LA FORMA DI JORDAN DI UN ENDOMORFISMO

FILIPPO BRACCI

1. POLINOMIO MINIMO E POLINOMIO CARATTERISTICO

Sia V uno spazio vettoriale di dimensione finita $N \geq 1$ su $\mathbb{K}$ con $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Indichiamo con End$(V)$ lo spazio delle applicazioni lineari da V in V. Sia $T \in$ End(V). Osserviamo che per un polinomio $p(x) \in \mathbb{K}[x]$ è ben definito $p(T)$ nel modo seguente: se $p(x) = a_0 + a_1 x + \ldots a_n x^n$ allora

$$p(T) = a_0 \mathrm{Id} + a_1 T + \ldots + a_n T^n \in \mathrm{End}(V),$$

dove $T^j = T \circ \ldots \circ T$ j-volte. Possiamo allora definire

$$I_T = \{p \in \mathbb{K}[x] : p(T) = 0\},$$

dove 0 è ovviamente l'operatore nullo che associa ad ogni $v \in V$ l'elemento neutro 0. Si prova subito che I_T è un ideale di $\mathbb{K}[x]$.

Lemma 1.1. *$I_T \neq \mathbb{K}[x]$ e $I_T \neq \{0\}$.*

Dimostrazione. È chiaro che $I_T \neq \mathbb{K}[x]$. Inoltre End(V) ha dimensione N^2. Dato che Id, $T, \ldots, T^{N^2}$ sono $N^2 + 1$ vettori di End(V) devono essere necessariamente linearmente indipendenti e dunque esistono $a_0, \ldots, a_{N^2} \in \mathbb{K}$ tali che

$$a_0 \mathrm{Id} + \ldots + a_{N^2} T^{N^2} = 0,$$

e pertanto il polinomio $a_0 + \ldots + a_{N^2} x^{N^2} \in I_T$ che dunque non è ridotto al solo 0. $\square$

Dato che $\mathbb{K}$ è un campo ne segue che $\mathbb{K}[x]$ è un anello euclideo e pertanto è a ideali principali (PID). In particolare esiste un (unico) polinomio monico $\mu_T(x) \in \mathbb{K}[x]$ tale che

$$I_T = \langle \mu_T(x) \rangle.$$

Definizione 1.2. Il polinomio $\mu_T(x)$ si dice il *polinomio minimo* di T.

Nota 1.3.
1. Sia R un automorfismo di V. Dato che $(RTR^{-1})^m = RT^m R^{-1}$, ne segue che $p(RTR^{-1}) = Rp(T)R^{-1}$. In particolare endomorfismi simili hanno lo stesso polinomio minimo.
2. Il polinomio minimo è caratterizzato dall'essere quel polinomio monico p di grado minimo tale che $p(T) = 0$.
3. End(V) è un $\mathbb{K}[x]$-modulo tramite la mappa $\mathbb{K}[x] \ni p \mapsto p(T) \in \mathrm{End}(V)$.

Fissiamo una base di V e associamo a T in tal base la $(N \times N)$-matrice A. In base all'osservazione 1.3.1 è possibile parlare di polinomio minimo della matrice A.

Definizione 1.4. Il *polinomio caratteristico* di un $T \in \mathrm{End}(V)$ è definito tramite
$$p_T(\lambda) = \det(\lambda \mathrm{Id} - T).$$
Le radici di $p_T(\lambda)$ si dicono gli *autovalori* di T.

Si dice che un vettore $v \in V \setminus \{0\}$ è un *autovettore* di T se $Tv = \lambda v$ per qualche $\lambda \in \mathbb{K}$.

Si noti che $v \in V \setminus \{0\}$ è un autovettore di T se e solo se $(T - \lambda \mathrm{Id})v = 0$ per qualche $\lambda \in \mathbb{K}$. In particolare $v \in V \setminus \{0\}$ è un autovettore di T se e solo se l'operatore $\lambda \mathrm{Id} - T$ non è invertibile, ovvero se e solo se $\det(\lambda \mathrm{Id} - T) = 0$. Pertanto λ è un autovalore di T se e solo se esiste un autovettore (non nullo) v di T tale che $Tv = \lambda v$. Si definisce
$$V_\lambda = \{v \in V : Tv = 0\}.$$
È facile verificare che $V_\lambda = \{0\}$ se e solo se λ *non* è un autovalore di T. Se $\lambda \in \mathbb{K}$ è un autovalore di T allora V_λ si chiamo l'*autospazio di T relativo all'autovalore* λ. Il nome "autospazio" suggerisce che effettivamente V_λ sia uno spazio vettoriale. In effetti vale

Teorema 1.5. *Sia $\lambda \in \mathbb{K}$ un autovalore di T. Allora V_λ è un sottospazio vettoriale di V.*

Dimostrazione. Occorre provare che V_λ è chiuso rispetto alla somma e al prodotto per uno scalare. Siano $v, w \in V_\lambda$. Allora
$$T(v + w) = T(v) + T(w) = \lambda v + \lambda w = \lambda(v + w)$$
e pertanto $v + w \in V_\lambda$. In modo simile si dimostra che V_λ è chiuso rispetto al prodotto per uno scalare. □

Definizione 1.6. Sia $\lambda \in \mathbb{K}$ un autovalore di $T \in \mathrm{End}(V)$. La *molteplicità algebrica* $\mathrm{ma}(\lambda)$ è definita come la molteplicità della radice λ nel polinomio caratteristico $p_T(x)$. In altri termini $\mathrm{ma}(\lambda) = \alpha$ se $(x - \lambda)^\alpha$ divide $p_T(x)$ e $(x - \lambda)^{\alpha+1}$ non divide $p_T(x)$.

La *molteplicità geometrica* di λ è definita da $\mathrm{mg}(\lambda) = \mathrm{dim} V_\lambda$.

Lemma 1.7. *Per un autovalore $\lambda \in \mathbb{K}$ di T vale $mg(\lambda) \leq ma(\lambda)$.*

Dimostrazione. Poniamo $\beta = \mathrm{mg}(\lambda)$. Allora esiste una base $\{v_1, \ldots, v_\beta, v_{\beta+1}, \ldots, v_N\}$ di V tale che $V_\lambda = \langle v_1, \ldots, v_\beta \rangle$. In tale base la matrice $A = (a_{ij})$ associata a T è tale che $a_{jj} = \lambda$ per $j = 1, \ldots, \beta$, $a_{ij} = 0$ per $1 \leq i < j \leq \beta$, $1 \leq j < i \leq \beta$. Pertanto un calcolo diretto mostra che $(x - \lambda)^\beta$ divide $p_T(x)$ e dunque $\beta \leq \mathrm{ma}(\lambda)$ come volevasi. □

Proposizione 1.8. *Siano $\alpha, \beta \in \mathbb{K}$ due autovalori di T. Se $\alpha \neq \beta$ allora $V_\alpha \cap V_\beta = \{0\}$. In particolare per ogni $v \in V_\alpha \setminus \{0\}$ e $w \in V_\beta \setminus \{0\}$ si ha che v e w sono linearmente indipendenti.*

Dimostrazione. Se $v \in V_\alpha \cap V_\beta$ allora $\alpha v = T(v) = \beta v$ e dunque $(\alpha - \beta)v = 0$. Dato che $\alpha - \beta \neq 0$ ciò implica che $v = 0$. □

Definizione 1.9. Sia $W \subset V$ un sottospazio vettoriale di V. Si dice che W è T*-invariante* se $Tw \in W$ per ogni $w \in W$.

Un sottospazio $W \subset V$ che sia T-invariante si dice T*-irriducibile* (o semplicemente *irriducibile*) se per ogni sottospazio $W' \subset W$ che sia T-invariante risulta $W' = W$ oppure $W' = \{0\}$.

Se v è un autovettore di T allora $\langle v\rangle$ è T-invariante (e irriducibile). Viceversa se W è un sottospazio unidimensionale di V T-invariante, allora ogni $w \in W \setminus \{0\}$ è un autovettore di T. Più in generale V_λ è T-invariante (ma può essere o meno irriducibile).

Definizione 1.10. Un'applicazione lineare $T \in \mathrm{End}(V)$ si dice *triangolarizzabile* se esiste una base $\{v_1, \ldots, v_N\}$ di V tale che gli spazi $V_j = \langle v_1, \ldots, v_j\rangle$ siano T-invarianti per ogni $j = 1, \ldots, N$.

Osserviamo che se T è triangolarizzabile, e la base $\{v_1, \ldots, v_N\}$ realizza tale triangolarizzazione, allora v_1 è un autovettore per T. Più in generale la matrice associata a T nell base $\{v_1, \ldots, v_N\}$ è una matrice triangolare superiore (ovvero tale che $A = (a_{ij})$ con $a_{ij} = 0$ per $i > j$, $i, j = 1, \ldots, N$).

Teorema 1.11. *Sono equivalenti:*

(1) *T è triangolarizzabile*
(2) *Il polinomio caratteristico $p_T(x)$ si scompone nel prodotto di fattori lineari.*

La dimostrazione è omessa in questa versione.
In particolare si ha la prima differenza fondamentale tra spazi vettoriali reali e complessi:

Corollario 1.12. *Se V è uno spazio vettoriale su $\mathbb{C}$ allora ogni $T \in$ End(V) è triangolarizzabile.*

Dimostrazione. Sia $T \in \mathrm{End}(V)$. Per il teorema 1.11 si ha che T è triangolarizzabile se e solo se

$$P_T(x) = (x - \lambda_1)^{\alpha_1} \cdots (x - \lambda_r)^{\alpha_r},$$

per qualche $r \in \mathbb{N}$, $\lambda_j \in \mathbb{C}$ e $\alpha_j \in \mathbb{N}$ tale che $\alpha_1 + \ldots \alpha_r = N$. Ma per il teorema fondamentale dell'algebra questo è sempre vero. □

Si ha inoltre

Teorema 1.13 (Hamilton-Caley)**.** $p_T(T) = 0$

La dimostrazione è omessa in questa versione.
Come conseguenza diretta si ha

Corollario 1.14. *Il polinomio minimo $\mu_T(x)$ divide il polinomio caratteristico $p_T(x)$. In particolare deg$(\mu_T(x)) \leq N$.*

Vediamo adesso le relazioni tra autovalori e polinomio minimo

Proposizione 1.15. *Se λ è un autovalore di $T \in$ End(V) allora $(x - \lambda)$ divide il polinomio minimo $\mu_T(x)$.*

Dimostrazione. Sia $v \in V \setminus \{0\}$ tale che $Tv = \lambda v$. Si consideri l'ideale di $\mathbb{K}[x]$ definito da

$$I_{T,v} := \{q \in \mathbb{K}[x] : q(T)v = 0\}.$$

Chiaramente $(x - \lambda) \in I_{T,v}$. Ma $\mathbb{K}[x]$ è un dominio di integrità ad ideali principali, ed essendo $(x - \lambda)$ il polinomio monico di grado minimo (le costanti non nulle non stanno in $I_{T,v}$) in $I_{T,v}$

ne segue che $(x-\lambda)$ è un generatore di $I_{T,v}$. Ora, per definizione, $\mu_T(T)v = 0$ e pertanto $\mu_T(x) \in I_{T,v}$ e dunque $(x-\lambda)$ divide $\mu_T(x)$. $\square$

Mettendo assieme il Corollario 1.14 e la Proposizione 1.15 si ottiene il seguente risultato:

Proposizione 1.16. *Sia $T \in End(V)$ e supponiamo che il polinomio caratteristico di T sia dato da*

$$p_T(x) = (x-\lambda_1)^{\alpha_1} \cdots (x-\lambda_r)^{\alpha_r}$$

dove $\lambda_j \in \mathbb{K}$,con $\lambda_j \neq \lambda_i$ se $i \neq j$, $r \leq N$, $1 \leq \alpha_j \in \mathbb{N}$ e $\alpha_1 + \ldots + \alpha_r = N$.

Allora il polinomio minimo $\mu_T(x)$ è dato da

$$\mu_T(x) = (x-\lambda_1)^{\beta_1} \cdots (x-\lambda_r)^{\beta_r}$$

con $1 \leq \beta_j \leq \alpha_j$.

Esempio 1.17. Sia $T : \mathbb{C}^2 \to \mathbb{C}^2$ definito da $T(x,y) = A(x,y)^t$, dove A è definita dalla seguente matrice:

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

Un calcolo diretto ci dà $p_T(\lambda) = (\lambda-1)^2$. Per il teorema 1.13 si ha che μ_T può essere $\lambda - 1$ oppure $(\lambda-1)^2$. Si scarta subito la prima possibilità dato che $T - \mathrm{Id} \neq 0$. Dunque $\mu_T(\lambda) = (\lambda-1)^2$.

2. Diagonalizzazione di endomorfismi

Definizione 2.1. Un operatore $T \in \mathrm{End}(V)$ si dice *diagonalizzabile* se esiste una base di V formata da autovettori di T.

Dunque T è diagonalizzabile se esistono $v_1, \ldots, v_N$ vettori linearmente indipendenti tali che $Tv_j = \lambda_j v_j$ per qualche $\lambda_j \in \mathbb{K}$. Nella base $\{v_1, \ldots, v_N\}$ che diagonalizza T la matrice associata a T è la matrice diagonale A che ha come entrate λ_j, ovvero $A = (a_{ij})$ con $a_{ij} = 0$ per $i \neq j$ e $a_{jj} = \lambda_j$ per $i, j = 1, \ldots, N$.

Teorema 2.2. (1) *Se T possiede N autovalori distinti allora T è diagonalizzabile.*

(2) *T è diagonalizzabile se e solo se esistono r autovalori distinti $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ tali che*

$$\sum_{i=1}^{r} dim V_{\lambda_i} = N.$$

(3) *T è diagonalizzabile se e solo se esistono r autovalori distinti $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ tali che*

$$V = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_r}.$$

(4) *T è diagonalizzabile se e solo se il polinomio caratteristico $p_T(x)$ si spezza nel prodotto di fattori lineari e la molteplicità geometrica di ogni autovalore eguaglia la molteplicità algebrica.*

Dimostrazione. 1. Siano $\lambda_1, \dots, \lambda_N$ gli N autovalori distinti. Allora dim$V_{\lambda_j} \geq 1$. Per $j = 1, \dots, N$, sia $v_j \in V_{\lambda_j}$ un autovettore (non nullo) di T. Occorre e basta provare che $\{v_1, \dots, v_N\}$ sono linearmente indipendenti. Ma questo segue subito dalla Proposizione 1.8. Infatti, supponiamo per assurdo $v_j = \sum_{i \neq j} a_i v_i$. Allora $v_j \in V_{\lambda_j}$ e d'altra parte $v_j \in \oplus_{i \neq j} V_{\lambda_i}$. Ma $V_{\lambda_j} \cap \oplus_{i \neq j} V_{\lambda_i} = \{0\}$ e pertanto $v_j = 0$, contraddizione.

2. Se T è diagonalizzabile è chiaro che la somma delle dimensioni degli autospazi è N. Viceversa per ogni j si sceglie una base di V_{λ_j} indicata da $\{v_1^j, \dots, v_{\beta_j}^j\}$ (dove $\beta_j =$ dimV_{λ_j}). Nuovamente per la Proposizione 1.8 i vettori $\{v_1^1, \dots, v_{\beta_1}^1, \dots, v_{\beta_r}^r\}$ sono linearmente indipendenti, ma dato che sono N sono una base di V e quindi T è diagonalizzabile.

3. Segue dal punto 2. precedente e dalla Proposizione 1.8.

4. Se il polinomio caratteristico si spezza nel prodotto di fattori lineari distinti significa che esistono $r \geq 1$ autovalori distinti $\lambda_1, \dots, \lambda_r$ tali che ma$(\lambda_1) + \dots +$ ma$(\lambda_r) = N$. Se mg$(\lambda_j) =$ ma(λ_j) per ogni j allora (per definizione di molteplicità geometrica) risulta $\sum$ dim$V_{\lambda_j} = N$ e si applica il punto 2. Il viceversa è ovvio. □

Un criterio più fine per la diagonalizzazione è il seguente, la cui dimostrazione richiede però il teorema di decomposizione primaria ed è posticipata alla sezione successiva:

Teorema 2.3. *T è diagonalizzabile se e solo se il polinomio minimo $\mu_T(x)$ si spezza nel prodotto di fattori lineari distinti, vale a dire se esistono $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ tutti distinti tali che*

$$\mu_T(x) = (x - \lambda_1) \cdots (x - \lambda_r).$$

3. Il teorema di decomposizione primaria nel caso di spezzamento lineare del polinomio caratteristico

In questa sezione supponiamo V uno spazio vettoriale di dimensione N su $\mathbb{K} = \mathbb{R}, \mathbb{C}$ e $T \in$ End(V). Se $p_T(x)$ è il polinomio caratteristico di T supponiamo che si spezzi in fattori lineari:

$$p_T(x) = (x - \lambda_1)^{\alpha_1} \cdots (x - \lambda_r)^{\alpha_r}$$

con r, α_j numeri naturali strettamente maggiori di zero tali che $\alpha_1 + \dots + \alpha_r = N$ e $\lambda_j \in \mathbb{C}$ con $\lambda_j \neq \lambda_i$ se $i \neq j$. Per la Proposizione 1.16 il polinomio minimo $\mu_T(x)$ è allora dato da

$$\mu_T(x) = (x - \lambda_1)^{\beta_1} \cdots (x - \lambda_r)^{\beta_r}$$

con $1 \leq \beta_j \leq \alpha_j$.

Osserviamo che se $\mathbb{K} = \mathbb{C}$ allora per il teorema fondamentale dell'algebra il polinomio caratteristico si spezza sempre in fattori lineari.

Definizione 3.1. L'*autospazio generalizzato* E_{λ_j} relativo all'autovalore λ_j di T è definito tramite

$$E_{\lambda_j} = \text{Ker}(\lambda_j \text{Id} - T)^{\alpha_j}.$$

Enunciamo adesso le proprietà basilari di E_{λ_j}:

Teorema 3.2 (Decomposizione Primaria)**.** *Se il polinomio caratteristico di T si spezza in fattori lineari valgono le seguenti:*

(1) E_{λ_j} *è un sottospazio* T*-invariante.*
(2) $V_{\lambda_j} \subseteq E_{\lambda_j}$.
(3) $V = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_r}$.
(4) *Il polinomio caratteristico di* $T|_{E_j}$ *è dato da* $(x-\lambda_j)^{\alpha_j}$.
(5) *dim*$E_{\lambda_j} = \alpha_j$.
(6) *Il polinomio minimo di* $T|_{E_{\lambda_j}}$ *è dato da* $(x-\lambda_j)^{\beta_j}$.

Dimostrazione. 1. Notiamo che

$$(\lambda_j \mathrm{Id} - T) \circ T = T \circ (\lambda_j \mathrm{Id} - T).$$

Pertanto iterando questa relazione si ottiene $(\lambda_j \mathrm{Id} - T)^s \circ T = T \circ (\lambda_j \mathrm{Id} - T)^s$ per ogni $s \geq 1$. Dunque se $v \in E_{\lambda_j}$ allora $(\lambda_j \mathrm{Id} - T)^s v = 0$ per qualche $1 \leq s \leq \alpha_j$. Pertanto per l'osservazione precedente si ha

$$(\lambda_j \mathrm{Id} - T)^s T v = T(\lambda_j \mathrm{Id} - T)^s v = T0 = 0,$$

e allora $Tv \in E_{\lambda_j}$.

2. Dato che $V_{\lambda_j} = \mathrm{Ker}(\lambda_j \mathrm{Id} - T) \subseteq \mathrm{Ker}(\lambda_j \mathrm{Id} - T)^s$ per ogni $s \geq 1$, allora $V_{\lambda_j} \subseteq E_{\lambda_j}$.

3.Proviamo che $E_{\lambda_1} + \ldots + E_{\lambda_r} = V$. Siano $p_j(x) = \prod_{i \neq j} (x - \lambda_i)^{\alpha_i}$ per $j = 1, \ldots, r$. I polinomi $p_j(x)$ per $j = 1, \ldots, r$ non hanno fattori comuni di grado positivo in $\mathbb{C}[x]$, ne segue che esistono $h_1(x), \ldots, h_r(x) \in \mathbb{C}[x]$ tali che

$$p_1(x)h_1(x) + \ldots + p_r(x)h_r(x) = 1.$$

Poniamo $V_j = p_j(T)V$. Dalla relazione precedente segue che $V = V_1 + \ldots + V_r$, dato che ogni $v \in V$ si può scrivere come

$$v = p_1(T)u_1 + \ldots + p_r(T)u_r,$$

con $u_j = h_j(T)v$. Dato che $p_T(T)v = 0$ per il Teorema 1.13, ne segue che $V_j \subset \mathrm{Ker}(\lambda_j \mathrm{Id} - T)^{\alpha_j} = E_{\lambda_j}$ e pertanto $E_{\lambda_1} + \ldots + E_{\lambda_r} = V$. Proviamo adesso che $E_{\lambda_j} \cap (\sum_{i \neq j} E_{\lambda_i}) = \{0\}$. Diamo la dimostrazione per $j = 1$. Sia $0 \neq v \in E_{\lambda_1}$ e supponiamo $v = u_2 + \ldots + u_r$ con $u_j \in E_{\lambda_j}$. Allora $p_1(T)u_k = 0$ per $k = 2, \ldots, r$ e pertanto deve essere $p_1(T)v = 0$. Per quanto visto in precedenza risulta

$$v = p_1(T)v_1 + \ldots + p_r(T)v_r,$$

dove $v_j = h_j(T)v$. Essendo però $p_j(T)v = 0$ per $j \neq 1$ (dato che $v \in E_{\lambda_1}$), ne segue che $p_l(T)v_l = 0$ e dunque $v = p_1(T)h_1(T)v = h_1(T)p_1(T)v = 0$ contro l'ipotesi su v.

4. Per semplificare le notazioni poniamo $T_j = T|_{E_{\lambda_j}}$. Per i punti 2. e 3. risulta che l'unico autovalore di T_j su E_{λ_j} è λ_j. Pertanto il polinomio caratteristico di T_j è dato da $p_{T_j}(x) = (x - \lambda_j)^{s_j}$ per qualche $s_j \geq 1$, e $s_j = \mathrm{dim}E_{\lambda_j}$. Vogliamo provare che $s_j = \alpha_j$. Per farlo scegliamo una base $\{v_1, \ldots, v_N\}$ di V in modo tale che i primi s_1 vettori siano una base di E_{λ_1}, i successivi s_2 vettori siano un base di E_{λ_2} e cosi via (è possibile farlo per il punto 4.). Inoltre utilizzando il Corollario 1.12 su ciascun T_j si può assumere che tale base triangolarizzi T. In tale base la matrice A associata a T è una matrice triangolare superiore tale che $a_{ii} = \lambda_1$ per

$i = 1, \ldots, s_1$, $a_{ii} = \lambda_2$ per $i = s_1 + 1, \ldots s_1 + s_2$, etc., e si ha

$$p_T(x) = \prod_{i=1}^{N} (x - a_{ii}).$$

Pertanto $s_j = \alpha_j$ per ogni j.

5. Segue subito dal punto 5. e dalla definizione di polinomio caratteristico.

6. Sia $\mu_j(x)$ il polinomio minimo di $T_j = T|_{E_{\lambda_j}}$. Per il punto 5. e per il Corollario 1.14 risulta che $\mu_j(x) = (x - \lambda_j)^{s_j}$ per qualche $1 \leq s_j \leq \alpha_j$. Dato che $\mu_T(T)v = 0$ per ogni $v \in V$, ne segue che $\mu_j(x)$ divide $\mu_T(x)$ per $j = 1, \ldots, r$ e dunque $s_j \leq \beta_j$. D'altra parte, poniamo $p(x) = \mu_1(x) \cdots \mu_r(x)$. Poichè ogni $v \in V$ si scrive come somma di $u_1 + \ldots + u_r$ per qualche $u_j \in E_{\lambda_j}$ risulta che $p(T)v = 0$ e dunque $\mu_T(x)$ divide $p(x)$ e pertanto $s_j \geq \beta_j$, da cui la tesi. $\square$

Possiamo adesso dimostrare il Teorema 2.3.

Dimostrazione del Teorema 2.3. Se T è diagonalizzabile la matrice A associata a T in una base che diagonalizza T è una matrice diagonale. Un calcolo diretto mostra allora che il polinomio minimo di T si spezza nel prodotto di fattori lineari distinti. Viceversa, supponiamo che $\mu_T(x)$ sia il prodotto di fattori lineari distinti. Per il teorema di decomposizione primaria possiamo scegliere una base $\{v_1, \ldots, v_N\}$ di V in modo che i primi α_1 vettori formino una base di E_{λ_1}, i successivi α_2 vettori formino una base di E_{λ_2}, etc.. Vogliamo (e basta) provare che ogni v_j è un autovettore di T. Ma il polinomio minimo di $T|_{E_{\lambda_j}}$ è $(x - \lambda_j)$ per il teorema di decomposizione primaria. Dunque per ogni $v \in E_{\lambda_j}$ vale $(T - \lambda_j \mathrm{Id})v = 0$. In particolare ogni v_j è un autovettore, come volevasi. $\square$

4. Decomposizione secondaria

In questa sezione ci preoccuperemo di trovare una decomposizione T-invariante di E_{λ_j} per $j = 1, \ldots, r$, nelle ipotesi del Teorema 3.2. Fissiamo dunque λ_j. Per semplificare le notazioni poniamo $\lambda = \lambda_j$, $\alpha = \alpha_j$ e $\beta = \beta_j$. Per il Teorema 3.2, si ha che il polinomio caratteristico di $T|_{E_\lambda}$ è $p_j(x) = (x - \lambda)^\alpha$ e il suo polinomio minimo è $\mu_j(x) = (x - \lambda)^\beta$.

Definizione 4.1. Un sottospazio $W \subset E_\lambda$ si dice un *sottospazio ciclico* se esiste $w_0 \in W$ tale che ogni altro $w \in W$ è della forma $p(T)w_0$ per qualche $p(x) \in \mathbb{K}[x]$. Il vettore w_0 si dice un *generatore ciclico* di W.

Vale il seguente risultato:

Teorema 4.2 (Teorema di decomposizione secondaria)**.** *Esiste una decomposizione T-invariante di E_λ data da*

$$E_\lambda = E_1^\lambda \oplus \ldots \oplus E_{\mathit{ma}(\lambda)}^\lambda, \tag{4.1}$$

tale che ogni E_j^λ è un sottospazio ciclico. Inoltre se $n_j = \dim E_j^\lambda$, e si ordinano in modo che $1 \leq n_{\mathit{ma}(\lambda)} \leq \ldots \leq n_1$, risulta che

(1) $n_1 = \beta$.

(2) $(x-\lambda)^{n_j}$ *è il polinomio minimo di* $T|_{E_j^\lambda}$.
(3) *Ogni altra decomposizione* T*-invariante di* E_λ *in sottospazi ciclici ha ma*(λ) *componenti e su ogni componente la restrizione di* T *ha polinomio minimo* $(x-\lambda)^{n_j}$ *per* $j = 1, \ldots,$ *ma*(λ).

Prima di dare la dimostrazione del Teorema 4.2, premettiamo alcune considerazioni:

Definizione 4.3. I polinomi $(x-\lambda)^{n_1}, \ldots, (x-\lambda)^{n_{\mathrm{ma}(\lambda)}}$ si dicono i *divisori elementari* di T su E_λ.

Più in generale, facendo variare λ_j si ottiene una collezione $\{(x-\lambda_1)^{n_1^1}, \ldots, (x-\lambda_r)^{n^r_{\mathrm{ma}(\lambda_r)}}\}$ di polinomi che sono detti i *divisori elementari* di T.

Il Teorema 3.2 e il Teorema 4.2 affermano che i divisori elementari di T sono univocamente determinati e determinano univocamente T a meno di coniugio con automorfismi di V. In altri termini

Proposizione 4.4. *Siano* T, T' *due endomorfismi di* V *tali che il polinomio caratteristico* $p_T(x)$ *di* T *e il polinomio caratteristico* $p_{T'}(x)$ *di* T' *si spezzano in fattori lineari.* T *e* T' *hanno gli stessi divisori elementari se e solo se esiste un automorfismo* S *di* V *tale che* $T' = S \circ T \circ S^{-1}$.

Dimostrazione. Se T e T' sono coniugati allora hanno ovviamente gli stessi divisori elementari. Viceversa, applicando dai Teoremi 3.2 e 4.2 si ottengono due decomposizioni di V per T e T'. L'isomorfismo S si realizza allora mandano ogni generatore ciclico della T-decomposizione nel corrispondente generatore ciclico della T'-decomposizione. I dettagli sono lasciati come esercizio. $\square$

Nota 4.5. Il Teorema 4.2 *non* afferma gli spazi E_j^λ sono univocamente determinati, visto che, come sarà chiaro dalla dimostrazione, dipendono dal generatore ciclico (che non è unico).

Procediamo adesso alla dimostrazione del Teorema 4.2. La dimostrazione è costruttiva e fornisce un metodo pratico per trovare la decomposizione secondaria e più in particolare sarà utilizzata per trovare la forma di Jordan di T.

Dimostrazione del Teorema 4.2. Se $\beta = 1$ allora T è diagonalizzabile su E_λ per il Teorema 2.3 e non c'è niente da dimostrare (ogni E_j^λ è generato da un autovettore di T). Supponiamo $\beta > 1$. Dato che $\mu_j(T) = (x-\lambda)^\beta$, esiste $v \neq 0$ tale che $(T - \lambda\mathrm{Id})^{\beta-1}(v) \neq 0$. Osserviamo che $(T-\lambda\mathrm{Id})^{\beta-1}(E_\lambda) \subset V_\lambda$ poichè $(T-\lambda\mathrm{Id})^\beta(E_\lambda) = \{0\}$. Dunque costruiamo la seguente catena di sottospazi:

$$\{0\} \subset \mathrm{Im}(T-\lambda\mathrm{Id})^{\beta-1}(E_\lambda) \cap V_\lambda \subseteq \mathrm{Im}(T-\lambda\mathrm{Id})^{\beta-2}(E_\lambda) \cap V_\lambda \subseteq \ldots \subseteq V_\lambda.$$

I. Posto $k_{\beta-1} = \mathrm{dimIm}(T-\lambda\mathrm{Id})^{\beta-1}(E_\lambda) \cap V_\lambda$, prendiamo una base $u^{\beta-1}_{1,\beta-1}, \ldots, u^{\beta-1}_{k_{\beta-1},\beta-1}$ di $\mathrm{Im}(T-\lambda\mathrm{Id})^{\beta-1}(E_\lambda) \cap V_\lambda$. Poniamo $S = (T-\lambda\mathrm{Id})$. Dato che ogni $u^{\beta-1}_{l,\beta-1}$ è immagine di $(T-\lambda\mathrm{Id})^{\beta-1}v = S^{\beta-1}v$ per qualche $v \in E_\lambda$ (che non sarà unico dato che ogni altra combinazione del tipo $v+u$ con $u \in V_\lambda$ risolve lo stesso sistema) si possono trovare $u^0_{1,\beta-1}, \ldots, u^0_{\beta-1,\beta-1} \in E_\lambda$ tali che per $l = 1, \ldots, k_{\beta-1}$

$$u^{\beta-1}_{l,\beta-1} = S^{\beta-1}(u^0_{l,\beta-1}).$$

Poniamo $u^j_{l,\beta-1} = S^j(u^0_{l,\beta-1})$ per $l = 1,\dots,k_{\beta-1}$ e $j = 1,\dots,\beta-2$. Si pone allora

$$E^\lambda_l = \langle u^0_{l,\beta-1},\dots,u^{\beta-1}_{l,\beta-1}\rangle$$

per $l = 1,\dots,k_{\beta-1}$.

(A_I) Si dimostra che dim$E^\lambda_l = \beta$ per $l = 1,\dots,k_{\beta-1}$ e $E^\lambda_i \cap E^\lambda_j = \{0\}$ per $i \neq j$.

Posticipiamo la dimostrazione di (A_I) e procediamo nella costruzione degli spazi E^λ_j. Poniamo $k_i = \text{dimIm}(T-\lambda\text{Id})^i(E_\lambda)\cap V_\lambda$ per $i = 0,\dots,\beta-2$.

II. Se $k_{\beta-2} > k_{\beta-1}$ allora esistono dei vettori $u^{\beta-2}_{k_{\beta-1}+1,\beta-1},\dots,u^{\beta-2}_{k_{\beta-2},\beta-1}$ tutti non nulli, tali che $u^{\beta-1}_{1,\beta-1},\dots,u^{\beta-1}_{k_{\beta-1},\beta-1},u^{\beta-2}_{k_{\beta-1}+1,\beta-1},\dots,u^{\beta-2}_{k_{\beta-2},\beta-1}$ sono una base di Im$(T-\lambda\text{Id})^{\beta-2}(E_\lambda)\cap V_\lambda$.

Come prima si possono trovare dei vettori $u^0_{k_{\beta-1}+1,\beta-2},\dots,u^0_{k_{\beta-2},\beta-2} \in E_\lambda$ tali che per $l = k_{\beta-1}+1,\dots,k_{\beta-2}$

$$u^{\beta-2}_{l,\beta-2} = S^{\beta-2}(u^0_{l,\beta-2}).$$

Poniamo $u^j_{l,\beta-2} = T^j(u^0_{l,\beta-2})$ per $l = k_{\beta-1}+1,\dots,k_{\beta-2}$ e $j = 1,\dots,\beta-2$. Si pone allora

$$E^\lambda_l = \langle u^0_{l,\beta-2},\dots,u^{\beta-2}_{l,\beta-2}\rangle$$

per $l = k_{\beta-1}+1,\dots,k_{\beta-2}$ e $j = 1,\dots,\beta-2$.

(A_{II}) Si dimostra che dim$E^\lambda_l = \beta-2$ per $l = k_{\beta-1}+1,\dots,k_{\beta-2}$ e $E^\lambda_i \cap E^\lambda_j = \{0\}$ per $i \neq j$.

La dimostrazione di (A_{II}) è posticipata. Se $k_{\beta-1} = k_{\beta-2}$ si salta il punto II e si passa la punto III seguente:

III. Se $k_{\beta-3} > k_{\beta-2}$ si ragiona come nel punto II costruendo $k_{\beta-3}-k_{\beta-2}$ catene di vettori $u^j_{l,\beta-3}$ per $j = 0,\dots,\beta-3$ e $l = k_{\beta-2}+1,\dots,k_{\beta_3}$. Si definiscono poi gli spazi

$$E^\lambda_l = \langle u^0_{l,\beta-3},\dots,u^{\beta-3}_{l,\beta-3}\rangle$$

per $l = k_{\beta-2}+1,\dots,k_{\beta-3}$ e $j = 1,\dots,\beta-3$. Si prova poi

(A_{III}) Si dimostra che dim$E^\lambda_l = \beta-3$ per $l = k_{\beta-2}+1,\dots,k_{\beta-3}$ e $E^\lambda_i \cap E^\lambda_j = \{0\}$ per $i \neq j$.

Se $k_{\beta-3} = k_{\beta-2}$ si salta il punto III. Si passa poi al punto IV ripetendo il punto II se $k_{\beta-4} > k_{\beta-3}$. Si prosegue in questo modo poi fino ad arrivare a confrontare $k_{\beta-(\beta-1)}$ e $k_0 = \text{ma}(\lambda)$.

Si nota che ogni dim$(E^\lambda_j \cap V_\lambda) = 1$ e pertanto gli E^λ_j sono ma(λ). Inoltre se $v \in E_\lambda$ risulta $(T-\lambda\text{Id})^{\beta-l}(v) = 0$ per qualche $l = 1,\dots,\beta-1$. Allora v appartiene allo spazio generato da $u^{\beta-l}_{\beta-1,1},\dots,u^{\beta-l}_{\beta-1,k_{\beta-1}},u^0_{\beta-l,k_{\beta-l-1}},\dots,u^0_{\beta-l,k_{\beta-l}}$ che è contenuto dunque nella somma degli E^λ_j.

Proviamo adesso (A_I), ..., (A_x). Per prima cosa, posto $u = u^0_{j,\beta-l}$ si vuole provare che $u, Su,\dots,S^{\beta-l}u$ sono linearmente indipendenti (qui $l = 1,\dots,\beta$ e $j = k_{\beta-l-1},\dots,k_{\beta-l}$). Infatti se $a_0u + \dots a_{\beta-l}S^{\beta-l}u = 0$ allora applicando $S^{\beta-l}$ si ottiene $a_0S^{\beta-l}u = 0$ ed essendo $S^{\beta-l}u \neq 0$ per costruzione, risulta $a_0 = 0$. Si applica poi $S^{\beta-l+1}$ e si ottiene $a_1 = 0$ e cosi via fino a $a_{\beta-l} = 0$. Questo dimostra che le dimensioni dei vari E^λ_j sono quelle volute. Proviamo poi che $E_i \cap E_j = \{0\}$ per $i \neq j$. Supponiamo che $E_i = \langle u, Su,\dots,S^{\beta-a-1}u\rangle$ e $E_j = \langle v, Sv,\dots,S^{\beta-b-1}v\rangle$ per certi $0 \leq a,b \leq \beta$, con $S^{\beta-a-1}u$ e $S^{\beta-b-1}v$ linearmente indipendenti. Possiamo anche suppore $a \leq b$. Si vede subito che $u,\dots,S^{b-a-1}u$ non possono appartenere a E_j poiché $S^{\beta-b}(S^j u) = 0$ solo se $j \geq b-a$. Poniamo $w = S^{\beta-a}u$ e $E_i' = \langle w, \ldots, S^{\beta-b-1}w\rangle$. Basta allora dimostrare che $S^j w \notin E_j$ per $j = 0, \ldots, \beta-b-1$. Sia $S^j w = a_0 v + \ldots + a_{\beta-b-1}S^{\beta-b-1}v$. Applicando $S^{\beta-b-1}$ a tale identità si ottiene $S^{j+\beta-b-1}w = a_0 S^{\beta-b-1}v$, da cui $a_0 = 0$ se $j > 0$ poichè $S^{j+\beta-b-1}w = 0$ e se $j = 0$ nuovamente $a_0 = 0$ poichè $S^{\beta-b-1}w = S^{\beta-a-1}u$ e $S^{\beta-b-1}v$ sono linearmente indipendenti. Si procede cosi applicando successivamente $S^{\beta-b-2}, \ldots, S, \mathrm{Id}$ per ottenere $a_1 = \ldots = a_{\beta-b-1} = 0$.

Per terminare la dimostrazione del teorema occorre verificare (1), (2) e (3). Ma (1) e (2) sono immediati dalla definizione di E_j^λ. Per la (3), sia $E_\lambda = W_1 \oplus \ldots \oplus W_m$ una decomposizione T-invariante in sottospazi ciclici. Sia w_1 un generatore ciclico di W_1. Allora esiste $1 \leq s \leq \beta$ tale che $(T-\lambda\mathrm{Id})^s w_1 = 0$ ma $(T-\lambda\mathrm{Id})^{s-1}w_1 \neq 0$. Vogliamo dimostrare che $W_1 = \langle w_1, \ldots, S^{s-1}w_1\rangle$. Sia $w \in W_1$. Allora $w = p(T)w_1$ per qualche $p(x) \in \mathbb{K}[x]$. È facile verificare che allora esiste $q(x) \in \mathbb{K}[x]$ tale che $w = q(S)w_1$. Infatti se $w = a_0 T w_1$ allora $w = a_0 S w_1 + \lambda w_1$. Per induzione supponiamo sia vero per tutti i polinomi fino al grado $m-1$. Allora dato che $S^m = (T-\lambda\mathrm{Id})^m = T^m + h(T)$ con deg$h(x) < m$ e dunque $T^m = S^m - h(T)$, se $w = p(T)w_1 = a_m T^m w_1 + \ldots + a_0 w_1$ si ha $w = a_m S^m w_1 + g(T)w_1$ con $g(x)$ di grado minore di m, e per induzione si ha il risultato. Dunque w_1 è un generatore S-ciclico di W_1. Dunque per ogni $w \in W_1$ esiste un polinomio $q(x)$ di grado m tale che $w = q(S)w_1 = a_0 w_1 + \ldots + a_m S^m w_1$. Ma poichè $S^j w_1 = 0$ per $j \geq s$ si ha $w = a_0 w_1 + \ldots + a_{s-1}S^{s-1}w_1$. Ragionando come in precedenza quando abbiamo trovato la dimensione degli E_j^λ, si vede che $w_1, Sw_1, \ldots, S^{s-1}w_1$ sono linearmente indipendenti. Si noti che il polinomio minimo di T ristretto a W_1 è $(x-\lambda)^s$ e dunque la dimensione di W_1 coincide con il grado del polinomio minimo di T ristretto a W_1. Si ripete il ragionamento per tutti i W_j. Si nota che per ogni W_j risulta $\dim(W_j \cap V_\lambda) = 1$ e quindi risulta $m = \mathrm{ma}(\lambda)$. Ordiniamo in modo decrescente i W_j in base alla loro dimensione. Sia η_k il numero di spazi W_j di dimensione k. Occorre e basta provare che gli η_k sono univocamente determinati da S. Dalla base di W_j si vede subito che

$$\dim(W_j \cap S^l(E_\lambda)) = 0 \text{ se } l \geq \dim W_j,$$

mentre

$$\dim(W_j \cap S^l(E_\lambda)) = \dim W_j - l \text{ se } l < \dim W_j.$$

Pertanto ogni spazio di dimensione $\beta - k$ (con $k = 0, \ldots, \beta$) contribuisce con $(\beta-k)-(\beta-j) = j-k$ vettori linearmente indipendenti ad una base di $\mathrm{Im}(S^{\beta-j})$ (dove $0 < j < k$). Pertanto si ha, per $0 \leq j < \beta$

$$\dim \mathrm{Im} S^{\beta-j} = j\eta_\beta + (j-1)\eta_{\beta-1} + \ldots + \eta_{j+1}.$$

Da qui si vede subito che η_β è determinato solo da S, dunque anche $\eta_{\beta-1}$ è determinato solo da S e cosi via tutti gli altri. □

5. LA FORMA DI JORDAN

Definizione 5.1. Una matrice A quadrata $m \times m$ si dice un *blocco di Jordan di dimensione* m relativo a $\lambda \in \mathbb{C}$ se $a_{ii} = \lambda$, $a_{i+1,i} = 1$ per $i = 1, \ldots, m$, e tutte le altre entrate $a_{ij} = 0$ se $i \neq j, j+1$. Una matrice $N \times N$ J si dice *matrice di Jordan* se è una matrice a blocchi, in cui ogni blocco è di Jordan.

Definizione 5.2. Sia $T \in \mathrm{End}(V)$. Si dice che T è riducibile in *forma di Jordan* se esiste una base (detta *base di Jordan* per T) di V in cui la matrice associata a T sia una matrice di Jordan.

Teorema 5.3 (Esistenza della forma di Jordan). *Sia $T \in End(V)$. L'endomorfismo T è riducibile in forma di Jordan se e solo se il suo polinomio caratteristico $p_T(x)$ si spezza nel prodotto di fattori lineari.*

Dimostrazione. Se T è riducibile in forma di Jordan basta calcolare $p_T(x)$ per la matrice associata a T nella base di Jordan per T e verificare che è prodotto di fattori lineari. Viceversa si applica il Teorema 3.2 e il Teorema 4.2. Dato che ogni $E_k^{\lambda_j}$ è ciclico, per ogni j, k si sceglie un generatore ciclico v_{jk} tale che, $(T - \lambda_j \mathrm{Id})^{n_{jk}} v_{jk} = 0$ (dove $n_{jk} = \mathrm{dim} E_k^{\lambda_j}$). Si pone allora $e_l^{jk} = (T - \lambda_j \mathrm{Id})^{n_{jk}-l} v_{jk}$, per $l = 1, \ldots, n_{jk}$. Per la scelta fatta risulta che $\{e_l^{jk}\}$, ordinata in modo che e_{l+1}^{jk} segua e_l^{jk} per ogni j, k, è una base di V, e si verifica subito che la matrice associata a T in tale base è una matrice di Jordan. □

Teorema 5.4 (Unicità della forma di Jordan). *La forma di Jordan di $T \in End(V)$ quando esiste è unica a meno di coniugio. In altri termini, se A è una matrice di Jordan di T in una data base di V e B è una matrice di Jordan di T in un altra base, allora esiste una matrice C invertibile tale che $B = C^{-1}AC$.*

Dimostrazione. Notiamo che se A è una matrice di Jordan associata a T in una qualche base, allora lo spazio V ha una decomposizione in sottospazi T-invarianti ciclici (su ognuno di tali sottospazi T è dato da un blocco di Jordan di A). Si vede subito che ognuno di tali sottospazi è contenuto in un (unico) autospazio generalizzato di T e dunque l'unicità segue dal Teorema 4.2. □

Per il teorema precedente è possibile parlare di "numero dei blocchi di Jordan di T di una certa dimensione relativi ad un autovalore". In effetti, come si vede dalla dimostrazione dei due precedenti teoremi, T ha un blocco di dimensione k relativo all'autovalore λ_j se e solo se nella decomposizione spettrale di V relativa a T esiste $E_s^{\lambda_j}$ di dimensione k. Poniamo

$$\eta_k^j = \text{numero di blocchi di Jordan di dimensione } k \text{ relativi a } \lambda_j.$$

Teorema 5.5 (Proprietà della forma di Jordan). *Sia $T \in End(V)$ tale che*

$$p_T(x) = (x - \lambda_1)^{\alpha_1} \cdots (x - \lambda_r)^{\alpha_r}$$

e

$$\mu_T(x) = (x - \lambda_1)^{\beta_1} \cdots (x - \lambda_r)^{\beta_r}.$$

Allora

(1) $\alpha_j = \eta_1^j + 2\eta_2^j + \ldots + \beta_j \eta_{\beta_j}^j$.
(2) $\eta_{\beta_j}^j \geq 1$.
(3) $\eta_1^j + \ldots + \eta_{\beta_j}^j = mg(\lambda_j) = dim V_{\lambda_j}$.
(4) $dim[Im(T - \lambda_j Id)^{\beta_j - k} \cap E_{\lambda_j}] = \sum_{l=0}^{k-1} (k - l)\eta_{\beta_j - l}^j$ *per* $k = 1, \ldots, \beta_j - 1$.
(5) $\eta_{\beta_j - k}^j = dim[Im(T - \lambda_j Id)^{\beta_j - k+1} \cap V_{\lambda_j}] - dim[Im(T - \lambda_j Id)^{\beta_j - k} \cap V_{\lambda_j}]$.

(6) $\eta_k^j = 2\textit{dimKer}(T-\lambda_j \textit{Id})^k - \textit{dimKer}(T-\lambda_j \textit{Id})^{k+1} - \textit{dimKer}(T-\lambda_j \textit{Id})^{k-1}$.

Dimostrazione. I punti dall'1 al 5 si ricavano direttamente dalla dimostrazione del Teorema 3.2 e 4.2 e dalle osservazioni precedenti. Per quanto riguarda il punto 6., si nota che

$$\begin{aligned}\dim[\mathrm{Im}(T-\lambda_j\mathrm{Id})^{\beta_j-k}\cap E_{\lambda_j}] &= \mathrm{rk}(T-\lambda_j\mathrm{Id})^{\beta_j-k}|_{E_{\lambda_j}}\\ &= \alpha_j - \mathrm{dimKer}(T-\lambda_j\mathrm{Id})^{\beta_j-k}|_{E_{\lambda_j}} = \alpha_j - \mathrm{dimKer}(T-\lambda_j\mathrm{Id})^{\beta_j-k},\end{aligned}$$

poichè l'operatore $(T-\lambda_j\mathrm{Id})^{\beta_j-k}$ è invertibile su $\oplus_{i\neq j}E_{\lambda_i}$. Dalla formula 4., tenendo conto che $\alpha_j = \mathrm{dimKer}(T-\lambda_j\mathrm{Id})^{\beta_j}$ si ha la formula.

$\square$