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

La forma canonica di Jordan:
motivazioni e applicazioni

Laureando: Mattia Montoncelli

Relatore: prof. M. Cristina Ronconi

Anno accademico 2010 - 2011
Padova, 27 settembre 2011

Indice

1 Introduzione **1**

2 Richiami di Algebra Lineare **3**
2.1 Applicazioni lineari e matrici 3
2.2 Autovalori e autovettori di un endomorfismo 6
2.3 Autovalori e autovettori di una matrice 9
2.4 Endomorfisimi diagonalizzabili 9
2.5 Matrici simili 11

3 La forma canonica di Jordan **12**
3.1 Matrice in forma canonica di Jordan 12
3.2 Autovettori generalizzati 13
3.3 Teorema di esistenza della forma canonica di Jordan 16

4 Esponenziale di una matrice **18**
4.1 Nozione di norma 18
4.2 Norma submoltiplicativa 19
4.3 La norma come metrica: convergenza 19
4.4 Esponenziale di matrice 20

5 Applicazioni **22**
5.1 Sistemi dinamici a tempo continuo 22
5.2 Matrice esponenziale e analisi modale 23

1 Introduzione

L'utilizzo di modelli matematici per lo studio dei fenomeni naturali o dei processi tecnologici è divenuto sempre più frequente negli ultimi anni, risultando essenziale quando l'analisi che si effettua deve dare dei risultati quantificabili e quando si vogliono confrontare gli effetti delle diverse strategie di controllo. Tra questi, il più elegante è senz'altro il *modello di stato lineare*, poiché permette di rappresentare un sistema fisico attraverso un sistema di equazioni differenziali lineari a coefficienti costanti, abitualmente riscritto in forma matriciale. Proprio per tali motivi la teoria delle matrici e delle applicazioni lineari trova vastissimo impiego in ogni settore della matematica e dell'ingegneria.

In questo articolo viene analizzata una particolare forma con cui rappresentare un endomorfismo di uno spazio vettoriale: *la forma canonica di Jordan* (dal lavoro dell'omonimo matematico francese Camille Jordan, *Traité des substitutions et des équations algébriques, Parigi, 1870).*

L'algebra lineare mostra come ad ogni endomorfismo L di uno spazio vettoriale V (di dimensione n) sia possibile associare delle matrici quadrate, una per ogni base considerata in V. Tutte queste matrici di L hanno una proprietà: sono simili tra di loro. La relazione di similarità è una relazione di equivalenza e quindi suddivide l'insieme delle matrici quadrate $M_n(\mathbb{C})$ in classi di similitudine, una per ogni endomorfismo L. In altre parole, fissata una di queste classi, le matrici quadrate in essa contenuta rappresentano, rispetto alle varie basi di V, una stessa applicazione lineare, mentre matrici di classi distinte rappresentano applicazioni lineari distinte.
Esistono opportune ipotesi affinché una classe di similitudine relativa ad un endomorfismo L contenga una matrice diagonale; si tratta della matrice associata a una base di autovettori, i cui elementi sulla diagonale sono gli autovalori di L. Qualora essa esista, possiamo sceglierla, vista la sua semplicità, come rappresentante della classe. Purtroppo esistono endomorfismi che non possiedono tale proprietà, e che per questo sono definiti come *non diagonalizzabili.* Tuttavia, estendendo il concetto di autovettore classico a quello di *autovettore generalizzato*, è sempre possibile determinare all'interno di una classe qualunque un elemento di forma "quasi diagonale". Questa matrice, detta forma canonica di Jordan, non è diagonale in senso stretto, ma bensì una matrice diagonale a blocchi, i cui blocchi sono matrici quadrate diagonali o sopradiagonali. La matrice di Jordan è unica nella classe di similitudine a meno di permutazioni dei vari blocchi e assume la forma diagonale nel caso in cui questi ultimi siano tutti di ordine uno.

Volendo proporre altre motivazioni che portano allo studio di tale forma canonica, basti ricordare che nelle applicazioni molti problemi vengono ricon-

dotti alla risoluzione di sistemi di equazioni differenziali lineari a coefficienti costanti.
Per esempio, consideriamo il modello di stato privo di ingresso $\dot{x}(t) = Fx(t)$, dove F è una matrice quadrata di $M_n(\mathbb{C})$ e il vettore $x(t)$ rappresenta lo stato. Il primo obiettivo consiste nel trovare la soluzione generale del sistema di equazioni differenziali e, nel caso si conoscano le condizioni iniziali date dal vettore $x(0)$, la soluzione particolare che le verifica.
In analogia al caso $n = 1$, è naturale indicare la soluzione che verifica il problema di Cauchy con il simbolo $x(t) = e^{Ft}x(0)$. Risulta così evidente la necessità di dare una appropriata definizione dell'esponenziale di una matrice, assicurandosi della relativa convergenza, per poi applicarla alla risoluzione di sistemi di equazioni differenziali.
Se F è diagonale si verifica il cosiddetto *disaccoppiamento delle variabili*, il quale rende immediata la risoluzione dell'intero sistema.
Se F non è diagonale, può essere conveniente effettuare un cambiamento di base del tipo $J = P^{-1}FP$, al fine di renderla in forma diagonale o, nel caso peggiore, in forma canonica di Jordan. Con tale cambiamento il sistema iniziale viene ricondotto a uno del tipo $\dot{y}(t) = Jy(t)$, con $x(t) = Py(t)$. Dalla matrice esponenziale e^{Jt} ad essa legata, si può allora determinare e^{Ft} in base alla relazione $e^{Ft} = Pe^{Jt}P^{-1}$.

La trattazione si snoda lungo quattro paragrafi. Nel primo vengono introdotti alcuni concetti base sull'algebra lineare, utili per la comprensione delle successive pagine. Il secondo paragrafo è incentrato sulla dimostrazione dell'esistenza, nella classe di similitudine di una qualsiasi matrice $A \in M_n(\mathbb{C})$, di una matrice in forma canonica di Jordan. La dimostrazione presentata è dovuta al matematico russo A.F.Filippov, che la tratta nella sua pubblicazione [2]. Essa è stata scelta per la sua particolare eleganza e semplicità, sebbene durante il suo svolgimento non presenti la costruzione della matrice di similitudine, perché condotta per induzione sull'ordine n di A. Nel paragrafo stesso vengono comunque esposte le linee guida sia per la determinazione di J, sia per quella della relativa matrice di similitudine. Nel paragrafo tre si parla invece dell'esponenenziale di matrice, strumento necessario per calcolare la soluzione dei sistemi dinamici presentati nell'ultimo paragrafo.

Nella stesura si sono seguite principalmente le linee guida presentate da [6], mentre per il paragrafo finale si è fatto riferimento a [1]. Utili suggerimenti sono stati inoltre tratti da [5] e [7].

2 Richiami di Algebra Lineare

In questo prima sezione ci occuperemo di elencare una serie di concetti e teoremi di base dell' Algebra Lineare che saranno poi utilizzati nei capitoli successivi. Tra le dimostrazioni vengono presentate solo quello più significative o che possono essere utili nel seguito, mentre per le rimanenti si rimanda ai testi di Algebra lineare. Daremo inizialmente la definizione di *applicazione lineare* tra due spazi vettoriali, evidenziando la relazione con le matrici. In seguito, dopo aver dato la definizione di *autovalori* e *autovettori* mostreremo come determinare le basi del dominio e del codominio di una funzione lineare tali da rendere la matrice ad esse relativa di forma particolarmente semplice.

2.1 Applicazioni lineari e matrici

Se V e V' sono due spazi vettoriali, un'*applicazione lineare* $L : V \to V'$ non è nient'altro che una funzione particolare tra elementi di V e elementi di V'. V è chiamato *spazio dominio* di L, mentre V' è il *codominio.*

Definizione 2.1. *Dati V e V' due spazi vettoriali su uno stesso campo $\mathbb{K}$, una applicazione (o funzione)*

$$L : \quad V \to V'$$

si dice 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 quelle definite su V'.

Osservazione 2.1. Sia $L : V \to V'$ una funzione tra due spazi vettoriali V e V' su uno stesso campo $\mathbb{K}$. Condizione necessaria e sufficiente affiné L sia lineare è che:

$$L(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 L(v_1) + \cdots + \alpha_n L(v_n).$$

□

Le applicazioni lineari vengono anche dette *omomorfismi* oppure *trasformazioni lineari.*
Una applicazione biietiiva $L : V \to V'$ si dice anche *isomorfismo*, mentre ogni applicazione di V in V viene chiamata *endomorfismo*; in particolare un isomorfismo di V in V si dice *automorfismo.* Diamo ora la definizione di *nucleo* e *immagine* di un'applicazione lineare:

Definizione 2.2. *Sia* $L : V \to V'$ *una applicazione lineare. L'insieme* $L(V) = \{v' \in V' | v' = L(v)$ *per qualche* $v \in V\} \subseteq V'$ *viene denominato* immagine *di* L *e indicato con il simbolo* $I_m(L)$. *Come si riconosce facilmente,* $I_m(L)$ *è un sottospazio vettoriale di* V' *e la sua dimensione viene detta* rango *di* L *ed indicata con* $\rho(L)$.
L'insieme $\{v \in V | L(v) = 0_{v'}\} \subseteq V$ *viene detto* nucleo *di* L *o* spazio nullo *di* L *e viene indicato con il simbolo* $N(L)$. *Si può verificare semplicemente che* $N(L)$ *risulta essere un sottospazio vettoriale di* V.

Proposizione 2.1. *Sia* $L : V \to V'$ *un'applicazione lineare.*

$$L \text{ è iniettiva} \Longleftrightarrow N(L) = \{0_V\}. \tag{1}$$

Osservazione 2.2. Sia $L : V \to V'$ un'applicazione lineare.

1. Se $W =< v_1, \ldots, v_r >\subset V$, allora $L(W) =< L(v_1), \ldots, L(v_r) >$.

2. Se $V =< v_1, \ldots, v_n >\subset V$, allora $L(V) = I_m(L) =< L(v_1), \ldots, L(v_n) >$.

In questo modo risulta molto semplice determinare l'immagine di uno spazio vettoriale noti i suoi generatori o eventualmente una sua base. □

Nelle applicazioni è spesso utile vedere se esiste, ed eventualmente determinarla, una funzione lineare che verifica assegnate condizioni, cioè che a prefissati vettori del dominio fa corrispondere assegnati vettori del codominio. Il teorema seguente fornisce una risposta nel caso particolare in cui i vettori assegnati nel dominio formino una base. Il problema generale può poi essere ricondotto a questo caso con opportune osservazioni.

Teorema 2.1. *Siano* V *e* V' *due spazi vettoriali su uno stesso campo* $\mathbb{K}$. *Sia inoltre* $[v_1, \ldots, v_n]$ *una base di* V *e* $[w_{1'}, \ldots, w_{n'}]$ *una famiglia di vettori di* V'. *Allora esiste una e una sola applicazione lineare* $L : V \to V'$ *tale che*

$$L(v_1) = w'_1, \ldots, L(v_n) = w'_n.$$

Inoltre

1. L *è iniettiva* $\Longleftrightarrow [w_{1'}, \ldots, w_{n'}]$ *è linearmente indipendente;*

2. L *è suriettiva* $\Longleftrightarrow [w_{1'}, \ldots, w_{n'}]$ *genera* V'.

Corollario 2.1. *Sia* $L : V \to V'$ *un'applicazione lineare e sia* $[v_1, \ldots, v_n]$ *una base di* V. *Allora*

1. L *è iniettiva* $\Longleftrightarrow [L(v_1), \ldots, L(v_n)]$ *è una base di* $I_m(L)$.

2. L *è un isomorfismo* $\Longleftrightarrow [L(v_1), \ldots, L(v_n)]$ *è una base di* V'.

Un risultato estremamente utile è che la somma del rango di un'applicazione lineare e della dimensione del suo nucleo è sempre uguale alla dimensione dello spazio dominio.

Teorema 2.2 (delle dimensioni)**.** *Sia $L : V \to V'$ un'applicazione lineare. Allora:*

$$\dim V = \dim I_m(L) + \dim N(L).$$

Dimostrazione. Sia $n = \dim V$ e $r = \dim N(L)$. Bisogna dimostrare che $\dim I_m(L) = n - r$.
Se $r = 0$, cioè se L è iniettiva, la tesi è una conseguenza del corollario 2.1 - 1. Se $r \neq 0$, si consideri una base $[v_1, \dots, v_r]$ di $N(L)$ e la si estenda (utilizzando il teorema della base incompleta) a una base di V, ottenendo così: $B = [w_1, \dots, w_r, v_{r+1}, \dots, v_n]$. Allora, dall'osservazione 2.2 - 2 $[L(w_1), \dots, L(w_r), L(v_{r+1}), \dots, L(v_n)]$ genera $I_m(L)$. Ricordando poi che $w_1, \dots, w_r \in N(L)$, si ha:

$$< L(v_{r+1}), \dots, L(v_n) >= I_m(L).$$

Passiamo ora a dimostrare che tali generatori sono linearmente indipendenti. Sia infatti:

$$b_{r+1}L(v_{r+1}) + \ldots + b_n L(v_n) = 0_{V'}.$$

Allora $L(b_{r+1}v_{r+1} + \cdots + b_n v_n) = 0_{V'}$ e quindi $b_{r+1}v_{r+1} + \cdots + b_n v_n \in N(L)$, cioè si può scrivere come combinazione lineare di $w_1, \cdots, w_r$:

$$b_{r+1}v_{r+1} + \cdots + b_n v_n = a_1 v_1 + \cdots + a_r v_r$$

Poiché $[w_1, \dots, w_r, v_{r+1}, \dots, v_n]$ è una base di V, dalla precedente uguaglianza si deduce: $a_1 = \cdots = a_r = 0 = b_{r+1} = \cdots = b_n$. Pertanto $[L(v_{r+1}), \dots, L(v_n)]$ è linearmente indipendente e quindi una base di $I_m(L)$. Allora $\dim I_m(L) = n - r$.

□

Consideriamo un'applicazione lineare $L : V \to V'$ tra due spazi vettoriali V e V' su uno stesso campo $\mathbb{K}$ e siano B e B' basi di V e V' rispettivamente. Cerchiamo con la successiva proposizione di determinare la relazione che intercorre tra l'ennupla delle coordinate di un vettore $v \in V$ (rispetto alla base B) e quella delle coordinate di $L(v)$ (rispetto alla base V').

Proposizione 2.2. *Sia $L : V \to V'$ una applicazione lineare tra due spazi vettoriali V e V' (su uno stesso campo $\mathbb{K}$) di dimensioni n e m rispettivamente. Siano inoltre B e B' basi di V e V'. Indichiamo con*

$$X_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \textit{e con} \quad X'_{B'} = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

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

$$X'_{B'} = AX_B$$

La matrice A è univocamente determinata da L, B e B'.

Con l'osservazione successiva vedremo che una matrice A qualsiasi rappresenta sempre una applicazione lineare tra due spazi vettoriali V e V' fissati.

Osservazione 2.3. Siano V e V' due spazi vettoriali su $\mathbb{K}$ e siano $B = [v_1, \ldots, v_n]$ e $B' = [v'_1, \ldots, v'_n]$ due loro basi.
Comunque si fissi $A \in M_{m,n}(\mathbb{K})$ si può considerare la funzione $L_A : V \to V'$ tra V e V' che al vettore $v = x_1v_1 + \cdots + x_nv_n$ fa corrispondere il vettore $v' = x'_1v'_1 + \cdots + x'_nv'_n$ le cui coordinate rispetto a B', sono date da:

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

Tale funzione risulta essere lineare. In particolare, se V coincide con K^n, V' con K^m e B e B' sono le basi canoniche di K^n e K^m rispettivamente, la funzione L_A viene detta **funzione lineare associata in modo canonico ad** A. Il nucleo di L_A, viene anche detto *spazio nullo* di A e indicato con $Null(A)$. □

Si vuole ora analizzare la relazione che intercorre tra le varie matrici quadrate di uno stesso endomorfismo:

Osservazione 2.4. Consideriamo un endomorfismo $L : V \to V$, siano A e D le matrici di L relative alle basi B e C rispettivamente (si ricorda che nel caso di endomorfisimi, si usa scegliere la stessa base nel dominio e nel codominio). Le due matrici dell' endomorfismo L sono legate da una relazione del tipo:

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

dove H e H^{-1} sono le due matrici di passaggio tra le basi B e C dello spazio vettoriale V. Un risultato analogo si ottiene anche per un qualsiasi *omomorfismo*.
□

2.2 Autovalori e autovettori di un endomorfismo

La teoria degli autovettori permette di determinare delle basi del dominio e del codominio di una applicazione lineare $L : V \to V'$ in grado di rendere la matrice di L ad esse relative di forma particolarmente semplice (per esempio diagonale). In questa sezione definiremo gli strumenti necessari al processo di *diagonalizzazione*.

Definizione 2.3. *Sia $L : V \to V$ un endomorifsmo di V, dove V è uno spazio vettoriale su $\mathbb{K}$. 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 2.4. *$\lambda \in \mathbb{K}$ si dice **autovalore** di L se esiste un vettore $v \in V$, $v \neq 0_V$, tale che*

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

Fissato $\lambda \in \mathbb{K}$, tutti i vettori non nulli per cui $L(v) = \lambda v$ vengono detti autovettori associati all'autovalore λ.

Cerchiamo ora, con la successiva osservazione, di intravedere il concetto di diagonalizzazione:

Osservazione 2.5. Consideriamo un endomorfismo L di V, sia $B = [v_1, \ldots, v_n]$ una base di V e sia A la matrice di L ad essa relativa. Se i primi r vettori $v_1, \ldots, v_r$ di B sono autovettori di L relativi agli autovalori $\lambda_1, \ldots, \lambda_r$, la matrice A di L relativa alla base B è del tipo:

$$\left(\begin{array}{cccc|c} \lambda_1 & 0 & \ldots & 0 & \\ 0 & \lambda_2 & \ldots & 0 & \\ \ldots & \ldots & \ldots & \ldots & C \\ 0 & 0 & \ldots & \lambda_r & \\ \hline & O & & & D \end{array}\right).$$

□

Definizione 2.5. *Fissato $\lambda \in \mathbb{K}$, l'insieme $V_\lambda = \{v \in V | L(v) = \lambda v\}$, che è un sottospazio di V, si dice **autospazio** di L relativo all'autovalore λ.*

Un metodo per riconoscere gli autovalori di un dato endomorfismo è illustrato nella proposizione successiva.

Proposizione 2.3. *Siano: L un endomorfismo di uno spazio vettoriale V su $\mathbb{K}$, B una base di V e $A \in M_n(\mathbb{K})$ la matrice di L relativa a B. Allora:*

$$\lambda \in \mathbb{K} \text{ è un autovalore di } L \iff det(A - \lambda I_n) = 0,$$

oppure, analogamente,

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

Sia L un endomorfismo e sia A la matrice di L relativa a una base B di V. Detta T una indeterminita su $\mathbb{K}$, si consideri il polinomio

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

La proposizione 2.3 può essere enunciata anche nel modo seguente:

$$\lambda \in \mathbb{K} \text{ è un autovalore di L } \iff \lambda \text{ è una radice del polinomio } P_L(T).$$

Poichè si può dimostrare che tale polinomio non dipende dalla matrice A, cioè dalla base scelta, le sue radici individuano tutti gli autovalori di L. Diamo un nome a questo polinomio:

Definizione 2.6. *Sia L un endomorfismo di V. Il polinomio di grado n*

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

*dove A è la matrice di L relativa a una base di V scelta arbitrariamente, si dice **polinomio caratteristico** di L. L'equazione*

$$P_L(T) = 0.$$

*si dice **equazione caratteristica** di L.*

Definizione 2.7. *Sia λ un autovalore di L, si dice **molteplicità algebrica** dell'autovalore λ, e si indica con $\mu(\lambda)$, la molteplicità di λ come radice del polinomio caratteristico $P_L(T)$.*
*La dimensione dell'autospazio V_λ viene invece detta **molteplicità geometrica** di λ.*

Osservazione 2.6. Dato un endomorfismo $L : V \to V$, poiché il grado del polinomio caratteristico $P_L(T)$ è pari alla dimensione di V, la somma delle molteplicità degli autovalori di L non supera dim V:

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

Nella relazione precedente vale l'uguaglianza solo se il polinomio $P_L(T)$ è scomponibile in fattori di primo grado. Questo è sempre possibile se il campo $\mathbb{K}$ degli scalari coincide con il campo $\mathbb{C}$ dei numeri complessi, mentre altrettanto non si può dire, ad esempio, se $\mathbb{K} = \mathbb{R}$. $\square$

Proposizione 2.4. *Sia λ un autovalore di L. Allora*

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

Proposizione 2.5. *Sia L un endomorfismo di V. Comunque fissati degli autovalori $\lambda_1, \ldots, \lambda_r$, con $\lambda_i \neq \lambda_j$ per $i \neq j$, la somma $V_{\lambda_1} + \cdots + V_{\lambda_r}$ è diretta:*

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

Corollario 2.2. *Siano $L, \lambda_1, \ldots, \lambda_r$ come nella proposizione 2.5 e siano $v_1, \ldots, v_r$ autovettori associati agli autovalori $\lambda_1, \ldots, \lambda_r$. Allora $v_1, \ldots, v_r$ sono linearmente indipendenti.*

La famiglia ottenuta riunendo basi dei singoli autospazi è, per la proposizione 2.5, una famiglia linearmente indipendente formata da autovettori. Il numero massimo di autovettori di L linearmente indipendenti coincide con la dimensione dello spazio somma di tutti i vari autospazi: $\dim(V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}) = \dim V_{\lambda_1} + \cdots + \dim V_{\lambda_r}$ per le proprietà delle somme dirette.
È possibile dunque ottenere una base di V composta interamente da autovettori di L solo se il polinomio caratteristico può essere scomposto in fattori di grado uno e se la dimensione di un autospazio coincide con la molteplicità algebrica del corrispondente autovalore.

2.3 Autovalori e autovettori di una matrice

Definizione 2.8. *Data una matrice $A \in M_n(\mathbb{K})$, si dice* ***autovettore*** *di A ogni ennupla non nulla $X \in K^n$ (pensata come matrice colonna) tale che*

$$AX = \lambda X$$

per qualche $\lambda \in K$. $\lambda \in K$ si dice ***autovalore*** *di A se esiste un'ennupla $X \neq O$ tale che*

$$AX = \lambda X.$$

Osservazione 2.7. Come già sappiamo dall'osservazione 2.3, ad ogni matrice $A \in M_n(\mathbb{K})$ resta associato in modo canonico un endomorfismo L_A di K^n. In questo modo gli autovettori, gli autovalori, gli autospazi, il polinomio caratteristico di A sono gli autovettori, gli autovalori, gli autospazi, il polinomio caratteristico dell'endomorfismo L_A.
Con questa semplice osservazione possiamo evitare di formulare nuovi teoremi o nuove proposizioni sugli autovalori di una matrice, in quanto basta utilizzare quelli del paragrafo precedente sugli autovalori di un endomorfismo. $\square$

2.4 Endomorfisimi diagonalizzabili

Lo scopo di questo paragrafo è quello di individuare le condizioni necessarie e sufficienti per stabilire se un dato endomorfismo è diagonalizzabile.

Definizione 2.9. *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 relativa ad essa sia* ***diagonale****.*

Proposizione 2.6. *Sia L un endomorfismo di uno spazio vettoriale V. Sono equivalenti:*

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

La semplice dimostrazione si basa sull'osservazione 2.5.

Definizione 2.10. *Una matrice $A \in M_n(\mathbb{K})$ si dice* ***diagonalizzabile*** *se esiste una matrice $H \in M_n(\mathbb{K})$ invertibile tale che*

$$\Delta = H^{-1}AH$$

sia ***diagonale****.*

Osservazione 2.8. Indicato con L_A l'endomorfismo di K^n associato ad A in modo canonico,

$$A \text{ risulta diagonalizzabile se e solo se lo è } L_A.$$

Supponiamo A diagonalizzabile; esiste allora, per definizione, una matrice invertibile H per cui $\Delta = H^{-1}AH$ è diagonale. Indicata con B la base di K^n formata dalle colonne di H, la matrice di L_A, relativa a B, coincide con Δ. L_A pertanto risulta diagonalizzabile.
Viceversa, se L_A è diagonalizzabile esiste una base C di K^n per cui la matrice D ad essa relativa è diagonale. poiché d'altra parte $D = K^{-1}AK$, con K matrice di passaggio tra la base C e la base canonica di K^n, si deduce che A è diagonalizzabile. $\square$

Abbiamo così dimostrato l'analogia tra la condizione di diagonalizzabilità di un endomorfismo e quella di una matrice. Si ottiene, ad esempio, il seguente corollario alla proposizione 2.6:

Corollario 2.3. *Sia $A \in M_n(\mathbb{K})$. Sono equivalenti:*

1. *A è diagonalizzabile;*
2. *esiste una base di K^n formata da autovettori di A.*

Proponiamo ora delle condizioni **necessarie e sufficienti** affinché una matrice o un endomorfismo risultino diagonalizzabili.

Proposizione 2.7. *Sia L un endomorfismo di uno spazio vettoriale V di dimensione n su $\mathbb{K}$. Siano inoltre $\lambda_1, \ldots, \lambda_r$ gli autovalori di L, $\mu(\lambda_1), \ldots, \mu(\lambda_r)$ le rispettive molteplicità e $V_{\lambda_1}, \ldots, V_{\lambda_r}$ gli autospazi associati agli autovalori. Sono equivalenti:*

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

Corollario 2.4. *Sia $A \in M_n(\mathbb{K})$, $\lambda_1, \ldots, \lambda_r$ gli autovalori di A, $\mu(\lambda_1), \ldots, \mu(\lambda_r)$ le rispettive molteplicità e $V_{\lambda_1}, \ldots, V_{\lambda_r}$ gli autospazi associati agli autovalori.*
Sono equivalenti:

1. *A è diagonalizzabile;*

2. *$K^n = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$;*

3. *$\sum_{i=1}^{r} \mu(\lambda_i) = n$ e $\mathrm{dim} V_{\lambda_i} = \mu(\lambda_i)$, per ogni $1 \leq i \leq r$.*

Vediamo come, data una matrice A diagonalizzabile, è possibile costruire una matrice invertibile H che la diagonalizza, cioé tale che $\Delta = H^{-1}AH$ sia diagonale.

Proposizione 2.8. *Se $A \in M_n(\mathbb{K})$ è diagonalizzabile e se Γ è una base di K^n formata da autovettori di A, una delle matrici che diagonalizzano A si ottiene ponendo lungo le sue colonne i vettori della base Γ.*

2.5 Matrici simili

Definizione 2.11. *Assegnate $A, B \in M_n(\mathbb{K})$, si dice che B è **simile** a A, e si scrive $A \sim B$, se esiste una matrice invertibile $H \in M_n(\mathbb{K})$ tale che*

$$B = H^{-1}AH. \tag{2}$$

*La relazione (2) si dice **trasformazione per similitudine**.*

Proposizione 2.9. *La relazione in $M_n(\mathbb{K})$: $R = \{(A; B) \mid A \sim B\}$ è una* relazione di equivalenza *su $M_n(\mathbb{K})$.*

Come una qualunque relazione di equivalenza, la relazione di similitudine partiziona lo spazio vettoriale $M_n(\mathbb{K})$ in classi di equivalenza disgiunte. Ogni classe di equivalenza è l'insieme di tutte le matrici in $M_n(\mathbb{K})$ simili a una matrice rappresentante della classe. Tutte le matrici in una classe di equivalenza sono simili, mentre due matrici di due classi differenti non lo sono.

Osservazione 2.9. *Dall'osservazione 2.4 del paragrafo 2.1 si deduce che le matrici di uno stesso endomorfismo relative a due basi diverse sono simili.*

Proposizione 2.10. *Matrici simili si possono interpretare come matrici di uno stesso endomorfismo (relativamente a basi diverse).*

L'osservazione 2.9 e la proposizione 2.10 garantiscono che due matrici sono simili se e solo se sono matrici di uno stesso endomorfismo.

Osservazione 2.10. *Due matrici simili sono o **entrambe diagonalizzabili** o **entrambe non diagonalizzabili** perché la relazione di similitudine è una relazione di equivalenza.*

3 La forma canonica di Jordan

Nella sezione precedente si è visto come ad ogni endomorfismo di uno spazio vettoriale V sia possibile associare delle matrici, una per ogni base di V. Tutte le matrici di un dato endomorfismo risultano simili tra loro e viceversa, matrici simili rappresentano, rispetto a basi diverse, lo stesso endomorfismo.
La relazione di similitudine è una relazione di equivalenza; per questo l'insieme delle matrici quadrate di un certo ordine può essere ripartito nelle varie classi di similitudine, una per ogni endomorfismo.
Sono state enunciate in 2.4 delle condizioni necessarie e sufficienti affinché una classe di similitudine contenga una matrice diagonale. Purtroppo non tutte le classi ammettono tale proprietà; è sempre tuttavia possibile determinare in ciascuna di esse, nel caso in cui il campo degli scalari sia il campo $\mathbb{C}$ dei numeri complessi, una matrice particolare di forma "quasi diagonale", detta *matrice di Jordan* o *forma canonica di Jordan.* In questo paragrafo e nei successivi si supporrà allora $\mathbb{K} = \mathbb{C}$.
All'interno della sezione mostreremo una tecnica per ottenere la matrice di similitudine, dimostrando poi per induzione l'esistenza della forma canononica.

3.1 Matrice in forma canonica di Jordan

Una matrice in *forma canonica di Jordan* è una matrice diagonale a blocchi, in cui ogni blocco è un *blocco di Jordan* relativo ad un particolare autovalore della matrice.
Nel seguito, la matrice quadrata nulla, ossia la matrice i cui coefficienti sono tutti zero, verrà denotata con il simbolo $\boldsymbol{O}$.
Iniziamo dando la definizione di *matrice nilpotente*:

Definizione 3.1. *Sia $A \in M_n(\mathbb{C})$ una matrice quadrata di ordine n. Essa si dice* nilpotente *se esiste un $m > 0$ tale che $A^m = \boldsymbol{O}$. Il più piccolo intero $m_0 > 0$ per cui $A^{m_0} = \boldsymbol{O}$ viene detto* grado di nilpotenza *di A.*

Definizione 3.2. *Sia m un intero maggiore di 0. La matrice quadrata $m \times m$:*

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

si dice blocco nilpotente elementare.

Proposizione 3.1. *La matrice N_m è nilpotente di ordine m.*

Ora che si dispone della definizione di matrice nilpotente è possibile definire il *blocco elementare di Jordan*:

Definizione 3.3. *Siano r un intero postivo e $\lambda \in \mathbb{C}$. Una matrice quadrata $A \in M_r(\mathbb{C})$ si dice* blocco elementare di Jordan *di ordine r relativamente a $\lambda \in \mathbb{C}$ se $A = J(\lambda)$ dove la notazione $J(\lambda)$ è definita dalla seguente uguaglianza:*

$$J(\lambda) = \lambda I_r + N_r,$$

essendo N_r una matrice nilpotente elementare.

Osservazione 3.1. Un blocco elementare di Jordan $J(\lambda)$ è una matrice triangolare superiore della forma:

$$J(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix} \tag{3}$$

e, quindi, il suo polinomio caratteristico è $(\lambda - T)^r$, da cui segue che $\lambda \in \mathbb{C}$ è l'unico autovalore di $J(\lambda)$. La molteplicità algebrica di λ è r, mentre la molteplicità geometrica è 1, perché il rango di $J(\lambda) - \lambda I_r$ è $r - 1$.
Si osservi inoltre che se $\lambda = 0$ allora:

$$J(0) = N_r.$$

□

Definizione 3.4. *Una matrice quadrata diagonale a blocchi del tipo:*

$$J = \begin{pmatrix} J_1(\lambda_1) & 0 & \dots & 0 & 0 \\ 0 & J_2(\lambda_2) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & J_{p-1}(\lambda_{p-1}) & 0 \\ 0 & 0 & \dots & 0 & J_p(\lambda_p) \end{pmatrix} \tag{4}$$

dove i blocchi $J_i(\lambda_i)$ sono blocchi di Jordan, viene detta matrice di Jordan *o* matrice in forma canonica di Jordan.

Osservazione 3.2. *Nella definizione precedente gli autovalori λ_i e λ_j relativi ai blocchi $J_i(\lambda_i)$ e $J_j(\lambda_j)$ possono anche coincidere.*
Si può osservare inoltre che, nel caso tutti i blocchi $J_i(\lambda_i)$ siano di ordine uno, la matrice J risulta diagonale.

3.2 Autovettori generalizzati

Prima di dimostrare il teorema di esistenza di una matrice in forma canonica di Jordan in ciascuna classe di similitudine, discuteremo di alcune implicazioni di base del teorema. I blocchi di Jordan $J_i(\lambda_i)$ sono matrici $r_i \times r_i$, dove l'ordine

r_i è un intero compreso tra 1 e n tale per cui $\sum_{i=1}^{p} r_i = n$. Si ottiene una matrice diagonale quando $p = n$ e $r_i = 1, i = 1, \ldots, n$, mentre nel caso $p = 1$ e $r_1 = n$ si ottiene un blocco di Jordan di ordine n. Tutte le possibilità tra questi due casi estremi sono possibili.
Abbiamo osservato precedentemente che la matrice in forma (3) ha un solo autovettore, ovvero $\mathbf{e_1}$, il vettore con 1 nella prima posizione e 0 nelle altre. Quindi, i vettori che hanno 1 nelle posizioni $1, r_1+1, r_1+r_2+1, r_1+\cdots+r_{p-1}+1$ rispettivamente e zero altrove sono autovettori linearmente indipendenti di J. Come conseguenza:

Proposizione 3.2. *La molteplicità geometrica di un autovalore λ di J è uguale al numero di blocchi di Jordan associati a λ.*

Sia ora A una matrice quadrata di ordine r simile ad una matrice J in forma canonica di Jordan e sia P una matrice invertibile tale che $P^{-1}AP = J$. Allora

$$AP = PJ. \tag{5}$$

Analizzando le colonne di P, si vede che quelle nelle posizioni $1, r_1+1, r_1+r_2+1, r_1+\cdots+r_{p-1}+1$ sono autovettori di A, linearmente indipendenti poiché P è non singolare.
I vettori (linearmente indipendenti) delle altre colonne di P sono chiamati *autovettori generalizzati* di A. Cerchiamo di entrare più in dettaglio per analizzarne le proprietà. Confrontando le prime r_1 colonne del primo e del secondo membro di (5), relative al primo blocco di Jordan $J_1(\lambda_1)$, e ricordando le proprietà del prodotto tra matrici si ottiene:

$$A\mathbf{p_1} = \lambda_1 \mathbf{p_1}, \qquad A\mathbf{p_i} = \lambda_1 \mathbf{p_i} + \mathbf{p_{i-1}}, \qquad i = 2, \ldots, r_1. \tag{6}$$

I vettori $\mathbf{p_1}, \ldots, \mathbf{p_{r_1}}$ formano una *catena di Jordan* di lunghezza r_1. Ogni $\mathbf{p_i}$, $i \geq 2$, viene detto autovettore generalizzato di *ordine* (o *grado*) i e $\mathbf{p_{r_1}}$ viene detto *direttore* della catena.

Osservazione 3.3. *Si nota da (6) che gli autovettori generalizzati $\mathbf{p_2}, \ldots, \mathbf{p_{r_1}}$ non soddisfano l'equazione $A\mathbf{p} = \lambda_1 \mathbf{p}$, ma qualcosa di molto simile ad essa.*

Ora possiamo riscrivere (6) nella forma

$$(A - \lambda_1 I)\mathbf{p_1} = \boldsymbol{O}, \qquad (A - \lambda_1 I)\mathbf{p_i} = \mathbf{p_{i-1}}, \qquad i = 2, \ldots, r_1 \tag{7}$$

la quale mostra chiaramente che $A - \lambda_1 I$ mappa $\mathbf{p_i}$ nel successivo elemento $\mathbf{p_{i-1}}$ della catena. In particolare l'elemento direttore $\mathbf{p_{r_1}}$ è in grado di generare l'intera catena attraverso ripetute moltiplicazioni per $A - \lambda_1 I$. Lo stesso ragionamento vale per gli altri autovalori di A:

Proposizione 3.3. *Se A è una matrice $n \times n$ tale per cui $A = PJP^{-1}$, dove J è in forma di Jordan (4), allora le colonne di P formano p catene di Jordan*

$$\{\mathbf{p_1}, \ldots, \mathbf{p_{r_1}}\}, \{\mathbf{p_{r_1+1}}, \ldots, \mathbf{p_{r_1+r_2}}\}, \ldots, \{\mathbf{p_{n-r_p+1}}, \ldots, \mathbf{p_n}\} \tag{8}$$

relative rispettivamente a $\lambda_1, \lambda_2, \ldots, \lambda_p$.

La relazione (7) implica poi che

$$(A - \lambda_1 I)^i \mathbf{p_j} = \boldsymbol{O}, \qquad i = 1, \ldots, r_1,\ j = 1, \ldots, i \tag{9}$$

Nel caso λ_1 sia distinto dagli altri autovalori $\lambda_2, \ldots, \lambda_p$, le formule (7) possono essere usate, almeno in via teorica per realizzare la prima catena di Jordan relativa all'autovalore λ_1.
Partendo dall'autovettore $\mathbf{p_1}$ infatti possiamo costruire un autovettore generalizzato $\mathbf{p_2}$ in grado di verificare l'uguaglianza $(A - \lambda_1 I_{r_1})\mathbf{p_2} = \mathbf{p_1}$ e così via. In molti casi particolari questa strada può essere seguita, anche se il calcolo risulta di regola complicato. In generale però questa tecnica non produce alcun risultato.
Volendo invece seguire una strada che consenta, qualunque sia la matrice, di costruire la catena $\{\mathbf{p_1}, \ldots, \mathbf{p_{r_1}}\}$, basta considerare i sottospazi W_i di $\mathbb{C}^n$ definiti da:

$$W_i = Null(A - \lambda_i I)^i,$$

ovvero dalle soluzioni del sistema omogeneo $(A - \lambda_i I)^i X = \boldsymbol{O}$.
Risulta: $W_1 =< \mathbf{p_1} >,\ W_2 =< \mathbf{p_1}, \mathbf{p_2} >,\ \ldots,\ W_{r_1} =< \mathbf{p_1}, \ldots, \mathbf{p_{r_1}} >$.
Ovviamente $W_i \subset W_{i+1}$ dove l'inclusione è stretta perché $\mathbf{p_i} \notin W_{i-1}$:

$$W_i \subsetneqq W_2 \subsetneqq \cdots \subsetneqq W_{r_1}.$$

Basta allora considerare per primo l'autovettore generalizzato $\mathbf{p_{r_1}} \in W_{r_1}$ non appartenente all'autospazio $W_{r_1 - 1}$, per poi determinare via via il resto della catena $\mathbf{p_{r_1-1}}, \ldots, \mathbf{p_1}$. Il vettore $\mathbf{p_1}$, determinato nell'ultimo passaggio, risulta essere un autovettore relativo a λ_1.
Anche nel caso in cui λ_1 coincida con uno o più degli altri autovalori, è lo studio dei sottospazi W_i che conduce alla determinazione delle varie catene di autovettori generalizzati relativi a λ_1. In questo caso però, più difficile del precedente, bisogna tener conto che la differenza $dimW_i - dimW_{i-1}$ può essere maggiore di uno.

Esempio 3.1. Si supponga a titolo di esempio che λ_1 coincida con λ_2 e che sia distinto dai rimanenti autovalori, ci siano cioé due blocchi di Jordan relativi a λ_1; si supponga inoltre che il primo sia di ordine 4 e l'altro di ordine 2. Ci sono allora due catene di Jordan:

$$\{\mathbf{p_1}, \mathbf{p_2}, \mathbf{p_3}, \mathbf{p_4}\} \quad e \quad \{\mathbf{p_5}, \mathbf{p_6}\}$$

una di lunghezza quattro e una di lunghezza due.
Risulterà di conseguenza

$$dimW_1 = 2,\ dimW_2 = 4,\ dimW_3 = 5,\ dimW_4 = 6.$$

Per determinare il vettore direttore della prima catena basta considerare una base di un qualsiasi complemento diretto di W_3 in W_4, un vettore $\mathbf{p_4}$ cioè tale che $\mathbf{p_4} \in W_4$ e $\mathbf{p_4} \notin W_3$. I vettori $\mathbf{p_3}, \mathbf{p_2}, \mathbf{p_1}$ vengono poi determinati come in precedenza con (7). Per determinare il vettore direttore della seconda catena basta considerare un complemento diretto di W_1 in W_2 contenente $\mathbf{p_2}$. L'ulteriore vettore $\mathbf{p_6}$ che, assieme a $\mathbf{p_2}$, ne costituisce una base è il vettore cercato. Infine $\mathbf{p_5} = (A - \lambda I)\mathbf{p_6}$. □

3.3 Teorema di esistenza della forma canonica di Jordan

L'obiettivo di questa sezione è di dimostrare l'esistenza, nella classe di similitudine di una qualsiasi matrice quadrata $A \in M_n(\mathbb{C})$, di una matrice in forma canonica di Jordan. La dimostrazione, dovuta a Filippov [2] e riportata nel libro [6] è molto semplice, in quanto si svolge per induzione sull'ordine della matrice. Si tratta però di una dimostrazione non costruttiva, poiché non illustra la tecnica necessaria per risalire alla forma canonica.

Teorema 3.1 (di esistenza della forma canonica di Jordan)**.** *Ogni matrice $A \in M_n(\mathbb{C})$ è simile ad una matrice in forma canonica di Jordan, ossia di forma (4), dove ogni blocco di Jordan $J_i(\lambda_i)$ è una matrice del $r_i \times r_i$ tipo (3), e $\sum_{i=1}^{p} r_i = n$. Equivalentemente, se $L : V \to V$ è un'applicazione lineare su uno spazio vettoriale V di dimensione finita, allora esiste una base di V tale che la matrice di L ad essa relativa sia del tipo (4).*

Dimostrazione. Come già anticipato, la dimostrazione procede per induzione sull' ordine n della matrice A o, equivalentemente, sulla dimensione n dello spazio vettoriale V, nel caso di un endomorfismo.
Qui si preferisce usare per maggiore semplicità e chiarezza la dimostrazione applicata agli endomorfismi.
Ovviamente, la tesi è vera nel caso $n = 1$.
Consideriamo ora un endomorfismo $L : V \to V$ di uno spazio vettoriale di dimensione n e sia λ un suo autovalore. Indicato con I l'endomorfismo identità su V, cioè l'endomorfismo che ad ogni vettore $v \in V$ associa il vettore stesso, allora $dim(L - \lambda I) = r < n$. Sia L_λ la restrizione di $L - \lambda I$ a $I_m(L - \lambda I)$ definita cioè da

$$L_\lambda x = (L - \lambda I)x \qquad \text{per ogni } x \in I_m(L - \lambda I)$$

dove $L_\lambda x$ e $(L - \lambda I)x$ stanno ad indicare per comodità $L_\lambda(x)$ e $(L - \lambda I)(x)$. Poiché la dimensione del dominio e del codominio di L_λ è $r < n$, per l'ipotesi induttiva, esistono r vettori linearmente indipendenti $\mathbf{p_i}$, che costituiscono una base di $I_m(L - \lambda I)$, tali che la matrice di L_λ ad essa relativa è in forma canonica di Jordan.
Si consideri ora $Q = N(L - \lambda I) \cap I_m(L - \lambda I)$ e sia $q = dimQ$. Ogni vettore di Q è un autovettore di L_λ corrispondente all'autovalore zero. Quindi, L_λ ha q catene di Jordan terminanti in autovettori linearmente indipendenti, dove questi ultimi vettori formano una base per Q. Supponiamo che tra i $\mathbf{p_i}$, i primi

q vettori $\mathbf{p_1}, \ldots, \mathbf{p_q}$ siano i direttori delle catene; essi, per costruzione, sono contenuti in $Im(L - \lambda I)$. Perciò esisteranno vettori $\mathbf{y_1}, \ldots, \mathbf{y_q}$ tali per cui

$$(L - \lambda I)\mathbf{y_i} = \mathbf{p_i} \qquad i = 1, \ldots, q.$$

Infine, per il teorema delle dimensioni, $dim N(L - \lambda I) = n - r$ e quindi $s = n - r - q$ è la dimensione di un qualsiasi complemento diretto di Q in $N(L - \lambda I)$. Siano $\mathbf{z_1}, \ldots, \mathbf{z_s}$ dei vettori di una base di tale complemento diretto. Allora gli n vettori $\mathbf{p_1}, \ldots, \mathbf{p_r}$, $\mathbf{y_1}, \ldots, \mathbf{y_q}$, $\mathbf{z_1}, \ldots, \mathbf{z_s}$ sono tutti autovettori o autovettori generalizzati di $(L - \lambda I)$ e, quindi, di L stessa. Ora, rimane solo dimostrare che tali vettori sono linearmente indipendenti.
Si suppone che:

$$\sum_{i=1}^{r} a_i \mathbf{p_i} + \sum_{i=1}^{q} b_i \mathbf{y_i} + \sum_{i=1}^{s} c_i \mathbf{z_i} = 0_V. \tag{10}$$

La definizione di famiglia linearmente indipendente esige che tutte le costanti a_i, b_i e c_i debbano essere zero. Per dimostrarlo, si considera l'immagine tramite $L - \lambda I$ del vettore al primo membro di (10).
Poichè i vettori $\mathbf{z_i}$ stanno in $N(L - \lambda I)$, si ha

$$\sum_{i=1}^{r} a_i (L - \lambda I)\mathbf{p_i} + \sum_{i=1}^{q} b_i (L - \lambda_I)\mathbf{y_i} = 0_V. \tag{11}$$

Per costruzione, $(L - \lambda I)\mathbf{y_i} = \mathbf{p_i}$, $i = 1, \ldots, q$; di conseguenza la seconda somma di (11) è $\sum_{i=1}^{q} b_i \mathbf{p_i}$, dove i vettori $\mathbf{p_i}$ corrispondono all'autovalore λ e sono i vettori direttori delle catene di Jordan. Guardando poi gli addendi $(L - \lambda I)\mathbf{p_i}$ della prima somma, si deduce che o $(L - \lambda I)\mathbf{p_i} = 0_V$ (se la catena di cui sono elementi direttori è di lunghezza 1) oppure $(L - \lambda I)\mathbf{p_i} = \mathbf{p_j}$ con $q + 1 \le j \le r$ perché $\mathbf{p_j}$ non è un elemento direttore di una catena. Così, nessuno dei $\mathbf{p_i}$ nella seconda somma di (11) appare nella prima somma.
La (11) è allora combinazione lineare tra alcuni dei $\mathbf{p_i}$, $1 \le i \le r$, e poiché questi sono linearmente indipendenti per costruzione, si deduce, in particolare, che tutti i coefficienti b_i devono essere nulli, e la (10) diventa:

$$\sum_{i=1}^{r} a_i \mathbf{p_i} + \sum_{i=1}^{s} c_i \mathbf{z_i} = 0_V.$$

Infine, visto che $\mathbf{z_1}, \ldots, \mathbf{z_s}$ appartengono per costruzione ad un complemento diretto di Q in $N(L - \lambda I)$ e dunque a un complemento diretto di $I_m(L - \lambda I)$ in V, allora $\sum_{i=1}^{s} c_i \mathbf{z_i} = 0_V$. Ricordando ora che i vettori z_i sono linearmente indipendenti, si deduce che tutti i coefficienti c_i sono nulli. Allora da $\sum_{i=1}^{r} a_i \mathbf{p_i} = 0_V$ segue che tutti gli a_i sono nulli. $\square$

4 Esponenziale di una matrice

In questo paragrafo ci occuperemo brevemente di uno strumento molto utilizzato nella risoluzione dei sistemi di equazioni differenziali lineari a coefficienti costanti: l'*esponenziale di una matrice*.
Volutamente saranno omesse le dimostrazioni per le quali si rimanda ai testi riportati nella bibliografia.

4.1 Nozione di norma

Definizione 4.1. *Sia V uno spazio vettoriale reale o complesso. Una* norma *su V è una funzione:*

$$||\cdot|| : V \longrightarrow \mathbb{R}$$

che verifica simultaneamente le seguenti proprietà:

1. $||\mathbf{x}|| \geq 0 \quad \forall \mathbf{x} \in V$ *e* $||\mathbf{x}|| = 0 \Longrightarrow \mathbf{x} = 0$*;*
2. $\forall \lambda \in \mathbb{K}, \quad \forall \mathbf{x} \in V, \quad ||\lambda \mathbf{x}|| = |\lambda|\, ||\mathbf{x}||$*;*
3. $\forall \mathbf{x}, \mathbf{y} \in V, \quad ||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||$.

La condizione 3. viene detta disuguaglianza triangolare.

Uno spazio vettoriale V (reale o complesso) dotato di una norma viene detto *spazio normato.* Se lo spazio vettoriale rimane invariato ma cambia la norma, allora cambia lo spazio normato. In altre parole lo spazio normato è una coppia $(V; ||\cdot||)$ formata da uno spazio vettoriale V e da una norma su V.

Riportiamo ora un importante risultato sulle norme in generale, dimostrato nel libro [5] a pag. 353:

Teorema 4.1. *Sia $[v_1, \dots, v_n]$ una base di uno spazio vettoriale V. Allora qualunque norma definita su V dipende con continuità dalle coordinate degli elementi di V rispetto alla base $[v_1, \dots, v_n]$. In altre parole se $\mathbf{x} = \sum_{j=1}^{n} x_j v_j, \quad \mathbf{y} = \sum_{j=1}^{n} y_j v_j$ sono elementi arbitrari di V e $||\cdot||$ è una norma su V, allora dato $\epsilon > 0$ esiste un $\delta > 0$ dipendente da ϵ tale che*

$$|\, ||\mathbf{x}|| - ||\mathbf{y}||\, | < \epsilon$$

quando $|x_j - y_j| < \delta$ per $j = 1, 2, \dots, n$.

In parole semplici esso afferma che piccole variazioni delle coordinate dell'elemento $\mathbf{x}$ in V danno luogo a piccole variazioni della sua norma.

4.2 Norma submoltiplicativa

Dopo aver dato la definizione di norma su uno spazio vettoriale V qualunque, focalizziamo la nostra attenzione sulle norme dello spazio vettoriale delle matrici quadrate $M_n(\mathbb{C})$. Ovviamente tutte le proprietà discusse nella sezione precedente rimangono valide anche in questo caso particolare.
La possibilità di moltiplicare due matrici qualunque $A, B \in M_n(\mathbb{C})$ comporta l'esigenza di considerare, in tale spazio vettoriale, norme di tipo particolare.

Definizione 4.2. *Sia* $||\cdot||$ *una norma definita in 4.1 sullo spazio vettoriale* $M_n(\mathbb{C})$*. Essa viene detta* norma submoltiplicativa *se verifica anche il seguente assioma:*

$$||AB|| \leq ||A||\ ||B||, \qquad \forall A, B \in M_n(\mathbb{C}).$$

E' importante ricordare che non tutte le norme su $M_n(\mathbb{C})$ sono submoltiplicative.

4.3 La norma come metrica: convergenza

La norma definita su uno spazio vettoriale V permette il confronto tra due qualsiasi elementi di V, consente cioè di precisare quale dei due si può considerare come "più piccolo". Oltre a questo, essa ci consente anche di definire in V una *distanza*.

Proposizione 4.1. *Sia* $(V; ||\cdot||)$ *uno spazio vettoriale normato, allora la funzione* $d : V \times V \to \mathbb{R}$ *che ad ogni coppia di vettori* $\mathbf{x}, \mathbf{y} \in V$ *associa* $||\mathbf{x} - \mathbf{y}||$ *è una distanza in* V*, essa cioè verifica le seguenti proprietà:*

1. $||\mathbf{x} - \mathbf{y}|| \geq 0$ *dove l'uguaglianza è verificata solo se* $\mathbf{x} = \mathbf{y}$*;*
2. $||\mathbf{x} - \mathbf{y}|| = ||\mathbf{y} - \mathbf{x}||$*;*
3. $||\mathbf{x} - \mathbf{y}|| \leq ||\mathbf{x} - \mathbf{z}|| + ||\mathbf{z} - \mathbf{y}||$.

Ora che i concetti di norma su uno spazio vettoriale V di dimensione finita e di distanza tra due elementi di V sono stati definiti, è possibile analizzare la convergenza di successioni (e quindi di serie) di elementi in V. Esistono due tipi di convegenza:

Definizione 4.3. *Una successione* $\{\mathbf{x_k}\}_{k=1}^{\infty}$ *in uno spazio vettoriale normato* $(V; ||\cdot||)$ *viene detta* convergente in norma *a* $\mathbf{x_0} \in V$ *se* $||\mathbf{x_k} - \mathbf{x_0}|| \to 0$ *per* $k \to \infty$.

Teorema 4.2. *Se una successione* $\{\mathbf{x_k}\}_{k=1}^{\infty}$ *converge a* $\mathbf{x_0}$ *secondo una norma, allora essa converge a* $\mathbf{x_0}$ *secondo qualsiasi altra norma.*

Dopo aver introdotto una base $[v_1, \ldots, v_n]$ di V, diamo l'altra definizione di convergenza:

Definizione 4.4. *Una successione* $\{\mathbf{x_k}\}_{k=1}^{\infty}$ *in uno spazio vettoriale* V*, dove ogni elemento ha la seguente rappresentazione*

$$\mathbf{x_k} = \sum_{j=1}^{n} \alpha_j^{(k)} \mathbf{v_j}, \qquad k = 1, 2, \ldots,$$

viene detta convergente coordinata per coordinata *ad un elemento che indicheremo con* $\mathbf{x_0} = \sum_{j=1}^{n} \alpha_j^{(0)} \mathbf{v_j}$ *se* $\alpha_j^{(k)} \to \alpha_j^{(0)}$ *quando* $k \to \infty$ *per ognuna delle* n *sequenze scalari ottenute ponendo* $j = 1, 2, \ldots, n$.

Un caso particolare della definizione precedente si verifica quando lo spazio V è lo spazio delle matrici quadrate:

Definizione 4.5. *Sia* $\{A_k\}_{k=1}^{\infty}$ *una sequenza a elementi in* $M_n(\mathbb{C})$ *e sia* $a_{ij}^{(k)}$ *l'elemento di posto* i, j *nella matrice* A_k *dove* $k = 1, 2, \ldots$ *La sequenza* converge *(coordinata per coordinata) alla matrice* $A \in M_n(\mathbb{C})$ *se gli elementi* a_{ij} *di* A *sono tali per cui* $a_{ij}^{(k)} \to a_{ij}$*, dove* $k \to \infty$ *per ogni coppia* i, j.

Il teorema successivo, dimostrato a pag. 357 di [5], mostra che la convergenza in coordinate e la convergenza in norma per successioni a elementi in V sono equivalenti. Il risultato non è però valido se lo spazio non è di dimensione finita.

Teorema 4.3. *Sia* $\{\mathbf{x_k}\}_{k=1}^{\infty}$ *una successione in uno spazio vettoriale* V *di dimensione finita. La successione converge coordinata per coordinata a* $\mathbf{x_0} \in V$ *se e solo se essa converge a* $\mathbf{x_0}$ *secondo una qualsiasi norma.*

La norma è dunque un utile strumento per studiare la convergenza di successioni e serie di elementi su uno spazio vettoriale arbitrario. La seguente proposizione non è altro che un caso particolare del teorema precedente, dove viene considerato lo spazio vettoriale $M_n(\mathbb{C})$:

Proposizione 4.2. *Sia* $\{A_k\}_{k=1}^{\infty}$ *una successione a elementi in* $M_n(\mathbb{C})$*. La successione converge coordinata per coordinata ad* $A \in M_n(\mathbb{C})$ *se e solo se* $||A_k - A|| \to 0$ *per* $k \to \infty$ *in una qualsiasi norma.*

4.4 Esponenziale di matrice

Concentreremo ora la nostra attenzione su una particolare serie a elementi in $M_n(\mathbb{C})$, che dimostreremo essere convergente. Stiamo parlando della *serie esponenziale*. Diamone la definizione:

Definizione 4.6 (Esponenziale di una matrice)**.** *Sia* $A \in M_n(\mathbb{C})$*. La* matrice esponenziale e^A *di* A *è definita come:*

$$e^A = \mathbf{I} + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{A^k}{k!}, \tag{12}$$

dove e rappresenta il numero di Nepero*. In modo analogo si pone*

$$e^{At} = \mathbf{I} + At + \frac{A^2t^2}{2!} + \frac{A^3t^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{(At)^k}{k!}. \tag{13}$$

Rimane da dimostrare che le due serie e^A ed e^{At} convergono. Prima però enunciamo una proposizione, la cui dimostrazione si trova in [4] a pag.118:

Proposizione 4.3. *Sia $\sum_{k=0}^{\infty} A_k$ una serie di matrici. Se la serie di numeri reali positivi:*

$$\sum_{k=0}^{\infty} ||A_k||$$

converge, allora anche

$$\sum_{k=0}^{\infty} A_k$$

converge.

Questa sarà utilizzata nella dimostrazione del successivo teorema, che mostrerà effettivamente la convergenza dell'esponenziale di una matrice.

Teorema 4.4. *La serie e^A converge per ogni A.*

Dimostrazione. Come conseguenza della definizione 4.2, segue che per ogni $k \geq 0$, $||A^k|| \leq ||A||^k$. Ora la serie:

$$\sum_{k=0}^{\infty} \frac{||A^k||}{k!},$$

è una serie di numeri reali positivi. Si ha, per la suddetta proprietà:

$$\sum_{k=0}^{\infty} \frac{||A^k||}{k!} \leq \sum_{k=0}^{\infty} \frac{||A||^k}{k!} = e^{||A||},$$

e per la proposizione 4.3 la serie (12) converge. La somma, come si è detto, verrà indicata con e^A. □

Corollario 4.1. *La serie e^{At} converge per ogni A e per ogni t.*

Proposizione 4.4. *Siano A e B due matrici quadrate di $M_n(\mathbb{C})$. Se A e B commutano (ovvero se $AB = BA$), vale la proprietà:*

$$e^{A+B} = e^A + e^B.$$

Nel paragrafo successivo mostreremo l'importanza dell'esponenziale di una matrice A, soprattutto nel caso in cui A si trovi in forma canonica di Jordan.

5 Applicazioni

In questo capitolo cercheremo di mettere in evidenza l'utilità della teoria delle matrici e in particolare della forma canonica di Jordan, la quale negli ultimi anni ha trovato vastissimo impiego in molti settori della matematica e dell'ingegneria.
Nella Teoria dei sistemi per esempio, una matrice $F \in M_n(\mathbb{C})$ rappresenta un *sistema dinamico autonomo.* Quest'ultimo è un sistema di n equazioni differenziali del primo ordine in grado di modellare in forma matematica un sistema fisico.
L'obiettivo dell'ingegnere è la risoluzione di queste equazioni rispetto alle condizioni iniziali, per essere così in grado di prevedere l'evoluzione libera del sistema. Ricondurremo il problema della risoluzione del sistema lineare autonomo al calcolo dell' esponenziale della matrice F, mostrando l'importanza di disporre di tale matrice nella forma *'più diagonale possibile'*.

5.1 Sistemi dinamici a tempo continuo

Nel linguaggio della Teoria dei sistemi il termine *sistema dinamico* è un modello matematico di un fenomeno naturale o artificiale, caratterizzato da un certo numero di **variabili descrittive**.
Quasi sempre il legame tra le variabili è orientato nel senso di causa-effetto. Infatti, tra loro individuiamo alcune variabili, dette **variabili di ingresso**, cui normalmente viene data l'interpretazione di variabili che, dall'esterno, agiscono sul sistema e rappresentano, quindi, le *cause* agenti sul sistema, e altre dette **variabili d'uscita**, che rappresentano le risposte del sistema a tali cause, ovvero gli *effetti*.
Naturalmente non tutte le variabili in gioco possono essere considerate ingressi o uscite.
Parlando di sistema dinamico, si può fare riferimento alla cosiddetta *rappresentazione esterna* (o *ingresso-uscita*) o alla *rappresentazione interna* (o *di stato*).
La rappresentazione interna è più complessa di quella esterna ma anche, per certi aspetti più interessante. In essa infatti, oltre alle funzioni di ingresso e uscita, compaiono anche le **variabili di stato**. Queste variabili, hanno nel loro insieme una caratteristica peculiare che viene detta *proprietà di separazione.* Preso cioè un generico istante di tempo, i valori assunti in tale istante da queste variabili contengono nel loro complesso tutta l'informazione sulla storia passata del sistena necessaria per valutare l'andamento futuro sia delle stesse variabili di stato che di quelle di uscita per tempi successivi all'istante considerato.
Alcune volte le variabili del sistema fisico possono essere rappresentate da funzioni che variano con continuità nel tempo, ad esempio la tensione ai capi di un'induttanza, la velocità di un pezzo meccanico, il livello di liquido di un

serbatoio, altre volte, come per la densità di una popolazione rilevata in successivi censimenti o il livello di un magazzino rilevato con inventari periodici, sono meglio rappresentate da successioni numeriche i cui elementi sono i valori assunti dalle variabili negli istanti in cui si considerano. Nel primo caso si parla di **sistemi continui** (nel tempo), nel secondo di **sistemi discreti** (nel tempo). A pag. 15 del libro [3] è possibile trovare un interessante esempio di modellizzazione di un sistema fisico.
Il modo più elegante e proficuo per rappresentare un sistema dinamico è far ricorso a quello che viene comunemente indicato con il termine di **modello di stato lineare**

$$\begin{cases} \dot{x}(t) = Fx(t) + Gu(t) \\ y(t) = Hx(t) + Ju(t) \end{cases}$$

dove, in generale, F è una matrice quadrata a coefficienti costanti di dimensioni $n \times n$, con n il numero delle *variabili di stato*, cioè delle componenti del vettore di stato x, G è una matrice $n \times m$, con m il numero di ingressi (anche l'ingresso in generale sarà un vettore), H una matrice $p \times n$, con p il numero di uscite e J una matrice $p \times m$.
Possiamo evidenziare il seguente caso particolare di interesse

$$\begin{cases} \dot{x}(t) = Fx(t) \\ y(t) = Hx(t) \end{cases}$$

che rappresenta un **sistema autonomo**, dove con il termine *autonomo* intendiamo l'assenza di ingressi che influiscono sul sistema. In tal caso il sistema dinamico evolve solo come conseguenza delle condizioni iniziali $x(0)$ che possono a priori essere qualsiasi.
Si noti che per costruzione la matrice F risulta sempre quadrata (di dimensioni $n \times n$), mentre le matrici G, H e J sono in generale rettangolari.

5.2 Matrice esponenziale e analisi modale

Il passo successivo alla costruzione di un modello è quello di utilizzarlo con riferimento a obiettivi riconducibili fondamentalmente alla soluzione di problemi di analisi.
In questo ambito, lo scopo è quello di valutare il comportamento dinamico del sistema conseguente alla scelta dello stato iniziale e delle funzioni di ingresso, oppure alla variazione di uno o più parmetri caratteristici del modello adottato. Per fare ciò abbiamo bisogno di risolvere le equazioni differenziali del modello di stato. Cerchiamo le soluzioni nel caso autonomo e lineare.
È ben noto come, nel caso $n = 1$ in cui F è uno scalare, la soluzione cercata ha la forma:

$$x(t) = e^{Ft}x(0)$$

dove x(0) è uno scalare che tiene conto delle condizioni iniziali. Nel caso $n > 1$ si vorrebbe che la soluzione generale assumesse la forma $x(t) = e^{Ft}C$

e che ci fosse una soluzione che risolve il problema ai valori iniziali. Iniziamo richiamando il concetto di di *esponenziale di una matrice*, dato nella definizione 4.6, in modo da poter utilizzare la formula precedente anche nel caso generale:

$$e^{Ft} = \mathbf{I} + Ft + \frac{(Ft)^2}{2!} + \cdots + \frac{(Ft)^n}{n!} + \cdots \tag{14}$$

Nel teorema 4.4 della sezione precedente abbiamo dimostrato che tale serie è convergente, ottenendo in questo modo una serie di potenze nella variabile $t \in \mathbb{R}$.
Abbiamo così ricondotto il problema iniziale della risoluzione del sistema lineare autonomo al calcolo di e^{Ft}, in quanto utilizzando la formula (14) (e notando che $e^{Ft} = I$ per $t = 0$) è immediato verificare che $x(t) = e^{Ft}x(0)$ è la soluzione cercata dell'equazione differenziale $\dot{x}(t) = Fx(t)$ con condizione iniziale $x(0)$ anche nel caso $n > 1$.
Cerchiamo ora di calcolare la matrice esponenziale, utilizzando alcune sue proprietà.

Proposizione 5.1. *Se F è diagonale, $F = diag\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, allora e^{Ft} è diagonale e risulta*

$$e^{Ft} = diag\{e^{\lambda_1 t}, e^{\lambda_2 t}, \ldots, e^{\lambda_n t}\}.$$

Quindi possiamo affermare che se F è diagonale lo è anche e^{Ft} e l'elemento i-esimo della sua diagonale è ottenuto calcolando semplicemente l'esponenziale del corrispondente elemento della matrice F. Questo spiega l'importanza di disporre di una matrice F diagonale.
Anche se la matrice F non è diagonale, ma è diagonalizzabile, il problema si può risolvere facilmente ricorrendo a una matrice diagonale J simile a F:

$$J = P^{-1}FP,$$

dove la matrice invertibile P rappresenta ovviamente un cambiamente di base in $\mathbb{C}^n$. La relazione precedente può essere invertita, per cui la matrice originale può essere espressa come:

$$F = PJP^{-1}$$

Utilizzando lo sviluppo in serie è possibile dimostrare facilmente la seguente relazione

$$e^{Ft} = Pe^{Jt}P^{-1}$$

che permette di calcolare e^{Ft} a partire dalla matrice diagonale J simile a F.
Purtroppo non tutti i sistemi dinamici sono diagonalizzabili, come mostra l'esempio 8 a pag. 14 del libro [1]. Nel caso in cui non sia possibile effettuare un cambio di base che diagonalizzi F, si cercherà comunque di ottenere un'espressione la più semplice possibile ("quasi diagonale"). Tale espressione è la

forma canonica di Jordan, che come abbiamo visto, si può sempre trovare. Consideriamo per ora il caso di un unico miniblocco di Jordan, del tipo:

$$J = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

che può essere riscritto come:

$$J = \lambda I + N,$$

dove N è un blocco nilpotente elementare $r \times r$ (r qui rappresenta l'ordine del miniblocco relativo all'autovalore considerato, ed è stato scelto per semplicità uguale a 4).
Per la proposizione 4.4 si ha facilmente

$$e^{Jt} = e^{(\lambda I + N)t} = e^{\lambda I t} e^{Nt}$$

dove l'ultimo passaggio, in generale non effettuabile, è qui corretto in quanto le matrici λI e N commutano. Sfruttando proprietà già viste delle matrici esponenziali si ha:

$$e^{Jt} = \begin{pmatrix} e^{\lambda t} & 0 & 0 & 0 \\ 0 & e^{\lambda t} & 0 & 0 \\ 0 & 0 & e^{\lambda t} & 0 \\ 0 & 0 & 0 & e^{\lambda t} \end{pmatrix} e^{Nt} = [e^{\lambda t} I] e^{Nt}$$

e quindi:

$$e^{Jt} = e^{\lambda t} e^{Nt}$$

Un'importante proprietà della matrice N, vista nella definizione 3.1, è che $N^m = \boldsymbol{O}$. Tale proprietà si mantiene per le successive potenze di N, per cui dopo m elevamenti a potenza otterremo la matrice nulla. Sviluppando in serie e^{Nt} otteremo quindi un numero finito di termini, essendo nulli tutti i termini dal grado m in poi. Pertanto:

$$e^{Nt} = \mathbf{I} + Nt + \frac{(Nt)^2}{2!} + \cdots + \frac{(Nt)^{m-1}}{(m-1)!}$$

e, tornando all'esempio 4×4 precedente:

$$e^{Nt} = \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \frac{t^3}{3!} \\ 0 & 1 & t & \frac{t^2}{2!} \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

che comunque è una matrice triangolare superiore. Alla luce di questi ultimi risultati si può esprimere e^{Jt} come

$$e^{Jt} = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} & \frac{t^3}{3!}e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} \\ 0 & 0 & e^{\lambda t} & te^{\lambda t} \\ 0 & 0 & 0 & e^{\lambda t} \end{pmatrix} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \frac{t^3}{3!} \\ 0 & 1 & t & \frac{t^2}{2!} \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{15}$$

ed è evidente come la matrice e^{Jt} contenga al suo interno elementi del tipo:

$$e^{\lambda t},\ te^{\lambda t},\ \frac{t^2}{2!}e^{\lambda t},\ \frac{t^3}{3!}e^{\lambda t}.$$

In generale, se l'ordine del miniblocco è r, si ottiene

$$e^{Jt} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{r-1}}{(r-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{r-2}}{(r-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & t \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \tag{16}$$

Il seguente teorema generalizza tutte le considerazioni fatte per un singolo miniblocco:

Teorema 5.1. *Sia $F = PJP^{-1}$ una matrice $n \times n$ con forma canonica di Jordan $J = diag(J_1, \dots, J_p)$, dove J_i è un blocco di Jordan $r_i \times r_i$ relativo all'autovalore λ_i. Allora, la matrice e^{Ft} è la*

$$e^{Ft} = Pe^{Jt}P^{-1} = P diag(e^{J_1 t}, e^{J_2 t}, \dots, e^{J_p t})P^{-1}$$

dove $e^{J_i t}$ ha la forma di (16).

In e^{Ft} compariranno i cosiddetti **modi (elementari) del sistema**:

$$e^{\lambda_i t},\ te^{\lambda_i t},\ t^2 e^{\lambda_i t}, \dots,\ t^{h_i} e^{\lambda_i t}$$

dove il massimo esponente in t coincide con l' ordine del miniblocco di Jordan considerato diminuito di 1; avendo più miniblocchi relativi allo stesso autovalore avremo pertanto delle ripetizioni dei modi e quindi il massimo esponente in t relativo ad un autovalore dipende dalla dimensione del miniblocco più grande ad esso associato.

Evidenziamo ora lo stretto legame che intercorre tra i sistemi di equazioni differenziali lineari a coefficienti costanti e l'esponenziale di una matrice. Iniziamo prendendo in esame un sistema omogeneo di n equazioni differenziali lineari a coefficienti costanti

$$\dot{x}(t) = Fx(t), \tag{17}$$

allo scopo di determinarne la **soluzione generale**.
Poniamo $y(t) = P^{-1}x(t)$ dove la matrice invertibile P sia tale che $P^{-1}FP = J$, in cui J è la forma canonica di Jordan associata a F.
Risulta allora

$$x(t) = Py(t) \quad e \quad \dot{x}(t) = P\dot{y}(t) \tag{18}$$

perché gli elementi di P sono delle costanti.
Allora da (17), con le sostituzioni (18), si ottiene il sistema equivalente

$$\dot{y}(t) = Jy(t) \tag{19}$$

in cui le singole equazioni differenziali sono di forma particolarmente semplice.
Osservando che le funzioni $x_i(t)$ che si riferiscono a uno dei blocchi sono diverse da quelle che si riferiscono ad un altro blocco, possiamo ridurre l'analisi al caso in cui J è un blocco elementare di Jordan di ordine n:

$$\begin{pmatrix} \dot{y_1}(t) \\ \vdots \\ \dot{y_n}(t) \end{pmatrix} = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix} \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} \tag{20}$$

In questo modo tutte le equazioni di (20) sono del tipo:

$$\dot{y_n}(t) = \lambda y_n(t), \tag{21}$$

dove compare una sola funzione incognita, oppure del tipo

$$\dot{y_i}(t) = \lambda y_i(t) + y_{i+1}(t) \quad 1 \le i \le n-1 \tag{22}$$

dove, oltre alla funzione y_i compare anche la y_{i+1}.
La soluzione generale di (21) è banalmente $y_n(t) = e^{\lambda t}c_n$.
Il risultato può servirci per risolvere anche la penultima equazione, in cui compare esplicitamente $y_n(t)$:

$$\dot{y_{n-1}}(t) = \lambda y_{n-1}(t) + y_n(t) = \lambda y_{n-1}(t) + e^{\lambda t}c_n$$

Ci troviamo così di fronte ad una equazione differenziale del primo ordine che coinvolge la sola funzione incognita y_{n-1} e che dunque si sa risolvere.
Iterando lo stesso procedimento sino alla prima equazione si ottiene la soluzione generale:

$$y(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} & \cdots & \frac{t^{n-1}}{(n-1)!}e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} & \cdots & \frac{t^{n-2}}{(n-2)!}e^{\lambda t} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & te^{\lambda t} \\ 0 & 0 & 0 & \cdots & e^{\lambda t} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}. \tag{23}$$

Osservando la matrice quadrata di (23), si nota che essa coincide con la matrice esponenziale e^{Jt}.

Il passo successivo è quello di risolvere il *problema di Cauchy*, ovvero provare l'esistenza di un'unica soluzione che verifichi assegnate condizioni iniziali; indichiamo queste ultime con:

$$y(0) = y_0 = \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}.$$

Dalla soluzione generale $y(t) = e^{Jt}C$ dove $C = \begin{pmatrix} c_1 & \cdots & c_n \end{pmatrix}^T$, si determina C perché e^{Jt} è invertibile per ogni $t \in \mathbb{C}$.
Ricordandoci infine del fatto che abbiamo lavorato su di un sistema equivalente, ritorniamo all'originario, sapendo che $x(t) = Py(t)$.
Alla luce di questi risultati enunciamo infine il seguente teorema:

Teorema 5.2. *L'unica soluzione del problema di Cauchy*

$$\begin{cases} \dot{x}(t) = Fx(t) \\ x(0) = x_0 \end{cases}$$

è la funzione:

$$x(t) = e^{Ft}x_0$$

definita nell'intero asse reale e con codominio $\mathbb{C}^n$.

Riferimenti bibliografici

[1] M. Bisiacco - S. Braghetto, *Teoria dei sistemi dinamici*, Esculapio, 2010.

[2] A. F. Filippov, *A short proof of the theorem on reduction of a matrix to Jordan form* (Russian) Vestnik Moskov. Univ. Ser. I Mat. Meh. 26 1971 no. 2, 18-19, tradotto in inglese in Mosc. Univ. Math. Bull. 26 (1971), No.1-2, 70-71.

[3] E. Fornasini - G. Marchesini, *Appunti di Teoria dei sistemi*, Padova: Libreria Progetto, 1994.

[4] L. Gatto, *Un'introduzione amichevole alla forma canonica di Jordan*, C.l.u.t. editrice, 1998.

[5] P. Lancaster - M. Tismenetsky, *The theory of matrices, Second edition with applications*, Academic press inc., 1985.

[6] J.M. Ortega,

Matrix theory a second course, Plenum Press, 1987.

[7] C. Ronconi, *Appunti di Geometria*, Univer editrice, 2009.