Rozdział 7

Wartości i wektory własne

Niech X będzie skończenie wymiarową przestrzenią liniową nad ciałem $\mathbb{F} = \mathbb{R}$ lub $\mathbb{F} = \mathbb{C}$. Niech $f : X \to X$ będzie **endomorfizmem**, tj. odwzorowaniem liniowym przekształającym przestrzeń liniową w nią samą.

Definicja 7.1. *Skalar $\lambda \in \mathbb{F}$ nazywamy* **wartością własną** *endomorfizmu f jeżeli istnieje niezerowy wektor $v \in X$, taki że*

$$f(v) = \lambda v; \tag{7.1}$$

wektor v nazywamy **wektorem własnym** *odpowiadającym wartości własnej λ.*

Zachodzi następujące

Twierdzenie 7.1. *Dla endomorfizmu $f : X \to X$ następujące warunki są równoważne:*

(a) *λ jest wartością własną f;*
(b) $\ker(f - \lambda\,\mathrm{id}_X) \neq \{0\}$;
(c) $\det(A_f - \lambda I) = 0$, *gdzie A_f jest macierzą endomorfizmu f (w dowolnej bazie przestrzeni X).*

Niech teraz $A = [a_{ij}] \in \mathbb{F}^{n\times n}$ będzie dowolną macierzą, a X n–wymiarową przestrzenią liniową o bazie $e_1, \ldots, e_n$. Możemy skonstruować endomorfizm $f : X \to X$, którego macierzą w ustalonej bazie $e_1, \ldots, e_n$ jest macierz A. Endomorfizm ten wystarczy zdefiniować na wektorach bazowych (co z innymi wektorami?) w następujący sposób:

$$f(e_i) = \sum_{j=1}^{n} a_{ji} e_j \qquad (i = 1, \ldots, n).$$

Pojęcia wartości własnej oraz wektora własnego w sposób naturalny przenoszą się więc na macierze.

Definicja 7.2. *Skalar $\lambda \in \mathbb{F}$ nazywamy* **wartością własną** *macierzy $A \in \mathbb{F}^{n\times n}$ jeżeli istnieje wektor $v \neq 0$, taki że*

$$Av = \lambda v;$$

wektor v nazywamy **wektorem własnym** *odpowiadającym wartości własnej λ.*

Zbiór wszystkich wartości własnych macierzy A oznaczamy $\sigma(A)$ i nazywamy jej **widmem**.
Prawdziwe jest następujące twierdzenie.

Twierdzenie 7.2. *Dla macierzy $A \in \mathbb{F}^{n\times n}$ następujące warunki są równoważne:*

(a) *λ jest wartością własną A;*
(b) *układ równań $(A-\lambda I)\,v=0$ ma niezerowe rozwiązanie;*
(c) $\det(A-\lambda I)=0$.

Dla dowolnej macierzy $A \in \mathbb{F}^{n\times n}$ odwzorowanie

$$\varphi_A : \lambda \to \det(A-\lambda I)$$

jest wielomianem stopnia n (ćwiczenie), którego pierwiastkami są wartości własne macierzy A. Wielomian φ_A nazywamy **wielomianem charakterystycznym** macierzy A.

Uwaga 7.1. *Jeżeli elementy macierzy $A \in \mathbb{F}^{n\times n}$ należą do ciała $\mathbb{F}$, które jest* **algebraicznie domknięte** *(tzn. każdy wielomian stopnia n o współczynnikach z ciała $\mathbb{F}$ ma n pierwiastków w ciele $\mathbb{F}$), to macierz A posiada n wartości własnych (liczonych z krotnościami). Jeżeli natomiast elementy macierzy są elementami ciała, które nie jest algebraicznie domknięte (takim ciałem jest na przykład ciało liczb rzeczywistych!), to macierz ta może nie mieć wartości własnych (zob. przykład 7.4 poniżej).*

Przypuśćmy, że $\lambda_1,\ldots,\lambda_n \in \sigma(A)$ są wartościami własnymi macierzy $A \in \mathbb{F}^{n\times n}$. Wówczas, wielomian charakterystyczny φ_A macierzy A możemy zapisać w postaci

$$\begin{aligned}\varphi_A(\lambda) &= a_n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0\\ &= a_n(\lambda-\lambda_1)\cdots(\lambda-\lambda_n).\end{aligned}$$

Łatwo wykazać, że $a_n=(-1)^n$ oraz, uwzględniając wzory Viète'a,

$$\begin{aligned}\lambda_1\cdot\ldots\cdot\lambda_n &= a_0 = \det A,\\ \lambda_1+\ldots+\lambda_n &= (-1)^{n+1}a_{n-1} = \operatorname{tr}(A),\end{aligned}$$

gdzie $\operatorname{tr}(A)=a_{11}+\ldots+a_{nn}$ to **ślad** macierzy A.

Własności widma macierzy $(A\in\mathbb{F}^{n\times n})$**:**

a) $\lambda\in\sigma(A),\ k\in\mathbb{N}\Rightarrow\lambda^k\in\sigma\left(A^k\right);$
b) $\lambda\in\sigma(A)$, $\det A\neq 0\Rightarrow\lambda^{-1}\in\sigma\left(A^{-1}\right);$
c) $\lambda\in\sigma(A)$, $\alpha\in\mathbb{F}\Rightarrow\alpha\lambda\in\sigma(\alpha A);$
d) $\lambda\in\sigma(A)\Rightarrow\overline{\lambda}\in\sigma(A^*)$; *w szczególności:* $\lambda\in\sigma(A)\Rightarrow\lambda\in\sigma\left(A^T\right)$.

Ćwiczenie Uzasadnić powyższe własności.

Przykład 7.1. *Wyznaczymy wartości oraz wektory własne macierzy*

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

Ponieważ $\varphi_A(\lambda)=\det(A-\lambda I)=-(1-\lambda)(2-\lambda)(1+\lambda)$, *zatem macierz A ma trzy różne wartości własne: $\lambda_1=-1,\lambda_2=1,\lambda_3=2$. Dla każdej z nich wyznaczymy wektor własny:*

- *dla $\lambda_1=-1$ mamy:*

$$\begin{bmatrix}2 & 2 & 0\\ 0 & 3 & 0\\ -2 & -2 & 0\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}2x+2y\\ 3y\\ -2x-2y\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$$

skąd otrzymujemy $(x,y,z)=(0,0,t)$, $t\in\mathbb{R}$; *przykładowy wektor własny* $v_{\lambda_1}=(0,0,1)$;

- *dla* $\lambda_2 = 1$ *mamy:*

$$\begin{bmatrix} 0 & 2 & 0 \\ 0 & 1 & 0 \\ -2 & -2 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2y \\ y \\ -2x-2y-2z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

skąd otrzymujemy $(x, y, z) = (t, 0, -t)$, $t \in \mathbb{R}$*; przykładowy wektor własny* $v_{\lambda_2} = (1, 0, -1)$;

- *dla* $\lambda_3 = 2$ *mamy:*

$$\begin{bmatrix} -1 & 2 & 0 \\ 0 & 0 & 0 \\ -2 & -2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x+2y \\ 0 \\ -2x-2y-3z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

skąd otrzymujemy $(x, y, z) = (2t, t, -2t)$, $t \in \mathbb{R}$*; przykładowy wektor własny* $v_{\lambda_3} = (2, 1, -2)$.

7.1. Twierdzenie Cayleya-Hamiltona

Rozważmy macierz kwadratową $A \in \mathbb{C}^{n\times n}$ oraz wielomian w postaci

$$w(s) = a_m s^m + \ldots + a_1 s + a_0.$$

Możemy wówczas utworzyć macierz $w(A)$ określoną jako

$$w(A) = a_m A^m + \ldots + a_1 A + a_0 I_n.$$

Definicja 7.3 (wielomian zerujący). *Jeżeli* $w(A) = 0_n$*, to wielomian* w *nazywamy* **wielomianem zerującym** *(wielomianem anulującym) macierzy* A.

Pytanie jakie może się nasunąć jest następujące – czy dla każdej macierzy istnieje wielomian zerujący? Odpowiedź jest zawarta w poniższym twierdzeniu.

Twierdzenie 7.3 (Cayley–Hamilton). *Wielomian charakterystyczny macierzy kwadratowej jest jej wielomianem zerującym.*

Przykład 7.2 (odwracanie macierzy). *Rozważmy macierz* $A \in \mathbb{R}^{3\times 3}$ *postaci*

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

Jej wielomian charakterystyczny ma postać

$$\varphi(\lambda) = \begin{vmatrix} 1-\lambda & 2 & -1 \\ -2 & -\lambda & 3 \\ 2 & 0 & -\lambda \end{vmatrix} = -\lambda^3 + \lambda^2 - 6\lambda + 12.$$

Ponieważ $\varphi(0) = 12 \neq 0$ *zatem macierz* A *jest nieosobliwa. Wyznaczymy teraz jej odwrotność. Na podstawie twierdzenie Cayleya–Hamiltona możemy napisać*

$$-A^3 + A^2 - 6A + 12I_3 = 0_3$$

lub równoważnie

$$\frac{1}{12} A\left(A^2 - A + 6I_3\right) = I_3.$$

Oznacza to, na podstawie definicji macierzy odwrotnej, że

$$A^{-1} = \frac{1}{12}\left(A^2 - A + 6I_3\right) = \frac{1}{12}\left(\begin{pmatrix} -5 & 2 & 5 \\ 4 & -4 & 2 \\ 2 & 4 & -2 \end{pmatrix} - \begin{pmatrix} 1 & 2 & -1 \\ -2 & 0 & 3 \\ 2 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix}\right)$$
$$= \frac{1}{12}\begin{pmatrix} 0 & 0 & 6 \\ 6 & 2 & -1 \\ 0 & 4 & 4 \end{pmatrix}.$$

7.2. Podprzestrzeń własna

Niech $A \in \mathbb{F}^{n\times n}$ oraz niech $\lambda \in \sigma(A)$. Zbiór

$$V_\lambda = \{v \in \mathbb{F}^n : Av = \lambda v\}$$

składa się z $\mathbf{0}$ oraz wszystkich wektorów własnych odpowiadających wartości własnej λ. Ponieważ

$$V_\lambda = \{v \in \mathbb{F}^n : (A - \lambda I)v = 0\} = \ker\{v \to (A - \lambda I)\, v\}$$

zatem zbiór ten – jako jądro endomorfizmu – jest podprzestrzenią liniową przestrzeni $\mathbb{F}^n$; jest to tak zwana **podprzestrzeń własna** macierzy A (ew. podprzestrzeń własna endomorfizmu wyznaczonego przez macierz A) odpowiadająca wartości własnej λ. Wymiar podprzestrzeni własnej V_λ to tzw. **krotność geometryczna wartości własnej** λ – jest to liczba odpowiadających jej liniowo niezależnych wektorów własnych macierzy A. Jak się wkrótce okaże, krotność geometryczna wartości własnej λ nigdy nie przekracza jej **krotności algebraicznej**, czyli krotności wartości własnej λ jako pierwiastka wielomianu charakterystycznego φ_A.

Przykład 7.3. *Ponieważ macierz z przykładu 7.1 ma trzy różne wartości własne, zatem możemy dla niej wyznaczyć trzy podprzestrzenie własne:*

$$\begin{aligned} V_{\lambda_1} &= \left\{(x,y,z) \in \mathbb{R}^3 : x = y = 0, z = t, t \in \mathbb{R}\right\}; \\ V_{\lambda_2} &= \left\{(x,y,z) \in \mathbb{R}^3 : x = t, y = 0, z = -t, t \in \mathbb{R}\right\}; \\ V_{\lambda_3} &= \left\{(x,y,z) \in \mathbb{R}^3 : x = 2t, y = t, z = -2t, t \in \mathbb{R}\right\}. \end{aligned}$$

Udowodnimy teraz twierdzenie, z którego wynika bardzo ważna własność podprzestrzeni własnych odpowiadających różnym wartościom własnym.

Twierdzenie 7.4. *Różnym wartościom własnym macierzy $A \in \mathbb{F}^{n\times n}$ odpowiadają liniowo niezależne wektory własne.*

Dowód: Niech $\lambda_1, \ldots, \lambda_k$ $(k \leqslant n)$ będą różnymi wartościami własnymi macierzy A, a $v_i \in V_{\lambda_i}$ $(i = 1, \ldots, k)$ odpowiadającymi im wektorami własnymi. Należy wykazać warunek

$$\alpha_1 v_1 + \ldots + \alpha_k v_k = 0 \Rightarrow \alpha_1 = \ldots = \alpha_k = 0. \tag{7.2}$$

Dowód poprowadzimy przez indukcję względem k. Dla $k = 1$ teza zachodzi (wektor zerowy, mimo że należy do każdej podprzestrzeni własnej, nie jest wektorem własnym). Załóżmy, że teza zachodzi dla dowolnych $k - 1$ wektorów własnych odpowiadających różnym wartościom własnym oraz, dla dowodu nie wprost, przypuśćmy, że warunek (7.2) nie jest spełniony. Oznacza to, że dla pewnego $i \in \{1, \ldots, k\}$:

$$\alpha_1 v_1 + \ldots + \alpha_k v_k = 0 \quad \text{oraz} \quad \alpha_i \neq 0.$$

Dla dowolnego $r \neq i$ mamy

$$\begin{aligned} 0 = &(A - \lambda_r I)(\alpha_1 v_1 + \ldots + \alpha_k v_k) = \alpha_1 (A - \lambda_r I) v_1 + \ldots + \alpha_k (A - \lambda_r I) v_k = \\ =&\alpha_1 (\lambda_1 - \lambda_r) v_1 + \ldots + \alpha_{r-1} (\lambda_{r-1} - \lambda_r) v_{r-1} + \alpha_{r+1} (\lambda_{r+1} - \lambda_r) v_{r+1} + \\ &+ \ldots + \alpha_k (\lambda_k - \lambda_r) v_k \end{aligned}$$

skąd, na podstawie założenia indukcyjnego, wynika, że wszystkie współczynniki $\alpha_m (\lambda_m - \lambda_r)$ są zerami; w szczególności $\alpha_i (\lambda_i - \lambda_r) = 0$. Ponieważ $\lambda_i \neq \lambda_r$, zatem $\alpha_i = 0$, wbrew założeniu. ■

7.3. Diagonalizowalność

Niech $f : X \to X$ będzie endomorfizmem.

Definicja 7.4. *Endomorfizm f jest* **diagonalizowalny**, *jeżeli istnieje baza przestrzeni X w której macierz tego endomorfizmu jest diagonalna.*

Pojęcie diagonalizowalności można również wprowadzić w zbiorze macierzy. Zanim to zrobimy wprowadzimy w zbiorze macierzy relację podobieństwa.

Definicja 7.5. *Niech $A, B \in \mathbb{F}^{n \times n}$. Mówimy, że macierz A jest podobna do macierzy B (ozn. $A \sim B$) jeżeli istnieje macierz nieosobliwa $P \in \mathbb{F}^{n \times n}$, taka że*

$$A = PBP^{-1}. \tag{7.3}$$

Łatwo wykazać (ćwiczenie), że relacja podobieństwa macierzy jest relacją równoważności, tzn. jest:

- zwrotna, tj. $A \sim A$;
- symetryczna, tj. $A \sim B \Rightarrow B \sim A$;
- przechodnia, tj. $A \sim B, B \sim C \Rightarrow A \sim C$.

Przypuśćmy teraz, że $A \sim B$. Oznacza to, że $A = PBP^{-1}$, dla pewnej macierzy nieosobliwej P. Mamy:

$$\begin{aligned} \varphi_A(\lambda) &= \det(A - \lambda I) = \det\left(PBP^{-1} - \lambda I\right) = \det P (B - \lambda I) P^{-1} = \\ &= \det P \det(B - \lambda I) \det P^{-1} = \varphi_B(\lambda) \end{aligned}$$

co oznacza, że **macierze podobne mają ten sam wielomian charakterystyczny**; w konsekwencji mają one również identyczne wartości własne.

Definicja 7.6. *Macierz $A \in \mathbb{F}^{n \times n}$ jest macierzą* **diagonalizowalną**, *jeżeli jest podobna do macierzy diagonalnej.*

Zanim podamy twierdzenie charakteryzujące macierze diagonalizowalne rozważmy następujący przykład.

Przykład 7.4. *Niech*

$$A_1 = I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Macierze te mają ten sam wielomian charakterystyczny

$$\varphi_{A_1}(\lambda) = \varphi_{A_2}(\lambda) = (1 - \lambda)^3.$$

Ponieważ widma macierzy A_1 oraz A_2 są jednoelementowe

$$\sigma(A_1) = \sigma(A_2) = \{1\},$$

możemy dla każdej z nich wyznaczyć po jednej przestrzeni własnej.

Dla macierzy A_1 *mamy:*

$$V^{(1)}_{\lambda=1} = \left\{x \in \mathbb{R}^3 : x = x\right\} = \mathbb{R}^3,$$

co oznacza, że możemy dla niej wybrać dokładnie trzy liniowo niezależne wektory własne odpowiadające jej jedynej wartości własnej $\lambda = 1$. *Macierz* A_1 *– jako macierz diagonalna – jest oczywiście macierzą diagonalizowalną.*

Z kolei dla macierzy A_2 *mamy*

$$V^{(2)}_{\lambda=1} = \left\{x \in \mathbb{R}^3 : A_2 x = x\right\} = \left\{x \in \mathbb{R}^3 : (A_2 - I)\, x = 0\right\} = \left\{(x_1, 0, 0) : x_1 \in \mathbb{R}\right\}.$$

Oznacza to, że dla macierzy A_2 *znajdziemy tylko jeden liniowo niezależny wektor własny odpowiadający jej jedynej wartości własnej* $\lambda = 1$. *Ponadto, macierz* A_2 *nie jest diagonalizowalna. Jedyną macierzą diagonalną, do której macierz* A_2 *mogłaby być podobna, jest macierz jednostkowa (macierze podobne mają te same wartości własne). Musiałaby więc istnieć macierz nieosobliwa* P *spełniająca warunek*

$$A_2 = P I_3 P^{-1} = I_3,$$

który nie jest prawdziwy.

Możliwość diagonalizacji macierzy A_1 oraz brak możliwości diagonalizacji macierzy A_2 jest wynikiem tego, że dla macierzy A_1 możemy wybrać tyle liniowo niezależnych wektorów własnych, ile wynosi krotność jej wartości własnej jako pierwiastka wielomianu charakterystycznego; dla macierzy A_2 warunek ten nie jest spełniony.

Prawdziwe jest następujące

Twierdzenie 7.5. *Macierz* $A \in \mathbb{F}^{n\times n}$ *jest diagonalizowalna wtedy i tylko wtedy, gdy*

(a) jej wielomian charakterystyczny ma n *pierwiastków w ciele* $\mathbb{F}$ *(liczonych z krotnościami);*

(b) dla każdej wartości własnej macierzy A *można wybrać tyle liniowo niezależnych wektorów własnych, ile wynosi krotność tej wartości własnej jako pierwiastka wielomianu charakterystycznego.*

Dowód: Przypuśćmy, że macierz A jest diagonalizowalna. Oznacza to, że istnieją macierz diagonalna $B = \operatorname{diag}(b_1, \ldots, b_n)$, $b_i \in \mathbb{F}$ oraz macierz nieosobliwa $P \in \mathbb{F}^{n\times n}$, dla których $A = PBP^{-1}$, lub równoważnie

$$AP = PB. \tag{7.4}$$

Niech $P = [p_1, \ldots, p_n]$, gdzie $p_i \in \mathbb{F}^n$ $(i = 1, \ldots, n)$. Ponieważ

$$AP = A\,[p_1, \ldots, p_n] = [Ap_1, \ldots, Ap_n]$$

oraz

$$PB = [p_1, \ldots, p_n] \operatorname{diag}(b_1, \ldots, b_n) = [b_1 p_1, \ldots, b_n p_n],$$

zatem z warunku (7.4) otrzymujemy, że

$$Ap_i = b_i p_i \quad (i = 1, \ldots, n). \tag{7.5}$$

Z warunku (7.5) wynika, że wektory $p_1, \ldots, p_n$ są liniowo niezależnymi wektorami własnymi macierzy A odpowiadającymi wartościom własnym $b_1, \ldots, b_n$.

Przypuśćmy teraz, że macierz A posiada n liniowo niezależnych wektorów własnych $v_1, \ldots, v_n \in \mathbb{F}^n$ odpowiadających wartościom własnym $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$. Niech $V = [v_1, \ldots, v_n] \in \mathbb{F}^{n\times n}$. Macierz V jest nieosobliwa (dlaczego?). Ponadto:

$$AV = A\,[v_1, \ldots, v_n] = [Av_1, \ldots, Av_n] = [\lambda_1 v_1, \ldots, \lambda_n v_n] = V \operatorname{diag}(\lambda_1, \ldots, \lambda_n),$$

skąd wynika, że $A = V \operatorname{diag}(\lambda_1, \ldots, \lambda_n) V^{-1}$. Macierz A jest więc diagonalizowalna. $\blacksquare$

Zanotujmy na koniec, że twierdzenie 7.4, mimo że sformułowane dla macierzy, pozostaje słuszne również dla dowolnego endomorfizmu $f : X \to X$, gdzie X jest skończenie wymiarową przestrzenią wektorową nad ciałem $\mathbb{F}$. W szczególności, jeżeli f jest endomorfizmem diagonalizowalnym to istnieje baza przestrzeni X złożona z wektorów własnych endomorfizmu f, przy której macierz tego endomorfizmu jest diagonalna.

Przykład 7.5. *Niech $A \in \mathbb{R}^{2\times 2}$ będzie macierzą postaci*

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

Ponieważ $\varphi_A(\lambda) = \lambda^2 + 1$, zatem macierz A nie ma rzeczywistych wartości własnych – nie jest więc diagonalizowalna w klasie macierzy $\mathbb{R}^{2\times 2}$. Ta sama macierz traktowana jako element przestrzeni $\mathbb{C}^{2\times 2}$ ma dwie różne wartości własne $\lambda_1 = i$, $\lambda_2 = -i$, którym odpowiadają liniowo niezależne wektory własne równe odpowiednio $v_1 = (-i, 1)$ oraz $v_2 = (i, 1)$. Z dowodu twierdzenia 7.4 wynika, że $A = P \operatorname{diag}(i, -i) P^{-1}$, gdzie

$$P = [v_1, v_2] = \begin{bmatrix} -i & i \\ 1 & 1 \end{bmatrix}.$$

Faktycznie, ponieważ $P^{-1} = \frac{1}{2} \begin{bmatrix} i & 1 \\ -i & 1 \end{bmatrix}$, zatem

$$P \operatorname{diag}(i, -i) P^{-1} = \frac{1}{2} \begin{bmatrix} -i & i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \begin{bmatrix} i & 1 \\ -i & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = A.$$