

GAL

J. Chaber, R.Pol
Wydział MIM UW

październik 2015

Wstęp

Materiały do zajęć z GAL-u są oparte na naszym wieloletnim doświadczeniu w prowadzeniu tych zajęć na Wydziale MIM UW i są dostosowane do obecnego programu tego przedmiotu.

Spośród wielu możliwych realizacji tego programu wybraliśmy ujęcie podkreślające rolę algebry macierzy i jej ścisłe związki z przekształceniami liniowymi, przy czym duży wpływ na nasz sposób prezentacji miały znakomite książki A. Białynickiego–Biruli, Algebra liniowa z geometrią, G. Birkhoffa i S. Mac Lane’a, Przegląd algebry współczesnej, A.I. Kostrykina, Wstęp do Algebry, tom 2 oraz G. Stranga, Linear algebra and its applications.

Należy wymienić tu też skrypt T. Koźniewskiego, Wykłady z algebry liniowej, który odgrywa bardzo ważną rolę w wykładach i ćwiczeniach z GAL-u na Wydziale MIM.

Aktualną wersję skryptu można znaleźć na stronie

<http://www.mimuw.edu.pl/~chaber/gal/>.

Spis treści

1	Układy równań liniowych	1
1.1	Układy równań, macierze.	1
1.2	Redukcja wierszowa macierzy.	2
1.3	Eliminacja Gaussa.	3
2	Liczby zespolone, ciała	5
2.1	Liczby zespolone.	5
2.2	Postać trygonometryczna.	6
2.3	Pierwiastki z jedności.	6
2.4	Ciała.	7
2.5	Ciała \mathbb{Z}_p	7
3	Przestrzenie liniowe	8
3.1	Przestrzenie liniowe.	8
3.2	Kombinacje liniowe.	9
3.3	Liniowa niezależność.	10
3.4	Baza i wymiar.	11
3.5	Rząd macierzy.	13
3.6	Suma prosta podprzestrzeni.	14
4	Przekształcenia liniowe	16
4.1	Przekształcenia liniowe.	16
4.2	Jądro i obraz, izomorfizmy.	17
4.3	Przekształcenia liniowe przestrzeni współrzędnych.	19
4.4	Przestrzenie przekształceń liniowych.	20
4.5	Izomorfizmy przestrzeni współrzędnych.	20
4.6	Macierz przekształcenia.	22
4.7	Przestrzeń funkcjonałów liniowych.	22
5	Wyznaczniki	25
5.1	Definicja i podstawowe własności.	25
5.2	Obliczanie wyznaczników.	27
5.3	Macierz stowarzyszona i wzory Cramera.	29
6	Endomorfizmy przestrzeni liniowych	30
6.1	Wielomian charakterystyczny, wektory własne.	30
6.2	Zasadnicze twierdzenie algebry.	32
6.3	Twierdzenie Jordana.	33
7	Liniowe przestrzenie euklidesowe	37
7.1	Iloczyn skalarny i norma.	37
7.2	Ortogonalność w przestrzeniach euklidesowych.	38
7.3	Wyznacznik Grama i objętość.	39
7.4	Orientacja i iloczyn wektorowy.	41
7.5	Prostokątne układy współrzędnych.	43
8	Endomorfizmy przestrzeni euklidesowych	44
8.1	Endomorfizmy samosprężone.	44
8.2	Izometrie liniowe.	45
8.3	*Sprężenie endomorfizmu przestrzeni euklidesowej.	48

9	Przestrzenie afiniczne	49
9.1	Struktura afiniczna przestrzeni współrzędnych.	49
9.2	Przestrzenie afiniczne.	50
9.3	Kombinacje afiniczne.	51
9.4	Układ bazowy, baza punktowa.	52
9.5	Przekształcenia afiniczne	53
9.6	Afiniczne układy współrzędnych.	54
10	Przestrzenie Euklidesowe	55
10.1	Afiniczne przestrzenie euklidesowe.	55
10.2	Izometrie przestrzeni euklidesowych.	55
11	Formy kwadratowe	57
11.1	Formy kwadratowe.	57
11.2	Algorytm Lagrange'a.	58
11.3	Warunek Jacobiego.	61
11.4	Przestrzenie z formą kwadratową.	61
11.5	Twierdzenie Sylwestera o bezwładności.	62
12	Funkcje kwadratowe	64
12.1	Funkcje kwadratowe na \mathbb{K}^n	64
12.2	Klasyfikacja funkcji kwadratowych na \mathbb{R}^n	65
13	Hiperpowierzchnie stopnia 2.	67
13.1	Hiperpowierzchnie stopnia 2.	67
13.2	Klasyfikacja hiperpowierzchni właściwych.	69
14	Uzupełnienia	71
14.1	Twierdzenie Weierstrassa.	71
14.2	Zespolona wartość własna macierzy rzeczywistej.	72
14.3	Orientacja liniowej przestrzeni euklidesowej.	72
14.4	Równania opisujące hiperpowierzchnię właściwą.	73

1 Układy równań liniowych

1.1 Układy równań, macierze.

W tej części opiszemy metodę rozwiązywania układów m równań z n niewiadomymi o współczynnikach rzeczywistych, tzn. układów

$$(*) \quad \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases},$$

gdzie $a_{ij} \in \mathbb{R}$ są stałymi współczynnikami, $b_i \in \mathbb{R}$ są stałymi wyrazami wolnymi, a symbole x_j oznaczają niewiadome.

Definicja 1.1.1 *Jeśli wszystkie wyrazy wolne b_i są zerami, to układ jest jednorodny.*

Współczynniki układu (*) można zapisać w postaci $(m \times n)$ -macierzy

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

współczynnik a_{ij} nazywamy (i, j) -tym wyrazem macierzy, a $(m \times 1)$ -macierze i $(1 \times n)$ -macierze

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{mj} \end{bmatrix} \quad \text{i} \quad [a_{i1}, a_{i2}, \dots, a_{in}]$$

nazywamy odpowiednio j -tą kolumną i i -tym wierszem macierzy.

Następujące dwie interpretacje będą odgrywały w przyszłości ważną rolę.

Niech \mathbb{R}^m będzie przestrzenią kolumn o m elementach, tzn. $(m \times 1)$ -macierzy. Elementy \mathbb{R}^m będziemy nazywali wektorami wymiaru m , a wyraz takiego wektora stojący w i -tym wierszu jego i -tą współrzędną. Wektor, który ma wszystkie współrzędne zerowe nazywamy *wektorem zerowym* i oznaczamy symbolem $\mathbf{0}$. Wektory z \mathbb{R}^m dodajemy, sumując i -te współrzędne i mnożymy przez liczby (lub symbole x), mnożąc każdą współrzędną osobno.

Tak więc układ (*) zapisuje się w postaci

$$(*w) \quad x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Określając iloczyn macierzy o n kolumnach przez wektor wymiaru n o współrzędnych x_i wzorem

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

i przyjmując oznaczenia

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

układ równań (*) można zapisać w postaci

$$(*m) \quad AX = B.$$

Tak więc rozwiązanie układu (*) polega na wyznaczeniu, o ile istnieją, wszystkich wektorów X takich, że wektor wyrazów wolnych B jest iloczynem macierzy współczynników A i wektora X .

Definicja 1.1.2 Układ równań $AX = B$, który nie ma rozwiązań nazywamy sprzecznym.

1.2 Redukcja wierszowa macierzy.

Definicja 1.2.1 Dwa układy m równań z n niewiadomymi nazywamy równoważnymi jeśli mają taki sam zbiór rozwiązań (w szczególności, jeśli oba są sprzeczne).

Opiszemy teraz trzy operacje na układach równań, które nie zmieniają zbioru rozwiązań i pozwalają zastąpić dany układ równań układem równoważnym o przejrzystej, “schodkowej” strukturze:

- (I) $_{a(i)+(k)}$ dodanie do k -tego równania i -tego równania pomnożone przez a ,
- (II) $_{(i)(k)}$ zamiana miejscami i -tego równania z k -tym,
- (III) $_{c(i)}$ pomnożenie i -tego równania przez liczbę $c \neq 0$.

Twierdzenie 1.2.2 Wykonanie na układzie równań liniowych jednej z wymienionych wyżej operacji nie zmienia zbioru rozwiązań tego układu.

Dowód. Teza jest oczywista dla operacji typu (II) i (III). Rozpatrzmy operację (I) $_{a(i)+(k)}$ przeprowadzającą układ (*) na układ (*)', w którym zmienia się jedynie k -te równanie, otrzymane w wyniku dodania stronami do k -tego równania układu (*) równania i -tego, obustronnie pomnożonego przez a .

Jest jasne, że każde rozwiązanie układu (*) jest także rozwiązaniem układu (*)' . Ponieważ operacja (I) $_{(-a)(i)+(k)}$ (odwrotna do (I) $_{a(i)+(k)}$) przeprowadza układ (*)' na układ (*), także rozwiązania układu (*)' są rozwiązaniami (*). To pokazuje równoważność obu układów. ■

Przy opisanych przekształceniach układu równań $AX = B$, celowe jest pomijanie zmiennych i wykonywanie operacji na wierszach macierzy rozszerzonej tego układu

$$[A|B] = [A_1, \dots, A_n|B],$$

gdzie A_j są kolejnymi kolumnami macierzy A , B jest dopisaną jako ostatnia kolumną wyrazów wolnych, a kreska oddzielająca B od poprzednich kolumn nie ma formalnego znaczenia i ma jedynie przypominać, że przy przejściu do układu równań, rola ostatniej kolumny jest inna niż pozostałych.

Operacjom na układach równań odpowiadają następujące operacje elementarne na wierszach macierzy:

- (I) $_{a(i)+(k)}$ dodanie do k -tego wiersza i -tego wiersza pomnożone przez a ,
- (II) $_{(i)(k)}$ zamiana miejscami i -tego wiersza z k -tym,
- (III) $_{c(i)}$ pomnożenie i -tego wiersza przez liczbę $c \neq 0$.

Opiszemy teraz pewne macierze o szczególnie prostej postaci i pokażemy, że każdą macierz można sprowadzić do macierzy takiej postaci operacjami typu (I) i (II), zob. Uwagę 1.2.5.

Definicja 1.2.3 Mówimy, że macierz jest w postaci schodkowej jeśli spełnione są dwa warunki:

(S1) "raden wiersz zerowy tej macierzy nie poprzedza wiersza niezerowego,

(S2) pierwsze niezerowe wyrazy (schodki) kolejnych niezerowych wierszy tej macierzy stoją w kolumnach o rosnących numerach.

Twierdzenie 1.2.4 Dowolną macierz można sprowadzić do postaci schodkowej operacjami elementarnymi typu (I) i (II) na wierszach tej macierzy.

Dowód. Niech A będzie $(m \times n)$ -macierzą mającą niezerowe wyrazy i niech j_1 będzie numerem pierwszej niezerowej kolumny A . Zamieniając w razie potrzeby wiersze macierzy A miejscami (operacja typu (II)) można otrzymać macierz \tilde{A} mającą niezerowy wyraz a_{1j_1} w pierwszym wierszu kolumny o numerze j_1 :

$$\tilde{A} = \begin{bmatrix} 0 & \dots & 0 & a_{1j_1} & \dots & a_{1n} \\ 0 & \dots & 0 & a_{2j_1} & \dots & a_{2n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & a_{mj_1} & \dots & a_{mn} \end{bmatrix}.$$

Odejmując kolejno, dla $i = 2, 3, \dots, m$, od i -tego wiersza macierzy \tilde{A} pierwszy wiersz pomnożony przez $a_i = \frac{a_{ij_1}}{a_{1j_1}}$ (czyli wykonując operację $(I)_{(-a_i)(1)+(i)}$) otrzymujemy macierz A' , której kolumny o numerach mniejszych niż j_1 są zerowe, a jedynym niezerowym wyrazem w kolumnie o numerze j_1 jest a_{1j_1} :

$$A' = \begin{bmatrix} 0 & \dots & 0 & a_{1j_1} & a_{1j_1+1} & \dots & a_{1n} \\ 0 & \dots & 0 & 0 & a'_{2j_1+1} & \dots & a'_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & a'_{mj_1+1} & \dots & a'_{mn} \end{bmatrix}.$$

W następnym kroku powtarzamy tę procedurę dla macierzy A' , z pominięciem pierwszego wiersza tej macierzy, który nie będzie już wykorzystywany w kolejnych operacjach. Znajdujemy $j_2 > j_1$ i macierz $A'' \in \mathbb{R}_n^m$ (pierwszy wiersz A' pozostaje pierwszym wierszem A'') taką, że wyraz drugiego wiersza kolumny o numerze j_2 jest niezerowy, a wszystkie wyrazy pod nim oraz wyrazy z wcześniejszych kolumn (z wyjątkiem wyrazów pominiętego pierwszego wiersza) są zerowe.

Po kolejnych analogicznych krokach dochodzimy do $(m \times n)$ -macierzy $A^{(r)}$ w postaci schodkowej mającej w r niezerowych wierszach pierwsze niezerowe wyrazy (schodki) w kolumnach o numerach $j_1 < \dots < j_r$. ■

Uwaga 1.2.5 W twierdzeniu 1.2.4 można ograniczyć się do operacji typu (I). Operację $(II)_{(1)(i)}$, którą stosowaliśmy przy przejściu od A do \tilde{A} w przypadku, gdy w kolumnie o numerze j_1 pierwszy wyraz jest zerowy, a i -ty różny od zera, można zastąpić operacją $(I)_{1(i)+(1)}$. Analogicznie można postępować w kolejnych krokach. □

1.3 Eliminacja Gaussa.

Metoda eliminacji Gaussa polega na wykorzystaniu Twierdzenia 1.2.4 do analizy układów równań liniowych.

Niech $AX = B$ będzie układem m równań z n niewiadomymi.

Zgodnie z Twierdzeniem 1.2.4, macierz rozszerzoną $[A|B]$ można zredukować operacjami elementarnymi typu (I) i (II) (ignorując kreskę oddzielającą A i B) do postaci schodkowej $[A'|B']$, przy czym, zgodnie z Twierdzeniem 1.2.2, układy równań $AX = B$ i $AX' = B'$ są równoważne.

Tak więc, należy ustalić, czy układ $A'X = B'$ jest niesprzeczny i jeśli tak – wyznaczyć wszystkie jego rozwiązania.

Założmy, że macierz $[A'|B']$ ma r niezerowych wierszy, których pierwsze niezerowe wyrazy stoją w kolumnach o numerach $j_1 < j_2 < \dots < j_r$.

Jeśli $j_r = n + 1$ (schodek ostatniego niezerowego wiersza macierzy $[A'|B']$ znajduje się w ostatniej kolumnie B' tej macierzy), to układ $A'X = B'$ (a więc i układ $AX = B$) jest sprzeczny.

W przeciwnym wypadku ($j_r < n + 1$) wszystkie rozwiązania układu $AX = B$ znajdujemy wyznaczając z układu $A'X = B'$ niewiadome $x_{j_1}, x_{j_2}, \dots, x_{j_r}$ (zmiennie zależne) w zależności od pozostałych niewiadomych, które mogą przyjmować dowolne wartości (zmiennie niezależne, parametry).

Kolejne zmienne zależne $x_{j_r}, x_{j_{r-1}}, \dots, x_{j_1}$ wyznaczamy wtedy z kolejnych równań układu $A'X = B'$, zaczynając od ostatniego niezerowego (tak więc, w pewnym sensie, wyznaczając kolejne zmienne zależne od ostatniej do pierwszej “wchodzimy po schodkach” układu równań).

Zmienna zależna x_{j_k} wyliczana z k -tego równania zależy wyłącznie od zmiennych niezależnych o numerach większych niż j_k (za zmienne zależne o numerach większych niż j_k podstawiamy znalezione wcześniej wartości).

Uwaga 1.3.1 Rozwiązanie $X \in \mathbb{R}^n$ zależy od $n - r$ współrzędnych X – zmiennych niezależnych i można je przedstawić w postaci (zwanej rozwiązaniem ogólnym układu $AX = B$)

$$X = X_0 + t_1 X_1 + t_2 X_2 + \dots + t_p X_p,$$

gdzie t_1, t_2, \dots, t_p są dowolnymi parametrami ($p = (n - r)$ jest liczbą kolumn bez schodków macierzy A').

W praktyce, zamiast wyliczać rozwiązanie ogólne X , wygodnie jest obliczyć wektory X_0, X_1, \dots, X_p występujące we wzorze na X , podstawiając za zmienne niezależne odpowiednie wartości:

X_0 jest rozwiązaniem $A'X = B'$ odpowiadającym parametrom $t_j = 0$ dla $j = 1, 2, \dots, p$,

X_k jest rozwiązaniem $A'X = \mathbf{0}$ odpowiadającym parametrom $t_k = 1$ oraz $t_j = 0$ dla $j \neq k$

($X_k = (X_0 + X_k) - X_0$ jest rozwiązaniem $A'X = \mathbf{0}$ jako różnica dwóch rozwiązań $A'X = B'$). \square

2 Liczby zespolone, ciała

Nasza dyskusja równań liniowych opierała się na jedynie na regułach arytmetyki liczb rzeczywistych i zbiór liczb rzeczywistych można tu zastąpić innymi obiektami algebraicznymi – ciałami, których elementy można dodawać i mnożyć zgodnie z analogicznymi regułami.

Z punktu widzenia tego wykładu najważniejszym, obok \mathbb{R} , ciałem jest ciało liczb zespolonych, które otrzymuje się dołączając do \mathbb{R} , w możliwie oszczędny sposób, rozwiązanie równania $x^2 = -1$ (którego nie ma w ciele \mathbb{R}).

Wspomnimy też jednak o ciałach zupełnie innego typu – ciałach skończonych \mathbb{Z}_p .

2.1 Liczby zespolone.

Liczby rzeczywiste \mathbb{R} rozszerzymy dołączając “liczbę urojoną” $\sqrt{-1}$ oznaczaną symbolem i , tak aby na otrzymanych “liczbach zespolonych” można było wykonywać algebraiczne operacje dodawania i mnożenia zgodnie ze standardowymi regułami arytmetyki liczb rzeczywistych.

W części 6.2 pokażemy, że dołączenie $\sqrt{-1}$ prowadzi do systemu liczbowego, w którym każdy wielomian stopnia dodatniego $a_0 + a_1x^1 + \dots + a_nx^n$ ma pierwiastek (zasadnicze twierdzenie algebry).

Definicja 2.1.1 *Liczbami zespolonymi będziemy nazywać wyrażenia postaci $a + ib$ (gdzie $i = \sqrt{-1}$ oraz $a + ib = c + id \Leftrightarrow a = c, b = d$) z następującymi operacjami dodawania \oplus i mnożenia \odot :*

$$(a + ib) \oplus (c + id) = (a + c) + i(b + d); \quad (a + ib) \odot (c + id) = (ac - bd) + i(ad + bc)$$

Zbiór liczb zespolonych z tak określonymi działaniami oznaczamy symbolem \mathbb{C} .

Uwaga 2.1.2 (a) Wyrażenie $a + i0$ zapisujemy jako a i utożsamiamy je z liczbą rzeczywistą a . W ten sposób $\mathbb{R} \subset \mathbb{C}$, przy czym działania \oplus i \odot pokrywają się na \mathbb{R} ze zwykłym dodawaniem i mnożeniem.

(b) Zamiast $0 + ib$, $b \neq 0$ piszemy ib (lub i , jeśli $b = 1$); w szczególności $i \odot i = -1$, tzn. $i^2 = -1$ w \mathbb{C} .

(c) Liczbę zespoloną $z = a + ib$ można interpretować jako punkt (a, b) płaszczyzny kartezjańskiej. Współrzędne a, b tego punktu będziemy nazywać odpowiednio *częścią rzeczywistą* $\operatorname{Re}z$ i *częścią urojoną* $\operatorname{Im}z$ liczby z . □

Uwaga 2.1.3 Dodawanie i mnożenie liczb zespolonych spełniają standardowe reguły arytmetyki liczb rzeczywistych, \mathbb{C} ma elementy neutralne ze względu na dodawanie i mnożenie (zero i jedynkę); w \mathbb{C} wykonalne są też operacje odejmowania i dzielenia przez liczby różne od zera:

- (1) przemienność $z_1 \oplus z_2 = z_2 \oplus z_1, \quad z_1 \odot z_2 = z_2 \odot z_1;$
- (2) łączność $(z_1 \oplus z_2) \oplus z_3 = z_1 \oplus (z_2 \oplus z_3), \quad (z_1 \odot z_2) \odot z_3 = z_1 \odot (z_2 \odot z_3);$
- (3) elementy neutralne 0 dla dodawania: $z \oplus 0 = z, \quad 1$ dla mnożenia: $1 \odot z = z;$
- (4) istnienie elementu przeciwnego $-z: z \oplus -z = 0$ odwrotnego z^{-1} , dla $z \neq 0: z \odot z^{-1} = 1$
 $-(a + ib) = (-a) + i(-b), \quad a + ib \neq 0, \text{ to } (a + ib)^{-1} = \left(\frac{a}{a^2 + b^2}\right) + i\left(\frac{-b}{a^2 + b^2}\right);$
- (5) rozdzielność mnożenia względem dodawania $z_1 \odot (z_2 \oplus z_3) = z_1 \odot z_2 \oplus z_1 \odot z_3.$ □

W dalszym ciągu zamiast \oplus i \odot będziemy używali zwykłych symboli dodawania i mnożenia. Odejmowanie definiujemy formalnie jako dodanie liczby przeciwnej $z_1 - z_2 = z_1 + (-z_2)$, a dzielenie jako mnożenie przez liczbę odwrotną $z_1 : z_2 = \frac{z_1}{z_2} = z_1(z_2^{-1})$. Na \mathbb{C} rozszerzamy zwykle konwencje dotyczące działań w \mathbb{R} , na przykład iloczyn n egzemplarzy liczby z (tzn. n -tą potęgę z) zapisujemy w postaci z^n , $z^0 = 1$ i $z^{-n} = (z^{-1})^n$.

2.2 Postać trygonometryczna.

Modułem liczby zespolonej $z = a + ib$ nazywamy liczbę $|z| = \sqrt{a^2 + b^2} \in \mathbb{R}$. Interpretując liczbę $z \neq 0$ jako punkt (a, b) płaszczyzny kartezjańskiej widzimy, że $|z|$ jest odległością z od 0 , a liczba $\frac{z}{|z|}$ odpowiadająca punktowi okręgu jednostkowego na płaszczyźnie ma postać $\frac{z}{|z|} = \cos \theta + i \sin \theta$, przy czym każde dwa takie kąty θ – *argumenty* z , różnią się o wielokrotność 2π ; zbiór argumentów liczby z oznaczamy symbolem $\arg z$.

Każdy zapis liczby zespolonej $z \neq 0$,

$$z = |z|(\cos \theta + i \sin \theta),$$

gdzie $|z|$ jest modułem z i $\theta \in \arg z$ nazywamy *postacią trygonometryczną* z .

Twierdzenie 2.2.1 *Niech $z_1 = |z_1|(\cos \theta_1 + i \sin \theta_1)$, $z_2 = |z_2|(\cos \theta_2 + i \sin \theta_2)$. Wtedy*

$$z_1 z_2 = |z_1| |z_2| (\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)),$$

tn. moduł iloczynu jest iloczynem modułów, a suma argumentów czynników jest argumentem iloczynu.

Dowód. $z_1 z_2 = |z_1|(\cos \theta_1 + i \sin \theta_1) |z_2|(\cos \theta_2 + i \sin \theta_2) = |z_1| |z_2| (\cos \theta_1 + i \sin \theta_1)(\cos \theta_2 + i \sin \theta_2) = |z_1| |z_2| (\cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2 + i(\sin \theta_1 \cos \theta_2 + \cos \theta_1 \sin \theta_2)) = |z_1| |z_2| (\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2))$. ■

Wniosek 2.2.2 (Formuła de Moivre'a). $(\cos \theta + i \sin \theta)^n = \cos(n\theta) + i \sin(n\theta)$.

Sprzężeniem liczby $z = a + ib$ nazywamy liczbę $\bar{z} = a - ib$. Dla $z \neq 0$ mamy $z\bar{z} = |z|^2$ i $\frac{1}{z} = \frac{\bar{z}}{|z|^2}$.

2.3 Pierwiastki z jednościami.

Ustalmy liczbę naturalną $n > 1$. *Pierwiastkiem stopnia n z jednościami* będziemy nazywać każdą liczbę zespoloną z taką, że $z^n = 1$.

Niech $\omega = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$. Z formuły de Moivre'a wynika natychmiast, że liczby $\omega^k = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}$, $k = 0, 1, \dots, n-1$, są wszystkimi pierwiastkami stopnia n z jednościami.

Punkty płaszczyzny kartezjańskiej odpowiadające pierwiastkom stopnia n z jednościami są wierzchołkami n -kąta foremnego wpisanego w okrąg jednostkowy, mającego wierzchołek w $\omega^0 = 1$.

Pierwiastek stopnia n z jednościami nazywamy *pierwotnym* jeśli nie jest pierwiastkiem z jednościami stopnia $< n$. Do scharakteryzowania pierwiastków pierwotnych skorzystamy z następującego faktu związanego z dzieleniem z resztą liczb naturalnych.

Lemat 2.3.1 *Dla względnie pierwszych liczb naturalnych $0 < k < n$ istnieją liczby całkowite l, t takie, że $lk + tn = 1$. Co więcej, można zakładać, że $0 < l < n$.*

Dowód. Niech d będzie najmniejszą liczbą dodatnią postaci $d = sk + tn$, gdzie s, t są całkowite. Wystarczy pokazać, że d jest dzielnikiem k i n . Dla reszty $r = k - qd$ z dzielenia k przez d mamy $r = k - q(sk + tn) = (1 - qs)k + (-t)n$, więc $r = 0$ z minimalności d . Analogicznie pokazuje się, że d dzieli n . Drugą część tezy otrzymujemy przyjmując za l resztę z dzielenia s przez n . Wtedy $s = qn + l$, więc $1 = sk + tn = (qn + l)k + tn = lk + (qk + t)n$ i w szczególności $l > 0$. ■

Twierdzenie 2.3.2 *Pierwiastek $\omega^k = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}$, $1 < k < n$ stopnia n z jednościami jest pierwotny wtedy i tylko wtedy, gdy k i n są względnie pierwsze.*

Dowód. Niech k i n będą względnie pierwsze. Z lematu istnieją l, t takie, że $1 = lk + tn$, a stąd $\omega = \omega^{lk+tn} = \omega^{lk}(\omega^n)^t = \omega^{lk}$. Jeśli $m > 0$ spełnia $(\omega^k)^m = 1$, to $\omega^m = (\omega^{lk})^m = (\omega^{km})^l = 1$, więc $m \geq n$, czyli pierwiastek ω^k jest pierwotny.

Załóżmy teraz, że $d > 1$ jest wspólnym dzielnikiem k i n , a q oraz m są takie, że $k = qd$ oraz $n = md$. Wtedy $(\omega^k)^m = (\omega^{qd})^m = (\omega^{md})^q = 1$, więc pierwiastek ω^k nie jest pierwotny. ■

2.4 Ciała.

Własności dodawania i mnożenia w \mathbb{R} i w \mathbb{C} zebrane w Uwadze 2.1.3, stanowią punkt wyjścia definicji ciała.

Definicja 2.4.1 Zbiór K z dwoma ustalonymi elementami: $0, 1$ ($0 \neq 1$) oraz dwoma działaniami: dodawania “+” i mnożenia “ \cdot ” nazywamy ciałem, jeśli dla dowolnych $a, b, c \in K$ spełnione są warunki (dziewięć aksjomatów ciała)

- | | | |
|--|--|--|
| (1) przemienność | $a + b = b + a,$ | $a \cdot b = b \cdot a;$ |
| (2) łączność | $(a + b) + c = a + (b + c),$ | $(a \cdot b) \cdot c = a \cdot (b \cdot c);$ |
| (3) elementy neutralne | 0 dla dodawania: $a + 0 = a,$ | 1 dla mnożenia: $1 \cdot a = a;$ |
| (4) istnienie elementu przeciwnego a' : | $a + a' = 0,$ | odwrotnego a^* , dla $a \neq 0$: $a \cdot a^* = 1;$ |
| (5) rozdzielność mnożenia względem dodawania | $a \cdot (b + c) = a \cdot b + a \cdot c.$ | |

Dla podkreślenia, że ciało to zbiór z wyróżnionymi zerem i jedynką oraz z ustalonymi działaniami, będziemy pisać \mathbb{K} zamiast K .

Równanie $x + a = b$ ma w ciele \mathbb{K} dokładnie jedno rozwiązanie, bo dodając do obu stron tego równania a' – ustalony element przeciwny do a otrzymujemy, po uporządkowaniu równoważne równanie $x = b + a'$.

W szczególności wynika stąd, że 0 i element przeciwny do a (oznaczany przez $-a$) są wyznaczone jednoznacznie. Analogiczne rozumowanie dla równania $x \cdot a = b$, gdzie $a \neq 0$, pokazuje że 1 i element odwrotny do a (oznaczany przez a^{-1}) są wyznaczone jednoznacznie. Ułamek $\frac{b}{a}$ oznacza iloczyn $b \cdot a^{-1}$.

Wszystko, co powiedzieliśmy w pierwszym rozdziale o układach równań liniowych o współczynnikach z ciała liczb rzeczywistych przenosi się bez zmian na układy o współczynnikach z dowolnego ciała, tzn. na układy postaci $AX = B$, gdzie A jest $(m \times n)$ -macierzą o wyrazach z \mathbb{K} , $B \in \mathbb{K}^m$.

W dowolnym ciele prawdziwe są dobrze znane własności działań w \mathbb{R} (piszemy często ab zamiast $a \cdot b$).

Uwaga 2.4.2 Dla dowolnych $a, b \in \mathbb{K}$:

- a) $a0 = 0$ (bo do obu stron $a0 + a0 = a(0 + 0) = a0$ można dodać $-(a0)$);
- b) $ab = 0$, to $a = 0$ lub $b = 0$ (bo $a \neq 0$, to obie strony można pomnożyć przez a^{-1});
- c) $(-1)a = -a$ (bo $a + (-1)a = (1 + (-1))a = 0a = 0$). □

2.5 Ciała \mathbb{Z}_p .

Ważne przykłady ciał, które określimy w tej części są, w odróżnieniu od ciała liczb rzeczywistych \mathbb{R} , ciała liczb wymiernych \mathbb{Q} i ciała liczb zespolonych \mathbb{C} – ciałami skończonymi.

Niech p będzie liczbą pierwszą i niech $\omega = \cos \frac{2\pi}{p} + i \sin \frac{2\pi}{p}$ będzie pierwiastkiem stopnia p z jedności. Wszystkie potęgi ω^n są również pierwiastkami stopnia p z jedności. Zbiór $\mathbb{Z}_p = \{\omega^0, \omega^1, \dots, \omega^{p-1}\}$ wszystkich pierwiastków stopnia p z jedności jest więc zamknięty ze względu na działania

$$\omega^k \oplus \omega^l = \omega^{k+l} \quad \text{oraz} \quad \omega^k \odot \omega^l = \omega^{kl}.$$

Twierdzenie 2.5.1 \mathbb{Z}_p z ustalonym elementem zerowym $\mathbf{0} = \omega^0$, jedynką $\mathbf{1} = \omega^1$ oraz działaniami dodawania \oplus i mnożenia \odot jest ciałem.

Dowód. Elementem przeciwnym do $\omega^k \in \mathbb{Z}_p$ jest ω^{p-k} , bo $\omega^k \oplus \omega^{p-k} = \omega^p = \mathbf{0}$. Jeśli $\omega^k \in \mathbb{Z}_p \setminus \{\mathbf{0}\}$, to z Lematu 2.3.1 dla $n = p$ istnieją l, t takie, że $1 = lk + tp$. Elementem odwrotnym do $\omega^k \neq \mathbf{0}$ jest wtedy ω^l , bo $\mathbf{1} = \omega^1 = \omega^{lk} \omega^{tp} = \omega^{lk} = \omega^k \odot \omega^l$. Pozostałe aksjomaty wynikają z odpowiednich własności dodawania i mnożenia liczb naturalnych. ■

3 Przestrzenie liniowe

3.1 Przestrzenie liniowe.

Dla dowolnego ciała \mathbb{K} , analogicznie jak to robiliśmy dla \mathbb{R} , wprowadza się operację dodawania wektorów – kolumn z \mathbb{K}^n i mnożenia tych wektorów przez elementy ciała – skalary.

Jeśli A jest $(m \times n)$ -macierzą o wyrazach z ciała \mathbb{K} , to zbiór V rozwiązań układu jednorodnego $AX = \mathbf{0}$ jest zamknięty ze względu na dodawanie wektorów i mnożenie wektorów przez skalary.

Podobnie, w zbiorze W wielomianów stopnia nie większego niż n o współczynnikach rzeczywistych, podzielnych przez wielomian $x^2 + 1$, określone jest naturalne działanie dodawania i mnożenia przez liczby.

Są to przykłady przestrzeni liniowych – obiektów algebraicznych złożonych ze zbioru wektorów, ciała skalarów oraz działań, które określa się następująco.

Definicja 3.1.1 *Zbiór V z ustalonym elementem $\mathbf{0}$ (wektor zerowy) i działaniem dodawania “+” nazywamy przestrzenią liniową nad ciałem \mathbb{K} , jeśli jest ustalone działanie mnożenia “ \cdot ” elementów V – wektorów przez elementy ciała \mathbb{K} – skalary, dające w wyniku elementy V , przy czym dla dowolnych $a, b \in \mathbb{K}$ i $v, w, u \in V$ spełnione są warunki (osiem aksjomatów przestrzeni liniowej)*

- | | | |
|-----|---|---|
| (1) | $v + w = w + v$ | przemienność dodawania wektorów, |
| (2) | $v + (w + u) = (v + w) + u$ | łączność dodawania wektorów, |
| (3) | $v + \mathbf{0} = v$ | wektor zerowy jest elementem neutralnym dodawania wektorów, |
| (4) | istnieje v' takie, że $v + v' = \mathbf{0}$ | element przeciwny dodawania wektorów, |
| (5) | $a \cdot (w + u) = a \cdot w + a \cdot u$ | rozdzielność mnożenia przez skalar względem dodawania wektorów, |
| (6) | $(a + b) \cdot u = a \cdot u + b \cdot u$ | rozdzielność mnożenia przez skalar względem dodawania skalarów, |
| (7) | $a(b \cdot v) = (ab) \cdot v$ | łączność mnożenia przez skalary, |
| (8) | $1 \cdot v = v$ | skalar 1 jest elementem neutralnym mnożenia. |

Jak zobaczymy później, przyjęte aksjomaty pozwalają utożsamiać, ze względu na strukturę algebraiczną, przestrzenie liniowe nad ciałem \mathbb{K} skończonego wymiaru (innych nie będziemy tu w zasadzie rozpatrywać) z przestrzeniami \mathbb{K}^n , a jednocześnie pozwalają operować na wektorach z V , bez konieczności przypisania im konkretnych współrzędnych.

Równanie $x + v = w$ ma w przestrzeni liniowej V dokładnie jedno rozwiązanie, bo dodając do obu stron v' – ustalony element przeciwny do v otrzymujemy, po uporządkowaniu równoważne równanie $x = w + v'$.

W szczególności $\mathbf{0}$ i wektor przeciwny do v (oznaczany przez $-v$) są wyznaczone jednoznacznie.

Iloczyn av (opuszczamy znak mnożenia) znaczy to samo co va (używa się jednak zazwyczaj zapisu av).

Uwaga 3.1.2 Dla dowolnych $a \in \mathbb{K}$, $v \in V$:

- a) $av = \mathbf{0}$ jeśli $a = 0$ lub $v = \mathbf{0}$ ($v = \mathbf{0}$, to do obu stron $a\mathbf{0} + a\mathbf{0} = a(\mathbf{0} + \mathbf{0}) = a\mathbf{0}$ dodajemy $-(a\mathbf{0})$
 $a = 0$, to do obu stron $0v + 0v = (0 + 0)v = 0v$ dodajemy $-(0v)$);
- b) $av = \mathbf{0}$, to $a = 0$ lub $v = \mathbf{0}$ (bo $a \neq 0$, to obie strony mnożymy z lewej przez a^{-1});
- c) $(-1)v = -v$ (bo $v + (-1)v = (1 + (-1))v = 0v = \mathbf{0}$). □

Podamy teraz kilka podstawowych przykładów przestrzeni liniowych nad \mathbb{K} , użytecznych przy ilustrowaniu wprowadzanych przez nas kolejnych pojęć.

Przykład 3.1.3 (a) Przestrzeń współrzędnych \mathbb{K}^m .

Elementami \mathbb{K}^m (wektorami z \mathbb{K}^m) są kolumny m skalarów (współrzędnych tego wektora). Wektor zerowy ma wszystkie współrzędne zerowe. Definiujemy działania “po współrzędnych”

$$(+)\quad \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ \vdots \\ a_m + b_m \end{bmatrix}; \quad (\cdot)\quad c \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} ca_1 \\ \vdots \\ ca_m \end{bmatrix}.$$

Aksjomaty przestrzeni liniowej wynikają z odpowiednich aksjomatów ciała.

(b) Przestrzeń macierzy \mathbb{K}_n^m .

Wektorami w \mathbb{K}_n^m są macierze o wyrazach z ciała \mathbb{K} mające m wierszy i n kolumn, zob. 1.1. Macierz zerowa ma wszystkie wyrazy zerowe. Wektory – macierze dodajemy sumując ich odpowiednie wyrazy i mnożymy przez skalary – elementy ciała \mathbb{K} , mnożąc przez skalar wszystkie wyrazy macierzy.

Często wygodnie jest myśleć o $(m \times n)$ -macierzy jako o układzie n kolumn – wektorów z \mathbb{K}^m . Jeśli $A = [A_1, \dots, A_n]$, $B = [B_1, \dots, B_n]$, $A_j, B_j \in \mathbb{K}^m$ oraz $c \in \mathbb{K}$, to $A + B = [A_1 + B_1, \dots, A_n + B_n]$ i $cA = [cA_1, \dots, cA_n]$.

(c) Przestrzeń wielomianów $\mathbb{K}[x]$.

Wielomianem o współczynnikach z ciała \mathbb{K} nazywamy wyrażenie $a_0 + a_1x^1 + \dots + a_nx^n$, gdzie $a_0, \dots, a_n \in \mathbb{K}$; każdy ze składników a_jx^j nazywamy jednomianem, a największe n takie, że $a_n \neq \mathbf{0}$ nazywamy stopniem wielomianu. Często pomijamy w takim wyrażeniu te jednomiany a_jx^j , dla których $a_j = 0$, a wielomian zerowy (bez niezerowych jednomianów) oznaczamy przez $\mathbf{0}$.

W zbiorze wielomianów $\mathbb{K}[x]$ określone są działania dodawania i mnożenia spełniające wszystkie aksjomaty ciała, poza aksjomatem o istnieniu elementu odwrotnego. W szczególności $\mathbb{K}[x]$ jest przestrzenią liniową nad ciałem \mathbb{K} , bo skalar $c \in \mathbb{K}$ można uważać za jednomian.

Definicja 3.1.4 Niech V będzie przestrzenią liniową nad ciałem \mathbb{K} . Podzbiór W zbioru wektorów zawierający wektor zerowy nazywamy podprzestrzenią V , jeśli W jest zamknięty za względu na działanie dodawania i mnożenia przez skalary, to znaczy spełnione są dwa warunki

$$(+)\quad v + w \in W \text{ dla } v, w \in W; \quad (\cdot)\quad cv \in W \text{ dla } c \in \mathbb{K}, v \in W.$$

Uwaga 3.1.5 Jeśli W jest podprzestrzenią przestrzeni liniowej V nad ciałem \mathbb{K} , to W z działaniami dodawania wektorów i mnożenia wektora przez skalar ograniczonymi do W jest przestrzenią liniową nad \mathbb{K} , bo dla $v \in W$ wektor przeciwny $-v = (-1)v$ też jest w W . \square

Każda przestrzeń V liniowa zawiera podprzestrzeń maksymalną i podprzestrzeń minimalną w sensie inkluzji (zwane niewłaściwymi): samą siebie i podprzestrzeń zerową $\{\mathbf{0}\}$. W następnej części podamy ogólną metodę generowania podprzestrzeni przestrzeni liniowych V .

3.2 Kombinacje liniowe.

Kombinacje liniowe pojawiły się już przy omawianiu układów równań liniowych.

Definicja 3.2.1 *Kombinacją liniową wektorów układu (v_1, \dots, v_n) z przestrzeni liniowej V nad ciałem \mathbb{K} o współczynnikach x_1, \dots, x_n (z \mathbb{K}) nazywamy wektor $\sum_{j=1}^n x_j v_j = x_1 v_1 + \dots + x_n v_n \in V$. Powłoką liniową układu (v_1, \dots, v_n) nazywamy zbiór $\text{lin}(v_1, \dots, v_n)$ wszystkich kombinacji liniowych tego układu.*

Uwaga 3.2.2 Wygodnie jest przyjąć, że jedyną kombinacją układu pustego (nie zawierającego żadnego wektora) jest wektor zerowy. W szczególności $\text{lin}(\emptyset) = \{\mathbf{0}\}$. \square

Uwaga 3.2.3 W definicji podprzestrzeni przestrzeni liniowej V warunki (+) i (\cdot) dla $W \subset V$ można zastąpić mocniejszym warunkiem

$$x_1 v_1 + \dots + x_n v_n \in W \quad \text{dla} \quad x_1, \dots, x_n \in \mathbb{K}, v_1, \dots, v_n \in W,$$

który wynika z (+) i (\cdot) przez indukcję ze względu na $n \geq 1$. \square

Twierdzenie 3.2.4 *Powłoka liniowa $\text{lin}(v_1, \dots, v_n)$ układu wektorów w przestrzeni V jest najmniejszą podprzestrzenią przestrzeni V zawierającą wektory v_j , $j = 1, \dots, n$.*

Dowód. Suma dwóch kombinacji liniowych wektorów v_1, \dots, v_n oraz wynik pomnożenia takiej kombinacji przez skalar jest kombinacją liniową wektorów v_1, \dots, v_n :

$$(+) \quad \sum_{j=1}^n x_j v_j + \sum_{j=1}^n y_j v_j = \sum_{j=1}^n (x_j + y_j) v_j; \quad (\cdot) \quad c \sum_{j=1}^n x_j v_j = \sum_{j=1}^n (c x_j) v_j.$$

Wynika stąd, że $\text{lin}(v_1, \dots, v_n)$ jest podprzestrzenią liniową V zawierającą wszystkie wektory v_j .

Z drugiej strony, jeśli podprzestrzeń liniowa W przestrzeni V zawiera v_1, \dots, v_n , to zawiera też wszystkie kombinacje liniowe tych wektorów, zob. 3.2.3, a więc $\text{lin}(v_1, \dots, v_n) \subset W$. ■

Iloczyn AX macierzy A i wektora X odpowiednich wymiarów wprowadziliśmy w 1.1, jednak ze względu na wagę tej operacji powtórzmy to określenie w sposób bardziej formalny.

Definicja 3.2.5 *Iloczynem macierzy $A = [A_1, \dots, A_n] \in \mathbb{K}_n^m$ (gdzie A_j jest j -tą kolumną A) i wektora $X \in \mathbb{K}^n$ o współrzędnych x_1, \dots, x_n nazywamy wektor $AX = \sum_{j=1}^n x_j A_j \in \mathbb{K}^m$.*

Uwaga 3.2.6 Operacja mnożenia macierzy i wektorów ma następujące własności (zob. dowód 3.2.4)

$$(+) \quad AX + AY = A(X + Y); \quad (\cdot) \quad c(AX) = A(cX),$$

tzn. w terminologii, którą uściślimy poniżej, operacja $X \rightarrow AX$ jest liniowa. □

Definicja 3.2.7 *Mówimy, że układ wektorów (v_1, \dots, v_n) z V rozpina V jeśli $V = \text{lin}(v_1, \dots, v_n)$.*

Uwaga 3.2.8 Układ wektorów (A_1, \dots, A_n) z przestrzeni \mathbb{K}^m rozpina \mathbb{K}^m wtedy i tylko wtedy, gdy dla każdego $B \in \mathbb{K}^m$ równanie $x_1 A_1 + \dots + x_n A_n = B$ jest niesprzeczne, a więc wtedy i tylko wtedy, gdy macierz otrzymana w wyniku redukcji $A = [A_1, \dots, A_n]$ do postaci schodkowej ma schodek w każdym wierszu. □

3.3 Liniowa niezależność.

Liniowa niezależność jest centralnym pojęciem związanym z przestrzeniami liniowymi.

Definicja 3.3.1 *Układ wektorów (v_1, \dots, v_k) w przestrzeni liniowej V nad ciałem \mathbb{K} nazywamy liniowo niezależnym jeśli z $a_1 v_1 + \dots + a_k v_k = \mathbf{0}$ wynika, że $a_1 = \dots = a_k = 0$. Układ, który nie jest liniowo niezależny nazywamy zależnym.*

Uwaga 3.3.2 Liniowa niezależność układu (v_1, \dots, v_k) oznacza, że każdy wektor $v \in \text{lin}(v_1, \dots, v_k)$ można zapisać w postaci kombinacji liniowej $v = a_1 v_1 + \dots + a_k v_k$ tylko w jeden sposób (później będziemy interpretowali współczynniki a_j jako współrzędne wektora v względem układu (v_1, \dots, v_k)). Istotnie, jeśli mamy także $v = b_1 v_1 + \dots + b_k v_k$, to $\mathbf{0} = (a_1 - b_1) v_1 + \dots + (a_k - b_k) v_k = \mathbf{0}$, a liniowa niezależność oznacza, że $\mathbf{0}$ może być zapisane tylko jako kombinacja liniowa v_j z zerowych współczynnikami. □

Twierdzenie 3.3.3 *Dla układu wektorów (v_1, \dots, v_k) w przestrzeni liniowej V nad ciałem \mathbb{K} następujące warunki są równoważne.*

- (i) *Układ (v_1, \dots, v_k) jest liniowo niezależny.*
- (ii) *Żaden z wektorów v_j nie jest kombinacją liniową pozostałych (to znaczy $v_j \notin \text{lin}(v_i)_{i \neq j}$ dla $j = 1, \dots, k$).*
- (iii) *Żaden z wektorów v_j nie jest kombinacją liniową poprzednich wektorów (to znaczy $v_1 \neq \mathbf{0}$ i $v_j \notin \text{lin}(v_1, \dots, v_{j-1})$ dla $j = 2, \dots, k$).*

Dowód. Dla dowodu implikacji (i) \Rightarrow (ii) założmy negację (ii), czyli istnienie $j \geq 1$ takiego, że $v_j = \sum_{i \neq j} a_i v_i$ dla pewnego układu skalarów $(a_i)_{i \neq j}$. Wtedy $-v_j + \sum_{i \neq j} a_i v_i = \mathbf{0}$ jest nietrywialnym przedstawieniem wektora zerowego, co przeczy (i).

Implikacja (ii) \Rightarrow (iii) jest oczywista.

Dla dowodu implikacji (iii) \Rightarrow (i) rozważmy kombinację $\sum_{i \leq k} a_i v_i = \mathbf{0}$. Gdyby nie wszystkie a_i były zerowe, to dla $j = \max\{i : a_i \neq 0\}$ mielibyśmy $v_j = \sum_{i < j} \frac{-a_i}{a_j} v_i$, co przeczyłoby (iii). ■

Uwaga 3.3.4 Liniowa niezależność układu (A_1, \dots, A_k) w \mathbb{K}^m oznacza, że równanie $\sum_{i=1}^k x_i A_i = \mathbf{0}$ ma dokładnie jedno rozwiązanie, czyli w wyniku redukcji macierzy $A = [A_1, \dots, A_k]$ do postaci schodkowej otrzymamy macierz A' , mającą schodek w każdej kolumnie (w szczególności $k \leq m$).

Równoważność warunków (i) oraz (iii) jest dla takiego układu oczywista, bo macierz A' ma schodek w j -tej kolumnie wtedy i tylko wtedy, gdy równanie $\sum_{i < j} x_i A_i = A_j$ jest sprzeczne, czyli $A_j \notin \text{lin}(A_i)_{i < j}$. □

3.4 Baza i wymiar.

Wyróżnienie n -elementowej bazy w przestrzeni liniowej V nad ciałem \mathbb{K} pozwala przypisać każdemu wektorowi $v \in V$ wektor z \mathbb{K}^n (wektor współrzędnych v w tej bazie) z zachowaniem operacji dodawania i mnożenia przez skalary.

Definicja 3.4.1 Układ wektorów (v_1, \dots, v_n) w przestrzeni liniowej V nad ciałem \mathbb{K} nazywamy bazą V jeśli układ (v_1, \dots, v_n) jest liniowo niezależny i rozpina V .

Uwaga 3.4.2 Jeśli układ (v_1, \dots, v_n) jest bazą V , to zgodnie z Uwagą 3.3.2, każdy wektor $v \in V$ daje się przedstawić jako kombinacja liniowa $v = a_1 v_1 + \dots + a_n v_n$ w dokładnie jeden sposób. Współczynniki tej kombinacji nazywamy *współrzędnymi wektora v w bazie (v_1, \dots, v_n)* . □

Przykład 3.4.3 (a) W przestrzeni \mathbb{K}^m położmy $E_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $E_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$, \dots , $E_m = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$.

Układ (E_1, E_2, \dots, E_m) jest bazą przestrzeni \mathbb{K}^m . Współrzędne wektora $X \in \mathbb{K}^m$ są identyczne ze współrzędnymi X w tej bazie. Bazę (E_1, E_2, \dots, E_m) nazywamy bazą standardową \mathbb{K}^m .

(b) W \mathbb{K}_2^2 położmy $E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$.

Układ $(E_{11}, E_{21}, E_{12}, E_{22})$ jest bazą przestrzeni \mathbb{K}_2^2 . Współrzędne macierzy $A \in \mathbb{K}_2^2$ w tej bazie są wyrazami tej macierzy w porządku, w jakim ustawiliśmy macierze E_{ij} .

Analogicznie w przestrzeni macierzy \mathbb{K}_n^m definiuje się bazę mającą $m \cdot n$ elementów $E_{kl} \in \mathbb{K}_n^m$, gdzie E_{kl} jest macierzą mającą na miejscu k, l jedynekę i wszystkie pozostałe wyrazy zerowe.

(c) Układ jednomianów (x^0, x^1, \dots, x^n) tworzy bazę podprzestrzeni $\mathbb{K}_n[x]$ wielomianów stopnia $\leq n$ przestrzeni $\mathbb{K}[x]$. Współrzędne wielomianu $w(x)$ w tej bazie są współczynnikami tego wielomianu.

Uwaga 3.4.4 Układ (A_1, \dots, A_n) w \mathbb{K}^m wyznacza macierz $A = [A_1, \dots, A_n] \in \mathbb{K}_n^m$. Jeśli w wyniku redukcji A do postaci schodkowej otrzymujemy macierz A' mającą schodki w kolumnach o numerach j_1, \dots, j_r , to układ $(A_{j_1}, \dots, A_{j_r})$ jest bazą $V = \text{lin}(A_1, \dots, A_n)$, bo dla każdego $B \in \mathbb{K}^m$ takiego, że układ $AX = B$ jest niesprzeczny, równanie $x_{j_1} A_{j_1} + \dots + x_{j_r} A_{j_r} = B$ ma dokładnie jedno rozwiązanie.

W szczególności, dla $n = m$ układ (A_1, \dots, A_n) w \mathbb{K}^n jest bazą \mathbb{K}^n wtedy i tylko wtedy, gdy macierz zredukowana A' ma n schodków (w każdej kolumnie i w każdym wierszu). □

Bazę $(A_{j_1}, \dots, A_{j_r})$ przestrzeni $\text{lin}(A_1, \dots, A_n) \subset \mathbb{K}^m$ otrzymujemy wybierając z układu (A_1, \dots, A_n) wektory, które nie są kombinacjami poprzednich, zob. Uwaga 3.3.4. Tak samo można postępować w przypadku ogólnym.

Twierdzenie 3.4.5 (o istnieniu bazy). *Jeśli z układu wektorów (v_1, \dots, v_n) rozpinającego przestrzeń V wybierzemy wszystkie wektory v_j takie, że $v_j \notin \text{lin}(v_i)_{i < j}$, to otrzymamy bazę $(v_{j_1}, \dots, v_{j_r})$ przestrzeni V .*

Dowód. Z Twierdzenia 3.3.3 (iii) układ $(v_{j_1}, \dots, v_{j_r})$ jest liniowo niezależny. Niech $W = \text{lin}(v_{j_1}, \dots, v_{j_r})$. Pokażemy, że $(v_{j_1}, \dots, v_{j_r})$ rozpinają V , czyli $W = V$. W tym celu wystarczy wykazać, że $v_i \in W$ dla $i \leq n$. Gdyby nie wszystkie v_i należały do W , to dla $j = \min\{i \leq n : v_i \notin W\}$ mielibyśmy $\text{lin}(v_i)_{i < j} \subset W$ oraz $v_j \notin W$. Zatem $v_j \notin \text{lin}(v_i)_{i < j}$, więc v_j byłby w W jako jeden z wybranych wektorów, co przeczy wyborowi j . ■

Z Twierdzenia 3.4.5 wynika.

Twierdzenie 3.4.6 (o rozszerzaniu układu liniowo niezależnego do bazy). *Jeśli układ wektorów (v_1, \dots, v_k) w przestrzeni liniowej V jest liniowo niezależny, a układ (w_1, \dots, w_m) rozpinają V , to układ (v_1, \dots, v_k) można rozszerzyć do bazy V wektorami z układu (w_1, \dots, w_m) .*

Dowód. Układ $(v_1, \dots, v_k, w_1, \dots, w_m)$ rozpinają V . Usuając z tego układu wszystkie wektory będące kombinacjami poprzednich otrzymamy, zgodnie z Twierdzeniem 3.4.5, bazę przestrzeni V , a z Twierdzenia 3.3.3 (iii) wynika, że nie usuniemy żadnego z wektorów v_j . ■

Zastosowane w tym dowodzie rozumowanie wykorzystamy też w dowodzie kolejnego twierdzenia, które pozwoli na określenie wymiaru przestrzeni liniowej.

Twierdzenie 3.4.7 (Steinitza o wymianie). *Jeśli układ wektorów (v_1, \dots, v_k) w przestrzeni liniowej V jest liniowo niezależny, a układ (w_1, \dots, w_m) rozpinają V , to $k \leq m$ oraz istnieją parami różne indeksy $i_1, \dots, i_k \leq m$ takie, że układ otrzymany z $(v_1, \dots, v_k, w_1, \dots, w_m)$ przez usunięcie wektorów w_{i_1}, \dots, w_{i_k} rozpinają V .*

Dowód. Nierówność $k \leq m$ wynika z drugiej części tezy, którą udowodnimy przez indukcję ze względu na $j \leq k$ dopisując na początku układu (w_1, \dots, w_m) kolejno wektory v_j i usuwając, za każdym razem, wektor w_{i_j} tak, by układ otrzymany po j wymianach pozostawał układem rozpinającym V .

W kroku indukcyjnym dodajemy do układu rozpinającego kolejny wektor v_j , bezpośrednio po wektorze v_{j-1} (v_1 wstawiamy na początku). Z warunku (ii) Twierdzenia 3.3.3 dostajemy układ liniowo zależny, a z warunku (iii) tego twierdzenia, jeden z pozostających w naszym układzie wektorów w_{i_j} jest kombinacją poprzednich wektorów, więc po jego usunięciu otrzymamy układ rozpinający V . ■

Przestrzeń liniowa może mieć wiele baz (zob. Uwaga 3.4.4). Jednakże z pierwszej części tezy twierdzenia Steinitza wynika, że w przestrzeni V z bazą mającą n wektorów, każdy układ liniowo niezależny ma $k \leq n$ wektorów, a każdy układ rozpinający ma $m \geq n$ wektorów. Tak więc, wszystkie bazy w V mają tyle samo elementów.

Definicja 3.4.8 *Wymiarem przestrzeni liniowej V mającej bazę skończoną nazywamy liczbę wektorów tej bazy, którą oznaczamy $\dim V$ ($\dim \{0\} = 0$). Jeśli V nie ma bazy skończonej, mówimy, że wymiar V jest nieskończony.*

Przykład 3.4.9 Z Przykładu 3.4.3 dostajemy

(a) $\dim \mathbb{K}^m = m,$

(b) $\dim \mathbb{K}_n^m = mn,$

(c) $\dim \mathbb{K}_n[x] = n + 1.$

Uwaga 3.4.10 Jeśli W jest podprzestrzenią przestrzeni V mającej skończony wymiar, to z twierdzenia Steinitza $\dim W \leq \dim V$. Co więcej, z $\dim W = \dim V$ wynika, że $W = V$, bo gdyby $W \neq V$, to bazę W można by było istotnie rozszerzyć do bazy V , zob. Twierdzenie 3.4.6. \square

Odnotujmy jednak, że przestrzenie wymiaru nieskończonego, na przykład $\mathbb{K}[x]$, mogą zawierać właściwe podprzestrzenie wymiaru nieskończonego. W dalszej części, jeśli nie powiemy wyraźnie inaczej, będziemy zakładać, że wszystkie rozważane przestrzenie mają wymiar skończony.

3.5 Rząd macierzy.

Z macierzą $A \in \mathbb{K}_n^m$ są związane trzy przestrzenie: podprzestrzeń rozpięta na kolumnach, podprzestrzeń rozpięta na wierszach i podprzestrzeń rozwiązań jednorodnego układu równań $AX = \mathbf{0}$. Dwie pierwsze mają taki sam wymiar – rząd macierzy A , a wymiar trzeciej jest różnicą n i rzędu A .

Przejdziemy teraz do systematycznego przedstawienia tych zagadnień.

Definicja 3.5.1 *Przestrzenią kolumn macierzy $A \in \mathbb{K}_n^m$ nazywamy podprzestrzeń $K(A)$ przestrzeni \mathbb{K}^m rozpiętą przez kolumny A .*

Z definicji mnożenia macierzy przez wektor (3.2.5) wynika, że przestrzeń kolumn $K(A)$ macierzy A jest zbiorem wszystkich wektorów B , dla których układ równań $AX = B$ jest niesprzeczny.

Definicja 3.5.2 *Rzędem $\text{rank } A$ macierzy A nazywamy $\dim K(A)$.*

Z Uwagi 3.4.4 wynika, że $\text{rank } A$ jest liczbą kolumn ze schodkami w macierzy A' otrzymanej w wyniku redukcji macierzy A do postaci schodkowej. Jeśli macierz \tilde{A} powstaje z A w wyniku operacji elementarnych na wierszach, to $K(\tilde{A})$ różni się na ogół od $K(A)$, ale $\text{rank } \tilde{A} = \text{rank } A$, bo \tilde{A} i A można zredukować do tej samej macierzy w postaci schodkowej.

Definicja 3.5.3 *Przestrzenią zerową macierzy A nazywamy podprzestrzeń $N(A)$ przestrzeni \mathbb{K}^n złożoną z rozwiązań jednorodnego układu równań $AX = \mathbf{0}$.*

Następne twierdzenie opisuje rozwiązania układu $AX = B$ w terminach zdefiniowanych wyżej pojęć.

Twierdzenie 3.5.4 (Kroneckera – Capelliego). *Niech $A \in \mathbb{K}_n^m$ i $B \in \mathbb{K}^m$. Układ równań $AX = B$ jest niesprzeczny wtedy i tylko wtedy, gdy $\text{rank } A = \text{rank } [A|B]$. Jeśli X_* jest rozwiązaniem tego układu, to zbiór wszystkich rozwiązań ma postać $X_* + N(A) = \{X_* + Z : Z \in N(A)\}$.*

Dowód. Pierwsza część tezy wynika z faktu, że niesprzeczność $AX = B$ jest równoważna warunkowi $B \in K(A)$. Druga część oznacza, że X jest rozwiązaniem wtedy i tylko wtedy, gdy $X - X_* \in N(A)$, a to wynika z równości $A(X - X_*) = AX - AX_* = AX - B$ (zob. wzory w Uwadze 3.2.6). \blacksquare

Opiszemy teraz wymiar $N(A)$ korzystając z faktu, że liczba zmiennych zależnych w rozwiązaniu ogólnym układu $AX = \mathbf{0}$ jest liczbą schodków macierzy A' otrzymanej w wyniku redukcji A do postaci schodkowej.

Twierdzenie 3.5.5 *Dla macierzy $A \in \mathbb{K}_n^m$, $\dim N(A) = n - \text{rank } A$.*

Dowód. Niech $p = n - \text{rank } A$ będzie liczbą zmiennych niezależnych układu $AX = \mathbf{0}$. Zgodnie z Uwagą 1.3.1 każde rozwiązanie tego układu jest wyznaczone przez wartości niezależnych parametrów t_1, \dots, t_p i ma postać $X = t_1 X_1 + \dots + t_p X_p$. Z Uwagi 3.3.2 wynika, że układ (X_1, \dots, X_p) jest bazą $N(A)$. \blacksquare

Podprzestrzeń V przestrzeni \mathbb{K}^n mająca bazę (A_1, \dots, A_r) jest przestrzenią kolumn macierzy $A = [A_1, \dots, A_r] \in \mathbb{K}_r^n$. Pokażemy, że V jest również przestrzenią zerową pewnej macierzy z \mathbb{K}_n^{n-r} .

Twierdzenie 3.5.6 *Jeśli $V \subset \mathbb{K}^n$, $\dim V = r$, to V jest przestrzenią zerową pewnej macierzy $C \in \mathbb{K}_n^{n-r}$.*

Dowód. Niech $A \in \mathbb{K}_r^n$ będzie macierzą, której kolumny tworzą bazę V , $V = K(A)$. Wektor $Y \in \mathbb{K}^n$ jest w V wtedy i tylko wtedy, gdy układ równań $AX = Y$ jest niesprzeczny, a więc gdy po redukcji macierzy $[A|Y]$ do postaci schodkowej otrzymana macierz $[A'|Y']$ nie ma schodka w ostatniej kolumnie. Ponieważ $\text{rank } A = \dim V = r$, to oznacza, że współrzędne Y' o indeksach $\geq r + 1$ są zerami.

Zauważmy, że wektor $Y \in \mathbb{K}^n$ jest jedynym rozwiązaniem układu równań $IX = Y$ o macierzy $I = [E_1, \dots, E_n] \in \mathbb{K}_n^n$, zob. Przykład 3.4.3 (a).

Wektory $Y \in V$ można więc opisać następująco. Niech I' będzie macierzą otrzymaną z macierzy I przez wykonanie na niej operacji elementarnych redukujących A do A' i niech $C \in \mathbb{K}_n^{n-r}$ będzie macierzą złożoną z ostatnich $n - r$ wierszy macierzy I' . Jedynym rozwiązaniem układu równań $I'X = Y'$ jest Y , bo ten układ jest równoważny układowi $IX = Y$, a więc $Y \in V$ wtedy i tylko wtedy, gdy $I'Y$ ma zera na ostatnich $n - r$ miejscach, tzn. gdy $CY = \mathbf{0}$. Zatem $V = N(C)$. ■

W praktyce macierz C układu równań opisującego przestrzeń $K(A) \subset \mathbb{K}^n$ wymiaru r wyznacza się redukując macierz $[A|I]$ do macierzy $[A'|I']$ takiej, że A' jest w postaci schodkowej (macierz C jest złożona z ostatnich $n - r$ wierszy macierzy I').

Wiersze macierzy $A \in \mathbb{K}_n^m$ należą do przestrzeni liniowej macierzy jednowierszowych \mathbb{K}_n^1 .

Definicja 3.5.7 *Przestrzenią wierszy macierzy $A \in \mathbb{K}_n^m$ nazywamy podprzestrzeń $W(A)$ przestrzeni \mathbb{K}_n^1 rozpiętą przez wiersze A .*

Twierdzenie 3.5.8 *Dla macierzy $A \in \mathbb{K}_n^m$, $\dim W(A) = \dim K(A)$.*

Dowód. Operacje elementarne na wierszach nie zmieniają przestrzeni wierszy. Jest to oczywiste dla operacji typu (II) i (III). Jeśli \tilde{A} powstaje z $A \in \mathbb{K}_n^m$ w wyniku zastosowania operacji $(I)_{a(i)+(k)}$, to oczywiście $W(\tilde{A}) \subset W(A)$. Równość wynika z faktu, że operacja $(I)_{(-a)(i)+(k)}$ prowadzi od \tilde{A} do A .

Wystarczy teraz pokazać, że dla macierzy A' w postaci schodkowej wymiar $\dim W(A')$ jest równy liczbie schodków tej macierzy. Istotnie, niezerowe wiersze (w'_1, \dots, w'_r) macierzy $W(A')$ są liniowo niezależne, bo po zmianie kolejności na $(w'_r, \dots, w'_2, w'_1)$ spełniają warunek (iii) Twierdzenia 3.3.3. ■

3.6 Suma prosta podprzestrzeni.

W klasie podprzestrzeni liniowych ustalonej przestrzeni liniowej są dwie naturalne operacje: przecięcia oraz sumy algebraicznej. Podamy pewne użyteczne fakty dotyczące tych operacji.

Uwaga 3.6.1 Jeśli V_1, V_2 są podprzestrzeniami przestrzeni liniowej V to część wspólna $V_1 \cap V_2 = \{v : v \in V_1 \text{ i } v \in V_2\}$ i suma $V_1 + V_2 = \{v_1 + v_2 : v_1 \in V_1, v_2 \in V_2\}$ są podprzestrzeniami V . □

Definicja 3.6.2 *Podprzestrzeń $V_1 + V_2$ przestrzeni V nazywamy sumą algebraiczną podprzestrzeni V_1, V_2 .*

Definicja 3.6.3 *Sumę algebraiczną $V_1 + V_2$ nazywamy sumą prostą jeśli dla dowolnie wybranych $v_j \in V_j$ z $v_1 + v_2 = \mathbf{0}$ wynika, że $v_1 = v_2 = \mathbf{0}$. Sumę prostą $V_1 + V_2$ oznaczamy $V_1 \oplus V_2$.*

Uwaga 3.6.4 Suma $V_1 + V_2$ jest sumą prostą wtedy i tylko wtedy, gdy każdy wektor $v \in V_1 + V_2$ daje się przedstawić jako suma $v = v_1 + v_2$, gdzie $v_j \in V_j$, w dokładnie jeden sposób (bo z $v = w_1 + w_2 = u_1 + u_2$ wynika, że $(w_1 - u_1) + (w_2 - u_2) = \mathbf{0}$). Wektory v_j nazywamy *składowymi wektora* $v \in V_1 \oplus V_2$. □

Twierdzenie 3.6.5 *Dla $V_1, V_2 \subset V$, $V_1 + V_2 = V_1 \oplus V_2$ wtedy i tylko wtedy, gdy $V_1 \cap V_2 = \{\mathbf{0}\}$.*

Dowód. Teza wynika z faktu, że $\mathbf{0} = v_1 + v_2 \in V_1 + V_2$ wtedy i tylko wtedy, gdy $v_2 = -v_1 \in V_1 \cap V_2$ ■

Twierdzenie 3.6.6 *Jeśli układ \mathcal{A}_j jest bazą przestrzeni $V_j \subset V$ dla $j = 1, 2$ i układ $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ powstaje przez dołączenie do \mathcal{A}_1 układu \mathcal{A}_2 , to układ \mathcal{A} jest bazą $V_1 + V_2$ wtedy i tylko wtedy, gdy $V_1 + V_2 = V_1 \oplus V_2$.*

Dowód. Układ \mathcal{A} oczywiście rozpiną $V_1 + V_2$. Każdy wektor $v_j \in V_j$ daje się jednoznacznie przedstawić jako kombinacją wektorów bazy \mathcal{A}_j , zob. Uwaga 3.3.2. Jednoznaczność rozkładu wektora $\mathbf{0}$ na składowe $v_j \in V_j$ jest więc równoważna z jednoznacznością zapisu $\mathbf{0}$ jako kombinacji wektorów układu \mathcal{A} . ■

Wniosek 3.6.7 $V_1 + V_2 = V_1 \oplus V_2$ wtedy i tylko wtedy, gdy $\dim V_1 + \dim V_2 = \dim(V_1 + V_2)$.

Ważną własnością przestrzeni liniowych jest fakt, że każdą podprzestrzeń przestrzeni liniowej można uzupełnić do sumy prostej.

Wniosek 3.6.8 *Dla dowolnej podprzestrzeni W przestrzeni liniowej V istnieje podprzestrzeń $U \subset V$ taka, że $W \oplus U = V$.*

Wniosek jest natychmiastową konsekwencją Twierdzenia 3.6.6 i twierdzenia o rozszerzaniu dowolnego układu liniowo niezależnego do bazy (zob. 3.4.6).

Wyprowadzimy stąd następującą formułę Grassmana.

Twierdzenie 3.6.9 *Jeśli V_1, V_2 są podprzestrzeniami przestrzeni liniowej V , to*

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

Dowód. Połóżmy $W = V_1 \cap V_2$ i niech U będzie podprzestrzenią V_2 taką, że $V_2 = W \oplus U$ ($U = V_2$, jeśli $W = \{\mathbf{0}\}$). Zauważmy, że $V_1 + V_2 = V_1 + U$, bo dla $v_1 + v_2 \in V_1 + V_2$ wektor $v_2 = w + u \in W \oplus U$, więc $v_1 + v_2 = v_1 + (w + u) = (v_1 + w) + u \in V_1 + U$.

Z $U \subset V_2$ mamy $V_1 \cap U = V_1 \cap V_2 \cap U = W \cap U = \{\mathbf{0}\}$, więc $V_1 + U = V_1 \oplus U$ z Twierdzenia 3.6.5. Z Wniosku 3.6.7 dostajemy $\dim(V_1 + V_2) = \dim(V_1 + U) = \dim V_1 + \dim U = \dim V_1 + \dim V_2 - \dim W$. ■

4 Przekształcenia liniowe

Obok przestrzeni liniowych, podstawowym obiektem algebry liniowej są przekształcenia liniowe. Rozpatrując przekształcenia liniowe między przestrzeniami liniowymi będziemy zawsze zakładać, że są to przestrzenie liniowe nad tym samym ciałem skalarów \mathbb{K} .

4.1 Przekształcenia liniowe.

Przekształcenia liniowe to funkcje między przestrzeniami liniowymi zgodne z ich strukturą algebraiczną. Dokładniej, przyjmujemy następującą definicję.

Definicja 4.1.1 Funkcję $T : V \rightarrow W$ nazywamy *przekształceniem liniowym* jeśli funkcja T jest addytywna i jednorodna (zachowuje dodawanie wektorów i mnożenie wektora przez skalar), to znaczy spełnione są dwa warunki

$$(+) \quad T(v_1 + v_2) = T(v_1) + T(v_2) \quad \text{dla } v_1, v_2 \in V; \quad (\cdot) \quad T(cv) = cT(v) \quad \text{dla } c \in \mathbb{K}, v \in V.$$

Uwaga 4.1.2 Jeśli $T : V \rightarrow W$ jest przekształceniem liniowym, to $T(\mathbf{0}_V) = T(0 \cdot \mathbf{0}_V) = 0T(\mathbf{0}_V) = \mathbf{0}_W$. \square

Identyczność $\text{id}_V : V \rightarrow V$, funkcja stałe równa zero $\mathbf{0} : V \rightarrow W$ (funkcja zerowa) i mnożenie przez niezerowy skalar $c \cdot \text{id}_V : V \rightarrow V$ są przekształceniami liniowymi.

Jak wyjaśnimy później, po ustaleniu baz w przestrzeniach liniowych, przekształcenia liniowe między tymi przestrzeniami można utożsamiać w naturalny sposób z macierzami. Na razie zauważmy, że macierze wyznaczają przekształcenia liniowe między przestrzeniami współrzędnych odpowiednich wymiarów.

Przykład 4.1.3 Macierz $A \in \mathbb{K}_n^m$ wyznacza przekształcenie liniowe $T_A : \mathbb{K}^n \rightarrow \mathbb{K}^m$ wzorem $T_A(X) = AX$ (zob. Uwaga 3.2.6).

Uwaga 4.1.4 Warunek zachowania dodawania wektorów i mnożenia wektora przez skalar można zastąpić warunkiem zachowywania kombinacji liniowych

$$T(x_1v_1 + \dots + x_nv_n) = x_1T(v_1) + \dots + x_nT(v_n) \quad \text{dla } x_1, \dots, x_n \in \mathbb{K}, v_1, \dots, v_n \in V,$$

który łatwo wyprowadza się z (+) i (\cdot) przez indukcję ze względu na $n \geq 1$. \square

Dowolna funkcja określona na bazie przestrzeni liniowej V o wartościach w przestrzeni liniowej W przedłuża się jednoznacznie do przekształcenia liniowego z V w W .

Twierdzenie 4.1.5 (o określaniu przekształceń liniowych na bazie). Niech V, W będą przestrzeniami liniowymi nad \mathbb{K} . Jeśli (v_1, \dots, v_n) jest bazą V , a (w_1, \dots, w_n) układem wektorów z W , to $T : V \rightarrow W$ określone formułą

$$T(x_1v_1 + \dots + x_nv_n) = x_1w_1 + \dots + x_nw_n$$

jest jedynym przekształceniem liniowym V w W takim, że $T(v_j) = w_j$ dla $j = 1, \dots, n$.

Dowód. Funkcja T jest dobrze określona, bo każdy wektor $v \in V$ jest kombinacją liniową wektorów bazy i współczynniki tej kombinacji są wyznaczone jednoznacznie. Z warunków (+), (\cdot) w dowodzie Twierdzenia 3.2.4 zastosowanych do obu stron formuły definiującej T wynika, że tak określone T jest przekształceniem liniowym.

Jednoznaczność wynika z Uwagi 4.1.4. \blacksquare

W szczególności, odnotujmy spostrzeżenie dotyczące przekształceń liniowych na sumach prostych (zob. Twierdzenie 3.6.6).

Uwaga 4.1.6 Jeśli $V = V_1 \oplus V_2$ oraz przekształcenia $T_i : V_i \rightarrow W$ są liniowe dla $i = 1, 2$, to istnieje dokładnie jedno przekształcenie liniowe $T : V \rightarrow W$ takie, że $T(v) = T_i(v)$ dla $v \in V_i$, $i = 1, 2$. Istotnie, dla wektora v mającego jednoznaczny rozkład na składowe $v = v_1 + v_2$ wystarczy zdefiniować $T(v) = T_1(v_1) + T_2(v_2)$. \square

Kończąc tę część, wskażemy dwa ważne typy przekształceń liniowych przestrzeni V w siebie.

Definicja 4.1.7 Niech $V = V_1 \oplus V_2$ i niech $v = v_1 + v_2$ będzie rozkładem $v \in V$ na składowe.

- (a) Przekształcenie liniowe $T : V \rightarrow V$ takie, że $T(v_1 + v_2) = v_1$ (T jest identycznością na V_1 i jest zerowe na V_2) nazywamy rzutem V na V_1 równoległym do V_2 .
- (b) Przekształcenie liniowe $S : V \rightarrow V$ takie, że $S(v_1 + v_2) = v_1 - v_2$ (S jest identycznością na V_1 i mnożeniem przez -1 na V_2) nazywamy symetrią V względem V_1 równoległą do V_2 ¹.

4.2 Jądro i obraz, izomorfizmy.

Przy opisie przekształcenia liniowego ważną rolę odgrywają dwie związane z nim podprzestrzenie liniowe: jądro i obraz.

Uwaga 4.2.1 Dla przekształcenia liniowego $T : V \rightarrow W$ i podprzestrzeni $V_0 \subset V$ oraz $W_0 \subset W$.

- (a) Obraz $T(V_0) = \{T(v) : v \in V_0\}$ podprzestrzeni $V_0 \subset V$ jest podprzestrzenią W .
- (b) Przeciwobraz $T^{-1}(W_0) = \{v : T(v) \in W_0\}$ podprzestrzeni $W_0 \subset W$ jest podprzestrzenią V . \square

Definicja 4.2.2 Niech $T : V \rightarrow W$ będzie przekształceniem liniowym.

- (a) Obrazem T nazywamy podprzestrzeń $\text{im } T = T(V) = \{T(v) : v \in V\}$ przestrzeni W . Wymiar obrazu $\dim \text{im } T$ nazywamy rzędem T i oznaczamy przez $\text{rank } T$.
- (b) Jądrem T nazywamy podprzestrzeń $\ker T = T^{-1}(\{\mathbf{0}\}) = \{v \in V : T(v) = \mathbf{0}\}$ przestrzeni V . Wymiar jądra $\dim \ker T$ nazywamy defektem T i oznaczamy przez $\text{def } T$.

Obraz i jądro przekształcenia liniowego wyznaczonego przez macierz mają ścisły związek z pojęciami wprowadzonymi w części 3.5.

Uwaga 4.2.3 Dla przekształcenia $T_A : \mathbb{K}^n \rightarrow \mathbb{K}^m$ wyznaczonego przez macierz $A \in \mathbb{K}_n^m$ (zob. Przykład 4.1.3) $\text{im } T_A = K(A)$, $\text{rank } T_A = \text{rank } A$, $\ker T_A = N(A)$ i z Twierdzenia 3.5.5, $\text{def } T_A = n - \text{rank } A$. \square

Następujące proste twierdzenie opisuje ważną własność przekształceń liniowych: trywialność jądra implikuje różnowartościowość.

Twierdzenie 4.2.4 Przekształcenie liniowe T jest różnowartościowe wtedy i tylko wtedy, gdy $\ker T = \{\mathbf{0}\}$.

Dowód. Niech $T : V \rightarrow W$. Dla $u, v \in V$ równość $T(u) = T(v)$ oznacza, że $T(u - v) = T(u) - T(v) = \mathbf{0}$, czyli $u - v \in \ker T$. Zatem $T(u) \neq T(v)$ dla różnych u, v wtedy i tylko wtedy, gdy $\ker T = \{\mathbf{0}\}$. \blacksquare

Wyróżnimy teraz trzy ważne klasy przekształceń liniowych.

Definicja 4.2.5 Przekształcenie liniowe $T : V \rightarrow W$ nazywamy

- (a) epimorfizmem jeśli $\text{im } T = W$,
- (b) monomorfizmem jeśli $\ker T = \{\mathbf{0}\}$,

¹zakładamy tu, że $-1 \neq 1$ w \mathbb{K} , czyli \mathbb{K} ma charakterystykę $\neq 2$

(c) izomorfizmem liniowym jeśli T jest epimorfizmem i monomorfizmem.

Twierdzenie 4.2.6 Funkcja odwrotna T^{-1} do izomorfizmu liniowego $T : V \longrightarrow W$ jest przekształceniem liniowym $T^{-1} : W \longrightarrow V$.

Dowód. Istnienie funkcji odwrotnej T^{-1} wynika z Twierdzenia 4.2.4. Dla sprawdzenia, że T^{-1} zachowuje dodawanie wektorów weźmy $w_i = T(v_i) \in W$ dla $i = 1, 2$. Wtedy $T(v_1 + v_2) = w_1 + w_2$ i przykładając do obu stron tej równości T^{-1} otrzymujemy $v_1 + v_2 = T^{-1}(w_1 + w_2)$, czyli $T^{-1}(w_1) + T^{-1}(w_2) = T^{-1}(w_1 + w_2)$. Analogicznie sprawdza się, że T^{-1} zachowuje mnożenie wektora przez skalar. ■

Z twierdzenia o określaniu przekształceń liniowych na bazie wynika, że własności przekształcenia liniowego $T : V \longrightarrow W$ są wyznaczone przez układ $(T(v_1), \dots, T(v_n))$ obrazów wektorów ustalonej bazy (v_1, \dots, v_n) przestrzeni V .

Twierdzenie 4.2.7 Niech $T : V \longrightarrow W$ będzie przekształceniem liniowym określonym na przestrzeni V z bazą (v_1, \dots, v_n) . Wtedy:

- (a) T jest epimorfizmem wtedy i tylko wtedy, gdy układ $(T(v_1), \dots, T(v_n))$ rozpiną W ;
- (b) T jest monomorfizmem wtedy i tylko wtedy, gdy układ $(T(v_1), \dots, T(v_n))$ jest liniowo niezależny;
- (c) T jest izomorfizmem wtedy i tylko wtedy, gdy układ $(T(v_1), \dots, T(v_n))$ jest bazą W .

Dowód. Część (a) wynika z równości $\text{im } T = T(\text{lin}(v_1, \dots, v_n)) = \text{lin}(T(v_1), \dots, T(v_n))$, część (b) z równoważności $\sum_{j=1}^n x_j T(v_j) = \mathbf{0} \Leftrightarrow \sum_{j=1}^n x_j v_j \in \ker T$, a część (c) jest konsekwencją (a) i (b). ■

Mówimy, że przestrzenie liniowe V, W nad \mathbb{K} są *izomorficzne* jeśli istnieje izomorfizm liniowy V na W . Z części (c) i z twierdzenia o określaniu przekształceń liniowych na bazie wynika następujący wniosek.

Wniosek 4.2.8 Przestrzenie liniowe V i W są izomorficzne wtedy i tylko wtedy, gdy $\dim V = \dim W$.

Szczególnie ważną rolę pełnią izomorfizmy przestrzeni V na przestrzeń współrzędnych wymiaru $\dim V$ – układy współrzędnych. Właściwy dobór układu współrzędnych znacznie upraszcza analizę wielu zagadnień algebry liniowej.

Ostatnie twierdzenie tej części można, przechodząc do przestrzeni współrzędnych, wyprowadzić z Twierdzenia 3.5.5 (por. Uwaga 4.2.3). Podamy jednak bezpośredni dowód, a systematyczne wykorzystanie układów współrzędnych poprzedzimy analizą przekształceń liniowych na przestrzeniach współrzędnych.

Twierdzenie 4.2.9 Jeżeli $T : V \longrightarrow W$ jest przekształceniem liniowym, to $\dim V = \text{def } T + \text{rank } T$.

Dowód. Niech U będzie podprzestrzenią V taką, że $V = \ker T \oplus U$ (zob. Wniosek 3.6.8) i niech $S = T|_U : U \longrightarrow W$ będzie obcięciem T do U ($S(u) = T(u)$ dla $u \in U$). Wtedy $\text{im } S = \text{im } T$, bo dla $v = w + u \in \ker T \oplus U$ mamy $T(v) = T(w) + T(u) = S(u)$.

Z Twierdzenia 3.6.5 $\ker T \cap U = \{\mathbf{0}\}$, więc S jest izomorfizmem U na $\text{im } T$ i z Wniosku 3.6.7 dostajemy $\dim V = \text{def } T + \dim U = \text{def } T + \text{rank } T$. ■

4.3 Przekształcenia liniowe przestrzeni współrzędnych.

Przekształcenie liniowe $T : \mathbb{K}^n \rightarrow \mathbb{K}^m$ jest jednoznacznie wyznaczone przez układ $(T(E_1), \dots, T(E_n))$ wartości T na wektorach bazy standardowej przestrzeni \mathbb{K}^n (zob. Twierdzenie 4.1.5).

Definicja 4.3.1 *Macierzą przekształcenia liniowego $T : \mathbb{K}^n \rightarrow \mathbb{K}^m$ nazywamy macierz $M(T) \in \mathbb{K}_n^m$ postaci $M(T) = [T(E_1), \dots, T(E_n)]$, gdzie (E_1, \dots, E_n) jest bazą standardową \mathbb{K}^n .*

Następne twierdzenie ustala podstawowe związki między przekształceniem liniowym przestrzeni współrzędnych i jego macierzą.

Twierdzenie 4.3.2 *Jeśli $T : \mathbb{K}^n \rightarrow \mathbb{K}^m$ jest przekształceniem liniowym, to $T(X) = M(T)X$ dla $X \in \mathbb{K}^n$. Co więcej:*

- (a) T jest epimorfizmem wtedy i tylko wtedy, gdy $\text{rank } M(T) = m$;
- (b) T jest monomorfizmem wtedy i tylko wtedy, gdy $\text{rank } M(T) = n$;
- (c) T jest izomorfizmem wtedy i tylko wtedy, gdy $\text{rank } M(T) = m = n$.

Dowód. Jeśli $X = \sum_{j=1}^n x_j E_j \in \mathbb{K}^n$, to $T(X) = T(\sum_{j=1}^n x_j E_j) = \sum_{j=1}^n x_j T(E_j) = [T(E_1), \dots, T(E_n)]X$. Druga część tezy wynika z Twierdzenia 4.2.7. ■

Z pierwszej części tezy wynika, że przyporządkowanie przekształceniu $T : \mathbb{K}^n \rightarrow \mathbb{K}^m$ jego macierzy $M(T) = [T(E_1), \dots, T(E_n)] \in \mathbb{K}_n^m$ jest operacją odwrotną do opisanego w Przykładzie 4.1.3 przyporządkowania macierzy $A \in \mathbb{K}_n^m$ przekształcenia $T_A : \mathbb{K}^n \rightarrow \mathbb{K}^m$.

Zdefiniujemy teraz operację mnożenia macierzy odpowiadającą składaniu przekształceń. Jeśli macierz $B = [B_1, \dots, B_k] \in \mathbb{K}_k^n$ wyznacza $T_B : \mathbb{K}^k \rightarrow \mathbb{K}^n$ (czyli $T_B(E_l) = B_l$ dla $l = 1, \dots, k$), a macierz $A \in \mathbb{K}_n^m$ wyznacza $T_A : \mathbb{K}^n \rightarrow \mathbb{K}^m$, to złożenie $T_A \circ T_B : \mathbb{K}^k \rightarrow \mathbb{K}^m$ jest przekształceniem liniowym i

$$M(T_A \circ T_B) = [T_A \circ T_B(E_1), \dots, T_A \circ T_B(E_k)] = [T_A(B_1), \dots, T_A(B_k)] = [AB_1, \dots, AB_k].$$

Definicja 4.3.3 *Wynikiem pomnożenia macierzy $A \in \mathbb{K}_n^m$ przez macierz $B = [B_1, \dots, B_k] \in \mathbb{K}_k^n$ jest macierz $AB = [AB_1, \dots, AB_k] \in \mathbb{K}_k^m$.*

Definicja 4.3.4 *Macierz $I_n = M(\text{id}_{\mathbb{K}^n}) = [E_1, \dots, E_n]$ nazywamy macierzą jednostkową.*

Uwaga 4.3.5 (a) Podobnie do złożenia funkcji $f \circ g$, które jest określone tylko wtedy, gdy dziedzina f jest przeciwdziedziną g , iloczyn macierzy AB ma sens tylko wtedy, gdy liczba kolumn A jest taka jak liczba wierszy B . Mówiąc o iloczynie macierzy zawsze zakładamy zgodność odpowiednich wymiarów.

(b) Jeśli $A \in \mathbb{K}_n^m$, to $AI_n = A = I_m A$ (bo $T_A \circ \text{id}_{\mathbb{K}^n} = T_A = \text{id}_{\mathbb{K}^m} \circ T_A$).

(c) Mnożenie macierzy jest łączne, $A(BC) = (AB)C$, co wynika z łączności składania funkcji.

(d) Mnożenie macierzy nie zawsze jest przemienne (nawet wtedy, gdy zmiana kolejności czynników ma sens).

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

(e) Iloczyn macierzy niezerowych może być macierzą zerową $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. □

4.4 Przestrzenie przekształceń liniowych.

Niech V i W będą przestrzeniami liniowymi nad \mathbb{K} . Zbiór $L(V, W)$ przekształceń liniowych z V do W będziemy rozpatrywać jako przestrzeń liniową nad \mathbb{K} , z przekształceniem zerowym $\mathbf{0}$ jako wektorem zerowym oraz naturalnymi działaniami dodawania i mnożenia przez skalary określonymi następująco: dla $T_1, T_2, T \in L(V, W)$, $c \in \mathbb{K}$ i $v \in V$

$$(+) \quad (T_1 + T_2)(v) = T_1(v) + T_2(v); \quad (\cdot) \quad (cT)(v) = c(T(v)).$$

Następujące twierdzenie pozwala utożsamić, z zachowaniem operacji algebraicznych, przestrzenie przekształceń $L(\mathbb{K}^n, \mathbb{K}^m)$ z przestrzeniami macierzy \mathbb{K}_n^m opisanymi w Przykładzie 3.1.3 (b).

Twierdzenie 4.4.1 *Przyporządkowanie przekształceniu liniowemu $T : \mathbb{K}^n \rightarrow \mathbb{K}^m$ jego macierzy $M(T)$ jest izomorfizmem liniowym przestrzeni $L(\mathbb{K}^n, \mathbb{K}^m)$ na przestrzeń macierzy \mathbb{K}_n^m .*

Dowód. Widzieliśmy, że przyporządkowanie macierzy $A \in \mathbb{K}_n^m$ przekształcenia $T_A \in L(\mathbb{K}^n, \mathbb{K}^m)$ jest odwróceniem funkcji $T \rightarrow M(T)$. Sprawdzimy, że ta funkcja zachowuje dodawanie i mnożenie przez skalar, czyli dla $T_1, T_2, T \in L(V, W)$, $c \in \mathbb{K}$ spełnione są warunki

$$(+) \quad M(T_1 + T_2) = M(T_1) + M(T_2); \quad (\cdot) \quad M(cT) = cM(T).$$

Istotnie, $M(T_1 + T_2) = [(T_1 + T_2)(E_1), \dots, (T_1 + T_2)(E_n)] = [T_1(E_1) + T_2(E_1), \dots, T_1(E_n) + T_2(E_n)] = [T_1(E_1), \dots, T_1(E_n)] + [T_2(E_1), \dots, T_2(E_n)] = M(T_1) + M(T_2)$ oraz $M(cT) = [(cT)(E_1), \dots, (cT)(E_n)] = [cT(E_1), \dots, cT(E_n)] = c[T(E_1), \dots, T(E_n)] = cM(T)$. ■

Uwaga 4.4.2 Przyporządkowanie przekształceniu liniowemu przestrzeni współrzędnych jego macierzy przeprowadza operację złożenia przekształceń na mnożenie macierzy, więc z łatwych do sprawdzenia własności przekształceń natychmiast wynikają następujące algebraiczne własności mnożenia macierzy (zakładamy zgodność wymiarów macierzy w odpowiednich działaniach)

$$(a) \quad (A_1 + A_2)B = A_1B + A_2B,$$

$$(b) \quad A(B_1 + B_2) = AB_1 + AB_2,$$

$$(c) \quad A(cB) = c(AB) = (cA)B. \quad \square$$

4.5 Izomorfizmy przestrzeni współrzędnych.

Macierz mającą m wierszy i m kolumn nazywamy *macierzą kwadratową*. Macierze kwadratowe z \mathbb{K}_m^m odpowiadają przekształceniom $T : \mathbb{K}^m \rightarrow \mathbb{K}^m$, przy czym macierze odpowiadające izomorfizmom są elementami odwracalnymi w \mathbb{K}_m^m , ze względu na operację mnożenia.

Definicja 4.5.1 *Macierz kwadratową $A \in \mathbb{K}_m^m$ nazywamy macierzą odwracalną, jeśli istnieje macierz kwadratowa $M \in \mathbb{K}_m^m$ taka, że $MA = I_m$*

Uwaga 4.5.2 Dla macierzy kwadratowych $A, M \in \mathbb{K}_m^m$ warunek $MA = I_m$ oznacza, że $T_M : \mathbb{K}^m \rightarrow \mathbb{K}^m$ jest epimorfizmem, a $T_A : \mathbb{K}^m \rightarrow \mathbb{K}^m$ jest monomorfizmem. Z Twierdzenia 4.3.2 wynika, że $\text{rank } M = \text{rank } A = m$, czyli T_A i T_M są wzajemnie odwrotnymi izomorfizmami \mathbb{K}^m oraz $AM = MA = I_m$. □

Definicja 4.5.3 *Jeśli macierz kwadratowa $A \in \mathbb{K}_m^m$ jest odwracalna, to macierz $M(T_A^{-1})$ izomorfizmu odwrotnego do T_A nazywamy macierzą odwrotną do A i oznaczamy przez A^{-1} .*

Podamy teraz metodę znajdowania macierzy odwrotnej korzystając z interpretacji operacji elementarnych na wierszach macierzy jako pewnych izomorfizmów przestrzeni współrzędnych.

Niech \mathcal{E} będzie operacją elementarną na wierszach macierzy z \mathbb{K}_n^m . Wynik operacji \mathcal{E} na macierzy A będziemy oznaczać przez $\mathcal{E}(A)$, niezależnie od liczby kolumn macierzy A (także dla macierzy jednokolumnowych). W szczególności, dla macierzy $A = [A_1, \dots, A_n]$ mamy $\mathcal{E}(A) = [\mathcal{E}(A_1), \dots, \mathcal{E}(A_n)]$.

Operacja \mathcal{E} na macierzach jednokolumnowych jest odwracalną funkcją $\mathcal{E} : \mathbb{K}^m \rightarrow \mathbb{K}^m$ zachowującą kombinacje liniowe (jeśli $x_1 A_1 + \dots + x_n A_n = B$, to $x_1 \mathcal{E}(A_1) + \dots + x_n \mathcal{E}(A_n) = \mathcal{E}(B)$, zob. Twierdzenie 1.2.2)

Definicja 4.5.4 *Izomorfizm $\mathcal{E} : \mathbb{K}^m \rightarrow \mathbb{K}^m$ wyznaczony przez operację elementarną na wierszach \mathcal{E} nazywamy izomorfizmem elementarnym, a jego macierz $M(\mathcal{E}) = [\mathcal{E}(E_1), \dots, \mathcal{E}(E_m)] = \mathcal{E}(I_m)$ nazywamy macierzą elementarną.*

Uwaga 4.5.5 Wykonanie operacji elementarnej \mathcal{E} na wierszach macierzy $A = [A_1, \dots, A_n] \in \mathbb{K}_n^m$ daje macierz $\mathcal{E}(A) = [\mathcal{E}(A_1), \dots, \mathcal{E}(A_n)] = [\mathcal{E}(T_A(E_1)), \dots, \mathcal{E}(T_A(E_n))] = M(\mathcal{E} \circ T_A) = M(\mathcal{E})A$, czyli odpowiada pomnożeniu macierzy A z lewej strony przez macierz elementarną $M(\mathcal{E})$ izomorfizmu \mathcal{E} . \square

Twierdzenie 4.5.6 *Macierz odwracalną $A \in \mathbb{K}_m^m$ można zredukować do macierzy jednostkowej I_m operacjami elementarnymi na wierszach. Jeśli $\mathcal{E}_p, \dots, \mathcal{E}_1$ są operacjami redukującymi A do I_m , to iloczyn macierzy elementarnych $M(\mathcal{E}_p) \cdot \dots \cdot M(\mathcal{E}_1)$ jest macierzą odwrotną do A .*

Dowód. Macierz A ma rząd m , więc redukując A do postaci schodkowej otrzymamy macierz A' mającą na przekątnej wyrazy niezerowe i zera pod przekątną. Wykonując operacje typu (I) z użyciem ostatniego wiersza macierzy A' można wyzerować wszystkie, prócz ostatniego, wyrazy ostatniej kolumny tej macierzy. Wykorzystując następnie przedostatni wiersz, w podobny sposób można wyzerować wszystkie wyrazy przedostatniej kolumny leżące w poprzednich wierszach i po kolejnych, analogicznych krokach otrzymać macierz diagonalną B (czyli macierz mającą zera poza przekątną). Na koniec, operacjami typu (III) można zamienić wszystkie wyrazy przekątnej B na jedynki.

Jeśli $\mathcal{E}_p, \dots, \mathcal{E}_1$ są operacjami redukującymi macierz A do I_m , to złożenie $T = \mathcal{E}_p \circ \dots \circ \mathcal{E}_1$ izomorfizmów elementarnych przeprowadza j -tą kolumnę macierzy A na j -tą kolumnę macierzy I_m , jest więc izomorfizmem odwrotnym do izomorfizmu T_A , a jego macierz $M(T) = M(\mathcal{E}_p) \cdot \dots \cdot M(\mathcal{E}_1)$ jest macierzą odwrotną do A . \blacksquare

Jeśli macierz $A \in \mathbb{K}_m^m$ jest odwracalna, to macierz A^{-1} też jest odwracalna i $(A^{-1})^{-1} = A$, więc z drugiej części tezy zastosowanej do macierzy A^{-1} , otrzymujemy następujący wniosek.

Wniosek 4.5.7 *Macierz odwracalna $A \in \mathbb{K}_m^m$ jest iloczynem skończenie wielu macierzy elementarnych.*

Uwaga 4.5.8 Niech $[A|I_m] \in \mathbb{K}_{m+m}^m$ będzie macierzą powstałą przez dopisanie do macierzy odwracalnej $A \in \mathbb{K}_m^m$ macierzy jednostkowej I_m . Redukując macierz $[A|I_m]$ do macierzy $[I_m|M]$ operacjami elementarnymi $\mathcal{E}_p, \dots, \mathcal{E}_1$ otrzymujemy w dopisanej części macierz złożenia $\mathcal{E}_p \circ \dots \circ \mathcal{E}_1$ redukujących izomorfizmów elementarnych, a więc $M = M(\mathcal{E}_p \circ \dots \circ \mathcal{E}_1) = A^{-1}$. \square

Przy odwracaniu iloczynu macierzy musimy zmienić kolejność czynników.

Twierdzenie 4.5.9 *Jeśli macierze $A, B \in \mathbb{K}_m^m$ są odwracalne i $c \in \mathbb{K}$ jest niezerowym skalarzem, to iloczyny AB i cA są macierzami odwracalnymi, $(AB)^{-1} = B^{-1}A^{-1}$ oraz $(cA)^{-1} = c^{-1}A^{-1}$.*

Dowód. Mnożenie macierzy jest łączne, więc $(B^{-1}A^{-1})AB = B^{-1}(A^{-1}A)B = B^{-1}B = I_m$. Analogicznie, $c^{-1}A^{-1}(cA) = c^{-1}cA^{-1}A = I_m$. \blacksquare

4.6 Macierz przekształcenia.

Opiszemy teraz przejście od dowolnych przestrzeni liniowych do przestrzeni współrzędnych, odwołując się do istnienia izomorfizmu przestrzeni liniowej na przestrzeń współrzędnych odpowiedniego wymiaru.

Dla wyróżnienia takich izomorfizmów, będziemy je oznaczali greckimi literami σ, τ .

Definicja 4.6.1 *Izomorfizmy n -wymiarowej przestrzeni V nad \mathbb{K} na przestrzeń \mathbb{K}^n nazywamy układami współrzędnych w V . Układem współrzędnych związanym z bazą (v_1, \dots, v_n) w V nazywamy izomorfizm $\sigma : V \rightarrow \mathbb{K}^n$ przeprowadzający v_j na E_j , $j = 1, \dots, n$.*

Uwaga 4.6.2 (a) Układ współrzędnych σ związany z bazą (v_1, \dots, v_n) przestrzeni V przeprowadza wektor $v = \sum_{j=1}^n x_j v_j \in V$ na wektor $X = \sum_{j=1}^n x_j E_j \in \mathbb{K}^n$ współrzędnych v w tej bazie.

(b) Dowolny układ współrzędnych $\sigma : V \rightarrow \mathbb{K}^n$ jest związany z bazą $(\sigma^{-1}(E_1), \dots, \sigma^{-1}(E_n))$.

(c) Jeśli $\sigma, \sigma' : V \rightarrow \mathbb{K}^n$ są układami współrzędnych w V , to złożenie $\sigma' \circ \sigma^{-1} : \mathbb{K}^n \rightarrow \mathbb{K}^n$ zmienia współrzędne $\sigma(v)$ wektora v w σ na współrzędne v w σ' (bo $\sigma' \circ \sigma^{-1}(\sigma(v)) = \sigma'(v)$). \square

Pokażemy teraz, jak wybór układów współrzędnych w V i W pozwala przyporządkować każdemu przekształceniu liniowemu $T : V \rightarrow W$ jego macierz w tych układach współrzędnych.

Definicja 4.6.3 *Niech $\sigma : V \rightarrow \mathbb{K}^n$ będzie układem współrzędnych w V , a $\tau : W \rightarrow \mathbb{K}^m$ układem współrzędnych w W . Macierzą przekształcenia liniowego $T : V \rightarrow W$ w układach σ, τ nazywamy macierz $M_\sigma^\tau(T) = M(\tau \circ T \circ \sigma^{-1})$, gdzie $\tau \circ T \circ \sigma^{-1} : \mathbb{K}^n \rightarrow \mathbb{K}^m$.*

Uwaga 4.6.4 Niech $T : V \rightarrow W$ będzie przekształceniem liniowym określonym na przestrzeni V z bazą (v_1, \dots, v_n) związaną z układem współrzędnych $\sigma : V \rightarrow \mathbb{K}^n$ i niech $\tau : W \rightarrow \mathbb{K}^m$ będzie układem współrzędnych w W . Wtedy:

(a) $M_\sigma^\tau(T) = [\tau(T(v_1)), \dots, \tau(T(v_n))]$ (bo $\tau \circ T \circ \sigma^{-1}(E_j) = \tau(T(v_j))$ dla $j = 1, \dots, n$);

(b) $M_\sigma^\tau(T) \cdot \sigma(v) = \tau(T(v))$ dla $v \in V$ (bo $\tau \circ T \circ \sigma^{-1}(\sigma(v)) = \tau(T(v))$). \square

Dla ustalonych układów współrzędnych $\sigma : V \rightarrow \mathbb{K}^n$ i $\tau : W \rightarrow \mathbb{K}^m$ przyporządkowanie przekształceniu liniowemu $T \in L(V, W)$ złożenia $\tau \circ T \circ \sigma^{-1} : \mathbb{K}^n \rightarrow \mathbb{K}^m$ jest izomorfizmem liniowym $L(V, W)$ na $L(\mathbb{K}^n, \mathbb{K}^m)$. Z Twierdzenia 4.4.1 wynika więc natychmiast następujący fakt.

Twierdzenie 4.6.5 *Dla ustalonych układów współrzędnych $\sigma : V \rightarrow \mathbb{K}^n$ i $\tau : W \rightarrow \mathbb{K}^m$ przyporządkowanie przekształceniu liniowemu $T \in L(V, W)$ jego macierzy $M_\sigma^\tau(T)$ jest izomorfizmem liniowym przestrzeni przekształceń $L(V, W)$ na przestrzeń macierzy \mathbb{K}_n^m .*

4.7 Przestrzeń funkcjonałów liniowych.

Niech V będzie n -wymiarową przestrzenią liniową nad ciałem \mathbb{K} . Przestrzeń $L(V, \mathbb{K})$ nazywamy *przestrzenią sprzężoną do przestrzeni V* , a jej elementy – przekształcenia liniowe $f : V \rightarrow \mathbb{K}$, nazywamy *funkcjonałami liniowymi*.

Uwaga 4.7.1 Niech $\mathcal{A} = (v_1, \dots, v_n)$ będzie bazą przestrzeni V i niech $f_i : V \rightarrow \mathbb{K}$ będzie jedynym funkcjonałem liniowym takim, że

$$(*) \quad f_i(v_j) = \begin{cases} 1 & \text{jeśli } i = j, \\ 0 & \text{jeśli } i \neq j. \end{cases}$$

Wówczas:

(a) $v = \sum_j f_j(v)v_j$ dla $v \in V$, czyli f_i przyporządkowuje wektorowi jego i -tą współrzędną w bazie \mathcal{A} , bo, dla $v = \sum_j x_j v_j$, $f_i(v) = f_i(\sum_j x_j v_j) = \sum_j x_j f_i(v_j) = x_i$;

- (b) $f = \sum_i f(v_i)f_i$ dla $f \in V^*$, przy czym przedstawienie jest jednoznaczne, bo, obie strony równości przyjmują na wektorze v_j bazy \mathcal{A} wartość $f(v_j)$, więc są równe i analogiczny argument pokazuje jednoznaczność przedstawienia f jako sumy $f = \sum_i a_i f_i$;
- (c) układ funkcjonałów $\mathcal{A}^* = (f_1, \dots, f_n)$ jest bazą V^* i wartość funkcjonału $f \in V^*$ na wektorze v_j jest j -tą współrzędną tego funkcjonału w bazie \mathcal{A}^* . \square

Definicja 4.7.2 Bazy (v_1, \dots, v_n) w V i (f_1, \dots, f_n) w V^* nazywamy *dualnymi*, jeśli spełniony jest warunek (*) w Uwadze 4.7.1.

Twierdzenie 4.7.3 Dla każdej bazy (f_1, \dots, f_n) w przestrzeni V^* istnieje dualna do niej baza (v_1, \dots, v_n) w przestrzeni V .

Dowód. Niech $V^{**} = (V^*)^*$ i niech $\mathcal{J} : V \rightarrow V^{**}$ będzie przekształceniem liniowym określonym wzorem

$$\mathcal{J}(v)(f) = f(v), \quad \text{dla } v \in V \text{ i } f \in V^*.$$

Jeśli $v \neq \mathbf{0}$, to istnieje $f \in V^*$ takie, że $f(v) \neq \mathbf{0}$, a więc $\mathcal{J}(v) \neq \mathbf{0}$ (w V^{**}). Zatem \mathcal{J} jest monomorfizmem i ponieważ (zob. Uwaga 4.7.1) $\dim V = \dim V^* = \dim V^{**}$, \mathcal{J} jest izomorfizmem.

Zgodnie z 4.7.1, dla bazy (f_1, \dots, f_n) przestrzeni V^* istnieje baza (u_1, \dots, u_n) w V^{**} spełniająca warunek

$$(**) \quad u_i(f_j) = \begin{cases} 1 & \text{jeśli } i = j, \\ 0 & \text{jeśli } i \neq j. \end{cases}$$

Jeśli (v_1, \dots, v_n) jest bazą V taką, że $\mathcal{J}(v_i) = u_i$, to $u_i(f_j) = \mathcal{J}(v_i)(f_j) = f_j(v_i)$, więc z (**) wynika, że bazy (v_1, \dots, v_n) oraz (f_1, \dots, f_n) są dualne. \blacksquare

Definicja 4.7.4 Niech $T : V \rightarrow W$ będzie przekształceniem liniowym. Przekształceniem sprzężonym do T nazywamy przekształcenie liniowe $T^* : W^* \rightarrow V^*$ określone formułą

$$T^*(g) = g \circ T, \quad \text{dla } g \in W^*.$$

Pokażemy, że przy wyborze baz dualnych V, V^* oraz W, W^* , macierze przekształceń T i T^* w związanych z tymi bazami układach współrzędnych powstają, jedna z drugiej, przez zamianę kolumn na wiersze, tzn. przez transponowanie.

Zacznijmy od wprowadzenia operacji transpozycji macierzy.

Definicja 4.7.5 Macierz transponowaną macierzy $A \in \mathbb{K}_n^m$ nazywamy macierz $A^T \in \mathbb{K}_m^n$, której kolejne kolumny są kolejnymi wierszami macierzy A .

Uwaga 4.7.6 Dla macierzy $A \in \mathbb{K}_n^m$:

- (a) $(A^T)^T = A$ (bo zamiana kolumn na wiersze zmienia wiersze na kolumny);
- (b) $\text{rank } A^T = \text{rank } A$ (zob. Twierdzenie 3.5.8);
- (c) $(AB)^T = B^T A^T$ (jeśli iloczyn AB ma sens, to iloczyn $B^T A^T$ również ma sens, a równość można sprawdzić porównując odpowiednie wyrazy tych iloczynów);
- (d) $(A^T)^{-1} = (A^{-1})^T$ jeśli A jest macierzą odwracalną (bo $(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$). \square

Twierdzenie 4.7.7 Niech $T : V \rightarrow W$ będzie przekształceniem liniowym i niech $\sigma : V \rightarrow \mathbb{K}^n$, $\tau : W \rightarrow \mathbb{K}^m$ będą układami współrzędnych związanymi z pewnymi bazami V i W , a $\hat{\sigma} : V^* \rightarrow \mathbb{K}^n$, $\hat{\tau} : W^* \rightarrow \mathbb{K}^m$ będą układami współrzędnych związanymi z bazami dualnymi do nich. Wówczas

$$M_{\hat{\tau}}^{\hat{\sigma}}(T^*) = M_{\sigma}^{\tau}(T)^T.$$

Dowód. Niech σ i τ będą związane z bazami (v_1, \dots, v_n) w V i (w_1, \dots, w_m) w W , a układy $\hat{\sigma}$ i $\hat{\tau}$, z bazami dualnymi (f_1, \dots, f_n) w V^* i (g_1, \dots, g_m) w W^* , odpowiednio.

Na mocy Uwagi 4.6.4 (a), i -ty wyraz j -tej kolumny macierzy $M_{\hat{\sigma}}^{\hat{\tau}}(T)$ jest i -tą współrzędną wektora $\tau(T(v_j))$, która zgodnie z Uwagą 4.7.1 (a) ma postać $a_{ij} = g_i(T(v_j))$.

Po transpozycji, a_{ij} staje się i -tym wyrazem j -tego wiersza macierzy transponowanej, a odpowiedni wyraz b_{ji} macierzy $M_{\hat{\tau}}^{\hat{\sigma}}(T^*)$ jest j -tą współrzędną i -tej kolumny tej macierzy mającą, zgodnie z Uwagą 4.7.1 (c) postać $b_{ji} = T^*(g_i)(v_j)$. Z definicji T^* , $b_{ji} = g_i \circ T(v_j) = a_{ij}$, co kończy dowód. ■

Z twierdzenia wyprowadzimy następujący wniosek.

Wniosek 4.7.8 *Niech $T : V \rightarrow W$ będzie przekształceniem liniowym. Wówczas:*

(a) T jest monomorfizmem $\Leftrightarrow T^*$ jest epimorfizmem;

(b) T jest epimorfizmem $\Leftrightarrow T^*$ jest monomorfizmem.

Dowód. Z Uwagi 4.7.6 (b) $\text{rank } T = \text{rank } T^*$, a stąd i z Twierdzenia 4.2.7 otrzymujemy równoważności:

T jest monomorfizmem $\Leftrightarrow \text{rank } T = \dim V \Leftrightarrow \text{rank } T^* = \dim V^* \Leftrightarrow T^*$ jest epimorfizmem

oraz

T jest epimorfizmem $\Leftrightarrow \text{rank } T = \dim W \Leftrightarrow \text{rank } T^* = \dim W^* \Leftrightarrow T^*$ jest monomorfizmem. ■

5 Wyznaczniki

Wyznacznik macierzy kwadratowych jest funkcją $\det : \mathbb{K}_n^n \rightarrow \mathbb{K}$, ($n = 1, 2, \dots$) przypisującą każdej macierzy kwadratowej skalar, liniowo ze względu na każdy wiersz osobno i zerującą się na macierzach mających dwa identyczne wiersze. Jak zobaczymy, te warunki i warunek $\det I_n = 1$ charakteryzują wyznacznik jednoznacznie, przy czym funkcję \det można określić przez indukcję ze względu na n .

Zanim przystąpimy do dokładnego opisu wyznacznika, podamy pewną interpretację geometryczną funkcji $\det : \mathbb{R}_3^3 \rightarrow \mathbb{R}$. Własności wyznacznika zapewniają, że moduł wyznacznika nie zmienia się przy operacjach elementarnych typu (I) i (II) na wierszach macierzy. Tak więc, jeśli $A \in \mathbb{R}_3^3$ jest macierzą odwracalną, a macierz B jest macierzą diagonalną otrzymaną w wyniku operacji elementarnych na wierszach A (zob. dowód Twierdzenia 4.5.6), to $|\det A| = |\det B|$.

Nietrudno sprawdzić, że operacje elementarne redukujące A do B nie zmieniają objętości równoległościanu rozpiętego na wierszach macierzy w trójwymiarowej przestrzeni euklidesowej. Zatem objętości równoległościanów rozpiętych na wierszach A i B są identyczne. Ponieważ wiersze B rozpinają prostopadłościan i $|\det B|$ jest iloczynem modułów wyrazów na przekątnej B – długości jego krawędzi, $|\det B|$ jest objętością tego prostopadłościanu.

W rezultacie widzimy, że $|\det A|$ jest objętością równoległościanu rozpiętego na wierszach macierzy A . Znak wyznacznika wiąże się z orientacją przestrzeni.

Do tej ważnej interpretacji geometrycznej wyznacznika nad ciałem liczb rzeczywistych powrócimy w dalszej części, po wprowadzeniu n -wymiarowych przestrzeni euklidesowych. Najpierw jednak skupimy się na własnościach algebraicznych wyznaczników nad dowolnym ciałem skalarów.

5.1 Definicja i podstawowe własności.

W tej części podamy dowód twierdzenia o istnieniu i jednoznaczności wyznacznika.

Twierdzenie 5.1.1 *Istnieje dokładnie jedna funkcja $\det : \mathbb{K}_n^n \rightarrow \mathbb{K}$ (zwana wyznacznikiem) taka, że*

$$(1) \det \begin{bmatrix} w_1 \\ \vdots \\ cw_k \\ \vdots \\ w_n \end{bmatrix} = c \det \begin{bmatrix} w_1 \\ \vdots \\ w_k \\ \vdots \\ w_n \end{bmatrix} \text{ dla } c \in \mathbb{K} \text{ (jednorodność względem } k\text{-tego wiersza, } k = 1, \dots, n),$$

$$(2) \det \begin{bmatrix} w_1 \\ \vdots \\ w'_k + w''_k \\ \vdots \\ w_n \end{bmatrix} = \det \begin{bmatrix} w_1 \\ \vdots \\ w'_k \\ \vdots \\ w_n \end{bmatrix} + \det \begin{bmatrix} w_1 \\ \vdots \\ w''_k \\ \vdots \\ w_n \end{bmatrix} \text{ (addytywność względem } k\text{-tego wiersza, } k = 1, \dots, n),$$

$$(3) \det A = 0, \text{ jeśli } A \text{ ma dwa sąsiednie wiersze równe,}$$

$$(4) \det I_n = 1.$$

Definicja 5.1.2 *Wartość $\det A$ funkcji \det na macierzy $A \in \mathbb{K}_n^n$ nazywamy wyznacznikiem A .*

Dowód przeprowadzimy określając najpierw indukcyjnie funkcję spełniającą warunki (1)–(4), a następnie upewniając się, że te warunki określają funkcję \det jednoznacznie.

Zacniemy od uwagi pokazującej, że warunek (3) można wzmocnić, żądając by wyznacznik zerował się na macierzach mających dwa równe wiersze. W definicji wyznacznika często podaje się taką mocniejszą wersję warunku (3), ale użycie w 5.1.1 słabszej wersji (3) upraszcza dowód istnienia funkcji \det .

Uwaga 5.1.3 Niech $\det : \mathbb{K}_n^n \rightarrow \mathbb{K}$ spełnia warunki (1)–(4) i $A \in \mathbb{K}_n^n$.

- (a) Ustalmy $k < l \leq n$. Jeśli $\det C = 0$ dla macierzy C takich, że k -ty wiersz C jest równy l -temu i $B \in \mathbb{K}_n^n$ powstaje z A w wyniku zamiany miejscami wiersza k -tego z l -tym, to $\det B = -\det A$.
- (b) $\det C = 0$ jeśli C ma dwa wiersze równe.

Uzasadnimy (a). Niech w_k i w_l będą k -tym oraz l -tym wierszem macierzy A . Rozpatrzmy macierze C' , C'' i C''' , mające k -ty oraz l -ty wiersz równy odpowiednio $w_k + w_l$, w_k i w_l , a pozostałe wiersze identyczne z wierszami A . W (a) zakładamy, że wyznaczniki tych macierzy się zerują, a z (2) dla wierszy k i l mamy $\det C' = \det C'' + \det A + \det B + \det C'''$, czyli $\det A + \det B = 0$, co dowodzi (a).

Z (3) wynika, że założenie w (a) jest spełnione dla $l = k + 1$. Oznacza to, że przestawienie dwóch sąsiednich wierszy zmienia znak wyznacznika. Jeśli C ma dwa wiersze równe, to kilkakrotnie zamieniając dwa sąsiednie wiersze miejscami możemy przekształcić C w macierz A mającą dwa sąsiednie wiersze równe. Z (3) mamy więc $\det C = \pm \det A = 0$. \square

(A) Istnienie funkcji \det .

Założmy, że istnieje funkcja $\det : \mathbb{K}_{(n-1)}^{(n-1)} \rightarrow \mathbb{K}$ spełniająca (1)–(4) (dla $n = 1$ przyjmujemy $\det[a] = a$).

Dla macierzy $A = [a_{ij}] \in \mathbb{K}_n^n$ oznaczmy przez A_{ij} macierz z $\mathbb{K}_{(n-1)}^{(n-1)}$ otrzymaną z A przez skreślenie i -tego wiersza i j -tej kolumny oraz przyjmijmy

$$d_j(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{ij}.$$

Ustalmy $j \leq n$. Pokażemy, że d_j spełnia warunki (1)–(4) na macierzach z \mathbb{K}_n^n (po upewnieniu się, że (1)–(4) określają wyznacznik jednoznacznie, będziemy także wiedzieć, że $d_j(A)$ nie zależy od j).

Istotnie, jednorodność i addytywność za względu na k -ty wiersz każdego składnika $(-1)^{i+j} a_{ij} \det A_{ij}$ wynikają z założenia indukcyjnego dla $i \neq k$, a jeśli $i = k$, to własności te wynikają z faktu, że zmiana k -tego wiersza nie zmienia $\det A_{kj}$, a jedynie a_{kj} .

Dla sprawdzenia własności (3) założmy, że k -ty i $(k+1)$ -szy wiersz macierzy A są identyczne. Z założenia indukcyjnego zerują się wtedy wszystkie $\det A_{ij}$ dla $i \notin \{k, k+1\}$, więc

$$d_j(A) = (-1)^{k+j} a_{kj} \det A_{kj} + (-1)^{(k+1)+j} a_{k+1j} \det A_{k+1j} = 0,$$

a ponieważ macierze A_{kj} i A_{k+1j} są identyczne, mamy $d_j(A) = 0$.

Własność (4) wynika z równości $d_j(I_n) = (-1)^{j+j} \det(I_n)_{jj} = \det I_{n-1} = 1$.

(B) Jednoznaczność funkcji \det na macierzach elementarnych.

Niech B będzie macierzą otrzymaną z macierzy A w wyniku operacji elementarnej na wierszach. Wyjaśnimy zależność między $\det A$ i $\det B$.

Warunek (1) oznacza, że $\det B = c \det A$ dla operacji mnożącej k -ty wiersz przez $c \neq 0$.

Z Uwagi 5.1.3 wynika, że $\det B = -\det A$ dla operacji zamieniającej dwa wiersze miejscami.

Pokażemy, że $\det B = \det A$ dla operacji dodającej do k -tego wiersza w_k macierzy A i -ty wiersz w_i tej macierzy pomnożony przez skalar a . Istotnie, z (2) i (1) wynika, że $\det B = \det A + a \det C$, gdzie C jest macierzą mającą k -ty i i -ty wiersz równy w_i , więc $\det C = 0$.

Przypomnijmy, że wykonanie operacji elementarnej na wierszach A daje iloczyn MA , gdzie M jest odpowiednią macierzą elementarną, zob. Uwaga 4.5.5. Dla macierzy $A \in \mathbb{K}_n^n$ i macierzy elementarnej $M \in \mathbb{K}_n^n$ mamy więc

$$(*) \quad \det MA = \begin{cases} \det A & \text{dla } M \text{ dodającej do wiersza inny wiersz pomnożony przez skalar,} \\ -\det A & \text{dla } M \text{ zamieniającej dwa wiersze miejscami,} \\ c \det A & \text{dla } M \text{ mnożącej wiersz przez } c \neq 0. \end{cases}$$

Zastępując A przez I_n , z warunku (4) dostajemy

$$\det M = \begin{cases} 1 & \text{dla } M \text{ dodającej do wiersza inny wiersz pomnożony przez skalar,} \\ -1 & \text{dla } M \text{ zamieniającej dwa wiersze miejscami,} \\ c & \text{dla } M \text{ mnożącej wiersz przez } c \neq 0, \end{cases}$$

czyli $\det MA = \det M \det A$ dla dowolnej macierzy A i macierzy elementarnej M .

(C) Jednoznaczność funkcji \det .

Jeśli $M_1, \dots, M_p \in \mathbb{K}_n^n$ są macierzami elementarnymi, to ze wzoru $\det MA = \det M \det A$ wynika (przez indukcję ze względu na p), że $\det(M_p \dots M_1 B) = \det M_p \dots \det M_1 \det B$ dla $B \in \mathbb{K}_n^n$.

Dla odwracalnej macierzy A wartość $\det A$ jest jednoznacznie wyznaczona i $\det A \neq 0$ (bo z Wniosku 4.5.7, A rozkłada się na iloczyn $A = M_p \dots M_1$ macierzy elementarnych i przyjmując $B = I_n$ dostajemy $\det A = \det M_p \dots \det M_1$). Ponadto, $\det AB = \det A \det B$ (bo $\det AB = \det(M_p \dots M_1 B) = \det M_p \dots \det M_1 \det B = \det A \det B$).

Dla macierzy A , która nie jest odwracalna, $\det A = 0$ (bo jeśli M jest iloczynem macierzy elementarnych odpowiadających operacjom redukującym A do postaci schodkowej, to ostatni wiersz MA jest zerowy i z (1) dla $k = n$, $c = 0$ mamy $\det(MA) = 0$, ale $\det(MA) = \det M \det A$ i $\det M \neq 0$, bo M jest odwracalna).

Wykazaliśmy więc jednoznaczność i zakończyliśmy dowód Twierdzenia 5.1.1. ■

Uwaga 5.1.4 (a) $\det A \neq 0$ wtedy i tylko wtedy, gdy A jest odwracalna (pokazaliśmy to w dowodzie jednoznaczności).

(b) **(Twierdzenie Cauchy’ego)** $\det AB = \det A \det B$ dla $A, B \in \mathbb{K}_n^n$ (pokazaliśmy to dla odwracalnej macierzy A ; w przeciwnym przypadku $\det AB = 0 = \det A \det B$).

(c) Jeśli macierz A jest odwracalna, to $\det(A^{-1}) = (\det A)^{-1}$ (bo $1 = \det I_n = \det(A \cdot A^{-1}) = \det A \cdot \det(A^{-1})$). □

5.2 Obliczanie wyznaczników.

Z jednoznaczności w Twierdzeniu 5.1.1 wynika, że $d_j(A) = \det A$ dla funkcji d_j zdefiniowanych w dowodzie istnienia, stąd otrzymujemy natychmiast następujące twierdzenie.

Twierdzenie 5.2.1 (o rozwinięciu Laplace’a względem j -tej kolumny). Dla $A = [a_{ij}] \in \mathbb{K}_n^n$,

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{ij}.$$

Przykład 5.2.2 Rozwijając względem pierwszej kolumny, dostajemy dla $n = 2$ wzór

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

a dla $n = 3$ wzór *Sarrusa*

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

Uwaga 5.2.3 Wyznacznik $(n \times n)$ -macierzy dla $n > 3$ można obliczyć zmniejszając wymiar macierzy przy pomocy rozwinięcia Laplace’a, lub redukując tę macierz do postaci schodkowej. Wzory (*) w części 5.1 pozwalają powiązać wyznacznik macierzy wyjściowej z wyznacznikiem macierzy zredukowanej, który łatwo obliczyć korzystając z punktu (a) poniżej.

(a) Wyznacznik macierzy *górnje trójkątnej* (czyli macierzy kwadratowej mającej zera pod przekątną) jest iloczynem wyrazów na przekątnej – dla macierzy $A = [a_{ij}]$ takiej, że $a_{ij} = 0$ dla $i > j$, $\det A = a_{11} \cdot \dots \cdot a_{nn}$. Istotnie, rozwijając $\det A$ względem pierwszej kolumny dostajemy wzór dający krok indukcyjny dowodu.

(b) Dla macierzy blokowo trójkątnej $A = \left[\begin{array}{c|c} A_1 & * \\ \hline \mathbf{0} & A_2 \end{array} \right]$, gdzie $A_i \in \mathbb{K}_n^{n_i}$, mamy $\det A = \det A_1 \det A_2$ (bo zgodnie z Uwagą 1.2.5 macierze $[A_1 \mid *] \in \mathbb{K}_n^{n_1}$ i $[\mathbf{0} \mid A_2] \in \mathbb{K}_n^{n_2}$ można doprowadzić do postaci schodkowej operacjami elementarnymi typu (I), które nie zmieniają wyznaczników). \square

Pokażemy teraz, że zmiana wierszy w kolumny macierzy kwadratowej nie zmienia wyznacznika.

Twierdzenie 5.2.4 Dla $A \in \mathbb{K}_n^n$, $\det A = \det A^T$.

Dowód. Jeśli $\text{rank } A < n$, to $\text{rank } A^T < n$ i oba wyznaczniki są zerami.

Jeśli $\text{rank } A = n$, to z Wniosku 4.5.7, A rozkłada się na iloczyn $A = M_p \dots M_1$ macierzy elementarnych. Zgodnie z Uwagą 4.7.6 (c), $A^T = M_1^T \dots M_p^T$, więc korzystając z Uwagi 5.1.4 (b), wystarczy zauważyć, że $\det M = \det M^T$ dla macierzy elementarnych M . \blacksquare

Rozwinięcie Laplace’a wyznacznika $\det A^T$ względem i -tej kolumny macierzy A^T daje wzór na rozwinięcie $\det A$ względem i -tego wiersza macierzy A .

Twierdzenie 5.2.5 (o rozwinięciu Laplace’a względem i -tego wiersza). Dla $A = [a_{ij}] \in \mathbb{K}_n^n$,

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A_{ij}.$$

Wyznacznik pozwala też określić znak permutacji, co prowadzi do formuły uogólniającej wzór Sarrusa.

Niech S_n będzie zbiorem wszystkich bijekcji zbioru $\{1, 2, \dots, n\}$ na siebie – *permutacji*. Każdej permutacji $\pi \in S_n$ odpowiada macierz $E_\pi = [E_{\pi(1)}, E_{\pi(2)}, \dots, E_{\pi(n)}]$, której wyznacznik $\det E_\pi \in \{1, -1\}$ nazywamy *znakiem permutacji* π i oznaczamy symbolem $\text{sgn}(\pi)$ (łatwo upewnić się, że znak π określa parzystość liczby transpozycji przeprowadzających π na identyczność – dla $\text{sgn}(\pi) = 1$ ta liczba jest parzysta, a dla $\text{sgn}(\pi) = -1$, nieparzysta).

Twierdzenie 5.2.6 Dla macierzy $A = [a_{ij}] \in \mathbb{K}_n^n$,

$$\det A = \sum_{\pi \in S_n} \text{sgn}(\pi) a_{\pi(1)1} a_{\pi(2)2} \dots a_{\pi(n)n}.$$

Dowód. Niech $A = [A_1, \dots, A_n] = [a_{ij}] \in \mathbb{K}_n^n$. Wtedy $A_j = \sum_{i=1}^n a_{ij} E_i$ i z liniowości wyznacznika względem kolejnych kolumn mamy

$$\begin{aligned} \det A &= \det \left[\sum_{i=1}^n a_{i1} E_i, \sum_{i=1}^n a_{i2} E_i, \dots, \sum_{i=1}^n a_{in} E_i \right] = \\ &= \sum_{i_1=1}^n a_{i_1 1} \det [E_{i_1}, \sum_{i=1}^n a_{i2} E_i, \dots, \sum_{i=1}^n a_{in} E_i] = \\ &= \sum_{i_1=1}^n \sum_{i_2=1}^n a_{i_1 1} a_{i_2 2} \det [E_{i_1}, E_{i_2}, \dots, \sum_{i=1}^n a_{in} E_i] = \\ &= \dots = \\ &= \sum_{i_1=1}^n \sum_{i_2=1}^n \dots \sum_{i_n=1}^n a_{i_1 1} a_{i_2 2} \dots a_{i_n n} \det [E_{i_1}, E_{i_2}, \dots, E_{i_n}]. \end{aligned}$$

Teza wynika teraz z faktu, że występujące we wzorze wyznaczniki $\det [E_{i_1}, E_{i_2}, \dots, E_{i_n}]$ są zerowe jeśli $i_j = i_k$ dla pewnych $j \neq k$, więc sumowanie można ograniczyć do ciągów różnowartościowych (i_1, \dots, i_n) , czyli do permutacji zbioru $\{1, \dots, n\}$. \blacksquare

5.3 Macierz stowarzyszona i wzory Cramera.

Ważną rolę (choć nie przy obliczeniach) odgrywa macierz $\text{adj}A$ stowarzyszona z macierzą kwadratową A , zdefiniowana przy pomocy wyznaczników, która po pomnożeniu przez A daje macierz $\det A \cdot I$. Przy pomocy macierzy stowarzyszonej otrzymuje się wzory Cramera opisujące w terminach wyznaczników rozwiązania układów równań $AX = B$ z macierzą odwracalną A (*układy Cramera*).

Ustalmy macierz $A = [A_1, \dots, A_n] = [a_{ij}] \in \mathbb{K}_n^n$ i przypomnijmy, że w części 5.1 zdefiniowaliśmy A_{ij} jako macierz otrzymaną z A przez skreślenie i -tego wiersza i j -tej kolumny.

Definicja 5.3.1 *Macierzą stowarzyszoną z A nazywamy macierz $\text{adj}A = [\hat{a}_{ij}]^T$, gdzie $\hat{a}_{ij} = (-1)^{i+j} \det A_{ij}$.*

Dla $B \in \mathbb{K}^n$ mamy

$$(**) \quad \text{adj}A \cdot B = \begin{bmatrix} \det[B, A_2, \dots, A_n] \\ \det[A_1, B, \dots, A_n] \\ \vdots \\ \det[A_1, A_2, \dots, B] \end{bmatrix}.$$

Wzór $(**)$ wynika z Twierdzenia 5.2.1, bo dla $B = [b_1, b_2, \dots, b_n]^T$

$$\begin{bmatrix} (-1)^{1+1} \det A_{11} & \dots & (-1)^{n+1} \det A_{n1} \\ (-1)^{1+2} \det A_{12} & \dots & (-1)^{n+2} \det A_{n2} \\ \vdots & & \vdots \\ (-1)^{1+n} \det A_{1n} & \dots & (-1)^{n+n} \det A_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n (-1)^{i+1} b_i \det A_{i1} \\ \sum_{i=1}^n (-1)^{i+2} b_i \det A_{i2} \\ \vdots \\ \sum_{i=1}^n (-1)^{i+n} b_i \det A_{in} \end{bmatrix} = \begin{bmatrix} \det[B, A_2, \dots, A_n] \\ \det[A_1, B, \dots, A_n] \\ \vdots \\ \det[A_1, A_2, \dots, B] \end{bmatrix}.$$

W szczególności $\text{adj}A \cdot A = [\text{adj}A A_1, \dots, \text{adj}A A_n] = (\det A)I_n$, więc dostajemy

Twierdzenie 5.3.2 *Jeśli A jest macierzą odwracalną, to $A^{-1} = \frac{1}{\det A} \text{adj}A$.*

Rozwiązanie $X = [x_1, \dots, x_n]^T$ układu Cramera $AX = B$ ma postać $X = A^{-1}B = \frac{1}{\det A}(\text{adj}A B)$. Zatem z $(**)$ dostajemy *wzory Cramera*:

$$x_1 = \frac{\det[B, A_2, \dots, A_n]}{\det A}, \quad x_2 = \frac{\det[A_1, B, \dots, A_n]}{\det A}, \quad \dots, \quad x_n = \frac{\det[A_1, A_2, \dots, B]}{\det A}.$$

6 Endomorfizmy przestrzeni liniowych

Przekształcenia liniowe przestrzeni liniowych w siebie nazywamy *endomorfizmami*.

Zauważmy, że wybór układu współrzędnych $\sigma : V \rightarrow \mathbb{K}^n$ w n -wymiarowej przestrzeni liniowej V nad \mathbb{K} prowadzi do izomorfizmu $T \rightarrow M_\sigma^\sigma(T)$ przestrzeni endomorfizmów $L(V, V)$ i przestrzeni macierzy \mathbb{K}_n^n , przy czym, przy tym izomorfizmie składaniu endomorfizmów odpowiada mnożenie macierzy. Analizując interesujące nas własności, czasem wygodniej jest rozpatrywać endomorfizmy, a czasem macierze.

Jednym z najważniejszych wyników tego rozdziału jest twierdzenie Jordana, podające klarowny opis struktury endomorfizmów przestrzeni nad ciałem liczb zespolonych. Dowód tego twierdzenia odbiega stopniem trudności od pozostałych rozumowań w skrypcie i choć podajemy go ze szczegółami (wybraliśmy rozumowanie, które prowadzi szybko do celu, ustalając po drodze fakty potrzebne przy zastosowaniach), należy przypomnieć, że nie wchodzi on w zakres materiału obowiązującego w obecnym programie GAL.

6.1 Wielomian charakterystyczny, wektory własne.

Pierwiastki wielomianu charakterystycznego macierzy, który definiujemy poniżej, pozwalają wyznaczać niezerowe wektory przechodzące pod działaniem tej macierzy na wektory proporcjonalne – wektory własne macierzy.

Definicja 6.1.1 *Wielomianem charakterystycznym macierzy kwadratowej $A \in \mathbb{K}_n^n$ nazywamy wielomian $w_A(x) \in \mathbb{K}[x]$ określony wzorem*

$$w_A(x) = \det(A - xI_n).$$

Definicja wielomianu charakterystycznego wymaga komentarza. Występujący w niej wzór najprościej można zinterpretować przy pomocy formuły z Twierdzenia 5.2.6 opisującej $\det(A - xI_n)$ jako sumę iloczynów, których pewne czynniki mają postać $a_{ii} - x$, gdzie a_{ii} są wyrazami z przekątnej A .

Po wymnożeniu i pogrupowaniu wynika stąd wzór

$$w_A(x) = a_0 + a_1(-x) \dots + a_{n-1}(-x)^{n-1} + (-x)^n$$

pokazujący, że $w_A(x)$ jest wielomianem. Łatwo zauważyć, że wyraz stały a_0 jest wyznacznikiem A , zaś współczynnik przy $(-x)^{n-1}$ jest sumą $a_{n-1} = \sum_{i=1}^n a_{ii}$ wyrazów stojących na głównej przekątnej A , którą nazywamy *śladem macierzy A* i oznaczamy $\text{tr } A$.

Jeśli \mathbb{K} jest ciałem nieskończonym (przypomnijmy, że w tym wykładzie najważniejsze są ciała \mathbb{R} i \mathbb{C}), $w_A(x)$ można utożsamiać z funkcją przypisującą skalarowi $\lambda \in \mathbb{K}$ skalar $\det(A - \lambda I_n)$ i w dalszym ciągu traktować będziemy $w_A(x)$ właśnie w taki sposób – jako funkcję $w_A : \mathbb{K} \rightarrow \mathbb{K}$.

Należy jednak zwrócić uwagę, że nad \mathbb{Z}_2 , wielomian $\det\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - x \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = -x + x^2 \in \mathbb{Z}_2[x]$ określa funkcję tożsamościowo równą zero, a więc nie jest obojętne jak interpretujemy wzór w 6.1.1.

Definicja 6.1.2 *Dla $A \in \mathbb{K}_n^n$ pierwiastki wielomianu charakterystycznego $w_A(x)$ nazywamy wartościami własnymi macierzy A , a ich zbiór $\text{Sp}(A) = \{\lambda \in \mathbb{K} : w_A(\lambda) = 0\}$ – spektrum A .*

Wartości własne macierzy $A \in \mathbb{K}_n^n$ są opisane równaniem $\det(A - xI_n) = 0$, a więc, zob. 3.5.3,

$$\lambda \in \text{Sp}(A) \Leftrightarrow N(A - \lambda I_n) \neq \{\mathbf{0}\}.$$

Definicja 6.1.3 *Niezerowe wektory z przestrzeni $N(A - \lambda I_n)$, tzn. niezerowe wektory $X \in \mathbb{K}^n$ takie, że $AX = \lambda X$, nazywamy wektorami własnymi A (odpowiadającymi wartości własnej λ).*

Określimy teraz ważną relację równoważności między macierzami kwadratowymi tego samego wymiaru – relację podobieństwa i sprawdzimy, że relacja podobieństwa zachowuje wielomian charakterystyczny.

Definicja 6.1.4 Macierze $A, B \in \mathbb{K}_n^n$ są podobne, jeśli istnieje macierz odwracalna $C \in \mathbb{K}_n^n$ taka, że $B = C^{-1}AC$.

Twierdzenie 6.1.5 Macierze podobne mają identyczne wielomiany charakterystyczne.

Dowód. Niech $B = C^{-1}AC$, gdzie $A, B, C \in \mathbb{K}_n^n$. Należy pokazać, że funkcje w_A i w_B na \mathbb{K} są identyczne. Istotnie, dla $\lambda \in \mathbb{K}$ mamy $w_B(\lambda) = \det(B - \lambda I_n) = \det(C^{-1}AC - \lambda I_n) = \det(C^{-1}(A - \lambda I_n)C) = \det C^{-1} \det(A - \lambda I_n) \det C = w_A(\lambda)$. ■

Jak jednak wyjaśnimy w dalszym ciągu, macierze z identycznym wielomianem charakterystycznym nie muszą być podobne.

Twierdzenie 6.1.5, w połączeniu z następną obserwacją, pozwala zdefiniować wielomian charakterystyczny endomorfizmu.

Uwaga 6.1.6 Jeśli $T : V \rightarrow V$ jest endomorfizmem przestrzeni liniowej V , a $\sigma, \tau : V \rightarrow \mathbb{K}^n$ są układami współrzędnych w V , to macierze $M_\tau^\tau(T)$ i $M_\sigma^\sigma(T)$ są podobne.

Istotnie, dla macierzy $C = M(\tau \circ \sigma^{-1})$ zmieniającej współrzędne $\sigma(v)$ wektora v na jego współrzędne $\tau(v)$, zob. Uwaga 4.6.2 (c), mamy $M_\sigma^\sigma(T) = M(\sigma \circ T \circ \sigma^{-1}) = M(\sigma \circ \tau^{-1} \circ \tau \circ T \circ \tau^{-1} \circ \tau \circ \sigma^{-1}) = M(\sigma \circ \tau^{-1})M(\tau \circ T \circ \tau^{-1})M(\tau \circ \sigma^{-1}) = C^{-1}M_\tau^\tau(T)C$. □

Definicja 6.1.7 Wielomianem charakterystycznym endomorfizmu $T : V \rightarrow V$ nazywamy wielomian $w_T(x) = \det(M_\sigma^\sigma(T) - xI_n)$, gdzie $\sigma : V \rightarrow \mathbb{K}^n$ jest dowolnym układem współrzędnych. Wyznacznikiem i śladem endomorfizmu T nazywamy wyraz wolny $\det T = \det M_\sigma^\sigma(T)$ i współczynnik $\text{tr} T = \text{tr} M_\sigma^\sigma(T)$ tego wielomianu, odpowiednio, a zbiór $\text{Sp}(T)$ pierwiastków $w_T(x)$ nazywamy spektrum endomorfizmu T .

Ponieważ $\ker(T - \text{id}) = \sigma^{-1}(N(M_\sigma^\sigma(T) - \lambda I_n))$, mamy

$$\lambda \in \text{Sp}(T) \Leftrightarrow \ker(T - \lambda \text{id}) \neq \{0\}.$$

Terminologię wprowadzoną w Definicji 6.1.3 przenosimy także na przypadek endomorfizmów.

Definicja 6.1.8 Dla endomorfizmu $T : V \rightarrow V$ niezerowe wektory z $\ker(T - \lambda \text{id})$, tzn. niezerowe wektory $v \in V$ takie, że $T(v) = \lambda v$, nazywamy wektorami własnymi T (odpowiadającymi wartości własnej λ).

Występujące w kolejnej uwadze operacje sumy algebraicznej i sumy prostej skończenie wielu podprzestrzeni przestrzeni liniowej V są naturalnymi uogólnieniami odpowiednich operacji dla dwóch składników wprowadzonych w Definicjach 3.6.2 i 3.6.3.

Uwaga 6.1.9 Dla endomorfizmu $T : V \rightarrow V$:

- (a) suma wektorów własnych odpowiadających parami różnym wartościom własnym T nie jest wektorem zerowym;
- (b) wektory własne odpowiadające parami różnym wartościom własnym T są liniowo niezależne;
- (c) suma algebraiczna $\sum_{\lambda \in \text{Sp}(T)} \ker(T - \lambda \text{id})$ jest sumą prostą $\bigoplus_{\lambda \in \text{Sp}(T)} \ker(T - \lambda \text{id})$.

Zdanie (c) orzeka, że dla dowolnego wyboru $v_\lambda \in \ker(T - \lambda \text{id})$, $\lambda \in \text{Sp}(T)$, z równości $\sum_{\lambda \in \text{Sp}(T)} v_\lambda = 0$ wynika, że wszystkie v_λ są zerowe, a to jest konsekwencją (a). Zdanie (b) również wynika z (a), bo jeśli v_λ jest wektorem własnym odpowiadającym wartości własnej λ i $a_\lambda \neq 0$, to $a_\lambda v_\lambda$ też jest wektorem własnym odpowiadającym λ .

Własność (a) udowodnimy przez indukcję ze względu na liczbę składników k . Załóżmy (a) dla sum mniej niż k wektorów własnych T i niech v_j będzie wektorem własnym odpowiadającym λ_j , $j = 1, \dots, k$, gdzie $\lambda_1, \dots, \lambda_k$ są parami różne. Jeśli $\sum_{j \leq k} v_j = 0$, to $0 = T(\sum_{j \leq k} v_j) - \lambda_k(\sum_{j \leq k} v_j) = \sum_{j < k} (\lambda_j - \lambda_k)v_j$, co przeczy założeniu indukcyjnemu, bo $\lambda_j - \lambda_k \neq 0$ dla $j < k$. □

Endomorfizmy, których wektory własne rozpinają całą przestrzeń mają wyjątkowo prostą strukturę. Zgodnie z Uwagą 6.1.9 klasę takich endomorfizmów można opisać następująco:

Definicja 6.1.10 *Endomorfizm $T : V \rightarrow V$ jest diagonalizowalny, jeśli $V = \bigoplus_{\lambda \in \text{Sp}(T)} \ker(T - \lambda \text{id})$. Macierz $A \in \mathbb{K}_n^n$ jest diagonalizowalna, jeśli A jest podobna do macierzy diagonalnej.*

Uwaga 6.1.11 Diagonalizowalność endomorfizmu $T : V \rightarrow V$ oznacza istnienie bazy (v_1, \dots, v_n) przestrzeni V złożonej z wektorów własnych T (por. Twierdzenie 3.6.6). W układzie współrzędnych $\sigma : V \rightarrow \mathbb{K}^n$ związanym z tą bazą, macierz $D = M_\sigma^\sigma(T)$ jest diagonalna – jeśli $T(v_j) = \lambda_j v_j$, to $\lambda_1, \dots, \lambda_n$ stoją na przekątnej D , zob. Uwaga 4.6.4 (a). W szczególności, wielomian charakterystyczny $w_T(x) = w_D(x) = (\lambda_1 - x) \cdot \dots \cdot (\lambda_n - x)$ ma rozkład na czynniki liniowe, a krotność każdego pierwiastka tego wielomianu jest równa $\dim[\ker(T - \lambda \text{id})]$ (te dwa warunki charakteryzują diagonalizowalność T , zob. Wniosek 3.6.7).

Z Uwagi 6.1.6 wynika, że endomorfizm T jest diagonalizowalny wtedy i tylko wtedy, gdy dla dowolnego układu współrzędnych $\tau : V \rightarrow \mathbb{K}^n$ macierz $M_\tau^\tau(T)$ jest diagonalizowalna. \square

Najważniejszą klasę endomorfizmów diagonalizowalnych poznamy w części 8.1 – są to endomorfizmy samosprężone na przestrzeniach euklidesowych.

6.2 Zasadnicze twierdzenie algebry.

Warunkiem koniecznym istnienia wektorów własnych jest istnienie pierwiastków wielomianu charakterystycznego. To wskazuje szczególną rolę ciała skalarów \mathbb{C} , bo jak wykażemy poniżej, każdy wielomian o współczynnikach zespolonych ma pierwiastek zespolony.

Twierdzenie 6.2.1 *Dla każdego wielomianu $w \in \mathbb{C}[x]$ stopnia dodatniego istnieje $\lambda \in \mathbb{C}$ takie, że $w(\lambda) = 0$.*

Dowód poprzedzimy dwoma lematami

Lemat 6.2.2 (d’Alemberta). *Niech $w(z) = a_0 + a_1 z + \dots + z^n$, $a_i \in \mathbb{C}$. Jeśli $w(a) \neq 0$, to istnieje $b \in \mathbb{C}$ takie, że $|w(b)| < |w(a)|$.*

Dowód. Mamy $w(a+z) = a_0 + a_1(a+z) + \dots + (a+z)^n = A_0 + A_1 z + \dots + z^n = w(a) + A_k z^k + \dots + z^n$, gdzie $A_k \neq 0$.

Niech $c \in \mathbb{C}$ będzie takie, że $c^k = -\frac{w(a)}{A_k}$ ($c = \sqrt[k]{r}(\cos \frac{\theta}{k} + i \sin \frac{\theta}{k})$, jeśli $-\frac{w(a)}{A_k} = r(\cos \theta + i \sin \theta)$, $r > 0$).

Dla $t \in \mathbb{R}$ z przedziału $[0, 1]$,

$$w(a+tc) = w(a) + A_k t^k c^k + \dots + t^n c^n = w(a) - w(a)t^k + B_{k+1} t^{k+1} \dots + B_n t^n = w(a)(1-t^k) + B_{k+1} t^{k+1} \dots + B_n t^n.$$

Tak więc

$$|w(a+tc)| \leq (1-t^k)|w(a)| + t^k(|B_{k+1}|t \dots + |B_n|t^{n-k}) = |w(a)| + t^k(-|w(a)| + |B_{k+1}|t \dots + |B_n|t^{n-k}).$$

Dobierzmy $t > 0$ tak małe, żeby wyrażenie w nawiasie było ujemne. Wówczas dla $b = a + tc$ mamy $|w(b)| < |w(a)|$. \blacksquare

Lemat 6.2.3 *Dla wielomianu zespolonego $w(z) = a_0 + a_1 z + \dots + z^n$ i $K = \{z \in \mathbb{C} : |\text{Re}z|, |\text{Im}z| \leq M\}$ istnieje $z_0 \in K$ takie, że $|w(z_0)| = \inf \{|w(z)| : z \in K\}$.*

Lemat jest wersją twierdzenia Weierstrassa dla funkcji $|w(z)|$ na kwadracie K , jego uzasadnienie podamy w uzupełnieniach, zob. Lemat 14.1.1. Z tych dwóch lematów łatwo wyprowadzimy teraz zasadnicze twierdzenie algebry.

Dowód Twierdzenia 6.2.1. Niech $w(z) = a_0 + a_1z + \dots + z^n$ będzie wielomianem zespolonym. Ponieważ

$$|w(z)| = \left| z^n \left(\frac{a_0}{z^n} + \dots + \frac{a_{n-1}}{z} + 1 \right) \right| \geq |z^n| \left(1 - \left(\frac{|a_0|}{|z^n|} + \dots + \frac{|a_{n-1}|}{|z|} \right) \right),$$

dla pewnego kwadratu K o środku w zerze i dostatecznie dużym boku, $\mu = \inf \{|w(z)| : z \in \mathbb{C}\} = \inf \{|w(z)| : z \in K\}$. Z Lematu 6.2.3 istnieje $z_0 \in K$ takie, że $|w(z_0)| = \mu$, a z Lematu d'Alemberta $\mu = 0$. ■

6.3 Twierdzenie Jordana.

Głównym celem tej części jest możliwie przejrzysty opis endomorfizmów przestrzeni liniowych nad ciałem liczb zespolonych. W języku macierzowym, pokażemy, że każda macierz z \mathbb{C}_n^n jest podobna do macierzy o szczególnie prostej postaci – macierzy zbudowanej z klatek Jordana.

Nasze rozważania podzielimy na dwie części. W pierwszej, nie nakładając na ciało żadnych ograniczeń, wyjaśnimy strukturę endomorfizmów, których pewna potęga jest zerem – *endomorfizmów nilpotentnych*. W drugiej ograniczymy się do przestrzeni nad ciałem liczb zespolonych i wykorzystamy zasadnicze twierdzenie algebry 6.2.1 – w połączeniu z wcześniejszą analizą endomorfizmów nilpotentnych, to łatwo doprowadzi do celu.

(A) Struktura endomorfizmów nilpotentnych.

Niech $T : V \rightarrow V$ będzie endomorfizmem przestrzeni liniowej nad dowolnym ciałem K .

Układ wektorów $(v, T(v), \dots, T^{k-1}(v))$ w V taki, że $T^{k-1}(v) \neq \mathbf{0}$ i $T^k(v) = \mathbf{0}$ będziemy nazywali *T-serią* o początku v i wysokości k . Przestrzeń $[v] = \text{lin}(v, T(v), \dots, T^{k-1}(v))$ jest *niezmiennicza dla T* (to znaczy $T[v] \subset [v]$) i seria $(v, T(v), \dots, T^{k-1}(v))$ jest bazą $[v]$: jeśli $\sum_{j=0}^{k-1} a_j T^j(v) = \mathbf{0}$, gdzie $T^0 = \text{id}$, to kolejno działając na obie strony tej równości endomorfizmami $T^{k-1}, T^{k-2}, \dots, T^0$, wnosimy, że $a_0 = 0, a_1 = 0, \dots, a_{k-1} = 0$.

Ponieważ $T(T^j(v)) = T^{j+1}(v)$, w układzie współrzędnych $\sigma : [v] \rightarrow \mathbb{K}^k$ związanym z *T-serią* o początku v macierz obciążenia T do $[v]$ ma bardzo prostą postać

$$M_\sigma^\sigma(T) = \begin{bmatrix} 0 & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & 0 \end{bmatrix} \in \mathbb{K}_k^k;$$

jest to klatka Jordana $J_k(0)$.

Twierdzenie 6.3.1 Niech $T : V \rightarrow V$ będzie endomorfizmem takim, że $T^m = \mathbf{0}$. Istnieją wówczas skończone (niektóre być może puste) zbiory wektorów $F_k \subset V$, $k = 1, \dots, m$ takie, że każde $v \in F_k$ jest początkiem *T-serii* o wysokości k oraz

$$(*) \quad V = \bigoplus_{k \leq m} \left(\bigoplus_{v \in F_k} [v] \right).$$

Przy tym, dla $k \leq m$ mamy

$$(**) \quad |F_k| = \text{rank } T^{k-1} - 2 \text{rank } T^k + \text{rank } T^{k+1}.$$

Dowód.* Połóżmy $W_k = \ker T^k$ dla $k \geq 0$ i rozważmy łańcuch $\{\mathbf{0}\} = W_0 \subset W_1 \subset \dots \subset W_{m-1} \subset W_m = V$ podprzestrzeni V . Endomorfizm T przeprowadza W_{k+1} w W_k , więc korzystając z Wniosku 3.6.8, znajdziemy (być może zerową) podprzestrzeń $U_k \subset W_k$ taką, że

$$(\dagger) \quad W_k = (W_{k-1} + T(W_{k+1})) \oplus U_k \text{ dla } k = 1, \dots, m.$$

Dla niezerowych przestrzeni U_k definiujemy F_k jako zbiór wektorów ustalonej bazy U_k i kładziemy $F_k = \emptyset$ jeśli U_k jest zerowa.

Przyjmując $[U_k] = U_k + T(U_k) + \dots + T^{k-1}(U_k)$, zauważmy, że dla $k = m, m-1, \dots, 1$

$$V = W_{k-1} + [U_k] + [U_{k+1}] + \dots + [U_m].$$

Istotnie, dla $k = m$ równość wynika z warunku (\dagger) , bo $W_{m+1} = W_m = V$ i $T(W_m) \subset W_{m-1}$. Załóżmy, że $V = W_k + [U_{k+1}] + \dots + [U_m]$. Wtedy $T(V) \subset T(W_k) + [U_{k+1}] + \dots + [U_m] \subset W_{k-1} + [U_{k+1}] + \dots + [U_m]$, więc, odwołując się znowu do (\dagger) , mamy $V = (W_{k-1} + T(W_{k+1}) + U_k) + [U_{k+1}] + \dots + [U_m] \subset W_{k-1} + T(V) + U_k + [U_{k+1}] + \dots + [U_m] \subset W_{k-1} + [U_k] + [U_{k+1}] + \dots + [U_m]$.

W szczególności, V jest sumą algebraiczną podprzestrzeni $[U_k]$, $k \leq m$. Ponieważ $[U_k]$ jest sumą algebraiczną podprzestrzeni $[v]$, $v \in F_k$, wynika stąd, że V jest sumą algebraiczną podprzestrzeni $[v]$, $v \in \bigcup_{k \leq m} F_k$. To, że jak orzeka (*), jest to suma prosta, wywnioskujemy z własności (**), do uzasadnienia której teraz przejdziemy.

Ustalmy $k \leq m$. Ponieważ obcięcie $T^{k-1}|U_k$ jest izomorfizmem, zob. (\dagger) , $|F_k| = \dim U_k = \text{rank}(T^{k-1}|U_k)$. Warunek (\dagger) i $W_{k-1} = \ker T^{k-1}$ implikują $T^{k-1}(W_k) = T^{k-1}(T(W_{k+1})) + T^{k-1}(U_k)$. Ponadto $T^{k-1}(W_k) = T^{k-1}(T(W_{k+1})) \oplus T^{k-1}(U_k)$, bo jeśli $w \in T(W_{k+1})$, $u \in U_k$ i $T^{k-1}(w) = T^{k-1}(u)$, to $u - w \in W_{k-1}$ i zgodnie z (\dagger) , z $(u - w) + w = u$ wynika $u = \mathbf{0}$. Zatem $|F_k| = \text{rank}(T^{k-1}|W_k) - \text{rank}(T^{k-1}|U_k)$. Z Twierdzenia 4.2.9 dostajemy $\text{rank}(T^{k-1}|W_k) = \dim W_k - \text{def}(T^{k-1}|W_k) = \text{def} T^k - \text{def} T^{k-1} = \text{rank} T^{k-1} - \text{rank} T^k$, a stąd $|F_k| = (\text{rank} T^{k-1} - \text{rank} T^k) - (\text{rank} T^k - \text{rank} T^{k+1})$, czyli (**).

Ponieważ, jak zauważyliśmy V jest sumą algebraiczną przestrzeni $[v]$, $v \in \bigcup_{k \leq m} F_k$, dla uzasadnienia (*) pozostaje wyprowadzić z (**), że suma wymiarów takich $[v]$ jest równa $\dim V$, zob. Wniosek 3.6.6.

Dla uproszczenia oznaczeń, przyjmujemy $r_j = \text{rank} T^j$ i zauważmy, że $r_0 = 0$, $r_m = r_{m+1} = 0$. Każdy wektor $v \in F_k$ jest początkiem T -serii wysokości k , zob. (\dagger) , więc suma wymiarów przestrzeni $[v]$, $v \in \bigcup_{k \leq m} F_k$ jest równa

$$\sum_{k=1}^m k|F_k| = \sum_{k=1}^m k[(r_{k-1} - r_k) - (r_k - r_{k+1})] = [\sum_{k=1}^m (r_{k-1} - r_k)] - m(r_m - r_{m+1}) = r_0,$$

co kończy dowód twierdzenia. ■

Twierdzenie 6.3.1 uzupełnimy obserwacją, z której skorzystamy w dowodzie twierdzenia Jordana.

Lemat 6.3.2 *Dla endomorfizmu $T : V \rightarrow V$ i wartości własnej $\lambda \in \text{Sp}(T)$ istnieje $m \geq 1$ takie, że*

$$V = \ker [(T - \lambda \text{id})^m] \oplus \text{im} [(T - \lambda \text{id})^m],$$

przy czym, jeśli $\mu \neq \lambda$, $k \geq 1$, to $\ker [(T - \mu \text{id})^k] \subset \text{im} [(T - \lambda \text{id})^m]$.

Dowód. Połóżmy $S = T - \lambda \text{id}$. Ponieważ $\dim V$ jest skończony, łańcuch $\{\mathbf{0}\} \neq \ker S \subset \ker S^2 \subset \dots$ musi się stabilizować, tzn. dla pewnego $m \geq 1$ mamy $\ker S^m = \ker S^{m+1} = \dots$.

Jeśli $w = S^m(v)$ oraz $S^m(w) = \mathbf{0}$, to $S^{2m}(v) = \mathbf{0}$, stąd $S^m(v) = \mathbf{0}$ i $w = \mathbf{0}$.

To pokazuje, że część wspólna $\ker S^m \cap \text{im} S^m$ jest przestrzenią zerową, a ponieważ suma wymiarów tych przestrzeni jest równa $\text{def} S^m + \text{rank} S^m = \dim V$, otrzymujemy pierwszą część tezy.

Dla dowodu drugiej części połóżmy dodatkowo $W = \ker [(T - \mu \text{id})^k]$. Dla $v \in W$, z przemienności endomorfizmów $S = (T - \lambda \text{id})$ i $(T - \mu \text{id})$ wynika, że $S(v) \in W$. Ponadto warunek $v \in W \cap \ker S$ implikuje $v = \mathbf{0}$. Istotnie, $v \in \ker S$ oznacza $T(v) = \lambda v$, więc $(T - \mu \text{id})(v) = (\lambda - \mu)v$, a stąd $(T - \mu \text{id})^k(v) = (\lambda - \mu)^k v$ i z $v \in W$ wynika wtedy $v = \mathbf{0}$. Tak więc obcięcie $S|_W : W \rightarrow W$ jest izomorfizmem, a zatem $W = S^m(W) \subset \text{im} [(T - \lambda \text{id})^m]$. ■

(B) Twierdzenie Jordana i postać Jordana macierzy.

Z zasadniczego twierdzenia algebry i z Lematu 6.3.2 wyprowadzimy teraz

Twierdzenie 6.3.3 *Jeśli $T : V \rightarrow V$ jest endomorfizmem przestrzeni liniowej V nad \mathbb{C} , to dla każdej wartości własnej $\lambda \in \text{Sp}(T)$ istnieje $m_\lambda \geq 1$ takie, że*

$$V = \bigoplus_{\lambda \in \text{Sp}(T)} \ker [(T - \lambda \text{id})^{m_\lambda}].$$

Dowód. Dowód przeprowadzimy indukcyjnie ze względu na wymiar przestrzeni. Załóżmy, że twierdzenie jest prawdziwe dla przestrzeni wymiaru mniejszego niż $\dim V$. Z zasadniczego twierdzenia algebry istnieje $\lambda \in \text{Sp}(T)$. Wówczas $\ker (T - \lambda \text{id}) \neq \{0\}$ i z Lematu 6.3.2 mamy $V = \ker [(T - \lambda \text{id})^{m_\lambda}] \oplus \text{im} [(T - \lambda \text{id})^{m_\lambda}]$ dla pewnego $m_\lambda \geq 1$.

Przyjmijmy $U = \text{im} [(T - \lambda \text{id})^{m_\lambda}]$. Mamy $\dim U < \dim V$ i endomorfizm T przeprowadza U w siebie. Założenie indukcyjne dla obcięcia $T|U$ daje więc rozkład $U = \bigoplus_{\mu \in \text{Sp}(T|U)} \ker [(T|U - \mu \text{id})^{m_\mu}]$ dla pewnych $m_\mu \geq 1$, $\mu \in \text{Sp}(T|U)$. Z drugiej części Lematu 6.3.2 wynika, że $\ker [(T|U - \mu \text{id})^{m_\mu}] = \ker [(T - \mu \text{id})^{m_\mu}]$ dla $\mu \in \text{Sp}(T|U)$ i $\text{Sp}(T|U) = \text{Sp}(T) \setminus \{\lambda\}$, a to już daje tezę. ■

Uwaga 6.3.4 Jeśli $T : V \rightarrow V$ i m_λ dla $\lambda \in \text{Sp}(T)$ są takie jak w Twierdzeniu 6.3.3, to dla każdego $\lambda \in \text{Sp}(T)$ obcięcie $(T - \lambda \text{id})$ do $\ker [(T - \lambda \text{id})^{m_\lambda}]$ jest endomorfizmem nilpotentnym $\ker [(T - \lambda \text{id})^{m_\lambda}]$, więc zgodnie z Twierdzeniem 6.3.1, istnieją skończone (niektóre być może puste) zbiory wektorów $F_k^\lambda \subset V$, $k = 1, \dots, m_\lambda$ takie, że każde $v \in F_k^\lambda$ jest początkiem $(T - \lambda \text{id})$ -serii o wysokości k oraz

$$(*) \quad \ker [(T - \lambda \text{id})^{m_\lambda}] = \bigoplus_{k \leq m_\lambda} \left(\bigoplus_{v \in F_k^\lambda} [v] \right).$$

Przy tym dla $k \leq m_\lambda$, z warunku $(**)$ w Twierdzeniu 6.3.1 mamy

$$(**) \quad |F_k^\lambda| = \text{rank} [(T - \lambda \text{id})^{k-1}] - 2 \text{rank} [(T - \lambda \text{id})^k] + \text{rank} [(T - \lambda \text{id})^{k+1}],$$

bo $(T - \lambda \text{id})$ jest izomorfizmem na $\text{im} [(T - \lambda \text{id})^{m_\lambda}]$.

Ustalmy $v \in F_k^\lambda$, rozpatrzmy $(T - \lambda \text{id})$ -serię $v_1 = v$, $v_2 = (T - \lambda \text{id})(v_1), \dots, v_k = (T - \lambda \text{id})^{k-1}(v_{k-1})$ i niech $\sigma_v : [v] \rightarrow \mathbb{C}^k$ będzie układem współrzędnych związanym z tą serią.

Ponieważ $T(v_j) = (T - \lambda \text{id})(v_j) + \lambda v_j = v_{j+1} + \lambda v_j$, macierz endomorfizmu $T|[v] : [v] \rightarrow [v]$ w układzie współrzędnych σ_v jest *klatką Jordana*

$$J_k(\lambda) = \begin{bmatrix} \lambda & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & \lambda \end{bmatrix} \in \mathbb{C}_k^k.$$

Tak więc, w układzie współrzędnych σ_λ na $\ker [(T - \lambda \text{id})^{m_\lambda}]$, który na $[v]$ pokrywa się z σ_v , macierz $M_{\sigma_\lambda}^{\sigma_\lambda}(T|_{\ker [(T - \lambda \text{id})^{m_\lambda}]})$ jest zbudowana z klatek Jordana $J_k(\lambda)$ na przekątnej, przy czym liczba $|F_k^\lambda|$ klatek $J_k(\lambda)$ jest opisana wzorem $(**)$.

Łącząc układy współrzędnych σ_λ na składnikach rozkładu z Twierdzenia 6.3.3, dostajemy układ współrzędnych $\sigma : V \rightarrow \mathbb{C}^n$, w którym macierz $M_\sigma^\sigma(T)$ endomorfizmu T jest zbudowana z klatek Jordana $J_k(\lambda)$, $\lambda \in \text{Sp}(T)$, przy czym liczba klatek $J_k(\lambda)$ jest opisana wzorem $(**)$. □

Bazę wyznaczającą układ współrzędnych σ w Uwadze 6.3.4 będziemy nazywać *bazą Jordana dla T* . Baza Jordana jest złożona z $(T - \lambda \text{id})$ -serii generujących składniki rozkładów $(*)$ dla $\lambda \in \text{Sp}(T)$.

Macierz zbudowaną z klatek Jordana leżących na jej przekątnej będziemy nazywać *macierzą Jordana*. W języku macierzowym, nasze ustalenia można sformułować w postaci wniosku:

Wniosek 6.3.5 *Każda macierz $A \in \mathbb{C}_n^n$ jest podobna do macierzy w postaci Jordana. Dokładniej, istnieje macierz odwracalna $C \in \mathbb{C}_n^n$ taka, że*

$$C^{-1}AC = \begin{bmatrix} J_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & J_s \end{bmatrix},$$

gdzie każda macierz J_l jest klatką Jordana $J_k(\lambda)$ dla pewnego $\lambda \in \text{Sp}(A)$, a liczba klatek $J_k(\lambda)$ jest dana wzorem $\text{rank}[(A - \lambda I_n)^{k-1}] - 2 \text{rank}[(A - \lambda I_n)^k] + \text{rank}[(A - \lambda I_n)^{k+1}]$. W szczególności, macierz Jordana podobna do A jest wyznaczona jednoznacznie z dokładnością do kolejności klatek Jordana.

Dowód. Z Uwagi 6.3.4 zastosowanej do endomorfizmu $T_A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ danego wzorem $T_A(Z) = AZ$ wynika, że istnieje układ współrzędnych $\sigma : \mathbb{C}^n \rightarrow \mathbb{C}^n$ taki, że $M_\sigma^\sigma(T_A) = M(\sigma \circ T_A \circ \sigma^{-1}) = M(\sigma)AM(\sigma^{-1})$ jest macierzą Jordana. Jako C wystarczy więc przyjąć macierz $M(\sigma^{-1})$. ■

7 Liniowe przestrzenie euklidesowe

W tym rozdziale rozpatrujemy wyłącznie skończone wymiarowe przestrzenie liniowe nad ciałem liczb rzeczywistych.

Wzorując się na własnościach algebraicznych iloczynu skalarnego wektorów rozpatrywanego w geometrii, wyróżnimy pewne przekształcenia $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$, które pozwalają na określenie prostokątności wektorów i długości wektorów, przy czym spełnione jest twierdzenie Pitagorasa o długości boków trójkątów prostokątnych. Wprowadzone w ten sposób liniowe przestrzenie euklidesowe $(V, \langle \cdot, \cdot \rangle)$, dla wymiarów 2 i 3 można utożsamiać, ze względu na własności algebraiczne i geometryczne, z przestrzenią wektorów płaszczyzny euklidesowej lub trójwymiarowej przestrzeni euklidesowej, odpowiednio.

7.1 Iloczyn skalarny i norma.

Definicja 7.1.1 Niech V będzie przestrzenią liniową nad \mathbb{R} . Przekształcenie $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ przy-
porządkowujące parze wektorów $v, w \in V$ skalar $\langle v, w \rangle$ nazywamy iloczynem skalarnym jeśli

- (1) $\langle v, w \rangle = \langle w, v \rangle$,
- (2) $\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle$,
- (3) $\langle av, w \rangle = a\langle v, w \rangle$,
- (4) $\langle v, v \rangle > 0$, dla $v \neq 0$.

Normę $\|v\|$ (długość) wektora $v \in V$ określamy wzorem $\|v\| = \sqrt{\langle v, v \rangle}$.

Przestrzeń liniową z wyróżnionym iloczynem skalarnym $(V, \langle \cdot, \cdot \rangle)$ nazywamy liniową przestrzenią euklidesową.

Przykład 7.1.2 (a) Kartezjańska przestrzeń euklidesowa $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$, gdzie $\langle X, Y \rangle = X^T Y = \sum_i x_i y_i$ jest sumą iloczynów odpowiednich współrzędnych X i Y .

(b) Przestrzeń rzeczywistych wielomianów stopnia $\leq n$ z iloczynem skalarnym $\langle w, u \rangle = \int_0^1 w(t)u(t)dt$.

W oznaczeniach z Przykładu 7.1.2 (a), znaną z Analizy nierówność Cauchy'ego można zapisać w postaci $(\langle X, Y \rangle)^2 \leq \langle X, X \rangle \langle Y, Y \rangle$. Jest to szczególny przypadek nierówności Schwarz'a, którą wykażemy w następującym twierdzeniu.

Twierdzenie 7.1.3 (nierówność Schwarz'a). W liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ prawdziwa jest nierówność

$$(*) \quad (\langle v, w \rangle)^2 \leq \langle v, v \rangle \langle w, w \rangle,$$

przy czym równość ma miejsce wtedy i tylko wtedy, gdy wektory v, w są liniowo zależne.

Dowód. Dla każdego $x \in \mathbb{R}$, zgodnie z własnościami (1)–(3) w Definicji 7.1.1, $w(x) = \langle xv + w, xv + w \rangle = x^2 \langle v, v \rangle + 2x \langle v, w \rangle + \langle w, w \rangle$. Z własności (4) mamy $w(x) \geq 0$, więc $\Delta = 4(\langle v, w \rangle)^2 - 4\langle v, v \rangle \langle w, w \rangle \leq 0$, co daje (*). Ponadto, równość w (*) oznacza, że dla pewnego $\lambda \in \mathbb{R}$, $w(\lambda) = 0$, a więc $\lambda v + w = \mathbf{0}$. ■

Uwaga 7.1.4 Odnotujmy, że w liniowych przestrzeniach euklidesowych $(V, \langle \cdot, \cdot \rangle)$ norma $\|v\| = \sqrt{\langle v, v \rangle}$ ma następujące własności

- (1) $\|v\| \geq 0$ i $\|v\| = 0$ tylko dla $v = \mathbf{0}$,
- (2) $\|av\| = |a| \|v\|$, $a \in \mathbb{R}$,
- (3) $\|v + w\| \leq \|v\| + \|w\|$.

Tylko ostatnia własność – nierówność trójkąta wymaga wyjaśnienia.

Z nierówności Schwarz'a, $|\langle v, w \rangle| \leq \|v\| \|w\|$, a więc $\|v + w\|^2 = \langle v + w, v + w \rangle = \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle \leq \|v\|^2 + 2\|v\| \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2$. □

7.2 Ortogonalność w przestrzeniach euklidesowych.

Dwa wektory w geometrii są prostopadłe jeśli ich iloczyn skalarny jest zerem. Tę własność przyjmujemy jako określenie prostopadłości wektorów w liniowych przestrzeniach euklidesowych, przy czym zgodnie z powszechnie przyjętą terminologią, będziemy raczej mówili o ortogonalności.

Definicja 7.2.1 *Wektory v, w w liniowej przestrzeni euklidesowej są ortogonalne (lub prostopadłe) jeśli $\langle v, w \rangle = 0$, co zapisujemy $v \perp w$.*

Ortogonalność podprzestrzeni U, W przestrzeni V , $U \perp W$, oznacza, że $u \perp w$ dla dowolnych $u \in U$ i $w \in W$. W szczególności, ortogonalność wektora u do podprzestrzeni W , $u \perp W$ oznacza, że $\text{lin}(u) \perp W$.

Uwaga 7.2.2 Jeśli $U \perp W$, to $U \cap W = \{0\}$, bo zgodnie z 7.1.1 (4), $v \perp v$ oznacza, że $v = 0$. Jeśli ponadto $V = U \oplus W$, to będziemy mówić, że V jest sumą ortogonalną podprzestrzeni U i W . \square

Z własności iloczynu skalarnego i określenia długości wektorów wynika natychmiast formuła Pitagorasa

$$v \perp w \Leftrightarrow \|v - w\|^2 = \|v\|^2 + \|w\|^2.$$

Bez założenia ortogonalności mamy $\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\langle v, w \rangle$. Definiując kąt między niezerowymi wektorami v, w jako liczbę $\theta \in [0, \pi]$ taką, że $\cos \theta = \frac{\langle v, w \rangle}{\|v\| \|w\|}$ (z nierówności Schwarza wynika, że ułamek po prawej stronie jest w przedziale $[-1, 1]$) otrzymujemy stąd twierdzenie cosinusów:

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2 \cos \theta \|v\| \|w\|.$$

Definicja 7.2.3 *Układ wektorów (v_1, \dots, v_n) w liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ jest ortogonalny jeśli $v_i \perp v_j$ dla $i \neq j$.*

Uwaga 7.2.4 Ortogonalny układ niezerowych wektorów (v_1, \dots, v_n) jest liniowo niezależny, bo jeśli $\sum_i a_i v_i = 0$, to dla każdego j , $0 = \langle \sum_i a_i v_i, v_j \rangle = \sum_i a_i \langle v_i, v_j \rangle = a_j \|v_j\|^2$, a więc $a_j = 0$. \square

Definicja 7.2.5 *Ortogonalny układ wektorów niezerowych (v_1, \dots, v_n) w przestrzeni $(V, \langle \cdot, \cdot \rangle)$ rozpinający V nazywamy bazą ortogonalną w V ; jeśli dodatkowo $\|v_j\| = 1$, $j = 1, \dots, n$, mówimy, że układ (v_1, \dots, v_n) jest bazą ortonormalną V .*

Pokażemy teraz, że każda liniowa przestrzeń euklidesowa ma bazę ortogonalną. Co więcej, opiszemy procedurę – ortogonalizację Grama-Schmidta pozwalającą przyporządkować każdemu układowi liniowo niezależnemu układ ortogonalny, bez zmiany powłoki liniowej.

Twierdzenie 7.2.6 (ortogonalizacja Grama-Schmidta). *Niech (v_1, \dots, v_n) będzie liniowo niezależnym układem wektorów w liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$. Określmy indukcyjnie wektory w_1, \dots, w_n formułą*

$$(1) \quad w_1 = v_1, \quad w_k = v_k - \sum_{i=1}^{k-1} \frac{\langle v_k, w_i \rangle}{\langle w_i, w_i \rangle} w_i.$$

Wówczas

$$(2) \quad w_i \perp w_j \text{ dla } i \neq j,$$

$$(3) \quad \text{lin}(w_1, \dots, w_j) = \text{lin}(v_1, \dots, v_j), \quad j = 1, \dots, n.$$

Dowód. Załóżmy, że układ (w_1, \dots, w_{k-1}) został już określony tak, że (3) jest spełnione dla $j < k$. Mamy pokazać, że dla w_k określonego formułą (1) układ (w_1, \dots, w_k) jest ortogonalny i spełnia (3) dla $j = k$. Ta ostatnia własność wynika z faktu, że $w_k - v_k \in \text{lin}(w_1, \dots, w_{k-1}) = \text{lin}(v_1, \dots, v_{k-1})$, zob. (1).

Niech $j < k$. Sprawdźmy, że $w_j \perp w_k$. Istotnie, z (1) i (2),

$$\langle w_j, w_k \rangle = \langle w_j, v_k - \sum_{i < k} \frac{\langle v_k, w_i \rangle}{\langle w_i, w_i \rangle} w_i \rangle = \langle w_j, v_k \rangle - \sum_{i < k} \frac{\langle v_k, w_i \rangle}{\langle w_i, w_i \rangle} \langle w_j, w_i \rangle = \langle w_j, v_k \rangle - \frac{\langle v_k, w_j \rangle}{\langle w_j, w_j \rangle} \langle w_j, w_j \rangle = 0 \blacksquare$$

Wyjaśnimy teraz, że w opisanej wyżej procedurze ortogonalizacji, na każdym kroku znajdujemy wektor w_k odejmując od wektora v_k rzut ortogonalny v_k na podprzestrzeń rozpiętą przez poprzednie wektory.

Twierdzenie 7.2.7 Niech W będzie podprzestrzenią liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$. Istnieje wówczas przekształcenie liniowe $P : V \rightarrow V$ – rzut ortogonalny V na W , takie, że

- (1) $P(w) = w$ dla $w \in W$
- (2) $v - P(v) \perp W$ dla $v \in V$.

Dowód. Zgodnie z Twierdzeniem 7.2.6 możemy wybrać układ ortogonalny w_1, \dots, w_m rozpinający W i niech

$$P(v) = \sum_{i=1}^m \frac{\langle v, w_i \rangle}{\langle w_i, w_i \rangle} w_i.$$

Podobnie jak w dowodzie 7.2.6, dla $j \leq m$,

$$\langle w_j, v - P(v) \rangle = \langle w_j, v - \sum_{i \leq m} \frac{\langle v, w_i \rangle}{\langle w_i, w_i \rangle} w_i \rangle = \langle w_j, v \rangle - \sum_{i \leq m} \frac{\langle v, w_i \rangle}{\langle w_i, w_i \rangle} \langle w_j, w_i \rangle = \langle w_j, v \rangle - \frac{\langle v, w_j \rangle}{\langle w_j, w_j \rangle} \langle w_j, w_j \rangle = 0,$$

a stąd $\langle w, v - P(v) \rangle = 0$ dla $w = \sum_{j \leq m} a_j w_j \in W$, czyli (2).

Jeśli $w \in W$, to $w - P(w) \in W$ i z (2), $\langle w - P(w), w - P(w) \rangle = 0$, co daje $w - P(w) = \mathbf{0}$, czyli (1). ■

Uwaga 7.2.8 Zauważmy, że w sytuacji opisanej w Twierdzeniu 7.2.7, dla $v \in V$, warunki $w \in W$ oraz $v - w \perp W$ charakteryzują wektor w jednoznacznie, co więcej jest to jedyny wektor z W taki, że $\|v - w\| = \inf \{\|v - u\| : u \in W\}$. Istotnie, dla $u \in W$, $u - w \in W$, więc $(v - w) \perp (u - w)$ i z formuły Piagorasa, $\|v - u\|^2 = \|(v - w) + (w - u)\|^2 = \|v - w\|^2 + \|w - u\|^2 \geq \|v - w\|^2$, przy czym dla $u \neq w$ nierówność jest ostra.

Tak więc rzut ortogonalny $P : V \rightarrow W$ można opisać jako przekształcenie, które dla każdego wektora $v \in V$ wybiera z W (jedyny) wektor $P(v)$ minimalizujący długości $\|v - u\|$, $u \in W$. □

Jak zobaczymy za chwilę, istnienie rzutu ortogonalnych pozwala uzupełnić każdą podprzestrzeń liniowej przestrzeni euklidesowej składnikiem prostym, ortogonalnym do tej przestrzeni.

Twierdzenie 7.2.9 Niech W będzie podprzestrzenią liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ i niech $W^\perp = \{v \in V : v \perp W\}$. Wówczas $V = W \oplus W^\perp$ jest sumą ortogonalną.

Dowód. Niech $P : V \rightarrow W$ będzie rzutem ortogonalnym na W . Zauważmy, że $\text{im } P = W$ oraz $\ker P = W^\perp$. Ponieważ $W \cap W^\perp = \{\mathbf{0}\}$ i $\dim V = \text{def } P + \text{rank } P$, więc $V = W \oplus W^\perp$ i oczywiście $W \perp W^\perp$. ■

Podprzestrzeń W^\perp nazywamy *dopełnieniem ortogonalnym* przestrzeni W w V . Zauważmy, że rzut ortogonalny V na W jest rzutem na W równoległym do W^\perp zdefiniowanym w 4.1.7. Symetrię względem W równoległą do W^\perp nazywamy *symetrią ortogonalną* względem W .

7.3 Wyznacznik Grama i objętość.

W tej części wyjaśnimy, że moduł wyznacznika macierzy kwadratowej można interpretować jako objętość równoległościanu rozpiętego na wierszach (równoważnie – kolumnach) tej macierzy w wielowymiarowej przestrzeni euklidesowej. Ważnym elementem naszych rozważań jest wyznacznik Grama wprowadzony poniżej.

Definicja 7.3.1 Macierzą Grama układu wektorów (v_1, \dots, v_m) w liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ nazywamy $(m \times m)$ -macierz

$$G(v_1, \dots, v_m) = \begin{bmatrix} \langle v_1, v_1 \rangle & \cdots & \langle v_1, v_m \rangle \\ \vdots & & \vdots \\ \langle v_m, v_1 \rangle & \cdots & \langle v_m, v_m \rangle \end{bmatrix}.$$

Wyznacznik $\Gamma(v_1, \dots, v_m) = \det(G(v_1, \dots, v_m))$ nazywamy *wyznacznikiem Grama* układu (v_1, \dots, v_m) .

Uwaga 7.3.2 Macierz Grama $G(v_1, \dots, v_m)$ układu wektorów w $(V, \langle \cdot, \cdot \rangle)$ wyznacza wartości iloczynu skalarnego na $\text{lin}(v_1, \dots, v_m)$: dla $v = \sum_i a_i v_i$, $w = \sum_j a_j v_j$,

$$\langle v, w \rangle = \left\langle \sum_i a_i v_i, \sum_j b_j v_j \right\rangle = \sum_{ij} a_i b_j \langle v_i, v_j \rangle = [a_1, \dots, a_m] G(v_1, \dots, v_m) \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}. \quad \square$$

Lemat 7.3.3 Niech (v_1, \dots, v_m) będzie układem wektorów w $(V, \langle \cdot, \cdot \rangle)$ i niech $w_k = \sum_i a_{ik} v_i$, $k \leq m$. Wówczas $\Gamma(w_1, \dots, w_m) = (\det A)^2 \Gamma(v_1, \dots, v_m)$, dla macierzy $A = [a_{ij}]_{i,j=1}^m$.

Dowód. Niech $A = [A_1, \dots, A_m]$ (tzn. A_k jest k -tą kolumną macierzy A) i połóżmy $G(v_1, \dots, v_m) = G$. Zgodnie z Uwagą 7.3.2, (k, l) -ty wyraz macierzy $G(w_1, \dots, w_m)$ ma postać $\langle w_k, w_l \rangle = A_k^T G A_l$, więc $G(w_1, \dots, w_m) = A^T G A$, a stąd $\Gamma(w_1, \dots, w_m) = \det A^T \det G \det A = (\det A)^2 \Gamma(v_1, \dots, v_m)$. ■

Uwaga 7.3.4 Jeśli układ (v_1, \dots, v_m) jest liniowo zależny, to $\Gamma(v_1, \dots, v_m) = 0$, bo wówczas wiersze macierzy $G(v_1, \dots, v_m)$ są liniowo zależne. □

Jak wynika w szczególności z kolejnego twierdzenia, wyznacznik Grama układu liniowo niezależnego jest dodatni.

Twierdzenie 7.3.5 Niech (v_1, \dots, v_m) będzie układem wektorów w liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$. Jeśli układ wektorów (w_1, \dots, w_m) otrzymuje się w procesie ortogonalizacji Grama-Schmidta z układu (v_1, \dots, v_m) , to $\Gamma(v_1, \dots, v_m) = \Gamma(w_1, \dots, w_m) = \|w_1\|^2 \cdot \dots \cdot \|w_m\|^2$.

Dowód. Zgodnie z formułą (1) w Twierdzeniu 7.2.6 $v_k = \sum_i a_{ik} w_i$, gdzie macierz $A = [a_{ij}]_{i,j=1}^m$ ma na głównej przekątnej jedynek, a pod główną przekątną zera. Zatem $\det A = 1$ i z Lematu 7.3.3 mamy $\Gamma(v_1, \dots, v_m) = \Gamma(w_1, \dots, w_m)$. Ponadto macierz $G(w_1, \dots, w_m)$ jest diagonalna i ma na głównej przekątnej wyrazy $\langle w_i, w_i \rangle = \|w_i\|^2$, więc $\Gamma(w_1, \dots, w_m) = \|w_1\|^2 \cdot \dots \cdot \|w_m\|^2$. ■

Wniosek 7.3.6 Niech W będzie podprzestrzenią liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ rozpiętą na liniowo niezależnym układzie (v_1, \dots, v_m) i niech $P : V \rightarrow W$ będzie rzutem ortogonalnym V na W . Wówczas dla $v \in V$,

$$\|v - P(v)\| = \sqrt{\frac{\Gamma(v_1, \dots, v_m, v)}{\Gamma(v_1, \dots, v_m)}}.$$

Dowód. Jeśli $v \in W$, to po obu stronach równości mamy zero, zob. Uwaga 7.3.4. Niech $v \notin W$ i niech układ (w_1, \dots, w_m, w) będzie wynikiem ortogonalizacji Grama-Schmidta układu (v_1, \dots, v_m, v) . Wówczas $w = v - P(v)$ oraz

$$\Gamma(v_1, \dots, v_m, v) = \|w_1\|^2 \cdot \dots \cdot \|w_m\|^2 \cdot \|w\|^2 \quad \text{i} \quad \Gamma(v_1, \dots, v_m) = \|w_1\|^2 \cdot \dots \cdot \|w_m\|^2,$$

co dowodzi tezy. ■

Uwaga 7.3.7 Wyznacznik Grama układu wektorów nie zależy od ich kolejności.

Istotnie, niech (v_1, \dots, v_m) będzie układem wektorów w $(V, \langle \cdot, \cdot \rangle)$ i $\pi : \{1, \dots, m\} \rightarrow \{1, \dots, m\}$ permutacją. Wówczas $v_{\pi(j)} = \sum_i a_{ij} v_i$, gdzie $A = [a_{ij}]_{i,j=1}^m$ jest macierzą powstałą z permutacji kolumn macierzy jednostkowej. Zatem $(\det A)^2 = 1$ i zgodnie z Lematem 7.3.3, $\Gamma(v_{\pi(1)}, \dots, v_{\pi(m)}) = \Gamma(v_1, \dots, v_m)$. □

Dla układu wektorów (v_1, \dots, v_m) w liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ równoległościaniem rozpiętym na tym układzie nazywamy zbiór

$$R(v_1, \dots, v_m) = \left\{ \sum_{i=1}^m t_i v_i : t_i \in [0, 1] \right\}.$$

Jeśli układ (v_1, v_2, v_3) jest liniowo niezależny, to $R(v_1, v_2)$ jest równoległobokiem na płaszczyźnie $\text{lin}(v_1, v_2)$, a $R(v_1, v_2, v_3)$ jest równoległościaniem w trójwymiarowej przestrzeni euklidesowej $\text{lin}(v_1, v_2, v_3)$.

Zgodnie z określeniem przyjętym w geometrii euklidesowej objętość równoległościanu $R(v_1, \dots, v_m)$ (dla $m = 2$, pole równoległoboku $R(v_1, v_2)$) powinna być równa iloczynowi objętości podstawy $R(v_1, \dots, v_{m-1})$ przez wysokość $\|v_m - P(v_m)\|$, gdzie $P(v_m)$ jest rzutem ortogonalnym v_m na przestrzeń $\text{lin}(v_1, \dots, v_{m-1})$ rozpiętą na podstawie.

Definicja 7.3.8 Objętością (n -wymiarową) układu (v_1, \dots, v_m) w liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ nazywamy liczbę

$$\text{vol}(v_1, \dots, v_m) = \sqrt{\Gamma(v_1, \dots, v_m)}.$$

Natychmiastową konsekwencją Wniosku 7.3.6 jest

Wniosek 7.3.9 Funkcja $\text{vol}(v_1, \dots, v_m)$ jest jedyną funkcją, która układowi wektorów (v_1, \dots, v_m) w liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ przyporządkowuje liczbę nieujemną w taki sposób, że

- (1) $\text{vol}(v) = \|v\|$
- (2) $\text{vol}(v_1, \dots, v_m) = \|v_m - P(v_m)\| \text{vol}(v_1, \dots, v_{m-1})$,

gdzie P jest rzutem ortogonalnym przestrzeni V na podprzestrzeń $\text{lin}(v_1, \dots, v_{m-1})$.

Uwaga 7.3.10 Rozważmy kartezjańską przestrzeń euklidesową z Przykładu 7.1.2 (a) i niech $A \in \mathbb{K}_n^n$ będzie macierzą o kolumnach A_1, \dots, A_n .

Wtedy $G(A_1, \dots, A_n) = A^T A$, więc $\Gamma(A_1, \dots, A_n) = (\det A)^2$ i mamy

$$\text{vol}(A_1, \dots, A_n) = |\det A|. \quad \square$$

Następująca obserwacja pokazuje, że moduł wyznacznika endomorfizmu liniowej przestrzeni euklidesowej można interpretować jako współczynnik zmiany objętości przy tym endomorfizmie pełnowymiarowych równoległościanów.

Twierdzenie 7.3.11 Niech $T : V \rightarrow V$ będzie endomorfizmem liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$. Dla dowolnej bazy (v_1, \dots, v_n) w V zachodzi wówczas równość

$$\text{vol}(T(v_1), \dots, T(v_n)) = |\det T| \cdot \text{vol}(v_1, \dots, v_n).$$

Dowód. Niech $A = [a_{ij}]_{i,j=1}^m$ będzie macierzą endomorfizmu T w układzie współrzędnych związanym z bazą (v_1, \dots, v_n) przestrzeni V , tzn. $T(v_j) = \sum_i a_{ij} v_i$. Z Lematu 7.3.3

$$\Gamma(T(v_1), \dots, T(v_n)) = (\det A)^2 \cdot \Gamma(v_1, \dots, v_n),$$

a ponieważ $\det T = \det A$, otrzymujemy tezę twierdzenia. ■

7.4 Orientacja i iloczyn wektorowy.

Każdy wybór układu współrzędnych $\sigma : V \rightarrow \mathbb{R}^n$ rozбивa układy współrzędnych w przestrzeni V na dwie klasy: układy współrzędnych $\tau : V \rightarrow \mathbb{R}^n$, dla których $\det(\tau \circ \sigma^{-1}) > 0$ oraz takie, dla których ten wyznacznik jest ujemny. To pozwala mówić o orientacji przestrzeni liniowej V nad \mathbb{R} . Wygodniej przy tym będzie skupić uwagę na bazach (v_1, \dots, v_n) w V , a nie na związanych z nimi układach współrzędnych $\sigma : V \rightarrow \mathbb{R}^n$, $\sigma(v_j) = E_j$.

Definicja 7.4.1 Dwie bazy (v_1, \dots, v_n) i (w_1, \dots, w_n) w przestrzeni liniowej V nad \mathbb{R} są zorientowane zgodnie (przeciwnie) jeśli dla związanych z nimi układów współrzędnych $\sigma : V \rightarrow \mathbb{R}^n$, $\tau : V \rightarrow \mathbb{R}^n$, $\det(\tau \circ \sigma^{-1}) > 0$ ($\det(\tau \circ \sigma^{-1}) < 0$). Wybór bazy (v_1, \dots, v_n) w V ustala orientację V – klasę wszystkich baz w V , które są z nią zgodnie zorientowane. Orientacja standardowa \mathbb{R}^n jest wyznaczona przez bazę standardową (E_1, \dots, E_n) .

W przestrzeniach euklidesowych $(V, \langle \cdot, \cdot \rangle)$ pojęcie orientacji ma klarowną interpretację geometryczną: bazy (v_1, \dots, v_n) i (w_1, \dots, w_n) mają zgodne orientacje jeśli dla każdego t z przedziału $[0, 1]$ istnieje baza $(u_1(t), \dots, u_n(t))$ taka, że $u_j(0) = v_j$, $u_j(1) = w_j$, oraz wektory $u_j(t)$ zależą od t w sposób ciągły, tzn. $t_m \rightarrow t_0$ implikuje $\|u_j(t_m) - u_j(t_0)\| \rightarrow 0$. Wyjaśnimy to w części 14.3 uzupełnień.

Własność opisana w następnym twierdzeniu wyróżnia klasę izomorfizmów zachowujących orientację V .

Twierdzenie 7.4.2 Niech $T : V \rightarrow V$ będzie izomorfizmem. Wówczas $\det T > 0$ wtedy i tylko wtedy, gdy pewna (równoważnie – każda) baza (v_1, \dots, v_n) w V jest zgodnie zorientowana z bazą $(T(v_1), \dots, T(v_n))$.

Dowód. Niech $\sigma : V \rightarrow \mathbb{R}^n$ będzie układem współrzędnych związanym z bazą $(T(v_1), \dots, T(v_n))$. Wówczas $\tau = \sigma \circ T : V \rightarrow \mathbb{R}^n$ jest układem współrzędnych związanym z bazą (v_1, \dots, v_n) . Ponieważ $\tau \circ \sigma^{-1} = \sigma \circ T \circ \sigma^{-1}$, więc $\det(\tau \circ \sigma^{-1}) = \det M_\sigma^\sigma(T)$. ■

Dla każdego liniowo niezależnego układu (v_1, \dots, v_{n-1}) , w n -wymiarowej liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ istnieją dokładnie dwa wektory o normie 1 prostopadłe do $\text{lin}(v_1, \dots, v_{n-1})$. Ustalona orientacja V pozwala wyróżnić jeden z nich, co prowadzi do ważnej operacji iloczynu wektorowego.

Definicja 7.4.3 Niech baza (u_1, \dots, u_n) wyznacza orientację w liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$. Dla każdego liniowo niezależnego układu (v_1, \dots, v_{n-1}) w V , iloczynem wektorowym wektorów tego układu nazywamy (jednoznacznie wyznaczony) wektor $v_1 \times \dots \times v_{n-1}$ spełniający następujące warunki:

- (1) $v_1 \times \dots \times v_{n-1} \in \text{lin}(v_1, \dots, v_{n-1})^\perp$,
- (2) $\|v_1 \times \dots \times v_{n-1}\| = \text{vol}(v_1, \dots, v_{n-1})$,
- (3) baza $(v_1, \dots, v_{n-1}, v_1 \times \dots \times v_{n-1})$ jest zorientowana zgodnie z (u_1, \dots, u_n) .

Jeśli układ (v_1, \dots, v_{n-1}) jest liniowo zależny, przyjmujemy $v_1 \times \dots \times v_{n-1} = \mathbf{0}$.

Twierdzenie 7.4.4 Niech $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ będzie kartezjańską przestrzenią euklidesową, $\langle X, Y \rangle = X^T Y$, z orientacją standardową. Wówczas dla dowolnego układu wektorów (A_1, \dots, A_{n-1}) w \mathbb{R}^n

$$(*) \quad A_1 \times \dots \times A_{n-1} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}, \quad \text{gdzie } b_i = (-1)^{i+n} \det A(i) \text{ dla macierzy } A(i) \text{ otrzymanej} \\ \text{z macierzy } [A_1, \dots, A_{n-1}] \text{ przez skreślenie } i\text{-tego wiersza.}$$

Dowód. Niech $B = [b_1, \dots, b_n]^T$. Wówczas $\langle A_i, B \rangle = A_i^T B = \det[A_1, \dots, A_{n-1}, A_i]$, co otrzymujemy rozwijając wyznacznik względem ostatniej kolumny, a ponieważ wyznacznik tej macierzy jest zerem, bo i -ta kolumna jest równa n -tej, $\langle A_i, B \rangle = 0$. Zatem $B \perp \text{lin}(A_1, \dots, A_{n-1})$.

Niech $M = \det[A_1, \dots, A_{n-1}, B]$. Wówczas $\det M = \text{vol}(A_1, \dots, A_{n-1}, B) = \text{vol}(A_1, \dots, A_{n-1}) \|B\|$. Rozwijając $\det M$ względem ostatniej kolumny M widzimy też, że $\det M = \langle B, B \rangle = \|B\|^2 > 0$. Tak więc $\|B\| = \text{vol}(A_1, \dots, A_{n-1})$ i układ kolumn M jest zorientowany zgodnie z bazą standardową \mathbb{R}^n . ■

Z wzoru (*) w Twierdzeniu 7.4.4 wynika, że iloczyn wektorowy w przestrzeni $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ ze standardową orientacją jest liniowy ze względu na każdy czynnik z osobna i antysymetryczny (przestawienie dwóch argumentów zmienia znak iloczynu).

Jak zobaczymy w następnej części, jest tak w dowolnej n -wymiarowej liniowej przestrzeni euklidesowej z orientacją, bo układ współrzędnych związany z ortonormalną bazą dodatnio zorientowaną identyfikuje tę przestrzeń z $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ z orientacją standardową.

7.5 Prostokątne układy współrzędnych.

Wśród układów współrzędnych na liniowej przestrzeni euklidesowej wyróżnioną rolę pełnią układy zachowujące iloczyn skalarny – prostokątne układy współrzędnych.

Definicja 7.5.1 Układ współrzędnych $\sigma : V \rightarrow \mathbb{R}^n$ w liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ związany z bazą ortonormalną V , będziemy nazywać prostokątnym układem współrzędnych.

Uwaga 7.5.2 Jeśli układ współrzędnych $\sigma : V \rightarrow \mathbb{R}^n$ w przestrzeni $(V, \langle \cdot, \cdot \rangle)$ jest związany z bazą ortonormalną (v_1, \dots, v_n) , to

$$(a) \quad \sigma(u) = \begin{bmatrix} \langle u, v_1 \rangle \\ \vdots \\ \langle u, v_n \rangle \end{bmatrix} \quad \text{dla } u \in V,$$

bo mnożąc $u = \sum_i x_i v_i$ obustronnie przez v_j dostajemy $\langle u, v_j \rangle = x_j$;

$$(b) \quad \langle u, v \rangle = \langle \sigma(u), \sigma(v) \rangle \quad \text{dla } u, v \in V,$$

gdzie po prawej stronie równości jest iloczyn skalarny w kartezjańskiej przestrzeni euklidesowej $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ opisanej w Przykładzie 7.1.2 (a).

Aby uzasadnić (b) rozpatrzmy $u = \sum_i x_i v_i$, $v = \sum_j y_j v_j$ i zauważmy, że z warunku $v_i \perp v_j$ dla $i \neq j$ oraz $\langle v_i, v_i \rangle = 1$ wynika $\langle u, v \rangle = \langle \sum_i x_i v_i, \sum_j y_j v_j \rangle = \sum_i x_i y_i = \langle \sigma(u), \sigma(v) \rangle$. \square

Zauważmy też, że warunek zachowania iloczynu skalarnego odnotowany w (b) charakteryzuje prostokątne układy współrzędnych.

Uwaga 7.5.3 Z istnienia prostokątnych układów współrzędnych wynika także, że dla każdych dwóch n -wymiarowych liniowych przestrzeni euklidesowych $(V, \langle \cdot, \cdot \rangle)$ i $(W, \langle \cdot, \cdot \rangle)$ istnieje izomorfizm $T : V \rightarrow W$ zachowujący iloczyn skalarny, tzn. $\langle u, v \rangle = \langle T(u), T(v) \rangle$.

Istotnie, dla prostokątnych układów współrzędnych $\sigma : V \rightarrow \mathbb{R}^n$ i $\tau : W \rightarrow \mathbb{R}^n$, z Uwagi 7.5.2 (b) wynika, że $T = \tau^{-1} \circ \sigma$ zachowuje iloczyn skalarny. \square

8 Endomorfizmy przestrzeni euklidesowych

W tym rozdziale omówimy dwie ważne klasy endomorfizmów $T : V \longrightarrow V$ liniowych przestrzeni euklidesowych $(V, \langle \cdot, \cdot \rangle)$ – endomorfizmy samosprężone (tzn. takie, że $\langle v, T(w) \rangle = \langle T(v), w \rangle$) oraz izometrie liniowe (spełniające warunek $\langle v, w \rangle = \langle T(v), T(w) \rangle$).

Głównymi wynikami są dwa twierdzenia, z których pierwsze mówi, że endomorfizmy samosprężone mają w pewnym prostokątnym układzie współrzędnych macierz diagonalną, a drugie, że izometrie liniowe mają w pewnym prostokątnym układzie współrzędnych macierz mającą na przekątnej albo ± 1 , albo też

macierze obrotu płaszczyzny $\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$.

Oba te fakty można wyprowadzić z zasadniczego twierdzenia algebry, przy czym macierz obrotu o kąt θ odpowiada zespolonej wartości własnej $\lambda = \cos \theta + i \sin \theta$ izometrii liniowej. Takie uzasadnienie naszkicujemy w części 14.2 uzupełnień.

Warto też jednak pokazać dowody tych twierdzeń nie wychodzące poza dziedzinę rzeczywistą i to podejście przedstawimy poniżej.

8.1 Endomorfizmy samosprężone.

Definicja 8.1.1 *Endomorfizm $T : V \longrightarrow V$ liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ jest samosprężony, jeśli $\langle v, T(w) \rangle = \langle T(v), w \rangle$ dla $v, w \in V$.*

Uwaga 8.1.2 Niech $A \in \mathbb{R}^n$. Endomorfizm $T(X) = AX$ kartezjańskiej przestrzeni euklidesowej $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ jest samosprężony wtedy i tylko wtedy, gdy macierz A jest symetryczna, czyli $A = A^T$.

Istotnie, warunek $\langle X, T(Y) \rangle = \langle T(X), Y \rangle$ oznacza, że $X^T A Y = (A X)^T Y = X^T A^T Y$, a więc, z dowolności $X, Y \in \mathbb{R}^n$, jest on równoważny symetrii macierzy A . \square

Dwie podstawowe własności endomorfizmów samosprężonych (charakteryzujące tę klasę), to istnienie wektorów własnych i zachowywanie przestrzeni ortogonalnych do podprzestrzeni niezmienniczych.

Drugą z tych własności bardzo łatwo uzasadnić.

Uwaga 8.1.3 Niech $T : V \longrightarrow V$ będzie endomorfizmem samosprężonym liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ i niech $W \subset V$ będzie podprzestrzenią taką, że $T(W) \subset W$. Wówczas $T(W^\perp) \subset W^\perp$.

Istotnie, jeśli $v \in W^\perp$, to dla dowolnego $w \in W$ mamy $0 = \langle v, T(w) \rangle = \langle T(v), w \rangle$, a więc $T(v) \in W^\perp$. \square

Przejdziemy teraz do dowodu istnienia wektorów własnych dla endomorfizmów samosprężonych.

Twierdzenie 8.1.4 *Dla każdego endomorfizmu samosprężonego $T : V \longrightarrow V$ liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ istnieje wektor własny $w \in V$ taki, że*

$$(*) \quad \|w\| = 1 \quad \text{oraz} \quad \langle w, T(w) \rangle = \sup \{ \langle v, T(v) \rangle : \|v\| = 1 \}.$$

Dowód. Istnienie wektora spełniającego (*) wynika z twierdzenia Weierstrassa zastosowanego do funkcji $v \longrightarrow \langle v, T(v) \rangle$ na sferze jednostkowej $\{v \in V : \|v\| = 1\}$, co wyjaśnimy w uzupełnieniach, zob. Lemat 14.1.2.

Dla dowodu twierdzenia wystarczy teraz pokazać, że wektor w spełniający (*) jest wektorem własnym T . Tak jest jeśli $T(w) = 0$. Załóżmy więc $T(w) \neq 0$ i dla t z przedziału $J = (-\frac{\|w\|}{\|T(w)\|}, \frac{\|w\|}{\|T(w)\|}) \subset \mathbb{R}$ połóżmy

$$v(t) = \frac{1}{\|w + tT(w)\|} (w + tT(w)).$$

Określmy funkcję $\phi : J \longrightarrow \mathbb{R}$ wzorem

$$\phi(t) = \langle v(t), T(v(t)) \rangle = \frac{\langle w, T(w) \rangle + 2t \langle T(w), T(w) \rangle + t^2 \langle T(w), T^2(w) \rangle}{\langle w, w \rangle + 2t \langle w, T(w) \rangle + t^2 \langle T(w), T(w) \rangle}$$

(w liczniku skorzystaliśmy z równości $\langle w, T^2(w) \rangle = \langle T(w), T(w) \rangle$ wynikającej z samosprężoności T).

Ponieważ $\|v(t)\| = 1$ i $v(0) = w$, z (*) wynika, że funkcja ϕ osiąga maksimum w punkcie $t = 0$, a więc $\phi'(0) = 0$. Z drugiej strony,

$$\phi'(0) = \frac{2\langle T(w), T(w) \rangle \langle w, w \rangle - 2\langle w, T(w) \rangle \langle w, T(w) \rangle}{\langle w, w \rangle^2},$$

więc $\langle w, T(w) \rangle^2 = \langle w, w \rangle \langle T(w), T(w) \rangle$ i z warunku równości w nierówności Schwarz'a, zob. 7.1.3, dostajemy $T(w) \in \text{lin}(w)$. ■

Łącząc Twierdzenie 8.1.4 i Uwagę 8.1.3 otrzymujemy główne twierdzenie tej części

Twierdzenie 8.1.5 *Dla każdego endomorfizmu samosprężonego $T : V \rightarrow V$ liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ istnieje baza ortonormalna V złożona z wektorów własnych T . W prostokątnym układzie współrzędnych związanym z tą bazą endomorfizm T ma macierz diagonalną.*

Dowód. Załóżmy, że pierwsza część tezy jest prawdziwa dla endomorfizmów samosprężonych na przestrzeniach euklidesowych wymiaru mniejszego niż $\dim V$.

Niech w będzie wektorem własnym dla T , $\|w\| = 1$ i niech $U = \text{lin}(w)^\perp$. Z Uwagi 8.1.3, obcięcie $T|_U$ jest endomorfizmem samosprężonym U i założenie indukcyjne zapewnia istnienie bazy ortonormalnej U złożonej z wektorów własnych T . Dołączając do tej bazy wektor w dostajemy ortonormalną bazę V złożoną z wektorów własnych T . Druga część tezy wynika z części pierwszej, zob. Uwaga 6.1.11. ■

Przed przeformułowaniem tego twierdzenia w języku macierzowym, wprowadzimy klasę macierzy ortogonalnych, ściśle związanych z izometriami liniowymi.

Definicja 8.1.6 *Macierz $C \in \mathbb{R}^n$ jest ortogonalna, jeśli $C^T = C^{-1}$.*

Zauważmy, że warunek ortogonalności, $C^T C = I_n$ oznacza, że kolumny macierzy C tworzą bazę ortonormalną w kartezjańskiej przestrzeni euklidesowej $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$. Ponadto, ortogonalność C jest równoważna ortogonalności macierzy C^T , bo warunki $C^T C = I_n$ i $CC^T = I_n$ są równoważne.

Wniosek 8.1.7 *Dla każdej macierzy symetrycznej $A \in \mathbb{R}^n$ istnieje macierz ortogonalna $C \in \mathbb{R}^n$ taka, że macierz $C^T A C$ jest diagonalna.*

Dowód. Rozpatrzmy kartezjańską przestrzeń euklidesową $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ i endomorfizm $T(X) = AX$ tej przestrzeni. Jak zauważyliśmy w Uwadze 8.1.2, endomorfizm T jest samosprężony i niech $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^n$ będzie prostokątnym układem współrzędnych związanym z ortonormalną bazą (C_1, \dots, C_n) w $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$, złożoną z wektorów własnych T . Macierz $M_\sigma^\sigma(T) = M(\sigma \circ T \circ \sigma^{-1}) = M(\sigma)AM(\sigma^{-1})$ jest diagonalna, a $C = M(\sigma^{-1}) = [C_1, \dots, C_n]$ jest macierzą ortogonalną, więc $C^T A C = C^{-1} A C = M_\sigma^\sigma(T)$. ■

8.2 Izometrie liniowe.

Definicja 8.2.1 *Izomorfizm $T : V \rightarrow W$ liniowych przestrzeni euklidesowych $(V, \langle \cdot, \cdot \rangle)_V$ i $(W, \langle \cdot, \cdot \rangle)_W$ nazywamy izometrią liniową, jeśli T zachowuje iloczyn skalarny, tzn. spełnia warunek*

$$\langle u, v \rangle_V = \langle T(u), T(v) \rangle_W \quad \text{dla } u, v \in V.$$

W szczególności, każdy prostokątny układ współrzędnych $\sigma : V \rightarrow \mathbb{R}^n$ jest izometrią liniową, a w 7.5.3 zauważyliśmy, że dla dowolnych dwóch n -wymiarowych liniowych przestrzeni euklidesowych $(V, \langle \cdot, \cdot \rangle_V)$ i $(W, \langle \cdot, \cdot \rangle_W)$ istnieje izometria liniowa V na W .

Uwaga 8.2.2 Endomorfizm $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ kartezyjskiej przestrzeni euklidesowej $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ jest izometrią liniową wtedy i tylko wtedy, gdy macierz $M(T)$ jest ortogonalna.

Istotnie, jeśli T jest izometrią liniową, to kolumny macierzy $M(T) = [T(E_1), \dots, T(E_n)]$ tworzą bazę ortonormalną \mathbb{R}^n jako obrazy wektorów ortonormalnej bazy standardowej, więc $M(T)^T M(T) = I_n$, czyli $M(T)$ jest macierzą ortogonalną. Jeśli $M = M(T)$ jest macierzą ortogonalną, to dla $X, Y \in \mathbb{R}^n$ mamy $\langle T(X), T(Y) \rangle = (MX)^T MY = X^T (M^T M) Y = X^T I_n Y = \langle X, Y \rangle$, więc T jest izometrią liniową. \square

Uwaga 8.2.3 Dla każdej izometrii liniowej $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dwuwymiarowej kartezyjskiej przestrzeni euklidesowej $(\mathbb{R}^2, \langle \cdot, \cdot \rangle)$ istnieje prostokątny układ współrzędnych $\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ taki, że

$$M_\sigma^\sigma(T) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (\text{i wówczas } T \text{ jest obrotem } \mathbb{R}^2 \text{ o kąt } \theta, \sigma = \text{id}_{\mathbb{R}^2})$$

lub też

$$M_\sigma^\sigma(T) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (\text{i wówczas } T \text{ jest symetrią ortogonalną } V \text{ względem pewnej prostej}).$$

Istotnie, macierz $M(T) = [T(E_1), T(E_2)]$ jest ortogonalna. Jeśli $T(E_1) = [a, b]^T$ dla $a, b \in \mathbb{R}^2$, to $M(T)$ ma postać

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \quad \text{lub} \quad \begin{bmatrix} a & b \\ b & -a \end{bmatrix}, \quad \text{gdzie } a^2 + b^2 = 1.$$

W pierwszym przypadku $\det T = 1$ i wówczas T jest obrotem \mathbb{R}^2 o kąt θ taki, że $a = \cos \theta$, $b = \sin \theta$.

W drugim przypadku $\det T = -1$, wielomian charakterystyczny $w_T(x) = x^2 - 1 = (1 - x)(-1 - x)$, więc \mathbb{R}^2 ma ortonormalną bazę (C_1, C_2) taką, że $T(C_1) = C_1$ i $T(C_2) = -C_2$. W prostokątnym układzie współrzędnych $\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ związanym z tą bazą $M_\sigma^\sigma(T)$ jest macierzą diagonalną z 1, -1 na przekątnej. \square

Z tej uwagi wynika, że każda izometria liniowa $S : W \rightarrow W$ płaszczyzny euklidesowej $(W, \langle \cdot, \cdot \rangle)$ jest obrotem tej płaszczyzny (jeśli $\det S = 1$) lub symetrią ortogonalną względem pewnej prostej $U \subset W$ (jeśli $\det S = -1$). Pokażemy, że izometrie liniowe przestrzeni euklidesowej wymiaru ≥ 2 są złożeniem obrotów we wzajemnie ortogonalnych płaszczyznach i symetrii ortogonalnej względem podprzestrzeni zawierającej te płaszczyzny.

Twierdzenie 8.2.4 Dla każdej izometrii liniowej $T : V \rightarrow V$ liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ wymiaru n istnieje prostokątny układ współrzędnych $\sigma : V \rightarrow \mathbb{R}^n$ taki, że $M_\sigma^\sigma(T)$ ma na przekątnej macierze obrotów $\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ lub skalary ± 1 .

W języku macierzowym Twierdzenie 8.2.4 formułuje się następująco.

Wniosek 8.2.5 Dla każdej macierzy ortogonalnej $A \in \mathbb{R}_n^n$ istnieje macierz ortogonalna C taka, że $C^T A C (= C^{-1} A C)$ ma na przekątnej macierze obrotów lub skalary ± 1 .

Do dowodu twierdzenia potrzebna będzie następująca obserwacja wynikająca z Twierdzenia 8.1.4.

Lemat 8.2.6 Dla każdej izometrii liniowej $T : V \rightarrow V$ liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ wymiaru ≥ 2 istnieje podprzestrzeń $W \subset V$ taka, że $T(W) \subset W$ oraz $\dim W \leq 2$.

Dowód. Zauważmy najpierw, że podstawiając $w = T(v)$ w równości $\langle u, v \rangle = \langle T(u), T(v) \rangle$, warunek zachowania iloczynów skalarnych można przepisać w postaci $\langle u, T^{-1}(w) \rangle = \langle T(u), w \rangle$. Zastępując w tym warunku izometrią liniową T przez $T^{-1} : V \rightarrow V$, dostajemy również $\langle u, T(w) \rangle = \langle T^{-1}(u), w \rangle$, a stąd

$$\langle u, (T + T^{-1})(w) \rangle = \langle u, T(w) \rangle + \langle u, T^{-1}(w) \rangle = \langle T^{-1}(u), w \rangle + \langle T(u), w \rangle = \langle (T^{-1} + T)(u), w \rangle,$$

co oznacza, że endomorfizm $T + T^{-1} : V \rightarrow V$ jest samosprzężony.

Z Twierdzenia 8.1.4 istnieje więc niezerowy wektor $v \in V$ i $\lambda \in \mathbb{R}$ takie, że $(T + T^{-1})(v) = \lambda v$. Zatem $T(T + T^{-1})(v) = \lambda T(v)$ i stąd $T^2(v) = \lambda T(v) - v$.

Niech $W = \text{lin}(v, T(v))$. Ponieważ, jak zauważyliśmy $T^2(v) \in W$, mamy $T(W) \subset W$. ■

Dowód Twierdzenia 8.2.4. Załóżmy, że teza jest prawdziwa dla izometrii liniowych przestrzeni euklidesowych wymiaru mniejszego niż $\dim V = n$ i niech $W \subset V$ będzie podprzestrzenią wymiaru $k \leq 2$ spełniającą warunki Lematu 8.2.6.

Jeśli $k = 1$, to $T|_W$ jest identycznością lub operacją mnożenia przez -1 . Jeśli $k = 2$, to $T|_W$ jest obrotem W lub symetrią ortogonalną W względem pewnej prostej, zob. Uwaga 8.2.3. Zatem w W istnieje prostokątny układ współrzędnych $\tau : W \rightarrow \mathbb{R}^k$ taki, że macierz $M_\tau^T(T|_W)$ jest macierzą obrotu lub macierzą diagonalną mającą na przekątnej $1, -1$.

Ponieważ T jest izometrią liniową i $T(W) = W$, mamy również $T(W^\perp) = W^\perp$. Ustalmy prostokątny układ współrzędnych $\tau_\perp : W^\perp \rightarrow \mathbb{R}^{n-k}$ w W^\perp dany przez założenie indukcyjne.

Dla prostokątnego układu współrzędnych $\sigma : V \rightarrow \mathbb{R}^n$ związanego z bazą ortonormalną powstałą przez dołączenie do bazy W^\perp związanej z τ_\perp , bazy W związanej z τ , macierz $M_\sigma^\sigma(T)$ jest postaci opisanej w twierdzeniu. ■

Na zakończenie tej części podamy Twierdzenie 8.2.8 charakteryzujące izometrie liniowe przy pomocy warunku słabszego niż warunek przyjęty w definicji, wyjaśniającego przy tym lepiej terminologię.

Uwaga 8.2.7 Dla izomorfizmu $T : V \rightarrow V$ liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ warunek zachowywania iloczynu skalarnego jest równoważny (formalnie słabszemu) warunkowi zachowywania normy

$$\|v\| = \|T(v)\| \quad \text{dla } v \in V,$$

Istotnie, z $\langle u - v, u - v \rangle = \langle u, u \rangle - 2\langle u, v \rangle + \langle v, v \rangle$ dla $u, v \in V$ mamy

$$(*) \quad \langle u, v \rangle = \frac{1}{2}(\|u\|^2 + \|v\|^2 - \|u - v\|^2),$$

więc iloczyn jest wyznaczony przez normę, a stąd zachowanie normy implikuje zachowanie iloczynu. □

Modyfikując założenie o zachowywaniu normy w Uwadze 8.2.7, możemy zrezygnować z założenia liniowości przekształcenia.

Twierdzenie 8.2.8 *Jeśli $\bar{f} : V \rightarrow V$ jest przekształceniem liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ takim, że*

$$(**) \quad \bar{f}(\mathbf{0}) = \mathbf{0} \quad \text{i} \quad \|u - v\| = \|\bar{f}(u) - \bar{f}(v)\| \quad \text{dla } u, v \in V,$$

to \bar{f} jest izometrią liniową.

Dowód. Niech $\sigma : V \rightarrow \mathbb{R}^n$ będzie prostokątnym układem współrzędnych związanym z ortonormalną bazą (v_1, \dots, v_n) w V . Z (** i (*) wynika, że \bar{f} zachowuje iloczyn skalarny, więc $(\bar{f}(v_1), \dots, \bar{f}(v_n))$ też jest bazą ortonormalną V . Oznaczmy przez τ prostokątny układ współrzędnych związany z tą bazą.

Z Uwagi 7.5.2 (a) mamy $\sigma(u) = [\langle u, v_1 \rangle, \dots, \langle u, v_n \rangle]^T = [\langle \bar{f}(u), \bar{f}(v_1) \rangle, \dots, \langle \bar{f}(u), \bar{f}(v_n) \rangle]^T = \tau(\bar{f}(u))$ dla $u \in V$, więc $\bar{f} = \tau^{-1} \circ \sigma$ jest izometrią liniową jako złożenie dwóch izometrii liniowych. ■

Udowodnimy jeszcze pewne wzmocnienie Twierdzenia 8.2.8.

Twierdzenie 8.2.9* *Niech $A \subset V$ będzie podzbiorem liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ zawierającym $\mathbf{0}$. Jeśli funkcja $\bar{f} : A \rightarrow V$ spełnia warunek (**) dla $u, v \in A$, to istnieje izometria liniowa $T : V \rightarrow V$ taka, że $T|_A = \bar{f}$*

Dowód. Niech (v_1, \dots, v_m) będzie liniowo niezależnym układem wektorów z A takim, że A jest zawarte w $U = \text{lin}(v_1, \dots, v_m)$. Podobnie jak w dowodzie 8.2.8,

$$\langle u, v \rangle = \langle \bar{f}(u), \bar{f}(v) \rangle \quad \text{dla } u, v \in A.$$

W szczególności, równe są macierze Grama $G(v_1, \dots, v_m) = G(\bar{f}(v_1), \dots, \bar{f}(v_m))$ i z Uwagi 7.3.4 wynika, że układ $(\bar{f}(v_1), \dots, \bar{f}(v_m))$ jest bazą przestrzeni $W = \text{lin}(\bar{f}(v_1), \dots, \bar{f}(v_m))$. Mamy też $\bar{f}(A) \subset W$, bo gdyby $\bar{f}(u) \notin W$ dla pewnego $u \in A$, to sprzecznie z Uwagą 7.3.4, układ liniowo zależny (v_1, \dots, v_m, u) przechodziłby na układ liniowo niezależny $(\bar{f}(v_1), \dots, \bar{f}(v_m), \bar{f}(u))$.

Przekształcenie liniowe $S : U \rightarrow W$ takie, że $S(v_j) = \bar{f}(v_j)$ jest izometrią liniową, bo równe są macierze Grama $G(v_1, \dots, v_m) = G(S(v_1), \dots, S(v_m))$, zob. Uwaga 7.3.2.

Dla $u \in A$,

$$\langle S(u), S(v_j) \rangle = \langle u, v_j \rangle = \langle \bar{f}(u), \bar{f}(v_j) \rangle = \langle \bar{f}(u), S(v_j) \rangle \quad \text{dla } j \leq m,$$

więc $\langle S(u) - \bar{f}(u), w \rangle = 0$ dla $w \in W$, a stąd $\|S(u) - \bar{f}(u)\| = 0$, czyli $S|_A = \bar{f}$.

Dla zakończenia dowodu wystarczy teraz przedłużyć S do izometrii liniowej $T : V \rightarrow V$ przy pomocy dowolnej izometrii liniowej przeprowadzającej U^\perp na W^\perp . ■

8.3 *Sprzężenie endomorfizmu przestrzeni euklidesowej.

W tej części opiszemy kanoniczny izomorfizm przestrzeni euklidesowej z przestrzenią sprzężoną i wyjaśnimy związek operacji sprzężenia endomorfizmu opisanej w Definicji 4.7.4 z pojęciem endomorfizmu samosprzężonego przestrzeni euklidesowej.

Niech $(V, \langle \cdot, \cdot \rangle)$ będzie liniową przestrzenią euklidesową wymiaru n . Każdemu wektorowi $v \in V$ przyporządkujemy funkcjonal $f_v \in V^*$ określony wzorem

$$(1) \quad f_v(u) = \langle u, v \rangle.$$

Przekształcenie $v \rightarrow f_v$ jest liniowe i jeśli $v \neq 0$, to $f_v \neq 0$ (w V^*), bo $f_v(v) = \|v\|^2 \neq 0$, a więc to przyporządkowanie jest izomorfizmem V na V^* . Ten kanoniczny izomorfizm pozwala utożsamiać V z V^* , przy ustalonym iloczynie skalarnym w V .

Zauważmy, że jeśli (v_1, \dots, v_n) jest bazą ortonormalną w V , to układ $(f_{v_1}, \dots, f_{v_n})$ jest bazą dualną do niej w V^* , bo $f_{v_i}(v_j) = \langle v_i, v_j \rangle$.

Niech $T : V \rightarrow V$ będzie endomorfizmem i niech $T^* : V^* \rightarrow V^*$ będzie endomorfizmem sprzężonym określonym w 4.7.4. Utożsamienie V^* z V pozwala interpretować T^* jako endomorfizm $T^* : V \rightarrow V$ (oznaczany tak samo jak endomorfizm na V^*) taki, że dla $v \in V$ mamy (w V^*)

$$f_{T^*(v)} = T^*(f_v),$$

gdzie po prawej stronie równości występuje endomorfizm $T^* : V^* \rightarrow V^*$, a po lewej jego odpowiednik $T^* : V \rightarrow V$. Z (1) obliczamy wartość funkcjonału $f_{T^*(v)}$ na wektorze $u \in V$ jako $f_{T^*(v)}(u) = \langle u, T^*(v) \rangle$, a z 4.7.4, prawa strona równości ma na u wartość $T^*(f_v)(u) = f_v(T(u)) = \langle T(u), v \rangle$.

Wykorzystując izomorfizm kanoniczny otrzymaliśmy więc endomorfizm $T^* : V \rightarrow V$ taki, że

$$(2) \quad \langle u, T^*(v) \rangle = \langle T(u), v \rangle.$$

Endomorfizm $T^* : V \rightarrow V$ opisany formułą (2) nazywamy *endomorfizmem sprzężonym do T ze względu na iloczyn skalarny $\langle \cdot, \cdot \rangle$* .

W szczególności, samosprzężoność T oznacza, że $T^* = T$, a warunek $T^* = T^{-1}$ określa izometrie liniowe.

9 Przestrzenie afiniczne

W tym rozdziale skupimy uwagę na geometrii – dokładniej, geometrii analitycznej. Zbliżając nasz język do języka geometrii, wprowadzimy pojęcie przestrzeni afinicznej, obiektu, w którym wyróżnia się przestrzeń punktów E i przestrzeń wektorów swobodnych \vec{E} nad E , działających na E jako bijekcje – przesunięcia, przy czym dodawaniu wektorów w \vec{E} odpowiada składanie przesunięć w E .

Zacniemy od omówienia struktury afinicznej przestrzeni \mathbb{K}^n .

9.1 Struktura afiniczna przestrzeni współrzędnych.

Struktura afiniczna przestrzeni \mathbb{K}^n związana jest z warstwami – przesunięciami podprzestrzeni liniowych tej przestrzeni oraz przekształceniami afinicznymi – złożeniami endomorfizmów \mathbb{K}^n z przesunięciami.

O ile wektor zerowy w \mathbb{K}^n pełni wyróżnioną rolę względem klasy przekształceń liniowych, \mathbb{K}^n jest przestrzenią jednorodną ze względu na przekształcenia afiniczne – klasa takich przekształceń nie wyróżnia żadnego elementu \mathbb{K}^n .

Definicja 9.1.1 *Warstwą podprzestrzeni $W \subset \mathbb{K}^n$ przechodzącą przez wektor $X_0 \in \mathbb{K}^n$ nazywamy zbiór $X_0 + W = \{X_0 + Z : Z \in W\} \subset \mathbb{K}^n$.*

Jeśli $X_0 \notin W$, to warstwa $X_0 + W$ nie jest podprzestrzenią liniową \mathbb{K}^n , bo $\mathbf{0} \notin X_0 + W$.

W Twierdzeniu 3.5.4 pokazaliśmy, że zbiór rozwiązań niesprzecznego układu równań $AX = B$, gdzie $A \in \mathbb{K}_n^m$, $B \in \mathbb{K}^m$, jest warstwą $X_* + N(A)$ podprzestrzeni $N(A)$ rozwiązań jednorodnego układu $AX = \mathbf{0}$ przechodzącą przez (dowolne) rozwiązanie X_* układu $AX = B$.

Zauważmy też, że dla macierzy $A \in \mathbb{K}_n^m$ warstwa $X_0 + N(A) \in \mathbb{K}^n$ jest zbiorem rozwiązań układu równań $AX = B$, gdzie $B = AX_0$. Z Twierdzenia 3.5.6 wynika więc, że każda warstwa w \mathbb{K}^n jest zbiorem rozwiązań pewnego niesprzecznego układu równań liniowych.

Zbiór R rozwiązań niesprzecznego układu równań liniowych $AX = B$ wyznacza zbiór $\{Y - X : X, Y \in R\}$ rozwiązań układu jednorodnego, więc każda warstwa R jednoznacznie wyznacza podprzestrzeń W taką, że $R = X_* + W$ dla dowolnego $X_* \in R$. Wymiar W nazywamy *wymiarem warstwy R* . Warstwy wymiaru 1 nazywamy *prostymi*, a warstwy wymiaru $n - 1$, *hiperpłaszczyznami* w \mathbb{K}^n .

Definicja 9.1.2 *Mówimy, że przekształcenie $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$ jest afiniczne, jeśli $f(X) = AX + B$ dla $A \in \mathbb{K}_n^m$, $B = f(\mathbf{0}) \in \mathbb{K}^m$. Przekształcenie afiniczne f wyznacza przekształcenie liniowe $T : \mathbb{K}^n \rightarrow \mathbb{K}^m$ dane wzorem $T(X) = f(X) - f(\mathbf{0}) = AX$, które nazywamy częścią liniową f i oznaczamy przez \vec{f} . Przekształcenie afiniczne $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$ nazywamy izomorfizmem afinicznym \mathbb{K}^n , jeśli macierz A jest odwracalna.*

Uwaga 9.1.3 Niech $f(X) = AX + B$ będzie przekształceniem afinicznym \mathbb{K}^n w \mathbb{K}^m .

(a) Dla dowolnego $X \in \mathbb{K}^n$

$$(*) \quad f(X + Z) = f(X) + \vec{f}(Z) \text{ dla } Z \in \mathbb{K}^n.$$

Istotnie, $f(X + Z) = A(X + Z) + B = AX + AZ + B = f(X) + AZ$.

(b) Jeśli $m = n$ i f jest izomorfizmem afinicznym \mathbb{K}^n , to f ma przekształcenie odwrotne, które też jest izomorfizmem afinicznym \mathbb{K}^n , bo z $Y = AX + B$ mamy $f^{-1}(Y) = A^{-1}(Y - B) = A^{-1}Y + A^{-1}(-B)$.

(c) Złożenie przekształcenia afinicznego f z przekształceniem afinicznym $g : \mathbb{K}^m \rightarrow \mathbb{K}^l$ jest przekształceniem afinicznym $g \circ f : \mathbb{K}^n \rightarrow \mathbb{K}^l$, bo dla $f(X) = AX + B$ i $g(Y) = CY + D$ złożenie $g \circ f$ jest dane wzorem $g \circ f(X) = C(AX + B) + D = (CA)X + (CB + D)$. \square

Odnotujmy na koniec, że przekształcenia afiniczne $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$ przekształcają warstwy w \mathbb{K}^n na warstwy w \mathbb{K}^m , zob. (*) w Uwadze 9.1.3.

9.2 Przestrzenie afiniczne.

Przed wprowadzeniem ogólnej definicji przestrzeni afinicznej, wyjaśnimy to pojęcie na podstawowym przykładzie kartezjańskiej przestrzeni afinicznej \mathbb{K}^n .

Przykład 9.2.1 Pisząc \mathbb{K}^n wskazujemy, że elementy n -tej potęgi kartezjańskiej \mathbb{K} traktujemy jako punkty, odróżniając \mathbb{K}^n od przestrzeni liniowej \mathbb{K}^n , której elementami są wektory–kolumny.

Punkty będziemy oznaczać literami p, q, \dots , a współrzędne punktu będziemy zapisywać w nawiasach okrągłych, $p = (x_1, \dots, x_n)$. Każda para punktów $p = (x_1, \dots, x_n)$, $q = (y_1, \dots, y_n)$ w \mathbb{K}^n wyznacza wektor $\vec{pq} = [y_1 - x_1, \dots, y_n - x_n]^T \in \mathbb{K}^n$.

Wektor $Z = [z_1, \dots, z_n]^T \in \mathbb{K}^n$ można zaczepić w dowolnym punkcie $p = (x_1, \dots, x_n) \in \mathbb{K}^n$, w wyniku czego otrzymuje się punkt $p + Z = (x_1 + z_1, \dots, x_n + z_n)$ (w szczególności, $p + \vec{pq} = q$). Tak więc, każdy wektor $Z \in \mathbb{K}^n$ wyznacza bijekcję $p \rightarrow p + Z$ przestrzeni punktów \mathbb{K}^n na siebie – przesunięcie o wektor Z , przy czym $p + (X + Y) = (p + X) + Y$, a więc operacji dodawania wektorów w \mathbb{K}^n odpowiada składanie przesunięć o te wektory w \mathbb{K}^n .

Pojęcie przestrzeni afinicznej pozwala uwolnić się od współrzędnych przypisanych punktom, jak to ma miejsce w Przykładzie 9.2.1.

Definicja 9.2.2 *Przestrzenią afiniczną nad ciałem \mathbb{K} nazywamy trójkę $(E, \vec{E}, \#)$, gdzie E jest niepustym zbiorem punktów, \vec{E} jest przestrzenią liniową nad \mathbb{K} , której elementy nazywamy wektorami swobodnymi nad E , a $\# : E \times \vec{E} \rightarrow E$ jest operacją zaczepiania wektorów swobodnych w punktach taką, że przyjmując oznaczenie $\#(p, u) = p + u$, mamy*

- (0) $p + \mathbf{0} = p$ dla $p \in E$,
- (1) $(p + u) + v = p + (u + v)$ dla $u, v \in \vec{E}$,
- (2) Dla każdej pary punktów $p, q \in E$ istnieje dokładnie jeden wektor $u \in \vec{E}$ taki, że $p + u = q$ (będziemy go oznaczać symbolem \vec{pq}).

Dla uproszczenia oznaczeń, zamiast $(E, \vec{E}, \#)$ będziemy często pisać E , a elementy \vec{E} będziemy nazywać po prostu wektorami nad E .

Wymiarem $\dim E$ przestrzeni afinicznej E będziemy nazywać wymiar przestrzeni \vec{E} .

Uwaga 9.2.3 (a) Warunek (2) w definicji oznacza, że dla ustalonego $p \in E$ (punktu początkowego) przyporządkowanie wektorowi $u \in \vec{E}$ punktu $p + u \in E$ (końca wektora u zaczepionego w p) ustala wzajemnie jednoznaczność między E i \vec{E} (operacją odwrotną jest przyporządkowanie punktowi q wektora \vec{pq}). Punkt początkowy p odpowiada wtedy wektorowi zerowemu.

- (b) Warunek (1) mówi jak wektor przyporządkowany punktowi zmienia się w wyniku zastąpienia punktu początkowego p przez inny punkt q : podstawiając $q = p + \vec{pq}$ do $r = q + \vec{qr}$ dostajemy $r = (p + \vec{pq}) + \vec{qr} = p + (\vec{pq} + \vec{qr})$, a stąd

$$\vec{pr} = \vec{pq} + \vec{qr}.$$

W szczególności, dla $r = p$ mamy $\mathbf{0} = \vec{pp} = \vec{pq} + \vec{qp}$, czyli $\vec{qp} = -\vec{pq}$. □

Przykład 9.2.4 Niech V będzie przestrzenią liniową nad \mathbb{K} . Podobnie jak w Przykładzie 9.2.1, pisząc \dot{V} rozpatrujemy V jako zbiór punktów i określamy przestrzeń afiniczną $(\dot{V}, V, \#)$ definiując operację zaczepienia wektora $u \in V$ w punkcie $p \in \dot{V}$ wzorem $\#(p, u) = p + u$, gdzie plus po prawej stronie jest operacją dodawania wektorów w V .

W przestrzeni afinicznej E wyróżnimy niepuste podzbiory odpowiadające warstwom w \mathbb{K}^n zdefiniowanym w pierwszej części tego rozdziału.

Uwaga 9.2.5 Niech $W \subset \vec{E}$ będzie podprzestrzenią liniową przestrzeni \vec{E} wektorów swobodnych nad E i niech $p+W = \{p+w : w \in W\} \subset E$ będzie zbiorem końców wektorów z W zaczepionych w ustalonym punkcie $p \in E$. Wówczas dla każdego punktu $q \in p+W$ mamy $q+W = p+W$.

Istotnie, $q \in p+W$ oznacza, że $q = p+u$ dla pewnego wektora $u \in W$. Wtedy $q+W = (p+u)+W = \{(p+u)+w : w \in W\} = \{p+(u+w) : w \in W\} = p+W$. \square

Z tej uwagi wynika, że zbiór $p+W$ wyznacza podprzestrzeń W wzorem $W = \{\vec{qr} : q, r \in p+W\}$ i jest zamknięty ze względu na operację zaczepiania wektorów z W w punktach $p+W$. Zatem zbiór $p+W \subset E$ dziedziczy w naturalny sposób strukturę afiniczną z przestrzeni $(E, \vec{E}, +)$: zbiorem wektorów swobodnych nad $p+W$ jest W , a operacja zaczepiania wektorów w punktach $p+W$ jest obcięciem operacji zaczepiania wektorów $\# : E \times \vec{E} \rightarrow E$ do $(p+W) \times W$.

Definicja 9.2.6 Zbiory postaci $p+W \subset E$, gdzie $p \in E$, a $W \subset \vec{E}$ jest podprzestrzenią przestrzeni wektorów swobodnych nad E nazywamy podprzestrzeniami afinicznymi E . Przestrzenią wektorów swobodnych nad $p+W$ jest przestrzeń $W = \{\vec{qr} : q, r \in p+W\}$, a $\dim W$ jest wymiarem $p+W$. Podprzestrzenie wymiaru 1 nazywamy prostymi, a podprzestrzenie wymiaru $\dim \vec{E} - 1$, hiperpłaszczyznami.

Na zakończenie tej części podamy przykład przestrzeni afinicznej innej niż wcześniej omawiane. Warto jednak podkreślić, że jak się okaże, każdą przestrzeń afiniczną można utożsamić, z zachowaniem struktury afinicznej, z przestrzenią afiniczną odpowiedniego wymiaru, opisaną w Przykładzie 9.2.1.

Przykład 9.2.7* Niech V będzie niezerową przestrzenią liniową nad \mathbb{K} i niech $f : V \rightarrow \mathbb{K}$ będzie niezerowym funkcjonałem liniowym. Jako E przyjmijmy zbiór jednowymiarowych podprzestrzeni V , nie leżących w $\ker f = \vec{E}$. Operację $\# : E \times \vec{E} \rightarrow E$ określamy następująco: jeśli $p \in E$ jest jednowymiarową podprzestrzenią V oraz $u \in \vec{E}$, wybieramy niezerowy wektor $v \in p$ i definiujemy

$$\#(p, u) = \text{lin}(v + f(v)u).$$

Bez trudu sprawdza się, że operacja $\#$ jest określona poprawnie (prawa strona równości nie zależy od wyboru $v \in p$) i spełnia warunki (0)–(2) Definicji 9.2.2. Zatem $(E, \vec{E}, \#)$ jest przestrzenią afiniczną.

9.3 Kombinacje afiniczne.

Przeniesiemy teraz do przestrzeni afinicznych operację środka ciężkości układu punktów z wagami, odgrywającą ważną rolę w geometrii.

Uwaga 9.3.1 Niech (p_0, \dots, p_n) będzie układem punktów w przestrzeni afinicznej E nad \mathbb{K} i niech $a_0, \dots, a_n \in \mathbb{K}$ będą skalarami o sumie $\sum_{j=0}^n a_j = 1$. Wówczas, dla dowolnego punktu początkowego $p \in E$, punkt

$$(*) \quad q = p + \sum_{j=0}^n a_j \overrightarrow{pp_j}$$

spełnia warunek

$$(**) \quad \sum_{j=0}^n a_j \overrightarrow{qp_j} = \mathbf{0},$$

przy czym warunek $(**)$ wyznacza punkt q jednoznacznie, więc q w $(*)$ nie zależy od wyboru p .

Zanim przejdziemy do uzasadnienia, zauważmy, że dla $E = \mathbb{R}^n$, jeśli skalary $a_j \geq 0$ interpretować jako masy rozmieszczone w punktach $p_j \in \mathbb{R}^n$, to punkt q spełniający $(**)$ jest punktem, względem którego taki układ pozostaje w równowadze, to znaczy jest środkiem ciężkości tego układu.

Aby sprawdzić, że q określone wzorem $(*)$ spełnia $(**)$ zauważmy, że ze wzoru w Uwadze 9.2.3 (b), $\overrightarrow{qp_j} = \overrightarrow{qp} + \overrightarrow{pp_j}$, a więc $\sum_{j=0}^n a_j \overrightarrow{qp_j} = \sum_{j=0}^n a_j (\overrightarrow{qp} + \overrightarrow{pp_j}) = (\sum_{j=0}^n a_j) \overrightarrow{qp} + \sum_{j=0}^n a_j \overrightarrow{pp_j} = \overrightarrow{qp} + \overrightarrow{pq} = \mathbf{0}$.

Dla sprawdzenia, że punkt q dany przez $(*)$ jest jedynym punktem spełniającym $(**)$, założymy, że punkt $r \in E$ spełnia $(**)$. Wówczas $\mathbf{0} = \sum_{j=0}^n a_j \overrightarrow{rp_j} - \sum_{j=0}^n a_j \overrightarrow{rp_j} = \sum_{j=0}^n a_j (\overrightarrow{qp_j} - \overrightarrow{rp_j}) = \sum_{j=0}^n a_j (\overrightarrow{qp_j} + \overrightarrow{p_j r}) = (\sum_{j=0}^n a_j) \overrightarrow{qr} = \overrightarrow{qr}$, a więc $r = q$. \square

W uzasadnieniu uwagi istotne było założenie, że wagi sumują się do jedności. Mówiąc o wagach a_0, \dots, a_n rozmieszczonych w punktach p_0, \dots, p_n będziemy zawsze zakładać, że $\sum_{j=0}^n a_j = 1$.

Definicja 9.3.2 *Kombinacją afiniczną układu punktów (p_0, \dots, p_n) w przestrzeni afinicznej E nad \mathbb{K} , z wagami $a_0, \dots, a_n \in \mathbb{K}$, $\sum_{j=0}^n a_j = 1$, nazywamy punkt q dany wzorem $(*)$ (gdzie punkt $p \in E$ jest dowolny) i przyjmujemy oznaczenie $q = \sum_{j=0}^n a_j p_j$.*

Zauważmy, że podprzestrzenie afiniczne $p + W$ w E są zamknięte ze względu na kombinacje afiniczne w E , co wynika natychmiast z $(*)$.

Definicja 9.3.3 *Powłoką afiniczną układu punktów (p_0, \dots, p_n) w przestrzeni afinicznej E nad \mathbb{K} nazywamy zbiór $\text{af}(p_0, \dots, p_n)$ wszystkich kombinacji afinicznych tego układu.*

Twierdzenie 9.3.4 *Powłoka afiniczna układu punktów (p_0, \dots, p_n) w przestrzeni afinicznej E jest podprzestrzenią afiniczną E postaci*

$$\text{af}(p_0, \dots, p_n) = p_0 + \text{lin}(\overrightarrow{p_0 p_1}, \dots, \overrightarrow{p_0 p_n}).$$

Co więcej, $\text{af}(p_0, \dots, p_n)$ jest najmniejszą podprzestrzenią afiniczną E zawierającą wszystkie punkty p_j .

Dowód. Ponieważ $\overrightarrow{p_0 p_0} = \mathbf{0}$, z wzoru $(*)$ dla $p = p_0$ dostajemy

$$\begin{aligned} \text{af}(p_0, \dots, p_n) &= \left\{ p_0 + \sum_{j=0}^n a_j \overrightarrow{p_0 p_j} : \sum_{j=0}^n a_j = 1 \right\} = \left\{ p_0 + (a_0 \overrightarrow{p_0 p_0} + \sum_{j=1}^n a_j \overrightarrow{p_0 p_j}) : \sum_{j=0}^n a_j = 1 \right\} = \\ &= \left\{ p_0 + \sum_{j=1}^n a_j \overrightarrow{p_0 p_j} : a_1, \dots, a_n \in \mathbb{K} \right\} = p_0 + \text{lin}(\overrightarrow{p_0 p_1}, \dots, \overrightarrow{p_0 p_n}). \end{aligned}$$

W szczególności, $p_0 \in \text{af}(p_0, \dots, p_n)$, a ponieważ powłoka afiniczna układu punktów nie zależy od ich kolejności, mamy stąd $p_j \in \text{af}(p_0, \dots, p_n)$ dla $j \leq n$. Podprzestrzenie afiniczne w E są zamknięte ze względu na kombinacje afiniczne, więc $\text{af}(p_0, \dots, p_n)$ jest najmniejszą podprzestrzenią afiniczną E zawierającą wszystkie p_j . ■

Definicja 9.3.5 *Mówimy, że układ punktów (p_0, \dots, p_n) w przestrzeni afinicznej E jest afinicznie niezależny jeśli podprzestrzeń $\text{af}(p_0, \dots, p_n)$ ma wymiar n .*

Uwaga 9.3.6 Z Twierdzenia 9.3.4 wynika, że afiniczna niezależność układu punktów (p_0, \dots, p_n) jest równoważna liniowej niezależności układu wektorów $(\overrightarrow{p_0 p_1}, \dots, \overrightarrow{p_0 p_n})$, przy czym p_0 można zastąpić dowolnym innym punktem układu, bo powłoka afiniczna nie zależy od kolejności punktów układu. □

9.4 Układ bazowy, baza punktowa.

W Uwadze 9.2.3 (a) odnotowaliśmy, że wybranie punktu początkowego $p \in E$ ustala wzajemnie jednoznacznie odpowiedniość punktów E z wektorami z \vec{E} , $q \longrightarrow \overrightarrow{p q}$. Dodatkowy wybór bazy w \vec{E} pozwala więc przyporządkować każdemu punktowi *współrzędne*.

Definicja 9.4.1 *Układem bazowym w przestrzeni afinicznej E nad \mathbb{K} nazywamy układ $(p; v_1, \dots, v_n)$ taki, że $p \in E$, a (v_1, \dots, v_n) jest bazą \vec{E} .*

Definicja 9.4.2 *Bazą punktową w przestrzeni afinicznej E nad \mathbb{K} nazywamy układ punktów (p_0, \dots, p_n) taki, że $(p_0; \overrightarrow{p_0 p_1}, \dots, \overrightarrow{p_0 p_n})$ jest układem bazowym w E .*

Jeśli (p_0, p_1, \dots, p_n) jest bazą punktową E , to punkt $q \in E$ można jednoznacznie zapisać w układzie bazowym $(p_0; \overrightarrow{p_0 p_1}, \dots, \overrightarrow{p_0 p_n})$ w postaci $q = p_0 + \sum_{j=1}^n a_j \overrightarrow{p_0 p_j}$. Wtedy $q = \sum_{j=0}^n a_j p_j$ jest kombinacją afiniczną (p_0, p_1, \dots, p_n) , w której waga $a_0 = 1 - \sum_{j=1}^n a_j$ jest dobrana tak, by suma wag była jednością. Wagi a_0, \dots, a_n nazywamy *współrzędnymi barycentrycznymi* punktu q .

Zbiór współrzędnych barycentrycznych, interpretowanych jako elementy \mathbb{K}^{n+1} , tworzy w tej przestrzeni hiperplaszczynę opisaną równaniem $\sum_{j=0}^n x_j = 1$.

9.5 Przekształcenia afiniczne

Rozpatrując przekształcenia między przestrzeniami punktów przestrzeni afinicznych będziemy zawsze zakładać, że dziedzina i przeciwdziedzina są przestrzeniami nad tym samym ciałem \mathbb{K} .

Przekształcenia afiniczne przestrzeni afinicznych, to przekształcenia, które w ustalonych układach bazowych są opisane przez przekształcenia afiniczne z \mathbb{K}^n w \mathbb{K}^m , postaci $Y = AX + B$, określone na początku tego rozdziału. Wygodniej jednak będzie przyjąć jako definicję warunek niezależny od wyboru układów bazowych.

Definicja 9.5.1 Niech E, F będą przestrzeniami afinicznymi nad \mathbb{K} . Funkcję $f : E \rightarrow F$ nazywamy *przekształceniem afinicznym*, jeśli dla pewnego $p \in E$ istnieje przekształcenie liniowe $\vec{f} : \vec{E} \rightarrow \vec{F}$, zwane *częścią liniową* f , spełniające warunek

$$(*) \quad f(p + v) = f(p) + \vec{f}(v) \text{ dla } v \in \vec{E}.$$

Uwaga 9.5.2 (a) W warunku (*) można, bez zmiany \vec{f} , zastąpić punkt p dowolnym punktem $q \in E$, bo dla $q = p + u$ mamy $f(q + v) = f(p + u + v) = f(p) + \vec{f}(u + v) = f(p) + \vec{f}(u) + \vec{f}(v) = f(q) + \vec{f}(v)$.

(b) f jest wyznaczone przez podanie obrazu $f(p)$ jakiegokolwiek punktu $p \in E$ i części liniowej \vec{f} .

(c) Złożenie przekształceń afinicznych $g \circ f$ jest przekształceniem afinicznym i $\overrightarrow{g \circ f} = \vec{g} \circ \vec{f}$.

Istotnie, $g(f(p + v)) = g(f(p) + \vec{f}(v)) = g(f(p)) + \vec{g}(\vec{f}(v))$. □

Niech $f : E \rightarrow E$ będzie przekształceniem afinicznym. Jeśli $\vec{f} = \mathbf{0}$, to f jest przekształceniem stałym.

Jeśli $\vec{f} = \text{id}$, to f jest *przesunięciem* $f(p + v) = f(p) + v = (p + \overrightarrow{pf(p)}) + v = (p + v) + \overrightarrow{pf(p)}$.

Jeśli $\vec{f} = c \cdot \text{id}$, $c \neq 1$, to f nazywamy *jednokładnością o skali c* (środkiem tej jednokładności jest punkt $q = p + (1 - c)^{-1} \overrightarrow{pf(p)}$ spełniający warunek $f(q) = q$).

Przekształcenie afiniczne $f : E \rightarrow E$ nazywamy *rzutem na podprzestrzeń $p + W$ równoległą do U* (symetrią względem $p + W$ równoległą do U), jeśli $f(p) = p$ i część liniowa f jest rzutem \vec{E} na W równoległą do U (symetrią względem W równoległą do U).

Definicja 9.5.3 Przekształcenie afiniczne $f : E \rightarrow F$ nazywamy *izomorfizmem afinicznym* jeśli część liniowa $\vec{f} : \vec{E} \rightarrow \vec{F}$ jest izomorfizmem liniowym.

Uwaga 9.5.4 Izomorfizm afiniczny $f : E \rightarrow F$ ma funkcję odwrotną $f^{-1} : F \rightarrow E$, która jest izomorfizmem afinicznym zadanym warunkami $f^{-1}(f(p)) = p$ i $\overrightarrow{(f^{-1})} = (\vec{f})^{-1} : \vec{F} \rightarrow \vec{E}$. □

Mówimy, że przestrzenie afiniczne E, F są *izomorficzne* jeśli istnieje izomorfizm afiniczny E na F , co zgodnie z Wnioskiem 4.2.8 jest równoważne równości wymiarów $\dim E = \dim F$.

Na koniec tej części odnotujemy ważną własność przekształceń afinicznych (w istocie, charakteryzującą tę klasę przekształceń).

Uwaga 9.5.5 Przekształcenia afiniczne zachowują kombinacje afiniczne.

Istotnie, dla przekształcenia afinicznego $f : E \rightarrow F$ warunek (*) oznacza, że $\vec{f}(\overrightarrow{pq}) = \overrightarrow{f(p)f(q)}$, $p, q \in E$. Stąd dla kombinacji afinicznej $\sum_{j=0}^n a_j p_j$ w E , gdzie $\sum_{j=0}^n a_j = 1$, $f(\sum_{j=0}^n a_j p_j) = f(p + \sum_{j=0}^n a_j \overrightarrow{pp_j}) = f(p) + \vec{f}(\sum_{j=0}^n a_j \overrightarrow{pp_j}) = f(p) + \sum_{j=0}^n a_j \vec{f}(\overrightarrow{pp_j}) = f(p) + \sum_{j=0}^n a_j \overrightarrow{f(p)f(p_j)} = \sum_{j=0}^n a_j f(p_j)$. □

9.6 Afiniczne układy współrzędnych.

Przekształcenie n -wymiarowej przestrzeni afinicznej E nad \mathbb{K} przyporządkowujące punktom E ich współrzędne w układzie bazowym przestrzeni E określa w E układ współrzędnych.

Definicja 9.6.1 *Afinicznym układem współrzędnych w przestrzeni afinicznej E nad \mathbb{K} (związany z układem bazowym $(p; v_1, \dots, v_n)$ w E) nazywamy przekształcenie $\sigma_p : E \rightarrow \mathbb{K}^n$ dane wzorem*

$$\sigma_p(p + v) = \sigma(v),$$

gdzie $p \in E$, a $\sigma : \vec{E} \rightarrow \mathbb{K}^n$ jest układem współrzędnych w \vec{E} (przyporządkowującym wektorowi $v \in \vec{E}$ jego współrzędne w bazie (v_1, \dots, v_n) przestrzeni \vec{E}).

Definicja afinicznego układu współrzędnych jest formalnie bardzo podobna do definicji izomorfizmu afinicznego – σ_p przeprowadza p na $\mathbf{0}$, a “część liniowa” σ jest izomorfizmem liniowym. W istocie, σ_p ustala odpowiedniość między “strukturami afinicznymi” przestrzeni punktów E i przestrzeni wektorów \mathbb{K}^n .

Uwaga 9.6.2 (a) Jeśli $\sigma_p : E \rightarrow \mathbb{K}^n$ jest afinicznym układem współrzędnych w przestrzeni afinicznej E , to dla dowolnego $q = p + u \in E$ mamy

$$(*) \quad \sigma_p(q + v) = \sigma_p(q) + \sigma(v) \text{ dla } v \in \vec{E}.$$

Istotnie, $\sigma_p(q + v) = \sigma_p(p + u + v) = \sigma(u + v) = \sigma(u) + \sigma(v) = \sigma_p(q) + \sigma(v)$.

(b) Jeśli $\sigma_p, \tau_q : E \rightarrow \mathbb{K}^n$ są afinicznymi układami współrzędnych w E , to złożenie $\sigma_p \circ (\tau_q)^{-1} : \mathbb{K}^n \rightarrow \mathbb{K}^n$ zmienia współrzędne $X = \tau_q(r)$ punktu $r \in E$ w na współrzędne $\sigma_p(r)$, bo $\sigma_p \circ (\tau_q)^{-1}(X) = \sigma_p \circ (\tau_q)^{-1}(\tau_q(r)) = \sigma_p(r)$. \square

Fakt, że jak wskazaliśmy, afiniczne układy współrzędnych w E pozwalają utożsamiać struktury afiniczne w E i w \mathbb{K}^n , prowadzi do następującego twierdzenia (implikacja (i) \Rightarrow (ii) w tezie mówi, że zmiana współrzędnych opisana w Uwadze 9.6.2 (c) jest izomorfizmem afinicznym \mathbb{K}^n).

Twierdzenie 9.6.3 *Niech $\sigma_p : E \rightarrow \mathbb{K}^n$ będzie afinicznym układem współrzędnych w przestrzeni afinicznej E . Dla odwracalnego przekształcenia $\phi : E \rightarrow \mathbb{K}^n$ następujące warunki są równoważne:*

(i) ϕ jest afinicznym układem współrzędnych w E ;

(ii) złożenie $\sigma_p \circ \phi^{-1} : \mathbb{K}^n \rightarrow \mathbb{K}^n$ jest afinicznym izomorfizmem \mathbb{K}^n ;

(iii) złożenie $\phi^{-1} \circ \sigma_p : E \rightarrow E$ jest afinicznym izomorfizmem E .

Dowód. Udowodnimy równoważność (i) \Leftrightarrow (ii). Dowód równoważności (i) \Leftrightarrow (iii) jest analogiczny.

(i) \Rightarrow (ii): jeśli $\phi = \tau_q$, to $\sigma_p \circ (\tau_q)^{-1}(X) = \sigma_p(q + \tau^{-1}(X)) = \sigma_p(q) + \sigma(\tau^{-1}(X)) = AX + B$,
gdzie $A = M(\sigma \circ \tau^{-1})$ i $B = \sigma_p(q)$.

(ii) \Rightarrow (i): jeśli $g = \sigma_p \circ \phi^{-1}$ jest izomorfizmem afinicznym \mathbb{K}^n i $q = (\sigma_p)^{-1}(g(\mathbf{0}))$, to z $\phi = g^{-1} \circ \sigma_p$ mamy

$$\phi(q + v) = g^{-1} \circ \sigma_p(q + v) = g^{-1}(\sigma_p(q) + \sigma(v)) = g^{-1}(g(\mathbf{0})) + \overrightarrow{g^{-1}}(\sigma(v)) = (\vec{g})^{-1} \circ \sigma(v),$$

więc $\phi = \tau_q$, gdzie $\tau = (\vec{g})^{-1} \circ \sigma$ jest układem współrzędnych w \vec{E} \blacksquare

Podobnie jak w Uwadze 9.5.5, sprawdza się, że każdy afiniczny układ współrzędnych w E przeprowadza kombinacje afiniczne punktów E na kombinacje liniowe ich obrazów w \mathbb{K}^n .

10 Przestrzenie Euklidesowe

W geometrii, przestrzeń euklidesowa, to zbiór punktów E z zadaną odległością $d(p, q)$ między punktami, która, po wprowadzeniu odpowiedniego układu współrzędnych w E , jest opisana formułą $d(p, q) = \sqrt{\sum_{j=1}^n (y_j - x_j)^2}$, $p = (x_1, \dots, x_n)$, $q = (y_1, \dots, y_n)$.

Pojęcie przestrzeni euklidesowej wprowadzimy w przyjętym przez nas formalizmie przestrzeni afinicznych nad ciałem liczb rzeczywistych, ustalając w przestrzeni wektorów swobodnych iloczyn skalarny.

10.1 Afiniczne przestrzenie euklidesowe.

Definicja 10.1.1 *Przestrzeń afiniczną $(E, (\vec{E}, \langle \cdot, \cdot \rangle), \oplus)$ nad \mathbb{R} , gdzie w skończeniu wymiarowej przestrzeni wektorów swobodnych \vec{E} nad E ustalony jest iloczyn skalarny $\langle \cdot, \cdot \rangle$, będziemy nazywali afiniczną przestrzenią euklidesową.*

W dalszym ciągu będziemy mówili po prostu o przestrzeni euklidesowej E , jeśli jasne jest jaki iloczyn skalarny został ustalony w \vec{E} .

Definicja 10.1.2 *Odległość między punktami przestrzeni euklidesowej określamy wzorem $d(p, q) = \|\vec{pq}\|$, gdzie $\|\vec{pq}\| = \sqrt{\langle \vec{pq}, \vec{pq} \rangle}$ jest normą wektora swobodnego \vec{pq} w $(\vec{E}, \langle \cdot, \cdot \rangle)$.*

Przykład 10.1.3 W przestrzeni euklidesowej \mathbb{R}^n , gdzie w przestrzeni wektorów swobodnych \mathbb{R}^n ustalony jest iloczyn skalarny $\langle X, Y \rangle = X^T Y$, odległość między punktami $p = (x_1, \dots, x_n)$, $q = (y_1, \dots, y_n)$ jest określona wzorem $d(p, q) = \sqrt{\sum_{j=1}^n (y_j - x_j)^2}$.

Definicja 10.1.4 *Rzutem ortogonalnym przestrzeni euklidesowej E na podprzestrzeń afiniczną $p+W \subset E$ nazywamy przekształcenie afiniczne $f : E \rightarrow E$ dane wzorem $f(p+v) = p + P(v)$, gdzie $P : \vec{E} \rightarrow W$ jest rzutem ortogonalnym \vec{E} na W . Symetrią ortogonalną E względem $p+W$ nazywamy przekształcenie afiniczne $g : E \rightarrow E$ określone wzorem $g(p+v) = p + \vec{g}(v)$, gdzie \vec{g} jest symetrią ortogonalną \vec{E} względem W .*

Niech $f : E \rightarrow E$ będzie rzutem ortogonalnym przestrzeni euklidesowej E na podprzestrzeń $p+W$. Zgodnie z Uwagą 7.2.8, rzut ortogonalny $r = f(q)$ punktu q jest jedynym punktem $p+W$ spełniającym warunek $d(r, q) = \inf \{d(\mathbf{x}, q) : \mathbf{x} \in p+W\}$.

Minimalna odległość $d(r, q)$ punktu q od $p+W$ jest normą rzutu ortogonalnego wektora \vec{pq} na dopełnienie ortogonalne W^\perp podprzestrzeni W w \vec{E} , nazywamy ją *odległością punktu q od podprzestrzeni $p+W$* .

Jeśli U i W są podprzestrzeniami \vec{E} , $p, q \in E$, to kres dolny odległości między punktami podprzestrzeni $q+U$ i $p+W$, $\inf \{d(p+w, q+u) : w \in W, u \in U\} = \inf \{d(p+w-u, q) : w \in W, u \in U\}$ jest odległością punktu q od podprzestrzeni $p+(W+U)$ równą normie rzutu wektora \vec{pq} na podprzestrzeń $(W+U)^\perp$ – jest to *odległość między podprzestrzeniami $q+U$ i $p+W$ przestrzeni E* .

Definicja 10.1.5 *Afiniczny układ współrzędnych $\sigma_p : E \rightarrow \mathbb{R}^n$ w przestrzeni euklidesowej E nazywamy prostokątnym, jeśli $\sigma : \vec{E} \rightarrow \mathbb{R}^n$ jest prostokątnym układem współrzędnych w \vec{E} .*

10.2 Izometrie przestrzeni euklidesowych.

Izometrie przestrzeni euklidesowej to przekształcenia $f : E \rightarrow E$ zachowujące odległość: dla $p, q \in E$,

$$(*) \quad d(f(p), f(q)) = d(p, q)$$

Każdy izomorfizm afiniczny $f : E \rightarrow E$, którego część liniowa jest izometrią liniową, jest oczywiście izometrią. Pokażemy, że wszystkie izometrie E są takiej postaci.

Twierdzenie 10.2.1 *Izometrie przestrzeni euklidesowej E są przekształceniami afinicznymi, których część liniowa jest izometrią liniową przestrzeni \vec{E} wektorów swobodnych nad E .*

Dowód. Ustalmy punkt $p \in E$ i połóżmy $\vec{f}(\overrightarrow{pq}) = \overrightarrow{f(p)f(q)}$. Wtedy $\vec{f} : \vec{E} \rightarrow \vec{E}$ jest dobrze określone i $f(q) = f(p) + \vec{f}(\overrightarrow{pq})$ dla $q \in E$, więc nasza teza mówi, że \vec{f} jest izometrią liniową. Wyprowadzimy ją z Twierdzenia 8.2.8 dowodząc, że \vec{f} spełnia warunek (**) podany w założeniach tego twierdzenia.

Warunek $\vec{f}(\mathbf{0}) = \mathbf{0}$ jest oczywiście spełniony. Rozważmy wektory $\overrightarrow{pq}, \overrightarrow{pr} \in \vec{E}$. Z (*) dla $q, r \in E$ mamy $\|\overrightarrow{qr}\| = \|\overrightarrow{f(q)f(r)}\|$, a z drugiej strony, $\overrightarrow{qr} = \overrightarrow{pr} - \overrightarrow{pq}$ i, analogicznie, $\overrightarrow{f(q)f(r)} = \overrightarrow{f(p)f(r)} - \overrightarrow{f(p)f(q)}$, czyli $\|\overrightarrow{pr} - \overrightarrow{pq}\| = \|\overrightarrow{f(p)f(r)} - \overrightarrow{f(p)f(q)}\|$, więc \vec{f} spełnia warunek (**) w 8.2.8. ■

Ważnym wnioskiem z Twierdzenia 10.2.1 jest fakt, że odległość w przestrzeni euklidesowej E pozwala scharakteryzować przesunięcia tej przestrzeni (a więc pozwala zdefiniować przestrzeń wektorów swobodnych nad E).

Wniosek 10.2.2 *Przekształcenie $f : E \rightarrow E$ przestrzeni euklidesowej E jest przesunięciem o pewien wektor wtedy i tylko wtedy, gdy f jest izometrią i $\sup \{d(q, f(q)) : q \in E\} < \infty$.*

Dowód. Oczywiście, przesunięcie $q \rightarrow q + v$ jest izometrią E i $d(q, q + v) = \|v\|$.

Na odwrót, niech f będzie izometrią taką, że $\sup \{d(q, f(q)) : q \in E\} < \infty$. Zgodnie z 10.2.1, f jest przekształceniem afinicznym. Ustalmy $p \in E$ i rozważmy przekształcenie afiniczne $g : E \rightarrow E$ będące złożeniem f z przesunięciem o wektor $\overrightarrow{f(p)p}$. Wtedy $g(p) = p$ i wystarczy pokazać, że $g = \text{id}_E$, bo stąd natychmiast wynika, że f jest przesunięciem o wektor $\overrightarrow{pf(p)}$.

Zauważmy, że $\sup \{d(q, g(q)) : q \in E\} < \infty$. Gdyby $g(p + v) \neq p + v$ dla pewnego wektora $v \in \vec{E}$, to z $g(p + v) = p + \vec{g}(v)$ mielibyśmy $\vec{g}(v) - v \neq \mathbf{0}$, czyli $\|\vec{g}(v) - v\| \neq 0$. Wtedy $\lim_{t \rightarrow \infty} \|\vec{g}(tv) - tv\| = \lim_{t \rightarrow \infty} t \|\vec{g}(v) - v\| = \infty$, czyli $\lim_{t \rightarrow \infty} d(p + tv, g(p + tv)) = \infty$, sprzecznie z założeniem. ■

Twierdzenie 8.2.9 pozwala wzmocnić Twierdzenie 10.2.1 w następujący sposób.

Powiemy, że przekształcenie $f : H \rightarrow E$ niepustego podzbioru $H \subset E$ przestrzeni euklidesowej E w E jest *izometrią*, jeśli f zachowuje odległości punktów z H , czyli spełnia warunek (*) dla $p, q \in H$.

Twierdzenie 10.2.3* *Dla każdej izometrii $f : H \rightarrow E$ podzbioru $H \subset E$ przestrzeni euklidesowej E istnieje izometria liniowa $T : \vec{E} \rightarrow \vec{E}$ taka, że $f(q) = f(p) + T(\overrightarrow{pq})$ dla $p, q \in H$.*

Dowód. Ustalmy punkt $p \in H$ i połóżmy $A = \{\overrightarrow{pq} : q \in H\}$. Tak jak w dowodzie Twierdzenia 10.2.1, sprawdza się, że funkcja $\vec{f} : A \rightarrow A$ dana przez $\vec{f}(\overrightarrow{pq}) = \overrightarrow{f(p)f(q)}$ dla $\overrightarrow{pq} \in A$, spełnia założenia Twierdzenia 8.2.9 i wyprowadza się stąd tezę twierdzenia. ■

11 Formy kwadratowe

W tym rozdziale rozpatrujemy wyłącznie przestrzenie liniowe nad ciałami \mathbb{K} charakterystyki różnej od 2 (tzn. $1 + 1 \neq 0$ w \mathbb{K}).

11.1 Formy kwadratowe.

Macierz symetryczna $A = A^T \in \mathbb{K}_n^n$ wyznacza formę kwadratową $Q : \mathbb{K}^n \rightarrow \mathbb{K}$ wzorem

$$(*) \quad Q(X) = X^T A X, \quad A = A^T \in \mathbb{K}_n^n.$$

Z formą Q związany jest funkcjonal dwuliniowy $h : \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}$ – forma biegunowa Q ,

$$(**) \quad h(X, Y) = X^T A Y, \quad Q(X) = h(X, X).$$

Złożenie formy kwadratowej Q wyznaczonej przez A z izomorfizmem liniowym $S : \mathbb{K}^n \rightarrow \mathbb{K}^n$ jest formą kwadratową wyznaczoną przez macierz symetryczną $B = C^T A C$, $C = M(S)$, tzn. $Q \circ S(X) = X^T B X$; w szczególności, macierze B i A są kongruentne, zob. 11.1.5.

Główny wynik tego rozdziału mówi, że izomorfizm S można zawsze dobrać tak, aby $Q \circ S(X) = \sum_j d_j x_j^2$ (x_j – współrzędne X) lub równoważnie, w języku macierzowym, że każda macierz symetryczna jest kongruentna z macierzą diagonalną.

Nasze rozważania będziemy prowadzić w ogólniejszym ujęciu. Zaczniemy od określenia symetrycznych funkcjonałów dwuliniowych na przestrzeni V i związanych z nimi form kwadratowych.

Definicja 11.1.1 *Niech V będzie przestrzenią liniową nad \mathbb{K} . Funkcjonał $h : V \times V \rightarrow \mathbb{K}$ jest symetrycznym funkcjonałem dwuliniowym jeśli $h(v, w) = h(w, v)$ i dla każdego $w \in V$ funkcjonal $v \rightarrow h(v, w)$ jest liniowy. Funkcjonał $Q : V \rightarrow \mathbb{K}$ dany wzorem $Q(v) = h(v, v)$ nazywamy formą kwadratową wyznaczoną przez h i mówimy, że h jest formą biegunową dla Q .*

W szczególności, iloczyn skalarny $\langle \cdot, \cdot \rangle$ w liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ nad \mathbb{R} jest symetrycznym funkcjonałem dwuliniowym – formą biegunową dla formy kwadratowej $\|v\|^2 = \langle v, v \rangle$.

Uwaga 11.1.2 Jeśli h jest formą biegunową dla formy kwadratowej Q , to z dwuliniowości h mamy $h(v + w, v + w) = h(v, v) + 2h(v, w) + h(w, w)$, więc

$$h(v, w) = \frac{1}{2}(Q(v + w) - Q(v) - Q(w)),$$

czyli forma kwadratowa wyznacza swoją formę biegunową jednoznacznie. □

Podobnie jak dla iloczynu skalarnego w przestrzeniach euklidesowych, symetryczny funkcjonal dwuliniowy $h : V \times V \rightarrow \mathbb{K}$ pozwala związać z każdym układem (v_1, \dots, v_m) wektorów w V symetryczną macierz Grama, zob. Definicja 7.3.1,

$$G_h(v_1, \dots, v_m) = [h(v_i, v_j)]_{i,j=1}^m,$$

która wyznacza wartości funkcjonału h na powłoce liniowej tego układu, zob. Uwaga 7.3.2.

Twierdzenie 11.1.3 *Niech $h : V \times V \rightarrow \mathbb{K}$ będzie symetrycznym funkcjonałem dwuliniowym, układ (v_1, \dots, v_n) bazą w V i niech $\sigma : V \rightarrow \mathbb{K}^n$ będzie układem współrzędnych związanym z tą bazą. Wówczas funkcjonal h jest opisany wzorem*

$$h(v, w) = X^T A Y \quad \text{dla } X = \sigma(v), Y = \sigma(w),$$

gdzie $A = G_h(v_1, \dots, v_n)$ jest macierzą Grama bazy (v_1, \dots, v_n) .

Dowód. Jak w Uwadze 7.3.2, dla $v = \sum_i x_i v_i$, $w = \sum_j y_j v_j$, z dwuliniowości h mamy $h(v, w) = h(\sum_i x_i v_i, \sum_j y_j v_j) = \sum_{i,j=1}^n x_i y_j h(v_i, v_j) = \sigma(v)^T G_h(v_1, \dots, v_n) \sigma(w)$. ■

Uwaga 11.1.4 (a) Niech $h : V \times V \rightarrow \mathbb{K}$ będzie symetrycznym funkcjonałem dwuliniowym i niech bazy (v_1, \dots, v_n) , (w_1, \dots, w_n) w V będą związane z układami współrzędnych $\sigma, \tau : V \rightarrow \mathbb{K}^n$, odpowiednio. Wówczas $G_h(w_1, \dots, w_n) = C^T A C$ dla $A = G_h(v_1, \dots, v_n)$ i odwracalnej macierzy $C = M(\sigma \circ \tau^{-1})$ (bo j -ta kolumna C ma postać $\sigma(\tau^{-1}(E_j)) = \sigma(w_j)$, zob. dowód Lematu 7.3.3).

(b) Niech $Q(X) = X^T A X$ będzie formą kwadratową na \mathbb{K}^n wyznaczoną przez macierz symetryczną A i niech $S : \mathbb{K}^n \rightarrow \mathbb{K}^n$ będzie izomorfizmem. Wówczas $Q \circ S(Y) = Y^T (C^T A C) Y$ dla odwracalnej macierzy $C = M(S)$ (wystarczy we wzorze na $Q(X)$ podstawić $X = CY$ i skorzystać ze wzoru $(CY)^T = Y^T C^T$ lub przyjmując w (a) $\sigma = \text{id}_{\mathbb{K}^n}$ i $\tau^{-1} = S$). □

Uwaga 11.1.4 prowadzi do ważnej relacji równoważności w zbiorze $(n \times n)$ -macierzy.

Definicja 11.1.5 Macierze $A, B \in \mathbb{K}_n^n$ są kongruentne, jeśli istnieje macierz odwracalna $C \in \mathbb{K}_n^n$ taka, że $B = C^T A C$.

Tak więc, zgodnie z Uwagą 11.1.4, jeśli Q jest formą kwadratową opisaną wzorem (*), macierze kongruentne do A , są to dokładnie symetryczne macierze opisujące formy kwadratowe $Q \circ S$, gdzie S jest izomorfizmem \mathbb{K}^n .

Zauważmy też, że macierze kongruentne mają równe rzędy, bo mnożenie przez macierz odwracalną nie zmienia rzędu macierzy.

Definicja 11.1.6 Rzędem rank(h) symetrycznego funkcjonału dwuliniowego $h : V \times V \rightarrow \mathbb{K}$ nazywamy rząd jego macierzy Grama w dowolnej bazie V . Rzędem rank(Q) formy kwadratowej $Q : V \rightarrow \mathbb{K}$ nazywamy rząd formy biegunowej Q .

Na zakończenie podamy dwa przydatne przykłady macierzy kongruentnych z macierzą $\begin{bmatrix} 0 & c \\ c & 0 \end{bmatrix} \in \mathbb{K}_2^2$.

Przykład 11.1.7 Dla $c \in \mathbb{K}$:

$$(a) \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & c \\ c & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -2c & 0 \\ 0 & 2c \end{bmatrix}; \quad (b) \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & c \\ c & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2c & c \\ c & 0 \end{bmatrix}.$$

11.2 Algorytm Lagrange'a.

Algorytm Lagrange'a jest użyteczną procedurą, którą wykorzystamy do dowodu następującego głównego wyniku w tej części.

Twierdzenie 11.2.1 Każda macierz symetryczna $A = A^T \in \mathbb{K}_n^n$ jest kongruentna z macierzą diagonalną, tzn. istnieje macierz odwracalna $C \in \mathbb{K}_n^n$ taka, że

$$C^T A C = \begin{bmatrix} d_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & d_n \end{bmatrix},$$

przy czym wyrazy niezerowe na przekątnej poprzedzają wyrazy zerowe

Zanim przystąpimy do dowodu tego twierdzenia, podamy jego interpretację w języku form kwadratowych na \mathbb{K}^n .

Wniosek 11.2.2 Dla każdej niezerowej formy kwadratowej $Q : \mathbb{K}^n \rightarrow \mathbb{K}$ istnieje izomorfizm $S : \mathbb{K}^n \rightarrow \mathbb{K}^n$ taki, że $Q \circ S(Y) = \sum_{j=1}^r d_j y_j^2$, dla $Y = [y_1, \dots, y_n]^T$, gdzie $d_j \neq 0$, $j \leq r = \text{rank}(Q)$.

W dowodzie twierdzenia będziemy wykonywać operacje elementarne na wierszach i kolumnach A . Zaważmy, że jeśli macierz elementarna M jest macierzą operacji elementarnej \mathcal{E} na wierszach (czyli zgodnie z Uwagą 4.5.5, iloczyn MA jest macierzą otrzymaną z A przez wykonanie operacji \mathcal{E} na jej wierszach) to iloczyn $AM^T = (MA^T)^T$ powstaje z A przez wykonanie odpowiedniej operacji na jej kolumnach.

Dowód Twierdzenia 11.2.1 (algorytm Lagrange’a). Niech $A = [a_{ij}]_{i,j=1}^n \in \mathbb{K}_n^n$ będzie niezerową macierzą symetryczną.

(I) Załóżmy, że $a_{11} \neq 0$. Wówczas odejmując pierwszy wiersz pomnożony przez $\frac{a_{i1}}{a_{11}}$ od i -tego, dla $i \geq 2$, wyzerujemy wszystkie wyrazy pierwszej kolumny pod a_{11} . Jeśli M jest iloczynem odpowiednich macierzy elementarnych, to z $A = A^T$ wynika, że $(MAM^T)^T = (M^T)^T A^T M^T = MAM^T$, więc iloczyn MAM^T jest symetryczną macierzą postaci

$$MAM^T = \left[\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array} \right],$$

(operowanie pierwszą kolumną MA na kolejnych kolumnach zeruje wyrazy $a_{12} = a_{21}, \dots, a_{1n} = a_{n1}$ pierwszej wiersza na prawo od a_{11}). W szczególności, macierz B jest symetryczna.

(II) Jeśli $a_{11} = 0$ ale $a_{ii} \neq 0$ dla pewnego $i > 1$, to dla macierzy M operacji elementarnej zamieniającej pierwszy wiersz z i -tym macierz symetryczna MAM^T ma w lewym górnym rogu wyraz a_{ii} , a więc spełnia warunek w (I).

(III) Jeśli $a_{ii} = 0$ dla $i = 1, \dots, n$, to z $A \neq \mathbf{0}$ i $A = A^T$ wynika, że istnieją $i > j$ takie, że $a_{ij} \neq 0$. Wzorujemy się wtedy na kongruencji z Przykładu 11.1.7 (a) przyjmując $M = [m_{kl}]_{k,l=1}^n$, gdzie wyrazy na przekątnej M są jedynekami, $m_{ij} = 1, m_{ji} = -1$, a pozostałe wyrazy M są zerowe. Macierz symetryczna MAM^T ma na przekątnej niezerowe wyrazy $\mp 2a_{ij}$, a więc spełnia warunek w (II) (lub w (I), jeśli $j = 1$).

Tak więc, mnożąc odpowiednie macierze opisane w (I),(II),(III) otrzymamy macierz odwracalną $M_1 \in \mathbb{K}_n^n$ taką, że

$$M_1AM_1^T = \left[\begin{array}{c|ccc} d_1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array} \right], \quad A_1 = A_1^T, \quad d_1 \neq 0.$$

Jeśli macierz $A_1 \neq \mathbf{0}$, to powtarzając tę procedurę dla A_1 otrzymamy macierz odwracalną $M_2 \in \mathbb{K}_n^n$ (postaci MM_1) taką, że

$$M_2AM_2^T = \left[\begin{array}{cc|ccc} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \hline \vdots & \vdots & A_2 & \\ 0 & 0 & & \end{array} \right], \quad A_2 = A_2^T, \quad d_1, d_2 \neq 0.$$

Po $r = \text{rank } A$ takich krokach znajdziemy macierz odwracalną $M_r \in \mathbb{K}_n^n$ taką, że

$$M_rAM_r^T = \left[\begin{array}{ccc|ccc} d_1 & & 0 & \mathbf{0} \\ & \ddots & & \\ 0 & & d_r & \mathbf{0} \\ \hline \mathbf{0} & & \mathbf{0} & \mathbf{0} \end{array} \right], \quad d_1, \dots, d_r \neq 0$$

i przyjmujemy $C = M_r^T$. ■

Uwaga 11.2.3 Macierz $M_r = C^T$ występującą w dowodzie twierdzenia można otrzymać, podobnie jak przy odwracaniu macierzy, z macierzy jednostkowej I_n wykonując na jej wierszach operacje takie same jak wykonywane w trakcie redukcji operacje na wierszach macierzy A (mnożenie z lewej strony przez macierz M opisaną w (III) odpowiada zastąpieniu j -tego wiersza przez różnicę j -tego i i -tego wiersza, zaś i -tego wiersza przez sumę tych wierszy, zob. Przykład 11.1.7 (a)). \square

Twierdzenie 11.2.1 można sformułować równoważnie w następujący sposób.

Twierdzenie 11.2.4 *Dla każdego symetrycznego funkcjonału dwuliniowego $h : V \times V \rightarrow \mathbb{K}$ istnieje baza (w_1, \dots, w_n) w V taka, że $h(w_i, w_j) = 0$ dla $i \neq j$ oraz $h(w_j, w_j) = 0$ dla $j > \text{rank}(h)$.*

Niezależnie od wyprowadzenia Twierdzenia 11.2.4 z 11.2.1, warto też podać bezpośrednie uzasadnienie, które jest interpretacją algorytmu Lagrange'a w języku funkcjonałów dwuliniowych, podkreślającą związek tego algorytmu z procedurą wykorzystaną przy ortogonalizacji Grama-Schmidta.

Dowód Twierdzenia 11.2.4. Załóżmy, że h jest funkcjonałem niezerowym i ustalmy dowolną bazę (v_1, \dots, v_n) w przestrzeni V .

(I) Jeśli $h(v_1, v_1) \neq 0$, to przyjmujemy

$$w_1 = v_1 \text{ i } v'_i = v_i - \frac{h(v_i, v_1)}{h(v_1, v_1)} v_1, \quad i = 2, \dots, n.$$

Wówczas dla $i \geq 2$ mamy $h(v'_i, w_1) = h(v_i - \frac{h(v_i, v_1)}{h(v_1, v_1)} v_1, v_1) = h(v_i, v_1) - h(v_i, v_1) = 0$.

(II) Jeśli $h(v_1, v_1) = 0$ ale $h(v_i, v_i) \neq 0$ dla pewnego $i > 0$, to zamieniamy w bazie v_1 z v_i miejscami i jesteśmy w sytuacji takiej jak w (I).

(III) Jeśli $h(v_i, v_i) = 0$ dla $i = 1, \dots, n$, to istnieją $i > j$ takie, że $h(v_i, v_j) = c \neq 0$. Wówczas, podobnie jak w Przykładzie 11.1.7 (a), $h(v_j - v_i, v_j - v_i) = -2c$ (a także $h(v_j + v_i, v_j + v_i) = 2c$ i $h(v_j - v_i, v_j + v_i) = 0$). Zastępując w bazie (v_1, \dots, v_n) wektor v_j przez $v_j - v_i$, a wektor v_i przez $v_j + v_i$ doprowadzamy do sytuacji takiej jak w (II) (lub (I), jeśli $j = 1$).

Tak więc, operacje opisane w (I),(II),(III) pozwalają przejść od bazy (v_1, \dots, v_n) do bazy (w_1, v'_2, \dots, v'_n) takiej, że $h(v'_i, w_1) = 0$ dla $i \geq 2$.

Stosując tę samą procedurę do bazy (v'_2, \dots, v'_n) przestrzeni $\text{lin}(v'_2, \dots, v'_n)$ dostajemy bazę $(w_1, w_2, v''_3, \dots, v''_n)$ przestrzeni V taką, że $h(w_i, w_j) = 0$ dla $i < j \leq 2$ oraz $h(v''_i, w_j) = 0$ dla $j = 1, 2, i \geq 3$.

Po $r = \text{rank}(h)$ krokach dostaniemy bazę (w_1, \dots, w_n) przestrzeni V taką, że $h(w_i, w_j) = 0$ dla $i < j \leq n$ oraz $h(v, w) = 0$ dla $v, w \in \text{lin}(w_{r+1}, \dots, w_n)$. \blacksquare

Uwaga 11.2.5 (a) Opisanym w dowodzie operacjom prowadzącym od bazy (v_1, \dots, v_n) do bazy (w_1, \dots, w_n) odpowiadają opisane w dowodzie Twierdzenia 11.2.1 operacje na wierszach i kolumnach macierzy Grama $A = [a_{ij}]_{i,j=1}^n = [h(v_i, v_j)]_{i,j=1}^n$ funkcjonału h w bazie (v_1, \dots, v_n) .

(b) Etapy (II) i (III) algorytmu można zastąpić operacją elementarną odpowiadającą macierzy elementarnej użytej w Przykładzie 11.1.7 (b). Jeśli $a_{11} = 0$ i $a_{i1} = 0$ dla $i > 1$, to w macierzy A pierwsza kolumna (i wiersz) są zerowe. Przechodzimy wtedy do redukcji macierzy Grama układu (v_2, \dots, v_n) (oznaczanej przez A_1 w dowodzie 11.2.1). Jeśli $a_{11} = 0$ i $a_{i1} \neq 0$ dla pewnego $i > 1$, to zastępujemy wektor v_1 przez $v_1 + v_i$ (do pierwszej wiersza A dodajemy i -ty i symetrycznie, do pierwszej kolumny dodajemy i -tą), co prowadzi do sytuacji takiej jak w (I). Po nie więcej niż n takich krokach dostajemy nową bazę (macierz diagonalną), która po zmianie kolejności wektorów (wierszy i kolumn) spełnia tezę Twierdzenia 11.2.4 (11.2.1). \square

11.3 Warunek Jacobiego.

Algorytm Lagrange'a przebiega szczególnie prosto dla macierzy symetrycznych spełniających następujący warunek.

Definicja 11.3.1 *Macierz symetryczna $A = [a_{ij}]_{i,j=1}^n$ spełnia warunek Jacobiego, jeśli dla $k = 1, \dots, n$, $\Delta_k = \det[a_{ij}]_{i,j=1}^k \neq 0$; wyznaczniki Δ_k nazywamy minorami wiodącymi macierzy A .*

Twierdzenie 11.3.2 *Jeśli $A = A^T \in \mathbb{K}_n^n$ jest macierzą symetryczną spełniającą warunek Jacobiego, to istnieje macierz odwracalna C taka, że iloczyn $C^T A C$ jest macierzą diagonalną mającą na przekątnej wyrazy $\Delta_1, \frac{\Delta_2}{\Delta_1}, \dots, \frac{\Delta_n}{\Delta_{n-1}}$.*

Dowód. Dla macierzy symetrycznej A spełniającej warunek Jacobiego, w algorytmie Lagrange'a przedstawionym w poprzedniej części, wykonuje się jedynie operacje opisane w (I), bo $a_{11} = \Delta_1 \neq 0$ i wykonywane w (I) operacje elementarne typu $(I)_{a(1)+(i)}$ nie zmieniają żadnego z minorów Δ_k (działamy tylko pierwszym wierszem lub kolumną na dalsze wiersze lub kolumny). W szczególności $d_1 = \Delta_1$ i $\Delta_2 = d_1 d_2$, gdzie d_2 jest wyrazem w lewym górnym rogu macierzy A_1 , co pokazuje, że $d_2 \neq 0$, więc w kolejnym kroku również można się ograniczyć do zachowujących minory Δ_k operacji elementarnych.

Po $n-1$ takich krokach dostajemy macierz M taką, że $M A M^T$ jest macierzą diagonalną mającą na przekątnej wyrazy d_1, d_2, \dots, d_n oraz minory wiodące takie jak odpowiednie minory wiodące macierzy A , a stąd $\Delta_k = d_1 \cdot \dots \cdot d_k$, czyli $d_k = \frac{\Delta_k}{\Delta_{k-1}}$, dla $k > 1$. ■

Zauważmy, że w dowodzie założenie $\Delta_n \neq 0$ nie było wykorzystywane, więc założenia twierdzenia możemy nieco osłabić, do warunku $\Delta_k \neq 0$ dla $k < n$.

11.4 Przestrzenie z formą kwadratową.

Wyróżnienie w przestrzeni liniowej V formy kwadratowej Q pozwala mówić o ortogonalności wektorów w (V, Q) .

Definicja 11.4.1 *Niech (V, Q) będzie przestrzenią liniową nad \mathbb{K} z wyróżnioną formą kwadratową $Q : V \rightarrow \mathbb{K}$ i niech $h : V \times V \rightarrow \mathbb{K}$ będzie formą biegunową dla Q , $Q(v) = h(v, v)$. Wektory $v, w \in V$ są ortogonalne, $v \perp w$, jeśli $Q(v+w) = Q(v) + Q(w)$ lub równoważnie, jeśli $h(v, w) = 0$, zob. Uwaga 11.1.2. Baza (w_1, \dots, w_n) w przestrzeni (V, Q) jest ortogonalna jeśli $w_i \perp w_j$ dla $i \neq j$.*

Zgodnie z Twierdzeniem 11.2.4, w każdej przestrzeni (V, Q) istnieje baza ortogonalna.

Definicja 11.4.2 *Niech $Q : V \rightarrow \mathbb{K}$ będzie formą kwadratową. Mówimy, że wektor $v \in V$ jest izotropowy dla Q , jeśli $Q(v) = 0$. Anihilatorem formy Q nazywamy podprzestrzeń liniową $An(Q) = \{v \in V : v \perp w \text{ dla każdego } w \in V\}$ przestrzeni V .*

Zilustrujemy te pojęcia na ważnym przykładzie płaszczyzny Minkowskiego.

Przykład 11.4.3 Płaszczyznę Minkowskiego $(\mathbb{R}^2, x_1^2 - x_2^2)$ nazywamy przestrzeń liniową \mathbb{R}^2 z formą kwadratową $Q(X) = x_1^2 - x_2^2$ dla $X = [x_1, x_2]^T$.

Jeśli wektor $Y = [a, b]^T$ jest taki, że $Q(Y) = a^2 - b^2 = 1$, to wektor $Z = [b, a]^T$ jest ortogonalny do Y na płaszczyźnie Minkowskiego (bo $Q(Z) = -1$, więc $Q(Y+Z) = (a+b)^2 - (a+b)^2 = 0 = Q(Y) + Q(Z)$). Każda taka para (Y, Z) jest bazą ortogonalną w $(\mathbb{R}^2, x_1^2 - x_2^2)$.

Wektory $[1, 1]^T, [1, -1]^T$ są izotropowe na płaszczyźnie Minkowskiego i rozpinają \mathbb{R}^2 , ale $An(Q) = \{\mathbf{0}\}$.

Uwaga 11.4.4 Jeśli (w_1, \dots, w_n) jest bazą ortogonalną w przestrzeni (V, Q) , to anihilator $An(Q)$ jest rozpięty na wektorach izotropowych z tej bazy i $\dim An(Q) = n - \text{rank}(Q)$, bo macierz Grama $G_h(w_1, \dots, w_n)$ formy biegunowej h dla Q , jako diagonalna macierz rzędu $\text{rank}(Q)$, ma $n - \text{rank}(Q)$ zerowych wierszy (i kolumn) odpowiadających wektorom izotropowym bazy (w_1, \dots, w_n) rozpinającym $An(Q)$. □

11.5 Twierdzenie Sylwestera o bezwładności.

Dla przestrzeni liniowych nad ciałem liczb rzeczywistych wzmocnimy Twierdzenie 11.2.4 i uzupełnimy w istotny sposób obserwację z Uwagi 11.4.4.

Uwaga 11.5.1 W przestrzeni liniowej (V, Q) nad \mathbb{R} istnieje baza ortogonalna (u_1, \dots, u_n) taka, że dla $r = \text{rank}(Q)$ i pewnego $s \leq r$

$$(*) \quad Q(\sum_{j=1}^n z_j u_j) = z_1^2 + \dots + z_s^2 - z_{s+1}^2 - \dots - z_r^2,$$

tnz. diagonalna macierz Grama $G_h(u_1, \dots, u_n)$ formy biegunowej dla Q ma na przekątnej s jedynek, $r - s$ minus jedynek i $n - r$ zer.

Istotnie, zgodnie z Twierdzeniem 11.2.4 i Uwagą 11.4.4, istnieje baza ortogonalna (w_1, \dots, w_n) w (V, Q) taka, że $An(Q) = \text{lin}(w_{r+1}, \dots, w_n)$. Zmieniając kolejność wektorów (w_1, \dots, w_r) możemy ponadto założyć, że $Q(w_j) > 0$ dla $j = 1, \dots, s$ i $Q(w_j) < 0$ dla $j = s + 1, \dots, r$.

Z ortogonalności bazy (w_1, \dots, w_n) wynika, że $Q(\sum_{j=1}^n y_j w_j) = \sum_{j=1}^n Q(y_j w_j) = \sum_{j=1}^r y_j^2 Q(w_j)$. Kładąc $u_j = \frac{1}{\sqrt{|Q(w_j)|}} w_j$ dla $j \leq r$ i $u_j = w_j$ dla $j > r$ otrzymujemy bazę ortogonalną (u_1, \dots, u_n) spełniającą $(*)$, bo $Q(\sum_{j=1}^n z_j u_j) = \sum_{j=1}^r Q(z_j \frac{1}{\sqrt{|Q(w_j)|}} w_j) + \sum_{j=r+1}^n Q(z_j w_j) = \sum_{j=1}^r z_j^2 \frac{1}{|Q(w_j)|} Q(w_j)$. \square

Pokażemy, że wskaźniki s i $r - s$ po prawej stronie wzoru $(*)$ nie zależą od wyboru bazy.

Twierdzenie 11.5.2 (o bezwładności). Niech (V, Q) będzie przestrzenią liniową nad \mathbb{R} z wyróżnioną formą kwadratową. Istnieje wówczas rozkład na sumę prostą $V = An(Q) \oplus U \oplus W$ taki, że $Q(u) > 0$ dla $u \in U \setminus \{0\}$, $Q(w) < 0$ dla $w \in W \setminus \{0\}$, przy czym wymiary $s_+(Q) = \dim U$, $s_-(Q) = \dim W$ nie zależą od wyboru U i W w tym rozkładzie.

Dowód. Jeśli (u_1, \dots, u_n) , r i s są takie jak w Uwadze 11.5.1, $V_+ = \text{lin}(u_1, \dots, u_s)$, $V_- = \text{lin}(u_{s+1}, \dots, u_r)$ oraz $An(Q) = \text{lin}(u_{r+1}, \dots, u_n)$, to z $(*)$ mamy

$$(**) \quad Q(u) > 0 \text{ dla } u \in V_+ \setminus \{0\} \text{ i } Q(w) < 0 \text{ dla } w \in V_- \setminus \{0\},$$

co pokazuje, że rozkład $V = An(Q) \oplus V_+ \oplus V_-$ spełnia warunki twierdzenia.

Niech $V = An(Q) \oplus U \oplus W$ będzie dowolnym rozkładem V takim, że forma Q jest dodatnia na $U \setminus \{0\}$ i ujemna na $W \setminus \{0\}$. Wtedy forma Q jest niedodatnia na $An(Q) \oplus W$ (i nieujemna na $An(Q) \oplus U$), bo dla $v + w \in An(Q) \oplus W$, z $v \perp w$ mamy $Q(v + w) = Q(v) + Q(w) \leq 0$ (i podobnie dla $v + u \in An(Q) \oplus U$).

Z $(**)$ mamy więc $V_+ \cap (An(Q) \oplus W) = \{0\}$, co daje $\dim V_+ \leq \dim U$ (i analogicznie, $\dim V_- \leq \dim W$, bo $V_- \cap (An(Q) \oplus U) = \{0\}$), a ponieważ $\dim V_+ + \dim V_- = \dim U + \dim W = n - r$, otrzymujemy stąd $\dim U = \dim V_+$ oraz $\dim W = \dim V_-$. \blacksquare

Twierdzenie o bezwładności pozwala, jak zobaczymy w 11.5.4, na prostą klasyfikację macierzy symetrycznych w \mathbb{R}_n^n z dokładnością do relacji kongruencji.

Definicja 11.5.3 Sygnaturą formy kwadratowej nazywa się liczbę $s(Q) = s_+(Q) - s_-(Q)$.

Zauważmy, że ponieważ $\text{rank}(Q) = s_+(Q) + s_-(Q)$,

$$s_+(Q) = \frac{1}{2}(\text{rank}(Q) + s(Q)), \quad s_-(Q) = \frac{1}{2}(\text{rank}(Q) - s(Q))$$

Uwaga 11.5.4 Niech $J_{s,r} \in \mathbb{R}_n^n$ będzie macierzą diagonalną mającą na przekątnej kolejno s jedynek, $r - s$ minus jedynek i $n - r$ zer. Macierz symetryczna $A = A^T \in \mathbb{R}_n^n$ jest kongruentna z macierzą $J_{s,r}$ wtedy i tylko wtedy, gdy dla formy kwadratowej $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ danej wzorem $Q(X) = X^T A X$ mamy $r = \text{rank}(Q)$ i $s = s_+(Q)$.

Istotnie, macierz A jest kongruentna z macierzą $J_{s,r}$ wtedy i tylko wtedy, gdy istnieje izomorfizm liniowy $S : \mathbb{R}^n \rightarrow \mathbb{R}^n$ taki, że $Q \circ S(Z) = Z^T J_{s,r} Z$ zob. Uwaga 11.1.4 (b). Interpretując S^{-1} jako układ współrzędnych $\tau : \mathbb{R}^n \rightarrow \mathbb{R}^n$ związany z bazą $(S(E_1), \dots, S(E_n))$, zob. Uwaga 4.6.2, widzimy, że jest to równoważne istnieniu opisanej w Uwadze 11.5.1 bazy ortogonalnej w przestrzeni (\mathbb{R}^n, Q) . Z definicji sygnatury, istnienie takiej bazy oznacza, że $\text{rank}(Q) = r$ i $s_+(Q) = s$. \square

Na zakończenie podamy ważną charakteryzację dodatniej określoności form kwadratowych – *kryterium Sylwestera*.

Uwaga 11.5.5 Niech $A = A^T \in \mathbb{R}_n^n$, niech $\Delta_1, \dots, \Delta_n$ będą minorami wiodącymi macierzy A (zob. 11.3.1) i niech $Q(X) = X^T A X$.

- (a) Forma kwadratowa Q jest *dodatnio określona*, tzn. ma sygnaturę $s(Q) = n$, wtedy i tylko wtedy, gdy wszystkie minory Δ_j są dodatnie i wówczas forma biegunowa dla Q , $h(X, Y) = X^T A Y$ jest iloczynem skalarnym w \mathbb{R}^n .
- (b) Forma kwadratowa Q jest *ujemnie określona*, tzn. ma sygnaturę $s(Q) = -n$, wtedy i tylko wtedy, gdy $(-1)^j \Delta_j > 0$ dla $j = 1, 2, \dots$.

Istotnie, jeśli forma Q jest dodatnio określona, to forma biegunowa $h(X, Y) = X^T A Y$ jest iloczynem skalarnym, więc wszystkie minory Δ_j są dodatnie jako wyznaczniki macierzy Grama $G_h(E_1, \dots, E_j)$, zob. Twierdzenie 7.3.5. Jeśli forma Q jest ujemnie określona, to $-h$ jest iloczynem skalarnym, więc wszystkie minory wiodące macierzy $-A$ są dodatnie. Implikacje odwrotne wynikają z Twierdzenia 11.3.2. \square

12 Funkcje kwadratowe

Także i w tym rozdziale utrzymujemy założenie, że ciało \mathbb{K} ma charakterystykę różną od 2.

12.1 Funkcje kwadratowe na \mathbb{K}^n .

Definicja 12.1.1 Funkcją kwadratową na \mathbb{K}^n nazywamy funkcję $f : \mathbb{K}^n \rightarrow \mathbb{K}$ postaci

$$(*) \quad f(X) = X^T A X + B^T X + c, \text{ gdzie } A = A^T \in \mathbb{K}_n^n, A \neq \mathbf{0}, B \in \mathbb{K}^n, c = f(\mathbf{0}) \in \mathbb{K}.$$

Niezerową formę kwadratową $Q(X) = X^T A X$ będziemy nazywać częścią kwadratową funkcji f .

Uwaga 12.1.2 Funkcja kwadratowa f wyznacza swoją część kwadratową. Istotnie, dla $X, Y \in \mathbb{K}^n$ mamy $f(X + Y) - f(X) - f(Y) + f(\mathbf{0}) = (X + Y)^T A(X + Y) + B^T(X + Y) + c - (X^T A X + B^T X + c) - (Y^T A Y + B^T Y + c) + c = X^T A Y + Y^T A X = 2X^T A Y$. \square

Dla $n = 1$ funkcję kwadratową $f(x) = ax^2 + bx + c = 0$, $a \neq 0$, redukuje się do postaci $ay^2 + c'$ podstawieniem $y = x + \frac{b}{2a}$ odpowiadającym złożeniu $f \circ g$ funkcji f z przesunięciem $g(y) = y - \frac{b}{2a}$. Ta obserwacja jest szczególnym przypadkiem głównego twierdzenia w tej części – twierdzenia o redukcji afinicznej.

Uwaga 12.1.3 Dla izomorfizmu afinicznego $g : \mathbb{K}^n \rightarrow \mathbb{K}^n$ danego wzorem $g(Z) = MZ + X_0$, zob. 9.1.2, i funkcji kwadratowej $f : \mathbb{K}^n \rightarrow \mathbb{K}$ opisanej formułą (*), złożenie $f \circ g : \mathbb{K}^n \rightarrow \mathbb{K}$ jest funkcją kwadratową opisaną wzorem

$$(**) \quad f \circ g(Z) = Z^T A' Z + B'^T Z + c', \text{ gdzie } A' = M^T A M, B' = (2X_0^T A + B^T)M, c' = f(X_0).$$

W szczególności, macierz A' części kwadratowej funkcji $f \circ g$ jest kongruentna z macierzą A .

Istotnie, podstawiając $g(Z) = MZ + X_0$ do (*) dostajemy $(MZ + X_0)^T A(MZ + X_0) + B^T(MZ + X_0) + c = (MZ)^T A(MZ) + ((MZ)^T A X_0 + X_0^T A M Z + B^T M Z) + (X_0^T A X_0 + B^T X_0 + c)$ i ponieważ składniki tej sumy są skalarami, a macierz A jest symetryczna, mamy $(MZ)^T A X_0 = X_0^T A M Z$, więc wyrażenie w pierwszym nawiasie można zapisać w postaci $(2X_0^T A + B^T)M Z$. \square

Pokażemy, że izomorfizm g można dobrać tak, by złożenie $f \circ g(Z)$ było dane prostym wzorem.

Twierdzenie 12.1.4 (o redukcji afinicznej). Dla każdej funkcji kwadratowej $f : \mathbb{K}^n \rightarrow \mathbb{K}$ istnieje izomorfizm afiniczny $g : \mathbb{K}^n \rightarrow \mathbb{K}^n$ taki, że $f \circ g(Z)$ jest dane jednym ze wzorów:

$$\begin{aligned} \text{(AI)}_r \quad & f \circ g(Z) = d_1 z_1^2 + \dots + d_r z_r^2 + c', \quad d_j \neq 0 \text{ dla } j \leq r, \\ \text{(AII)}_r \quad & f \circ g(Z) = d_1 z_1^2 + \dots + d_r z_r^2 + z_n, \quad d_j \neq 0 \text{ dla } j \leq r < n, \end{aligned}$$

gdzie $Z = [z_1, \dots, z_n]^T$, a $r = \text{rank}(Q)$ jest rzędem części kwadratowej f .

Dowód twierdzenia. Niech f będzie funkcją daną wzorem (*). Z Twierdzenia 11.2.1 istnieje macierz odwracalna $M \in \mathbb{K}_n^n$ taka, że $M^T A M$ jest macierzą diagonalną mającą na przekątnej niezerowe wyrazy d_1, \dots, d_r na kolejnych pierwszych r miejscach oraz $n - r$ zer na pozostałych miejscach.

Dla izomorfizmu liniowego $g_1(Y) = MY$, gdzie $Y = [y_1, \dots, y_n]^T$, złożenie $f \circ g_1$ jest opisane wzorem $f \circ g_1(Y) = Y^T M^T A M Y + B^T M Y + c$ i przyjmując $B^T M = [a_1, \dots, a_n]$, dostajemy $f \circ g_1(Y) = d_1 y_1^2 + \dots + d_r y_r^2 + a_1 y_1 + \dots + a_n y_n + c = d_1 (y_1 + \frac{a_1}{2d_1})^2 + \dots + d_r (y_r + \frac{a_r}{2d_r})^2 + a_{r+1} y_{r+1} + \dots + a_n y_n + c'$.

Izomorfizm g_2 odwrotny do podstawienia $z_j = y_j + \frac{a_j}{2d_j}$ dla $j \leq r$ i $z_j = y_j$ dla $j > r$ jest przesunięciem takim, że złożenie $f \circ g_1 \circ g_2$ jest opisane wzorem $f \circ g_1 \circ g_2(Z) = d_1 z_1^2 + \dots + d_r z_r^2 + a_{r+1} z_{r+1} + \dots + a_n z_n + c'$, $Z = [z_1, \dots, z_n]^T$.

(I) Jeśli wszystkie współczynniki a_{r+1}, \dots, a_n są zerowe, to dla $g = g_1 \circ g_2$ mamy wzór (AI)_r.

(II) Jeśli dla pewnego $i > r$ współczynnik $a_i \neq 0$, to podstawienie $\tilde{z}_i = a_{r+1} z_{r+1} + \dots + a_n z_n + c'$ oraz $\tilde{z}_j = z_j$ dla $j \neq i$ jest odwracalne, a izomorfizm afiniczny g_3 odwrotny do tego podstawienia

proceedzi do wzoru $f \circ g_1 \circ g_2 \circ g_3(\tilde{Z}) = d_1 \tilde{z}_1^2 + \dots + d_r \tilde{z}_r^2 + \tilde{z}_i$, $\tilde{Z} = [\tilde{z}_1, \dots, \tilde{z}_n]^T$, z którego po zmianie współrzędnej \tilde{z}_i z \tilde{z}_n , określającej izomorfizm g_4 , dostajemy dla $g = g_1 \circ g_2 \circ g_3 \circ g_4$ wzór (AII)_r. ■

Uwaga 12.1.5 Jeśli, w oznaczeniach Twierdzenia 12.1.4, układ równań $2AX + B = \mathbf{0}$ ma rozwiązanie $X_0 \in \mathbb{K}^n$, to $\mathbf{0}^T = (2AX_0 + B)^T = 2X_0^T A + B^T$, więc z (**) dla izomorfizmu afinicznego $g(Z) = MZ + X_0$ (gdzie macierz M jest taka jak w dowodzie twierdzenia), złożenie $f \circ g$ jest opisane wzorem (AI)_r. □

Uwaga 12.1.5 wskazuje własność wyróżniającą funkcje kwadratowe redukujące się do postaci (AI)_r.

Definicja 12.1.6 Wektor $X_0 \in \mathbb{K}^n$ nazywamy *środkiem symetrii funkcji* $f : \mathbb{K}^n \rightarrow \mathbb{K}$, jeśli $f(X_0 + X) = f(X_0 - X)$ dla $X \in \mathbb{K}^n$. Zbiór środków symetrii funkcji f oznaczamy przez $C(f)$.

Pokażemy, że układ równań z Uwagi 12.1.5 opisuje zbiór środków symetrii funkcji kwadratowej f .

Uwaga 12.1.7 Dla funkcji kwadratowej f danej wzorem (*) równość $f(X_0 + X) = f(X_0 - X)$ oznacza (ze wzoru (**)) dla $M = \pm I_n$, że $X^T AX + (2X_0^T A + B^T)X + c' = (-X)^T A(-X) + (2X_0^T A + B^T)(-X) + c'$, czyli $(2X_0^T A + B^T)X = 0$, a stąd zbiór $C(f)$ jest zbiorem rozwiązań układu równań $2AX + B = \mathbf{0}$. □

Uwaga 12.1.8 Niech f i g będą takie jak w Twierdzeniu 12.1.4 i załóżmy, że $f' = f \circ g$ jest dana wzorem (AI)_r. Zbiór $C(f')$ jest opisany układem równań $z_1 = \dots = z_r = 0$ i f' jest stale równa c' na tym zbiorze. Co więcej, $C(f) = g(C(f'))$ (bo dla $X_0 = g(Z_0)$ i $X = \tilde{g}(Z)$ mamy $f(X_0 \pm X) = f(g(Z_0 \pm Z))$), więc $X_0 \in C(f)$ jeśli $Z_0 \in C(f')$ i f jest stale równa c' na $C(f)$. Funkcja f'' dana wzorem (AII)_r nie ma środków symetrii, więc nie może być zredukowana do żadnej funkcji danej wzorem (AI)_r. □

12.2 Klasyfikacja funkcji kwadratowych na \mathbb{R}^n .

Dla $\mathbb{K} = \mathbb{R}$ tezę Twierdzenia 12.1.4 można wzmocnić, korzystając z Uwagi 11.5.1.

Twierdzenie 12.2.1 Dla każdej funkcji kwadratowej $f : \mathbb{R}^n \rightarrow \mathbb{R}$ istnieje izomorfizm afiniczny $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ taki, że złożenie $f \circ g(Z)$ jest opisane jednym ze wzorów:

$$\begin{aligned} \text{(AI)}_{s,r} \quad f \circ g(Z) &= z_1^2 + \dots + z_s^2 - z_{s+1}^2 - \dots - z_r^2 + c', & s \leq r \leq n, \\ \text{(AII)}_{s,r} \quad f \circ g(Z) &= z_1^2 + \dots + z_s^2 - z_{s+1}^2 - \dots - z_r^2 + z_n, & s \leq r < n, \end{aligned}$$

gdzie $Z = [z_1, \dots, z_n]^T$, a $r = \text{rank}(Q)$ i $s = s_+(Q)$ dla formy Q będącej częścią kwadratową f .

Definicja 12.2.2 Powiemy, że funkcje $f, f' : \mathbb{R}^n \rightarrow \mathbb{R}$ są *afinicznie równoważne* jeśli istnieje izomorfizm afiniczny $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ taki, że $f' = f \circ g$.

Z Uwag 12.1.8 i 11.5.4 wynika następująca obserwacja.

Uwaga 12.2.3 Każda funkcja kwadratowa na \mathbb{R}^n jest afinicznie równoważna dokładnie jednej z funkcji typu (AI)_{s,r} lub (AII)_{s,r} z Twierdzenia 12.2.1. □

W kartezjańskiej przestrzeni euklidesowej $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ redukcję opisaną w Twierdzeniu 12.1.4, z pewnymi niewielkimi zmianami, można zrealizować przy pomocy izometrii (czyli izomorfizmów afinicznych \mathbb{R}^n , których część liniowa jest opisana macierzą ortogonalną, zob. Twierdzenie 10.2.1).

Twierdzenie 12.2.4 (o redukcji euklidesowej). Dla każdej funkcji kwadratowej $f : \mathbb{R}^n \rightarrow \mathbb{R}$ na kartezjańskiej przestrzeni euklidesowej $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ istnieje izometria $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ taka, że złożenie $f \circ g(Z)$ jest opisane jednym ze wzorów:

$$\begin{aligned} \text{(EI)}_{\lambda_1, \dots, \lambda_r} \quad f \circ g(Z) &= \lambda_1 z_1^2 + \dots + \lambda_r z_r^2 + c', & \lambda_1 \geq \dots \geq \lambda_r \text{ i } \lambda_j \neq 0 \text{ dla } j \leq r, \\ \text{(EII)}_{\lambda_1, \dots, \lambda_r} \quad f \circ g(Z) &= \lambda_1 z_1^2 + \dots + \lambda_r z_r^2 + bz_n, & \lambda_1 \geq \dots \geq \lambda_r \text{ i } \lambda_j \neq 0 \text{ dla } j \leq r < n, \end{aligned}$$

gdzie $Z = [z_1, \dots, z_n]^T$, a $\lambda_1, \dots, \lambda_r$ są niezerowymi wartościami własnymi symetrycznej macierzy $A \in \mathbb{R}_n^n$ takiej, że forma $Q(X) = X^T AX$ jest częścią kwadratową f .

Dowód. W dowodzie 12.1.4 trzeba jedynie dokonać następujących modyfikacji.

Izomorfizm g_1 należy zastąpić izometrią liniową $g_1(X) = MX$, gdzie M jest macierzą ortogonalną otrzymaną z występującej we Wniosku 8.1.7 macierzy ortogonalnej C , której kolumny zostały uporządkowane tak, aby odpowiadający im ciąg wartości własnych miał na początkowych $r = \text{rank } A$ miejscach nieronący ciąg wyrazów niezerowych i $n - r$ zer na końcu.

Izomorfizm $g_3 \circ g_4$ w kroku (II) należy zastąpić izometrią liniową $g_3 = \tau^{-1}$, gdzie $\tau : \mathbb{R}^n \rightarrow \mathbb{R}^n$ jest prostokątnym układem współrzędnych nie zmieniającym pierwszych r współrzędnych wektora $Z = [z_1, \dots, z_n]^T$ i zastępującego ostatnią współrzędną sumą $\frac{1}{b}(a_{r+1}z_{r+1} + \dots + a_n z_n + c')$, gdzie $b = \sqrt{\sum_{j=r+1}^n a_j^2}$. Macierz ortogonalna $M(g_3) = M(\tau)^{-1} = M(\tau)^T$ ma kolumny $g_3(E_j) = E_j$ dla $j \leq r$, a jej ostatnią kolumną jest $g_3(E_n) = \frac{1}{b} \sum_{j=r+1}^n a_j E_j$. ■

Definicja 12.2.5 Powiemy, że funkcje kwadratowe $f, f' : \mathbb{R}^n \rightarrow \mathbb{R}$ na kartezjańskiej przestrzeni euklidesowej $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ są euklidesowo równoważne, jeśli istnieje izometria $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ taka, że $f' = f \circ g$.

Z Twierdzenia 12.2.4, podobnie jak w Uwadze 12.2.3, wynika następująca obserwacja.

Uwaga 12.2.6 Każda funkcja kwadratowa na kartezjańskiej przestrzeni euklidesowej $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ jest euklidesowo równoważna dokładnie jednej z funkcji typu (EI) $_{\lambda_1, \dots, \lambda_r}$ lub (EII) $_{\lambda_1, \dots, \lambda_r}$ z Twierdzenia 12.2.4. □

13 Hiperpowierzchnie stopnia 2.

W dalszym ciągu zakładamy, że ciało \mathbb{K} ma charakterystykę różną od 2.

13.1 Hiperpowierzchnie stopnia 2.

Hiperpowierzchnie stopnia 2 w przestrzeni afinicznej E są niepustymi zbiorami punktów w E opisanymi w ustalonym afinicznym układzie współrzędnych równaniem kwadratowym, to znaczy równaniem postaci $f(X) = 0$, gdzie f jest funkcją kwadratową na \mathbb{K}^n . Dodatkowo uznajemy, że hiperpłaszczyzn w E , które można opisać równaniem liniowym $B^T X + c = 0$ (a także równaniem kwadratowym $(B^T X + c)^2 = 0$) nie będziemy uważać za hiperpowierzchnie stopnia 2. Dokładniej, przyjmujemy następującą definicję.

Definicja 13.1.1 *Niech E będzie przestrzenią afiniczną. Niepusty zbiór $H \in E$, który nie jest hiperpłaszczyzną, jest hiperpowierzchnią stopnia 2, jeśli dla pewnego afinicznego układu współrzędnych $\sigma_p : E \rightarrow \mathbb{K}^n$ i funkcji kwadratowej $f : \mathbb{K}^n \rightarrow \mathbb{K}$,*

$$H = \{\mathbf{x} \in E : f(X) = 0 \text{ dla } X = \sigma_p(\mathbf{x})\}.$$

Będziemy wtedy mówili, że równanie $f(X) = 0$ opisuje H w układzie współrzędnych σ_p .

Twierdzenie 13.1.2 *Dla hiperpowierzchni H stopnia 2 w przestrzeni afinicznej E nad \mathbb{K} można wybrać afiniczny układ współrzędnych $\tau_q : E \rightarrow \mathbb{K}^n$ taki, że hiperpowierzchnia H jest opisana w tym układzie jednym z równań*

$$\begin{aligned} \text{(AI)}_r & \quad d_1 z_1^2 + \dots + d_r z_r^2 + c' = 0, & d_j \neq 0 \text{ dla } j \leq r, \\ \text{(AII)}_r & \quad d_1 z_1^2 + \dots + d_r z_r^2 + z_n = 0, & d_j \neq 0 \text{ dla } j \leq r < n, \end{aligned}$$

gdzie $\tau_q(\mathbf{x}) = [z_1, \dots, z_n]^T$.

Dowód. Ustalmy afiniczny układ współrzędnych $\sigma_p : E \rightarrow \mathbb{K}^n$ i funkcję kwadratową $f : \mathbb{K}^n \rightarrow \mathbb{K}$ taką, że równanie $f(X) = 0$ opisuje H w układzie σ_p . Z twierdzenia o redukcji afinicznej, istnieje izomorfizm afiniczny $g : \mathbb{K}^n \rightarrow \mathbb{K}^n$ taki, że złożenie $f \circ g$ jest dane jednym z wzorów (AI) $_r$ lub (AII) $_r$. Wówczas

$$H = \{\mathbf{x} \in E : f \circ \sigma_p(\mathbf{x}) = 0\} = \{\mathbf{x} \in E : (f \circ g) \circ (g^{-1} \circ \sigma_p)(\mathbf{x}) = 0\},$$

a więc, korzystając z równoważności (i) \Leftrightarrow (ii) w Twierdzeniu 9.6.3, możemy przyjąć $\tau_q = g^{-1} \circ \sigma_p$. ■

W równaniu (AI) $_r$ można ponadto zakładać, że c' jest zerem lub jedynką, bo jeśli $c' \neq 0$, to obie strony (AI) $_r$ można pomnożyć przez $(c')^{-1}$, nie zmieniając zbioru rozwiązań.

W dalszej części będziemy rozważać wyłącznie hiperpowierzchnie w przestrzeni afinicznej E nad ciałem liczb rzeczywistych. Dla $\mathbb{K} = \mathbb{R}$, zastępując w powyższym dowodzie twierdzenie o redukcji afinicznej Twierdzeniem 12.2.1, można dodatkowo uprościć równania (AI) $_r$ i (AII) $_r$.

Twierdzenie 13.1.3 *Dla hiperpowierzchni H stopnia 2 w przestrzeni afinicznej E nad \mathbb{R} można wybrać afiniczny układ współrzędnych $\tau_q : E \rightarrow \mathbb{R}^n$ taki, że hiperpowierzchnia H jest opisana w tym układzie jednym z równań*

$$\begin{aligned} \text{(AI)}_{s,r} & \quad z_1^2 + \dots + z_s^2 - z_{s+1}^2 - \dots - z_r^2 = 0, & 0 < \frac{r}{2} \leq s \leq r \leq n, \\ \text{(AI)}_{s,r} & \quad z_1^2 + \dots + z_s^2 - z_{s+1}^2 - \dots - z_r^2 + 1 = 0, & 0 \leq s \leq r \leq n, \\ \text{(AII)}_{s,r} & \quad z_1^2 + \dots + z_s^2 - z_{s+1}^2 - \dots - z_r^2 + z_n = 0, & 0 < \frac{r}{2} \leq s \leq r < n, \end{aligned}$$

gdzie $\tau_q(\mathbf{x}) = [z_1, \dots, z_n]^T$.

Dowód. Jak w dowodzie Twierdzenia 13.1.2, założmy, że $f : \mathbb{R}^n \rightarrow \mathbb{R}$ jest funkcją kwadratową taką, że równanie $f(X) = 0$ opisuje H w afinicznym układzie współrzędnych $\sigma_p : E \rightarrow \mathbb{R}^n$.

Jeśli f ma środek symetrii X_0 , to możemy założyć, że f przyjmuje w X_0 wartość 0 lub 1, bo jeśli $f(X_0) \neq 0$, to f można zastąpić przez $\frac{1}{f(X_0)} f$.

Jeśli f zeruje się w swoim środku symetrii, lub też f nie ma środka symetrii, to zastępując w razie potrzeby f przez $-f$, możemy założyć, że część kwadratowa Q funkcji f spełnia warunek $s_+(Q) \geq s_-(Q)$. Dla $s = s_+(Q)$ i $r = \text{rank}(Q) = s_+(Q) + s_-(Q)$ mamy wtedy $r \leq 2s$, czyli $\frac{r}{2} \leq s$.

Dla tak zmodyfikowanej funkcji kwadratowej f , opisującej H w układzie σ_p , istnieje izomorfizm afiniczny $g: \mathbb{R}^n \rightarrow \mathbb{R}^n$ taki, że złożenie $f \circ g$ jest dane jednym z wzorów (AI) $_r$ lub (AII) $_r$ z Twierdzenia 12.2.1. Podobnie jak w dowodzie Twierdzenia 13.1.2, dostajemy stąd tezę przyjmując $\tau_q = g^{-1} \circ \sigma_p$. ■

Występujące w Twierdzeniu 13.1.3 równania będziemy nazywali *równaniami kanonicznymi*.

Równania (AI) $_{1,r,r}$ (postaci $z_1^2 + \dots + z_r^2 + 1 = 0$) są sprzeczne, a równania (AI) $_{0,r,r}$ (postaci $z_1^2 + \dots + z_r^2 = 0$) opisują podprzestrzenie afiniczne wymiaru $n-r$ w E (zbiory rozwiązań układu równań $z_1 = \dots = z_r = 0$).

Definicja 13.1.4 Niech E będzie przestrzenią afiniczną nad \mathbb{R} . Mówimy, że $H \subset E$ jest hiperpowierzchnią właściwą w E , jeśli H jest hiperpowierzchnią stopnia 2, która nie jest podprzestrzenią afiniczną E .

Z kolejnej uwagi wynika, że wszystkie pozostałe równania kanoniczne opisują hiperpowierzchnie właściwe.

Uwaga 13.1.5 Jeśli hiperpowierzchnia H stopnia 2 w przestrzeni afinicznej E nad \mathbb{R} jest, w układzie współrzędnych τ_q związanym z układem bazowym $(q; w_1, \dots, w_n)$ w E , opisana jednym z równań kanonicznych, z wyłączeniem równań typu (AI) $_{0,r,r}$ i (AI) $_{1,r,r}$, to istnieje prosta $L_0 \subset E$ przecinająca H w dokładnie dwóch punktach.

Istotnie, dla każdego z takich równań kanonicznych możemy zdefiniować prostą L_0 w układzie współrzędnych τ_q następująco:

- prosta $z_r = 1$ i $z_j = 0$ dla pozostałych $j > 2$ ($L_0 = q + w_r + \text{lin}(w_1)$) dla równań (AI) $_{s,r}$, $s < r$,
- prosta $z_j = 0$ dla $j \neq r$ ($L_0 = q + \text{lin}(w_r)$) dla równań (AI) $_{s,r}$, $s < r$,
- prosta $z_{r+1} = -1$ i $z_j = 0$ dla pozostałych $j > 2$ ($L_0 = q + (-w_{r+1}) + \text{lin}(w_1)$) dla równań (AII) $_{s,r}$. □

Podamy teraz przykłady hiperpowierzchni właściwych w \mathbb{R}^n dla $n = 2, 3$ uwzględniając, dla każdego z typów równań kanonicznych, wskazane przez nas ograniczenia na indeksy s i r :

- (AI) $_{s,r}$ – $0 < \frac{r}{2} \leq s < r \leq n$,
- (AI) $_{1,s,r}$ – $0 \leq s < r \leq n$,
- (AII) $_{s,r}$ – $0 < \frac{r}{2} \leq s \leq r < n$.

Hiperpowierzchnie właściwe w \mathbb{R}^2 nazywamy krzywymi stopnia 2.

grupa	r	s	równanie kanoniczne	krzywa
(AI) $_0$	2	1	$z_1^2 - z_2^2 = 0$	para przecinających się prostych ($z_1 = \pm z_2$)
(AI) $_1$	2	1	$z_1^2 - z_2^2 + 1 = 0$	hiperbola
–''–	2	0	$-z_1^2 - z_2^2 + 1 = 0$	elipsa
–''–	1	0	$-z_1^2 + 1 = 0$	para prostych równoległych ($z_1 = \pm 1$)
(AII)	1	1	$z_1^2 + z_2 = 0$	parabola

Hiperpowierzchnie właściwe w \mathbb{R}^3 nazywamy powierzchniami stopnia 2. W identyfikacji powierzchni pomaga analiza jej przekrojów płaszczyznami (na przykład takimi jak wskazane w tabeli).

grupa	r	s	równanie kanoniczne	przekroje	powierzchnia
(AI ₀)	3	2	$z_1^2 + z_2^2 - z_3^2 = 0$	$z_3 = c$	stożek
—''—	2	1	$z_1^2 - z_2^2 = 0$		para przecinających się płaszczyzn ($z_1 = \pm z_2$)
(AI ₁)	3	2	$z_1^2 + z_2^2 - z_3^2 + 1 = 0$	$z_3 = c$	hiperboloida dwupowłokowa
—''—	3	1	$z_1^2 - z_2^2 - z_3^2 + 1 = 0$	$z_1 = c$	hiperboloida jednopowłokowa
—''—	3	0	$-z_1^2 - z_2^2 - z_3^2 + 1 = 0$		elipsoida
—''—	2	1	$z_1^2 - z_2^2 + 1 = 0$		walec hiperboliczny
—''—	2	0	$-z_1^2 - z_2^2 + 1 = 0$		walec eliptyczny
—''—	1	0	$-z_1^2 + 1 = 0$		para płaszczyzn równoległych ($z_1 = \pm 1$)
(AII)	2	2	$z_1^2 + z_2^2 + z_3 = 0$	$z_3 = c$	paraboloida eliptyczna
—''—	2	1	$z_1^2 - z_2^2 + z_3 = 0$		paraboloida hiperboliczna
—''—	1	1	$z_1^2 + z_2 = 0$		walec paraboliczny

13.2 Klasyfikacja hiperpowierzchni właściwych.

Jeśli hiperpowierzchnia H stopnia 2 w E jest podprzestrzenią afiniczną E wymiaru $n - r$ opisaną w pewnym afinicznym układzie współrzędnych równaniem kanonicznym $z_1^2 + \dots + z_r^2 = 0$, to $r > 1$ (bo H nie jest hiperpłaszczyzną) i H jest też opisana równaniem $2z_1^2 + \dots + z_r^2 = 0$, które nie jest krotnością równania kanonicznego. Natomiast hiperpowierzchnia właściwa wyznacza swoje równanie z dokładnością do stałej.

Twierdzenie 13.2.1 *Niech H będzie hiperpowierzchnią właściwą w przestrzeni afinicznej E nad \mathbb{R} . Jeśli $f, f_* : \mathbb{R}^n \rightarrow \mathbb{R}$ są funkcjami kwadratowymi takimi, że równania $f(X) = 0$ i $f_*(X) = 0$ opisują H w tym samym afinicznym układzie współrzędnych w E , to $f_* = \lambda f$ dla pewnego $\lambda \in \mathbb{R}$.*

Dowód tego twierdzenia podamy w części 14.4 uzupełnień.

Wniosek 13.2.2 *Hiperpowierzchnia właściwa w przestrzeni afinicznej E nad \mathbb{R} wyznacza swoje równanie kanoniczne.*

Dowód. Niech $\sigma_p, \tau_q : E \rightarrow \mathbb{R}^n$ będą afinicznymi układami współrzędnych w E , $f, f_* : \mathbb{R}^n \rightarrow \mathbb{R}$ funkcjami kwadratowymi i założmy, że równanie $f(X) = 0$ opisuje H w układzie σ_p , a równanie $f_*(Z) = 0$ jest kanoniczne i opisuje H w układzie τ_q . Pokażemy, że redukcja równania $f(X) = 0$ prowadzi do równania kanonicznego takiego samego typu i z takimi samymi indeksami r, s jak równanie $f_*(Z) = 0$.

Mamy $\mathbf{x} \in H \Leftrightarrow f \circ \sigma_p(\mathbf{x}) = 0 \Leftrightarrow f \circ \sigma_p \circ ((\tau_q)^{-1} \circ \tau_q)(\mathbf{x}) = 0$. Zatem równanie $f \circ (\sigma_p \circ (\tau_q)^{-1})(Z) = 0$ opisuje hiperpowierzchnię H w układzie τ_q i z Twierdzenia 13.2.1 istnieje $\lambda \in \mathbb{R}$ takie, że $f_* = \lambda f \circ (\sigma_p \circ (\tau_q)^{-1})$. Zgodnie z Twierdzeniem 9.6.3, złożenie $\sigma_p \circ (\tau_q)^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ jest izomorfizmem afinicznym \mathbb{R}^n , więc funkcje f_* oraz λf są afinicznie równoważne i teza wynika z Uwagi 12.2.3. ■

Definicja 13.2.3 *Mówimy, że hiperpowierzchnie właściwe H, H' w przestrzeni afinicznej E nad \mathbb{R} mają ten sam typ afiniczny, jeśli istnieje izomorfizm afiniczny $g : E \rightarrow E$ taki, że $g(H) = H'$.*

Twierdzenie 13.2.4 *Hiperpowierzchnie właściwe H, H' w przestrzeni afinicznej E nad \mathbb{R} mają ten sam typ afiniczny wtedy i tylko wtedy, gdy wyznaczają to samo równanie kanoniczne.*

Dowód. Niech $g : E \rightarrow E$ będzie izomorfizmem afinicznym. Jeśli $g(H)$ jest hiperpowierzchnią stopnia 2 opisaną w układzie τ_q równaniem $f(Z) = 0$, to $\mathbf{x} \in H \Leftrightarrow g(\mathbf{x}) \in g(H) \Leftrightarrow f \circ \tau_q(g(\mathbf{x})) = 0$, więc hiperpowierzchnia H jest opisana równaniem $f(X) = 0$ w układzie $\sigma_p = \tau_q \circ g$, zob. Twierdzenie 9.6.3.

Jeśli $\sigma_p, \tau_q : E \rightarrow \mathbb{R}^n$ są afinicznymi układami współrzędnych w E takimi, że dla funkcji kwadratowej $f : \mathbb{R}^n \rightarrow \mathbb{R}$ równanie $f(X) = 0$ opisuje H w układzie σ_p , a równanie $f(Z) = 0$ opisuje H' w układzie τ_q , to $\mathbf{x} \in H \Leftrightarrow f \circ \sigma_p(\mathbf{x}) = 0 \Leftrightarrow f \circ (\tau_q \circ (\tau_q)^{-1}) \circ \sigma_p(\mathbf{x}) = 0 \Leftrightarrow f \circ \tau_q \circ ((\tau_q)^{-1} \circ \sigma_p)(\mathbf{x}) = 0 \Leftrightarrow (\tau_q)^{-1} \circ \sigma_p(\mathbf{x}) \in H'$. Izomorfizm afiniczny $(\tau_q)^{-1} \circ \sigma_p : E \rightarrow E$ przeprowadza więc H na H' , zob. Twierdzenie 9.6.3. ■

Geometryczną własnością odróżniającą hiperpowierzchnie opisane równaniami typu (AI)₀, (AI)₁ od hiperpowierzchni opisanych równaniami typu (AII) jest istnienie środka symetrii hiperpowierzchni.

Definicja 13.2.5 Niech H będzie niepustym podzbiorem przestrzeni afinicznej E . Punkt $p \in E$ nazywamy środkiem symetrii zbioru H jeśli dla $v \in \vec{E}$ warunki $p + v \in E$ i $p + (-v) \in E$ są równoważne.

Twierdzenie 13.2.6 Niech H będzie hiperpowierzchnią właściwą w przestrzeni afinicznej E nad \mathbb{R} . Jeśli $f : \mathbb{R}^n \rightarrow \mathbb{R}$ jest funkcją kwadratową taką, że równanie $f(Y) = 0$ opisuje H w układzie współrzędnych $\sigma_p : E \rightarrow \mathbb{R}^n$, to punkt $q \in E$ jest środkiem symetrii H wtedy i tylko wtedy, gdy $\sigma_p(q)$ jest środkiem symetrii funkcji f .

Dowód. Ustalmy $q \in E$, $X_0 = \sigma_p(q)$ i niech $X = \sigma(v)$ dla $v \in \vec{E}$. Wtedy

$$\sigma_p(q + v) = X_0 + X \quad \text{oraz} \quad \sigma_p(q + (-v)) = X_0 - X.$$

Zatem q jest środkiem symetrii hiperpowierzchni H , jeśli X_0 jest środkiem symetrii funkcji f .

Założmy, że q jest środkiem symetrii H i niech $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ będzie symetrią afiniczną \mathbb{R}^n względem X_0 , tzn. $g(X_0 + X) = X_0 - X$. Złożenie $f_* = f \circ g : \mathbb{R}^n \rightarrow \mathbb{R}$ jest funkcją kwadratową taką, że równanie $f_*(Y) = 0$ opisuje H w układzie σ_p , bo $q + v \in H \Leftrightarrow q - v \in H \Leftrightarrow f(X_0 - X) = 0 \Leftrightarrow f_*(X_0 + X) = 0$.

Z Twierdzenia 13.2.1 istnieje $\lambda \in \mathbb{R}$ takie, że $f = \lambda f_*$, czyli $f(X_0 + X) = \lambda f(X_0 - X)$ dla $X \in \mathbb{R}^n$ i pozostaje wykazać, że $\lambda = 1$. Jeśli funkcja f jest dana wzorem (*) z Definicji 12.1.1, to rachując jak w Uwadze 12.1.7, dostajemy $X^T A X + (2X_0^T A + B^T)X + c' = (-X)^T \lambda A (-X) + \lambda(2X_0^T A + B^T)(-X) + \lambda c'$, a ponieważ f wyznacza swoją część kwadratową, $A = \lambda A$ i z $A \neq \mathbf{0}$ wnioskujemy, że $\lambda = 1$. ■

Wniosek 13.2.7 Niech H będzie hiperpowierzchnią właściwą w przestrzeni afinicznej E nad \mathbb{R} .

- (I)₀ Jeśli H jest opisana równaniem kanonicznym typu (AI)₀, to H ma środek symetrii i zawiera wszystkie swoje środki symetrii.
- (I)₁ Jeśli H jest opisana równaniem kanonicznym typu (AI)₁, to H ma środek symetrii i nie zawiera żadnego swojego środka symetrii.
- (II) Jeśli H jest opisana równaniem kanonicznym typu (AII), to H nie ma środków symetrii.

Na zakończenie tej części podamy kilka uwag dotyczących hiperpowierzchni właściwych w przestrzeni euklidesowej E .

Podobnie jak w Twierdzeniu 13.1.2, dowodzi się, że każda taka hiperpowierzchnia jest opisana w pewnym prostokątnym układzie współrzędnych τ_q równaniem $f(Z) = 0$, gdzie f jest jedną z funkcji występujących w Twierdzeniu 12.2.4. Z Twierdzenia 13.2.1 wynika, że takie równanie jest wyznaczone z dokładnością do stałej $\lambda \in \mathbb{R}$.

Uznając dwa takie proporcjonalne równania za równoważne, dowodzi się, podobnie jak w Twierdzeniu 13.2.4, następujące twierdzenie.

Twierdzenie 13.2.8 Hiperpowierzchnie właściwe H, H' w przestrzeni euklidesowej E są izometryczne wtedy i tylko wtedy, gdy w pewnych prostokątnych układach współrzędnych są opisane równoważnymi równaniami.

14 Uzupełnienia

14.1 Twierdzenie Weierstrassa.

W tej części wyjaśnimy pominięte w tekście elementy dowodów związane z twierdzeniem Weierstrassa o istnieniu ekstremów funkcji ciągłych.

Zacznijmy od Lematu 6.2.3 użytego w dowodzie zasadniczego twierdzenia algebry.

Lemat 14.1.1 *Dla wielomianu zespolonego $w(z) = a_0 + a_1z + \dots + z^n$ i $K = \{z \in \mathbb{C} : |\operatorname{Re}z|, |\operatorname{Im}z| \leq M\}$ istnieje $z_0 \in K$ takie, że $|w(z_0)| = \inf \{|w(z)| : z \in K\}$.*

Dowód. Zacznijmy od warunku (ustalającego ciągłość funkcji $|w(z)|$ zmiennej zespolonej). Dla każdego ciągu $(z_m)_m$ liczb zespolonych

$$(*) \quad \lim_m |z_m - z_0| = 0 \Rightarrow \lim_m |w(z_m)| = |w(z_0)|.$$

Istotnie, dla $h \in \mathbb{C}$ mamy $w(z_0 + h) - w(z_0) = A_1h + \dots + A_nh^n$, stąd $||w(z_0 + h)| - |w(z_0)|| \leq |w(z_0 + h) - w(z_0)| \leq |h|(|A_1| + \dots + |A_n||h|^{n-1})$.

Podstawiając w tej nierówności za h różnicę $z_m - z_0$ dostajemy (*).

Niech

$$\mu = \inf \{|w(z)| : z \in K\}.$$

Dla $m = 1, 2, \dots$ wybierzmy $z_m = a_m + ib_m \in K$ tak, by ciąg $(|w(z_m)|)_m$ zbiegał do μ . Z twierdzenia Bolzano-Weierstrassa możemy założyć, przechodząc w razie potrzeby do podciągu, że ciąg $(a_m)_m$ jest zbieżny do pewnego a_0 z przedziału $[-M, M]$. Przechodząc ponownie do podciągu możemy ponadto założyć, że ciąg $(b_m)_m$ jest zbieżny do pewnego $b_0 \in [-M, M]$. Dla $z_0 = a_0 + ib_0$ mamy $|z_m - z_0| = \sqrt{(a_m - a_0)^2 + (b_m - b_0)^2}$, więc $\lim_m |z_m - z_0| = 0$, a stąd i z (*) dostajemy $\mu = \lim_m |w(z_m)| = |w(z_0)|$. ■

W dowodzie Twierdzenia 8.1.4 korzystaliśmy z pewnej własności funkcji na sferze jednostkowej w liniowej przestrzeni euklidesowej, analogicznej do udowodnionej w Lemacie 14.1.1. Podamy teraz uzasadnienie tej własności, przy czym podobieństwo do dowodu Lematu 14.1.1 nie jest przypadkowe – obie własności wynikają z twierdzenia Weierstrassa o funkcjach ciągłych na zwartych przestrzeniach metrycznych.

Lemat 14.1.2 *Dla endomorfizmu $T : V \rightarrow V$ liniowej przestrzeni euklidesowej $(V, \langle \cdot, \cdot \rangle)$ istnieje wektor $w \in V$ taki, że $\|w\| = 1$ oraz $\langle w, T(w) \rangle = \sup \{\langle v, T(v) \rangle : \|v\| = 1\}$.*

Dowód. Połóżmy $Q(v) = \langle v, T(v) \rangle$ dla $v \in V$. Sprawdźmy najpierw, że dla każdego ciągu wektorów $(v_m)_m$ takich, że $\|v_m\| = 1$

$$(*) \quad \lim_m \|v_m - w\| = 0 \Rightarrow \lim_m Q(v_m) = Q(w).$$

Ustalmy bazę ortonormalną (u_1, \dots, u_n) w V , połóżmy $M = \sum_j \|T(u_j)\|$ i niech $u = \sum_j a_j u_j \in V$. Mamy

$$\|T(u)\| = \|T(\sum_j a_j u_j)\| \leq \sum_j |a_j| \|T(u_j)\| \leq \|u\| M.$$

Warunek (*) wynika teraz z oszacowania

$$\begin{aligned} |Q(v_m) - Q(w)| &= |\langle v_m, T(v_m) \rangle - \langle w, T(w) \rangle| \leq |\langle v_m, T(v_m) \rangle - \langle w, T(v_m) \rangle| + |\langle w, T(v_m) \rangle - \langle w, T(w) \rangle| = \\ &= |\langle v_m - w, T(v_m) \rangle| + |\langle w, T(v_m) - T(w) \rangle| \leq \|v_m - w\| \|T(v_m)\| + \|w\| \|T(v_m - w)\| \leq \\ &\quad \|v_m - w\| M \|v_m\| + \|w\| M \|v_m - w\|. \end{aligned}$$

Niech teraz

$$\mu = \sup \{Q(v) : \|v\| = 1\}.$$

Dla $m = 1, 2, \dots$ wybierzmy wektory v_m o normie 1 tak, by ciąg $(Q(v_m))_m$ zbiegał do μ . Podobnie jak w dowodzie Lematu 14.1.1 wystarczy teraz pokazać, że dla pewnego podciągu $(v_{m_k})_k$ istnieje $w \in V$ takie, że $\lim_k \|v_{m_k} - w\| = 0$, bo z warunku $|\|v_m\| - \|w\|| \leq \|v_m - w\|$ wynika, że wtedy także $\|w\| = 1$.

Udowodnimy przez indukcję ze względu na $\dim V$, że dla każdego ciągu wektorów V o normie ≤ 1 można znaleźć taki podciąg $(v_{m_k})_k$ i wektor w . Dla przestrzeni jednowymiarowych teza wynika z twierdzenia Bolzano-Weierstrassa. Załóżmy, że $\dim V > 1$ i teza jest prawdziwa dla przestrzeni wymiaru mniejszego niż $\dim V$.

Niech $V = W \oplus L$ będzie sumą ortogonalną, gdzie $\dim L = 1$ i niech $v_m = w_m + u_m$ będzie rozkładem v_m na ortogonalne składowe. Z założenia indukcyjnego wynika, że ciąg $(w_m)_m$ zawiera podciąg $(w_{m_j})_j$ taki, że $\lim_j \|w_{m_j} - w_0\| = 0$ dla pewnego $w_0 \in W$. Z twierdzenia Bolzano-Weierstrassa możemy wybrać z ciągu $(u_{m_j})_j$ podciąg $(u_{m_{j_k}})_{j_k}$ zbieżny do pewnego $u_0 \in L$. Dla wektora $w = w_0 + u_0$ mamy wówczas $\lim_k \|v_{m_{j_k}} - w\| = 0$. ■

14.2 Zespólona wartość własna macierzy rzeczywistej.

Dla macierzy $A \in \mathbb{R}^n$ oznaczmy przez T endomorfizm $T(X) = AX$ przestrzeni \mathbb{R}^n wyznaczony przez A i załóżmy, że $Z = X + iY \in \mathbb{C}^n$, gdzie $X, Y \in \mathbb{R}^n$, jest wektorem własnym A odpowiadającym zespolonej wartości własnej $\lambda = a + ib \notin \mathbb{R}$. Pokażemy, że

- (a) wektory X, Y rozpinają w \mathbb{R}^n płaszczyznę $W = \text{lin}(X, Y)$ taką, że $T(W) = W$ i macierz endomorfizmu $T|_W : W \rightarrow W$ w układzie współrzędnych związanym z bazą (X, Y) płaszczyzny W ma postać $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$;
- (b) $A^T \neq A$, por. Twierdzenie 8.1.4;
- (c) jeśli A jest macierzą ortogonalną, to $|\lambda| = 1$ i w kartezjańskiej przestrzeni euklidesowej, zob. 7.1.2 (a), $\|X\| = \|Y\|$ oraz $X \perp Y$. W szczególności, $T|_W$ jest obrotem W , por. Lemat 8.2.6.

Istotnie, z równości $AZ = \lambda Z$ mamy $AX + iAY = (a + ib)(X + iY) = (aX - bY) + i(bX + aY)$. Porównując części rzeczywiste i urojone po obu stronach dostajemy stąd wzory

$$(*) \quad AX = aX - bY \quad \text{oraz} \quad AY = bX + aY$$

pokazujące, że $T(W) \subset W$.

Położmy $\bar{Z} = X - iY$. Z (*) mamy $A\bar{Z} = AX - iAY = (aX - bY) - i(bX + aY) = (a - ib)(X - iY)$, czyli wektor \bar{Z} jest wektorem własnym A odpowiadającym wartości własnej $\bar{\lambda}$. Zatem wektory Z, \bar{Z} rozpinają płaszczyznę $\text{lin}(Z, \bar{Z}) \subset \mathbb{C}^n$. Wektory $X = \frac{1}{2}(Z + \bar{Z})$ i $Y = -\frac{i}{2}(Z - \bar{Z})$ są bazą tej płaszczyzny (nad \mathbb{C}), więc są również liniowo niezależne nad \mathbb{R} .

Stąd wynika (a), bo (*) określa postać macierzy $T|_W$ w tej bazie, a z $\det(T|_W) = a^2 + b^2 = |\lambda| > 0$ mamy $T(W) = W$.

Dla dowodu (b) załóżmy, że macierz A jest symetryczna. Wtedy $(AZ)^T \bar{Z} = Z^T (A\bar{Z})$, więc $\lambda(Z^T \bar{Z}) = (AZ)^T \bar{Z} = Z^T (A\bar{Z}) = Z^T (\bar{\lambda} \bar{Z}) = \bar{\lambda}(Z^T \bar{Z})$. Ponieważ $Z^T \bar{Z} = (X + iY)^T (X - iY) = \|X\|^2 + \|Y\|^2 > 0$, dostajemy $\lambda = \bar{\lambda}$, sprzecznie z założeniem, że $\lambda \notin \mathbb{R}$.

Jeśli A jest macierzą ortogonalną, to T jest izometrią liniową kartezjańskiej przestrzeni euklidesowej, więc $\det(T|_W) = 1 = |\lambda|$, co daje pierwszą część (c). Mamy też $AZ = \lambda Z$ oraz $A^T Z = A^{-1} Z = \lambda^{-1} Z$. Zatem $\lambda(Z^T Z) = (AZ)^T Z = Z^T (A^T Z) = Z^T (\lambda^{-1} Z) = \bar{\lambda}(Z^T Z)$. Ponieważ $\lambda \neq \bar{\lambda}$, dostajemy stąd $Z^T Z = 0$, czyli $0 = (X + iY)^T (X + iY) = (\|X\|^2 - \|Y\|^2) + 2i(X^T Y)$, co dowodzi (c).

14.3 Orientacja liniowej przestrzeni euklidesowej.

Powiemy, że macierz $A \in \mathbb{R}^n$ jest *necosobliwa* jeśli $\det A \neq 0$ (równoważnie, macierz A jest odwracalna, zob. Uwaga 5.1.4).

Drogą nieosobliwą w \mathbb{R}_n^n łączącą macierze nieosobliwe $A, B \in \mathbb{R}_n^n$ będziemy nazywali przyporządkowanie każdemu t z przedziału $[0, 1] \subset \mathbb{R}$ macierzy nieosobliwej $A(t) = [a_{ij}(t)]_{ij=1}^n$ takie, że $A(0) = A$, $A(1) = B$ i każda z funkcji a_{ij} jest ciągła.

Funkcja $\det A(t)$ jest wtedy ciągła i nie zeruje się na $[0, 1]$, więc $\det A$, $\det B$ mają ten sam znak.

Pokażemy, że z równości znaków wyznaczników wynika istnienie takiej drogi: jeśli $\det A > 0$, to istnieje nieosobliwa droga łącząca macierz A z macierzą jednostkową I_n ; jeśli dodatkowo macierz A jest ortogonalna, to istnieje taka droga złożona z macierzy ortogonalnych.

Założmy najpierw, że A jest macierzą ortogonalną. Z Wniosku 8.2.5 istnieje macierz ortogonalna C taka, że $J = C^T A C$ ma na przekątnej macierze obrotów i jedynki (2 wyrazy -1 na przekątnej J określają obrót o kąt π w odpowiedniej płaszczyźnie). Ponieważ macierz obrotu \mathbb{R}^2 o kąt θ można połączyć (mnożąc kąt obrotu przez $1-t$) z macierzą I_2 , istnieją macierze ortogonalne $J(t)$ tworzące drogę łączącą J z I_n . Macierze $CJ(t)C^T$ są wtedy ortogonalne i określają drogę łączącą A z I_n .

Dla odwracalnej macierzy A nietrudno zauważyć, interpretując geometrycznie kolejne kroki procesu ortonormalizacji Grama-Schmidta układu kolumn A , że powstającą w tym procesie macierz ortogonalną można połączyć drogą nieosobliwą z macierzą A . Wraz z wcześniejszą obserwacją daje to nieosobliwą drogę łączącą macierz A z I_n .

Istnienie takich dróg pozwala na następującą interpretację orientacji w liniowej przestrzeni euklidesowej \mathbb{R}^n : baza (A_1, \dots, A_n) jest zorientowana zgodnie z bazą standardową (E_1, \dots, E_n) wtedy i tylko wtedy, gdy dla każdego t z przedziału $[0, 1]$ istnieje baza $(A_1(t), \dots, A_n(t))$ taka, że $A_j(0) = A_j$, $A_j(1) = E_j$ oraz współrzędne wektorów $A_j(t)$ zależą od t w sposób ciągły. Co więcej, dla ortonormalnej bazy (A_1, \dots, A_n) istnieją takie bazy $(A_1(t), \dots, A_n(t))$, które są dodatkowo ortonormalne.

14.4 Równania opisujące hiperpowierzchnię właściwą.

W tej części podamy dowód Twierdzenia 13.2.1 mówiącego, że hiperpowierzchnia właściwa H stopnia 2 w przestrzeni afinicznej E nad \mathbb{R} wyznacza swoje równanie z dokładnością do stałej.

Założmy więc, że $f, f_* : \mathbb{R}^n \rightarrow \mathbb{R}$ są funkcjami kwadratowymi takimi, że równania $f(X) = 0$ i $f_*(X) = 0$ opisują H w afinicznym układzie współrzędnych $\sigma_p : E \rightarrow \mathbb{R}^n$. Dla uproszczenia oznaczeń połączmy $\phi = f \circ \sigma_p$ i $\phi_* = f_* \circ \sigma_p$. Pokażemy, że $\phi_* = \lambda\phi$ dla pewnego $\lambda \in \mathbb{R}$.

Niech funkcja f będzie dana wzorem

$$f(X) = X^T A X + B^T X + c, \text{ gdzie } A = A^T \in \mathbb{R}_n^n, B \in \mathbb{R}^n, c \in \mathbb{R}$$

i oznaczymy przez Q część kwadratową f , $Q(X) = X^T A X$.

Jeśli $L = q + \text{lin}(w)$ jest prostą w E , $X = \sigma_p(q)$ i $Z = \sigma(w)$, to $\sigma_p(q + tw) = \sigma_p(q) + \sigma(tw) = X + tZ$, więc wartości funkcji ϕ na prostej L są dane wzorem $\phi(q + tw) = f(X + tZ) = t^2 Z^T A Z + t(2X^T A + B^T)Z + f(X)$ (zob. Uwaga 12.1.3 dla $g(Z) = tZ + X$). Zatem

$$(\curvearrowright) \quad \phi(q + tw) = at^2 + bt + c, \text{ gdzie } a = Q \circ \sigma(w) \text{ i } c = \phi(q).$$

Zgodnie z Uwagą 13.1.5, istnieje prosta $L_0 = q_0 + \text{lin}(w_0)$ przecinająca H w dokładnie dwóch punktach. Z (\curvearrowright) wynika, że $a = Q(\sigma(w_0)) \neq 0$. Bez zmniejszenia ogólności możemy założyć, że $a > 0$ (zastępując w razie potrzeby ϕ przez $-\phi$ oraz $\phi(q_0) < 0$ (przesuwając w razie potrzeby punkt początkowy q_0 na L_0).

Położmy $\lambda = \frac{\phi_*(q_0)}{\phi(q_0)}$ i niech $E_* = \{q \in E : \phi_*(q) = \lambda\phi(q)\}$. Oczywiście $H \cup \{q_0\} \subset E_*$. Z (\curvearrowright) dla funkcji ϕ_* i funkcji $\lambda\phi$ wynika, że każda prosta w E zawierająca trzy punkty zbioru E_* zawiera się w E_* .

Niech $L = q_0 + \text{lin}(w)$ będzie prostą taką, że $Q(\sigma(w)) > 0$. Z (\curvearrowright) dla ϕ na L wnioskujemy, że zbiór $L \cap H$ jest dwupunktowy, a $q_0 \in L \setminus H$ daje wtedy $L \subset E_*$. Pokazaliśmy więc, że $\{q_0 + w : Q(\sigma(w)) > 0\} \subset E_*$.

Weźmy teraz dowolne $q = q_0 + u \in E$. Punkty prostej $L_q = q + \text{lin}(w_0)$ możemy przedstawić w postaci $q_0 + (u + tw_0)$. Powtarzając uzasadnienie (\curvearrowright) widzimy, że $Q(\sigma(u + tw_0)) = t^2 Q(\sigma(w_0)) + tb + Q(\sigma(u)) > 0$ dla dostatecznie dużych t (bo $Q(\sigma(w_0)) > 0$). Dla takich t punkty $q_0 + (u + tw_0)$ prostej L_q są w E_* , więc cała prosta L_q jest w E_* i stąd $q \in L_q \subset E_*$. Mamy więc $E_* = E$, co daje tezę.

Skorowidz

- śląd
 - endomorfizmu, 31
 - macierzy, 30
- środek symetrii
 - funkcji, 65
 - hiperpowierzchni, 70
- anihilator formy kwadratowej, 61
- baza
 - dualna, 23
 - Jordana, 35
 - ortogonalna, 38
 - ortogonalna w (V, Q) , 61
 - ortonormalna, 38
 - przestrzeni liniowej, 11
 - punktowa, 52
 - sprzężona, 23
 - standardowa w \mathbb{K}^n , 11
- ciało, 7
 - \mathbb{Z}_p , 7
 - liczb zespolonych, 5
- część liniowa przekształcenia afinicznego, 53
- część wspólna podprzestrzeni, 14
- dopełnienie
 - ortogonalne, 39
 - podprzestrzeni, 15
- endomorfizm, 30
 - diagonalizowalny, 32
 - nilpotentny, 33
 - samosprzężony, 44
 - sprzężony, 48
- epimorfizm, 17
- forma kwadratowa, 57
 - dodatnio określona, 63
 - ujemnie określona, 63
- formuła
 - de Moivre'a, 6
 - Grassmana, 15
 - Pitagorasa, 38
- funkcja kwadratowa, 64
- funkcjonał
 - dwuliniowy symetryczny, 57
 - liniowy, 22
- hiperpłaszczyzna
 - w \mathbb{K}^n , 49
 - w przestrzeni afinicznej, 51
- hiperpowierzchnia
 - stopnia 2, 67
 - właściwa, 68
- iloczyn
 - macierzy, 19
 - macierzy przez wektor, 10
 - skalarny, 37
 - wektorowy, 42
- izometria, 56
 - przestrzeni euklidesowej, 55
 - liniowa, 45
- izomorfizm
 - afiniczny, 53
 - afiniczny na \mathbb{K}^n , 49
 - elementarny, 21
 - liniowy, 18
 - zachowujący orientację, 42
- jądro przekształcenia liniowego, 17
- jednokładność, 53
- kąt między wektorami, 38
- klatka Jordana, 35
- kombinacja
 - afiniczna, 52
 - liniowa, 9
- kryterium Sylwestera, 63
- macierz
 - diagonalizowalna, 32
 - diagonalna, 21
 - elementarna, 21
 - górnio trójkątna, 28
 - Gram, 39
 - jednostkowa, 19
 - Jordana, 35
 - kwadratowa, 20
 - niesobliwa, 72
 - odwracalna, 20
 - odwrotna, 20
 - ortogonalna, 45
 - przekształcenia, 19
 - przekształcenia w układach σ, τ , 22
 - rozszerzona układu równań, 2
 - stowarzyszona, 29
 - symetryczna, 44
 - transponowana, 23
 - w postaci schodkowej, 3
- macierze

- kongruentne, 58
 - podobne, 31
- minor wiodący, 61
- monomorfizm, 17
- nierówność
 - Schwarza, 37
 - trójkąta, 37
- niezależność
 - afiniczna, 52
 - liniowa, 10
- norma wektora, 37
- objętość układu wektorów, 41
- odległość
 - podprzestrzeni, 55
 - punktów, 55
 - punktu od podprzestrzeni, 55
- operacje elementarne na wierszach, 2
- orientacja, 42
- ortogonalizacja Grama-Schmidta, 38
- ortogonalne podprzestrzenie, 38
- permutacja, 28
- podprzestrzeń
 - afiniczna, 51
 - liniowa, 9
 - niezmiennicza, 33
- powłoka
 - afiniczna, 52
 - liniowa, 9
- prosta
 - afiniczna, 51
 - w \mathbb{K}^n , 49
- przekształcenie
 - afiniczne, 53
 - afiniczne na \mathbb{K}^n , 49
 - defekt, 17
 - jądro, 17
 - liniowe, 16
 - obraz, 17
 - rząd, 17
 - sprężone, 23
- przestrzeń
 - $\mathbb{K}[x]$ wielomianów, 9
 - \mathbb{K}^m współrzędnych, 8
 - \mathbb{K}_n^m macierzy, 9
 - afiniczna, 50
 - euklidesowa, 55
 - kolumn macierzy, 13
 - liniowa, 8
 - liniowa euklidesowa, 37
 - przekształceń, 20
 - sprężona, 22
 - wierszy macierzy, 14
 - zerowa macierzy, 13
- przesunięcie, 53
- równanie kanoniczne hiperpowierzchni, 68
- równoległoscian, 41
- równoważność
 - afiniczna funkcji, 65
 - euklidesowa funkcji, 66
- rząd
 - funkcjonału kwadratowego, 58
 - macierzy, 13
- rzut
 - afiniczny, 53
 - na podprzestrzeń, 17
 - ortogonalny, 39
 - ortogonalny przestrzeni euklidesowej, 55
- składowe wektora, 14
- spektrum
 - endomorfizmu, 31
 - macierzy, 30
- suma
 - algebraiczna podprzestrzeni, 14
 - ortogonalna, 38
 - prosta podprzestrzeni, 14
- sygnatura, 62
- symetria
 - afiniczna, 53
 - ortogonalna, 39
 - ortogonalna przestrzeni euklidesowej, 55
 - względem podprzestrzeni, 17
- twierdzenie
 - Cauchy'ego, 27
 - cosinusów, 38
 - Kroneckera – Capelliego, 13
 - o istnieniu bazy, 12
 - o określaniu przekształceń liniowych, 16
 - o równoważności układów równań, 2
 - o redukcji wierszowej, 3
 - o rozszerzaniu do bazy, 12
 - o rozwinięciu Laplace'a względem kolumny, 27
 - o rozwinięciu Laplace'a względem wiersza, 28
 - Steinitza o wymianie, 12
 - Sylwestera o bezwładności, 62
- typ afiniczny hiperpowierzchni, 69
- układ
 - bazowy, 52
 - ortogonalny, 38
 - rozpinający, 10

- układ równań
 - Cramera, 29
 - jednorodny, 1
 - równoważny, 2
 - sprzeczny, 2
- układ współrzędnych, 22
 - afiniczny, 54
 - prostokątny, 43
 - prostokątny afiniczny, 55
- warstwa podprzestrzeni, 49
- wartości własne, 30
- wektor
 - własny endomorfizmu, 31
 - własny macierzy, 30
 - zerowy, 1
- wektory
 - izotropowe formy kwadratowej, 61
 - ortogonalne, 38
 - ortogonalne w (V, Q) , 61
- wielomian charakterystyczny
 - endomorfizmu, 31
 - macierzy, 30
- współrzędne
 - barycentryczne, 52
 - punktu, 52
 - wektora, 11
- wymiar
 - przestrzeni afinicznej, 50
 - przestrzeni liniowej, 12
 - przestrzeni wierszy, 14
 - przestrzeni zerowej macierzy, 13
 - sumy prostej, 15
 - warstwy, 49
- wyznacznik
 - endomorfizmu, 31
 - Gramma, 39
 - macierzy, 25
- wzór Sarrusa, 27
- wzory Cramera, 29

- zmiennie
 - niezależne, 4
 - zależne, 4
- znak permutacji, 28