Capitolo 5. La forma canonica di un operatore lineare.[1]

0. Richiami sui polinomi.

Sia
$$p(t) = a_0 + a_1 t + \cdots + a_n t^n$$
un polinomio a coefficienti complessi, di grado $\deg(p(t)) = n > 0$. Quindi $a_n \neq 0$. Il coefficiente a_n si dice *il coefficiente direttore di* $p(t)$. Se $a_n = 1$ si dice che $p(t)$ e' *monico.*

Nell'insieme dei polinomi $\mathbf{C}[\,t\,]$, oltre all'addizione interna ed alla moltiplicazione esterna per uno scalare $a \in \mathbf{C}$, possiamo definire anche una moltiplicazione interna. Dati due polinomi $p(t) = a_0 + a_1 t + \cdots + a_n t^n$ e $q(t) = b_0 + b_1 t + \cdots + b_m t^m$ si definisce il loro prodotto *per distributivita'*, cioe' ponendo:

$$p(t) \cdot q(t) := a_0 b_0 + (a_0 b_1 + a_1 b_0)t + (a_0 b_2 + a_1 b_1 + a_2 b_0)t^2 + \cdots + (a_n b_m)t^{n+m}.$$

In altri termini, se c_j e' il coefficiente di t^j in $p(t) \cdot q(t)$, allora

$$c_j := \sum_{i=0}^{j} a_i b_{j-i}$$

(in tale formula si assume che $a_i = 0$ per $i > n$ e $b_{j-i} = 0$ per $j - i > m$). Quindi, come nel caso delle matrici quadrate, possiamo riguardare l'insieme dei polinomi come un'algebra
$$(\mathbf{C}[\,t\,], +\,, .\,, .).$$
La moltiplicazione interna e' associativa, e' commutativa, il polinomio costante 1 funge da elemento neutro, e un polinomio $p(t)$ e' invertibile (cioe' esiste un polinomio $q(t)$ tale che $p(t) \cdot q(t) \equiv 1$) se e solo se $p(t)$ e' un polinomio costante non nullo. Dalla stessa definizione di prodotto interno segue che $p(t) \cdot q(t)$ e' il polinomio nullo se e solo se $p(t)$ oppure $q(t)$ e' il polinomio nullo. E si ha anche:

$$\deg(p(t) \cdot q(t)) = \deg(p(t)) + \deg(q(t)), \quad \deg(p(t) + q(t)) \leq \max\{\deg(p(t)), \deg(q(t))\}.$$

Dato un numero $c \in \mathbf{C}$ possiamo valutare il polinomio $p(t)$ in c definendo $p(c) := a_0 + a_1 c + \cdots + a_n c^n$. Se $p(c) = 0$ allora diremo che c e' *una radice (o uno zero)* per $p(t)$. E' un problema classico quello di calcolare le radici di un polinomio. Il seguente importante teorema ci dice che conoscere una radice di un polinomio $p(t)$ equivale a conoscerne un fattore di primo grado.

Teorema di Ruffini. Se c e' una radice di $p(t)$ allora esiste un polinomio $q(t)$ tale che $p(t) = q(t)(t - c)$.

La dimostrazione del Teorema di Ruffini e' una immediata conseguenza del cosiddetto

[1] *Ultimo aggiornamento: 25 dicembre 2022*

Algoritmo della divisione. *Siano $f(t)$ e $g(t)$ polinomi, con $g(t)$ polinomio non nullo. Allora esistono e sono unici polinomi $q(t)$ ed $r(t)$ tali che*

$$f(t) = q(t)g(t) + r(t),$$

con $r(t)$ polinomio nullo oppure non nullo ma con il grado strettamente minore del grado di $g(t)$.

I polinomi $q(t)$ ed $r(t)$ sono detti *il quoziente ed il resto della divisione tra $f(t)$ e $g(t)$*. Il motivo di tale definizione risiede nel fatto che l'Algoritmo della divisione si dimostra proprio effettuando la divisione tra $f(t)$ e $g(t)$.

Ritornando al Teorema di Ruffini, possiamo allora dividere $p(t)$ per $t - c$ e scrivere $p(t) = q(t)(t - c) + r(t)$, con $q(t)$ ed $r(t)$ quoziente e resto della divisione. Poiche' il grado di $t - c$ e' 1, allora il resto $r(t)$ e' un polinomio costante. Se c e' una radice di $p(t)$, sostituendo in $p(t) = q(t)(t - c) + r(t)$ otteniamo

$$0 = p(c) = q(c)(c - c) + r(c).$$

Vediamo quindi che deve essere $r(c) = 0$, per cui $r(t)$ e' il polinomio nullo e si ha $p(t) = q(t)(t - c)$. Cio' prova il Teorema di Ruffini. Si osservi che se $p(t) = q(t)(t - c)$ allora lo studio dell'equazione $p(t) = 0$ equivale allo studio dell'equazione $q(t) = 0$, che e' piu' semplice in quanto $\deg q(t) = \deg p(t) - 1$.

Teorema Fondamentale dell'Algebra. *Sia $p(t) = a_0 + a_1 t + \cdots + a_n t^n$ un polinomio a coefficienti complessi, di grado $\deg(p(t)) = n > 0$. Allora esiste una radice $c \in \mathbf{C}$ per $p(t)$.*

Applicando il Teorema Fondamentale dell'Algebra ed il Teorema di Ruffini deduciamo il seguente

Corollario. *Sia $p(t) = a_0 + a_1 t + \cdots + a_n t^n$ un polinomio a coefficienti complessi, di grado $\deg(p(t)) = n > 0$. Allora esistono numeri $c_1, \ldots, c_n \in \mathbf{C}$ tali che*

$$p(t) = a_n(t - c_1) \cdot \ldots \cdot (t - c_n).$$

Cioe' ogni polinomio si puo' decomporre nel prodotto di polinomi complessi di primo grado (si puo' anche provare che tale decomposizione e' unica, a parte l'ordine con cui appaiono i fattori).

Nella decomposizione

$$p(t) = a_n(t - c_1) \cdot \ldots \cdot (t - c_n).$$

potrebbero apparire fattori ripetuti. Raggruppando i fattori uguali possiamo scrivere

$$p(t) = a_n(t - c_1)^{k_1} \cdot \ldots \cdot (t - c_l)^{k_l}$$

dove con $c_1, \dots, c_l$ denotiamo tutte le radici distinte di $p(t)$, e con k_j la molteplicita' con cui appare la radice c_j nella fattorizzazione di $p(t)$. Il numero k_j si chiama *la molteplicita' algebrica della radice* c_j. Denoteremo con $Spec(p(t))$ lo *spettro di* $p(t)$, cioe' l'insieme delle sue radici $c_1, \dots, c_l$.

Il Teorema Fondamentale dell'Algebra ha una interessante conseguenza per cio' che riguarda la fattorizzazione di polinomi reali tramite polinomi reali.

A tale proposito consideriamo un polinomio reale, cioe' con coefficienti a_i numeri reali. Sia c una radice di $p(t)$. Poiche' il coniugato di una somma e' uguale alla somma dei coniugati, ed il coniugato di un prodotto e' uguale al prodotto dei coniugati, abbiamo:

$$0 = p(c) = \overline{p(c)} = \overline{a_0 + a_1 c + \dots + a_n c^n} = \overline{a_0} + \overline{a_1} \cdot \overline{c} + \dots + \overline{a_n} \cdot \overline{c}^n.$$

Poiche' i coefficienti sono reali allora $\overline{a_j} = a_j$ e percio'

$$0 = p(c) = \overline{a_0} + \overline{a_1} \cdot \overline{c} + \dots + \overline{a_n} \cdot \overline{c}^n = a_0 + a_1 \cdot \overline{c} + \dots + a_n \cdot \overline{c}^n = p(\overline{c}).$$

In altre parole, *se* $p(t)$ *e' un polinomio a coefficienti reali e se* c *e' una radice di* $p(t)$ *allora anche* $\overline{c}$ *e' una radice di* $p(t)$. *Quindi le radici di* $p(t)$ *si ripartiscono nell'insieme delle radici reali, diciamo* $c_1, \dots, c_a$*, e nell'insieme delle radici complesse non reali, diciamo* $d_1, \overline{d_1}, \dots, d_b, \overline{d_b}$*: queste ultime appaiono in coppie, ognuna insieme alla sua coniugata (e si puo' provare che appaiono con la stessa molteplicita').*

Pertanto possiamo riscrivere la fattorizzazione di $p(t)$ nel seguente modo:

$$p(t) = a_n (t - c_1)^{k_1} \cdot \ldots \cdot (t - c_a)^{k_a} \cdot (t - d_1)^{h_1} \cdot (t - \overline{d_1})^{h_1} \cdot \ldots \cdot (t - d_b)^{h_b} \cdot (t - \overline{d_b})^{h_b}.$$

Osserviamo adesso che

$$(t - d_j) \cdot (t - \overline{d_j}) = t^2 - (d_j + \overline{d_j}) t + d_j \overline{d_j}$$

e' un polinomio a coefficienti reali (con discriminante negativo) in quanto $d_j + \overline{d_j}$ e $d_j \overline{d_j}$ sono sempre numeri reali.

In conclusione possiamo riassumere quanto detto nella seguente

Proposizione. *Sia* $p(t)$ *un polinomio non nullo a coefficienti reali. Allora* $p(t)$ *ammette una fattorizzazione del tipo*

$$p(t) = a q_1(t) \cdot q_2(t) \cdot \ldots \cdot q_m(t)$$

in cui a *e' un numero reale, ed i polinomi* $q_j(t)$ *sono monici, con coefficienti reali, di grado* 1*, oppure di grado* 2 *con discriminante negativo.*

Si puo' provare che la decomposizione e' unica, a meno dell'ordine con cui appaiono i fattori.

Esempio. Consideriamo il polinomio a coefficienti reali $p(t) = t^3 - 1$. Se vogliamo decomporre $p(t)$ in polinomi reali piu' semplici, possiamo scrivere

$$t^3 - 1 = (t - 1)(t^2 + t + 1).$$

Con polinomi reali non possiamo fare di piu'. Se vogliamo una fattorizzazione in polinomi di primo grado dobbiamo ricorrere ai numeri complessi:

$$t^3 - 1 = (t-1)(t - \frac{-1 + i\sqrt{3}}{2})(t - \frac{-1 - i\sqrt{3}}{2}). \quad \blacksquare$$

Concludiamo il paragrafo con una osservazione circa gli operatori lineari. Nel caso degli spazi vettoriali su $\mathbf{R}$ sappiamo che vale il seguente:

Teorema. *Sia $f : V \to V$ un operatore lineare definito su uno spazio vettoriale reale V. Allora f e' diagonalizzabile se e solo se valgono le seguenti due proprieta'.*

(1) Il polinomio caratteristico di f si decompone completamente nel prodotto di polinomi di primo grado (reali).

(2) Per ogni autovalore λ di f si ha $m_g(\lambda) = m_a(\lambda)$.

Lo stesso Teorema vale anche per spazi vettoriali complessi (cioe' per spazi dove gli scalari sono numeri complessi). La dimostrazione e' identica. In questo caso pero', grazie al Teorema Fondamentale dell'Algebra, sappiamo che ogni polinomio si puo' fattorizzare in polinomi di primo grado (con coefficienti complessi). Quindi nell'enunciato precedente la condizione (1) e' sovrabbondante, e possiamo riscrivere l'enunciato nella seguente forma piu' semplice:

Teorema. *Sia $f : V \to V$ un operatore lineare definito su uno spazio vettoriale complesso V. Allora f e' diagonalizzabile se e solo se per ogni autovalore λ di f si ha $m_g(\lambda) = m_a(\lambda)$.*

1. Matrici diagonali a blocchi.

Con delle linee orizzontali e verticali tratteggiate possiamo ripartire una matrice A in tante matrici piu' piccole, chiamate *blocchi* di A. La matrice A viene allora chiamata *matrice a blocchi*. Una data matrice puo' essere ripartita in blocchi in vari modi.

Esempio 1.

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 5 & 3 & 2 \\ 1 & 7 & 2 & 3 & 1 \end{bmatrix} = \left[\begin{array}{cc|cc|c} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 5 & 3 & 2 \\ \hline 1 & 7 & 2 & 3 & 1 \end{array}\right] = \left[\begin{array}{ccc|cc} 1 & 2 & 3 & 4 & 5 \\ \hline 4 & 1 & 5 & 3 & 2 \\ \hline 1 & 7 & 2 & 3 & 1 \end{array}\right].$$

In questo esempio la matrice A e' una matrice 3×5; nella prima decomposizione A appare come una matrice a blocchi 2×3, cioe' appare nella forma

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

dove $A_{11} = \begin{bmatrix} 1 & 2 \\ 4 & 1 \end{bmatrix}$, $A_{12} = \begin{bmatrix} 3 & 4 \\ 5 & 3 \end{bmatrix}$, $A_{13} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$, $A_{21} = [1 \quad 7]$, $A_{22} = [2 \quad 3]$, $A_{23} = [1]$. *Invece nella seconda decomposizione A appare come una matrice a blocchi 3×2.* $\blacksquare$

La convenienza della partizione in blocchi consiste nel fatto che (almeno in certi casi) il risultato di operazioni su matrici a blocchi si puo' ottenere portando avanti il calcolo sui blocchi stessi, proprio come se essi fossero gli elementi effettivi della matrice. Cio' conduce ad una semplificazione dei calcoli. Questo avviene soprattutto in presenza di matrici diagonali a blocchi. Una matrice A si dice *matrice diagonale a blocchi* se soddisfa le seguenti condizioni: A e' una matrice quadrata ripartita in blocchi, i blocchi formano a loro volta una matrice quadrata, i blocchi disposti lungo la diagonale principale sono quadrati, e tutti i blocchi fuori della diagonale sono nulli.

Esempio 2. *La seguente matrice*

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 & 6 \end{bmatrix}$$

e' diagonale a blocchi, con i blocchi lungo la diagonale principale dati da $A_{11} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, $A_{22} = [\,3\,]$ *ed* $A_{33} = \begin{bmatrix} 1 & 5 \\ 1 & 6 \end{bmatrix}$. *Si puo' rappresentare* A *anche sotto la forma*

$$A = \begin{bmatrix} A_{11} & & \\ & A_{22} & \\ & & A_{33} \end{bmatrix},$$

in cui si indicano solo i blocchi disposti lungo la diagonale principale, e gli spazi bianchi stanno ad indicare i blocchi nulli. ■

Piu' in generale una matrice diagonale a blocchi si puo' indicare sotto la forma:

$$\begin{bmatrix} A_{11} & & & & \\ & A_{22} & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & & A_{kk} \end{bmatrix}.$$

E' evidente che la somma di matrici diagonali a blocchi (con blocchi corrispondenti della stessa grandezza) e' ancora una matrice diagonale a blocchi. Cio' vale anche per il prodotto per uno scalare. Inoltre se $A_{11}, \ldots, A_{kk}$ denotano i blocchi sulla diagonale di una matrice A diagonale a blocchi, allora il determinante di A e' uguale al prodotto dei determinanti dei singoli blocchi, cioe' $\det(A) = \det(A_{11}) \cdot \ldots \cdot \det(A_{kk})$. Infine A e' invertibile se e solo se i singoli blocchi sono invertibili, e l'inversa di A e' quella matrice diagonale a blocchi con i blocchi dati dalle matrici inverse dei blocchi di A.

Esempio 3. *Con riferimento all'esempio precedente, abbiamo:*

$$A^2 = \begin{bmatrix} A_{11}^2 & & \\ & A_{22}^2 & \\ & & A_{33}^2 \end{bmatrix} = \begin{bmatrix} 5 & 3 & 0 & 0 & 0 \\ 3 & 2 & 0 & 0 & 0 \\ 0 & 0 & 9 & 0 & 0 \\ 0 & 0 & 0 & 6 & 35 \\ 0 & 0 & 0 & 7 & 41 \end{bmatrix},$$

ed anche

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} & & \\ & A_{22}^{-1} & \\ & & A_{33}^{-1} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 6 & -5 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}.$$

Infine

$$\det(A) = \det(A_{11}) \cdot \det(A_{22}) \cdot \det(A_{33}) = 1 \cdot 3 \cdot 1 = 3. \quad \blacksquare$$

2. Matrici diagonali a blocchi e sottospazi invarianti di un operatore lineare.

Ora vogliamo mostrare che in certe situazioni la matrice rappresentativa di un operatore si puo' presentare come una matrice diagonale a blocchi. Il concetto che consente di tradurre la nozione di matrice a blocchi nel contesto degli operatori lineari e' quello di *sottospazio invariante.* Sia $f : V \to V$ un operatore lineare di uno spazio vettoriale V. Un sottospazio U di V si dice *sottospazio invariante per* f se i vettori di U sono trasformati tramite f ancora in vettori di U, cioe' se $f(U) \subseteq U$. In tal caso allora l'operatore f induce per restrizione un operatore di U, cioe' l'operatore $\mathbf{u} \in U \to f(\mathbf{u}) \in U$. Si osservi anche che sia V che $\{\mathbf{0}\}$ sono sempre invarianti. Altri esempi di sottospazi invarianti sono gli autospazi. La seguente osservazione e' utile per riconoscere se un sottospazio e' oppure no invariante.

Proposizione 1. *Supponiamo che* U *sia generato dai vettori* $\mathbf{u}_1, \dots, \mathbf{u}_h$. *Allora* U *e' invariante per* f *se e solo se i vettori* $f(\mathbf{u}_1), \dots, f(\mathbf{u}_h)$ *appartengono ancora ad* U.

Quanto segue mostra il legame con le matrici diagonali a blocchi. Supponiamo che V sia somma diretta di due sottospazi invarianti per f, U e W:

$$V = U \oplus W.$$

Denotiamo con $\mathcal{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_h\}$ una base per U e con $\mathcal{W} = \{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ una base per W. Sappiamo che $\mathcal{B} := \mathcal{U} \cup \mathcal{W}$ e' una base per V. Sia $\varphi : U \to U$ l'operatore indotto da f su U, e sia $\psi : W \to W$ quello indotto su W. Ebbene vale la seguente

Proposizione 2. *La matrice rappresentativa $M_{\mathcal{B}}^{\mathcal{B}}(f)$ di f rispetto alla base $\mathcal{B}$ e' diagonale a blocchi, con blocchi dati dalla matrice $M_{\mathcal{U}}^{\mathcal{U}}(\varphi)$ di φ rispetto ad $\mathcal{U}$ e dalla matrice $M_{\mathcal{W}}^{\mathcal{W}}(\psi)$ di ψ rispetto a $\mathcal{W}$, cioe':*

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{bmatrix} M_{\mathcal{U}}^{\mathcal{U}}(\varphi) & \\ & M_{\mathcal{W}}^{\mathcal{W}}(\psi) \end{bmatrix}.$$

Vale anche il viceversa, e cioe': sia $f : V \to V$ un operatore lineare per il quale esiste una base $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ tale che la matrice rappresentativa di f rispetto a $\mathcal{B}$ sia diagonale a blocchi del tipo

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{bmatrix} A_{11} & \\ & A_{22} \end{bmatrix},$$

con A_{11} di ordine h e A_{22} di ordine $k = n - h$. Allora posto $\mathcal{U} := \{\mathbf{b}_1, \dots, \mathbf{b}_h\}$, $U := Span(\mathcal{U})$, $\mathcal{W} := \{\mathbf{b}_{h+1}, \dots, \mathbf{b}_n\}$ e $W := Span(\mathcal{W})$, si ha che U e W sono sottospazi di V invarianti per f e $V = U \oplus W$. Inoltre, denotata con $\varphi : U \to U$ la restrizione di f su U e con $\psi : W \to W$ la restrizione di f su W, si ha che

$$M_{\mathcal{U}}^{\mathcal{U}}(\varphi) = A_{11} \quad e \quad M_{\mathcal{W}}^{\mathcal{W}}(\varphi) = A_{22}.$$

Quindi la conoscenza di una decomposizione di V in sottospazi invarianti permette di rappresentare il dato operatore con una matrice diagonale a blocchi. Cio' consente di semplificare lo studio dell'operatore. Ad esempio il polinomio caratteristico di f si decompone nel prodotto del polinomio di φ e del polinomio di ψ, quindi per calcolare gli autovalori di f ci si riconduce a calcolare le radici di polinomi aventi grado piu' piccolo di quello di f. Una ovvia generalizzazione si ha quando V ammette una decomposizione in piu' di due sottospazi invarianti.

Esempio 4. *Sia $f : \mathbf{C}^4 \to \mathbf{C}^4$ l'operatore lineare definito ponendo $f(x,y,z,t) := (x+y+z, t, -y, x+z)$. Consideriamo i seguenti sottospazi $U := Span(2\mathbf{e}_1+\mathbf{e}_2-\mathbf{e}_3+\mathbf{e}_4)$, $W := Span(\mathbf{e}_1-\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_4)$. Poniamo $\mathcal{U} := \{2\mathbf{e}_1+\mathbf{e}_2-\mathbf{e}_3+\mathbf{e}_4\}$ e $\mathcal{W} := \{\mathbf{e}_1-\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_4\}$. E' chiaro che $\mathcal{U}$ e' una base per U, $\mathcal{W}$ e' una base per W, e che $\mathcal{B} := \mathcal{U} \cup \mathcal{W}$ e' una base per $\mathbf{C}^4$. Dunque $\mathbf{C}^4 = U \oplus W$. Proviamo che U e W sono invarianti per f. Infatti $f(2\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3 + \mathbf{e}_4) = 2\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3 + \mathbf{e}_4 \in U$, e cio' per la Proposizione 1 comporta che U e' invariante. Analogamente abbiamo: $f(\mathbf{e}_1 - \mathbf{e}_3) = \mathbf{0} \in W$, $f(\mathbf{e}_2) = \mathbf{e}_1 - \mathbf{e}_3 \in W$ e $f(\mathbf{e}_4) = \mathbf{e}_2 \in W$, quindi anche W e' invariante. Con le stesse notazioni della Proposizione 2, andiamo a studiare gli operatori indotti. Studiare l'operatore indotto su U significa calcolare la matrice rappresentativa di φ rispetto alla base $\mathcal{U}$. Nel nostro caso $M_{\mathcal{U}}^{\mathcal{U}}(\varphi) = [\,1\,]$ (in particolare $\varphi = id_U$). Invece $M_{\mathcal{W}}^{\mathcal{W}}(\psi) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ (quindi se*

(y', z', t') sono le coordinate di W rispetto alla base $\mathcal{W}$ si ha $\psi(y', z', t') = (z', t', 0)$). In base alla Proposizione 2 possiamo anche dire che

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Quindi se (x', y', z', t') sono le coordinate di $\mathbf{C}^4$ rispetto alla base $\mathcal{B}$ si ha $f(x', y', z', t') = (x', z', t', 0)$, ed inoltre $p_f(t) = p_\varphi(t) \cdot p_\psi(t) = (t-1)t^3$. ■

3. Blocchi di Jordan e stringhe di un operatore lineare.

Tra i possibili blocchi di una matrice, di grande interesse sono i *blocchi di Jordan.* Fissiamo un numero complesso $\lambda \in \mathbf{C}$ ed un intero $p \in \mathbf{N}$. Si definisce *blocco di Jordan di ordine p relativo all'autovalore λ* la matrice quadrata $J_{\lambda,p}$ di ordine p che ha sulla diagonale principale tutte le entrate uguali a λ, sopra la diagonale principale tutte le entrate uguali ad 1, e tutte le rimanenti entrate nulle. Scriviamo esplicitamente tali blocchi per $1 \leq p \leq 4$.

$$J_{\lambda,1} = [\,\lambda\,], \quad J_{\lambda,2} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \quad J_{\lambda,3} = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}, \quad J_{\lambda,4} = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}.$$

Si osservi che $J_{\lambda,p} = \lambda I_p + J_{0,p}$. Quindi, come diremo meglio in seguito, lo studio di un blocco di Jordan si puo' sempre ricondurre a quello di un blocco relativo all'autovalore 0. Il polinomio caratteristico di $J_{\lambda,p}$ e' $p_{J_{\lambda,p}}(t) = (-1)^p(t-\lambda)^p$. Dunque $J_{\lambda,p}$ ha soltanto l'autovalore λ. Inoltre, quando $p \geq 2$ $J_{\lambda,p}$ non e' diagonalizzabile. Infatti $m_a(\lambda) = p$ mentre $m_g(\lambda) = p - rk(J_{\lambda,p} - \lambda I_p) = p - (p-1) = 1$.

La nozione di blocco di Jordan corrisponde, nell'ambito degli operatori lineari, a quella di *stringa.* Se $f : V \to V$ e' un operatore lineare ed $\mathcal{S} = \{\mathbf{b}_1, \dots, \mathbf{b}_p\}$ e' un sistema di p vettori di V, allora $\mathcal{S}$ si dice che e' una *λ-stringa di lunghezza p per f* se $\mathbf{b}_1 \neq \mathbf{0}$ ed inoltre:

$$\begin{cases} f(\mathbf{b}_p) = \lambda \mathbf{b}_p + \mathbf{b}_{p-1} \\ f(\mathbf{b}_{p-1}) = \lambda \mathbf{b}_{p-1} + \mathbf{b}_{p-2} \\ \dots \\ f(\mathbf{b}_1) = \lambda \mathbf{b}_1. \end{cases}$$

Possiamo rappresentare una stringa nel seguente modo:

$$\mathbf{b}_p \xrightarrow{f-\lambda id_V} \mathbf{b}_{p-1} \xrightarrow{f-\lambda id_V} \mathbf{b}_{p-2} \xrightarrow{f-\lambda id_V} \dots \xrightarrow{f-\lambda id_V} \mathbf{b}_1 \xrightarrow{f-\lambda id_V} \mathbf{0}.$$

Il grafico precedente mostra che una λ-stringa di lunghezza p per f e' anche una 0-stringa di lunghezza p per l'operatore $f - \lambda id_V$, e viceversa. Una stringa di lunghezza 1 altro non e' che un autovettore. Osserviamo anche che la condizione $\mathbf{b}_1 \neq \mathbf{0}$ comporta che i vettori di una stringa siano linearmente indipendenti. Possiamo provare cio' per induzione sulla lunghezza p della stringa. Se $p = 1$ l'asserto e' vero proprio perche' $\mathbf{b}_1 \neq 0$. Supponiamo allora $p \geq 2$. E' lecito assumere che la stringa sia relativa all'autovalore nullo. Sia $a_1\mathbf{b}_1 + a_2\mathbf{b}_2 + a_3\mathbf{b}_3 + \cdots + a_p\mathbf{b}_p = \mathbf{0}$ una relazione tra i vettori della stringa. Applicando l'operatore f otteniamo $a_2\mathbf{b}_1 + a_3\mathbf{b}_2 + \cdots + a_p\mathbf{b}_{p-1} = \mathbf{0}$. Poiche' i vettori $\mathbf{b}_1, \ldots, \mathbf{b}_{p-1}$ formano una stringa di ordine $p-1$, per induzione deduciamo che $a_2 = a_3 = \cdots = a_p = 0$. Sara' anche $a_1 = 0$, perche' allora $a_1\mathbf{b}_1 = \mathbf{0}$. E cosi' tutti i pesi devono essere nulli.

Il nesso tra la nozione di blocco di Jordan e quella di stringa e' fornita dalla seguente

Proposizione 3. *Sia $f : V \to V$ un operatore lineare ed $\mathcal{S}$ un sistema di vettori di V. Allora $\mathcal{S}$ e' una λ-stringa di lunghezza p per f se e solo se $\mathcal{S}$ e' linearmente indipendente, $Span(\mathcal{S})$ e' invariante per f, e la matrice rappresentativa $M^{\mathcal{S}}_{\mathcal{S}}(f|Span(\mathcal{S}) \to Span(\mathcal{S}))$ rispetto alla base $\mathcal{S}$ dell'operatore f ristretto su $Span(\mathcal{S})$ e' il blocco di Jordan $J_{\lambda,p}$.*

Esempio 5. *(i) Nell'Esempio 4 sia $\mathcal{U}$ che $\mathcal{W}$ sono stringhe per f. La prima e' una stringa di lunghezza 1 relativa all'autovalore 1, mentre la seconda e' una stringa di lunghezza 3 relativa all'autovalore 0;*

(ii) La base canonica di $\mathbf{C}^4$ e' una stringa per l'operatore $f(x, y, z, t) = (\lambda x + y, \lambda y + z, \lambda z + t, \lambda t)$.

(iii) In base alla Proposizione 3, se la matrice rappresentativa rispetto alla base canonica e' un blocco di Jordan, questo significa proprio che la base canonica e' una stringa. $\blacksquare$

4. Il Teorema di Jordan sulla forma canonica.

Siamo in condizione di enunciare il il seguente Teorema, che e' il risultato principale di questo capitolo.

Teorema (di Jordan). *Sia $f : V \to V$ un operatore lineare definito su uno spazio vettoriale complesso V di dimensione finita. Allora esiste una base $\mathcal{S}$ di V, formata da stringhe, e detta base a stringhe, tale che la matrice rappresentativa $J := J_f := M^{\mathcal{S}}_{\mathcal{S}}(f)$ e' una matrice diagonale a blocchi, con blocchi di Jordan. Tale matrice si chiama la forma canonica di Jordan di f. Essa e' unica a meno dell'ordine con cui figurano i blocchi. Gli autovalori di f coincidono con gli autovalori dei blocchi di J. Per ciascun autovalore λ di f, $m_g(\lambda)$ e' uguale al numero dei blocchi relativi a λ che appaiono in J, mentre $m_a(\lambda)$ e' uguale alla somma delle grandezze dei blocchi relativi a λ che appaiono in J. In particolare f e' diagonalizzabile se e solo se tutti i blocchi che appaiono in J hanno grandezza 1.*

Dal Teorema di Jordan deduciamo dunque che ogni operatore lineare si puo' rappresentare con una matrice diagonale a blocchi di Jordan. I blocchi di Jordan sono percio' i "mattoni" con cui si possono costruire tutte le matrici, a meno di similitudine.

Per una data matrice quadrata A di ordine n, denoteremo con J_A la forma canonica dell'operatore $\mathbf{x} \in \mathbf{C}^n \to A \cdot \mathbf{x} \in \mathbf{C}^n$. La matrice J_A si chiama anche *la forma canonica di* A. Per la stessa definizione, A e J_A sono simili. Inoltre, se A e' simile a B, allora per transitivita' J_A e' simile a J_B. Per l'unicita' della forma canonica segue che $J_A = J_B$ (a meno dell'ordine con cui figurano i blocchi). Viceversa, se $J_A = J_B$, sempre per transitivita' e' chiaro che A e' simile a B. Ne risulta pertanto il seguente criterio per stabilire completamente se due matrici sono simili oppure no:

Corollario. *Due matrici sono simili se e solo se hanno la stessa forma canonica di Jordan.*

Esempio 6. *(i) Supponiamo che f sia un operatore la cui matrice rappresentativa rispetto ad una certa base $\mathcal{S}$ sia fatta cosi':*

$$\begin{bmatrix}
0 & 1 & & & & & & & & & & & & & & \\
0 & 0 & & & & & & & & & & & & & & \\
 & & 0 & 1 & 0 & & & & & & & & & & & \\
 & & 0 & 0 & 1 & & & & & & & & & & & \\
 & & 0 & 0 & 0 & & & & & & & & & & & \\
 & & & & & 0 & 1 & 0 & 0 & & & & & & & \\
 & & & & & 0 & 0 & 1 & 0 & & & & & & & \\
 & & & & & 0 & 0 & 0 & 1 & & & & & & & \\
 & & & & & 0 & 0 & 0 & 0 & & & & & & & \\
 & & & & & & & & & \sqrt{5} & 1 & 0 & & & & \\
 & & & & & & & & & 0 & \sqrt{5} & 1 & & & & \\
 & & & & & & & & & 0 & 0 & \sqrt{5} & & & & \\
 & & & & & & & & & & & & \sqrt{5} & 1 & & \\
 & & & & & & & & & & & & 0 & \sqrt{5} & & \\
 & & & & & & & & & & & & & & \sqrt{5} & \\
 & & & & & & & & & & & & & & & -4
\end{bmatrix}.$$

Allora possiamo dire che lo spazio su cui agisce f ha dimensione 16, *e che la base $\mathcal{S}$ e' una base a stringhe. Gli autovalori di f sono $0, \sqrt{5}, -4$, e le rispettive molteplicita' sono: $m_a(0) = 9$, $m_g(0) = 3$, $m_a(\sqrt{5}) = 6$, $m_g(\sqrt{5}) = 3$, $m_a(-4) = 1$, $m_g(-4) = 1$. Infine il polinomio caratteristico di f e' $p_f(t) = t^9(t - \sqrt{5})^6(t + 4)$.*

(ii) Nell'esempio 4 la base $\mathcal{S}$ e' una base a stringhe. ■

Ora passiamo alla dimostrazione del Teorema di Jordan. Potremo utilizzare tale dimostrazione come una sorta di algoritmo per calcolare esplicitamente una base a stringhe. Dal punto di vista delle matrici si tratta di risolvere il seguente problema: assegnata una matrice quadrata A trovare una matrice invertibile P tale che la matrice $J = P^{-1}AP$ sia una matrice a blocchi di Jordan. Le colonne della matrice P formeranno una base a stringhe.

L'algoritmo consta di tre passi.

Passo 1*: La separazione degli autovalori.*

Consideriamo il nostro dato iniziale, cioe' l'operatore lineare $f : V \to V$. Siano $\lambda_1, \dots, \lambda_k$ gli autovalori distinti di f, e sia $p_f(t)$ il polinomio caratteristico di f, che, poiche' stiamo usando i numeri complessi, si potra' scrivere sotto la forma:

$$p_f(t) = (t-\lambda_1)^{m_a(\lambda_1)} \cdot (t-\lambda_2)^{m_a(\lambda_2)} \cdot \dots \cdot (t-\lambda_k)^{m_a(\lambda_k)}.$$

Per ciascun autovalore λ, accanto all'autospazio V_λ si puo' definire *l'autospazio generalizzato* $\widetilde{V}_\lambda$:

$$\widetilde{V}_\lambda := \ker(f - \lambda id_V)^{m_a(\lambda)}.$$

Se A e' una matrice rappresentativa di f rispetto a certe coordinate $\mathbf{x}$, allora il calcolo esplicito dell'autospazio generalizzato consiste nel risolvere il sistema lineare

$$\widetilde{V}_\lambda \cong \{\mathbf{x} : (A - \lambda I)^{m_a(\lambda)} \cdot \mathbf{x} = \mathbf{0}\}.$$

Si osservi che se $m_a(\lambda) = 1$ allora $\widetilde{V}_\lambda = V_\lambda$. Circa l'autospazio generalizzato sussistono le seguenti proprieta'.

Lemma 1. *Le proprieta' dell'autospazio generalizzato.*

1) $V_\lambda \subseteq \widetilde{V}_\lambda$, cioe' l'autospazio e' sempre contenuto nell'autospazio generalizzato corrispondente;

2) ciascun autospazio generalizzato e' invariante per f;

3) $\dim \widetilde{V}_\lambda = m_a(\lambda)$;

4) l'operatore f ristretto sull'autospazio generalizzato $\widetilde{V}_\lambda$ ha come polinomio caratteristico il polinomio $(t-\lambda)^{m_a(\lambda)}$;

5) $V = \widetilde{V}_{\lambda_1} \oplus \dots \oplus \widetilde{V}_{\lambda_k}$, cioe' lo spazio ambiente V e' la somma diretta dei suoi autospazi generalizzati (cioe' ogni vettore $\mathbf{u}$ di V si puo' scrivere, in modo unico, come una somma del tipo $\mathbf{u}_1 + \mathbf{u}_2 + \dots + \mathbf{u}_k$, con $\mathbf{u}_i \in \widetilde{V}_{\lambda_i}$).

Dimostrazione del Lemma 1. 1). Fissiamo il primo autovalore λ_1, e chiamiamolo λ. Per ogni intero $h \geq 0$, poniamo:

$$N_h = \ker(f - \lambda id_V)^h, \quad T_h = \Im(f - \lambda id_V)^h.$$

Osserviamo che

$$(f - \lambda id_V)^h(\mathbf{u}) = \mathbf{0} \implies (f - \lambda id_V)(f - \lambda id_V)^h(\mathbf{u}) = \mathbf{0} \implies (f - \lambda id_V)^{h+1}(\mathbf{u}) = \mathbf{0}.$$

Percio' abbiamo la sequenza crescente:

$$\{\mathbf{0}\} = N_0 \subseteq N_1 \subseteq N_2 \subseteq \dots$$

da cui, in particolare, deduciamo la proprieta' 1), tenuto conto che

$$V_\lambda = N_1 \quad \text{e che} \quad \widetilde{V}_\lambda = N_{m_a(\lambda)}.$$

Similmente, si prova che i sottospazi T_h formano una sequenza decrescente:

$$V = T_0 \supseteq T_1 \supseteq T_2 \supseteq \dots .$$

2). Sia $\mathbf{u} \in N_h$. Allora

$$\mathbf{0} = f(\mathbf{0}) = f(f - \lambda id_V)^h(\mathbf{u}) = (f - \lambda id_V)^h f(\mathbf{u}).$$

Quindi $f(\mathbf{u}) \in N_h$. Cio' prova che ogni N_h e' f-invariante. In particolare, anche $\widetilde{V}_\lambda$ $(=N_{m_a(\lambda)})$ e' invariante per f. Similmente, si prova che tutti i sottospazi T_h sono invarianti per f.

3) e 4). Sia p il primo intero per cui $N_p = N_{p+1}$. Naturalmente, per il Teorema della dimensione, p coincide anche con il primo intero per cui $T_p = T_{p+1}$. Andiamo a provare che per ogni $h \geq p$ si ha:

$$N_p = N_h \quad e \quad T_p = T_h$$

(per tale motivo, l'intero p si chiama *l'indice di stabilizzazione dell'autovalore* λ). Sia infatti $\mathbf{u} \in N_h$, con $h \geq p+1$. Allora

$$\mathbf{0} = (f - \lambda id_V)^h(\mathbf{u}) = (f - \lambda id_V)^{p+1}(f - \lambda id_V)^{h-p-1}(\mathbf{u}).$$

Percio' $(f - \lambda id_V)^{h-p-1}(\mathbf{u}) \in N_{p+1} = N_p$, quindi

$$\mathbf{0} = (f - \lambda id_V)^p(f - \lambda id_V)^{h-p-1}(\mathbf{u}) = (f - \lambda id_V)^{h-1}(\mathbf{u}).$$

Sicche' $\mathbf{u} \in N_{h-1}$. Cosi' continuando, si riesce a provare che $\mathbf{u} \in N_p$. Percio' le sequenze precedenti possono essere rappresentate anche cosi':

$$\{\mathbf{0}\} = N_0 \subset N_1 \subset N_2 \subset \dots \subset N_p = N_{p+1} = N_{p+2} = \dots$$

e

$$V = T_0 \supset T_1 \supset T_2 \supset \dots \supset T_p = T_{p+1} = T_{p+2} = \dots .$$

In particolare, $p \leq \dim N_p$. Il fatto che tali sequenze ad un certo punto si stabilizzano (simultaneamente), ci consente di provare che

$$V = N_p \oplus T_p.$$

A tale scopo, sia $\mathbf{u} \in N_p \cap T_p$. Andiamo a provare che $\mathbf{u}$ *deve* essere il vettore nullo. Infatti, da una parte sappiamo che $(f - \lambda id_V)^p(\mathbf{u}) = \mathbf{0}$, e dall'altra che esiste un vettore $\mathbf{v}$ tale che $\mathbf{u} = (f - \lambda id_V)^p(\mathbf{v})$. Ma allora $(f - \lambda id_V)^{2p}(\mathbf{v}) = \mathbf{0}$. Da cui $\mathbf{v} \in N_{2p} = N_p$.

Percio' $(f - \lambda id_V)^p(\mathbf{v}) = \mathbf{0}$. Cioe' $\mathbf{u} = \mathbf{0}$. Avendo provato che $N_p \cap T_p = \{\mathbf{0}\}$, ne consegue che $V = N_p \oplus T_p$.

Quindi V e' la somma diretta di due sottospazi invarianti, cioe' N_p e T_p. Sia allora $\varphi : N_p \to N_p$ la restrizione di $f : V \to V$ ad N_p, e sia $\psi : T_p \to T_p$ la restrizione di f a T_p. Sappiamo che

$$p_f(t) = p_\varphi(t) \cdot p_\psi(t).$$

Ora andiamo a provare che λ *(cioe'* λ_1*) e' l'unico autovalore di* φ*, e che* λ *non e' un autovalore per* ψ. Infatti, innanzitutto sia μ un autovalore per φ. Allora esiste un vettore non nullo $\mathbf{u} \in N_p$ tale che $f(\mathbf{u}) = \mu\mathbf{u}$. Da cui $(\varphi - \lambda id_{N_p})(\mathbf{u}) = (\mu - \lambda)\mathbf{u}$. Ma allora

$$\mathbf{0} = (f - \lambda id_V)^p(\mathbf{u}) = (\varphi - \lambda id_{N_p})^p(\mathbf{u}) = (\mu - \lambda)^p\mathbf{u}.$$

Ne consegue $(\mu - \lambda)^p = 0$, cioe' $\mu = \lambda$. Cio' prova che λ e' l'unico autovalore di φ. Ora supponiamo per assurdo che λ sia un autovalore anche per ψ, cioe' che esista un vettore non nullo $\mathbf{u} \in T_p$ tale che $\psi(\mathbf{u}) = \lambda\mathbf{u}$. Quindi $f(\mathbf{u}) = \lambda\mathbf{u}$, percio' $\mathbf{u} \in N_1 \subseteq N_p$. Cio' contraddice il fatto che $V = N_p \oplus T_p$. Dalla proprieta' appena provata, segue che

$$p_\varphi(t) = (t - \lambda_1)^{m_a(\lambda_1)},$$

e che

$$p_\psi(t) = (t - \lambda_2)^{m_a(\lambda_2)} \cdot (t - \lambda_3)^{m_a(\lambda_3)} \cdot \ldots \cdot (t - \lambda_k)^{m_a(\lambda_k)}.$$

Deduciamo che la dimensione di N_p e' $m_a(\lambda)$, e poiche' $p \leq \dim N_p$, ne consegue che $N_p = N_{m_a(\lambda)}$, cioe' $N_p = \widetilde{V}_\lambda$. Con cio' abbiamo provato 3) e 4).[2]

5). Infine, dalla decomposizione appena provata, sappiamo che

$$V = N_p \oplus T_p = \widetilde{V}_{\lambda_1} \oplus T_p.$$

Adesso, tenuto conto che tutti gli autovalori diversi da λ_1 sono concentrati in ψ, possiamo riapplicare il ragionamento precedente sull'operatore ψ, con l'autovalore λ_2. Cosi' continuando si ottiene la decomposizione voluta.

Con cio' si conclude la dimostrazione del Lemma 1.

In base a quanto abbiamo detto a proposito delle matrici rappresentative diagonali a blocchi, dalle proprieta' precedenti segue che *per dimostrare il Teorema di Jordan e' sufficiente saperlo dimostrare per ciascuna restrizione* $f|\widetilde{V}_{\lambda_j} \to \widetilde{V}_{\lambda_j}$*. Cioe' e' sufficiente saperlo dimostrare quando l'operatore* f *ha un unico autovalore.* Questo e' quanto andremo ad imparare nel prossimo passo.

Passo 2: Riduzione all'autovalore nullo.

[2] Poiche' l'indice di stabilizzazione di λ e' $m_a(\lambda)$, allora $N_{m_a(\lambda)}$, che coincide con $\widetilde{V}_\lambda$, contiene tutti gli N_h. Quindi i vettori di una stringa per f relativa all'autovalore λ appartengono a $\widetilde{V}_\lambda$. Per tale motivo, i vettori di una stringa sono detti *autovettori generalizzati.*

Supponiamo allora che l'operatore $f : V \to V$ abbia polinomio caratteristico del tipo $p_f(t) = (t-\lambda)^n$, con $\dim V = n$. Ora effettuando la "traslazione" $g := f - \lambda id_V$ si ottiene un operatore g che ha polinomio caratteristico $p_g(t) = t^n$. Chiaramente una base a stringhe per g e' anche una base a stringhe per f. Pertanto *per dimostrare il Teorema di Jordan e' sufficiente saperlo dimostrare per un operatore che abbia soltanto l'autovalore* 0. Un siffatto operatore dicesi *operatore nilpotente*. A proposito degli operatori nilpotenti, si puo' provare che:

Lemma 2. *Per un operatore $f : V \to V$ definito su uno spazio complesso di dimensione n sono equivalenti le seguenti proprieta':*

(i) f e' nilpotente;

(ii) $p_f(t) = t^n$;

(iii) f^n e' l'operatore nullo;

(iv) esiste un intero $h > 0$ tale che f^h e' l'operatore nullo.

Dimostrazione del Lemma 2.

(i) $\implies$ (ii). Se f e' nilpotente, ha soltanto l'autovalore nullo. D'altra parte, poiche' i coefficienti sono complessi, il polinomio caratteristico si decompone completamente. Percio' $p_f(t) = t^n$.

(ii) $\implies$ (iii). Se $p_f(t) = t^n$, allora $V = \widetilde{V}_0$. Percio' f^n e' l'operatore nullo.

(iii) $\implies$ (iv). Ovvia.

(iv) $\implies$ (i). Sia λ un autovalore per f, e sia $\mathbf{u}$ un vettore non nullo tale che $f(\mathbf{u}) = \lambda\mathbf{u}$. Allora $\mathbf{0} = f^h(\mathbf{u}) = \lambda^h\mathbf{u}$. Per la legge di annullamento del prodotto, segue che $\lambda = 0$. Quindi f ha solo l'autovalore nullo, cioe' f e' nilpotente.

Cio' conclude la dimostrazione del Lemma 2.

Per concludere la dimostrazione del Teorema di Jordan, dobbiamo solo capire come si trova una base a stringhe per un operatore nilpotente.

Passo 3*: La forma canonica di Jordan per un operatore nilpotente.*

Sia $f : V \to V$ un operatore nilpotente. Innanzitutto occorre individuare *l'indice di nilpotenza di* f, che e' definito come il minimo intero positivo p tale che $f^p \equiv \mathbf{0}$. Tale indice si calcola andando a vedere qual e' la minima potenza che annulla una matrice rappresentativa di f (e coincide con l'indice di stabilizzazione dell'autovalore $\lambda = 0$).

Esempio 7. *(i) Gli unici blocchi di Jordan nilpotenti sono quelli con l'autovalore nullo, cioe' i blocchi della forma $J_{0,p}$. Inoltre l'indice di nilpotenza di un blocco di Jordan del tipo $J_{0,p}$ e' proprio p. Visualizziamo nel caso $p = 4$.*

$$J_{0,4} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad J_{0,4}^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$J_{0,4}^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad J_{0,4}^4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Questo comportamento vale in generale. Cioe' la diagonale di "1" di $J_{0,p}$ sale di una diagonale nella matrice $J_{0,p}^2$, e cio' accade ogni volta che si esegue una nuova potenza, finche' non si arriva a $J_{0,p}^p$, che e' la prima potenza in cui si annulla il blocco. Ne consegue che l'indice di nilpotenza di una matrice diagonale a blocchi di Jordan nilpotenti e' l'ordine massimo dei blocchi che appaiono nella matrice.

(ii) Per esempio la matrice

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

ha indice di nilpotenza $p = 4$. ■

Una volta conosciuto l'indice di nilpotenza abbiamo la seguente informazione.

Proposizione 4. *Sia $f : V \to V$ un operatore nilpotente con indice p. Allora*

1) esiste almeno una 0-stringa per f di lunghezza p;

2) ogni 0-stringa di f ha lunghezza minore o uguale a p.

Cio' equivale a dire che nella forma canonica di un operatore nilpotente con indice p ci deve essere almeno un blocco di grandezza p, e tutti gli altri avranno grandezza minore o uguale a p.

Dimostrazione della Proposizione 4.

1). Poiche' l'indice e' p, allora $Imf^{p-1} \neq \{\mathbf{0}\}$, e $Imf^p = \{\mathbf{0}\}$. Sia $\mathbf{u}$ un vettore in Imf^{p-1} non nullo. Sia $\mathbf{w} \in V$ un vettore tale che $\mathbf{u} = f^{p-1}(\mathbf{w})$. Abbiamo allora la seguente sequenza (si tenga presente che $f(\mathbf{u}) = \mathbf{0}$ perche' $f(\mathbf{u}) = f^p(\mathbf{w})$):

$$\mathbf{w} \xrightarrow{\quad f \quad} f(\mathbf{w}) \xrightarrow{\quad f \quad} \ldots \xrightarrow{\quad f \quad} f^{p-2}(\mathbf{w}) \xrightarrow{\quad f \quad} f^{p-1}(\mathbf{w}) = \mathbf{u} \xrightarrow{\quad f \quad} \mathbf{0},$$

Cio' significa, per la stessa definizione di stringa, che la sequenza:

$$\mathbf{u} = f^{p-1}(\mathbf{w}), \quad f^{p-2}(\mathbf{w}), \quad \ldots \quad , \quad f(\mathbf{w}), \quad \mathbf{w}$$

e' una stringa per f, di lunghezza p, relativa all'autovalore nullo.

2). Sia $\mathbf{b}_1, \ldots, \mathbf{b}_h$ una stringa per f, relativa all'autovalore nullo, di lunghezza h. Allora $\mathbf{b}_1 = f^{h-1}(\mathbf{b}_h)$. Poiche' $\mathbf{b}_1 \neq \mathbf{0}$, allora f^{h-1} non e' l'operatore nullo. Percio' $h - 1 < p$, cioe' $h \leq p$.

Cio' conclude la dimostrazione della Proposizione 4.

Siamo ora in grado di calcolare una base a stringhe per un operatore nilpotente con indice p.

Innanzitutto ci andiamo a calcolare le stringhe di lunghezza massima p. Si procede nel seguente modo. Poiche' p e' l'indice di nilpotenza, allora l'operatore f^{p-1} non e' nullo, cioe' $Im(f^{p-1})$ ha dimensione positiva. Sia $\mathbf{u}_1, \ldots, \mathbf{u}_l$ una base per $Im(f^{p-1})$ (nella pratica, se f e' un operatore di $\mathbf{C}^n$ ed A e' la matrice di f rispetto alla base canonica, allora $\mathbf{u}_1, \ldots, \mathbf{u}_l$ e' una base per lo spazio delle colonne di A^{p-1}, che e' una matrice non nulla). Sia $\mathbf{u}$ uno dei vettori di tale base. Per definizione esiste un vettore $\mathbf{w}$ tale che $f^{p-1}(\mathbf{w}) = \mathbf{u}$ (con riferimento alla matrice A, se $\mathbf{u}$ e' la colonna di posto j di A^{p-1} allora possiamo prendere $\mathbf{w} = \mathbf{e}_j$). Consideriamo i seguenti vettori: $\mathbf{w}$, $f(\mathbf{w})$, $f^2(\mathbf{w})$, ... , $f^{p-2}(\mathbf{w})$, $f^{p-1}(\mathbf{w}) = \mathbf{u}$ (cioe' $\mathbf{e}_j$, la colonna di posto j di A, la colonna di posto j di A^2, ..., la colonna di posto j di A^{p-2}, la colonna di posto j di A^{p-1}). Essi, presi nell'ordine opposto, formano una 0-stringa di lunghezza p in quanto $\mathbf{u} \neq \mathbf{0}$, e $f(\mathbf{u}) = f^p(\mathbf{w}) = \mathbf{0}$ perche' p e' l'indice di nilpotenza. Graficamente abbiamo:

$$\mathbf{w} \xrightarrow{\ f\ } f(\mathbf{w}) \xrightarrow{\ f\ } \ldots \xrightarrow{\ f\ } f^{p-2}(\mathbf{w}) \xrightarrow{\ f\ } f^{p-1}(\mathbf{w}) = \mathbf{u} \xrightarrow{\ f\ } \mathbf{0},$$

o anche, in termini di matrici,

$$\mathbf{e}_j \xrightarrow{\ A\ } A\mathbf{e}_j \xrightarrow{\ A\ } \ldots \xrightarrow{\ A\ } A^{p-2}\mathbf{e}_j \xrightarrow{\ A\ } A^{p-1}\mathbf{e}_j = \mathbf{u} \xrightarrow{\ A\ } \mathbf{0}.$$

Facendo variare $\mathbf{u}$ nella base di $Im(f^{p-1})$ costruiamo tante stringhe di lunghezza p quant'e' la dimensione di $Im(f^{p-1})$.

In modo analogo si possono costruire delle stringhe di lunghezza $p - 1$. Cioe' si fissa una base di $Im(f^{p-2})$, ed in corrispondenza di un vettore $\mathbf{u}$ di tale base possiamo considerare la sequenza $\mathbf{w}$, $f(\mathbf{w})$, $f^2(\mathbf{w})$, ... , $f^{p-3}(\mathbf{w})$, $f^{p-2}(\mathbf{w}) = \mathbf{u}$, dove appunto $\mathbf{w}$ e' un vettore tale che $f^{p-2}(\mathbf{w}) = \mathbf{u}$. Questa volta pero' non e' piu' scontato che tale sequenza sia una stringa in quanto non e' detto che $f(\mathbf{u}) = f^{p-1}(\mathbf{w})$ sia $\mathbf{0}$. Per ovviare a questa difficolta' occorre fissare non una base di $Im(f^{p-2})$, ma una base di $Im(f^{p-2}) \cap \ker f$. In realta' anche questo accorgimento non e' sufficiente in quanto le stringhe cosi' costruite potrebbero essere parte delle stringhe gia' costruite prima. Poiche' stiamo cercando di costruire una base non possiamo accettare vettori ripetuti. Per ovviare anche a questa difficolta' si puo' considerare un complementare $\mathcal{K}^{(p-1)}$ di $Im(f^{p-1})$ in $Im(f^{p-2}) \cap \ker f$:

$$Im(f^{p-2}) \cap \ker f = Im(f^{p-1}) \oplus \mathcal{K}^{(p-1)}.$$

Applicando ad una fissata base di $\mathcal{K}^{(p-1)}$ lo stesso argomento con cui abbiamo costruito stringhe di lunghezza p si costruiscono delle stringhe di lunghezza $p - 1$ indipendenti dalle stringhe di lunghezza p calcolate prima (nella pratica una base di $\mathcal{K}^{(p-1)}$ si ottiene estendendo a base di $Im(f^{p-2}) \cap \ker f$ una base del sottospazio $Im(f^{p-1})$).

Similmente, per costruire stringhe di lunghezza $p-2$ indipendenti dalle precedenti, si parte da una base di un fissato complementare $\mathcal{K}^{(p-2)}$ di $Im(f^{p-2}) \cap \ker f$ in $Im(f^{p-3}) \cap \ker f$.

Cosi' continuando si costruisce un certo numero di stringhe di lunghezza $p-3$, poi di lunghezza $p-4$, fino ad arrivare a stringhe di lunghezza 1.

Possiamo riepilogare e precisare quanto appena detto nel seguente modo.

Supponiamo che $f : V \to V$ sia un operatore nilpotente, con indice p. Consideriamo i seguenti sottospazi $\mathcal{K}^{(p)}, \dots, \mathcal{K}^{(1)}$ di V, definiti dalle seguenti uguaglianze (si tenga presente che $Imf \supset Imf^2 \supset Imf^3 \supset \cdots \supset Imf^{p-1} \supset \{\mathbf{0}\} = Imf^p = Imf^{p+1} = \dots$):

$$Imf^{p-1} = \mathcal{K}^{(p)}$$

$$Imf^{p-2} \cap Kerf = Imf^{p-1} \oplus \mathcal{K}^{(p-1)}$$

$$Imf^{p-3} \cap Kerf = (Imf^{p-2} \cap Kerf) \oplus \mathcal{K}^{(p-2)}$$

$$\dots\dots$$

$$Imf \cap Kerf = (Imf^2 \cap Kerf) \oplus \mathcal{K}^{(2)}$$

$$Imf^0 \cap Kerf = (Imf \cap Kerf) \oplus \mathcal{K}^{(1)}.$$

In una forma piu' breve, possiamo scrivere:

$$Imf^{i-1} \cap Kerf = (Imf^i \cap Kerf) \oplus \mathcal{K}^{(i)} \quad \text{per ogni } 1 \leq i \leq p.$$

Per costruire lo spazio $\mathcal{K}^{(i)}$ si puo'procedere cosi'. Si calcola una base di $Imf^i \cap Kerf$. Poiche' $Imf^i \cap Kerf$ e' un sottospazio di $Imf^{i-1} \cap Kerf$, possiamo estendere tale base a base di tutto lo spazio $Imf^{i-1} \cap Kerf$. Allora possiamo prendere come $\mathcal{K}^{(i)}$ lo spazio generato dai vettori che abbiamo dovuto aggiungere durante l'estensione a base.

Poniamo $k_i := \dim \mathcal{K}^{(i)}$. Come abbiamo visto prima, per la stessa definizione dei sottospazi $\mathcal{K}^{(i)}$, possiamo costruire tante stringhe di lunghezza i quant'e' la dimensione di ciascun $\mathcal{K}^{(i)}$, cioe' k_i stringhe di lunghezza i. Tali stringhe, riunite insieme, formano un sistema $\mathcal{S}$ di vettori linearmente indipendenti. Infatti si ha il seguente lemma:

Lemma 3 *Sia $f : V \to V$ un operatore lineare, e siano $\mathcal{S}_1, \dots, \mathcal{S}_k$ stringhe per f, con autovettori iniziali $\mathbf{s}_1, \dots, \mathbf{s}_k$. Se $\mathbf{s}_1, \dots, \mathbf{s}_k$ sono linearmente indipendenti, allora lo e' anche il sistema $\mathcal{S}_1 \cup \cdots \cup \mathcal{S}_k$.*

Dimostrazione del Lemma 3. I vettori di una stringa relativa all'autovalore λ appartengono all'autospazio generalizzato $\widetilde{V}_\lambda$. Poiche' V e' la somma diretta degli autospazi generalizzati, allora e' lecito supporre che le stringhe in questione siano relative tutte allo stesso autovalore. E possiamo supporre anche che tale autovalore sia quello nullo. Sia ora ℓ la lunghezza massima tra le lunghezze delle stringhe assegnate. Se $\ell = 1$ l'asserto e' vero per ipotesi. Possiamo allora supporre $\ell \geq 2$, e ragionare per induzione su ℓ. La dimostrazione adesso procede allo stesso modo con cui si prova che i vettori di una stringa sono linearmente indipendenti.

Fine della dimostrazione del Lemma 3.

La dimostrazione del Teorema di Jordan si conclude osservando che il numero di vettori che stanno in $\mathcal{S}$ e' $n\,(=\dim V)$, cioe' osservando che:

$$\sum_{i=1}^{p} ik_i = n.$$

Per provare cio', abbiamo bisogno della seguente osservazione, che dimostreremo subito dopo:

Lemma 4. *Per ogni* $1 \leq i \leq p$ *si ha*

$$k_i = \dim Imf^{i-1} - 2\dim Imf^{i} + \dim Imf^{i+1}.$$

Sulla base di questa osservazione, allora abbiamo:

$$\sum_{i=1}^{p} ik_i = \sum_{i=1}^{p} i\left(\dim Imf^{i-1} - 2\dim Imf^{i} + \dim Imf^{i+1}\right).$$

Ora, per ogni $1 \leq i \leq p-1$, nella somma precedente, $\dim Imf^{i}$ appare nell'espressione di k_{i+1} con coefficiente $i+1$, nell'espressione di k_i con coefficiente $-2i$, e nell'espressione di k_{i-1} (solo quando $i \geq 2$) con coefficiente $i-1$. Poiche' $(i+1)-2i+(i-1)=0$, allora al termine della somma precedente non apparira' nessun $\dim Imf^{i}$, con $1 \geq 1$. Cioe' nella somma precedente, rimane solo $\dim Imf^{0}$, che e' n. Cio' prova che $\sum_{i=1}^{p} ik_i = n$.

Ora passiamo alla dimostrazione del Lemma 4.

Dimostrazione del Lemma 4. Ragioniamo per induzione su i.

Supponiamo innanzitutto $i=1$. In tal caso abbiamo

$$\ker f = (Imf \cap \ker f) \oplus \mathcal{K}^{(1)}.$$

Percio'

$$k_1 = \dim\ker f - \dim(Imf \cap \ker f).$$

Ora osserviamo che $Imf \cap \ker f$ e' il nucleo della restrizione $\varphi : Imf \to Imf$ di f su Imf. Percio', per il Teorema della dimensione applicato a φ, abbiamo:

$$\dim(Imf \cap \ker f) = \dim\ker\varphi = \dim Imf - \dim Im\varphi = \dim Imf - \dim Imf^{2}.$$

Sostituendo, otteniamo:

$$k_1 = \dim\ker f - \dim(Imf \cap \ker f) = \dim\ker f - (\dim Imf - \dim Imf^{2})$$

$$= \dim Imf^{0} - 2\dim Imf + \dim Imf^{2}.$$

Ora assumiamo $i \geq 2$. Abbiamo:

$$k_i = \dim(Imf^{i-1} \cap \ker f) - \dim(Imf^i \cap \ker f)$$

e

$$k_{i-1} = \dim(Imf^{i-2} \cap \ker f) - \dim(Imf^{i-1} \cap \ker f).$$

Sommando, otteniamo:

$$k_i + k_{i-1} = \dim(Imf^{i-2} \cap \ker f) - \dim(Imf^i \cap \ker f).$$

Come prima, possiamo interpretare $Imf^i \cap \ker f$ come il nucleo della retrizione $\varphi : Imf^i \to Imf^i$ di f ad Imf^i (che ha come immagine Imf^{i+1}), e $Imf^{i-2} \cap \ker f$ come il nucleo della retrizione $\psi : Imf^{i-2} \to Imf^{i-2}$ di f ad Imf^{i-2} (che ha come immagine Imf^{i-1}). Applicando il Teorema della dimensione a φ ed a ψ, otteniamo:

$$k_i + k_{i-1} = (\dim Imf^{i-2} - \dim Imf^{i-i}) - (\dim Imf^i - \dim Imf^{i+i}).$$

Infine, applicando l'ipotesi induttiva a k_{i-1}, si ha:

$$\begin{aligned} k_i &= (\dim Imf^{i-2} - \dim Imf^{i-i}) - (\dim Imf^i - \dim Imf^{i+i}) - k_{i-1} \\ &= (\dim Imf^{i-2} - \dim Imf^{i-i}) - (\dim Imf^i - \dim Imf^{i+i}) \\ &\quad -(\dim Imf^{i-2} - 2\dim Imf^{i-1} + \dim Imf^i) = \\ &\quad \dim Imf^{i-1} - 2\dim Imf^i + \dim Imf^{i+1}. \end{aligned}$$

Il Lemma 4 e' dimostrato.

Cio' conclude l'algoritmo per costruire una base a stringhe $\mathcal{S}$ per f, e conclude la dimostrazione del Teorema di Jordan, per quel che riguarda l'esistenza di una base a stringhe. Nel prossimo paragrafo dimostreremo l'unicita' della matrice J_f.

Esempio 8. *Determinare la forma canonica di Jordan J ed una base a stringhe $\mathcal{S}$ per l'operatore $f : \mathbf{C}^4 \to \mathbf{C}^4$ definito ponendo $f(x, y, z, t) = (x + y + z, t, -y, x + z)$.*

Svolgimento. Cominciamo con l'osservare che la matrice rappresentativa A di f rispetto alla base canonica e'

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

Il polinomio caratteristico di f e' $p_f(t) = t^3(t-1)$. Quindi f possiede solo due autovalori, cioe' 0 e 1. Un calcolo prova che $m_a(1) = m_g(1) = 1$, $m_a(0) = 3$ e $m_g(0) = 1$. In base al Teorema di Jordan gia' siamo in grado di dire che

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

Infatti l'autovalore 1 *deve apparire con un solo blocco (perche'* $m_g(1) = 1$*) di grandezza* 1 *(perche'* $m_a(1) = 1$*), mentre l'autovalore* 0 *deve apparire anch'esso con un solo blocco (perche'* $m_g(0) = 1$*) ma di grandezza* 3 *(perche'* $m_a(0) = 3$*). Come vedremo in seguito questo e' un fatto di carattere generale, cioe' e' possibile calcolare la forma canonica di un operatore senza calcolarne esplicitamente una base a stringhe. Chiaramente la conoscenza di una base a stringhe fornisce piu' informazioni, e pertanto ha un costo maggiore in termini di calcolo. Ora andiamo a calcolare una base a stringhe.*

Sappiamo che $\mathbf{C}^4 = \widetilde{V}_1 \oplus \widetilde{V}_0$*.*

Per cio' che concerne $\widetilde{V}_1$ *sappiamo che, poiche'* $m_a(1) = m_g(1) = 1$*, allora* $\widetilde{V}_1 = V_1$*, ed una base a stringhe per l'operatore* f *ristretto su* $\widetilde{V}_1$ *e' data semplicemente da un autovettore relativo all'autovalore* 1*. Un calcolo mostra che* $\widetilde{V}_1 = Span((2, 1, -1, 1))$*.*

Andiamo ora a studiare l'operatore f *ristretto su* $\widetilde{V}_0$*. Chiamiamo* $\varphi : \widetilde{V}_0 \to \widetilde{V}_0$ *tale operatore. Dovendo l'autospazio* $\widetilde{V}_0$ *contribuire con il blocco di* J *relativo all'autovalore* 0*, sappiamo a priori che* φ *e' nilpotente con indice* $p = 3$*. Per poter studiare esplicitamente* φ *occorre calcolarne una matrice rappresentativa. Quindi innanzitutto ci calcoliamo una base di* $\widetilde{V}_0$*, semplicemente ricordando dalla definizione che* $\widetilde{V}_0$ *e' il nucleo della matrice*

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

Cioe' $\widetilde{V}_0$ *e' definito dall'equazione* $x + z = 0$*. Risolvendo tale equazione otteniamo una base* $\mathcal{B}$ *di* $\widetilde{V}_0$*, per esempio* $\mathcal{B} = \{\mathbf{e}_2, \mathbf{e}_3 - \mathbf{e}_1, \mathbf{e}_4\}$ *(tale risultato e' coerente con la proprieta' 2) degli autospazi generalizzati, che dice che* $\dim \widetilde{V}_0 = m_a(0)$*). La matrice rappresentativa di* φ *rispetto alla base* $\mathcal{B}$ *e' allora:*

$$M := M^{\mathcal{B}}_{\mathcal{B}}(\varphi) = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Si osservi che $\mathcal{B}$ *non e' una base a stringhe. Ora dobbiamo trovare una sola stringa di lunghezza* $p = 3$*. Possiamo applicare quanto imparato per calcolare le stringhe di lunghezza massima* p*. Cioe' dobbiamo calcolare una base per lo spazio delle colonne di* $M^{p-1} = M^2$*. Poiche'*

$$M^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

allora una base per lo spazio delle colonne di M^2 *e' costituita dalla terza colonna di* M^2*:*

$$\mathbf{u} := (0, -1, 0)^T.$$

Poiche' $M^2 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \mathbf{u}$, *allora la stringa cercata e'*

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \xrightarrow{M} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \xrightarrow{M} \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \xrightarrow{M} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Il calcolo precedente e' stato fatto utilizzando le coordinate rispetto alla base $\mathcal{B}$. *Quindi occorre tener presente che* $(0,0,1)^T$ *corrisponde ad* $\mathbf{e}_4$, $(1,0,0)^T$ *corrisponde ad* $\mathbf{e}_2$, *e che* $(0,-1,0)^T$ *corrisponde a* $\mathbf{e}_1 - \mathbf{e}_3$.

In conclusione una base a stringhe per f *e' data dal sistema di vettori:*

$$\mathcal{S} = \{2\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3 + \mathbf{e}_4,\, \mathbf{e}_1 - \mathbf{e}_3,\, \mathbf{e}_2,\, \mathbf{e}_4\}$$

(si osservi che abbiamo dovuto invertire l'ordine con cui appaiono i vettori nel grafico precedente).

Per svolgere una verifica si puo' considerare la matrice P *costituita dai vettori di* $\mathcal{S}$ *messi in colonna*

$$P := \begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix},$$

ed andare a vedere che $J = P^{-1}AP$ *(o anche che* $\det(P) \neq 0$, *e che* $AP = PJ$*).* ■

Esempio 9. *Determinare la forma canonica di Jordan* J *ed una base a stringhe* $\mathcal{S}$ *per l'operatore* $f : \mathbf{C}^4 \to \mathbf{C}^4$ *definito ponendo* $f(x,y,z,t) = (2z+t, 2y+z+t, 2z+t, 2t)$.

Svolgimento. La matrice rappresentativa A *di* f *rispetto alla base canonica e'*

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

Il polinomio caratteristico di f *e'* $p_f(t) = t(t-2)^3$. *Quindi gli autovalori di* f *sono* 0 *e* 2. *Poiche'* $m_a(0) = m_g(0) = 1$, $m_a(2) = 3$ *e* $m_g(2) = 1$ *possiamo dire che*

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

Sappiamo che $\mathbf{C}^4 = \widetilde{V}_0 \oplus \widetilde{V}_2$.

Poiche' $\widetilde{V}_0 = V_0$ *allora una base a stringhe per l'operatore* f *ristretto su* $\widetilde{V}_0$ *e' data semplicemente da un autovettore relativo all'autovalore* 0. *E' evidente che* $\widetilde{V}_0 = Span((1,0,0,0))$.

Andiamo a studiare l'operatore f *ristretto su* $\widetilde{V}_2$. *Sappiamo che tale operatore ha soltanto l'autovalore* 2. *Quindi, per ridurci al caso nilpotente, invece di studiare* f, *andiamo a studiare l'operatore* $f - 2id_V$ *ristretto su* $\widetilde{V}_2$. *Chiamiamo* $\varphi : \widetilde{V}_2 \to \widetilde{V}_2$ *tale operatore. Dovendo l'autospazio* $\widetilde{V}_2$ *contribuire con il blocco di* J *relativo all'autovalore* 2, *sappiamo a priori che* φ *e' nilpotente con indice* 3. *Per poter studiare esplicitamente* φ *occorre calcolarne una matrice rappresentativa. Quindi innanzitutto ci calcoliamo una base di* $\widetilde{V}_2$, *ricordando dalla definizione che* $\widetilde{V}_2$ *e' il nucleo della matrice*

$$(A-2I)^3 = \begin{bmatrix} -8 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Cioe' $\widetilde{V}_2$ *e' definito dall'equazione* $x - z = 0$. *Risolvendo tale equazione otteniamo una base* $\mathcal{B}$ *di* $\widetilde{V}_2$, *per esempio* $\mathcal{B} = \{\mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_3, \mathbf{e}_4\}$. *La matrice rappresentativa di* φ *rispetto alla base* $\mathcal{B}$ *e':*

$$M := M_{\mathcal{B}}^{\mathcal{B}}(\varphi) = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Ora dobbiamo trovare una sola stringa di lunghezza $p = 3$. *Cioe' dobbiamo calcolare una base per lo spazio delle colonne di* $M^{p-1} = M^2$. *Poiche'*

$$M^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

allora una base per lo spazio delle colonne di M^2 *e' costituita dalla terza colonna di* M^2:

$$\mathbf{u} := (1,0,0)^T.$$

Poiche' $M^2 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \mathbf{u}$, *allora la stringa cercata e'*

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \xrightarrow{M} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \xrightarrow{M} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \xrightarrow{M} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Il calcolo precedente e' stato fatto utilizzando le coordinate rispetto alla base $\mathcal{B}$. *Quindi occorre tener presente che* $(0,0,1)^T$ *corrisponde ad* $\mathbf{e}_4$, $(1,1,0)^T$ *corrisponde ad* $\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$, *e che* $(1,0,0)^T$ *corrisponde a* $\mathbf{e}_2$.

In conclusione una base a stringhe per f *e' data dal sistema di vettori:*

$$\mathcal{S} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_4\}$$

(si osservi che abbiamo dovuto invertire l'ordine con cui appaiono i vettori nel grafico precedente). ■

Esempio 10. *Determinare la forma canonica di Jordan* J *ed una base a stringhe* $\mathcal{S}$ *per l'operatore* $f : \mathbf{C}^5 \to \mathbf{C}^5$ *definito ponendo* $f(x, y, z, t, w) = (-3x + y, -3y, 7z + t, 7t + w, 7w)$.

Svolgimento. La matrice rappresentativa A *di* f *rispetto alla base canonica e'*

$$A = \begin{bmatrix} -3 & 1 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & 7 & 1 & 0 \\ 0 & 0 & 0 & 7 & 1 \\ 0 & 0 & 0 & 0 & 7 \end{bmatrix}.$$

Poiche' tale matrice e' gia' una matrice a blocchi di Jordan allora essa e' la forma canonica J *cercata, ed una base a stringhe e' la base canonica (in questo caso dunque e' inutile attivare l'algoritmo imparato in precedenza).* ■

Esempio 11. *Determinare la forma canonica di Jordan* J *ed una base a stringhe* $\mathcal{S}$ *per l'operatore* $f : \mathbf{C}^3 \to \mathbf{C}^3$ *definito ponendo* $f(x, y, z) = (-4x + z, -2x - 3y + 2z, -x - 2z)$.

Svolgimento. La matrice rappresentativa A *di* f *rispetto alla base canonica e'*

$$A = \begin{bmatrix} -4 & 0 & 1 \\ -2 & -3 & 2 \\ -1 & 0 & -2 \end{bmatrix}.$$

Il polinomio caratteristico di f *e'* $p_f(t) = -(t + 3)^3$. *Quindi* f *ha soltanto l'autovalore* -3. *Inoltre si ha* $m_a(-3) = 3$ *e* $m_g(-3) = 2$. *Quindi la forma canonica e' formata da due blocchi, cioe'*

$$J = \begin{bmatrix} -3 & 0 & 0 \\ 0 & -3 & 1 \\ 0 & 0 & -3 \end{bmatrix}.$$

Osserviamo che $\mathbf{C}^3 = \widetilde{V}_{-3}$, *ed il passo della separazione degli autovalori in questo caso e' inutile. Invece per rendere l'operatore nilpotente effettuiamo la traslazione* $g := f + 3id_{\mathbf{C}^3}$. *La matrice rappresentativa di* g *rispetto alla base canonica e':*

$$M := A + 3I = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}.$$

Tale matrice e' nilpotente con indice $p = 2$. *Cominciamo col determinare una stringa di lunghezza massima* $p = 2$. *Occorre fissare una base per lo spazio delle colonne di*

$M^{p-1} = M$: possiamo prendere la prima colonna $(-1,-2,-1)^T$. Allora la stringa cercata e':

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \xrightarrow{M} \begin{bmatrix} -1 \\ -2 \\ -1 \end{bmatrix} \xrightarrow{M} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

A questo punto occorre determinare una stringa di lunghezza $p-1=1$, cioe' un vettore del nucleo di M che sia linearmente indipendente dai due vettori trovati prima. Possiamo prendere $(0,1,0)^T$. In conclusione una base a stringhe per f e' data dal sistema di vettori:

$$\mathcal{S} = \{\mathbf{e}_2, -\mathbf{e}_1 - 2\mathbf{e}_2 - \mathbf{e}_3, \mathbf{e}_1\}. \blacksquare$$

Esempio 12. *Determinare la forma canonica di Jordan J ed una base a stringhe $\mathcal{S}$ per l'operatore di derivazione $D : \mathbf{C}[t]_{\leq 3} \to \mathbf{C}[t]_{\leq 3}$ definito ponendo $Dp(t) := p'(t)$.*

Svolgimento. L'operatore D ammette la stringa

$$t^3 \xrightarrow{D} 3t^2 \xrightarrow{D} 6t \xrightarrow{D} 6 \xrightarrow{D} 0.$$

Quindi una base a stringhe per D e' data da $\mathcal{S} = \{6, 6t, 3t^2, t^3\}$ e $J = J_{0,4}$. $\blacksquare$

5. Calcolo della forma canonica di Jordan senza conoscere una base a stringhe.

Come abbiamo gia' visto negli esempi precedenti, e' possibile calcolare la forma canonica di Jordan senza calcolare esplicitamente una base a stringhe. Questo nuovo algoritmo si basa sul seguente ragionamento. Possiamo limitarci a considerare il caso di un operatore nilpotente $f : V \to V$, con $\dim V = n$ ed indice di nilpotenza p. Ora consideriamo il seguente diagramma:

$$
\begin{array}{ccccccccccc}
\mu_1 & \mu_2 & . & . & . & . & . & . & . & . & \mu_p \\
* & * & . & . & . & . & . & . & . & . & * \\
* & * & . & . & . & . & . & . & . & . & * \\
. & . & . & . & . & . & . & . & . & . & . \\
* & * & . & . & . & . & . & . & . & . & * \\
 & * & * & . & . & . & . & . & . & . & * \\
 & * & * & . & . & . & . & . & . & . & * \\
 & . & . & . & . & . & . & . & . & . & . \\
 & * & * & . & . & . & . & . & . & . & * \\
 & & & & & & & & & & \\
 & & & . & . & . & . & . & . & . & . \\
 & & & & & & & & & & \\
 & & & & & & & & & * & * \\
 & & & & & & & & & * & * \\
 & & & & & & & & & . & . \\
 & & & & & & & & & * & * \\
 & & & & & & & & & & * \\
 & & & & & & & & & & * \\
 & & & & & & & & & & . \\
 & & & & & & & & & & *
\end{array}
$$

ottenuto nel seguente modo. Gli asterischi denotano vettori appartenenti ad una fissata (ma qualunque) base a stringhe $\mathcal{S}$. Essi sono disposti in riga, ciascuna riga rappresentando una stringa di data lunghezza. La stringa e' disposta all'incontrario, cioe' l'ultimo vettore sulla destra denota l'autovettore. Percio' l'ultima colonna del diagramma e' costituita da vettori che si annullano tramite f. Si comincia dall'alto disponendo in riga le stringhe di lunghezza massima p. Poi quelle di lunghezza $p-1$, e cosi' via fino a quelle di lunghezza 1. Denotiamo con μ_1, μ_2, $\dots$, μ_p il numero di vettori che appaiono in ciascuna *colonna*, e con $\sharp\, J_{0,i}$ il numero di blocchi di grandezza i che compaiono nella forma canonica J_f di f. Si osservi che per definizione si ha:

$$
(1) \qquad \mu_1 + \mu_2 + \cdots + \mu_p = n, \quad \text{e} \quad \mu_1 \leq \mu_2 \leq \cdots \leq \mu_p.
$$

Poiche' $\mathcal{S}$ genera V, allora $f(\mathcal{S})$ genera l'immagine di f. Quando applichiamo f ai vettori dell'ultima colonna, questi si annullano in quanto sono ultimi vettori di stringhe. I vettori delle altre colonne si spostano di un posto. Percio' $f(\mathcal{S})\backslash\{\mathbf{0}\}$ e' contenuto in $\mathcal{S}$, dunque e' un sistema libero, quindi una base per l'immagine. Inoltre, in $f(\mathcal{S})\backslash\{\mathbf{0}\}$ ci sono tutti i vettori di $\mathcal{S}$, tranne i primi vettori a sinistra con cui cominciano le stringhe nel diagramma, che sono tanti quanto il numero totale delle stringhe, cioe' μ_p. Cio' implica che $\dim Im(f) = n - \mu_p = \mu_1 + \mu_2 + \cdots + \mu_{p-1}$. Analogamente si ha

$\dim Im(f^2) = \mu_1 + \mu_2 + \cdots + \mu_{p-2}$ e cosi' continuando si ottengono le seguenti formule:

$$(2) \qquad \begin{cases} \mu_1 + \mu_2 + \cdots + \mu_{p-1} + \mu_p = n \\ \mu_1 + \mu_2 + \cdots + \mu_{p-1} = \dim Im(f) \\ \cdots\cdots\cdots\cdots\cdots \\ \mu_1 + \cdots + \mu_i = \dim Im(f^{p-i}) \\ \cdots\cdots \\ \mu_1 = \dim Im(f^{p-1}). \end{cases}$$

Risolvendo questo sistema lineare si calcolano i numeri μ_i $(1 \leq i \leq p)$:

$$\mu_i = \dim Imf^{p-i} - \dim Imf^{p-i+1}.$$

Dopodiche' sono determinati i numeri $\sharp\, J_{0,i}$ dei blocchi di Jordan di ordine i che appaiono in $M_{\mathcal{S}}^{\mathcal{S}}(f)$, in quanto, per ogni $1 \leq i \leq p$, si ha:

$$\sharp\, J_{0,i} = \mu_{p-i+1} - \mu_{p-i} \qquad (\mu_0 := 0).$$

Quindi

$$\sharp\, J_{0,i} = \dim Imf^{i-1} - 2\dim Imf^{i} + \dim Imf^{i+1}.$$

A questa formula siamo arrivati a partire da una *qualunque* base a stringhe $\mathcal{S}$ per f. Percio' il numero dei blocchi di Jordan $\sharp\, J_{0,i}$ di ordine i che appaiono nella matrice rappresentativa $M_{\mathcal{S}}^{\mathcal{S}}(f)$ non dipende dalla base a stringhe $\mathcal{S}$, ma solo da f. Cio' vuol dire proprio che la forma canonica di f e' unica, a meno dell'ordine con cui appaiono i blocchi (la dimostrazione di questo fatto era rimasta in sospeso).

Esempio 13. *Determinare la forma canonica di Jordan J della matrice*

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

Svolgimento. Il polinomio caratteristico della matrice e' $p_A(t) = (t+1)^4$. Quindi la matrice $B := A + I$ e' nilpotente, e poiche' $B^2 = \mathbf{0}$, ha indice $p = 2$. Le formule precedenti (2) diventano in questo caso:

$$\begin{cases} \mu_1 + \mu_2 = 4 \\ \mu_1 = 1. \end{cases}$$

Quindi per i blocchi di B si ha $\sharp\, J_{0,2} = 1$ e $\sharp\, J_{0,1} = 2$, cioe' la forma canonica di B e'

$$\begin{bmatrix} 0 & 1 & & \\ 0 & 0 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}.$$

E quella di A e'

$$\begin{bmatrix} -1 & 1 & & \\ 0 & -1 & & \\ & & -1 & \\ & & & -1 \end{bmatrix}. \quad \blacksquare$$

Esempio 14. *Una matrice A e' 10×10 ed ha indice di nilpotenza $p = 4$. Sapendo che A^2 ha rango 3, calcolare le possibili forme canoniche per A.*

Svolgimento. Per cio' che concerne le formule (2) possiamo dire che devono essere del tipo:

$$\begin{cases} \mu_1 + \mu_2 + \mu_3 + \mu_4 = 10 \\ \mu_1 + \mu_2 + \mu_3 = rk(A) \\ \mu_1 + \mu_2 = 3 \\ \mu_1 = rk(A^3). \end{cases}$$

Occorre determinare $rk(A)$ e $rk(A^3)$. Ricordiamo che i numeri μ_i sono interi e crescenti rispetto all'indice (si veda (1)). Quindi dovendo essere $1 \leq \mu_1 \leq \mu_2$ e $\mu_1 + \mu_2 = 3$ segue che $\mu_1 = 1$ e che $\mu_2 = 2$. Quindi nella forma canonica di A deve esserci un solo blocco di grandezza 4 ed un solo blocco di grandezza 3. Con un ragionamento analogo si vede che μ_3 puo' essere 2 oppure 3. Cioe' i blocchi rimanenti o sono tre blocchi di grandezza 1, oppure ce n'e' uno di grandezza 2 ed uno di grandezza 1. In conclusione le possibili forme canoniche per A sono date dalle matrici:

$$\begin{bmatrix} 0 & 1 & 0 & 0 & & & & & & \\ 0 & 0 & 1 & 0 & & & & & & \\ 0 & 0 & 0 & 1 & & & & & & \\ 0 & 0 & 0 & 0 & & & & & & \\ & & & & 0 & 1 & 0 & & & \\ & & & & 0 & 0 & 1 & & & \\ & & & & 0 & 0 & 0 & & & \\ & & & & & & & 0 & 1 & \\ & & & & & & & 0 & 0 & \\ & & & & & & & & & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 & 0 & & & & & & \\ 0 & 0 & 1 & 0 & & & & & & \\ 0 & 0 & 0 & 1 & & & & & & \\ 0 & 0 & 0 & 0 & & & & & & \\ & & & & 0 & 1 & 0 & & & \\ & & & & 0 & 0 & 1 & & & \\ & & & & 0 & 0 & 0 & & & \\ & & & & & & & 0 & & \\ & & & & & & & & 0 & \\ & & & & & & & & & 0 \end{bmatrix}. \quad \blacksquare$$

6. Il Teorema di Cayley-Hamilton.

Una interessante conseguenza del Teorema di Jordan e' il famoso Teorema di Cayley-Hamilton. Prima di enunciare tale teorema ci occorre qualche nozione preliminare.

Innanzitutto fissiamo un polinomio $\varphi(t) = a_0 + a_1 t + \cdots + a_h t^h$. Assegnato un numero $c \in \mathbf{C}$ possiamo valutare il polinomio in c. Si ottiene un altro numero, che si indica con $\varphi(c)$. Ad esempio, se $\varphi(t) = 3 + t - 2t^2$ allora $\varphi(-2) = -7$. In modo analogo possiamo valutare il polinomio in una matrice quadrata A, ponendo

$$\varphi(A) := a_0 I + a_1 A + \cdots + a_h A^h.$$

E, assegnato un operatore $f : V \to V$, possiamo definire

$$\varphi(f) := a_0 id_V + a_1 f + \cdots + a_h f^h$$

come quell'operatore $\varphi(f) : V \to V$ che al vettore $\mathbf{u} \in V$ associa il vettore $\varphi(f)(\mathbf{u}) = a_0\mathbf{u} + a_1 f(\mathbf{u}) + \cdots + a_h f^h(\mathbf{u})$, essendo $f^r := f \circ f \circ \cdots \circ f$ la composizione di f con se' stesso r volte. Si osservi che se A e' una matrice che rappresenta f rispetto a qualche base, allora $\varphi(A)$ rappresenta $\varphi(f)$ rispetto alla stessa base. Inoltre, comunque si assegnino polinomi $\varphi(t)$ e $\psi(t)$, una matrice A ed un operatore f, valgono le seguenti proprieta' di calcolo:

$$(\varphi + \psi)(A) = \varphi(A) + \psi(A), \quad (\varphi \cdot \psi)(A) = \varphi(A) \cdot \psi(A) = \psi(A) \cdot \varphi(A)$$

$$(\varphi + \psi)(f) = \varphi(f) + \psi(f), \quad (\varphi \cdot \psi)(f) = \varphi(f) \cdot \psi(f) = \psi(f) \cdot \varphi(f).$$

Possiamo ora enunciare il:

Teorema (di Cayley-Hamilton). *Sia $f : V \to V$ un operatore lineare definito su uno spazio vettoriale complesso V di dimensione finita. Sia $p_f(t)$ il polinomio caratteristico di f. Allora $p_f(f)$ e' l'operatore nullo, cioe'*

$$p_f(f) = \mathbf{0}.$$

Se $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$ rappresenta f, allora $p_f(f) = \mathbf{0}$ equivale a dire che $p_f(A) = \mathbf{0}$.

Esempio 15. *Il caso di dimensione 2 si puo' esaminare con un calcolo diretto. Sia $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ una qualunque matrice 2×2. In questo caso si ha*

$$p_A(t) = \det \begin{bmatrix} a-t & b \\ c & d-t \end{bmatrix} = ad - bc - (a+d)t + t^2.$$

Il Teorema di Cayley-Hamilton ci dice che

$$(ad - bc) \cdot I - (a+d)A + A^2 = \mathbf{0}.$$

Verifichiamo questa uguaglianza:

$$(ad - bc) \cdot I - (a+d)A + A^2 = (ad - bc) \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - (a+d) \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix}^2$$

$$= \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} - \begin{bmatrix} a^2+ad & ab+db \\ ac+dc & ad+d^2 \end{bmatrix} + \begin{bmatrix} a^2+bc & ab+bd \\ ac+dc & bc+d^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \quad \blacksquare$$

Ora andiamo a dimostrare il Teorema di Cayley-Hamilton nel caso generale, cioe' di dimensione qualunque. La dimostrazione si puo' vedere come una applicazione del Teorema di Jordan. Piu' precisamente come una conseguenza della decomposizione di V come somma dei suoi autospazi generalizzati. Vediamo.

Siano $\lambda_1, \lambda_2, \ldots, \lambda_k$ gli autovalori distinti di f, e

$$p_f(t) = (t - \lambda_1)^{p_1} \cdot (t - \lambda_2)^{p_2} \cdot \ldots \cdot (t - \lambda_k)^{p_k}$$

la corrispondente decomposizione di $p_f(t)$, dove $p_i = m_a(\lambda_i)$. Consideriamo anche la decomposizione di V nella somma diretta dei suoi autospazi generalizzati

$$V = \widetilde{V}_{\lambda_1} \oplus \cdots \oplus \widetilde{V}_{\lambda_k}.$$

Sia $\mathbf{u}$ un qualunque vettore di V. Noi vogliamo provare che $p_f(f)(\mathbf{u}) = \mathbf{0}$. In virtu' della decomposizione in autospazi generalizzati, possiamo scrivere

$$\mathbf{u} = \mathbf{u}_1 + \cdots + \mathbf{u}_k$$

secondo opportuni autovettori generalizzati $\mathbf{u}_i \in \widetilde{V}_{\lambda_i}$. Poiche' $p_f(f)(\mathbf{u}) = p_f(f)(\mathbf{u}_1) + \cdots + p_f(f)(\mathbf{u}_k)$ sara' sufficiente provare che $p_f(f)(\mathbf{u}_i) = \mathbf{0}$ per ogni $i = 1, \ldots, k$.

Adesso, tenendo presente che $\widetilde{V}_{\lambda_i} = \ker(f - \lambda_i id_V)^{p_i}$, si ha:

$$p_f(f)(\mathbf{u}_i) = (f - \lambda_1 id_V)^{p_1} \cdot (f - \lambda_2 id_V)^{p_2} \cdot \ldots \cdot (f - \lambda_k id_V)^{p_k}(\mathbf{u}_i)$$

$$= \left[\prod_{j \neq i} (f - \lambda_j id_V)^{p_j} \right] \cdot (f - \lambda_i id_V)^{p_i}(\mathbf{u}_i) = \mathbf{0}.$$

Cio' conclude la dimostrazione del Teorema di Cayley-Hamilton. ■

7. Il polinomio minimo di un operatore.

Sia $f : V \to V$ un operatore lineare definito su uno spazio vettoriale di dimensione finita $n \geq 1$. Denotiamo con $\mathcal{I}_f$ l'insieme di tutti i polinomi non nulli $\varphi(t)$ tali che $\varphi(f) = \mathbf{0}$. In altre parole $\mathcal{I}_f$ e' l'insieme delle equazioni algebriche (non banali) in una variabile, soddisfatte da f. In base al Teorema di Cayley-Hamilton sappiamo che $\mathcal{I}_f$ non e' vuoto, cioe' che esistono polinomi non nulli $\varphi(t)$ tali che $\varphi(f) = \mathbf{0}$. Per esempio, il polinomio caratteristico di f. Ha senso pertanto considerare, tra tutti questi polinomi, quelli di grado minimo (tale grado sara' allora $\leq n$). Sia $m_f(t)$ un tale polinomio. Dividendo eventualmente per il coefficiente direttore, possiamo sempre supporre che $m_f(t)$ sia monico. Un siffatto polinomio si dice *polinomio minimo per* f, e si puo' dimostrare che esso e' unico. Cio' segue dalla seguente Proposizione 5. Similmente si puo' definire l'insieme $\mathcal{I}_A$ per una data matrice quadrata A, e si puo' definire il polinomio minimo $m_A(t)$ di A. Se A rappresenta l'operatore f e' chiaro che $\mathcal{I}_f = \mathcal{I}_A$. Per cui $m_f(t) = m_A(t)$, e matrici simili hanno lo stesso polinomio minimo. In particolare $m_A(t) = m_{J_A}(t)$.

Proposizione 5. *Sia $\varphi(t)$ un polinomio che si annulla sull'operatore f, tale cioe' che $\varphi(f) = \mathbf{0}$. Allora il polinomio minimo di f e' un fattore di $\varphi(t)$, cioe' esiste un polinomio $q(t)$ tale che $\varphi(t) = q(t) \cdot m_f(t)$. In particolare il polinomio minimo di f e' unico.*

Dimostrazione. Dividendo $\varphi(t)$ per $m_f(t)$ possiamo scrivere $\varphi(t) = q(t)m_f(t) + r(t)$, dove $q(t)$ ed $r(t)$ denotano quoziente e resto. Noi vogliamo provare che il resto e' il polinomio nullo. Sostituendo alla variabile t l'operatore f in $\varphi(t) = q(t)m_f(t) + r(t)$, deduciamo $r(f) = \mathbf{0}$ (si tenga presente che $m_f(f) = \mathbf{0}$). Questo implica che $r(t)$ e' il polinomio nullo, altrimenti sarebbe un polinomio non nullo di grado strettamente minore del grado di $m_f(t)$, che si annulla in f: questo contraddice il fatto che $m_f(t)$ e' il polinomio minimo.

Per concludere la dimostrazione dobbiamo provare che il polinomio minimo e' unico. Siano $m_1(t)$ ed $m_2(t)$ due polinomi minimi per f. Per quanto appena provato possiamo scrivere $m_2(t) = q(t)m_1(t)$. Poiche' $m_1(t)$ ed $m_2(t)$ hanno lo stesso grado allora $q(t)$ e' un polinomio costante, e deve essere la costante 1 perche' $m_1(t)$ ed $m_2(t)$ sono monici. Dunque $q(t) \equiv 1$, cioe' $m_2(t) = m_1(t)$. ■

In altre parole, il polinomio minimo $m_f(t)$ di un operatore f e' l'unico polinomio monico avente grado minimo rispetto alla condizione $m_f(f) = \mathbf{0}$, ed ogni polinomio che si annulla in f e' multiplo di $m_f(t)$. Il calcolo esplicito del polinomio minimo e' immediato una volta che sia nota la forma canonica di Jordan di f. Per esempio, supponiamo che f sia nilpotente di indice p. Allora e' chiaro che $m_f(t) = t^p$. Infatti, sappiamo che $f^p = \mathbf{0}$, percio', per la Proposizione precedente, il polinomio minimo deve essere un fattore di t^p. D'altra parte, per $q < p$, si ha $f^q \neq \mathbf{0}$, altrimenti p non sarebbe l'indice di nilpotenza di f. Percio' il polinomio minimo di un endomorfismo f nilpotente con indice di nilpotenza p e' $m_f(t) = t^p$. Questo fatto si puo' generalizzare. Infatti vale la seguente proposizione, la cui dimostrazione si riconduce facilmente al caso nilpotente.

Proposizione 6. *Siano $\lambda_1, \lambda_2, \dots, \lambda_k$ gli autovalori distinti di f. Per ogni autovalore λ_i sia h_i la grandezza massima tra i blocchi di Jordan relativi all'autovalore λ_i che appaiono nella forma canonica di f. Allora*

$$m_f(t) = (t - \lambda_1)^{h_1} \cdot (t - \lambda_2)^{h_2} \cdot \dots \cdot (t - \lambda_k)^{h_k}.$$

Dimostrazione. Per ogni $j = 1, \dots, k$, sia φ_j la restrizione di f su $\widetilde{V}_{\lambda_j}$, e $\psi_j := \varphi_j - \lambda_j \mathrm{id}_{\widetilde{V}_{\lambda_j}}$. Ora, ψ_j e' nilpotente con indice h_j. Quindi il polinomio minimo di ψ_j e' t^{h_j}. Ne consegue che il polinomio minimo di φ_j e' $(t - \lambda_j)^{h_j}$. In particolare, $(t - \lambda_j)^{h_j}$ si annulla in φ_j. Ma allora, come nella dimostrazione del Teorema di Cayley-Hamilton, il polinomio $m(t) := (t - \lambda_1)^{h_1} \cdot (t - \lambda_2)^{h_2} \cdot \dots \cdot (t - \lambda_k)^{h_k}$ si annulla in f. Quindi, per la Proposizione 5, $m_f(t)$ e' un divisore di $m(t)$, cioe' e' della forma $m_f(t) = (t - \lambda_1)^{q_1} \cdot (t - \lambda_2)^{q_2} \cdot \dots \cdot (t - \lambda_k)^{q_k}$ con $q_j \leq h_j$. D'altra parte, poiche' $m_f(t)$ si annulla in f, si deve annullare anche in ciascuna restrizione φ_j. Allora, come prima, ciascun $(t - \lambda_j)^{h_j}$ deve essere un fattore di $m_f(t)$, e percio' $h_j \leq q_j$ per ogni j. ■

Come conseguenze della proposizione precedente osserviamo che: (i) lo spettro del polinomio minimo coincide con quello di f, cioe' $Spec(m_f(t)) = Spec(p_f(t))$; (ii) f e' diagonalizzabile se e solo se tutte le radici del polinomio minimo di f hanno molteplicita' algebrica 1 per $m_f(t)$ (sappiamo gia' che se il polinomio caratteristico soddisfa questa proprieta' allora f e' diagonalizzabile, ma non vale il viceversa, per esempio id_V e' diagonalizzabile, ma $p_{id_V}(t) = (t-1)^{\dim V}$).

Esempio 16. *Sia f un operatore che ammette come matrice rappresentativa la seguente matrice*

$$J = \begin{bmatrix} \sqrt{2} & 1 & & & & & & & & \\ 0 & \sqrt{2} & & & & & & & & \\ & & \sqrt{2} & & & & & & & \\ & & & 0 & 1 & 0 & & & & \\ & & & 0 & 0 & 1 & & & & \\ & & & 0 & 0 & 0 & & & & \\ & & & & & & 0 & 1 & 0 & 0 \\ & & & & & & 0 & 0 & 1 & 0 \\ & & & & & & 0 & 0 & 0 & 1 \\ & & & & & & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Allora abbiamo

$$p_f(t) = t^7(t-\sqrt{2})^3, \qquad m_f(t) = t^4(t-\sqrt{2})^2. \quad \blacksquare$$

Esempio 17. *Dimostrare che una matrice A tale che $A^2 = I$ e' diagonalizzabile.*

Svolgimento. La matrice si annulla nel polinomio t^2-1. Quindi il polinomio minimo di A e' uno dei fattori di tale polinomio. Per cui $m_A(t)$ e' uguale a $t-1$, a $t+1$ oppure a $t^2-1 = (t-1)(t+1)$. In ogni caso $m_A(t)$ e' prodotto di polinomi distinti di primo grado. Cio' prova che A e' diagonalizzabile. Si osservi che se $m_A(t) = t-1$ allora $A = I$, se $m_A(t) = t+1$ allora $A = -I$, e che se $m_A(t) = t^2-1$ allora la forma canonica di A e' la matrice diagonale D che ha sulla diagonale principale solo 1 e -1, ed il numero delle entrate uguali ad 1 e' pari a $n - rk(A-I)$, dove n e' la grandezza della matrice A. In particolare tutte e sole le matrici che soddisfano l'equazione $A^2 = I$ sono le matrici della forma PDP^{-1}, con P matrice invertibile. $\blacksquare$

Esempio 18. *Dimostrare che una matrice A tale che $A^2 = A$ e' diagonalizzabile.*

Svolgimento. La matrice si annulla nel polinomio t^2-t. Quindi il polinomio minimo di A e' uno dei fattori di tale polinomio. Per cui $m_A(t)$ e' uguale a $t-1$, a t oppure a $t^2-t = t(t-1)$. In ogni caso $m_A(t)$ e' prodotto di polinomi distinti di primo grado. Cio' prova che A e' diagonalizzabile. Si osservi che se $m_A(t) = t-1$ allora $A = I$, se $m_A(t) = t$ allora $A = \mathbf{0}$, e che se $m_A(t) = t^2-t$ allora la forma canonica di A e' la matrice diagonale D che ha sulla diagonale principale solo 1 e 0, ed il numero *delle entrate uguali ad* 1 *e' pari al rango di* A*. In particolare tutte e sole le matrici che soddisfano l'equazione* $A^2 = A$ *sono le matrici della forma* PDP^{-1}*, con* P *matrice invertibile. Per esempio tutte e sole le matrici* A 3×3*, aventi rango* 2 *e tali che* $A^2 = A$ *sono le matrici*

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

al variare di P *nell'insieme delle matrici invertibili.* ■

Esempio 19. *Sia* A *una matrice* 4×4 *tale che* $A^4 = 2A^3 - A^2$ *e* $A^4 = A^2$*. Determinare le possibili forme canoniche per* A*.*

Svolgimento. La matrice si annulla sia nel polinomio $t^4 - 2t^3 + t^2 = t^2(t-1)^2$*, sia nel polinomio* $t^4 - t^2 = t^2(t-1)(t+1)$*. Poiche'* $m_A(t)$ *e' un fattore di entrambi i polinomi, il polinomio minimo di* A *e' un fattore del massimo comun divisore dei due polinomi, cioe'* $m_A(t)$ *deve essere un fattore del polinomio* $t^2(t-1)$*. Cioe'* $m_A(t)$ *deve essere uno dei seguenti polinomi:* t*,* $t-1$*,* t^2*,* $t(t-1)$*,* $t^2(t-1)$*. Se* $m_A(t) = t$ *allora* $A = \mathbf{0}$*, se* $m_A(t) = t-1$ *allora* $A = I$*, e se* $m_A(t) = t^2$ *allora la forma canonica di* A *puo' essere una soltanto delle seguenti due matrici:*

$$\begin{bmatrix} 0 & 1 & & \\ 0 & 0 & & \\ & & 0 & 1 \\ & & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & & \\ 0 & 0 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}.$$

Poi se $m_A(t) = t(t-1)$ *allora la forma canonica di* A *puo' essere una soltanto delle seguenti tre matrici:*

$$\begin{bmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{bmatrix}.$$

Infine se $m_A(t) = t^2(t-1)$ *allora la forma canonica di* A *puo' essere una soltanto delle seguenti due matrici:*

$$\begin{bmatrix} 0 & 1 & & \\ 0 & 0 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & & \\ 0 & 0 & & \\ & & 1 & \\ & & & 0 \end{bmatrix}. \quad ■$$

Bibilografia consigliata per questo capitolo:

S. Abeasis, *Complementi di algebra lineare*, Ed. Zanichelli.

K. Hoffman-R. Kunze, *Linear Algebra*, Ed. Prentice-Hall.

S. Lipschutz, *Algebra lineare*, Ed. Schaum.

S. Mac Lane-G. Birkhoff, *Algebra*, Ed. Mursia.

A. Silva, *Algebra lineare*, Ed. Nuova Cultura