OPERATORY LINIOWE

Zwykle przedmiotem zainteresowania nie są same przestrzenie liniowe, ale przekształcenia liniowe pomiędzy nimi. Przykłady takich przekształceń to obroty, symetrie i jednokładności w $\mathbb{R}^3_{\mathrm{fiz}}$, a także operacje różniczkowania i całkowania w analizie. Na początku skupimy się na najbardziej ogólnych własnościach operatorów liniowych.

§ 1. PRZEKSZTAŁCENIA LINIOWE PRZESTRZENI LINIOWYCH

1. Język przekształceń liniowych. Jak wiemy z części I (rozdz. 2, § 3), każdej macierzy rzeczywistej A wymiarów $m \times n$ odpowiada przekształcenie liniowe $\varphi_A : \mathbb{R}^n \to \mathbb{R}^m$. Aksjomatyzując jego własności, wprowadzamy następującą definicję.

DEFINICJA 1. Niech V i W będą przestrzeniami liniowymi nad tym samym ciałem $\mathfrak{K}$. Odwzorowanie $F : V \to W$ nazywamy *przekształceniem liniowym*, jeśli

$$f(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}), \qquad f(\lambda \mathbf{x}) = \lambda f(\mathbf{x})$$

dla dowolnych $\mathbf{x}, \mathbf{y} \in V$ i $\lambda \in \mathfrak{K}$. Innymi słowy, $f(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y})$ dla $\mathbf{x}, \mathbf{y} \in V$ i $\alpha, \beta \in \mathfrak{K}$.

Szczególnym przypadkiem przekształceń liniowych są formy (funkcjonały) liniowe $f : V \to \mathfrak{K}$, szczegółowo omówione w rozdziale 1.

Zbiór wszystkich przekształceń liniowych $f : V \to W$, oznaczany przez $\mathcal{L}(V, W)$ lub $\mathrm{Hom}(V, W)$, stanowi przestrzeń liniową z naturalnymi działaniami: jeśli $f, g \in \mathcal{L}(V, W)$ i $\nu, \mu \in \mathfrak{K}$, to definiujemy

$$(\nu f + \mu g)(\mathbf{x}) = \nu f(\mathbf{x}) + \mu g(\mathbf{x}) \quad \text{dla } \mathbf{x} \in V.$$

Łatwo sprawdzić, że $\mathcal{L}(V, W)$ istotnie spełnia wszystkie aksjomaty przestrzeni liniowej (rozdz. 1, § 1).

Z każdym przekształceniem liniowym $f : V \to W$ wiążemy dwie podprzestrzenie liniowe: *jądro*

$$\operatorname{Ker} f = \{\mathbf{v} \in V \mid f(\mathbf{v}) = \mathbf{0}\}$$

i *obraz*

$$\operatorname{Im} f = \{\mathbf{w} \in W \mid \mathbf{w} = f(\mathbf{v}) \text{ dla pewnego } \mathbf{v} \in V\}.$$

Nie są to dla nas pojęcia nowe — teraz warto jedynie podkreślić, że są to rzeczywiście podprzestrzenie liniowe w (odpowiednio) V i W (dla $\operatorname{Ker} f$ jest to oczywiste; łatwe sprawdzenie dla $\operatorname{Im} f$ przeprowadzimy poniżej). Dla każdej podprzestrzeni $U \subset V$ będziemy pisać $f(U) = \{f(\mathbf{u}) \mid \mathbf{u} \in U\}$. Jeśli $\mathbf{u}_1, \mathbf{u}_2 \in U$ i $\nu_1, \nu_2 \in \mathfrak{K}$, to

$$\nu_1 f(\mathbf{u}_1) + \nu_2 f(\mathbf{u}_2) = f(\nu_1 \mathbf{u}_1 + \nu_2 \mathbf{u}_2) \in f(U),$$

gdyż $\nu_1 \mathbf{u}_1 + \nu_2 \mathbf{u}_2 \in U$. Zatem $f(U)$ jest podprzestrzenią liniową w W. W szczególności dotyczy to $\operatorname{Im} f = f(V)$.

Zauważmy, że injektywność przekształcenia f jest równoważna równości $\operatorname{Ker} f = \{\mathbf{0}\}$. Istotnie, jeśli $\mathbf{x} \neq \mathbf{y}$ i $f(\mathbf{x}) = f(\mathbf{y})$, to $\mathbf{0} \neq \mathbf{x} - \mathbf{y} \in \operatorname{Ker} f$. Na odwrót, jeśli $\mathbf{0} \neq \mathbf{x} \in \operatorname{Ker} f$, to $f(\mathbf{x}) = \mathbf{0} = f(\mathbf{0})$.

Twierdzenie 1. *Niech $f : V \to W$ będzie przekształceniem liniowym. Jeśli $U = \langle \mathbf{e}_1, \ldots, \mathbf{e}_s \rangle \subset V$, to $f(U) = \langle f(\mathbf{e}_1), \ldots, f(\mathbf{e}_s) \rangle \subset W$. W szczególności*

$$\dim f(U) \leqslant \dim U.$$

Dowód. Na mocy założenia każdy wektor $\mathbf{u} \in U$ można zapisać w postaci $\mathbf{u} = \alpha_1 \mathbf{e}_1 + \ldots + \alpha_s \mathbf{e}_s$, a stąd $f(\mathbf{u}) = \alpha_1 f(\mathbf{e}_1) + \ldots + \alpha_s f(\mathbf{e}_s)$, co właśnie oznacza, że $f(U) = \langle f(\mathbf{e}_1), \ldots, f(\mathbf{e}_s) \rangle$. Zauważmy, że jeśli układ $(\mathbf{e}_1, \ldots, \mathbf{e}_s)$ jest bazą w U (czyli $\dim U = s$), to $(f(\mathbf{e}_1), \ldots, f(\mathbf{e}_s))$ może już nie być bazą w $f(U)$, a więc mamy jedynie nierówność $\dim f(U) \leqslant s = \dim U$. Może się np. zdarzyć, że $U \subset \operatorname{Ker} f$, i wtedy $f(U) = \{\mathbf{0}\}$. ■

2. Macierz przekształcenia liniowego. Niech $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ i $(\mathbf{w}_1, \ldots, \mathbf{w}_m)$ będą bazami przestrzeni odpowiednio V i W. Każdy wektor z $\operatorname{Im} f \subset W$ jest kombinacją liniową wektorów $\mathbf{w}_i$, w szczególności

$$\begin{aligned} f(\mathbf{v}_1) &= a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \ldots + a_{m1}\mathbf{w}_m, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f(\mathbf{v}_n) &= a_{1n}\mathbf{w}_1 + a_{2n}\mathbf{w}_2 + \ldots + a_{mn}\mathbf{w}_m. \end{aligned} \tag{1}$$

Na odwrót, podanie wektorów $\mathbf{w}'_1 = f(\mathbf{v}_1), \ldots, \mathbf{w}'_n = f(\mathbf{v}_n)$ przestrzeni W całkowicie wyznacza przekształcenie liniowe f: mianowicie obrazem wektora $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n$ jest wektor $\mathbf{w} = \alpha_1 \mathbf{w}'_1 + \ldots + \alpha_n \mathbf{w}'_n$. Istotnie, łatwo sprawdzić, że tak określone odwzorowanie jest liniowe: jeśli $\mathbf{v}' = \alpha'_1 \mathbf{v}_1 + \ldots + \alpha'_n \mathbf{v}_n$ oraz $\nu, \nu' \in \mathfrak{K}$, to

$$\begin{aligned} f(\nu \mathbf{v} + \nu' \mathbf{v}') &= f((\nu\alpha_1 + \nu'\alpha_1')\mathbf{v}_1 + \ldots + (\nu\alpha_n + \nu'\alpha_n')\mathbf{v}_n) \\ &= (\nu\alpha_1 + \nu'\alpha_1')\mathbf{w}_1' + \ldots + (\nu\alpha_n + \nu'\alpha_n')\mathbf{w}_n' \\ &= \nu(\alpha_1\mathbf{w}_1' + \ldots + \alpha_n\mathbf{w}_n') + \nu'(\alpha_1'\mathbf{w}_1' + \ldots + \alpha_n'\mathbf{w}_n') \\ &= \nu f(\mathbf{v}) + \nu' f(\mathbf{v}'). \end{aligned}$$

Macierz

$$M_f = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{Vmatrix} \tag{2}$$

nosi nazwę *macierzy przekształcenia liniowego* $f : V \to W$ w bazach $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ i $(\mathbf{w}_1, \ldots, \mathbf{w}_m)$ (lub krótko $(\mathbf{v}_j)$ i $(\mathbf{w}_i)$). Jeśli ustalimy bazy w V i W, to różnym macierzom odpowiadają różne przekształcenia liniowe.

Zauważmy, że współrzędne wektora $f(\mathbf{v}_j)$ w bazie $(\mathbf{w}_i)$ tworzą j-tą kolumnę macierzy M_f. Stąd $\operatorname{rank}(f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n)) = \operatorname{rank} M_f$, a ponieważ $\operatorname{rank}(f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n)) = \dim\langle f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n)\rangle = \dim \operatorname{Im} f$, więc $\dim \operatorname{Im} f = \operatorname{rank} M_f$.

DEFINICJA 2. Wymiar przestrzeni liniowej $\operatorname{Im} f$ nazywamy *rzędem* przekształcenia liniowego f (oznaczenie: $\operatorname{rank} f$).

Wymiar ten nie zależy oczywiście od wyboru jakiejkolwiek bazy. Udowodniliśmy zatem

TWIERDZENIE 2.

(i) *Niech $V = \langle \mathbf{v}_1, \ldots, \mathbf{v}_n\rangle$ i $W = \langle \mathbf{w}_1, \ldots, \mathbf{w}_m\rangle$ będą przestrzeniami liniowymi z ustalonymi bazami. Wtedy istnieje odpowiedniość wzajemnie jednoznaczna między przekształceniami liniowymi z V do W a macierzami $m \times n$ o wyrazach z ciała $\mathfrak{K}$.*

(ii) *Dla każdego układu wektorów $\mathbf{w}_1', \ldots, \mathbf{w}_n' \in W$ istnieje dokładnie jedno przekształcenie liniowe $f : V \to W$ takie, że $f(\mathbf{v}_i) = \mathbf{w}_i'$ dla $i = 1, \ldots, n$.*

(iii) *Rząd przekształcenia liniowego $f : V \to W$ jest równy rzędowi odpowiadającej mu macierzy M_f (przy dowolnym wyborze baz w V i W).* ■

Odpowiedniość pomiędzy macierzami i przekształceniami liniowymi można wykorzystać do innego dowodu znanego twierdzenia o rzędzie iloczynu macierzy (część I, rozdz. 2, § 3). W języku przekształceń liniowych odpowiada mu

TWIERDZENIE 3. *Niech $f \circ g$ będzie złożeniem przekształceń liniowych*

$$U \xrightarrow{g} V \xrightarrow{f} W.$$

Wtedy

(i) $\dim \operatorname{Im}(f \circ g) \leqslant \dim \operatorname{Im} f$;

(ii) $\dim \operatorname{Im}(f \circ g) \leqslant \dim \operatorname{Im} g$.

Dowód. Nierówność (i) jest oczywista, gdyż $\operatorname{Im}(f \circ g) \subset \operatorname{Im} f$. Aby udowodnić nierówność (ii), zauważmy, że $\operatorname{Im}(f \circ g) = f(\operatorname{Im} g)$. Wystarczy teraz skorzystać z twierdzenia 1. ∎

Niech $\mathbf{x} = \sum_{j=1}^{n} x_j \mathbf{v}_j$ będzie wektorem z V, a $\mathbf{y} = f(\mathbf{x}) = \sum_{i=1}^{m} y_i \mathbf{w}_i$ jego obrazem przy przekształceniu $f : V \to W$ o macierzy M_f postaci (2) względem baz $(\mathbf{v}_j)$ i $(\mathbf{w}_i)$. Wtedy zgodnie z (1) mamy

$$f(\mathbf{x}) = \sum_{j=1}^{n} x_j f(\mathbf{v}_j) = \sum_{j=1}^{n} x_j \Big(\sum_{i=1}^{m} a_{ij} \mathbf{w}_i \Big) = \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} a_{ij} x_j \Big) \mathbf{w}_i = \sum_{i=1}^{m} y_i \mathbf{w}_i.$$

Stąd $y_i = \sum_{j=1}^{n} a_{ij} x_j$ dla $i = 1, \ldots, m$ lub krócej

$$Y = M_f \cdot X, \tag{3}$$

gdzie $X = [x_1, \ldots, x_n]$ i $Y = [y_1, \ldots, y_m]$ są kolumnami współrzędnych wektorów $\mathbf{x} \in V$ i $\mathbf{y} = f(\mathbf{x}) \in W$. Na wzór (3) możemy patrzeć jak na definicję przekształcenia przestrzeni kartezjańskich $\mathfrak{K}^n \to \mathfrak{K}^m$, odpowiadającego przekształceniu f.

Niech teraz f i g będą dwoma przekształceniami liniowymi z V do W. Ustalając w tych przestrzeniach bazy $(\mathbf{v}_j)$ i $(\mathbf{w}_i)$, przechodzimy de facto do przekształceń $f : X \mapsto M_f X$ i $g : X \mapsto M_g X$ z $\mathfrak{K}^n$ do $\mathfrak{K}^m$. Zgodnie z wcześniejszymi ustaleniami (część I, rozdz. 2, § 3), przekształceniu liniowemu $\nu f + \mu g$ (zob. p. 1) odpowiada macierz

$$M_{\nu f + \mu g} = \nu M_f + \mu M_g.$$

Analogicznie, przy ustalonych bazach w przestrzeniach U, V i W, złożeniu przekształceń $U \xrightarrow{g} V \xrightarrow{f} W$ odpowiada macierz

$$M_{f \circ g} = M_f M_g.$$

Jak wiadomo, macierze $m \times n$ o wyrazach z ciała $\mathfrak{K}$ tworzą przestrzeń liniową nad $\mathfrak{K}$ wymiaru mn; bazę tej przestrzeni stanowią np. macierze E_{ij} składające się z samych zer poza jedną jedynką na przecięciu i-tego wiersza i j-tej kolumny. Z powyższych rozważań wynika, że przestrzeń $\mathcal{L}(V, W)$ przekształceń liniowych jest izomorficzna z tą przestrzenią macierzy (rozdz. 1, § 2); w szczególności zachodzi równość

$$\dim \mathcal{L}(V, W) = (\dim V)(\dim W).$$

Przypomnieliśmy tu jedynie znane fakty, chcąc jeszcze raz podkreślić pełną zgodność przekształceń liniowych i macierzy.

3. Wymiar jądra i obrazu. Zachodzi następujące

TWIERDZENIE 4. *Jeśli przestrzeń V jest skończenie wymiarowa i $f : V \to W$ jest przekształceniem liniowym, to przestrzenie* $\operatorname{Ker} f$ *i* $\operatorname{Im} f$ *są też skończenie wymiarowe oraz*

$$\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim V.$$

Dowód (por. część I, rozdz. 2, § 3, twierdzenie 7). Ponieważ $\operatorname{Ker} f \subset V$, więc $\dim \operatorname{Ker} f \leqslant \dim V < \infty$. Wybierzmy bazę $(\mathbf{e}_1, \ldots, \mathbf{e}_k)$ w $\operatorname{Ker} f$ i uzupełnijmy ją do bazy $(\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{e}_{k+1}, \ldots, \mathbf{e}_n)$ w V (rozdz. 1, § 2, twierdzenie 3). Każdy wektor z $\operatorname{Im} f$ jest postaci

$$f\Big(\sum_{i=1}^{n} \alpha_i \mathbf{e}_i\Big) = \sum_{i=k+1}^{n} \alpha_i f(\mathbf{e}_i), \qquad \alpha_i \in \mathfrak{K},$$

czyli wektory $f(\mathbf{e}_{k+1}), \ldots, f(\mathbf{e}_n)$ generują $\operatorname{Im} f$. Wystarczy wykazać, że wektory te są liniowo niezależne.

Przypuśćmy, że $\sum_{i=k+1}^{n} \lambda_i f(\mathbf{e}_i) = \mathbf{0}$. Wtedy $f(\sum_{i=k+1}^{n} \lambda_i \mathbf{e}_i) = \mathbf{0}$, co oznacza, że $\sum_{i=k+1}^{n} \lambda_i \mathbf{e}_i \in \operatorname{Ker} f$, a więc $\sum_{i=k+1}^{n} \lambda_i \mathbf{e}_i = \sum_{j=1}^{k} \lambda_j \mathbf{e}_j$. Zależność liniowa między wektorami bazowymi $\mathbf{e}_1, \ldots, \mathbf{e}_n$ musi być jednak trywialna, w szczególności $\lambda_{k+1} = \ldots = \lambda_n = 0$. Oznacza to, że wektory $f(\mathbf{e}_{k+1}), \ldots, f(\mathbf{e}_n)$ są liniowo niezależne i $\dim \operatorname{Im} f = n - k$. ■

WNIOSEK. *Jeśli* $\dim V < \infty$, *to następujące własności przekształcenia liniowego* $f : V \to W$ *są równoważne:*

(i) f *jest różnowartościowe;*

(ii) $\dim V = \dim \operatorname{Im} f$.

Dowód. Z twierdzenia 4 wynika, że $\dim V = \dim \operatorname{Im} f$ wtedy i tylko wtedy, gdy $\dim \operatorname{Ker} f = 0$, czyli $\operatorname{Ker} f = \{\mathbf{0}\}$, a wiemy już, że zerowe jądro oznacza różnowartościowość. ■

Uwaga. Jeśli $\dim V = \dim W < \infty$ i $f : V \to W$ jest przekształceniem liniowym, to zarówno injektywność ($\operatorname{Ker} f = \{\mathbf{0}\}$), jak i surjektywność ($\operatorname{Im} f = W$) implikują już bijektywność przekształcenia f, czyli f jest w takim wypadku izomorfizmem.

ĆWICZENIA

1. Zapisując kolumnę współrzędnych $X = [x_1, x_2, x_3, x_4]$ w postaci macierzy $\left\| \begin{smallmatrix} x_1 & x_2 \\ x_3 & x_4 \end{smallmatrix} \right\| \in M_2(\mathfrak{K})$, a następnie biorąc dowolną macierz $A = \left\| \begin{smallmatrix} a_1 & a_2 \\ a_3 & a_4 \end{smallmatrix} \right\| \in M_2(\mathfrak{K})$, określamy dwa przekształcenia liniowe

$$f_L : X \mapsto A \begin{Vmatrix} x_1 & x_2 \\ x_3 & x_4 \end{Vmatrix} = \begin{Vmatrix} x_1' & x_2' \\ x_3' & x_4' \end{Vmatrix} = X',$$

$$f_R : X \mapsto \begin{Vmatrix} x_1 & x_2 \\ x_3 & x_4 \end{Vmatrix} A = \begin{Vmatrix} x_1'' & x_2'' \\ x_3'' & x_4'' \end{Vmatrix} = X''.$$

Sprawdzić, że macierze tych przekształceń mają postać

$$M_{f_L} = \begin{Vmatrix} a_1 & 0 & a_2 & 0 \\ 0 & a_1 & 0 & a_2 \\ a_3 & 0 & a_4 & 0 \\ 0 & a_3 & 0 & a_4 \end{Vmatrix}, \quad M_{f_R} = \begin{Vmatrix} a_1 & a_3 & 0 & 0 \\ a_2 & a_4 & 0 & 0 \\ 0 & 0 & a_1 & a_3 \\ 0 & 0 & a_2 & a_4 \end{Vmatrix}.$$

2. Sprawdzić liniowość następujących odwzorowań oraz wyznaczyć $\operatorname{Ker} f$ i $\operatorname{rank} f$:

(a) V jest przestrzenią liniową, $L \subset V$ — podprzestrzenią, $W = V/L$ – przestrzenią ilorazową; odwzorowanie $f : V \to W$ przyporządkowuje wektorowi $\mathbf{x} \in V$ jego warstwę $\overline{\mathbf{x}} = \mathbf{x} + L$;

(b) $f : P_n \to P_n$ jest określone wzorem $f(u(t)) = tu'(t) - u(t)$.

3. Wykazać, że odwzorowanie $f_C : M_n(\mathfrak{K}) \to M_n(\mathfrak{K})$ określone wzorem $f_C(X) = C^{-1}XC$, gdzie $C \in M_n(\mathfrak{K})$ jest macierzą nieosobliwą, jest liniowe i spełnia warunek $f_C(XY) = f_C(X)f_C(Y)$.

§ 2. ALGEBRA OPERATORÓW LINIOWYCH

1. Definicje i przykłady. Ciało skalarów $\mathfrak{K}$ pozostaje na razie dowolne. W przypadku $V = W$ elementy przestrzeni liniowej $\mathcal{L}(V, V)$, oznaczanej częściej przez $\mathcal{L}(V)$ lub $\operatorname{End} V$, nazywa się zwykle *operatorami liniowymi* lub *endomorfizmami* przestrzeni V. Operatory takie będziemy oznaczać literami kaligraficznymi $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \ldots$, a macierze odpowiadające im w ustalonej bazie $(\mathbf{e}_i)$ przestrzeni V — odpowiednimi literami $A, B, C, D \ldots$. W innej bazie $(\mathbf{e}'_i)$ tym samym operatorom $\mathcal{A}, \mathcal{B}, \ldots$ będą odpowiadały macierze $A', B', \ldots$ Symbol $\mathcal{E} = \mathrm{Id}$ oznacza zawsze *operator tożsamościowy* (*jednostkowy*) $\mathbf{x} \mapsto \mathbf{x}$, o macierzy $E = (\delta_{ij})$. Z zasady wartość operatora liniowego $\mathcal{A}$ na wektorze $\mathbf{x}$ będziemy oznaczać bez nawiasów: $\mathcal{A}\mathbf{x}$ (choć niekiedy piszemy również $\mathcal{A}(\mathbf{x})$).

Operator liniowy $\mathcal{B}$ nazywamy *odwrotnym* do $\mathcal{A}$, jeśli $\mathcal{AB} = \mathcal{BA} = \mathcal{E}$. Zgodnie z ogólnymi rezultatami (część I, rozdz. 1, § 5), jeśli operator odwrotny do $\mathcal{A}$ istnieje, to jest wyznaczony jednoznacznie; oznaczamy go przez $\mathcal{A}^{-1}$. W myśl wniosku z twierdzenia 4 w § 1, jeśli $\dim V < \infty$, to istnienie $\mathcal{A}^{-1}$ jest równoważne każdemu z warunków $\operatorname{Ker} \mathcal{A} = \{\mathbf{0}\}$ lub $\dim V = \dim \operatorname{Im} \mathcal{A}$. Ogólnie, liczbę $\dim \operatorname{Ker} \mathcal{A}$ nazywamy *defektem* operatora $\mathcal{A}$. Tak więc operatory o defekcie zero są odwracalne, i tylko one. Ponadto odwracalność oznacza, że równanie $\mathcal{A}\mathbf{x} = \mathbf{b}$ ma rozwiązanie dla każdego $\mathbf{b} \in V$. Przypomnijmy jeszcze, że $\operatorname{rank} \mathcal{A} = \dim \operatorname{Im} \mathcal{A} = \dim V - \dim \operatorname{Ker} \mathcal{A}$ to rząd operatora $\mathcal{A}$. Wszystkie te pojęcia i warunki są nam dobrze znane w języku macierzy; pojęcie operatora liniowego jest jednak bardziej podstawowe, jako niezwiązane z wyborem bazy przestrzeni.

Oto kilka przykładów operatorów liniowych.

Przykład 1. *Operator zerowy* $\mathcal{O}$ przeprowadza każdy wektor w wektor zerowy; $\operatorname{rank}\mathcal{O} = 0$.

Przykład 2. Jednokładność o środku w zerze: $\mathcal{A}\mathbf{x} = \lambda\mathbf{x}$, gdzie λ jest ustalonym skalarem.

Przykład 3. Operator $\mathcal{A}$ obrotu płaszczyzny $\mathbb{R}^2$ o kąt α dookoła początku układu najprościej zrealizować, interpretując $\mathbb{R}^2$ jako płaszczyznę liczb zespolonych z bazą (rzeczywistą) $(1, i)$. Wtedy oczywiście $\mathcal{A} : z = x + iy \mapsto e^{i\alpha}z$ jest operatorem mnożenia przez liczbę $e^{i\alpha} = \cos\alpha + i\sin\alpha$, w szczególności $\mathcal{A}i = -\sin\alpha + i\cos\alpha$. Zatem macierzą operatora $\mathcal{A}$ w bazie $(1, i)$ jest

$$A = \begin{Vmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{Vmatrix}.$$

Przykład 4. Niech $V = U \oplus W$ będzie sumą prostą dwóch podprzestrzeni liniowych. Każdy wektor $\mathbf{x} \in V$ ma jednoznaczny rozkład $\mathbf{x} = \mathbf{x}_U + \mathbf{x}_W$, gdzie $\mathbf{x}_U \in U$ i $\mathbf{x}_W \in W$. Operator $\mathcal{P} : V \to V$ określony wzorem $\mathcal{P}\mathbf{x} = \mathbf{x}_U$ nazywamy *operatorem rzutowania* (lub *projekcją* albo po prostu *rzutem*) na podprzestrzeń U równolegle do W (lub wzdłuż W). Zauważmy, że $\mathcal{P}^2 := \mathcal{P} \circ \mathcal{P} = \mathcal{P}$ (zapis $a := b$ oznacza „a jest z definicji równe b").

Przykład 5. Jeśli $P_n = \langle 1, t, \ldots, t^{n-1} \rangle$ jest przestrzenią wielomianów stopni $\leqslant n-1$ o współczynnikach w $\mathfrak{K}$, to $\mathcal{D}_t = d/dt$ oznacza operator różniczkowania względem t: $\mathcal{D}_t f(t) = f'(t)$.

Warto przestrzec Czytelnika przed możliwą fałszywą interpretacją związku

$$\dim\operatorname{Ker}\mathcal{A} + \dim\operatorname{Im}\mathcal{A} = \dim V \tag{1}$$

(§ 1, twierdzenie 4). Otóż z równości tej nie wynika bynajmniej, że $V = \operatorname{Ker}\mathcal{A} + \operatorname{Im}\mathcal{A}$, o czym świadczy choćby przykład operatora $\mathcal{D}_t$ z przykładu 5 (dla $n \geqslant 2$):

$$\operatorname{Ker}\mathcal{D}_t = \langle 1 \rangle = \mathfrak{K} \cdot 1 \subset \langle 1, t, \ldots, t^{n-2} \rangle = \operatorname{Im}\mathcal{D}_t.$$

2. Algebra operatorów. Wiemy już, że zbiór $\mathcal{L}(V)$ operatorów liniowych na przestrzeni liniowej V stanowi przestrzeń liniową wymiaru

$$\dim\mathcal{L}(V) = (\dim V)^2. \tag{2}$$

Przypominamy, że jeśli $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V)$ i $\mathbf{x} \in V$, to z definicji

$$(\mathcal{A} + \mathcal{B})\mathbf{x} = \mathcal{A}\mathbf{x} + \mathcal{B}\mathbf{x}, \qquad (\lambda\mathcal{A})\mathbf{x} = \lambda(\mathcal{A}\mathbf{x}), \qquad (\mathcal{A}\mathcal{B})\mathbf{x} = \mathcal{A}(\mathcal{B}\mathbf{x})$$

(złożenie $\mathcal{A} \circ \mathcal{B}$ operatorów liniowych zapisuje się zwykle bez kółka, jako $\mathcal{A}\mathcal{B}$).

Łatwo sprawdzić, że tak zdefiniowane operatory spełniają warunki

$$\begin{aligned} \alpha(\mathcal{A}+\mathcal{B}) &= \alpha\mathcal{A}+\alpha\mathcal{B}, \\ (\alpha+\beta)\mathcal{A} &= \alpha\mathcal{A}+\beta\mathcal{A}, \\ (\alpha\beta)\mathcal{A} &= \alpha(\beta\mathcal{A}), \\ 1\cdot\mathcal{A} &= \mathcal{A}; \end{aligned} \tag{3'}$$

$$\begin{aligned} \mathcal{A}(\mathcal{B}\mathcal{C}) &= (\mathcal{A}\mathcal{B})\mathcal{C} \text{ (łączność)}, \\ \mathcal{A}(\mathcal{B}+\mathcal{C}) &= \mathcal{A}\mathcal{B}+\mathcal{A}\mathcal{C}, \\ (\mathcal{A}+\mathcal{B})\mathcal{C} &= \mathcal{A}\mathcal{C}+\mathcal{B}\mathcal{C} \text{ (rozdzielność)}; \end{aligned} \tag{3''}$$

$$\lambda(\mathcal{A}\mathcal{B}) = (\lambda\mathcal{A})\mathcal{B} = \mathcal{A}(\lambda\mathcal{B}). \tag{3'''}$$

Widzimy więc, że zbiór $\mathcal{L}(V)$ operatorów liniowych jest jednocześnie przestrzenią liniową nad ciałem $\mathfrak{K}$ (związki (3′)) i pierścieniem przemiennym (związki (3″)); ostatnia relacja (3‴) to dodatkowy związek mnożenia przez skalary i składania operatorów.

DEFINICJA 1. Pierścień P będący jednocześnie przestrzenią liniową nad ciałem $\mathfrak{K}$ spełniającą warunek $\lambda(ab) = (\lambda a)b = a(\lambda b)$ dla dowolnych $\lambda \in \mathfrak{K}$ i $a, b \in P$ nazywamy *algebrą* nad ciałem $\mathfrak{K}$. Wymiar P jako przestrzeni liniowej nad $\mathfrak{K}$ nazywamy *wymiarem algebry* P nad $\mathfrak{K}$. Każdą podprzestrzeń liniową $Q \subset P$ zamkniętą względem mnożenia w P (piszemy $Q \cdot Q \subset Q$) nazywamy *podalgebrą* algebry P; jeśli P ma jedynkę, to dodatkowo zakładamy, że $1 \in Q$.

Mówiąc o algebrach, mamy najczęściej na myśli algebry łączne ($a(bc) = (ab)c$) i z jedynką ($1 \cdot x = x \cdot 1 = x$ dla $x \in P$). Algebra $\mathcal{L}(V)$ operatorów liniowych ma obie te własności. Jej wariant macierzowy $M_n(\mathfrak{K})$ wystąpił w części I (rozdz. 2, § 3), gdzie sformułowaliśmy odpowiedniki własności (3′)–(3‴) dla macierzy. Choć o związku macierzy i operatorów liniowych była już mowa nie raz (m.in. w § 1), podkreślamy jeszcze raz prosty, ale podstawowy fakt: *jeśli*

$$\mathcal{A} : \mathbf{e}_k \mapsto \mathcal{A}\mathbf{e}_k = \sum_{i=1}^{n} a_{ik}\mathbf{e}_i, \qquad \mathcal{B} : \mathbf{e}_j \mapsto \mathcal{B}\mathbf{e}_j = \sum_{k=1}^{n} b_{kj}\mathbf{e}_k$$

są operatorami liniowymi o macierzach $A = (a_{ik})$ i $B = (b_{kj})$ w bazie $(\mathbf{e}_i)$ przestrzeni V, to macierzą operatora $\mathcal{A}\mathcal{B}$ w tej bazie jest $C = AB$.

Istotnie, jeśli $(\mathcal{A}\mathcal{B})\mathbf{e}_j = \sum_i c_{ij}\mathbf{e}_i$, to

$$\begin{aligned} \sum_i c_{ij}\mathbf{e}_i &= \mathcal{A}(\mathcal{B}\mathbf{e}_j) = \mathcal{A}\Big(\sum_k b_{kj}\mathbf{e}_k\Big) = \sum_k b_{kj}\mathcal{A}\mathbf{e}_k \\ &= \sum_k b_{kj}\sum_i a_{ik}\mathbf{e}_i = \sum_i\Big(\sum_k a_{ik}b_{kj}\Big)\mathbf{e}_i, \end{aligned}$$

tj. $c_{ij} = \sum_k a_{ik}b_{kj}$, czyli $(c_{ij}) = AB$.

Pierścienie najbardziej interesujące z punktu widzenia zastosowań okazują się być algebrami. Zbiór wielomianów $\mathfrak{K}[t]$ jest najprostszym przykładem algebry nieskończenie wymiarowej. W algebrze $\mathcal{L}(V)$ na szczególną uwagę zasługują podalgebry generowane przez jeden operator. Jeśli mianowicie $\mathcal{A} \in \mathcal{L}(V)$, to *podalgebra generowana przez* $\mathcal{A}$, oznaczana przez $\mathfrak{K}[\mathcal{A}]$, jest najmniejszą podalgebrą (z jedynką) w $\mathcal{L}(V)$, zawierającą $\mathcal{A}$. Elementami tej podalgebry są wszystkie potęgi operatora $\mathcal{A}$:

$$\mathcal{A}^0 = \mathcal{E},\ \mathcal{A},\ \mathcal{A}^2 = \mathcal{A}\mathcal{A},\ \ldots,\ \mathcal{A}^k = \underbrace{\mathcal{A}\mathcal{A}\ldots\mathcal{A}}_{k},\ \ldots,$$

oraz ich kombinacje liniowe. Inaczej mówiąc, najogólniejszą postacią elementu z $\mathfrak{K}[\mathcal{A}]$ jest

$$f(\mathcal{A}) = a_0\mathcal{A}^m + a_1\mathcal{A}^{m-1} + \ldots + a_{m-1}\mathcal{A} + a_m\mathcal{E}, \tag{4}$$

gdzie

$$f(t) = a_0t^m + a_1t^{m-1} + \ldots + a_{m-1}t + a_m \in \mathfrak{K}[t].$$

Przyjmując funkcyjny punkt widzenia, można powiedzieć, że $f(\mathcal{A})$ jest wartością wielomianu $f \in \mathfrak{K}[t]$ dla $t = \mathcal{A}$. Operator liniowy $f(\mathcal{A})$ działa na wektory $\mathbf{x} \in V$ w naturalny sposób:

$$f(\mathcal{A})\mathbf{x} = a_0\mathcal{A}^m\mathbf{x} + a_1\mathcal{A}^{m-1}\mathbf{x} + \ldots + a_{m-1}\mathcal{A}\mathbf{x} + a_m\mathbf{x}.$$

Algebra $\mathfrak{K}[\mathcal{A}]$ jest przemienna: $f(\mathcal{A})\cdot g(\mathcal{A}) = g(\mathcal{A})\cdot f(\mathcal{A})$, co wynika z przemienności potęg: $\mathcal{A}^k\mathcal{A}^l = \mathcal{A}^{k+l} = \mathcal{A}^l\mathcal{A}^k$. Jaki jest jej wymiar? Przekonamy się później, że

$$\dim \mathfrak{K}[\mathcal{A}] \leqslant \dim V, \tag{5}$$

ale to już rezultat nieco bardziej subtelny. Na razie ograniczymy się do kilku uwag wstępnych.

Definicja 2. Mówimy, że operator liniowy $\mathcal{A}$ jest *pierwiastkiem* wielomianu $f(t)$, jeśli $f(\mathcal{A}) = \mathcal{O}$. Wielomian unormowany (tzn. o najwyższym współczynniku 1) najniższego stopnia o tej własności nazywamy *wielomianem minimalnym* operatora $\mathcal{A}$.

Niech

$$\mu_{\mathcal{A}}(t) = t^m + \mu_1t^{m-1} + \ldots + \mu_{m-1}t + \mu_m \tag{6}$$

będzie wielomianem minimalnym operatora $\mathcal{A}$. Wówczas operatory $\mathcal{E}, \mathcal{A}, \mathcal{A}^2, \ldots, \mathcal{A}^{m-1}$ są liniowo niezależne, ponieważ nietrywialna zależność liniowa $\sum_{i=0}^{m-1}\lambda_i\mathcal{A}^i = \mathcal{O}$ oznaczałaby, że $\mathcal{A}$ jest pierwiastkiem niezerowego wielomianu $\sum_{i=0}^{m-1}\lambda_it^i$, choć stopień tego wielomianu jest mniejszy od m. Na odwrót, jeśli operatory $\mathcal{E}, \mathcal{A}, \mathcal{A}^2, \ldots, \mathcal{A}^{m-1}$ są liniowo niezależne (jako wektory przestrzeni $\mathcal{L}(V)$), a operator $\mathcal{A}^m$ jest ich kombinacją liniową, to stopień wielomianu minimalnego operatora $\mathcal{A}$ wynosi m. Z zawierania $\mathfrak{K}[\mathcal{A}] \subset \mathcal{L}(V)$ wynika, że m o powyższej własności zawsze istnieje oraz $m \leqslant n^2 = \dim\mathcal{L}(V)$. W ten sposób udowodniliśmy część poniższego twierdzenia.

TWIERDZENIE 1. *Każdy operator liniowy $\mathcal{A}$ ma wielomian minimalny $\mu_{\mathcal{A}}(t)$. Stopień tego wielomianu jest równy wymiarowi algebry $\mathfrak{K}[\mathcal{A}]$. Operator $\mathcal{A}$ jest odwracalny wtedy i tylko wtedy, gdy wyraz wolny μ_m wielomianu* (6) *jest różny od zera.*

Dowód. Uzasadnienie ostatniego stwierdzenia jest równie proste jak powyższy dowód pierwszych dwóch stwierdzeń. Jeśli mianowicie $\mu_m = 0$, to

$$\mathcal{O} = \mu_{\mathcal{A}}(\mathcal{A}) = \mathcal{A}(\mathcal{A}^{m-1} + \mu_1 \mathcal{A}^{m-2} + \ldots + \mu_{m-1}\mathcal{E}).$$

Operator w nawiasie jest niezerowy (na mocy minimalności $\mu_{\mathcal{A}}(t)$), a więc operator $\mathcal{A}$ jest dzielnikiem zera w pierścieniu $\mathcal{L}(V)$, czyli nie może być odwracalny. Na odwrót, jeśli $\mu_m \neq 0$, to zależność

$$\mathcal{A}(-\mu_m^{-1}\mathcal{A}^{m-1} - \mu_m^{-1}\mu_1\mathcal{A}^{m-2} - \ldots - \mu_m^{-1}\mu_{m-1}\mathcal{E}) = \mathcal{E},$$

wynikająca z równości $\mu_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$, daje jawny wzór na operator odwrotny do $\mathcal{A}$. ∎

TWIERDZENIE 2. *Jeśli operator $\mathcal{A}$ jest pierwiastkiem wielomianu $f(t)$, to wielomian ten dzieli się przez wielomian minimalny $\mu_{\mathcal{A}}(t)$* ([1]).

Dowód. Podzielmy wielomian $f(t)$ przez $\mu_{\mathcal{A}}(t)$ z resztą: $f(t) = q(t)\mu_{\mathcal{A}}(t) + r(t)$. Wtedy

$$\mathcal{O} = f(\mathcal{A}) = q(\mathcal{A}) \cdot \mathcal{O} + r(\mathcal{A}),$$

czyli $r(\mathcal{A}) = \mathcal{O}$. Mamy jednak $\deg r(t) < \deg \mu_{\mathcal{A}}(t)$, więc na mocy definicji wielomianu minimalnego $r(t) = 0$. ∎

DEFINICJA 3. Operator liniowy $\mathcal{A}$ nazywamy *nilpotentnym*, jeśli $\mathcal{A}^m = \mathcal{O}$ dla pewnego $m > 0$; najmniejszą liczbę naturalną m o tej własności nazywamy *stopniem nilpotentności* operatora $\mathcal{A}$.

Jeśli $\mathcal{A}$ jest operatorem nilpotentnym stopnia m, to oczywiście $\mu_{\mathcal{A}}(t) = t^m$; jeśli $\mathcal{A}$ jest nietrywialnym (czyli $\neq \mathcal{O}, \mathcal{E}$) operatorem spełniającym warunek $\mathcal{A}^2 = \mathcal{A}$, to $\mu_{\mathcal{A}}(t) = t^2 - t$; ponadto $\mu_{\mathcal{O}}(t) = t$ i $\mu_{\mathcal{E}}(t) = t - 1$. Typowym przykładem operatora nilpotentnego stopnia n jest operator różniczkowania $\mathcal{D}_t$ na przestrzeni P_n wielomianów stopni $\leqslant n-1$. Operator rzutowania $\mathcal{P}$ (przykład 4 z p. 1) spełnia warunek $\mathcal{P}^2 = \mathcal{P}$. Przykłady te będą wielokrotnie wykorzystywane.

3. Macierze operatora liniowego w różnych bazach. Niech V będzie n-wymiarową przestrzenią liniową nad ciałem $\mathfrak{K}$, a $\mathcal{A} : V \to V$ — operatorem

---

([1]) Ponieważ każdy operator jest pierwiastkiem swojego wielomianu minimalnego, więc z twierdzenia tego wynika, że wielomian minimalny operatora jest wyznaczony jednoznacznie; fakt ten będzie później wykorzystywany (*przyp. tłum.*).

liniowym. Wybierając w V bazę $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$, możemy określić $\mathcal{A}$ za pomocą macierzy $A = (a_{ki})$:

$$\mathcal{A}\mathbf{e}_i = \sum_k a_{ki}\mathbf{e}_k. \tag{7}$$

Ten sam operator w innej bazie $(\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ ma jednak inną macierz $A' = (a'_{kj})$:

$$\mathcal{A}\mathbf{e}'_j = \sum_k a'_{kj}\mathbf{e}'_k. \tag{7'}$$

Jeśli $B = (b_{ij})$ jest macierzą przejścia od bazy $(\mathbf{e}_i)$ do $(\mathbf{e}'_j)$, to wzory $\mathbf{e}'_j = \sum_i b_{ij}\mathbf{e}_i$ dla $j = 1, \ldots, n$ sugerują wprowadzenie operatora $\mathcal{B}$:

$$\mathcal{B}\mathbf{e}_j = \mathbf{e}'_j, \tag{8}$$

o macierzy B w bazie $(\mathbf{e}_i)$. Operator $\mathcal{B}$ jest oczywiście odwracalny (operator do niego odwrotny przeprowadza $\mathbf{e}'_j$ na $\mathbf{e}_j$).

Rozważmy jeszcze pomocniczy operator $\mathcal{A}'$, mający w bazie $(\mathbf{e}_i)$ tę samą macierz A', co operator $\mathcal{A}$ w bazie $(\mathbf{e}'_i)$. Innymi słowy,

$$\mathcal{A}'\mathbf{e}_j = \sum_i a'_{ij}\mathbf{e}_i. \tag{9}$$

Taki operator istnieje, ponieważ przy ustalonej bazie mamy odpowiedniość wzajemnie jednoznaczną pomiędzy operatorami a macierzami. Wykorzystując (7) i (8), przepiszemy teraz (7′) w postaci

$$\mathcal{A}\mathcal{B}\mathbf{e}_j = \mathcal{A}\mathbf{e}'_j = \sum_i a'_{ij}\mathbf{e}'_i = \sum_i a'_{ij}\mathcal{B}\mathbf{e}_i = \mathcal{B}\Big(\sum_i a'_{ij}\mathbf{e}_i\Big).$$

Na podstawie odwracalności $\mathcal{B}$ i wzoru (9) wynika stąd, że

$$\mathcal{B}^{-1}\mathcal{A}\mathcal{B}\mathbf{e}_j = \mathcal{A}'\mathbf{e}_j, \qquad j = 1, \ldots, n. \tag{10}$$

Rozpatrując macierze wszystkich operatorów $\mathcal{A}$, $\mathcal{B}$, $\mathcal{A}'$ w tej samej bazie $(\mathbf{e}_i)$, otrzymujemy z (10) równość macierzową

$$A' = B^{-1}AB. \tag{11}$$

Do związku (11) można dojść w sposób bardziej bezpośredni, stosując współrzędne wektorów. Niech, jak zwykle, $\mathbf{x} = \sum_i x_i\mathbf{e}_i = \sum_i x'_i\mathbf{e}'_i$ będą zapisami tego samego wektora $\mathbf{x} \in V$ w obu bazach, a $X = [x_1, \ldots, x_n]$ i $X' = [x'_1, \ldots, x'_n]$ odpowiednimi kolumnami współrzędnych. Niech ponadto $Y = AX$ i $Y' = A'X'$, gdzie A i A' są macierzami określonymi przez (7) i (7′). Ponieważ $X = BX'$ i $Y = BY'$ (rozdz. 1, § 2, (4′)), więc

$$ABX' = AX = Y = BY' = BA'X'.$$

Ponieważ wektor kolumnowy X' jest dowolny, wynika stąd, że $AB = BA'$, czyli $A' = B^{-1}AB$.

W ten sposób dwukrotnie przekonaliśmy się, że zachodzi

TWIERDZENIE 3. *Macierz A' operatora liniowego $\mathcal{A}$ w bazie $(\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ otrzymuje się z macierzy A tego operatora w bazie $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ za pomocą wzoru $A' = B^{-1}AB$, gdzie B jest macierzą przejścia od bazy $(\mathbf{e}_i)$ do $(\mathbf{e}'_i)$.* ■

DEFINICJA 4. Mówimy, że macierz A' jest *podobna* do macierzy A (piszemy $A' \sim A$), jeśli istnieje taka macierz odwracalna B, że $A' = B^{-1}AB$. (Zakładamy, że wszystkie te macierze są macierzami kwadratowymi tego samego stopnia nad jednym i tym samym ciałem $\mathfrak{K}$).

Jasne jest, że zawsze $A \sim A$ (wystarczy wziąć $B = E$). Ponadto związek (11) można przepisać w postaci $A = B_1^{-1}A'B_1$, gdzie $B_1 = B^{-1}$, skąd wynika, że relacja $\sim$ jest symetryczna: $A' \sim A \Rightarrow A \sim A'$. Jest ona również przechodnia: jeśli $A' = B^{-1}AB$ i $A'' = C^{-1}A'C$, to $A'' = (BC)^{-1}A(BC)$, tj. $(A'' \sim A'$ i $A' \sim A) \Rightarrow A'' \sim A$. Tak więc podobieństwo macierzy jest relacją równoważności i zbiór wszystkich macierzy kwadratowych stopnia n dzieli się na rozłączne klasy podobieństwa (część I, rozdz. 2, § 3, p. 6). Twierdzenie 3 głosi, że każdemu operatorowi liniowemu odpowiada dokładnie jedna klasa macierzy podobnych; na odwrót, wszystkie macierze z tej samej klasy podobieństwa są macierzami tego samego operatora w różnych bazach.

Język operatorów liniowych jest wygodny w rozważaniach teoretycznych, jednak konkretne obliczenia przeprowadza się na ogół na macierzach. Z tego względu tak ważna od strony praktycznej jest klasyfikacja macierzy z dokładnością do podobieństwa. Jeśli chcemy np. obliczyć potęgę A^k macierzy A stopnia $n > 1$ (albo, powiedzmy, stopnia $n > 100$) dla dużego wykładnika $k \geqslant 1000$ (a takie zagadnienia pojawiają się w praktyce), to wygodnie jest znaleźć najpierw macierz $A' \sim A$, której potęgi łatwo obliczać. Gdyby się np. udało dobrać $A' = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ (o ile oczywiście taka macierz istnieje w klasie podobieństwa macierzy A), to $A = BA'B^{-1}$ i $A^k = B(A')^k B^{-1} = B \operatorname{diag}(\lambda_1^k, \ldots, \lambda_n^k)B^{-1}$.

Przykładem jest macierz $A = \begin{Vmatrix} 0 & 1 \\ 1 & 1 \end{Vmatrix}$, związana z ciągiem Fibonacciego; była już o niej mowa w części I (rozdz. 2, § 3, p. 5, przykład 3). Dla niej

$$A' = \begin{Vmatrix} (1+\sqrt{5})/2 & 0 \\ 0 & (1-\sqrt{5})/2 \end{Vmatrix}.$$

Dodajmy jeszcze, że umiejętność obliczania potęg pozwala znajdować wartości

$$f(A) = a_0A^m + a_1A^{m-1} + \ldots + a_{m-1}A + a_mE$$

dla dowolnego wielomianu $f(t) = a_0t^m + a_1t^{m-1} + \ldots + a_m$.

4. Wyznacznik i ślad operatora liniowego. Niech $\mathcal{A}$ będzie operatorem liniowym na przestrzeni liniowej V. Jego *wyznacznikiem* $\det \mathcal{A}$ nazywamy wyznacznik $\det A$ macierzy A operatora $\mathcal{A}$ w jakiejkolwiek bazie przestrzeni V. Ponieważ $\det(B^{-1}AB) = \det A$, więc wyznacznik operatora $\mathcal{A}$ jest dobrze określony.

Macierzom odwracalnym odpowiadają operatory odwracalne, zatem operator $\mathcal{A}$ jest odwracalny dokładnie wtedy, gdy $\det \mathcal{A} \neq 0$. Jeśli $\det \mathcal{A} = 0$, to operator $\mathcal{A}$ nazywamy *osobliwym*.

Śladem operatora $\mathcal{A}$ (ang. *trace*) nazywamy skalar

$$\operatorname{tr} \mathcal{A} = \operatorname{tr} A = \sum_{i=1}^{n} a_{ii}.$$

Jak już wiemy i jak łatwo sprawdzić,

$$\operatorname{tr} AB = \operatorname{tr} BA \tag{12}$$

dla dowolnych macierzy kwadratowych tego samego stopnia. Jeśli w szczególności macierz B jest odwracalna, to

$$\operatorname{tr}(B^{-1}AB) = \operatorname{tr}(B \cdot B^{-1}A) = \operatorname{tr} A,$$

skąd wynika, że ślad operatora jest dobrze określony, tzn. nie zależy od wyboru bazy. Odpowiednikiem (12) dla operatorów jest równość

$$\operatorname{tr} \mathcal{AB} = \operatorname{tr} \mathcal{BA}. \tag{12'}$$

Obie funkcje $\det, \operatorname{tr} : \mathcal{L}(V) \to \mathfrak{K}$ odgrywają ważną rolę. Funkcja det jest multiplikatywna ($\det(\mathcal{AB}) = (\det \mathcal{A})(\det \mathcal{B})$) i pozwala wyodrębnić grupę $\operatorname{Aut} V$ automorfizmów przestrzeni V; mianowicie operator liniowy $\mathcal{A}$ jest automorfizmem wtedy i tylko wtedy, gdy jest *nieosobliwy*, tj. gdy $\det \mathcal{A} \neq 0$. Wybór bazy w V pozwala utożsamić grupę $\operatorname{Aut} V$ ze znaną z części I grupą $\mathrm{GL}_n(\mathfrak{K})$ macierzy odwracalnych stopnia $n = \dim_{\mathfrak{K}} V$ o elementach z ciała $\mathfrak{K}$. Dokładniej: ma miejsce izomorfizm grup $\operatorname{Aut} V \cong \mathrm{GL}_n(\mathfrak{K})$.

Funkcja tr z kolei jest liniowa:

$$\operatorname{tr}(\alpha \mathcal{A} + \beta \mathcal{B}) = \alpha \operatorname{tr} \mathcal{A} + \beta \operatorname{tr} \mathcal{B}$$

(jak łatwo sprawdzić) i fakt ten jest szeroko wykorzystywany w matematyce; np. teoria charakterów grup (zob. część III) jest w całości oparta na pojęciu śladu. Na razie rozpatrzymy skromniejsze zastosowania tego pojęcia.

Przykład 6 (algebry Liego). Algebry, podobnie jak pierścienie, nie zawsze są łączne. Ważnymi przykładami algebr niełącznych są tzw. *algebry Liego* (na cześć Sophusa Liego (1842–1899)). Mnożenie $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} * \mathbf{y}$ w algebrze Liego L spełnia dwa aksjomaty:

(i) $\mathbf{x} * \mathbf{x} = 0$; stąd $(\mathbf{x} + \mathbf{y}) * (\mathbf{x} + \mathbf{y}) = 0$, a więc $\mathbf{x} * \mathbf{y} = -\mathbf{y} * \mathbf{x}$ (*antyprzemienność*);

(ii) $(\mathbf{x} * \mathbf{y}) * \mathbf{z} + (\mathbf{y} * \mathbf{z}) * \mathbf{x} + (\mathbf{z} * \mathbf{x}) * \mathbf{y} = 0$ (*tożsamość Jacobiego*).

Działanie $\mathbf{x} * \mathbf{y}$ zapisuje się na ogół jako $[\mathbf{x}, \mathbf{y}]$ i nazywa *nawiasem Liego* lub *komutatorem* elementów $\mathbf{x}$ i $\mathbf{y}$. Przestrzeń $L = \mathbb{R}^3 = \langle \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \rangle$ jest trójwymiarową algebrą Liego z komutatorem $[\mathbf{x}, \mathbf{y}]$ określonym jako iloczyn wektorowy wektorów $\mathbf{x}$ i $\mathbf{y}$: jeśli

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3, \qquad \mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + y_3\mathbf{e}_3,$$

to

$$[\mathbf{x}, \mathbf{y}] = (x_2y_3 - x_3y_2)\mathbf{e}_1 + (x_3y_1 - x_1y_3)\mathbf{e}_2 + (x_1y_2 - x_2y_1)\mathbf{e}_3.$$

Ponadto łatwo sprawdzić (każdy matematyk powinien to zrobić przynajmniej raz w życiu!), że jeśli w zbiorze $L = \mathcal{L}(V)$, gdzie V jest dowolną przestrzenią liniową, określimy komutator wzorem

$$[\mathcal{A}, \mathcal{B}] = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}, \tag{13}$$

to aksjomaty (i) i (ii) będą spełnione, czyli możemy traktować L jako algebrę Liego. Oznaczamy ją wtedy przez $\mathfrak{gl}(V)$. Jeśli $\dim V < \infty$, to algebra Liego $\mathfrak{gl}(V)$ jest oczywiście izomorficzna z algebrą wszystkich macierzy $n \times n$, w której określono komutator macierzy wzorem analogicznym do (13); tę ostatnią algebrę Liego oznacza się przez $\mathfrak{gl}_n(\mathfrak{K})$. Istnieje głębokie twierdzenie głoszące, że każda skończenie wymiarowa algebra Liego nad $\mathfrak{K}$ jest (izomorficzna z) podalgebrą algebry $(\mathcal{L}(V); [\ ,\ ])$ dla pewnej przestrzeni V; przypominamy, że przez podalgebrę rozumiemy podprzestrzeń liniową zamkniętą względem działania $[\ ,\ ]$.

Algebry Liego, zarówno skończenie jak i nieskończenie wymiarowe, odgrywają istotną rolę w mechanice kwantowej (zob. podręcznik [2] w spisie literatury uzupełniającej). Rzecz polega na tym, że tzw. zmienne dynamiczne w teorii kwantowej są nieprzemienne i stopień ich nieprzemienności mierzy się właśnie komutatorem (13). Nietrywialny i w pewnym sensie bliski mechanice kwantowej przykład tzw. relacji komutacji otrzymamy, biorąc jako V nieskończenie wymiarową przestrzeń wszystkich wielomianów nad ciałem $\mathfrak{K}$. Niech $\mathcal{D}_t = d/dt$ będzie operatorem różniczkowania względem t, a $\mathcal{F}_t$ — operatorem mnożenia przez t: $\mathcal{D}_t(f) = f'$, $\mathcal{F}_t(f) = t \cdot f$. Łatwo sprawdzić, że

$$[\mathcal{D}_t, \mathcal{F}_t] = \mathcal{D}_t\mathcal{F}_t - \mathcal{F}_t\mathcal{D}_t = \mathcal{E} \tag{14}$$

jest operatorem tożsamościowym na $V = \mathfrak{K}[t]$.

Powstaje pytanie: czy relacja $[\mathcal{A}, \mathcal{B}] = \mathcal{E}$ typu (14) może zachodzić w skończenie wymiarowej algebrze $\mathcal{L}(V)$? Okazuje się, że odpowiedź zależy od charakterystyki ciała skalarów $\mathfrak{K}$. Jeśli $\mathfrak{K} = \mathbb{C}$ lub $\mathfrak{K} = \mathbb{R}$ (przypadki najbardziej interesujące), to natychmiast dochodzimy do sprzeczności:

$$0 = \operatorname{tr}\mathcal{A}\mathcal{B} - \operatorname{tr}\mathcal{B}\mathcal{A} = \operatorname{tr}[\mathcal{A}, \mathcal{B}] = \operatorname{tr}\mathcal{E} = n = \dim V.$$

Jeśli jednak $\operatorname{char}\mathfrak{K} = p$ oraz $p \mid n$, to sprzeczności już nie ma i np. macierze

$$J_p = \begin{Vmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \hdotsfor{6} \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{Vmatrix}, \quad N_p = \begin{Vmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 2 & \dots & 0 & 0 \\ \hdotsfor{5} \\ 0 & 0 & \dots & p-1 & 0 \end{Vmatrix} \tag{15}$$

stopnia p nad $\mathfrak{K}$ spełniają warunek $[J_p, N_p] = E_p$.

ĆWICZENIA

1. Sprawdzić, że obie macierze w (15) są nilpotentne: $J_p^p = N_p^p = 0$.

2. Wykazać, że jeśli A, B, C są macierzami wymiarów odpowiednio $n \times p$, $p \times q$, $q \times n$, to $\operatorname{tr}(ABC) = \operatorname{tr}(BCA) = \operatorname{tr}(CAB)$.

3. Znaleźć rząd grupy $\mathrm{GL}_n(\mathbb{F}_p)$, interpretując ją jako grupę automorfizmów $\operatorname{Aut} V$ przestrzeni liniowej V wymiaru n nad $\mathbb{F}_p$.

4. Wykazać, że zbiór $\mathfrak{sl}(V)$ operatorów liniowych o śladzie zero jest podalgebrą kowymiaru 1 algebry Liego $\mathcal{L}(V) = \mathfrak{gl}(V)$.

5. Udowodnić, że dla dowolnych operatorów liniowych $\mathcal{A}, \mathcal{B} : V \to V$ zachodzi równość

$$\operatorname{rank} \mathcal{A} = \operatorname{rank} \mathcal{B} + \dim(\operatorname{Im} \mathcal{A} \cap \operatorname{Ker} \mathcal{B}).$$

6. Wykorzystując ćwiczenie 5, udowodnić, że dla dowolnych operatorów liniowych $\mathcal{A}, \mathcal{B}, \mathcal{C} : V \to V$ zachodzi *nierówność Frobeniusa*

$$\operatorname{rank} \mathcal{BA} + \operatorname{rank} \mathcal{AC} \leqslant \operatorname{rank} \mathcal{A} + \operatorname{rank} \mathcal{BAC}.$$

7. Udowodnić, że dla dowolnego operatora liniowego $\mathcal{A} : V \to V$ i każdego $i \geqslant 1$ zachodzi wzór

$$\dim(\operatorname{Im} \mathcal{A}^{i-1} \cap \operatorname{Ker} \mathcal{A}) = \dim \operatorname{Ker} \mathcal{A}^i - \dim \operatorname{Ker} \mathcal{A}^{i-1}$$

(dla $i = 1$ wzór jest oczywisty — pamiętajmy, że z definicji $\mathcal{A}^0 = \mathcal{E}$).

8. Udowodnić, że jeśli dwie macierze $A, B \in M_n(\mathbb{R})$ są podobne nad ciałem liczb zespolonych, to są również podobne nad ciałem liczb rzeczywistych.

9. Uogólniając definicję 2, powiemy, że operator liniowy $\mathcal{A}$ jest pierwiastkiem wielomianu $f(t)$ *względem wektora* $\mathbf{v} \in V$, jeśli $f(\mathcal{A})\mathbf{v} = \mathbf{0}$. Unormowany wielomian najniższego stopnia o tej własności nazywamy *wielomianem minimalnym operatora* $\mathcal{A}$ *względem* $\mathbf{v}$ i oznaczamy przez $\mu_{\mathcal{A},\mathbf{v}}(t)$. Załóżmy, że ciało $\mathfrak{K}$ jest nieskończone. Udowodnić, że:

(a) $\mu_{\mathcal{A},\mathbf{v}}(t)$ dzieli $\mu_{\mathcal{A}}(t)$;
(b) istnieje taki wektor $\mathbf{a} \in V$, że $\mu_{\mathcal{A},\mathbf{a}}(t) = \mu_{\mathcal{A}}(t)$.

10. Niech V będzie przestrzenią liniową, a U i W — dwiema jej podprzestrzeniami liniowymi. Załóżmy, że istnieją rozkłady na sumy proste
$$V = V_1 \oplus V_2, \qquad W = W_1 \oplus W_2,$$
przy czym $W_i \subset V_i$ dla $i = 1, 2$. Niech $\mathcal{P}_i : V \to V$ będzie operatorem rzutowania na V_i równolegle do V_j $(i \neq j)$. Udowodnić, że:

(a) jeśli
$$V_1 = W_1 + U \cap V_1, \qquad V_2 = W_2 + \mathcal{P}_2(U), \tag{$*$}$$
to $V = W + U$;

(b) jeśli $V = W + U$ i $\mathcal{P}_2(U) \cap W_2 = \{\mathbf{0}\}$, to zachodzą równości $(*)$, przy czym $W \cap V = W_1 \cap U$.

11. Udowodnić, że jeśli $\operatorname{char} \mathfrak{K} = 0$, to każda macierz $A \in M_n(\mathfrak{K})$ o zerowym śladzie jest podobna do pewnej macierzy A' z zerami na głównej przekątnej ($A' = (a'_{ij})$, $a'_{11} = a'_{22} = \ldots = a'_{nn} = 0$).

12. Czy założenie $\operatorname{char} \mathfrak{K} = 0$ w ćwiczeniu 11 jest istotne?

§ 3. PODPRZESTRZENIE NIEZMIENNICZE I WEKTORY WŁASNE

1. Rzuty. W przykładzie 4 z § **2** (p. 1) z każdym rozkładem przestrzeni V na sumę prostą dwóch podprzestrzeni związaliśmy dwa operatory rzutowania $\mathcal{P}$, spełniające warunek $\mathcal{P}^2 = \mathcal{P}$. Na odwrót: każdy operator o tej własności (operatory takie nazywamy *idempotentnymi*) jest rzutem na pewną podprzestrzeń. Udowodnimy stwierdzenie ogólniejsze.

Niech $V = W_1 \oplus \ldots \oplus W_m$ będzie rozkładem przestrzeni V na sumę prostą m podprzestrzeni liniowych (rozdz. 1, § 2, p. 5). Wtedy każdy wektor $\mathbf{x} \in V$ można jednoznacznie zapisać w postaci
$$\mathbf{x} = \mathbf{x}_1 + \ldots + \mathbf{x}_m, \qquad \mathbf{x}_i \in W_i,$$
a odwzorowanie $\mathcal{P}_i : \mathbf{x} \mapsto \mathbf{x}_i$ jest operatorem liniowym na V. Ponadto
$$\mathcal{P}_1 + \ldots + \mathcal{P}_m = \mathcal{E},$$
przy czym $\mathcal{P}_i\mathcal{P}_j = \mathcal{O}$ dla $i \neq j$ oraz $\mathcal{P}_i^2 = \mathcal{P}_i$. Wreszcie
$$W_i = \mathcal{P}_i V = \{\mathbf{x} \in V \mid \mathcal{P}_i \mathbf{x} = \mathbf{x}\}, \qquad K_i := \operatorname{Ker} \mathcal{P}_i = W_1 + \ldots + \widehat{W_i} + \ldots + W_m$$
i $\mathcal{P}_i$ jest operatorem rzutowania przestrzeni V na W_i wzdłuż K_i.

TWIERDZENIE 1. *Niech $\mathcal{P}_1, \ldots, \mathcal{P}_m : V \to V$ będzie skończonym układem operatorów liniowych spełniających warunki*
$$\sum_{i=1}^{m} \mathcal{P}_i = \mathcal{E}; \qquad \mathcal{P}_i^2 = \mathcal{P}_i, \quad i = 1, \ldots, m; \qquad \mathcal{P}_i\mathcal{P}_j = \mathcal{O}, \quad i \neq j. \tag{1}$$

Wtedy

$$V = W_1 \oplus \ldots \oplus W_m, \quad gdzie \quad W_i = \operatorname{Im} \mathcal{P}_i.$$

Dowód. Z założenia dla każdego $\mathbf{x} \in V$ mamy

$$\mathbf{x} = \mathcal{E}\mathbf{x} = \sum \mathcal{P}_i \mathbf{x} = \mathbf{x}_1 + \ldots + \mathbf{x}_m, \qquad \mathbf{x}_i \in W_i.$$

Zatem $V = W_1 + \ldots + W_m$. Suma ta jest prosta; aby się o tym przekonać, zastosujemy kryterium podane w rozdziale 1 (§ 2, p. 5, twierdzenie 7). Załóżmy mianowicie, że $\mathbf{x} \in W_j \cap (\sum_{i \neq j} W_i)$. Ponieważ $W_i = \operatorname{Im} \mathcal{P}_i$, więc istnieją takie wektory $\mathbf{y}_1, \ldots, \mathbf{y}_m \in V$, że

$$\mathbf{x} = \mathcal{P}_j(\mathbf{y}_j) = \sum_{i \neq j} \mathcal{P}_i(\mathbf{y}_i).$$

Działając na tę równość operatorem $\mathcal{P}_j$ i wykorzystując własności $\mathcal{P}_j^2 = \mathcal{P}_j$ oraz $\mathcal{P}_j \mathcal{P}_i = \mathcal{O}$ dla $i \neq j$, otrzymujemy

$$\mathbf{x} = \mathcal{P}_j(\mathbf{y}_j) = \mathcal{P}_j^2(\mathbf{y}_j) = \sum_{i \neq j} \mathcal{P}_j \mathcal{P}_i(\mathbf{y}_i) = \mathbf{0}.$$

Na mocy wspomnianego kryterium suma $V = \sum W_i$ jest więc prosta oraz $\mathcal{P}_i$ jest rzutem przestrzeni V na W_i wzdłuż $K_i = \operatorname{Ker} \mathcal{P}_i = \sum_{j \neq i} W_j$. ∎

Dodajmy, że jeśli $\mathcal{P}^2 = \mathcal{P}$ i $V = U \oplus W$ jest rozkładem związanym z rzutem $\mathcal{P}$, tzn. $U = \operatorname{Im} \mathcal{P} = \langle \mathbf{e}_1, \ldots, \mathbf{e}_r \rangle$ i $W = \operatorname{Ker} \mathcal{P} = \langle \mathbf{e}_{r+1}, \ldots, \mathbf{e}_n \rangle$, to w bazie $(\mathbf{e}_i)$ operator $\mathcal{P}$ ma macierz

$$P = \left\| \begin{matrix} E_r & 0 \\ 0 & 0 \end{matrix} \right\|, \qquad r = \operatorname{rank} \mathcal{P}. \tag{2}$$

W szczególności każda macierz A wymiarów $n \times n$ rzędu r o własności $A^2 = A$ jest podobna do macierzy postaci (2), tj. $B^{-1}AB = P$ dla pewnej macierzy nieosobliwej B; ponadto $\operatorname{rank} A = \operatorname{tr} A$.

Uwaga. Jeśli operatory $\mathcal{P}_1, \ldots, \mathcal{P}_m$ spełniają warunki

$$\mathcal{P}_i \mathcal{P}_j = \delta_{ij} \mathcal{P}_i, \qquad i, j = 1, \ldots, m,$$

to często mówimy, że stanowią one *układ ortogonalny operatorów idempotentnych*; ich macierze tworzą *układ ortogonalny macierzy idempotentnych*. Jeśli spełnione są wszystkie warunki (1), to układ taki nazywamy *zupełnym*.

2. Podprzestrzenie niezmiennicze. Każdy operator liniowy działa nie tylko na pojedyncze wektory $\mathbf{x} \in V$, ale i na podprzestrzenie liniowe $U \subset V$: definiujemy $\mathcal{A}U = \{\mathcal{A}\mathbf{x} \mid \mathbf{x} \in U\}$. W tym kontekście dużą rolę odgrywa pojęcie niezmienniczości.

DEFINICJA 1. Podprzestrzeń liniowa $U \subset V$ jest *niezmiennicza* dla operatora liniowego $\mathcal{A} : V \to V$ (lub $\mathcal{A}$-*niezmiennicza*), jeśli $\mathcal{A}U \subset U$.

Na przykład $\operatorname{Ker} \mathcal{A}$ i $\operatorname{Im} \mathcal{A}$ są zawsze podprzestrzeniami niezmienniczymi dla $\mathcal{A}$, chociaż mogą być trywialne, tzn. równe $\{0\}$ lub V. Dla operatora $\mathcal{D}_t$ różniczkowania na przestrzeni P_n wielomianów stopni $\leqslant n-1$ widać od razu ciąg podprzestrzeni niezmienniczych

$$\{0\} \subset V_1 \subset \ldots \subset V_n = V, \tag{3}$$

gdzie V_i składa się z wielomianów stopni $\leqslant i-1$ dla $i = 1, \ldots, n$.

Rodzina rzutów $\mathcal{P}_1, \ldots, \mathcal{P}_m$ na składniki rozkładu $V = W_1 \oplus \ldots \oplus W_m$, rozpatrzona w punkcie 1, ma tę własność, że wszystkie podprzestrzenie postaci

$$W_{i_1} \oplus \ldots \oplus W_{i_k}, \qquad \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, m\}$$

są niezmiennicze dla każdego z operatorów $\mathcal{P}_1, \ldots, \mathcal{P}_m$ (wykorzystujemy tu oczywisty fakt, że suma i część wspólna podprzestrzeni $\mathcal{A}$-niezmienniczych są też $\mathcal{A}$-niezmiennicze).

W pewnym sensie przeciwną własność mają macierze J_p i N_p, zdefiniowane w końcu § 2. Wyznaczone przez nie operatory na przestrzeni $V = \mathfrak{K}^p$ (gdzie $\mathfrak{K}$ jest ciałem o charakterystyce $p > 0$) nie mają żadnej wspólnej nietrywialnej podprzestrzeni niezmienniczej. Można wskazać jedną istotną przyczynę tej różnicy: $[\mathcal{P}_i, \mathcal{P}_j] = \mathcal{O}$, podczas gdy $[J_p, N_p] \neq 0$.

Podobne zjawisko występuje również dla ciała $\mathbb{R}$. Operator $\mathcal{A}$ obrotu płaszczyzny dookoła początku układu o kąt α, gdzie $0 < \alpha < \pi$ (§ 2, p. 1, przykład 3), nie ma nietrywialnych podprzestrzeni niezmienniczych — żadna prosta nie przechodzi na siebie przy działaniu $\mathcal{A}$.

Jeśli istnieje właściwa podprzestrzeń niezmiennicza $U \subset V$, to w odpowiednio dobranej bazie macierz operatora $\mathcal{A}$ ma prostszą postać. Bierzemy mianowicie dowolną bazę $(\mathbf{e}_1, \ldots, \mathbf{e}_m)$ w U i uzupełniamy ją do bazy $(\mathbf{e}_1, \ldots, \mathbf{e}_m, \mathbf{e}_{m+1}, \ldots, \mathbf{e}_n)$ w V; wtedy z warunku $\mathcal{A}\mathbf{e}_i \in U$ dla $i = 1, \ldots, m$ wynika, że macierz operatora $\mathcal{A}$ w tej bazie ma postać

$$A = \begin{Vmatrix} A_1 & A_0 \\ 0 & A_2 \end{Vmatrix}, \tag{4}$$

gdzie A_1 jest macierzą $m \times m$, A_2 — macierzą $(n-m) \times (n-m)$, a A_0 macierzą $m \times (n-m)$. Na macierz A_1 możemy patrzeć jak na macierz operatora $\mathcal{A}_U : U \longrightarrow U$ wyznaczonego przez $\mathcal{A}$ (piszemy $A_1 = A_U$).

Załóżmy na chwilę, że macierz A_0 jest zerowa. Wtedy oczywiście podprzestrzeń $W = \langle \mathbf{e}_{m+1}, \ldots, \mathbf{e}_n \rangle$ jest też $\mathcal{A}$-niezmiennicza oraz A_2 jest macierzą operatora $\mathcal{A}_W : W \longrightarrow W$. W takim wypadku mówimy o *sumie prostej operatorów*

$$\mathcal{A} = \mathcal{A}_U \dotplus \mathcal{A}_W, \tag{5}$$

odpowiadającej rozkładowi $V = U \oplus W$ na sumę prostą podprzestrzeni niezmienniczych. Macierz sumy prostej operatorów ma strukturę *blokowo-diagonalną*:

$$A = \begin{Vmatrix} A_U & 0 \\ 0 & A_W \end{Vmatrix} = A_U \dotplus A_W. \tag{5'}$$

Udowodniliśmy zatem

TWIERDZENIE 2. *Przestrzeń liniowa V jest sumą prostą dwóch nietrywialnych podprzestrzeni niezmienniczych operatora $\mathcal{A} : V \to V$ wtedy i tylko wtedy, gdy macierz tego operatora w pewnej bazie ma postać blokowo-diagonalną* (5'). ■

Twierdzenie to ma oczywiste uogólnienie na przypadek dowolnej liczby m podprzestrzeni niezmienniczych ($2 \leqslant m \leqslant n$), których sumą prostą jest V. Jeśli $m = n = \dim V$, to macierz operatora w pewnej bazie jest diagonalna.

Może się zdarzyć, że macierz A_0 w równości (4) jest zawsze niezerowa, niezależnie od tego, jak wybierzemy wektory $\mathbf{e}_{m+1}, \ldots, \mathbf{e}_n$ uzupełniające bazę $(\mathbf{e}_1, \ldots, \mathbf{e}_m)$ podprzestrzeni niezmienniczej U do bazy w V. Oznacza to, że chociaż przestrzeń V można rozłożyć na sumę prostą $U \oplus W$ na wiele sposobów (rozdz. 1, § 2, twierdzenie 9), to żadna z podprzestrzeni dopełniających W nie jest $\mathcal{A}$-niezmiennicza. Jako dobry przykład takiej sytuacji może służyć operator różniczkowania $\mathcal{D}_t$: jeśli V_i jest jedną z podprzestrzeni niezmienniczych w łańcuchu (3) i $V = V_i \oplus W_i$, to zawsze $\mathcal{D}_t(W_i) \cap V_i \neq \{0\}$.

Zauważmy na koniec, że jeśli $\mathcal{A}U \subset U$ i $\mathcal{B}U \subset U$, tj. U jest wspólną podprzestrzenią niezmienniczą dla operatorów $\mathcal{A}$ i $\mathcal{B}$, to U jest też podprzestrzenią niezmienniczą dla ich kombinacji liniowych $\alpha\mathcal{A} + \beta\mathcal{B}$ oraz złożeń $\mathcal{AB}$ i $\mathcal{BA}$. W szczególności

$$\mathcal{A}U \subset U \;\Rightarrow\; f(\mathcal{A})U \subset U$$

dla każdego wielomianu $f \in \mathfrak{K}[t]$.

3. Wektory własne. Wielomian charakterystyczny. Na szczególną uwagę zasługują jednowymiarowe podprzestrzenie niezmiennicze.

DEFINICJA 2. Każdy niezerowy wektor należący do jednowymiarowej podprzestrzeni niezmienniczej operatora $\mathcal{A}$ nazywamy *wektorem własnym* operatora $\mathcal{A}$. Jeśli $\mathbf{x}$ jest takim wektorem, tzn.

$$\mathbf{x} \neq \mathbf{0} \quad \text{oraz} \quad \mathcal{A}\mathbf{x} = \lambda\mathbf{x} \quad \text{dla pewnego } \lambda \in \mathfrak{K},$$

to skalar λ nazywamy *wartością własną* operatora $\mathcal{A}$, odpowiadającą wektorowi własnemu $\mathbf{x}$.

Zauważmy, że

$$\mathcal{A}\mathbf{x} = \lambda\mathbf{x} \;\Rightarrow\; \mathcal{A}^k\mathbf{x} = \lambda^k\mathbf{x} \text{ dla każdego } k \in \mathbb{N},$$

a stąd

$$f(\mathcal{A})\mathbf{x} = f(\lambda)\mathbf{x} \tag{6}$$

dla dowolnego wielomianu $f \in \mathfrak{K}[t]$. W szczególności

$$f(\mathcal{A}) = \mathcal{O} \;\Rightarrow\; f(\lambda) = 0 \tag{7}$$

dla każdej wartości własnej λ operatora $\mathcal{A}$.

Niech
$$V^\lambda = \{\mathbf{v} \in V \mid \mathcal{A}\mathbf{v} = \lambda\mathbf{v}\};$$
zbiór ten składa się z $\mathbf{0}$ i wszystkich wektorów własnych odpowiadających wartości własnej λ. Z oczywistej implikacji
$$(\mathcal{A}\mathbf{x} = \lambda\mathbf{x} \text{ i } \mathcal{A}\mathbf{y} = \lambda\mathbf{y}) \Rightarrow \mathcal{A}(\alpha\mathbf{x} + \beta\mathbf{y}) = \lambda(\alpha\mathbf{x} + \beta\mathbf{y})$$
wynika, że zbiór V^λ jest podprzestrzenią liniową; możemy więc przyjąć następującą definicję:

DEFINICJA 3. Zbiór V^λ nazywamy *podprzestrzenią własną* operatora $\mathcal{A}$, odpowiadającą wartości własnej λ.

Warunek istnienia wektora własnego (czyli $V^\lambda \neq \{\mathbf{0}\}$) można zapisać w postaci
$$(\mathcal{A} - \lambda\mathcal{E})\mathbf{x} = 0, \qquad \mathbf{x} \neq \mathbf{0}, \tag{8}$$
tj.
$$\operatorname{Ker}(\mathcal{A} - \lambda\mathcal{E}) \neq \{\mathbf{0}\}.$$
Oznacza to, że operator $\mathcal{A} - \lambda\mathcal{E}$ jest osobliwy:
$$\det(\mathcal{A} - \lambda\mathcal{E}) = 0. \tag{9}$$
Jeśli A jest macierzą operatora $\mathcal{A}$ w pewnej bazie $(\mathbf{e}_i)$, to macierzą operatora $\mathcal{A} - \lambda\mathcal{E}$ jest $A - \lambda E$, a więc warunek (9) można przepisać w postaci
$$\det(A - \lambda E) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0. \tag{9'}$$
Niech $\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n$ będzie wektorem własnym operatora $\mathcal{A}$, odpowiadającym wartości własnej λ. We współrzędnych równość (8) przybiera postać
$$\begin{aligned} (a_{11} - \lambda)x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0, \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + \ldots + a_{2n}x_n &= 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + (a_{nn} - \lambda)x_n &= 0. \end{aligned}$$
Jest to układ równań liniowych jednorodnych o zerowym wyznaczniku (9′). Zbiór jego rozwiązań można utożsamić z podprzestrzenią własną V^λ. Jak wiemy z części I, wymiar $\dim V^\lambda$ tego zbioru rozwiązań (zwany *krotnością geometryczną* wartości własnej λ) wynosi $n - r$, gdzie r oznacza rząd macierzy $A - \lambda E$.

Rozwijając wyznacznik $\det(tE - A) = (-1)^n \det(A - tE)$ według wzoru z rozdziału 3 części I (§ 1, (3)):
$$\det(tE - A) = \sum_{\pi \in S_n} \varepsilon_\pi(\delta_{1,\pi 1}t - a_{1,\pi 1})(\delta_{2,\pi 2}t - a_{2,\pi 2}) \ldots (\delta_{n,\pi n}t - a_{n,\pi n}),$$

otrzymamy unormowany wielomian

$$\chi_A(t) = \det(tE - A) = t^n + \chi_1 t^{n-1} + \ldots + \chi_{n-1} t + \chi_n \tag{10}$$

stopnia n zmiennej t o współczynnikach $\chi_i \in \mathfrak{K}$.

Definicja 4. Wielomian (10) nazywamy *wielomianem charakterystycznym* macierzy A. Równanie $\chi_A(t) = 0$ to *równanie charakterystyczne* tej macierzy.

W rzeczywistości można mówić o wielomianie (i równaniu) charakterystycznym operatora liniowego $\mathcal{A}$, ponieważ z twierdzenia 3 w § 2 wynika proste

Twierdzenie 3. *Macierze podobne mają te same wielomiany charakterystyczne.*

Dowód. Jeśli $A' = C^{-1}AC$, to również $tE - A' = C^{-1}(tE - A)C$, więc $\det(tE - A') = \det(tE - A)$. ∎

Możemy więc określić

$$\chi_{\mathcal{A}}(t) = \chi_A(t).$$

Z definicji (10) wynika, że skalar $\lambda \in \mathfrak{K}$ jest wartością własną operatora $\mathcal{A}$ wtedy i tylko wtedy, gdy $\chi_{\mathcal{A}}(\lambda) = 0$, tj. gdy λ jest pierwiastkiem wielomianu charakterystycznego. Jeśli wielomian ten nie ma pierwiastków w $\mathfrak{K}$, to operator $\mathcal{A}$ nie ma wektorów własnych. Z zasadniczego twierdzenia algebry wynika, że *każdy operator liniowy na zespolonej przestrzeni liniowej ma wektory własne.*

Definicja 5. Krotność λ jako pierwiastka wielomianu charakterystycznego nazywamy *krotnością algebraiczną* wartości własnej λ operatora $\mathcal{A}$.

Twierdzenie 4. *Krotność geometryczna wartości własnej nie przekracza jej krotności algebraicznej.*

Dowód. Z definicji krotność geometryczna to wymiar m przestrzeni V^λ rozwiązań równania $\mathcal{A}\mathbf{x} = \lambda\mathbf{x}$. Przestrzeń V^λ jest oczywiście $\mathcal{A}$-niezmiennicza i jeśli oznaczymy przez $\mathcal{A}'$ operator na V^λ wyznaczony przez $\mathcal{A}$, to $\det(t\mathcal{E}' - \mathcal{A}') = (t - \lambda)^m$, przy czym $\chi_{\mathcal{A}}(t) = (t - \lambda)^m q(t)$ dla pewnego wielomianu $q(t) \in \mathfrak{K}[t]$. Przypuśćmy, że λ jest pierwiastkiem wielomianu $q(t)$ o krotności $k \geqslant 0$. Wtedy wartość własna λ ma krotność algebraiczną $m + k$. ∎

4. Kryterium diagonalizowalności. Pierwiastki wielomianu charakterystycznego (zwane też *pierwiastkami charakterystycznymi*) tworzą zbiór niosący ważne informacje o operatorze $\mathcal{A}$. Z oczywistych powodów nie wszystkie te pierwiastki są równoprawne.

DEFINICJA 6. Zbiór wszystkich wartości własnych operatora liniowego $\mathcal{A}$ nazywamy *widmem* (lub *spektrum*) tego operatora i oznaczamy symbolem $\operatorname{Spec} \mathcal{A}$: każdy element widma posiada pewną krotność geometryczną. Analogicznie definiujemy widmo $\operatorname{Spec} A$ macierzy A. Wartości własne o krotności geometrycznej 1 nazywamy *prostymi*. Jeśli widmo składa się z n różnych wartości własnych, gdzie $n = \dim V$ (wtedy wszystkie wartości własne są proste), to całe widmo też nazywamy *prostym*.

Pamiętajmy, że do widma należą tylko te pierwiastki wielomianu charakterystycznego, które leżą w ciele skalarów $\mathfrak{K}$ (a nie w jakimś jego rozszerzeniu); dlatego widmo może być zbiorem pustym, jak w przypadku operatora obrotu na płaszczyźnie rzeczywistej. Z drugiej strony, jeśli ciało $\mathfrak{K}$ jest algebraicznie domknięte, to każdy pierwiastek charakterystyczny jest wartością własną.

LEMAT 1. *Wektory własne odpowiadające różnym wartościom własnym są liniowo niezależne. Suma $\sum_{\lambda \in \operatorname{Spec} \mathcal{A}} V^{\lambda}$ jest prosta (ale na ogół nie jest równa V).*

Dowód. Niech $\lambda_1, \ldots, \lambda_m$ będą różnymi wartościami własnymi, a $\mathbf{e}_i \in V^{\lambda_i}$ — odpowiadającymi im wektorami własnymi. Niezależność liniową wektorów $\mathbf{e}_1, \ldots, \mathbf{e}_m$ udowodnimy przez indukcję względem m. Dla $m = 1$ teza zachodzi. Załóżmy, że zachodzi ona dla $m - 1$ wektorów własnych, i przypuśćmy, że istnieje nietrywialna zależność liniowa

$$\alpha_1 \mathbf{e}_1 + \ldots + \alpha_m \mathbf{e}_m = \mathbf{0}.$$

Niech np. $\alpha_1 \neq 0$. Działamy operatorem $\mathcal{A}$ na obie strony równości; ponieważ $\mathcal{A}\mathbf{e}_i = \lambda_i \mathbf{e}_i$, więc

$$\alpha_1 \lambda_1 \mathbf{e}_1 + \ldots + \alpha_m \lambda_m \mathbf{e}_m = \mathbf{0}.$$

Mnożąc pierwszą zależność przez λ_m i odejmując drugą, otrzymamy zależność liniową $m - 1$ wektorów

$$\alpha_1(\lambda_m - \lambda_1)\mathbf{e}_1 + \ldots + \alpha_{m-1}(\lambda_m - \lambda_{m-1})\mathbf{e}_{m-1} = \mathbf{0}.$$

Na mocy założenia indukcyjnego $\alpha_i(\lambda_m - \lambda_i) = 0$ dla $i = 1, \ldots, m - 1$. Ale to nie jest prawdą dla $i = 1$, gdyż $\alpha_1 \neq 0$ i $\lambda_m \neq \lambda_1$.

Z udowodnionej liniowej niezależności wynika, że $V^{\lambda_i} \cap \sum_{j \neq i} V^{\lambda_j} = \{0\}$. Oznacza to, że $\sum_{\lambda \in \operatorname{Spec} \mathcal{A}} V^{\lambda}$ jest sumą prostą (rozdz. 1, § 2, twierdzenie 7). ∎

DEFINICJA 7. Operator liniowy $\mathcal{A}$ na przestrzeni n-wymiarowej V nazywamy *diagonalizowalnym*, jeśli istnieje baza $(\mathbf{e}_i)$, w której $\mathcal{A}$ ma macierz diagonalną

$$A = \left\| \begin{matrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{matrix} \right\|.$$

TWIERDZENIE 5. *Jeśli operator $\mathcal{A}$ ma widmo proste, to jest diagonalizowalny.*

Dowód. Z założenia operator $\mathcal{A}$ ma $n = \dim V$ różnych wartości własnych $\lambda_1, \ldots, \lambda_n$, którym odpowiadają wektory własne $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Na mocy lematu 1 wektory te są liniowo niezależne, więc $V = \langle \mathbf{e}_1, \ldots, \mathbf{e}_n \rangle$, a ponieważ $\mathcal{A}\mathbf{e}_i = \lambda_i \mathbf{e}_i$, zatem $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. ∎

Prostota widma jest warunkiem dostatecznym diagonalizowalności, ale w żadnym razie nie koniecznym; np. każdy operator rzutowania jest diagonalizowalny (zob. (2)), chociaż jego widmo dla $n > 2$ nie jest proste. Warunek konieczny i dostateczny podaje

TWIERDZENIE 6. *Operator liniowy $\mathcal{A}$ na skończenie wymiarowej przestrzeni V nad ciałem $\mathfrak{K}$ jest diagonalizowalny wtedy i tylko wtedy, gdy spełnione są następujące warunki:*

(i) *wielomian charakterystyczny $\chi_{\mathcal{A}}(t)$ rozkłada się na czynniki liniowe nad ciałem $\mathfrak{K}$;*
(ii) *krotność geometryczna każdej wartości własnej jest równa jej krotności algebraicznej.*

Dowód. Przypuśćmy, że warunki (i) i (ii) są spełnione. Jeśli $\lambda_1, \ldots, \lambda_m$ są wszystkimi różnymi pierwiastkami wielomianu $\chi_{\mathcal{A}}(t)$ o krotnościach (geometrycznych i algebraicznych) $k_1, \ldots, k_m$, to

$$\dim V^{\lambda_i} = k_i, \qquad k_1 + \ldots + k_m = n. \tag{11}$$

Na mocy lematu 1 suma $V^{\lambda_1} + \ldots + V^{\lambda_m}$ jest prosta, a z porównania wymiarów (zob. (11)) wynika, że

$$V = V^{\lambda_1} + \ldots + V^{\lambda_m}. \tag{12}$$

Wystarczy teraz rozpatrzyć bazę w V będącą sumą baz w przestrzeniach V^{λ_i}. Baza ta składa się z wektorów własnych operatora $\mathcal{A}$; istnienie takiej bazy jest równoważne diagonalizowalności operatora.

Na odwrót, załóżmy, że operator $\mathcal{A}$ jest diagonalizowalny. Możemy założyć, że jego macierz w pewnej bazie ma postać

$$A = \mathrm{diag}(\underbrace{\lambda_1, \ldots, \lambda_1}_{l_1}; \ldots; \underbrace{\lambda_m, \ldots, \lambda_m}_{l_m}), \qquad l_1 + \ldots + l_m = n, \tag{13}$$

dla pewnych różnych skalarów $\lambda_i \in \mathfrak{K}$. Wtedy

$$\chi_A(t) = \det(tE - A) = (t - \lambda_1)^{l_1} \ldots (t - \lambda_m)^{l_m},$$

a więc warunek (i) jest spełniony. Z drugiej strony wiemy już, że suma $V^{\lambda_1} + \ldots + V^{\lambda_m}$ jest prosta (lemat 1), a ponadto równa V, gdyż skoro V ma bazę złożoną z wektorów własnych, to $V^{\lambda_1}, \ldots, V^{\lambda_m}$ generują V. Niech $k_i = \dim V^{\lambda_i}$. Mamy więc $k_1 + \ldots + k_m = n$, $l_1 + \ldots + l_m = n$ oraz $k_i \leqslant l_i$ dla każdego i (lemat 4). Jest to możliwe tylko wtedy, gdy $k_i = l_i$ dla każdego i, czyli warunek (ii) jest spełniony (¹).

5. Istnienie podprzestrzeni niezmienniczych. Wszystkie ogólne rozważania dotyczące podprzestrzeni niezmienniczych, wartości własnych i wektorów własnych są zasadniczo słuszne dla dowolnego ciała skalarów. Istnienie tych obiektów zależy jednak w sposób decydujący od tego ciała, jak się można przekonać na przykładzie najważniejszych ciał $\mathbb{R}$ i $\mathbb{C}$.

TWIERDZENIE 7. *Każdy operator liniowy $\mathcal{A}$ na zespolonej (odpowiednio rzeczywistej) przestrzeni liniowej V ma jednowymiarową (odpowiednio jedno- lub dwuwymiarową) podprzestrzeń niezmienniczą.*

Dowód. W przypadku zespolonym wielomian $\chi_{\mathcal{A}}(t)$ ma pierwiastki zespolone, a więc operator $\mathcal{A}$ ma wartości własne; każdy wektor własny generuje jednowymiarową podprzestrzeń niezmienniczą.

W przypadku ciała $\mathbb{R}$ rozważmy wielomian minimalny $\mu_{\mathcal{A}}(t) \in \mathbb{R}[t]$ operatora $\mathcal{A}$ (§ 2, definicja 2). Jeśli $\mu_{\mathcal{A}}(t)$ ma pierwiastek rzeczywisty α, to

$$\mu_{\mathcal{A}}(t) = (t - \alpha)g(t) \quad \text{dla pewnego } g(t) \in \mathbb{R}[t].$$

Ponieważ $g(\mathcal{A}) \neq \mathcal{O}$ (na mocy minimalności $\mu_{\mathcal{A}}(t)$), więc $\mathbf{v} := g(\mathcal{A})\mathbf{u} \neq \mathbf{0}$ dla pewnego $\mathbf{u} \in V$. Ale

$$(\mathcal{A} - \alpha\mathcal{E})\mathbf{v} = (\mathcal{A} - \alpha\mathcal{E})g(\mathcal{A})\mathbf{u} = \mu_{\mathcal{A}}(\mathcal{A})\mathbf{u} = \mathbf{0},$$

czyli $\mathcal{A}\mathbf{v} = \alpha\mathbf{v}$, tj. $\mathbf{v}$ jest wektorem własnym.

Załóżmy teraz, że $\mathcal{A}$ nie ma rzeczywistych wartości własnych; z powyższego rozumowania wynika, że wtedy $\mu_{\mathcal{A}}(t)$ nie ma pierwiastków rzeczywistych. Wówczas na mocy twierdzenia o rozkładzie wielomianów rzeczywistych na czynniki nieprzywiedlne (część I, rozdz. 6, § 4, twierdzenie 1)

$$\mu_{\mathcal{A}}(t) = (t^2 - \alpha t - \beta)h(t)$$

dla pewnych $\alpha, \beta \in \mathbb{R}$ i $h(t) \in \mathbb{R}[t]$. Znowu $\mathbf{v} := h(\mathcal{A})\mathbf{u} \neq \mathbf{0}$ dla pewnego $\mathbf{u} \in V$ oraz

$$\mathcal{A}^2\mathbf{v} - \alpha\mathcal{A}\mathbf{v} - \beta\mathbf{v} = \mu_{\mathcal{A}}(\mathcal{A})\mathbf{u} = \mathbf{0}.$$

Stąd $\mathcal{A}^2\mathbf{v} = \alpha\mathcal{A}\mathbf{v} + \beta\mathbf{v}$, a więc $L = \langle \mathbf{v}, \mathcal{A}\mathbf{v} \rangle$ jest podprzestrzenią $\mathcal{A}$-niezmienniczą; ponadto $\dim L = 2$, ponieważ $\mathcal{A}\mathbf{v} \neq \lambda\mathbf{v}$ ($\mathcal{A}$ nie ma wektorów własnych). ■

6. Operator sprzężony. Zbadajmy teraz związki operatorów z przestrzenią dualną. Niech V będzie przestrzenią liniową nad $\mathfrak{K}$, V^* — przestrzenią dualną,

---

(¹) Warunek (ii) można też łatwo udowodnić we współrzędnych: jeśli $X = [x_1, \ldots, x_n]$, to z (13) wynika, że np. $AX = \lambda_1 X \Leftrightarrow x_{l_1+1} = \ldots = x_n = 0$, czyli $\dim V^{\lambda_1} = l_1$ (*przyp. tłum.*).

a $\mathcal{A}$ — operatorem liniowym na V. Dla każdej ustalonej formy liniowej $f \in V^*$ odwzorowanie $\mathbf{x} \mapsto (f, \mathcal{A}\mathbf{x}) = f(\mathcal{A}\mathbf{x})$ (przy oznaczeniach z rozdziału 1, § 3, p. 2) jest znowu elementem przestrzeni V^*, czyli formą liniową:

$$(f, \mathcal{A}(\alpha\mathbf{x} + \beta\mathbf{y})) = (f, \alpha\mathcal{A}\mathbf{x} + \beta\mathcal{A}\mathbf{y}) = \alpha(f, \mathcal{A}\mathbf{x}) + \beta(f, \mathcal{A}\mathbf{y}).$$

Oznaczmy tę formę liniową przez $\mathcal{A}^*f$; tak więc

$$(\mathcal{A}^*f, \mathbf{x}) := (f, \mathcal{A}\mathbf{x}). \tag{14}$$

Ponieważ f jest dowolne, otrzymujemy w ten sposób odwzorowanie $\mathcal{A}^* : V^* \to V^*$, $f \mapsto \mathcal{A}^*f$, które jest liniowe:

$$\begin{aligned}(\mathcal{A}^*(\alpha f + \beta g), \mathbf{x}) &= (\alpha f + \beta g, \mathcal{A}\mathbf{x}) = \alpha(f, \mathcal{A}\mathbf{x}) + \beta(g, \mathcal{A}\mathbf{x}) \\ &= \alpha(\mathcal{A}^*f, \mathbf{x}) + \beta(\mathcal{A}^*g, \mathbf{x}) = (\alpha\mathcal{A}^*f + \beta\mathcal{A}^*g, \mathbf{x}),\end{aligned}$$

a więc $\mathcal{A}^* \in \mathcal{L}(V^*)$.

DEFINICJA 8. Operator liniowy $\mathcal{A}^* \in \mathcal{L}(V^*)$ określony wzorem (14) nazywamy *operatorem sprzężonym* do $\mathcal{A} \in \mathcal{L}(V)$.

Otrzymujemy w ten sposób odwzorowanie $\mathcal{L}(V) \to \mathcal{L}(V^*)$, $\mathcal{A} \mapsto \mathcal{A}^*$. Wprost z definicji wynikają następujące własności tego odwzorowania:

$$\begin{aligned}&\mathcal{O}_V^* = \mathcal{O}_{V^*}, \quad \mathcal{E}_V^* = \mathcal{E}_{V^*}, \quad (\alpha\mathcal{A})^* = \alpha\mathcal{A}^*, \\ &(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*, \quad (\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*.\end{aligned} \tag{15}$$

Na przykład ostatnią równość w (15) uzasadnia się tak:

$$((\mathcal{A}\mathcal{B})^*f, \mathbf{x}) = (f, (\mathcal{A}\mathcal{B})\mathbf{x}) = (f, \mathcal{A}(\mathcal{B}\mathbf{x})) = (\mathcal{A}^*f, \mathcal{B}\mathbf{x}) = (\mathcal{B}^*\mathcal{A}^*f, \mathbf{x}).$$

Aby znaleźć postać macierzową operatora $\mathcal{A}^*$, wygodnie jest wybrać w V i V^* bazy dualne $(\mathbf{e}_i)$ i (e^i). Jeśli $\mathcal{A}\mathbf{e}_j = \sum_{k=1}^n a_{kj}\mathbf{e}_k$, to

$$(e^i, \mathcal{A}\mathbf{e}_j) = \sum_{k=1}^n a_{kj}(e^i, \mathbf{e}_k) = \sum_{k=1}^n a_{kj}\delta_{ik} = a_{ij}.$$

Jeśli teraz przyjmiemy

$$\mathcal{A}^*e^i = \sum_{k=1}^n a_{ki}^* e^k,$$

to $(\mathcal{A}^*e^i, \mathbf{e}_j) = \sum_{k=1}^n a_{ki}^*(e^k, \mathbf{e}_j) = a_{ji}^*$. Wobec tego $a_{ji}^* = a_{ij}$. Udowodniliśmy

TWIERDZENIE 8. *Jeśli operator $\mathcal{A}$ ma w pewnej bazie macierz A, to operator $\mathcal{A}^*$ ma w bazie dualnej macierz $A^* = {}^tA$.* ∎

Zauważmy, że jeśli utożsamimy przestrzenie V i V^{**} za pomocą odwzorowania $\mathbf{x} \mapsto \varepsilon_{\mathbf{x}}$, gdzie $\varepsilon_{\mathbf{x}}(f) = (f, \mathbf{x})$ (rozdz. 1, § 3, twierdzenie 2), to dla każdego operatora $\mathcal{A} \in \mathcal{L}(V)$ zachodzi równość

$$\mathcal{A}^{**} = \mathcal{A}. \tag{16}$$

Równość tę należy rozumieć w ten sposób, że jeśli $\mathcal{A}\mathbf{x} = \mathbf{y}$, to $\mathcal{A}^{**}\varepsilon_{\mathbf{x}} = \varepsilon_{\mathbf{y}}$. Istotnie, dla każdego $f \in V^*$ mamy

$$(\mathcal{A}^{**}\varepsilon_{\mathbf{x}}, f) = (\varepsilon_{\mathbf{x}}, \mathcal{A}^* f) = (\mathcal{A}^* f, \mathbf{x}) = (f, \mathcal{A}\mathbf{x}) = (f, \mathbf{y}) = (\varepsilon_{\mathbf{y}}, f).$$

Wynika stąd, że przekształcenie liniowe $\mathcal{A} \mapsto \mathcal{A}^*$ z $\mathcal{L}(V)$ do $\mathcal{L}(V^*)$ o własnościach (15) jest wzajemnie jednoznaczne; ze względu na ostatnią równość w (15) mówimy, że przekształcenie to jest *antyizomorfizmem algebr* $\mathcal{L}(V)$ i $\mathcal{L}(V^*)$.

Jednoczesne rozpatrywanie par $(V, \mathcal{A})$ i $(V^*, \mathcal{A}^*)$ często daje praktyczne korzyści. Jako przykład niech posłuży dowód następującego faktu.

TWIERDZENIE 9. *Każdy operator liniowy na przestrzeni zespolonej ma hiperpłaszczyznę niezmienniczą.*

Dowód. Niech $\dim V = n$ i $\mathcal{A} \in \mathcal{L}(V)$. Wiemy, że $\dim \operatorname{Ker} f = n - 1$ dla dowolnej formy liniowej $f \neq 0$ na V. Weźmy teraz jako f dowolny wektor własny operatora $\mathcal{A}^*$ na V^*. Wektor taki istnieje na mocy twierdzenia 7; jeśli λ jest odpowiadającą mu wartością własną, to na mocy (14) mamy

$$\mathbf{x} \in \operatorname{Ker} f \quad \Rightarrow \quad (f, \mathcal{A}\mathbf{x}) = (\mathcal{A}^* f, \mathbf{x}) = (\lambda f, \mathbf{x}) = \lambda(f, \mathbf{x}) = 0 \quad \Rightarrow \quad \mathcal{A}\mathbf{x} \in \operatorname{Ker} f,$$

co oznacza, że $\operatorname{Ker} f$ jest szukaną hiperpłaszczyzną. ■

7. Operator ilorazowy. Niech $L \subset V$ będzie podprzestrzenią niezmienniczą dla operatora $\mathcal{A} \in \mathcal{L}(V)$. Uważając V i L za ustalone, będziemy oznaczać przestrzeń ilorazową V/L (określoną w rozdziale 1, § 2, p. 6) symbolem $\overline{V}$, a jej element $\mathbf{x} + L$ przez $\overline{\mathbf{x}}$.

DEFINICJA 9. Wzór

$$\overline{\mathcal{A}}\overline{\mathbf{x}} = \overline{\mathcal{A}\mathbf{x}}, \qquad \overline{\mathbf{x}} \in \overline{V},$$

określa operator liniowy $\overline{\mathcal{A}} : \overline{V} \to \overline{V}$, zwany *operatorem ilorazowym*. Inaczej mówiąc, $\overline{\mathcal{A}}(\mathbf{x} + L) = \mathcal{A}\mathbf{x} + L$.

Powyższa definicja jest poprawna, tzn. nie zależy od wyboru reprezentanta $\mathbf{x}$: jeśli $\mathbf{x} + L = \mathbf{x}' + L$, to $\mathbf{x} - \mathbf{x}' = \mathbf{y} \in L$ i $\mathcal{A}\mathbf{x} - \mathcal{A}\mathbf{x}' = \mathcal{A}(\mathbf{x} - \mathbf{x}') = \mathcal{A}\mathbf{y} \in L$ (na mocy niezmienniczości L). Stąd $\mathcal{A}\mathbf{x} + L = \mathcal{A}\mathbf{x}' + L$. Gdyby podprzestrzeń L nie była $\mathcal{A}$-niezmiennicza, definicja $\overline{\mathcal{A}}$ nie miałaby sensu.

Załóżmy teraz, że $V = L \oplus M$ jest sumą prostą podprzestrzeni $\mathcal{A}$-niezmienniczych. Jak wiemy, wtedy $\mathcal{A} = \mathcal{A}_L \dotplus \mathcal{A}_M$ jest sumą prostą swoich ograniczeń do L i M. Jeśli $f : \mathbf{u} \mapsto \mathbf{u} + L$ jest izomorfizmem między M i $\overline{V} = V/L$ (rozdz. 1, § 2, twierdzenie 10), to

$$(f \circ \mathcal{A}_M)\mathbf{y} = f(\mathcal{A}_M\mathbf{y}) = \mathcal{A}_M\mathbf{y} + L = \overline{\mathcal{A}}(\mathbf{y} + L) = \overline{\mathcal{A}}(f\mathbf{y}), \qquad \mathbf{y} \in M,$$

skąd
$$f \circ \mathcal{A}_M = \overline{\mathcal{A}} \circ f. \tag{17}$$
Oznacza to, że z dokładnością do izomorfizmu f działanie $\overline{\mathcal{A}}$ na przestrzeni $\overline{V}$ pokrywa się z działaniem $\mathcal{A}_M$ na M. Mówimy też w takim wypadku o *równoważności* operatorów $\overline{\mathcal{A}}$ i $\mathcal{A}_M$.

Przykład. Jeśli $\mathbf{e}_1$ jest wektorem własnym operatora $\mathcal{B}$ w przestrzeni dwuwymiarowej $V = \langle \mathbf{e}_1, \mathbf{e}_2 \rangle$, to macierz tego operatora w bazie $(\mathbf{e}_1, \mathbf{e}_2)$ jest macierzą trójkątną $\left\| \begin{smallmatrix} \lambda & \sigma \\ 0 & \mu \end{smallmatrix} \right\|$. Na przestrzeni $\overline{V} = V/\langle \mathbf{e}_1 \rangle = \langle \overline{\mathbf{e}}_2 \rangle$ mamy $\overline{\mathcal{A}}\overline{\mathbf{e}}_2 = \mu \overline{\mathbf{e}}_2$.

Jeśli $V = \langle \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \rangle$ jest trójwymiarową przestrzenią liniową nad $\mathbb{C}$, przy czym $\mathbf{e}_1$ jest wektorem własnym operatora $\mathcal{A} : V \to V$ (czyli $\mathcal{A}\mathbf{e}_1 = \alpha \mathbf{e}_1$ dla pewnego $\alpha \in \mathbb{C}$), to w dwuwymiarowej przestrzeni $\overline{V} = V/\langle \mathbf{e}_1 \rangle = \langle \overline{\mathbf{e}}_2, \overline{\mathbf{e}}_3 \rangle$ działa operator ilorazowy, który (w odpowiedniej bazie) ma macierz trójkątną $\left\| \begin{smallmatrix} \beta & \delta \\ 0 & \gamma \end{smallmatrix} \right\|$. Przypuśćmy dla uproszczenia, że jest to baza $(\overline{\mathbf{e}}_2, \overline{\mathbf{e}}_3)$. Wtedy
$$\overline{\mathcal{A}}\overline{\mathbf{e}}_2 = \beta \overline{\mathbf{e}}_2, \quad \text{tj.} \quad \mathcal{A}\mathbf{e}_2 = \beta \mathbf{e}_2 + \nu \mathbf{e}_1,$$
$$\overline{\mathcal{A}}\overline{\mathbf{e}}_3 = \gamma \overline{\mathbf{e}}_3 + \delta \overline{\mathbf{e}}_2, \quad \text{tj.} \quad \mathcal{A}\mathbf{e}_3 = \gamma \mathbf{e}_3 + \delta \mathbf{e}_2 + \mu \mathbf{e}_1.$$
Tak więc
$$A = \begin{Vmatrix} \alpha & \nu & \mu \\ 0 & \beta & \delta \\ 0 & 0 & \gamma \end{Vmatrix}.$$
Dla przestrzeni zespolonej większego wymiaru rozumowanie jest analogiczne.

ĆWICZENIA

1. Niech $\{A_i \mid i = 1, \ldots, m-1\}$ będzie układem ortogonalnym macierzy idempotentnych (zob. uwagę w końcu p. 1). Wykazać, że jeśli $A = A_1 + \ldots + A_{m-1}$, to $A^2 = A$ oraz $AA_i = A_iA = A_i$ dla $i = 1, \ldots, m-1$; jeśli ponadto przyjmiemy $A_m = E - A$, to układ idempotentów $\{A_i \mid i = 1, \ldots, m\}$ jest zupełny.

2. Udowodnić, że jeśli operator liniowy $\mathcal{D} : M_n(\mathfrak{K}) \to M_n(\mathfrak{K})$ na przestrzeni macierzy kwadratowych jest multiplikatywny:
$$\mathcal{D}(AB) = \mathcal{D}(A)\mathcal{D}(B) \quad \text{dla dowolnych } A, B \in M_n(\mathfrak{K})$$
oraz $\mathcal{D} \neq \mathcal{O}$, to $\mathcal{D} = f_C$ dla pewnej macierzy nieosobliwej C, gdzie $f_C(A) = C^{-1}AC$ (§ 1, ćwiczenie 3).

3. Niech $\mathcal{A} : V \to V$ będzie operatorem liniowym o tej własności, że $\operatorname{Im} \mathcal{A}^p = \operatorname{Im} \mathcal{A}^{p+1}$ dla pewnej liczby naturalnej p. Wykazać, że wtedy $V = \operatorname{Ker} \mathcal{A}^p \oplus \operatorname{Im} \mathcal{A}^p$ jest sumą prostą dwóch podprzestrzeni $\mathcal{A}$-niezmienniczych.

4. Udowodnić, że jeśli operatory $\mathcal{E}, \mathcal{A}, \mathcal{A}^2, \ldots, \mathcal{A}^{n-1}$ działające na n-wymiarowej przestrzeni V są liniowo niezależne, to istnieje taki wektor $\mathbf{v} \in V$, że

$$V = \langle \mathbf{v}, \mathcal{A}\mathbf{v}, \mathcal{A}^2\mathbf{v}, \ldots, \mathcal{A}^{n-1}\mathbf{v} \rangle$$

(mówimy w tym przypadku, że przestrzeń V jest *cykliczna*).

5. Niech A będzie rzeczywistą macierzą $n \times n$ bez rzeczywistych wartości własnych; w szczególności, n jest parzyste i macierz A jest odwracalna. Wykazać, że istnieje taka macierz rzeczywista B, że $AB = BA$ i $B^2 = -E$ (D. Djoković).

6. Udowodnić, że dla dowolnych $A, B \in M_n(\mathbb{R})$ macierze AB i BA mają ten sam wielomian charakterystyczny.

7. Znaleźć pierwiastki charakterystyczne macierzy „cyklicznej"

$$A = \begin{Vmatrix} a_0 & a_1 & a_2 \\ a_2 & a_0 & a_1 \\ a_1 & a_2 & a_0 \end{Vmatrix},$$

wykorzystując łatwy do sprawdzenia związek

$$A = a_0 E + a_1 B + a_2 B^2, \qquad B = \begin{Vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{Vmatrix}.$$

8. Udowodnić, że przestrzeń $\mathrm{SMag}_n(\mathbb{Q})$ kwadratów półmagicznych (rozdz. 1, § 1, przykład 8) jest podalgebrą w $M_n(\mathbb{Q})$.

9. (Amer. Math. Monthly 98 (1991), 131–133). Niech $A, B \in M_n(\mathfrak{K})$, gdzie $\operatorname{char} \mathfrak{K} \neq 2$. Piszemy $B \sim A$, jeśli $B = DA$ dla pewnej macierzy $D = \operatorname{diag}(\theta_1, \ldots, \theta_n)$, gdzie $\theta_i = \pm 1$; jest to oczywiście relacja równoważności. Jeśli $S(A)$ jest klasą równoważności macierzy A, to $\operatorname{Card} S(A) \leqslant 2^n$. Udowodnić, że co najmniej jedna macierz w każdej klasie $S(A)$ nie ma wartości własnej 1.

10. Niech $\mathcal{A}$ będzie operatorem liniowym na n-wymiarowej przestrzeni V. Wykazać, że

$$\mathcal{A}^2 = \mathcal{A} \Leftrightarrow \operatorname{rank} \mathcal{A} + \operatorname{rank}(\mathcal{E} - \mathcal{A}) = n.$$

§ 4. POSTAĆ KANONICZNA JORDANA

Gdy chcemy zbadać działanie danego operatora liniowego $\mathcal{A} : V \to V$, naturalne jest poszukiwanie takiej bazy w V, która byłaby możliwie najbardziej dopasowana do $\mathcal{A}$. Inaczej mówiąc, w klasie macierzy podobnych $C^{-1}AC$ odpowiadających operatorowi $\mathcal{A}$ należy znaleźć macierz możliwie najprostszej postaci. Ze zrozumiałych względów rozwiązanie tego zadania w istotny sposób zależy od ciała skalarów $\mathfrak{K}$. W dalszym ciągu przyjmujemy, że $\mathfrak{K}$ jest ciałem $\mathbb{C}$ liczb zespolonych, choć w zasadzie mogłoby to być dowolne ciało algebraicznie domknięte.

1. Twierdzenie Hamiltona–Cayleya. Następujący prosty fakt okazuje się bardzo użyteczny:

TWIERDZENIE 1. *Macierz operatora liniowego $\mathcal{A}$ można zawsze sprowadzić (w sensie podobieństwa) do postaci trójkątnej.*

Dowód. Najprościej udowodnić to przez indukcję. Zgodnie z twierdzeniem 9 z § 3 przestrzeń V zawiera hiperpłaszczyznę $\mathcal{A}$-niezmienniczą U. Na mocy założenia indukcyjnego w U istnieje baza $(\mathbf{e}_1, \dots, \mathbf{e}_{n-1})$ taka, że $\mathcal{A}\mathbf{e}_i = \lambda_i \mathbf{e}_i + \mathbf{v}_i$, przy czym $\mathbf{v}_i \in \langle \mathbf{e}_1, \dots, \mathbf{e}_{i-1} \rangle$. Mamy $V = \langle U, \mathbf{e}_n \rangle$, gdzie $\mathbf{e}_n$ jest dowolnym wektorem z $V \setminus U$. Niech $\mathcal{A}\mathbf{e}_n = \lambda_n \mathbf{e}_n + \mathbf{u}$, gdzie $\mathbf{u} \in U$. Wtedy w bazie $(\mathbf{e}_1, \dots, \mathbf{e}_{n-1}, \mathbf{e}_n)$ macierz operatora A ma żądaną postać

$$A = \left\| \begin{matrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{matrix} \right\| . \quad \blacksquare \tag{1}$$

Dość łatwo można teraz udowodnić ważne

TWIERDZENIE 2 (Hamiltona–Cayleya). *Każdy operator liniowy $\mathcal{A}$ jest pierwiastkiem swojego wielomianu charakterystycznego:*

$$\chi_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}.$$

Dowód. Wykorzystamy bazę $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ skonstruowaną w dowodzie twierdzenia 1. Jeśli oznaczymy $V_k = \langle \mathbf{e}_1, \dots, \mathbf{e}_{n-k} \rangle$, to

$$V = V_0 \supset V_1 \supset \ldots \supset V_{n-1} \supset V_n = \{\mathbf{0}\},$$

każda podprzestrzeń V_i jest $\mathcal{A}$-niezmiennicza oraz

$$(\mathcal{A} - \lambda_{n-k}\mathcal{E})\mathbf{e}_{n-k} \in V_{k+1}.$$

Wobec tego

$$(\mathcal{A} - \lambda_{n-k}\mathcal{E})V_k \subset V_{k+1},$$

a stąd

$$\begin{aligned} \chi_{\mathcal{A}}(\mathcal{A})V &= (\mathcal{A} - \lambda_1\mathcal{E}) \cdots (\mathcal{A} - \lambda_n\mathcal{E})V \\ &= (\mathcal{A} - \lambda_1\mathcal{E}) \ldots (\mathcal{A} - \lambda_n\mathcal{E})V_0 \subset (\mathcal{A} - \lambda_1\mathcal{E}) \ldots (\mathcal{A} - \lambda_{n-1}\mathcal{E})V_1 \\ &\subset (\mathcal{A} - \lambda_1\mathcal{E}) \ldots (\mathcal{A} - \lambda_{n-2}\mathcal{E})V_2 \subset \ldots \subset (\mathcal{A} - \lambda_1\mathcal{E})V_{n-1} = \{\mathbf{0}\}. \end{aligned}$$

Zatem obraz operatora $\chi_{\mathcal{A}}(\mathcal{A})$ jest podprzestrzenią zerową, czyli $\chi_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$. ■

Wniosek. *Wielomian minimalny $\mu_{\mathcal{A}}(t)$ operatora $\mathcal{A}$ jest dzielnikiem wielomianu charakterystycznego $\chi_{\mathcal{A}}(t)$ i dzieli się przez wszystkie czynniki liniowe postaci $t - \lambda$, gdzie $\lambda \in \operatorname{Spec} \mathcal{A}$.*

Dowód. Ponieważ operator $\mathcal{A}$ jest pierwiastkiem wielomianu $\chi_{\mathcal{A}}(t)$ (twierdzenie 2), więc podzielność $\chi_{\mathcal{A}}(t)$ przez $\mu_{\mathcal{A}}(t)$ wynika z twierdzenia 2 w § 2. Jeśli ponadto λ jest wartością własną operatora $\mathcal{A}$, to $\mathcal{A}\mathbf{v} = \lambda\mathbf{v}$ dla pewnego $\mathbf{v} \neq \mathbf{0}$, a więc $\mathbf{0} = \mu_{\mathcal{A}}(\mathcal{A})\mathbf{v} = \mu_{\mathcal{A}}(\lambda)\mathbf{v}$, czyli $\mu_{\mathcal{A}}(\lambda) = 0$, skąd $(t - \lambda) \mid \mu_{\mathcal{A}}(t)$ (powtórzyliśmy tu uzasadnienie implikacji (7) w § 3). ∎

Uwaga. Można by dowodzić twierdzenia Hamiltona-Cayleya tak: $\det(tE - A)|_{t=A} = \det(AE - A) = \det 0 = 0$, co kończy dowód. Dlaczego rozumowanie to jest zupełnie błędne?

Choć twierdzenie Hamiltona-Cayleya ma rozliczne zastosowania, będziemy je na razie wykorzystywać tylko w najbardziej bezpośredniej formie.

Przykład 1. Niech $\mathcal{A} : V \to V$ będzie operatorem nilpotentnym stopnia m (§ 2, definicja 3); tak więc $\mu_{\mathcal{A}}(t) = t^m$. Przypuśćmy, że $\mathcal{A}^{m-1}\mathbf{v} \neq \mathbf{0}$. Wtedy wektory $\mathbf{v}, \mathcal{A}\mathbf{v}, \ldots, \mathcal{A}^{m-1}\mathbf{v}$ są liniowo niezależne. Istotnie, każdą nietrywialną zależność liniową między nimi można sprowadzić do postaci

$$\mathcal{A}^k\mathbf{v} + \alpha_1\mathcal{A}^{k+1}\mathbf{v} + \ldots + \alpha_{m-1-k}\mathcal{A}^{m-1}\mathbf{v} = \mathbf{0} \quad \text{dla pewnego } k \ (0 \leqslant k \leqslant m-1).$$

Zastosowanie operatora $\mathcal{A}^{m-1-k}$ do obu stron tej równości prowadzi do wniosku, że $\mathcal{A}^{m-1}\mathbf{v} = \mathbf{0}$, wbrew wyborowi $\mathbf{v}$.

Tak więc stopień nilpotentności m operatora $\mathcal{A}$ nie przekracza $n = \dim V$, co ma się rozumieć wynika też z twierdzenia Hamiltona-Cayleya. Załóżmy teraz, że $m = n$ i $\mathcal{A}^{n-1}\mathbf{e} \neq \mathbf{0}$. Wyżej wykazaliśmy, że wtedy następujące wektory tworzą bazę w V:

$$\mathbf{e}_1 = \mathcal{A}^{n-1}\mathbf{e},\ \mathbf{e}_2 = \mathcal{A}^{n-2}\mathbf{e},\ \ldots,\ \mathbf{e}_{n-1} = \mathcal{A}\mathbf{e},\ \mathbf{e}_n = \mathbf{e}.$$

Wówczas $\mathcal{A}\mathbf{e}_1 = \mathbf{0}$, $\mathcal{A}\mathbf{e}_k = \mathbf{e}_{k-1}$ dla $k > 1$ i macierzą operatora $\mathcal{A}$ w bazie $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ jest klatka Jordana $J_n(0)$, zdefiniowana poniżej.

Jeśli np. $V = \langle 1, t, \ldots, t^{n-1}\rangle$ jest przestrzenią wielomianów zespolonych stopni $< n$ i $\mathcal{D} = d/dt$ jest operatorem różniczkowania, to jego macierzą w bazie $(\mathbf{e}_i)$, gdzie $\mathbf{e}_{i+1} = (1/i!)t^i$ dla $i = 0, \ldots, n-1$, jest właśnie klatka Jordana $J_n(0)$.

Definicja 1. (a) (*Górną*) *klatką Jordana* (na cześć matematyka francuskiego Camille'a Jordana (1838–1922)) wymiarów $m \times m$ odpowiadającą wartości własnej λ nazywamy macierz

$$J_m(\lambda) = \begin{Vmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{Vmatrix}.$$

(b) *Macierz Jordana* to macierz zbudowana z bloków $J_{m_i}(\lambda_i)$ wzdłuż głównej przekątnej i zer poza tymi blokami:

$$J = \begin{Vmatrix} J_{m_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{m_s}(\lambda_s) \end{Vmatrix}. \qquad (2)$$

(c) *Bazą Jordana* dla operatora liniowego $\mathcal{A} : V \to V$ nazywamy taką bazę przestrzeni V, w której macierz operatora $\mathcal{A}$ jest macierzą Jordana; inaczej mówimy, że w takiej bazie macierz operatora ma *postać kanoniczną Jordana* (w skrócie PKJ) $J(\mathcal{A})$.

(d) *Sprowadzenie macierzy do postaci kanonicznej Jordana* to rozwiązanie równania macierzowego $X^{-1}AX = J(A)$, gdzie X jest (nieznaną) macierzą nieosobliwą, a $J(A)$ — (również nieznaną) macierzą Jordana.

Zauważmy, że $J_m(\lambda) - \lambda E = J_m(0)$ jest macierzą nilpotentną. W szczególności $(t-\lambda)^m$ jest wielomianem minimalnym klatki Jordana $J_m(\lambda)$ i λ jest jej jedyną wartością własną: $\operatorname{Spec} J_m(\lambda) = \{\lambda\}$.

Przykład 2. Niech $D_n(\lambda)$ będzie przestrzenią liniową funkcji zespolonych postaci $e^{\lambda t}f(t)$, gdzie $\lambda \in \mathbb{C}$ i $f(t)$ przebiega wielomiany stopni $\leqslant n-1$. Ponieważ

$$\frac{d}{dt}(e^{\lambda t}f(t)) = e^{\lambda t}(\lambda f(t) + f'(t)),$$

więc różniczkowanie $\mathcal{D} = d/dt$ wyznacza operator liniowy na $D_n(\lambda)$. Jeśli przyjmiemy $\mathrm{e}_{i+1} = e^{\lambda t}t^i/i!$ dla $i = 0, \dots, n-1$, to

$$\mathcal{D}\mathrm{e}_{i+1} = \frac{t^{i-1}}{(i-1)!}e^{\lambda t} + \lambda\frac{t^i}{i!}e^{\lambda t} = \mathrm{e}_i + \lambda\mathrm{e}_{i+1}$$

($0! = 1$; dla $i = 0$ nie ma pierwszego składnika). Zatem (e_i) jest bazą Jordana dla operatora $\mathcal{D}$ w $D_n(\lambda)$ oraz $J(\mathcal{D}) = J_n(\lambda)$.

Przykład ten wskazuje na szczególną rolę macierzy Jordana w teorii liniowych równań różniczkowych. Do tego tematu jeszcze wrócimy.

Przykład 3. Jeśli $f(t)$ jest dowolnym wielomianem, to

$$f(J_m(\lambda)) = \begin{Vmatrix} f(\lambda) & f'(\lambda)/1! & f''(\lambda)/2! & \dots & f^{(m-1)}(\lambda)/(m-1)! \\ 0 & f(\lambda) & f'(\lambda)/1! & \dots & f^{(m-2)}(\lambda)/(m-2)! \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & f(\lambda) \end{Vmatrix},$$

a więc klatkami Jordana znacznie łatwiej operować niż dowolnymi macierzami.

2. Postać kanoniczna Jordana: twierdzenie i wnioski

TWIERDZENIE (podstawowe o postaci kanonicznej Jordana). *Każdą macierz kwadratową A stopnia n nad ciałem algebraicznie domkniętym $\mathfrak{K}$ (w szczególności nad $\mathbb{C}$) można sprowadzić do postaci kanonicznej Jordana. Istnieje macierz nieosobliwa C, dla której $C^{-1}AC = J(A) = J$ jest macierzą postaci (2), wyznaczoną jednoznacznie z dokładnością do porządku klatek.*

Ponieważ macierze podobne mają te same wielomiany minimalne, więc z tego twierdzenia i z uwag dotyczących klatek Jordana wynika, że

$$\mu_A(t) = (t - \lambda_{i_1})^{m_{i_1}} \dots (t - \lambda_{i_p})^{m_{i_p}}, \tag{3}$$

gdzie $\lambda_{i_1}, \dots, \lambda_{i_p}$ są wszystkimi różnymi wartościami własnymi macierzy A oraz m_{i_k} jest największym stopniem klatki Jordana odpowiadającej wartości własnej λ_{i_k} i występującej w $J(A)$.

Jasne jest, że macierz A jest diagonalizowalna (tzn. podobna do macierzy diagonalnej $\mathrm{diag}(\lambda_1, \dots, \lambda_n)$) wtedy i tylko wtedy, gdy w $J(A)$ nie ma klatek stopnia > 1. Jeśli uwzględnimy wzór (3), prowadzi to do następującego wygodnego kryterium:

WNIOSEK. *Zespolona macierz kwadratowa jest diagonalizowalna wtedy i tylko wtedy, gdy jej wielomian minimalny nie ma pierwiastków wielokrotnych.*

Kryterium to jest efektywne, ponieważ do znalezienia $\mu_A(t)$ nie ma potrzeby sprowadzania macierzy A do postaci kanonicznej Jordana.

Dowód podstawowego twierdzenia dzieli się na trzy części, przedstawione w punktach 3–5 poniżej. Po drodze podajemy szereg praktycznych zaleceń, pomocnych przy sprowadzaniu macierzy do PKJ; następnie szkicujemy inne metody dowodu.

3. Podprzestrzenie pierwiastkowe

DEFINICJA 2. Zbiór wektorów

$$V(\lambda) = \{\mathbf{v} \in V \mid (\mathcal{A} - \lambda\mathcal{E})^k \mathbf{v} = \mathbf{0} \text{ dla pewnego } k \in \mathbb{N}\}$$

nazywamy *podprzestrzenią pierwiastkową* odpowiadającą wartości własnej $\lambda \in \operatorname{Spec} \mathcal{A}$.

Łatwo sprawdzić, że $V(\lambda)$ jest podprzestrzenią liniową: jeśli $\mathbf{u}, \mathbf{v} \in V(\lambda)$, $(\mathcal{A} - \lambda\mathcal{E})^s\mathbf{u} = \mathbf{0}$, $(\mathcal{A} - \lambda\mathcal{E})^t\mathbf{v} = \mathbf{0}$ i $m = \max(s,t)$, to dla dowolnych $\alpha, \beta \in \mathbb{C}$ mamy

$$(\mathcal{A} - \lambda\mathcal{E})^m(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha(\mathcal{A} - \lambda\mathcal{E})^m\mathbf{u} + \beta(\mathcal{A} - \lambda\mathcal{E})^m\mathbf{v} = \mathbf{0},$$

czyli $\alpha\mathbf{u} + \beta\mathbf{v} \in V(\lambda)$. Ponieważ do $V(\lambda)$ należą wektory własne odpowiadające λ, więc $V(\lambda) \neq \{\mathbf{0}\}$. Mamy zatem $V^\lambda \subset V(\lambda)$, ale równości może nie być, jak pokazuje przykład operatora nilpotentnego stopnia $m > 1$: $\lambda = 0$ jest jedyną wartością własną, $\dim V^0 = 1$, ale $V(0) = V$.

Ponieważ $\dim V(\lambda) \leqslant n$ i ograniczenie operatora $\mathcal{A} - \lambda\mathcal{E}$ do $V(\lambda)$ jest operatorem nilpotentnym, więc (zob. przykład 1)

$$V(\lambda) = \{\mathbf{v} \in V \mid (\mathcal{A} - \lambda\mathcal{E})^n\mathbf{v} = \mathbf{0}\}.$$

TWIERDZENIE 3. *Niech $\mathcal{A} : V \to V$ będzie operatorem liniowym z wielomianem charakterystycznym*

$$\chi_{\mathcal{A}}(t) = \prod_{i=1}^{p}(t - \lambda_i)^{n_i}, \quad \textit{gdzie} \quad \lambda_i \neq \lambda_j \ \textit{dla}\ i \neq j.$$

Wówczas przestrzeń V jest sumą prostą podprzestrzeni pierwiastkowych $V(\lambda_i)$; każda z nich jest $\mathcal{A}$-niezmiennicza i ma wymiar n_i. Operator $\mathcal{A} - \lambda_i\mathcal{E}$, nilpotentny na $V(\lambda_i)$, jest izomorfizmem na podprzestrzeni dopełniającej

$$V_i = V(\lambda_1) \oplus \ldots \oplus V(\lambda_{i-1}) \oplus V(\lambda_{i+1}) \oplus \ldots \oplus V(\lambda_p).$$

Ponadto λ_i jest jedyną wartością własną operatora $\mathcal{A}_{V(\lambda_i)}$.

Dowód. Żaden z czynników $t - \lambda_k$ nie dzieli jednocześnie wszystkich wielomianów

$$\chi_i(t) = \prod_{j \neq i}(t - \lambda_j)^{n_j}, \qquad i = 1, \ldots, p,$$

wobec tego $\operatorname{NWD}(\chi_1(t), \ldots, \chi_p(t)) = 1$. Istnieją zatem wielomiany $f_1(t), \ldots, f_p(t) \in \mathbb{C}[t]$ spełniające równość

$$\sum_{i=1}^{p} \chi_i(t) f_i(t) = 1. \tag{4}$$

Podprzestrzenie

$$W_i = \chi_i(\mathcal{A}) f_i(\mathcal{A}) V = \{\chi_i(\mathcal{A}) f_i(\mathcal{A})\mathbf{v} \mid \mathbf{v} \in V\}, \qquad i = 1, \ldots, p,$$

są $\mathcal{A}$-niezmiennicze:

$$\mathcal{A}W_i = \chi_i(\mathcal{A}) f_i(\mathcal{A}) \mathcal{A}V \subset \chi_i(\mathcal{A}) f_i(\mathcal{A}) V = W_i.$$

Ponadto

$$(\mathcal{A}-\lambda_i\mathcal{E})^{n_i}W_i = \chi_{\mathcal{A}}(\mathcal{A})f_i(\mathcal{A})V = \{\mathbf{0}\}$$

(gdyż $\chi_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$ na mocy twierdzenia 2), więc

$$W_i \subset V(\lambda_i). \tag{5}$$

Równość (4), przepisana w postaci

$$\mathcal{E} = \sum_{i=1}^{p} \chi_i(\mathcal{A})f_i(\mathcal{A}),$$

daje rozkład

$$V = \sum_{i=1}^{p} W_i$$

i tym bardziej (dzięki zawieraniu (5))

$$V = \sum_{i=1}^{p} V(\lambda_i).$$

Załóżmy, że $\mathbf{v} \in V(\lambda_i)\cap V_i$, gdzie $V_i = \sum_{j\neq i} V(\lambda_j)$, jak w sformułowaniu twierdzenia. Wówczas $(\mathcal{A}-\lambda_i\mathcal{E})^n\mathbf{v} = \mathbf{0}$, a ponieważ $\mathbf{v} = \sum_{j\neq i}\mathbf{v}_j$ oraz $(\mathcal{A}-\lambda_j\mathcal{E})^n\mathbf{v}_j = \mathbf{0}$, więc $[\prod_{j\neq i}(\mathcal{A}-\lambda_j\mathcal{E})^n]\mathbf{v} = \mathbf{0}$. Z drugiej strony, wielomiany $(t-\lambda_i)^n$ i $c(t) = \prod_{j\neq i}(t-\lambda_j)^n$ są względnie pierwsze, czyli istnieją wielomiany $a(t)$ i $b(t)$ takie, że

$$a(t)(t-\lambda_i)^n + b(t)c(t) = 1.$$

Wówczas

$$\mathbf{v} = a(\mathcal{A})(\mathcal{A}-\lambda_i\mathcal{E})^n\mathbf{v} + b(\mathcal{A})\Big\{\prod_{j\neq i}(\mathcal{A}-\lambda_j\mathcal{E})^n\Big\}\mathbf{v} = \mathbf{0},$$

tj. podprzestrzenie $V(\lambda_i)$ i V_i mają trywialną część wspólną. Otrzymaliśmy zatem rozkład

$$V = V(\lambda_1)\oplus\ldots\oplus V(\lambda_p) \tag{6}$$

na sumę prostą podprzestrzeni $\mathcal{A}$-niezmienniczych.

Z zawierania (5) i rozkładu (6) wynika, że $W_i = V(\lambda_i)$. Mamy więc jawne wyrażenie na $V(\lambda_i)$:

$$V(\lambda_i) = \chi_i(\mathcal{A})f_i(\mathcal{A})V,$$

gdzie wielomiany $\chi_i(t)$ i $f_i(t)$ pochodzą z tożsamości (4). W szczególności

$$(\mathcal{A}-\lambda_i\mathcal{E})^n V(\lambda_i) = \{\mathbf{0}\}.$$

Wielomianem minimalnym dla $\mathcal{A}$ na $V(\lambda_i)$ jest więc pewien dzielnik wielomianu $(t-\lambda_i)^{n_i}$. Wynika stąd po pierwsze, że λ_i jest jedyną wartością własną operatora $\mathcal{A}_{V(\lambda_i)}$. Ponadto w bazie będącej sumą baz podprzestrzeni $V(\lambda_i)$ operator $\mathcal{A}$ ma macierz

$$A = \begin{Vmatrix} A_1 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & A_p \end{Vmatrix},$$

gdzie A_i jest macierzą stopnia $n'_i = \dim V(\lambda_i)$ z jedyną wartością własną λ_i i wielomianem charakterystycznym

$$\chi_{A_i}(t) = (t - \lambda_i)^{n'_i}, \quad \text{przy czym} \quad n'_i \leqslant n_i.$$

Ale $\chi_A(t) = \prod_{i=1}^p \chi_{A_i}(t)$, a stąd $n = n'_1 + \ldots + n'_p$ i $n_i = n'_i$.

Pozostało nam do udowodnienia, że $(\mathcal{A} - \lambda_i\mathcal{E})_{V_i}$ jest izomorfizmem. To jednak jest oczywiste: w przeciwnym razie $\mathcal{A}\mathbf{v} - \lambda_i\mathbf{v} = \mathbf{0}$ dla pewnego $\mathbf{0} \neq \mathbf{v} \in V_i$, a z drugiej strony wielomianem charakterystycznym dla $\mathcal{A}$ na V_i jest $\chi_i(t) = \prod_{j\neq i}(t - \lambda_j)^{n_j}$ i λ_i nie jest jego pierwiastkiem. ∎

4. Przypadek operatora nilpotentnego. Skoro twierdzenie 3 jest już udowodnione, to zagadnienie wyboru najprostszej macierzy dla operatora liniowego $\mathcal{A} : V \to V$ sprowadza się do przypadku, gdy $\mathcal{A}$ ma jedyną wartość własną λ i $\mathcal{B} = \mathcal{A} - \lambda\mathcal{E}$ jest operatorem nilpotentnym z macierzą nilpotentną B. Niech m będzie ich stopniem nilpotentności; $m \leqslant \dim V$.

W tej sytuacji naturalne jest przyjęcie następującej definicji:

DEFINICJA 3. Podprzestrzeń liniową

$$\mathfrak{K}[\mathcal{B}]\mathbf{v} = \langle \mathbf{v}, \mathcal{B}\mathbf{v}, \ldots, \mathcal{B}^{m'-1}\mathbf{v}\rangle$$

nazywamy *podprzestrzenią cykliczną*, stowarzyszoną z operatorem nilpotentnym $\mathcal{B}$ stopnia m i wektorem $\mathbf{v}$. Tutaj $m' \leqslant m$ oznacza najmniejszą liczbę naturalną, dla której $\mathcal{B}^{m'}\mathbf{v} = \mathbf{0}$.

TWIERDZENIE 4. *Każdą macierz nilpotentną B (nad dowolnym ciałem skalarów) można sprowadzić do postaci kanonicznej Jordana.*

Dowód. Z przykładu 1 i definicji 1 wiemy już, że każdej podprzestrzeni cyklicznej odpowiada klatka Jordana. Wystarczy więc wykazać, że całą przestrzeń można rozłożyć na sumę prostą podprzestrzeni cyklicznych. Wykażemy to przez indukcję względem wymiaru przestrzeni.

Na mocy twierdzenia 1 macierz B można sprowadzić do postaci górnotrójkątnej z zerami na głównej przekątnej. Niech U będzie podprzestrzenią liniową rozpiętą na pierwszych $n-1$ wektorach bazowych. Wtedy $\mathcal{B}V \subset U$ i z założenia indukcyjnego istnieje rozkład

$$U = \mathfrak{K}[\mathcal{B}]\mathbf{e}_1 \oplus \ldots \oplus \mathfrak{K}[\mathcal{B}]\mathbf{e}_s, \tag{7}$$

gdzie

$$\mathfrak{K}[\mathcal{B}]\mathbf{e}_i = \langle \mathbf{e}_i, \mathcal{B}\mathbf{e}_i, \ldots, \mathcal{B}^{m_i-1}\mathbf{e}_i\rangle, \qquad \mathcal{B}^{m_i}\mathbf{e}_i = \mathbf{0}.$$

Możemy przyjąć, że

$$m_1 \geqslant \ldots \geqslant m_s. \tag{8}$$

Ponadto $V = \langle \mathbf{v}, U \rangle$ dla dowolnego $\mathbf{v} \in V \setminus U$, a więc $\mathcal{B}\mathbf{v} = \sum_i \alpha_i \mathbf{e}_i + \mathcal{B}\mathbf{u}$ dla pewnego $\mathbf{u} \in U$. Biorąc $\mathbf{v}' = \mathbf{v} - \mathbf{u}$, otrzymamy

$$V = \langle \mathbf{v}', U \rangle, \qquad \mathcal{B}\mathbf{v}' = \sum_{i=1}^{s} \alpha_i \mathbf{e}_i.$$

Jeśli $\alpha_i = 0$ dla $i = 1, \ldots, s$, to do klatek Jordana $J_{m_1}(0), \ldots, J_{m_s}(0)$ dojdzie klatka $J_1(0)$ odpowiadająca podprzestrzeni cyklicznej $\langle \mathbf{v}' \rangle$, tj.

$$B \sim J(B) = \operatorname{diag}(J_{m_1}(0), \ldots, J_{m_s}(0), J_1(0)).$$

Załóżmy teraz, że dla pewnego $r \geqslant 1$ mamy

$$\alpha_1 = \ldots = \alpha_{r-1} = 0, \qquad \alpha_r \neq 0, \qquad \mathcal{B}\mathbf{v}' = \sum_{i=r}^{s} \alpha_i \mathbf{e}_i.$$

Przyjmijmy

$$\mathbf{e}_i' = \mathbf{e}_i \text{ dla } i \neq r, \qquad \mathbf{e}_r' = \frac{1}{\alpha_r}\mathbf{v}', \qquad \beta_i = \frac{\alpha_i}{\alpha_r}.$$

Wtedy

$$\mathcal{B}\mathbf{e}_r' = \mathbf{e}_r + \sum_{i=r+1}^{s} \beta_i \mathbf{e}_i =: \mathbf{f}_r.$$

Ponieważ $\mathcal{B}^{m_r}\mathbf{f}_r = \mathbf{0}$ na mocy nierówności (8) oraz suma (7) jest prosta, wynika stąd, że $\mathcal{B}^{m_r - 1}\mathbf{f}_r \neq \mathbf{0}$, niezależnie od tego, jakie są współczynniki β_i. Ponadto nietrudno wykazać, że suma

$$\sum_{i \neq r} \mathfrak{K}[\mathcal{B}]\mathbf{e}_i' + \mathfrak{K}[\mathcal{B}]\mathbf{f}_r$$

jest też prosta i równa U. Teraz jednak podprzestrzeń cykliczną $\mathfrak{K}[\mathcal{B}]\mathbf{f}_r$ można rozszerzyć do $\mathfrak{K}[\mathcal{B}]\mathbf{e}_r'$, gdzie $\mathbf{e}_r' \notin U$, i otrzymujemy sumę prostą

$$V = \bigoplus_{i=1}^{s} \mathfrak{K}[\mathcal{B}]\mathbf{e}_i',$$

odpowiadającą wykładnikom $m_1', \ldots, m_s'$, gdzie $m_i' = m_i$ dla $i \neq r$ oraz $m_r' = m_r + 1$. Zatem

$$B \sim \operatorname{diag}(J_{m_1'}(0), \ldots, J_{m_s'}(0))$$

(liczba klatek nie uległa zmianie, tylko rozmiar jednej z nich zwiększył się o 1). Ciąg $(m_1', \ldots, m_s')$ nie jest na ogół uporządkowany; jeśli chcemy to osiągnąć, wystarczy przenumerować wektory bazowe. ■

5. Jednoznaczność. Udowadniając jednoznaczność, wskażemy jednocześnie praktyczne postępowanie przy sprowadzaniu danej macierzy A do postaci kanonicznej Jordana.

Należy mianowicie umieć znajdować liczbę $N(m, \lambda)$ klatek Jordana stopnia m odpowiadających wartości własnej λ. Przyporządkowujemy jak zwykle macierzy A operator $\mathcal{A}$ działający w n-wymiarowej przestrzeni liniowej V, którą rozkładamy na sumę prostą

$$V = V(\lambda) \oplus V', \tag{9}$$

gdzie

$$V(\lambda) = \bigoplus_{j=1}^{s} \langle \mathbf{e}_j, (\mathcal{A} - \lambda\mathcal{E})\mathbf{e}_j, \ldots, (\mathcal{A} - \lambda\mathcal{E})^{m_j-1}\mathbf{e}_j \rangle,$$

$$V' = \sum_{\lambda' \neq \lambda} V(\lambda').$$

Obliczymy rząd r_t macierzy $(A - \lambda E)^t$ lub, co na jedno wychodzi, wymiar przestrzeni $(\mathcal{A} - \lambda\mathcal{E})^t V$. Nie zależy on oczywiście od wyboru bazy w V. Obie przestrzenie w rozkładzie (9) są niezmiennicze względem $(\mathcal{A} - \lambda\mathcal{E})^t$, a więc

$$\dim(\mathcal{A} - \lambda\mathcal{E})^t V = \sum \dim(\mathcal{A} - \lambda\mathcal{E})^t \mathbb{C}[\mathcal{A}]\mathbf{e}_j + \dim(\mathcal{A} - \lambda\mathcal{E})^t V'.$$

Załóżmy dla ustalenia uwagi, że $m_1 \leqslant \ldots \leqslant m_s$. Jeśli $m_j \leqslant t$, to $(\mathcal{A} - \lambda\mathcal{E})^t \mathbb{C}[\mathcal{A}]\mathbf{e}_j = \{\mathbf{0}\}$. Dla $m_j > t$ mamy

$$(\mathcal{A} - \lambda\mathcal{E})^t \mathbb{C}[\mathcal{A}]\mathbf{e}_j = \langle (\mathcal{A} - \lambda\mathcal{E})^t \mathbf{e}_j, (\mathcal{A} - \lambda\mathcal{E})^{t+1}\mathbf{e}_j, \ldots, (\mathcal{A} - \lambda\mathcal{E})^{m_j-1}\mathbf{e}_j \rangle,$$

czyli

$$\dim(\mathcal{A} - \lambda\mathcal{E})^t \mathbb{C}[\mathcal{A}]\mathbf{e}_j = m_j - t.$$

Operator $\mathcal{A} - \lambda\mathcal{E}$ jest izomorfizmem na podprzestrzeni V' (twierdzenie 1), zatem

$$\dim(\mathcal{A} - \lambda\mathcal{E})^t V' = \dim V'.$$

Stąd

$$r_t = \sum_{m_j > t} (m_j - t) + \dim V',$$

a wobec tego

$$\begin{aligned}
r_t - r_{t+1} &= \sum_{m_j > t} (m_j - t) - \sum_{m_j > t+1} (m_j - t - 1) \\
&= \sum_{m_j > t} (m_j - t) - \sum_{m_j > t+1} (m_j - t) + \sum_{m_j > t+1} 1 \\
&= \sum_{m_j = t+1} 1 + \sum_{m_j > t+1} 1 = N(t+1, \lambda) + N(t+2, \lambda) + \ldots
\end{aligned}$$

Wynika stąd, że

$$\begin{aligned}
r_{m-1} - r_m - (r_m - r_{m+1}) &= \{N(m, \lambda) + N(m+1, \lambda) + \ldots\} \\
&\quad - \{N(m+1, \lambda) + N(m+2, \lambda) + \ldots\} \\
&= N(m, \lambda)
\end{aligned}$$

i ostatecznie

$$N(m,\lambda) = r_{m-1} - 2r_m + r_{m+1}, \tag{10}$$

gdzie

$$m \geqslant 1, \quad r_t = \operatorname{rank}(\mathcal{A} - \lambda\mathcal{E})^t, \quad r_0 = n.$$

Zauważmy, że r_t jest niezmiennikiem podobieństwa (tzn. zależy tylko od klasy podobieństwa macierzy A). Ze wzoru (10) wynika więc jednoznaczność postaci kanonicznej Jordana. ■

Nie mówiliśmy dotychczas nic o macierzy C ustalającej podobieństwo:

$$J(A) = C^{-1}AC.$$

Skoro jednak obie macierze A i $J(A)$ są teraz znane, to $C = (c_{ij})$ możemy znaleźć z równania macierzowego

$$XJ(A) - AX = 0,$$

o którym wspominaliśmy w definicji 1(d) i które jest równoważne układowi liniowemu jednorodnemu o n^2 niewiadomych. Niech $C_1, \dots, C_r$ będzie bazą przestrzeni rozwiązań tego równania. Na ogół nie wszystkie macierze C_i są nieosobliwe, ale ponieważ postać kanoniczna Jordana istnieje, więc

$$\det(\alpha_1 C_1 + \ldots + \alpha_r C_r) \neq 0$$

dla pewnych współczynników $\alpha_1, \dots, \alpha_r \in \mathbb{C}$. Wówczas $C = \alpha_1 C_1 + \ldots + \alpha_r C_r$ jest szukaną macierzą. Nie jest ona, rzecz jasna, wyznaczona jednoznacznie, nawet przy unormowaniu $\det C = 1$.

Znajdowanie w ten sposób macierzy C, realizującej sprowadzenie do postaci kanonicznej Jordana, nie jest szczególnie praktyczne, ale nie przedstawia istotnych trudności.

Przykład 4. Wielomianem minimalnym macierzy

$$S = \sum_{i,j} E_{ij} = \left\| \begin{matrix} 1 & \dots & 1 \\ \dots & \dots & \dots \\ 1 & \dots & 1 \end{matrix} \right\|,$$

spełniającej warunek $S^2 = nS$, jest oczywiście $\mu_S(t) = t^2 - nt$, tj. $\lambda_1 = n$ i $\lambda_2 = 0$ są wartościami własnymi macierzy S o krotnościach odpowiednio 1 i $n-1$. Ponieważ $\operatorname{rank} S = 1$ i macierz S nie jest nilpotentna (ma inny wielomian minimalny), więc jej postacią Jordana może być tylko $J(S) = \operatorname{diag}(n, 0, \dots, 0)$. Nie jest to dziwne, gdyż $P = (1/n)S$ jest macierzą operatora rzutowania. Rozwiązaniem równania macierzowego

$$\left\| \begin{matrix} x_{11} & \dots & x_{1n} \\ \dots & \dots & \dots \\ x_{n1} & \dots & x_{nn} \end{matrix} \right\| \cdot \left\| \begin{matrix} n & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & 0 \end{matrix} \right\| = \left\| \begin{matrix} 1 & \dots & 1 \\ \dots & \dots & \dots \\ 1 & \dots & 1 \end{matrix} \right\| \cdot \left\| \begin{matrix} x_{11} & \dots & x_{1n} \\ \dots & \dots & \dots \\ x_{n1} & \dots & x_{nn} \end{matrix} \right\|$$

jest np.

$$C = \begin{Vmatrix} 1/n & -1 & \dots & -1 \\ 1/n & n-1 & \dots & -1 \\ \dots & \dots & \dots & \dots \\ 1/n & -1 & \dots & n-1 \end{Vmatrix}.$$

6. Inne podejścia do PKJ. Efektywnych metod znajdowania postaci kanonicznej Jordana i bazy Jordana dostarcza ogólna teoria modułów nad pierścieniami ideałów głównych; jedną z jej odmian stanowi dobrze rozwinięta teoria λ-macierzy. Choć teoria ta jest uniwersalna i prowadzi do innych ważnych wniosków, jej przedstawienie w tej książce byłoby nużące i nie wydaje się wskazane. Z drugiej strony, bezpośrednia metoda geometryczna, którą przyjęliśmy, jest dostatecznie poglądowa i zasługuje na przedstawienie jeszcze jednego jej wariantu, łączącego twierdzenia 3 i 4 (zob. np. podręczniki [2] i [9]).

Niech więc $\mathcal{A} : V \to V$ będzie zespolonym operatorem liniowym, a λ — jedną z jego wartości własnych. Ponieważ baza Jordana dla $\mathcal{A} - \lambda\mathcal{E}$ jest jednocześnie bazą Jordana dla $\mathcal{A}$, możemy założyć, że $\lambda = 0$. Wtedy operator $\mathcal{A}$ jest osobliwy, zawieranie $\operatorname{Im}\mathcal{A} \subset V$ nie jest równością i możemy rozumować przez indukcję względem $n = \dim V$. Rozpatrzmy ciąg podprzestrzeni

$$\operatorname{Im}\mathcal{A}^0 \supset \operatorname{Im}\mathcal{A}^1 \supset \operatorname{Im}\mathcal{A}^2 \supset \dots \supset \operatorname{Im}\mathcal{A}^{p-1} \supset \operatorname{Im}\mathcal{A}^p = \operatorname{Im}\mathcal{A}^{p+1} = \dots,$$

stabilizujący się na p-tym wyrazie (dla pewnego p), czyli

$$\operatorname{Im}\mathcal{A}^p \cap \operatorname{Ker}\mathcal{A} = \{\mathbf{0}\}, \qquad \operatorname{Im}\mathcal{A}^{p-1} \cap \operatorname{Ker}\mathcal{A} \neq \{\mathbf{0}\}.$$

Zgodnie z ćwiczeniem 3 w § 3 mamy rozkład

$$V = \operatorname{Ker}\mathcal{A}^p \oplus \operatorname{Im}\mathcal{A}^p$$

na sumę prostą podprzestrzeni $\mathcal{A}$-niezmienniczych. Jeśli $\dim \operatorname{Im}\mathcal{A}^p > 0$, to na mocy założenia indukcyjnego oba składniki mają bazy Jordana; ich suma jest bazą Jordana w V.

Pozostaje rozpatrzyć sytuację, gdy

$$V = \operatorname{Ker}\mathcal{A}^p,$$

tj. $\mathcal{A}^p = \mathcal{O}$, $\mathcal{A}^{p-1} \neq \mathcal{O}$. Tak więc — omijając twierdzenie 3 (ważne jednak samo w sobie) — sprowadziliśmy zadanie do przypadku operatora nilpotentnego. Nie odwołując się do twierdzenia 4, postąpimy następująco. Dla wygody wprowadźmy oznaczenie

$$V_i = \operatorname{Im}\mathcal{A}^{i-1} \cap \operatorname{Ker}\mathcal{A}.$$

Wtedy

$$\operatorname{Ker}\mathcal{A} = V_1 \supset V_2 \supset \dots \supset V_p \neq \{\mathbf{0}\}, \qquad V_{p+1} = \{\mathbf{0}\}.$$

W przestrzeni V_p wybierzmy bazę $(\mathbf{a}_i^1 \mid i = 1, \dots, n_p)$. Ponieważ $\mathbf{a}_i^1 \in \operatorname{Im}\mathcal{A}^{p-1}$, więc $\mathbf{a}_i^1 = \mathcal{A}^{p-1}\mathbf{a}_i^p$ dla pewnego wektora $\mathbf{a}_i^p$. Definiujemy $\mathbf{a}_i^k = \mathcal{A}^{p-k}\mathbf{a}_i^p$ dla

$k = 1, \dots, p$. Uzupełniamy wektory $\mathbf{a}_i^l$ do bazy podprzestrzeni V_{p-1}, dodając n_{p-1} wektorów $\mathbf{b}_j^1$; przedstawiamy je w postaci $\mathbf{b}_j^1 = \mathcal{A}^{p-2}\mathbf{b}_j^{p-1}$ i definiujemy $\mathbf{b}_j^l = \mathcal{A}^{p-l-1}\mathbf{b}_j^{p-1}$ dla $l = 1, \dots, p-1$. Następnie uzupełniamy wektory $\mathbf{a}_i^1, \mathbf{b}_j^1$ do bazy w V_{p-2}, dodając n_{p-2} wektorów $\mathbf{c}_k^1$ itd. Cała ta procedura przedstawiona jest na poniższym diagramie:

$$
\begin{array}{cccccc}
\mathbf{a}_i^p & & & & & \\
\downarrow & & & & & \\
\mathbf{a}_i^{p-1} & \mathbf{b}_j^{p-1} & & & & \\
\downarrow & \downarrow & & & & \\
\mathbf{a}_i^{p-2} & \mathbf{b}_j^{p-2} & \mathbf{c}_k^{p-2} & & & \\
\downarrow & \downarrow & \downarrow & & & \\
\multicolumn{6}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\
\downarrow & \downarrow & \downarrow & & & \\
\mathbf{a}_i^2 & \mathbf{b}_j^2 & \mathbf{c}_k^2 & \dots & \mathbf{d}_s^2 & \\
\downarrow & \downarrow & \downarrow & \dots & \downarrow & \\
\mathbf{a}_i^1 & \mathbf{b}_j^1 & \mathbf{c}_k^1 & \dots & \mathbf{d}_s^1 & \mathbf{e}_t^1 \\
\downarrow & \downarrow & \downarrow & \dots & \downarrow & \downarrow \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0}
\end{array}
$$

Przypuśćmy, że wektory w tym diagramie są liniowo zależne:

$$\sum_i \alpha_i^{(p)}\mathbf{a}_i^p + \sum_i \alpha_i^{(p-1)}\mathbf{a}_i^{p-1} + \ldots + \sum_j \beta_j^{(p-1)}\mathbf{b}_j^{p-1} + \ldots + \sum_t \delta_t \mathbf{e}_t^1 = \{\mathbf{0}\}. \quad (11)$$

Działając na (11) operatorem $\mathcal{A}^{p-1}$, otrzymamy

$$\sum_i \alpha_i^{(p)}\mathbf{a}_i^1 = \mathbf{0},$$

skąd

$$\alpha_i^{(p)} = 0, \qquad i = 1, \dots, n_p.$$

Działając teraz na (11) operatorem $\mathcal{A}^{p-2}$, dochodzimy do wniosku, że

$$\sum_i \alpha_i^{(p-1)}\mathbf{a}_i^1 + \sum_j \beta_j^{(p-1)}\mathbf{b}_j^1 = \mathbf{0},$$

co wobec wyboru wektorów daje $\alpha_i^{(p-1)} = 0 = \beta_j^{(p-1)}$. Kontynuując to postępowanie, przekonujemy się, że zależność (11) jest trywialna. Z drugiej strony, łączna liczba wektorów w diagramie wynosi

$$
\begin{aligned}
|\{\mathbf{a}_i^k\}| + |\{\mathbf{b}_j^l\}| + \ldots + |\{\mathbf{e}_t\}| \\
&= pn_p + (p-1)n_{p-1} + \ldots + 2n_2 + n_1 \\
&= (n_1 + \ldots + n_p) + (n_2 + \ldots + n_p) + \ldots + (n_{p-1} + n_p) + n_p
\end{aligned}
$$

$$= \dim V_1 + \dim V_2 + \ldots + \dim V_{p-1} + \dim V_p$$
$$= \sum_{i=1}^{p} \dim(\operatorname{Im} \mathcal{A}^{i-1} \cap \operatorname{Ker} \mathcal{A}) = \sum_i (\dim \operatorname{Ker} \mathcal{A}^i - \dim \operatorname{Ker} \mathcal{A}^{i-1})$$
$$= \dim \operatorname{Ker} \mathcal{A}^p = \dim V$$

(w związku z tą równością zob. ćwiczenie 7 w § 2).

Wektory w diagramie tworzą więc bazę przestrzeni V; na mocy konstrukcji jest to baza Jordana.

Co się tyczy jednoznaczności PKJ, to stosując oznaczenia punktu 5, mamy

$$\begin{aligned}
N(m, 0) &= \dim V_m - \dim V_{m+1} \\
&= (\dim \operatorname{Ker} \mathcal{A}^m - \dim \operatorname{Ker} \mathcal{A}^{m-1}) - (\dim \operatorname{Ker} \mathcal{A}^{m+1} - \dim \operatorname{Ker} \mathcal{A}^m) \\
&= 2 \dim \operatorname{Ker} \mathcal{A}^m - \dim \operatorname{Ker} \mathcal{A}^{m-1} - \dim \operatorname{Ker} \mathcal{A}^{m+1} \\
&= \operatorname{rank} \mathcal{A}^{m-1} - 2 \operatorname{rank} \mathcal{A}^m + \operatorname{rank} \mathcal{A}^{m+1} = r_{m-1} - 2r_m + r_{m+1},
\end{aligned}$$

a o tej wielkości wiemy już, że jest niezmiennicza.

7. Inne postaci normalne. W tym punkcie omówimy pokrótce inne postaci normalne macierzy, wygodne w szczególności dla ciał, które nie są algebraicznie domknięte.

(a) *Podprzestrzenie cykliczne i klatki cykliczne.* Rozwiniemy nieco definicję 3. Przestrzeń V wymiaru n nad $\mathfrak{K}$ nazywamy *cykliczną* względem operatora liniowego $\mathcal{A} : V \to V$, jeśli w V istnieje wektor $\mathbf{v}$ (również zwany *cyklicznym*) taki, że

$$V = \langle \mathbf{e}_1, \ldots, \mathbf{e}_n \rangle, \quad \text{gdzie} \quad \mathbf{e}_i = \mathcal{A}^{n-i}\mathbf{v} \text{ dla } i = 1, \ldots, n.$$

Ponieważ $(\mathbf{e}_i)$ jest bazą w V, więc

$$\mathcal{A}^n \mathbf{v} = \sum_{i=0}^{n-1} \alpha_i \mathcal{A}^i \mathbf{v}$$

dla jednoznacznie wyznaczonych współczynników $\alpha_i \in \mathfrak{K}$; oznacza to, że macierzą operatora $\mathcal{A}$ w tej bazie jest tzw. *klatka cykliczna*

$$A = \begin{Vmatrix} \alpha_{n-1} & 1 & 0 & \ldots & 0 & 0 \\ \alpha_{n-2} & 0 & 1 & \ldots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ \alpha_1 & 0 & 0 & \ldots & 0 & 1 \\ \alpha_0 & 0 & 0 & \ldots & 0 & 0 \end{Vmatrix}. \tag{12}$$

Na odwrót, jeśli macierzą operatora $\mathcal{A}$ w bazie $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ jest klatka cykliczna, to wektor $\mathbf{v} = \mathbf{e}_n$ jest cykliczny dla $\mathcal{A}$, przy czym $\mathbf{e}_i = \mathcal{A}^{n-i}\mathbf{e}_n$ (dowód przez indukcję „w dół” względem i).

Pokażemy, że postać klatki cyklicznej odpowiadającej $\mathcal{A}$ nie zależy od wyboru wektora cyklicznego. W tym celu sprawdzimy, że pierwsza kolumna macierzy (12) składa się ze współczynników wielomianu minimalnego operatora $\mathcal{A}$:

$$\mu_{\mathcal{A}}(t) = f(t) = t^n - \sum_{i=0}^{n-1} \alpha_i t^i.$$

Istotnie, $f(\mathcal{A}) = \mathcal{O}$, ponieważ

$$f(\mathcal{A})(\mathcal{A}^i \mathbf{v}) = \mathcal{A}^i(f(\mathcal{A})\mathbf{v}) = \mathbf{0},$$

a wektory $\mathcal{A}^i\mathbf{v}$ generują V. Z drugiej strony, $g(\mathcal{A}) \neq \mathcal{O}$ dla każdego niezerowego wielomianu $g(t)$ stopnia $< n$, gdyż w przeciwnym razie działając operatorem $g(\mathcal{A}) = \mathcal{O}$ na wektor cykliczny $\mathbf{v}$, otrzymalibyśmy nietrywialną zależność liniową między wektorami bazowymi $\mathcal{A}^i\mathbf{v}$.

(b) *Kryterium cykliczności przestrzeni.* Zgodnie z powyższymi rozważaniami, jeśli przestrzeń V jest cykliczna względem $\mathcal{A}$, to jej wymiar jest stopniem wielomianu minimalnego operatora $\mathcal{A}$, skąd wynika, że wielomian minimalny jest równy wielomianowi charakterystycznemu. Zachodzi również implikacja odwrotna (§ 3, ćwiczenie 4).

(c) *Każda macierz jest podobna do sumy prostej klatek cyklicznych.* Dowód można przeprowadzić analogicznie do dowodu twierdzenia o PKJ. Zamiast czynników $(t-\lambda_i)^{n_i}$ wielomianu charakterystycznego należy rozpatrzyć czynniki $p_i(t)^{r_i}$, gdzie $p_i(t)$ są dzielnikami wielomianu charakterystycznego, nieprzywiedlnymi nad ciałem $\mathfrak{K}$. Jednoznaczność ma również miejsce pod warunkiem, że ograniczymy się do klatek cyklicznych, których wielomiany minimalne są nieprzywiedlne. Bez tego założenia jednoznaczności nie ma: przestrzeń cykliczna może być sumą prostą dwóch podprzestrzeni cyklicznych, dla których wielomiany minimalne są względnie pierwsze.

ĆWICZENIA

1. Wykorzystując macierz S z przykładu 4 i wzór na $J(S)$, znaleźć wyznacznik macierzy

$$A = \begin{Vmatrix} m & -1 & \dots & -1 \\ -1 & m & \dots & -1 \\ \dots & \dots & \dots & \dots \\ -1 & -1 & \dots & m \end{Vmatrix},$$

przedstawiając go w postaci $\det A = \chi_S(m+1)$.

2. Z dokładnością do podobieństwa istnieją cztery macierze nilpotentne 4×4:

$$A_1 = J_2(0) \dot{+} J_1(0) \dot{+} J_1(0), \qquad A_2 = J_2(0) \dot{+} J_2(0),$$
$$A_3 = J_3(0) \dot{+} J_1(0), \qquad A_4 = J_4(0).$$

Macierze

$$\begin{Vmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{Vmatrix}, \quad \begin{Vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{Vmatrix}, \quad \begin{Vmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{Vmatrix}, \quad \begin{Vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{Vmatrix},$$

są oczywiście nilpotentne. Do których spośród macierzy A_i są one podobne?

3. (a) Znaleźć $J(\mathcal{A})$, wiedząc, że $\chi_{\mathcal{A}}(t) = (t-3)^4(t+2)$ i $\mathrm{rank}(\mathcal{A} - 3\mathcal{E}) = 2$.
(b) Czy $J(\mathcal{A})$ będzie wyznaczone jednoznacznie, jeśli $\mathrm{rank}(\mathcal{A} - 3\mathcal{E}) = 1, 3, 4$?

4. (a) Wykazać, że macierze

$$A = \begin{Vmatrix} 6 & 2 & -2 \\ -2 & 2 & 2 \\ 2 & 2 & 2 \end{Vmatrix}, \qquad B = \begin{Vmatrix} 6 & 2 & 2 \\ -2 & 2 & 0 \\ 0 & 0 & 2 \end{Vmatrix}$$

mają jednakowe wielomiany charakterystyczne.
(b) Znaleźć $\mu_A(t)$ i $\mu_B(t)$.
(c) Znaleźć $J(A)$ i $J(B)$.

5. Niech $A \in M_n(\mathfrak{K})$, gdzie $\mathrm{char}\,\mathfrak{K} = 0$. Udowodnić, że macierz A jest nilpotentna dokładnie wtedy, gdy $\mathrm{tr}(A^k) = 0$ dla $k = 1, \ldots, n$.

6. Udowodnić, że macierze $A \in M_n(\mathbb{C})$ i ${}^tA$ są podobne.

7. Wykazać, że jeśli $A \in M_n(\mathbb{C})$, to $A^N = E$ wtedy i tylko wtedy, gdy macierz A jest diagonalizowalna i jej wartości własne są pierwiastkami z jedynki stopnia N.

8. Łatwo sprawdzić, że jeśli

$$A = \begin{Vmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{Vmatrix} \in \mathrm{Mag}_3(\mathbb{Q}),$$

to $A^2 \notin \mathrm{Mag}_3(\mathbb{Q})$, tj. kwadraty magiczne, w odróżnieniu od półmagiczny (§ 3, ćwiczenie 8), nie stanowią pierścienia. Tym bardziej nieoczekiwany j fakt, że jeśli $A \in \mathrm{Mag}_3(\mathbb{Q})$, to $A^m \in \mathrm{Mag}_3(\mathbb{Q})$ dla każdego nieparzyst $m \geqslant 1$. Udowodnić to, wykorzystując twierdzenie Hamiltona-Cayleya.

9. Sprawdzić, że

$$A = \begin{Vmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{Vmatrix} \in \mathrm{Mag}_4(\mathbb{Q}),$$

ale macierz A^m nie jest magiczna dla żadnego $m \geqslant 2$. Wykorzystując n A, wykazać, że dla każdego $n \geqslant 4$ istnieje macierz magiczna, której potęga (o wykładniku $m \geqslant 2$) nie jest magiczna.

10. Zapiszmy macierz $A = J_1(\lambda) \dot{+} J_2(\mu)$, gdzie $\lambda \neq \mu$, w postaci $A = S + N$, gdzie $S = \operatorname{diag}(\lambda, \mu, \mu)$:

$$A = \begin{Vmatrix} \lambda & 0 & 0 \\ 0 & \mu & 1 \\ 0 & 0 & \mu \end{Vmatrix}, \quad S = \begin{Vmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{Vmatrix}, \quad N = \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{Vmatrix}.$$

Wyrazić S i N jako wielomiany $s(A)$ i $n(A)$ od macierzy A (na mocy twierdzenia Hamiltona–Cayleya można uważać, że $\deg s(t) \leqslant 2$ i $\deg n(t) \leqslant 2$).

11. Niech V będzie n-wymiarową przestrzenią zespoloną, a $\mathcal{A}$ — operatorem liniowym na V. Udowodnić, że $\mathcal{A}$ można jednoznacznie przedstawić w postaci

$$\mathcal{A} = \mathcal{S} + \mathcal{N}, \qquad \mathcal{S}\mathcal{N} = \mathcal{N}\mathcal{S},$$

gdzie operator $\mathcal{S}$ jest diagonalizowalny, $\mathcal{N}$ — nilpotentny, a oba są wielomianami od $\mathcal{A}$ (operator $\mathcal{S}$ nazywa się *składową półprostą*, a $\mathcal{N}$ — *składową nilpotentną* operatora $\mathcal{A}$).

12. Obliczyć $(J_n(\lambda))^k$ dla dowolnej liczby naturalnej k.

13. Udowodnić, że $[[X, Y]^2, Z] = 0$ dla dowolnych trzech macierzy $X, Y, Z \in M_2(\mathfrak{K})$, gdzie $[X, Y] = XY - YX$ oznacza komutator macierzy.