UNIVERSITÀ DEGLI STUDI DI PADOVA
Facoltà di Ingegneria
Corso di Laurea in Ingegneria dell'Informazione

LA FORMA CANONICA DI JORDAN

Laureanda: Nicoletta Bof

Relatore: prof. M. Cristina Ronconi

Anno accademico 2010 - 2011
Padova, 27 settembre 2011

Indice

Introduzione

L'argomento trattato in questa tesina è la Forma Canonica di Jordan, un importante risultato degli studi di Camille Jordan, matematico francese del XIX secolo, da lui riportato nel trattato *Traité des substitutions et des équations algébriques*, opera scritta a Parigi nel 1870.

L'importanza di questa forma canonica, è strettamente legata al concetto di similitudine, una particolare relazione che esiste tra matrici quadrate di uno stesso ordine. La similitudine è una relazione di equivalenza, per cui l'insieme di tutte le matrici di un dato ordine viene partizionato in classi di equivalenza. Ogni singola classe di similitudine può essere identificata da un rappresentante, per cui la ricerca, tra tutte le matrici di una data classe, di un rappresentante semplice è importante. Dal momento infatti che la similitudine è la relazione che lega tutte le matrici di un dato endomorfismo $L : V \to V$, trovare una matrice di L che sia semplice e le cui proprietà siano subito evidenti, risulta molto utile per lo studio dell'endomorfismo stesso.
Se la classe di equivalenza è composta da matrici diagonalizzabili, allora il rappresentante più semplice per questa classe risulta essere la matrice diagonale che si ottiene attraverso la similitudine. Questa matrice inoltre risulta unica, a meno di permutazioni degli elementi della diagonale principale.
Se però la classe di equivalenza è composta da matrici non diagonalizzabili, è necessario cercare un rappresentante di forma semplice, anche se non più diagonale, e tale da essere anch'esso determinato in modo unico. La forma canonica di Jordan risponde a queste esigenze: si può infatti dimostrare che ogni matrice è simile a una matrice in forma canonica di Jordan, che, essendo diagonale a blocchi, con i singoli blocchi sopradiagonali, ha una forma decisamente semplice; è inoltre possibile provare che nella classe di similitudine c'è un solo rappresentante di questo tipo, a meno della permutazione dei blocchi della diagonale. Si nota anche che la forma canonica di Jordan coincide con la forma diagonale quando la matrice è diagonalizzabile, quindi può essere considerata una generalizzazione della forma diagonale.
Lo scopo di questa tesina è allora dimostrare che, comunque sia assegnata una matrice quadrata A, esiste una matrice A' in forma canonica di Jordan ad essa simile, e che questa matrice A' è univocamente determinata, a meno di permutazioni dei blocchi elementari di Jordan.

Oltre ad essere un rappresentante privilegiato di una classe di similitudine, la forma canonica di Jordan risulta anche particolarmente utile nella risoluzione di sistemi di equazioni differenziali a coefficienti costanti del primo ordine; questi sistemi sono diffusi in ambito ingegneristico, per cui una loro semplificazione è importante.
In alcuni corsi viene infatti richiesta la modellizzazione di sistemi di varia natura (biologici, fisici,etc.), operazione che porta alla descrizione di questi tramite sistemi di equazioni differenziali in forma di stato. Questi sistemi, dopo una opportuna linearizzazione, si presentano sotto forma di equazioni matriciali del tipo $X' = AX + GU$ (che rappresentano sistemi di equazioni differenziali lineari a coefficienti costanti del primo ordine), con X vettore colonna che rappresenta gli stati (X' la derivata del vettore X), U il vettore che rappresenta gli ingressi del sistema e A e G matrici che rappresentano il legame tra gli stati, gli ingressi e le derivate degli stati. La soluzione di questi sistemi risulta relativamente semplice se la matrice A che descrive il sistema stesso è in forma canonica di Jordan, e d'altra parte, come si vedrà successivamente, è sempre possibile ricondursi a questo caso semplice qualsiasi sia il sistema assegnato.

La tesina si svolge in 4 paragrafi: nel primo si riprendono concetti di teoria già trattati nel corso di Algebra Lineare e Geometria e se ne introducono alcuni strettamente legati a questi; questa parte non conterrà molte dimostrazioni, dal momento che contiene nozioni ampiamente note. Nel secondo si analizza in modo abbastanza approfondito la composizione reiterata di un endomorfismo con se stesso e le matrici nilpotenti; con questo paragrafo si introducono concetti che risulteranno fondamentali per la trattazione seguente. Il terzo è totalmente dedicato alla Forma Canonica di Jordan, in cui si dimostra che ogni matrice risulta simile ad una e una sola matrice in forma canonica di Jordan, e si illustra anche un metodo per ottenere la matrice che, per similitudine, rende una data matrice nella sua forma canonica. L'ultimo paragrafo è infine dedicato alle equazioni differenziali e al loro svolgimento con le matrici, e serve per mostrare l'utilità operativa della forma canonica studiata.

Per lo svolgimento del primo paragrafo è stato utilizzato in particolare [8], per la seconda e terza parte [1] e [3], per la quarta parte la linea guida è stata nuovamente [8] e qualche spunto si è preso anche da [2] e [6]. Per consultazione sono risultati utili anche [5], [4] e [7].

1 Richiami di algebra lineare

In questo paragrafo si riprendono alcuni concetti fondamentali dell'Algebra lineare, omettendo gran parte delle dimostrazioni, che possono essere trovate nei testi di base.

1.1 Applicazioni lineari

Definizione 1.1 (Applicazione lineare) *Siano V e V' due spazi vettoriali su uno stesso campo $\mathbb{K}$. Una applicazione (o funzione)*

$$L : V \to V'$$

si dice $\mathbb{K}$ lineare o semplicemente ***lineare*** *se soddisfa le seguenti condizioni:*

1. *$L(u + v) = L(u) + L(v)$ per ogni u, $v \in V$,*
2. *$L(\alpha v) = \alpha L(v)$ per ogni $v \in V$ e $\alpha \in \mathbb{K}$,*

dove le operazioni al primo membro di 1. e 2. sono la somma e il prodotto per uno scalare definite su V, e quelle al secondo membro sono le operazioni definite su V'.

Le applicazioni lineari vengono anche dette **trasformazioni lineari**. Un **isomorfismo** è una applicazione lineare biiettiva, cioè una funzione iniettiva e suriettiva.
Le applicazioni lineari di V in V di dicono **endomorfismi**; un endomorfismo biiettivo si dice anche **automorfismo**.

Proposizione 1.1 *Sia $L : V \to V'$ una applicazione lineare tra due spazi vettoriali V e V' su uno stesso campo $\mathbb{K}$. Allora:*

1. *l'immagine (tramite L) di un sottospazio W di V è un sottospazio di V';*
2. *l'antimmagine (tramite L) di un sottospazio W' di V' (cioè l'insieme $W = \{v \in V | L(v) \in W'\}$) è un sottospazio di V.*

Definizione 1.2 *Data $L : V \to V'$, l'immagine di V, $L(V)$, è secondo la proposizione precedente un sottospazio di V' e si dice* ***immagine*** *di V secondo L e viene indicata con $Im(L)$; la dimensione di $Im(L)$ si dice* ***rango*** *e si indica con $\mathrm{rg}(L)$.*
Si definisce inoltre il **nucleo** *di L come l'antimmagine di $\{0_{V'}\}$ e si indica con $ker(L)$; dalla proposizione precedente si può dedurre che il nucleo è un sottospazio di V.*

Può essere utile capire quando una data funzione lineare sia iniettiva o suriettiva e con la prossima osservazione introduciamo un criterio per stabilirlo

Osservazione 1.1 Sia $L : V \to V'$ una applicazione lineare; allora

- L è iniettiva $\Leftrightarrow ker(L) = \{0_V\}$;
- L è suriettiva $\Leftrightarrow Im(L) = V'$.

$\square$

Altra utile considerazione può essere fatta riguardo alle dimensioni dei diversi spazi vettoriali coinvolti in una applicazione lineare:

Proposizione 1.2 *Sia $L : V \to V'$ una applicazione lineare. Allora:*

$$\dim V = \dim Im(L) + \dim ker(L).$$

1.2 Relazioni tra applicazioni lineari e matrici

Esiste un forte legame tra le applicazioni lineari e le matrici; data infatti una funzione lineare $L : V \to V'$ e date $\mathcal{B}$ e $\mathcal{B}'$ basi rispettivamente di V e V', possiamo trovare la relazione che lega l'ennupla delle coordinate rispetto a $\mathcal{B}$ di un vettore $v \in V$ con le coordinate rispetto a $\mathcal{B}'$ di $L(v) \in V'$ attraverso l'uso di una matrice $m \times n$, dove $n = \dim V$ e $m = \dim V'$.

Proposizione 1.3 *Sia $L : V \to V'$ una applicazione lineare tra due spazi vettoriali V e V' (su uno stesso campo $\mathbb{K}$) di dimensione n e m rispettivamente. Siano inoltre $\mathcal{B}$ e $\mathcal{B}'$ basi di V e V'. Indicate con*

$$X_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ e con } X'_{\mathcal{B}'} = \begin{pmatrix} x'_1 \\ \vdots \\ x'_m \end{pmatrix}$$

le matrici colonna i cui elementi coincidono con le coordinate di v e di $L(v)$ rispetto a $\mathcal{B}$ e $\mathcal{B}'$, esiste una matrice $A \in M_{m,n}(\mathbb{K})$ tale che

$$X'_{\mathcal{B}'} = AX_{\mathcal{B}}.$$

Inoltre questa matrice A è univocamente determinata da L, $\mathcal{B}$ e $\mathcal{B}'$ e possiamo denotarla con $A^{\mathcal{B}'}_{L,\mathcal{B}}$.

Si noti che, detto v_j il j-esimo vettore della base $\mathcal{B}$ e chiamata $\mathcal{C}_j$ la colonna j-esima di A, questa contiene le coordinate di $L(v_j)$ rispetto alla base $\mathcal{B}'$. In particolare, se gli spazi vettoriali V e V' corrispondono rispettivamente a $\mathbb{K}^n$ e $\mathbb{K}^m$ e le basi $\mathcal{B}$ e $\mathcal{B}'$ corrispondono alle basi canoniche di $\mathbb{K}^n$ e $\mathbb{K}^m$, allora le colonne della matrice A generano l'immagine della funzione L.

Osservazione 1.2 Sia $L : V \to V'$ una funzione lineare e sia $A \in M_{m,n}$ una delle matrici di L. Allora

$$\mathrm{rg}\,(A) = \dim Im(L).$$

$\square$

Allo stesso modo, data una matrice A, si può trovare una funzione lineare ad essa associata:

Osservazione 1.3 Dati V e V' spazi vettoriali su $\mathbb{K}$ e $\mathcal{B} = (v_1, \dots, v_n)$ e $\mathcal{B}' = (v'_1, \dots, v'_m)$ due loro basi, comunque si fissi $A \in M_{m,n}(\mathbb{K})$ si può considerare la funzione $L_A : V \to V'$ che fa corrispondere al vettore $v = x_1 v_1 + \cdots + x_n v_n$ il vettore $v' = x'_1 v'_1 + \cdots + x'_m v'_m$ le cui coordinate rispetto a $\mathcal{B}'$ sono date da

$$\begin{pmatrix} x'_1 \\ \vdots \\ x'_m \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Si può verificare che L_A è una funzione lineare. A risulta dunque la matrice di L_A rispetto alle basi $\mathcal{B}$ e $\mathcal{B}'$.
Nel caso particolare in cui gli spazi vettoriali V e V' corrispondano rispettivamente a $\mathbb{K}^n$ e $\mathbb{K}^m$ e le basi $\mathcal{B}$ e $\mathcal{B}'$corrispondano alle basi canoniche di $\mathbb{K}^n$ e $\mathbb{K}^m$, allora L_A viene detta **funzione lineare associata in modo canonico ad** A. $\square$

Analizzando la matrice A della funzione $L : V \to V'$ rispetto a due basi qualsiasi del dominio e codominio, possiamo analizzare alcune proprietà di L: dati infatti n e m corrispondenti alle dimensioni di V e V'

- L è iniettiva $\Leftrightarrow \mathrm{rg}(A) = n$;
- L è suriettiva $\Leftrightarrow \mathrm{rg}(A) = m$;

- L è biiettiva $\Leftrightarrow \mathrm{rg}(A) = n = m$.

La matrice associata a una data funzione L dipende dalle basi che vengono scelte per V e V', per cui se abbiamo le basi $\mathcal{B}$ e $\mathcal{C}$ in V, e $\mathcal{B}'$ e $\mathcal{C}'$ in V', chiamate $X_{\mathcal{B}}$ e $Y_{\mathcal{C}}$ le matrici colonna delle coordinate di un vettore $v \in V$ rispetto alle basi $\mathcal{B}$ e $\mathcal{C}$, e chiamate $X'_{\mathcal{B}'}$ e $Y'_{\mathcal{C}'}$ le coordinate di $L(v) \in V'$ rispetto alle basi $\mathcal{B}'$ e $\mathcal{C}'$, esisteranno 2 matrici, $A^{\mathcal{B}'}_{L,\mathcal{B}}$ e $D^{\mathcal{C}'}_{L,\mathcal{C}}$, tali che

$$X'_{\mathcal{B}'} = AX_{\mathcal{B}} \tag{1.1}$$

$$Y'_{\mathcal{C}'} = DY_{\mathcal{C}} \tag{1.2}$$

Il legame che esiste tra $X_{\mathcal{B}}$ e $Y_{\mathcal{C}}$ si ottiene, ad esempio, con la matrice di passaggio dalla base $\mathcal{C}$ alla base $\mathcal{B}$; indicata questa con $K \in M_n$ ($K = K^{\mathcal{B}}_{\mathcal{C}}$) allora risulta $X_{\mathcal{B}} = KY_{\mathcal{C}}$.
Allo stesso modo il legame tra $X'_{\mathcal{B}'}$ e $Y'_{\mathcal{C}'}$ può essere rappresentato dalla matrice $H \in M_m$ ($H = H^{\mathcal{B}'}_{\mathcal{C}'}$), che permette il passaggio dalla base $\mathcal{C}'$ a quella $\mathcal{B}'$; usando H si ottiene infatti $X'_{\mathcal{B}'} = HY'_{\mathcal{C}'}$.
Facendo in (1.1) le dovute sostituzioni si ottiene $HY'_{\mathcal{C}'} = AKY_{\mathcal{C}}$ e quindi

$$Y'_{\mathcal{C}'} = H^{-1}AKY_{\mathcal{C}}.$$

Confrontando con la (1.2), risulta $D = H^{-1}AK$ (si ricorda che l'invertibilità di H è assicurata dal suo essere una matrice di cambio base). Per le nozioni note relative alle matrici per il cambio di base $H^{-1} = \left(H^{\mathcal{B}'}_{\mathcal{C}'}\right)^{-1} = H^{\mathcal{C}'}_{\mathcal{B}'}$, quindi

$$D^{\mathcal{C}'}_{L,\mathcal{C}} = H^{\mathcal{C}'}_{\mathcal{B}'} A^{\mathcal{B}'}_{L,\mathcal{B}} K^{\mathcal{B}}_{\mathcal{C}}.$$

Avendo una matrice di una data funzione lineare L, le altre matrici relative a L si possono quindi ottenere moltiplicando la matrice data per opportune matrici di passaggio da una base ad un'altra: la moltiplicazione a destra corrisponde a un cambio di base nel dominio, quella a sinistra a uno nel codominio.
Nel caso particolare di endomorfismi, normalmente si usa la stessa base $\mathcal{B}$ per il dominio e il codominio (in questo caso la matrice $A^{\mathcal{B}}_{L,\mathcal{B}}$ si dice matrice dell'endomorfismo L relativa alla base $\mathcal{B}$), per cui le diverse matrici di un endomorfismo sono del tipo

$$D = H^{-1}AH$$

con H matrice di passaggio dalla base $\mathcal{C}$ alla base $\mathcal{B}$ (ricordiamo infatti che $\mathcal{B} = \mathcal{B}'$ e $\mathcal{C} = \mathcal{C}'$).

1.3 Composizione di applicazioni lineari

Nel capitolo relativo alla Forma Canonica di Jordan saranno importanti le nozioni che vengono qui introdotte, cioè le composizioni tra funzioni e in particolare la composizione di una applicazione lineare con se stessa, eventualmente ripetuta più volte; in questa parte vengono riportate anche alcune dimostrazioni.

Definizione 1.3 *Siano V, V', V'' tre spazi vettoriali su uno stesso campo $\mathbb{K}$ e siano:*

$$L_1 : V \to V' \quad L_2 : V' \to V''$$

due applicazioni lineari. L'applicazione:

$$\begin{array}{cccc} L & : & V & \to & V'' \\ & & v & \mapsto & L_2(L_1(v)) \end{array}$$

*viene detta **applicazione composta** di L_1 e L_2 e denotata con $L = L_2 \circ L_1$.*

Proposizione 1.4 *La funzione che risulta dalla composizione di due funzioni lineari è lineare.*

Dimostrazione:

$$\begin{array}{lll} L(u+v) & = L_2(L_1(u+v)) = & \text{(per la linearità di } L_1) \\ & = L_2(L_1(u)+L_1(v)) = & \text{(per la linearità di } L_2) \\ & = L_2(L_1(u)) + L_2(L_1(v)) = & \\ & = L(u)+L(v), & \end{array}$$

$\forall u, v \in V$.

Allo stesso modo

$$\begin{array}{lll} L(\alpha u) & = L_2(L_1(\alpha u)) = & \text{(per la linearità di } L_1) \\ & = L_2(\alpha L_1(u)) = & \text{(per la linearità di } L_2) \\ & = \alpha L_2(L_1(u)) = & \\ & = \alpha L(u), & \end{array}$$

$\forall \alpha \in \mathbb{K}$ e $\forall u \in V$. □

Osservazione 1.4 Data una funzione lineare $L = L_2 \circ L_1$, con $L_1 : V \to V'$, $L_2 : V' \to V''$, $L : V \to V''$, allora si ha che

1. $Im(L) \subseteq Im(L_2)$;
2. $ker(L) \supseteq ker(L_1)$.

□

Dimostrazione:

1. se $v'' \in Im(L) \Rightarrow \exists\, v \in V \,|\, L_2(L_1(v)) = v''$ ma allora, detto $v' = L_1(v) \in V'$, $L_2(v') = v'' \Rightarrow v'' \in Im(L_2)$; tutto questo vale $\forall\, v'' \in Im(L)$;

2. se $v \in ker(L_1) \Rightarrow L_1(v) = 0_{V'}$; dal momento che $L_2(0_{V'}) = 0_{V''}$ allora $L(v) = L_2(L_1(v)) = L_2(0_{V'}) = 0_{V''} \Rightarrow v \in ker(L)$, $\forall\, v \in ker(L_1)$. □

Osservazione 1.5 Se consideriamo le matrici relative alle funzioni lineari, la composizione di funzioni comporta la moltiplicazione tra le matrici che le rappresentano. Fissate infatti delle basi in V, V' e V'' e indicate con A la matrice della funzione $L_1 : V \to V'$ e con B quella di $L_2 : V' \to V''$, relative alle rispettive basi fissate, allora la matrice che rappresenta la funzione composta $L = L_2 \circ L_1$ rispetto alle basi di V e V'' fissate è

$$C = BA.$$

Come si nota è di fondamentale importanza l'ordine in cui le matrici vengono moltiplicate. □

1.4 Autovettori e diagonalizzazione

Una matrice diagonalizzabile A è una particolare matrice per la quale esiste una matrice H invertibile (di cambio base) tale che $H^{-1}AH$ è una matrice in forma diagonale. Questa tipologia di matrici viene qui trattata in modo esplicito ed approfondito per le ragioni esposte nell'introduzione. Alla fine della trattazione risulterà chiaro che non tutte le matrici risultano diagonalizzabili e proprio per questo è necessaria l'introduzione della forma canonica di Jordan, che, sebbene non sempre diagonale, è la forma il più possibile simile a questa.

Definizione 1.4 *Sia V uno spazio vettoriale su $\mathbb{K}$ e sia L un endomorfismo di V. Si dice* **autovettore** *di L ogni vettore $v \in V, v \neq 0_V$, tale che*

$$L(v) = \lambda v$$

per qualche $\lambda \in \mathbb{K}$.

Definizione 1.5 *$\lambda \in \mathbb{K}$ si dice* **autovalore** *di L se esiste un vettore $v \in V, v \neq 0_V$, tale che*

$$L(v) = \lambda v.$$

Fissato $\lambda \in \mathbb{K}$, i vettori v non nulli e tali che $L(v) = \lambda v$ si dicono autovettori associati all'autovalore λ.

Osservazione 1.6 $\lambda = 0$ è un autovalore $\Leftrightarrow ker(L) \neq \{0_V\}$. $\square$

Osservazione 1.7 Ogni autovettore può essere associato ad un unico autovalore $\lambda \in \mathbb{K}$. $\square$

Possono essere fatte alcune considerazioni generali sulle matrici relative ad endomorfismi e sugli autovalori e autovettori dell'endomorfismo stesso. Data infatti la matrice $A \in M_n(\mathbb{K})$, relativa all'endomorfismo L e alla base $\mathcal{B}$:

1. Se uno dei vettori di $\mathcal{B}$, per esempio v_i, è un autovettore di L associato a λ, la matrice A risulta essere del tipo

$$\begin{pmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ a_{21} & \cdots & 0 & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & \cdots & \lambda & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{pmatrix} \text{dove} \begin{pmatrix} 0 \\ 0 \\ \cdots \\ \lambda \\ \cdots \\ 0 \end{pmatrix} \text{è la colonna i} - \text{esima;}$$

   questa infatti rappresenta le coordinate di $L(v_i)$ rispetto alla base $\mathcal{B}$ stessa.

2. Allo stesso modo, se la colonna j-esima della matrice A è nulla ad eccezione eventualmente dell'elemento a_{jj}, allora il j-esimo vettore della base $\mathcal{B}$ è un autovettore di L relativo all'autovalore a_{jj}.

Definizione 1.6 *Sia $\lambda \in \mathbb{K}$ un autovalore dell'endomorfismo $L : V \to V$, l'insieme $V_\lambda = \{v \in V | L(v) = \lambda v\}$ è un sottospazio vettoriale di V e si dice* **autospazio** *relativo all'autovalore λ.*

Per determinare gli autovalori di un dato endomorfismo si può utilizzare la seguente

Proposizione 1.5 *Dati $L : V \to V$ e $A \in M_n(\mathbb{K})$ la matrice di L relativa alla base $\mathcal{B}$, allora*

$$\lambda \in \mathbb{K} \textit{ è un autovalore di } L \Leftrightarrow \det(A - \lambda I_n) = 0.$$

Un metodo equivalente per il calcolo degli autovalori prende in considerazione il rango della matrice $A - \lambda I_n$; se A è la matrice relativa all'endomorfismo L, $A - \lambda I_n$ rappresenta quella relativa alla funzione

$L_\lambda : V \to V$ che a un generico $v \in V$ associa $L(v) - \lambda v$. Risulta allora che il nucleo di L_λ coincide con l'autospazio V_λ di L relativo a λ. Si può quindi concludere che

$$\lambda \in \mathbb{K} \text{ è un autovalore di L} \Leftrightarrow \mathrm{rg}(A - \lambda I_n) < n.$$

Come si deduce dalla proposizione 1.2 e dall'osservazione 1.2, risulta inoltre

$$\dim V_\lambda = n - \mathrm{rg}(A - \lambda I_n).$$

Data una matrice A dell'endomorfismo L relativa a una base $\mathcal{B}$, e detta T una indeterminata in $\mathbb{K}$, possiamo considerare il polinomio $P_L(T) = \det(A - TI_n)$, che risulta essere di grado n; introducendo questo polinomio la proposizione 1.5 può essere rivista ottenendo che

$$\lambda \in \mathbb{K} \text{ è un autovalore di } L \Leftrightarrow \lambda \text{ è una radice del polinomio } \det(A - TI_n).$$

Questo polinomio risulta utile perché, come dimostrato dalla proposizione seguente, esso non dipende dalla particolare matrice di L scelta, cioè non dipende dalla base fissata in V.

Proposizione 1.6 *Siano A e A' matrici di uno stesso endomorfismo L di V, relative alle basi $\mathcal{B}$ e $\mathcal{B}'$ rispettivamente. Allora i due polinomi*

$$P_L(T) = \det(A - TI_n) \quad \text{e} \quad P'_L(T) = \det(A' - TI_n)$$

coincidono.

Definizione 1.7 *Dato L un endomorfismo di V e detta A la matrice di L relativa a una base qualsiasi di V, il polinomio*

$$P_L(T) = \det(A - TI_n)$$

si dice **polinomio caratteristico** *di L. L'equazione $P_L(T) = 0$ si dice* **equazione caratteristica** *di L.*

Definizione 1.8 *Dato un autovalore λ di L, si dice* **molteplicità algebrica** *di λ, e si indica con $\mu(\lambda)$, la molteplicità di λ come radice del polinomio caratteristico $P_L(T)$. Se un autovalore ha molteplicità algebrica pari a 1 si dice che è un autovalore semplice.*
La dimensione di V_λ si dice **molteplicità geometrica** *di λ.*

Dal momento che $P_L(T)$ è un polinomio di grado n, si ha che

$$\sum \mu(\lambda_i) \leq \dim V = n.$$

La disuguaglianza risulta stretta se il polinomio ha dei fattori irriducibili di grado maggiore o uguale a due, altrimenti è una uguaglianza. Gli autovalori di un endomorfismo dipendono quindi dal campo $\mathbb{K}$ degli scalari; se $\mathbb{K}$ coincide con $\mathbb{C}$, per il teorema fondamentale dell'algebra, $P_L(T)$ è scomponibile in fattori di primo grado, ma se coincide con $\mathbb{R}$ questo non è sempre detto. Per questo motivo il lavoro condotto da Camille Jordan si concentra sullo studio di matrici con elementi in $\mathbb{C}$.

Proposizione 1.7 *Sia λ un autovalore dell'endomorfismo L. Allora*

$$\dim V_\lambda \leq \mu(\lambda).$$

Come conseguenza della proposizione precedente e della definizione di autovalore si ha che se $\mu(\lambda) = 1$ allora $\dim V_\lambda = 1$.

Caratteristica fondamentale degli autospazi è che la loro somma risulta diretta, come si afferma nella prossima

Proposizione 1.8 *Sia L un endomorfismo di V; comunque vengano fissati degli autovalori $\lambda_1, \dots, \lambda_r$, con $\lambda_i \neq \lambda_j$ per $i \neq j$, la somma degli autospazi $V_{\lambda_1} + \cdots + V_{\lambda_r}$ è diretta:*

$$V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$$

Come conseguenza di questo fatto, la famiglia di vettori che si ottiene riunendo basi dei singoli autospazi è una famiglia linearmente indipendente formata da autovettori. Il massimo numero di autovettori linearmente indipendenti è quindi pari alla somma delle dimensioni di tutti i singoli autospazi.

Tutte le considerazioni fatte finora sugli autovettori e autovalori di un endomorfismo possono essere estese anche alle matrici. In questo caso, data $A \in M_n(\mathbb{K})$, una ennupla non nulla $X \in \mathbb{K}^n$ si dice **autovettore** di A se $AX = \lambda X$ per qualche $\lambda \in \mathbb{K}$; allo stesso modo $\lambda \in \mathbb{K}$ si dice **autovalore** di A se esiste un'ennupla $X \neq 0$ tale che $AX = \lambda X$.
Inoltre, come è noto dall'osservazione 1.2, esiste la funzione lineare associata in modo canonico ad $A \in M_n(\mathbb{K})$ (che in questo caso sarà un endomorfismo di $\mathbb{K}^n$), quindi possiamo affermare che gli autovettori, gli

autovalori, gli autospazi e il polinomio caratteristico di A coincidono con quelli dell'endomorfismo L_A, per cui tutte le considerazioni fatte precedentemente relativamente a un endomorfismo L possono essere pensate anche rispetto a una matrice quadrata A qualsiasi.

Definizione 1.9 *Sia L un endomorfismo di uno spazio vettoriale V di dimensione n su $\mathbb{K}$. L si dice* **diagonalizzabile** *se esiste una base di V tale che la matrice di L ad essa relativa sia diagonale (cioè abbia gli eventuali elementi diversi da 0 solo sulla diagonale principale).*

Proposizione 1.9 *Sia L un endomorfismo di V; allora sono equivalenti:*

1. *L è diagonalizzabile;*
2. *esiste una base di V formata da autovettori di L.*

Sfruttando le proposizioni 1.7, 1.8 e quella precedente, possiamo affermare che gli endomorfismi di V, con V di dimensione n, aventi n autovalori semplici, sono diagonalizzabili.

Definizione 1.10 *Una matrice $A \in M_n(\mathbb{K})$ si dice* **diagonalizzabile** *se esiste una matrice $H \in M_n(\mathbb{K})$ invertibile e tale che $\Delta = H^{-1}AH$ sia diagonale.*

Osservazione 1.8 Dato L_A l'endomorfismo associato in modo canonico alla matrice A, A è diagonalizzabile $\Leftrightarrow$ L_A è diagonalizzabile. $\square$

Possiamo anche in questo caso, sfruttando l'osservazione precedente, riscrivere la proposizione 1.9 per le matrici quadrate:

Proposizione 1.10 *Sia $A \in M_n(\mathbb{K})$; allora sono equivalenti:*

1. *A è diagonalizzabile;*
2. *esiste una base di $\mathbb{K}^n$ formata da autovettori di A.*

Inoltre, come per gli endomorfismi, se A ha n autovalori semplici allora è diagonalizzabile.

Nelle prossime proposizioni enunciamo delle condizioni necessarie e sufficienti affinché un endomorfismo o una matrice sia digonalizzabile:

Proposizione 1.11 *Sia L un endomorfismo di uno spazio vettoriale V di dimensione n su $\mathbb{K}$; siano $\lambda_1, \dots, \lambda_r$ gli autovalori di L, $\mu(\lambda_1), \dots, \mu(\lambda_r)$ le rispettive molteplicità e $V_{\lambda_1}, \dots, V_{\lambda_r}$ gli autospazi relativi a questi autovalori. Allora sono equivalenti:*

1. *L è diagonalizzabile;*
2. *$V = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_r}$;*
3. *$\sum_{i=1}^{r} \mu(\lambda_i) = n$ e $\dim V_{\lambda_i} = \mu(\lambda_i)$, per ogni $1 \leq i \leq r$.*

Proposizione 1.12 *Sia $A \in M_n(\mathbb{K})$ e siano $\lambda_1, \dots, \lambda_r$ gli autovalori di A, $\mu(\lambda_1), \dots, \mu(\lambda_r)$ le rispettive molteplicità e $V_{\lambda 1}, \dots, V_{\lambda r}$ gli autospazi relativi a questi autovalori. Allora sono equivalenti:*

1. *A è diagonalizzabile;*
2. *$\mathbb{K}^n = V_{\lambda 1} \oplus \dots \oplus V_{\lambda r}$;*
3. *$\sum_{i=1}^{r} \mu(\lambda_i) = n$ e $\dim V_{\lambda i} = \mu(\lambda_i)$, per ogni $1 \leq i \leq r$.*

Viste queste proposizioni possiamo affermare che non tutte le matrici quadrate sono diagonalizzabili: infatti le matrici per cui il polinomio caratteristico non è scomponibile in fattori di primo grado non possono essere messe in forma diagonale, e nemmeno quelle in cui per uno o più autovalori la molteplicità algebrica è strettamente superiore alla molteplicità geometrica.

Avendo una matrice $A \in M_n(\mathbb{K})$ diagonalizzabile, può essere utile trovare la matrice H che la diagonalizza. Dal momento che A risulta diagonalizzabile, allora esiste una base Γ di $\mathbb{K}^n$ formata da soli autovettori di A. Come abbiamo già visto, la matrice Δ di L_A relativa alla base Γ è diagonale; si ha quindi che, chiamata E_n la base canonica di $\mathbb{K}^n$ e $H_{\Gamma}^{E_n}$ matrice di passaggio dalla base Γ alla base canonica E_n

$$\Delta = \Delta_{L_A,\Gamma}^{\Gamma} = H_{E_n}^{\Gamma} A_{L_A,E_n}^{E_n} H_{\Gamma}^{E_n}.$$

$H_{\Gamma}^{E_n}$, per come è definita, ha come colonne i vettori della base Γ e quindi una base di $\mathbb{K}^n$ formata da autovettori di A; l'ordine in cui vengono disposti gli autovettori come colonne di H influisce sull'ordine degli elementi nella diagonale principale di Δ (se la colonna i-esima contiene l'autovettore relativo all'autovalore λ_i, l'elemento di Δ in posizione i, i corrisponde a λ_i).

1.5 Matrici simili

Definizione 1.11 *Assegnate* $A,\ B \in M_n(\mathbb{K})$*, A si dice* **simile** *a B, e si indica con $A \sim B$, se esiste una matrice $H \in M_n(\mathbb{K})$ invertibile e tale che*

$$B = H^{-1}AH.$$

Proposizione 1.13 *La relazione in $M_n(\mathbb{K})\colon R = \{(A, B)|A \sim B\}$ è una relazione di equivalenza.*

Osservazione 1.9 Date due matrici simili possiamo fare le seguenti considerazioni:

1. $A \sim B \Rightarrow \mathrm{rg}(A) = \mathrm{rg}(B)$,
   ma $\mathrm{rg}(A) = \mathrm{rg}(B) \not\Rightarrow A \sim B$.

2. $A \sim B \Rightarrow A$ e B hanno lo stesso polinomio caratteristico
   ma A e B con lo stesso polinomio caratteristico $\not\Rightarrow A \sim B$.

$\square$

Per la definizione di similitudine e per le considerazioni fatte rispetto alle matrici che rappresentano lo stesso endomorfismo usando basi diverse (si confronti l'ultima parte del sottoparagrafo 1.2), si può fare la seguente:

Osservazione 1.10 Matrici di uno stesso endomorfismo relative a due basi diverse sono simili.
Al contrario, matrici simili possono essere interpretate come matrici di uno stesso endomorfismo relative a basi diverse. $\square$

Sfruttando quest'ultima osservazione possiamo affermare che due matrici simili sono entrambe diagonalizzabili o entrambe non diagonalizzabili.

2 Endomorfismi: catene di immagini e endomorfismi nilpotenti

In questo capitolo riprendiamo il concetto di composizione di funzioni, applicandolo però alla composizione di un endomorfismo con se stesso (eventualemente ripetuta più volte), e ne evidenziamo qualche proprietà; introduciamo inoltre la nozione di trasformazione e matrice nilpotente, concetto che risulterà utile per la trattazione della forma canonica di Jordan.

2.1 Catene di immagini

Volendo in questa parte studiare in modo approfondito la composizione di un endomorfismo con se stesso, si fanno alcune premesse sulla notazione: assegnato un endomorfismo L non nullo di uno spazio vettoriale V su $\mathbb{C}$, indichiamo con L^m l'endomorfismo di V

$$L^m = L \circ \cdots \circ L \quad (m \text{ volte}),$$

con m intero positivo.
Il simbolo L^m può essere anche esteso al caso $m = 0$, intendendo per definizione che L^0 indichi l'endomorfismo identità, cioè quello che a ogni vettore $v \in V$ associa il vettore v stesso.
La prossima osservazione è una riscrittura della 1.3 e tratta nello specifico la composizione di un endomorfismo con se stesso, anche più volte:

Osservazione 2.11 Dato un endomorfismo $L : V \to V$ e indicata con A la sua matrice rispetto a una data base, la matrice che corrisponde a L^m è la matrice

$$A^m = A \cdots A \quad (m \text{ volte}).$$

□

Indicata, come di consueto, con A^0 la matrice identità, all'endomorfismo L^0 resta associata allora la matrice A^0. Le considerazioni fatte nell'osservazione 1.3 possono essere adattate al caso di endomorfismi e risulta:

Osservazione 2.12 Dato un endomorfismo $L : V \to V$ si ha che:

1. $\forall k \geq 1 \;\; Im(L^k) \subseteq Im(L^{k-1})$, quindi

$$Im(L) \supseteq Im(L^2) \supseteq Im(L^3) \supseteq \cdots;$$

2. $\forall k \geq 1 \;\; ker(L^k) \supseteq ker(L^{k-1})$, quindi

$$ker(L) \subseteq ker(L^2) \subseteq ker(L^3) \subseteq \cdots$$

$\square$

Indicata con A una qualsiasi delle matrici di L, dalla catena di immagini e dall'osservazione 1.2 possiamo concludere anche che

$$\mathrm{rg}\,(A) \geq \mathrm{rg}\,(A^2) \geq \mathrm{rg}\,(A^3) \geq \cdots \tag{2.1}$$

Inoltre i sottospazi $Im\,(L^k)$ e $ker\,(L^k)$ indicati nell'osservazione precedente, sono sottospazi L-stabili:

$$L(Im\,(L^k)) \subseteq Im\,(L^k)$$
$$L(ker\,(L^k)) \subseteq ker\,(L^k)$$

cioè la restrizione di L a $Im\,(L^k)$ induce su $Im\,(L^k)$ un endomorfismo e la restrizione a $ker\,(L^k)$ ne induce uno su $ker\,(L^k)$.

Dimostrazione: Sia $v \in Im\,(L^k)$; allora $v' = L(v) \in Im\,(L^{k+1})$, ma dato che $Im\,(L^{k+1}) \subseteq Im\,(L^k)$, si ha che $v' \in Im\,(L^k)$.
Nel secondo invece, se $v \in ker\,(L^k)$ allora $L^k(v) = 0_V$. Poichè $L^k(L(v)) = L(L^k(v)) = L(0_V) = 0_V$, si deduce che $L(v) \in ker\,(L^k)$. $\square$

Proposizione 2.1 *Se per un dato esponente h si ha $Im\,(L^h) = Im\,(L^{h+1})$, allora $Im\,(L^h) = Im\,(L^i)$ per ogni $i > h$.*

Dimostrazione: Per ogni vettore $u \in Im\,(L^h)$, cioè del tipo $u = L^h(v)$ per qualche $v \in V$, esiste per ipotesi w tale che $u = L^{h+1}(w)$, relazione che possiamo riscrivere come $u = Lu'$, intendendo con u' il vettore $u' = L^h(w)$; questo implica che l'endomorfismo indotto da L su $Im\,(L^k)$ è suriettivo e quindi anche iniettivo, per cui è un isomorfismo. Possiamo quindi concludere che per ogni $j \geq 1$ si ha $L^j(Im\,(L^h)) = Im\,(L^{j+h}) = Im\,(L^h)$. $\square$

Analizzando la catena (2.1), possiamo affermare che per ogni endomorfismo L esiste un indice minimo k_0 tale che $Im\,(L^{k_0}) = Im\,(L^{k_0+1})$. Infatti gli interi in (2.1) sono necessariamente non negativi, per cui per qualche k_0 deve essere $\mathrm{rg}\,(A^{k_0}) = \mathrm{rg}\,(A^{k_0+1})$; a questo punto basta ricordare che $Im(L^{i+1}) \subseteq Im(L^i)$ per ogni i, e che $\mathrm{rg}\,(A^i) = \dim L^i$.

Posto allora $B = A^{k_0}$, e dato F un endomorfismo che ha come matrice rappresentativa B rispetto a una data base, si ha che $Im\,(F) = Im\,(F^2)$; sfruttando questa nuova matrice possiamo enunciare la seguente:

Proposizione 2.2 *Se un endomorfismo $F : V \to V$ è tale che $\dim Im(F) = \dim Im(F^2)$, allora $ker\,(F)$ ed $Im\,(F)$ decompongono V in somma diretta:*

$$V = ker\,(F) \oplus Im\,(F).$$

Dimostrazione: Dalle considerazioni e dalla proposizione precedenti si ha che $Im\,(F) = Im\,(F^2)$, quindi la restrizione di F a $\mathrm{Im}\,(F)$ è biiettiva, per cui si ha che $Im\,(F) \cap ker\,(F) = \{0_V\}$. Bisogna infine mostrare che la somma dei due sottospazi di V coincide con V, dimostrazione immediata dal momento che per la proposizione 1.2, si ha $\dim ker(F) + \dim Im(F) = n$. □

Si introduce qui una proposizione riguardante somme dirette di sottospazi invarianti, che permetterà successivamente di fare importanti considerazioni su L^{k_0}:

Proposizione 2.3 *Sia $L : V \to V$ una trasformazione lineare e siano U e W due sottospazi di V L-stabili tali che $V = U \oplus W$. Allora:*

1. *esiste una base di V rispetto alla quale la matrice di L ad essa relativa risulta diagonale a blocchi (di ordini le dimensioni di U e W):*
$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix};$$
2. *gli autovalori di L sono dati dagli autovalori di L ristretta ad U e da quelli di L ristretta a W;*
3. *se λ è un autovalore di L di molteplicità algebrica μ e se μ_1 e μ_2 sono le molteplicità algebriche di λ delle restrizioni di L a U e a W rispettivamente, allora $\mu = \mu_1 + \mu_2$.*

Dimostrazione:

1. Se $\dim V = n$, $\dim U = h$, $\dim W = n - h$, prendiamo come base di V la famiglia di vettori $v_1, v_2, \dots, v_n$ ottenuta riunendo una base $v_1, \dots, v_h$ di U e una base $v_{h+1}, \dots, v_n$ di W. Dal momento che U e W sono L-stabili, $L(v_i)$, per $1 \leq i \leq h$ (quindi $v_i \in U$), si scrive come combinazione dei soli vettori della base di U, $v_1, \dots, v_h$. Allo stesso modo $L(v_j)$, per $h+1 \leq j \leq n$ ($v_j \in W$), risulta combinazione lineare dei soli vettori $v_{h+1}, \dots, v_n$. I coefficienti di queste combinazioni lineari determinano le matrici A_1 e A_2.
2. Per come abbiamo costruito la matrice A, si ha che $P_A(T) = P_{A_1}(T)P_{A_2}(T)$, quindi gli zeri del polinomio $P_A(T)$, corrispondenti agli autovalori di A (e di L), sono gli zeri di $P_{A_1}(T)$ e quelli di $P_{A_2}(T)$.
3. Per il punto precedente la molteplicità di un dato autovalore λ di L è data dalla somma della molteplicità di λ delle restrizioni di L a U e a W.

□

Proposizione 2.4 *Per ogni endomorfismo* $L : V \to V$ *sia* $k_0 > 0$ *l'indice minimo tale che* $Im(L^{k_0}) = Im(L^{k_0+1})$. *Allora:*

1. $Im(L^{k_0})$ *e* $ker(L^{k_0})$ *decompongono lo spazio* V *in somma diretta;*

2. *l'unico autovalore della restrizione di* L *a* $ker(L^{k_0})$ *è lo zero, mentre gli autovalori della restrizione a* $Im(L^{k_0})$ *sono tutti diversi da zero.*

3. *la molteplicità algebrica* $\mu(0)$ *è uguale alla dimensione di* $ker(L^{k_0})$ *e risulta* $\mu(0) = \dim ker(L^{k_0}) \geq k_0$.

Dimostrazione:

1. Prendendo $F = L^{k_0}$ l'affermazione è conseguenza della proposizione precedente.

2. Per la proposizione 2.1 la funzione L ristretta a $Im(L^{k_0})$ è un isomorfismo, dunque i suoi autovalori sono tutti non nulli.
   Sia inoltre λ un autovalore di L e v l'autovettore relativo a λ, quindi $L(v) = \lambda v$. Se $v \in ker(L^{k_0})$, si ha $L^{k_0}(v) = 0_V$, ma anche $L^{k_0}(v) = \lambda^{k_0} v$. Dal momento che v è un autovettore, v risulta diverso dal vettore nullo e quindi si ha necessariamente $\lambda = 0$.

3. Analizzando la catena $0_V \subseteq ker(L) \subseteq ker(L^2) \subseteq \cdots \subseteq ker(L^{k_0})$ ci si accorge che le inclusioni devono essere tutte strette, altrimenti avremmo che $Im(L^i) = Im(L^{i+1})$ con $i < k_0$; questo implica che $\dim ker(L^{k_0}) \geq k_0$. D'altra parte si è visto che $ker(L^{k_0})$ e $Im(L^{k_0})$ sono L-stabili, e per il primo punto di questa proposizione decompongono V in somma diretta. Allora per la proposizione 2.3 si ha che $\mu(0) = \dim ker(L^{k_0})$. □

Da questa proposizione otteniamo che k_0 è sempre minore o uguale della dimensione del $ker(L^{k_0})$; dal momento che la dimensione del nucleo di una data funzione risulta sempre minore o uguale di n (dimensione dello spazio vettoriale V), si ricava che $k_0 \leq n$, e quindi

- $Im(L^{k_0}) = Im(L^n)$ e in ogni caso dalla potenza n-esima di L la catena di immagini è stabile;

- $ker(L^{k_0}) = ker(L^n)$ e da $ker(L^n)$ la catena di nuclei è stabile.

Possiamo inoltre affermare che $ker(L^{k_0})$ è il massimo sottospazio L-stabile rispetto all'inclusione, su cui L ha autovalore 0, mentre $Im(L^k)$ è il massimo sottospazio L-stabile su cui L non ha autovalori nulli.

2.2 Endomorfismi e matrici nilpotenti

Definizione 2.1 *Un endomorfismo $N : V \to V$ si dice* **nilpotente** *se esiste un intero $m > 0$ tale che*

$$N^m(v) = 0_V \quad \forall\ v \in V$$

cioè se esiste un $m > 0$ per il quale N^m è l'endomorfismo nullo.
Il più piccolo indice positivo $m_0 > 0$ per cui N^{m_0} è l'endomorfismo nullo si dice **indice di nilpotenza** *di N.*

Una analoga definizione può essere fatta per una matrice.

Si noti che se N è un endomorfismo nilpotente di indice m_0, allora ogni matrice A di N è tale per cui A^{m_0} è la matrice nulla, cioè quella formata da soli elementi nulli.

Se una trasformazione è nilpotente, allora gode di certe caratteristiche, come viene affermato nella prossima:

Proposizione 2.5 *Sia N un endomorfismo in uno spazio vettoriale di dimensione n, allora le seguenti affermazioni sono equivalenti:*

1. *N è nilpotente;*
2. *N ammette il solo autovalore nullo ;*
3. *N^n è l'endomorfismo nullo.*

Dimostrazione: Sia A la matrice di N rispetto a una data base.

$1 \Rightarrow 2$ per ipotesi A^k è la matrice nulla per qualche k. Allora tutti gli autovalori di A^k sono nulli; preso un autovalore λ di A, λ^k è un autovalore di A^k, quindi $\lambda = 0$.

$2 \Rightarrow 3$ se A ha il solo autovalore nullo, questo ha molteplicità n (si ricordi che il campo è $\mathbb{C}$ e quindi la somma delle molteplicità algebriche degli autovalori, in questo caso solo 0, deve essere pari all'ordine della matrice); preso k il minimo intero per cui $Im(A^k) = Im(A^{k+1})$, per il punto 3 della proposizione 2.4 si ha che $n = \dim ker(A^k)$, quindi $\operatorname{rg}(A^k) = 0$, per cui A^k è la matrice nulla. Dal momento che per la stessa proposizione risulta $k \leq n$, si ha che A^n è la matrice nulla

$3 \Rightarrow 1$ dall'ipotesi segue che N risulta nilpotente secondo la definizione 2.1. $\square$

Definizione 2.2 *Sia m un intero maggiore di 0. La matrice quadrata $m \times m$ del tipo*

$$N_m = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

si dice **blocco nilpotente elementare.**

3 Forma canonica di Jordan

Nel paragrafo 1.4 sono state trattate le matrici diagonalizzabili, simili cioè a matrici diagonali, evidenziando però che non tutte le matrici lo sono. Vengono allora introdotte delle matrici di forma particolarmente semplice, dette in forma canonica di Jordan, perché, come si vedrà, ogni matrice risulta simile ad una di esse. Come caso particolare, le matrici in forma canonica di Jordan possono assumere la forma diagonale, rappresentando in questo caso la classe di similitudine di matrici diagonalizzabili.
Questo capitolo, riprendendo nozioni da quelli precedenti, illustrerà la forma canonica di Jordan e ne dimostrerà l'esistenza e l'unicità per ogni classe di equivalenza di matrici.

In tutto il paragrafo assumiamo che lo spazio vettoriale V sia uno spazio vettoriale complesso di dimensione n.

3.1 Il teorema e la strategia della dimostrazione

In questo sottoparagrafo enunceremo il teorema relativo alla forma canonica di Jordan, illustrando brevemente le linee della dimostrazione che verrà esposta nel dettaglio nei paragrafi successivi.

Definizione 3.1 *La matrice $J_{\lambda,k}$ quadrata di ordine k del tipo*

$$J_{\lambda,k} = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

si dice **blocco di Jordan** *relativo all'autovalore λ o* **blocco elementare di Jordan**.
Un blocco elementare di Jordan $J_{\lambda,1}$ corrisponde alla matrice con un solo elemento, cioè alla matrice (λ).

Secondo la definizione appena esposta, un blocco nilpotente elementare (confrontare la definizione 2.2) è un blocco elementare di Jordan relativo all'autovalore nullo.
Si può anche notare che un blocco di Jordan può essere scomposto nella somma di un blocco nilpotente elementare e di una matrice diagonale:

$$J_{\lambda,k} = \lambda I_k + N_k$$

intendendo con I_k la matrice identità di ordine k.

Prendendo la funzione lineare associata in modo canonico a $J_{\lambda,k}$, con $k > 1$, questa ha il solo autovalore λ la cui molteplicità algebrica $\mu(\lambda)$ è pari all'ordine della matrice quadrata, quindi $\mu(\lambda) = k$. Inoltre si nota che l'autospazio V_λ ha dimensione pari a 1, perché i soli autovettori sono della forma $v = (\alpha, 0, \dots, 0) \in \mathbb{C}^k$, $\alpha \neq 0$. Si può così concludere che tale endomorfismo non è diagonalizzabile.

Definizione 3.2 *Una matrice diagonale a blocchi, in cui ogni blocco risulta essere un blocco elementare di Jordan, si dice* **in forma canonica di Jordan**.

Se nella matrice J in forma canonica di Jordan i singoli blocchi elementari sono tutti di ordine uno, allora J risulta essere in forma diagonale, per cui si può affermare che la forma canonica di Jordan generalizza la forma diagonale.

Il prossimo teorema è quello che assicura che ogni matrice risulta simile a una matrice in forma canonica di Jordan e che questa matrice è (a meno di differenze non sostanziali) unica.
In questo contesto, se non ci sono ambiguità, per indicare la molteplicità algebrica dell'i-esimo autovalore λ_i, useremo semplicemente il simbolo μ_i.

Teorema 3.1 *Sia $L : V \to V$ una trasformazione lineare e siano $\lambda_1, \dots, \lambda_h$ i suoi autovalori distinti, di molteplicità algebrica $\mu_1, \dots, \mu_h$. Esiste allora almeno una base di V rispetto alla quale la matrice di L è la matrice diagonale a blocchi J*

$$J = \begin{pmatrix} J_{\lambda_1,n_{11}} & & & & & & \\ & \ddots & & & & & \\ & & J_{\lambda_1,n_{1t}} & & & & \\ & & & \ddots & & & \\ & & & & J_{\lambda_h,n_{h1}} & & \\ & & & & & \ddots & \\ & & & & & & J_{\lambda_h,n_{hs}} \end{pmatrix},$$

dove le $J_{\lambda_i,n_{ij}}$ sono blocchi elementari di Jordan e gli interi n_{ij} sono univocamente determinati da L.
La matrice J descritta risulta evidentemente in forma canonica di Jordan, come si deduce dalla definizione 3.2.

L'ordine n_{ij} dei singoli blocchi elementari di J è tale per cui

$$n_{11} + n_{12} + \cdots + n_{1t} = \mu_1, \quad \dots \quad , n_{h1} + \cdots + n_{hs} = \mu_h.$$

La matrice J risulta unica a meno di permutazioni dei singoli blocchi elementari di Jordan, permutazioni che però non influiscono in modo sostanziale sulla matrice stessa.

Questo teorema può essere rivisto anche per le matrici; data infatti una matrice quadrata $A \in M_n(\mathbb{C})$, possiamo interpretarla come matrice di una funzione lineare L rispetto a una base $\mathcal{B}$ di V, per cui esisterà sempre una matrice J in forma canonica di Jordan simile ad A, cioè tale che $J = H^{-1}AH$; basta infatti prendere H come la matrice di passaggio da una base $\mathcal{C}$, rispetto a cui la matrice di L risulti in forma canonica di Jordan, alla base $\mathcal{B}$. (Risulta infatti $J^{\mathcal{C}}_{L,\mathcal{C}} = H^{\mathcal{C}}_{\mathcal{B}} A^{\mathcal{B}}_{L,\mathcal{B}} H^{\mathcal{B}}_{\mathcal{C}}$).

Il teorema 3.1 si può dimostrare procedendo in tre passi distinti, qui esposti in maniera sintetica:

I Passo: procedimento di separazione degli autospazi
Si dimostra che per ogni endomorfismo L esistono dei sottospazi $U_1, \dots, U_h$ L-invarianti ($L(U_i) \subseteq U_i, \forall\ 1 \leq i \leq h$), tali per cui $V = U_1 \oplus \cdots \oplus U_h$ e per cui ogni U_i è il massimo sottospazio di V L-stabile in cui l'endomorfismo ammette il solo autovalore λ_i. Ciò comporta che, a patto di scegliere una opportuna base per V, la matrice che rappresenta L risulta in forma diagonale a blocchi del tipo:

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_h \end{pmatrix}$$

dove ogni A_i è di ordine μ_i e presenta il solo autovalore λ_i.

II Passo: riduzione all'autovalore nullo
Prendendo in considerazione ogni singola matrice A_i come descritta nel passo precedente, ci si riporta al caso dell'autovalore nullo considerando la matrice $A_i - \lambda_i I$, con I matrice identità della stesso ordine di A_i. Si dimostra inoltre che se esiste una matrice C invertibile tale per cui $C^{-1}(A_i - \lambda_i I)C$ è in forma canonica di Jordan, allora, usando la stessa matrice C, si ha anche che $C^{-1}A_iC$ è in forma canonica di Jordan.

III Passo: analisi delle matrici nilpotenti
Prendendo la matrice con il solo autovalore nullo, ottenuta dal passo precedente, si lavora su questa matrice nilpotente (si confronti la proposizione 2.5) e si dimostra per questa il teorema.

I prossimi tre paragrafi sono dedicati ad esporre in maniera dettagliata i passi appena descritti, che sono stati brevemente enunciati per permettere di avere già un'idea generale di come si vuole arrivare alla dimostrazione vera e propria del teorema.

3.2 Procedimento di separazione degli autospazi

Definizione 3.3 *Si chiama* **decomposizione di Fitting** *dello spazio V relativo all'endomorfismo L la decomposizione in somma diretta:*

$$V = ker(L^k) \oplus Im(L^k) = V_0 \oplus V_0'$$

dove $Im(L^k) = Im(L^{k+1})$, $V_0 = ker(L^k)$, $V_0' = Im(L^k)$.

La proposizione 2.4 assicura che se $Im(L^k) = Im(L^{k+1})$ allora la somma in questione tra l'immagine e il nucleo di L^k è diretta.
Conoscendo la molteplicità algebrica μ dell'autovalore nullo di L, per la proposizione 2.4 si ha che $V_0 = ker(L^\mu)$.Tuttavia il sottospazio V_0 può essere determinato anche non conoscendo il valore μ, dal momento che $V_0 = ker(L^k) = ker(L^n)$ come abbiamo già visto nel paragrafo 2.1.

Tutte le considerazioni fatte nel caso dell'autovalore nullo possono essere estese anche ad un autovalore λ qualsiasi, considerando l'endomorfismo L_λ di V, con $L_\lambda = L - \lambda I$, dove I indica in questo caso l'endomorfismo identità. Infatti, se λ è autovalore di L, allora 0 è autovalore di L_λ, di molteplicità algebrica e geometrica pari a quella che λ ha in L. Si può quindi fare la decomposizione di Fitting usando i due sottospazi $V_\lambda = ker((L - \lambda I)^k)$ e $V_\lambda' = Im((L - \lambda I)^k)$ ottenendo

$$V = V_\lambda \oplus V_\lambda'. \tag{3.1}$$

V_λ' risulta un sottospazio vettoriale L-stabile e su questo ci sono solo autovalori diversi da λ.

Definizione 3.4 *Dato l'endomorfismo $L : V \to V$, sia λ un suo autovalore e k_0 il minimo esponente per cui $Im((L - \lambda I)^{k_0}) = Im((L-\lambda I)^{k_0+1})$. Allora il sottospazio vettoriale*

$$V_\lambda = ker((L - \lambda I)^{k_0}) = \Big\{x \in V \mid (L - \lambda I)^{k_0}(x) = 0_V\Big\}$$

si dice **autospazio generalizzato** *della trasformazione L relativo all'autovalore λ.*
I vettori non nulli che lo compongono si chiamano **autovettori generalizzati** *relativi all'autovalore λ.*

Se non si conosce k_0, per determinare l'autospazio generalizzato relativo a λ, si può usare come esponente la molteplicità algebrica dell'autovalore, seguendo la proposizione 2.4.
Anche V_λ, come V'_λ, risulta un sottospazio L-invariante; inoltre la sua dimensione corrisponde alla dimensione del nucleo di $(L - \lambda I)^{k_0}$ e quindi alla molteplicità algebrica di λ in L (tutto questo sempre per la proposizione 2.4).
Si nota che l'autospazio generalizzato V_λ contiene anche l'autospazio (si confronti la definizione 1.6) e quindi gli autovettori relativi a λ sono alcuni dei suoi elementi (per questo la notazione usata è in effetti la stessa).

Teorema 3.2 (di decomposizione in autospazi generalizzati): *sia $L : V \to V$ una trasformazione lineare di autovalori distinti $\lambda_1, \lambda_2, \ldots, \lambda_h$ con rispettive molteplicità algebriche $\mu_1, \mu_2, \ldots, \mu_h$. Allora si ha la decomposizione:*

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_h}$$

dove V_{λ_i} è il massimo sottospazio L-stabile di V su cui L ha solo l'autovalore λ_i. Inoltre risulta $\dim V_{\lambda_i} = \mu_i$.

Dimostrazione: prendendo in considerazione l'autovalore λ_1 si suddivide V usando (3.1) ottenendo $V = V_{\lambda_1} \oplus V'_{\lambda_1}$. Usando la proposizione 2.3 possiamo trovare una base per V tale che la matrice di L sia diagonale a blocchi del tipo

$$\begin{pmatrix} A_1 & 0 \\ 0 & B \end{pmatrix}$$

dove A_1, di ordine μ_1, è la matrice della restrizione di L a V_{λ_1} e B quella della restrizione a V'_{λ_1}. Gli autovalori di B risultano, per costruzione, $\lambda_2, \ldots, \lambda_h$. A questo punto si itera nuovamente il ragionamento su V'_{λ_1} e la matrice B; questo ragionamento si può applicare ripetutamente fino ad analizzare anche l'autospazio generalizzato V_{λ_h}, ottenendo che la matrice A rispetto ad una opportuna base diventi

$$\begin{pmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & A_h \end{pmatrix}.$$

□

La ricerca degli autospazi generalizzati V_{λ_i} si effettua conoscendo gli autovalori e le rispettive molteplicità algebriche μ_i della funzione lineare data, dal momento che per determinare l'autospazio i-esimo basta risolvere il sistema $(A - \lambda_i I)^{\mu_i} x = \underline{0}$ (con $\underline{0}$ indicante il vettore nullo di $\mathbb{C}^n$), dove A è una qualsiasi matrice di L.

A questo punto è importante notare che se per uno o più autovalori si ha che l'autospazio generalizzato coincide con l'autospazio, la matrice A_i relativa a questo autovalore risulta essere diagonale (dunque già in forma canonica di Jordan) perché esiste una base di V_λ formata da autovettori.

Si è così dimostrata la possibilità di riportare la matrice dell'endomorfismo in una forma diagonale a blocchi, in cui ad ogni blocco è associato un unico autovalore, come ci proponevamo nel I PASSO.

3.3 Riduzione all'autovalore nullo

In questa breve parte si analizzano matrici quadrate aventi un unico autovalore, come le matrici A_i ottenute nel sottoparagrafo precedente. Lo scopo è dimostrare che ogni matrice invertibile C che, per similitudine, semplifica $A_i - \lambda I$ trovandone la forma canonica di Jordan, semplifica anche A_i.

Proposizione 3.1 *Data una matrice quadrata $A \in M_n(\mathbb{C})$ avente come unico autovalore λ, se C è una matrice invertibile di ordine n tale che $J = C^{-1}(A - \lambda I)C$ è una matrice in forma canonica di Jordan, allora la stessa matrice C è tale per cui $C^{-1}AC$ risulta anch'essa in forma di Jordan.*

Dimostrazione: sia A una matrice con un solo autovalore λ, e si consideri la matrice $(A - \lambda I)$, che ha 0 come unico autovalore. Sia C una matrice quadrata invertibile tale che $C^{-1}(A - \lambda I)C$ sia in forma canonica di Jordan:

$$C^{-1}(A - \lambda I)C = \begin{pmatrix} J_{0,t_1} & 0 & \dots & 0 \\ 0 & J_{0,t_2} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & J_{0,t_s} \end{pmatrix}.$$

Allora se C viene applicata a A otteniamo

$$C^{-1}AC = C^{-1}(A-\lambda I+\lambda I)C = C^{-1}(A-\lambda I)C+C^{-1}\lambda IC = C^{-1}(A-\lambda I)C+\lambda C^{-1}C =$$

$$= \begin{pmatrix} J_{0,t_1} & 0 & \dots & 0 \\ 0 & J_{0,t_2} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & J_{0,t_s} \end{pmatrix} + \lambda I = \begin{pmatrix} J_{\lambda,t_1} & 0 & \dots & 0 \\ 0 & J_{\lambda,t_2} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & J_{\lambda,t_s} \end{pmatrix}$$

che è una matrice in forma canonica di Jordan avente il solo autovalore λ. $\square$

Come abbiamo già evidenziato nella dimostrazione, la matrice $(A - \lambda I)$ presenta solo l'autovalore nullo e quindi, per la proposizione 2.5, questa matrice è nilpotente.
Con questa proposizione anche il secondo passo della dimostrazione è stato completato.

3.4 Analisi delle matrici nilpotenti

Con questa parte si conclude la dimostrazione in 3 passi del teorema 3.1; questo ultimo passaggio studia le matrici nilpotenti e spiega come queste possano essere ridotte in forma canonica di Jordan. Questo procedimento risulta importante perché, come si è visto nel secondo passo, esso consente di risalire alla forma di Jordan di una matrice avente un solo autovalore anche se questo è diverso da 0.

Nel seguito si introduce il concetto di stringa, che risulterà importante per lo sviluppo successivo.
Sia $J_{0,n}$ un blocco di Jordan di ordine n relativo all'autovalore 0; esso è anche un blocco nilpotente elementare (si confronti la definizione 2.2). Allora se $J_{0,n}$ è la matrice di un endomorfismo T di V relativa alla base formata dai vettori $v_1, u_2, \dots, u_n$, si ha che:

$$T(u_i) = u_{i-1}, \text{ per } 3 \leq \text{i} \leq \text{n}, \quad T(u_2) = v_1, \quad T(v_1) = 0_V. \tag{3.2}$$

La stessa trasformazione T può essere rappresentata con il seguente diagramma:

$$\overset{u_n}{*} \longrightarrow \overset{u_{n-1}}{*} \longrightarrow \cdots \longrightarrow \overset{u_i}{*} \longrightarrow \overset{u_{i-1}}{*} \longrightarrow \cdots \longrightarrow \overset{u_2}{*} \longrightarrow \overset{v_1}{*}$$

Quella appena illustrata si definisce **stringa**; ogni vertice del diagramma è rappresentato con $*$ e indica un vettore della base di V e le $n-1$ frecce che legano i differenti vettori rappresentano il legame tra gli stessi, quindi il modo in cui l'endomorfismo T opera. Si può notare che l'ultimo vettore della stringa, v_1, è un autovettore relativo allo 0, e che gli altri vettori della stringa sono autovettori generalizzati relativi allo 0, dal momento che, preso il vettore u_i, $T^i(u_i)$ risulta essere il vettore nullo. Dal diagramma si può anche concludere che T^n è la trasformazione nulla.
La lettura della stringa fornisce ulteriori informazioni:

- $Im(T^i) = < u_{n-i}, \dots, u_2, v_1 >$;

- $\dim Im(T^i) = n - i = \mathrm{rg}(T^i)$;
- $ker(T^i) = < v_1, u_2, \ldots, u_i >$.

È inoltre importante notare che la trasformazione T è univocamente determinata dalla stringa data, esiste cioè un solo endomorfismo di V che verifica le condizioni (3.2).
Data una matrice J in forma canonica di Jordan, avente tutti gli r blocchi di autovalore 0, cioè formata da blocchi $J_{0,n_1}, \ldots, J_{0,n_r}$, a questa J possiamo associare un diagramma formato da r stringhe, in cui la stringa relativa al blocco J_{0,n_i}, con $1 \leq i \leq r$, presenta n_i vertici e $n_i - 1$ frecce.

Considerata adesso una trasformazione nilpotente $T : V \to V$, cerchiamo una base di V rispetto alla quale la matrice di T risulti in forma canonica di Jordan. La matrice J di T in forma di Jordan risulterà formata da uno o più blocchi di Jordan che presentano il solo autovalore nullo (si confronti la proposizione 2.5), per cui è sufficiente costruire delle stringhe di questa matrice J e dimostrare che i vettori delle stringhe formano una base di V.

Per la costruzione delle stringhe si considerino i seguenti sottospazi di V:

$$K_0 = ker(T), \quad K_i = ker(T) \cap Im(T^i), \quad 1 \leq i \leq n.$$

Dal momento che l'osservazione 2.1 vale per tutti gli endomorfismi si ha che:

$$\{0_V\} = K_n \subseteq K_{n-1} \subseteq \cdots \subseteq K_i \subseteq K_{i-1} \subseteq \cdots \subseteq K_1 \subseteq K_0 = ker(T),$$

dove non è detto che tutti i K_i siano sottospazi distinti. Si possono allora considerare solo alcuni dei K_j, in particolare quelli per cui $K_j \neq \{0_V\}$ e $K_j \neq K_{j-1}$. Sia j_1 il massimo intero per cui $K_{j_1} \neq \{0_V\}$, j_2 il massimo intero per cui $K_{j_2} \neq K_{j_1}$, e così via fino a j_s, il massimo intero per cui $K_{j_s} \neq K_{j_{s-1}}$ e $K_{j_s} = ker(T)$. Si ottiene così la sequenza di indici $j_1 > j_2 > \cdots > j_s$ e la catena di inclusioni (questa volta proprie):

$$\{0_V\} \subset K_{j_1} \subset K_{j_2} \subset \cdots \subset K_{j_s} = ker(T).$$

Denotiamo anche queste grandezze:

$$n_1 = \dim K_{j_1},\ n_2 = \dim K_{j_2} - \dim K_{j_1}, \ldots,\ n_s = \dim K_{j_s} - \dim K_{j_{s-1}}$$

la cui somma $n_1 + n_2 + \cdots + n_s = \dim K_{j_s} = \dim ker(T) = m$, che è anche la dimensione dell'autospazio (come definito nel paragrafo 1) relativo all'autovalore 0, dal momento che il nucleo di T coincide con questo autospazio.

Il passo successivo è la costruzione di una base di $ker(T)$, che avviene in più momenti qui presentati.

a$_1$) Siano v_1, $v_2, \ldots$, v_{n_1} una base di K_{j_1} arbitrariamente scelta.

a$_2$) Si completi la base trovata per K_{j_1} ad una base di K_{j_2}, scegliendo arbitrariamente i vettori $v_{n_1+1}, \ldots$, $v_{n_1+n_2}$.

$\vdots$

a$_s$) Si completi infine la base di $K_{j_{s-1}}$ a una base di K_{j_s}, scegliendo a piacere i vettori $v_{n_1+\cdots+n_{s-1}+1}, \ldots$, v_m.

Con questo procedimento è stata definita una base di $ker(T)$ e di conseguenza anche dell'autospazio relativo a 0, per cui tutti i v_i possono essere intesi come i vertici terminali di una stringa.

A questo punto si individua una stringa per ciascun vettore della base del nucleo. Scriveremo, per semplicità, $v_{n_1+\cdots+n_{h-1}+i} = v_i^{(h)}$.

b$_1$) Per ogni vettore v_i, $1 \leq i \leq n_1$, si sceglie un vettore u_i in modo che $T^{j_1}(u_i) = v_i = v_i^{(1)}$; il vettore u_i esiste certamente dal momento che $v_i \in ker(T) \cap Im(T^{j_1})$. Il vettore u_i è quindi il vettore iniziale della stringa

$$\overset{u_i}{*} \longrightarrow \overset{T(u_i)}{*} \longrightarrow \overset{T^2(u_i)}{*} \longrightarrow \cdots \longrightarrow \overset{T^{j_1-1}(u_i)}{*} \longrightarrow \overset{v_i^{(1)}}{*}$$

Si ottengono alla fine n_1 stringhe di lunghezza j_1 (con j_1 frecce e j_1+1 vertici).

b$_2$) Per ogni vettore v_{n_1+i}, $1 \leq i \leq n_2$, troviamo un w_i tale che $T^{j_2}(w_i) = v_{n_1+i} = v_i^{(2)}$, vettore sempre esistente dal momento che $v_i^{(2)} \in ker(T) \cap Im(T^{j_2})$. Possiamo anche in questo caso determinare una stringa che inizia con w_i:

$$\overset{w_i}{*} \longrightarrow \overset{T(w_i)}{*} \longrightarrow \overset{T^2(w_i)}{*} \longrightarrow \cdots \longrightarrow \overset{T^{j_2-1}(w_i)}{*} \longrightarrow \overset{v_i^{(2)}}{*} \ .$$

Ci saranno quindi n_2 stringhe di lunghezza j_2.

$\vdots$

b$_s$) Come ultimo passo per ogni $v_{n_1+\cdots+n_{s-1}+i} = v_i^{(s)}, 1 \leq i \leq n_s$, si trova un vettore z_i in modo che $T^{j_s}(z_i) = v_i^{(s)}$, la cui esistenza può essere provata come nei passi precedenti. Allora z_i è il vettore iniziale di una stringa che termina con $v_i^{(s)}$:

$$\overset{z_i}{*} \longrightarrow \overset{T(z_i)}{*} \longrightarrow \overset{T^2(z_i)}{*} \longrightarrow \cdots \longrightarrow \overset{T^{j_s-1}(z_i)}{*} \longrightarrow \overset{v_i^{(s)}}{*} \; .$$

Si ottengono quindi n_s stringhe di lunghezza j_s.

Nel seguito si dimostra che la famiglia di vettori che si ottiene riunendo i vettori delle stringhe ottenuti nei passi precedenti $\mathrm{b}_1, \ldots, \mathrm{b}_\mathrm{s}$ è una base di V, cioè questi vettori sono linearmente indipendenti e generano V.

Indipendenza: disponiamo in una tabella i vettori trovati con il procedimento appena spiegato

$$\begin{array}{cccccccccc}
u_1 & T(u_1) & \ldots & * & * & \ldots & * & * & \ldots & T^{j_1}(u_1) = v_1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
u_{n_1} & T(u_{n_1}) & \ldots & * & * & \ldots & * & * & \ldots & T^{j_1}(u_{n_1}) = v_{n_1} \\
 & & & w_1 & T(w_1) & \ldots & * & * & \ldots & T^{j_2}(w_1) = v_{n_1+1} \\
 & & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\
 & & & w_{n_2} & T(w_{n_2}) & \ldots & * & * & \ldots & T^{j_2}(w_{n_2}) = v_{n_1+n_2} \\
 & & & & & & \vdots & \vdots & & \vdots \\
 & & & & & & z_{n_s} & T(z_{n_s}) & \ldots & T^{j_s}(z_{n_s}) = v_m
\end{array} \tag{3.3}$$

Risulta evidente che l'operazione T trasforma un vettore di una riga in quello successivo, e che l'ultimo viene trasformato in 0_V.
Si prenda una combinazione lineare formata da tutti questi vettori il cui risultato sia il vettore nullo. Allora se si applica la trasformazione T^{j_1} tutti i vettori si annullano, tranne quelli della prima colonna che creano quindi una relazione di dipendenza lineare tra $v_1, \ldots, v_{n_1}$ con gli stessi coefficienti che avevano in precedenza $u_1, \ldots, u_{n_1}$. I vettori $v_1, \ldots, v_{n_1}$ sono tra di loro linearmente indipendenti dal momento che formano una base di K_{j_1}, e quindi gli unici coefficienti della combinazione lineare che danno come risultato 0_V sono i coefficienti tutti nulli. Resta così dimostrato che, nella combinazione lineare dell'intera famiglia, i coefficienti relativi a $u_1, \ldots, u_{n_1}$ sono tutti nulli. A questo punto si procede sulla combinazione lineare degli altri vettori restanti usando la trasformazione T^{j_1-1} e con lo stesso ragionamento si possono eliminare anche i vettori $Tu_1, \ldots, Tu_{n_1}$ dalla combinazione lineare di partenza. In questo modo, usando le potenze decrescenti di T, si dimostra la lineare indipendenza di tutti i vettori trovati, dal momento che nella combinazione iniziale tutti i coefficienti sono obbligariamente nulli.

Spazio vettoriale generato: preso $v \in V$, $v \neq 0_V$ esiste un esponente h tale che $T^h(v) = 0_V$ e $T^{h-1}(v) \neq 0_V$, dal momento che l'endomorfismo T è nilpotente. Allora $T^{h-1}(v) \in ker(T) \cap Im(T^{h-1}) = K_{h-1}$, e quindi si può scrivere come combinazione lineare w dei vettori della base di K_{h-1}, cioè $T^{h-1}(v) = w$; w è a sua volta strettamente legato ai vettori che sono stati trovati durante la costruzione delle stringhe, esiste infatti una combinazione lineare w' di questi vettori tale che $w = T^{h-1}(w')$. Sfruttando le uguaglianze $T^{h-1}(v) = w$ e $w = T^{h-1}(w')$ si ha che $T^{h-1}(v) = T^{h-1}(w')$ cioè $T^{h-1}(v - w') = 0_V$. A questo punto si ha un nuovo vettore $v - w'$ che si annulla se si applica su questo la trasformazione T^{h-1}. Si può quindi reiterare su questo nuovo vettore il procedimento appena svolto, trovando dopo $h-1$ passi un vettore x per cui $T(x) = 0_V$. Il vettore x che si ottiene dopo tutte le precedenti iterazioni, risulta combinazione lineare dei vettori della base del nucleo di T, elementi presenti nella lista, e si ha quindi che x e tutti i vettori precedenti, anche v, risultano combinazione lineare dei vettori trovati con le stringhe. Dal momento che il vettore v è un qualunque vettore di V, la famiglia di vettori trovati genera tutto lo spazio vettoriale V.

Abbiamo così dimostrato che i vettori trovati nei passi $\mathrm{b}_1, \dots, \mathrm{b}_\mathrm{s}$ sono linearmente indipendenti. Ogni stringa individua univocamente un blocco di Jordan; in generale la stringa relativa al vettore $v_i^{(h)}$, con $1 \leq i \leq n_h$ e $1 \leq h \leq s$ individua un blocco di Jordan di ordine $j_h + 1$ di autovalore 0.
Inoltre, dal momento che riuniti i vettori di tutte le stringhe questi generano l'intero spazio vettoriale V, si ha che la famiglia di vettori è una base di V e che la matrice relativa a questa base trovata è una matrice in forma canonica di Jordan, perchè ad ogni stringa è legato un blocco elementare di Jordan indipendente da quelli delle altre stringhe, per cui la matrice risultante è diagonale a blocchi, con tutti blocchi elementari di Jordan.

Per scrivere la forma canonica di Jordan di una matrice nilpotente è in realtà importante determinare i valori $j_1 + 1, \dots, j_s + 1$ che indicano il numero di vertici delle diverse stringhe (e quindi l'ordine dei singoli blocchi elementari di Jordan aventi il solo autovalore nullo), e i valori $n_1, \dots, n_s$ che determinano il numero di blocchi elementari di quell'ordine (quindi se abbiamo j_i e n_i con $1 \leq i \leq s$, nella matrice in forma

canonica di Jordan ci saranno n_i blocchi di Jordan con autovalore nullo di ordine $j_i + 1$).
I suddetti valori possono essere determinati se prima è stata determinata la base con il procedimento spiegato; analizzando infatti la tabella (3.3), i valori $j_i + 1$ possono essere trovati contando il numero di vettori sulle righe (e quindi la lunghezza delle stringhe) e i valori n_i contando quante stringhe di lunghezza $j_i + 1$ sono presenti.

Gli stessi valori possono però essere determinati senza costruire la base che rende la matrice in forma canonica di Jordan, ma solo analizzando la matrice nilpotente, e in particolare i ranghi delle sue potenze. Per fare questo chiamiamo m_i il numero di elementi contenuti nella i-esima colonna della tabella (3.3), con l'indice i che parte da 1 e cresce andando verso destra, e prendiamo in considerazione una qualsiasi matrice A dell'endomorfismo nilpotente avente indice di nilpotenza k.
I valori cercati possono essere ottenuti usando le seguenti uguaglianze:

$$\begin{gathered}
\operatorname{rg}(A^k) = 0 \\
\operatorname{rg}(A^{k-1}) = m_1 \\
\operatorname{rg}(A^{k-2}) = m_1 + m_2 \\
\vdots \\
\operatorname{rg}(A^2) = m_1 + m_2 + \cdots + m_{k-2} \\
\operatorname{rg}(A) = m_1 + m_2 + \cdots + m_{k-2} + m_{k-1} \\
n = m_1 + m_2 + \cdots + m_{k-2} + m_{k-1} + m_k
\end{gathered}$$

dove $n = \dim V$ è l'ordine della matrice A e m_k è la dimensione del nucleo di L.
Con questo procedimento siamo in grado di determinare la forma della tabella (3.3), e quindi i valori cercati.

Dal momento che i valori $m_1, \ldots, m_k$ si trovano analizzando il rango delle potenze di A, e il rango è invariante per similitudine, allora anche questi valori non dipendono dalla particolare matrice A scelta per rappresentare L.
Questo comporta che nella classe di similitudine delle matrici che rappresentano L, ce ne è una sola (a meno della disposizione dei singoli blocchi elementari) in forma canonica di Jordan, nel senso che il numero dei blocchi e il loro tipo è univocamente determinato da L.

Termina così la dimostrazione del teorema 3.1.

3.5 Esempio di riduzione in forma canonica di Jordan

In questo ultimo sottoparagrafo illustriamo un breve esempio di riduzione in forma canonica di Jordan di una matrice di ordine 4.

Consideriamo la matrice A di cui vogliamo trovare la forma canonica di Jordan

$$A = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 3 & 9 & 9 & 0 \\ -2 & -4 & -3 & 0 \\ -8 & -12 & -16 & 1 \end{pmatrix}$$

Il calcolo di det $(A - TI)$ dà il seguente risultato: $P_A(T) = (3-T)^3(1-T)$. Gli autovalori di A sono quindi

$$\lambda_1 = 3 \text{ con } \mu_1 = 3 \quad \text{e} \quad \lambda_2 = 1 \text{ con } \mu_2 = 1$$

Applichiamo il I PASSO suggerito nella dimostrazione, e cerchiamo gli autospazi V_{λ_1} e V_{λ_2}.
$V_{\lambda_1} = ker(A - 3I)^3$; svolgendo il calcolo indicato si ottiene
$V_{\lambda_1} =< (1, 0, 0, -3), (0, 1, 0, -4), (0, 0, 1, -5) >$.
Per calcolare V_{λ_2} sfruttiamo il fatto che $\mu_2 = 1$ e dunque la dimensione dell'autospazio è 1. D'altra parte, analizzando la matrice, si vede facilmente che $(0, 0, 0, 1)$ è un autovettore. Si ha quindi $V_{\lambda_2} =< (0, 0, 0, 1) >$. Abbiamo in questo modo ottenuto una base come indicato nella proposizione 2.3, e quindi possiamo fare un cambio base usando la matrice

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -3 & -4 & -5 & 1 \end{pmatrix} \text{ ottenendo } B = H^{-1}AH = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 3 & 9 & 9 & 0 \\ -2 & -4 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

La matrice di V_{λ_1} è B_1 e quella relativa a V_{λ_2} è B_2, con

$$B_1 = \begin{pmatrix} 3 & 0 & 0 \\ 3 & 9 & 9 \\ -2 & -4 & -3 \end{pmatrix} \quad \text{e} \quad B_2 = (1).$$

B_2 è già in forma canonica di Jordan.

Prendiamo in considerazione B_1 e svolgiamo su questa i passi 2 e 3. Per il secondo passo prendiamo in considerazione la matrice $C = B_1 - 3I$, che evidentemente risulta nilpotente:

$$C = \begin{pmatrix} 0 & 0 & 0 \\ 3 & 6 & 9 \\ -2 & -4 & -6 \end{pmatrix}$$

La matrice C risulta nilpotente di ordine 2 (C^2 risulta essere la matrice nulla di ordine 3). A questo punto si può calcolare la forma canonica di Jordan di questa matrice C anche senza calcolare la base per la quale assume questa forma, usando il rango di C e delle sue potenze:
$\mathrm{rg}\,(C^2) = 0, \quad \mathrm{rg}\,(C) = m_1 = 1, \quad n = 3 = m_1 + m_2 \Rightarrow m_2 = 3 - 1 = 2.$
Si hanno quindi 2 stringhe, una di lunghezza 1 e una di lunghezza 0, perciò alla prima stringa corrisponde un blocco di Jordan di ordine 2 e alla seconda un blocco di ordine 1. La matrice C è quindi simile alla seguente forma canonica di Jordan:

$$J_C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ o equivalentemente a } J'_C = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

La differenza tra J_C e J'_C sta nell'ordine in cui sono messi i blocchi elementari di Jordan.
Se invece risulta necessario trovare la matrice S tale per cui $S^{-1}CS$ è in forma canonica di Jordan, allora per prima cosa calcoliamo il nucleo di C e l'immagine di C e C^2:
$ker(C) =< (-2, 1, 0), (0, 3, -2) >, \quad Im(C) =< (0, 3, -2) >$
$Im(C^2) =< (0, 0, 0) >.$
Per cui si ha $K_1 =< (0, 3, -2) >$ e $K_0 =< (0, 3, -2), (-2, 1, 0) >$. Troviamo adesso la stringa relativa a $v_1 = (0, 3, -2)$, cioè un u_1 tale che $Cu_1 = v_1$; svolgendo i calcoli si ottiene $u_1 = (-1, 1, 0)$. Nel caso del vettore $v_2 = (-2, 1, 0)$ questo è l'unico elemento della sua stringa.
Allora le matrici S e S'

$$S = \begin{pmatrix} 0 & -1 & -2 \\ 3 & 1 & 1 \\ -2 & 0 & 0 \end{pmatrix} \quad S' = \begin{pmatrix} -2 & 0 & -1 \\ 1 & 3 & 1 \\ 0 & -2 & 0 \end{pmatrix}$$

rendono C simile a J_C e J'_C rispettivamente. A questo punto la forma canonica di Jordan di B_1 si può trovare usando la proposizione 3.1 e quindi si ha

$$J_{B_1} = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{per cui} \quad J_B = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Quest'ultima matrice trovata si può ottenere attraverso la similitudine $J_B = Y^{-1}BY$, con Y matrice diagonale a blocchi avente come primo blocco S e come secondo blocco una matrice di ordine 1 avente come elemento 1 (si ricorda che la matrice relativa a V_{λ_2} è già in forma canonica di Jordan e quindi non va modificata):

$$Y = \begin{pmatrix} 0 & -1 & -2 & 0 \\ 3 & 1 & 1 & 0 \\ -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Data quindi la matrice di partenza A, abbiamo che

$$J_B = J_A = Y^{-1}H^{-1}AHY,$$

per cui abbiamo trovato una matrice simile ad A in forma canonica di Jordan, e anche la matrice $K = HY$ di similitudine.

4 Risoluzione di equazioni differenziali

In questo paragrafo vengono trattate le equazioni differenzali, concentrando in particolare l'attenzione sui sistemi di equazioni lineari del primo ordine a coefficienti costanti. Per questi sistemi infatti l'uso delle matrici è naturale e risulterà evidente che la soluzione di questi è semplice se si usano le matrici in forma canonica di Jordan.

Si indicherà con I un intervallo di $\mathbb{R}$.

4.1 Le equazioni differenziali

Definizione 4.1 *Una relazione tra una funzione $y(t) = y$ e le sue derivate $y'(t) = y'$, $y^{(2)}(t) = y^{(2)}, \ldots, y^{(n)}(t) = y^{(n)}$ del tipo*

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + a_{n-2}(t)y^{(n-2)} + \cdots + a_1(t)y' + a_0(t)y = b(t) \quad (4.1)$$

dove $a_{n-1}(t), \ldots, a_0(t)$ e *$b(t)$ sono funzioni date e continue in un intervallo I, si dice* **equazione differenziale lineare di ordine** *n.*

La stessa equazione (4.1) viene detta **ordinaria** perché y(t) è una funzione di una sola variabile.
Se le funzioni $a_i(t)$ sono costanti, l'equazione (4.1) si dice equazione differenziale lineare di ordine n a **coefficienti costanti.**
Se $b(t)$ è la funzione identicamente nulla, la suddetta equazione si dice **omogenea** e l'equazione

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + a_{n-2}(t)y^{(n-2)} + \cdots + a_1(t)y' + a_0(t)y = 0 \quad (4.2)$$

viene detta equazione differenziale **omogenea associata** a (4.1).

Definizione 4.2 *Data l'equazione differenziale (4.1), ogni funzione $\varphi(t)$ che insieme a tutte le sue prime n derivate è continua in I e tale che*

$$\varphi^{(n)} + a_{n-1}(t)\varphi^{(n-1)} + a_{n-2}(t)\varphi^{(n-2)} + \cdots + a_1(t)\varphi' + a_0(t)\varphi = b(t)$$

per ogni $t \in I$, si dice **soluzione (o integrale)** *dell'equazione (4.1).*

Data un'equazione differenziale, risolverla significa trovare tutte le sue soluzioni.
L'insieme di tutte le soluzioni si dice **soluzione generale**, mentre una singola soluzione si dice **soluzione particolare**.

Osservazione 4.13 Ricordando che una funzione si dice di classe $C^i(B)$ se è continua insieme a tutte le sue prime i derivate nell'intevallo B, indicati con $V = C^n(I)$ e $W = C^0(I)$ gli spazi vettoriali su $\mathbb{R}$ che contengono rispettivamente le funzioni di classe C^n e quelle di classe C^0 nell'intervallo I, si può pensare la funzione $\gamma : V \to W$ del tipo

$$\gamma(y) = y^{(n)} + a_{n-1}(t)y^{(n-1)} + a_{n-2}(t)y^{(n-2)} + \cdots + a_1(t)y' + a_0(t)y$$

funzione che risulta essere lineare. Inoltre il nucleo della funzione γ coincide con la soluzione generale dell'equazione omogenea associata a (4.1). $\square$

Ricordando che, data una funzione lineare T, la controimmagine di un vettore appartenente a $Im(T)$ coincide con la somma di un vettore particolare di tale controimmagine con $ker(T)$, si può allora concludere che per trovare la soluzione generale di (4.1) basta sommare a una soluzione particolare $\varphi_0(t)$ la soluzione generale dell'equazione differenziale omogenea associata a (4.1).

Definizione 4.3 *Data l'equazione differenziale (4.1) e assegnati* $t_0 \in I$ *e* $(z_0, \dots, z_{n-1}) \in \mathbb{R}^n$*, il problema di trovare una soluzione* $\varphi(t)$ *di (4.1) che soddisfi le condizioni*

$$\begin{cases} \varphi(t_0) = z_0 \\ \varphi'(t_0) = z_1 \\ \dots \\ \varphi^{(n-1)}(t_0) = z_{n-1} \end{cases} \tag{4.3}$$

viene detto **problema ai valori iniziali** *o* **problema di Cauchy**.
Le condizioni sopra esposte sono dette **condizioni iniziali** *o* **condizioni di Cauchy**.

Il prossimo teorema che viene enunciato risulta valido in generale per le equazioni differenziali lineari, ma la dimostrazione verrà svolta nei sottoparagrafi successivi solo per le equazioni differenziali lineari a coefficienti costanti. La dimostrazione nel caso generale si può trovare nei testi di Analisi Matematica.

Teorema 4.1 *Data l'equazione differenziale lineare*

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + a_{n-2}(t)y^{(n-2)} + \cdots + a_1(t)y' + a_0(t)y = b(t)$$

con $a_i(t)$ *e* $b(t)$ *funzioni continue in* I*, e assegnato* $t_0 \in I$ *e* $(z_0, \dots, z_{n-1}) \in \mathbb{R}^n$*, esiste una sola soluzione* $\varphi(t)$ *dell'equazione differenziale che soddisfa le condizioni iniziali.*

4.2 Equazioni differenziali del primo ordine

Da questo sottoparagrafo tratteremo solo equazioni differenziali a coefficienti costanti, per cui gli $a_i(t)$ che comparivano in (4.1) sono termini costanti. Scriveremo allora semplicemente $a_i(t) = a_i$.

L'argomento di questa parte sono le equazioni differenziali di primo ordine. Vengono trattate inizialmente le equazioni omogenee, quindi quelle del tipo

$$y' = ay \quad \text{con } a \in \mathbb{R}. \tag{4.4}$$

Una soluzione di (4.4) è evidentemente la funzione e^{at}, che è definita e continua su tutto $\mathbb{R}$; come si può facilmente notare anche le funzioni ce^{at}, con $c \in \mathbb{R}$ sono soluzioni.
Con la prossima proposizione si dimostra che ce^{at} sono tutte e sole le soluzioni di (4.4):

Proposizione 4.1 *Lo spazio vettoriale su $\mathbb{R}$ delle funzioni soluzioni di (4.4) è di dimensione uno e la funzione e^{at} ne è una base.*

Dimostrazione: Sia $\varphi(t)$ una soluzione qualsiasi di (4.4), per cui$\varphi'(t) = a\varphi(t)$. Presa la funzione $\psi(t) = \frac{\varphi(t)}{e^{at}} = \varphi(t)e^{-at}$ se si calcola la sua derivata si ottiene

$$\psi'(t) = \varphi'(t)e^{-at} - a\varphi(t)e^{-at} = a\varphi(t)e^{-at} - a\varphi(t)e^{-at} = 0, \forall\, t \in \mathbb{R}$$

Dal momento che $\psi'(t) = 0$ per ogni $t \in \mathbb{R}$, $\psi(t)$ è una funzione costante, $\psi(t) = c$, $c \in \mathbb{R}$, quindi necessariamente $\varphi(t) = ce^{at}$, per cui ogni soluzione è di questo tipo.
Se si cerca l'equazione particolare per cui $\varphi(t_0) = z_0$, con $t_0 \in I$ e $z_0 \in \mathbb{R}$, bisogna determinare il parametro c tenendo presente l'uguaglianza $\varphi(t_0) = ce^{at_0} = z_0$, da cui $c = z_0e^{-at_0}$. Allora l'unica soluzione che soddisfa la condizione iniziale è $\varphi(t) = z_0e^{-at_0}e^{at}$. $\square$

Si consideri adesso l'equazione differenziale non omogenea

$$y' = ay + b(t) \quad \text{con} \quad a \neq 0 \in \mathbb{R} \text{ e } b(t) \in C^0(I) \tag{4.5}$$

e si prenda una qualsiasi primitiva di $e^{-at}b(t)$, per esempio $F(t) = \int_{t_0}^{t} e^{-a\tau}b(\tau)d\tau$. La funzione

$$\varphi_0(t) = e^{at}F(t) \tag{4.6}$$

è una soluzione particolare di (4.5), infatti

$$\varphi_0'(t) = ae^{at}F(t) + e^{at}e^{-at}b(t) = ae^{at}F(t) + b(t) = a\varphi_0(t) + b(t).$$

Per le considerazioni fatte nel sottoparagrafo precedente, e dal momento che la soluzione generale dell'equazione omogenea associata a (4.5) è ce^{at}, la soluzione generale dell'equazione (4.5) è:

$$\varphi(t) = ce^{at} + e^{at}F(t). \tag{4.7}$$

Le soluzioni dipendono quindi dal parametro $c \in \mathbb{R}$, la cui determinazione si ottiene imponendo la condizione iniziale $\varphi(t_0) = z_0$: basta infatti porre $t = t_0$ in (4.7) e risolvere l'equazione in c che ne deriva. La determinazione di c risulta sempre possibile perché il coefficiente che lo moltiplica (e^{at}) risulta diverso da 0 per ogni $t \in \mathbb{R}$ e l'equazione che si ottiene ponendo t uguale a t_0 ha dunque una sola soluzione.
Proprio quest'ultima affermazione giustifica il teorema 4.1 che qui di seguito viene enunciato nel caso particolare di equazioni differenziali lineari di primo ordine a coefficienti costanti:

Teorema 4.2 *Assegnata l'equazione $y' = ay+b(t)$ e dati $t_0 \in I$ e $z_0 \in \mathbb{R}$, esiste una e una sola soluzione $\varphi(t)$ per questa equazione che verifica la condizione iniziale $\varphi(t_0) = z_0$.*

4.3 Sistemi di equazioni differenziali del primo ordine

Nella modellizzazione di sistemi fisici o biologici (per esempio la descrizione del moto di un corpo, o l'assorbimento di un dato farmaco nel corpo) si perviene spesso a relazioni tra più funzioni incognite e le loro derivate, oppure equazioni differenziali lineari di ordine n. Nel primo caso le relazioni possono essere facilmente visualizzate attraverso sistemi di equazioni differenziali del primo ordine, e anche nel secondo caso ci si può ricondurre a sistemi di questo tipo introducendo apposite funzioni ausiliarie, come si vedrà nel seguito.
Lo studio di questi sistemi risulta quindi importante e un modo facile per rappresentare queste relazioni sono le matrici; è in questo sottoparagrafo che risulta evidente l'utilità della forma canonica di Jordan.

Definizione 4.4 *Un sistema del tipo*

$$\begin{cases} y_1' = a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n + b_1(t) \\ y_2' = a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n + b_2(t) \\ \vdots \\ y_n' = a_{n1}y_1 + a_{n2}y_2 + \cdots + a_{nn}y_n + b_n(t) \end{cases} \tag{4.8}$$

con $b_1(t), \dots, b_n(t)$ funzioni date continue nell'intervallo I, e $y_1, \dots, y_n$ con le loro derivate prime funzioni incognite, si dice **sistema di n equazioni differenziali del primo ordine**.

Se tutte le $b_i(t)$ sono identicamente nulle allora il sistema si dice **omogeneo**, *e il sistema (4.8) in cui tutte le $b_i(t)$ sono identicamente nulle si dice sistema* **omogeneo associato** *a (4.8).*

Una **soluzione** di (4.8) è un'ennupla $(\varphi_1(t), \dots, \varphi_n(t))$ di funzioni continue in I che soddisfano contemporaneamente, insieme alle loro derivate prime, tutte le equazioni del sistema (4.8).
Usando la seguente notazione:

$$Y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix}, \quad Y'(t) = \begin{pmatrix} y_1'(t) \\ y_2'(t) \\ \vdots \\ y_n'(t) \end{pmatrix}, \quad B(t) = \begin{pmatrix} b_1(t) \\ b_2(t) \\ \vdots \\ b_n(t) \end{pmatrix}$$

il sistema può essere riscritto

$$Y'(t) = AY(t) + B(t) \tag{4.9}$$

con A indicante la matrice quadrata di ordine n contenente i coefficienti a_{ij} che compaiono nel sistema (4.8).

Se l'equazione differenziale che bisogna risolvere è una equazione differenziale lineare a coefficienti costanti di ordine n, cioè del tipo

$$y^{(n)} + a_{n-1}y^{(n-1)} + a_{n-2}y^{(n-2)} + \cdots + a_1y' + a_0y = b(t) \tag{4.10}$$

questa può essere trasformata in un sistema differenziale di primo ordine con l'uso delle seguenti funzioni:

$$z_1(t) = y(t), \quad z_2(t) = y'(t), \quad z_3(t) = y^{(2)}(t), \; \dots, \; z_n(t) = y^{(n-1)}(t).$$

Si ottiene infatti che l'equazione (4.10) è equivalente al sistema

$$\begin{cases} z_1' = z_2 \\ z_2' = z_3 \\ \vdots \\ z_{n-1}' = z_n \\ z_n' = -a_{n-1}z_n - an - 2z_{n-1} - \cdots - a_1z_2 - a_0z_1 - b(t) \end{cases} \tag{4.11}$$

che può essere scritto nella forma compatta

$$\begin{pmatrix} z_1' \\ z_2' \\ \vdots \\ z_{n-1}' \\ z_n' \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}. \tag{4.12}$$

Quest'ultima scrittura è particolarmente interessante perché la matrice che vi compare è una matrice che si definisce **compagna**, e il calcolo del polinomio caratteristico di questa matrice è immediato, infatti $P_A(T) = (-1)^n(T^n + a_{n-1}T^{n-1} + \cdots + a_1T + a_0)$, risultato che si può ottenere per induzione sull'ordine n della matrice.

Attraverso lo studio dei sistemi differenziali di primo grado si possono quindi risolvere anche le equazioni differenziali lineari di ordine n.

Osservazione 4.14 Come nel caso dell'osservazione 4.1, possiamo definire una funzione $\Gamma : V^n \to W^n$, con V lo spazio vettoriale delle funzioni $C^1(I)$ e W spazio vettoriale delle funzioni $C^0(I)$. La funzione Γ è definita nel modo seguente, $\Gamma(Y) = AY$ ed è una funzione lineare. La soluzione del sistema omogeneo associato a (4.8) coincide con il nucleo di Γ e la soluzione generale del sistema è la controimmagine secondo Γ di $B(t)$. Anche nel caso dei sistemi differenziali lineari, se $\Psi_0(t)$ è una soluzione particolare del sistema di equazioni (4.8), allora la soluzione generale di questo sistema si ottiene sommando a $\Psi_0(t)$ la soluzione generale del sistema omogeneo associato. $\square$

Nella prossima osservazione si cerca il legame esistente tra un sistema (4.8) esprimibile con una matrice quadrata A di tipo qualsiasi e il sistema differenziale in cui la matrice che lo esprime risulta simile ad A ed in forma canonica di Jordan.

Osservazione 4.15 Sia (4.9) l'espressione che esprime un sistema del tipo (4.8). Indicata con $J = S^{-1}AS$ la matrice simile ad A in forma canonica di Jordan (sempre esistente per il teorema 3.1), se si considerano le funzioni $z_1(t), \ldots, z_n(t)$ tali per cui

$$Z(t) = \begin{pmatrix} z_1(t) \\ \vdots \\ z_n(t) \end{pmatrix} = S^{-1}Y(t)$$

si ha anche che

$$Y(t) = SZ(t) \quad \text{e} \quad Y'(t) = SZ'(t)$$

dal momento che la matrice S è formata da elementi costanti.
Sostituendo in (4.9) si ottiene il sistema

$$SZ'(t) = ASZ(t) + B(t) \text{ e quindi}$$

$$Z'(t) = S^{-1}ASZ(t) + S^{-1}B(t) = JZ(t) + D(t), \qquad (4.13)$$

con J matrice in forma canonica di Jordan e $D(t) = S^{-1}B(t)$ una ennupla di funzioni assegnate. In questa nuova forma le singole equazioni differenziali sono particolarmente semplici perché J è una matrice quadrata in forma di Jordan.
Trovata un'ennupla $\Psi_z(t)$ che soddisfa l'equazione (4.13), per trovare l'ennupla di funzioni $\Psi_y(t)$ che soddisfa (4.9) basta moltiplicare $\Psi_z(t)$ per S, cioè $\Psi_y(t) = S\Psi_z(t)$. $\square$

Anche per i sistemi di equazioni differenziali esiste il problema ai valori iniziali, o **problema di Cauchy**; avendo infatti un sistema (4.8), e assegnati $t_0 \in I$ e $(k_1, \ldots, k_n) \in \mathbb{R}^n$, il problema di Cauchy coincide con la ricerca di una soluzione formata da un'ennupla di funzioni $(\phi_1, \ldots, \phi_n)$ che soddisfino le condizioni

$$\begin{cases} \phi_1(t_0) = k_1 \\ \phi_2(t_0) = k_2 \\ \qquad \vdots \\ \phi_n(t_0) = k_n \end{cases} .$$

Se le condizioni iniziali di (4.9) sono date dal vettore colonna Y_0 formato dall'ennupla $(k_1, \ldots, k_n)$, allora le condizioni iniziali del sistema (4.13) sono date dal vettore colonna Z_0 formato dall'ennupla $(k'_1, \ldots, k'_n)$, dove $Z_0 = S^{-1}Y_0$.

Per la risoluzione del generico sistema di equazioni differenziali $Y'(t) = AY(t) + B(t)$ è molto conveniente trovare l'equivalente sistema $Z'(t) = JZ(t)+D(t)$, con J matrice in forma canonica di Jordan. Trovare l'equivalente del sistema di partenza è sempre possibile e la sua costruzione è stata esposta nell'osservazione 4.3; nella stessa osservazione si mostra anche come ottenere la soluzione per il sistema di partenza dopo aver trovato la soluzione per il sistema equivalente.
Cerchiamo quindi la soluzione per sistemi del tipo (4.13); in questi sistemi le equazioni possono essere esclusivamente di due tipi:

$$z'_i = \lambda_i z_i + d_i(t), \tag{4.14}$$

dove è presente una sola funzione incognita, oppure

$$z'_i = \lambda_i z_i + z_{i+1} + d_i(t), \tag{4.15}$$

dove sono presenti due funzioni incognite: z_i e z_{i-1}.
L'ultima equazione del sistema, quella relativa alla funzione z_n è sempre

del tipo (4.14), per come è strutturata la matrice J. Per un'equazione di questo tipo possiamo trovare la soluzione generale seguendo la metodologia esposta nel paragrafo 4.2 (si veda (4.6) per una soluzione particolare e (4.7) per la soluzione generale).
La penultima equazione del sistema può essere sia del tipo (4.14), sia del tipo (4.15); nel primo caso la soluzione si trova con lo stesso metodo usato precedentemente, nell'altro caso, cioè se l'equazione è del tipo $z'_{n-1} = \lambda_{n-1} z_{n-1} + z_n + d_{n-1}(t)$, si sostituista a z_n l'espressione determinata nel passaggio precedente, ottenendo così una equazione differenziale lineare del primo ordine nella sola funzione incognita z_{n-1}, riportandosi così al caso precedente.
Tutte le altre equazioni possono essere risolte applicando il ragionamento appena svolto dal momento che in una matrice in forma di Jordan, per ogni blocco elementare di autovalore λ_i all'ultima riga corrisponde un'equazione del tipo (4.14), e a quelle precedenti corrispondono tutte equazioni del tipo (4.15).
Si ottiene cosi un'ennupla di funzioni $\Psi(t) = (\psi_1(t), \dots, \psi_n(t))$, ennupla che rappresenta la soluzione generale del sistema (4.13).

Volendo entrare in ulteriori dettagli, si svolge nel seguito la risoluzione del sistema omogeneo $Z'(t) = JZ(t)$; si può anche, senza perdita di generalità, supporre che J sia un blocco elementare di Jordan, dal momento che le funzioni coinvolte in un dato blocco elementare non compaiono mai nelle equazioni relative a blocchi diversi, per come è strutturata la forma canonica di Jordan.
Si analizza allora il sistema $Z'(t) = JZ(t)$, con

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

L'ultima equazione risulta $z'_n = \lambda z_n$ e quindi ha come soluzione generale $\psi_n(t) = c_n e^{\lambda t}$.

La penultima equazione diventa così $z'_{n-1} = \lambda z_{n-1} + c_n e^{\lambda t}$, per cui la soluzione generale risulta essere $\psi_{n-1} = c_{n-1} e^{\lambda t} + c_n t e^{\lambda t}$, dal momento che una primitiva $F(t)$ di $e^{-\lambda t} c_n e^{\lambda t} = c_n$ è $c_n t$ (si confronti (4.7)).

L'equazione relativa a z_{n-2} è $z'_{n-2} = \lambda z_{n-2} + c_{n-1} e^{\lambda t} + c_n t e^{\lambda t}$ e come si può notare facilmente si ha che una primitiva $F(t)$ di

$e^{-\lambda t}(c_{n-1}e^{\lambda t} + c_n te^{\lambda t}) = c_{n-1} + c_n t$ risulta $c_{n-1}t + c_n t^2/2$. Quindi la soluzione generale è $\psi_{n-2} = c_{n-2}e^{\lambda t} + c_{n-1}te^{\lambda t} + c_n \frac{t^2}{2}e^{\lambda t}$.

Procedendo con questo ragionamento si ottiene che la soluzione generale relativa all'i-esima equazione è

$$\psi_i(t) = c_i e^{\lambda t} + c_{i+1}te^{\lambda t} + \cdots + c_{i+j}\frac{t^j}{j!}e^{\lambda t} + \cdots + c_{n-1}\frac{t^{n-1-i}}{(n-1-i)!}e^{\lambda t} + c_n\frac{t^{n-i}}{(n-i)!}e^{\lambda t}.$$

La soluzione dell'equazione omogenea $Z'(t) = JZ(t)$ può così essere espressa in maniera compatta:

$$\Psi(t) = e^{\lambda t}\begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{n-2}}{(n-2)!} & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{n-3}}{(n-3)!} & \frac{t^{n-2}}{(n-2)!} \\ \vdots & \ddots & \ddots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & t \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_{n-1} \\ c_n \end{pmatrix}. \tag{4.16}$$

Supponiamo adesso che la matrice J di $Z'(t) = JZ(t)$ sia una matrice quadrata di ordine n formata dai blocchi di Jordan $J_1, \dots, J_r$. Considerato l'i-esimo blocco elementare di Jordan $J_i = J_{\lambda_i,n_i}$, $1 \le i \le r$, la soluzione $\Psi_i(t)$ si ottiene sfruttando la (4.16), con le dovute sostituzioni riguardanti l'autovalore e l'ordine della matrice (quindi $\lambda = \lambda_i$ e $n = n_i$). Chiamata $K_i(t)$ la matrice che si ottiene moltiplicando $e^{\lambda_i t}$ per la matrice quadrata che compare nella soluzione $\Psi_i(t)$, e chiamata C_i la matrice colonna contenente i coefficienti $c_{i1}, \dots, c_{in_i}$, si ha che $\Psi_i(t) = K_i(t)C_i(t)$.

Si ottiene allora che la soluzione generale del sistema può essere scritta nella forma

$$\Psi(t) = \begin{pmatrix} K_1(t) & 0 & \cdots & 0 \\ 0 & K_2(t) & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & K_r(t) \end{pmatrix}\begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ C_r \end{pmatrix}, \tag{4.17}$$

dove $K_i(t)$ e C_i si riferiscono all'i-esimo blocco elementare J_{λ_i,n_i} di ordine n_i che compare in J, $C_i \in M_{n_i,1}(\mathbb{C})$ e $n_1 + \cdots + n_r = n$.

L'ennupla di funzioni $\Psi(t)$ dipende da n parametri che possono essere sempre determinati sfruttando l'ennupla delle condizioni iniziali Z_0 corrispondenti ai valori che le n funzioni devono assumere in t_0. Infatti la matrice

$$K(t) = \begin{pmatrix} K_1(t) & & \\ & \ddots & \\ & & K_r(t) \end{pmatrix}$$

è invertibile per ogni valore di t, dal momento che è sopradiagonale e con elementi diversi da zero sulla diagonale principale. Dalle condizioni $\Psi(t_0) = Z_0 = K(t_0)C$ si può quindi determinare C (e dunque i parametri): risulta infatti $C = (K(t_0))^{-1}Z_0$.
I valori dei parametri sono quindi univocamente determinati, per cui si può affermare che, dato un problema di Cauchy, esiste una e una sola soluzione del sistema $Z'(t) = JZ(t)$ che soddisfa le condizioni iniziali date.

Riferimenti bibliografici

[1] S. Abeasis, *Algebra lineare e geometria*, Zanichelli (1990).

[2] M. Bisiacco, S. Braghetto, *Teoria dei sistemi dinamici*, Esculapio Editrice (2010).

[3] L. Gatto, *Un'introduzione amichevole alla forma canonica di Jordan*, Clut Editrice (1998).

[4] R. Horn, C. Johnson, *Topics in matrix analysis*, Cambridge University Press (1994).

[5] P. Lancaster, M. Tismenetsky, *The Theory of Matrices Second Edition*, Academic Press inc (1985).

[6] B. Noble, *Applied linear Algebra*, Prentice Hall (1985).

[7] J. Ortega, *Matrix Theory a second course*, Plenum Press (1969).

[8] C. Ronconi, *Appunti di geometria*, Univer Editrice (2009).