

Forma canonica di Jordan di un operatore

Introduzione

Sia V uno spazio vettoriale non nullo finitamente generato, di dimensione $\dim V = n$, su un campo \mathbb{K} e sia $f : V \rightarrow V$ un *endomorfismo*, o *operatore lineare*, di V . Come è noto, sotto certe condizioni (espresse nel Teorema Spettrale) f è diagonalizzabile, ovvero esiste una base \mathcal{D} di V , detta *spettrale*, costituita da autovettori. La matrice $D = M_{\mathcal{D}}^{\mathcal{D}}(f)$, associata ad f rispetto a tale base spettrale, risulta diagonale e si ha

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

dove gli scalari $\lambda_1, \lambda_2, \dots, \lambda_n$ che compaiono sulla diagonale sono gli autovalori (non necessariamente tutti distinti) di f . Se invece f non è diagonalizzabile, ma vale comunque la prima parte della condizione (d) del Teorema Spettrale, è possibile trovare una base \mathcal{C} costituita da alcuni autovettori e da altri vettori, detti *autovettori generalizzati*, tale che la matrice $C = M_{\mathcal{C}}^{\mathcal{C}}(f)$ risulti *quasi* diagonale, nel senso che possono comparire, oltre agli autovalori già menzionati sulla diagonale, anche degli 1 sulla sopradiagonale, che è costituita dagli elementi di posizione $i, i + 1$, ($i = 1, \dots, n - 1$). Nelle applicazioni risulta spesso utile determinare una tale base \mathcal{C} .

Osserviamo esplicitamente che la prima parte della condizione (d) del Teorema Spettrale equivale al fatto che il polinomio caratteristico di $p_f(x)$ di f sia completamente riducibile, ovvero che, detti $\lambda_1, \lambda_2, \dots, \lambda_k$ gli autovalori distinti di f ed indicata con il simbolo $m_a(\lambda_j)$ la molteplicità algebrica di λ_j , si abbia

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

e tale condizione è sempre verificata se il campo \mathbb{K} è algebricamente chiuso, ad esempio se $\mathbb{K} = \mathbb{C}$.

Matrici diagonali a blocchi e sottospazi invarianti

Consideriamo una matrice quadrata A di ordine n suddivisa in blocchi quadrati:

$$A = \left(\begin{array}{c|c|c|c} A_{1,1} & A_{1,2} & \cdots & A_{1,k} \\ \hline A_{2,1} & A_{2,2} & \cdots & A_{2,k} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline A_{k,1} & A_{k,2} & \cdots & A_{k,k} \end{array} \right).$$

1 DEFINIZIONE. La matrice A si dice quadrata a blocchi se i blocchi $A_{i,j}$, con $i \neq j$, sono nulli. Scriveremo allora, per semplicità,

$$A = \begin{pmatrix} A_{1,1} & & & \\ & A_{2,2} & & \\ & & \ddots & \\ & & & A_{k,k} \end{pmatrix} \quad (1)$$

In tale situazione è chiaro che n è la somma degli ordini dei blocchi $A_{1,1}, A_{2,2}, \dots, A_{k,k}$.

2 ESEMPIO. La matrice

$$A = \begin{pmatrix} A_{1,1} & \\ & A_{2,2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 5 & & \\ 1 & 5 & 4 & & \\ 8 & 8 & 7 & & \\ & & & 1 & 5 \\ & & & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 5 & 0 & 0 \\ 1 & 5 & 4 & 0 & 0 \\ 8 & 8 & 7 & 0 & 0 \\ 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 & 4 \end{pmatrix}$$

è diagonale a blocchi, di ordine 5, e i suoi blocchi diagonali sono di ordine 3 e 2.

Sia ora A come in (1), e sia B dello stesso tipo, ovvero

$$B = \begin{pmatrix} B_{1,1} & & & \\ & B_{2,2} & & \\ & & \ddots & \\ & & & B_{k,k} \end{pmatrix}$$

dove $A_{i,i}$ e $B_{i,i}$ sono dello stesso ordine, per ogni i .

3 LEMMA. Valgono le seguenti proprietà:

1. $\det A = \det A_{1,1} \cdot \det A_{2,2} \cdot \dots \cdot \det A_{k,k}$.
2. $A \cdot B$ e $A + B$ sono ancora diagonali a blocchi dello stesso tipo, e si ha

$$A \cdot B = \begin{pmatrix} A_{1,1} \cdot B_{1,1} & & & \\ & A_{2,2} \cdot B_{2,2} & & \\ & & \ddots & \\ & & & A_{k,k} \cdot B_{k,k} \end{pmatrix} ;$$

$$A + B = \begin{pmatrix} A_{1,1} + B_{1,1} & & & & \\ & A_{2,2} + B_{2,2} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & A_{k,k} + B_{k,k} \end{pmatrix}.$$

3. A è invertibile se e solo se lo è ogni suo blocco diagonale, e si ha che

$$A^{-1} = \begin{pmatrix} A_{1,1}^{-1} & & & & \\ & A_{2,2}^{-1} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & A_{k,k}^{-1} \end{pmatrix}$$

Introduciamo ora il concetto di sottospazio invariante rispetto ad un endomorfismo.

4 DEFINIZIONE. Sia $f : V \rightarrow V$ un endomorfismo dello spazio vettoriale V . Un sottospazio $U \leq V$ si dice f -invariante, o, più semplicemente, invariante, se $f(U) \leq U$, ovvero per ogni $\mathbf{u} \in U$ si ha anche che $f(\mathbf{u}) \in U$.

Ad esempio, i sottospazi banali $\{\mathbf{0}\}$ e V sono sempre invarianti. Se λ è un autovalore, l'autospazio V_λ è invariante. Il congiungente di sottospazi invarianti è ancora invariante. Se U un sottospazio invariante, f induce, per restrizione, un endomorfismo $f_U : U \rightarrow U$ di U .

5 TEOREMA. Siano $U_1, \dots, U_k \leq V$ sottospazi invarianti rispetto all'endomorfismo f e sia $V = U_1 \oplus \dots \oplus U_k$. Consideriamo, per ogni h , una base \mathcal{B}_h di U_h e la base $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ di V . La matrice $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$ è diagonale a blocchi, come in (1), dove $A_{h,h} = M_{\mathcal{B}_h}^{\mathcal{B}_h}(f_{U_h})$, per ogni h .

In una matrice quadrata $A = (a_{i,j})$ è utile, talvolta, considerare gli elementi posizionati immediatamente al di sopra della diagonale, ovvero gli elementi del tipo $a_{i,i+1}$ ($1 \leq i < n$):

$$A = \begin{pmatrix} * & a_{1,2} & * & \cdots & * \\ * & * & a_{2,3} & \cdots & * \\ * & * & * & \cdots & * \\ \vdots & \vdots & \vdots & & \vdots \\ * & * & * & \cdots & a_{n-1,n} \\ * & * & * & \cdots & * \end{pmatrix},$$

dove gli asterischi, in questo caso, indicano scalari qualunque. Chiameremo *sopradiagonale* il vettore numerico $(a_{1,2}, a_{2,3}, \dots, a_{n-1,n})$, che ha ordine $n - 1$.

6 DEFINIZIONE. Sia $\lambda \in \mathbb{K}$, $p \in \mathbb{N}$. Si dice *blocco di Jordan di ordine p relativo a λ* la matrice quadrata $J_{\lambda,p}$ di ordine p con lo scalare λ sulla diagonale e lo scalare 1 sulla sopradiagonale:

$$J_{\lambda,p} = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}.$$

Ad esempio

$$J_{\lambda,3} = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}; \quad J_{0,3} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}; \quad J_{0,1} = (0); \quad J_{0,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Osserviamo che $J_{\lambda,p} = \lambda \cdot I_p + J_{0,p}$, e che $\det J_{\lambda,p} = \lambda^p$. Pertanto $J_{\lambda,p}$ è invertibile se e solo se $\lambda \neq 0$. Il polinomio caratteristico di $J_{\lambda,p}$ è $(\lambda - x)^p$ e λ è l'unico autovalore di $J_{\lambda,p}$, di molteplicità algebrica $m_a(\lambda) = p$ e geometrica $m_g(\lambda) = \dim V_\lambda = 1$.

7 ESERCIZIO. a)

$$J_{\lambda,p}^2 = \begin{pmatrix} \lambda^2 & 2\lambda & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \lambda^2 & 2\lambda & 1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \lambda^2 & 2\lambda & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda^2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \lambda^2 & 2\lambda & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda^2 & 2\lambda \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \lambda^2 \end{pmatrix};$$

- b) $J_{0,p}^2 = (c_{i,j})$, dove $c_{i,j} = \delta_{i+2,j}$;
c) $J_{0,p}^3 = (c_{i,j})$, dove $c_{i,j} = \delta_{i+3,j}$;
d) $J_{0,p}^h = (c_{i,j})$, dove $c_{i,j} = \delta_{i+h,j}$; $J_{0,p}^h = O$ se $h \geq p$.

Ad esempio

$$J_{0,4} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} ; J_{0,4}^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} ;$$

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

8 ESEMPIO. Sia $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ una base ordinata di V , λ uno scalare, e definiamo un endomorfismo $f : V \rightarrow V$ ponendo

$$f(\mathbf{e}_1) = \lambda \mathbf{e}_1, f(\mathbf{e}_2) = \mathbf{e}_1 + \lambda \mathbf{e}_2, f(\mathbf{e}_3) = \mathbf{e}_2 + \lambda \mathbf{e}_3, \dots, f(\mathbf{e}_n) = \mathbf{e}_n + \lambda \mathbf{e}_{n-1},$$

ed estendendo poi per linearità. In tale situazione si ha che $M_{\mathcal{B}}^{\mathcal{B}}(f) = J_{\lambda,n}$, e $V_{\lambda} = \mathcal{L}(\mathbf{e}_1)$.

9 DEFINIZIONE. Sia $f : V \rightarrow V$ un endomorfismo e $\mathcal{C} = (\mathbf{u}_1, \dots, \mathbf{u}_p)$ una p -pla di vettori, con $\mathbf{u}_1 \neq \mathbf{0}$. \mathcal{C} è una λ -stringa (di lunghezza p) se

$$f(\mathbf{u}_1) = \lambda \mathbf{u}_1, f(\mathbf{u}_2) = \mathbf{u}_1 + \lambda \mathbf{u}_2, f(\mathbf{u}_3) = \mathbf{u}_2 + \lambda \mathbf{u}_3, \dots, f(\mathbf{u}_p) = \mathbf{u}_{p-1} + \lambda \mathbf{u}_p,$$

ovvero $f(\mathbf{u}_h) = \mathbf{u}_{h-1} + \lambda \mathbf{u}_h$, dove $h = 1, \dots, p$ e si intende $\mathbf{u}_0 = \mathbf{0}$.

In tale situazione è chiaro che \mathbf{u}_1 è un autovettore, associato all'autovalore λ , e si scrive talvolta

$$\mathbf{u}_p \xrightarrow{f - \lambda \text{id}_V} \mathbf{u}_{p-1} \xrightarrow{f - \lambda \text{id}_V} \mathbf{u}_{p-2} \xrightarrow{f - \lambda \text{id}_V} \dots \xrightarrow{f - \lambda \text{id}_V} \mathbf{u}_1 \xrightarrow{f - \lambda \text{id}_V} \mathbf{0} .$$

Il vettore \mathbf{u}_p si dice *leader* della stringa.

10 PROPOSIZIONE. Sia \mathcal{C} una λ -stringa di lunghezza p . Si ha che \mathcal{C} è indipendente.

Dimostrazione. Sia $\mathcal{C} = (\mathbf{u}_1, \dots, \mathbf{u}_p)$. Poiché $\mathbf{u}_1 \neq \mathbf{0}$, è chiaro che $\mathbf{u}_h \neq \mathbf{0}$ per ogni h . Procediamo per induzione. L'asserto è banalmente vero se $p = 1$. Sia ora $p > 1$. Supponiamo induttivamente che l'asserto sia vero per stringhe di lunghezza minore di p e consideriamo una combinazione nulla di \mathcal{C} :

$$a_1 \mathbf{u}_1 + \dots + a_{p-1} \mathbf{u}_{p-1} + a_p \mathbf{u}_p = \mathbf{0} .$$

Se fosse $a_p \neq 0$ avremmo

$$\mathbf{u}_p = \left(-\frac{a_1}{a_p}\right) \mathbf{u}_1 + \cdots + \left(-\frac{a_{p-1}}{a_p}\right) \mathbf{u}_{p-1} .$$

Pertanto

$$\begin{aligned} \mathbf{0} &= f(\mathbf{0}) = f(a_1 \mathbf{u}_1 + \cdots + a_{p-1} \mathbf{u}_{p-1} + a_p \mathbf{u}_p) \\ &= a_1 f(\mathbf{u}_1) + \cdots + a_{p-1} f(\mathbf{u}_{p-1}) + a_p f(\mathbf{u}_p) \\ &= a_1 \lambda \mathbf{u}_1 + a_2 \lambda \mathbf{u}_2 + \cdots + a_{p-1} \lambda \mathbf{u}_{p-1} + a_p \lambda \mathbf{u}_p \\ &\quad + a_2 \mathbf{u}_1 + \cdots + a_{p-1} \mathbf{u}_{p-2} + a_p \mathbf{u}_{p-1} \\ &= a_1 \lambda \mathbf{u}_1 + a_2 \lambda \mathbf{u}_2 + \cdots + a_{p-1} \lambda \mathbf{u}_{p-1} \\ &\quad + a_2 \mathbf{u}_1 + \cdots + a_{p-1} \mathbf{u}_{p-2} + a_p \mathbf{u}_{p-1} \\ &\quad + a_p \lambda \left(-\frac{a_1}{a_p}\right) \mathbf{u}_1 + a_p \lambda \left(-\frac{a_2}{a_p}\right) \mathbf{u}_2 + \cdots + a_p \lambda \left(-\frac{a_{p-1}}{a_p}\right) \mathbf{u}_{p-1} \\ &= a_2 \mathbf{u}_1 + \cdots + a_{p-1} \mathbf{u}_{p-2} + a_p \mathbf{u}_{p-1} . \end{aligned}$$

Essendo $a_p \neq 0$, questa sarebbe una relazione di dipendenza per \mathcal{C}' , contro l'ipotesi induttiva. Deve quindi essere $a_p = 0$. La combinazione si riduce quindi a $a_1 \mathbf{u}_1 + \cdots + a_{p-1} \mathbf{u}_{p-1} = \mathbf{0}$, e poiché $\mathcal{C}' = (\mathbf{u}_1, \dots, \mathbf{u}_{p-1})$ è una stringa di lunghezza $p - 1$, per ipotesi induttiva i vettori $\mathbf{u}_1, \dots, \mathbf{u}_{p-1}$ sono indipendenti e deve essere $a_1 = \cdots = a_{p-1} = 0$. \square

11 ESERCIZIO. Si provi che \mathcal{C} è una λ -stringa (di lunghezza p) se e solo se \mathcal{C} è indipendente, il sottospazio $U = \mathcal{L}(\mathcal{C})$ è invariante, e $M_{\mathcal{C}}^{\mathcal{C}}(f|_U) = J_{\lambda, p}$.

Il caso nilpotente

Sia $f : V \rightarrow V$ un endomorfismo e poniamo $f^1 = f$ e, induttivamente, $f^{n+1} = f \circ f^n$ per ogni $n \geq 1$. Ricordiamo che un endomorfismo f si dice *nilpotente* se esiste $q \in \mathbb{N}$ tale che $f^q = \mathbf{0}$. Abbiamo qui indicato con $\mathbf{0}$ l'endomorfismo nullo. Il minimo di tali esponenti si dice *indice di nilpotenza* e si indica con $\iota(f)$. In modo del tutto analogo, diremo che una matrice quadrata A è nilpotente se una sua potenza è nulla, ed in tal caso il minimo di tali esponenti, $\iota(A)$, prende il nome di indice di nilpotenza di A . È chiaro che tale indice è necessariamente maggiore di 1 quando l'endomorfismo f non è nullo, o quando la matrice A non è nulla.

12 ESERCIZIO. Si provi che un endomorfismo nilpotente (o anche una matrice nilpotente) ammette $\mathbf{0}$ come unico suo autovalore.

Osserviamo esplicitamente che se \mathcal{C} è una base, mod W , di U e $\mathcal{C}' = [\mathbf{w}_1, \dots, \mathbf{w}_t]$ è una base di W , allora $\mathcal{B}_U = \mathcal{C} \cup \mathcal{C}' = [\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{w}_1, \dots, \mathbf{w}_t]$ è una base di U .

19 LEMMA. Sia $f : V \rightarrow V$ un endomorfismo nilpotente e sia $\iota(f) = p \geq 2$. Se i vettori $\mathbf{u}_1, \dots, \mathbf{u}_k \in \ker f^i$ sono indipendenti, mod $\ker f^{i-1}$, allora $f(\mathbf{u}_1), \dots, f(\mathbf{u}_k)$ sono indipendenti, mod $\ker f^{i-2}$.

Dimostrazione. È chiaro che se $\mathbf{u} \in \ker f^i$ allora $f(\mathbf{u}) \in \ker f^{i-1}$, in quanto

$$\mathbf{0} = f^i(\mathbf{u}) = f^{i-1}(f(\mathbf{u})) .$$

Proviamo dunque che i vettori $f(\mathbf{u}_1), \dots, f(\mathbf{u}_k)$ sono indipendenti, mod $\ker f^{i-2}$. Siano a_1, \dots, a_k scalari tali che $a_1 f(\mathbf{u}_1) + \dots + a_k f(\mathbf{u}_k) = f(a_1 \mathbf{u}_1 + \dots + a_k \mathbf{u}_k) \in \ker f^{i-2}$. Abbiamo quindi che

$$f^{i-1}(a_1 \mathbf{u}_1 + \dots + a_k \mathbf{u}_k) = f^{i-2} f(a_1 \mathbf{u}_1 + \dots + a_k \mathbf{u}_k) = \mathbf{0} ,$$

e cioè $a_1 \mathbf{u}_1 + \dots + a_k \mathbf{u}_k \in \ker f^{i-1}$. Poiché $\mathbf{u}_1, \dots, \mathbf{u}_k$ sono indipendenti, mod $\ker f^{i-1}$, avremo che $a_1 = \dots = a_k = 0$. \square

20 TEOREMA DI JORDAN PER GLI OPERATORI NILPOTENTI. Sia $\mathbb{K} = \mathbb{C}$ ed $f : V \rightarrow V$ un endomorfismo nilpotente, e sia $\iota(f) = p$. Esiste una base \mathcal{B} di V che è unione di stringhe (detta base a stringhe). La matrice $J = M_{\mathcal{B}}^{\mathcal{B}}(f)$ è una matrice diagonale a blocchi di Jordan nilpotenti. Tale matrice, detta forma canonica di Jordan per f , è unica, a meno dell'ordine in cui compaiono i blocchi sulla diagonale. La molteplicità geometrica di 0 è pari al numero di blocchi di Jordan che compaiono sulla diagonale di J .

(Per inciso, ricordiamo che nell'ipotesi del teorema appena enunciato la molteplicità algebrica di 0 è necessariamente n).

Dimostrazione. Poniamo $n_i = \dim \ker f^i$, $q_i = n_i - n_{i-1}$, per ogni $i \in \{1, \dots, p\}$ (e, convenzionalmente, poniamo $n_0 = 0$). Sia $\mathcal{B}^p = [\mathbf{u}_1^p, \dots, \mathbf{u}_{q_p}^p]$ una base di $V = \ker f^p$, mod $\ker f^{p-1}$. Poniamo $\mathbf{u}_h^{p-1} = f(\mathbf{u}_h^p)$ ($1 \leq h \leq q_p$). I vettori così definiti appartengono, come osservato, a $\ker f^{p-1}$, e $[\mathbf{u}_1^{p-1}, \dots, \mathbf{u}_{q_p}^{p-1}]$ è un sistema indipendente, mod $\ker f^{p-2}$. È possibile quindi determinare ulteriori vettori $\mathbf{u}_{q_p+1}^{p-1}, \dots, \mathbf{u}_{q_{p-1}}^{p-1}$ (qualora fosse $q_p < q_{p-1}$) ed ottenere una base $\mathcal{B}^{p-1} = [\mathbf{u}_1^{p-1}, \dots, \mathbf{u}_{q_{p-1}}^{p-1}]$ di $\ker f^{p-1}$, mod $\ker f^{p-2}$. Si procede in modo analogo, nel senso che, avendo trovato una base $\mathcal{B}^{i+1} = [\mathbf{u}_1^{i+1}, \dots, \mathbf{u}_{q_{i+1}}^{i+1}]$ di $\ker f^{i+1}$, mod $\ker f^i$, si costruisce una base $\mathcal{B}^i = [\mathbf{u}_1^i, \dots, \mathbf{u}_{q_i}^i]$ di $\ker f^i$, mod $\ker f^{i-1}$. Nell'ultimo passaggio di questo procedimento, si determina una base $\mathcal{B}^1 = [\mathbf{u}_1^1, \dots, \mathbf{u}_{q_1}^1]$ di $\ker f$ (i cui

elementi sono quindi autovettori associati all'unico autovalore 0). È chiaro che $\mathcal{B}^1 \cup \mathcal{B}^2$ è una base di $\ker f^2$, $\mathcal{B}^1 \cup \mathcal{B}^2 \cup \mathcal{B}^3$ è una base di $\ker f^3$, e così via. Infine $\mathcal{B} = \mathcal{B}^1 \cup \dots \cup \mathcal{B}^p$ è una base di V . I vettori di \mathcal{B}^p sono leader di stringhe di lunghezza p (necessariamente associate all'autovalore 0) contenute in \mathcal{B} . Infatti, per costruzione,

$$\mathbf{u}_h^p \xrightarrow{f} \mathbf{u}_h^{p-1} \xrightarrow{f} \mathbf{u}_h^{p-2} \xrightarrow{f} \dots \xrightarrow{f} \mathbf{u}_h^1 \xrightarrow{f} \mathbf{0} \quad (1 \leq h \leq q_p).$$

Gli eventuali vettori $\mathbf{u}_{q_p+1}^{p-1}, \dots, \mathbf{u}_{q_{p-1}}^{p-1}$ di $\ker f^{p-1}$ (presenti quando $q_p < q_{p-1}$), sono leader di stringhe di lunghezza $p-1$ (necessariamente associate all'autovalore 0) contenute in \mathcal{B} , e così via. Quindi \mathcal{B} è una base di stringhe. Se si sceglie un opportuno ordine dei vettori di \mathcal{B} , la matrice $J = M_{\mathcal{B}}^{\mathcal{B}}(f)$ è del tipo richiesto. Più precisamente, si dovranno ordinare i vettori di \mathcal{B} stringa per stringa, in ordine inverso a quello trovato. Ad esempio, abbiamo

$$\mathbf{u}_1^{p-1} = f(\mathbf{u}_1^p), \quad \mathbf{u}_1^{p-2} = f(\mathbf{u}_1^{p-1}), \quad \dots, \quad \mathbf{u}_1^1 = f(\mathbf{u}_1^2)$$

e in \mathcal{B} consideriamo tali vettori in ordine inverso: $\mathcal{B} = (\mathbf{u}_1^1, \mathbf{u}_1^2, \dots, \mathbf{u}_1^{p-1}, \mathbf{u}_1^p, \dots)$. Avremo

$$J = \begin{pmatrix} J_{0,p} & & \\ & \ddots & \\ & & \dots \end{pmatrix}.$$

Procedendo in modo analogo nell'ordinare i vettori di \mathcal{B} , stringa per stringa, otterremo che la matrice associata ad f rispetto a \mathcal{B} sarà diagonale a blocchi, il blocco $J_{0,p}$ comparirà q_p volte, il blocco $J_{0,p-1}$ comparirà q_{p-1} volte, e così via. Si osserva esplicitamente che possono comparire blocchi del tipo $J_{0,1}$. È chiaro che se si cambia l'ordine delle stringhe in \mathcal{B} , può cambiare l'ordine dei blocchi di Jordan lungo la diagonale di J . Si può verificare anche l'unicità di tale matrice, a meno di permutazione dell'ordine in cui compaiono i blocchi, ma ce ne asterremo. \square

Il caso generale

Premettiamo il seguente lemma, di carattere generale.

21 LEMMA. *Siano $g, h : V \rightarrow V$ due endomorfismi che commutano, ovvero tali che $h \circ g = g \circ h : V \rightarrow V$. Allora*

- (i) $\ker g$ e $\operatorname{im} g$ sono h -invarianti.
- (ii) $\ker h$ e $\operatorname{im} h$ sono g -invarianti.

Dimostrazione. Proviamo la (i) (la (ii) è analoga). Sia $\mathbf{v} \in \ker g$. Allora $g(h(\mathbf{v})) = h(g(\mathbf{v})) = \mathbf{0}$, e quindi $h(\mathbf{v}) \in \ker g$. Sia ora $\mathbf{v} \in \operatorname{im} g$. Esiste $\mathbf{u} \in V$ tale che $g(\mathbf{u}) = \mathbf{v}$. Quindi $h(\mathbf{v}) = h(g(\mathbf{u})) = g(h(\mathbf{u}))$, cioè $h(\mathbf{v}) \in \operatorname{im} g$. \square

Sia ora $f : V \rightarrow V$ un endomorfismo, λ un suo autovalore e V_λ il corrispondente autospazio. Abbiamo che $\mathbf{u} \in V_\lambda \iff f(\mathbf{u}) = \lambda\mathbf{u}$, ovvero $(f - \lambda \text{id}_V)(\mathbf{u}) = \mathbf{0}$. In altre parole $V_\lambda = \ker(f - \lambda \text{id}_V)$. Abbiamo quindi una catena

$$\{\mathbf{0}\} < \ker(f - \lambda \text{id}_V) \leq \ker(f - \lambda \text{id}_V)^2 \leq \dots \leq V$$

e, come già osservato, vale la disuguaglianza stretta fino ad un certo esponente, diciamo n_λ , ovvero

$$\{\mathbf{0}\} < V_\lambda = \ker(f - \lambda \text{id}_V) < \ker(f - \lambda \text{id}_V)^2 < \dots < \ker(f - \lambda \text{id}_V)^{n_\lambda} \leq V .$$

22 DEFINIZIONE. Il sottospazio $\tilde{V}_\lambda := \ker(f - \lambda \text{id}_V)^{n_\lambda}$ si dice autospazio generalizzato di f , associato all'autovalore λ .

Analogamente, abbiamo che

$$\text{im}(f - \lambda \text{id}_V) \supset \text{im}(f - \lambda \text{id}_V)^2 \supset \dots \supset \text{im}(f - \lambda \text{id}_V)^{n_\lambda} .$$

Poniamo, per semplicità, $\hat{V}_\lambda := \text{im}(f - \lambda \text{id}_V)^{n_\lambda}$. Osserviamo che $f \circ (f - \lambda \text{id}_V) = (f - \lambda \text{id}_V) \circ f$, e quindi anche $f \circ (f - \lambda \text{id}_V)^k = (f - \lambda \text{id}_V)^k \circ f$, per ogni $k \in \mathbb{N}$.

23 LEMMA. I sottospazi \tilde{V}_λ e \hat{V}_λ sono invarianti e complementari (ovvero la loro somma è diretta e coincide con V).

Dimostrazione. Poiché f e $(f - \lambda \text{id}_V)^{n_\lambda}$ commutano, in base al lemma precedente \tilde{V}_λ e \hat{V}_λ sono invarianti. Per provare l'ultima parte dell'enunciato, osserviamo che, in base al Teorema dell'Equazione Dimensionale, si ha che $n = \dim V = \dim \tilde{V}_\lambda + \dim \hat{V}_\lambda$. Proveremo che $\tilde{V}_\lambda \cap \hat{V}_\lambda = \{\mathbf{0}\}$ e quindi il sottospazio congiungente $\tilde{V}_\lambda + \hat{V}_\lambda$ è una somma diretta, ed ha dimensione n , ovvero $V = \tilde{V}_\lambda \oplus \hat{V}_\lambda$. Sia dunque $\mathbf{v} \in \tilde{V}_\lambda \cap \hat{V}_\lambda$. Poiché $\mathbf{v} \in \hat{V}_\lambda$, esiste un vettore \mathbf{u} tale che $\mathbf{v} = (f - \lambda \text{id}_V)^{n_\lambda}(\mathbf{u})$. D'altra parte, poiché $\mathbf{v} \in \tilde{V}_\lambda := \ker(f - \lambda \text{id}_V)^{n_\lambda}$, si ha che

$$\mathbf{0} = (f - \lambda \text{id}_V)^{n_\lambda}(\mathbf{v}) = (f - \lambda \text{id}_V)^{n_\lambda}(f - \lambda \text{id}_V)^{n_\lambda}(\mathbf{u}) = (f - \lambda \text{id}_V)^{2n_\lambda}(\mathbf{u}) .$$

Quindi $\mathbf{u} \in \ker(f - \lambda \text{id}_V)^{2n_\lambda} = \ker(f - \lambda \text{id}_V)^{n_\lambda}$ (in quanto $2n_\lambda > n_\lambda$). Pertanto $\mathbf{v} = (f - \lambda \text{id}_V)^{n_\lambda}(\mathbf{u}) = \mathbf{0}$, come richiesto. \square

24 TEOREMA DI JORDAN. Sia $f : V \rightarrow V$ un endomorfismo, dove V è uno spazio vettoriale di dimensione n sul campo \mathbb{K} , e supponiamo che il suo polinomio caratteristico $p(x)$ sia

completamente riducibile. Esiste allora una base \mathcal{B} a stringhe, tale che $J = M_{\mathcal{B}}^{\mathcal{B}}(f)$ sia una matrice diagonale a blocchi di Jordan. La matrice J , detta forma canonica di Jordan di f , è unica, a meno di permutazioni dei suoi blocchi di Jordan. Gli autovalori di f coincidono con gli autovalori di tali blocchi. Per ognuno di tali autovalori, λ , avremo che $m_g(\lambda) = \dim V_{\lambda}$ è il numero di blocchi di Jordan di J relativi a λ , mentre $m_a(\lambda)$ è la somma degli ordini di tali blocchi, e si ha $m_a(\lambda) = \dim \tilde{V}_{\lambda}$.

Dimostrazione. Consideriamo la restrizione di f al sottospazio invariante \widehat{V}_{λ} . Per tale restrizione λ non è un autovalore, in quanto $V_{\lambda} \subseteq \widehat{V}_{\lambda}$ e $\tilde{V}_{\lambda} \cap \widehat{V}_{\lambda} = \{\mathbf{0}\}$. In \tilde{V}_{λ} non compaiono autovettori di f associati ad autovalori diversi da λ . Infatti, se $\mu \neq \lambda$ è un altro autovalore di f , e \mathbf{u} è un autovettore ad esso associato, si ha

$$(f - \lambda \text{id}_V)(\mathbf{u}) = f(\mathbf{u}) - \lambda \mathbf{u} = \mu \mathbf{u} - \lambda \mathbf{u} = (\mu - \lambda) \mathbf{u} \neq \mathbf{0} .$$

Pertanto

$$(f - \lambda \text{id}_V)^{n_{\lambda}}(\mathbf{u}) = (\mu - \lambda)^{n_{\lambda}} \mathbf{u} \neq \mathbf{0} ,$$

e quindi $\mathbf{u} \notin \tilde{V}_{\lambda} = \ker(f - \lambda \text{id}_V)^{n_{\lambda}}$. Dunque la restrizione $f_{\tilde{V}_{\lambda}}$ di f a \tilde{V}_{λ} non ha autovalori diversi da λ e, posto $\dim \tilde{V}_{\lambda} = d_{\lambda}$, il polinomio caratteristico di $f_{\tilde{V}_{\lambda}}$ è $(\lambda - x)^{d_{\lambda}}$. Osserviamo esplicitamente che $p_f(x)$ è prodotto di $(\lambda - x)^{d_{\lambda}}$ e del polinomio caratteristico della restrizione $f_{\widehat{V}_{\lambda}}$ di f , e quest'ultimo non ammette λ come radice. Avremo quindi che $d_{\lambda} = m_a(\lambda)$. L'endomorfismo $f_{\tilde{V}_{\lambda}} - \lambda \text{id}_{\tilde{V}_{\lambda}} = (f - \lambda \text{id}_V)_{\tilde{V}_{\lambda}}$ è nilpotente, con indice di nilpotenza al più n_{λ} , e quindi \tilde{V}_{λ} ha una base a stringhe rispetto a $f_{\tilde{V}_{\lambda}} - \lambda \text{id}_{\tilde{V}_{\lambda}}$ e al suo unico autovalore 0. Ma se, ad esempio, $\mathcal{C} = (\mathbf{u}_1, \dots, \mathbf{u}_p)$ è una tale stringa, si ha che

$$(f - \lambda \text{id}_V)(\mathbf{u}_p) = \mathbf{0} \mathbf{u}_p + \mathbf{u}_{p-1} = \mathbf{u}_{p-1}; \dots; (f - \lambda \text{id}_V)(\mathbf{u}_2) = \mathbf{u}_1; (f - \lambda \text{id}_V)(\mathbf{u}_1) = \mathbf{0}$$

ovvero

$$f(\mathbf{u}_p) = \lambda \mathbf{u}_p + \mathbf{u}_{p-1}; \dots; f(\mathbf{u}_2) = \lambda \mathbf{u}_2 + \mathbf{u}_1; f(\mathbf{u}_1) = \mathbf{0}$$

e quindi \mathcal{C} è una stringa per f associata all'autovalore λ . In altre parole, \tilde{V}_{λ} ammette una base di stringhe. Il ragionamento portato avanti finora in relazione ad un singolo autovalore λ sarà utilizzato per ogni autovalore. Supponiamo che $\lambda_1, \dots, \lambda_t$ siano gli autovalori di f . Per quanto detto fin qui, abbiamo che

$$V = \tilde{V}_{\lambda_1} \oplus W$$

$W = \text{im}(f - \lambda_1 \text{id}_V)^{n_{\lambda_1}}$ è invariante per f e $f|_W$ non ammette λ_1 come autovalore. Il polinomio caratteristico $p_f(x) = (\lambda_1 - x)^{m_a(\lambda_1)} \cdot (\lambda_2 - x)^{m_a(\lambda_2)} \cdot \dots \cdot (\lambda_t - x)^{m_a(\lambda_t)}$ si decompone in $p_f(x) = (\lambda_1 - x)^{m_a(\lambda_1)} \cdot p_{f_W}(x)$. Quindi $p_{f_W}(x)$ non ammette λ_1 come radice e si ha $p_{f_W}(x) = (\lambda_2 - x)^{m_a(\lambda_2)} \cdot \dots \cdot (\lambda_t - x)^{m_a(\lambda_t)}$. Ragioniamo come in precedenza

ed otteniamo che $W = \widetilde{W}_{\lambda_2} \oplus Z$, dove $\widetilde{W}_{\lambda_2}$ è l'autospazio generalizzato di f_W relativo all'autovalore λ_2 , Z è un sottospazio invariante e la restrizione di f_W e quindi anche di f , a Z non ammette λ_1, λ_2 come autovalori. È chiaro che $\widetilde{W}_{\lambda_2} \leq \widetilde{V}_{\lambda_2}$ e $\dim \widetilde{W}_{\lambda_2} = \dim \widetilde{V}_{\lambda_2} = m_a(\lambda_2)$, quindi $\widetilde{W}_{\lambda_2} = \widetilde{V}_{\lambda_2}$ e $V = \widetilde{V}_{\lambda_1} \oplus \widetilde{V}_{\lambda_2} \oplus Z$. Dopo $t - 1$ passaggi si ottiene la decomposizione

$$V = \widetilde{V}_{\lambda_1} \oplus \widetilde{V}_{\lambda_2} \oplus \cdots \oplus \widetilde{V}_{\lambda_t},$$

e poiché ogni autospazio generalizzato $\widetilde{V}_{\lambda_j}$ ammette una base a stringhe \mathcal{B}_j , ponendo $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_t$ si ottiene una base a stringhe. È poi chiaro che la matrice $J = M_{\mathcal{B}}^{\mathcal{B}}(f)$ è del tipo richiesto. Ci asterremo, come nel caso nilpotente, dal provare l'unicità di J . \square

25 ESEMPIO. Sia V uno spazio vettoriale reale e sia $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ una sua base ordinata. Sia $f : V \rightarrow V$ l'endomorfismo definito ponendo

$$f(\mathbf{e}_1) = \mathbf{e}_1 + 3\mathbf{e}_2; \quad f(\mathbf{e}_2) = -\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3; \quad f(\mathbf{e}_3) = -\mathbf{e}_1 + 3\mathbf{e}_2 + 2\mathbf{e}_3$$

ed estendendo poi per linearità. Si ha che $A = M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} 1 & -1 & -1 \\ 3 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix}$ e $p(x) = 8x - 1 \cdot (x - 2)^2$. Gli autovalori di f sono quindi $\lambda = 1$, semplice, e $\lambda = 2$ doppio, con $\dim V_2 = 1$. Pertanto f non è diagonalizzabile, ma il polinomio caratteristico è completamente riducibile ed è possibile determinare una base a stringhe. Un autovettore per $\lambda = 1$ è il vettore $\mathbf{u}_1 = -2\mathbf{e}_1 - 3\mathbf{e}_2 + 3\mathbf{e}_3$, che genera V_1 . Un autovettore per $\lambda = 2$ è il vettore $\mathbf{u}_2 = -\mathbf{e}_1 + \mathbf{e}_3$, che genera V_2 . La dimensione dell'autospazio generalizzato \widetilde{V}_2 è 2, e quindi bisognerà trovare un ulteriore autovettore generalizzato $\mathbf{u}_3 \in \widetilde{V}_2$ che formi una stringa con \mathbf{u}_2 . Studiamo allora $\widetilde{V}_2 = \ker(f - \text{id}_V)^{n_2}$. In questo contesto è chiaro che $n_2 = 2$. Ma ragioniamo in modo più generale. Vediamo a che punto la sequenza

$$\{\mathbf{0}\} < \ker(f - 2\text{id}_V) < \ker(f - 2\text{id}_V)^2 \leq \dots < V$$

si stabilizza. Si ha, in effetti, che

$$\{\mathbf{0}\} < \ker(f - 2\text{id}_V) < \ker(f - 2\text{id}_V)^2 = \ker(f - 2\text{id}_V)^3 = \ker(f - 2\text{id}_V)^4 = \dots < V,$$

e cioè $n_2 = 2$. Infatti, la matrice B associata a $f - 2\text{id}_V$ è $B = \begin{pmatrix} -1 & -1 & -1 \\ 3 & 0 & 3 \\ 0 & 1 & 0 \end{pmatrix}$ e si ha che $\rho(B) = 2$ e $\dim V_2 = \dim \ker(f - 2\text{id}_V) = 3 - 2 = 1$. Inoltre $B^2 = \begin{pmatrix} -2 & 0 & -2 \\ -3 & 0 & -3 \\ 3 & 0 & 3 \end{pmatrix}$, $\rho(B^2) = 1$ e $\dim \ker(f - 2\text{id}_V)^2 = 3 - 1 = 2$, e poi $B^3 = -B^2$, $B^4 = -B^3$, e così via. Quindi $\ker(f - 2\text{id}_V)^2 = \ker(f - 2\text{id}_V)^3 = \ker(f - 2\text{id}_V)^4 = \dots < V$. Si ha che

$$\widetilde{V}_2 = \ker(f - \text{id}_V)^2 : B^2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Tale sistema si riduce ad un'unica equazione:

$$\widetilde{V}_2 : x + z = 0$$

che ammette, come insieme delle soluzioni,

$$\text{Sol} = \{(-a, b, a) \mid a, b \in \mathbb{R}\}.$$

Quindi

$$\widetilde{V}_2 = \{-a\mathbf{e}_1 + b\mathbf{e}_2 + a\mathbf{e}_3 \mid a, b \in \mathbb{R}\}.$$

Una base di \tilde{V}_2 è $\tilde{\mathcal{B}}_2 = \{-\mathbf{e}_1 + \mathbf{e}_3, \mathbf{e}_2\}$. Come si vede, il primo vettore di tale base è un autovettore dell'autovalore $\lambda = 2$; il secondo è invece un suo autovettore generalizzato. Si ha che

$$f(\mathbf{e}_2) = -\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3 = (-\mathbf{e}_1 + \mathbf{e}_3) + 2\mathbf{e}_2 ; f(-\mathbf{e}_1 + \mathbf{e}_3) = 2(-\mathbf{e}_1 + \mathbf{e}_3) ,$$

e quindi $(\mathbf{e}_2, -\mathbf{e}_1 + \mathbf{e}_3)$ è una stringa di lunghezza 2 associata all'autovalore $\lambda = 2$. Pertanto, posto $\mathbf{u}_3 = \mathbf{e}_2$, la base a stringhe cercata è

$$\mathcal{C} = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$$

e si ha che

$$J = M_{\mathcal{C}}^{\mathcal{C}}(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} J_{1,1} & & \\ & J_{2,2} & \\ & & \end{pmatrix} .$$