1 Struktura endomorfizmu (operatora) liniowego

1.1 Algebra endomorfizmów. Wielomian od operatora

Definicja. *Algebrą* nad ciałem $\mathbf{K}$ nazywamy zbiór $\mathcal{A}$ wyposażony w strukturę przestrzeni wektorowej nad $\mathbf{K}$ oraz mnożenie:

$$\mathcal{A} \times \mathcal{A} \ni (a, b) \to a \cdot b \in \mathcal{A}$$

które jest *łączne*, tzn.

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c \quad \forall\, a, b, c \in \mathcal{A}$$

oraz *dwuliniowe* (por. definicję k-formy), tzn.

$$a\cdot(\lambda b+\mu c) = \lambda(a\cdot b)+\mu(a\cdot c), \quad (\lambda b+\mu c)\cdot a = \lambda(b\cdot a)+\mu(c\cdot a) \quad \forall\, a, b, c \in \mathcal{A},\ \lambda, \mu \in \mathbf{K}.$$

Mówimy, że algebra $\mathcal{A}$ *ma jedynkę*, jeśli istnieje element $I \in \mathcal{A}$ taki, że $I \cdot a = a = a \cdot I\ \forall a \in \mathcal{A}$.

Mówimy, że algebra $\mathcal{A}$ *jest przemienna*, jeśli $a \cdot b = b \cdot a\ \forall a, b \in \mathcal{A}$.

Przykłady:

1. Zbiór wielomianów $\mathbf{K}[\cdot]$ jest algebrą przemienną z jedynką.

2. Odwzorowanie liniowe przestrzeni wektorowej V *w siebie* nazywamy *endomorfizmem* przestrzeni V (lub *operatorem* w V). Zbiór endomorfizmów przestrzeni wektorowej V:

$$\mathrm{End}V \equiv L(V, V) \equiv L(V)$$

   jest algebrą nieprzemienną z jedynką. W szczególności, zbiór $\mathrm{End}\mathbf{K}^n \cong \mathbf{K}^n_n$ macierzy kwadratowych jest algebrą (nieprzemienną z jedynką).

Dla wielomianu $w \in \mathbf{K}[\cdot]$,

$$w(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \ldots + a_n\lambda^n$$

oraz operatora $T \in \mathrm{End}V$ można utworzyć operator

$$w(T) = a_0 I + a_1 T + a_2 T^2 + \ldots + a_n T^n.$$

(odtąd mnożenie w algebrze będziemy zapisywac bez użycia $\cdot$).

To formalne podstawienie T w miejsce λ ma sens dla elementów każdej algebry z jedynką (nad $\mathbf{K}$), w szczególności dla operatorów. Operację przejścia od T do $w(T)$ nazywamy *wzięciem wielomianu w od operatora T*.

1.2 Wektory własne, wartości własne i wielomian charakterystyczny operatora. Podprzestrzenie niezmiennicze

Definicja. Niech $T \in \text{End}V$. Wektor $x \in V$ taki, że $Tx = \lambda x$ dla pewnej liczby $\lambda \in \mathbf{K}$, nazywa się *wektorem własnym* operatora T.

Definicja. Dla niezerowego wektora własnego x, liczba λ taka, że $Tx = \lambda x$, jest wyznaczona jednoznacznie. Nazywamy ją *wartością własną* operatora T, odpowiadającą wektorowi x.

Definicja. Zbiór wartości własnych operatora T nazywamy *widmem* operatora T i oznaczamy jako $\text{Sp}\, T$: $\text{Sp}\, T = \{\lambda_1, \ldots, \lambda_r\}$.

Jeśli x jest wektorem własnym dla T, to $Tx = \lambda x$, $T^2 x = \lambda^2 x$, $T^3 x = \lambda^3 x$, $\ldots, T^n x = \lambda^n x$ i w konsekwencji $w(T)x = w(\lambda)x$ dla dowolnego wielomianu w. Pokazuje to użyteczność wektorów własnych T przy analizie $w(T)$.

Jak znajdować wektory własne dla danego operatora T?

Zakładamy, że V (przestrzeń w której działa operator T) ma wymiar skończony: $\dim V = n$. Skorzystamy teraz z faktu, iż równanie

$$(T - \lambda I)x = \mathbf{0}$$

ma niezerowe rozwiązanie (na x) stedy i tylko wtedy, gdy odwzorowanie: $T - \lambda I$ jest nieodwracalne, czyli wtedy i tylko wtedy gdy

$$\det(T - \lambda I) = 0,$$

gdzie przez wyznacznik operatora rozumiemy wyznacznik macierzy tego operatora w dowolnej bazie. Aby to określenie miało sens, musimy się upewnić, iż nie zależy ono od bazy; tak jest w istocie:

Stwierdzenie. Wyznacznik operatora nie zależy od bazy.

Dowód: Pamiętamy, zę przy zamianie bazy $\{e_i\} \to \{f_i\}$ macierz operatora A zmienia się w następujący sposób:

$$[A]^f{}_f = [Id]^f{}_e [A]^e{}_e [Id]^e{}_f$$

gdzie macierze $[Id]^f{}_e$ i $[Id]^e{}_f$ są macierzami zamian bazy; są one odwrotne względem siebie. Mamy więc:

$$\det[A]^f{}_f = \det([Id]^f{}_e [A]^e{}_e [Id]^e{}_f) = \det([Id]^f{}_e)\det([A]^e{}_e)\det([Id]^e{}_f) = \det([A]^e{}_e)$$

(ostatnia równość wynika z faktu iż $\det([Id]^f{}_e) = 1/\det([Id]^e{}_f)$).

Definicja. Funkcję zmiennej λ:

$$w_T(\lambda) := \det(T - \lambda I); \tag{1}$$

nazywamy *wielomianem charakterystycznym* operatora T (widać bowiem, że funkcja $w_T(\lambda)$ jest wielomianem stopnia n w zmiennej λ).

W myśl powyższego rozumowania, pierwiastki wielomianu charakterystycznego pokrywają się z wartościami własnymi operatora T.

Stąd mamy proste

Stwierdzenie: Widmo dowolnego operatora w przestrzeni skończenie wymiarowej posiada nie więcej niż n wartości.

Definicja. Podprzestrzeń V_1 przestrzeni V nazywa się *podprzestrzenią niezmienniczą* dla T, jeżeli $T(V_1) \subset V_1$. Znajdowaniem podprzestrzeni niezmienniczych dla operatorów zajmiemy się dalej.

Przykłady.

1. Rozpatrzmy: $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. Mamy: $w_A(\lambda) = \det \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} = \lambda^2 - 4\lambda + 3 = (\lambda-1)(\lambda-3)$, tak więc pierwiastkami równania charakterystycznego są: $\lambda_1 = 1, \lambda_2 = 3$, a odpowiadającymi im wektorami własnymi: $v_1 = \mathrm{Ker}(A - \lambda_1 I) = [1,-1]^T$, $v_2 = \mathrm{Ker}(A - \lambda_2 I) = [1,1]^T$. Tutaj wektory własne rozpinają całą przestrzeń. Podprzestrzenie rozpięte przez wektory v_1, v_2 są też podprzestrzeniami niezmienniczymi operatora A.

2. Rozpatrzmy: $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$. Mamy:
$w_A(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & 1 & 2 \\ 0 & 1-\lambda & 1 \\ 0 & 0 & 2-\lambda \end{bmatrix} = (1-\lambda)^2(2-\lambda)$, pierwiastkami równania charakterystycznego są więc: $\lambda_1 = 1$ (dwukrotny), $\lambda_2 = 2$ (jednokrotny). Wektorami własnymi, odpowiadającymi tym wartościom własnym, są: Dla λ_1: $v_1 = \mathrm{Ker}(A - \lambda_1 I) = \langle [1,0,0]^T \rangle$, oraz dla λ_2: $v_2 = \mathrm{Ker}(A - \lambda_2 I) = \langle [3,1,1]^T \rangle$. Tutaj wektory własne nie rozpinają całej przestrzeni. Operator A posiada dwie podprzestrzenie niezmiennicze: jedna jest rozpięta przez wektor v_1, a druga - przez wektor v_2 i dodatkowo przez wektor $v_2' = [0,1,0]^T$.

1.3 Rzuty

Niech V będzie sumą prostą

$$V = V_1 \oplus V_2.$$

Znaczy to, że każdy $v \in V$ posiada jednoznaczny rozkład $v = v_1 + v_2$, gdzie $v_i \in V_i, i = 1, 2$. Przyporządkowanie: $v \to P_1 v := v_1$ jest odwzorowaniem liniowym. Nazywa się ono *rzutem na* V_1 *wzdłuż* V_2. Podobnie mamy: $v \to P_2 v := v_2$ – rzut na V_2 wzdłuż V_1. Operatory P_1, P_2 mają własności:

$$P_1 + P_2 = I, \quad P_1^2 = P_1, \quad P_2^2 = P_2, \quad P_1 P_2 = P_2 P_1 = \mathbf{0}. \tag{2}$$

Na odwrót, jeśli P_1 i P_2 są jakimiś operatorami spełniającymi warunki (2), to $V = \mathrm{Im}\, P_1 \oplus \mathrm{Im}\, P_2$, i przy tym P_1, P_2 są rzutami związanymi z tym rozkładem. Zachodzi bowiem:

1. Każdy $v \in V$ ma rozkład: $v = P_1 v + P_2 v$ (wynika to z pierwszej równości spośród (2)).

2. Rozkład ten jest jednoznaczny: jeśli $\mathbf{0} = v_1 + v_2$, gdzie $v_1 \in \operatorname{Im} P_1$, $v_2 \in \operatorname{Im} P_2$, to

$$\mathbf{0} = P_1(v_1 + v_2) = P_1 v_1 = v_1, \mathbf{0} = P_2(v_1 + v_2) = P_2 v_2 = v_2.$$

Powyższa argumentacja pokazuje, że mamy wzajemnie jednoznaczną odpowiedniość między rozkądami przestrzeni V na sumę prostą dwóch podprzestrzeni, $V = V_1 \oplus V_2$, a parami operatorów P_1, P_2 spełniających warunki (2).

Uwaga: Jeśli operator $P \in \operatorname{End} V$ spełnia równość $P^2 = P$, to operatory $P_1 := P$, $P_2 := I - P$ spełniają (2).

Operator $P \in \operatorname{End} V$, spełniający warunek $P^2 = P$, nazywamy *operatorem rzutowym.*

Powyższą konstrukcję operatorów rzutowych łatwo można uogólnić. Niech V będzie sumą prostą:

$$V = V_1 \oplus \ldots \oplus V_k,$$

tzn. każdy $v \in V$ ma jednoznaczny rozkład $v = v_1 + \ldots + v_k$, gdzie $v_i \in V_i$, $i = 1, \ldots, k$. Przyporządkowanie $v \to P_i v := v_i$ jest rzutem. Operatory $P_1, \ldots, P_k$ mają własności:

$$I = P_1 + \ldots + P_k, \quad P_i^2 = P_i \ \ (i = 1, \ldots, k); \quad P_i P_j = \mathbf{0} \ \ (i \neq j). \tag{3}$$

Na odwrót, jeśli $P_1, \ldots, P_k$ są operatorami spełniającymi (3), to $V = \operatorname{Im} P_1 \oplus \ldots \oplus \operatorname{Im} P_k$ (dowód jest analogiczny jak dla $k = 2$).

Uwagi.

1. Warunki (3) nie są niezależne. Np. warunki $P_i^2 = P_i$ wynikają z pozostałych:
$$P_i = P_i I = P_i(P_1 + \ldots + P_k) = P_i P_i.$$

2. Niektóre z rzutów P_i mogą być zerami (tzn. operatorami zerowymi); wtedy odpowiadające im podprzestrzenie V_i są podprzestrzeniami zerowymiarowymi.

Układ operatorów spełniających (3) nazywamy *rozkładem jedynki* (na rzuty).

Przykład. Niech $V = \mathbf{R}^3$, V_1 – podprzestrzeń wektorów, których składowe spełniają równanie: $x_1 + x_2 + x_3 = 0$, zaś V_2 – podprzestrzeń wektorów, których składowe spełniają warunki: $x_1 = x_2 = x_3$. Czytelnik może przekonać się (zachęcam do tego), że:

• podprzestrzeń V_1 jest rozpięta przez wektory: $E_1 = [1, -1, 0]^T$, $E_2 = [1, 0, -1]^T$, zaś V_2 – przez wektor $E_3 = [1, 1, 1]^T$;

• • $V_1 \cap V_2 = \mathbf{0}$,

zatem $V = V_1 \oplus V_2$. Znajdźmy rzuty P_1, P_2 na podprzestrzenie V_1, V_2.

Spróbujmy najpierw zgadnąć, jak wygląda rozwiązanie problemu w jakiejś szczególnej bazie. Gdybyśmy wzięli: $e_1 = [1, 0, 0]^T, e_2 = [0, 1, 0]^T, e_3 = [0, 0, 1]^T$, $\tilde{V}_1 = \langle e_1, e_2 \rangle, \tilde{V}_2 = \langle e_3 \rangle$, to rzuty $\tilde{P}_1, \tilde{P}_2$ na podprzestrzenie $\tilde{V}_1, \tilde{V}_2$ odpowiednio miałyby postać:

$$\tilde{P}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \tilde{P}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Gdybyśmy teraz umieli przejść z bazy e do bazy E, to nasze zadanie byłoby rozwiązane. Weźmy bowiem transformację S wektorów bazy: $SE_1 = e_1, SE_2 = e_2, SE_3 = e_3$, a następnie operatorów P_1, P_2: $\tilde{P}_1 = SP_1S^{-1}, \tilde{P}_2 = SP_2S^{-1}$.

Przejście od bazy E do bazy e możemy zapisać: $S[E_1, E_2, E_3] = [e_1, e_2, e_3]$, czyli: $S \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, skąd mamy: $S = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}^{-1} = \frac{1}{3} \begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & -2 \\ 1 & 1 & 1 \end{bmatrix}$; stąd:

$$P_1 = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & -2 \\ 1 & 1 & 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix},$$

$$P_2 = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & -2 \\ 1 & 1 & 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Czytelnik zechce sprawdzić, że są to rzeczywiście rzuty na przestrzenie V_1, V_2 i że warunki (2) są spełnione.

1.4 Twierdzenie Cayleya – Hamiltona

Mamy ważne twierdzenie, bardzo ułatwiające rachunki na macierzach.

Twierdzenie Cayleya – Hamiltona (w skrócie: CHam'a). Niech $T \in L(V, V)$ i niech $w_T(\lambda)$ będzie wielomianem charakterystycznym T. Wówczas $w_T(T) = 0$.

Dowód. Oznaczmy: $A := [T]^e_e$. Przypomnijmy sobie (z miejsca, gdzie było o dopełnieniu algebraicznym):

$$(A - \lambda I)(A - \lambda I)^D = \det(A - \lambda I)I. \tag{4}$$

Z definicji dopełnienia algebraicznego, wyrazy macierzy $(A - \lambda I)^D$ są wielomianami zmiennej λ stopnia co najwyżej $n - 1$. Możemy więc zapisać:

$$(A - \lambda I)^D = B_0 + \lambda B_1 + \ldots + \lambda^{n-1} B_{n-1}$$

dla pewnych macierzy $B_0, B_1, \ldots, B_{n-1}$. Zapiszmy teraz równość (4) jako wielomian w zmiennej λ (o współczynnikach *macierzowych*):

$$(A-\lambda I)(A-\lambda I)^D = AB_0 + \lambda(AB_1 - B_0) + \lambda^2(AB_2 - B_1) + \ldots + \lambda^{n-1}(AB_{n-1} - B_{n-2}) - \lambda^n B_{n-1}$$

A teraz wstawmy do powyższej równości $\lambda = A$:

$$(A - \lambda I)(A - \lambda I)^D|_{\lambda=A} = \det(A - \lambda I)I|_{\lambda=A}$$

$$= AB_0 + A^2B_1 - AB_0 + A^3B_2 - A^2B_1 + \ldots + A^nB_{n-1} - A^{n-1}B_{n-2} - A^nB_{n-1} = \mathbf{0}.$$

Pierwsze ważne zastosowania tw. CHam'a poznamy już za chwilę.

1.5 Rozkład na podprzestrzenie pierwiastkowe

Uwaga: Na razie będziemy zakładać, że pracujemy z ciałem *liczb zespolonych* $\mathbf{C}$.

Mamy do czynienia z operatorem T, posiadającym wielomian charakterystyczny $w_T(\lambda)$. Zakładamy, że pierwiastki tego wielomianu to λ_1, λ_r, a ich krotności to $n_1, \ldots, n_r$ odpowiednio. Mamy więc

$$w_T(\lambda) = \prod_{k=1}^{r}(\lambda_k - \lambda)^{n_k}.$$

Niech λ_i będzie pierwiastkiem $w_T(\lambda)$ o krotności n_i. Zdefiniujmy wielomian ω_i:

$$\omega_i(\lambda) := \prod_{k\neq i}(\lambda_k - \lambda)^{n_k} = \frac{w_T(\lambda)}{(\lambda_i - \lambda)^{n_i}}$$

Mamy proste

Stwierdzenie. $\mathrm{NWD}(\omega_1, \ldots, \omega_r) = 1$. Wynika stąd (p. rozdział o wielomianach), że istnieją takie wielomiany $u_1, \ldots, u_r$, że

$$\sum_{i=1}^{r} u_i\omega_i = 1.$$

Zdefiniujmy endomorfizmy $P_1, \ldots, P_r$ przestrzeni V wzorami:

$$P_i := u_i(T)\omega_i(T).$$

Mają one następujące własności:

1. $\sum_{i=1}^{r} u_i\omega_i = 1$, więc $\sum_{i=1}^{r} P_i = I$.

2. $P_iP_j = \mathbf{0}$ dla $i \neq j$, ponieważ $u_i\omega_i u_j\omega_j = w_T\cdot$(pewien wielomian), zaś $w_T(T) = \mathbf{0}$ z twierdzenia CHam'a.

3. $P_iP_i = P_i$, bo $Id_V = \sum_{j=1}^{r} P_j$ oraz, z poprzedniego punktu, $P_i = P_iI = \sum_{j=1}^{r} P_iP_j = P_i^2$. Zatem P_i jest operatorem rzutowym.

4. Oznaczmy: $V_i = \mathrm{Im}P_i$. Pokażemy, że $V_i = \mathrm{Ker}(T - \lambda_i I)^{n_i}$.
   Istotnie: jeżeli $x \in \mathrm{Ker}(T-\lambda_i I)^{n_i}$, to $P_jx = \mathbf{0}$ dla $j \neq i$ (jest tak, ponieważ $P_j = u_j(T)\omega_j(T)$, a $\omega_j(T)$ jest iloczynem czynników, z których jednym jest $(T - \lambda_i I)^{n_i}$, jeśli $i \neq j$) i w konsekwencji: $x = Ix = \sum_k P_kx = P_ix \in V_i$.
   Na odwrót, jeżeli $x \in \mathrm{Im}P_i$, to istnieje taki wektor v, że $x = P_iv$, czyli

   $$(T - \lambda_i I)^{n_i}x = (T - \lambda_i I)^{n_i}P_iv = (-1)^{n_i}\omega_T(T)v = \mathbf{0}$$

   czyli, innymi słowy, $x \in \mathrm{Ker}(T - \lambda_i I)^{n_i} = V_i$.

5. V_i jest podprzestrzenią niezmienniczą dla T.
Istotnie, zachodzi: $TP_i = P_iT$ (ponieważ P_i jest wielomianem w T), więc dla $x = P_iv$ mamy: $Tx = T(P_iv) = P_i(Tv) \in V_i$.

6. Jeżeli $i \neq j$, to $V_i \cap V_j = \{\mathbf{0}\}$, bo dla $x = P_iv_i = P_jv_j$ (v_i, v_j – jakieś wektory) to mamy z 2)

$$x = P_iv_i = P_i(P_iv_i) = P_ix = P_iP_jv_j = \mathbf{0}.$$

7. $V = \bigoplus_{i=1}^{r}$, ponieważ $x = (\sum_i P_i)x = \sum_i P_ix$, a $V_i = \mathrm{Im}P_i$. Z 2) wynika, że rozkład ten jest jednoznaczny: jeżeli $x = \sum_i y_i$, gdzie $y_i \in V_i$, to $P_ix = P_iy_i = y_i$.

8. Oznaczmy: $d_i = \mathrm{dim}V_i$. Niech wektory: $(e_{i,1}, \ldots, e_{i,d_i})$ tworzą bazę V_i. Z 7), kolekcja wszystkich takich zbiorów:
$(e_{1,1}, \ldots, e_{1,d_1}; e_{2,1}, \ldots, e_{2,d_2}; \ldots, e_{r,1}, \ldots, e_{r,d_r})$ tworzy bazę w całej przestrzeni V; oznaczmy ją e. Na mocy 4), macierz $[T]^e_e$ jest postaci

$$\begin{bmatrix} \boxed{A_1} & & & & \\ & \ddots & & \mathbf{0} & \\ & & \boxed{A_i} & & \\ & \mathbf{0} & & \ddots & \\ & & & & \boxed{A_r} \end{bmatrix} \tag{5}$$

gdzie i-ta klatka (podmacierz) A_i ma rozmiar $d_i \times d_i$ (poza tymi klatkami wszędzie są zera). Stąd mamy:

$$\det(T - \lambda I) = \prod_{i=1}^{r}(A_i - \lambda I).$$

Zatem pierwiastek równania $\det(A_i - \lambda I) = 0$ jest wartością własną T. Niech będzie to λ_j. Istnieje więc $\mathbf{0} \neq v \in V_i$ taki, że $Fv = \lambda_j v$. Zatem $v \in V_i \cap V_j$ i z 6) wynika, że $j = i$. Tak więc $\det(A_i - \lambda I) = (\lambda_i - \lambda)^{d_i}$ i, w konsekwencji,

$$\omega_T(\lambda) = \prod_{i=1}^{r}(\lambda_i - \lambda)^{n_i} = \det(T - \lambda I) = \prod_{i=1}^{r}(A_i - \lambda I) = \prod_{i=1}^{r}(\lambda_i - \lambda)^{d_i}.$$

Stąd $d_i = n_i$, tzn $\mathrm{dim}V_i = n_i$.

Definicja. Podprzestrzeń V_i nazywamy *przestrzenią pierwiastkową* wartości własnej λ_i.

Wszystkie powyższe fakty można zebrać w następujące twierdzenie.
Twierdzenie. (*O rozkładzie na przestrzenie pierwiastkowe*).

Niech $T \in L(V, V)$. Wówczas V rozkłada się na sumę prostą: $V = \bigoplus_{i=1}^{r} V_i$, gdzie $V_i = \mathrm{Ker}(T - \lambda I)^{n_i}$ są podprzestrzeniami niezmienniczymi dla T. Ponadto $\mathrm{dim} V_i = n_i$, gdzie n_i jest krotnością wartości własnej λ_i.

Uwagi.

1. Każda podprzestrzeń pierwiastkowa V_i zawiera wektor własny odpowiadający wartości własnej λ_i.

2. W przypadku gdy wszystkie pierwiastki wielomianu charakterystycznego są różne (jest ich $n = \mathrm{dim} V$), to mamy n wektorów własnych. W konsekwencji, istnieje baza w przestrzeni V złożona z wektorów własnych. W bazie tej, operator T ma postać *diagonalną*:

$$\begin{bmatrix} \lambda_1 & & & & \\ & \ddots & & \mathbf{0} & \\ & & \lambda_i & & \\ & \mathbf{0} & & \ddots & \\ & & & & \lambda_n \end{bmatrix} \tag{6}$$

3. Co się może zdarzyć w przypadku, gdy mamy pierwiastki wielokrotne? Mogą wówczas zajść dwa przypadki:

   - Wszystkie niezerowe wektory przestrzeni pierwiastkowych są wektorami własnymi i istnieje baza przestrzeni V złożona z wektorów własnych; w konsekwencji, w bazie wektorów własnych, operator T ma postać *diagonalną* – analogiczną do (6) z tą różnicą, że niektóre z wartości własnych są równe. Takie operatory nazywamy *diagonalizowalnymi.*

   - Może się też zdarzyć, że ilość wektorów własnych T jest mniejsza niż wymiar przestrzeni V. Wtedy operatora T *nie można* doprowadzić do postaci diagonalnej. Istnieje jednak wtedy inna "standardowa" postać, do której można ten operator doprowadzić – tzw. *postać Jordana.* Rozpatrzmy postać (5) macierzy operatora w bazie wektorów, tworzących bazy podprzestrzeni pierwiastkowych. Weźmy "klatkę" A_i, odpowiadającą działaniu operatora w podprzestrzeni V_i. Okazuje się, że przez odpowiedni dobór bazy można macierz A_i przedstawić w postaci blokowo-diagonalnej

$$\begin{bmatrix} \boxed{J_1} & & & & \\ & \ddots & & \mathbf{0} & \\ & & \boxed{J_m} & & \\ & \mathbf{0} & & \ddots & \\ & & & & \boxed{J_k} \end{bmatrix},$$

gdzie J_m jest *klatką Jordana*, tzn. podmacierzą postaci:

$$\begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ 0 & \dots & \ddots & \dots & 0 \\ 0 & \dots & 0 & \lambda_i & 1 \\ 0 & \dots & 0 & 0 & \lambda_i \end{bmatrix}$$

Jaka jest ilość klatek Jordana oraz jakie są ich rozmiary? Aby odpowiedzieć na te pytania, trzeba obliczyć wymiary kolejnych podprzestrzeni: $\mathrm{Ker}(A - \lambda_i I), \mathrm{Ker}(A - \lambda_i I)^2, \dots, \mathrm{Ker}(A - \lambda_i I)^{n_i} \equiv V_i$.

Czytelnik zainteresowany dowodami oraz większą ilością szczegółów może o tym przeczytać np. w skrypcie S. Zakrzewskiego, "Algebra $\sqrt{-1}$ geometria", lub książce: A. I. Kostrikin, J. I. Manin, "Algebra liniowa i geometria".

Przykład. Operator różniczkowania D w przestrzeni wielomianów $\mathbf{R}_{n-1}[\cdot]$. Ma on jedną, $(n-1)$-krotną wartość własną równą zeru. Czytelnik sprawdzi, że w bazie $e = \{1, x, x^2/2!, \dots, x^n/n!\}$ macierz operatora jest

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

a więc jest od razu w postaci Jordana.

1.6 Funkcje od operatora

Umiejętność liczenia funkcji od operatora jest konieczna w wielu zastosowaniach; zacznijmy od przykładów.

Przykłady.

1. *Ciąg Fibonacciego.* Jest to ciąg $\{u_n\}$ określony rekurencyjnie wzorami:

$$u_0 = u_1 = 1, \qquad u_n = u_{n-1} + u_{n-2}.$$

Naszym celem jest wyprowadzenie jawnego (nierekurencyjnego) wyrażenia na u_n. Aby to zrobić, zauważmy, że powyższą równość określającą u_n można zapisać w postaci *macierzowej*:

$$\begin{bmatrix} u_n \\ u_{n-1} \end{bmatrix} = \begin{bmatrix} u_{n-1} + u_{n-2} \\ u_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_{n-1} \\ u_{n-2} \end{bmatrix}$$

Oznaczając:

$$U_n = \begin{bmatrix} u_{n-1} \\ u_{n-2} \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \tag{7}$$

powyższą równość możemy przepisać jako: $U_n = A \cdot U_{n-1}$ i dalej: $U_n = A^2 \cdot U_{n-2} = \ldots = A^{n-1}U_1$, gdzie $U_1 = [1\,1]^T$. Zadanie będziemy mieli więc rozwiązane, jeśli będziemy umieli obliczyć dowolną potęgę macierzy A, tzn. funkcję $f(A) = A^n$.

2. Układ (jednorodny) równań różniczkowych liniowych o stałych współczynnikach można zapisać w postaci:

$$\frac{dx(t)}{dt} = A \cdot x(t), \tag{8}$$

gdzie $x \in \mathbf{R}^n$, zaś A jest macierzą $n \times n$. W postaci tej można zapisać również dowolne równanie różniczkowe liniowe o stałych współczynnikach. Rozpatrzmy np. równanie ruchu oscylatora harmonicznego: $\ddot{x} = \omega^2 x$ (kropka oznacza różniczkowanie po czasie). Wprowadzając zmienną: $y = \dot{x}$, możemy zapisać równanie ruchu oscylatora harmonicznego jako:

$$\frac{d}{dt}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} y \\ -\omega^2 x \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Rozwiązaniem układu rówań (8) z warunkiem początkowym: $x(0) = x_0$ jest: $x(t) = e^{At}x_0$. Podamy *formalny* dowód tego faktu. Funkcja wykładnicza od argumentu macierzowego jest określona jako: $e^A = I + A + A^2/2! + \ldots + A^n/n! + \ldots$. Mamy zatem:

$$\begin{aligned} \frac{dx(t)}{dt} &= (\frac{d}{dt}I + At + A^2t^2/2! + A^3t^3/3! + \ldots)x_0 = \\ &= (A + A^2t + A^3t^2/2! + \ldots)x_0 = \\ A(I &+ At + A^2t^2/2! + A^3t^3/3! + \ldots)x_0 = e^{At}x_0. \end{aligned}$$

spełnione jest zatem równanie (8); ponieważ $e^{At}|_{t=0} = I$, to widać, że spełniony jest również warunek początkowy. Aby dowód był pełny, należy udowodnić zbieżność powyższego szeregu i poprawność różniczkowania go wyraz za wyrazem; będzie to przedstawione na analizie.

Konkluzja powyższych rozważań jest taka, że aby rozwiązać układ równań różniczkowych, trzeba umieć obliczyć funkcję wykładniczą od argumentu macierzowego.

Wszędzie dalej, gdy będziemy liczyć $f(A)$, zakładamy o funkcji f, że jest ona analityczna (tzn. rozwijalna w szereg Taylora).

Metoda 1. (Znalezienie reszty z dzielenia przez wielomian charakterystyczny). Załóżmy na początek, że $f(x)$ jest wielomianem. Z rozdziału o wielomianach przypomnijmy sobie, że dzieląc $f(x)$ przez $w_A(x)$ otrzymujemy:

$$f(x) = w_A(x)q(x) + r(x), \tag{9}$$

gdzie $\deg r < \deg w_A$. Podstawiając za x naszą macierz A otrzymujemy

$$f(A) = w_A(A)q(A) + r(A) = r(A)$$

(z tw. Cayleya-Hamiltona). W ten sposób, wielomian dowolnego stopnia od A sprowadza się do obliczenia wielomianu stopnia nie większego niż $n - 1$ od A.

Jak znaleźć współczynniki wielomianu r? Podstawmy do równości (9) kolejne pierwiastki wielomianu w_A: $\lambda_1, \ldots, \lambda_r$, tzn. wartości własne operatora A. Mamy wtedy równości:

$$f(\lambda_i) = r(\lambda_i), \qquad i = 1, \ldots, r \tag{10}$$

(gdy mamy pierwiastki wielokrotne, np. λ_i ma krotność n_i, to wstawiamy je do równości (10) różniczkowanej kolejno aż do $n_i - 1$).

Uwaga. *Metoda pracuje także dla funkcji analitycznych f nie będących wielomianami.* Dowód nie jest trudny, ale wymaga składników analitycznych i dlatego go pominiemy.

Przykład. Obliczmy n-tą potęgę macierzy A zdefiniowanej przez (7). Mamy: $w_A(\lambda) = \det(A - \lambda I) = \lambda^2 - \lambda - 1$, którego pierwiastki są: $\lambda_1 = \frac{1-\sqrt{5}}{2}, \lambda_2 = \frac{1+\sqrt{5}}{2}$. Wymiar macierzy (= stopień wielomianu charakterystycznego) jest równy 2. Możemy więc napisać:

$$A^n = a\,I + b\,A,$$

a układ równań (10) na współczynniki a, b będzie miał postać:

$$\begin{cases} \lambda_1^n &= a + b\,\lambda_1 \\ \lambda_2^n &= a + b\,\lambda_2 \end{cases}$$

Rozwiązaniem powyższego układu równań jest: $a = \frac{\Lambda_{n-1}}{\sqrt{5}}, b = \frac{\Lambda_n}{\sqrt{5}}$, gdzie oznaczyliśmy: $\Lambda^n = \lambda_2^n - \lambda_1^n$. Mamy więc:

$$A^n = a\,I + b\,A = \frac{\Lambda_{n-1}}{\sqrt{5}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{\Lambda_{n-1}}{\sqrt{5}} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} \Lambda_n + \Lambda_{n-1} & \Lambda_n \\ \Lambda_n & \Lambda n - 1 \end{bmatrix}$$

Zastosujmy otrzymane wyrażenie do znalezienia ogólnego wyrazu ciągu Fibonacciego. Mamy:

$$\begin{bmatrix} u_{n+1} \\ u_n \end{bmatrix} = A^n \begin{bmatrix} u_1 \\ u_0 \end{bmatrix} = A^n \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2\Lambda_n + \Lambda_{n-1} \\ \Lambda_n + \Lambda_{n-1} \end{bmatrix}$$

Biorąc np. drugą składową powyższego wektora, otrzymujemy po kilku przekształceniach:

$$u_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^{n+1} - \left(\frac{1-\sqrt{5}}{2} \right)^{n+1} \right]$$

Metoda 2. *Rozkład dowolnego wektora na składowe z podprzestrzeni pierwiastkowych.*

Niech przestrzeniami pierwiastkowymi operatora A będą $V_1, \ldots, V_r$ – każda z nich odpowiadająca wartości własnej λ_i. Chcemy znaleźć działanie funkcji od operatora na wektor: $f(A)v$. W tym celu, rozłóżmy najsampierw wektor na składowe należące do poszczególnych podprzestrzeni pierwiastkowych:

$$v = v_1 + v_2 + \ldots + v_r.$$

Mamy:

$$f(A)v = f(A)v_1 + f(A)v_2 + \ldots + f(A)v_r = \sum_{i=1}^{r} f(A)v_i. \tag{11}$$

Zapiszmy każdy ze składników tej sumy w następujący sposób:

$$f(A)v_i = f(A - \lambda_i I + \lambda_i I)v_i.$$

Rozwińmy powyższe wyrażenie w szereg Taylora wokół $\lambda_i I$. W tym celu, napiszmy najpierw szereg dla zespolonych wartości argumentu $\lambda \in \mathbf{C}$:

$$f(\lambda) = f((\lambda - \lambda_i) + \lambda_i) = f(\lambda_i) + f'(\lambda_i)(\lambda - \lambda_i) + \frac{f''(\lambda_i)}{2!}(\lambda - \lambda_i)^2 + \ldots,$$

a następnie zapiszmy analogiczne wyrażenie dla argumentu macierzowego:

$$f(A)v_i = f((A-\lambda_i I)+\lambda_i I)v_i = f(\lambda_i I)v_i + f'(\lambda_i I)(A-\lambda_i I)v_i + \frac{f''(\lambda_i I)}{2!}(A-\lambda_i I)^2 v_i + \dots$$

$$+\frac{f^{(n_i-1)}(\lambda_i I)}{(n_i-1)!}(A-\lambda_i I)^{n_i-1}v_i + \frac{f^{(n_i)}(\lambda_i I)}{n_i!}(A-\lambda_i I)^{n_i}v_i + \dots; \tag{12}$$

patrząc na powyższe wyrażenie, widzimy, iż $(A-\lambda_i I)^{n_i}v_i$ znika, ponieważ $v_i \in \mathrm{Ker}(A-\lambda_i I)^{n_i}$, tak więc szereg (12) obrywa się na (n_i-1)-wszej potędze operatora $(A-\lambda_i I)$.

Uwagi.

1. Jeśli A jest diagonalizowalna, to

$$f(A)v = \sum_{i=1}^{r} f(\lambda_i)v_i.$$

2. Metoda jest szczególnie efektywna w przypadku, gdy mamy do czynienia z koniecznością policzenia działania $f(A)$ na konkretny wektor.

Przykład. Obliczmy działanie macierzy e^{At} ($t \in \mathbf{R}$) na wektor $x_0 = [1,1,1]^T$, gdzie A jest macierzą

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

(gdy to zrobimy, będziemy mogli od razu napisać rozwiązanie równania różniczkowego $\dot{x} = Ax$ z warunkiem początkowym $x(0) = x_0$).

Wartości własne macierzy At są: $\lambda_1 = t, \lambda_2 = -t$ o krotnościach $n_1 = 1, n_2 = 2$. Podprzestrzeń pierwiastkowa V_1 odpowiadająca λ_1 jest rozpięta przez wektor $v_1 = [-1,-2,1]^T$ (jest to jednocześnie wektor własny), zaś podprzestrzeń pierwiastkowa V_2 odpowiadająca λ_2 jest rozpięta przez wektory: $v_2 = [1,0,0]^T, v_3 = [0,-1,1]^T$.

Rozkładamy wektor x_0 na składowe x_1, x_2 należące do podprzestrzeni V_1, V_2 odpowiednio: $x_1 = -2v_1 = [2,4,-2]^T$, $x_2 = -1v_2 + 3v_3 = [-1,-3,3]^T$. Mamy ogólnie, dla dowolnej funkcji analitycznej f:

$$f(At)x_0 = f(At)x_1 + f(At)x_2 = f(\lambda_1 I + (At-\lambda_1 I))x_1 + f(\lambda_2 I + (At-\lambda_2 I))x_2$$
$$= f(\lambda_1 I)x_1 + [f(\lambda_2 I) + f'(\lambda_2 I)(At-\lambda_2 I)]x_2$$

(wszystkie wyższe człony rozwinięcia będą znikać z uwagi na fakt, iż $x_1 \in \mathrm{Ker}(At-\lambda_1 I)$, $x_2 \in \mathrm{Ker}(At-\lambda_2 I)^2$). Biorąc teraz pod uwagę, że $f(\lambda_i I) = f(\lambda_i)I$ oraz konkretyzując teraz postać funkcji f: $f(x) = e^x$, a także fakt, iż $(e^x)' = e^x$, otrzymujemy:

$$e^{At}x_0 = e^{\lambda_1 t}v_1 + e^{\lambda_2 t}v_2 + e^{\lambda_2 t}(At-\lambda_2 I)v_2$$

$$= e^t v_1 + e^{-t}v_2 + e^{-t}\begin{bmatrix} 1 & 2t & 3t \\ t & 1+3t & 5t \\ -t & -2t & 1-4t \end{bmatrix} v_2 = \begin{bmatrix} 2e^t + e^{-t}(-2+3t) \\ 4e^t + e^{-t}(-6+5t) \\ -2e^t + e^{-t}(6-5t) \end{bmatrix}$$