

Spis treści

1	Przestrzenie i przekształcenia	3
1.1	Przestrzeń liniowa	3
1.2	Przekształcenia liniowe	12
2	Macierze i układy równań	19
2.1	Wyznacznik macierzy	19
2.2	Macierz odwrotna	33
2.3	Układy równań liniowych	43
3	Diagonalizacja macierzy	57
3.1	Wartości i wektory własne (rzeczywiste)	57
3.2	Wartości i wektory własne (zespolone)	67
3.3	Twierdzenie Jordana	79
3.4	Zastosowania diagonalizacji	86
4	Przestrzeń euklidesowa	95
4.1	Iloczyn skalarny	95
4.2	Forma kwadratowa	108
4.3	Twierdzenie spektralne	119
4.4	Rozkład SVD	127

Rozdział 1

Przestrzenie i przekształcenia

1.1 Przestrzeń liniowa

Przestrzeń liniowa (inaczej: *przestrzeń wektorowa*) to dowolny zbiór obiektów, które będziemy nazywać *elementami przestrzeni liniowej* lub *wektorami* (nawet jeśli nie będą miały nic wspólnego z wektorami na płaszczyźnie lub w przestrzeni), wyposażony w dwa działania: dodawanie i mnożenie przez skalar (tj. liczbę rzeczywistą), które to działania mają podobne własności jak dodawanie i mnożenie liczb rzeczywistych. Precyzyjniej:

Definicja 1.1

Przestrzenią liniową (nad $\mathbb{R}$) nazywamy zbiór V z dwoma działaniami:

- dodawaniem elementów V ,
- mnożeniem elementów V przez skalar (liczbę rzeczywistą),

oraz wyróżnionym elementem 0 (elementem neutralnym dodawania) i przypisaniem każdemu elementowi $v \in V$ elementu $-v \in V$ (element przeciwny) tak, że dla dowolnych $u, v, w \in V$ oraz $s, t \in \mathbb{R}$ zachodzą następujące warunki:

1. $u + v = v + u$ (+ jest przemienne)
2. $(u + v) + w = u + (v + w)$ (+ jest łączne)
3. $0 + v = v + 0 = v$ (0 jest elementem neutralnym +)
4. $v + (-v) = (-v) + v = 0$ ($-v$ jest elementem przeciwnym do v)
5. $(s + t) \cdot u = s \cdot u + t \cdot u$ ($\cdot$ jest rozdzielne względem dodawania skalarów)
6. $t \cdot (u + v) = t \cdot u + t \cdot v$ ($\cdot$ jest rozdzielne względem dodawania wektorów)
7. $1 \cdot v = v$

Przykład 1

Przestrzeń $\mathbb{R}^2$ z dodawaniem wektorów i mnożeniem wektorów przez skalar jest przestrzenią liniową. Elementem neutralnym dodawania jest $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, a elementem przeciwnym do $\begin{pmatrix} x \\ y \end{pmatrix}$ jest element $\begin{pmatrix} -x \\ -y \end{pmatrix}$. Podobnie pokazujemy strukturę przestrzeni liniowej na $\mathbb{R}^3$.

Przykład 2

Przestrzeń $\mathbb{R}^n$ układów n liczb rzeczywistych $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ wraz z dodawaniem i mnożeniem przez skalar określonymi następująco:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \quad \alpha \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha \cdot x_1 \\ \vdots \\ \alpha \cdot x_n \end{pmatrix}$$

to przestrzeń liniowa. Element 0 to $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, a element przeciwny do $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ to $\begin{pmatrix} -x_1 \\ \vdots \\ -x_n \end{pmatrix}$.

Przykład 3

Przestrzeń $M_{m \times n}(\mathbb{R})$ macierzy rozmiaru $m \times n$ o wyrazach rzeczywistych (będziemy używać też oznaczenia $M_{m \times n}$) wraz z dodawaniem macierzy i mnożeniem macierzy przez skalar jest przestrzenią liniową. Elementem 0 jest macierz $\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$, a elementem przeciwnym do $\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ jest $\begin{pmatrix} -a_{11} & \cdots & -a_{1n} \\ \vdots & \ddots & \vdots \\ -a_{m1} & \cdots & -a_{mn} \end{pmatrix}$.

Przykład 4

Przestrzeń wielomianów $\mathbb{R}_n[x]$ stopnia co najwyżej n o współczynnikach rzeczywistych z działaniami dodawania wielomianów i mnożenia wielomianów przez skalar oraz przestrzeń $\mathbb{R}[x]$ wszystkich wielomianów o współczynnikach rzeczywistych. Elementem 0 jest wielomian zerowy (tzn. wielomian o wszystkich współczynnikach równych zero), a elementem przeciwnym do wielomianu $a_k x^k + \cdots + a_1 x + a_0$ jest wielomian $-a_k x^k - \cdots - a_1 x - a_0$.

Przykład 5

Przestrzeń funkcji ciągłych $C(\mathbb{R})$ o dziedzinie $\mathbb{R}$ z działaniem dodawania funkcji i mnożenia funkcji przez skalar:

$$(f + g)(x) = f(x) + g(x) \quad (\alpha \cdot f)(x) = \alpha \cdot f(x)$$

Elementem 0 jest funkcja zerowa, tj. funkcja f taka, że $f(x) = 0$ dla każdego $x \in \mathbb{R}$. Elementem przeciwnym do funkcji f jest funkcja $-f$, gdzie $(-f)(x) = -f(x)$. Podobnie definiujemy:

- przestrzeń $C^1(\mathbb{R})$ funkcji różniczkowalnych o ciągłej pochodnej,
- przestrzeń $C^2(\mathbb{R})$ funkcji dwukrotnie różniczkowalnych o ciągłej drugiej pochodnej,
- przestrzeń $C^\infty(\mathbb{R})$ funkcji różniczkowalnych dowolną liczbę razy.

Przykład 6

Przestrzeń wszystkich ciągów o wyrazach rzeczywistych z dodawaniem ciągów i mnożeniem ciągów przez skalar. Podobnie:

- przestrzeń ℓ^∞ wszystkich ciągów ograniczonych o wyrazach rzeczywistych,
- przestrzeń c wszystkich ciągów zbieżnych (o wyrazach rzeczywistych),
- przestrzeń c_0 wszystkich ciągów zbieżnych do 0 (o wyrazach rzeczywistych).

Przykład 7

Zbiór $\mathbb{C}$ można traktować jako przestrzeń liniową (nad liczbami rzeczywistymi).

Definicja 1.2

Podprzestrzenią W przestrzeni liniowej V (ozn. $W < V$) nazywamy taki **niepusty** podzbiór $W \subset V$, który jest:

- 1) zamknięty na dodawanie, tzn. dla dowolnych $u, v \in W$ zachodzi $u + v \in W$ oraz
- 2) zamknięty na mnożenie przez skalar, tzn. dla dowolnych $u \in W, t \in \mathbb{R}$ zachodzi $tu \in W$.

Podprzestrzeń przestrzeni liniowej sama również jest przestrzenią liniową (z działaniami zdefiniowanymi tak samo jak dla V).

Fakt 1.3

Podprzestrzeń $W < V$ zawsze zawiera element 0.

Dowód. Weźmy dowolny wektor $v \in W$ (z definicji W jest niepusty) oraz skalar $0 \in \mathbb{R}$. Zgodnie z definicją podprzestrzeni element $0 \cdot v = 0 \in W$. $\square$

Przykład 8

Kompletna lista podprzestrzeni $\mathbb{R}^2$ to:

(a) zbiór jednoelementowy $\{0\}$, (b) dowolna prosta przechodząca przez 0, (c) całe $\mathbb{R}^2$.

Przykład 9

Kompletna lista podprzestrzeni $\mathbb{R}^3$ to:

(a) zbiór jednoelementowy $\{0\}$, (b) dowolna prosta przechodząca przez 0, (c) dowolna płaszczyzna przechodząca przez 0, (d) całe $\mathbb{R}^3$.

Przykład 10

Dla dowolnego n zachodzi: $\mathbb{R}_n[x] < \mathbb{R}[x]$. Podobnie $\mathbb{R}_m[x] < \mathbb{R}_n[x]$, o ile $m \leq n$. Natomiast zbiór wielomianów stopnia *dokładnie* n nie jest podprzestrzenią $\mathbb{R}_n[x]$.

Przykład 11

Dla dowolnej przestrzeni liniowej V zachodzi $\{0\} < V$ oraz $V < V$.

Przykład 12

Każdy z następujących podzbiorów zbioru $M_{3 \times 3}$ jest podprzestrzenią:

- zbiór macierzy diagonalnych rozmiaru 3×3 ,
- zbiór macierzy symetrycznych rozmiaru 3×3 ,
- zbiór macierzy górnotrójkątnych rozmiaru 3×3 ,

natomiast żaden z poniższych podzbiorów nie jest podprzestrzenią $M_{3 \times 3}$:

- zbiór macierzy odwracalnych rozmiaru 3×3 ,
- zbiór macierzy o wyrazach 0 i 1.

Przykład 13

Każdy z następujących podzbiorów zbioru $C(\mathbb{R})$ jest podprzestrzenią:

- $C^1(\mathbb{R}), C^2(\mathbb{R}), \dots, C^\infty(\mathbb{R})$,
- zbiór ograniczonych funkcji ciągłych,
- zbiór funkcji ciągłych mających granicę 0 w $\pm\infty$.

natomiast żaden z poniższych podzbiorów nie jest podprzestrzenią:

- zbiór nieograniczonych funkcji ciągłych,
- zbiór funkcji ciągłych mających granicę 1 w $\pm\infty$.

Przykład 14

Każdy z następujących podzbiorów zbioru wszystkich ciągów jest podprzestrzenią:

- zbiór ℓ^∞ ciągów ograniczonych,
- zbiór c ciągów zbieżnych,
- zbiór c_0 ciągów zbieżnych do 0.

natomiast żaden z poniższych podzbiorów nie jest podprzestrzenią:

- zbiór ciągów rozbieżnych,
- zbiór ciągów zbieżnych do 1.

Intuicyjnie, *wymiarem* przestrzeni liniowej nazywamy liczbę parametrów potrzebnych do opisanie elementu tej przestrzeni, np.

- do opisanie elementu $\mathbb{R}^3$ potrzeba 3 parametrów (współrzędnych),
- do opisanie elementu $M_{3 \times 3}$ potrzeba 9 elementów (wyrazów macierzy),
- do opisanie elementu $\mathbb{R}_5[x]$ potrzeba 6 parametrów (współczynników wielomianu),
- do opisanie elementu $\mathbb{C}$ potrzeba 2 parametrów rzeczywistych (część rzeczywista i urojona).

Formalna definicja wymiaru wymaga wprowadzenia pojęcia *bazy*.

Definicja 1.4

Niech V będzie przestrzenią liniową. *Kombinacją liniową* wektorów $v_1, \dots, v_n \in V$ ze współczynnikami $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ nazywamy wektor:

$$\alpha_1 v_1 + \dots + \alpha_n v_n$$

Bazą przestrzeni liniowej nazywamy taki skończony ciąg wektorów $b_1, \dots, b_n \in V$, że każdy wektor $v \in V$ przedstawia się *jednoznacznie* w postaci kombinacji liniowej:

$$v = \alpha_1 b_1 + \dots + \alpha_n b_n$$

lub taki nieskończony zbiór wektorów $B \subset V$, że każdy wektor $v \in V$ przedstawia się jednoznacznie w postaci kombinacji liniowej skończonej liczby elementów z B .

W niniejszym skrypcie będziemy wykorzystywać pojęcie bazy wyłącznie w odniesieniu do przestrzeni skończenie wymiarowych (tzn. posiadających skończoną bazę).

Definicja wymiaru wymaga następującego twierdzenia, które pozostawiamy bez dowodu:

Twierdzenie 1.5

Każda przestrzeń liniowa ma bazę. Ponadto wszystkie bazy ustalonej przestrzeni liniowej są równoliczne.

Definicja 1.6

Wymiarem przestrzeni liniowej V (ozn. $\dim V$) nazywamy liczbę elementów bazy V .

W świetle Twierdzenia 1.5 powyższa definicja ma sens, gdyż pomimo iż przestrzeń liniowa ma bardzo wiele różnych baz, to wszystkie one mają taką samą liczbę elementów.

Przykład 15

Standardową bazą $\mathbb{R}^3$ nazywamy układ wektorów $\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right)$. Ogólnie: bazą $\mathbb{R}^3$ będzie dowolny układ wektorów (b_1, b_2, b_3) taki, że $\det(b_1, b_2, b_3) \neq 0$.

Przykład 16

Przykładem bazy przestrzeni $M_{2 \times 2}$ jest układ $\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right)$. Nietrudno przekonać się, że każdy element przestrzeni $M_{2 \times 2}$ przedstawia się jednoznacznie w postaci kombinacji liniowej tych wektorów:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Innym przykładem bazy jest układ $\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right)$ (sprawdzenie pozostawiamy czytelnikowi).

Przykład 17

Przykładem bazy przestrzeni $\mathbb{R}_3[x]$ jest układ $(1, x, x^2, x^3)$. Nietrudno pokazać, że każdy element przestrzeni $\mathbb{R}_3[x]$ jest kombinacją liniową tych wektorów:

$$ax^3 + bx^2 + cx + d = d \cdot 1 + c \cdot x + b \cdot x^2 + a \cdot x^3$$

Inny przykład bazy to $(x^3, x^2, x, 1)$ (baza jest ciągiem wektorów, więc zmiana ich kolejności oznacza zmianę bazy). Jeszcze inny przykład to układ $(x^3 + x^2 + x + 1, x^2 + x + 1, x + 1, 1)$ (sprawdzenie pozostawiamy czytelnikowi).

Przykład 18

Przykładem bazy przestrzeni $\mathbb{C}$ jest układ $(1, i)$. Inne przykłady bazy tej przestrzeni to $(1, -i)$, $(1 + i, 1 - i)$.

Przykład 19

Bazą przestrzeni macierzy symetrycznych rozmiaru 2×2 jest układ $\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right)$. Nietrudno się przekonać, że każdy element tej przestrzeni przedstawia się jednoznacznie w postaci kombinacji liniowej podanych wektorów:

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} = a \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + c \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Przykład 20

Bazą przestrzeni (nieskończenie wymiarowej) $\mathbb{R}[x]$ jest układ $(1, x, x^2, x^3, \dots)$. Zauważmy, że (zgodnie z definicją bazy) każdy element $\mathbb{R}[x]$ jest kombinacją liniową pewnego **skończonego** podzbioru bazy.

Definicja 1.7

Współrzednymi wektora $v \in V$ w bazie $\mathcal{B} = (b_1, \dots, b_n)$ nazywamy ciąg współczynników $\alpha_1, \dots, \alpha_n$, gdzie

$$v = \alpha_1 b_1 + \dots + \alpha_n b_n$$

Współrzedne wektora v w bazie $\mathcal{B}$ oznaczamy $[v]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$.

Powyższa definicja wyjaśnia dlaczego baza to **ciąg**, a nie **zbiór** – pozycja wektora w bazie odpowiada pozycji współczynnika w wektorze współrzednych.

Przykład 21 (zbiór za duży na bazę)

...ale można go pomniejszyć do bazy (zbiór generujący).

Dane są wektory $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in \mathbb{R}^3$. Nie stanowią one bazy $\mathbb{R}^3$, gdyż każda baza $\mathbb{R}^3$ składa się z trzech wektorów. Ale można sprawdzić, że każdy element $\mathbb{R}^3$ daje się przedstawić jako kombinację liniową tych czterech wektorów (tzn. wygenerować przy użyciu tych czterech wektorów). Najłatwiej uzasadnić to w następujący sposób:

- wektor bazy standardowej $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ można wygenerować jako $e_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix}$;
- wektor e_1 można wygenerować (wykorzystując otrzymany już e_2) jako $e_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} - 2e_2$;
- wektor e_3 można wygenerować (wykorzystując otrzymany już e_2) jako $e_3 = \begin{pmatrix} 0 \\ 2 \end{pmatrix} - 2e_2$;
- wygenerowawszy wektory e_1, e_2, e_3 możemy następnie wygenerować dowolny wektor $\mathbb{R}^3$.

Wobec tego dowolny wektor $\mathbb{R}^3$ można przedstawić w postaci kombinacji liniowej wyjściowych czterech wektorów, ale przedstawienie to nie będzie jednoznaczne. Oznacza to, że w układzie tym jest „za dużo” wektorów. Na przykład wektor $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ można wygenerować z pozostałych trzech:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{4} \begin{pmatrix} 0 \\ 2 \end{pmatrix} + \frac{3}{4} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

W związku z tym wektor $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ można usunąć, a pozostały układ trzech wektorów wciąż będzie generował całą przestrzeń $\mathbb{R}^3$. Można sprawdzić, że pozostały układ $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in \mathbb{R}^3$ jest już bazą $\mathbb{R}^3$.

Przykład 22 (zbiór za mały na bazę)

...ale można go powiększyć do bazy (zbiór liniowo niezależny).

Dane są wektory $1, 1+x, 1+x^2 \in \mathbb{R}_3[x]$. Każdy wektor, który przedstawia się jako kombinacja liniowa tych wektorów, przedstawia się jednoznacznie – jeżeli:

$$ax^3 + bx^2 + cx + d = \alpha \cdot 1 + \beta \cdot (1+x) + \gamma \cdot (1+x^2)$$

to $\gamma = b, \beta = c, \alpha = d - b - c$ – ale nie każdy wektor z $\mathbb{R}_3[x]$ daje się przedstawić w tej postaci (np. nie jest to możliwe dla wektora x^3). Dołączając wektor x^3 do początkowego układu otrzymujemy układ $1, 1+x, 1+x^2, x^3 \in \mathbb{R}_3[x]$, który – co łatwo sprawdzić – jest bazą.

Powyższe dwa przykłady stanowią motywację do podzielenia warunku bycia bazą (każdy element przedstawia się jednoznacznie w postaci kombinacji liniowej) na dwa warunki:

- każdy element przedstawia się (na co najmniej jeden sposób),
- każdy element przedstawia się na co najwyżej jeden sposób.

Definicja 1.8

Zbiór wektorów $v_1, \dots, v_n \in V$ nazywamy:

- *zbiorem generującym* V , jeśli każdy element $v \in V$ można przedstawić (na co najmniej jeden sposób) w postaci kombinacji liniowej elementów $v_1, \dots, v_n$
- *zbiorem liniowo niezależnym (lnz)*, jeśli każdy element $v \in V$ można przedstawić na co najwyżej jeden sposób w postaci kombinacji liniowej elementów $v_1, \dots, v_n$. Zbiór, który nie jest liniowo niezależny, nazywamy zbiorem *liniowo zależnym (lż)*.

Porównanie Definicji 1.8 z Definicją 1.4 daje natychmiastowy wniosek:

Fakt 1.9

Baza przestrzeni liniowej to zbiór, który jest równocześnie liniowo niezależny i generujący.

Sprawdzanie, że dany zbiór wektorów jest generujący lub liniowo niezależny zazwyczaj wykonujemy odwołując się do następujących charakterystyk:

Fakt 1.10: Charakterystyka zbioru generującego

Zbiór wektorów $v_1, \dots, v_n \in V$ generuje V wtedy i tylko wtedy, gdy można z niego wygenerować (przy pomocy kombinacji liniowych) pewną bazę V .

Dowód. Ze zbioru generującego można wygenerować każdy wektor, w szczególności wektory dowolnej bazy. Jeśli natomiast można wygenerować wektory pewnej bazy, to z nich z kolei (przy użyciu kombinacji liniowej) można już wygenerować dowolny wektor. $\square$

Fakt 1.11: charakterystyka liniowej niezależności

Zbiór wektorów $v_1, \dots, v_n \in V$ jest liniowo niezależny (lnz) wtedy i tylko wtedy, gdy:

$$\forall \alpha_1, \dots, \alpha_n \in \mathbb{R} (\alpha_1 v_1 + \dots + \alpha_n v_n = 0 \Rightarrow \alpha_1 = \dots = \alpha_n = 0) \quad (1.1)$$

(tzn. jedyną zerującą się kombinacją liniową $v_1, \dots, v_n$ to kombinacja o wszystkich współczynnikach równych 0).

Dowód. Ponieważ wektor 0 daje się zawsze przedstawić jako kombinacja liniowa dowolnych wektorów:

$$0 = 0 \cdot v_1 + \dots + 0 \cdot v_n$$

(tzw. kombinacja trywialna), więc jeśli układ $v_1, \dots, v_n$ jest lnz, to innej kombinacji nie ma, czyli spełniony jest warunek (1.1).

Jeśli układ $v_1, \dots, v_n$ jest lż, to istnieje wektor, który przedstawia się na dwa różne sposoby w postaci kombinacji liniowej $v_1, \dots, v_n$, czyli:

$$\alpha_1 v_1 + \dots + \alpha_n v_n = \beta_1 v_1 + \dots + \beta_n v_n$$

gdzie przynajmniej dla jednego indeksu i mamy $\alpha_i \neq \beta_i$. Stąd:

$$(\alpha_1 - \beta_1)v_1 + \cdots + (\alpha_n - \beta_n)v_n = 0$$

czyli istnieje zerująca się kombinacja liniowa $v_1, \dots, v_n$, której nie wszystkie współczynniki są zerami. $\square$

Fakt 1.12: charakteryzacja liniowej zależności

Zbiór wektorów $v_1, \dots, v_n \in V$ jest liniowo zależny (lz) wtedy i tylko wtedy, gdy jeden z tych wektorów jest kombinacją liniową pozostałych.

Dowód. Zgodnie z Faktem 1.11 układ $v_1, \dots, v_n$ jest liniowo zależny wtedy i tylko wtedy, gdy istnieje zerująca się kombinacja liniowa $v_1, \dots, v_n$, która ma przynajmniej jeden niezerowy współczynnik:

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0$$

Przyjmijmy, że $\alpha_1 \neq 0$ (rozumowanie w pozostałych przypadkach jest analogiczne). Wówczas:

$$v_1 = -\frac{\alpha_2}{\alpha_1}v_2 - \cdots - \frac{\alpha_n}{\alpha_1}v_n$$

czyli jeden z wektorów jest kombinacją liniową pozostałych. $\square$

Przykład 23

Sprawdź, czy układ wektorów $1+x, 1+2x-x^2, 1+x^3 \in \mathbb{R}_3[x]$ jest liniowo niezależny.

Rozwiązanie. Rozważmy zerującą się kombinację liniową tych wektorów:

$$\alpha(1+x) + \beta(1+2x-x^2) + \gamma(1+x^3) = 0 \quad \text{czyli} \quad \gamma x^3 - \beta x^2 + (2\beta + \alpha)x + (\alpha + \beta + \gamma) = 0$$

Powyższa równość to równość wielomianów, a zero po prawej stronie oznacza wielomian zerowy. Wielomian jest wielomianem zerowym wtedy i tylko wtedy, gdy wszystkie jego współczynniki są zerami, czyli:

$$\begin{cases} \gamma = 0 \\ -\beta = 0 \\ 2\beta + \alpha = 0 \\ \alpha + \beta + \gamma = 0 \end{cases}$$

Stąd otrzymujemy $\alpha = \beta = \gamma = 0$, czyli jedyna zerująca się kombinacja liniowa to kombinacja o wszystkich współczynnikach równych 0. Badany układ wektorów jest więc liniowo niezależny.

Opisana w Przykładach 21 i 22 procedura to dość typowy sposób wyznaczania bazy:

- weź dowolny zbiór generujący i zmniejsz go do bazy (wyrzucając, po jednym, elementy, które są kombinacją liniową pozostałych) lub
- weź dowolny zbiór liniowo niezależny i powiększ go do bazy (dodając, po jednym, elementy, które nie są kombinacją liniową pozostałych).

Fakt 1.13

Niech V będzie skończenie wymiarową przestrzenią liniową. Wówczas:

- 1) z każdego (skończonego) zbioru generującego V można wybrać bazę, usuwając z niego (po jednym) elementy będące kombinacją liniową pozostałych
[tzn. baza to minimalny zbiór generujący]
- 2) Każdy zbiór liniowo niezależny w V można uzupełnić do bazy, dodając do niego (po jednym) elementy, które nie są kombinacją liniową pozostałych
[tzn. baza to maksymalny zbiór liniowo niezależny]

Fakt 1.14

Niech V będzie skończenie wymiarową przestrzenią liniową. Wówczas każda podprzestrzeń $W < V$ spełnia warunek $\dim W \leq \dim V$, przy czym równość zachodzi wtedy i tylko wtedy, gdy $W = V$.

Dowód. Baza podprzestrzeni W (czyli układ wektorów liniowo niezależnych i generujących W) potraktowana jako układ wektorów V przestaje już być zbiorem generującym (chyba, że $W = V$), ale pozostaje zbiorem liniowo niezależnym. Zgodnie z Faktem 1.13 można go więc rozszerzyć do bazy V , skąd $\dim V \geq \dim W$. Równość zachodzi wtedy i tylko wtedy, gdy baza W jest równocześnie bazą V , czyli $W = V$. $\square$

Wniosek 1.15

Jeśli V jest skończenie wymiarową przestrzenią liniową, to układ wektorów $(v_1, \dots, v_n)$ przestrzeni V , który spełnia dowolne dwa z poniższych warunków jest bazą V i spełnia również trzeci z warunków:

- 1) wektory $v_1, \dots, v_n$ są lnz,
- 2) wektory $v_1, \dots, v_n$ generują V ,
- 3) $n = \dim V$.

Dowód. Zgodnie z Faktem 1.9 baza to układ wektorów spełniający warunki (1) i (2). Dla skończenie wymiarowych przestrzeni baza spełnia również warunek (3) (zgodnie z definicją wymiaru). Pozostaje pokazać, że układ spełniający warunki (1) i (3) lub spełniający warunki (2) i (3) również jest bazą.

Niech układ $(v_1, \dots, v_n)$ generuje V (warunek (2)) oraz $n = \dim V$ (warunek (3)). Zgodnie z Faktem 1.13 z układu tego można wybrać bazę V , ale nie może to powodować zmniejszenia liczby wektorów (bo $n = \dim V$). Wobec tego układ ten musi być bazą.

Rozważmy teraz sytuację, gdy układ $(v_1, \dots, v_n)$ jest liniowo niezależny V (warunek (1)) oraz $n = \dim V$ (warunek (3)). Zgodnie z Faktem 1.13 układ ten można uzupełnić do bazy V , ale nie może to powodować zwiększenia liczby wektorów (bo $n = \dim V$). Wobec tego układ ten musi być bazą. $\square$

Fakt 1.16

Jeśli $W_1, W_2 < V$ są podprzestrzeniami liniowymi, to $W_1 \cap W_2$ też jest podprzestrzenią V .

Dowód. Wystarczy sprawdzić, że $W_1 \cap W_2$ jest zamknięty na dodawanie wektorów i mnożenie przez skalar. Rozważmy wektory $w, w' \in W_1 \cap W_2$. Wobec tego $w, w' \in W_1$, a ponieważ W_1 jest podprzestrzenią V , więc $w + w' \in W_1$. Podobnie $w + w' \in W_2$, a stąd dostajemy $w + w' \in W_1 \cap W_2$, czyli zamkniętość na dodawanie. Analogicznie dowodzimy zamkniętości na mnożenie przez skalar. $\square$

1.2 Przekształcenia liniowe

Definicja 1.17

Niech V i W będą przestrzeniami liniowymi. Przekształcenie $F : V \rightarrow W$ nazywamy *przekształceniem liniowym* (inaczej: *homomorfizmem*), jeśli dla dowolnych $u, v \in V$ oraz $t \in \mathbb{R}$ spełnione są warunki:

1. $F(u + v) = F(u) + F(v)$ (addytywność)
2. $F(t \cdot u) = t \cdot F(u)$ (jednorodność)

Zbiór wszystkich przekształceń liniowych z V do W oznaczamy $\text{Hom}(V, W)$.

Przypominamy, że terminy *funkcja*, *przekształcenie*, *odwzorowanie*, *operator* to synonimy, aczkolwiek istnieją pewne ustalone konwencje dotyczące ich używania (np. gdy dziedziną jest $\mathbb{R}$ zwykle używamy terminu *funkcja*, a gdy dziedziną jest przestrzeń funkcji, jak $C(\mathbb{R})$, to zwykle używamy terminu *operator*). Termin *homomorfizm* jest synonimem terminu *przekształcenie liniowe*.

Przykład 1

Każde przekształcenie $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$ zadane wzorem:

$$F(X) = AX \quad \text{czyli} \quad F\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

gdzie $A \in M_{m \times n}$ (mnożenie macierzy było w pierwszym semestrze definiowane dla macierzy dowolnego rozmiaru) jest przekształceniem liniowym. W szczególności, wszystkie przekształcenia liniowe rozważane w poprzednim semestrze (gdy $m, n \leq 3$) pozostają przekształceniami liniowymi (choć nowa definicja jest inna niż ta z pierwszego semestru). Sprawdzamy, korzystając z praw działań na macierzach:

$$F(X + Y) = A(X + Y) = AX + AY = F(X) + F(Y)$$

$$F(tX) = A(tX) = t \cdot (AX) = t \cdot F(X)$$

W szczególności przekształcenie $F : \mathbb{R}^n \rightarrow \mathbb{R}$ postaci: $F\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = a_1x_1 + \cdots + a_nx_n$ jest liniowe.

Przykład 2

Przekształcenie $D : C^1(\mathbb{R}) \rightarrow C(\mathbb{R})$ zdefiniowane: $D(f) = f'$ (tzn. operator pochodnej) jest liniowe, zgodnie z twierdzeniem z analizy matematycznej:

$$(f + g)' = f' + g' \quad (t \cdot f)' = t \cdot f'$$

Przykład 3

Przekształcenie $I : C(\mathbb{R}) \rightarrow C^1(\mathbb{R})$ zdefiniowane: $I(f) = \int f(x)dx$ (tzn. operator całki nieoznaczonej) byłoby liniowe, zgodnie z twierdzeniem z analizy matematycznej:

$$\int (f(x) + g(x))dx = \int f(x)dx + \int g(x)dx$$

$$\int t \cdot f(x)dx = t \cdot \int f(x)dx$$

gdyby nie to, że jest ono źle zdefiniowane (całka nieoznaczona daje w wyniku funkcję z dokładnością do stałej, zwaną stałą całkowania, czyli cały zbiór funkcji). Z problemem poprawnego zdefiniowania operatora całki nieoznaczonej (który faktycznie okaże się operatorem liniowym) zmierzmy się ponownie, gdy poznamy pojęcie *warstwy podprzestrzeni*.

Przykład 4

Niech a i b będą ustalonymi liczbami rzeczywistymi. Przekształcenie $I : C(\mathbb{R}) \rightarrow \mathbb{R}$ zdefiniowane: $I(f) = \int_a^b f(x)dx$ (tzn. operator całki oznaczonej) jest liniowe, zgodnie z twierdzeniem z analizy matematycznej:

$$\int_a^b (f(x) + g(x))dx = \int_a^b f(x)dx + \int_a^b g(x)dx$$

$$\int_a^b t \cdot f(x)dx = t \cdot \int_a^b f(x)dx$$

Przykład 5

Przekształcenie $F : C(\mathbb{R}) \rightarrow \mathbb{R}$ zdefiniowane $F(f) = f(3)$ (czyli przekształcenie przypisujące każdej funkcji jej wartość w punkcie 3) jest liniowe, bo:

$$(f + g)(3) = f(3) + g(3) \quad (t \cdot f)(3) = t \cdot f(3)$$

Podobnie przekształcenie $G : \mathbb{R}_n[x] \rightarrow \mathbb{R}$ zdefiniowane $G(P) = P(4)$ (czyli przypisujące każdemu wielomianowi jego wartość w punkcie 4).

Przykład 6

Przekształcenie $F : \mathbb{R}_3[x] \rightarrow \mathbb{R}_4[x]$ zdefiniowane $F(P)(x) = xP(x)$ (czyli mnożące każdy wielomian przez x) jest przekształceniem liniowym, bo:

$$x(P(x) + Q(x)) = xP(x) + xQ(x) \quad x(t \cdot P(x)) = t \cdot xP(x)$$

Definicja 1.18

Niech V i W będą przestrzeniami liniowymi, zaś $F : V \rightarrow W$ przekształceniem liniowym. Wówczas:

- *obrazem* przekształcenia liniowego F nazywamy zbiór $\text{Im}F = \{F(v) : v \in V\} \subset W$,
- *jądrem* przekształcenia liniowego F nazywamy zbiór $\ker F = \{v \in V : F(v) = 0\}$.

Fakt 1.19

Niech V i W będą przestrzeniami liniowymi, zaś $F : V \rightarrow W$ przekształceniem liniowym. Wówczas:

- 1) $\ker F$ jest podprzestrzenią dziedziny V (tzn. $\ker F < V$),
- 2) $\operatorname{Im} F$ jest podprzestrzenią przeciwdziedziny W (tzn. $\operatorname{Im} F < W$).

Dowód. Wystarczy sprawdzić, że $\ker F$ oraz $\operatorname{Im} F$ są zamknięte na dodawanie wektorów oraz na mnożenie przez skalar. Niech $v, v' \in \ker F$, czyli $F(v) = F(v') = 0$. Wówczas, z uwagi na addytywność F :

$$F(v + v') = F(v) + F(v') = 0 + 0 = 0$$

czyli $v + v' \in \ker F$, co oznacza zamkniętość na dodawanie. Zamkniętość na mnożenie przez skalar sprawdzamy podobnie.

Niech teraz $w, w' \in \operatorname{Im} F$. Oznacza to, że $w = F(v)$ oraz $w' = F(v')$ dla pewnych wektorów $v, v' \in V$. Wówczas, z addytywności F :

$$w + w' = F(v) + F(v') = F(v + v')$$

czyli $w + w' \in \operatorname{Im} F$, co oznacza zamkniętość na dodawanie. Zamkniętość na mnożenie przez skalar sprawdzamy podobnie. $\square$

Fakt 1.20

Niech V i W będą przestrzeniami liniowymi, zaś $F : V \rightarrow W$ przekształceniem liniowym. Wówczas:

- 1) F jest różnowartościowe $\Leftrightarrow \ker F = \{0\}$,
- 2) F jest „na” $\Leftrightarrow \operatorname{Im} F = W$.

Dowód. Część (2) to przeformułowana definicja funkcji „na”. Udowodnimy część (1).

$\Rightarrow$ Załóżmy, że F jest różnowartościowe. Wówczas przeciwobraz każdego punktu $w \in W$ (w szczególności przeciwobraz $0 \in W$, który oznaczamy $\ker F$) jest zbiorem pustym lub jednoelementowym. Ponieważ $F(0) = 0$ (z liniowości), więc $F^{-1}[0] = \ker F = \{0\}$.

$\Leftarrow$ Załóżmy, że $\ker F = \{0\}$. Niech $v, v' \in V$ będą takimi punktami, że $F(v) = F(v')$. Wówczas z liniowości F :

$$F(v - v') = F(v) - F(v') = 0$$

Zatem $v - v' \in \ker F$, czyli $v - v' = 0$, a zatem $v = v'$. Dowodzi to różnowartościowości F (różne punkty nie mogą mieć jednakowych obrazów). $\square$

Twierdzenie 1.21: Twierdzenie o indeksie

Niech V i W będą przestrzeniami liniowymi, przy czym $\dim V < \infty$, zaś $F : V \rightarrow W$ przekształceniem liniowym. Wówczas:

$$\dim \ker F + \dim \operatorname{Im} F = \dim V$$

Wniosek 1.22

Niech V i W będą skończone wymiarowymi przestrzeniami liniowymi tego samego wymiaru (tzn. $\dim V = \dim W < \infty$), zaś $F : V \rightarrow W$ – przekształceniem liniowym. Wówczas następujące warunki są równoważne:

- 1) F jest różnowartościowe
- 2) F jest „na”
- 3) F jest bijekcją

Dowód. Zauważmy, że $\dim \ker F \geq 0$ oraz $\dim \operatorname{Im} F \leq \dim V$ (gdyż $\operatorname{Im} F \subset W$, czyli $\dim \operatorname{Im} F \leq \dim W = \dim V$). Z twierdzenia o indeksie wiemy, że:

$$\dim \ker F + \dim \operatorname{Im} F = \dim V$$

a wobec tego $\dim \ker F = 0 \iff \dim \operatorname{Im} F = \dim V$. Wynika stąd, że jeśli F jest różnowartościowe, to jest „na”, a zatem jest bijekcją oraz, że jeśli F jest „na” to jest różnowartościowe, czyli też jest bijekcją. $\square$

Zwróćmy uwagę, że powyższy wniosek prawdziwy jest wyłącznie w przypadku, gdy dziedzina i przeciwdziedzina są tego samego (skończonego) wymiaru.

Definicja 1.23

Niech $F : V \rightarrow W$ będzie przekształceniem liniowym, a $\mathcal{B} = (b_1, \dots, b_n)$ i $\mathcal{C} = (c_1, \dots, c_m)$ bazami, odpowiednio V i W . Wówczas macierz:

$$m_{\mathcal{C}}^{\mathcal{B}}(F) = ([F(b_1)]_{\mathcal{C}}, \dots, [F(b_n)]_{\mathcal{C}}) \quad (1.2)$$

nazywamy *macierzą przekształcenia F w bazach $\mathcal{B}$ i $\mathcal{C}$* .

Fakt 1.24

Dla przekształcenia liniowego $F : V \rightarrow W$ oraz baz $\mathcal{B}$ i $\mathcal{C}$ zachodzi:

$$[F(v)]_{\mathcal{C}} = m_{\mathcal{C}}^{\mathcal{B}}(F) \cdot [v]_{\mathcal{B}} \quad (1.3)$$

Dowód. Jeśli $[v]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$, to $v = \alpha_1 b_1 + \dots + \alpha_n b_n$. Zauważmy, że dla dowolnej macierzy $A = (A_1, \dots, A_n)$ (czyli macierzy o kolumnach $A_1, \dots, A_n$) zachodzi:

$$A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (A_1, \dots, A_n) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 A_1 + \dots + x_n A_n$$

Zgodnie ze wzorem (1.2) dostajemy zatem:

$$m_{\mathcal{C}}^{\mathcal{B}}(F) \cdot [v]_{\mathcal{B}} = ([F(b_1)]_{\mathcal{C}}, \dots, [F(b_n)]_{\mathcal{C}}) \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha_1 [F(b_1)]_{\mathcal{C}} + \dots + \alpha_n [F(b_n)]_{\mathcal{C}}$$

$$= [\alpha_1 F(b_1) + \dots + \alpha_n F(b_n)]_{\mathcal{C}} = [F(\alpha_1 b_1 + \dots + \alpha_n b_n)]_{\mathcal{C}} = [F(v)]_{\mathcal{C}}$$

$\square$

Wniosek 1.25

Dla baz $\mathcal{B}$ i $\mathcal{C}$ przestrzeni V zachodzi:

$$[v]_{\mathcal{C}} = m_{\mathcal{C}}^{\mathcal{B}}(\text{id}) \cdot [v]_{\mathcal{B}} \quad (1.4)$$

Macierz $m_{\mathcal{C}}^{\mathcal{B}}(\text{id})$ nazywamy *macierzą zmiany bazy*.

Dowód. Jest to szczególna wersja Faktu 1.24, dla $F = \text{id}$. □

Fakt 1.26

Dla przekształceń liniowych $F : U \rightarrow V$ i $G : V \rightarrow W$ oraz baz $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ odpowiednio U , V , W zachodzi:

$$m_{\mathcal{D}}^{\mathcal{B}}(G \circ F) = m_{\mathcal{D}}^{\mathcal{C}}(G) \cdot m_{\mathcal{C}}^{\mathcal{B}}(F) \quad (1.5)$$

Dowód. Zauważmy, że zgodnie z Faktem 1.24:

$$m_{\mathcal{D}}^{\mathcal{B}}(G \circ F) \cdot [b_1]_{\mathcal{B}} = [(G \circ F)(b_1)]_{\mathcal{D}} = [G(F(b_1))]_{\mathcal{D}}$$

$$m_{\mathcal{D}}^{\mathcal{C}}(G) \cdot m_{\mathcal{C}}^{\mathcal{B}}(F) \cdot [b_1]_{\mathcal{B}} = m_{\mathcal{D}}^{\mathcal{C}}(G) \cdot [F(b_1)]_{\mathcal{C}} = [G(F(b_1))]_{\mathcal{D}}$$

Ponieważ $[b_1]_{\mathcal{B}} = \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}$, więc oznacza to, że pierwsza kolumna macierzy $m_{\mathcal{D}}^{\mathcal{B}}(G \circ F)$ jest identyczna z pierwszą kolumną macierzy $m_{\mathcal{D}}^{\mathcal{C}}(G) \cdot m_{\mathcal{C}}^{\mathcal{B}}(F)$. Podobnie pokazujemy (zastępując b_1 przez b_i), że pozostałe kolumny tych macierzy też są równe. □

Wniosek 1.27

Dla baz $\mathcal{B}$ i $\mathcal{C}$ przestrzeni V zachodzi:

$$m_{\mathcal{C}}^{\mathcal{B}}(\text{id}) = (m_{\mathcal{B}}^{\mathcal{C}}(\text{id}))^{-1} \quad (1.6)$$

Dowód. Stosując Fakt 1.26 dla przekształceń $F = G = \text{id}$, gdzie $\text{id} : V \rightarrow V$ jest przekształceniem identycznościowym i przyjmując bazę $\mathcal{D} = \mathcal{B}$ otrzymujemy:

$$I = m_{\mathcal{B}}^{\mathcal{B}}(\text{id}) = m_{\mathcal{B}}^{\mathcal{B}}(\text{id} \circ \text{id}) = m_{\mathcal{B}}^{\mathcal{C}}(\text{id}) \cdot m_{\mathcal{C}}^{\mathcal{B}}(\text{id})$$

Stąd macierze $m_{\mathcal{B}}^{\mathcal{C}}(\text{id})$ i $m_{\mathcal{C}}^{\mathcal{B}}(\text{id})$ są wzajemnie odwrotne. □

Wniosek 1.28

Dla baz $\mathcal{B}$ i $\mathcal{C}$ przestrzeni V oraz przekształcenia $F : V \rightarrow V$ zachodzi:

$$m_{\mathcal{C}}^{\mathcal{C}}(F) = m_{\mathcal{C}}^{\mathcal{B}}(\text{id}) \cdot m_{\mathcal{B}}^{\mathcal{B}}(F) \cdot m_{\mathcal{B}}^{\mathcal{C}}(\text{id}) \quad (1.7)$$

Dowód. Stosując dwukrotnie Fakt 1.26 otrzymujemy:

$$m_{\mathcal{C}}^{\mathcal{B}}(\text{id}) \cdot m_{\mathcal{B}}^{\mathcal{B}}(F) \cdot m_{\mathcal{B}}^{\mathcal{C}}(\text{id}) = m_{\mathcal{C}}^{\mathcal{B}}(\text{id}) \cdot m_{\mathcal{B}}^{\mathcal{C}}(F \circ \text{id}) = m_{\mathcal{C}}^{\mathcal{C}}(\text{id} \circ F \circ \text{id}) = m_{\mathcal{C}}^{\mathcal{C}}(F)$$

□

Definicja 1.29

Przekształcenie liniowe $F : V \rightarrow W$ nazywamy *odwracalnym* (lub *izomorfizmem*), jeśli istnieje przekształcenie liniowe $G : W \rightarrow V$ takie, że $F \circ G = \text{id}_W$ i $G \circ F = \text{id}_V$. Takie G jest najwyżej jedno i oznaczamy je $G = F^{-1}$ (i nazywamy *przekształceniem odwrotnym* do F). Przestrzenie liniowe V i W nazywamy wówczas *izomorficznymi* (i oznaczamy $V \cong W$).

Fakt 1.30

Przekształcenie $F : V \rightarrow W$ jest izomorfizmem wtedy i tylko wtedy, gdy jest bijekcją.

Dowód. Zwróćmy uwagę, że warunek bijektywności $F : V \rightarrow W$ oznacza istnienie przekształcenia odwrotnego $F^{-1} : W \rightarrow V$, ale nie gwarantuje natychmiast, że F^{-1} jest przekształceniem liniowym. Zatem pozostaje sprawdzić, że F^{-1} jest addytywne i jednorodne. Weźmy dowolne wektory $w, w' \in W$. Ponieważ F jest bijekcją, więc istnieją $v, v' \in V$ takie, że $w = F(v)$ i $w' = F(v')$, czyli $F^{-1}(w) = v$ i $F^{-1}(w') = v'$. Wówczas:

$$F(v + v') = F(v) + F(v') = w + w' \quad \text{czyli} \quad F^{-1}(w + w') = v + v' = F^{-1}(w) + F^{-1}(w')$$

czyli F^{-1} jest addytywne. Podobnie sprawdzamy jednorodność F^{-1} . □

Fakt 1.31

Jeśli V jest przestrzenią liniową wymiaru n , to $V \cong \mathbb{R}^n$.

Dowód. Odwzorowanie $F : V \rightarrow \mathbb{R}^n$ zdefiniowane jako $F(v) = [v]_{\mathcal{B}}$, gdzie $\mathcal{B}$ jest pewną ustaloną bazą, jest przekształceniem liniowym (dodawania wektorów odpowiada dodawaniu współrzędnych w bazie $\mathcal{B}$, podobnie mnożenie przez skalar) oraz jest bijekcją. Zatem F jest izomorfizmem. □

Rozdział 2

Macierze i układy równań

2.1 Wyznacznik macierzy

Definicja 2.1

Rozstawieniem wyrazów macierzy $n \times n$ nazywamy każdy taki wybór zbioru n wyrazów tej macierzy:

$$a_{i_1 j_1} \quad a_{i_2 j_2} \quad \dots \quad a_{i_n j_n}$$

że w każdym wierszu i w każdej kolumnie znajduje się dokładnie jeden z wybranych wyrazów (tzn. indeksy $i_1, \dots, i_n$ są parami różne oraz indeksy $j_1, \dots, j_n$ są parami różne). *Rozstawieniem diagonalnym* nazwiemy rozstawienie zawierające wszystkie wyrazy przekątnej macierzy:

$$a_{11} \quad a_{22} \quad \dots \quad a_{nn}$$

Rozstawienie wyrazów macierzy $n \times n$ jest matematyczną wersją szachowego zadania: ustawić n wież na szachownicy $n \times n$ w taki sposób, by żadne dwie z nich się nie atakowały.

Pojęcie rozstawienia zostało wprowadzone wyłącznie na potrzeby niniejszego skryptu i nie jest ogólnie przyjętą terminologią. Klasyczne podręczniki algebry liniowej zamiast posługiwać się pojęciem rozstawienia wyrazów macierzy używają pojęcia permutacji liczb $1, 2, \dots, n$, czyli bijekcji:

$$\sigma : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$$

Używając tej terminologii, rozstawienie to zbiór wyrazów macierzy postaci:

$$a_{1\sigma(1)} \quad a_{2\sigma(2)} \quad \dots \quad a_{n\sigma(n)}$$

dla pewnej permutacji σ . Rozstawienie diagonalne odpowiada permutacji identycznościowej, tzn. permutacji σ takiej, że $\sigma(k) = k$ dla dowolnego k .

Fakt 2.2

Dla macierzy $n \times n$ istnieje dokładnie $n!$ rozstawień.

Dowód. Z pierwszego wiersza macierzy możemy wybrać dowolny z n wyrazów. Z drugiego wiersza możemy wybrać dowolny z $n - 1$ wyrazów (jedna kolumna jest już „zajęta” przez wyraz wybrany z pierwszego wiersza). Z trzeciego wiersza możemy wybrać dowolny z $n - 2$ wyrazów (dwie kolumny są już „zajęte” przez wyrazy wybrane z pierwszego i drugiego wiersza), itd. Stąd liczba możliwych rozstawień to:

$$n \cdot (n - 1) \cdot \dots \cdot 2 \cdot 1 = n!$$

□

Fakt 2.3

Każde rozstawienie wyrazów macierzy $n \times n$ możemy przekształcić w rozstawienie diagonalne poprzez wielokrotne powtórzenie operacji zamiany miejscami dwóch kolumn macierzy. Rozstawienie nazywamy:

- *parzystym*, jeśli można to zrobić przy użyciu parzystej liczby zamian,
- *nieparzystym*, jeśli można to zrobić przy użyciu nieparzystej liczby zamian.

Żadne rozstawienie nie może być równocześnie parzyste i nieparzyste.

Pojęcie parzystego/nieparzystego rozstawienia odpowiada pojęciom parzystej/nieparzystej permutacji.

Dowód. Fakt, że każde rozstawienie można przekształcić w rozstawienie diagonalne przez wielokrotne zamiany miejscami dwóch kolumn jest oczywisty, choć przekształcenie to może być wykonywane na wiele różnych sposobów. Głównym elementem powyższego Faktu jest stwierdzenie, że parzystość liczby zamian jest jednakowa dla każdego z tych sposobów. Poniżej przedstawiamy szkic dowodu tego faktu. Dowód ten nie jest istotny dla dalszego toku wykładu i mniej zaawansowany czytelnik może go pominąć.

Przy ustalonym rozstawieniu łączymy odcinkiem każde dwa z wyrazów tego rozstawienia (otrzymując $\binom{n}{2}$ odcinków). Każdy z tych odcinków ma kierunek $\nearrow$ lub $\searrow$ (nie będzie odcinków pionowych ani poziomych, gdyż żadne dwa wyrazy nie leżą w tej samej kolumnie ani wierszu). Nazwijmy liczbę odcinków o kierunku $\nearrow$ *liczbą nieporządków* rozstawienia (np. rozstawienie diagonalne ma liczbę nieporządków 0). Zauważmy teraz, że:

- zamiana miejscami dwóch sąsiednich kolumn zawsze zmienia liczbę nieporządków o 1,
- zamianę miejscami kolumn odległych o k można przedstawić w postaci sekwencji $2k + 1$ zamian miejscami sąsiednich kolumn, czyli zmienia ona liczbę nieporządków o $2k + 1$ (np. zamianę miejscami kolumny pierwszej i trzeciej można przedstawić jako ciąg 3 zamian sąsiednich kolumn: pierwsza–druga, druga–trzecia, pierwsza–druga).

Wobec tego zamiana miejscami dowolnych dwóch kolumn zmienia parzystość liczby nieporządków. Ponieważ rozstawienie diagonalne ma liczbę nieporządków 0, więc rozstawienie o parzystej liczbie nieporządków wymaga parzystej liczby zamian kolumn, a rozstawienie o nieparzystej liczbie nieporządków wymaga nieparzystej liczby zamian kolumn i nie może istnieć rozstawienie, które jest równocześnie parzyste i nieparzyste. $\square$

Fakt 2.4

Dla $n \geq 2$, wśród wszystkich $n!$ rozstawień wyrazów macierzy $n \times n$, połowa rozstawień jest parzysta, a połowa – nieparzysta.

Dowód. Rozstawienia możemy połączyć w pary, łącząc każde rozstawienie z rozstawieniem powstałym z niego przez zamianę miejscami pierwszej i drugiej kolumny. W ten sposób w każdej parze mamy jedno rozstawienie parzyste i jedno nieparzyste, czyli rozstawień parzystych i nieparzystych jest tyle samo. $\square$

Definicja 2.5

Wyznacznik macierzy $A = (a_{ij})$ rozmiaru $n \times n$ to:

$$\det A = \sum \pm a_{i_1 j_1} \cdot a_{i_2 j_2} \cdots a_{i_n j_n} \quad (2.1)$$

gdzie sumujemy względem wszystkich możliwych rozstawień wyrazów macierzy A , a wybór znaku jest zgodny z parzystością rozstawienia (+ dla parzystych, a – dla nieparzystych).

Przykład 1

Oblicz wyznacznik macierzy $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Rozwiązanie. Nietrudno zauważyć, że możliwe są tylko dwa rozstawienia:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$+ad \quad -bc$$

Drugie rozstawienie wymaga jednej zamiany kolumn, by sprowadzić je do diagonalnego (czyli jest nieparzyste, stąd znak –). Pierwsze rozstawienie już jest diagonalne, czyli parzyste (wymaga 0 zamian kolumn), stąd znak +. Otrzymujemy wzór:

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

który pokrywa się ze wzorem z ubiegłego semestru. Zauważmy, że otrzymana suma ma 2! składników, z których każdy jest iloczynem 2 wyrazów macierzy oraz że połowa składników występuje ze znakiem +, a połowa ze znakiem –.

Przykład 2

Oblicz wyznacznik macierzy $\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$.

Rozwiązanie. Nietrudno zauważyć, że możliwe jest sześć rozstawień:

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

$$+a_1 b_2 c_3 \quad +b_1 c_2 a_3 \quad +c_1 a_2 b_3 \quad -c_1 b_2 a_3 \quad -b_1 a_2 c_3 \quad -a_1 b_2 c_3$$

Pierwsze rozstawienie jest diagonalne (czyli parzyste). Drugie rozstawienie można przekształcić w diagonalne wykonując dwie zamiany kolumn (zamieniając pierwszą kolumnę z drugą, a następnie drugą z trzecią), czyli również jest parzyste. Czwarte rozstawienie można przekształcić w diagonalne przez jedną zamianę kolumn (zamieniając pierwszą kolumnę z trzecią), czyli jest nieparzyste. Podobnie sprawdzamy znaki dla pozostałych rozstawień, otrzymując wzór:

$$\det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = a_1 b_2 c_3 + b_1 c_2 a_3 + c_1 a_2 b_3 - c_1 b_2 a_3 - b_1 a_2 c_3 - a_1 b_2 c_3$$

który pokrywa się z wzorem z ubiegłego semestru. Zauważmy, że otrzymana suma ma 3! składników, z których każdy jest iloczynem 3 wyrazów macierzy oraz że połowa składników występuje ze znakiem +, a połowa ze znakiem –.

W praktyce rzadko będziemy wyliczać wyznacznik korzystając bezpośrednio z Definicji 2.5, ale definicja ta pozwoli nam zrozumieć wiele własności wyznacznika, które znacząco ułatwią jego wyliczanie.

Fakt 2.6

Dla dowolnej macierzy $A \in M_{n \times n}$ zachodzi $\det A^\top = \det A$.

Dowód. Każdemu rozstawieniu wyrazów macierzy A odpowiada identyczne rozstawienie wyrazów macierzy $A^\top$, powstałe przez „odbicie” rozstawienia względem przekątnej macierzy. W związku z tym składniki sumy (2.1) dla wyznacznika $\det A$ i dla wyznacznika $\det A^\top$ są jednakowe, o ile tylko każde rozstawienie ma tę samą parzystość, co jego „odbicie” względem przekątnej.

Definicja parzystości rozstawienia wyrazów macierzy zawarta w sformułowaniu Faktu 2.3 odwołuje się do zamian miejscami kolumn macierzy, a dowód tego faktu bazuje na obserwacji, że zamiana miejscami dwóch kolumn zmienia parzystość liczby nieporządków. Ponieważ zamiana miejscami dwóch wierszy również zmienia parzystość liczby nieporządków (uzasadnienie jest analogiczne jak dla kolumn), więc parzystość rozstawienia można równie dobrze określić przy pomocy liczby zamian wierszy, które prowadzą do rozstawienia diagonalnego.

Wobec tego, ponieważ transpozycja zamienia kolumny na wiersze, więc rozstawienie i jego „odbicie” względem przekątnej macierzy mają taką samą parzystość, co dowodzi, że $\det A = \det A^\top$. $\square$

Wniosek 2.7

Wyznacznik macierzy $A \in M_{n \times n}$ zmienia znak przy zamianie miejscami dwóch kolumn lub dwóch wierszy.

Dowód. Zamiana miejscami dwóch kolumn macierzy zamienia każde rozstawienie na rozstawienie o tym samym iloczynie, ale przeciwnej parzystości, czyli zmienia znak każdego składnika sumy (2.1), a wobec tego zmienia znak wyznacznika.

Zamiana miejscami dwóch wierszy macierzy A odpowiada zamianie miejscami dwóch kolumn macierzy $A^\top$, czyli zmienia znak wyznacznika $\det A^\top$. Ponieważ, zgodnie z Faktem 2.6, $\det A = \det A^\top$, więc zmienia również znak wyznacznika $\det A$. $\square$

Wniosek 2.8

Jeśli macierz $A \in M_{n \times n}$ ma dwie jednakowe kolumny (lub dwa jednakowe wiersze), to $\det A = 0$.

Dowód. Jeśli zamienimy miejscami dwie jednakowe kolumny macierzy A , to z jednej strony wyznacznik się nie zmieni (bo macierz pozostanie niezmienną), a z drugiej strony wyznacznik zmieni się na przeciwny (zgodnie z Wnioskiem 2.7). Wobec tego:

$$\det A = -\det A \quad \text{czyli} \quad \det A = 0$$

$\square$

Fakt 2.9

Wyznacznik macierzy diagonalnej to iloczyn wyrazów na przekątnej, tzn.

$$\det \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & a_n \end{pmatrix} = a_1 \cdot \dots \cdot a_n$$

Dowód. Jedyne rozstawienie wyrazów macierzy diagonalnej, które nie zawiera wyrazu 0, to rozstawienie złożone z wyrazów na przekątnej. Stąd suma (2.1) dla rozważanej macierzy diagonalnej redukuje się do jednego składnika: $+a_1 \cdot \dots \cdot a_n$. $\square$

Fakt 2.10

Wyznacznik macierzy górnotrójkątnej to iloczyn wyrazów na przekątnej, tzn.

$$\det \begin{pmatrix} a_1 & * & \dots & * \\ 0 & a_2 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & a_n \end{pmatrix} = a_1 \cdot \dots \cdot a_n$$

Dowód. Jedyne rozstawienie wyrazów macierzy górnotrójkątnej, które nie zawiera wyrazu 0, to rozstawienie złożone z wyrazów na przekątnej. Stąd suma (2.1) dla rozważanej macierzy górnotrójkątnej redukuje się do jednego składnika: $+a_1 \cdot \dots \cdot a_n$. $\square$

Przykład 3

$$\det \begin{pmatrix} 2 & 1 & -1 & 0 \\ 0 & 3 & 1 & 2 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 7 \end{pmatrix} = 2 \cdot 3 \cdot 1 \cdot 7 = 42$$

Fakt 2.11

Wyznacznik macierzy kwadratowej złożonej z dwóch *klatek* jest iloczynem wyznaczników obu klatek:

$$\det \begin{pmatrix} a_{11} & \dots & a_{1n} & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & 0 & \dots & 0 \\ a_{n1} & \dots & a_{nn} & 0 & \dots & 0 \\ 0 & \dots & 0 & b_{11} & \dots & b_{1m} \\ 0 & \dots & 0 & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & b_{m1} & \dots & b_{mm} \end{pmatrix} = \det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \cdot \det \begin{pmatrix} b_{11} & \dots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{m1} & \dots & b_{mm} \end{pmatrix}$$

Powyższy wzór w uproszczeniu zapisujemy jako:

$$\det \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \det A \cdot \det B$$

gdzie A i B to (kwadratowe) klatki.

Dowód. We wzorze (2.1) wystarczy sumować względem tych rozstawień macierzy, które nie zawierają zer (rozstawienie zawierające zero daje zerowy składnik sumy). W przypadku rozważanej macierzy postaci $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ rozstawienia nie zawierające zer to dokładnie te rozstawienia, które pochodzą od połączenia dowolnego rozstawienia macierzy A z dowolnym rozstawieniem macierzy B . $\square$

Przykład 4

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

Dla większej czytelności w zapisie dużych macierzy często pomija się bloki zer (zostawiając puste miejsca), np. powyższe wyliczenie zapisuje się w postaci:

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

Fakt 2.11 można uogólnić na macierze kwadratowe złożone z większej liczby klatek (kwadratowych macierzy ułożonych wzdłuż przekątnej), jak w następującym przykładzie:

Przykład 5

$$\det \begin{pmatrix} 1 & 5 & & & \\ 3 & 1 & & & \\ & & 7 & 2 & \\ & & 2 & 0 & \\ & & & & 3 & 1 \\ & & & & 1 & 1 \end{pmatrix} = \det \begin{pmatrix} 1 & 5 \\ 3 & 1 \end{pmatrix} \cdot \det \begin{pmatrix} 7 & 2 \\ 2 & 0 \end{pmatrix} \cdot \det \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix} = (-14) \cdot (-4) \cdot 2 = 112$$

Zauważmy, że powyższy wzór stanowi uogólnienie Faktu 2.9, dotyczącego wyznacznika macierzy diagonalnej (wyrazy na przekątnej macierzy diagonalnej można traktować jako klatki rozmiaru 1×1).

Definicja 2.12

Niech V i W będą przestrzeniami liniowymi. Odwzorowanie $F : V \times \cdots \times V \rightarrow W$ (przebiegające układ n wektorów $(v_1, \dots, v_n)$ przestrzeni V na jeden wektor przestrzeni W) nazywamy *liniowym względem każdej współrzędnej* (albo *n -liniowym*) jeśli dla dowolnego indeksu k :

- 1) $F(v_1, \dots, v_k + v'_k, \dots, v_n) = F(v_1, \dots, v_k, \dots, v_n) + F(v_1, \dots, v'_k, \dots, v_n)$
(addytywność względem k -tej współrzędnej)
- 2) $F(v_1, \dots, t \cdot v_k, \dots, v_n) = t \cdot F(v_1, \dots, v_k, \dots, v_n)$
(jednorodność względem k -tej współrzędnej)

Wyznacznik macierzy rozmiaru $n \times n$ to odwzorowanie

$$\det : M_{n \times n} \rightarrow \mathbb{R}$$

Ponieważ macierz $A \in M_{n \times n}$ możemy traktować jako układ n kolumn (czyli układ n wektorów przestrzeni $\mathbb{R}^n$), co zapisujemy $A = (A_1, \dots, A_n)$, gdzie $A_1, \dots, A_n \in \mathbb{R}^n$ to kolumny macierzy A , więc odwzorowanie $\det$ możemy traktować jako odwzorowanie:

$$\det : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$$

Dla takiego odwzorowania możemy rozważać pytanie o liniowość względem każdej współrzędnej.

Fakt 2.13

Wyznacznik macierzy, traktowany jako odwzorowanie $\det : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$ jest liniowy względem każdej współrzędnej, tzn. dla dowolnego indeksu k :

- 1) $\det(A_1, \dots, A_k + A'_k, \dots, A_n) = \det(A_1, \dots, A_k, \dots, A_n) + \det(A_1, \dots, A'_k, \dots, A_n)$
(addytywność względem k -tej współrzędnej)
- 2) $\det(A_1, \dots, t \cdot A_k, \dots, A_n) = t \cdot \det(A_1, \dots, A_k, \dots, A_n)$
(jednorodność względem k -tej współrzędnej)

Dowód. (1) Dla uproszczenia zapisu pokażemy liniowość względem pierwszej współrzędnej. Dowód liniowości względem pozostałych współrzędnych jest analogiczny. Wprowadźmy oznaczenia:

$$A = (A_1, A_2, \dots, A_n) \quad A' = (A'_1, A_2, \dots, A_n) \quad S = (A_1 + A'_1, A_2, \dots, A_n)$$

Każde rozstawienie wyrazów macierzy możemy uporządkować tak, by wypisywać wyrazy zgodnie z kolejnością kolumn, z których zostały wybrane. Każde rozstawienie wyrazów macierzy S ma więc postać:

$$a_{i_1 1} + a'_{i_1 1} \quad a_{i_2 2} \quad \dots \quad a_{i_n n}$$

a składnik sumy (2.1) związany z tym rozstawieniem to:

$$\pm(a_{i_1 1} + a'_{i_1 1}) \cdot a_{i_2 2} \cdot \dots \cdot a_{i_n n} = \pm a_{i_1 1} \cdot a_{i_2 2} \cdot \dots \cdot a_{i_n n} + \pm a'_{i_1 1} \cdot a_{i_2 2} \cdot \dots \cdot a_{i_n n}$$

czyli suma odpowiedniego składnika sumy (2.1) dla wyznacznika $\det A$ i odpowiedniego składnika sumy (2.1) dla wyznacznika $\det A'$. Stąd:

$$\det S = \det A + \det A'$$

co należało udowodnić.

(2) Pomnożenie wszystkich wyrazów pewnej kolumny przez t powoduje przemnożenie przez t wszystkich składników sumy (2.1), czyli przemnożenie przez t wyznacznika. $\square$

Wniosek 2.14

Wyznacznik macierzy:

- 1) mnoży się przez t , jeśli pomnożymy wybraną kolumnę macierzy przez t ;
- 2) nie zmienia się, jeśli dodamy k -tą kolumnę macierzy do innej kolumny.

Powyższe własności prawdziwe są również, jeśli słowo *kolumna* zastąpimy słowem *wiersz*.

Dowód. (1) to inne sformułowanie Faktu 2.13(2).

(2) Oznaczmy macierz $A = (A_1, \dots, A_n)$, gdzie $A_1, \dots, A_n$ to kolumny macierzy. Dla uproszczenia zapisu rozważymy sytuację, gdy t -krotność pierwszej kolumny dodajemy do drugiej kolumny, czyli gdy macierz:

$$A = (A_1, A_2, A_3, \dots, A_n)$$

przekształcamy w macierz:

$$A' = (A_1, A_2 + tA_1, A_3, \dots, A_n)$$

(dowód w pozostałych przypadkach jest analogiczny). Zgodnie z Faktem 2.13(2) (liniowość względem drugiej współrzędnej) otrzymujemy:

$$\det A' = \det(A_1, A_2 + tA_1, A_3, \dots, A_n) = \det(A_1, A_2, A_3, \dots, A_n) + \det(A_1, tA_1, A_3, \dots, A_n)$$

Zgodnie z udowodnionym punktem (1) oraz Wnioskiem 2.8 (macierz o dwóch jednakowych kolumnach ma zerowy wyznacznik):

$$= \det(A_1, A_2, A_3, \dots, A_n) + t \cdot \det(A_1, A_1, A_3, \dots, A_n) = \det A + 0 = \det A$$

co należało udowodnić.

Z uwagi na Fakt 2.6 transpozycja macierzy nie zmienia wyznacznika, a ponieważ transpozycja zamienia kolumny na wiersze, a wiersze na kolumny, więc punkty (1) i (2) są prawdziwe również dla wierszy macierzy. $\square$

Przykład 6

Obliczyć wyznacznik macierzy:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 4 & 2 & 2 & 4 \\ 0 & 2 & 1 & 1 \\ 3 & 3 & 3 & 2 \end{pmatrix}$$

Rozwiązanie. Zgodnie z Wnioskiem 2.14 dodanie lub odjęcie krotności jednego wiersza od innego wiersza nie wpływa na wyznacznik macierzy. W związku z tym odejmując 4-krotność pierwszego wiersza od drugiego i 3-krotność pierwszego wiersza od czwartego, a następnie dodając drugi wiersz do trzeciego otrzymujemy:

$$\det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 4 & 2 & 2 & 4 \\ 0 & 2 & 1 & 1 \\ 3 & 3 & 3 & 2 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & -2 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & -2 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix} = -2$$

gdzie ostatni wyznacznik jest prosty do wyliczenia jako wyznacznik macierzy górnotrójkątnej.

W tym momencie potrafimy wyliczać wyznacznik macierzy $n \times n$:

- wypisując (zgodnie z definicją) wszystkie $n!$ rozstawień wyrazów macierzy (co jest w praktyce bardzo uciążliwe, chyba że, jak w przypadku macierzy górnotrójkątnej, niewiele jest rozstawień nie zawierających zera);
- dla pewnych macierzy pewnych specjalnych postaci, jak macierz diagonalna, górnotrójkątna lub macierz składająca się z klatek;
- sprowadzając macierz, przy pomocy operacji nie zmieniających wyznacznika, do macierzy górnotrójkątnej (lub innej postaci, dla której wyznacznik łatwo jest obliczyć).

Wciąż brakuje nam jednak prostego algorytmu pozwalającego sprawnie wyliczać wyznacznik dowolnie skomplikowanej macierzy. Tym algorytmem jest przedstawione poniżej rozwinięcie wyznacznika względem dowolnie wybranego wiersza lub kolumny (tzw. rozwinięcie Laplace'a).

Fakt 2.15: Rozwinięcie Laplace'a

Jeśli $A = (a_{ij})$ jest macierzą $n \times n$, a przez A_{ij} oznaczmy macierz $(n-1) \times (n-1)$ powstałą z macierzy A przez usunięcie i -tego wiersza i j -tej kolumny, to prawdziwe są następujące wzory:

- 1) $\det A = \sum_j (-1)^{i+j} a_{ij} \det A_{ij}$ (rozwinięcie wyznacznika względem i -tego wiersza)
- 2) $\det A = \sum_i (-1)^{i+j} a_{ij} \det A_{ij}$ (rozwinięcie wyznacznika względem j -tej kolumny)

W skrócie powyższy algorytm można opisać następująco:

1. Wybierz dowolny wiersz lub kolumnę macierzy.
2. Dla każdego elementu a_{ij} wybranego wiersza/kolumny pomnóż a_{ij} przez wyznacznik macierzy powstałej przez usunięcie wiersza i kolumny zawierającej a_{ij} i dopisz znak odpowiadający pozycji a_{ij} na „szachownicy znaków”:

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

3. Dodaj wszystkie składniki otrzymane w punkcie 2.

Dowód. Dla uproszczenia rozumowania udowodnimy wzór dla rozwinięcia względem pierwszego wiersza. Dowód dla rozwinięcia względem innego wiersza lub względem kolumny jest analogiczny.

Zauważmy, że każde rozstawienie wyrazów macierzy A składa się z elementu a_{1k} (wyraz z pierwszego wiersza) oraz pewnego rozstawienia wyrazów macierzy A_{1k} . odpowiada rozstawienie dla macierzy A zawierające wyraz a_{ij} . Stąd zarówno lewa jak i prawa strona równości:

$$\det A = \sum_k (-1)^{k+1} a_{1k} \det A_{1k}$$

jest sumą iloczynów wszystkich rozstawień wyrazów macierzy A . Do sprawdzenia pozostaje jedynie, że znaki składników po lewej i po prawej stronie równości są jednakowe. Tę część dowodu pominiemy. $\square$

Dla uproszczenia zapisu, wyznacznik macierzy A zamiast $\det A$ oznaczany jest czasem $|A|$. Jest to szczególnie wygodne przy wyliczaniu wyznacznika metodą Laplace'a.

Przykład 7

Oblicz następujące wyznaczniki:

$$\det \begin{pmatrix} 1 & 2 & 5 & 1 \\ 2 & 0 & 1 & 2 \\ 1 & 3 & 1 & 0 \\ 3 & 2 & 1 & 1 \end{pmatrix}, \quad \det \begin{pmatrix} 7 & 0 & 5 & 0 \\ 4 & 1 & 0 & 0 \\ 2 & 2 & 0 & 1 \\ 0 & 1 & 0 & 3 \end{pmatrix}.$$

Rozwiązanie. Rozwijając względem pierwszej kolumny otrzymujemy:

$$\begin{vmatrix} 1 & 2 & 5 & 1 \\ 2 & 0 & 1 & 2 \\ 1 & 3 & 1 & 0 \\ 3 & 2 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 2 \\ 3 & 1 & 0 \\ 2 & 1 & 1 \end{vmatrix} - 2 \begin{vmatrix} 2 & 5 & 1 \\ 3 & 1 & 0 \\ 2 & 1 & 1 \end{vmatrix} + \begin{vmatrix} 2 & 5 & 1 \\ 0 & 1 & 2 \\ 2 & 1 & 1 \end{vmatrix} - 3 \begin{vmatrix} 2 & 5 & 1 \\ 0 & 1 & 2 \\ 3 & 1 & 0 \end{vmatrix}$$

a następnie rozwijając każdy z otrzymanych wyznaczników 3×3 względem pierwszej kolumny (i pomijając składniki zerowe):

$$\begin{aligned}
 &= -3 \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} + 2 \begin{vmatrix} 1 & 2 \\ 1 & 0 \end{vmatrix} - 4 \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} + 6 \begin{vmatrix} 5 & 1 \\ 1 & 1 \end{vmatrix} - 4 \begin{vmatrix} 5 & 1 \\ 1 & 0 \end{vmatrix} + 2 \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} + 2 \begin{vmatrix} 5 & 1 \\ 1 & 2 \end{vmatrix} - 6 \begin{vmatrix} 1 & 2 \\ 1 & 0 \end{vmatrix} - 9 \begin{vmatrix} 5 & 1 \\ 1 & 2 \end{vmatrix} \\
 &= 3 - 4 - 4 + 24 + 4 - 2 + 18 + 12 - 81 = -30
 \end{aligned}$$

Jak widać obliczenia istotnie upraszczają się, ilekroć natrafiamy na wyraz 0. Wykorzystamy to przy obliczaniu drugiego wyznacznika wybierając rozwinięcie względem trzeciej kolumny (zawierającej aż trzy zera):

$$\begin{vmatrix} 7 & 0 & 5 & 0 \\ 4 & 1 & 0 & 0 \\ 2 & 2 & 0 & 1 \\ 0 & 1 & 0 & 3 \end{vmatrix} = 5 \cdot \begin{vmatrix} 4 & 1 & 0 \\ 2 & 2 & 1 \\ 0 & 1 & 3 \end{vmatrix} = 5 \cdot \left(4 \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} - 2 \begin{vmatrix} 1 & 0 \\ 1 & 3 \end{vmatrix} \right) = 70$$

Zwróćmy uwagę, że rozwinięcie Laplace'a zamienia wyznacznik $n \times n$ na n wyznaczników $(n-1) \times (n-1)$, które z kolei w ten sam sposób zamienia na $n(n-1)$ wyznaczników $(n-2) \times (n-2)$ itd. ostatecznie zamieniając na $n!$ wyznaczników 1×1 (czyli $n!$ składników, z których każdy jest iloczynem n liczb, podobnie jak we wzorze (2.1)).

Dlatego, aby zmniejszyć liczbę składników (i uprościć obliczenia), warto przed zastosowaniem rozwinięcia Laplace'a uprościć macierz wielokrotnie dodając krotność jednego wiersza/kolumny do innego wiersza/kolumny (co zgodnie z Wnioskiem 2.14 nie zmienia wartości wyznacznika). Pokazuje to poniższy przykład.

Przykład 8

Obliczmy raz jeszcze wyznacznik:

$$\det \begin{pmatrix} 1 & 2 & 5 & 1 \\ 2 & 0 & 1 & 2 \\ 1 & 3 & 1 & 0 \\ 3 & 2 & 1 & 1 \end{pmatrix}$$

Rozwiązanie. Odejmując ostatnią kolumnę od pierwszej, a następnie odejmując podwojony trzeci wiersz od ostatniego i dodając podwojony pierwszy wiersz do ostatniego otrzymujemy:

$$\begin{vmatrix} 1 & 2 & 5 & 1 \\ 2 & 0 & 1 & 2 \\ 1 & 3 & 1 & 0 \\ 3 & 2 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 2 & 5 & 1 \\ 0 & 0 & 1 & 2 \\ 1 & 3 & 1 & 0 \\ 2 & 2 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 2 & 5 & 1 \\ 0 & 0 & 1 & 2 \\ 1 & 3 & 1 & 0 \\ 0 & -4 & -1 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 2 & 5 & 1 \\ 0 & 0 & 1 & 2 \\ 1 & 3 & 1 & 0 \\ 0 & 0 & 9 & 3 \end{vmatrix}$$

Teraz możemy rozwinąć względem pierwszej kolumny:

$$= \begin{vmatrix} 2 & 5 & 1 \\ 0 & 1 & 2 \\ 0 & 9 & 3 \end{vmatrix} = 2 \begin{vmatrix} 1 & 2 \\ 9 & 3 \end{vmatrix} = -30$$

Poznamy jeszcze sposób na obliczanie wyznacznika jednej bardzo szczególnej macierzy – macierzy Vandermonde'a. Jest ona o tyle ważna, że pojawia się w wielu zastosowaniach, dlatego poniższy wzór zaoszczędzi wielu obliczeń:

Fakt 2.16: Wyznacznik Vandermonde'a

Dla dowolnych liczb rzeczywistych $a_1, \dots, a_n$ zachodzi warunek:

$$\det \begin{pmatrix} 1 & 1 & \dots & 1 \\ a_1 & a_2 & \dots & a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & \dots & a_n^{n-1} \end{pmatrix} = \prod_{i < j} (a_j - a_i) \quad (2.2)$$

gdzie iloczyn po prawej stronie wyliczany jest względem wszystkich par indeksów (i, j) , gdzie $i < j$. W szczególności, powyższy wyznacznik jest niezerowy, o ile liczby $a_1, \dots, a_n$ są parami różne.

Dowód. Zauważmy, że dla ustalonych liczb $a_1, \dots, a_{n-1}$ rozwijając względem ostatniej kolumny wyznacznik:

$$P(x) = \det \begin{pmatrix} 1 & 1 & \dots & 1 \\ a_1 & a_2 & \dots & x \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & \dots & x^{n-1} \end{pmatrix}$$

otrzymujemy wielomian zmiennej x stopnia $n - 1$ o współczynniku przy najwyższej potędze wynoszącym:

$$\det \begin{pmatrix} 1 & 1 & \dots & 1 \\ a_1 & a_2 & \dots & a_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-2} & a_2^{n-2} & \dots & a_{n-1}^{n-2} \end{pmatrix}$$

Co więcej, liczby $a_1, \dots, a_{n-1}$ są pierwiastkami wielomianu P , gdyż $P(a_i) = 0$ jako wyznacznik macierzy o dwóch jednakowych kolumnach (Wniosek 2.8). W związku z tym:

$$P(x) = \det \begin{pmatrix} 1 & 1 & \dots & 1 \\ a_1 & a_2 & \dots & a_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-2} & a_2^{n-2} & \dots & a_{n-1}^{n-2} \end{pmatrix} \cdot (x - a_1) \cdot \dots \cdot (x - a_{n-1})$$

czyli wstawiając $x = a_n$ otrzymujemy:

$$\det \begin{pmatrix} 1 & 1 & \dots & 1 \\ a_1 & a_2 & \dots & a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & \dots & a_n^{n-1} \end{pmatrix} = \det \begin{pmatrix} 1 & 1 & \dots & 1 \\ a_1 & a_2 & \dots & a_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-2} & a_2^{n-2} & \dots & a_{n-1}^{n-2} \end{pmatrix} \cdot \prod_i (a_{n+1} - a_i)$$

Wykorzystując powyższy wzór do przeprowadzenia dowodu indukcyjnego (indukcja względem rozmiaru macierzy), otrzymujemy wzór (2.2). $\square$

Przykład 9

Oblicz

$$\det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 5 \\ 1 & 4 & 9 & 25 \\ 1 & 8 & 27 & 125 \end{pmatrix}$$

Rozwiązanie. Powyższy wyznacznik to wyznacznik Vandermonde'a, który zgodnie ze wzorem

(2.2) jest równy:

$$(5-1)(5-2)(5-3)(3-2)(3-1)(2-1) = 4 \cdot 3 \cdot 2 \cdot 1 \cdot 2 \cdot 1 = 48$$

Fakt 2.17

Niech $A \in M_{n \times n}$. Wówczas $\det A \neq 0 \iff$ kolumny macierzy A są lnz.

Dowód. Niech kolumny macierzy A będą liniowo zależne, tzn. jedna z nich (dla ustalenia uwagi przyjmijmy, że pierwsza) jest kombinacją liniową pozostałych. Odejmując od tej kolumny odpowiednie krotności pozostałych kolumn (co zgodnie z Wnioskiem 2.14 nie zmienia wyznacznika), zamieniamy macierz na macierz z całą kolumną zer, pokazując że $\det A = 0$:

$$\det A = \det(\alpha_2 A_2 + \alpha_3 A_3 + \dots + \alpha_n A_n, A_2, A_3, \dots, A_n) =$$

$$\det(\alpha_3 A_3 + \dots + \alpha_n A_n, A_2, A_3, \dots, A_n) =$$

$$\dots = \det(\alpha_n A_n, A_2, A_3, \dots, A_n) = \det(0, A_2, A_3, \dots, A_n) = 0$$

Niech teraz kolumny macierzy A będą liniowo niezależne. Wówczas stanowią one bazę prze-

strzeni $\mathbb{R}^n$, w szczególności wektor $e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ jest ich kombinacją liniową:

$$e_1 = \alpha_1 A_1 + \dots + \alpha_n A_n$$

Przynajmniej jeden ze współczynników $\alpha_1, \dots, \alpha_n$ jest niezerowy. Przyjmijmy, że jest to α_1 (rozumowanie w pozostałych przypadkach jest podobne). Oznacza to, że mnożąc pierwszą kolumnę macierzy A przez $\alpha_1 \neq 0$ (co mnoży wyznacznik przez α_1), a następnie dodając odpowiednie krotności pozostałych kolumn (co nie zmienia wyznacznika) otrzymujemy macierz

$$A^{(1)} = (e_1, A_2, A_3, \dots, A_n)$$

której kolumny również są liniowo niezależne (bo możemy z nich wygenerować bazę $A_1, \dots, A_n$), a ponadto $\det A^{(1)} \neq 0 \iff \det A \neq 0$. Postępując analogicznie z wektorami $e_2, \dots, e_n$ otrzymujemy ciąg macierzy:

$$A^{(i)} = (e_1, \dots, e_i, A_{i+1}, \dots, A_n)$$

których kolumny są liniowo niezależne oraz

$$\det A \neq 0 \iff \det A^{(1)} \neq 0 \iff \dots \iff \det A^{(n)} \neq 0$$

Ponieważ $A^{(n)} = (e_1, \dots, e_n) = I$, czyli $\det A^{(n)} = \det I = 1 \neq 0$, więc $\det A \neq 0$. □

Przykład 10

Sprawdź, czy wektory $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ są liniowo niezależne.

Rozwiązanie. Zgodnie z Faktem 2.17 wystarczy sprawdzić, czy następująca macierz ma nieze-

rowy wyznacznik:

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

Ponieważ macierz ta ma dwa jednakowe wiersze, więc jej wyznacznik jest zerowy, czyli kolumny są liniowo zależne.

Fakt 2.18: Wzory Cramera

Dany jest układ n równań liniowych z niewiadomymi $x_1, \dots, x_n$:

[illegible]

Macierz $A = (a_{ij}) \in M_{n \times n}$ nazywamy *macierzą główną układu*. Wówczas układ (2.3) ma jednoznaczne rozwiązanie wtedy i tylko wtedy, gdy $\det A \neq 0$. Wtedy rozwiązanie to dane jest wzorem:

$$x_i = \frac{\det A^{(i)}}{\det A}, \quad i = 1, 2, \dots, n \quad (2.4)$$

gdzie $A^{(i)}$ jest macierzą powstałą przez zastąpienie i -tej kolumny A przez $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$.

Dowód. Zapiszmy układ (2.3) w postaci:

$$x_1A_1 + \cdots + x_nA_n = b \quad (2.5)$$

Wówczas:

$$\det A^{(1)} = \det(b, A_2, \dots, A_n) = \det(x_1 A_1 + \dots + x_n A_n, A_2, \dots, A_n)$$

i zgodnie z Faktem 2.13 (liniowość względem pierwszej współrzędnej):

$$\det A^{(1)} = x_1 \det(A_1, A_2, \dots, A_n) + x_2 \det(A_2, A_2, \dots, A_n) + \dots + x_n \det(A_n, A_2, \dots, A_n)$$

a ponieważ wyznacznik macierzy o dwóch jednakowych kolumnach wynosi 0, więc

$$\det A^{(1)} = x_1 \det(A_1, A_2, \dots, A_n) = x_1 \det A \quad (2.6)$$

co w przypadku $\det A \neq 0$ dowodzi wzoru (2.4) dla $i = 1$. Podobnie dowodzimy wzoru dla $i = 2, \dots, n$. W szczególności dla $\det A \neq 0$ mamy dokładnie jedno rozwiązanie.

W przypadku $\det A = 0$ kolumny $A_1, \dots, A_n$ są, zgodnie z Faktem 2.17, liniowo zależne. W związku z tym wektor b albo nie przedstawia się w postaci kombinacji liniowej $A_1, \dots, A_n$ (tzn. układ (2.5) jest sprzeczny), albo przedstawia się na więcej niż jeden sposób (tzn. układ (2.5) jest nieoznaczony). $\square$

Wzory Cramera pozwolą nam łatwo sprawdzać liniową niezależność układu wektorów w przestrzeni $C^\infty(\mathbb{R})$ (umiejętność ta będzie regularnie wykorzystywana w ramach przedmiotu Równania różniczkowe 1, dla którego przestrzeń liniowa $C^\infty(\mathbb{R})$ jest głównym obiektem zainteresowania).

2.2 Macierz odwrotna

Definicja 2.20

Macierzą odwrotną do macierzy $A \in M_{n \times n}$ nazywamy taką macierz $B \in M_{n \times n}$, że

$$AB = BA = I$$

Macierz odwrotną do A oznaczamy A^{-1} . Macierz, dla której istnieje macierz odwrotna nazywamy *macierzą odwracalną*.

Tylko macierze kwadratowe (i to nie wszystkie) są odwracalne.

Fakt 2.21

Dla sprawdzenia, że macierz $B \in M_{n \times n}$ jest odwrotną do macierzy $A \in M_{n \times n}$ wystarczy sprawdzić tylko jeden z warunków

$$AB = I \quad \text{oraz} \quad BA = I$$

(tzn. spełnienie jednego z warunków pociąga spełnienie drugiego).

Dowód. Załóżmy, że spełniony jest warunek $AB = I$. Oznaczmy przez $F_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ przekształcenie o macierzy A , tzn. $F_A(X) = AX$. Podobnie oznaczmy F_B . Wówczas

$$F_A \circ F_B = \text{id}$$

gdyż $A(BX) = (AB)X = IX = X$. Wobec tego F_A jest „na”, czyli zgodnie z Wnioskiem 1.22 jest bijekcją, a więc jest odwracalne. Skoro F_A jest odwracalne, to macierz A jest odwracalna. Mnożąc równanie $AB = I$ z lewej strony przez A^{-1} otrzymujemy $B = A^{-1}$. Podobnie rozumujemy w przypadku, gdy $BA = I$. $\square$

Powyższy fakt można wysłowić w terminach odwzorowań: jeśli V jest **skończenie wymiarową** przestrzenią, zaś $F : V \rightarrow V$ oraz $G : V \rightarrow V$ są przekształceniami liniowymi takimi, że $F \circ G = \text{id}$, to F i G są wzajemnie odwrotne. Dlaczego ważne jest założenie o skończonym wymiarze V pokazuje poniższy przykład.

Przykład 1

Rozważmy przekształcenie $F : \mathbb{R}[x] \rightarrow \mathbb{R}[x]$ zdefiniowane:

$$F(a_n x^n + \cdots + a_1 x + a_0) = a_n x^{n-1} + \cdots + a_1$$

oraz przekształcenie $G : \mathbb{R}[x] \rightarrow \mathbb{R}[x]$ zdefiniowane:

$$G(a_n x^n + \cdots + a_1 x + a_0) = a_n x^{n+1} + \cdots + a_1 x^2 + a_0 x$$

Nietrudno zauważyć, że $F \circ G = \text{id}$, ale $G \circ F \neq \text{id}$, w szczególności F i G nie są odwracalne.

Fakt 2.22

Niech V i W będą skończenie wymiarowymi przestrzeniami liniowymi, zaś $\mathcal{B}$ i $\mathcal{C}$ ich bazami. Wówczas przekształcenie liniowe $F : V \rightarrow W$ jest odwracalne wtedy i tylko wtedy, gdy macierz $m_{\mathcal{C}}^{\mathcal{B}}(F)$ jest odwracalna. Ponadto dla odwracalnego F zachodzi wzór:

$$m_{\mathcal{B}}^{\mathcal{C}}(F^{-1}) = (m_{\mathcal{C}}^{\mathcal{B}}(F))^{-1}$$

Dowód. Zgodnie z Faktem 1.26 mamy

$$m_{\mathcal{C}}^{\mathcal{B}}(F) \circ m_{\mathcal{B}}^{\mathcal{C}}(F^{-1}) = m_{\mathcal{C}}^{\mathcal{C}}(F \circ F^{-1}) = m_{\mathcal{C}}^{\mathcal{C}}(\text{id}) = I$$

Wobec tego macierze (kwadratowe) $m_{\mathcal{B}}^{\mathcal{C}}(F^{-1})$ oraz $m_{\mathcal{C}}^{\mathcal{B}}(F)$ są wzajemnie odwrotne. $\square$

Fakt 2.23: Macierz odwrotna (podejście pierwsze)

Macierz $A \in M_{n \times n}$ jest odwracalna wtedy i tylko wtedy, gdy $\det A \neq 0$. Ponadto, jeżeli:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

to

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} +\det A_{11} & -\det A_{12} & +\det A_{13} & \cdots & \pm\det A_{1n} \\ -\det A_{21} & +\det A_{22} & -\det A_{23} & \cdots & \mp\det A_{2n} \\ +\det A_{31} & -\det A_{32} & +\det A_{33} & \cdots & \pm\det A_{3n} \\ \cdots & \cdots & \cdots & \ddots & \vdots \\ \pm\det A_{n1} & \mp\det A_{n2} & \pm\det A_{n3} & \cdots & +\det A_{nn} \end{pmatrix}^{\top}$$

gdzie A_{ij} to macierz powstała z macierzy A przez usunięcie wiersza i kolumny zawierających wyraz a_{ij} .

Powyższy wzór można przedstawić w postaci 4-krokowego algorytmu:

1. Oblicz wyznacznik $\det A$ macierzy $n \times n$.
2. Policz wszystkie wyznaczniki $\det A_{ij}$ macierzy $(n-1) \times (n-1)$ powstałe z A przez usunięcie jednego wiersza i jednej kolumny i napisz macierz tych wyznaczników.
3. Dopisz znaki, zgodnie z „szachownicą znaków”:

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

4. Transponuj otrzymaną macierz.

Dowód. Szukamy macierzy $B = (b_{ij})$ takiej, że $A \cdot B = I$, czyli

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Wobec tego:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} \\ \vdots \\ b_{n1} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}$$

Traktując to równanie macierzowe jako układ równań z niewiadomymi $b_{11}, \dots, b_{n1}$, zgodnie z wzorami Cramera otrzymujemy

$$b_{i1} = \frac{1}{\det A} \cdot (-1)^{1+i} \det A_{1i}$$

co mieliśmy uzasadnić. Podobnie wyliczenia dla kolejnych kolumn macierzy B pozwalają wyliczyć pozostałe wyrazy. $\square$

Przykład 2

Wyznaczyć macierz odwrotną do macierzy $A = \begin{pmatrix} 2 & 1 & 0 & 1 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

Rozwiązanie.

$$A^{-1} = \frac{1}{24} \begin{pmatrix} +12 & -0 & +0 & -0 \\ -3 & +6 & -0 & +0 \\ +1 & -2 & +8 & -0 \\ -12 & +0 & -0 & +24 \end{pmatrix}^T = \begin{pmatrix} \frac{1}{2} & -\frac{1}{8} & \frac{1}{24} & -\frac{1}{2} \\ 0 & \frac{1}{4} & -\frac{1}{12} & 0 \\ 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Drugi sposób wyliczania macierzy odwrotnej będzie odwoływał się do tzw. operacji elementarnych. Operacje elementarne będą również ważne przy rozwiązywaniu dowolnych układów równań liniowych, dlatego definiujemy je od razu dla wszystkich macierzy prostokątnych.

Definicja 2.24

Operacją elementarną wierszową na macierzy prostokątnej A nazywamy każdą z następujących operacji:

- 1) zamiana miejscami dwóch wierszy,
- 2) przemnożenie wybranego wiersza przez niezerową liczbę,
- 3) dodanie do wybranego wiersza niezerowej krotności innego wiersza.

Fakt 2.25

Każda operacja elementarna wierszowa zamienia macierz A na iloczyn $E \cdot A$, gdzie E jest pewną macierzą kwadratową. Macierz E nazywamy *macierzą elementarną*.

Pominiemy formalny dowód tego faktu, zamiast tego pokazując jego działanie dla trzech przykładów operacji elementarnych na następującej macierzy A :

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

1. Zamiana miejscami drugiego i czwartego wiersza macierzy A jest realizowana przy pomocy macierzy elementarnej, która powstaje z macierzy identycznościowej I przez zmianę położenia drugiej i czwartej jedynki na przekątnej:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \color{red}{0} & 0 & \color{red}{1} \\ 0 & 0 & 1 & 0 \\ 0 & \color{red}{1} & 0 & \color{red}{0} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ \color{red}{a_{41}} & \color{red}{a_{42}} & \color{red}{a_{43}} & \color{red}{a_{44}} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ \color{red}{a_{21}} & \color{red}{a_{22}} & \color{red}{a_{23}} & \color{red}{a_{24}} \end{pmatrix}$$

2. Przemnożenie drugiego wiersza macierzy A przez 3 jest realizowane przy pomocy macierzy elementarnej, która powstaje z macierzy identycznościowej I przez zmianę drugiej jedynki

na przekątnej na 3:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \color{red}{3} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ \color{red}{3}a_{21} & \color{red}{3}a_{22} & \color{red}{3}a_{23} & \color{red}{3}a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

3. Dodanie do trzeciego wiersza macierzy A 4-krotności drugiego wiersza jest realizowane przy pomocy macierzy elementarnej, która powstaje z macierzy identycznościowej I przez zamianę zera na przecięciu trzeciego wiersza i drugiej kolumny na 4:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \color{red}{4} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} + \color{red}{4}a_{21} & a_{32} + \color{red}{4}a_{22} & a_{33} + \color{red}{4}a_{23} & a_{34} + \color{red}{4}a_{24} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

Fakt 2.26: Wyznaczanie macierzy odwrotnej

Jeśli ciąg operacji elementarnych wierszowych przeprowadza macierz kwadratową A na macierz identycznościową I , to ten sam ciąg operacji przeprowadza I na A^{-1} .

Dowód. Wykonując kolejne operacje elementarne wierszowe, którym odpowiadają macierze elementarne $E_1, \dots, E_n$, otrzymujemy kolejno macierze:

$$A, \quad E_1 A, \quad E_2 E_1 A, \quad E_3 E_2 E_1 A, \quad \dots, \quad E_n \dots E_3 E_2 E_1 A$$

Jeśli powyższy ciąg operacji elementarnych prowadzi do macierzy identycznościowej I , czyli:

$$E_n \dots E_3 E_2 E_1 A = (E_n \dots E_3 E_2 E_1) \cdot A = I$$

to macierz $E_n \dots E_3 E_2 E_1$ jest macierzą odwrotną do A , czyli:

$$A^{-1} = E_n \dots E_3 E_2 E_1 = E_n \dots E_3 E_2 E_1 \cdot I$$

Oznacza to, że macierz A^{-1} powstaje przez zastosowanie tego samego ciągu operacji do macierzy identycznościowej I . $\square$

Wniosek 2.27: Macierz odwrotna (podejście drugie)

Jeśli z prawej strony macierzy $A \in M_{n \times n}$ dopiszemy macierz identycznościową $I \in M_{n \times n}$ otrzymując macierz rozmiaru $2n \times n$:

$$(A|I)$$

a następnie przy pomocy elementarnych operacji wierszowych wykonywanych na powstałej macierzy $2n \times n$ zamienimy lewą połowę macierzy na macierz identycznościową I , to prawa połowa macierzy zamieni się na A^{-1} .

Dowód. Wykonywanie elementarnych operacji wierszowych na macierzy $(A|I)$ oznacza równoczesne wykonywanie tych samych operacji na macierzy A i na macierzy I . Zgodnie z Faktem 2.26 te same elementarne operacje wierszowe, które prowadzą do zamiany A na I , prowadzą też do zamiany I na A^{-1} , czyli:

$$(A|I) \longrightarrow (I|A^{-1})$$

$\square$

Przykład 3

Wyznaczyć macierz odwrotną do macierzy $A = \begin{pmatrix} 1 & 2 & 0 & 3 \\ 1 & 1 & 1 & 2 \\ 2 & 1 & 1 & 2 \\ 3 & 1 & 1 & 1 \end{pmatrix}$

Rozwiązanie. Zgodnie z Wnioskiem 2.27 dopisujemy z prawej strony rozważanej macierzy macierz identycznościową, a następnie wykonujemy serię operacji wierszowych. Najpierw odejmujemy odpowiednie krotności pierwszego wiersza od pozostałych wierszy:

$$\begin{array}{l} -w_1 \\ -2w_1 \\ -3w_1 \end{array} \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 3 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 2 & 0 & 1 & 0 & 0 \\ 2 & 1 & 1 & 2 & 0 & 0 & 1 & 0 \\ 3 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array} \right)$$

Następnie odejmujemy krotności drugiego wiersza od wierszy trzeciego i czwartego:

$$\begin{array}{l} -3w_2 \\ -5w_2 \end{array} \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 3 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & -3 & 1 & -4 & -2 & 0 & 1 & 0 \\ 0 & -5 & 1 & -8 & -3 & 0 & 0 & 1 \end{array} \right)$$

Następnie odejmujemy krotność trzeciego wiersza do ostatniego wiersza:

$$-2w_3 \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 3 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & -2 & -1 & 1 & -3 & 1 & 0 \\ 0 & 0 & -4 & -3 & 2 & -5 & 0 & 1 \end{array} \right)$$

W ten sposób sprowadziliśmy wyjściową macierz do postaci macierzy górnotrójkątnej (tzn. wyzerowaliśmy wszystkie wyrazy poniżej przekątnej). Teraz wyzerujemy wszystkie wyrazy powyżej przekątnej. Najpierw odejmujemy odpowiednie krotności ostatniego wiersza od pozostałych wierszy:

$$\begin{array}{l} +3w_4 \\ -w_4 \\ -w_4 \end{array} \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 3 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & -2 & -1 & 1 & -3 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 & -2 & 1 \end{array} \right)$$

Następnie dodajemy odpowiednie krotności trzeciego wiersza do wierszy drugiego i pierwszego:

$$+\frac{1}{2}w_3 \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 0 & 1 & 3 & -6 & 3 \\ 0 & -1 & 1 & 0 & -1 & 0 & 2 & -1 \\ 0 & 0 & -2 & 0 & 1 & -4 & 3 & -1 \\ 0 & 0 & 0 & -1 & 0 & 1 & -2 & 1 \end{array} \right)$$

Wreszcie dodajemy odpowiednią krotność drugiego wiersza do pierwszego wiersza:

$$+2w_2 \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 0 & 1 & 3 & -6 & 3 \\ 0 & -1 & 0 & 0 & -\frac{1}{2} & -2 & \frac{7}{2} & -\frac{3}{2} \\ 0 & 0 & -2 & 0 & 1 & -4 & 3 & -1 \\ 0 & 0 & 0 & -1 & 0 & 1 & -2 & 1 \end{array} \right)$$

W ten sposób otrzymaliśmy macierz diagonalną. Na końcu mnożymy każdy z wierszy przez odpowiednią (niezerową) liczbę tak, by otrzymać macierz identycznościową.

$$\begin{array}{l} \cdot(-1) \\ \cdot(-\frac{1}{2}) \\ \cdot(-1) \end{array} \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 0 & -\frac{1}{2} & -2 & \frac{7}{2} & -\frac{3}{2} \\ 0 & 0 & -2 & 0 & 1 & -4 & 3 & -1 \\ 0 & 0 & 0 & -1 & 0 & 1 & -2 & 1 \end{array} \right)$$

co prowadzi do macierzy odwrotnej:

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & 2 & -\frac{7}{2} & \frac{3}{2} \\ 0 & 0 & 1 & 0 & -\frac{1}{2} & 2 & -\frac{3}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 & -1 & 2 & -1 \end{array} \right)$$

czyli

$$A^{-1} = \begin{pmatrix} 0 & -1 & 1 & 0 \\ \frac{1}{2} & 2 & -\frac{7}{2} & \frac{3}{2} \\ -\frac{1}{2} & 2 & -\frac{3}{2} & \frac{1}{2} \\ 0 & -1 & 2 & -1 \end{pmatrix}$$

Pionowa kreska w powyższym zapisie nie ma żadnego formalnego znaczenia, pomaga jedynie oddzielić obie części powstałej macierzy. Zwróćmy uwagę, że operacje wierszowe należy wykonywać **pojedynczo**, aczkolwiek często na jednym diagramie zobrazować kilka kolejnych operacji.

Jak widać sposób postępowania w powyższym przykładzie jest dość algorytmiczny:

1. Sprowadzamy macierz do macierzy górnotrójkątnej z niezerowymi wyrazami na przekątnej, zerując wyrazy w kolejnych kolumnach (od pierwszej do ostatniej).
2. Sprowadzamy macierz do macierzy diagonalnej z niezerowymi wyrazami na przekątnej, zerując wyrazy w kolejnych kolumnach (od ostatniej do pierwszej).
3. Mnożąc każdy wiersz przez odpowiednią liczbę otrzymujemy macierz identycznościową.

W trakcie wykonywania powyższego algorytmu mogą pojawić się dwa problemy:

- (A) Natrafimy na macierz, której jeden z wyrazów na przekątnej jest zerem.
- (B) Natrafimy na macierz, która ma cały wiersz zer.

Z problemem (A) łatwo sobie poradzić wykonując zamianę miejscami wierszy (jedna z operacji wierszowych), jak pokazuje kolejny przykład.

Przykład 4

Wyznaczyć macierz odwrotną do macierzy $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$

Rozwiązanie.

$$\begin{array}{l} -w_1 \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \\ -w_1 \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & -1 & 1 & -1 & 0 & 1 \end{array} \right) \end{array}$$

W tym miejscu pojawia się problem (A), który rozwiązujemy zamieniając miejscami drugi i trzeci wiersz:

$$-w_3 \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 0 & 1 \\ 0 & 0 & 1 & -1 & 1 & 0 \end{array} \right)$$

$$\begin{aligned}
& +w_2 \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 & 1 & 0 \end{array} \right) \\
& \cdot (-1) \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 1 \\ 0 & -1 & 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 & 1 & 0 \end{array} \right) \\
& \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 1 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 & 1 & 0 \end{array} \right)
\end{aligned}$$

czyli

$$A^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ -1 & 1 & 0 \end{pmatrix}$$

Problem (B) jest większym problemem i oznacza, że macierz, którą badamy nie jest odwracalna, co pokażemy w następnych faktach.

Fakt 2.28

Jeśli macierze $A, B \in M_{n \times n}$ są odwracalne, to macierz AB też jest odwracalna oraz:

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

Dowód. Korzystając z prawa łączności mnożenia macierzy. Zauważmy, że:

$$(AB) \cdot (B^{-1}A^{-1}) = A \cdot (B \cdot B^{-1}) \cdot A^{-1} = A \cdot I \cdot A^{-1} = A \cdot A^{-1} = I$$

czyli macierze AB i $B^{-1}A^{-1}$ są wzajemnie odwrotne. $\square$

Wniosek 2.29

Macierz kwadratową $A \in M_{n \times n}$ można przy pomocy operacji elementarnych wierszowych sprowadzić:

- 1) do macierzy identycznościowej $\iff A$ jest odwracalna,
- 2) do macierzy zawierającej wiersz samych zer $\iff A$ nie jest odwracalna.

Dowód. Opisany wcześniej algorytm pozwala sprowadzić dowolną macierz A do macierzy identycznościowej albo do macierzy zawierającej wiersz samych zer. W pierwszym przypadku (zgodnie z Wnioskiem 2.27) macierz A jest odwracalna. W drugim wypadku mamy:

$$B = E_n \cdot \dots \cdot E_2 \cdot E_1 \cdot A$$

gdzie E_i to macierze elementarne (w szczególności odwracalne), zaś B to macierz z wierszem samych zer (w szczególności nieodwracalna). Gdyby macierz A była odwracalna, to na mocy Faktu 2.28 odwracalna byłaby również macierz B (jako iloczyn macierzy odwracalnych), co prowadziło do sprzeczności. Wobec tego macierz A , którą można sprowadzić przy pomocy operacji wierszowych do macierzy z wierszem samych zer nie jest odwracalna. $\square$

Wniosek 2.30

Macierz jest odwracalna $\iff$ ma niezerowy wyznacznik.

Dowód. Operacje elementarne wierszowe zastosowane do macierzy o niezerowym wyznaczniku dają macierz o niezerowym wyznaczniku, zaś zastosowane do macierzy o zerowym wyznaczniku dają macierz o zerowym wyznaczniku. Ponieważ, zgodnie z Wnioskiem 2.29, macierz odwracalną można w ten sposób sprowadzić do macierzy I (o niezerowym wyznaczniku), zaś macierz nieodwracalną – do macierzy o wierszu samych zer (czyli o zerowym wyznaczniku), więc macierz jest odwracalna wtedy i tylko wtedy, gdy ma niezerowy wyznacznik. $\square$

Wniosek 2.31

Dla macierzy $A, B \in M_{n \times n}$ zachodzi:

- 1) $\det(AB) = \det A \cdot \det B$,
- 2) $\det(A^{-1}) = \frac{1}{\det A}$

Dowód. Nietrudno sprawdzić, że (1) jest prawdą w sytuacji, gdy A jest macierzą elementarną. Jeśli A jest odwracalna, to można ją przedstawić w postaci iloczynu macierzy elementarnych:

$$A = E_n \dots E_2 E_1$$

Wówczas:

$$\begin{aligned} \det(AB) &= \det(E_n \dots E_2 E_1 B) = \det E_n \det(E_{n-1} \dots E_2 E_1 B) \\ &= \dots = \det E_n \det E_{n-1} \dots \det E_2 \det E_1 \det B = \det(E_n \dots E_2 E_1) \det B = \det A \det B \end{aligned}$$

Dowód w przypadku, gdy A nie jest odwracalna pomijamy.

Wzór (2) jest szczególnym przypadkiem wzoru (1), dla $B = A^{-1}$. $\square$

Analogicznie do operacji elementarnych wierszowych można zdefiniować operacje elementarne kolumnowe.

Definicja 2.32

Operacjami elementarnymi kolumnowymi na macierzy prostokątnej A nazywamy każdą z następujących operacji:

- 1) zamiana miejscami dwóch kolumn,
- 2) przemnożenie wybranej kolumny przez niezerową liczbę,
- 3) dodanie do wybranej kolumny niezerowej krotności innej kolumny.

Fakt 2.33

Każda operacja elementarna kolumnowa zamienia macierz A na iloczyn macierzy $A \cdot E$, gdzie E jest pewną macierzą kwadratową. Macierz E nazywamy *macierzą elementarną*.

Pominiemy formalny dowód tego faktu, zamiast tego pokazując jego działanie na trzech przykładach związanych z macierzą

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

1. Zamiana miejscami drugiej i czwartej kolumny jest realizowana przez macierz elementarną, która powstaje z macierzy identycznościowej przez zmianę położenia drugiej i czwartej

jedynki na przekątnej:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \mathbf{0} & 0 & \mathbf{1} \\ 0 & 0 & 1 & 0 \\ 0 & \mathbf{1} & 0 & \mathbf{0} \end{pmatrix} = \begin{pmatrix} a_{11} & \mathbf{a_{14}} & a_{13} & \mathbf{a_{12}} \\ a_{21} & \mathbf{a_{24}} & a_{23} & \mathbf{a_{22}} \\ a_{31} & \mathbf{a_{34}} & a_{33} & \mathbf{a_{32}} \\ a_{41} & \mathbf{a_{44}} & a_{43} & \mathbf{a_{42}} \end{pmatrix}$$

2. Przemnożenie drugiej kolumny przez 3 jest realizowane przez macierz elementarną, która powstaje z macierzy identycznościowej przez zamianę drugiej jedynki na przekątnej na 3:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \mathbf{3} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & \mathbf{3a_{12}} & a_{13} & a_{14} \\ a_{21} & \mathbf{3a_{22}} & a_{23} & a_{24} \\ a_{31} & \mathbf{3a_{32}} & a_{33} & a_{34} \\ a_{41} & \mathbf{3a_{42}} & a_{43} & a_{44} \end{pmatrix}$$

3. Dodanie do trzeciej kolumny 4-krotności drugiej kolumny jest realizowane przez macierz elementarną, która powstaje z macierzy identycznościowej przez zamianę zera na przecięciu drugiej wiersza i trzeciej kolumny na 4:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \mathbf{4} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} + \mathbf{4a_{12}} & a_{14} \\ a_{21} & a_{22} & a_{23} + \mathbf{4a_{22}} & a_{24} \\ a_{31} & a_{32} & a_{33} + \mathbf{4a_{32}} & a_{34} \\ a_{41} & a_{42} & a_{43} + \mathbf{4a_{42}} & a_{44} \end{pmatrix}$$

Jak widać macierze elementarne w Fakcie 2.33 (operacje kolumnowe) i macierze elementarne w Fakcie 2.33 (operacje wierszowe) to te same macierze. Mnożenie przez macierz elementarną z lewej strony prowadzi do operacji wierszowej, natomiast mnożenie z prawej strony – do operacji kolumnowej.

Fakt 2.34

Operacje elementarne (zarówno wierszowe, jak i kolumnowe) nie wpływają na zerowanie się wyznacznika.

W kolejnym rozdziale będziemy stosowali operacje elementarne (zarówno wierszowe, jak i kolumnowe) nie tylko do macierzy kwadratowych, ale również do prostokątnych. Będziemy potrzebowali następującego pojęcia:

Fakt 2.35

Dla dowolnych odwracalnych macierzy kwadratowych A i B (niekoniecznie tego samego rozmiaru) zachodzi:

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{pmatrix}$$

Dowód. Nietrudno zauważyć, że jeśli $A, A' \in M_{m \times m}$ oraz $B, B' \in M_{n \times n}$, to

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \cdot \begin{pmatrix} A' & 0 \\ 0 & B' \end{pmatrix} = \begin{pmatrix} AA' & 0 \\ 0 & BB' \end{pmatrix}$$

Stąd

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \cdot \begin{pmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} = I$$

□

Definicja 2.36

Rzędem (kolumnowym) macierzy $A \in M_{m \times n}$ nazywamy wymiar podprzestrzeni $\mathbb{R}^n$ rozpiętej przez kolumny A .

Rzędem (wierszowym) macierzy A nazywamy wymiar podprzestrzeni $\mathbb{R}^m$ rozpiętej przez wiersze A .

Powyższą definicję można wyrazić również następująco: rząd wierszowy (kolumnowy) macierzy to maksymalna liczba liniowo niezależnych wierszy (kolumn) tej macierzy. Okazuje się, że rząd wierszowy jest tym samym co rząd kolumnowy (który to fakt zostawimy bez dowodu):

Fakt 2.37

Dla każdej macierzy prostokątnej A , rząd wierszowy i rząd kolumnowy są równe. Każdą z tych liczb nazywamy *rzędem* macierzy A i oznaczamy: $\text{rank } A$.

Wniosek 2.38

Jeśli $A \in M_{m \times n}$, to $\text{rank } A \leq m$ oraz $\text{rank } A \leq n$.

Dowód. Rząd macierzy jest równy zarówno jej rzędowi kolumnowemu, jak i rzędowi wierszowemu, czyli nie przekracza liczby jej wierszy, ani liczby jej kolumn. $\square$

Definicja 2.39

Minorem stopnia k macierzy A nazywamy wyznacznik dowolnej macierzy $k \times k$ powstałej przez usunięcie pewnej liczby wierszy i kolumn A .

Przykład 5

Jednym z minorów stopnia 3 macierzy

$$\begin{pmatrix} 3 & 7 & 1 & 0 & 2 \\ 2 & 8 & 1 & 0 & 2 \\ 1 & 1 & 5 & 1 & 2 \\ 8 & 2 & 4 & 1 & 3 \end{pmatrix}$$

jest minor

$$\begin{vmatrix} 7 & 0 & 2 \\ 8 & 0 & 2 \\ 2 & 1 & 3 \end{vmatrix} = 2$$

Fakt 2.40

Rząd macierzy (prostokątnej) A to największy stopień niezerowego minora tej macierzy.

Dowód. Wystarczy pokazać, że:

$$\text{rank } A \geq k \iff \text{macierz } A \text{ ma niezerowy minor stopnia } k$$

Ponieważ rząd macierzy to maksymalna liczba liniowo niezależnych kolumn, więc $\text{rank } A \geq k$ wtedy i tylko wtedy, gdy macierz A ma k liniowo niezależnych kolumn $A_{i_1}, \dots, A_{i_k}$, tzn. $\text{rank}(A_{i_1}, \dots, A_{i_k}) = k$. Z kolei macierz $(A_{i_1}, \dots, A_{i_k})$ ma rząd k wtedy i tylko wtedy, gdy ma k liniowo niezależnych wierszy, czyli można z niej wybrać niezerowy minor stopnia k . Otrzymany minor jest równocześnie minorem macierzy A , co kończy dowód. $\square$

Wektory $v_1, \dots, v_k \in \mathbb{R}^n$ są liniowo niezależne wtedy i tylko wtedy, gdy istnieje niezerowy minor stopnia k macierzy $(v_1, \dots, v_k)$.

Wniosek 2.41 ma wersję dotyczącą wektorów dowolnej przestrzeni liniowej:

Niech V będzie skończenie wymiarową przestrzenią liniową, zaś $\mathcal{B}$ jej dowolną bazą. Wówczas wektory $v_1, \dots, v_k \in V$ są liniowo niezależne wtedy i tylko wtedy, gdy istnieje niezerowy minor stopnia k macierzy $([v_1]_{\mathcal{B}}, \dots, [v_k]_{\mathcal{B}})$.

Operacje elementarne (zarówno wierszowe jak i kolumnowe) nie zmieniają rzędu macierzy.

2.3 Układy równań liniowych

Jednorodnym układem m równań liniowych z niewiadomymi $x_1, \dots, x_n$ nazywamy następujący układ (gdzie a_{ij} to dowolne liczby rzeczywiste):

[illegible]

Niejednorodnym układem m równań liniowych z niewiadomymi $x_1, \dots, x_n$ nazywamy następujący układ (gdzie a_{ij} oraz b_i to dowolne liczby rzeczywiste):

[illegible]

Układ (2.9) nazywamy układem jednorodnym *związany* z układem niejednorodnym (2.10). W obu przypadkach macierz $A = (a_{ij})$ nazywamy *macierzą główną układu*. Macierz $(A|b)$ (macierz A z dopisaną kolumną wyrazów wolnych) nazywamy *macierzą rozszerzoną układu* (2.10).

Przykład 2

Opisać warstwy podprzestrzeni $W = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : 3x - 4y + z = 0 \right\} < \mathbb{R}^3$.

Rozwiązanie. Przestrzeń W to płaszczyzna o równaniu parametrycznym $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = t \begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix}$. Każda warstwa tej podprzestrzeni to zbiór punktów postaci

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = t \begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

czyli przesunięcie równoległe płaszczyzny W o wektor $v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$. Zamieniając na równanie ogólne otrzymujemy płaszczyznę o równaniu $3x - 4y + z + D = 0$ (dla każdej wartości D otrzymujemy inną warstwę).

Przykład 3

Opisać warstwy podprzestrzeni $W = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \frac{x}{2} = \frac{y}{3} = \frac{z}{5} = 0 \right\} < \mathbb{R}^3$.

Rozwiązanie. Przestrzeń W to prosta o równaniu kierunkowym $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = t \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}$. Każda warstwa tej podprzestrzeni to zbiór punktów postaci $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = t \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} + \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, czyli przesunięcie równoległe prostej W o wektor $v = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$. Zamieniając na postać kierunkową otrzymujemy prostą $\frac{x-a}{2} = \frac{y-b}{3} = \frac{z-c}{5} = 0$.

Przykład 4

Opisać warstwy podprzestrzeni $W = \{f : \mathbb{R} \rightarrow \mathbb{R} : f \text{ stała}\} < C(\mathbb{R})$

Rozwiązanie. Każda warstwa przestrzeni funkcji stałych jest postaci $f(x) + C$ (tzn. ustalona funkcja f + dowolna funkcja stała). Jest to dobrze znany przykład z analizy, gdzie całka nieoznaczona funkcji nie jest **funkcją**, tylko **warstwą podprzestrzeni funkcji stałych** (tzn. funkcją z dokładnością do dodania stałej). Zmienną C używaną przy wyliczaniu całki nieoznaczonej należy w tym kontekście interpretować nie jako liczbę, ale jako przestrzeń liniową (przestrzeń funkcji stałych).

Przykład 5

Szczególным przypadkiem warstwy podprzestrzeni $W < V$ jest sama podprzestrzeń W , którą traktujemy jako warstwę $0 + W$ (czyli W przesunięta o wektor zerowy).

Przykład 6

Opisać warstwy podprzestrzeni $\{0\} < V$ oraz $V < V$.

Rozwiązanie. Warstwa $v + \{0\}$ to zbiór $\{v + 0\}$, czyli zbiór jednopunktowy. Dla każdego elementu $v \in V$ otrzymujemy zatem warstwę $\{v\}$. Warstwa podprzestrzeni V jest tylko jedna i jest nią V .

Interpretacja warstw podprzestrzeni $W < V$ jako „równoległych przesunięć” podprzestrzeni W wyjaśnia dlaczego warstwy są parami rozłączne i łącznie pokrywają całą przestrzeń V . Formalizuje to poniższy fakt:

Fakt 2.47

Każdy element $v \in V$ należy do dokładnie jednej warstwy W (jest to warstwa $v + W$).

Fakty 2.45 i 2.48 pokazują, że niejednorodny układ równań liniowych można rozwiązywać według następującego algorytmu:

1. Znaleźć przestrzeń $\ker A$ rozwiązań układu jednorodnego związanego z badanym układem niejednorodnym.
2. Znaleźć jakiegokolwiek (jedno) rozwiązanie v badanego układu niejednorodnego.
3. Zbiorem rozwiązań badanego układu niejednorodnego będzie warstwa $v + \ker A$.

Algorytm ten bywa wygodny w sytuacji, gdy mamy rozwiązać dużą liczbę układów równań liniowych różniących się jedynie prawymi stronami.

Przykład 7

Rozwiązać następujące układy równań:

$$\begin{cases} 2x_1 + 3x_2 - x_3 + x_4 = 1 \\ 3x_1 + x_2 - x_3 + 2x_4 = 2 \end{cases} \quad \begin{cases} 2x_1 + 3x_2 - x_3 + x_4 = 3 \\ 3x_1 + x_2 - x_3 + 2x_4 = 1 \end{cases} \quad \begin{cases} 2x_1 + 3x_2 - x_3 + x_4 = 2 \\ 3x_1 + x_2 - x_3 + 2x_4 = 2 \end{cases}$$

Rozwiązanie. Rozwiązujemy układ jednorodny związany z powyższymi układami niejednorodnymi:

$$\begin{cases} 2x_1 + 3x_2 - x_3 + x_4 = 0 \\ 3x_1 + x_2 - x_3 + 2x_4 = 0 \end{cases}$$

otrzymując rozwiązanie:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} t \\ s \\ 5s + t \\ 2s - t \end{pmatrix}$$

Zgadujemy jakiegokolwiek rozwiązanie szczególne każdego z układów równań, np. (odpowiednio) $\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$. Stąd dostajemy rozwiązania ogólne każdego z układów (jako różne warstwy tej samej podprzestrzeni):

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} t \\ s \\ 5s + t \\ 2s - t + 1 \end{pmatrix} \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} t \\ s + 1 \\ 5s + t \\ 2s - t \end{pmatrix} \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} t + 1 \\ s + 1 \\ 5s + t + 1 \\ 2s - t + 1 \end{pmatrix}$$

Fakt 2.49

Dana jest przestrzeń liniowa V oraz jej podprzestrzeń $W < V$. Na zbiorze wszystkich warstw podprzestrzeni W możemy wprowadzić działania:

$$(v + W) + (v' + W) = (v + v') + W \quad \text{oraz} \quad \alpha(v + W) = \alpha v + W$$

(tzn. dodanie dwóch warstw to dodanie ich dowolnie wybranych reprezentantów, a mnożenie warstwy przez skalar, to mnożenie przez skalar dowolnego reprezentanta). Zbiór warstw z tak określonymi działaniami tworzy przestrzeń liniową zwaną *przestrzenią ilorazową* i oznaczaną V/W . Wektorem zerowym przestrzeni ilorazowej jest warstwa $0 + W$, czyli W .

Powyższe stwierdzenie podajemy jako fakt, a nie jako definicję, gdyż sprawdzenia wymagają dwie rzeczy:

- podane działania są dobrze określone, tzn. wynik dodawania dwóch warstw (mnożenia warstwy przez skalar) nie zależy od wyboru reprezentantów (reprezentanta);
- działania spełniają aksjomaty przestrzeni liniowej z Definicji 1.1.

Sprawdzenie powyższych elementów pozostawimy czytelnikowi.

Przykład 8

Opisać przestrzeń ilorazową $\mathbb{R}^2/W$, gdzie $W = \{(x, y) \in \mathbb{R}^2 : 2x + 3y = 0\}$.

Rozwiązanie. Elementy przestrzeni ilorazowej to warstwy podprzestrzeni W , czyli zgodnie z Przykładem 1 proste o równaniach parametrycznych $\begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} 3 \\ -2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$. Suma warstw $\begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} 3 \\ -2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ oraz $\begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} 3 \\ -2 \end{pmatrix} + \begin{pmatrix} v'_1 \\ v'_2 \end{pmatrix}$ to warstwa $\begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} 3 \\ -2 \end{pmatrix} + \begin{pmatrix} v_1+v'_1 \\ v_2+v'_2 \end{pmatrix}$, a iloczyn warstwy $\begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} 3 \\ -2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ przez liczbę α to $\begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} 3 \\ -2 \end{pmatrix} + \begin{pmatrix} \alpha v_1 \\ \alpha v_2 \end{pmatrix}$.

Przykład 9

Całka nieoznaczona to przekształcenie liniowe: $F : C(\mathbb{R}) \rightarrow C(\mathbb{R})/C$, gdzie $C \subset C(\mathbb{R})$ to podprzestrzeń funkcji stałych. Innymi słowy, całka nieoznaczona każdej funkcji ciągłej f przyporządkowuje **warstwę** podprzestrzeni funkcji stałych, np. $F(\cos x) = \sin x + C$, gdzie C to podprzestrzeń (a nie liczba).

Fakt 2.50

Niech $W < V$ będzie podprzestrzenią przestrzeni liniowej V . Wówczas wymiar przestrzeni ilorazowej V/W wynosi:

$$\dim V/W = \dim V - \dim W \quad (2.12)$$

Dowód. Zauważmy, że przekształcenie $F : V \rightarrow V/W$ zdefiniowane jako $F(v) = v + W$ (przypisujące wektorowi $v \in V$ warstwę podprzestrzeni W zawierającą v) jest przekształceniem liniowym. Przekształcenie F jest „na” (czyli $\text{Im} F = V/W$) oraz $\ker F = W$, więc zgodnie z twierdzeniem o indeksie:

$$\dim V = \dim \ker F + \dim \text{Im} F = \dim W + \dim V/W$$

co dowodzi wzoru (2.12) □

Niniejszy rozdział pokazuje metodę rozwiązywania dowolnego układu równań liniowych. Ponieważ w praktyce często nie jest potrzebne pełne rozwiązanie, a jedynie informacja o istnieniu i liczbie rozwiązań, podzielimy zadanie na trzy problemy:

Pytanie 1 Czy układ jest niesprzeczny (tzn. ma przynajmniej jedno rozwiązanie)?

Pytanie 2 Ile rozwiązań ma układ? Ponieważ niesprzeczny układ równań liniowych ma albo 1 albo ∞ rozwiązań, więc precyzyjniejsze pytanie brzmi: Jaki jest wymiar przestrzeni (lub warstwy) wszystkich rozwiązań?

Pytanie 3 Jakie wygląda rozwiązanie ogólne układu?

Dotychczasowy rezultat (wzory Cramera) daje częściową odpowiedź w odniesieniu do układów n równań z n niewiadomymi. Oznaczając macierz główną takiego układu przez A , wiemy, że:

1. układ jest niesprzeczny $\iff \det A \neq 0$ (nie wiemy co w przypadku $\det A = 0$);
2. układ ma 1 rozwiązanie (0-wymiarowa przestrzeń/warstwa), gdy $\det A \neq 0$ (w przypadku $\det A = 0$ układ jest sprzeczny lub ma ∞ rozwiązań);

Dowód. Zgodnie z Faktem 2.48 zbiór rozwiązań jest warstwą przestrzeni $\ker A$. Zgodnie z twierdzeniem o indeksie zastosowanym do przekształcenia $F_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$:

$$n = \dim \ker F_A + \dim \operatorname{Im} F_A = \dim \ker A + \operatorname{rank} A$$

czyli

$$\dim \ker A = n - \operatorname{rank} A$$

Zbiór rozwiązań układu jest warstwą przestrzeni $\ker A$, która, jak wynika z powyższego rachunku, jest przestrzenią wymiaru $n - \operatorname{rank} A$. $\square$

Zwróćmy uwagę, że powyższy rachunek odnosi się wyłącznie do układu, o którym uprzednio przekonaliśmy się (np. przez zastosowanie Twierdzenia Kroneckera-Capellego), że jest niesprzeczny.

Przykład 12

Wyznacz wymiar zbioru rozwiązań następującego układu równań liniowych:

$$\begin{cases} x_1 + x_2 + 2x_3 + 3x_4 + 3x_5 = 1 \\ 2x_1 + 3x_2 + 3x_3 + 5x_4 + 6x_5 = 2 \\ 2x_1 + 5x_2 + x_3 + 3x_4 + 6x_5 = 2 \end{cases}$$

Rozwiązanie. Zaczynamy od ustalenia rzędów macierzy głównej i macierzy rozszerzonej układu, aby ustalić czy układ jest niesprzeczny:

$$\begin{aligned} \operatorname{rank} \begin{pmatrix} 1 & 1 & 2 & 3 & 3 \\ 2 & 3 & 3 & 5 & 6 \\ 2 & 5 & 1 & 3 & 6 \end{pmatrix} &= \operatorname{rank} \begin{pmatrix} 1 & 1 & 2 & 3 & 3 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 3 & -3 & -3 & 0 \end{pmatrix} = \operatorname{rank} \begin{pmatrix} 1 & 1 & 2 & 3 & 3 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = 2 \\ \operatorname{rank} \begin{pmatrix} 1 & 1 & 2 & 3 & 3 & | & 1 \\ 2 & 3 & 3 & 5 & 6 & | & 2 \\ 2 & 5 & 1 & 3 & 6 & | & 2 \end{pmatrix} &= \operatorname{rank} \begin{pmatrix} 1 & 1 & 2 & 3 & 3 & | & 1 \\ 0 & 1 & -1 & -1 & 0 & | & 0 \\ 0 & 3 & -3 & -3 & 0 & | & 0 \end{pmatrix} = \operatorname{rank} \begin{pmatrix} 1 & 1 & 2 & 3 & 3 & | & 1 \\ 0 & 1 & -1 & -1 & 0 & | & 0 \\ 0 & 0 & 0 & 0 & 0 & | & 0 \end{pmatrix} = 2 \end{aligned}$$

Ponieważ $\operatorname{rank} A = \operatorname{rank}(A|b) = 2$, więc zgodnie z Twierdzeniem 2.51 układ jest niesprzeczny. Wobec tego zgodnie z Twierdzeniem 2.53 wymiar zbioru (warstwy) rozwiązań wynosi:

$$5 - \operatorname{rank} A = 5 - 2 = 3$$

Odpowiedzi na Pytanie 3 dostarczy metoda eliminacji Gaussa, opisana w kolejnym twierdzeniu. Przed jej wprowadzeniem potrzebujemy opisać macierze pewnej szczególnej postaci.

Definicja 2.54

Macierz (prostokątną) nazywamy *macierzą schodkową*, jeśli każdy wiersz tej macierzy zaczyna się większą liczbą zer niż wiersz poprzedni (za wyjątkiem pierwszego wiersza, który nie ma poprzednika oraz wierszy złożonych z samych zer, których dowolna liczba może znajdować się na końcu macierzy). Pierwszy niezerowy wyraz w każdym niezerowym wierszu nazywamy *wyrazem wiodącym*.

szerzonej układu do postaci schodkowej. Zaznaczone na czerwono wyrazy, to wyrazy wiodące:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 2 & 3 & 3 & 1 \\ 2 & 3 & 3 & 5 & 6 & 2 \\ 2 & 5 & 1 & 3 & 6 & 2 \end{array}\right) \rightarrow \left(\begin{array}{ccccc|c} 1 & 1 & 2 & 3 & 3 & 1 \\ 0 & 1 & -1 & -1 & 0 & 0 \\ 0 & 3 & -3 & -3 & 0 & 0 \end{array}\right) \rightarrow \left(\begin{array}{ccccc|c} \color{red}{1} & 1 & 2 & 3 & 3 & 1 \\ 0 & \color{red}{1} & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Następnie zapisujemy nowy układ równań, dla którego otrzymana macierz to macierz rozszerzona układu:

$$\begin{cases} \color{red}{x_1} + x_2 + 2x_3 + 3x_4 + 3x_5 = 1 \\ \color{red}{x_2} - x_3 - x_4 = 0 \\ 0 = 0 \end{cases}$$

Z każdego z równań wyliczamy zaznaczoną zmienną zależną (poczynając od ostatniej, a kończąc na pierwszej):

$$\begin{cases} \color{red}{x_2} = x_3 + x_4 \\ \color{red}{x_1} = -x_2 - 2x_3 - 3x_4 - 3x_5 + 1 = -3x_3 - 4x_4 - 3x_5 + 1 \end{cases}$$

Stąd (przyjmując zmienne wolne x_3, x_4 i x_5 za parametry s, t, v) otrzymujemy rozwiązanie ogólne:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -3s - 4t - 3v + 1 \\ s + t \\ s \\ t \\ v \end{pmatrix} = s \begin{pmatrix} -3 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -4 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + v \begin{pmatrix} -3 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Zwróćmy uwagę, że do zastosowania metody eliminacji Gaussa nie trzeba wcześniej sprawdzać czy układ jest niesprzeczny. Gdyby układ był niesprzeczny, podana metoda wykaże to w kroku 2, tak jak w poniższym przykładzie:

Przykład 15

Znaleźć rozwiązanie ogólne następującego układu równań:

$$\begin{cases} x_1 + x_2 + 2x_3 + 3x_4 + 3x_5 = 1 \\ 2x_1 + 3x_2 + 3x_3 + 5x_4 + 6x_5 = 2 \\ 2x_1 + 5x_2 + x_3 + 3x_4 + 6x_5 = 4 \end{cases}$$

Rozwiązanie. Zgodnie z metodą eliminacji Gaussa zaczynamy od sprowadzenia macierzy rozszerzonej układu do postaci schodkowej:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 2 & 3 & 3 & 1 \\ 2 & 3 & 3 & 5 & 6 & 2 \\ 2 & 5 & 1 & 3 & 6 & 4 \end{array}\right) \rightarrow \left(\begin{array}{ccccc|c} 1 & 1 & 2 & 3 & 3 & 1 \\ 0 & 1 & -1 & -1 & 0 & 0 \\ 0 & 3 & -3 & -3 & 0 & 2 \end{array}\right) \rightarrow \left(\begin{array}{ccccc|c} 1 & 1 & 2 & 3 & 3 & 1 \\ 0 & 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{array}\right)$$

Następnie zapisujemy nowy układ równań, dla którego otrzymana macierz to macierz rozszerzona układu:

$$\begin{cases} x_1 + x_2 + 2x_3 + 3x_4 + 3x_5 = 1 \\ x_2 - x_3 - x_4 = 0 \\ 0 = 2 \end{cases}$$

Otrzymaliśmy układ sprzeczny, a zatem wyjściowy układ równań również jest sprzeczny.

Definicja 2.56

Fundamentalnym układem rozwiązań układu jednorodnego nazywamy bazę przestrzeni rozwiązań tego układu (tzn. bazę $\ker A$, gdzie A jest macierzą główną układu).

Przykład 16

Znaleźć fundamentalny układ rozwiązań następującego jednorodnego układu równań:

$$\begin{cases} x_1 + x_2 + 2x_3 + 3x_4 + 3x_5 = 0 \\ 2x_1 + 3x_2 + 3x_3 + 5x_4 + 6x_5 = 0 \\ 2x_1 + 5x_2 + x_3 + 3x_4 + 6x_5 = 0 \end{cases}$$

Rozwiązanie. Postępując podobnie jak w poprzednich przykładach wyznaczamy:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -3s - 4t - 3v \\ s + t \\ s \\ t \\ v \end{pmatrix} = s \begin{pmatrix} -3 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -4 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + v \begin{pmatrix} -3 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

czyli fundamentalny układ rozwiązań stanowią wektory:

$$\begin{pmatrix} -3 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -4 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -3 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Fakt 2.57

Każdą k -wymiarową podprzestrzeń $\mathbb{R}^n$ można opisać przy pomocy jednorodnego układu $n - k$ równań liniowych.

Pominiemy dowód powyższego faktu, pokazując zamiast tego jego działanie na konkretnym przykładzie. Zwróćmy jedynie uwagę, że w przypadku układu z n niewiadomymi, składającego się z $n - k$ liniowo niezależnych równań przestrzeń rozwiązań ma wymiar $n - (n - k) = k$.

Definicja 2.58

Podprzestrzeń $W < V$ przestrzeni liniowej V generowaną przez wektory $w_1, \dots, w_k \in W$ oznaczamy:

$$W = \text{Lin}\{w_1, \dots, w_k\}$$

(tzn. $\text{Lin}\{w_1, \dots, w_k\}$ to zbiór wszystkich kombinacji liniowych wektorów $w_1, \dots, w_k$).

Przykład 17

Opisz przy pomocy układu równań liniowych podprzestrzeń $W = \text{Lin}\left\{\begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 3 \\ 2 \\ 1 \end{pmatrix}\right\}$ przestrzeni $\mathbb{R}^5$.

Rozwiązanie. Podane wektory są liniowo niezależne (co można sprawdzić znajdując dla poniż-

szej macierzy niezerowy minor stopnia 3):

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

wobec czego $\dim W = 3$. Zgodnie z Faktem 2.57 podprzestrzeń W można opisać przy pomocy jednorodnego układu $5 - 3 = 2$ równań liniowych. Każde z tych równań jest postaci:

$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 = 0$$

i musi być spełnione przez każdy z wektorów:

$$\begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 3 \\ 2 \\ 1 \end{pmatrix}$$

Współczynniki każdego szukanego równania spełniają zatem warunki:

$$\begin{cases} a_1 + 2a_2 + a_3 + a_4 + 3a_5 = 0 \\ a_1 + a_2 - a_3 + a_4 + a_5 = 0 \\ a_2 + 3a_3 + 2a_4 + a_5 = 0 \end{cases}$$

Potrzebujemy znaleźć 2 (liniowo niezależne) rozwiązania powyższego układu – otrzymamy w ten sposób współczynniki 2 szukanых równań. Rozwiązujemy powyższy układ metodą eliminacji Gaussa:

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 3 & 0 \\ 1 & 1 & -1 & 1 & 1 & 0 \\ 0 & 1 & 3 & 2 & 1 & 0 \end{array} \right) \rightarrow \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 3 & 0 \\ 0 & -1 & -2 & 0 & -2 & 0 \\ 0 & 1 & 3 & 2 & 1 & 0 \end{array} \right) \rightarrow \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 3 & 0 \\ 0 & -1 & -2 & 0 & -2 & 0 \\ 0 & 0 & 1 & 2 & -1 & 0 \end{array} \right)$$

Otrzymujemy układ:

$$\begin{cases} a_1 + 2a_2 + a_3 + a_4 + 3a_5 = 0 \\ -a_2 - 2a_3 - 2a_5 = 0 \\ a_3 + 2a_4 - a_5 = 0 \end{cases} \quad \text{czyli} \quad \begin{cases} a_3 = -2a_4 + a_5 \\ a_2 = -2a_3 - 2a_5 = 4a_4 - 4a_5 \\ a_1 = -2a_2 - a_3 - a_4 - 3a_5 = -7a_4 + 4a_5 \end{cases}$$

Stąd

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix} = \begin{pmatrix} -7s + 4t \\ 4s - 4t \\ -2s + t \\ s \\ t \end{pmatrix} = s \begin{pmatrix} -7 \\ 4 \\ -2 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 4 \\ -4 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

Fundamentalnym układem rozwiązań są wektory $\begin{pmatrix} -7 \\ 4 \\ -2 \\ 1 \\ 0 \end{pmatrix}$ oraz $\begin{pmatrix} 4 \\ -4 \\ 1 \\ 0 \\ 1 \end{pmatrix}$, które dają szukane dwa równania opisujące podprzestrzeń W :

$$\begin{cases} -7x_1 + 4x_2 - 2x_3 + x_4 = 0 \\ 4x_1 - 4x_2 + x_3 + x_5 = 0 \end{cases}$$

Rozdział 3

Diagonalizacja macierzy

W tym rozdziale będziemy zajmować się wyłącznie przekształceniami liniowymi $F : V \rightarrow V$ (dla których dziedzina i przeciwdziedzina są jednakowe). Przekształcenia takie nazywamy *endomorfizmami* (przypomnijmy, że dowolne przekształcenie liniowe nazywamy *homomorfizmem*). Odwracalny endomorfizm nazywamy *automorfizmem* (przypomnijmy, że dowolne odwracalne przekształcenie liniowe nazywamy *izomorfizmem*).

3.1 Wartości i wektory własne (rzeczywiste)

W niniejszym podrozdziale będziemy rozpatrywać jedynie przypadek rzeczywistych wartości własnych. Przypadek zespolonych wartości własnych rozpatrzymy w kolejnym podrozdziale.

Definicja 3.1

Niech V będzie przestrzenią liniową. *Wektor własny* przekształcenia liniowego $F : V \rightarrow V$ to wektor v spełniający warunek:

$$F(v) = \lambda \cdot v$$

dla pewnej liczby rzeczywistej λ . Jeśli $v \neq 0$, to liczbę λ nazywamy *wartością własną* przekształcenia F , zaś v *wektorem własnym dla wartości własnej λ* . Wektor $v = 0$ jest wektorem własnym dla każdej wartości własnej λ . Zbiór wszystkich wektorów własnych dla wartości własnej λ nazywamy *przestrzenią własną* dla λ i oznaczamy V^λ .

Każda macierz $A \in M_{n \times n}$ wyznacza przekształcenie liniowe

$$F_A : \mathbb{R}^n \rightarrow \mathbb{R}^n, \text{ gdzie } F_A(X) = A \cdot X$$

Dla uproszczenia zapisu, wartości własne i wektory własne przekształcenia F_A będziemy nazywać wartościami własnymi i wektorami własnymi macierzy A .

Przykład 1

Znajdź wartości własne, wektory własne i przestrzeń własną przekształcenia liniowego $F : M_{2 \times 2} \rightarrow M_{2 \times 2}$ zdefiniowanego $F(A) = A^\top$.

Rozwiązanie. Szukamy takich macierzy $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, że dla pewnego $\lambda \in \mathbb{R}$ zachodzi warunek $A^\top = \lambda A$, czyli

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \lambda \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{tzn.} \quad \begin{cases} a = \lambda a \\ b = \lambda c \\ c = \lambda b \\ d = \lambda d \end{cases}$$

Rozwiązując powyższy układ równań otrzymujemy:

- wektory własne $\begin{pmatrix} a & b \\ b & d \end{pmatrix}$ dla wartości własnej $\lambda = 1$ (tzn. $V^1 = \left\{ \begin{pmatrix} a & b \\ b & d \end{pmatrix} : a, b, d \in \mathbb{R} \right\}$),
- wektory własne $\begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}$ dla wartości własnej $\lambda = -1$ (tzn. $V^{-1} = \left\{ \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} : b \in \mathbb{R} \right\}$).

Przykład 2

Znajdź wartości własne, wektory własne i przestrzenie własne przekształcenia liniowego $F : \mathbb{R}_2[x] \rightarrow \mathbb{R}_2[x]$ zdefiniowanego $F(P)(x) = xP'(x)$.

Rozwiązanie. Jeśli oznaczmy $P(x) = ax^2 + bx + c$, to $F(P)(x) = 2ax^2 + bx$. Wektorem własnym jest wielomian spełniający dla pewnej niezerowej liczby λ warunek:

$$F(ax^2 + bx + c) = \lambda(ax^2 + bx + c) \quad \text{czyli} \quad 2ax^2 + bx = \lambda ax^2 + \lambda bx + \lambda c$$

Powyższa równość to równość wielomianów, skąd

$$\begin{cases} 2a = \lambda a \\ b = \lambda b \\ 0 = \lambda c \end{cases}$$

Rozwiązując powyższy układ otrzymujemy:

- wektory własne $P(x) = c$ dla wartości własnej $\lambda = 0$ (tzn. $V^0 = \{c : c \in \mathbb{R}\}$),
- wektory własne $P(x) = bx$ dla wartości własnej $\lambda = 1$ (tzn. $V^1 = \{bx : b \in \mathbb{R}\}$),
- wektory własne $P(x) = ax^2$ dla wartości własnej $\lambda = 2$ (tzn. $V^2 = \{ax^2 : a \in \mathbb{R}\}$).

Przykład 3

Znajdź wartości własne, wektory własne i przestrzenie własne przekształcenia $F : \mathbb{C} \rightarrow \mathbb{C}$ zdefiniowanego $F(z) = \bar{z}$.

Rozwiązanie. Szukamy takich liczb $a, b, \lambda \in \mathbb{R}$, że

$$a - bi = \lambda(a + bi) \quad \text{czyli} \quad a - bi = (\lambda a) + (\lambda b)i$$

Porównujemy części rzeczywiste i urojone obu stron równania:

$$\begin{cases} a = \lambda a \\ -b = \lambda b \end{cases}$$

Rozwiązując powyższy układ otrzymujemy:

- wektory własne $z = a$ dla wartości własnej $\lambda = 1$ (tzn. $V^1 = \{a : a \in \mathbb{R}\}$),
- wektory własne $z = ib$ dla wartości własnej $\lambda = -1$ (tzn. $V^{-1} = \{ib : b \in \mathbb{R}\}$).

Fakt 3.2

Niech $F : V \rightarrow V$ będzie przekształceniem liniowym, a λ wartością własną F . Wówczas przestrzeń własna V^λ jest podprzestrzenią V (tzn. $V^\lambda \leq V$).

Dowód. Musimy pokazać zamkniętość V^λ na dodawanie i mnożenie przez skalar. Dla $v_1, v_2 \in V^\lambda$:

$$F(v_1) = \lambda v_1 \quad \text{oraz} \quad F(v_2) = \lambda v_2$$

Wówczas z addytywności F :

$$F(v_1 + v_2) = F(v_1) + F(v_2) = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2)$$

czyli $v_1 + v_2 \in V^\lambda$. Podobnie sprawdzamy zamkniętość na mnożenie przez skalar. $\square$

Fakt 3.3

Jeśli $\lambda_1 \neq \lambda_2$ są wartościami własnymi przekształcenia liniowego $F : V \rightarrow V$, to

$$V^{\lambda_1} \cap V^{\lambda_2} = \{0\}$$

Dowód. Jeśli $v \in V^{\lambda_1} \cap V^{\lambda_2}$, to $F(v) = \lambda_1 v$ oraz $F(v) = \lambda_2 v$, czyli $\lambda_1 v = \lambda_2 v$, skąd (wobec $\lambda_1 \neq \lambda_2$) dostajemy $v = 0$. $\square$

Poniższy fakt pokazuje praktyczny sposób wyznaczania wartości własnych przekształcenia.

Fakt 3.4

Wartości własne przekształcenia liniowego $F : V \rightarrow V$ to pierwiastki *wielomianu charakterystycznego* tego przekształcenia, czyli wielomianu:

$$\chi_F(\lambda) = \det(m_{\mathcal{B}}^{\mathcal{B}}(F) - \lambda I)$$

gdzie $\mathcal{B}$ jest dowolną bazą V . Wielomian charakterystyczny nie zależy od wyboru bazy $\mathcal{B}$.

Dowód. Niech $\mathcal{B}$ będzie dowolną bazą przestrzeni V (jednakową dla dziedziny i przeciwdziedziny). Wówczas warunek $F(v) = \lambda v$ możemy (zgodnie z Faktem 1.24) zapisać we współrzędnych jako:

$$m_{\mathcal{B}}^{\mathcal{B}}(F) \cdot [v]_{\mathcal{B}} = [\lambda v]_{\mathcal{B}} \quad (3.1)$$

lub, jeśli oznaczymy $A = m_{\mathcal{B}}^{\mathcal{B}}(F)$ oraz $[v]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, w postaci:

$$A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix} \quad \text{czyli} \quad (A - \lambda I) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0 \quad (3.2)$$

Liczba λ jest wartością własną F wtedy i tylko wtedy, gdy równanie (3.2) ma niezerowe rozwiązanie. Równanie to zawsze ma rozwiązanie zerowe, które zgodnie ze wzorem Cramera jest jedynym rozwiązaniem, gdy $\det(A - \lambda I) \neq 0$. W przeciwnym razie równanie ma więcej niż jedno rozwiązanie (w szczególności ma rozwiązanie niezerowe). Wobec tego wartości własne F to liczby λ spełniające warunek:

$$\det(A - \lambda I) = 0 \quad \text{czyli} \quad \det(m_{\mathcal{B}}^{\mathcal{B}}(F) - \lambda I) = 0$$

Z powyższego rozumowania wynika, że pierwiastki wielomianu charakterystycznego nie zależą od bazy, ale to jeszcze nie dowodzi, że sam wielomian nie zależy od bazy (np. mogłoby się zdarzyć, że krotności pierwiastków zależą od bazy). Dlatego przytoczymy dodatkowy argument, pokazujący niezależność wielomianu charakterystycznego od wyboru bazy. Jeśli $\mathcal{B}$ i $\mathcal{C}$ są różnymi bazami przestrzeni V , to zgodnie ze wzorami (1.4) oraz (1.6) zachodzi:

$$m_{\mathcal{C}}^{\mathcal{C}}(F) = m_{\mathcal{C}}^{\mathcal{B}}(\text{id}) \cdot m_{\mathcal{B}}^{\mathcal{B}}(F) \cdot m_{\mathcal{B}}^{\mathcal{C}}(\text{id}) = P \cdot (m_{\mathcal{B}}^{\mathcal{B}}(F)) \cdot P^{-1}$$

gdzie $P = m_{\mathcal{C}}^{\mathcal{B}}(\text{id})$. Wobec tego:

$$\det(m_{\mathcal{C}}^{\mathcal{C}}(F)) = \det(P \cdot m_{\mathcal{B}}^{\mathcal{B}}(F) \cdot P^{-1})$$

czyli na mocy Wniosku 2.31:

$$= \det P \cdot \det m_{\mathcal{B}}^{\mathcal{B}}(F) \cdot \det P^{-1} = \det P \cdot \det m_{\mathcal{B}}^{\mathcal{B}}(F) \cdot \frac{1}{\det P} = \det m_{\mathcal{B}}^{\mathcal{B}}(F)$$

tzn. wyznacznik nie zależy od bazy. □

Przykład 4

Wyznaczyć wartości własne i wektory własne dla przekształcenia $F : \mathbb{R}_3[x] \rightarrow \mathbb{R}_3[x]$ zdefiniowanego $F(P) = P'$.

Rozwiązanie. W bazie $\mathcal{B} = (1, x, x^2, x^3)$ przekształcenie F ma macierz:

$$m_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Wobec tego wielomian charakterystyczny przekształcenia to:

$$\chi_F(\lambda) = \det \begin{pmatrix} -\lambda & 1 & 0 & 0 \\ 0 & -\lambda & 2 & 0 \\ 0 & 0 & -\lambda & 3 \\ 0 & 0 & 0 & -\lambda \end{pmatrix} = \lambda^4$$

Stąd jedyną wartością własną F jest $\lambda = 0$. Wektory własne dla $\lambda = 0$ to wielomiany spełniające warunek $P' = 0$, czyli wielomiany stałe.

Przykład 5

Wyznaczyć wartości własne i wektory własne dla przekształcenia $F : M_{2 \times 2} \rightarrow M_{2 \times 2}$ zdefiniowanego $F(A) = 2A + 3A^{\top}$.

Rozwiązanie. W bazie $\mathcal{B} = ((\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}), (\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}), (\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}), (\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}))$ przekształcenie F ma macierz:

$$m_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 2 & 3 & 0 \\ 0 & 3 & 2 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

Wobec tego wielomian charakterystyczny przekształcenia to:

$$\chi_F(\lambda) = \det \begin{pmatrix} 5 - \lambda & 0 & 0 & 0 \\ 0 & 2 - \lambda & 3 & 0 \\ 0 & 3 & 2 - \lambda & 0 \\ 0 & 0 & 0 & 5 - \lambda \end{pmatrix} = (\lambda - 5)^3(\lambda + 1)$$

czyli wartościami własnymi F są $\lambda_1 = 5$ i $\lambda_2 = -1$. Wektory własne dla λ_1 to macierze spełniające warunek $2A + 3A^\top = 5A$, czyli $A = A^\top$ (macierze symetryczne), zaś wektory własne dla $\lambda_2 = -1$ to macierze spełniające warunek $2A + 3A^\top = -A$, czyli $A = -A^\top$ (macierze antysymetryczne).

Wniosek 3.5

Niech V będzie przestrzenią liniową wymiaru n , zaś $F : V \rightarrow V$ przekształceniem liniowym. Wówczas F ma co najwyżej n rzeczywistych wartości własnych (licząc z krotnościami).

Dowód. Zgodnie z Faktem 3.4 wartości własne F to pierwiastki wielomianu charakterystycznego χ_F , który jest wielomianem stopnia $n = \dim V$. Zgodnie z Zasadniczym Twierdzeniem Algebry, taki wielomian ma dokładnie n pierwiastków zespolonych (licząc z krotnościami), czyli co najwyżej n pierwiastków rzeczywistych. $\square$

Fakt 3.6

Jeśli $v_1, \dots, v_k$ są niezerowymi wektorami własnymi przekształcenia $F : V \rightarrow V$ dla parami różnych wartości własnych $\lambda_1, \dots, \lambda_k \in \mathbb{R}$, to wektory $v_1, \dots, v_k$ są liniowo niezależne.

Dowód. Rozważmy zerującą się kombinację liniową wektorów $v_1, \dots, v_k$:

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0$$

Chcemy udowodnić, że $\alpha_1 = \dots = \alpha_k = 0$. Ponieważ F jest odwzorowaniem liniowym, zaś $v_1, \dots, v_k$ jego wektorami własnymi dla wartości własnych $\lambda_1, \dots, \lambda_k$, więc:

$$0 = F(0) = F(\alpha_1 v_1 + \cdots + \alpha_k v_k) = \alpha_1 F(v_1) + \cdots + \alpha_k F(v_k) = \alpha_1 \lambda_1 v_1 + \cdots + \alpha_k \lambda_k v_k$$

czyli

$$\alpha_1 \lambda_1 v_1 + \cdots + \alpha_k \lambda_k v_k = 0$$

Podobnie:

$$0 = F(0) = F(\alpha_1 \lambda_1 v_1 + \cdots + \alpha_k \lambda_k v_k) = \alpha_1 \lambda_1 F(v_1) + \cdots + \alpha_k \lambda_k F(v_k) = \alpha_1 \lambda_1^2 v_1 + \cdots + \alpha_k \lambda_k^2 v_k$$

Kontynuując to postępowanie dostajemy układ równań (wektorowych):

[illegible]

Jeśli zapiszemy współrzędne wektorów $v_1, \dots, v_k$ w dowolnie wybranej bazie $\mathcal{B}$:

$$[v_1]_{\mathcal{B}} = \begin{pmatrix} v_{11} \\ \vdots \\ v_{n1} \end{pmatrix}, \quad [v_2]_{\mathcal{B}} = \begin{pmatrix} v_{12} \\ \vdots \\ v_{n2} \end{pmatrix}, \quad \dots \quad [v_k]_{\mathcal{B}} = \begin{pmatrix} v_{1k} \\ \vdots \\ v_{nk} \end{pmatrix}$$

to układ (3.3) w wersji macierzowej będzie miał postać:

$$\begin{pmatrix} \alpha_1 v_{11} & \alpha_2 v_{12} & \cdots & \alpha_k v_{1k} \\ \alpha_1 v_{21} & \alpha_2 v_{22} & \cdots & \alpha_k v_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1 v_{n1} & \alpha_2 v_{n2} & \cdots & \alpha_k v_{nk} \end{pmatrix} \cdot \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{k-1} \\ 1 & \lambda_3 & \lambda_3^2 & \cdots & \lambda_3^{k-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_k & \lambda_k^2 & \cdots & \lambda_k^{k-1} \end{pmatrix} = 0$$

Ponieważ druga macierz jest kwadratową macierzą odwracalną (jest to transponowana macierz Vandermonde'a), więc:

$$\begin{pmatrix} \alpha_1 v_{11} & \alpha_2 v_{12} & \cdots & \alpha_k v_{1k} \\ \alpha_1 v_{21} & \alpha_2 v_{22} & \cdots & \alpha_k v_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1 v_{n1} & \alpha_2 v_{n2} & \cdots & \alpha_k v_{nk} \end{pmatrix} = 0 \cdot \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{k-1} \\ 1 & \lambda_3 & \lambda_3^2 & \cdots & \lambda_3^{k-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_k & \lambda_k^2 & \cdots & \lambda_k^{k-1} \end{pmatrix}^{-1} = 0$$

skąd (wobec tego, że wektory $v_1, \dots, v_k$ są niezerowe) otrzymujemy $\alpha_1 = \dots = \alpha_k = 0$. $\square$

Twierdzenie 3.7: o diagonalizacji macierzy (wersja 1)

Jeśli macierz $A \in M_{n \times n}$ ma (parami różne) wartości własne $\lambda_1, \dots, \lambda_n$, zaś $v_1, \dots, v_n$ są niezerowymi wektorami własnymi, odpowiednio, dla $\lambda_1, \dots, \lambda_n$, to:

$$A = PDP^{-1}, \text{ gdzie } P = (v_1, \dots, v_n), \quad D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

Dowód. Oznaczmy współrzędne wektorów własnych jako

$$v_1 = \begin{pmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{pmatrix}, \quad v_2 = \begin{pmatrix} v_{12} \\ v_{22} \\ \vdots \\ v_{n2} \end{pmatrix}, \quad \dots, \quad v_n = \begin{pmatrix} v_{1n} \\ v_{2n} \\ \vdots \\ v_{nn} \end{pmatrix}$$

Wówczas

$$P = \begin{pmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{n1} & \cdots & v_{nn} \end{pmatrix}$$

Sprawdzamy rachunkowo, że macierze AP i PD mają jednakowe pierwsze kolumny:

$$AP \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = A \begin{pmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{pmatrix} = Av_1 = \lambda_1 v_1 = P \begin{pmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = PD \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Podobnie sprawdzamy, że drugie, trzecie itd. kolumny macierzy AP i PD są jednakowe. Stąd

$$AP = PD$$

Na mocy Faktu 3.6 wektory $v_1, \dots, v_n$ są liniowo niezależne, czyli macierz P jest odwracalna. Stąd

$$A = PDP^{-1}$$

$\square$

Twierdzenie 3.8: o diagonalizacji macierzy (wersja 2)

Macierz $A \in M_{n \times n}$ można zapisać w postaci diagonalnej:

$$A = PDP^{-1}, \text{ gdzie } P = (v_1, \dots, v_n), \quad D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

wtedy i tylko wtedy, gdy wektory $v_1, \dots, v_n$ tworzą bazę $\mathbb{R}^n$ złożoną z wektorów własnych macierzy A (gdzie $\lambda_1, \dots, \lambda_n$ to odpowiadające tym wektorom wartości własne). Macierz taką nazywamy *diagonalizowalną*.

Dowód. Jedyne miejsce w dowodzie wersji 1 twierdzenia, gdzie korzystaliśmy z założenia o tym, że $v_1, \dots, v_n$ są parami różne była uwaga, że wektory $v_1, \dots, v_n$ są liniowo niezależne (co implikowało odwracalność macierzy P). W związku z tym przy założeniu, że $v_1, \dots, v_n$ tworzą bazę $\mathbb{R}^n$, poprzedni dowód przepisuje się bez żadnych zmian. $\square$

Sprawdzanie, że ciąg wektorów własnych jest bazą znacznie ułatwia następujące uogólnienie Faktu 3.6. Dowód tego faktu wykracza poza ramy niniejszego skryptu.

Fakt 3.9

Jeśli $\lambda_1, \dots, \lambda_k$ to parami różne wartości własne macierzy A , zaś $\mathcal{B}_i$ to baza przestrzeni V^{λ_i} (oczywiście złożona z wektorów własnych), to zbiór $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ jest liniowo niezależnym zbiorem wektorów.

Przykład 6

Zdiagonalizować następujące macierze:

$$\begin{pmatrix} -1 & 1 & 3 \\ -4 & 3 & 4 \\ -2 & 1 & 4 \end{pmatrix} \quad \begin{pmatrix} 3 & 0 & 0 \\ 0 & 5 & 2 \\ 0 & 2 & 5 \end{pmatrix}$$

Rozwiązanie. Wielomian charakterystyczny pierwszej macierzy to $-(\lambda-1)(\lambda-2)(\lambda-3)$. Dla wartości własnych $\lambda_1 = 1$, $\lambda_2 = 2$ i $\lambda_3 = 3$ wyznaczamy przestrzenie wektorów własnych – są to jednowymiarowe przestrzenie rozpinane przez wektory $v_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Stąd dostajemy diagonalizację:

$$\begin{pmatrix} -1 & 1 & 3 \\ -4 & 3 & 4 \\ -2 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}^{-1}$$

Dla drugiej macierzy wyznaczamy wielomian charakterystyczny $-(\lambda-3)^2(\lambda-7)$. Wyznaczamy przestrzenie własne: $V^3 = \text{Lin}\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}\right\}$ oraz $V^7 = \text{Lin}\left\{\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right\}$. Zgodnie z Faktem 3.9 dostajemy diagonalizację:

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 5 & 2 \\ 0 & 2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 7 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix}^{-1}$$

Głównym celem zapisu macierzy w postaci diagonalnej jest ułatwienie potęgowania macierzy, co rozwija poniższy fakt.

Fakt 3.10

Jeśli macierz $A \in M_{n \times n}$ diagonalizuje się, tzn. $A = PDP^{-1}$ dla pewnej odwracalnej macierzy P oraz diagonalnej macierzy $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$, to dla dowolnej liczby naturalnej k zachodzi wzór:

$$A^k = P \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{pmatrix} P^{-1}$$

Dowód. Dla dowolnej macierzy D (nawet nie diagonalnej) zachodzi wzór:

$$(PDP^{-1})^k = \underbrace{PDP^{-1} \cdot PDP^{-1} \cdot \dots \cdot PDP^{-1}}_k = P \underbrace{D \cdot D \cdot \dots \cdot D}_k P^{-1} = PD^k P^{-1}$$

Reszta wynika ze wzoru na potęgowanie macierzy diagonalnej. □

Przykład 7

Obliczyć potęgę macierzy:

$$\begin{pmatrix} -1 & 1 & 3 \\ -4 & 3 & 4 \\ -2 & 1 & 4 \end{pmatrix}^n \quad \begin{pmatrix} 3 & 0 & 0 \\ 0 & 5 & 2 \\ 0 & 2 & 5 \end{pmatrix}^{10}$$

Rozwiązanie. Wykorzystując diagonalizację z poprzedniego przykładu otrzymujemy:

$$\begin{aligned} \begin{pmatrix} -1 & 1 & 3 \\ -4 & 3 & 4 \\ -2 & 1 & 4 \end{pmatrix}^n &= \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1^n & 0 & 0 \\ 0 & 2^n & 0 \\ 0 & 0 & 3^n \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 + 2^{n+1} - 2 \cdot 3^n & 3^n - 2^n & 2 \cdot 3^n - 2^n - 1 \\ 2 - 2 \cdot 3^n & 3^n & 2 \cdot 3^n - 2 \\ 2^{n+1} - 2 \cdot 3^n & 3^n - 2^n & 2 \cdot 3^n - 2^n \end{pmatrix} \\ \begin{pmatrix} 3 & 0 & 0 \\ 0 & 5 & 2 \\ 0 & 2 & 5 \end{pmatrix}^{10} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 3^{10} & 0 & 0 \\ 0 & 3^{10} & 0 \\ 0 & 0 & 7^{10} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 3^{10} & 0 & 0 \\ 0 & 3^{10} + 7^{10} & -3^{10} + 7^{10} \\ 0 & -3^{10} + 7^{10} & 3^{10} + 7^{10} \end{pmatrix} \end{aligned}$$

Zauważmy, że wyrazy macierzy A^n są kombinacjami liniowymi n -tych potęg wartości własnych macierzy A .

Fakt 3.11

Jeśli macierz $A \in M_{n \times n}$ jest diagonalizowalna, a jej wartościami własnymi (niekoniecznie różnymi) są $\lambda_1, \dots, \lambda_n$, to każdy wyraz macierzy A^k jest postaci:

$$\alpha_1 \lambda_1^k + \dots + \alpha_n \lambda_n^k$$

przy czym współczynniki $\alpha_1, \dots, \alpha_n$ nie zależą od wykładnika k .

Dowód. Zgodnie z Faktem 3.10:

$$A^k = P \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{pmatrix} P^{-1}$$

Wymnażając powyższe trzy macierze otrzymujemy macierz o wyrazach postaci $\alpha_1 \lambda_1^k + \dots + \alpha_n \lambda_n^k$, przy czym współczynniki $\alpha_1, \dots, \alpha_n$ zależą wyłącznie od wyrazów macierzy P i P^{-1} , które są niezależne od wykładnika k . $\square$

Definicja 3.12

Krotnością wartości własnej λ przekształcenia liniowego $F : V \rightarrow V$ nazywamy krotność pierwiastka λ jako pierwiastka wielomianu charakterystycznego $\chi_F(\lambda)$.

Fakt 3.13

Niech V będzie skończenie wymiarową przestrzenią liniową. Jeśli λ jest wartością własną przekształcenia liniowego $F : V \rightarrow V$, to:

$$1 \leq \dim V^\lambda \leq \text{krotność wartości własnej } \lambda$$

W szczególności, jeśli λ jest jednokrotną wartością własną, to $\dim V^\lambda = 1$.

Dowód. Niech $\mathcal{B} = (b_1, \dots, b_k)$ będzie bazą przestrzeni V^λ . Zatem $b_1, \dots, b_k$ to wektory liniowo niezależne w V , czyli można je uzupełnić do bazy $\mathcal{B}' = (b_1, \dots, b_k, b_{k+1}, \dots, b_n)$ przestrzeni V . W bazie $\mathcal{B}'$ przekształcenie F ma macierz składającą się z następujących klatek:

$$m_{\mathcal{B}'}^{\mathcal{B}'}(F) = \begin{pmatrix} \lambda I & A \\ 0 & B \end{pmatrix}$$

gdzie $\lambda I \in M_{k \times k}$ oraz $B \in M_{(n-k) \times (n-k)}$ to macierze kwadratowe. Wielomian charakterystyczny powyższej macierzy ma (co najmniej) k -krotny pierwiastek λ , więc krotność λ jest równa przynajmniej $k = \dim V^\lambda$. $\square$

Wniosek 3.14

Niech V będzie skończenie wymiarową przestrzenią liniową. Wówczas przekształcenie liniowe $F : V \rightarrow V$ jest diagonalizowalne (nad liczbami rzeczywistymi) wtedy i tylko wtedy, gdy dla wszystkie pierwiastki wielomianu charakterystycznego χ_F są rzeczywiste oraz dla każdego z nich zachodzi:

$$\text{krotność pierwiastka (wartości własnej) } \lambda = \dim V^\lambda$$

Dowód. Niech $\lambda_1, \dots, \lambda_k$ będą wartościami własnymi F . Wielomian charakterystyczny χ_F ma stopień $n = \dim V$, więc

$$n = \sum_i (\text{krotność } \lambda_i)$$

Ponieważ, zgodnie z Faktem 3.13 dla każdego i zachodzi $(\text{krotność } \lambda_i) \geq \dim V^{\lambda_i}$ więc

$$n \geq \sum_i \dim V^{\lambda_i}$$

Zgodnie z Twierdzeniem 3.8 przekształcenie F jest diagonalizowalne wtedy i tylko wtedy, gdy V ma bazę złożoną z wektorów własnych, czyli:

$$n = \dim V = \sum_i \dim V^{\lambda_i}$$

Wobec tego przekształcenie jest diagonalizowalne wtedy i tylko wtedy, gdy we wszystkich wcześniejszych nierównościach zachodzi równość, czyli $(\text{krotność } \lambda_i) = \dim V^{\lambda_i}$ dla $i = 1, \dots, k$. $\square$

Przykład 8

Wyznaczyć krotności wartości własnych oraz wymiary przestrzeni własnych macierzy

$$\begin{pmatrix} 3 & 2 & 2 & -1 \\ 0 & 5 & 2 & -1 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix} \quad \begin{pmatrix} 3 & 2 & 2 & -1 \\ 0 & 5 & 2 & 1 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

oraz ustalić czy macierze te są diagonalizowalne.

Rozwiązanie. W obu przypadkach wielomian charakterystyczny macierzy to $(\lambda - 3)^2(\lambda - 5)^2$, czyli każda z wartości własnych $\lambda_1 = 3$ i $\lambda_2 = 5$ jest podwójną wartością własną. W pierwszym przypadku otrzymujemy:

$$\dim V^3 = 2 = \text{krotność wartości } 3 \quad \text{oraz} \quad \dim V^5 = 2 = \text{krotność wartości } 5$$

czyli macierz jest diagonalizowalna. W drugim przypadku otrzymujemy:

$$\dim V^3 = 2 = \text{krotność wartości } 3 \quad \text{oraz} \quad \dim V^5 = 1 < \text{krotność wartości } 5$$

czyli macierz nie jest diagonalizowalna.

Twierdzenie 3.15: o diagonalizacji przekształcenia

Niech $\mathcal{B} = (v_1, \dots, v_n)$ będzie bazą skończonej wymiarowej przestrzeni V . Macierz $m_{\mathcal{B}}^{\mathcal{B}}(F)$ przekształcenia liniowego $F : V \rightarrow V$ w bazie $\mathcal{B}$ jest diagonalna wtedy i tylko wtedy, gdy $\mathcal{B}$ jest złożona z wektorów własnych F . Przekształcenie takie nazywamy *diagonalizowalnym*. Ponadto dla dowolnej bazy $\mathcal{C}$ zachodzi:

$$m_{\mathcal{C}}^{\mathcal{C}}(F) = PDP^{-1}$$

dla $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ oraz $P = ([v_1]_{\mathcal{C}}, \dots, [v_n]_{\mathcal{C}})$, gdzie $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ to wartości własne F odpowiadające wektorom własnym $v_1, \dots, v_n \in V$.

Dowód. Skoro wektory bazy $\mathcal{B}$ są wektorami własnymi, to $F(v_i) = \lambda_i v_i$, skąd:

$$m_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = D$$

Zgodnie ze wzorem (1.7) zachodzi:

$$m_{\mathcal{C}}^{\mathcal{C}}(F) = m_{\mathcal{C}}^{\mathcal{B}}(\text{id}) \cdot m_{\mathcal{B}}^{\mathcal{B}}(F) \cdot m_{\mathcal{B}}^{\mathcal{C}}(\text{id}) = PDP^{-1}$$

gdzie $P = m_{\mathcal{C}}^{\mathcal{B}}(\text{id}) = ([v_1]_{\mathcal{C}}, \dots, [v_n]_{\mathcal{C}})$. $\square$

Przykład 9

Zdiagonalizuj macierz przekształcenia $F : M_{2 \times 2} \rightarrow M_{2 \times 2}$ zdefiniowanego $F(A) = A^\top$, w bazie $\mathcal{C} = ((\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}), (\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}), (\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}), (\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}))$.

Rozwiązanie. Przekształcenie to było już analizowane w poprzednich przykładach. Ustaliliśmy, że ma dwie wartości własne $\lambda_1 = 1$ i $\lambda_2 = -1$, a ponadto

$$V^1 = \text{Lin}\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right\} \quad \text{oraz} \quad V^{-1} = \text{Lin}\left\{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right\}$$

Ponieważ $\dim V^1 + \dim V^{-1} = \dim M_{2 \times 2}$, więc przekształcenie jest diagonalizowalne. Zgodnie z Twierdzeniem 3.15:

$$m_{\mathcal{C}}^{\mathcal{C}}(F) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 \end{pmatrix}^{-1}$$

Przykład 10

Ustal czy przekształcenie $F : \mathbb{R}_3[x] \rightarrow \mathbb{R}_3[x]$ zdefiniowane $F(P) = P'$ jest diagonalizowalne.

Rozwiązanie. Przekształcenie to było już analizowane w poprzednich przykładach. Ponieważ ma ono 4-krotną wartość własną 0, ale $\dim V^0 = 1 < 4$, więc nie jest ono diagonalizowalne.

3.2 Wartości i wektory własne (zespolone)

W poprzednim semestrze wykorzystywaliśmy liczby zespolone do diagonalizacji takich macierzy A o wyrazach rzeczywistych, których wielomian charakterystyczny χ_A miał pierwiastki rzeczywiste. Przeprowadzone rozumowania były poprawne rachunkowo, ale nie było jasne skąd ta poprawność się bierze – otrzymywane nierzeczywiste wektory własne nie były elementami rozważanych przestrzeni liniowych $\mathbb{R}^2$ i $\mathbb{R}^3$ (z formalnego punktu widzenia nie były to zatem wektory własne). Rozwiązanie tej kwestii wymaga wprowadzenia pojęcia *zespolonej* przestrzeni liniowej.

Definicja 3.16

Zespoloną przestrzeń liniową lub *przestrzeń liniową nad $\mathbb{C}$* nazywamy zbiór V z dwoma działaniami:

- dodawaniem elementów V ,
- mnożeniem elementów V przez liczby zespolone (które będziemy nazywać *skalarami*)

oraz wyróżnionym elementem 0 (elementem neutralnym dodawania) i przypisaniem każdemu elementowi $v \in V$ elementu $-v \in V$ (element przeciwny) tak, że dla dowolnych wektorów $u, v, w \in V$ oraz skalarów $s, t \in \mathbb{C}$ zachodzą warunki (1)–(7) z Definicji 1.1.

Przykład 1

Przestrzeń $\mathbb{C}^n$ układów n liczb zespolonych $\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$ wraz z dodawaniem i mnożeniem przez skalar

określonymi następująco:

$$\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} + \begin{pmatrix} z'_1 \\ \vdots \\ z'_n \end{pmatrix} = \begin{pmatrix} z_1 + z'_1 \\ \vdots \\ z_n + z'_n \end{pmatrix} \quad \alpha \cdot \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} \alpha \cdot z_1 \\ \vdots \\ \alpha \cdot z_n \end{pmatrix}$$

to zespolona przestrzeń liniowa. Element 0 to $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, a element przeciwny do $\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$ to $\begin{pmatrix} -z_1 \\ \vdots \\ -z_n \end{pmatrix}$.

Przykład 2

Przestrzeń $M_{m \times n}(\mathbb{C})$ macierzy rozmiaru $m \times n$ o wyrazach zespolonych wraz z dodawaniem macierzy i mnożeniem macierzy przez skalar jest zespoloną przestrzenią liniową. Elementem 0 jest macierz $\begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}$, a elementem przeciwnym do $\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$ jest $\begin{pmatrix} -a_{11} & \dots & -a_{1n} \\ \vdots & \ddots & \vdots \\ -a_{m1} & \dots & -a_{mn} \end{pmatrix}$.

Przykład 3

Przestrzeń $\mathbb{C}_n[x]$ wielomianów stopnia co najwyżej n o współczynnikach zespolonych z działaniami dodawania wielomianów i mnożenia wielomianów przez skalar oraz przestrzeń $\mathbb{C}[x]$ wszystkich wielomianów o współczynnikach zespolonych. Elementem 0 jest wielomian zerowy (tzn. wielomian o wszystkich współczynnikach równych zero), a elementem przeciwnym do wielomianu $a_k x^k + \dots + a_1 x + a_0$ jest wielomian $-a_k x^k - \dots - a_1 x - a_0$.

Przykład 4

Przestrzeń funkcji zespolonych, czyli wszystkich funkcji $f : \mathbb{C} \rightarrow \mathbb{C}$ z działaniem dodawania funkcji i mnożenia funkcji przez skalar:

$$(f + g)(x) = f(x) + g(x) \quad (\alpha \cdot f)(x) = \alpha \cdot f(x)$$

Elementem 0 jest funkcja zerowa, tj. funkcja f taka, że $f(x) = 0$ dla każdego $x \in \mathbb{C}$. Elementem przeciwnym do funkcji f jest funkcja $-f$, gdzie $(-f)(x) = -f(x)$.

Przykład 5

Przestrzeń wszystkich ciągów o wyrazach zespolonych z dodawaniem ciągów i mnożeniem ciągów przez skalar.

Podobnie w miejsce pojęcia *przekształcenie liniowe* wprowadzamy pojęcie *zespolone przekształcenie liniowe*.

Definicja 3.17

Niech V i W będą zespolonymi przekształceniami liniowymi. Odwzorowanie $F : V \rightarrow W$ nazywamy *zespolonym przekształceniem liniowym*, jeśli dla dowolnych $u, v \in V$ oraz $t \in \mathbb{C}$ spełnione są warunki:

1. $F(u + v) = F(u) + F(v)$ (addytywność)
2. $F(t \cdot u) = t \cdot F(u)$ (jednorodność)

Przykład 6

Dane są dwa przekształcenia:

- $F : \mathbb{C}^2 \rightarrow \mathbb{C}$, gdzie $F\left(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}\right) = z_1 + z_2$,
- $G : \mathbb{C}^2 \rightarrow \mathbb{C}$, gdzie $G\left(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}\right) = z_1 + \bar{z}_2$.

Ustal czy są to: (a) rzeczywiste przekształcenia liniowe, (b) zespolone przekształcenia liniowe.

Rozwiązanie. Nietrudno sprawdzić addytywność obu przekształceń:

$$F\left(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\right) = F\left(\begin{pmatrix} z_1 + w_1 \\ z_2 + w_2 \end{pmatrix}\right) = z_1 + w_1 + z_2 + w_2 = F\left(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}\right) + F\left(\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\right)$$

$$G\left(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\right) = G\left(\begin{pmatrix} z_1 + w_1 \\ z_2 + w_2 \end{pmatrix}\right) = z_1 + w_1 + \bar{z}_2 + \bar{w}_2 = G\left(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}\right) + G\left(\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\right)$$

Przekształcenie F spełnia warunek jednorodności zarówno dla skalarów rzeczywistych, jak i zespolonych, bo dla dowolnego $\alpha \in \mathbb{C}$ zachodzi:

$$F\left(\alpha \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}\right) = F\left(\begin{pmatrix} \alpha z_1 \\ \alpha z_2 \end{pmatrix}\right) = \alpha z_1 + \alpha z_2 = \alpha(z_1 + z_2) = \alpha F\left(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}\right)$$

Natomiast przekształcenie G spełnia warunek jednorodności dla skalarów rzeczywistych, ale nie dla skalarów zespolonych, bo dla dowolnego $a \in \mathbb{R}$ zachodzi:

$$G\left(\alpha \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}\right) = G\left(\begin{pmatrix} \alpha z_1 \\ \alpha z_2 \end{pmatrix}\right) = \alpha z_1 + \bar{\alpha} \bar{z}_2 = \alpha z_1 + \alpha \bar{z}_2 = \alpha G\left(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}\right)$$

ale np. dla $\alpha = i$, $z_1 = 0$, $z_2 = 2 + i$ dostajemy:

$$G\left(i \begin{pmatrix} 0 \\ 2 + i \end{pmatrix}\right) = G\left(\begin{pmatrix} 0 \\ -1 + 2i \end{pmatrix}\right) = -1 - 2i \neq 1 + 2i = i G\left(\begin{pmatrix} 0 \\ 2 + i \end{pmatrix}\right)$$

Wobec tego oba przekształcenia są rzeczywistymi przekształceniami liniowymi, ale jedynie F jest zespolonym przekształceniem liniowym.

Wszystkie fakty, wnioski i twierdzenia z poprzednich rozdziałów pozostają prawdziwe (a ich dowody nie ulegają istotnym zmianom), jeśli w ich sformułowaniach oraz we wszystkich definicjach:

- pojęcie *przestrzeni liniowej* zamienimy na pojęcie *zespolonej przestrzeni liniowej*,
- pojęcie *przekształcenia liniowego* zamienimy na pojęcie *zespolonego przekształcenia liniowego*,
- zbiór $\mathbb{R}$ oraz przestrzeń $\mathbb{R}^n$ zamienimy, odpowiednio, na zbiór $\mathbb{C}$ oraz przestrzeń $\mathbb{C}^n$,
- przestrzeń $M_{m \times n}(\mathbb{R})$ (oznaczaną dotychczas $M_{m \times n}$) zamienimy na przestrzeń $M_{m \times n}(\mathbb{C})$.

Przykład 7

Zbiór $\mathbb{C}^3$ jest zespoloną przestrzenią liniową wymiaru 3. Jej baza to:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

bo każdy element $\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$, $z_1, z_2, z_3 \in \mathbb{C}$ można przedstawić jednoznacznie w postaci kombinacji liniowej powyższych elementów ze współczynnikami zespolonymi:

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = z_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + z_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z_3 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Ten sam zbiór możemy również traktować jako **rzeczywistą** przestrzeń liniową i wówczas ma ona wymiar 6, a jej bazą jest:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} i \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ i \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ i \end{pmatrix}$$

bo każdy element $\begin{pmatrix} a_1+ib_1 \\ a_2+ib_2 \\ a_3+ib_3 \end{pmatrix}$ można przedstawić jednoznacznie w postaci kombinacji liniowej powyższych elementów ze współczynnikami rzeczywistymi:

$$\begin{pmatrix} a_1+ib_1 \\ a_2+ib_2 \\ a_3+ib_3 \end{pmatrix} = a_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b_1 \cdot \begin{pmatrix} i \\ 0 \\ 0 \end{pmatrix} + a_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + b_2 \cdot \begin{pmatrix} 0 \\ i \\ 0 \end{pmatrix} + a_3 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + b_3 \cdot \begin{pmatrix} 0 \\ 0 \\ i \end{pmatrix}$$

Ogólnie: zbiór $\mathbb{C}^n$ można traktować jako zespoloną przestrzeń liniową wymiaru n lub jako rzeczywistą przestrzeń liniową wymiaru $2n$.

W szczególności prawdziwa jest zespolona wersja Twierdzenia 3.8 oraz Wniosku 3.14:

Twierdzenie 3.18: o diagonalizacji macierzy (wersja zespolona)

Macierz $A \in M_{n \times n}(\mathbb{C})$ można zapisać w postaci diagonalnej:

$$A = PDP^{-1}, \text{ gdzie } P = (v_1, \dots, v_n), \quad D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

wtedy i tylko wtedy, gdy wektory $v_1, \dots, v_n$ tworzą bazę $\mathbb{C}^n$ złożoną z wektorów własnych macierzy A , a $\lambda_1, \dots, \lambda_n$ to odpowiadające tym wektorom wartości własne. Taka baza istnieje wtedy i tylko wtedy, gdy dla każdej wartości własnej macierzy A zachodzi warunek:

$$\text{krotność wartości własnej } \lambda = \dim V^\lambda$$

Zwróćmy uwagę, że każdą macierz $A \in M_{n \times n}(\mathbb{R})$ (o wyrazach rzeczywistych) możemy traktować jako macierz $A \in M_{n \times n}(\mathbb{C})$ (o wyrazach zespolonych) co uzasadnia rachunki z ubiegłego semestru dotyczące diagonalizacji macierzy rzeczywistych z wykorzystaniem liczb zespolonych. Ponieważ będzie to sytuacja często spotykana w praktyce, nasza uwaga będzie się koncentrować przede wszystkim na diagonalizacji (w zespolonej przestrzeni liniowej) macierzy o wyrazach rzeczywistych.

Fakt 3.19

Jeżeli macierz $A \in M_{n \times n}(\mathbb{R})$ potraktujemy jako macierz o wyrazach zespolonych, to:

- 1) jeśli $\lambda \notin \mathbb{R}$ jest wartością własną A , to $\bar{\lambda}$ też jest wartością własną A oraz krotności i $\bar{\lambda}$ są jednakowe;
- 2) jeśli $v \in \mathbb{C}^n$ jest wektorem własnym dla λ , to $\bar{v} \in \mathbb{C}^n$ jest wektorem własnym dla $\bar{\lambda}$ (w szczególności $\dim V^\lambda = \dim V^{\bar{\lambda}}$).

Dowód. Zgodnie z faktem z ubiegłego semestru, wielomian charakterystyczny $\chi_A(x)$ jako wielomian o współczynnikach rzeczywistych, rozkłada się na iloczyn wielomianów pierwszego i drugiego stopnia:

$$\chi_A(\lambda) = \pm(\lambda - \alpha_1) \cdot \dots \cdot (\lambda - \alpha_k) \cdot (\lambda^2 + b_1\lambda + c_1) \cdot \dots \cdot (\lambda^2 + b_l\lambda + c_l)$$

gdzie $\alpha_i \in \mathbb{R}$, $i = 1, \dots, k$ oraz $b_i, c_i \in \mathbb{R}$, $i = 1, \dots, l$. Możemy założyć, że czynniki drugiego stopnia nie mają pierwiastków rzeczywistych. Wówczas każdy czynnik $\lambda^2 + b_i\lambda + c_i$ ma dwa nierzeczywiste sprzężone pierwiastki:

$$\lambda = -\frac{b_i}{2} + i\frac{\sqrt{|\Delta_i|}}{2} \quad \text{oraz} \quad \bar{\lambda} = -\frac{b_i}{2} - i\frac{\sqrt{|\Delta_i|}}{2}$$

gdzie $\Delta_i = b_i^2 - 4c_i < 0$. Stąd dla każdego pierwiastka λ istnieje pierwiastek $\bar{\lambda}$ tej samej krotności.

Niech teraz v będzie wektorem własnym dla wartości własnej λ , czyli:

$$Av = \lambda v$$

Sprzęgając równanie stronami i wykorzystując własności sprzężenia otrzymujemy:

$$\bar{A}\bar{v} = \bar{\lambda}\bar{v}$$

gdzie $\bar{A}$ powstaje z macierzy A przez zamianę każdego wyrazu na jego sprzężenie. Ponieważ A ma wszystkie wyrazy rzeczywiste, więc $\bar{A} = A$, czyli

$$A\bar{v} = \bar{\lambda}\bar{v}$$

co oznacza, że $\bar{v}$ jest wektorem własnym dla $\bar{\lambda}$ □

Przykład 8

Wyznaczyć potęgę macierzy:

$$\begin{pmatrix} 1 & -5 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 2 & 2 \end{pmatrix}^{20}$$

Rozwiązanie. Macierz ma wielomian charakterystyczny $\chi(\lambda) = (\lambda^2 - 4\lambda + 8)^2$, czyli wartości własne $\lambda_1 = \lambda_2 = 2 + 2i$ oraz $\lambda_3 = \lambda_4 = 2 - 2i$. Przestrzeń własna to:

$$V^{2+2i} = \text{Lin}\left\{\begin{pmatrix} 0 \\ 0 \\ i \\ 1 \end{pmatrix}, \begin{pmatrix} -1+2i \\ 1 \\ 0 \\ 0 \end{pmatrix}\right\} \quad \text{oraz} \quad V^{2-2i} = \text{Lin}\left\{\begin{pmatrix} 0 \\ 0 \\ -i \\ 1 \end{pmatrix}, \begin{pmatrix} -1-2i \\ 1 \\ 0 \\ 0 \end{pmatrix}\right\}$$

Zwróćmy uwagę, że zgodnie z Faktem 3.19 wystarczy wyliczyć przestrzeń własną dla wartości własnej $2 + 2i$. Wektory przestrzeni własnej V^{2-2i} to sprzężenia wektorów przestrzeni własnej V^{2+2i} . Stąd otrzymujemy diagonalizację badanej macierzy:

$$\begin{pmatrix} 0 & -1+2i & 0 & -1-2i \\ 0 & 1 & 0 & 1 \\ i & 0 & -i & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2+2i & & & \\ & 2+2i & & \\ & & 2-2i & \\ & & & 2-2i \end{pmatrix} \begin{pmatrix} 0 & -1+2i & 0 & -1-2i \\ 0 & 1 & 0 & 1 \\ i & 0 & -i & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}^{-1}$$

Ponieważ

$$(2+2i)^{20} = (2\sqrt{2}(\cos \frac{\pi}{4} + i \sin \frac{\pi}{4}))^{20} = (2\sqrt{2})^{20}(\cos \frac{20\pi}{4} + i \sin \frac{20\pi}{4}) = -2^{30}$$

$$(2-2i)^{20} = \overline{(2+2i)^{20}} = \overline{-2^{30}} = -2^{30}$$

więc szukana potęga macierzy to:

$$\begin{pmatrix} 0 & -1+2i & 0 & -1-2i \\ 0 & 1 & 0 & 1 \\ i & 0 & -i & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} -2^{30} & & & \\ & -2^{30} & & \\ & & -2^{30} & \\ & & & -2^{30} \end{pmatrix} \begin{pmatrix} 0 & -1+2i & 0 & -1-2i \\ 0 & 1 & 0 & 1 \\ i & 0 & -i & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} -2^{30} & 0 & 0 & 0 \\ 0 & -2^{30} & 0 & 0 \\ 0 & 0 & -2^{30} & 0 \\ 0 & 0 & 0 & -2^{30} \end{pmatrix}$$

Zwróćmy uwagę, że pomimo rachunków wykorzystujących liczby zespolone, ostateczny wynik jest macierzą o wyrazach rzeczywistych. Jest to spodziewana przez nas sytuacja, skoro jest to potęga macierzy, której wszystkie wyrazy są rzeczywiste.

Przykład 9

Dane jest przekształcenie liniowe $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ o macierzy

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

Wyznacz macierz przekształcenia $F^{10} = \underbrace{F \circ \dots \circ F}_{10}$.

Rozwiązanie. Wielomian charakterystyczny przekształcenia to $\chi_F(\lambda) = -(\lambda - 2)(\lambda - 2\lambda + 4)$, czyli wartości własne to $\lambda_1 = 2$, $\lambda_2 = 1 + i\sqrt{3}$, $\lambda_3 = 1 - i\sqrt{3}$. Wyznaczamy wektory własne: $v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1+i\sqrt{3} \\ 1 \\ 1 \end{pmatrix}$ oraz $v_3 = \overline{v_2} = \begin{pmatrix} 1-i\sqrt{3} \\ 1 \\ 1 \end{pmatrix}$. Macierz przekształcenia F (w bazie standardowej $\mathcal{E}$) ma zatem postać:

$$m_{\mathcal{E}}^{\mathcal{E}}(F) = \begin{pmatrix} 1 & 1+i\sqrt{3} & 1-i\sqrt{3} \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1+i\sqrt{3} & 0 \\ 0 & 0 & 1-i\sqrt{3} \end{pmatrix} \begin{pmatrix} 1 & 1+i\sqrt{3} & 1-i\sqrt{3} \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}^{-1}$$

zaś macierz przekształcenia F^{10} ma postać:

$$m_{\mathcal{E}}^{\mathcal{E}}(F^{10}) = \begin{pmatrix} 1 & 1+i\sqrt{3} & 1-i\sqrt{3} \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2^{10} & 0 & 0 \\ 0 & (1+i\sqrt{3})^{10} & 0 \\ 0 & 0 & (1-i\sqrt{3})^{10} \end{pmatrix} \begin{pmatrix} 1 & 1+i\sqrt{3} & 1-i\sqrt{3} \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}^{-1}$$

Uwzględniając, że:

$$(1+i\sqrt{3})^{10} = (2(\cos \frac{\pi}{3} + i \sin \frac{\pi}{3}))^{10} = 2^{10}(\cos \frac{10\pi}{3} + i \sin \frac{10\pi}{3}) = -2^9(1+i\sqrt{3})$$

$$(1-i\sqrt{3})^{10} = \overline{(1-i\sqrt{3})^{10}} = \overline{(1-i\sqrt{3})^{10}} = -2^9(1-i\sqrt{3})$$

otrzymujemy

$$m_{\mathcal{E}}^{\mathcal{E}}(F^{10}) = 2^9 \begin{pmatrix} 0 & 2 & -6 \\ 1 & 1 & -3 \\ 1 & -1 & -1 \end{pmatrix}$$

Diagonalizacja macierzy rzeczywistych przy pomocy liczb zespolonych ma tę wadę, że otrzymywane wektory własne są elementami przestrzeni $\mathbb{C}^n$, a nie $\mathbb{R}^n$. O ile więc liczby zespolone pozwalają wyznaczyć potęgę macierzy lub złożenie przekształcenia $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ (jak w powyższych przykładach), to nie pomagają w zrozumieniu działania pojedynczego przekształcenia F (otrzymane zespolone wektory własne nie są elementami dziedziny $\mathbb{R}^n$ przekształcenia F). Kolejne twierdzenie ma na celu zaradzenie temu problemowi.

Twierdzenie 3.20: klatkowa diagonalizacja nad $\mathbb{R}$

Jeśli macierz $A \in M_{n \times n}(\mathbb{R})$, traktowana jako macierz o wyrazach zespolonych, jest diagonalizowalna, to można ją przedstawić w postaci:

$$A = PDP^{-1} \text{ dla } D = \begin{pmatrix} D_1 & & \\ & \ddots & \\ & & D_k \end{pmatrix} \quad (3.4)$$

gdzie klatki D_i są dwojakiego rodzaju:

- każdej rzeczywistej wartości własnej λ macierzy A odpowiada klatka $D_i = (\lambda)$ (rozmiaru 1×1), a odpowiadająca jej kolumna macierzy P to wektor własny dla λ ,
- każdej parze sprzężonych wartości własnych $\lambda = a + bi$, $\bar{\lambda} = a - bi$ odpowiada klatka $D_i = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ (rozmiaru 2×2), a odpowiadające jej dwie kolumny macierzy P to wektory $u, w \in \mathbb{R}^n$, gdzie $u + wi$ oraz $u - wi$ to wektory własne, odpowiednio, dla λ i $\bar{\lambda}$.

Dowód. Pokażemy dowód dla sytuacji $n = 2$. Dowód w ogólnej sytuacji jest analogiczny. Macierz A ma wówczas 2 zespolone wartości własne: λ_1 i λ_2 . Jeśli $\lambda_1, \lambda_2 \in \mathbb{R}$, to przedstawienie (3.4) jest tożsame z diagonalizacją macierzy. W przeciwnym przypadku, zgodnie z Faktem 3.19, wartości własne to $\lambda_1 = a + bi$ oraz $\lambda_2 = \bar{\lambda}_1 = a - bi$, zaś wektory własne to $v_1 = u + iw$ oraz $v_2 = \bar{v}_1 = u - iw$, gdzie $u, w \in \mathbb{R}^2$. Zgodnie z definicją wektorów własnych:

$$Av_1 = \lambda_1 v_1 \quad \text{czyli} \quad A(u + iw) = (a + ib)(u + iw) = (au - bw) + i(bu + aw)$$

$$Av_2 = \lambda_2 v_2 \quad \text{czyli} \quad A(u - iw) = (a - ib)(u - iw) = (au - bw) - i(bu + aw)$$

Stąd otrzymujemy:

$$Au = A(\frac{1}{2}v_1 + \frac{1}{2}v_2) = \frac{1}{2}Av_1 + \frac{1}{2}Av_2 = au - bw$$

$$Aw = A(\frac{1}{2i}v_1 - \frac{1}{2i}v_2) = \frac{1}{2i}Av_1 - \frac{1}{2i}Av_2 = bu + aw$$

Wobec, traktując macierz A jako macierz przekształcenia $F_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, w bazie $\mathcal{B} = (u, w)$ dostajemy macierz przekształcenie:

$$m_{\mathcal{B}}^{\mathcal{B}}(F_A) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

czyli w bazie standardowej otrzymujemy:

$$A = m_{\mathcal{E}}^{\mathcal{E}}(F_A) = m_{\mathcal{E}}^{\mathcal{B}}(\text{id}) \cdot m_{\mathcal{B}}^{\mathcal{B}}(F_A) \cdot m_{\mathcal{B}}^{\mathcal{E}}(\text{id}) = P \cdot \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \cdot P^{-1}$$

gdzie

$$P = m_{\mathcal{E}}^{\mathcal{B}}(\text{id}) = ([b_1]_{\mathcal{E}}, [b_2]_{\mathcal{E}}) = \begin{pmatrix} u_1 & w_1 \\ u_2 & w_2 \end{pmatrix}$$

□

Przykład 10

Napisać klatkową rzeczywistą postać diagonalną macierzy obrotu o kąt θ wokół punktu O :

$$R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

Rozwiązanie. Zauważmy, że macierz R jest już w postaci $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, dla $a = \cos \theta$, $b = -\sin \theta$, więc przedstawienie (3.4) powinno przyjąć postać:

$$R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1}$$

Sprawdźmy, że algorytm z Twierdzenia 3.20 faktycznie prowadzi do takiego przedstawienia. Wielomian charakterystyczny macierzy R to $\chi_R(\lambda) = \lambda^2 - 2\lambda \cos \theta + 1$, a jego pierwiastkami są:

$$\lambda = \frac{1}{2}(2 \cos \theta \pm \sqrt{4 \cos^2 \theta - 4}) = \frac{1}{2}(2 \cos \theta \pm \sqrt{-4 \sin^2 \theta}) = \frac{1}{2}(2 \cos \theta \pm 2i \sin \theta) = \cos \theta \pm i \sin \theta$$

Wobec tego wartości własne macierzy R to:

$$\lambda_1 = \cos \theta - i \sin \theta \quad \lambda_2 = \cos \theta + i \sin \theta$$

Nietrudno wyliczyć, że odpowiadające im wektory własne to (z dokładnością do skalowania):

$$v_1 = \begin{pmatrix} 1 \\ i \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + i \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad v_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - i \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Stąd:

$$R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1}$$

Zauważmy, że jeśli zamienimy kolejność wartości (i równocześnie wektorów) własnych, otrzymamy postać:

$$R = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^{-1}$$

która również wypełnia polecenie z przykładu.

Przykład 11

Napisać klatkową rzeczywistą postać diagonalną macierzy:

$$A = \begin{pmatrix} 1 & -5 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 2 & 2 \end{pmatrix} \quad \text{oraz} \quad B = \begin{pmatrix} 0 & 2 & 2 \\ -1 & 3 & -1 \\ -1 & 1 & 1 \end{pmatrix}$$

Rozwiązanie. Wyznaczone w Przykładzie 8 zespolone wartości własne i wektory własne macierzy A tworzą dwie pary sprzężonych wartości i wektorów własnych:

$$\lambda = 2 \pm 2i \quad \text{wraz z} \quad v_1 = \begin{pmatrix} 0 \\ 0 \\ \pm i \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \pm i \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\lambda = 2 \pm 2i \quad \text{wraz z} \quad v_1 = \begin{pmatrix} -1 \pm 2i \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \pm i \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Stąd klatkowa rzeczywista postać diagonalna to:

$$A = \begin{pmatrix} 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ -2 & 2 \\ & 2 & 2 \\ & -2 & 2 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}^{-1}$$

Wyznaczone w Przykładzie 9 zespolone wartości własne i wektory własne macierzy B są następujące:

$$\lambda_1 = 2, \quad \lambda_2 = 1 + i\sqrt{3}, \quad \lambda_3 = \bar{\lambda}_2 = 1 - i\sqrt{3}$$

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 + i\sqrt{3} \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + i \begin{pmatrix} \sqrt{3} \\ 0 \\ 0 \end{pmatrix}, \quad v_3 = \bar{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - i \begin{pmatrix} \sqrt{3} \\ 0 \\ 0 \end{pmatrix}$$

Stąd dostajemy klatkową rzeczywistą postać diagonalną:

$$B = \begin{pmatrix} 1 & 1 & \sqrt{3} \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & & \\ & 1 & \sqrt{3} \\ & -\sqrt{3} & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & \sqrt{3} \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}^{-1}$$

Zamianę postaci zespolonej diagonalizacji na postać rzeczywistą można łatwiej zrozumieć, jeśli zauważymy, że bazuje on na utożsamieniu liczby zespolonej $a + ib \in \mathbb{C}$ z macierzą postaci $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, gdzie dodawanie i mnożenie liczb zespolonych odpowiada dodawaniu i mnożeniu macierzy:

$$(a + ib) + (c + id) = (a + c) + i(b + d) \quad \text{oraz} \quad \begin{pmatrix} a & b \\ -b & a \end{pmatrix} + \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} a + c & b + d \\ -(b + d) & a + c \end{pmatrix}$$

$$(a + ib) \cdot (c + id) = (ac - bd) + i(bc + ad) \quad \text{oraz} \quad \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \cdot \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} ac - bd & bc + ad \\ -(bc + ad) & ac - bd \end{pmatrix}$$

Klatkowa rzeczywista postać diagonalna nie będzie bardzo przydatna przy potęgowaniu macierzy (do tego celu znacznie lepiej nadaje się stosowana dotychczas diagonalizacja zespolona), ale pozwoli lepiej opisać i zrozumieć pojedyncze odwzorowanie $F : V \rightarrow V$. Do tego przydatne będą pojęcia podprzestrzeni niezmienniczych oraz sumy prostej podprzestrzeni.

Definicja 3.21

Dana jest (rzeczywista lub zespolona) przestrzeń liniowa V oraz (rzeczywiste lub zespolone) przekształcenie liniowe $F : V \rightarrow V$. Podprzestrzeń $W < V$ nazywamy *podprzestrzenią F -niezmienniczą*, jeśli $F(w) \in W$ dla każdego $w \in W$ (tzn. podprzestrzeń W jest zamknięta na działanie przekształcenia F).

Jeśli w powyższej sytuacji ograniczymy dziedzinę przekształcenia F do podprzestrzeni W , to otrzymamy przekształcenie $F|_W : W \rightarrow W$, które będziemy nazywać *obcięciem przekształcenia F do podprzestrzeni W* .

W sytuacji, gdy z kontekstu będzie jasno wynikać o jakie przekształcenie liniowe chodzi, zamiast *podprzestrzeń F -niezmiennicza* będziemy mówić po prostu *podprzestrzeń niezmiennicza*.

Przykład 12

Jeśli λ jest wartością własną przekształcenia liniowego $F : V \rightarrow V$, to przestrzeń własna V^λ jest podprzestrzenią F -niezmienniczą. Dla sprawdzenia tego faktu zauważmy, że jeśli $v \in V^\lambda$, to $F(v) = \lambda v$, czyli $F(v) \in V^\lambda$ (bo V^λ , jako podprzestrzeń, jest zamknięta na mnożenie przez skalar).

Przykład 13

Dla każdego przekształcenia liniowego $F : V \rightarrow V$ podprzestrzenie $\{0\}$ i V są podprzestrzeniami niezmienniczymi.

Przykład 14

Znajdź podprzestrzenie niezmiennicze dla przekształcenia liniowego $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, które jest obrotem wokół prostej ℓ przechodzącej przez 0.

Rozwiązanie. Podprzestrzenią niezmienniczą jest oczywiście prosta ℓ (jest to przestrzeń własna V^1) oraz płaszczyzna π prostopadła do ℓ i przechodząca przez punkt 0. Oczywiście podprzestrzeniami niezmienniczymi (jak dla każdego przekształcenia liniowego) są oczywiście podprzestrzenie zerowa oraz cała przestrzeń $\mathbb{R}^3$.

Przykład 15

Dane jest przekształcenie liniowe $F : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ zadane (w bazie standardowej) macierzą:

$$A = \begin{pmatrix} 2 & 3 & 0 & 0 \\ 1 & 5 & 0 & 0 \\ 0 & 0 & 7 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Wyznacz (nietrywialne) podprzestrzenie niezmiennicze $\mathbb{R}^4$, a następnie wyznacz macierze obcięć tych przekształceń do wyznaczonych podprzestrzeni (w odpowiednio dobranych bazach).

Rozwiązanie. Nietrudno zauważyć, że podprzestrzenią niezmienniczą jest każda z podprzestrzeni:

$$U = \left\{ \begin{pmatrix} x \\ y \\ 0 \\ 0 \end{pmatrix} : x, y \in \mathbb{R} \right\} \quad \text{oraz} \quad W = \left\{ \begin{pmatrix} 0 \\ 0 \\ z \\ t \end{pmatrix} : z, t \in \mathbb{R} \right\}$$

Dla wyznaczenia macierzy przekształceń $F|_U : U \rightarrow U$ oraz $F|_W : W \rightarrow W$ potrzebujemy wybrać bazy U i W . Jeśli wybierzemy odpowiednie elementy bazy standardowej $\mathbb{R}^4$, tzn. $\mathcal{B} = (e_1, e_2)$ dla U oraz $\mathcal{C} = (e_3, e_4)$ dla W , to:

$$m_{\mathcal{B}}^{\mathcal{B}}(F|_U) = \begin{pmatrix} 2 & 3 \\ 1 & 5 \end{pmatrix} \quad \text{oraz} \quad m_{\mathcal{C}}^{\mathcal{C}}(F|_W) = \begin{pmatrix} 7 & 3 \\ 1 & 1 \end{pmatrix}$$

Zauważmy, że są to dokładnie klatki macierzy A .

Powyższy przykład pokazuje, że jeśli przestrzeń V rozłożymy na podprzestrzenie niezmiennicze $V_1, \dots, V_n$, to analizę przekształcenia liniowego $F : V \rightarrow V$ możemy przeprowadzać osobno dla każdej podprzestrzeni V_i , zaś macierz przekształcenia F (w odpowiednio dobranej bazie) będzie macierzą klatkową złożoną z klatek – macierzy przekształceń $F|_{V_i}$. Poniższa definicja formalizuje co znaczy „rozłożyć” przestrzeń na podprzestrzenie.

Definicja 3.22

Przestrzeń liniowa V jest *sumą prostą* swoich podprzestrzeni $V_1, \dots, V_n < V$, co oznaczamy:

$$V = V_1 \oplus \dots \oplus V_n$$

jeśli każdy wektor $v \in V$ przedstawia się jednoznacznie w postaci

$$v = v_1 + \dots + v_n, \quad \text{gdzie} \quad v_i \in V_i, \quad \text{dla} \quad i = 1, \dots, n \quad (3.5)$$

Pojęcie sumy prostej (ozn. $\oplus$) należy odróżniać od pojęcia sumy kompleksowej (ozn. $+$) zdefiniowanej poniżej:

Definicja 3.23

Sumą kompleksową podzbiorów $A, B \subset V$ nazywamy zbiór:

$$A + B = \{a + b : a \in A, b \in B\}$$

Przykładem sumy kompleksowej jest poznana wcześniej warstwa podprzestrzeni. Warstwa $v + W$ to formalnie suma kompleksowa $\{v\} + W$. Suma kompleksowa $V_1 + V_2$ podprzestrzeni $V_1, V_2 \subset V$ to podprzestrzeń generowana przez V_1 i V_2 . Suma prosta to taka suma kompleksowa, w której składniki spełniają pewne dodatkowe założenia (innymi słowy: jeśli $V = V_1 \oplus V_2$, to $V = V_1 + V_2$, ale nie na odwrót).

Przykład 16

Pokazać, że jeśli $(b_1, \dots, b_n)$ jest bazą przestrzeni liniowej V , to V rozkłada się na sumę prostą jednowymiarowych podprzestrzeni:

$$V = \text{Lin}\{b_1\} \oplus \dots \oplus \text{Lin}\{b_n\}$$

Rozwiązanie. Zgodnie z definicją bazy każdy element $v \in V$ przedstawia się jednoznacznie w postaci kombinacji liniowej:

$$v = \alpha_1 b_1 + \dots + \alpha_n b_n$$

gdzie $\alpha_i b_i \in \text{Lin}\{b_i\}$, $i = 1, \dots, n$.

Przykład 17

Pokazać, że przestrzeń $\mathbb{R}^3$ można rozłożyć na sumę prostą w następujący sposób:

$$\mathbb{R}^3 = V_1 \oplus V_2, \quad \text{gdzie} \quad V_1 = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : x, y \in \mathbb{R} \right\}, \quad V_2 = \left\{ \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} : z \in \mathbb{R} \right\}$$

Rozwiązanie. Nietrudno zauważyć, że każdy wektor $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ przedstawia się jednoznacznie w postaci:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$$

Przykład 16 stanowi motywację do następującej charakteryzacji sumy prostej podprzestrzeni:

Fakt 3.24

Dane jest (skończenie wymiarowa) przestrzeń liniowa V oraz podprzestrzenie $V_1, \dots, V_n \subset V$, zaś $\mathcal{B}_1, \dots, \mathcal{B}_n$ niech będą bazami odpowiednio $V_1, \dots, V_n$. Wówczas $V = V_1 \oplus \dots \oplus V_n$ wtedy i tylko wtedy, gdy $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_n$ jest bazą V .

Dowód. Rozpatrzmy przypadek $n = 2$ (przypadek $n \geq 3$ jest analogiczny). Jeśli $\mathcal{B}_1 = (b_1, \dots, b_k)$ oraz $\mathcal{B}_2 = (b'_1, \dots, b'_l)$ to bazy odpowiednio V_1 i V_2 , to przedstawienie wektora $v \in V$ w postaci $v = v_1 + v_2$, gdzie $v_1 \in V_1$, $v_2 \in V_2$ oznacza przedstawienie:

$$v = (\alpha_1 b_1 + \dots + \alpha_k b_k) + (\alpha'_1 b'_1 + \dots + \alpha'_l b'_l)$$

które jest jednoznaczne wtedy i tylko wtedy, gdy $(b_1, \dots, b_k, b'_1, \dots, b'_l)$ jest bazą V . $\square$

Przykład 18

Pokazać, że $\mathbb{R}^3 = V_1 \oplus V_2$, jeśli $V_1 < \mathbb{R}^3$ jest 1-wymiarową podprzestrzenią (tzn. prostą przechodzącą przez 0), zaś $V_2 < \mathbb{R}^3$ jest 2-wymiarową podprzestrzenią (tzn. płaszczyzną przechodzącą przez 0) taką, że prosta V_1 nie leży na płaszczyźnie V_2 .

Rozwiązanie. Jeśli b_1 jest wektorem kierunkowym prostej V_1 , zaś wektory b_2 i b_3 rozpinają płaszczyznę V_2 , to układ (b_1, b_2, b_3) jest bazą $\mathbb{R}^3$.

Przykład 19

Pokazać, że przestrzeń $M_{2 \times 2}(\mathbb{R})$ można rozłożyć na sumę prostą:

$$M_{2 \times 2}(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ b & d \end{pmatrix} : a, b, d \in \mathbb{R} \right\} \oplus \left\{ \begin{pmatrix} 0 & c \\ -c & 0 \end{pmatrix} : c \in \mathbb{R} \right\}$$

Rozwiązanie. Bazą pierwszej przestrzeni jest $((\frac{1}{0} \ 0), (\frac{0}{1} \ 0), (\frac{0}{0} \ 1))$, zaś bazą drugiej przestrzeni jest $((\frac{0}{-1} \ 0))$. Nietrudno sprawdzić, że $((\frac{1}{0} \ 0), (\frac{0}{1} \ 0), (\frac{0}{0} \ 1), (\frac{0}{-1} \ 0))$ jest bazą przestrzeni $M_{2 \times 2}$.

Fakt 3.25

Niech V będzie skończenie wymiarową przestrzenią liniową, zaś $F : V \rightarrow V$ – przekształceniem liniowym. Jeżeli dla pewnych F -niezmienniczych podprzestrzeni $V_1, \dots, V_n$ zachodzi:

$$V = V_1 \oplus \dots \oplus V_n$$

to macierz F ma postać klatkową:

$$m_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} M_1 & & \\ & \ddots & \\ & & M_n \end{pmatrix}, \quad \text{gdzie} \quad M_i = m_{\mathcal{B}_i}^{\mathcal{B}_i}(F|_{V_i})$$

dla baz $\mathcal{B}_1, \dots, \mathcal{B}_n$ przestrzeni (odpowiednio) $V_1, \dots, V_n$ oraz bazy $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_n$ przestrzeni V .

Dowód. Przedstawimy dowód jedynie w szczególnym przypadku, gdy $V = V_1 \oplus V_2$ oraz $\dim V_1 = \dim V_2 = 2$. Dowód w przypadku ogólnym jest analogiczny.

Niech $\mathcal{B}_1 = (b_1, b_2)$ oraz $\mathcal{B}_2 = (b_3, b_4)$ będą bazami, odpowiednio, V_1 i V_2 . Wówczas $\mathcal{B} = (b_1, b_2, b_3, b_4)$ jest bazą V . Ponieważ V_1 jest przestrzenią F -niezmienniczą, więc $F(b_1), F(b_2) \in V_1$, czyli

$$F(b_1) = a_{11}b_1 + a_{21}b_2 \quad \text{oraz} \quad F(b_2) = a_{12}b_1 + a_{22}b_2$$

Podobnie, ponieważ V_2 jest przestrzenią F -niezmienniczą, więc $F(b_3), F(b_4) \in V_2$, czyli

$$F(b_3) = a'_{11}b_3 + a'_{21}b_4 \quad \text{oraz} \quad F(b_4) = a'_{12}b_3 + a'_{22}b_4$$

W tej sytuacji:

$$m_{\mathcal{B}_1}^{\mathcal{B}_1}(F|_{V_1}) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{oraz} \quad m_{\mathcal{B}_2}^{\mathcal{B}_2}(F|_{V_2}) = \begin{pmatrix} a'_{11} & a'_{12} \\ a'_{21} & a'_{22} \end{pmatrix}$$

natomiast

$$m_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ 0 & 0 & a'_{11} & a'_{12} \\ 0 & 0 & a'_{21} & a'_{22} \end{pmatrix}$$

□

Przykład 20

Napisać macierz przekształcenia $F : \mathbb{R}_3[x] \rightarrow \mathbb{R}_3[x]$ takiego, że $F(x^2 + 1) = x^2 + x + 1$, $F(x) = 3x$ oraz $F(x^3 + 1) = F(1) = x^3$:

- 1) w dowolnie wybranej bazie,
- 2) w bazie $\mathcal{C} = (1, x, x^2, x^3)$.

Rozwiązanie. Zauważmy, że podprzestrzenie $U = \text{Lin}\{x, x^2 + 1\}$ oraz $W = \text{Lin}\{x^3 + 1, 1\}$ są F -niezmiennicze. Wybierzmy bazę $\mathcal{B}_1 = (x, x^2 + 1)$ dla przestrzeni U oraz bazę $\mathcal{B}_2 = (x^3 + 1, 1)$ dla przestrzeni W . Zauważmy, że $\mathcal{B} = (x, x^2 + 1, x^3 + 1, 1)$ jest bazą przestrzeni $\mathbb{R}_3[x]$. Wobec tego:

$$\mathbb{R}_3[x] = U \oplus W$$

oraz

$$m_{\mathcal{B}_1}^{\mathcal{B}_1}(F|_U) = \begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{i} \quad m_{\mathcal{B}_2}^{\mathcal{B}_2}(F|_W) = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$$

Wobec tego:

$$m_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} 3 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & -1 & -1 \end{pmatrix}$$

Macierz przekształcenia F w bazie $\mathcal{C}$ otrzymujemy z następującego wzoru:

$$m_{\mathcal{C}}^{\mathcal{C}}(F) = m_{\mathcal{C}}^{\mathcal{B}}(\text{id}) m_{\mathcal{B}}^{\mathcal{B}}(F) m_{\mathcal{B}}^{\mathcal{C}}(\text{id}) = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & -1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}^{-1}$$

Fakt 3.26

Niech V będzie skończenie wymiarową przestrzenią liniową. Wówczas dla dowolnej podprzestrzeni $W < V$ istnieje podprzestrzeń $W' < V$ (zwana *podprzestrzenią dopełniczą do W*), dla której

$$V = W \oplus W'$$

Zazwyczaj jest wiele różnych podprzestrzeni dopełniczych do ustalonej podprzestrzeni.

Dowód. Wybierzmy bazę $\mathcal{B} = (b_1, \dots, b_k)$ przestrzeni W . Jest to zbiór liniowo niezależny w V , więc można go uzupełnić do bazy $\mathcal{B}' = (b_1, \dots, b_k, b_{k+1}, \dots, b_n)$ przestrzeni V . Wówczas przestrzeń $W' = \text{Lin}\{b_{k+1}, \dots, b_n\}$ jest, zgodnie z Faktem 3.24, podprzestrzenią dopełniczą do W . Ponieważ uzupełnienie bazy W do bazy V można wykonać na wiele różnych sposobów, więc podprzestrzeni dopełniczych również (zazwyczaj) jest wiele. $\square$

3.3 Twierdzenie Jordana

Jak wiemy, efektywne wyliczanie potęg macierzy, wymaga zapisania macierzy w postaci diagonalnej. Niestety, nie wszystkie macierze są diagonalizowalne, nawet przy użyciu liczb zespolonych (pamiętamy, że dzieje się tak zawsze w przypadku, gdy wymiar jakiegokolwiek przestrzeni własnej V^λ jest mniejszy od krotności wartości własnej λ). Niniejszy rozdział poświęcony będzie postępowaniu z takimi właśnie macierzami. Rozpoczniemy od pewnych szczególnych przykładów macierzy niediagonalizowalnych.

Przykład 1

Znajdź wartości i wektory własne macierzy:

$$\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \quad \begin{pmatrix} 5 & 1 \\ 0 & 5 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \quad \begin{pmatrix} 7 & 1 & 0 & 0 \\ 0 & 7 & 1 & 0 \\ 0 & 0 & 7 & 1 \\ 0 & 0 & 0 & 7 \end{pmatrix}$$

Rozwiązanie. Wszystkie rozważane macierze są macierzami górnotrójkątnymi, więc każda z nich ma tylko jedną wartość własną (odpowiednio: 3, 5, 2 i 7). Nietrudno sprawdzić, że każda z nich ma 1-wymiarową przestrzeń własną (generowaną odpowiednio przez: $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$).

Przykład 2

Niech λ będzie pewną liczbą rzeczywistą. Znajdź potęgi macierzy:

$$A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \quad B = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \quad C = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

Rozwiązanie. Każda z tych macierzy ma jedną wartość własną λ i 1-wymiarową przestrzeń własną, nie jest więc diagonalizowalna. Wyliczając kolejne potęgi:

$$\begin{aligned} A^2 &= \begin{pmatrix} \lambda^2 & 2\lambda \\ 0 & \lambda^2 \end{pmatrix} & B^2 &= \begin{pmatrix} \lambda^2 & 2\lambda & 1 \\ 0 & \lambda^2 & 2\lambda \\ 0 & 0 & \lambda^2 \end{pmatrix} & C^2 &= \begin{pmatrix} \lambda^2 & 2\lambda & 1 & 0 \\ 0 & \lambda^2 & 2\lambda & 1 \\ 0 & 0 & \lambda^2 & 2\lambda \\ 0 & 0 & 0 & \lambda^2 \end{pmatrix} \\ A^3 &= \begin{pmatrix} \lambda^3 & 3\lambda^2 \\ 0 & \lambda^3 \end{pmatrix} & B^3 &= \begin{pmatrix} \lambda^3 & 3\lambda^2 & 3\lambda \\ 0 & \lambda^3 & 3\lambda^2 \\ 0 & 0 & \lambda^3 \end{pmatrix} & C^3 &= \begin{pmatrix} \lambda^3 & 3\lambda^2 & 3\lambda & 1 \\ 0 & \lambda^3 & 3\lambda^2 & 3\lambda \\ 0 & 0 & \lambda^3 & 3\lambda^2 \\ 0 & 0 & 0 & \lambda^3 \end{pmatrix} \\ A^4 &= \begin{pmatrix} \lambda^4 & 4\lambda^3 \\ 0 & \lambda^4 \end{pmatrix} & B^4 &= \begin{pmatrix} \lambda^4 & 4\lambda^3 & 6\lambda^2 \\ 0 & \lambda^4 & 4\lambda^3 \\ 0 & 0 & \lambda^4 \end{pmatrix} & C^4 &= \begin{pmatrix} \lambda^4 & 4\lambda^3 & 6\lambda^2 & 4\lambda \\ 0 & \lambda^4 & 4\lambda^3 & 6\lambda^2 \\ 0 & 0 & \lambda^4 & 4\lambda^3 \\ 0 & 0 & 0 & \lambda^4 \end{pmatrix} \end{aligned}$$

nietrudno znaleźć ogólny wzór:

$$A^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix} \quad B^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} \\ 0 & \lambda^n & n\lambda^{n-1} \\ 0 & 0 & \lambda^n \end{pmatrix} \quad C^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \binom{n}{3}\lambda^{n-3} \\ 0 & \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} \\ 0 & 0 & \lambda^n & n\lambda^{n-1} \\ 0 & 0 & 0 & \lambda^n \end{pmatrix}$$

który następnie można udowodnić indukcyjnie. Przeprowadzenie dowodu indukcyjnego pozostawiamy czytelnikowi.

Można zauważyć, że wyrazy pojawiające się w potęgach powyższych macierzy, to kolejne

składniki rozwinięcia dwumianu Newtona:

$$(\lambda + 1)^n = \lambda^n + n\lambda^{n-1} + \binom{n}{2}\lambda^{n-2} + \binom{n}{3}\lambda^{n-3} + \cdots + n\lambda + 1$$

W związku z tym nietrudno uogólnić powyższe wzory na macierze rozmiaru 5×5 i większe. Macierze z powyższego przykładu nazywane są *klatkami Jordana*. Dzięki temu, że istnieje prosty wzór na n -tą potęgę macierzy będącej klatką Jordana, będzie ona dobrym substytutem macierzy diagonalnej, wykorzystywanym w następującym twierdzeniu:

Twierdzenie 3.27: Twierdzenie Jordana (wersja macierzowa)

Każdą macierz $A \in M_{n \times n}(\mathbb{C})$ możemy przedstawić w postaci (tzw. *postaci Jordana*):

$$A = P \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_n \end{pmatrix} P^{-1}, \text{ gdzie } J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}$$

Macierze J_i nazywamy *klatkami Jordana*.

Twierdzenie 3.28: Twierdzenie Jordana (wersja przekształceniowa)

Niech V będzie skończenie wymiarową zespoloną przestrzenią liniową, zaś $F : V \rightarrow V$ przekształceniem liniowym. Wówczas istnieje baza $\mathcal{B}$ przestrzeni V taka, że:

$$m_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_n \end{pmatrix}, \text{ gdzie } J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}$$

Macierze J_i nazywamy *klatkami Jordana*, a bazę $\mathcal{B}$ – bazą jordanowską. Wówczas dla dowolnej bazy $\mathcal{C}$ zachodzi:

$$m_{\mathcal{C}}^{\mathcal{C}}(F) = P \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_n \end{pmatrix} P^{-1}, \text{ gdzie } P = m_{\mathcal{C}}^{\mathcal{B}}(\text{id})$$

Dowód Twierdzenia Jordana jest skomplikowany i wykracza poza ramy niniejszego skryptu.

Przykład 3

Przykłady postaci Jordana to:

$$P \begin{pmatrix} 2 & 1 & & \\ 0 & 2 & & \\ & & 3 & 1 \\ & & 0 & 3 \end{pmatrix} P^{-1} \quad P \begin{pmatrix} 3 & 1 & & \\ 0 & 3 & & \\ & & 3 & 1 \\ & & 0 & 3 \end{pmatrix} P^{-1} \quad P \begin{pmatrix} 5 & 1 & & \\ 0 & 5 & & \\ & & 5 & \\ & & & 7 \end{pmatrix} P^{-1}$$

Zauważmy, że postać diagonalna macierzy PDP^{-1} , gdzie D jest macierzą diagonalną, jest bardzo szczególną postacią Jordana (wszystkie klatki Jordana mają rozmiar 1×1). Z tego powodu Twierdzenie Jordana można traktować jako uogólnienie twierdzenia o diagonalizacji macierzy.

Dalsza część tego rozdziału będzie poświęcona sposobom wyznaczania postaci Jordana macierzy, które nie są diagonalizowalne. Wyznaczanie przedstawienia Jordana PJP^{-1} rozpoczniemy od macierzy J :

Fakt 3.29

Jeśli $A = PJP^{-1}$ jest przedstawieniem Jordana macierzy A , to dla dowolnej wartości własnej λ macierzy A :

- 1) liczba wystąpień λ na przekątnej macierzy J jest równa krotności wartości własnej λ (innymi słowy: wyrazy na przekątnej macierzy J to pierwiastki wielomianu charakterystycznego χ_A liczone z krotnościami);
- 2) liczba klatek Jordana dla λ (wliczając w to ewentualne klatki rozmiaru 1×1) jest równa $\dim V^\lambda$.

Dowód. Zauważmy, że:

$$A - \lambda I = PJP^{-1} - P(\lambda I)P^{-1} = P(J - \lambda I)P^{-1}$$

czyli

$$\chi_A(\lambda) = \det(A - \lambda I) = \det(P(J - \lambda I)P^{-1}) = \det P \cdot \det(J - \lambda I) \cdot (\det P)^{-1} = \det(J - \lambda I) = \chi_J(\lambda)$$

Wobec tego wielomian charakterystyczny macierzy A jest identyczny z wielomianem charakterystycznym macierzy J . Z kolei macierz J jest macierzą górnątrójkątną, a wyrazy na przekątnej to dokładnie pierwiastki jego wielomianu charakterystycznego (z krotnościami), co dowodzi (1).

Dla dowodu (2) zauważmy, że każda klatka Jordana dla λ wyznacza jeden (z dokładnością do skalowania) wektor własny dla λ . Stąd liczba klatek dla λ to wymiar przestrzeni własnej dla λ . $\square$

Fakt 3.30

Dla danej macierzy A , macierz J w przedstawieniu Jordana $A = PJP^{-1}$ jest wyznaczona jednoznacznie z dokładnością do zamiany kolejności klatek Jordana.

Z Faktu 3.29 wynika, że wyrazy na przekątnej macierzy J oraz liczba klatek dla λ jest niezależna wyboru przedstawienia. To jest pierwszy krok w kierunku dowodu powyższego faktu, aczkolwiek niewystarczający, co pokazuje poniższy przykład (obie macierze mają jednakowe przekątne i jednakową liczbę klatek Jordana dla 2):

$$\begin{pmatrix} 2 & & & \\ & 2 & 1 & 0 \\ & 0 & 2 & 1 \\ & 0 & 0 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 & & \\ 0 & 2 & & \\ & & 2 & 1 \\ & & 0 & 2 \end{pmatrix}$$

Pełny dowód powyższego faktu wykracza poza ramy niniejszego skryptu.

Przykład 4

Wyznacz wszystkie możliwe macierze J z przedstawienia Jordana macierzy A , której wielomian charakterystyczny to:

$$\chi_A(\lambda) = -(\lambda - 2)^2(\lambda - 3)^3(\lambda - 5)$$

Rozwiązanie. Przekątna macierzy J zawiera pierwiastki wielomianu charakterystycznego A , czyli dwie 2, trzy 3 i jedną 5. Liczba klatek Jordana dla 2 i dla 3 zależy od $\dim V^2$ oraz $\dim V^3$. Dla 5 mamy jedną klatkę Jordana (gdyż $\dim V^5 = 1$ to jedyna możliwość). Zatem

macierz J jest (z dokładnością do zamiany kolejności klatek) jedną z następujących macierzy:

$$\begin{pmatrix} 2 & & & & \\ & 2 & & & \\ & & 3 & & \\ & & & 3 & \\ & & & & 5 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 & & & \\ 0 & 2 & & & \\ & & 3 & & \\ & & & 3 & \\ & & & & 5 \end{pmatrix} \quad \begin{pmatrix} 2 & & & & \\ & 2 & & & \\ & & 3 & 1 & \\ & & 0 & 3 & \\ & & & & 3 \\ & & & & & 5 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 & & & \\ 0 & 2 & & & \\ & & 3 & 1 & \\ & & 0 & 3 & \\ & & & & 3 \\ & & & & & 5 \end{pmatrix} \quad \begin{pmatrix} 2 & & & & \\ & 2 & & & \\ & & 3 & 1 & 0 \\ & & 0 & 3 & 1 \\ & & 0 & 0 & 3 \\ & & & & & 5 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 & & & \\ 0 & 2 & & & \\ & & 3 & 1 & 0 \\ & & 0 & 3 & 1 \\ & & 0 & 0 & 3 \\ & & & & & 5 \end{pmatrix}$$

Przykład 5

Wyznacz macierz J z przedstawienia Jordana macierzy A , której wielomian charakterystyczny to: $\chi_A(\lambda) = -(\lambda-2)^3(\lambda-5)^2$, a przestrzenie własne mają wymiary: $\dim V^2 = 2$ i $\dim V^3 = 1$.
Rozwiązanie. Przekątna macierzy J zawiera pierwiastki wielomianu charakterystycznego A , czyli trzy 2 i dwie 5. Liczba klatek Jordana dla 2 wynosi $2 = \dim V^2$, natomiast liczba klatek dla 5 wynosi $1 = \dim V^3$. Stąd J (z dokładnością do zamiany kolejności klatek) jest równa:

$$\begin{pmatrix} 2 & 1 & & & \\ 0 & 2 & & & \\ & & 2 & & \\ & & & 5 & 1 \\ & & & 0 & 5 \end{pmatrix}$$

Przykład 6

Wyznacz macierz J z przedstawienia Jordana każdej z następujących macierzy:

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 1 & 2 \\ -2 & 4 & 3 \\ 0 & 0 & 3 \end{pmatrix}$$

Rozwiązanie. Dla macierzy A wyznaczamy wielomian charakterystyczny $\chi_A(\lambda) = (\lambda-2)^2$ oraz wartości własne dla jedynej wartości własnej:

$$V^2 = \left\{ \begin{pmatrix} x \\ x \end{pmatrix}, x \in \mathbb{R} \right\}$$

Ponieważ $\dim V^2 = 1$, więc macierz J składa się z jednej klatki Jordana, skąd $J = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$.

Dla macierzy B wielomian charakterystyczny wynosi $\chi_B(\lambda) = -(\lambda-3)^2(\lambda-2)$. Wymiaru V^2 nie potrzeba wyliczać (wiadomo, że dla jednokrotnej wartości własnej mamy $\dim V^2 = 1$). Wyliczając wymiar V^3 otrzymujemy:

$$V^3 = \left\{ \begin{pmatrix} x \\ 2x \\ 0 \end{pmatrix}, x \in \mathbb{R} \right\}$$

czyli $\dim V^3 = 1$. Stąd macierz J ma jedną klatkę Jordana dla 3 i jedną klatkę dla 2, czyli

$$J = \begin{pmatrix} 3 & 1 & \\ 0 & 3 & \\ & & 2 \end{pmatrix}$$

Trudniejszą kwestią od wyznaczania macierzy J jest wyznaczanie macierzy P z postaci PJP^{-1} . W ogólnym przypadku algorytm wyznaczania P (czyli bazy jordanowskiej) jest trudny. W niniejszym skrypcie ograniczymy się do sytuacji, gdy dla każdej wartości λ mamy tylko jedną klatkę Jordana i ma ona rozmiar najwyżej 3×3 .

Przykład 7

Przekształcenie $F_A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ zadane macierzą $A = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 3 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ ma w pewnej bazie macierz

$m_{\mathcal{B}}^{\mathcal{B}}(F) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. Wyznacz bazę $\mathcal{B}$ o takiej własności (tzn. bazę jordanowską).

Rozwiązanie. Jeśli $\mathcal{B} = (b_1, b_2, b_3)$ to podaną macierz $m_{\mathcal{B}}^{\mathcal{B}}(F)$ można zinterpretować jako:

$$F_A(b_1) = 2b_1 + 0b_2 + 0b_3 = 2b_1$$

$$F_A(b_2) = 1b_1 + 2b_2 + 0b_3 = b_1 + 2b_2$$

$$F_A(b_3) = 0b_1 + 1b_2 + 2b_3 = b_2 + 2b_3$$

Wyznaczanie bazy $\mathcal{B}$ rozpoczniemy od wektora b_1 (wektora własnego). Nietrudno wyliczyć

$$V^2 = \left\{ \begin{pmatrix} x \\ 0 \\ x \end{pmatrix} : x \in \mathbb{R} \right\}$$

skąd możemy przyjąć np. $b_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. Uwzględniając ten wybór, z drugiego równania wyznaczamy $b_2 = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$:

$$\begin{pmatrix} 2 & 1 & 0 \\ -1 & 3 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + 2 \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$$

skąd otrzymujemy $b_2 = \begin{pmatrix} x_2 \\ 1 \\ x_2 - 1 \end{pmatrix}$. Możemy przyjąć np. $b_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$. Uwzględniając ten wybór, z trzeciego równania wyznaczamy $b_3 = \begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix}$:

$$\begin{pmatrix} 2 & 1 & 0 \\ -1 & 3 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix}$$

skąd otrzymujemy $b_3 = \begin{pmatrix} x_3 \\ 1 \\ x_3 \end{pmatrix}$. Możemy przyjąć np. $b_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. Stąd (przykładową) bazą spełniającą warunki zadania (bazą jordanowską) jest baza $\mathcal{B} = \left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right)$.

Przykład 8

Wyznacz postać Jordana i bazę jordanowską każdej z następujących macierzy:

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 1 & 2 \\ -2 & 4 & 3 \\ 0 & 0 & 3 \end{pmatrix}$$

Rozwiązanie. Zgodnie z wyliczeniami z Przykładu 6:

$$A = P \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} P^{-1}$$

Potraktujmy macierz A jako macierz przekształcenia $F_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Wówczas dla bazy jordanowskiej $\mathcal{B} = (b_1, b_2)$ mamy (podobnie jak w poprzednim przykładzie) $m_{\mathcal{B}}^{\mathcal{B}}(F_A) = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$.

Wobec tego:

$$Ab_1 = 2b_1$$

$$Ab_2 = b_1 + 2b_2$$

Wektor b_1 jest wektorem własnym dla $\lambda = 2$, więc możemy przyjąć (zgodnie z wyliczeniem z Przykładu 6) $b_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Wówczas z drugiego równania wyliczamy $b_2 = \begin{pmatrix} x \\ y \end{pmatrix}$:

$$\begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 2 \begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{skąd} \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ x+1 \end{pmatrix}$$

Możemy więc przyjąć $b_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, co daje nam przedstawienie Jordana:

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{-1}$$

Zgodnie z wyliczeniami z Przykładu 6:

$$B = P \begin{pmatrix} 3 & 1 \\ 0 & 3 \\ & & 2 \end{pmatrix} P^{-1}$$

Potraktujmy macierz B jako macierz przekształcenia $F_B : \mathbb{R}^3 \rightarrow \mathbb{R}^3$. Wówczas dla bazy jordanowskiej $\mathcal{C} = (c_1, c_2, c_3)$ mamy (podobnie jak w poprzednim przykładzie) $m_{\mathcal{C}}^{\mathcal{C}}(F_B) = \begin{pmatrix} 3 & 1 \\ 0 & 3 \\ & & 2 \end{pmatrix}$. Wobec tego:

$$Bc_1 = 3c_1$$

$$Bc_2 = c_1 + 3c_2$$

$$Bc_3 = 2c_3$$

Zatem c_1 i c_3 to wektory własne. Zgodnie z wyliczeniami z Przykładu 6 możemy przyjąć $c_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$. Nietrudno wyliczyć, że $c_3 = \begin{pmatrix} x \\ x \\ 0 \end{pmatrix}$, np. $c_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$. Wektor $c_2 = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ wyliczamy z drugiego równania:

$$\begin{pmatrix} 1 & 1 & 2 \\ -2 & 4 & 3 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \text{czyli} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ 2x-1 \\ 1 \end{pmatrix}$$

Stąd możemy przyjąć np. $c_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, co daje przedstawienie Jordana:

$$B = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 0 & 3 \\ & & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}^{-1}$$

Fakt 3.31

Jeśli $A = PJP^{-1}$, gdzie $A, J, P \in M_{n \times n}(\mathbb{C})$, to $A^k = PJ^kP^{-1}$ dla dowolnej liczby naturalnej k .

Dowód. Dla dowolnej macierzy J zachodzi wzór:

$$(PJP^{-1})^k = \underbrace{PJP^{-1} \cdot PJP^{-1} \cdot \dots \cdot PJP^{-1}}_k = P \underbrace{J \cdot J \cdot \dots \cdot J}_k P^{-1} = PJ^kP^{-1}$$

□

Przykład 9

Oblicz:

$$\begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}^n \quad \begin{pmatrix} 1 & 1 & 2 \\ -2 & 4 & 3 \\ 0 & 0 & 3 \end{pmatrix}^{10}$$

Rozwiązanie. Korzystając z przedstawień Jordana z poprzedniego przykładu dostajemy:

$$\begin{aligned} \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}^n &= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}^n \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2^n & n2^{n-1} \\ 0 & 2^n \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 2^n - n2^{n-1} & n2^{n-1} \\ -n2^{n-1} & 2^n + n2^{n-1} \end{pmatrix} \\ \begin{pmatrix} 1 & 1 & 2 \\ -2 & 4 & 3 \\ 0 & 0 & 3 \end{pmatrix}^{20} &= \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 0 & 3 \\ & 2 \end{pmatrix}^{10} \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3^{10} & 10 \cdot 3^9 & \\ 0 & 3^{10} & \\ & & 2^{10} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} -3^{10} + 2^{11} & 3^{10} - 2^{10} & 3^{10} + 10 \cdot 3^9 - 2^{10} \\ -2 \cdot 3^{10} + 2^{11} & 2 \cdot 3^{10} - 2^{10} & 3^{10} + 20 \cdot 3^9 - 2^{10} \\ 0 & 0 & 3^{10} \end{pmatrix} \end{aligned}$$

3.4 Zastosowania diagonalizacji**Przykład 1**

Wyznacz zwarty wzór ciągu a_n danego wzorem rekurencyjnym:

$$\begin{cases} a_0 = 2 \\ a_1 = 3 \\ a_{n+1} = 5a_n - 6a_{n-1}, \quad n \geq 1 \end{cases}$$

Rozwiązanie. Rekurencję tę możemy zapisać w postaci:

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} 5 & -6 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix}$$

Stosując wielokrotnie powyższy wzór otrzymujemy:

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} 5 & -6 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} a_1 \\ a_0 \end{pmatrix} = \begin{pmatrix} 5 & -6 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

Wyznaczamy diagonalizację macierzy:

$$\begin{pmatrix} 5 & -6 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}^{-1}$$

Skąd

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2^n & 0 \\ 0 & 3^n \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \cdot 2^{n+1} - 3^{n+1} \\ 3 \cdot 2^n - 3^n \end{pmatrix}$$

czyli $a_n = 3 \cdot 2^n - 3^n$. Zwróćmy uwagę, że powyższy algorytm dla dowolnych wartości a_0 i a_1 prowadzi do wzoru postaci $a_n = \alpha \cdot 2^n + \beta \cdot 3^n$.

Przykład 2

Wyznacz zwarty wzór ciągu b_n danego wzorem rekurencyjnym:

$$\begin{cases} b_0 = 3 \\ b_1 = 8 \\ b_{n+1} = 4b_n - 4b_{n-1}, \quad n \geq 1 \end{cases}$$

Rozwiązanie. Rekurencję tę możemy zapisać w postaci:

$$\begin{pmatrix} b_{n+1} \\ b_n \end{pmatrix} = \begin{pmatrix} 4 & -4 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} b_n \\ b_{n-1} \end{pmatrix}$$

Stosując wielokrotnie powyższy wzór otrzymujemy:

$$\begin{pmatrix} b_{n+1} \\ b_n \end{pmatrix} = \begin{pmatrix} 4 & -4 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} b_1 \\ b_0 \end{pmatrix} = \begin{pmatrix} 4 & -4 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} 8 \\ 3 \end{pmatrix}$$

Tym razem macierz nie jest diagonalizowalna, ale możemy wyznaczyć jej postać Jordana:

$$\begin{pmatrix} 4 & -4 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}^{-1}$$

Skąd

$$\begin{pmatrix} b_{n+1} \\ b_n \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2^n & n2^{n-1} \\ 0 & 2^n \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 8 \\ 3 \end{pmatrix} = \begin{pmatrix} (n+4) \cdot 2^{n+1} \\ (n+3) \cdot 2^n \end{pmatrix}$$

czyli $b_n = (n+3) \cdot 2^n$. Zwróćmy uwagę, że powyższy algorytm dla dowolnych wartości b_0 i b_1 prowadzi do wzoru postaci $b_n = \alpha \cdot 2^n + \beta \cdot n2^n$.

Przykład 3

Wyznacz zwarty wzór ciągu c_n danego wzorem rekurencyjnym:

$$\begin{cases} c_0 = 2 \\ c_1 = 5 \\ c_2 = 15 \\ c_{n+1} = 5c_n - 8c_{n-1} + 4c_{n-2}, \quad n \geq 2 \end{cases}$$

Rozwiązanie. Rekurencję tę możemy zapisać w postaci:

$$\begin{pmatrix} c_{n+1} \\ c_n \\ c_{n-1} \end{pmatrix} = \begin{pmatrix} 5 & -8 & 4 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} c_n \\ c_{n-1} \\ c_{n-2} \end{pmatrix}$$

Stosując wielokrotnie powyższy wzór otrzymujemy:

$$\begin{pmatrix} c_{n+1} \\ c_n \\ c_{n-1} \end{pmatrix} = \begin{pmatrix} 5 & -8 & 4 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}^{n-1} \begin{pmatrix} c_2 \\ c_1 \\ c_0 \end{pmatrix} = \begin{pmatrix} 5 & -8 & 4 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}^{n-1} \begin{pmatrix} 15 \\ 5 \\ 2 \end{pmatrix}$$

Powyższa macierz nie jest diagonalizowalna, ale możemy wyznaczyć jej postać Jordana:

$$\begin{pmatrix} 5 & -8 & 4 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 4 & 1 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & \\ 0 & 2 & \\ & & 1 \end{pmatrix} \begin{pmatrix} 4 & 4 & 1 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}^{-1}$$

Skąd

$$\begin{pmatrix} c_{n+1} \\ c_n \\ c_{n-1} \end{pmatrix} = \begin{pmatrix} 4 & 4 & 1 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2^n & n2^{n-1} & \\ 0 & 2^n & \\ & & 1^n \end{pmatrix} \begin{pmatrix} 4 & 4 & 1 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 15 \\ 5 \\ 2 \end{pmatrix} = \begin{pmatrix} (2n+1) \cdot 2^{n+1} + 3 \\ (2n-1) \cdot 2^n + 3 \\ (2n-3) \cdot 2^{n-1} + 3 \end{pmatrix}$$

czyli $c_n = (2n-1) \cdot 2^n + 3$. Zwróćmy uwagę, że powyższy algorytm dla dowolnych wartości c_0, c_1 i c_2 prowadzi do wzoru postaci $c_n = \alpha \cdot 2^n + \beta \cdot n2^n + \gamma \cdot 1^n$.

Przykład 4

Wyznacz zwarty wzór ciągu d_n danego wzorem rekurencyjnym:

$$\begin{cases} d_0 = 2 \\ d_1 = 0 \\ d_2 = 0 \\ d_{n+1} = 6d_n - 12d_{n-1} + 8d_{n-2}, \quad n \geq 2 \end{cases}$$

Rozwiązanie. Rekurencję tę możemy zapisać w postaci:

$$\begin{pmatrix} d_{n+1} \\ d_n \\ d_{n-1} \end{pmatrix} = \begin{pmatrix} 6 & -12 & 8 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} d_n \\ d_{n-1} \\ d_{n-2} \end{pmatrix}$$

Stosując wielokrotnie powyższy wzór otrzymujemy:

$$\begin{pmatrix} d_{n+1} \\ d_n \\ d_{n-1} \end{pmatrix} = \begin{pmatrix} 6 & -12 & 8 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}^{n-1} \begin{pmatrix} d_2 \\ d_1 \\ d_0 \end{pmatrix} = \begin{pmatrix} 6 & -12 & 8 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}^{n-1} \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$$

Powyższa macierz nie jest diagonalizowalna, ale możemy wyznaczyć jej postać Jordana:

$$\begin{pmatrix} 6 & -12 & 8 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 4 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 4 & 4 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}^{-1}$$

Skąd

$$\begin{pmatrix} d_{n+1} \\ d_n \\ d_{n-1} \end{pmatrix} = \begin{pmatrix} 4 & 4 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2^n & n2^{n-1} & \binom{n}{2}2^{n-2} \\ 0 & 2^n & n2^{n-1} \\ 0 & 0 & 2^n \end{pmatrix} \begin{pmatrix} 4 & 4 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} (n^2 - n) \cdot 2^{n+1} \\ (n^2 - 3n + 2) \cdot 2^n \\ (n^2 - 5n + 6) \cdot 2^{n-1} \end{pmatrix}$$

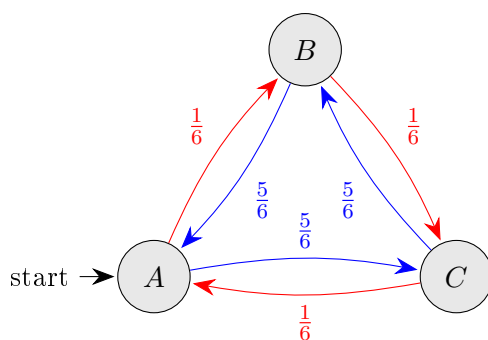
czyli $d_n = (n^2 - 3n + 2) \cdot 2^n$. Zwróćmy uwagę, że powyższy algorytm dla dowolnych wartości d_0 , d_1 i d_2 prowadzi do wzoru postaci $d_n = \alpha \cdot 2^n + \beta \cdot n2^n + \gamma \cdot n^2 2^n$.

Kolejne dwa przykłady pokazują zadania z rachunku prawdopodobieństwa, które prowadzą najpierw do układu rekurencji liniowych, a następnie do potęgowania macierzy. Schematy tego typu jak w tych przykładach nazywamy *tańcuchami Markowa* albo *procesami Markowa z czasem dyskretnym*.

Przykład 5

Na planszy są trzy pola: A , B i C , będące wierzchołkami trójkąta równobocznego. Początkowo pionek stoi na polu A . W każdym ruchu rzucamy kostką i przesuwamy pionek o jedno pole w kierunku zgodnym z ruchem wskazówek zegara, gdy wypadnie 6, a w kierunku przeciwnym do ruchu wskazówek zegara, gdy wypadnie inna liczba. Jakie jest prawdopodobieństwo, że po dokładnie 10 ruchach pionek znajdzie się znów na polu A ?

Rozwiązanie. Powyższą sytuację można zobrazować na następującym diagramie:



gdzie każdy ruch oznaczony jest strzałką opisaną prawdopodobieństwem jego wykonania. Niech a_n , b_n , c_n będą prawdopodobieństwami znalezienia się na, odpowiednio, na polach A , B , C po n rzutach. Wówczas:

$$\begin{cases} a_{n+1} = \frac{5}{6}b_n + \frac{1}{6}c_n \\ b_{n+1} = \frac{1}{6}a_n + \frac{5}{6}c_n \\ c_{n+1} = \frac{5}{6}a_n + \frac{1}{6}b_n \end{cases}$$

co można zapisać w postaci wektorowej:

$$\begin{pmatrix} a_{n+1} \\ b_{n+1} \\ c_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & \frac{5}{6} & \frac{1}{6} \\ \frac{1}{6} & 0 & \frac{5}{6} \\ \frac{5}{6} & \frac{1}{6} & 0 \end{pmatrix} \begin{pmatrix} a_n \\ b_n \\ c_n \end{pmatrix}$$

Pytanie z zadania dotyczy prawdopodobieństwa a_{10} . Można je wyliczyć przy pomocy potęgowania macierzy:

$$\begin{pmatrix} a_{10} \\ b_{10} \\ c_{10} \end{pmatrix} = \begin{pmatrix} 0 & \frac{5}{6} & \frac{1}{6} \\ \frac{1}{6} & 0 & \frac{5}{6} \\ \frac{5}{6} & \frac{1}{6} & 0 \end{pmatrix}^{10} \begin{pmatrix} a_0 \\ b_0 \\ c_0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{5}{6} & \frac{1}{6} \\ \frac{1}{6} & 0 & \frac{5}{6} \\ \frac{5}{6} & \frac{1}{6} & 0 \end{pmatrix}^{10} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \approx \begin{pmatrix} 0.30 \\ 0.38 \\ 0.32 \end{pmatrix}$$

Stąd szukane prawdopodobieństwo wynosi $a_{10} \approx 30\%$.

Przykład 6

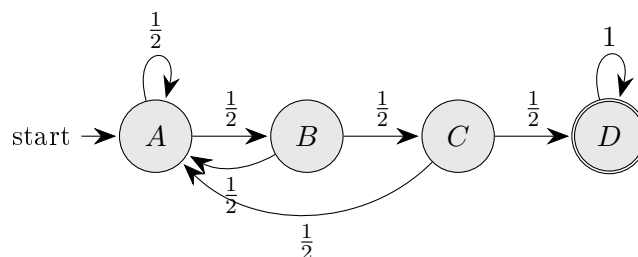
Jakie jest prawdopodobieństwo, że w 20 rzutach monetą wyrzucimy (przynajmniej raz) 3 orły pod rząd?

Rozwiązanie. Po każdym rzucie monetą znajdujemy się w jednej z następujących sytuacji:

(A) w ostatnim rzucie wypadła reszka (lub nie rzuciliśmy jeszcze ani razu),

- (B) w ostatnim rzucie wypadł orzeł, a w przedostatnim – reszka,
- (C) w ostatnich dwóch rzutach wypadły orły, a przed nimi – reszka,
- (D) wyrzuciliśmy już (kiedykolwiek) trzy orły pod rząd.

Na początku znajdujemy się w sytuacji A , naszym celem jest dotarcie do sytuacji D , a każdy rzut przenosi nas do kolejnej sytuacji (w zależności od wyniku). Możemy to zobrazować na następującym rysunku:



gdzie każdy rzut monetą oznaczony jest strzałką opisaną odpowiednim prawdopodobieństwem. Oznaczmy przez a_n, b_n, c_n, d_n prawdopodobieństwo znalezienia się po n rzutach, odpowiednio, w stanie A, B, C, D . Wówczas:

$$\begin{cases} a_{n+1} = \frac{1}{2}a_n + \frac{1}{2}b_n + \frac{1}{2}c_n \\ b_{n+1} = \frac{1}{2}a_n \\ c_{n+1} = \frac{1}{2}b_n \\ d_{n+1} = \frac{1}{2}c_n + d_n \end{cases}$$

co można zapisać w postaci wektorowej:

$$\begin{pmatrix} a_{n+1} \\ b_{n+1} \\ c_{n+1} \\ d_{n+1} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} a_n \\ b_n \\ c_n \\ d_n \end{pmatrix}$$

Pytanie z zadania dotyczy prawdopodobieństwa d_{20} . Można je wyliczyć przy pomocy potęgi macierzy:

$$\begin{pmatrix} a_{20} \\ b_{20} \\ c_{20} \\ d_{20} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 1 \end{pmatrix}^{20} \begin{pmatrix} a_0 \\ b_0 \\ c_0 \\ d_0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 1 \end{pmatrix}^{20} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \approx \begin{pmatrix} 0.115 \\ 0.062 \\ 0.034 \\ 0.787 \end{pmatrix}$$

jako że przed wykonaniem jakiegokolwiek rzutu znajdujemy się (z prawdopodobieństwem 1) w sytuacji A . Zatem szukane prawdopodobieństwo to $d_{20} \approx 78.7\%$.

Macierze, które pojawiają się w powyższych dwóch przykładach to tzw. *(lewe) macierze stochastyczne*. Zauważmy, że takie macierze mają wyrazy nieujemne (wyrazy oznaczają prawdopodobieństwa przejścia), zaś suma wyrazów w każdej kolumnie wynosi 1 (jest to łączne prawdopodobieństwo „wyjścia” z ustalonego stanu).

Ostatni przykład zastosowań diagonalizacji macierzy (w tym twierdzenia Jordana) dotyczy rozwiązywania układów liniowych równań różniczkowych. Zaczniemy od przypomnienia następującego faktu z analizy matematycznej:

Fakt 3.32

Dla dowolnego $x \in \mathbb{R}$ zachodzi równość:

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n$$

Nazwiązując do tego faktu zdefiniujemy exponens macierzy:

Definicja 3.33

Exponensem macierzy kwadratowej $A \in M_{n \times n}$ (oznaczanym $\exp A$ lub e^A) nazywamy następującą macierz:

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n$$

Przykład 7

Obliczyć e^A oraz e^B dla następujących macierzy:

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

Rozwiązanie.

$$\begin{aligned} e^A &= \sum_{n=0}^{\infty} \frac{1}{n!} A^n = \sum_{n=0}^{\infty} \frac{1}{n!} \begin{pmatrix} 2^n & 0 \\ 0 & 3^n \end{pmatrix} = \begin{pmatrix} \sum_{n=0}^{\infty} \frac{2^n}{n!} & \sum_{n=0}^{\infty} \frac{0}{n!} \\ \sum_{n=0}^{\infty} \frac{0}{n!} & \sum_{n=0}^{\infty} \frac{3^n}{n!} \end{pmatrix} = \begin{pmatrix} e^2 & 0 \\ 0 & e^3 \end{pmatrix} \\ e^B &= \sum_{n=0}^{\infty} \frac{1}{n!} B^n = \sum_{n=0}^{\infty} \frac{1}{n!} \begin{pmatrix} 2^n & n2^{n-1} \\ 0 & 2^n \end{pmatrix} = \begin{pmatrix} \sum_{n=0}^{\infty} \frac{2^n}{n!} & \sum_{n=0}^{\infty} \frac{n2^{n-1}}{n!} \\ \sum_{n=0}^{\infty} \frac{0}{n!} & \sum_{n=0}^{\infty} \frac{2^n}{n!} \end{pmatrix} = \begin{pmatrix} \sum_{n=0}^{\infty} \frac{2^n}{n!} & \sum_{n=1}^{\infty} \frac{2^{n-1}}{(n-1)!} \\ \sum_{n=0}^{\infty} \frac{0}{n!} & \sum_{n=0}^{\infty} \frac{2^n}{n!} \end{pmatrix} \\ &= \begin{pmatrix} e^2 & e^2 \\ 0 & e^2 \end{pmatrix} \end{aligned}$$

Fakt 3.34

Jeśli $A \in M_{n \times n}$ oraz $A = PJP^{-1}$, gdzie $P, J \in M_{n \times n}$, to:

$$e^A = e^{PJP^{-1}} = P \cdot e^J \cdot P^{-1}$$

Dowód.

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n = \sum_{n=0}^{\infty} \frac{1}{n!} (PJP^{-1})^n = \sum_{n=0}^{\infty} \frac{1}{n!} P J^n P^{-1} = P \cdot \left(\sum_{n=0}^{\infty} \frac{1}{n!} J^n \right) \cdot P^{-1} = P \cdot e^J \cdot P^{-1}$$

□

Przykład 8

Wyznaczyć *exponens* następujących macierzy:

$$A = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}$$

Rozwiązanie. Ponieważ $A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}^{-1}$ oraz $B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{-1}$, więc zgodnie z Faktem 3.34 oraz Przykładem 7:

$$e^A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} e^{\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}} \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} e^2 & 0 \\ 0 & e^3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 2e^3 - e^2 & e^2 - e^3 \\ 2e^3 - 2e^2 & 2e^2 - e^3 \end{pmatrix}$$

$$e^B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} e^{\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} e^2 & e^2 \\ 0 & e^2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 2e^2 & -e^2 \\ e^2 & 0 \end{pmatrix}$$

Pojęcie *exponensa* macierzy gra istotną rolę przy rozwiązywaniu równań różniczkowych. Zrozumienie poniższego przykładu **nie jest wymagane** w ramach kursu algebry liniowej, należy go traktować jedynie jako motywację dla zajmowania się tak nietypową operacją na macierzach jak *exponens*.

Najprostsze równanie różniczkowe to równanie postaci:

$$f'(x) = 2f(x)$$

a jego rozwiązanie to (nietrudno zgadnąć) $f(x) = e^{2x} \cdot C$ (gdzie C jest dowolną liczbą rzeczywistą). W podobny sposób rozwiązujemy liniowe układy równań różniczkowych:

Przykład 9

Rozwiązać układ równań różniczkowych:

$$\begin{cases} f'(x) = 4f(x) - g(x) \\ g'(x) = 2f(x) + g(x) \end{cases}$$

(tzn. wyznaczyć funkcje f i g spełniające to równanie).

Rozwiązanie. Powyższy układ można zapisać w postaci wektorowej:

$$\begin{pmatrix} f(x) \\ g(x) \end{pmatrix}' = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} f(x) \\ g(x) \end{pmatrix}$$

gdzie pochodną wektora obliczamy licząc pochodną każdej współrzędnej z osobna. Wówczas rozwiązanie dane jest wzorem:

$$\begin{pmatrix} f(x) \\ g(x) \end{pmatrix} = e^{x \cdot \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}} \cdot C = e^{\begin{pmatrix} 4x & -x \\ 2x & x \end{pmatrix}} \cdot C$$

gdzie C jest dowolnym wektorem z $\mathbb{R}^2$ (tak, by powyższe mnożenie miało sens). Diagonalizując macierz $A = \begin{pmatrix} 4x & -x \\ 2x & x \end{pmatrix}$ otrzymujemy:

$$\begin{pmatrix} 4x & -x \\ 2x & x \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2x & 0 \\ 0 & 3x \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}^{-1}$$

skąd, zgodnie z Faktem 3.34, dostajemy:

$$e^{\begin{pmatrix} 4x & -x \\ 2x & x \end{pmatrix}} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \cdot e^{\begin{pmatrix} 2x & 0 \\ 0 & 3x \end{pmatrix}} \cdot \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} e^{2x} & 0 \\ 0 & e^{3x} \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} 2e^{3x} - e^{2x} & e^{2x} - e^{3x} \\ 2e^{3x} - 2e^{2x} & 2e^{2x} - e^{3x} \end{pmatrix}$$

czyli (przyjmując $C = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$):

$$\begin{pmatrix} f(x) \\ g(x) \end{pmatrix} = \begin{pmatrix} \alpha(2e^{3x} - e^{2x}) + \beta(e^{2x} - e^{3x}) \\ \alpha(2e^{3x} - 2e^{2x}) + \beta(2e^{2x} - e^{3x}) \end{pmatrix}$$

Nietrudno sprawdzić, że tak określone funkcje f i g spełniają badany układ równań różniczkowych.

Rozdział 4

Przestrzeń euklidesowa

4.1 Iloczyn skalarny

Definicja 4.1

Przestrzenią euklidesową nazywamy (rzeczywistą) przestrzeń liniową V z dodatkową operacją zwaną *iloczynem skalarnym*, przypisującą (uporządkowanej) parze wektorów (u, v) liczbę $\langle u, v \rangle \in \mathbb{R}$ spełniającą dla dowolnych $u, v, w \in V$ oraz dowolnego $t \in \mathbb{R}$ następujące warunki:

- 1) $\langle u, v \rangle = \langle v, u \rangle$ (symetryczność)
- 2) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$
 $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$
 $\langle tu, v \rangle = \langle u, tv \rangle = t\langle u, v \rangle$ (liniowość względem każdej współrzędnej)
- 3) $\langle v, v \rangle \geq 0$, przy czym $\langle v, v \rangle = 0 \iff v = 0$ (dodania określoność)

Liniowość względem każdej z dwóch współrzędnych nazywamy również *dwuliniowością*. Innymi słowy: iloczyn skalarny to funkcja dwuliniowa.

Przykład 1

Sprawdź, że przestrzeń $\mathbb{R}^n$ z dodatkową operacją:

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = x_1 y_1 + \cdots + x_n y_n$$

(nazywanym *standardowym iloczynem skalarnym na $\mathbb{R}^n$*) jest przestrzenią euklidesową.

Rozwiązanie. Sprawdzenie symetryczności i dwuliniowości pozostawiamy czytelnikowi. Dodania określoność wynika z faktu, że:

$$x_1^2 + \cdots + x_n^2 \geq 0$$

dla dowolnych $x_1, \dots, x_n \in \mathbb{R}$, przy czym równość zachodzi jedynie gdy $x_1 = \cdots = x_n = 0$.

Przykład 2

Sprawdź, że operacja na przestrzeni $\mathbb{R}^2$ zdefiniowana następująco:

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = x_1 y_1 + x_1 y_2 + x_2 y_1 + 2x_2 y_2$$

jest iloczynem skalarnym (i tym samym wyznacza na $\mathbb{R}^2$ strukturę przestrzeni euklidesowej).

Rozwiązanie. Sprawdzenie symetryczności i dwuliniowości pozostawiamy czytelnikowi. Dodatnia określoność wynika z faktu, że:

$$x_1^2 + 2x_1x_2 + 2x_2^2 = (x_1 + x_2)^2 + x_2^2 \geq 0$$

dla dowolnych $x_1, x_2 \in \mathbb{R}$, przy czym równość zachodzi jedynie gdy $x_1 + x_2 = x_2 = 0$, czyli $x_1 = x_2 = 0$.

Przykład 3

Zdefiniowana na przestrzeni $\mathbb{R}^n$ dwuargumentowa funkcja:

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{1 \leq i, j \leq n} a_{ij} x_i y_j$$

inaczej zapisywana jako:

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = (x_1 \quad \dots \quad x_n) \cdot \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = X^T A Y$$

spełnia warunki (1) i (2) Definicji 4.49 dla dowolnych $a_{ij} \in \mathbb{R}$ takich, że $a_{ij} = a_{ji}$ (tzn. dla dowolnej macierzy symetrycznej A). Funkcja ta jest zatem iloczynem skalarnym na $\mathbb{R}^n$ wtedy i tylko wtedy, gdy macierz A jest symetryczna oraz dodatnio określona (tzn. spełniony jest warunek (3) Definicji 4.49). Metodę szybkiego ustalania czy macierz jest dodatnio określona poznamy w kolejnych rozdziałach.

Przykład 4

Sprawdź, że funkcja:

$$\langle P, Q \rangle = \int_0^1 P(x)Q(x)dx$$

jest iloczynem skalarnym na przestrzeni $\mathbb{R}_n[x]$.

Rozwiązanie. Symetryczność jest oczywista. Addytywność względem pierwszej współrzędnej wynika z własności całki:

$$\langle P + P', Q \rangle = \int_0^1 (P(x) + P'(x))Q(x)dx = \int_0^1 P(x)Q(x)dx + \int_0^1 P'(x)Q(x)dx = \langle P, Q \rangle + \langle P', Q \rangle$$

Podobnie dowodzimy addytywności względem drugiej współrzędnej oraz jednorodności względem każdej współrzędnej. Dla sprawdzenia dodatniej określoności zauważmy, że:

$$\langle P, P \rangle = \int_0^1 P^2(x)dx \geq 0$$

gdyż całka z nieujemnej funkcji zawsze jest nieujemna. Ponadto powyższa całka jest równa zero wtedy i tylko wtedy, gdy całkujemy funkcję zerową, czyli $P(x) = 0$ dla każdego $x \in [0, 1]$, ale jedyny wielomian, który przyjmuje wartość 0 w każdym punkcie przedziału $[0, 1]$ to wielomian zerowy.

Przykład 5

Sprawdź, że funkcja:

$$\langle P, Q \rangle = \int_{-1}^1 P(x)Q(x)dx$$

jest iloczynem skalarnym na przestrzeni $\mathbb{R}_n[x]$.

Rozwiązanie. Rozumowanie jest analogiczne jak w poprzednim przykładzie.

Przykład 6

Sprawdź, że funkcja:

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx$$

jest iloczynem skalarnym na przestrzeni $C[a, b]$.

Rozwiązanie. Rozumowanie jest takie jak w poprzednim przykładzie. Jedyna różnica polega na tym, że w odróżnieniu od niezerowego wielomianu, niezerowa funkcja może przyjmować wartość 0 na całym przedziale. Stąd należy osobno uzasadnić spełnienie warunku:

$$\langle v, v \rangle = 0 \iff v = 0$$

Dla funkcji ciągłej f równość $\int_a^b f^2(x)dx = 0$ zachodzi wtedy i tylko wtedy, gdy $f(x) = 0$ dla każdego $x \in [a, b]$. W przestrzeni $C[a, b]$ oznacza to, że f jest funkcją zerową.

Przykład 7

Sprawdź, że funkcja:

$$\langle f, g \rangle = \int_0^1 f(x)g(x)dx$$

nie jest iloczynem skalarnym na przestrzeni $C(\mathbb{R})$.

Rozwiązanie. Jedyny warunek iloczynu skalarnego, który nie jest spełniony, to warunek

$$\langle v, v \rangle = 0 \iff v = 0$$

Nietrudno wskazać ciągłą funkcję $f : \mathbb{R} \rightarrow \mathbb{R}$, która na całym przedziale $[0, 1]$ przyjmuje wartość 0, a poza tym przedziałem wartości niezerowe. Wówczas $\int_0^1 f^2(x)dx = 0$, ale $f \neq 0$.

Iloczyn skalarny pozwala zdefiniować takie pojęcia jak kąt między wektorami, prostopadłość (ortogonalność) wektorów oraz długość wektora.

Definicja 4.2

Dana jest przestrzeń euklidesowa V . *Długością* wektora $v \in V$ nazywamy liczbę nieujemną:

$$|v| = \sqrt{\langle v, v \rangle}$$

Odległością punktów $u, v \in V$ nazywamy liczbę nieujemną:

$$d(v, u) = |u - v| = \sqrt{\langle u - v, u - v \rangle}$$

Zwróćmy uwagę, że długość wektora jest dobrze zdefiniowana dzięki własności dodatniej określoności iloczynu skalarnego. Zauważmy też, że powyższa definicja jest motywowana analogicznym faktem dotyczącym przestrzeni $\mathbb{R}^2$ i $\mathbb{R}^3$. Podobnie motywacją dla definicji miary kąta

między (niezerowymi) wektorami dowolnej przestrzeni euklidesowej jest znany nam fakt dotyczący przestrzeni $\mathbb{R}^2$ i $\mathbb{R}^3$:

$$\cos \angle(u, v) \cdot |u| \cdot |v| = \langle u, v \rangle$$

Powyższy wzór można przyjąć za definicję miary kąta między wektorami dowolnej przestrzeni euklidesowej, o ile sprawdzimy, że iloczyn skalarny spełnia następujący warunek (w przeciwnym razie powyższy wzór prowadziłby do sprzeczności, gdyż $\cos \angle(u, v)$ miałby wartość spoza przedziału $[-1, 1]$):

Fakt 4.3: Nierówność Schwarza

Dla dowolnych wektorów $u, v \in V$ przestrzeni euklidesowej V zachodzi:

$$|\langle u, v \rangle| \leq |u| \cdot |v|$$

przy czym równość zachodzi wtedy i tylko wtedy, gdy u i v są współliniowe.

Dowód. Iloczyn skalarny jest dodatnio określony, więc dla dowolnego $t \in \mathbb{R}$ zachodzi:

$$0 \leq \langle u - tv, u - tv \rangle = \langle u, u \rangle - 2t\langle u, v \rangle + t^2\langle v, v \rangle$$

Podstawiając $t = \frac{\langle u, v \rangle}{\langle v, v \rangle}$ otrzymujemy:

$$0 \leq \langle u, u \rangle - 2 \frac{\langle u, v \rangle^2}{\langle v, v \rangle} + \frac{\langle u, v \rangle^2}{\langle v, v \rangle}$$

czyli po przekształceniu

$$\langle u, v \rangle^2 \leq \langle u, u \rangle \cdot \langle v, v \rangle$$

a więc

$$|\langle u, v \rangle| \leq |u| \cdot |v|$$

Równość zachodzi wtedy i tylko wtedy, gdy w wyjściowej nierówności zachodzi równość, tzn. $0 = \langle u - tv, u - tv \rangle$. Z uwagi na dodatnią określoność iloczynu skalarnego zachodzi to wtedy i tylko wtedy, gdy $u - tv = 0$, tzn. wektory u i v są współliniowe. $\square$

Definicja 4.4

Dana jest przestrzeń euklidesowa V . Kątem między niezerowymi wektorami $u, v \in V$ nazywamy liczbę θ z przedziału $[0, \pi]$ daną wzorem:

$$\theta = \arccos \frac{\langle u, v \rangle}{|u| \cdot |v|}$$

Wektory u i v nazywamy *ortogonalnymi* (lub *prostopadłymi*) jeśli $\theta = \frac{\pi}{2}$, czyli $\langle u, v \rangle = 0$.

Nierówność Schwarza zapewnia, że jest to poprawna definicja (tzn. $\frac{\langle u, v \rangle}{|u| \cdot |v|}$ jest liczbą z przedziału $[-1, 1]$). Drugim wnioskiem z nierówności Schwarza jest nierówność trójkąta (której również oczekivalibyśmy od każdej sensownej definicji długości):

Fakt 4.5: Nierówność trójkąta

Dla dowolnych wektorów $u, v \in V$ przestrzeni euklidesowej V zachodzi:

$$|u + v| \leq |u| + |v|$$

przy czym równość zachodzi wtedy i tylko wtedy, gdy $u = tv$ lub $v = tu$ dla pewnego $t \geq 0$.

Dowód.

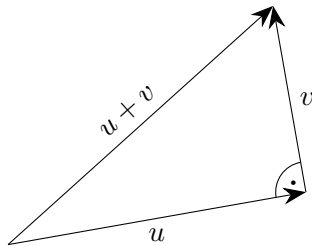
$$\begin{aligned} (|u+v|)^2 &= \langle u+v, u+v \rangle = \langle u, u \rangle + 2\langle u, v \rangle + \langle v, v \rangle = |u|^2 + 2\langle u, v \rangle + |v|^2 \\ &\leq |u|^2 + 2|u| \cdot |v| + |v|^2 = (|u| + |v|)^2 \end{aligned}$$

skąd

$$|u+v| \leq |u| + |v|$$

□

W dowolnej przestrzeni euklidesowej prawdziwe jest również Twierdzenie Pitagorasa. Jego sformułowanie jest motywowane następującą sytuacją z $\mathbb{R}^2$ i $\mathbb{R}^3$:



Fakt 4.6: Twierdzenie Pitagorasa

W przestrzeni euklidesowej V dla dowolnych wektorów $u, v \in V$ takich, że $u \perp v$, zachodzi:

$$|u+v|^2 = |u|^2 + |v|^2$$

Dowód. Korzystając z liniowości iloczynu skalarnego względem każdej współrzędnej oraz faktu $\langle u, v \rangle = 0$ otrzymujemy:

$$|u+v|^2 = \langle u+v, u+v \rangle = \langle u, u \rangle + 2\langle u, v \rangle + \langle v, v \rangle = \langle u, u \rangle + \langle v, v \rangle = |u|^2 + |v|^2$$

□

Przykład 8

Wyznacz długości boków AB i AC oraz miarę kąta przy wierzchołku A trójkąta w przestrzeni $\mathbb{R}^4$ (ze standardowym iloczynem skalarnym) o wierzchołkach $A = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, $C = \begin{pmatrix} 2 \\ 4 \\ 0 \\ 2 \end{pmatrix}$.

Rozwiązanie. Boki trójkąta wyznaczone są przez wektory:

$$\overrightarrow{AB} = B - A = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{oraz} \quad \overrightarrow{AC} = C - A = \begin{pmatrix} 1 \\ 3 \\ -1 \\ 1 \end{pmatrix}$$

Długości tych boków to:

$$|AB| = \sqrt{\left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle} = \sqrt{0^2 + 1^2 + 0^2 + 1^2} = \sqrt{2}$$

$$|AC| = \sqrt{\left\langle \begin{pmatrix} 1 \\ 3 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ -1 \\ 1 \end{pmatrix} \right\rangle} = \sqrt{1^2 + 3^2 + (-1)^2 + 1^2} = \sqrt{12}$$

Miara kąta przy wierzchołku A wynosi:

$$\arccos \frac{\langle \overrightarrow{AB}, \overrightarrow{AC} \rangle}{|\overrightarrow{AB}| \cdot |\overrightarrow{AC}|} = \arccos \frac{0 \cdot 1 + 1 \cdot 3 + 0 \cdot (-1) + 1 \cdot 1}{\sqrt{2} \cdot \sqrt{12}} = \arccos \frac{\sqrt{6}}{3}$$

Przykład 9

W przestrzeni euklidesowej $\mathbb{R}[x]$ z iloczynem skalarnym danym wzorem $\langle P, Q \rangle = \int_0^1 P(x)Q(x)dx$ dane są wektory $P(x) = x$ oraz $Q(x) = x^2 - 1$. Wyznacz długości tych wektorów oraz sprawdź, czy są one prostopadłe.

Rozwiązanie.

$$|P| = \sqrt{\langle P, P \rangle} = \sqrt{\int_0^1 x^2 dx} = \frac{\sqrt{3}}{3}$$

$$|Q| = \sqrt{\langle Q, Q \rangle} = \sqrt{\int_0^1 (x^2 - 1)^2 dx} = \frac{2\sqrt{30}}{15}$$

Wektory te **nie** są prostopadłe, gdyż:

$$\langle P, Q \rangle = \int_0^1 x(x^2 - 1)dx = -\frac{1}{4} \neq 0$$

Przykład 10

Rozważmy te same pytania co w poprzednim przykładzie, ale tym razem w odniesieniu do $\mathbb{R}[x]$ z iloczynem skalarnym danym wzorem $\langle P, Q \rangle = \int_{-1}^1 P(x)Q(x)dx$.

Rozwiązanie.

$$|P| = \sqrt{\langle P, P \rangle} = \sqrt{\int_{-1}^1 x^2 dx} = \frac{2\sqrt{3}}{3}$$

$$|Q| = \sqrt{\langle Q, Q \rangle} = \sqrt{\int_{-1}^1 (x^2 - 1)^2 dx} = \frac{2\sqrt{30}}{15}$$

Wektory te **są** prostopadłe, gdyż:

$$\langle P, Q \rangle = \int_{-1}^1 x(x^2 - 1)dx = 0$$

Przykład 11

W przestrzeni euklidesowej $C[0, 2\pi]$ z iloczynem skalarnym danym wzorem

$$\langle f, g \rangle = \int_0^{2\pi} f(x)g(x)dx$$

dane są wektory $f_n(x) = \cos nx$ oraz $g_n(x) = \sin nx$ (w tym $f_0(x) = 1$). Wyznacz długości tych wektorów oraz uzasadnij, że są one parami prostopadłe.

Rozwiązanie. Dla $n > 0$ mamy:

$$|f_n| = \sqrt{\langle \cos nx, \cos nx \rangle} = \sqrt{\int_0^{2\pi} \cos^2 nx dx} = \sqrt{\int_0^{2\pi} \frac{1}{2}(1 + \cos 2nx) dx} = \sqrt{\pi}$$

$$|g_n| = \sqrt{\langle \sin nx, \sin nx \rangle} = \sqrt{\int_0^{2\pi} \sin^2 nx dx} = \sqrt{\int_0^{2\pi} \frac{1}{2}(1 - \cos 2nx) dx} = \sqrt{\pi}$$

Natomiast dla $n = 0$:

$$|f_0| = \sqrt{\langle 1, 1 \rangle} = \sqrt{\int_0^{2\pi} 1 dx} = \sqrt{2\pi}$$

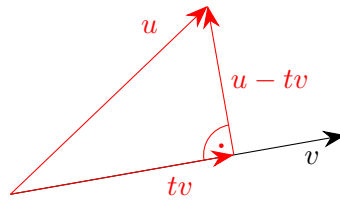
Prostopadłość (ortogonalność) wynika z następującego rachunku (gdzie $n \neq m$):

$$\langle f_n, f_m \rangle = \sqrt{\langle \cos nx, \cos mx \rangle} = \sqrt{\int_0^{2\pi} \frac{1}{2}(\cos(n-m)x + \cos(n+m)x)dx} = 0$$

$$\langle g_n, g_m \rangle = \sqrt{\langle \sin nx, \sin mx \rangle} = \sqrt{\int_0^{2\pi} \frac{1}{2}(\cos(n-m)x - \cos(n+m)x)dx} = 0$$

$$\langle f_n, g_m \rangle = \sqrt{\langle \cos nx, \sin mx \rangle} = \sqrt{\int_0^{2\pi} \frac{1}{2}(\sin(n+m)x - \sin(n-m)x)dx} = 0$$

Dysponując pojęciem ortogonalności (prostopadłości), możemy zdefiniować rzut ortogonalny (prostopadły) wektora na wektor. Definicja poniższa motywowana jest znaną nam sytuacją z przestrzeni $\mathbb{R}^2$ i $\mathbb{R}^3$, którą przypominamy na poniższym rysunku:



Definicja 4.7

Dana jest przestrzeń euklidesowa V . *Rzutem ortogonalnym (prostopadłym)* wektora $u \in V$ na wektor $v \in V$ (albo na podprzestrzeń $\text{Lin}\{v\} < V$) nazywamy taki wektor tv że $u - tv \perp v$. Rzut wektora u na wektor v oznaczamy $p_v(u)$.

Fakt 4.8

Jeśli $v \in V$ jest niezerowym wektorem przestrzeni euklidesowej V , to dla każdego wektora $u \in V$ istnieje dokładnie jeden wektor $p_v(u)$ i dany jest on wzorem:

$$p_v(u) = \frac{\langle u, v \rangle}{\langle v, v \rangle} v$$

Ponadto przekształcenie $p_v : V \rightarrow V$ jest przekształceniem liniowym

Dowód. Ponieważ $u - tv \perp v$, więc:

$$0 = \langle u - tv, v \rangle = \langle u, v \rangle - t\langle v, v \rangle$$

skąd

$$t = \frac{\langle u, v \rangle}{\langle v, v \rangle}$$

Z powyższego wzoru wynika addytywność p_v :

$$p_v(u + u') = \frac{\langle u + u', v \rangle}{\langle v, v \rangle} v = \frac{\langle u, v \rangle + \langle u', v \rangle}{\langle v, v \rangle} v = \frac{\langle u, v \rangle}{\langle v, v \rangle} v + \frac{\langle u', v \rangle}{\langle v, v \rangle} v$$

Podobnie sprawdzamy jednorodność p_v , co łącznie dowodzi liniowości przekształcenia p_v . $\square$

Fakt 4.9

Dana jest przestrzeń euklidesowa V . Rzut wektora $u \in V$ na niezerowy wektor $v \in V$ to (jedyny) taki punkt podprzestrzeni $\text{Lin}\{v\} < V$, który jest najbliższy punktowi u .

Dowód. Dla ułatwienia rachunków, zamiast znajdować minimum odległości, możemy znaleźć minimum kwadratu odległości:

$$d(u, tv)^2 = \langle u - tv, u - tv \rangle = \langle u, u \rangle - 2t\langle u, v \rangle + t^2\langle v, v \rangle$$

Powyższa funkcja jest funkcją kwadratową zmiennej $t \in \mathbb{R}$ i ma dodatni współczynnik przy t^2 , więc minimum przyjmuje w dokładnie jednym punkcie:

$$t = \frac{\langle u, v \rangle}{\langle v, v \rangle}$$

Otrzymany punkt (najbliższy punktowi u punkt podprzestrzeni $\text{Lin}\{v\}$) pokrywa się z $p_v(u)$. $\square$

Definicja 4.10

Niech V będzie przestrzenią euklidesową, zaś $A \subset V$. Wówczas *dopełnieniem ortogonalnym* zbioru A nazywamy:

$$A^\perp = \{v \in V : \forall a \in A \langle v, a \rangle = 0\}$$

Fakt 4.11

Dla dowolnego podzbioru $A \subset V$ przestrzeni euklidesowej zachodzi

$$A^\perp = (\text{Lin } A)^\perp$$

W szczególności dopełnienie ortogonalne podprzestrzeni $W < V$ pokrywa się z dopełnieniem ortogonalnym bazy W .

Dowód. Oczywiście jest, że $A^\perp \supset (\text{Lin } A)^\perp$. Dla udowodnienia zawierania w drugą stronę zauważmy, że każdy element $v \in \text{Lin } A$ jest postaci $v = \alpha_1 v_1 + \dots + \alpha_n v_n$, gdzie $v_1, \dots, v_n \in A$. Jeśli $w \in A^\perp$ to w szczególności $\langle w, v_i \rangle = 0$ dla każdego $v_i \in A$, czyli

$$\langle w, v \rangle = \langle w, \alpha_1 v_1 + \dots + \alpha_n v_n \rangle = \alpha_1 \langle w, v_1 \rangle + \dots + \alpha_n \langle w, v_n \rangle = 0$$

czyli $w \in (\text{Lin } A)^\perp$. $\square$

Fakt 4.12

Niech V będzie skończenie wymiarową przestrzenią euklidesową. Wówczas:

- 1) dla dowolnych $A \subset B \subset V$ zachodzi $A^\perp \supset B^\perp$;
- 2) dla dowolnego podzbioru $A \subset V$ zbiór $A^\perp$ jest podprzestrzenią V ;
- 3) dla dowolnej podprzestrzeni $W < V$ zachodzi $V = W \oplus W^\perp$.

Dowód. (1) Jeśli $v \in V$ jest ortogonalny do każdego wektora z B , to tym bardziej jest ortogonalny do każdego wektora z $A \subset B$, czyli $B^\perp \subset A^\perp$.

(2) Potrzebujemy sprawdzić zamkniętość na dodawanie i na mnożenie przez skalar. Jeśli $v, w \in A^\perp$, to dla każdego $a \in A$ zachodzi:

$$0 = \langle v, a \rangle + \langle w, a \rangle = \langle v + w, a \rangle$$

czyli $v + w \in A^\perp$. Podobnie sprawdzamy zamkniętość na mnożenie przez skalar.

(3) Niech $(w_1, \dots, w_k)$ będzie bazą przestrzeni W , zaś $(b_1, \dots, b_l)$ bazą przestrzeni $W^\perp$. Aby pokazać, że $V = W \oplus W^\perp$ musimy pokazać, że układ $(w_1, \dots, w_k, b_1, \dots, b_l)$ jest bazą V . Układ ten jest zbiorem liniowo niezależnym w V , gdyż dla dowolnej zerującej się kombinacji liniowej:

$$\alpha_1 w_1 + \dots + \alpha_k w_k + \beta_1 b_1 + \dots + \beta_l b_l = 0$$

otrzymujemy:

$$\alpha_1 w_1 + \dots + \alpha_k w_k = -\beta_1 b_1 - \dots - \beta_l b_l$$

Lewa strona tej nierówności należy do W , a prawa do $W^\perp$. Ponieważ $W \cap W^\perp = \{0\}$, więc obie strony są równe 0, skąd $\alpha_1 = \dots = \alpha_k = 0$ (bo $w_1, \dots, w_k$ jest lnz) oraz $\beta_1 = \dots = \beta_l = 0$ (bo $b_1, \dots, b_l$ jest lnz). Pozostaje wykazać, że $\dim V = k + l$. Przekształcenie $F : V \rightarrow \mathbb{R}^k$ zdefiniowane

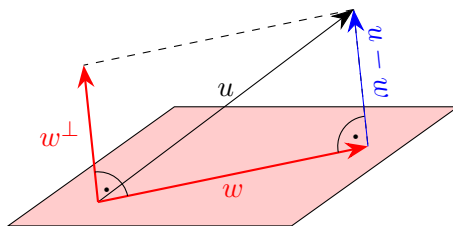
$$F(v) = \begin{pmatrix} \langle v, w_1 \rangle \\ \dots \\ \langle v, w_k \rangle \end{pmatrix}$$

jest liniowe oraz $\ker F = W^\perp$. Wobec tego z twierdzenia o indeksie otrzymujemy:

$$\dim V = \dim \ker F + \dim \operatorname{Im} F = \dim W^\perp + \dim \operatorname{Im} F \leq k + l$$

Ponieważ w V znaleźliśmy układ $k + l$ wektorów lnz, więc $\dim V \geq k + l$, co łącznie daje $\dim V = k + l$. $\square$

W Definicji 4.7 rozważaliśmy rzut ortogonalny na 1-wymiarową podprzestrzeń przestrzeni euklidesowej. Kolejna definicja opisuje rzut ortogonalny na podprzestrzeń dowolnego wymiaru. Motywacją dla tej definicji jest obrazek związany z rzutem ortogonalnym na płaszczyznę w $\mathbb{R}^3$:



Definicja 4.13

Niech V będzie skończenie wymiarową przestrzenią euklidesową, zaś $W < V$ jej podprzestrzenią. *Rzutem ortogonalnym (prostokątnym)* wektora $u \in V$ na podprzestrzeń W nazywamy taki wektor $w \in W$, że $u - w \in W^\perp$. Rzut wektora u na podprzestrzeń W oznaczamy $p_W(u)$.

Inny sposób określenia rzutu na podprzestrzeń W (również motywowany powyższym rysunkiem) jest następujący: ponieważ $V = W \oplus W^\perp$ (Fakt 4.12), więc wektor $u \in V$ rozkłada się jednoznacznie na sumę:

$$u = w + w^\perp, \quad \text{gdzie } w \in W, w^\perp \in W^\perp$$

Przy takim przedstawieniu $P_W(u) = w$.

Fakt 4.14

Niech $W < V$ jest podprzestrzenią skończenie wymiarowej przestrzeni euklidesowej. Wówczas dla każdego wektora u istnieje dokładnie jeden wektor $p_W(u)$. Co więcej, rzut ortogonalny na podprzestrzeń $p_W : V \rightarrow V$ jest odwzorowaniem liniowym.

Dowód. Definicja rzutu ortogonalnego na W sprowadza się do znajdowania rozkładu wektora u na sumę:

$$u = w + w^\perp, \quad \text{gdzie } w \in W, w^\perp \in W^\perp$$

Skoro $V = W \oplus W^\perp$, to dla danego wektora u , taki rozkład jest dokładnie jeden, co oznacza, że wektor $w = p_W(u)$ jest jednoznacznie wyznaczony.

Dla dowodu addytywności odwzorowania p_W założymy, że $p_W(u_1) = w_1$ oraz $p_W(u_2) = w_2$. Oznacza to, że:

$$u_1 = w_1 + w_1^\perp \quad \text{oraz} \quad u_2 = w_2 + w_2^\perp$$

Stąd

$$(u_1 + u_2) = (w_1 + w_2) + (w_1^\perp + w_2^\perp)$$

Ponieważ $w_1, w_2 \in W$, więc $w_1 + w_2 \in W$ (jako że W jest podprzestrzenią). Podobnie $w_1^\perp + w_2^\perp \in W^\perp$. Zatem powyższe równanie oznacza:

$$p_W(u_1 + u_2) = w_1 + w_2 = p_W(u_1) + p_W(u_2)$$

Podobnie sprawdzamy jednorodność p_W . □

Fakt 4.15

Jeśli $W < V$ jest podprzestrzenią przestrzeni euklidesowej V , zaś $v \in V$ dowolnym wektorem, to $P_W(v)$ jest punktem W najbliższym punktowi v .

Dowód. Zgodnie z definicją rzutu ortogonalnego $P_W(v) \in W$ oraz $v - P_W(v) \in W^\perp$. Ponieważ dla dowolnego $w \in W$ zachodzi $P_W(v) - w \in W$, więc $(v - P_W(v)) \perp (P_W(v) - w)$. Stąd, zgodnie z Twierdzeniem Pitagorasa, dostajemy:

$$|v - w|^2 = |(v - P_W(v)) + (P_W(v) - w)|^2 = |v - P_W(v)|^2 + |P_W(v) - w|^2 \geq |v - P_W(v)|^2$$

przy czym równość zachodzi wtedy i tylko wtedy, gdy $w = P_W(v)$. Oznacza to, że $P_W(v)$ to punkt podprzestrzeni W najbliższy punktowi v . □

Najwygodniejszym sposobem wyznaczania rzutu na podprzestrzeń jest sumowanie rzutów na wektory bazy podprzestrzeni. Jest to jednak możliwe wyłącznie wówczas, gdy wektory bazowe są parami prostopadłe:

Definicja 4.16

Bazę $\mathcal{B} = (b_1, \dots, b_n)$ przestrzeni euklidesowej nazywamy:

- *ortogonalną*, jeśli wektory $b_1, \dots, b_n$ są parami prostopadłe,
- *ortonormalną*, jeśli wektory $b_1, \dots, b_n$ są parami prostopadłe oraz każdy z nich ma długość 1.

Fakt 4.17

Niech V będzie przestrzenią euklidesową, $W < V$ jej podprzestrzenią, a $(b_1, \dots, b_k)$ bazą ortogonalną W . Wówczas:

$$P_W(v) = \sum_{i=1}^k p_{b_i}(v) = \sum_{i=1}^k \frac{\langle v, b_i \rangle}{\langle b_i, b_i \rangle} b_i \quad (4.1)$$

Jeśli, dodatkowo, baza $(b_1, \dots, b_k)$ jest ortonormalna, to:

$$P_W(v) = \sum_{i=1}^k \langle v, b_i \rangle b_i \quad (4.2)$$

Dowód. Oznaczmy

$$v' = \sum_{i=1}^k \frac{\langle v, b_i \rangle}{\langle b_i, b_i \rangle} b_i$$

Zgodnie z Definicją 4.13 wystarczy sprawdzić, że $v' \in W$ oraz $v - v' \in W^\perp$. Pierwszy z tych warunków jest natychmiastowy (v' to kombinacja liniowa bazy W). Dla sprawdzenia drugiego warunku musimy pokazać, że dla dowolnego $w \in W$ zachodzi $\langle v - v', w \rangle = 0$. Zaczniemy od pomocniczego rachunku wyliczającego $\langle v', b_j \rangle$. Z uwagi na liniowość iloczynu skalarnego względem pierwszej współrzędnej oraz ortogonalność bazy $(b_1, \dots, b_k)$ otrzymujemy:

$$\langle v', b_j \rangle = \left\langle \sum_{i=1}^k \frac{\langle v, b_i \rangle}{\langle b_i, b_i \rangle} b_i, b_j \right\rangle = \sum_{i=1}^k \frac{\langle v, b_i \rangle}{\langle b_i, b_i \rangle} \langle b_i, b_j \rangle = \frac{\langle v, b_j \rangle}{\langle b_j, b_j \rangle} \langle b_j, b_j \rangle = \langle v, b_j \rangle$$

Każdy element $w \in W$ jest postaci $w = \alpha_1 b_1 + \dots + \alpha_k b_k$, więc z uwagi na liniowość iloczynu skalarnego względem każdej współrzędnej, uwzględniając powyższy rachunek, otrzymujemy:

$$\langle v - v', w \rangle = \langle v - v', \alpha_1 b_1 + \dots + \alpha_k b_k \rangle = (\alpha_1 \langle v, b_1 \rangle + \dots + \alpha_k \langle v, b_k \rangle) - (\alpha_1 \langle v', b_1 \rangle + \dots + \alpha_k \langle v', b_k \rangle) = 0$$

co dowodzi (4.1). Wzór (4.2) jest szczególnym przypadkiem wzoru (4.1) w sytuacji, gdy $\langle b_i, b_i \rangle = 1$ dla $i = 1, \dots, k$. $\square$

Sprawdzenie, że podany układ parami ortogonalnych wektorów to baza jest wyjątkowo proste dzięki następującej obserwacji:

Fakt 4.18

Jeśli $v_1, \dots, v_n \in V$ to parami ortogonalne niezerowe wektory przestrzeni euklidesowej V , to są one liniowo niezależne.

Dowód. Rozważmy zerującą się kombinację liniową wektorów $v_1, \dots, v_n$:

$$\alpha_1 v_1 + \dots + \alpha_n v_n = 0$$

Korzystając z własności iloczynu skalarnego oraz wzajemnej ortogonalności rozważanych wektorów otrzymujemy:

$$0 = \langle 0, v_1 \rangle = \langle \alpha_1 v_1 + \dots + \alpha_n v_n, v_1 \rangle = \alpha_1 \langle v_1, v_1 \rangle + \alpha_2 \langle v_2, v_1 \rangle + \dots + \alpha_n \langle v_n, v_1 \rangle = \alpha_1 \langle v_1, v_1 \rangle = \alpha_1 |v_1|^2$$

Ponieważ v_1 jest niezerowy, więc $\alpha_1 = 0$. Podobnie dowodzimy, że $\alpha_2 = \dots = \alpha_n = 0$, co dowodzi liniowej niezależności $v_1, v_2, \dots, v_n$. $\square$

Przykład 12

W przestrzeni euklidesowej $C[0, 2\pi]$ z iloczynem skalarnym danym wzorem

$$\langle f, g \rangle = \frac{1}{\pi} \int_0^{2\pi} f(x)g(x)dx$$

dane są wektory $\frac{1}{\sqrt{2}}, \cos x, \sin x, \cos 2x, \sin 2x, \dots, \cos nx, \sin nx$. Uzasadnij, że wektory te stanowią bazę ortonormalną pewnej podprzestrzeni $W < C[0, 2\pi]$.

Rozwiązanie. Rozważany iloczyn skalarny różni się od iloczynu rozważanego w Przykładzie 11 jedynie współczynnikiem $\frac{1}{\pi}$. Wobec tego wszystkie iloczyny skalarne w bieżącym przykładzie będą π razy mniejsze od iloczynów skalarnych z Przykładu 11 (w szczególności wszystkie długości będą $\sqrt{\pi}$ razy mniejsze. Stąd rozważane wektory są parami ortogonalne i mają długości 1. Stanowią wobec tego bazę ortonormalną pewnej podprzestrzeni $W < C[0, 2\pi]$.

Przykład 13 (Szeregi Fouriera)

W przestrzeni euklidesowej $C[0, 2\pi]$ z iloczynem skalarnym danym wzorem

$$\langle f, g \rangle = \frac{1}{\pi} \int_0^{2\pi} f(x)g(x)dx$$

wyznacz rzut ortogonalny wektora $f \in C[0, 2\pi]$ na podprzestrzeń

$$W = \text{Lin}\left\{\frac{1}{\sqrt{2}}, \cos x, \sin x, \cos 2x, \sin 2x, \dots, \cos nx, \sin nx\right\}$$

Rozwiązanie. Zgodnie z poprzednim przykładem, wektory $\frac{1}{\sqrt{2}}, \cos x, \sin x, \cos 2x, \sin 2x, \dots, \cos nx, \sin nx$ stanowią bazę ortonormalną W . Wówczas dla dowolnego $f \in C[0, 2\pi]$, zgodnie ze wzorem (4.2):

$$p_W(f)(x) = \frac{a_0}{\sqrt{2}} + a_1 \cos x + b_1 \sin x + \dots + a_n \cos nx + b_n \sin nx$$

gdzie

$$a_0 = \langle f, \frac{1}{\sqrt{2}} \rangle = \frac{1}{\pi} \int_0^{2\pi} \frac{1}{\sqrt{2}} f(x) dx$$

$$a_k = \langle f, \cos kx \rangle = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx dx, \quad k = 1, 2, \dots, n$$

$$b_k = \langle f, \sin kx \rangle = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx dx, \quad k = 1, 2, \dots, n$$

Co jest dobrze znanym wzorem z analizy (przybliżeniem funkcji f zwanym *szeregiem Fouriera*).

Układ $\frac{1}{\sqrt{2}}, \cos x, \sin x, \cos 2x, \sin 2x, \dots$ nie stanowi bazy przestrzeni $C[0, 2\pi]$, gdyż wymiar tej przestrzeni jest nieprzeliczalny. Rozważania dotyczące baz przestrzeni nieskończenie wymiarowych wykraczają jednak poza zakres niniejszego skryptu.

Fakt 4.19: Ortogonalizacja Grama–Schmidta

Każda skończona wymiarowa przestrzeń euklidesowa V ma bazę ortonormalną. Bazę taką można wyznaczyć według następującej procedury (zwanej *ortogonalizacją Grama–Schmidta*):

1. wyznaczamy **dowolną** bazę $(b_1, \dots, b_n)$;
2. wyznaczamy bazę **ortogonalną** $(b'_1, \dots, b'_n)$ według następującego wzoru:

$$\begin{cases} b'_1 = b_1 \\ b'_2 = b_2 - p_{b'_1}(b_2) \\ b'_3 = b_3 - p_{b'_1}(b_3) - p_{b'_2}(b_3) \\ \dots \\ b'_n = b_n - p_{b'_1}(b_n) - p_{b'_2}(b_n) - \dots - p_{b'_{n-1}}(b_n) \end{cases}$$

czyli (rozwijając wzór na rzut wektora na wektor):

$$\begin{cases} b'_1 = b_1 \\ b'_2 = b_2 - \frac{\langle b_2, b'_1 \rangle}{\langle b'_1, b'_1 \rangle} b'_1 \\ b'_3 = b_3 - \frac{\langle b_3, b'_1 \rangle}{\langle b'_1, b'_1 \rangle} b'_1 - \frac{\langle b_3, b'_2 \rangle}{\langle b'_2, b'_2 \rangle} b'_2 \\ \dots \\ b'_n = b_n - \frac{\langle b_n, b'_1 \rangle}{\langle b'_1, b'_1 \rangle} b'_1 - \frac{\langle b_n, b'_2 \rangle}{\langle b'_2, b'_2 \rangle} b'_2 - \dots - \frac{\langle b_n, b'_{n-1} \rangle}{\langle b'_{n-1}, b'_{n-1} \rangle} b'_{n-1} \end{cases}$$

3. wyznaczamy bazę **ortonormalną** $(b''_1, \dots, b''_n)$ według następującego wzoru:

$$b''_1 = \frac{b'_1}{|b'_1|} \quad b''_2 = \frac{b'_2}{|b'_2|} \quad \dots \quad b''_n = \frac{b'_n}{|b'_n|}$$

Dowód. Oznaczmy

$$W_1 = \text{Lin}\{b_1\}, \quad W_2 = \text{Lin}\{b_1, b_2\}, \quad \dots, \quad W_n = \text{Lin}\{b_1, b_2, \dots, b_n\} = V$$

Ponieważ wektor b'_k jest kombinacją liniową wektorów $b'_1, \dots, b'_{k-1}, b_k$, więc wybór wektorów $b'_1, \dots, b'_n$ prowadzi do nowych baz przestrzeni $W_1, \dots, W_n$:

$$W_1 = \text{Lin}\{b'_1\}, \quad W_2 = \text{Lin}\{b'_1, b'_2\}, \quad \dots, \quad W_n = \text{Lin}\{b'_1, b'_2, \dots, b'_n\} = V$$

Pozostaje sprawdzenie, że wektory $b'_1, \dots, b'_n$ są parami ortogonalne. Przy pomocy indukcji względem k dowodzimy, że $b'_1, \dots, b'_k$ są parami ortogonalne:

1) Dla $k = 1$ jest to oczywiste.

2) Jeśli $b'_1, \dots, b'_k$ są parami ortogonalne, to zgodnie z Faktem 4.17 mamy $b'_{k+1} = b_{k+1} - p_{W_k}(b_{k+1})$, więc $b'_{k+1} \in W_k^\perp$, czyli b'_{k+1} jest ortogonalny do każdego z wektorów $b'_1, \dots, b'_k$. $\square$

Przykład 14

W przestrzeni euklidesowej $\mathbb{R}_4$ ze standardowym iloczynem skalarnym dana jest podprzestrzeń

$$W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4 : x_1 + 2x_2 + x_3 - x_4 = 0 \right\}. \text{ Wyznacz:}$$

- a) bazę ortogonalną przestrzeni W .
- b) bazę ortonormalną przestrzeni W .

Rozwiązanie. Zaczynamy od wyznaczenia jakiegokolwiek bazy, np.

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

Zgodnie z metodą Grama-Schmidta wyznaczamy:

$$\begin{cases} b'_1 = b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\ b'_2 = b_2 - \frac{\langle b_2, b'_1 \rangle}{\langle b'_1, b'_1 \rangle} b'_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix} - \frac{2}{2} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \\ b'_3 = b_3 - \frac{\langle b_3, b'_1 \rangle}{\langle b'_1, b'_1 \rangle} b'_1 - \frac{\langle b_3, b'_2 \rangle}{\langle b'_2, b'_2 \rangle} b'_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{6} \\ -\frac{1}{3} \\ 1 \\ \frac{1}{6} \end{pmatrix} \end{cases}$$

Otrzymując w ten sposób bazę ortogonalną. Dla wyznaczenia bazy ortonormalnej, normalizujemy otrzymane wektory:

$$b''_1 = \frac{b'_1}{|b'_1|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad b''_2 = \frac{b'_2}{|b'_2|} = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \quad b''_3 = \frac{b'_3}{|b'_3|} = \frac{1}{\sqrt{\frac{7}{6}}} \begin{pmatrix} -\frac{1}{6} \\ -\frac{1}{3} \\ 1 \\ \frac{1}{6} \end{pmatrix}$$

Przykład 15

Wyznacz (a) bazę ortogonalną, (b) bazę ortonormalną, przestrzeni euklidesowej $\mathbb{R}_2[x]$ z iloczynem skalarnym $\langle P, Q \rangle = \int_0^1 P(x)Q(x)dx$.

Rozwiązanie. Zaczynamy od wyznaczenia jakiegokolwiek bazy, np. $b_1 = 1$, $b_2 = x$, $b_3 = x^2$. Zgodnie z metodą Grama-Schmidta wyznaczamy:

$$\begin{cases} b'_1 = b_1 = 1 \\ b'_2 = b_2 - \frac{\langle b_2, b'_1 \rangle}{\langle b'_1, b'_1 \rangle} b'_1 = x - \frac{\int_0^1 x \cdot 1 dx}{\int_0^1 1^2 dx} \cdot 1 = x - \frac{1}{2} \\ b'_3 = b_3 - \frac{\langle b_3, b'_1 \rangle}{\langle b'_1, b'_1 \rangle} b'_1 - \frac{\langle b_3, b'_2 \rangle}{\langle b'_2, b'_2 \rangle} b'_2 = x^2 - \frac{\int_0^1 x^2 \cdot 1 dx}{\int_0^1 1^2 dx} \cdot 1 - \frac{\int_0^1 x^2 \cdot (x - \frac{1}{2}) dx}{\int_0^1 (x - \frac{1}{2})^2 dx} \cdot (x - \frac{1}{2}) = x^2 - x + \frac{1}{6} \end{cases}$$

Otrzymując w ten sposób bazę ortogonalną. Dla wyznaczenia bazy ortonormalnej, normalizujemy otrzymane wektory:

$$\begin{cases} b''_1 = \frac{b'_1}{|b'_1|} = \frac{1}{\sqrt{\int_0^1 1^2 dx}} \cdot 1 = 1 \\ b''_2 = \frac{b'_2}{|b'_2|} = \frac{1}{\sqrt{\int_0^1 (x - \frac{1}{2})^2 dx}} \cdot (x - \frac{1}{2}) = \sqrt{3}(2x - 1) \\ b''_3 = \frac{b'_3}{|b'_3|} = \frac{1}{\sqrt{\int_0^1 (x^2 - x + \frac{1}{6})^2 dx}} \cdot (x^2 - x + \frac{1}{6}) = \sqrt{5}(6x^2 - 6x + 1) \end{cases}$$

4.2 Forma kwadratowa

W niniejszym rozdziale zajmiemy się (nietypowo dla algebry liniowej) funkcjami kwadratowymi, a dokładnie wielomianami kwadratowymi wielu zmiennych. Zaczniemy od zdefiniowania wielomianu wielu zmiennych.

Definicja 4.20

(Rzeczywistym) wielomianem zmiennych $x_1, \dots, x_n$ nazywamy funkcję $P : \mathbb{R}^n \rightarrow \mathbb{R}$, gdzie $P(x_1, \dots, x_n)$ jest kombinacją liniową jednomianów, tzn. funkcji postaci

$$\alpha \cdot x_1^{k_1} x_2^{k_2} \cdot \dots \cdot x_n^{k_n}$$

gdzie $k_1, \dots, k_n \in \mathbb{N}$ oraz $\alpha \in \mathbb{R}$. *Stopniem* wielomianu wielu zmiennych nazywamy najwyższy ze stopni tworzących go jednomianów, gdzie stopień jednomianu to $k_1 + \dots + k_n$.

Przykład 1

Ustal stopnie każdego z następujących wielomianów wielu zmiennych:

$$P(x, y, z) = -2x^3 + 3x^2yz - 5x^2z^3, \quad Q(x, y) = x^4 + 7x^3y^2 - y^3, \quad R(x, y, z, t) = 2xyzt - x - y - z$$

Rozwiązanie. Wielomian P ma stopień 5 (jednomian najwyższego stopnia to $-5x^2z^3$). Wielomian Q ma stopień 5 (jednomian najwyższego stopnia to $7x^3y^2$). Wielomian R ma stopień 4 (jednomian najwyższego stopnia to $2xyzt$).

Przykład 2

Ogólna postać wielomianu stopnia pierwszego dwóch zmiennych to:

$$P(x, y) = ax + by + c, \quad a, b, c \in \mathbb{R}$$

Ogólna postać wielomianu stopnia drugiego dwóch zmiennych to:

$$P(x, y) = ax^2 + bxy + cy^2 + dx + ey + f, \quad a, b, c, d, e, f \in \mathbb{R}$$

Definicja 4.21

Wielomian wielu zmiennych nazywamy *wielomianem jednorodnym*, jeśli jest sumą wyłącznie jednomianów tego samego stopnia.

Przykład 3

Ogólna postać **jednorodnego** wielomianu stopnia pierwszego trzech zmiennych to:

$$P(x, y, z) = ax + by + cz, \quad a, b, c \in \mathbb{R}$$

Ogólna postać **jednorodnego** wielomianu stopnia drugiego trzech zmiennych to:

$$P(x, y, z) = ax^2 + by^2 + cz^2 + dxy + eyz + fzx, \quad a, b, c, d, e, f \in \mathbb{R}$$

Poniższy warunek jest uogólnieniem własności jednorodności (odnoszonej do przekształceń liniowych) na wielomiany wyższych stopni:

Fakt 4.22

Jeśli $P(x_1, \dots, x_n)$ jest wielomianem jednorodnym stopnia k zmiennych $x_1, \dots, x_n$, to dla dowolnej liczby rzeczywistej t zachodzi:

$$P(tx_1, \dots, tx_n) = t^k \cdot P(x_1, \dots, x_n)$$

W niniejszym rozdziale będziemy zajmować się jednorodnymi wielomianami wielu zmiennych stopnia drugiego. Zauważmy, że:

Fakt 4.23

Jeśli $P : \mathbb{R}^n \rightarrow \mathbb{R}$ jest jednorodnym wielomianem stopnia drugiego zmiennych $x_1, \dots, x_n$, to jest on postaci:

$$P(X) = X^\top \cdot A \cdot X$$

dla pewnej symetrycznej macierzy A . Macierz A jest przy tym wyznaczona jednoznacznie przez wielomian P .

Dowód. Ogólna postać wielomianu jednorodnego stopnia drugiego to:

$$P(x_1, \dots, x_n) = \sum_{1 \leq i, j \leq n} a_{ij} x_i x_j$$

który to warunek można zapisać jako:

$$P(x_1, \dots, x_n) = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

przy czym współczynniki a_{ij} nie są wyznaczone jednoznacznie przez wielomian P , gdyż wyrazy $a_{ij}x_i x_j$ oraz $a_{ji}x_j x_i$ są nierozróżnialne z uwagi na przemienność mnożenia liczb rzeczywistych. Niemniej sumy współczynników $a_{ij} + a_{ji}$ (dla $i \neq j$) oraz współczynniki a_{ii} są już wyznaczone jednoznacznie. Stąd przy dodatkowym założeniu, że macierz $A = (a_{ij})$ jest macierzą symetryczną (tzn. $a_{ij} = a_{ji}$) macierz ta jest wyznaczona jednoznacznie. $\square$

Przykład 4

Zapisz wielomian $P(x, y, z) = 2x^2 - 6xy + 8yz + y^2 + 3z^2$ w postaci macierzowej.

Rozwiązanie.

$$P(x, y, z) = \begin{pmatrix} x & y & z \end{pmatrix} \cdot \begin{pmatrix} 2 & -3 & 0 \\ -3 & 1 & 4 \\ 0 & 4 & 3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Powyższa postać stanowi motywację do następującej definicji formy kwadratowej na dowolnej przestrzeni liniowej.

Definicja 4.24

Formą kwadratową na (skończenie wymiarowej) przestrzeni liniowej V nazywamy funkcję $Q : V \rightarrow \mathbb{R}$, która dla pewnej bazy $\mathcal{B}$ przestrzeni V ma postać:

$$Q(v) = [v]_{\mathcal{B}}^\top \cdot A \cdot [v]_{\mathcal{B}}$$

dla pewnej symetrycznej macierzy $A \in M_{n \times n}(\mathbb{R})$, gdzie $n = \dim V$. Macierz A nazywamy *macierzą formy Q w bazie $\mathcal{B}$* i oznaczamy $m^{\mathcal{B}\mathcal{B}}(Q) = A$.

Powyższą definicję można przeformułować w następujący sposób: forma kwadratowa to jednorodny wielomian drugiego stopnia, którego zmiennymi są współrzędne w pewnej bazie. Z tego też powodu będziemy (dla uproszczenia) posługiwać się również zapisem $Q(x_1, \dots, x_n)$ dla oznaczenia $Q(x_1 b_1 + \dots + x_n b_n)$, gdzie $\mathcal{B} = (b_1, \dots, b_n)$ będzie ustaloną bazą przestrzeni V .

Przykład 5

Sprawdź, że odwzorowanie $Q : \mathbb{R}_2[x] \rightarrow \mathbb{R}$ zdefiniowane $Q(P) = (P(2))^2$ jest formą kwadratową.

Rozwiązanie. Przekształcenie Q jest zdefiniowane następująco:

$$Q(ax^2 + bx + c) = (4a + 2b + c)^2 = 16a^2 + 4b^2 + c^2 + 16ab + 4bc + 8ca$$

czyli jest jednorodnym wielomianem kwadratowym zmiennych a, b, c , które są współrzędnymi w bazie $\mathcal{B} = (1, x, x^2)$. Inaczej:

$$Q(ax^2 + bx + c) = \begin{pmatrix} a & b & c \end{pmatrix} \cdot \begin{pmatrix} 16 & 8 & 4 \\ 8 & 4 & 2 \\ 4 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = [ax^2 + bx + c]_{\mathcal{B}}^{\top} \cdot \begin{pmatrix} 16 & 8 & 4 \\ 8 & 4 & 2 \\ 4 & 2 & 1 \end{pmatrix} \cdot [ax^2 + bx + c]_{\mathcal{B}}$$

Przykład 6

Sprawdź, że odwzorowanie $Q : M_{2 \times 2}(\mathbb{R}) \rightarrow \mathbb{R}$ zdefiniowane $Q(A) = \det A$ jest formą kwadratową.

Rozwiązanie. Przekształcenie Q jest zdefiniowane następująco:

$$Q\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = ad - bc$$

czyli jest jednorodnym wielomianem kwadratowym zmiennych a, b, c, d które są współrzędnymi w bazie $\mathcal{B} = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right)$. Inaczej:

$$Q\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right]_{\mathcal{B}}^{\top} \cdot \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \end{pmatrix} \cdot \left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right]_{\mathcal{B}}$$

Fakt 4.25

Niech $\mathcal{B}$ będzie bazą (skończenie wymiarowej) przestrzeni liniowej V , zaś $Q : V \rightarrow \mathbb{R}$ formą kwadratową. Wówczas macierz $A = m^{\mathcal{B}\mathcal{B}}(Q)$ formy Q w bazie $\mathcal{B}$ jest jednoznacznie wyznaczona przez Q .

Dowód. Oznaczmy $A = (a_{ij})$. Wówczas:

$$Q(b_i) = [b_i]_{\mathcal{B}}^{\top} \cdot A \cdot [b_i] = (0 \dots 0 \ 1 \ 0 \dots 0) \cdot A \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = a_{ii}$$

co jednoznacznie wyznacza wyrazy na przekątnej macierzy A . Ponadto dla $i \neq j$:

$$Q(b_i + b_j) = [b_i + b_j]_{\mathcal{B}}^{\top} \cdot A \cdot [b_i + b_j] = (0 \dots 0 \ 1 \ 0 \dots 0 \ 1 \ 0 \dots 0) \cdot A \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = a_{ii} + a_{jj} + a_{ij} + a_{ji}$$

co wobec wcześniejszego wyznaczenia wyrazów na przekątnej macierzy A , jednoznacznie wyznacza sumę $a_{ij} + a_{ji}$. Dzięki założeniu o symetryczności macierzy A (które oznacza $a_{ij} = a_{ji}$) to jednoznacznie wyznacza wyrazy $a_{ij} = a_{ji}$. $\square$

Wygodniejszym sposobem definiowania formy kwadratowej oraz badania jej macierzy jest odwołanie się do pojęcia formy dwuliniowej.

Definicja 4.26

Formę dwuliniową na przestrzeni liniowej V nazywamy odwzorowanie $\Phi : V \times V \rightarrow \mathbb{R}$, które jest liniowe względem każdej współrzędnej. Formę taką nazywamy *symetryczną*, jeśli $\Phi(v, w) = \Phi(w, v)$ dla dowolnych $v, w \in V$.

Fakt 4.27

Jeśli $\Phi : V \times V \rightarrow \mathbb{R}$ jest formą dwuliniową na przestrzeni liniowej V , zaś $\mathcal{B} = (b_1, \dots, b_n)$ bazą przestrzeni V , to:

$$\Phi(v, w) = [v]_{\mathcal{B}}^{\top} \cdot A \cdot [w]_{\mathcal{B}}$$

dla macierzy $A = (a_{ij}) \in M_{n \times n}(\mathbb{R})$, gdzie $a_{ij} = \Phi(b_i, b_j)$.

Dowód. Niech $\mathcal{B} = (b_1, \dots, b_n)$. Jeśli $v = x_1 b_1 + \dots + x_n b_n$ oraz $w = y_1 b_1 + \dots + y_n b_n$, to z uwagi na dwuliniowość (liniowość względem każdej współrzędnej) otrzymujemy:

$$\Phi(v, w) = \Phi(x_1 b_1 + \dots + x_n b_n, y_1 b_1 + \dots + y_n b_n) = \sum_{1 \leq i, j \leq n} \Phi(b_i, b_j) x_i y_j$$

co, przyjmując oznaczenia $A = (a_{ij})$, gdzie $a_{ij} = \Phi(b_i, b_j)$, można zapisać jako:

$$\Phi(v, w) = [v]_{\mathcal{B}}^{\top} \cdot A \cdot [w]_{\mathcal{B}}$$

$\square$

Wniosek 4.28

Forma dwuliniowa $\Phi : V \times V \rightarrow \mathbb{R}$ jest symetryczna wtedy i tylko wtedy, gdy jej macierz $m^{\mathcal{B}\mathcal{B}}(\Phi)$ jest symetryczna.

Wniosek 4.29

Jeśli $\Phi : V \times V \rightarrow \mathbb{R}$ jest symetryczną formą dwuliniową na przestrzeni liniowej V , to $Q : V \rightarrow \mathbb{R}$ zdefiniowane $Q(v) = \Phi(v, v)$ jest formą kwadratową. Formę dwuliniową Φ nazywamy *stowarzyszoną* z formą kwadratową Q .

Symetryczne formy dwuliniowe oraz formy kwadratowe są na tyle mocno związane ze sobą, że często używamy tego samego oznaczenia na formę kwadratową oraz stowarzyszoną z nią formę kwadratową, stosując równocześnie zapis $Q(v)$ (w odniesieniu do formy kwadratowej) oraz $Q(v, w)$ (w odniesieniu do stowarzyszonej z nią formy dwuliniowej).

Przykład 7

Dana jest przestrzeń euklidesowa $\mathbb{R}_2[x]$, gdzie iloczyn skalarny to $\langle P_1, P_2 \rangle = \int_0^1 P_1(x)P_2(x)dx$. Sprawdź, że odwzorowanie $Q : \mathbb{R}_2[x] \rightarrow \mathbb{R}$ zdefiniowane $Q(P) = |P|^2$ jest formą kwadratową. *Rozwiązanie.* Odwzorowanie $Q : \mathbb{R}_2[x] \times \mathbb{R}_2[x] \rightarrow \mathbb{R}$ zdefiniowane $Q(P_1, P_2) = \langle P_1, P_2 \rangle$ jest symetryczną formą dwuliniową. Wobec tego $Q(P) = \langle P, P \rangle = |P|^2$ jest stowarzyszoną z nią formą kwadratową.

Przykład 8

Wyznacz symetryczną formę dwuliniową $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ stowarzyszoną z formą kwadratową $Q : \mathbb{R}^2 \rightarrow \mathbb{R}$ zdefiniowaną:

$$Q\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = ax_1^2 + bx_2^2 + cx_1x_2$$

Rozwiązanie. Forma kwadratowa oraz stowarzyszona z nią symetryczna forma dwuliniowa mają taką samą macierz. Macierzą formy kwadratowej Q jest:

$$\begin{pmatrix} a & \frac{c}{2} \\ \frac{c}{2} & b \end{pmatrix}$$

Symetryczna forma dwuliniowa stowarzyszona z Q ma taką samą macierz, czyli:

$$\Phi\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a & \frac{c}{2} \\ \frac{c}{2} & b \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = ax_1y_1 + bx_2y_2 + \frac{c}{2}x_1y_2 + \frac{c}{2}x_2y_1$$

Fakt 4.30

Dla dowolnej formy kwadratowej $Q : V \rightarrow \mathbb{R}$ istnieje dokładnie jedna symetryczna forma dwuliniowa $\Phi : V \times V \rightarrow \mathbb{R}$ stowarzyszona z Q , tzn $Q(v) = \Phi(v, v)$ dla każdego $v \in V$.

Dowód. Niech $\mathcal{B} = (b_1, \dots, b_n)$ będzie bazą przestrzeni V . Jeśli:

$$Q(v) = [v]_{\mathcal{B}}^{\top} \cdot A \cdot [v]_{\mathcal{B}}$$

dla pewnej symetrycznej macierzy $A \in M_{n \times n}(\mathbb{R})$, to

$$\Phi(v, w) = [v]_{\mathcal{B}}^{\top} \cdot A \cdot [w]_{\mathcal{B}}$$

jest szukaną symetryczną formą dwuliniową. □

Fakt 4.31

Niech $\mathcal{B}$ i $\mathcal{C}$ będą bazami przestrzeni liniowej V , zaś $Q : V \rightarrow \mathbb{R}$ formą dwuliniową. Wówczas:

$$m^{\mathcal{C}\mathcal{C}}(Q) = (m_{\mathcal{B}}^{\mathcal{C}}(\text{id}))^{\top} \cdot m^{\mathcal{B}\mathcal{B}}(Q) \cdot m_{\mathcal{B}}^{\mathcal{C}}(\text{id})$$

Dowód. Zauważmy, że:

$$Q(v) = [v]_{\mathcal{B}}^{\top} \cdot m^{\mathcal{B}\mathcal{B}}(Q) \cdot [v]_{\mathcal{B}} = (m_{\mathcal{B}}^{\mathcal{C}}(\text{id}))^{\top} \cdot [v]_{\mathcal{C}}^{\top} \cdot m^{\mathcal{B}\mathcal{B}}(Q) \cdot (m_{\mathcal{B}}^{\mathcal{C}}(\text{id})) \cdot [v]_{\mathcal{C}}$$

$$= ([v]_{\mathcal{C}})^{\top} ((m_{\mathcal{B}}^{\mathcal{C}}(\text{id}))^{\top} m^{\mathcal{B}\mathcal{B}}(Q) \cdot m_{\mathcal{B}}^{\mathcal{C}}(\text{id})) \cdot [v]_{\mathcal{C}}$$

Porównując ten wzór ze wzorem na macierz formy Q w bazie $\mathcal{C}$:

$$Q(v) = [v]_{\mathcal{C}}^{\top} \cdot m^{\mathcal{C}\mathcal{C}}(Q) \cdot [v]_{\mathcal{C}}$$

otrzymujemy:

$$m^{\mathcal{C}\mathcal{C}}(Q) = (m_{\mathcal{B}}^{\mathcal{C}}(\text{id}))^{\top} \cdot m^{\mathcal{B}\mathcal{B}}(Q) \cdot m_{\mathcal{B}}^{\mathcal{C}}(\text{id})$$

□

Twierdzenie 4.32: Diagonalizacja formy kwadratowej metodą Lagrange'a

Niech V będzie skończenie wymiarową przestrzenią liniową, zaś $Q : V \rightarrow \mathbb{R}$ formą kwadratową. Wówczas istnieje taka baza $\mathcal{B} = (b_1, \dots, b_n)$, w której forma Q diagonalizuje się, tzn. $m^{\mathcal{B}\mathcal{B}}(Q)$ jest macierzą diagonalną. Innymi słowy:

$$Q(x_1 b_1 + \dots + x_n b_n) = \alpha_1 x_1^2 + \dots + \alpha_n x_n^2$$

Zgodnie ze wcześniejszymi uwagami dotyczącymi oznaczeń, powyższy wzór będziemy również zapisywać jako:

$$Q(x_1, \dots, x_n) = \alpha_1 x_1^2 + \dots + \alpha_n x_n^2$$

Zamiast formalnego dowodu pokażemy na przykładzie działanie metody Lagrange'a. Dana jest forma kwadratowa $Q : \mathbb{R}^3 \rightarrow \mathbb{R}$, która w bazie standardowej $\mathcal{B}$ ma wzór:

$$Q(x, y, z) = 2x^2 + 3y^2 - z^2 + 4xy + 6yz + 8zx$$

Zaczynamy od „zwinienia w kwadrat” wszystkich składników zawierających x :

$$= 2(x + y + 2z)^2 - 2y^2 - 8z^2 - 8yz + 3y^2 - z^2 + 6yz = 2(x + y + 2z)^2 + y^2 - 9z^2 - 2yz$$

wprowadzamy nową współrzędną $x' = x + y + 2z$:

$$= 2(x')^2 + y^2 - z^2 - 2yz$$

Następnie powtarzamy tę samą procedurę ze zmienną y :

$$= 2(x')^2 + (y - z)^2 - z^2 - 9z^2 = 2(x')^2 + (y')^2 - 10(z')^2$$

gdzie $y' = y - z$ oraz $z' = z$. W ten sposób znaleźliśmy nowe współrzędne (x', y', z') , w których forma Q ma postać diagonalną. Pozostaje pytanie jak przetłumaczyć zamianę współrzędnych na zamianę bazy. Zauważmy, że nowe współrzędne można wyrazić przy pomocy starych w następujący sposób:

$$\begin{cases} x' = x + y + 2z \\ y' = y - z \\ z' = z \end{cases} \quad \text{czyli} \quad \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Powyższą równość można w następujący sposób zapisać przy użyciu baz:

$$[v]_{\mathcal{B}'} = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \cdot [v]_{\mathcal{B}} \quad \text{czyli} \quad [v]_{\mathcal{B}} = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \cdot [v]_{\mathcal{B}'} = \begin{pmatrix} 1 & -1 & -3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot [v]_{\mathcal{B}'}$$

Wektory nowej bazy b'_1, b'_2, b'_3 mają zatem współrzędne:

$$[b'_1]_{\mathcal{B}} = \begin{pmatrix} 1 & -1 & -3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{czyli} \quad b'_1 = b_1$$

$$[b'_2]_{\mathcal{B}} = \begin{pmatrix} 1 & -1 & -3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad \text{czyli} \quad b'_2 = -b_1 + b_2$$

$$[b'_3]_{\mathcal{B}} = \begin{pmatrix} 1 & -1 & -3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \\ 1 \end{pmatrix} \quad \text{czyli} \quad b'_3 = -3b_1 + b_2 + b_3$$

Przedstawioną powyżej metodę Lagrange'a można podsumować jako wielokrotne powtarzanie następującego kroku:

(1) Zwijamy w kwadrat wszystkie składniki zawierające x_i , po czym wprowadzamy nową współrzędną x'_i (która występuje już tylko w jednym składniku – składniku $(x'_i)^2$).

Krok ten powtarzamy dla kolejnych współrzędnych aż do momentu, gdy pozbędziemy się wszystkich wyrazów mieszanych. Jedynym problemem może być sytuacja, gdy w pewnym momencie zmienna x_i nie występuje w wyrazie kwadratowym x_i^2 , natomiast występuje w wyrazach mieszanych. W takiej sytuacji wykonujemy drugi krok metody Lagrange'a:

(2) Jeśli x_i ani x_j nie występują w wyrazach kwadratowych, natomiast występują w wyrazie mieszanym $x_i x_j$, wykonujemy podstawienie:

$$\begin{cases} x_i = x'_i + x'_j \\ x_j = x'_i - x'_j \end{cases} \quad \text{czyli} \quad \begin{cases} x'_i = \frac{1}{2}x_i + \frac{1}{2}x_j \\ x'_j = \frac{1}{2}x_i - \frac{1}{2}x_j \end{cases}$$

Stąd dostajemy $x_i x_j = (x'_i)^2 - (x'_j)^2$ i wracamy do kroku (1).

Diagonalizację formy z użyciem drugiego kroku pokazuje poniższy przykład.

Przykład 9

Zdiagonalizuj formę kwadratową $Q : \mathbb{R}^3 \rightarrow \mathbb{R}$, która w bazie standardowej ma wzór $Q(x, y, z) = xy + yz$ i podaj bazę, w której przyjmuje wyznaczoną postać diagonalną.

Rozwiązanie. Ponieważ krok (1) metody Lagrange'a jest niewykonalny z uwagi na brak składników kwadratowych, wykonujemy podstawienie z drugiego kroku, tzn. $x = x' + y'$, $y = x' - y'$, czyli $x' = \frac{1}{2}x + \frac{1}{2}y$, $y' = \frac{1}{2}x - \frac{1}{2}y$:

$$Q(x, y, z) = xy + yz = (x' + y')(x' - y') + (x' - y')z = (x')^2 - (y')^2 + x'z - y'z$$

Następnie wykonujemy krok (1) metody Lagrange'a, tzn. zwijamy w kwadrat wyrazy zawierające x' :

$$= (x' + \frac{1}{2}z)^2 - \frac{1}{4}z^2 - (y')^2 - y'z$$

wykonujemy podstawienie $x'' = x' + \frac{1}{2}z$:

$$= (x'')^2 - \frac{1}{4}z^2 - (y')^2 - y'z$$

i ponownie wykonujemy krok (1), podstawiając $y'' = y' + \frac{1}{2}z$:

$$= (x'')^2 - (y' + \frac{1}{2}z)^2 = (x'')^2 - (y'')^2$$

Nowe współrzędne można w następujący sposób wyrazić przy pomocy starych (przyjmujemy

$z'' = z' = z$):

$$\begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Wobec tego:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix}$$

Wobec tego wektory nowej bazy b_1'', b_2'', b_3'' to:

$$b_1'' = b_1 + b_2$$

$$b_2'' = b_1 - b_2$$

$$b_3'' = -b_1 + b_3$$

Definicja 4.33

Niech V będzie (skończenie wymiarową) przestrzenią liniową. Formę kwadratową $Q : V \rightarrow \mathbb{R}$ nazywamy:

- *dodatnio określoną*, jeśli $Q(v) > 0$ dla każdego $v \neq 0$,
- *ujemnie określoną*, jeśli $Q(v) < 0$ dla każdego $v \neq 0$,
- *dodatnio półokreśloną*, jeśli $Q(v) \geq 0$ dla każdego v ,
- *ujemnie półokreśloną*, jeśli $Q(v) \leq 0$ dla każdego v .

Macierz $A \in M_{n \times n}(\mathbb{R})$ nazywamy *macierzą dodatnio/ujemnie określoną/półokreśloną*, jeśli odpowiednią własność ma forma kwadratowa $Q : \mathbb{R}^n \rightarrow \mathbb{R}$, dla której A jest macierzą w bazie standardowej.

Przykład 10

Sprawdź czy forma kwadratowa $Q : \mathbb{R}_2[x] \rightarrow \mathbb{R}$ zdefiniowana $Q(P) = (P(2))^2$ jest dodatnio określona.

Rozwiązanie. Zauważmy, że $\Phi : \mathbb{R}_2[x] \times \mathbb{R}_2[x] \rightarrow \mathbb{R}$ zdefiniowane $\Phi(P_1, P_2) = P_1(2) \cdot P_2(2)$ jest symetryczną formą dwuliniową, a zatem $Q(P) = \Phi(P, P)$ jest formą kwadratową. Dla dowolnego P zachodzi $Q(P) \geq 0$, ale warunek $Q(P) = 0$ zachodzi dla niektórych niezerowych wielomianów. Wobec tego forma Q nie jest dodatnio określona, ale jest dodatnio półokreślona.

Fakt 4.34

Niech V będzie (skończenie wymiarową) przestrzenią liniową, zaś $Q : V \rightarrow \mathbb{R}$ formą kwadratową. Jeśli $\mathcal{B}$ jest taką bazą, że:

$$m^{\mathcal{B}\mathcal{B}}(Q) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

to forma Q jest:

- dodatnio określona $\iff \lambda_1, \dots, \lambda_n > 0$,
- ujemnie określona $\iff \lambda_1, \dots, \lambda_n < 0$,
- dodatnio półokreślona $\iff \lambda_1, \dots, \lambda_n \geq 0$,
- ujemnie półokreślona $\iff \lambda_1, \dots, \lambda_n \leq 0$.

Dowód. Niech $\mathcal{B} = (b_1, \dots, b_n)$. Wówczas dla dowolnego wektora $v = \alpha_1 b_1 + \dots + \alpha_n b_n$ otrzymujemy:

$$Q(v) = [v]_{\mathcal{B}}^T \cdot m^{\mathcal{B}\mathcal{B}}(Q) \cdot [v]_{\mathcal{B}} = (\alpha_1 \ \dots \ \alpha_n) \cdot \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \lambda_1 \alpha_1^2 + \dots + \lambda_n \alpha_n^2$$

Forma Q jest dodatnio określona wtedy i tylko wtedy, gdy $Q(v) > 0$ dla dowolnego niezerowego wektora v , czyli

$$\lambda_1 \alpha_1^2 + \dots + \lambda_n \alpha_n^2 > 0$$

dla dowolnych $\alpha_1, \dots, \alpha_n \in \mathbb{R}$, które nie są jednocześnie równe 0. To jest prawdą wtedy i tylko wtedy, gdy $\lambda_1, \dots, \lambda_n > 0$. Podobnie dowodzimy pozostałych warunków. $\square$

Wygodną metodą sprawdzania dodatniej/ujemnej określoności macierzy jest kryterium Sylwestera, opierające się na wyznaczeniu minorów głównych macierzy formy (w dowolnie wybranej bazie).

Definicja 4.35

Minorem głównym stopnia k macierzy A nazywamy minor powstały przez wybranie pierwszych k wierszy i pierwszych k kolumn macierzy A .

Przykład 11

Wyznacz wszystkie minory główne macierzy $\begin{pmatrix} 1 & 2 & 0 & 3 \\ 1 & 4 & 9 & -1 \\ 2 & 4 & 3 & 1 \\ 5 & 7 & 1 & 0 \end{pmatrix}$.

Rozwiązanie. Minory główne powyższej macierzy to:

$$\det(1) = 1, \quad \det \begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix} = 2, \quad \det \begin{pmatrix} 1 & 2 & 0 \\ 1 & 4 & 9 \\ 2 & 4 & 3 \end{pmatrix} = 6, \quad \det \begin{pmatrix} 1 & 2 & 0 & 3 \\ 1 & 4 & 9 & -1 \\ 2 & 4 & 3 & 1 \\ 5 & 7 & 1 & 0 \end{pmatrix} = 19$$

Fakt 4.36: kryterium Sylwestera

Macierz $A \in M_{n \times n}(\mathbb{R})$ jest:

- 1) dodatnio określona $\iff$ wszystkie minory główne macierzy A są dodatnie,
- 2) ujemnie określona $\iff$ minory główne stopni 1, 3, 5, ... macierzy A są ujemne, natomiast minory główne stopni 2, 4, 6, ... są dodatnie.

Fakt 4.37

Niech V będzie przestrzenią liniową. Wówczas forma kwadratowa $Q : V \rightarrow \mathbb{R}$ jest dodatnio (ujemnie) określona, jeśli jej macierz $m^{\mathcal{B}\mathcal{B}}(Q)$ w dowolnie wybranej bazie $\mathcal{B}$ jest dodatnio (ujemnie) określona.

Kryterium Sylwestera jest natychmiastowe dla bazy $\mathcal{B}$, w której forma kwadratowa Q się diagonalizuje (tzn. macierz $m^{\mathcal{B}\mathcal{B}}(Q)$ jest diagonalna). Dowód kryterium dla dowolnej bazy jest trudniejszy i pominiemy go.

Przykład 12

Sprawdź, które z poniższych macierzy są dodatnio określone, a które są ujemnie określone:

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 0 \\ 1 & 0 & 7 \end{pmatrix} \quad \begin{pmatrix} -2 & 3 & 0 \\ 3 & -5 & 1 \\ 0 & 1 & -6 \end{pmatrix} \quad \begin{pmatrix} 7 & 3 & 2 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

Rozwiązanie. Stosujemy kryterium Sylwestera. Dla pierwszej macierzy otrzymujemy:

$$\det(1) = 1 > 0 \quad \det \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} = 1 > 0 \quad \det \begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 0 \\ 1 & 0 & 7 \end{pmatrix} = 2 > 0$$

czyli jest to macierz dodatnio określona.

Dla drugiej macierzy otrzymujemy:

$$\det(-2) = -2 < 0 \quad \det \begin{pmatrix} -2 & 3 \\ 3 & -5 \end{pmatrix} = 1 > 0 \quad \det \begin{pmatrix} -2 & 3 & 0 \\ 3 & -5 & 1 \\ 0 & 1 & -6 \end{pmatrix} = -4 < 0$$

czyli jest to macierz ujemnie określona.

Dla trzeciej macierzy otrzymujemy:

$$\det(7) = 7 < 0 \quad \det \begin{pmatrix} 7 & 3 \\ 3 & 1 \end{pmatrix} = -2 < 0$$

czyli nie jest to ani macierz dodatnio określona, ani ujemnie określona.

Definicja 4.38

Niech V będzie skończenie wymiarową przestrzenią liniową oraz $n = \dim V$. *Sygnaturą* formy kwadratowej $Q : V \rightarrow \mathbb{R}$ nazywamy parę liczb naturalnych (p, q) , gdzie p i q to odpowiednio liczba dodatnich i ujemnych wyrazów na przekątnej macierzy formy $m^{\mathcal{B}\mathcal{B}}(Q)$ w takiej bazie $\mathcal{B}$, że macierz ta jest diagonalna, tzn.

$$m^{\mathcal{B}\mathcal{B}}(Q) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

gdzie $\lambda_1, \dots, \lambda_p > 0$, $\lambda_{p+1}, \dots, \lambda_{p+q} < 0$ oraz $\lambda_{p+q+1} = \dots = \lambda_n = 0$.

Fakt 4.39

Sygnatura formy kwadratowej nie zależy od wyboru bazy.

Przykład 13

Wyznacz sygnaturę następujących form kwadratowych:

$$Q_1(x, y, z) = 2x^2 + 3y^2 - z^2 + 4xy + 6yz + 8zx$$

$$Q_2(x, y, z) = xy + yz$$

Rozwiązanie. W poprzednich przykładach znaleźliśmy bazy w których formy Q_1 i Q_2 są diagonalne:

$$Q_1(x', y', z') = 2(x')^2 + (y')^2 - 10(z')^2$$

$$Q_2(x'', y'', z'') = (x'')^2 - (y'')^2$$

Stąd sygnatura formy Q_1 to $(2, 1)$, zaś sygnatura formy Q_2 to $(1, 1)$.

4.3 Twierdzenie spektralne**Definicja 4.40**

Izometrią liniową przestrzeni euklidesowej V nazywamy przekształcenie liniowe $F : V \rightarrow V$, które zachowuje odległości, tzn. dla dowolnych $u, v \in V$ zachodzi:

$$d(u, v) = d(F(u), F(v))$$

Fakt 4.41

Przekształcenie liniowe $F : V \rightarrow V$ jest izometrią wtedy i tylko wtedy, gdy zachowuje długości wektorów, tzn.

$$|v| = |F(v)|$$

dla dowolnego wektora $v \in V$.

Dowód. Izometria liniowa (jak każde przekształcenie liniowe) spełnia warunek $F(0) = 0$, więc:

$$|F(v)| = d(F(v), 0) = d(F(v), F(0)) = d(v, 0) = |v|$$

czyli F zachowuje długości wektorów.

Z drugiej strony, jeśli F jest liniowe i zachowuje długości wektorów, to:

$$d(u, v) = |u - v| = |F(u - v)| = |F(u) - F(v)| = d(F(u), F(v))$$

czyli jest izometrią. □

Fakt 4.42

Dla dowolnych wektorów $u, v \in V$ przestrzeni euklidesowej V zachodzi warunek:

$$\langle u, v \rangle = \frac{1}{4}(|u + v|^2 - |u - v|^2)$$

Dowód. Korzystając z własności iloczynu skalarnego otrzymujemy:

$$\begin{aligned} |u + v|^2 - |u - v|^2 &= \langle u + v, u + v \rangle - \langle u - v, u - v \rangle \\ &= (\langle u, u \rangle + 2\langle u, v \rangle + \langle v, v \rangle) - (\langle u, u \rangle - 2\langle u, v \rangle + \langle v, v \rangle) = 4\langle u, v \rangle \end{aligned}$$

□

Fakt 4.43

Izometria liniowa $F : V \rightarrow V$ zachowuje iloczyn skalarny oraz zachowuje miary kątów.

Dowód. Skoro F zachowuje długości wektorów, to zachowuje również iloczyn skalarny, gdyż zgodnie z Faktem 4.42:

$$\begin{aligned} \langle F(u), F(v) \rangle &= \frac{1}{4}(|F(u) + F(v)|^2 - |F(u) - F(v)|^2) \\ &= \frac{1}{4}(|F(u + v)|^2 - |F(u - v)|^2) = \frac{1}{4}(|u + v|^2 - |u - v|^2) = \langle u, v \rangle \end{aligned}$$

Skoro F zachowuje długości wektorów oraz iloczyn skalarny, to zachowuje również miary kątów (w szczególności ortogonalność wektorów), gdyż:

$$\angle(F(u), F(v)) = \arccos \frac{\langle F(u), F(v) \rangle}{|F(u)| \cdot |F(v)|} = \arccos \frac{\langle u, v \rangle}{|u| \cdot |v|} = \angle(u, v)$$

□

Przykład 1

Dana jest przestrzeń euklidesowa $\mathbb{R}^4$ ze standardowym iloczynem skalarnym. Sprawdź, czy przekształcenie liniowe $F : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ zdefiniowane:

$$F\left(\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}\right) = \begin{pmatrix} y \\ z \\ t \\ x \end{pmatrix}$$

jest izometrią.

Rozwiązanie. Dla dowolnego wektora $v = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$ zachodzi:

$$|v| = \sqrt{x^2 + y^2 + z^2 + t^2} = |F(v)|$$

Wobec tego przekształcenie F zachowuje długości wektorów, czyli jest izometrią.

Przykład 2

Dana jest przestrzeń euklidesowa $\mathbb{R}_3[x]$ z iloczynem skalarnym $\langle P, Q \rangle = \int_0^1 P(x)Q(x)dx$. Sprawdź, czy przekształcenie liniowe $F : \mathbb{R}_3[x] \rightarrow \mathbb{R}_3[x]$ zdefiniowane:

$$F(ax^2 + bx + c) = bx^2 + cx + a$$

jest izometrią.

Rozwiązanie. Zauważmy np. że dla $P(x) = x^2$ mamy $|P| = \int_0^1 x^4 dx = \frac{1}{5}$, ale $|F(P)| = \int_0^1 1 dx = 1$, czyli F nie zachowuje długości wektorów. Nie jest zatem izometrią.

Definicja 4.44

Macierzą izometrii albo *macierzą ortogonalną* nazywamy taką macierz $A \in M_{n \times n}(\mathbb{R})$, że przekształcenie $F_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ (przestrzeni $\mathbb{R}^n$ ze standardowym iloczynem skalarnym) jest izometrią.

Fakt 4.45

Dla dowolnej macierzy kwadratowej $A \in M_{n \times n}(\mathbb{R})$ następujące warunki są równoważne:

- 1) A jest macierzą ortogonalną (macierzą izometrii),
- 2) kolumny macierzy A tworzą bazę ortonormalną $\mathbb{R}^n$,
- 3) wiersze macierzy A tworzą bazę ortonormalną $\mathbb{R}^n$,
- 4) A jest odwracalna i $A^{-1} = A^\top$.

Dowód. Oznaczmy $A_1, \dots, A_n$ kolumny macierzy A . Zauważmy, że warunek (4) można zapisać jako $A^\top A = I$. Ponieważ każdy wyraz macierzy $A^\top A$ jest postaci $A_i^\top \cdot A_j = \langle A_i, A_j \rangle$, więc warunek $A^\top A = I$ jest równoważny stwierdzeniu:

$$\langle A_i, A_j \rangle = \begin{cases} 0, & \text{gdy } i \neq j \\ 1, & \text{gdy } i = j \end{cases}$$

czyli (4) $\iff$ (2). Analogicznie (przekształcając warunek (4) do postaci $AA^\top = I$) dowodzimy (4) $\iff$ (3).

Pozostaje udowodnić (1) $\iff$ (4). Jeśli macierz A jest macierzą izometrii, to zachowuje długości wektorów (Fakt 4.41) i miary kątów (Fakt 4.43), czyli jej kolumny (jako obrazy wektorów bazy standardowej) są długości 1 i parami prostopadłe (bo takie warunki spełniają wektory bazy standardowej). To dowodzi (1) $\Rightarrow$ (4). Z drugiej strony, jeśli kolumny macierzy A tworzą bazę ortonormalną, to dla każdego wektora $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ zachodzi:

$$\begin{aligned} \left\langle A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\rangle &= \langle x_1 A_1 + \dots + x_n A_n, x_1 A_1 + \dots + x_n A_n \rangle \\ &= \sum_{1 \leq i, j \leq n} x_i x_j \langle A_i, A_j \rangle = \sum_{1 \leq i \leq n} x_i^2 = \left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\rangle \end{aligned}$$

czyli A zachowuje długości wektorów, a zatem jest macierzą izometrii. To dowodzi (4) $\Rightarrow$ (1). $\square$

Przykład 3

Uzupełnij następującą macierz do macierzy ortogonalnej:

$$\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & a \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & b \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & c \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & d \end{pmatrix}$$

Rozwiązanie. Pierwsze trzy kolumny macierzy mają długości 1 oraz są parami ortogonalne. Macierz będzie zatem ortogonalna, jeśli czwarta kolumna będzie miała długość 1 i będzie ortogonalna do każdej z trzech pierwszych kolumn. Wobec tego:

$$\begin{cases} \frac{1}{2}a + \frac{1}{2}b + \frac{1}{2}c + \frac{1}{2}d = 0 \\ -\frac{1}{2}a - \frac{1}{2}b + \frac{1}{2}c + \frac{1}{2}d = 0 \\ \frac{1}{2}a - \frac{1}{2}b - \frac{1}{2}c + \frac{1}{2}d = 0 \\ a^2 + b^2 + c^2 + d^2 = 1 \end{cases}$$

Rozwiązując powyższy układ równań dostajemy następujące rozwiązania:

$$\begin{cases} a = \frac{1}{2} \\ b = -\frac{1}{2} \\ c = \frac{1}{2} \\ d = -\frac{1}{2} \end{cases} \quad \text{lub} \quad \begin{cases} a = -\frac{1}{2} \\ b = \frac{1}{2} \\ c = -\frac{1}{2} \\ d = \frac{1}{2} \end{cases}$$

Fakt 4.46

Dla dowolnej macierzy $A \in M_{n \times n}(\mathbb{R})$ oraz dowolnych wektorów $u, v \in \mathbb{R}^n$ zachodzi:

$$\langle Au, v \rangle = \langle u, A^\top v \rangle$$

gdzie $\langle \cdot, \cdot \rangle$ oznacza standardowy iloczyn skalarny w przestrzeni euklidesowej $\mathbb{R}^n$.

Dowód. Zauważmy, że standardowy iloczyn skalarny można zapisać w postaci:

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = (x_1 \ \cdots \ x_n) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}^\top \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

gdzie mnożenie po prawej stronie oznacza mnożenie macierzy. Wobec tego:

$$\langle Au, v \rangle = (Au)^\top \cdot v = (u^\top \cdot A^\top) \cdot v = u^\top \cdot (A^\top \cdot v) = \langle u, A^\top v \rangle$$

□

Wniosek 4.47

Dla macierzy **symetrycznej** $A \in M_{n \times n}(\mathbb{R})$ oraz wektorów $u, v \in \mathbb{R}^n$ zachodzi:

$$\langle Au, v \rangle = \langle u, Av \rangle$$

gdzie $\langle \cdot, \cdot \rangle$ oznacza standardowy iloczyn skalarny w przestrzeni euklidesowej $\mathbb{R}^n$.

Fakt 4.48

Jeśli λ i μ są **różnymi** wartościami własnymi symetrycznej macierzy $A \in M_{n \times n}(\mathbb{R})$, zaś v i w są wektorami własnymi, odpowiednio dla λ i μ , to $v \perp w$ (względem standardowego iloczynu skalarnego na $\mathbb{R}^n$).

Dowód. Ponieważ v i w są wektorami własnymi, więc $Av = \lambda v$ oraz $Aw = \mu w$. Z Wniosku 4.47 otrzymujemy:

$$\lambda \langle v, w \rangle = \langle \lambda v, w \rangle = \langle Av, w \rangle = \langle v, Aw \rangle = \langle v, \mu w \rangle = \mu \langle v, w \rangle$$

Ponieważ $\lambda \langle v, w \rangle = \mu \langle v, w \rangle$ oraz $\lambda \neq \mu$, więc $\langle v, w \rangle = 0$, czyli $v \perp w$. $\square$

Definicja 4.49

Zespoloną przestrzeń euklidesową nazywamy zespoloną przestrzeń liniową V z dodatkową operacją zwaną *zespolonym iloczynem skalarnym*, przypisującą (uporządkowanej) parze wektorów (u, v) liczbę $\langle u, v \rangle \in \mathbb{C}$ spełniającą dla dowolnych $u, v, w \in V$ oraz dowolnego $t \in \mathbb{C}$ następujące warunki:

- 1) $\langle u, v \rangle = \langle v, u \rangle$ (symetryczność)
- 2a) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$
- 2b) $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$
- 2c) $\langle tu, v \rangle = t \langle u, v \rangle$
- 2d) $\langle u, tv \rangle = \bar{t} \langle u, v \rangle$ (półtoraliniowość)
- 3) $\langle v, v \rangle \geq 0$, przy czym $\langle v, v \rangle = 0 \iff v = 0$ (dodania określoność)

Zwróćmy uwagę na sprzężenie we własności (2d):

$$\langle u, tv \rangle = \bar{t} \langle u, v \rangle \quad (4.3)$$

które odróżnia definicję zespolonego iloczynu skalarnego od rzeczywistego iloczynu skalarnego. W związku z tą modyfikacją zespolony iloczyn skalarny nie jest dwuliniowy (bo nie jest jednorodny względem drugiej współrzędnej). Ponieważ z czterech warunków dwuliniowości spełnia on trzy warunki, mówimy czasem, że jest on *półtoraliniowy*. Powodem modyfikacji (4.3) jest wymóg (3), który ma zapewniać, że długość wektora w zespolonej przestrzeni euklidesowej była (nieujemną) liczbą rzeczywistą (podobnie jak moduł liczby zespolonej). Gdyby warunek (2d) był sformułowany tak jak dla rzeczywistej przestrzeni euklidesowej, mielibyśmy dla dowolnego $z \in \mathbb{C}$:

$$|zv|^2 = \langle zv, zv \rangle = z^2 \langle v, v \rangle = z^2 \cdot |v|^2$$

co stałoby w sprzeczności z założeniem, że zarówno $|v|$, jak i $|zv|$ są dodatnimi liczbami rzeczywistymi. Dzięki sprzężeniu w warunku (2d) otrzymujemy:

$$|zv|^2 = \langle zv, zv \rangle = z \cdot \bar{z} \langle v, v \rangle = |z|^2 \cdot |v|^2$$

co pozostaje w zgodzie z założeniem, że $|v|$ i $|zv|$ są dodatnimi liczbami rzeczywistymi.

Przykład 4

Sprawdź, że zdefiniowana na przestrzeni $\mathbb{C}^n$ funkcja:

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{i=1}^n x_i \bar{y}_i$$

jest zespolonym iloczynem skalarnym.

Rozwiązanie. Symetryczność oraz liniowość względem pierwszej współrzędnej jest oczywista. Sprawdźmy pozostałe warunki:

$$\begin{aligned}\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1+y'_1 \\ \vdots \\ y_n+y'_n \end{pmatrix} \right\rangle &= \sum_{i=1}^n x_i \cdot \overline{y_i + y'_i} = \sum_{i=1}^n x_i \overline{y_i} + \sum_{i=1}^n x_i \overline{y'_i} = \left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle + \left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y'_1 \\ \vdots \\ y'_n \end{pmatrix} \right\rangle \\ \left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} ty_1 \\ \vdots \\ ty_n \end{pmatrix} \right\rangle &= \sum_{i=1}^n x_i \overline{ty_i} = \bar{t} \cdot \sum_{i=1}^n x_i \overline{y_i} = \bar{t} \cdot \left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle\end{aligned}$$

Dodatnia określoność wynika z faktu, że $z \cdot \bar{z} = |z|^2$ jest dodatnią liczbą rzeczywistą dla każdej liczby zespolonej $z \neq 0$ oraz jest równa zero dla $z = 0$:

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\rangle = \sum_{i=1}^n x_i \bar{x}_i = \sum_{i=1}^n |x_i|^2$$

Przykład 5

Sprawdź, że zdefiniowana na przestrzeni $\mathbb{C}_n[x]$ funkcja:

$$\langle P, Q \rangle = \int_0^1 P(x) \overline{Q(x)} dx$$

jest zespolonym iloczynem skalarnym.

Rozwiązanie. Sprawdzenie warunków (1) i (2) Definicji 4.49 pozostawiamy czytelnikowi. Sprawdźmy warunek dodatniej określoności (3):

$$\langle P, P \rangle = \int_0^1 P(x) \overline{P(x)} dx = \int_0^1 |P(x)|^2 dx$$

Powyższa całka to całka z nieujemnej funkcji rzeczywistej, jest więc ona zawsze nieujemną liczbą rzeczywistą, przy czym jest równa 0 wtedy i tylko wtedy, gdy $P(x) = 0$ dla każdego $0 \leq x \leq 1$. W przypadku wielomianu jest to równoważne stwierdzeniu, że P jest wielomianem zerowym.

W poprzednich rozdziałach zauważaliśmy, że diagonalizując macierz $A \in M_{n \times n}(\mathbb{R})$, której wielomian ma nierzeczywiste pierwiastki możemy ją potraktować jako macierz $A \in M_{n \times n}(\mathbb{C})$, zamieniając rzeczywistą przestrzeń liniową $\mathbb{R}^n$ na zespoloną przestrzeń liniową $\mathbb{C}^n$. Podobne uogólnienie jest możliwe również w odniesieniu do przestrzeni euklidesowych. Aby się o tym przekonać wystarczy sprawdzić, że rzeczywisty iloczyn skalarny wektorów $u, v \in \mathbb{R}^n$ jest taki sam jak zespolony iloczyn skalarny wektorów u i v traktowanych jako elementy zespolonej przestrzeni euklidesowej $\mathbb{C}^n$. W tym celu zauważamy, że definicja zespolonego iloczynu skalarnego (Definicja 4.49) poktywa się z definicją rzeczywistego iloczynu skalarnego, jeśli tylko $t \in \mathbb{R}$ oraz $u, v \in \mathbb{R}^n$. Jest to prawdą, gdyż jedyna różnica dotyczy warunku (2d), w którym $\bar{t} = t$ dla liczb rzeczywistych $t \in \mathbb{R}$. Podejście take będzie istotne przy dowodzeniu poniższego twierdzenia.

Twierdzenie 4.50: Twierdzenie spektralne (wersja macierzowa)

Jeśli $A \in M_{n \times n}(\mathbb{R})$ jest macierzą symetryczną, to

$$A = UDU^{-1} = UDU^{\top}$$

gdzie $D \in M_{n \times n}(\mathbb{R})$ jest macierzą diagonalną, a $U \in M_{n \times n}(\mathbb{R})$ jest macierzą ortogonalną.

Dowód. Niech $\lambda \in \mathbb{C}$ będzie wartością własną macierzy $A \in M_{n \times n}(\mathbb{R})$ (traktowanej jako macierz o wyrazach zespolonych), zaś $v \in \mathbb{C}^n$ wektorem własnym dla λ . Wówczas $Av = \lambda v$, skąd na mocy Wniosku 4.47 (macierz A jest symetryczną macierzą o wyrazach rzeczywistych):

$$\lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle Av, v \rangle = \langle v, Av \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \langle v, v \rangle$$

Ponieważ $\langle v, v \rangle \neq 0$ (wektor v jest niezerowy), więc $\lambda = \bar{\lambda}$, czyli $\lambda \in \mathbb{R}$. Oznacza to, że wszystkie wartości własne macierzy A są rzeczywiste, czyli A diagonalizuje się nad $\mathbb{R}$. Zgodnie z Faktem 4.48 baza złożona z wektorów własnych A jest bazą ortogonalną (w przypadku wielokrotnych wartości własnych wymaga to dodatkowego argumentu, który pomijamy). Odpowiednio skalując wektory z tej bazy otrzymujemy ortonormalną bazę wektorów własnych, co doprowadza A do postaci:

$$A = UDU^{-1}$$

gdzie U jest macierzą ortogonalną (macierzą izometrii). Ponieważ dla macierzy ortogonalnych $U^{-1} = U^{\top}$, więc

$$A = UDU^{\top}$$

□

Twierdzenie 4.51: Twierdzenie spektralne (wersja przekształceniowa)

Jeśli $A \in M_{n \times n}(\mathbb{R})$ jest macierzą symetryczną, to przekształcenie $F_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ diagonalizuje się nad $\mathbb{R}$ w pewnej bazie ortonormalnej, tzn. $m_{\mathcal{B}}^{\mathcal{B}}(F)$ jest macierzą diagonalną dla pewnej bazy ortonormalnej $\mathcal{B}$.

Dowód. Jest to bezpośredni wniosek z Twierdzenia 4.50, gdzie elementami bazy $\mathcal{B}$ są kolumny $U_1, \dots, U_n$ macierzy U . □

Przed sformułowaniem trzeciej wersji twierdzenia spektralnego, musimy przyjrzeć się mnożeniu „pionowych” i „poziomych” wektorów:

Fakt 4.52

Niech $u, v \in \mathbb{R}^n$. Wówczas:

- 1) $u^{\top} \cdot v = \langle u, v \rangle$ jest liczbą rzeczywistą,
- 2) $u \cdot v^{\top}$ jest macierzą rozmiaru $n \times n$, która to macierz ma rząd 1 (o ile $u \neq 0$ i $v \neq 0$).

Dowód. Niech $u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ oraz $v = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. Wówczas:

$$u^{\top} \cdot v = (x_1 \quad \cdots \quad x_n) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \cdots + x_n y_n = \langle u, v \rangle \in \mathbb{R}$$

$$u \cdot v^\top = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot (y_1 \quad \cdots \quad y_n) = \begin{pmatrix} x_1 y_1 & \cdots & x_1 y_n \\ \vdots & \ddots & \vdots \\ x_n y_1 & \cdots & x_n y_n \end{pmatrix}$$

Zwróćmy uwagę, że w tym ostatnim przypadku każda kolumna otrzymanej macierzy jest skalarną krotnością wektora u , co oznacza, że rząd tej macierzy to 1 (przy założeniu $u \neq 0$ i $v \neq 0$). $\square$

Fakt 4.53

Niech $A, B \in M_{n \times n}(\mathbb{R})$ oraz $A = (A_1, \dots, A_n)$ i $B = (B_1, \dots, B_n)$ będą przedstawieniami macierzy w postaci ciągu kolumn (tzn. $A_i, B_i \in \mathbb{R}^n$). Wówczas:

$$A \cdot B^\top = \sum_{i=1}^n A_i \cdot B_i^\top$$

Dowód. Oznaczmy przez $e_1, \dots, e_n \in \mathbb{R}^n$ wektory bazy standardowej $\mathbb{R}^n$. Zauważmy, że $A_i \cdot e_i^\top$ jest macierzą $n \times n$, której i -tą kolumną jest A , natomiast pozostałe kolumny są zerowe. Wobec tego:

$$A = A_1 \cdot e_1^\top + \cdots + A_n \cdot e_n^\top = \sum_{i=1}^n A_i \cdot e_i^\top$$

Analogicznie:

$$B = B_1 \cdot e_1^\top + \cdots + B_n \cdot e_n^\top$$

skąd dostajemy:

$$B^\top = (B_1 \cdot e_1^\top)^\top + \cdots + (B_n \cdot e_n^\top)^\top = e_1 \cdot B_1^\top + \cdots + e_n \cdot B_n^\top = \sum_{j=1}^n e_j \cdot B_j^\top$$

Wobec tego:

$$A \cdot B = \sum_{i=1}^n A_i e_i^\top \cdot \sum_{j=1}^n e_j B_j^\top = \sum_{1 \leq i, j \leq n} A_i \cdot e_i^\top e_j \cdot B_j^\top$$

Z uwagi na to, że $e_i^\top e_j = 0$, gdy $i \neq j$ oraz $e_i^\top e_i = 1$, otrzymujemy:

$$A \cdot B = \sum_{i=1}^n A_i \cdot B_i^\top$$

$\square$

Przykład 6

Niech $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ i $B = \begin{pmatrix} 5 & 1 \\ 2 & 6 \end{pmatrix}$. Zapisz iloczyn $AB^\top$ w takiej postaci jak w powyższym fakcie.

Rozwiązanie.

$$A = (A_1, A_2), \quad \text{gdzie} \quad A_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

$$B = (B_1, B_2), \quad \text{gdzie} \quad B_1 = \begin{pmatrix} 5 \\ 2 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 1 \\ 6 \end{pmatrix}$$

Zgodnie z Faktem 3.13:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 5 & 2 \\ 1 & 6 \end{pmatrix} = AB^\top = A_1 B_1^\top + A_2 B_2^\top = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \cdot (5 \quad 2) + \begin{pmatrix} 2 \\ 4 \end{pmatrix} \cdot (1 \quad 6)$$

Twierdzenie 4.54: Twierdzenie spektralne (trzecia wersja)

Jeśli $A \in M_{n \times n}(\mathbb{R})$ jest macierzą symetryczną, to:

$$A = \lambda_1 u_1 u_1^\top + \cdots + \lambda_n u_n u_n^\top$$

gdzie $\lambda_1, \dots, \lambda_n$ to wartości własne A (z krotnościami), zaś $u_1, \dots, u_n \in \mathbb{R}^n$ to odpowiadające im wektory własne z pewnej ortonormalnej bazy.

Dowód. Zgodnie z pierwszą wersją Twierdzenia spektralnego (Twierdzenie 4.50) mamy:

$$A = UDU^\top = (UD) \cdot U^\top = (\lambda_1 u_1, \dots, \lambda_n u_n) \cdot (u_1, \dots, u_n)^\top$$

dla $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ oraz $U = (u_1, \dots, u_n)$, gdzie $u_1, \dots, u_n \in \mathbb{R}^n$ stanowią bazę ortonormalną.

Wobec tego, zgodnie z Faktem 3.13:

$$A = \lambda_1 u_1 \cdot u_1^\top + \cdots + \lambda_n u_n \cdot u_n^\top$$

□

Twierdzenie 4.55: Diagonalizacja formy kwadratowej w bazie ortonormalnej

Niech V będzie skończenie wymiarową przestrzenią euklidesową, zaś $Q : V \rightarrow \mathbb{R}$ formą kwadratową. Wówczas istnieje taka **ortonormalna** baza $\mathcal{B} = (b_1, \dots, b_n)$, w której forma Q diagonalizuje się, tzn. $m^{\mathcal{B}\mathcal{B}}(Q)$ jest macierzą diagonalną.

Dowód. Niech $\mathcal{C}$ będzie dowolną bazą przestrzeni Q . Wówczas macierz $A = m^{\mathcal{C}\mathcal{C}}(Q)$ jest macierzą symetryczną. Zgodnie z twierdzeniem spektralnym:

$$A = UDU^{-1} \quad \text{czyli} \quad D = U^{-1}AU$$

dla pewnej macierzy ortogonalnej U . Ponieważ U jest ortogonalna, więc $U^\top = U^{-1}$, czyli macierz $U^\top AU$ jest macierzą diagonalną. Niech $\mathcal{B}$ będzie taką bazą, że $m_{\mathcal{C}}^{\mathcal{B}}(\text{id}) = U$ (tzn. kolumny macierzy U wyznaczają współrzędne nowej bazy $\mathcal{B}$ w starej bazie $\mathcal{C}$). Wówczas macierz:

$$U^{-1}AU = U^\top AU = (m_{\mathcal{C}}^{\mathcal{B}}(\text{id}))^\top \cdot m^{\mathcal{C}\mathcal{C}}(Q) \cdot m_{\mathcal{C}}^{\mathcal{B}}(\text{id}) = m^{\mathcal{B}\mathcal{B}}(Q)$$

jest macierzą diagonalną, czyli forma kwadratowa Q diagonalizuje się w bazie ortonormalnej $\mathcal{B}$. □

4.4 Rozkład SVD

Dotychczas zajmowaliśmy się przede wszystkim badaniem macierzy kwadratowych, jako że nasze główne narzędzie – diagonalizacja – stosowało się jedynie do takich macierzy. Niniejszy rozdział pokazuje jak wykorzystać diagonalizację oraz twierdzenie spektralne do analizy macierzy prostokątnych. Pierwszym krokiem w tym kierunku jest zamiana analizowanej macierzy prostokątnej na pewną macierz kwadratową (najlepiej symetryczną), którą umiemy diagonalizować.

Fakt 4.56

Dla dowolnej macierzy $A \in M_{m \times n}(\mathbb{R})$ macierze $AA^\top \in M_{m \times m}(\mathbb{R})$ oraz $A^\top A \in M_{n \times n}(\mathbb{R})$ są macierzami symetrycznymi i wszystkie ich wartości własne są nieujemne. Co więcej, niezerowe wartości własne (uwzględniając krotności) są takie same dla obu tych macierzy.

W kolejnych rozdziałach macierz o takich własnościach (symetryczna o nieujemnych wartościach własnych) będziemy nazywać *macierzą nieujemnie określoną*.

Dowód. Symeryczność macierzy wynika z następującego rachunku:

$$(AA^T)^T = (A^T)^T \cdot A^T = AA^T$$

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

W związku z tym, że macierze AA^T i $A^T A$ są symetryczne, mają one tylko rzeczywiste wartości własne. Niech $\lambda \in \mathbb{R}$ będzie wartością własną AA^T , a odpowiadający jej (niezerowy) wektor własny to v . Wówczas $AA^T v = \lambda v$ i zgodnie z Faktem ?? (przyjmując $u = A^T v$) dostajemy:

$$\lambda|v|^2 = \lambda\langle v, v \rangle = \langle \lambda v, v \rangle = \langle AA^T v, v \rangle = \langle A^T v, A^T v \rangle = |A^T v|^2$$

Stąd

$$\lambda = \frac{|A^T v|^2}{|v|^2} \geq 0$$

Podobnie uzasadniamy analogiczną własność dla macierzy $A^T A$.

Niech teraz v będzie wektorem własnym macierzy AA^T dla wartości własnej $\lambda > 0$. Oznacza to, że:

$$AA^T v = \lambda v$$

Wówczas:

$$(A^T A) \cdot A^T v = A^T (AA^T) v = A^T \cdot \lambda v = \lambda \cdot A^T v$$

czyli wektor $A^T v$ jest wektorem własnym macierzy $A^T A$ dla wartości własnej λ . Wektor $A^T v$ jest wektorem niezerowym, gdyż w przeciwnym razie mielibyśmy $\lambda v = AA^T v = A \cdot 0 = 0$, wbrew założeniu. Wobec tego λ jest wartością własną dla $A^T A$. Podobnie pokazujemy, że każda niezerowa wartość własna macierzy $A^T A$ jest również wartością własną macierzy AA^T . Sprawdzenie, że każda niezerowa wartość własna λ ma taką samą krotność jako wartość własna $A^T A$, co jako wartość własna AA^T , pomijamy. $\square$

Przykład 1

Dla każdej z podanych macierzy A wylicz macierze AA^T oraz $A^T A$, a następnie wyznacz ich wartości własne.

$$\begin{pmatrix} 2 & 1 & 3 \end{pmatrix} \quad \begin{pmatrix} 2 & 2 \\ -1 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & -6 \\ 4 & 0 \\ -5 & 3 \end{pmatrix}$$

Rozwiązanie. (a) Dla $A = \begin{pmatrix} 2 & 1 & 3 \end{pmatrix}$ mamy $AA^T = (14)$ oraz $A^T A = \begin{pmatrix} 4 & 2 & 6 \\ 2 & 1 & 3 \\ 6 & 3 & 9 \end{pmatrix}$. Jedyną wartością własną macierzy 1×1 jest jej jedyny wyraz, czyli $\lambda_1 = 14$. W przypadku drugiej macierzy otrzymujemy $\lambda_1 = 14$, $\lambda_2 = \lambda_3 = 0$.

(b) dla $A = \begin{pmatrix} 2 & 2 \\ -1 & 2 \end{pmatrix}$ mamy $AA^T = \begin{pmatrix} 8 & 2 \\ 2 & 5 \end{pmatrix}$ oraz $A^T A = \begin{pmatrix} 5 & 2 \\ 2 & 8 \end{pmatrix}$. Obie macierze mają wartości własne $\lambda_1 = 9$ i $\lambda_2 = 4$.

(c) dla $A = \begin{pmatrix} 2 & -6 \\ 4 & 0 \\ -5 & 3 \end{pmatrix}$ mamy $AA^T = \begin{pmatrix} 40 & 8 & -28 \\ 8 & 16 & -20 \\ -28 & -20 & -34 \end{pmatrix}$ oraz $A^T A = \begin{pmatrix} 45 & -27 \\ -27 & 45 \end{pmatrix}$.

Pierwsza z macierzy ma wartości własne $\lambda_1 = 72$, $\lambda_2 = 18$ i $\lambda_3 = 0$, zaś druga ma wartości własne $\lambda_1 = 72$, $\lambda_2 = 18$.

Definicja 4.57

Prostokątną macierz diagonalną nazywamy taką macierz $A = (a_{ij}) \in M_{m \times n}(\mathbb{R})$, że $a_{ij} = 0$ zawsze gdy $i \neq j$.

Rozważaną dotychczas (kwadratową) macierz diagonalną będziemy traktować jako szczególny przypadek prostokątnej macierzy diagonalnej. Inne przykłady prostokątnych macierzy diagonalnych to:

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \quad \begin{pmatrix} 5 & 0 \\ 0 & 7 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Niniejszy rozdział poświęcony będzie uogólnieniu Twierdzeń 4.50, 4.51 i 4.54 na macierze prostokątne. Prostokątna macierz diagonalna będzie w tych uogólnieniach pełniła taką rolę, jak macierz diagonalna w przywołanych twierdzeniach.

Fakt 4.58

Niech $\Sigma \in M_{m \times n}(\mathbb{R})$ będzie prostokątną macierzą diagonalną, a niezerowe wyrazy na jej przekątnej to $\lambda_1, \dots, \lambda_k$. Wówczas każda z macierzy $\Sigma\Sigma^\top$ oraz $\Sigma^\top\Sigma$ jest (kwadratową) macierzą diagonalną, a niezerowe wyrazy na jej przekątnej to $\lambda_1^2, \dots, \lambda_k^2$.

Dowód. Wykonując mnożenie otrzymujemy:

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n & 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} \lambda_1^2 & & & & & & \\ & \lambda_2^2 & & & & & \\ & & \ddots & & & & \\ & & & \lambda_n^2 & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}$$

Podobna sytuacja ma miejsce, gdy wymnożymy czynniki w odwrotnej kolejności. $\square$

Poniższe twierdzenie można traktować jako uogólnienie twierdzenia spectralnego (Twierdzenie 4.50) na przypadek macierzy prostokątnych. Rozkład SVD to skrót od *Singular Value Decomposition*, czyli *rozkład według wartości osobliwych*.

Twierdzenie 4.59: Rozkład SVD (wersja macierzowa)

Dla dowolnej macierzy (prostokątnej) $A \in M_{m \times n}(\mathbb{R})$ istnieje rozkład:

$$A = U\Sigma V^\top$$

gdzie Σ to prostokątna macierz diagonalna, zaś $U \in M_{m \times m}(\mathbb{R})$ i $V \in M_{n \times n}(\mathbb{R})$ to (kwadratowe) macierze ortogonalne. Co więcej:

- niezerowe wyrazy na przekątnej macierzy Σ to pierwiastki z niezerowych wartości własnych macierzy $AA^\top$ (i równocześnie $A^\top A$),
- kolumny macierzy U to wektory własne macierzy $AA^\top$,
- kolumny macierzy V to wektory własne macierzy $A^\top A$.

Niezerowe wyrazy na przekątnej macierzy Σ nazywamy *wartościami osobliwymi*.

Dowód. Szczegółowa konstrukcja rozkładu SVD dla zadanej macierzy A (w tym dowód istnienia takiego rozkładu) wykracza poza ramy niniejszego skryptu. Pokażemy jedynie wymienione własności macierzy U , Σ i V , które pozwalają wyznaczyć macierz Σ oraz kandydatów na macierze U i V . Szczegóły wyboru macierzy U i V pominiemy.

Zgodnie z Faktem 4.56 macierze $AA^\top$ i $A^\top A$ są symetryczne, mają jednakowe niezerowe wartości własne i wszystkie ich wartości własne są nieujemne. Niech $A = U\Sigma V^\top$ będzie rozkładem SVD macierzy A . Macierze U i V są macierzami ortogonalnymi, więc $U^\top = U^{-1}$ oraz $V^\top = V^{-1}$, skąd:

$$AA^\top = (U\Sigma V^\top)(U\Sigma V^\top)^\top = U\Sigma(V^\top V)\Sigma^\top U^\top = UDU^\top = UDU^{-1} \quad (4.4)$$

$$A^\top A = (U\Sigma V^\top)^\top(U\Sigma V^\top) = V\Sigma^\top(U^\top U)\Sigma V^\top = VD'V^\top = VD'V^{-1} \quad (4.5)$$

gdzie (zgodnie z Faktem 4.58) macierze $D \in M_{m \times m}(\mathbb{R})$ oraz $D' \in M_{n \times n}(\mathbb{R})$ to (kwadratowe) macierze diagonalne, których niezerowe wyrazy to kwadraty niezerowych wyrazów macierzy Σ . Wobec tego (4.4) i (4.5) to diagonalizacje odpowiednio macierzy $AA^\top$ i $A^\top A$, w szczególności:

- niezerowe wyrazy na przekątnej macierzy Σ to pierwiastki z niezerowych wartości własnych macierzy $AA^\top$ (i równocześnie $A^\top A$),
- kolumny macierzy U to wektory własne macierzy $AA^\top$,
- kolumny macierzy V to wektory własne macierzy $A^\top A$.

□

Przykład 2

Dla każdej z poniższych macierzy wyznacz macierz Σ z rozkładu SVD:

$$\begin{pmatrix} 2 & 1 & 3 \end{pmatrix} \quad \begin{pmatrix} 2 & 2 \\ -1 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & -6 \\ 4 & 0 \\ -5 & 3 \end{pmatrix}$$

Rozwiązanie. Jeśli $A = U\Sigma V^\top$ jest rozkładem SVD, to macierz Σ ma ten sam rozmiar co macierz A , jest prostokątną macierzą diagonalną, a niezerowe wyrazy na jej przekątnej to pierwiastki z wartości własnych macierzy $AA^\top$ (a tym samym macierzy $A^\top A$). Wobec tego, zgodnie z wyliczonymi w Przykładzie 1 wartościami własnymi dostajemy:

$$\begin{aligned} \begin{pmatrix} 2 & 1 & 3 \end{pmatrix} &= U_1 \cdot (\sqrt{14} \ 0 \ 0) \cdot V_1^\top \\ \begin{pmatrix} 2 & 2 \\ -1 & 2 \end{pmatrix} &= U_2 \cdot \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \cdot V_2^\top \\ \begin{pmatrix} 2 & -6 \\ 4 & 0 \\ -5 & 3 \end{pmatrix} &= U_3 \cdot \begin{pmatrix} 6\sqrt{2} & 0 \\ 0 & 3\sqrt{2} \\ 0 & 0 \end{pmatrix} \cdot V_3^\top \end{aligned}$$

Przykład 3

Sprawdź, które z poniższych rozkładów to rozkłady SVD zadanych macierzy:

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 3 & -2 & 1 \\ 3 & 1 & -2 \\ 3 & 3 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 \\ 0 & 5 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}^\top \quad B = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix}^\top$$

Rozwiązanie. W każdym z podanych rozkładów mnożenie macierzy jest wykonalne. Ponadto

środkowa macierzy (potencjalna macierz Σ) jest prostokątną macierzą diagonalną o nieujemnych wyrazach, a skrajne macierze (potencjalne U i $V^\top$) są macierzami kwadratowymi. Pozostaje do sprawdzenia, czy skrajne macierze są macierzami ortogonalnymi, tzn. mają kolumny/wiersze parami prostopadłe i długości 1. Jest to prawdą w odniesieniu do rozkładu macierzy A (zatem jest to rozkład SVD), natomiast nie jest to prawdą w odniesieniu do rozkładu macierzy B (zatem nie jest to rozkład SVD).

Poniższe twierdzenie jest drugą wersją Twierdzenia 4.59, uogólniającą Twierdzenie 4.51.

Twierdzenie 4.60: Rozkład SVD (wersja przekształceniowa)

Dla dowolnego przekształcenia liniowego $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$ istnieją bazy ortonormalne $\mathcal{B}$ przestrzeni $\mathbb{R}^m$ oraz $\mathcal{C}$ przestrzeni $\mathbb{R}^n$ takie, że $m_{\mathcal{B}}^{\mathcal{C}}(F)$ jest prostokątną macierzą diagonalną.

Dowód. Niech $A = m_{\mathcal{E}'}^{\mathcal{E}}(F)$ będzie macierzą przekształcenia F w bazach standardowych $\mathcal{E}$ przestrzeni $\mathbb{R}^m$ i $\mathcal{E}'$ przestrzeni $\mathbb{R}^n$. Zgodnie z Twierdzeniem 4.59 $A = U\Sigma V^\top$. Niech $\mathcal{B}$ będzie bazą $\mathbb{R}^m$ złożoną z kolumn macierzy U , zaś $\mathcal{C}$ – bazą $\mathbb{R}^n$ złożoną z kolumn macierzy V . Wówczas:

$$m_{\mathcal{C}'}^{\mathcal{C}}(\text{id}) = U \quad \text{oraz} \quad m_{\mathcal{E}}^{\mathcal{B}}(\text{id}) = V$$

Zgodnie z Faktem ??:

$$m_{\mathcal{C}}^{\mathcal{B}}(F) = m_{\mathcal{C}'}^{\mathcal{C}}(\text{id}) m_{\mathcal{E}'}^{\mathcal{E}}(F) m_{\mathcal{E}}^{\mathcal{B}}(\text{id}) = U^{-1} \cdot U \Sigma V^\top \cdot V = \Sigma \cdot V^{-1} V = \Sigma$$

co oznacza, że $m_{\mathcal{C}}^{\mathcal{B}}(F)$ jest prostokątną macierzą diagonalną. $\square$

Fakty 4.52(2) oraz 4.53 można uogólnić na macierze prostokątne. Dowody tych uogólnień jest analogiczny do dowodów tamtych faktów, dlatego je pominiemy:

Fakt 4.61

Niech $u \in \mathbb{R}^m$ oraz $v \in \mathbb{R}^n$. Wówczas $u \cdot v^\top$ jest macierzą rozmiaru $m \times n$, która to macierz ma rząd 1 (o ile $u \neq 0$ i $v \neq 0$).

Fakt 4.62

Niech $A \in M_{m \times k}$ oraz $B \in M_{n \times k}(\mathbb{R})$ oraz $A = (A_1, \dots, A_k)$ i $B = (B_1, \dots, B_k)$ będą przedstawieniami macierzy w postaci ciągu kolumn (tzn. $A_i \in \mathbb{R}^m$ oraz $B_i \in \mathbb{R}^n$). Wówczas:

$$A \cdot B^\top = \sum_{i=1}^n A_i \cdot B_i^\top$$

Twierdzenie 4.63: Rozkład SVD (trzecia wersja)

Dana jest macierz $A \in M_{m \times n}(\mathbb{R})$. Niech $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > 0$ będą wszystkimi niezerowymi wartościami własnymi $AA^\top$ (i równocześnie $A^\top A$). Wówczas:

$$A = \sqrt{\lambda_1} u_1 v_1^\top + \dots + \sqrt{\lambda_k} u_k v_k^\top$$

gdzie $u_1, \dots, u_k$ oraz $v_1, \dots, v_k$ są parami prostopadłymi wektorami własnymi długości 1 macierzy $AA^\top$ i $A^\top A$ odpowiadającymi, odpowiednio, wartościom własnym $\lambda_1, \dots, \lambda_k$. Wartości $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_k}$ to (uporządkowane malejąco) wartości osobliwe macierzy A .

Dowód. Niech $A = U\Sigma V^\top$ będzie rozkładem SVD macierzy A . Oznaczmy przez $u_1, \dots, u_m$ kolumny macierzy U (tzn. $U = (u_1, \dots, u_m)$), zaś przez $v_1, \dots, v_n$ – kolumny macierzy V (tzn. $V = (v_1, \dots, v_n)$). Wówczas:

$$U\Sigma = (\sqrt{\lambda_1}u_1, \dots, \sqrt{\lambda_k}u_k, 0, \dots, 0)$$

a zgodnie z Faktem 4.62 dostajemy:

$$\begin{aligned} A &= U\Sigma \cdot V^\top = (\sqrt{\lambda_1}u_1, \dots, \sqrt{\lambda_k}u_k, 0, \dots, 0) \cdot (v_1, \dots, v_k, \dots, v_n) \\ &= \sqrt{\lambda_1}u_1v_1^\top + \dots + \sqrt{\lambda_k}u_kv_k^\top + 0 + \dots + 0 \end{aligned}$$

□

Fakt 4.61 można w następujący sposób uogólnić:

Fakt 4.64

Jeśli $u_1, \dots, u_k \in \mathbb{R}^m$ oraz $v_1, \dots, v_k \in \mathbb{R}^n$ są dowolnymi wektorami, to:

$$P = u_1v_1^\top + \dots + u_kv_k^\top \in M_{m \times n}(\mathbb{R})$$

jest macierzą rzędu co najwyżej k .

Dowód. Każda kolumna macierzy $u_iv_i^\top$ jest krotnością wektora u_i . Wobec tego jeśli:

$$P = u_1v_1^\top + \dots + u_kv_k^\top$$

to każda kolumna P jest kombinacją liniową wektorów $u_1, \dots, u_k$. Zatem każda kolumna macierzy P jest elementem k -wymiarowej przestrzeni $\text{Lin}\{u_1, \dots, u_k\}$, czyli $\text{rank } P \leq k$. □

Powyższy fakt częściowo tłumaczy następujące twierdzenie (które pozostawimy bez dowodu):

Twierdzenie 4.65: Eckarta–Younga o aproksymacji macierza niskiego rzędu

Dana jest macierz $A \in M_{m \times n}(\mathbb{R})$. Niech $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > 0$ będą wszystkimi niezerowymi wartościami własnymi $AA^\top$ (i równocześnie $A^\top A$). Wówczas najlepszym przybliżeniem macierzy A macierzą rzędu k jest:

$$A \approx \sqrt{\lambda_1}u_1v_1^\top + \dots + \sqrt{\lambda_r}u_rv_r^\top$$

gdzie $u_1, \dots, u_r$ oraz $v_1, \dots, v_r$ są parami prostopadłymi wektorami własnymi długości 1 macierzy $AA^\top$ i $A^\top A$ odpowiadającymi, odpowiednio, wartościom własnym $\lambda_1, \dots, \lambda_r$. Wartości $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_r}$ to (uporządkowane malejąco) największe spośród wartości osobliwych macierzy A .

Przykład 4

Dany jest rozkład SVD macierzy $A = \begin{pmatrix} 10 & 8 & 5 \\ -6 & -12 & -3 \\ -2 & -4 & -13 \\ -2 & 8 & 11 \end{pmatrix}$:

$$A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} 24 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 6 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{pmatrix}^\top$$

Znajdź najlepsze przybliżenie A macierzą rzędu 2.

Rozwiązanie. Powyższy rozkład SVD można zapisać w postaci:

$$A = 24 \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{pmatrix}^\top + 12 \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{pmatrix}^\top + 6 \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}^\top$$

Zgodnie z Faktem ?? otrzymujemy przybliżenie macierzą rzędu 2:

$$A \approx 24 \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{pmatrix}^\top + 12 \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{pmatrix}^\top = \begin{pmatrix} 8 & 10 & 4 \\ -8 & -10 & -4 \\ 0 & -6 & -12 \\ 0 & 6 & 12 \end{pmatrix}$$

Zauważmy, że żaden wyraz otrzymanego przybliżenia nie różni się o więcej niż 2 od oryginalnej macierzy A .

Drugim zastosowaniem rozkładu SVD jest następujący problem: szukamy takiego wektora, który jest najbardziej wydłużany przez przekształcenie $F_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ zadane macierzą $A \in M_{m \times n}(\mathbb{R})$, tzn. szukamy niezerowego wektora $v \in \mathbb{R}^n$, dla którego $\frac{|Av|}{|v|}$ jest największe. Zauważmy, że współczynnik ten jest jednakowy dla współliniowych wektorów, gdyż:

$$\frac{|A(tv)|}{|tv|} = \frac{|t \cdot Av|}{|t \cdot v|} = \frac{|t| \cdot |Av|}{|t| \cdot |v|} = \frac{|Av|}{|v|}$$

wobec czego tak naprawdę szukamy kierunku, w którym wektory wydłużane są najbardziej. Odpowiedź na to pytanie daje następujący fakt:

Fakt 4.66

Dana jest macierz $A \in M_{m \times n}(\mathbb{R})$. Jeśli $A = U\Sigma V^\top$ jest rozkładem SVD macierzy A , to iloraz $\frac{|Av|}{|v|}$ osiąga maksimum dla wektorów współliniowych z wektorem $v_1 \in \mathbb{R}^n$ (tzn. wektory równoległe do v_1 to wektory najbardziej wydłużane przez przekształcenie $F_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$). Co więcej, owo maksimum jest równe największej spośród wartości osobliwych macierzy A .

Przykład 5

Znajdź wektor najbardziej wydłużany przez przekształcenie $F_A : \mathbb{R}^3 \rightarrow \mathbb{R}^4$, gdzie macierz A jest macierzą z Przykładu 4.

Rozwiązanie. Wykorzystując rozkład SVD z Przykładu 4 otrzymujemy, że szukany wektorem jest każdy wektor równoległy do $v_1 = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{pmatrix}$ (tzn. do wektora macierzy V odpowiadający największej z wartości osobliwych), czyli tym samym każdy wektor równoległy do $\begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix}$. Zauważmy, że zgodnie z Faktem 4.66:

$$|A \cdot \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix}| = \left| \begin{pmatrix} 36 \\ -36 \\ -36 \\ 36 \end{pmatrix} \right| = 72 = 24 \cdot \left| \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix} \right|$$

czyli znaleziony wektor jest wydłużany przez przekształcenie F_A dokładnie 24 razy (gdzie 24 to największa z wartości osobliwych macierzy A).

Jako ostatnie zastosowanie SVD w niniejszym skrypcie przedstawimy przybliżone rozwiązywanie liniowych układów równań. Rozważmy następujący przykład: odbiornik GPS próbuje

wyznaczyć swoje położenie w oparciu o sygnał z satelitów. W pewnym uproszczeniu działa on w ten sposób, że znając położenie $\begin{pmatrix} a_i \\ b_i \\ c_i \end{pmatrix} \in \mathbb{R}^3$ i -tego satelity ($i = 1, \dots, n$) oraz wyznaczając (w oparciu o sygnał z satelity) odległość d_i od i -tego satelity, wyznacza swoje położenie $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ rozwiązując następujący układ równań:

$$\begin{cases} (x - a_1)^2 + (y - b_1)^2 + (z - c_1)^2 = d_1^2 \\ (x - a_2)^2 + (y - b_2)^2 + (z - c_2)^2 = d_2^2 \\ \dots \\ (x - a_n)^2 + (y - b_n)^2 + (z - c_n)^2 = d_n^2 \end{cases}$$

Jest to układ równań kwadratowych, ale odejmując stronami pierwsze równanie od każdego z pozostałych otrzymujemy układ:

$$\begin{cases} (x - a_1)^2 + (y - b_1)^2 + (z - c_1)^2 = d_1^2 \\ 2(a_1 - a_2)x + 2(b_1 - b_2)y + 2(c_1 - c_2)z = (d_2^2 - a_2^2 - b_2^2 - c_2^2) - (d_1^2 - a_1^2 - b_1^2 - c_1^2) \\ \dots \\ 2(a_1 - a_n)x + 2(b_1 - b_n)y + 2(c_1 - c_n)z = (d_n^2 - a_n^2 - b_n^2 - c_n^2) - (d_1^2 - a_1^2 - b_1^2 - c_1^2) \end{cases}$$

Pomińmy chwilowo pierwsze równanie i skupmy się na pozostałym układzie $(n - 1)$ równań z 3 niewiadomymi. Zauważmy, że o ile dla $n - 1 = 3$ nasz układ (najprawdopodobniej) ma jednoznaczne rozwiązanie, natomiast dla $n - 1 > 3$ z dużym prawdopodobieństwem jest to układ sprzeczny (problemem są niedokładności pomiarów). Moglibyśmy pominąć „nadmiarowe” równania, ale oznaczałoby to pozbycie się części zebranych informacji (podczas gdy oczekujemy, że położenie wyznaczone w oparciu o sygnał z większej liczby satelitów będzie dokładniejsze). Problem sprowadza się do rozwiązania układu równań liniowych:

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{cases} \quad \text{czyli} \quad AX = b$$

w sytuacji, gdy $m > n$ (równań jest więcej niż niewiadomych), a równania są prawdziwe „z pewną dokładnością” (co jest typową sytuacją w praktyce). Wobec „przybliżonej” prawdziwości równań nie szukamy zmiennych $X = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}$ spełniających badany układ (z uwagi na dużą liczbę równań układ najpewniej jest sprzeczny), tylko zmiennych „najlepiej pasujących” do tego układu. Formalizując: zamiast szukać X spełniającego układ:

$$AX - b = 0$$

szukamy X , dla którego długość $|AX - b|$ jest najmniejsza.

Fakt 4.67

Dana jest macierz $A \in M_{m \times n}(\mathbb{R})$ oraz wektor $b \in \mathbb{R}^m$. Niech

$$A = U\Sigma V^\top, \quad \text{gdzie} \quad \Sigma = \begin{pmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_k & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}$$

Wówczas wektor $X \in \mathbb{R}^n$, dla którego długość $|AX - b|$ jest najmniejsza, to wektor postaci

$$X = V\Sigma^+U^\top \cdot b, \quad \text{gdzie} \quad \Sigma^+ = \begin{pmatrix} \sigma_1^{-1} & & & & \\ & \ddots & & & \\ & & \sigma_k^{-1} & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}$$

Macier $A^+ = V\Sigma^+U^\top$ nazywamy *pseudo-odwrotnością* macierzy A .

Macierz Σ^+ nazywamy *pseudo-odwrotnością* macierzy Σ . Zauważmy, że jeśli Σ jest odwracalną macierzą kwadratową (tzn. macierzą kwadratową o niezerowych wyrazach na przekątnej), to $\Sigma^+ = \Sigma^{-1}$. Zauważmy również, że w przypadku, gdy Σ jest odwracalna, rozwiązanie podane w powyższym fakcie to:

$$X = V\Sigma^{-1}U^\top \cdot b = V\Sigma^{-1}U^{-1} \cdot b = (U\Sigma V^{-1})^{-1} \cdot b = (U\Sigma V^\top)^{-1} \cdot b = A^{-1}b$$

co jest rozwiązaniem równania $AX = b$.

Przykład 6

Traktując układ równań:

$$\begin{cases} 10x + 8y + 5z = 9 \\ -6x - 12y - 3z = 0 \\ -2x - 4y - 13z = 0 \\ -2x + 8y + 11z = -6 \end{cases}$$

jako przybliżony układ równań, znajdź „najlepiej pasujące” rozwiązanie $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

Rozwiązanie. Zauważmy, że układ jest sprzeczny. Układ złożony z pierwszych trzech równań ma jedno rozwiązanie: $x = 1.5$, $y = -0.75$, $z = 0$, ale rozwiązanie to nie spełnia czwartego równania.

Aby znaleźć „najlepiej pasujące” rozwiązanie, wyznaczamy rozkład SVD macierzy A :

$$A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} 24 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 6 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{pmatrix}^\top$$

oraz pseudo-odwrotność macierzy A :

$$A^+ = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{24} & 0 & 0 & 0 \\ 0 & \frac{1}{12} & 0 & 0 \\ 0 & 0 & \frac{1}{6} & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}^\top = \frac{1}{144} \begin{pmatrix} 13 & 3 & -5 & -11 \\ -4 & -12 & 8 & 8 \\ 2 & 6 & -10 & 2 \end{pmatrix}$$

Stąd otrzymujemy:

$$X = A^+b = A^+ \cdot \begin{pmatrix} 9 \\ 0 \\ 0 \\ 18 \end{pmatrix} \approx \begin{pmatrix} 1.271 \\ -0.583 \\ 0.042 \end{pmatrix}$$

Zauważmy, że dla $x = 1.271$, $y = -0.583$, $z = 0.042$ otrzymujemy:

$$\begin{cases} 10x + 8y + 5z = 8.25 \\ -6x - 12y - 3z = -0.75 \\ -2x - 4y - 13z = -0.75 \\ -2x + 8y + 11z = -6.75 \end{cases}$$