

INTRODUZIONE AI SISTEMI DINAMICI
Equazioni differenziali ordinarie,
analisi qualitativa e alcune applicazioni

Guido Gentile

Indice generale

Prefazione	v
Capitolo 1. Operatori lineari	1
1. Spazi vettoriali e operatori lineari	1
2. Esponenziale di un operatore lineare	16
3. Operatori non diagonalizzabili	21
4. Forma canonica di Jordan e forma canonica reale	31
Nota bibliografica	33
Esercizi	33
Capitolo 2. Equazioni differenziali lineari	35
5. Equazioni differenziali lineari omogenee del primo ordine	35
6. Sistemi planari lineari	38
7. Soluzioni di sistemi lineari del primo ordine	45
8. Equazioni differenziali ordinarie di ordine n	66
9. Equazioni differenziali lineari non omogenee del primo ordine	69
Nota bibliografica	72
Esercizi	72

Capitolo 3. Equazioni differenziali ordinarie	75
10. Problema di Cauchy: esistenza e unicità della soluzione	75
11. Dipendenza dai dati iniziali	89
12. Prolungamento delle soluzioni	95
13. Sistemi non autonomi di equazioni differenziali	103
14. Sistemi di equazioni differenziali di ordine qualsiasi	107
15. Dipendenza differenziabile dai dati iniziali	109
Nota bibliografica	114
Esercizi	115
Capitolo 4. Analisi qualitativa del moto e stabilità	119
16. Stabilità secondo Ljapunov	119
17. Linearizzazione	131
18. Teoremi di stabilità	140
19. Teorema della scatola di flusso	149
Nota bibliografica	153
Esercizi	154
Capitolo 5. Alcuni esempi di analisi qualitativa	157
20. Sistemi planari	157
21. Sistemi gradiente	174
22. Equazioni di Lotka-Volterra	177
23. Il pendolo matematico	181
24. Un altro esempio di analisi qualitativa di un sistema planare	191
Nota bibliografica	197
Esercizi	197
Capitolo 6. Sistemi a un grado di libertà	205
25. Sistemi a un grado di libertà	205
26. Analisi qualitativa delle traiettorie: aspetti preliminari	209
27. Analisi qualitativa delle traiettorie: struttura delle orbite	213
28. Periodi e tempi di percorrenza delle orbite	224
29. Stime di periodi	230

Nota bibliografica	234
Esercizi	234
Capitolo 7. Moti centrali e problema dei due corpi	237
30. Moti centrali e problema dei due corpi	237
31. Campo centrale armonico e campo centrale coulombiano	250
32. Teorema di Bertrand	266
Nota bibliografica	273
Esercizi	274
Capitolo 8. Moti relativi e sistemi di riferimento	277
33. Moto in un sistema mobile di coordinate	277
34. Forze d'inerzia	291
Nota bibliografica	299
Esercizi	299
Capitolo 9. Vincoli e sistemi rigidi	305
35. Vincoli	305
36. Sistemi rigidi	310
37. Caratteristiche cinematiche dei sistemi rigidi	319
38. Caratteristiche dinamiche dei sistemi rigidi	324
39. Principio di d'Alembert	326
40. Principio di d'Alembert e vincoli rigidi	330
41. Vincoli anolonomi	334
Nota bibliografica	340
Esercizi	341
Capitolo 10. Proprietà dei sistemi rigidi	343
42. Momenti d'inerzia ed ellissoide d'inerzia	343
43. Angoli di Eulero	356
44. Equazioni di Eulero	365
45. Descrizione del moto secondo Poincot	373
46. Integrabilità di un sistema rigido con un punto fisso	377

iv

Nota bibliografica	385
Esercizi	385
Bibliografia ragionata	389
Indice analitico	393

Prefazione

Lo studio dei sistemi dinamici costituisce una vasta area dell'analisi e della fisica matematica. Lo scopo principale del presente testo è di fornire una trattazione introduttiva per chi si avvicini per la prima volta all'argomento, con lo scopo di fornire le basi necessarie per poter poi passare a problemi di ricerca più complessi.

Il punto di vista adottato è essenzialmente quello della fisica matematica: si è quindi voluto evidenziare come risultati teorici e matematici studiati precedentemente, sia dal punto di vista cronologico sia da quello concettuale, spesso in un'ottica più astratta (cioè prescindendo in sostanza da eventuali applicazioni), possano essere utilizzati e messi in relazione tra loro con lo scopo di affrontare un qualche problema concreto di interesse fisico, costruire un "modello matematico" che possa descriverlo in maniera soddisfacente e quindi cercare di trovarne la soluzione. Si è data perciò grande importanza alla "parte pratica" costituita da esercizi e problemi, di molti dei quali sono forniti la soluzione completa o almeno un suggerimento su come procedere.

Si è voluto espressamente mantenere il livello della discussione a un livello "elementare" (intendendo con ciò che si è ridotta al minimo la richiesta di prerequisiti necessari per seguire il testo), in modo da favorire un primo approccio allo studio dei sistemi dinamici anche da parte di studenti che non abbiano ancora alcuna familiarità con l'argomento. È comunque importante sottolineare che aver tentato di mantenere elementare la presentazione non vuol dire che ci si è limitati a esporre risultati banali: semplicemente si è cercato di chiarire tutti i requisiti indispensabili e fornire tutti i dettagli delle dimostrazioni.

Partendo ab initio, è quindi naturale scegliere come primo argomento i fondamenti della teoria delle equazioni differenziali ordinarie, cominciando dal caso dei sistemi lineari (sempre con l'intento di considerare casi concreti in cui si possano calcolare esplicitamente le soluzioni), per poi affrontare la teoria generale, concettualmente più sottile: tali argomenti coprono i primi tre capitoli del libro (capitoli 1, 2 e 3).

Tali risultati sono poi applicati allo studio di quelle proprietà dei sistemi dinamici che non dipendono in modo essenziale dai dettagli dei singoli sistemi (è quella che viene comunemente chiamata "analisi qualitativa"), di nuovo con particolare attenzione ai casi più adatti all'intuizione e alla visualizzazione, quali i sistemi planari e i sistemi meccanici conservativi a un grado di libertà: a tali argomenti sono dedicati i tre capitoli successivi (capitoli 4, 5 e 6).

La parte restante del testo è quindi rivolta all'applicazione a problemi più complessi

e fisicamente rilevanti, quali i moti centrali (capitolo 7), i cambiamenti di sistemi di riferimento nella descrizione del moto (capitolo 8) e alcune proprietà dei moti di sistemi rigidi (capitoli 9 e 10), introducendo di volta in volta le nozioni fisiche e matematiche necessarie, quali ad esempio quelle di momento angolare o di sistemi vincolati.

Si è preferito privilegiare tali ultimi argomenti rispetto ad altri ugualmente (e forse più) interessanti, quali il formalismo lagrangiano e il formalismo hamiltoniano, per una serie di motivi. In primo luogo si tratta di problemi che si prestano anche allo studio con le sole tecniche introdotte nei primi capitoli, e quindi si evidenzia l'aspetto applicativo dei risultati astratti discussi precedentemente. In secondo luogo, prima di introdurre nuovi concetti e nuovi formalismi, è naturale studiare un po' più in dettaglio quelli già introdotti e discuterne le applicazioni. In terzo luogo, anche solo per meglio apprezzare l'efficacia della teoria dei sistemi lagrangiani e hamiltoniani e i vantaggi che esse offrono, ha ovviamente senso vedere cosa succede "al di fuori" di tali teorie. In quarto luogo una trattazione esauriente, e non semplicemente schematica, dei formalismi lagrangiano e hamiltoniano richiederebbe molto più spazio di quanto occupato dagli ultimi capitoli del libro: abbiamo quindi preferito rimandare la discussione di tali argomenti in altra sede, in modo da avere una trattazione il più organica e autoconsistente possibile.

I libri sugli argomenti trattati non mancano: un breve elenco (tutt'altro che esaustivo!) si può trovare nelle indicazioni bibliografiche alla fine. Il presente testo non nasce ovviamente con l'intento di sostituirli, ma, piuttosto, si prefigge, come anticipato sopra, di offrire una trattazione organica ed elementare. Anzi in molte parti abbiamo tenuto presente altri testi, sempre doverosamente citati, cercando sempre però di rendere omogenea e autoconsistente la discussione: è con questo intento che si è cercato di evitare per quanto possibile di fare riferimento a risultati discussi in altri libri, senza darne una dimostrazione. In ogni caso si è cercato di dare i riferimenti bibliografici nel modo più completo possibile, rendendo quindi estremamente agevole la ricerca di quei risultati elementari (per lo più di Analisi e Geometria) che, pur non essendo esplicitamente dimostrati, sono utilizzati nel corso delle dimostrazioni.

Per la stesura del testo sono state fondamentali le lunghe e proficue discussioni con A. Berretti, G. Gallavotti, V. Mastropietro e E. Scoppola. In particolare ringrazio A. Berretti per avermi mostrato le sue note delle lezioni del corso di Meccanica Razionale, e soprattutto G. Gallavotti per l'interesse mostrato.

Capitolo 1. Operatori lineari

1. Spazi vettoriali e operatori lineari

1.1. Introduzione. Nel seguito considereremo sempre equazioni differenziali in \mathbb{R}^n ; per sottolineare la struttura di spazio vettoriale di \mathbb{R}^n , indicheremo spesso \mathbb{R}^n con E , e chiameremo E uno spazio vettoriale reale. Vedremo che in realtà non è restrittivo identificare E con \mathbb{R}^n ; cfr. §1.26.

Considereremo anche spazi vettoriali complessi, dal momento che tale nozione sorge spontaneamente quando si studino le proprietà spettrali degli operatori lineari definiti su spazi vettoriali (reali). Sarà sempre inteso che considereremo solo spazi vettoriali finito-dimensionali, *i.e.* spazi vettoriali in cui una base (cfr. la definizione 1.4) ha solo un numero finito di elementi.

Nel presente paragrafo richiameremo alcune nozioni basilari sugli spazi vettoriali e sugli operatori lineari su spazi vettoriali, che saranno utilizzate per risolvere sistemi di equazioni lineari, rimandando a testi specifici sull'argomento (cfr. la nota bibliografica alla fine) per un eventuale approfondimento.

1.2. DEFINIZIONE (SPAZIO VETTORIALE). *Definiremo spazio vettoriale reale un insieme non vuoto E in cui siano state introdotte due operazioni, di addizione e di moltiplicazione per uno scalare in $I = \mathbb{R}$, tali che*

$$\begin{aligned}
 (1) \quad & x + y = y + x, \quad \forall x, y \in E, \\
 (2) \quad & x + (y + z) = (x + y) + z, \quad \forall x, y, z \in E, \\
 (3) \quad & x + 0 = x, \quad 0 \in E, \quad \forall x \in E, \\
 (4) \quad & x + (-x) = 0, \quad \forall x \in E, \\
 (5) \quad & \lambda(x + y) = \lambda x + \lambda y, \quad \forall x, y \in E, \quad \forall \lambda \in I, \\
 (6) \quad & (\lambda + \mu)x = \lambda x + \mu x, \quad \forall x \in E, \forall \lambda, \mu \in I, \\
 (7) \quad & 1x = x, \quad 1 \in I, \quad \forall x \in E, \\
 (8) \quad & 0x = 0 \in E, \quad 0 \in I, \quad \forall x \in E,
 \end{aligned} \tag{1.1}$$

Se $I = \mathbb{C}$ diremo che E è uno spazio vettoriale complesso. Chiameremo vettori gli elementi di E .

1.3. DEFINIZIONE (SOTTOSPAZIO). *Un sottoinsieme $F \subset E$ è chiamato sottospazio*

2 CAPITOLO 1. OPERATORI LINEARI

di E se è chiuso rispetto alle operazioni di addizione e di moltiplicazione per uno scalare, i.e. se

$$\lambda x + \mu y \in F, \quad \forall x, y \in F, \quad \forall \lambda, \mu \in I. \quad (1.2)$$

Se $F \neq E$, diremo che F è un sottospazio proprio.

1.4. Dati due spazi vettoriali E e F , indicheremo con $L(E, F)$ l'insieme degli operatori lineari definiti in E a valori in F , i.e. l'insieme degli operatori $T: E \rightarrow F$ tali che

$$T(\lambda x + \mu y) = \lambda Tx + \mu Ty, \quad \forall x, y \in E, \quad \forall \lambda, \mu \in I. \quad (1.3)$$

Dato un operatore $T \in L(E, F)$, definiamo il *nucleo* di T come

$$\text{Ker}(T) = \{x \in E : Tx = 0\} = T^{-1}(0), \quad (1.4)$$

e l'*immagine* di T come

$$\text{Im}(T) = \{y \in F : \exists x \in E \text{ tale che } Tx = y\} = T(E); \quad (1.5)$$

è immediato verificare che i due insiemi (1.4) e (1.5) sono sottospazi vettoriali di E e di F , rispettivamente (cfr. l'esercizio 1). Nel seguito considereremo prevalentemente il caso di operatori $T: E \rightarrow E$, e scriveremo, in tal caso, per semplicità, $L(E, E) \equiv L(E)$.

1.5. DEFINIZIONE (BASE). Diremo che i vettori e_1, \dots, e_n in E generano E , se ogni vettore $x \in E$ si può scrivere come combinazione lineare dei vettori e_1, \dots, e_n , i.e. se per ogni $x \in E$ esistono n numeri x_1, \dots, x_n tali che

$$x = \sum_{i=1}^n x_i e_i. \quad (1.6)$$

Diremo che n vettori e_1, \dots, e_n sono linearmente indipendenti se

$$\sum_{i=1}^n t_i e_i = 0 \quad (1.7)$$

è possibile solo se $t_i = 0 \forall i = 1, \dots, n$ (i.e. se nessuno dei vettori e_1, \dots, e_n si può scrivere come combinazione lineare degli altri). Un insieme di vettori $\{e_1, \dots, e_n\}$ linearmente indipendenti che generano E si dirà una base di E .

1.6. DEFINIZIONE (BASE STANDARD). La base standard in \mathbb{R}^n è definita come la base

$$\begin{aligned} e_1 &= (1, 0, 0, \dots, 0, 0), \\ e_2 &= (0, 1, 0, \dots, 0, 0), \\ &\dots\dots\dots \\ e_n &= (0, 0, 0, \dots, 0, 1), \end{aligned} \quad (1.8)$$

i.e. tale che $(e_j)_i = \delta_{ij}$, dove δ_{ij} è il simbolo di Kronecker, i.e. $\delta_{ij} = 0$ per $i \neq j$ e $\delta_{ii} = 1$.

1.7. LEMMA. *Un sistema di n equazioni lineari omogenee in $n+1$ incognite ammette sempre una soluzione non banale (i.e. diversa dalla soluzione nulla).*

1.8. Dimostrazione del lemma 1.7. La dimostrazione si può fare per induzione.

Il caso $n = 1$ si verifica facilmente. Infatti si ha in tal caso

$$a_1x_1 + a_2x_2 = 0, \quad (1.9)$$

dove almeno uno dei due coefficienti a_1, a_2 è non nullo: supponiamo si abbia $a_1 \neq 0$. Allora si ha

$$x_1 = -\frac{a_2}{a_1}x_2, \quad (1.10)$$

che fornisce una soluzione non banale (\bar{x}_1, \bar{x}_2) , non appena sia $\bar{x}_2 \neq 0$.

Assumiamo dunque che l'affermazione sia vera per qualche n : vogliamo allora far vedere che essa deve essere vera anche per $n+1$. Consideriamo dunque il sistema di $n+1$ equazioni

$$\begin{cases} a_{1,1}x_1 + \dots + a_{1,n+2}x_{n+2} = 0, \\ \dots \\ a_{n+1,1}x_1 + \dots + a_{n+1,n+2}x_{n+2} = 0; \end{cases} \quad (1.11)$$

usiamo la prima equazione per esprimere una variabile in funzione delle altre. Questo è possibile perché almeno uno dei coefficienti $a_{1,j}$, con $j = 1, \dots, n+2$, deve essere diverso da zero: se $a_{1,m} \neq 0$, per qualche m , possiamo allora scrivere

$$x_m = -\sum_{j=1}^{m-1} \frac{a_j}{a_m} x_j - \sum_{j=m+1}^{n+2} \frac{a_j}{a_m} x_j, \quad (1.12)$$

che, sostituita nelle restanti n equazioni in (1.11), produce un sistema di n equazioni nelle $n+1$ incognite $x_1, \dots, x_{m-1}, x_{m+1}, \dots, x_{n+1}$. Per l'ipotesi induttiva tale sistema ammette una soluzione non banale

$$(\bar{x}_1, \dots, \bar{x}_{m-1}, \bar{x}_{m+1}, \dots, \bar{x}_{n+1}). \quad (1.13)$$

Introdotta nella (1.12) tale soluzione determina un valore \bar{x}_m : quindi

$$(\bar{x}_1, \dots, \bar{x}_{m-1}, \bar{x}_m, \bar{x}_{m+1}, \dots, \bar{x}_{n+1}) \quad (1.14)$$

costituisce una soluzione non banale del sistema (1.11). ■

1.9. LEMMA. *Sia $\{e_1, \dots, e_n\}$ una base di uno spazio vettoriale E . Se v_1, \dots, v_m sono m vettori di E linearmente indipendenti, allora $m \leq n$.*

1.10. Dimostrazione del lemma 1.9. Ragioniamo per assurdo. Se $m \geq n+1$ possiamo allora considerare $n+1$ vettori tra i vettori dati v_1, \dots, v_m , per esempio i primi $n+1$, che saranno ancora linearmente indipendenti.

4 CAPITOLO 1. OPERATORI LINEARI

Poiché $\{e_1, \dots, e_n\}$ è una base, si ha

$$v_i = \sum_{j=1}^n v_{ij} e_j, \quad i = 1, \dots, n+1, \quad (1.15)$$

per un'opportuna scelta dei coefficienti v_{ij} . Consideriamo allora il sistema di equazioni

$$\sum_{k=1}^{n+1} v_{kj} x_k = 0, \quad j = 1, \dots, n; \quad (1.16)$$

per il lemma 1.7 tale sistema ammette una soluzione non banale $\bar{x}_1, \dots, \bar{x}_{n+1}$.

Si ha allora

$$\sum_{i=1}^{n+1} \bar{x}_i v_i = \sum_{i=1}^{n+1} \bar{x}_i \sum_{j=1}^n v_{ij} e_j = \sum_{j=1}^n \left(\sum_{i=1}^{n+1} v_{ij} \bar{x}_i \right) e_j = 0, \quad (1.17)$$

per la (1.16): poiché i coefficienti \bar{x}_i non sono tutti nulli, la (1.17) contraddice l'indipendenza lineare dei vettori v_1, \dots, v_{n+1} . ■

1.11. DEFINIZIONE (INSIEME MASSIMALE DI VETTORI LINEARMENTE INDIPENDENTI). Dato un insieme di vettori $\{v_1, \dots, v_m\}$ di uno spazio vettoriale E linearmente indipendenti, diremo che esso è massimale se per ogni $v \in E$ i vettori $\{v, v_1, \dots, v_m\}$ sono linearmente dipendenti.

1.12. LEMMA. Un insieme massimale di vettori linearmente indipendenti di uno spazio vettoriale E costituisce una base di E .

1.13. Dimostrazione del lemma 1.12. Sia $\{v_1, \dots, v_m\}$ un insieme massimale di vettori linearmente indipendenti di E e sia $v \in E \setminus \{v_1, \dots, v_m\}$ non nullo: dobbiamo dimostrare che allora v è combinazione lineare dei vettori v_1, \dots, v_m . Per la definizione 1.11 i vettori v, v_1, \dots, v_m sono linearmente dipendenti, quindi esistono coefficienti x, x_1, \dots, x_m non tutti nulli tali che

$$xv + \sum_{j=1}^m x_j v_j = 0. \quad (1.18)$$

Notiamo innanzitutto che si deve avere $x \neq 0$, altrimenti si ha $x_j = 0 \forall j = 1, \dots, m$, poiché i vettori v_1, \dots, v_m sono linearmente indipendenti. Quindi

$$v = - \sum_{j=1}^m \frac{x_j}{x} v_j. \quad (1.19)$$

che esprime v come combinazione lineare dei vettori v_1, \dots, v_m . ■

1.14. PROPOSIZIONE. *Se uno spazio vettoriale E ha una base $\{e_1, \dots, e_n\}$, ogni altra base in E ha lo stesso numero n di elementi.*

1.15. Dimostrazione della proposizione 1.14. Sia $\{f_1, \dots, f_m\}$ un'altra base di E : per il lemma 1.9 si deve avere $m \leq n$. D'altra parte possiamo scambiare il ruolo delle due basi e riapplicare il lemma 1.9 per concludere che si deve avere $n \leq m$: dalle due disequaglianze si ottiene $m = n$. ■

1.16. DEFINIZIONE (DIMENSIONE). *Si definisce dimensione di uno spazio vettoriale E il numero n di elementi di una base in E ; scriveremo $n = \dim(E)$.*

1.17. Osservazione. Stante la proposizione 1.14 la definizione di dimensione è ben posta, *i.e.* non dipende dalla particolare base scelta.

1.18. PROPOSIZIONE. *Siano E e F due spazi vettoriali e sia $T \in L(E, F)$ un operatore lineare. Si ha allora*

$$\dim(E) = \dim(\text{Ker}(T)) + \dim(\text{Im}(T)). \quad (1.20)$$

1.19. Dimostrazione della proposizione 1.18. Poniamo $n = \dim(E)$, $k = \dim(\text{Im}(T))$. Sia $\{f_1, \dots, f_k\}$ una base di $\text{Im}(T)$. Siano e_1, \dots, e_k i vettori di cui f_1, \dots, f_k sono l'immagine, *i.e.* $f_j = Te_j$, con $j = 1, \dots, k$. Sia $\{g_1, \dots, g_m\}$ una base del sottospazio vettoriale $\text{Ker}(T)$, con $m = \dim(\text{Ker}(T))$. Vogliamo dimostrare che

$$\{e_1, \dots, e_k, g_1, \dots, g_m\} \quad (1.21)$$

costituisce una base di E : da questo segue la (1.20) per la definizione 1.16.

In virtù della definizione 1.5 di base, dobbiamo dimostrare quindi che i vettori in (1.21) sono linearmente indipendenti e generano E .

Consideriamo la combinazione lineare

$$\sum_{j=1}^k \lambda_j e_j + \sum_{i=1}^m \mu_i g_i \quad (1.22)$$

e imponiamo che sia nulla, *i.e.*

$$\sum_{j=1}^k \lambda_j e_j + \sum_{i=1}^m \mu_i g_i = 0, \quad (1.23)$$

Applicando T al vettore (1.22) otteniamo dunque, per la (1.23) e per definizione di $\text{Ker}(T)$,

$$T\left(\sum_{j=1}^k \lambda_j e_j + \sum_{i=1}^m \mu_i g_i\right) = \sum_{j=1}^k \lambda_j Te_j = \sum_{j=1}^k \lambda_j f_j = 0, \quad (1.24)$$

6 CAPITOLO 1. OPERATORI LINEARI

così che l'indipendenza lineare dei vettori f_1, \dots, f_k implica che i coefficienti $\lambda_1, \dots, \lambda_k$ devono essere tutti nulli. Quindi il vettore (1.22), per soddisfare la (1.23), deve essere della forma

$$\sum_{i=1}^m \mu_i g_i; \quad (1.25)$$

tale vettore può essere nullo e quindi soddisfare la (1.23) solo se $\mu_i = 0 \forall i = 1, \dots, m$, poiché i vettori g_1, \dots, g_m sono linearmente indipendenti. In conclusione i vettori in (1.21) sono linearmente indipendenti.

Per completare la dimostrazione occorre dimostrare che i vettori in (1.21) generano E . Sia $x \in E$. Poiché $Tx \in \text{Im}(T)$ e $\{f_1, \dots, f_k\}$ è una base di $\text{Im}(T)$, possiamo scrivere

$$Tx = \sum_{j=1}^k v_j f_j. \quad (1.26)$$

Definiamo

$$v = \sum_{j=1}^k v_j e_j; \quad (1.27)$$

per costruzione $Tv = Tx$. Possiamo allora scrivere $x = v + (x - v)$: il vettore v appartiene al sottospazio di E generato dai vettori e_1, \dots, e_k , per la (1.27), mentre $x - v \in \text{Ker}(T)$, poiché $T(x - v) = 0$, quindi $x - v$ appartiene al sottospazio di E generato dai vettori g_1, \dots, g_k . Questo dimostra quindi che i vettori in (1.21) generano E . ■

1.20. LEMMA. *Sia E uno spazio vettoriale. Dato $T \in L(E, F)$, se $\text{Im}(T) = F$ l'operatore T è invertibile (i.e. esiste $S \in L(F, E)$ tale che $STx = x \forall x \in E$ e $TSy = y \forall y \in F$) se e solo se $\text{Ker}(T) = \{0\}$.*

1.21. Dimostrazione del lemma 1.20. Dimostriamo che l'operatore T è iniettivo se e solo se $\text{Ker}(T) = \{0\}$. Se T è iniettivo e $x \in \text{Ker}(T)$, si ha $Tx = 0 = T0$ e quindi $x = 0$; viceversa se $\text{Ker}(T) = \{0\}$ e $Tx = Ty$, allora $T(x - y) = 0$, i.e. $x - y \in \text{Ker}(T)$, quindi $x - y = 0$, i.e. $x = y$ e quindi T è iniettivo. Poiché per ipotesi T è anche suriettivo ($\text{Im}(T) = F$), la condizione $\text{Ker}(T) = \{0\}$ è quindi necessaria e sufficiente per garantire l'invertibilità di T . ■

1.22. Indichiamo con $M(n)$ l'insieme delle matrici $n \times n$. Se $A, B \in M(n)$ denotiamo con $[A, B]$ il commutatore delle due matrici: $[A, B] = AB - BA$.

1.23. LEMMA. *Fissata una base $\{e_1, \dots, e_n\}$ nello spazio vettoriale E , a ogni operatore $T \in L(E)$ corrisponde una matrice $A \in M(n)$ e viceversa. Diremo che A è la matrice che rappresenta T nella base $\{e_1, \dots, e_n\}$.*

1.24. Dimostrazione del lemma 1.23. Data una matrice A , ovviamente l'applicazione $x \mapsto Ax$ è lineare, quindi definisce un operatore in $L(E)$. Viceversa, dato un operatore

T , definiamo la matrice A i cui elementi A_{ij} sono dati da

$$A_{ij} = (Te_j)_i. \quad (1.28)$$

La (1.28) definisce la matrice che corrisponde a T . Infatti per ogni $x \in E$, se $x = x_1e_1 + \dots + x_n e_n$, si ha, per la linearità di T (cfr. la (1.3)),

$$(Tx)_i = \sum_{j=1}^n x_j (Te_j)_i \equiv \sum_{j=1}^n A_{ij} x_j, \quad (1.29)$$

i.e., in forma più compatta, se identifichiamo x con le sue componenti, $x = (x_1, \dots, x_n)$, possiamo scrivere $Tx = Ax$. ■

1.25. Data una base $\{e_1, \dots, e_n\}$ in uno spazio vettoriale E , indichiamo con $\varphi(z)$ l'operatore che associa al vettore z le sue *componenti* nella base $\{e_1, \dots, e_n\}$, *i.e.* se

$$z = \sum_{i=1}^n x_i e_i, \quad (1.30)$$

allora

$$\varphi(z) = (x_1, \dots, x_n). \quad (1.31)$$

Si verifica subito, a partire dalla definizione (1.3) che l'operatore (1.31) è lineare. Chiameremo le (1.31) le *coordinate cartesiane* (o *coordinate tout court*) di z nella base $\{e_1, \dots, e_n\}$.

Se definiamo $x_i(z) = x_i$ per $i = 1, \dots, n$, allora risulta

$$\varphi(z) = (x_1(z), \dots, x_n(z)), \quad x_i(e_j) = \delta_{ij}; \quad (1.32)$$

diremo che $x = (x_1, \dots, x_n)$ è un *sistema di coordinate* per E nella base $\{e_1, \dots, e_n\}$.

Date due basi $\{e_1, \dots, e_n\}$ e $\{f_1, \dots, f_n\}$ in E , siano $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$, rispettivamente, i sistemi di coordinate degli elementi di E nelle due basi. Quindi $x_i(e_j) = \delta_{ij}$ e $y_i(f_j) = \delta_{ij}$, per costruzione.

1.26. Se l'operatore $T \in L(E, F)$ è invertibile, diremo che T è un *isomorfismo* tra gli spazi vettoriali E e F , e che E e F sono *isomorfi*. Si può facilmente dimostrare che ogni spazio vettoriale di dimensione n è isomorfo a \mathbb{R}^n ; cfr. l'esercizio 2.

1.27. Osservazione. Date due basi $\{e_1, \dots, e_n\}$ e $\{f_1, \dots, f_n\}$ in E , è sempre possibile scrivere

$$f_i = \sum_{j=1}^n P_{ij} e_j, \quad i = 1, \dots, n, \quad (1.33)$$

8 CAPITOLO 1. OPERATORI LINEARI

per qualche matrice P di elementi P_{ij} (*matrice di cambiamento di base*). Infatti ogni f_i è un vettore in E e quindi può essere scritto come combinazione lineare dei vettori della base $\{e_1, \dots, e_n\}$. In particolare, poiché sia i vettori e_1, \dots, e_n sia i vettori f_1, \dots, f_n sono linearmente indipendenti, risulta

$$\det P \neq 0, \quad (1.34)$$

i.e. la trasformazione che fa passare da una base all'altra è invertibile.

Se $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$ sono i sistemi di coordinate di E nelle due basi, rispettivamente, $\{e_1, \dots, e_n\}$ e $\{f_1, \dots, f_n\}$, allora esiste un operatore lineare (invertibile) Q tale che

$$y_i = \sum_{j=1}^n Q_{ij} x_j, \quad i = 1, \dots, n; \quad (1.35)$$

la matrice Q prende il nome di *matrice di cambiamento di coordinate*. Infatti, per ogni $z \in E$, gli operatori $\varphi(z) = (x_1, \dots, x_n)$ e $\psi(z) = (y_1, \dots, y_n)$, che associano a z le sue componenti nelle rispettive basi, sono lineari (invertibili), quindi

$$(y_1(z), \dots, y_n(z)) = \psi(z) = \psi \circ \varphi^{-1}(x_1(z), \dots, x_n(z)), \quad (1.36)$$

e $\psi \circ \varphi^{-1}$ definisce un operatore lineare (invertibile), poiché composizione di operatori lineari (invertibili).

1.28. LEMMA. *Siano $\{e_1, \dots, e_n\}$ e $\{f_1, \dots, f_n\}$ due basi in E , e sia P la matrice tale che*

$$f_i = \sum_{j=1}^n P_{ij} e_j. \quad (1.37)$$

Sia Q la matrice che esprime le coordinate y in termini delle coordinate x (cfr. la (1.35)). Allora

$$Q = (P^T)^{-1}, \quad (1.38)$$

dove P^T indica la matrice trasposta di P .

1.29. Dimostrazione del lemma 1.28. Si ha, utilizzando la linearità delle trasformazioni che associano a ogni vettore le sue componenti in una data base (cfr. la (1.32)),

$$\begin{aligned} \delta_{ij} = y_i(f_j) &= \sum_{k=1}^n Q_{ik} x_k \left(\sum_{s=1}^n P_{js} e_s \right) = \sum_{k=1}^n \sum_{s=1}^n Q_{ik} P_{js} x_k(e_s) \\ &= \sum_{k=1}^n \sum_{s=1}^n Q_{ik} P_{js} \delta_{ks} = \sum_{k=1}^n Q_{ik} P_{jk} = \sum_{k=1}^n Q_{ik} P_{kj}^T, \end{aligned} \quad (1.39)$$

che, scritta, in forma compatta, si legge

$$\mathbb{1} = QP^T, \quad (1.40)$$

e quindi segue la (1.38). ■

1.30. LEMMA. Siano $\{e_1, \dots, e_n\}$ e $\{f_1, \dots, f_n\}$ due basi in E , e siano A e B le matrici che rappresentano un operatore lineare $T \in L(E)$, rispettivamente nelle due basi. Si ha allora

$$B = QAQ^{-1} = (P^T)^{-1}AP^T. \quad (1.41)$$

dove Q e P sono le matrici definite nel lemma 1.28.

1.31. Dimostrazione del lemma 1.30. Si ha $y = Qx$, se x e y sono le coordinate che rappresentano uno stesso vettore di E nelle due basi. Analogamente, se $y' = By$ e $x' = Ax$, risulta $y' = Qx'$. Quindi

$$By = QAx = QAQ^{-1}y; \quad (1.42)$$

poiché la (1.42) vale per ogni x (e quindi per ogni y), segue la prima uguaglianza in (1.41). La seconda segue dal lemma 1.28. ■

1.32. DEFINIZIONE (SOMMA DIRETTA). Dato uno spazio vettoriale E , scriveremo che E è la somma diretta di r sottospazi E_1, \dots, E_r ,

$$E = E_1 \oplus E_2 \oplus \dots \oplus E_r = \bigoplus_{i=1}^r E_i, \quad (1.43)$$

se per ogni $x \in E$ si può scrivere in modo unico $x = x_1 + \dots + x_r$, con $x_i \in E_i$, $\forall i = 1, \dots, r$. Diremo in tal caso che un operatore $T \in L(E)$ è la somma diretta di r operatori lineari T_1, \dots, T_r , con $T_i \in L(E_i)$, $i = 1, \dots, r$,

$$T = T_1 \oplus T_2 \oplus \dots \oplus T_r = \bigoplus_{i=1}^r T_i, \quad (1.44)$$

se

- (1) $TE_i \subset E_i \forall i = 1, \dots, r$ (i.e. il sottospazio E_i è invariante sotto l'azione di T), e
- (2) $Tx = T_i x$ per ogni $x \in E_i$ e per ogni $i = 1, \dots, r$.

1.33. Osservazione. Se è possibile decomporre un operatore $T \in L(E)$ nella somma diretta (1.44), allora deve esistere in E una base $\{e_1, \dots, e_n\}$, tale che

- (1) i primi n_1 elementi e_1, \dots, e_{n_1} costituiscono una base in E_1 , i successivi n_2 elementi $e_{n_1+1}, \dots, e_{n_1+n_2}$ costituiscono una base in E_2 , ..., gli ultimi n_r elementi

10 CAPITOLO 1. OPERATORI LINEARI

e_{n-n_r+1}, \dots, e_n costituiscono una base in E_r ;

(2) in tale base l'operatore $T = T_1 \oplus \dots \oplus T_r$ è rappresentato dalla *matrice a blocchi*

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_r \end{pmatrix}, \quad (1.45)$$

dove A_i è una matrice $n_i \times n_i$, $i = 1, \dots, r$.

1.34. Dato un operatore T , siano $\lambda_1, \dots, \lambda_n$ i suoi *autovalori*: essi saranno le radici del polinomio (di ordine n in λ)

$$p_n(\lambda) = \det(T - \lambda \mathbf{1}), \quad (1.46)$$

che prende il nome di *polinomio caratteristico*. L'insieme degli autovalori di T , che sarà indicato con $\Sigma(T) = \{\lambda_1, \dots, \lambda_n\}$, prende il nome di *spettro* dell'operatore T .

Dato un autovalore λ_i , sia v_i l'*autovettore* corrispondente: v_i è soluzione dell'equazione

$$(T - \lambda_i \mathbf{1}) v_i = 0; \quad (1.47)$$

in virtù della (1.46) la soluzione v_i di (1.47) sarà un vettore non banale, *i.e.* $v_i \neq 0$ per ogni autovalore λ_i .

1.35. DEFINIZIONE (OPERATORE DIAGONALIZZABILE) *Dato uno spazio vettoriale reale E e un operatore lineare $T \in L(E)$, se esiste una base $\{e_1, \dots, e_n\}$ in cui l'operatore è diagonale, *i.e.* T è rappresentato da una matrice diagonale $A_{ij} = \lambda_i \delta_{ij}$, diremo che l'operatore T è diagonalizzabile.*

1.36. Un operatore $T \in L(E)$ avrà sempre n autovalori se $\dim(E) = n$ (poiché n sono le radici del polinomio caratteristico, come conseguenza del teorema fondamentale dell'algebra; cfr. la nota bibliografica). Tuttavia alcuni autovalori possono essere coincidenti, dal momento che un polinomio di ordine n può avere radici coincidenti. Il numero di autovettori (indipendenti) può anche essere minore di n : in generale la (1.47) implica solo che a ogni autovalore λ_i corrisponde un autovettore v_i , ma non permette di concludere se gli autovettori sono linearmente indipendenti o no. Vale tuttavia il seguente risultato.

1.37. TEOREMA. *Se un operatore $T \in L(E)$ ha autovalori reali distinti $\lambda_1, \dots, \lambda_n$, allora è diagonalizzabile (nella base degli autovettori $\{v_1, \dots, v_n\}$).*

1.38. *Dimostrazione del teorema 1.37.* Se $\lambda_1, \dots, \lambda_n$ sono gli autovalori di T e v_1, \dots, v_n gli autovettori corrispondenti, dobbiamo dimostrare che $\{v_1, \dots, v_n\}$ costituisce una base in E . Infatti in tal caso segue dalla (1.28) e dalla (1.47) che, nella base degli autovettori,

$$(Tv_j)_i = \lambda_j (v_j)_i = \lambda_j \delta_{ij}, \quad (1.48)$$

e quindi nella base $\{v_1, \dots, v_n\}$ l'operatore T è diagonale.

Supponiamo per assurdo che $\{v_1, \dots, v_n\}$ non costituisca una base. Sarà allora possibile esprimere (almeno) uno dei vettori, *e.g.* v_n , come combinazione lineare di m degli altri linearmente indipendenti, con $m \leq n - 1$,

$$v_n = \sum_{i=1}^m t_i v_i, \quad (1.49)$$

con i coefficienti t_i non tutti nulli, *i.e.* $(t_1, \dots, t_m) \neq (0, \dots, 0)$. Si avrà allora, per la (1.47),

$$\begin{aligned} 0 &= (T - \lambda_n \mathbf{1}) v_n = (T - \lambda_n \mathbf{1}) \sum_{i=1}^m t_i v_i \\ &= \sum_{i=1}^m (T - \lambda_n \mathbf{1}) t_i v_i = \sum_{i=1}^m (\lambda_i - \lambda_n) t_i v_i. \end{aligned} \quad (1.50)$$

dove si è usata la linearità di T . Per ipotesi $\lambda_i \neq \lambda_j$ per ogni $i \neq j$, quindi l'ultima formula in (1.50) esprime una combinazione lineare a coefficienti non tutti nulli dei vettori v_1, \dots, v_m . Poiché tali vettori sono stati supposti linearmente indipendenti, tale combinazione lineare non può essere nulla, e quindi si è arrivati a una contraddizione. ■

1.39. COROLLARIO. *Se l'operatore $T \in L(E)$ ha autovalori reali distinti e A è la matrice che lo rappresenta in una base qualsiasi $\{e_1, \dots, e_n\}$; allora esiste una matrice Q tale che la matrice*

$$QAQ^{-1} = D \quad (1.51)$$

è la matrice diagonale

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \quad (1.52)$$

e rappresenta T nella base degli autovettori $\{v_1, \dots, v_n\}$.

1.40. Dimostrazione del Corollario 1.39. Che D rappresenti T nella base $\{v_1, \dots, v_n\}$ degli autovettori segue dalla (1.48); basta allora definire Q come la matrice che fa passare dal sistema di coordinate nella base $\{e_1, \dots, e_n\}$ al sistema di coordinate nella base $\{v_1, \dots, v_n\}$ e applicare il lemma 1.30. ■

1.41. Un discorso analogo vale qualora gli autovalori di un operatore T definito in uno spazio vettoriale reale siano sempre distinti, ma complessi (e non più necessariamente reali). Tuttavia in tal caso occorre considerare uno spazio vettoriale esteso, in cui siano compresi anche vettori complessi e l'operazione di moltiplicazione sia definita per scalari complessi.

1.42. DEFINIZIONE (COMPLESSIFICAZIONE DI UNO SPAZIO VETTORIALE). *Sia E uno spazio vettoriale reale; definiremo la sua complessificazione come lo spazio vettoriale*

12 CAPITOLO 1. OPERATORI LINEARI

complesso

$$E_C = \{z : z = \sum_i \lambda_i x_i, \quad x_i \in E, \quad \lambda_i \in \mathbb{C}\}, \quad (1.53)$$

i.e. come lo spazio vettoriale i cui vettori si possono scrivere come combinazione lineare dei vettori di E a coefficienti complessi.

1.43. Sia σ l'operatore di *coniugazione complessa*. Se F è uno spazio vettoriale complesso (con una sua base) e $z = (z_1, \dots, z_n) \in F$, allora

$$\sigma(z) \equiv \sigma(z_1, \dots, z_n) = (\bar{z}_1, \dots, \bar{z}_n) \equiv \bar{z}. \quad (1.54)$$

Ovviamente σ è un'*involutione* in E , *i.e.* $\sigma^2 = \mathbb{1}$ (ovvero $\sigma^{-1} = \sigma$).

1.44. DEFINIZIONE (DECOMPLESSIFICAZIONE DI UNO SPAZIO VETTORIALE). *Sia F uno spazio vettoriale complesso tale che $\sigma(F) \subset F$; allora definiremo la sua decomplessificazione come lo spazio vettoriale reale*

$$F_R = \{z \in F : \sigma(z) = z\}. \quad (1.55)$$

i.e. come l'insieme dei vettori reali in F .

1.45. Osservazione. La condizione $\sigma(F) \subset F$ è essenziale per poter definire la decomplessificazione di uno spazio vettoriale complesso: infatti se $z = x + iy \in F$, dove i è l'unità immaginaria, allora solo se z e $\bar{z} = x - iy$ appartengono a F si ha che

$$x = \frac{z + \bar{z}}{2} \in F, \quad (1.56)$$

in quanto combinazione lineare di vettori dello spazio vettoriale F , e quindi F_R è un sottoinsieme di F che ha ancora struttura di spazio vettoriale.

1.46. Osservazione. Ovviamente la decomplessificazione della complessificazione di uno spazio vettoriale reale E è lo spazio vettoriale E stesso, *i.e.* $E_{CR} = E$.

1.47. DEFINIZIONE (COMPLESSIFICAZIONE DI UN OPERATORE). *Sia $T \in L(E)$ e E_C la complessificazione di E . Definiremo la complessificazione $T_C \in L(E_C)$ dell'operatore T nel modo seguente: se $z \in E_C$, e quindi*

$$z = \sum_i \lambda_i x_i, \quad (1.57)$$

allora

$$T_C z = \sum_i \lambda_i T x_i. \quad (1.58).$$

1.48. PROPOSIZIONE. *Sia E uno spazio vettoriale reale ed E_C la sua complessificazione. Se $Q \in L(E_C)$ esiste un operatore $T \in L(E)$ tale che $Q = T_C$ se e solo se*

$$\sigma Q = Q\sigma, \quad (1.59)$$

dove σ è la coniugazione $\sigma: E_C \rightarrow E_C$.

1.49. Dimostrazione della proposizione 1.48. Se $Q = T_C$, per qualche $T \in L(E)$, allora se z è della forma (1.57) si ha

$$\begin{aligned} Q\sigma(z) &= Q\sigma\left(\sum_i \lambda_i x_i\right) = Q\left(\sum_i \bar{\lambda}_i x_i\right) = T_C\left(\sum_i \bar{\lambda}_i x_i\right) = \sum_i \bar{\lambda}_i T x_i \\ &= \sigma\left(\sum_i \lambda_i T x_i\right) = \sigma(Qz) = \sigma Qz, \end{aligned} \quad (1.60)$$

e quindi $Q\sigma = \sigma Q$.

Viceversa, se $Q\sigma = \sigma Q$, si ha, per $x \in E$,

$$\sigma(Qx) = Q\sigma(x) = Qx, \quad (1.61)$$

così che

$$Qx \in \{y \in E_C : \sigma(y) = y\} = E_{CR} = E, \quad (1.62)$$

i.e. $QE \subset E$. Quindi $T \equiv Q|_E$, la *restrizione* di Q a E , è un operatore lineare in $L(E)$, e possiamo quindi porre $Q = T_C$. ■

1.50. Osservazione. Il significato della proposizione 1.48 è il seguente. Dato un operatore $T \in L(E)$ è sempre possibile considerarne la complessificazione T_C , mentre dato un operatore $Q \in L(E_C)$ non è detto che esista un operatore $T \in L(E)$ di cui Q sia la complessificazione. Anzi in generale questo non è vero: una condizione necessaria e sufficiente perché questo sia possibile è che Q commuti con l'operatore di coniugazione complessa σ .

1.51. LEMMA. *Sia $T \in L(E)$ e T_C la sua complessificazione. Allora T e T_C hanno gli stessi autovalori.*

1.52. Dimostrazione del lemma 1.51. Basta notare che T e T_C hanno lo stesso polinomio caratteristico (1.46): infatti, data una base, T e T_C sono rappresentati dalla stessa matrice (l'unica differenza è che il primo agisce sui vettori di E e l'altro sui vettori di E_C). Poiché gli autovalori sono le radici del polinomio caratteristico, T e T_C devono avere gli stessi autovalori. ■

1.53. TEOREMA. *Se un operatore $T \in L(E)$ ha autovalori complessi distinti $\lambda_1, \dots, \lambda_n$, allora $T_C \in L(E_C)$ è in forma diagonale nella base degli autovettori $\{v_1, \dots, v_n\}$.*

14 CAPITOLO 1. OPERATORI LINEARI

Diremo in tal caso che l'operatore T è semisemplice.

1.54. Dimostrazione del teorema 1.53. La dimostrazione è identica a quella del teorema 1.37. L'unica accortezza è che dobbiamo considerare la complessificazione T_C di T : allora $T_C \in L(E_C)$ ha autovalori distinti e quindi, ragionando come in §1.38, si trova che è in forma diagonale nella base degli autovettori (complessi). ■

1.55. Osservazione. Si noti che un operatore $T \in L(E)$ è semisemplice se la sua complessificazione $T_C \in L(E_C)$ è diagonalizzabile.

1.56. PROPOSIZIONE. *Sia T un operatore lineare in uno spazio vettoriale reale E di dimensione n . I suoi autovalori si possono dividere in due classi, la prima costituita da autovalori reali e la seconda costituita da coppie di autovalori complessi coniugati:*

$$\begin{cases} \lambda_1, \dots, \lambda_s & \text{reali,} \\ \mu_1, \bar{\mu}_1, \dots, \mu_r, \bar{\mu}_r & \text{non reali,} \end{cases} \quad (1.63)$$

con $s + 2r = n$. Gli autovettori corrispondenti a una coppia di autovalori complessi coniugati sono a loro volta complessi coniugati.

1.57. Dimostrazione della proposizione 1.56. Basta dimostrare che lo spettro di T è preservato dall'operazione di coniugazione σ introdotto in §1.43. Sia T_C la complessificazione di T , e siano μ un autovalore complesso di T_C (cfr. il lemma 1.51) e φ l'autovettore associato a μ . Notiamo innanzitutto che essendo μ complesso deve essere φ complesso. Si ha inoltre

$$T_C \bar{\varphi} = T_C \sigma(\varphi) = \sigma(T_C \varphi) = \sigma(\mu \varphi) = \bar{\mu} \bar{\varphi}, \quad (1.64)$$

dove si è usata la proposizione 1.48, e quindi $\bar{\mu}$ è un autovalore di T e $\bar{\varphi}$ è l'autovettore associato a $\bar{\mu}$, per il lemma 1.51. ■

1.58. Dato un operatore $T \in L(E)$ con autovalori distinti, in parte reali in parte complessi (non reali), ordinati come in (1.63), possiamo decomporre lo spazio vettoriale E nella somma diretta di due sottospazi $E_a \oplus E_b$, generati, rispettivamente, dagli autovettori associati agli autovalori reali e dagli autovettori associati agli autovalori complessi. Corrispondentemente risulterà determinata una decomposizione dell'operatore T nella somma diretta di due operatori, $T = T_a \oplus T_b$. Infatti $T E_a \subset E_a$ e $T E_b \subset E_b$, poiché l'azione di T_C su un autovettore è semplicemente la moltiplicazione per il corrispondente autovalore, e quindi combinazioni lineari di autovettori vengono trasformati da T_C in combinazioni lineari degli stessi autovettori: possiamo quindi definire $T_a = T|_{E_a}$ e $T_b = T|_{E_b}$.

In particolare se consideriamo la restrizione di T al sottospazio vettoriale generato dagli autovettori associati a una coppia di autovalori complessi coniugati, otteniamo un operatore invariante.

Quindi possiamo scrivere $E_a = E_{a1} \oplus \dots \oplus E_{ar}$ e, analogamente, $E_b = E_{b1} \oplus \dots \oplus E_{bs}$, dove E_{ai} è il sottospazio generato dall'autovettore v_i associato all'autovalore λ_i , con $i = 1, \dots, r$, ed E_{bi} è il sottospazio generato dagli autovettori φ_i e $\bar{\varphi}_i$ associati agli autovalori μ_i e $\bar{\mu}_i$, con $i = 1, \dots, s$. Si ha quindi $T_a = T_{a1} \oplus \dots \oplus T_{ar}$ e $T_b = T_{b1} \oplus \dots \oplus T_{bs}$. L'operatore T_{ai} agisce (sui vettori in E_{ai}) come un operatore di moltiplicazione per lo scalare λ_i , mentre l'operatore T_{bi} agisce (sui vettori in E_{bi}) come una matrice 2×2 , la cui forma è data dal lemma seguente.

1.59. LEMMA. *Sia E uno spazio vettoriale reale bidimensionale e sia $T \in L(E)$ un operatore lineare in E con autovalori non reali $\mu, \bar{\mu}$, dove $\mu = a + ib$. Esiste allora una base $\{v, u\}$ in cui la matrice A che rappresenta T ha la forma*

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad (1.65)$$

e u, v sono tali che $\varphi = u + iv$ è l'autovettore associato a μ .

1.60. Dimostrazione del lemma 1.59. Dati gli autovalori $\mu, \bar{\mu}$ siano $\varphi, \bar{\varphi}$ gli autovettori associati. Sia E_C la complessificazione di E : si ha allora $\varphi, \bar{\varphi} \in E_C$. Poniamo $\varphi = u + iv$: poiché

$$u = \frac{\varphi + \bar{\varphi}}{2}, \quad v = \frac{\varphi - \bar{\varphi}}{2i}, \quad (1.66)$$

risulta $u, v \in E$. Si ha inoltre

$$T_C \varphi = \mu \varphi = (a + ib)(u + iv) = (au - bv) + i(bu + av); \quad (1.67)$$

D'altra parte

$$T_C \varphi = T_C(u + iv) = Tu + iTv, \quad (1.68)$$

così che, uguagliando parte reale e parte immaginaria delle (1.67) e (1.68), otteniamo

$$Tv = bu + av, \quad Tu = -bv + au. \quad (1.69)$$

Poiché $\varphi, \bar{\varphi}$ sono linearmente indipendenti, lo sono anche v, u : quindi $\{v, u\}$ è una base di E . In tale base l'operatore T è rappresentato dalla matrice di elementi

$$\begin{aligned} A_{11} &= (Tv)_1 = a, & A_{12} &= (Tu)_1 = -b, \\ A_{21} &= (Tv)_2 = b, & A_{22} &= (Tu)_2 = a, \end{aligned} \quad (1.70)$$

(cfr. la (1.28)), da cui segue la (1.65). ■

1.61. Osservazione. Dato un operatore T definito in uno spazio vettoriale reale con autovalori complessi coniugati $\mu, \bar{\mu}$, con $\mu = a + ib$, $b \neq 0$, sia $\{\varphi, \bar{\varphi}\}$ la base degli autovettori, con $\varphi = u + iv$. La matrice P che esprime i vettori della base $\{\varphi, \bar{\varphi}\}$ in funzione della base $\{v, u\}$ è data da

$$\begin{pmatrix} \varphi \\ \bar{\varphi} \end{pmatrix} = P \begin{pmatrix} v \\ u \end{pmatrix}, \quad P = \begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix}, \quad (1.71)$$

16 CAPITOLO 1. OPERATORI LINEARI

così che, ricordando la (1.38),

$$Q^{-1} = P^T = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}, \quad Q = \frac{1}{2i} \begin{pmatrix} 1 & i \\ -1 & i \end{pmatrix}, \quad (1.72)$$

Quindi se $y = (y_1, y_2)$ è il sistema di coordinate nella base $\{\varphi, \bar{\varphi}\}$ e $z = (z_1, z_2)$ è il sistema di coordinate nella base $\{v, u\}$, si ha

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = Q \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \frac{1}{2i} \begin{pmatrix} 1 & i \\ -1 & i \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \quad (1.73)$$

e

$$\begin{pmatrix} \mu & 0 \\ 0 & \bar{\mu} \end{pmatrix} = Q \begin{pmatrix} a & -b \\ b & a \end{pmatrix} Q^{-1} = \frac{1}{2i} \begin{pmatrix} 1 & i \\ -1 & i \end{pmatrix} \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}. \quad (1.74)$$

La (1.74) permette dunque di esprimere la matrice (1.65) in termini della matrice diagonale che rappresenta l'operatore T nella base degli autovettori (complessi).

2. Esponenziale di un operatore lineare

2.1. Introduzione. Vedremo nel prossimo capitolo che, dato un qualsiasi sistema di equazioni differenziali lineari (le definizioni precise saranno date più avanti), esso ammette sempre una soluzione, che si può esplicitamente calcolare purché si sappia calcolare l'esponenziale di una matrice.

Nel presente paragrafo ci occuperemo di definire l'esponenziale di un operatore lineare, di vederne alcune proprietà fondamentali e di calcolarlo in alcuni casi semplici (essenzialmente quando l'operatore è diagonalizzabile o semisemplice). Più avanti vedremo come si calcola anche in casi più complicati.

Ricordiamo che dato uno spazio vettoriale complesso E , si definisce *prodotto scalare* un'applicazione $\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathbb{C}$ tale che (1) $\forall x, y \in E$ si ha $\langle x, y \rangle = \overline{\langle y, x \rangle}$, (2) $\forall x, y, z \in E$ si ha $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$, (3) $\forall x, y \in E$ e $\forall \lambda \in \mathbb{C}$ si ha $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$, (4) $\forall x \in E$ si ha $\langle x, x \rangle \geq 0$, e vale il segno uguale se e solo se $x = 0$. A volte si definisce il prodotto scalare se valgono le prime tre condizioni, e si dice che tale prodotto scalare è *non degenero* se (4') $\langle x, y \rangle = 0 \forall y \in E$ implica $x = 0$, ed è *definito positivo* se invece vale la proprietà (4). Osserviamo anche che a volte la definizione sopra viene data se E è uno spazio vettoriale reale, mentre si parla di *prodotto hermitiano* se E è complesso. Noi, nel seguito, intenderemo comunque il prodotto scalare sempre nel senso della prima definizione data sopra.

Si definisce invece *norma* un'applicazione $|\cdot| : E \rightarrow \mathbb{R}$ tale che (1) $\forall x \in E$ si ha $|x| \geq 0$, e vale il segno uguale se e solo se $x = 0$, (2) $\forall x \in E$ e $\forall \lambda \in \mathbb{C}$ si ha $|\lambda x| = |\lambda| |x|$, dove $|\lambda|$ indica il modulo di λ , (3) $\forall x, y \in E$ si ha $|x + y| \leq |x| + |y|$ (disuguaglianza triangolare).

Ogni prodotto scalare definito su uno spazio vettoriale E induce una norma, attraverso la definizione $|x| = \sqrt{\langle x, x \rangle}$, come è immediato verificare (cfr. l'esercizio 3). Fissata una base $\{e_1, \dots, e_n\}$ in E , chiameremo *prodotto scalare standard* il prodotto scalare $\langle x, y \rangle = x_1 \bar{y}_1 + \dots + x_n \bar{y}_n$, dove (x_1, \dots, x_n) e (y_1, \dots, y_n) sono le componenti di x e y , rispettivamente, nella base $\{e_1, \dots, e_n\}$.

In generale diremo che una base $\{f_1, \dots, f_n\}$ è *ortogonale* se $\langle f_i, f_j \rangle = 0 \forall i, j = 1, \dots, n$ con $i \neq j$; una base ortogonale si dice *ortonormale* se inoltre $\langle f_i, f_j \rangle = \delta_{ij} \forall i, j = 1, \dots, n$.

Uno spazio vettoriale E di dimensione n munito di un prodotto scalare $\langle \cdot, \cdot \rangle$ è chiamato *spazio euclideo*. Il prodotto scalare induce una norma $|\cdot|$: la coppia $(E, |\cdot|)$ costituisce uno *spazio normato*. Se $\langle \cdot, \cdot \rangle$ è il prodotto scalare standard la corrispondente norma $|\cdot|$ prende il nome di *norma euclidea standard* (o semplicemente *norma euclidea*): si ha $|x|^2 = x_1 \bar{x}_1 + \dots + x_n \bar{x}_n \equiv |x_1|^2 + \dots + |x_n|^2$. Se E è uno spazio vettoriale reale, per ogni $x \in E$ si ha $|x|^2 = x_1^2 + \dots + x_n^2$.

2.2. DEFINIZIONE (NORMA UNIFORME). *Sia E uno spazio vettoriale reale e $|\cdot|$ una norma in E (e.g. la norma euclidea standard se E è lo spazio euclideo). Sia $T \in L(E)$. Definiamo la norma uniforme di T in E come*

$$\|T\| \equiv \max_{|x| \leq 1} |Tx|. \quad (2.1)$$

2.3. LEMMA. *Sia E uno spazio vettoriale reale e T, S operatori lineari in E . Risulta:*

$$\begin{aligned} (1) \quad & |Tx| \leq \|T\| |x|, \quad \forall x \in E, \\ (2) \quad & \|ST\| \leq \|S\| \|T\|, \\ (3) \quad & \|T^m\| \leq \|T\|^m, \quad \forall m \in \mathbb{N}. \end{aligned} \quad (2.2)$$

2.4. Dimostrazione del lemma 2.3. (1) Se $x = 0$ si ha $Tx = 0$, e quindi la prima di (2.2) è ovviamente soddisfatta. Supponiamo dunque $x \neq 0$. Possiamo allora definire $y = x/|x|$, così che $|y| = 1$. Si ha allora, dalla definizione (2.1), $\|T\| \geq |Ty| \geq |Tx|/|x|$ (per le proprietà della norma), da cui segue la (1) in (2.2). ■

(2) Si ha $|STx| \leq \|S\| |Tx| \leq \|S\| \|T\| |x|$, e quindi

$$\|ST\| \equiv \max_{|x| \leq 1} |STx| \leq \max_{|x| \leq 1} \|S\| \|T\| |x| = \|S\| \|T\|, \quad (2.3)$$

da cui segue la (2) in (2.2). ■

(3) Si può procedere per induzione. Per $m = 1$ la relazione è ovviamente soddisfatta. Supponendo che essa valga per $m' < m$, si ottiene allora dalla (2)

$$\|T^m\| = \|T^{m-1}T\| \leq \|T^{m-1}\| \|T\| \leq \|T\|^{m-1} \|T\|, \quad (2.4)$$

18 CAPITOLO 1. OPERATORI LINEARI

che dà subito la (3) in (2.2) anche per m . ■

2.5. DEFINIZIONE (ESPONENZIALE DI UN OPERATORE LINEARE). *Definiamo l'esponenziale di un operatore lineare $T \in L(E)$ come*

$$\exp T \equiv e^T \equiv \sum_{k=0}^{\infty} \frac{T^k}{k!}, \quad (2.5)$$

dove $T^0 = \mathbb{1}$ per ogni $T \in L(E)$.

2.6. PROPOSIZIONE. *La serie che definisce l'esponenziale di un operatore lineare è assolutamente convergente per ogni T .*

2.7. Dimostrazione della proposizione 2.6. Per ogni $x \in E$ si ha

$$e^T x = \sum_{k=0}^{\infty} \frac{T^k x}{k!} \quad (2.6)$$

e quindi

$$|e^T x| \leq \sum_{k=0}^{\infty} \frac{|T^k x|}{k!} \leq \sum_{k=0}^{\infty} \frac{\|T\|^k |x|}{k!}, \quad (2.7)$$

per le disuguaglianze (1) e (3) in (2.2). Quindi

$$\|e^T\| \leq \sum_{k=0}^{\infty} \frac{\|T\|^k}{k!}, \quad (2.8)$$

che converge per ogni $\|T\| \in \mathbb{R}$. La (2.8) esprime la convergenza assoluta (*i.e.* in norma) della serie (2.5). ■

2.8. LEMMA. *Siano A, B, Q, T, S operatori in $L(E)$, e sia Q invertibile. Valgono le seguenti identità:*

$$\begin{aligned} (1) \quad & \text{se } B = Q A Q^{-1}, \quad \text{allora } e^B = Q e^A Q^{-1}; \\ (2) \quad & \text{se } S T = T S, \quad \text{allora } e^{S+T} = e^S e^T; \\ (3) \quad & e^{-S} = (e^S)^{-1}; \\ (4) \quad & \text{se } A_{ij} = \lambda_i \delta_{ij}, \quad \text{allora } (e^A)_{ij} = e^{\lambda_i} \delta_{ij}; \\ (5) \quad & \text{se } A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad \text{allora } e^A = e^a \begin{pmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{pmatrix}. \end{aligned} \quad (2.9)$$

2.9. Dimostrazione del lemma 2.8. (1) Dalla definizione 2.5 si ha

$$e^{Q A Q^{-1}} = \sum_{k=0}^{\infty} \frac{(Q A Q^{-1})^k}{k!}, \quad (2.10)$$

dove

$$(QAQ^{-1})^k = (QAQ^{-1})(QAQ^{-1}) \dots (QAQ^{-1}) = QA^kQ^{-1}, \quad (2.11)$$

poiché ovviamente $QQ^{-1} = \mathbb{1}$. Quindi la (2.9) dà

$$e^{QAQ^{-1}} = \sum_{k=0}^{\infty} \frac{QA^kQ^{-1}}{k!} = Q \sum_{k=0}^{\infty} \frac{A^k}{k!} Q^{-1} = Qe^A Q^{-1}, \quad (2.12)$$

che è la (1) in (2.9). ■

(2) Dalla definizione 2.5 si ha

$$e^{S+T} = \sum_{k=0}^{\infty} \frac{(S+T)^k}{k!}, \quad (2.13)$$

dove, utilizzando la *formula del binomio di Newton* e la commutatività degli operatori S e T , risulta

$$(S+T)^k = \sum_{j=0}^k \binom{k}{j} S^j T^{k-j}, \quad (2.14)$$

così che

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{(S+T)^k}{k!} &= \sum_{k=0}^{\infty} \sum_{j=0}^k \frac{S^j T^{k-j}}{j!(k-j)!} = \sum_{k=0}^{\infty} \sum_{k_1=0}^k \frac{S^{k_1} T^{k-k_1}}{k_1! (k-k_1)!} \\ &= \sum_{k_1=0}^{\infty} \sum_{k=k_1}^{\infty} \frac{S^{k_1} T^{k-k_1}}{k_1! (k-k_1)!} = \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \frac{S^{k_1} T^{k_2}}{k_1! k_2!} = e^S e^T, \end{aligned} \quad (2.15)$$

di nuovo, per definizione di esponenziale di un operatore. ■

(3) Se nella (2) in (2.9) si pone $T = -S$, otteniamo

$$e^S e^{-S} = e^{S-S} = e^0 = \mathbb{1} + \sum_{k=1}^{\infty} \frac{0^k}{k!} = \mathbb{1}, \quad (2.16)$$

e quindi

$$e^S e^{-S} = \mathbb{1}, \quad (2.17)$$

da cui segue la (3) in (2.9). ■

(4) Se A è una matrice diagonale, *i.e.*

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \quad (2.18)$$

20 CAPITOLO 1. OPERATORI LINEARI

risulta, come è immediato verificare (per induzione; cfr. l'esercizio 4),

$$A^k = \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n^k \end{pmatrix}, \quad (2.19)$$

così che, applicando la definizione 2.5, troviamo

$$\begin{aligned} e^A &= \sum_{k=0}^{\infty} \frac{A^k}{k!} = \begin{pmatrix} \sum_{k=0}^{\infty} \frac{\lambda_1^k}{k!} & 0 & \dots & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{\lambda_2^k}{k!} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sum_{k=0}^{\infty} \frac{\lambda_n^k}{k!} \end{pmatrix} \\ &= \begin{pmatrix} e^{\lambda_1} & 0 & \dots & 0 \\ 0 & e^{\lambda_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & e^{\lambda_n} \end{pmatrix}, \end{aligned} \quad (2.20)$$

che prova la (4) in (2.9). ■

(5) Per dimostrare l'ultima delle (2.9) utilizziamo la (1) e la (4). In particolare la seconda delle due implica che, nella base degli autovettori $\{\varphi, \bar{\varphi}\}$, con $\varphi = u + iv$, in cui l'operatore è rappresentato dalla matrice diagonale

$$D = \begin{pmatrix} \mu & 0 \\ 0 & \bar{\mu} \end{pmatrix}, \quad (2.21)$$

con $\mu = a + ib$, risulta

$$e^D = \begin{pmatrix} e^{\mu} & 0 \\ 0 & e^{\bar{\mu}} \end{pmatrix}. \quad (2.22)$$

Quindi, utilizzando la matrice del cambiamento di base definita nell'osservazione 1.61, troviamo

$$e^T = \frac{1}{2i} \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} \begin{pmatrix} e^{\mu} & 0 \\ 0 & e^{\bar{\mu}} \end{pmatrix} \begin{pmatrix} 1 & i \\ -1 & i \end{pmatrix}. \quad (2.23)$$

che implica, per calcolo diretto, la (5) in (2.9). ■

2.10. Osservazione. Il lemma 2.8 permette di calcolare l'esponenziale di qualsiasi operatore lineare diagonalizzabile (o semisemplice). Supponiamo infatti che T sia un operatore lineare diagonalizzabile. Esiste allora una matrice invertibile Q tale che $D = QTQ^{-1}$ è diagonale: possiamo quindi applicare la (4) di (2.9) e calcolare e^D . D'altra parte, per la (1) di (2.9), si deve avere $e^T = Qe^DQ^{-1}$, e quindi abbiamo ottenuto l'esponenziale di T .

2.11. DEFINIZIONE (OPERATORE LINEARE NILPOTENTE). Dato un operatore lineare $T \in L(E)$ diremo che T è nilpotente di ordine k se $T^k = 0$, i.e. se la potenza k -esima

di T è l'operatore nullo.

2.12. ESEMPIO. Sia $T = a\mathbb{1} + B$ una matrice nello spazio vettoriale reale bidimensionale E , dove

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ b & 0 \end{pmatrix}. \quad (2.24)$$

Ovviamente, $a\mathbb{1}$ e B commutano, quindi possiamo applicare la (2) in (2.8),

$$e^T = e^{a\mathbb{1}}e^B = e^a\mathbb{1}e^B = e^ae^B = e^a \sum_{k=0}^{\infty} \frac{B^k}{k!} = e^a(\mathbb{1} + B) = \begin{pmatrix} e^a & 0 \\ e^{ab} & e^a \end{pmatrix}, \quad (2.25)$$

dove si è usato che la matrice B è nilpotente di ordine 2, *i.e.* $B^2 = 0$, così che dell'intera somma in (2.25) solo i termini con $k = 0, 1$ sopravvivono (ovviamente, se $B^2 = 0$, allora $B^k = 0$ per ogni $k \geq 2$).

3. Operatori non diagonalizzabili

3.1. Introduzione. Finora (con l'eccezione del caso (4) in §6.2) abbiamo considerato solo operatori lineari diagonalizzabili (semisemplici se visti nello spazio reale). Ovviamente non tutti gli operatori lineari sono diagonalizzabili. Un caso notevole di operatori diagonalizzabili che abbiamo incontrato è rappresentato dagli operatori il cui spettro è costituito da autovalori tutti distinti; un altro caso importante è dato dagli operatori simmetrici (cfr. gli esercizi 6÷10). Se tuttavia l'operatore considerato non è diagonalizzabile, la trattazione dei paragrafi precedenti non è più sufficiente. Vedremo nel presente paragrafo come la discussione si estende al caso di operatori non diagonalizzabili; a una prima lettura si possono omettere i sottoparagrafi 3.4÷3.14, e passare direttamente da §3.3 al teorema 3.15 (trascurandone ovviamente la dimostrazione, che è basata sui lemmi precedenti).

3.2. Dato uno spazio vettoriale complesso E , di dimensione $\dim(E) = n$, un operatore $T \in L(E)$ può avere alcuni autovalori coincidenti. In generale il polinomio caratteristico si scriverà

$$p_n(\lambda) = \prod_{i=1}^r (\lambda - \lambda_i)^{n_i}, \quad (3.1)$$

dove

$$n_1 + \dots + n_r = n, \quad (3.2)$$

se n_i è la *molteplicità algebrica* (o, nel seguito, semplicemente *molteplicità*) dell'autovalore λ_i , $i = 1, \dots, r$.

22 CAPITOLO 1. OPERATORI LINEARI

Se gli autovalori sono tutti distinti, *i.e.* $r = n$ e $n_i = 1 \forall i$, T è diagonalizzabile: allora possiamo scrivere E come somma diretta di n sottospazi E_i , $i = 1, \dots, n$, dove E_i è il sottospazio generato da

$$\text{Ker}(T - \lambda_i \mathbb{1}); \quad (3.3)$$

corrispondentemente si può scrivere T come somma diretta di n operatori T_i (cfr. §1.58).

Anche nel caso di autovalori coincidenti il sottospazio associato all'autovalore λ_i è dato da (3.3). Tuttavia E non si potrà più scrivere come somma diretta di tali sottospazi. In tal caso vale però il teorema 3.15 più avanti.

3.3. DEFINIZIONE (AUTOSPAZIO GENERALIZZATO). *Dato uno spazio vettoriale complesso E e dato un operatore $T \in L(E)$, definiremo autospazio generalizzato associato all'autovalore λ_i di T il sottospazio invariante*

$$E(T, \lambda_i) = \text{Ker}(T - \lambda_i \mathbb{1})^{n_i}, \quad (3.4)$$

dove n_i è la molteplicità di λ_i .

3.4. LEMMA. *Sia E uno spazio vettoriale. Dati $x \in E$ e $T \in L(E)$, se i vettori $x, Tx, \dots, T^p x$ sono tutti non nulli e $T^{p+1}x = 0$, allora i vettori $x, Tx, \dots, T^p x$ sono linearmente indipendenti.*

3.5. Dimostrazione del lemma 3.4. Supponiamo per assurdo che si abbia

$$\sum_{k=0}^p a_k T^k x = 0, \quad (3.5)$$

con i coefficienti a_k non tutti nulli. Sia allora

$$m = \min\{k \in \{0, 1, \dots, p\} : a_k \neq 0\}. \quad (3.6)$$

Si ha allora

$$\begin{aligned} 0 &= T^{p-m} \left(\sum_{k=0}^p a_k T^k x \right) = T^{p-m} \left(\sum_{k=m}^p a_k T^k x \right) \\ &= a_m T^p x + \left(\sum_{k=m+1}^p a_k T^{p+(k-m)} x \right) = a_m T^p x, \end{aligned} \quad (3.7)$$

poiché $p + (k - m) \geq p + 1$ per $k \geq m + 1$ e $T^{p'} x = 0$ per ogni $p' \geq p + 1$ (si ha infatti $T^{p+1} x = 0$ per ipotesi). Quindi si è trovato $T^p x = 0$ dal momento che $a_m \neq 0$ per la (3.6): ciò è in contraddizione con l'ipotesi. ■

3.6. LEMMA. *Sia E uno spazio vettoriale. Dati $y \in E$ e $T \in L(E)$, se esistono p vettori x_1, \dots, x_p in E tali che, ponendo $y = x_0$, si abbia $x_{i-1} = Tx_i$, per $i = 1, \dots, p$,*

e non esiste alcun vettore $x_{p+1} \in E$ tale che $x_p = Tx_{p+1}$, allora i vettori x_0, x_1, \dots, x_p sono linearmente indipendenti.

3.7. Dimostrazione del lemma 3.6. Supponiamo per assurdo che si abbia

$$\sum_{k=0}^p a_k x_k = 0, \quad (3.8)$$

con i coefficienti a_k non tutti nulli. Sia allora

$$M = \max\{k \in \{0, 1, \dots, p\} : a_k \neq 0\}. \quad (3.9)$$

Si ha allora $y = T^k x_k$ e $x_k = T^{p-k} x_p$, per ogni $k = 0, \dots, n$, come è immediato verificare, e quindi

$$\begin{aligned} 0 &= T^M \left(\sum_{k=0}^p a_k x_k \right) = T^M \left(\sum_{k=0}^M a_k T^{p-k} x \right) = a_M T^p x_n + \left(\sum_{k=0}^{M-1} a_k T^{p+(M-k)} x \right) \\ &= a_M T^p x_p + T^{p+1} \left(\sum_{k=0}^{M-1} a_k T^{M-k-1} x \right) \equiv a_M y + T^{p+1} \tilde{x}_{p+1}, \end{aligned} \quad (3.10)$$

se definiamo

$$\tilde{x}_{p+1} = \sum_{k=0}^{M-1} a_k T^{M-k-1} x_p. \quad (3.11)$$

Il vettore $x_{p+1} = -\tilde{x}_{p+1}/a_M$ costituisce dunque un vettore tale che $T^{p+1} x_{p+1} = y$ contro l'ipotesi che non esistesse alcun vettore per il quale valesse tale proprietà. ■

3.8. LEMMA. Dati uno spazio vettoriale E e un operatore $T \in L(E)$, definiamo, per $j \in \mathbb{Z}_+$,

$$\begin{aligned} K_j(T) &= \text{Ker}(T^j), & \mathcal{K}(T) &= \bigcup_{j \geq 0} K_j(T), \\ L_j(T) &= \text{Im}(T^j), & \mathcal{L}(T) &= \bigcap_{j \geq 0} L_j(T). \end{aligned} \quad (3.12)$$

Si hanno allora le relazioni di inclusione

$$\begin{aligned} \{0\} &= K_0 \subset K_1 \subset K_2 \subset \dots \subset K_j \subset K_{j+1} \subset \dots \subset \mathcal{K}(T), \\ E &= L_0 \supset L_1 \supset L_2 \supset \dots \supset L_j \supset L_{j+1} \supset \dots \supset \mathcal{L}(T); \end{aligned} \quad (3.13)$$

inoltre esistono due interi k, ℓ tali che

$$\begin{aligned} K_j(T) &= K_k(T) & \forall j \geq k, \\ L_j(T) &= L_\ell(T) & \forall j \geq \ell. \end{aligned} \quad (3.14)$$

24 CAPITOLO 1. OPERATORI LINEARI

Si ha $\max\{k, \ell\} \leq \dim(E)$.

3.9. Dimostrazione del lemma 3.8. Sia $j \in \mathbb{Z}_+$. Se $x \in K_j(T)$ si ha $T^j x = 0$: quindi $T^{j+1}x = TT^j x = 0$, i.e. $x \in K_{j+1}(T)$. Analogamente se $x \in L_{j+1}(T)$ allora esiste $z \in E$ tale che $x = T^{j+1}z$: quindi $x = T^j y$ se $y = Tz$, i.e. $x \in L_j(T)$. Per $j = 0$ si ha $T^0 = \mathbb{1}$: ovviamente $\text{Ker}(\mathbb{1}) = 0$ e $\text{Im}(\mathbb{1}) = E$. Questo dimostra la (3.13).

La dimostrazione delle (3.14) è immediata. Supponiamo che non esista alcun intero $k \leq n$ per cui la prima delle (3.14) sia soddisfatta. Questo vuol dire che ogni spazio $K_j(T)$ è strettamente contenuto in $K_{j+1}(T)$, quindi $\dim(K_j(T)) \leq \dim(K_{j+1}(T)) + 1$. D'altra parte $\dim(\mathcal{K}(T)) \leq n$, quindi le dimensioni degli insiemi $K_j(T)$ non possono crescere indefinitamente all'aumentare di n . Analogamente si ragiona per la seconda delle (3.14).

3.10. Osservazione. Una dimostrazione alternativa (più complicata) della (3.14) si può ottenere applicando il lemma 3.4 per la prima delle (3.14) e il lemma 3.8 per la seconda, nel modo seguente. Sia $n = \dim(E)$. Supponiamo che non esista alcun intero $k \leq n$ per cui la prima di (3.14) sia soddisfatta. Allora deve esistere un intero $n_1 \geq n$ e un vettore $x \in E$ tali che i vettori $x, Tx, \dots, T^{n_1}x$ sono tutti non nulli, mentre $T^{n_1+1}x = 0$. Per il lemma 3.4 i vettori $x, Tx, \dots, T^{n_1}x$ sono $n_1 + 1$ vettori linearmente indipendenti: poiché $n_1 \geq n$ l'affermazione contraddice il lemma 1.9.

Analogamente se non esiste alcun intero $\ell \leq n$ per cui la seconda di (3.14) sia soddisfatta, allora deve esistere un intero $n_2 \geq n$ e un vettore $x \in E$ tali che $x = z_0 = Tz_1 = \dots = T^{n_2}z_{n_2}$ per opportuni $z_1, \dots, z_{n_2} \in E$ e non esiste alcun vettore $z_{n_2+1} \in E$ tale che $x = T^{n_2+1}z_{n_2+1}$. Per il lemma 3.6 i vettori z_0, z_1, \dots, z_{n_2} sono $n_2 + 1$ vettori linearmente indipendenti, di nuovo in contraddizione con il lemma 1.9.

3.11. LEMMA. Dato uno spazio vettoriale E , si ha

$$E = \mathcal{K}(T) \oplus \mathcal{L}(T), \quad (3.15)$$

per ogni operatore $T \in L(E)$.

3.12. Dimostrazione del lemma 3.11. Sia $T \in L(E)$. Per il lemma 3.8 esistono $k, \ell \leq n$ tali che si ha $K_k(T) = \mathcal{K}(T)$ e $L_\ell(T) = \mathcal{L}(T)$. Poniamo $\mathcal{K} \equiv \mathcal{K}(T)$ e $\mathcal{L} \equiv \mathcal{L}(T)$ per semplicità notazionale.

Si ha $T\mathcal{L} = TL_\ell(T) = L_{\ell+1}(T) = L_\ell(T) = \mathcal{L}$. Indicando con $T|_{\mathcal{L}}$ la restrizione di T a \mathcal{L} , risulta $\text{Ker}(T|_{\mathcal{L}}) = \{0\}$: quindi, per il lemma 1.20, $T|_{\mathcal{L}}$ è invertibile.

Inoltre, dal momento che $T^j\mathcal{L} = \mathcal{L}$ per ogni $j \geq 0$, si ha, in particolare, $T^k\mathcal{L} = \mathcal{L}$ e $T^k x \neq 0$ per $x \in \mathcal{L}$ non nullo. Al contrario $T^k\mathcal{K} = 0$. Quindi $\mathcal{K} \cap \mathcal{L} = \{0\}$.

Sia $x \in E$. Si ha $y = T^\ell x \in \mathcal{L}$ (poiché $y \in \text{Im}(T^\ell)$). Per l'appena dimostrata invertibilità di $T|_{\mathcal{L}}$, dato $y \in \mathcal{L}$ esiste $z \in \mathcal{L}$ tale che $y = T^\ell z$. Scriviamo quindi $x = (x - z) + z$. Si ha $T^\ell z \in \mathcal{L}$ e $T^\ell(x - z) = T^\ell x - T^\ell z = y - y = 0$, i.e. $x - z \in \mathcal{K}$. Quindi si può scrivere x come somma di due vettori, uno in \mathcal{L} e uno in \mathcal{K} . Poiché

questo vale per ogni vettore x in E , possiamo concludere che vale la (3.15). ■

3.13. LEMMA. *Sia E uno spazio vettoriale. Dato $T \in L(E)$, sia $\Sigma(T) = \{\lambda_1, \dots, \lambda_r\}$ lo spettro di T . Definiamo, per $k = 1, \dots, r$,*

$$\begin{aligned}\mathcal{K}_k &= \mathcal{K}(T - \lambda_k \mathbb{1}) = \bigcup_{j \geq 0} \text{Ker}(T - \lambda_k \mathbb{1})^j, \\ \mathcal{L}_k &= \mathcal{L}(T - \lambda_k \mathbb{1}) = \bigcap_{j \geq 0} \text{Im}(T - \lambda_k \mathbb{1})^j.\end{aligned}\tag{3.16}$$

Si ha allora

$$E = \mathcal{K}_1 \oplus \dots \oplus \mathcal{K}_r.\tag{3.17}$$

3.14. Dimostrazione del lemma 3.13. La dimostrazione si può fare per induzione sulla dimensione n dello spazio vettoriale E .

Il caso $n = 1$ è ovvio: in tal caso T è l'operatore di moltiplicazione per uno scalare λ_1 . Quindi $(T - \lambda_1)x = 0$ per ogni $x \in E = \mathbb{R}$, *i.e.* $\mathcal{K}_1 = E$.

Supponiamo quindi che la (3.17) valga se $\dim(E) \leq n$ e dimostriamo che allora essa deve valere anche se $\dim(E) = n + 1$.

Per il lemma 3.11, con l'operatore T sostituito con $T - \lambda_1 \mathbb{1}$, si ha

$$E = \mathcal{K}_1 \oplus \mathcal{L}_1,\tag{3.18}$$

avendo utilizzato le definizioni (3.16).

Dobbiamo quindi dimostrare che si ha

$$\mathcal{L}_1 = \mathcal{K}_2 \oplus \dots \oplus \mathcal{K}_r.\tag{3.19}$$

Basterà quindi far vedere che

(1) l'operatore $T|_{\mathcal{L}_1}$ ha autovalori $\lambda_2, \dots, \lambda_r$ e

(2) risulta

$$\mathcal{K}(T - \lambda_k \mathbb{1}|_{\mathcal{L}_1}) = \mathcal{K}(T - \lambda_k \mathbb{1}) \equiv \mathcal{K}_k \quad \forall 1 < k \leq r.\tag{3.20}$$

Questo implicherà la (3.19). Infatti, poiché per la proposizione 1.18 si ha $\dim(\mathcal{L}_1) + \dim(\mathcal{K}_1) = \dim(E)$, così che, tenendo conto che $\dim(\mathcal{K}_1) \geq 1$, risulta $\dim(\mathcal{L}_1) < \dim(E) = n + 1$, *i.e.* $\dim(\mathcal{L}_1) \leq n$, allora, per l'ipotesi induttiva, per $T|_{\mathcal{L}_1}$ si ha

$$\mathcal{L}_1 = \mathcal{K}(T - \lambda_2 \mathbb{1}|_{\mathcal{L}_1}) \oplus \dots \oplus \mathcal{K}(T - \lambda_r \mathbb{1}|_{\mathcal{L}_1}),\tag{3.21}$$

i.e. per $|\mathcal{L}_1$, utilizzando la (3.20), vale la decomposizione (3.19).

Dimostriamo preliminarmente che

$$\text{Ker}(T - \lambda_1 \mathbb{1}|_{\mathcal{K}_k}) = \{0\} \quad \forall 1 < k \leq r.\tag{3.22}$$

26 CAPITOLO 1. OPERATORI LINEARI

Infatti se $(T - \lambda_1 \mathbb{1})x = 0$ per $x \neq 0$, si ha per tale x

$$(T - \lambda_k \mathbb{1})x = (\lambda_1 - \lambda_k)x \quad \forall 1 < k \leq r, \quad (3.23)$$

quindi

$$(T - \lambda_k \mathbb{1})^j x = (\lambda_1 - \lambda_k)^j x \neq 0 \quad \forall 1 < k \leq r, \quad \forall j \geq 0, \quad (3.24)$$

così che si deve avere $x \notin \mathcal{K}_k \quad \forall 1 < k \leq r$, da cui segue la (3.22).

Inoltre se $x \in \mathcal{K}_k$, *i.e.* se esiste $j \geq 0$ tale che

$$(T - \lambda_k \mathbb{1})^j x = 0, \quad (3.25)$$

allora

$$(T - \lambda_k \mathbb{1})^j (T - \lambda_1 \mathbb{1})x = (T - \lambda_1 \mathbb{1})(T - \lambda_k \mathbb{1})^j x = 0, \quad (3.26)$$

quindi \mathcal{K}_k è invariante sotto l'azione di $T - \lambda_1 \mathbb{1}$, *i.e.*

$$(T - \lambda_1 \mathbb{1})\mathcal{K}_k \subset \mathcal{K}_k \quad \forall 1 < k \leq r. \quad (3.27)$$

Questo in particolare implica

$$\mathcal{K}_k \subset \text{Im}(T - \lambda_1 \mathbb{1})^j \quad \forall 1 < k \leq r, \quad \forall j \geq 0, \quad (3.28)$$

dunque

$$\mathcal{K}_k \subset \mathcal{L}_1 \quad \forall 1 < k \leq r, \quad (3.29)$$

per la definizione di \mathcal{L}_1 in (3.16). La (3.20) è perciò dimostrata.

In conclusione $\lambda_2, \dots, \lambda_r$ devono essere autovalori di $T|_{\mathcal{L}_1}$ poiché se $\lambda_k \in \Sigma(T)$, con $k > 1$, esiste un autovettore $x \in E$ tale che $(T - \lambda_k \mathbb{1})x = 0$, *i.e.* $x \in \mathcal{K}_k$ e quindi, per la (3.29), $x \in \mathcal{L}_1$.

Al contrario λ_1 non è un autovalore di $T|_{\mathcal{L}_1}$, poiché la (3.18) implica che l'autovettore di T corrispondente a λ_1 , dovendo appartenere a \mathcal{K}_1 , non può essere in \mathcal{L}_1 .

Infine notiamo che ogni autovalore λ di $T|_{\mathcal{L}_1}$ deve essere anche un autovalore di T : quindi gli autovalori di $T|_{\mathcal{L}_1}$ devono essere elementi dell'insieme $\Sigma(T) \setminus \{\lambda_1\}$, *i.e.* devono essere esattamente $\{\lambda_2, \dots, \lambda_r\}$. ■

3.15. TEOREMA (DECOMPOSIZIONE PRIMARIA). *Sia E uno spazio vettoriale complesso e sia T un operatore in $L(E)$ con autovalori $\lambda_1, \dots, \lambda_r$ di molteplicità, rispettivamente, n_1, \dots, n_r . Allora E è dato dalla somma diretta degli autospazi generalizzati (3.4), $i = 1, \dots, r$. La dimensione di ogni sottospazio (3.4) è n_i .*

3.16. Dimostrazione del teorema 3.15. Per il lemma 3.13 vale la (3.17). L'operatore $T|_{\mathcal{K}_k}$ ha solo l'autovalore λ_k . Infatti supponiamo per assurdo che per qualche $x \in \mathcal{K}_k$ (*i.e.* tale che $(T - \lambda_k \mathbb{1})^j x = 0$ per qualche $j \geq 0$) si abbia $Tx = \lambda_p x$, con $\lambda_p \neq \lambda_k$: allora

$$(T - \lambda_p \mathbb{1})^{n_i} x = (\lambda_k - \lambda_p)^{n_i} x \neq 0, \quad (3.30)$$

quindi x non può essere un autovettore di T corrispondente all'autovalore λ_p .

Inoltre λ_k , come autovalore di $T|_{\mathcal{K}_k}$, ha la stessa molteplicità n_k che ha come autovalore di T : quindi il polinomio caratteristico di $T|_{\mathcal{K}_k}$ ha grado n_k , *i.e.* $\dim(\mathcal{K}_k) = n_k$.

Poiché $E_k \equiv E(T, \lambda_k) = \text{Ker}(T - \lambda_k \mathbb{1})^{n_k}$ (cfr. la definizione 3.3), si ha $E_k \subset \mathcal{K}_k$.

D'altra parte, dato $x \in E_k$, se

$$m(x) = \min\{n : (T - \lambda_k \mathbb{1})^n x = 0\}, \quad (3.31)$$

si ha, ovviamente,

$$\begin{aligned} (T - \lambda_k \mathbb{1})^n x &= 0 & \forall n \geq m(x), \\ (T - \lambda_k \mathbb{1})^n x &\neq 0 & \forall n < m(x); \end{aligned} \quad (3.32)$$

inoltre i vettori $x, (T - \lambda_k \mathbb{1})x, \dots, (T - \lambda_k \mathbb{1})^{m(x)-1}x$ sono linearmente indipendenti per il lemma 3.4: quindi, tenuto conto che $\dim(\mathcal{K}_k) = n_k$,

(1) si deve avere $m(x) \leq n_k$ per ogni $x \in E_k$ per il lemma 3.8 e

(2) deve esistere $x \in E_k$ tale che $m(x) = n_k$.

Riassumendo si deve avere

$$E_k = (T - \lambda_k \mathbb{1})^{n_k} \mathcal{K}_k = \bigcup_{j=0}^{n_k} \text{Ker}(T - \lambda_k \mathbb{1})^j = \bigcup_{j=0}^{\infty} \text{Ker}(T - \lambda_k \mathbb{1})^j = \mathcal{K}_k, \quad (3.33)$$

che dimostra che di fatto si ha $E_k = \mathcal{K}_k$.

Quindi dalla (3.17) del lemma 3.13 concludiamo che si ha

$$E = E(T, \lambda_1) \oplus \dots \oplus E(T, \lambda_r), \quad (3.34)$$

così che il teorema è dimostrato. ■

3.17. TEOREMA. *Sia E uno spazio vettoriale complesso e sia $T \in L(E)$. Allora è sempre possibile scrivere T nella forma*

$$T = S + N, \quad [S, N] = 0, \quad (3.35)$$

dove S è un operatore diagonalizzabile e N un operatore nilpotente. Tale decomposizione è unica.

3.18. Dimostrazione del teorema 3.17. Poniamo $E_k = E(T, \lambda_k)$ come in §3.16. Definiamo $T_k = T|_{E_k}$: per la (3.34) e, tenendo conto che $T|_{E_k} \subset E_k$ (poiché $(T - \lambda_i \mathbb{1})^{n_i} T v = (T - \lambda_i \mathbb{1})^{n_i+1} v = 0$ se $v \in E_k$), per la definizione 1.32 abbiamo

$$T = T_1 \oplus \dots \oplus T_r. \quad (3.36)$$

Per costruzione ogni $T_k : E_k \rightarrow E_k$ ha il solo autovalore λ_k con molteplicità n_k . Possiamo allora definire i due operatori in $L(E_k)$

$$S_k = \lambda_k \mathbb{1}, \quad N_k = T_k - S_k, \quad (3.37)$$

28 CAPITOLO 1. OPERATORI LINEARI

dove S_k è diagonalizzabile e N_k è nilpotente di ordine n_k : infatti S_k è proporzionale all'identità e l'operatore $N_k^{n_k} = (T - \lambda_k \mathbb{1})^{n_k}$, applicato a qualsiasi vettore $x \in E_k$, dà il vettore nullo. Inoltre $[S_k, N_k] = 0$, di nuovo poiché S_k è proporzionale all'identità.

Se allora definiamo

$$S = S_1 \oplus \dots \oplus S_r, \quad N = N_1 \oplus \dots \oplus N_r, \quad (3.38)$$

si ha $T = S + N$. L'operatore S è quindi diagonalizzabile, poiché è diagonale in una base di E costituita dagli elementi delle basi degli autospazi generalizzati, mentre N è nilpotente di ordine $q = \max\{n_1, \dots, n_r\}$, poiché

$$N^q = N_1^q \oplus \dots \oplus N_r^q = 0. \quad (3.39)$$

Inoltre $[S, N] = 0$ poiché per ogni $k = 1, \dots, r$ si ha $[S_k, N_k] = 0$.

Per concludere la dimostrazione del teorema occorre ancora far vedere che la decomposizione $T = S + N$ è unica.

Supponiamo dunque di avere una decomposizione $T = S + N$: vogliamo mostrare che allora S e N devono essere i due operatori dati dalla (3.38).

Notiamo innanzitutto che gli autospazi E_k sono invarianti sotto l'azione di ogni operatore in $L(E)$ che commuta con T : infatti se $x \in E_k$ e A è un operatore in $L(E)$ tale che $[A, T] = 0$, si ha $(T - \lambda_k \mathbb{1})^{n_k} Ax = A(T - \lambda_k \mathbb{1})^{n_k} x = 0$ per ogni $x \in E_k$ e quindi $Ax \in E_k$ se $x \in E_k$. In particolare per ogni $k = 1, \dots, r$ il sottospazio E_k è invariante sotto l'azione di S e N , poiché $[S, T] = [N, T] = 0$: segue che i due operatori $S|_{E_k}$ e $N|_{E_k}$ sono operatori lineari in E_k . Se ricordiamo la definizione 1.32 di somma diretta di operatori lineari, abbiamo quindi dimostrato che S e N si possono scrivere come somme dirette degli operatori $S|_{E_1}, \dots, S|_{E_r}$ e, rispettivamente, degli operatori $N|_{E_1}, \dots, N|_{E_r}$.

Definiamo $S_k \equiv \lambda_k \mathbb{1} \in L(E_k)$ e $N_k \equiv T_k - \lambda_k \mathbb{1}$, come in (3.37). Dobbiamo dunque dimostrare che si ha $S|_{E_k} = S_k$: da questo segue immediatamente l'unicità dell'operatore S (e quindi anche di N), poiché la decomposizione in somma diretta è unica.

Poiché S è diagonalizzabile, lo è anche $S|_{E_k}$: quindi $S|_{E_k} - S_k$ è diagonalizzabile, come si deduce immediatamente tenendo conto che l'operatore $S_k \in L(E_k)$ è proporzionale all'identità in E_k . Si ha

$$S|_{E_k} - S_k = (T|_{E_k} - N|_{E_k}) - (T_k - N_k) = N_k - N|_{E_k}, \quad (3.40)$$

dove $N|_{E_k}$ commuta con $S_k = \lambda_k \mathbb{1}$ e con T_k , quindi anche con N_k . Inoltre, utilizzando il fatto che $N|_{E_k}$ e N_k sono entrambi nilpotenti, è facile verificare che anche l'operatore (3.40) è nilpotente. Infatti, se $(N_k)^m = 0$ e $(N|_{E_k})^\ell = 0$, poniamo $M = \max\{m, \ell\}$: si ha allora

$$(N_k - N|_{E_k})^{2M} = \sum_{p=0}^{2M} \binom{2M}{p} (N_k)^p (N|_{E_k})^{2M-p} = 0, \quad (3.41)$$

poiché $\max_{0 \leq p \leq 2M} \{p, 2M - p\} \geq M$. Per la (3.40) l'operatore $S|E_k - S_k$ è rappresentato da una matrice diagonale nilpotente: l'unica matrice che verifica tali condizioni è la matrice nulla 0. Quindi $S|E_k = \lambda_k \mathbb{1}$, così che si ottiene l'unicità della decomposizione $T = S + N$ e il teorema è dimostrato. ■

3.19. Osservazione. La dimostrazione del teorema 3.17 mostra anche che, nella decomposizione (3.35), l'operatore nilpotente N è tale che $N^q = 0$, con $q = \max\{n_1, \dots, n_r\}$, se n_k è la molteplicità dell'autovalore λ_k , $k = 1, \dots, r$ (cfr. la (3.39)).

3.20. TEOREMA (CAYLEY-HAMILTON). *Sia E uno spazio vettoriale complesso di dimensione n , e sia $T \in L(E)$. Sia*

$$p_n(\lambda) = \prod_{i=1}^r (\lambda - \lambda_i)^{n_i} = \sum_{i=0}^n a_i \lambda^i \quad (3.42)$$

il polinomio caratteristico di T . Allora

$$p_n(T) = 0, \quad (3.43)$$

i.e.

$$\sum_{i=0}^n a_i T^i x = 0, \quad (3.44)$$

per ogni $x \in E$.

3.21. Dimostrazione del teorema 3.20. Occorre dimostrare che per ogni $x \in E$ si ha $p_n(T)x = 0$. Per il teorema 3.15 possiamo scrivere

$$E = E(T, \lambda_1) \oplus \dots \oplus E(T, \lambda_r), \quad (3.45)$$

dove $\Sigma(T) = \{\lambda_1, \dots, \lambda_r\}$ e $E(T, \lambda_1), \dots, E(T, \lambda_r)$ sono gli autospazi generalizzati associati agli autovalori di T . Ne segue che per ogni $x \in E$ esiste $\lambda_k \in \Sigma(T)$ tale che $x \in E(T, \lambda_k)$: quindi $p_n(\lambda_k) = 0$ ed esiste un polinomio $q(\lambda)$ di grado $n - n_k$ tale che

$$p_n(\lambda) = q(\lambda) (\lambda - \lambda_k)^{n_k}, \quad (3.46)$$

se n_k è la molteplicità di λ_k . Si ha allora

$$p_n(T)x = q(\lambda) (\lambda - \lambda_k)^{n_k} x = 0, \quad (3.47)$$

poiché $x \in E(T, \lambda_k)$, i.e. $(T - \lambda_k \mathbb{1})^{n_k} x = 0$. ■

3.22. TEOREMA. *Sia E uno spazio vettoriale reale e sia $T \in L(E)$. Allora è sempre possibile scrivere T nella forma*

$$T = S + N, \quad [S, N] = 0, \quad (3.48)$$

30 CAPITOLO 1. OPERATORI LINEARI

dove S è un operatore semisemplice e N un operatore nilpotente. Tale decomposizione è unica.

3.23. *Dimostrazione del teorema 3.22.* Sia $T \in L(E)$ e T_C la sua complessificazione in E_C . Per il teorema 3.17 esistono due operatori (unici) S_0 e N_0 tali che

- (1) $T_C = S_0 + N_0$,
- (2) $[S_0, N_0] = 0$ e
- (3) S_0 è diagonalizzabile e N_0 è nilpotente.

Se indichiamo con σ l'operatore di coniugazione complessa (cfr. §1.43), è facile vedere che S_0 e N_0 commutano con σ , *i.e.*

$$S_0\sigma = \sigma S_0, \quad N_0\sigma = \sigma N_0. \quad (3.49)$$

Infatti se poniamo

$$S_1 = \sigma S_0 \sigma^{-1}, \quad N_1 = \sigma N_0 \sigma^{-1}, \quad (3.50)$$

e ricordiamo che $T_C\sigma = \sigma T_C$, *i.e.* $T_C = \sigma T_C \sigma^{-1}$, (per la proposizione 1.48), si ha

$$T_C = \sigma T_C \sigma^{-1} = \sigma S_0 \sigma^{-1} + \sigma N_0 \sigma^{-1} = S_1 + N_1. \quad (3.51)$$

Inoltre dalle definizioni (3.50) si ha che $[S_1, N_1] = 0$, S_1 è diagonalizzabile (in E_C) e N_1 è nilpotente. Infatti

$$\begin{aligned} S_1 N_1 - N_1 S_1 &= \sigma S_0 \sigma^{-1} \sigma N_0 \sigma^{-1} - \sigma N_0 \sigma^{-1} \sigma S_0 \sigma^{-1} \\ &= \sigma S_0 N_0 \sigma^{-1} - \sigma N_0 S_0 \sigma^{-1} = \sigma [S_0, N_0] \sigma^{-1} = 0, \end{aligned} \quad (3.52)$$

e, se q è tale che $N_0^q = 0$,

$$N_1^q = \sigma (\sigma N_0 \sigma^{-1}) \dots (\sigma N_0 \sigma^{-1}) = \sigma N_0^q \sigma^{-1} = 0; \quad (3.53)$$

infine, se Q è la matrice del cambiamento di coordinate che diagonalizza S_0 (*i.e.* tale che $Q S_0 Q^{-1}$ è diagonale), si vede immediatamente che la matrice $Q \sigma^{-1}$ diagonalizza S_1 .

Poiché la rappresentazione $T_C = S_0 + N_0$ è unica, si deve quindi avere, per confronto diretto con la (3.51), $S_0 = S_1$ e $N_0 = N_1$. Ma questo equivale a dire che S_0 e N_0 commutano entrambi con σ ; infatti, se $S_0 = S_1$, la definizione di S_1 implica $S_0\sigma = S_1\sigma = \sigma S_0$, e analogamente si ragiona per N_0 . Per la proposizione 1.48 possiamo concludere che sia S_0 sia N_0 sono la complessificazione di operatori in $L(E)$, *i.e.* esistono due operatori $S, N \in L(E)$ tali che

$$S_0 = S_C, \quad N_0 = N_C. \quad (3.54)$$

Si ha

$$(SN - NS)_C = S_0 N_0 - N_0 S_0 = 0, \quad (3.55)$$

quindi $[S, N] = 0$; analogamente si vede che S è diagonalizzabile e N è nilpotente. Questo completa la dimostrazione del teorema. ■

3.24. DEFINIZIONE (PARTE SEMISEMPLICE E PARTE NILPOTENTE). *Se $T \in L(E)$, nella decomposizione (3.35) S è chiamato la parte semisemplice di T e N la sua parte nilpotente.*

4. Forma canonica di Jordan e forma canonica reale

4.1. Introduzione. Nel presente paragrafo ci limitiamo a enunciare alcuni risultati, senza darne la dimostrazione (per la quale rimandiamo ai testi citati nella nota bibliografica).

Nei precedenti paragrafi abbiamo visto come, data una qualsiasi matrice A , se ne possa calcolare l'esponenziale. Questo esaurisce completamente il problema di calcolare esplicitamente la soluzione di un sistema di equazioni differenziali ordinarie omogenee.

Tuttavia può essere interessante, a volte, studiare proprietà generali o qualitative delle soluzioni, prescindendo dalla forma esatta della soluzione. I seguenti risultati sono allora di grande utilità perché mostrano che, dato un qualsiasi operatore lineare T , esiste sempre una base in cui esso può essere rappresentato da una matrice A con una forma ben precisa. Più precisamente essa avrà una struttura a blocchi: quello che cambierà a seconda dell'operatore considerato sono le dimensioni dei blocchi e i valori di alcuni elementi che in essi appaiono (oltre, ovviamente, la forma della matrice di cambiamento di base che porta T nella forma A). Quindi, per discutere alcune proprietà dei sistemi lineari, ci si può sempre mettere nella base in cui T ha la forma A , per ridurre il numero dei casi da considerare; questo per esempio è quanto sarà fatto in §6 per discutere i possibili casi che si presentano nello studio dei sistemi planari lineari.

4.2. DEFINIZIONE (BLOCCO ELEMENTARE DI JORDAN). *Sia $\lambda \in \mathbb{C}$; la matrice $r \times r$*

$$J = \begin{pmatrix} \lambda & 0 & 0 & \dots & 0 & 0 \\ 1 & \lambda & 0 & \dots & 0 & 0 \\ 0 & 1 & \lambda & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & \lambda \end{pmatrix} \quad (4.1)$$

prende il nome di blocco elementare di Jordan di ordine r . Se $r = 1$, $J = \lambda$.

4.3. DEFINIZIONE (MATRICE IN FORMA CANONICA DI JORDAN). *Una matrice si chiama matrice in forma canonica di Jordan (o in forma normale di Jordan) se è*

32 CAPITOLO 1. OPERATORI LINEARI

costituita da blocchi elementari di Jordan.

4.4. Osservazione. In una matrice di Jordan possono esserci blocchi elementari con lo stesso λ . Per esempio la matrice

$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \quad (4.2)$$

è costituita da due blocchi elementari di Jordan, rispettivamente di ordine 2 e di ordine 1.

4.5. TEOREMA. Sia E uno spazio vettoriale complesso e sia $T \in L(E)$. Allora esiste una base in cui T è rappresentata da una matrice in forma canonica di Jordan, tale che i numeri λ che compaiono nei blocchi elementari costituenti sono gli autovalori di T .

4.6. LEMMA. Sia $\lambda = \mu = a + ib$, con $b \neq 0$, un autovalore di molteplicità p di un operatore $T \in L(E)$. Per il Teorema 4.5 esiste una base in cui la restrizione di T_C allo spazio generalizzato associato a μ , i.e. $T_C|_{\text{Ker}(T_C - \mu)^p}$, è rappresentata da una matrice $p \times p$ costituita da blocchi elementari di Jordan, ciascuno della forma (4.1), con $\lambda = \mu$. Per la proposizione 1.56, anche $\bar{\mu} = a - ib$ sarà un autovalore di T , con la stessa molteplicità p . Se consideriamo il sottospazio

$$E_\mu \equiv \text{Ker}(T - \mu)^p \oplus \text{Ker}(T - \bar{\mu})^p, \quad (4.3)$$

allora esiste una base in cui $T_C|_{E_\mu}$ è rappresentato da una matrice $2p \times 2p$ costituita da blocchi della forma

$$J_0 = \begin{pmatrix} D & 0 & 0 & \dots & 0 & 0 \\ I & D & 0 & \dots & 0 & 0 \\ 0 & I & D & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & I & D \end{pmatrix} \quad (4.4)$$

dove D , I e 0 sono le matrici 2×2

$$D = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad 0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad (4.5)$$

Se $p = 1$, $J_0 = D$.

4.7. DEFINIZIONE (MATRICE IN FORMA CANONICA REALE). Una matrice si chiama matrice in forma canonica reale se è costituita da blocchi elementari di Jordan (4.1) corrispondentemente ad autovalori reali e da blocchi della forma (4.4) corrispondentemente ad autovalori complessi coniugati.

4.8. TEOREMA. Sia E uno spazio vettoriale reale e $T \in L(E)$. Allora esiste una base in cui T è rappresentata da una matrice in forma canonica reale, tale che i numeri

λ che compaiono nei blocchi elementari costituenti sono gli autovalori reali di T e i numeri a e b che compaiono nei blocchi della forma (4.4) sono parte reale e parte immaginaria degli autovalori complessi coniugati.

Nota bibliografica

Nel presente capitolo abbiamo seguito (quasi pedissequamente) la trattazione svolta in [Hirsch-Smale], Capp. 3÷6. Definizioni e proprietà fondamentali di spazi vettoriali, prodotti scalari e norme, quali quelle richiamate nel paragrafo §2, possono essere trovate anche in qualsiasi libro di algebra lineare, *e.g.* [Lang], Capp. 2 e 7.

Per una trattazione esaustiva delle matrici canoniche di Jordan, alle quali si è accennato nel paragrafo §4 (e per la dimostrazione dei risultati ivi riportati) si può vedere, oltre che [Hirsch-Smale], Cap. 6, anche [Kuroš], Cap. XIII.

Per le proprietà degli operatori lineari simmetrici, discusse negli esercizi, si può vedere [Lang], Cap. 11. Per la dimostrazione del teorema fondamentale dell'algebra cfr. *e.g.* [Kuroš], Cap. V.

Altri esempi di matrici di cui si può calcolare l'esponenziale si possono trovare negli esercizi del capitolo 2.

Esercizi

Esercizio 1. Dimostrare che, dati due spazi vettoriali E e F e un operatore lineare $T \in L(E, F)$, allora $\text{Ker}(T)$ e $\text{Im}(T)$ sono sottospazi vettoriali di E e F , rispettivamente. [Suggerimento. Applicare la definizione di spazio vettoriale.]

Esercizio 2. Dimostrare che due spazi vettoriali sono isomorfi se e solo se hanno la stessa dimensione. Dedurne che ogni spazio vettoriale di dimensione n è isomorfo a \mathbb{R}^n . [Soluzione. Siano E e F isomorfi: se $\{e_1, \dots, e_n\}$ è una base per E allora $\{Te_1, \dots, Te_n\}$ è una base per F (per la proposizione 1.18 con $\text{Ker}(T) = \emptyset$ e per il lemma 1.20), quindi $\dim(E) = \dim(F)$. Viceversa sia $\dim(E) = \dim(F)$: se $\{e_1, \dots, e_n\}$ è una base per E e $\{f_1, \dots, f_n\}$ è una base per F , e se l'operatore $T \in L(E, F)$ è definito per linearità a partire dalle relazioni $Te_i = f_i$, per $i = 1, \dots, n$, allora per ogni $x = \sum_{i=1}^n x_i e_i$ si ha $Tx = \sum_{i=1}^n x_i f_i$, così che T è sia suriettivo (poiché f_1, \dots, f_n generano F) sia iniettivo (poiché $\text{Ker}(T) = \emptyset$ essendo i vettori f_1, \dots, f_n linearmente indipendenti), quindi T è un isomorfismo.]

Esercizio 3. Dato uno spazio vettoriale E munito di un prodotto scalare $\langle \cdot, \cdot \rangle$ dimostrare che le proprietà (1)÷(3) del prodotto scalare nel sottoparagrafo §2.1 implicano $\langle x, \lambda y \rangle = \bar{\lambda} \langle x, y \rangle \forall x, y \in E$ e $\forall \lambda \in \mathbb{C}$. [Soluzione. Si ha $\langle x, \lambda y \rangle = \overline{\langle \lambda y, x \rangle} = \overline{\lambda \langle y, x \rangle} = \bar{\lambda} \overline{\langle y, x \rangle} = \bar{\lambda} \langle x, y \rangle$.]

Esercizio 4. Si verifichi che, dato uno spazio vettoriale E e ivi definito un prodotto scalare $\langle \cdot, \cdot \rangle$, l'applicazione che associa a ogni $x \in E$ la quantità $|x| = \langle x, x \rangle$ effettivamente definisce una norma.

Esercizio 5. Sia A la matrice diagonale (2.18); si dimostri per induzione che la matrice A^k , con $k \in \mathbb{N}$, è data dalla (2.19).

Esercizio 6. Un operatore lineare $A \in T(E)$ si dice *simmetrico* se esiste una base in cui si ha

34 CAPITOLO 1. OPERATORI LINEARI

$\langle x, Ay \rangle = \langle Ax, y \rangle \forall x, y \in E$. Dimostrare che se A è simmetrico allora, nella base fissata, la matrice A che lo rappresenta è tale che $A = A^T$, i.e. $A_{ij} = A_{ji}$.

Esercizio 7. Dimostrare che gli autovalori di un operatore lineare simmetrico A devono essere tutti reali. [Soluzione. Se $Av = \lambda v$, si ha $\bar{\lambda} = \langle v, \lambda v \rangle = \langle v, Av \rangle = \langle Av, v \rangle = \langle \lambda v, v \rangle = \lambda \langle v, v \rangle$.]

Esercizio 8. Dimostrare che un operatore lineare simmetrico ha autovettori reali non nulli. [Soluzione. Siano λ un autovalore di A e v l'autovettore corrispondente: quindi $Av = \lambda v$. Poiché v è non nullo, se scriviamo $v = a + ib$, con a, b vettori reali, almeno uno tra a e b deve essere non nullo. Per linearità $Av = Aa + iAb = \lambda a + i\lambda b$, poiché λ è reale, quindi $Aa = \lambda a$ e $Ab = \lambda b$, i.e. a e b sono entrambi autovettori di A , e almeno uno di essi è non nullo.]

Esercizio 9. Dimostrare che gli autovettori di un operatore lineare simmetrico $A \in T(E)$ formano una base ortogonale in E . [Soluzione. Per induzione sulla dimensione di E . Il caso $n = \dim(E) = 1$ è ovvio. Se $n > 1$ sia $\lambda_1 \in \Sigma(A)$ tale che $Av_1 = \lambda_1 v_1$ con v_1 reale non nullo (per l'esercizio 7 questo è sempre possibile). Sia $W = \{w \in E : \langle w, v_1 \rangle = 0\}$ il complemento ortogonale di v_1 in E . Quindi $\dim(W) = n - 1$. Inoltre se $w \in W$ anche $Aw \in W$ (poiché $\langle Aw, v_1 \rangle = \langle w, Av_1 \rangle = \lambda_1 \langle w, v_1 \rangle = 0$): W è invariante sotto l'azione di A . L'operatore $A|_W$ (i.e. A ristretto a W) è quindi un operatore lineare simmetrico in $T(W)$, quindi, per l'ipotesi induttiva, i suoi autovettori costituiscono una base ortogonale $\{v_2, \dots, v_n\}$. D'altra parte v_1 e W generano E , sono ovviamente linearmente indipendenti e $\langle v_j, v_1 \rangle = 0$ per definizione di W .]

Esercizio 10. Dimostrare che se A è una matrice simmetrica esiste una matrice U ortogonale (i.e. tale che $U^T = U^{-1}$) che la diagonalizza. [Soluzione. Nella base degli autovettori la matrice D , che rappresenta l'operatore che è rappresentato dalla matrice A nella base standard, è diagonale, quindi esiste una matrice invertibile Q tale che $QAQ^{-1} = D$, dove $P = (Q^T)^{-1}$ è la matrice del cambiamento di base. Se gli autovettori sono normalizzati (cosa che si può sempre supporre) si ha $\langle v_i, v_j \rangle = \delta_{ij}$, quindi, come è immediato verificare, $P^T P = \mathbb{1}$.]

Esercizio 11. Sia T un operatore lineare in \mathbb{R}^2 non diagonalizzabile. Dimostrare che esiste una base $\{v_1, v_2\}$ in cui l'operatore T è rappresentato dalla matrice in forma canonica di Jordan (4.1) con $n = 2$.

Esercizio 12. Sia A la matrice

$$A = \begin{pmatrix} 3 & -2 \\ 2 & -1 \end{pmatrix}.$$

Se ne calcoli l'esponenziale. Trovare la matrice P del cambiamento di base che porta la matrice A nella forma canonica di Jordan.

Esercizio 13. Si calcoli l'esponenziale della matrice

$$A = \begin{pmatrix} 1 & -2 \\ 4 & 1 \end{pmatrix}.$$

Esercizio 14. Si calcoli l'esponenziale della matrice

$$A = \begin{pmatrix} 2 & 1 & 1 \\ -4 & 2 & 2 \\ 2 & 1 & 2 \end{pmatrix}.$$

Esercizio 15. Si calcoli l'esponenziale della matrice

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

Esercizio 16. Si calcoli l'esponenziale della matrice

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