Forma Canonica di Jordan
Complementi di algebra lineare per Geometria IV

Rocco Chirivì

`versione 09.02.16`

Queste note sono basate sulle lezioni tenute per il corso di Geometria IV, per il secondo anno di Matematica presso il Dipartimento di Matematica e Fisica "Ennio De Giorgi" dell'Università del Salento. Si tratta di appunti per gli studenti e, come tali, non hanno alcuna pretesa di completezza; il lettore dovrebbe confrontare quanto segue con un qualsiasi testo di algebra lineare.

Invito chiunque trovi degli errori a voler essere così gentile da segnalarmeli mandando un e–mail a `rocco.chirivi@unisalento.it`.

1 Introduzione alla forma canonica di Jordan

Definizione 1. Sia V uno spazio vettoriale sul campo $\mathbb{K}$ e sia $f : V \longrightarrow V$ un endomorfismo. Un sottospazio $U \subseteq V$ si dice *f–invariante* se $f(U) \subseteq U$.

Lo spazio vettoriale V si dice *f–irriducibile* se $V \neq 0$ e 0 e V sono gli unici sottospazi f–invarianti. In altre parole, V è irriducibile se ha esattamente due sottospazi invarianti, quelli banali. Se V non è irriducibile allora si dice *riducibile*.

Lo spazio vettoriale V si dice *f–decomponibile* se esistono sottospazi U_1, U_2 di V che siano f–invarianti non banali, cioè diversi da 0 e V, e tali che $V = U_1 \oplus U_2$. Se V non è decomponibile si dice *indecomponibile*.

Esempio 2. Consideriamo il seguente endomorfismo

$$V = \mathbb{K}^3 \ni \begin{pmatrix} x \\ y \\ z \end{pmatrix} \stackrel{f}{\longmapsto} \begin{pmatrix} -y \\ x \\ z \end{pmatrix} \in V$$

La matrice associata ad f nella base canonica di V è

$$\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Siano $U_1 = \mathcal{L}(e_1, e_2)$ e $U_2 = \mathcal{L}(e_3)$. Abbiamo $f(e_1) = e_2 \in U_1$ e $f(e_2) = -e_1 \in U_1$ e quindi U_1 è f–invariante; inoltre anche U_2 è f–invariante visto che $f(e_3) = e_3$. I sottospazi U_1 e U_2 sono generati da sottoinsiemi complementari della base canonica di V, risulta quindi $V = U_1 \oplus U_2$, cioè V è decomponibile.

Osserviamo ora che U_2 ha dimensione 1 e quindi i suoi unici sottospazi sono 0 e U_2. È chiaro che U_2 è irriducibile, come tutti i sottospazi f–invarianti di dimensione 1.

Ci proponiamo ora di decidere se U_1 è decomponibile o irriducibile. Avendo U_1 dimensione 2, un sottospazio invariante non banale deve avere dimensione 1, in particolare sarà generato da un autovettore di f su U_1. Il polinomio caratteristico di $f_{|U_1}$ è $t^2 + 1$.

Quindi, se $i = \sqrt{-1} \notin \mathbb{K}$ l'endomorfismo $f_{|U_1}$ non ha alcun autovalore e quindi U_1 è irriducibile.

Se invece $i \in \mathbb{K}$, ad esempio se $\mathbb{K} = \mathbb{C}$ o anche $\mathbb{K} = \mathbb{F}_{13}$, allora $\pm i$ sono autovalori per $f_{|U_1}$ e quindi esistono autovettori corrispondenti. In particolare $U_1 = \mathbb{K}u \oplus \mathbb{K}v$ con u autovettore per i e v autovettore per $-i$; ad esempio possiamo prendere $u = e_1 - ie_2$ e $v = e_1 + ie_2$. Quindi U_1 è decomponibile e nella base u, v, e_3 la matrice associata ad f è

$$\begin{pmatrix} i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Esempio 3. Consideriamo l'endomorfismo

$$V = \mathbb{K}^2 \ni \begin{pmatrix} x \\ y \end{pmatrix} \stackrel{f}{\longmapsto} \begin{pmatrix} x+y \\ y \end{pmatrix} \in V$$

Abbiamo $f(e_1) = e_1$, da cui $U = \mathcal{L}(e_1)$ è f–invariante e anche irriducibile avendo dimensione 1; quindi V è riducibile. Osserviamo però che V è indecomponibile. Infatti se esistesse W sottospazio f–invariante di V per cui $V = U \oplus W$ allora W dovrebbe avere dimensione 1. Ma allora W dovrebbe essere generato da un autovettore di f.

Il polinomio caratteristico di f è $(1-t)^2$ e quindi f ha solo 1 come autovalore. Inoltre l'autospazio $V(1)$ di 1, cioè $\mathrm{Ker}(f - \mathrm{Id})$, il nucleo di $f - \mathrm{Id}$, ha dimensione 1. Ma allora $V(1) = U$ visto che e_1 è un autovettore per 1. Questo finisce la dimostrazione che V è indecomponibile.

Come si vede dagli esempi appena considerati, uno spazio vettoriale può essere riducibile senza essere decomponibile rispetto ad un endomorfismo. Invece, è ovvio che irriducibile implichi indecomponibile.

Definizione 4. Sia $\lambda \in \mathbb{K}$ e sia $n \in \mathbb{N}$ un intero positivo, chiamiamo λ–*blocco* $n \times n$ di Jordan la matrice

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \cdots & & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \in \mathbb{K}^{n,n}$$

che ha λ sulla diagonale, 1 sulla sovradiagonale e 0 nei restanti ingressi.

Ricordiamo che data una matrice $A \in \mathbb{K}^{n,n}$ indichiamo con $f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ l'endomorfismo $f_A(X) = A \cdot X$. In particolare sia $f_{n,\lambda} = f_{J_n(\lambda)}$. Risulta

$$f_{n,\lambda} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 + x_2 \\ \lambda x_2 + x_3 \\ \vdots \\ \lambda x_{n-1} + x_n \\ \lambda x_n \end{pmatrix}$$

o, in termini dei vettori della base canonica,

$$\begin{array}{lcl} f_{n,\lambda}(e_1) & = & \lambda e_1 \\ f_{n,\lambda}(e_2) & = & \lambda e_2 + e_1 \\ & \vdots & \\ f_{n,\lambda}(e_{n-1}) & = & \lambda e_{n-1} + e_{n-2} \\ f_{n,\lambda}(e_n) & = & \lambda e_n + e_{n-1} \end{array}$$

Quindi e_1 è un autovettore di autovalore λ per $f_{n,\lambda}$. Vediamo ora che questo è l'unico autovalore. Infatti il polinomio caratteristico di $f_{n,\lambda}$ è $p_{f_{n,\lambda}}(t) = \det(f_{n,\lambda} - t\,\mathrm{Id}) = (\lambda - t)^n$ in quanto la matrice $J_n(\lambda) - t\,\mathrm{Id}$ è triangolare superiore. Inoltre la molteplicità algebrica $\mathrm{ma}(\lambda)$ di λ è n.

Invece per la molteplicità geometrica di λ abbiamo

$$\mathrm{mg}(\lambda) = \dim \mathrm{Ker}(f_{n,\lambda} - \lambda\,\mathrm{Id}) = n - rkJ_n(0) = n - rk \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \cdots & & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} = n - (n-1) = 1$$

In particolare l'autospazio di λ è $\mathcal{L}(e_1)$. Riepiloghiamo quanto appena visto nella seguente osservazione.

Osservazione 5. L'endomorfismo $f_{n,\lambda}$ di $\mathbb{K}^n$ associato al blocco di Jordan $J_n(\lambda)$ ha solo λ come autovalore. Inoltre λ ha molteplicità algebrica n, geometrica 1 e autospazio $\mathcal{L}(e_1)$.

Vogliamo ora studiare quali sono i sottospazi di $V = \mathbb{K}^n$ invarianti per $f_{n,\lambda}$. Dalle formule che abbiamo scritto sopra si vede subito che la seguente *bandiera*, cioè successione di sottospazi, è formata da sottospazi invarianti

$$V_0 = 0 \subsetneq V_1 = \mathcal{L}(e_1) \subsetneq V_2 = \mathcal{L}(e_1, e_2) \subsetneq \cdots \subsetneq V_k = \mathcal{L}(e_1, e_2, \ldots, e_k) \subsetneq \cdots \subsetneq V_n = V$$

e si ha $\dim V_k = k$ per $k = 0, 1, \ldots, n$. In realtà non ci sono altri spazi invarianti, e questo è il contentuto della seguente osservazione.

Osservazione 6. Se $U \subseteq V$ è $f_{n,\lambda}$–invariante allora $U = V_k$ per qualche $0 \leq k \leq n$.

Dimostrazione. Procediamo per induzione su n. Se $n = 1$ allora $V = \mathbb{K}$ e non ci sono altri sottospazi oltre a $V_0 = 0$ e $V_1 = \mathbb{K}$, quindi non ci sono neanche altri spazi $f_{n,\lambda}$–invarianti, cioè $U = V_0$ o $U = V_1$.

Sia ora $n > 1$. Se $U \subseteq V_{n-1}$ allora, visto che $f_{|V_{n-1}} = f_{n-1,\lambda}$, per induzione abbiamo $U = V_k$ per qualche $0 \leq k \leq n-1$. Supponiamo allora che $U \not\subseteq V_{n-1}$.

Esiste quindi $v_n = a_1 e_1 + a_2 e_2 + \cdots + a_{n-1} e_{n-1} + a_n e_n \in U$ con $a_n \neq 0$. Dividendo per a_n possiamo supporre $a_n = 1$. Osserviamo che U è $(f_{n,\lambda} - \lambda\,\mathrm{Id})$–invariante; infatti se $u \in U$ allora $f_{n,\lambda}(u) \in U$ e quindi anche $f_{n,\lambda}(u) - \lambda u \in U$. Inoltre $f_{n,\lambda} - \lambda\,\mathrm{Id} = f_{n,0}$.

Allora i seguenti vettori sono tutti elementi di U

$$\begin{array}{rcccl} v_{n-1} & = & f_{n,0}(v_n) & = & a_2 e_1 + a_3 e_2 + \cdots + a_{n-2} e_{n-3} + a_{n-1} e_{n-2} + e_{n-1} \\ v_{n-2} & = & f_{n,0}(v_{n-1}) & = & a_3 e_1 + a_4 e_2 + \cdots + a_{n-1} e_{n-3} + e_{n-2} \\ & & \vdots & & \\ v_2 & = & f_{n,0}(v_3) & = & a_{n-1} e_1 + e_2 \\ v_1 & = & f_{n,0}(v_2) & = & e_1 \end{array}$$

In particolare la matrice che ha per colonne le coordinate di $v_1, v_2, \ldots, v_{n-1}, v_n$ rispetto alla base canonica è triangolare superiore con tutti 1 sulla diagonale. Tale matrice ha determinante non nullo e quindi gli n vettori $v_1, v_2, \ldots, v_n$ di U sono linearmente indipendenti. In particolare U ha dimensione almeno n, ma essendo un sottospazio dello spazio vettoriale V di dimensione n, esso risulta uguale a $V = V_n$. $\square$

Corollario 7. *Lo spazio V è $f_{n,\lambda}$–indecomponibile ma riducibile non appena $n > 1$.*

Dimostrazione. Supponiamo che $V = U_1 \oplus U_2$ con U_1 e U_2 $f_{n,\lambda}$–invarianti. Dall'osservazione precedente abbiamo $U_1 = V_k$ e $U_2 = V_h$ per qualche $0 \leq h, k \leq n$ e possiamo supporre, senza perdita di generalità, che $h \leq k$. Ma allora $U_1 \subseteq U_2$ e quindi è impossibile che la somma sia diretta con U_1 e U_2 non banali. Questo prova che V è indecomponibile.

Infine se $n \geq 2$ abbiamo in V il sottospazio $f_{n,\lambda}$–invariante $0 \subsetneq V_1 \subsetneq V$ non banale; quindi V non è irriducibile. □

Vediamo ora un lemma sulla relazione tra il polinomio caratteristico di un endomorfismo e i sottospazi invarianti.

Lemma 8. *Sia $f : V \longrightarrow V$ un endomorfismo e sia $U \subseteq V$ un sottospazio f–invariante; allora $p_{f_{|U}}(t)$ divide $p_f(t)$. Inoltre se $V = U_1 \oplus U_2$ è una decomposizione in sottospazi invarianti allora $p_f(t) = p_{f_{|U_1}}(t) \cdot p_{f_{|U_2}}(t)$.*

Dimostrazione. Sia $\mathcal{B}$ una base di U e sia $\mathcal{C}$ un suo completamento ad una base di V. Allora nella base $\mathcal{C}$ la matrice associata ad f ha la seguente forma a blocchi

$$M_{\mathcal{C}}(f) = \begin{pmatrix} M_{\mathcal{B}}(f_{|U}) & A \\ 0 & B \end{pmatrix}$$

per opportune sottomatrici A e B. Da cui $p_f(t) = \det(M_{\mathcal{C}}(f) - t\,\mathrm{Id}) = \det(M_{\mathcal{B}}(f_{|U}) - t\,\mathrm{Id})\det(B - t\,\mathrm{Id}) = p_{f_{|U}}(t)\det(B - t\,\mathrm{Id})$ e il primo asserto segue. Allo stesso modo, se prendiamo come base per V l'unione di una base $\mathcal{B}_1$ di U_1 e di una base $\mathcal{B}_2$ di U_2, la matrice associata ad f è

$$M_{\mathcal{C}}(f) = \begin{pmatrix} M_{\mathcal{B}_1}(f_{|U_1}) & 0 \\ 0 & M_{\mathcal{B}_2}(f_{|U_2}) \end{pmatrix}$$

e quindi si ha tesi visto che $p_f(t) = \det(M_{\mathcal{C}}(f) - t\,\mathrm{Id}) = \det(M_{\mathcal{B}_1}(f_{|U_1}) - t\,\mathrm{Id})\det(M_{\mathcal{B}_2}(f_{|U_2}) - t\,\mathrm{Id}) = p_{f_{|U_1}}(t) \cdot p_{f_{|U_2}}(t)$. □

2 La forma canonica di Jordan

Se $A \in \mathbb{K}^{n,n}$ e $B \in \mathbb{K}^{m,m}$ sono due matrici allora con $A \oplus B$ indichiamo la matrice a blocchi

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \in \mathbb{K}^{n+m,n+m}$$

Definizione 9. Si chiama *matrice di Jordan* una matrice a blocchi $J = J_{n_1}(\lambda_1) \oplus J_{n_2}(\lambda_2) \oplus \cdots \oplus J_{n_r}(\lambda_r)$, dove l'intero positivo r è il numero di blocchi e gli interi positivi $n_1, n_2, \ldots, n_r$ sono le dimensioni dei blocchi di Jordan relativi a $\lambda_1, \lambda_2, \ldots, \lambda_r \in \mathbb{K}$.

Ad esempio, la matrice

$$\begin{pmatrix} \lambda & 1 & & & & \\ & \lambda & & & & \\ & & \mu & 1 & & \\ & & & \mu & 1 & \\ & & & & \mu & \\ & & & & & \mu \end{pmatrix}$$

è la matrice di Jordan $J_2(\lambda) \oplus J_3(\mu) \oplus J_1(\mu)$ formata da un λ–blocco 2×2, e due μ–blocchi di dimensioni 3×3 e 1×1 rispettivamente.

Sia ora $f : V \longrightarrow V$ un endomorfismo e sia $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ una base di V rispetto a cui la matrice associata ad f è $A = J_{n_1}(\lambda_1) \oplus J_{n_2}(\lambda_2) \oplus \cdots \oplus J_{n_r}(\lambda_r)$, cioè in forma di Jordan.

Alla decomposizione di A come somma di blocchi di Jordan corrisponde una decomposizione $V_1 \oplus V_2 \oplus \cdots \oplus V_r$ di V, dove V_1 è generato dai primi n_1 vettori della base, V_2 dai seguenti n_2 vettori della base e così via, fino a V_r che è generato dagli ultimi n_r vettori della base.

Se l'addendo diretto V_h è generato dal sottoinsieme $v_i, v_{i+1}, \ldots, v_{j-1}, v_j$ della base $\mathcal{B}$ allora chiamiamo $(v_i, v_{i+1}, \ldots, v_{j-1}, v_j)$ la *catena* di Jordan relativa al blocco h–esimo $J_{n_h}(\lambda_h)$ e, inoltre, v_j è detto *vettore iniziale* della catena e v_i è il *vettore finale* della catena. Si noti che si avrà $j - i + 1 = n_h$.

Ancora alcune notazioni. Per un campo $\mathbb{K}$, scriviamo $\overline{\mathbb{K}}$ per la chiusura algebrica di $\mathbb{K}$ e, dato un endomorfismo f, indichiamo con $\Sigma(f) \subseteq \overline{\mathbb{K}}$ lo *spettro* di f, cioè l'insieme delle radici del polinomio caratteristico di f, o, in altri termini, gli autovalori di f in $\overline{\mathbb{K}}$.

Teorema 10 (Forma canonica di Jordan)**.** *Sia V uno spazio vettoriale di dimensione finita su un campo $\mathbb{K}$ e sia $f : V \longrightarrow V$ un suo endomorfismo. Se lo spettro di f è contenuto in $\mathbb{K}$, allora esiste una base di V rispetto a cui la matrice associata ad f è in forma di Jordan.*

Dimostrazione. Per induzione su $n = \dim V$. Se $n = 1$ allora in qualsiasi base $\mathcal{B} = (v_1)$ si ha $M_{\mathcal{B}}(f) = (\lambda) = J_1(\lambda)$ e la tesi è vera.

Sia ora n qualsiasi e sia λ un autovalore di f; si noti che un tale λ esiste in quanto il polinomio caratteristico ha grado n e tutte le sue radici sono in $\mathbb{K}$ per ipotesi. Consideriamo il sottospazio vettoriale $U = \operatorname{Im}(f - \lambda \operatorname{Id})$ di V e sia, per brevità, f_λ l'endomorfismo $f - \lambda \operatorname{Id}$.

Essendo λ un autovalore, il nucleo $\operatorname{Ker} f_\lambda$ è non banale; in particolare $\dim U = \dim \operatorname{Im} f_\lambda = n - \dim \operatorname{Ker} f_\lambda < n$. Osserviamo inoltre che U è f–invariante: infatti se $u \in U$ allora esiste $v \in V$ per cui $u = f_\lambda(v)$ e quindi $f(u) = (f \circ f_\lambda)v = (f \circ f - \lambda f)v = f \circ (f - \lambda \operatorname{Id})v = (f_\lambda \circ f)v = f_\lambda(f(v)) \in \operatorname{Im} f_\lambda = U$.

Dal lemma 8 sappiamo che il polinomio caratteristico $p_{f_{|U}}(t)$ divide il polinomio caratteristico di f, in particolare $\Sigma(f_{|U}) \subseteq \Sigma(f) \subseteq \mathbb{K}$. Abbiamo quindi tutte le condizioni per applicare l'ipotesi induttiva all'endomorfismo $f_{|U} : U \longrightarrow U$. Esiste allora una base $\mathcal{B} = (v_1, v_2, \ldots, v_m)$ di U, dove $m = \dim U$, in cui $A = M_{\mathcal{B}}(f_{|U})$ è una matrice di Jordan.

Il nostro obiettivo è ora completare $\mathcal{B}$ ad una base $\mathcal{C}$ di V in modo da avere $M_{\mathcal{C}}(f)$ in forma di Jordan. Per fare questo aumenteremo ogni blocco di Jordan relativo a λ in A e aggiungeremo, eventualmente, nuovi blocchi 1×1 relativi a λ. Invece, tutti i blocchi relativi ad autovalori diversi da λ in A resteranno invariati nella forma di Jordan di f.

A meno di permutare tra loro i blocchi di Jordan in A, cioè di permutare tra loro i vettori della base $\mathcal{B}$, possiamo supporre che i primi $v_1, v_2, \ldots, v_p$ vettori di $\mathcal{B}$ siano l'unione di tutte le λ–catene in A. Sia ora $Z = U \cap \operatorname{Ker} f_\lambda$; tale spazio è l'autospazio associato all'autovalore λ di $f_{|U}$. Sia $s = \dim Z$.

Per l'osservazione 5, ogni blocco di Jordan in A aumenta di 1 la molteplicità geometrica di λ come autovalore di $f_{|U}$. Quindi nella base $\mathcal{B}$ ci sono esattamente s catene relative a λ e, per la discussione subito prima dell'osservazione 5, possiamo prendere i vettori finali $v_{i_1}, v_{i_2}, \ldots, v_{i_s}$ di tali catene come base di Z.

Siano inoltre $w_1, w_2, \ldots, w_t \in V$ tali che $v_{i_1}, v_{i_2}, \ldots, v_{i_s}, w_1, w_2, \ldots, w_t$ è un completamento della base di Z ad una base di $\operatorname{Ker} f_\lambda$, cioè dell'autospazio di λ in V.

Ora indichiamo con $v_{j_1}, v_{j_2}, \ldots, v_{j_s}$ i vettori iniziali delle λ–catene in A. Visto che i vettori $v_{j_1}, v_{j_2}, \ldots, v_{j_s}$ sono in $\mathcal{B} \subseteq U = \operatorname{Im} f_\lambda$, esistono $u_1, u_2, \ldots, u_s \in V$ per cui $f_\lambda(u_h) = v_{j_h}$, per $h = 1, 2, \ldots, s$.

Il nostro obiettivo è ora dimostrare che

$$\mathcal{C} = (v_1, v_2, \ldots, v_m, u_1, u_2, \ldots, u_s, w_1, w_2, \ldots, w_t)$$

è una base di V. Per prima cosa proviamo la lineare indipendenza; sia quindi

$$\sum_{j=1}^{m} a_j v_j + \sum_{h=1}^{s} b_h u_h + \sum_{k=1}^{t} c_k w_k = 0$$

una combinazione lineare nulla. Applicando f_λ ad ambo i lati otteniamo

$$\sum_{j=1}^{p} a_j f_\lambda(v_j) + \sum_{j=p+1}^{m} a_j f_\lambda(v_j) + \sum_{h=1}^{s} b_h f_\lambda(u_h) + \sum_{k=1}^{t} c_k f_\lambda(w_k) = 0 \tag{1}$$

dove, nella prima somma, abbiamo diviso i termini relativi alle λ–catene nella base $\mathcal{B}$ da quelli relativi alle catene di autovalori diversi da λ. Per prima cosa osserviamo che $f_\lambda(w_k) = 0$ per ogni $k = 1, 2, \ldots, t$ visto che $w_1, w_2, \ldots, w_t \in \operatorname{Ker} f_\lambda$ per costruzione; quindi l'ultima sommatoria nella precedente equazione (1) è nulla.

Notiamo anche che la prima somma è combinazione lineare dei vettori *non* iniziali delle λ–catene in A (si confronti con la discussione che precede l'osservazione 5). La terza somma è per costruzione

$$\sum_{h=1}^{s} b_h v_{j_h}$$

ed è quindi una combinazione lineare dei *soli* vettori iniziali delle λ–catene in A.

Infine se μ è un autovalore di $f_{|U}$ diverso da λ allora il sottospazio vettoriale generato dalle μ–catene è f–invariante e, quindi, anche f_λ invariante. In particolare la seconda somma nell'equazione (1) è combinazione lineare delle catene *non* relative a λ.

Da quanto osservato otteniamo che tutti e tre i primi termini dell'equazione sono nulli in quanto combinazioni lineari di sottoinsiemi disgiunti di una base di U. Quindi abbiamo $\sum_{h=1}^{s} b_h v_{j_h} = 0$ da cui $b_1 = b_2 = \cdots = b_s = 0$ essendo quest'ultima una combinazione lineare di vettori di una base di U. Allora otteniamo

$$\sum_{j=1}^{m} a_j v_j = -\sum_{k=1}^{t} c_k w_k$$

ma il vettore a sinistra è in Z mentre quello di destra è una combinazione lineare di vettori linearmente indipendenti da Z visto che $w_1, w_2, \ldots, w_t$ completano una base di Z ad una base di $\operatorname{Ker} f_\lambda$. Ne segue che entrambi i vettori sono nulli, quindi $a_1 = a_2 = \cdots = a_m = 0$ e $c_1 = c_2 = \cdots c_t = 0$ trattandosi di combinazioni lineari di vettori in una base. Questo termina la dimostrazione della lineare indipendenza dei vettori in $\mathcal{C}$.

Ma i vettori in $\mathcal{C}$ sono in numero

$$m + s + t = \dim U + \dim Z + t = \dim \operatorname{Im} f_\lambda + \dim \operatorname{Ker} f_\lambda = \dim V$$

e questo prova che essi sono anche generatori. Concludiamo che $\mathcal{C}$ è una base di V.

Ci resta ora solo da osservare che, a meno di permutare i vettori in $\mathcal{C}$ otteniamo una base di V in cui la matrice associata ad f è di Jordan. Ma questo segue subito dalle nostre costruzioni: le catene di λ sono $v_{i_h}, v_{i_h+1}, \ldots, v_{j_h-1}, v_{j_h}, u_h$ per $h = 1, \ldots, s$ a cui si aggiungono le nuove catene lunghe 1 date da w_1, w_2, ..., w_t; le catene relative ad autovettori diversi da λ restano invariate. □

Corollario 11. *Sia V uno spazio vettoriale di dimensione finita sul campo $\mathbb{K}$. Se $\mathbb{K}$ è algebricamente chiuso allora ogni endomorfismo di V ammette una forma di Jordan.*

3 Unicità della forma di Jordan

In questa sezione vogliamo studiare se e in che senso possiamo considerare unica la forma canonica di Jordan di un endomorfismo.

È chiaro che cambiando base è possibile permutare tra loro i blocchi che formano una matrice di Jordan. Però, come proveremo nel seguito, *a meno dell'ordine* la forma di Jordan di un endomorfismo, quando esiste, è unica. In particolare vedremo che se f ammette una forma di Jordan, allora la struttura in blocchi di Jordan può essere ricavata dalle dimensioni degli autospazi generalizzati definiti come segue.

Sia $f : V \longrightarrow V$ un endomorfismo e, dato $\lambda \in \Sigma(f)$, chiamiamo $V_k(\lambda) = \operatorname{Ker}(f - \lambda \operatorname{Id})^k$ il k–esimo *autospazio generalizzato* di f relativo all'autovalore λ.

Proposizione 12. *Gli autospazi generalizzati relativi ad un autovalore formano una catena, cioè*

$$0 = V_0(\lambda) \subseteq V_1(\lambda) \subseteq \cdots \subseteq V_k(\lambda) \subseteq V_{k+1}(\lambda) \subseteq \cdots$$

e tale catena si stabilizza. Inoltre, se $V_k(\lambda) = V_{k+1}(\lambda)$ allora $V_h(\lambda) = V_k(\lambda)$ per ogni $h \geq k$ o, in altre parole, se la catena è stabile in un passo allora è stabile da quel passo in poi.

Dimostrazione. Se $u \in V_k(\lambda)$ allora $(f - \lambda \operatorname{Id})^k(u) = 0$ per definizione di $V_k(\lambda)$, quindi anche $(f - \lambda \operatorname{Id})^{k+1}(u) = (f - \lambda \operatorname{Id})(f - \lambda \operatorname{Id})^k(u) = (f - \lambda \operatorname{Id})(0) = 0$. Questo prova che gli autospazi generalizzati formano una catena.

È chiaro che tale catena non può crescere infinite volte in quanto lo spazio vettoriale V è di dimensione finita. Essa sarà quindi definitivamente stabile. Proviamo, infine, che se $V_k(\lambda) = V_{k+1}(\lambda)$ allora la catena è stabile dal passo k in poi: procediamo per induzione su $h - k$ per dimostrare che $V_h(\lambda) = V_k(\lambda)$ per ogni $h \geq k + 1$.

Il passo base $h = k + 1$ è la nostra ipotesi. Sia ora $h > k + 1$ qualsiasi e sia $u \in V_{h+1}(\lambda)$ e $v = (f - \lambda \operatorname{Id})(u)$. Allora si ha $(f - \lambda \operatorname{Id})^h(v) = (f - \lambda \operatorname{Id})^{h+1}(u) = 0$ e quindi $v \in V_h(\lambda)$. Ma $V_h(\lambda) = V_{h-1}(\lambda)$ per induzione e quindi $v \in V_{h-1}(\lambda)$, cioè $(f - \lambda \operatorname{Id})^{h-1}(v) = 0$. Abbiamo $(f - \lambda \operatorname{Id})^h(u) = (f - \lambda \operatorname{Id})^{h-1}(v) = 0$, cioè $u \in V_h(\lambda) = V_k(\lambda)$. □

Il seguente lemma sarà fondamentale per i calcoli che seguono.

Lemma 13. *Sia $A = J_n(0)$ un blocco di Jordan $n \times n$ relativo a 0. Allora il rango delle potenze di A è*

$$\operatorname{rk} A^k = \max(n - k, 0) = \begin{cases} n - k & \textit{se } 0 \leq k \leq n \\ 0 & \textit{altrimenti.} \end{cases}$$

Dimostrazione. Sia $f : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ l'endomorfismo f_A associato ad A. Osserviamo che le immagini dei vettori della base canonica $e_1, e_2, \ldots, e_n$ di $\mathbb{K}^n$ secondo f sono

$$e_n \xmapsto{f} e_{n-1} \xmapsto{f} e_{n-2} \xmapsto{f} \cdots \xmapsto{f} e_2 \xmapsto{f} e_1 \xmapsto{f} 0$$

È allora chiaro che le immagini rispetto all'endomorfismo f^k degli stessi vettori della base sono

$$f^k(e_h) = \begin{cases} e_{h-k} & \text{se } h > k \\ 0 & \text{altrimenti.} \end{cases}$$

Ma f^k è l'endomorfismo associato ad A^k e quindi la matrice A^k è formata da tutti zeri tranne degli 1 in posizione (i, j) con $j - i = k$; in altri termini, gli uno sulla sovradiagonale di A si spostano via via in alto a destra quando l'esponente k aumenta. Inoltre abbiamo subito $\operatorname{rk} A^k = \dim \operatorname{Im} f^k = \mathcal{L}(e_1, e_2, \ldots, e_{n-k}) = n - k$ se $k \leq n$, mentre, per $k > n$, $f^k(e_h) = 0$ per ogni h e quindi $\operatorname{rk} A^k = 0$. □

Vogliamo ora calcolare le dimensioni degli autospazi generalizzati supponendo che lo spettro dell'endomorfismo f sia contenuto in $\mathbb{K}$, in particolare f ammette una forma di Jordan J. La matrice associata a $f - \lambda \operatorname{Id}$ è $A = J - \lambda \operatorname{Id}$, una matrice a blocchi con la stessa struttura di J; però un blocco di Jordan $J_h(\mu)$ di J diventa il blocco $J_h(\mu - \lambda)$ in A. Se $\mu \neq \lambda$ allora il blocco $J_h(\mu - \lambda)$ ha determinante diverso da zero.

In particolare per calcolare le dimensioni $n_{k,\lambda} = \dim V_k(\lambda)$ dei λ–autospazi generalizzati, basta considerare *solo* i blocchi di λ visto che tutti i blocchi relativi ad autovalori diversi da λ sono invertibili e quindi non danno alcun contributi al nucleo di $(f - \lambda \operatorname{Id})^k$.

Possiamo quindi supporre, senza perdita di generalità, $J = J_{n_1}(\lambda) \oplus J_{n_2}(\lambda) \oplus \cdots \oplus J_{n_r}(\lambda)$ e quindi $A = J_{n_1}(0) \oplus J_{n_2}(0) \oplus \cdots \oplus J_{n_r}(0)$.

Lemma 14. *Vale* $n_{k,\lambda} = \sum_{h=1}^{r} \min(n_h, k)$.

Dimostrazione. Segue subito dalla struttura a blocchi di A e dal lemma precedente osservando che $n_h - \max(n_h - k, 0) = \min(n_h, k)$. □

Indichiamo ora con $m_{k,\lambda}$ il numero di blocchi di Jordan $k \times k$ relativi a λ in J.

Osservazione 15. Per ogni $k \geq 1$ si ha

$$n_{k,\lambda} = m_{1,\lambda} + 2m_{2,\lambda} + \cdots + (k-1)m_{k-1,\lambda} + km_{k,\lambda} + km_{k+1,\lambda} + km_{k+2,\lambda} + \cdots$$

da cui

$$m_{k,\lambda} = -n_{k-1,\lambda} + 2n_{k,\lambda} - n_{k+1,\lambda}$$

Dimostrazione. Consideriamo un λ–blocco $h \times h$ in A; dal lemma precedente sappiamo che esso contribuisce ad aumentare la dimensione del nucleo di A^k di $\min(h, k)$. Quindi tutti i blocchi $h \times h$ contribuiscono con $\min(h, k)m_h$; cioè con hm_h se $h \leq k$ e con km_h per $h > k$. Questo prova la prima formula. La seconda segue subito dalla prima. □

Corollario 16. *Se un endomorfismo ammette una forma di Jordan allora questa è unica a meno dello scambio delle posizioni dei blocchi.*

Dimostrazione. Grazie all'osservazione precedente il numero dei blocchi di ogni dimensione e di ogni autovalore, cioè i numeri $m_{k,\lambda}$, può essere calcolato dalle dimensioni degli autospazi generalizzati, cioè dai numeri $n_{k,\lambda}$. Questi ultimi sono chiaramente definiti in modo intrinseco da f e non dipendono dalla forma di Jordan scelta. □

Osserviamo, infine, che $n_{1,\lambda} = \operatorname{mg}(\lambda) = m_{1,\lambda} + m_{2,\lambda} + \cdots$ è il numero totale di λ–blocchi in una forma di Jordan.

Corollario 17. *Sia $\mathbb{K}$ un campo algebricamente chiuso. Due matrici A, $B \in \mathbb{K}^{n,n}$ sono simili se e solo se hanno la stessa forma di Jordan (a meno dell'ordine dei blocchi). Equivalentemente, esiste una bigezione*

$$\{\textit{classi di similitudine di } \mathbb{K}^{n,n}\} \quad \longleftrightarrow \quad \{\textit{forme di Jordan su } \mathbb{K}\}$$

Quindi, se assumiamo $\mathbb{K}$ algebricamente chiuso, le forme di Jordan classificano le classi di similitudine di $\mathbb{K}^{n,n}$. Si osservi che ad una matrice possiamo associare altri *invarianti* per similitudine: la traccia, il determinante, lo spettro, il polinomio caratteristico, il polinomio minimo, ...

È chiaro che due matrici per cui anche uno solo di questi invarianti è diverso *non* sono coniugate. Viceversa però, nessuno di questi invarianti permette di distinguere in maniera completa le classi di similitudine. Ad esempio, esistono matrici non simili con lo stesso polinomio caratteristico.

In generale solo la forma di Jordan discrimina tra le classi di coniugio.

4 Polinomio minimo e teorema di Cayley–Hamilton

Fissiamo una matrice $A \in \mathbb{K}^{n,n}$ e consideriamo la mappa

$$\mathbb{K}[t] \ni p(t) \xmapsto{v_A} p(A) \in \mathbb{K}^{n,n}$$

di *valutazione* dei polinomi nella matrice A. Ad esempio, per il polinomio $p(t) = t^2 - 3t + 1$, si ha $v_A(p(t)) = A^2 - 3A + \mathrm{Id}$.

Proposizione 18. *La mappa v_A è un omomorfismo di anelli ed ha nucleo non banale.*

Dimostrazione. L'anello dei polinomi $\mathbb{K}[t]$ è generato come $\mathbb{K}$–algebra da t, cioè ogni polinomio è combinazione lineare di potenze di t, e la mappa v_A è l'unica estensione ad un omomorfismo dell'assegnazione $t \longmapsto A$.

Per provare che esiste un polinomio non banale che si annulla in A, consideriamo le matrici $\mathrm{Id} = A^0, A, A^2, \dots, A^{n^2}$. Si tratta di $n^2 + 1$ matrici nello spazio vettoriale $\mathbb{K}^{n,n}$ che ha dimensione n^2; esse non possono quindi essere linearmente indipendenti, deve, cioè, esistere una combinazione lineare nulla e non banale $a_0 \mathrm{Id} + a_1 A + a_2 A^2 + \cdots + a_{n^2} A^{n^2} = 0$. Ma allora il polinomio $a_0 + a_1 t + a_2 t^2 + \cdots + a_{n^2} t^{n^2}$ si annulla su A ed è quindi un elemento non banale del nucleo di v_A. □

Definizione 19. Visto che $\mathbb{K}[t]$ è un dominio ad ideali principali, l'ideale non banale $\operatorname{Ker} v_A$ è generato da un polinomio. A meno di dividere per il coefficiente del termine principale, possiamo scegliere un generatore monico; tale polinomio si chiama *polinomio minimo* di A e si indica con $\mu_A(t)$. Equivalentemente $\mu_A(t)$ è il polinomio monico di grado minore tra tutti i polinomi che si annullano su A.

Dalla dimostrazione della proposizione precedente sappiamo che il polinomio minimo di A ha grado minore uguale ad $n^2 + 1$. Il nostro prossimo obiettivo è mostrare che, in realtà, μ_A ha grado minore o uguale ad n. Esiste infatti un polinomio di grado n che si annulla su A, tale polinomio è $p_A(t)$, il polinomio caratteristico di A. Per dimostrare questo teorema procederemo in vari passi.

Lemma 20. *Se $C \in \mathsf{GL}_n(\mathbb{K})$ è una matrice invertibile allora $v_{CAC^{-1}}(p) = Cv_A(p)C^{-1}$ per ogni polinomio $p(t) \in \mathbb{K}[t]$.*

Dimostrazione. Sia $\varphi : \mathbb{K}[t] \longrightarrow \mathbb{K}^{n,n}$ l'applicazione $\varphi(p) = Cv_A(p)C^{-1}$. Osserviamo che φ è lineare; infatti, usando che v_A è lineare, abbiamo

$$\begin{aligned}
\varphi(\alpha p + \beta q) &= Cv_A(\alpha p + \beta q)C^{-1} \\
&= C(\alpha v_A(p) + \beta v_A(q))C^{-1} \\
&= \alpha C v_A(p) C^{-1} + \beta C v_A(q) C^{-1} \\
&= \varphi(p) + \varphi(q)
\end{aligned}$$

Inoltre φ è anche un omomorfismo di anelli visto che lo è v_A, infatti abbiamo

$$\begin{aligned}
\varphi(p \cdot q) &= Cv_A(p \cdot q)C^{-1} \\
&= Cv_A(p)v_A(q)C^{-1} \\
&= Cv_A(p)C^{-1}Cv_A(q)C^{-1} \\
&= \varphi(p)\varphi(q)
\end{aligned}$$

Ma allora $\varphi = v_{CAC^{-1}}$ visto che $\varphi(t) = Cv_A(t)C^{-1} = CAC^{-1} = v_{CAC^{-1}}(t)$ e $v_{CAC^{-1}}$ è l'unico omomorfismo di anelli che mandi t in CAC^{-1}. □

Corollario 21. *Il polinomio minimo è invariante per coniugio, cioè per ogni $A \in \mathbb{K}^{n,n}$ e $C \in \mathsf{GL}_n(\mathbb{K})$, i polinomi minimi di A e di CAC^{-1} sono uguali.*

Dimostrazione. Dal lemma precedente un polinomio si annulla in A se e solo se si annula in CAC^{-1}. Quindi $(\mu_A) = \operatorname{Ker} v_A = \operatorname{Ker} v_{CAC^{-1}} = (\mu_{CAC^{-1}})$ ed essendo μ_A e $\mu_{CAC^{-1}}$ entrambi monici abbiamo la tesi. □

Lemma 22. *Per ogni polinomio $p(t) \in \mathbb{K}[t]$ e coppia di matrici quadrate A, B si ha $p(A \oplus B) = p(A) \oplus p(B)$. In particolare $p(A \oplus B) = 0$ se e solo se $p(A) = 0$ e $p(B) = 0$.*

Dimostrazione. Sia $p(t) = \sum_{h=0}^{r} a_h t^h$ con $a_0, a_1, \ldots, a_r \in \mathbb{K}$. Osservando che la somma ed il prodotto di matrici a blocchi si eseguono blocco per blocco, abbiamo

$$\begin{aligned} p(A \oplus B) &= \textstyle\sum_{h=0}^{r} a_h (A \oplus B)^h \\ &= \textstyle\sum_{h=0}^{r} a_h (A^h \oplus B^h) \\ &= (\textstyle\sum_{h=0}^{r} a_h A^h) \oplus (\textstyle\sum_{h=0}^{r} a_h B^h) \\ &= p(A) \oplus p(B) \end{aligned}$$

La seconda parte della tesi segue allora subito, visto che una matrice a blocchi è nulla se e solo tutti i blocchi sono nulli. □

Corollario 23. *Per il polinomio minimo di una matrice a blocchi abbiamo*

$$\mu_{A \oplus B}(t) = \operatorname{mcm}(\mu_A(t), \mu_B(t))$$

Dimostrazione. Sia $p(t) = \operatorname{mcm}(\mu_A(t), \mu_B(t))$. Visto che $p(t)$ è un multiplo di $\mu_A(t)$ e $\mu_B(t)$, usando il lemma 22, abbiamo $p(A \oplus B) = p(A) \oplus p(B) = 0 \oplus 0 = 0$. Quindi $\mu_{A \oplus B}(t)$ divide $p(t)$.

D'altra parte, sempre usando il lemma 22, $0 = \mu_{A \oplus B}(A \oplus B) = \mu_{A \oplus B}(A) \oplus \mu_{A \oplus B}(B)$ e quindi $\mu_{A \oplus B}(A) = 0$ e $\mu_{A \oplus B}(B) = 0$. Allora $\mu_A(t)$ e $\mu_B(t)$ dividono $\mu_{A \oplus B}(t)$. Da cui $p(t)$ divide $\mu_{A \oplus B}(t)$ per definizione di minimo comune multiplo.

Per concludere basta allora osservare che $p(t)$ e $\mu_{A \oplus B}(t)$ sono entrambi monici. □

Lemma 24. *Un blocco di Jordan soddisfa il proprio polinomio caratteristico.*

Dimostrazione. Sia $A = J_n(\lambda)$ e osserviamo che, essendo A triangolare superiore, si ha $p_A(t) = (\lambda - t)^n$. Quindi $p_A(A) = (\lambda \operatorname{Id} - J_n(\lambda))^n = (-J_n(0))^n = 0$ per il lemma 13. □

Teorema 25 (Cayley–Hamilton)**.** *Ogni matrice soddisfa il proprio polinomio caratteristico, cioè $p_A(A) = 0$.*

Dimostrazione. Supponiamo che A sia in forma di Jordan; tratteremo il caso generale riconducendoci alla forma di Jordan.

Sia quindi $A = \oplus_{h=0}^{r} J_{n_h}(\lambda_h)$ la decomposizione di A in blocchi di Jordan. Essendo $A - t \operatorname{Id}$ una matrice triangolare superiore, troviamo subito $p_A(t) = \det(A - t \operatorname{Id}) = \prod_{h=1}^{r} (\lambda_h - t)^{n_h}$. In particolare si noti che $p_A(t)$ è un multiplo dei polinomi caratteristici dei blocchi. Allora, per il lemma 22, abbiamo $p_A(A) = \oplus_{h=0}^{r} p_A(J_{n_h}(\lambda_h))$.

Ora $p_A(J_{n_h}(\lambda)) = 0$, per ogni $h = 0, 1, \ldots, r$, visto che $p_A(t)$ è un multiplo dei polinomi caratteristici di ogni blocco e, per il lemma precedente, ogni blocco di Jordan annulla il proprio polinomio caratteristico.

Sia ora A una matrice generica in $\mathbb{K}^{n,n}$; possiamo considerare che $A \in \overline{\mathbb{K}}^{n,n}$. Visto che $\Sigma(f) \subseteq \overline{\mathbb{K}}$ esiste $C \in \mathsf{GL}_n(\overline{\mathbb{K}})$ per cui $B = CAC^{-1}$ è in forma di Jordan. Allora applicando il lemma 20, l'invarianza del polinomio caratteristico per coniugio e quanto sopra provato per una matrice in forma di Jordan abbiamo

$$\begin{aligned} p_A(A) &= v_A(p_A) \\ &= C^{-1}v_B(p_A)C \\ &= C^{-1}v_B(p_B)C \\ &= 0 \end{aligned}$$

□

Corollario 26. *Il polinomio caratteristico p_A è un elemento del nucleo $\operatorname{Ker} v_A$ della valutazione in A. In particolare $\mu_A(t)$ divide $p_A(t)$ e quindi il grado di $\mu_A(t)$ è minore o uguale ad n.*

Dimostrazione. Il primo asserto è il teorema di Cayley–Hamilton. Inoltre è chiaro che $p_A(t) = \det(A - t\,\mathrm{Id})$ ha grado n in quanto, sviluppando il determinante, il grado massimo si ottiene quando si moltiplicano tutti gli elementi sulla diagonale e questi sono in numero di n. □

Ci proponiamo ora di usare la forma canonica di Jordan per calcolare il polinomio minimo di una matrice $A \in \mathbb{K}^{n,n}$ o, equivalentemente, di un endomorfismo.

Considerando A come una matrice a coefficienti in $\overline{\mathbb{K}}$ possiamo trovare una matrice $C \in \mathsf{GL}_n(\overline{\mathbb{K}})$ per cui $J = CAC^{-1}$ è in forma di Jordan. Visto che sappiamo dal corollario 21 che il polinomio minimo è invariante per coniugio, si avrà $\mu_A(t) = \mu_J(t)$. Questo ci dice che possiamo limitarci a descrivere il polinomio minimo della matrice J in forma di Jordan.

Sia ora $\Sigma(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_s\}$ lo spettro di A in $\overline{\mathbb{K}}$. Scriviamo la decomposizione di J in blocchi di Jordan come $J = \oplus_{i=1}^{s} J_i$, con $J_i = \oplus_{j=1}^{r_i} J_{n_{i,j}}(\lambda_i)$, mettiamo cioè insieme tutti i blocchi relativi allo stesso autovalore.

Vogliamo provare la seguente formula

Osservazione 27. $\mu_J(t) = \prod_{i=1}^{s}(t - \lambda_i)^{n_i}$ con $n_i = \max(n_{i,1}, n_{i,2}, \ldots, n_{i,r_i})$.

Dimostrazione. Visto che $p_J(t) \in \operatorname{Ker} v_J = (\mu_J(t))$ dal corollario 26, si deve avere $\mu_J(t) = \prod_{i=1}^{s}(t - \lambda_i)^{m_i}$ per certi interi non negativi $m_1, m_2, \ldots, m_s$.

Abbiamo $0 = \mu_J(J) = \oplus_{i=1}^{s} \oplus_{j=1}^{r_i} \mu_J(J_{n_{i,j}}(\lambda_i))$ e quindi, per il lemma 22, $\mu_J(J_{n_{i,j}}(\lambda_i)) = 0$ per ogni i, j. Inoltre

$$\mu_J(J_{n_{i,j}}(\lambda_i)) = \prod_{h=1}^{s}(J_{n_{i,j}}(\lambda_i) - \lambda_h\,\mathrm{Id})^{m_h} = \prod_{h=1}^{s}(J_{n_{i,j}}(\lambda_i - \lambda_h))^{m_h}$$

Osserviamo ora che se $h \neq i$ allora $J_{n_{i,j}}(\lambda_i - \lambda_h)$ è una matrice triangolare superiore con sulla diagonale tutti $\lambda_i - \lambda_h \neq 0$. Si tratta quindi di una matrice invertibile. Allora bisogna che sia necessariamente il termine del prodotto corrispondente a $h = i$ ad essere nullo, cioè $J_{n_{i,j}}(0)^{m_i} = 0$ per ogni $j = 1, 2, \ldots, r_i$.

Ma per avere ciò è necessario e sufficiente che $m_i \geq \max(n_{i,1}, n_{i,2}, \ldots, n_{i,r_i})$ usando il lemma 13. Concludiamo che deve essere $m_i = \max(n_{i,1}, n_{i,2}, \ldots, n_{i,r_i})$ visto che il polinomio minimo è il polinomio di grado più basso che si annulla in J. □

5 Decomposizione di Wedderburn

In questa sezione assumiamo fissato un campo $\mathbb{K}$ algebricamente chiuso di caratteristica zero. In particolare, ogni endomorfismo ammette una forma di Jordan; nel seguito useremo questa esistenza senza fare esplicito riferimento alle ipotesi su $\mathbb{K}$.

5.1 Matrici ed endomorfismi semisemplici

Definizione 28. Diciamo che una matrice in $\mathbb{K}^{n,n}$ è *semisemplice* se è diagonalizzabile, cioè se esiste una matrice diagonale nella sua classe di coniugio. Analogamente diciamo che un endomorfismo di una spazio vettoriale è semisemplice se la matrice ad esso associata in una qualche base è semisemplice. Ovviamente tale definizione non dipende dalla base scelta essendo tutte le matrici associate ad uno stesso endomorfismo coniugate tra loro.

Proposizione 29. *Una matrice S è semisemplice se e solo se il suo polinomio minimo non ha radici multiple.*

Dimostrazione. È chiaro che una matrice è semisemplice se e solo se la sua forma di Jordan è diagonale. Ma dall'osservazione 27 sappiamo che il polinomio minimo di S ha radici multiple se e solo se nella sua forma di Jordan esiste un blocco almeno 2×2. $\square$

Proposizione 30. *Sia $f : V \longrightarrow V$ un endomorfismo semisemplice e sia $V = U \oplus W$ una decomposizione di V in sottospazi f–invarianti. Allora $f_{|U}$ e $f_{|W}$ sono semisemplici.*

Dimostrazione. Il polinomio minimo di f non ha radici multiple visto che f è semisemplice. Ma dal corollario 23 abbiamo $\mu_f(t) = \mathrm{mcm}(\mu_{f_{|U}}(t), \mu_{f_{|W}}(t))$ e quindi neanche $\mu_{f_{|U}}(t)$ e $\mu_{f_{|W}}(t)$ hanno radici multiple. Ne segue che $f_{|U}$ e $f_{|W}$ sono semisemplici. $\square$

Teorema 31. *Se $f, g : V \longrightarrow V$ sono due endomorfismi semisemplici che commutano allora essi sono diagonalizzabili contemporaneamente, esiste cioè una base di V rispetto a cui le matrici associate ad f e a g sono* entrambe *diagonali.*

Dimostrazione. Dato λ autovalore di f, indichiamo con $V_f(\lambda)$ l'autospazio di λ rispetto ad f. Essendo f semisemplice abbiamo la decomposizone $V = \oplus_{\lambda\in\Sigma(f)} V_f(\lambda)$.

Se $v \in V_f(\lambda)$ allora, usando la commutatività di f e g, si ha $f(g(v)) = (f \circ g)(v) = (g \circ f)(v) = g(f(v)) = g(\lambda v) = \lambda g(v)$. Quindi $g(v) \in V_f(\lambda)$. Questo prova che $V_f(\lambda)$ è uno spazio g–invariante per ogni autovalore λ.

Sia ora $\mathcal{B}$ una base in cui A, la matrice associata ad f, è diagonale. Per quanto appena visto, la matrice B associata a g in tale base è a blocchi, un blocco per ogni $V_f(\lambda)$.

Inoltre dalla proposizione 30 concludiamo che $g_{|V_f(\lambda)}$ è un endomorfismo semisemplice di $V_f(\lambda)$. Quindi ogni blocco di B è diagonalizzabile. Esiste allora una matrice a blocchi $C = \oplus C_\lambda$ tale che CBC^{-1} è diagonale.

Ma C commuta con A in quanto in ogni blocco A è una matrice diagonale e le matrici diagonali commutano con ogni matrice. Quindi $A = CAC^{-1}$ e nella base $C \cdot \mathcal{B}$ le matrici associate ad f e g sono diagonali. $\square$

Corollario 32. *Se S_1 e S_2 sono due matrici semisemplici che commutano allora anche $S_1 + S_2$ è semisemplice.*

Dimostrazione. Visto che S_1 e S_2 commutano esiste una matrice invertibile C per cui CS_1C^{-1} e CS_2C^{-1} sono entrambe diagonali. Allora anche $C(S_1+S_2)C^{-1} = CS_1C^{-1}+CS_2C^{-1}$ è diagonale in quanto somma di due matrici diagonali. Abbiamo provato che $S_1 + S_2$ è diagonalizzabile. $\square$

5.2 Matrici ed endomorfismi nilpotenti

Definizione 33. Una matrice N si dice *nilpotente* se esiste un intero non negativo h per cui $N^h = 0$. Inoltre il minimo di tali interi si chiama *indice di nilpotenza* di N.

Osserviamo che se N è nilpotente e $C \in \mathsf{GL}(\mathbb{K})$ allora anche CNC^{-1} è nilpotente: se N ha indice di nilpotenza h allora $(CNC^{-1})^h = CN^hC^{-1} = 0$. Quindi tutte le matrici nella classe di similitudine di una matrice nilpotente sono nilpotenti. Questo ci permette di definire un endomorfismo nilpotente come un endomorfismo la cui matrice associata rispetto ad una qualche base è nilpotente.

Proposizione 34. *Una matrice N è nilpotente se e solo se $\mu_N(t) = t^h$, inoltre h è l'indice di nilpotenza.*

Dimostrazione. La matrice N è nilpotente se e solo se $N^k = 0$ per qualche k, quindi $t^k \in \operatorname{Ker} v_N$. Allora $\mu_K(t) = t^h$ per qualche h essendo $\mu_K(t)$ il generatore di $\operatorname{Ker} v_N$. È ora chiaro anche che h è l'indice di nilpotenza visto che $\mu_N(t)$ è il polinomio di grado minimo che si annulla su N. □

Proposizione 35. *Sia $f : V \longrightarrow V$ un endomorfismo nilpotente. Se $V = U \oplus W$ è una decomposizione in sottospazi f–invarianti, allora $f_{|U}$ e $f_{|W}$ sono nilpotenti.*

Dimostrazione. Segue subito dal lemma precedente e dal corollario 23 che $\mu_{f_{|U}}(t)$ e $\mu_{f_{|W}}(t)$ sono potenze di t visto che lo è μ_f. □

Proposizione 36. *Se N_1, N_2 sono due matrici nilpotenti che commutano allora anche $N_1 + N_2$ è nilpotente.*

Dimostrazione. Siano h_1 e h_2 tali che $N_1^{h_1} = 0$ e $N_2^{h_2} = 0$. Le due matrici N_1 e N_2 commutano, possiamo quindi applicare la regola del binomio di Newton

$$(N_1 + N_2)^{h_1+h_2} = \sum_{k=0}^{h_1+h_2} \binom{h_1 + h_2}{k} N_1^k N_2^{h_1+h_2-k}$$

Osserviamo ora che se $k \geq h_1$ allora $N_1^{h_1} = 0$; mentre se $k < h_1$ allora $h_1 + h_2 - k > h_2$ e quindi $N_2^{h_1+h_2-k} = 0$. In conclusione, ogni termine della precedente somma è zero. Questo prova che $(N_1 + N_2)^{h_1+h_2} = 0$, cioè che $N_1 + N_2$ è una matrice nilpotente. □

5.3 Decomposizione di Wedderburn

Possiamo ora usare i risultati delle sezioni precedenti per dimostrare il seguente

Teorema 37 (Decomposizione di Wedderburn)**.** *Data $A \in \mathbb{K}^{n,n}$ esistono uniche due matrici S, N, con S semisemplice ed N nilpotente che commutano per cui $A = S + N$. Inoltre esistono $p(t)$ e $q(t)$ in $\mathbb{K}[t]$ tali che $S = p(A)$ e $N = q(A)$. In particolare N e S commutano con ogni matrice che commuti con A.*

Dimostrazione. Sia $C \in \mathsf{GL}_n(\mathbb{K})$ tale che $J = CAC^{-1}$ è in forma di Jordan e sia $\tilde{S}$ la parte diagonale di J e $\tilde{N} = J - \tilde{S}$. Allora $\tilde{S}$ è semisemplice e $\tilde{N}$ è nilpotente essendo nilpotente in ogni blocco di Jordan di J. Inoltre $\tilde{S}$ e $\tilde{N}$ commutano tra loro in quanto commutano in ogni blocco. Quindi ponendo $S = C^{-1}\tilde{S}C$ e $N = C^{-1}\tilde{N}C$ abbiamo $A = S + N$, S semisemplice, N nilpotente con S e N che commutano.

Non dimostriamo la parte relativa all'esistenza di $p(t)$ e $q(t)$.

Visto che S e N sono dei polinomi in A esse commutano con ogni matrice che commuti con A, come si prova subito.

Occupiamoci ora dell'unicità: siano S' e N' due matrici con S' semisemplice, N' nilpotente tali che S' e N' commutano e $A = S' + N'$. La matrice S' commuta con N' e quindi commuta anche con $A = S' + N'$ commutando, chiaramente, con se stessa. In particolare S' commuta con S. Quindi $S - S'$ è ancora semisemplice dal corollario 32. Allo stesso modo si ragiona per provare che $N' - N$ è nilpotente.

Ma abbiamo $S' + N' = A = S + N$ e quindi $S - S' = N' - N$ è una matrice semisemplice e nilpotente. Ma solo la matrice nulla è contemporaneamente semisemplice e nilpotente come segue dalle proposizioni 29 e 34. Quindi $S' = S$ e $N' = N$ e l'unicità è provata. $\square$

- `versione base.`
- `versione 29.05.13:`

    1. aggiunta data della versione nella prima pagina,
    2. corretti errori di battitura nell'esempio 2: U_1 era erroneamente indicato con U_2 nella parte finale,
    3. corretto errore di battitura nella quinta riga della dimostrazione della proposizione 18: A^{n^2} e non a^{n^2},
    4. corretto errore di battitura nella seconda riga della dimostrazione della proposizione 30: $f_{|W}$ e non $f_{|V}$,
    5. nella dimostrazione della proposizione 36 la formula del binomio di Newton è ora in formato display e non in linea.

- `versione 29.06.13:`

    corretta disuguaglianza nella quinta riga dell'enunciato della proposizione 12: $h \geq k$ e non $k \geq h$.

- `versione 04.07.13:`

    corretto errore di battitura nell'enunciato della proposizione 36 a pagina 13.

- `versione 20.02.15:`

    corretto errore di battitura nella dimostrazione del lemma 13 a pagina 7.

- `versione 09.02.16:`

    revisione generale.