STRUKTURA ENDOMORFIZMU C.D.

POWTÓRZENIE Z POPRZEDNICH ZAJĘĆ: Wykład zaczniemy od przypomnienia pojęć, które pojawiły się w przykładzie omawianym w trakcie poprzedniego wykładu. Dodamy także kilka definicji, które poprzednio się nie pojawiły.

Pracujemy w przestrzeni endomorfizmów liniowych, tzn. odwzorowań liniowych z przestrzeni V do przestrzeni V. Dla wygody wprowadzamy skrócone oznaczenie $L(V,V) = End(V)$. Zbiór $End(V)$ ma strukturę nieprzemiennej algebry z jedynką. Algebra jest to przestrzeń wektorowa z działaniem (zazwyczaj nazywanym mnożeniem), które jest liniowe ze względu na każdy argument. Zapiszmy warunki: $\mathcal{A}$ - przestrzeń wektorowa, $\circ : \mathcal{A} \times \mathcal{A} \longrightarrow \mathcal{A}$ mnożenie spełniające

$$(\lambda_1 a_1 + \lambda_2 a_2) \circ b = \lambda_1 a_1 \cdot b + \lambda_2 a_2 \cdot b, \quad a(\lambda_1 b_1 + \lambda_2 b_2) = \lambda_1 a b_1 + \lambda_2 a b_2$$

dla dowolnych $a, a_1, a_2, b, b_1, b_2 \in \mathcal{A}$, $\lambda_1, \lambda_2 \in \mathbb{K}$. Mówimy, że algebra jest przemienna jeśli $ab = ba$. Algebra jest algebrą z jedynką jeśli istnieje element $1 \in \mathcal{A}$ taki, że $1 \cdot a = a \cdot 1 = a$ dla dowolnego $a \in \mathcal{A}$.

Działaniem mnożenia w $End(V)$ jest składanie odwzorowań, jedynką jest odwzorowanie identycznościowe. Przestrzeń macierzy kwadratowych $n \times n$ o współczynnikach z ciała $\mathbb{K}$ to oczywiście $End(\mathbb{K}^n) = \mathbb{K}^n_n$. Wiadomo, że jeśli w przestrzeni V wybierzemy bazę e to każdemu elementowi $T \in End(V)$ możemy przypisać macierz $[T]^e_e \in End(\mathbb{K}^n)$. Odwzorowanie

$$End(V) \ni T \longmapsto [T]^e_e \in End(\mathbb{K}^n)$$ jest homomorfizmem algebr, tzn. jest liniowe i zachowuje mnożenie $[T \cdot S]^e_e = [T]^e_e [S]^e_e$ i jedynkę $[id]^e_e = \mathbb{1}$.

Zauważmy, że nie wszystkie algebry mają jedynkę. W $\mathbb{K}^n_n$ możemy wprowadzić inną strukturę algebry: $(X, Y) \longmapsto [X, Y] = XY - YX$.

komutator macierzy

$(\mathbb{K}^n_n, [\cdot,\cdot])$ jest algebrą, choć w tym przypadku działanie nazywa się raczej nawiasem niż mnożeniem. Nawias jest antysymetryczny, tzn $[X,Y]=-[Y,X]$. Algebra $(\mathbb{K}^n_n, \cdot)$ jest łączna, tzn $(X\cdot Y)\cdot Z = X\cdot(Y\cdot Z)$. Algebra $(\mathbb{K}^n, [\cdot,\cdot])$ nie jest łączna. nawias spełnia za to warunek $[X,[Y,Z]] = [[X,Y],Z] + [Y,[X,Z]]$, zwany tożsamością Jacobiego. Jeśli działanie w algebrze spełnia tożsamość Jacobiego i jest antysymetryczne nazywamy tę algebrę algebrą Liego.

Wracamy do algebry endomorfizmów z mnożeniem. Elementy tej algebry nadają się dobrze do podstawiania jako argumenty wielomianu nad odpowiednim ciałem:

$w \in \mathbb{K}_r[\cdot]$ $\quad w(t) = a_r t^r + a_{r-1}t^{r-1} + \dots + a_1 t + a_0$

$$T \in \mathrm{End}(V) \quad w(T) = a_r T^r + a_{r-1}T^{r-1} + \dots + a_1 T + a_0 \cdot \mathbb{1}$$

Wektorem własnym operatora $T \in \mathrm{End}(V)$ nazywamy niezerowy wektor $v \in V$ taki, że dla pewnej liczby $\lambda \in \mathbb{K}$ $\; Tv = \lambda v$. Liczbę tę nazywamy wartością własną operatora T. Zbiór wszystkich wartości własnych T nazywamy spektrum T i oznaczamy $Sp(T)$.

Wartości własne znajdujemy rozwiązując równanie $\det(T - \lambda\mathbb{1}) = 0$. równanie to nazywa się równaniem charakterystycznym operatora T. Równanie to ma charakter wielomianowy. Wielomian $w_T(\lambda) = \det(T-\lambda\mathbb{1})$ jest wielomianem stopnia równego wymiarowi przestrzeni V. Uwaga: na pierwszy rzut oka nie jest w ogóle jasne, że pojęcie wyznacznika ma sens dla operatorów w przestrzeni innej niż $\mathbb{K}^n$. Przecież zmiana bazy w której zapisujemy macierz operatora powoduje zmianę tej macierzy. Jeśli wyznacznik też się zmienia, to definicja w_T nie ma sensu! Tak jednak nie jest: Niech f i e będą dwiema bazami w V a $\mathcal{U} = [id]^f_e$ macierzą przejścia. Wówczas

$$[T]^f_f = [id]^f_e [T]^e_e [id]^e_f = \mathcal{U}\cdot[T]^e_e\cdot\mathcal{U}^{-1}$$

$$\det\left([T]^f_f\right) = \det\left(\mathcal{U}\cdot[T]^e_e\cdot\mathcal{U}^{-1}\right) = \det\mathcal{U}\cdot\det[T]^e_e\cdot\det\mathcal{U}^{-1} = \det\mathcal{U}\cdot\frac{1}{\det\mathcal{U}}\cdot\det[T]^e_e =$$

$$= \det[T]^e_e$$

Wyznacznik operatora jest dobrze zdefiniowany, tzn nie zależy od wyboru bazy!

Wielomian charakterystyczny $w_T(\lambda) = \det(T-\lambda \mathbb{1})$ jest wielomianem stopnia równego wymiarowi przestrzeni w której działa operator T. Oznacza to, że operator T ma n wartości własnych licząc z krotnościami, tzn.: krotności

$$w_T(\lambda) = (\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k_2}\dots(\lambda-\lambda_r)^{k_r} \quad i \quad k_1+k_2+\dots+k_r = n.$$

Jeśli operator T działa na przestrzeni zespolonej, wówczas zarówno współczynniki w_T jak i elementy Sp_T są liczbami zespolonymi. Jeśli zaś T działa na przestrzeni rzeczywistej, to współczynniki w_T są rzeczywiste ale wartości własne mogą nadal być zespolone. Prowadzi to do pewnych komplikacji. Jeśli jednak T jest rzeczywisty, a pewna jego wartość własna jest zespolona, to także liczba do niej sprzężona jest wartością własną T z tą samą krotnością.

Wektory własne odpowiadające wartości własnej λ znajdujemy rozwiązując równanie $(T-\lambda\mathbb{1})v=0$, czyli

$$V_\lambda = \ker(T-\lambda\mathbb{1})$$

wartość własna

przestrzeń własna

Z definicji $Sp(T)$ wynika, że $\dim V_\lambda \geq 1$ jeśli T zespolony i jeśli T rzeczywisty i λ rzeczywista. Jeśli T rzeczywista i λ jest zespolona wektorów własnych w przestrzeni V odpowiadających λ nie ma.

Podprzestrzenie własne są niezmiennicze, tzn $TV_\lambda = V_\lambda$. Dokładniej, podprzestrzenią niezmienniczą dla T nazywamy podprzestrzeń $U \subset V$ taką, że $TU \subseteq U$. Dla podprzestrzeni własnych mamy równość.

Naszym celem w dalszym ciągu będzie znalezienie rozkładu V na sumę prostą podprzestrzeni niezmienniczych dla danego operatora T. Będziemy też nalegać, żeby podprzestrzenie te były możliwie małego wymiaru. Po co to? Załóżmy, że $V = U_1 \oplus U_2$ i U_1, U_2 są niezmiennicze. Weźmy też bazę e przestrzeni V zgodną z powyższym rozkładem tzn. $e = (\underbrace{e_1, \dots, e_k}_{U_1}, \underbrace{e_{k+1} \dots e_n}_{U_2})$ wtedy macierz $[T]_e^e$ ma postać blokową:

$$[T]_e^e = \begin{array}{cc} & \begin{array}{cc} \overbrace{\quad\quad}^{k} & \quad\quad \end{array} \\ \begin{array}{c} k\{ \\ \\ \end{array} & \left[\begin{array}{c|c} * & 0 \\ \hline 0 & * \end{array}\right] \begin{array}{c} \\ \}\,n-k \end{array} \\ & \begin{array}{cc} \quad\quad & \underbrace{\quad\quad}_{n-k} \end{array} \end{array}$$

Znalezienie rozkładu na podprzestrzenie niezmiennicze jest więc krokiem w kierunku znalezienia bazy w której operator ma prostą postać. W najprostszym przypadku – gdy każda z podprzestrzeni jest jednowymiarowa – operator jest diagonalny.

Przygotujemy się teraz do omówienia sytuacji ogólnej, tzn. gdy nie istnieje baza złożona z wektorów własnych operatora T.

OPERATORY RZUTOWE:

Załóżmy, że $V = U_1 \oplus U_2$. Z takim rozkładem stowarzyszyć można dwa endomorfizmy $P_1, P_2 \in \operatorname{End}(V)$. Jeśli $V = U_1 \oplus U_2$, to każdy wektor $v \in V$ daje się jednoznacznie rozłożyć na składowe $v_1 \in U_1$ i $v_2 \in U_2$, $v = v_1 + v_2$. Definiujemy

$$P_1(v) = v_1, \quad P_2(v) = v_2$$

rzut na U_1 wzdłuż U_2

rzut na U_2 wzdłuż U_1

Endomorfizmy P_1 i P_2 mają następujące własności:

$$P_1 + P_2 = \mathbb{1}, \quad P_1^2 = P_1, \quad P_2^2 = P_2, \quad P_1 P_2 = P_2 P_1 = 0$$

Własności te można sprawdzić bezpośrednim rachunkiem. Okazuje się, że jest też odwrotnie: Para operatorów P_1 i P_2 spełniająca powyższe warunki prowadzi do rozkładu przestrzeni V na sumę prostą. Widać, że tak naprawdę wystarczy zadać jeden operator $P \in \operatorname{End}(V)$ spełniający warunek $P^2 = P$.

FAKT: Niech $P \in \operatorname{End}(V)$ spełnia $P^2 = P$. Wtedy także $\mathbb{1} - P$ spełnia $(\mathbb{1} - P)^2 = (\mathbb{1} - P)$. Ponadto $P(\mathbb{1} - P) = (\mathbb{1} - P)P = 0$ oraz $V = U_1 \oplus U_2$ gdzie $U_1 = \operatorname{im} P = \ker(\mathbb{1} - P)$, $U_2 = \operatorname{im}(\mathbb{1} - P) = \ker P$.

DOWÓD: Sprawdźmy najpierw warunki algebraiczne:

$$P(\mathbb{1}-P)=P-P^2=P-P=0,\quad (\mathbb{1}-P)P=P-P^2=P-P=0$$

$$(\mathbb{1}-P)^2=(\mathbb{1}-P)(\mathbb{1}-P)=\mathbb{1}-P-P+P^2=\mathbb{1}-P-P+P=\mathbb{1}-P.$$

Oznaczmy teraz $U_1=\operatorname{im} P$ i $U_2=\operatorname{im}(\mathbb{1}-P)$. Mamy $\mathbb{1}=P+(\mathbb{1}-P)$, czyli

$$v=\mathbb{1}\cdot v=(P+(\mathbb{1}-P))v=\underbrace{Pv}_{\in U_1}+\underbrace{(\mathbb{1}-P)v}_{\in U_2}$$

wynika z tego, że $V=U_1+U_2$.

Niech teraz $u\in U_1\cap U_2$. Oznacza to, że istnieje $v_1\in V$ takie, że $u=Pv_1$ oraz istnieje $v_2\in V$ takie, że $u=(\mathbb{1}-P)v_2$

$$u=\underbrace{Pv_1=(\mathbb{1}-P)v_2}$$

$$\underbrace{P^2v_1}_{Pv_1}=\overbrace{P(\mathbb{1}-P)}^{0}v_2=0$$

$$\Rightarrow \underbrace{Pv_1}_{=u}=0$$

zatem $U_1\cap U_2=\{0\}$

Istotnie więc $V=U_1\oplus U_2$. Dodatkowo, jeśli $u\in U_2$, tzn $u=(\mathbb{1}-P)v$ to $Pu=P(\mathbb{1}-P)v=0$ zatem $u\in\ker P$. Wykazaliśmy, że $U_2\subset\ker P$. Z drugiej strony jeśli $u\in\ker P$, tzn $Pu=0$ to $u=u-0=u-Pu=(\mathbb{1}-P)u$, tzn $u\in\operatorname{im}(\mathbb{1}-P)=U_2$. Wykazaliśmy, że $\ker P\subset U_2$. Ostatecznie więc $\ker P=U_2$. Podobnie $\ker(\mathbb{1}-P)=U_1$. ■

Powyższy fakt stanowi uzasadnienie definicji pojęcia **operatora rzutowego**. Operator $P\in\operatorname{End}(V)$ nazywa się **operatorem rzutowym** lub **rzutem** jeśli $P^2=P$.

Powyższy fakt można uogólnić na większą liczbę składowych:

FAKT: Niech $P_1,\dots,P_k$ będą operatorami spełniającymi $P_1+P_2+\dots+P_k=\mathbb{1}$, $P_iP_j=0$ dla $i\neq j$. Wtedy P_i są rzutami oraz

$$V=V_1\oplus V_2\oplus\dots\oplus V_k \quad \text{dla } V_i=\operatorname{im}P_i$$

DOWÓD: Udowodnimy jedynie, że P_i są rzutami. Reszta dowodu jest jak dla $k=2$.

$$\mathbb{1} = P_1 + P_2 + \dots + P_i + \dots + P_k \qquad P_i = P_i \cdot \mathbb{1} = P_i\left(P_1 + \dots P_i + \dots P_k\right) = \underbrace{P_iP_1}_{0} + \underbrace{P_iP_2}_{0} + \dots +$$

$$+ P_i^2 + \underbrace{P_iP_{i+1}}_{0} + \dots + \underbrace{P_iP_k}_{0} = P_i^2 \quad \blacksquare$$

Układ operatorów rzutowych spełniających warunki $P_iP_j = 0$, $i \neq j$ oraz $\mathbb{1} = P_1 + \dots + P_k$ nazywamy **rozkładem jedynki.**

Wracamy teraz do naszego operatora T i poszukiwania jego podprzestrzeni niezmienniczych. Wiemy już, że rozkład na sumę prostą podprzestrzeni odpowiada rzutowemu rozkładowi jedynki. Ciekawe jak w języku rzutów wyrazić warunek, że podprzestrzenie są niezmiennicze dla danego operatora?

Niech więc $T \in \mathrm{End}(V)$, $V = U_1 \oplus \dots \oplus U_k$, U_i niezmiennicze, P_i - odpowiedni układ rzutów. Niech także $v_i \in U_i$

Weźmy dowolny wektor v i rozłóżmy według sumy prostej:

$$v = v_1 + v_2 + \dots + v_k$$

Liczymy

$$P_iTv = PT(v_1 + v_2 + \dots + v_i + \dots + v_k) = P(\underset{\in U_1}{Tv_1} + \dots + \underset{\in U_i}{Tv_i} + \dots \underset{\in U_k}{Tv_k}) =$$

$$0 + \dots + Tv_i + \dots + 0 = Tv_i$$

$$TP_iv = Tv_i$$

$$P_iT(v) = TP_i(v)$$

Wektor $v \in V$ był dowolny, zatem P_i i T są przemienne, tzn **komutują.** $[P_i, T] = 0.$

Odwrotnie, weźmy teraz rzutowy rozkład jedności taki, że $[P_i,T]=0$ dla wszystkich $i\in\{1,\dots,k\}$, niech $U_i = \operatorname{im} P_i$. Jeśli $v_i \in U_i$ to:

$$0=[P_i,T]u_i = P_i T(u_i) - T\underbrace{P_i(u_i)}_{u_i} = P_i(T(u_i)) - T(u_i)$$

$$P_i(T(u_i)) = T(u_i) \quad \text{tzn} \quad T(u_i)\in \operatorname{im} P_i = U_i$$

Przestrzeń U_i jest więc niezmiennicza dla T. Warunek, aby dany rozkład na sumę prostą był rozkładem na podprzestrzenie niezmiennicze wyrażony w języku rzutów ma postać $[T,P_i]=0$.

PRZYKŁAD:

$$T = \begin{bmatrix} 3 & -2 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix} \qquad T \in \mathrm{End}(\mathbb{R}^3) \qquad \varphi(\lambda) = (\lambda-1)^2(\lambda-2)$$

Operator T spełnia warunek $\varphi(T)=0$. φ jest z dokładnością do znaku równy wielomianowi charakterystycznemu operatora T. Twierdzenie, które mówi, że $w_T(T)=0$ nosi nazwę Twierdzenia Cayleya Hamiltona. Dowodem tego twierdzenia zajmiemy się później.

Oznaczamy $\varphi_1(\lambda) = (\lambda-1)^2$, $\varphi_2(\lambda) = (\lambda-2)$. Wielomiany te są względnie pierwsze, zatem korzystając z algorytmu Euklidesa możemy wyznaczyć takie wielomiany w_1 i w_2 że $w_1\varphi_1 + w_2\varphi_2 = 1$.

$\lambda^2-2\lambda+1$	$\lambda-2$	1
λ		
λ	1	0

$$1\cdot(\lambda^2-2\lambda+1) - \lambda(\lambda-2) = 1$$

$$w_1(\lambda) = 1 \quad w_2(\lambda) = -\lambda$$

Do równania wielomianowego $\varphi_1 w_1 + \varphi_2 w_2 = 1$ podstawmy operator T:

$$\mathbb{1} = w_1(T)\varphi_1(T) + w_2(T)\varphi_2(T) = \underbrace{(T-\mathbb{1})^2}_{P_1} + \underbrace{(-T)(T-2\mathbb{1})}_{P_2} = P_1 + P_2$$

$$P_1 = (T-\mathbb{1})^2$$
$$P_2 = 2T - T^2$$

Zbierzmy własności operatorów P_1, P_2: (1) $P_1+P_2 = \mathbb{1}$, (2) $P_1 \cdot P_2 = w_1(T)\varphi_1(T)\cdot w_2(T)\varphi_2(T) = w_1(T)w_2(T)\varphi_1(T)\varphi_2(T) = w_1(T)w_2(T)\varphi(T) = 0$ podobnie $P_2P_1 = 0$. W takim przypadku zgodnie z udowodnionym przez nas faktem rozkład $P_1+P_2=\mathbb{1}$ jest rzutowym rozkładem jedności. Dodatkowo, ponieważ P_1 i P_2 są wielomianami od T, to P_1 i P_2 są przemienne z T. Rozkład $U_1 \oplus U_2 = \mathbb{R}^3$ odpowiadający $P_1+P_2=\mathbb{1}$ jest więc rozkładem na podprzestrzenie niezmiennicze względem T.

Wyznaczmy operatory P_1 i P_2 explicite:

$$\underbrace{(T-\mathbb{1})^2}_{P_1} + \underbrace{(-T)(T-2\mathbb{1})}_{P_2} = P$$

$$P_1 = (T-\mathbb{1})^2 = \begin{bmatrix} 2 & -2 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & 0 \end{bmatrix}^2 = \begin{bmatrix} 3 & -3 & 0 \\ 2 & -2 & 0 \\ 1 & -1 & 0 \end{bmatrix}$$

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

Przestrzenie $U_1 = \operatorname{im} P_1$ i $U_2 = \operatorname{im} P_2$ można znaleźć bezpośrednio badając odpowiednie operatory.

$$U_1 = \operatorname{im} P_1 = \left\langle \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \right\rangle \qquad \operatorname{im} P_2 = \left\langle \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\rangle = U_2$$

Rozkład $\mathbb{R}^3 = U_1 \oplus U_2$ jest rozkładem $\mathbb{R}^3$ na podprzestrzenie niezmiennicze dla operatora T. U_1 jest podprzestrzenią własną V_2, natomiast U_2 nie jest podprzestrzenią własną. Zawiera jednak podprzestrzeń własną V_1.

Istnieje szybszy sposób znajdowania podprzestrzeni niezmienniczych bez konieczności wypisywania postaci rzutów. Rozpatrzmy następującą sytuację:

$$\varphi(T) = 0 \qquad \varphi(\lambda) = (\lambda-\lambda_1)^{k_1} \cdots (\lambda-\lambda_r)^{k_r}$$

$$\varphi_i(\lambda) = \frac{\varphi(\lambda)}{(\lambda-\lambda_i)^{k_i}} \qquad 1 = w_1(\lambda)\varphi_1(\lambda) + \dots + w_r(\lambda)\varphi_r(\lambda)$$

$P_i = w_i(T)\varphi_i(T)$ FAKT: $\operatorname{im} P_i = \mathcal{U}_i = \ker(T-\lambda_i \mathbb{1})^{k_i}$

„P_i"

Dowód:

Dowodzimy, że $\operatorname{im} w_i(T)\varphi_i(T) = \ker(T-\lambda_i\mathbb{1})^{k_i}$

Niech więc $u \in \operatorname{im} P_i$. Oznacza to, że $u = P_i(v)$ dla pewnego $v \in V$. Obliczmy $(T-\lambda_i\mathbb{1})^{k_i}(u)$:

$$(T-\lambda_i\mathbb{1})^{k_i}(u) = (T-\lambda_i\mathbb{1})^{k_i} w_i(T)\varphi_i(T)v = 0$$

$\varphi(T) = 0$

zatem $\operatorname{im} P_i \subset \ker(T-\lambda\mathbb{1})^{k_i}$

Z drugiej strony niech $u \in \ker(T-\lambda\mathbb{1})^{k_i}$, tzn $(T-\lambda\mathbb{1})^{k_i}u = 0$. Wiadomo jednak, że $\mathbb{1} = P_1 + \dots + P_r$

zatem $u = P_1 u + P_2 u + \dots + P_i u + \dots + P_r u$

Weźmy $j \neq i$ wtedy $P_j = w_j(T)\varphi_j(T)$ a φ_j zawiera czynnik $(\lambda-\lambda_i)^{k_i}$, wobec tego $P_j u = 0$. W powyższej sumie zostanie jedynie i-ty składnik, tzn

$u = P_i u$, tzn $u \in \operatorname{im} P_i$

wykazaliśmy więc, że $\ker(T-\lambda_i\mathbb{1})^{m_i} \subset \operatorname{im} P_i$.

Ponieważ zachodzę zawieranie w obie strony, to znaczy, że mamy do czynienia z równością. ∎

Podsumujmy: Mając wielomian zerujący dany operator T możemy znaleźć rozkład V na podprzestrzenie niezmiennicze w przypadku, kiedy (1) przestrzeń V jest nad $\mathbb{C}$, (2) przestrzeń V jest nad $\mathbb{R}$ i wszystkie pierwiastki wielomianu są rzeczywiste. Mamy wtedy:

$$\varphi(T)=0 \qquad \varphi(\lambda)=(\lambda-\lambda_1)^{k_1}\cdot\ldots\cdot(\lambda-\lambda_r)^{k_r}$$

$$\varphi_i(\lambda)=\frac{\varphi(\lambda)}{(\lambda-\lambda_i)^{k_i}} \qquad 1=\varphi_1(\lambda)\omega_1(\lambda)+\cdots+\varphi_r(\lambda)\omega_r(\lambda)$$

$$P_i=\omega_i(T)\varphi_i(T) \qquad \mathbb{1}=P_1+P_2+\cdots+P_r \qquad P_iP_j=0 \;;\; P_i^2=P_i$$

$$\mathcal{U}_i=\operatorname{im}P_i=\ker(T-\lambda_i\mathbb{1})^{k_i} \qquad V=\mathcal{U}_1\oplus\mathcal{U}_2\oplus\cdots+\mathcal{U}_r$$

Macierz T w bazie dostosowanej do rozkładu ma postać blokową.

Definicja: Jeśli $\varphi(\lambda)=\omega_T(\lambda)$ to podprzestrzeń $\mathcal{U}_i$ utworzona w powyższy sposób nazywa się podprzestrzenią pierwiastkową dla wartości własnej λ_i tzn

$$W_{\lambda_i}=\ker(T-\lambda_i\mathbb{1})^{k_i} \quad \text{gdzie } \lambda_i\in Sp(T)\; k_i \text{ jest krotnością.}$$

Do załatwienia zostało: (1) Dowód tw. Cayley'a Hamiltona, (2) Zbadanie działania T na każdej z podprzestrzeni niezmienniczych, (3) Zbadanie przypadku operatora na rzeczywistej przestrzeni mającego zespolone wartości własne.

TWIERDZENIE CAYLEY'A HAMILTONA

Niech $\omega_T(\lambda) = \det(T - \lambda \mathbb{1})$ wtedy $\omega_T(T) = 0$

DOWÓD: Dowód przeprowadzimy dla macierzy. Jest to wystarczające, ponieważ pokazaliśmy wcześniej, że wyznacznik endomorfizmu jest dobrze określony (nie zależy od wyboru bazy, w której zapisujemy endomorfizm jako macierz).

Wiadomo, że dla $X \in \mathbb{K}^n_n$ zachodzi wzór $X^D \cdot X = (\det X)\mathbb{1}$

Jako X weźmiemy $T - \lambda\mathbb{1}$:

$X = T - \lambda \mathbb{1}$
$$\underbrace{(T - \lambda \mathbb{1})^D}_{\text{macierz, której wyrazy są wielomianami stopnia } n-1} \cdot (T - \lambda\mathbb{1}) = \underbrace{\omega_T(\lambda)}_{\text{wielomian stopnia } n} \mathbb{1}$$

Oto przykładowa macierz z wielomianowymi wyrazami:

$$\begin{bmatrix} \lambda^2+1 & 2\lambda^2+\lambda-3 \\ \lambda+2 & \lambda^2-\lambda \end{bmatrix}$$ można ją przepisać następująco:

$$\lambda^2 \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} + 1 \cdot \begin{bmatrix} 1 & -3 \\ 2 & 0 \end{bmatrix}$$ tzn. jako wielomian o współczynnikach macierzowych

Wobec tego

$$(T - \lambda\mathbb{1})^D = B_0 + B_1\lambda + B_2\lambda^2 + \dots + B_{n-1}\lambda^{n-1}$$

Liczymy dalej:

$$(T-\lambda\mathbb{1})^D(T-\lambda\mathbb{1}) = (B_0 + B_1\lambda + \dots + B_{n-1}\lambda^{n-1})(T - \lambda\mathbb{1}) = B_0T + B_1T\lambda + \dots + B_{n-1}T\lambda^{n-1}$$
$$- B_0\lambda - B_1\lambda^2 - \dots - B_{n-1}\lambda^n =$$

$$(T-\lambda \mathbb{1})^D (T-\lambda \mathbb{1}) = \left(B_0 + B_1\lambda + \dots + B_{n-1}\lambda^{n-1}\right)(T-\lambda \mathbb{1}) = B_0T + B_1T\lambda + \dots + B_{n-1}T\lambda^{n-1}$$

$$- B_0\lambda - B_1\lambda^2 - \dots - B_{n-1}\lambda^n =$$

$$B_0T + \lambda(B_1T - B_0) + \lambda^2(B_2T - B_1) + \dots + \lambda^{n-1}(B_{n-1}T - B_{n-2}) - \lambda^n B_{n-1} =$$

$$c_0\mathbb{1} + c_1\lambda\mathbb{1} + c_2\lambda^2\mathbb{1} + \dots + c_n\lambda^n\mathbb{1} = \omega_T(\lambda)\cdot\mathbb{1}$$

porównujemy wyrazy przy jednakowych potęgach

$$\begin{aligned}
B_0T &= c_0\mathbb{1} \\
B_1T - B_0 &= c_1\mathbb{1} \qquad /\cdot T \\
B_2T - B_1 &= c_2\mathbb{1} \qquad /\cdot T^2 \\
&\vdots \\
B_{n-1}T - B_{n-2} &= c_{n-1}\mathbb{1} \qquad /\cdot T^{n-1} \\
-B_{n-1} &= c_n\mathbb{1} \qquad /\cdot T^n
\end{aligned}$$

$$\begin{aligned}
B_0T &= c_0\mathbb{1} \\
B_1T^2 - B_0T &= c_1T \\
B_2T^3 - B_1T^2 &= c_2T^2 \\
\vdots \qquad &\vdots \qquad \vdots \\
B_{n-1}T^n - B_{n-2}T^{n-1} &= c_{n-1}T^{n-1} \\
+ \qquad -B_{n-1}T^n &= c_nT^n \\
\hline
\end{aligned}$$

$$0 = c_0\mathbb{1} + c_1T + \dots + c_{n-1}T^{n-1} + c_nT^n = \omega_T(T)$$

∎

We wszystkich rachunkach wymagających użycia wielomianu zerującego operator można więc użyć wielomianu charakterystycznego.

W niektórych szczególnych przypadkach istnieje wielomian zerujący niższego stopnia. O tym innym razem.

DZIAŁANIE OPERATORA NA PODPRZESTRZENI PIERWIASTKOWEJ. OPERATORY NILPOTENTNE.

Potrafimy już w wielu przypadkach podzielić przestrzeń na sumę prostą podprzestrzeni pierwiastkowych, niezmienniczych dla operatora. Dalej zajmiemy się poszukiwaniem postaci operatora obciętego do każdej z podprzestrzeni pierwiastkowych. W szczególności oznacza to także znalezienie bazy w której operator ma prostą postać.

Potrzebne do tego będzie pojęcie operatora nilpotentnego.

Definicja: Niech V będzie przestrzenią wektorową wymiaru n. Operator $N \in End(V)$ nazywamy nilpotentnym, jeśli $N^k = 0$ dla pewnego $k \in \mathbb{Z}$.

Fakt: Wielomian charakterystyczny operatora nilpotentnego ma postać $\omega_N(\lambda) = (-1)^n \lambda^n$.

Dowód: Udowodnimy to dla przestrzeni zespolonych. Dowód opiera się na istnieniu wektora własnego dla każdej wartości własnej, a to ma miejsce jedynie nad $\mathbb{C}$. Twierdzenie jest prawdziwe także dla rzeczywistych przestrzeni, jednak dowód wymaga dodatkowego namysłu.

Rozważmy więc $N \in End(V)$, V nad $\mathbb{C}$, $N^k = 0$. Jeśli $\lambda \in sp(N)$ to istnieje $v \neq 0$ i taki, że $Nv = \lambda v$, tzn $N^k v = \lambda^k v$ ale $N^k = 0$, więc $\lambda^k v = 0$. Wynika z tego, że $\lambda = 0$. Spektrum operatora nilpotentnego składa się więc jedynie z wartości 0: $Sp(N) = \{0\}$. Wobec tego $\omega_N(\lambda) = a_n \lambda^n$. Wiadomo jednak, że wyraz przy najwyższej potędze w wielomianie charakterystycznym to $(-1)^n$. ∎

Weźmy teraz dowolny operator T, którego wielomian charakterystyczny rozkłada się na czynniki liniowe nad $\mathbb{K}$. Znajdźmy rozkład V na podprzestrzenie pierwiastkowe i odpowiedni rzutowy rozkład jedności.

$$V = W_{\lambda_1} \oplus \cdots \oplus W_{\lambda_r}, \quad \mathbb{1} = P_1 + \cdots + P_r \qquad W_{\lambda_i} = \ker(T - \lambda_i \mathbb{1})^{k_i}$$

Zadaniem naszym jest znaleźć $\dim W_{\lambda_i}$. Operator $T - \lambda \mathbb{1}$ w bazie dostosowanej do rozkładu $V = W_{\lambda_1} \oplus \cdots \oplus W_{\lambda_r}$ ma postać blokową. Jego wyznacznik jest więc iloczynem wyznaczników podmacierzy:

$$w_T(\lambda) = \det(T - \lambda \mathbb{1}) = \prod_{i=1}^{r} \det\left(T|_{W_{\lambda_i}} - \lambda \mathbb{1}_{W_{\lambda_i}}\right) \qquad (*)$$

Oznaczmy dla skrótu $T_i = T|_{W_{\lambda_i}}$ i $\mathbb{1}_i = \mathbb{1}|_{W_{\lambda_i}}$. Obliczmy $\det(T_i - \lambda \mathbb{1}_{W_i})$. Uważamy teraz T_i za element $\operatorname{End}(W_{\lambda_i})$. Wiadomo, że $T_i - \lambda_i \mathbb{1}_i$ jest nilpotentny na W_{λ_i}. Z definicji bowiem $(T_i - \lambda_i \mathbb{1}_i)^{k_i} = 0$.

$$T_i - \lambda \mathbb{1}_i = T_i - \lambda_i \mathbb{1} + \lambda_i \mathbb{1}_i - \lambda \mathbb{1}_i = (T_i - \lambda_i \mathbb{1}) - (\lambda - \lambda_i)\mathbb{1}_i$$

$$\det(T_i - \lambda \mathbb{1}_i) = \det\left((T_i - \lambda_i \mathbb{1}) - (\lambda - \lambda_i)\mathbb{1}_i\right) = (\lambda - \lambda_i)^{n_i} (-1)^{n_i} \qquad n_i = \dim W_{\lambda_i}$$

wielomian charakterystyczny operatora nilpotentnego

wstawiamy do $(*)$

$$w_T(\lambda) = \prod_{i=1}^{r} \det(W_i - \lambda \mathbb{1}_i) = \prod_{i=1}^{r} (\lambda - \lambda_i)^{n_i} (-1)^{n_i} =$$

$$= (-1)^n (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \cdot \ldots \cdot (\lambda - \lambda_r)^{n_r}$$

Z porównania postaci $w_T(\lambda) = (-1)^n (\lambda - \lambda_1)^{k_1} \cdots (\lambda - \lambda_r)^{k_r}$ wynika, że $n_i = k_i$, czyli wymiar podprzestrzeni pierwiastkowej równy jest krotności wartości własnej.

Obserwacja: Niech $Sp(T)=\{\lambda_1,\dots,\lambda_r\}$ i niech V będzie nad $\mathbb{C}$ lub V nad $\mathbb{R}$ i wszystkie $\lambda_i \in \mathbb{R}$. Wtedy operator

$$N = T - \lambda_1 P_1 - \lambda_2 P_2 - \dots - \lambda_r P_r \quad \text{jest nilpotentny.}$$

Istotnie, $N|_{W_{\lambda_i}} = T_i - \lambda_i \mathbb{1}_i$ gdyż $P_i|_{W_{\lambda_i}} = \mathbb{1}_i$ zatem $N_i = N|_{W_{\lambda_i}}$ jest nilpotentny na W_{λ_i}. Podprzestrzenie W_{λ_i} są niezmiennicze dla N, zatem $N^{\max\{k_i\}} = 0$. Oznaczmy $D = \lambda_1 P_1 + \dots + \lambda_r P_r$. Operator D jest diagonalizowalny, a podprzestrzenie W_{λ_i} są dla niego podprzestrzeniami własnymi. Mamy więc

$$T = T - D + D = N + D \quad = \text{nilpotentny + diagonalizowalny.}$$

Ponadto $[N,D]=0$. Istotnie, T komutuje z każdym rzutem P_i, zatem komutuje także z ich kombinacją liniową. Naturalnie $[D,D]=0$ więc $[\underset{\overset{\uparrow}{N}}{T-D},D]=0$

Skonstruowaliśmy rozkład dowolnego operatora zespolonego i niektórych rzeczywistych na sumę komutujących nilpotentnego i diagonalizowalnego. To bywa wygodne.

W każdej przestrzeni W_{λ_i} znajdziemy teraz bazę wygodną dla operatora T.