macierz f w tej bazie. Dla endomorfizmu f określamy

- *ślad* $\operatorname{tr} f$ jako ślad macierzy $M_{f,\mathbf{b}}$:
$$\operatorname{tr} f = \operatorname{tr}(M_{f,\mathbf{b}}).$$
- *wyznacznik* $\det f$ jako wyznacznik macierzy $M_{f,\mathbf{b}}$:
$$\det f = \det(M_{f,\mathbf{b}}).$$
- *wielomian charakterystyczny* $p_f(\lambda)$ jako wielomian charakterystyczny macierzy $M_{f,\mathbf{b}}$:
$$p_f(\lambda) = p_{M_{f,\mathbf{b}}}(\lambda).$$

Zanotujmy następujący fakt:

Stwierdzenie 11.6. *Niech X będzie przestrzenią liniową wymiaru $n < \infty$ nad ciałem $\mathbb{K}$ i $f \in L(X)$. Następujące warunki są równoważne:*

(i) $\operatorname{im} f \subsetneq X$,

(ii) $\ker f \neq \{0\}$,

(iii) $\det f = 0$.

Zadanie 11.6. Udowodnij stw. 11.6.

11.5 Wektory i wartości własne

Definicja 11.7. Załóżmy, że $f \in L(X)$ jest endomorfizmem przestrzeni liniowej X nad ciałem $\mathbb{K}$.

- Powiemy, że skalar $\lambda \in \mathbb{K}$ jest *wartością własną* endomorfizmu f, jeżeli istnieje niezerowy wektor $v \in X$ taki, że
$$f(v) = \lambda v.$$
- Taki wektor v nazywamy *wektorem własnym* endomorfizmu f, *odpowiadającym wartości własnej* λ.
- Jeżeli $\lambda \in \mathbb{K}$ jest wartością własną endomorfizmu f, to zbiór
$$V_\lambda = \{v \in X : f(v) = \lambda v\}$$
nazywamy *podprzestrzenią własną* endomorfizmu f, *odpowiadającą* wartości własnej λ.

Uwaga 11.3. Ponieważ $V_\lambda = \ker(f - \lambda \mathrm{id}_X)$, więc zbiór ten jest podprzestrzenią liniową w X.

Definicja 11.8. *Widmem* endomorfizmu $f \in L(X)$ nazywamy zbiór wszystkich wartości własnych f. Zbiór ten oznaczamy symbolem $\sigma(f)$.

Definicja 11.9. Dowolną macierz $A \in \mathbb{K}^{n,n}$ możemy interpretować jako endomorfizm przestrzeni $\mathbb{K}^n$: $f(\vec{x}) = A\vec{x}$. *Wartość własną*, *wektor własny* i *podprzestrzeń własną* macierzy A definiujemy jako wartość, wektor i podprzestrzeń własne tak zdefiniowanego endomorfizmu f. Tak samo definiujemy widmo macierzy A, które oznaczamy $\sigma(A)$.

Przykład 11.4. Dany jest endomorfizm $F \in L(\mathbb{R}[x]_n)$,
$$F(w)(t) = t \cdot w'(t).$$
Wówczas wielomiany $1, t, t^2, \ldots, t^n$ są wektorami własnymi endomorfizmu f odpowiadające (kolejno) wartościom własnym $0, 1, \ldots, n$.

Przykład 11.5. Niech $A = \begin{bmatrix} -5 & -4 \\ 8 & 7 \end{bmatrix}$. Można obliczyć, że
$$A \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad A \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix} = 3\begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$
Zatem A ma wartości własne -1, 3, a odpowiadające im wektory własne to $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ i $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$.

Przykład 11.6. Jeżeli $A \in \mathbb{K}^{n,n}$ jest macierzą diagonalną,

$$A = \begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{bmatrix},$$

to liczby $a_1, a_2, \ldots, a_n$ są wartościami własnymi macierzy A, a odpowiadające im wektory własne to $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$.

Zauważmy, że zachodzi następujące

Stwierdzenie 11.7. *Niech $f \in L(X)$. Wówczas $\lambda \in \mathbb{K}$ jest wartością własną endomorfizmu f wtedy i tylko wtedy, gdy* $\dim \ker(f - \lambda \mathrm{id}_X) > 0$.

Ze stw. 11.6 wynika teraz, że $\lambda \in \mathbb{K}$ jest wartością własną endomorfizmu f wtedy i tylko, gdy $\det(f - \lambda \mathrm{id}_X) = 0$. Zatem

Twierdzenie 11.8. *Niech* $\dim X < \infty$ *i $f \in L(X)$. Wówczas $\lambda \in \mathbb{K}$ jest wartością własną f wtedy i tylko wtedy, gdy $p_f(\lambda) = 0$.*

Inaczej mówiąc, wartości własne endomorfizmu (macierzy) to pierwiastki jego (jej) wielomianu charakterystycznego.

Przykład 11.7. Wielomian charakterystyczny macierzy $A = \begin{bmatrix} -5 & -4 \\ 8 & 7 \end{bmatrix}$ z przykładów 11.1 i 11.5 to $p_A(\lambda) = (\lambda + 1)(\lambda - 3)$. Zatem macierz ta ma dokładnie dwie wartości własne: -1 oraz 3.

Przykład 11.8. Wyznaczymy wartości własne macierzy $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Jako endomorfizm $\mathbb{R}^2$ jest to obrót o kąt $\pi/2$. Mamy $p_A(\lambda) = \lambda^2 + 1$, więc A nie ma rzeczywistych wartości własnych. Jednak, jeżeli A potraktujemy jako endomorfizm $\mathbb{C}^2$, to $p_A(\lambda) = (\lambda + i)(\lambda - i)$. Wówczas A ma dwie zespolone wartości własne: $\lambda_1 = i$, $\lambda_2 = -i$.

Zadanie 11.7. Wyznacz (zespolone) wektory własne macierzy A z przykładu 11.8.

Stwierdzenie 11.9. *Jeżeli $\lambda_1, \lambda_2, \ldots, \lambda_k$ są różnymi wartościami własnymi endomorfizmu $f \in L(X)$ i $v_j \in X$ oznacza wektor własny odpowiadający wartości własnej λ_j, to wektory $v_1, v_2, \ldots, v_k$ są liniowo niezależne.*

Dowód. Stosujemy indukcję po k. Dla $k = 1$ teza jest oczywista. Załóżmy, że teza zachodzi dla $k-1$ i niech

$$\alpha_1 v_1 + \ldots + \alpha_k v_k = 0. \tag{11.2}$$

Zatem

$$f(\alpha_1 v_1 + \ldots + \alpha_k v_k) = \alpha_1\lambda_1 v_1 + \alpha_2\lambda_2 v_2 + \ldots + \alpha_k\lambda_k v_k = 0$$

Mnożąc (11.2) stronami przez λ_1 dostajemy

$$\alpha_1\lambda_1 v_1 + \alpha_2\lambda_1 v_2 + \ldots + \alpha_k\lambda_1 v_k = 0,$$

więc po odjęciu ostatnich dwóch równości stronami, otrzymujemy

$$\alpha_2(\lambda_2 - \lambda_1)v_2 + \ldots + \alpha_k(\lambda_k - \lambda_1)v_k = 0.$$

Wobec założenia indukcyjnego $\alpha_2 = \ldots = \alpha_k = 0$, więc także $\alpha_1 = 0$ i pokazaliśmy liniową niezależność wektorów $v_1, \ldots, v_k$. □

Jako wnioski ze stw. 11.9 otrzymujemy:

Stwierdzenie 11.10. *Załóżmy, że $f \in L(X)$ i $\lambda_1, \lambda_2 \in \sigma(f)$, przy czym $\lambda_1 \neq \lambda_2$. Wówczas $V_{\lambda_1} \cap V_{\lambda_2} = \{0\}$.*

Dowód. Jeżeli $x \in V_{\lambda_1} \cap V_{\lambda_2}$, to $f(x) = \lambda_1 x = \lambda_2 x$, więc $(\lambda_1 - \lambda_2)x = 0$. Ponieważ $\lambda_1 \neq \lambda_2$, więc $x = 0$. □

Stwierdzenie 11.11. *Załóżmy, że $f \in L(X)$ i $\lambda_1, \ldots, \lambda_k \in \sigma(f)$, przy czym $\lambda_i \neq \lambda_j$ dla $i \neq j$. Wówczas*

$$V_{\lambda_1} + \ldots + V_{\lambda_k} = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_k}.$$

Uwaga 11.4. Zapis $U = U_1 \oplus U_2 \oplus \ldots \oplus U_n$ oznacza, że każdy wektor $u \in U$ ma jednoznaczne przedstawienie postaci $u = u_1 + u_2 + \ldots + u_n$, gdzie $u_j \in U_j$ dla $j = 1, 2, \ldots, n$.

Dowód. Stosujemy indukcję po k. Dla $k = 2$ mamy $V_{\lambda_1} \cap V_{\lambda_2} = \{0\}$, więc $V_{\lambda_1} + V_{\lambda_2} = V_{\lambda_1} \oplus V_{\lambda_2}$. Załóżmy więc, że teza zachodzi dla $k-1$. Mamy

$$(V_{\lambda_1} + \ldots + V_{\lambda_{k-1}}) + V_{\lambda_k} = (V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_{k-1}}) + V_{\lambda_k}.$$

Niech $x \in (V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_{k-1}}) \cap V_{\lambda_k}$. Wówczas $x = \alpha_1 v_1 + \ldots + \alpha_{k-1}v_{k-1}$. Gdy $x \neq 0$, to jest to wektor własny f odpowiadającym wartości własnej λ_k, który przedstawiony został jako kombinacja liniowa wektorów własnych odpowiadających innym wartościom własnym. Uzyskujemy w ten sposób sprzeczność ze stw. 11.9. Zatem $(V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_{k-1}}) \cap V_{\lambda_k} = \{0\}$ i tym samym

$$(V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_{k-1}}) + V_{\lambda_k} = (V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_{k-1}}) \oplus V_{\lambda_k}.$$

□

11.6 Krotność wartości własnych

Jeżeli macierz $A \in \mathbb{K}^{n,n}$ jest diagonalizowalna i podobna do macierzy $D = \operatorname{diag}(\alpha_1, \dots, \alpha_n)$, to

$$p_A(x) = p_D(x) = (x - \alpha_1) \dots (x - \alpha_n),$$

skąd wynika, że krotność każdej wartości własnej macierzy A jako pierwiastka wielomianu charakterystycznego jest równa wymiarowi odpowiedniej podprzestrzeni własnej. Inaczej mówiąc, jeżeli

$$p_A(x) = (x - \lambda_1)^{r_1} \dots (x - \lambda_k)^{r_k}, \qquad \lambda_i \neq \lambda_j \text{ dla } i \neq j, \quad r_1 + \dots + r_k = n,$$

to, gdy macierz A jest diagonalizowalna,

$$\dim V_{\lambda_j} = r_j, \qquad j = 1, 2, \dots, k.$$

Jednak przykład 11.13 pokazuje, że krotność wartości własnej $\lambda \in \mathbb{K}$ jako pierwiastka wielomiany charakterystycznego nie musi być równa wymiarowi podprzestrzeni własnej.

Definicja 11.10. Załóżmy, że $\lambda \in \mathbb{K}$ jest wartością własną macierzy $A \in \mathbb{K}^{n,n}$ (lub endomorfizmu $f \in L(X)$, gdzie X jest przestrzenią liniową nad $\mathbb{K}$).

- *krotność algebraiczna* wartości własnej λ jest to krotność λ jako pierwiastka wielomianu charakterystycznego macierzy A (lub endomorfizmu f),

- *krotność geometryczna* wartości własnej λ jest to wymiar podprzestrzeni własnej V_λ.

11.7 Podprzestrzenie niezmiennicze

Definicja 11.11. Podprzestrzeń $U \subset X$ nazywamy *podprzestrzenią niezmienniczą* endomorfizmu $f \in L(X)$, jeżeli $f(U) \subset U$.

Przykład 11.9. Jeżeli $D \in L(\mathbb{K}[t]_n)$, $D(p) = p'$, to D ma podprzestrzenie niezmiennicze $\operatorname{span}(1, \dots, t^k)$ dla $k = 0, 1, \dots, n$.

Przykład 11.10. Rozważmy endomorfizm $f \in L(\mathbb{R}^5)$, $f(\vec{x}) = A\vec{x}$, gdzie

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

Endomorfizm f ma podprzestrzenie niezmiennicze

$$U = \operatorname{span}(\vec{e}_1, \vec{e}_2), \quad V = \operatorname{span}(\vec{e}_3), \quad W = \operatorname{span}(\vec{e}_4, \vec{e}_5).$$

Natomiast podprzestrzenie własne f to

$$V_5 = \operatorname{span}(\vec{e}_1) \subset U, \quad V_1 = \operatorname{span}(\vec{e}_3) \subset V, \quad V_2 = \operatorname{span}(\vec{e}_4) \subset W.$$

Stwierdzenie 11.12. *Jeżeli λ jest wartością własną endomorfizmu f, to podprzestrzeń własna V_λ jest podprzestrzenią niezmienniczą dla f.*

Zadanie 11.8. Niech $f \in L(X)$. Pokaż, że podprzestrzenie $\operatorname{im} f$ oraz $\ker f$ są niezmiennicze dla f.

Stwierdzenie 11.13. *Jeżeli $U \subset X$ jest podprzestrzenią niezmienniczą endomorfizmu $f \in L(X)$, to wielomian charakterystyczny endomorfizmu $f|_U \in L(U)$ jest dzielnikiem wielomianu charakterystycznego endomorfizmu f.*

Dowód. Niech $u_1, \ldots, u_k$ to baza podprzestrzeni U, $u_{k+1}, \ldots, u_n$ to jej dopełnienie do bazy całej przestrzeni X. W tej bazie macierz f ma postać

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$

wobec czego, z lematu 8.9

$$\begin{aligned} p_f(\lambda) = \det(A - \lambda I_n) &= \det \begin{bmatrix} B - \lambda I_k & C \\ 0 & D - \lambda I_{n-k} \end{bmatrix} \\ &= \det(B - \lambda I_k) \cdot \det(D - \lambda I_{n-k}) \\ &= p_{f|_U}(\lambda) \cdot \det(D - \lambda I_{n-k}) \end{aligned}$$

□

Zadanie 11.9. Pokaż, że krotność geometryczna wartości własnej endomorfizmu $f \in L(X)$ nigdy nie jest większa od jej krotności algebraicznej.

11.8 Twierdzenie Jordana

Definicja 11.12. *Klatka Jordana* jest to macierz postaci

$$J_{k,\lambda} = \begin{bmatrix} \lambda & 1 & & & & \\ & \lambda & 1 & & & \\ & & \lambda & \ddots & & \\ & & & \ddots & 1 & \\ & & & & \lambda & 1 \\ & & & & & \lambda \end{bmatrix} \in \mathbb{K}^{k,k}.$$

W szczególności $J_{1,\lambda} = [\lambda]$.

Macierz Jordana jest to macierz postaci

$$J = \begin{bmatrix} J_{k_1,\lambda_1} & & & \\ & J_{k_2,\lambda_2} & & \\ & & \ddots & \\ & & & J_{k_r,\lambda_r} \end{bmatrix} \in \mathbb{K}^{n,n}, \tag{11.3}$$

gdzie $k_1 + \ldots + k_r = n$.

Powiemy, że baza $w_1, \cdots w_n$ przestrzeni X jest *bazą Jordana* dla endomorfizmu $f \in L(X)$, jeżeli macierz f w tej bazie jest macierzą Jordana.

Zadanie 11.10. Pokaż, że klatka Jordana $J_{k,\lambda}$ dla $k > 1$ nie jest diagonalizowalna.

Twierdzenie 11.14 (Twierdzenie Jordana)**.** *Jeżeli wielomian charakterystyczny endomorfizmu $f \in L(X)$ jest iloczynem czynników liniowych, to w X istnieje baza Jordana dla endomorfizmu f i macierz f w tej bazie jest macierzą Jordana.*

Uwaga 11.5. Jeżeli macierz $A \in \mathbb{K}^{n,n}$ ma wielomian charakterystyczny, który jest iloczynem czynników liniowych, to możemy zastosować twierdzenie Jordana do endomorfizmu $f \in L(\mathbb{K}^n)$ danego wzorem $f(\vec{x}) = A\vec{x}$. Teza twierdzenia Jordana mówi wówczas, że macierz A jest podobna do pewnej macierzy Jordana $J \in \mathbb{K}^{n,n}$.

Uwaga 11.6. W macierzy Jordana dla danego endomorfizmu f (macierzy A), liczba klatek Jordana $J_{k,\lambda}$ dla ustalonej wartości własnej λ jest równa krotności geometrycznej wartości własnej λ, czyli wymiarowi podprzestrzeni własnej V_λ.

Uwaga 11.7. Ponieważ wielomian charakterystyczny jest niezmiennikiem relacji podobieństwa macierzy, więc z postaci wielomianu charakterystycznego wynika, jakie są elementy na przekątnej macierzy Jordana J dla danego endomorfizmu f (macierzy A): są to wszystkie pierwiastki tego wielomianu, z uwzględnieniem ich krotności.

Dowód twierdzenia Jordana. W ramach przygotowania poczynimy kilka spostrzeżeń dotyczących endomorfizmu $h \in L(X)$, o którym zakładamy, że w bazie $w_1, \ldots, w_n$ ma macierz Jordana J, taką jak w (11.3).

Niech $l_1 = 0$, $l_2 = k_1$, $l_3 = k_1 + k_2$, $\ldots$ $l_r = k_1 + \ldots + k_{r-1}$, $l_{r+1} = k_1 + \ldots + k_r = n$. Wówczas:

1. $\mathrm{span}(w_{l_s+1}, \ldots, w_{l_s+k_s})$ to podprzestrzeń niezmiennicza h, odpowiadająca blokowi J_{λ_s,k_s} macierzy J.

2. w_{l_s+1} to wektor własny h odpowiadający wartości własnej λ_s, w szczególności $w_{l_s+1} \in \operatorname{im} h$.

3. Jeżeli $\lambda_s = 0$, to $w_{l_s+1} \in \ker h$. Ponadto $\ker h = \operatorname{span}\{w_{l_s+1} : \lambda_s = 0\}$.

4. $h(w_{l_s+j}) = w_{l_s+j-1} + \lambda_s w_{l_s+j}$ dla $j = 2, 3, \ldots, k_s$ (wynika to z postaci klatki Jordana J_{k_s,λ_s}).

5. Podprzestrzeń $\operatorname{im} h$ jest rozpięta przez kombinacje liniowe wektorów w_j, a współczynniki każdej takiej kombinacji to wyrazy pewnej niezerowej kolumny macierzy J. Dla $\lambda_s = 0$ wektory $w_{l_s+k_s}$ (czyli ostatnie elementy ciągu wektorów bazowych $(w_{l_s+1}, \ldots w_{l_s+k_s})$ odpowiadającego klatce Jordana J_{k_s,λ_s}) jako jedyne nie występują w tych kombinacjach liniowych. Jest ich $n - \dim \operatorname{im} h = \dim \ker h$ i są one liniowo niezależne, więc

$$X = \operatorname{im} h \oplus \operatorname{span}\{w_{l_s+k_s} : \lambda_s = 0\}.$$

Dowodząc twierdzenie Jordana, stosujemy indukcję po $n = \dim X$. Jeżeli $n = 1$ to twierdzenie jest oczywiste, gdyż każda macierz 1×1 jest macierzą Jordana.

Załóżmy, że twierdzenie zachodzi dla dowolnej przestrzeni wymiaru mniejszego niż n.

I. Rozważmy sytuację, gdy $\dim \operatorname{im} f < n$. Niech

$$g = f|_{\operatorname{im} f} \in L(\operatorname{im} f).$$

Wielomian charakterystyczny p_g jest dzielnikiem wielomianu p_f, na mocy stwierdzenia 11.13. Wobec tego p_g jest iloczynem czynników liniowych i możemy stosować założenie indukcyjne do endomorfizmu g.

W podprzestrzeni $\operatorname{im} f$ istnieje baza Jordana dla g: $w_1, \ldots, w_m$, $m = \dim \operatorname{im} f$.

Jeżeli $\operatorname{im} f \oplus \ker f = X$, to bazę $w_1, \ldots w_m$ rozszerzamy o dowolną bazę $\ker f$, otrzymując bazę Jordana dla f.

W przeciwnym przypadku $\operatorname{im} f \cap \ker f \neq \{0\}$. Mamy też

$$\ker g = \ker f \cap \operatorname{im} f.$$

Niech $t = \dim \ker g$. Możemy przyjąć, że pierwsze t klatek Jordana dla g odpowiada wartości własnej 0, czyli $\lambda_1 = \ldots = \lambda_t = 0$. Wtedy

$$\ker g = \operatorname{span}(w_{l_1+1}, \ldots, w_{l_t+1}).$$

Ponadto, $w_{l_1+k_1}, \ldots, w_{l_t+k_t} \in \operatorname{im} f$, więc istnieją wektory $v_1, \ldots, v_t \in X$ takie, że

$$f(v_1) = w_{l_1+k_1}, \ldots, f(v_t) = w_{l_t+k_t}.$$

Układ $w_{l_1+1}, \ldots, w_{l_t+1}$ (bazę $\ker g$) uzupełniamy wektorami $u_1, \ldots, u_{n-m-t}$ do bazy $\ker f$.

Pokażemy, że układ wektorów

$$w_1, \ldots, w_m, v_1, \ldots, v_t, u_1, \ldots, u_{n-m-t} \tag{11.4}$$

jest liniowo niezależny, czyli jest bazą X. Załóżmy, że

$$\alpha_1 w_1 + \ldots \alpha_m w_m + \beta_1 v_1 + \ldots + \beta_m v_t + \gamma_1 u_1 + \ldots + \gamma_{n-m-t} u_{n-m-t} = 0.$$

Na obie strony tej równości działamy endomorfizmem f, pamiętając, że $f|_{\operatorname{im} f} = g$, $f(v_s) = w_{l_s+k_s}$, $f(u_i) = 0$. Dostajemy

$$\alpha_1 g(w_1) + \ldots + \alpha_m g(w_m) = -\beta_1 w_{l_1+k_1} - \ldots - \beta_t w_{l_t+k_t}.$$

Lewa strona tej równości to wektor z $\operatorname{im} g$, co, wobec spostrzeżenia 5. powyżej, oznacza, że jest ona równa 0. Zatem $\beta_1 = \ldots = \beta_s = 0$. Mamy więc

$$\alpha_1 w_1 + \ldots \alpha_m w_m + \gamma_1 u_1 + \ldots + \gamma_{n-m-t} u_{n-m-t} = 0,$$

więc $u = \gamma_1 u_1 + \ldots + \gamma_{n-m-t} u_{n-m-t} \in \operatorname{im} f \cap \ker f$. Wobec wyboru wektorów u_i jako uzupełnienia bazy $\operatorname{im} f \cap \ker f$ do bazy $\ker f$ oznacza to, że $u = 0$, więc $\gamma_1 = \ldots = \gamma_{n-m-t} = 0$. Teraz z liniowej niezależności $w_1, \ldots, w_m$ dostajemy $\alpha_1 = \ldots = \alpha_m = 0$.

Układ (11.4) ustawiamy w następującej kolejności:

$$u_1, \ldots, u_{n-m-t}, \quad w_{l_1+1}, \ldots, w_{l_1+k_1}, v_1, \quad w_{l_2+1}, \ldots w_{l_2+k_2}, v_2, \quad \ldots,$$
$$w_{l_t+1}, \ldots, w_{l_t+k_t}, v_t, \quad w_{l_{t+1}+1}, \ldots, w_m.$$

Wobec sposobu określenia wektorów v_s i u_j, jest to baza Jordana dla endomorfizmu f.

II. Jeżeli $\dim \operatorname{im} f = n$, postępujemy następująco: niech λ będzie dowolną wartością własną f (jest $\lambda \neq 0$) i niech $F = f - \lambda \mathrm{id}_X$. Do F stosujemy część **I** dowodu. Otrzymana baza Jordana dla F jest też bazą Jordana dla $f = F + \lambda \mathrm{id}_X$. Macierz Jordana dla f otrzymujemy z macierzy Jordana dla F dodając do niej λI_n. $\square$

Odnotujmy jeszcze następujące wnioski z twierdzenia Jordana:

Wniosek 11.15. *Jeżeli X jest przestrzenią liniową nad $\mathbb{C}$ i $f \in L(X)$, to dla f istnieje baza Jordana.*

Wniosek 11.16. *Jeżeli $A \in \mathbb{K}^{n,n}$ i wielomian charakterystyczny $p_A(\lambda)$ jest iloczynem czynników liniowych, to macierz A jest podobna do pewnej macierzy Jordana.*

Dowód. Należy wyznaczyć bazę Jordana $(\vec{u}_1, \ldots, \vec{u}_n)$ i odpowiadającą jej macierz Jordana J dla endomorfizmu $f \in L(\mathbb{K}^n)$ danego wzorem $f(\vec{x}) = A\vec{x}$. Wtedy, dla $U = [\vec{u}_1, \ldots, \vec{u}_n]$

$$A = UJU^{-1}.$$

□

Przykład 11.11. Wyznaczymy macierz Jordana podobną do macierzy

$$A = \begin{bmatrix} 3 & -1 & 0 & 1 \\ 0 & 3 & 4 & 4 \\ 0 & 0 & -5 & -8 \\ 0 & 0 & 4 & 7 \end{bmatrix}.$$

Obliczamy wielomian charakterystyczny: $p_A(x) = (x-3)^3(x+1)$. Jest on czynnikiem iloczynów liniowych, więc zachodzi teza twierdzenia Jordana, czyli istnieje macierz Jordana J podobna do macierzy A.

Ponadto, macierz A ma 2 wartości własne: $\lambda_1 = 3$ oraz $\lambda_2 = -1$. Krotność algebraiczna wartości własnej 3 jest równa 3, a krotność algebraiczna wartości własnej -1 wynosi 1, więc jesteśmy w stanie określić, jak wygląda przekątna macierzy J:

$$J = \begin{bmatrix} 3 & ? & 0 & 0 \\ 0 & 3 & ? & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

Wyznaczamy podprzestrzenie własne:

$$V_3 = \ker(A - 3I) = \ker \begin{bmatrix} 0 & -1 & 0 & 1 \\ 0 & 0 & 4 & 4 \\ 0 & 0 & -8 & -8 \\ 0 & 0 & 4 & 8 \end{bmatrix} = \operatorname{span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ 1 \end{bmatrix} \right),$$

$$V_{-1} = \ker(A + I) = \ker \begin{bmatrix} 4 & -1 & 0 & 1 \\ 0 & 4 & 4 & 4 \\ 0 & 0 & -4 & -8 \\ 0 & 0 & 4 & 4 \end{bmatrix} = \operatorname{span}\left(\begin{bmatrix} 0 \\ 1 \\ -2 \\ 1 \end{bmatrix} \right).$$

Zatem $\dim V_3 = 2$ oraz $\dim V_{-1} = 1$. Są więc 2 klatki Jordana dla wartości własnej 3: $J_{1,3}$ oraz $J_{2,3}$, jak i jedna klatka Jordana dla wartości własnej -1: $J_{1,-1}$. Zatem

$$J = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

11.9 Diagonalizacja

Definicja 11.13. Powiemy, że macierz $A \in \mathbb{K}^{n,n}$ jest *diagonalizowalna*, jeżeli jest ona podobna do pewnej macierzy diagonalnej.

Możemy rozważać endomorfizmy diagonalizowalne:

Definicja 11.14. X jest przestrzenią liniową skończonego wymiaru nad ciałem $\mathbb{K}$. Powiemy, że endomorfizm $f \in L(X)$ jest diagonalizowalny, jeżeli w pewnej bazie f ma macierz diagonalną.

Zajmiemy się teraz charakteryzacją macierzy i endomorfizmów diagonalizowalnych. Najpierw pokażemy kilka faktów pomocniczych.

Twierdzenie 11.17. *X jest przestrzenią liniową wymiaru n nad ciałem $\mathbb{K}$, $F \in L(X)$ i $p_F \in \mathbb{K}[\lambda]_n$ jest iloczynem czynników stopnia 1. Wówczas endomorfizm F jest diagonalizowalny wtedy i tylko wtedy, gdy krotności algebraiczne wszystkich jego wartości własnych są równe ich krotnościom geometrycznym.*

Dowód. Załóżmy, że F jest diagonalizowalny i w bazie $u_1, \ldots, u_n$ ma macierz diagonalną

$$M = \begin{bmatrix} \lambda_1 I_{k_1} & & & \\ & \lambda_2 I_{k_2} & & \\ & & \ddots & \\ & & & \lambda_r I_{k_r} \end{bmatrix},$$

gdzie $\lambda_i \neq \lambda_j$ dla $i \neq j$, $k_1 + \ldots + k_r = n$. Wtedy $\lambda_1, \ldots, \lambda_r$ to różne wartości własne F i $p_F(\lambda) = \prod_{s=1}^{r} (\lambda_s - \lambda)^{k_s}$. Zatem krotność algebraiczne wartości własnej λ_s jest równa k_s. Ponadto, podprzestrzeń własna odpowiadająca wartości własnej λ_s to $V_{\lambda_s} = \ker(F - \lambda_s \cdot \mathrm{id}_X)$ i z postaci macierzy M wynika, że $\dim V_{\lambda_s} = k_s$. Zatem krotność geometryczna wartości własnej λ_s też jest równa k_s.

Załóżmy teraz, że $\lambda_1, \ldots, \lambda_s$ to wszystkie różne wartości własne F i ich krotności algebraiczne są równe krotnościom geometrycznym. Wówczas $p_F(\lambda) = \prod_{s=1}^{r} (\lambda_s - \lambda)^{k_s}$, więc $k_1 + \ldots + k_s = n$. Ponadto

$$\dim V_{\lambda_s} = \dim(\ker(f - \lambda_s \cdot \mathrm{id}_X)) = k_s.$$

Niech układ $v_1^s, \ldots v_{k_2}^s \in X$ będzie bazą przestrzeni V_{lambda_s}. Wówczas układ $v_1^1, \ldots v_{k_1}^s, \ldots, v_1^r, \ldots v_{k_2}^r$ jest liniowo niezależny (co wynika stw. 11.9) i składa się z $n = \dim X$ wektorów, jest więc bazą przestrzeni X, w której F ma macierz diagonalną M. $\square$

Z powyższego twierdzenia można bez trudu wywnioskować następujące

Stwierdzenie 11.18. *Załóżmy, że* $\dim X = n < \infty$ *i* $F \in L(X)$*. Niech* $\lambda_1, \ldots, \lambda_r$ *oznaczają wszystkie wartości własne endomorfizmu* F*, przy czym* $\lambda_i \neq \lambda_j$ *dla* $i \neq j$ *i niech* V_{λ_j}*,* $j = 1, \ldots, r$ *to podprzestrzenie własne odpowiadające tym wartościom własnym. Wówczas endomorfizm* F *jest diagonalizowalny wtedy i tylko wtedy, gdy*

$$\sum_{j=1}^{k} \dim V_{\lambda_j} = n.$$

Uwaga 11.8. Ponieważ każdą macierz $A \in \mathbb{K}^{n,n}$ możemy traktować jako endomorfizm przestrzeni $\mathbb{K}^n$, więc twierdzenie 11.17 można także stosować w celu rozstrzygania, czy dana macierz jest diagonalizowalna.

Przykład 11.12. Zbadamy diagonalizowalność endomorfizmu $f \in L(\mathbb{R}^2)$, $f(\vec{x}) = A\vec{x}$, zadanego macierzą $A = \begin{bmatrix} -5 & -4 \\ 8 & 7 \end{bmatrix}$. Z wcześniejszych przykładów wiemy już, że endomorfizm f (jak również macierz A) ma wartości własne $\lambda_1 = -1$ i $\lambda_2 = 3$, a odpowiadające im wektory własne to

$$\vec{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad \vec{v}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

Zatem f ma 2 podprzestrzenie własne, $V_{-1} = \operatorname{span}(\vec{v}_1)$, $V_3 = \operatorname{span}(\vec{v}_2)$, obie wymiaru 1. Z tw. 11.17 wynika, że endomorfizm f jest diagonalizowalny i macierz A (czyli macierz f w bazie $\vec{e}_1, \vec{e}_2$) jest podobna do macierzy $\begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}$. Zauważmy, że

$$\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} -5 & -4 \\ 8 & 7 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

skąd dostajemy

$$\begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}^{-1} \cdot \begin{bmatrix} -5 & -4 \\ 8 & 7 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}.$$

W iloczynie powyżej

- macierz $\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$ zmienia bazę z $\vec{v}_1, \vec{v}_2$ na $\vec{e}_1, \vec{e}_2$,
- macierz $\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}^{-1}$ zmienia bazę $\vec{e}_1, \vec{e}_2$ z powrotem na bazę $\vec{v}_1, \vec{v}_2$.

Iloczyn macierzy daje więc macierz f w bazie $\vec{v}_1, \vec{v}_2$. Z drugiej strony, wyznaczając wartości i podprzestrzenie własne f pokazaliśmy, że macierz f w bazie $\vec{v}_1$, $\vec{v}_2$ jest równa $\begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}$. Macierz $\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}^{-1}$ jest dla endomorfizmu f macierzą zmiany bazy $\vec{e}_1, \vec{e}_2$ na bazę własną.

Przykład 11.13. Zbadamy, czy macierz $B = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ jest diagonalizowalna. W pierwszym kroku wyznaczamy wartości własne, znajdując pierwiastki wielomianu charakterystycznego:

$$p_B(x) = \det(B - xI_2) = \det \begin{bmatrix} 3-x & 1 \\ 0 & 3-x \end{bmatrix} = (3-x)^2.$$

Macierz B ma jedną wartość własną $\lambda = 3$.

W kolejnym kroku wyznaczamy bazę podprzestrzeni własnej odpowiadającej wartości własnej 3. Jest to podprzestrzeń

$$V_3 = \ker(B - 3I_2) = \ker \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \operatorname{span}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right).$$

Zatem $\dim V_3 = 1$. Z twierdzenia 11.17 wynika, że macierz B nie jest diagonalizowalna.

Zadanie 11.11. Pokaż, że macierz $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ jest diagonalizowalna nad $\mathbb{C}$ (jako element przestrzeni $\mathbb{C}^{2,2}$) i nie jest diagonalizowalna nad $\mathbb{R}$.

11.10 Macierze ortogonalne i unitarne

Definicja 11.15. Powiemy, że macierz $U \in \mathbb{R}^{n,n}$ jest *ortogonalna*, jeżeli $U^T U = I_n$.

Powiemy, że macierz $U \in \mathbb{C}^{n,n}$ jest *unitarna*, jeżeli $U^H U = I_n$.

Uwaga 11.9. Z definicji powyższej wynika, że każda macierz ortogonalna lub unitarna U jest nieosobliwa i $U^{-1} = U^T$ (gdy U jest ortogonalna) lub $U^{-1} = U^H$ (gdy U jest unitarna).

Przykład 11.14. Macierz

$$U = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \in \mathbb{K}^{4,4}$$

jest ortogonalna (dla $\mathbb{K} = \mathbb{R}$) i unitarna (dla $\mathbb{K} = \mathbb{C}$).

Przykład 11.15. Macierz $A \in \mathbb{R}^{2,2}$ obrotu o kąt θ jest ortogonalna.

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

i

$$\begin{aligned} A^T A &= \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\theta + \sin^2\theta & -\cos\theta\sin\theta + \sin\theta\cos\theta \\ -\sin\theta\cos\theta + \cos\theta\sin\theta & \sin^2\theta + \cos^2\theta \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

Przykład 11.16. Niech $z_1, \ldots, z_n \in \mathbb{C}$ i $|z_j| = 1$ dla $j = 1, \ldots, n$. Wówczas macierz $\operatorname{diag}(z_1, \ldots, z_n) \in \mathbb{C}^{n,n}$ jest unitarna, $U^H = \operatorname{diag}(\overline{z}_1, \ldots, \overline{z}_n)$.

Stwierdzenie 11.19. *Macierz $U = [\vec{u}_1, \ldots, \vec{u}_n] \in \mathbb{K}^{n,n}$, $\vec{u}_j \in \mathbb{K}^n$, jest ortogonalna/unitarna wtedy i tylko wtedy, gdy jej kolumny $\vec{u}_j$ tworzą bazę ortonormalną w $\mathbb{K}^n$ ze standardowym iloczynem skalarnym $\langle \vec{x}, \vec{y} \rangle = \vec{x}^H \vec{y}$.*

Dowód. Macierz $U^H U$ to macierz Grama układu $\vec{u}_1, \ldots, \vec{u}_n$. Mamy

$$U^H U = [\langle \vec{u}_i, \vec{u}_j \rangle]_{i,j=1}^n = [\vec{u}_i^H \vec{u}_j]_{i,j=1}^n = I_n.$$

□

Zadanie 11.12. Macierz $U \in \mathbb{K}^{n,n}$ jest ortogonalna/unitarna, $\langle \cdot, \cdot \rangle$ to standardowy iloczyn skalarny na $\mathbb{K}^n$. Pokaż, że dla dowolnych wektorów $\vec{x}, \vec{y} \in \mathbb{K}^n$ zachodzi

$$\langle \vec{x}, \vec{y} \rangle = \langle U\vec{x}, U\vec{y} \rangle.$$

Uwaga 11.10. Na przestrzeni $\mathbb{K}^n$ rozważamy standardowy iloczyn skalarny i, $U \in \mathbb{K}^{n,n}$ jest macierzą unitarną / ortogonalną. Rozważamy endomorfizm $f \in L(\mathbb{K}^n)$ zadany wzorem $f(\vec{x}) = U\vec{x}$. Wówczas f zachowuje iloczyny skalarne i normy wektorów – mówimy, że f jest *izometrią* przestrzeni euklidesowej/unitarnej $\mathbb{K}^n$.

Zadanie 11.13. Załóżmy, że macierz $U \in \mathbb{C}^{n,n}$ jest unitarna. Pokaż, że

(a) jeżeli $\lambda_1 \in \mathbb{C}$ jest wartością własną U, to $|\lambda_1| = 1$,

(b) jeżeli λ_1 i λ_2 są różnymi wartościami własnymi, a $\vec{v}_1$ i $\vec{v}_2$ to odpowiadające im wektory własne, to $\vec{v}_1 \perp \vec{v}_2$ (czyli $\vec{v}_1^H \vec{v}_2 = 0$).

Zadanie 11.14. Załóżmy, że macierz $U \in \mathbb{C}^{n,n}$ jest unitarna. Udowodnij, że istnieje macierz diagonalna

$$D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^{n,n}, \quad \text{gdzie } |\lambda_j| \text{ dla } j = 1, \ldots, n,$$

oraz macierz unitarna $V \in \mathbb{C}^{n,n}$ takie, że

$$U = V^H D V.$$

Uwaga 11.11. Inaczej mówiąc, macierz unitarna U jest diagonalizowalna nad $\mathbb{C}$, jej wartości własne leżą na okręgu jednostkowym, i istnieje baza własna dla U, która jest ortonormalna.

11.11 Diagonalizacja macierzy hermitowskich

W tym podrozdziale $\mathbb{K}$ oznacza ciało $\mathbb{R}$ lub $\mathbb{C}$. Zakładamy także, że w przestrzeni $\mathbb{K}^n$ mamy standardowy iloczyn skalarny

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^H \vec{y}.$$

Jak pamiętamy, macierz $A \in \mathbb{K}^{n,n}$ jest hermitowska, jeżeli $A^H = A$. W przypadku, gdy $\mathbb{K} = \mathbb{R}$, są to macierze symetryczne, czyli takie, że $A^T = A$.

Stwierdzenie 11.20. *Załóżmy, że $A \in \mathbb{K}^{n,n}$ jest macierzą hermitowską (gdy $\mathbb{K} = \mathbb{C}$) lub symetryczną (gdy $\mathbb{K} = \mathbb{R}$). Wówczas*

(i) dla dowolnych wektorów $\vec{x}, \vec{y} \in \mathbb{K}^n$ zachodzi

$$\langle A\vec{x}, \vec{y} \rangle = \langle \vec{x}, A\vec{y} \rangle,$$

(ii) jeżeli λ_1 jest wartością własną macierzy A, to $\lambda_1 \in \mathbb{R}$ (także w sytuacji, gdy $\mathbb{K} = \mathbb{C}$),

(iii) jeżeli λ_1 i λ_2 to dwie różne wartości własne macierzy A, to odpowiadające im podprzestrzenie własne V_{λ_1} i V_{λ_2} są ortogonalne, czyli dla $\vec{v}_1 \in V_{\lambda_1}$, $\vec{v}_2 \in V_{\lambda_2}$ mamy $\langle \vec{v}_1, \vec{v}_2 \rangle = 0$.

Dowód. (i): sprawdzamy, że $\langle A\vec{x}, \vec{y} \rangle = (A\vec{x})^H \vec{y} = \vec{x}^H A\vec{y} = \langle \vec{x}, A\vec{y} \rangle$

(ii): niech $\vec{x}_1 \in \mathbb{K}^n$ będzie takim wektorem, że $\|\vec{x}_1\| = 1$ i $A\vec{x}_1 = \lambda_1 \vec{x}_1$. Wtedy

$$\begin{aligned}\lambda_1 = \lambda_1 \cdot \langle \vec{x}_1, \vec{x}_1 \rangle = \langle \vec{x}_1, \lambda_1 \vec{x}_1 \rangle = \langle \vec{x}_1, A\vec{x}_1 \rangle \\ = \langle A\vec{x}_1, \vec{x}_1 \rangle = \langle \lambda_1 \vec{x}_1, \vec{x}_1 \rangle = \bar{\lambda}_1 \langle \vec{x}_1, \vec{x}_1 \rangle = \bar{\lambda}_1,\end{aligned}$$

czyli $\lambda_1 \in \mathbb{R}$.

(iii): możemy założyć, że $\lambda_1 \neq 0$. Wobec podpunktów (i) i (ii):

$$\begin{aligned}\langle \vec{v}_1, \vec{v}_2 \rangle = \frac{1}{\lambda_1}\langle \lambda_1 \vec{v}_1, \vec{v}_2 \rangle = \frac{1}{\lambda_1}\langle A\vec{v}_1, v_2 \rangle \\ = \frac{1}{\lambda_1}\langle v_1, A\vec{v}_2 \rangle = \frac{1}{\lambda_1}\langle v_1, \lambda_2 \vec{v}_2 \rangle = \frac{\lambda_2}{\lambda_1}\langle \vec{v}_1, \vec{v}_2 \rangle.\end{aligned}$$

Zatem $\left(1 - \frac{\lambda_2}{\lambda_1}\right)\langle \vec{v}_1, \vec{v}_2 \rangle = 0$. Ponieważ $\lambda_1 \neq \lambda_2$, więc $\langle \vec{v}_1, \vec{v}_2 \rangle = 0$. □

Twierdzenie 11.21. *Dla dowolnej macierz hermitowskiej $A = A^H \in \mathbb{C}^{n,n}$ istnieje macierz ortogonalna/unitarna $U \in \mathbb{K}^{n,n}$ oraz macierz diagonalna $D \in \mathbb{R}^{n,n}$ takie, że*

$$A = U^H D U = U^{-1} D U.$$

Uwaga 11.12. Inaczej mówiąc, macierz A jest podobna do macierzy diagonalnej, a macierz podobieństwa jest unitarna. W szczególności, w $\mathbb{C}^n$ ze standardowym iloczynem skalarnym istnieje baza ortonormalna złożona z wektorów własnych macierzy A.

Dowód. Z tw. Jordana (11.14)wynika, że istnieje macierz $J \in \mathbb{C}^{n,n}$ podobna do macierzy A

$$J = \begin{bmatrix} J_{k_1,\lambda_1} & & & \\ & J_{k_2,\lambda_2} & & \\ & & \ddots & \\ & & & J_{k_r,\lambda_r} \end{bmatrix}.$$

Wobec stw. 11.20 $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$. Pokażemy, że wszystkie klatki Jordana w macierzy J mają rozmiar 1×1. Niech $\vec{u}_1, \ldots, \vec{u}_n$ to baza, w której endomorfizm $F \in L(\mathbb{C}^n)$, $F(\vec{x}) = A\vec{x}$, ma macierz J i $\vec{u}_i, \ldots, \vec{u}_{i+k-1}$ to podukład tej bazy odpowiadający klatce $J_{k,\lambda}$. Załóżmy, że $k > 1$. Wtedy $A\vec{u}_i = \lambda\vec{u}_i$ i $A\vec{u}_{i+1} = \vec{u}_i + \lambda\vec{u}_{i+1}$. Ze stw. 11.20 $\langle A\vec{u}_i, \vec{u}_{i+1} \rangle = \langle \vec{u}_i, A\vec{u}_{i+1} \rangle$, więc $\langle \lambda\vec{u}_i, \vec{u}_{i+1} \rangle = \langle \vec{u}_i, \vec{u}_i + \lambda\vec{u}_{i+1} \rangle$, skąd

$$\lambda\langle \vec{u}_i, \vec{u}_{i+1} \rangle = \langle \vec{u}_i, \vec{u}_i \rangle + \lambda\langle \vec{u}_i, \vec{u}_{i+1} \rangle.$$

Zatem $\langle \vec{u}_i, \vec{u}_i \rangle = 0$, czyli $\vec{u}_i = 0$, co jest niemożliwe. Zatem $k = 1$. Skoro wszystkie klatki Jordana mają rozmiar 1×1, to macierz A jest diagonalizowalna i jest ona podobna do macierzy

$$D = \begin{bmatrix} \lambda_1 I_{l_1} & & & \\ & \lambda_2 I_{l_2} & & \\ & & \ddots & \\ & & & \lambda_p I_{l_p} \end{bmatrix},$$

gdzie $\lambda_i \neq \lambda_j$ dla $i \neq j$. Niech $\vec{v}_1^{(i)}, \ldots \vec{v}_{k_i}^{(i)}$ to ortonormalna baza podprzestrzeni własnej V_{λ_i}. Wówczas $(\vec{u}_1, \ldots, \vec{u}_n) = (\vec{v}_1^{(1)}, \ldots \vec{v}_{k_1}^{(1)}, \ldots, \vec{v}_1^{(p)}, \ldots \vec{v}_{k_p}^{(p)})$ to baza ortonormalna $\mathbb{C}^n$, w której F ma macierz D. Jeżeli $U = [\vec{u}_1, \ldots, \vec{u}_n]^{-1}$ (czyli jest to macierz zmiany bazy $\vec{e}_1, \ldots, \vec{e}_n$ na $\vec{u}_1, \ldots, \vec{u}_n$), to

$$A = U^{-1}DU = U^H DU.$$

□

Twierdzenie 11.22. *Dla dowolnej macierz symetrycznej $A = A^H \in \mathbb{R}^{n,n}$ istnieje macierz ortogonalna $U \in \mathbb{K}^{n,n}$ oraz macierz diagonalna $D \in \mathbb{R}^{n,n}$ takie, że*

$$A = U^H DU = U^{-1}DU.$$

Dowód. Macierz A jest także hermitowska, więc z tw. 11.21 wynika, że p_A jest iloczynem czynników stopnia 1. Ze stw. 11.20 wynika, że czynniki te są rzeczywiste. Dalej dowód przebiega tak samo jak w przypadku tw. 11.21. □

11.12 Określoność macierzy symetrycznych

Definicja 11.16. Macierz $A \in \mathbb{R}^{n,n}$ jest symetryczna. Powiemy, że A jest

(i) *dodatnio określona*, jeżeli dla każdego $\vec{x} \in \mathbb{R}^n \backslash \{0\}$ zachodzi $\vec{x}^T A\vec{x} > 0$ (ozn. $A > 0$);

(ii) *ujemnie określona*, jeżeli dla każdego $\vec{x} \in \mathbb{R}^n \backslash \{0\}$ zachodzi $\vec{x}^T A\vec{x} < 0$ (ozn. $A < 0$);

(iii) *nieujemnie określona*, jeżeli dla każdego $\vec{x} \in \mathbb{R}^n$ zachodzi $\vec{x}^T A\vec{x} \geqslant 0$ (ozn. $A \geqslant 0$);

(iv) *niedodatnio określona*, jeżeli dla każdego $\vec{x} \in \mathbb{R}^n$ zachodzi $\vec{x}^T A\vec{x} \leqslant 0$ (ozn. $A \leqslant 0$);

(v) *nieokreślona*, jeżeli istnieją $\vec{x}, \vec{y} \in \mathbb{R}^n$ takie, że $\vec{x}^T A\vec{x} > 0$ i $\vec{y}^T A\vec{y} < 0$.

Stwierdzenie 11.23. *Macierz $A \in \mathbb{R}^n$ jest symetryczna i $\lambda_1, \ldots, \lambda_n$ to jej wartości własne (z uwzględnieniem ich krotności). Wówczas*

(i) A jest dodatnio określona wtedy i tylko wtedy, gdy $\lambda_1, \ldots, \lambda_n > 0$;

(ii) A jest ujemnie określona wtedy i tylko wtedy, gdy $\lambda_1, \ldots, \lambda_n < 0$;

(iii) A jest nieujemnie określona wtedy i tylko wtedy, gdy $\lambda_1, \ldots, \lambda_n \geqslant 0$;

(iv) *A jest nieujemnie określona wtedy i tylko wtedy, gdy* $\lambda_1, \ldots, \lambda_n \leqslant 0$*;*

(v) *A jest nieokreślona, wtedy i tylko wtedy, gdy dla pewnych* i, j $\lambda_i > 0$ *i* $\lambda_j < 0$.

.

Dowód. Udowodnimy podpunkt (i). Dowody (ii) - (iv) są podobne, a (v) wynika z (i) - (iv).

Załóżmy, że $A > 0$ i $\vec{v}_i$ jest wektorem własnym dla wartości własnej λ_i. Wówczas $\vec{v}_i^T A\vec{v}_i = \vec{v}_i(\lambda_i\vec{v}_i) = \lambda_i(\vec{v}_i^T\vec{v}_i) > 0$, więc $\lambda_i > 0$.

Załóżmy, że wszystkie wartości własne A są dodatnie. Z tw. 11.22, istnieje baza ortonormalna $\vec{v}_1, \ldots, \vec{v}_n$ taka, że $A\vec{v}_i = \lambda_i\vec{v}_i$. Jeżeli $\vec{x} \in \mathbb{R}^n\backslash\{0\}$, to istnieją $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$, nie wszystkie równe 0, takie, że $\vec{x} = \alpha_1\vec{v}_1 + \ldots + \alpha_n\vec{v}_n$. Wówczas

$$
\begin{aligned}
\vec{x}^T A\vec{x} &= (\alpha_1\vec{v}_1^T + \ldots + \alpha_n\vec{v}_n^T) \cdot A \cdot (\alpha_1\vec{v}_1 + \ldots + \alpha_n\vec{v}_n) \\
&= (\alpha_1\vec{v}_1^T + \ldots + \alpha_n\vec{v}_n^T) \cdot (\lambda_1\alpha_1\vec{v}_1 + \ldots + \lambda_n\alpha_n\vec{v}_n) \\
&= \lambda_1\alpha_1^2 \cdot \vec{x}_1^T\vec{x}_1 + \ldots + \lambda_n\alpha_n^2 \cdot \vec{x}_n^T\vec{x}_n \\
&> 0.
\end{aligned}
$$

□

Jeżeli chcemy sprawdzić tylko czy macierz jest dodatnio lub ujemnie określona, możemy skorzystać z poniższego kryterium:

Stwierdzenie 11.24 (Kryterium Sylwestera)**.** *Załóżmy, że* $A = A^T \in \mathbb{R}^n$, $A = [a_{i,j}]_{i,j=1}^n$. *Niech* $A_k = [a_{i,j}]_{i,j=1}^k$. *Wówczas*

(i) $A > 0$ *wtedy i tylko wtedy, gdy* $\det_k A_k > 0$ *dla* $k = 1, 2, \ldots, n$*;*

(ii) $A < 0$ *wtedy i tylko wtedy, gdy* $(-1)^k \det_k A_k > 0$ *dla* $k = 1, 2, \ldots, n$.

Dowód tego faktu znajduje się w kolejnym rozdziale (zob. tw. 12.8).

Zadanie 11.15. Załóżmy, że macierz symetryczna $A \in \mathbb{R}^n$ jest dodatnio określona. Pokaż, że istnieje macierz symetryczna i dodatnio określona $B \in \mathbb{R}^n$ taka, że

$$B^2 = A.$$

Rozdział 12

Formy hermitowskie i kwadratowe

W tym rozdziale zakładamy, że $\mathbb{K}$ oznacza ciało $\mathbb{R}$ lub $\mathbb{C}$.

12.1 Formy hermitowskie

Definicja 12.1. Funkcję $\phi : X \times X \to \mathbb{K}$ nazywamy *formą hermitowską* na X, jeżeli

(i) dla dowolnych $x, y_1, y_2 \in X$ oraz $\alpha_1, \alpha_2 \in \mathbb{K}$

$$\phi(x, \alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 \phi(x, y_1) + \alpha_2 \phi(x, y_2),$$

(ii) dla dowolnych $x, y \in X$

$$\phi(x, y) = \overline{\phi(y, x)}.$$

Uwaga 12.1. Widzimy, że forma hermitowska to taki iloczyn skalarny bez warunku dodatniej określoności, czyli nie żądamy, aby $\phi(x, x) > 0$ dla $x \neq 0$. Analogicznie jak w przypadku iloczynu skalarnego

- dla dowolnych $x_1, x_2, y \in X$ oraz $\alpha_1, \alpha_2 \in \mathbb{K}$

$$\phi(\alpha_1 x_1 + \alpha_2 x_2, y) = \bar{\alpha}_1 \phi(x_1, y) + \bar{\alpha}_2 \phi(x_2, y);$$

- jeżeli $\mathbb{K} = \mathbb{R}$, to forma ϕ jest liniowa względem obu argumentów i symetryczna ($\phi(x, y) = \phi(y, x)$). Takie funkcje są nazywane formami dwuliniowymi symetrycznymi. Dla uproszczenie, tutaj będziemy posługiwać się terminem „forma hermitowska” także w przypadku $\mathbb{K} = \mathbb{R}$;

- z warunku (ii) definicji wynika, że $\phi(x,x) \in \mathbb{R}$ dla każdego $x \in X$.

Przykład 12.1. Każdy iloczyn skalarny na przestrzeni liniowej X jest formą hermitowską.

Przykład 12.2. Dla $\vec{x}, \vec{y} \in \mathbb{C}^3$ niech

$$\phi(\vec{x}, \vec{y}) = \bar{x}_1 y_2 + \bar{x}_2 y_1 + 3\bar{x}_3 y_3 - i\bar{x}_2 y_3 + i\bar{x}_3 y_2.$$

Funkcja $\phi : \mathbb{C}^3 \times \mathbb{C}^3 \to \mathbb{C}$ jest formą hermitowską.

Przykład 12.3. Dla $\vec{x}, \vec{y} \in \mathbb{R}^3$ niech

$$\psi(\vec{x}, \vec{y}) = x_1 y_1 - 2x_2 y_2 + 3x_1 y_2 + 3y_1 x_2 - x_1 y_3 - x_3 y_1$$

To jest forma hermitowska na $\mathbb{R}^4$.

Stwierdzenie 12.1. *Jeżeli $A \in \mathbb{K}^{n,n}$ jest macierzą hermitowską, to odwzorowanie $\phi : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ dane wzorem*

$$\phi(\vec{x}, \vec{y}) = \vec{x}^H A \vec{y}$$

jest formą hermitowską na $\mathbb{K}^n$.

Dowód. Sprawdzamy warunki z definicji 12.1:

$$\begin{aligned}
\phi(x, \alpha_1 \vec{y}_1 + \alpha_2 \vec{y}_2) &= \vec{x}^H A(\alpha_1 \vec{y}_1 + \alpha_2 \vec{y}_2) \\
&= \alpha_1 \vec{x}^H A \vec{y}_1 + \alpha_2 \vec{x}^H A \vec{y}_2 = \alpha_1 \phi(\vec{x}, \vec{y}_1) + \alpha_2 \phi(\vec{x}, \vec{y}_2), \\
\overline{\phi(\vec{y}, \vec{x})} &= \overline{\vec{y}^H A \vec{x}} = (\vec{y}^H A \vec{x})^H = \vec{x}^H A^H (\vec{y}^H)^H = \vec{x}^H A \vec{y} \\
&= \phi(\vec{x}, \vec{y}).
\end{aligned}$$

□

Uwaga 12.2. W szczególności, gdy $\mathbb{K} = \mathbb{R}$, każda macierz symetryczna $A \in \mathbb{R}^{n,n}$ zadaje formę hermitowską na $\mathbb{R}^n$:

$$\phi(\vec{x}, \vec{y}) = \vec{x}^T A \vec{y}.$$

Przykład 12.4. Formę ϕ z przykładu 12.2 można zadać wzorem:

$$\phi(\vec{x}, \vec{y}) = \vec{x}^H \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & -i \\ 0 & i & 3 \end{bmatrix} \vec{y}.$$

Przykład 12.5. Formę ψ z przykładu 12.3 można zadać wzorem:

$$\psi(\vec{x}, \vec{y}) = [x_1, x_2, x_3] \cdot \begin{bmatrix} 1 & 3 & -1 \\ 3 & -2 & 0 \\ -1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}.$$

Zwróćmy uwagę na to, że element w i-tym wierszu i j-tej kolumnie macierzy, to współczynnik w wielomianie zadającym ψ, stojący przed $x_i y_j$. Jest on także równy $\phi(\vec{e}_i, \vec{e}_j)$.

12.2 Macierz formy hermitowskiej

Ustalając bazę w przestrzeni liniowej X, każdej formie hermitowskiej na X można przyporządkować pewną macierz hermitowską / symetryczną.

Definicja 12.2. Macierz formy hermitowskiej $\phi : X \times X \to \mathbb{K}$ w bazie $x_1, x_2, \ldots, x_n$ przestrzeni X jest zdefiniowana jako

$$M = [\phi(x_i, x_j)]_{i,j=1}^{n}.$$

Przykład 12.6. Macierze w bazie standardowej formy ϕ i ψ z przykładów 12.2 i 12.3 są podane w przykładach 12.4 i 12.5.

Uwaga 12.3. Z warunku (ii) w definicji formy hermitowskiej wynika, że macierz M jest hermitowska, czyli $M^H = M$ (lub symetryczna w przypadku rzeczywistym)

Stwierdzenie 12.2. *M jest macierzą formy hermitowskiej $\phi : X \times X \to \mathbb{K}$ w bazie $x_1, \ldots, x_n$ wtedy i tylko wtedy, gdy dla dowolnych $x, y \in X$,*

$$x = \sum_{i=1}^{n} \alpha_i x_i, \quad y = \sum_{j=1}^{n} \beta_j x_j,$$

gdzie $\vec{\alpha} = [\alpha_1, \ldots, \alpha_n]^T$, $\vec{\beta} = [\beta_1, \ldots, \beta_n]^T$, zachodzi

$$\phi(x, y) = \vec{\alpha}^H M \vec{\beta}. \tag{12.1}$$

Dowód. Załóżmy, że M jest macierzą formy ϕ. Obliczamy

$$\phi(x, y) = \phi\Big(\sum_i \alpha_i x_i, \sum_j \beta_j x_j\Big) = \sum_{i,j} \bar{\alpha}_i \beta_j \phi(x_i, x_j) = \vec{\alpha}^H M \vec{\beta}.$$

Z drugiej strony, jeżeli zachodzi (12.1), to $\phi(x_i, x_j) = \vec{e}_i^H M \vec{e}_j$ – jest to element w i-tym wierszu i j-tej kolumnie macierzy M. □

Stwierdzenie 12.3. *Jeżeli M jest macierzą formy hermitowskiej $\phi : X \times X \to \mathbb{K}$ w bazie $x_1, \ldots, x_n$, natomiast $y_1, \ldots, y_n$ jest inną bazą przestrzeni X, to macierz formy ϕ w bazie $y_1, \ldots, y_n$ jest równa*

$$C^H M C,$$

gdzie C to macierz zmiany bazy z $x_1, \ldots, x_n$ na $y_1, \ldots, y_n$.

Dowód. Jeżeli $x = \sum_{i=1}^n \xi_i y_i$, $y = \sum_{j=1}^n \eta_j y_j$, to $x = \sum_{i=1}^n \alpha_i x_i$, $y = \sum_{j=1}^n \beta_j x_j$ dla $\vec{\alpha} = C\vec{\xi}$, $\vec{\beta} = C\vec{\eta}$ oraz

$$\phi(x, y) = \vec{\alpha}^H M \vec{\beta} = (C\vec{\xi})^H M (C\vec{\eta}) = \vec{\xi}^H (C^H M C)\vec{\eta}.$$

□

Przykład 12.7. Jeżeli $A = A^H \in \mathbb{K}^{n,n}$, to $\phi(\vec{x}, \vec{y}) = \vec{x}^H A \vec{y}$ jest formą hermitowską na $\mathbb{K}^n$. Zarazem A to macierz formy ϕ w bazie standardowej.

Jako uzupełnienie stwierdzenia 12.2 sformułujemy następującą definicję:

Definicja 12.3. Powiemy, że macierz hermitowskie $A, B \in \mathbb{K}^{n,n}$ są *przystające (kongruentne)*, jeżeli istnieje macierz nieosobliwa $C \in \mathbb{K}^{n,n}$ taka, że

$$B = C^H A C.$$

Uwaga 12.4. Można pokazać, że przystawanie macierzy jest relacją równoważności na zbiorze macierzy hermitowskich. W podrozdziale 12.5 zajmiemy się sposobem określenia klasy abstrakcji tej relacji dla danej macierzy hermitowskiej.

12.3 Formy kwadratowe

Definicja 12.4. Załóżmy, że X jest przestrzenią liniową nad ciałem $\mathbb{K}$ i $\phi : X \times X \to \mathbb{K}$ jest formą hermitowską. Wówczas funkcję

$$\Phi : X \to \mathbb{R} \qquad \Phi(x) = \phi(x, x)$$

nazywamy *formą kwadratową* na X.

Uwaga 12.5. Ponieważ $\phi(x, x) = \overline{\phi(x, x)}$, więc $\Phi(x) \in \mathbb{R}$ także w przypadku, gdy $\mathbb{K} = \mathbb{C}$.

Przykład 12.8. Na $\mathbb{C}^3$ dana jest funkcja

$$\Phi(\vec{x}) = 2\mathrm{Re}(\bar{x}_1 x_2) + 3|x_3|^2 + 2\mathrm{Im}(\bar{x}_2 x_3).$$

Można (i warto) sprawdzić, że $\Phi(\vec{x}) = \phi(\vec{x}, \vec{x})$, gdzie ϕ jest formą hermitowską z przykładu 12.2. Zatem Φ jest formą kwadratową na Φ.

Przykład 12.9. Na $\mathbb{R}^3$ dana jest funkcja

$$\Psi(\vec{x}) = x_1^2 - 2x_2^2 + 6x_1 x_2 - 2x_1 x_3.$$

Jest to forma kwadratowa, gdyż $\Psi(\vec{x}) = \psi(\vec{x}, \vec{x})$ dla formy hermitowskiej ψ z przykładu 12.2.

Przykład 12.10. Jeżeli $(X, \langle\cdot,\cdot\rangle)$ jest przestrzenią z iloczynem skalarnym, to $\Phi(x) = \|x\|^2 = \langle x, x\rangle$ jest formą kwadratową na X zadaną przez formę hermitowską $\phi(x, y) = \langle x, y\rangle$.

Przykład 12.11. Jeżeli $A = A^H \in \mathbb{K}^{n,n}$, to funkcja $\Phi : \mathbb{K}^n \to \mathbb{K}$ zadana wzorem

$$\Phi(\vec{x}) = \vec{x}^H A\vec{x}$$

jest formą kwadratową na $\mathbb{K}^n$. Odpowiada jej forma hermitowska $\phi(\vec{x}, \vec{y}) = \vec{x}^H A\vec{y}$.

12.4 Określoność form

Definicja 12.5. Powiemy, że forma kwadratowa $\Phi : X \to \mathbb{K}$ jest

(i) *dodatnio określona*, jeżeli $\Phi(x) > 0$ dla każdego $x \in X\backslash\{0\}$,

(ii) *ujemnie określona*, jeżeli $\Phi(x) < 0$ dla każdego $x \in X\backslash\{0\}$,

(iii) *nieujemnie określona*, jeżeli $\Phi(x) \geqslant 0$ dla każdego $x \in X$,

(iv) *niedodatnio określona*, jeżeli $\Phi(x) \leqslant 0$ dla każdego $x \in X$,

(v) *nieokreślona*, jeżeli $\Phi(x) > 0$ dla pewnego $x \in X$ oraz $\Phi(y) < 0$ dla pewnego $y \in X$.

Definicja 12.6. Powiemy, że forma hermitowska $\phi : X \times X \to \mathbb{K}$ jest *dodatnio, ujemnie, niedodatnio, nieujemnie określona*, lub *nieokreślona*, jeżeli forma kwadrtowa $\Phi(x) = \phi(x, x)$ jest odpowiednio *dodatnio, ujemnie, niedodatnio, nieujemnie określona*, lub *nieokreślona*.

Uwaga 12.6. Każda dodatnio określona forma hermitowska $\phi : X \times X \to \mathbb{K}$ jest iloczynem skalarnym na X.

Definicja 12.7. Powiemy, że macierz $A = A^H \in \mathbb{K}^{n,n}$ jest *dodatnio, ujemnie, niedodatnio, nieujemnie określona*, lub *nieokreślona*, jeżeli forma kwadrtowa $\Phi(x) = \vec{x}^H A\vec{x}$ jest odpowiednio *dodatnio, ujemnie, niedodatnio, nieujemnie określona*, lub *nieokreślona*.

Uwaga 12.7. Określoność formy hermitowskiej ϕ jest taka sama jak określoność macierzy tej formy w pewnej (każdej) bazie.

Zadanie 12.1. Niech $C \in \mathbb{K}^{n,n}$. Pokaż, że macierz $A = C^H C$ jest nieujemnie określona. Sformułuj warunek konieczny i dostateczny na to, aby macierz A była dodatnio określona.

Przykład 12.12. Na $\mathbb{R}^2$ forma $\phi(\vec{x}, \vec{y}) = \alpha x_1 y_1 + \beta x_2 y_2$ jest

- dodatnio określona dla $\alpha, \beta > 0$,
- ujemnie określona dla $\alpha, \beta < 0$,
- nieujemnie określona dla $\alpha, \beta \geqslant 0$,
- niedodatnio określona dla $\alpha, \beta \leqslant 0$,
- nieokreślona dla $\alpha > 0$ i $\beta < 0$.

12.5 Twierdzenie o bezwładności

Wygodnym sposobem badania określoności formy hermitowskiej jest badanie jej macierzy. Z twierdzenia 11.21 o diagonalizacji macierzy hermitowskich otrzymujemy następujący wniosek:

Wniosek 12.4. *Jeżeli $\phi : X \times X \to \mathbb{K}$ jest formą hermitowską na X, to istnieje baza $x_1, \ldots, x_n$ przestrzeni X, w której macierz formy ϕ jest diagonalna i rzeczywista.*

Dowód. Niech A to macierz formy ϕ w pewnej bazie $y_1, \ldots, y_n$. Wtedy $A = C^H DC$, gdzie D jest diagonalna i rzeczywista, natomiast C to macierz unitarna / ortogonalna, która jest macierzą zmiany bazy z $y_1, \ldots, y_n$ na $x_1, \ldots, x_n$. □

Z wniosku 12.4 wynika, że macierz A formy hermitowskiej ϕ przystaje do pewnej macierzy diagonalnej, czyli

$$D = C^H AC,$$

gdzie macierz C jest nieosobliwa. Z twierdzenia 11.21 wiemy, ża macierz C może być ortogonalna / unitarna. Jednak w definicji 12.3 przystawania macierzy wymagamy jedynie, aby macierz C była nieosobliwa. To pozwala nam jeszcze bardziej uprościć macierz D.

Przekształcając dalej macierz D, stosujemy operacje elementarne na kolumnach i wierszach. Aby w wyniku takich przekształceń otrzymać ciąg macierzy przystających, każdej operacji elementarnej na kolumnach musi odpowiadać „hermitowsko sprzężona" operacja na wierszach, a dokładniej:

(I) operacji przestawienia kolumn k_i i k_j odpowiada operacja przestawienia wierszy w_i i w_j,

(II) operacji pomnożenia kolumny przez skalar α, czyli $k_i \mapsto \alpha k_i$ odpowiada operacja $w_i \mapsto \bar{\alpha} w_i$,

(III) operacji dodania do kolumny innej pomnożonej przez skalar α, czyli $k_i \mapsto k_i + \alpha k_j$ odpowiada operacja $w_i \mapsto w_i + \bar{\alpha} w_j$.

Jeżeli E jest macierzą danej operacji elementarnej na kolumnach, to E^H jest operacją odpowiadającej jej „hermitowsko sprzężonej" operacji na wierszach.

Niech D' oznacza macierz otrzymaną z D za pomocą ciągu takich przekształceń. Wówczas

$$D' = E_r^H \dots E_2^H E_1^H D E_1 E_2 \dots E_r = K^H D K = L^H A L,$$

gdzie $K = E_1 E_2 \dots E_r$, $L = CK$. Macierz A przystaje więc do macierzy D'. (Jednak nie jest już prawdą, że macierze A i D' są podobne!)

Przekształcimy teraz macierz hermitowską A do możliwie prostej macierzy przystającej. Z wniosku 12.4 wynika, że macierz A przystaje do macierzy

$$D = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}, \qquad \text{gdzie } d_i \in \mathbb{R}, \quad i = 1, \dots, n.$$

Teraz przestawiamy elementy na przekątnej macierzy D tak, abyśmy najpierw mieli r liczb dodatnich, potem s liczb ujemnych, a na końcu $n - r - s$ zer. Możemy to zrobić za pomocą par opercji przestawiania wierszy i kolumn: aby przestawić elementy d_i oraz d_j, najpierw przestawiamy kolumny i oraz j, a potem wiersze i oraz j. W ten sposób otrzymamy macierz D' przystającą do macierzy D:

$$D' = \operatorname{diag}(d'_1, \dots, d'_r, d'_{r+1}, \dots, d'_{r+s}, 0, \dots, 0),$$

gdzie $d'_1, \dots, d'_r > 0$, $d'_{r+1}, \dots, d'_{r+s} < 0$. Następnie, dla $i = 1, \dots, r + s$, i-tą kolumnę oraz i-ty wiersz mnożymy przez $|d_i|^{-1/2}$, otrzymując macierz D'' przystająca do D', D i A:

$$D'' = \operatorname{diag}(\underbrace{1, \dots, 1}_{r \text{ razy}}, \underbrace{-1, \dots, -1}_{s \text{ razy}}, 0, \dots, 0).$$

Udowodniliśmy

Twierdzenie 12.5 (Twierdzenie Sylwestera - Jacobiego o bezwładności)**.** *Załóżmy, że X jest przestrzenią liniową wymiaru n nad ciałem liczb rzeczywistych lub zespolonych $\mathbb{K}$. Niech $\phi : X \times X \to \mathbb{K}$ będzie formą hermitowską na X. Wówczas w X istnieje baza, w której ϕ ma macierz diagonalną*

$$D = \begin{bmatrix} I_r & & \\ & -I_s & \\ & & 0 \end{bmatrix}. \tag{12.2}$$

Ponieważ każda macierz hermitowska zadaje formę hermitowską, więc natychmiast otrzymujemy twierdzenie o bezwładności dla macierzy hermitowskich:

Wniosek 12.6. *Każda macierz hermitowska $A \in \mathbb{K}^{n,n}$ przystaje do macierzy diagonalnej*

$$D = \begin{bmatrix} I_r & & \\ & -I_s & \\ & & 0 \end{bmatrix}.$$

Definicja 12.8. Liczbę r we wzorze (12.2) nazywamy *dodatnim indeksem bezwładności* formy hermitowskiej ϕ.

Liczbę s we wzorze (12.2) nazywamy *ujemnym indeksem bezwładności* formy hermitowskiej ϕ.

Analogicznie definiujemy dodatni i ujemny indeks bezwładności macierzy hermitowskiej.

Indeksy bezwładności formy / macierzy hermitowskiej jednoznacznie określają klasę abstrakcji relacji przystawania macierzy, do której należą macierze danej formy i dają nam pełną informację o jej określoności:

Twierdzenie 12.7. *Niech r i s oznaczają odp. dodatni i ujemny indeks bezwładności formy hermitowskiej ϕ na n-wymiarowej przestrzei liniowej X (lub macierzy hermitowskiej $A \in \mathbb{K}^{n,n}$). Wówczas forma ϕ (macierz A) jest*

(a) dodatanio określona wtedy i tylko wtedy, gdy $r = n$,

(b) ujemnie określona wtedy i tylko wtedy, gdy $s = n$,

(c) nieujemnie określona wtedy i tylko wtedy, gdy $s = 0$,

(d) niedodatnio określona wtedy i tylko wtedy, gdy $r = 0$,

(e) nieokreślona wtedy i tylko wtedy, gdy $r > 0$ i $s > 0$.

Zadanie 12.2. Udowodnij twierdzenie 12.7.

Zadanie 12.3. Wyznacz indeksy bezwładności macierzy $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Zadanie 12.4. Pokaż, że suma dodatniego i ujemnego indeksu bezwładności macierzy hermitowskiej $A \in \mathbb{K}^{n,n}$ jest równa n wtedy i tylko wtedy, gdy macierz A jest nieosobliwa.

12.6 Kryterium Sylwestera

Na koniec udowodnimy proste kryterium, które pozwala rozstrzygać o dodatniej i ujemnej określoności form i macierzy hermitowskich niewielkich rozmiarów.

Twierdzenie 12.8 (Kryterium Sylwestera)**.** *Załóżmy, że $\phi : X \times X \to \mathbb{K}$ jest formą hermitowską na przestrzeni liniowej X, $x_1, \ldots, x_n$ to baza X i $A_k = [\phi(x_i, x_j)]_{i,j=1}^k$ dla $k = 1, \ldots, n$. Wówczas forma ϕ jest*

(a) dodatnio określona wtedy i tylko wtedy, gdy $\det_k A_k > 0$ dla $k = 1, \ldots, n$,

(b) ujemnie określona wtedy i tylko wtedy, gdy $(-1)^k \det_k A_k > 0$ dla $k = 1, \ldots, n$.

Dowód. W dowodzie wykorzystamy twierdzenie 12.5. Załóżmy, że forma ϕ jest dodatnio określona. Niech $V_k = \operatorname{span}(x_1, \ldots, x_k)$. Wówczas forma $\phi_k = \phi|_{V_k \times V_k}$ też jest dodatnio określona, więc w pewnej innej bazie ma macierz identycznościową I_k, przy czym macierze A_k i I_k są przystające:

$$A_k = C_k^H I C_k = C_k^H C_k,$$

skąd $\det A_k = \det C_k^H \cdot \det C_k = \det \overline{C}_k \cdot \det C_k = \overline{\det C_k} \cdot \det C_k = |\det C_k|^2 > 0$.

Załóżmy teraz, że $\det_k A_k > 0$ dla $k = 1, \ldots, n$. Pokażemy przez indukcję, że każda z form ϕ_k jest dodatnio określona (na V_k). Dla $k = n$ otrzymamy tezę. Dla $k = 1$ $\dim V_k = 1$ i ϕ_1 ma macierz $A_1 = [a_{1,1}]$, $a_{1,1} > 0$. Zatem ϕ_1 jest dodatnio określona.

Załóżmy teraz, że forma ϕ_k jest dodatnio określona (na podprzestrzeni $V_k = \operatorname{span}(x_1, \ldots, x_k)$). A_k to macierz formy ϕ_k w bazie $x_1, \ldots, x_n$. Forma ϕ_k zadaje iloczyn skalarny $\langle u, v \rangle = \phi_k(u, v) = \phi(u, v)$ na podprzestrzeni V_k.

Niech $y_1, \ldots y_k$ to baza podprzestrzeni V_k, w ktorej forma ϕ_k ma macierz identycznościową I_k. Względem iloczynu skalarnego zadanego przez ϕ_k jest to baza ortonormalna. Określmy

$$y_{k+1} = x_{k+1} - \sum_{j=1}^{k} \phi(y_j, x_{k+1}) y_j.$$

Zauważmy, że $y_{k+1} \in V_{k+1} \backslash V_k$, wiec układ $y_1, \ldots, y_k, y_{k+1}$ jest bazą podprzestrzeni V_{k+1}. Ponadto, dla $j = 1, \ldots, k$, z ortogonalności układu $y_1, \ldots, y_k$

$$\begin{aligned}
\phi(y_j, y_{k+1}) &= \phi\left(y_j, x_{k+1} - \sum_{i=1}^{k} \phi(y_i, x_{k+1}) y_i\right) \\
&= \phi(y_j, x_{k+1}) - \sum_{i=1}^{k} \phi(y_i, x_{k+1}) \phi(y_j, y_i) \\
&= \phi(y_j, x_{k+1}) - \phi(y_j, x_{k+1}) \\
&= 0,
\end{aligned}$$

Zatem macierz formy ϕ w bazie $y_1, \ldots, y_{k+1}$ to

$$D = [\phi(y_i, y_j)]_{i,j=1}^{n} = \operatorname{diag}(1, \ldots, 1, \phi(y_{k+1}, y_{k+1})).$$

Ponieważ A_{k+1} to macierz ϕ w bazie $x_1, \ldots, x_{k+1}$, więc $A_n = C^H D C$, gdzie C jest macierzą zmiany bazy z $x_1, \ldots, x_{k+1}$ na $y_1, \ldots, y_{k+1}$. Ponadto

$$\begin{aligned}
0 < \det A_k = \det(C^H D C) &= \det(C^H) \cdot \det C \cdot \det D \\
&= \overline{\det C} \cdot \det C \cdot \det D = |\det C|^2 \cdot \phi(y_{k+1}, y_{k+1}),
\end{aligned}$$

więc $\phi(y_{k+1}, y_{k+1}) > 0$, czyli wszystkie wyrazy na przekątnej macierzy D są dodatnie – macierz ta przystaje więc do macierzy I_{k+1}, skąd wynika, że forma ϕ_k jest dodatnio określona.

W przypadku formy ϕ ujemnie określonej, wystarczy zauważyć, że forma $-\phi$ jest dodatnio określona i skorzystać z własności wyznacznika. $\square$