

SERGE LANG

ALGEBRA LINEARE

PROGRAMMA DI MATEMATICA, FISICA, ELETTRONICA

Collezione diretta da

Emilio Gatti, Francesco Gherardelli, Luigi Radicati, Edoardo Vesentini

Mario Ageno, *Elementi di fisica*

T. M. Apostol, *Calcolo*

Vol. 1 *Analisi I*

Vol. 2 *Geometria*

Vol. 3 *Analisi 2*

Max Born, *Fisica atomica*

Francesco Carassa, *Comunicazioni elettriche*

P. A. M. Dirac, *I principi della meccanica quantistica*

Albert Einstein, *Il significato della relatività*

Enrico Fermi, *Termodinamica*

Bruno Ferretti, *Le radici classiche della meccanica quantica*

Giorgio Franceschetti, *Campi elettromagnetici*

Giovanni Gallavotti, *Meccanica elementare*

Enrico Giusti, *Analisi matematica I*

Enrico Giusti, *Analisi matematica 2*

Werner Heisenberg, *I principi fisici della teoria dei quanti*

Gerhard Herzberg, *Spettri atomici e struttura atomica*

Charles Kittel, *Introduzione alla fisica dello stato solido*

Serge Lang, *Algebra lineare*

Giorgio Letta, *Teoria elementare dell'integrazione*

P. F. Manfredi, Piero Maranesi e Tiziana Tacchi,

L'amplificatore operazionale

Jacob Millman e C. C. Halkias, *Dispositivi e circuiti elettronici*

Jacob Millman e C. C. Halkias, *Microelettronica*

R. S. Muller e T. I. Kamins, *Dispositivi elettronici nei circuiti integrati*

Athanasios Papoulis, *Probabilità, variabili aleatorie e processi stocastici*

Wolfgang Pauli, *Teoria della relatività*

Giovanni Prodi, *Analisi matematica*

Antonio Ruberti e Alberto Isidori, *Teoria dei sistemi*

Walter Rudin, *Analisi reale e complessa*

H. H. Schaefer, *Introduzione alla teoria spettrale*

I. M. Singer e J. A. Thorpe, *Lezioni di topologia elementare e di geometria*

W. V. Smith e P. P. Sorokin, *Il laser*

Bruno Touschek e Giancarlo Rossi, *Meccanica statistica*

© 1970 Editore Boringhieri *società per azioni*
Torino, corso Vittorio Emanuele 86
Stampato in Italia dalla tipografia Gravinese di Torino Gennaio 1984
CL 74-7739-2

Prima edizione 1970
Settima impressione 1984

Titolo originale
Linear Algebra
Addison-Wesley Publishing Company - Reading, Mass. - 1966

Traduzione di Salvatore Ciampa

Indice

Prefazione, 7

1. Vettori in \mathbb{R}^n , 11

- 1. Definizione di punto in un n -spazio
- 2. Vettori applicati
- 3. Prodotto scalare
- 4. Norma di un vettore
- 5. Rette e piani
- 6. Numeri complessi

2. Spazi vettoriali, 35

- 7. Terminologia
- 8. Definizioni
- 9. Basi
- 10. Dimensione di uno spazio vettoriale
- 11. Somme e somme dirette

3. Matrici, 55

- 12. Lo spazio delle matrici
- 13. Equazioni lineari
- 14. Moltiplicazione tra matrici
- 15. Appendice: eliminazione

4. Applicazioni lineari, 76

- 16. Applicazioni
- 17. Applicazioni lineari
- 18. Nucleo e immagine di un'applicazione lineare
- 19. Dimensione del nucleo e dell'immagine
- 20. Composizione di applicazioni lineari

5. Applicazioni lineari e matrici, 103

- 21. Applicazione lineare associata a una matrice
- 22. Matrice associata a un'applicazione lineare
- 23. Composizione di applicazioni lineari

6. Determinanti, 117

- 24. Determinanti del secondo ordine
- 25. Proprietà dei determinanti
- 26. Regola di Cramer
- 27. Esistenza dei determinanti
- 28. Permutazioni
- 29. Unicità
- 30. Determinante della trasposta di una matrice
- 31. Determinante di un prodotto
- 32. Inversa di una matrice
- 33. Determinante di un'applicazione lineare

- 7. Prodotti scalari e ortogonalità, 149**
34. Prodotti scalari 35. Prodotti definiti positivi 36. Basi ortogonali nel caso generale
37. Spazio duale 38. Caratteristica di una matrice e sistemi di equazioni lineari
- 8. Matrici e applicazioni bilineari, 182**
39. Forme bilineari 40. Forme quadratiche 41. Operatori simmetrici 42. Operatori hermitiani
43. Operatori unitari 44. Teorema di Sylvester
- 9. Polinomi e matrici, 213**
45. Polinomi 46. Polinomi di matrici e di applicazioni lineari 47. Autovettori e autovalori
48. Polinomio caratteristico
- 10. Triangolazione delle matrici e delle applicazioni lineari, 232**
49. Esistenza delle triangolazioni 50. Teorema di Hamilton-Cayley 51. Diagona-
lizzazione di applicazioni unitarie
- 11. Il teorema spettrale, 241**
52. Autovettori di applicazioni lineari simmetriche 53. Il teorema spettrale
54. Caso complesso
- 12. Polinomi e decomposizioni primarie, 254**
55. L'algoritmo euclideo 56. Massimo comun divisore 57. Fattorizzazione unica
58. Gli interi 59. Applicazione alla scomposizione di uno spazio vettoriale
60. Il lemma di Schur 61. Sviluppi α -adici di un polinomio
- 13. Prodotti multilinearari, 280**
62. Prodotto tensoriale 63. Isomorfismi di prodotti tensoriali 64. Prodotti alter-
nanti: caso particolare 65. Prodotti alternanti: caso generale 66. Appendice:
lo spazio vettoriale generato da un insieme
- 14. Gruppi, 305**
67. Gruppi ed esempi 68. Semplici proprietà dei gruppi 69. Classi laterali e
sottogruppi normali 70. Gruppi ciclici 71. Gruppi abeliani liberi
- 15. Anelli, 332**
72. Anelli e ideali 73. Omomorfismi 74. Moduli 75. Moduli quoziente
- Appendice 1. Insiemi convessi, 351**
- A. Definizioni B. Iperpiani separanti C. Punti estremi e iperpiani radenti
D. Teorema di Krein-Milman
- Appendice 2. Complementi, 361**
- E. Induzione F. Chiusura algebrica del corpo dei numeri complessi G. Rela-
zioni di equivalenza
- Indice analitico, 369*

Prefazione

Questo libro è concepito come testo per un corso di algebra lineare per i primi anni universitari. Esso contiene materiale sufficiente per un corso annuale: con opportune omissioni, può esser facilmente adattato a un corso semestrale.

Nell'ultimo decennio il corso istituzionale di algebra si è andato identificando sempre più con un corso di algebra lineare. Ciò è dovuto in parte al riconoscimento che questa parte dell'algebra è di acquisizione più immediata di altre parti (perché è meno astratta e, comunque, è motivabile direttamente con la geometria dello spazio) e in parte alla vastità delle sue applicazioni. Ho pertanto creduto opportuno aprire la trattazione con la fondamentale nozione di vettori in uno spazio euclideo reale, dando così uno schema generale per molte delle parti seguenti. A causa delle strette relazioni con l'algebra lineare sono stati inclusi i capitoli sui gruppi e sugli anelli: il gruppo delle applicazioni lineari (o matrici) invertibili e l'anello delle applicazioni lineari di uno spazio vettoriale sono forse tra gli esempi più importanti di gruppi e anelli. In un corso di algebra lineare val la pena di far vedere che uno spazio vettoriale può essere riguardato come un modulo sull'anello dei suoi endomorfismi, ma, d'altra parte, una completa trattazione di questi argomenti, senz'altro possibile, non sembrava conciliarsi con lo spirito generale di tutta l'opera.

La notevole importanza dei prodotti tensoriali, e più ancora di quelli alternanti, nei corsi di analisi matematica ha reso inderogabile l'inserimento di questi argomenti nella presente trattazione, avendo sempre presenti le loro applicazioni. I limitati scopi della trattazione hanno d'altra parte reso possibile un'esposizione concreta e semplice.

L'appendice sugli insiemi convessi vuole continuare il discorso geometrico iniziato nel primo capitolo, supponendo la conoscenza di alcuni fatti fondamentali sulle funzioni continue definite sui compatti, sulla chiusura degli insiemi ecc. L'appendice sulla convessità può essere letta subito dopo il primo capitolo, avendo cura di vedere prima la definizione di applicazione lineare. Vari complementi (tra cui una dimostrazione della chiusura algebrica del corpo complesso) si trovano nell'appendice 2: il loro inserimento in un programma d'insegnamento è lasciato al giudizio dell'insegnante.

S. L.

Algebra lineare

Capitolo 1

Vettori in \mathbb{R}^n

Per fornire una motivazione algebrica e geometrica alle nozioni di cui tratteremo in seguito con maggiore generalità, cominciamo a considerarne in questo capitolo dei casi particolari. Assumiamo nel lettore la conoscenza dei numeri reali: in tutto questo capitolo, con l'eccezione del paragrafo 6, *numero* significherà sempre *numero reale*.

1. DEFINIZIONE DI PUNTO IN UN n -SPAZIO

Fissata un'unità per le lunghezze, i numeri possono essere adoperati per rappresentare i punti di una retta.

Una coppia (ordinata) di numeri (x, y) può essere adoperata per rappresentare un punto di un piano.

Osserviamo poi che una terna di numeri (x, y, z) può servire a rappresentare un punto dello spazio, detto spazio tridimensionale, o 3-spazio. Basta introdurre un nuovo asse, come è illustrato nella figura 1.1.

Invece di scrivere (x, y, z) possiamo anche scrivere (x_1, x_2, x_3) . La retta può essere chiamata un 1-spazio, il piano un 2-spazio, concludendo che un numero rappresenta un punto dell'1-spazio, una coppia di numeri rappresenta un punto del 2-spazio, una terna di numeri rappresenta un punto del 3-spazio.

Sebbene non possiamo disegnare una figura appropriata, nulla ci vieta di considerare quaterne di numeri

$$(x_1, x_2, x_3, x_4),$$

e chiamarle, per definizione, punti del 4-spazio. Analogamente,

potremo considerare una quintupla di numeri come punto del 5-spazio e così continuare con sestuple, settuple ecc.

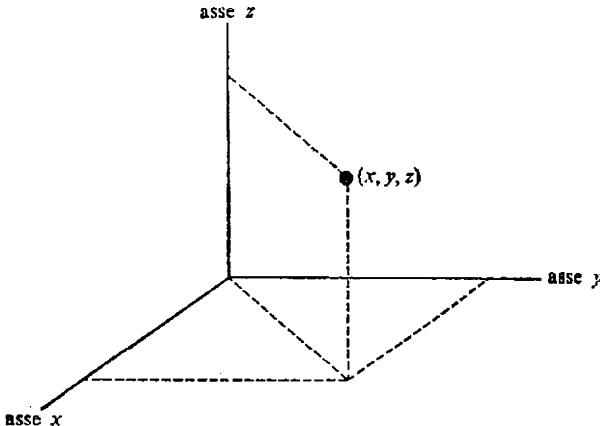


Figura 1.1

Allora, se n è un intero positivo, possiamo senz'altro *definire punto dell' n -spazio* una n -upla di numeri

$$(x_1, x_2, \dots, x_n).$$

Denoteremo ognuna di queste n -uple con una X maiuscola, cercando di riservare le lettere minuscole per denotare numeri e quelle maiuscole per denotare punti. I numeri x_1, \dots, x_n sono chiamati le *coordinate* del punto X .

Vogliamo ora definire un'addizione tra i punti. Siano A, B due punti:

$$A = (a_1, \dots, a_n), \quad B = (b_1, \dots, b_n);$$

definiamo $A + B$ come il punto le cui coordinate sono

$$(a_1 + b_1, \dots, a_n + b_n).$$

Per esempio, nel piano, se $A = (1, 2)$ e $B = (-3, 5)$, allora $A + B = (-2, 7)$. Nel 3-spazio, se $A = (-1, \pi, 3)$ e $B = (\sqrt{2}, 7, -2)$, allora:

$$A + B = (\sqrt{2} - 1, \pi + 7, 1).$$

Inoltre, se c è un numero qualsiasi, *definiamo* cA come il punto

le cui coordinate sono:

$$(ca_1, \dots, ca_n).$$

Se $A = (2, -1, 5)$ e $c = 7$, avremo $cA = (14, -7, 35)$.

Osserviamo che le seguenti uguaglianze sono soddisfatte:

$$1) (A + B) + C = A + (B + C);$$

$$2) A + B = B + A;$$

$$3) c(A + B) = cA + cB;$$

4) se c_1, c_2 sono numeri, allora

$$(c_1 + c_2)A = c_1A + c_2A \quad \text{e} \quad (c_1c_2)A = c_1(c_2A);$$

5) indicato con $O = (0, \dots, 0)$ il punto le cui coordinate sono tutte nulle, per ogni A si ha $O + A = A + O = A$;

6) $1 \cdot A = A$; se con $-A$ si denota la n -upla $(-1)A$, allora

$$A + (-A) = O.$$

[Invece di scrivere $A + (-B)$, scriveremo, solitamente, $A - B$.]

Tutte queste proprietà si provano facilmente; suggeriamo al lettore di verificarle su qualche esempio.

Diamo per esempio la dimostrazione della proprietà 3.

Sia $A = (a_1, \dots, a_n)$ e $B = (b_1, \dots, b_n)$, allora

$$A + B = (a_1 + b_1, \dots, a_n + b_n)$$

e

$$\begin{aligned} c(A + B) &= (c(a_1 + b_1), \dots, c(a_n + b_n)) = \\ &= (ca_1 + cb_1, \dots, ca_n + cb_n) = \\ &= cA + cB, \end{aligned}$$

L'ultimo passaggio essendo giustificato dalla definizione di addizione tra n -uple.

Le altre dimostrazioni sono lasciate come esercizio.

Attenzione. Non confondere il numero 0 e la n -upla $(0, \dots, 0)$. Solitamente denotiamo questa n -upla con O e la chiamiamo zero, in pratica ciò non causa alcuna confusione.

Daremo ora un'interpretazione geometrica, nel piano, dell'addizione e della moltiplicazione per numeri (il lettore può, contemporaneamente, raffigurarsi la stessa situazione nel 3-spazio).

Sia, ad esempio, $A = (2, 3)$ e $B = (-1, 1)$. Allora $A + B = (1, 4)$.

Questa operazione viene rappresentata graficamente da un parallelogramma (vedi fig. 1.2).

Altro esempio: sia $A = (3, 1)$ e $B = (1, 2)$, allora, $A + B = (4, 3)$. Nella figura 1.3, possiamo di nuovo notare che la rappresentazione geometrica dell'addizione è un parallelogramma.

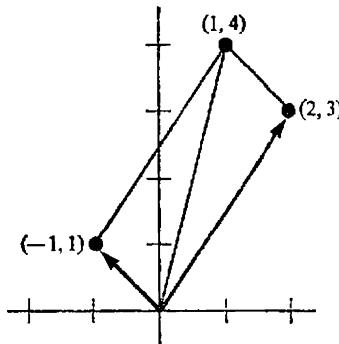


Figura 1.2

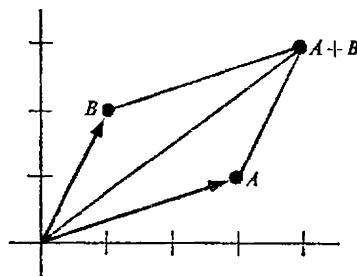
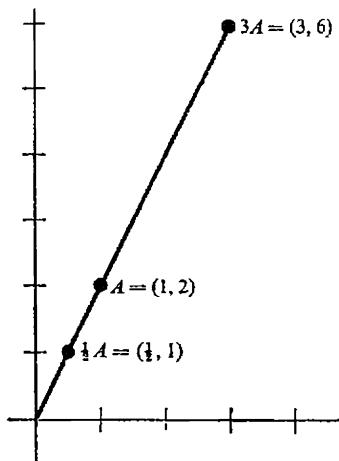
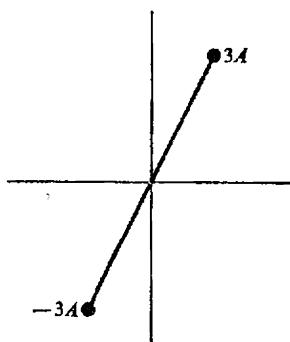


Figura 1.3

Quale sarà la rappresentazione geometrica della moltiplicazione per un numero? Sia $A = (1, 2)$ e $c = 3$, allora $cA = (3, 6)$ (vedi fig. 1.4a).



(a)



(b)

Figura 1.4

La moltiplicazione per 3 produce un allungamento di A di 3 volte; analogamente, $\frac{1}{2}A$ significa allungare A mezza volta, cioè dimezzarlo. In generale, se t è un numero, $t > 0$, interpretiamo tA come il punto che si trova sulla stessa semiretta passante per l'origine su cui si trova A , t volte più distante dall'origine.

La moltiplicazione per un numero negativo inverte la direzione (cioè, fa cambiare semiretta). Nella figura 1.4b, per esempio, è rappresentato il punto $-3A$.

Esercizi

Trovare $A + B$, $A - B$, $3A$, $-2B$ in ognuno dei seguenti casi:

1. $A = (2, -1)$, $B = (-1, 1)$.
2. $A = (-1, 3)$, $B = (0, 4)$.
3. $A = (2, -1, 5)$, $B = (-1, 1, 1)$.
4. $A = (-1, -2, 3)$, $B = (-1, 3, -4)$.
5. $A = (\pi, 3, -1)$, $B = (2\pi, -3, 7)$.
6. $A = (15, -2, 4)$, $B = (\pi, 3, -1)$.
7. Su un foglio di carta quadrettata disegnare i punti degli esercizi 1, 2, 3, 4.
8. Siano A , B come nell'esercizio 1. Su un foglio di carta quadrettata disegnare i punti $A + 2B$, $A + 3B$, $A - 2B$, $A - 3B$, $A + \frac{1}{2}B$.

2. VETTORI APPLICATI

Chiameremo *vettore applicato* una coppia (ordinata) di punti che denoteremo con \overrightarrow{AB} . (Attenzione: questo *non* è un prodotto.) Graficamente lo denoteremo con una freccia da A a B che sono chiamati, rispettivamente, l'*origine* e la *fine* del vettore applicato (vedi fig. 2.1).

Come si ottengono le coordinate di B da quelle di A ? Osserviamo dapprima che, nel piano,

$$b_1 = a_1 + (b_1 - a_1)$$

e

$$b_2 = a_2 + (b_2 - a_2).$$

Questo significa che:

$$B = A + (B - A).$$

Siano \vec{AB} e \vec{CD} due vettori applicati. Diremo che essi sono *equivalenti* se $B - A = D - C$. Ogni vettore applicato \vec{AB} ha un equivalente che inizia nell'origine, \vec{AB} infatti è equivalente a $\vec{O(B-A)}$. E questo, naturalmente, è l'unico vettore applicato equivalente ad \vec{AB} che inizia nell'origine. Se si osserva la figura

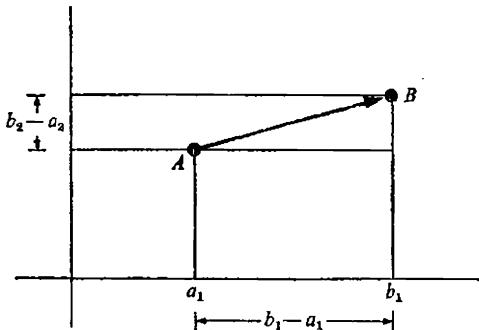


Figura 2.1

della regola del parallelogramma nel piano, risulta chiaro che l'equivalenza di due vettori applicati può essere interpretata geometricamente dicendo che i segmenti determinati dalle due coppie di punti hanno uguale lunghezza e hanno la stessa "direzione".

Nella figura 2.2 sono tracciati i vettori applicati $\vec{O(B-A)}$ e \vec{AB} .

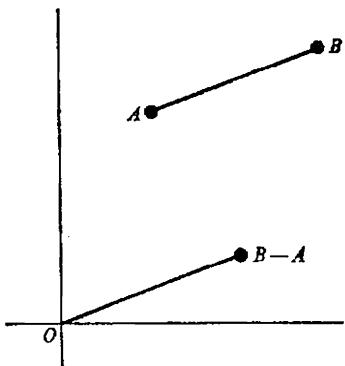


Figura 2.2

Si consideri ora un vettore applicato \vec{OC} che inizia nell'origine, diciamo allora che il vettore è *applicato nell'origine*. Considerato

un qualsiasi vettore applicato \vec{AB} diremo che il vettore è *applicato in A*.

Osservato che un vettore applicato nell'origine è perfettamente determinato dalla sua fine, chiameremo una n -upla di numeri un punto o un vettore a seconda dell'interpretazione che vogliamo dare.

Due vettori applicati \vec{AB} e \vec{PQ} si dicono *parallelî* se esiste un numero $c \neq 0$ tale che $B - A = c(Q - P)$. Diremo che essi hanno la *stessa direzione* oppure *direzioni opposte* secondo se il numero c è, rispettivamente, maggiore o minore di zero. In modo analogo, ogni definizione relativa alle n -uple di numeri può essere trasferita ai vettori applicati. Per esempio, nel paragrafo successivo,

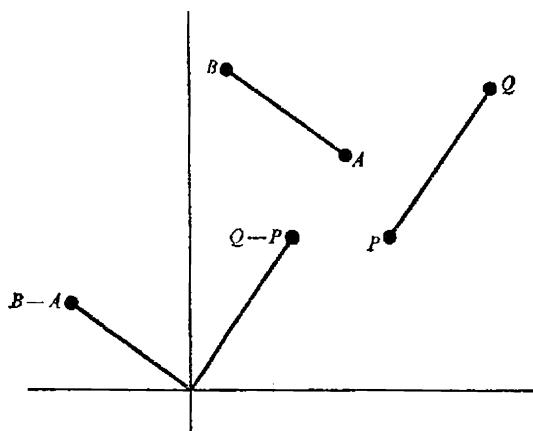


Figura 2.3

definiremo la perpendicolarità tra n -uple di numeri. Allora potremo anche dire che due vettori applicati \vec{AB} e \vec{PQ} sono perpendicolari se $B - A$ e $Q - P$ lo sono. Nella figura 2.3 è rappresentata questa situazione nel piano.

3. PRODOTTO SCALARE

Conveniamo di considerare in tutta la discussione vettori nello stesso spazio n -dimensionale.

Siano $A = (a_1, \dots, a_n)$ e $B = (b_1, \dots, b_n)$ due vettori. Definiamo

come loro *prodotto scalare* $A \cdot B$:

$$a_1 b_1 + \dots + a_n b_n.$$

Notiamo che questo prodotto è un *numero*. Per esempio, se

$$A = (1, 3, -2) \quad \text{e} \quad B = (-1, 4, -3),$$

allora

$$A \cdot B = -1 + 12 + 6 = 17.$$

Per ora non diamo nessuna interpretazione geometrica al prodotto scalare; la daremo più tardi. Troviamo dapprima alcune sue proprietà importanti. Quelle principali sono:

PS 1. *Se A, B sono vettori, allora $A \cdot B = B \cdot A$.*

PS 2. *Se A, B, C sono vettori, allora*

$$A \cdot (B + C) = A \cdot B + A \cdot C = (B + C) \cdot A.$$

PS 3. *Se x è un numero, allora*

$$(xA) \cdot B = x(A \cdot B) \quad \text{e} \quad A \cdot (xB) = x(A \cdot B).$$

PS 4. *Se $A = O$ è il vettore nullo, allora $A \cdot A = 0$; in ogni altro caso abbiamo $A \cdot A > 0$.*

Diamo ora la dimostrazione di queste proprietà.

Nei riguardi della prima si ha

$$a_1 b_1 + \dots + a_n b_n = b_1 a_1 + \dots + b_n a_n,$$

perché per ogni coppia di numeri a, b , abbiamo $ab = ba$.

Rimane così provata la prima proprietà.

Per PS 2, sia $C = (c_1, \dots, c_n)$. Allora

$$B + C = (b_1 + c_1, \dots, b_n + c_n)$$

e

$$\begin{aligned} A \cdot (B + C) &= a_1(b_1 + c_1) + \dots + a_n(b_n + c_n) = \\ &= a_1 b_1 + a_1 c_1 + \dots + a_n b_n + a_n c_n. \end{aligned}$$

Riordinando i termini si può scrivere

$$a_1 b_1 + \dots + a_n b_n + a_1 c_1 + \dots + a_n c_n,$$

che non è altro che $A \cdot B + A \cdot C$. Si ottiene così l'uguaglianza da provare.

La dimostrazione della proprietà PS 3 è lasciata per esercizio. Infine, nei riguardi di PS 4, osserviamo che se il vettore A ha una coordinata a_i diversa da zero, allora nel prodotto scalare

$$A \cdot A = a_1^2 + \dots + a_n^2$$

vi è un termine $a_i^2 > 0$; essendo ogni altro termine maggiore o uguale a zero, segue che la somma è maggiore di zero, come volevasi dimostrare.

Nei servirci in seguito dei vettori adopereremo soltanto le ordinarie proprietà dell'addizione e della moltiplicazione per numeri e le quattro proprietà del prodotto scalare. Di queste sarà fatta una più approfondita discussione più tardi. Per il momento osserviamo che ci sono altri oggetti a noi familiari e che possono essere addizionati, sottratti, moltiplicati per numeri; per esempio le funzioni continue definite in un intervallo $[a, b]$ (vedi esercizio 5).

Invece di scrivere $A \cdot A$ per il prodotto scalare di un vettore per sé stesso è certe volte conveniente scrivere A^2 . (Si tenga presente che questo è l'unico caso in cui adoperiamo una tale notazione: A^3 , per esempio, non ha significato.) Per esercizio il lettore verifichi le seguenti identità:

$$(A + B)^2 = A^2 + 2A \cdot B + B^2,$$

$$(A - B)^2 = A^2 - 2A \cdot B + B^2.$$

Diremo che due vettori A, B sono *perpendicolari* (oppure *ortogonali*) se il prodotto scalare $A \cdot B$ è uguale a zero. Per il momento non è ancora chiaro che, per i vettori del piano, questa definizione coincide con la nozione intuitiva di perpendicolarità geometrica. Nel prossimo paragrafo ci convinceremo del susseguente di questa coincidenza.

Esercizi

1. Calcolare $A \cdot A$ per ognuna delle n -uple degli esercizi da 1 a 6 del paragrafo 1.
2. Calcolare $A \cdot B$ per le n -uple degli esercizi da 1 a 6 del paragrafo 1.
3. Adoperando soltanto le quattro proprietà del prodotto scalare, dimostrare in tutti i dettagli le uguaglianze che danno $(A + B)^2$ e $(A - B)^2$.

4. Quali, tra le seguenti, sono coppie di vettori perpendicolari?
- a) $(1, -1, 1)$ e $(2, 1, 5)$.
 - b) $(1, -1, 1)$ e $(2, 3, 1)$.
 - c) $(-5, 2, 7)$ e $(3, -1, 2)$.
 - d) $(\pi, 2, 1)$ e $(2, -\pi, 0)$.

5. Considerate le funzioni continue nell'intervallo $[-1, 1]$, si definisca il prodotto scalare tra due tali funzioni f, g il numero

$$\int_{-1}^{+1} f(x)g(x) dx.$$

Si denoti questo integrale con $\langle f, g \rangle$. Verificare che valgono le quattro proprietà del prodotto scalare; dimostrare, cioè, che

- PS 1. $\langle f, g \rangle = \langle g, f \rangle$.
- PS 2. $\langle f, g + h \rangle = \langle f, g \rangle + \langle f, h \rangle$.
- PS 3. $\langle cf, g \rangle = c\langle f, g \rangle$.
- PS 4. Se $f = 0$, allora $\langle f, f \rangle = 0$ e, se $f \neq 0$, allora $\langle f, f \rangle > 0$.
- 6. Se $f(x) = x$ e $g(x) = x^2$, quanto valgono $\langle f, f \rangle$, $\langle g, g \rangle$ e $\langle f, g \rangle$?

7. Si considerino le funzioni continue nell'intervallo $[-\pi, \pi]$, e si definisca per queste funzioni un prodotto scalare analogo al precedente estendendo l'integrale a questo intervallo. Provare che, se m ed n sono interi, le funzioni $\sin mx$ e $\cos mx$ risultano ortogonali rispetto a questo prodotto scalare.

8. Dimostrare che se un vettore A è perpendicolare ad ogni vettore X , allora $A = O$.

4. NORMA DI UN VETTORE

La diseguaglianza che ora proveremo è chiamata *diseguaglianza di Schwarz* ed è di fondamentale importanza nella teoria dei vettori.

TEOREMA 1 *Siano A, B due vettori, allora*

$$(A \cdot B)^2 \leq (A \cdot A)(B \cdot B).$$

Dimostrazione. Sia $x = B \cdot B$ e $y = -A \cdot B$; da PS 4 segue allora

$$0 \leq (xA + yB) \cdot (xA + yB).$$

Sviluppando il secondo membro di questa diseguaglianza si ha

$$0 \leq x^2(A \cdot A) + 2xy(A \cdot B) + y^2(B \cdot B).$$

Sostituendo i valori di x e y si ottiene

$$0 < (B \cdot B)^2 (A \cdot A) - 2(B \cdot B)(A \cdot B)^2 + (A \cdot B)^2 (B \cdot B).$$

Se $B = O$, la diseguaglianza da provare è ovvia essendo nulli entrambi i membri. Se $B \neq O$, allora $B \cdot B \neq 0$ e possiamo quindi dividere l'ultima espressione per $B \cdot B$ ottenendo

$$0 < (A \cdot A)(B \cdot B) - (A \cdot B)^2.$$

La dimostrazione si conclude portando il termine $-(A \cdot B)^2$ nel primo membro.

Noi definiamo *norma* (o *modulo*), oppure *lunghezza*, del vettore A , e la denotiamo con $\|A\|$, il numero

$$\|A\| = \sqrt{A \cdot A}.$$

Osserviamo che, poiché $A \cdot A > 0$, si può estrarre la radice quadrata; si vede subito che se $A \neq O$, anche $\|A\| \neq 0$.

Esprimendo il modulo mediante coordinate si ottiene

$$\|A\| = \sqrt{a_1^2 + \dots + a_n^2},$$

e si vede quindi che, quando $n = 2$ oppure $n = 3$, il modulo coincide con la nostra intuitiva nozione di lunghezza (dedotta dal teorema di Pitagora).

Con questa nostra definizione, la diseguaglianza del teorema 1, estraendo la radice quadrata da entrambi i membri, può venire così riscritta:

$$|(A \cdot B)| < \|A\| \|B\|.$$

La useremo sotto questa forma nella dimostrazione del teorema seguente.

TEOREMA 2 *Siano A , B vettori, allora*

$$\|A + B\| < \|A\| + \|B\|.$$

Dimostrazione. Osserviamo dapprima che i due membri di questa diseguaglianza sono positivi o nulli. È quindi sufficiente dimostrare la diseguaglianza per i loro quadrati, cioè dimostrare che

$$(A + B) \cdot (A + B) < (\|A\| + \|B\|)^2.$$

A tal fine osserviamo che

$$(A + B) \cdot (A + B) = A \cdot A + 2A \cdot B + B \cdot B.$$

Questa quantità risulta, per quanto abbiamo già dimostrato, minore o uguale a

$$\|A\|^2 + 2\|A\|\|B\| + \|B\|^2,$$

che non è altro che

$$(\|A\| + \|B\|)^2.$$

Il nostro teorema è così dimostrato.

Il teorema 2 è conosciuto come la *disegualanza triangolare* (vedi esercizio 11).

TEOREMA 3 *Sia x un numero, allora*

$$\|xA\| = |x|\|A\|$$

(*valore assoluto di x moltiplicato per la lunghezza di A*).

Dimostrazione. Per definizione si ha

$$\|xA\|^2 = (xA) \cdot (xA),$$

che è uguale a

$$x^2(A \cdot A)$$

per le proprietà del prodotto scalare. Per concludere la dimostrazione, basta estrarre la radice quadrata.

Diremo che U è un *vettore unità* se U è un vettore di lunghezza 1. Per ogni vettore A , se $a = \|A\| \neq 0$,

$$\frac{1}{a}A$$

è un vettore unità perché

$$\left\| \frac{1}{a}A \right\| = \frac{1}{a}a = 1.$$

Diremo che due vettori A, B (nessuno dei quali sia nullo) hanno la stessa direzione se esiste un numero $c > 0$ tale che $cA = B$. Da questa definizione segue che il vettore

$$\frac{1}{\|A\|}A$$

è un vettore unità che ha la stessa direzione di A (supposto che $A \neq O$). È appena il caso di menzionare che due vettori A, B (nessuno dei quali nullo) hanno *direzioni opposte* se esiste un numero $c < 0$ tale che $cA = B$.

Siano A, B due n -uple. Noi definiamo *distanza* tra A e B il numero $\|A - B\| = \sqrt{(A - B) \cdot (A - B)}$. Questa definizione coincide con la nostra intuizione geometrica quando A, B sono punti del piano (vedi fig. 4.1). Siamo ora in condizioni di giustificare

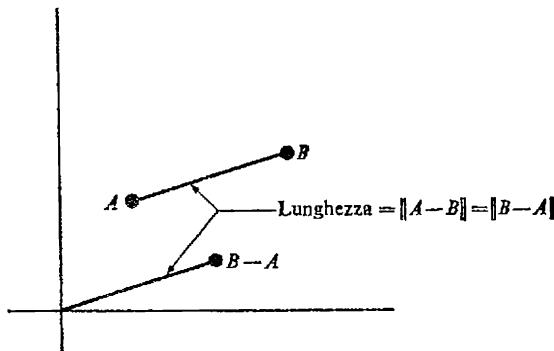


Figura 4.1

la nostra definizione di perpendicolarità. Dati A, B nel piano, la condizione

$$\|A + B\| = \|A - B\|$$

(illustrata nella fig. 4.2b) traduce la proprietà geometrica della perpendicolarità di A con B . Questa condizione è equivalente a

$$(A + B) \cdot (A + B) = (A - B) \cdot (A - B)$$

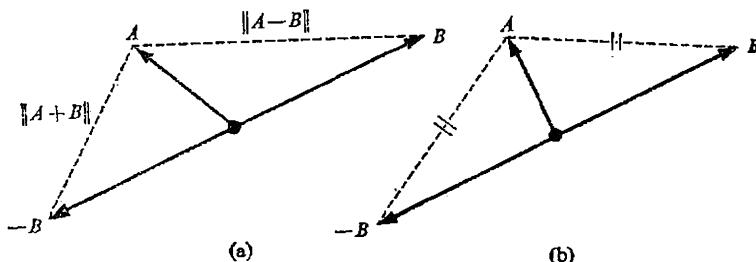


Figura 4.2

(innalzare al quadrato ognuno dei due membri) e a sua volta, svolgendo i calcoli, questa uguaglianza è equivalente a

$$A \cdot A + 2A \cdot B + B \cdot B = A \cdot A - 2A \cdot B + B \cdot B.$$

Eliminando i termini simili nei due membri, otteniamo ancora

$$4A \cdot B = 0$$

e anche

$$A \cdot B = 0.$$

Siano A, B due vettori e sia $B \neq O$. Supponiamo di poter trovare un numero c tale che $A - cB$ sia perpendicolare a B , in altre parole

$$(A - cB) \cdot B = 0;$$

abbiamo allora

$$A \cdot B = cB \cdot B,$$

da cui

$$c = \frac{A \cdot B}{B \cdot B}.$$

Quindi il numero c è univocamente determinato dalla nostra condizione di perpendicolarità. Viceversa il numero c ora determinato è tale che $(A - cB) \cdot B = 0$.

Definiamo cB come la *proiezione di A su B* . Se B è un vettore unità, allora, più semplicemente, abbiamo

$$c = A \cdot B.$$

La nostra costruzione ha un'immediata interpretazione nel piano e questa dà, a sua volta, un'interpretazione geometrica al prodotto scalare (vedi fig. 4.3). Infatti, considerato un vettore $A \neq O$, si consideri l'angolo θ tra A e B . Dalla geometria piana abbiamo

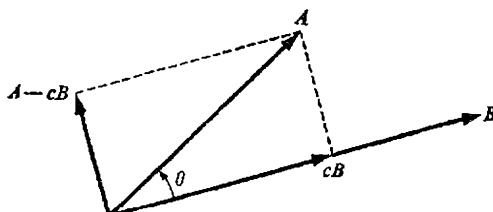


Figura 4.3

allora che

$$\cos \theta = \frac{c\|B\|}{\|A\|},$$

e sostituendo a c il valore prima ottenuto,

$$A \cdot B = \|A\|\|B\| \cos \theta.$$

Sappiamo che, per il teorema 1, nell' n -spazio il numero

$$\frac{A \cdot B}{\|A\|\|B\|}$$

ha valore assoluto non maggiore di 1. Di conseguenza,

$$-1 \leq \frac{A \cdot B}{\|A\|\|B\|} \leq 1,$$

e quindi esiste un unico angolo θ tale che $0 < \theta < \pi$ e inoltre tale che

$$\cos \theta = \frac{A \cdot B}{\|A\|\|B\|}.$$

Definiamo questo angolo come l'*angolo tra A e B*.

Esercizi

1. Trovare la lunghezza del vettore A negli esercizi da 1 a 6 del paragrafo 1.
2. Trovare la lunghezza del vettore B negli esercizi da 1 a 6 del paragrafo 1.
3. Trovare la proiezione di A su B negli esercizi da 1 a 6 del paragrafo 1.
4. Trovare la proiezione di B su A in questi stessi esercizi.
5. Nell'esercizio 6 del paragrafo 3, trovare la proiezione di f su g e la proiezione di g su f , adoperando la definizione di proiezione data nel testo (senza far riferimento alle coordinate).
6. Trovare il modulo delle funzioni $\sin 3x$ e $\cos x$ rispetto al prodotto scalare definito dall'integrale esteso all'intervallo $[-\pi, \pi]$.
7. Trovare il modulo della funzione costante 1 nell'intervallo $[-\pi, \pi]$.
8. Trovare il modulo della funzione costante 1 nell'intervallo $[-1, +1]$.
9. Siano A_1, \dots, A_r vettori non nulli e mutuamente perpendicolari, in

altre parole $A_i \cdot A_j = 0$ se $i \neq j$. Siano c_1, \dots, c_r numeri tali che

$$c_1 A_1 + \dots + c_r A_r = 0.$$

Dimostrare che tutti i numeri c_i sono nulli.

10. Siano A, B due vettori non nulli nell' n -spazio. Sia θ il loro angolo. Supposto $\cos \theta = 1$, dimostrare che A e B hanno la stessa direzione. Se invece $\cos \theta = -1$, dimostrare che A e B hanno direzioni opposte.

11. Siano A, B due vettori nell' n -spazio, si denoti con $d(A, B)$ la distanza tra A e B , cioè $d(A, B) = \|B - A\|$. Dimostrare che $d(A, B) = d(B, A)$ e che, comunque si prendano tre vettori A, B, C , si ha

$$d(A, B) \leq d(A, C) + d(B, C).$$

5. RETTE E PIANI

Definiamo equazione parametrica della retta passante per un punto P e avente la direzione di un vettore $A \neq O$ la seguente

$$X = P + tA,$$

dove t assume tutti i valori reali (vedi fig. 5.1).

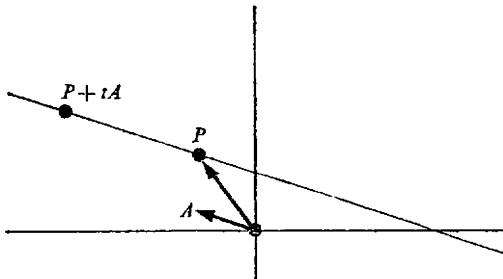


Figura 5.1

Consideriamo ora il piano e scriviamo le coordinate di un punto X nella forma (x, y) . Sia $P = (p, q)$ e $A = (a, b)$. Allora, riferendoci alle coordinate, possiamo scrivere

$$x = p + ta, \quad y = q + tb.$$

L'equazione solita in x e y si ottiene eliminando il parametro t .

Per esempio, sia $P = (2, 1)$ e $A = (-1, 5)$. L'equazione parametrica della retta per P nella direzione di A fornisce

$$x = 2 - t, \quad y = 1 + 5t.$$

Moltiplicando la prima equazione per 5 e sommando si ottiene

$$5x + y = 11,$$

che è la forma familiare dell'equazione.

In casi di spazi con maggior numero di dimensioni, *non* possiamo eliminare t in questo modo e l'equazione parametrica è l'unica disponibile per descrivere una retta.

Tuttavia con un'equazione singola, come quella della retta, possiamo descrivere un piano. Procediamo come segue.

Siano P un punto e N un vettore non nullo. Noi definiamo *iperpiano* passante per P e perpendicolare a N l'insieme di tutti i punti X tali che $X - P$ è perpendicolare a N , quindi

$$(X - P) \cdot N = 0,$$

che possiamo anche scrivere

$$X \cdot N = P \cdot N.$$

La figura 5.2 riporta una tipica situazione nel 3-spazio.

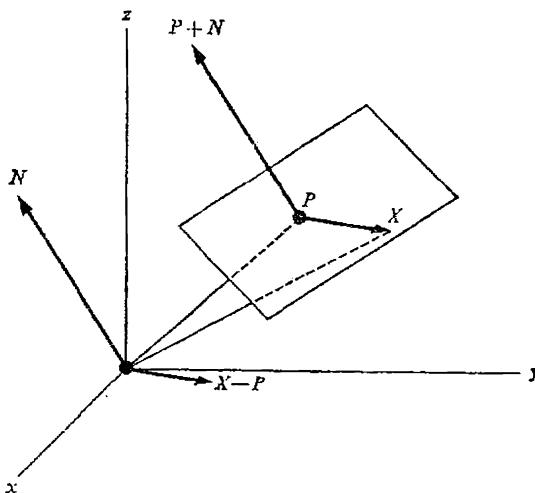


Figura 5.2

Invece di dire che N è perpendicolare al piano, si dice anche che N è *normale* al piano.

Se t è un numero diverso da zero, l'insieme dei punti X tali

che

$$(X - P) \cdot N = 0$$

coincide con l'insieme dei punti X tali che

$$(X - P) \cdot tN = 0.$$

Possiamo quindi dire che il nostro piano è quello passante per P e perpendicolare alla retta nella direzione di N . Per trovare l'equazione del piano possiamo usare qualunque vettore tN (con $t \neq 0$) invece di N .

Nel 3-spazio otteniamo un piano ordinario. Per esempio, siano $P = (2, 1, -1)$ e $N = (-1, 1, 3)$. Allora l'equazione del piano passante per P e perpendicolare a N è

$$-x + y + 3z = -2 + 1 - 3$$

oppure

$$-x + y + 3z = -4.$$

Osserviamo che nel 2-spazio, con $X = (x, y)$, siamo condotti all'equazione di una retta in senso ordinario. Per esempio, l'equazione della retta passante per $(4, -3)$ e perpendicolare a $(-5, 2)$ è

$$-5x + 2y = -20 - 6 = -26.$$

Siamo ora in grado di interpretare i coefficienti $(-5, 2)$ di x e di y di questa equazione. Essi danno luogo a un vettore perpendicolare alla retta. In ogni equazione

$$ax + by = c$$

il vettore (a, b) è perpendicolare alla retta determinata dall'equazione. Analogamente, nel 3-spazio, il vettore (a, b, c) è perpendicolare al piano determinato dall'equazione

$$ax + by + cz = d.$$

Due vettori A, B si dicono *parallel*i se esiste un numero $c \neq 0$ tale che $cA = B$. Due rette si dicono parallele se, presi due punti distinti P_1, Q_1 sulla prima retta e P_2, Q_2 sulla seconda retta, i vettori

$$P_1 - Q_1 \quad \text{e} \quad P_2 - Q_2$$

sono paralleli.

Due piani si dicono *paralleli* (nel 3-spazio) se i vettori ad essi rispettivamente normali sono paralleli. I piani sono detti *perpendicolari* se i loro vettori normali lo sono. L'*angolo* tra i due piani è definito come l'angolo tra i loro vettori normali.

Esempio. Trovare il coseno dell'angolo tra due piani

$$2x - y + z = 0,$$

$$x + 2y - z = 1.$$

Il coseno da trovare è quello dell'angolo tra $(2, -1, 1)$ e $(1, 2, -1)$ ed è quindi uguale a $-\frac{1}{6}$.

Esercizi

Trovare un'equazione parametrica della retta passante per i seguenti punti:

1. $(1, 1, -1)$ e $(-2, 1, 3)$. 2. $(-1, 5, 2)$ e $(3, -4, 1)$.

Trovare l'equazione della retta nel 2-spazio perpendicolare ad A e passante per P , per i seguenti valori di A e P :

3. $A = (1, -1)$, $P = (-5, 3)$. 4. $A = (-5, 4)$, $P = (3, 2)$.

5. Dimostrare che le rette $3x - 5y = 1$, $2x + 3y = 5$ non sono perpendicolari.

6. Quali tra le seguenti sono coppie di rette perpendicolari?

- a) $3x - 5y = 1$ e $2x + y = 2$.
- b) $2x + 7y = 1$ e $x - y = 5$.
- c) $3x - 5y = 1$ e $5x + 3y = 7$.
- d) $-x + y = 2$ e $x + y = 9$.

7. Trovare l'equazione del piano perpendicolare al vettore N e passante per il punto P , dove:

- a) $N = (1, -1, 3)$, $P = (4, 2, -1)$.
- b) $N = (-3, -2, 4)$, $P = (2, \pi, -5)$.
- c) $N = (-1, 0, 5)$, $P = (2, 3, 7)$.

8. Trovare l'equazione del piano passante per i tre punti seguenti:

- a) $(2, 1, 1)$, $(3, -1, 1)$, $(4, 1, -1)$.
- b) $(-2, 3, -1)$, $(2, 2, 3)$, $(-4, -1, 1)$.
- c) $(-5, -1, 2)$, $(1, 2, -1)$, $(3, -1, 2)$.

9. Trovare un vettore perpendicolare a $(1, 2, -3)$ e $(2, -1, 3)$ e un altro vettore perpendicolare a $(-1, 3, 2)$ e $(2, 1, 1)$.

10. Siano P il punto $(1, 2, 3, 4)$ e Q il punto $(4, 3, 2, 1)$. Sia A il vettore $(1, 1, 1, 1)$. Sia L la retta passante per P e parallela ad A .

a) Dato un punto X sulla retta L , determinare la distanza di X da Q (in funzione del parametro t).

b) Dimostrare che esiste un unico punto X_0 sulla retta L avente distanza minima da Q e che questa distanza minima è $2\sqrt{5}$.

c) Dimostrare che il vettore $X_0 - Q$ è perpendicolare alla retta L .

11. Siano P il punto $(1, -1, 3, 1)$ e Q il punto $(1, 1, -1, 2)$. Sia A il vettore $(1, -3, 2, 1)$. Risolvere con questi dati lo stesso problema dell'esercizio precedente, mostrando che, in questo caso, la distanza minima è $\sqrt{146/15}$.

12. Determinare un vettore parallelo alla retta di intersezione dei due piani

$$2x - y + z = 1, \quad 3x + y + z = 2.$$

13. Risolvere lo stesso problema per i piani

$$2x + y + 5z = 2, \quad 3x - 2y + z = 3.$$

14. Trovare un'equazione parametrica di ognuna delle rette di intersezione dei piani degli esercizi 12 e 13.

15. Trovare il coseno dell'angolo tra i seguenti piani

a) $x + y + z = 1,$ b) $2x + 3y - z = 2,$
 $x - y - z = 5.$ $x - y + z = 1.$

c) $x + 2y - z = 1,$ d) $2x + y + z = 3,$
 $-x + 3y + z = 2.$ $-x - y + z = \pi.$

16. Sia $X \cdot N = P \cdot N$ l'equazione di un piano nel 3-spazio. Sia Q un punto non appartenente al piano. Dimostrare che esiste un unico numero t tale che $Q + tN$ giaccia nel piano (cioè, soddisfi l'equazione del piano). Esprimere questo numero t mediante P, Q, N .

17. Sia $Q = (1, -1, 2)$, $P = (1, 3, -2)$, $N = (1, 2, 2)$. Trovare il punto comune alla retta per P nella direzione di N e al piano per Q perpendicolare a N .

18. Siano P e Q due punti e N un vettore nel 3-spazio. Sia P' il punto comune alla retta per P nella direzione di N e al piano per Q perpendicolare a N . Definiamo la *distanza* tra P e questo piano come la distanza tra P e P' . Determinare questa distanza nel caso in cui

$$P = (1, 3, 5), \quad Q = (-1, 1, 7), \quad N = (-1, 1, -1).$$

19. Sia $P = (1, 3, 5)$, $A = (-2, 1, 1)$. Determinare l'intersezione della retta per P nella direzione di A con il piano

$$2x + 3y - z = 1.$$

20. Determinare la distanza tra il punto $(1, 1, 2)$ e il piano

$$3x + y - 5z = 2.$$

21. Sia $P = (1, 3, -1)$, $Q = (-4, 5, 2)$. Determinare le coordinate dei seguenti punti: a) punto medio del segmento PQ ; b) i due punti su questo segmento situati, a partire da P , a un terzo e due terzi della distanza da P verso Q .

22. Determinare la formula generale per le coordinate del punto medio del segmento PQ passante per due punti qualsiasi P e Q nell' n -spazio.

6. NUMERI COMPLESSI

I numeri complessi sono un insieme di oggetti che possono essere addizionati e moltiplicati; somme e prodotti di numeri complessi sono ancora numeri complessi e inoltre sono soddisfatte le seguenti condizioni:

- 1) Ogni numero reale è un numero complesso, e, se α, β sono numeri reali, la loro somma e il loro prodotto come numeri complessi coincidono, rispettivamente, con la loro somma e il loro prodotto come numeri reali.
- 2) Esiste un numero complesso, denotato con i , tale che $i^2 = -1$.
- 3) Ogni numero complesso può essere scritto, in modo unico, nella forma $a + bi$, a e b essendo numeri reali.
- 4) Le usuali proprietà aritmetiche dell'addizione e della moltiplicazione continuano a valere: ne diamo qui di seguito l'elenco.

Se α, β, γ sono numeri complessi, allora

$$(\alpha\beta)\gamma = \alpha(\beta\gamma) \quad \text{e} \quad (\alpha + \beta) + \gamma = \alpha + (\beta + \gamma).$$

Si ha $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ e $(\beta + \gamma)\alpha = \beta\alpha + \gamma\alpha$.

Si ha $\alpha\beta = \beta\alpha$ e $\alpha + \beta = \beta + \alpha$.

Se 1 è il numero reale uno, allora $1\alpha = \alpha$.

Se 0 è il numero reale zero, allora $0\alpha = 0$.

Si ha, infine, $\alpha + (-1)\alpha = 0$.

Mostreremo ora alcune conseguenze di queste proprietà. Ad ogni numero complesso $\alpha + bi$ associamo nel piano il vettore (a, b) . Siano $\alpha = a_1 + a_2 i$ e $\beta = b_1 + b_2 i$ due numeri complessi, allora

$$\alpha + \beta = a_1 + b_1 + (a_2 + b_2)i.$$

Si può quindi dire che l'addizione tra i numeri complessi si esegue « componente per componente » e che essa corrisponde all'addizione dei vettori nel piano. Per esempio

$$(2 + 3i) + (-1 + 5i) = 1 + 8i.$$

Nella moltiplicazione dei numeri complessi l'uguaglianza $i^2 = -1$ è utile per semplificare il prodotto e per scriverlo nella forma $a + bi$. Per esempio, sia $\alpha = 2 + 3i$, $\beta = 1 - i$. Allora

$$\begin{aligned}\alpha\beta &= (2 + 3i)(1 - i) = 2(1 - i) + 3i(1 - i) \\ &= 2 - 2i + 3i - 3i^2 \\ &= 2 + i - 3(-1) \\ &= 2 + 3 + i \\ &= 5 + i.\end{aligned}$$

Sia $\alpha = a + bi$ un numero complesso: denotiamo con $\bar{\alpha}$ il numero complesso $a - bi$. Quindi se $\alpha = 2 + 3i$, allora $\bar{\alpha} = 2 - 3i$. Il numero complesso $\bar{\alpha}$ è chiamato il *conjugato* di α . Si vede subito che

$$\alpha\bar{\alpha} = a^2 + b^2.$$

Interpretando i numeri complessi come vettori, vediamo che $\alpha\bar{\alpha}$ è il quadrato della distanza del punto (a, b) dall'origine.

Vediamo ora un'altra importante proprietà dei numeri complessi che ci permetterà la divisione per i numeri complessi non nulli.

Se $\alpha = a + bi$ è un numero complesso non nullo e se poniamo

$$\lambda = \frac{\bar{\alpha}}{a^2 + b^2}$$

allora $\alpha\lambda = \lambda\alpha = 1$.

La dimostrazione di questo fatto è un'immediata conseguenza della regola di moltiplicazione tra i numeri complessi, infatti

$$\alpha \frac{\bar{\alpha}}{a^2 + b^2} = \frac{\alpha\bar{\alpha}}{a^2 + b^2} = 1.$$

Il numero λ ricavato sopra è chiamato l'*inverso* di α ed è denotato con α^{-1} o $1/\alpha$. Se α, β sono numeri complessi, spesso scriveremo β/α invece di $\alpha^{-1}\beta$ (oppure $\beta\alpha^{-1}$), analogamente a quanto facciamo con i numeri reali. Vediamo quindi che la divisione per numeri complessi non nulli è sempre possibile.

Definiamo *valore assoluto* del numero complesso $\alpha = a_1 + ia_2$ il numero

$$|\alpha| = \sqrt{a_1^2 + a_2^2}.$$

Questo valore assoluto non è altro che la lunghezza del vettore (a_1, a_2) . Adoperando il valore assoluto, possiamo scrivere

$$\alpha^{-1} = \frac{\bar{\alpha}}{|\alpha|^2}$$

supposto $\alpha \neq 0$.

La diseguaglianza triangolare per le lunghezze dei vettori può essere ora interpretata con numeri complessi. Se α, β sono numeri complessi allora

$$|\alpha + \beta| \leq |\alpha| + |\beta|.$$

Un'altra proprietà del valore assoluto è data nell'esercizio 5.

Esercizi

1. Scrivere i seguenti numeri complessi nella forma $x+iy$, dove x e y sono numeri reali.

- | | |
|---------------------------------|------------------------------|
| a) $(-1 + 3i)^{-1}$. | b) $(1 + i)(1 - i)$. |
| c) $(1 + i)i(2 - i)$. | d) $(i - 1)(2 - i)$. |
| e) $(7 + \pi i)(\pi + i)$. | f) $(2i + 1)\pi i$. |
| g) $(\sqrt{2} + i)(\pi + 3i)$. | h) $(i + 1)(i - 2)(i + 3)$. |

2. Scrivere i numeri complessi seguenti nella forma $x+iy$, dove x e y sono numeri reali.

- | | | | |
|------------------------|------------------------|----------------------------|-------------------------|
| a) $(1 + i)^{-1}$. | b) $\frac{1}{3 + i}$. | c) $\frac{2 + i}{2 - i}$. | d) $\frac{1}{2 - i}$. |
| e) $\frac{1 + i}{i}$. | f) $\frac{i}{1 + i}$. | g) $\frac{2i}{3 - i}$. | h) $\frac{1}{-1 + i}$. |

3. Sia α un numero complesso diverso da zero. Qual è il valore assoluto di $\alpha/\bar{\alpha}$? Quale numero è $\bar{\bar{\alpha}}$?

4. Siano α, β due numeri complessi. Dimostrare che $\overline{\alpha\beta} = \bar{\alpha}\bar{\beta}$ e che

$$\overline{\alpha + \beta} = \bar{\alpha} + \bar{\beta}$$

5. Provare che $|\alpha\beta| = |\alpha||\beta|$.

6. Definire l'addizione tra n -uple di numeri complessi componente per componente e la moltiplicazione di una n -upla di numeri complessi per un numero complesso componente per componente. Siano $A = (\alpha_1, \dots, \alpha_n)$ e $B = (\beta_1, \dots, \beta_n)$ n -uple di numeri complessi, si definisca il loro prodotto scalare $\langle A, B \rangle$ come il numero

$$\alpha_1\bar{\beta}_1 + \dots + \alpha_n\bar{\beta}_n$$

(notare che i secondi fattori sono i coniugati dei numeri β_i). Dimostrare allora le seguenti proprietà:

PS 1. $\langle A, B \rangle = \overline{\langle B, A \rangle}$.

PS 2. $\langle A, B + C \rangle = \langle A, B \rangle + \langle A, C \rangle$.

PS 3. Se α è un numero complesso, allora

$$\langle \alpha A, B \rangle = \alpha \langle A, B \rangle \quad \text{e} \quad \langle A, \alpha B \rangle = \bar{\alpha} \langle A, B \rangle.$$

PS 4. Se $A = O$ allora $\langle A, A \rangle = 0$; in ogni altro caso $\langle A, A \rangle > 0$.

7. In questo esercizio è presupposta la conoscenza delle funzioni seno e coseno e delle relative formule di addizione. Sia θ un numero reale:

a) Definire

$$e^{i\theta} = \cos \theta + i \sin \theta.$$

e dimostrare che se θ_1 e θ_2 sono numeri reali, allora

$$e^{i(\theta_1 + \theta_2)} = e^{i\theta_1} e^{i\theta_2}.$$

Dimostrare che ogni numero complesso di valore assoluto uguale a 1 può essere scritto nella forma e^{it} , con un opportuno numero reale t .

b) Dimostrare che ogni numero complesso può essere scritto nella forma $r e^{i\theta}$, con opportuni numeri reali r, θ , essendo $r \geq 0$.

c) Se $z_1 = r_1 e^{i\theta_1}$ e $z_2 = r_2 e^{i\theta_2}$, dove r_1, r_2 sono numeri reali non negativi e θ_1, θ_2 numeri reali, dimostrare che

$$z_1 z_2 = r_1 r_2 e^{i(\theta_1 + \theta_2)}.$$

d) Se z è un numero complesso e n è un intero maggiore di zero, dimostrare che esiste un numero complesso w tale che $w^n = z$. Dimostrare che, in effetti, di tali numeri complessi w ne esistono n distinti (se $z \neq 0$). [Se $z = r e^{i\theta}$, si consiglia di considerare dapprima $r^{1/n} e^{i\theta/n}$.]

Capitolo 2

Spazi vettoriali

7. TERMINOLOGIA

Adoperando un termine entrato ormai nell'uso comune, chiameremo *insieme* ogni collezione di oggetti: un oggetto della collezione viene detto *elemento* dell'insieme. È utile in pratica adoperare brevi simboli per denotare certi insiemi. Per esempio, denoteremo con \mathbb{R} l'insieme dei numeri reali e con \mathbb{C} l'insieme dei numeri complessi. Dire quindi che "x è un numero reale" oppure che "x è un elemento di \mathbb{R} " è la stessa cosa. L'insieme di tutte le n -uple (ordinate) di numeri reali sarà denotato con \mathbb{R}^n . Perciò "X è un elemento di \mathbb{R}^n " e "X è una n -upla di numeri reali" significano la stessa cosa.

Invece di dire che u è un elemento di un insieme S , molto spesso diremo anche che u *appartiene a* S e scriveremo $u \in S$. Se S ed S' sono insiemi, se ogni elemento di S' è anche un elemento di S , diremo che S' è un *sottoinsieme* di S . Quindi l'insieme dei numeri reali è un sottoinsieme dell'insieme dei numeri complessi. Dire che S' è un sottoinsieme di S equivale a dire che S' è una parte di S . Si osservi che la nostra definizione di sottoinsieme non esclude la possibilità che S' coincida con S . Se S' è un sottoinsieme di S , ma S' non coincide con S , diremo che S' è un *sottoinsieme proprio* di S . Quindi, \mathbb{C} è un sottoinsieme di \mathbb{C} , ma \mathbb{R} è un sottoinsieme proprio di \mathbb{C} . Per denotare il fatto che S' è un sottoinsieme di S , scriveremo $S' \subset S$ e diremo anche che S' è *contenuto in* S .

Se S_1 , S_2 sono insiemi, l'*intersezione* di S_1 con S_2 , denotata con $S_1 \cap S_2$, è l'insieme degli elementi che appartengono tanto

a S_1 quanto a S_2 . L'unione di S_1 e S_2 , denotata con $S_1 \cup S_2$, è invece l'insieme degli elementi che appartengono ad almeno uno dei due insiemi S_1 , S_2 .

8. DEFINIZIONI

Sia K un sottoinsieme dell'insieme \mathbb{C} dei numeri complessi. Noi diremo che K è un *corpo* se sono soddisfatte le condizioni seguenti:¹

- a) Se x , y sono elementi di K , allora $x+y$ e xy sono elementi di K .
- b) Se $x \in K$, allora $-x$ è un elemento di K ; se inoltre $x \neq 0$, allora anche x^{-1} è un elemento di K .
- c) I numeri 0 e 1 sono elementi di K .

Osserviamo subito che \mathbb{R} e \mathbb{C} sono entrambi corpi.

Denotiamo con \mathbb{Q} l'insieme dei numeri razionali, cioè l'insieme di tutte le frazioni m/n dove m , n sono interi e $n \neq 0$. È facile verificare allora che \mathbb{Q} è un corpo.

Sia \mathbb{Z} l'insieme di tutti i numeri interi. Poiché la condizione b) (sopra riportata) non è verificata, \mathbb{Z} non è un corpo. Infatti se n è un intero non nullo, allora $n^{-1} = 1/n$ non è un intero (con l'ovvia eccezione del caso $n = 1$ o $n = -1$). Per esempio, $\frac{1}{2}$ non è un intero.

La cosa essenziale a proposito dei corpi è che essi sono insiemi di elementi che si possono addizionare e moltiplicare tra loro, in modo che l'addizione e la moltiplicazione soddisfino le usuali leggi aritmetiche e inoltre in modo che sia sempre possibile la divisione per elementi non nulli. È possibile trattare questa nozione assiomaticamente, ma noi lo faremo soltanto più tardi per evitare discussioni astratte che invece divengono molto semplici quando il lettore ha acquisito la necessaria maturità matematica. In vista di questa possibile generalizzazione, dovremmo dire che

¹ La traduzione italiana del termine inglese *field* è *campo*. Abbiamo tuttavia preferito adoperare il termine italiano *corpo* perché questo è l'uso nella letteratura più recente. Il termine *corpo* traduce più propriamente l'inglese *field*, e indica un corpo in cui la moltiplicazione può non essere commutativa; quando si vuol porre l'accento sul fatto che la moltiplicazione è commutativa, si dice *corpo commutativo*. In questo caso, trattandosi di insiemi di numeri complessi, è chiaro che *corpo* significa *corpo commutativo*. [N.d.T.]

L'oggetto sopra definito è un corpo di numeri (complessi). Tuttavia anche questi speciali corpi saranno chiamati semplicemente corpi.

In tutta questa trattazione dell'algebra lineare, il lettore può, per sua comodità, limitarsi a considerare il corpo dei numeri reali e quello dei numeri complessi. Poiché sarà necessario considerarli entrambi, ci converrà scegliere per denotarli una stessa lettera K .

Siano K, L corpi e si supponga K contenuto in L (cioè che K sia un sottoinsieme di L). Diremo allora che K è un *sottocorpo* di L : ognuno dei corpi che stiamo considerando è quindi un sottocorpo del corpo complesso. In particolare possiamo dire che \mathbf{R} è un sottocorpo di \mathbf{C} e che \mathbf{Q} è un sottocorpo di \mathbf{R} .

Se K è un corpo, i suoi elementi saranno anche chiamati *numeri* (senza altra specificazione) se il riferimento a K è chiaro dal contesto, altrimenti li chiameremo *scalari*.

Uno *spazio vettoriale* V sul corpo K è un insieme di oggetti che possono essere addizionati tra loro e moltiplicati per elementi di K in modo che la somma di due elementi di V sia ancora un elemento di V , il prodotto di un elemento di V per un elemento di K sia un elemento di V e siano soddisfatte le seguenti proprietà:

SV 1. *Comunque si prendano gli elementi u, v, w di V si ha*

$$(u + v) + w = u + (v + w).$$

SV 2. *Esiste un elemento di V , indicato con O , tale che*

$$O + u = u + O = u$$

per ogni elemento u di V .

SV 3. *Per ogni elemento u di V , l'elemento $(-1)u$ è tale che*

$$u + (-1)u = O.$$

SV 4. *Per tutti gli elementi u, v di V , si ha*

$$u + v = v + u.$$

SV 5. *Se c è un numero, allora $c(u+v) = cu+cv$.*

SV 6. *Se a, b sono due numeri, allora $(a+b)v = av+bv$.*

SV 7. *Se a, b sono due numeri, allora $(ab)v = a(bv)$.*

SV 8. *Per ogni elemento u di V si ha $1 \cdot u = u$ (1 è il numero uno).*

Altre proprietà che possono essere facilmente dedotte da queste sono riportate negli esercizi; esse, d'ora in poi, saranno assunte come note.

La somma $u + (-1)v$ è usualmente scritta $u - v$; analogamente scriveremo $-v$ invece di $(-1)v$.

Noi adopereremo 0 per indicare il numero zero e riserveremo O per denotare l'elemento di un qualsiasi spazio vettoriale definito nella proprietà SV 2. Anche questo elemento sarà chiamato zero, ma non vi sarà mai possibilità di confusione. È utile osservare che questo elemento O è univocamente determinato dalla condizione SV 2 (vedi esercizio 5).

È possibile addizionare più elementi di uno spazio vettoriale: supponiamo di voler addizionare i quattro elementi u, v, w, z . Noi addizioniamo prima due di essi, poi aggiungiamo il terzo e infine aggiungiamo il rimanente.

Le proprietà SV 1 e SV 4 permettono di eseguire queste successive addizioni in un ordine qualsiasi: questa è proprio la stessa situazione che avevamo con i vettori. Per esempio noi abbiamo

$$\begin{aligned} ((u+v)+w)+z &= (u+(v+w))+z \\ &= ((v+w)+u)+z \\ &= (v+w)+(u+z) \\ &\quad \text{ecc.} \end{aligned}$$

È uso comune tralasciare le parentesi e scrivere semplicemente

$$u+v+w+z.$$

La stessa osservazione si può fare nei riguardi dell'addizione di un qualunque numero di elementi di V ; una dimostrazione formale si può facilmente conseguire per induzione.

Sia V uno spazio vettoriale e W un sottoinsieme di V . Si supponga che W abbia le seguenti proprietà:

- 1) Se v, w sono elementi di W , la loro somma $v+w$ è di nuovo un elemento di W .
- 2) Se v è un elemento di W e c è un numero, cv è un elemento di W .
- 3) L'elemento O di V appartiene a W .

Allora W è esso stesso uno spazio vettoriale: infatti, le proprietà da SV 1 a SV 8 sono soddisfatte per gli elementi di V e quindi, a maggior ragione, lo sono anche per gli elementi di W . Diremo allora che W è un *sottospazio* di V .

Come immediata conseguenza delle definizioni, osserviamo che, se W_1 e W_2 sono sottospazi di V , la loro intersezione $W_1 \cap W_2$ è ancora un sottospazio di V .

Esempio 1. Sia $V = \mathbb{R}^n$. Allora V è uno spazio vettoriale sul corpo reale, poiché abbiamo già osservato, considerando le n -uple di numeri reali, che l'addizione tra esse e la loro moltiplicazione per numeri reali soddisfano certe proprietà che noi ora riconosciamo essere quelle che definiscono uno spazio vettoriale.

Più in generale, se K è un corpo, indichiamo con K^n l'insieme delle n -uple di elementi di K , cioè l'insieme degli elementi

$$X = (x_1, \dots, x_n)$$

con $x_i \in K$ per $i = 1, \dots, n$. Noi quindi definiamo l'addizione tra queste n -uple componenti per componente, esattamente come abbiamo fatto per le n -uple di numeri reali. Quindi se $Y = (y_1, \dots, y_n)$ con $y_i \in K$, allora

$$X + Y = (x_1 + y_1, \dots, x_n + y_n).$$

Se $c \in K$, noi definiamo $cX = (cx_1, \dots, cx_n)$. Allora verifichiamo immediatamente che gli assiomi che definiscono uno spazio vettoriale sono soddisfatti da queste operazioni, cioè verifichiamo che K^n è uno spazio vettoriale su K .

Quindi \mathbb{C}^n è uno spazio vettoriale su \mathbb{C} , e \mathbb{Q}^n è uno spazio vettoriale su \mathbb{Q} . Notiamo però che \mathbb{R}^n non è uno spazio vettoriale su \mathbb{C} . Quindi, trattando degli spazi vettoriali, avremo sempre cura di specificare il corpo su cui consideriamo lo spazio vettoriale. Quando noi scriviamo K^n sarà sempre sottinteso che si tratta di uno spazio vettoriale su K . Gli elementi di K^n saranno chiamati vettori, ed è anche usuale chiamare vettori gli elementi di ogni spazio vettoriale.

Esempio 2. Sia $V = \mathbb{R}^n$ e sia W l'insieme dei vettori di V la cui ultima coordinata è 0. Allora W è un sottospazio di V che noi possiamo identificare con \mathbb{R}^{n-1} .

Esempio 3. Sia V uno spazio vettoriale qualsiasi, siano v_1, \dots, v_n elementi di V , siano x_1, \dots, x_n numeri. Allora l'espressione

$$x_1v_1 + \dots + x_nv_n,$$

è chiamata *combinazione lineare* di v_1, \dots, v_n . Se W è l'insieme di tutte le combinazioni lineari di v_1, \dots, v_n , W è un sottospazio di V .

Dimostrazione. Siano y_1, \dots, y_n numeri, allora

$$\begin{aligned} (x_1v_1 + \dots + x_nv_n) + (y_1v_1 + \dots + y_nv_n) &= \\ &= (x_1 + y_1)v_1 + \dots + (x_n + y_n)v_n, \end{aligned}$$

quindi la somma di due elementi di W è di nuovo un elemento di W , cioè è di nuovo una combinazione lineare di v_1, \dots, v_n . Inoltre se c è un numero, abbiamo che l'espressione

$$c(x_1v_1 + \dots + x_nv_n) = cx_1v_1 + \dots + cx_nv_n$$

è una combinazione lineare di v_1, \dots, v_n , e quindi è un elemento di W . Infine, l'uguaglianza

$$O = 0v_1 + \dots + 0v_n$$

dice che O è un elemento di W . Abbiamo così provato che W è un sottospazio di V .

Nell'esempio 3, W è chiamato il sottospazio *generato* da v_1, \dots, v_n . Se $W = V$, cioè se ogni elemento di V è combinazione lineare di v_1, \dots, v_n , diciamo che i vettori v_1, \dots, v_n *generano* V su K . Quando ci riferiamo a un fissato corpo K , diciamo senz'altra specificazione che i vettori v_1, \dots, v_n *generano* V oppure che sono *generatori* di V .

Esempio 4. Sia S un insieme e K un corpo. Dicendo *funzione* di S in K intenderemo un modo di associare ad ogni elemento di S un elemento di K . Perciò, se f è una funzione di S in K , noi esprimiamo questo con il simbolo

$$f: S \rightarrow K.$$

Diremo anche che f è una funzione a valori in K . Sia V l'insieme di tutte le funzioni di S in K . Se f, g sono due tali funzioni, noi possiamo parlare della loro somma definendola come la funzione

che associa all'elemento x di S l'elemento $f(x) + g(x)$ di K . E scriviamo allora

$$(f + g)(x) = f(x) + g(x).$$

Se $c \in K$, definiamo cf come la funzione tale che

$$(cf)(x) = cf(x).$$

Quindi, il valore di cf in x è $cf(x)$. È molto facile dimostrare che V è uno spazio vettoriale su K . Lasciamo questa dimostrazione al lettore. Osserviamo soltanto che l'elemento zero di V è la funzione 0, cioè la funzione f per cui $f(x) = 0$ per ogni $x \in S$. Questa funzione sarà denotata con 0.

Esempio 5. Sia V l'insieme di tutte le funzioni di \mathbb{R} in \mathbb{R} . Sappiamo allora che V è uno spazio vettoriale su \mathbb{R} . Sia W il sottoinsieme di V costituito dalle funzioni continue. Se f, g sono funzioni continue, anche $f + g$ è una funzione continua. Se c è un numero reale, la funzione cf è continua. La funzione 0 è evidentemente continua. Possiamo allora dire che W è un sottospazio dello spazio vettoriale di tutte le funzioni di \mathbb{R} in \mathbb{R} , cioè W è un sottospazio di V .

Sia U l'insieme di tutte le funzioni differenziabili di \mathbb{R} in \mathbb{R} . Se f, g sono funzioni differenziabili, anche la loro somma è una funzione differenziabile; se c è un numero reale, la funzione cf è differenziabile. La funzione 0 è differenziabile. Possiamo quindi dire che U è un sottospazio di V . In effetti, U è un sottospazio di W , giacché ogni funzione differenziabile è anche continua.

Esempio 6. Consideriamo nuovamente V , spazio vettoriale (su \mathbb{R}) di tutte le funzioni di \mathbb{R} in \mathbb{R} . Consideriamo le funzioni e^t, e^{2t} (per essere precisi, avremmo dovuto dire: consideriamo le due funzioni f, g tali che, per ogni $t \in \mathbb{R}, f(t) = e^t$ e $g(t) = e^{2t}$). Queste funzioni generano un sottospazio dello spazio di tutte le funzioni differenziabili. La funzione $3e^t + 2e^{2t}$ è un elemento di questo sottospazio e così anche la funzione $2e^t + \pi e^{2t}$.

Esempio 7. Osserviamo che i numeri complessi costituiscono uno spazio vettoriale sul corpo reale. Questa affermazione segue immediatamente dalle regole che abbiamo date nel paragrafo 6 del primo capitolo a proposito dell'addizione e della moltiplicazione tra numeri complessi.

Esercizi

1. Sia V uno spazio vettoriale, facendo uso delle proprietà SV 1, ..., SV 8 dimostrare che se v è un elemento di V e 0 è il numero reale zero, allora $0v = O$.
2. Siano c un numero diverso da zero e v un elemento di V . Dimostrare che se $cv = O$ allora $v = O$.
3. Nello spazio vettoriale delle funzioni, qual è la funzione che soddisfa la condizione SV 2?
4. Sia V uno spazio vettoriale e siano v, w due elementi di V . Se $v + w = O$, dimostrare che $w = -v$.
5. Sia V uno spazio vettoriale e siano v, w due elementi di V tali che $v + w = v$. Dimostrare che $w = O$.
6. Sia K un sottocorpo del corpo L : dimostrare che L è uno spazio vettoriale su K . In particolare \mathbf{C} e \mathbf{R} sono spazi vettoriali su \mathbf{Q} .
7. Sia K l'insieme di tutti i numeri reali che possono essere scritti nella forma $a + b\sqrt{2}$, con a, b numeri razionali. Dimostrare che K è un corpo.
8. Sia K l'insieme di tutti i numeri che possono essere scritti nella forma $a + bi$, con a, b numeri razionali. Dimostrare che K è un corpo.
9. Sia c un numero razionale maggiore di zero, sia γ un numero reale tale che $\gamma^2 = c$. Dimostrare che l'insieme di tutti i numeri che possono essere scritti nella forma $a + b\gamma$, con a, b numeri razionali, è un corpo.

9. BASI

Sia V uno spazio vettoriale sul corpo K e siano v_1, \dots, v_n elementi di V . Noi diremo che v_1, \dots, v_n sono *linearmente dipendenti su K* se in K esistono n elementi a_1, \dots, a_n , non tutti nulli, tali che

$$a_1v_1 + \dots + a_nv_n = O.$$

Se elementi siffatti non esistono, diciamo che v_1, \dots, v_n sono *linearmente indipendenti su K*. Spesso la specificazione "su K " viene omessa.

Esempio 1. Sia $V = \mathbf{R}^n$ e si considerino i vettori

$$E_1 = (1, 0, \dots, 0)$$

⋮

$$E_n = (0, 0, \dots, 1).$$

Allora E_1, \dots, E_n sono linearmente indipendenti. Infatti, siano

a_1, \dots, a_n numeri tali che $a_1E_1 + \dots + a_nE_n = O$, poiché

$$a_1E_1 + \dots + a_nE_n = (a_1, \dots, a_n),$$

segue che ogni a_i è nullo.

Esempio 2. Sia V lo spazio vettoriale di tutte le funzioni della variabile reale t . Siano $f_1(t), \dots, f_n(t)$ n funzioni. Dire che esse sono linearmente dipendenti significa affermare che esistono n numeri reali a_1, \dots, a_n , non tutti nulli, tali che

$$a_1f_1(t) + \dots + a_nf_n(t) = 0$$

per ogni valore di t .

Le due funzioni e^t, e^{2t} sono linearmente indipendenti. Per dimostrarlo, supponiamo che vi siano due numeri a, b tali che

$$ae^t + be^{2t} = 0$$

(per ogni valore di t). Derivando otteniamo

$$ae^t + 2be^{2t} = 0.$$

Sottraendo la prima relazione dalla seconda, otteniamo $be^t = 0$ e quindi $b = 0$. Dalla prima relazione segue allora che $ae^t = 0$ e quindi che anche a è nullo. Si conclude perciò che e^t ed e^{2t} sono linearmente indipendenti.

Consideriamo nuovamente uno spazio vettoriale arbitrario V sul corpo K . Siano v_1, \dots, v_n elementi di V linearmente indipendenti. Siano x_1, \dots, x_n e y_1, \dots, y_n due n -uple di numeri tali che

$$x_1v_1 + \dots + x_nv_n = y_1v_1 + \dots + y_nv_n.$$

In altre parole, supponiamo che due combinazioni lineari di v_1, \dots, v_n siano uguali. Allora deve necessariamente essere $x_i = y_i$ per ogni $i = 1, \dots, n$. Infatti, sottraendo il membro destro da quello sinistro, otteniamo

$$x_1v_1 - y_1v_1 + \dots + x_nv_n - y_nv_n = O.$$

Questa relazione può essere anche scritta nella forma

$$(x_1 - y_1)v_1 + \dots + (x_n - y_n)v_n = O.$$

Per definizione, allora, dobbiamo avere $x_i - y_i = 0$ per ogni $i = 1, \dots, n$; provando così la nostra asserzione.

Noi definiamo *base* di V su K una successione di elementi $\{v_1, \dots, v_n\}$ che generano V e che sono linearmente indipendenti.

Nell'esempio 1, i vettori E_1, \dots, E_n costituiscono una base di R^n su R . Sia W lo spazio vettoriale su R generato dalle due funzioni e^t, e^{2t} : allora $\{e^t, e^{2t}\}$ è una base di W su R .

Sia V uno spazio vettoriale e $\{v_1, \dots, v_n\}$ ne sia una base. Gli elementi di V possono essere rappresentati da n -uple di numeri relativamente a questa base nel modo seguente. Se un elemento v di V è scritto come una combinazione lineare

$$v = x_1 v_1 + \dots + x_n v_n$$

di elementi della base, diciamo che (x_1, \dots, x_n) sono le *coordinate* di v rispetto alla base data, e chiamiamo x_i la i -esima coordinata. Diciamo allora che la n -upla $X = (x_1, \dots, x_n)$ è il *vettore delle coordinate* di v rispetto alla base $\{v_1, \dots, v_n\}$.

Per esempio, sia V lo spazio vettoriale di funzioni generato dalle due funzioni e^t, e^{2t} . Allora le coordinate della funzione

$$3e^t + 5e^{2t}$$

rispetto alla base $\{e^t, e^{2t}\}$ sono $(3, 5)$.

Esempio 3. Dimostrare che i vettori $(1, 1)$ e $(-3, 2)$ sono linearmente indipendenti su R .

Siano a, b due numeri reali tali che

$$a(1, 1) + b(-3, 2) = O.$$

Riscrivendo questa uguaglianza separando le componenti, troviamo

$$a - 3b = 0,$$

$$a + 2b = 0.$$

Questo è un sistema di due equazioni che risolviamo rispetto ad a e b . Sottraendo la seconda uguaglianza dalla prima otteniamo $-5b = 0$, da cui $b = 0$. Sostituendo b in una delle due equazioni troviamo $a = 0$. Quindi a e b sono entrambi nulli e i vettori dati sono perciò indipendenti.

Esempio 4. Determinare le coordinate di $(1, 0)$ rispetto ai due vettori $(1, 1)$ e $(-1, 2)$.

Dobbiamo trovare due numeri a, b tali che

$$a(1, 1) + b(-1, 2) = (1, 0).$$

Riscrivendo questa uguaglianza separando le coordinate, troviamo

$$a - b = 1,$$

$$a + 2b = 0.$$

Ricavando nel solito modo a e b , otteniamo $b = -\frac{1}{3}$ e $a = \frac{2}{3}$. Quindi, le coordinate di $(1, 0)$ rispetto a $(1, 1)$ e $(-1, 2)$ sono $(\frac{2}{3}, -\frac{1}{3})$.

Sia $\{v_1, \dots, v_n\}$ un insieme di elementi di uno spazio vettoriale V su un corpo K . Sia r un intero positivo non maggiore di n . Diremo che $\{v_1, \dots, v_r\}$ è un sottoinsieme *massimale* di elementi linearmente indipendenti se v_1, \dots, v_r sono linearmente indipendenti e inoltre, per ogni elemento v_i con $i > r$, gli elementi v_1, \dots, v_r, v_i risultano linearmente dipendenti.

Il teorema che segue fornisce un utile criterio per decidere se un insieme di elementi di uno spazio vettoriale è una base.

TEOREMA 1 *Sia $\{v_1, \dots, v_n\}$ un insieme di generatori di uno spazio vettoriale V . Sia $\{v_1, \dots, v_r\}$ un sottoinsieme massimale di elementi linearmente indipendenti. Allora, $\{v_1, \dots, v_r\}$ è una base dello spazio V .*

Dimostrazione. Dobbiamo provare che i vettori v_1, \dots, v_r generano lo spazio V . Proveremo dapprima che ogni v_i (con $i > r$) è una combinazione lineare di v_1, \dots, v_r . Per ipotesi, fissato v_i , esistono i numeri x_1, \dots, x_r, y , non tutti nulli, per cui

$$x_1 v_1 + \dots + x_r v_r + y v_i = O.$$

Inoltre, y non è nullo, altrimenti avremmo una relazione di dipendenza lineare tra gli elementi v_1, \dots, v_r . Possiamo allora ricavare v_i ottenendo

$$v_i = \frac{x_1}{-y} v_1 + \dots + \frac{x_r}{-y} v_r,$$

e provando così che v_i è una combinazione lineare degli elementi v_1, \dots, v_r .

Consideriamo ora un elemento qualsiasi v di V . Esistono allora

i numeri c_1, \dots, c_n in modo che

$$v = c_1 v_1 + \dots + c_n v_n.$$

In questa relazione possiamo sostituire ad ogni v_i (con $i > r$) una combinazione lineare di v_1, \dots, v_r . Così facendo, dopo aver raccolto i termini simili, troviamo che v rimane espresso come combinazione lineare di v_1, \dots, v_r . Si prova così che v_1, \dots, v_r generano lo spazio V e quindi che ne costituiscono una base.

Esercizi

1. Dimostrare che i seguenti vettori sono linearmente indipendenti, tanto sul corpo reale \mathbb{R} quanto su quello complesso \mathbb{C} .

- a) $(1, 1, 1)$ e $(0, 1, -1)$.
- b) $(1, 0)$ e $(1, 1)$.
- c) $(-1, 1, 0)$ e $(0, 1, 2)$.
- d) $(2, -1)$ e $(1, 0)$.
- e) $(\pi, 0)$ e $(0, 1)$.
- f) $(1, 2)$ e $(1, 3)$.
- g) $(1, 1, 0), (1, 1, 1)$ e $(0, 1, -1)$.
- h) $(0, 1, 1), (0, 2, 1)$ e $(1, 5, 3)$.

2. Esprimere l'assegnato vettore X come combinazione lineare dei vettori A, B e trovare le coordinate di X rispetto ad A, B .

- a) $X = (1, 0), A = (1, 1), B = (0, 1)$.
- b) $X = (2, 1), A = (1, -1), B = (1, 1)$.
- c) $X = (1, 1), A = (2, 1), B = (-1, 0)$.
- d) $X = (4, 3), A = (2, 1), B = (-1, 0)$.

(Si possono riguardare i vettori assegnati come elementi di \mathbb{R}^2 o di \mathbb{C}^2 : le coordinate risultano le stesse nei due casi.)

3. Trovare le coordinate del vettore X rispetto ai vettori A, B, C .

- a) $X = (1, 0, 0), A = (1, 1, 1), B = (-1, 1, 0), C = (1, 0, -1)$.
- b) $X = (1, 1, 1), A = (0, 1, -1), B = (1, 1, 0), C = (1, 0, 2)$.
- c) $X = (0, 0, 1), A = (1, 1, 1), B = (-1, 1, 0), C = (1, 0, -1)$.

4. Siano (a, b) e (c, d) due vettori del piano. Se $ad - bc = 0$, dimostrare che essi sono linearmente dipendenti. Se $ad - bc \neq 0$ dimostrare che essi sono linearmente indipendenti.

5. Considerato lo spazio vettoriale di tutte le funzioni reali della variabile reale t , dimostrare che le seguenti coppie sono costituite da funzioni linearmente indipendenti.

- a) $1, t$.
- b) t, t^2 .
- c) t, t^4 .
- d) e^t, t .
- e) te^t, e^{2t} .
- f) $\sin t, \cos t$.
- g) $t, \sin t$.
- h) $\sin t, \sin 2t$.
- i) $\cos t, \cos 3t$.

6. Si consideri lo spazio vettoriale di tutte le funzioni reali definite per $t > 0$. Dimostrare che le seguenti coppie sono costituite da funzioni linearmente indipendenti.

- a) $t, 1/t$. b) $e^t, \log t$.

7. Quali sono le coordinate della funzione $f(t) = 3 \sin t + 5 \cos t$ rispetto alla base $\{\sin t, \cos t\}$?

8. Si indichi con D la derivazione d/dt . Sia $f(t)$ la funzione dell'esercizio 7. Quali sono le coordinate della funzione $Df(t)$ rispetto alla base indicata nello stesso esercizio 7?

9. Siano A_1, \dots, A_r vettori di \mathbb{R}^n e si supponga che essi siano mutuamente perpendicolari (cioè, due qualunque di essi, se di indici diversi, siano perpendicolari tra loro) e che nessuno di essi sia uguale a O . Dimostrare che questi vettori sono linearmente indipendenti.

10. Sia V uno spazio vettoriale delle funzioni reali continue definite nell'intervallo $[-\pi, \pi]$. Se f, g sono due funzioni continue su questo intervallo, si definisca il loro prodotto scalare $\langle f, g \rangle$ ponendo

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(t)g(t)dt.$$

Dimostrare che le funzioni $\sin nt$ ($n = 1, 2, 3, \dots$) sono mutuamente perpendicolari, cioè, il prodotto scalare di due di esse, con diversi n , è nullo.

11. Dimostrare che le funzioni $\sin t, \sin 2t, \sin 3t, \dots, \sin nt$ sono linearmente indipendenti su \mathbb{R} , qualunque sia l'intero $n \geq 1$.

10. DIMENSIONE DI UNO SPAZIO VETTORIALE

Il principale risultato di questo paragrafo è che due basi qualsiasi di uno spazio vettoriale hanno lo stesso numero di elementi. Per dimostrare ciò, stabiliamo dapprima il seguente:

TEOREMA 2 *Sia V uno spazio vettoriale sul corpo K . Sia $\{v_1, \dots, v_m\}$ una base di V su K . Siano w_1, \dots, w_n elementi di V . Se n è maggiore di m , i vettori w_1, \dots, w_n sono linearmente dipendenti.*

Dimostrazione. Osserviamo intanto che l'affermazione da provare è vera se gli m elementi w_1, \dots, w_m sono linearmente dipendenti. Basta quindi dare una dimostrazione del teorema sotto l'ulteriore ipotesi che i vettori w_1, \dots, w_m siano linearmente indi-

pendenti (e perciò, osserviamo esplicitamente, nessuno di essi può essere il vettore zero).

Poiché $\{v_1, \dots, v_m\}$ è una base, in K esistono gli elementi a_1, \dots, a_m per cui

$$w_1 = a_1 v_1 + \dots + a_m v_m.$$

Per ipotesi sappiamo che $w_1 \neq 0$ e quindi sappiamo che qualche a_i è diverso da zero. Dopo aver numerato diversamente gli elementi v_1, \dots, v_m (se è necessario), senza ledere la generalità possiamo assumere che sia $a_1 \neq 0$. Possiamo allora ricavare v_1 ottenendo

$$a_1 v_1 = w_1 - a_2 v_2 - \dots - a_m v_m,$$

$$v_1 = a_1^{-1} w_1 - a_1^{-1} a_2 v_2 - \dots - a_1^{-1} a_m v_m.$$

Il sottospazio di V generato da w_1, v_2, \dots, v_m contiene v_1 , e quindi deve coincidere con l'intero spazio V giacché v_1, v_2, \dots, v_m generano appunto V . L'idea della dimostrazione consiste ora nel proseguire questa procedura sostituendo successivamente v_2, v_3, \dots con w_2, w_3, \dots finché tutti gli elementi v_1, \dots, v_m risultano sostituiti, ottenendo nel contempo che w_1, \dots, w_m generano V . Assumiamo ora, procedendo per induzione, che ci sia un intero r per cui $1 \leq r < m$ e tale che, dopo una eventuale opportuna rinumerazione di v_1, \dots, v_m , gli elementi $w_1, \dots, w_r, v_{r+1}, \dots, v_m$ generino V . In K esistono allora gli elementi $b_1, \dots, b_r, c_{r+1}, \dots, c_m$ tali che

$$w_{r+1} = b_1 w_1 + \dots + b_r w_r + c_{r+1} v_{r+1} + \dots + c_m v_m.$$

Osserviamo che non può essere $c_j = 0$ per $j = r+1, \dots, m$, perché, altrimenti, otterremmo una relazione di dipendenza lineare tra w_1, \dots, w_{r+1} , contro quanto abbiamo supposto. Rinumerando, se necessario, v_{r+1}, \dots, v_m , senza ledere la generalità possiamo assumere che sia $c_{r+1} \neq 0$. Allora otteniamo

$$c_{r+1} v_{r+1} = w_{r+1} - b_1 w_1 - \dots - b_r w_r - c_{r+2} v_{r+2} - \dots - c_m v_m.$$

Dividendo per c_{r+1} concludiamo che v_{r+1} appartiene al sottospazio generato da $w_1, \dots, w_{r+1}, v_{r+2}, \dots, v_m$. Dall'ipotesi di induzione segue allora che $w_1, \dots, w_{r+1}, v_{r+2}, \dots, v_m$ generano V . Per induzione, rimane così provato che w_1, \dots, w_m generano V . Se n supera m , in K esistono degli elementi d_1, \dots, d_m tali che

$$w_n = d_1 w_1 + \dots + d_m w_m,$$

provando così che w_1, \dots, w_n sono linearmente dipendenti. Questo conclude la dimostrazione del teorema.

TEOREMA 3 *Sia V uno spazio vettoriale e si supponga che esista una base costituita da n elementi, mentre esiste un'altra base costituita da m elementi. Allora $m = n$.*

Dimostrazione. Applichiamo il teorema 2 alle due basi. Questo teorema implica che entrambe le alternative $n > m$ e $m > n$ sono impossibili, perciò deve risultare $m = n$.

Sia V uno spazio vettoriale avente una base costituita da n elementi. Diciamo allora che n è la *dimensione* di V . Se V consiste del solo O , allora V non ha basi e noi diciamo che V ha *dimensione zero*.

Esempio 1. Lo spazio vettoriale \mathbb{R}^n ha dimensione n su \mathbb{R} , lo spazio vettoriale \mathbb{C}^n ha dimensione n su \mathbb{C} ; più in generale, per ogni corpo K , lo spazio vettoriale K^n ha dimensione n su K . Infatti gli n vettori

$$(1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, \dots, 0, 1)$$

costituiscono una base di K^n su K .

La dimensione di uno spazio vettoriale V su K sarà denotata con $\dim_K V$ o, più semplicemente, con $\dim V$.

Gli spazi vettoriali aventi una base costituita da un numero finito di elementi e lo spazio costituito dal solo elemento O sono detti spazi di *dimensione finita*. Gli altri spazi vettoriali sono detti di *dimensione infinita*. È possibile dare una definizione di base costituita da infiniti elementi. Il lettore può trovarla in trattazioni più approfondite. In questo libro, quando nel seguito si parlerà di dimensione di uno spazio vettoriale, sarà *sempre sottinteso* che lo spazio vettoriale ha dimensione finita.

Esempio 2. Sia K un corpo. Allora K è uno spazio vettoriale su sé stesso ed è di dimensione 1. L'elemento 1 di K , infatti, costituisce una base di K su sé stesso perché ogni elemento x appartenente a K è univocamente esprimibile come $x = x \cdot 1$.

Daremo ora dei criteri che ci permetteranno di dire quando alcuni elementi di uno spazio vettoriale ne costituiscono una base.

Siano v_1, \dots, v_n elementi linearmente indipendenti di uno spazio

vettoriale V . Diremo che essi costituiscono un *insieme massimale di elementi linearmente indipendenti* di V se, quale che sia l'elemento w di V , gli elementi w, v_1, \dots, v_n risultano linearmente dipendenti.

TEOREMA 4 *Sia V uno spazio vettoriale e $\{v_1, \dots, v_n\}$ sia un insieme massimale di elementi linearmente indipendenti di V . Allora $\{v_1, \dots, v_n\}$ è una base di V .*

Dimostrazione. Dobbiamo dimostrare che v_1, \dots, v_n genera V , cioè che ogni elemento di V può essere espresso come combinazione lineare di v_1, \dots, v_n . Sia w un elemento di V . Poiché gli elementi w, v_1, \dots, v_n di V sono per ipotesi linearmente dipendenti, esistono numeri x_0, x_1, \dots, x_n non tutti nulli tali che

$$x_0w + x_1v_1 + \dots + x_nv_n = O.$$

Osserviamo che non può essere $x_0 = 0$ perché, se così fosse, avremmo ottenuto una relazione di dipendenza lineare tra v_1, \dots, v_n . Possiamo perciò esprimere w mediante v_1, \dots, v_n , abbiamo cioè:

$$w = -\frac{x_1}{x_0}v_1 - \dots - \frac{x_n}{x_0}v_n.$$

Questo dimostra che w è una combinazione lineare di v_1, \dots, v_n e quindi che $\{v_1, \dots, v_n\}$ è una base di V .

TEOREMA 5 *Sia V uno spazio vettoriale di dimensione n , e siano v_1, \dots, v_n elementi linearmente indipendenti di V . Allora i vettori v_1, \dots, v_n costituiscono una base di V .*

Dimostrazione. A causa del teorema 2, $\{v_1, \dots, v_n\}$ è un insieme massimale di elementi linearmente indipendenti di V . Questo insieme è quindi una base di V per il teorema 4.

COROLLARIO *Sia V uno spazio vettoriale e sia W un suo sottospazio. Se $\dim V = \dim W$ allora $V = W$.*

Dimostrazione. Sia $\{w_1, \dots, w_n\}$ una base di W . Poiché $n = \dim V$, invocando il teorema 5 possiamo concludere che $V = W$.

TEOREMA 6 *Sia V uno spazio vettoriale avente una base costituita da n elementi. Sia W un sottospazio di V non costituito dal solo O . Allora W ha una base e la sua dimensione non supera n .*

Dimostrazione. Sia w_1 un elemento non nullo di W . Se $\{w_1\}$ non è un insieme massimale di elementi linearmente indipendenti di W , possiamo trovare un elemento w_2 di W tale che w_1, w_2 risultino linearmente indipendenti. Continuando ad aggiungere in questo modo un elemento per volta, deve esistere un intero $m < n$ tale che si siano potuti trovare gli elementi w_1, w_2, \dots, w_m linearmente indipendenti e costituenti un insieme massimale di elementi linearmente indipendenti di W (il teorema 2 assicura che non è possibile continuare a trovare indefinitamente elementi linearmente indipendenti dai precedenti, giacché il numero di elementi siffatti che si possono trovare è al massimo n). Il teorema 4 permette allora di concludere che $\{w_1, \dots, w_m\}$ è una base di W .

TEOREMA 7 *Sia V uno spazio vettoriale di dimensione n e siano v_1, \dots, v_r elementi linearmente indipendenti di V . È possibile allora trovare in V degli elementi v_{r+1}, \dots, v_n in modo che $\{v_1, \dots, v_n\}$ sia una base di V .*

Dimostrazione. Se $r < n$, per la definizione di dimensione, v_1, \dots, v_r non possono costituire una base di V e quindi non possono generare V . Deve quindi esistere un elemento v_{r+1} di V che non appartiene al sottospazio generato da v_1, \dots, v_r . Allora gli elementi v_1, \dots, v_{r+1} sono linearmente indipendenti giacché, se valesse un'uguaglianza

$$a_1 v_1 + \dots + a_r v_r + a_{r+1} v_{r+1} = O,$$

con $a_i \in K$, a_{r+1} non potrebbe essere nullo (altrimenti avremmo ottenuto una relazione di dipendenza lineare tra gli elementi v_1, \dots, v_r). Ricavando allora v_{r+1} in termini di v_1, \dots, v_r , scrivendo cioè

$$v_{r+1} = -a_{r+1}^{-1}(a_1 v_1 + \dots + a_r v_r),$$

si otterebbe una contraddizione col fatto che v_{r+1} non appartiene al sottospazio generato da v_1, \dots, v_r .

Supponiamo quindi di aver trovato gli elementi v_{r+1}, \dots, v_s in modo che v_1, \dots, v_s risultino linearmente indipendenti. Per il teorema 2, s non supera n . Rendendo s il più grande possibile (continuando cioè ad aggiungere elementi v_t indipendenti dai precedenti finché è possibile) l'argomentazione precedente prova che $s = n$ e quindi che $\{v_1, \dots, v_n\}$ è una base di V .

11. SOMME E SOMME DIRETTE

Sia V uno spazio vettoriale sul corpo K . Siano U, W sottospazi di V . Definiamo *somma* di U e W il sottoinsieme di V costituito da tutte le somme $u + w$ con $u \in U$ e $w \in W$. Denotiamo questo insieme con $U + W$. Osserviamo che si tratta di un sottospazio di V , infatti, se u_1, u_2 appartengono a U e w_1, w_2 appartengono a W ,

$$(u_1 + w_1) + (u_2 + w_2) = u_1 + u_2 + w_1 + w_2 \in U + W.$$

Se $c \in K$, allora

$$c(u_1 + w_1) = cu_1 + cw_1 \in U + W.$$

infine, $O + O \in U + W$. Ciò prova che $U + W$ è un sottospazio di V .

Diremo poi che V è *somma diretta* di U e W se ogni elemento v di V può essere scritto in modo unico come somma di un elemento $u \in U$ e di un elemento $w \in W$.

TEOREMA 8 *Sia V uno spazio vettoriale sul corpo K , siano U, W suoi sottospazi. Se $U + W = V$ e se $U \cap W = \{O\}$, allora V è somma diretta di U e W .*

Dimostrazione. Se $v \in V$, per la prima parte dell'ipotesi esistono due elementi $u \in U, w \in W$ in modo che $v = u + w$. Quindi V è somma dei sottospazi U e W . Per dimostrare che la somma è diretta, dobbiamo dimostrare che gli elementi u e w trovati sono univocamente determinati. Supponiamo che esistano elementi $u' \in U$ e $w' \in W$ tali che $v = u' + w'$. Allora

$$u + w = u' + w'.$$

E quindi

$$u - u' = w' - w.$$

Ma $u - u' \in U$ e $w' - w \in W$; per la seconda ipotesi possiamo concludere che $u - u' = O$ e $w' - w = O$, da cui segue che $u = u'$ e $w = w'$, provando così il teorema.

La notazione adoperata per indicare che V è la somma diretta dei sottospazi U, W è la seguente

$$V = U \oplus W.$$

TEOREMA 9 *Sia V uno spazio vettoriale di dimensione finita sul corpo K . Sia W un suo sottospazio. Esiste allora in V un sottospazio U tale che V sia somma diretta di W e U .*

Dimostrazione. Scegliamo una base di W ed estendiamola a una base di V , con il teorema 7 del paragrafo 10. L'affermazione del teorema è allora evidente. Con le notazioni del teorema citato, se $\{v_1, \dots, v_r\}$ è una base di W , basta indicare con U lo spazio generato dai vettori $\{v_{r+1}, \dots, v_n\}$.

È bene osservare che per ogni sottospazio W possono esistere molti sottospazi U tali che V sia somma diretta di W ed U . (Per vederne esempi, guardare gli esercizi.) Nei paragrafi in cui, più avanti, discuteremo dell'ortogonalità, ne faremo uso per determinare questi sottospazi U .

TEOREMA 10 *Sia V uno spazio vettoriale di dimensione finita sul corpo K , se V è somma diretta dei sottospazi U, W allora:*

$$\dim V = \dim U + \dim W.$$

Dimostrazione. Sia $\{u_1, \dots, u_r\}$ una base di U e sia $\{w_1, \dots, w_s\}$ una base di W . Ogni elemento di U è univocamente esprimibile come una combinazione lineare $x_1u_1 + \dots + x_ru_r$, con $x_i \in K$, e ogni elemento di W è univocamente esprimibile come una combinazione lineare $y_1w_1 + \dots + y_sw_s$, con $y_j \in K$. Per definizione, perciò, ogni elemento di V ha un'unica espressione come combinazione lineare

$$x_1u_1 + \dots + x_ru_r + y_1w_1 + \dots + y_sw_s,$$

e questo prova che gli elementi $u_1, \dots, u_r, w_1, \dots, w_s$ costituiscono una base di V . Così il nostro teorema è dimostrato.

Osservazione. Possiamo anche definire V come somma diretta di più di due sottospazi. Siano W_1, \dots, W_r sottospazi di V : diremo che V è loro somma diretta se ogni elemento di V può essere espresso in un sol modo come una somma del tipo

$$v = w_1 + \dots + w_r$$

con $w_i \in W_i$.

Supponiamo ora che U, W siano spazi vettoriali qualsiasi sul corpo K (cioè U, W non sono necessariamente sottospazi di uno stesso spazio vettoriale). Indicato con $U \times W$ l'insieme di tutte le coppie ordinate (u, w) la cui prima componente è un elemento u di U e la cui seconda componente è un elemento w di W , definiamo l'addizione per queste coppie componenti per componente, cioè, se $(u_1, w_1) \in U \times W$ e $(u_2, w_2) \in U \times W$, definiamo

$$(u_1, w_1) + (u_2, w_2) = (u_1 + u_2, w_1 + w_2).$$

Se $c \in K$, definiamo il prodotto $c(u_1, w_1)$ come segue:

$$c(u_1, w_1) = (cu_1, cw_1).$$

Si verifica allora immediatamente che $U \times W$ è uno spazio vettoriale, chiamato *prodotto diretto* di U e W . Quando tratteremo delle applicazioni lineari, stabiliremo un confronto tra prodotto diretto e somma diretta.

Se n è un intero positivo che è somma di due altri interi positivi, $n = r + s$, si vede che K^n è il prodotto diretto $K^r \times K^s$.

Osserviamo poi che $\dim(U \times W) = \dim U + \dim W$. La dimostrazione, che è piuttosto semplice, è lasciata al lettore.

Esercizi

1. Sia $V = \mathbb{R}^3$, e sia W il sottospazio generato da $(2, 1)$. Sia U il sottospazio generato da $(0, 1)$. Dimostrare che V è somma diretta di W e U . Se poi U' è il sottospazio generato da $(1, 1)$ dimostrare che V è anche somma diretta di W e U' .
2. Sia $V = K^3$ per un certo corpo K . Sia W il sottospazio generato da $(1, 0, 0)$, sia U il sottospazio generato dagli elementi $(1, 1, 0)$ e $(0, 1, 1)$. Dimostrare che V è somma diretta di W e U .
3. Siano A, B due vettori di \mathbb{R}^2 e si assuma che nessuno di essi sia nullo. Se non esiste alcun numero c tale che $cA = B$, dimostrare che A, B costituiscono una base di \mathbb{R}^2 e che \mathbb{R}^2 stesso è somma diretta dei sottospazi generati rispettivamente da A e da B .

Capitolo 3

Matrici

12. LO SPAZIO DELLE MATRICI

Vogliamo ora considerare un nuovo tipo di oggetti: le matrici. Sia K un corpo. Siano n, m due interi positivi; mn numeri di K , disposti nel modo seguente

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix},$$

costituiscono una *matrice in K* . Possiamo abbreviare la notazione di questa matrice scrivendola (a_{ij}) , $i = 1, \dots, m$ e $j = 1, \dots, n$. Diciamo che si tratta di una matrice m per n o una matrice $m \times n$. La matrice sopra scritta ha *m righe e n colonne*. Per esempio, la sua prima colonna è

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$$

e la seconda riga è $(a_{21}, a_{22}, \dots, a_{2n})$. Chiamiamo a_{ij} il *termine ij* o la *componente ij* della matrice. Se denotiamo con A la matrice sopra scritta, allora la i -esima riga è denotata con A_i , ed è definita ponendo

$$A_i = (a_{i1}, a_{i2}, \dots, a_{in}).$$

La j -esima colonna è denotata con A^j ed è definita ponendo

$$A^j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Esempio 1. La seguente scrittura rappresenta una matrice 2×3

$$\begin{pmatrix} 1 & 1 & -2 \\ -1 & 4 & -5 \end{pmatrix}.$$

Essa ha due righe e tre colonne. Le righe sono $(1, 1, -2)$ e $(-1, 4, -5)$. Le colonne sono

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 4 \end{pmatrix}, \quad \begin{pmatrix} -2 \\ -5 \end{pmatrix}.$$

Quindi le righe di una matrice possono essere riguardate come n -uple e le colonne possono essere riguardate come m -uple verticali. Una m -upla verticale è chiamata talvolta un *vettore colonna*.

Un vettore (x_1, \dots, x_n) è una matrice $1 \times n$. Un vettore colonna

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

è una matrice $n \times 1$.

Quando scriviamo una matrice nella forma abbreviata (a_{ij}) , con i denotiamo la riga e con j denotiamo la colonna. Nell'esempio 1, per esempio, abbiamo $a_{11} = 1$, $a_{23} = -5$.

Un elemento singolo di K , (a) , può essere considerato una matrice 1×1 .

Sia (a_{ij}) , $i = 1, \dots, m$ e $j = 1, \dots, n$, una matrice; se $m = n$, diciamo che si tratta di una matrice *quadrata*. Per esempio

$$\begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} 1 & -1 & 5 \\ 2 & 1 & -1 \\ 3 & 1 & -1 \end{pmatrix}$$

sono entrambe matrici quadrate.

Esiste anche una *matrice zero*: quella in cui, per ogni i e j ,

$a_{ij} = 0$. Essa si scrive quindi come segue:

$$\begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

La denoteremo con O . Osserviamo che finora abbiamo incontrato il numero zero, il vettore zero e la matrice zero.

Vogliamo ora definire l'addizione tra matrici e la moltiplicazione di una matrice per uno scalare.

Definiamo l'addizione tra matrici soltanto quando esse hanno le stesse dimensioni. Siano quindi m e n due interi positivi fissati. Siano $A = (a_{ij})$ e $B = (b_{ij})$ due matrici $m \times n$. Definiamo $A + B$ come la matrice il cui termine sulla i -esima riga e la j -esima colonna è $a_{ij} + b_{ij}$. In altre parole, noi addizioniamo matrici di uguali dimensioni componente per componente.

Esempio 2. Sia

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 3 & 4 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 5 & 1 & -1 \\ 2 & 1 & -1 \end{pmatrix}.$$

Allora

$$A + B = \begin{pmatrix} 6 & 0 & -1 \\ 4 & 4 & 3 \end{pmatrix}.$$

Se A , B sono entrambe matrici $1 \times n$, cioè n -uple, osserviamo che l'addizione tra matrici che ora abbiamo definito coincide con l'addizione tra n -uple che abbiamo definito in precedenza.

Se O è la matrice zero, allora per ogni matrice A (avente le stesse dimensioni, naturalmente) noi abbiamo $O + A = A + O = A$, e questo si verifica immediatamente.

Andiamo ora a definire la moltiplicazione di una matrice per uno scalare. Sia c uno scalare e $A = (a_{ij})$ una matrice. Definiamo cA come la matrice la cui componente ij è ca_{ij} . Scriviamo $cA = (ca_{ij})$. In altre parole moltiplichiamo ogni componente di A per c .

Esempio 3. Siano A , B come nell'esempio 2. Sia $c = 2$. Allora

$$2A = \begin{pmatrix} 2 & -2 & 0 \\ 4 & 6 & 8 \end{pmatrix} \quad \text{e} \quad 2B = \begin{pmatrix} 10 & 2 & -2 \\ 4 & 2 & -2 \end{pmatrix}.$$

Abbiamo inoltre

$$(-1)A = -A = \begin{pmatrix} -1 & 1 & 0 \\ -2 & -3 & -4 \end{pmatrix}.$$

Per ogni matrice A , è facile verificare che $A + (-1)A = O$.

Lasciamo come esercizio la verifica di tutte le proprietà SV 1 ... SV 8 tenendo conto delle regole date per l'addizione tra matrici e la moltiplicazione di una matrice per un numero. La cosa più importante da osservare qui è che l'addizione tra matrici è definita mediante quella delle componenti, e per l'addizione delle componenti le corrispondenti proprietà SV 1 ... SV 4 valgono in quanto si tratta delle usuali proprietà dei numeri. Analogamente le proprietà SV 5 ... SV 8 sono valide per la moltiplicazione di matrici per numeri in quanto sono valide le proprietà corrispondenti per la moltiplicazione tra numeri.

Possiamo quindi riconoscere che le matrici (di date dimensioni $m \times n$) in K formano uno spazio vettoriale su K , che noi denoteremo $\mathcal{M}_{m,n}(K)$ o, se il riferimento a K è chiaro, semplicemente con $\mathcal{M}_{m,n}$.

Definiamo ora un'altra nozione relativa alle matrici. Sia $A = (a_{ij})$ una matrice $m \times n$. La matrice $n \times m$ $B = (b_{ji})$ tale che $b_{ji} = a_{ij}$ è chiamata la *trasposta* di A ed è denotata con $'A$. Considerare la trasposta di una matrice significa scambiarne le righe con le colonne. Se A è la matrice che abbiamo scritta all'inizio di questo paragrafo, allora $'A$ è la matrice

$$\begin{pmatrix} a_{11} & a_{21} & a_{31} & \dots & a_{m1} \\ a_{12} & a_{22} & a_{32} & \dots & a_{m2} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \dots & a_{mn} \end{pmatrix}.$$

Per considerare un caso particolare, si osservi che

$$\text{se } A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 5 \end{pmatrix} \text{ allora } 'A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \\ 0 & 5 \end{pmatrix}.$$

Se $A = (2, 1, -4)$ è un vettore riga, allora

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

è un vettore colonna.

Una matrice A si dice *simmetrica* quando coincide con la sua trasposta, cioè quando ${}^t A = A$. Una matrice simmetrica è necessariamente quadrata. Per esempio la matrice

$$\begin{pmatrix} 1 & -1 & 2 \\ -1 & 0 & 3 \\ 2 & 3 & 7 \end{pmatrix}$$

è simmetrica.

Sia $A = (a_{ij})$ una matrice quadrata. Chiamiamo a_{11}, \dots, a_{nn} le sue componenti diagonali. Una matrice quadrata si dice matrice *diagonale* se tutte le sue componenti, con l'eventuale eccezione di quelle diagonali, sono nulle, cioè se $a_{ij} = 0$ quando $i \neq j$. Ogni matrice diagonale è una matrice simmetrica. Una matrice diagonale è scritta nel modo seguente

$$\begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_n \end{pmatrix}.$$

Noi definiamo la *matrice unità* $n \times n$ come la matrice quadrata le cui componenti sono tutte nulle escluse quelle diagonali che, invece, sono uguali a uno. Questa matrice unità è denotata con I_n , o semplicemente con I quando non c'è necessità di specificare il numero n . Quindi

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Esercizi

1. Sia

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 2 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} -1 & 5 & -2 \\ 2 & 2 & -1 \end{pmatrix}.$$

Determinare $A + B$, $3B$, $-2B$, $A + 2B$, $2A + B$, $A - B$, $A - 2B$, $B - A$.

2. Sia

$$A = \begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} -1 & 1 \\ 0 & -3 \end{pmatrix}.$$

Determinare $A + B$, $3B$, $-2B$, $A + 2B$, $A - B$, $B - A$.

3. Per le matrici dell'esercizio 1 determinare ${}^t A$ e ${}^t B$.
 4. Per le matrici dell'esercizio 2 determinare ${}^t A$ e ${}^t B$.
 5. Siano A, B matrici arbitrarie $m \times n$; dimostrare allora che

$${}^t(A+B) = {}^t A + {}^t B.$$

6. Se c è un numero, dimostrare che si ha

$${}^t(cA) = c {}^t A.$$

7. Se $A = (a_{ij})$ è una matrice quadrata, allora gli elementi a_{ii} sono chiamati gli elementi diagonali. In che cosa differiscono gli elementi diagonali delle matrici A e ${}^t A$?

8. Per le matrici dell'esercizio 2 determinare ${}^t(A+B)$ e ${}^t A + {}^t B$.
 9. Per le matrici dell'esercizio 2 determinare $A + {}^t A$ e $B + {}^t B$.
 10. Dimostrare che per ogni matrice quadrata A , la matrice $A + {}^t A$ è simmetrica.
 11. Scrivere i vettori riga e i vettori colonna delle matrici A, B dell'esercizio 1.
 12. Scrivere i vettori riga e i vettori colonna delle matrici A, B dell'esercizio 2.
 13. Qual è la dimensione dello spazio vettoriale $\mathfrak{M}_{m,n}(K)$ delle matrici $m \times n$ sul corpo K ?
 14. Sia e_{ij} ($i = 1, \dots, m$ e $j = 1, \dots, n$) la matrice la cui componente ij è uguale a 1 mentre tutte le altre componenti sono nulle. Dimostrare che queste matrici e_{ij} costituiscono una base di $\mathfrak{M}_{m,n}(K)$ su K .
 15. Le matrici diagonali $n \times n$ su un corpo K formano uno spazio vettoriale. Quale ne è la dimensione?
 16. Una matrice quadrata $n \times n$ si chiama *triangolare superiore* se tutte le sue componenti al di sotto della diagonale sono uguali a zero. Una tale matrice si scrive quindi come segue

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}.$$

Qual è la dimensione dello spazio di tutte le matrici triangolari superiori di dimensioni $n \times n$?

13. EQUAZIONI LINEARI

Applicheremo ora i teoremi sulla dimensione alla risoluzione di equazioni lineari.

Sia K un corpo e sia $A = (a_{ij})$, $i = 1, \dots, m$ e $j = 1, \dots, n$, una matrice in K . Siano b_1, \dots, b_m elementi di K . Equazioni della forma

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= b_1 \\ &\dots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= b_m \end{aligned} \quad [1]$$

sono chiamate equazioni lineari. Diremo anche che [1] è un *sistema* di equazioni lineari. Il sistema si dice *omogeneo* se tutti i numeri b_1, \dots, b_m sono uguali a zero. Il numero n è chiamato numero delle *incognite*, il numero m è chiamato numero delle equazioni. La matrice (a_{ij}) è chiamata la matrice dei *coefficientsi*.

Il sistema di equazioni

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= 0 \\ &\dots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= 0 \end{aligned} \quad [2]$$

si chiama *sistema omogeneo associato* al sistema [1].

Il sistema [2] ha sempre soluzione, infatti basta porre ogni $x_j = 0$. Questa soluzione viene detta soluzione *banale*. Una soluzione (x_1, \dots, x_n) , in cui qualche x_j è diverso da zero, sarà detta *non banale*.

Consideriamo dapprima il sistema omogeneo [2], lo possiamo intanto riscrivere nel seguente modo

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = O,$$

oppure, con la notazione dei vettori colonna della matrice $A = (a_{ij})$,

$$x_1 A^1 + \dots + x_n A^n = O.$$

Una soluzione non banale $X = (x_1, \dots, x_n)$ del nostro sistema [2] è quindi nient'altro che una n -upla $X \neq O$ che stabilisce una relazione di dipendenza lineare tra le colonne A^1, \dots, A^n . Vedremo

che questo modo di riscrivere il sistema ci sarà molto utile e ci permetterà di utilizzare il teorema 2 del capitolo 2 (§ 10). I vettori colonna sono elementi di K^m , la cui dimensione su K è m . Possiamo quindi affermare che:

TEOREMA 1 *Sia*

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= 0 \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= 0 \end{aligned}$$

un sistema omogeneo di m equazioni lineari in n incognite, i coefficienti appartengano a un corpo K . Si supponga $n > m$. Allora il sistema dato possiede soluzioni non banali in K .

Dimostrazione. Per il teorema 2 del capitolo 2, siamo sicuri che i vettori A^1, \dots, A^n sono linearmente dipendenti.

Naturalmente, per risolvere esplicitamente un sistema di equazioni lineari per ora non abbiamo altro modo che quello elementare, appreso negli studi precedenti, dell'eliminazione successiva delle incognite. Ne diamo ora un esempio.

Supponiamo dapprima di avere una sola equazione, sia

$$2x + y - 4z = 0.$$

Per trovarne una soluzione non banale, assegniamo a tutte le variabili tranne la prima un valore numerico diverso da zero, per esempio $y = 1$, $z = 1$. Risolviamo quindi rispetto ad x . Troviamo $2x = -y + 4z = 3$, da cui $x = \frac{3}{2}$.

Consideriamo poi un sistema di due equazioni, sia

$$2x + 3y - z = 0, \quad [3]$$

$$x + y + z = 0. \quad [4]$$

Riduciamo il problema di risolvere queste due equazioni simultanee al caso, precedentemente trattato, di una sola equazione eliminando una variabile. Per esempio, moltiplicando l'equazione [4] per 2 e sottraendola dalla [3], otteniamo

$$y - 3z = 0. \quad [5]$$

Troviamo così un'equazione con più di una incognita. Fissiamo per z un valore non nullo, per esempio $z = 1$, e risolviamo ri-

spetto a y , ottenendo $y = 3$. Possiamo allora risolvere rispetto a x l'equazione [4] ottenendo $x = -4$. I valori che abbiamo ottenuto per x , y , z sono anche soluzioni della prima equazione perché questa è (in un senso ovvio) la somma dell'equazione [4] moltiplicata per 2 e dell'equazione [5]. Una descrizione di questo procedimento relativo al caso generale è riportata nell'appendice a questo capitolo, dove viene inoltre riportata una dimostrazione diretta del teorema 1.

Daremo in seguito dei metodi più efficienti per trovare le soluzioni di equazioni lineari.

Consideriamo ora l'originario sistema di equazioni [1]. Sia B il vettore colonna

$$B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Allora possiamo riscrivere il sistema [1] nella forma

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

o, più brevemente, adoperando i vettori colonna di A

$$x_1 A^1 + \dots + x_n A^n = B.$$

Dire che $X = (x_1, \dots, x_n)$ è una soluzione del sistema di equazioni lineari dato significa dire che B è una combinazione lineare dei vettori A^1, \dots, A^n .

TEOREMA 2 *Nel sistema dato [1] si assuma che m e n coincidano e che i vettori A^1, \dots, A^n siano linearmente indipendenti. Allora il sistema [1] ha in K un'unica soluzione.*

Dimostrazione. Poiché i vettori A^1, \dots, A^n sono linearmente indipendenti, essi costituiscono una base di K^n . Ne segue che ogni vettore B è esprimibile in modo unico come combinazione lineare

$$B = x_1 A^1 + \dots + x_n A^n,$$

con $x_i \in K$, quindi $X = (x_1, \dots, x_n)$ è l'unica soluzione del sistema.

Esercizi

1. Sia [2] un sistema di equazioni lineari omogenee in un corpo K e sia $m = n$. Si supponga poi che i vettori colonna dei coefficienti siano linearmente indipendenti. Si dimostri che il sistema ha come unica soluzione quella banale.
2. Sia [2] un sistema di equazioni lineari omogenee nel corpo K , con n incognite. Dimostrare che l'insieme delle sue soluzioni $X = (x_1, \dots, x_n)$ è uno spazio vettoriale su K .
3. Risolvere i seguenti sistemi di equazioni lineari in \mathbb{R} :
- | | |
|------------------|----------------------|
| a) $2x + 3y = 5$ | b) $2x + 3y + z = 0$ |
| $4x - y = 7$. | $x - 2y - z = 1$ |
| | $x + 4y + z = 2$. |
4. Risolvere i seguenti sistemi di equazioni lineari in \mathbb{C} :
- | | |
|---------------------|---------------------------|
| a) $ix - 2y = 1$ | b) $2x + iy - (1+i)z = 1$ |
| $x + iy = 2$. | $x - 2y + iz = 0$ |
| | $-ix + y - (2-i)z = 1$. |
| c) $(1+i)x - y = 0$ | d) $ix - (2+i)y = 1$ |
| $ix + y = 3 - i$. | $x + (2-i)y = 1 + i$. |
5. Siano A^1, \dots, A^n vettori colonna di dimensione m . Si supponga che essi abbiano coefficienti reali e che siano linearmente indipendenti su \mathbb{R} . Dimostrare che i vettori dati sono linearmente indipendenti anche su \mathbb{C} .
6. Sia [2] un sistema di equazioni lineari omogenee con coefficienti reali. Dimostrare che se questo sistema ha una soluzione non banale in \mathbb{C} , ne ha una non banale anche in \mathbb{R} .

14. MOLTIPLICAZIONE TRA MATRICI

Considereremo ora matrici su un corpo K fissato. Cominciamo col notare che la definizione di prodotto scalare data nel primo capitolo per vettori a coefficienti reali si può applicare anche ai vettori con componenti in K . Quindi, se $A = (a_1, \dots, a_n)$ e $B = (b_1, \dots, b_n)$ sono in K^n , definiamo

$$A \cdot B = a_1 b_1 + \dots + a_n b_n.$$

Questo è un elemento di K e le proprietà fondamentali continuano a valere.

PS 1. Se A, B sono in K^n , abbiamo $A \cdot B = B \cdot A$.

PS 2. Se A, B, C sono in K^n , allora

$$A \cdot (B + C) = A \cdot B + A \cdot C = (B + C) \cdot A.$$

PS 3. Se $x \in K$, allora

$$(xA) \cdot B = x(A \cdot B) \quad e \quad A \cdot (xB) = x(A \cdot B).$$

Osserviamo però che la proprietà di positività non vale in generale: per esempio, se $K = \mathbb{C}$, se $A = (1, i)$, allora $A \neq O$ ma

$$A \cdot A = 1 + i^2 = 0.$$

In molte applicazioni la proprietà di positività non è necessaria e in sua vece si può adoperare una proprietà che chiameremo di *non degenerazione*, cioè:

Se $A \in K^n$ e se per ogni $X \in K^n$ $A \cdot X = 0$, allora $A = O$.

La dimostrazione della validità di questa proprietà è ovvia, giacché si ha $A \cdot E_i = 0$ per ogni vettore unitario $E_i = (0, \dots, 0, 1, 0, \dots, 0)$ con 1 nel posto i -esimo e 0 altrove: ma $A \cdot E_i = a_i$, quindi, per ogni i , $a_i = 0$ e questo significa che $A = O$.

Passiamo ora a definire il prodotto tra matrici.

Sia $A = (a_{ij})$, $i = 1, \dots, m$ e $j = 1, \dots, n$, una matrice $m \times n$.
Sia $B = (b_{jk})$, $j = 1, \dots, n$ e $k = 1, \dots, s$, una matrice $n \times s$.

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & \dots & b_{1s} \\ \vdots & \ddots & \vdots \\ b_{n1} & \dots & b_{ns} \end{pmatrix}.$$

Definiamo il prodotto AB come la matrice $m \times s$ la cui coordinata ik è

$$\sum_{j=1}^n a_{ij} b_{jk} = a_{i1} b_{1k} + a_{i2} b_{2k} + \dots + a_{in} b_{nk}$$

Se A_1, \dots, A_m sono i vettori riga della matrice A e se B^1, \dots, B^s sono i vettori colonna della matrice B , allora la coordinata ik del prodotto AB è anche uguale ad $A_i \cdot B^k$. Quindi

$$AB = \begin{pmatrix} A_1 \cdot B^1 & \dots & A_1 \cdot B^s \\ \vdots & \ddots & \vdots \\ A_m \cdot B^1 & \dots & A_m \cdot B^s \end{pmatrix}.$$

Di qui appare che la moltiplicazione tra matrici è una generalizzazione del prodotto scalare.

Esempio 1. Sia

$$A = \begin{pmatrix} 2 & 1 & 5 \\ 1 & 3 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 4 \\ -1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Allora AB è una matrice 2×2 e il calcolo mostra che

$$AB = \begin{pmatrix} 2 & 1 & 5 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 3 & 4 \\ -1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 15 & 15 \\ 4 & 12 \end{pmatrix}.$$

Esempio 2. Sia

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

Siano A , B le matrici dell'esempio 1. Allora

$$BC = \begin{pmatrix} 3 & 4 \\ -1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 5 \\ -3 & -5 \\ 1 & 5 \end{pmatrix}$$

e

$$A(BC) = \begin{pmatrix} 2 & 1 & 5 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} -1 & 5 \\ -3 & -5 \\ 1 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 30 \\ -8 & 0 \end{pmatrix}.$$

Si calcoli $(AB)C$. Che cosa si trova?

Sia A una matrice $m \times n$ e sia B una matrice $n \times 1$, cioè un vettore colonna; allora AB è di nuovo un vettore colonna.

Sia A una matrice $1 \times n$, cioè un vettore riga, sia B una matrice $n \times s$; allora AB è un vettore riga.

TEOREMA 3 *Siano A , B , C matrici. Si assuma che A e B possano essere moltiplicate, che A e C possano essere moltiplicate, che B e C possano essere addizionate. Allora le matrici A , $B + C$ possono essere moltiplicate e si ha:*

$$A(B + C) = AB + AC.$$

Se x è un numero, allora

$$A(xB) = x(AB).$$

Dimostrazione. Sia A_i la i -esima riga di A e siano B^k , C^k le colonne k -esime di B e C rispettivamente. Allora $B^k + C^k$ è la k -esima colonna di $B + C$. Per definizione, la componente ik di AB è $A_i \cdot B^k$, la componente ik di AC è $A_i \cdot C^k$, la componente ik di $A(B + C)$ è $A_i \cdot (B^k + C^k)$. Poiché

$$A_i \cdot (B^k + C^k) = A_i \cdot B^k + A_i \cdot C^k,$$

la prima affermazione del teorema è provata. Nei riguardi della seconda, si osservi che la k -esima colonna della matrice xB è xB^k . Poiché

$$A_i \cdot xB^k = x(A_i \cdot B^k),$$

la seconda affermazione è provata.

TEOREMA 4 *Siano A , B , C matrici tali che A , B possano essere moltiplicate e B , C possano essere moltiplicate. Allora le matrici A , BC possono essere moltiplicate e così anche le matrici AB , C , e si ha*

$$(AB)C = A(BC).$$

Dimostrazione. Sia $A = (a_{ij})$ una matrice $m \times n$, sia $B = (b_{jk})$ una matrice $n \times r$, sia $C = (c_{kl})$ una matrice $r \times s$; allora il prodotto AB è una matrice $m \times r$ la cui componente ik è uguale alla somma

$$a_{i1}b_{1k} + a_{i2}b_{2k} + \dots + a_{in}b_{nk}.$$

Scriveremo abbreviatamente questa somma nel modo seguente

$$\sum_{j=1}^n a_{ij}b_{jk}.$$

Per definizione, la componente il -esima di $(AB)C$ è uguale a

$$\sum_{k=1}^r \left[\sum_{j=1}^n a_{ij}b_{jk} \right] c_{kl} = \sum_{k=1}^r \left[\sum_{j=1}^n a_{ij}b_{jk} c_{kl} \right].$$

La somma nel secondo membro può essere anche riguardata come

la somma di tutti i termini

$$a_{ij}b_{jk}c_{kl},$$

dove j, k sono interi tali che $1 \leq j < n$ e $1 \leq k < r$.

Se fossimo partiti con la componente jl di BC e avessimo quindi calcolato la componente il -esima di $A(BC)$, avremmo trovato esattamente la stessa somma; il teorema è così dimostrato.

Sia A una matrice quadrata $n \times n$. Diremo che A è *invertibile* oppure *non singolare* se esiste una matrice $n \times n$ B tale che

$$AB = BA = I_n.$$

Una tale matrice B è univocamente determinata da A in quanto se C è una matrice per cui $AC = CA = I_n$, allora

$$B = BI_n = B(AC) = (BA)C = I_n C = C.$$

(Vedi esercizio 1.) Questa matrice B sarà chiamata l'*inversa* di A e sarà denotata con A^{-1} . Dopo aver studiato i determinanti, saremo in grado di poterla scrivere esplicitamente, se esiste.

Sia A una matrice quadrata. Possiamo allora formare il prodotto di A con sé stessa, cioè AA , o formare il prodotto

$$A \dots A$$

prendendo A m volte. Definiremo, per ogni intero $m \geq 1$, A^m come il prodotto $A \dots A$ con m fattori. Definiamo poi $A^0 = I$ (la matrice unità di dimensioni uguali a quella di A) e $A^1 = A$. La regola solita, $A^{r+s} = A^r A^s$, vale per ogni coppia di interi non negativi r, s .

Il risultato seguente mette in relazione la trasposizione con la moltiplicazione di matrici.

TEOREMA 5 *Siano A, B matrici che possono essere moltiplicate. Allora anche le matrici ${}^t B$, ${}^t A$ possono essere moltiplicate e si ha*

$${}^t(AB) = {}^tB {}^tA.$$

Dimostrazione. Sia $A = (a_{ij})$, $B = (b_{jk})$. Sia $AB = C$. Allora

$$c_{ik} = \sum_{j=1}^n a_{ij} b_{jk}.$$

Sia ${}^t B = (b'_{kj})$, ${}^t A = (a'_{ji})$. Allora la componente ki del prodotto

tB^tA è per definizione

$$\sum_{j=1}^n b'_{kj} a'_{jt}.$$

Poiché $b'_{kj} = b_{jk}$ e $a'_{jt} = a_{ij}$, vediamo che quest'ultima espressione è uguale a

$$\sum_{j=1}^n b_{jk} a_{ij} = \sum_{j=1}^n a_{ij} b_{jk}.$$

Per definizione, questa è la componente k di tC , come era da dimostrare.

Esercizi

1. Sia I la matrice unità $n \times n$. Sia A una matrice $n \times r$. Che cosa è IA ? Se A è una matrice $n \times n$, che cosa è AI ?

2. Sia O la matrice le cui coordinate sono tutte nulle. Sia A una matrice di dimensioni tali che il prodotto AO sia definito. Che cosa è AO ?

3. In ognuno dei casi seguenti trovare $(AB)C$ e $A(BC)$.

a) $A = \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$.

b) $A = \begin{pmatrix} 2 & 1 & -1 \\ 3 & 1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 3 & -1 \end{pmatrix}$, $C = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

c) $A = \begin{pmatrix} 2 & 4 & 1 \\ 3 & 0 & -1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & -1 \\ 3 & 1 & 5 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 2 \\ 3 & 1 \\ -1 & 4 \end{pmatrix}$.

4. Siano A , B matrici quadrate di uguali dimensioni e si supponga $AB = BA$. Dimostrare che $(A + B)^2 = A^2 + 2AB + B^2$ e che

$$(A + B)(A - B) = A^2 - B^2,$$

adoperando le proprietà delle matrici enunciate nel teorema 3.

5. Sia

$$A = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}.$$

Trovare AB e BA .

6. Sia $C = \begin{pmatrix} 7 & 0 \\ 0 & 7 \end{pmatrix}$. Siano A , B le matrici dell'esercizio 5. Trovare CA ,

AC , CB e BC . Enunciare la regola generale di cui questo esercizio è un caso particolare.

7. Sia $X = (1, 0, 0)$ e sia

$$A = \begin{pmatrix} 3 & 1 & 5 \\ 2 & 0 & 1 \\ 1 & 1 & 7 \end{pmatrix},$$

cosa è XA ?

8. Sia $X = (0, 1, 0)$ e sia A un'arbitraria matrice 3×3 . Come si potrebbe descrivere XA ? E se $X = (0, 0, 1)$? Generalizzare considerando matrici $n \times n$ e i loro prodotti con vettori unità.

9. Siano A, B le matrici dell'esercizio 3a. Verificare, eseguendo i calcoli, che ${}^t(AB) = {}^tB{}^tA$. Fare lo stesso nei riguardi degli esercizi 3b e 3c. Se A, B, C sono matrici che possono essere moltiplicate, dimostrare che ${}^t(ABC) = {}^tC{}^tB{}^tA$.

10. Sia M una matrice $n \times n$ tale che ${}^tM = M$. Assegnati due vettori colonna nell' n -spazio, siano essi A e B , si definisca $\langle A, B \rangle$ come il prodotto tAMB . (Si identifichi una matrice 1×1 con un numero.) Dimostrare che le proprietà del prodotto scalare sono valide, con la possibile eccezione della proprietà di positività. Dare un esempio di una matrice M e di due vettori A, B tali che tAMB sia negativo (considerando il caso $n = 2$).

11. Sia A la matrice

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

Trovare A^2 e A^3 . Generalizzare alle matrici 4×4 .

12. Sia A la matrice

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}.$$

Trovare A^2, A^3, A^n per ogni intero positivo n .

13. Sia A una matrice invertibile $n \times n$. Dimostrare che

$${}^t(A^{-1}) = ({}^tA)^{-1}.$$

È quindi possibile scrivere ${}^tA^{-1}$ senza pericolo di confusioni.

14. Sia A una matrice a coefficienti complessi, $A = (a_{ij})$, e sia $\bar{A} = (\bar{a}_{ij})$, avendo indicato con il soprassegno il numero complesso coniugato. Dimostrare che

$${}^t(\bar{A}) = \bar{{}^tA}.$$

Quindi possiamo semplicemente scrivere ${}^t\bar{A}$.

15. Sia A una matrice diagonale:

$$A = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_n \end{pmatrix}.$$

Se per ogni i , $a_i \neq 0$, dimostrare che A è invertibile. Qual è la sua inversa?

16. Sia A una matrice *strettamente triangolare superiore* $n \times n$, cioè una matrice quadrata (a_{ij}) in cui tutte le componenti sulla diagonale e sotto di essa sono uguali a 0. Possiamo esprimere tutto ciò scrivendo $a_{ij} = 0$, se $i > j$:

$$A = \begin{pmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & 0 & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Dimostrare che $A^n = O$. (Se si ritiene conveniente, si può fare la dimostrazione soltanto nei casi $n = 2, 3, 4$. Il caso generale può essere trattato con il metodo di induzione.)

17. Sia A una matrice triangolare $n \times n$ le cui componenti sulla diagonale siano uguali ad uno:

$$A = \begin{pmatrix} 1 & a_{12} & \dots & a_{1n} \\ 0 & 1 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & a_{n-1,n} \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

Sia $N = A - I_n$. Dimostrare che $N^{n+1} = O$. Si noti che $A = I + N$. Dimostrare che la matrice A è invertibile e che la sua inversa è data da:

$$(I + N)^{-1} = I - N + N^2 - \dots + (-1)^n N^n.$$

18. Se N è una matrice quadrata tale che $N^{r+1} = O$ per un certo intero positivo r , dimostrare che la matrice $I - N$ è invertibile e che la sua inversa è $I + N + \dots + N^r$.

19. Sia A una matrice triangolare:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}.$$

Si supponga che nessun elemento diagonale sia nullo e sia

$$B = \begin{pmatrix} a_{11}^{-1} & 0 & \dots & 0 \\ 0 & a_{22}^{-1} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{nn}^{-1} \end{pmatrix}.$$

Dimostrare che BA e AB sono matrici triangolari le cui componenti diagonali sono uguali a uno.

20. Sia $A = (a_{ij})$ una matrice quadrata $n \times n$. Si definisca sua *traccia*, denotata con $\text{tr}(A)$, la somma dei suoi elementi diagonali,

$$\text{tr}(A) = a_{11} + \dots + a_{nn}.$$

Per esempio, la traccia di

$$\begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$$

è uguale a $1 + 5 = 6$.

Determinare la traccia delle matrici A , B , AB , e BA dell'esercizio 5.

21. a) Siano A, B matrici $n \times n$. Dimostrare che $\text{tr}(AB) = \text{tr}(BA)$.
 b) Se B è invertibile, dimostrare che $\text{tr}(B^{-1}AB) = \text{tr}(A)$.

22. Una matrice quadrata A viene detta *nilpotente* se esiste un intero positivo r tale che $A^r = O$. Siano A, B matrici nilpotenti aventi le stesse dimensioni e tali che $AB = BA$. Dimostrare che le matrici AB e $A + B$ sono nilpotenti.

15. APPENDICE: ELIMINAZIONE

Nella dimostrazione del teorema 1 relativo alla risoluzione di un sistema di equazioni lineari omogenee, abbiamo fatto uso del teorema che afferma che n elementi di uno spazio vettoriale di dimensione m sono sicuramente linearmente dipendenti se n supera m . Viceversa, è possibile dare una dimostrazione diretta del teorema 1 adoperando il metodo di eliminazione e dimostrare quindi il teorema sulla dimensione. Daremo ora i dettagli di questo procedimento come alternativa a quello dato nel testo.

Consideriamo nuovamente il sistema di equazioni omogenee [2]. Siano A_1, \dots, A_m i vettori riga della matrice (a_{ij}) . Allora possiamo scrivere le equazioni del sistema [2] nella forma

$$\begin{aligned} A_1 \cdot X &= 0 \\ &\vdots \\ A_m \cdot X &= 0. \end{aligned} \tag{2a}$$

Faremo uso della notazione del prodotto scalare, perché così sarà facilitata l'esposizione della dimostrazione del nostro teorema.

TEOREMA *Sia*

$$\begin{aligned} a_{11} x_1 + \dots + a_{1n} x_n &= 0 \\ &\vdots \\ a_{m1} x_1 + \dots + a_{mn} x_n &= 0 \end{aligned}$$

un sistema di m equazioni lineari in n incognite e si supponga che n sia maggiore di m . Il sistema ammette allora soluzioni non banali.

Dimostrazione. Nella dimostrazione faremo uso del metodo di induzione, seguiremo cioè un processo iterativo.

Si consideri dapprima il caso di una equazione in n incognite, $n > 1$:

$$a_1 x_1 + \dots + a_n x_n = 0.$$

Se tutti i coefficienti a_1, \dots, a_n sono nulli, allora ogni valore sostituito alle variabili darà una soluzione e quindi una soluzione non banale esiste certamente. Supponiamo ora che qualche coefficiente a_i sia diverso da zero. Numerando diversamente le variabili e i coefficienti, possiamo senz'altro supporre che si tratti di a_1 . Diamo allora a x_2, \dots, x_n valori arbitrari, per esempio poniamo $x_2 = \dots = x_n = 1$, e risolviamo rispetto a x_1 ottenendo

$$x_1 = \frac{-1}{a_1} (a_2 + \dots + a_n).$$

In questo modo otteniamo una soluzione sicuramente non banale del nostro sistema di equazioni.

Assumiamo ora che il nostro teorema sia vero per un sistema di $m - 1$ equazioni in più di $m - 1$ incognite. Proveremo che il teorema è vero per un sistema di m equazioni in n incognite quando n supera m .

Consideriamo il sistema [2a].

Se tutti i coefficienti (a_{ij}) sono nulli, per ottenere una soluzione basta assegnare alle variabili valori arbitrari (quindi anche non nulli). Se qualche coefficiente non è nullo, mutando la numerazione delle equazioni e delle variabili possiamo assumere che si tratti di a_{11} . Sottrarremo un multiplo della prima equazione dalle altre per eliminare x_1 . In altre parole consideriamo il si-

stema di equazioni

$$\begin{aligned} \left(A_2 - \frac{a_{21}}{a_{11}} A_1 \right) \cdot X &= 0 \\ &\vdots \\ \left(A_m - \frac{a_{m1}}{a_{11}} A_1 \right) \cdot X &= 0, \end{aligned}$$

che può essere anche scritto nella forma

$$\begin{aligned} A_2 \cdot X - \frac{a_{21}}{a_{11}} A_1 \cdot X &= 0 \\ &\vdots \\ A_m \cdot X - \frac{a_{m1}}{a_{11}} A_1 \cdot X &= 0. \end{aligned} \quad [6]$$

In questo sistema il coefficiente di x_1 è nullo. Possiamo quindi considerare [6] come un sistema di $m-1$ equazioni in $n-1$ incognite, sapendo che $n-1$ è maggiore di $m-1$.

Per l'ipotesi fatta, questo sistema ammette una soluzione non banale (x_2, \dots, x_n). Possiamo allora risolvere rispetto a x_1 la prima equazione, ottenendo

$$x_1 = \frac{-1}{a_{11}} (a_{12}x_2 + \dots + a_{1n}x_n).$$

Abbiamo così trovato una soluzione di $A_1 \cdot X = 0$. Ma, tenendo conto di [6], abbiamo

$$A_i \cdot X = \frac{a_{i1}}{a_{11}} A_1 \cdot X$$

per $i = 2, \dots, m$. Quindi $A_i \cdot X = 0$ anche per $i = 2, \dots, m$: abbiamo così trovato una soluzione non banale del nostro sistema originario [2a].

Il procedimento che abbiamo ora seguito si può iterare permettendoci di passare da una a due equazioni, poi da due a tre, e così via.

Il teorema è così dimostrato.

COROLLARIO *Sia V uno spazio vettoriale e $\{v_1, \dots, v_m\}$ ne sia una base. Siano w_1, \dots, w_n elementi di V e si assuma che n sia mag-*

giore di m . Allora gli elementi w_1, \dots, w_n sono linearmente dipendenti.

Dimostrazione. Poiché $\{v_1, \dots, v_m\}$ è una base, esistono i numeri (a_{ij}) in modo da poter scrivere

$$\begin{aligned} w_1 &= a_{11}v_1 + \dots + a_{m1}v_m \\ &\vdots \\ w_n &= a_{1n}v_1 + \dots + a_{mn}v_m. \end{aligned}$$

Se x_1, \dots, x_n sono numeri, allora

$$\begin{aligned} x_1w_1 + \dots + x_nw_n &= \\ &= (x_1a_{11} + \dots + x_na_{1n})v_1 + \dots + (x_1a_{m1} + \dots + x_na_{mn})v_m \end{aligned}$$

(per ottenere questa uguaglianza basta sommare in colonna i coefficienti di v_1, \dots, v_n). Il teorema precedentemente dimostrato permette di affermare che il sistema di equazioni

$$\begin{aligned} x_1a_{11} + \dots + x_na_{1n} &= 0 \\ &\vdots \\ x_1a_{m1} + \dots + x_na_{mn} &= 0 \end{aligned}$$

ha una soluzione non banale perché n supera m . Secondo quanto abbiamo precedentemente osservato, una siffatta soluzione (x_1, \dots, x_n) è tale che

$$x_1w_1 + \dots + x_nw_n = O,$$

come si voleva dimostrare.

Capitolo 4

Applicazioni lineari

Definiremo dapprima la nozione generale di applicazione di cui la nozione di funzione è un caso particolare. Tra tutte le applicazioni quelle lineari sono le più importanti. Un bel po' di matematica è dedicata alla riduzione di questioni concernenti applicazioni arbitrarie a questioni concernenti applicazioni lineari. Da una parte esse sono interessanti per sé stesse, e molte applicazioni sono lineari. D'altra parte, nel calcolo infinitesimale, le applicazioni lineari sono adoperate per approssimare le applicazioni differenziabili.

16. APPLICAZIONI

Siano S , S' due insiemi. Un'applicazione da S in S' è un modo di associare ad ogni elemento di S un elemento di S' . Invece di dire che F è un'applicazione da S in S' , spesso noi scriveremo $F: S \rightarrow S'$.

Una funzione è un'applicazione di tipo particolare, si tratta infatti dell'applicazione di un insieme nel corpo considerato K .

Alcuni termini adoperati per le funzioni sono estesi alle applicazioni. Per esempio, se $T: S \rightarrow S'$ è un'applicazione e u è un elemento di S , denotiamo con $T(u)$, o semplicemente Tu , l'elemento di S' associato a u da T . Noi chiamiamo $T(u)$ il valore dell'applicazione T in u , oppure l'immagine di u attraverso T . Il simbolo $T(u)$ si legge "T di u". L'insieme di tutti gli elementi $T(u)$, quando u varia nell'insieme S , è chiamato l'immagine di T . Se W è un sottoinsieme di S , l'insieme degli elementi $T(w)$, quando w varia nell'insieme W , viene chiamato l'immagine di W attraverso T e viene denotato con $T(W)$.

Esempio 1. Siano S e S' entrambi uguali a \mathbb{R} . Sia $f: \mathbb{R} \rightarrow \mathbb{R}$ la funzione $f(x) = x^2$ (cioè la funzione il cui valore nel numero x è x^2). Allora f è un'applicazione di \mathbb{R} in \mathbb{R} .

Esempio 2. Sia S l'insieme dei numeri non negativi e sia $S' = \mathbb{R}$. Sia $g: S \rightarrow S'$ la funzione tale che $g(x) = x^{\frac{1}{2}}$. Allora g è un'applicazione di S in \mathbb{R} .

Esempio 3. Sia S l'insieme delle funzioni aventi derivate di ogni ordine nell'intervallo $0 < t < 1$ e sia $S' = S$. Allora la derivazione $D = d/dt$ è un'applicazione di S in S . Infatti, la derivazione D associa ad ogni funzione f di S la funzione $df/dt = Df$. Seguendo la terminologia adottata, diremo che Df è il valore dell'applicazione D in f .

Esempio 4. Sia S l'insieme delle funzioni continue sull'intervallo $[0, 1]$ e sia S' l'insieme delle funzioni differenziabili sullo stesso intervallo. Definiamo l'applicazione $\mathcal{I}: S \rightarrow S'$ assegnando il suo valore per ogni funzione f appartenente a S . Definiamo $\mathcal{I}f$ (oppure $\mathcal{I}(f)$) come la funzione il cui valore in x è

$$\int_0^x f(t) dt.$$

Allora $\mathcal{I}(f)$ è una funzione differenziabile.

Esempio 5. Sia S l'insieme \mathbb{R}^3 , cioè l'insieme delle terne (ordinate) di numeri reali. Sia $A = (2, 3, -1)$. Sia $L: \mathbb{R}^3 \rightarrow \mathbb{R}$ l'applicazione che ad ogni vettore $X = (x, y, z)$ associa il numero $A \cdot X$. Allora $L(X) = A \cdot X$. Se $X = (1, 1, -1)$, allora il valore di L in X è 6.

Si vede quindi che definiamo un'applicazione, come abbiamo fatto per le funzioni, dandone i valori. Per esempio, nell'esempio 5, invece di enunciare la frase che descrive l'applicazione L , possiamo anche dire $L: \mathbb{R}^3 \rightarrow \mathbb{R}$ è l'applicazione $L(X) = A \cdot X$. Questo modo di procedere è più breve ma meno corretto, in pratica però non dà luogo a confusioni.

Esempio 6. Sia $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ l'applicazione definita da

$$F(x, y) = (2x, 2y).$$

Determinare l'immagine attraverso F dei punti appartenenti alla circonferenza $x^2 + y^2 = 1$.

Sia (x, y) un punto di questa circonferenza.

Poniamo $u = 2x$ e $v = 2y$; allora u e v soddisfano la relazione

$$(u/2)^2 + (v/2)^2 = 1$$

o, in altre parole,

$$\frac{u^2}{4} + \frac{v^2}{4} = 1.$$

Quindi (u, v) è un punto della circonferenza di raggio 2. Quindi l'immagine della circonferenza di raggio 1 attraverso l'applicazione F è un sottoinsieme della circonferenza di raggio 2. Viceversa, dato un punto (u, v) tale che

$$u^2 + v^2 = 4,$$

sia $x = u/2$ e $y = v/2$. Si vede allora che il punto (x, y) soddisfa l'equazione $x^2 + y^2 = 1$ e perciò appartiene alla circonferenza di raggio 1. Inoltre $F(x, y) = (u, v)$. Possiamo allora dire che ogni punto della circonferenza di raggio 2 è immagine di qualche punto della circonferenza di raggio 1. Concludiamo quindi affermando che l'immagine della circonferenza di raggio 1 attraverso F è precisamente la circonferenza di raggio 2.

Osservazione. Siano S , S' due insiemi. Per provare che essi coincidono, molto spesso si prova che S è un sottoinsieme di S' e che, viceversa, S' è un sottoinsieme di S . Questo è quanto abbiamo fatto nell'esempio 6.

Esempio 7. Sia S un insieme e sia V uno spazio vettoriale sul corpo K . Siano F , G applicazioni di S in V . Possiamo definire l'applicazione somma $F+G$ come quella il cui valore nell'elemento t di S è $F(t) + G(t)$. Definiamo anche il prodotto di un elemento c di K per F come l'applicazione il cui valore nell'elemento t di S è $cF(t)$. È facile verificare che le proprietà SV 1 ... SV 8 sono soddisfatte.

Esempio 8. Sia S un insieme e sia $F: S \rightarrow K^n$ un'applicazione. Per ogni elemento t di S , il valore di F in t è un vettore $F(t)$. Le coordinate di $F(t)$ dipendono da t , vi sono quindi n funzioni f_1, \dots, f_n di S in K tali che

$$F(t) = (f_1(t), \dots, f_n(t)).$$

Queste funzioni sono chiamate le *funzioni coordinate* di F . Per

esempio, se $K = \mathbb{R}$ e se S è un intervallo di numeri reali che noi denotiamo con J , allora un'applicazione

$$F: J \rightarrow \mathbb{R}^n$$

è anche chiamata una *curva* (in forma parametrica) nell' n -spazio.

Sia di nuovo S un insieme qualsiasi, siano $F, G: S \rightarrow K^n$ applicazioni di S in K^n . Siano f_1, \dots, f_n le funzioni coordinate di F , siano g_1, \dots, g_n le funzioni coordinate di G . Allora $G(t) = (g_1(t), \dots, g_n(t))$, per ogni $t \in S$. Inoltre

$$(F + G)(t) = F(t) + G(t) = (f_1(t) + g_1(t), \dots, f_n(t) + g_n(t)),$$

e per ogni $c \in K$,

$$(cF)(t) = cF(t) = (cf_1(t), \dots, cf_n(t)).$$

Vediamo, in particolare, che le funzioni coordinate di $F + G$ sono

$$f_1 + g_1, \dots, f_n + g_n.$$

Siano U, V, W insiemi. Siano $F: U \rightarrow V$ e $G: V \rightarrow W$ applicazioni. Possiamo allora considerare l'applicazione composta di U in W , denotata con $G \circ F$. Si tratta dell'applicazione definita ponendo

$$(G \circ F)(t) = G(F(t))$$

per ogni $t \in U$.

Esempio 9. Se f, g sono funzioni di \mathbb{R} in \mathbb{R} , allora $g \circ f$ è la funzione composta.

Esempio 10. Sia $F: \mathbb{R} \rightarrow \mathbb{R}^2$ l'applicazione definita ponendo

$$F(t) = (t, t^2),$$

e sia $G: \mathbb{R}^2 \rightarrow \mathbb{R}$ l'applicazione definita ponendo $G(x, y) = xy$. Allora, essendo $G(F(t)) = tt^2 = t^3$, si ha $(G \circ F)(t) = t^3$.

Esempio 11. Sia $X: \mathbb{R} \rightarrow \mathbb{R}^3$ l'applicazione definita ponendo

$$X(t) = (t, e^t, 2 \sin t).$$

Sia $F: \mathbb{R}^3 \rightarrow \mathbb{R}$ l'applicazione (che è poi una funzione) definita ponendo

$$F(x, y, z) = x^2y + z.$$

Allora $F(X(t))$ è uguale a $t^2 e^t + 2 \sin t$ e la possiamo anche scrivere $(F \circ X)(t)$.

Esempio 12. Sia U lo spazio vettoriale delle funzioni differenziabili di una variabile t , sia $U = V = W$. Sia $F: U \rightarrow U$ l'applicazione che ad ogni funzione f associa il suo quadrato (cioè $F(f) = f^2$) e sia $D: U \rightarrow U$ la derivazione. Allora, per ogni funzione differenziabile f , noi abbiamo

$$(D \circ F)(f) = 2ff',$$

dove f' denota la derivata di f .

Esempio 13. Sia $U = V = W$ lo spazio vettoriale delle funzioni aventi derivate di ogni ordine. Sia D la derivazione. Allora

$$(D \circ D)(f) = f''$$

è la derivata seconda. Analogamente, $(D \circ D \circ D)(f) = f''' = f^{(3)}$ è la derivata terza. In generale, possiamo scrivere $D^n f = f^{(n)}$. Quindi D^n è la iterazione di D fatta n volte.

La seguente proposizione esprime un'importante proprietà delle applicazioni.

Siano U, V, W, S insiemi e siano

$$F: U \rightarrow V, \quad G: V \rightarrow W, \quad H: W \rightarrow S$$

applicazioni. Allora

$$H \circ (G \circ F) = (H \circ G) \circ F.$$

Dimostrazione. La dimostrazione è molto semplice. Per definizione, per ogni elemento u di U , abbiamo le uguaglianze:

$$(H \circ (G \circ F))(u) = H((G \circ F)(u)) = H(G(F(u))).$$

D'altra parte

$$((H \circ G) \circ F)(u) = (H \circ G)(F(u)) = H(G(F(u))).$$

E questo, per definizione, significa

$$(H \circ G) \circ F = H \circ (G \circ F).$$

Concludiamo questo paragrafo introducendo un'utile notazione. Siano S, S' insiemi e sia $F: S \rightarrow S'$ un'applicazione. Per

denotare l'immagine di un elemento $x \in S$ attraverso F useremo una speciale freccia, cioè

$$x \mapsto F(x).$$

Possiamo quindi dire che F è l'applicazione tale che $x \mapsto F(x)$. Per esempio, invece di dire: "sia $F: \mathbb{R} \rightarrow \mathbb{R}$ tale che $F(x) = x^2$ " noi possiamo anche dire: "sia $F: \mathbb{R} \rightarrow \mathbb{R}$ l'applicazione $x \mapsto x^2$."

Esercizi

1. Con riferimento all'esempio 3, scrivere Df quando f è la funzione:

- a) $f(x) = \sin x$, b) $f(x) = e^x$, c) $f(x) = \log x$.

2. Con riferimento all'esempio 4, scrivere $\mathcal{J}(f)$ quando f è la funzione:

- a) $f(x) = e^x$, b) $f(x) = \frac{1}{1+x^2}$, c) $f(x) = \cos x$.

3. Con riferimento all'esempio 5, scrivere $L(X)$ quando X è il vettore:

- a) $(1, 2, -3)$, b) $(-1, 5, 0)$, c) $(2, 1, 1)$.

4. Sia $F: \mathbb{R} \rightarrow \mathbb{R}^2$ l'applicazione tale che $F(t) = (e^t, t)$. Che cosa sono $F(1)$, $F(0)$, $F(-1)$?

5. Sia $G: \mathbb{R} \rightarrow \mathbb{R}^2$ l'applicazione tale che $G(t) = (t, 2t)$. Sia F l'applicazione definita nell'esercizio 4. Che cosa sono $(F+G)(1)$, $(F+G)(2)$, $(F+G)(0)$?

6. Sia F l'applicazione definita nell'esercizio 4. Che cosa sono $(2F)(0)$, $(\pi F)(1)$?

7. Sia $A = (1, 1, -1, 3)$. Sia $F: \mathbb{R}^4 \rightarrow \mathbb{R}$ l'applicazione tale che per ogni vettore $X = (x_1, x_2, x_3, x_4)$ abbiamo $F(X) = X \cdot A + 2$. Qual è il valore di $F(X)$ quando: a) $X = (1, 1, 0, -1)$, b) $X = (2, 3, -1, 1)$?

(Negli esercizi da 8 a 12, si faccia riferimento all'esempio 6. In ciascun caso, per provare che l'immagine è uguale a un certo insieme S , si deve provare che l'immagine è contenuta in S e anche che ogni elemento di S è nell'immagine.)

8. Sia $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ l'applicazione definita ponendo $F(x, y) = (2x, 3y)$. Descrivere l'immagine dei punti appartenenti alla circonferenza $x^2 + y^2 = 1$.

9. Sia $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ l'applicazione definita ponendo $F(x, y) = (xy, y)$. Descrivere l'immagine attraverso F della retta $x = 2$.

10. Sia F l'applicazione definita ponendo $F(x, y) = (e^x \cos y, e^x \sin y)$. Descrivere l'immagine attraverso F della retta $x = 1$. Descrivere, più in generale, l'immagine attraverso F di una retta $x = c$, dove c è una costante.

11. Sia F l'applicazione definita ponendo $F(t, u) = (\cos t, \sin t, u)$. Descrivere geometricamente l'immagine del piano (t, u) attraverso F .

12. Sia F l'applicazione definita ponendo $F(x, y) = (x/3, y/4)$. Qual è l'immagine attraverso F dell'ellisse

$$\frac{x^2}{9} + \frac{y^2}{16} = 1?$$

17. APPLICAZIONI LINEARI

Siano V, V' spazi vettoriali sul corpo K . Un'applicazione lineare

$$F: V \rightarrow V'$$

è un'applicazione per cui valgono le seguenti proprietà:

AL 1. *Comunque si prendano gli elementi u, v in V abbiamo*

$$F(u + v) = F(u) + F(v).$$

AL 2. *Per ogni c in K e ogni v in V abbiamo*

$$F(cv) = cF(v).$$

Se desideriamo specificare il corpo K , possiamo dire che l'applicazione F è *K -lineare*. Poiché, di solito, considereremo un corpo fissato K , diremo semplicemente *lineare* omettendo il prefisso K .

Esempio 1. Sia V uno spazio vettoriale di dimensione finita su K e sia $\{v_1, \dots, v_n\}$ una base di V . Definiamo l'applicazione

$$F: V \rightarrow K^n$$

associando ad ogni elemento $v \in V$ il vettore X delle sue coordinate rispetto alla base fissata. Quindi, se

$$v = x_1 v_1 + \dots + x_n v_n,$$

con $x_i \in K$, noi poniamo

$$F(v) = (x_1, \dots, x_n).$$

Vogliamo ora far vedere che F è un'applicazione lineare. Se

$$w = y_1 v_1 + \dots + y_n v_n,$$

col vettore delle coordinate $Y = (y_1, \dots, y_n)$, allora

$$v + w = (x_1 + y_1)v_1 + \dots + (x_n + y_n)v_n,$$

da cui $F(v + w) = X + Y = F(v) + F(w)$. Se $c \in K$, allora

$$cv = cx_1v_1 + \dots + cx_nv_n,$$

e quindi $F(cv) = cX = cF(v)$. Questo prova che F è lineare.

Esempio 2. Sia $V = \mathbb{R}^3$ lo spazio vettoriale (su \mathbb{R}) dei vettori del 3-spazio. Sia $V' = \mathbb{R}^2$ lo spazio vettoriale dei vettori del 2-spazio. Possiamo definire un'applicazione

$$F: \mathbb{R}^3 \rightarrow \mathbb{R}^2$$

per proiezione, cioè $F(x, y, z) = (x, y)$. Lasciamo al lettore la verifica delle condizioni AL 1 e AL 2.

Più in generale, se r, n sono interi positivi, $r < n$, possiamo definire un'applicazione di proiezione

$$F: K^n \rightarrow K^r$$

ponendo

$$F(x_1, \dots, x_n) = (x_1, \dots, x_r).$$

Una facile verifica prova che questa applicazione è lineare.

Esempio 3. Sia $A = (1, 2, -1)$. Siano $V = \mathbb{R}^3$ e $V' = \mathbb{R}$. Possiamo definire un'applicazione $L = L_A: \mathbb{R}^3 \rightarrow \mathbb{R}$ ponendo $X \mapsto X \cdot A$, cioè

$$L(X) = X \cdot A$$

per ogni vettore X del 3-spazio. Il fatto che L sia un'applicazione lineare riassume due note proprietà del prodotto scalare; precisamente, comunque si prendano i vettori X, Y in \mathbb{R}^3 , abbiamo

$$(X + Y) \cdot A = X \cdot A + Y \cdot A,$$

$$(cX) \cdot A = c(X \cdot A).$$

Più in generale, sia K un corpo e A sia un vettore fissato in K^n . Allora ponendo $L_A(X) = X \cdot A$ per ogni $X \in K^n$, otteniamo un'applicazione lineare (cioè, un'applicazione K -lineare)

$$L_A: K^n \rightarrow K.$$

Possiamo anche generalizzare questa situazione al caso delle matrici. Sia A una matrice $m \times n$ nel corpo K . Ponendo

$$L_A(X) = AX$$

per ogni vettore colonna X in K^n , otteniamo un'applicazione lineare

$$L_A: K^n \rightarrow K^m.$$

Anche in questo caso la linearità segue dalle proprietà della moltiplicazione tra matrici. Se $A = (a_{ij})$ allora AX si scrive nel modo seguente:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Questo tipo di moltiplicazione sarà incontrato molte volte in seguito.

Esempio 4. Sia V uno spazio vettoriale. L'applicazione che associa ad ogni elemento u di V lo stesso elemento u è evidentemente un'applicazione lineare, chiamata *applicazione identica*. La denoteremo con id o semplicemente con I . Quindi $id(u) = u$.

Esempio 5. Siano V, V' spazi vettoriali sul corpo K . L'applicazione che associa l'elemento O di V' ad ogni elemento u di V è chiamata *applicazione nulla* ed è ovviamente lineare.

Esempio 6. Spazio delle applicazioni lineari. Siano V, V' due spazi vettoriali sul corpo K . Vogliamo considerare l'insieme di tutte le applicazioni lineari di V in V' , e lo denoteremo con $\mathcal{L}(V, V')$ o, più semplicemente, \mathcal{L} se il riferimento agli spazi V e V' è chiaro. Faremo vedere che è possibile definire un'addizione tra applicazioni lineari e una moltiplicazione per numeri in modo che \mathcal{L} divenga uno spazio vettoriale.

Siano $T: V \rightarrow V'$ ed $F: V \rightarrow V'$ due applicazioni lineari. Noi definiamo loro somma $T + F$ l'applicazione il cui valore in un elemento u di V è $T(u) + F(u)$. Quindi possiamo scrivere

$$(T + F)(u) = T(u) + F(u).$$

L'applicazione $T + F$ risulta lineare. Infatti è molto semplice verificare che le due condizioni caratteristiche affinché un'applica-

zione sia lineare sono soddisfatte. Comunque si prendano gli elementi u, v di V abbiamo

$$\begin{aligned}(T + F)(u + v) &= T(u + v) + F(u + v) \\&= T(u) + T(v) + F(u) + F(v) \\&= T(u) + F(u) + T(v) + F(v) \\&= (T + F)(u) + (T + F)(v).\end{aligned}$$

Inoltre, se $c \in K$, allora

$$\begin{aligned}(T + F)(cu) &= T(cu) + F(cu) \\&= cT(u) + cF(u) \\&= c[T(u) + F(u)] \\&= c[(T + F)(u)].\end{aligned}$$

Quindi $T + F$ è un'applicazione lineare.

Se $a \in K$ e se $T: V \rightarrow V'$ è un'applicazione lineare, definiamo un'applicazione aT da V in V' assegnando il suo valore nell'elemento u di V mediante l'uguaglianza $(aT)(u) = aT(u)$. Si verifica facilmente allora che aT è un'applicazione lineare. Lasciamo questa verifica per esercizio.

Abbiamo quindi definito le operazioni di addizione e moltiplicazione per numeri nel nostro insieme \mathcal{L} . Inoltre, se $T: V \rightarrow V'$ è un'applicazione lineare, cioè un elemento di \mathcal{L} , allora possiamo definire $-T$ come $(-1)T$, cioè come il prodotto del numero -1 per l'applicazione T . Osserviamo infine che abbiamo anche un'applicazione zero che è quella che ad ogni elemento di V associa l'elemento O di V' . Abbiamo allora che \mathcal{L} è uno spazio vettoriale. In altre parole, ripetiamo, l'insieme delle applicazioni lineari di V in V' è esso stesso uno spazio vettoriale. La verifica che le proprietà SV 1 ... SV 8 della definizione di spazio vettoriale sono soddisfatte è molto semplice ed è lasciata al lettore.

Esempio 7. Sia $V = V'$ lo spazio vettoriale delle funzioni reali di una variabile reale aventi derivate di ogni ordine. Sia D l'operazione di derivazione. Allora $D: V \rightarrow V'$ è un'applicazione lineare. Questa affermazione è nient'altro che un modo breve per riassumere le note proprietà della derivazione, precisamente:

$$D(f + g) = Df + Dg \quad \text{e} \quad D(cf) = cDf$$

per ogni coppia di funzioni differenziabili f, g e ogni costante c .

Se f è in V e I è l'applicazione identica, allora

$$(D + I)f = Df + f.$$

Quindi, se f è la funzione definita da $f(x) = e^x$ allora $(D + I)f$ è la funzione il cui valore in x è $e^x + e^x = 2e^x$.

Se $f(x) = \sin x$, allora $((D + I)f)(x) = \cos x + \sin x$.

Sia $T: V \rightarrow V'$ un'applicazione lineare; siano u, v, w elementi di V , allora

$$T(u + v + w) = T(u) + T(v) + T(w).$$

Questo può esser visto usando due volte successivamente la definizione di applicazione lineare, cioè

$$T(u + v + w) = T(u + v) + T(w) = T(u) + T(v) + T(w).$$

Analogamente, considerata una somma di più di tre elementi, una proprietà simile è soddisfatta. Per esempio, se u_1, \dots, u_n sono elementi di V allora

$$T(u_1 + \dots + u_n) = T(u_1) + \dots + T(u_n).$$

L'addizione nel secondo membro può esser fatta considerando gli addendi in un ordine qualunque. Una dimostrazione formale può facilmente conseguirsi col metodo di induzione, noi la omettiamo.

Se a_1, \dots, a_n sono numeri, allora

$$T(a_1 u_1 + \dots + a_n u_n) = a_1 T(u_1) + \dots + a_n T(u_n).$$

Facciamo la dimostrazione nel caso che si tratti di tre elementi

$$\begin{aligned} T(a_1 u + a_2 v + a_3 w) &= T(a_1 u) + T(a_2 v) + T(a_3 w) \\ &= a_1 T(u) + a_2 T(v) + a_3 T(w). \end{aligned}$$

Il teorema seguente mostra come un'applicazione lineare sia determinata dai suoi valori sugli elementi di una base.

TEOREMA 1 *Siano V e W spazi vettoriali sul corpo K . Sia $\{v_1, \dots, v_n\}$ una base di V e siano w_1, \dots, w_n elementi arbitrari di W . Allora esiste un'unica applicazione lineare $T: V \rightarrow W$ tale che $T(v_1) = w_1, \dots, T(v_n) = w_n$. Se x_1, \dots, x_n sono scalari, allora*

$$T(x_1 v_1 + \dots + x_n v_n) = x_1 w_1 + \dots + x_n w_n.$$

Dimostrazione. Dimostriamo intanto che un'applicazione lineare T soddisfacente le condizioni richieste esiste. Sia v un elemento di V e siano x_1, \dots, x_n gli unici numeri per cui $v = x_1 v_1 + \dots + x_n v_n$. Definiamo allora

$$T(v) = x_1 w_1 + \dots + x_n w_n.$$

L'applicazione T di V in W così definita risulta lineare. Infatti, se v' è un elemento di V e se $v' = y_1 v_1 + \dots + y_n v_n$, allora

$$v + v' = (x_1 + y_1)v_1 + \dots + (x_n + y_n)v_n.$$

Per definizione abbiamo

$$\begin{aligned} T(v + v') &= (x_1 + y_1)w_1 + \dots + (x_n + y_n)w_n \\ &= x_1 w_1 + y_1 w_1 + \dots + x_n w_n + y_n w_n \\ &= T(v) + T(v'). \end{aligned}$$

Se $c \in K$ allora $cv = cx_1 v_1 + \dots + cx_n v_n$ da cui

$$T(cv) = cx_1 w_1 + \dots + cx_n w_n = cT(v).$$

Abbiamo così provato che l'applicazione T è lineare e quindi che esiste un'applicazione lineare come è asserito nell'enunciato del teorema.

Una tale applicazione è poi unica giacché per ogni elemento $x_1 v_1 + \dots + x_n v_n$ di V , una qualunque applicazione lineare $F: V \rightarrow W$ tale che $F(v_i) = w_i$ ($i = 1, \dots, n$) deve anche essere tale che

$$\begin{aligned} F(x_1 v_1 + \dots + x_n v_n) &= x_1 F(v_1) + \dots + x_n F(v_n) \\ &= x_1 w_1 + \dots + x_n w_n. \end{aligned}$$

Così la dimostrazione è completa.

Esercizi

1. Dire quali, tra le seguenti applicazioni F , sono lineari:

- a) $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definita ponendo $F(x, y, z) = (x, z)$.
- b) $F: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ definita ponendo $F(X) = -X$.
- c) $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definita ponendo $F(X) = X + (0, -1, 0)$.
- d) $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definita ponendo $F(x, y) = (2x + y, y)$.
- e) $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definita ponendo $F(x, y) = (2x, y - x)$.
- f) $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definita ponendo $F(x, y) = (y, x)$.
- g) $F: \mathbb{R}^2 \rightarrow \mathbb{R}$ definita ponendo $F(x, y) = xy$.

2. Sia $T: V \rightarrow W$ un'applicazione lineare tra due spazi vettoriali. Dimostrare che $T(O) = O$.
3. Sia T come nell'esercizio 2. Siano u, v elementi di V e sia $Tu = w$. Se $Tv = O$, dimostrare che $T(u + v)$ è uguale a w .
4. Determinare tutti gli elementi z di V tali che $Tz = w$.
5. Sia $T: V \rightarrow W$ un'applicazione lineare. Sia v un elemento di V . Dimostrare che $T(-v) = -T(v)$.
6. Sia V uno spazio vettoriale su \mathbb{R} e siano $f: V \rightarrow \mathbb{R}$, $g: V \rightarrow \mathbb{R}$ due applicazioni lineari. Sia $F: V \rightarrow \mathbb{R}^2$ l'applicazione definita da
- $$F(v) = (f(v), g(v)).$$
- Dimostrare che F è lineare. Generalizzare.
7. Siano V, W due spazi vettoriali su K e sia $F: V \rightarrow W$ un'applicazione lineare. Sia U il sottoinsieme di V costituito da tutti gli elementi v tali che $F(v) = O$. Dimostrare che U è un sottospazio di V .
8. Quali tra le applicazioni definite negli esercizi 4, 7, 8, 9 del paragrafo 16 sono lineari?
9. Sia $F: \mathbb{R}^3 \rightarrow \mathbb{R}^4$ un'applicazione lineare. Sia P un punto di \mathbb{R}^2 e A sia un elemento non nullo di \mathbb{R}^3 . Descrivere l'immagine della retta $P + tA$ attraverso F . [Distinguere i due casi $F(A) = O$ e $F(A) \neq O$.]
- Sia V uno spazio vettoriale su \mathbb{R} e siano v_1, v_2 elementi di V linearmente indipendenti. L'insieme degli elementi di V che possono essere scritti nella forma $t_1v_1 + t_2v_2$ con numeri t_1, t_2 tali che $0 \leq t_1 \leq 1$ e $0 \leq t_2 \leq 1$ è chiamato il *parallelogramma* generato da v_1, v_2 .
10. Siano V e W spazi vettoriali su \mathbb{R} , sia $F: V \rightarrow W$ un'applicazione lineare. Siano v_1, v_2 elementi di V linearmente indipendenti e si assuma che anche $F(v_1), F(v_2)$ siano elementi linearmente indipendenti. Dimostrare che l'immagine attraverso F del parallelogramma generato da v_1 e v_2 è il parallelogramma generato da $F(v_1), F(v_2)$.
11. Sia F un'applicazione lineare di \mathbb{R}^2 in sé stesso tale che
- $$F(E_1) = (1, 1) \quad \text{e} \quad F(E_2) = (-1, 2).$$
- Sia S il quadrato i cui vertici sono nei punti $(0, 0), (1, 0), (1, 1), (0, 1)$. Dimostrare che l'immagine di questo quadrato attraverso F è un parallelogramma.
12. Siano A, B due vettori non nulli nel piano. Essi siano tali che non esiste nessuna costante non nulla c tale che $B = cA$. Sia T un'applicazione lineare del piano in sé stesso tale che $T(E_1) = A$ e $T(E_2) = B$. Descrivere l'immagine attraverso T del rettangolo i cui vertici sono nei punti $(0, 1), (3, 0), (0, 0), (3, 1)$.
13. Siano A, B due vettori non nulli del piano tali che non esiste alcuna costante non nulla c tale che $B = cA$. Descrivere geometricamente l'insieme dei punti $tA + uB$ per valori di t e u tali che $0 \leq t \leq 5$ e $0 \leq u \leq 2$.

14. Sia S un insieme contenuto in \mathbb{R}^n . Diremo che S è *convesso* se, comunque si prendano due punti P, Q in S , il segmento di retta congiungente P e Q è contenuto in S . [I punti di questo segmento di retta sono quelli che possono essere scritti nella forma $tP + (1-t)Q$, con $0 < t < 1$.]

Sia $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$ un'applicazione lineare. Dimostrare che l'immagine di un insieme convesso attraverso L è di nuovo un insieme convesso.

15. Sia $L: \mathbb{R}^n \rightarrow \mathbb{R}$ un'applicazione lineare. Sia S l'insieme di tutti i punti A di \mathbb{R}^n tali che $L(A) \geq 0$. Dimostrare che S è un insieme convesso.

16. Sia $L: \mathbb{R}^n \rightarrow \mathbb{R}$ un'applicazione lineare e sia c un numero. Dimostrare che l'insieme S costituito dai punti A di \mathbb{R}^n tali che $L(A) > c$ è convesso.

17. Siano S_1, S_2 due insiemi convessi di \mathbb{R}^n . Dimostrare che l'insieme S dei punti comuni a entrambi è convesso.

18. Sia $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$ un'applicazione lineare. Sia S un insieme convesso di \mathbb{R}^m . Sia S' l'insieme degli elementi $X \in \mathbb{R}^n$ tali che $L(X)$ sia in S . Dimostrare che anche l'insieme S' è convesso.

19. Sia A un vettore non nullo di \mathbb{R}^n e sia c un numero. Dimostrare che l'insieme dei punti X tali che $X \cdot A \geq c$ è convesso.

20. Siano A, B, C tre punti distinti di \mathbb{R}^n tali che i vettori $B - A$ e $C - A$ siano linearmente indipendenti. Dimostrare che questa condizione equivale al fatto che i tre punti non siano allineati.

21. Siano A, B, C tre punti di \mathbb{R}^n tali che $B - A$ e $C - A$ siano linearmente indipendenti. L'insieme dei punti del tipo

$$t_1A + t_2B + t_3C$$

dove t_i sono numeri non negativi per $i = 1, 2, 3$ e tali che

$$t_1 + t_2 + t_3 = 1$$

è chiamato il *triangolo* determinato dai punti A, B, C .

a) Dimostrare che ogni triangolo è convesso.

b) Dimostrare che ogni insieme convesso contenente i punti A, B, C deve contenere anche il triangolo da essi determinato.

c) Sia $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ un'applicazione lineare tale che $F(A), F(B), F(C)$ siano distinti e non allineati. Dimostrare che l'immagine attraverso F del triangolo determinato da A, B, C è il triangolo determinato da $F(A), F(B), F(C)$.

22. Siano V, W due spazi vettoriali su K e sia $F: V \rightarrow W$ un'applicazione lineare. Siano w_1, \dots, w_n elementi di W linearmente indipendenti; siano v_1, \dots, v_n elementi di V tali che $F(v_i) = w_i$ per $i = 1, \dots, n$. Dimostrare che anche gli elementi v_1, \dots, v_n sono linearmente indipendenti.

23. Sia V uno spazio vettoriale sul corpo K e sia $F: V \rightarrow K$ un'applicazione lineare. Sia W il sottoinsieme di V costituito da tutti gli elementi v tali che $F(v) = 0$. Si assuma che W sia distinto da V e sia v_0 un elemento

di V che non appartiene a W . Dimostrare che ogni elemento di V può essere scritto come somma $w + cv_0$, per un opportuno w in W e un opportuno scalare c .

24. Con riferimento all'esercizio 23, dimostrare che W è un sottospazio di V . Sia $\{v_1, \dots, v_n\}$ una base di W . Dimostrare che $\{v_0, v_1, \dots, v_n\}$ è una base di V .

18. NUCLEO E IMMAGINE DI UN'APPLICAZIONE LINEARE

Siano V , W spazi vettoriali sul corpo K , sia $F: V \rightarrow W$ un'applicazione lineare. Vogliamo dimostrare che le seguenti due proposizioni sono equivalenti:

- 1) Se v è un elemento di V tale che $F(v) = O$, allora $v = O$.
- 2) Se v , w sono elementi di V tali che $F(v) = F(w)$, allora $v = w$.

Per dimostrare la nostra affermazione assumiamo dapprima che per F sia vera la prima proposizione; siano poi v , w elementi di V tali che $F(v) = F(w)$. Allora

$$F(v - w) = F(v) - F(w) = O.$$

Per ipotesi, $v - w = O$ e quindi $v = w$.

Viceversa, supposta vera la seconda proposizione, se l'elemento v è tale che $F(v) = F(O) = O$, possiamo concludere senz'altro che $v = O$.

Sia $F: V \rightarrow W$ un'applicazione lineare. L'insieme degli elementi v di V tali che $F(v) = O$ si chiama il *nucleo* di F . La dimostrazione che il nucleo di un'applicazione lineare F è un sottospazio di V è lasciata per esercizio.

TEOREMA 2 *Sia $F: V \rightarrow W$ un'applicazione lineare il cui nucleo sia $\{O\}$. Se v_1, \dots, v_n sono elementi linearmente indipendenti di V , allora $F(v_1), \dots, F(v_n)$ sono elementi linearmente indipendenti di W .*

Dimostrazione. Siano x_1, \dots, x_n numeri tali che

$$x_1 F(v_1) + \dots + x_n F(v_n) = O.$$

Per la linearità dell'applicazione F , otteniamo

$$F(x_1 v_1 + \dots + x_n v_n) = O.$$

Quindi $x_1v_1 + \dots + x_nv_n = O$. Poiché v_1, \dots, v_n sono linearmente indipendenti, segue che $x_i = 0$ per $i = 1, \dots, n$. Questo dimostra il nostro teorema.

Esempio. Siano A, B due vettori linearmente indipendenti dell' n -spazio. Sia P un punto dell' n -spazio. L'insieme dei punti X tali che

$$X = P + tA + uB,$$

dove t, u assumono tutti i valori reali, è chiamato il *piano* passante per P e *parallelo* ad A e B .

Sia $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ un'applicazione lineare e si supponga che il nucleo di T sia $\{O\}$. Allora l'immagine di un piano attraverso T è di nuovo un piano. Infatti, considerato il piano costituito da tutti i punti $X = P + tA + uB$, l'immagine di ogni tale punto X è

$$T(P) + tT(A) + uT(B),$$

e poiché il nucleo di T è $\{O\}$, $T(A)$ e $T(B)$ sono linearmente indipendenti. Conseguentemente l'immagine del piano dato è il piano passante per $T(P)$, parallelo ai vettori $T(A)$ e $T(B)$.

L'equazione $X = P + tA + uB$ è talvolta chiamata *equazione parametrica* di un piano.

Sia $F: V \rightarrow W$ un'applicazione lineare. Si definisce *immagine* di F l'insieme degli elementi w di W per i quali esiste un elemento v di V tale che $F(v) = w$. L'immagine di F è un sottospazio di W . Per dimostrarlo si osservi dapprima che $F(O) = O$ e perciò O appartiene all'immagine di F . Si supponga poi che w_1, w_2 siano nell'immagine. Ciò significa che in V esistono due elementi v_1, v_2 tali che $F(v_1) = w_1$ e $F(v_2) = w_2$. Allora $F(v_1 + v_2) = F(v_1) + F(v_2) = w_1 + w_2$, provando così che $w_1 + w_2$ appartiene all'immagine. Infine, se c è uno scalare, allora

$$F(cv_1) = cF(v_1) = cw_1.$$

Quindi cw_1 è un elemento dell'immagine. Abbiamo così provato che l'immagine di F è un sottospazio di W .

Esercizi

1. Siano A, B due vettori di \mathbb{R}^2 che ne costituiscono una base. Sia $F: \mathbb{R}^2 \rightarrow \mathbb{R}^n$ un'applicazione lineare. Dimostrare che o gli elementi $F(A), F(B)$ sono linearmente indipendenti, o l'immagine di F ha dimensione 1, o l'immagine di F è $\{O\}$.

2. Trovare un'equazione parametrica del piano di \mathbb{R}^4 passante per i tre punti seguenti: $(1, 1, 0, -1)$, $(2, -1, 1, 3)$ e $(4, -2, 1, -1)$.

3. Sia $F: V \rightarrow W$ un'applicazione lineare e si supponga che il nucleo di F sia $\{O\}$. Dimostrare che se gli spazi V e W hanno la stessa dimensione n , allora l'immagine di F è l'intero spazio W .

4. Sia $F: V \rightarrow W$ un'applicazione lineare e si supponga che l'immagine di F sia l'intero spazio W . Dimostrare che se gli spazi V e W hanno la stessa dimensione n , allora il nucleo di F è $\{O\}$.

5. Sia $L: V \rightarrow W$ un'applicazione lineare. Sia w un elemento di W , sia v_0 un elemento di V tale che $L(v_0) = w$. Dimostrare allora che ogni soluzione dell'equazione $L(X) = w$ è del tipo $v_0 + u$ dove u è un elemento del nucleo di L .

6. Sia V lo spazio vettoriale delle funzioni che hanno derivate di ogni ordine, sia $D: V \rightarrow V$ la derivazione. Qual è il nucleo di D ?

7. Sia D^n la derivazione di secondo ordine (cioè l'iterazione di D fatta due volte). Qual è il nucleo di D^n ? In generale, qual è il nucleo dell'applicazione D^n (derivazione di ordine n)?

8. Sia V lo spazio considerato nell'esercizio 6. Scriviamone gli elementi come funzioni della variabile t e sia $D = d/dt$. Siano a_1, \dots, a_m numeri e g un elemento di V . Dire in che modo il problema della determinazione di una soluzione dell'equazione differenziale

$$a_m \frac{d^m f}{dt^m} + a_{m-1} \frac{d^{m-1} f}{dt^{m-1}} + \dots + a_0 f = g$$

può essere interpretato nel contesto della situazione astratta descritta nell'esercizio 5.

9. Siano V, D come nell'esercizio 6. Sia $L = D - I$, dove I è l'applicazione identica dello spazio V . Qual è il nucleo di L ?

10. Stessa domanda se $L = D - aI$, dove a è un numero.

19. DIMENSIONE DEL NUCLEO E DELL'IMMAGINE

Il teorema principale relativo all'immagine e al nucleo di un'applicazione lineare è il seguente.

TEOREMA 3 *Sia V uno spazio vettoriale di dimensione finita. Sia $L: V \rightarrow W$ un'applicazione lineare di V in un altro spazio W . Siano n la dimensione di V , q la dimensione del nucleo di L , s la dimensione dell'immagine di L . Allora $n = q + s$.*

Dimostrazione. Se l'immagine di L è il solo O , l'affermazione da provare è ovvia. Possiamo quindi assumere $s > 0$. Sia $\{w_1, \dots, w_s\}$ una base dell'immagine di L . Siano v_1, \dots, v_s elementi di V tali che $L(v_i) = w_i$ per $i = 1, \dots, s$. Se il nucleo di L non è $\{O\}$, sia $\{u_1, \dots, u_q\}$ una base del nucleo. Se il nucleo di L è $\{O\}$, si deve intendere omesso nel seguito ogni riferimento agli elementi $\{u_1, \dots, u_q\}$. Vogliamo mostrare che $\{v_1, \dots, v_s, u_1, \dots, u_q\}$ è una base di V . Questo sarà sufficiente a dimostrare il nostro teorema. Sia v un elemento di V . Esistono allora s scalari x_1, \dots, x_s tali che

$$L(v) = x_1 w_1 + \dots + x_s w_s,$$

perché $\{w_1, \dots, w_s\}$ è una base dell'immagine di L . Per la linearità dell'applicazione L , abbiamo

$$L(v) = L(x_1 v_1 + \dots + x_s v_s),$$

e, sottraendo il secondo membro dal primo, sempre per l'ipotesi di linearità, segue che

$$L(v - x_1 v_1 - \dots - x_s v_s) = O.$$

Quindi l'elemento $v - x_1 v_1 - \dots - x_s v_s$ appartiene al nucleo di L ed esistono quindi q scalari y_1, \dots, y_q in modo che

$$v - x_1 v_1 - \dots - x_s v_s = y_1 u_1 + \dots + y_q u_q.$$

Quindi abbiamo

$$v = x_1 v_1 + \dots + x_s v_s + y_1 u_1 + \dots + y_q u_q$$

che è una combinazione lineare di $v_1, \dots, v_s, u_1, \dots, u_q$. È così dimostrato che questi $s + q$ elementi di V generano V stesso.

Dimostriamo ora che essi sono linearmente indipendenti e quindi che costituiscono una base. Supponiamo che sussista una relazione lineare

$$x_1 v_1 + \dots + x_s v_s + y_1 u_1 + \dots + y_q u_q = O.$$

Applicando L a entrambi i membri di questa relazione e tenendo conto del fatto che $L(u_j) = O$ per $j = 1, \dots, q$ otteniamo

$$x_1 L(v_1) + \dots + x_s L(v_s) = O.$$

Ma, $L(v_1), \dots, L(v_s)$ non sono altro che w_1, \dots, w_s , elementi che

sono per ipotesi linearmente indipendenti. Deve perciò essere $x_i = 0$ per $i = 1, \dots, s$. Quindi

$$y_1 u_1 + \dots + y_q u_q = 0.$$

Ma u_1, \dots, u_q costituiscono una base del nucleo di L e, in particolare, sono elementi linearmente indipendenti. Perciò ogni y_j , per $j = 1, \dots, q$, deve essere nullo. Si conclude così la dimostrazione della nostra asserzione.

COROLLARIO *Sia $L: V \rightarrow W$ un'applicazione lineare e si assuma che*

$$\dim V = \dim W.$$

Si supponga poi che il nucleo di L sia $\{O\}$. Allora l'immagine di L è l'intero spazio W .

Dimostrazione. Per il teorema precedente l'immagine di L ha dimensione uguale a $\dim V = \dim W$. Poiché l'immagine è un sottospazio di W , essa deve coincidere con W stesso.

Introdurremo ora alcuni termini che sono molto utili nel trattare con applicazioni lineari, o anche con applicazioni qualsiasi.

Siano S, S' due insiemi, sia $F: S \rightarrow S'$ un'applicazione. Noi diremo che F è *iniettiva* se, comunque si prendano elementi distinti t, u in S , risulta $F(t) \neq F(u)$. Diremo poi che l'applicazione F è *surgettiva* se l'immagine di F è l'intero insieme S' , in altre parole se, per ogni elemento w di S' , esiste un elemento v in S tale che $F(v) = w$. Noi diciamo anche, in questo caso, che F applica S su S' (invece di dire in S').

Se $F: V \rightarrow W$ è un'applicazione lineare, ricordando l'osservazione fatta all'inizio del paragrafo 18 prima della definizione del nucleo di F , possiamo affermare che F è *iniettiva se, e solo se, il nucleo di F è $\{O\}$* .

Per esempio, sia $F: \mathbb{R} \rightarrow \mathbb{R}^2$ l'applicazione tale che $F(x) = (x, 0)$. Si vede subito che F è un'applicazione lineare iniettiva di \mathbb{R} in \mathbb{R}^2 . Questa applicazione però non è surgettiva, giacché la sua immagine consiste di tutti i vettori di \mathbb{R}^2 la cui seconda coordinata è uguale a 0, quindi l'immagine stessa non è uguale a \mathbb{R}^2 . Più in generale, se (α, β) è un vettore non nullo di \mathbb{R}^2 , con α, β appartenenti a \mathbb{R} , l'applicazione $G: \mathbb{R} \rightarrow \mathbb{R}^2$ tale che $G(x) = (x\alpha, x\beta)$ è un'applicazione lineare iniettiva di \mathbb{R} in \mathbb{R}^2 .

Osserviamo che la proiezione $F: \mathbb{R}^2 \rightarrow \mathbb{R}$ definita da

$$(x, y) \mapsto x$$

è surgettiva, ma non iniettiva. Infatti essa è un'applicazione lineare il cui nucleo è lo spazio di tutti i vettori di \mathbb{R}^2 la cui prima coordinata è 0.

L'applicazione $F: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ definita da $F(x, y, z) = (x, 0)$ è lineare, ma non è iniettiva né surgettiva.

Il corollario del teorema 3 può essere espresso dicendo che se $L: V \rightarrow W$ è un'applicazione lineare, se $\dim V = \dim W$, se L è iniettiva, allora L è anche surgettiva. Altri esempi di applicazioni lineari iniettive o surgettive si trovano negli esercizi.

Siano V, W spazi vettoriali sul corpo K . Sia $F: V \rightarrow W$ un'applicazione lineare. Diciamo che F è invertibile quando esiste un'applicazione lineare $G: W \rightarrow V$ tale che $F \circ G = id$ e $G \circ F = id$. (Si noti che quando scriviamo $F \circ G = id$, questo id denota l'applicazione identica di W , mentre quando scriviamo $G \circ F = id$, id denota l'applicazione identica di V .) Quindi diciamo che G è un'inversa di F e scriviamo $G = F^{-1}$, o $F = G^{-1}$. Osserviamo che un'inversa, se esiste, è unica. Infatti, siano $G_1, G_2: W \rightarrow V$ due inverse di F . Osserviamo allora che si può scrivere

$$G_1 = id \circ G_1 = (G_2 \circ F) \circ G_1 = G_2 \circ (F \circ G_1) = G_2,$$

ottenendo, come richiesto, l'uguaglianza $G_1 = G_2$.

L'inversa dell'applicazione F è solitamente denotata con F^{-1} .

TEOREMA 4 *Sia $F: V \rightarrow W$ un'applicazione lineare, si assuma F iniettiva e surgettiva. Allora, l'applicazione F è invertibile.*

Dimostrazione. Noi dobbiamo definire la sua inversa $G: W \rightarrow V$. Sia $w \in W$. Poiché F è surgettiva, esiste in V un elemento v tale che $F(v) = w$, e poiché F è iniettiva, questo elemento v è univocamente determinato. Allora, possiamo definire $G(w)$ come l'unico elemento $v \in V$ tale che $F(v) = w$. Resta da provare che G è lineare.

Siano w_1, w_2 in W e v_1, v_2 in V tali che $F(v_1) = w_1$, $F(v_2) = w_2$. Allora $F(v_1 + v_2) = w_1 + w_2$. Per definizione,

$$G(w_1 + w_2) = v_1 + v_2 = G(w_1) + G(w_2).$$

Inoltre, se c è un numero, allora $F(cv_1) = cF(v_1) = cw_1$. Quindi,

$$G(cw_1) = cv_1 = cG(w_1),$$

provando così il nostro teorema.

Un'applicazione lineare che sia iniettiva e surgettiva (e quindi invertibile) è chiamata un *isomorfismo*. Se, V , W sono spazi vettoriali su un corpo K e desideriamo specificare K nella terminologia, diciamo allora che l'applicazione lineare considerata è un K -*isomorfismo*. Dal corollario del teorema 3 e dal teorema 4 traiamo il:

COROLLARIO *Sia $F: V \rightarrow W$ un'applicazione lineare tra spazi vettoriali di uguale dimensione (finita). Se il nucleo di F è $\{O\}$, allora F è un isomorfismo.*

Esempio 1. Siano U , W sottospazi di uno spazio vettoriale V , e si supponga V somma diretta di U e W . L'applicazione

$$U \times W \rightarrow V$$

definita da

$$(u, w) \mapsto u + w$$

è lineare, come si verifica immediatamente, e ha per nucleo $\{O\}$. Questa applicazione dà quindi un isomorfismo del prodotto diretto $U \times W$ su V .

Esempio 2. Sia V uno spazio vettoriale di dimensione n su K e sia $\{v_1, \dots, v_n\}$ una sua base. Ad ogni elemento $v \in V$ associamo in K^n il vettore delle sue coordinate rispetto a questa base. In tal modo otteniamo un'applicazione lineare

$$F: V \rightarrow K^n,$$

e questa applicazione lineare è un isomorfismo. Notiamo esplicitamente che questo isomorfismo dipende dalla scelta della base.

Esempio 3. Sia $ay'' + by' + cy = 0$ un'equazione differenziale lineare con coefficienti complessi costanti a , b , c , con $a \neq 0$. Le sue soluzioni allora costituiscono uno spazio vettoriale V sul corpo complesso. Se f è una soluzione, associamo ad essa l'elemento di C^2 di coordinate $(f(0), f'(0))$. In tal modo otteniamo un'applicazione di V in C^2 , cioè $f \mapsto (f(0), f'(0))$. Si vede facilmente che questa applicazione è lineare. Il teorema di unicità della teoria delle equazioni differenziali lineari dice che il suo

nucleo è 0, mentre il teorema di esistenza dice che, per ogni vettore complesso $(\alpha, \beta) \in \mathbb{C}^2$, esiste una soluzione f tale che $f(0) = \alpha$ e $f'(0) = \beta$, dice cioè che la nostra applicazione è surgettiva. Abbiamo così ottenuto un isomorfismo tra V e \mathbb{C}^2 .

Sia $F: V \rightarrow V'$ un isomorfismo tra spazi vettoriali. Possiamo allora dire, parlando alla buona, che V e V' "sembrano gli stessi" dal punto di vista delle proprietà algebriche. Per esempio, *V ha dimensione finita se, e solo se, V' ha dimensione finita, in tal caso hanno la stessa dimensione.* Per dimostrare ciò, sia $\{v'_1, \dots, v'_n\}$ una base di V' . Siano v_1, \dots, v_n gli elementi di V tali che $F(v_i) = v'_i$ ($i = 1, \dots, n$). Questi elementi esistono perché F è surgettiva. Siano x_i scalari tali che

$$x_1 v_1 + \dots + x_n v_n = O,$$

Applicando F , troviamo:

$$O = x_1 F(v_1) + \dots + x_n F(v_n) = x_1 v'_1 + \dots + x_n v'_n,$$

ed essendo $\{v'_1, \dots, v'_n\}$ una base di V' , segue che, per ogni i , $x_i = 0$, pertanto gli elementi v_1, \dots, v_n sono linearmente indipendenti. Sia $G: V' \rightarrow V$ l'inversa di F . Allora, $G(v'_i) = v_i$. Ogni elemento $v \in V$ può essere scritto $G(v')$ per un opportuno $v' \in V'$ e quindi esistono gli scalari y_1, \dots, y_n tali che

$$v = G(y_1 v'_1 + \dots + y_n v'_n) = y_1 G(v'_1) + \dots + y_n G(v'_n).$$

Questo prova che gli elementi v_1, \dots, v_n generano V e che quindi, come richiesto, costituiscono una base di V .

Consideriamo di nuovo un isomorfismo $F: V \rightarrow V'$. Sia W un sottospazio di V . Allora F induce un'applicazione lineare di W nella sua immagine, essa sarà denotata con F_W ; cioè

$$F_W: W \rightarrow F(W).$$

L'applicazione F_W è ovviamente surgettiva (per definizione), e poiché il nucleo di F (in V) è costituito dal solo elemento O , segue che il nucleo di F_W (in W) è costituito dal solo elemento O . Quindi, F_W è un isomorfismo di W sulla sua immagine.

Esercizi

1. Siano V, W spazi vettoriali sul corpo K e si supponga $\dim V = \dim W$. Sia $F: V \rightarrow W$ un'applicazione lineare. Dimostrare che se F è surgettiva, F è un isomorfismo.

2. Sia V uno spazio vettoriale su \mathbb{R} . Sia $v \in V$ un elemento non nullo di V . Dimostrare che l'applicazione

$$x \mapsto xv$$

è lineare e iniettiva da \mathbb{R} in V e che ogni applicazione lineare iniettiva di \mathbb{R} in V è di questo tipo.

3. Siano V, W spazi vettoriali su un corpo K e si supponga che la dimensione di V sia minore di quella di W . Dimostrare che nessuna applicazione lineare $F: V \rightarrow W$ può essere surgettiva.

4. Sia $F: V \rightarrow W$ un'applicazione lineare invertibile. Dimostrare che F è un isomorfismo.

5. Siano V, W spazi vettoriali su un corpo K e si supponga che $\dim V$ sia minore di $\dim W$. Dimostrare che nessuna applicazione lineare $G: W \rightarrow V$ può essere iniettiva.

6. Sia $f: \mathbb{R} \rightarrow \mathbb{R}$ l'applicazione che a x associa x^2 . L'applicazione f è iniettiva? È surgettiva?

7. Sia $f: \mathbb{C} \rightarrow \mathbb{C}$ l'applicazione che a z associa $z+1$. L'applicazione f è iniettiva? È surgettiva?

8. Sia S l'insieme dei numeri reali positivi. Sia $f: S \rightarrow S$ l'applicazione che a x associa x^2 . L'applicazione f è iniettiva? È surgettiva?

9. Sia $f: \mathbb{R} \rightarrow \mathbb{R}$ l'applicazione definita da $f(x) = x^3$. Determinare se f è surgettiva e/o se è iniettiva.

10. Sia V la somma diretta di due sottospazi U, W . Dimostrare in tutti i dettagli che, come asserito in un esempio del testo, il prodotto diretto $U \times W$ è isomorfo a V .

11. Sia V uno spazio vettoriale. Dimostrare che l'applicazione $F: V \rightarrow V \times V$ definita da $f(v) = (v, v)$ è lineare e iniettiva.

12. Siano U, W sottospazi dello spazio vettoriale V . Dimostrare che l'applicazione del prodotto diretto $U \times W$ in V definita da

$$(u, w) \mapsto u - w$$

è lineare. Dimostrare che l'immagine di questa applicazione lineare è $U + W$. Dimostrare poi che il suo nucleo è isomorfo al sottospazio $U \cap W$.

13. Con riferimento alle notazioni introdotte nell'esercizio precedente, dimostrare che se U, W hanno dimensione finita, allora si ha

$$\dim U + \dim W = \dim(U \cap W) + \dim(U + W).$$

14. Sia V la somma diretta di due sottospazi U, W . Si definisca un'applicazione $P_1: V \rightarrow V$ come segue: se $v \in V$ è espresso univocamente come somma $u + w$ con $u \in U$ e $w \in W$, allora $P_1(v) = u$. Si definisca poi l'applicazione $P_2: V \rightarrow V$ come $P_2(v) = w$.

cazione $P_2: V \rightarrow V$ ponendo $P_2(v) = w$. Dimostrare che P_1 e P_2 sono applicazioni lineari e che $P_1 + P_2 = I$, applicazione identica. Le applicazioni P_1 , P_2 sono chiamate rispettivamente le *proiezioni* sul primo e sul secondo fattore.

20. COMPOSIZIONE DI APPLICAZIONI LINEARI

In precedenza abbiamo già dato la definizione di composizione di applicazioni qualsiasi. Nel caso particolare di applicazioni lineari abbiamo il seguente:

TEOREMA 5 *Siano U , V , W spazi vettoriali su un corpo K . Siano*

$$F: U \rightarrow V \quad e \quad G: V \rightarrow W$$

applicazioni lineari. Allora l'applicazione composta $G \circ F$ è anch'essa lineare.

Dimostrazione. L'affermazione si prova facilmente. Siano u , v elementi di U . Poiché F è lineare, abbiamo $F(u+v) = F(u) + F(v)$. Quindi

$$(G \circ F)(u+v) = G(F(u+v)) = G(F(u) + F(v)).$$

E poiché G è lineare, otteniamo

$$G(F(u) + F(v)) = G(F(u)) + G(F(v)).$$

Da cui

$$(G \circ F)(u+v) = (G \circ F)(u) + (G \circ F)(v).$$

Sia poi c uno scalare. Allora

$$\begin{aligned} (G \circ F)(cu) &= G(F(cu)) \\ &= G(cF(u)) \quad (\text{perché } F \text{ è lineare}) \\ &= cG(F(u)) \quad (\text{perché } G \text{ è lineare}). \end{aligned}$$

Questo prova che $G \circ F$ è un'applicazione lineare.

Il teorema seguente dice che alcune delle regole di aritmetica concernenti somme e prodotti di numeri continuano a valere per somme e composizioni di applicazioni lineari.

TEOREMA 6 *Siano U , V , W spazi vettoriali sul corpo K . Sia*

$$F: U \rightarrow V$$

un'applicazione lineare e siano G, H due applicazioni lineari di V in W . Allora

$$(G + H) \circ F = G \circ F + H \circ F.$$

Se $c \in K$, allora

$$(cG) \circ F = c(G \circ F).$$

Se $T: U \rightarrow V$ è un'applicazione lineare di U in V , allora

$$G \circ (F + T) = G \circ F + G \circ T.$$

Le dimostrazioni sono tutte molto semplici. Noi dimostreremo soltanto la prima asserzione, lasciando le altre come esercizio.

Sia u un elemento di U . Abbiamo allora

$$\begin{aligned} ((G + H) \circ F)(u) &= (G + H)(F(u)) = G(F(u)) + H(F(u)) \\ &= (G \circ F)(u) + (H \circ F)(u). \end{aligned}$$

Da cui segue, per definizione,

$$(G + H) \circ F = G \circ F + H \circ F.$$

Consideriamo ora il caso $U = V = W$. Siano

$$F: U \rightarrow U \quad \text{e} \quad G: U \rightarrow U$$

due applicazioni lineari. Allora esistono le applicazioni $F \circ G$ e $G \circ F$. Non sempre è vero che queste due applicazioni composte coincidono. Per esempio, sia $U = \mathbb{R}^3$. Sia F l'applicazione lineare definita da

$$F(x, y, z) = (x, y, 0)$$

e sia G l'applicazione lineare definita da

$$G(x, y, z) = (x, z, 0).$$

Allora $(G \circ F)(x, y, z) = (x, 0, 0)$, mentre $(F \circ G)(x, y, z) = (x, z, 0)$.

Sia $F: V \rightarrow V$ un'applicazione lineare di uno spazio vettoriale in sé stesso. Una tale applicazione è chiamata talvolta *operatore*. Noi possiamo considerare l'applicazione composta $F \circ F$ che, di nuovo, è un'applicazione lineare di V in sé stesso. Analogamente, possiamo considerare la composizione di F con sé stessa n volte,

per ogni n intero positivo:

$$F \circ F \circ \dots \circ F.$$

Denotiamo questa applicazione composta con F^n . Se $n = 0$, definiamo $F^0 = I$ (applicazione identica). Allora la regola

$$F^{r+s} = F^r \circ F^s$$

vale per ogni coppia di interi non negativi r, s .

Per semplicità di notazione spesso omettiamo il piccolo cerchietto tra le applicazioni lineari scrivendo soltanto FG invece di $F \circ G$.

Esercizi

1. Sia V uno spazio vettoriale. Sia $P: V \rightarrow V$ un'applicazione lineare tale che $P \circ P = P$. Siano U l'immagine di P e W il nucleo di P . Dimostrare che V è la somma diretta di U e W . [Per dimostrare che V è una somma, si suggerisce di scrivere un elemento v di V nella forma $v - Pv + Pv$.]
2. Sia V uno spazio vettoriale e siano P_1, P_2 applicazioni lineari di V in sé stesso. Si supponga che le seguenti condizioni siano soddisfatte:
 - a) $P_1 + P_2 = I$ (applicazione identica).
 - b) $P_1 \circ P_2 = O$ e $P_2 \circ P_1 = O$.
 - c) $P_1 \circ P_1 = P_1$ e $P_2 \circ P_2 = P_2$.

Dimostrare che V è somma diretta delle immagini di P_1 e P_2 .

3. Con riferimento alle notazioni introdotte nell'esercizio 2, dimostrare che l'immagine di P_1 è uguale al nucleo di P_2 .

[Si suggerisce di dimostrare le due proposizioni seguenti:

L'immagine di P_1 è contenuta nel nucleo di P_2 .

Il nucleo di P_2 è contenuto nell'immagine di P_1 .]

4. Siano $F: V \rightarrow W$ e $G: W \rightarrow U$ isomorfismi tra spazi vettoriali su K . Dimostrare che $G \circ F$ è invertibile e che

$$(G \circ F)^{-1} = F^{-1} \circ G^{-1}.$$

5. Siano $F: V \rightarrow W$ e $G: W \rightarrow U$ isomorfismi tra spazi vettoriali su K . Dimostrare che anche l'applicazione composta $G \circ F: V \rightarrow U$ è un isomorfismo.

6. Siano V, W due spazi vettoriali su K di uguale dimensione finita n . Dimostrare che V e W sono isomorfi.

7. Sia A un'applicazione lineare di uno spazio vettoriale in sé stesso. Si supponga che

$$A^2 - A + I = O$$

(dove I è l'applicazione identica). Dimostrare che l'applicazione inversa A^{-1} esiste ed è uguale a $I - A$. Generalizzare (vedi l'es. 18, § 14, cap. 3).

8. Siano A, B applicazioni lineari di uno spazio vettoriale in sé stesso. Se il nucleo di A è $\{O\}$ e il nucleo di B è $\{O\}$, dimostrare che anche il nucleo di AB è $\{O\}$.

9. Siano A, B applicazioni lineari di uno spazio vettoriale in sé stesso. Si assuma che AB coincida con BA . Dimostrare che

$$(A + B)^2 = A^2 + 2AB + B^2$$

e

$$(A + B)(A - B) = A^2 - B^2.$$

10. Più in generale, siano $A: V \rightarrow W$ e $B: W \rightarrow U$ applicazioni lineari. Si supponga che il nucleo di A sia $\{O\}$ e che il nucleo di B sia $\{O\}$. Dimostrare che anche il nucleo dell'applicazione BA è $\{O\}$.

11. Con riferimento alle notazioni introdotte nell'esercizio 10, si assuma che le applicazioni A, B siano surgettive. Dimostrare allora che anche l'applicazione BA è surgettiva.

Capitolo 5

Applicazioni lineari e matrici

Una volta fissate le basi, è possibile rappresentare ogni applicazione lineare con una matrice e, viceversa, ad ogni matrice corrisponde un'applicazione lineare. Vedremo che l'addizione e la moltiplicazione di applicazioni lineari corrispondono all'addizione e moltiplicazione di matrici.

21. APPLICAZIONE LINEARE ASSOCIATA A UNA MATRICE

Siano V e W spazi vettoriali su K , siano $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_m\}$ loro rispettive basi. Sia

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ & \ddots & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

una matrice $m \times n$ in K . Vogliamo definire un'applicazione lineare

$$L_A: V \rightarrow W,$$

dipendente da A e dalle basi scelte in V e W .

Siano A_1, \dots, A_m i vettori riga della matrice A . Sia cioè

$$A_1 = (a_{11}, \dots, a_{1n}),$$

...

$$A_m = (a_{m1}, \dots, a_{mn}).$$

Sia v un elemento di V e sia X il vettore delle sue coordinate rispetto alla base fissata. Poiché è importante distinguere tra vettore riga e vettore colonna, d'ora in poi converremo di considerare il vettore delle coordinate come un *vettore colonna*.

Possiamo allora scrivere

$${}^t X = (x_1, \dots, x_n).$$

Definiamo ora l'applicazione $L_A: V \rightarrow W$ (dipendente dalle basi scelte) ponendo

$$L_A(v) = (A_1 \cdot X)w_1 + \dots + (A_m \cdot X)w_m.$$

In altre parole, possiamo dire che la i -esima coordinata di $L_A(v)$ è $A_i \cdot X$. Quindi L_A è un'applicazione di V in W .

TEOREMA 1 *L'applicazione L_A è lineare.*

Dimostrazione. Siano u, v elementi di V . Esistono allora, e sono unici, gli scalari y_1, \dots, y_n e x_1, \dots, x_n in modo che possiamo scrivere

$$u = y_1v_1 + \dots + y_nv_n, \quad v = x_1v_1 + \dots + x_nv_n.$$

Quindi

$$u + v = (x_1 + y_1)v_1 + \dots + (x_n + y_n)v_n.$$

Abbiamo allora:

$$L_A(u + v) = (A_1 \cdot (X + Y))w_1 + \dots + (A_m \cdot (X + Y))w_m.$$

In altre parole, la i -esima coordinata di $L_A(u + v)$ è $A_i \cdot (X + Y)$. Ma sappiamo che $A_i \cdot (X + Y) = A_i \cdot X + A_i \cdot Y$, perciò

$$\begin{aligned} L_A(u + v) &= (A_1 \cdot X)w_1 + \dots + (A_m \cdot X)w_m + \\ &\quad + (A_1 \cdot Y)w_1 + \dots + (A_m \cdot Y)w_m \\ &= L_A(u) + L_A(v). \end{aligned}$$

Sia c uno scalare, allora

$$\begin{aligned} L_A(cu) &= (A_1 \cdot cX)w_1 + \dots + (A_m \cdot cX)w_m \\ &= c(A_1 \cdot X)w_1 + \dots + c(A_m \cdot X)w_m \\ &= cL_A(u). \end{aligned}$$

E questo prova il nostro teorema.

Esempio 1. Sia $V = K^n$ e $W = K^m$; se le basi sono costituite dagli usuali vettori unità, se A è una matrice $m \times n$ in K , vediamo

che

$$L_A(X) = AX$$

dove AX indica la moltiplicazione tra matrici:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Se $L_A(X) = Y$, allora

$$\begin{aligned} y_1 &= a_{11}x_1 + \dots + a_{1n}x_n, \\ \vdots &\quad \vdots \quad \dots \quad \vdots \\ y_m &= a_{m1}x_1 + \dots + a_{mn}x_n. \end{aligned}$$

Questo significa che rispetto alla base standard di K^n , e considerando X come un *vettore colonna*, abbiamo

$$Y = AX,$$

l'applicazione lineare essendo ora definita dalla moltiplicazione tra matrici.

Siano A e B matrici $m \times n$. È naturale chiedersi quando esse danno luogo alla stessa applicazione lineare, cioè quando $L_A = L_B$. Il teorema seguente risponde a questa domanda.

TEOREMA 2 *Siano A, B matrici $m \times n$. Siano V, W spazi vettoriali e, come in precedenza, fissiamo in essi due basi. Se $L_A = L_B$, allora $A = B$. In altre parole, se le matrici danno luogo alla stessa applicazione lineare, esse coincidono.*

Dimostrazione. Siano $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_m\}$ basi rispettive di V e W . Si assuma che L_A coincida con L_B . Siano A_1, \dots, A_m i vettori riga di A e siano B_1, \dots, B_m i vettori riga di B . Per ogni n -upla $X = (x_1, \dots, x_n)$, se $v = x_1v_1 + \dots + x_nv_n$, le espressioni di $L_A(v)$ e $L_B(v)$ come combinazioni lineari di w_1, \dots, w_m sono uguali. Quindi sono uguali anche le loro coordinate rispetto alla base $\{w_1, \dots, w_m\}$. Quindi

$$A_i \cdot X = B_i \cdot X$$

per ogni $i = 1, \dots, m$. Perciò $(A_i - B_i) \cdot X = 0$ per ogni i e ogni X . Allora $A_i - B_i = O$ e $A_i = B_i$ per ogni i . Da cui, infine, $A = B$.

Un buon esercizio consiste nel provare il teorema seguente.

TEOREMA 3 Siano V, W spazi vettoriali e siano $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_m\}$ loro rispettive basi. Siano A, B due matrici $m \times n$ e siano L_A, L_B le applicazioni lineari di V in W ad esse associate relativamente alle basi scelte. Allora, per ogni elemento v di V e ogni numero c , si ha

$$L_{A+B}(v) = L_A(v) + L_B(v),$$

$$L_{cA}(v) = cL_A(v).$$

Potremmo omettere l'indicazione di v nell'enunciato del teorema 3 e semplicemente scrivere

$$L_{A+B} = L_A + L_B,$$

$$L_{cA} = cL_A.$$

Possiamo dire che l'associazione

$$A \mapsto L_A$$

definisce un'applicazione线are dello spazio delle matrici nello spazio $\mathcal{L}(V, W)$.

22. MATRICE ASSOCIATA A UN'APPlicazione LINEARE

Siano V, W spazi vettoriali. Sia $F: V \rightarrow W$ un'applicazione lineare di V in W . Vedremo ora in che modo si può associare una matrice a F . Una tale matrice dipenderà dalla scelta delle basi in V e W .

Siano $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_m\}$ rispettivamente basi di V e W . Ognuno degli elementi $F(v_1), \dots, F(v_n)$ appartiene a W . Ognuno di essi, quindi, può essere scritto come combinazione lineare di w_1, \dots, w_m . Cioè

$$F(v_1) = a_{11}w_1 + \dots + a_{m1}w_m,$$

...

$$F(v_n) = a_{1n}w_1 + \dots + a_{mn}w_m.$$

I numeri a_{ij} così disposti

$$\begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}$$

formano una matrice. La *trasposta* di questa matrice sarà chiamata la *matrice associata* all'applicazione lineare F (relativa alla nostra scelta delle basi).

La trasposta della matrice sopra scritta è quindi

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

La ragione della considerazione della matrice trasposta sarà chiara tra breve.

Esempio 1. Sia F l'applicazione lineare tale che

$$F(v_1) = 3w_1 - w_2 + 17w_3,$$

$$F(v_2) = w_1 + w_2 - w_3,$$

avendo assunto $\dim V = 2$ e $\dim W = 3$. Allora la matrice associata a F è

$$\begin{pmatrix} 3 & 1 \\ -1 & 1 \\ 17 & -1 \end{pmatrix}$$

uguale alla trasposta della matrice

$$\begin{pmatrix} 3 & -1 & 17 \\ 1 & 1 & -1 \end{pmatrix}.$$

Consideriamo di nuovo il caso generale.

Sia $v = x_1v_1 + \dots + x_nv_n$ un elemento di V . Per la linearità dell'applicazione F , abbiamo

$$F(v) = x_1F(v_1) + \dots + x_nF(v_n).$$

Tenendo conto dell'espressione di $F(v_1), \dots, F(v_n)$ mediante w_1, \dots, w_m scritta sopra, abbiamo che

$F(v) = x_1(a_{11}w_1 + \dots + a_{m1}w_m) + \dots + x_n(a_{1n}w_1 + \dots + a_{mn}w_m)$,
e, raccogliendo i coefficienti dei termini w_1, \dots, w_m , possiamo riscrivere questa espressione nella forma

$$(a_{11}x_1 + \dots + a_{1n}x_1)w_1 + \dots + (a_{m1}x_1 + \dots + a_{mn}x_n)w_m.$$

E questo è precisamente uguale a $L_A(v)$. Quindi $F = L_A$.

In altre parole, se v è un elemento di V , se X è il vettore (verticale) delle sue coordinate relative alla base $\{v_1, \dots, v_n\}$, se A è la matrice associata a F relativa alle basi scelte, allora il vettore delle coordinate di $F(v)$ relative alla base $\{w_1, \dots, w_m\}$ è AX .

Come conseguenza del teorema 2 del paragrafo precedente, vediamo che la nostra matrice A è l'unica matrice tale che $F = L_A$. L'unicità della matrice A potrebbe essere dimostrata direttamente tenendo conto del fatto che i valori di F sono determinati dai suoi valori sugli elementi di una base.

Possiamo ora precisare che l'associazione $A \mapsto L_A$ fornisce un *isomorfismo* tra lo spazio delle matrici $m \times n$ e lo spazio delle applicazioni lineari $\mathcal{L}(V, W)$.

La ragione per cui abbiamo considerato la matrice trasposta nella definizione di matrice associata a un'applicazione lineare sta nel fatto che, rappresentando i vettori con le coordinate, il valore dell'applicazione lineare è rappresentato dall'ordinario prodotto tra la matrice associata A e il vettore colonna X . Ciò non sarebbe stato possibile, in questo modo, se non avessimo considerato la matrice trasposta. Per questa ragione, d'ora in poi, considereremo *sempre* il vettore delle coordinate di un elemento $v \in V$ come un *vettore colonna*.

Esempio 2. Sia $F: K^3 \rightarrow K^2$ la proiezione, in altre parole consideriamo l'applicazione definita da $F(x_1, x_2, x_3) = (x_1, x_2)$. Allora la matrice associata a F relativa alle usuali basi è

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

Esempio 3. Sia $F: K^n \rightarrow K^n$ l'applicazione identica. Allora la matrice associata a F relativa alle usuali basi è la matrice

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

le cui componenti sono uguali a 1 sulla diagonale e uguali a 0 altrove.

Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base dello spazio vettoriale V , e sia $\mathcal{B}' = \{w_1, \dots, w_m\}$ una base dello spazio vettoriale W . Denoteremo

con

$$M_{\mathcal{B}}^{\mathcal{B}'}(F)$$

la matrice associata all'applicazione lineare F di V in W , relativa alle basi \mathcal{B} e \mathcal{B}' . Se queste basi sono fissate in tutta la discussione possiamo anche scrivere, più semplicemente, $M(F)$.

Attenzione. Assumiamo $V = W$, ma consideriamo due basi distinte di V , \mathcal{B} e \mathcal{B}' . Allora la matrice associata all'applicazione identica di V in sé stesso, relativa a queste due diverse basi, non è la matrice unità!

Esempio 4. Rotazioni. Vogliamo considerare ora due diverse situazioni. Dapprima sceglieremo due diversi sistemi di coordinate differenti per una rotazione. L'applicazione identica, allora, avrà come matrice associata una matrice che non è la matrice unità. Successivamente, discuteremo la matrice associata a una rotazione rispetto a una base fissata.

Caso 1. Partiamo nel piano col nostro solito sistema di coordinate. Siano $E_1 = (1, 0)$ ed $E_2 = (0, 1)$ i vettori unità. Consideriamo poi un altro sistema di coordinate ottenuto da quello ora fissato con una rotazione in senso antiorario di ampiezza θ .

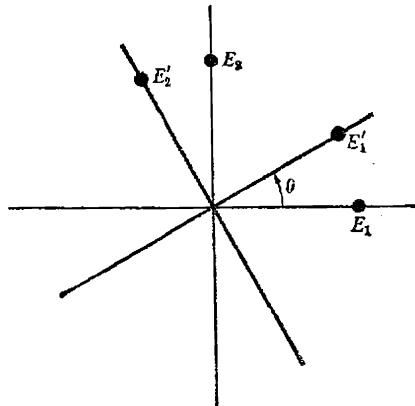


Figura 22.1

Allora (vedi fig. 22.1) i vettori unità sono trasformati in due nuovi vettori unità E'_1 ed E'_2 .

Dalla figura vediamo che

$$\begin{aligned}E'_1 &= (\cos \theta)E_1 + (\sin \theta)E_2, \\E'_2 &= (-\sin \theta)E_1 + (\cos \theta)E_2.\end{aligned}$$

E se moltiplichiamo la prima equazione per $\cos \theta$ e la seconda per $-\sin \theta$, addizionando, otteniamo:

$$E_1 = (\cos \theta)E'_1 - (\sin \theta)E'_2.$$

Analogamente otteniamo:

$$E_2 = (\sin \theta)E'_1 + (\cos \theta)E'_2.$$

Sia $id: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ l'applicazione identica. Siano poi $\mathcal{B} = \{E_1, E_2\}$ e $\mathcal{B}' = \{E'_1, E'_2\}$. Allora

$$\begin{aligned}id(E_1) &= E_1 = (\cos \theta)E'_1 + (-\sin \theta)E'_2, \\id(E_2) &= E_2 = (\sin \theta)E'_1 + (\cos \theta)E'_2.\end{aligned}$$

Conseguentemente, la matrice associata all'applicazione identica relativa alle basi \mathcal{B} e \mathcal{B}' è

$$\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.$$

Caso 2. Ritorniamo ora al nostro sistema di coordinate di base $\mathcal{B} = \{E_1, E_2\}$. Sia $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ l'applicazione del piano in sé realizzata con una rotazione in senso antiorario di ampiezza θ . Possiamo scrivere $F = F_\theta$. Allora:

$$\begin{aligned}F(E_1) &= E'_1 = (\cos \theta)E_1 + (\sin \theta)E_2, \\F(E_2) &= E'_2 = (-\sin \theta)E_1 + (\cos \theta)E_2.\end{aligned}$$

Quindi la matrice associata a F relativa alle basi $\mathcal{B}, \mathcal{B}'$ è la trasposta della matrice trovata nel caso 1, cioè:

$$M_{\mathcal{B}}^{\mathcal{B}'}(F_\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Non è possibile evitare il fatto che la matrice del caso 1 sia la trasposta di quella del caso 2. È quindi necessario far sempre attenzione alla scelta delle basi per determinare la matrice associata a un'assegnata applicazione lineare.

Esempio 5. Siano $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{B}' = \{w_1, \dots, w_n\}$ basi di uno stesso spazio vettoriale V . Esiste allora una matrice $A = (a_{ij})$ tale che

$$\begin{aligned} w_1 &= a_{11}v_1 + \dots + a_{1n}v_n, \\ \vdots &\quad \vdots \quad \dots \quad \vdots \\ w_n &= a_{n1}v_1 + \dots + a_{nn}v_n. \end{aligned}$$

Allora per ogni $i = 1, \dots, n$ vediamo che $w_i = id(w_i)$. Per definizione quindi,

$$M_{\mathcal{B}'}^{\mathcal{B}}(id) = {}^t A.$$

D'altra parte, esiste un'unica applicazione lineare $F: V \rightarrow V$ tale che

$$F(v_1) = w_1, \dots, F(v_n) = w_n.$$

Di nuovo, per definizione, abbiamo:

$$M_{\mathcal{B}}^{\mathcal{B}}(F) = {}^t A.$$

Il teorema seguente è l'analogo del teorema 3 del paragrafo 21:

TEOREMA 4 *Siano V, W spazi vettoriali. Siano \mathcal{B} una base di V e \mathcal{B}' una base di W . Siano f, g due applicazioni lineari di V in W . Allora*

$$M(f+g) = M(f) + M(g).$$

Se c è un numero, allora

$$M(cf) = cM(f).$$

(Le matrici associate sono sempre prese relativamente alle basi fissate $\mathcal{B}, \mathcal{B}'$.)

La dimostrazione è lasciata per esercizio.

Esercizi

(Tutti i vettori dell' n -spazio devono essere considerati come vettori colonna. Per comodità tipografica continueremo a scriverli come righe, omettendo il simbolo t della trasposizione.)

1. Si considerino gli spazi $\mathbb{R}^n, \mathbb{R}^m$ con le usuali basi. Determinare la matrice associata alle applicazioni lineari seguenti.

- a) $F: \mathbb{R}^4 \rightarrow \mathbb{R}^2$ definita da $F(x_1, x_2, x_3, x_4) = (x_1, x_2)$ (cioè la proiezione).
- b) La proiezione da \mathbb{R}^4 in \mathbb{R}^3 .

- c) $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definita da $F(x, y) = (3x, 3y)$.
d) $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ definita da $F(X) = 7X$.
e) $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ definita da $F(X) = -X$.
f) $F: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ definita da $F(x_1, x_2, x_3, x_4) = (x_1, x_2, 0, 0)$.

2. Sia $\mathcal{B} = \{E_1, E_2\}$ la solita base di \mathbb{R}^3 e sia \mathcal{B}' la base ottenuta da questa ruotando il sistema di coordinate di un angolo θ . Trovare la matrice associata all'applicazione identica relativa alle basi $\mathcal{B}, \mathcal{B}'$ per ognuno dei seguenti valori di θ .

- a) $\pi/2$. b) $\pi/4$. c) π . d) $-\pi$.
e) $-\pi/3$. f) $\pi/6$. g) $5\pi/4$.

3. In generale, sia $\theta > 0$. Qual è la matrice associata all'applicazione identica, relativamente a una base fissata e alla base ottenuta ruotando questa di un angolo $-\theta$ (cioè rotazione in senso orario di ampiezza θ)?

4. Sia $X = (1, 2)$ un punto del piano. Si indichi con F la rotazione di ampiezza $\pi/4$. Quali sono le coordinate del punto $F(X)$ rispetto alla base solita $\{E_1, E_2\}$?

5. Stessa domanda quando $X = (-1, 3)$ e F denota la rotazione di ampiezza $\pi/2$.

6. In generale, sia F la rotazione di ampiezza θ . Sia (x, y) un punto del piano riferito al solito sistema di coordinate. Siano (x', y') le coordinate dello stesso punto rispetto al sistema di coordinate ruotato dell'angolo θ . Esprimere x' e y' mediante x, y, θ .

7. Sia $X = (x, y)$ e sia F la rotazione di ampiezza θ . Dimostrare che $\|X\| = \|F(X)\|$ (cioè dimostrare che F conserva il modulo).

8. In ognuno dei casi seguenti, si denoti con $D = d/dt$ la derivazione. Diamo un insieme \mathcal{B} di funzioni linearmente indipendenti: queste generano uno spazio vettoriale V e D è un'applicazione lineare di questo spazio V in sé stesso. Trovare la matrice associata all'applicazione D relativa alle basi \mathcal{B}, \mathcal{B} .

- a) $\{e^t, e^{st}\}$. b) $\{1, t\}$. c) $\{e^t, te^t\}$.
d) $\{1, t, t^2\}$. e) $\{1, t, e^t, e^{st}, te^{st}\}$. f) $\{\sin t, \cos t\}$.

23. COMPOSIZIONE DI APPLICAZIONI LINEARI

Siano U, V, W insiemi. Siano $F: U \rightarrow V$ e $G: V \rightarrow W$ applicazioni. Allora, come abbiamo già detto in precedenza, possiamo considerare l'applicazione composta $G \circ F$ da U in W .

TEOREMA 5 Siano V, W, U spazi vettoriali. Siano $\mathcal{B}, \mathcal{B}', \mathcal{B}''$ loro rispettive basi. Siano

$$F: V \rightarrow W \quad e \quad G: W \rightarrow U$$

applicazioni lineari. Allora

$$M_{\mathcal{B}'}^{\mathcal{B}''}(G) M_{\mathcal{B}'}^{\mathcal{B}}(F) = M_{\mathcal{B}''}^{\mathcal{B}}(G \circ F).$$

(Nota. Relativamente alla nostra scelta delle basi, il teorema afferma la corrispondenza della composizione delle applicazioni alla moltiplicazione delle matrici.)

Dimostrazione. Faremo vedere che il teorema 5 segue da quanto si è detto nel capitolo 3. Sia A la matrice associata a F relativamente alle basi $\mathcal{B}, \mathcal{B}'$ e sia B la matrice associata a G relativamente alle basi $\mathcal{B}', \mathcal{B}''$. Se v è un elemento di V e X è il vettore (colonna) delle sue coordinate rispetto alla base \mathcal{B} , allora il vettore delle coordinate di $F(v)$ rispetto alla base \mathcal{B}' è AX . Per definizione, il vettore delle coordinate di $G(F(v))$ rispetto alla base \mathcal{B}'' è $B(AX)$ che, per quanto si è detto nel capitolo 3, coincide con $(BA)X$. Ma $G(F(v)) = (G \circ F)(v)$. Quindi il vettore delle coordinate di $(G \circ F)(v)$ rispetto alla base \mathcal{B}'' è $(BA)X$. Per definizione, questo significa che BA è la matrice associata all'applicazione $G \circ F$. Il nostro teorema è così dimostrato.

Osservazione. In molti casi si considerano applicazioni lineari di uno spazio vettoriale V in sé stesso. Se in V è stata già fissata una base \mathcal{B} e se $F: V \rightarrow V$ è un'applicazione lineare, allora la matrice

$$M_{\mathcal{B}}^{\mathcal{B}}(F)$$

è solitamente chiamata la matrice associata a F relativamente alla base \mathcal{B} (invece di dire relativamente alle basi \mathcal{B}, \mathcal{B}). Dalle definizioni segue che

$$M_{\mathcal{B}}^{\mathcal{B}}(id) = I,$$

dove I è la matrice unità. Come diretta conseguenza del teorema 5 otteniamo il seguente:

COROLLARIO *Sia V uno spazio vettoriale e siano $\mathcal{B}, \mathcal{B}'$ sue basi, allora*

$$M_{\mathcal{B}'}^{\mathcal{B}}(\text{id}) M_{\mathcal{B}'}^{\mathcal{B}'}(\text{id}) = I = M_{\mathcal{B}'}^{\mathcal{B}'}(\text{id}) M_{\mathcal{B}}^{\mathcal{B}}(\text{id}).$$

In particolare $M_{\mathcal{B}'}^{\mathcal{B}}(\text{id})$ è invertibile.

Dimostrazione. Basta, nel teorema 5, assumere $V = W = U$ e $F = G =$ applicazione identica.

La formula generale espressa nel teorema 5 ci permette di descrivere come la matrice associata a un'applicazione lineare muti quando cambiamo le basi.

TEOREMA 6 *Sia $F: V \rightarrow V$ un'applicazione lineare e siano $\mathcal{B}, \mathcal{B}'$ basi dello spazio V . Esiste allora una matrice invertibile N tale che*

$$M_{\mathcal{B}'}^{\mathcal{B}}(F) = N^{-1} M_{\mathcal{B}}^{\mathcal{B}}(F) N.$$

Infatti, basta assumere

$$N = M_{\mathcal{B}'}^{\mathcal{B}'}(\text{id}).$$

Dimostrazione. Applicando più volte successivamente il teorema 5 otteniamo che

$$M_{\mathcal{B}'}^{\mathcal{B}'}(F) = M_{\mathcal{B}'}^{\mathcal{B}}(\text{id}) M_{\mathcal{B}}^{\mathcal{B}}(F) M_{\mathcal{B}}^{\mathcal{B}'}(\text{id}).$$

Il corollario del teorema 5 implica l'asserzione da dimostrare.

Sia V uno spazio vettoriale di dimensione finita su K e sia $F: V \rightarrow V$ un'applicazione lineare. Si dice che una base \mathcal{B} di V *diagonalizza* F se la matrice associata a F relativa alla base \mathcal{B} è una matrice diagonale. Se una base che diagonalizza F esiste, si dice che l'applicazione F è *diagonalizzabile*. Non è sempre vero che un'applicazione lineare sia diagonalizzabile. Più avanti in questo libro troveremo delle condizioni sufficienti perché ciò accada. Se A è una matrice $n \times n$ in K , noi diciamo che A può essere *diagonalizzata* (in K) se l'applicazione lineare su K^n rappresentata da A può essere diagonalizzata. Il teorema 6 permette di concludere immediatamente che:

TEOREMA 7 *Sia V uno spazio vettoriale di dimensione finita su K . Sia $F: V \rightarrow V$ un'applicazione lineare; sia M la matrice ad essa associata relativamente a una base \mathcal{B} . Allora F (o M) può*

essere diagonalizzata (in K) se, e solo se, esiste una matrice invertibile N in K tale che $N^{-1}MN$ sia una matrice diagonale.

A causa dell'importanza dell'applicazione $M \mapsto N^{-1}MN$, le assegniamo una speciale denominazione. Due matrici M, M' sono dette *simili* (in un corpo K) se esiste una matrice invertibile N in K tale che $M' = N^{-1}MN$.

Il teorema seguente dà un'altra formula per descrivere il cambio delle coordinate:

TEOREMA 8 *Siano V, W spazi vettoriali su K e sia $F: V \rightarrow W$ un'applicazione lineare. Sia \mathcal{B} una base di V e \mathcal{B}' sia una base di W . Se $v \in V$, denotiamo con $M_{\mathcal{B}'}(v)$ il vettore colonna delle coordinate di v rispetto alla base \mathcal{B} .*

Allora,

$$M_{\mathcal{B}'}(F(v)) = M_{\mathcal{B}'}^{\mathcal{B}}(F) M_{\mathcal{B}}(v).$$

Dimostrazione. L'uguaglianza da provare non è altro che una riformulazione della definizione di matrice associata a un'applicazione lineare.

COROLLARIO *Sia V uno spazio vettoriale e siano $\mathcal{B}, \mathcal{B}'$ sue basi. Sia $v \in V$. Allora*

$$M_{\mathcal{B}'}(v) = M_{\mathcal{B}'}^{\mathcal{B}}(id) M_{\mathcal{B}}(v).$$

Questo corollario esprime in forma abbreviata il modo in cui le coordinate di un vettore cambiano quando cambiamo la base dello spazio vettoriale.

Le coordinate di un vettore rispetto a una data base possono essere trovate risolvendo, in un senso intuitivo, un sistema di equazioni lineari come abbiamo visto nel paragrafo 9 (cap. 2). La matrice

$$M_{\mathcal{B}'}^{\mathcal{B}}(id)$$

può essere trovata esattamente nello stesso modo.

Esercizi

- Per ogni numero reale θ sia $F_\theta: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ l'applicazione lineare rappresentata dalla matrice

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Dimostrare che, se θ e θ' sono numeri reali, $F_\theta F_{\theta'} = F_{\theta+\theta'}$. (Nella dimostrazione adoperare la formula di addizione per il seno e il coseno.) Dimostrare anche che $F_{\bar{\theta}}^{-1} = F_{-\theta}$.

2. Sia F_θ come nell'esercizio 1. Dimostrare che per ogni vettore X in \mathbb{R}^2 , $\|F_\theta(X)\| = \|X\|$.

3. Determinare in ognuno dei seguenti casi la matrice $M_{\mathcal{B}'}^{\mathcal{B}}(id)$. In ognuno dei casi lo spazio vettoriale è sempre \mathbb{R}^3 .

a) $\mathcal{B} = \{(1, 1, 0), (-1, 1, 1), (0, 1, 2)\}$,

$$\mathcal{B}' = \{(2, 1, 1), (0, 0, 1), (-1, 1, 1)\}.$$

b) $\mathcal{B} = \{(3, 2, 1), (0, -2, 5), (1, 1, 2)\}$,

$$\mathcal{B}' = \{(1, 1, 0), (-1, 2, 4), (2, -1, 1)\}.$$

Capitolo 6

Determinanti

La maggior parte dei capitoli seguenti è indipendente dalla teoria dei determinanti. Conseguentemente, questo capitolo può essere omesso senza inficiare in modo essenziale la comprensione di quei capitoli. Comunque, i determinanti costituiscono un mezzo di calcolo molto efficiente, utile per vari scopi, principalmente per determinare quando più vettori sono linearmente indipendenti; è quindi molto conveniente poter disporre di essi in simili calcoli. Perciò, il lettore che non desidera approfondirne la teoria e che conosce il modo di calcolare un determinante, può saltare questo capitolo oppure leggere soltanto gli enunciati delle proprietà dei determinanti.

La trattazione qui seguita per i determinanti si ispira a quella data anni fa da Artin.

24. DETERMINANTI DEL SECONDO ORDINE

Prima di enunciare le proprietà generali di un determinante arbitrario, considereremo un caso particolare.

Sia

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

una matrice 2×2 in un corpo K . Definiamo suo determinante il numero $ad - cb$. Il determinante quindi è un elemento di K .

Il determinante può esser visto come una funzione della matrice A . Esso può essere anche visto come una funzione delle sue due colonne. Indicandole, come al solito, con A^1 e A^2 , possiamo indicare il determinante in uno dei seguenti modi

$$D(A), \quad \text{Det}(A), \quad \text{oppure} \quad D(A^1, A^2).$$

Le proprietà che seguono si dimostrano facilmente eseguendo i calcoli, e invitiamo il lettore a eseguirli.

Come funzione dei vettori colonne, il determinante è lineare.
Questo significa: se b' , d' sono due numeri, allora

$$\text{Det} \begin{pmatrix} a & b + b' \\ c & d + d' \end{pmatrix} = \text{Det} \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \text{Det} \begin{pmatrix} a & b' \\ c & d' \end{pmatrix}.$$

Inoltre, se t è un numero, allora

$$\text{Det} \begin{pmatrix} a & tb \\ c & td \end{pmatrix} = t \text{Det} \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Proprietà analoghe valgono anche nei riguardi della prima colonna.

Se due colonne sono uguali, allora il determinante è uguale a 0.

Se A è la matrice unità

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

allora $\text{Det}(A) = 1$.

Il determinante ha inoltre le seguenti proprietà.

Se si addiziona a una colonna un multiplo di un'altra colonna, il valore del determinante non cambia.

In altre parole, se t è un numero, il determinante della matrice

$$\begin{pmatrix} a + tb & b \\ c + td & d \end{pmatrix}$$

è $D(A)$; lo stesso accade quando addizioniamo alla seconda colonna un multiplo della prima.

Se due colonne sono scambiate tra loro, il determinante cambia segno.

In altre parole, abbiamo

$$\text{Det} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = - \text{Det} \begin{pmatrix} b & a \\ d & c \end{pmatrix}.$$

Il determinante della matrice A è uguale al determinante della sua trasposta, cioè, $D(A) = D({}^t A)$.

Esplicitamente questo significa:

$$\text{Det} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \text{Det} \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

I vettori $\begin{pmatrix} a \\ c \end{pmatrix}$ e $\begin{pmatrix} b \\ d \end{pmatrix}$ sono linearmente dipendenti se, e soltanto se, il determinante $ad - bc$ è uguale a zero.

Nel paragrafo successivo considereremo il determinante di matrice $n \times n$, le proprietà analoghe a quelle già viste ci forniranno un metodo generale per il loro calcolo.

25. PROPRIETÀ DEI DETERMINANTI

Sia A una matrice $n \times n$. Sarebbe possibile definire il suo determinante come una somma, come abbiamo fatto nella definizione di determinante di una matrice 2×2 . Tuttavia, scrivere per esteso una tale somma non è molto agevole, tanto più che per trovare il valore di un determinante non è necessario conoscere l'espressione di questa somma. Ciò che è necessario è la conoscenza delle proprietà che possono agevolarne il calcolo.

Alcune di queste proprietà sono espresse nel seguente:

TEOREMA 1 *Ad ogni matrice $n \times n A$ nel corpo K , possiamo associare un elemento di K , chiamato il suo determinante e denotato con $D(A)$ oppure con $D(A^1, \dots, A^n)$, se A^1, \dots, A^n sono le colonne di A , in modo che le seguenti proprietà risultino soddisfatte:*

1) Il determinante, come funzione di ogni vettore colonna, è lineare, cioè, se la j -esima colonna A^j è la somma di due vettori colonne, per esempio $A^j = C + C'$, allora

$$\begin{aligned} D(A^1, \dots, C + C', \dots, A^n) &= \\ &= D(A^1, \dots, C, \dots, A^n) + D(A^1, \dots, C', \dots, A^n). \end{aligned}$$

Inoltre, se $t \in K$, allora

$$D(A^1, \dots, tA^j, \dots, A^n) = tD(A^1, \dots, A^j, \dots, A^n).$$

2) Se due colonne contigue sono uguali, cioè $A^j = A^{j+1}$ per qualche valore di j tra 1 e $n-1$, allora il determinante $D(A)$ è nullo.

3) Se I è la matrice unità, allora $D(I) = 1$.

I determinanti sono univocamente determinati da queste tre proprietà.

Nel paragrafo 27 proveremo che effettivamente i determinanti esistono. In questo paragrafo ci limitiamo a dimostrare semplici proprietà che discendono dalle tre su esposte. Adopereremo la notazione $\text{Det}(A)$ invece di $D(A)$. Inoltre, il determinante della matrice (a_{ij}) è anche denotato con gli elementi che costituiscono le matrici racchiusi tra due linee verticali:

$$D(A) = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

4) *Sia j un intero positivo e minore di n . Se le colonne j -esima e $(j+1)$ -esima sono scambiate, allora il determinante cambia di segno.*

Dimostrazione. Nella matrice A , sostituiamo alle colonne j -esima e $(j+1)$ -esima il vettore colonna $A^j + A^{j+1}$. Otteniamo una matrice con due colonne contigue uguali e per la proprietà 2, abbiamo

$$0 = D(\dots, A^j + A^{j+1}, A^j + A^{j+1}, \dots).$$

Se sviluppiamo il secondo membro adoperando ripetutamente la proprietà 1 otteniamo:

$$\begin{aligned} 0 = D(\dots, A^j, A^j, \dots) &+ D(\dots, A^{j+1}, A^j, \dots) + \\ &+ D(\dots, A^j, A^{j+1}, \dots) + D(\dots, A^{j+1}, A^{j+1}, \dots). \end{aligned}$$

Per la proprietà 2, possiamo affermare che due di questi quattro termini sono nulli e quindi che

$$0 = D(\dots, A^{j+1}, A^j, \dots) + D(\dots, A^j, A^{j+1}, \dots).$$

Da quest'ultima uguaglianza segue che, come volevamo dimostrare, uno dei termini è opposto all'altro.

5) *Se due colonne A^j, A^i della matrice A sono uguali, se $j \neq i$, allora il determinante di A è uguale a zero.*

Dimostrazione. Si supponga che due colonne della matrice A siano uguali. Possiamo cambiare la matrice data scambiando successivamente colonne adiacenti finché non otteniamo una matrice con due colonne contigue uguali. (Una dimostrazione formale di questa possibilità può farsi per induzione.) Ogni volta che facciamo uno di tali scambi tra colonne contigue, il determinante cambia segno, una cosa che non muta il suo essere nullo o no.

Perciò, servendoci della proprietà 2, possiamo concludere che il determinante di A è nullo se in A due colonne sono uguali.

6) *Se si addiziona a una colonna un multiplo scalare di un'altra, il valore del determinante non cambia.*

Dimostrazione. Fissiamo l'attenzione su due diverse colonne, per esempio le colonne k -esima e j -esima, A^k e A^j , con $k \neq j$. Sia t uno scalare. Aggiungiamo alla colonna A^k il vettore colonna tA^j . Per la proprietà 1 il determinante della matrice così ottenuta si scrive:

$$D(\dots, A^k + tA^j, \dots) = D(\dots, A^k, \dots) + D(\dots, tA^j, \dots)$$

\uparrow \uparrow \uparrow
 k k k

(il k in basso indica la k -esima colonna).

In entrambi gli addendi del secondo membro, la colonna indicata si trova al k -esimo posto: ma $D(\dots, A^k, \dots)$ non è altro che $D(A)$. Inoltre,

$$D(\dots, tA^j, \dots) = tD(\dots, A^j, \dots)$$

\uparrow \uparrow
 k k

Poiché $k \neq j$, il determinante del secondo membro viene ad avere due colonne uguali perché A^j si trova tanto nel k -esimo posto quanto nel j -esimo posto. Pertanto il determinante nel secondo membro è nullo, e quindi abbiamo

$$D(\dots, A^k + tA^j, \dots) = D(\dots, A^k, \dots),$$

provando così la proprietà 6.

Esercizio

1. Siano c un elemento di K e A una matrice $n \times n$. Dimostrare che

$$D(cA) = c^n D(A).$$

26. REGOLA DI CRAMER

Le proprietà viste nel paragrafo precedente sono già sufficienti per dimostrare una ben nota regola adoperata nella risoluzione di sistemi di equazioni lineari.

TEOREMA 2 Siano A^1, \dots, A^n vettori colonna in K^n tali che
 $D(A^1, \dots, A^n) \neq 0$.

Sia B un vettore colonna di K^n . Se x_1, \dots, x_n sono elementi di K tali che

$$x_1 A^1 + \dots + x_n A^n = B,$$

allora, per ogni $j = 1, \dots, n$, abbiamo:

$$x_j = \frac{D(A^1, \dots, B, \dots, A^n)}{D(A^1, \dots, A^n)},$$

dove B occupa la j -esima colonna invece di A^j . In altre parole,

$$x_j = \frac{\begin{vmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ a_{21} & \dots & b_2 & \dots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{vmatrix}}.$$

(Il numeratore è ottenuto dalla matrice A sostituendovi la j -esima colonna A^j con B . Il denominatore è il determinante della matrice A .)

Il teorema 2 fornisce esplicitamente i valori degli elementi x_i mediante A^1, \dots, A^n . Nel linguaggio delle equazioni lineari, il teorema 2 ci permette di risolvere esplicitamente, adoperando i determinanti, il sistema di n equazioni lineari in n incognite:

$$x_1 a_{11} + \dots + x_n a_{1n} = b_1,$$

$$\cdots$$

$$x_1 a_{n1} + \dots + x_n a_{nn} = b_n.$$

Sia B scritto come nell'enunciato del teorema e consideriamo il determinante della matrice ottenuta sostituendo la j -esima colonna di A con B . Allora

$$D(A^1, \dots, B, \dots, A^n) = D(A^1, \dots, x_1 A^1 + \dots + x_n A^n, \dots, A^n).$$

Adoperando la proprietà 1 otteniamo l'espressione:

$$D(A^1, \dots, x_1 A^1, \dots, A^n) + \dots + D(A^1, \dots, x_j A^j, \dots, A^n) +$$

$$+ \dots + D(A^1, \dots, x_n A^n, \dots, A^n),$$

che diviene, sfruttando nuovamente la proprietà 1,

$$x_1 D(A^1, \dots, A^1, \dots, A^n) + \dots + x_j D(A^1, \dots, A^n) + \\ + \dots + x_n D(A^1, \dots, A^n, \dots, A^n).$$

In ogni addendo di questa somma, tranne che nel j -esimo, due vettori colonna coincidono: ogni addendo quindi, tranne il j -esimo, è nullo per la proprietà 5. Il j -esimo addendo è uguale a:

$$x_j D(A^1, \dots, A^n),$$

e coincide quindi col determinante da cui siamo partiti, cioè $D(A^1, \dots, B, \dots, A^n)$. Possiamo allora risolvere rispetto a x_j , ottenendo precisamente l'espressione data nell'enunciato del teorema.

La regola espressa nel teorema 2, che fornisce la soluzione del sistema di equazioni lineari per mezzo dei determinanti, è conosciuta sotto il nome di *regola di Cramer*.

Esempio. Risolvere il seguente sistema di equazioni lineari:

$$3x + 2y + 4z = 1,$$

$$2x - y + z = 0,$$

$$x + 2y + 3z = 1.$$

Abbiamo:

$$x = \frac{\begin{vmatrix} 1 & 2 & 4 \\ 0 & -1 & 1 \\ 1 & 2 & 3 \\ 3 & 2 & 4 \\ 2 & -1 & 1 \\ 1 & 2 & 3 \end{vmatrix}}{\begin{vmatrix} 3 & 1 & 4 \\ 2 & 0 & 1 \\ 1 & 1 & 3 \\ 3 & 2 & 4 \\ 2 & -1 & 1 \\ 1 & 2 & 3 \end{vmatrix}}, \quad y = \frac{\begin{vmatrix} 3 & 1 & 4 \\ 2 & 0 & 1 \\ 1 & 1 & 3 \\ 3 & 2 & 4 \\ 2 & -1 & 1 \\ 1 & 2 & 3 \end{vmatrix}}{\begin{vmatrix} 3 & 2 & 1 \\ 2 & -1 & 0 \\ 1 & 2 & 1 \\ 3 & 2 & 4 \\ 2 & -1 & 1 \\ 1 & 2 & 3 \end{vmatrix}}, \quad z = \frac{\begin{vmatrix} 3 & 2 & 1 \\ 2 & -1 & 0 \\ 1 & 2 & 1 \\ 3 & 2 & 4 \\ 2 & -1 & 1 \\ 1 & 2 & 3 \end{vmatrix}}{\begin{vmatrix} 3 & 2 & 1 \\ 2 & -1 & 0 \\ 1 & 2 & 1 \\ 3 & 2 & 4 \\ 2 & -1 & 1 \\ 1 & 2 & 3 \end{vmatrix}}.$$

Si osservi come la colonna

$$B = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

passa dal primo posto quando ricaviamo x , al secondo quando ricaviamo y , al terzo quando ricaviamo z . Nelle tre espressioni il denominatore è sempre lo stesso, cioè è il determinante della matrice dei coefficienti delle equazioni.

Nel paragrafo seguente daremo un metodo per il calcolo dei determinanti. Si troverà allora $x = -\frac{1}{5}$, $y = 0$, $z = \frac{2}{5}$.

I determinanti ci sono utili anche per sapere quando dei vettori dati sono linearmente indipendenti.

TEOREMA 3 *Siano A^1, \dots, A^n vettori colonna (di dimensione n). Se essi risultano linearmente dipendenti, allora*

$$D(A^1, \dots, A^n) = 0.$$

Se $D(A^1, \dots, A^n)$ è diverso da 0, allora A^1, \dots, A^n sono linearmente indipendenti.

Dimostrazione. La seconda asserzione è semplicemente una diversa formulazione della prima. Sarà quindi sufficiente provare la prima. Si supponga quindi che i vettori A^1, \dots, A^n siano linearmente dipendenti. Possiamo allora trovare dei numeri x_1, \dots, x_n , non tutti nulli, tali che

$$x_1 A^1 + \dots + x_n A^n = O.$$

Supponiamo che sia $x_j \neq 0$. Allora

$$x_j A^j = -x_1 A^1 - \dots - x_{j-1} A^{j-1} - x_{j+1} A^{j+1} - \dots - x_n A^n = \sum_{k \neq j} x_k A^k,$$

avendo così indicato che nel secondo membro non compare il termine j -esimo. Dividendo per x_j , otteniamo A^j come combinazione lineare di A^1, \dots, A^n (escluso A^j). In altre parole, ci sono dei numeri y_1, \dots, y_n (manca y_j) tali che

$$A^j = y_1 A^1 + \dots + y_n A^n = \sum_{k \neq j} y_k A^k,$$

dove, al solito, il termine j -esimo non compare. Otteniamo allora:

$$D(A^1, \dots, A^n) = D(A^1, \dots, y_1 A^1 + \dots + y_n A^n, \dots, A^n),$$

che, sviluppato con l'uso della proprietà 1, dà:

$$y_1 D(A^1, \dots, A^1, \dots, A^n) + \dots + y_n D(A^1, \dots, A^n, \dots, A^n).$$

Qui, nuovamente, manca il termine j -esimo. In tutti i termini che compaiono ci sono sempre due colonne uguali e quindi, per la proprietà 5, ognuno di essi è nullo. Questo dimostra il teorema 3.

COROLLARIO Siano A^1, \dots, A^n vettori colonna di K^n tali che $D(A^1, \dots, A^n) \neq 0$. Sia B un vettore colonna di K^n , allora esistono in K n elementi x_1, \dots, x_n tali che

$$x_1 A^1 + \dots + x_n A^n = B.$$

Dimostrazione. Per il teorema già visto, i vettori A^1, \dots, A^n sono linearmente indipendenti e quindi formano una base di K^n . È quindi chiaro che ogni vettore di K^n può essere scritto come combinazione lineare di A^1, \dots, A^n .

Esercizio

1. Risolvere i seguenti sistemi di equazioni lineari (dopo aver letto il paragrafo seguente).

a) $3x + y - z = 0,$

$x + y + z = 0,$

$y - z = 1.$

b) $2x - y + z = 1,$

$x + 3y - 2z = 0,$

$4x - 3y + z = 2.$

27. ESISTENZA DEI DETERMINANTI

Veniamo ora alla questione dell'esistenza. I determinanti esistono? La risposta è affermativa e li definiremo per induzione.

Quando $n=1$, trattiamo di matrici 1×1 e tutte le nostre proprietà sono ovviamente soddisfatte se definiamo $\text{Det}(a) = a$, per ogni scalare a .

Procedendo per induzione, supponiamo di essere riusciti a definire i determinanti, in modo che le nostre proprietà siano soddisfatte, per le matrici quadrate di ordine minore di n . Mostriremo ora come sia possibile definire il determinante per le matrici $n \times n$ facendo sì che le nostre proprietà rimangano soddisfatte. In effetti la nostra definizione farà intervenire un'altra proprietà dei determinanti che, d'altra parte, riesce molto utile nel loro calcolo. Prima di procedere dobbiamo introdurre alcune notazioni.

Siano i, j due interi tra 1 e n . Se cancelliamo la i -esima riga e la j -esima colonna in una matrice $n \times n$ A , otteniamo una matrice $(n-1) \times (n-1)$, che denoteremo con A_{ij} . Questa matrice si

scrive come segue:

$$i \left(\begin{array}{ccc|cc} a_{11} & \dots & & \dots & a_{1n} \\ \vdots & & & & \vdots \\ \hline & & a_{ij} & & \\ \vdots & & & & \vdots \\ a_{n1} & \dots & & \dots & a_{nn} \end{array} \right).$$

Esempio 1. Sia

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 4 \\ -3 & 2 & 5 \end{pmatrix}.$$

Allora

$$A_{11} = \begin{pmatrix} 1 & 4 \\ 2 & 5 \end{pmatrix}, \quad A_{22} = \begin{pmatrix} 2 & 0 \\ -3 & 5 \end{pmatrix}, \quad A_{31} = \begin{pmatrix} 1 & 0 \\ 1 & 4 \end{pmatrix}.$$

Daremo ora un'espressione del determinante di una matrice $n \times n$ adoperando determinanti di matrici $(n-1) \times (n-1)$.

Sia A una matrice $n \times n$, $A = (a_{ij})$. Sia i un intero compreso tra 1 e n . Definiamo

$$D(A) = (-1)^{i+1} a_{1i} \text{Det}(A_{11}) + \dots + (-1)^{i+n} a_{in} \text{Det}(A_{in}).$$

Notiamo che ogni A_{ij} è una matrice $(n-1) \times (n-1)$.

Questa somma può essere descritta in altre parole come segue. Per ogni elemento della i -esima riga abbiamo un addendo nella somma; questo addendo è uguale al prodotto dell'elemento stesso per il determinante della matrice ottenuta da A cancellandovi la i -esima riga e la colonna in cui si trova l'elemento in considerazione, il prodotto è preso col proprio segno o con quello opposto secondo la regola della scacchiera;

$$\begin{pmatrix} + & \dots & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \\ & & & \dots & \end{pmatrix}.$$

Questa somma è chiamata lo *sviluppo del determinante secondo la riga i -esima*. Dimostreremo che questa funzione D ha le proprietà 1, 2, 3.

Notiamo intanto che $D(A)$ è una somma di termini del tipo

$$(-1)^{i+j} a_{ij} \operatorname{Det}(A_{ij})$$

al variare di j da 1 a n .

1) Consideriamo D come funzione della k -esima colonna e consideriamo un addendo qualsiasi

$$(-1)^{i+j} a_{ij} \operatorname{Det}(A_{ij}).$$

Sia j diverso da k , allora a_{ij} non dipende dalla k -esima colonna e $\operatorname{Det}(A_{ij})$ dipende linearmente dalla k -esima colonna. Se j coincide con k , allora a_{ij} dipende linearmente dalla k -esima colonna, mentre $\operatorname{Det}(A_{ij})$ non ne dipende. In ogni caso l'addendo scelto dipende linearmente dalla k -esima colonna. Poiché $D(A)$ è somma di addendi siffatti, anche esso dipende linearmente dalla k -esima colonna e così è dimostrata la proprietà 1.

2) Supponiamo che due colonne contigue della matrice A siano uguali, $A^k = A^{k+1}$. Sia j un indice non coincidente con k né con $k+1$. Allora la matrice A_{ij} ha due colonne contigue uguali e quindi il suo determinante è nullo. Quindi, nella somma che definisce $D(A)$, gli addendi con indice diverso da k o da $k+1$ sono nulli. Gli altri due addendi possono essere scritti come segue

$$(-1)^{i+k} a_{ik} \operatorname{Det}(A_{ik}) + (-1)^{i+k+1} a_{i,k+1} \operatorname{Det}(A_{i,k+1}).$$

Le due matrici A_{ik} e $A_{i,k+1}$ sono uguali per l'ipotesi della coincidenza delle colonne k -esima e $(k+1)$ -esima nella matrice A . Per la stessa ragione, $a_{ik} = a_{i,k+1}$. Quindi questi due addendi si elidono perché sono opposti. Così anche la proprietà 2 è dimostrata.

3) Sia A la matrice unità. Allora a_{ij} è nullo eccetto il caso in cui i e j coincidono: allora $a_{ii} = 1$. Ogni A_{ij} è la matrice unità $(n-1) \times (n-1)$. L'unico termine della somma che definisce $D(A)$ e che dà contributo non nullo è

$$(-1)^{i+i} a_{ii} \operatorname{Det}(A_{ii}),$$

ed è uguale a 1. Questo dimostra la proprietà 3.

Esempio 2. Lo sviluppo secondo una riga fornisce un modo esplicito per il calcolo di un determinante. Per esempio, per calcolare il determinante

$$\begin{vmatrix} 1 & 2 & 1 \\ -1 & 3 & 1 \\ 0 & 1 & -5 \end{vmatrix}$$

possiamo adoperare lo sviluppo secondo la terza riga (perché vi è uno zero), e quindi soltanto due termini compaiono nella somma:

$$(-1) \begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} + (-5) \begin{vmatrix} 1 & 2 \\ -1 & 3 \end{vmatrix}.$$

Possiamo calcolare esplicitamente i determinanti 2×2 come nel paragrafo 24; otteniamo così -27 come valore del determinante della nostra matrice 3×3 .

In uno dei paragrafi successivi mostreremo che il determinante di una matrice è uguale al determinante della matrice trasposta. Da questo risultato seguirà allora il seguente:

TEOREMA 4 *I determinanti si possono sviluppare secondo righe oppure secondo colonne. Per ogni colonna A^j della matrice $A = (a_{ij})$, abbiamo*

$$D(A) = (-1)^{1+j} a_{1j} D(A_{1j}) + \dots + (-1)^{n+j} a_{nj} D(A_{nj}).$$

In pratica il calcolo di un determinante è sempre fatto sviluppando secondo una riga o una colonna.

Esempio 3. Scriveremo ora lo sviluppo del determinante di una matrice 3×3 secondo la prima colonna.

Sia

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Allora $D(A)$ è uguale alla somma

$$a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}.$$

Dopo quanto abbiamo detto, possiamo ora calcolare i determinanti abbastanza rapidamente. Nel far questo dobbiamo tener

presenti le operazioni descritte nella proprietà 6 per ridurre a zero quanti più elementi è possibile nella matrice data. È particolarmente utile tentare di ridurre a zero tutti gli elementi, uno escluso, di una riga o di una colonna e quindi sviluppare il determinante secondo quella riga o quella colonna. In tal modo lo sviluppo conterrà soltanto un termine e il calcolo del determinante sarà quindi ridotto al calcolo di un altro determinante di dimensioni più piccole.

Esempio 4. Calcolare il valore del determinante

$$\begin{vmatrix} 3 & 0 & 1 \\ 1 & 2 & 5 \\ -1 & 4 & 2 \end{vmatrix}.$$

Osserviamo che nella prima riga vi è già uno zero. Se alla terza riga sommiamo la seconda moltiplicata per -2 , il nostro determinante si scrive:

$$\begin{vmatrix} 3 & 0 & 1 \\ 1 & 2 & 5 \\ -3 & 0 & -8 \end{vmatrix}.$$

Sviluppiamo allora secondo la seconda colonna. Nello sviluppo compare un solo termine, col segno $+$, ed è:

$$2 \begin{vmatrix} 3 & 1 \\ -3 & -8 \end{vmatrix}.$$

Il determinante 2×2 che ivi compare può essere calcolato con la definizione che abbiamo già dato $ad - bc$, e così troviamo $2(-24 - (-3)) = -42$.

Esercizi

1. Calcolare il valore dei seguenti determinanti.

- | | | |
|---|---|---|
| a) $\begin{vmatrix} 2 & 1 & 2 \\ 0 & 3 & -1 \\ 4 & 1 & 1 \end{vmatrix}$. | b) $\begin{vmatrix} 3 & -1 & 5 \\ -1 & 2 & 1 \\ -2 & 4 & 3 \end{vmatrix}$. | c) $\begin{vmatrix} 2 & 4 & 3 \\ -1 & 3 & 0 \\ 0 & 2 & 1 \end{vmatrix}$. |
| d) $\begin{vmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 0 & 2 & 7 \end{vmatrix}$. | e) $\begin{vmatrix} -1 & 5 & 3 \\ 4 & 0 & 0 \\ 2 & 7 & 8 \end{vmatrix}$. | |

2. Calcolare il valore dei seguenti determinanti.

$$\text{a) } \begin{vmatrix} 1 & 1 & -2 & 4 \\ 0 & 1 & 1 & 3 \\ 2 & -1 & 1 & 0 \\ 3 & 1 & 2 & 5 \end{vmatrix}. \quad \text{b) } \begin{vmatrix} -1 & 1 & 2 & 0 \\ 0 & 3 & 2 & 1 \\ 0 & 4 & 1 & 2 \\ 3 & 1 & 5 & 7 \end{vmatrix}.$$

$$\text{c) } \begin{vmatrix} 3 & 1 & 1 \\ 2 & 5 & 5 \\ 8 & 7 & 7 \end{vmatrix}. \quad \text{d) } \begin{vmatrix} 4 & -9 & 2 \\ 4 & -9 & 2 \\ 3 & 1 & 0 \end{vmatrix}.$$

3. Il lettore costruisca da sé delle matrici e ne calcoli il determinante finché non si accorge di poterlo fare rapidamente.

4. a) Scrivere lo sviluppo di un determinante 3×3 secondo la seconda riga analogamente a quanto fatto nell'esempio 3.

b) Scrivere la formula generale per lo sviluppo del determinante di una matrice $n \times n$ secondo la i -esima riga.

5. a) Siano x_1, x_2, x_3 numeri: dimostrare che

$$\begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix} = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2).$$

b) Siano x_1, \dots, x_n numeri: dimostrare, per induzione, che

$$\begin{vmatrix} 1 & x_1 & \dots & x_1^{n-1} \\ 1 & x_2 & \dots & x_2^{n-1} \\ \dots & \dots & \dots & \dots \\ 1 & x_n & \dots & x_n^{n-1} \end{vmatrix} = \prod_{i < j} (x_j - x_i)$$

Il simbolo nel secondo membro significa che si tratta del prodotto di tutti i termini $x_j - x_i$ con $i < j$, i e j interi compresi tra 1 e n . Questo determinante è detto di *Vandermonde* ed è indicato con V_n . Per conseguire facilmente la dimostrazione per induzione, moltiplicare ogni colonna per x_1 e sottrarla dalla successiva, partendo dalla penultima. Si troverà che

$$V_n = (x_n - x_1) \dots (x_2 - x_1) V_{n-1}.$$

6. Sia A una matrice triangolare $n \times n$, per esempio una matrice le cui componenti al disotto della diagonale sono nulle

$$A = \begin{pmatrix} a_{11} & & & & \\ 0 & a_{22} & & * & \\ 0 & 0 & \ddots & . & \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & & & a_{nn} \end{pmatrix}.$$

Quanto vale $D(A)$?

7. Se $a(t), b(t), c(t), d(t)$ sono funzioni di t , possiamo considerare il determinante

$$\begin{vmatrix} a(t) & b(t) \\ c(t) & d(t) \end{vmatrix},$$

definito come abbiamo fatto con i numeri. Scrivere per esteso il determinante

$$\begin{vmatrix} \sin t & \cos t \\ -\cos t & \sin t \end{vmatrix}.$$

8. Scrivere per esteso il determinante

$$\begin{vmatrix} t+1 & t-1 \\ t & 2t+5 \end{vmatrix}.$$

9. Siano $f(t), g(t)$ due funzioni aventi derivate di ogni ordine. Sia $\varphi(t)$ la funzione ottenuta sviluppando il determinante seguente

$$\varphi(t) = \begin{vmatrix} f(t) & g(t) \\ f'(t) & g'(t) \end{vmatrix}.$$

Dimostrare che

$$\varphi'(t) = \begin{vmatrix} f(t) & g(t) \\ f''(t) & g''(t) \end{vmatrix}$$

(cioè, la derivata si ottiene eseguendo la derivazione delle funzioni dell'ultima riga).

10. Generalizzare l'esercizio 9 al caso di tre funzioni e quindi al caso di n funzioni. [Si suggerisce di sviluppare il determinante secondo la prima riga.]

$$\begin{vmatrix} f_1 & \dots & f_n \\ f'_1 & \dots & f'_n \\ \vdots & & \vdots \\ f^{(n-1)}_1 & \dots & f^{(n-1)}_n \end{vmatrix}.$$

11. Siano $\alpha_1, \dots, \alpha_n$ numeri distinti e non nulli. Dimostrare che le funzioni

$$e^{\alpha_1 t}, \dots, e^{\alpha_n t}$$

sono linearmente indipendenti sul corpo complesso. [Suggerimento: si supponga di avere una relazione lineare

$$c_1 e^{\alpha_1 t} + \dots + c_n e^{\alpha_n t} = 0$$

con c_i costanti, valida per ogni t . Se qualche c_i è diverso da zero, senza ledere la generalità possiamo ritenere che nessuno di essi sia zero. Derivando questa relazione $n-1$ volte, si ottengono altre $n-1$ relazioni lineari. Il determinante dei coefficienti del sistema costituito da queste n relazioni lineari deve essere zero. (Perché?) Da ciò dedurre una contraddizione.]

12. In questo esercizio si suppone la conoscenza dei polinomi. Se il lettore non ne ha conoscenza può rimandare la risoluzione di questo esercizio a quando sarà giunto ad essi nel corso del suo studio (cap. 9).

a) Siano P_{ij} , ($i, j = 1, \dots, n$) polinomi. Si supponga che, disposti i polinomi dati come segue

$$\begin{pmatrix} P_{11} & \dots & P_{1n} \\ \vdots & & \vdots \\ P_{n1} & \dots & P_{nn} \end{pmatrix},$$

quelli appartenenti a una stessa colonna abbiano lo stesso grado; siano d_1, \dots, d_n questi gradi. Si indichi poi con c_{ij} il primo coefficiente (non nullo) del polinomio P_{ij} . Sia Q il determinante della matrice sopra scritta. Dimostrare che Q ha un'espressione del tipo

$$Q(t) = ct^d + \text{termini di grado minore di } d$$

dove $c = \text{Det}(c_{ij})$. Quindi, se $\text{Det}(c_{ij})$ è diverso da zero, troviamo che Q è diverso da zero. (Se lo si ritiene conveniente, risolvere l'esercizio dapprima nel caso $n = 2$ e poi per $n = 3$ sviluppando il determinante secondo una colonna. Il caso generale può essere trattato ricorrendo al procedimento di induzione. Analogamente, nelle parti successive di questo esercizio, si può assumere $n = 2$ oppure $n = 3$.)

b) Si denoti con D la derivazione $D = d/dt$. Sia P un polinomio e sia α un numero diverso da zero. Dimostrare allora che

$$D(P(t)e^{\alpha t}) = (D + \alpha)P(t)e^{\alpha t},$$

e, per induzione, che

$$D^k(P(t)e^{\alpha t}) = (D + \alpha)^k P(t)e^{\alpha t}.$$

c) Siano $\alpha_1, \dots, \alpha_n$ numeri distinti non nulli. Dimostrare che le funzioni $e^{\alpha_1 t}, \dots, e^{\alpha_n t}$ sono linearmente indipendenti rispetto ai polinomi, cioè dimostrare che se P_1, \dots, P_n sono polinomi per cui

$$P_1(t)e^{\alpha_1 t} + \dots + P_n(t)e^{\alpha_n t} = 0$$

per ogni t , allora P_1, \dots, P_n coincidono col polinomio nullo. [Suggerimento: derivare l'espressione $n - 1$ volte. Provare che, se il numero α è diverso da zero, allora

$$\deg(D + \alpha)^k P = \deg P$$

per ogni polinomio P e ogni intero k non negativo. Si ottiene così un sistema di equazioni lineari]

$$P_{k1}(t)e^{\alpha_1 t} + \dots + P_{kn}(t)e^{\alpha_n t} = 0$$

con $k = 0, \dots, n - 1$. Quindi il determinante $\text{Det}(P_{kj})$ deve essere nullo. Servirsi poi della parte a) per dedurre una contraddizione. Riconoscere che il determinante dei primi coefficienti dei polinomi è di tipo speciale.]

28. PERMUTAZIONI

(*Nota.* Il lettore che non sopporta le dimostrazioni combinatorie potrà limitarsi a studiare soltanto gli enunciati delle proposizioni, omettendone le dimostrazioni.)

Nel seguito tratteremo soltanto di permutazioni dell'insieme di interi $\{1, \dots, n\}$, che denoteremo con J_n . Per definizione, una *permutazione* di questo insieme è un'applicazione

$$\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$$

di J_n in sé stesso tale che, se i, j sono elementi distinti di J_n , allora $\sigma(i)$ è diverso da $\sigma(j)$. Se σ è una di queste permutazioni, allora l'insieme di interi

$$\{\sigma(1), \dots, \sigma(n)\}$$

ha n elementi distinti e quindi contiene gli stessi interi $1, \dots, n$, in ordine diverso. Quindi, per ogni intero $j \in J_n$ esiste ed è unico l'intero k tale che $\sigma(k) = j$. Possiamo definire la *permutazione inversa*, denotata con σ^{-1} , come l'applicazione

$$\sigma^{-1}: J_n \rightarrow J_n$$

che associa a k l'unico intero $j \in J_n$ tale che $\sigma(j) = k$. Se σ, τ sono permutazioni dell'insieme J_n , possiamo considerare l'applicazione composta

$$\sigma \circ \tau,$$

e questa applicazione sarà ancora una permutazione. Solitamente ometteremo il piccolo circoletto e scriveremo $\sigma\tau$ per indicare l'applicazione composta. Quindi

$$(\sigma\tau)(i) = \sigma(\tau(i)).$$

Per definizione, per ogni permutazione σ , abbiamo

$$\sigma\sigma^{-1} = id \quad \text{e} \quad \sigma^{-1}\sigma = id,$$

ove id denota la permutazione identica, cioè quella per cui $id(i) = i$ per ogni $i = 1, \dots, n$.

Se $\sigma_1, \dots, \sigma_r$ sono permutazioni dell'insieme J_n , allora l'in-

versa dell'applicazione composta

$$\sigma_1 \dots \sigma_r$$

è la permutazione

$$\sigma_r^{-1} \dots \sigma_1^{-1}.$$

Ciò si vede immediatamente eseguendo la moltiplicazione.

Una *trasposizione* è una permutazione che scambia due numeri e lascia gli altri al loro posto. Evidentemente, l'inversa di una trasposizione è ancora una trasposizione.

PROPOSIZIONE 1 *Ogni permutazione dell'insieme J_n può essere espressa come prodotto di trasposizioni.*

Dimostrazione. Daremo la dimostrazione procedendo per induzione su n . Se $n = 1$ non vi è nulla da dimostrare. Sia n maggiore di 1 e si supponga vera l'affermazione per $n - 1$. Sia σ una permutazione dell'insieme J_n . Sia $\sigma(n) = k$ e sia τ la trasposizione dell'insieme J_n tale che $\tau(k) = n$, $\tau(n) = k$. Allora la permutazione $\tau\sigma$ è tale che

$$\tau\sigma(n) = \tau(k) = n.$$

In altre parole, $\tau\sigma$ lascia fisso n . Possiamo quindi considerare $\tau\sigma$ come una permutazione dell'insieme J_{n-1} , allora, per l'ipotesi di induzione, esistono le trasposizioni τ_1, \dots, τ_s di J_{n-1} , che quindi lasciano fisso n , tali che

$$\tau\sigma = \tau_1 \dots \tau_s.$$

Possiamo quindi scrivere

$$\sigma = \tau^{-1} \tau_1 \dots \tau_s,$$

dimostrando così la nostra proposizione.

PROPOSIZIONE 2 *Ad ogni permutazione σ dell'insieme J_n è possibile associare il numero 1 oppure -1 , denotato con $\varepsilon(\sigma)$, in modo che le seguenti condizioni risultino soddisfatte:*

- a) *Se τ è una trasposizione, allora $\varepsilon(\tau) = -1$.*
- b) *Se σ e σ' sono permutazioni di J_n , allora*

$$\varepsilon(\sigma\sigma') = \varepsilon(\sigma)\varepsilon(\sigma').$$

Dimostrazione. Siano x_1, \dots, x_n variabili e sia σ una permutazione dell'insieme J_n . Sia P^+ l'insieme delle coppie (i, j) in cui $1 < i < j < n$ e inoltre $\sigma(i) < \sigma(j)$. Sia P^- l'insieme delle coppie (i, j) in cui $1 < i < j < n$ e inoltre $\sigma(i) > \sigma(j)$. Sia m il numero delle coppie appartenenti a P^- . Questo numero m è chiamato il numero delle *inversioni* di σ . Infine, sia P l'insieme di tutte le coppie (i, j) in cui $1 < i < j < n$. Cioè, P è l'unione di P^- e P^+ , P^+ e P^- non hanno elementi in comune.

Consideriamo ora l'espressione

$$\Delta_\sigma = \prod_{(i,j) \in P} [x_{\sigma(j)} - x_{\sigma(i)}].$$

Il simbolo di prodotto sta a indicare che dobbiamo eseguire il prodotto sui fattori $x_{\sigma(j)} - x_{\sigma(i)}$, per tutte le coppie (i, j) in P . Possiamo allora scomporre questo prodotto in due prodotti parziali; il primo esteso a tutte le coppie appartenenti a P^+ , il secondo esteso a tutte le coppie appartenenti a P^- . Il prodotto può essere quindi scritto come segue

$$\Delta_\sigma = \prod_{(i,j) \in P^+} [x_{\sigma(j)} - x_{\sigma(i)}] \prod_{(k,l) \in P^-} [x_{\sigma(l)} - x_{\sigma(k)}].$$

Invertendo ogni fattore nel secondo prodotto, otteniamo

$$\Delta_\sigma = \prod_{(i,j) \in P^+} [x_{\sigma(j)} - x_{\sigma(i)}] (-1)^m \prod_{(k,l) \in P^-} [x_{\sigma(k)} - x_{\sigma(l)}].$$

Le coppie $(\sigma(i), \sigma(j))$ con $(i, j) \in P^+$ sono tutte distinte e similmente sono distinte le coppie $(\sigma(l), \sigma(k))$ con $(k, l) \in P^-$ perché σ è una permutazione.

Inoltre nessuna coppia $(\sigma(l), \sigma(k))$ con $(k, l) \in P^-$ può essere uguale a una coppia $(\sigma(i), \sigma(j))$ con $(i, j) \in P^+$, giacché, altrimenti, dovrebbe essere $l = i$ e $k = j$, e questo è in contraddizione col fatto che k è minore di l .

Quindi l'insieme di coppie

$$S = \begin{cases} (\sigma(i), \sigma(j)) & \text{con } (i, j) \in P^+ \\ (\sigma(k), \sigma(l)) & \text{con } (k, l) \in P^- \end{cases}$$

ha tanti elementi quanti ne ha l'unione $P = P^+ \cup P^-$. Ma, $P^+ \cup P^-$ è l'insieme di tutte le coppie d'interi (i, j) con $1 < i < j < n$. Perciò $S = P$.

Definiamo $\varepsilon(\sigma)$ come $(-1)^m$, m essendo il numero delle inversioni di σ . Allora il nostro prodotto Δ_σ può essere riscritto come segue

$$\prod_{(i,j) \in P} [x_{\sigma(j)} - x_{\sigma(i)}] = \varepsilon(\sigma) \prod_{(i,j) \in P} [x_j - x_i].$$

Sia σ' una permutazione. Sostituiamo $\sigma'(\lambda)$ a x_λ per ogni valore $\lambda = 1, \dots, n$. Allora, da una parte abbiamo

$$\begin{aligned} \prod_{i < j} [\sigma'(\sigma(j)) - \sigma'(\sigma(i))] &= \varepsilon(\sigma) \prod_{i < j} [\sigma'(j) - \sigma'(i)] = \\ &= \varepsilon(\sigma)\varepsilon(\sigma') \prod_{i < j} (j - i), \end{aligned}$$

e, d'altra parte, il prodotto del primo membro è uguale a

$$\varepsilon(\sigma') \prod_{i < j} (j - i).$$

Possiamo quindi concludere che $\varepsilon(\sigma'\sigma) = \varepsilon(\sigma')\varepsilon(\sigma)$.

Per determinare il segno di una trasposizione, consideriamo l'insieme delle coppie P^- associate a una trasposizione σ . Supponiamo che $1 < \alpha < \beta < n$ e che σ sia una trasposizione tale che $\sigma(\alpha) = \beta$ e $\sigma(\beta) = \alpha$. Allora P^- è costituita da tutte le coppie

$$\begin{aligned} (\alpha, l) &\quad \text{con} \quad \alpha + 1 \leq l \leq \beta, \\ (k, \beta) &\quad \text{con} \quad \alpha + 1 \leq k \leq \beta - 1. \end{aligned}$$

Perciò P^- contiene un numero dispari di coppie e quindi il segno delle trasposizioni è -1 . Questo dimostra la nostra proposizione.

COROLLARIO 1 Se una permutazione σ dell'insieme J_n è espressa come prodotto di trasposizioni:

$$\sigma = \tau_1 \dots \tau_s,$$

dove ogni τ_i è una trasposizione, allora il numero s è pari o dispari secondo se $\varepsilon(\sigma) = 1$ oppure -1 .

Dimostrazione. Avendosi

$$\varepsilon(\sigma) = (-1)^s,$$

la nostra asserzione è immediata.

COROLLARIO 2 Se σ è una permutazione dell'insieme J_n , allora

$$\varepsilon(\sigma) = \varepsilon(\sigma^{-1}).$$

Dimostrazione. Abbiamo

$$1 = \varepsilon(id) = \varepsilon(\sigma\sigma^{-1}) = \varepsilon(\sigma)(\sigma^{-1}).$$

Quindi, $\varepsilon(\sigma)$ ed $\varepsilon(\sigma^{-1})$ o sono entrambi uguali a 1 o sono entrambi uguali a -1 , come era da dimostrare.

Terminiamo con una questione di terminologia: una permutazione è chiamata *pari* se il suo segno è 1, è chiamata *dispari* se il suo segno è -1 . In particolare, ogni trasposizione è dispari.

Esercizi

1. Una permutazione σ degli interi $1, \dots, n$ è talvolta indicata con $\begin{bmatrix} 1 & \dots & n \\ \sigma(1) & \dots & \sigma(n) \end{bmatrix}$. Per esempio, $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix}$ denota la permutazione σ tale che $\sigma(1) = 2$, $\sigma(2) = 1$, $\sigma(3) = 3$.

Questa permutazione è, in effetti, una trasposizione. Determinare il segno di ognuna delle seguenti permutazioni ed esprimerele come prodotto di trasposizioni.

- | | | |
|---|---|---|
| a) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$. | b) $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}$. | c) $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}$. |
| d) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{bmatrix}$. | e) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{bmatrix}$. | f) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{bmatrix}$. |

2. In ognuno dei casi riportati nell'esercizio 1, scrivere la permutazione inversa.

3. Dimostrare che il numero delle permutazioni dispari dell'insieme $\{1, \dots, n\}$, per n maggiore di uno, è uguale al numero delle permutazioni pari.

29. UNICITÀ

Prima di affrontare la parte centrale della nostra argomentazione, facciamo alcune considerazioni sullo sviluppo delle applicazioni lineari come abbiamo visto nella proprietà 1. Consideriamo dapprima il caso $n = 2$. Se dobbiamo sviluppare

$$D(3A + 5B, 2A - B),$$

dove A, B sono due vettori di uno spazio bidimensionale, allora,

servendoci della proprietà 1, otteniamo una somma di quattro termini, cioè

$$\begin{aligned} D(3A, 2A - B) + D(5B, 2A - B) &= \\ &= D(3A, 2A) + D(3A, -B) + D(5B, 2A) + D(5B, -B) = \\ &= 6D(A, A) - 3D(A, B) + 10D(B, A) - 5D(B, B). \end{aligned}$$

Osserviamo che abbiamo fatto uso del fatto che quando t coincide con -1 , nella proprietà 1, accade che

$$D(A, -B) = -D(A, B).$$

In questa espressione notiamo che $D(A, A) = 0$ e $D(B, B) = 0$. Quindi il possibile contributo non nullo può venir dato soltanto da due termini.

Per considerare un altro esempio, sviluppiamo il determinante $D(2A + B - C, 3E + F)$ dove A, B, C, E, F sono vettori. Servendoci ripetutamente della proprietà 1, otteniamo una somma di sei termini, cioè

$$6D(A, E) + 2D(A, F) + 3D(B, E) + D(B, F) - 3D(C, E) - D(C, F).$$

In generale, consideriamo il determinante

$$D(x_1 A^1 + \dots + x_n A^n, y_1 B^1 + \dots + y_m B^m)$$

dove $x_1, \dots, x_n, y_1, \dots, y_m$ sono numeri e $A^1, \dots, A^n, B^1, \dots, B^m$ sono vettori. Servendoci ripetutamente della proprietà 1, possiamo scrivere questo determinante come segue

$$D(x_1 A^1, y_1 B^1 + \dots + y_m B^m) + \dots + D(x_n A^n, y_1 B^1 + \dots + y_m B^m).$$

Ogni addendo può essere ulteriormente sviluppato e così il nostro determinante è espresso dalla somma doppia

$$\begin{aligned} D(x_1 A^1, y_1 B^1) + \dots + D(x_1 A^1, y_m B^m) + D(x_n A^n, y_1 B^1) + \\ + \dots + D(x_n A^n, y_m B^m). \end{aligned}$$

Lo schema che si segue nello sviluppare un tale tipo di determinante è il seguente. Si fa la somma dei termini ottenuti scegliendo, in tutti i modi possibili, un termine nella prima somma e un termine nella seconda. La nostra somma doppia può quindi

essere così scritta

$$\sum_{i=1}^n \sum_{j=1}^m x_i y_j D(A^i, B^j).$$

Nella stessa maniera si può procedere quando si abbia a che fare con sviluppi più complessi.

Vogliamo trovare ora un'espressione esplicita di un determinante, per cui siano soddisfatte le condizioni 1, 2, 3, mediante le componenti della matrice. Osserviamo che possiamo usare liberamente le proprietà 4, 5, 6 in quanto queste conseguono dalle proprietà 1, 2, 3.

Sia A una matrice $n \times n$, siano A^1, \dots, A^n i suoi vettori colonna. Allora possiamo scrivere

$$\begin{aligned} A^1 &= a_{11}E^1 + \dots + a_{n1}E^n, \\ &\vdots \quad \vdots \quad \vdots \\ A^n &= a_{1n}E^1 + \dots + a_{nn}E^n, \end{aligned}$$

avendo denotato con E^1, \dots, E^n i vettori colonna unità. Allora

$$D(A^1, \dots, A^n) = D(a_{11}E^1 + \dots + a_{n1}E^n, \dots, a_{1n}E^1 + \dots + a_{nn}E^n).$$

E per la proprietà 1 il secondo membro può essere scritto come somma di termini del tipo

$$D(a_{\sigma(1),1}E^{\sigma(1)}, \dots, a_{\sigma(n),n}E^{\sigma(n)}),$$

dove $\sigma(1), \dots, \sigma(n)$ sono n interi distinti scelti tra 1 e n . Quindi σ è un'applicazione dell'insieme di interi $\{1, \dots, n\}$ in sé stesso. Adoperando di nuovo la proprietà 1, ognuno dei termini riportati sopra può essere così riscritto

$$a_{\sigma(1),1} \dots a_{\sigma(n),n} D(E^{\sigma(1)}, \dots, E^{\sigma(n)}).$$

Se qualche applicazione σ assegna lo stesso intero a due distinti valori i, j tra 1 e n , il determinante a destra viene ad avere due colonne uguali e quindi è nullo. Conseguentemente, possiamo limitarci a considerare nella nostra somma soltanto quelle applicazioni σ per cui $i \neq j$ implica $\sigma(i) \neq \sigma(j)$: cioè le *permutazioni*. Invece di dire che noi consideriamo la somma estesa a tutte le permutazioni σ , riassumiamo questa frase nella notazione con

l'usuale segno \sum . Cioè scriviamo

$$D(A^1, \dots, A^n) = \sum_{\sigma} a_{\sigma(1),1} \dots a_{\sigma(n),n} D(E^{\sigma(1)}, \dots, E^{\sigma(n)}).$$

I vettori unità $E^{\sigma(1)}, \dots, E^{\sigma(n)}$ costituiscono una permutazione rispetto all'ordine naturale E^1, \dots, E^n . Con successivi scambi di colonne contigue possiamo riportare i vettori nell'ordine naturale. Ogni volta che scambiamo due colonne contigue, il determinante cambia segno. Se $m(\sigma)$ è il numero delle trasposizioni di vettori colonna contigui che occorrono per riportare i vettori stessi nell'ordine naturale, allora

$$D(E^{\sigma(1)}, \dots, E^{\sigma(n)}) = (-1)^{m(\sigma)} D(E^1, \dots, E^n) = (-1)^{m(\sigma)}.$$

Il segno $(-1)^{m(\sigma)}$ è il segno della permutazione σ . Perciò, finalmente, possiamo scrivere

$$D(A^1, \dots, A^n) = \sum_{\sigma} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n},$$

intendendo la somma estesa a tutte le permutazioni dell'insieme di interi $\{1, \dots, n\}$. Questa espressione prova che il valore del determinante è univocamente determinato dalle proprietà 1, 2, 3. In altre parole, abbiamo dimostrato il seguente:

TEOREMA 5 *I determinanti sono univocamente determinati dalle proprietà 1, 2, 3. Il determinante può essere espresso come segue*

$$D(A^1, \dots, A^n) = \sum_{\sigma} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n},$$

intendendo la somma estesa a tutte le permutazioni dell'insieme di interi $\{1, \dots, n\}$.

L'espressione mediante la somma data nel teorema 5 è praticamente inutile nei calcoli, essa si dimostra invece conveniente in certe considerazioni di carattere teorico: per esempio nel paragrafo successivo nel dimostrare che il determinante della matrice è uguale al determinante della matrice trasposta.

Osservazione. Nel caso $n = 2$, la somma che compare nel teorema 5 è niente altro che la formula $ad - bc$ del paragrafo 24.

Esercizi

- Siano A^1, \dots, A^n vettori colonna di dimensione n e si supponga che essi siano linearmente indipendenti. Dimostrare che $D(A^1, \dots, A^n)$ è diverso

da zero. [Si suggerisce di esprimere ognuno dei vettori unità E^1, \dots, E^n , considerato come vettore colonna, come combinazione lineare dei vettori A^1, \dots, A^n . Usando il fatto che $D(E^1, \dots, E^n) = 1$ e le proprietà 1 e 2, concludere la dimostrazione.]

2. Siano A^1, \dots, A^n vettori colonna di dimensione n aventi componenti reali. Dimostrare che se essi sono linearmente indipendenti su \mathbb{R} , lo sono anche sul corpo complesso \mathbb{C} . Dimostrare anche il viceversa. In generale, se le componenti dei vettori A^1, \dots, A^n sono in un corpo K , dimostrate che i vettori A^1, \dots, A^n sono linearmente indipendenti su K se, e soltanto se, essi lo sono sul corpo complesso \mathbb{C} .

30. DETERMINANTE DELLA TRASPOSTA DI UNA MATRICE

TEOREMA 6 *Sia A una matrice quadrata. Allora $\text{Det}(A) = \text{Det}({}^t A)$.*

Dimostrazione. Nel teorema 5 abbiamo visto che

$$\text{Det}(A) = \sum_{\sigma} \varepsilon(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n}. \quad [1]$$

Sia σ una permutazione dell'insieme $\{1, \dots, n\}$. Se $\sigma(j) = k$, allora $\sigma^{-1}(k) = j$. Possiamo quindi scrivere

$$a_{\sigma(j),j} = a_{k,\sigma^{-1}(k)}.$$

In ogni prodotto

$$a_{\sigma(1),1} \dots a_{\sigma(n),n}$$

ogni intero k compreso tra 1 e n compare esattamente una volta tra gli interi $\sigma(1), \dots, \sigma(n)$. Questo prodotto quindi può essere così riscritto

$$a_{1,\sigma^{-1}(1)} \dots a_{n,\sigma^{-1}(n)},$$

e la nostra somma [1] è uguale a

$$\sum_{\sigma} \varepsilon(\sigma^{-1}) a_{1,\sigma^{-1}(1)} \dots a_{n,\sigma^{-1}(n)},$$

perché $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$. In questa somma ogni addendo corrisponde a una permutazione σ ; inoltre, al variare di σ nell'insieme di tutte le permutazioni, anche σ^{-1} varia nell'insieme di tutte le permutazioni perché ogni permutazione determina univocamente la sua inversa. Perciò la nostra somma è uguale a

$$\sum_{\sigma} \varepsilon(\sigma) a_{1,\sigma(1)} \dots a_{n,\sigma(n)}. \quad [2]$$

Ma la somma [2] coincide con lo sviluppo del determinante della matrice trasposta di A . Abbiamo così dimostrato quanto avevamo asserito.

31. DETERMINANTE DI UN PRODOTTO

Vogliamo ora dimostrare la seguente importante regola:

TEOREMA 7 *Siano A , B due matrici $n \times n$. Allora*

$$\text{Det}(AB) = \text{Det}(A) \text{ Det}(B).$$

Il determinante di un prodotto è uguale al prodotto dei determinanti.

Dimostrazione. Sia $A = (a_{ij})$ e $B = (b_{jk})$:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1k} & \dots & b_{1n} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \dots & b_{nk} & \dots & b_{nn} \end{pmatrix}.$$

Sia $AB = C$ e sia C^k la k -esima colonna della matrice C . Allora, per definizione,

$$C^k = b_{1k} A^1 + \dots + b_{nk} A^n.$$

Quindi

$$\begin{aligned} D(AB) &= D(C^1, \dots, C^n) = \\ &= D(b_{11} A^1 + \dots + b_{n1} A^n, \dots, b_{1n} A^1 + \dots + b_{nn} A^n). \end{aligned}$$

Sviluppando questa espressione, con l'ausilio della proprietà I, troviamo la somma

$$\begin{aligned} \sum_{\sigma} D(b_{\sigma(1),1} A^{\sigma(1)}, \dots, b_{\sigma(n),n} A^{\sigma(n)}) &= \\ &= \sum_{\sigma} b_{\sigma(1),1} \dots b_{\sigma(n),n} D(A^{\sigma(1)}, \dots, A^{\sigma(n)}) = \\ &= \sum_{\sigma} \epsilon(\sigma) b_{\sigma(1),1} \dots b_{\sigma(n),n} D(A^1, \dots, A^n). \end{aligned}$$

Ricordando la formula per lo sviluppo di un determinante che abbiamo trovata, questa somma è uguale a $D(B)D(A)$; come volevamo dimostrare.

COROLLARIO 1 *Sia A una matrice invertibile $n \times n$, allora*

$$\text{Det}(A^{-1}) = \text{Det}(A)^{-1}.$$

Dimostrazione. Si ha $1 = D(I) = D(AA^{-1}) = D(A)D(A^{-1})$. Questo prova quanto volevamo.

COROLLARIO 2 *Siano X^1, \dots, X^n vettori colonna di K^n ; sia A una matrice $n \times n$ in K . Allora*

$$\text{Det}(AX^1, \dots, AX^n) = \text{Det}(A) \text{ Det}(X^1, \dots, X^n).$$

Dimostrazione. Si tratta di una semplice riformulazione del teorema: basta indicare con

$$B = X = (X^1, \dots, X^n)$$

la matrice le cui colonne sono i vettori X^1, \dots, X^n .

Nel corollario 2 possiamo interpretare A come un'applicazione lineare di K^n in sé stesso: il corollario dice come calcolare un determinante se i suoi vettori colonna vengono trasformati mediante un'applicazione lineare.

C'è un altro modo di procedere nella dimostrazione del teorema 7 che, forse, può riuscire più chiara. Siano v_1, \dots, v_n vettori di K^n e sia $A = (a_{ij})$ una matrice con elementi in K . Possiamo allora considerare il prodotto

$$A \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

nello stesso modo in cui noi moltiplichiamo le matrici, cioè

$$w_1 = a_{11}v_1 + \dots + a_{1n}v_n,$$

$$\dots$$

$$w_n = a_{n1}v_1 + \dots + a_{nn}v_n.$$

È di nuovo semplice verificare il permanere della proprietà associativa: se B è un'altra matrice $n \times n$ in K , allora

$$B \left(A \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right) = (BA) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Sviluppando al solito modo, troviamo la relazione

$$\begin{aligned} D(w_1, \dots, w_n) &= D(a_{11}v_1 + \dots + a_{1n}v_n, \dots, a_{n1}v_1 + \dots + a_{nn}v_n) = \\ &= \text{Det}(A) D(v_1, \dots, v_n). \end{aligned} \quad [1]$$

Per l'associatività e per la relazione [1], possiamo concludere che

$$D(AB)D(v_1, \dots, v_n) = D(A)D(B)D(v_1, \dots, v_n).$$

Ora, per ottenere la regola di moltiplicazione tra determinanti, basta supporre che v_1, \dots, v_n siano i vettori unità e^1, \dots, e^n .

32. INVERSA DI UNA MATRICE

Sia A una matrice $n \times n$. Se B è una matrice tale che $AB = I$ e $BA = I$ (I = matrice unità $n \times n$), allora diciamo che B è un'inversa di A e scriviamo $B = A^{-1}$. Se esiste un'inversa della matrice A , essa è unica. Infatti, sia C un'inversa di A . Allora $CA = I$ e moltiplicando per B a destra, otteniamo $CAB = B$. Ma $CAB = C(AB) = CI = C$. Quindi $C = B$. Un'argomentazione simile vale nel caso $AC = I$.

TEOREMA 8 *Sia $A = (a_{ij})$ una matrice $n \times n$ e si assuma che $D(A)$ non sia nullo. Allora A è invertibile. Sia E^j il j -esimo vettore unità colonna e sia*

$$b_{ij} = \frac{D(A^1, \dots, E^j, \dots, A^n)}{D(A)},$$

dove E^j compare all'i-esimo posto. Allora la matrice $B = (b_{ij})$ è l'inversa di A .

Dimostrazione. Sia $X = (x_{ij})$ una matrice $n \times n$ a elementi ignoti. Vogliamo trovare le componenti x_{ij} in modo che sia vera l'uguaglianza $AX = I$. Per la definizione di prodotto tra matrici questo significa che per ogni j dobbiamo risolvere l'equazione

$$E^j = x_{1j}A^1 + \dots + x_{nj}A^n.$$

In effetti si tratta di un sistema di equazioni lineari che può essere univocamente risolto facendo uso della regola di Cramer; otteniamo così

$$x_{ij} = \frac{D(A^1, \dots, E^j, \dots, A^n)}{D(A)},$$

che è la formula data nell'enunciato del teorema.

Dobbiamo ancora provare che $XA = I$. Notiamo intanto che

$D({}^t A)$ non è nullo. Quindi, per quanto abbiamo già dimostrato, possiamo trovare una matrice Y tale che ${}^t A Y = I$. Considerando le matrici trasposte otteniamo ${}^t Y A = I$. Quindi abbiamo

$$I = {}^t Y(AX)A = {}^t YA(XA) = XA,$$

provando così quanto volevamo, cioè che $X = B$ è l'inversa di A .

Possiamo scrivere esplicitamente le componenti della matrice B nel teorema 8 come segue:

$$b_{ij} = \frac{\begin{vmatrix} a_{11} & \dots & 0 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{j1} & \dots & 1 & \dots & a_{jn} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & 0 & \dots & a_{nn} \end{vmatrix}}{\text{Det}(A)}.$$

Se sviluppiamo il determinante che compare nel numeratore secondo la i -esima colonna, allora tutti i termini tranne uno sono nulli e quindi otteniamo il numeratore di b_{ij} come un sottodeterminante di $\text{Det}(A)$. Indichiamo con A_{ij} la matrice ottenuta da A cancellandovi la i -esima riga e la j -esima colonna. Allora

$$b_{ij} = \frac{(-1)^{i+j} \text{Det}(A_{ji})}{\text{Det}(A)}$$

(si noti lo scambio degli indici) e quindi concludiamo con la formula

$$A^{-1} = \text{trasposta di } \left(\frac{(-1)^{i+j} \text{Det}(A_{ij})}{\text{Det}(A)} \right).$$

Esempio. Sia

$$A = \begin{pmatrix} 1 & 2 & 5 \\ 2 & 3 & 1 \\ -1 & 1 & 1 \end{pmatrix}.$$

Sia $A^{-1} = (b_{ij})$. Per esempio, vogliamo calcolare b_{12} . Eseguendo i calcoli si trova che $\text{Det}(A) = 21$. Ora sappiamo che

$$b_{12} = \frac{(-1)^{1+2} \text{Det}(A_{21})}{\text{Det}(A)}.$$

La matrice A_{21} è uguale a

$$A_{21} = \begin{pmatrix} 2 & 5 \\ 1 & 1 \end{pmatrix}$$

(nella matrice A abbiamo cancellato la seconda riga e la prima colonna). Quindi

$$\text{Det}(A_{21}) = -3.$$

Infine otteniamo

$$b_{12} = -(-3/21) = 3/21.$$

La matrice A^{-1} si può quindi scrivere

$$A^{-1} = \begin{pmatrix} b_{11} & \frac{b_{21}}{b_{11}} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}.$$

Per esercizio, calcolarne tutti gli altri termini.

Esercizi

1. Trovare le inverse delle matrici nell'esercizio 1, paragrafo 27.

2. Sapendo che, se A, B sono due matrici $n \times n$ allora

$$\text{Det}(AB) = \text{Det}(A)\text{Det}(B),$$

dimostrare che una matrice A con determinante nullo può non avere inversa.

3. Scrivere esplicitamente l'inversa della matrice 2×2

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

4. Qual è l'inversa della matrice $\begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}$, e quella della matrice $\begin{pmatrix} b & 0 \\ 0 & c \end{pmatrix}$ dove b, c sono numeri non nulli?

33. DETERMINANTE DI UN'APPLICAZIONE LINEARE

Sia V uno spazio vettoriale di dimensione finita su un corpo K , sia $A: V \rightarrow V$ un'applicazione lineare di V in sé stesso. Sia \mathcal{B} una base di V e sia

$$M = M_{\mathcal{B}}^{\mathcal{B}}(A)$$

La matrice che rappresenta A rispetto alla base \mathcal{B} . Sia \mathcal{B}' un'altra base di V , e sia M' la matrice che rappresenta A rispetto a questa nuova base \mathcal{B}' , sappiamo allora che esiste una matrice invertibile N tale che $M' = N^{-1}MN$. Conseguentemente, ricordando il corollario del teorema 7, possiamo scrivere:

$$\text{Det}(M') = D(N^{-1})D(M)D(N) = D(M).$$

Il determinante è quindi indipendente dalla scelta della base e della matrice che rappresenta l'applicazione lineare.

Perciò, possiamo definire *determinante di un'applicazione lineare* A il numero $\text{Det}(M)$ di una qualunque matrice M scelta come si è detto sopra.

Alcune proprietà dei determinanti di matrici si traducono immediatamente in proprietà di determinanti di applicazioni lineari. Enunciamo ora queste proprietà lasciando le loro (semplici) dimostrazioni per esercizio.

Se A, B sono applicazioni lineari di V in sé stesso, allora $D(AB) = D(A)D(B)$.

Se A è invertibile, allora $D(A^{-1}) = D(A)^{-1}$.

Se I è l'applicazione identica, allora $D(I) = 1$.

Un'applicazione lineare dello spazio V in sé stesso è invertibile se, e solo se, il suo determinante è diverso da zero.

Esercizi

1. Sia V lo spazio vettoriale delle matrici $n \times n$ sul corpo K . Sia B un elemento di V e sia $\varphi_B: V \rightarrow V$ l'applicazione definita da

$$\varphi_B(A) = AB - BA.$$

Dimostrare che l'applicazione φ_B è lineare e che il suo determinante è nullo.

2. Siano A_1, \dots, A_m matrici quadrate di dimensioni rispettive $d_1 \times d_1, \dots, d_m \times d_m$. Dimostrare che il determinante della matrice

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

costituita dai blocchi A_1, \dots, A_m sulla diagonale è uguale al prodotto

$$D(A) = D(A_1) \dots D(A_m).$$

3. Sia V lo spazio delle matrici 2×2 sul corpo K , e sia $B \in V$. Sia $L_B: V \rightarrow V$ l'applicazione definita da $L_B(A) = BA$. Dimostrare che l'applicazione L_B è lineare e che

$$\text{Det}(L_B) = (\text{Det } B)^2.$$

[Si suggerisce di considerare l'applicazione di L_B alla base di V costituita dalle matrici $\{e_{11}, e_{21}, e_{12}, e_{22}\}$, dove e_{ij} è la matrice avente per componente 1 nel posto ij -esimo e zero altrove.]

4. Generalizzare l'esercizio 3 al caso di matrici $n \times n$. In altre parole, sia V lo spazio delle matrici $n \times n$ sul corpo K , e sia $B \in V$. Sia L_B l'applicazione lineare dello spazio V in sé stesso definita da $L_B(A) = BA$. Dimostrare che $\text{Det}(L_B) = (\text{Det } B)^n$.

5. Sia A una matrice 2×2 sul corpo K e sia $A^2 = O$. Dimostrare che per ogni scalare c , abbiamo $\text{Det}(cI - A) = c^2$.

6. Sia V lo spazio vettoriale sul corpo reale generato dalle due funzioni $\{\sin t, \cos t\}$ che consideriamo come una base di V . Qual è il determinante dell'applicazione lineare su questo spazio definita dalla derivazione?

7. Sia V lo spazio vettoriale sul corpo reale generato dalle funzioni $\{e^{\alpha t}, e^{\beta t}\}$ dove α, β sono due numeri reali distinti e consideriamo queste due funzioni come una base di V . Qual è il determinante dell'applicazione lineare definita su questo spazio dalla derivazione?

Capitolo 7

Prodotti scalari e ortogonalità

34. PRODOTTI SCALARI

Sia V uno spazio vettoriale su un corpo K . Un *prodotto scalare* su V è un modo di associare ad ogni coppia di elementi v, w appartenenti a V uno scalare, indicato con $\langle v, w \rangle$, oppure con $v \cdot w$, in modo che le seguenti proprietà siano soddisfatte:

PS 1. *Per ogni coppia di elementi v, w in V , abbiamo $\langle v, w \rangle = \langle w, v \rangle$.*

PS 2. *Se u, v, w sono elementi di V , allora*

$$\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle.$$

PS 3. *Se $x \in K$ allora*

$$\langle xu, v \rangle = x \langle u, v \rangle \quad e \quad \langle u, xv \rangle = x \langle u, v \rangle.$$

Il prodotto scalare si dice *non degenere* se è soddisfatta anche la proprietà seguente:

Se v è un elemento di V e $\langle v, w \rangle = 0$ per ogni w appartenente a V , allora $v = O$.

Esempio 1. Sia $V = K^n$. Allora l'applicazione

$$(X, Y) \mapsto X \cdot Y,$$

che agli elementi X, Y di K^n associa il loro prodotto scalare come è stato definito in precedenza, è un prodotto scalare anche secondo la definizione ora data.

Esempio 2. Sia V lo spazio delle funzioni reali continue definite nell'intervallo $[0, 1]$. Se f, g sono in V , definiamo

$$\langle f, g \rangle = \int_0^1 f(t)g(t) dt.$$

Semplici proprietà dell'integrale mostrano che si tratta di un prodotto scalare.

In entrambi gli esempi dati il prodotto scalare è non degenere. Questo lo abbiamo già rilevato parlando di prodotto scalare tra vettori in K^n . Nel secondo esempio la cosa si vede facilmente avendo presenti semplici proprietà dell'integrale.

Sia V uno spazio vettoriale in cui è definito un prodotto scalare. Come al solito, definiamo due elementi v, w di V *ortogonali* o *perpendicolari* e scriviamo $v \perp w$ se $\langle v, w \rangle = 0$. Se S è un sottoinsieme di V , denotiamo con S^\perp l'insieme di tutti gli elementi $w \in V$ che risultano perpendicolari ad ogni elemento di S , cioè $\langle w, v \rangle = 0$ per ogni $v \in S$. Dalle proprietà PS 2 e PS 3 segue immediatamente che S^\perp è un sottospazio di V , chiamato il *complemento ortogonale di S* . Se w è perpendicolare a S , scriviamo anche $w \perp S$. Sia U il sottospazio di V generato dagli elementi di S . Se w è perpendicolare a S , se v_1, v_2 sono in S , allora

$$\langle w, v_1 + v_2 \rangle = \langle w, v_1 \rangle + \langle w, v_2 \rangle = 0.$$

Se c è uno scalare, allora

$$\langle w, cv_1 \rangle = c\langle w, v_1 \rangle.$$

Quindi w è perpendicolare ad ogni combinazione lineare di elementi di S , perciò w risulta perpendicolare anche a U .

Esempio 3. Sia (a_{ij}) una matrice $m \times n$ in K e siano A_1, \dots, A_m i suoi vettori riga. Sia $X = (x_1, \dots, x_n)$. Allora il sistema di equazioni lineari omogenee

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= 0 \\ &\dots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= 0 \end{aligned} \tag{1}$$

può essere anche scritto in forma abbreviata come segue

$$A_1 \cdot X = 0, \dots, A_m \cdot X = 0.$$

L'insieme delle soluzioni X di questo sistema omogeneo è uno spazio vettoriale su K . Infatti, sia W lo spazio generato dai vettori A_1, \dots, A_m e sia U lo spazio costituito da tutti i vettori di K^n perpendicolari ad A_1, \dots, A_m . Allora U è precisamente lo spazio vettoriale delle soluzioni del sistema [1]. Osserviamo che i vettori A_1, \dots, A_m possono non essere linearmente indipendenti. Notiamo poi che la dimensione di W non supera m ; chiamiamo

$$\dim U = \dim W^\perp$$

la dimensione dello spazio delle soluzioni del sistema di equazioni lineari dato. Nel seguito discuteremo più approfonditamente questa dimensione.

Sia di nuovo V uno spazio vettoriale sul corpo K in cui sia definito un prodotto scalare.

Sia $\{v_1, \dots, v_n\}$ una base di V . Diremo che questa è una *base ortogonale* se, tutte le volte che $i \neq j$, $\langle v_i, v_j \rangle = 0$. Faremo vedere più avanti che se V è uno spazio vettoriale di dimensione finita in cui è definito un prodotto scalare, in V è sempre possibile trovare una base ortogonale. Cominciamo col discutere dapprima alcuni casi particolari importanti, relativi al corpo dei numeri reali e a quello dei numeri complessi.

35. PRODOTTI DEFINITI POSITIVI

Sia V uno spazio vettoriale sul corpo reale nel quale sia definito un prodotto scalare. Chiameremo questo prodotto scalare *definito positivo* se, per ogni $v \in V$, $\langle v, v \rangle > 0$ e, per ogni v non nullo, $\langle v, v \rangle > 0$. L'ordinario prodotto scalare tra vettori di R^n è definito positivo, e lo stesso accade per il prodotto scalare definito nell'esempio 2 sopra considerato (§ 34).

Sia V uno spazio vettoriale sul corpo reale in cui sia definito un prodotto scalare definito positivo, denotato al solito con \langle , \rangle . Sia W un sottospazio. Allora in W è definito un prodotto scalare con la stessa legge con la quale è definito in tutto V . In altre parole, se w, w' sono elementi di W , possiamo considerare il loro prodotto $\langle w, w' \rangle$. Questo è un prodotto scalare su W ed è, ovviamente, definito positivo.

Per esempio, se W è il sottospazio di R^3 generato dai due vettori $(1, 1, 2)$ e $(\pi, -1, 0)$, allora W è esso stesso uno spazio

vettoriale e possiamo considerare il prodotto scalare tra due vettori di W come definizione di un prodotto scalare definito positivo in W stesso. Spesso dovremo considerare siffatti sottospazi e questa è una delle ragioni per cui sviluppiamo la nostra teoria per spazi vettoriali arbitrari (di dimensione finita) sul corpo \mathbb{R} nei quali si è definito un prodotto scalare definito positivo, invece di considerare soltanto lo spazio \mathbb{R}^n con l'ordinario prodotto scalare. Un'altra ragione sta nel fatto che vogliamo applicare la nostra teoria alla situazione descritta nell'esempio 2 del paragrafo 34.

Analogamente a quanto abbiamo visto con gli ordinari vettori nell' n -spazio, abbiamo qui di nuovo la nozione di proiezione. Sia V uno spazio vettoriale su \mathbb{R} , sia definito in V un prodotto scalare definito positivo. Siano v, w elementi di V e sia $w \neq O$. Definiamo *coefficiente di Fourier* di v rispetto a w il numero

$$c = \frac{\langle v, w \rangle}{\langle w, w \rangle},$$

e definiamo *proiezione di v lungo w* il vettore cw . Allora i vettori $v - cw$ e w sono perpendicolari, giacché

$$\langle v - cw, w \rangle = \langle v, w \rangle - \frac{\langle v, w \rangle}{\langle w, w \rangle} \langle w, w \rangle = 0.$$

Sia $\{v_1, \dots, v_n\}$ una base ortogonale dello spazio V . Sia $v \in V$. Esistono allora numeri x_1, \dots, x_n tali che

$$v = x_1 v_1 + \dots + x_n v_n.$$

Eseguendo il prodotto scalare con v_i , per $i = 1, \dots, n$, otteniamo

$$\langle v, v_i \rangle = x_i \langle v_i, v_i \rangle$$

perché $\langle v_i, v_j \rangle = 0$ se $i \neq j$. Quindi x_i è il coefficiente di Fourier di v rispetto a v_i , cioè

$$x_i = \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle}.$$

Quindi, quando abbiamo un prodotto scalare definito positivo e una base ortogonale del nostro spazio, le coordinate di un elemento di V rispetto a questa base sono semplicemente i suoi coefficienti di Fourier.

Passiamo a formulare un teorema analogo al teorema 7 del capitolo 2 (§ 10) avendo a disposizione un prodotto scalare.

TEOREMA 1 *Sia V uno spazio vettoriale di dimensione finita sul corpo reale, in esso sia definito un prodotto scalare definito positivo. Sia W un sottospazio di V e sia $\{w_1, \dots, w_m\}$ una base ortogonale di W . Se W non coincide con V , allora esistono in V degli elementi w_{m+1}, \dots, w_n in modo che $\{w_1, \dots, w_n\}$ sia una base ortogonale di V .*

Dimostrazione. Il metodo di dimostrazione di questo teorema è altrettanto importante quanto il teorema stesso ed è chiamato il processo di *ortogonalizzazione di Gram-Schmidt*. Dal teorema 7 del capitolo 2 traiamo l'esistenza degli elementi v_{m+1}, \dots, v_n di V tali che

$$\{w_1, \dots, w_m, v_{m+1}, \dots, v_n\}$$

sia una base di V . Naturalmente, non ci aspettiamo che questa sia una base ortogonale. Sia W_{m+1} lo spazio generato dai vettori w_1, \dots, w_m, v_{m+1} . Costruiremo dapprima una base ortogonale per W_{m+1} . L'idea che seguiremo consiste nel considerare v_{m+1} e sottrarlo dalla sua proiezione lungo w_1, \dots, w_m . Quindi, consideriamo i numeri

$$c_1 = \frac{\langle v_{m+1}, w_1 \rangle}{\langle w_1, w_1 \rangle}, \dots, c_m = \frac{\langle v_{m+1}, w_m \rangle}{\langle w_m, w_m \rangle},$$

e poniamo

$$w_{m+1} = v_{m+1} - c_1 w_1 - \dots - c_m w_m.$$

Allora, w_{m+1} è perpendicolare ad ognuno dei vettori w_1, \dots, w_m perché, per ogni intero i tale che $1 < i < m$, abbiamo

$$\langle w_{m+1}, w_i \rangle = \langle v_{m+1}, w_i \rangle - \langle c_i w_i, w_i \rangle = 0.$$

Inoltre, w_{m+1} è diverso da zero (altrimenti v_{m+1} dipenderebbe linearmente dai vettori w_1, \dots, w_m) e w_{m+1} appartiene allo spazio generato da w_1, \dots, w_{m+1} , perché

$$v_{m+1} = w_{m+1} + c_1 w_1 + \dots + c_m w_m.$$

Perciò, $\{w_1, \dots, w_{m+1}\}$, è una base ortogonale di W_{m+1} . Possiamo ora procedere per induzione, mostrando che lo spazio W_{m+s} generato dai vettori $w_1, \dots, w_m, v_{m+1}, \dots, v_{m+s}$ ha una base orto-

gonale

$$\{w_1, \dots, w_{m+1}, \dots, w_{m+s}\}$$

per $s = 1, \dots, n-m$, e concludere così la dimostrazione.

COROLLARIO *Sia V uno spazio vettoriale di dimensione finita sul corpo reale in cui sia dato un prodotto scalare definito positivo. Si supponga che V non consiste del solo elemento zero. Allora, V ha una base ortogonale.*

Dimostrazione. Per ipotesi, esiste un elemento v di V tale che $v \neq O$. Sia W il sottospazio generato da v , allora basta applicare il teorema per ottenere la base richiesta.

Esponiamo ancora una volta, riassumendola, la procedura adoperata nel dimostrare il teorema 1. Supponiamo che sia data una base arbitraria $\{v_1, \dots, v_n\}$ dello spazio V . Vogliamo renderla ortogonale. Procediamo allora così. Poniamo

$$\begin{aligned} v'_1 &= v_1, \\ v'_2 &= v_2 - \frac{\langle v_2, v'_1 \rangle}{\langle v'_1, v'_1 \rangle} v'_1, \\ v'_3 &= v_3 - \frac{\langle v_3, v'_2 \rangle}{\langle v'_2, v'_2 \rangle} v'_2 - \frac{\langle v_3, v'_1 \rangle}{\langle v'_1, v'_1 \rangle} v'_1, \\ &\vdots \\ v'_n &= v_n - \frac{\langle v_n, v'_{n-1} \rangle}{\langle v'_{n-1}, v'_{n-1} \rangle} v'_{n-1} - \dots - \frac{\langle v_n, v'_1 \rangle}{\langle v'_1, v'_1 \rangle} v'_1. \end{aligned}$$

Allora $\{v'_1, \dots, v'_n\}$ è una base ortogonale.

Le definizioni del paragrafo 4, capitolo 1 e le proprietà qui dimostrate valgono per un arbitrario prodotto scalare *definito positivo*. Per esempio, possiamo definire la *norma* di un elemento v di V , ponendo

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

Allora valgono le tre proprietà seguenti:

Per ogni v in V , abbiamo $\|v\| \geq 0$ e l'uguaglianza sussiste se, e solo se, $v = O$.

Per ogni numero reale x abbiamo $\|xv\| = |x|\|v\|$.

Per ogni coppia di elementi v, w di V , abbiamo $\|v + w\| \leq \|v\| + \|w\|$.

Quest'ultima proprietà è chiamata la *disuguaglianza triangolare*. La sua dimostrazione si consegue adoperando la *disuguaglianza di Schwarz*, cioè

$$|\langle v, w \rangle| \leq \|v\| \|w\|,$$

che, a sua volta, si dimostra esattamente come il teorema 1 del paragrafo 4 del capitolo 1. Si osservi, infatti, che nel paragrafo 4 i teoremi 1, 2, 3 sono stati dimostrati senza far uso delle coordinate e quindi le loro dimostrazioni valgono senz'altro nel contesto più generale nel quale ora ci troviamo.

Sia di nuovo V uno spazio vettoriale sul corpo reale, nel quale sia dato un prodotto scalare definito positivo. Un elemento $v \in V$ è chiamato *vettore unità* se la sua norma è uguale a 1. Due vettori non nulli v, w di V si dicono avere la *stessa direzione* se esiste un numero reale positivo c tale che $cv = w$; v e w hanno invece *direzioni opposte* se esiste un numero reale negativo c in modo che $cv = w$. Se v è un vettore arbitrario di V , allora

$$\frac{v}{\|v\|}$$

è un vettore unità avente la stessa direzione di v .

Si supponga ora lo spazio V di dimensione finita. Una base $\{v_1, \dots, v_n\}$ di V si dice *ortonormale* se essa è ortogonale e se ogni v_i ($i = 1, \dots, n$) è un vettore unità. Da una base ortogonale di V possiamo sempre trarre una base ortonormale dividendo ogni elemento della base data per la sua norma. Otteniamo allora vettori di norma 1 che sono ancora mutuamente perpendicolari. Possiamo quindi enunciare il teorema 1 e il suo corollario per basi ortonormali.

TEOREMA 1' *Sia V uno spazio vettoriale di dimensione finita sul corpo reale, in cui sia dato un prodotto scalare definito positivo. Sia W un sottospazio di V e sia $\{w_1, \dots, w_m\}$ una base ortonormale di W . Se W non coincide con V , allora esistono in V degli elementi w_{m+1}, \dots, w_n in modo che $\{w_1, \dots, w_n\}$ sia una base ortonormale di V .*

COROLLARIO *Sia V uno spazio vettoriale di dimensione finita sul corpo reale, in cui sia dato un prodotto scalare definito positivo. Se V non consiste del solo elemento O , allora V ha una base ortonormale.*

Esempio 4. Trovare una base ortonormale dello spazio vettoriale sul corpo reale generato dai vettori $(1, 1, 0, 1)$, $(1, -2, 0, 0)$, $(1, 0, -1, 2)$.

Denotiamo questi vettori con A , B , C e poniamo

$$B' = B - \frac{B \cdot A}{A \cdot A} A.$$

In altre parole, sottraiamo da B la sua proiezione lungo A . Allora B' risulta perpendicolare ad A e troviamo:

$$B' = \frac{1}{6}(4, -5, 0, 1).$$

Ora sottraiamo da C le sue proiezioni rispettivamente lungo A e B' , ottenendo

$$C' = C - \frac{C \cdot A}{A \cdot A} A - \frac{C \cdot B'}{B' \cdot B'} B'.$$

Poiché A e B' sono perpendicolari, considerando il prodotto scalare di C' con A e B' ci accorgiamo che C' è perpendicolare tanto ad A quanto a B' . Troviamo così:

$$C' = \frac{1}{7}(-4, -2, -1, 6).$$

I vettori A , B' , C' sono non nulli e mutuamente perpendicolari. Essi appartengono allo spazio generato dai vettori A , B , C . Essi perciò costituiscono una base ortogonale di questo spazio. Se noi desideriamo avere una base ortonormale basta dividere questi vettori per la loro norma e così otteniamo i vettori

$$\frac{A}{\|A\|} = \frac{1}{\sqrt{3}}(1, 1, 0, 1),$$

$$\frac{B'}{\|B'\|} = \frac{1}{\sqrt{42}}(4, -5, 0, 1),$$

$$\frac{C'}{\|C'\|} = \frac{1}{\sqrt{57}}(-4, -2, -1, 6),$$

che costituiscono una base ortonormale.

Sia V uno spazio vettoriale di dimensione finita su \mathbb{R} , in esso sia definito un prodotto scalare definito positivo. Sia $\{e_1, \dots, e_n\}$ una sua base ortonormale. Siano v, w elementi di V . Allora esiste

stono dei numeri $x_1, \dots, x_n \in \mathbb{R}$ e $y_1, \dots, y_n \in \mathbb{R}$ tali che

$$v = x_1 e_1 + \dots + x_n e_n$$

e

$$w = y_1 e_1 + \dots + y_n e_n.$$

Allora

$$\langle v, w \rangle = \langle x_1 e_1 + \dots + x_n e_n, y_1 e_1 + \dots + y_n e_n \rangle =$$

$$= \sum_{i,j=1}^n x_i y_j \langle e_i, e_j \rangle = x_1 y_1 + \dots + x_n y_n.$$

Quindi, adoperando questa base ortonormale, se X, Y sono rispettivamente i vettori delle coordinate di v e w , il loro prodotto scalare è uguale all'ordinario prodotto scalare $X \cdot Y$ dei vettori delle coordinate. Questo, però, certamente non accade quando abbiamo a che fare con una base non ortonormale. Se $\{v_1, \dots, v_n\}$ è una base qualsiasi di V e se noi scriviamo

$$v = x_1 v_1 + \dots + x_n v_n,$$

$$w = y_1 v_1 + \dots + y_n v_n$$

in termini degli elementi di base, allora

$$\langle v, w \rangle = \sum_{i,j=1}^n x_i y_j \langle v_i, v_j \rangle.$$

Ogni elemento $\langle v_i, v_j \rangle$ è un numero che indichiamo con a_{ij} . Allora

$$\langle v, w \rangle = \sum_{i,j=1}^n a_{ij} x_i y_j.$$

Prodotti hermitiani

Tratteremo ora di una modifica necessaria per adattare i risultati precedenti agli spazi vettoriali sul corpo complesso. Vogliamo fare in modo da conservare il più possibile la nozione di prodotto scalare definito positivo. Osserviamo che il prodotto scalare di un vettore a componenti complesse per sé stesso può risultare nullo senza che il vettore lo sia, dobbiamo quindi cambiare qualcosa nella definizione. Vedremo che i cambiamenti necessari sono molto lievi.

Sia V uno spazio vettoriale sul corpo complesso. Un *prodotto*

hermitiano in V è un modo di associare ad ogni coppia di vettori v, w di V un numero complesso, come prima denotato con $\langle v, w \rangle$, in modo che le proprietà seguenti siano soddisfatte:

PH 1. Per ogni v, w in V abbiamo $\langle v, w \rangle = \overline{\langle w, v \rangle}$. (La sopra-lineatura denota il numero complesso coniugato.)

PH 2. Se u, v, w sono elementi di V , allora

$$\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle.$$

PH 3. Se α è un numero complesso, allora

$$\langle \alpha u, v \rangle = \alpha \langle u, v \rangle \quad \text{e} \quad \langle u, \alpha v \rangle = \bar{\alpha} \langle u, v \rangle.$$

Il prodotto hermitiano si dice *definito positivo* se $\langle v, v \rangle$ risulta non negativo per ogni $v \in V$, mentre se $v \neq O$, $\langle v, v \rangle > 0$.

Le nozioni di *ortogonalità*, *perpendicolarità*, *basi ortogonali* e *complemento ortogonale* sono definite come nel caso precedente. Non vi è nulla da cambiare neppure nella definizione di *coefficienti di Fourier*, né in quella di *proiezione del vettore v lungo il vettore w* o nelle osservazioni che abbiamo fatte a proposito di queste nozioni.

Esempio 5. Sia $V = \mathbb{C}^n$. Se $X = (x_1, \dots, x_n)$ e $Y = (y_1, \dots, y_n)$ sono vettori di \mathbb{C}^n , definiamo come loro prodotto hermitiano il numero

$$\langle X, Y \rangle = x_1 \bar{y}_1 + \dots + x_n \bar{y}_n.$$

Le tre condizioni sopra scritte sono immediatamente verificate. Questo prodotto risulta inoltre definito positivo perché se $X \neq O$, allora qualcuno dei numeri x_i è diverso da 0 e perciò $x_i \bar{x}_i$ è positivo e positivo risulta anche $\langle X, X \rangle$.

Esempio 6. Sia V lo spazio delle funzioni continue a valori complessi definite nell'intervallo $[-\pi, \pi]$. Se $f, g \in V$, definiamo

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(t) \overline{g(t)} dt.$$

Note proprietà dell'integrale permettono anche in questo caso di verificare che si tratta di un prodotto hermitiano definito positivo.

Si indichi con f_n la funzione definita ponendo

$$f_n(t) = e^{int}.$$

Una semplice verifica prova che per ogni coppia n, m d'interi distinti le funzioni f_n e f_m sono ortogonali. Abbiamo inoltre:

$$\langle f_n, f_n \rangle = \int_{-\pi}^{\pi} e^{int} e^{-int} dt = 2\pi.$$

Se $f \in V$, il suo coefficiente di Fourier rispetto alla funzione f_n si scrive allora

$$\frac{\langle f, f_n \rangle}{\langle f_n, f_n \rangle} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt,$$

espressione ben nota al lettore provvisto di conoscenze di analisi.

Ritorniamo ora alla discussione generale sui prodotti hermitiani definiti. Il teorema 1 e il suo corollario hanno un analogo per i prodotti hermitiani definiti positivi, cioè:

TEOREMA 2 *Sia V uno spazio vettoriale di dimensione finita sul corpo complesso. In V sia dato un prodotto hermitiano definito positivo. Sia W un sottospazio di V e $\{w_1, \dots, w_m\}$ sia una base ortogonale di W . Se W non coincide con V , è possibile trovare degli elementi w_{m+1}, \dots, w_n in V in modo che $\{w_1, \dots, w_n\}$ sia una base ortogonale di V .*

COROLLARIO *Sia V uno spazio vettoriale di dimensione finita sul corpo complesso, in esso sia dato un prodotto hermitiano definito positivo. Se V non consiste del solo elemento O , allora V ha una base ortogonale.*

Le dimostrazioni sono esattamente uguali a quelle date in precedenza per il caso reale; non vi è quindi bisogno di ripeterle.

Vogliamo ora trattare delle norme. Sia V uno spazio vettoriale sul corpo complesso, in V sia dato un prodotto hermitiano definito positivo. Se $v \in V$, noi definiamo sua *norma* il numero

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

Osserviamo che $\langle v, v \rangle$ è un numero reale non negativo e quindi

il segno di radice quadrata indica, come al solito, l'unico numero reale non negativo il cui quadrato è $\langle v, v \rangle$.

Vale ancora la *disuguaglianza di Schwarz*:

$$|\langle v, w \rangle| \leq \|v\| \|w\|,$$

e ne diamo di nuovo la dimostrazione perché vi è qualche differenza con quella precedente, dovuta al comparire di numeri complessi coniugati.

Sia $\alpha = \langle w, w \rangle$ e $\beta = -\langle v, w \rangle$. Abbiamo allora

$$\begin{aligned} 0 &\leq \langle \alpha v + \beta w, \alpha v + \beta w \rangle \\ &= \langle \alpha v, \alpha v \rangle + \langle \beta w, \alpha v \rangle + \langle \alpha v, \beta w \rangle + \langle \beta w, \beta w \rangle \\ &= \alpha \bar{\alpha} \langle v, v \rangle + \beta \bar{\alpha} \langle w, v \rangle + \alpha \bar{\beta} \langle v, w \rangle + \beta \bar{\beta} \langle w, w \rangle. \end{aligned}$$

Sostituendo ora ad α e β i rispettivi valori e tenendo conto che

$$\alpha = \langle w, w \rangle = \|w\|^2$$

è un numero reale, otteniamo

$$0 \leq \|w\|^4 \|v\|^2 - 2\|w\|^2 \langle v, w \rangle \overline{\langle v, w \rangle} + \|w\|^2 \langle v, w \rangle \overline{\langle v, w \rangle}.$$

Ma $\langle v, w \rangle \overline{\langle v, w \rangle} = |\langle v, w \rangle|^2$. Quindi

$$\|w\|^2 |\langle v, w \rangle|^2 \leq \|w\|^4 \|v\|^2.$$

Se $w = O$ la disuguaglianza è ovvia, nel caso contrario, dividiamo la disuguaglianza ottenuta per $\|w\|^2$ e otteniamo la disuguaglianza

$$|\langle v, w \rangle|^2 \leq \|w\|^2 \|v\|^2.$$

La disuguaglianza di Schwarz si ottiene allora estraendo la radice quadrata da entrambi i membri.

Le tre proprietà della norma, già viste nel caso reale, continuano a valere:

Per ogni $v \in V$, abbiamo $\|v\| \geq 0$, e $\|v\| = 0$ se, e solo se, $v = O$.

Per ogni numero complesso α , abbiamo $\|\alpha v\| = |\alpha| \|v\|$.

Per ogni coppia di elementi v, w in V abbiamo $\|v + w\| \leq \|v\| + \|w\|$.

Le dimostrazioni, anche in questo caso, sono molto semplici, lasciamo le prime due come esercizio. Mostreremo invece come la terza proprietà possa dedursi dalla disuguaglianza di Schwarz.

Osserviamo intanto che è sufficiente dimostrare che

$$\|v + w\|^2 \leq (\|v\| + \|w\|)^2.$$

A tal fine, osserviamo che

$$\|v + w\|^2 = \langle v + w, v + w \rangle = \langle v, v \rangle + \langle w, v \rangle + \langle v, w \rangle + \langle w, w \rangle.$$

Ma $\langle w, v \rangle + \langle v, w \rangle = \overline{\langle v, w \rangle} + \langle v, w \rangle \leq 2|\langle v, w \rangle|$. Quindi, ricordando la diseguaglianza di Schwarz,

$$\begin{aligned} \|v + w\|^2 &\leq \|v\|^2 + 2|\langle v, w \rangle| + \|w\|^2 \leq \\ &\leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2. \end{aligned}$$

La diseguaglianza da provare si ottiene estraendo la radice quadrata dal primo e dall'ultimo membro.

Un elemento v di V è detto *vettore unità* se, come nel caso reale, la sua norma è 1. Una base ortogonale $\{v_1, \dots, v_n\}$ si dice *ortonormale* se tutti i suoi vettori sono unità. Come nel caso precedente, una base ortonormale si ottiene da una ortogonale dividendone ciascun vettore per la propria norma.

Sia $\{e_1, \dots, e_n\}$ una base ortonormale di V . Siano v, w elementi di V . Esistono allora numeri complessi $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n$ in modo che

$$v = \alpha_1 e_1 + \dots + \alpha_n e_n \quad .$$

e

$$w = \beta_1 e_1 + \dots + \beta_n e_n.$$

Allora

$$\begin{aligned} \langle v, w \rangle &= \langle \alpha_1 e_1 + \dots + \alpha_n e_n, \beta_1 e_1 + \dots + \beta_n e_n \rangle = \\ &= \sum_{i,j=1}^n \alpha_i \bar{\beta}_j \langle e_i, e_j \rangle = \\ &= \alpha_1 \bar{\beta}_1 + \dots + \alpha_n \bar{\beta}_n. \end{aligned}$$

Quindi, con riferimento a questa base ortonormale, se A, B sono i rispettivi vettori delle coordinate di v e w , il prodotto hermitiano di questi vettori coincide con il prodotto descritto nell'esempio 5, cioè $A \cdot \bar{B}$.

Abbiamo ora un teorema che enunciamo contemporaneamente per il caso reale e per quello complesso.

TEOREMA 3 *Sia V uno spazio vettoriale sul corpo reale \mathbb{R} , in cui sia dato un prodotto scalare definito positivo oppure uno spazio vet-*

toriale sul corpo complesso \mathbf{C} , in cui sia dato un prodotto hermitiano definito positivo. Sia n la dimensione di V . Sia W un sottospazio di V avente dimensione r . Sia U il sottospazio di V costituito da tutti gli elementi di V che sono perpendicolari a W , sia cioè $U = W^\perp$. Allora la dimensione di U è $n - r$. In altre parole

$$\dim W + \dim W^\perp = \dim V.$$

Dimostrazione. Se W è costituito dal solo elemento O oppure se $W = V$, allora la nostra asserzione è ovvia. Assumiamo quindi che W non coincida né con $\{O\}$ né con V . Sia $\{w_1, \dots, w_r\}$ una base ortonormale di W . I teoremi 1' e 2' (formulabile sostituendo nell'enunciato del teorema 2 la parola *ortogonale* con *ortonormale* [N.d.T.]), assicurano l'esistenza di elementi u_{r+1}, \dots, u_n in V in modo che $\{w_1, \dots, w_r, u_{r+1}, \dots, u_n\}$ sia una base ortonormale di V . Proveremo che $\{u_{r+1}, \dots, u_n\}$ è una base ortonormale di U .

Sia u un elemento di U . Esistono allora numeri x_1, \dots, x_n in modo che

$$u = x_1 w_1 + \dots + x_r w_r + x_{r+1} u_{r+1} + \dots + x_n u_n.$$

Poiché u è perpendicolare a W , eseguendone il prodotto con un qualsiasi elemento w_i , $i = 1, \dots, r$, troviamo

$$0 = \langle u, w_i \rangle = x_i \langle w_i, w_i \rangle = x_i.$$

Quindi, per $i = 1, \dots, r$, $x_i = 0$ e perciò u è combinazione lineare di u_{r+1}, \dots, u_n .

Viceversa, se $u = x_{r+1} u_{r+1} + \dots + x_n u_n$ è una combinazione lineare di u_{r+1}, \dots, u_n , considerandone il prodotto con ogni w_i si ottiene zero. Perciò u è perpendicolare ad ogni w_i ($i = 1, \dots, r$) e quindi è perpendicolare a tutto W . È così provato che gli elementi u_{r+1}, \dots, u_n generano U . Essi formano una base ortonormale di U perché sono mutuamente ortogonali e di norma unitaria. Il sottospazio U viene quindi ad avere la dimensione $n - r$, come si doveva dimostrare.

Esempio. Consideriamo nello spazio vettoriale \mathbb{R}^3 due vettori linearmente indipendenti A, B . Allora lo spazio dei vettori che sono perpendicolari contemporaneamente ad A e B ha dimensione 1. Se $\{N\}$ è una base di questo spazio, ogni altra base è del tipo $\{tN\}$, dove t è un numero non nullo.

Consideriamo, nuovamente in \mathbb{R}^3 , un vettore non nullo N .

Lo spazio dei vettori perpendicolari a N ha dimensione 2, si tratta cioè di un piano passante per l'origine O .

TEOREMA 4 *Sia V uno spazio vettoriale sul corpo reale \mathbf{R} , in cui sia dato un prodotto scalare definito positivo oppure uno spazio vettoriale sul corpo complesso \mathbf{C} , in cui sia dato un prodotto hermitiano definito positivo. Si assuma V di dimensione finita. Se W è un sottospazio di V , allora V risulta somma diretta di W e W^\perp .*

Dimostrazione. La stessa argomentazione adoperata nel dimostrare il teorema 3 serve a dimostrare questo teorema. Infatti, riferendosi alle notazioni introdotte in quella dimostrazione, vediamo che ogni elemento di V è univocamente esprimibile come combinazione lineare

$$x_1 w_1 + \dots + x_r w_r + x_{r+1} u_{r+1} + \dots + x_n u_n,$$

e quindi è univocamente esprimibile come somma $w + u$, con $w \in W$ e $u \in U$.

Esercizi

1. Determinare una base ortonormale del sottospazio di \mathbf{R}^3 generato dai vettori seguenti:
 - a) $(1, 1, -1)$ e $(1, 0, 1)$.
 - b) $(2, 1, 1)$ e $(1, 3, -1)$.
2. Trovare una base ortonormale del sottospazio di \mathbf{R}^4 generato dai vettori seguenti:
 - a) $(1, 2, 1, 0)$ e $(1, 2, 3, 1)$.
 - b) $(1, 1, 0, 0)$, $(1, -1, 1, 1)$ e $(-1, 0, 2, 1)$.
3. Negli esercizi che seguono consideriamo lo spazio vettoriale delle funzioni reali continue definite nell'intervallo $[0, 1]$. Definiamo un prodotto scalare per due tali funzioni f, g ponendo

$$\langle f, g \rangle = \int_0^1 f(t)g(t) dt.$$

Adoperando note proprietà dell'integrale, dimostrare che si tratta effettivamente di un prodotto scalare.

4. Sia V il sottospazio di funzioni generato dalle funzioni f, g definite, rispettivamente, da $f(t) = t$ e $g(t) = t^n$. Trovare una base ortonormale di V .

5. Sia V il sottospazio generato dalle tre funzioni $1, t, t^2$ (1 indica la funzione costantemente uguale a 1). Trovare una base ortonormale di V .

6. Trovare una base ortonormale del sottospazio di \mathbf{C}^3 generato dai seguenti vettori:

$$\text{a) } (1, i, 0) \text{ e } (1, 1, 1). \quad \text{b) } (1, -1, -i) \text{ e } (i, 1, 2).$$

7. Sia V lo spazio vettoriale di tutte le matrici $n \times n$ a elementi reali. Si definisca un prodotto scalare per due matrici A, B ponendo

$$\langle A, B \rangle = \text{tr}(AB),$$

dove "tr" indica la traccia (somma degli elementi diagonali). Dimostrare che effettivamente si tratta di un prodotto scalare non degenere.

8. Con riferimento all'esercizio 7, descrivere il complemento ortogonale del sottospazio delle matrici diagonali. Qual è la dimensione di tale complemento ortogonale?

9. Sia V uno spazio vettoriale di dimensione finita sul corpo reale, in cui sia dato un prodotto scalare definito positivo. Sia $\{v_1, \dots, v_m\}$ un insieme di elementi di V mutuamente ortogonali e di norma unitaria (cioè, $\langle v_i, v_j \rangle = 0$ se $i \neq j$, $\langle v_i, v_i \rangle = 1$ se $i = j$). Si assuma che per ogni $v \in V$ si abbia

$$\|v\|^2 = \sum_{i=1}^m \langle v, v_i \rangle^2.$$

Dimostrare che $\{v_1, \dots, v_m\}$ è una base di V .

10. Sia V uno spazio vettoriale di dimensione finita sul corpo reale, in cui sia definito un prodotto scalare definito positivo. Provare che per ogni coppia di elementi v, w in V vale la seguente uguaglianza (detta *uguaglianza del parallelogramma*)

$$\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

36. BASI ORTOGONALI NEL CASO GENERALE

Sia V uno spazio vettoriale di dimensione finita sul corpo K , in V sia definito un prodotto scalare. Questo prodotto scalare può non essere definito positivo e interessanti esempi di siffatti prodotti si trovano persino nei numeri reali. Per esempio, si può definire il prodotto di due vettori $X = (x_1, x_2)$ e $Y = (y_1, y_2)$ come il numero $x_1 y_1 - x_2 y_2$. Allora

$$\langle X, X \rangle = x_1^2 - x_2^2.$$

Questo prodotto viene considerato in parecchie applicazioni, in fisica per esempio, quando si tratta di un prodotto di vettori nel

4-spazio tale che se $X = (x, y, z, t)$, allora

$$\langle X, X \rangle = x^2 + y^2 + z^2 - t^2.$$

In questo paragrafo vogliamo vedere quanto di ciò che si è detto a proposito delle basi ortogonali può continuare a valere.

Sia V uno spazio vettoriale di dimensione finita sul corpo K , in V sia definito un prodotto scalare. Se W è un sottospazio, non sempre accade che V è somma diretta di W e W^\perp . Questo dipende dal fatto che in V possano trovarsi vettori non nulli v tali che $\langle v, v \rangle = 0$. Per esempio, in uno spazio sul corpo complesso, $(1, i)$ può essere uno di questi vettori. Il teorema che afferma l'esistenza di una base ortogonale, tuttavia, continua a valere, e ne daremo una dimostrazione apportando le modificazioni del caso ai ragionamenti fatti nel paragrafo precedente.

Cominciamo col fare alcune osservazioni. Si supponga dapprima che per ogni elemento u di V si abbia $\langle u, u \rangle = 0$. Allora il prodotto scalare viene detto *nullo* e V è chiamato uno *spazio nullo*. La ragione di questa denominazione sta nel fatto che, per ogni coppia di vettori v, w appartenenti a V , $\langle v, w \rangle = 0$. Abbiamo infatti

$$\langle v, w \rangle = \frac{1}{2}[\langle v + w, v + w \rangle - \langle v, v \rangle - \langle w, w \rangle].$$

Per ipotesi, il secondo membro di questa uguaglianza è zero, come si vede immediatamente sviluppando il primo addendo. Si può allora dire che ogni base di V è una base ortogonale.

TEOREMA 5 *Sia V uno spazio vettoriale di dimensione finita sul corpo K . In V sia definito un prodotto scalare. Se V non consiste del solo elemento O , allora V ha una base ortogonale.*

Dimostrazione. Procediamo per induzione sulla dimensione dello spazio V . Se V ha dimensione 1, allora ogni elemento non nullo di V è una base ortogonale di V e perciò il teorema è vero.

Supponiamo ora che la dimensione n dello spazio V sia maggiore di 1. Possono presentarsi due casi.

Caso 1. Per ogni elemento $u \in V$, abbiamo $\langle u, u \rangle = 0$. Allora, come abbiamo già osservato, ogni base di V è una base ortogonale.

Caso 2. Esiste in V un elemento v_1 tale che $\langle v_1, v_1 \rangle \neq 0$. Possiamo allora procedere con lo stesso metodo seguito nel caso del

prodotto scalare definito positivo; ci riferiamo al processo di ortogonalizzazione di Gram-Schmidt. Proveremo infatti che se v_1 è un elemento di V tale che $\langle v_1, v_1 \rangle \neq 0$ e se V_1 è lo spazio di dimensione 1 generato dal vettore v_1 , allora V è somma diretta di V_1 e V_1^\perp . Sia $v \in V$ e sia, come al solito,

$$c = \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle}.$$

Allora $v - cv_1$ appartiene allo spazio V_1^\perp e quindi l'espressione

$$v = (v - cv_1) + cv_1$$

prova che V è somma di V_1 e V_1^\perp . Questa somma poi è diretta giacché $V_1 \cap V_1^\perp$ è un sottospazio di V_1 , che non può coincidere con V_1 stesso (perché $\langle v_1, v_1 \rangle \neq 0$) e quindi deve consistere del solo elemento O , perché V_1 ha dimensione 1. Poiché $\dim V_1^\perp < \dim V$, possiamo ripetere l'intero procedimento riferendoci allo spazio V_1 , cosa possibile per l'ipotesi di induzione. In altre parole, troviamo una base ortogonale di V_1 , sia ad esempio $\{v_2, \dots, v_n\}$. È immediato concludere allora che $\{v_1, \dots, v_n\}$ è una base ortogonale di V .

Esempio 1. Nello spazio \mathbb{R}^2 , siano $X = (x_1, x_2)$ e $Y = (y_1, y_2)$. Si definisca come segue il loro prodotto scalare

$$\langle X, Y \rangle = x_1 y_1 - x_2 y_2.$$

Si vede allora che i vettori $(1, 0)$ e $(0, 1)$ costituiscono una base ortogonale anche rispetto al prodotto ora definito. Invece, i vettori $(1, 2)$ e $(2, 1)$ costituiscono una base ortogonale per il prodotto ora definito, ma non costituiscono una base ortogonale rispetto all'ordinario prodotto scalare.

Esempio 2. Sia V il sottospazio di \mathbb{R}^3 generato dai due vettori $A = (1, 2, 1)$ e $B = (1, 1, 1)$. Se $X = (x_1, x_2, x_3)$ e $Y = (y_1, y_2, y_3)$ sono vettori di \mathbb{R}^3 , definiamo loro prodotto scalare il numero

$$\langle X, Y \rangle = x_1 y_1 - x_2 y_2 - x_3 y_3.$$

Vogliamo ora trovare una base ortogonale di V rispetto a questo prodotto. Osserviamo che $\langle A, A \rangle = 1 - 4 - 1 = -4 \neq 0$. Poniamo $v_1 = A$. Il processo di ortogonalizzazione applicato a B dà

per A il coefficiente

$$c = \frac{\langle B, A \rangle}{\langle A, A \rangle} = \frac{1}{2}.$$

Definiamo allora $v_2 = B - \frac{1}{2}A$. Allora $\{v_1, v_2\}$ è una base ortogonale di V rispetto al prodotto scalare sopra definito.

Esercizi

1. Trovare una base ortogonale del sottospazio di \mathbb{R}^3 generato dai vettori dati A, B rispetto al prodotto scalare indicato con $X \cdot Y$.

a) $A = (1, 1, 1)$, $B = (1, -1, 2)$;

$$X \cdot Y = x_1y_1 + 2x_2y_2 + x_3y_3.$$

b) $A = (1, -1, 4)$, $B = (-1, 1, 3)$;

$$X \cdot Y = x_1y_1 - 3x_2y_2 + x_3y_3 - x_3y_2.$$

2. Trovare una base ortogonale dello spazio \mathbb{C}^2 sul corpo complesso, se il prodotto scalare è definito da $X \cdot Y = x_1y_2 - ix_2y_1$.

3. Stessa domanda dell'esercizio 2, quando il prodotto scalare è invece definito da

$$X \cdot Y = x_1y_2 + 4x_2y_1.$$

37. SPAZIO DUALE

Sia V uno spazio vettoriale su un corpo K . Indichiamo con V^* l'insieme di tutte le applicazioni lineari di V in K (considerato come uno spazio vettoriale di dimensione 1 su sé stesso). Noi già sappiamo che V^* stesso è uno spazio vettoriale su K , abbiamo già visto infatti che le applicazioni lineari possono essere addizionate fra loro e moltiplicate per scalari. Gli elementi di V^* sono chiamati *funzionali* (su V) e V^* è chiamato lo *spazio duale* (di V).

Sia φ un elemento di V^* e sia v un elemento di V . È conveniente indicare $\varphi(v)$ col simbolo $\langle \varphi, v \rangle$. L'utilità di questa notazione appare evidente considerando che, se φ_1, φ_2 sono in V^* , allora $(\varphi_1 + \varphi_2)(v) = \varphi_1(v) + \varphi_2(v)$ e se $c \in K$, allora

$$(c\varphi)(v) = c\varphi(v).$$

In altre parole,

$$\langle \varphi_1 + \varphi_2, v \rangle = \langle \varphi_1, v \rangle + \langle \varphi_2, v \rangle,$$

$$\langle c\varphi, v \rangle = c\langle \varphi, v \rangle.$$

Inoltre, se v_1 e v_2 sono in V , allora

$$\langle \varphi, v_1 + v_2 \rangle = \langle \varphi, v_1 \rangle + \langle \varphi, v_2 \rangle,$$

$$\langle \varphi, cv \rangle = c\langle \varphi, v \rangle.$$

Queste due ultime proprietà dicono, in altre parole, che φ è lineare. Abbiamo quindi una perfetta analogia col formalismo adoperato per i prodotti scalari, con l'unica differenza che nel simbolo $\langle \varphi, v \rangle$ i due argomenti non appartengono allo stesso spazio.

Esempio 1. Sia $V = K^n$. Sia $\varphi: K^n \rightarrow K$ la proiezione nel primo fattore, cioè

$$\varphi(x_1, \dots, x_n) = x_1.$$

Allora φ è un funzionale. Analogamente, per ogni $i = 1, \dots, n$ abbiamo il funzionale φ_i tale che

$$\varphi_i(x_1, \dots, x_n) = x_i.$$

Esempio 2. Sia V uno spazio vettoriale sul corpo K , in V sia definito un prodotto scalare. Sia v_0 un elemento di V , allora l'applicazione definita da

$$v \mapsto \langle v, v_0 \rangle, \quad v \in V$$

è un funzionale, come immediatamente segue dalla definizione di prodotto scalare.

Esempio 3. Sia V lo spazio vettoriale delle funzioni continue reali definite nell'intervallo $[0, 1]$. Possiamo definire un funzionale su V mediante l'uguaglianza

$$L(f) = \int_0^1 f(t) dt$$

per ogni $f \in V$. Note proprietà dell'integrale provano che si tratta di un'applicazione lineare. Se f_0 è un fissato elemento di V , allora l'applicazione definita da

$$f \mapsto \int_0^1 f_0(t) f(t) dt$$

è un altro funzionale su V .

Esempio 4. Sia V come nell'esempio 3. Sia $\delta: V \rightarrow \mathbb{R}$ l'applicazione tale che $\delta(f) = f(0)$. Allora δ è un funzionale su V , chiamato *funzionale di Dirac*.

Sia V uno spazio vettoriale sul corpo complesso e si supponga che in V sia definito un prodotto hermitiano. Sia v_0 un elemento di V , allora l'applicazione definita da

$$v \mapsto \langle v, v_0 \rangle, \quad v \in V$$

è un funzionale. Tuttavia, non è vero che l'applicazione definita da $v \mapsto \langle v_0, v \rangle$ sia un funzionale. Per ogni $\alpha \in \mathbb{C}$, infatti, abbiamo

$$\langle v_0, \alpha v \rangle = \bar{\alpha} \langle v_0, v \rangle.$$

quindi quest'ultima applicazione *non* è lineare. Essa è talvolta chiamata *antilineare* oppure *semilineare*.

TEOREMA 6 *Sia V uno spazio vettoriale di dimensione finita sul corpo K . Allora lo spazio duale V^* ha dimensione finita e si ha $\dim V = \dim V^*$.*

Dimostrazione. Sia $\{v_1, \dots, v_n\}$ una base di V . Vogliamo trovare una base di V^* . Per il teorema 1, paragrafo 17 (cap. 4), per ogni $i = 1, \dots, n$ esiste un funzionale che denoteremo con v_i^* tale che

$$\langle v_i^*, v_j \rangle = \begin{cases} 1 & \text{se } i = j, \\ 0 & \text{se } i \neq j. \end{cases}$$

(Il teorema cui ci riferiamo è quello che afferma l'esistenza di un'applicazione lineare quando siano dati i valori di essa sugli elementi di una base.) Vogliamo provare che $\{v_1^*, \dots, v_n^*\}$ è una base di V^* .

Sia $\varphi \in V^*$. Sia $c_i = \langle \varphi, v_i \rangle$. Affermiamo che

$$\varphi = c_1 v_1^* + \dots + c_n v_n^*.$$

Per ogni i , infatti, abbiamo

$$\langle c_1 v_1^* + \dots + c_n v_n^*, v_i \rangle = c_i \langle v_i^*, v_i \rangle = c_i.$$

E poiché $c_i = \varphi(v_i)$, segue che φ e $c_1 v_1^* + \dots + c_n v_n^*$ assumono gli stessi valori su ciascun elemento della base $\{v_1, \dots, v_n\}$. Quindi esse assumono lo stesso valore su ciascuna combinazione lineare

di questi elementi della base e perciò coincidono. Ne segue che v_1^*, \dots, v_n^* generano V^* .

Per vedere che questi funzionali sono linearmente indipendenti, supponiamo che

$$x_1 v_1^* + \dots + x_n v_n^* = 0,$$

per opportuni elementi x_i appartenenti a K .

Calcolando il valore di questa applicazione per il vettore v_i , troviamo

$$0 = \langle x_1 v_1^* + \dots + x_n v_n^*, v_i \rangle = x_i \langle v_i^*, v_i \rangle = x_i.$$

Quindi ogni x_i deve essere nullo e questo prova quanto volevamo.

La base $\{v_1^*, \dots, v_n^*\}$ di V^* definita nella dimostrazione del teorema precedente si chiama *base duale* di $\{v_1, \dots, v_n\}$.

Adopereremo anche in questo caso la terminologia relativa alla nozione di perpendicolarità già usata per il prodotto scalare. Quindi, se S è un sottoinsieme di V e se $\varphi \in V^*$, diremo che φ è *ortogonale* oppure *perpendicolare* a S , se

$$\varphi(v) = \langle \varphi, v \rangle = 0$$

per ogni $v \in S$. L'insieme degli elementi $\varphi \in V^*$ che sono ortogonali a S è un sottospazio V^* che denoteremo di nuovo con S^\perp (se dal contesto risulta chiaro che non vi è altro prodotto scalare in V che può dar luogo ad ambiguità). Si vede subito che ogni elemento di S^\perp è perpendicolare al sottospazio di V generato da S .

TEOREMA 7 *Sia V uno spazio vettoriale di dimensione finita sul corpo K . Sia W un sottospazio. Allora*

$$\dim W + \dim W^\perp = \dim V.$$

Dimostrazione. Sia $\{w_1, \dots, w_r\}$ una base di W . Adoperando il teorema 7, paragrafo 10 (cap. 2), aggiungiamo a questa base degli elementi fino a ottenere una base $\{w_1, \dots, w_n\}$ dell'intero spazio V . Sia $\{w_1^*, \dots, w_n^*\}$ la sua base duale. Vogliamo dimostrare che $\{w_{r+1}^*, \dots, w_n^*\}$ è una base di W^\perp . Questo, evidentemente, implica il teorema che stiamo dimostrando. Sarà sufficiente dimostrare che gli elementi w_{r+1}^*, \dots, w_n^* generano W^\perp essendo già linearmente indipendenti.

Sia φ un elemento di W^\perp . Esistono allora numeri c_1, \dots, c_n in K tali che

$$\varphi = c_1 w_1^* + \dots + c_n w_n^*.$$

poiché $\varphi \in W^\perp$, per ogni $i = 1, \dots, r$, abbiamo

$$0 = \varphi(w_i) = c_i \langle w_i^*, w_i \rangle = c_i$$

per la definizione di base duale. Quindi

$$\varphi = c_{r+1} w_{r+1}^* + \dots + c_n w_n^*.$$

Questo prova che W^\perp è contenuto nel sottospazio generato dai funzionali w_{r+1}^*, \dots, w_n^* . Viceversa, se $r+1 < j < n$, allora

$$\langle w_j^*, w_i \rangle = 0, \quad i = 1, \dots, r.$$

Quindi il funzionale w_j^* appartiene a W^\perp . Rimane così provato che W^\perp contiene lo spazio generato da w_{r+1}^*, \dots, w_n^* e la dimostrazione è quindi conclusa.

Dai ragionamenti finora fatti si vede chiaramente che tra un prodotto scalare in uno spazio vettoriale e lo spazio duale vi è una stretta relazione, e ne vogliamo ora trattare più approfonditamente.

Sia V uno spazio vettoriale sul corpo K , in V sia definito un prodotto scalare. Ad ogni elemento v di V possiamo associare un funzionale L_v , appartenente allo spazio duale, definito ponendo

$$L_v(w) = \langle v, w \rangle$$

per ogni $w \in V$. Se v_1, v_2 sono elementi di V , allora $L_{v_1+v_2} = L_{v_1} + L_{v_2}$. Se poi c è uno scalare, $L_{cv} = cL_v$. Queste uguaglianze sono, in sostanza, una riformulazione della definizione di prodotto scalare. Possiamo allora dire che l'applicazione definita da

$$v \mapsto L_v$$

è un'applicazione lineare di V nello spazio duale V^* . A questo proposito il seguente teorema è molto importante:

TEOREMA 8 *Sia V uno spazio vettoriale di dimensione finita sul corpo K , in V sia definito un prodotto scalare non degenere. Dato*

un funzionale $L: V \rightarrow K$, esiste un unico elemento $v \in V$ tale che

$$L(w) = \langle v, w \rangle$$

per ogni $w \in V$.

Dimostrazione. Consideriamo l'insieme di tutti i funzionali su V che sono del tipo L_v , per qualche v appartenente a V . Questo insieme è in effetti un sottospazio di V^* perché il funzionale nullo vi appartiene e inoltre valgono le uguaglianze

$$L_{v_1} + L_{v_2} = L_{v_1 + v_2} \quad \text{e} \quad L_{cv} = cL_v.$$

Inoltre, se $\{v_1, \dots, v_n\}$ è una base di V , allora L_{v_1}, \dots, L_{v_n} sono linearmente indipendenti. Infatti, se x_1, \dots, x_n sono elementi di K tali che

$$x_1 L_{v_1} + \dots + x_n L_{v_n} = 0,$$

allora

$$L_{x_1 v_1} + \dots + L_{x_n v_n} = 0,$$

e quindi

$$L_{x_1 v_1 + \dots + x_n v_n} = 0.$$

Osserviamo ora, dato che il prodotto scalare non è degenere, che se $v \in V$ e L_v è il funzionale nullo, allora $v = O$. Quindi

$$x_1 v_1 + \dots + x_n v_n = O,$$

e perciò $x_1 = \dots = x_n = 0$, provando così la nostra asserzione. Concludiamo allora che lo spazio dei funzionali del tipo L_v ($v \in V$) è un sottospazio di V^* , della stessa dimensione di V^* , perciò coincidente con V^* . Il nostro teorema è così dimostrato.

Con riferimento al teorema precedente, diciamo che il vettore v rappresenta il funzionale L , rispetto al prodotto scalare non degenere definito in V .

La dimostrazione del teorema 8 può essere semplificata provando che:

TEOREMA 8' *L'applicazione di V in V^* definita da $v \mapsto L_v$ è un isomorfismo.*

Dimostrazione. Il nucleo dell'applicazione considerata contiene il solo elemento O perché il prodotto scalare non è degenere. Poiché gli spazi V e V^* hanno la stessa dimensione, possiamo

concludere la dimostrazione ricorrendo al corollario del teorema 3, paragrafo 19 (cap. 4).

La dimostrazione più lunga del teorema 8 è stata data perché è analoga al ragionamento che faremo in seguito, considerando una situazione simile a proposito dei numeri complessi.

Esempi. Consideriamo lo spazio $V = K^n$ con il solito prodotto scalare

$$X \cdot Y = x_1 y_1 + \dots + x_n y_n,$$

che sappiamo essere non degenere. Se

$$\varphi: V \rightarrow K$$

è un'applicazione lineare, allora esiste un unico vettore $A \in K^n$ tale che, per ogni $H \in K^n$, si ha

$$\varphi(H) = A \cdot H.$$

Possiamo quindi rappresentare il *funzionale* φ con il vettore A .

Per considerare un altro esempio, sia $V = R^n$ con l'usuale prodotto scalare. Sia $f: R^n \rightarrow R$ una funzione differenziabile. Nel calcolo differenziale la derivata di f nel punto X è definita come un'applicazione lineare $\varphi: R^n \rightarrow R$. Il vettore che rappresenta φ rispetto al prodotto scalare ordinario è chiamato il *gradiente* di f in X ed è indicato con $(\text{grad } f)(X)$ oppure $\nabla f(X)$. Per definizione quindi abbiamo

$$\varphi(H) = \nabla f(X) \cdot H$$

per ogni $H \in R^n$.

Come applicazione dei risultati ottenuti a proposito dello spazio duale, possiamo ottenere un analogo del teorema 3 per un prodotto scalare non degenere qualsiasi, non necessariamente definito positivo.

TEOREMA 9 *Sia V uno spazio vettoriale di dimensione finita sul corpo K , in V sia definito un prodotto scalare non degenere. Se W è un sottospazio di V , se U è il sottospazio di V ortogonale a W , allora $\dim U = \dim V - \dim W$.*

Dimostrazione. Poiché avremo a che fare simultaneamente col prodotto scalare definito in V e con lo spazio duale di V , deno-

teremo con

$$\text{Perp}_V(W)$$

lo spazio degli elementi $v \in V$ tali che $\langle v, w \rangle = 0$ per ogni $w \in W$, mentre con

$$\text{Perp}_{V^*}(W)$$

denoteremo lo spazio degli elementi $\varphi \in V^*$ tali che $\varphi(w) = 0$ per ogni $w \in W$. Per definizione, quindi, un elemento v di V appartiene a $\text{Perp}_V(W)$ se, e solo se, il corrispondente funzionale L_v appartiene a $\text{Perp}_{V^*}(W)$. Perciò l'applicazione definita da $v \mapsto L_v$ induce un isomorfismo tra gli spazi $\text{Perp}_V(W)$ e $\text{Perp}_{V^*}(W)$ che quindi, vengono ad avere la stessa dimensione. La dimostrazione si conclude ricorrendo al teorema 7.

Osservazione. Sottolineiamo il fatto che il teorema 9 è vero nonostante V possa non essere la somma diretta di W e $\text{Perp}_V(W)$. Per esempio, se $V = \mathbb{C}^2$, consideriamo il sottospazio di dimensione 1, W , generato dal vettore $(1, i)$. Vogliamo determinare W^\perp in \mathbb{C}^2 , intendendo la perpendicolarità rispetto all'ordinario prodotto scalare tra vettori. Certamente, allora, W^\perp contiene W poiché $(1, i) \cdot (1, i) = 0$. D'altra parte, per il teorema dimostrato, $\dim W^\perp = \dim \mathbb{C}^2 - \dim W = 1$. Perciò $W^\perp = W$.

Il teorema che abbiamo dimostrato verrà applicato nel prossimo paragrafo, fornendo così illustrazione di altri esempi.

Esercizi

1. Siano A, B due vettori linearmente indipendenti di \mathbb{R}^n . Qual è la dimensione dello spazio perpendicolare a entrambi i vettori A e B ?
2. Siano A, B due vettori linearmente indipendenti in \mathbb{C}^n . Qual è la dimensione del sottospazio di \mathbb{C}^n perpendicolare a entrambi i vettori A, B ? (La perpendicolarità è riferita all'ordinario prodotto scalare tra vettori di \mathbb{C}^n .)
3. Sia V uno spazio vettoriale di dimensione finita sul corpo K . Siano U, W suoi sottospazi e si supponga che V sia somma diretta di U e W . Dimostrare che V^* è uguale alla somma diretta di U^\perp e W^\perp .
4. Sia W il sottospazio di \mathbb{C}^3 generato dal vettore $(1, i, 0)$. Determinare una base di W^\perp in \mathbb{C}^3 (rispetto all'ordinario prodotto scalare tra vettori).
5. Sia V uno spazio vettoriale di dimensione n sul corpo K . Sia φ un funzionale su V , e si assuma $\varphi \neq 0$. Qual è la dimensione del nucleo di φ ? Come si dimostra?

6. Sia V uno spazio vettoriale di dimensione n sul corpo K . Siano ψ e φ due funzionali non nulli su V . Si supponga che non esista alcuno scalare c non nullo e tale che $\psi = c\varphi$. Dimostrare allora che lo spazio

$$(\text{nucleo } \varphi) \cap (\text{nucleo } \psi)$$

ha dimensione $n - 2$.

7. Sia V uno spazio di dimensione n sul corpo K . Sia V^{**} lo spazio duale di V^* . Dimostrare che ad ogni elemento $v \in V$ si può associare un elemento $\lambda_v \in V^{**}$ in modo che l'applicazione definita da $v \mapsto \lambda_v$ sia un isomorfismo tra V e V^{**} .

8. Sia V uno spazio vettoriale di dimensione finita sul corpo K . In V sia definito un prodotto scalare non degenere. Dimostrare che, se W è un sottospazio di V , allora $W^{\perp\perp} = W$.

9. Dimostrare che la conclusione del precedente esercizio è valida anche quando W^\perp indica il complemento ortogonale di W nello spazio duale V^* .

10. Sia V uno spazio vettoriale di dimensione finita sul corpo K , ne siano W_1 e W_2 sottospazi. Esprimere $(W_1 + W_2)^\perp$ mediante W_1^\perp e W_2^\perp . Esprimere poi, sempre mediante W_1^\perp e W_2^\perp , lo spazio $(W_1 \cap W_2)^\perp$.

38. CARATTERISTICA DI UNA MATRICE E SISTEMI DI EQUAZIONI LINEARI

Sia $A = (a_{ij})$ una matrice $m \times n$ in un corpo K . Le colonne A^1, \dots, A^n della matrice A generano un sottospazio di K^m la cui dimensione è chiamata la *caratteristica per colonne* di A . Analogamente, le righe A_1, \dots, A_m di A generano un sottospazio di K^n la cui dimensione è chiamata *caratteristica per righe* di A . Possiamo anche dire che la caratteristica per colonne di A è il numero massimo di colonne linearmente indipendenti e che la caratteristica per righe è il numero massimo di righe linearmente indipendenti.

Ricordiamo che abbiamo considerato un sistema di equazioni lineari

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= 0 \\ &\dots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= 0 \end{aligned} \tag{1}$$

sotto due aspetti. In un caso abbiamo considerato una sua soluzione $X = (x_1, \dots, x_n)$ come una relazione lineare tra i vettori colonna:

$$x_1 A^1 + \dots + x_n A^n = O.$$

Una seconda volta, la soluzione X è stata interpretata come un vettore ortogonale a tutti i vettori riga (e quindi ortogonale allo spazio generato dai vettori riga):

$$A_1 \cdot X = 0, \dots, A_m \cdot X = 0.$$

L'insieme delle soluzioni del sistema è quindi un sottospazio di K^n .

TEOREMA 10 *Sia $A = (a_{ij})$ una matrice $m \times n$ su un corpo K . Allora la caratteristica per righe e la caratteristica per colonne di A sono uguali a uno stesso numero r . Inoltre lo spazio delle soluzioni del sistema di equazioni lineari [1] ha dimensione $n - r$.*

Dimostrazione. Dimostreremo le due asserzioni contemporaneamente. Consideriamo l'applicazione

$$L: K^n \rightarrow K^m$$

definita da

$$L(X) = x_1 A^1 + \dots + x_n A^n.$$

Evidentemente si tratta di un'applicazione lineare. La sua immagine coincide con lo spazio generato dai vettori colonna della matrice A . Il suo nucleo è, per definizione, lo spazio delle soluzioni del sistema di equazioni lineari prima considerato. Per il teorema 3, paragrafo 19 (cap. 4), otteniamo

caratteristica per colonne +

$$+ \text{dimensione dello spazio delle soluzioni} = n.$$

D'altra parte, interpretando lo spazio delle soluzioni come lo spazio ortogonale ai vettori riga e ricordando che l'ordinario prodotto scalare in K^n non è degenere, possiamo applicare il teorema 9 del paragrafo precedente, ottenendo

caratteristica per righe +

$$+ \text{dimensione dello spazio delle soluzioni} = n.$$

Questo prova entrambe le nostre asserzioni.

Le nostre equazioni omogenee sono suscettibili ancora di una nuova interpretazione. Sia $A = (a_{ij})$ la matrice dei coefficienti, consideriamo X come un vettore colonna. Allora lo spazio delle soluzioni è costituito da quei vettori X che sono nel nucleo del-

l'applicazione lineare

$$F: K^n \rightarrow K^m$$

tale che $F(X) = AX$ (moltiplicazione tra matrici). Quindi lo spazio delle soluzioni del sistema omogeneo considerato può essere interpretato anche come il nucleo dell'applicazione lineare F .

Passiamo ora a discutere le equazioni non omogenee. Sia

$$B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

un vettore colonna di K^m . Consideriamo il sistema di equazioni lineari

$$\begin{aligned} A_1 \cdot X &= b_1, \\ &\vdots \\ A_m \cdot X &= b_m, \end{aligned} \tag{2}$$

che interpreteremo ancora in un altro modo. Se consideriamo l'applicazione lineare $F: K^n \rightarrow K^m$ definita da

$$F(X) = \begin{pmatrix} A_1 \cdot X \\ \vdots \\ A_m \cdot X \end{pmatrix} = AX,$$

lo spazio delle soluzioni del sistema [1] coincide col nucleo di questa applicazione lineare. L'insieme delle soluzioni del sistema non omogeneo [2] coincide con l'insieme degli elementi X di K^n tali che $F(X) = B$. È un semplice esercizio (vedi esercizio 5, § 18, cap. 4) dimostrare che se X_0 è una soluzione del sistema [2], cioè se

$$F(X_0) = B,$$

allora ogni altra soluzione è del tipo $X_0 + Y$, dove $Y \in$ nucleo F . Naturalmente, è possibile che il sistema non omogeneo [2] non abbia soluzioni, può accadere cioè che le equazioni del sistema siano incompatibili tra loro. Per esempio, il sistema

$$2x + 3y - z = 1,$$

$$2x + 3y - z = 2$$

non ha soluzioni. Comunque, se almeno una soluzione del si-

stema [2] esiste, noi possiamo di nuovo parlare della *dimensione* dell'insieme delle soluzioni. Per definizione essa è la dimensione dello spazio delle soluzioni del sistema *omogeneo* [1].

Esempio 1. Determinare la caratteristica della matrice

$$\begin{pmatrix} 1 & 5 & -7 \\ 2 & 3 & 1 \end{pmatrix}$$

sul corpo dei numeri reali.

Poiché vi sono due righe, la caratteristica non supera 2. D'altra parte, le prime due colonne sono linearmente indipendenti, perché il determinante

$$\begin{vmatrix} 1 & 5 \\ 2 & 3 \end{vmatrix} = 3 - 10$$

è diverso da 0. Concludiamo perciò che la caratteristica cercata è 2.

Osserviamo che la caratteristica rimane la stessa anche se consideriamo la matrice dell'esempio 1 sul corpo dei numeri complessi. Possiamo infatti determinare la caratteristica adoperando soltanto i determinanti che si calcolano mediante una stessa formula in cui intervengono soltanto le componenti della matrice. Questo fatto è molto utile.

Esempio 2. Determinare la dimensione dell'insieme delle soluzioni del seguente sistema di equazioni e determinare questo insieme in \mathbb{R}^3 :

$$2x + y + z = 1,$$

$$y - z = 0.$$

È immediatamente chiaro che almeno una soluzione esiste, cioè $x = \frac{1}{2}$, $y = z = 0$. La caratteristica della matrice

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

è uguale a 2. Perciò la dimensione dell'insieme delle soluzioni è 1. Lo spazio vettoriale delle soluzioni del sistema omogeneo ha dimensione 1 e una di queste soluzioni, come si verifica facil-

mente, è

$$y = z = 1 \quad \text{e} \quad x = -\frac{1}{2}.$$

Si può perciò concludere che l'insieme di tutte le soluzioni del sistema dato coincide con l'insieme di tutti i vettori

$$\left(\frac{1}{2}, 0, 0\right) + t\left(-\frac{1}{2}, 1, 1\right),$$

dove t assume tutti i valori reali. Osserviamo che il nostro insieme di soluzioni è una retta (vedi cap. 1).

Nello schema seguente riassumiamo le tre interpretazioni che abbiamo date del sistema di equazioni lineari omogenee la cui matrice dei coefficienti sia A .

a) *Lo spazio delle soluzioni è costituito da quei vettori X che danno relazioni lineari*

$$x_1 A^1 + \dots + x_n A^n = O$$

tra le colonne di A .

b) *Le soluzioni X costituiscono lo spazio ortogonale ai vettori riga di A .*

c) *Le soluzioni costituiscono il nucleo dell'applicazione lineare rappresentata da A , cioè sono le soluzioni dell'equazione $AX = O$.*

TEOREMA 11 *Sia A una matrice $n \times n$. Allora, le proposizioni seguenti sono equivalenti:*

- a) *Le colonne di A sono linearmente indipendenti.*
- b) *Le righe di A sono linearmente indipendenti.*
- c) *La matrice A è invertibile.*

Dimostrazione. Poiché la caratteristica per righe di A coincide con la sua caratteristica per colonne, le prime due condizioni sono equivalenti. Nei riguardi della terza, si osservi che se A è invertibile, l'unica soluzione dell'equazione $AX = O$ è $X = O$ perché, se B è un'inversa di A , dall'uguaglianza $AX = O$ segue che $X = BAX = BO = O$. Quindi lo spazio delle soluzioni $AX = O$ ha dimensione 0, cioè le colonne di A sono linearmente indipendenti.

La dimostrazione dell'implicazione inversa è lasciata al lettore.

Esercizi

1. Trovare la caratteristica delle matrici seguenti sul corpo dei numeri reali:

a) $\begin{pmatrix} 2 & 1 & 3 \\ 7 & 2 & 0 \end{pmatrix}$. b) $\begin{pmatrix} -1 & 2 & -2 \\ 3 & 4 & -5 \end{pmatrix}$. c) $\begin{pmatrix} 1 & 2 & 7 \\ 2 & 4 & -1 \end{pmatrix}$.

d) $\begin{pmatrix} 1 & 2 & -3 \\ -1 & -2 & 3 \\ 4 & 8 & -12 \\ 0 & 0 & 0 \end{pmatrix}$.

2. Siano A, B due matrici che possono essere moltiplicate fra loro. Dimostrare che la caratteristica della matrice AB non supera né la caratteristica di A né la caratteristica di B .

3. Sia A una matrice triangolare su un corpo K :

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ 0 & a_{22} & \vdots \\ \vdots & \ddots & \ddots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix},$$

e si supponga che nessuno degli elementi diagonali sia nullo. Qual è la caratteristica di A ?

4. Trovare la dimensione sul corpo reale dello spazio delle soluzioni dei seguenti sistemi di equazioni lineari. Di ognuno di questi spazi di soluzioni trovare anche una base.

a) $2x + y - z = 0,$
 $y + z = 0.$

c) $4x + 7y - \pi z = 0,$
 $2x - y + z = 0.$

b) $x - y + z = 0.$

$x - y = 0,$
 $y + z = 0.$

5. Determinare la dimensione sul corpo reale dello spazio delle soluzioni dei seguenti sistemi di equazioni lineari:

a) $2x - 3y + z = 0,$
 $x + y - z = 0.$

b) $2x + 7y = 0,$
 $x - 2y + z = 0.$

c) $2x - 3y + z = 0,$
 $x + y - z = 0,$
 $3x + 4y = 0,$
 $5x + y + z = 0.$

d) $x + y + z = 0,$
 $2x + 2y + 2z = 0.$

6. Sia A un vettore non nullo dello spazio \mathbb{R}^n . Sia P un punto dell' n -spazio. Qual è la dimensione dell'insieme delle soluzioni dell'equazione seguente

$$X \cdot A = P \cdot A ?$$

7. Determinare la dimensione sul corpo C dello spazio delle soluzioni dei seguenti sistemi di equazioni lineari. Di ognuno di questi spazi di soluzioni determinare anche una base.

a) $ix + y - z = 0,$
 $iy + z = 0.$

b) $(1+i)x - iy + 2z = 0,$
 $ix + 2iy - iz = 0.$

8. Determinare la dimensione dell'insieme delle soluzioni in \mathbb{R}^3 dei seguenti sistemi di equazioni. Per ogni sistema determinare una base dello spazio delle soluzioni del corrispondente sistema omogeneo e determinare una soluzione, se esiste, del sistema non omogeneo.

a) $2x + 3y - z = 1.$

b) $2x - y + z = 0,$
 $2x + y + z = 0.$

c) $-x + 4y + z = 2,$
 $3x + y - z = 0.$

d) $x - y + z = 1,$
 $2x - 3y + z = 0,$
 $x + y - z = 5.$

9. Dimostrare che la caratteristica di una matrice non cambia se sulla matrice stessa vengono effettuate le operazioni seguenti: scambio di due righe della matrice, addizione di un multiplo scalare di una riga a un'altra riga, stesse operazioni sulle colonne invece che sulle righe.

10. Prendendo spunto da un'osservazione fatta alla fine del capitolo, provare in generale che: se una matrice A ha coefficienti reali, allora la sua caratteristica è la stessa, tanto se A viene considerata come matrice in \mathbb{R} quanto se viene considerata come matrice in C .

Capitolo 8

Matrici e applicazioni bilineari

39. FORME BILINEARI

Sia K un corpo e siano V, W spazi vettoriali su K . Un'applicazione $g: V \times W \rightarrow K$ viene detta *bilineare* se essa ha le proprietà seguenti:

BI 1. *Comunque si prendano i vettori v_1, v_2 in V e w in W , si ha*

$$g(v_1 + v_2, w) = g(v_1, w) + g(v_2, w)$$

e comunque si prendano i vettori v in V e w_1, w_2 in W , si ha

$$g(v, w_1 + w_2) = g(v, w_1) + g(v, w_2).$$

BI 2. *Comunque si scelgano c in K , v in V , w in W , si ha*

$$g(cv, w) = cg(v, w) = g(v, cw).$$

Possiamo quindi dire che un'applicazione bilineare è un'applicazione tale che, per ogni $v \in V$, l'applicazione definita da $w \mapsto g(v, w)$ è lineare e per ogni $w \in W$, l'applicazione definita da $v \mapsto g(v, w)$ è lineare. Se non vi è necessità di fare riferimento esplicito all'applicazione g , solitamente scriveremo

$$g(v, w) = \langle v, w \rangle.$$

Le proprietà sopradette sono quindi simili alle analoghe proprietà del prodotto scalare.

Nel caso in cui gli spazi V e W coincidono, e in questo caso g

è un'applicazione di $V \times V$ in K , diremo che g è una *forma bilineare* su V . In effetti sono queste le applicazioni che considereremo in tutto questo capitolo.

Se $g, g': V \times V \rightarrow K$ sono forme bilineari su V , possiamo considerare la loro somma $g + g'$. Si tratta di nuovo di una forma bilineare. Analogamente, se $c \in K$, allora cg è una forma bilineare. Possiamo quindi dire che l'insieme delle forme bilineari su V è uno spazio vettoriale su K .

Esempio 1. Sia $C = (c_{ij})$ una matrice $n \times n$ in K . Possiamo associare a C un'applicazione bilineare $K^n \times K^n \rightarrow K$ come segue: se ${}^t X = (x_1, \dots, x_n)$ e ${}^t Y = (y_1, \dots, y_n)$ sono due vettori (consideriamo X, Y come vettori colonna in K^n), associamo ad essi la matrice $1 \times 1 {}^t X C Y$, considerata come un elemento di K . In altre parole consideriamo l'applicazione

$$(X, Y) \mapsto {}^t X C Y.$$

La bilinearità di questa applicazione è una diretta conseguenza delle proprietà della moltiplicazione tra matrici. Possiamo anche scrivere, esplicitando la notazione matriciale:

$${}^t X C Y = (x_1, \dots, x_n) \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \dots & c_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

oppure, esprimendo tutto mediante le coordinate e le somme,

$${}^t X C Y = \sum_{i,j=1}^n x_i c_{ij} y_j = \sum_{i,j=1}^n c_{ij} x_i y_j.$$

L'applicazione bilineare su K^n associata alla matrice C può essere denotata g_C .

Se C, C' sono due matrici $n \times n$, allora $g_{C+C'} = g_C + g_{C'}$. Se $b \in K$, allora $g_{bC} = bg_C$. Questo significa che l'applicazione

$$C \mapsto g_C$$

è un'applicazione bilineare definita sullo spazio vettoriale delle matrici $n \times n$ in K nello spazio vettoriale delle forme bilineari su K^n . Questa è la riformulazione di alcune proprietà della moltiplicazione tra matrici.

plicazione tra matrici, precisamente

$${}^t X(C + C')Y = {}^t XCY + {}^t XC'Y,$$

$${}^t X(bC)Y = b {}^t XCY.$$

La forma bilineare g_0 è la forma nulla se, e soltanto se, C è la matrice nulla. Infatti, se $g_0 = 0$ allora ${}^t XCY = 0$ per tutti i vettori X, Y di K^n . Deve quindi essere necessariamente $CY = 0$ per ogni Y in K^n , altrimenti se $CY \neq 0$, deve esistere un vettore X in K^n tale che

$${}^t XCY \neq 0$$

(a causa della non degenerazione dell'ordinario prodotto scalare tra n -uple). Per la stessa ragione allora, deve essere $C = O$.

In particolare, se C, C' sono due matrici $n \times n$ tali che $g_0 = g_{0'}$, allora $C = C'$. Infatti, in tal caso abbiamo

$${}^t XCY = {}^t XC'Y$$

per ogni coppia di vettori X, Y appartenenti a K^n , segue allora ${}^t XCY - {}^t XC'Y = 0$ e quindi

$${}^t X(C - C')Y = 0$$

per ogni coppia di vettori X, Y appartenenti a K^n . Conseguentemente $C - C' = O$ e infine $C = C'$.

La matrice C dell'esempio 1 viene detta la matrice che *rappresenta* la forma bilineare.

Una forma bilineare, denotata con $\langle \cdot, \cdot \rangle$, dà quindi luogo a un certo tipo di prodotto tra elementi di V . Spesso sarà utile considerare tali prodotti per fattori che sono somme. Siano v_1, \dots, v_n e w_1, \dots, w_m elementi di V . Possiamo allora sviluppare il prodotto

$$\langle v_1 + \dots + v_n, w_1, \dots, w_m \rangle =$$

$$= \langle v_1, w_1 + \dots + w_m \rangle + \dots + \langle v_n, w_1 + \dots + w_m \rangle =$$

$$= \langle v_1, w_1 \rangle + \dots + \langle v_1, w_m \rangle + \dots + \langle v_n, w_1 \rangle + \dots + \langle v_n, w_m \rangle.$$

In questa somma ogni addendo è del tipo $\langle v_i, w_j \rangle$ con $i = 1, \dots, n$ e $j = 1, \dots, m$. Questa somma può essere, quindi, così abbreviata

$$\sum_{i=1}^n \sum_{j=1}^m \langle v_i, w_j \rangle,$$

o abbreviata ancora di più come segue

$$\sum_{i,j} \langle v_i, w_j \rangle,$$

se il riferimento agli indici i e j è chiaro dal contesto.

Per esempio

$$\langle v + w, v + w \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle.$$

Può accadere che $\langle v, w \rangle$ sia diverso da $\langle w, v \rangle$ e quindi, in generale, non si può ulteriormente semplificare.

Una forma bilineare $g: V \times V \rightarrow K$ si dice *simmetrica* quando

$$g(v, w) = g(w, v)$$

per ogni coppia di elementi v, w appartenenti a V . Quindi una forma bilineare simmetrica è nient'altro che un prodotto scalare così come è stato definito nel capitolo 7.

TEOREMA 1 Una matrice $n \times n$ C in K rappresenta una forma bilineare simmetrica se, e solo se, essa è una matrice simmetrica.

Dimostrazione. Si assuma C simmetrica, cioè $C = {}^t C$. Poiché, comunque si prendano i vettori X, Y in K^n , la matrice ${}^t X C Y$ è una matrice 1×1 , cioè un elemento di K , essa risulta uguale alla sua trasposta. Perciò

$${}^t X C Y = {}^t ({}^t X C Y) = {}^t Y {}^t C {}^t X = {}^t Y C X.$$

Questo dimostra che C rappresenta una forma bilineare simmetrica.

Viceversa, supponiamo che C rappresenti una forma bilineare simmetrica, supponiamo cioè che ${}^t X C Y = {}^t Y C X$ comunque si prendano X, Y in K^n . Essendo

$${}^t Y C X = {}^t ({}^t Y C X) = {}^t X {}^t C {}^t Y = {}^t X {}^t C Y,$$

otteniamo l'uguaglianza ${}^t X C Y = {}^t X {}^t C Y$ comunque si prendano X, Y in K^n , conseguentemente $C = {}^t C$, cioè la matrice C risulta simmetrica.

Esempio 2. La matrice

$$C = \begin{pmatrix} 1 & -2 & 3 \\ -2 & 1 & 1 \\ 3 & 1 & 4 \end{pmatrix}$$

è una matrice simmetrica. Sia ${}^tX = (x_1, x_2, x_3)$ e ${}^tY = (y_1, y_2, y_3)$. Se indichiamo con g la forma bilineare ad essa associata, abbiamo

$$\begin{aligned} g(X, Y) &= x_1y_1 - 2x_2y_1 + 3x_3y_1 - 2x_1y_2 + x_2y_2 + x_3y_2 + \\ &\quad + 3x_1y_3 + x_2y_3 + 4x_3y_3 = \\ &= g(Y, X). \end{aligned}$$

Sia V uno spazio vettoriale di dimensione finita sul corpo K e sia $g: V \times V \rightarrow K$ una forma bilineare. Sia $\{v_1, \dots, v_n\}$ una base di V . Siano v, w elementi di V , scriviamo questi elementi mediante gli elementi della base:

$$v = x_1v_1 + \dots + x_nv_n,$$

$$w = y_1v_1 + \dots + y_nv_n,$$

allora

$$g(v, w) = \sum_{i,j=1}^n x_i y_j g(v_i, v_j),$$

e ponendo $c_{ij} = g(v_i, v_j)$ si ha

$$g(v, w) = \sum_{i,j=1}^n c_{ij} x_i y_j;$$

quindi la forma bilineare g può essere espressa mediante i vettori delle coordinate di X e Y e la matrice $C = (c_{ij})$ esattamente come nell'esempio 1.

Si supponga ora che la forma bilineare g sia simmetrica.

Se $\{v_1, \dots, v_n\}$ è una base ortogonale, allora $g(v_i, v_j) = 0$ tutte le volte che i e j sono diversi. Conseguentemente la matrice (c_{ij}) è una matrice diagonale, per esempio

$$\begin{pmatrix} c_1 & 0 & \dots & 0 \\ 0 & c_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & c_n \end{pmatrix},$$

e la forma bilineare si dice allora *diagonalizzata*. Il prodotto scalare è allora esprimibile mediante i vettori delle coordinate rispetto alla base ortogonale semplicemente come segue

$$g(v, w) = c_1 x_1 y_1 + \dots + c_n x_n y_n.$$

Considerando il corpo reale, se $\{v_1, \dots, v_n\}$ è una base *ortonormale*, e se inoltre $g(v_i, v_i) = 1$, allora, in questo caso,

$$g(v, w) = x_1 y_1 + \dots + x_n y_n$$

è semplicemente il prodotto scalare.

Concludiamo questo paragrafo con una discussione sul cambiamento di coordinate.

Sia C una matrice quadrata che rappresenta una forma bilineare g . Allora la forma bilineare si può esprimere come

$$g(X, Y) = {}^t X C Y$$

mediante i vettori delle coordinate X e Y . Supponiamo ora di cambiare la base del nostro spazio vettoriale. Esiste allora una matrice non singolare N tale che $X = NX'$ e $Y = NY'$ se X' e Y' sono i vettori delle coordinate rispetto alla nuova base. In questo caso otteniamo

$${}^t X C Y = {}^t (N X') C N Y' = {}^t X' {}^t N C N Y'.$$

Conseguentemente, rispetto ai vettori delle coordinate X' e Y' , la matrice che rappresenta la forma diviene

$$\boxed{{}^t N C N}.$$

Quindi, diversamente da quanto accade nel caso della matrice di un'applicazione lineare (che viene mutata attraverso l'*inversa*) la matrice di una forma bilineare viene mutata tramite la *trasposta* di N . Con riferimento alle notazioni del paragrafo 23 (cap. 5), vediamo che la matrice N è niente altro che

$$\boxed{N = M_{\mathcal{B}}^{\mathcal{B}'}(\text{id})}.$$

Esercizi

1. Sia g , indicata con \langle , \rangle , la forma bilineare su \mathbb{R}^3 la cui matrice associata rispetto alla usuale base è

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

Trovare due vettori di \mathbb{R}^3 , X e Y , tali che $\langle X, Y \rangle \neq \langle Y, X \rangle$.

2. Trovare la matrice associata alla forma g considerata nell'esercizio 1 rispetto alla base $\{(1, 1, 0), (0, 1, 0), (1, 1, 1)\}$.

3. a) Sia V uno spazio vettoriale sul corpo K e si supponga che V sia la somma diretta di due suoi sottospazi W_1, W_2 : $V = W_1 \oplus W_2$. Siano g_1 e g_2 forme bilineari simmetriche su W_1 e W_2 rispettivamente. Dimostrare che su V esiste un'unica forma bilineare g tale che se $v = v_1 + v_2$ e $w = w_1 + w_2$ sono elementi di V , con v_1 e w_1 in W_1 e v_2 e w_2 in W_2 , allora $g(v, w) = g_1(v_1, w_1) + g_2(v_2, w_2)$.

b) Se \mathcal{B}_1 è una base di W_1 e \mathcal{B}_2 è una base di W_2 , quale forma assumerà la matrice della forma bilineare g rispetto alla base $\mathcal{B}_1 \cup \mathcal{B}_2$ dello spazio V ?

4. Sia V lo spazio vettoriale sul corpo reale costituito da tutti i polinomi di grado non superiore a n . Siano f, g elementi di V e sia

$$\langle f, g \rangle = \int_0^1 f(t)g(t) dt.$$

Trovare la matrice di questo prodotto scalare rispetto alla base $\{1, t, \dots, t^n\}$.

5. Sia V lo spazio vettoriale sul corpo reale una cui base è $\{\sin t, \cos t\}$. Si definisca un prodotto scalare ponendo

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(t)g(t) dt.$$

Qual è la matrice di questo prodotto scalare rispetto alla base data?

6. Sia V uno spazio vettoriale di dimensione finita sul corpo K e sia g una forma bilineare su V , indicata con \langle , \rangle .

a) Dimostrare che per ogni $w \in V$ l'applicazione definita da $v \mapsto \langle v, w \rangle$ è un funzionale L_w su V e che l'applicazione definita da $w \mapsto L_w$ è un'applicazione lineare di V nel suo spazio duale V^* .

b) Dimostrare che le seguenti proposizioni sono equivalenti:

- 1) il nucleo dell'applicazione L sopradescritta è $\{O\}$;
- 2) l'applicazione L è un isomorfismo tra V e V^* ;
- 3) se C è la matrice che rappresenta la forma bilineare g rispetto a una base di V , allora $\text{Dct}(C) \neq 0$.

Una forma bilineare per cui valgono le tre proposizioni precedenti viene detta *non degenera*.

7. Generalizzare i risultati di questo paragrafo nel senso seguente.

a) Sia C una matrice $m \times n$ in K . Dimostrare che C dà luogo a un'applicazione bilineare $g: K^m \times K^n$ quando si ponga

$$g(X, Y) = {}^t X C Y, \quad X \in K^m, Y \in K^n.$$

b) Viceversa, siano V, W spazi vettoriali di dimensione finita sul corpo K . Sia $g: V \times W \rightarrow K$ un'applicazione bilineare denotata con $g(v, w) = \langle v, w \rangle$. Dimostrare che ad ogni coppia di basi \mathcal{B} di V e \mathcal{B}' di W possiamo associare una matrice C tale che, se $v \in V$ e $w \in W$ e se $X = M_{\mathcal{B}}(v)$ e $Y = M_{\mathcal{B}'}(w)$, allora

$$g(v, w) = {}^t X C Y.$$

c) Sia $g: V \times W \rightarrow K$ un'applicazione bilineare denotata, come al solito, con \langle , \rangle . Per ogni $w \in W$ dimostrare che l'applicazione di V in K definita da

$$L_w: v \mapsto \langle v, w \rangle$$

è una funzionale su V e che l'applicazione lineare definita da $w \mapsto L_w$ è un'applicazione lineare di W nello spazio duale V^* .

d) Dimostrare che le proposizioni seguenti sono equivalenti:

- 1) l'applicazione $L: W \rightarrow V^*$ è iniettiva (cioè il suo nucleo è $\{O\}$);
- 2) se C è una matrice associata alla forma g come si è detto al capoverso b), allora la caratteristica di C coincide con la dimensione di W .

40. FORME QUADRATICHE

Sia V uno spazio vettoriale di dimensione finita sul corpo K . Sia $g = \langle , \rangle$ una forma bilineare simmetrica su V . Si chiama *forma quadratica* determinata da g la funzione

$$f: V \rightarrow K$$

tale che $f(v) = g(v, v) = \langle v, v \rangle$.

Esempio 1. Se $V = K^n$, allora $f(X) = X \cdot X = x_1^2 + \dots + x_n^2$ è la forma quadratica determinata dal prodotto scalare ordinario.

In generale, se $V = K^n$ e se C è una matrice simmetrica in K che rappresenta una forma bilineare simmetrica, allora la forma quadratica è data, come funzione di X , dall'uguaglianza

$$f(X) = {}^t X C X = \sum_{i,j=1}^n c_{ij} x_i x_j.$$

Se C è una matrice diagonale, per esempio

$$C = \begin{pmatrix} c_1 & 0 & \dots & 0 \\ 0 & c_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & c_n \end{pmatrix},$$

Allora la forma quadratica assume l'espressione più semplice:

$$f(X) = c_1 x_1^2 + \dots + c_n x_n^2.$$

Sia di nuovo V uno spazio vettoriale di dimensione finita sul corpo K . Sia g una forma bilineare simmetrica e sia f la sua forma quadratica. Allora i valori che g assume possono essere ricavati completamente da quelli di f , giacché, per ogni coppia di elementi v, w in V ,

$$\langle v, w \rangle = \frac{1}{4} [\langle v+w, v+w \rangle - \langle v-w, v-w \rangle]$$

oppure, adoperando g e f ,

$$g(v, w) = \frac{1}{4} [f(v+w) - f(v-w)].$$

Inoltre vale anche l'uguaglianza seguente

$$\langle v, w \rangle = \frac{1}{2} [\langle v+w, v+w \rangle - \langle v, v \rangle - \langle w, w \rangle].$$

Le dimostrazioni si conseguono facilmente sviluppando i prodotti e adoperando la bilinearità. Per esempio, nei riguardi della seconda formula, abbiamo

$$\begin{aligned} \langle v+w, v+w \rangle - \langle v, v \rangle - \langle w, w \rangle &= \\ &= \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle - \langle v, v \rangle - \langle w, w \rangle = \\ &= 2\langle v, w \rangle. \end{aligned}$$

Lo sviluppo della prima lo lasciamo come esercizio.

Esempio 2. Sia $V = \mathbb{R}^2$ e denotiamo con $X = (x, y)$ gli elementi di \mathbb{R}^2 . La funzione f definita da

$$f(x, y) = 2x^2 + 3xy + y^2$$

è una forma quadratica. Vogliamo trovare la matrice della sua forma bilineare simmetrica g . Se indichiamo questa matrice con

$$C = \begin{pmatrix} a & b \\ b & d \end{pmatrix},$$

dobbiamo avere

$$f(x, y) = (x, y) \begin{pmatrix} a & b \\ b & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

o, in altre parole,

$$2x^2 + 3xy + y^2 = ax^2 + 2bxy + dy^2.$$

Quindi otteniamo $a = 2$, $2b = 3$, $d = 1$. La matrice in conclusione si scrive

$$C = \begin{pmatrix} 2 & \frac{3}{2} \\ \frac{3}{2} & 1 \end{pmatrix}.$$

Esercizi

1. Sia V uno spazio vettoriale di dimensione finita sul corpo K . Sia $f: V \rightarrow K$ una funzione e si assuma che la funzione g definita da

$$g(v, w) = f(v + w) - f(v) - f(w)$$

sia bilineare. Si assuma inoltre che, per ogni $v \in V$ e ogni $a \in K$, $f(av) = a^2 f(v)$. Dimostrare che f è una forma quadratica e determinare la forma bilineare da cui deriva. Dimostrare che questa forma bilineare è unica.

2. Qual è la matrice associata alla forma quadratica

$$f(X) = x^2 - 3xy + 4y^2$$

se $X = (x, y, z)$?

3. Sia $f: \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione con derivate continue prime e seconde, tale che $f(tX) = t^2 f(X)$ per ogni $X \in \mathbb{R}^n$. Dimostrare che f è una forma quadratica. (Sono necessarie alcune formule del calcolo differenziale delle funzioni di più variabili.)

4. Siano x_1, x_2, x_3, x_4 le coordinate di un vettore X , e y_1, y_2, y_3, y_4 le coordinate di un altro vettore Y . Mediante queste coordinate esprimere le forme bilineari associate a ciascuna delle seguenti forme quadratiche:

- a) $x_1 x_2$. b) $x_1 x_3 + x_4^2$. c) $2x_1 x_2 - x_3 x_4$. d) $x_1^2 - 5x_4 x_3 + x_4^2$.

5. Dimostrare che se f_1 è la forma quadratica dedotta dalla forma bilineare g_1 e se f_2 è la forma quadratica dedotta dalla forma bilineare g_2 , allora alla forma bilineare $g_1 + g_2$ è associata la forma quadratica $f_1 + f_2$.

41. OPERATORI SIMMETRICI

Sia V uno spazio vettoriale di dimensione finita su un corpo K . In tutto questo paragrafo supporremo fissata in V una forma bilineare simmetrica non degenere, denotata con \langle , \rangle . Se il lettore vuole, può riferirsi al caso in cui V coincide con K^n e la forma bilineare data coincide con l'ordinario prodotto scalare

$$\langle X, Y \rangle = {}^t X Y,$$

Un'applicazione lineare dello spazio V in sé stesso sarà chiamata *operatore*. Sia $A: V \rightarrow V$ un operatore. Allora, mediante A , possiamo definire una forma bilineare su V definendo l'applicazione

$$(v, w) \mapsto \langle Av, w \rangle$$

per ogni coppia di elementi v, w in V . La verifica che si tratta effettivamente di un'applicazione bilineare è immediata. Viceversa, ogni forma bilineare può essere così rappresentata:

TEOREMA 2. *Sia V uno spazio vettoriale di dimensione finita sul corpo K , su V sia data una forma bilineare simmetrica non degenera $\langle \cdot, \cdot \rangle$. Sia g una qualsiasi forma bilineare su V . Esistono allora, e sono unici, due operatori A e B definiti in V tali che*

$$g(v, w) = \langle Av, w \rangle = \langle v, Bw \rangle$$

per ogni coppia di elementi v, w in V .

Dimostrazione. Per ogni $w \in V$ l'applicazione

$$L_w: v \mapsto g(v, w)$$

è un'applicazione lineare di V in K , cioè è un funzionale. Nel teorema 8, paragrafo 37 (cap. 7), abbiamo visto che ogni funzionale su V può essere rappresentato univocamente da un elemento di V , in altre parole esiste un *unico* elemento w' di V tale che, per ogni $v \in V$, abbiamo

$$L_w(v) = \langle v, w' \rangle.$$

L'associazione $w \mapsto w'$ è quindi un'applicazione di V in se stesso, e noi la denotiamo con B . Possiamo allora scrivere $w' = Bw$ e abbiamo la formula

$$g(v, w) = \langle v, Bw \rangle$$

per ogni coppia di elementi v, w in V . L'applicazione B è lineare. Per vederlo, si supponga che gli elementi w_1, w_2, w'_1, w'_2 di V siano tali che

$$g(v, w_1) = \langle v, w'_1 \rangle \quad \text{e} \quad g(v, w_2) = \langle v, w'_2 \rangle$$

per ogni $v \in V$. Allora

$$\begin{aligned} g(v, w_1 + w_2) &= g(v, w_1) + g(v, w_2) = \\ &= \langle v, w'_1 \rangle + \langle v, w'_2 \rangle = \\ &= \langle v, w'_1 + w'_2 \rangle. \end{aligned}$$

Da cui

$$B(w_1 + w_2) = Bw_1 + Bw_2,$$

Inoltre, se $c \in K$ abbiamo

$$g(v, cw_1) = cg(v, w_1) = c\langle v, w'_1 \rangle = \langle v, cw'_1 \rangle,$$

e perciò $B(cw_1) = cB(w_1)$. Questo prova che B è un'applicazione lineare. Il fatto che B sia univocamente determinata è dovuto alla esistenza, per ogni w , di un *unico* elemento $w' = Bw$ tale che $g(v, w) = \langle v, Bw \rangle$ per ogni $v \in V$.

Data la simmetria della nostra forma \langle , \rangle possiamo considerare nello stesso modo l'applicazione lineare $w \mapsto g(v, w)$ per ogni $v \in V$ e provare quindi l'esistenza di un unico operatore A tale che

$$g(v, w) = \langle Av, w \rangle$$

per ogni coppia di elementi v, w in V . Questo dimostra il nostro teorema.

Con riferimento alle notazioni del teorema precedente, noi diciamo che l'operatore A *rappresenta* la forma g .

Per definizione, l'operatore B di cui si è parlato nel teorema precedente, viene chiamato il *trasposto* di A e viene indicato con $'A$. L'operatore A viene detto poi *simmetrico* (rispetto alla forma bilineare simmetrica non degenere \langle , \rangle fissata) se $'A = A$.

Per ogni operatore A definito in V , per definizione, abbiamo

$$\langle Av, w \rangle = \langle v, 'Aw \rangle$$

per ogni coppia di elementi v, w in V . Se A è simmetrico, allora $\langle Av, w \rangle = \langle v, Aw \rangle$ e viceversa.

Esempio 1. Sia $V = K^n$ e si consideri in V come forma bilineare simmetrica l'ordinario prodotto scalare. Allora possiamo considerare A come una matrice in K e gli elementi di K^n come vettori colonna X, Y . Il loro prodotto scalare può essere scritto

come prodotto di matrici,

$$\langle X, Y \rangle = {}^t X Y.$$

Abbiamo allora:

$$\langle AX, Y \rangle = {}^t(AX)Y = {}^tX {}^tAY = \langle X, {}^tAY \rangle,$$

dove ora con tA indichiamo la matrice trasposta della matrice A . Quindi, quando noi consideriamo l'ordinario prodotto scalare di n -uple, il trasposto di un operatore è rappresentato dalla trasposta della matrice associata. Questa è la ragione per cui in entrambi i casi abbiamo adoperato per la trasposizione la stessa notazione.

Un operatore simmetrico A dà luogo a una forma quadratica f tale che

$$f(v) = \langle Av, v \rangle,$$

chiamata la *forma quadratica determinata* da A .

Esempio 2. Sia $V = \mathbb{R}^2$ e sia A la matrice

$$A = \begin{pmatrix} 3 & 5 \\ 5 & 13 \end{pmatrix}.$$

Allora la forma quadratica determinata da A (considerata come un'applicazione lineare di \mathbb{R}^2 in sé stesso) è data da

$$f(x, y) = (x, y) \begin{pmatrix} 3 & 5 \\ 5 & 13 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 3x^2 + 10xy + 13y^2.$$

Infine, l'operazione di trasposizione soddisfa le seguenti regole formali:

TEOREMA 3 *Sia V uno spazio vettoriale di dimensione finita sul corpo K , in V sia definita una forma bilineare simmetrica non degenera \langle , \rangle . Siano A, B due operatori definiti in V e sia $c \in K$. Allora*

$${}^t(A + B) = {}^tA + {}^tB, \quad {}^t(AB) = {}^tB {}^tA,$$

$${}^t(cA) = c {}^tA, \quad {}^tA = A.$$

Dimostrazione. Dimostriamo soltanto la seconda formula. Per ogni scelta di v, w in V , abbiamo

$$\langle ABv, w \rangle = \langle Bv, {}^tAw \rangle = \langle v, {}^tB {}^tAw \rangle.$$

Per definizione questo significa che ${}^t(AB) = {}^tB{}^tA$. La dimostrazione delle altre formule è altrettanto semplice.

Esercizi

1. a) Una matrice A è chiamata *semi-simmetrica* se ${}^tA = -A$. Dimostrare che ogni matrice M può essere espressa come somma di una matrice simmetrica e di una matrice semi-simmetrica; queste due matrici sono univocamente determinate. [Si suggerisce di scrivere $A = \frac{1}{2}(M + {}^tM)$.]
 - b) Se A è semi-simmetrica, allora A^2 è simmetrica.
 - c) Sia A una matrice semi-simmetrica. Dimostrare che $\text{Det}(A)$ è uguale a zero se A è una matrice quadrata $n \times n$, con n dispari.
2. Sia A una matrice invertibile simmetrica. Dimostrare che la matrice A^{-1} è anch'essa simmetrica.
3. Dimostrare che una matrice simmetrica triangolare è diagonale.
4. Dimostrare che gli elementi diagonali di una matrice semi-simmetrica sono nulli.
5. Sia V uno spazio vettoriale di dimensione finita sul corpo K , in V sia definito un prodotto scalare non degenere. Siano v_0, w_0 due elementi di V e sia $A: V \rightarrow V$ l'applicazione lineare tale che $A(v) = \langle v_0, v \rangle w_0$. Descrivere l'operatore trasposto tA .
6. Sia V lo spazio vettoriale sul corpo reale costituito da polinomi di grado non superiore a 5. Il prodotto scalare sia definito come al solito da

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt.$$
- Desrivere l'operatore trasposto della derivazione D , rispetto a questo prodotto scalare.
7. Sia V uno spazio vettoriale di dimensione finita sul corpo K , in V sia definito un prodotto scalare non degenere. Sia $A: V \rightarrow V$ un'applicazione lineare. Dimostrare che l'immagine dell'operatore trasposto tA è lo spazio ortogonale al nucleo di A .
8. Sia V uno spazio vettoriale di dimensione finita sul corpo reale R , in V sia dato un prodotto scalare definito positivo. Sia $P: V \rightarrow V$ un'applicazione lineare tale che $PP = P$. Dimostrare che, se ${}^tPP = P{}^tP$, allora $P = {}^tP$.
9. Una matrice reale simmetrica $n \times n$ A viene detta *definita positiva* se, per ogni $X \neq O$, tXAX è positivo. Se A, B sono matrici reali simmetriche quadrate (con la stessa dimensione), definiamo $A < B$ quando $B - A$ è definita positiva. Dimostrare che se $A < B$ e $B < C$, allora $A < C$.

10. Sia V uno spazio vettoriale di dimensione finita sul corpo reale \mathbb{R} : in V sia dato un prodotto scalare definito positivo $\langle \cdot, \cdot \rangle$. Un operatore A definito su V viene detto *positivo* quando, per ogni vettore non nullo $v \in V$, si ha $\langle Av, v \rangle > 0$. Si supponga che V sia somma diretta del sottospazio W e del suo ortogonale W^\perp . Sia P la proiezione su W e si assuma $W \neq \{O\}$. Dimostrare che P è un operatore simmetrico e positivo.

11. Con riferimento alle notazioni introdotte nell'esercizio 10, siano c numero reale e A l'operatore definito da

$$Av = cw.$$

Se possiamo scrivere $v = w + w'$ con $w \in W$ e $w' \in W^\perp$, dimostrare che l'operatore A è simmetrico.

12. Con riferimento alle notazioni introdotte nell'esercizio 10, se P indica di nuovo la proiezione sullo spazio W , dimostrare che esiste un operatore simmetrico A tale che $A^2 = I + P$.

13. Sia $A = (a_{ij})$ una matrice $m \times n$ a elementi reali o complessi. Si definisca una generalizzazione del valore assoluto come segue

$$|A| = mn \cdot \max |a_{ij}|.$$

(Non vi sarà confusione col determinante perché in questo contesto non interviene.) Se A, B sono due matrici delle quali si può fare la somma, dimostrare che

$$|A + B| \leq |A| + |B|.$$

Se le matrici possono essere moltiplicate, dimostrare che

$$|AB| \leq |A||B|.$$

Se c è un numero dimostrare che

$$|cA| = |c||A|.$$

14. Sia A una matrice reale simmetrica. Dimostrare che esiste un numero reale c tale che l'operatore $A + cI$ sia positivo.

15. Sia V uno spazio vettoriale di dimensione finita sul corpo K . In V sia data una forma bilineare non degenere $\langle \cdot, \cdot \rangle$. Se $A: V \rightarrow V$ è un'applicazione lineare tale che

$$\langle Av, Aw \rangle = \langle v, w \rangle$$

per ogni scelta di v, w in V , dimostrare che $\text{Det}(A) = \pm 1$.

16. Sia $Q = (q_{ij})$ ($i, j = 1, \dots, n$) una matrice quadrata di funzioni reali di una variabile reale con derivate di ogni ordine. Sia $Y = {}^t(y_1, \dots, y_n)$ un vettore colonna di funzioni. Sia S_Q l'insieme di tutti i vettori Y tale che

$$Y' = QY.$$

Dimostrare che S_q è uno spazio vettoriale sul corpo reale. Sia $\Phi = (\varphi_1, \dots, \varphi_n)$ un vettore colonna di funzioni e sia

$$\langle \Phi, Y \rangle = \varphi_1 y_1 + \dots + \varphi_n y_n.$$

Indicata con D l'operazione di derivazione, si estenda D a una matrice di funzioni definendo la derivata di una matrice come la matrice ottenuta applicando D ad ogni componente. Dimostrare che

$$D\langle \Phi, Y \rangle = \langle D\Phi, Y \rangle + \langle \Phi, DY \rangle = \langle (D + {}^t Q)\Phi, Y \rangle.$$

Dimostrare, per induzione, che per ogni intero positivo k , si ha

$$D^k \langle \Phi, Y \rangle = \langle (D + {}^t Q)^k \Phi, Y \rangle.$$

17. Sia A una matrice reale simmetrica non nulla. Dimostrare che $\text{tr}(AA) > 0$.

18. Siano A, B matrici simmetriche con le stesse dimensioni sul corpo K . Dimostrare che la matrice AB è simmetrica se, e soltanto se, $AB = BA$.

42. OPERATORI HERMITIANI

Sia V uno spazio vettoriale di dimensione finita sul corpo complesso. Supponiamo che in V sia fissato un prodotto hermitiano definito positivo, denotato con $\langle \cdot, \cdot \rangle$. Un prodotto hermitiano verrà anche chiamato *forma hermitiana*. Se per il lettore è più conveniente, si può assumere $V = \mathbb{C}^n$ e considerare come prodotto hermitiano fissato l'ordinario prodotto

$$\langle X, Y \rangle = {}^t X \bar{Y},$$

dove X, Y sono vettori colonna di \mathbb{C}^n .

Sia $A: V \rightarrow V$ un operatore, cioè un'applicazione lineare di V in sé stesso. Per ogni $w \in V$, l'applicazione

$$L_w: V \rightarrow \mathbb{C}$$

tale che

$$L_w(v) = \langle Av, w \rangle$$

per ogni $v \in V$, è un funzionale.

TEOREMA 4 *Sia V uno spazio vettoriale di dimensione finita su \mathbb{C} , in V sia data una forma hermitiana definita positiva $\langle \cdot, \cdot \rangle$. Considerato un funzionale L definito su V , esiste in V un unico elemento w' tale che, per ogni $v \in V$, $L(v) = \langle v, w' \rangle$.*

Dimostrazione. Il ragionamento è simile a quello fatto nel caso reale, cioè nel teorema 8, paragrafo 37 (cap. 7), e lo lasciamo al lettore.

Dal teorema 4 possiamo concludere che per ogni w esiste un unico w' tale che

$$\langle Av, w \rangle = \langle v, w' \rangle$$

per ogni $v \in V$.

Osservazione. L'associazione $w \mapsto L_w$ non è un isomorfismo di V col suo spazio duale. Infatti, se $\alpha \in \mathbb{C}$, allora $L_{\alpha w} = \bar{\alpha} L_w$. Comunque questo è irrilevante per l'esistenza dell'elemento w' .

L'applicazione di V in sé stesso definita da $w \mapsto w'$ viene denominata con A^* . Per definizione, abbiamo l'uguaglianza

$$\langle Av, w \rangle = \langle v, A^* w \rangle$$

per ogni scelta di v, w in V . L'applicazione A^* è *lineare* e il lettore è consigliato di scrivere tutti i dettagli della dimostrazione di questo fatto. La dimostrazione è simile a quella fatta nel caso delle forme simmetriche e non vi sono operazioni di coniugio che possano infirmare la linearità di A^* . L'operatore A^* viene chiamato *aggiunto* di A . Esso è l'unico operatore per cui sia valida l'ultima uguaglianza scritta.

Esempio. Sia $V = \mathbb{C}^n$ e come forma hermitiana si consideri quella ordinaria definita da

$$(X, Y) \mapsto {}^t X \bar{Y} = \langle X, Y \rangle,$$

dove X, Y sono vettori colonna di \mathbb{C}^n . Allora, per ogni matrice A rappresentante un'applicazione lineare di V in sé stesso, abbiamo

$$\langle AX, Y \rangle = {}^t(AX) \bar{Y} = {}^t X {}^t A \bar{Y} = {}^t X (\bar{{}^t A} \bar{Y}).$$

Per definizione, inoltre, il prodotto $\langle AX, Y \rangle$ è uguale a

$$\langle X, A^* Y \rangle = {}^t X (\bar{A^*} \bar{Y}).$$

E questo significa che

$$A^* = {}^t \bar{A}.$$

Vediamo allora che non sarebbe stato ragionevole usare lo stesso

simbolo t per l'aggiunto di un operatore nel caso complesso come invece abbiamo fatto per l'operatore trasposto nel caso reale.

Un operatore A è chiamato *hermitiano* (oppure *auto-aggiunto*) se $A^* = A$. Questo significa che, comunque si prendano gli elementi v, w in V , abbiamo

$$\langle Av, w \rangle = \langle v, Aw \rangle.$$

Tenendo conto dell'esempio precedente, una matrice quadrata di numeri complessi A è detta *hermitiana* quando ${}^t\bar{A} = A$ oppure, cosa equivalente, ${}^tA = \bar{A}$. Se A è una matrice hermitiana, possiamo definire una forma hermitiana su \mathbb{C}^n ponendo

$$(X, Y) \mapsto {}^t(AX) \bar{Y}.$$

(Si verifichi in tutti i dettagli che questa applicazione è un prodotto hermitiano.)

L'operazione indicata con l'asterisco soddisfa regole analoghe a quelle dell'operazione di trasposizione, cioè:

TEOREMA 5 *Sia V uno spazio vettoriale di dimensione finita su \mathbb{C} , in V sia data una forma hermitiana definita positiva \langle , \rangle . Siano A, B operatori definiti su V e sia $\alpha \in \mathbb{C}$. Allora*

$$\begin{aligned} (A + B)^* &= A^* + B^*, & (AB)^* &= B^*A^* \\ (\alpha A)^* &= \bar{\alpha}A^*, & A^{**} &= A. \end{aligned}$$

Dimostrazione. Proveremo la terza uguaglianza lasciando la dimostrazione delle altre al lettore. Per ogni scelta di v, w in V , abbiamo:

$$\langle \alpha Av, w \rangle = \alpha \langle Av, w \rangle = \alpha \langle v, A^*w \rangle = \langle v, \bar{\alpha}A^*w \rangle.$$

Quest'ultima espressione è anche uguale, per definizione, a

$$\langle v, (\alpha A)^*w \rangle$$

e conseguentemente $(\alpha A)^* = \bar{\alpha}A^*$, come volevamo dimostrare.

Possiamo anche stabilire un analogo del teorema 2 per operatori hermitiani:

TEOREMA 6 *Sia V uno spazio vettoriale di dimensione finita sul corpo complesso, in V sia data una forma hermitiana definita posi-*

tiva $\langle \cdot, \cdot \rangle$. Sia g un'altra forma hermitiana su V . Esiste allora un unico operatore hermitiano A definito su V tale che, comunque si prendano v, w in V , si abbia

$$g(v, w) = \langle Av, w \rangle.$$

Dimostrazione. Il ragionamento è del tutto simile a quello fatto nella dimostrazione del teorema 2 e lo lasciamo al lettore.

Con riferimento alle notazioni del teorema 6, diciamo che l'operatore A rappresenta la forma g .

Vediamo ora quale sia l'analogo delle forme quadratiche per le forme hermitiane. Sia g una forma hermitiana, rappresentata dall'operatore A . Definiamo l'analogo della forma quadratica come la funzione f tale che

$$f(v) = \langle Av, v \rangle.$$

Vale la seguente uguaglianza, detta *identità di polarizzazione*,

$$\boxed{\langle A(v+w), v+w \rangle - \langle A(v-w), v-w \rangle = 2[\langle Aw, v \rangle + \langle Av, w \rangle]}$$

comunque si prendano v, w in V ; oppure anche

$$\boxed{\langle A(v+w), v+w \rangle - \langle Av, v \rangle - \langle Aw, w \rangle = \langle Av, w \rangle + \langle Aw, v \rangle}.$$

La verifica della validità di queste uguaglianze è molto semplice: basta sviluppare i primi membri.

TEOREMA 7 *Sia V come nel teorema precedente. Sia A un operatore tale che, per ogni $v \in V$, $\langle Av, v \rangle = 0$. Allora $A = O$.*

Dimostrazione. Il primo membro della identità di polarizzazione è uguale a zero comunque si prendano in V gli elementi v, w . Quindi otteniamo

$$\langle Aw, v \rangle + \langle Av, w \rangle = 0$$

per ogni scelta di v, w in V . Sostituendo v con iv e ricordando le regole del prodotto hermitiano, possiamo scrivere

$$-i\langle Aw, v \rangle + i\langle Av, w \rangle = 0$$

da cui

$$-\langle Aw, v \rangle + \langle Av, w \rangle = 0.$$

Aggiungendo questa uguaglianza alla prima ottenuta, si ottiene

$$2\langle Av, w \rangle = 0$$

da cui $\langle Av, w \rangle = 0$. Quindi $A = O$, come si voleva dimostrare.

TEOREMA 8 *Sia V come nel teorema precedente. Sia A un operatore. A è hermitiano se, e soltanto se, $\langle Av, v \rangle$ è reale per ogni $v \in V$.*

Dimostrazione. Sia A un operatore hermitiano. Allora

$$\langle Av, v \rangle = \langle v, Av \rangle = \overline{\langle Av, v \rangle}.$$

Poiché un numero complesso che coincide col suo coniugato deve essere reale, concludiamo che $\langle Av, v \rangle$ è reale. Viceversa, si supponga che $\langle Av, v \rangle$ sia reale per ogni $v \in V$. Allora

$$\langle Av, v \rangle = \langle Av, v \rangle = \langle v, Av \rangle = \langle A^*v, v \rangle.$$

E quindi $\langle (A - A^*)v, v \rangle = 0$ per ogni $v \in V$; il teorema 7 permette di concludere allora che $A - A^* = O$, cioè $A = A^*$, come si voleva dimostrare.

Esercizi

1. Sia A una matrice hermitiana invertibile. Dimostrare che anche A^{-1} è hermitiana.
2. Dimostrare che l'analogo del teorema 7, quando V è uno spazio vettoriale di dimensione finita sul corpo reale, è falso. In altre parole, può accadere che Av sia perpendicolare a v per ogni $v \in V$ senza che necessariamente A sia l'applicazione nulla.
3. Dimostrare che l'analogo del teorema 7, quando V è uno spazio vettoriale di dimensione finita sul corpo reale, è vero se si aggiunge l'ipotesi della simmetria dell'operatore A .

4. Dire quali delle seguenti matrici sono hermitiane:

$$\text{a)} \begin{pmatrix} 2 & i \\ -i & 5 \end{pmatrix}. \quad \text{b)} \begin{pmatrix} 1+i & 2 \\ 2 & 5i \end{pmatrix}. \quad \text{c)} \begin{pmatrix} 1 & 1+i & 5 \\ 1-i & 2 & i \\ 5 & -i & 7 \end{pmatrix}.$$

5. Dimostrare che gli elementi diagonali di una matrice hermitiana sono reali.

6. Dimostrare che una matrice hermitiana triangolare è diagonale.

7. Siano A, B matrici hermitiane con le stesse dimensioni. Dimostrare che la matrice $A+B$ è hermitiana. Se inoltre $AB=BA$, dimostrare che la matrice AB è hermitiana.

8. Sia V uno spazio vettoriale di dimensione finita sul corpo complesso C , in V sia data una forma hermitiana definita positiva. Sia $A: V \rightarrow V$ un operatore hermitiano. Dimostrare che $I+iA$ e $I-iA$ sono invertibili. [Suggerimento: se $v \neq 0$, dimostrare che $\|(I+iA)v\| \neq 0$.]

9. Sia A una matrice hermitiana. Dimostrare che le matrici A e \bar{A} sono hermitiane. Se A è invertibile, dimostrare che A^{-1} è hermitiana.

10. Sia V uno spazio vettoriale di dimensione finita sul corpo C , in V sia data una forma hermitiana definita positiva \langle , \rangle . Sia $A: V \rightarrow V$ un'applicazione lineare. Dimostrare che le seguenti proposizioni sono equivalenti:

1) si ha $AA^* = A^*A$;

2) per ogni $v \in V$, $\|Av\| = \|A^*v\|$ (dove $\|v\| = \sqrt{\langle v, v \rangle}$);

3) si può scrivere $A = B + iC$, dove B, C sono hermitiane e $BC = CB$.

11. Sia A una matrice hermitiana non nulla. Dimostrare che $\text{tr}(AA^*) > 0$.

43. OPERATORI UNITARI

Sia V uno spazio vettoriale di dimensione finita sul corpo reale R , in V sia dato un prodotto scalare definito positivo. Sia $A: V \rightarrow V$ un'applicazione lineare. Diremo che A è un'applicazione unitaria reale se

$$\langle Av, Aw \rangle = \langle v, w \rangle$$

per ogni scelta di v, w in V . Possiamo anche dire che A unitaria significa che A conserva il prodotto. In altri libri un'applicazione unitaria reale è anche chiamata applicazione ortogonale. Noi preferiamo adoperare il termine applicazione unitaria per la ragione esposta nel teorema che segue.

TEOREMA 9 *Sia V lo spazio sopra considerato. Sia $A: V \rightarrow V$ un'applicazione lineare. Allora le seguenti proposizioni concernenti A sono equivalenti:*

1) A è un'applicazione unitaria.

2) A conserva la norma dei vettori, cioè per ogni vettore $v \in V$, abbiamo

$$\|Av\| = \|v\|.$$

3) Se $v \in V$ è un vettore di norma 1, anche il vettore Av ha norma 1.

Dimostrazione. Lasciamo al lettore la cura di dimostrare l'equivalenza delle proposizioni 2 e 3. Che 1 implica 2 è ovvio perché il quadrato della norma $\langle Av, Av \rangle$ è un caso speciale di prodotto. Proviamo, viceversa, che 2 implica 1. Intanto abbiamo

$$\langle A(v + w), A(v + w) \rangle - \langle A(v - w), A(v - w) \rangle = 2\langle Av, Aw \rangle.$$

Adoperando l'ipotesi 2 e osservando che il primo membro consiste di quadrati di lunghezze, possiamo scrivere il primo membro della nostra uguaglianza

$$\langle v + w, v + w \rangle - \langle v - w, v - w \rangle$$

che è anche uguale a $2\langle v, w \rangle$. Il nostro teorema allora è immediato.

Il teorema 9 fa vedere perché abbiamo preferito la denominazione "applicazioni unitarie"; queste sono caratterizzate dal fatto che associano vettori di lunghezza 1 e vettori di lunghezza 1.

Naturalmente un'applicazione unitaria U conserva anche la perpendicolarità, cioè, se v, w sono perpendicolari allora anche Uv e Uw lo sono, giacché

$$\langle Uv, Uw \rangle = \langle v, w \rangle = 0.$$

D'altra parte, non è vero che un'applicazione che conservi la perpendicolarità sia necessariamente unitaria. Per esempio, con riferimento al corpo reale, l'applicazione che associa al vettore v il vettore $2v$ conserva la perpendicolarità, ma non è unitaria. Sfortunatamente, è molto comune chiamare ortogonali le applicazioni unitarie reali. Sottolineiamo che queste applicazioni non conservano soltanto la perpendicolarità, ma anche la norma.

TEOREMA 10 *Sia V uno spazio vettoriale di dimensione finita sul corpo reale. In V sia dato un prodotto scalare definito positivo. Un'applicazione lineare $A: V \rightarrow V$ è unitaria se, e soltanto se,*

$$^tAA = I.$$

Dimostrazione. L'operatore A è unitario se, e soltanto se,

$$\langle Av, Aw \rangle = \langle v, w \rangle$$

per ogni scelta di v e w in V . Questa condizione equivale a

$$\langle {}^t A A v, w \rangle = \langle v, w \rangle$$

per ogni v, w in V e perciò equivale anche a ${}^t A A = I$.

Non rimane che interpretare, con il linguaggio delle matrici, la condizione che A sia unitaria. Osserviamo subito che un'applicazione unitaria è invertibile: infatti, se A è unitaria e $A v = O$, allora $v = O$ perché A conserva la norma.

Se nel teorema 10 consideriamo $V = \mathbb{R}^n$, e come prodotto scalare consideriamo l'usuale prodotto tra vettori, allora possiamo rappresentare A con una matrice reale. È naturale quindi definire *unitaria* (oppure *ortogonale*) una matrice reale A tale che ${}^t A A = I_n$, o, cosa equivalente,

$${}^t A = A^{-1}.$$

Esempio. Le uniche applicazioni unitarie del piano \mathbb{R}^2 in sé stesso sono le applicazioni aventi matrici del tipo

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{oppure} \quad \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

Se il determinante di una tale applicazione è uguale a 1, allora la matrice che rappresenta l'applicazione rispetto a una base ortonormale è necessariamente del primo tipo e allora l'applicazione è chiamata una *rotazione*. Tracciando un'appropriata figura si vede immediatamente come questa terminologia sia giustificata. Alcune proprietà relative alle applicazioni unitarie del piano si trovano negli esercizi. Queste proprietà sono semplici da ricavare e forniscono una buona pratica di calcolo che sarebbe inopportuno d'altra parte inserire nel testo. Questi esercizi possono essere considerati come ulteriori esempi dei concetti introdotti in questo paragrafo.

Caso complesso

Come al solito esistono nozioni analoghe nel caso complesso. Sia V uno spazio vettoriale di dimensione finita sul corpo complesso \mathbb{C} , in V sia dato un prodotto hermitiano definito positivo. Sia $A: V \rightarrow V$ un'applicazione lineare. Noi definiamo A *unitaria complessa* se

$$\langle A v, A w \rangle = \langle v, w \rangle$$

per ogni scelta di v, w in V . L'analogo del teorema 9 è vero parola per parola: l'applicazione A è unitaria se, e soltanto se, conserva la norma e se, e soltanto se, associa a vettori di norma 1 vettori di norma 1. Ne lasciamo la dimostrazione per esercizio.

L'analogo del teorema 10, invece, è:

TEOREMA 11 *Sia V uno spazio vettoriale di dimensione finita sul corpo complesso C , in V sia dato un prodotto hermitiano definito positivo. Un'applicazione lineare $A: V \rightarrow V$ è unitaria se, e soltanto se,*

$$A^*A = I.$$

Anche di questo teorema lasciamo la dimostrazione come esercizio.

Considerando $V = C^n$ con la forma hermitiana usuale, definita da

$$\langle X, Y \rangle = x_1\bar{y}_1 + \dots + x_n\bar{y}_n,$$

noi possiamo rappresentare A con una matrice complessa. È quindi naturale definire *unitaria* una matrice complessa A se ${}^t\bar{A}A = I_n$, oppure

$${}^t\bar{A} = A^{-1}.$$

Esercizi

1. a) Sia V uno spazio vettoriale di dimensione finita sul corpo reale R , in V sia dato un prodotto scalare definito positivo. Siano $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_n\}$ due basi ortonormali di V . Sia $A: V \rightarrow V$ un operatore tale che $Av_i = w_i$. Dimostrare che A è unitario reale.

b) Enunciare e dimostrare l'analogo risultato nel caso complesso.

2. Sia V lo spazio introdotto nell'esercizio 1. Sia $\{v_1, \dots, v_n\}$ una base ortonormale di V . Sia A un operatore unitario definito in V . Dimostrare che $\{Av_1, \dots, Av_n\}$ è una base ortonormale di V .

3. Sia A una matrice unitaria reale. a) Dimostrare che la matrice ${}^t\bar{A}$ è unitaria. b) Dimostrare che A^{-1} esiste ed è unitaria. c) Se B è una matrice unitaria reale, dimostrare che anche AB è unitaria e così anche $B^{-1}AB$.

4. Sia A una matrice complessa unitaria. a) Dimostrare che ${}^t\bar{A}$ è unitaria. b) Dimostrare che A^{-1} esiste ed è unitaria. c) Se B è un'altra matrice unitaria complessa, provare che anche le matrici AB e $B^{-1}AB$ lo sono.

5. a) Sia V uno spazio vettoriale di dimensione finita sul corpo reale R , in V sia dato un prodotto scalare definito positivo; siano $\mathcal{B} = \{v_1, \dots, v_n\}$

e $\mathcal{B}' = \{w_1, \dots, w_n\}$ due basi ortonormali di V . Dimostrare che la matrice $M_{\mathcal{B}'}^{\mathcal{B}}(id)$ è unitaria reale. [Si suggerisce di tener conto del fatto che $\langle w_i, w_i \rangle = 1$ e $\langle w_i, w_j \rangle = 0$ se $i \neq j$, come pure della possibilità di esprimere $w_i = \sum a_{ij} v_j$, con opportuni $a_{ij} \in \mathbb{R}$.]

b) Sia $F: V \rightarrow V$ un'applicazione tale che $F(v_i) = w_i$ per ogni i . Dimostrare allora che la matrice $M_{\mathcal{B}'}^{\mathcal{B}}(F)$ è unitaria.

6. Dimostrare che il valore assoluto del determinante di una matrice unitaria reale è uguale a 1. Dedurre che se A è una matrice unitaria reale, allora $\text{Det}(A) = 1$ oppure -1 .

7. Sia A una matrice quadrata complessa, dimostrare allora che $\text{Det}(\bar{A}) = \overline{\text{Det}(A)}$. Se ne deduca che il valore assoluto del determinante di una matrice unitaria complessa è uguale a 1.

8. Sia A una matrice diagonale unitaria reale. Dimostrare che gli elementi diagonali di A sono uguali a 1 oppure a -1 .

9. Sia A una matrice diagonale unitaria complessa. Dimostrare che ogni elemento diagonale ha valore assoluto uguale a 1 e quindi è del tipo $e^{i\theta}$, con θ reale.

Gli esercizi che seguono descrivono varie proprietà delle applicazioni unitarie reali del piano \mathbb{R}^2 .

10. Sia V uno spazio vettoriale di dimensione 2 sul corpo reale \mathbb{R} , in V sia dato un prodotto scalare definito positivo. Sia A un'applicazione unitaria reale di V in sé stesso. Siano $\{v_1, v_2\}$ e $\{w_1, w_2\}$ due basi ortonormali di V tali che $Av_i = w_i$ per $i = 1, 2$. Siano a, b, c, d numeri reali tali che

$$w_1 = av_1 + bv_2,$$

$$w_2 = cv_1 + dv_2.$$

Dimostrare che $a^2 + b^2 = 1$, $c^2 + d^2 = 1$, $ac + bd = 0$, $a^3 = d^3$ e $c^3 = b^3$.

11. Dimostrare che il determinante $ad - bc$ è uguale a 1 oppure a -1 . (Dimostrare che il suo quadrato è uguale a 1.)

12. Si definisca *rotazione* di V un'applicazione unitaria reale A di V il cui determinante sia 1. Dimostrare che la matrice di A relativa a una base ortogonale di V è della forma

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

per opportuni numeri reali a, b tali che $a^2 + b^2 = 1$. Dimostrare anche che, viceversa, ogni applicazione lineare di V in sé stesso rappresentata da una siffatta matrice rispetto a una base ortogonale, è unitaria e ha determinante 1. Con l'ausilio del calcolo infinitesimale, si può allora concludere che esiste un numero θ tale che $a = \cos \theta$ e $b = \sin \theta$.

13. Dimostrare che esiste una matrice unitaria complessa U tale che se

$$A = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad e \quad B = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$$

si ha $U^{-1}AU = B$.

14. Si consideri $V = \mathbb{C}$ come uno spazio vettoriale di dimensione 2 sul corpo reale \mathbb{R} . Se $\alpha \in \mathbb{C}$, sia $L_\alpha : \mathbb{C} \rightarrow \mathbb{C}$ l'applicazione definita da $z \mapsto \alpha z$. Dimostrare che L_α è un'applicazione di V in sé stesso \mathbb{R} -lineare. Per quali numeri complessi α l'applicazione L_α risulta unitaria rispetto al prodotto scalare $\langle z, w \rangle = \operatorname{Re}(z\bar{w})$? Qual è la matrice di L_α rispetto alla base $\{1, i\}$ di \mathbb{C} su \mathbb{R} ?

44. TEOREMA DI SYLVESTER

Sia V uno spazio vettoriale di dimensione finita sul corpo reale, V sia di dimensione non nulla. Sia $\langle \cdot, \cdot \rangle$ un prodotto scalare su V (cioè una forma bilineare simmetrica). Sappiamo che, per il teorema 5 del paragrafo 41, è sempre possibile trovare una base ortogonale. Poiché la nostra forma non è necessariamente definita positiva, può accadere che esista un vettore $v \in V$ tale che $\langle v, v \rangle = 0$ oppure $\langle v, v \rangle = -1$.

Esempio. Sia $V = \mathbb{R}^2$ e vi si consideri la forma rappresentata dalla matrice

$$C = \begin{bmatrix} -1 & +1 \\ +1 & -1 \end{bmatrix}.$$

Allora i vettori

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad e \quad v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

costituiscono una base ortogonale rispetto alla forma data e abbiamo

$$\langle v_1, v_1 \rangle = -1 \quad \text{come pure} \quad \langle v_2, v_2 \rangle = 0.$$

Per esempio, riferendoci alle coordinate rispetto alla base usuale di \mathbb{R}^2 data da $\{e^1, e^2\}$, se $X = (1, -1)$ è il vettore delle coordinate di v_2 , un semplice calcolo direttamente mostra che

$$\langle X, X \rangle = {}^t X C X = 0.$$

Ci proponiamo in questo paragrafo di analizzare la situazione

in generale, considerando uno spazio di dimensione arbitraria.
Sia $\{v_1, \dots, v_n\}$ una base ortogonale di V . Poniamo

$$c_i = \langle v_i, v_i \rangle.$$

Numerando diversamente gli elementi della nostra base, se è necessario, possiamo assumere che la base $\{v_1, \dots, v_n\}$ abbia gli elementi ordinati in modo che

$$c_1, \dots, c_r > 0,$$

$$c_{r+1}, \dots, c_s < 0,$$

$$c_{s+1}, \dots, c_n = 0.$$

Siamo interessati a conoscere il numero degli elementi positivi, negativi, nulli tra i "quadrati" $\langle v_i, v_i \rangle$: in altre parole, siamo interessati ai numeri r e s . Vedremo nel seguito che questi numeri non dipendono dalla scelta della base ortogonale.

Se X è il vettore delle coordinate di un elemento di V rispetto alla nostra base, se f è la forma quadratica associata al nostro prodotto scalare, allora, adoperando le coordinate, potremo scrivere

$$f(X) = c_1 x_1^2 + \dots + c_r x_r^2 + c_{r+1} x_{r+1}^2 + \dots + c_s x_s^2.$$

Vediamo che nell'espressione di f mediante le coordinate vi sono esattamente r termini positivi e $s - r$ termini negativi. Inoltre, $n - s$ termini sono scomparsi.

Tutto questo diviene ancora più chiaro considerando una base normalizzata.

Generalizzeremo la prima nozione già data in base ortonormale. Diremo che una base ortogonale $\{v_1, \dots, v_n\}$ è *ortonormale* se, per ogni i , abbiamo

$$\langle v_i, v_i \rangle = 1 \text{ oppure } \langle v_i, v_i \rangle = -1 \text{ oppure } \langle v_i, v_i \rangle = 0.$$

Se $\{v_1, \dots, v_n\}$ è una base ortogonale, procedendo come nel caso definito positivo possiamo ottenere da essa una base ortonormale. Poniamo $c_i = \langle v_i, v_i \rangle$. Se $c_i = 0$, definiamo $v'_i = v_i$. Se $c_i > 0$, definiamo

$$v'_i = \frac{v_i}{\sqrt{c_i}}.$$

Se infine, $c_i < 0$ definiamo

$$v'_i = \frac{v_i}{\sqrt{-c_i}}.$$

Allora $\{v'_1, \dots, v'_n\}$ è una base ortonormale.

Sia $\{v_1, \dots, v_n\}$ una base ortonormale dello spazio V rispetto al prodotto scalare di V . Se X è il vettore delle coordinate di un elemento di V , allora, con riferimento alla nostra base ortonormale,

$$f(X) = x_1^2 + \dots + x_s^2 - x_{s+1}^2 - \dots - x_n^2.$$

Perciò, rispetto a una base ortonormale, il numero dei termini positivi e il numero dei termini negativi appaiono con particolare chiarezza. Per dimostrare che il numero di questi termini non dipende dalla scelta della base ortonormale, considereremo dapprima il numero dei termini che scompaiono, dandone un'interpretazione geometrica.

TEOREMA 12 *Sia V uno spazio vettoriale di dimensione finita sul corpo reale \mathbb{R} , in V sia dato un prodotto scalare. Si assuma $\dim V > 0$.*

Sia V_0 il sottospazio di V costituito da tutti i vettori $v \in V$ tali che, per ogni $w \in V$, $\langle v, w \rangle = 0$. Sia infine $\{v_1, \dots, v_n\}$ una base ortogonale di V . Allora, il numero degli interi i tali che $\langle v_i, v_i \rangle = 0$ coincide con la dimensione di V_0 .

Dimostrazione. Supponiamo di avere ordinato i termini della base $\{v_1, \dots, v_n\}$ in modo che

$$\langle v_1, v_1 \rangle \neq 0, \dots, \langle v_s, v_s \rangle \neq 0 \quad \text{ma} \quad \langle v_i, v_i \rangle = 0 \quad \text{se} \quad i > s.$$

Poiché $\{v_1, \dots, v_n\}$ è una base ortogonale, è chiaro che v_{s+1}, \dots, v_n sono in V_0 . Sia v un elemento di V_0 e scriviamo

$$v = x_1 v_1 + \dots + x_s v_s + \dots + x_n v_n$$

dove $x_i \in \mathbb{R}$. Considerandone il prodotto scalare con ogni v_j per $j \leq s$, troviamo

$$0 = \langle v, v_j \rangle = x_j \langle v_j, v_j \rangle.$$

E poiché $\langle v_j, v_j \rangle \neq 0$, ne segue che $x_j = 0$. Perciò v appartiene allo spazio generato da v_{s+1}, \dots, v_n . Possiamo quindi concludere che v_{s+1}, \dots, v_n costituiscono una base di V_0 .

(4)

Con riferimento al teorema 12, la dimensione di V_0 è chiamata l'indice di nullità della forma f . Noi vediamo che la forma stessa è non degenere se, e soltanto se, il suo indice di nullità è 0.

TEOREMA DI SYLVESTER *Sia V uno spazio vettoriale di dimensione finita sul corpo reale \mathbb{R} , in V sia dato un prodotto scalare. Esiste allora un intero r non negativo avente la seguente proprietà: se $\{v_1, \dots, v_n\}$ è una base ortogonale di V , allora vi sono esattamente r interi i tali che $\langle v_i, v_i \rangle > 0$.*

Dimostrazione. Siano $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_n\}$ due basi ortogonali di V . Supponiamo di avere ordinato i loro elementi in modo che

$$\begin{array}{lll} \langle v_i, v_i \rangle > 0 & \text{se} & 1 \leq i \leq r \\ \hline \langle v_i, v_i \rangle < 0 & \text{se} & r+1 \leq i \leq s \\ \hline \langle v_i, v_i \rangle = 0 & \text{se} & s+1 \leq i \leq n \end{array}$$

e analogamente,

$$\begin{array}{lll} \langle w_i, w_i \rangle > 0 & \text{se} & 1 \leq i \leq r' \\ \hline \langle w_i, w_i \rangle < 0 & \text{se} & r'+1 \leq i \leq s' \\ \hline \langle w_i, w_i \rangle = 0 & \text{se} & s'+1 \leq i \leq n. \end{array}$$

Proviamo dapprima che i vettori

$$v_1, \dots, v_r, w_{r'+1}, \dots, w_n$$

sono linearmente indipendenti.

Supponiamo di avere la relazione lineare

$$x_1 v_1 + \dots + x_r v_r + y_{r'+1} w_{r'+1} + \dots + y_n w_n = 0.$$

Allora

$$x_1 v_1 + \dots + x_r v_r = -(y_{r'+1} w_{r'+1} + \dots + y_n w_n).$$

Poniamo $c_i = \langle v_i, v_i \rangle$ e $d_i = \langle w_i, w_i \rangle$ per ogni i . Nell'ultima uguaglianza ottenuta, moltiplichiamo scalarmente ogni membro per sé stesso, ottenendo

$$c_1 x_1^2 + \dots + c_r x_r^2 = d_{r'+1} y_{r'+1}^2 + \dots + d_s y_s^2.$$

Il primo membro non è negativo, il secondo membro non è positivo: è quindi necessario che entrambi siano nulli e questo accade soltanto quando

$$x_1 = \dots = x_r = 0.$$

Poiché i vettori w_{r+1}, \dots, w_n sono linearmente indipendenti, ne segue che anche i coefficienti y_{r+1}, \dots, y_n sono nulli.

Poiché la dimensione dello spazio V è n , noi possiamo concludere che

$$r + n - r' \leq n$$

oppure, in altre parole, che r non supera r' . Ma le due basi hanno un ruolo simmetrico (cioè si può ripetere tutto il ragionamento fatto partendo dai vettori $w_1, \dots, w_r, v_{r+1}, \dots, v_n$) e si può quindi dedurre anche che r' non supera r . Si conclude quindi che r e r' coincidono, provando così il teorema di Sylvester.

L'intero r che interviene nel teorema di Sylvester è chiamato *indice di positività* della forma.

Esercizi

1. Determinare l'indice di nullità e l'indice di positività per ognuna delle forme determinate dalle seguenti matrici simmetriche in \mathbb{R}^2

$$\text{a) } \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}. \quad \text{b) } \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \quad \text{c) } \begin{pmatrix} 1 & -3 \\ -3 & 2 \end{pmatrix}.$$

2. Sia V uno spazio vettoriale di dimensione finita sul corpo reale \mathbb{R} . Sia \langle , \rangle un prodotto scalare in V . Dimostrare che V si può scomporre nella somma diretta

$$V = V^+ \oplus V^- \oplus V_0,$$

dove V_0 è lo spazio definito nel teorema 12, V^+ è lo spazio dove la forma è definita positiva, V^- è lo spazio dove la forma è definita negativa. (Questo significa che

$$\langle v, v \rangle > 0 \quad \text{per ogni } v \in V^+$$

$$\langle v, v \rangle < 0 \quad \text{per ogni } v \in V^-.)$$

Dimostrare che le dimensioni degli spazi V^+, V^- sono le stesse in ciascuna di tali decomposizioni.

3. Sia V lo spazio vettoriale sul corpo reale \mathbb{R} costituito dalle matrici reali simmetriche 2×2 .

a) Dimostrare che la funzione f definita nello spazio V in modo che $f(A) = \text{Det}(A)$ è una forma quadratica su V .

b) Sia W il sottospazio di V costituito da tutte le matrici A tali che $\text{tr}(A) = 0$. Dimostrare che la forma bilineare associata alla forma quadratica f è definita negativa in W .

Capitolo 9

Polinomi e matrici

45. POLINOMI

Sia K un corpo. Per *polinomio su K* intendiamo una funzione di K in sé stesso tale che esistano in K elementi a_0, \dots, a_n in modo che

$$f(t) = a_n t^n + \dots + a_0$$

per ogni $t \in K$. Sia

$$g(t) = b_m t^m + \dots + b_0$$

un altro polinomio con $b_j \in K$: possiamo considerare la loro somma $f + g$. Se, per esempio, $n \geq m$, possiamo scrivere $b_j = 0$ se $j > m$,

$$g(t) = 0t^n + \dots + b_m t^m + \dots + b_0,$$

e quindi possiamo scrivere i valori della somma $f + g$ come

$$(f + g)(t) = (a_n + b_n)t^n + \dots + (a_0 + b_0),$$

quindi $f + g$ è di nuovo polinomio. Se $c \in K$ allora

$$(cf)(t) = ca_n t^n + \dots + ca_0,$$

e quindi cf è un polinomio. Questo significa che i polinomi costituiscono uno spazio vettoriale sul corpo K .

Possiamo anche considerare il prodotto dei due polinomi, fg , e allora

$$(fg)(t) = (a_n b_m)t^{n+m} + \dots + a_0 b_0.$$

E si vede che fg è di nuovo un polinomio. Infatti, se noi scriviamo

$$(fg)(t) = c_{n+m} t^{n+m} + \dots + c_0,$$

allora

$$c_k = \sum_{i=0}^k a_i b_{k-i} = a_0 b_k + a_1 b_{k-1} + \dots + a_k b_0.$$

Tutto quanto abbiamo richiamato ora era già noto, probabilmente, al lettore: abbiamo preferito ripetere queste cose per avviare il discorso.

TEOREMA 1 *Siano f, g due polinomi tali che, per ogni $t \in K$, $f(t) = g(t)$. Se possiamo scrivere*

$$f(t) = a_n t^n + \dots + a_0,$$

$$g(t) = b_n t^n + \dots + b_0.$$

Allora, per ogni i , $a_i = b_i$.

Dimostrazione. Diamo qui la dimostrazione soltanto nel caso in cui K coincide con \mathbb{R} oppure con \mathbb{C} . Consideriamo il polinomio $h = f - g$. Allora, per ogni x in K , $h(x) = 0$. Dobbiamo dimostrare che, per ogni i , $a_i - b_i = 0$. Supponiamo che non sia così e supponiamo che r sia il più grande intero per cui a_r e b_r sono diversi. Allora per ogni t , possiamo scrivere

$$0 = (a_r - b_r)t^r + \dots + (a_0 - b_0).$$

Dividendo per t^r , otteniamo

$$0 = a_r - b_r + \frac{a_{r-1} - b_{r-1}}{t} + \dots + \frac{a_0 - b_0}{t^r}.$$

Ora immaginiamo di dare a t valore molto grande (intendiamo sostituire t con numeri reali). Allora tutti gli addendi dopo $a_r - b_r$ tendono a zero. Ne segue che $a_r - b_r$ è nullo, contro la nostra ipotesi.

Abbiamo così dimostrato il teorema.

Osservazione. Nella dimostrazione ora fatta abbiamo adoperato la nozione di limite. Nel capitolo 12, paragrafo 55, daremo di questo stesso teorema una dimostrazione con metodi puramente

algebrici, adatta anche al caso di corpi più generali. In questo capitolo, nostro scopo principale è la discussione di semplici proprietà dei polinomi relative al corpo complesso e al corpo reale, in vista delle applicazioni che ne faremo nei successivi due capitoli. Perciò non rifuggiremo dal fare uso di qualche nozione tratta dall'analisi.

Il teorema 1 dice che quando scriviamo un polinomio f sotto la forma

$$f(t) = a_n t^n + \dots + a_0$$

con $a_i \in K$, allora i numeri a_0, \dots, a_n sono univocamente determinati. Essi sono chiamati i *coefficients* del polinomio. Se n è il più grande intero per cui a_n risulta diverso da zero, noi diciamo che n è il *grado* di f e scriviamo $n = \deg f$. Il numero a_n è anche chiamato il *primo coefficiente* di f , mentre a_0 è chiamato il *termine costante* di f . Se f è il polinomio nullo, per convenzione porremo $\deg f = -\infty$. A questo proposito poniamo anche le seguenti convenzioni

$$-\infty + -\infty = -\infty,$$

$$-\infty + a = -\infty, \quad -\infty < a,$$

per ogni intero a e conveniamo esplicitamente che *nessun'altra operazione con $-\infty$ è definita*.

Le convenzioni ora fatte servono a rendere vero senza eccezioni il teorema seguente:

TEOREMA 2 *Siano f, g polinomi con coefficients nel corpo K . Allora*

$$\deg(fg) = \deg f + \deg g.$$

Dimostrazione. Sia

$$f(t) = a_n t^n + \dots + a_0 \quad \text{e} \quad g(t) = b_m t^m + \dots + b_0,$$

dove i numeri a_n e b_m non sono nulli. Allora, moltiplicando f per g , vediamo che

$$f(t)g(t) = a_n b_m t^{n+m} + \text{termini di grado minore}$$

e $a_n b_m \neq 0$. Quindi $\deg(fg) = n + m = \deg f + \deg g$. Se f oppure g è il polinomio nullo, allora la nostra convenzione circa l'uso di $-\infty$ giustifica l'asserzione fatta.

Un polinomio di grado 1 è anche chiamato un polinomio *lineare*.

Ogni numero α tale che $f(\alpha) = 0$ sarà chiamato una *radice* di f . Ammettiamo, senza dimostrarla, la seguente affermazione:

TEOREMA 3 *Sia f un polinomio con coefficienti complessi di grado maggiore di zero. Allora in \mathbb{C} esiste qualche radice di f .*

Una dimostrazione di questo teorema, adoperando delle nozioni tratte dell'analisi, si trova nell'appendice.

TEOREMA 4 *Sia f un polinomio con coefficienti complessi, di grado positivo e con primo coefficiente uguale a 1. Esistono allora i numeri $\alpha_1, \dots, \alpha_n$ tali che*

$$f(t) = (t - \alpha_1) \dots (t - \alpha_n).$$

I numeri $\alpha_1, \dots, \alpha_n$ sono univocamente determinati a meno di permutazioni. Ogni radice α di f è uguale a uno degli α_i e viceversa.

Dimostrazione. Dimostreremo completamente il teorema 4 nel capitolo 12 (dando come vero il teorema 3). Poiché in questo capitolo e nei due successivi non avremo bisogno di conoscere nient'altro a proposito dei polinomi se non gli enunciati dati in questo paragrafo, riteniamo preferibile rimandarne la dimostrazione al capitolo successivo. Aggiungiamo che il proseguimento della teoria dei polinomi, sviluppata nel capitolo 12, avrà ulteriori applicazioni alla teoria delle applicazioni lineari e delle matrici.

A proposito di terminologia, se $\alpha_1, \dots, \alpha_r$ sono le radici distinte del polinomio f nel corpo complesso \mathbb{C} , possiamo scrivere

$$f(t) = (t - \alpha_1)^{m_1} \dots (t - \alpha_r)^{m_r},$$

dove gli interi m_1, \dots, m_r sono positivi e univocamente determinati. Diciamo allora che m_i è la *multiplicità* della radice α_i di f .

46. POLINOMI DI MATRICI E DI APPLICAZIONI LINEARI

L'insieme di tutti i polinomi con coefficienti nel corpo K viene denotato col simbolo $K[t]$.

Sia A una matrice quadrata con coefficienti in K . Sia

$f \in K[t]$ e scriviamo

$$f(t) = a_n t^n + \dots + a_0$$

con $a_i \in K$. Definiamo allora

$$f(A) = a_n A^n + \dots + a_0 I.$$

Esempio 1. Siano $f(t) = 3t^2 - 2t + 5$, e $A = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}$.

$$\text{Allora } f(A) = 3 \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}^2 - \begin{pmatrix} 2 & -2 \\ 4 & 0 \end{pmatrix} + \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 4 & -5 \\ 0 & -1 \end{pmatrix}.$$

TEOREMA 5 Siano f, g polinomi di $K[t]$. Sia A una matrice quadrata con coefficienti in K . Allora

$$(f + g)(A) = f(A) + g(A),$$

$$(fg)(A) = f(A)g(A).$$

Se $c \in K$, allora $(cf)(A) = cf(A)$.

Dimostrazione. Siano i polinomi $f(t)$ e $g(t)$ scritti sotto la forma

$$f(t) = a_n t^n + \dots + a_0$$

e

$$g(t) = b_m t^m + \dots + b_0$$

con a_i, b_j in K . Allora

$$(fg)(t) = c_{m+n} t^{m+n} + \dots + c_0$$

dove

$$c_k = \sum_{i=0}^k a_i b_{k-i}.$$

Per definizione,

$$(fg)(A) = c_{m+n} A^{m+n} + \dots + c_0 I.$$

D'altra parte,

$$f(A) = a_n A^n + \dots + a_0 I$$

e

$$g(A) = b_m A^m + \dots + b_0 I.$$

Da cui

$$f(A)g(A) = \sum_{i=0}^n \sum_{j=0}^m a_i A^i b_j A^j = \sum_{i=0}^n \sum_{j=0}^m a_i b_j A^{i+j} = \sum_{k=0}^{m+n} c_k A^k.$$

E quindi $f(A)g(A) = (fg)(A)$.

Nei riguardi della somma, supposto $n \geq m$, si ponga $b_j = 0$ se $j > m$. Abbiamo allora

$$\begin{aligned} (f+g)(A) &= (a_n + b_n)A^n + \dots + (a_0 + b_0)I \\ &= a_n A^n + b_n A^n + \dots + a_0 I + b_0 I \\ &= f(A) + g(A). \end{aligned}$$

Se $c \in K$, allora

$$(cf)(A) = ca_n A^n + \dots + ca_0 I = cf(A).$$

E questo prova il nostro teorema.

Esempio 2. Sia $f(t) = (t-1)(t+3) = t^2 + 2t - 3$. Allora

$$f(A) = A^2 + 2A - 3I = (A-I)(A+3I).$$

Se eseguiamo la moltiplicazione in quest'ultimo prodotto usando le regole della moltiplicazione tra matrici, otteniamo in effetti

$$A^2 - IA + 3AI - 3I^2 = A^2 + 2A - 3I.$$

Esempio 3. Siano $\alpha_1, \dots, \alpha_n$ numeri e si ponga

$$f(t) = (t - \alpha_1) \dots (t - \alpha_n).$$

Allora

$$f(A) = (A - \alpha_1 I) \dots (A - \alpha_n I).$$

Sia V uno spazio vettoriale sul corpo K e sia $A: V \rightarrow V$ un operatore (cioè un'applicazione lineare di V in sé stesso). Allora possiamo considerare $A^2 = A \circ A = AA$ e, più in generale, $A^n =$ = iterazione di A presa n volte, per ogni intero positivo n . Definiamo infine $A^0 = I$ (dove I ora denota l'applicazione identica). Abbiamo allora

$$A^{m+n} = A^m A^n$$

per ogni coppia di interi non negativi m, n . Se f è un polinomio in $K[t]$, noi possiamo considerare $f(A)$ come abbiamo fatto per

le matrici, allora valgono le stesse proprietà stabilite nel teorema 5. La dimostrazione è la stessa: i punti essenziali della dimostrazione fatta erano le ordinarie proprietà dell'addizione e della moltiplicazione, e queste valgono anche per le applicazioni lineari.

TEOREMA 6 *Sia A una matrice $n \times n$ in un corpo K . Allora esiste in $K[t]$ un polinomio non nullo f tale che $f(A) = O$.*

Dimostrazione. Lo spazio vettoriale delle matrici $n \times n$ sul corpo K ha dimensione finita, precisamente ha dimensione n^2 . Perciò le successive potenze

$$I, A, A^2, \dots, A^N$$

sono linearmente dipendenti quando N supera n^2 . Questo significa che nel corpo K esistono dei numeri a_0, \dots, a_N non tutti nulli e tali che

$$a_N A^N + \dots + a_0 I = O.$$

Per ottenere quanto vogliamo, basta allora porre $f(t) = a_N t^N + \dots + a_0$.

Come abbiamo visto per il teorema 5, anche qui notiamo che il teorema 6 vale per un'applicazione lineare A definita sullo spazio vettoriale di dimensione finita sul corpo K . La dimostrazione, anche in questo caso, è la stessa: noi ci riferiremo al teorema 6 indifferentemente trattando di matrici o di applicazioni lineari.

Più tardi, nel paragrafo 50 (cap. 10), daremo la costruzione esplicita di un polinomio $P(t)$ tale che $P(A) = O$.

Se dividiamo il polinomio f di cui si parla nel teorema 6 per il suo primo coefficiente otteniamo un polinomio g con primo coefficiente uguale a 1 e tale che $g(A) = O$. È solitamente conveniente considerare polinomi con primo coefficiente uguale a 1, questo può servire a semplificare le notazioni.

Esercizi

1. Calcolare $f(A)$ quando $f(t) = t^3 - 2t + 1$ e $A = \begin{pmatrix} -1 & 1 \\ 2 & 4 \end{pmatrix}$.

2. Sia A una matrice simmetrica e sia f un polinomio con coefficienti reali. Dimostrare che anche la matrice $f(A)$ è simmetrica.

3. Sia A una matrice hermitiana, sia f un polinomio con coefficienti reali. Dimostrare che anche la matrice $f(A)$ è hermitiana.

4. Siano A, B due matrici $n \times n$ nel corpo K e si supponga la matrice B invertibile. Dimostrare allora che

$$(B^{-1}AB)^n = B^{-1}A^nB$$

per ogni intero positivo n .

5. Sia f un polinomio in $K[t]$. Siano A, B due matrici come nell'esercizio 4. Dimostrare allora che

$$f(B^{-1}AB) = B^{-1}f(A)B.$$

47. AUTOVETTORI E AUTOVALORI

Sia V uno spazio vettoriale sul corpo K , sia

$$A: V \rightarrow V$$

un operatore di V (cioè un'applicazione lineare di V in sé stesso). Un elemento v di V è chiamato *autovettore* di A se in K esiste un elemento λ tale che $Av = \lambda v$. Se v non è il vettore nullo allora il numero λ è *univocamente determinato* giacché $\lambda_1 v = \lambda_2 v$ implica $\lambda_1 = \lambda_2$. In questo caso diciamo che λ è un *autovalore* di A relativo all'autovettore v . Diciamo anche che v è un autovettore avente come autovalore λ . Insieme alle denominazioni autovettore e autovalore si adoperano anche i termini *vettore caratteristico* e *valore caratteristico*.

Se A è una matrice quadrata $n \times n$ con coefficienti in K , allora un *autovettore* di A è, per definizione, un autovettore dell'applicazione lineare di K^n in sé stesso rappresentata da questa matrice relativamente alla base usuale. Quindi un autovettore X di A è un vettore (colonna) di K^n per cui esiste in K un numero λ tale che $AX = \lambda X$.

Esempio 1. Sia V lo spazio vettoriale sul corpo reale R generato da tutte le funzioni infinitamente differenziabili. Sia $\lambda \in R$. Allora la funzione f tale che $f(t) = e^{\lambda t}$ è un autovettore dell'operatore di derivazione d/dt perché $df/dt = \lambda e^{\lambda t}$.

Esempio 2. Sia

$$A = \begin{pmatrix} a_1 & & & \\ & \ddots & & 0 \\ & & \ddots & \\ 0 & & & a_n \end{pmatrix}$$

una matrice diagonale in K . Allora ogni vettore unità e^i ($i = 1, \dots, n$) è un autovettore di A . Infatti abbiamo $Ae^i = a_i e^i$:

$$\begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_n \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ a_i \\ \vdots \\ 0 \end{pmatrix}.$$

Esempio 3. Se $A: V \rightarrow V$ è un'applicazione lineare e se v è un autovettore di A , allora, per ogni scalare non nullo c , anche cv è un autovettore di A e ha lo stesso autovalore.

TEOREMA 7 *Sia V uno spazio vettoriale sul corpo K , sia $A: V \rightarrow V$ un'applicazione lineare. Sia λ un elemento di K . Sia infine V_λ il sottospazio di V generato da tutti gli autovettori di A che hanno λ come autovalore. Allora ogni elemento non nullo di V_λ è un autovettore di A che ha λ come autovalore.*

Dimostrazione. Siano v_1, v_2 vettori di V tali che $Av_1 = \lambda v_1$ e $Av_2 = \lambda v_2$. Allora

$$A(v_1 + v_2) = Av_1 + Av_2 = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2).$$

Se $c \in K$, allora $A(cv_1) = cAv_1 = c\lambda v_1 = \lambda cv_1$. Questo prova il nostro teorema.

Il sottospazio V_λ definito nel teorema 7 è chiamato l'*auto-spazio* di A relativo all'autovalore λ .

Osservazione. Se v_1, v_2 sono autovettori di A con autovalori diversi λ_1, λ_2 , allora è chiaro che $v_1 + v_2$ non è autovettore di A . Abbiamo infatti il seguente:

TEOREMA 8 *Sia V uno spazio vettoriale sul corpo K e sia $A: V \rightarrow V$ un operatore. Siano v_1, \dots, v_m autovettori di A , siano $\lambda_1, \dots, \lambda_m$ i loro rispettivi autovalori. Si assuma che questi autovalori siano tutti distinti, cioè*

$$\lambda_i \neq \lambda_j \quad \text{se} \quad i \neq j.$$

Allora i vettori v_1, \dots, v_m sono linearmente indipendenti.

Dimostrazione. Induzione rispetto a m . Per $m = 1$, un elemento di V , $v_1 \neq 0$, è linearmente indipendente. Si assuma ora

$m > 1$. Supponiamo di avere una relazione lineare

$$c_1 v_1 + \dots + c_m v_m = O \quad [1]$$

con $c_i \in K$. Vogliamo dimostrare che tutti gli scalari c_i sono nulli. Moltiplicando i membri della relazione [1] per λ_1 otteniamo

$$c_1 \lambda_1 v_1 + \dots + c_m \lambda_1 v_m = O.$$

Applichiamo ora A ai due membri della relazione [1]. A causa della proprietà di linearità otteniamo

$$c_1 \lambda_1 v_1 + \dots + c_m \lambda_m v_m = O;$$

sottraendo membro a membro queste due ultime espressioni, si ha

$$c_2 (\lambda_2 - \lambda_1) v_1 + \dots + c_m (\lambda_m - \lambda_1) v_m = O.$$

E poiché $\lambda_j - \lambda_1 \neq 0$ se $j = 2, \dots, m$, concludiamo, per l'ipotesi di induzione, che $c_2 = \dots = c_m = 0$. Ritornando alla nostra relazione originaria, vediamo che $c_1 v_1 = O$, da cui traiamo $c_1 = 0$ dimostrando così il nostro teorema.

Esempio 4. Sia V lo spazio vettoriale sul corpo complesso costituito da tutte le funzioni differenziabili a valori complessi di una variabile reale t . Siano $\alpha_1, \dots, \alpha_m$ numeri complessi distinti. Allora le funzioni

$$e^{\alpha_1 t}, \dots, e^{\alpha_m t}$$

sono autovettori dell'operatore di derivazione, essi hanno autovalori distinti $\alpha_1, \dots, \alpha_m$ e quindi sono linearmente indipendenti. In analisi è molto frequente la considerazione delle funzioni $e^{2\pi n i t}$, dove $n = 1, 2, \dots$

TEOREMA 9 *Sia V uno spazio vettoriale di dimensione finita sul corpo K e sia $\lambda \in K$. Sia $A: V \rightarrow V$ un'applicazione lineare. Allora λ è un autovalore di A se, e soltanto se, l'applicazione $A - \lambda I$ non è invertibile.*

Dimostrazione. Si assuma che λ sia un autovalore di A . Esiste allora un vettore non nullo $v \in V$ tale che $Av = \lambda v$. Allora $Av - \lambda v = O$ e anche $(A - \lambda I)v = O$. Ne segue che l'operatore $A - \lambda I$ ha un nucleo non nullo e quindi non può essere invertibile.

Viceversa, si assuma che $A - \lambda I$ non sia invertibile. Dal corollario del teorema 4, paragrafo 19 (cap. 4), deduciamo che $A - \lambda I$ deve avere un nucleo non nullo, significando con ciò che esiste un vettore non nullo $v \in V$ tale che $(A - \lambda I)v = O$. Quindi $Av - \lambda v = O$ e infine $Av = \lambda v$. Si conclude quindi che λ è un autovalore di A . Questo dimostra il nostro teorema.

TEOREMA 10 *Sia V uno spazio vettoriale di dimensione finita sul corpo complesso e sia $\dim V > 1$. Sia $A: V \rightarrow V$ un'applicazione lineare. Allora esiste un autovettore non nullo di A .*

Dimostrazione. A causa del teorema 6 (§ 46), esiste un polinomio non nullo $f \in \mathbb{C}[x]$ tale che $f(A) = O$, e possiamo inoltre assumere che f ha per primo coefficiente 1. Il teorema 4 (§ 45) afferma allora l'esistenza dei numeri complessi $\lambda_1, \dots, \lambda_n$ tali che

$$f(t) = (t - \lambda_1) \dots (t - \lambda_n).$$

Allora concludiamo che

$$O = f(A) = (A - \lambda_1 I) \dots (A - \lambda_n I).$$

Non tutte le applicazioni lineari $A - \lambda_i I$ sono invertibili perché, altrimenti, l'operatore composto

$$(A - \lambda_1 I) \dots (A - \lambda_n I)$$

sarebbe anch'esso invertibile. Quindi, per qualche i deve esistere un vettore non nullo $v \in V$ tale che

$$(A - \lambda_i I)v = O.$$

Ma allora $Av = \lambda_i v$ e quindi abbiamo trovato l'autovettore cercato.

Osservazione. L'affermazione contenuta nel teorema 10 è falsa se il nostro spazio vettoriale V ha il corpo reale invece di quello complesso come corpo degli scalari. Per esempio, una rotazione di un angolo θ in \mathbb{R}^2 non ha autovettori reali se θ non è multiplo intero di π . (Vedi l'esercizio 3.)

TEOREMA 11 *Sia V uno spazio vettoriale sul corpo K e sia $A: V \rightarrow V$ un'applicazione lineare. Si assuma che esista una base di V $\{v_1, \dots, v_n\}$ costituita da autovettori di A aventi rispettivamente*

gli autovalori $\lambda_1, \dots, \lambda_n$. Allora la matrice associata ad A rispetto a questa base è la matrice diagonale

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Dimostrazione. Evidente.

Abbiamo già visto, considerando l'esempio 2, il viceversa del teorema 11. È quindi importante saper determinare quando uno spazio vettoriale ha una base costituita da autovettori di un operatore dato. Questi sono precisamente i casi in cui l'operatore può essere diagonalizzato. Per esempio, ora dal teorema 8 possiamo trarre i corollari seguenti:

COROLLARIO 1 *Sia V uno spazio vettoriale di dimensione finita sul corpo K , sia n la dimensione di V . Sia $A: V \rightarrow V$ un operatore. Si assume che A abbia n autovalori distinti. Allora V ha una base costituita da autovettori di A e quindi A stesso può essere diagonalizzato.*

Oppure con il linguaggio delle matrici:

COROLLARIO 2 *Sia A una matrice $n \times n$ su un corpo K . Si assume che A abbia n autovalori distinti in K . Allora esiste una matrice B non singolare e a coefficienti in K , tale che la matrice $B^{-1}AB$ sia diagonale.*

Negli esercizi considereremo il caso delle matrici 2×2 . Diamo ora un esempio.

Esempio 5. Trovare gli autovalori e gli autovettori della matrice

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

Se esistono un autovettore $\begin{pmatrix} x \\ y \end{pmatrix}$ e un autovalore λ , allora

$$x + 2y = \lambda x,$$

$$-x + y = \lambda y,$$

e conseguentemente, sottraendo il primo membro dal secondo, abbiamo

$$x(\lambda - 1) - 2y = 0,$$

$$x + (\lambda - 1)y = 0.$$

Poiché x e y non possono essere entrambi nulli, il determinante

$$\begin{vmatrix} \lambda - 1 & -2 \\ 1 & \lambda - 1 \end{vmatrix}$$

è nullo e quindi λ è una radice del polinomio

$$(\lambda - 1)^2 + 2,$$

cioè $\lambda^2 - 2\lambda + 3 = 0$. Le radici sono $1 + i\sqrt{2}$ e $1 - i\sqrt{2}$.

Viceversa, se λ è una radice del polinomio precedente, allora il determinante prima scritto deve essere nullo e quindi esiste una soluzione del sistema di equazioni

$$x(\lambda - 1) - 2y = 0$$

$$x + (\lambda - 1)y = 0$$

costituita da numeri x , y non entrambi nulli. In questo modo troviamo un autovalore λ e un autovettore $\begin{pmatrix} x \\ y \end{pmatrix}$. Il calcolo può essere svolto esplicitamente. Siano $\lambda_1 = 1 + i\sqrt{2}$ e $\lambda_2 = 1 - i\sqrt{2}$. Risolvendo il sistema di equazioni lineari ponendovi $\lambda = \lambda_1$ si ottiene l'autovettore

$$v_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 1 \\ i\sqrt{2} \end{pmatrix}.$$

Risolvendo poi lo stesso sistema ponendovi $\lambda = \lambda_2$ si ottiene l'altro autovettore

$$v_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -i\sqrt{2} \end{pmatrix}.$$

Osservazione. Ogni autovettore relativo all'autovalore λ_1 è uguale a cv_1 , dove c è un numero non nullo. Ogni autovettore relativo all'autovalore λ_2 è uguale a $c'v_2$, dove c' è un numero non nullo. Osserviamo che gli autovalori e gli autovettori sono complessi e che non ci sono autovalori reali.

Esercizi

1. Sia

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

una matrice 2×2 in un corpo K . Dimostrare che ogni autovalore λ di questa matrice è una radice del polinomio

$$\begin{vmatrix} t-a & -b \\ -c & t-d \end{vmatrix} = (t-a)(t-d) - bc.$$

2. Viceversa, dimostrare che ogni radice del polinomio

$$\begin{vmatrix} t-a & -b \\ -c & t-d \end{vmatrix}$$

è un autovalore della matrice data.

3. Dimostrare che se θ è un numero reale non multiplo intero di π , allora la matrice

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

non ha autovettori non nulli in \mathbb{R}^2 .

4. Dimostrare che se θ è un numero reale allora la matrice

$$A = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$$

ha sempre un autovettore in \mathbb{R}^2 e che infatti esiste un vettore v_1 tale che $Av_1 = v_1$. [Si suggerisce di assumere come prima componente di v_1 , se $\cos \theta \neq 1$,

$$x = \frac{\sin \theta}{1 - \cos \theta}$$

e risolvere poi rispetto a y . Che cosa succede quando $\cos \theta = 1$?]

5. Con riferimento all'esercizio 4, sia v_2 un vettore di \mathbb{R}^2 perpendicolare al vettore v_1 trovato in quell'esercizio. Dimostrare che $Av_2 = -v_2$. In questo caso A viene chiamata riflessione.

6. Dimostrare che ogni matrice reale unitaria 2×2 con determinante uguale a -1 può essere espressa come prodotto

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

per un opportuno numero reale θ .

7. Sia V uno spazio vettoriale di dimensione finita sul corpo K , sia $A: V \rightarrow V$ un'applicazione lineare. Sia $v \in V$ un autovettore di A e sia $Av = \lambda v$. Se f è un polinomio in $K[t]$, dimostrare che $f(A)v = f(\lambda)v$.

8. Sia V uno spazio vettoriale di dimensione finita su un corpo K . Siano A, B applicazioni lineari dello spazio V in sé stesso. Si assuma che AB coincida con BA . Dimostrare allora che se v è un autovettore di A con autovalore λ , allora, se Bv non è il vettore nullo, anche Bv è un autovettore di A con autovalore λ .

9. Sia V uno spazio vettoriale di dimensione finita sul corpo complesso. Siano A, B due applicazioni lineari dello spazio V in sé stesso. Dimostrare che A e B hanno un autovettore in comune. [Suggerimento: se λ è un autovalore di A , si consideri lo spazio V_λ costituito da tutti i vettori v tali che $Av = \lambda v$ e si dimostri che B applica questo spazio in sé stesso. Poi proseguire da soli.]

10. Sia V uno spazio vettoriale di dimensione finita sul corpo K . Siano A, B due applicazioni lineari dello spazio V in sé stesso tali che $AB = BA$. Si supponga che esistano una base di V costituita da autovettori di A e una base di V costituita da autovettori di B . Dimostrare allora che esiste una base i cui elementi sono contemporaneamente autovettori di A e di B (cioè A e B possono essere diagonalizzate simultaneamente).

48. POLINOMIO CARATTERISTICO

Negli esercizi del paragrafo precedente abbiamo determinato per ogni matrice 2×2 un polinomio le cui radici erano gli autovalori della matrice. L'abbiamo fatto con l'usuale metodo di eliminazione. Vogliamo ora considerare il caso generale di matrici qualsiasi.

Sia A una matrice $n \times n$ in un corpo K , $A = (a_{ij})$. Definiamo *polinomio caratteristico* P_A di A il determinante

$$P_A(t) = \text{Det}(tI - A),$$

o, scritto per esteso,

$$P(t) = \begin{vmatrix} t - a_{11} & -a_{12} & \dots & -a_{1n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \dots & t - a_{nn} \end{vmatrix} = \begin{vmatrix} t - a_{11} & & & \\ -a_{12} & \ddots & & -a_{1n} \\ & \ddots & \ddots & \\ -a_{n1} & -a_{n2} & \dots & t - a_{nn} \end{vmatrix}.$$

Esempio 1. Il polinomio caratteristico della matrice

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

è

$$\begin{vmatrix} t-1 & 1 & -3 \\ 2 & t-1 & -1 \\ 0 & -1 & t+1 \end{vmatrix}$$

che noi possiamo sviluppare secondo la prima colonna trovando

$$P_A(t) = t^3 - t^2 - 4t + 6.$$

Per ogni arbitraria matrice $A = (a_{ij})$ il polinomio caratteristico può essere trovato sviluppando il determinante secondo la prima colonna e quindi consisterà sempre di una somma del tipo

$$(t - a_{11}) \dots (t - a_{nn}) + \dots$$

Ogni termine diverso da quello che abbiamo scritto ha grado minore di n . Il polinomio caratteristico quindi è del tipo

$$P_A(t) = t^n + \text{termini di grado minore di } n.$$

TEOREMA 12 *Sia A una matrice $n \times n$ in un corpo K . Un elemento $\lambda \in K$ è un autovalore di A se, e soltanto se, λ è una radice del polinomio caratteristico di A .*

Dimostrazione. Sia λ un autovalore di A . Allora, per il teorema 9, $\lambda I - A$ non è invertibile e quindi, per il teorema 8, paragrafo 32 (cap. 6), $\text{Det}(\lambda I - A) = 0$. Conseguentemente λ è una radice del polinomio caratteristico. Viceversa, se λ è una radice del polinomio caratteristico, allora $\text{Det}(\lambda I - A) = 0$ e quindi, sempre per il teorema 8 paragrafo 32, possiamo concludere che $\lambda I - A$ non è invertibile. Quindi, per il teorema 9, λ è un autovalore di A .

Il teorema 12 indica una via esplicita per determinare gli autovalori di una matrice, supponendo di potere determinare esplicitamente le radici del suo polinomio caratteristico. Questo talvolta riesce facilmente, specialmente negli esercizi alla fine dei capitoli, dove le matrici sono scelte in modo che si possano determinare le radici a prima vista o con semplici artifici. La cosa riesce considerevolmente più difficile in altri casi.

Per esempio, per determinare le radici del polinomio considerato nell'esempio 1, si dovrebbe prima sviluppare la teoria dei polinomi di terzo grado. Questo può essere fatto ma richiede delle formule che sono un po' più complicate della formula occorrente per risolvere un'equazione di secondo grado. Si può anche stu-

diare il modo di determinare le radici per approssimazione. In ogni caso lo studio di questi metodi appartiene a un ordine di idee diverso da quello in cui ci si muove in questo capitolo.

Esempio 2. Determinare gli autovalori della matrice

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

Calcoliamo il polinomio caratteristico, che è il determinante

$$\begin{vmatrix} t-1 & -1 & 1 \\ 0 & t-1 & 0 \\ -1 & 0 & t-1 \end{vmatrix}$$

facilmente scrivibile nella forma

$$P(t) = (t-1)(t^2 - 2t + 2).$$

Le sue radici sono $1, 1+i, 1-i$. Queste sono le radici caratteristiche. Osserviamo che vi è una sola radice caratteristica reale.

Esempio 3. Il polinomio caratteristico della matrice

$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 5 & -1 \\ 0 & 0 & 7 \end{pmatrix}$$

è $(t-1)(t-5)(t-7)$. Il lettore vede un modo di generalizzare questa situazione?

TEOREMA 13 *Siano A, B due matrici $n \times n$ e si supponga B invertibile. Allora il polinomio caratteristico di A coincide con il polinomio caratteristico della matrice $B^{-1}AB$.*

Dimostrazione. Dalla definizione e dalle proprietà del determinante otteniamo

$$\begin{aligned} \text{Det}(tI - A) &= \text{Det}(B^{-1}(tI - A)B) = \text{Det}(tB^{-1}B - B^{-1}AB) = \\ &= \text{Det}(tI - B^{-1}AB). \end{aligned}$$

E questo prova quanto volevamo.

Vogliamo ora definire il polinomio caratteristico di un'applicazione lineare. Seguiamo il solito metodo, considerando la sua rappresentazione mediante matrici.

Sia $A: V \rightarrow V$ un'applicazione lineare definita in uno spazio

vettoriale di dimensione finita sul corpo K . Sia \mathcal{B} una base di V e sia M la matrice che rappresenta A rispetto a \mathcal{B} . Sia \mathcal{B}' un'altra base, esiste allora una matrice invertibile N in K tale che la matrice M' di A rispetto alla base \mathcal{B}' si può scrivere

$$M' = N^{-1}MN.$$

Per il teorema 13 il polinomio caratteristico di M coincide con quello di M' . Definiamo questo polinomio caratteristico come il *polinomio caratteristico* dell'applicazione lineare A .

Se λ è un autovalore di A , la *moltiplicità* di λ è definita come la sua molteplicità quale radice del polinomio caratteristico di A .

Esercizi

1. Sia A la matrice diagonale

$$A = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_n \end{pmatrix}.$$

- a) Qual è il polinomio caratteristico di A ?
- b) Quali sono i suoi autovalori?

2. Sia A la matrice triangolare

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Qual è il polinomio caratteristico di A e quali sono i suoi autovalori?

3. Determinare il polinomio caratteristico delle seguenti matrici e determinare anche, nel corpo complesso, gli autovalori e gli autovettori

a) $\begin{pmatrix} 1 & i \\ i & -2 \end{pmatrix}$. b) $\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$. c) $\begin{pmatrix} 1 & 2i \\ 0 & 2 \end{pmatrix}$.

d) $\begin{pmatrix} 2 & 4 \\ 5 & 3 \end{pmatrix}$. e) $\begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$. f) $\begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix}$.

g) $\begin{pmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{pmatrix}$. h) $\begin{pmatrix} 3 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & -1 \end{pmatrix}$.

4. Sia A una matrice 4×4 con polinomio caratteristico uguale a

$$(t-2)(t+1)(t+i)(t+2i).$$

Dimostrare che A può essere diagonalizzata nel corpo complesso.

5. Sia V uno spazio vettoriale di dimensione n sul corpo complesso e si assuma che il polinomio caratteristico di un'applicazione lineare $A: V \rightarrow V$ abbia n radici distinte. Dimostrare che V ha una base costituita da autovettori di A .

6. Sia A la matrice

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

dove a è un numero complesso non nullo. Dimostrare che A non può essere diagonalizzata.

7. Adoperando il teorema del valor medio per le funzioni continue, dimostrare che ogni matrice reale $n \times n$ ha un autovettore non nullo, reale se n è dispari.

8. Sia A una matrice quadrata. Dimostrare che gli autovalori della matrice $'A'$ sono uguali a quelli della matrice A . Gli autovettori di $'A'$ sono uguali a quelli della matrice A ?

9. Dimostrare che gli autovalori della matrice

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

nel corpo complesso sono i numeri ± 1 e $\pm i$.

10. Sia V lo spazio generato sul corpo reale \mathbb{R} dalle due funzioni $\sin t$ e $\cos t$. L'operatore di derivazione (considerato come applicazione lineare di V in sé stesso) ha in V autovettori non nulli? In caso affermativo, quali sono?

11. Determinare gli autovalori e gli autovettori della matrice

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

Dimostrare che gli autovettori formano uno spazio di dimensione 1.

12. Determinare gli autovalori e gli autovettori delle matrici seguenti:

a) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. b) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. c) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

Capitolo 10

Triangolazione delle matrici e delle applicazioni lineari

49. ESISTENZA DELLE TRIANGOLAZIONI

Sia V uno spazio vettoriale di dimensione finita sul corpo K e si assuma che la dimensione n di V sia diversa da zero. Sia $A: V \rightarrow V$ un'applicazione lineare. Si chiama *ventaglio per A* (in V) una successione di sottospazi $\{V_1, \dots, V_n\}$ tale che V_i è contenuto in V_{i+1} per ogni $i = 1, \dots, n-1$, inoltre tale che $\dim V_i = i$, e infine tale che A applica ogni V_i in sé stesso. In altre parole, se v appartiene a V_i allora anche Av è contenuto in V_i . Diciamo brevemente in questo caso che AV_i è contenuto in V_i .

Osserviamo che le dimensioni dei sottospazi V_1, \dots, V_n aumentano di un'unità da ogni sottospazio a quello che segue. Inoltre V_n coincide con V .

Daremo ora una interpretazione dei ventagli mediante le matrici. Sia $\{V_1, \dots, V_n\}$ un ventaglio per A . Dicendo *base a ventaglio* intenderemo una base $\{v_1, \dots, v_n\}$ di V tale che $\{v_1, \dots, v_i\}$ sia una base per V_i . Si vede immediatamente che una base a ventaglio esiste. Per esempio, sia $\{v_1\}$ una base di V_1 . Estendiamo $\{v_1\}$ a una base $\{v_1, v_2\}$ di V_2 (cosa possibile per un vecchio teorema) e quindi estendiamo ancora ad una base $\{v_1, v_2, v_3\}$ di V_3 e così induttivamente fino a una base $\{v_1, \dots, v_n\}$ di V_n .

TEOREMA 1 *Sia $\{v_1, \dots, v_n\}$ una base a ventaglio per A . Allora la matrice associata ad A relativa a questa base è una matrice triangolare superiore.*

Dimostrazione. Poiché AV_i è contenuto in V_i per ogni $i=1, \dots, n$, esistono i numeri a_{ij} in modo che

$$\begin{aligned} Av_1 &= a_{11}v_1, \\ Av_2 &= a_{12}v_1 + a_{22}v_2, \\ &\vdots \\ Av_i &= a_{1i}v_1 + a_{2i}v_2 + \dots + a_{ii}v_i, \\ &\vdots \\ Av_n &= a_{1n}v_1 + a_{2n}v_2 + \dots + a_{nn}v_n. \end{aligned}$$

Ciò significa che la matrice associata ad A rispetto alla nostra base è la matrice triangolare

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix},$$

come si doveva dimostrare.

Osservazione. Sia A una matrice triangolare superiore come quella già considerata. Riguardiamo A come un'applicazione lineare di K^n in sé stesso. Allora i vettori colonna unità e^1, \dots, e^n costituiscono una base a ventaglio per A . Se indichiamo con V_i lo spazio generato dai vettori e^1, \dots, e^i allora $\{V_1, \dots, V_n\}$ è il ventaglio corrispondente. Quindi il viceversa del teorema 1 è immediatamente chiaro.

Ricordiamo che non è sempre possibile trovare un autovettore (oppure autovalore) di un'applicazione lineare se il corpo assegnato K non è quello complesso. Analogamente, non è sempre possibile trovare un ventaglio per un'applicazione lineare quando K è il corpo reale. Se $A: V \rightarrow V$ è un'applicazione lineare e se esiste una base di V rispetto alla quale la matrice associata ad A è triangolare, diciamo che A è *triangolabile*. Analogamente, se A è una matrice $n \times n$ sul corpo K , noi diciamo che A è *triangolabile su K* se essa è triangolabile considerata come applicazione lineare di K^n in sé stesso. Questo equivale a dire che esiste una matrice non singolare B in K tale che $B^{-1}AB$ sia una matrice triangolare superiore.

Tenendo conto dell'esistenza di autovettori sul corpo complesso, dimostreremo che ogni matrice o applicazione lineare può essere triangolata sul corpo complesso.

TEOREMA 2 *Sia V uno spazio vettoriale di dimensione finita sul corpo complesso e si assuma che la dimensione di V sia diversa da zero. Sia $A: V \rightarrow V$ un'applicazione lineare. Allora in V esiste un ventaglio per A .*

Dimostrazione. Procederemo per induzione. Se $\dim V = 1$, non c'è nient'altro da provare. Assumiamo allora che il teorema sia vero quando $\dim V = n - 1$, se n è positivo. Per il teorema 10, paragrafo 47 (cap. 9), esiste un autovettore non nullo v_1 di A . Sia V_1 il sottospazio di dimensione 1 generato da v_1 . Possiamo allora scrivere V come somma diretta $V = V_1 \oplus W$ per un opportuno sottospazio W (per il teorema 9, § 10, cap. 2, che essenzialmente afferma che ogni insieme di vettori linearmente indipendenti può essere esteso fino a una base). L'inconveniente ora è che A non applica W in sé stesso. Sia P_1 la proiezione di V su V_1 , sia P_2 la proiezione di V su W . Allora $P_2 A$ è un'applicazione lineare di V in V che applica W in W (perché P_2 applica ogni elemento di V in un elemento di W). Consideriamo quindi $P_2 A$ come un'applicazione lineare di W in sé stesso. Per l'ipotesi di induzione, esiste un ventaglio di $P_2 A$ in W , sia questo $\{W_1, \dots, W_{n-1}\}$. Poniamo allora

$$V_i = V_1 + W_{i-1}$$

per $i = 2, \dots, n$. Allora V_i è contenuto in V_{i+1} per ogni $i = 1, \dots, n$ ed è facile verificare che $\dim V_i = i$.

(Se $\{u_1, \dots, u_{n-1}\}$ è una base di W tale che $\{u_1, \dots, u_j\}$ è una base di W_j , allora $\{v_1, u_1, \dots, u_{i-1}\}$ è una base di V_i per $i = 2, \dots, n$.)

Per provare che $\{V_1, \dots, V_n\}$ è un ventaglio per A in V , sarà sufficiente provare che AV_i è contenuto in V_i . A questo scopo, osserviamo che

$$A = IA = (P_1 + P_2)A = P_1 A + P_2 A.$$

Sia $v \in V_i$; possiamo allora scrivere $v = cv_1 + w_i$, con $c \in \mathbb{C}$ e $w_i \in W_i$. Allora $P_1 Av = P_1(Av)$ è contenuto in V_1 e quindi in V_i . Inoltre

$$P_2 Av = P_2 A(cv_1) + P_2 Aw_i.$$

Poiché $P_2A(cv_1) = cP_2Av_1$ e poiché v_1 è un autovettore di A , sia per esempio $Av_1 = \lambda_1 v_1$, troviamo che $P_2A(cv_1) = P_2(c\lambda_1 v_1) = O$. Per l'ipotesi di induzione, P_2A applica W_1 in sé stesso e quindi P_2Aw_i appartiene a W_1 . Perciò P_2Av appartiene a $V_1 + W_1 = V_1$. Osserviamo, per concludere, che $PAv = P_1Av + P_2Av$ appartiene a V_1 , provando così il nostro teorema.

COROLARIO 1 *Sia V uno spazio vettoriale di dimensione finita sul corpo complesso, si assuma che la dimensione di V non sia nulla. Sia $A: V \rightarrow V$ un'applicazione lineare. Allora esiste una base di V rispetto alla quale la matrice associata ad A è triangolare.*

Dimostrazione. Abbiamo già esposto il ragionamento da fare prima del teorema 1.

COROLARIO 2 *Sia M una matrice di numeri complessi. Esiste allora una matrice non singolare B tale che la matrice $B^{-1}MB$ è triangolare.*

Dimostrazione. Si tratta della usuale interpretazione del cambiamento di matrici quando si cambia la base, applicata al caso trattato nel corollario 1.

Esercizi .

1. Sia A una matrice triangolare superiore

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}.$$

Considerando A come un'applicazione lineare, quali sono gli autovalori di A^3, A^r , in generale di A^r , essendo r un intero positivo?

2. Sia A una matrice quadrata. Diciamo che A è *nilpotente* se esiste un intero r positivo tale che $A^r = O$. Dimostrare che se A è nilpotente, allora tutti gli autovalori di A sono nulli.

3. Sia V uno spazio vettoriale di dimensione finita sul corpo complesso, sia $A: V \rightarrow V$ un'applicazione lineare. Si assuma che tutti gli autovalori di A siano nulli. Dimostrare allora che la matrice A è nilpotente.

(Dapprima svolgere dettagliatamente questi due ultimi esercizi considerando matrici 2×2 .)

4. Facendo uso della nozione di ventaglio, dimostrare che l'inversa di una matrice triangolare invertibile è di nuovo triangolare. Dimostrare infatti

che se V è uno spazio vettoriale di dimensione finita, se $A: V \rightarrow V$ è un'applicazione lineare invertibile, allora se $\{V_1, \dots, V_n\}$ è un ventaglio per A , lo è anche per A^{-1} .

5. Sia A una matrice quadrata di numeri complessi tale che, per un certo intero positivo r , $A^r = I$. Se α è un autovalore di A , dimostrare che $\alpha^r = 1$.

6. Determinare una base a ventaglio per le applicazioni lineari definite in \mathbb{C}^2 e rappresentate dalle matrici

$$\text{a) } \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \quad \text{b) } \begin{pmatrix} 1 & i \\ 1 & i \end{pmatrix}. \quad \text{c) } \begin{pmatrix} 1 & 2 \\ i & i \end{pmatrix}.$$

50. TEOREMA DI HAMILTON-CAYLEY

Sia V uno spazio vettoriale di dimensione finita su un corpo K e sia $A: V \rightarrow V$ un'applicazione lineare. Si assuma che V abbia una base costituita da autovettori di A , sia questa $\{v_1, \dots, v_n\}$. Siano $\{\lambda_1, \dots, \lambda_n\}$ i corrispondenti autovalori. Allora il polinomio caratteristico di A si scrive

$$P(t) = (t - \lambda_1) \dots (t - \lambda_n),$$

e

$$P(A) = (A - \lambda_1 I) \dots (A - \lambda_n I).$$

Se ora applichiamo $P(A)$ ad ogni vettore v_i , il fattore $A - \lambda_i I$ farà scomparire v_i , in altre parole, $P(A)v_i = O$. Conseguentemente, $P(A) = O$.

In generale non sarà possibile trovare una base come quella precedente. Tuttavia, adoperando i ventagli, possiamo arrivare a una generalizzazione dell'argomentazione appena fatta per il caso diagonale.

TEOREMA 3 *Sia V uno spazio vettoriale di dimensione finita sul corpo complesso, la dimensione di V sia diversa da zero. Sia $A: V \rightarrow V$ un'applicazione lineare e sia P il suo polinomio caratteristico. Allora $P(A) = O$.*

Dimostrazione. Sappiamo dal teorema 2 che possiamo trovare un ventaglio per A , $\{V_1, \dots, V_n\}$. Sia

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ 0 & \dots & a_{2n} \\ \vdots & & \vdots \\ 0 & \dots & a_{nn} \end{pmatrix}$$

la matrice associata ad A rispetto a una base a ventaglio $\{v_1, \dots, v_n\}$. Allora

$$Av_i = a_{ii}v_i + \text{un elemento di } V_{i-1}$$

oppure, in altre parole, dato che $(A - a_{ii}I)v_i = Av_i - a_{ii}v_i$, troviamo che

$$(A - a_{ii}I)v_i \text{ appartiene a } V_{i-1}.$$

Inoltre il polinomio caratteristico di A è dato da

$$P(t) = (t - a_{11}) \dots (t - a_{nn}),$$

e allora

$$P(A) = (A - a_{11}I) \dots (A - a_{nn}I).$$

Proveremo per induzione che

$$(A - a_{11}I) \dots (A - a_{ii}I)v = O$$

per ogni v in V_i , $i = 1, \dots, n$. Quando $i = n$, ritroviamo il nostro teorema.

Sia $i = 1$. Allora $(A - a_{11}I)v_1 = Av_1 - a_{11}v_1 = O$ e l'asserzione è provata.

Sia $i > 1$ e si supponga l'asserzione provata per $i-1$. Ogni elemento di V_i può essere scritto come una somma $v' + cv_i$, dove v' è in V_{i-1} e c è un opportuno scalare. Osserviamo che $(A - a_{ii}I)v'$ appartiene a V_{i-1} perché AV_{i-1} è contenuto in V_{i-1} e altrettanto accade per $a_{ii}v'$. Per l'ipotesi di induzione

$$(A - a_{11}I) \dots (A - a_{i-1,i-1}I)(A - a_{ii}I)v' = O.$$

D'altra parte, $(A - a_{ii}I)cv_i$ appartiene a V_{i-1} e quindi, per induzione,

$$(A - a_{11}I) \dots (A - a_{i-1,i-1}I)(A - a_{ii}I)cv_i = O.$$

Infine, se v è in V_i , abbiamo

$$(A - a_{11}I) \dots (A - a_{ii}I)v = O,$$

provando così il nostro teorema.

COROLLARIO 1 *Sia A una matrice $n \times n$ di numeri complessi e sia P il suo polinomio caratteristico. Allora $P(A) = O$.*

Dimostrazione. Basta considerare A come un'applicazione lineare di \mathbb{C}^n in sé stesso e tenere conto del teorema ora dimostrato.

COROLLARIO 2 *Sia V uno spazio vettoriale di dimensione finita sul corpo K , sia $A: V \rightarrow V$ un'applicazione lineare. Sia infine P il polinomio caratteristico di A . Allora $P(A) = O$.*

Dimostrazione. Fissata in V una base, sia M la matrice che rappresenta A rispetto a questa base. Allora $P_M = P_A$ e basterà quindi dimostrare che $P_M(M) = O$. Ricorrendo allora al teorema 3, si conclude la dimostrazione.

Osservazione. La dimostrazione del teorema 3 si può fare anche adoperando la nozione di continuità. Considerata una matrice complessa A , si può provare, in vari modi sui quali tuttavia non ci soffermiamo, che esistono matrici Z aventi la stessa dimensione di A , vicine quanto si vuole ad A (cioè ogni componente di Z è vicina alla corrispondente di A) e tali che P_Z ha tutte le sue radici di molteplicità 1. Infatti, i polinomi complessi aventi radici di molteplicità superiore a 1 sono distribuiti tra gli altri polinomi in modo rarefatto. Ora, se Z è uno dei polinomi con le proprietà dette sopra, l'applicazione lineare rappresentata da Z è diagonalizzabile (perché Z ha gli autovalori tutti distinti tra loro) e allora, ovviamente, $P_Z(Z) = O$, come abbiamo notato all'inizio di questo paragrafo. D'altra parte, $P_Z(Z)$ tende a $P_A(A)$ al tendere di Z ad A . Si conclude perciò che $P_A(A) = O$.

51. DIAGONALIZZAZIONE DI APPLICAZIONI UNITARIE

TEOREMA 4 *Sia V uno spazio vettoriale di dimensione finita sul corpo dei numeri complessi, sia $\dim V \geq 1$. In V sia dato un prodotto hermitiano definito positivo. Sia $A: V \rightarrow V$ un'applicazione unitaria. Esiste allora una base ortonormale di V costituita di autovettori di A .*

Dimostrazione. Osserviamo dapprima che se w è un autovettore di A con autovalore λ , allora $Aw = \lambda w$ e λ non può essere nullo perché A conserva la norma.

Per quanto abbiamo visto nel teorema 2, possiamo trovare un ventaglio in A , $\{V_1, \dots, V_n\}$, e conseguentemente una base a ventaglio $\{v_1, \dots, v_n\}$. Adoperando il processo di Gram-Schmidt pos-

siamo ortogonalizzare questa base. Questo processo, ricordiamo, consiste nel porre successivamente:

$$\begin{aligned} v'_1 &= v_1 \\ v'_2 &= v_2 - \frac{\langle v_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1. \\ &\dots \end{aligned}$$

Dalla costruzione fatta, vediamo che $\{v'_1, \dots, v'_n\}$ è una base ortogonale che è ancora una base a ventaglio giacché $\{v'_1, \dots, v'_i\}$ è una base dello stesso spazio V_i di cui era base $\{v_1, \dots, v_i\}$. Per ottenere una base a ventaglio ortonormale $\{w_1, \dots, w_n\}$ basta dividere ogni vettore v'_i per la sua norma. Vogliamo ora far vedere che ogni w_i è un autovettore dell'operatore A . Procediamo per induzione. Appartenendo Aw_1 al sottospazio V_1 , deve esistere uno scalare λ_1 tale che $Aw_1 = \lambda_1 w_1$: w_1 è quindi un autovettore e λ_1 non è nullo. Supponiamo ora di avere già provato che w_1, \dots, w_{i-1} sono autovettori con autovalori non nulli. Osserviamo che esistono degli scalari c_1, \dots, c_i tali che

$$Aw_i = c_1 w_1 + \dots + c_i w_i.$$

E poiché A conserva la perpendicolarità, Aw_i risulta perpendicolare ad ogni Aw_k per ogni k minore di i . Ma $Aw_k = \lambda_k w_k$, ne segue allora che Aw_i è perpendicolare anche a w_k e si conclude che c_k deve essere nullo. Allora $Aw_i = c_i w_i$ e inoltre c_i non può essere nullo perché A conserva la norma. Possiamo quindi andare da 1 a n e provare il nostro teorema.

COROLLARIO *Sia A una matrice unitaria complessa. Esiste allora una matrice unitaria U tale che $U^{-1}AU$ sia una matrice diagonale.*

Dimostrazione. Sia $\mathcal{B} = \{e^1, \dots, e^n\}$ la base ortonormale ordinaria di \mathbb{C}^n e sia $\mathcal{B}' = \{w_1, \dots, w_n\}$ una base ortonormale rispetto alla quale A , considerata come un'applicazione lineare di \mathbb{C}^n in sé stesso, sia diagonalizzabile. Si ponga

$$U = M_{\mathcal{B}}^{\mathcal{B}'}(id).$$

Allora U è unitaria (si veda l'esercizio 5 del § 43, cap. 8) e, se M'

è la matrice di A relativa alla base \mathcal{B}' , allora

$$M' = U^{-1}AU.$$

Ne segue che anche M' è unitaria.

Esercizi

1. Sia A una matrice complessa unitaria. Dimostrare che ogni autovalore di A si può scrivere come $e^{i\theta}$, per un opportuno numero reale θ .
2. Sia A una matrice complessa unitaria. Dimostrare che esistono una matrice diagonale B e una matrice complessa unitaria U in modo che A sia uguale a $U^{-1}BU$.

Capitolo 11

Il teorema spettrale

S2. AUTOVETTORI DI APPLICAZIONI LINEARI SIMMETRICHE

In tutto questo capitolo avremo a che fare con autovettori e autovalori di speciali tipi di operatori, cioè operatori simmetrici (o, nel caso complesso, hermitiani). Cominciamo col considerare il caso di operatori simmetrici sul corpo dei numeri reali.

Sia V uno spazio vettoriale sul corpo reale, di dimensione finita e maggiore di zero. Supponiamo che in V sia fissato un prodotto scalare definito positivo e simmetrico, indicato con $\langle \cdot, \cdot \rangle$. Il lettore può pensare a V come lo spazio \mathbb{R}^n con l'ordinario prodotto scalare, noi però considereremo anche sottospazi di V , e in questi sottospazi il prodotto scalare sarà quello indotto dal prodotto scalare di V ; essi però non saranno uguali a \mathbb{R}^m , saranno soltanto isomorfi a \mathbb{R}^m , per opportuni m , quando le basi siano scelte convenientemente. La cosa più importante in questo capitolo sarà la scelta di basi soddisfacenti a varie condizioni.

Sia $A: V \rightarrow V$ un'applicazione lineare. Ricordiamo che A si dice simmetrica quando

$$\langle Av, w \rangle = \langle v, Aw \rangle = \langle Aw, v \rangle$$

per ogni scelta di v, w in V . Se V coincide con \mathbb{R}^n e il prodotto scalare è quello ordinario, rappresentando A con una matrice rispetto alla base usuale, A è simmetrica se, e soltanto se, la matrice che la rappresenta è simmetrica. Nel caso generale, se scegliamo una base ortonormale di V , rappresentando A con una matrice rispetto a questa base, di nuovo si trova che A è simmetrica se, e soltanto se, la matrice trovata è simmetrica.

In effetti è talvolta conveniente riferirsi agli spazi \mathbb{R}^n , \mathbb{C}^n in qualche argomentazione esplicita. Per esempio consideriamo il teorema che segue, essenzialmente già dimostrato in precedenza, ma che ripetiamo qui in tutti i dettagli per comodità del lettore.

TEOREMA 1 *Sia A una matrice reale simmetrica $n \times n$. Se λ è un autovalore di A in \mathbb{C} , allora λ è reale.*

Dimostrazione. Ci riferiamo al prodotto hermitiano in \mathbb{C}^n definito in modo che se Z, Z' sono in \mathbb{C}^n , allora

$$\langle Z, Z' \rangle = {}^t Z \bar{Z}' = z_1 \bar{z}'_1 + \dots + z_n \bar{z}'_n.$$

Sia Z un autovettore avente λ come autovalore; allora $Z \neq 0$ e $AZ = \lambda Z$. Abbiamo poi

$$\langle AZ, Z \rangle = \langle \lambda Z, Z \rangle = \lambda \langle Z, Z \rangle.$$

D'altra parte, essendo A un operatore simmetrico e reale, abbiamo anche

$$\langle AZ, Z \rangle = \langle Z, {}^t AZ \rangle = \langle Z, AZ \rangle = \overline{\langle AZ, Z \rangle}.$$

Perciò $\lambda \langle Z, Z \rangle$ coincide col suo complesso coniugato ed è quindi reale. Che λ stesso sia reale segue dall'essere $\langle Z, Z \rangle$ reale e non nullo.

Sia Z un vettore di \mathbb{C}^n . Possiamo allora scrivere Z , in un unico modo, come somma $Z = X + iY$, X e Y essendo vettori reali di \mathbb{R}^n . Questo infatti è vero per ogni componente di Z e quindi anche per Z stesso. Per esempio,

$$\begin{pmatrix} 3+2i \\ 1-i \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} + i \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

TEOREMA 2 *Sia A una matrice reale simmetrica $n \times n$. Allora A possiede autovettori reali non nulli.*

Dimostrazione. Teniamo conto del fatto che, se consideriamo A come un'applicazione lineare di \mathbb{C}^n in sé stesso, A ha certamente un autovettore complesso non nullo Z . Esiste perciò un numero complesso λ tale che $AZ = \lambda Z$. Il teorema 1 precisa che questo numero λ è reale. Scrivendo

$$Z = X + iY$$

dove X, Y sono vettori reali, otteniamo

$$AZ = AX + iAY$$

e quindi

$$AZ = \lambda Z = \lambda X + i\lambda Y.$$

Essendo λ un numero reale e poiché le parti reale e immaginaria di un vettore complesso sono univocamente determinate, concludiamo che

$$AX = \lambda X \quad \text{e} \quad AY = \lambda Y.$$

Poiché almeno uno dei due vettori X, Y non è nullo, uno di essi è quindi un autovettore reale non nullo, come volevamo dimostrare.

COROLLARIO *Sia V uno spazio vettoriale sul corpo reale, di dimensione finita e maggiore di zero. In V sia dato un prodotto scalare definito positivo. Se $A: V \rightarrow V$ è un'applicazione lineare simmetrica, allora A possiede in V autovettori non nulli.*

Dimostrazione. Fissiamo in V una base ortonormale. Allora, rispetto a questa base, A è rappresentato da una matrice reale simmetrica. Se X è un autovettore non nullo di A in \mathbb{R}^n , allora l'elemento v di V , avente X come vettore delle coordinate rispetto alla base fissata, è un autovettore non nullo di A in V .

Nella dimostrazione appena fatta del teorema 2 abbiamo fatto uso di una proprietà dei numeri complessi, cioè abbiamo supposto l'esistenza di un autovettore complesso per una matrice complessa. Adoperando un'altra nozione tratta dal calcolo infinitesimale, è istruttivo dare un'altra dimostrazione del teorema 2 basandola sulla considerazione del massimo di una opportuna funzione.

Sia S la sfera unitaria di \mathbb{R}^n , cioè l'insieme di tutti i vettori X di \mathbb{R}^n tali che $\|X\| = 1$.

Ad ogni matrice reale simmetrica A associamo la forma quadratica $f: \mathbb{R}^n \rightarrow \mathbb{R}$ definita da $f(X) = \langle AX, X \rangle$. Consideriamo ora i valori che la funzione f assume sulla sfera unitaria.

TEOREMA 3 *Sia A una matrice reale simmetrica $n \times n$. Sia*

$$f(X) = \langle AX, X \rangle$$

per ogni $X \in \mathbb{R}^n$. Sia v un vettore della sfera unitaria tale che $f(v) > f(X)$ per ogni X sulla sfera stessa. Allora v è un autovettore di A .

Dimostrazione. Sia w un vettore unità di \mathbb{R}^n , perpendicolare al vettore v . Si vede facilmente allora che esiste una curva differentiabile

$$C: (a, b) \rightarrow \mathbb{R}^n$$

giacente sulla sfera unitaria, definita in un intervallo aperto contenente 0, tale che $C(0) = v$ e tale che il vettore tangente di C in 0 sia w , in altre parole $C'(0) = w$. [Per esempio, la curva

$$C(t) = (\cos t)v + (\sin t)w$$

soddisfa queste condizioni.] Poiché $f(v) = f(C(0))$ è un massimo della funzione f sulla sfera S , è anche un massimo per $f \circ C$ nell'intervallo (a, b) e quindi $(f \circ C)'(0) = 0$. Ricordando la regola di derivazione di un prodotto otteniamo

$$\begin{aligned} (f \circ C)'(t) &= \frac{d}{dt} \langle AC(t), C(t) \rangle = \\ &= \langle AC'(t), C(t) \rangle + \langle AC(t), C'(t) \rangle = \\ &= \langle C'(t), AC(t) \rangle + \langle C'(t), AC(t) \rangle = \\ &= 2\langle AC(t), C'(t) \rangle, \end{aligned}$$

da cui

$$0 = (f \circ C)'(0) = 2\langle AC(0), C'(0) \rangle = 2\langle Av, w \rangle.$$

Quindi Av è perpendicolare ad ogni vettore w che sia perpendicolare a v e segue allora che Av giace nello spazio generato da v : in altre parole, esiste un numero reale λ tale che $Av = \lambda v$. Il teorema è così dimostrato.

Richiamando all'analisi il fatto che la sfera è un insieme compatto e quindi che ogni funzione continua su essa è definita da un massimo, concludiamo che esiste sempre un vettore con le proprietà espresse nel teorema 3. Abbiamo quindi dato un'altra dimostrazione dell'esistenza di un autovettore della matrice A .

Esercizi

1. Sia A una matrice reale simmetrica 3×3 . Sia, per ogni $X \in \mathbb{R}^n$, $f(X) = \langle AX, X \rangle$. Sia v un vettore nell'ellissoide definito dall'equazione

$$3x^2 + 4y^2 + 5z^2 = 1,$$

tale che $f(v) \geq f(X)$ per ogni X appartenente all'ellissoide. Dimostrare allora che v è un autovettore di A .

2. Più in generale, sia $\varphi: \mathbb{R}^8 \rightarrow \mathbb{R}$ una funzione con derivata continua. Sia S la superficie definita dall'equazione $\varphi(X) = 0$. Si assuma che per ogni punto P appartenente a S , il gradiente di $\varphi(P)$ non sia nullo. Si assuma inoltre la superficie S chiusa e limitata. Sia A una matrice reale simmetrica 3×3 e sia $f(X) = \langle AX, X \rangle$. Sia v un punto di S tale che $f(v) > f(X)$ per ogni X su S . Dimostrare allora che v è un autovettore di A .

3. Sia A una matrice reale $n \times n$ (non si assume la simmetria). Si assuma che tutti gli autovalori di A siano reali. Dimostrare allora che A ha un autovettore reale e non nullo.

53. IL TEOREMA SPETTRALE

Ritorniamo ora alle considerazioni algebriche.

In tutto questo paragrafo facciamo l'ipotesi che V sia uno spazio vettoriale di dimensione finita sul corpo reale \mathbb{R} , che inoltre $\dim V > 1$ e che in V sia dato un prodotto scalare definito positivo.

TEOREMA 4 *Sia $A: V \rightarrow V$ un'applicazione lineare simmetrica. Sia v un autovettore non nullo di A . Se w è un elemento di V perpendicolare a v , allora anche il vettore Aw è perpendicolare a v .*

Dimostrazione. Si tratta di una semplicissima osservazione:

$$\langle Aw, v \rangle = \langle w, Av \rangle = \langle w, \lambda v \rangle = \lambda \langle w, v \rangle = 0.$$

TEOREMA 5 *Sia $A: V \rightarrow V$ un'applicazione lineare simmetrica. Allora si può trovare una base ortogonale di V costituita da autovettori di A .*

Dimostrazione. Procediamo per induzione sulla dimensione di V . Se $\dim V = 1$, non vi è nulla da provare. Sia $\dim V > 1$. Il teorema 2 afferma l'esistenza in V di un autovettore non nullo v_1 di A . Sia $W = v_1^\perp$ lo spazio ortogonale di v_1 . Allora

$$\dim W = \dim V - 1.$$

Per il teorema 4, A applica lo spazio W in sé stesso. Osserviamo che W ha un prodotto scalare definito positivo: quello indotto dal prodotto scalare di V . Inoltre A , considerata come applicazione lineare di W in sé stessa, è simmetrica. Per l'ipotesi di induzione, esiste allora una base ortogonale di W costituita da autovettori di A , sia $\{v_2, \dots, v_n\}$. Allora, per ogni i maggiore di 1,

i vettori v_0 e v_1 sono perpendicolari e quindi $\{v_1, \dots, v_n\}$ è la base la cui esistenza è asserita nel teorema.

Esempio. Sia A la matrice

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

Vogliamo trovare una base ortogonale di \mathbb{R}^2 costituita da autovettori di A .

Gli autovalori di A sono

$$\frac{5 \pm \sqrt{5}}{2}.$$

(Questo risultato si trova con equazioni lineari oppure calcolando le radici del polinomio caratteristico che è $t^2 - 5t + 5$.) Per trovare un autovettore, dobbiamo risolvere le equazioni

$$2x + y = \frac{5 + \sqrt{5}}{2} x,$$

$$x + 3y = \frac{5 + \sqrt{5}}{2} y.$$

E allora troviamo $x = 2$, $y = 1 + \sqrt{5}$. Concludiamo che

$$v_1 = \begin{pmatrix} 2 \\ 1 + \sqrt{5} \end{pmatrix}$$

è un autovettore. Lo spazio ortogonale a v_1 ha dimensione 1 e perciò consiste di tutti i multipli reali di un vettore perpendicolare a v_1 . Possiamo per esempio supporre

$$v_2 = \begin{pmatrix} 2 \\ 1 - \sqrt{5} \end{pmatrix}.$$

Questo è necessariamente un autovettore di A giacché A applica lo spazio ortogonale di v_1 in sé stesso. Allora $\{v_1, v_2\}$ è la base da noi richiesta.

COROLLARIO *Sia A una matrice reale simmetrica $n \times n$. Esiste allora una matrice unitaria reale $n \times n$ U tale che ${}^t UAU = U^{-1}AU$ sia una matrice diagonale.*

Dimostrazione. Consideriamo A come la matrice associata all'applicazione lineare $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$, relativamente alla base usuale $\mathcal{B} = \{e^1, \dots, e^n\}$. Il teorema 5 assicura la possibilità di trovare una base ortonormale di \mathbb{R}^n , $\mathcal{B}' = \{w_1, \dots, w_n\}$, tale che $M_{\mathcal{B}'}^{\mathcal{B}'}(F)$ sia diagonale. (Vedi esercizio 2 alla fine del paragrafo.) Sia $U = M_{\mathcal{B}'}^{\mathcal{B}'}(\text{id})$. Allora $U^{-1}AU$ è una matrice diagonale. Inoltre U è unitaria: sia infatti $U = (c_{ij})$. Allora

$$w_i = \sum_{s=1}^n c_{is} e^s, \quad i = 1, \dots, n.$$

Le uguaglianze $w_i \cdot w_i = 1$ e $w_i \cdot w_j = 0$ se $i \neq j$, significano, come si vede immediatamente, che ${}^t U U = I$, cioè che ${}^t U = U^{-1}$. Questo conclude la dimostrazione.

Osservazione 1. Nel teorema 5 noi consideriamo due forme sullo spazio vettoriale V . Dapprima consideriamo la forma definita positiva $\langle \cdot, \cdot \rangle$, poi la forma g definita da $g(v, w) = \langle Av, w \rangle$. Osserviamo che una base ortogonale $\{v_1, \dots, v_n\}$ di V costituita da autovettori di A è una base ortogonale anche per la seconda forma. Infatti:

$$g(v_i, v_j) = \langle Av_i, v_j \rangle = \langle \lambda_i v_i, v_j \rangle = \lambda_i \langle v_i, v_j \rangle.$$

Quest'ultima espressione è ovviamente nulla quando i e j non coincidono: questo prova che la nostra base è ortogonale anche relativamente alla seconda forma.

Osservazione 2. L'assegnato prodotto scalare definito positivo su V dà luogo a una forma quadratica $f_1: V \rightarrow \mathbb{R}$ tale che

$$f_1(v) = \langle v, v \rangle.$$

L'operatore A dà luogo a una seconda forma quadratica (non necessariamente definita positiva!), precisamente la forma f_2 definita da

$$f_2(v) = \langle Av, v \rangle.$$

Sia $\{v_1, \dots, v_n\}$ una base ortogonale di V costituita da autovettori di A . Sia $c_i = \langle v_i, v_i \rangle$. Sia infine

$$v = x_1 v_1 + \dots + x_n v_n, \quad x_i \in \mathbb{R}$$

dove X è il vettore delle coordinate di v rispetto alla nostra base.
Allora

$$Av = x_1 \lambda_1 v_1 + \dots + x_n \lambda_n v_n,$$

avendo indicato con $\lambda_1, \dots, \lambda_n$ gli autovalori di A corrispondenti rispettivamente agli autovettori v_1, \dots, v_n . Ne segue allora che

$$f_1(v) = c_1 x_1^2 + \dots + c_n x_n^2 \quad \text{e} \quad f_2(v) = \lambda_1 c_1 x_1^2 + \dots + \lambda_n c_n x_n^2.$$

Interpretiamo queste uguaglianze dicendo che le nostre due forme quadratiche sono *simultaneamente diagonalizzate*. Il teorema 5 può allora enunciarsi come segue: *due forme reali quadratiche, una delle quali sia definita positiva, possono sempre essere simultaneamente diagonalizzate*.

Una base avente le proprietà espresse nel teorema 5 si chiama una *base spettrale* per l'operatore A .

Esercizi

1. Determinare una base ortogonale di \mathbb{R}^2 costituita da autovettori delle seguenti matrici

$$\begin{array}{lll} \text{a)} \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}. & \text{b)} \begin{pmatrix} -1 & 1 \\ 1 & 2 \end{pmatrix}. & \text{c)} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}. \\ \text{d)} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. & \text{e)} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. & \text{f)} \begin{pmatrix} 2 & -3 \\ -3 & 1 \end{pmatrix}. \end{array}$$

2. Sia A una matrice reale simmetrica $n \times n$. Dimostrare che si può trovare una base *ortonormale* di \mathbb{R}^n costituita da autovettori di A .

3. Sia V uno spazio vettoriale con le proprietà descritte nel paragrafo 53. Sia $A: V \rightarrow V$ un'applicazione lineare simmetrica. Siano v_1, v_2 autovettori di A aventi rispettivamente gli autovalori λ_1, λ_2 . Dimostrare che, se λ_1 e λ_2 non coincidono, allora i vettori v_1 e v_2 sono perpendicolari tra loro.

4. Sia A una matrice reale simmetrica 2×2 . Dimostrare che se gli autovalori di A sono distinti, allora i loro autovettori costituiscono una base ortogonale di \mathbb{R}^2 .

5. Sia V uno spazio vettoriale con le proprietà descritte nel paragrafo 53. Sia $A: V \rightarrow V$ un'applicazione lineare avente un solo autovalore. Dimostrare allora che *ogni* base ortogonale di V è costituita da autovettori di A .

6. Sia V uno spazio vettoriale con le proprietà descritte nel paragrafo 53. Sia $A: V \rightarrow V$ un'applicazione lineare simmetrica. Sia $\dim V = n$ e si supponga che A abbia n autovalori distinti. Dimostrare allora che i corrispondenti autovettori costituiscono una base ortogonale di V .

7. Sia V uno spazio vettoriale con le proprietà descritte nel paragrafo 53, e sia $A: V \rightarrow V$ un'applicazione lineare simmetrica. Dimostrare che se il nucleo di A è $\{O\}$ allora nessun autovalore di A è nullo e viceversa.

8. Dimostrare che ogni matrice reale simmetrica può essere scritta nella forma $'UBU'$, dove B è una matrice diagonale e U è una matrice unitaria reale.

9. Sia V uno spazio vettoriale con le proprietà descritte nel paragrafo 53, e sia $A: V \rightarrow V$ un'applicazione lineare simmetrica. Dimostrare che le seguenti proprietà di A sono equivalenti:

- a) Tutti gli autovalori di A sono positivi.
- b) Per ogni vettore non nullo v di V , si ha $\langle Av, v \rangle > 0$.

Ogni applicazione A avente queste proprietà viene chiamata *definita positiva*. La stessa definizione si dà a proposito delle matrici reali simmetriche considerate come applicazioni lineari di R^n in sé stesso. Per esempio, la seconda proprietà, espressa mediante i vettori delle coordinate, dice che:

- b') Per ogni vettore non nullo X in R^n , abbiamo

$$'XAX > 0.$$

10. Determinare quali tra le seguenti matrici sono definite positive.

a) $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. b) $\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$. c) $\begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix}$.

d) $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 3 & 1 & 1 \end{pmatrix}$. e) $\begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}$.

11. Dimostrare che le seguenti proprietà di una matrice reale simmetrica sono equivalenti:

- a) Tutti gli autovalori di A sono negativi.
- b) Per ogni vettore non nullo X in R^n , abbiamo $'XAX < 0$.

12. Sia A una matrice reale simmetrica non singolare $n \times n$. Dimostrare le seguenti proposizioni:

- a) Ogni autovalore di A è diverso da zero.
- b) Se λ è un autovalore di A , allora λ^{-1} è un autovalore di A^{-1} .
- c) Le matrici A e A^{-1} hanno gli stessi autovettori.

13. Sia A una matrice reale simmetrica definita positiva. Dimostrare che A^{-1} esiste ed è definita positiva.

14. Sia A una matrice reale simmetrica i cui autovalori sono tutti non negativi. Dimostrare che allora esiste una matrice reale simmetrica B tale che $B^2 = A$ e $AB = BA$.

15. Dimostrare che una matrice reale simmetrica A è definita positiva se, e soltanto se, esiste una matrice reale non singolare N tale che $A = {}^t NN$.

[Si suggerisce di adoperare il corollario del teorema 5 e scrivere tUAU come il quadrato di una matrice diagonale, per esempio B^2 . Porre allora $N = UB^{-1}$.]

16. Sia V uno spazio vettoriale con le proprietà descritte nel paragrafo 53. Sia $A: V \rightarrow V$ un'applicazione lineare simmetrica. Riferendosi al teorema di Sylvester dimostrato nei paragrafi precedenti, dimostrare che l'indice di nullità della forma definita da

$$(v, w) \mapsto \langle Av, w \rangle$$

è uguale alla dimensione del nucleo di A . Dimostrare poi che l'indice di positività è uguale al numero degli autovettori di una base spettrale aventi un autovalore positivo.

17. Sia V uno spazio vettoriale con le proprietà descritte nel paragrafo 53. Siano A e B due operatori simmetrici di V tali che $AB = BA$. Dimostrare che esiste allora una base ortogonale di V che è una base spettrale contemporaneamente per A e per B , costituita cioè da autovettori di A e B contemporaneamente. [Suggerimento: Se λ è un autovalore di A e se V_λ è costituito da tutti i vettori v di V tali che $Av = \lambda v$, dimostrare che BV_λ è contenuto in V_λ . Questa osservazione riduce il problema al caso in cui A sia della forma λI .]

18. Sia V uno spazio vettoriale con le proprietà descritte nel paragrafo 53 e sia $A: V \rightarrow V$ un operatore simmetrico. Siano $\lambda_1, \dots, \lambda_r$ gli autovalori distinti di A . Se λ è un autovalore di A , si indichi con $V_\lambda(A)$ l'insieme di tutti i vettori v di V tali che $Av = \lambda v$.

a) Dimostrare che $V_\lambda(A)$ è un sottospazio di V e che A applica $V_\lambda(A)$ in sé stesso.

b) Dimostrare che V è la somma diretta degli spazi

$$V = V_{\lambda_1}(A) \oplus \dots \oplus V_{\lambda_r}(A)$$

e che questi spazi sono a due a due ortogonali.

Lo spazio $V_\lambda(A)$ si chiama l'autospazio di A appartenente a λ .

19. Con riferimento alle notazioni introdotte nell'esercizio precedente, si supponga A definito positivo. Dimostrare allora che esiste un operatore simmetrico definito positivo B definito su V tale che $B^2 = A$ e dimostrare che questo operatore B è univocamente determinato. [Suggerimento: dimostrare dapprima che A e B hanno gli stessi autospazi, procedendo come segue. Siano μ_1, \dots, μ_s gli autovalori distinti di B e sia

$$V = V_{\mu_1}(B) \oplus \dots \oplus V_{\mu_s}(B)$$

la decomposizione di V in somma diretta di autospazi di B . Dimostrare allora che ogni $V_{\mu_i}(B)$ è un autospazio di A appartenente a un certo autovalore λ_i di A e provare poi che, se μ_i e μ_j non coincidono, neppure λ_i e λ_j coincidono. Concludere che $s = r$, $V_{\mu_i}(B) = V_{\lambda_i}(A)$ e quindi che l'effetto dell'operatore B è determinato univocamente da quello di A .]

20. Sia V uno spazio vettoriale con le proprietà descritte nel paragrafo 53 e sia $A: V \rightarrow V$ un arbitrario operatore invertibile definito su V . Dimostrare allora che esistono un operatore reale unitario U e un operatore simmetrico definito positivo P tali che $A = UP$, U e P essendo univocamente determinati. [Suggerimento: sia P l'operatore simmetrico definito positivo, definito dalla condizione $P^2 = {}^tAA$. Posto $U = AP^{-1}$, si provi che U è unitario, così si arriva all'esistenza. Per provare l'unicità, si supponga $A = U_1 P_1$, U_1 essendo un operatore unitario e P_1 essendo un operatore simmetrico definito positivo. Sia $U_2 = P_1^{-1}U_1$. Allora $I = {}^tU_2 U_2$ (perché?) e quindi $P^2 = P_1^2$. Ricorrere all'esercizio 19 per concludere che $P = P_1$ e finalmente che $U = U_1$.]

21. Se P_1 e P_2 sono due matrici reali simmetriche definite positive (della stessa dimensione) se t e u sono numeri reali positivi, dimostrare che la matrice $tP_1 + uP_2$ è simmetrica e definita positiva.

22. Sia V uno spazio vettoriale con le proprietà descritte nel paragrafo 53 e sia $A: V \rightarrow V$ un operatore simmetrico. Siano $\lambda_1, \dots, \lambda_r$ gli autovalori distinti di A . Dimostrare allora che

$$(A - \lambda_1 I) \dots (A - \lambda_r I) = O.$$

23. Sia V uno spazio vettoriale con le proprietà descritte nel paragrafo 53 e sia $A: V \rightarrow V$ un operatore simmetrico. Un sottospazio W di V viene detto *invariante per A* se, per ogni w in W , anche Aw appartiene a W , in altre parole quando $AW \subseteq W$. Dimostrare che se A non ha sottospazi invarianti oltre $\{O\}$ e V , allora esiste un numero λ tale che $A = \lambda I$. [Suggerimento: dimostrare dapprima che A ha un solo autovalore.]

54. CASO COMPLESSO

Come al solito, abbiamo un analogo del teorema spettrale nel caso complesso.

TEOREMA 6 *Sia V uno spazio vettoriale di dimensione finita maggiore di zero sul corpo dei numeri complessi. Sia \langle , \rangle una forma hermitiana definita positiva assegnata su V . Sia $A: V \rightarrow V$ un'applicazione lineare hermitiana. Esiste allora una base ortogonale di V costituita da autovettori di A .*

Dimostrazione. Stessa argomentazione fatta per dimostrare il teorema 5.

Come ulteriore esercizio, si facciano tutte le osservazioni relative al caso hermitiano analoghe a quelle fatte nel paragrafo precedente relativamente al caso simmetrico. Si osservi anche che se $\{v_1, \dots, v_n\}$ è una base di quelle descritte nel teorema, allora la matrice dell'operatore A relativa a questa base è una matrice

diagonale *reale*. Questo significa che la teoria degli operatori hermitiani (o delle matrici hermitiane) può essere trattata esattamente come nel caso reale.

Esercizi

In tutti questi esercizi noi supponiamo che V sia uno spazio vettoriale di dimensione finita maggiore di zero sul corpo complesso \mathbb{C} , assumiamo anche che in V sia fissato un prodotto hermitiano definito positivo.

1. Un operatore $A: V \rightarrow V$ viene detto *normale* quando $AA^* = A^*A$. Enunciare e dimostrare il teorema spettrale per un operatore A normale. [Suggerimento per la dimostrazione: trovare un autovettore comune ad A e ad A^* .]
2. Se f è un polinomio a coefficienti reali e se A è un operatore hermitiano, dimostrare che l'operatore $f(A)$ è anch'esso hermitiano.
3. Dare una definizione di operatore hermitiano definito positivo e dimostrare le proposizioni analoghe a quelle contenute negli esercizi 9 e 13 del paragrafo 53.
4. Sia A un operatore hermitiano definito positivo, dimostrare allora che esiste un operatore hermitiano B tale che $B^2 = A$ e $AB = BA$. L'operatore B è univocamente determinato?
5. Un operatore hermitiano A viene chiamato *positivo* (non necessariamente definito positivo) quando nessun suo autovalore è negativo. Dimostrare che questa proprietà equivale a chiedere che, per ogni $v \in V$, $\langle Av, v \rangle \geq 0$.
6. Dimostrare che un operatore hermitiano positivo A ha una radice quadrata, cioè esiste un operatore hermitiano B , univocamente determinato, tale che $B^2 = A$.
7. Dimostrare che la matrice

$$A = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$$
 è positiva. Se ne trovi la radice quadrata.
8. Sia A un operatore hermitiano. Dimostrare che esistono due operatori hermitiani positivi P_1, P_2 tali che $A = P_1 - P_2$.
9. Sia A un operatore invertibile. Dimostrare che esistono un operatore unitario complesso U e un operatore hermitiano definito positivo P , entrambi univocamente determinati, tali che $A = UP$. [Suggerimento: si proceda come nell'esercizio 20 del § 53.]
10. Siano A, B operatori normali tali che $AB = BA$. Dimostrare che l'operatore AB è anch'esso normale.

11. Sia A una matrice hermitiana $n \times n$. Dimostrare che esiste una matrice complessa unitaria $n \times n$ U tale che U^*AU sia una matrice diagonale. Trovare una tale matrice U nel caso in cui A sia uguale a:

a) $\begin{pmatrix} 2 & 1+i \\ 1-i & 1 \end{pmatrix}$. b) $\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$.

12. Sia A una matrice complessa non singolare. Dimostrare che A è hermitiana e definita positiva se, e soltanto se, esiste una matrice non singolare N tale che $A = N^*N$.

Capitolo 12

Polinomi e decomposizioni primarie

55. L'ALGORITMO EUCLIDEO

Nel capitolo 9 abbiamo già parlato di polinomi e del loro grado. In questo capitolo tratteremo di altre importanti proprietà dei polinomi. Quella più importante è l'algoritmo euclideo, oppure divisione lunga, insegnata, presumiamo, in tutte le scuole medie.

TEOREMA 1 *Siano f, g polinomi sul corpo K , cioè polinomi appartenenti a $K[t]$. Si assuma g di grado maggiore o uguale a zero. Esistono allora in $K[t]$ i polinomi q, r tali che*

$$f(t) = q(t)g(t) + r(t),$$

e $\deg r < \deg g$. I polinomi q e r sono univocamente determinati da queste condizioni.

Dimostrazione. Sia $m = \deg g \geq 0$. Sia

$$f(t) = a_n t^n + \dots + a_0,$$

$$g(t) = b_m t^m + \dots + b_0,$$

dove $b_m \neq 0$. Se $n < m$, si ponga $q = 0, r = f$. Se $n \geq m$, si ponga

$$f_1(t) = f(t) - a_n b_m^{-1} t^{n-m} g(t).$$

(Questo è il primo passo nel procedimento della divisione lunga.) Allora $\deg f_1 < \deg f$. Continuando allo stesso modo o, più formalmente, procedendo per induzione su n , possiamo trovare due

polinomi q_1, r tali che

$$f_1 = q_1 g + r,$$

con $\deg r < \deg g$. Allora

$$\begin{aligned} f(t) &= a_n b_m^{-1} t^{n-m} g(t) + f_1(t) = \\ &= a_n b_m^{-1} t^{n-m} g(t) + q_1(t)g(t) + r(t) = \\ &= (a_n b_m^{-1} t^{n-m} + q_1)g(t) + r(t), \end{aligned}$$

e quindi troviamo che il nostro polinomio è espresso nella forma richiesta. Per provare l'unicità, si supponga che

$$f = q_1 g + r_1 = q_2 g + r_2,$$

con $\deg r_1 < \deg g$ e $\deg r_2 < \deg g$. Allora

$$(q_1 - q_2)g = r_2 - r_1.$$

Se il primo membro non è zero, il suo grado non è inferiore a quello di g . Il secondo membro, d'altra parte, se non è zero ha grado inferiore a quello di g . L'unica possibilità è quindi che essi siano entrambi zero e questo implica che

$$q_1 = q_2 \quad \text{e} \quad r_1 = r_2,$$

come si voleva dimostrare.

COROLLARIO 1 *Sia f un polinomio non nullo appartenente a $K[t]$. Sia $\alpha \in K$ tale che $f(\alpha) = 0$. Allora esiste in $K[t]$ un polinomio $q(t)$ tale che*

$$f(t) = (t - \alpha)q(t).$$

Dimostrazione. Intanto possiamo scrivere

$$f(t) = q(t)(t - \alpha) + r(t),$$

con $\deg r < \deg(t - \alpha)$. Ma $\deg(t - \alpha) = 1$. Perciò r è costante. Osservato poi che

$$0 = f(\alpha) = q(\alpha)(\alpha - \alpha) + r(\alpha) = r(\alpha),$$

si conclude, come richiesto, che r è nullo.

COROLLARIO 2 *Sia K un corpo tale che ogni polinomio non costante appartenente a $K[t]$ abbia una radice in K . Sia f un siffatto polinomio. In K esistono allora gli elementi $\alpha_1, \dots, \alpha_n$ e c in modo che*

$$f(t) = c(t - \alpha_1) \dots (t - \alpha_n).$$

Dimostrazione. Nel corollario 1 si osservi che $\deg q = \deg f - 1$. Si consideri il corollario 1 nel caso in cui $\alpha = \alpha_1$. Per ipotesi, se q non è costante, esiste una radice α_2 di q e quindi si può scrivere

$$f(t) = q_2(t)(t - \alpha_1)(t - \alpha_2).$$

Procedendo per induzione, si continui questo processo fino a che q_{n+1} non risulta costante.

Poiché supponiamo che il corpo dei numeri complessi soddisfi l'ipotesi del corollario 2, vediamo che in tal modo abbiamo provato l'esistenza di una fattorizzazione di un polinomio sul corpo complesso in fattori di grado 1. L'unicità di questa fattorizzazione sarà dimostrata nel paragrafo successivo.

COROLLARIO 3 *Sia f un polinomio di grado n appartenente a $K[t]$. Allora in K esistono al più n radici di f .*

Dimostrazione. Nel caso contrario, se $\alpha_1, \dots, \alpha_m$, con $m > n$, fossero radici distinte di f in K , si avrebbe

$$f(t) = (t - \alpha_1) \dots (t - \alpha_m)g(t)$$

con un opportuno polinomio g , si dedurrebbe allora la contraddizione $\deg f > m$.

Esercizi

1. In ognuno dei seguenti casi, scrivere $f = qg + r$ con $\deg r < \deg g$.

- a) $f(t) = t^3 - 2t + 1$, $g(t) = t - 1$.
- b) $f(t) = t^3 + t - 1$, $g(t) = t^2 + 1$.
- c) $f(t) = t^3 + t$, $g(t) = t$.
- d) $f(t) = t^3 - 1$, $g(t) = t - 1$.

2. Se $f(t)$ ha coefficienti interi, se $g(t)$ ha coefficienti interi e primo coefficiente uguale a 1, dimostrare che quando scriviamo $f = qg + r$, con $\deg r < \deg g$, anche i polinomi q e r hanno i coefficienti interi.

3. Richiamando il teorema del valor medio dal calcolo infinitesimale, dimostrare che ogni polinomio di grado dispari sul corpo reale ha una radice reale.

4. Sia $f(t) = t^n + \dots + a_0$ un polinomio con coefficienti complessi, di grado n , e sia α una sua radice. Dimostrare che $|\alpha| < n \cdot \max_i |a_i|$. [Suggerimento: scrivere $-\alpha^n = a_{n-1}\alpha^{n-1} + \dots + a_0$. Se $|\alpha| > n \cdot \max_i |a_i|$, si divida per α^n e si consideri il valore assoluto. Una semplice maggiorazione darà una contraddizione.]

56. MASSIMO COMUN DIVISORE

Vogliamo ora definire una nozione che, relativamente all'insieme dei polinomi $K[t]$, si comporta come un sottospazio relativamente a uno spazio vettoriale.

Chiameremo *ideale di $K[t]$* , oppure *ideale polinomiale* o, più brevemente, *ideale*, ogni insieme J contenuto in $K[t]$ soddisfacente le condizioni seguenti:

Il polinomio nullo appartiene a J ; se f, g sono in J , allora anche $f + g$ è in J ; se f è in J e se g è un polinomio arbitrario, allora gf è in J .

Osserviamo che da quest'ultima condizione segue che se $c \in K$ e se $f \in J$, allora cf appartiene a J . Quindi ogni ideale è anche uno spazio vettoriale sul corpo K . Ma è qualcosa di più, in considerazione del fatto che è chiuso rispetto alla moltiplicazione per elementi arbitrari di $K[t]$, non soltanto costanti.

Esempio 1. Siano f_1, \dots, f_n polinomi appartenenti a $K[t]$. Sia J l'insieme di tutti i polinomi che possono venire scritti sotto la forma

$$g = g_1 f_1 + \dots + g_n f_n$$

dove ogni g_i appartiene a $K[t]$. Allora J è un ideale. Infatti, se

$$h = h_1 f_1 + \dots + h_n f_n$$

dove $h_i \in K[t]$, allora

$$g + h = (g_1 + h_1) f_1 + \dots + (g_n + h_n) f_n$$

appartiene anch'esso all'insieme J . Inoltre, $0 = 0f_1 + \dots + 0f_n$ appartiene a J . Se f è un arbitrario polinomio in $K[t]$, allora anche il polinomio

$$fg = (fg_1) f_1 + \dots + (fg_n) f_n$$

appartiene a J . Abbiamo così verificato tutte le condizioni.

L'ideale J considerato nell'esempio 1 si dice *generato* dai polinomi f_1, \dots, f_n e diciamo che f_1, \dots, f_n ne costituiscono un *insieme di generatori*.

Osserviamo che ogni polinomio f_i appartiene all'ideale J considerato nell'esempio 1. Per esempio,

$$f_1 = 1 \cdot f_1 + 0 \cdot f_2 + \dots + 0 \cdot f_n.$$

Esempio 2. L'insieme costituito dal solo elemento 0 è un ideale. Anche $K[t]$ è un ideale. Notiamo che il polinomio 1 è un generatore di $K[t]$ che perciò viene detto *ideale unità*.

Esempio 3. Consideriamo l'ideale generato dai due polinomi $t-1$ e $t-2$. Vogliamo far vedere che si tratta dell'ideale unità. Basta osservare, infatti, che il polinomio 1 si può scrivere

$$(t-1) - (t-2) = 1$$

e quindi appartiene all'ideale. Si vede così che può accadere che un ideale sia generato da più elementi e che, d'altra parte, se ne possa trovare un insieme di generatori costituito da un solo elemento. Questa situazione viene chiarita nei teoremi che seguono.

TEOREMA 2 *Sia J un ideale di $K[t]$. Esiste allora un polinomio g che genera J .*

Dimostrazione. Supponiamo che J non sia l'ideale costituito dal solo 0. Sia g un polinomio di J non nullo e di grado minimo (tra quelli dei polinomi di J). Vogliamo far vedere che g è un generatore di J . Sia f un qualsiasi elemento di J . L'algoritmo euclideo ci fa trovare due polinomi q, r tali che

$$f = qg + r$$

ed inoltre $\deg r < \deg g$. Allora $r = f - qg$ e quindi, per definizione di ideale, anche r appartiene a J . Poiché $\deg r < \deg g$, deve necessariamente essere $r = 0$. Perciò $f = qg$ e g genera quindi J , come volevamo.

Osservazione. Sia g_1 un generatore non nullo di un ideale J , generato anche da un altro polinomio g_2 . Allora esiste un poli-

polinomio q tale che $g_1 = qg_2$. Essendo poi

$$\deg g_1 = \deg q + \deg g_2,$$

segue che $\deg g_2 < \deg g_1$. Scambiando il ruolo di g_1 con quello di g_2 si conclude che deve essere necessariamente

$$\deg g_2 = \deg g_1.$$

Quindi q è costante e possiamo scrivere

$$g_1 = cg_2$$

per un'opportuna costante c . Scrivendo

$$g_2(t) = a_n t^n + \dots + a_0$$

dove $a_n \neq 0$, ponendo $b = a_n^{-1}$, si vede che bg_2 è ancora un generatore di J e ha primo coefficiente uguale a 1. È chiaro inoltre che un siffatto generatore è univocamente determinato.

Siano f, g due polinomi non nulli. Diciamo che g divide f , e scriviamo $g|f$ se esiste un polinomio q tale che $f = gg$. Siano f_1, f_2 polinomi non nulli. Definiamo *massimo comun divisore* di f_1, f_2 un polinomio g che divide entrambi i polinomi f_1 e f_2 e inoltre tale che, se h divide f_1 e f_2 , allora h divide anche g .

TEOREMA 3 *Siano f_1, f_2 due polinomi non nulli in $K[t]$. Sia g un generatore dell'ideale generato da f_1, f_2 , allora g è un massimo comune divisore di f_1 e f_2 .*

Dimostrazione. Poiché f_1 appartiene all'ideale generato da f_1, f_2 , deve esistere un polinomio q_1 tale che

$$f_1 = q_1 g,$$

perciò g divide f_1 . Per la stessa ragione, g divide anche f_2 . Sia per esempio

$$f_1 = h_1 h \quad \text{e} \quad f_2 = h_2 h$$

per opportuni polinomi h_1 e h_2 . Poiché g appartiene all'ideale generato da f_1, f_2 , esistono due polinomi g_1, g_2 tali che $g = g_1 f_1 + g_2 f_2$ da cui

$$g = g_1 h_1 h + g_2 h_2 h = (g_1 h_1 + g_2 h_2) h.$$

Perciò h divide g e il nostro teorema è dimostrato.

Osservazione 1. Il massimo comun divisore è determinato a meno di una costante moltiplicativa non nulla. Se sceglieremo un massimo comun divisore con primo coefficiente uguale a 1, questo è univocamente determinato.

Osservazione 2. La stessa dimostrazione vale anche nel caso in cui si tratti di più di due polinomi. Per esempio, se f_1, \dots, f_n sono polinomi non nulli e se g è un generatore dell'ideale generato da f_1, \dots, f_n , allora g è un massimo comune divisore dei polinomi f_1, \dots, f_n .

Se i polinomi f_1, \dots, f_n hanno 1 come massimo comune divisore, essi sono detti *relativamente primi*.

Esercizi

1. Dimostrare che il polinomio $t^n - 1$ è divisibile per $t - 1$.
2. Dimostrare che $t^4 + 4$ può essere scomposto nel prodotto di due polinomi di secondo grado con coefficienti interi.
3. Se n è un numero intero dispari, determinare il quoziente della divisione di $t^n + 1$ per $t + 1$.
4. Sia A una matrice $n \times n$ sul corpo K e sia J l'insieme di tutti i polinomi $f(t)$ in $K[t]$ tali che $f(A) = O$. Dimostrare che J è un ideale.

57. FATTORIZZAZIONE UNICA

Un polinomio p in $K[t]$ verrà detto *irriducibile* (su K) se è di grado maggiore o uguale a 1 e se, per ogni fattorizzazione $p = fg$ con f e g in $K[t]$, il grado di f o quello di g è nullo (cioè, uno dei due polinomi f , g è costante). Quindi, a meno di un fattore costante non nullo, gli unici divisori di p sono 1 e p stesso.

Esempio 1. Gli unici polinomi irriducibili sul corpo dei numeri complessi sono quelli di primo grado, cioè i multipli secondo una costante non nulla dei polinomi del tipo $t - \alpha$, con $\alpha \in \mathbb{C}$.

Esempio 2. Il polinomio $t^2 + 1$ è irriducibile sul corpo reale \mathbb{R} .

TEOREMA 4 *Ogni polinomio in $K[t]$ di grado maggiore di zero può essere espresso come prodotto $p_1 \dots p_m$ di polinomi irriducibili. In questa scomposizione i polinomi p_1, \dots, p_m sono univocamente determinati a meno dell'ordine e di fattori costanti non nulli.*

Dimostrazione. Proviamo dapprima l'esistenza di una fattorizzazione in un prodotto di polinomi irriducibili. Sia f un polinomio di grado maggiore di zero appartenente a $K[t]$. Se f non è già irriducibile, possiamo scrivere

$$f = gh$$

dove $\deg g < \deg f$ e $\deg h < \deg f$. Se g, h sono irriducibili la dimostrazione è finita, in caso contrario fattorizziamo ulteriormente g e h in prodotti di polinomi di grado più basso. Questo processo non può essere continuato indefinitamente e quindi troviamo alla fine una fattorizzazione di f in polinomi irriducibili. (È facile vedere che questa dimostrazione è un procedimento di induzione.)

Resta da provare ora l'unicità della scomposizione. Premettiamo dapprima un lemma:

LEMMA *Sia p un polinomio irriducibile in $K[t]$. Siano f, g polinomi non nulli in $K[t]$. Se p divide il prodotto fg , allora p divide uno dei due fattori f, g .*

Dimostrazione. Supponiamo che p non divida f . Allora il massimo comun divisore di p, f è 1 e quindi in $K[t]$ esistono due polinomi h_1 e h_2 tali che

$$1 = h_1 p + h_2 f.$$

(Questo si vede ricorrendo al teorema 3.) Moltiplicando per g otteniamo

$$g = gh_1 p + h_2 fg.$$

Per ipotesi esiste un polinomio h_3 in modo che $fg = ph_3$, perciò si può scrivere

$$g = (gh_1 + h_2 h_3)p,$$

e si può concludere che p divide g , come si voleva.

Il lemma ora dimostrato sarà applicato nel caso in cui p divide un prodotto di polinomi irriducibili $q_1 \dots q_s$. In questo caso, o p divide q_1 oppure p divide il prodotto $q_2 \dots q_s$. Perciò o esiste una costante c tale che $p = cq_1$ oppure p divide il prodotto $q_2 \dots q_s$. In quest'ultimo caso procediamo per induzione. Pos-

siamo allora concludere che in ogni caso esiste un indice i tale che i polinomi p e q_i differiscono per un fattore costante.

Supponiamo ora di avere due prodotti uguali di polinomi irriducibili

$$p_1 \cdots p_r = q_1 \cdots q_s.$$

Numerando diversamente, se necessario, i polinomi q_i , possiamo ritenere che esista una costante c_1 in modo che $p_1 = c_1 q_1$. Cancellando da entrambi i membri q_1 otteniamo

$$c_1 p_2 \cdots p_r = q_2 \cdots q_s.$$

Procedendo per induzione, concludiamo che, con un'eventuale permutazione di q_1, \dots, q_s , per ogni indice i è possibile trovare una costante c_i tale che $p_i = c_i q_i$. Viene così dimostrata l'unicità richiesta.

COROLLARIO 1 *Sia f un polinomio in $K[t]$ di grado maggiore di zero. Allora è possibile scomporre f in un prodotto $cp_1 \cdots p_s$ dove p_1, \dots, p_s sono polinomi irriducibili con primo coefficiente uguale a 1, univocamente determinati a meno di una permutazione.*

COROLLARIO 2 *Sia f un polinomio in $C[t]$ di grado maggiore di zero. Allora f può essere fattorizzato nel prodotto*

$$f(t) = c(t - \alpha_1) \cdots (t - \alpha_n),$$

con $\alpha_i \in C$ e $c \in C$. I fattori $t - \alpha_i$ sono univocamente determinati a meno di una permutazione.

Nel seguito considereremo principalmente polinomi con primo coefficiente uguale a 1. Sia f un siffatto polinomio con grado maggiore di zero. Siano p_1, \dots, p_r i polinomi irriducibili distinti (con primo coefficiente uguale a 1) che compaiono nella sua scomposizione. Allora possiamo esprimere f come un prodotto

$$f = p_1^{i_1} \cdots p_r^{i_r}$$

dove i_1, \dots, i_r sono interi positivi determinati univocamente dai polinomi p_1, \dots, p_r . Questa fattorizzazione sarà chiamata una fattorizzazione normalizzata di f . In particolare, sul corpo dei numeri complessi, possiamo scrivere

$$f(t) = (t - \alpha_1)^{i_1} \cdots (t - \alpha_r)^{i_r}.$$

Un polinomio con primo coefficiente uguale a 1 viene talvolta detto *monico*.

Se p è irriducibile e $f = p^m g$, dove p non divide g e m è un intero non negativo, diciamo che m è la *molteplicità* di p in f . (Definiamo p^0 uguale a 1.) Inoltre denoteremo questa molteplicità con $\text{ord}_p f$, e la chiameremo anche *ordine* di f in p .

Se α è una radice di f e se

$$f(t) = (t - \alpha)^m g(t),$$

con $g(\alpha) \neq 0$, allora $t - \alpha$ non divide $g(t)$ e m è la molteplicità di $t - \alpha$ in f . Diciamo anche che m è la molteplicità di α in f .

C'è un semplice criterio, che fa uso delle derivate, per stabilire quando m supera 1.

Sia $f(t) = a_n t^n + \dots + a_0$ un polinomio. Definiamo sua derivata (formale) il polinomio

$$Df(t) = f'(t) = n a_n t^{n-1} + (n-1) a_{n-1} t^{n-2} + \dots + a_1.$$

Allora le seguenti proposizioni sono vere e ne lasciamo la dimostrazione come esercizio.

a) Se f, g sono polinomi, allora

$$(f + g)' = f' + g',$$

e anche

$$(fg)' = f'g + fg'.$$

Se c è una costante, allora $(cf)' = cf'$.

b) Sia α una radice di f e si supponga che il grado di f sia positivo. Dimostrare allora che la molteplicità di α in f supera 1 se, e soltanto se, $f'(\alpha) = 0$. Quindi se $f'(\alpha)$ non è nullo, la molteplicità di α è 1.

Esercizi

1. Sia f un polinomio di secondo grado su un corpo K . Dimostrare che o f è irriducibile su K oppure f ha una fattorizzazione in fattori lineari su K .
2. Sia f un polinomio di terzo grado su un corpo K . Se f non è irriducibile su K , dimostrare che f ha una radice in K .
3. Sia $f(t)$ un polinomio irriducibile con primo coefficiente uguale a 1 sul corpo reale. Sia $\deg f = 2$. Dimostrare che $f(t)$ può essere scritto nella

forma

$$f(t) = (t-a)^2 + b^2$$

per opportuni numeri reali a, b e b diverso da zero. Dimostrare, viceversa, che ogni siffatto polinomio è irriducibile sul corpo reale \mathbb{R} .

4. Sia f un polinomio a coefficienti complessi, per esempio

$$f(t) = \alpha_n t^n + \dots + \alpha_0.$$

Si definisca suo complesso coniugato il polinomio

$$\bar{f}(t) = \bar{\alpha}_n t^n + \dots + \bar{\alpha}_0$$

ottenuto mutando ogni coefficiente nel suo complesso coniugato. Dimostrare che se f e g sono in $\mathbb{C}[t]$, allora

$$\overline{(f+g)} = \bar{f} + \bar{g}, \quad \overline{(fg)} = \bar{f}\bar{g},$$

e se $\beta \in \mathbb{C}$ allora $\overline{(\beta f)} = \bar{\beta}f$.

5. Sia $f[t]$ un polinomio a coefficienti reali. Dimostrare che se il numero α è una radice di f anche il suo coniugato $\bar{\alpha}$ è una radice di f .

6. Con riferimento alle notazioni dell'esercizio precedente, dimostrare che α e $\bar{\alpha}$ hanno in f la stessa molteplicità.

7. Sia A una matrice $n \times n$ in un corpo K . Sia J l'insieme dei polinomi f in $K[t]$ tali che $f(A) = O$. Dimostrare che J è un ideale. Il generatore monico di J è chiamato il *polinomio minimo* di A su K . Una definizione analoga viene data se A è un'applicazione lineare di uno spazio vettoriale di dimensione finita V in sé stesso.

8. Sia V uno spazio vettoriale di dimensione finita sul corpo K . Sia $A: V \rightarrow V$ un'applicazione lineare. Sia f il suo polinomio minimo. Dimostrare allora che se A può essere diagonalizzata (cioè se esiste una base di V costituita da autovettori di A), il polinomio minimo è uguale al prodotto

$$(t - \alpha_1) \dots (t - \alpha_r),$$

dove $\alpha_1, \dots, \alpha_r$ sono gli autovalori distinti di A .

9. Dimostrare che i seguenti polinomi non hanno radici multiple in \mathbb{C} .

a) $t^4 + t$.

b) $t^5 - 5t + 1$.

c) Ogni polinomio $t^2 + bt + c$ se b, c sono numeri tali che $b^2 - 4c$ non sia nullo.

10. Dimostrare che il polinomio $t^n - 1$ non ha radici multiple nel corpo complesso \mathbb{C} . Il lettore saprebbe determinarne tutte le radici e darne una fattorizzazione in fattori di primo grado?

11. Siano f, g due polinomi relativamente primi in $K[t]$. Dimostrare che è possibile trovare due polinomi f_1, g_1 tali che il determinante

$$\begin{vmatrix} f & g \\ f_1 & g_1 \end{vmatrix}$$

sia uguale a 1.

12. Siano f_1, f_2, f_3 polinomi in $K[t]$ e si assuma che l'ideale da essi generato sia l'ideale unità. Dimostrare allora che è possibile trovare in $K[t]$ i polinomi f_{ij} in modo che il determinante

$$\begin{vmatrix} f_1 & f_2 & f_3 \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix}$$

sia uguale a 1.

13. Sia α un numero complesso e sia J l'insieme di tutti i polinomi $f(t)$ in $K[t]$ tali che $f(\alpha) = 0$. Dimostrare che J è un ideale. Dimostrare che se J non è l'ideale zero, il generatore monico di J è irriducibile.

14. Siano f, g due polinomi scritti nella forma

$$f = p_1^{i_1} \dots p_r^{i_r}$$

e

$$g = p_1^{j_1} \dots p_r^{j_r}$$

dove i_s, j_s sono interi non negativi, mentre p_1, \dots, p_r sono polinomi irriducibili distinti.

a) Dimostrare che il massimo comune divisore di f e g può essere espresso come prodotto $p_1^{k_1}, \dots, p_r^{k_r}$ dove k_1, \dots, k_r sono interi non negativi. Esprimere k_s mediante i_s e j_s .

b) Si definisca il minimo comune multiplo di polinomi e si esprima il minimo comune multiplo di f e g come prodotto $p_1^{k_1}, \dots, p_r^{k_r}$ dove k_s sono interi non negativi. Esprimere infine k_s mediante i_s e j_s .

15. Scrivere il massimo comune divisore e il minimo comune multiplo delle seguenti coppie di polinomi:

a) $(t-2)^3(t-3)^4(t-i)$ e $(t-1)(t-2)(t-3)^3$.

b) $(t^2+1)(t^2-1)$ e $(t+i)^3(t^3-1)$.

58. GLI INTERI

La teoria della fattorizzazione per i numeri interi è molto vicina alla corrispondente teoria per i polinomi su un corpo. Sia \mathbb{Z} l'insieme di numeri interi relativi. Consideriamo l'algoritmo euclideo.

TEOREMA 1' *Siano m, n interi non negativi, m sia positivo. Allora esistono due interi non negativi q, r , con $0 \leq r < m$, tali che*

$$n = qm + r.$$

Gli interi q, r sono univocamente determinati da queste condizioni.

Dimostrazione. Se $n < m$, basta porre $q = 0$ e $r = n$. Se $n \geq m$ allora $0 \leq n - m < n$. Con un procedimento di induzione possiamo trovare due interi non negativi q_1, r con $r < m$ in modo che

$$n - m = q_1 m + r.$$

Allora

$$n = m + q_1 m + r = (1 + q_1)m + r.$$

E così si prova l'esistenza. L'unicità è lasciata per esercizio.

Definiamo *ideale* J di interi un sottoinsieme di \mathbb{Z} avente le seguenti proprietà:

Il numero 0 appartiene a J . Se m, n appartengono a J , allora anche $m + n$ vi appartiene. Se m appartiene a J , se n è un intero arbitrario, allora il prodotto nm appartiene a J .

Come per i polinomi, definiamo che cosa significa che un ideale è generato dagli interi m_1, \dots, m_n . Anche in questo caso abbiamo l'ideale 0 e l'ideale unità (cioè \mathbb{Z} stesso). Abbiamo allora:

TEOREMA 2' *Sia J un ideale di \mathbb{Z} . Esiste allora un intero d che genera J .*

Dimostrazione. Del tutto simile alla dimostrazione del teorema 2. La lasciamo al lettore come facile esercizio. [Suggerimento: al posto di un polinomio di grado minimo, considerare il più piccolo intero positivo nell'ideale.]

La nozione di divisibilità viene definita come nel caso dei polinomi.

TEOREMA 3' *Siano m_1, m_2 interi positivi. Sia d un generatore positivo dell'ideale generato da m_1, m_2 . Allora d è il massimo comune divisore di m_1 e m_2 .*

Dimostrazione. Lasciamo al lettore il piacere di sviluppare i dettagli. In effetti non si tratta che di copiare, *mutatis mutandis*, la dimostrazione fatta per il teorema 3.

Definiamo p numero primo se p è maggiore di 1 e se, per ogni fattorizzazione $p = mn$ con interi positivi m, n , si ha $m = 1$ oppure $n = 1$.

TEOREMA 4' *Ogni intero positivo n maggiore di 1 può essere espresso come prodotto di numeri primi,*

$$n = p_1 \dots p_r ,$$

univocamente determinati a meno di una permutazione.

Dimostrazione. Copiare la dimostrazione del teorema 4, omettendo le irrilevanti considerazioni inerenti ai fattori costanti.

Esercizi

1. Sia $f(t) = t^n + a_{n-1}t^{n-1} + \dots + a_0$, $a_0 \neq 0$, un polinomio con coefficienti interi. Dimostrare che se α è una radice di f in \mathbb{Z} allora α divide a_0 .

2. Sia $f(t)$ un polinomio a coefficienti interi di grado positivo e avente il primo coefficiente uguale a 1. Dimostrare che ogni radice di f che sia un numero razionale, in effetti è un numero intero.

3. Dimostrare che i seguenti polinomi sono irriducibili sul corpo dei numeri razionali.

- | | | |
|--------------------|---------------------|---------------------|
| a) $t^2 + 1$. | b) $t^2 - 2t + 2$. | c) $t^2 - t + 4$. |
| d) $t^3 - t + 1$. | e) $t^3 + 3t - 1$. | f) $t^3 - 4t + 5$. |

4. Siano a, b due interi relativamente primi. Dimostrare allora che esistono due interi c, d in modo che il determinante

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

sia uguale a 1.

5. (Euclide.) Dimostrare che esistono infiniti numeri primi. [Suggerimento: partendo da n numeri primi distinti p_1, \dots, p_n , se ne costruisca un altro come segue. Si ponga $a = p_1 \dots p_n + 1$ e si dimostri che ogni numero primo p se divide a non può coincidere con nessuno dei numeri p_i .]

6. Enunciare e dimostrare la proprietà dei numeri interi positivi analoga a quella dell'esercizio 14 del paragrafo precedente.

7. Determinare il massimo comune divisore e il minimo comune multiplo delle seguenti coppie di numeri interi positivi.

- | | |
|-----------------------------|--------------|
| a) $5^3 2^6 3$ e $5^2 17$. | b) 248 e 28. |
|-----------------------------|--------------|

59. APPLICAZIONE ALLA SCOMPOSIZIONE DI UNO SPAZIO VETTORIALE

Sia V uno spazio vettoriale sul corpo K e sia $A: V \rightarrow V$ un operatore definito in V . Sia W un sottospazio di V . Diremo che W è un *sottospazio invariante* rispetto ad A se Aw appartiene a W per ogni w in W cioè se AW è contenuto in W .

Esempio 1. Sia v_1 un autovettore non nullo di A e sia V_1 lo spazio unidimensionale generato da v_1 . Allora V_1 è un sottospazio invariante rispetto ad A .

Esempio 2. Sia λ un autovalore di A e sia V_λ il sottospazio di V costituito da tutti i vettori v di V tale che $Av = \lambda v$. Allora V_λ è un sottospazio invariante rispetto ad A , chiamato l'*autospazio* di λ .

Esempio 3. Sia $f(t) \in K[t]$ un polinomio e sia W il nucleo di $f(A)$. Allora W è un sottospazio invariante rispetto ad A .

Dimostrazione. Supponiamo che $f(A)w = O$. Poiché $tf(t) = f(t)t$, possiamo scrivere

$$Af(A) = f(A)A,$$

da cui

$$f(A)(Aw) = f(A)Aw = Af(A)w = O.$$

Quindi anche Aw è nel nucleo di $f(A)$, provando così la nostra affermazione.

Si osservi che, in generale, per ogni coppia di polinomi f, g abbiamo

$$f(A)g(A) = g(A)f(A)$$

perché $f(t)g(t) = g(t)f(t)$. Nel seguito adopereremo frequentemente questa proprietà.

Vogliamo ora far vedere come la fattorizzazione di un polinomio in due fattori di massimo comune divisore uguale a 1 dia luogo a una decomposizione dello spazio vettoriale V in somma diretta di sottospazi invarianti.

TEOREMA 5 *Sia $f(t) \in K[t]$ un polinomio e si supponga che $f = f_1f_2$ dove f_1, f_2 sono polinomi di grado maggiore di zero e aventi un massimo comune divisore uguale a 1. Sia $A: V \rightarrow V$ un opera-*

tore e si supponga che $f(A) = O$. Posto allora

$$W_1 = \text{nucleo di } f_1(A) \quad \text{e} \quad W_2 = \text{nucleo di } f_2(A),$$

lo spazio V risulta somma diretta di W_1 e W_2 .

Dimostrazione. Per l'ipotesi fatta, esistono due polinomi g_1, g_2 in modo che

$$g_1(t)f_1(t) + g_2(t)f_2(t) = 1,$$

e quindi

$$g_1(A)f_1(A) + g_2(A)f_2(A) = I. \quad [1]$$

Sia $v \in V$ allora

$$v = g_1(A)f_1(A)v + g_2(A)f_2(A)v.$$

Il primo termine di questa somma appartiene a W_1 giacché

$$f_2(A)g_1(A)f_1(A)v = g_1(A)f_1(A)f_2(A)v = g_1(A)f(A)v = O.$$

Analogamente si vede che il secondo termine di questa somma appartiene a W_2 . È quindi provato che V è somma di W_1 e W_2 .

Per dimostrare che si tratta di somma diretta, dobbiamo dimostrare che nell'espressione

$$v = w_1 + w_2$$

con $w_1 \in W_1$ e $w_2 \in W_2$, w_1 e w_2 sono univocamente determinati da v . Applicando $g_1(A)f_1(A)$ a entrambi i membri di questa somma troviamo

$$g_1(A)f_1(A)v = g_1(A)f_1(A)w_2,$$

perché $f_1(A)w_1 = O$. Applicando entrambi i membri dell'espressione [1] a w_2 stesso troviamo

$$w_2 = g_1(A)f_1(A)w_2$$

giacché $f_2(A)w_2 = O$. Conseguentemente

$$w_2 = g_1(A)f_1(A)v,$$

e quindi w_2 è univocamente determinato. Analogamente si trova che $w_1 = g_2(A)f_2(A)v$ è univocamente determinato e la somma è quindi diretta. Il teorema è così dimostrato.

Il teorema 5 vale anche nel caso in cui f sia espresso come prodotto di più fattori. Enunciamo questo risultato riferendoci al corpo dei numeri complessi.

TEOREMA 6 *Sia V uno spazio vettoriale sul corpo complesso \mathbb{C} e sia $A: V \rightarrow V$ un operatore. Sia $P(t)$ un polinomio tale che $P(A) = O$, e sia*

$$P(t) = (t - \alpha_1)^{m_1} \dots (t - \alpha_r)^{m_r}$$

la sua fattorizzazione, $\alpha_1, \dots, \alpha_r$ essendo le sue radici distinte. Indicando con U_i il nucleo dell'applicazione $(A - \alpha_i I)^{m_i}$, V risulta somma diretta dei sottospazi U_1, \dots, U_r .

Dimostrazione. Si può procedere per induzione considerando separatamente, uno per uno, i fattori $(t - \alpha_1)^{m_1}, (t - \alpha_2)^{m_2}, \dots$. Quindi otteniamo dapprima una decomposizione come somma diretta del nucleo U_1 di $(A - \alpha_1 I)^{m_1}$, e del nucleo W di

$$(A - \alpha_2 I)^{m_2} \dots (A - \alpha_r I)^{m_r}.$$

Ora, per induzione, possiamo ritenere che W sia espresso come somma diretta

$$W = U_2 \oplus \dots \oplus U_r$$

dove U_j ($j = 2, \dots, r$) è il nucleo di $(A - \alpha_j I)^{m_j}$ in W . Allora

$$V = U_1 \oplus U_2 \oplus \dots \oplus U_r$$

è una somma diretta. Rimane da provare che U_j ($j = 2, \dots, r$) è il nucleo di $(A - \alpha_j I)^{m_j}$ in V . Sia

$$v = u_1 + u_2 + \dots + u_r$$

un elemento di V , dove $u_i \in U_i$, appartenente al nucleo di $(A - \alpha_j I)^{m_j}$, allora, in particolare, v appartiene anche al nucleo di

$$(A - \alpha_2 I)^{m_2} \dots (A - \alpha_r I)^{m_r},$$

ne segue che v deve appartenere a W e conseguentemente che u_1 è nullo. Poiché v appartiene a W , possiamo concludere ora che v coincide con u_j giacché W è la somma diretta di U_2, \dots, U_r .

Esempio 4. Sia V lo spazio vettoriale delle soluzioni (infinitamente differenziabili) dell'equazione differenziale

$$D^n f + a_{n-1} D^{n-1} f + \dots + a_0 f = 0,$$

con coefficienti complessi costanti a_i . Si ponga

$$P(t) = t^n + a_{n-1} t^{n-1} + \dots + a_0.$$

Scomponiamo $P(t)$ in fattori come nel teorema 6, ottenendo

$$P(t) = (t - \alpha_1)^{m_1} \dots (t - \alpha_r)^{m_r}.$$

Allora V è somma diretta degli spazi delle soluzioni delle singole equazioni differenziali

$$(D - \alpha_i I)^{m_i} f = 0,$$

per $i = 1, \dots, r$. Quindi lo studio dell'originaria equazione differenziale è ricondotto allo studio dell'equazione molto più semplice

$$(D - \alpha I)^m f = 0.$$

Le soluzioni di questa equazione si trovano facilmente. Per ogni numero complesso α , infatti, abbiamo

$$(D - \alpha I)^m f = e^{\alpha t} D^m (e^{-\alpha t} f).$$

(La dimostrazione si consegna facilmente per induzione.) Conseguentemente, f appartiene al nucleo di $(D - \alpha I)^m$ se, e soltanto se,

$$D^m (e^{-\alpha t} f) = 0.$$

Le uniche funzioni la cui derivata m -esima è nulla, sono i polinomi di grado inferiore a m . Quindi lo spazio delle soluzioni dell'equazione $(D - \alpha I)^m f = 0$ è quello generato dalle funzioni

$$e^{\alpha t}, t e^{\alpha t}, \dots, t^{m-1} e^{\alpha t}.$$

Si verifica facilmente che queste funzioni sono linearmente indipendenti e quindi che lo spazio delle soluzioni ha dimensione finita, precisamente m .

Esercizio

- Con riferimento al teorema 5, dimostrare che l'immagine di $f_1(A)$ coincide con il nucleo di $f_2(A)$.

60. IL LEMMA DI SCHUR

Sia V uno spazio vettoriale sul corpo K e sia S un insieme di operatori definiti su V . Sia W un sottospazio di V . Diremo che W è un sottospazio S -invariante se, per ogni operatore B in S , BW è contenuto in W . Diremo poi che V è un S -spazio semplice se V non consiste del solo vettore nullo e se gli unici sottospazi S -invarianti sono V stesso e il sottospazio costituito dal solo zero.

Osservazione 1. Sia $A: V \rightarrow V$ un operatore tale che $AB = BA$ per ogni B appartenente a S . Allora l'immagine e il nucleo di A sono sottospazi di V S -invarianti.

Dimostrazione. Sia w un elemento dell'immagine di A . Esiste cioè un vettore v in V in modo che $w = Av$. Allora $Bw = BA v = ABv$. Questo prova che anche B è nell'immagine di A e quindi che l'immagine di A è S -invariante. Se u è un elemento del nucleo di A , allora $ABu = BAu = O$. Perciò Bu è nel nucleo che quindi viene a essere un sottospazio S -invariante.

Osservazione 2. Consideriamo nuovamente S e sia $A: V \rightarrow V$ un operatore. Si supponga che, per ogni $B \in S$, AB e BA coincidano. Se f è un polinomio in $K[t]$, allora $f(A)B = Bf(A)$, per ogni $B \in S$. La dimostrazione è un semplice esercizio.

TEOREMA 7 *Sia V uno spazio vettoriale sul corpo K e sia S un insieme di operatori definiti su V . Si supponga V un S -spazio semplice. Se $A: V \rightarrow V$ è un'applicazione lineare tale che $AB = BA$ per ogni B in S , allora o A è invertibile o A è l'applicazione nulla.*

Dimostrazione. Sia $A \neq O$. Per l'osservazione 1, il nucleo di A è costituito dal solo $\{O\}$ e la sua immagine è l'intero spazio V . L'operatore A è quindi invertibile.

TEOREMA 8 *Sia V uno spazio vettoriale di dimensione finita sul corpo dei numeri complessi. Sia S un insieme di operatori definiti su V e si supponga V un S -spazio semplice. Se $A: V \rightarrow V$ è un'applicazione lineare tale che $AB = BA$ per ogni $B \in S$, allora esiste un numero λ tale che $A = \lambda I$.*

Dimostrazione. Sia J l'ideale dei polinomi f in $\mathbb{C}[t]$ tali che $f(A) = O$. Sia g un generatore di questo ideale, avente primo coefficiente uguale a 1. Allora g non è nullo e vogliamo provare che è irriducibile. In caso contrario, possiamo scrivere $g = h_1 h_2$ dove i polinomi h_1 e h_2 hanno grado minore del grado di g . Ne segue che $h_1(A) \neq O$ e, per il teorema 7 e le osservazioni 1 e 2, concludiamo che $h_1(A)h_2(A)$ risulta invertibile e questo è impossibile: rimane così provato che g deve essere irriducibile. Ma gli unici polinomi irriducibili nel corpo complesso sono quelli di primo grado e perciò deve esistere in \mathbb{C} un λ in modo che $g(t) = t - \lambda$. Ricordando poi che $g(A) = O$, concludiamo che $A - \lambda I = O$ da cui, come dovevamo dimostrare, $A = \lambda I$.

Esercizi

1. Sia V uno spazio vettoriale di dimensione finita sul corpo K e sia S l'insieme di tutte le applicazioni lineari di V in sé stesso. Dimostrare che V è un S -spazio semplice.

2. Sia $V = \mathbb{R}^3$, sia S l'insieme costituito dalla matrice $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ considerata come applicazione lineare di V in sé stesso. Considerando a numero reale non nullo fissato, determinare tutti i sottospazi di V S -invarianti.

3. Sia V uno spazio vettoriale sul corpo K e sia $\{v_1, \dots, v_n\}$ una base di V . Per ogni permutazione σ di $\{1, \dots, n\}$ sia $A_\sigma: V \rightarrow V$ l'applicazione lineare tale che

$$A_\sigma(v_i) = v_{\sigma(i)}.$$

a) Dimostrare che se σ, τ sono permutazioni, allora abbiamo

$$A_\sigma A_\tau = A_{\sigma\tau},$$

e inoltre $A_{\text{id}} = I$.

b) Dimostrare che il sottospazio generato da $v = v_1 + \dots + v_n$ è un sottospazio invariante per l'insieme S_n costituito da tutte le applicazioni A_σ .

c) Dimostrare che il vettore v considerato nella parte b) è un autovettore di ogni A_σ . Qual è l'autovalore di A_σ appartenente a v ?

d) Sia $n = 2$ e sia σ la permutazione non identica. Dimostrare allora che $v_1 - v_2$ genera un sottospazio di dimensione 1 che è invariante rispetto ad A_σ . Dimostrare poi che $v_1 - v_2$ è un autovettore di A_σ . Qual è il suo autovalore?

4. Sia V uno spazio vettoriale sul corpo K e sia $A: V \rightarrow V$ un operatore. Supponiamo che esista un intero positivo r tale che $A^r = I$. Sia $T = I + A + \dots + A^{r-1}$ e sia v_0 un elemento di V . Dimostrare che lo spazio generato da Tv_0 è un sottospazio invariante rispetto ad A e che Tv_0 è un autovettore di A . Se $Tv_0 \neq O$, qual è il suo autovalore?

5. Sia V uno spazio vettoriale sul corpo K e sia S un insieme di operatori definiti su V . Siano U, W sottospazi di V S -invarianti. Dimostrare che anche i sottospazi $U+W$ e $U \cap W$ sono S -invarianti.

61. SVILUPPI α -ADICI DI UN POLINOMIO

TEOREMA 9 *Sia $f(t)$ un polinomio in $K[t]$ e sia $\alpha \in K$. Allora in K esistono delle costanti c_0, \dots, c_n in modo che*

$$f(t) = c_0 + c_1(t-\alpha) + \dots + c_n(t-\alpha)^n,$$

e queste costanti c_0, \dots, c_n sono univocamente determinate.

Dimostrazione. Possiamo supporre $f \neq 0$. Sia $c_0 = f(\alpha)$. Allora il polinomio $f(t) - c_0$ ha α per radice e quindi può essere scritto nella forma

$$f(t) - c_0 = f_1(t)(t-\alpha),$$

oppure

$$f(t) = c_0 + f_1(t)(t-\alpha).$$

Ripetiamo ora la stessa argomentazione nei riguardi di f_1 . Sia $c_1 = f_1(\alpha)$ e scriviamo

$$f_1(t) = c_1 + f_2(t)(t-\alpha).$$

Allora sostituendo f_1 nella precedente espressione di f otteniamo

$$f(t) = c_0 + c_1(t-\alpha) + f_2(t)(t-\alpha).$$

Ripetendo questo procedimento tante volte quanto è il grado di f , otteniamo

$$f(t) = c_0 + c_1(t-\alpha) + \dots + f_n(t)(t-\alpha)^n.$$

Osservando il secondo membro concludiamo che il grado di f_n è zero e quindi $f_n(t)$ è una costante. Così si prova l'esistenza della scomposizione. Per provare l'unicità, supponiamo che f possa scriversi anche come segue

$$f(t) = b_0 + b_1(t-\alpha) + \dots + b_n(t-\alpha)^n$$

dove b_i sono le costanti appartenenti a K . Supponiamo anche che non per tutti gli indici accada che $a_i = b_i$. Sottraendo

membro a membro le due espressioni di f otteniamo

$$0 = a_0 - b_0 + (a_1 - b_1)(t - \alpha) + \dots + (a_n - b_n)(t - \alpha)^n.$$

Sia r il più piccolo intero per cui $a_r - b_r$ non è nullo, possiamo allora scrivere

$$\begin{aligned} 0 &= (a_r - b_r)(t - \alpha)^r + \dots + (a_n - b_n)(t - \alpha)^n = \\ &= (t - \alpha)^r [(a_r - b_r) + \dots + (a_n - b_n)(t - \alpha)^{n-r}]. \end{aligned}$$

Da cui

$$0 = a_r - b_r + \dots + (a_n - b_n)(t - \alpha)^{n-r}.$$

E sostituendo t con α troviamo che $a_r - b_r$ è nullo e concludiamo perciò che un tale r non può esistere, in altre parole, $a_i = b_i$ per ogni i , come volevamo dimostrare.

COROLLARIO *Sia $f(t)$ un polinomio in $K[t]$ e sia $\alpha \in K$. Sia m un intero positivo. Allora in K esistono degli elementi b_1, \dots, b_m e in $K[t]$ esiste un polinomio $g(t)$ in modo che*

$$\frac{f(t)}{(t - \alpha)^m} = \frac{b_m}{(t - \alpha)^m} + \frac{b_{m-1}}{(t - \alpha)^{m-1}} + \dots + \frac{b_1}{(t - \alpha)} + g(t).$$

Dimostrazione. Nell'espressione di f ottenuta nel teorema precedente basta dividere ogni termine per $(t - \alpha)^m$ e fare la somma di quello che si ottiene.

Per esercizio si può facilmente provare che le costanti b_1, \dots, b_m sono univocamente determinate.

L'espressione ottenuta nel teorema 9 è chiamata *sviluppo α -adico* di f . È utile nella decomposizione di frazioni parziali, come ora vedremo.

Un quoziente di polinomi

$$\frac{h(t)}{f(t)}$$

(con $f \neq 0$) è chiamato una *funzione razionale*. Diciamo che si tratta di una funzione razionale in K se i coefficienti di h e di f sono in K . Se esprimiamo h e f come prodotti di polinomi irriducibili, possiamo cancellare in questo quoziente il massimo comun

divisore g di h e f , e se scriviamo

$$h = gh_1, \quad f = gf_1$$

allora

$$h/f = h_1/f_1.$$

E inoltre h_1 e f_1 sono primi tra loro. Un quoziente h_1/f_1 in cui h_1 e f_1 sono relativamente primi viene detto *quoziente in forma ridotta*.

Sia $f = f_1f_2$ un polinomio in $K[t]$, espresso come prodotto di due polinomi f_1, f_2 di grado positivo e primi tra loro. Allora esistono due polinomi g_1, g_2 in modo che

$$g_1f_1 + g_2f_2 = 1.$$

Da cui

$$\frac{1}{f} = \frac{1}{f_1f_2} = \frac{g_1f_1 + g_2f_2}{f_1f_2} = \frac{g_1}{f_2} + \frac{g_2}{f_1}.$$

In questo modo abbiamo ottenuto la decomposizione della frazione $1/f$ nella somma di frazioni i cui denominatori sono rispettivamente f_1 e f_2 .

Esempio 1. Abbiamo

$$\frac{1}{(t-1)(t-2)} = \frac{-1}{t-1} + \frac{1}{t-2},$$

che non è altro che la decomposizione adoperata nel calcolo infinitesimale quando si voglia integrare una funzione razionale.

Sia h un polinomio e sia, come nel caso precedente, $f = f_1f_2$. Allora

$$\frac{h}{f} = \frac{hg_1}{f_2} + \frac{hg_2}{f_1}.$$

Quindi una funzione razionale può anche essere scomposta in una somma di frazioni i cui denominatori sono rispettivamente f_1, f_2 .

TEOREMA 10 *Sia $R(t) = h(t)/f(t)$ una funzione razionale, espressa con un quoziente di due polinomi h e f a coefficienti complessi. Sia*

$$f(t) = (t - \alpha_1)^{m_1} \dots (t - \alpha_r)^{m_r}$$

la fattorizzazione di f dove $\alpha_1, \dots, \alpha_r$ sono le sue radici distinte. Allora in $\mathbb{C}[t]$ esistono dei polinomi h_1, \dots, h_r in modo che

$$R(t) = \frac{h_1(t)}{(t - \alpha_1)^{m_1}} + \dots + \frac{h_r(t)}{(t - \alpha_r)^{m_r}}.$$

Dimostrazione. Poniamo

$$f_1(t) = (t - \alpha_1)^{m_1} \quad \text{e} \quad f_2(t) = (t - \alpha_2)^{m_2} \dots (t - \alpha_r)^{m_r}.$$

Possiamo allora applicare la decomposizione di cui abbiamo trattato prima del teorema, ottenendo

$$R(t) = \frac{h_1(t)}{(t - \alpha_1)^{m_1}} + R_2(t)$$

dove $R_2(t)$ è una funzione razionale il cui denominatore è $f_2(t)$. Possiamo allora ripetere l'argomentazione separando un ulteriore termine

$$\frac{h_2(t)}{(t - \alpha_2)^{m_2}}$$

e concludere la dimostrazione col procedimento di induzione.

Ogni addendo della somma che rappresenta $R(t)$ nel teorema 10 può essere sviluppato come nel corollario del teorema 9. Quello che si ottiene si chiama la decomposizione di R in frazioni parziali. Per trovarle si può ricorrere al solito metodo esplicito della risoluzione di equazioni lineari.

Sia h/f una funzione razionale. Dopo avere applicato l'algoritmo euclideo

$$h = qf + h_1$$

dove $\deg h_1 < \deg f$, otteniamo

$$h/f = q + h_1/f.$$

Abbiamo così ridotto il problema al caso in cui il grado del nominatore sia minore di quello del denominatore. Allora scriviamo la decomposizione in frazioni parziali con coefficienti incogniti e risolviamo rispetto a questi.

Esempio 2. Riprendiamo l'esempio 1. Per trovare i coefficienti, scriviamo

$$\frac{1}{(t-1)(t-2)} = \frac{a}{t-1} + \frac{b}{t-2},$$

e dobbiamo risolvere rispetto ad a e b . Portando il secondo membro a un comune denominatore, otteniamo

$$\frac{1}{(t-1)(t-2)} = \frac{a(t-2) + b(t-1)}{(t-1)(t-2)}.$$

Dall'uguaglianza dei numeratori abbiamo ancora

$$a + b = 0,$$

$$-2a - b = 1.$$

Che ora risolviamo rispetto ad a e b , ottenendo $a = -1$, $b = 1$.

Esercizi

1. Determinare lo sviluppo 2-adico dei polinomi seguenti.

- a) $t^2 - 1$. b) $t^3 + t - 1$. c) $t^2 + 3$. d) $t^4 + 2t^3 - t + 5$.

2. Dei polinomi considerati nell'esercizio 1 trovare lo sviluppo 3-adico.

3. In questo esercizio si può supporre che ogni polinomio irriducibile sul corpo reale abbia grado 1 oppure 2. Sia p un polinomio di secondo grado irriducibile sul corpo reale R . Dimostrare che ogni polinomio f in $R[t]$ può essere scritto nella forma

$$f(t) = c_0 + c'_0 t + (c_1 + c'_1 t)p(t) + \dots + (c_m + c'_m t)p(t)^m$$

dove c_i , c'_i sono numeri reali e m è un intero. Dimostrare che i numeri c_i , c'_i sono univocamente determinati.

4. a) Enunciare e dimostrare nel caso dei numeri reali la proposizione analoga al corollario del teorema 9.

b) Enunciare e dimostrare nel caso dei numeri reali l'analogia della decomposizione in frazioni parziali.

5. Si generalizzi come segue il teorema 9. Sia g un fissato polinomio di grado positivo in $K[t]$. Dato un polinomio f in $K[t]$, dimostrare che si può scrivere

$$f = h_0 + h_1 g + \dots + h_n g^n$$

dove h_0, \dots, h_n sono polinomi di grado minore di quello di g . Dimostrare che questi polinomi h_i sono univocamente determinati.

6. Sia a un intero positivo e sia d un intero maggiore di 1. Dimostrare che esistono gli interi c_0, \dots, c_n tali che $0 < c_i < d - 1$ e

$$a = c_0 + c_1 d + \dots + c_n d^n.$$

Dimostrare che questi numeri c_i sono univocamente determinati. L'espressione di a ora ottenuta viene detta suo sviluppo d -adico. Il lettore riconosce in ciò qualcosa di familiare del caso $d = 10$?

7. Determinare lo sviluppo 2-adico di 25, 100 e 293.

8. Determinare lo sviluppo 3-adico dei numeri 25, 100 e 293.

9. Sia a un numero razionale positivo e sia $a = b/c$, b e c essendo interi positivi. Sia $c = p_1^{m_1} \dots p_r^{m_r}$ la fattorizzazione di c in fattori primi distinti. Dimostrare allora che esistono gli interi non negativi b_1, \dots, b_r in modo che

$$a = \frac{b_1}{p_1^{m_1}} + \dots + \frac{b_r}{p_r^{m_r}}.$$

10. Sia K un corpo. Si definisca un polinomio in n variabili su K come una funzione f di n variabili che può essere scritta nella forma

$$f(t_1, \dots, t_n) = \sum c_{i_1 \dots i_n} t_1^{i_1} \dots t_n^{i_n}$$

dove la somma è estesa a tutte le n -uple di interi (i_1, \dots, i_n) tali che $0 < i_e \in K$, $c_{(i)}$ è nullo con l'eccezione di un numero finito di n -uple (i) .

a) Dimostrare che ogni siffatto polinomio può essere scritto nella forma

$$f(t_1, \dots, t_n) = f_d(t_1, \dots, t_{n-1}) t_n^d + \dots + f_0(t_1, \dots, t_{n-1})$$

dove f_0, \dots, f_d sono polinomi in $n-1$ variabili.

b) Dimostrare col procedimento di induzione che se $f(t_1, \dots, t_n) = 0$ comunque si prendano in K t_1, \dots, t_n , allora ogni $c_{(i)}$ è nullo.

Capitolo 13

Prodotti multilinearari

62. PRODOTTO TENSORIALE

Siano V, W spazi vettoriali su un corpo K . Vogliamo definire un nuovo tipo di prodotto tra elementi di V e di W . I valori che questo prodotto assume devono essere in uno spazio vettoriale e, detto alla buona, vogliamo che non vi siano tra essi altre relazioni se non quelle bilineari. In altre parole, se denotiamo con $v \otimes w$ il prodotto degli elementi $v \in V$ e $w \in W$, allora vorremmo avere soltanto le seguenti relazioni:

Se v_1, v_2 sono in V , se w è in W , allora

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w.$$

Se w_1, w_2 sono in W , se v è in V , allora

$$v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2.$$

Se $c \in K$, allora

$$(cv) \otimes w = c(v \otimes w) = v \otimes cw.$$

Costruiremo ora un prodotto siffatto e ne dimostreremo diverse proprietà.

Siano U, V, W spazi vettoriali sul corpo K . Per *applicazione bilineare*

$$g: V \times W \rightarrow U$$

noi intendiamo un'applicazione che ad ogni coppia di elementi

(v, w) con v in V e w in W associa un elemento $g(v, w)$ di U avente la seguente proprietà:

Per ogni v in V , l'applicazione di W in U definita da $w \mapsto g(v, w)$ è lineare e, per ogni $w \in W$, l'applicazione di V in U definita da $v \mapsto g(v, w)$ è lineare.

Vediamo quindi che un'applicazione bilineare è definita in modo del tutto simile alle forme bilineari, l'unica differenza consiste nel fatto che ammettiamo che i valori dell'applicazione possono essere in uno spazio vettoriale anziché nel corpo K .

TEOREMA 1 Siano V, W spazi vettoriali di dimensione finita sul corpo K . Esiste allora uno spazio vettoriale T di dimensione finita su K e un'applicazione bilineare $V \times W \rightarrow T$ denotata con

$$(v, w) \mapsto v \otimes w,$$

soddisfacente le seguenti proprietà.

PT 1. Se U è uno spazio vettoriale su K e $g: V \times W \rightarrow U$ è un'applicazione bilineare, allora esiste un'unica applicazione lineare

$$g_*: T \rightarrow U$$

tale che, per ogni coppia (v, w) con v in V e w in W , si abbia

$$g(v, w) = g_*(v \otimes w).$$

PT 2. Se $\{v_1, \dots, v_n\}$ è una base di V e se $\{w_1, \dots, w_m\}$ è una base di W , allora gli elementi

$$v_i \otimes w_j \quad (i = 1, \dots, n \text{ e } j = 1, \dots, m)$$

costituiscono una base di T .

Dimostrazione. Sia $\{v_1, \dots, v_n\}$ una base di V e sia $\{w_1, \dots, w_m\}$ una base di W . Per ogni coppia (i, j) con $1 \leq i \leq n$ e $1 \leq j \leq m$, sia t_{ij} una lettera. Come è spiegato nell'appendice di questo capitolo, definiamo T come lo spazio vettoriale su K costituito da tutte le combinazioni lineari formali di questi elementi t_{ij} a coefficienti in K , in tal modo questi elementi costituiscono una base di T su K . Quindi gli elementi di T sono le combinazioni lineari

$$\sum_{i=1}^n \sum_{j=1}^m c_{ij} t_{ij}$$

dove $c_{ij} \in K$.

Se $v = x_1 v_1 + \dots + x_n v_n$ e $w = y_1 w_1 + \dots + y_m w_m$, con x_i, y_j in K , definiamo $v \otimes w$ come l'elemento

$$v \otimes w = \sum_{i=1}^n \sum_{j=1}^m x_i y_j t_{ij}$$

di T . In particolare, $v_i \otimes w_j = t_{ij}$. Vogliamo ora dimostrare che il nostro prodotto $v \otimes w$ ha tutti i requisiti richiesti.

Dimostrazione di PT 1. Per semplicità abbreviamo le somme

$$\sum_{i=1}^n \sum_{j=1}^m \quad \text{con} \quad \sum_i \sum_j .$$

Dimostriamo dapprima la bilinearità dell'applicazione definita da $(v, w) \mapsto v \otimes w$. Sia

$$v' = x'_1 v_1 + \dots + x'_n v_n$$

e siano v e w espressi come combinazioni lineari degli elementi delle rispettive basi sopra menzionate. Allora

$$v + v' = (x_1 + x'_1)v_1 + \dots + (x_n + x'_n)v_n .$$

Per definizione,

$$\begin{aligned} (v + v') \otimes w &= \sum_i \sum_j (x_i + x'_i) y_j t_{ij} = \\ &= \sum_i \sum_j (x_i y_j + x'_i y_j) t_{ij} = \\ &= \sum_i \sum_j (x_i y_j t_{ij} + x'_i y_j t_{ij}) = \\ &= \sum_i \sum_j x_i y_j t_{ij} + \sum_i \sum_j x'_i y_j t_{ij} = \\ &= v \otimes w + v' \otimes w . \end{aligned}$$

La dimostrazione della distributività dall'altro lato è del tutto simile e la omettiamo. Se $c \in K$, allora

$$\begin{aligned} (cv) \otimes w &= \sum_i \sum_j (cx_i) y_j t_{ij} = \\ &= \sum_i \sum_j cx_i y_j t_{ij} = \\ &= c \sum_i \sum_j x_i y_j t_{ij} = \\ &= c(v \otimes w) . \end{aligned}$$

Così si dimostra che il nostro prodotto \otimes è bilineare.

Sia $g: V \times W \rightarrow U$ un'applicazione lineare. Ricordando il teorema che afferma che si possono prescrivere arbitrariamente i valori di un'applicazione lineare sugli elementi di una base (teorema 1, cap. 4, § 17) concludiamo che esiste un'unica applicazione lineare

$$g_*: T \rightarrow U$$

tale che

$$g_*(t_{ij}) = g(v_i, w_j).$$

Allora per ogni scelta dei vettori v, w , espressi come si è già detto come combinazioni lineari di elementi delle due basi scelte, si ha

$$\begin{aligned} g(v, w) &= g\left(\sum_i x_i v_i, \sum_j y_j w_j\right) = \\ &= \sum_i \sum_j x_i y_j g(v_i, w_j) = \\ &= g_*(v \otimes w). \end{aligned}$$

Quindi l'applicazione che cercavamo g_* esiste ed è univocamente determinata.

Dimostrazione di PT 2. Siano $\{v'_1, \dots, v'_n\}$ e $\{w'_1, \dots, w'_m\}$ basi rispettive di V e di W . Dobbiamo provare che gli elementi $v'_i \otimes w'_j$ costituiscono una base di T . Gli elementi generici $v \in V$ e $w \in W$ possono essere scritti come combinazioni lineari

$$v = x'_1 v'_1 + \dots + x'_n v'_n$$

e

$$w = y'_1 w'_1 + \dots + y'_m w'_m$$

con x'_i e y'_j in K . Allora

$$v \otimes w = \sum_i \sum_j x'_i y'_j (v'_i \otimes w'_j).$$

Quindi gli elementi

$$v'_i \otimes w'_j$$

generano T su K . Questi elementi sono in numero di mn e, se fossero linearmente dipendenti, la dimensione di T risulterebbe minore di mn , in contraddizione col fatto che gli elementi t_{ij} costituiscono una base di T . Questo dimostra il nostro teorema.

Lo spazio T trovato nel teorema 1 viene chiamato il *prodotto tensoriale* di V e W ed è denotato con $V \otimes W$. Osserviamo che la sua dimensione è data da

$$\dim(V \otimes W) = (\dim V)(\dim W).$$

L'elemento $v \otimes w$ associato alla coppia (v, w) è anche chiamato *prodotto tensoriale* di v e w .

Esercizi

1. Siano V, W spazi vettoriali di dimensione finita sul corpo K . Sia $V \otimes W \rightarrow U$ un'applicazione lineare. Dimostrare che l'applicazione definita da

$$(v, w) \mapsto F(v \otimes w)$$

è un'applicazione bilineare di $V \times W$ in U .

2. Dimostrare che la corrispondenza $g \mapsto g_*$ stabilita nel teorema 1 è un isomorfismo tra lo spazio delle applicazioni bilineari di $V \times W$ in U e lo spazio delle applicazioni lineari $\mathcal{L}(V \otimes W, U)$.

3. Sia V uno spazio vettoriale di dimensione finita sul corpo K . Sia $A: V \rightarrow V$ un'applicazione lineare. Dimostrare che esiste un'unica applicazione lineare: $F: V \otimes V \rightarrow V \otimes V$ tale che

$$F(v \otimes w) = Av \otimes Aw$$

per ogni scelta di v, w in V . Questa applicazione viene denotata con $A \otimes A$.

4. Generalizzare l'esercizio precedente al caso del prodotto tensoriale $V \otimes W$. Siano $A: V \rightarrow V$ e $B: W \rightarrow W$ due applicazioni lineari. Far vedere come si può definire l'applicazione lineare

$$A \otimes B: V \otimes W \rightarrow V \otimes W.$$

63. ISOMORFISMI DI PRODOTTI TENSORIALI

Capita di frequente di voler considerare il prodotto tensoriale di più di due spazi. Vogliamo far vedere che il prodotto è associativo.

TEOREMA 2 *Siano U, V, W spazi vettoriali di dimensione finita sul corpo K . Esiste allora un unico isomorfismo*

$$U \otimes (V \otimes W) \rightarrow (U \otimes V) \otimes W$$

tale che

$$u \otimes (v \otimes w) \mapsto (u \otimes v) \otimes w$$

per ogni scelta di u in U , di v in V , di w in W .

Dimostrazione. Siano $\{u_i\}$, $\{v_j\}$, $\{w_k\}$ basi rispettive di U , V , W . Allora gli elementi

$$(u_i \otimes v_j) \otimes w_k$$

costituiscono una base di $(U \otimes V) \otimes W$, mentre gli elementi

$$u_i \otimes (v_j \otimes w_k)$$

costituiscono una base di $U \otimes (V \otimes W)$. Dal teorema generale sull'esistenza e unicità di applicazioni lineari, deduciamo che esiste un'unica applicazione lineare

$$F: U \otimes (V \otimes W) \rightarrow (U \otimes V) \otimes W$$

che associa $(u_i \otimes v_j) \otimes w_k$ a $u_i \otimes (v_j \otimes w_k)$. Si verifica facilmente con il solito sviluppo in combinazioni lineari che per ogni elemento $u \otimes (v \otimes w)$ di $U \otimes (V \otimes W)$ l'applicazione F si comporta come vogliamo. Poiché F associa una base di $U \otimes (V \otimes W)$ a una base di $(U \otimes V) \otimes W$, segue che F è un isomorfismo.

Il teorema 2 permette di omettere le parentesi nell'indicazione del prodotto tensoriale di diversi fattori. Quindi, se V_1, \dots, V_r sono spazi vettoriali su K , noi possiamo considerare il loro prodotto tensoriale

$$V_1 \otimes V_2 \otimes \dots \otimes V_r,$$

e il prodotto tensoriale

$$v_1 \otimes v_2 \otimes \dots \otimes v_r$$

di elementi v_i in V_i .

I teoremi 1 e 2 danno proprietà generali dei prodotti tensoriali che sono spesso utili. Esistono però altri interessanti isomorfismi che possono essere considerati a proposito di prodotti tensoriali. Qui ne mostreremo soltanto uno, frequentemente usato nei calcoli che intervengono in geometria differenziale.

TEOREMA 3 *Sia V uno spazio vettoriale di dimensione finita sul corpo K . Sia V^* il suo spazio duale e sia $\mathcal{L}(V, V)$ lo spazio delle applicazioni lineari di V in sé stesso. Esiste allora un unico isomorfismo*

$$V^* \otimes V \rightarrow \mathcal{L}(V, V),$$

che associa ad ogni elemento $\varphi \otimes v$ (dove $\varphi \in V^*$ e $v \in V$) l'applicazione lineare $L_{\varphi \otimes v}$ tale che

$$L_{\varphi \otimes v}(w) = \varphi(w)v.$$

Dimostrazione. Associamo ad ogni coppia (φ, v) nel prodotto diretto $V^* \times V$ l'applicazione lineare $L_{\varphi, v}$ tale che

$$L_{\varphi, v}(w) = \varphi(w)v.$$

Si verifica subito che l'associazione

$$(\varphi, v) \mapsto L_{\varphi, v}$$

è un'applicazione bilineare di $V^* \times V$ in $\mathcal{L}(V, V)$. Conseguentemente, per il teorema 1, esiste un'unica applicazione lineare di $V^* \otimes V$ in $\mathcal{L}(V, V)$ che ad ogni elemento $\varphi \otimes v$ associa la nostra applicazione lineare $L_{\varphi, v}$. Dobbiamo ora dimostrare che l'applicazione definita da

$$\varphi \otimes v \mapsto L_{\varphi, v}$$

è un isomorfismo di $V^* \otimes V$ in $\mathcal{L}(V, V)$. Sia $\{v_1, \dots, v_n\}$ una base di V e sia $\{\varphi_1, \dots, \varphi_n\}$ la sua base duale. Allora $\varphi_i(v_k)$ è nullo se gli indici i e k sono diversi, è invece uguale a 1 se gli indici coincidono. Per semplificare le notazioni, poniamo

$$L_{ij} = L_{\varphi_i, v_j}.$$

Vogliamo ora far vedere che gli elementi L_{ij} ($i = 1, \dots, n$ e $j = 1, \dots, n$) sono linearmente indipendenti. Si supponga che esista una relazione lineare

$$\sum_j \sum_i c_{ij} L_{ij} = O$$

dove $c_{ij} \in K$. Applicando il primo membro ad ogni v_k otteniamo

$$0 = \sum_j \sum_i c_{ij} L_{ij}(v_k).$$

In questa somma, $L_{ij}(v_k)$ è nullo eccetto il caso in cui $i = k$ e allora è uguale a v_j . Abbiamo perciò

$$O = \sum_j c_{kj} v_j.$$

Possiamo concludere, a causa dell'indipendenza lineare di v_1, \dots, v_n , che c_{kj} è nullo per ogni j e per ogni k : provando così che le applicazioni lineari L_{ij} sono linearmente indipendenti. Il numero di queste applicazioni è n^2 e la dimensione dello spazio $\mathcal{L}(V, V)$ è precisamente n^2 . Le applicazioni L_{ij} , pertanto, costituiscono una base di $\mathcal{L}(V, V)$. Poiché l'applicazione

$$V^* \otimes V \rightarrow \mathcal{L}(V, V)$$

associa una base $\{\varphi_i \otimes v_j\}$ di $V^* \otimes V$ a una base $\{L_{ij}\}$ di $\mathcal{L}(V, V)$, si può concludere, come volevamo, che questa applicazione è un isomorfismo.

Esercizi

1. Siano V, W spazi vettoriali di dimensione finita sul corpo K . Dimostrare che esiste un unico isomorfismo di $V \otimes W$ su $W \otimes V$ che, per ogni v in V e ogni w in W , associa $v \otimes w$ a $w \otimes v$.
2. Siano V, W come nell'esercizio precedente. Dimostrare che esiste un unico isomorfismo

$$V^* \otimes W \rightarrow \mathcal{L}(V, W)$$

definito in modo che

$$\varphi \otimes w \mapsto L_{\varphi, w},$$

dove $L_{\varphi, w}$ è un'applicazione lineare tale che $L_{\varphi, w}(v) = \varphi(v)w$.

3. Siano V, W come nell'esercizio 1. Dimostrare che esiste un unico isomorfismo

$$V^* \otimes W^* \rightarrow (V \otimes W)^*$$

che ad ogni prodotto tensoriale $\varphi \otimes \psi$ ($\varphi \in V^*$ e $\psi \in W^*$) associa un funzionale $L_{\varphi, \psi}$ di $V \otimes W$ avente la proprietà

$$L_{\varphi, \psi}(v \otimes w) = \varphi(v)\psi(w).$$

Descrivere questo isomorfismo in termini di basi e basi duali.

64. PRODOTTI ALTERNANTI: CASO PARTICOLARE

Vogliamo ora considerare un'altra specie di prodotto, continuamente adoperato nella teoria delle forme differenziali che si incontra nel calcolo infinitesimale. Poiché questa nozione, in tutta la sua generalità, non risulta molto semplice, dedicheremo un po'

di tempo allo studio di un caso particolare che si riduce al prodotto vettoriale tra vettori. Comunque, ne tratteremo in modo da poterlo generalizzare al caso di uno spazio di dimensione qualsiasi.

Sia V uno spazio vettoriale di dimensione 3 sul corpo K . Sia $f: V \times V \rightarrow U$ un'applicazione bilineare di $V \times V$ in un certo spazio vettoriale U sul corpo K . Diremo che f è *alternante* quando, per ogni vettore v in V , $f(v, v)$ è nullo. (Questa condizione somiglia molto a quella che abbiamo già incontrato studiando i determinanti.) Vogliamo definire un prodotto tra elementi di V aventi valore in uno spazio vettoriale e tale che le uniche relazioni cui esso deve soddisfare sono quelle dovute alla bilinearità e all'alternanza, cioè, se denotiamo questo prodotto tra due elementi di V con $v \wedge w$, allora dobbiamo avere $v \wedge v = 0$. Supponiamo di aver trovato un tale prodotto. Allora, comunque si prendano in V gli elementi v, w , si ha

$$(v + w) \wedge (v + w) = 0.$$

A causa della bilinearità, il primo membro si può anche scrivere

$$v \wedge v + w \wedge v + v \wedge w + w \wedge w.$$

Quindi la relazione di alternanza implica che

$$v \wedge w = -w \wedge v. \quad [1]$$

Sia ora $\{v_1, v_2, v_3\}$ una base di V . Comunque si prendano in V due elementi v, w , questi possono essere scritti come combinazione lineare degli elementi di questa base, sia per esempio

$$v = x_1 v_1 + x_2 v_2 + x_3 v_3,$$

$$w = y_1 v_1 + y_2 v_2 + y_3 v_3.$$

Se ora $f: V \times V \rightarrow U$ è un'applicazione bilineare alternante, possiamo scrivere, a causa della bilinearità,

$$\begin{aligned} f(v, w) &= x_1 y_1 f(v_1, v_1) + x_1 y_2 f(v_1, v_2) + x_1 y_3 f(v_1, v_3) + \\ &\quad + x_2 y_1 f(v_2, v_1) + x_2 y_2 f(v_2, v_2) + x_2 y_3 f(v_2, v_3) + \\ &\quad + x_3 y_1 f(v_3, v_1) + x_3 y_2 f(v_3, v_2) + x_3 y_3 f(v_3, v_3) = \\ &= \sum_i \sum_j x_i y_j f(v_i, v_j). \end{aligned}$$

Tenendo conto poi delle relazioni d'alternanza, vediamo che tre

addendi sono nulli e che gli altri possono essere espressi come combinazioni lineari di $f(v_1, v_2)$, $f(v_1, v_3)$, $f(v_2, v_3)$. Basta osservare semplicemente che

$$f(v_2, v_1) = -f(v_1, v_2), \quad f(v_3, v_1) = -f(v_1, v_3),$$

e che

$$f(v_3, v_2) = -f(v_2, v_3).$$

Perciò

$$\begin{aligned} f(v, w) &= (x_1 y_2 - x_2 y_1) f(v_1, v_2) + (x_1 y_3 - x_3 y_1) f(v_1, v_3) \\ &\quad + (x_2 y_3 - x_3 y_2) f(v_2, v_3). \end{aligned} \quad [2]$$

In conclusione, se abbiamo un'applicazione bilineare alternante f , lo spazio generato da tutti i valori $f(v, w)$ per ogni v, w in V , viene ad avere dimensione non superiore a 3.

Vogliamo far vedere ora che effettivamente *esiste un prodotto alternante definito in $V \times V$ a valori in uno spazio denotato con $V \wedge V$, generato da tutti i prodotti $v \wedge w$ con v e w scelti in V , e tale che $V \wedge V$ abbia dimensione precisamente 3.*

Fissiamo tre lettere t_{12} , t_{23} , t_{13} come elementi di una base di questo spazio. Se v, w sono elementi di V , espressi, come sopra si è detto, mediante gli elementi della base v_1, v_2, v_3 , definiamo loro prodotto $v \wedge w$ l'espressione

$$(x_1 y_2 - x_2 y_1) t_{12} + (x_1 y_3 - x_3 y_1) t_{13} + (x_2 y_3 - x_3 y_2) t_{23}.$$

Osserviamo che, se i è minore di j , $v_i \wedge v_j$ coincide con t_{ij} . È molto facile ora verificare con un semplice calcolo che il prodotto da noi definito è bilineare e alternante. La relazione di alteranza è particolarmente semplice da dimostrare giacché se $v = w$, ogni coefficiente nello sviluppo del prodotto è del tipo

$$x_i x_j - x_j x_i = 0.$$

Sia $f: V \times V \rightarrow U$ un'applicazione bilineare alternante. Allora esiste un'unica applicazione lineare

$$f_*: V \wedge V \rightarrow U$$

tale che, per ogni coppia (v, w) di elementi di V , si abbia

$$f(v, w) = f_*(v \wedge w).$$

Dimostrazione. Per il teorema relativo all'esistenza e unicità di un'applicazione lineare avente valori assegnati sugli elementi di una base, sappiamo che esiste un'unica applicazione lineare $f_*: V \wedge V \rightarrow U$ tale che

$$f_*(t_{ij}) = f(v_i, v_j)$$

per ogni coppia di indici i, j tali che $1 < i < j < 3$. Dalla relazione [2] concludiamo immediatamente che $f(v, w)$ coincide con $f_*(v \wedge w)$, per ogni scelta di v e w in V .

In conclusione, se lo spazio V ha dimensione 3, abbiamo provato per i prodotti alternanti l'analogo del teorema 1 per le applicazioni bilineari alternanti.

A questo punto si può chiedere se non sia possibile considerare prodotti di più fattori: la risposta è affermativa e se ne tratterà negli esercizi 1, 2, 3.

Esempio 1. La teoria dei prodotti alternanti è molto usata nel calcolo delle forme differenziali e perciò ne vogliamo trattare in questo esempio.

Sia $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ un'applicazione differenziabile. Per definizione, per ogni X in \mathbb{R}^3 , esiste un'applicazione lineare di \mathbb{R}^3 in \mathbb{R} , denominata con $df(X)$ tale che per vettori piccoli H ,

$$f(X + H) = df(X)H + o(H).$$

(La notazione $o(H)$ è presa in prestito dall'analisi e qui non la definiremo.) Quindi df associa ad ogni punto X di \mathbb{R}^3 un funzionale $df(X)$ appartenente a $\mathcal{L}(\mathbb{R}^3, \mathbb{R})$. Ponendo $\mathbb{R}^3 = V$, $\mathcal{L}(\mathbb{R}^3, \mathbb{R})$ è nient'altro che lo spazio duale V^* .

Indicate con x, y, z le coordinate di X , possiamo considerare le tre funzioni coordinate f_1, f_2, f_3 definite da

$$f_1(X) = x, \quad f_2(X) = y, \quad f_3(X) = z.$$

Allora la notazione usuale è la seguente:

$$df_1(X) = dx, \quad df_2(X) = dy, \quad df_3(X) = dz.$$

Dalle definizioni si vede immediatamente che dx, dy, dz , costituiscono una base di V^* e, più precisamente, costituiscono la base duale della base usuale costituita dai vettori unità $\{e^1, e^2, e^3\}$.

Diremo *forma differenziale su \mathbb{R}^3 di grado 2* ogni applicazione

(non necessariamente lineare)

$$\omega: \mathbb{R}^3 \rightarrow V^* \wedge V^*$$

definita in \mathbb{R}^3 e con valori nel prodotto alternante nello spazio duale con sé stesso. Poiché per ogni X in \mathbb{R}^3 dx, dy, dz , costituiscono una base di V^* , ne segue che $dx \wedge dy, dx \wedge dz, dy \wedge dz$ costituiscono una base di $V^* \wedge V^*$. Conseguentemente, esistono le funzioni

$$\omega_{ij}: \mathbb{R}^3 \rightarrow \mathbb{R} \quad (1 < i < j < 3)$$

in modo che $\omega(X)$ si può esprimere come combinazione lineare

$$\omega(X) = \omega_{12}(X)dx \wedge dy + \omega_{13}(X)dx \wedge dz + \omega_{23}(X)dy \wedge dz.$$

Queste funzioni ω_{ij} non sono altro che le funzioni coordinate di ω rispetto alla summenzionata base di $V^* \wedge V^*$.

Esempio 2. Sia $V = \mathbb{R}^3$. Se $X = (x_1, x_2, x_3)$ e $Y = (y_1, y_2, y_3)$ sono elementi di \mathbb{R}^3 , si definisce loro *prodotto vettoriale* l'espressione

$$X \times Y = (x_1 y_2 - x_2 y_1, x_1 y_3 - x_3 y_1, x_2 y_3 - x_3 y_2).$$

Il lettore non avrà difficoltà a riconoscere che si tratta essenzialmente del prodotto alternante $X \wedge Y$ espresso mediante le sue coordinate. Nel paragrafo successivo vedremo come questo può essere generalizzato al caso di dimensioni superiori, dando le coordinate del prodotto alternante mediante determinanti di ordine più elevato.

Esercizi

1. Sia V uno spazio vettoriale tridimensionale sul corpo K . Si definisca $V \times V \times V$ come l'insieme di tutte le terne ordinate (u, v, w) di elementi di V . Un'applicazione trilineare

$$f: V \times V \times V \rightarrow U$$

in uno spazio vettoriale U sullo stesso corpo K è un'applicazione che è lineare in ogni componente. Diciamo poi che un'applicazione trilineare f è *alternante* quando

$$f(u, v, w) = O$$

tutte le volte che $u = v$ oppure $v = w$. Dimostrare che, se f è alternante, allora $f(u, v, w) = O$ se $u = w$.

2. Sia $\{v_1, v_2, v_3\}$ una base di V . Sia $V \wedge V \wedge V$ lo spazio vettoriale di dimensione uno su K generato da una singola lettera t_{123} . Se X, Y, Z sono i vettori delle coordinate degli elementi u, v, w di V rispetto alla base fissata, si definisca il prodotto

$$u \wedge v \wedge w = \text{Det}(X, Y, Z) t_{123}.$$

Dimostrare che questo prodotto è trilineare e alternante, ricorrendo alla definizione di determinante. Si noti anche che $t_{123} = v_1 \wedge v_2 \wedge v_3$.

3. Sia $f: V \times V \times V \rightarrow U$ un'applicazione trilineare alternante in uno spazio vettoriale U sullo stesso corpo K . Dimostrare che esiste un'unica applicazione lineare $f_*: V \wedge V \wedge V \rightarrow U$ tale che

$$f(u, v, w) = f_*(u \wedge v \wedge w)$$

per ogni scelta di u, v, w in V .

4. Sia V uno spazio vettoriale di dimensione n sul corpo K . Si definisca $V \wedge V$ in modo analogo a quello seguito nel caso tridimensionale. Dimostrare le proprietà analoghe a quelle trovate in questo caso speciale. [Suggerimento: usare tutti i determinanti $2 \times 2 x_i y_j - x_j y_i$, con $1 < i < j \leq n$.]

65. PRODOTTI ALTERNANTI: CASO GENERALE

Questo paragrafo è di carattere piuttosto astratto: per alcuni lettori può essere preferibile ometterlo.

Per definire il prodotto alternante nel caso di un arbitrario numero di fattori prenderemo spunto dal prodotto alternante considerato nel precedente paragrafo. Osserviamo che ogni espressione

$$x_i y_j - x_j y_i$$

con $i < j$ è un determinante e che tutti questi determinanti compaiono come coefficienti di $v_i \wedge v_j$. Ora generalizzeremo questa situazione in modo naturale.

Sia V uno spazio vettoriale sul corpo K . Sia r un intero maggiore di zero. Indicheremo abbreviatamente con $V^{(r)}$ l'insieme di tutte le r -uple di elementi di V , cioè porremo

$$V^{(r)} = \overbrace{V \times \dots \times V}^r.$$

Un elemento di $V^{(r)}$ è quindi una r -upla (w_1, \dots, w_r) dove $w_i \in V$. Ogni componente dell' r -upla è quindi un elemento di V .

Sia U uno spazio vettoriale sul corpo K . Dicendo *applicazione*

r-multilineare di V in U intenderemo un'applicazione

$$f: V \times \dots \times V \rightarrow U$$

di $V^{(r)}$ in U che sia lineare in ogni componente. In altre parole, per ogni $i = 1, \dots, r$, si ha

$$f(w_1, \dots, w_i + w'_i, \dots, w_r) = f(w_1, \dots, w_r) + f(w_1, \dots, w'_i, \dots, w_r),$$

$$f(w_1, \dots, cw_i, \dots, w_r) = cf(w_1, \dots, w_r)$$

per ogni scelta di w_i e w'_i in V e di c in K . Diremo che un'applicazione multilineare f come definita sopra è *alternante* se, inoltre, per essa abbiamo

$$f(w_1, \dots, w_r) = O$$

tutte le volte che due componenti contigue sono uguali, cioè tutte le volte che esiste un indice j minore di r tale che w_j coincide con w_{j+1} .

Osserviamo che un'applicazione multilineare alternante soddisfa delle condizioni del tutto simili alle prime due proprietà soddisfatte per i determinanti. Quindi le applicazioni multilineari alternanti possono essere considerate come una generalizzazione dei determinanti. In effetti, possiamo dire ora che un determinante è un'applicazione multilineare alternante definita su K^n avente l'ulteriore proprietà

$$\text{Det}(e^1, \dots, e^n) = 1,$$

se $\{e^1, \dots, e^n\}$ è la base naturale di K^n .

Se ricordiamo le proprietà 4, 5, 6 dei determinanti, vediamo che le loro dimostrazioni dipendono soltanto dalle proprietà 1, 2. Ne segue che queste dimostrazioni continuano a valere anche nella presente situazione più generale. Quindi, per esempio, se noi scambiamo due componenti contigue, w_j e w_{j+1} , il valore dell'applicazione multilineare cambia il segno. Se due componenti distinte qualsiasi w_i e w_j coincidono (con $i \neq j$), allora

$$f(w_1, \dots, w_r) = O.$$

Queste proprietà sono costantemente usate nel calcolare i valori delle applicazioni alternanti.

Il nostro problema è definire un prodotto tra r elementi di V in modo che siano soddisfatte soltanto le relazioni di multilinearità e di alternanza. La risoluzione di questo problema costituirà il teorema 6. Prima di far questo, tuttavia, vogliamo dedurre alcune conseguenze da queste relazioni.

TEOREMA 4 *Siano V , U spazi vettoriali sul corpo K e sia*

$$f: V^{(r)} \rightarrow U$$

un'applicazione r -multilineare alternante. Siano w_1, \dots, w_r elementi di V , sia $A = (a_{ij})$ una matrice $r \times r$ in K . Sia

$$\begin{aligned} u_1 &= a_{11}w_1 + \dots + a_{1r}w_r, \\ &\dots \\ u_r &= a_{r1}w_1 + \dots + a_{rr}w_r. \end{aligned}$$

Allora

$$f(u_1, \dots, u_r) = \text{Det}(A)f(w_1, \dots, w_r).$$

Dimostrazione. Intanto abbiamo

$$f(u_1, \dots, u_r) = f(a_{11}w_1 + \dots + a_{1r}w_r, \dots, a_{r1}w_1 + \dots + a_{rr}w_r).$$

Sviluppando per multilinearità si ottiene una somma di termini

$$\sum_{\sigma} f(a_{1,\sigma(1)}w_{\sigma(1)}, \dots, a_{r,\sigma(r)}w_{\sigma(r)})$$

estesa a tutte le possibili scelte $\sigma(1), \dots, \sigma(r)$, cioè estesa a tutte le possibili permutazioni $\sigma: \{1, \dots, r\} \rightarrow \{1, \dots, r\}$. Questa somma risulta uguale a

$$\sum_{\sigma} a_{1,\sigma(1)} \dots a_{r,\sigma(r)} f(w_{\sigma(1)}, \dots, w_{\sigma(r)})$$

estraendo tutti i fattori scalari dagli argomenti dell'applicazione f . Se σ non è una permutazione di $\{1, \dots, r\}$, allora due componenti distinte nella r -upla

$$(w_{\sigma(1)}, \dots, w_{\sigma(r)})$$

risultano uguali e il termine corrispondente nella somma è quindi nullo. Perciò possiamo estendere la somma soltanto alle permutazioni σ di $\{1, \dots, r\}$.

Se facciamo questo e riportiamo i termini della r -upla

$(w_{\sigma(1)}, \dots, w_{\sigma(r)})$ nell'ordinamento naturale (w_1, \dots, w_r) , allora il segno di ogni addendo viene mutato secondo il segno della permutazione σ . Perciò, infine, si può scrivere

$$f(u_1, \dots, u_r) = \sum_{\sigma} \varepsilon(\sigma) a_{1,\sigma(1)} \dots a_{r,\sigma(r)} f(w_1, \dots, w_r),$$

che è uguale a

$$\text{Det}(A)f(w_1, \dots, w_r)$$

ricordando una delle espressioni che abbiamo ottenute per i determinanti. Questo dimostra il nostro teorema.

Per prodotti alternanti di più fattori, abbiamo ancora bisogno di una formula più generale per sviluppare un'applicazione alternante, cioè ne abbiamo bisogno nei casi in cui $A = (a_{ij})$ è una matrice $r \times n$ con r diverso da n . Per esempio, nel precedente paragrafo abbiamo considerato il caso $r = 2$ e $n = 3$. Dobbiamo quindi introdurre ancora qualche notazione. Sia r tale che $1 < r < n$.

Sia S un sottoinsieme dell'insieme di interi $\{1, \dots, n\}$ costituito da r elementi. Il numero dei possibili sottoinsiemi siffatti è uguale al coefficiente binomiale

$$\binom{n}{r}.$$

Gli elementi di un tale insieme S possono inoltre essere ordinati in modo che, se i_1, \dots, i_r sono questi elementi, allora $i_1 < \dots < i_r$. Sia

$$\sigma: \{1, \dots, r\} \rightarrow S$$

un'applicazione, cioè un modo di associare ad ogni intero tra 1 e r un elemento di S . Si assuma inoltre che, se i non coincide con j , allora neppure $\sigma(i)$ coincide con $\sigma(j)$. Possiamo allora considerare σ come una permutazione di S . Infatti, se i_1, \dots, i_r sono gli elementi di S e sono ordinati in modo che

$$i_1 < \dots < i_r,$$

allora σ definisce una permutazione denotata simbolicamente con

$$\begin{bmatrix} i_1 & \dots & i_r \\ \sigma(1) & \dots & \sigma(r) \end{bmatrix}.$$

Quindi questa permutazione è l'associazione

$$i_1 \mapsto \sigma(1), \quad i_2 \mapsto \sigma(2), \quad \dots, \quad i_r \mapsto \sigma(r).$$

Il segno di questa permutazione verrà denotato con $\varepsilon_S(\sigma)$.

Esempio. Siano $n = 4$ e $r = 3$. Sia $S = \{1, 3, 4\}$. Sia σ definita da

$$\sigma(1) = 4, \quad \sigma(3) = 1, \quad \sigma(4) = 3.$$

Allora σ definisce la permutazione

$$\begin{bmatrix} 1 & 3 & 4 \\ 4 & 1 & 3 \end{bmatrix}$$

dell'insieme $\{1, 3, 4\}$. Il suo segno è

$$\varepsilon_S(\sigma) = +1.$$

Per stabilire una notazione, denotiamo con $P(S)$ l'insieme delle applicazioni

$$\sigma: \{1, \dots, r\} \rightarrow S$$

tali che $\sigma(i) \neq \sigma(j)$ se $i \neq j$. Quindi $P(S)$ è essenzialmente l'insieme delle permutazioni di S .

Sia $A = (a_{ij})$ una matrice $r \times n$ in K . Per ogni sottoinsieme S di $\{1, \dots, n\}$ costituito esattamente da r elementi, possiamo considerare il minore $r \times r$ di A costituito da quegli elementi a_{ij} tali che $j \in S$. Denotiamo con

$$\text{Det}_S(A)$$

il determinante di questo minore. Questo stesso determinante sarà anche chiamato il sottodeterminante di A relativo all'insieme S . Allora possiamo scrivere

$$\text{Det}_S(A) = \sum_{\sigma \in P(S)} \varepsilon_S(\sigma) a_{1,\sigma(1)} \dots a_{r,\sigma(r)},$$

dove la somma è estesa a tutte le applicazioni σ dell'insieme $P(S)$. Si tratta di una semplice riformulazione dell'espressione di un determinante, tenuto conto delle notazioni ora introdotte.

Siano v_1, \dots, v_n elementi di V . Per ognuno dei sottoinsiemi S prima considerati, denotiamo con v_S la r -upla

$$v_S = (v_{i_1}, \dots, v_{i_r}),$$

dove i_1, \dots, i_r sono gli elementi di S ordinati in modo che

$$i_1 < \dots < i_r.$$

Abbiamo ora tutte le notazioni necessarie per enunciare la generalizzazione del teorema 4 che cercavamo.

TEOREMA 5 *Siano V , U spazi vettoriali sul corpo K . Sia*

$$f: V^{(r)} \rightarrow U$$

un'applicazione r -multilineare alternante. Siano v_1, \dots, v_n elementi di V e $A = (a_{ij})$ sia una matrice $r \times n$ in K . Sia infine

$$\begin{aligned} u_1 &= a_{11}v_1 + \dots + a_{1n}v_n, \\ &\quad \cdots \\ u_r &= a_{r1}v_1 + \dots + a_{rn}v_n. \end{aligned}$$

Allora

$$f(u_1, \dots, u_r) = \sum_S \text{Dets}_S(A) f(v_S),$$

dove la somma è estesa a tutti i sottoinsiemi S di $\{1, \dots, n\}$ costituiti esattamente da r elementi.

Dimostrazione. Abbiamo intanto

$$f(u_1, \dots, u_r) = f(a_{11}v_1 + \dots + a_{1n}v_n, \dots, a_{r1}v_1 + \dots + a_{rn}v_n).$$

Sviluppando per multilinearità, otteniamo una somma

$$\sum_{\sigma} a_{1,\sigma(1)} \dots a_{r,\sigma(r)} f(v_{\sigma(1)}, \dots, v_{\sigma(r)})$$

estesa a tutte le possibili scelte σ che assegnano ad ogni intero tra 1 e r un intero tra 1 e n . Quindi la somma è estesa a tutte le applicazioni

$$\sigma: \{1, \dots, r\} \rightarrow \{1, \dots, n\}.$$

Come abbiamo già fatto in precedenza, anche ora osserviamo che se per qualche coppia di indici diversi i, j , $\sigma(i)$ coincide con $\sigma(j)$, allora il corrispondente termine nella somma è nullo perché f è alternante. Quindi, nella nostra somma, possiamo considerare soltanto quelle applicazioni σ tali che $\sigma(i) \neq \sigma(j)$ se $i \neq j$.

Possiamo ora decomporre la nostra somma in una doppia

somma raggruppando insieme tutte le applicazioni che mandano $\{1, \dots, r\}$ in uno stesso insieme S e quindi considerando la somma estesa a tutti questi possibili insiemi S . Quindi, simbolicamente, possiamo scrivere

$$\sum_{\sigma} = \sum_S \sum_{\sigma \in P(S)} .$$

In ognuna delle somme interne

$$\sum_{\sigma \in P(S)} a_{1,\sigma(1)} \dots a_{r,\sigma(r)} f(v_{\sigma(1)}, \dots, v_{\sigma(r)})$$

riordiniamo la r -upla $(v_{\sigma(1)}, \dots, v_{\sigma(r)})$ in modo da riportarne gli elementi nella posizione usuale $(v_{i_1}, \dots, v_{i_r})$, dove i_1, \dots, i_r sono gli elementi di S ordinati in modo che $i_1 < \dots < i_r$. Allora f muta secondo il segno $\varepsilon_S(\sigma)$ e conseguentemente ogni somma interna viene scritta come segue

$$\sum_{\sigma \in P(S)} \varepsilon_S(\sigma) a_{1,\sigma(1)} \dots a_{r,\sigma(r)} f(v_S) .$$

Considerando ora la somma di tutti questi termini estesa a tutti i possibili insiemi S , otteniamo esattamente la formula enunciata nel teorema.

Il teorema che segue tratta dei prodotti alternanti nel caso generale:

TEOREMA 6 *Sia V uno spazio vettoriale di dimensione finita sul corpo K ; sia n questa dimensione. Sia r un intero tale che $1 \leq r \leq n$. Esiste allora uno spazio vettoriale di dimensione finita su K , denotato con $\wedge^r V$ e un'applicazione r -multilineare alternante $V^{(r)} \rightarrow \wedge^r V$, denominata con*

$$(u_1, \dots, u_r) \mapsto u_1 \wedge \dots \wedge u_r ,$$

soddisfacente le seguenti proprietà.

PA 1. *Se U è uno spazio vettoriale sul corpo K e se $g: V^{(r)} \rightarrow U$ è un'applicazione r -multilineare alternante, allora esiste un'unica applicazione lineare*

$$g_*: \wedge^r V \rightarrow U$$

tale che, per ogni scelta di u_1, \dots, u_r in V , si abbia

$$g(u_1, \dots, u_r) = g_*(u_1 \wedge \dots \wedge u_r) .$$

PA 2. Se $\{v_1, \dots, v_n\}$ è una base di V , allora l'insieme degli elementi

$$\{v_{i_1} \wedge \dots \wedge v_{i_r}\} \quad (1 < i_1 < \dots < i_r < n)$$

è una base di $\wedge^r V$.

Dimostrazione. Per ogni sottoinsieme S di $\{1, \dots, n\}$ costituito da r elementi, fissiamo una lettera t_S . Queste lettere t_S costituiscono una base di uno spazio vettoriale sul corpo K la cui dimensione è uguale al coefficiente binomiale $\binom{n}{r}$. Denotiamo questo spazio con $\wedge^r V$. Sia $\{v_1, \dots, v_n\}$ una base di V . Siano u_1, \dots, u_r elementi di V . Sia $A = (a_{ij})$ la matrice con elementi di K tale che

$$\begin{aligned} u_1 &= a_{11}v_1 + \dots + a_{1n}v_n, \\ &\dots \\ u_r &= a_{r1}v_1 + \dots + a_{rn}v_n. \end{aligned}$$

Si definisca infine

$$u_1 \wedge \dots \wedge u_r = \sum_S \text{Det}_S(A) t_S.$$

Vogliamo far vedere che questo prodotto ha tutti i requisiti richiesti.

Proviamo dapprima la multilinearità. Si tratta essenzialmente di un'osservazione tecnica, cioè si tratta di far vedere che $\text{Det}_S(A)$ è multilineare se considerato come funzione delle righe di A . Supponiamo che

$$u'_i = a'_{i1}v_1 + \dots + a'_{in}v_n.$$

Siano A_1, \dots, A_n le righe di A e sia $A'_i = (a'_{i1}, \dots, a'_{in})$. Scrivendo il determinante in funzione delle righe, abbiamo

$$\begin{aligned} \text{Det}_S(A_1, \dots, A_i + A'_i, \dots, A_n) &= \\ &= \text{Det}_S(A_1, \dots, A_n) + \text{Det}_S(A_1, \dots, A'_i, \dots, A_n) \end{aligned}$$

e, se c è in K ,

$$\text{Det}_S(A_1, \dots, cA_i, \dots, A_n) = c \text{Det}_S(A_1, \dots, A_n).$$

Queste uguaglianze seguono direttamente dalla definizione di determinante. Conseguentemente

$$u_1 \wedge \dots \wedge (u_i + u'_i) \wedge \dots \wedge u_r = (u_1 \wedge \dots \wedge u_r) + (u_1 \wedge \dots \wedge u'_i \wedge \dots \wedge u_r),$$

e anche

$$u_1 \wedge \dots \wedge (cu_i) \wedge \dots \wedge u_r = cu_1 \wedge \dots \wedge u_r.$$

Il prodotto è alternante perché, se esiste un indice i per cui $u_i = u_{i+1}$, allora due righe contigue della matrice A risultano uguali. Perciò per ogni S sono uguali due righe contigue nel minore di A corrispondente a S e quindi $\text{Det}_S(A) = 0$.

Osserviamo poi che

$$t_S = v_{i_1} \wedge \dots \wedge v_{i_r}$$

se i_1, \dots, i_r sono gli elementi di S ordinati in modo che

$$i_1 < \dots < i_r.$$

Dal teorema relativo all'esistenza e unicità di applicazioni lineari aventi prefissati valori sugli elementi di una base, traiamo la conclusione che, se $g: V^{(r)} \rightarrow U$ è un'applicazione multilineare alternante, allora esiste un'unica applicazione lineare

$$g_*: \bigwedge^r V \rightarrow U$$

tale che, per ogni insieme S , si abbia

$$g_*(t_S) = g(v_S) = g(v_{i_1}, \dots, v_{i_r}),$$

se i_1, \dots, i_r sono scelti nel modo già detto. Dal teorema 5 segue allora che

$$g(u_1, \dots, u_r) = g_*(u_1 \wedge \dots \wedge u_r)$$

per tutti gli elementi u_1, \dots, u_r di V . Così si prova PA 1.

Per provare PA 2, sia $\{w_1, \dots, w_n\}$ una base di V . Per lo sviluppo dato nel teorema 5, segue che gli elementi

$$\{f(w_S)\},$$

cioè gli elementi

$$\{w_{i_1} \wedge \dots \wedge w_{i_r}\},$$

per tutte le possibili scelte di r -uple (i_1, \dots, i_r) tali che

$$i_1 < \dots < i_r,$$

sono generatori di $\bigwedge^r V$. Poiché il numero di questi elementi è

esattamente $\binom{n}{r}$, essi risultano realmente indipendenti e costituiscono quindi una base di $\wedge^r V$, come si doveva dimostrare.

La notazione v_s si è mostrata utile per abbreviare le espressioni che abbiamo incontrato finora. Tuttavia, la somma viene anche scritta frequentemente conservando le indicazioni degli indici i_1, \dots, i_r . Perciò, se $\{v_1, \dots, v_n\}$ è una base di V , allora ogni elemento di $\wedge^r V$ è univocamente esprimibile come combinazione lineare

$$\sum_{i_1 < \dots < i_r} c_{i_1 \dots i_r} v_{i_1} \wedge \dots \wedge v_{i_r},$$

la somma essendo estesa a tutte le r -uple (i_1, \dots, i_r) di interi compresi tra 1 e n e tali che

$$i_1 < \dots < i_r.$$

Questa notazione può essere abbreviata ponendo $(i) = (i_1, \dots, i_r)$ e allora la somma scritta sopra diviene

$$\sum_{(i)} c_{(i)} v_{i_1} \wedge \dots \wedge v_{i_r}.$$

Esercizi

1. Sia V uno spazio vettoriale di dimensione n sul corpo K . Dimostrare che $\wedge^n V$ ha dimensione 1 sul corpo K .

2. Sia V come nell'esercizio 1. Sia $\{v_1, \dots, v_n\}$ una base di V e sia $A = (a_{ij})$ una matrice $n \times n$ in K . Si ponga

$$\begin{aligned} u_1 &= a_{11}v_1 + \dots + a_{1n}v_n, \\ &\dots \\ u_n &= a_{n1}v_1 + \dots + a_{nn}v_n. \end{aligned}$$

Esprimere allora $u_1 \wedge \dots \wedge u_n$ mediante $v_1 \wedge \dots \wedge v_n$.

3. Sia V uno spazio vettoriale di dimensione n sul corpo K e sia r un intero maggiore di n . Dimostrare che ogni applicazione r -multilineare alternante

$$f: V^{(r)} \rightarrow U$$

in uno spazio vettoriale U è l'applicazione nulla.

4. Sia A una matrice semisimmetrica $n \times n$ in K . Sia cioè ${}^t A = -A$. Dimostrare che l'associazione $(X, Y) \mapsto {}^t XAY$ definisce una forma alternante su K^n .

5. Sia V uno spazio vettoriale di dimensione n sul corpo K . Sia $\{v_1, \dots, v_n\}$ una base di V . Sia $c \in K$.

a) Dimostrare che esiste un'unica forma n -multilineare alternante $f_c: V^{(n)} \rightarrow K$ tale che

$$f_c(v_1, \dots, v_n) = c.$$

b) Sia f_1 l'unica forma n -multilineare alternante di V in K tale che

$$f_1(v_1, \dots, v_n) = 1.$$

Se g è una forma n -multilineare alternante in V e tale che $g(v_1, \dots, v_n) = c \in K$, dimostrare che g coincide con cf_1 .

6. Sia V uno spazio vettoriale di dimensione n sul corpo K . Sia $A: V \rightarrow V$ un'applicazione lineare. Sia $f: V^{(n)} \rightarrow K$ una forma n -multilineare alternante. Sia $g: V^{(n)} \rightarrow K$ definita da

$$g(w_1, \dots, w_n) = f(Aw_1, \dots, Aw_n).$$

Dimostrare che g è una forma n -multilineare alternante.

7. Più in generale, siano V, W spazi vettoriali sul corpo K . Sia $A: V \rightarrow W$ un'applicazione lineare. Sia $f: W^{(n)} \rightarrow U$ un'applicazione n -multilineare alternante di W in uno spazio vettoriale U . Sia $g: V^{(n)} \rightarrow U$ definita da

$$g(v_1, \dots, v_n) = f(Av_1, \dots, Av_n).$$

Dimostrare allora che g è un'applicazione n -multilineare alternante.

8. Sia V uno spazio vettoriale di dimensione n sul corpo K . Sia $A: V \rightarrow V$ un'applicazione lineare. Siano w_1, \dots, w_n elementi di V e sia $f: V^{(n)} \rightarrow U$ un'applicazione n -multilineare alternante. Dimostrare allora che

$$f(Aw_1, \dots, Aw_n) = \text{Det}(A)f(w_1, \dots, w_n).$$

66. APPENDICE: LO SPAZIO VETTORIALE GENERATO DA UN INSIEME

Sia K un corpo e sia S un insieme finito di oggetti. Per semplicità numeriamo gli elementi di S ; siano quindi

$$s_1, \dots, s_n$$

questi elementi. Vogliamo ora definire che cosa intendiamo dicendo spazio vettoriale T delle combinazioni lineari "formali"

$$c_1s_1 + \dots + c_ns_n$$

di elementi di S a coefficienti c_i in K . Se vogliamo essere del tutto precisi, dobbiamo prima descrivere gli elementi dello spazio

vettoriale T e quindi definire l'addizione tra essi. Altrimenti il segno “+” non ha significato. Questo non significa che non si possa semplicemente ignorare il problema e procedere come se tutto fosse chiaro. Molti, infatti, preferiscono ogni volta regalarsi in questo modo, e non sembrano preoccuparsene.

Comunque, il proponimento di questa appendice è precisamente quello di mostrare come si fa a essere precisi in questa circostanza. In effetti è una cosa molto semplice. Che cosa vogliamo sia una “somma” come la seguente

$$c_1 s_1 + \dots + c_n s_n ?$$

Vogliamo che essa sia completamente determinata dai “coefficienti” c_1, \dots, c_n e che ogni “coefficiente” c_i sia associato all’elemento s_i dell’insieme S . Ma un’associazione non è altro che una funzione, e questo ci suggerisce come definire gli elementi del nostro spazio T .

Per ogni elemento s_i di S e ogni elemento c di K definiamo il simbolo

$$cs_i$$

come la funzione che associa c a s_i e 0 a s_j , se $j \neq i$. Se $a \in K$, si ha evidentemente

$$a(cs_i) = (ac)s_i \quad \text{e} \quad (c + c')s_i = cs_i + c's_i .$$

Definiamo T come l’insieme di tutte le funzioni di S in K che possono essere scritte nella forma

$$c_1 s_1 + \dots + c_n s_n$$

scegliendo c_i in K . Facendo uso delle precedenti immediate proprietà, si vede subito che T è uno spazio vettoriale sul corpo K . (Osserviamo che non c’è nessun problema nel prendere le somme, giacché sappiamo che cosa significa addizionare funzioni di S in K .)

Vogliamo ora far vedere che le funzioni

$$1s_1, \dots, 1s_n$$

sono linearmente indipendenti e che perciò costituiscono una base di T su K .

Per dimostrare ciò, supponiamo che in K esistano elementi c_1, \dots, c_n tali che

$$c_1 s_1 + \dots + c_n s_n = 0 \quad (\text{cioè la funzione zero}) .$$

Allora, per definizione, il primo membro assume il valore c_i in s_i e perciò c_i è nullo. Questo prova, come volevamo, l'indipendenza lineare.

In pratica è conveniente abbreviare la notazione e scrivere soltanto s_i invece di $1s_i$. Gli elementi di T sono allora chiamati combinazioni lineari formali degli elementi di S . Abbiamo così completamente giustificato questa terminologia.

Capitolo 14

Gruppi

67. GRUPPI ED ESEMPI

Facciamo ora un piccolo passo verso una maggiore astrazione e definiamo una nozione della quale molti esempi precedenti appaiono casi particolari.

Un *gruppo* G è un insieme in cui è assegnata una regola (chiamata legge di composizione) che permette di associare ad ogni coppia di elementi x, y di G un elemento di G stesso, denotato con xy e avente le seguenti proprietà.

GR 1. *Comunque si prendano x, y, z in G vale la proprietà associativa, cioè*

$$(xy)z = x(yz).$$

GR 2. *In G esiste un elemento e tale che, per ogni x in G , $ex = xe = x$.*

GR 3. *Per ogni elemento x di G esiste un elemento y di G tale che $xy = yx = e$.*

Più precisamente chiamiamo G un gruppo *moltiplicativo*. Se l'elemento di G associato alla coppia (x, y) viene denotato con $x + y$, allora GR 1 assume la forma

$$(x + y) + z = x + (y + z),$$

GR 2 viene espressa dicendo che esiste un elemento 0 tale che

$$0 + x = x + 0 = x$$

per ogni x in G e infine GR 3 viene espressa dicendo che per ogni $x \in G$ esiste un elemento y di G tale che

$$x + y = y + x = 0.$$

Con queste notazioni, G viene chiamato gruppo *additivo*. Noi adopereremo la notazione “+” soltanto quando nel gruppo è soddisfatta l'ulteriore proprietà

$$x + y = y + x$$

comunque si scelgano gli elementi x, y in G . Nella notazione moltiplicativa questa proprietà si scrive $xy = yx$ per ogni scelta di x, y in G . Quando il gruppo G ha questa proprietà, G viene detto *commutativo* oppure *abeliano*.

Esempio 1. I numeri razionali costituiscono un gruppo rispetto all'addizione. Lo stesso accade per i numeri reali e anche per i numeri complessi. In effetti, in ogni corpo K gli elementi di K costituiscono un gruppo rispetto all'addizione.

Esempio 2. I numeri razionali non nulli costituiscono un gruppo rispetto alla moltiplicazione. Lo stesso accade per i numeri reali non nulli, per i numeri complessi non nulli e per gli elementi non nulli di un qualsiasi corpo K .

Esempio 3. I numeri complessi aventi valore assoluto uguale a 1 costituiscono un gruppo rispetto alla moltiplicazione.

Esempio 4. Le permutazioni di $\{1, \dots, n\}$ costituiscono un gruppo rispetto alla moltiplicazione (cioè alla composizione delle applicazioni), detto *gruppo simmetrico* S_n su n elementi.

Esempio 5. Gli elementi di uno spazio vettoriale costituiscono un gruppo rispetto all'addizione.

Esempio 6. Le matrici $m \times n$ su un corpo K costituiscono un gruppo rispetto all'addizione.

Gruppi moltiplicativi di matrici

Esempio 7. a) Le matrici invertibili $n \times n$ su un corpo K costituiscono un gruppo rispetto alla moltiplicazione, denotato con $GL_n(K)$ oppure $GL(n, K)$, e chiamato *gruppo lineare generale*.

b) Sia V uno spazio vettoriale sul corpo K . Le applicazioni lineari invertibili di V in sé stesso costituiscono un gruppo rispetto alla moltiplicazione (composizione di applicazioni), denotato con $GL(V)$.

Esempio 8. Le matrici reali unitarie $n \times n$ costituiscono un gruppo rispetto alla moltiplicazione, denotato con $U_n(\mathbb{R})$ oppure $U(n, \mathbb{R})$.

Esempio 9. Le matrici unitarie complesse $n \times n$ costituiscono un gruppo rispetto alla moltiplicazione, denotato con $U_n(\mathbb{C})$ oppure con $U(n, \mathbb{C})$.

Esempio 10. Le matrici invertibili superiormente triangolari $n \times n$ sul corpo K costituiscono un gruppo rispetto alla moltiplicazione.

Esempio 11. Le matrici invertibili $n \times n$ su un corpo K aventi determinante uguale a 1 costituiscono un gruppo rispetto alla moltiplicazione. Questo gruppo è chiamato il *gruppo lineare speciale*.

Terminologia. I gruppi additivi di matrici o di applicazioni lineari non presentano interesse. Perciò in tutta la matematica, a meno che non sia detto il contrario, un gruppo di matrici è sempre inteso come un gruppo *moltiplicativo* di matrici. Lo stesso deve intendersi per i gruppi di applicazioni lineari, la legge di composizione essendo la composizione di applicazioni lineari, frequentemente denotata come una moltiplicazione.

Si osservi che i gruppi moltiplicativi di matrici considerati nei vari esempi precedenti non sono commutativi quando n supera 1. In ogni caso la verifica che l'insieme considerato è un gruppo si riduce a una ripetizione sommaria di note proprietà degli oggetti in considerazione.

Un gruppo costituito da un solo elemento viene detto *banale*. In generale un gruppo può avere infiniti elementi o solamente un numero finito. Se G ha soltanto un numero finito di elementi, G stesso è allora chiamato *gruppo finito* e il numero degli elementi di G viene detto il suo *ordine*. Il gruppo di permutazioni dell'esempio 4 è un gruppo finito. La determinazione del suo ordine costituirà un esercizio. Il gruppo i cui elementi sono 1 e -1 (e in cui la legge di composizione è la moltiplicazione) ha ordine 2.

Esempio 12. Prodotto diretto. Siano G, G' due gruppi. Sia $G \times G'$ l'insieme costituito da tutte le coppie ordinate (x, x') con $x \in G$ e $x' \in G'$. Se (x, x') e (y, y') sono due di tali coppie, definiamo loro prodotto la coppia $(xy, x'y')$. Allora $G \times G'$ è un gruppo. Per esercizio si verifichi in tutti i dettagli che tutte le condizioni sono soddisfatte. Chiamiamo $G \times G'$ il *prodotto diretto* di G e G' .

Esercizi

1. Dimostrare che l'ordine del gruppo simmetrico S_2 è 2. Dimostrare che l'ordine del gruppo simmetrico S_3 è 6. Dimostrare per induzione che, in generale, l'ordine del gruppo simmetrico S_n è $n!$.
2. Per ognuno degli esempi considerati nel testo, dire esplicitamente qual è l'unità del gruppo. (L'elemento unità è quello la cui esistenza è assurta nella proposizione GR 2. Nel paragrafo successivo dimostreremo che questo elemento è univocamente determinato.)
3. Dimostrare che l'insieme dei numeri complessi che sono radici del polinomio $f(t) = t^n - 1$ è un gruppo rispetto alla moltiplicazione. Qual è l'ordine di questo gruppo?
4. Sia \mathcal{S} un insieme contenente almeno un elemento. Sia G l'insieme di tutte le applicazioni $f: \mathcal{S} \rightarrow \mathcal{S}$ che sono iniettive e surgettive. Dimostrare che G è un gruppo, la legge di composizione essendo la composizione delle applicazioni. (La condizione GR 1 è già nota come la legge associativa delle applicazioni.) Questo gruppo G è chiamato il gruppo di tutte le *applicazioni invertibili* di \mathcal{S} in sé stesso. Esso costituisce una generalizzazione della nozione di gruppo delle permutazioni su n elementi.
5. Sia V uno spazio vettoriale su un corpo K e sia \langle , \rangle un prodotto scalare su V , cioè una forma bilineare simmetrica. Per *automorfismo* della forma intenderemo un'applicazione lineare invertibile $A: V \rightarrow V$ tale che $\langle Av, Aw \rangle = \langle v, w \rangle$ per ogni scelta di v e w in V . Dimostrare che l'insieme degli automorfismi della forma è un gruppo.
6. Sia G un gruppo e siano a, b, c elementi di G . Dimostrare che, se $ab = ac$, allora $b = c$.
7. Siano G, G' gruppi finiti di rispettivi ordini m, n . Qual è l'ordine del prodotto diretto $G \times G'$?
8. Sia G un gruppo abeliano finito di ordine n e siano a_1, \dots, a_n i suoi elementi. Dimostrare che il prodotto $a_1 \dots a_n$ è un elemento il cui quadrato è l'elemento unità.

68. SEMPLICI PROPRIETÀ DEI GRUPPI

Vogliamo ora dimostrare varie semplici proposizioni che sono valide per tutti i gruppi. Sia G un gruppo. L'elemento e di G ,

la cui esistenza è asserita nella proposizione GR 2, è univocamente determinato, giacché, se e ed e' soddisfano entrambi questa condizione, allora

$$e' = ee' = e.$$

Questo elemento viene chiamato elemento *unità* di G . Nel caso additivo esso viene invece chiamato l'elemento *zero*.

Sia x un elemento di G . L'elemento y tale che $yx = xy = e$ è univocamente determinato perché se z è un elemento per cui $zx = xz = e$, allora

$$z = ez = (yx)z = y(xz) = ye = y.$$

Chiamiamo y l'*inverso* di x e lo denotiamo con x^{-1} . Nella notazione additiva scriviamo invece $y = -x$.

Siano G un gruppo e H un sottoinsieme di G . Diciamo che H è un *sottogruppo* se ad esso appartiene l'elemento unità e se, comunque si prendano x e y in H , anche gli elementi xy ed x^{-1} appartengono a H . (Nella notazione additiva scriviamo $x + y \in H$ e $-x \in H$.) Allora anche H è un gruppo, la legge di composizione in H essendo uguale a quella in G . L'elemento unità di G costituisce un sottogruppo e G è un sottogruppo di sé stesso.

Esempio 1. Un sottospazio W di uno spazio vettoriale V è in particolare un sottogruppo (del gruppo additivo dei vettori).

Esempio 2. Il gruppo delle matrici complesse unitarie è un sottogruppo del gruppo di tutte le matrici complesse invertibili (di fissata dimensione), ecc.

C'è un modo generale per ottenere sottogruppi di un gruppo. Sia S un sottoinsieme di un gruppo G avente almeno un elemento. Sia H l'insieme degli elementi di G costituito da tutti i prodotti $x_1 \dots x_n$ tali che, per ogni indice i , x_i oppure x_i^{-1} è un elemento di S . Allora H è evidentemente un sottogruppo di G , chiamato il sottogruppo *generato* da S .

Diciamo anche che S è un insieme di *generatori* di H . Questa nozione è analoga alla nozione di insieme di generatori di uno spazio vettoriale, incontrata nelle parti precedenti di questo libro. Esempi di generatori saranno dati negli esercizi.

Siano G , G' due gruppi. Un *omomorfismo*

$$f: G \rightarrow G'$$

di G in G' è un'applicazione avente la seguente proprietà: comunque si prendano x e y in G , abbiamo

$$f(xy) = f(x)f(y)$$

(e, nella notazione additiva, $f(x+y) = f(x) + f(y)$).

Esempio 3. Un'applicazione lineare è un omomorfismo.

Esempio 4. Sia K un corpo e si denoti con K^\times il gruppo moltiplicativo dei suoi elementi non nulli. Sia G il gruppo delle matrici invertibili $n \times n$ in K . Allora

$$\text{Det}: G \rightarrow K^\times$$

è un omomorfismo. Ciò non è altro che la regola di moltiplicazione dei determinanti.

Esempio 5. L'applicazione

$$z \mapsto |z|$$

è un omomorfismo del gruppo moltiplicativo dei numeri complessi non nulli nel gruppo moltiplicativo dei numeri complessi non nulli (più precisamente nel gruppo moltiplicativo dei numeri reali positivi).

Esempio 6. L'applicazione

$$x \mapsto e^x$$

è un omomorfismo del gruppo additivo dei numeri reali nel gruppo moltiplicativo dei numeri reali positivi. La sua applicazione inversa, il logaritmo, è ancora un omomorfismo.

Esempio 7. Sia G un gruppo. Sia x un elemento di G . Se n è un intero positivo, definiamo x^n come

$$xx \dots x$$

nel prodotto comprendendo n fattori. Se $n = 0$, definiamo $x^n = e$. Se $n = -m$, m essendo un intero positivo, definiamo

$$x^{-m} = (x^{-1})^m.$$

Si può allora facilmente verificare che la regola

$$x^{m+n} = x^m x^n$$

vale comunque si prendano gli interi m, n . Poiché questa verifica è un po' noiosa noi la omettiamo. Osserviamo però che in conseguenza di questa proprietà, l'applicazione

$$n \mapsto x^n$$

è un omomorfismo del gruppo additivo degli interi \mathbb{Z} in G . Quando G è scritto additivamente, scriviamo nx invece di x^n .

Per brevità talvolta diciamo: "Sia $f: G \rightarrow G'$ un omomorfismo di gruppi", invece di dire: "Siano G, G' due gruppi e sia f un omomorfismo di G in G' ".

Sia $f: G \rightarrow G'$ un omomorfismo di gruppi e siano e, e' gli elementi unità rispettivamente di G e G' . Allora $f(e) = e'$.

Dimostrazione. Abbiamo $f(e) = f(ee) = f(e)f(e)$. Moltiplicando ambo i membri per $f(e)^{-1}$ si ottiene il risultato richiesto.

Sia $f: G \rightarrow G'$ un omomorfismo di gruppi. Sia $x \in G$. Allora

$$f(x^{-1}) = f(x)^{-1}.$$

Dimostrazione. Abbiamo infatti

$$e' = f(e) = f(xx^{-1}) = f(x)f(x^{-1}).$$

Siano $f: G \rightarrow G'$ e $g: G' \rightarrow G''$ due omomorfismi di gruppi. Allora l'applicazione composta $g \circ f$ è un omomorfismo del gruppo G nel gruppo G'' .

Dimostrazione. Abbiamo infatti

$$(g \circ f)(xy) = g(f(xy)) = g(f(x)f(y)) = g(f(x))g(f(y)).$$

Sia $f: G \rightarrow G'$ un omomorfismo di gruppi. Definiamo *nucleo* di f l'insieme di tutti gli elementi $x \in G$ tali che $f(x) = e'$. Si vede immediatamente che il *nucleo* è un sottogruppo di G . (Al nucleo appartiene e perché abbiamo già visto che $f(e) = e'$. Si dimostrino le altre proprietà come semplice esercizio.)

Osserviamo che il nucleo di un'applicazione lineare tra spazi vettoriali coincide col nucleo della stessa applicazione se considerata come omomorfismo di gruppi (cioè omomorfismo di gruppi additivi).

Ricordiamo che un'applicazione $f: S \rightarrow S'$ di un insieme in un

altro viene detta *iniettiva* se $f(x)$ e $f(y)$ sono distinti ogni volta che x e y lo sono.

Sia $f: G \rightarrow G'$ un omomorfismo di gruppi. Se il nucleo di f è costituito soltanto dall'elemento unità e , allora f è iniettiva.

Dimostrazione. Il lettore provi a farla da solo, adattando al caso presente la dimostrazione che abbiamo precedentemente data per la dimostrazione dell'analogia proposizione relativa alle applicazioni lineari.

Sia $f: G \rightarrow G'$ un omomorfismo di gruppi. L'immagine di f è un sottogruppo di G' .

Dimostrazione. Se $x' = f(x)$ con $x \in G$, se $y' = f(y)$ con $y \in G$, allora

$$x'y' = f(xy) = f(x)f(y)$$

è di nuovo un elemento dell'immagine. Inoltre e' appartiene all'immagine e anche $x'^{-1} = f(x^{-1})$ è nell'immagine. Perciò questa immagine è un sottogruppo.

Sia $f: G \rightarrow G'$ un omomorfismo di gruppi. Diciamo che f è un *isomorfismo* (o, più precisamente, un isomorfismo di gruppi) se esiste un omomorfismo $g: G' \rightarrow G$ tale che $f \circ g$ e $g \circ f$ sono le applicazioni identiche di G' e G , rispettivamente.

Esempio 8. La funzione esponenziale è un isomorfismo del gruppo additivo dei numeri reali sul gruppo moltiplicativo dei numeri reali positivi. Il suo inverso è il logaritmo.

Esempio 9. Due spazi vettoriali, se sono isomorfi come spazi vettoriali, sono isomorfi anche come gruppi additivi.

Un omomorfismo di gruppi $f: G \rightarrow G'$ che sia iniettivo e surgettivo (cioè tale che l'immagine di f sia G') è un isomorfismo.

Dimostrazione. Dobbiamo definire l'inversa di f . Per ogni $x' \in G'$, indichiamo con $g(x')$ l'unico elemento x tale che $f(x) = x'$. Questo elemento esiste perché f è surgettiva ed è unico perché f è iniettiva. Resta da dimostrare che g è un omomorfismo. Siano x', y' elementi di G' , siano x, y elementi di G tali che $f(x) = x'$ e $f(y) = y'$. Allora $f(xy) = x'y'$. Perciò per definizione,

$$g(x'y') = xy = g(x')g(y').$$

E questo prova la nostra asserzione.

La dimostrazione della proposizione precedente coincide essenzialmente con la dimostrazione dell'analogia proposizione relativa alle applicazioni lineari, ma è scritta in notazione moltiplicativa.

Si chiama *automorfismo* di un gruppo un isomorfismo di un gruppo con sé stesso.

Esempio 10. Sia V uno spazio vettoriale di dimensione finita sul corpo K e sia $A: V \rightarrow V$ un'applicazione lineare invertibile. Allora A può essere considerata un automorfismo del gruppo additivo di V .

Esempio 11. Sia G un gruppo commutativo. Allora l'applicazione

$$x \mapsto x^{-1}$$

è un automorfismo di G . Se ne faccia la dimostrazione per esercizio. Come si scrive questo automorfismo nella notazione additiva?

Esercizi

1. Sia \mathbb{R}^\times il gruppo moltiplicativo dei numeri reali non nulli. Si descriva esplicitamente il nucleo dell'omomorfismo valore assoluto

$$x \mapsto |x|$$

di \mathbb{R}^\times in sé stesso. Qual è l'immagine di questo omomorfismo?

2. Sia \mathbb{C}^\times il gruppo moltiplicativo dei numeri complessi non nulli. Qual è il nucleo dell'omomorfismo valore assoluto

$$z \mapsto |z|$$

di \mathbb{C}^\times in \mathbb{R}^\times ?

3. Sia S l'insieme di tutte le applicazioni di \mathbb{R}^n in sé stesso che sono o applicazioni reali unitarie oppure traslazioni. (Una *traslazione* $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ è un'applicazione per cui esiste in \mathbb{R}^n un vettore B in modo che, per ogni $X \in \mathbb{R}^n$, $T(X) = X + B$.) Sia G il gruppo generato dagli elementi di S . Questo gruppo G è chiamato il gruppo dei *movimenti rigidi* di \mathbb{R}^n . Dimostrare che se F è un movimento rigido, allora F conserva la distanza, cioè, comunque si prendano in \mathbb{R}^n X e Y , la distanza tra $F(X)$ e $F(Y)$ è uguale alla distanza tra X e Y . Il gruppo unitario è un sottogruppo di G .

4. a) Sia G l'insieme di tutte le applicazioni di \mathbb{R} in sé stesso del tipo $x \mapsto ax + b$, dove $a \in \mathbb{R}$, $a \neq 0$ e $b \in \mathbb{R}$. Dimostrare che G è un gruppo. Denotiamo l'applicazione ora considerata con $\sigma_{a,b}$. Pertanto $\sigma_{a,b}(x) = ax + b$.

b) Associamo ad ogni applicazione $\sigma_{a,b}$ il numero a . Dimostrare che questa associazione

$$\sigma_{a,b} \mapsto a$$

è un omomorfismo di G in \mathbb{R}^\times . Descriverne il nucleo.

5. Sia G l'insieme di tutte le applicazioni di \mathbb{R}^n in sé stesso che sono del tipo

$$\sigma_{A,B}: X \mapsto AX + B,$$

dove A è una matrice invertibile $n \times n$ e B appartiene a \mathbb{R}^n . Dimostrare che G è un gruppo. Dimostrare che l'applicazione definita da

$$\sigma_{A,B} \mapsto A$$

è un omomorfismo di G nel gruppo lineare generale. Descriverne il nucleo.

6. Sia V uno spazio vettoriale di dimensione n sul corpo K . Dimostrare che i gruppi $GL(n, K)$ e $GL(V)$ sono isomorfi.

7. Dimostrare che il gruppo simmetrico S_3 è generato dalle permutazioni

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix}.$$

8. Sia G un gruppo e sia a un suo elemento. Sia

$$\sigma_a: G \rightarrow G$$

l'applicazione tale che

$$\sigma_a(x) = axa^{-1}.$$

Dimostrare che l'insieme di tutte queste applicazioni, per ogni $a \in G$, è un gruppo.

9. Dimostrare che l'insieme degli automorfismi di un gruppo G è esso stesso un gruppo, denotato con $\text{Aut}(G)$.

10. Con riferimento alle notazioni introdotte nell'esercizio 8, dimostrare che l'associazione $a \mapsto \sigma_a$ definisce un omomorfismo di G in $\text{Aut}(G)$. L'immagine di questo omomorfismo è chiamato il gruppo degli *automorfismi interni* di G . Pertanto, un automorfismo interno di G è uno di quelli tali da essere uguali a qualche σ_a per un opportuno a di G .

11. Sia K un corpo. Dimostrare che il gruppo additivo di K è isomorfo al gruppo (moltiplicativo) delle matrici del tipo

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

con $a \in K$.

12. Sia G il gruppo di tutte le matrici

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

dove a, b, d sono elementi di un corpo K e ad non è nullo. Dimostrare che l'applicazione definita da

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \mapsto (a, d)$$

è un omomorfismo di G sul prodotto diretto $K^\times \times K^\times$ (dove K^\times è il gruppo moltiplicativo degli elementi non nulli di K). Se ne descriva il nucleo. Possiamo considerare anche il nostro omomorfismo come avente valori nel gruppo delle matrici diagonali

$$\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}$$

che è isomorfo al gruppo $K^\times \times K^\times$.

13. a) Sia K un corpo e sia $G = G_0$ il gruppo delle matrici 3×3 superiormente triangolari in K , costituito dalle matrici invertibili

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}.$$

Sia G_1 l'insieme di tutte le matrici

$$\begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & 1 & a_{23} \\ 0 & 0 & 1 \end{pmatrix}.$$

Dimostrare che G_1 è un sottogruppo di G e che G_1 è il nucleo dell'omomorfismo che ad ogni matrice triangolare T associa la matrice diagonale contenente gli elementi diagonali di T .

b) Sia G_2 l'insieme delle matrici

$$\begin{pmatrix} 1 & 0 & a_{13} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Dimostrare che G_2 è un sottogruppo di G_1 .

c) Dimostrare che l'applicazione definita da

$$\begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & 1 & a_{23} \\ 0 & 0 & 1 \end{pmatrix} \mapsto (a_{12}, a_{23})$$

è un omomorfismo del gruppo G_1 sul prodotto diretto del gruppo additivo di K con sé stesso. (Denotiamo questo gruppo con $K \times K$). Qual è il suo nucleo?

d) Dimostrare che il gruppo G_2 è isomorfo al gruppo additivo di K .

14. Generalizzare i risultati dell'esercizio 13 dapprima al caso di matrici 4×4 e poi al caso di matrici $n \times n$.

15. a) Sia V uno spazio vettoriale sul corpo K , e sia v_1 un elemento di V . Sia G l'insieme di tutte le applicazioni lineari invertibili A di V in sé stesso tali che $Av_1 = v_1$. Dimostrare che G è un gruppo.

b) Più in generale, sia $\{v_1, \dots, v_i\}$ un sottoinsieme di V . Sia G l'insieme di tutti gli operatori invertibili A definiti su V e tali che

$$Av_1 = v_1, \dots, Av_i = v_i.$$

Dimostrare che G è un gruppo.

c) Siano S un insieme e S' un suo sottoinsieme. Sia G l'insieme di tutte le applicazioni di S in sé stesso invertibili e tali che, per ogni $x \in S'$, $f(x) = x$.
Dimostrare che G è un gruppo.

16. Sia V uno spazio vettoriale di dimensione n sul corpo K . Sia $\{V_1, \dots, V_n\}$ una successione di sottospazi tale che $\dim V_i = i$ e inoltre V_i sia contenuto in V_{i+1} . Sia G l'insieme di tutti gli operatori invertibili definiti su V per i quali $\{V_1, \dots, V_n\}$ è un ventaglio. Sia G_i il sottoinsieme di G costituito da tutti gli operatori A tali che $Av = v$ per ogni $v \in V_i$. Dimostrare che G , G_1, \dots, G_n sono gruppi e che G_{i+1} è un sottogruppo di G_i . Si confronti questa descrizione geometrica con i gruppi di matrici considerati negli esercizi 13 o 14.

17. Sia G un gruppo, sia V uno spazio vettoriale di dimensione finita sul corpo K . Una rappresentazione di G su V è un omomorfismo $\varrho: G \rightarrow GL(V)$ di G nel gruppo delle applicazioni lineari invertibili di V in sé stesso. La rappresentazione viene detta fedele se il nucleo di ϱ è l'elemento unità di G , cioè se ϱ è iniettivo. Se fissiamo in V una base, allora ϱ diviene un omomorfismo di G nel gruppo di matrici $GL(n, K)$, n essendo la dimensione di V . Dimostrare, seguendo il procedimento qui appresso indicato, che ogni gruppo finito G ha sempre qualche rappresentazione. Sia V lo spazio vettoriale delle combinazioni lineari formali di elementi di G , costruito come è stato spiegato nell'appendice al capitolo 13 (§ 66). Gli elementi di G , pertanto, costituiscono una base di questo spazio, siano essi $\{\sigma_1, \dots, \sigma_n\}$. Per ogni elemento σ di G , sia A_σ l'applicazione lineare di V in sé stesso tale che

$$A_\sigma(\sigma_i) = \sigma\sigma_i.$$

Dimostrare che l'associazione $\sigma \mapsto A_\sigma$ è un omomorfismo iniettivo di G in $GL(V)$.

18. Fare degli esempi della situazione descritta nell'esercizio 17. a) Assumere come G il gruppo costituito da due elementi $\{e, \sigma\}$ con $\sigma^2 = e$. b) Si assuma come G il gruppo simmetrico S_3 . In ciascun caso scrivere la matrice associata ad ogni elemento di G , dopo aver fissato un ordinamento tra gli elementi di G stesso.

19. Con riferimento all'esercizio 17, se σ è un elemento di G distinto dall'elemento unità, dimostrare che tutti gli elementi diagonali della matrice A_σ sono nulli. Sempre con riferimento all'esercizio 17, qual è la matrice A_ϵ ?

69. CLASSI LATERALI E SOTTOGRUPPI NORMALI

Siano G un gruppo e H un suo sottogruppo. Sia a un elemento di G . L'insieme di tutti gli elementi ax con $x \in H$ viene chiamato *classe laterale di H in G* . La denotiamo con aH .

Nella notazione additiva, una classe laterale di H sarebbe scritta $a + H$.

Esempio 1. Sia A un fissato vettore in \mathbb{R}^n considerato come gruppo additivo e sia W un sottospazio di \mathbb{R}^n . Allora l'insieme di tutti i vettori $A + X$ con $X \in W$ è una classe laterale di W in \mathbb{R}^n . Possiamo quindi considerare $A + W$ come la traslazione di W secondo il vettore A . Questa situazione è illustrata nella figura 69.1, quando W è una retta di \mathbb{R}^2 .

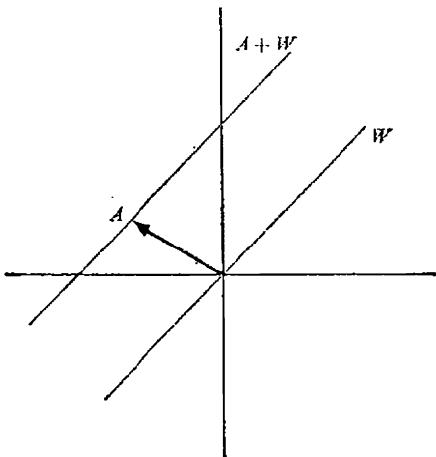


Figura 69.1

Poiché il gruppo G può non essere commutativo, dobbiamo dire che aH è una classe laterale *sinistra* di H . Analogamente possiamo definire le classi laterali *destre*: nel seguito tuttavia, a meno che non sia specificato il contrario, classe laterale significherà sempre classe laterale sinistra.

TEOREMA 1 *Siano aH e bH classi laterali di H in un gruppo G . Allora queste classi laterali, se non coincidono, non hanno elementi in comune.*

Dimostrazione. Supponiamo che esista un elemento appartenente tanto ad aH quanto a bH . Vogliamo far vedere che allora queste classi laterali coincidono. Siano x, y due elementi di H tali che $ax = by$. Allora $a = byx^{-1}$ e perciò, se ax' è un arbitrario elemento di aH , con x' in H , abbiamo

$$ax' = b(yx^{-1})x'.$$

Ma yx^{-1} è un elemento di H e perciò anche $(yx^{-1})x'$ appartiene ad H : possiamo allora concludere che ax' appartiene alla classe laterale bH . Con la stessa argomentazione si prova che la classe laterale bH è contenuta in aH e si conclude quindi che le due classi laterali coincidono.

TEOREMA 2 *Siano G un gruppo e H un suo sottogruppo finito. Allora il numero degli elementi di una qualunque classe laterale aH è uguale al numero degli elementi di H .*

Dimostrazione. Siano x, x' elementi distinti di H . Allora anche gli elementi ax e ax' sono distinti giacché, se $ax = ax'$, moltiplicando a sinistra per a^{-1} si trova $x = x'$. Perciò, se x_1, \dots, x_n sono gli elementi (distinti) che costituiscono H , ax_1, \dots, ax_n sono gli elementi, distinti, che costituiscono aH , e la nostra asserzione è dimostrata.

Se G è un gruppo e H è un suo sottogruppo, il numero delle classi laterali distinte di H in G si chiama l'*indice* di H in G . Questo indice può, naturalmente, non essere finito. Se G è un gruppo finito, allora l'indice di ogni suo sottogruppo è finito. L'indice di un sottogruppo H è denotato con $(G:H)$.

COROLLARIO *Siano G un gruppo finito e H un suo sottogruppo. Allora*

$$\text{ordine di } G = (G:H) \text{ (ordine di } H\text{)}.$$

Dimostrazione. Ogni elemento di G appartiene a qualche classe laterale (infatti, a appartiene alla classe laterale aH , essendo $a = ae$). Il teorema 1 assicura poi che ogni elemento non

può appartenere a più di una classe laterale, mentre il teorema 2 afferma che due classi laterali hanno lo stesso numero di elementi. A questo punto l'uguaglianza da dimostrare è evidente.

Il corollario dice anche che l'ordine di un sottogruppo di un gruppo finito divide l'ordine del gruppo.

Sia G un gruppo. Un suo sottogruppo H viene chiamato *normale* se esso è il nucleo di un omomorfismo di G in qualche gruppo.

TEOREMA 3 *Sia $f: G \rightarrow G'$ un omomorfismo di gruppi. Sia H il suo nucleo e sia a' un elemento di G' appartenente all'immagine di f : ad esempio, sia $a' = f(a)$ per un certo a in G . Allora l'insieme degli elementi x di G tali che $f(x) = a'$ è precisamente la classe laterale aH .*

Dimostrazione. Sia x un elemento di aH , cioè esista h in H in modo che $x = ah$. Allora

$$f(x) = f(a)f(h) = f(a).$$

Viceversa, supponiamo che x appartenga a G e che $f(x) = a'$. Allora

$$f(a^{-1}x) = f(a)^{-1}f(x) = a'^{-1}a' = e'.$$

Perciò $a^{-1}x$ appartiene al nucleo H , esiste cioè h in H in modo che $a^{-1}x = h$. Allora $x = ah$, come si doveva dimostrare.

Sia $f: S \rightarrow S'$ un'applicazione. Se x' è un elemento di S' , denotiamo con $f^{-1}(x')$ l'insieme di tutti gli x di S tali che $f(x) = x'$ e chiamiamo questo insieme *immagine inversa* di x' attraverso f . Solitamente essa consiste di più di un elemento. Nel teorema 3 possiamo dire che l'immagine inversa di un elemento a' di G' è una classe laterale di G .

Esempio 2. Sia A una matrice $n \times n$ in un corpo K e sia $L_A: K^n \rightarrow K^n$ l'applicazione lineare ad essa associata. Sia B un elemento di K^n , allora

$$L_A^{-1}(B)$$

è l'insieme delle soluzioni dell'equazione lineare $AX = B$. Il teorema 3 generalizza la circostanza che questo insieme di soluzioni è una classe laterale del nucleo di L_A , nel caso in cui esiste almeno

una soluzione. Infatti, se X_0 è una soluzione e W è il nucleo di L_A , allora

$$L_A^{-1}(B) = X_0 + W.$$

Abbiamo già visto tutto ciò discutendo delle equazioni lineari: ora abbiamo una nuova denominazione per questa circostanza. L'ipotesi dell'esistenza di una soluzione almeno, si traduce nel chiedere che il vettore B appartenga all'immagine di L_A .

Vogliamo ora dare un criterio semplice per riconoscere se un sottogruppo è normale. Abbiamo bisogno di una notazione più conveniente. Siano S, S' sottoinsiemi di un gruppo G . Definiamo SS' come l'insieme di tutti gli elementi xx' con $x \in S$ e $x' \in S'$. È facile allora verificare che se S_1, S_2, S_3 sono tre sottoinsiemi di G , allora

$$(S_1 S_2) S_3 = S_1 (S_2 S_3).$$

Questo prodotto è costituito semplicemente da tutti gli elementi xyz , con $x \in S_1, y \in S_2, z \in S_3$. Se H è un sottogruppo di G , si verifica immediatamente che $HH = H$.

TEOREMA 4 *Sia G un gruppo e sia H un suo sottogruppo tale che, per ogni x di G , $xH = Hx$. Se aH e bH sono classi laterali di H , allora il prodotto $(aH)(bH)$ è ancora una classe laterale e l'insieme di tutte le classi laterali è un gruppo rispetto all'operazione di prodotto più sopra definita.*

Dimostrazione. Abbiamo intanto $(aH)(bH) = aHbH = abH$. Questo dice che il prodotto di due classi laterali è una classe laterale. La condizione GR 1 è soddisfatta a causa delle osservazioni precedenti relative alla moltiplicazione tra sottoinsiemi di G . La condizione GR 2 è anch'essa soddisfatta; l'elemento unità risulta essere la classe laterale $eH = H$. (Eseguire la verifica in dettaglio.) La condizione GR 3 è anch'essa soddisfatta, l'inversa della classe laterale aH è la classe laterale $a^{-1}H$. (Anche in questo caso eseguire la verifica in dettaglio.) Abbiamo così provato il teorema.

Il gruppo delle classi laterali considerato nel teorema 4 è chiamato *gruppo quoziante* di G per H ed è denotato con G/H . Osserviamo che si tratta di un gruppo di classi laterali sinistre o destre, non essendovi tra loro alcuna differenza per l'ipotesi

fatta su H . Mettiamo l'accento sul fatto che è proprio questa ipotesi che ci ha permesso di definire la moltiplicazione tra classi laterali. Se la condizione $xH = Hx$ per ogni x di G non è soddisfatta, non è possibile definire un gruppo con le classi laterali.

COROLLARIO *Siano G un gruppo e H un suo sottogruppo avente la proprietà che $xH = Hx$ per ogni x di G . Sia G/H il gruppo quoziante e sia*

$$f: G \rightarrow G/H$$

l'applicazione che associa ad ogni a di G la classe laterale $f(a) = aH$. Allora f è un omomorfismo e il suo nucleo è precisamente H . Perciò H è un sottogruppo normale.

Dimostrazione. L'essere f un omomorfismo è nient'altro che una ripetizione della definizione di prodotto tra classi laterali. Relativamente al suo nucleo, è chiaro che ogni elemento di H è nel nucleo. Viceversa, se x è in G e se $f(x) = xH$ è l'elemento unità di G/H , si tratta della classe laterale H stessa, perciò $xH = H$. Questo significa che $xe = x$ è un elemento di H per cui H è il nucleo di f , come si voleva.

Esercizi

1. Sia $f: G \rightarrow G'$ un omomorfismo di nucleo H . Si assuma G finito. Dimostrare allora che

$$\text{ordine di } G = (\text{ordine dell'immagine di } f) (\text{ordine di } H).$$

Si confronti col teorema analogo sulle dimensioni delle applicazioni lineari.

2. a) Sia H un sottogruppo normale di un gruppo G . Dimostrare che se x è in H e se a è in G , allora anche axa^{-1} è in H .

- b) Sia H un sottogruppo normale di G . Dimostrare che la classe laterale sinistra aH coincide con la classe laterale destra Ha .

3. Sia H un sottogruppo di G e si assuma che, per ogni x di G , $xHx^{-1} = H$. Allora, per ogni x di G , $x^{-1}Hx = H$ e anche $Hx = xH$.

4. Sia G un gruppo e H un suo sottogruppo. Dimostrare che H è normale se, e soltanto se, $xHx^{-1} = H$ per ogni x in G .

5. Dimostrare che se G è commutativo allora ogni sottogruppo è normale.

6. Siano H_1, H_2 due sottogruppi normali di G . Dimostrare che anche il sottogruppo $H_1 \cap H_2$ è normale.

7. Sia $f: G \rightarrow G'$ un omomorfismo e sia H' un sottogruppo di G' . Dimostrare che $f^{-1}(H')$ è un sottogruppo di G . Se H' è un sottogruppo normale di G' , dimostrare che $f^{-1}(H')$ è un sottogruppo normale di G .

8. Sia $f: G \rightarrow G'$ un omomorfismo surgettivo. Sia H un sottogruppo normale di G . Dimostrare che $f(H)$ è un sottogruppo normale di G' .

9. In ognuno dei seguenti casi specifichiamo un gruppo e un suo sottogruppo. Dire se il sottogruppo è normale nel gruppo dato.

a) $G = GL_n(K)$ e H = gruppo delle matrici $n \times n$ in K con determinante uguale a 1.

b) $G = GL_n(\mathbb{R})$ e $H = U_n(\mathbb{R})$.

c) $G = GL_n(K)$ e H = gruppo delle matrici diagonali (invertibili) in K .

d) G = gruppo delle matrici superiormente triangolari in K e H = gruppo delle matrici superiormente triangolari in K aventi gli elementi diagonali tutti uguali a 1.

e) G = gruppo simmetrico S_n e H = gruppo delle permutazioni pari.

f) G = gruppo simmetrico S_n e H = sottogruppo delle permutazioni che lasciano fisso l'intero n (cioè tali che $\sigma(n) = n$).

10. Sia G_0 il gruppo di tutte le matrici

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

dove a, b, c, d sono numeri interi e aventi determinante uguale a 1. Scrivere tre elementi di questo gruppo.

11. Sia G il gruppo di tutte le matrici

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

in cui a, b, c, d sono numeri interi, e aventi determinante uguale a 1 oppure a -1 . Dimostrare che il gruppo considerato nell'esercizio 10 è un sottogruppo normale di G . Dimostrare che il gruppo quoziente G/G_0 ha ordine 2.

12. Sia G un gruppo. Si definisca *centro* di G l'insieme di tutti gli elementi a in G tali che, per ogni x di G , $ax = xa$. Dimostrare che il centro è un sottogruppo e, più precisamente, che è un sottogruppo normale. Dimostrare che il centro è il nucleo dell'omomorfismo considerato nell'esercizio 10 del paragrafo 68.

13. Sia G un gruppo commutativo, sia H un suo sottogruppo. Dimostrare che G/H è commutativo.

14. Sia G un gruppo finito con n elementi a_1, \dots, a_n . Dimostrare che, per ogni x di G , gli elementi xa_1, \dots, xa_n sono distinti e quindi che costituiscono una permutazione di a_1, \dots, a_n . Perciò, ad ogni x di G possiamo asso-

ciare una permutazione σ_x dell'insieme $\{1, \dots, n\}$ tale che

$$xa_i = a_{\sigma_x(i)}$$

per $i = 1, \dots, n$. Dimostrare che l'applicazione definita da $x \mapsto \sigma_x$ è un omomorfismo iniettivo di G sul gruppo simmetrico S_n . In questo modo possiamo considerare ogni gruppo G come sottogruppo di un gruppo di permutazioni.

15. Sia $f: G \rightarrow G'$ un omomorfismo di gruppi e sia H il suo nucleo. Si assuma che G' sia l'immagine di f . Dimostrare che i gruppi G/H e G' sono isomorfi.

16. Siano G un gruppo e H un suo sottogruppo. Sia N_H l'insieme di tutti gli x di G tali che $xHx^{-1} \subset H$. Dimostrare che N_H è un gruppo contenente H e che H è un sottogruppo normale di N_H .

17. Siano G un gruppo, H un suo sottogruppo, N un suo sottogruppo normale. Dimostrare che NH è un sottogruppo e che $NH = HN$.

70. GRUPPI CICLICI

L'insieme degli interi \mathbb{Z} è un gruppo rispetto all'addizione. Ne vogliamo determinare tutti i sottogruppi. Sia H un sottogruppo di \mathbb{Z} . Se H non è banale, sia a il più piccolo intero positivo di H . Vogliamo far vedere che H è costituito da tutti gli elementi na , con n in \mathbb{Z} . Per dimostrare questo, sia $y \in H$ e sia y positivo. Esistono allora due interi n, r tali che $0 < r < a$ e inoltre

$$y = na + r.$$

Poiché H è un sottogruppo e $r = y - na$, vediamo che r appartiene a H e perciò r deve essere nullo. Se y è negativo ripetiamo la dimostrazione ora fatta per $-y$ che, di nuovo, appartiene a H poiché H è un sottogruppo.

Sia G un gruppo. Diremo che G è *ciclico* se esiste un elemento a di G tale che ogni elemento x di G possa essere scritto nella forma a^n per un opportuno intero n . (Questo equivale a dire che l'applicazione $f: \mathbb{Z} \rightarrow G$ tale che $f(n) = a^n$ è surgettiva.) Un siffatto elemento a di G è chiamato un *generatore* di G .

Sia G un gruppo e sia a un suo elemento. L'insieme di tutti gli elementi a^n ($n \in \mathbb{Z}$) è evidentemente un sottogruppo ciclico di G . Se m è un intero tale che $a^m = e$ e m è positivo, diciamo allora che m è un *esponente* di a .

Siano G un gruppo e a un suo elemento. Sia $f: \mathbb{Z} \rightarrow G$ l'omo-

morfismo tale che $f(n) = a^n$ e sia H il nucleo di f . Allora due casi sono possibili:

1) Il nucleo è banale. Allora f è un isomorfismo di \mathbb{Z} sul sottogruppo ciclico di G generato da a , poiché f è iniettiva e l'immagine di f è precisamente uguale a questo sottogruppo. Inoltre questo sottogruppo è ciclico e infinito. Se a genera G , allora G è ciclico. Diciamo allora che a ha *periodo infinito*.

Esempio 1. Il numero 2 genera un sottogruppo ciclico infinito del gruppo moltiplicativo dei numeri complessi non nulli. I suoi elementi sono

$$\dots, 2^{-5}, 2^{-4}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1, 2, 4, 8, 2^4, 2^5, \dots$$

2) Il nucleo non è banale. Sia allora d il più piccolo intero positivo appartenente a questo nucleo. Allora d è chiamato il *periodo* di a . Se m è un intero tale che $a^m = e$, allora $m = ds$ per un opportuno intero s , per quanto abbiamo dimostrato all'inizio del paragrafo. Osserviamo anche che gli elementi

$$e, a, \dots, a^{d-1}$$

sono distinti. Infatti, sia $a^r = a^s$ con $0 < r < d-1$ e

$$0 < s < d-1,$$

se, inoltre, per esempio, $r < s$, allora $a^{s-r} = e$. Poiché

$$0 < s - r < d,$$

deve essere $s - r = 0$, cioè $r = s$. Concludiamo affermando che il gruppo ciclico generato da a in questo caso ha ordine d .

Esempio 2. Il gruppo moltiplicativo $\{1, -1\}$ è ciclico e di ordine 2.

Esempio 3. I numeri complessi $\{1, i, -1, -i\}$ costituiscono un gruppo ciclico di ordine 4 (rispetto alla moltiplicazione). Il numero i ne è un generatore.

Esempio 4. La matrice

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

è un generatore di un gruppo ciclico di ordine 4 (con la moltiplicazione come legge di composizione). Verificarlo in dettaglio.

TEOREMA 4 *Sia G un gruppo finito e sia a un suo elemento. Allora il periodo di a divide l'ordine di G .*

Dimostrazione. L'ordine del sottogruppo generato da a è uguale a d . Ora basta applicare il corollario del teorema 2 del paragrafo 69.

TEOREMA 5 *Sia G un gruppo ciclico. Allora ogni sottogruppo di G è ciclico.*

Dimostrazione. Sia a un generatore di G , allora possiamo considerare un omomorfismo surgettivo

$$f: \mathbb{Z} \rightarrow G$$

tale che $f(n) = a^n$. Sia H un sottogruppo di G : allora $f^{-1}(H)$ (cioè l'insieme degli interi n tali che $f(n) \in H$) è un sottogruppo A di \mathbb{Z} , e perciò è ciclico. Infatti sappiamo che esiste un unico intero positivo d tale che $f^{-1}(H)$ sia costituito da tutti gli interi che possono essere scritti nella forma md con $m \in \mathbb{Z}$. Poiché f è surgettiva, segue che f applica A su tutto H , cioè ogni elemento di H è della forma a^{md} per qualche intero m . Segue così che H è ciclico, e infatti è generato da a^d .

Esercizi

1. Una radice dell'unità nei numeri complessi è un numero ω tale che $\omega^n = 1$ per qualche intero positivo n . Diciamo allora che ω è una radice n -esima dell'unità. Descrivere l'insieme delle radici n -esime dell'unità in \mathbb{C} . Dimostrare che si tratta di un gruppo ciclico di ordine n .

2. Determinare i periodi delle matrici seguenti

$$\text{a)} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{b)} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Generalizzare il risultato al caso di matrici $n \times n$.

3. Sia G un gruppo finito. Dimostrare che ogni elemento di G ha periodo finito.

4. Siano $\omega_1, \dots, \omega_n$ radici dell'unità e si supponga che la matrice

$$A = \begin{pmatrix} \omega_1 & 0 & \dots & 0 \\ 0 & \omega_2 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \omega_n \end{pmatrix}$$

abbia periodo d . Sia B una matrice invertibile $n \times n$. Dimostrare che anche la matrice $B^{-1}AB$ ha periodo d .

5. In generale, siano A, B matrici invertibili $n \times n$. Dimostrare che le matrici A e $B^{-1}AB$ hanno lo stesso periodo.

6. Sia $f: G \rightarrow G'$ un isomorfismo di gruppi. Sia a un elemento di G . Dimostrare che gli elementi a e $f(a)$ hanno lo stesso periodo.

7. Si consideri il gruppo additivo degli interi \mathbb{Z} . Dimostrare che esso ha soltanto due generatori, precisamente 1 e -1 . Dimostrare in generale che ogni gruppo ciclico infinito ha soltanto due generatori.

8. Sia S_3 il gruppo simmetrico e sia $\varepsilon: S_3 \rightarrow \{1, -1\}$ l'omomorfismo definito dal segno della permutazione. Qual è l'ordine del nucleo di ε ?

9. Stessa domanda relativa al gruppo S_n invece che S_3 .

10. Dimostrare che un gruppo finito il cui ordine è un numero primo è necessariamente ciclico.

11. Dimostrare che un gruppo di ordine 4 o è ciclico oppure contiene due elementi distinti a, b tali che $a^2 = b^2 = e$, e inoltre $ab = ba$.

12. Dimostrare che se A è una matrice $n \times n$ avente periodo finito, allora tutti gli autovalori di A sono radici dell'unità.

13. Siano A, B due gruppi ciclici finiti di rispettivi ordini m e n , si assuma che m e n siano primi tra loro. Dimostrare che $A \times B$ è un gruppo ciclico. Qual è il suo ordine?

14. Sia G un gruppo ciclico e sia $f: G \rightarrow G'$ un omomorfismo. Dimostrare che anche l'immagine di f è un gruppo ciclico.

15. Sia G un gruppo ciclico finito di ordine n . Dimostrare che per ogni intero positivo d divisore di n esiste un sottogruppo di ordine d .

16. Con riferimento all'esercizio 15, dimostrare che il sottogruppo di ordine d è univocamente determinato.

17. Sia G un gruppo ciclico finito di ordine n . Sia a un suo generatore e sia r un intero non nullo primo con n . Dimostrare che a^r è un generatore di G . Dimostrare che ogni generatore di G può essere scritto in questa forma.

18. Sia G un gruppo ciclico di ordine p , p essendo numero primo. Quanti generatori possiede G ?

19. Siano A, B due gruppi abeliani. Dimostrare che gli omomorfismi di A in B costituiscono un gruppo. (Se scriviamo i gruppi A, B additivamente, e se $f, g: A \rightarrow B$ sono omomorfismi, si definisca $f+g$ come l'applicazione tale che $(f+g)(x) = f(x) + g(x)$.) Qual è l'elemento unità?

20. Supponiamo che G sia un gruppo ciclico di ordine n e che Z_n sia un altro gruppo ciclico di ordine n . Dimostrare che il gruppo degli omomorfismi di G in Z_n è anch'esso ciclico di ordine n .

21. Sia A un gruppo abeliano, denotato additivamente, e sia n un intero positivo tale che, per ogni x in A , $nx = 0$. Supponiamo di potere scrivere $n = rs$, con r e s interi positivi primi tra loro. Se A_r è costituito da tutti gli x di A tali che $rx = 0$ e se A_s è costituito da tutti gli x di A tali che $sx = 0$, dimostrare che ogni elemento a di A può essere univocamente espresso nella forma $a = b + c$, con $b \in A_r$ e $c \in A_s$. (Questo risultato è l'analogo del teorema 5, cap. 12, § 59.)

22. Sia A un gruppo abeliano additivo e siano B, C suoi sottogruppi. Si indichi con $B+C$ l'insieme di tutte le somme $b+c$, con b in B e c in C . Dimostrare che $B+C$ è un sottogruppo, chiamato la *somma* di B e C . Si definisca analogamente la somma di un numero finito di sottogruppi.

Diciamo che A è *somma diretta* di B e C se ogni elemento x di A può essere univocamente espresso nella forma $x = b+c$ con b in B e c in C , analogamente nel caso di più sottogruppi.

23. Dimostrare che il gruppo abeliano additivo A è la somma diretta dei sottogruppi B e C se, e soltanto se, $A = B+C$ e $B \cap C = \{0\}$.

24. Sia A un gruppo abeliano finito di ordine n . Sia

$$n = p_1^{r_1} \dots p_t^{r_t}$$

la decomposizione di n in potenze di numeri primi p_i distinti. Dimostrare che A è una somma diretta $A = A_1 \oplus \dots \oplus A_t$, dove ogni elemento A_i ha periodo che divide $p_i^{r_i}$.

71. GRUPPI ABELIANI LIBERI

In questo paragrafo tratteremo soltanto di gruppi commutativi. Vogliamo analizzare le condizioni sotto le quali sia possibile definire qualcosa di analogo alle basi per questi gruppi.

Sia A un gruppo abeliano. Una *base* di A è definita come un insieme di elementi $\{v_1, \dots, v_n\}$ ($n > 1$) di A tali che ogni elemento di A si possa univocamente esprimere come somma

$$c_1 v_1 + \dots + c_n v_n$$

dove i c_i sono interi. Si vede quindi che una base di un gruppo

abeliano è definita in modo del tutto analogo a una base di uno spazio vettoriale, con la differenza che ora i coefficienti c_1, \dots, c_n devono essere interi.

TEOREMA 6 *Sia A un gruppo abeliano e sia $\{v_1, \dots, v_n\}$ una sua base. Sia B un altro gruppo abeliano e siano w_1, \dots, w_n elementi di B . Esiste allora un unico omomorfismo di gruppi $f: A \rightarrow B$ tale che $f(v_i) = w_i$ per ogni $i = 1, \dots, n$.*

Dimostrazione. Basta ricopiare la dimostrazione fatta per l'analogia proposizione relativa agli spazi vettoriali, omettendo le costanti inutili.

I teoremi relativi alla possibilità di estendere una base di un sottogruppo non sono più validi quando si tratta di gruppi abeliani. Qualcosa tuttavia continua a valere.

Per evitare possibili confusioni quando si considerino basi di gruppi abeliani nel senso sopraspecificato e basi di spazi vettoriali, chiameremo *Z-basi* quelle dei gruppi abeliani.

In accordo con lo spirito geometrico di questo libro, nella dimostrazione dei nostri risultati adopereremo talvolta argomentazioni di carattere geometrico.

TEOREMA 7 *Sia A un sottogruppo non banale di \mathbb{R}^n . Si assuma che in ogni regione limitata dello spazio esista soltanto un numero finito di elementi di A . Sia m il numero massimo di elementi di A che sono linearmente indipendenti su \mathbb{R} . Allora è possibile trovare in A m elementi linearmente indipendenti su \mathbb{R} e costituenti una Z-base di A .*

Dimostrazione. Sia $\{w_1, \dots, w_m\}$ un insieme massimale di elementi appartenenti ad A , linearmente indipendenti su \mathbb{R} . Sia V lo spazio vettoriale generato da questi elementi e sia V_{m-1} lo spazio generato dagli elementi w_1, \dots, w_{m-1} . Sia A_{m-1} l'intersezione di A con V_{m-1} . Allora, certamente, in ogni regione dello spazio esiste soltanto un numero finito di elementi di A_{m-1} . Perciò, se m supera 1, possiamo, procedendo per induzione, assumere che $\{w_1, \dots, w_{m-1}\}$ sia una Z-base di A .

Si consideri ora l'insieme S di tutti gli elementi di A che possono essere scritti nella forma

$$t_1 w_1 + \dots + t_m w_m$$

con $0 < t_i < 1$ se $i = 1, \dots, m-1$, mentre $0 < t_m < 1$. Questo insieme S è certamente limitato e quindi contiene soltanto un numero finito di elementi (tra cui w_m). Scegliamo in questo insieme un elemento v_m in modo che la sua ultima coordinata t_m sia quella positiva minima. Vogliamo dimostrare che l'insieme

$$\{w_1, \dots, w_{m-1}, v_m\}$$

è una Z-base di A . Scriviamo intanto v_m come combinazione lineare di w_1, \dots, w_m a coefficienti reali:

$$v_m = c_1 w_1 + \dots + c_m w_m, \quad 0 < c_m < 1.$$

Sia v un elemento di A e sia

$$v = x_1 w_1 + \dots + x_m w_m$$

con $x_i \in \mathbb{R}$. Sia q_m l'intero per cui si ha

$$q_m c_m \leq x_m < (q_m + 1) c_m.$$

Allora l'ultima coordinata del vettore $v - q_m v_m$ rispetto alla base $\{w_1, \dots, w_m\}$ è uguale a $x_m - q_m c_m$ e quindi

$$0 < x_m - q_m c_m < (q_m + 1) c_m - q_m c_m = c_m < 1.$$

Siano ora q_i ($i = 1, \dots, m-1$) interi tali che

$$q_i \leq x_i < q_i + 1.$$

Allora

$$v - q_m v_m - q_1 w_1 - \dots - q_{m-1} w_{m-1} \quad [1]$$

è un elemento di S . Se la sua ultima coordinata non fosse nulla, esso sarebbe un elemento con ultima coordinata minore di c_m , in contrasto con la scelta di v_m . La sua ultima coordinata quindi deve essere nulla e si conclude che l'elemento [1] appartiene a V_{m-1} . Per l'ipotesi di induzione questo elemento può essere scritto come combinazione lineare a coefficienti interi di w_1, \dots, w_{m-1} . È inoltre chiaro che gli elementi w_1, \dots, w_{m-1}, v_m sono linearmente indipendenti su \mathbb{R} , essi quindi soddisfano tutti i requisiti del nostro teorema.

Siamo ora in grado di estendere il nostro teorema a gruppi più generali. Sia A un gruppo additivo e sia $f: A \rightarrow A'$ un iso-

morfismo di A con un altro gruppo A' . Se A' ha una base, per esempio $\{v'_1, \dots, v'_n\}$, e se v_i è l'elemento di A tale che $f(v_i) = v'_i$, è allora immediatamente chiaro che $\{v_1, \dots, v_n\}$ è una base di A .

TEOREMA 8 *Sia A un gruppo additivo avente una base costituita da n elementi. Sia B un suo sottogruppo non costituito dal solo $\{0\}$. Allora B ha una base contenente al più n elementi.*

Dimostrazione. Sia $\{v_1, \dots, v_n\}$ una base di A . Il teorema 6 afferma l'esistenza di un omomorfismo

$$f: A \rightarrow \mathbb{R}^n$$

tale che $f(v_i) = e^i$ per $i = 1, \dots, n$ e questo omomorfismo è ovviamente iniettivo. Si può allora dire che f stabilisce un isomorfismo tra A e la sua immagine in \mathbb{R}^n . D'altra parte è immediato verificare che in ogni regione limitata di \mathbb{R}^n vi sono soltanto un numero finito di elementi appartenenti all'immagine $f(A)$, giacché in ogni regione limitata le componenti di un vettore

$$(c_1, \dots, c_n)$$

sono limitate. Il teorema 7 permette di concludere che $f(B)$ ha una \mathbb{Z} -base e perciò anche B ha una \mathbb{Z} -base.

TEOREMA 9 *Sia A un gruppo additivo avente una base costituita da n elementi. Allora tutte le basi di A sono costituite da questo stesso numero di elementi.*

Dimostrazione. Consideriamo di nuovo l'omomorfismo $f: A \rightarrow \mathbb{R}^n$ già visto nella dimostrazione del teorema 8. Sia $\{w_1, \dots, w_m\}$ una base di A . Ogni v_i è una combinazione lineare a coefficienti interi di w_1, \dots, w_m . Perciò $f(v_i) = e^i$ è una combinazione lineare a coefficienti interi di $f(w_1), \dots, f(w_m)$. Allora e^1, \dots, e^n appartengono allo spazio generato da $f(w_1), \dots, f(w_m)$. Dai teoremi sulle basi degli spazi vettoriali segue allora che n non supera m e si conclude (per il teorema 8) che m e n sono uguali.

Esercizi

1. Sia A un gruppo abeliano con un numero finito di generatori e si supponga che in A non esista alcun elemento diverso dall'unità avente periodo finito. Conveniamo di scrivere A con la notazione additiva. Se d è un

numero intero positivo, dimostrare che l'applicazione definita da $x \mapsto dx$ è un omomorfismo iniettivo di A in sé stesso avente immagine isomorfa ad A .

2. Con riferimento alle notazioni introdotte nell'esercizio precedente, sia $\{a_1, \dots, a_m\}$ un insieme di generatori di A . Sia $\{a_1, \dots, a_r\}$ un sottoinsieme massimale linearmente indipendente su \mathbb{Z} . Sia B il sottogruppo generato dagli elementi a_1, \dots, a_r . Dimostrare che esiste un intero positivo d tale che, per ogni x in A , dx appartiene a B . Tenendo conto del teorema 7 e dell'esercizio precedente, concludere che il gruppo A ha una base.

Capitolo 15

Anelli

Nelle parti svolte finora quasi mai abbiamo avuto occasione di considerare insieme all'addizione e alla moltiplicazione anche la divisione. Può quindi riuscire utile assiomatizzare la struttura che riguarda soltanto le addizioni e le moltiplicazioni.

Molti esempi e molte proposizioni di questo capitolo sono dati come semplici esercizi: invitiamo il lettore a trattarli in tutti i dettagli.

72. ANELLI E IDEALI .

Un anello R è un insieme i cui elementi possono essere addizionati e moltiplicati (cioè sono assegnate le applicazioni $(x, y) \mapsto x + y$ e $(x, y) \mapsto xy$ che ad ogni coppia ordinata di elementi di R associano un elemento di R stesso), in modo che siano soddisfatte le condizioni seguenti:

AN 1. *Rispetto all'addizione, R è un gruppo additivo (abeliano).*

AN 2. *Per ogni scelta di x, y, z in R abbiamo*

$$x(y+z) = xy + xz \quad \text{e} \quad (y+z)x = yx + zx.$$

AN 3. *Comunque si scelgano x, y, z in R , abbiamo $(xy)z = x(yz)$.*

AN 4. *Esiste in R un elemento e tale che, per ogni x in R abbiamo $ex = xe = x$.*

Esempio 1. Sia R l'insieme dei numeri interi relativi \mathbb{Z} . Allora (rispetto alle usuali operazioni) R è un anello.

Esempio 2. Ogni corpo è anche un anello.

Esempio 3. Sia R l'insieme delle funzioni reali continue definite nell'intervallo $[0, 1]$. La somma e il prodotto di due funzioni f, g sono definiti nel modo solito: $(f+g)(t)=f(t)+g(t)$, e $(fg)(t)=f(t)g(t)$. Rispetto a queste operazioni R è un anello.

L'esempio 3 può essere immediatamente generalizzato considerando funzioni definite in un insieme qualsiasi e aventi valori in un corpo oppure in un anello.

Esempio 4. Se K è un corpo, allora l'insieme dei polinomi di $K[t]$ è un anello (rispetto alle usuali operazioni tra polinomi).

Esempio 5. Anche i polinomi in più variabili $K[t_1, \dots, t_n]$ costituiscono un anello.

Esempio 6. Le matrici $n \times n$, $\text{Mat}_n(K)$, su un corpo K costituiscono un anello: si tratta di un'altra formulazione di alcune proprietà concernenti l'addizione e la moltiplicazione tra matrici.

Esempio 7. Sia V uno spazio vettoriale sul corpo K . Allora gli operatori di V in sé stesso costituiscono un anello $\mathcal{L}(V, V)$. In questo caso la moltiplicazione è la composizione delle applicazioni. Anche qui, il fatto che gli assiomi che definiscono un anello siano soddisfatti è nient'altro che una riformulazione delle proprietà concernenti la composizione tra applicazioni lineari.

In un anello le usuali leggi dell'aritmetica possono essere tutte facilmente dedotte dagli assiomi. Per esempio, abbiamo che, per ogni x in R , $0x = 0$ e $(-e)x = -x$. Ne lasciamo le dimostrazioni per esercizio. Abbiamo anche $(-e)(-e) = e$. Per dimostrarlo, partendo dall'uguaglianza $e + (-e) = 0$ e moltiplicando per $-e$ otteniamo

$$-e + (-e)(-e) = 0.$$

Addizionando a entrambi i membri e si ottiene $(-e)(-e) = e$, come volevamo.

Osservazione. Può talvolta riuscire utile considerare soltanto i primi due assiomi AN 1 e AN 2, specialmente nel caso in cui si tratti di uno spazio vettoriale. Più precisamente, se V è uno spazio vettoriale su un corpo K e se è assegnata un'applicazione bilineare $V \times V \rightarrow V$, che consideriamo come una moltiplicazione, diciamo che V è una K -algebra per distinguere V dagli anelli

perché non abbiamo presupposto né l'associatività della moltiplicazione né l'esistenza di un elemento unità.

Esempio 8. Sia V lo spazio vettoriale delle funzioni continue definite sulla retta reale e periodiche di periodo 2π . Per ogni f, g in V , definiamo il prodotto $f * g$ con la formula

$$(f * g)(t) = \int_{-\pi}^{\pi} f(t-u)g(u) du.$$

Semplici proprietà dell'integrale mostrano che le condizioni AN 1, 2, 3 sono soddisfatte. Il prodotto ora definito prende nome di *convoluzione*.

Esempio 9. Sia G un gruppo finito. Sia K un corpo e sia V lo spazio vettoriale delle funzioni di G in K . Se f, g sono due tali funzioni, definiamo il loro prodotto di convoluzione $f * g$ mediante la formula

$$(f * g)(x) = \sum_{y \in G} f(xy^{-1})g(y).$$

È di nuovo un semplice esercizio verificare che le condizioni AN 1, 2, 3 sono soddisfatte. In questo caso, inoltre, anche la condizione AN 4 è soddisfatta. Qual è l'elemento unità?

Un esempio in cui non vale la proprietà associativa della moltiplicazione è quello trattato nell'esercizio 8.

Come nel caso dei gruppi, si verifica immediatamente che l'elemento e di un anello è unico. Viene chiamato *elemento unità* e spesso viene denotato con 1.

¹ Se in un anello vale l'uguaglianza $xy = yx$ per ogni scelta di x, y in R , allora l'anello considerato viene detto *commutativo*.

Sia R un anello. Un *ideale sinistro* di R è un sottoinsieme J di R avente la proprietà seguente: se x, y sono in J allora $x + y$ è in J , l'elemento zero appartiene a J , se x è in J , allora, per ogni a in R , ax è in J .

Dalle proprietà dell'elemento $-e$ segue che se J è un ideale sinistro e se $x \in J$, allora anche $-x$ è in J giacché $-x = (-e)x$. Ne segue che gli elementi di un ideale sinistro costituiscono un sottogruppo additivo di R e possiamo anche dire che un ideale sinistro è un sottogruppo additivo J di R tale che, se $x \in J$ e $a \in R$, allora $ax \in J$.

Osserviamo che R stesso è un ideale sinistro, chiamato *ideale unità*, e così pure è un ideale sinistro il sottoinsieme di R costituito dal solo 0.

In modo del tutto analogo possono venir definiti gli *ideali destri* e gli *ideali bilateri*. Un ideale bilatero J , allora, è per definizione un sottogruppo additivo di R tale che, se x è in J e se a è in R , allora tanto ax quanto xa sono in J .

Esempio 10. Sia R l'anello delle funzioni reali continue definite nell'intervallo $[0, 1]$. Sia J il sottoinsieme costituito dalle funzioni f tali che $f(\frac{1}{2}) = 0$. Allora J è un ideale (bilatero, giacché R è commutativo).

Esempio 11. Sia R l'anello degli interi relativi \mathbf{Z} . Allora gli interi pari, cioè gli interi del tipo $2n$ con $n \in \mathbf{Z}$, costituiscono un ideale. Gli interi dispari costituiscono anch'essi un ideale?

Esempio 12. Sia R un anello e a sia un suo elemento. L'insieme degli elementi xa , con $x \in R$, è un ideale sinistro, chiamato l'ideale sinistro *principale* generato da a . (Verificare in dettaglio che effettivamente si tratta di un ideale sinistro.) Questo ideale viene denotato con (a) . Più in generale, se a_1, \dots, a_n sono elementi di R , l'insieme di tutti gli elementi

$$x_1 a_1 + \dots + x_n a_n$$

con $x_i \in R$ è un ideale sinistro denotato con (a_1, \dots, a_n) . Gli elementi a_1, \dots, a_n sono chiamati *generatori* di questo ideale.

Esempio 13. Sia R un anello, siano L, M ideali sinistri. Denotiamo con LM l'insieme di tutti gli elementi $x_1 y_1 + \dots + x_n y_n$ con $x_i \in L$ e $y_i \in M$. È un semplice esercizio per il lettore verificare che anche LM è un ideale sinistro. Verificare anche che se L, M, N sono ideali sinistri, allora $(LM)N = L(MN)$.

Esempio 14. Siano L, M ideali sinistri. Definiamo $L + M$ come il sottoinsieme costituito da tutti gli elementi $x + y$ con $x \in L$ e $y \in M$. Allora $L + M$ è un ideale sinistro. Dopo aver verificato ciò in dettaglio, dimostrare anche che, se L, M, N sono ideali sinistri, allora

$$L(M + N) = LM + LN.$$

Enunciare e dimostrare inoltre gli analoghi degli esempi 13 e 14 per ideali destri e bilateri.

Esempio 15. Sia L un ideale sinistro e si denoti con LR l'insieme di tutti gli elementi $x_1y_1 + \dots + x_ny_n$ con $x_i \in L$ e $y_i \in R$. Allora LR è un ideale bilatero. La dimostrazione viene lasciata per esercizio.

Sia R un anello. *Sottoanello* R' è un sottoinsieme di R tale che l'elemento unitario R è in R' , e che, se x e y sono in R' , a R' appartengono anche gli elementi $-x$, $x + y$, xy . Si vede allora che, con le operazioni di addizione e moltiplicazione coincidenti con quelle definite in R , R' è un anello.

Esempio 16. Le funzioni reali differenziabili definite in R costituiscono un sottoanello dell'anello di tutte le funzioni continue su R .

Esempio 17. Sia V uno spazio vettoriale sul corpo K e sia A un operatore definito in V . L'insieme di tutti gli operatori definiti in V del tipo $f(A)$, dove f è un polinomio in $K[t]$, è un sottoanello dell'anello di tutte le applicazioni lineari di V in sé stesso.

Osservazioni sui corpi astratti. Un anello commutativo K tale che $1 \neq 0$ e tale che l'insieme dei suoi elementi non nulli costituisca un gruppo rispetto all'operazione di moltiplicazione, viene chiamato un *corpo astratto (commutativo)*. La seconda condizione può essere espressa anche dicendo che ogni elemento non nullo dell'anello ha un inverso rispetto alla moltiplicazione. La maggior parte dei teoremi dell'algebra lineare continuano a valere anche se si considerano spazi vettoriali su corpi astratti commutativi. Gli unici casi in cui abbiamo fatto uso di proprietà specifiche dei numeri reali o complessi sono i seguenti:

- 1) Studiando i prodotti scalari definiti positivi abbiamo fatto uso delle proprietà di ordinamento.
- 2) Abbiamo fatto uso della proprietà che ogni polinomio di grado positivo a coefficienti complessi ha almeno una radice nel corpo complesso. Tuttavia, ciò che abbiamo fatto in questo caso continuerebbe a valere considerando un corpo astratto avente questa stessa proprietà, chiamato allora *corpo algebricamente chiuso*.

3) Incidentalmente, trattando del criterio per le radici multiple che fa uso della derivazione, abbiamo fatto uso della proprietà che il corpo da noi considerato conteneva i numeri razionali come sottocorpo. Si è trattato, in effetti, di un punto di scarsa importanza in tutta la teoria.

Lo studente incline alla matematica più astratta potrebbe ora leggere il libro interpretando la parola corpo come corpo astratto e noterebbe soltanto le tre eccezioni qui sopra riportate. In realtà, considerando il livello a cui questo corso si svolge, tutto ciò ha importanza secondaria, perciò si è preferito evitare al lettore il disagio di questa ulteriore astrazione.

Esercizi

1. Dimostrare che in un corpo non esistono ideali bilateri che non siano lo zero oppure l'ideale unità.
2. In un anello R , può succedere che un prodotto xy sia nullo senza che lo sia almeno uno dei fattori. Dare un esempio di questo fatto nell'anello delle matrici e anche nell'anello delle funzioni continue definite nell'intervallo $[0, 1]$.
3. Sia R un anello commutativo. Se M è un ideale, si scriva M^2 invece di MM . Dimostrare che se M_1, M_2 sono ideali tali che $M_1 + M_2 = R$, allora anche $M_1^2 + M_2^2 = R$.
4. Sia R un anello e siano J_1, J_2 ideali sinistri. Dimostrare che anche $J_1 \cap J_2$ è un ideale sinistro.
5. Sia R l'anello delle matrici $n \times n$ su un corpo K . Dimostrare che l'insieme delle matrici del tipo

$$\begin{pmatrix} a_1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ a_n & 0 & \dots & 0 \end{pmatrix}$$

aventi cioè componenti nulle con l'eventuale eccezione di quelle nella prima colonna, è un ideale sinistro di R . Dimostrare la proposizione analoga relativa all'insieme delle matrici aventi componenti nulle con l'eventuale eccezione di quelle nella colonna j -esima.

6. Siano A, B matrici $n \times n$ su un corpo K aventi tutte le componenti nulle con l'eventuale eccezione di quelle nella prima colonna. Sia $A \neq O$. Dimostrare allora che esiste una matrice $n \times n$ C in K tale che $CA = B$.

Suggerimento: considerare dapprima il caso particolare in cui

$$A = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

7. Sia V uno spazio vettoriale di dimensione finita sul corpo K . Sia R l'anello delle applicazioni K -lineari di V in sé stesso. Dimostrare che R non ha ideali bilateri diversi da $\{O\}$ e da sé stesso. [Suggerimento: sia $A \in R$, $A \neq O$. Sia v_1 un elemento non nullo di V tale che $Av_1 \neq O$. Estendere $\{v_1\}$ a una base $\{v_1, \dots, v_n\}$ di V . Siano $\{w_1, \dots, w_n\}$ elementi qualsiasi di V . Per ogni $i = 1, \dots, n$ esiste $B_i \in R$ tale che

$$B_i v_i = v_i \quad \text{e} \quad B_i v_j = O \quad \text{se} \quad j \neq i$$

ed esiste inoltre $C_i \in R$ tale che $C_i A v_i = w_i$ (giustificare in dettaglio queste due proposizioni esistenziali). Sia $F = C_1 A B_1 + \dots + C_n A B_n$. Dimostrare allora che per ogni $i = 1, \dots, n$ $F(v_i) = w_i$. Concludere allora che l'ideale bilatero generato da A coincide con l'intero anello R .]

8. Sia R un anello comutativo. Un'applicazione $D: R \rightarrow R$ viene chiamata una *derivazione* se, per ogni x, y in R , $D(x+y) = Dx + Dy$ e $D(xy) = -(Dx)y + x(Dy)$. Se D_1, D_2 sono derivazioni, si definisca il prodotto

$$[D_1, D_2] = D_1 \circ D_2 - D_2 \circ D_1.$$

Dimostrare che $[D_1, D_2]$ è una derivazione.

Esempio. Sia R l'anello delle funzioni reali infinitamente differenziabili di due variabili reali, per esempio. Ogni operatore differenziale

$$f(x, y) \frac{\partial}{\partial x} \quad \text{oppure} \quad g(x, y) \frac{\partial}{\partial y}$$

con coefficienti f, g funzioni infinitamente differenziabili, è una derivazione su R . Ne segue che l'insieme di tutti i prodotti e dei prodotti iterati di questi operatori è un'algebra che però non è associativa: come si dimostra?

9. Sia R un anello in cui ogni elemento x è tale che $x^2 = x$. Dimostrare che R è comutativo.

73. OMOMORFISMI

Siano R, R' due anelli. Dicesi *omomorfismo di anelli* ogni applicazione $f: R \rightarrow R'$ avente le proprietà seguenti: per ogni scelta di x e y in R , si ha

$$f(x+y) = f(x) + f(y), \quad f(xy) = f(x)f(y), \quad f(e) = e'$$

(e cd e' sono rispettivamente gli elementi unità di R e R').

Dicesi *nucleo* dell'omomorfismo di anelli $f: R \rightarrow R'$ il nucleo di f considerato come omomorfismo tra i gruppi additivi, il nucleo cioè è l'insieme di tutti gli elementi x di R tali che $f(x) = 0$.
Esercizio: Dimostrare che il nucleo di f è un ideale bilatero di R .

Esempio 1. Sia R l'anello delle funzioni complesse definite nell'intervallo $[0, 1]$. L'applicazione che associa ad ogni funzione $f \in R$ il numero $f(\frac{1}{2})$ è un omomorfismo di anelli di R in \mathbb{C} .

Esempio 2. Sia R l'anello delle funzioni reali definite nell'intervallo $[0, 1]$. Sia R' l'anello delle funzioni reali definite nell'intervallo $[0, \frac{1}{2}]$. Ogni funzione f di R può essere considerata come una funzione definita in $[0, \frac{1}{2}]$, così facendo, si viene a considerare ciò che si chiama la *restrizione* di f all'intervallo $[0, \frac{1}{2}]$. Più in generale, siano S un insieme e S' un suo sottoinsieme. Sia R l'anello delle funzioni reali definite su S . Per ogni f di R , denotiamo con $f|S'$ la funzione definita su S' il cui valore per ogni $x \in S'$ è $f(x)$. Allora la funzione $f|S'$ è chiamata la *restruzione di f ad S'*. Sia R' l'anello delle funzioni reali definite in S' . Allora l'applicazione definita da

$$f \mapsto f|S'$$

è un omomorfismo di anelli di R in R' .

Esempio 3. Sia K un corpo e sia $R = K[t]$ l'anello dei polinomi a coefficienti in K . Sia A una matrice appartenente a $\text{Mat}_n(K)$, allora l'applicazione definita da

$$f \mapsto f(A)$$

è un omomorfismo di anelli di R in $\text{Mat}_n(K)$. Ciò non è altro che una riformulazione delle proprietà dell'applicazione definita da $f \mapsto f(A)$ già considerata in precedenza.

In ognuno degli esempi precedenti si descriva esplicitamente il nucleo dell'omomorfismo di anelli considerato.

Poiché il nucleo di un omomorfismo di anelli è definito soltanto mediante i gruppi additivi degli anelli considerati, ne deduciamo che un omomorfismo di anelli che abbia nucleo banale è iniettivo.

Sia $f: R \rightarrow R'$ un omomorfismo di anelli. Se esiste un omo-

morfismo di anelli $g: R' \rightarrow R$ tale che $g \circ f$ e $f \circ g$ sono le rispettive applicazioni identiche, diciamo che f è un *isomorfismo di anelli*. Come nel caso degli spazi vettoriali e dei gruppi, anche in questo caso è vero che un omomorfismo di anelli f che sia iniettivo e surgettivo è un isomorfismo di anelli. Ne lasciamo la dimostrazione per esercizio.

Vogliamo ora definire anche per gli anelli una nozione simile a quella di gruppo quoziante.

Sia R un anello e sia M un suo ideale bilatero. Se x e y sono elementi di R , si definisca x congruo a y mod M quando $x - y$ appartiene a M . Scriviamo questa relazione nella forma

$$x \equiv y \pmod{M}.$$

È allora molto facile dimostrare le proposizioni seguenti, cosa di cui lasciamo la cura al lettore.

- a) Si ha $x \equiv x \pmod{M}$.
- b) Se $x \equiv y$ e $y \equiv z \pmod{M}$, allora $x \equiv z \pmod{M}$.
- c) Se $x \equiv y$, allora $y \equiv x \pmod{M}$.
- d) Se $x \equiv y \pmod{M}$, e se $z \in R$, allora $xz \equiv yz \pmod{M}$ e anche $zx \equiv zy \pmod{M}$.
- e) Se $x \equiv y$ e $x' \equiv y' \pmod{M}$, allora $xx' \equiv yy' \pmod{M}$. Inoltre, $x + x' \equiv y + y' \pmod{M}$.

Le dimostrazioni delle affermazioni ora fatte sono tutte molto semplici. Per esempio diamo la dimostrazione di e). L'ipotesi significa che possiamo scrivere

$$x = y + z \quad \text{e} \quad x' = y' + z'$$

con z, z' in M . Allora

$$xx' = (y + z)(y' + z') = yy' + zy' + yz' + zz'.$$

E poiché M è un ideale bilatero, ognuno degli addendi zy' , yz' , zz' appartiene a M e allora anche la loro somma vi appartiene. Perciò $xx' \equiv yy' \pmod{M}$, come si doveva dimostrare.

Se $x \in R$, si indichi con \bar{x} l'insieme di tutti gli elementi di R congrui a $x \pmod{M}$. Ricordando la definizione di gruppo quoziante, vediamo che \bar{x} non è altro che la classe laterale additiva di x relativa a M , $x + M$. Ogni elemento di questa classe laterale (detta anche *classe di congruenza* di x mod M) viene detto un rappresentante della classe laterale stessa.

Sia \bar{R} l'insieme di tutte le classi di congruenza di R mod M . In altre parole, poniamo $\bar{R} = R/M$, gruppo quoziante additivo di R modulo M . Quindi sappiamo già che \bar{R} è un gruppo additivo. Vogliamo ora definire una moltiplicazione che farà di \bar{R} un anello.

Se \bar{x} e \bar{y} sono classi laterali additive relative a M , definiamo loro prodotto la classe laterale cui appartiene xy cioè la classe laterale \bar{xy} . Tenendo conto della condizione e) sopramenzionata, vediamo facilmente che questa classe laterale non dipende dalla scelta dei rappresentanti x in \bar{x} e y in \bar{y} . Quindi la nostra moltiplicazione risulta perfettamente definita dalla regola

$$(x + M)(y + M) = (xy + M).$$

La verifica degli assiomi della definizione di anello è immediata. In particolare, l'elemento unità di \bar{R} è $1 + M$, se 1 è l'elemento unità di R .

È altresì immediato verificare che l'applicazione

$$f: R \rightarrow \bar{R}$$

tale che $f(x) = \bar{x} = x + M$, è un omomorfismo di anelli di R su \bar{R} avente per nucleo M . Anche questa verifica viene lasciata al lettore.

Esercizi

1. Sia $f: R \rightarrow R'$ un omomorfismo di anelli. Dimostrare che l'immagine di f è un sottoanello di R' .
2. Sia V uno spazio vettoriale di dimensione finita sul corpo K . Sia \mathcal{B} una base di V . Dimostrare che l'applicazione definita da $f \mapsto M_{\mathcal{B}}^{\mathcal{B}}(f)$ è un isomorfismo dell'anello $\mathcal{L}(V, V)$ sull'anello delle matrici $n \times n$ (se n è la dimensione di V). (Questo esercizio non è altro che un'interpretazione nel linguaggio ora introdotto di proposizioni già viste a proposito di $M_{\mathcal{B}}^{\mathcal{B}}(f)$.)
3. Sia R un anello, sia G l'insieme degli elementi x di R tali che esiste un elemento $y \in R$ per cui $xy = yx = e$. Dimostrare che G è un gruppo, detto *gruppo delle unità* di R . (Se R è l'anello delle matrici $n \times n$ su un corpo K , allora G non è altro che il gruppo delle matrici invertibili.)
4. Dimostrare che un omomorfismo di anelli definito in un corpo K o è l'applicazione nulla o è un isomorfismo di K sulla sua immagine.
5. Consideriamo un esempio di classi di congruenza. Sia $R = \mathbb{Z}$ l'anello degli interi relativi. Sia M un ideale diverso da zero e dall'ideale unità. Allora

sappiamo che M ha un unico generatore positivo n e quindi $M = (n)$. Spesso scriveremo $\text{mod } n$ invece di $\text{mod}(n)$.

- Dimostrare che ogni intero x è congruo a un unico intero m tale che $0 < m < n$.
 - Dimostrare che ogni intero x non nullo e primo relativamente a n è congruo a un unico intero m primo relativamente a n e tale che $0 < m < n$.
 - Dimostrare che se x è un intero non nullo e primo relativamente a n , allora $x^{\varphi(n)} \equiv 1 \pmod{n}$, dove $\varphi(n)$ è il numero degli interi m primi relativamente a n e tali che $0 < m < n$.
 - Se p è un numero primo, che cosa è $\varphi(p)$?
 - Determinare $\varphi(n)$ per ogni intero n positivo e non superiore a 10.
6. a) Sia p un numero primo. Dimostrare che nell'anello $\mathbb{Z}/(p)$ ogni elemento non nullo ha un inverso rispetto alla moltiplicazione e che tutti gli elementi non nulli costituiscono un gruppo rispetto alla moltiplicazione.
b) Se a è un intero non congruo a zero \pmod{p} , dimostrare che

$$a^{p-1} \equiv 1 \pmod{p}.$$

7. Siano n, n' due interi positivi primi tra loro. Siano a, b due interi. Dimostrare che le congruenze

$$\begin{aligned} x &\equiv a \pmod{n}, \\ x &\equiv b \pmod{n'} \end{aligned}$$

possono essere simultaneamente risolte da qualche x scelto in \mathbb{Z} .

8. Enunciare gli analoghi degli esercizi 5a e 5b per l'anello dei polinomi su un corpo X .

9. Sia R un anello e siano M, M' suoi ideali bilateri. Si assuma che M contenga M' . Per ogni x in R si denoti con $x(M)$ la sua classe di congruenza \pmod{M} . Dimostrare che esiste un unico omomorfismo di anelli

$$R/M' \rightarrow R/M$$

che associa a $x(M')$ la classe $x(M)$.

10. Se n, m sono due interi non nulli e tali che n divide m , utilizzare l'esercizio 9 per ottenere un omomorfismo di anelli $\mathbb{Z}/(m) \rightarrow \mathbb{Z}/(n)$.

11. Siano R, R' due anelli. Sia $R \times R'$ l'insieme di tutte le coppie (x, x') con $x \in R$ e $x' \in R'$. Dimostrare che definendo addizione e moltiplicazione componente per componente, $R \times R'$ diviene un anello. In particolare, qual è l'elemento unità di $R \times R'$?

12. Siano m, n due interi positivi relativamente primi. Dimostrare che l'applicazione definita da

$$x \pmod{mn} \mapsto (x \pmod{n}, x \pmod{m})$$

è un isomorfismo di anelli di $\mathbb{Z}/(mn)$ su $\mathbb{Z}/(n) \times \mathbb{Z}/(m)$.

13. Dimostrare che se m e n sono interi positivi relativamente primi, allora

$$\varphi(mn) = \varphi(m)\varphi(n).$$

14. Sia $f: R \rightarrow R'$ un omomorfismo di anelli. Sia J' un ideale bilatero di R' e sia J l'insieme degli elementi x di R tali che $f(x)$ appartiene a J' . Dimostrare che J è un ideale bilatero di R .

15. Sia R un anello commutativo e sia N l'insieme degli elementi x di R tali che $x^n = 0$ per qualche intero positivo n . Dimostrare che N è un ideale.

16. Con riferimento all'esercizio 15, dimostrare che se \bar{x} è un elemento di R/N e se esiste un intero positivo n tale che $\bar{x}^n = 0$, allora $\bar{x} = 0$.

17. Sia R un anello e sia Z l'insieme degli elementi $x \in R$ tali che per ogni y di R , $xy = yx$. Dimostrare allora che Z è un sottoanello e che Z stesso è commutativo.

74. MODULI

Possiamo considerare una generalizzazione della nozione di spazio vettoriale su un corpo, cioè considerare un modulo su un anello: se R è un anello, un *modulo (sinistro) su R* , detto anche *R -modulo*, è un gruppo additivo M insieme a un'applicazione $R \times M \rightarrow M$ che ad ogni coppia (x, v) con $x \in R$ e $v \in M$ associa un elemento xv di M in modo che le seguenti condizioni siano soddisfatte.

MOD 1. Per ogni scelta di x, y in R e di v, w in M , abbiamo

$$(x + y)v = xv + yv, \quad x(v + w) = xv + xw.$$

MOD 2. Abbiamo anche $(xy)v = x(yv)$.

MOD 3. Se e è l'elemento unità di R , allora $ev = v$.

Esempio 1. Ogni ideale sinistro di R è un modulo.

Esempio 2. Sia M un modulo su un anello R . Allora M è anche un modulo su un qualsiasi sottoanello R' di R .

Esempio 3. Sia V uno spazio vettoriale sul corpo K e sia $R = \mathcal{L}(V, V)$ l'anello delle applicazioni K -lineari di V in sé stesso. Allora, definendo il prodotto di un elemento $A \in R$ per un vettore $v \in V$ come l'elemento $Av = A(v)$ di V , V stesso diviene un modulo su R . Le proprietà delle applicazioni lineari e la definizione di somma di applicazioni lineari servono a dimostrare che gli assiomi MOD 1, 2, 3 sono soddisfatti.

Esempio 4. Sia K un corpo. Allora K^n è un modulo sull'anello delle matrici $n \times n$ in K . Infatti se A è una matrice $n \times n$ in K e se X appartiene a K^n , allora l'ordinario prodotto tra matrici AX soddisfa le condizioni MOD 1, 2, 3.

Sia R un anello e siano M, M' due R -moduli. L'applicazione $f: M \rightarrow M'$ dicesi *R-lineare* (oppure un *R-omomorfismo*) se per ogni scelta di x in R e di v e w in M si ha

$$f(xv) = xf(v), \quad f(v+w) = f(v) + f(w).$$

Quindi la nozione di applicazione *R-lineare* è la generalizzazione di quella di applicazione *K-lineare* quando il modulo è uno spazio vettoriale su un corpo.

L'insieme di tutte le applicazioni *R-lineari* di M in M' viene denotato con $\mathcal{L}_R(M, M')$.

Esempio 5. Siano M, M', M'' tre R -moduli. Se

$$f: M \rightarrow M' \quad \text{e} \quad g: M' \rightarrow M''$$

sono applicazioni *R-lineari*, allora anche l'applicazione composta $g \circ f$ è *R-lineare*. Se

$$f_1, f_2 \in \mathcal{L}_R(M, M'),$$

allora

$$g \circ (f_1 + f_2) = g \circ f_1 + g \circ f_2.$$

Se g_1, g_2 sono in $\mathcal{L}_R(M', M'')$, allora

$$(g_1 + g_2) \circ f = g_1 \circ f + g_2 \circ f.$$

Provare per esercizio tutte queste uguaglianze.

Come nel caso degli spazi vettoriali, molto spesso occorre considerare l'insieme delle applicazioni *R-lineari* di un modulo M in sé stesso, ed è quindi utile dare un nome a queste applicazioni. Esse sono chiamate *R-endomorfismi* di M . L'insieme di tutti gli *R-endomorfismi* di M è denotato con $\text{End}_R(M)$.

Esempio 6. Sia R un anello e M un suo ideale sinistro. Sia y un elemento di M . L'applicazione

$$r_y: M \rightarrow M$$

definita in modo che

$$r_y(x) = xy$$

è un'applicazione R -lineare di M in sé stesso. Infatti, se x è in M allora xy appartiene a M giacché M è un ideale sinistro: le condizioni di R -linearità non sono che riformulazioni delle definizioni (quali?).

Chiamiamo r_y *moltiplicazione destra* per y . Quindi r_y è un R -endomorfismo di M .

Esempio 7. Sia V uno spazio vettoriale su un corpo K e sia A un'applicazione K -lineare di V in sé stesso. Sia R l'anello costituito da tutti gli elementi $f(A)$, f essendo un polinomio in $K[t]$. Allora V è un R -modulo. Se B è una qualsiasi applicazione K -lineare di V in sé stesso, B sarà un R -endomorfismo di V se, e soltanto se, $AB = BA$. In tal caso infatti, come abbiamo visto, per ogni polinomio f abbiamo

$$f(A)B = Bf(A).$$

Esempio 8. Siano R un anello e M un R -modulo. Sia $R' = \text{End}_R(M)$. Allora R' è un anello. Il fatto che gli assiomi che definiscono un anello sono soddisfatti segue dalle definizioni e dall'esempio 5. È importante notare che M può essere considerato anche un R' -modulo. Infatti, se $f \in R'$ e $v \in M$, alla coppia (f, v) possiamo associare l'elemento $f(v)$: è allora semplice verificare che tutte le condizioni che fanno di M un R' -modulo sono soddisfatte.

Vogliamo ora approfondire l'esame della situazione descritta nell'esempio 8. Ad ogni elemento x di R possiamo associare l'applicazione

$$\lambda_x: M \rightarrow M,$$

cioè l'applicazione definita in modo che

$$\lambda_x(v) = xv.$$

Allora per ogni scelta di v e w in M , abbiamo

$$\lambda_x(v + w) = x(v + w) = xv + xw = \lambda_x(v) + \lambda_x(w).$$

Inoltre, se $f \in R' = \text{End}_R(M)$, per definizione abbiamo

$$f(xv) = xf(v),$$

e quindi

$$f \circ \lambda_x(v) = \lambda_x \circ f(v).$$

Perciò λ_x è un'applicazione R' -lineare di M in sé stesso, cioè è un elemento di $\text{End}_{R'}(M)$. L'applicazione definita dall'associazione

$$\lambda: x \mapsto \lambda_x$$

è un omomorfismo di anelli di R in $\text{End}_{R'}(M)$, come è facile verificare. (Scrivere tutti i dettagli di questa verifica.)

TEOREMA 1 *Siano R un anello e M un R -modulo. Sia J l'insieme degli elementi x di R tali che, per ogni v di M , xv è nullo. Allora J è un ideale bilatero di R .*

Dimostrazione. Se x, y sono in J , allora, per ogni v in M , $(x + y)v = xv + yv = 0$. Se $a \in R$, allora

$$(ax)v = a(xv) = 0 \quad \text{e} \quad (xa)v = x(av) = 0$$

per ogni v in M . Si dimostra così il nostro teorema.

Osserviamo che l'ideale bilatero del teorema 1 non è altro che il nucleo dell'omomorfismo di anelli definito da

$$x \mapsto \lambda_x$$

e descritto nella dimostrazione precedente.

TEOREMA 2 (Wedderburn-Rieffel.) *Sia R un anello e sia L un ideale sinistro non costituito dal solo zero, considerato come un R -modulo. Siano $R' = \text{End}_R(L)$ e $R'' = \text{End}_{R'}(L)$. Sia poi*

$$\lambda: R \rightarrow R''$$

l'omomorfismo di anelli tale che, per ogni x di R e ogni y di L , $\lambda_x(y) = xy$. Si assume che in R non esistano ideali bilateri diversi da zero e da R stesso. Allora λ è un isomorfismo di anelli.

Dimostrazione. (Rieffel.) L'iniettività di λ è una conseguenza del teorema 1 e dell'ipotesi che L non sia costituito dal solo zero. L'unica cosa che rimane da provare quindi è la surgettività di λ . Dall'esempio 15 del paragrafo 72, sappiamo che LR è un ideale bilatero, non costituito dal solo zero perché in R vi è l'elemento unità: per l'ipotesi fatta questo ideale coincide con R . Allora

$$\lambda(LR) = \lambda(L)\lambda(R) = \lambda(R).$$

Vogliamo far vedere ora che $\lambda(L)$ è un ideale sinistro di R'' . Per dimostrare ciò, consideriamo un elemento f di R'' e un elemento x di L . Nell'esempio 6 abbiamo visto che per ogni y in L , r_y appartiene a R' e perciò abbiamo che

$$f \circ r_y = r_y \circ f.$$

Questo significa che $f(xy) = f(x)y$. Questa uguaglianza può essere riscritta nella forma

$$f \circ \lambda_x(y) = \lambda_{f(x)}(y).$$

Perciò $f \circ \lambda_x$ è un elemento di $\lambda(L)$, precisamente $\lambda_{f(x)}$. Questo dimostra che $\lambda(L)$ è un ideale sinistro di R'' . Ma allora

$$R''\lambda(R) = R''\lambda(L)\lambda(R) = \lambda(L)\lambda(R) = \lambda(R).$$

Poiché a $\lambda(R)$ appartiene l'applicazione identica, sia essa e , segue che per ogni f in R'' l'applicazione $f \circ e = f$ appartiene a $\lambda(R)$: segue cioè che R'' è contenuto in $\lambda(R)$ e quindi che $R'' = \lambda(R)$, come volevamo dimostrare.

Esercizi

1. Siano R un anello e M un R -modulo. Dimostrare che se v è in M , allora $0v=0$ (il primo zero è quello di R , il secondo è quello di M).
2. Sia V uno spazio vettoriale su un corpo K e sia R un sottoanello di $\text{End}_K(V)$ cui appartengono tutte le applicazioni scalari, cioè tutte le applicazioni cI con $c \in K$. Sia L un ideale sinistro di R e sia LV l'insieme di tutti gli elementi $A_1v_1 + \dots + A_nv_n$ con $A_i \in L$ e $v_i \in V$. Dimostrare che LV è un sottospazio R -invariante di V .
3. Sia R un'algebra sul corpo dei numeri complessi. Assumiamo che la moltiplicazione di R sia associativa e che in R esista l'elemento unità (e allora R è un anello) e infine che in R non vi siano ideali bilateri distinti da zero e da R stesso. Supponiamo inoltre che R abbia dimensione finita non nulla sul corpo complesso C . Sia L un ideale sinistro di R avente la minima dimensione positiva possibile su C .
 - a) Dimostrare che $\text{End}_R(L) = C$ (cioè le uniche applicazioni R -lineari di L sono le moltiplicazioni per numeri complessi). [Suggerimento: si tenga presente il lemma di Schur.]
 - b) Dimostrare che vi è un isomorfismo di anelli di R nell'anello delle applicazioni C -lineari di L in sé stesso.
4. Si definisca la nozione di sottomodulo di un modulo. Si definisca la nozione di isomorfismo fra moduli.

5. Sia R un anello. Un R -modulo viene detto semplice se M non consiste del solo zero e se i suoi unici sottomoduli sono soltanto $\{0\}$ e M stesso. Siano M, M' R -moduli semplici e sia $f: M \rightarrow M'$ un R -omomorfismo. Dimostrare che se f non è l'applicazione nulla, allora è un isomorfismo.

6. Siano R un anello, M un modulo, L un ideale sinistro. Dimostrare che LM è un sottomodulo di M . Si assuma ora che L e M siano entrambi semplici. Dimostrare allora che LM o consiste del solo zero o coincide con M .

75. MODULI QUOZIENTE

Abbiamo già studiato i gruppi quoziante e gli anelli modulo un ideale bilatero. Vogliamo ora introdurre l'analogia nozione a proposito dei moduli.

Siano R un anello e M un R -modulo. Diremo che N è un *sottomodulo* di M se N è un sottogruppo additivo di M tale che per ogni x in R e v in N , si abbia $xv \in N$. Quindi N stesso risulta un modulo (cioè R -modulo).

Già sappiamo come costruire il gruppo quoziante M/N . Osserviamo che M è un gruppo abeliano e quindi automaticamente N è normale in M . Si tratta quindi di una cosa che abbiamo già vista. Gli elementi del gruppo quoziante sono le classi laterali $v + N$ con v in M . Vogliamo ora definire una moltiplicazione di queste classi laterali per elementi di R . È una cosa che procede naturalmente: se x è in R , definiamo $x(v + N)$ come la classe laterale $xv + N$. Se v_1 è un altro rappresentante della classe laterale $v + N$, allora possiamo scrivere $v_1 = v + w$ per un opportuno w in N . Perciò

$$xv_1 = xv + xw,$$

e quindi xw appartiene a N . Ne segue che $xv_1 + N = xv + N$. La nostra definizione, quindi, non dipende dalla scelta del rappresentante v nella classe laterale $v + N$. È molto semplice verificare che per questa moltiplicazione sono soddisfatte tutte le condizioni della definizione di modulo.

Chiamiamo M/N il *modulo quoziante* di M per N oppure diciamo M *modulo* N .

Potremmo anche usare la notazione delle congruenze. Se v, v' sono elementi di M , possiamo scrivere

$$v \equiv v' \pmod{N}$$

per indicare che $v - v'$ è un elemento di N . Questo significa che le classi laterali $v + N$, $v' + N$ coincidono. Quindi una classe laterale $v + N$ è nient'altro che una classe di congruenza di elementi di M congrui a $v \text{ mod } N$. Possiamo riformulare l'affermazione che la moltiplicazione di una classe laterale per x è ben definita come segue: se $v \equiv v' \pmod{N}$, allora per ogni x in R abbiamo $xv \equiv xv' \pmod{N}$.

Esempio 1. Sia V uno spazio vettoriale sul corpo K . Sia W un suo sottospazio. Allora il modulo quoziante V/W è chiamato in questo caso *spazio quoziante*.

Siano M un R -modulo e N un suo sottomodulo. L'applicazione

$$f: M \rightarrow M/N$$

che ad ogni v di M associa la sua classe di congruenza $f(v) = v + N$ è evidentemente un R -omomorfismo giacché, per definizione, $f(xv) = xv + N = x(v + N)$. Esso viene chiamato l'*omomorfismo canonico*: il suo nucleo è N .

Esempio 2. Sia V uno spazio vettoriale sul corpo K . Sia W un sottospazio. Allora l'omomorfismo canonico $f: V \rightarrow V/W$ è una applicazione lineare, evidentemente surgettiva. Si supponga ora che V abbia dimensione finita su K e che W' sia un sottospazio di V tale che V stesso possa esprimersi come somma diretta $V = W \oplus W'$. Se $v \in V$ e se scriviamo $v = w + w'$ con w in W e w' in W' , allora $f(v) = f(w) + f(w') = f(w')$. Consideriamo ora l'applicazione f ristretta a W' e indichiamo questa nuova applicazione con f' . Allora, per ogni w' in W' , abbiamo per definizione $f'(w') = f(w')$. Allora f' applica W' su V/W e il nucleo di f' è $\{O\}$, poiché $W \cap W' = \{O\}$. Quindi $f': W' \rightarrow V/W$ è un isomorfismo tra il sottospazio supplementare W' di W e lo spazio quoziante V/W . Un tale isomorfismo esiste per ogni scelta del sottospazio supplementare W' .

Esercizi

1. Sia V uno spazio vettoriale di dimensione finita sul corpo K e sia W un suo sottospazio. Sia $\{v_1, \dots, v_r\}$ una base di W e la si estenda a una base $\{v_1, \dots, v_n\}$ di V . Sia $f: V \rightarrow V/W$ l'applicazione canonica. Dimostrare allora che

$$\{f(v_{r+1}), \dots, f(v_n)\}$$

è una base di V/W .

2. Siano V, W come nell'esercizio 1. Sia $A: V \rightarrow V$ un'applicazione lineare e si assuma che AW sia contenuto in W (cioè: $Aw \in W$ per ogni w in W). Sia $\{v_1, \dots, v_n\}$ la base di V definita nell'esercizio 1. Dimostrare che rispetto a questa base la matrice di A è del tipo

$$\begin{pmatrix} M_1 & M_2 \\ O & M_2 \end{pmatrix}$$

dove M_1 è una matrice quadrata $r \times r$, mentre M_2 è una matrice quadrata $(n-r) \times (n-r)$.

3. Con riferimento alle notazioni introdotte negli esercizi 1 e 2, dimostrare che mediante A è possibile definire un'applicazione lineare $\bar{A}: V/W \rightarrow V/W$ definendo

$$\bar{A}(v + W) = Av + W.$$

(Nella terminologia delle congruenze, se $v \equiv v' \pmod{W}$, allora $Av \equiv Av' \pmod{W}$.) Si scriva \bar{v} invece di $f(v)$. Dimostrare che la matrice di \bar{A} rispetto a $\{\bar{v}_{r+1}, \dots, \bar{v}_n\}$ è precisamente la matrice M_2 dell'esercizio 2. Chiamiamo \bar{A} l'applicazione lineare indotta da A sullo spazio quoziante.

4. Sia V lo spazio vettoriale generato su \mathbb{R} dalle funzioni $1, t, t^2, e^t, te^t, t^2e^t$. Sia W il sottospazio generato da $1, t, t^2, e^t, te^t$. Sia D l'operazione di derivazione.

a) Dimostrare che D applica W in sé stesso.

b) Qual è l'applicazione lineare \bar{D} indotta da D sullo spazio quoziante V/W ?

5. Sia V uno spazio vettoriale su \mathbb{R} costituito da tutti i polinomi di grado non superiore a n (per un certo intero n positivo). Sia W il sottospazio costituito da tutti i polinomi di grado non superiore a $n-1$. Qual è l'applicazione lineare \bar{D} indotta dall'operazione di derivazione D sullo spazio quoziante V/W ?

Appendice 1

Insiemi convessi

A. DEFINIZIONI

Sia S un sottoinsieme di \mathbb{R}^m . Diciamo che S è *convesso* se, comunque si prendano i punti P, Q in S , il segmento di retta congiungente P a Q è contenuto in S .

Ricordiamo che il segmento di retta congiungente P a Q è l'insieme di tutti i punti $P + t(Q - P)$ dove $0 < t < 1$. Quindi si tratta dell'insieme di punti

$$(1-t)P + tQ,$$

dove $0 < t < 1$.

TEOREMA 1 *Siano P_1, \dots, P_n punti di \mathbb{R}^m . Ad ogni insieme convesso cui appartengono i punti P_1, \dots, P_n appartengono anche tutte le combinazioni lineari*

$$x_1P_1 + \dots + x_nP_n,$$

tali che $0 < x_i < 1$ per ogni i , mentre $x_1 + \dots + x_n = 1$.

Dimostrazione. Questo è un bell'esercizio e non vogliamo togliere al lettore il piacere di risolverlo sviluppando tutti i dettagli. Ci limitiamo a dare un avvio: si proceda per induzione tenendo conto del fatto che se x_n è diverso da 1, allora la combinazione lineare scritta sopra è uguale a

$$(1-x_n) \left(\frac{x_1}{1-x_n} P_1 + \dots + \frac{x_{n-1}}{1-x_n} P_{n-1} \right) + x_n P_n.$$

TEOREMA 2 *Siano P_1, \dots, P_n punti dello spazio \mathbb{R}^m . L'insieme di tutte le combinazioni lineari*

$$x_1P_1 + \dots + x_nP_n$$

in cui $0 \leq x_i \leq 1$ e $x_1 + \dots + x_n = 1$ è un insieme convesso.

Dimostrazione. Si tratta di un esercizio molto semplice.

Per i teoremi 1 e 2 possiamo affermare che l'insieme delle combinazioni lineari descritte in questi teoremi è il più piccolo insieme convesso cui appartengono tutti i punti P_1, \dots, P_n .

Abbiamo già incontrato le proposizioni che seguono negli esercizi, le richiamiamo di nuovo per completezza.

1) *Se S e S' sono insiemi convessi, anche la loro intersezione $S \cap S'$ lo è.*

2) *Sia $F: \mathbb{R}^m \rightarrow \mathbb{R}^n$ un'applicazione lineare. Se S è un insieme convesso di \mathbb{R}^m , allora $F(S)$ (immagine di S attraverso F) è un insieme convesso di \mathbb{R}^n .*

3) *Sia $F: \mathbb{R}^m \rightarrow \mathbb{R}^n$ un'applicazione lineare. Sia S' un insieme convesso di \mathbb{R}^n . Sia $S = F^{-1}(S')$ l'insieme di tutti gli elementi X di \mathbb{R}^m per cui $F(X)$ appartiene a S' . Allora S è un insieme convesso.*

Esempi. Sia A un vettore dello spazio \mathbb{R}^n . L'applicazione F tale che $F(X) = A \cdot X$ è lineare. Osservato che un punto $c \in \mathbb{R}$ è un insieme convesso, l'iperpiano H costituito da tutti gli X tali che $A \cdot X = c$ è un insieme convesso.

Risulta anche convesso l'insieme S' costituito da tutti i numeri reali x maggiori di c . Perciò l'insieme di tutti i vettori X di \mathbb{R}^n tali che $A \cdot X > c$ è convesso. Questo insieme è chiamato *semispazio aperto*. Analogamente, l'insieme dei punti X di \mathbb{R}^n tali che $A \cdot X > c$ è chiamato un *semispazio chiuso*.

Nella figura A. 1 abbiamo illustrato un iperpiano (cioè una retta) di \mathbb{R}^2 insieme a uno dei semispazi che esso determina.

La retta è definita dall'equazione $3x - 2y = -1$, passa per il punto $P = (1, 2)$ ed è perpendicolare al vettore $N = (3, -2)$. È tratteggiata la parte di spazio corrispondente al semispazio costituito dai punti X tali che $X \cdot N \leq -1$.

Vediamo quindi che un iperpiano di equazione $X \cdot N = c$ determina due semispazi chiusi, cioè quelli definiti rispettivamente

dalle equazioni

$$X \cdot N > c \quad \text{e} \quad X \cdot N < c,$$

e analogamente per i semispazi aperti.

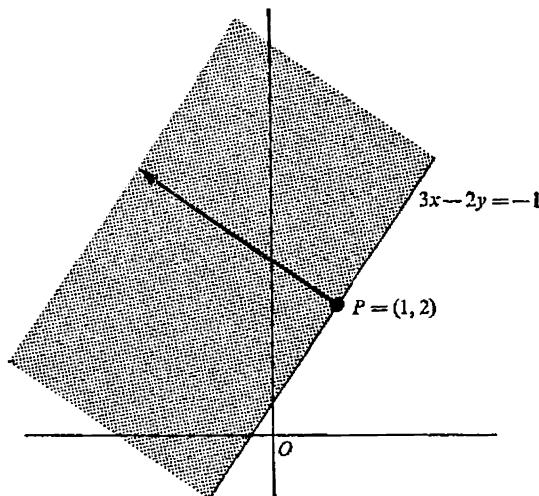


Figura A.1

Poiché l'intersezione di insiemi convessi è di nuovo un insieme convesso, l'intersezione di un numero finito di semispazi è un insieme convesso. Nelle figure A. 2 e A. 3 abbiamo tratteggiato intersezioni di un numero finito di semipiani. Queste intersezioni

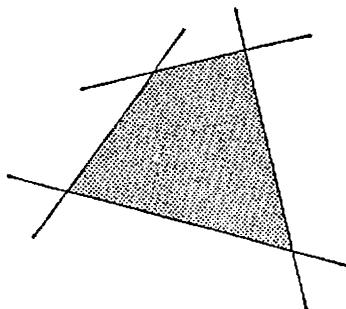


Figura A.2

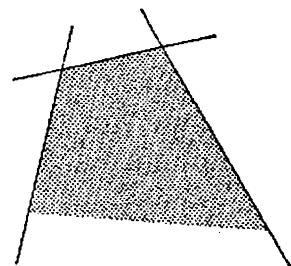


Figura A.3

possono risultare limitate o illimitate. (Ricordiamo che un sottoinsieme S di \mathbb{R}^n si dice *limitato* quando esiste un numero positivo c tale che $\|X\| \leq c$ per ogni X di S .)

B. IPERPIANI SEPARANTI

TEOREMA 3 *Sia S un insieme convesso e chiuso di \mathbb{R}^n . Sia P un punto di \mathbb{R}^n . Allora o P appartiene a S oppure esiste un iper piano H cui P appartiene e tale che l'insieme S sia contenuto in uno dei semispazi aperti determinati da H .*

Dimostrazione. Faremo uso di una proposizione tratta dal calcolo infinitesimale. Supponiamo che P non appartenga a S . Consideriamo la funzione definita sull'insieme chiuso S in modo che

$$f(X) = \|X - P\|.$$

Si dimostra nei corsi di calcolo infinitesimale (con gli ε e i δ) che questa funzione ha un minimo su S . Sia Q un punto di S tale che

$$\|Q - P\| < \|X - P\|,$$

per ogni X in S . Poniamo

$$N = Q - P.$$

Poiché P non appartiene a S , $Q - P$, e quindi N , non è il vettore nullo. Vogliamo far vedere che l'iperpiano passante per P e perpendicolare a N è quello che cerchiamo. Sia Q' un qualsiasi punto di S e sia Q' distinto da Q . Allora per ogni numero reale positivo t che non supera 1, abbiamo

$$\|Q - P\| < \|Q + t(Q' - Q) - P\| = \|(Q - P) + t(Q' - Q)\|.$$

Elevando al quadrato, otteniamo

$$(Q - P)^2 < (Q - P)^2 + 2t(Q - P) \cdot (Q' - Q) + t^2(Q' - Q)^2.$$

Riducendo i termini simili e dividendo per $2t$ otteniamo ancora

$$0 < 2(Q - P) \cdot (Q' - Q) + t(Q' - Q)^2.$$

Facendo ora tendere t a zero, abbiamo

$$\begin{aligned} 0 &\leq (Q - P) \cdot (Q' - Q) \leq \\ &\leq N \cdot (Q' - P) + N \cdot (P - Q) \leq \\ &\leq N \cdot (Q' - P) - N \cdot N. \end{aligned}$$

Ma $N \cdot N$ è un numero positivo e perciò

$$Q' \cdot N > P \cdot N.$$

Questo dimostra che S è contenuto nel semispazio aperto definito dalla condizione $X \cdot N > P \cdot N$.

Sia S un insieme convesso contenuto in \mathbb{R}^n . Allora la chiusura di S (denotata con \bar{S}) è un insieme convesso.

La dimostrazione è semplice: se P, Q sono punti della chiusura, possiamo trovare in S due successioni di punti $\{P_k\}, \{Q_k\}$ che, rispettivamente, tendono al limite P, Q . Allora, per ogni numero t tale che $0 < t < 1$, il punto

$$tP_k + (1-t)Q_k$$

tende a $tP + (1-t)Q$, che perciò appartiene alla chiusura di S .

Sia S un insieme convesso contenuto in \mathbb{R}^n . Sia P un punto di frontiera di S (un punto tale che per ogni ϵ positivo la sfera aperta di centro P e raggio ϵ di \mathbb{R}^n contenga dei punti che sono in S e dei punti che non sono in S). Un iperpiano H viene detto un *iperpiano radente S in P* se P appartiene a H e se S è contenuto in uno dei due semispazi chiusi determinati da H .

TEOREMA 4 *Sia S un insieme convesso contenuto in \mathbb{R}^n e sia P un punto frontiera di S . Esiste allora un iperpiano radente S in P .*

Dimostrazione. Sia \bar{S} la chiusura di S . Abbiamo già visto che anche \bar{S} è un insieme convesso e che P è un punto frontiera anche di \bar{S} . Se possiamo dimostrare la nostra affermazione per \bar{S} , essa varrà certamente anche per S . Quindi, senza ledere la generalità, possiamo ritenere che l'insieme di S sia chiuso.

Per ogni intero $k > 2$, possiamo trovare un punto P_k non appartenente a S e avente da P distanza minore di $1/k$. Dimostrando il teorema 3 abbiamo visto che esiste un punto Q_k di S avente distanza minima da P_k : poniamo $N_k = Q_k - P_k$. Sia N'_k il vettore

tore di lunghezza 1 avente direzione uguale a quella di N_k . La successione di vettori $\{N'_k\}$ ha un punto di accumulazione sulla sfera di raggio 1, sia esso N' , perché la sfera è un insieme compatto. Per il teorema 3, abbiamo allora che, per ogni X di S ,

$$X \cdot N_k > P_k \cdot N_k$$

per ogni k , e quindi, dividendo entrambi i membri per la lunghezza di N_k , otteniamo

$$X \cdot N'_k > P_k \cdot N'_k$$

per ogni k . Poiché N' è un punto di accumulazione di $\{N'_k\}$ e poiché P è un punto limite di $\{P_k\}$, per continuità segue che per ogni X in S ,

$$X \cdot N' > P \cdot N',$$

e questo dimostra il nostro teorema.

Osservazione. Sia S un insieme convesso e sia H un iperpiano definito dall'equazione

$$X \cdot N = a.$$

Supponiamo che, per ogni X in S , si abbia $X \cdot N > a$. Se P è un punto di S appartenente all'iperpiano, allora P deve essere un punto frontiera di S . Se così non fosse, per ε positivo e sufficientemente piccolo, $P - \varepsilon N$ apparterrebbe a S e si avrebbe allora

$$(P - \varepsilon N) \cdot N = P \cdot N - \varepsilon N \cdot N = a - \varepsilon N \cdot N < a,$$

contro l'ipotesi. Concludiamo perciò che H è un iperpiano radente S in P .

C. PUNTI ESTREMI E IPERPIANI RADENTI

Sia S un insieme convesso e sia P un suo punto. Diremo che P è un *punto estremo* di S se non è possibile trovare in S due punti distinti Q_1, Q_2 in modo che P possa venire scritto nella forma

$$P = tQ_1 + (1-t)Q_2 \quad \text{con} \quad 0 < t < 1.$$

In altre parole, P non può appartenere a un segmento contenuto in S se non nel caso in cui è un estremo del segmento stesso.

Definiamo un insieme S *inferiormente limitato* quando esiste un vettore $B = (b_1, \dots, b_n)$ tale che, per ogni $X = (x_1, \dots, x_n)$ in S , si abbia $x_i \geq b_i$ per $i = 1, \dots, n$.

TEOREMA 5 *Sia S un insieme convesso, chiuso, inferiormente limitato. Allora ogni iperpiano radente S ne contiene un punto estremo.*

Dimostrazione. Sia H un iperpiano radente S in un punto di frontiera P_0 , definito dall'equazione $X \cdot N = P_0 \cdot N$ e sia, per esempio, $X \cdot N > P_0 \cdot N$ per ogni punto X di S . Sia T l'intersezione di S con questo iperpiano. Allora T risulta convesso, chiuso, inferiormente limitato. Vogliamo far vedere che un punto estremo di T è anche un punto estremo di S . In tal modo il nostro problema sarà ridotto a trovare punti estremi di T . Per dimostrare la nostra asserzione, sia P un punto estremo di T e supponiamo di poter scrivere

$$P = tQ_1 + (1-t)Q_2, \quad 0 < t < 1.$$

Moltiplicando scalarmente per N e tenendo conto del fatto che P appartiene all'iperpiano, e che quindi $P \cdot N = P_0 \cdot N$, otteniamo

$$P_0 \cdot N = tQ_1 \cdot N + (1-t)Q_2 \cdot N. \quad [1]$$

Poiché Q_1 e Q_2 appartengono a S , tanto $Q_1 \cdot N$ quanto $Q_2 \cdot N$ non sono minori di $P_0 \cdot N$. Se uno di questi prodotti supera $P_0 \cdot N$, per esempio se $Q_1 \cdot N > P_0 \cdot N$, il secondo membro dell'equazione [1] risulta maggiore di

$$tP_0 \cdot N + (1-t)P_0 \cdot N = P_0 \cdot N,$$

cosa impossibile. Ne segue che Q_1 e Q_2 debbono entrambi appartenere all'iperpiano, contraddicendo così l'ipotesi che P sia un punto estremo di T .

Troviamo ora un punto estremo di T . Tra tutti i punti di T ne esiste almeno uno la cui prima coordinata sia la minima possibile, giacché T è un insieme chiuso e inferiormente limitato. (Proiettiamo sul primo spazio coordinato. L'immagine di T attraverso questa proiezione ha un estremo inferiore che appartiene a T , essendo T stesso chiuso.) Sia T_1 il sottoinsieme di T costituito da tutti i punti la cui prima coordinata è quella minima. Allora anche T_1 è un insieme chiuso e inferiormente limitato. Possiamo quindi

trovare in T_1 un punto la cui seconda coordinata sia la più piccola tra quelle degli altri punti di T_1 ; l'insieme T_2 di tutti i punti di T_1 aventi questa seconda coordinata è chiuso e inferiormente limitato. Procediamo in questo modo fino a trovare un punto P di T avente successivamente prima, seconda, ..., n -esima coordinata minima. Vogliamo far vedere che P è un punto estremo di T . Sia $P = (p_1, \dots, p_n)$.

Supponiamo di poter scrivere

$$P = tX + (1-t)Y, \quad 0 < t < 1$$

dove i punti $X = (x_1, \dots, x_n)$ e $Y = (y_1, \dots, y_n)$ appartengono a T . Allora x_1 e y_1 non sono minori di p_1 e inoltre

$$p_1 = tx_1 + (1-t)y_1.$$

Se uno dei numeri x_1 , y_1 supera p_1 , allora

$$tx_1 + (1-t)y_1 > tp_1 + (1-t)p_1 = p_1,$$

e questo è impossibile. Perciò $x_1 = y_1 = p_1$. Procedendo per induzione, supponendo di avere dimostrato che $x_i = y_i = p_i$ se $i = 1, \dots, r$, allora, se r è minore di n ,

$$p_{r+1} = tx_{r+1} + (1-t)y_{r+1},$$

e possiamo ripetere la dimostrazione precedente. Ne segue che

$$X = Y = P,$$

e quindi P è un punto estremo e il nostro teorema è dimostrato.

D. TEOREMA DI KREIN-MILMAN

Sia E un insieme non vuoto di punti appartenenti a \mathbb{R}^n . Vogliamo caratterizzare il più piccolo insieme convesso che contiene E . Possiamo senz'altro dire che si tratta dell'intersezione di tutti gli insiemi convessi contenenti E , infatti questa intersezione è un insieme convesso ed è chiaramente il più piccolo.

Possiamo però descrivere questo più piccolo insieme convesso in un altro modo. Sia E^c l'insieme di tutte le combinazioni lineari

$$t_1 P_1 + \dots + t_m P_m$$

di punti P_1, \dots, P_m di E a coefficienti reali t_i tali che

$$0 < t_i < 1 \quad \text{e} \quad t_1 + \dots + t_m = 1.$$

Allora l'insieme E^c è convesso: ne lasciamo la verifica, molto semplice, al lettore. Ogni insieme convesso contenente E deve contenere E^c e quindi E^c è il più piccolo insieme convesso contenente E . Chiamiamo E^c la *chiusura convessa* di E .

Sia S un insieme convesso e sia E l'insieme dei suoi punti estremi. Allora E^c è contenuto in S . Ci chiediamo sotto quali condizioni E^c coincide con S .

In linguaggio geometrico, punti estremi possono essere tanto quelli sul guscio di un uovo quanto quelli che sono nei vertici di un poligono, ad esempio si vedano le figure D. 1 e D. 2.

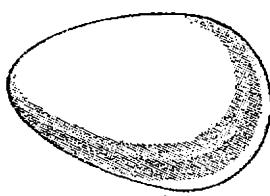


Figura D.1

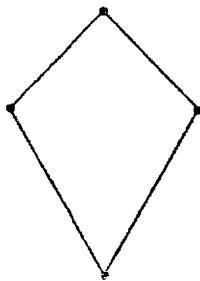


Figura D.2

Un insieme convesso non limitato non è necessariamente la chiusura convessa dell'insieme dei suoi punti estremi, si pensi per esempio al semipiano superiore chiuso. Gli unici punti estremi sono quelli dell'asse orizzontale. Analogamente, un insieme convesso aperto non è necessariamente la chiusura convessa dell'insieme dei suoi punti estremi (l'interno di un uovo non possiede punti estremi). Il teorema di Krein-Milman stabilisce che, eliminate queste due possibilità, non possono sorgere altri casi.

TEOREMA 6 *Sia S un insieme convesso, limitato, chiuso. Allora S è la chiusura convessa dell'insieme dei suoi punti estremi.*

Dimostrazione. Sia S' la chiusura convessa dell'insieme dei punti estremi di S . Basta far vedere che S è contenuto in S' . Sia $P \in S$ e si supponga che P non appartenga a S' . Il teorema 3 assicura l'esistenza di un iperpiano H passante per P , definito

da un'equazione

$$X \cdot N = c,$$

in modo che, per ogni X in S' , $X \cdot N$ sia maggiore di c . Sia $L: \mathbb{R}^n \rightarrow \mathbb{R}$ l'applicazione lineare tale che $L(X) = X \cdot N$. Allora $L(P) = c$ e quindi $L(P)$ non appartiene a $L(S')$. Poiché S è chiuso e limitato, anche la sua immagine $L(S)$ è chiusa e limitata, questa immagine è inoltre convessa. Perciò $L(S)$ è un intervallo chiuso, per esempio $[a, b]$, cui c appartiene. Perciò $a < c < b$. Sia H_a l'iperpiano definito dall'equazione

$$X \cdot N = a.$$

L'osservazione fatta dopo il teorema 4 dice che H_a è un iperpiano radente S . Il teorema 5 permette poi di concludere che a H_a appartiene un punto estremo di S . Questo punto estremo deve appartenere a S' . Otteniamo allora una contraddizione del fatto che $X \cdot N > c > a$ per ogni X in S' . Così il teorema di Krein-Milman è dimostrato.

Esercizi

1. Sia A un vettore di \mathbb{R}^n . Sia $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ la traslazione

$$F(X) = X + A.$$

Dimostrare che se S è un insieme convesso di \mathbb{R}^n , allora anche $F(S)$ lo è.

2. Sia c un numero positivo e sia P un punto di \mathbb{R}^n . Sia S l'insieme dei punti X tali che $\|X - P\| < c$. Dimostrare che S è un insieme convesso. Analogamente, dimostrare che l'insieme dei punti X tali che $\|X - P\| \leq c$ è convesso.

3. Disegnare la chiusura convessa dei seguenti insiemi di punti

- a) $(1, 2), (1, -1), (1, 3), (-1, 1)$.
- b) $(-1, 2), (2, 3), (-1, -1), (1, 0)$.

4. Sia $L: \mathbb{R}^n \rightarrow \mathbb{R}^n$ un'applicazione lineare invertibile. Sia S un insieme convesso di \mathbb{R}^n e sia P un punto estremo di S . Dimostrare che $L(P)$ è un punto estremo di $L(S)$. Questa affermazione continua a essere vera se l'applicazione L non è invertibile?

5. Dimostrare che l'intersezione di un numero finito di semispazi chiusi di \mathbb{R}^n non può avere che un numero finito di punti estremi.

6. Sia B un vettore colonna di \mathbb{R}^n e sia A una matrice $n \times n$. Dimostrare che l'insieme delle soluzioni dell'equazione lineare $AX = B$ è un insieme convesso di \mathbb{R}^n .

Appendice 2

Complementi

In questa appendice richiamiamo dapprima il procedimento di induzione. Ne abbiamo fatto uso più volte in tutto il libro e la maggior parte degli studenti lo conosce già: può tuttavia essere utile per qualcuno richiamare la sua esatta formulazione. Passeremo poi a dimostrare che il corpo dei numeri complessi è algebricamente chiuso, facendo uso soltanto di un fatto elementare dell'analisi. Infine accenneremo alla nozione di relazione di equivalenza.

E. INDUZIONE

Per dimostrare le proposizioni passo per passo, spesso abbiamo proceduto per induzione. Vogliamo ora enunciare precisamente la proprietà dei numeri interi chiamata induzione.

Supponiamo che per ogni intero positivo n sia assegnata un'affermazione $A(n)$. Vogliamo dimostrare che tutte le asserezioni $A(n)$, per $n = 1, 2, \dots$, sono vere. Supponiamo di poter dimostrare i due seguenti fatti:

- 1) L'asserzione $A(1)$ è vera.
- 2) Per ogni intero positivo n , se supponiamo vera l'asserzione $A(n)$, allora anche l'asserzione $A(n + 1)$ risulta vera.

Allora, la *proprietà di induzione* afferma che tutte le asserezioni $A(n)$ sono vere per ogni intero positivo n .

Esempio 1. Vogliamo dimostrare che per ogni intero positivo n , vale l'uguaglianza $A(n)$:

$$1 + 2 + \dots + n = \frac{n(n + 1)}{2}.$$

Essa certamente vale se n è uguale a 1, essendo $1 = 1(1+1)/2$. Supponiamo ora che la nostra uguaglianza sia vera per un intero positivo n . Allora

$$\begin{aligned} 1 + \dots + n + (n+1) &= \frac{n(n+1)}{2} + (n+1) = \frac{n(n+1) + 2(n+1)}{2} = \\ &= \frac{n^2 + n + 2n + 2}{2} = \frac{(n+1)(n+2)}{2}. \end{aligned}$$

Abbiamo quindi dimostrato che valgono le due proprietà 1), 2) relative alle proposizioni denotate con $A(n)$ e possiamo concludere per induzione che $A(n)$ è vera per ogni intero positivo n .

Mettiamo l'accento sul fatto che la proprietà di induzione è un assioma che riguarda i numeri interi. Noi lo consideriamo come una proprietà dei numeri interi: comunque si tratta di una proprietà molto plausibile.

F. CHIUSURA ALGEBRICA DEL CORPO DEI NUMERI COMPLESSI

Adoperando alcuni risultati elementari di analisi, vogliamo dimostrare che il corpo dei numeri complessi è algebricamente chiuso, in altre parole, che *ogni polinomio $f \in \mathbb{C}[t]$ di grado positivo ha una radice in \mathbb{C}* .

Possiamo scrivere

$$f(t) = a_n t^n + a_{n-1} t^{n-1} + \dots + a_0$$

con a_n non nullo. Per ogni numero reale positivo R , la funzione $|f|$ definita da

$$t \mapsto |f(t)|$$

è continua sul disco chiuso di raggio R e quindi su questo disco ammette un minimo. D'altra parte, dall'espressione

$$f(t) = a_n t^n \left(1 + \frac{a_{n-1}}{a_n t} + \dots + \frac{a_0}{a_n t^n} \right)$$

vediamo che quando $|t|$ diviene molto grande, anche $|f(t)|$ diviene molto grande, cioè, per ogni C positivo esiste un numero positivo R tale che da $|t| > R$ segue $|f(t)| > C$. Conseguentemente, esiste un numero positivo R_0 tale che, se z_0 è un punto di minimo

di $|f|$ sul disco chiuso di raggio R_0 , allora

$$|f(t)| \geq |f(z_0)|$$

per ogni numero complesso t . In altre parole, z_0 è un punto di minimo assoluto per $|f|$. Vogliamo dimostrare che $f(z_0) = 0$.

Scriviamo f nella forma

$$f(t) = c_0 + c_1(t - z_0) + \dots + c_n(t - z_0)^n$$

dove c_i sono delle costanti. (L'abbiamo già fatto in precedenza, ma si può anche vedere scrivendo $t = z_0 + (t - z_0)$ e andando a sostituire direttamente in $f(t)$.) Se $f(z_0)$ non è nullo, allora neppure $c_0 = f(z_0)$ è nullo. Posto $z = t - z_0$, sia m il più piccolo intero positivo tale che $c_m \neq 0$. Un tale valore m esiste perché abbiamo supposto f di grado positivo. Possiamo allora scrivere

$$f(t) = f_1(z) = c_0 + c_m z^m + z^{m+1} g(z)$$

per opportuni polinomi g e f_1 (quest'ultimo ottenuto da f con un cambiamento di variabile). Sia z_1 un numero complesso tale che $z_1^m = -c_0/c_m$, e consideriamo i valori di z del tipo

$$z = \lambda z_1$$

dove λ è reale e $0 < \lambda < 1$. Abbiamo allora

$$\begin{aligned} f(t) = f_1(\lambda z_1) &= c_0 - \lambda^m c_0 + \lambda^{m+1} z_1^{m+1} g(\lambda z_1) = \\ &= c_0 [1 - \lambda^m + \lambda^{m+1} z_1^{m+1} c_0^{-1} g(\lambda z_1)]. \end{aligned}$$

Esiste un numero positivo C tale che per ogni λ per cui $0 < \lambda < 1$ abbiamo $|z_1^{m+1} c_0^{-1} g(\lambda z_1)| \leq C$, e quindi

$$|f_1(\lambda z_1)| \leq |c_0|(1 - \lambda^m + C\lambda^{m+1}).$$

Ora, se noi possiamo dimostrare che per valori di λ sufficientemente piccoli e tali che $0 < \lambda < 1$ si ha

$$0 < 1 - \lambda^m + C\lambda^{m+1} < 1,$$

allora per ogni tale λ abbiamo anche $|f_1(\lambda z_1)| < |c_0|$, contraddicendo così l'ipotesi che $|f(z_0)|$ non supera $|f(t)|$ per nessun numero complesso t . La diseguaglianza sinistra è immediata giacché $0 < \lambda < 1$. La diseguaglianza a destra equivale a $C\lambda^{m+1} < \lambda^m$

o, cioè che è lo stesso, a $C\lambda < 1$, certamente soddisfatta per valori di λ opportunamente piccoli. Questo conclude la nostra dimostrazione.

Esercizio

- Partendo dal risultato appena dimostrato a proposito dei numeri complessi, dimostrare che ogni polinomio irriducibile a coefficienti reali ha grado 1 o 2. [Suggerimento: si scomponga il polinomio nel corpo complesso e si associno le coppie di fattori contenenti radici complesse coniugate.]

G. RELAZIONI DI EQUIVALENZA

Sia S un insieme. Dicesi *relazione di equivalenza* in S una relazione, denotata con $x \sim y$, tra certe coppie di elementi di S soddisfacente le seguenti condizioni:

RE 1. *Per ogni x di S , abbiamo $x \sim x$.*

RE 2. *Se $x \sim y$ e $y \sim z$, allora $x \sim z$.*

RE 3. *Se $x \sim y$, allora $y \sim x$.*

Supponiamo di avere una relazione di equivalenza in S . Per ogni elemento x di S , si indicherà con C_x l'insieme di tutti gli elementi di S che sono equivalenti a x . Allora tutti gli elementi di C_x sono equivalenti fra loro, come segue immediatamente dalle nostre tre proprietà. (Eseguire la verifica in dettaglio.) Inoltre si verifica immediatamente che se x e y sono elementi di S , allora o $C_x = C_y$ oppure C_x e C_y non hanno elementi in comune. Ogni C_x è chiamato una *classe di equivalenza*. Si vede quindi che la nostra relazione di equivalenza determina una decomposizione di S in classi di equivalenza a due a due disgiunte. Ogni elemento di una classe viene chiamato un *rappresentante* di quella classe.

Esempio 1. Siano G un gruppo e H un suo sottogruppo. Se x, y sono elementi di G , definiamo $x \sim y$ quando $xH = yH$. Si vede subito che si tratta di una relazione di equivalenza. Le classi di equivalenza sono state chiamate classi laterali rispetto a H (più precisamente classi laterali sinistre).

Esempio 2. Siano R un anello e M un suo ideale bilatero. Allora la congruenza modulo M è una relazione di equivalenza tra elementi di R .

Esempio 3. Sia S l'insieme di tutte le figure geometriche del piano \mathbb{R}^2 (cioè l'insieme di tutti i triangoli, quadrati, rettangoli ecc.). Sia G il gruppo generato da tutte le traslazioni e da tutte le applicazioni reali unitarie di \mathbb{R}^2 . Noi chiamiamo G il gruppo dei movimenti rigidi del piano. (Vedi esercizio 3 del cap. 14, § 68). Se α, β sono elementi di S , definiamo α equivalente a β e scriviamo $\alpha \sim \beta$, se esiste in G un elemento T tale che $T(\alpha) = \beta$. *Dimostrare che questa è una relazione di equivalenza.* Si tratta della relazione di equivalenza della geometria piana. Quando due figure del piano sono dette nella scuola elementare "uguali", si tratta di un grosso abuso di linguaggio. Per esempio, i due triangoli della figura G. 1 decisamente non sono uguali, cioè essi non sono lo stesso triangolo. Essi però sono equivalenti secondo la nostra definizione. Nella geometria piana quindi la regola di sostituzione di cose uguali a cose uguali è in effetti una regola che riguarda elementi equivalenti e non dice niente di più della proprietà RE 2.

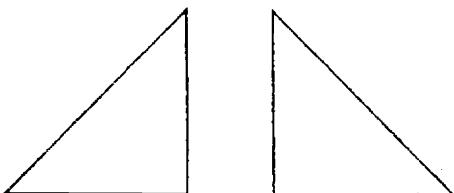


Figura G. 1

In matematica, la parola *uguale* significa *la stessa cosa*. Per esempio quando noi scriviamo

$$1 + 2 = 4 - 1$$

abbiamo lo stesso numero in ognuno dei due membri di questa uguaglianza, cioè il numero 3. Naturalmente il numero 3 è rappresentato in due modi diversi, ma il valore di $1 + 2$ è lo stesso del valore di $4 - 1$, cioè 3.

Naturalmente, l'uguaglianza è un caso particolare di relazione di equivalenza, quando in ogni classe di equivalenza vi è precisamente un elemento.

Esempio 4. Nel testo abbiamo menzionato le funzioni razionali, cioè i "quoienti" di polinomi, f/g . Se c è un numero tale che

$g(c) = 0$, allora un tale quoziente non è definito in c . Quindi una funzione razionale non può essere considerata come una funzione definita per tutti i numeri e vogliamo quindi affrontare il problema di definire rigorosamente le funzioni razionali. Si tratta in effetti di una cosa semplice. Sia K un corpo. Siano (f, g) e (f_1, g_1) coppie ordinate di polinomi in K tali che né g né g_1 siano il polinomio nullo. Noi diciamo che queste due coppie sono equivalenti quando $fg_1 = f_1g$. Si provi per esercizio che questa è una relazione di equivalenza. La classe di equivalenza cui appartiene la coppia (f, g) è denotata con f/g ed è chiamata una funzione razionale. Osserviamo che la nostra definizione di equivalenza è stata data in modo da far valere la regola del "prodotto in croce". Possiamo ora definire addizione e moltiplicazione di funzioni razionali. Se f/g e f_1/g_1 sono funzioni razionali, definiamo

$$\begin{aligned} f/g + f_1/g_1 &= (fg_1 + gf_1)/gg_1, \\ (f/g)(f_1/g_1) &= (ff_1)/(gg_1). \end{aligned}$$

È di nuovo un semplice esercizio dimostrare che questa somma e questo prodotto non dipendono dalla scelta dei rappresentanti (f, g) e (f_1, g_1) rispettivamente nelle funzioni razionali f/g e f_1/g_1 . Le funzioni razionali, quindi, costituiscono un anello e, di nuovo, si provi in tutti i dettagli per esercizio che la nostra addizione e moltiplicazione soddisfano tutti gli assiomi nella definizione di anello.

Osserviamo che, secondo la nostra definizione, se h è un polinomio non nullo, allora $fh/gh = f/g$. Infatti, questo significa semplicemente che, $fhg = ghf$, cosa certamente vera. Quindi la ordinaria regola di cancellazione è valida.

Esercizi

1. Sia V uno spazio vettoriale di dimensione finita sul corpo K e sia $\varrho: G \rightarrow GL(V)$ un omomorfismo di gruppi di un gruppo G nel gruppo delle applicazioni lineari invertibili di V . Se v, w sono elementi di V , definiamo $v \sim w$ se in G esiste un elemento σ tale che $\varrho(\sigma)v = w$. Dimostrare che questa è una relazione di equivalenza.

2. Sia V uno spazio vettoriale di dimensione finita sul corpo K . Se v, w sono elementi di V , definiamo $v \sim w$ se esiste un'applicazione lineare invertibile $A: V \rightarrow V$ tale che $Av = w$. Dimostrare che per questa relazione esistono soltanto due classi di equivalenza: una costituita dal solo 0 , l'altra costituita da tutti gli elementi non nulli di V .

3. Sia G un gruppo. Definiamo due elementi a, b di G *coniugati* e scriviamo $a \sim b$, se in G esiste un elemento x in modo che $xax^{-1} = b$. Dimostrare che questa è una relazione di equivalenza.

4. Sia G il gruppo simmetrico su tre elementi. Usando la relazione di equivalenza definita nell'esercizio precedente, determinare le classi di equivalenza in G . Dimostrare che due permutazioni appartenenti alla stessa classe di equivalenza sono dello stesso segno.

Indice analitico

- Abeliano, 306
- Aggiunto, 198
- Algebra, 333
- Anello, 332
- Angolo tra vettori, 25
- Antilinearità, 169
- Applicazione, 76
 - bilineare, 182, 280
 - hermitiana, 199
 - lineare, 82, 344
 - lineare simmetrica, 193
 - multilineare, 293
 - unitaria, 202
- Autoaggiunto, 199
- Automorfismo, 313
 - interno, 314
- Autospazio, 221, 250, 268
- Autovalore, 220
- Autovettore, 220
- Base:
 - a ventaglio, 232
 - di un gruppo abeliano, 327
 - di uno spazio vettoriale, 44
 - duale, 170
 - ortogonale, 151
 - spettrale, 248
- Caratteristica:
 - secondo le colonne, 175
 - secondo le righe, 175
- Chiusura:
 - algebrica, 336
 - convessa, 359
- Classe residua, 340
- Classi laterali, 317
- Coefficiente di Fourier, 152
- Coefficienti:
 - di una matrice, 55
 - di un polinomio, 215
- Colonna, 55
- Commutatività, 306
- Componenti di una matrice, 55
- Congruenza, 340
- Coniugazione, 367
- Convoluzione, 334
- Coordinate:
 - di un vettore, 12, 44
 - rispetto a una base, 44
- Corpo, 36
 - astratto, 336
- Curva, 79
- Definito negativo, 249
- Determinante, 117, 146
 - di Vandermonde, 130
- Diagonalizzabilità, 114, 186
- Dimensione, 49
 - finita, 49
- Dipendenza lineare, 42
- Direzione, 17, 155
- Distanza, 23
- Disuguaglianza:
 - di Schwarz, 20, 155
 - triangolare, 22, 154, 160
- Divisione, 259
- Elementi:
 - coniugati, 367
 - diagonali, 59

- Elemento unità**, 309, 334
Endomorfismo, 344
Equazioni:
 differenziali (lineari), 271
 omogenee, 61
Esponente di un elemento, 323
- Fedeltà**, 316
Forma:
 bilineare, 183
 hermitiana, 197
 nulla, 165
 quadratica, 189, 194
 ridotta, 276
 simmetrica, 185
Frazioni parziali, 277
Funzionale, 167
Funzione razionale, 275
Funzioni coordinate, 78
- Generatori**:
 di un gruppo, 309, 323
 di un ideale sinistro, 258
 di uno spazio vettoriale, 39
Generazione, 258
Gradiente, 173
Grado di un polinomio, 215
Gruppo, 305
 banale, 307
 ciclico, 323
 finito, 307
 lineare generale, 306
 quoziente, 320
 simmetrico, 306
 speciale lineare, 307
- Ideale**, 257, 266, 334
 bilatero, 335
 destro, 335
 sinistro, 334
 sinistro principale, 335
 unità, 258
Immagine, 76
 inversa, 319
- Indice**:
 di nullità, 210
 di positività, 211
 di un sottogruppo, 318
Induzione, 361
Iniettività, 94
Insieme:
 convesso, 352
 massimale di elementi linearmente
 indipendenti, 45, 50
Intersezione, 35
Inversa, 68, 91, 144
Invertibilità, 68, 91
Iperpiano, 27
 radente, 355
 separante, 354
Irriducibilità, 260
Isomorfismo:
 di gruppi, 312
 di spazi vettoriali, 96
- Lemma di Schur**, 272
Limitazione inferiore, 357
- Massimo comun divisore**, 259
Matrice, 55
 diagonale, 59
 hermitiana, 199
 nulla (o zero), 56
 quadrata, 56
 semi-simmetrica, 195
 simmetrica, 59
 triangolare superiore, 60
 unitaria, 204
Matrici simili, 115
Modulo, 343
 quoziente, 348
 un ideale, 340
Molteplicità:
 di una radice, 216
 di un autovalore, 230
Monico, 263
Movimento rigido, 313

- Nilpotenza, 72, 235
- Non degenerazione, 65, 149
- Non singolarità, 68
- Norma di un vettore, 21, 154
- Normalità, 27, 252
- Nucleo, 90
 - di un omomorfismo di anelli, 311
- Numeri complessi, 31
- Numero primo, 267
- Omomorfismo, 309, 344
 - canonico, 349
 - di anelli, 338
- Operatore, 100, 192
 - definito positivo, 196
- Ordine, 307
- Ortogonalità, 150, 170, 202
- Ortogonalizzazione di Gram-Schmidt, 153
- Ortonormalità, 155, 161, 208
- Parallelismo, 17
- Periodo, 324
- Permutazione, 133
 - dispari, 137
 - pari, 137
- Perpendicolarità, 19, 150
- Piano, 91
- Polarizzazione, 200
- Polinomio, 213
 - caratteristico, 227, 230
- Primo coefficiente, 215
- Prodotto:
 - alternante, 288
 - definito positivo, 151
 - diretto, 54
 - hermitiano, 157
 - scalare, 18
 - tensoriale, 284
- Proiezione, 24, 152
- Punto:
 - estremo, 356
 - fine, 15
 - origine, 15
- Radice, 216
 - dell'unità, 325
- Rappresentante, 340
- Rappresentare, 172, 193
- Rappresentazione, 316
- Regola di Cramer, 121
- Relativamente primi, 260
- Relazione di equivalenza, 364
- Riga, 55
- Rotazione, 109
- Scalare, 37
- Segno di una permutazione, 134
- Semispazio, 352
- Soluzione banale, 61
- Somma:
 - diretta, 52
 - di sottospazi, 52
- Sottoanello, 336
- Sottocorpo, 37
- Sottogruppo, 309
- Sottoinsieme, 35
 - proprio, 35
- Sottomodulo, 348
- Sottospazio, 39
 - invariante, 272
 - semplice, 272, 348
- Spazio:
 - duale, 167
 - nullo, 165
 - vettoriale, 37
- Surgettività, 94
- Sviluppo:
 - adico, 275
 - di un determinante, 126
- Teorema:
 - di Hamilton-Cayley, 236
 - di Krein-Milman, 358
 - di Sylvester, 207
 - di Wedderburn, 346
- Termine costante, 215
- Traccia, 72
- Traslazione, 313

- Trasposizione, 134
Trasposta:
di un'applicazione lineare, 193
di una matrice, 58
Triangolabilità, 233
Triangolarità superiore stretta, 71
Unione, 36
- Valore:
assoluto, 33
caratteristico, 220
Ventaglio, 232
Vettore:
applicato, 15
colonna, 56
unità, 22, 155

