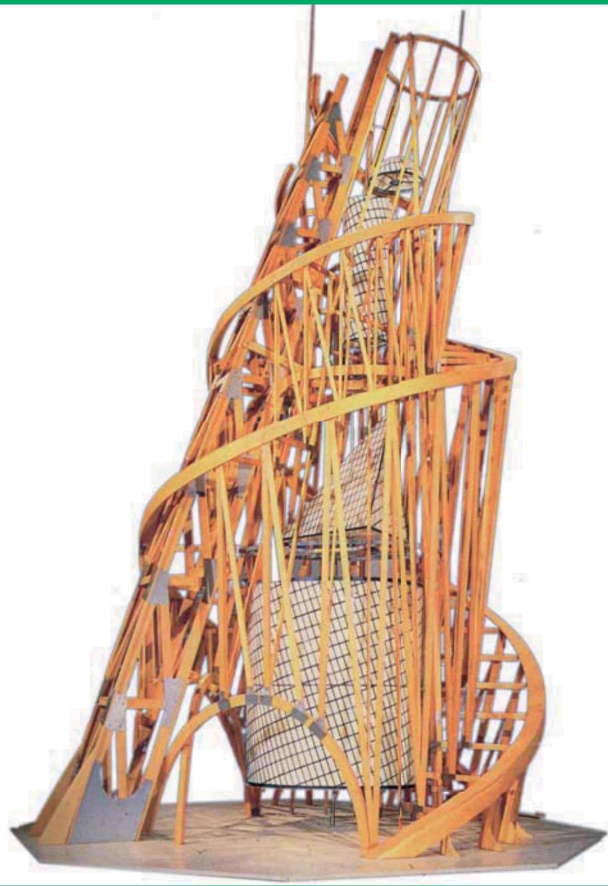
4

Matrici e operatori lineari



APOGEO SPICCHI

M. Conti, D. L. Ferrario, S. Terracini, G. Verzini, V. Barutello

ANALISI MATEMATICA

VOL. 2 con elementi di geometria e calcolo vettoriale

Analisi matematica

Con elementi di geometria e calcolo vettoriale

Volume 2

Vivina Barutello
Monica Conti
Davide L. Ferrario
Susanna Terracini
Gianmaria Verzini

APOGEO

Analisi matematica. Con elementi di geometria e calcolo vettoriale
II.4 Matrici e operatori lineari

Autori:
Vivina Barutello, Monica Conti, Davide L. Ferrario, Susanna Terracini, Gianmaria Verzini

Socio Unico Giangiacomo Feltrinelli Editore s.r.l.
E-mail education@apogeonline.com
U.R.L. www.apogeonline.com

ISBN **978-88-503-1470-6**

Composizione realizzata mediante LaTeX

APOGEO SPICCHI

Componi il tuo libro

Apogeo presenta un progetto di lettura modulare dedicato a lettori, studenti e docenti.
Spesso, di un volume composito e che tratta approfonditamente i differenti aspetti di una disciplina ci interessano solo alcune parti, alcuni argomenti.
In università non è raro che i docenti suggeriscano lo studio di un sottoinsieme di capitoli, ricomponendo in questo modo un percorso didattico adatto alle esigenze di tempi e opportunità inerenti il loro corso.
Per andare incontro alle esigenze formative proprie del mondo accademico e dei lettori che intendono fruire i testi in maniera personalizzata creando un proprio percorso di lettura, Apogeo Education ha realizzato il progetto “Spicchi”.
Grazie alla maggiore versatilità e facilità di distribuzione dei testi in formato elettronico è possibile scegliere all’interno del nostro catalogo e acquistare i singoli capitoli o piccole parti dei nostri testi, a un prezzo decisamente “piccolo”.

Prefazione

In questo secondo volume si prosegue (com'è ovvio) il percorso di avvicinamento all'analisi, e piú in generale, alla Matematica iniziato nel primo volume, per quei corsi di studio in cui questa viene usata per risolvere problemi scientifici e applicativi. Mentre però lo scopo del primo volume era quello di introdurre i concetti e gli strumenti fondamentali dell'analisi infinitesimale (differenziale e integrale), cercando di contestualizzare i vari aspetti della materia, in questo secondo volume l'orizzonte si amplia in modo significativo, dal momento che lo scopo non è solo quello di approfondire e generalizzare a piú variabili quei metodi, ma anche di studiare nel dettaglio alcune importanti applicazioni classiche dei metodi analitici. Si tratta sí di un libro di analisi in piú variabili, con i contenuti tipici di un secondo corso universitario, ma si è voluto fornire anche alcuni strumenti per approfondire e motivare la prospettiva unitaria ch'era alla base del primo volume (riquadri di approfondimento, intermezzi storici, esercizi e problemi teorici guidati, ma anche interi capitoli dedicati ad argomenti importanti, come le equazioni differenziali o l'algebra lineare). Questo ha fatto sí che le dimensioni dell'opera potrebbero risultare eccessive per un normale corso semestrale, e al lettore è richiesto un certo impegno per poterla considerare *in toto*. La distribuzione degli argomenti nei vari capitoli è stata perciò pensata al fine di massimizzare la libertà di scelta di possibili sotto-percorsi, pur conservando una certa consistenza logica (nei prerequisiti e nelle applicazioni). All'inizio di ogni capitolo si è descritto brevemente quali sono gli argomenti che sono necessari per affrontarlo, e quali capitoli sono ragionevolmente affrontabili in seguito. Come nel primo volume, ci sono piú livelli di percorso: un livello base, e un livello piú avanzato, contrassegnato con il simbolo ☢. Questo rivela una discontinuità di livello *locale*, cioè relativa al capitolo in cui ci si trova, e non assoluta. Questo perché alcuni capitoli sono *in sé* piú densi e impegnativi degli altri (per esempio i capitoli IV, VII e VIII).

Per permettere questo accesso a vari livelli, abbiamo scelto alcune soluzioni non convenzionali. Le equazioni differenziali sono affrontate in due capitoli, a due livelli diversi: nel Capitolo I la trattazione avviene principalmente attraverso esempi ed esercizi, mentre la sistematizzazione teorica e gli approfondimenti sono rimandati al Capitolo VIII, che richiede una certa consapevolezza degli strumenti dell'analisi in piú variabili. In un corso di analisi non ridotto ai minimi termini i due capitoli possono essere anche svolti contemporaneamente (l'uno fornendo materiale per le esercitazioni dell'altro). Similmente, l'approccio al calcolo differenziale per funzioni in piú variabili è sviluppato ad un livello elementare (limitato a funzioni di due variabili reali e focalizzato alle applicazioni) nel Capitolo VI, mentre le strutture astratte che permettono una trattazione piú rigorosa sono oggetto del Capitolo VII. Di nuovo, tempo permettendo, i due capitoli possono essere svolti in contemporanea. Viceversa, laddove la minimalità del corso impedisca tali approfondimenti, abbiamo voluto lasciare al lettore interessato la possibilità di completare in modo autonomo la propria formazione.

Per quanto riguarda gli elementi di geometria e di algebra lineare, che solitamente non vengono affrontati nei testi di analisi, abbiamo voluto inserirli proprio per cercare di ricomporre la prospettiva unitaria menzionata sopra (anche perché gli esempi significativi di spazio lineare sono proprio quelli di spazi di funzioni e di soluzioni di equazioni differenziali). Essi si trovano nel Capitolo III ad un livello molto elementare, e nel Capitolo IV ad un livello significativamente piú impegnativo. Completano il volume quattro capitoli con una esposizione piú tradizionale dei classici argomenti del secondo corso, con una particolare attenzione all'interdisciplinarietà: il Capitolo II (successioni e serie di funzioni), il Capitolo V (curve e superfici nello spazio), il Capitolo IX (integrali multipli) e il Capitolo X (integrali di linea e superficie e calcolo vettoriale).

Alla fine di ogni capitolo sono stati selezionati un certo numero di esercizi, rappresentativi o di particolare interesse, da guardare prima di dedicarsi ai numerosi esercizi da svolgere (tutti con soluzione dettagliata in appendice). Rispetto al primo volume, la variabilità della natura e della difficoltà degli esercizi è aumentata, per cui i simboli che segnalano alcune tipologie di esercizi non standard non sono solo la ☆ (per gli esercizi di maggiore difficoltà e interesse teorico), ma anche ☕ (per gli esercizi di natura prevalntemente teorica, che richiedono un certo tempo di riflessione), ⚽ (esercizi di tipo “enigmistico–ludico”), ✍ (esercizi di calcolo), ⌨ (esercizi al calcolatore – pochi in realtà). In questo modo si è cercato di fornire qualche indicazione utile per poter fruire del testo, sia da parte del docente, che da parte del lettore.

ti trancio un tetrastico tetro,
due distici è un torpido metro:
così, un verso avanti, uno indietro,
li incido con schegge di vetro:

EDOARDO SANGUINETI (1930–)

CAPITOLO IV

Matrici e operatori lineari

Nel Capitolo III abbiamo studiato i vettori, e le costruzioni geometriche e algebriche che con essi si possono fare. Quello che faremo in questo capitolo è di considerare tutti i vettori nel loro insieme, *cioè di considerare l'intero* spazio vettoriale, *formato dall'insieme di tutti i vettori. Questa visione "olistica" consente di introdurre le trasformazioni lineari e l'algebra delle matrici, ed alcuni attributi degli spazi vettoriali di notevole importanza: il prodotto scalare, le proiezioni e il determinante. Vedremo alcune applicazioni importanti, che consentiranno di dimostrare in modo generale alcune affermazioni che avevamo anticipato nel Capitolo III. Concluderemo il capitolo con due argomenti (legati tra loro) di notevole rilievo per il seguito del libro, che possono però presentare qualche difficoltà tecnica. Si tratta dello studio delle* forme quadratiche *e degli* autovalori. *Stiamo affrontando un capitolo inizialmente semplice, ma che si farà via via sempre più impegnativo, e finirà con un arrivo in salita decisamente faticoso. Cerchiamo quindi di iniziare in discesa, con una poco sensata parafrasi/parodia di un famoso* incipit.

Alla prima lezione sulle matrici cui andai mi aspettavo di essere annoiato a morte o forse nauseato da quello che mi avevano detto sarebbe capitato ai presenti. Ogni cosa che avevo letto a proposito delle lezioni universitarie insisteva su quel punto: la maggior parte delle persone che ne scrivevano condannavano apertamente le dimostrazioni come un affare inutile, stupido e perfino brutale, ma anche coloro che parlavano bene delle matrici come esibizione di abilità e come strumento utile deploravano l'uso delle dimostrazioni e si mostravano apertamente spiacenti della loro presenza. Il continuo esibire dimostrazioni nelle aule universitarie era considerato deplorevole. Suppongo, da un punto di vista moderno, cioè, il punto di vista delle applicazioni immediate, che l'intera storia delle matrici, o anche tutta la matematica, non sia difendibile; c'è certamente molta crudeltà nel dimostrare l'esistenza e le proprietà del determinante, c'è sempre il pericolo, cercato o capitato, di perdere tempo in speculazioni inutili, e arriva quasi sempre il momento in cui ci si trova a ragionare in astratto o per induzione. Ora cercherei non di difendere quest'impostazione, ma solo di spiegare onestamente quelle poche cose che ho imparato su questo argomento. Per far questo dovrò essere complessivamente chiaro, o almeno cercare di esserlo, e se coloro che leggono decidono con disgusto che questo testo è scritto da qualcuno che manca della loro, cioè dei Lettori, finezza di sentimenti e opinioni, posso solamente ammettere che in effetti potrebbe essere vero. Ma chiunque legga questo capitolo può esprimere un giudizio solo quando ha visto e capito le cose di cui si parla e sa davvero quale potrebbe essere la propria reazione ad esse. Ricordo che un giorno Gertrude Stein, a proposito di operatori lineari, parlò della sua ammirazione per Frobenius e mostrò alcune sue foto durante una lezione, con lei e Alice Toklas nella prima fila dei banchi di legno dell'aula di Valencia, con Frobenius e suo fratello Perron di sotto, e io ero appena tornato dal Medio Oriente, e ricordo di averle detto che non amavo le matrici per colpa delle dimostrazioni. In realtà la cosa più difficile, a parte sapere davvero cosa si prova e come ci si confronta con l'astratto, piuttosto che quello che ci si aspetta di provare o che ci è stato insegnato a giudicare vero, è mettere giu i fatti per come sono: che cosa è realmente uno spazio vettoriale, cosa è una matrice, a cosa serve, quali sono le proprietà alla radice dei risultati che si sperimentano veri.[10]

§ 1. Spazi vettoriali

Nel Capitolo III abbiamo dato la definizione di vettore più semplice possibile: una n-upla di numeri su cui è possibile applicare due operazioni (prodotto per uno scalare e somma). Cerchiamo di formalizzare, cioè di rendere un poco più astratta, quella nozione. Prendiamo in considerazione solo le *proprietà* elencate nella Proposizione (III.4): se $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ sono tre vettori della stessa dimensione e c, d sono due scalari, allora

▶ Vedi § 1 a pagina 99.

① (commutativa) $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}, \quad c\boldsymbol{u} = \boldsymbol{u}c.$

② (associativa) $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}), \quad (cd)\boldsymbol{u} = c(d\boldsymbol{u}).$

③ (distributiva) $c(\boldsymbol{u} + \boldsymbol{w}) = c\boldsymbol{u} + c\boldsymbol{w}, \ (c + d)\boldsymbol{u} = c\boldsymbol{u} + d\boldsymbol{u}.$

④ (vettore zero) Esiste un elemento $\mathbf{0}$ tale che $\boldsymbol{u} + \mathbf{0} = \boldsymbol{u}, \ \boldsymbol{u} - \boldsymbol{u} = \mathbf{0}.$

⑤ (uno) $1\boldsymbol{u} = \boldsymbol{u}, \ 0\boldsymbol{u} = \mathbf{0}.$

Definiremo ogni insieme di oggetti su cui si possano definire una somma e un prodotto per uno scalare con queste proprietà uno *spazio vettoriale*. Più precisamente:

(IV.1) Definizione. Sia V un insieme, su cui sono definite due operazioni: una di somma e una di moltiplicazione per un numero reale. Se queste due operazioni soddisfano tutte le proprietà sopra elencate, allora V si dice *spazio vettoriale* (o, equivalentemente, *spazio lineare*).

▶ Oltre che i numeri reali $\mathbb{R}$, anche i numeri complessi $\mathbb{C}$ possono servire da coefficienti di uno spazio vettoriale. In questo caso si parla di *spazio vettoriale complesso*. Si veda il riquadro a pagina 215.

(IV.2) Esempio. Per ogni $n \geq 1$ l'insieme di tutte le n-uple di numeri reali (l'insieme $\mathbb{R}^n$ definito a pagina 100) con la somma e il prodotto per uno scalare definiti in (III.2) e (III.1) costituisce l'esempio canonico di spazio vettoriale di dimensione n. Lo indicheremo ancora con il simbolo $\mathbb{R}^n$. ■

(IV.3) Esempio. Ricordiamo che su $\mathbb{R}^n$ si può definire il prodotto scalare, come in (III.16). Consideriamo quindi l'insieme V di tutti i vettori $\boldsymbol{v} \in \mathbb{R}^n$ ortogonali ad un vettore $\boldsymbol{n}$ dato, cioè quelli per cui $\boldsymbol{v} \cdot \boldsymbol{n} = 0$. La somma e il prodotto per uno scalare saranno quelli di $\mathbb{R}^n$, e rimane da controllare che in effetti se $\boldsymbol{v} \in V$ e $\boldsymbol{w} \in V$ sono due vettori di V e c è uno scalare, allora sia $\boldsymbol{v} + \boldsymbol{w}$ che $c\boldsymbol{v}$ sono ancora elementi di V: ma $\boldsymbol{v} \cdot \boldsymbol{n} = \boldsymbol{w} \cdot \boldsymbol{n} = 0 \implies (\boldsymbol{v} + \boldsymbol{w}) \cdot \boldsymbol{n} = \boldsymbol{v} \cdot \boldsymbol{n} + \boldsymbol{w} \cdot \boldsymbol{n} = 0$, e $(c\boldsymbol{v}) \cdot \boldsymbol{n} = c(\boldsymbol{v} \cdot \boldsymbol{n}) = 0$, dunque V è uno spazio vettoriale. ■

▶ Da non confondere con il prodotto *per uno scalare*.

(IV.4) Esempio. Sia V l'insieme di tutte le funzioni continue definite su un intervallo e a valori in $\mathbb{R}$. La somma di due funzioni $f, g \in V$ si definisce ponendo

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

per ogni x, e analogamente il prodotto per uno scalare è

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

Si può vedere che V è uno spazio vettoriale, in cui la funzione ovunque zero fa da $\mathbf{0}$. ■

▶ Lo spazio vettoriale delle funzioni continue su $[a, b]$ viene solitamente indicato con

$$C([a,b]).$$

▶ Il Lettore dovrebbe provare a dimostrare questo fatto per esercizio.

(IV.5) Esempio. L'elenco di spazi vettoriali è praticamente senza fine. I seguenti sono alcuni tra gli spazi vettoriali più diffusi.

① Lo spazio $\mathbb{R}^n$.

② Lo spazio di tutte le matrici $n \times m$ a coefficienti reali.

▶ Che definiremo tra poco, nel § 2.

③ Lo spazio dei polinomi di grado al più n.

④ L'insieme di tutte le funzioni definite su un intervallo (a, b).

⑤ L'insieme di tutte le funzioni derivabili in (a, b), oppure integrabili in (a, b).

⑥ L'insieme di tutte le soluzioni di una equazione differenziale lineare omogenea del secondo ordine:

$$y''(x) + a(x)y'(x) + b(x)y(x) = 0.$$

▶ Si veda il § 4 a pagina 13.

⑦ I polinomi trigonometrici (di grado al più n).

⑧ Le funzioni periodiche, le funzioni pari, le funzioni dispari, le funzioni a media nulla.

⑨ Le successioni di numeri reali convergenti.

⑩ I numeri complessi $\mathbb{C}$, con il campo degli scalari $\mathbb{R}$, il prodotto consueto di un numero reale per un numero complesso e la somma di numeri complessi. ■

▶ Si veda il § 4 a pagina 53, e la Definizione (II.23).

Ricordiamo la nozione di combinazione lineare di k vettori $v_1, v_2, \ldots, v_k$ con coefficienti $c_1, c_2, \ldots, c_k$, che avevamo dato in (III.3):

$$c_1 v_1 + c_2 v_2 + \ldots + c_k v_k.$$

Non c'è traccia esplicita delle componenti dei vettori $v_1, v_2, \ldots, v_k$, e quindi si può parlare di combinazione lineare di vettori in spazi vettoriali astratti. Intenderemo sempre che le combinazioni lineari sono combinazioni di un numero *finito* di vettori.

Un'altra definizione che può tranquillamente essere utilizzata è quella di dipendenza o indipendenza lineare, che avevamo dato in (III.8): k vettori $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$ sono linearmente (in)dipendenti se l'equazione

$$\boldsymbol{a}_1 x_1 + \boldsymbol{a}_2 x_2 + \ldots + \boldsymbol{a}_k x_k = \mathbf{0} \tag{1}$$

(non) ammette soluzioni non nulle $(x_1, x_2, \ldots, x_k) \neq \mathbf{0}$. Possiamo dire che i vettori di un insieme (anche infinito) S sono dipendenti se per almeno un k esistono k vettori $\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_k$ di S linearmente dipendenti. Naturalmente i vettori di un insieme (anche infinito) S sono indipendenti se non sono dipendenti.

(IV.6) Definizione. Sia S un insieme di elementi di uno spazio vettoriale V. Allora si indica con $\operatorname{Span}(S)$ l'insieme di tutte le combinazioni lineari (finite) di elementi di S, e si chiama lo *spazio lineare generato da* S. Poniamo $\operatorname{Span}(\emptyset) = \{\mathbf{0}\}$. Gli elementi di S si dicono *generatori* di $\operatorname{Span}(S)$. Se S ha un numero finito di elementi $\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_k$, allora si indica anche con $\operatorname{Span}(S) = \operatorname{Span}(\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_k)$.

Lo *span* di un insieme di vettori di V è anch'esso uno spazio vettoriale (con le operazioni ereditate da V), e viene anche chiamato un *sottospazio vettoriale* di V. Lo si provi a dimostrare per esercizio (Esercizio 46). Possiamo riscrivere la proprietà di avere o non avere soluzioni per un sistema di equazioni lineari nel seguente modo.

Figura IV.1: Hermann Günter Grassmann (1809–1877), che introdusse la nozione di spazio vettoriale astratto.

(IV.7) *Siano $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$ e $\boldsymbol{b}$ vettori di $\mathbb{R}^n$. Il sistema di equazioni*

$$\boldsymbol{a}_1 x_1 + \boldsymbol{a}_2 x_2 + \ldots + \boldsymbol{a}_k x_k = \boldsymbol{b}$$

ha soluzioni se e soltanto se

$$\boldsymbol{b} \in \operatorname{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k).$$

(IV.8) Definizione. Sia V uno spazio vettoriale. Un insieme di elementi $S \subset V$ tale che $\mathrm{Span}(S) = V$ è detto insieme di *generatori* di V. In altre parole, S è un insieme di generatori per V quando ogni elemento di V si può scrivere come combinazione lineare di elementi di S

$$v = c_1 s_1 + c_2 s_2 + \ldots + c_k s_k.$$

(IV.9) Definizione. Chiamiamo *base* dello spazio vettoriale V una n-upla di vettori di V

$$(e_1, e_2, \ldots, e_n)$$

tale che gli e_i sono indipendenti e $\mathrm{Span}(e_1, e_2, \ldots, e_n) = V$, cioè

① i vettori che la compongono siano indipendenti e

② siano generatori per V.

▶ Più propriamente, questa sarebbe la definizione di *base ordinata*. In generale si intende per base l'*insieme* dei vettori e_1, e_2, ..., e_n e non la n-upla (che è un insieme ordinato). Per quanto riguarda gli scopi di questo testo ci è sembrato più semplice dare questa definizione.

(IV.10) Esempio. I vettori $e_1 = i$, $e_2 = j$ e $e_3 = k$ sono una base di $\mathbb{R}^3$. In generale la cosiddetta *base canonica* di $\mathbb{R}^n$ è costituita dai vettori

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad e_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

■

(IV.11) Esempio. I vettori $s_1 = (1, 0)$ e $s_2 = (1, 1)$ costituiscono una base per $\mathbb{R}^2$: infatti sono linearmente indipendenti, e quindi per il Teorema (III.23) ogni vettore $v \in \mathbb{R}^2$ si può scrivere come combinazione lineare di $(1, 0)$ e $(1, 1)$. ■

▶ Per i Teoremi (III.23) e (III.29), ogni volta che si prendono due vettori in $\mathbb{R}^2$ o tre vettori in $\mathbb{R}^3$, se sono linearmente indipendenti allora costituiscono una base, dato che generano tutto lo spazio vettoriale. Torneremo su questa proprietà nella Proposizione (IV.17) a pagina 172.

(IV.12) Esempio. Lo spazio vettoriale costituito da tutti i polinomi in x a coefficienti reali di grado al più n ha $n+1$ generatori

$$1, x, x^2, \ldots, x^n.$$

Sono indipendenti, dato che una loro combinazione lineare

$$a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

è uguale al polinomio nullo se e solo se tutti i coefficienti a_i sono zero. ■

Consideriamo una base $(b_1, b_2, \ldots, b_n)$ di uno spazio vettoriale V. Allora, per ogni $x \in V$, l'equazione

(2) $\quad x_1 b_1 + x_2 b_2 + \ldots + x_n b_n = x$

ha almeno una soluzione▶. Ma visto che i $\boldsymbol{b}_i$ sono linearmente indipendenti, questa soluzione è unica. Se così non fosse, esisterebbero due n-uple $(x_1, \dots, x_n)$ e $(x'_1, \dots, x'_n)$ tali che

▶ Dato che $\operatorname{Span}(\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n) = V$.

$$x_1\boldsymbol{b}_1 + \dots + x_n\boldsymbol{b}_n = \boldsymbol{x} = x'_1\boldsymbol{b}_1 + \dots + x'_n\boldsymbol{b}_n,$$

da cui $(x_1 - x'_1)\boldsymbol{b}_1 + \dots + (x_n - x'_n)\boldsymbol{b}_n = \boldsymbol{0}$. Ma visto che i vettori $\boldsymbol{b}_i$ sono linearmente indipendenti, si avrebbe $x_i - x'_i = 0$ per ogni i, cioè gli x_i non possono essere diversi dagli x'_i.

Se vale la (2), i numeri x_1, x_2, $\dots$, x_n sono detti le *componenti* del vettore $\boldsymbol{x}$ nella base $(\boldsymbol{b}_1, \dots, \boldsymbol{b}_n)$.

▶ Segue da questa definizione che le componenti di un vettore di $\mathbb{R}^n$ rispetto alla base standard sono esattamente le sue componenti, quando considerato una semplice n-upla. Cioè, un vettore di $\mathbb{R}^n$ ha n componenti rispetto alla base standard (in quanto vettore), e n componenti (in quanto n-upla) che coincidono con le precedenti. Per esempio, $(1,2) = \boldsymbol{x} = \boldsymbol{i} + 2\boldsymbol{j} \in \mathbb{R}^2$ ha componenti $(1,2)$ rispetto alla base canonica di $\mathbb{R}^2$. Se però invece della base canonica usiamo quella vista nell'Esempio (IV.11), allora $-\boldsymbol{s}_1 + 2\boldsymbol{s}_2 = -(1,0) + 2(1,1) = (1,2) = \boldsymbol{x}$, *laonde per cui* $\boldsymbol{x} = -\boldsymbol{s}_1 + 2\boldsymbol{s}_2$ ha componenti $(-1,2)$ rispetto alla nuova base.

Osserviamo una cosa importante: se c è un numero, allora $c\boldsymbol{x}$ è un elemento di V (risultato del prodotto scalare-vettore proprio di V), che a sua volta ha componenti nella base $(\boldsymbol{b}_1, \dots, \boldsymbol{b}_n)$. Ma dato che

$$cx_1\boldsymbol{b}_1 + cx_2\boldsymbol{b}_2 + \dots + cx_n\boldsymbol{b}_n = c\boldsymbol{x}$$

e la soluzione è unica, le componenti di $\boldsymbol{x}$ in effetti si ottengono moltiplicando lo scalare c per il vettore $(x_1, x_2, \dots, x_n)$. Lo stesso per la somma: se $\boldsymbol{y}$ è un altro vettore, con componenti $y_1, y_2, \dots, y_n$, allora

$$(x_1 + y_1)\boldsymbol{b}_1 + (x_2 + y_2)\boldsymbol{b}_2 + \dots + (x_n + y_n)\boldsymbol{b}_n = \boldsymbol{x} + \boldsymbol{y},$$

e quindi le componenti di $\boldsymbol{x} + \boldsymbol{y}$ (la cui somma è la somma definita in V) si ottengono sommando i vettori delle componenti $(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n)$. La scelta di una base per V ci consente quindi di parametrizzare gli elementi di V con gli n parametri x_1, x_2, $\dots$, x_n in modo tale che le operazioni su V corrispondono alle operazioni delle n-uple. Questa corrispondenza si scrive anche come $V \cong \mathbb{R}^n$.

▶ In linea di principio si potrebbe pensare che il numero n (la dimensione) dipenda dalla scelta fatta della base. Invece, vedremo subito sotto che ciò non accade. L'esistenza di una parametrizzazione è importante, e verrà sottolineata ancora nel Corollario (IV.16).

(IV.13) *Se $(\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n)$ ed $(\boldsymbol{e}'_1, \boldsymbol{e}'_2, \dots, \boldsymbol{e}'_m)$ sono due basi per lo spazio vettoriale V, allora $n = m$. In altre parole, le basi di un dato spazio vettoriale hanno tutte lo stesso numero di elementi (posto che ne esista almeno una).*

Dimostrazione. Dato che $(\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n)$ è una base per V, possiamo scrivere i vettori delle componenti degli elementi $\boldsymbol{e}'_1$, $\boldsymbol{e}'_2$, $\dots$, $\boldsymbol{e}'_m$ in questa base, e li denoteremo con i simboli $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\dots$, $\boldsymbol{a}_m$; cioè per ogni $i = 1 \dots m$ il vettore $\boldsymbol{a}_i \in \mathbb{R}^n$ è l'unico vettore di dimensione n che risolve l'equazione

$$\boldsymbol{e}'_i = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \dots + x_n\boldsymbol{e}_n$$

nelle incognite $x_1, \dots, x_n$. Abbiamo ancora il problema dei due indici, ma se concordiamo di scrivere, per ogni i, $\boldsymbol{a}_i = (a_{1i}, a_{2i}, \dots, a_{ni})$, possiamo riscrivere anche $\boldsymbol{e}'_i = a_{1i}\boldsymbol{e}_1 + a_{2i}\boldsymbol{e}_2 + \dots + a_{ni}\boldsymbol{e}_n$. Ora, supponiamo che $m > n$. Vogliamo mostrare che gli $\boldsymbol{e}'_i$ sono linearmente dipendenti. Dimostreremo cioè il seguente lemma.

▶ Lo risolveremo tra non molto, introducendo finalmente il concetto di *matrice*.

(IV.14) Lemma. *Se V ha una base con n vettori, allora $m > n$ vettori di V sono linearmente dipendenti.*

Ma questo è equivalente a chiedere che gli m vettori delle componenti $\boldsymbol{a}_1$, ..., $\boldsymbol{a}_m$ siano linearmente dipendenti, cioè che esista una soluzione non nulla $(x_1, x_2, \ldots, x_m)$ dell'equazione $x_1 e'_1 + x_2 e'_2 + \ldots + x_m e'_m = \mathbf{0}$, che può essere scritta come

▶ Sostituendo le espressioni per gli e'_i.

$$x_1(a_{11}e_1 + a_{21}e_2 + \ldots) + x_2(a_{12}e_1 + a_{22}e_2 + \ldots) + \ldots + x_m(a_{1m}e_1 + a_{2m}e_2 + \ldots) = \mathbf{0}.$$

Raccogliendo (per le proprietà algebriche dei vettori), si ottiene l'equazione

$$\begin{aligned}(x_1a_{11} + x_2a_{12} + \ldots + x_ma_{1m})e_1 + (x_1a_{21} + x_2a_{22} + \ldots + x_ma_{2m})e_2 + \\ \ldots + (x_1a_{n1} + x_2a_{n2} + \ldots + x_ma_{nm})e_n = \mathbf{0}.\end{aligned}$$

Quest'ultima, dato che $e_1, e_2, \ldots, e_n$ sono linearmente indipendenti, è equivalente al sistema di equazioni

▶ Costituiscono una base!

$$\begin{cases} x_1a_{11} + x_2a_{12} + \ldots + x_ma_{1m} = & 0 \\ x_1a_{21} + x_2a_{22} + \ldots + x_ma_{2m} = & 0 \\ \qquad \ldots & \\ x_1a_{n1} + x_2a_{n2} + \ldots + x_ma_{nm} = & 0 \end{cases}$$

cioè $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \ldots + x_m\boldsymbol{a}_m = \mathbf{0}$. Questo è un sistema di n equazioni in m incognite, che ha sempre la soluzione $(x_1, x_2, \ldots, x_m) = \mathbf{0}$. Ora immaginiamo di procedere con l'eliminazione successiva delle variabili, Il procedimento di eliminazione si conclude quando non si possono più eliminare variabili: nel nostro caso con n equazioni possiamo eliminare n variabili al massimo. Le variabili non eliminate (le cosiddette variabili libere) saranno quindi almeno $m - n \geq 1$; dato che almeno una soluzione c'è (quella nulla), non potrà accadere di terminare su un sistema senza soluzioni, e quindi si può concludere che se $m > n$ i vettori $e'_1, \ldots e'_m$ sono linearmente dipendenti. Abbiamo concluso la dimostrazione del Lemma (IV.14). Ma per ipotesi questi vettori costituiscono una base, quindi non può essere che $m > n$, ovvero deve essere $m \leq n$. Adesso riprendiamo il ragionamento da capo, scambiando i ruoli di $(e_1, \ldots, e_n)$ e $(e'_1, \ldots, e'_m)$, per dedurne che è vero anche che $n \leq m$. Ma allora non può essere che $n = m$. ■

▶ Si veda il § 2 del Capitolo III a pagina 101. Potremmo non riuscire ad eliminarle tutte perché qualche variabile potrebbe perdersi nel procedimento — si veda l'Esercizio (III.2) a pagina 144.

(IV.15) Definizione. Il numero di elementi di una qualsiasi base di uno spazio vettoriale V è la *dimensione* di V (indicata con $\dim V$). Se V non ha una base finita, si dice che ha *dimensione infinita*.

▶ Non tutti gli spazi vettoriali hanno una base finita, cioè ci sono spazi vettoriali di dimensione infinita. Per esempio, possiamo estendere l'Esempio (IV.12) e considerare lo spazio vettoriale di tutti i polinomi, di ogni grado. Questo è generato da tutti gli infiniti monomi x^n, con $n \geq 0$, ma non può essere generato da un insieme finito di generatori. Molti altri spazi di dimensione infinita (spazi di funzioni continue, derivabili, con sviluppi in serie di Taylor, ...) sono di grande importanza (si vedano anche l'Esempio (IV.50) e il riquadro a pagina 190).

Nella dimostrazione di (IV.13) è di fatto dimostrato anche il seguente Corollario (cfr. la Definizione (IV.15)).

(IV.16) Corollario. *Ogni spazio vettoriale di dimensione n è parametrizzato con n parametri liberi: $V \cong \mathbb{R}^n$.*

(IV.17) *Consideriamo n vettori $e_1, e_2, \ldots, e_n$ di uno spazio vettoriale V di dimensione n. Allora:*

① *$(e_1, e_2, \ldots, e_n)$ è una base se e solo se i vettori e_i sono linearmente indipendenti;*

② $(e_1, e_2, \dots, e_n)$ *è una base se e solo se* $\mathrm{Span}(e_1, e_2, \dots, e_n) = V$.

Dimostrazione. In entrambe, la parte "solo se" è ovvia. Per la ①, dobbiamo mostrare che se gli n vettori e_i sono indipendenti, allora $\mathrm{Span}(e_1, e_2, \dots, e_n) = V$. Sia $v \in V$ un vettore qualsiasi non nullo. Per il Lemma (IV.14), $(v, e_1, \dots, e_n)$ sono $n+1 > n$ vettori linearmente dipendenti (dato che V ha certamente una base con n vettori), cioè esistono $n+1$ coefficienti c_i tali che

▶ Dato che per definizione i vettori di una base devono essere linearmente indipendenti e generare V.

$$c_0 v + c_1 e_1 + c_2 e_2 + \dots c_n e_n = \mathbf{0}.$$

Il primo coefficiente, c_0, non può essere zero▶, e quindi, dividendo per $-c_0$ otteniamo

▶ Dato che gli e_i sono linearmente indipendenti.

$$v = -(c_1 e_1 + c_2 e_2 + \dots c_n e_n)/c_0 \in \mathrm{Span}(e_1, e_2, \dots, e_n).$$

Per la ②, dobbiamo mostrare che se $\mathrm{Span}(e_1, e_2, \dots, e_n) = V$ allora gli e_i sono linearmente indipendenti. Se per assurdo non lo fossero, sarebbe possibile scrivere uno di loro come combinazione lineare degli altri. Supponiamo (a meno di cambiare gli indici) che sia

$$e_n = c_1 e_1 + c_2 e_2 + \dots + c_{n-1} e_{n-1}.$$

Ma allora

$$V = \mathrm{Span}(e_1, e_2, \dots, e_n) = \mathrm{Span}(e_1, e_2, \dots, e_{n-1}).$$

▶ Se e_n è combinazione lineare degli altri, allora ogni combinazione lineare dei e_i con $i = 1, \dots, n$ può sempre scriversi come combinazione lineare dei $e_1, \dots, e_{n-1}$, sostituendo a e_n la sua espressione lineare in funzione degli altri.

Proseguiamo così fino a che non abbiamo $k < n$ vettori linearmente indipendenti (a meno di permutazioni degli indici) $e_1, e_2, \dots, e_k$ che generano V. Ma allora V ha dimensione $k < n$: assurdo, perché per la Proposizione (IV.13) dovrebbe essere $k = n$. ∎

(IV.18) *La dimensione dello spazio V generato da m vettori $v_1, v_2, \dots, v_m$ (non necessariamente indipendenti) non può essere maggiore di m, ed è il massimo numero degli v_i linearmente indipendenti.*

▶ Tra un po' (a pagina 196) torneremo su questa proprietà, quando ci occuperemo del *rango* di una matrice.

Dimostrazione. Come per la dimostrazione di ② nella Proposizione (IV.17), togliendo vettori uno per volta si arriva ad un insieme di k vettori linearmente indipendenti che generano V, con $k \le m$. Se se ne tolgono altri, e si arriva ad un insieme di k' vettori con la stessa proprietà, per (IV.13) deve essere $k = k'$. Inoltre, per (IV.14), se un sottoinsieme di k vettori linearmente indipendenti scelto tra i v_i genera V, allora ogni insieme di $k' > k$ vettori non può essere linearmente indipendente (e quindi la dimensione k è il massimo numero di vettori v_i linearmente indipendenti). ∎

§ 2. Matrici e trasformazioni lineari

Ci è capitato di incontrare più volte casi in cui si considerano m-uple di vettori di dimensione n

$$(a_1, a_2, \dots, a_m),$$

in cui per ogni $j = 1, \dots, m$ il vettore $\boldsymbol{a}_j$ ha a sua volta n componenti $(a_{1j}, a_{2j}, \dots, a_{nj})$. Il modo più sensato per disporre i coefficienti a_{ij} è quello della *matrice* $n \times m$:

$$A = (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m) = \left(\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, \cdots, \begin{bmatrix} a_{1m} \\ a_{2m} \\ \cdots \\ a_{nm} \end{bmatrix} \right) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

in cui le m *colonne* (vettori di dimensione n) sono le componenti degli m vettori $\boldsymbol{a}_1, \dots, \boldsymbol{a}_m$, mentre le n *righe* sono vettori (chiamati vettori riga) di dimensione m, tutte le j-esime componenti degli m vettori. Si dice che la matrice A indicata sopra ha *componenti* o *coefficienti* a_{ij} (che a volte, per maggior chiarezza, scriveremo anche come $a_{i,j}$) ed è una matrice di n righe per m colonne (in breve: $n \times m$). A parte la rappresentazione "geometrica" bidimensionale, una matrice quindi può essere vista semplicemente come una funzione $A \colon \{1, \dots, n\} \times \{1, \dots, m\} \to \mathbb{R}$, in cui si scrive $A(i, j) = a_{ij}$, oppure come un vettore che ha vettori per componenti. Se $n = m$ la matrice si dice *quadrata*.

Figura IV.2: la *Melanconia* di Albrecht Dürer (1471–1528), in cui in alto a destra (poco visibile) compare un quadrato magico, cioè la matrice 4×4 (con alcune proprietà magiche)

$$\begin{bmatrix} 16 & 3 & 2 & 13 \\ 5 & 10 & 11 & 8 \\ 9 & 6 & 7 & 12 \\ 4 & 15 & 14 & 1 \end{bmatrix}.$$

(IV.19) Esempio (Matrice nulla e identica)**.** La matrice I_n (identica) e la matrice 0_n (nulla), che vengono spesso indicate per semplicità con I e 0 con abuso di notazione, sono le seguenti matrici $n \times n$

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}, \quad 0_n = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$ ■

Le due operazioni dei vettori (III.1) e (III.2) possono essere facilmente estese alle matrici, ponendo

$$c \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1m} \\ ca_{21} & ca_{22} & \cdots & ca_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{n1} & ca_{n2} & \cdots & ca_{nm} \end{bmatrix},$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1m}+b_{1m} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2m}+b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \cdots & a_{nm}+b_{nm} \end{bmatrix}.$$

Il fatto di avere due indici consente di definire un certo numero di altre operazioni tra matrici, oltre che queste due. La prima operazione, chiamata *prodotto righe per colonne*, è anche la più importante in assoluto per le matrici. Invece che definirlo subito, torniamo al concetto importante di *combinazione lineare* di m vettori di dimensione n, che abbiamo

▶ Come vedremo, queste operazioni portano ad una formidabile quantità di conseguenze.

già visto:

$$\boldsymbol{a}_1 x_1 + \boldsymbol{a}_2 x_2 + \ldots + \boldsymbol{a}_m x_m.$$

Il mattone fondamentale per il prodotto righe per colonne è proprio la combinazione lineare: iniziamo con il definire il prodotto di una matrice per un solo vettore, che immaginiamo disposto come *colonna*.

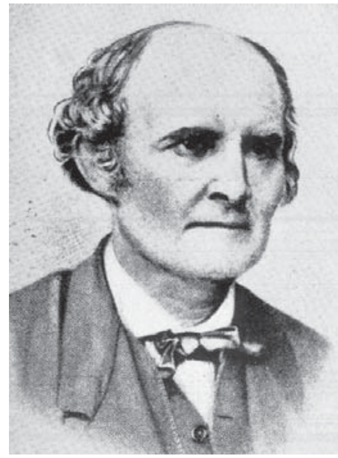


Figura IV.3: Arthur Cayley (1821–1895), il cui nome è legato alla teoria delle matrici.

(IV.20) Definizione. Il *prodotto righe per colonne* di una matrice per un vettore colonna è definito come segue:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} x_2 + \cdots + \begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{bmatrix} x_m$$

$$= \boldsymbol{a}_1 x_1 + \boldsymbol{a}_2 x_2 + \ldots + \boldsymbol{a}_m x_m = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m \end{bmatrix}$$

La definizione di prodotto di una matrice di n righe e m colonne per un vettore (colonna) di dimensione m ci porta subito a poter definire il prodotto di una matrice di n righe e m colonne ($n \times m$) per un'altra matrice di m righe e k colonne, semplicemente giustapponendo i prodotto della matrice per i vettori colonna, cioè se $B = (\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_k)$ è una seconda matrice formata da k vettori (colonna) di dimensione m, si ha

(3) $\quad AB = A(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n) = (A\boldsymbol{b}_1, A\boldsymbol{b}_2, \ldots, A\boldsymbol{b}_k).$

Il prodotto è definito solo quando il numero di colonne della matrice di sinistra A è uguale al numero di righe della matrice di destra B.

(IV.21) Esempio.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} = \begin{bmatrix} a+2d & b+2e & c+2f \\ 3a+4d & 3b+4e & 3c+4f \end{bmatrix}.$$

▶ Si può vedere che il termine di posto (i, j) della matrice prodotto è uguale al prodotto scalare della i-esima riga del primo fattore per la j-esima colonna del secondo (nel senso delle m-uple di numeri):

$$(AB)_{ij} = (i\text{-esima riga di } A)\cdot \\ \cdot (j\text{-esima colonna di } B).$$

È quasi facile dimostrare le seguenti proprietà del prodotto righe per colonne (tralasciamo di elencare le proprietà della somma e del prodotto per uno scalare, che sono identiche a quelle dei vettori).

(IV.22) *Se A, B sono matrici, $\boldsymbol{x}$ e $\boldsymbol{y}$ due vettori della stessa dimensione (che possono essere anche pensati come vettori colonna) e c uno scalare, allora (a patto che le operazioni siano ben definite, a seconda del numero di righe/colonne):*

① $A(\boldsymbol{x} + \boldsymbol{y}) = A\boldsymbol{x} + A\boldsymbol{y}$, $(A + B)\boldsymbol{x} = A\boldsymbol{x} + B\boldsymbol{x}$.

② $A(c\boldsymbol{x}) = c(A\boldsymbol{x}) = (cA)\boldsymbol{x}$.

③ $(AB)\boldsymbol{x} = A(B\boldsymbol{x})$.

Seguono subito[11], per la definizione di prodotto tra matrici, le proprietà seguenti: se A, B e C sono matrici e c uno scalare, allora

① $A(B + C) = AB + AC$, $(A + B)C = AC + BC$.

② $A(cB) = c(AB) = (cA)B$.

③ $(AB)C = A(BC)$.

▶ A differenza del prodotto tra numeri, del prodotto scalare, e del prodotto scalare–vettore, il prodotto tra matrici in generale non è commutativo, cioè non è detto che valga $AB = BA$. Innanzitutto perché non sempre entrambi i prodotti sono ben definiti (affinché lo siano, se A è $n \times m$, B deve essere $m \times n$) ma anche se lo fossero potrebbe capitare che $AB \neq BA$. Per esempio

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

mentre

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Per questo motivo si deve sempre fare attenzione se si tratta di una moltiplicazione *a destra* o *a sinistra* (cfr. Esercizio 19 a pag. 226).

Con questa nuova notazione possiamo riscrivere il sistema di equazioni

$$\begin{cases} x_1a_{11} + x_2a_{12} + \ldots + x_ma_{1m} = & b_1 \\ x_1a_{21} + x_2a_{22} + \ldots + x_ma_{2m} = & b_2 \\ \qquad\qquad \ldots & \\ x_1a_{n1} + x_2a_{n2} + \ldots + x_ma_{nm} = & b_n \end{cases} \iff \boldsymbol{a}_1x_1 + \boldsymbol{a}_2x_2 + \ldots + \boldsymbol{a}_mx_m = \boldsymbol{b}$$

semplicemente come

$$A\boldsymbol{x} = \boldsymbol{b},$$

dove A è la matrice con coefficienti a_{ij},

$$\boldsymbol{x} = (x_1, x_2, \ldots, x_m) \quad \text{e} \quad \boldsymbol{b} = (b_1, b_2, \ldots, b_n).$$

Osserviamo che $\boldsymbol{x} \in \mathbb{R}^m$ e $\boldsymbol{b} \in \mathbb{R}^n$, e che la matrice A consente di definire una *funzione* (equivalentemente, una *applicazione*, *operatore* o *trasformazione*) $A\colon \mathbb{R}^m \to \mathbb{R}^n$, definita ponendo

(4) $\mathbb{R}^m \ni \boldsymbol{x} \mapsto A\boldsymbol{x} \in \mathbb{R}^n$.

▶ Usiamo lo stesso simbolo per denotare la funzione e la matrice che la rappresenta. Non è proprio corretto, come vedremo anche in seguito, perché ci potrebbe essere una funzione rappresentata da diverse matrici (al variare delle basi scelte).

(IV.23) Esempio. Consideriamo il sistema

$$\begin{cases} x_1 + x_2 = y_1 \\ \qquad x_1 = y_2, \end{cases}$$

che possiamo riscrivere come

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

La matrice in questione rappresenta quindi anche una funzione $\mathbb{R}^2 \to \mathbb{R}^2$ definita ponendo

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

■

▶ Forse non tutti sanno che questa matrice si chiama *Q–matrice di Fibonacci*, e viene indicata con la lettera Q. Ricordiamo che i numeri di Fibonacci sono la successione definita per ricorrenza da $F_2 = F_1 = 1$, e $F_{n+1} = F_n + F_{n-1}$ per ogni $n \geq 2$. Chissà chi è in grado di dimostrare che

$$Q^n = \begin{bmatrix} F_{n+1} & F_n \end{bmatrix}$$

Se A è una matrice $n \times m$, la funzione $A\colon \mathbb{R}^m \to \mathbb{R}^n$ definita come sopra da $\boldsymbol{x} \mapsto A\boldsymbol{x}$ gode delle due proprietà importanti

(5) $A(\boldsymbol{x}+\boldsymbol{y}) = A\boldsymbol{x} + A\boldsymbol{y}$

(6) $A(c\boldsymbol{x}) = c(A\boldsymbol{x})$,

che sono state enunciate nella Proposizione (IV.22). Torniamo un momento nel contesto astratto degli spazi vettoriali, e diamo un nome a queste proprietà.

(IV.24) Definizione. Se V e W sono due spazi vettoriali, una funzione $L\colon V \to W$ si dice *lineare* quando per ogni $\boldsymbol{x}, \boldsymbol{y} \in V$ e per ogni numero c si ha

$$L(\boldsymbol{x}+\boldsymbol{y}) = L(\boldsymbol{x}) + L(\boldsymbol{y}) \qquad \text{e} \qquad L(c\boldsymbol{x}) = cL(\boldsymbol{x}).$$

Se $V = W$, una funzione lineare $f\colon V \to V$ viene a volte detta *operatore* su V.

▶ Altri quasi-sinonimi di «funzione» sono: trasformazione, applicazione, operatore, mappa. Hanno accezioni leggermente diverse a seconda del contesto. A volte si usa il termine «operatore» solo per funzioni lineari da V in sé.

▶ Quando una funzione L è lineare spesso si semplifica la scrittura $L(x)$ in Lx. Infatti le proprietà formali sono proprio quelle di «moltiplicazione da sinistra» e proprietà distributiva.

Con questa nozione possiamo rivedere molti dei concetti sotto una nuova luce, quella della *linearità*. Per prima cosa, la funzione $A\colon \mathbb{R}^m \to \mathbb{R}^n$ definita in (4) (moltiplicazione a sinistra con una matrice) è una funzione lineare. Vale anche il viceversa: se V e W sono spazi vettoriali di dimensione finita allora ogni applicazione lineare da V a W, una volta scelte delle basi per V e W, si può scrivere come prodotto per una matrice, come in (4).

(IV.25) (Matrice associata) *Siano $(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m)$ una base per lo spazio vettoriale V, ed $(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)$ una base per lo spazio vettoriale W. Indichiamo con $B\colon \mathbb{R}^m \to V$ e $E\colon \mathbb{R}^n \to W$ le parametrizzazioni indotte dalle basi. Se $f\colon V \to W$ è una funzione lineare, allora esiste una matrice F con n righe e m colonne tale che, nelle coordinate delle basi sopra indicate, f è uguale alla funzione $F\colon \boldsymbol{x} \mapsto F\boldsymbol{x}$, come indicato nel seguente diagramma*

$$\begin{tikzcd} V \arrow[r, "f"] & W \\ \mathbb{R}^m \arrow[u, "B", "\cong"'] \arrow[r, "F"'] & \mathbb{R}^n. \arrow[u, "\cong", "E"'] \end{tikzcd}$$

Dimostrazione. La funzione F è ben definita, dato che $B\colon \mathbb{R}^m \to V$ è ben definita e $E\colon \mathbb{R}^n \to W$ è biunivoca (e quindi invertibile). Prendiamo un vettore $\boldsymbol{x} \in \mathbb{R}^m$. Il corrispondente vettore di V è $B(\boldsymbol{x})$, definito da

$$B(\boldsymbol{x}) = x_1\boldsymbol{b}_1 + x_2\boldsymbol{b}_2 + \ldots + x_m\boldsymbol{b}_m.$$

La sua immagine in W è, per linearità

$$f(B(\boldsymbol{x})) = x_1 f(\boldsymbol{b}_1) + x_2 f(\boldsymbol{b}_2) + \ldots + x_m f(\boldsymbol{b}_m).$$

Dato che E è biunivoca, esistono m vettori di $\mathbb{R}^n$, che chiamiamo $\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_m$, tali che $E(\boldsymbol{f}_i) = f(\boldsymbol{b}_i)$, e risulta

▶ I vettori della base standard di $\mathbb{R}^n$ sono in corrispondenza, attraverso E, con gli n vettori $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ di W.

$$f(B(\boldsymbol{x})) = x_1 E(\boldsymbol{f}_1) + x_2 E(\boldsymbol{f}_2) + \ldots + x_m E(\boldsymbol{f}_m).$$

Dato che E è lineare, anche E^{-1} (l'inversa di E) è lineare, e quindi

$$F(\boldsymbol{x}) = E^{-1}(f(B(\boldsymbol{x}))) = x_1 \boldsymbol{f}_1 + x_2 \boldsymbol{f}_2 + \ldots + x_m \boldsymbol{f}_m = F\boldsymbol{x}$$

se F è la matrice con colonne gli $\boldsymbol{f}_i$. ■

(IV.26) Definizione. La matrice, la cui esistenza è provata nella Proposizione (IV.25), si dice *matrice associata alla trasformazione lineare* f.

(IV.27) Esempio. Consideriamo la trasformazione del piano in sé data dalla rotazione di $\pi/2$, in senso antiorario, attorno all'origine. Non è difficile convincersi che è lineare e▶ perciò, per il teorema precedente, deve essere rappresentata da una matrice. Il ruotato del vettore con componenti (x, y) è il vettore $f(x, y) = (-y, x)$. Otteniamo dunque

▶ Infatti il ruotato di una somma di vettori è la somma dei vettori ruotati, mentre il ruotato di un multiplo di un vettore è il multiplo del ruotato.

$$f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -y \\ x \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = A\begin{bmatrix} x \\ y \end{bmatrix}.$$ ■

▶ Per la rotazione di un generico angolo ϑ si veda l'Esercizio 21 a pagina 226.

(IV.28) Definizione. Se L è una funzione lineare $L\colon V \to W$ (con V e W spazi vettoriali), lo spazio di tutte le soluzioni $\boldsymbol{x} \in V$ dell'equazione

$$L\boldsymbol{x} = \boldsymbol{0}$$

è detto il *nucleo* di L e indicato con Nucleo(L). Il nucleo di L è a sua volta uno spazio vettoriale, la cui dimensione è anche detta *nullità* di L. L'*immagine* di L è l'insieme di tutti i vettori del tipo $L\boldsymbol{x}$, con $\boldsymbol{x} \in V$ (è l'*immagine* della funzione $L\colon V \to W$).

▶ Sinonimi di nucleo sono *kernel* (indicato con $\ker L$), e *nullspace*.

▶ L'equazione $L\boldsymbol{x} = \boldsymbol{0}$ non è altro che un sistema di n equazioni in m incognite, se L è una matrice che rappresenta una trasformazione lineare $L\colon \mathbb{R}^n \to \mathbb{R}^m$. Per quanto visto sopra, ogni funzione lineare può essere rappresentata in questo modo. In generale, perché la definizione di nullità, in (IV.28), sia ben posta occorre che il nucleo di un operatore lineare sia uno spazio vettoriale, in modo da poter usare la definizione (IV.15). È un simpatico esercizio, lasciato al Lettore (Esercizio 9 a pagina 225).

Nella prossima proposizione dimostriamo il fatto importante che la composizione di due applicazioni lineari, scritta in coordinate, non è altro che il prodotto righe per colonne di due matrici, definito come in (3).

(IV.29) (Prodotto righe per colonne) *Consideriamo due trasformazioni lineari*

$$f\colon V \to W \qquad e \qquad g\colon W \to Y.$$

Siano $(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m)$ *una base per lo spazio vettoriale* V, $(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)$ *una base per lo spazio vettoriale* W *e* $(\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_l)$ *una base per lo spazio* *vettoriale Y. Indichiamo con $B\colon \mathbb{R}^m \to V$, $E\colon \mathbb{R}^n \to W$ e $H\colon \mathbb{R}^l \to Y$ le parametrizzazioni indotte dalle basi. Sia F la matrice con n righe e m colonne che rappresenta f, e G la matrice con m righe e l colonne che rappresenta g (nelle coordinate delle basi sopra indicate). Allora la matrice che rappresenta la composizione $g \circ f$*

$$\begin{tikzcd} V \arrow[r, "f"'] \arrow[rr, bend left, "g \circ f"] & W \arrow[r, "g"'] & Y \end{tikzcd}$$

è la matrice prodotto GF.

▶ Possiamo rappresentare la proposizione nel seguente diagramma, in modo molto sintetico:

$$\begin{tikzcd} V \arrow[r, "f"'] \arrow[rr, bend left, "g \circ f"] & W \arrow[r, "g"'] & Y \\ \mathbb{R}^m \arrow[u, "B", "\cong"'] \arrow[r, "F"] \arrow[rr, bend right, "GF"'] & \mathbb{R}^n \arrow[u, "\cong", "E"'] \arrow[r, "G"] & \mathbb{R}^l \arrow[u, "\cong", "H"'] \end{tikzcd}$$

Dimostrazione. Vediamo dal diagramma a fianco che la matrice associata alla composizione $g \circ f$ non è altro che l'applicazione lineare ottenuta componendo G con F. Dobbiamo perciò dimostrare che la composizione delle due funzioni lineari G e F è il prodotto delle due matrici. Ma se $\boldsymbol{x} \in \mathbb{R}^m$, si ha $F(\boldsymbol{x}) = x_1\boldsymbol{f}_1 + x_2\boldsymbol{f}_2 + \ldots + x_m\boldsymbol{f}_m$, e quindi per linearità

$$\begin{aligned} G(F(\boldsymbol{x})) &= x_1G(\boldsymbol{f}_1) + x_2G(\boldsymbol{f}_2) + \ldots + x_mG(\boldsymbol{f}_m) \\ &= (G\boldsymbol{f}_1, G\boldsymbol{f}_2, \ldots, G\boldsymbol{f}_m)\boldsymbol{x} = (GF)\boldsymbol{x}. \end{aligned}$$ ■

▶ Il prodotto di matrici corrisponde dunque alla *composizione* dei due operatori (funzioni lineari) corrisponenti. Il fatto, già visto, che in generale $AB \neq BA$ corrisponde al fatto, molto più generale, che la *composizione* di funzioni non è una operazione commutativa: fermare la macchina e poi scendere non è sempre equivalente a scendere dalla macchina in corsa e dopo (cercare di) fermarla: basta fare una prova in autostrada per convincersene.

Un caso particolare si ha quando si rappresenta la funzione identica $1_V\colon V \to V$ (quella che manda $\boldsymbol{v} \in V$ in $\boldsymbol{v} \in V$) in due basi diverse:

$$\begin{tikzcd} V \arrow[r, "1_V"] & V \\ \mathbb{R}^n \arrow[u, "B", "\cong"'] \arrow[r, "S"'] & \mathbb{R}^n. \arrow[u, "\cong", "E"'] \end{tikzcd}$$

La matrice S è una matrice quadrata, detta matrice del *cambio di base*, o del *cambio di coordinate*. Se le due basi coincidono $B = E$, allora la matrice è quella identica.▶ Le sue colonne sono i vettori $E^{-1}(\boldsymbol{b}_i)$, cioè sono le *componenti dei vettori $\boldsymbol{b}_i$ scritti nella base $\boldsymbol{e}_j$* (quella della parametrizzazione E). Ovviamente potremmo pensare di scambiare il ruolo delle due basi, ottenendo di fatto la funzione inversa di S, a sua volta rappresentata da una matrice.

▶ Quella dell'Esempio (IV.19).

(IV.30) Definizione. Una matrice quadrata A si dice *invertibile* se esiste una matrice A^{-1} (l'*inversa* di A) tale che

(7) $\quad A^{-1}A = AA^{-1} = I.$

▶ Non è difficile mostrare che se esiste una inversa per la matrice A allora è unica. Infatti, se B e C sono entrambe inverse di A, cioè $BA = AB = I$ e $CA = AC = I$, allora $B = BI = B(AC) = (BA)C = IC = C$.

Ne segue che se S è la matrice di cambiamento di base da B ad E allora S^{-1} è quella del cambiamento da E a B, e $S^{-1}S = SS^{-1} = I$.

(IV.31) Esempio. Sia $\boldsymbol{x} = (x_1, x_2)$ un vettore di $\mathbb{R}^2$. Siano $(\hat{x}_1, \hat{x}_2)$ le componenti di $\boldsymbol{x}$ rispetto alla base $\boldsymbol{s}_1 = (1,0)$, $\boldsymbol{s}_2 = (1,1)$ La relazione tra le coordinate è semplicemente la

▶ Cfr. anche l'Esempio (IV.11) a pag. 170.

$$\boldsymbol{x} = \hat{x}_1\boldsymbol{s}_1 + \hat{x}_2\boldsymbol{s}_2 \iff \begin{cases} x_1 = \hat{x}_1 + \hat{x}_2 \\ x_2 = \hat{x}_2. \end{cases} \iff \boldsymbol{x} = S\hat{\boldsymbol{x}},$$

dove $S = (\boldsymbol{s}_1, \boldsymbol{s}_2) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Per calcolare la matrice inversa S^{-1} occorre semplicemente esprimere le $\hat{\boldsymbol{x}}$ in funzione di $\boldsymbol{x}$:

▶ Alternativamente, si può risolvere il sistema di quattro equazioni nelle quattro incognite a, b, c, d

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & a+b \\ c & c+d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

e ottenere $a = 1$, $b = -1$, $c = 0$, $d = 1$.

$$\begin{cases} x_1 = \hat{x}_1 + \hat{x}_2 \\ x_2 = \hat{x}_2. \end{cases} \iff \begin{cases} \hat{x}_1 = x_1 - x_2 \\ \hat{x}_2 = x_2, \end{cases} \iff \hat{\boldsymbol{x}} = S^{-1}\boldsymbol{x}$$

e quindi la matrice inversa S^{-1} è $S^{-1} = \begin{bmatrix} 1 & -1 \\ 0,1 & \end{bmatrix}$. ■

Vediamo ora come cambiano le matrici associate ad una trasformazione lineare, se si cambia la base di riferimento.

(IV.32) Definizione. Si dice che due matrici quadrate $n \times n$ A e B sono *simili* se esiste una matrice invertibile S tale che

$$S^{-1}AS = B.$$

Due matrici sono simili se rappresentano il medesimo operatore in due basi diverse. Infatti, se $A\colon \mathbb{R}^n \to \mathbb{R}^n$ è l'operatore lineare dato dalla matrice A (rispetto ad una certa base), ed $S\colon \mathbb{R}^n \to \mathbb{R}^n$ è la parametrizzazione indotta da un'altra base, allora $S^{-1}AS = B$, come si evince dal seguente diagramma (si ricordi che la composizione di operatori corrisponde al prodotto di matrici):

▶ In coordinate: se $\boldsymbol{x}$, $\boldsymbol{y}$ rappresentano le componenti di due vettori rispetto alla prima base e $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$ quelle rispetto alla seconda, allora

$$\boldsymbol{x} = S\hat{\boldsymbol{x}}, \quad \boldsymbol{y} = S\hat{\boldsymbol{y}}.$$

Quindi

$$\boldsymbol{y} = A\boldsymbol{x} \iff S\hat{\boldsymbol{y}} = AS\hat{\boldsymbol{x}} \\ \iff \hat{\boldsymbol{y}} = S^{-1}AS\hat{\boldsymbol{x}}$$

(nell'ultimo passaggio abbiamo moltiplicato a sinistra per S^{-1}). Quindi la matrice che rappresenta la trasformazione lineare nelle seconde coordinate è $B = S^{-1}AS$.

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^n \\ \mathbb{R}^n \arrow[u, "S", "\cong"'] \arrow[r, "B"'] & \mathbb{R}^n. \arrow[u, "S"', "\cong"] \end{tikzcd}$$

(IV.33) Esempio. Consideriamo l'operatore $Q\colon \mathbb{R}^2 \to \mathbb{R}^2$ dell'Esempio (IV.23), definito da

$$(*) \quad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Rispetto alla base $\boldsymbol{s}_1 = (1,0)$, $\boldsymbol{s}_2 = (1,1)$▶ l'operatore Q si scrive

▶ La matrice associata alla base è $S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, con inversa $S^{-1} = \begin{bmatrix} 1 & -1 \\ 0,1 & \end{bmatrix}$ (cfr. Esempio (IV.31) a pag. 180).

$$\hat{Q} = S^{-1}QS = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \ldots = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

Alternativamente, direttamente nelle coordinate, da

$$\begin{cases} \hat{x}_1 + \hat{x}_2 = x_1 \\ \hat{x}_2 = x_2 \end{cases} \quad \text{e} \quad \begin{cases} \hat{y}_1 + \hat{y}_2 = y_1 \\ \hat{y}_2 = y_2, \end{cases}$$

sostituendo nella (*) otteniamo

$$\begin{cases} \hat{y}_1 + \hat{y}_2 = \hat{x}_1 + \hat{x}_2 + \hat{x}_2 \\ \hat{y}_2 = \hat{x}_1 + \hat{x}_2 \end{cases} \iff \begin{cases} \hat{y}_1 = \hat{x}_2 \\ \hat{y}_2 = \hat{x}_1 + \hat{x}_2 \end{cases} \iff \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}.$$

Com'era da aspettarsi, la forma di $\hat{Q}$ è la stessa. ■

§ 3. Prodotto scalare e isometrie

(IV.34) Definizione (Prodotto scalare). Un *prodotto scalare* definito su uno spazio vettoriale V è un prodotto che associa ad ogni coppia di vettori $\boldsymbol{x}, \boldsymbol{y} \in V$ un numero reale, indicato con $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ (oppure, equivalentemente, con $\boldsymbol{x} \cdot \boldsymbol{y}$) con le seguenti proprietà: se $\boldsymbol{x}$, $\boldsymbol{y}$ e $\boldsymbol{z}$ sono tre vettori di V e $c \in \mathbb{R}$ è uno scalare,

① (commutativo/simmetrico) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$;

② (distributivo/lineare) $\langle \boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{x}, \boldsymbol{z} \rangle$;

③ (omogeneo) $\langle c\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, c\boldsymbol{y} \rangle = c\langle \boldsymbol{x}, \boldsymbol{y} \rangle$;

④ (definito positivo) $\boldsymbol{x} \neq \boldsymbol{0} \implies \langle \boldsymbol{x}, \boldsymbol{x} \rangle > 0$.

La radice quadrata del prodotto positivo $\langle \boldsymbol{x}, \boldsymbol{x} \rangle$ è detta la *norma* del vettore $\boldsymbol{x} \in V$, e viene indicata con il simbolo

$$||\boldsymbol{x}|| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}.$$

▶ Il prodotto scalare può essere definito anche per spazi vettoriali complessi, come vedremo nel riquadro a pagina 215

▶ Questa è la linearità rispetto al secondo fattore. Dalla ① segue anche quella rispetto al primo.

▶ Mentre, per linearità, $\langle \boldsymbol{0}, \boldsymbol{0} \rangle = 0$.

(IV.35) Esempio. In $\mathbb{R}^n$ si può definire il prodotto scalare, ponendo

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x} \cdot \boldsymbol{y} = \sum_{i=1}^{n} x_i y_i.$$

Viene chiamato *prodotto scalare standard* in $\mathbb{R}^n$. ■

▶ Cfr. la Definizione (III.16) a pagina 111

(IV.36) Esempio. Nello spazio vettoriale $C[a, b]$ di tutte le funzioni continue sull'intervallo $[a, b]$ il prodotto

$$\langle f, g \rangle = \int_a^b f(x)g(x)\,\mathrm{d}x$$

è un prodotto scalare. ■

Le prossime proprietà del prodotto scalare sono generalizzazioni delle corrispondenti (omonime) proprietà del prodotto scalare standard di $\mathbb{R}^n$.

▶ Le abbiamo viste nella Proposizione (III.19) a pagina 112

(IV.37) (Disuguaglianza di Cauchy–Schwarz) *Ogni prodotto scalare soddisfa la seguente disuguaglianza: per ogni x, y di V*

$$\langle x, y\rangle \leq \|x\|\|y\|,$$

e vale l'uguaglianza se e solo se x e y sono paralleli (dipendenti).

Dimostrazione. Per la proprietà ④ del prodotto scalare (Definizione (IV.34)) si ha che $\langle z, z\rangle \geq 0$ per ogni z e se $z \neq \mathbf{0}$ si ha $\langle z, z\rangle > 0$. Se $y = \mathbf{0}$, allora $(\langle x, \mathbf{0}\rangle)^2 = 0 \leq 0 = \|x\|^2\|\mathbf{0}\|^2$, quindi la disuguaglianza è vera. Se invece $y \neq \mathbf{0}$, si ha che per ogni $t \in \mathbb{R}$ deve essere

$$0 \leq \|x + ty\|^2 = \langle x + ty, x + ty\rangle = \|x\|^2 + t^2\|y\|^2 + 2t\langle x, y\rangle.$$

Ora, la funzione $f(t) = \|x\|^2 + t^2\|y\|^2 + 2t\langle x, y\rangle$ ha per grafico una parabola con la concavità verso l'alto (dato che il coefficiente di t^2 è positivo) ed è per definizione una funzione positiva; ha il minimo in quel t_0 tale che $f'(t_0) = 0$, cioè

$$f'(t_0) = 2t_0\|y\|^2 + 2\langle x, y\rangle = 0 \iff t_0 = -\frac{\langle x, y\rangle}{\|y\|^2}.$$

Il valore di $f(t_0)$ è

$$f(t_0) = \|x\|^2 + \left[-\frac{\langle x, y\rangle}{\|y\|^2}\right]^2 \|y\|^2 - 2\left[\frac{\langle x, y\rangle}{\|y\|^2}\right]\langle x, y\rangle = \|x\|^2 - \frac{(\langle x, y\rangle)^2}{\|y\|^2}.$$

Dato che deve essere $f(t_0) \geq 0$, si ha

$$\frac{(\langle x, y\rangle)^2}{\|y\|^2} \leq \|x\|^2,$$

da cui segue la tesi. Vale l'uguaglianza solo quando $f(t_0) = 0$, cioè quando

$$0 = f(t_0) = \|x + t_0 y\|^2 \implies x = -t_0 y,$$

cioè x e y sono paralleli. ■

Le proprietà della norma seguono facilmente da quelle del prodotto scalare, come nella dimostrazione della seconda parte della Proposizione (III.19).

(IV.38) (Proprietà della norma) *Valgono le seguenti proprietà della norma: per ogni x, y, z in V e per ogni scalare $c \in \mathbb{R}$,*

① *(positività)* $x \neq 0 \implies \|x\| > 0$, $\|x\| = 0 \iff x = \mathbf{0}$.

② *(omogeneità)* $\|cx\| = |c|\,\|x\|$.

③ *(disuguaglianza triangolare)* $\|x + y\| \leq \|x\| + \|y\|$.

Il termine *norma* in matematica ha un significato ben preciso, che puntualizziamo con la definizione seguente.

(IV.39) Definizione. Una *norma* su uno spazio vettoriale V è una funzione da V a $\mathbb{R}$ tale che valgono le proprietà ①, ② e ③ della proposizione (IV.38).

▶ Come vedremo in seguito, non tutte le norme sono definite a partire da un prodotto scalare. Si veda in proposito anche il Capitolo VII, il riquadro a pagina 342.

La Proposizione (IV.38) può quindi essere sintetizzata nell'affermazione *"la funzione $\|v\| = \sqrt{v \cdot v}$ è una norma"*.

Torniamo al prodotto scalare. Una volta che si è definito, è possibile parlare di *ortogonalità*.

▶ Esattamente come è stato fatto nella Proposizione (III.18) a pagina 111.

(IV.40) Definizione. Due vettori v e w sono *ortogonali* (rispetto al prodotto scalare in oggetto) quando $\langle v, w \rangle = 0$.

(IV.41) *Se k vettori v_1, v_2, ..., v_k non nulli di V sono ortogonali, allora sono linearmente indipendenti.*

Dimostrazione. È analoga a quella illustrata nell'Esercizio (III.20) a pagina 158 per tre vettori. Lasciamo al Lettore i dettagli. ■

(IV.42) (Teorema di Pitagora) *Se v e w sono due vettori ortogonali, allora*

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

Dimostrazione. Se $\langle v, w \rangle = 0$, si ha

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle \\ &= \|v\|^2 + \|w\|^2. \end{aligned}$$ ■

L'aggettivo "ortogonale" si usa anche rispetto alle basi e alle matrici.

(IV.43) Definizione. Diciamo che una base è *ortogonale* se consiste di vettori tutti ortogonali tra di loro. Diciamo che è *ortonormale* se è ortogonale, e in più tutti i vettori sono unitari.
Diciamo che una *matrice* è *ortogonale* se le sue colonne costituiscono una base ortogonale (cioè sono tutte ortogonali tra di loro) e sono vettori unitari.

▶ Cioè hanno norma uno (si dice anche che sono dei *versori*).

▶ In altre parole, una matrice è *ortogonale* se e solo se i suoi i vettori colonna costituiscono una base *ortonormale*.

La Definizione (III.50) di proiezione ortogonale su una retta o un piano può essere data anche per sottospazi vettoriali. Diciamo che un vettore v è *ortogonale* ad un sottospazio W di V se per ogni $w \in W$ i due vettori v e w sono ortogonali, cioè se

per ogni $w \in W$, $\langle v, w \rangle = 0$.

Osserviamo che se W è generato da k vettori e_1, e_2, ..., e_k, allora un vettore v è ortogonale a W se e solo se è ortogonale a tutti i k vettori e_i ▶.

▶ Infatti, se $w \in W$ e $\langle v, e_i \rangle = 0$ per tutti gli i, si ha

$$\begin{aligned} w &= c_1 e_1 + c_2 e_2 + \ldots + c_k e_k \\ \langle v, w \rangle &= \langle v, c_1 e_1 + c_2 e_2 + \ldots + c_k e_k \rangle \\ &= c_1 \langle v, e_1 \rangle + c_2 \langle v, e_2 \rangle + \ldots + c_k \langle v, e_k \rangle \\ &= 0. \end{aligned}$$

(IV.44) Definizione. Se $W \subset V$ è un sottospazio vettoriale di V, allora la *proiezione* ortogonale di un vettore $v \in V$ su W, indicata con $\operatorname{proj}_W v$, è quell'unico vettore $\bar{v} \in W$ tale che $v - \bar{v}$ è ortogonale a W.

Perché questa definizione sia ben posta, occorre però dimostrare che il vettore $\bar{v} = \mathrm{proj}_W v$ esiste ed è unico. Cominciamo a mostrare che se il sottospazio W ha una base ortogonale, allora il vettore $\bar{v}$, proiezione di v su W, esiste davvero.

▶ Ricordiamo come con l'equazione (14) a pagina 113 abbiamo definito la proiezione di un vettore su un altro vettore (meglio: sulla retta da lui generata). Si veda l'Esercizio (III.14) a pagina 154.

(IV.45) Teorema. *Se $e_1, e_2, \ldots, e_k$ sono k vettori ortogonali di uno spazio vettoriale V, e $v \in V$, allora esiste una unica* proiezione *di v sul sottospazio $W = \mathrm{Span}(e_1, e_2, \ldots, e_k)$, che è uguale a*

$$\text{(8)} \quad \mathrm{proj}_W(v) = \sum_{i=1}^{k} c_i e_i = \sum_{i=1}^{k} \frac{\langle v, e_i \rangle}{\|e_i\|^2} e_i.$$

Il vettore $\bar{v} = \mathrm{proj}_W(v)$ è quello che minimizza la distanza tra v e W, cioè per ogni $w \in W$

$$\|v - w\| \geq \|v - \bar{v}\|,$$

e vale l'uguaglianza solo se $w = \bar{v}$.

▶ La proiezione quindi è una funzione $\mathrm{proj}_W \colon V \to W \subset V$. Non è difficile mostrare che è una funzione lineare (si veda l'Esercizio 26 a pagina 226).

Dimostrazione. Gli elementi di W sono le combinazioni lineari

$$w = c_1 e_1 + c_2 e_2 + \ldots + c_k e_k,$$

con $c_i \in \mathbb{R}$ per ogni i. Occorre determinare, se esistono, i valori della k-upla $(c_1, c_2, \ldots, c_k)$ per cui il corrispondente w soddisfi l'equazione

$$\langle v - w, e_i \rangle = 0 \quad \iff \quad \langle v, e_i \rangle = \langle w, e_i \rangle$$

per ogni $i = 1, \ldots, k$. Ma

$$\begin{aligned} \langle w, e_i \rangle &= \langle c_1 e_1 + c_2 e_2 + \ldots + c_k e_k, e_i \rangle \\ &= c_1 \langle e_1, e_i \rangle + c_2 \langle e_2, e_i \rangle + \ldots + c_k \langle e_k, e_i \rangle = c_i \langle e_i, e_i \rangle = c_i \|e_i\|^2, \end{aligned}$$

e quindi l'unica soluzione è quella per cui i coefficienti c_i sono definiti da

$$\langle v, e_i \rangle = c_i \|e_i\|^2 \quad \Longrightarrow \quad c_i = \frac{\langle v, e_i \rangle}{\|e_i\|^2}.$$

Non resta che mostrare che il vettore $\bar{v} = \mathrm{proj}_W(v)$ realizza la distanza minima tra v e W, cioè che se $\bar{w} \in W$ è un vettore diverso da $\bar{v}$ (e quindi con $\|\bar{v} - \bar{w}\| > 0$), allora la distanza tra $\bar{w}$ e v è strettamente maggiore di quella tra $\bar{v}$ e v. Ma

$$v - \bar{w} = (v - \bar{v}) + (\bar{v} - \bar{w}),$$

dove $v - \bar{v}$ è ortogonale a W (per definizione) e $\bar{v} - \bar{w}$ è un vettore di W (perché somma di due vettori in W). Ma allora per la Proposizione (IV.42)

$$\|v - \bar{w}\| = \|v - \bar{v}\|^2 + \|\bar{v} - \bar{w}\|^2 > \|v - \bar{v}\|^2.$$

Quod erat demonstrandum. ■

Figura IV.4: David Hilbert (1862–1943), figura di spicco della matematica a cavallo tra l'Ottocento e il Novecento.

Cosa succede se si calcola la proiezione su un sottospazio W di un vettore $\boldsymbol{v}$ che già appartiene a W?

(IV.46) Corollario. *Se, nelle ipotesi del Teorema (IV.45), il vettore $\boldsymbol{v}$ appartiene già a $W = \operatorname{Span}(\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_k)$, allora $\operatorname{proj}_W(\boldsymbol{v}) = \boldsymbol{v}$, e i coefficienti $c_i = \dfrac{\langle \boldsymbol{v}, \boldsymbol{e}_i \rangle}{\|\boldsymbol{e}_i\|^2}$ sono le componenti di $\boldsymbol{v}$ nella base $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_k$ di W. Se inoltre $\|\boldsymbol{e}_i\| = 1$ per ogni i (cioè se $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_k$ è una base ortonormale per W), allora*

$$\boldsymbol{v} = \sum_{i=1}^{k} \langle \boldsymbol{v}, \boldsymbol{e}_i \rangle \boldsymbol{e}_i \quad e \quad \|\boldsymbol{v}\|^2 = \sum_{i=1}^{k} c_i^2 = \sum_{i=1}^{k} (\langle \boldsymbol{v}, \boldsymbol{e}_i \rangle)^2.$$

Dimostrazione. Se $\boldsymbol{v} \in W$, allora se $\bar{\boldsymbol{v}} \in W$ è un vettore tale che $\boldsymbol{v} - \bar{\boldsymbol{v}}$ è ortogonale a W (cioè ortogonale a tutti i vettori di W), allora $\boldsymbol{v} - \bar{\boldsymbol{v}}$ dovrebbe essere ortogonale anche a sé stesso, cioè avere norma zero. Ma questo è possibile solo se $\boldsymbol{v} = \bar{\boldsymbol{v}}$, cioè

$$\boldsymbol{v} \in W \implies \operatorname{proj}_W(\boldsymbol{v}) = \boldsymbol{v}.$$

Possiamo comunque applicare il Teorema (IV.45), ed ottenere

$$\operatorname{proj}_W(\boldsymbol{v}) = \boldsymbol{v} = \sum_{i=1}^{k} \frac{\langle \boldsymbol{v}, \boldsymbol{e}_i \rangle}{\|\boldsymbol{e}_i\|^2} \boldsymbol{e}_i.$$

Se aggiungiamo l'ipotesi che $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_k$ è una base ortonormale per W, e poniamo $c_i = \langle \boldsymbol{v}, \boldsymbol{e}_i \rangle$, si ha

$$\|\boldsymbol{v}\|^2 = \left\langle \sum_{i=1}^{k} c_i \boldsymbol{e}_i, \sum_{j=1}^{k} c_j \boldsymbol{e}_j \right\rangle = \sum_{i,j=1}^{k} c_i c_j \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = \sum_{i=1}^{k} c_i^2. \qquad \blacksquare$$

Quando si scrive un vettore $\boldsymbol{v}$ in una base di vettori ortogonali, le componenti del vettore $\boldsymbol{v}$ in questa base sono quindi proprio le *proiezioni* di $\boldsymbol{v}$ sugli assi, così come si fa per determinare le coordinate cartesiane di un punto di $\mathbb{R}^n$.

Ortogonalizzazione di Gram–Schmidt

Supponiamo di avere un insieme $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_k$ di vettori linearmente indipendenti di uno spazio vettoriale V (su cui sia definito un prodotto scalare). Dato che i $\boldsymbol{v}_i$ sono linearmente indipendenti, non sono nulli e quindi $\|\boldsymbol{v}_i\| > 0$ per ogni i.

Definiamo per ricorsione una successione $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_k$ di vettori ortogonali tra di loro nel modo seguente. Per $i = 1$ si pone

$$\boldsymbol{u}_1 = \boldsymbol{v}_1.$$

▶ I due matematici cui si fa riferimento sono Jorgen Pedersen Gram (1850–1916) e Erhard Schmidt (1876–1959), anche se l'algoritmo di ortogonalizzazione è stato certamente usato da altri prima di loro (Cauchy, Laplace).

Successivamente, per $i > 1$ si definisce prima la proiezione di v_i sul sottospazio $\mathrm{Span}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{i-1}) \subset V$, che come abbiamo visto sopra si scrive

▶ Posto che gli $\boldsymbol{u}_i$ siano ortogonali, ma questo lo vedremo subito.

$$\bar{v}_i = \sum_{j=1}^{i-1} \frac{\langle v_i, \boldsymbol{u}_j \rangle}{\|\boldsymbol{u}_j\|^2} \boldsymbol{u}_j,$$

da cui definiamo finalmente il vettore $\boldsymbol{u}_i$

$$\boldsymbol{u}_i = v_i - \bar{v}_i = v_i - \sum_{j=1}^{i-1} \frac{\langle v_i, \boldsymbol{u}_j \rangle}{\|\boldsymbol{u}_j\|^2} \boldsymbol{u}_j.$$

Per definizione di proiezione ortogonale, $\boldsymbol{u}_i$ risulta ortogonale a $\boldsymbol{u}_j$ per $j = 1, \dots, i-1$, dato che è ortogonale a $\mathrm{Span}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{i-1})$. Segue che con questa costruzione tutti gli $\boldsymbol{u}_j$ sono ortogonali tra di loro. Si ottiene quindi un sistema triangolare del tipo

$$\begin{aligned}
\boldsymbol{u}_1 &= v_1 \\
\boldsymbol{u}_2 &= v_2 - \frac{\langle v_2, \boldsymbol{u}_1 \rangle}{\|\boldsymbol{u}_1\|^2} \boldsymbol{u}_1 \\
\boldsymbol{u}_3 &= v_3 - \frac{\langle v_3, \boldsymbol{u}_1 \rangle}{\|\boldsymbol{u}_1\|^2} \boldsymbol{u}_1 - \frac{\langle v_3, \boldsymbol{u}_2 \rangle}{\|\boldsymbol{u}_2\|^2} \boldsymbol{u}_2 \\
&\vdots \\
\boldsymbol{u}_k &= v_k - \frac{\langle v_k, \boldsymbol{u}_1 \rangle}{\|\boldsymbol{u}_1\|^2} \boldsymbol{u}_1 - \frac{\langle v_k, \boldsymbol{u}_2 \rangle}{\|\boldsymbol{u}_2\|^2} \boldsymbol{u}_2 - \dots - \frac{\langle v_k, \boldsymbol{u}_{k-1} \rangle}{\|\boldsymbol{u}_{k-1}\|^2} \boldsymbol{u}_{k-1}.
\end{aligned}$$

Si può naturalmente scrivere anche come

$$\begin{aligned}
v_1 &= \boldsymbol{u}_1 \\
v_2 &= \frac{\langle v_2, \boldsymbol{u}_1 \rangle}{\|\boldsymbol{u}_1\|^2} \boldsymbol{u}_1 + \boldsymbol{u}_2 \\
v_3 &= \frac{\langle v_3, \boldsymbol{u}_1 \rangle}{\|\boldsymbol{u}_1\|^2} \boldsymbol{u}_1 + \frac{\langle v_3, \boldsymbol{u}_2 \rangle}{\|\boldsymbol{u}_2\|^2} \boldsymbol{u}_2 + \boldsymbol{u}_3 \\
&\vdots \\
v_k &= \frac{\langle v_k, \boldsymbol{u}_1 \rangle}{\|\boldsymbol{u}_1\|^2} \boldsymbol{u}_1 + \frac{\langle v_k, \boldsymbol{u}_2 \rangle}{\|\boldsymbol{u}_2\|^2} \boldsymbol{u}_2 + \dots + \frac{\langle v_k, \boldsymbol{u}_{k-1} \rangle}{\|\boldsymbol{u}_{k-1}\|^2} \boldsymbol{u}_{k-1} + \boldsymbol{u}_k
\end{aligned}$$

Supponiamo che V sia lo spazio Euclideo standard $\mathbb{R}^n$. Nella notazione matriciale, se chiamiamo M la matrice che ha per colonne i vettori v_i, si

▶ È possibile modificare leggermente la successione dei vettori $\boldsymbol{u}_i$ ponendo

$$q_i = \frac{\boldsymbol{u}_i}{\|\boldsymbol{u}_i\|}$$

per ogni i: in questo modo i vettori q_i hanno norma uno e sono ortogonali. La corrispondente decomposizione della matrice M che si ottiene è nota come la *decomposizione QR*, in cui Q è una matrice ortogonale (cioè con colonne ortogonali e può scrivere

$$M = (u_1, u_2, \dots, u_k) \begin{bmatrix} 1 & \frac{\langle v_2, u_1 \rangle}{\|u_1\|^2} & \frac{\langle v_3, u_1 \rangle}{\|u_1\|^2} & \dots & \frac{\langle v_k, u_1 \rangle}{\|u_1\|^2} \\ 0 & 1 & \frac{\langle v_3, u_2 \rangle}{\|u_2\|^2} & \dots & \frac{\langle v_k, u_2 \rangle}{\|u_2\|^2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & \frac{\langle v_k, u_{k-1} \rangle}{\|u_{k-1}\|^2} \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}$$

cioè $M = PR$, dove $P = (u_1, u_2, \dots, u_k)$ è una matrice di n righe e k colonne (ortogonali), e R è una matrice $k \times k$ triangolare superiore (con tutti 1 sulla diagonale).

Un corollario immediato del metodo di ortogonalizzazione è che ogni spazio vettoriale di dimensione finita ha almeno una base ortogonale▶.

▶ E quindi, normalizzando, una base ortonormale.

(IV.47) *Se V è uno spazio vettoriale di dimensione finita n e con prodotto scalare, allora esiste una base ortonormale e_1, e_2, ..., e_n di V.*

Dimostrazione. Se v_1, v_2, ..., v_n è una base per V, applicando il procedimento di ortogonalizzazione si ottiene una base ortogonale u_1, u_2, ..., u_n tale che per ogni $k = 1 \dots n$

$$\operatorname{Span}(u_1, \dots, u_k) = \operatorname{Span}(v_1, \dots, v_k).$$

In particolare, per $k = n$ si ottiene una base ortogonale per V. La base ortonormale cercata si ottiene *normalizzando* i vettori, cioè ponendo $e_i = u_i / \|u_i\|$ in modo tale che gli e_i non sono solo ortogonali tra di loro, ma anche unitari. ■

(IV.48) Corollario. *Se $W \subset V$ è un sottospazio vettoriale di dimensione finita di uno spazio vettoriale V con prodotto scalare, allora esiste la proiezione ortogonale di V su W, $P_W : V \to W$, definita come nell'equazione (8) a pagina 184.*

Dimostrazione. Il Teorema (IV.45) assicura che esiste la proiezione

$$\operatorname{proj}_W : V \to W,$$

a patto che W abbia una base ortogonale. Ma per la Proposizione (IV.47) ogni spazio di dimensione finita ha una base ortogonale. ■

Possiamo finalmente completare ed estendere la dimostrazione della Proposizione (III.49) a pagina 137.

(IV.49) Corollario. *Sia p un sottospazio lineare affine dello spazio euclideo cartesiano $\mathbb{R}^n$ (con il prodotto scalare standard) e Q un punto qualsiasi. Allora esiste unico il punto Q' che realizza la distanza tra punto Q e p (cioè* $\operatorname{distanza}(Q, p) = |Q' - Q|$*), che chiamiamo ancora la* proiezione *ortogonale di Q su p.*

Dimostrazione. Il sottospazio p di $\mathbb{R}^n$ per definizione si scrive come

$$p = \{A + t_1 v_1 + t_2 v_2 + \ldots + t_d v_d : t_1, t_2, \ldots, t_d \in \mathbb{R}\},$$

per un certo punto A e d vettori linearmente indipendenti. Se indichiamo con

$$W = \mathrm{Span}(v_1, \ldots, v_d),$$

notiamo che un punto $P \in \mathbb{R}^n$ appartiene a p se e solo se $P - A \in W$. Quindi un punto Q' in p si scrive come $Q' = A + w'$ per un certo $w' \in W$, e il vettore $Q' - Q$ è ortogonale a W se e solo se $A + w' - Q$ è ortogonale a W, ma questo accade se e solo se

$$\mathrm{proj}_W(Q - A) = w'.$$

Quindi esiste la proiezione ortogonale di Q su p, che è definita da

$$Q' = A + \mathrm{proj}_W(\overrightarrow{AQ}).$$ ■

(IV.50) Esempio. Nello spazio vettoriale $V = C(0, 2\pi)$ delle funzioni continue sull'intervallo $(0, 2\pi)$, con il prodotto scalare dell'Esempio (IV.36)

$$\langle f, g\rangle = \int_0^{2\pi} f(x)g(x)\,\mathrm{d}x,$$

le funzioni trigonometriche

$$(*) \qquad 1, \cos x, \sin x, \cos 2x, \sin 2x, \cos 3x, \sin 3x, \ldots$$

sono elementi (vettori) di V.

Come abbiamo visto nella la Proposizione (II.27), se $n \neq m$

$$\langle \cos nx, \cos mx\rangle = \int_0^{2\pi} \cos nx \cos mx\,\mathrm{d}x = 0$$
$$\langle \cos nx, \sin mx\rangle = \int_0^{2\pi} \cos nx \sin mx\,\mathrm{d}x = 0$$
$$\langle \sin nx, \sin mx\rangle = \int_0^{2\pi} \sin nx \sin mx\,\mathrm{d}x = 0.$$

▶ Forse adesso risulta chiaro perché sono chiamate *formule di ortogonalità*, o *relazioni di ortogonalità*. Gli estremi di integrazione sono assolutamente ininfluenti, dato che

$$\int_{-\pi}^{\pi} p(x)\,\mathrm{d}x = \int_0^{2\pi} p(x)\,\mathrm{d}x$$

se $p(x)$ è periodica di periodo 2π.

Inoltre per ogni $m \geq 1$

$$\langle 1, \cos mx\rangle = \int_0^{2\pi} \cos mx\,\mathrm{d}x = 0 = \langle 1, \sin mx\rangle = \int_0^{2\pi} \sin mx\,\mathrm{d}x,$$

quindi possiamo dire che le funzioni (*) costituiscono un insieme ortogonale. Se ne calcoliamo le norme abbiamo

▶ Quindi questa norma è la stessa con cui avevamo introdotto lo scarto quadratico medio, approfondimento a pagina 63.

$$\|1\|^2 = \langle 1, 1\rangle = \int_0^{2\pi} \mathrm{d}x = 2\pi.$$
$$\|\cos nx\|^2 = \langle \cos nx, \cos nx\rangle = \int_0^{2\pi} \cos^2 nx\,\mathrm{d}x = \pi$$
$$\|\sin nx\|^2 = \langle \sin nx, \sin nx\rangle = \int_0^{2\pi} \sin^2 nx\,\mathrm{d}x = \pi.$$

Quindi le funzioni (trigonometriche) normalizzate

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}}, \quad n = 1, 2, 3, \ldots$$

sono vettori unitari di V, e costituiscono un insieme ortonormale di vettori di V. ■

(IV.51) Esempio (Serie di Fourier)**.** Lo spazio delle funzioni continue $C(0, 2\pi)$ dell'Esempio (IV.50) contiene i sottospazi di dimensione finita

$$S_n = \mathrm{Span}(1, \cos x, \sin x, \ldots, \cos nx, \sin nx),$$

formati da tutte le combinazioni lineari delle funzioni $1, \cos kx, \sin kx$ con $k \leq n$. La dimensione di S_n è $2n + 1$, dato che i generatori sono linearmente indipendenti (sono ortogonali!). Per il Teorema (IV.45) esiste la proiezione di una funzione $f \in V$ sul sottospazio W, cioè la proiezione $\mathrm{proj}_W \colon V \to W$. Risulta definita da

▶ Cioè dai *polinomi trigonometrici* di grado n.

$$\begin{aligned}
\mathrm{proj}_W(f) &= \sum_{i=1}^{2n+1} \frac{\langle f, \boldsymbol{e}_i \rangle}{\|\boldsymbol{e}_i\|^2} \boldsymbol{e}_i \\
&= \frac{\langle f, 1 \rangle}{\|1\|^2} + \frac{\langle f, \cos x \rangle}{\|\cos x\|^2} \cos x + \frac{\langle f, \sin x \rangle}{\|\sin x\|^2} \sin x + \ldots + \frac{\langle f, \cos nx \rangle}{\|\cos nx\|^2} \cos nx + \frac{\langle f, \sin nx \rangle}{\|\sin nx\|^2} \sin nx \\
&= \frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \ldots + a_n \cos nx + b_n \sin nx,
\end{aligned}$$

dove si è posto

$$\begin{aligned}
a_k &= \frac{\langle f(x), \cos kx \rangle}{\|\cos kx\|^2} = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx \, \mathrm{d}x, \\
b_k &= \frac{\langle f(x), \sin kx \rangle}{\|\sin kx\|^2} = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx \, \mathrm{d}x.
\end{aligned}$$

I coefficienti c_i del Teorema (IV.45), che qui abbiamo chiamato a_i e b_i, sono quindi proprio i *coefficienti di Fourier* di f. Il polinomio trigonometrico $\bar{f}$ di grado n sopra trovato è quindi la migliore *approssimazione* di f, nel senso che minimizza la distanza

$$\|f - \bar{f}\|^2 = \int_0^{2\pi} (f(x) - \bar{f}(x))^2 \, \mathrm{d}x.$$ ■

▶ Consideriamo lo spazio V di tutti i polinomi nella variabile reale x, con il prodotto scalare

$$\langle f, g \rangle = \int_{-1}^{1} f(x) g(x) \, \mathrm{d}x,$$

come nell'Esempio (IV.36). Se si applica il procedimento di ortogonalizzazione di Gram–Schmidt alla successione di polinomi

$$1, x, x^2, \ldots, x^n, \ldots$$

si ottiene una successione di polinomi

$$f_0(x), f_1(x), \ldots, f_n(x), \ldots$$

chiamati i *polinomi di Legendre* (normalizzati).

§ 4. Il determinante

Riconsideriamo ora le proprietà delle aree e volumi con segno (cioè della funzione determinante in dimensione 2 e dimensione 3), trattate nei § 4

Spazi di Hilbert

Come abbiamo visto nella nota a margine della definizione (IV.15) o nell'Esempio (IV.50), gli spazi vettoriali non sempre hanno dimensione finita. Alcuni spazi infinito-dimensionali non solo hanno un prodotto scalare, ma in più godono anche di una proprietà molto importante: tutte le successioni di Cauchy ammettono un limite, cioè sono *completi* (si veda a pagina 352 e seguenti). Sono detti *spazi di Hilbert*.

(IV.52) Esempio (l^2). Sia l^2 l'insieme di tutte le successioni $\boldsymbol{x} = \{x_k\}_{k=1}^{\infty}$ di numeri reali a quadrati sommabili, cioè

$$l^2 = \left\{ \boldsymbol{x} = \{x_k\}_{k=1}^{\infty} : x_k \in \mathbb{R}, \sum_{k=1}^{\infty} x_k^2 < \infty \right\}.$$

La somma e il prodotto per uno scalare sono definite per componenti, come per il caso delle n-uple. Le successioni non sono altro che n-uple con $n = \infty$. È possibile definire un prodotto scalare su l^2, ponendo

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i,j=1}^{\infty} x_i y_i \implies \langle \boldsymbol{x}, \boldsymbol{x} \rangle = \|\boldsymbol{x}\|^2 = \sum_{i=1}^{\infty} x_i^2. \quad ■$$

(IV.53) Esempio (L^2). Se $(a,b) \subset \mathbb{R}$ è un intervallo,

$$L^2(a,b) = \{ f : (a,b) \to \mathbb{R} : \langle f, f \rangle < \infty \},$$

dove

$$\langle f, g \rangle = \int_a^b f(x) g(x) \, dx \Rightarrow \langle f, f \rangle = \|f\|^2 = \int_a^b f^2(x) \, dx.$$

Si tratta di uno spazio di dimensione infinita, costituito da tutte le funzioni con quadrato sommabile, con prodotto scalare (e conseguente norma). In realtà questa definizione non è ben posta (una funzione non nulla può avere norma nulla), e per renderla consistente occorre prendere non funzioni, ma *classi di equivalenza* di funzioni, definite a meno di un *insieme di misura di Lebesgue nulla*. Si veda a tal proposito il riquadro a pagina 488. ■

Le operazioni di proiezione ortogonale, distanza, e approssimazione che abbiamo visto per gli spazi euclidei $\mathbb{R}^n$ possono essere applicate con successo anche a questi spazi di funzioni. Entrambi gli spazi di Hilbert sopra descritti sono modelli di *spazio euclideo infinito-dimensionale*.

e § 5 del Capitolo III. A ben vedere, tutto ruota attorno al concetto di *linearità*, insieme al fatto che $\det(\boldsymbol{a}, \boldsymbol{a}) = 0$ (o, equivalentemente, come abbiamo visto, che $\det(\boldsymbol{a}, \boldsymbol{b}) = -\det(\boldsymbol{b}, \boldsymbol{a})$). Infatti, possiamo riscrivere le Proprietà (A1) e (A2) nella seguente unica affermazione:

▶ A pagina 115.

(L2) $\det(\boldsymbol{a}, \boldsymbol{b})$ è lineare sia nella variabile $\boldsymbol{a}$ che nella variabile $\boldsymbol{b}$ (cioè è lineare in ogni sua variabile).

Lo stesso per le Proprietà (V1) e (V2): possono essere scritte in modo più sintetico come

▶ A pagina 118.

(L3) $\det(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$ è lineare nella variabile $\boldsymbol{a}$, nella variabile $\boldsymbol{b}$ e nella variabile $\boldsymbol{c}$ (cioè è lineare in ogni sua variabile).

Riscrivendo la funzione determinante come funzione non di una n-upla di vettori ma come una funzione della matrice corrispondente, si può dire

(L2') La funzione $\det\left(\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}\right)$ è lineare nelle sue colonne.

(L3') La funzione $\det\left(\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}\right)$ è lineare nelle sue colonne.

Le Proprietà (A3') e (V3') possono leggersi come

(A) Se la matrice A' è la matrice ottenuta scambiando due colonne della matrice A, allora $\det(A') = -\det(A)$. Cioè il determinante cambia di segno se si scambiano due colonne di A.

Mostriamo che queste (che, come abbiamo visto, sono proprietà *caratterizzanti* per la funzione determinante) sono solo alcune delle proprietà. La forma più generale con cui consideriamo la funzione determinante è quella data dal seguente teorema.

(IV.54) Teorema. *Per ogni $n \geq 1$, esiste una unica funzione determinante* det*, che associa ad ogni matrice A (quadrata $n \times n$) un numero* $\det(A)$*, con le proprietà:*

① *Il determinante* det *è lineare nelle colonne di A.*

② *Il determinante* det *cambia segno se si scambiano due colonne.*

③ *Il determinante della matrice identica è* $\det(I_n) = 1$.

Risulta poi che det *è lineare anche nelle righe e che cambia di segno se si scambiano due righe tra loro.*

Prima di procedere con la dimostrazione, definiamo la *trasposta* di una matrice $n \times m$.

(IV.55) Definizione. Sia A una matrice $n \times m$. La trasposta di A, indicata con ${}^{t}\!A$, è la matrice $m \times n$ ottenuta scambiando le righe e le colonne di A, cioè la matrice che ha per colonne le righe di A

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \implies {}^{t}\!A = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{bmatrix}$$

▶ È ovvio che se A è quadrata, anche la trasposta ${}^{t}\!A$ è quadrata. La trasposta della matrice identica è ancora la matrice identica: ${}^{t}I_n = I_n$.

Dimostrazione del Teorema (IV.54). Per $n = 1$ la proposizione è vera: dato che una matrice 1×1 ha soltanto un coefficiente reale $\begin{bmatrix} a_{11} \end{bmatrix}$, basta porre $\det(\begin{bmatrix} a_{11} \end{bmatrix}) = a_{11}$. Esiste una unica riga e una unica colonna (di dimensione 1), ed è semplice verificare che questa funzione è lineare. Per $n = 2$, sappiamo dalla Proposizione (III.21) di pagina 116 che esiste unica una funzione determinante che soddisfa le proprietà in questione: esiste una unica funzione lineare nelle colonne, che cambia di segno se si scambiano due colonne e che vale uno sulla matrice identica. Quindi, se intendiamo dimostrare che il teorema è vero per la dimensione 2, dobbiamo verificare che la funzione determinante della Proposizione (III.21) è anche lineare per righe e cambia di segno se si scambiano due righe. Ma questo è facile: per l'unicità, è necessario che sia

▶ Queste sono le proprietà che mancano.

$$\det\left(\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}\right) = a_1 b_2 - a_2 b_1,$$

e le due proprietà che mancano sono vere, dato che

$$\det\left(\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}\right) = a_1 b_2 - a_2 b_1 = \det\left(\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}\right),$$

ovvero le proprietà aggiuntive seguono da $\det({}^{t}A) = \det(A)$.

Bene, adesso passiamo al caso $n = 3$: la Proposizione (III.24) ci assicura che c'è già una funzione del genere, ed è proprio il determinante in dimensione 3, che abbiamo definito nell'equazione (23) in modo ricorsivo come

▶ A pagina 121.

$$(9)\qquad \det\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = a_1 \det\begin{bmatrix} b_2 & c_2 \\ b_3 & c_3 \end{bmatrix} - b_1 \det\begin{bmatrix} a_2 & c_2 \\ a_3 & c_3 \end{bmatrix} + c_1 \det\begin{bmatrix} a_2 & b_2 \\ a_3 & b_3 \end{bmatrix}.$$

Sappiamo però che è lineare nelle colonne e che cambia di segno se si scambiano le colonne: dobbiamo mostrare che entrambe le proprietà sono vere per le righe. Per far questo ritorniamo alla dimostrazione della Proposizione (III.24), in cui abbiamo mostrato che vale l'equazione (22), che riscriviamo con la notazione matriciale come

$$(10)\qquad \det\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = a_1 \det\begin{bmatrix} b_2 & c_2 \\ b_3 & c_3 \end{bmatrix} - a_2 \det\begin{bmatrix} b_1 & c_1 \\ b_3 & c_3 \end{bmatrix} + a_3 \det\begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix}.$$

L'equazione (9) è uno *sviluppo* del determinante secondo la prima riga; l'equazione (10) è invece lo *sviluppo* del determinante secondo la prima colonna. Nella dimostrazione di (III.24) abbiamo usato lo sviluppo secondo la prima riga per *definire* il determinante e mostrare che ha le proprietà cercate per colonne, mentre abbiamo usato lo sviluppo secondo la prima colonna per mostrarne l'unicità. Ma lo sviluppo secondo la prima colonna di A non è altro che lo sviluppo secondo la prima riga della matrice trasposta ${}^{t}A$, e viceversa lo sviluppo secondo la prima riga di A è lo sviluppo secondo la prima colonna della trasposta ${}^{t}A$: dato che una funzione è lineare sulle colonne (risp. righe) di A se e solo se è lineare sulle righe (risp. colonne) della sua trasposta: possiamo ripetere il ragionamento sulla matrice ${}^{t}A$ e mostrare che det è lineare sulle righe e cambia di segno se si scambiano due righe qualsiasi. Per il caso $n \geq 4$, si dovrà procedere per induzione, nello stesso modo con cui si è passati da 2 a 3. Lasceremo solo qualche indizio su come iniziare la dimostrazione, negli Esercizi **55**, **56**, tralasciando i dettagli che possono essere completati dal Lettore. ■

▶ Abbiamo segnalato alcuni testi di algebra lineare nei Riferimenti bibliografici in appendice.

La seguente proposizione segue direttamente dal modo con cui abbiamo definito il determinante.

(IV.56) (Formula di Laplace) *Se A è una matrice quadrata $n \times n$*

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix},$$

allora il suo determinante si sviluppa nella prima colonna come

$$\text{(11)}\quad \det(A) = a_{1,1} \det \begin{bmatrix} a_{2,2} & \cdots & a_{2,n} \\ a_{3,2} & \cdots & a_{3,n} \\ \vdots & \ddots & \vdots \\ a_{n,2} & \cdots & a_{n,n} \end{bmatrix} - a_{2,1} \det \begin{bmatrix} a_{1,2} & \cdots & a_{1,n} \\ a_{3,2} & \cdots & a_{3,n} \\ \vdots & \ddots & \vdots \\ a_{n,2} & \cdots & a_{n,n} \end{bmatrix} + \ldots + (-1)^{n+1} a_{n,1} \det \begin{bmatrix} a_{1,2} & \cdots & a_{1,n} \\ a_{2,2} & \cdots & a_{2,n} \\ \vdots & \ddots & \vdots \\ a_{n-1,2} & \cdots & a_{n-1,n} \end{bmatrix}.$$

Possiamo scrivere l'equazione (11) anche come segue (sostituendo al simbolo "det" due barre verticali che racchiudono i coefficienti della matrice: questa notazione è molto comoda, specie per gli esercizi; inoltre scriviamo esplicitamente i coefficienti che abbiamo ignorato nel prendere i minori):

$$\det(A) = a_{1,1} \begin{vmatrix} \cancel{a_{1,1}} & \cancel{a_{1,2}} & \cancel{\cdots} & \cancel{a_{1,n}} \\ \cancel{a_{2,1}} & a_{2,2} & \cdots & a_{2,n} \\ \cancel{\vdots} & \vdots & \ddots & \vdots \\ \cancel{a_{n,1}} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix} - a_{2,1} \begin{vmatrix} \cancel{a_{1,1}} & a_{1,2} & \cdots & a_{1,n} \\ \cancel{a_{2,1}} & \cancel{a_{2,2}} & \cancel{\cdots} & \cancel{a_{2,n}} \\ \cancel{\vdots} & \vdots & \ddots & \vdots \\ \cancel{a_{n,1}} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix} + \ldots + (-1)^{n-1} a_{n,1} \begin{vmatrix} \cancel{a_{1,1}} & a_{1,2} & \cdots & a_{1,n} \\ \cancel{\vdots} & \vdots & \ddots & \vdots \\ \cancel{a_{n-1,1}} & a_{n-1,2} & \cdots & a_{n-1,n} \\ \cancel{a_{n,1}} & \cancel{a_{n,2}} & \cancel{\cdots} & \cancel{a_{n,n}} \end{vmatrix}.$$

(IV.57) Corollario. *Il determinante della trasposta di A è uguale al determinante di A:*

$$\det({}^{t}A) = \det(A).$$

Dimostrazione. È una conseguenza diretta del Teorema (IV.54), nella parte dell'unicità: entrambe sono funzioni lineari nelle righe e colonne, cambiano di segno se si scambiano righe o colonne, e valgono 1 sulla matrice identica I_n. Dato che può esistere una unica funzione con queste proprietà, $\det({}^{t}A)$ e $\det(A)$ devono essere uguali. ■

Vediamo ora una semplice proprietà della trasposta di una matrice, che ci tornerà utile in seguito.

(IV.58) *La trasposta del prodotto di due matrici A e B è uguale al prodotto delle trasposte nell'ordine inverso*

$${}^{t}(AB) = {}^{t}B\,{}^{t}A.$$

Dimostrazione. Se a_{ij} indicano i coefficienti di A, b_{jk} i coefficienti di B, e $\bar{a}_{ji} = a_{ij}$ e $\bar{b}_{kj} = b_{jk}$ i coefficienti delle trasposte di A e B, allora il coefficiente (i,k) del prodotto AB è

$$\sum_{j=1}^{m} a_{ij} b_{jk},$$

che quindi è uguale al coefficiente (k,i) di ${}^{t}(AB)$. Ma questo è uguale al coefficiente (k,i) del prodotto ${}^{t}B\,{}^{t}A$, che è

$$\sum_{j=1}^{m} \bar{b}_{kj} \bar{a}_{ji} = \sum_{j=1}^{m} b_{jk} a_{ij}.$$ ■

Il Teorema che segue è la generalizzazione a dimensione arbitraria dei Teoremi (III.23) e (III.29) del Capitolo III.

(IV.59) Teorema. *Sia A una matrice quadrata $n \times n$. Allora le seguenti affermazioni sono equivalenti:*

① *Le colonne di A sono linearmente indipendenti, cioè non esiste un vettore $\boldsymbol{x} \neq \boldsymbol{0}$, $\boldsymbol{x} \in \mathbb{R}^n$ tale che*

$$A\boldsymbol{x} = \boldsymbol{0}.$$

② $\det(A) \neq 0$

③ *Per ogni $\boldsymbol{c} \in \mathbb{R}^n$ il sistema di n equazioni in n incognite*

$$A\boldsymbol{x} = \boldsymbol{c}$$

ha una e una sola soluzione.

④ *Le righe di A sono linearmente indipendenti.*

▶ Le negazioni delle quattro affermazioni (che a loro volta saranno equivalenti) sono:

① Esiste $\boldsymbol{x} \neq \boldsymbol{0}$ tale che $A\boldsymbol{x} = \boldsymbol{0}$ (cioè le colonne di A sono linearmente dipendenti).

② $\det(A) = 0$.

③ Esiste $\boldsymbol{c} \in \mathbb{R}^n$ per cui il sistema di equazioni $A\boldsymbol{x} = \boldsymbol{c}$ ammette più di una soluzione oppure nessuna soluzione.

④ Le colonne di A sono linearmente dipendenti.

Dimostrazione. Cominciamo a dimostrare per induzione, così come abbiamo dimostrato il Teorema (III.29) a partire dal Teorema (III.23), che ① $\iff$ ② $\iff$ ③. Supponiamo quindi che questa equivalenza valga per $n-1$, e mostriamo che vale per n. Eliminando una variabile dal sistema di n equazioni in n incognite $A\boldsymbol{x} = \boldsymbol{0}$, otteniamo un sistema di $n-1$ equazioni in $n-1$ incognite con matrice che ha ancora determinante zero, se $\det A = 0$, e che quindi ha soluzioni non nulle (① $\implies$ ②). Per mostrare ② $\implies$ ③, osserviamo che, se eliminiamo una variabile come sopra, il sistema di $n-1$ equazioni in $n-1$ incognite ha matrice dei coefficienti con determinante diverso da zero, e quindi ha sempre una e una sola soluzione (per induzione), da cui possiamo sostituire all'indietro e trovare la soluzione cercata. Per mostrare ③ $\implies$ ①, basta osservare che se per ogni $\boldsymbol{c} \in \mathbb{R}^n$ il sistema ha una e una sola soluzione, allora in particolare per $\boldsymbol{x} = \boldsymbol{0}$ c'è una e una sola soluzione di $A\boldsymbol{x} = \boldsymbol{0}$. Dato che $\boldsymbol{x} = \boldsymbol{0}$ è una soluzione, non possono essercene altre cioè le colonne di A sono linearmente indipendenti. Ora, la dimostrazione è completata semplicemente notando che $\det({}^t\!A) = \det(A)$. ■

▶ Descriviamo la dimostrazione per sommi capi, lasciando al Lettore di riempire i dettagli mancanti.

(IV.60) Corollario (Formula di Binet)**.** *Il determinante del prodotto di due matrici quadrate A e B è uguale al prodotto dei loro determinanti:*

$$\det(AB) = \det(A)\det(B).$$

▶ Perciò $\det(AB) = \det(BA)$ anche se $AB \neq BA$.

Dimostrazione. Deduciamo questa proprietà dall'unicità della funzione determinante e dalla definizione di prodotto di matrici. Supponiamo per prima cosa $\det(A) \neq 0$. Allora possiamo scrivere la funzione di B

$$f(B) = \frac{\det(AB)}{\det(A)},$$

o anche

$$f(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n) = \frac{\det(A\boldsymbol{b}_1, A\boldsymbol{b}_2, \ldots, A\boldsymbol{b}_n)}{\det(A)},$$

se vogliamo insistere sulla struttura in colonne di B. È facile vedere che f è lineare nelle sue variabili $\boldsymbol{b}_i$ e che cambia di segno se ne scambiano due. Non resta che vedere quanto vale sulla matrice identica I_n:

$$f(I_n) = \frac{\det(AI_n)}{\det(A)} = \frac{\det(A)}{\det(A)} = 1,$$

dato che $AI_n = A$. Per l'unicità, deve quindi essere

$$f(B) = \det(B) \implies \frac{\det(AB)}{\det(A)} = \det(B) \implies \det(AB) = \det(A)\det(B).$$

Se invece $\det(A) = 0$, l'affermazione è vera se riusciamo a mostrare che $\det(AB) = 0$. Ora, per il Teorema (IV.59), $\det(A) = 0$ se e soltanto se le sue righe sono linearmente dipendenti, cioè se esiste $\boldsymbol{x} \neq \boldsymbol{0}$ tale che

$${}^t\!A\boldsymbol{x} = \boldsymbol{0}.$$

Analogamente, $\det(AB) = 0$ se e solo se le sue righe sono linearmente indipendenti, cioè se esiste $\boldsymbol{y} \neq \boldsymbol{0}$ tale che

$${}^t(AB)\boldsymbol{y} = \boldsymbol{0}.$$

Ma per (IV.58), ${}^t(AB) = {}^t\!B\,{}^t\!A$, e dunque basta porre $\boldsymbol{y} = \boldsymbol{x}$ per ottenere

$${}^t(AB)\boldsymbol{x} = ({}^t\!B\,{}^t\!A)\boldsymbol{x} = {}^t\!B({}^t\!A\boldsymbol{x}) = {}^t\!B\boldsymbol{0} = \boldsymbol{0}.$$ ■

Aggiungiamo una affermazione al Teorema (IV.59).

▶ Ricordiamo che l'*inversa* di una matrice A è una matrice A^{-1} che soddisfa l'equazione (7), cioè tale che $A^{-1}A = AA^{-1} = I$.

(IV.61) (Inversa di una matrice) *Esiste la matrice inversa A^{-1} della matrice A se e soltanto se* $\det(A) \neq 0$.

Dimostrazione. Supponiamo che esista l'inversa A^{-1}: allora

$$AA^{-1} = I \implies \det(AA^{-1}) = \det(A)\det(A^{-1}) = \det(I) = 1,$$

e quindi né $\det(A)$ né $\det(A^{-1})$ possono essere zero, da cui $\det(A) \neq 0$. Viceversa, se $\det(A) \neq 0$, per il Teorema (IV.59) per ogni $\boldsymbol{c} \in \mathbb{R}^n$ l'equazione

$$(*) \qquad A\boldsymbol{x} = \boldsymbol{c}$$

ha una e una sola soluzione. Siano $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ i vettori della base standard di $\mathbb{R}^n$. Allora per ogni $i = 1, \ldots, n$ esiste una e una sola soluzione $\boldsymbol{x}$ dell'equazione $(*)$ con $\boldsymbol{c} = \boldsymbol{e}_i$, e quindi sono univocamente determinati n vettori $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n$ tali che per ogni i,

$$A\boldsymbol{b}_i = \boldsymbol{e}_i,$$

da cui se B è la matrice con colonne $\boldsymbol{b}_i$ si deduce

$$AB = A(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n) = (A\boldsymbol{b}_1, A\boldsymbol{b}_2, \ldots, A\boldsymbol{b}_n) = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n) = I,$$

cioè $AB = I$. Analogamente, $\det({}^t\!A) \neq 0$ e quindi esiste una matrice B' tale che ${}^t\!AB' = I \implies {}^t\!B'A = I$. Moltiplicando da destra per B si ha

$${}^t\!B'A = I \implies B = IB = ({}^t\!B'A)B = {}^t\!B'(AB) = {}^t\!B'I = {}^t\!B',$$

cioè $B = {}^t\!B'$ e non solo $AB = I$, ma anche $BA = I$, cioè B è l'inversa di A. ■

▶ Ci sono molti metodi di calcolare l'inversa di una matrice: risolvere il sistema $(*)$ nei $\boldsymbol{c} = \boldsymbol{e}_i$ con eliminazione di Gauss, usare le formule di Laplace per determinare una formula ricorsiva per l'inversa, oppure usare una generalizzazione del metodo che abbiamo usato nella dimostrazione del Teorema (III.29), chiamato *metodo di Cramer*. Per dimostrare che l'inversa A^{-1} esiste non abbiamo però bisogno di costruirla esplicitamente. Si veda anche l'Esercizio (IV.3) a pagina 216.

Terminiamo questa sezione con una significativa estensione del Teorema (IV.59). Occorre prima definire il *rango* di una matrice e le sottomatrici quadrate di una matrice. Cominciamo con le *sottomatrici* quadrate di una matrice: se A è una matrice $n \times m$ e si scelgono s indici $i_1, i_2, \ldots, i_s$ e s indici $j_1, j_2, \ldots, j_s$ con

$$1 \leq i_1 < i_2 < \ldots < i_s \leq n, \qquad 1 \leq j_1 < j_2 < \ldots < j_s \leq m.$$

▶ Ricordate le sottosuccessioni? È la stessa cosa! Se le matrici sono successioni doppie $a_{i,j}$, una sottomatrice è una sottosuccessione di questa sottosuccessione doppia... (qui $0 < s \leq \min(n, m)$).

La sottomatrice quadrata di A corrispondente a questa scelta è la matrice $s \times s$

$$\begin{bmatrix} a_{i_1,j_1} & a_{i_1,j_2} & \cdots & a_{i_1,j_s} \\ a_{i_2,j_1} & a_{i_2,j_2} & \cdots & a_{i_2,j_s} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_s,j_1} & a_{i_s,j_2} & \cdots & a_{i_s,j_s} \end{bmatrix}.$$

(IV.62) Esempio. La sottomatrice corrispondente agli indici $(i_1, i_2) = (2, 3)$ e $(j_1, j_2) = (1, 3)$ della matrice indicata nel diagramma è

$$\begin{bmatrix} 1 & 2 & 3 \\ \boxed{4} & 5 & \boxed{6} \\ \boxed{7} & 8 & \boxed{9} \end{bmatrix} \Longrightarrow \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix}.$$

■

(IV.63) Definizione. Sia $A = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m)$ una matrice $n \times m$. Il *rango* (per colonne) di A è la dimensione dello spazio generato dalle sue colonne

$$\operatorname{Rango}(A) = \dim\left(\operatorname{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m)\right),$$

che come abbiamo visto nella Proposizione (IV.18) è il massimo numero di colonne $\boldsymbol{a}_i$ linearmente indipendenti.

▶ Il rango è anche la dimensione dell'*immagine* della funzione lineare $A \colon \mathbb{R}^m \to \mathbb{R}^n$ associata alla matrice A.

▶ Vedremo tra pochissimo che $\operatorname{Rango}(A) = \operatorname{Rango}({}^tA)$, cioè che il rango per colonne è uguale al rango per righe.

(IV.64) Teorema. *Sia A una matrice $n \times m$. Allora le seguenti affermazioni sono equivalenti:*

① *(rango per colonne)* $\operatorname{Rango}(A) = r$.

② *(rango per righe)* $\operatorname{Rango}({}^tA) = r$.

③ *(sottomatrici orlate — Kronecker) Esiste almeno una sottomatrice $r \times r$ di A con determinante diverso da zero, ma tutte le sottomatrici quadrate di dimensione $r+1$ che la estendono hanno determinante zero.*

④ *(nullità più rango) Il nucleo di A ha dimensione $m - r$, cioè*

$$\dim(\operatorname{Nucleo}(A)) + r = m.$$

Dimostrazione. Cominciamo con l'osservazione, molto facile da dimostrare, che il rango di una matrice non cambia se si permutano le sue colonne né se si permutano le sue righe. Allora possiamo sempre permutare righe e colonne di A e metterla nella seguente forma,

▶ Dato che lo Span non dipende dall'ordine con cui sono presi i vettori Inoltre permutare due componenti x_i e x_j dei vettori colonna corrisponde semplicemente ad una permutazione dei vettori della base canonica.

$$(12)\quad \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,c} & a_{1,c+1} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,c} & a_{2,c+1} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{r,1} & a_{r,2} & \cdots & a_{r,c} & a_{r,c+1} & \cdots & a_{r,m} \\ a_{r+1,1} & a_{r+1,2} & \cdots & a_{r+1,c} & a_{r+1,c+1} & \cdots & a_{r+1,m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,c} & a_{n,c+1} & \cdots & a_{n,m} \end{bmatrix}$$

in cui le prime c colonne sono indipendenti mentre le successive $m-c$ sono combinazioni lineari delle prime c, e le prime r righe sono indipendenti, mentre le successive $n-r$ sono combinazioni lineari delle prime r. Ora, sia $\hat{A}$ la sottomatrice di A delimitata dal tratteggio nella (12). Allora le sue colonne

▶ Quindi c è il rango per colonne, r è il rango per righe.

$$\hat{\boldsymbol{a}}_1, \hat{\boldsymbol{a}}_2, \ldots, \hat{\boldsymbol{a}}_c$$

(formate da vettori di dimensione r) non possono essere linearmente dipendenti: infatti, se esistessero $x_1, x_2, \ldots, x_c$ non tutti nulli tali che

$$(13)\quad x_1 \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{r,1} \end{bmatrix} + x_2 \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{r,2} \end{bmatrix} + \ldots + x_c \begin{bmatrix} a_{1,c} \\ a_{2,c} \\ \vdots \\ a_{r,c} \end{bmatrix} = \mathbf{0},$$

▶ Questo è un sistema di r equazioni in c incognite.

allora anche le prime c colonne $\boldsymbol{a}_i$ di A soddisferebbero l'equazione

$$(14)\quad x_1 \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{r,1} \\ a_{r+1,1} \\ \vdots \\ a_{n,1} \end{bmatrix} + x_2 \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{r,2} \\ a_{r+1,2} \\ \vdots \\ a_{n,2} \end{bmatrix} + \ldots + x_c \begin{bmatrix} a_{1,c} \\ a_{2,c} \\ \vdots \\ a_{r,c} \\ a_{r+1,c} \\ \vdots \\ a_{n,c} \end{bmatrix} = \mathbf{0}$$

▶ Questo è un sistema di n equazioni in c incognite.

visto che le ultime $n-r$ equazioni dipendono linearmente dalle prime r per ipotesi. Ma c vettori di $\mathbb{R}^r$ possono essere linearmente indipendenti solo quando $c \leq r$. Ragionando allo stesso modo sulla trasposta della matrice A si può concludere che $r \leq c$, e quindi in definitiva che r e c sono uguali. Ma c è per definizione il rango di A e r il rango della sua trasposta: abbiamo appena dimostrato l'equivalenza ① $\iff$ ②, cioè che

▶ In altre parole, il sistema di equazioni (14) è equivalente al sistema (13), dato che è ottenuto aggiungendo combinazioni lineari delle prime r equazioni.

$$\mathrm{Rango}(A) = \mathrm{Rango}({}^{t}\!A).$$

Da questo segue un fatto che tra breve ci tornerà utile, di cui lasciamo la dimostrazione per esercizio:

il rango di qualsiasi sottomatrice di A non può essere maggiore del rango di A.

▶ Esercizio 47 a pagina 228.

Torniamo alla matrice nella forma (12), in cui possiamo essere certi che $c = r$, cioè che la sottomatrice $\hat{A}$ è sempre quadrata. Non solo: dal momento che $\hat{A}$ ha colonne (righe) linearmente indipendenti, $\det(\hat{A}) \neq 0$. Inoltre ogni volta che la *orliamo*, cioè ogni volta che le aggiungiamo una riga e una colonna, e definiamo una sottomatrice quadrata di ordine $r+1$ che contiene $\hat{A}$, non possiamo far altro che ottenere matrici con determinante zero. Se così non fosse, esisterebbe una sottomatrice $(r+1) \times (r+1)$ di A con determinante diverso da zero, quindi con $r+1$ righe e $r+1$ colonne indipendenti, il cui rango sarebbe perciò $r+1$. Ma il rango di qualsiasi sottomatrice di A, abbiamo detto poco fa, non può superare r. Abbiamo perciò concluso la dimostrazione dell'implicazione ① $\implies$ ③.

Ora, supponiamo invece vera ③, cioè di aver trovato una sottomatrice $p \times p$, che chiameremo $\hat{B}$, con determinante diverso da zero e tale che tutte sue le orlate di dimensione $p+1$ in A abbiano determinante zero. È possibile che p non sia il rango? Il rango di $\hat{B}$ è p, visto che ha determinante diverso da zero, e quindi deve essere $p \leq r$. A meno di permutare righe e colonne di A, la sottomatrice $\hat{B}$ è quella tratteggiata nel seguente diagramma.

▶ Per l'osservazione fatta poco fa.

$$(15)\qquad \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,p} & a_{1,p+1} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,p} & a_{2,p+1} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{p,1} & a_{p,2} & \cdots & a_{p,p} & a_{p,p+1} & \cdots & a_{p,m} \\ a_{p+1,1} & a_{p+1,2} & \cdots & a_{p+1,p} & a_{p+1,p+1} & \cdots & a_{p+1,m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,p} & a_{n,p+1} & \cdots & a_{n,m} \end{bmatrix}$$

Dato che le p colonne di $\hat{B}$ sono linearmente indipendenti, anche le prime p colonne di A devono essere linearmente indipendenti. La condizione sui determinanti implica che la $(p+1)$-esima colonna di A, punteggiata in figura, è dipendente dalle prime p: infatti, visto che le p colonne di $\hat{B}$ sono una base per $\mathbb{R}^p$, esistono unici i p coefficienti $x_1, x_2, \ldots, x_p$ tali che

$$(16)\qquad x_1 \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{p,1} \end{bmatrix} + x_2 \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{p,2} \end{bmatrix} + \ldots + x_p \begin{bmatrix} a_{1,p} \\ a_{2,p} \\ \vdots \\ a_{p,p} \end{bmatrix} = \begin{bmatrix} a_{1,p+1} \\ a_{2,p+1} \\ \vdots \\ a_{p,p+1} \end{bmatrix}.$$

Ora consideriamo, per $j \geq p+1$, la matrice orlata con la j-esima riga e la $(p+1)$-esima colonna,

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,p} & a_{1,p+1} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,p} & a_{2,p+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{p,1} & a_{p,2} & \cdots & a_{p,p} & a_{p,p+1} \\ a_{j,1} & a_{j,2} & \cdots & a_{j,p} & a_{j,p+1} \end{bmatrix}$$

la cui ultima riga è combinazione lineare delle prime p per ipotesi. Ma allora la medesima combinazione lineare delle p equazioni del sistema (16), si scrive

$$x_1 a_{j,1} + x_2 a_{j,2} + \ldots + x_p a_{j,p} = a_{j,p+1}.$$

Questo vale per tutti i j, quindi la $(p+1)$-esima colonna è davvero combinazione lineare delle prime p.

Se ripetiamo questo ragionamento per tutte le altre colonne, concludiamo che il rango di A non può superare p, e quindi che $p = \mathrm{Rango}(A)$, cioè ③ $\implies$ ①.

Per concludere la dimostrazione non ci resta che mostrare ② $\iff$ ④, cioè che il rango di A più la nullità di A è uguale al numero di colonne di A. Ricordiamo che per risolvere il sistema $A\boldsymbol{x} = \mathbf{0}$ di n equazioni in m incognite x_i con il metodo di eliminazione delle variabili operiamo *trasformazioni elementari* del sistema. Quando eliminiamo una variabile, passiamo da un sistema di n equazioni in m incognite ad un sistema di $n-1$ equazioni in $m-1$ incognite, lasciando invariata la nullità, visto che l'insieme delle soluzioni, cioè il nucleo di A, non cambia. Dal punto di vista delle matrici, si passa dalla matrice $n \times m$ A ad una matrice $(n-1) \times (m-1)$, che chiamiamo $\hat{A}$. Cosa succede del rango? È fondamentale capire che il suo rango diminuisce di uno. Infatti, supponiamo di voler eliminare x_1 utilizzando la prima equazione. Allora sommando alle altre righe opportuni multipli della prima equazione, si elimina x_1 da tutte le equazioni tranne la prima. Lo Span delle righe non cambia quando ad una riga si somma la combinazione di altre righe, e quindi il rango di A non cambia. Ci si trova ad un certo punto ad avere la matrice A della forma

Le trasformazioni elementari sono le seguenti: permutare le equazioni, sommare ad una equazione una opportuna combinazione lineare delle altre equazioni e quando capita cancelare le equazioni del tipo $0 = 0$.

$$\begin{bmatrix} 1 & a_{1,2} & \cdots & a_{1,m} \\ 0 & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n,2} & \cdots & a_{n,m} \end{bmatrix},$$

mentre la matrice $\hat{A}$ del sistema dopo l'eliminazione di x_1 è la sottomatrice tratteggiata. La prima riga è sempre linearmente indipendente dalle altre $n-1$; tra queste $n-1$, è facile vedere che il massimo numero di linearmente indipendenti è uguale al massimo numero di righe indipendenti della sottomatrice $\hat{A}$, e quindi

$$\mathrm{Rango}(\hat{A}) + 1 = \mathrm{Rango}(A).$$

Cioè ad ogni eliminazione si diminuisce di uno il numero di variabili e di uno il rango: la differenza $m - \mathrm{Rango}$ quindi rimane costante. Il procedimento termina quando non ci sono più variabili da eliminare, cioè quando si rimane con nessuna equazione nelle d variabili rimaste. È chiaro quindi che il numero d è proprio la dimensione del nucleo di A, che può essere parametrizzato con queste ultime d variabili. In questo caso (terminale) la differenza tra il numero di variabili (d) e rango della matrice (0, visto che non ci sono equazioni) è uguale a d. Ma nemmeno d cambia ad ogni passo dell'eliminazione, e quindi

$$\dim(\mathrm{Nucleo}(A)) = m - \mathrm{Rango}(A). \qquad \blacksquare$$

(IV.65) Corollario (Teorema di Rouché–Capelli). *Il sistema di n equazioni in m incognite*

$$Ax = \boldsymbol{b}$$

ha soluzioni se e soltanto se il rango di A e il rango della matrice $A|\boldsymbol{b}$ ottenuta aggiungendo la colonna $\boldsymbol{b}$ ad A coincidono:

$$\mathrm{Rango}(A) = \mathrm{Rango}(A|\boldsymbol{b}).$$

In questo caso, lo spazio di tutte le soluzioni ha dimensione uguale a quella del nucleo di A, che è uguale a $m -$ Rango(A).

▶ La matrice $A|\boldsymbol{b}$ è detta anche *matrice aumentata* o *matrice completa*.

Dimostrazione. Sappiamo che $A\boldsymbol{x} = \boldsymbol{b}$ ha soluzioni se e soltanto se $\boldsymbol{b}$ è combinazione lineare dei vettori colonna $(\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m)$ di A, cioè se e solo se $\boldsymbol{b} \in \mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m)$. Ma

$$\begin{aligned}
&\boldsymbol{b} \in \mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m)\\
\iff &\mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m) = \mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m, \boldsymbol{b}),\\
\iff &\dim \mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m) = \dim \mathrm{Span}(\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m, \boldsymbol{b})\\
\iff &\mathrm{Rango}(A) = \mathrm{Rango}(A|\boldsymbol{b}).
\end{aligned}$$

Ora, se $\bar{\boldsymbol{x}}$ è una qualsiasi soluzione del sistema di equazioni $A\boldsymbol{x} = \boldsymbol{b}$, allora con il cambio di variabili $\boldsymbol{x} = \boldsymbol{y} + \bar{\boldsymbol{x}}$ il sistema diventa

$$A(\boldsymbol{y} + \bar{\boldsymbol{x}}) = \boldsymbol{b} \quad \iff \quad A\boldsymbol{y} + A\bar{\boldsymbol{x}} = \boldsymbol{b} \quad \iff \quad A\boldsymbol{y} = \boldsymbol{0}.$$

L'insieme delle soluzioni di $A\boldsymbol{x} = \boldsymbol{b}$ è perciò parametrizzato dalle soluzioni $\boldsymbol{y}$ dell'equazione omogenea $A\boldsymbol{y} = \boldsymbol{0}$, che costituiscono uno spazio vettoriale di dimensione $\dim \mathrm{Nucleo}(A) = m - \mathrm{Rango}(A)$. ■

▶ È curioso che questo teorema, o la sua prima parte, è attribuito a matematici diversi, nelle diverse parti del mondo. Pare un esercizio di analisi combinatoria. In Italia è noto come il *Teorema di Rouché–Capelli*. Nei paesi di lingua spagnola, è il *Teorema di Rouché–Frobenius*. Per alcuni paesi dell'Est europa è il *Teorema di Kronecker–Capelli*. In alcuni testi di lingua frances è il *Teorema di Rouché–Fontené*. Naturalmente poi ci sono tutte le varianti e combinazioni: *Rouché–Frobenius–Kronecker–Capelli*, *Kronecker–Rouché–Capelli*, ...

§ 5. Forme quadratiche

(IV.66) Definizione. Una *forma quadratica* $q(\boldsymbol{x})$ definita sui vettori $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ di $\mathbb{R}^n$ è un polinomio omogeneo di secondo grado nelle variabili x_i.

Osserviamo che se indichiamo con a_{ii} il coefficiente di x_i^2 in $q(\boldsymbol{x})$ e con $a_{ij} = a_{ji}$ la metà del coefficiente di $x_i x_j$, si ha

$$(17) \quad q(\boldsymbol{x}) = \sum_{i,j} a_{ij} x_i x_j = {}^t\boldsymbol{x} A \boldsymbol{x},$$

dove A è la matrice degli a_{ij}, $\boldsymbol{x}$ è il vettore colonna con componenti x_i e A la matrice (simmetrica) con componenti a_{ij}, che è detta *matrice associata* alla forma quadratica q. Infatti, nella somma (17) i termini x_i^2 hanno

▶ Il termine *forma* è molto usato in matematica (nei primi capitoli è stato usato 41 volte, e verrà usato 86 volte in questo capitolo). In tutto il volume compare 215 volte. Ma cosa è in generale una forma? Cosa significa? Forma *lineare*, forma *bilineare*, forma *quadratica*, forma *differenziale* sono solo alcuni dei molti modi di incontrare il termine. Termini derivati e composti sono: «formaggio» (principale sostentamento di molti matematici, insieme al pane), «formola» (anche nella versione un po' sciapa moderna di «formula», sostentamento intellettuale del matematico, insieme al formaggio), «formosa», «uniforme», «trasformare», «conforme», ... Purtroppo le origini di questo uso, così come le origini del formaggio e delle formole, sono a noi del tutto ignote.

coefficiente a_{ii} come voluto, mentre i termini $x_i x_j$ compaiono sia come $a_{ij}x_i x_j$ che come $a_{ji}x_j x_i = a_{ij}x_i x_j$. Dunque se $i \neq j$ il coefficiente di $x_i x_j$ nella somma (17) è $a_{ji} + a_{ij} = 2a_{ij}$.

(IV.67) Esempio. Se $n = 1$, la generica forma quadratica sarà del tipo $q(x_1) = ax_1^2$, con $a \in \mathbb{R}$. Per $n = 2$, ci saranno tre coefficienti a, b e c tali che

$$q(\boldsymbol{x}) = q(x_1, x_2) = ax_1^2 + bx_2^2 + cx_1x_2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b/2 \\ b/2c & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

▶ Utilizzando il prodotto scalare standard, a volte scriveremo anche

$$q(\boldsymbol{x}) = {}^t\boldsymbol{x}A\boldsymbol{x} = (A\boldsymbol{x}) \cdot \boldsymbol{x} = A\boldsymbol{x} \cdot \boldsymbol{x}.$$

Per $n = 3$, il generico polinomio omogeneo di grado due è somma di 6 monomi, e può essere scritto come

$$q(\boldsymbol{x}) = q(x_1, x_2, x_3) = ax_1^2 + bx_2^2 + cx_3^2 + dx_1x_2 + ex_1x_3 + fx_2x_3.$$

■

▶ Cioè

$$q(\boldsymbol{x}) = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} a & d/2 & e/2 \\ d/2 & b & f/2 \\ e/2 & f/2 & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

In generale, si può mostrare che ci sono $n(n+1)/2$ monomi di grado due in n variabili (Esercizio 49 a pagina 228), che sono anche i parametri liberi che descrivono le matrici simmetriche.

(IV.68) Esempio. Le seguenti sono forme quadratiche di $\mathbb{R}^3$, nelle coordinate (x, y, z).

① $q_1 = x^2 + y^2 + z^2$;

② $q_2 = x^2 - z^2$;

③ $q_3 = xy$;

④ $q_4 = x^2 + 2y^2 - 2xy + z^2$. ■

Ricordiamo però che è sempre possibile scegliere in $\mathbb{R}^n$ altri sistemi di coordinate (si veda a pagina 179): basta scegliere una matrice S invertibile e definire le coordinate

$$\boldsymbol{x} = S\hat{\boldsymbol{x}} \qquad \left(\iff \hat{\boldsymbol{x}} = S^{-1}\boldsymbol{x} \right).$$

Una qualsiasi forma quadratica q nelle variabili x_i cambierà corrispondentemente in un'altra forma $\hat{q}$ nelle nuove n variabili $\hat{x}_i$, definita semplicemente ponendo

$$\hat{q}(\hat{\boldsymbol{x}}) = q(\boldsymbol{x}).$$

▶ Se pensiamo la forma quadratica scritta in forma matriciale

$$q(\boldsymbol{x}) = {}^t\boldsymbol{x}A\boldsymbol{x},$$

con A matrice simmetrica associata a q, allora se $\boldsymbol{x} = S\hat{\boldsymbol{x}}$ la forma $\hat{q}(\hat{\boldsymbol{x}})$ si scriverà come

$$\hat{q}(\hat{\boldsymbol{x}}) = {}^t(S\hat{\boldsymbol{x}})A(S\hat{\boldsymbol{x}}) = {}^t\hat{\boldsymbol{x}}\,{}^tSAS\hat{\boldsymbol{x}},$$

con matrice associata

$$\hat{A} = {}^tSAS.$$

In questo caso si dirà che $\hat{A}$ è *congruente* a A. Si deve fare attenzione: *sembra* che la $\hat{A}$ sia *simile* a A, ma non lo è in generale. Perché $\hat{A}$ sia simile a A è necessario che $\hat{A} = S^{-1}AS$, e ${}^tSAS = S^{-1}AS$ se e soltanto se $S^{-1} = {}^tS$, cioè se e soltanto se S è una matrice *ortogonale* (si ricordi la Definizione (IV.43)). Infatti, i coefficienti della matrice ${}^tAA$ non sono altro che i prodotti scalari $\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle$, con $i, j = 1, \dots, n$. La matrice è ortogonale se e soltanto se

$$\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j, \end{cases}$$

cioè se ${}^tAA$ ha 0 fuori dalla diagonale e 1 sulla diagonale, ovverosia se è la matrice identica I. Due matrici M_1 e M_2 possono essere simili o congruenti, e queste due relazioni sono indipendenti: M_1 e M_2 possono essere simili ma non congruenti, congruenti ma non simili, sia congruenti che simili o né congruenti né simili.

(IV.69) Esempio. Osserviamo come si trasformano le forme quadratiche dell'Esempio (IV.68) mediante la trasformazione

$$\begin{bmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

cioè

$$\begin{cases} \hat{x} = x + y \\ \hat{y} = y \\ \hat{z} = z. \end{cases} \iff \begin{cases} x = \hat{x} - \hat{y} \\ y = \hat{y} \\ z = \hat{z}. \end{cases}$$

Nelle nuove coordinate, le forme quadratiche q_i, $i = 1, 2, 3, 4$ si scrivono quindi come

① $\hat{q}_1 = (\hat{x} - \hat{y})^2 + \hat{y}^2 + \hat{z}^2 = \hat{x}^2 + 2\hat{y}^2 - 2\hat{x}\hat{y} + \hat{z}^2$;

② $\hat{q}_1 = (\hat{x} - \hat{y})^2 - \hat{z}^2 = \hat{x}^2 + \hat{y}^2 - 2\hat{x}\hat{y} - \hat{z}^2$;

③ $\hat{q}_3 = (\hat{x} - \hat{y})\hat{y} = \hat{x}\hat{y} - \hat{y}^2$;

④ $\hat{q}_4 = (\hat{x} - \hat{y})^2 + 2\hat{y}^2 - 2(\hat{x} - \hat{y})\hat{y} + \hat{z}^2 = \hat{x}^2 + 5\hat{y}^2 - 4\hat{x}\hat{y} + \hat{z}^2$. ■

Con il prossimo teorema vediamo che tutte le forme quadratiche possono essere scritte senza termini misti, in una forma che è detta *diagonale*. I cambi di variabili che useremo saranno scelti fra i tre seguenti (in una opportuna successione):

▶ Vedremo poco sotto il motivo di questo termine.

- Permutazione delle variabili x_i (per esempio $x_1 = y_2$, $x_2 = y_1$, $x_3 = y_3$).
- Sostituire a x_1 la somma $y_1 + by_2$ per un certo b: $\begin{cases} x_1 = y_1 + by_2 \\ x_2 = y_2 \\ \vdots \quad \vdots \\ x_n = y_n \end{cases}$
- Sostituire a x_1, x_2 i termini $y_1 + y_2, y_1 - y_2$: $\begin{cases} x_1 = y_1 + y_2 \\ x_2 = y_1 - y_2 \\ x_3 = y_3 \\ \vdots \quad \vdots \\ x_n = y_n \end{cases}$

(IV.70) Teorema (Completamento dei quadrati di Lagrange)**.** *Se q è una forma quadratica $q(\boldsymbol{x})$ su $\mathbb{R}^n$, allora con una successione di cambi di variabili dei tre tipi elencati sopra è possibile renderla diagonale, cioè trovare variabili $\boldsymbol{y} = (y_1, y_2, \dots, y_n)$ tali che*

(18) $\hat{q} = c_1 y_1^2 + c_2 y_2^2 + \ldots + c_n y_n^2$

per certi coefficienti $c_1, c_2, \dots, c_n$.

▶ Si dice anche *diagonalizzare* q.

Dimostrazione. Procediamo ricorsivamente: mostriamo che con siffatti cambi di variabili si può scrivere ogni forma quadratica in n variabili $q(\boldsymbol{x})$ nella forma

$$c_1 y_1^2 + q_2(y_2, \dots, y_n),$$

dove q_2 è una forma quadratica nelle $n-1$ variabili $y_2, \dots, y_n$. Applicando $n-1$ volte questa procedura otteniamo la diagonalizzazione cercata. Iniziamo considerando tutti i termini di $q(\boldsymbol{x})$ che contengono la variabile x_1, cioè

(19) $q_1 = a_{1,1}x_1^2 + 2a_{1,2}x_1x_2 + \ldots + 2a_{1,n}x_1x_n$

Se operiamo il cambio di variabili

$$\begin{cases} x_1 = y_1 + by_2 \\ x_i = y_i \text{ per } i \geq 2, \end{cases}$$

l'espressione (19) cambia in

$$\begin{aligned} q_1 &= a_{1,1}(y_1 + by_2)^2 + 2a_{1,2}(y_1 + by_2)y_2 + \ldots + 2a_{1,n}(y_1 + by_2)y_n \\ &= a_{1,1}y_1^2 + 2ba_{1,1}y_1y_2 + b^2a_{1,1}y_2^2 + 2a_{1,2}y_1y_2 + 2ba_{1,2}y_2^2 + \ldots + 2a_{1,n}y_1y_n + 2ba_{1,n}y_2y_n \\ &= a_{1,1}y_1^2 + 2(ba_{1,1} + a_{1,2})y_1y_2 + \ldots + 2a_{1,n}y_1y_n + \left[(b^2a_{1,1} + 2ba_{1,2})y_2^2 + \ldots + 2ba_{1,n}y_2y_n\right]. \end{aligned}$$

Se $a_{1,1} \neq 0$, allora ponendo $b = -\dfrac{a_{1,2}}{a_{1,1}}$ si ottiene $ba_{1,1}+a_{1,2} = 0$, ovvero il coefficiente di y_1y_2 si annulla e quindi

$$q_1 = a_{1,1}y_1^2 + 2a_{1,3}y_1y_3 + \ldots + 2a_{1,n}y_1y_n + \left[ba_{1,2}y_2^2 + \ldots + 2ba_{1,n}y_2y_n.\right].$$

► Se $b = -\dfrac{a_{1,2}}{a_{1,1}}$, *alóra* il coefficiente di y_2^2 diventa

$$b^2a_{1,1} + 2ba_{1,2} = ba_{1,2}.$$

Operando in questo modo $n-1$ volte è possibile quindi far scomparire tutti i termini misti x_1x_i di $q(\boldsymbol{x})$, con $i \geq 2$, e cioè trovare nuove variabili $\hat{\boldsymbol{x}}$ per cui

$$q(\boldsymbol{x}) = \hat{q}(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n) = a_{1,1}\hat{x}_1^2 + q_2(\hat{x}_2, \ldots, \hat{x}_n),$$

dove q_2 è una forma quadratica nelle rimanenti $n-1$ variabili $\hat{x}_2$, $\hat{x}_3$, $\ldots$, $\hat{x}_n$. Naturalmente questo a patto che $a_{1,1} \neq 0$.

Se invece $a_{1,1} = 0$, questo cambio di variabili non funziona. Con una permutazione delle variabili possiamo però cercare un $a_{i,i} \neq 0$, scambiare x_i con x_1 e procedere come sopra. Se invece per ogni $i = 1, \ldots, n$ si ha che $a_{i,i} = 0$, occorre procedere in altro modo. Se $a_{1,j} = 0$ per ogni $j = 2 \ldots n$, allora x_1 non compare in $q(\boldsymbol{x})$ e quindi l'asserto è vero con $c_1 = 0$:

$$q = 0x_1^2 + q_2(y_2, y_3, \ldots, y_n).$$

Se almeno uno degli $a_{1,j}$ è diverso da zero, allora con uno scambio tra due variabili x_1 e x_j possiamo supporre che sia $a_{1,2} \neq 0$. Poniamo $x_1 = y_1 + y_2$, $x_2 = y_1 - y_2$, e $x_i = y_i$ per $i \geq 3$, e la forma quadratica che nelle x_i è

$$\begin{array}{rcccccccccc} q = 0x_1^2 & + & 2a_{1,2}x_1x_2 & + & 2a_{1,3}x_1x_3 & + \ldots + & 2a_{1,n-1}x_1x_{n-1} & + & 2a_{1,n}x_1x_n & + \\ & + & 0x_2^2 & + & 2a_{2,3}x_2x_3 & + \ldots + & 2a_{2,n-1}x_2x_{n-1} & + & 2a_{2,n}x_2x_n & + \\ & & & + & 0x_3^2 & + \ldots + & 2a_{3,n-1}x_3x_{n-1} & + & 2a_{3,n}x_3x_n & + \\ & & & & & & \ldots & & \ldots & \\ & & & & & + & 0x_{n-1}^2 & + & 2a_{n-1,n}x_{n-1}x_n & + \\ & & & & & & & + & 0x_n^2 & \end{array}$$

nelle nuove variabili y_i diventa

$$\begin{aligned} \hat{q} &= 0 + 2a_{1,2}(y_1 + y_2)(y_1 - y_2) + 2(y_1 + y_2)(a_{1,3}y_3 + \ldots + a_{1,n-1}y_{n-1} + a_{1,n}y_n) + \\ &+ 2(y_1 - y_2)(a_{2,3}y_3 + \ldots + a_{2,n-1}y_{n-1} + a_{2,n}x_n) + \ldots + 2a_{3,n-1}y_3y_{n-1} + 2a_{3,n}y_3y_n + \ldots + 2a_{n-1,n}y_{n-1}y_n. \end{aligned}$$

Con la forma quadratica in questa forma, notiamo che y_1^2 compare con coefficiente $2a_{1,2} \neq 0$, e quindi possiamo riprendere a eliminare i prodotti y_1y_j esattamente come abbiamo fatto per il caso $a_{1,1} \neq 0$: nuovamente possiamo trovare alla fine di tutto un sistema di coordinate $\hat{x}_i$ in cui

$$\hat{q} = c_1\hat{x}_1^2 + q_2(\hat{x}_2, \ldots, \hat{x}_n),$$

dove q_2 dipende dalle $n-1$ variabili $\hat{x}_2, \ldots, \hat{x}_n$. ■

Osserviamo che è possibile cambiare ulteriormente le coordinate e rendere l'identità (18) ancora più semplice. Permutando le variabili, possiamo sempre supporre che per certi interi p e r (con $0 \le p \le r \le n$) si abbia

$$\begin{cases} c_i > 0 & \text{se } i = 1, \dots, p \\ c_i < 0 & \text{se } i = p+1, \dots, r \\ c_i = 0 & \text{se } i = r+1, \dots, n. \end{cases}$$

Supponiamo di scrivere la forma quadratica (18) diagonale

$$q = c_1 x_1^2 + c_2 x_2^2 + \dots + c_n x_n^2.$$

Se si cambiano le variabili nel modo seguente:

$$y_i = \begin{cases} \sqrt{c_i} x_i & \text{se } i = 1, \dots, p \\ \sqrt{-c_i} x_i & \text{se } i = p+1, \dots, r \\ x_i & \text{se} i = r+1, \dots, n, \end{cases}$$

la q diventa

$$\hat{q} = y_1^2 + y_2^2 + \dots + y_p^2 - y_{p+1}^2 - \dots - y_r^2.$$

Nel prossimo teorema dimostreremo che questa forma è *canonica*, cioè che è unicamente determinata da q e che lo rappresenta fedelmente.

▶ Per *forma canonica* si intende, in matematica, una forma che assume un oggetto matematico, dopo opportune trasformazioni, che è univocamente determinata dall'oggetto e che rappresenta fedelmente le proprietà dell'oggetto in questione. In un certo senso una "rappresentazione standard" dell'oggetto.

(IV.71) Teorema (Legge di inerzia di Sylvester). *Se q è una forma quadratica su $\mathbb{R}^n$, allora esistono due interi p e r, con $0 \le p \le r \le n$, ed esiste un cambio di variabili $\boldsymbol{y} = B\boldsymbol{x}$ di $\mathbb{R}^n$ tale che nelle nuove coordinate*

(20) $$\hat{q} = y_1^2 + y_2^2 + \dots + y_p^2 - y_{p+1}^2 - \dots - y_r^2.$$

L'intero p è detto indice di positività*, $r - p$ è detto* indice di negatività *e r è il* rango *della forma quadratica. L'indice di positività e il rango della forma quadratica q non dipendono dal sistema di riferimento di $\mathbb{R}^n$.*

▶ Dato che entrambi non dipendono dalla base, è molto semplice vedere che il rango di una forma quadratica qui definito è uguale al rango della matrice che lo rappresenta, Definizione (IV.63), rispetto a qualunque base.

Dimostrazione. Abbiamo visto sopra che ogni forma quadratica si può scrivere come nell'equazione (20). Mostriamo quindi che se la forma quadratica è diagonalizzata in due sistemi di riferimento $\boldsymbol{y}$ e $\boldsymbol{z}$, l'indice di positività e il rango coincidono. In altre parole, vogliamo convincerci che se $\boldsymbol{z} = B\boldsymbol{y}$ per una matrice B invertibile, e

(21) $$y_1^2 + y_2^2 + \dots + y_p^2 - y_{p+1}^2 - \dots - y_r^2 = z_1^2 + z_2^2 + \dots + z_{p'}^2 - z_{p'+1}^2 - \dots - z_{r'}^2,$$

allora $p = p'$ e $r = r'$. Supponiamo invece che sia $p > p'$. Cerchiamo un vettore di $\mathbb{R}^n$ le cui coordinate nelle variabili $\boldsymbol{y}$ soddisfino le $n - p$ uguaglianze $y_{p+1} = y_{p+2} = \dots = y_r = \dots = y_n = 0$, mentre le sue coordinate nelle variabili $\boldsymbol{z}$ soddisfino $z_1 = z_2 = \dots = z_{p'} = 0$. Esiste? Dato che le z_i sono funzioni lineari delle y_j, si tratta di un sistema di $(n-p)+p'$ equazioni nelle n incognite y_i. Ma se $p' < p$ allora $n-p+p' < n$ e il sistema ha certamente almeno una soluzione $\bar{\boldsymbol{y}} \neq \mathbf{0}$. Sia $\bar{\boldsymbol{z}}$ il vettore corrispondente $\bar{\boldsymbol{z}} = B\bar{\boldsymbol{y}}$. Dato che $\bar{y}_{p+1} = \bar{y}_{p+2} = \ldots = \bar{y}_r = \ldots = \bar{y}_n = 0$,

$$(\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_p) \neq \mathbf{0} \implies \bar{y}_1^2 + \bar{y}_2^2 + \ldots + \bar{y}_p^2 > 0.$$

Inoltre per costruzione $\bar{z}_1^2 + \bar{z}_2^2 + \ldots + \bar{z}_{p'}^2 = 0$, e quindi per l'equazione (21) si ha

$$\begin{aligned} 0 < \bar{y}_1^2 + \bar{y}_2^2 + \ldots + \bar{y}_p^2 &= \bar{y}_1^2 + \bar{y}_2^2 + \ldots + \bar{y}_p^2 - \bar{y}_{p+1}^2 - \ldots - \bar{y}_r^2 \\ &= \bar{z}_1^2 + \bar{z}_2^2 + \ldots + \bar{z}_{p'}^2 - \bar{z}_{p'+1}^2 - \ldots - \bar{z}_{r'}^2 \\ &= -\bar{z}_{p'+1}^2 - \ldots - \bar{z}_{r'}^2 \leq 0. \end{aligned}$$

Questo è assurdo, e quindi non può essere $p' < p$. Il ragionamento può essere ripetuto *paro paro* scambiando i ruoli di $\boldsymbol{y}$ e $\boldsymbol{z}$, da cui si deduce che $p = p'$. Per concludere la dimostrazione, osserviamo che se $q(\boldsymbol{x})$ è diagonale, allora $q(\boldsymbol{x})$ e $-q(\boldsymbol{x})$ hanno stesso rango r, mentre l'indice di positività di $-q(\boldsymbol{x})$ è uguale all'indice di negatività $r-p$ di $q(\boldsymbol{x})$. Se q_1 e q_2 sono due diagonalizzazioni della forma q, allora $-q_1$ e $-q_2$ sono due diagonalizzazioni della forma $-q$ e quindi hanno lo stesso indice di positività. Ma allora q_1 e q_2 hanno gli stessi indici di positività e di negatività, e di conseguenza anche lo stesso rango. ■

(IV.72) Definizione. Si dice che una forma quadratica $q(\boldsymbol{x})$ è *definita positiva* se per ogni $\boldsymbol{x} \neq \mathbf{0}$ si ha $q(\boldsymbol{x}) > 0$. È *definita negativa* se per ogni $\boldsymbol{x} \neq \mathbf{0}$ si ha $q(\boldsymbol{x}) < 0$. Se invece della disuguaglianza stretta si pone $q(\boldsymbol{x}) \geq 0$ (risp. $q(\boldsymbol{x}) \leq 0$), allora q è *semidefinita positiva* (risp. *semidefinita negativa*). Altrimenti si dice *indefinita*.

(IV.73) Nota. Se $q(\boldsymbol{x})$ è definita positiva, allora i termini sulla diagonale sono tutti strettamente positivi (cioè $a_{ii} > 0$), e quindi mai nulli. Nel procedimento di diagonalizzazione usato nella dimostrazione del Teorema (IV.70) quindi bastano le permutazioni e le sostituzioni del tipo $x_i = y_i + by_j$. Analogamente se q è definita negativa $a_{ii} < 0$ per ogni i, e se è semidefinita valgono le disuguaglianze larghe. ■

▶ Infatti, basta valutare q nel punto $(1, 0, \ldots, 0)$ per ottenere a_{11}, che deve essere positivo. Lo stesso per tutti gli altri termini sulla diagonale.

▶ Non vale però il viceversa! Non è possibile stabilire queste proprietà conoscendo solo i termini sulla diagonale.

Data una matrice A, le sue sottomatrici quadrate che non contengono le ultime righe e colonne si dicono *sottomatrici principali*. Esse sono

$$A_1 = (a_{11}) \quad A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \ldots$$

▶ A volte le sottomatrici principali sono anche dette *sottomatrici nord-ovest*.

I loro determinanti sono detti *minori principali*:

$$d_1 = a_{11} \quad d_2 = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad d_3 = \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \ldots$$

Un criterio per stabilire la positività di una forma, data la matrice associata, è il seguente.

(IV.74) Proposizione (Criterio di Sylvester). *Una forma quadratica q su $\mathbb{R}^n$ è definita positiva se e soltanto se tutti i* minori principali *$d_1, d_2, \dots, d_n$ della matrice associata a q (cioè i determinanti delle sue sottomatrici principali) sono positivi*

$$d_1 > 0, d_2 > 0, d_3 > 0, \dots$$

È invece definita negativa se

$$d_1 < 0, d_2 > 0, d_3 < 0, \dots$$

Dimostrazione. La dimostrazione non è particolarmente difficile, ma un po' laboriosa. Verrà svolta dal volonteroso (e laborioso) Lettore negli Esercizi **53** e **54** a pagina 229. ∎

§ 6. Autovalori e autovettori

Nel paragrafo precedente abbiamo trovato una *forma canonica* per ogni forma quadratica. Il Teorema (IV.70) infatti, garantisce l'esistenza di una opportuna base di $\mathbb{R}^n$ nella quale la matrice corrispondente alla forma quadratica sia *diagonale*, cioè abbia tutti i coefficienti nulli al di fuori della diagonale principale. Allo stesso modo possiamo porci il problema di *diagonalizzare* una matrice; come abbiamo visto nella Definizione (IV.32), quando esiste una matrice invertibile B tale che $M_2 = B^{-1}M_1B$, le due matrici M_1 e M_2 si dicono *simili*.

(IV.75) Definizione. Una matrice M si dice *diagonalizzabile* se è simile ad una matrice diagonale, cioè se esiste una matrice invertibile B (detta *diagonalizzante*) per cui

$$D = B^{-1}MB \quad \text{è una matrice diagonale.}$$

Vediamo ora cosa significa *diagonalizzare* una funzione lineare f da $\mathbb{R}^n$ a $\mathbb{R}^n$, vale a dire una funzione che si può scrivere come $f\colon \boldsymbol{x} \mapsto M\boldsymbol{x}$ per una opportuna matrice $n \times n$. In modo equivalente, consideriamo uno spazio vettoriale V di dimensione n e $f\colon V \to V$ un operatore lineare. Cercheremo una base $(\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n)$ per V (con la corrispondente parametrizzazione $B\colon \mathbb{R}^n \to V$) tale che la matrice associata a f▶ sia una matrice diagonale D.

▶ Come nella Proposizione (IV.25) di pagina 177.

$$\begin{tikzcd}
V \arrow[r, "f"] & V \\
\mathbb{R}^n \arrow[u, "B", "\cong"'] \arrow[r, "D"'] & \mathbb{R}^n. \arrow[u, "\cong", "B"']
\end{tikzcd}$$

Quando $V = \mathbb{R}^n$, anche B può essere rappresentato come matrice, e quindi diagonalizzare f significa trovare una matrice invertibile B (le cui colonne sono gli n vettori $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, ..., $\boldsymbol{b}_n$) per cui le funzioni lineari associate alle matrici sono rappresentabili come nel seguente diagramma

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "M"] & \mathbb{R}^n \\ \mathbb{R}^n \arrow[u, "B"', "\cong"] \arrow[r, "D"'] & \mathbb{R}^n, \arrow[u, "B", "\cong"'] \end{tikzcd}$$

cioè trovare una matrice invertibile B tale che $D = B^{-1}MB$ è una matrice diagonale. Se è possibile diagonalizzare un operatore f con la scelta di una base $(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n)$, esso sarà detto *diagonalizzabile* e la base verrà detta *base diagonalizzante* (o, se $V = \mathbb{R}^n$, che f è diagonalizzato dalla matrice B che ha per colonne i vettori della base $\boldsymbol{b}_i$).

(IV.76) Esempio. Riprendiamo l'operatore $Q\colon \mathbb{R}^2 \to \mathbb{R}^2$ degli Esempi (IV.23) e (IV.33):

$$(*) \qquad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Come possiamo diagonalizzarlo? Occorre trovare una matrice invertibile $S = (\boldsymbol{b}_1, \boldsymbol{b}_2)$ tale che

$$S^{-1}QS = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$$

per certi coefficienti d_1 e d_2, o equivalentemente trovare una base $(\boldsymbol{b}_1, \boldsymbol{b}_2)$ di $\mathbb{R}^2$ per cui l'operatore è rappresentato da una matrice diagonale D. Ma $S^{-1}QS = D$ se e solo se $QS = SD$, cioè se e solo se $(Q\boldsymbol{b}_1, Q\boldsymbol{b}_2) = (Sd_1\boldsymbol{e}_1, Sd_2\boldsymbol{e}_2)$, dove $\boldsymbol{e}_1 = (1,0)$ e $\boldsymbol{e}_2 = (0,1)$. Dato che $S\boldsymbol{e}_1 = \boldsymbol{b}_1$ e $S\boldsymbol{e}_2 = \boldsymbol{b}_2$, dobbiamo trovare perciò due vettori indipendenti $\boldsymbol{b}_1$ e $\boldsymbol{b}_2$ che soddisfino l'equazione $Q\boldsymbol{b}_i = d_i\boldsymbol{b}_i$ per $i = 1, 2$. Prima di proseguire con questo esempio, introduciamo un po' di nozioni e notazioni. ■

(IV.77) Definizione. Sia V uno spazio vettoriale e $L\colon V \to V$ un operatore lineare. Un vettore $\boldsymbol{v} \in V$, $\boldsymbol{v} \neq \boldsymbol{0}$ è detto *autovettore* di L se esiste uno scalare $\lambda \in \mathbb{R}$ per cui $L(\boldsymbol{v}) = \lambda\boldsymbol{v}$. Lo scalare λ si dice *autovalore* di L relativo a $\boldsymbol{v}$. L'insieme $\{\lambda_1, \lambda_2, \ldots\}$ di tutti gli autovalori di f è lo *spettro* di f.

▶ Un autovalore di L è quindi uno scalare per cui esista un autovettore di cui è l'autovalore, cioè uno scalare λ per cui esista $\boldsymbol{v} \neq \boldsymbol{0}$ tale che $L(\boldsymbol{v}) = \lambda\boldsymbol{v}$. Non è difficile dimostrare che per ogni autovettore esiste un unico autovalore: si veda l'Esercizio 32.

Nell'Esempio (IV.76), ci basta quindi trovare due autovettori linearmente indipendenti di Q. Più in generale, consideriamo una matrice diagonale $n \times n$

$$D = \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{bmatrix}$$

Allora tutti i vettori della base standard $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ sono autovettori con autovalori $d_{11}, d_{22}, \ldots, d_{nn}$:

$$D\boldsymbol{e}_1 = \begin{bmatrix} d_{11} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = d_{11}\boldsymbol{e}_1, \quad D\boldsymbol{e}_2 = \begin{bmatrix} 0 \\ d_{22} \\ \vdots \\ 0 \end{bmatrix} = d_{22}\boldsymbol{e}_2, \quad \ldots \quad D\boldsymbol{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ d_{nn} \end{bmatrix} = d_{nn}\boldsymbol{e}_n.$$

La seguente proposizione illustra il passo chiave per comprendere il legame tra il problema della diagonalizzazione e quello degli autovettori/autovalori.

(IV.78) Proposizione. *Un operatore lineare $L\colon V \to V$ è diagonalizzabile se e soltanto se esiste in V una base $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n$ di autovettori per L. In tal caso, gli autovalori di L sono proprio i coefficienti della matrice diagonale, che si scrive come*

$$\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

► Come si può vedere ad esempio nell'Esercizio (IV.5) a pagina 218, gli operatori lineari (a differenza delle forme quadratiche) non sono tutte diagonalizzabili.

Dimostrazione. Basta mostrare che una base è diagonalizzante se e solo se è formata da autovettori. Ma se l'operatore L nella base $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, …, $\boldsymbol{b}_n$ è diagonalizzato, allora $L\boldsymbol{b}_i = d_{ii}\boldsymbol{b}_i$ (cioè i $\boldsymbol{b}_i$ sono autovettori rispetto agli autovalori $\lambda_i = d_{ii}$, i termini sulla diagonale della matrice associata), e quindi la sua matrice associata è la matrice con gli autovalori λ_i sulla diagonale e zero altrove. Viceversa, se gli autovettori $\boldsymbol{b}_i$ costituiscono una base, questa è chiaramente diagonalizzante. ■

Ora riprendiamo l'Esempio (IV.76), e cerchiamo gli autovettori, cioè le soluzioni dell'equazione

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \Longleftrightarrow \quad \begin{cases} x_1 + x_2 = \lambda x_1 \\ \qquad x_1 = \lambda x_2. \end{cases}$$

Eliminiamo x_1 con la seconda, e sostituendo nella prima otteniamo

$$\lambda x_2 + x_2 = \lambda^2 x_2 \iff (\lambda + 1)x_2 = \lambda^2 x_2 \iff (1 + \lambda - \lambda^2)x_2 = 0.$$

Se $1 + \lambda - \lambda^2 \neq 0$, allora l'unica soluzione è $x_2 = 0$, da cui segue $x_1 = 0$: niente autovettori. Se invece $1 + \lambda - \lambda^2 = 0$, allora si può porre $x_2 = 1$, e $x_1 = \lambda$ per risolvere le equazioni. Abbiamo due valori λ_1 e λ_2 che annullano il polinomio $1 + \lambda - \lambda^2$:

$$\lambda_1 = (1 + \sqrt{5})/2, \quad \lambda_2 = (1 - \sqrt{5})/2.$$

Seguono i due corrispondenti autovettori

$$\boldsymbol{b}_1 = ((1 + \sqrt{5})/2, 1), \quad \boldsymbol{b}_2 = ((1 - \sqrt{5})/2, 1).$$

In effetti si tratta di una base diagonalizzante per Q, cioè se $S = (\boldsymbol{b}_1, \boldsymbol{b}_2)$ si ha

$$S^{-1}QS = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

▶ Il Lettore è invitato di svolgere i calcoli e verificare quanto scritto. In premio potrebbe trovare alcune interessanti relazioni con i numeri di Fibonacci e le potenze di Q, se prima riesce a dimostrare per induzione che per ogni $n \geq 1$

$$S^{-1}Q^nS = (S^{-1}QS)^n.$$

Procediamo nel generalizzare questa procedura e dimostrare alcune proposizioni.

(IV.79) *Se $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, …, $\boldsymbol{v}_k$ sono k autovettori di un operatore L corrispondenti a k autovalori* distinti*, allora $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, …, $\boldsymbol{v}_k$ sono linearmente indipendenti.*

Dimostrazione. Per induzione: se $k = 1$, un vettore è linearmente indipendente se è diverso da zero, e questo è vero dato che gli autovettori sono per definizione diversi da zero. Supponiamo quindi che $k - 1$ autovettori di L corrispondenti a $k - 1$ autovalori distinti siano sempre linearmente indipendenti. Allora se

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_k\boldsymbol{v}_k = \boldsymbol{0}, \tag{22}$$

visto che L è lineare,

$$\begin{aligned} \boldsymbol{0} &= L(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_k\boldsymbol{v}_k) = c_1L(\boldsymbol{v}_1) + c_2L(\boldsymbol{v}_2) + \ldots + c_kL(\boldsymbol{v}_k) \\ &= c_1\lambda_1\boldsymbol{v}_1 + c_2\lambda_2\boldsymbol{v}_2 + \ldots + c_k\lambda_k\boldsymbol{v}_k, \end{aligned}$$

dove $\lambda_1, \ldots, \lambda_k$ sono i k autovalori distinti. Moltiplicando la (22) per λ_k e sottraendola a quest'ultime, si ottiene

$$c_1(\lambda_1 - \lambda_k)\boldsymbol{v}_1 + c_2(\lambda_2 - \lambda_k)\boldsymbol{v}_2 + \ldots + c_k(\lambda_k - \lambda_k)\boldsymbol{v}_k = \boldsymbol{0},$$

cioè, ponendo $c'_i = (\lambda_i - \lambda_k)c_i$,

$$c'_1\boldsymbol{v}_1 + c'_2\boldsymbol{v}_2 + \ldots + c'_{k-1}\boldsymbol{v}_{k-1} = \boldsymbol{0}.$$

Per l'ipotesi di induzione, gli autovettori $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, …,$\boldsymbol{v}_{k-1}$ sono linearmente indipendenti, e quindi per $i = 1, \ldots, k - 1$ si ha $c'_i = 0$. Ma i λ_i sono distinti, e quindi $\lambda_i - \lambda_k \neq 0$ per $i \leq k - 1$, e quindi $c_i = c'_i/(\lambda_i - \lambda_k) = 0$. Se tutti i c_i sono zero fino a $k - 1$, anche c_k è zero, visto che $\boldsymbol{v}_k \neq \boldsymbol{0}$. ■

(IV.80) Corollario. *Un operatore lineare su $\mathbb{R}^n$ non può avere più di n autovalori distinti.*

Dimostrazione. Non ci possono essere più di n vettori linearmente indipendenti in uno spazio vettoriale di dimensione n, come è stato provato nel Lemma (IV.14). ∎

(IV.81) Corollario. *Se un operatore lineare su $\mathbb{R}^n$ ha n autovalori distinti, allora è diagonalizzabile.*

Dimostrazione. Se ha n autovalori distinti, allora ha n autovettori linearmente indipendenti in $\mathbb{R}^n$, che costituiscono una base e quindi è diagonalizzabile. ∎

Il prossimo teorema è alla base del metodo per determinare gli autovalori per matrici abbastanza piccole.

(IV.82) Teorema (Polinomio caratteristico)**.** *Sia $L\colon V \to V$ un operatore, e M la matrice associata a L rispetto ad una certa base. Allora la funzione*

$$p_M(\lambda) = \det(M - \lambda I) = \det \begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{bmatrix}$$

è un polinomio di n-esimo grado, detto polinomio caratteristico *di M. Gli autovalori di L sono tutte e sole le radici del polinomio caratteristico $p_M(\lambda)$. Matrici simili hanno il medesimo polinomio caratteristico, e quindi il polinomio caratteristico dipende solo da L e non dalla base scelta.*

▶ Se la matrice M è triangolare (inferiore, superiore o addirittura diagonale) allora segue da (IV.82) che i suoi autovalori sono tutti e soli i termini m_{ii} sulla diagonale. Infatti se M è triangolare anche $M - \lambda I$ lo è e quindi il suo determinante è il prodotto dei binomi $(m_{ii} - \lambda)$. Inoltre è possibile dimostrare (ragionando con la dovuta cautela, almeno nel caso in cui il polinomio caratteristico abbia radici complesse non reali) che il prodotto degli autovalori di M è uguale al suo determinate, mentre la loro somma è uguale alla somma degli elementi sulla diagonale (la cosiddetta *traccia* di M — cfr. 38 a pagina 227). Si vedano in particolare gli Esercizi 35 (pagina 227) e 37 (pagina 227).

☢ *Dimostrazione.* Una parte della dimostrazione è piuttosto semplice:

$$\begin{aligned} \lambda \text{ è un autovalore } &\iff Mv = \lambda I v \text{ ha soluzioni non nulle} \\ &\iff (M - \lambda I)v = \mathbf{0} \text{ ha soluzioni non nulle} \\ &\iff \det(M - \lambda I) = 0 \end{aligned}$$

(abbiamo usato l'equivalenza ① $\iff$ ② nel Teorema (IV.59)). La parte difficile è lasciata per esercizio. ∎

▶ Una dimostrazione assistita completa di questo teorema si trova negli Esercizi 35 e 37 a pagina 227.

Diagonalizzazione mediante trasformazioni ortogonali

Alcune volte si riesce a diagonalizzare una matrice con un cambio di coordinate che non solo è invertibile, ma è anche una trasformazione ortogonale. Quando? Il Teorema che risponde a questa domanda è il seguente.

(IV.83) Teorema (Spettrale). *Una matrice M è diagonalizzabile mediante trasformazioni ortogonali se e soltanto se è simmetrica. In altre parole, esiste almeno una matrice ortogonale B tale che $D = B^{-1}MB$ è diagonale se e solo se M è simmetrica.*

▶ Ricordiamo che stiamo sempre considerando coefficienti reali e non complessi.

Dimostrazione. Cominciamo a dimostrare la parte "soltanto se": se $B^{-1}MB$ è diagonale e B è una matrice ortogonale, cioè $B^{-1} = {}^tB$, allora

$$ {}^tM = {}^t\left(BDB^{-1}\right) = {}^t\left(BD{}^tB\right) = B{}^tD{}^tB = BDB^{-1} = M, $$

e quindi M è simmetrica.

Per la parte "se", procediamo per induzione. Se $n = 1$, ogni matrice è diagonale, quindi non c'è nulla da provare. Supponiamo che $M = (m_{ij})$ sia simmetrica $n \times n$, e allo stesso tempo che ogni matrice $k \times k$ con $k < n$ e simmetrica sia diagonalizzabile con una matrice ortogonale $k \times k$. Supponiamo per il momento che esista un autovettore $\boldsymbol{b}_1$ di M con $\|\boldsymbol{b}_1\|^2 = 1$ e autovalore λ_1; dovremmo da qui costruire una matrice ortogonale B che abbia $\boldsymbol{b}_1$ come prima colonna (cioè tale che $B\boldsymbol{e}_1 = \boldsymbol{b}_1$), come abbiamo visto sopra con il procedimento di Gram–Schmidt. Allora $B^{-1}MB$ ha per prima colonna $\lambda_1\boldsymbol{e}_1 = (\lambda_1, 0, \dots, 0)$, dato che

$$ B^{-1}MB\boldsymbol{e}_1 = B^{-1}M\boldsymbol{b}_1 = B^{-1}\lambda_1\boldsymbol{b}_1 = \lambda_1\boldsymbol{e}_1. $$

Non solo, dato che $B^{-1} = {}^tB$ (B è ortogonale) e M è simmetrica, anche $B^{-1}MB$ è simmetrica e quindi anche la prima riga di $B^{-1}MB$ è composta da $\lambda_1{}^t\boldsymbol{e}_1$, cioè $B^{-1}MB$ è del tipo

$$ \begin{bmatrix} \boxed{\lambda_1} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} $$

Ma allora per ipotesi di induzione esiste una matrice $P = (p_{ij})$ ortogonale $(n-1) \times (n-1)$ tale che

$$ P^{-1} \begin{bmatrix} a_{22} & a_{23} & \cdots & a_{1n} \\ a_{32} & a_{33} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} P = \begin{bmatrix} \lambda_2 & 0 & \cdots & 0 \\ 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} $$

La matrice ottenuta da P nel modo seguente

$$ \hat{P} = \begin{bmatrix} 1 & {}^t\mathbf{0} \\ \mathbf{0} & P \end{bmatrix} = \begin{bmatrix} \boxed{1} & 0 & 0 & \cdots & 0 \\ 0 & p_{1,1} & p_{1,2} & \cdots & a_{2,n-1} \\ 0 & p_{2,1} & p_{22} & \cdots & p_{2,n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & p_{n-1,1} & p_{n-1,2} & \cdots & p_{n-1,n-1} \end{bmatrix} $$

ha la proprietà cercata di diagonalizzare la matrice $B^{-1}MB$

$$\hat{P}^{-1}B^{-1}MB\hat{P} = \text{diagonale}.$$

La trasformazione ortogonale cercata è la composizione di B e $\hat{P}$, che è ancora ortogonale in quanto prodotto di trasformazioni ortogonali. Non rimane che mostrare che esiste un autovettore $\boldsymbol{b}_1$: questa è una conseguenza del prossimo Lemma (IV.84). Infatti, il Teorema Fondamentale dell'Algebra ci assicura che ogni polinomio, e quindi in particolare il polinomio caratteristico $p_M(\lambda)$, ha almeno una radice λ_1 in $\mathbb{C}$. Ora, se questa è in $\mathbb{R} \subset \mathbb{C}$, allora si tratta di λ_1, autovalore di M, cui deve corrispondere un autovettore $\boldsymbol{x}_1$. ■

(IV.84) Lemma. *Il polinomio caratteristico di una matrice simmetrica M ha tutte le radici reali.*

☢ *Dimostrazione.* La matrice M ha coefficienti reali, ma può essere utilizzata per definire un operatore $\mathbb{C}^n \to \mathbb{C}^n$. Se λ è una radice in $\mathbb{C}$ del polinomio caratteristico della matrice simmetrica $M = {}^t\!M$, cioè $\det(M - \lambda I) = 0$, allora esiste un autovettore $\boldsymbol{x} \in \mathbb{C}^n$ tale che $M\boldsymbol{x} = \lambda\boldsymbol{x}$. Ora, questo vuol dire che le componenti di $M\boldsymbol{x}$ sono uguali alle componenti di $\lambda\boldsymbol{x}$, cioè per ogni $i = 1, \ldots, n$

▶ Si veda il riquadro a pagina 215. per qualche breve dettaglio in più.

$$\lambda x_i = \sum_{j=1}^{n} m_{ij} x_j.$$

Moltiplichiamo per il coniugato di x_i e sommiamo su i:

$$\lambda \sum_{i=1}^{n} \underbrace{\|x_i\|^2}_{\in\mathbb{R}} = \sum_{i,j=1}^{n} m_{ij} x_j \bar{x}_i = \sum_{i=1}^{n} \underbrace{m_{ii}\|x_i\|^2}_{\in\mathbb{R}} + \sum_{1\le i<j\le n} \underbrace{m_{ij}(x_j\bar{x}_i + \bar{x}_j x_i)}_{\in\mathbb{R}} \in \mathbb{R}.$$

▶ Se $z = x_j\bar{x}_i \in \mathbb{C}$, allora $z + \bar{z} \in \mathbb{R}$.

Segue pertanto che $\lambda \in \mathbb{R}$. ■

(IV.85) Corollario. *Se $q(\boldsymbol{x})$ è una forma quadratica su $\mathbb{R}^n$, allora $q(\boldsymbol{x})$ può essere diagonalizzata con una trasformazione ortogonale, e i coefficienti sulla diagonale sono gli autovalori della matrice associata a q.*

Dimostrazione. Sa A la matrice (simmetrica) associata a q, così che $q(\boldsymbol{x}) = {}^t\boldsymbol{x}A\boldsymbol{x}$. Se S è una matrice ortogonale tale che $S^{-1}AS$ è diagonale, allora $S^{-1} = {}^tS$ ▶ e quindi ${}^tSAS$ è diagonale, cioè nelle coordinate $\hat{\boldsymbol{x}}$ tali che $\boldsymbol{x} = S\hat{\boldsymbol{x}}$ la forma quadratica $q(\boldsymbol{x})$ è diagonale: $\hat{q}(\hat{\boldsymbol{x}}) = d_1\hat{x}_1^2 + \ldots + d_n\hat{x}_n^2$. Gli autovalori di A sono per definizione i termini della matrice diagonale $S^{-1}AS$, e quindi i coefficienti $d_1, \ldots, d_n$ sono proprio gli autovalori di A. ■

▶ Se S è ortogonale, allora $S{}^tS = {}^tSS = I \implies S^{-1} = {}^tS$.

(IV.86) Corollario. *Una forma quadratica $q(\boldsymbol{x})$ è definita positiva (risp. negativa) se e soltanto se i suoi autovalori* ▶ *sono tutti positivi (risp. negativi). È semidefinita positiva (risp. negativa) se e soltanto se i suoi autovalori sono tutti positivi (risp. negativi) o nulli. È indefinita se ci sono autovalori sia positivi che negativi.*

▶ Naturalmente qui si intende che gli autovalori di una forma quadratica sono gli autovalori della matrice simmetrica associata.

Dimostrazione. Sa A la matrice (simmetrica) associata a q. Per il corollario precedente, in una opportuna base ortogonale la forma $q(\boldsymbol{x})$ si diagonalizza come

$$\hat{q}(\hat{\boldsymbol{x}}) = d_1\hat{x}_1^2 + \ldots + d_n\hat{x}_n^2,$$

dove i $d_1, \ldots, d_n$ sono gli autovalori di A. Allora se sono tutti positivi (risp. negativi), q è definita positiva (risp. negativa). Se sono positivi (risp. negativi) o nulli, q è semidefinita positiva (risp. negativa). Se ce ne sono sia positivi che negativi, allora è indefinita. ■

(IV.87) Osservazione. Il fatto fondamentale che consente di dimostrare i Corollari (IV.85) e (IV.86) è che ogni forma quadratica ha una matrice simmetrica A associata, ed A è sempre diagonalizzabile mediante una trasformazione ortogonale (che è la matrice le cui colonne sono appunti gli autovettori di A). Allora ogni forma quadratica ha associata non solo una matrice A, ma in effetti un sistema completo di autovalori ed autovettori, e se $q(\boldsymbol{x})$ è in forma diagonale

$$q(\boldsymbol{x}) = d_1x_1^2 + d_2x_2^2 + \ldots + d_nx_n^2$$

i suoi *coefficienti* d_i sono anche i suoi *autovalori*. ■

Concludiamo questo paragrafo con una definizione: un operatore L definito da V in sé si dice *auto-aggiunto* (o *simmetrico*) rispetto ad un prodotto scalare $\langle \cdot, \cdot \rangle$ su V se per ogni $\boldsymbol{x}, \boldsymbol{y} \in V$ si ha

$$\langle L\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, L\boldsymbol{y} \rangle.$$

► Non è difficile vedere che un operatore $\mathbb{R}^n \to \mathbb{R}^n$ è auto-aggiunto se e soltanto se è rappresentato da una matrice simmetrica: si veda anche la nota a pagina 224.

Per $\mathbb{R}^n$ la seguente proposizione è una conseguenza diretta del Teorema Spettrale. Ne diamo invece una dimostrazione diretta, che vale anche in spazi di dimensione infinita (si veda l'Esempio (IV.89).

(IV.88) *Se $\lambda_1 \neq \lambda_2$ sono due autovalori distinti di un operatore auto-aggiunto (simmetrico) $L\colon V \to V$, allora i corrispondenti autovettori $\boldsymbol{x}_1$ e $\boldsymbol{x}_2$ sono ortogonali.*

Dimostrazione. Si ha

$$\lambda_1(\boldsymbol{x}_1 \cdot \boldsymbol{x}_2) = (\lambda_1\boldsymbol{x}_1) \cdot \boldsymbol{x}_2 = (L\boldsymbol{x}_1) \cdot \boldsymbol{x}_2 = \boldsymbol{x}_1 \cdot (L\boldsymbol{x}_2) = \boldsymbol{x}_1 \cdot (\lambda_2\boldsymbol{x}_2) = \lambda_2(\boldsymbol{x}_1 \cdot \boldsymbol{x}_2),$$

e dunque

$$(\lambda_1 - \lambda_2)(\boldsymbol{x}_1 \cdot \boldsymbol{x}_2) = 0 \implies \boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = 0,$$

visto che $\lambda_1 \neq \lambda_2$. ■

(IV.89) Esempio. Sia $C(0, \pi)$ lo spazio delle funzioni continue sull'intervallo $[0, \pi]$, con prodotto scalare

$$\langle f, g \rangle = \int_0^\pi f(x)g(x)\,\mathrm{d}x.$$

Sia V il sottospazio di tutte le funzioni C^2 su $[0,\pi]$ e tali che $f(0) = f(\pi) = 0$. Consideriamo l'operatore $L\colon V \to C(0,\pi)$ di derivazione seconda:

$$L\colon f \in V \mapsto L(f) = f''.$$

È un operatore lineare auto-aggiunto. Infatti

$$\begin{aligned}\langle f, Lg\rangle &= \int_0^\pi f(x)g''(x)\,\mathrm{d}x = [f(x)g'(x)]_0^\pi - \int_0^\pi f'(x)g'(x)\,\mathrm{d}x \\ &= -\int_0^\pi f'(x)g'(x)\,\mathrm{d}x \\ &= [f'(x)g(x)]_0^\pi - \int_0^\pi f'(x)g'(x)\,\mathrm{d}x = \int_0^\pi f''(x)g(x)\,\mathrm{d}x \\ &= \langle Lf, g\rangle.\end{aligned}$$

Quali sono gli autovalori/autovettori di L? Le soluzioni dell'equazione

$$y''(x) = \lambda y(x)$$

con la condizione al bordo $y(0) = y(\pi) = 0$, cioè le funzioni

$$y_n(x) = \sin nx, \quad \text{per } n \in \mathbb{Z} \setminus \{0\}.$$

L'autovalore corrispondente a $y_n(x) = \sin nx$ è $\lambda_n = -n^2$. Ora, se $n^2 \neq m^2$, i due autovalori λ_n e λ_m sono distinti, per cui la Proposizione (IV.88) ci permette di concludere $y_n(x)$ e $y_m(x)$ sono ortogonali, cioè che

$$0 = \langle \sin nx, \sin mx\rangle = \int_0^\pi \sin nx \sin mx\,\mathrm{d}x. \qquad ■$$

► Queste sono alcune delle relazioni di ortogonalità dell'Esempio (IV.50) e della Proposizione (II.27).

Esercizi

(IV.1) Mostrare che se $b_1, b_2, \ldots, b_n$ sono numeri reali distinti, allora le funzioni

$$f_1(x) = e^{b_1 x}, \quad f_2(x) = e^{b_2 x}, \quad \ldots, f_n(x) = e^{b_n x},$$

sono indipendenti (nello spazio vettoriale di tutte le funzioni reali).

Soluzione. A meno di permutare gli indici, possiamo supporre che $b_1 \leq b_2 \leq \ldots \leq b_n$. Supponiamo quindi per assurdo che ci siano n coefficienti non tutti nulli $a_1, a_2, \ldots, a_n$ tali che

$$a_1 e^{b_1 x} + a_2 e^{b_2 x} + \ldots + a_n e^{b_n x} = 0.$$

Spazi vettoriali complessi

Uno spazio vettoriale complesso è uno spazio vettoriale in cui il campo degli scalari, invece che $\mathbb{R}$, è il campo dei numeri complessi $\mathbb{C}$. In altre parole, è un insieme di vettori di cui è possibile calcolare la *somma* e il *prodotto* per uno scalare complesso, in modo tale che le proprietà di pagina 167 (la Definizione (IV.1)) siano verificate. Lo spazio di tutte le n-uple di numeri complessi $\mathbb{C}^n$ è uno spazio vettoriale complesso. Praticamente tutto quello che è stato detto per operatori lineari, matrici, dipendenza lineare, sistemi di equazioni, autovalori, autovettori, determinante, . . . può essere ripetuto parola per parola nel campo complesso. C'è una differenza invece a proposito di prodotto scalare e norma. In uno spazio vettoriale complesso, il prodotto scalare deve verificare le seguenti proprietà, invece che quelle per gli spazi con coefficienti reali, della Definizione (IV.34) (le parti contrassegnate con un asterisco sono quelle diverse): il prodotto scalare $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ di due vettori di uno spazio vettoriale complesso è un numero complesso, tale che

- * ① (simmetria Hermitiana) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle}$;
- ② (distributivo/lineare) $\langle \boldsymbol{x}, \boldsymbol{x}+\boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{x}, \boldsymbol{z} \rangle$;
- * ③ (omogeneo) $\langle c\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \bar{c}\boldsymbol{y} \rangle = c\langle \boldsymbol{x}, \boldsymbol{y} \rangle$;
- ④ (definito positivo) $\boldsymbol{x} \neq \boldsymbol{0} \implies \langle \boldsymbol{x}, \boldsymbol{x} \rangle > 0$.

Si può ancora definire la norma $\|x\|^2 = \langle \boldsymbol{x}, \boldsymbol{x} \rangle$, che è reale dato che per la simmetria Hermitiana deve essere $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = \overline{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$. Un prodotto di questo tipo si dice *prodotto scalare hermitiano*, e lo spazio V su cui è definito si dice *spazio hermitiano* o *unitario*. Il prodotto scalare standard di $\mathbb{C}^n$ è

$$\boldsymbol{x} \cdot \boldsymbol{y} = \langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n.$$

Gli operatori lineari su spazi vettoriali complessi sono molto importanti per la meccanica quantistica e le svariate applicazioni. Vedremo alcuni esempi negli esercizi a pagina 229.

Ma allora, se dividiamo per $e^{b_n x}$ e consideriamo il limite per $x \to +\infty$, si ha che

$$\lim_{x\to+\infty} \left[a_1 e^{(b_1-b_n)x} + a_2 e^{(b_2-b_n)x} + \ldots + a_n\right] = 0,$$

cioè $a_n = 0$, visto che tutti i termini precedenti tendono a zero. Ripetendo (in modo ricorsivo) il ragionamento sugli $n-1$ coefficienti $a_1, \ldots, a_{n-1}$, si arriva a mostrare che devono essere tutti nulli.

FIGURA IV.5: Vladimir Andreevich Steklov ([illegible], 1864–1926), i cui contributi scientifici principali sono legati a insiemi ortogonali di funzioni

(IV.2) Sia $V = C(0,1)$ lo spazio delle funzioni continue sull'intervallo $[0,1]$. Definiamo il prodotto scalare

$$\langle f, g \rangle = \int_0^1 f(x)g(x)x \,\mathrm{d}x$$

Calcolare $\|f\|$, dove $f(x) = ax+b$ è un polinomio di primo grado. Determinare poi un polinomio di primo grado ortogonale alla funzione 1 (costante).

Soluzione. La norma al quadrato di f è

$$\|f\|^2 = \langle f, f \rangle = \int_0^1 (f(x))^2 \,\mathrm{d}x = \int_0^1 (ax+b)^2 \,\mathrm{d}x$$
$$= \frac{1}{3a}\left[(a+b)^3 - b^3\right] = \frac{1}{3}a^2 + ab + b^2,$$

e quindi $\|f\| = \sqrt{\frac{1}{3}a^2 + ab + b^2}$. I polinomi di primo grado ortogonali alla funzione 1 sono quelli per cui

$$0 = \langle f(x), 1\rangle = \int_0^1 (ax+b)\,\mathrm{d}x = \frac{a}{2} + b,$$

cioè $f(x) = ax - a/2 = a(x - 1/2)$, che sono i multipli di $x - 1/2$.

(IV.3) Determinare per quali valori dei parametri a, b e c esiste l'inversa della matrice

$$A = \begin{bmatrix} a & b & c \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix},$$

e quando esiste calcolarla.

Soluzione. Per la Proposizione (IV.61), l'inversa esiste se e soltanto se $\det(A) \neq 0$, cioè se e soltanto se

$$0 \neq \begin{vmatrix} a & b & c \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{vmatrix} = a \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} - b \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} + c \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix} = -a + b - c.$$

Supponiamo quindi che $a - b + c \neq 0$, e vediamo di costruire l'inversa, che sappiamo esistere. Possiamo fare uso dell'ultima parte della dimostrazione del Teorema (III.29), in cui viene usato il *metodo di Cramer* per risolvere il sistema $A\boldsymbol{x} = \boldsymbol{b}$, osservando che se B è l'inversa di A, allora $AB = I$ e quindi le sue colonne $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{b}_3$ verificano le equazioni $A\boldsymbol{b}_i = \boldsymbol{e}_i$ per $i = 1, 2, 3$. Se $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ e $\boldsymbol{a}_3$ sono le colonne di A, si ha

▶ Dove gli $\boldsymbol{e}_i$ sono i vettori della base standard di $\mathbb{R}^3$. Stiamo dicendo in generale che invertire una matrice $n{\times}n$ A significa *risolvere gli n sistemi lineari $A\boldsymbol{x} = \boldsymbol{e}_i$*, per $i = 1, \ldots, n$.

$$x\boldsymbol{a}_1 + y\boldsymbol{a}_2 + z\boldsymbol{a}_3 = \boldsymbol{e}_i \iff \begin{cases} x = \dfrac{\det(\boldsymbol{e}_i, \boldsymbol{a}_2, \boldsymbol{a}_3)}{\det A} \\ y = \dfrac{\det(\boldsymbol{a}_1, \boldsymbol{e}_i, \boldsymbol{a}_3)}{\det A} \\ z = \dfrac{\det(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{e}_i)}{\det A}, \end{cases}$$

per $i = 1, 2, 3$. Si avrà quindi che le componenti della prima colonna di B

sono

$$b_{11} = \frac{\begin{vmatrix} 1 & b & c \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{vmatrix}}{\det(A)} = (-1)^0 \frac{\begin{vmatrix} 1 & b & c \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{vmatrix}}{\det A} = -\frac{1}{\det(A)}$$

$$b_{21} = \frac{\begin{vmatrix} a & 1 & c \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix}}{\det A} = (-1)^1 \frac{\begin{vmatrix} a & 1 & c \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix}}{\det A} = \frac{1}{\det A}$$

$$b_{31} = \frac{\begin{vmatrix} a & b & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{vmatrix}}{\det A} = (-1)^2 \frac{\begin{vmatrix} a & b & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{vmatrix}}{\det A} = -\frac{1}{\det A}$$

Le componenti della seconda e della terza colonna si calcolano in modo analogo. Si trova infine

$$B = A^{-1} = \frac{1}{-a+b-c} \begin{bmatrix} -1 & c & b-c \\ 1 & -c & -a \\ -1 & b-a & a \end{bmatrix}$$

▶ Il *pattern* che speriamo sia riuscito ad osservare è il seguente: il coefficiente b_{ij} dell'inversa è uguale al determinante della matrice che si ottiene cancellando la j-esima riga e la i-esima colonna, moltiplicato per $(-1)^{i+j}$ e diviso per il determinante $\det A$. Si può ricavare facilmente una formula per l'inversa di una matrice. Ci sono molti altri modi per calcolare A^{-1}, e si invita il Lettore interessato a cercarne qualcun altro.

☆ **(IV.4)** Mostrare che se P è una matrice ortogonale $n \times n$ e M una matrice quadrata $n \times n$, allora la somma dei quadrati dei coefficienti di M è uguale alla somma dei quadrati dei coefficienti di PM o MP:

$$||M||^2 = \sum_{i,j} m_{ij}^2 \qquad \Longrightarrow \qquad ||M||^2 = ||PM||^2 = ||MP||^2.$$

Soluzione. Osserviamo che $||M||^2$ è la somma delle norme al quadrato delle colonne di M. Se quindi $\boldsymbol{m}_1, \boldsymbol{m}_2, \ldots, \boldsymbol{m}_n$ sono le colonne di M, si ha

$$||M||^2 = ||\boldsymbol{m}_1||^2 + ||\boldsymbol{m}_2||^2 + \ldots + ||\boldsymbol{m}_n||^2.$$

Se P è una matrice ortogonale, allora le sue colonne costituiscono una base ortonormale per $\mathbb{R}^n$, cioè ${}^tPP = P\,{}^tP = I$. Quindi per ogni $\boldsymbol{m}_i$ si ha

$$\begin{aligned} ||\boldsymbol{m}_i||^2 &= \boldsymbol{m}_i \cdot \boldsymbol{m}_i = {}^t\boldsymbol{m}_i\boldsymbol{m}_i \\ ||P\boldsymbol{m}_i||^2 &= (P\boldsymbol{m}_i)\cdot(P\boldsymbol{m}_i) = {}^t(P\boldsymbol{m}_i)(P\boldsymbol{m}_i) = {}^t\boldsymbol{m}_i\,{}^tPP\boldsymbol{m}_i = {}^t\boldsymbol{m}_i\boldsymbol{m}_i, \end{aligned}$$

cioè $||\boldsymbol{m}_i||^2 = ||P\boldsymbol{m}_i||^2$. Ma allora

$$\begin{aligned} ||PM||^2 &= ||P\boldsymbol{m}_1||^2 + ||P\boldsymbol{m}_2||^2 + \ldots + ||P\boldsymbol{m}_n||^2 \\ &= ||\boldsymbol{m}_1||^2 + ||\boldsymbol{m}_2||^2 + \ldots + ||\boldsymbol{m}_n||^2 = ||M||^2, \end{aligned}$$

come volevasi dimostrare. Ora, per quanto riguarda MP, basta osservare che

$$||MP||^2 = ||{}^t(MP)||^2 = ||{}^tP\,{}^tM||^2 = ||{}^tM||^2 = ||M||^2,$$

dato che ${}^tP$ è anch'essa ortogonale.

(IV.5) Determinare gli autovalori e gli autovettori dell'operatore $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ definito dalla seguente matrice:

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 3 & -2 \\ 2 & 2 & 1 \end{bmatrix},$$

e dedurre se è diagonalizzabile.

Soluzione. Il polinomio caratteristico è

$$\begin{aligned}
p_A(\lambda) = \begin{vmatrix} 1-\lambda & 0 & 2 \\ 2 & 3-\lambda & -2 \\ 2 & 2 & 1-\lambda \end{vmatrix} &= (1-\lambda)\begin{vmatrix} 1-\lambda & 0 & 2 \\ 2 & 3-\lambda & -2 \\ 2 & 2 & 1-\lambda \end{vmatrix} + 2\begin{vmatrix} 1-\lambda & 0 & 2 \\ 2 & 3-\lambda & -2 \\ 2 & 2 & 1-\lambda \end{vmatrix} \\
&= (1-\lambda)((3-\lambda)(1-\lambda)+4) + 2(4-2(3-\lambda)) \\
&= (1-\lambda)(3-4\lambda+\lambda^2+4) - 4 + 4\lambda \\
&= (1-\lambda)(\lambda^2-4\lambda+7) - 4(1-\lambda) \\
&= (1-\lambda)(\lambda^2-4\lambda+3) = (1-\lambda)(\lambda-3)(\lambda-1).
\end{aligned}$$

Le sue radici sono $\lambda_1 = 1$ (con molteplicità algebrica due, cioè compare due volte nella fattorizzazione di $p_A(\lambda)$) e $\lambda_3 = 3$. Gli autovettori corrispondenti a λ_1 sono le soluzioni del sistema

▶ Non abbiamo tre autovalori distinti, per cui non siamo in grado di applicare la Proposizione (IV.81): non sappiamo ancora se è diagonalizzabile.

$$\begin{bmatrix} 1-1 & 0 & 2 \\ 2 & 3-1 & -2 \\ 2 & 2 & 1-1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \begin{cases} 2z = 0 \\ 2x+2y-2z = 0 \\ 2x+2y = 0. \end{cases}$$

Eliminiamo subito $z = 0$, e otteniamo il sistema

$$\begin{cases} 2x+2y = 0 \\ 2x+2y = 0, \end{cases}$$

che ha per soluzione tutti i multipli $(x, y, z) = c(1, -1, 0)$. Gli autovettori corrispondenti a λ_3 si trovano in modo analogo:

$$\begin{bmatrix} 1-3 & 0 & 2 \\ 2 & 3-3 & -2 \\ 2 & 2 & 1-3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \begin{cases} -2x+2z = 0 \\ 2x-2z = 0 \\ 2x+2y-2z = 0. \end{cases}$$

Eliminiamo $z = x$ (dalla prima e dalla seconda, che sono uguali), e rimane la relazione $2y = 0$. Dunque gli autovettori sono i multipli di $(1, 0, 1)$. In conclusione non ci sono più di due autovettori linearmente indipendenti, e la matrice non è quindi diagonalizzabile.

▶ Il Lettore può provare a completare i due vettori con un terzo vettore linearmente indipendente, ed ottenere una base in cui scrivere l'operatore L. Per esempio, nella base

$$\boldsymbol{b}_1 = \begin{bmatrix}1\\0\\1\end{bmatrix}, \boldsymbol{b}_2 = \begin{bmatrix}1\\-1\\0\end{bmatrix}, \boldsymbol{b}_3 = \begin{bmatrix}1\\-1\\1\end{bmatrix}$$

l'operatore L ha matrice associata

$$S^{-1}AS = \begin{bmatrix}3 & 0 & 0\\0 & 1 & 2\\0 & 0 & 1\end{bmatrix},$$

che non è diagonale ma quasi.

(IV.6) Sia V lo spazio delle funzioni continue su $[-1, 1]$, e $\langle f, g\rangle$ il prodotto scalare

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,\mathrm{d}x.$$

Determinare l'area del triangolo formato dalle tre funzioni 1, x e x^2.

Soluzione. Usiamo la formula di Erone, vista nell'Esercizio **49** a pagina 164:

$$\text{Area} = \sqrt{s(s-a)(s-b)(s-c)},$$

dove s è il semiperimetro e a, b e c le lunghezze dei lati. Allora risulta

$$\begin{aligned} a^2 &= \|x - x^2\|^2 = \int_{-1}^{1} (x-x^2)^2\,\mathrm{d}x = \frac{16}{15}\\ b^2 &= \|1 - x^2\|^2 = \int_{-1}^{1} (1-x^2)^2\,\mathrm{d}x = \frac{16}{15}\\ c^2 &= \|1 - x\|^2 = \int_{-1}^{1} (1-x)^2\,\mathrm{d}x = \frac{8}{3}. \end{aligned}$$

Quindi

$$a = b = \frac{4\sqrt{15}}{15}, \quad c = \frac{2\sqrt{6}}{3}, \quad s-a = s-b = \frac{\sqrt{6}}{3}$$
$$s = \frac{1}{2}(2a+c), \quad s-c = \frac{1}{2}(2a-c).$$

da cui segue che

$$\begin{aligned} \text{Area} &= (s-a)\sqrt{s(s-c)} = \frac{\sqrt{6}}{3}\frac{1}{2}\sqrt{4a^2-c^2}\\ &= \frac{\sqrt{6}}{6}\sqrt{4\frac{16}{15}-\frac{8}{3}} = \frac{\sqrt{6}}{6}\sqrt{\frac{8}{5}} = \frac{2\sqrt{15}}{15}. \end{aligned}$$

(IV.7) Sia V lo spazio vettoriale di tutte le matrici quadrate $n \times n$ (rispetto al prodotto per uno scalare e alla somma di matrici). Definiamo la norma di una matrice A con coefficienti a_{ij}:

$$\|A\| = \sqrt{\sum_{i,j=1}^{n} a_{ij}^2}.$$

Mostrare che in effetti è una norma e che verifica le ulteriori proprietà:

① se $v \in \mathbb{R}^n$, allora $\|Av\| \leq \|A\|\|v\|$;

② se A e B sono matrici, $\|AB\| \leq \|A\|\,\|B\|$.

(Questa norma è detta norma di Frobenius*, ed è anche indicata con* $\|A\|_2$*.)*

▶ Si veda la Definizione (IV.39) a pagina 183. Si veda anche l'Esercizio 40.

Soluzione. Lo spazio delle matrici $n \times n$ può essere identificato con $\mathbb{R}^{n^2}$ (parametrizzato con gli n^2 coefficienti delle matrici), e quindi la norma di una matrice non è altro che la norma standard del corrispondente vettore di $\mathbb{R}^{n^2}$. Osserviamo che se $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ sono i vettori colonna di una matrice A, allora

▶ Si veda la Proposizione (IV.38), e disuguaglianza triangolare della Proposizione (III.19) (pagina 112).

$$\|A\|^2 = \|\boldsymbol{a}_1\|^2 + \|\boldsymbol{a}_2\|^2 + \ldots + \|\boldsymbol{a}_n\|^2.$$

Lo stesso per i vettori riga (si vede facilmente che $\|A\| = \|{}^{t}A\|$). Per dimostrare la ①, ricordiamo che per la disuguaglianza di Cauchy–Schwartz se $\boldsymbol{a}$ e $\boldsymbol{b}$ sono due vettori di $\mathbb{R}^n$ allora $(\boldsymbol{a} \cdot \boldsymbol{b})^2 \leq \|\boldsymbol{a}\|^2\|\boldsymbol{b}\|^2$. Ora, se $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n$ indicano i vettori *riga* di una matrice A $(n \times n)$ e $v \in \mathbb{R}^n$ è un vettore di dimensione n, allora le componenti del vettore prodotto Av sono i prodotti scalari dei vettori riga per v $(\boldsymbol{r}_1 \cdot v, \boldsymbol{r}_2 \cdot v, \ldots, \boldsymbol{r}_n \cdot v)$, e quindi

$$\begin{aligned}
\|Av\|^2 &= (\boldsymbol{r}_1 \cdot v)^2 + (\boldsymbol{r}_2 \cdot v)^2 + \ldots + (\boldsymbol{r}_n \cdot v)^2 \\
&\leq \|\boldsymbol{r}_1\|^2\|v\|^2 + \|\boldsymbol{r}_2\|^2\|v\|^2 + \ldots + \|\boldsymbol{r}_n\|^2\|v\|^2 \\
&= \left(\|\boldsymbol{r}_1\|^2 + \|\boldsymbol{r}_2\|^2 + \ldots + \|\boldsymbol{r}_n\|^2\right)\|v\|^2 = \|A\|^2\|v\|^2.
\end{aligned}$$

Possiamo ora dedurre la ②, dato che (se $\boldsymbol{b}_i$ sono i vettori colonna di B)

$$\begin{aligned}
\|AB\|^2 &= \|(A\boldsymbol{b}_1, A\boldsymbol{b}_2, \ldots, A\boldsymbol{b}_n)\|^2 \\
&= \|A\boldsymbol{b}_1\|^2 + \|A\boldsymbol{b}_2\|^2 + \ldots + \|A\boldsymbol{b}_n\|^2 \\
&\leq \|A\|^2\|\boldsymbol{b}_1\|^2 + \|A\|^2\|\boldsymbol{b}_2\|^2 + \ldots + \|A\|^2\|\boldsymbol{b}_n\|^2 \\
&= \|A\|^2\left(\|\boldsymbol{b}_1\|^2 + \|\boldsymbol{b}_2\|^2 + \ldots + \|\boldsymbol{b}_n\|^2\right) = \|A\|^2\|B\|^2.
\end{aligned}$$

☆ **(IV.8)** *(Diagonalizzazione di Jacobi)* Sia M una matrice simmetrica e $q(\boldsymbol{x})$ la forma quadratica associata. Si cerchi un procedimento di diagonalizzazione di M (equivalentemente, di q) mediante permutazioni delle variabili e rotazioni tra le variabili x_1 e x_2, del tipo

$$\begin{cases} x_1 = cy_1 - sy_2 \\ x_2 = sy_1 + cy_2 \\ x_i = y_i \text{ per } i \geq 3, \end{cases}$$

con $c = \cos\vartheta$ e $s = \sin\vartheta$ per un opportuno angolo ϑ.

Soluzione. Sia

$$q(\boldsymbol{x}) = {}^t\boldsymbol{x}M\boldsymbol{x} = \sum_{i,j=1}^{n} m_{ij}x_ix_j$$

la forma quadratica associata a M. Cercheremo di diagonalizzare M con un procedimento analogo a quello usato nella dimostrazione del Teorema (IV.70), annullando uno a uno i termini al di fuori della diagonale. La forma quadratica $q(\boldsymbol{x})$ può essere scritta in modo più sintetico come

$$\begin{aligned} q = {} & m_{1,1}x_1^2 + 2m_{1,2}x_1x_2 + m_{2,2}x_2^2 + \\ & + x_1l_1(x_3,\dots,x_n) + x_2l_2(x_3,\dots,x_n) + q_3(x_3,\dots,x_n), \end{aligned}$$

dove l_1, l_2 sono espressioni *lineari* delle sole variabili $x_3,\dots,x_n$ e q_3 una forma quadratica nelle $x_3,\dots,x_n$. Dopo il cambio di variabili indicato, q si trasforma nella seguente espressione

$$\begin{aligned} q = {} & m_{1,1}(cy_1 - sy_2)^2 + 2m_{1,2}(cy_1 - sy_2)(sy_1 + cy_2) + m_{2,2}(sy_1 + cy_2)^2 + \\ & +(cy_1 - sy_2)l_1(y_3,\dots,y_n) + (sy_1 + cy_2)l_2(y_3,\dots,y_n) + q_3(y_3,\dots,y_n), \end{aligned} \tag{23}$$

con le stesse l_1, l_2 e q_3. Il coefficiente di y_1y_2 è

$$2\hat{m}_{1,2} = -2m_{1,1}cs + 2m_{1,2}(c^2 - s^2) + 2m_{2,2}cs,$$

che si può annullare se

$$\frac{c^2 - s^2}{cs} = \frac{m_{1,1} - m_{2,2}}{m_{1,2}}. \tag{24}$$

Ma $\frac{\cos^2\vartheta - \sin^2\vartheta}{\sin\vartheta\cos\vartheta} = \frac{2\cos 2\vartheta}{\sin 2\vartheta} = 2/\tan 2\vartheta$, e quindi per ogni $m \in \mathbb{R}$ esiste uno ed un solo valore $\vartheta_m \in (-\pi/4, \pi/4]$ per cui $2/\tan 2\vartheta_m = m$. A questo valore ϑ_m corrispondono due valori $c = \cos\vartheta_m$ e $s = \sin\vartheta_m$ che risolvono l'equazione (24), che quindi annullano il coefficiente di y_1y_2 nella (23) se $m = \frac{m_{1,1} - m_{2,2}}{m_{1,2}}$. I coefficienti dei termini y_1^2 e y_2^2 possono essere facilmente calcolati e risultano

$$\begin{cases} \hat{m}_{1,1} = c^2m_{1,1} + 2m_{1,2}cs + m_{2,2}s^2 = m_{1,1} + m_{1,2}\dfrac{s}{c} \\ \hat{m}_{2,2} = m_{1,1}s^2 - 2m_{1,2}cs + m_{2,2}c^2 = m_{2,2} - m_{1,2}\dfrac{s}{c}. \end{cases} \tag{25}$$

Quelli dei termini $y_1 y_i$ con $i \geq 3$ compaiono nell'espressione

$$y_1(cl_1(y_3, \ldots, y_n) + sl_2(y_3, \ldots, y_n)),$$

quelli dei termini $y_2 y_i$ con $i \geq 3$ nell'espressione

$$y_2(-sl_1(y_3, \ldots, y_n) + cl_2(y_3, \ldots, y_n)).$$

Il problema è che iterando questa procedura per cancellare tutti i coefficienti degli $x_1 x_j$, a ogni passo i coefficienti che si erano annullati prima possono tornare ad essere diversi da zero! Però accade che la somma dei quadrati degli elementi della matrice $m_{i,j}$ non cambia (perché si moltiplica per matrici ortogonali, si veda l'Esercizio (IV.4)), e quindi se si indica con $\hat{M}$ la matrice con coefficienti $\hat{m}_{ij}$ ottenuta con il cambio di coordinate, si ha

$$\sum_{i,j} \hat{m}_{ij}^2 = \|\hat{M}\|^2 = \|M\|^2 = \sum_{i,j} m_{ij}^2.$$

Indichiamo con S e $\hat{S}$ le somme dei quadrati dei termini al di fuori della diagonale:

$$S = \sum_{i \neq j} m_{ij}^2 = 2 \sum_{i<j} m_{ij}^2 = \|M\|^2 - \sum_i m_{ii}^2, \quad \hat{S} = \sum_{i \neq j} \hat{m}_{ij}^2 = 2 \sum_{i<j} \hat{m}_{ij}^2 = \|\hat{M}\|^2 - \sum_i \hat{m}_{ii}^2.$$

Dato che i termini sulla diagonale che vengono modificati sono solo m_{11} e m_{22}, si ha (sostituendo i valori di (25))

$$\begin{aligned} S - \hat{S} &= \hat{m}_{11}^2 - m_{11}^2 + \hat{m}_{22}^2 - m_{22}^2 = (m_{11} + m_{12}\frac{s}{c})^2 - m_{11}^2 + (m_{22} - m_{12}\frac{s}{c})^2 - m_{22}^2 \\ &= 2m_{12}m_{11}\frac{s}{c} + m_{12}^2\frac{s^2}{c^2} - 2m_{12}m_{22}\frac{s}{c} + m_{12}^2\frac{s^2}{c^2} = 2m_{12}(m_{11} - m_{22})\frac{s}{c} + 2m_{12}^2\frac{s^2}{c^2} \\ &= 2m_{12}^2\left[\frac{c^2 - s^2}{cs}\frac{s}{c} + \frac{s^2}{c^2}\right] = 2m_{12}^2. \end{aligned}$$

▶ Sostituiamo $m_{11} - m_{22} = m_{12}(c^2 - s^2)/(cs)$.

Ora immaginiamo di applicare questa procedura nel modo seguente: ad ogni passo selezioniamo gli indici $i < j$ in modo tale che il coefficiente della matrice al posto i, j sia il massimo tra quelli fuori dalla diagonale, e trasformiamo la matrice M con il cambio di coordinate appena visto se questo non è zero. Si otterrà una successione di matrici

$$M = M_0, M_1, M_2, \ldots, M_k, \ldots$$

in cui la somma S_k dei quadrati dei termini al di fuori della diagonale della matrice k-esima soddisfa l'uguaglianza

(26) $S_{k+1} = S_k - 2m_k^2,$

dove m_k^2 è appunto il massimo tra i quadrati dei termini al di fuori della diagonale della matrice M_k. Il procedimento termina quando M_k è in forma diagonale. Altrimenti $\{S_k\}$ è una successione di numeri positivi monotona decrescente, che deve convergere a un limite positivo o nullo

$$S_\infty = \lim_{k \to +\infty} S_k.$$

Osserviamo che da (26) segue che

$$S_k = S_{k-1} - 2m_{k-1}^2 = S_{k-2} - 2m_{k-2}^2 - 2m_{k-1}^2 = \ldots = S_0 - 2\sum_{j=0}^{k-1} m_j^2,$$

e quindi se la successione S_k converge, allora la successione m_j deve convergere a zero per $j \to \infty$, e dato che

$$(27) \quad 0 \le 2m_k^2 \le S_k \le n(n-1)m_k^2$$

(visto che m_k^2 è il massimo degli $n(n-1)$ termini di S_k) anche S_k deve convergere a zero. Dalla disuguaglianza (27) possiamo anche dedurre anche che

$$\frac{2}{n(n-1)} S_k \le 2m_k^2 \le S_k \implies S_{k+1} = S_k - 2m_k^2 \le S_k - \frac{2}{n(n-1)} S_k = \alpha S_k$$

per $\alpha = 1 - \frac{2}{n(n-1)} < 1$, da cui

$$S_k \le \alpha^k S_0 \implies m_k \le \beta^k \frac{\sqrt{S_0}}{2}$$

per $\beta = \sqrt{\alpha} \in [0,1)$. Possiamo dedurre che anche i coefficienti della diagonale di M_k abbiano un limite? Ad ogni passo del tipo (25) si aggiungono alla diagonale termini del tipo $\pm m_k \frac{\sin \vartheta_k}{\cos \vartheta_k}$ per un certo angolo $\vartheta_k \in (-\pi/4, \pi/4]$, in cui m_k tende a zero al massimo come β^k e $\tan \vartheta_k = \frac{\sin \vartheta_k}{\cos \vartheta_k} \in [-1,1]$. Ma allora non solo la serie $\sum_k m_k^2$, ma anche $\sum_k |m_k|$ converge, e quindi i termini sulla diagonale in effetti convergono. La successione di matrici M_k pertanto converge ad una matrice diagonale D. Non rimane che mostrare che i termini sulla diagonale di D sono proprio gli autovalori di M (o di ogni M_k, visto che le M_k sono simili a M e quindi hanno gli stessi autovalori), e che M è simile a D mediante una trasformazione ortogonale. Per far questo abbiamo bisogno di un po' di strumenti che per ora ci mancano, per cui lasciamo al Lettore per esercizio i dettagli di quest'ultima parte, tra qualche capitolo (Esercizio **41** a pagina 396).

► È chiaro che abbiamo già occupato abbastanza spazio con questa diagonalizzazione di Jacobi, e in questo margine non vi è abbastanza spazio per procedere oltre.(12)

☆ **(IV.9)** Sullo spazio euclideo $V = \mathbb{R}^n$ con prodotto scalare standard sia $L \colon V \to V$ un operatore lineare *auto-aggiunto*, cioè tale che $\langle Lx, y\rangle = \langle x, Ly\rangle$. Sia definita la funzione $q \colon V \setminus \{0\} \to \mathbb{R}$, ponendo

$$q(x) = \frac{\langle Lx, x\rangle}{\langle x, x\rangle}.$$

Mostrare che se z è un vettore tale che $q(x) \geq q(z)$ per ogni $x \in V$, allora z è un autovettore per L, $q(z)$ è il corrispondente autovalore, e questo è il minimo tra tutti gli autovalori di L.

▶ La matrice associata a L è simmetrica: se $\langle Lx, y\rangle = \langle x, Ly\rangle$, cioè ${}^t(Lx)y = {}^txLy$, si ha ${}^tx{}^tLy = {}^txLy$. Ora se si prendono $x = e_i$ e $y = e_j$ si ha che le componenti di ${}^tL$ devono essere uguali a quelle di L, cioè che L è simmetrica.

▶ La funzione $q(x)$ è il quoziente di due forme quadratiche, e viene chiamata il *quoziente di Rayleigh*.

Soluzione. Per il Teorema Spettrale (IV.83) la matrice L è diagonalizzabile mediante una matrice ortogonale Q, cioè esiste una matrice ortogonale Q tale che $D = Q^{-1}LQ$ è diagonale. Gli elementi sulla diagonale di D sono gli autovalori $\lambda_1, \lambda_2, \ldots, \lambda_n$ di L, e possiamo supporre (a meno di cambiare le coordinate) che

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n.$$

I corrispondenti autovettori di L (le colonne di Q) formano un sistema di riferimento ortonormale. Chiamiamo con y le coordinate di $\mathbb{R}^n$ relative a questo sistema di riferimento, cioè $x = Qy$. Dato che $Q^{-1} = {}^tQ$, si ha

$$\begin{aligned} \langle Lx, x\rangle &= {}^txLx = {}^t(Qy)L(Qy) = {}^ty{}^tQLQy = {}^tyDy \\ \langle x, x\rangle &= {}^txx = {}^ty{}^tQQy = {}^tyy \\ \implies q(x) &= \frac{\langle Lx, x\rangle}{\langle x, x\rangle} = \frac{{}^tyDy}{{}^tyy}. \end{aligned}$$

Ora, sia $q_1 = q(z)$. Si ha che $q(x) \geq q_1$ per ogni $x \neq \mathbf{0}$ se e soltanto se per ogni $y \neq \mathbf{0}$

$$\begin{aligned} & {}^tyDy \geq q_1{}^tyy \iff \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2 \geq q_1(y_1^2 + y_2^2 + \ldots + y_n^2) \\ (*) \quad & \iff (\lambda_1 - q_1)y_1^2 + (\lambda_2 - q_1)y_2^2 + \ldots + (\lambda_n - q_1)y_n^2 \geq 0, \end{aligned}$$

cioè se e solo se la forma quadratica scritta è semi-definita positiva, e dunque deve essere

$$q_1 \leq \lambda_1 \leq \lambda_2 \ldots \leq \lambda_n.$$

ma per il vettore z si ha

$$(**) \quad (\lambda_1 - q_1)z_1^2 + (\lambda_2 - q_1)z_2^2 + \ldots + (\lambda_n - q_1)z_n^2 = 0.$$

da cui segue che nella (**) si deve avere $\lambda_i = q_1$ per qualche i. L'unico i possibile è $i = 1$ (o un i successivo tale che $\lambda_1 = \lambda_2 = \ldots = \lambda_i$). Quindi $q_1 = q(x) = \lambda_1$, che è il più piccolo tra gli autovalori. Per la (**) le componenti di z corrispondenti ad autovalori diversi da q_1 devono essere zero, e quindi z è autovettore per λ_1.

Diamoci da fare… *(Soluzioni a pagina 586)*

1 Se $\boldsymbol{i}$ e $\boldsymbol{j}$ sono due vettori linearmente indipendenti di uno spazio vettoriale, si determinino due coefficienti x e y tali che $x(2\boldsymbol{i}+\boldsymbol{j})+y(\boldsymbol{i}+\boldsymbol{j})=\boldsymbol{i}-\boldsymbol{j}$

2 Siano $\boldsymbol{a}=(2,-1,1)$, $\boldsymbol{b}=(1,2,-1)$ e $\boldsymbol{c}=(2,-11,7)$. Sono linearmente dipendenti? Se sì, scriverne uno come combinazione lineare degli altri due.

3 Determinare tutti i valori del parametro t per cui i vettori (t,t^2) e $(1,t)$ sono paralleli, e quelli per cui sono ortogonali.

4 Siano $\boldsymbol{a}$, $\boldsymbol{b}$ e $\boldsymbol{c}$ tre vettori di $\mathbb{R}^n$. Mostrare che $\boldsymbol{a}$, $\boldsymbol{a}+\boldsymbol{b}$ e $\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}$ sono linearmente indipendenti se e solo se lo sono $\boldsymbol{a}$, $\boldsymbol{b}$ e $\boldsymbol{c}$. Cosa si può dire dell'indipendenza dei tre vettori $\boldsymbol{a}+\boldsymbol{b}$, $\boldsymbol{b}+\boldsymbol{c}$, $\boldsymbol{c}+\boldsymbol{a}$?

5 Si considerino in $\mathbb{R}^n$ (con base standard $\boldsymbol{e}_1,\boldsymbol{e}_2,\dots,\boldsymbol{e}_n$) i vettori $\boldsymbol{b}_1=\boldsymbol{e}_1$, $\boldsymbol{b}_2=\boldsymbol{e}_1+\boldsymbol{e}_2$, $\boldsymbol{b}_3=\boldsymbol{e}_1+\boldsymbol{e}_2+\boldsymbol{e}_3$, …, $\boldsymbol{b}_n=\boldsymbol{e}_1+\boldsymbol{e}_2+\ldots+\boldsymbol{e}_n$. Si mostri che costituiscono una base. Si scrivano le coordinate di $\boldsymbol{x}=(x_1,x_2,\dots,x_n)$ in questa base.

6 Determinare tutti i valori del parametro t per cui i tre vettori $(1,t,0)$, $(t,0,1)$ e $(0,1,t)$ sono linearmente dipendenti.

7 Siano S_1 e S_2 due sottoinsiemi di vettori di uno spazio vettoriale. Mostrare che

- **(a)** Se $S_1\subseteq S_2$, allora $\operatorname{Span}(S_1)\subseteq\operatorname{Span}(S_2)$.
- **(b)** $\operatorname{Span}(S_1\cap S_2)\subseteq\operatorname{Span}(S_1)\cap\operatorname{Span}(S_2)$.
- **(c)** Può essere $\operatorname{Span}(S_1\cap S_2)\neq\operatorname{Span}(S_1)\cap\operatorname{Span}(S_2)$?

8 Sia V lo spazio lineare di tutte le funzioni da $\mathbb{R}$ in $\mathbb{R}$. Calcolare le dimensioni dei seguenti Span:

- **(a)** $\operatorname{Span}(1,e^{ax},e^{bx})$;
- **(b)** $\operatorname{Span}(\sin x,\cos x,\sin^2 x,\cos^2 x)$;
- **(c)** $\operatorname{Span}(1,e^{ax},xe^{ax},x^2e^{ax})$;
- **(d)** $\operatorname{Span}(1,\cos 2x,\sin^2 x)$.

9 Si dimostri che il nucleo di una funzione lineare $L\colon V\to W$ è sempre uno spazio vettoriale (rispetto alle operazioni definite su V).

10 Quali delle seguenti funzioni $\mathbb{R}^2\to\mathbb{R}^2$ sono lineari? Di quelle lineari si calcolino la dimensione del nucleo e dell'immagine.

- **(a)** $f(x,y)=(y,x)$;
- **(b)** $f(x,y)=(x+y,x-y)$;
- **(c)** $f(x,y)=(x,2x)$;
- **(d)** $f(x,y)=(x^2,y^2-x)$;
- **(e)** $f(x,y)=(x+y,2x+2y)$.

11 Delle seguenti trasformazioni $\mathbb{R}^2\to\mathbb{R}^2$, si calcoli il nucleo e l'immagine, se sono lineari.

- **(a)** Rotazione attorno all'origine di angolo ϑ;
- **(b)** Riflessione rispetto alla retta di equazione $y=x$.
- **(c)** Riflessione rispetto alla retta di equazione $y=x+1$.
- **(d)** Omotetia di ragione 2 con centro in $(1,0)$.
- **(e)** Traslazione di vettore $(1,1)$.

12 Si calcolino il rango e la nullità delle seguenti trasformazioni $\mathbb{R}^3\to\mathbb{R}^3$, quando sono lineari.

- **(a)** $f(x,y,z)=(2x,3y,3z)$;
- **(b)** $f(x,y,z)=(x+y,y+z,z+x)$;
- **(c)** $f(x,y,z)=(1+x,1+y,1+z)$;
- **(d)** $f(x,y,z)=(x,y,1)$;
- **(e)** $f(x,y,z)=(x,y^2,z^3)$.

13 Si dica quali delle seguenti trasformazioni $V\to V$ sono lineari, dove V è lo spazio di tutti i polinomi di grado al più n, e di quelle lineari si calcoli il nucleo.

- **(a)** Se $p\in V$, $T(p)$ è il polinomio $q(x)$ definito da $q(x)=p(x+1)$.
- **(b)** Se $p\in V$, $T(p)(x)=(x-1)p'(x)$.
- **(c)** Se $p\in V$, $T(p)(x)=p'(x)$.
- **(d)** Se $p\in V$, $T(p)(x)=p''(x)+2p'(x)+p(x)$.
- **(e)** Se $p\in V$, $T(p)(x)=x^2p''(x)-2xp'(x)+p(x)$.

14 Sia V lo spazio vettoriale di tutte le successioni reali $\{x_n\}$ convergenti (con somma $\{x_n\}+\{y_n\}=\{x_n+y_n\}$ e prodotto per uno scalare $c\{x_n\}=\{cx_n\}$). Sia $T\colon V\to V$ definita come segue

$$T(\{x_n\})=\{x_n-\lim_{n\to\infty}x_n\}.$$

È un operatore lineare? Quale è il suo nucleo?

15 Sia V lo spazio di tutti i polinomi, $D\colon V\to V$ l'operatore di derivazione $D\colon p(x)\mapsto p'(x)$ e $I\colon V\to V$ l'operatore di integrazione $I\colon p(x)\mapsto\int_0^x p(t)\,dt$. Mostrare che $DI(p)=p$ per ogni $p\in V$, ma che $ID(p)$ può non essere uguale a p.

16 Determinare le matrici associate alle seguenti funzioni lineari.

(a) $f\colon \mathbb{R}^3 \to \mathbb{R}^3$, $f(\boldsymbol{x}) = 2\boldsymbol{x}$.

(b) $f\colon \mathbb{R}^3 \to \mathbb{R}^3$, $f(\boldsymbol{x}) = \boldsymbol{a} \times \boldsymbol{x}$ per un certo $\boldsymbol{a} \in \mathbb{R}^3$ fissato.

(c) $f\colon \mathbb{R}^3 \to \mathbb{R}$, $f(\boldsymbol{x}) = \boldsymbol{a}\cdot\boldsymbol{x}$ per un certo $\boldsymbol{a} \in \mathbb{R}^3$ fissato.

(d) $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ tale che $f(\boldsymbol{i}+\boldsymbol{j}) = \boldsymbol{k}$, $f(\boldsymbol{j}+\boldsymbol{k}) = \boldsymbol{i}$, e $f(\boldsymbol{k}+\boldsymbol{i}) = \boldsymbol{j}$.

17 Per ogni V indicato, determinare la matrice associata all'operatore $D\colon V \to V$ di derivazione.

(a) $V = \mathrm{Span}(\cos x, \sin x)$;

(b) $V = \mathrm{Span}(1, e^x, xe^x)$;

(c) $V = \mathrm{Span}(e^x \cos x, e^x \sin x)$;

(d) $V = \mathrm{Span}(1, x, x^2, x^3)$.

18 Siano $A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$, $C = \begin{bmatrix} 3 & 4 \\ 5 & 6 \\ 1 & 2 \end{bmatrix}$. Si calcolino $B+C$, AB, AC, BA, CA, $A(B+C)$.

19 Si calcoli $AB-BA$, dove $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. In generale, mostrare che una matrice Z 2×2 ha la proprietà che $BZ = ZB$ per ogni matrice 2×2 B (cioè *commuta* con B) se e soltanto se commuta con ognuna delle seguenti matrici:

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$
$$E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Come sono fatte questa matrici?

20 Sia $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Calcolare, per ogni $n \in \mathbb{Z}$, A^n.

21 Verificare che la matrice associata alla rotazione del piano di un angolo ϑ in senso antiorario attorno all'origine è $R_\vartheta = \begin{bmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{bmatrix}$.

22 Sia R_ϑ la matrice $R_\vartheta = \begin{bmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{bmatrix}$. Mostrare che $R_\vartheta R_{\vartheta'} = R_{\vartheta+\vartheta'}$, e calcolare $(R_\vartheta)^n$ per ogni n.

⚽ **23** Una *matrice di Hadamard* è una matrice $n\times n$ con le seguenti proprietà:

① Gli elementi della matrice sono tutti 1 o -1.

② Il prodotto scalare di due colonne distinte è zero.

Determinare tutte le matrici di Hadamard 2×2 (sono otto). Mostrare che possono esserci matrici di Hadamard con $n > 2$ solo se n è un multiplo di 4. *(Suggerimento: mostrare che $(\boldsymbol{x}+\boldsymbol{y})\cdot(\boldsymbol{x}+\boldsymbol{z}) = \|\boldsymbol{x}\|^2 = n$ se $\boldsymbol{x}$, $\boldsymbol{y}$ e $\boldsymbol{z}$ sono tre colonne ortogonali. Dal fatto che $(x_i+y_i)(x_i+z_i)$ è 0 o 4 per ogni i, dedurre che n è divisibile per 4.)*

24 Sia P_n lo spazio vettoriale di tutti i polinomi reali di grado $\le n$. Poniamo

$$\langle f, g\rangle = \sum_{i=0}^{n} f\left(\frac{i}{n}\right) g\left(\frac{i}{n}\right).$$

(a) Mostrare che è un prodotto scalare.

(b) Calcolare $\langle 1, x\rangle$.

25 Nello spazio vettoriale di tutti i polinomi, definiamo

$$\langle f, g\rangle = \int_0^\infty e^{-x} f(x) g(x)\, \mathrm{d}x.$$

(a) Mostrare che è ben definito (l'integrale esiste) ed è un prodotto scalare.

(b) I vettori x^n, al variare di n, sono ortogonali? Calcolare i prodotti scalari $\langle x^n, x^m\rangle$.

(c) Determinare quali polinomi di primo grado sono ortogonali a $1+x$.

26 Si mostri che se $W \subset V$ è un sottospazio di dimensione finita di uno spazio vettoriale V con prodotto scalare, allora la proiezione ortogonale $P_W\colon V \to W$ di V su W è una trasformazione lineare.

☆ **27** Sia l^2 l'insieme di tutte le successioni $\boldsymbol{x} = \{x_n\}_{n=1}^\infty$ tali che la serie $\sum_{n=1}^\infty x_n^2$ è convergente. Mostrare che è uno spazio vettoriale, e che il prodotto

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \sum_{n=1}^\infty x_n y_n$$

è un prodotto scalare. Calcolare l'angolo e la distanza tra le successioni $x_n = 1/n$ e $y_n = 1/(n+1)$.

⚽ **28** Calcolare il determinante delle seguenti matrici.

(a) $\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{bmatrix}$;

(b) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{bmatrix}$;

(c) $\begin{bmatrix} x & 1 & 0 & 0 \\ 1 & x & 1 & 0 \\ 0 & 1 & x & 1 \\ 0 & 0 & 1 & x \end{bmatrix}$;

(d) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

29 Mostrare che se M è una matrice del tipo

$$M = \begin{bmatrix} A_{k\times k} & 0_{k\times k} \\ B_{k\times k} & C_{k\times k} \end{bmatrix},$$

in cui i blocchi indicati sono matrici $k \times k$, come per esempio

$$M = \begin{bmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ b_{11} & b_{12} & c_{11} & c_{12} \\ b_{21} & b_{22} & c_{21} & c_{22} \end{bmatrix},$$

allora $\det(M) = \det(A)\det(C)$.

30 Calcolare l'inversa delle seguenti matrici.

(a) $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$;

(b) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$;

(c) $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

31 Fissate tre funzioni $a(x)$, $b(x)$, $c(x)$, si considerino le matrici

$$W(a,b;x) = \begin{bmatrix} a(x) & b(x) \\ a'(x) & b'(x) \end{bmatrix},$$

$$W(a,b,c;x) = \begin{bmatrix} a(x) & b(x) & c(x) \\ a'(x) & b'(x) & c'(x) \\ a''(x) & b''(x) & c''(x) \end{bmatrix}$$

(matrice *wronskiana*, in onore di Josef-Maria Hoëné de Wronski (1778–1853)).(13) Mostrare che se le funzioni $a(x)$, $b(x)$ (e $c(x)$) sono linearmente dipendenti (nello spazio vettoriale di tutte le funzioni derivabili due volte) allora $\det(W(a,b;x)) \equiv 0$ (risp. $\det(W(a,b,c;x)) \equiv 0$). Considerare le seguenti funzioni

$$a(x) = x^3; \quad b(x) = |x|^3,$$

per verificare se è vero il viceversa, cioè se $W(a,b;x) \equiv 0$ implica o no che $a(x)$, $b(x)$ sono dipendenti.

32 Dimostrare che ogni autovettore corrisponde ad uno e un solo autovalore.

☆ **33** Si determinino gli autovalori e gli autovettori delle matrici

$$A = \begin{bmatrix} 6 & -3 & 3 \\ 1 & 4 & 3 \\ -1 & 1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 1 & -2 \\ 2 & -1 & 0 \end{bmatrix}.$$

34 Sia $V = C^\infty(\mathbb{R})$ lo spazio vettoriale di tutte le funzioni derivabili infinite volte in $\mathbb{R}$. Si considerino gli operatori lineari $D\colon V \to V$ e $T\colon V \to V$ definiti da $D(f(x)) = f'(x)$ e $T(f(x)) = xf'(x)$. Determinare gli autovalori e autovettori (autofunzioni) di T e D.

35 Mostrare che la funzione $p_M(\lambda) = \det(M - \lambda I)$ è un polinomio di grado n in λ, se M è una matrice $n \times n$ (il *polinomio caratteristico*). *(Suggerimento: mostrare prima per induzione che il determinante è un polinomio omogeneo di grado n nei coefficienti della matrice M che contiene il termine $m_{11}m_{22}\dots m_{nn}$.)*

36 Mostrare che il polinomio caratteristico di una matrice A coincide con quello della sua trasposta ${}^{t}A$.

☆ **37** Mostrare che gli autovalori di una matrice M sono tutte e sole le radici del polinomio caratteristico $p_M(\lambda)$, e che i polinomi caratteristici $p_{M_1}(\lambda)$ e $p_{M_2}(\lambda)$ di due matrici simili coincidono.

Prima di proseguire con gli esercizi, diamo una definizione: la somma $\sum_{i=1}^{n} a_{ii}$ dei termini sulla diagonale di una matrice A si chiama *traccia* di A, e si indica con tr.

☆ **38** Mostrare che i coefficienti del polinomio caratteristico

$$p_A(\lambda) = c_0 + c_1\lambda + c_2\lambda^2 + \dots + c_{n-1}\lambda^{n-1} + c_n\lambda^n$$

di una matrice quadrata A soddisfano le uguaglianze

$$c_0 = \det(A), \quad c_{n-1} = (-1)^{n-1}\operatorname{tr} A, \quad c_n = (-1)^n.$$

Dedurre che se B e A sono matrici simili, allora $\operatorname{tr} A = \operatorname{tr} B$, e che

① la traccia di una matrice è la somma dei suoi autovalori: $\operatorname{tr} A = \sum_{i=1}^{n} \lambda_i$.

② il determinante di una matrice è il prodotto dei suoi autovalori: $\det A = \prod_{i=1}^{n} \lambda_i$.

39 Dimostrare le seguenti proprietà della traccia:

① $\operatorname{tr}(A + B) = \operatorname{tr}(A) + \operatorname{tr}(B)$;

② $\operatorname{tr}(cA) = c\operatorname{tr}(A)$;

③ $\operatorname{tr}({}^{t}A) = \operatorname{tr}(A)$.

④ $\operatorname{tr}(AB) = \operatorname{tr}(BA)$;

40 Sia V lo spazio vettoriale di dimensione nm delle matrici $n \times m$. Mostrare che la funzione

$$\langle A, B\rangle = \operatorname{tr}({}^{t}AB)$$

è un prodotto scalare su V (secondo la Definizione (IV.34) a pagina 181).

41 Quali tra le seguenti matrici hanno gli stessi autovalori? Quali sono simili?

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

42 Determinare quali delle seguenti matrici sono diagonalizzabili.

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 1 & 3 \end{bmatrix}, C = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, D = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}.$$

43 Sia V uno spazio vettoriale con prodotto scalare $\langle -, - \rangle$. Diciamo che un operatore lineare $L^*\colon V \to V$ è un operatore *aggiunto* di un operatore $L\colon V \to V$ se per ogni $\boldsymbol{x}$, $\boldsymbol{y}$ in V

$$\langle L^* \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, L\boldsymbol{y} \rangle.$$

Diciamo che un operatore lineare $L\colon V \to V$ è *auto-aggiunto* se

$$\langle L\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, L\boldsymbol{y} \rangle$$

per ogni $\boldsymbol{x}$, $\boldsymbol{y}$ in V. Mostrare che se $V \cong \mathbb{R}^n$ con il prodotto scalare standard e L è una matrice, allora esiste un unico aggiunto $L^* = {}^tL$, e gli operatori auto-aggiunti sono esattamente le matrici simmetriche.

44 Si associ al *grafo* (un grafo è un insieme di vertici e di segmenti che li collegano) della Figura IV.6

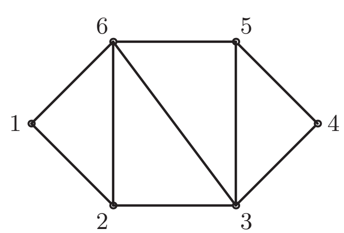


Figura IV.6: Grafo

la matrice A, detta *matrice di adiacenza*, definita come segue: $a_{ij} = 1$ se e solo se il vertice i e il vertice j sono collegati da un segmento, altrimenti $a_{ij} = 0$. Si scriva la matrice A, e si mostri per induzione che per ogni $n \geq 1$ la potenza A^n è una matrice che ha nella (i, j)-esima posizione il numero di percorsi diversi che vanno in n passi dall'i-esimo vertice al j-esimo vertice.

45 Siano $\boldsymbol{a}$ e $\boldsymbol{b}$ due vettori di $\mathbb{R}^3$. Allora possono essere pensati come matrici $A = (\boldsymbol{a})$ e $B = (\boldsymbol{b})$ con una sola colonna e tre righe. Mostrare che il determinante della matrice 3x3 $A\,{}^tB$ è sempre zero. Si rappresentino geometricamente i suoi tre vettori colonna.

46 Dimostrare che lo span di un insieme qualsiasi di vettori $S \subset V$ di uno spazio vettoriale V è uno spazio vettoriale rispetto alle operazioni di somma e prodotto per uno scalare ereditate da V.

47 Si mostri che il rango di una sottomatrice $\hat{A}$ di una matrice A non può superare il rango di A: $\text{Rango}(\hat{A}) \leq \text{Rango}(A)$.

48 Estendere l'Esercizio **54** a pagina 164 a quattro punti: dimostrare che quattro punti $A = (a_1, a_2, a_3)$, $B(b_1, b_2, b_3)$, $C = (c_1, c_2, c_3)$ e $D = (d_1, d_2, d_3)$ di $\mathbb{R}^3$ sono complanari se e soltanto se la matrice

$$\begin{bmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

ha determinante nullo.

49 Dimostrare che ci sono $n(n+1)/2$ distinti monomi di grado due in n variabili.

Diciamo che una matrice Q è *triangolare superiore* se ha tutti termini nulli sotto la diagonale

$$U = \begin{bmatrix} q_{11} & q_{12} & q_{13} & \cdots & q_{1n} \\ 0 & q_{22} & q_{23} & \cdots & q_{2n} \\ 0 & 0 & q_{33} & \cdots & q_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & q_{nn} \end{bmatrix}.$$

Diciamo che una matrice Q è *unitriangolare superiore* se è triangolare superiore e se i termini sulla diagonale sono 1, cioè se ha tutti 1 sulla diagonale e tutti 0 sotto la diagonale

$$Q = \begin{bmatrix} 1 & q_{12} & q_{13} & \cdots & q_{1n} \\ 0 & 1 & q_{23} & \cdots & q_{2n} \\ 0 & 0 & 1 & \cdots & q_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & 1 \end{bmatrix}.$$

50 Siano U, V due matrici triangolari superiori (cioè tali che $u_{ij} = 0$ se $i > j$). Mostrare che allora anche $U+V$, UV e (quando esiste) U^{-1} sono triangolari superiori.

51 Mostrare che l'inversa di una matrice unitriangolare superiore esiste sempre ed è unitriangolare.

☆ **52** Mostrare che se A e Q sono matrici $n \times n$ e Q è unitriangolare superiore allora i minori principali di A, AQ, ${}^tQA$, ${}^tQAQ$ coincidono.

☕ **53** Mostrare che se una forma quadratica q su $\mathbb{R}^n$ è definita positiva, allora tutti i *minori principali* $d_1, d_2, \dots, d_n$ della matrice associata a q (cioè i determinanti delle sue sottomatrici principali) sono positivi.

☆ **54** Dedurre dall'Esercizio **52** l'inversa dell'Esercizio **53**: si mostri per induzione che se tutti i *minori principali* $d_1, d_2, \dots, d_n$ della matrice associata ad una forma quadratica q sono positivi, allora q è definita positiva. *(Suggerimento: si considerino i cambi di coordinate usati nella dimostrazione del Teorema (IV.70), e si noti che deve essere $a_{11} > 0$.)*

☆ **55** Si supponga di aver definito il determinante solo per matrici $(n-1) \times (n-1)$, con le proprietà del Teorema (IV.54). Dimostrare che se $F(A)$ è una funzione lineare nelle colonne di A (matrice $n \times n$), che cambia di segno se si scambiano due colonne, e infine che vale 1 su I_n, allora

$$F\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{bmatrix} = a_{11} \det \begin{bmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn}. \end{bmatrix}$$

Dimostrare analogamente che se G è un'altra funzione, con le stesse proprietà per le righe invece che per le colonne, allora

$$G\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = a_{11} \det \begin{bmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn}. \end{bmatrix}$$

☆ **56** Mostrare che se $F(A)$ è una funzione lineare nelle colonne di A (matrice $n \times n$), che cambia di segno se si scambiano due colonne di A e tale che $F(I_n) = 1$, allora se A ha una colonna in cui tutti i coefficienti sono zero tranne uno, che è uguale a 1,

$$A = \left[\begin{array}{cccc|c|ccc} a_{11} & a_{12} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,j-1} & 0 & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i1} & a_{i2} & \cdots & a_{i,j-1} & 1 & a_{i,j+1} & \cdots & a_{in} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{nn} \end{array}\right]$$

allora si ha che $(-1)^{i+j}F(A)$ è uguale alla funzione

$$\det\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{bmatrix}.$$

Spazi vettoriali complessi

57 Siano $\boldsymbol{a} = (1, i)$, $\boldsymbol{b} = (-i, i)$ e $\boldsymbol{c} = (2i, -1)$ tre vettori di $\mathbb{C}^2$. Si calcolino i seguenti prodotti scalari (con il prodotto scalare hermitiano standard di $\mathbb{C}^n$).

(a) $\boldsymbol{a} \cdot \boldsymbol{b}$;

(b) $\boldsymbol{b} \cdot \boldsymbol{a}$;

(c) $\boldsymbol{a} \cdot (i\boldsymbol{b}) + (i\boldsymbol{a}) \cdot (\boldsymbol{b} + \boldsymbol{c})$;

(d) $(\boldsymbol{a} + i\boldsymbol{b}) \cdot (\boldsymbol{a} - i\boldsymbol{b})$.

58 Mostrare che per ogni $\boldsymbol{a}$, $\boldsymbol{b}$ di uno spazio vettoriale complesso con prodotto scalare hermitiano si ha

① $\|\boldsymbol{a} + \boldsymbol{b}\|^2 = \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2 + \boldsymbol{a} \cdot \boldsymbol{b} + \overline{\boldsymbol{a} \cdot \boldsymbol{b}}$;

② $\|\boldsymbol{a} + \boldsymbol{b}\|^2 - \|\boldsymbol{a} - \boldsymbol{b}\|^2 = 2(\boldsymbol{a} \cdot \boldsymbol{b} + \overline{\boldsymbol{a} \cdot \boldsymbol{b}})$;

③ $\|\boldsymbol{a} + \boldsymbol{b}\|^2 + \|\boldsymbol{a} - \boldsymbol{b}\|^2 = 2\|\boldsymbol{a}\|^2 + 2\|\boldsymbol{b}\|^2$.

④ $\boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{b} \cdot \boldsymbol{a} \in \mathbb{R} \subset \mathbb{C}$.

59 Mostrare che è possibile definire un unico angolo $\gamma \in [0, \pi]$ tra due vettori $\boldsymbol{a}$ e $\boldsymbol{b}$ di $\mathbb{C}^n$ ponendo

$$\cos\gamma = \frac{\boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{b} \cdot \boldsymbol{a}}{2\|\boldsymbol{a}\|\|\boldsymbol{b}\|}.$$

60 Le matrici

$$P_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad P_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

hanno un ruolo nella meccanica quantistica e sono dette le *matrici di Pauli*. Determinarne gli autovalori e gli autovettori.

Laboratorio...

Nei prossimi esercizi si chiede di dimostrare (in genere per induzione/ricorsione) alcune proposizioni, e di implementare la dimostrazione con un algoritmo al calcolatore. Si tratta di esercizi in genere impegnativi ma non troppo.

61 Si calcoli, a partire da ogni permutazione sugli n elementi $1, 2, \ldots, n$ una matrice che permuta i vettori della base nello stesso modo. Una matrice di questo tipo si dice *matrice di permutazione*.

62 *(LU decomposition, row echelon form)* Implementare l'algoritmo di eliminazione di Gauss per ridurre una matrice A alla sua forma di *echelon* U (cioè in ogni riga di U la successione di zeri iniziali è più lunga di quella della riga precedente): se A è una matrice $n \times m$, allora esistono una matrice di permutazione P ($n \times n$), una matrice unitriangolare inferiore L ($n \times n$) e una matrice triangolare superiore U ($n \times m$) tali che $PA = LU$.

63 *(QR decomposition)* Mostrare che se A è una matrice $n \times m$ di rango massimo, con $m \leq n$, il procedimento di ortogonalizzazione di Gram–Schmidt permette di scrivere $A = QR$ dove Q è una matrice ortogonale e R è una matrice triangolare superiore.

64 *(Diagonalizzazione di una forma quadratica)* Implementare la dimostrazione del Teorema (IV.70) a pagina 202 (si veda la nota (IV.73), e anche gli Esercizi **53** e **54** a pagina 229) nel caso in cui A è una matrice simmetrica definita positiva $n \times n$: costruire una matrice unitriangolare superiore U tale che ${}^tUAU$ è diagonale.

65 *(Cholesky decomposition)* Mostrare che se A è una matrice simmetrica definita positiva $n \times n$, allora esiste una matrice triangolare inferiore L con elementi sulla diagonale > 0 tale che $A = L{}^tL$ (equivalentemente, esiste una matrice triangolare superiore U con elementi sulla diagonale > 0 tale che $A = {}^tUU$. Con la matrice U è possibile scrivere $B = QR$ per una matrice non singolare B, dove Q è ortogonale e R è triangolare superiore: infatti, se B non è singolare allora la matrice simmetrica ${}^tBB$ è definita positiva, e quindi si può scrivere ${}^tBB = {}^tUU$ con U triangolare superiore. Se quindi si pone $Q = BU^{-1}$ e $R = U$ si ha $B = QR$ con R triangolare superiore e ${}^tQQ = {}^tU^{-1}{}^tBBU^{-1} = {}^tU^{-1}({}^tUU)U^{-1} = I$, cioè Q ortogonale). *(Suggerimento: si usi l'Esercizio* ***64****.)*

66 *(Least squares solution)* A volte invece che risolvere il sistema $A\boldsymbol{x} = \boldsymbol{b}$ (che potrebbe anche non essere risolvibile, per esempio se A ha più righe che colonne), si cerca una soluzione *ai minimi quadrati*, cioè una soluzione che minimizza la norma $\|A\boldsymbol{x} - \boldsymbol{b}\|^2$, dove A è una matrice $n \times m$. È chiaro che una soluzione vera è anche una soluzione ai minimi quadrati, ma in genere non è vero il viceversa. Osservare che al variare di $\boldsymbol{x} \in \mathbb{R}^m$ i punti $A\boldsymbol{x} - \boldsymbol{b} \in \mathbb{R}^n$ possono scriversi come combinazioni delle colonne di A

$$-\boldsymbol{b} + x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \ldots + x_n\boldsymbol{a}_n,$$

e determinare la soluzione $\boldsymbol{x}$ ortogonalizzando i vettori $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ con l'ortogonalizzazione di Gram–Schmidt (supponendo che i $\boldsymbol{a}_i$ siano linearmente indipendenti).

❊ ❊ ❊

Soluzioni degli esercizi

Capitolo IV

1 Raccogliendo si ottiene

$$(2x+y-1)\boldsymbol{i}+(x+y+1)\boldsymbol{j}=\boldsymbol{0} \implies \begin{cases} 2x+y=1 \\ x+y=-1, \end{cases}$$

che ha unica soluzione $(x, y) = (2, -3)$.

2 Sono linearmente dipendenti se e solo se esiste una soluzione non nulla del sistema $x\boldsymbol{a}+y\boldsymbol{b}+z\boldsymbol{c} = \boldsymbol{0}$. Si ottiene:

$$\begin{cases} 2x+y+2z=0 \\ -x+2y-11z=0 \\ x-y+7z=0 \end{cases} \iff \begin{bmatrix} x \\ y \\ z \end{bmatrix} = t\begin{bmatrix} -3 \\ 4 \\ 1 \end{bmatrix},$$

cioè $-3\boldsymbol{a}+4\boldsymbol{b}+\boldsymbol{c}=\boldsymbol{0} \implies \boldsymbol{c}=3\boldsymbol{a}-4\boldsymbol{b}$. Altro metodo (che però non consente di scriverne uno in funzione degli altri due) è mostrare che il determinante della matrice che ha per colonne i vettori $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ è nullo.

3 I due vettori $(t, t^2) = t(1, t)$ e $(1, t)$ sono sempre paralleli. Se $t = 0$, si ha $(0, 0)$ e $(1, 0)$, che sono paralleli in senso lato.

4 I tre vettori $\boldsymbol{a}$, $\boldsymbol{b}$ e $\boldsymbol{c}$ sono linearmente dipendenti se e solo se esiste $(x, y, z) \neq (0, 0, 0)$ tale che $x\boldsymbol{a}+y\boldsymbol{b}+z\boldsymbol{c}=\boldsymbol{0}$, mentre $\boldsymbol{a}$, $\boldsymbol{b}+\boldsymbol{a}$, e $\boldsymbol{c}$ sono linearmente indipendenti se e solo se esiste $(\hat{x}, \hat{y}, \hat{z}) \neq (0, 0, 0)$ tale che $\hat{x}\boldsymbol{a}+\hat{y}(\boldsymbol{a}+\boldsymbol{b})+\hat{z}\boldsymbol{c}=\boldsymbol{0}$, cioè $(\hat{x}+\hat{y})\boldsymbol{a}+\hat{y}\boldsymbol{b}+\hat{z}\boldsymbol{c}=\boldsymbol{0}$. Basta quindi porre $\hat{x}+\hat{y}=x$, $\hat{y}=y$, $\hat{z}=z$, e osservare che esistono (x, y, z) non nulli se e solo se esistono $(\hat{x}, \hat{y}, \hat{z})$. Cioè, se si somma $\boldsymbol{a}$ (uno dei tre vettori) ad un altro, la terna di vettori è indipendente se e solo se lo era la prima. Sommando quindi $\boldsymbol{a}$ a $\boldsymbol{b}$ si ottiene $\boldsymbol{a}, \boldsymbol{a}+\boldsymbol{b}, \boldsymbol{c}$. Sommando $\boldsymbol{a}+\boldsymbol{b}$ al terzo, si ottiene $\boldsymbol{a}, \boldsymbol{a}+\boldsymbol{b}, \boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}$. Altro modo:

$$\begin{aligned} &\det(\boldsymbol{a}, \boldsymbol{a}+\boldsymbol{b}, \boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}) \\ &= \det(\boldsymbol{a}, \boldsymbol{a}, \boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}) + \det(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}) \\ &= \det(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a}) + \det(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{b}) + \det(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = \det(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}). \end{aligned}$$

Quindi i due determinanti coincidono. Per i tre vettori $\boldsymbol{a}+\boldsymbol{b}$, $\boldsymbol{b}+\boldsymbol{c}$, $\boldsymbol{c}+\boldsymbol{a}$, osserviamo che

$$\begin{aligned} &\det(\boldsymbol{a}+\boldsymbol{b}, \boldsymbol{b}+\boldsymbol{c}, \boldsymbol{c}+\boldsymbol{a}) \\ &= \det(\boldsymbol{a}, \boldsymbol{b}+\boldsymbol{c}, \boldsymbol{c}+\boldsymbol{a}) + \det(\boldsymbol{b}, \boldsymbol{b}+\boldsymbol{c}, \boldsymbol{c}+\boldsymbol{a}) \\ &= \det(\boldsymbol{a}, \boldsymbol{c}, \boldsymbol{c}+\boldsymbol{a}) + \det(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}+\boldsymbol{a}) + \det(\boldsymbol{b}, \boldsymbol{c}, \boldsymbol{c}+\boldsymbol{a}) \\ &= 0 + \det(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) + \det(\boldsymbol{b}, \boldsymbol{c}, \boldsymbol{a}) = 2\det(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}). \end{aligned}$$

Quindi sono linearmente indipendenti se e solo se lo sono $\boldsymbol{a}$, $\boldsymbol{b}$ e $\boldsymbol{c}$.

5 Basta mostrare che sono linearmente indipendenti, cioè che $\det(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n) \neq 0$. Ma sottraendo $\boldsymbol{e}_1$ dalle restanti colonne, $\boldsymbol{e}_2$ dalle colonne $i = 3, \ldots, n$, e così via fino alla n-esima, si ha

$$\begin{aligned} &\det(\boldsymbol{e}_1, \boldsymbol{e}_1+\boldsymbol{e}_2, \boldsymbol{e}_1+\boldsymbol{e}_2+\boldsymbol{e}_3, \ldots, \boldsymbol{e}_1+\boldsymbol{e}_2+\ldots+\boldsymbol{e}_n) = \\ &\det(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_2+\boldsymbol{e}_3, \ldots, \boldsymbol{e}_2+\ldots+\boldsymbol{e}_n) = \\ &\ldots = \det(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \ldots, \boldsymbol{e}_n) = 1, \end{aligned}$$

e quindi sono linearmente indipendenti. Le coordinate di $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ in questa base sono $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$, con $y_1\boldsymbol{b}_1 + y_2\boldsymbol{b}_2 + \ldots + y_n\boldsymbol{b}_n = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \ldots + x_n\boldsymbol{e}_n$, e quindi

$$\begin{cases} x_1 = y_1+y_2+\ldots+y_n \\ x_2 = y_2+y_3+\ldots+y_n \\ \ldots \\ x_{n-1} = y_{n-1}+y_n \\ x_n = y_n, \end{cases} \implies \begin{cases} y_1 = x_1-x_2 \\ y_2 = x_2-x_3 \\ \ldots \\ y_{n-1} = x_{n-1}-x_n \\ y_n = x_n. \end{cases}$$

6 Il determinante della matrice che ha come colonne i tre vettori vale $-(1+t^3)$, sono quindi linearmente dipendenti se e solo se $t = -1$.

7 Per dimostrare **(a)**, se $\boldsymbol{x} \in \mathrm{Span}(S_1)$, allora $\boldsymbol{x} = c_1\boldsymbol{s}_1 + c_2\boldsymbol{s}_2 + \ldots + c_n\boldsymbol{s}_n$ con $\boldsymbol{s}_i \in S_1$, e quindi $\boldsymbol{x} = c_1\boldsymbol{s}_1 + c_2\boldsymbol{s}_2 + \ldots + c_n\boldsymbol{s}_n$ con $\boldsymbol{s}_i \in S_2$.

Per dimostrare **(b)**, osserviamo che $S_1 \cap S_2 \subseteq S_1$, e quindi per il punto precedente $\mathrm{Span}(S_1 \cap S_2) \subseteq \mathrm{Span}(S_1)$. Analo-

gamente, $S_1 \cap S_2 \subseteq S_2$ e quindi $\mathrm{Span}(S_1 \cap S_2) \subseteq \mathrm{Span}(S_2)$, da cui $\mathrm{Span}(S_1 \cap S_2) \subseteq \mathrm{Span}(S_1) \cap \mathrm{Span}(S_2)$.
Infine, per **(c)** , consideriamo lo spazio vettoriale $\mathbb{R}$, e $S_1 = \{1\}$, $S_2 = \{2\}$. Allora $S_1 \cap S_2 = \emptyset$, e quindi $\mathrm{Span}(S_1 \cap S_2) = 0$, mentre $\mathrm{Span}(S_1) = \mathrm{Span}(S_2) = \mathbb{R}$, per cui $\mathrm{Span}(S_1) \cap \mathrm{Span}(S_2) = \mathbb{R}$.

8 **(a)** Se $a \neq b$, $a \neq 0$, $b \neq 0$, la dimensione è 3 (si veda la soluzione dell'Esercizio (IV.1) a pagina 214). Se $a = 0$ o $b = 0$, ovviamente la corrispondente funzione è uguale a 1, e se $a = b$, le due funzioni e^{ax} e e^{bx} coincidono. **(b)** Supponiamo che esistano quattro coefficienti a, b, c, d tali che

$$a \sin x + b \cos x + c \sin^2 x + d \cos^2 x = 0.$$

Siccome $\sin^2 x + \cos^2 x = 1$, è equivalente che

$$(*) \qquad a \sin x + b \cos x + (c - d) \sin^2 x + d = 0.$$

Deriviamo rispetto a x:

$$\begin{aligned} &a \cos x - b \sin x + 2(c - d) \sin x \cos x = 0 \\ &- a \sin x - b \cos x + 2(c - d)(2 \cos^2 x - 1) = 0 \\ &- a \cos x + b \sin x + 2(c - d)(-4 \sin x \cos x) = 0. \end{aligned}$$

Sommiamo la prima e la terza: $c - d = 0$. Sommiamo la (*) alla seconda: $d = 0$. Rimane $a \sin x + b \cos x = 0$, che implica $a = 0$ (sostituendo $x = \pi/2$) e $b = 0$ (sostituendo $x = 0$). Quindi sono linearmente indipendenti e la dimensione è 4. **(c)** Mostriamo che sono linearmente indipendenti, cioè che la dimensione è 4. Infatti

$$c_1 + c_2 e^{ax} + c_3 x e^{ax} + c_4 x^2 e^{ax} = 0 \implies -c_4 = \frac{c_1 + c_2 e^{ax} + c_3 x e^{ax}}{x^2 e^{ax}}.$$

Prendendo il limite per $x \to +\infty$ (se $a \geq 0$) o per $x \to -\infty$ (se $a < 0$), si ha $c_4 = 0$. Allo stesso modo,

$$-c_3 = \frac{c_1 + c_2 e^{ax}}{x e^{ax}}$$

e prendendo il limite si ha $c_3 = 0$. Continuando così, anche $c_2 = 0$, e quindi $c_1 = 0$. **(d)** Si ha $\cos 2x = \cos^2 x - \sin^2 x = 1 - 2\sin^2 x$. Quindi $2 \sin^2 x = 1 - \cos 2x$, e $\sin^2 x$ è combinazione lineare di 1 e $\cos 2x$. La dimensione non è maggiore di 2. Ma 1 e $\cos 2x$ sono linearmente indipendenti (dato che $\cos 2x$ non è costante), e quindi la dimensione cercata è esattamente 2.

9 Basta dimostrare che le operazioni di somma e di prodotto per uno scalare sono definite anche sul nucleo. Infatti, se $\boldsymbol{v}, \boldsymbol{w} \in \mathrm{Nucleo}(L)$, cioè se $L(\boldsymbol{v}) = L(\boldsymbol{w}) = \mathbf{0}$, allora per linearità $L(\boldsymbol{v} + \boldsymbol{w}) = L(\boldsymbol{v}) + L(\boldsymbol{w}) = \mathbf{0}$ e $L(c\boldsymbol{v}) = cL(\boldsymbol{v}) = c\mathbf{0} = \mathbf{0}$ per ogni $c \in \mathbb{R}$.

10 **(a)** $\dim \mathrm{Nucleo}(f) = 0$, $\dim \mathrm{Immagine}(f) = 2$; **(b)** $\dim \mathrm{Nucleo}(f) = 0$, $\dim \mathrm{Immagine}(f) = 2$; **(c)** Nucleo $= \{(x, y) \in \mathbb{R}^2 : x = 0\} \implies \dim \mathrm{Nucleo}(f) = 1$, $\dim \mathrm{Immagine}(f) = 1$; **(d)** Non è lineare: $f(2, 2) = (4, 2) \neq 2f(1, 1) = 2(1, 0)$; **(e)** Nucleo $= \{(x, y) \in \mathbb{R}^2 : x + y = 0\} \implies \dim \mathrm{Nucleo}(f) = 1$, $\dim \mathrm{Immagine}(f) = 1$.

11 **(a)** Il nucleo è $\{0\}$, l'immagine è $\mathbb{R}^2$. **(b)** Il nucleo è $\{0\}$, l'immagine è $\mathbb{R}^2$. **(c)** Non è lineare: lo $\mathbf{0}$ non va nello $\mathbf{0}$. **(d)** Non è lineare: lo $\mathbf{0}$ non va nello $\mathbf{0}$. **(e)** Non è lineare: lo $\mathbf{0}$ non va nello $\mathbf{0}$.

12 **(a)** rango = 3, nullità = 0; **(b)** rango = 3, nullità = 0. Infatti $x + y = y + z = z + x = 0 \implies x - z = z + x = 0 \implies x = z = 0$; **(c)** Non è lineare: $f(0, 0, 0) \neq (0, 0, 0)$; **(d)** Non è lineare: $f(0, 0, 0) \neq (0, 0, 0)$; **(e)** Non è lineare: $f(0, 2, 0) \neq 2f(0, 1, 0)$.

13 **(a)** Se $p_1, p_2 \in V$, allora $T(p_1 + p_2)(x) = (p_1 + p_2)(x + 1) = p_1(x + 1) + p_2(x + 1) = T(p_1) + T(p_2)$. Se $c \in \mathbb{R}$, allora $T(cp)(x) = cp(x + 1) = cT(p)$. $\implies$ è lineare. $p(x + 1) \equiv 0$ se e soltanto se $p \equiv 0$, cioè il nucleo è banale. **(b)** È lineare, con nucleo i polinomi di grado zero. **(c)** È lineare, con nucleo i polinomi di grado zero. **(d)** Lineare. Il nucleo: 0. Infatti, se $p(x)$ è un polinomio di grado k per cui $p''(x) + 2p'(x) + p(x) = 0$, allora $p''(x)$ (di grado $k - 2$) deve essere uguale a $-2p'(x) - p(x)$, che ha grado k. **(e)** Lineare. Il nucleo: se un polinomio $p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots$ soddisfa l'equazione $x^2 p''(x) - 2xp'(x) + p(x) = 0$, allora

$$\begin{aligned} &(a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots + a_n x^n) \\ &-2x(a_1 + 2a_2 x + 3a_3 x^2 + \ldots + n a_n x^{n-1}) \\ &+x^2(2a_2 + 6a_3 x + \ldots + n(n-1)a_n x^{n-2}), \end{aligned}$$

e quindi

$$a_0 + (-a_1)x + (-a_2)x^2 + \ldots + (1 - 2n + n(n-1))a_n x^n = 0.$$

Ma allora $a_0 = a_1 = \ldots = a_n = 0$, ossia $p = 0$.

14 È lineare:

$$\begin{aligned} T(\{x_n\} + \{y_n\}) &= T(\{x_n + y_n\}) = \\ &= \{x_n + y_n - \lim_{n\to\infty}(x_n + y_n)\} = \{x_n + y_n - \lim_{n\to\infty} x_n - \lim_{n\to\infty} y_n\} \\ &= \{x_n - \lim_{n\to\infty} x_n\} + \{y_n - \lim_{n\to\infty} y_n\} = T(\{x_n\}) + T(\{y_n\}). \end{aligned}$$

$$T(c\{x_n\}) = \{cx_n - \lim_{n\to\infty} cx_n\} = cT(\{x_n\}).$$

Il nucleo: le successioni per cui per ogni n

$$x_n - \lim_{n\to\infty} x_n = 0,$$

cioè le successioni costanti.

15 Si ha, per il teorema fondamentale del calcolo integrale,

$$\frac{d}{dx}\left(\int_0^x p(t)\,dt\right) = p(x),$$

e dunque $DI(p) = p$. Invece

$$ID(p)(x) = \int_0^x p'(t)\,\mathrm{d}t = p(x) - p(0),$$

che è uguale a $p(x)$ se e solo se $p(0) = 0$.

16 **(a)** $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$; **(b)** $\begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$; **(c)** $\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}$; **(d)** L'inversa di $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ ha matrice

$$f^{-1} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \implies f = \begin{bmatrix} -1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 \end{bmatrix}.$$

Infatti, se poniamo $\boldsymbol{i}' = f(\boldsymbol{i})$, $\boldsymbol{j}' = f(\boldsymbol{j})$ e $\boldsymbol{k}' = f(\boldsymbol{k})$, abbiamo

$$\boldsymbol{i}' + \boldsymbol{j}' = \boldsymbol{k}, \quad \boldsymbol{j}' + \boldsymbol{k}' = \boldsymbol{i}, \quad \boldsymbol{k}' + \boldsymbol{i}' = \boldsymbol{j}$$
$$\implies \boldsymbol{i}' + \boldsymbol{j}' = \boldsymbol{k}, \quad \boldsymbol{i}' - \boldsymbol{j}' = \boldsymbol{j} - \boldsymbol{i} \implies 2\boldsymbol{i}' = -\boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k},$$

e analogamente per $\boldsymbol{j}'$ e $\boldsymbol{k}'$.

17 **(a)** $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$; **(b)** $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$; **(c)** $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$; **(d)** $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

18 $B + C = \begin{bmatrix} 4 & 6 \\ 8 & 10 \\ 6 & 8 \end{bmatrix}$, $AB = \begin{bmatrix} 22 & 28 \\ -22 & -28 \end{bmatrix}$, $AC = \begin{bmatrix} 16 & 22 \\ -16 & -22 \end{bmatrix}$, $BA = CA = \begin{bmatrix} -1 & -2 & -3 \\ -1 & -2 & -3 \\ -1 & -2 & -3 \end{bmatrix}$, $A(B + C) = \begin{bmatrix} 38 & 50 \\ -38 & -50 \end{bmatrix}$.

19 $AB - BA = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$. Per dimostrare la seconda affermazione, osserviamo che se Z commuta con *ogni* matrice B, allora in particolare commuta con le quattro matrici elencate. Viceversa, se $ZE_{ij} = E_{ij}Z$ per ogni i, j, allora se B è una matrice arbitraria,

$$\begin{aligned} B &= b_{11}E_{11} + b_{12}E_{12} + b_{21}E_{21} + b_{22}E_{22} \\ \implies ZB &= Z(b_{11}E_{11} + b_{12}E_{12} + b_{21}E_{21} + b_{22}E_{22}) \\ &= b_{11}ZE_{11} + b_{12}ZE_{12} + b_{21}ZE_{21} + b_{22}ZE_{22} \\ &= b_{11}E_{11}Z + b_{12}E_{12}Z + b_{21}E_{21}Z + b_{22}E_{22}Z \\ &= (b_{11}E_{11} + b_{12}E_{12} + b_{21}E_{21} + b_{22}E_{22})Z = BZ. \end{aligned}$$

Risolvendo il sistema di equazioni $E_{ij}Z = ZE_{ij}$ (per tutti gli i, j), si trova che le matrici Z con questa proprietà sono tutte e sole quelle del tipo $cI = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix}$, per qualche $c \in \mathbb{R}$.

20 Mostriamo per induzione che $A^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$, per ogni $n \in \mathbb{Z}$. Se $n = 0$, è vera (supposto di porre $A^0 = I$. Per $n = 1$, anche. Supponiamo che sia vera per $n - 1$: allora

$$A^n = A^{n-1}A = \begin{bmatrix} 1 & n-1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}.$$

Ora, l'inversa A^{-1} è proprio $\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ e analogamente, per $n > 0$,

$$(A^{-1})^n = (A^{-1})^{n-1}A^{-1} = \begin{bmatrix} 1 & -(n-1) \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -n \\ 0 & 1 \end{bmatrix}.$$

21 Invece di fare una dimostrazione trigonometrica (che il Lettore può svolgere autonomamente) identifichiamo $\mathbb{R}^2$ con il piano complesso $\mathbb{C}$. Come sappiamo dal primo volume, la rotazione di angolo ϑ corrisponde allora al prodotto per l'esponenziale complesso $e^{i\vartheta} = \cos\vartheta + i\sin\vartheta$. Posto $(x, y) = x + iy = z$ otteniamo subito $e^{i\vartheta}z = (x\cos\vartheta - y\sin\vartheta) + i(x\sin\vartheta + y\cos\vartheta)$, da cui la tesi.

22 Si ha

$$\begin{aligned} R_\vartheta R_{\vartheta'} &= \begin{bmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{bmatrix}\begin{bmatrix} \cos\vartheta' & -\sin\vartheta' \\ \sin\vartheta' & \cos\vartheta' \end{bmatrix} \\ &= \begin{bmatrix} \cos\vartheta\cos\vartheta' - \sin\vartheta\sin\vartheta' & -\cos\vartheta\sin\vartheta' - \sin\vartheta\cos\vartheta' \\ \sin\vartheta\cos\vartheta' + \cos\vartheta\sin\vartheta' & -\sin\vartheta\sin\vartheta' + \cos\vartheta\cos\vartheta' \end{bmatrix} \\ &= \begin{bmatrix} \cos(\vartheta + \vartheta') & -\sin(\vartheta + \vartheta') \\ \sin(\vartheta + \vartheta') & \cos(\vartheta + \vartheta') \end{bmatrix}. \end{aligned}$$

Si ha quindi $R_\vartheta^n = \begin{bmatrix} \cos n\vartheta & -\sin n\vartheta \\ \sin n\vartheta & \cos n\vartheta \end{bmatrix}$.

23 Se $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, ci sono $2^4 = 16$ modi di avere elementi $a_{ij} = \pm 1$. Siano $\boldsymbol{a}_1$ e $\boldsymbol{a}_2$ le due colonne.

$$\begin{aligned} \boldsymbol{a}_1 = (1, 1) &\implies \boldsymbol{a}_2 = (-1, 1) \text{ oppure } \boldsymbol{a}_2 = (1, -1) \\ \boldsymbol{a}_1 = (-1, -1) &\implies \boldsymbol{a}_2 = (-1, 1) \text{ oppure } \boldsymbol{a}_2 = (1, -1) \\ \boldsymbol{a}_1 = (-1, 1) &\implies \boldsymbol{a}_2 = (1, 1) \text{ oppure } \boldsymbol{a}_2 = (-1, -1) \\ \boldsymbol{a}_1 = (1, -1) &\implies \boldsymbol{a}_2 = (1, 1) \text{ oppure } \boldsymbol{a}_2 = (-1, -1). \end{aligned}$$

In totale, sono otto. Ora, se $\boldsymbol{x}$, $\boldsymbol{y}$ e $\boldsymbol{z}$ sono tre vettori ortogonali, si ha

$$(\boldsymbol{x} + \boldsymbol{y}) \cdot (\boldsymbol{x} + \boldsymbol{z}) = \boldsymbol{x} \cdot \boldsymbol{x} + \boldsymbol{x} \cdot \boldsymbol{z} + \boldsymbol{y} \cdot \boldsymbol{x} + \boldsymbol{y} \cdot \boldsymbol{z} = \|\boldsymbol{x}\|^2.$$

Se quindi sono colonne di dimensione n con termini ± 1, $\|\boldsymbol{x}\|^2 = n$. Per ogni $i = 1, \ldots, n$, $x_i + y_i$ e $x_i + z_i$ possono essere solo uguali a 0 oppure a ± 2 (basta enumerare i quattro casi possibili), e quindi il loro prodotto può essere uguale a 0 o a ±4. Ma allora si ha

$$n = (\boldsymbol{x}+\boldsymbol{y})\cdot(\boldsymbol{x}+\boldsymbol{z}) = \sum_{i=1}^{n}(x_i+y_i)(x_i+z_i),$$

cioè n è uguale alla somma di n termini uguali a 0 o a ±4, e quindi deve essere divisibile per 4.

24 Per mostrare **(a)** , le proprietà commutativa, distributiva e omogenea sono semplici e omettiamo i dettagli per non occupare troppo spazio. Rimane invece da mostrare che è definito positivo, cioè che $\langle f,f\rangle \geq 0$ e $f \neq 0 \implies \langle f,f\rangle > 0$. È chiaro che $\langle f,f\rangle \geq 0$. Supponiamo quindi che $f(x)$ è un polinomio di grado $\leq n$ tale che

$$\sum_{i=0}^{n} f^2(\frac{i}{n}) = 0 \text{ e quindi che } f(\frac{i}{n}) = 0 \text{ per } i = 0,\dots,n.$$

Quindi f è un polinomio di grado n che si annulla negli $n+1$ punti i/n: deve essere il polinomio nullo.

Calcoliamo $\langle 1,x\rangle$, per rispondere a **(b)** :

$$\langle 1,x\rangle = \sum_{i=0}^{n}\frac{i}{n} = \frac{1}{n}\sum_{i=1}^{n} i = \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2}.$$

25 Se $f(x)$ e $g(x)$ sono polinomi, allora il prodotto è un polinomio $p(x)$, e

$$\left|\int_0^\infty e^{-x}p(x)\,\mathrm{d}x\right| < \infty.$$

L'unica proprietà non banalmente verificabile è che è definito positivo: se

$$0 = \langle f,f\rangle = \int_0^\infty e^{-x}f^2(x)\,\mathrm{d}x,$$

allora l'integranda deve essere nulla (perché positiva), e quindi $f(x)\equiv 0$. Calcoliamo il prodotto scalare $\langle x^n,x^m\rangle$.

$$\begin{aligned}\int_0^\infty e^{-x}x^nx^m\,\mathrm{d}x &= [-e^{-x}x^{n+m}]_0^\infty + \int_0^\infty (n+m)e^{-x}x^{n+m-1}\,\mathrm{d}x\\ &= (n+m)\int_0^\infty e^{-x}x^{n+m-1}\,\mathrm{d}x\\ &= \cdots = (n+m)!\int_0^\infty e^{-x}\,\mathrm{d}x = (n+m)!\end{aligned}$$

e quindi $\langle x^n,x^m\rangle = (n+m)!$.

Un polinomio di primo grado $ax+b$ è ortogonale a $1+x$ se e soltanto se

$$0 = \int_0^\infty e^{-x}(ax+b)(1+x)\,\mathrm{d}x = \ldots = 2b+3a.$$

Quindi i polinomi di primo grado ortogonali a $1+x$ sono quelli del tipo

$$p(x) = a(x-3/2).$$

26 Se $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$ è una base ortonormale di W, per il Teorema (IV.45) la proiezione si scrive

$$\mathrm{proj}_W(\boldsymbol{v}) = \sum_{i=1}^{k}\langle \boldsymbol{v},\boldsymbol{e}_i\rangle\boldsymbol{e}_i,$$

da cui segue che

$$\begin{aligned}\mathrm{proj}_W(\boldsymbol{v}+\boldsymbol{w}) &= \sum_{i=1}^{k}\langle \boldsymbol{v},\boldsymbol{e}_i\rangle\boldsymbol{e}_i + \sum_{i=1}^{k}\langle \boldsymbol{w},\boldsymbol{e}_i\rangle\boldsymbol{e}_i\\ &= \sum_{i=1}^{k}\langle \boldsymbol{v}+\boldsymbol{w},\boldsymbol{e}_i\rangle\boldsymbol{e}_i = \mathrm{proj}_W(\boldsymbol{v})+\mathrm{proj}_W(\boldsymbol{w})\\ \mathrm{proj}_W(c\boldsymbol{v}) &= \sum_{i=1}^{k}\langle c\boldsymbol{v},\boldsymbol{e}_i\rangle\boldsymbol{e}_i = c\sum_{i=1}^{k}\langle \boldsymbol{v},\boldsymbol{e}_i\rangle\boldsymbol{e}_i = c\,\mathrm{proj}_W(\boldsymbol{v})\end{aligned}$$

27 Per la disuguaglianza di Cauchy–Schwarz, per ogni N si ha

$$\left(\sum_{n=1}^{N}x_ny_n\right)^2 \leq \left(\sum_{n=1}^{N}x_n^2\right)\left(\sum_{n=1}^{N}y_n^2\right),$$

e per la disuguaglianza triangolare

$$\sum_{n=1}^{N}(x_n+y_n)^2 \leq \left(\sqrt{\sum_{n=1}^{N}x_n^2} + \sqrt{\sum_{n=1}^{N}y_n^2}\right)^2.$$

Allora se $\{x_n\}$ e $\{y_n\}$ hanno quadrati sommabili, anche la somma $\{x_n+y_n\}$ ha quadrato sommabile, ed è definito il prodotto scalare $\sum_{n=1}^\infty x_ny_n$ definito sopra. Per finire di mostrare che l^2 è uno spazio vettoriale, basta osservare che $\sum_{n=1}^\infty (cx_n)^2 = c^2\sum_{n=1}^\infty x_n^2$, e quindi anche il prodotto per uno scalare è ben definito. Per mostrare che è il prodotto scalare definito sopra è ben definito, dobbiamo mostrare le proprietà commutativa (ovvia), distributiva (ovvia) e omogenea (ovvia) e che è definito positivo (ovvia). Le ovvietà, com'è ovvio, sono lasciate al Lettore. L'angolo γ tra le due successioni $x_n = 1/n$ e $y_n = 1/(n+1)$:

$$\langle\{x_n\},\{y_n\}\rangle = \sum_{n=1}^{\infty}\frac{1}{n(n+1)} = \sum_{n=1}^{\infty}\frac{1}{n} - \frac{1}{n+1} = 1;$$

$$\|\{x_n\}\|^2 = \sum_{n=1}^{\infty}\frac{1}{n^2} = \frac{\pi^2}{6}\|\{y_n\}\|^2 = \sum_{n=1}^{\infty}\frac{1}{(n+1)^2} = \frac{\pi^2}{6}-1,$$

e quindi

$$\cos\gamma = \frac{\langle\{x_n\},\{y_n\}\rangle}{\|\{x_n\}\|\|\{y_n\}\|} = \frac{1}{\pi/\sqrt{6}\sqrt{\pi^2/6-1}},$$

da cui $\gamma \cong 0.2418$. La distanza tra le due successioni è

$$\|\{x_n - y_n\}\|^2 = \sum_{n=1}^{\infty}(1/n - 1/n+1)^2 = \frac{\pi^2}{3} - 3.$$

28 $$\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = \begin{vmatrix} 1 & 1-1 & 1-1 \\ a & b-a & c-a \\ a^2 & b^2-a^2 & c^2-a^2 \end{vmatrix}$$
$$= \begin{vmatrix} b-a & c-a \\ b^2-a^2 & c^2-a^2 \end{vmatrix} = (b-a)(c-a)\begin{vmatrix} 1 & 1 \\ b+a & c+a \end{vmatrix}$$
$$= (b-a)(c-a)(c+a-b-a) = (b-a)(c-a)(c-b).$$

Per la **(b)** , sottraiamo alla seconda riga a volte la prima, alla terza riga a volte la seconda e alla quarta a volte la terza; otteniamo

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ a-a & b-a & c-a & d-a \\ a^2-a^2 & b^2-ab & c^2-ac & d^2-ad \\ a^3-a^3 & b^3-ab^2 & c^3-ac^2 & d^3-ad^2 \end{vmatrix}$$

$$= \begin{vmatrix} b-a & c-a & d-a \\ b(b-a) & c(c-a) & d(d-a) \\ b^2(b-a) & c^2(c-a) & d^2(d-a) \end{vmatrix} = (b-a)(c-a)(d-a)$$
$$\begin{vmatrix} 1 & 1 & 1 \\ b & c & d \\ b^2 & c^2 & d^2 \end{vmatrix} = (b-a)(c-a)(d-a)(c-b)(d-b)(d-c).$$

Queste matrici sono dette *matrici di Vandermonde*. Passiamo alla **(c)** :

$$\begin{vmatrix} x & 1 & 0 & 0 \\ 1 & x & 1 & 0 \\ 0 & 1 & x & 1 \\ 0 & 0 & 1 & x \end{vmatrix} = x\begin{vmatrix} x & 1 & 0 \\ 1 & x & 1 \\ 0 & 1 & x \end{vmatrix} - \begin{vmatrix} x & 1 \\ 1 & x \end{vmatrix}$$
$$= x^2(x^2-2) - (x^2-1) = x^4 - 3x^2 + 1.$$

Il determinante della matrice **(d)** vale invece 1.

29 Nella matrice di esempio i blocchi sono

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

e quindi

$$M = \begin{bmatrix} A & 0_{k\times k} \\ B & C \end{bmatrix}.$$

Supponiamo inizialmente che $B = 0$. Allora

$$M = \begin{bmatrix} A & 0_{k\times k} \\ B & C \end{bmatrix} = \begin{bmatrix} A & 0_{k\times k} \\ 0_{k\times k} & I_k \end{bmatrix}\begin{bmatrix} I_k & 0_{k\times k} \\ 0_{k\times k} & C \end{bmatrix},$$

e quindi per il Teorema di Binet

$$\det(M) = \det\begin{bmatrix} A & 0_{k\times k} \\ 0_{k\times k} & I_k \end{bmatrix}\det\begin{bmatrix} I_k & 0_{k\times k} \\ 0_{k\times k} & C \end{bmatrix}.$$

Ora, sia per l'unicità del determinante che per lo sviluppo di Laplace, si ha

$$\det\begin{bmatrix} A & 0_{k\times k} \\ 0_{k\times k} & I_k \end{bmatrix} = \det A, \quad \det\begin{bmatrix} I_k & 0_{k\times k} \\ 0_{k\times k} & C \end{bmatrix} = \det C,$$

e quindi se $B = 0$ $\det M = \det A \det C$.
Supponiamo invece che $B \neq 0$, e scriviamo M come

$$M = \begin{bmatrix} A & \mathbf{0}_{k\times k} \\ \mathbf{0}_{k\times k} & I_k \end{bmatrix}\begin{bmatrix} I_k & \mathbf{0}_{k\times k} \\ B & I_k \end{bmatrix}\begin{bmatrix} I_k & \mathbf{0}_{k\times k} \\ \mathbf{0}_{k\times k} & C \end{bmatrix}.$$

Allora la tesi segue se riusciamo a mostrare che

$$\begin{vmatrix} I_k & \mathbf{0}_{k\times k} \\ B & I_k \end{vmatrix} = 1 \text{ che se } k = 2 \text{ è } \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ b_{11} & b_{12} & 1 & 0 \\ b_{21} & b_{22} & 0 & 1 \end{vmatrix} = 1.$$

Ma questa è una matrice triangolare (inferiore), il cui determinante è uguale al prodotto degli elementi sulla diagonale (dimostrarlo!), che è uno.

30 **(a)** $1/2\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$; **(b)** $\begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}$; **(c)** $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

31 Supponiamo che esistano α, $\beta \in \mathbb{R}$ non entrambi nulli per cui $\alpha a(x) + \beta b(x) \equiv 0$.
Allora, derivando una volta otteniamo $\alpha a'(x) + \beta b'(x) \equiv 0$, e quindi

$$\alpha\begin{bmatrix} a(x) \\ a'(x) \end{bmatrix} + \beta\begin{bmatrix} b(x) \\ b'(x) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

per cui $W(a, b; x) = 0$.
Analogamente, per a, b, c, si deriva due volte e si ottiene il sistema

$$\alpha a(x) + \beta b(x) + \gamma c(x) = 0$$
$$\alpha a'(x) + \beta b'(x) + \gamma c'(x) = 0$$
$$\alpha a''(x) + \beta b''(x) + \gamma c''(x) = 0,$$

e quindi i tre vettori

$$\begin{bmatrix} a(x) \\ a'(x) \\ a''(x) \end{bmatrix}, \quad \begin{bmatrix} b(x) \\ b'(x) \\ b''(x) \end{bmatrix}, \quad \begin{bmatrix} c(x) \\ c'(x) \\ c''(x) \end{bmatrix}$$

sono linearmente dipendenti. Il determinante della matrice $W(a, b, c; x)$ è quindi identicamente nullo. Consideriamo le funzioni suggerite $a(x) = x^3$ $b(x) = |x|^3$, per cui il determinante della matrice wronskiana è

$$\det(W(a, b; x)) = \begin{cases} \begin{vmatrix} x^3 & x^3 \\ 3x^2 & 3x^2 \end{vmatrix} = 0 \text{ se } x \geq 0, \\ \begin{vmatrix} x^3 & -x^3 \\ 3x^2 & -3x^2 \end{vmatrix} = 0 \text{ se } x < 0, \end{cases}$$

cioè è identicamente nullo. Ma $a(x)$ e $b(x)$ non sono linearmente dipendenti: se esistessero α, β tali che $\alpha a(x) + \beta b(x) = 0$, deve essere $\alpha x^3 + \beta |x|^3 = 0$, e quindi se $x = 1$ si ha $\alpha + \beta = 0$, mentre per $x = -1$ si ha $\beta - \alpha = 0$: segue che $\alpha = \beta = 0$.

32 Se $Lv = \lambda_1 v$ e $Lv = \lambda_2 v$, per $v \neq 0$ e $\lambda_1, \lambda_2 \in \mathbb{R}$, allora $\lambda_1 v = \lambda_2 v$, e quindi $(\lambda_1 - \lambda_2) v = \mathbf{0}$. Dato che $v \neq \mathbf{0}$, deve quindi essere $\lambda_1 - \lambda_2 = 0$.

33 Il polinomio caratteristico di A è

$$\det(A - tI) = \begin{vmatrix} 6-t & -3 & 3 \\ 1 & 4-t & 3 \\ -1 & 1 & 2-t \end{vmatrix} = (6-t) \begin{vmatrix} 4-t & 3 \\ 1 & 2-t \end{vmatrix}$$

$$- \begin{vmatrix} -3 & 3 \\ 1 & 2-t \end{vmatrix} - \begin{vmatrix} -3 & 3 \\ 4-t & 3 \end{vmatrix} = 60 - 47t + 12t^2 - t^3.$$

Come si trovano le radici dei polinomi di terzo grado? Sappiamo che ci deve essere almeno una radice, e *per tentativi* arriviamo alle tre soluzioni $t = 3, 4, 5$ *(non è vero che le abbiamo trovate per tentativi: ci sono molti metodi, tra cui anche una formula risolutiva, analoga a quella per l'equazione di secondo grado)*. Quindi ci sono tre autovalori distinti $\lambda_1 = 3$, $\lambda_2 = 4$ e $\lambda_3 = 5$. Risolviamo i tre sistemi per trovare gli autovettori corrispondenti:

$$(A - 3I)\boldsymbol{x} = \mathbf{0} \iff \boldsymbol{x}_1 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \ (A - 4I)\boldsymbol{x} = \mathbf{0} \iff \boldsymbol{x}_2 = \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}$$

$$(A - 5I)\boldsymbol{x} = \mathbf{0} \iff \boldsymbol{x}_3 = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}.$$

La matrice diagonalizzante è quindi

$$S = \begin{bmatrix} 2 & 3 & 3 \\ 1 & 1 & 0 \\ -1 & -1 & -1 \end{bmatrix} \text{ con inversa } \begin{bmatrix} -1 & 0 & -3 \\ 1 & 1 & 3 \\ 0 & -1 & -1 \end{bmatrix}$$

Il Lettore è invitato a controllare che:

$$S^{-1}AS = \ldots = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix}.$$

In modo analogo si trovano gli autovalori e gli autovettori di B: il polinomio caratteristico è $\det(B - tI) = -t^3 + 2t^2 + t - 2 = (1 - t)(t^2 - t - 2)$ da cui si trovano gli autovalori $\lambda_1 = -1$, $\lambda_2 = 1$, $\lambda_3 = 2$ e gli autovettori $\boldsymbol{x}_1 = (0, 1, 1)$, $\boldsymbol{x}_2 = (1, 1, 1)$, $\boldsymbol{x}_3 = (1, 0, 1)$. Quindi

$$S = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{e} \quad S^{-1}BS = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

34 Si ha $D(f(x)) = \lambda f(x)$ se e soltanto se

$$f'(x) = \lambda f(x) \iff \frac{f'(x)}{f(x)} = \lambda \iff \frac{d}{dx} \log|f(x)| = \frac{d}{dx} \lambda x,$$

e quindi gli autovettori corrispondenti agli autovalori $\lambda \neq 0$ sono

$$f_\lambda(x) = ce^{\lambda x}$$

per qualche $c \in \mathbb{R}$, mentre quelli corrispondenti all'autovalore $\lambda = 0$ sono gli elementi del nucleo di D, cioè le funzioni costanti.
Gli autovettori di T sono le funzioni $f(x)$ tali che

$$xf'(x) = \lambda f(x),$$

cioè per $\lambda \neq 0$

$$\frac{d}{dx} \log \frac{f'(x)}{|f(x)|} = \frac{\lambda}{x} \implies f(x) = c|x|^\lambda.$$

I valori di λ per cui sono funzioni derivabili infinite volte su $\mathbb{R}$ sono solo gli interi, per cui gli autovalori sono $\lambda_n = n$, con $n \in \mathbb{N}$,e gli autovettori corrispondenti sono i monomi x^n.

35 Sia M una matrice $n \times n$. Mostriamo per induzione che il determinante è un polinomio di grado n nei coefficienti m_{ij} di M. Se $n = 1$, allora $M = (m_{11})$ e $\det M = m_{11}$, quindi l'affermazione è vera. Supponiamo che sia vero per $n - 1$. Allora, sviluppando il determinante di M nella prima colonna come in (IV.56), a pagina 192, si ha

$$\det M = m_{1,1}M_{1,1} - m_{2,1}M_{2,1} + \ldots + (-1)^{n-1} m_{n,1}M_{n,1},$$

dove i $M_{i,1}$ sono i minori (cioè i determinanti delle sottomatrici $n - 1 \times n - 1$) ottenuti cancellando la prima colonna e la i-esima riga. Per ipotesi induttiva ogni minore $M_{i,1}$ è un polinomio di grado $n - 1$ nei coefficienti della matrice, e quindi $\det M$ è un polinomio di grado n. In altre parole, è somma di prodotti dei m_{ij} e non ci sono più di n termini m_{ij} in ogni addendo. Con la stessa tecnica è possibile vedere che è un polinomio *omogeneo* di grado n nelle variabili m_{ij}, cioè tutti i monomi hanno grado n. Ora, il polinomio caratteristico si ottiene sostituendo a m_{11} il polinomio di primo grado $m_{11} - \lambda$, a m_{22} il polinomio $m_{22} - \lambda$, ... La composizione di due polinomi è ancora un polinomio, e quindi il $\det(M - \lambda I)$ è un polinomio di λ. Il suo grado non può superare n, naturalmente. Ora, il suo grado sarà n se nell'espressione del determinante compare il prodotto $m_{11}m_{22} \ldots m_{nn}$. Come sopra, è possibile dimostrarlo per induzione utilizzando lo sviluppo sulla prima colonna.

36 Sappiamo che $\det M = \det {}^tM$, e quindi

$$p_A(\lambda) = \det(A - \lambda I) = \det({}^tA - \lambda {}^tI) = \det({}^tA - \lambda I) = p_{{}^tA}(\lambda).$$

37 Uno scalare $\lambda \in \mathbb{R}$ è autovalore se e soltanto se esiste un autovettore, cioè se e soltanto se l'equazione $M\boldsymbol{x} = \lambda \boldsymbol{x}$ ha una soluzione non nulla. Se la riscriviamo come $(M - \lambda I)\boldsymbol{x} = \mathbf{0}$, allora λ è un autovalore se e soltanto se la matrice $M - \lambda I$ ha determinante zero (si applichi il Teorema (IV.59)), cioè se e soltanto se λ annulla il polinomio caratteristico $p_M(\lambda) = \det(M - \lambda I)$. Ora, se M_1 e M_2 sono simili (cioè se esiste una matrice non singolare B tale che $M_2 = B^{-1}M_1B$), allora

$$\begin{aligned}\det(M_2 - \lambda I) &= \det(B^{-1}M_1B - \lambda I)\\ &= \det(B^{-1}M_1B - B^{-1}\lambda IB)\\ &= \det(B^{-1}(M_1 - \lambda I)B) = \det(B^{-1})\det(M_1 - \lambda I)\det(B)\\ &= \det(B)^{-1}\det(B)\det(M_1 - \lambda I) = \det(M_1 - \lambda I),\end{aligned}$$

cioè $p_{M_1}(\lambda) = p_{M_2}(\lambda)$.

38 Valutando $p_A(\lambda)$ in $\lambda = 0$ si ha subito $c_0 = p_A(0) = \det(A)$. Per procedere, ricordiamo che in $\mathbb{C}$ ogni polinomio di grado n si fattorizza in n fattori lineari. Per l'Esercizio **35** il polinomio $p_A(\lambda)$ ha grado esattamente n, e quindi si fattorizza come

$$(+) \qquad p_A(\lambda) = c_n(\lambda - \lambda_1)(\lambda - \lambda_2)\dots(\lambda - \lambda_n),$$

dove i λ_i sono le n radici (contate con la loro molteplicità) di $p_A(\lambda)$. Le radici di $p_A(\lambda)$ sono gli n autovalori (complessi) di A, quindi i λ_i sono gli autovalori di A. Mostriamo ora per induzione che

$$p_A(\lambda) = \det(A) + \dots + \left[(-1)^{n-1}\sum_{i=1}^{n} a_{ii}\right]\lambda^{n-1} + (-1)^n\lambda^n,$$

cioè che $c_n = (-1)^n$ e $c_{n-1} = (-1)^{n-1}\sum_{i=1}^n a_{ii}$. Se $n = 2$, allora $\det(A - \lambda I) = (a_{11}a_{22} - a_{12}a_{21}) - (a_{11} + a_{22})\lambda + \lambda^2$, e quindi la proposizione è vera. Supponiamola vera per $n - 1$, e sviluppiamo lungo la prima colonna il determinante $\det(A - \lambda I)$ (Proposizione (IV.56)):

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22}-\lambda & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = \\ (a_{11}-\lambda)\begin{vmatrix} a_{22}-\lambda & \cdots & a_{2n}\\ a_{32} & \cdots & a_{3n}\\ \vdots & \ddots & \vdots\\ a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} \\ -a_{2,1}\begin{vmatrix} a_{12} & \cdots & a_{1n}\\ a_{32} & \cdots & a_{3n}\\ \vdots & \ddots & \vdots\\ a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} + \dots,$$

cioè

$$(*) \qquad p_A(\lambda) = (a_{11} - \lambda)p_{\hat{A}}(\lambda) + R(\lambda),$$

dove $\hat{A}$ è la sottomatrice quadrata ottenuta cancellando la prima riga e la prima colonna da A, mentre il resto $R(\lambda)$ è un polinomio di grado $n - 2$. Per ipotesi di induzione

$$p_{\hat{A}}(\lambda) = \hat{c}_0 + \dots + \left[(-1)^{n-2}\sum_{i=1}^{n-1}\hat{a}_{ii}\right]\lambda^{n-2} + (-1)^{n-1}\lambda^{n-1},$$

dove $\hat{a}_{ii} = a_{i+1,i+1}$ per ogni i, e quindi per la (*)

$$\begin{aligned}p_A(\lambda) &= (a_{11} - \lambda)p_{\hat{A}}(\lambda) + R(\lambda)\\ &= (a_{11} - \lambda)(\dots + \hat{c}_{n-2}\lambda^{n-2} + \hat{c}_{n-1}\lambda^{n-1}) + R(\lambda)\\ &= \dots + (a_{11}\hat{c}_{n-1} - \hat{c}_{n-2})\lambda^{n-1} - \hat{c}_{n-1}\lambda^n.\end{aligned}$$

Quindi $c_n = -\hat{c}_{n-1} = -(-1)^{n-1} = (-1)^n$ e

$$\begin{aligned}c_{n-1} &= a_{11}\hat{c}_{n-1} - \hat{c}_{n-2} = (-1)^{n-1}a_{11} - \left[(-1)^{n-2}\sum_{i=1}^{n-1}\hat{a}_{ii}\right]\\ &= (-1)^{n-1}\left[a_{11} + \sum_{i=1}^{n-1}\hat{a}_{ii}\right] = (-1)^{n-1}\sum_{i=1}^{n}a_{ii}.\end{aligned}$$

Ora, se A e B sono matrici simili, allora i polinomi caratteristici coincidono $p_A(\lambda) = p_B(\lambda)$, per cui i coefficienti dei polinomi sono uguali, e in particolare le tracce $\operatorname{tr} A$ e $\operatorname{tr} B$ (che compaiono come coefficienti di λ^{n-1}) sono uguali. Inoltre, dalla $(+)$ deduciamo che

$$\begin{aligned}p_A(\lambda) &= (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)\dots(\lambda - \lambda_n)\\ &= (-1)^n\big(\lambda^n - (\lambda_1 + \lambda_2 + \dots + \lambda_n)\lambda^{n-1} + \dots\\ &\quad \dots + (-1)^n\lambda_1\lambda_2\dots\lambda_n\big),\end{aligned}$$

da cui possiamo finalmente dedurre che

① La traccia è la somma degli autovalori $\operatorname{tr} A = \sum_{i=1}^n \lambda_i$.

② Il determinante è il prodotto degli autovalori $\det A = \prod_{i=1}^n \lambda_i$.

39 Siano a_{ij} e b_{ij} i coefficienti delle matrici A e B. Allora

$$\operatorname{tr}(A + B) = \sum_{i=1}^n (a_{ii} + b_{ii}) = \sum_{i=1}^n a_{ii} + \sum_{i=1}^n b_{ii} = \operatorname{tr}(A) + \operatorname{tr}(B).$$

$$\operatorname{tr}(cA) = \sum_{i=1}^n ca_{ii} = c\sum_{i=1}^n a_{ii} = c\operatorname{tr}(A).$$

$$\operatorname{tr}({}^{t}A) = \sum_{i=1}^n a_{ii} = \operatorname{tr}(A).$$

$$\operatorname{tr}(AB) = \sum_{i=1}^n\sum_{k=1}^n a_{ik}b_{ki} = \sum_{k=1}^n\sum_{i=1}^n b_{ki}a_{ik} = \operatorname{tr}(BA).$$

40 La traccia di una matrice quadrata coincide con la traccia della sua trasposta, per cui si ha

$$\langle A, B\rangle = \operatorname{tr}({}^{t}AB) = \operatorname{tr}({}^{t}({}^{t}AB)) = \operatorname{tr}({}^{t}BA) = \langle B, A\rangle.$$

Si ha inoltre $\operatorname{tr}(cA) = c \operatorname{tr}(A)$, per cui si vede facilmente che

$$\langle cA, B\rangle = \langle A, cB\rangle = c\langle A, B\rangle.$$

Si conclude osservando che se $A \neq 0$, allora $A \neq \mathbf{0} \implies \langle A, A\rangle > 0$. Infatti, se $A = (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m)$ e $B = (\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_m)$ (dove $\boldsymbol{a}_i, \boldsymbol{b}_i \in \mathbb{R}^n$), l'elemento al posto (i, j) della matrice $m \times m$ $C = {}^{\mathrm{t}}AB$ è il prodotto scalare standard

$$c_{ij} = \boldsymbol{a}_i \cdot \boldsymbol{b}_j = {}^{\mathrm{t}}\boldsymbol{a}_i \boldsymbol{b}_j,$$

e dunque se $A \neq 0$

$$\langle A, A\rangle = \operatorname{tr} {}^{\mathrm{t}}AA = \sum_{i=1}^{m} c_{ii} = \sum_{i=1}^{m} \boldsymbol{a}_i \cdot \boldsymbol{a}_i = \sum_{i=1}^{m} ||\boldsymbol{a}_i||^2 > 0.$$

41 Hanno tutte solo l'autovalore 1, dato che i polinomi caratteristici sono tutti uguali a $(1-\lambda)^2$. Calcoliamo gli autovettori corrispondenti. Sappiamo che per B tutti i vettori sono autovettori relativi all'autovalore 1. Non può essere la stessa cosa per A e C, dato che non sono la matrice identica. Controlliamo, comunque:

$$(A - I)\boldsymbol{x} = \mathbf{0} \iff \begin{cases} x_2 = 0 \\ 0 = 0. \end{cases}$$

L'unico autovettore di A è (a meno di multipli scalari) $\boldsymbol{x}_A = (1, 0)$.

$$(C - I)\boldsymbol{x} = \mathbf{0} \iff \begin{cases} 0 = 0 \\ x_1 = 0. \end{cases}$$

L'unico autovettore di C è $\boldsymbol{x}_C = (0, 1)$. Quindi sicuramente B non è simile né ad A né a C, e A e C potrebbero essere simili. Cerchiamo un cambio di coordinate, cioè una matrice invertibile S tale che $S^{-1}AS = C$. La trasformazione S dovrebbe mandare l'autovettore di A in quello di C, per cui proviamo con una matrice del tipo

$$S = \begin{bmatrix} 0 & 1 \\ 1 & s \end{bmatrix} \implies S^{-1} = \begin{bmatrix} -s & 1 \\ 1 & 0 \end{bmatrix}$$

Si vede che

$$S^{-1}AS = \dots = \begin{bmatrix} 1-s & -s^2 \\ 1 & 1+s \end{bmatrix},$$

e quindi per $s = 0$ si ha proprio $S^{-1}AS = C$, cioè A e C sono simili.

42 Il polinomio caratteristico di A è

$$p_A(\lambda) = -t^3 + t^2 + t - 1 = -(t-1)^2(t+1).$$

Gli autovalori sono 1 e −1. Cerchiamo una base di autovettori: l'autovettore dell'autovalore −1 risolve

$$(A + I)\boldsymbol{x} = \mathbf{0} \implies \boldsymbol{x}_{-1} = (-1, 0, 1).$$

Cerchiamo gli autovettori relativi all'autovalore 1:

$$(A - I)\boldsymbol{x} = \mathbf{0} \implies \boldsymbol{x}_1 = (0, 1, 0), \boldsymbol{x}_{1'} = (1, 0, 1).$$

Abbiamo trovato due autovalori, ma tre autovettori linearmente indipendenti, e quindi la matrice è diagonalizzabile (la base diagonalizzante è la terna di autovettori).
La matrice B ha due autovalori distinti (1 e 3), e quindi per il Corollario (IV.81) è diagonalizzabile.
La matrice C invece ha un solo autovalore $\lambda = 2$. Cerchiamone gli autovettori:

$$(C - 2I)\boldsymbol{x} = \mathbf{0} \implies \boldsymbol{x}_2 = (1, 0).$$

Ne abbiamo trovato uno solo, e per fare una base ne servono due: niente da fare, non è diagonalizzabile.
Il polinomio caratteristico di D è $t^2 - 2t + 5$, che non ha radici reali. Quindi D non ha autovalori, né tantomeno può avere autovettori (reali), e quindi non è diagonalizzabile.

43 Se si pone $\boldsymbol{x} = \boldsymbol{e}_i$ e $\boldsymbol{y} = \boldsymbol{e}_j$ (dove gli $\boldsymbol{e}_i$ sono i vettori della base standard), allora

$$\langle L^*\boldsymbol{e}_i, \boldsymbol{e}_j\rangle = {}^{\mathrm{t}}\boldsymbol{e}_i {}^{\mathrm{t}}L^* \boldsymbol{e}_j = l^*_{ji}, \quad \langle \boldsymbol{e}_i, L\boldsymbol{e}_j\rangle = {}^{\mathrm{t}}\boldsymbol{e}_i L \boldsymbol{e}_j = l_{ij}.$$

44 La matrice di adiacenza del grafo è

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

Supponiamo per induzione che il coefficiente a'_{ij} di A^{n-1} sia il numero di percorsi diversi che vanno in $n-1$ passi dal vertice i al vertice j: allora il coefficiente (i, j)-esimo del prodotto $A^n = A^{n-1}A$ è uguale a

$$\sum_{k=1}^{6} a'_{ik} a_{kj},$$

e dato che a_{kj} è uno quando k è collegato a j e 0 altrimenti, il numero di percorsi diversi di n passi è la somma di quelli in $n-1$ passi fino a k_1, quello in $n-1$ passi fino a k_2, ..., quello in $n-1$ passi fino a k_l, dove l è il numero di vertici collegati con j. Per esempio, si può calcolare A^{10} *(con un piccolo algoritmo da programmarsi o con un CAS)*, e vedere che ci sono 8759 modi diversi di partire e arrivare a 1 in 10 passi.

45 Si ha

$$A = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \; B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \text{ quindi } A{}^{\mathrm{t}}B = (Ab_1, Ab_2, Ab_3),$$

e quindi i tre vettori colonna di A^tB sono $b_1\boldsymbol{a}, b_2\boldsymbol{a}, b_3\boldsymbol{a}$, che chiaramente sono linearmente dipendenti.

47 Osserviamo innanzitutto che, in generale, la dimensione di $\mathrm{Span}(v_1, v_2, \dots, v_k)$ è il massimo numero dei vettori v_i linearmente indipendenti. Infatti, supponiamo che i primi d vettori siano linearmente indipendenti ma se ne aggiungiamo uno non sono più linearmente indipendenti: quest'ultimo non può che essere combinazione lineare dei precedenti. Ragionando così con tutti i vettori che mancano, mostriamo che si può generare $\mathrm{Span}(v_1, v_2, \dots, v_k)$ con i summenzionati primi d, cioè che questi costituiscono una base. Se k colonne di $\hat{A}$ sono linearmente indipendenti, allora le corrispondenti k colonne di A (che le estendono) devono essere linearmente indipendenti, e quindi il massimo numero di colonne di $\hat{A}$ linearmente indipendenti non può superare il massimo numero di colonne linearmente indipendenti di A, da cui la tesi.

48 Mostriamo per prima cosa il corrispondente dell'Esercizio **51** di pagina 164, cioè che quattro punti A, B, C, D di $\mathbb{R}^n$ sono complanari se e solo se esistono a, b, c, d non tutti nulli tali che $a+b+c+d = 0$ e $aA+bB+cC+dD = \mathbf{0}$. Infatti, supponiamo che siano complanari. Se ce ne sono tre allineati (per esempio, A, B e C), allora esistono $(a, b, c) \neq (0, 0, 0)$ tali che $a + b + c = 0$ e $aA + bB + cC = \mathbf{0}$, e quindi ponendo $d = 0$ si ha $a + b + c + d = 0$ e $aA + bB + cC + dD = \mathbf{0}$. Quindi supponiamo che non ce ne siano tre allineati: esiste un unico piano per A, B e C. Il quarto punto appartiene a tale piano se e solo se esistono (s, t) tali che $D = A + s(B - A) + t(C - A)$, cioè

$$(1 - s - t)A + sB + tC - D = \mathbf{0},$$

che è $aA+bB+cC+dD = \mathbf{0}$ con $(a, b, c, d) = (1-s-t, s, t, -1)$. Viceversa, se $aA + bB + cC + dD = \mathbf{0}$ e $a + b + c + d = 0$, ma uno dei coefficienti a, b, c o d è zero (per esempio d) allora gli altri tre punti sono allineati (in questo esempio, A, B, e C) e dunque esiste un piano che contiene i quattro punti. Se nessuno dei coefficienti a, b, c e d è zero, allora per esempio dividendo per $-d$ si ha $D = a'A + b'B + c'C$. con $a' = -a/d$, $b' = -b/d$ e $c' = -c/d$, e $a' + b' + c' = 1$. Ma allora ponendo $s = b'$, $t = c'$ si ha $a' = 1 - b' - c' = 1 - s - t$ e

$$D = (1 - s - t)A + sB + tC = A + s(B - A) + t(C - A),$$

cioè i quattro punti sono complanari. Ora, in $\mathbb{R}^3$ esiste una soluzione $(a, b, c, d) \neq (0, 0, 0, 0)$ del sistema

$$a\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} + b\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} + c\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} + d\begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$
$$a + b + c + d = 0,$$

che può essere scritto anche come

$$a\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ 1 \end{bmatrix} + b\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ 1 \end{bmatrix} + c\begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ 1 \end{bmatrix} + d\begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

oppure

$$\begin{cases} aa_1 + bb_1 + cc_1 + dd_1 = 0 \\ aa_2 + bb_2 + cc_2 + dd_2 = 0 \\ aa_3 + bb_3 + cc_3 + dd_3 = 0 \\ a + b + c + d = 0 \end{cases}$$

Ma per il Teorema (IV.59), questo sistema ha soluzioni non nulle se e solo se

$$\det\begin{bmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ 1 & 1 & 1 & 1 \end{bmatrix} = 0.$$

Oppure, potremmo usare il fatto che i quattro punti sono complanari se e solo se il volume del parallelepipedo che generano i vettori differenza $B - A$, $C - A$, $D - A$ è nullo (perché?) e sottrarre la prima colonna alle altre colonne della matrice, per ottenere un determinante $3x3$ che sappiamo essere il volume con segno di un parallelepipedo.

49 Per induzione: se $n = 1$, allora l'unico polinomio è x_1^2. Supponiamo quindi che ci siano $(n-1)n/2$ monomi di grado due in $n-1$ variabili $x_1, \dots, x_{n-1}$. I monomi di grado due nelle n variabili $x_1, \dots, x_n$ sono quindi $(n-1)n/2$ (quelli in cui x_n non compare) più quelli in cui compare x_n, che sono esattamente n (x_n per una delle n variabili), cioè $(n-1)n/2 + n = n(n+1)/2$.

50 Se $u_{ij} = 0$ quando $i > j$ e $v_{ij} = 0$ quando $i > j$, allora $u_{ij} + v_{ij} = 0$ quando $i > j$, cioè $U + V$ è triangolare superiore se lo sono U e V. Per il prodotto, il coefficiente al posto (i, j) di UV è

$$(*) \qquad \sum_{k=1}^{n} u_{ik}v_{kj}.$$

Sia $i > j$: il termine u_{ik} è nullo quando $k < i$, il termine v_{kj} è nullo quando $k > j$. Ma se $k \leq j < i$ sicuramente $k < i$, quindi almeno uno dei due termini u_{ik} e v_{kj} è zero, per cui il prodotto è zero.
Supponiamo ora che esista l'inversa U^{-1}, tale che $U^{-1}U = UU^{-1} = I$. Siano v_{ij} i coefficienti di U^{-1} Per ogni $j = 1, \dots, n-1$ si deve avere

$$(*) \qquad \sum_{k=1}^{n} u_{nk}v_{kj} = u_{nn}v_{nj} = 0 \text{ e } \sum_{k=1}^{n} u_{nk}v_{kn} = u_{nn}v_{nn} = 1,$$

dato che $u_{nk} = 0$ se $k < n$. Dunque $u_{nn} \neq 0$ e $v_{nn} = 1/u_{nn} \neq$

0, e per la (*) $v_{nj} = 0$ se $j = 1, \dots, n-1$. Ora, scriviamo U e V a blocchi

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1,n-1} & u_{1n} \\ 0 & u_{22} & \cdots & u_{2,n-1} & u_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & u_{n-1,n-1} & u_{n-1,n} \\ 0 & 0 & \cdots & 0 & u_{nn} \end{bmatrix}$$

$$U^{-1} = V = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1,n-1} & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2,n-1} & v_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ v_{n-1,1} & v_{n-1,2} & \cdots & v_{n-1,n-1} & v_{n-1,n} \\ 0 & 0 & \cdots & 0 & v_{nn} \end{bmatrix},$$

ossia

$$U = \begin{bmatrix} \boxed{\hat{U}} & \hat{\boldsymbol{u}} \\ {}^t\mathbf{0} & u_{nn} \end{bmatrix}, \quad U^{-1} = V = \begin{bmatrix} \boxed{\hat{V}} & \hat{\boldsymbol{v}} \\ {}^t\mathbf{0} & v_{nn} \end{bmatrix},$$

dove $\hat{U}$ e $\hat{V}$ sono le sottomatrici $n-1 \times n-1$ nel riquadro, $\hat{\boldsymbol{u}}$ e $\hat{\boldsymbol{v}}$ i vettori colonna di dimensione $n-1$, e ${}^t\mathbf{0}$ la riga zero di dimensione $n-1$. Non è difficile vedere che il prodotto può essere eseguito a blocchi, cioè

$$\begin{aligned} UV &= \begin{bmatrix} \hat{U} & \hat{\boldsymbol{u}} \\ {}^t\mathbf{0} & u_{nn} \end{bmatrix} \begin{bmatrix} \hat{V} & \hat{\boldsymbol{v}} \\ {}^t\mathbf{0} & v_{nn} \end{bmatrix} \\ &= \begin{bmatrix} \hat{U}\hat{V} + \hat{\boldsymbol{u}}{}^t\mathbf{0} & \hat{U}\hat{\boldsymbol{v}} + \hat{\boldsymbol{u}}v_{nn} \\ {}^t\mathbf{0}\hat{V} + u_{nn}{}^t\mathbf{0} & {}^t\mathbf{0}\hat{\boldsymbol{v}} + u_{nn}v_{nn} \end{bmatrix} = \begin{bmatrix} \hat{U}\hat{V} & \hat{U}\hat{\boldsymbol{v}} + \hat{\boldsymbol{u}}v_{nn} \\ {}^t\mathbf{0} & u_{nn}v_{nn} \end{bmatrix}. \end{aligned}$$

Dato che $UV = I_n$, $\hat{U}\hat{V}$ è uguale a I_{n-1}, e quindi $\hat{V}$ è l'inversa di $\hat{U}$ (che è triangolare superiore). Se supponiamo per induzione che $\hat{V}$ è triangolare superiore, segue subito che anche V è triangolare superiore.

51 Abbiamo visto nell'Esercizio **50** che l'inversa di una matrice triangolare superiore, se esiste, è triangolare superiore. Il prodotto a blocchi che abbiamo usato nella dimostrazione, se U è unitriangolare, si scrive

$$UV = \begin{bmatrix} \hat{U} & \hat{\boldsymbol{u}} \\ {}^t\mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \hat{V} & \hat{\boldsymbol{v}} \\ {}^t\mathbf{0} & v_{nn} \end{bmatrix} = \begin{bmatrix} \hat{U}\hat{V} & \hat{U}\hat{\boldsymbol{v}} + \hat{\boldsymbol{u}}v_{nn} \\ {}^t\mathbf{0} & v_{nn} \end{bmatrix},$$

da cui (ancora per induzione) si può dedurre che U è invertibile e $v_{ii} = 1$, cioè l'inversa è a sua volta unitriangolare.

52 Supponiamo che Q sia unitriangolare superiore e A una matrice. Allora, per definizione,

$$Q = \begin{bmatrix} 1 & q_{12} & q_{13} & \cdots & q_{1n} \\ 0 & 1 & q_{23} & \cdots & q_{2n} \\ 0 & 0 & 1 & \cdots & q_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

e se $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n$ indicano le colonne di A si ha

$$\begin{aligned} AQ = (&\boldsymbol{a}_1, q_{12}\boldsymbol{a}_1 + \boldsymbol{a}_2, q_{13}\boldsymbol{a}_1 + q_{23}\boldsymbol{a}_2 + \boldsymbol{a}_3, \dots \\ &\dots, q_{1n}\boldsymbol{a}_1 + q_{2n}\boldsymbol{a}_2 + \dots + \boldsymbol{a}_n). \end{aligned}$$

Ma allora, quando $q_{ij} \neq 0$, è possibile sommare il vettore $-q_{ij}\boldsymbol{a}_i$ alla j-esima colonna di AQ, ed eliminare il termine $q_{ij}\boldsymbol{a}_i$. Utilizziamo questo procedimento (e il fatto che il determinante non cambia se si somma un multiplo di una colonna ad un'altra colonna). Per i minori di ordine uno, è facile: le prime colonne di A e di AQ coincidono (sono uguali a $\boldsymbol{a}_1$), e quindi i minori principali di ordine 1 coincidono. Per i minori di ordine due, si elimina $q_{12}\boldsymbol{a}_1$ dalla seconda colonna, e quindi i minori principali di ordine 2 coincidono. Si procede eliminando $q_{13}\boldsymbol{a}_1 + q_{23}\boldsymbol{a}_2$ dalla terza colonna, e quindi i minori principali di ordine 3 coincidono. Si procede in questo modo, per arrivare all'eliminazione di tutti i $q_{ij}\boldsymbol{a}_i$, e ottenere che i determinanti $\det(A)$ e $\det(AQ)$ coincidono (ma questo era facile: $\det(A) = \det(A)\det(Q)$, e $\det(Q) = 1$). Abbiamo così mostrato che tutti i minori principali di A e di AQ coincidono, per ogni matrice A, se Q è unitriangolare.
Ora, i minori principali di una matrice coincidono con quelli della sua trasposta, e quindi i minori principali delle trasposte ${}^tB$ e ${}^tQ{}^tB$ coincidono. Se poniamo $B = {}^tA$, otteniamo che i minori principali di ${}^tQA$ coincidono con quelli di A. Ma allora moltiplicando ancora a destra per Q otteniamo una matrice ${}^tQAQ$ che, per quanto visto sopra, ha minori principali uguali a quelli di ${}^tQA$, che li ha uguali a A.

53 Sia $A = (a_{ij})$ la matrice simmetrica associata alla forma quadratica $q(\boldsymbol{x})$. Cominciamo a mostrare che $\det A > 0$. Per la Legge di Inerzia di Sylvester (IV.71), visto che una forma definita positiva ha indice di positività $i_p = n$, esiste un cambio di coordinate $\boldsymbol{x} = B\boldsymbol{y}$ tale che nelle $\boldsymbol{y}$ la forma è diagonale e tutti i coefficienti uguali a uno, cioè $y_1^2 + y_2^2 + \dots + y_n^2$. Ma allora

$${}^tBAB = I_n \implies \det({}^tBAB) = \det(I_n) = 1,$$

e dato che

$$\begin{aligned} 1 &= \det({}^tBAB) = \det{}^tB \det A \det B = \\ &= \det B \det A \det B = (\det B)^2 \det A, \end{aligned}$$

si ha che $\det A > 0$. Ora, i minori principali non sono altro che le matrici associate alle forme quadratiche definite sulle prime k coordinate ottenute annullando le restanti $n-k$ coordinate

$$q(x_1,0,\dots,0), q(x_1,x_2,0,\dots,0), \dots q(x_1,x_2,\dots,x_n-1,0),$$

che naturalmente sono definite positive. Dunque tutti i minori principali sono strettamente positivi.

54 Sia $A = a_{i,j}$ la matrice simmetrica associata alla forma quadratica $q(\boldsymbol{x})$. Vogliamo mostrare per induzione su n che se i minori principali sono positivi allora la forma quadratica definita da A è definita positiva. Se $n = 1$, allora $a_{11} > 0$ in quanto unico minore principale, e quindi $q(\boldsymbol{x}) = a_{11}x_1^2$ è definita positiva. Supponiamo, come ipotesi di induzione, che ogni matrice simmetrica di ordine k (con $k < n$) e minori principali tutti positivi produca una forma quadratica definita positiva. Allora procediamo come nella dimostrazione del Teorema (IV.70), osservando che per ipotesi $a_{11} > 0$ (è uno dei minori). L'eliminazione di tutti i termini x_1x_j con $j = 2,\dots,n$ viene ottenuta componendo sostituzioni del tipo $x_1 \mapsto x_1 + b_jx_j$ per opportuno b_j, e quindi possiamo considerare il cambio di coordinate risultante $x_1 = y_1 + b_2y_2 + \dots + b_ny_n$, con $b_j = -a_{1j}/a_{11}$. La matrice corrispondente (per cui $\boldsymbol{x} = B\boldsymbol{y}$) è

$$B = \begin{bmatrix} 1 & b_2 & b_3 & \cdots & b_n \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Ma questa è una matrice unitriangolare superiore! Allora per l'Esercizio **52** la matrice ${}^tBAB$ ha gli stessi minori principali di A, che quindi sono tutti positivi. Ma ${}^tBAB$ è la matrice ottenuta dopo il cambio di variabili, come nella dimostrazione del Teorema (IV.70), e quindi non ha termini diversi da zero nella prima riga e nella prima colonna, salvo $a_{11} > 0$,

$${}^tBAB = \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & \hat{a}_{22} & \hat{a}_{23} & \cdots & \hat{a}_{2n} \\ 0 & \hat{a}_{32} & \hat{a}_{33} & \cdots & \hat{a}_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \hat{a}_{n2} & \hat{a}_{n3} & \cdots & \hat{a}_{nn} \end{bmatrix}.$$

Ora, siano $\hat{d}_1$, $\hat{d}_2$, $\dots$, $\hat{d}_{n-1}$, i minori principali della matrice $\hat{A}$ di ordine $n-1$ indicata dal tratteggio. I minori principali di ${}^tBAB$ (che coincidono con quelli di A e sono tutti positivi per ipotesi) sono a_{11}, $a_{11}\hat{d}_1$, $a_{11}\hat{d}_2$, $\dots$, $a_{11}\hat{d}_{n-1}$. Ma allora, visto che $a_{11} > 0$, tutti i $\hat{d}_j$ sono positivi, e per l'ipotesi di induzione la forma quadratica $\hat{q}_{\hat{A}}(y_2,\dots,y_n)$ associata alla matrice $\hat{A}$ è definita positiva. È chiaro quindi che anche $q(\boldsymbol{x})$ è definita positiva, poiché nelle coordinate $\boldsymbol{y}$ si scrive come somma delle due forme definite positive

$$q(\boldsymbol{y}) = a_{11}y_1^2 + \hat{q}_{\hat{A}}(y_2,\dots,y_n).$$

55 Per la linearità di F sulla prima colonna, si ha

$$F\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{bmatrix} = a_{11}F\begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Se sottraiamo alla seconda colonna a_{12} per la prima, alla terza a_{13} per la prima, etc. etc. si ha per la linearità

$$F\begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{bmatrix} = F\begin{bmatrix} 1 & a_{12}-a_{12} & \cdots & a_{1n}-a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$= F\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Ma quest'ultima, come funzione della matrice evidenziata, è lineare nelle colonne, cambia di segno allo scambiare le colonne, e vale 1 sulla matrice identica: deve quindi essere il determinante, come volevasi dimostrare. La dimostrazione per le righe è esattamente la stessa, e la lasciamo al Lettore.

56 La matrice si ottiene cancellando la i-esima riga e la j-esima colonna, e la tesi si riscrive come $(-1)^{i+k}F(A) =$

$$\det\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,j-1} & 0 & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i1} & a_{i2} & \cdots & a_{i,j-1} & 1 & a_{i,j+1} & \cdots & a_{in} \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{nn} \end{bmatrix}$$

Dopo $i-1$ scambi di righe (e ogni volta si moltiplica per (-1)) e $j-1$ scambi di colonne (e ogni volta si moltiplica per (-1)), si arriva alla forma

$$\begin{bmatrix} 1 & * & \cdots & * \\ 0 & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n1} & \cdots & a_{nn} \end{bmatrix},$$

il cui valore di F per l'Esercizio **55** (con $a_{11} = 1$) è uguale al determinante, moltiplicato per $(-1)^{i-1}(-1)^{j-1} = (-1)^{i+j-2} = (-1)^{i+j}$.

57 **(a)** $\boldsymbol{a} \cdot \boldsymbol{b} = a_1\bar{b}_1 + a_2\bar{b}_2 = i + 1$; **(b)** $\boldsymbol{a} \cdot \boldsymbol{b} = 1 - i$; **(c)** $1 - i + 2 + i = 3$; **(d)** $2i$.

58 Per ①, ricordando le proprietà formali di un prodotto scalare hermitiano (pagina 215) si ha

$$\begin{aligned} \|\boldsymbol{a}+\boldsymbol{b}\|^2 &= (\boldsymbol{a}+\boldsymbol{b}) \cdot (\boldsymbol{a}+\boldsymbol{b}) = \boldsymbol{a}\cdot\boldsymbol{a} + \boldsymbol{a}\cdot\boldsymbol{b} \\ &+ \boldsymbol{b}\cdot\boldsymbol{a} + \boldsymbol{b}\cdot\boldsymbol{b} = \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2 + \boldsymbol{a}\cdot\boldsymbol{b} + \overline{\boldsymbol{a}\cdot\boldsymbol{b}}. \end{aligned}$$

Segue che

$$\begin{aligned} \|\boldsymbol{a}+\boldsymbol{b}\|^2 - \|\boldsymbol{a}-\boldsymbol{b}\|^2 &= \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2 + \boldsymbol{a}\cdot\boldsymbol{b} + \overline{\boldsymbol{a}\cdot\boldsymbol{b}} \\ &- \|\boldsymbol{a}\|^2 - \|\boldsymbol{b}\|^2 + \boldsymbol{a}\cdot\boldsymbol{b} + \overline{\boldsymbol{a}\cdot\boldsymbol{b}} = 2(\boldsymbol{a}\cdot\boldsymbol{b} + \overline{\boldsymbol{a}\cdot\boldsymbol{b}}) \end{aligned}$$

e anche

$$\begin{aligned} \|\boldsymbol{a}+\boldsymbol{b}\|^2 + \|\boldsymbol{a}-\boldsymbol{b}\|^2 &= \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2 + \boldsymbol{a}\cdot\boldsymbol{b} + \overline{\boldsymbol{a}\cdot\boldsymbol{b}} \\ &+ \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2 - \boldsymbol{a}\cdot\boldsymbol{b} - \overline{\boldsymbol{a}\cdot\boldsymbol{b}} = 2\|\boldsymbol{a}\|^2 + 2\|\boldsymbol{b}\|^2. \end{aligned}$$

Per l'ultima, se z è un numero complesso $z \in \mathbb{C}$, si ha $z + \bar{z} \in \mathbb{R}$, e quindi $\boldsymbol{a}\cdot\boldsymbol{b} + \boldsymbol{b}\cdot\boldsymbol{a} = \boldsymbol{a}\cdot\boldsymbol{b} + \overline{\boldsymbol{a}\cdot\boldsymbol{b}} \in \mathbb{R} \subset \mathbb{C}$.

59 Scriviamo $\boldsymbol{a} = \boldsymbol{x} + i\boldsymbol{y}$ e $\boldsymbol{b} = \boldsymbol{u} + i\boldsymbol{v}$, dove $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{u}$ e $\boldsymbol{v}$ sono vettori di $\mathbb{R}^n$. Allora

$$\boldsymbol{a}\cdot\boldsymbol{b} = (\boldsymbol{x}+i\boldsymbol{y})\cdot(\boldsymbol{u}+i\boldsymbol{v}) = \boldsymbol{x}\cdot\boldsymbol{u} + \boldsymbol{y}\cdot\boldsymbol{v} + i(-\boldsymbol{x}\cdot\boldsymbol{v} + \boldsymbol{y}\cdot\boldsymbol{u})$$

e quindi

$$\begin{aligned} \boldsymbol{a}\cdot\boldsymbol{b} + \boldsymbol{b}\cdot\boldsymbol{a} &= \boldsymbol{x}\cdot\boldsymbol{u} + \boldsymbol{y}\cdot\boldsymbol{v} + i(-\boldsymbol{x}\cdot\boldsymbol{v} + \boldsymbol{y}\cdot\boldsymbol{u}) \\ &+ \boldsymbol{u}\cdot\boldsymbol{x} + \boldsymbol{v}\cdot\boldsymbol{y} + i(-\boldsymbol{u}\cdot\boldsymbol{y} + \boldsymbol{v}\cdot\boldsymbol{x}) = 2(\boldsymbol{u}\cdot\boldsymbol{x} + \boldsymbol{v}\cdot\boldsymbol{y}). \end{aligned}$$

Ma anche, $\|\boldsymbol{a}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2$, $\|\boldsymbol{b}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2$. Allora se consideriamo i due vettori $\boldsymbol{X}$ e $\boldsymbol{Y}$ di dimensione $2n$ con componenti $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$ e $u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_n$, si ha

$$\|\boldsymbol{a}\|^2 = \|\boldsymbol{X}\|^2, \quad \|\boldsymbol{b}\|^2 = \|\boldsymbol{Y}\|^2, \quad \frac{\boldsymbol{a}\cdot\boldsymbol{b} + \boldsymbol{b}\cdot\boldsymbol{a}}{2} = \boldsymbol{X}\cdot\boldsymbol{Y},$$

e quindi la tesi segue dalla disuguaglianza di Cauchy–Schwarz.

60 La matrice P_1 ha autovalori 1 e -1, cui corrispondono gli autovettori $\boldsymbol{x}_1 = (1,1)$ e $\boldsymbol{x}_{-1} = (-1,1)$. Anche P_2 ha autovalori 1 e -1, con autovettori $\boldsymbol{x}_1 = (-i,1)$ e $\boldsymbol{x}_{-1} = (i,1)$. Così come P_3, i cui autovalori sono sempre 1 e -1, con autovettori $\boldsymbol{x}_1 = (1,0)$ e $\boldsymbol{x}_{-1} = (0,1)$.

❊ ❊ ❊

Tra le moltitudini di linguaggi adatti a fare qualche esempio di algoritmi per l'algebra lineare, abbiamo scelto il `python`,[23] *come per i pochi esempi del primo volume. Qui sotto c'è una possibile classe per rappresentare le matrici, e rendere il codice nelle soluzioni degli esercizi sufficientemente leggibile. Non si tratta assolutamente di programmi scritti bene, ma solo di esempi che abbiamo ritenuto di qualche valore didascalico. Soluzioni degne e analoghe possono essere prodotte in qualsiasi altro linguaggio di programmazione non troppo limitato.*

```
class Matrice:
 def __init__(self,data,uno=1.0):
  """Inizializzazione di una matrice, in cui le righe
vengono completate da zeri se necessario.
Il parametro opzionale uno=1.0 oppure uno=1
è per avere interi, oppure float, oppure...
Esempio: Matrice([[1,2],[1]])
"""
  if type(data) != type([]):
   data=[[data]]
  self.ismatrix=True
  self.data=data
  self.uno=uno
  self.num_righe=len(data)
  self.num_colonne=max([len(r) for r in data])
  for d in self.data:
   if len(d) < self.num_colonne:
    d += ([0] * (self.num_colonne-len(d)))

 def __repr__(self):
  res="[\n"+("\n".join([str(riga) for riga in self.data]))+"\n]"
  return str(res)

 def __str__(self):
  return  self.__repr__()

 def __getitem__(self,key):
   "per avere A[i,j]"
   return self.data[key[0]][key[1]]

 def __setitem__(self,key,value):
   "quando A[i,j] = 2"
   self.data[key[0]][key[1]]=value

 def colonna(self,j):
  return [ self[i,j] for i in range(self.num_righe) ]

 def riga(self,i):
  return [ self[i,j] for j in range(self.num_colonne) ]

 def sottomatrice(self,ii,jj):
  """sottomatrice (ii,jj)-esimo"""
  return Matrice ( [ [ self.data[i][j]
    for j in range(self.num_colonne) if j != jj
  ] for i in range(self.num_righe) if i != ii ])

 def __or__(self,other):
  """ A|B : matrice a blocchi aumentata """
  if self.num_righe != other.num_righe:
   raise Exception ("numero righe non coincide")
  return Matrice (
  [self.data[i] + other.data[i] for i in range(self.num_righe)])

 def __div__(self,other):
  """ A/B: matrice a blocchi uno sopra l'altro"""
  if self.num_colonne != other.num_colonne:
   raise Exception("numero di colonne non coincide")
  return Matrice (self.data + other.data)

 def t(self):
  """trasposta"""
  return Matrice ( [ [ self.data[i][j]
```

```
          for i in range(self.num_righe)
        ] for j in range(self.num_colonne) ])

 def __add__(self,other):
  return Matrice ( [ [
        (self.data[i][j] + other.data[i][j])
        for j in range (self.num_colonne)
      ] for i in range(self.num_righe) ])

 def __sub__(self,other):
  """A-B"""
  return Matrice ([ [ (self.data[i][j] - other.data[i][j])
       for j in range (self.num_colonne)
                     ] for i in range(self.num_righe) ])

 def __mul__(self,other):
  """ self * other : prodotto di matrici, righe per colonne """
  return Matrice (
                 [ [ sum ([self.data[i][k] * other.data[k][j]
     for k in range(self.num_colonne)
          ])  for j in range(other.num_colonne)
                  ] for i in range(self.num_righe) ])

 def __rmul__(self,other):
  """numero * matrice """
  return Matrice(
                [ [ self.data[i][j] * other
    for j in range(self.num_colonne)
  ] for i in range(self.num_righe) ])

 def __neg__(self):
  """- matrice """
  return Matrice(
              [ [ - self.data[i][j]
   for j in range(self.num_colonne)
 ] for i in range(self.num_righe) ])


 def inversa(self):
  """inversa"""
  if self.num_colonne==1:
   return self.uno / self[1,1]
  else:
   return self.uno/det(self) * \
    trasposta(
     Matrice( [ [(-1)**(i+j) * det(self.sottomatrice(i,j))
         for j in range(self.num_colonne)
       ] for i in range(self.num_righe) ] ))


 def __pow__(self,other):
  """ A ** n potenza. """
  result= MatriceIdentica(self.num_righe)
  if other >= 0:
   for i in range(other):
    result = self * result
   return result
  else:
   invself=self.inversa()
   for i in range(-other):
    result = invself  * result
   return result

 def __xor__(self,other):
    """ A ^ n : potenza (sinonimo di A^n ) """
    print "attenzione alle precedenze su ^!"
    return self.__pow__(other)


def MatriceIdentica(n):
 x=MatriceNulla(n,n)
 for i in range(x.num_colonne):
  x.data[i][i] = 1
 return x

def MatriceNulla(n,m):
 return Matrice([[0 for j in range(m)] for i in range(n)])

def MatriceDiagonale(l):
 le=len(l)
 x=MatriceNulla(le,le)
 for i in range(le):
  x.data[i][i] = l[i]
 return x

def trasposta(m):
  return m.t()

def det(x):
  """determinante di x"""
  if x.num_colonne!=x.num_righe:
   raise Exception("Matrice non quadrata")
  if x.num_colonne==1: return x.data[0][0]
  return sum ([(-1)**i * x.data[0][i]*det(x.sottomatrice(0,i))
              for i in range(x.num_colonne)  ])

def Estendi(m):
 return ( Matrice([ [1] ]) | MatriceNulla(1,m.num_colonne)
        ) / ( MatriceNulla(m.num_righe,1) | m )
```

61

```
from matrice import *

def MatricePermutazione( nupla ):
 n=len(nupla)
 m=MatriceNulla(n,n)
 m.uno=1
 for i in range(n): m[i,nupla[i]] = 1
 return m
```

62 La funzione che riportiamo è ricorsiva, così come il tipo di eliminazione di Gauss che abbiamo introdotto:

```
from matrice import *
from permutazione import *

# def Estendi(m):
#  return ( Matrice([ [1] ]) | MatriceNulla(1,m.num_colonne)
#        ) / ( MatriceNulla(m.num_righe,1) | m )

def LUdec(m):
 n=m.num_righe
 if n==1:
  return MatriceIdentica(1), MatriceIdentica(1), m
 else:
  slice_prima_colonna=[x**2 for x in m.colonna(0)[1:] ]
  max_entry=max(slice_prima_colonna)
  if max_entry >0:
   posizione_pivot=slice_prima_colonna.index(max_entry )+1
   perm=range(n)
   perm[0]=posizione_pivot
   perm[posizione_pivot]=0
   P=MatricePermutazione(perm)
   Pm = P * m
   pivot=Pm[0,0]
   L=MatriceIdentica(n)
   for i in range(1,n): L[i,0] =  Pm[i,0] * m.uno / pivot
  else:
   L=MatriceIdentica(n)
   P=MatricePermutazione(range(n))
   Pm = m
  U = L**(-1) * Pm
  AA=U.sottomatrice(0,0)
  PP,LL,UU = LUdec(AA)
  PP= Estendi(PP)
  LL= Estendi(LL)
  U=  Matrice( [Pm.riga(0)]) / (
    MatriceNulla(n-1,1) | UU )
  P=PP * P
  L=PP *  L * (PP **(-1) ) *   LL
  return P,L,U
```

63 Abbiamo implementato le formule di pagina 186, come indicato nella nota a margine. Per prima cosa ci serve una piccola classe per le operazioni sui vettori (`vettore.py`).

```
from math import sqrt

class Vettore:
 def __init__(self,data):
  if type(data) != type([]): data=[[data]]
  self.isvector=True
  self.data=data
  self.dim=len(data)
 def __getitem__(self,key):
  return self.data[key]
 def __setitem__(self,key,value):
  self.data[key]=value

 def  __repr__(self):
  return str(self)
 def __str__(self):
  return str(self.data)

 def __add__(self,other):
  return Vettore([ (self.data[i] + other.data[i])
   for i in range(self.dim) ])
 def __sub__(self,other):
  return Vettore([ (self.data[i] - other.data[i])
   for i in range(self.dim) ])
 def __rmul__(self,other):
  return Vettore([ (self.data[i] * other)
   for i in range(self.dim) ])
 def __neg__(self):
  return Vettore([(- self.data[i]) for i in range(self.dim)] )
 def __len__(self):
  return self.dim


def dotprod(u,v):
 return sum ( [ u[i] * v[i] for i in range(u.dim) ])

def norma(u):
 return sqrt(dotprod(u,u))

def normalizzato(u):
 n=norma(u)
 return  (1.0 / n ) * u
```

Ed ecco l'ortogonalizzazione di Gram–Schmidt e la decomposizione QR.

```
from vettore import *
from matrice import *

def GramSchmidt(kupla):
 n=kupla[0].dim
 q=[ normalizzato( kupla[0] ) ]
 R=MatriceIdentica(len(kupla) )
 R[0,0]=norma(kupla[0])
 for i in range(1,len(kupla)):
  u = kupla[i]
  for j in range(i):
   u = u - dotprod( kupla[i], q[j]) * q[j]
  q += [ normalizzato (u) ]
  R[i,i]= norma(u)
 for i in range(len(kupla) ):
  for j in range(i+1,len(kupla)):
   R[i,j] = dotprod(kupla[j],q[i])
 return q,R

def QRdec(A):
 Qs,R=GramSchmidt([ Vettore(A.colonna(i))  \
  for i in range(A.num_colonne) ] )
 return Matrice([Qs[i].data  \
  for i in range(len(Qs)) ]).t(), R
```

64 Una possibile implementazione è la seguente (attenzione alla nota (IV.73): se gli elementi sulla diagonale sono nulli l'algoritmo non funziona).

```
from matrice import *

def tBAB(A):
 n=A.num_righe
 if n==1: return Matrice([[1]])
 B= MatriceIdentica(n)
 for j in range(1,n):
  B[0,j] = -A[0,j] * A.uno / A[0,0]
 Ahat=(B.t() * A * B ).sottomatrice(0,0)
 Bhat = tBAB( Ahat )
 return B * Estendi(Bhat)
```

65 La matrice ${}^t\!BB$ è definita positiva perché ${}^t\boldsymbol{x}{}^t\!BB\boldsymbol{x}$ non è altro che la norma al quadrato $\|B\boldsymbol{x}\|^2$, che è > 0 se e solo se $\boldsymbol{x} \neq \boldsymbol{0}$. Sia P la matrice unitriangolare superiore tale che $D = {}^t\!PAP$ è diagonale. La matrice D ha coefficienti tutti positivi sulla diagonale, e quindi esiste una matrice $\sqrt{D}$ tale che $\sqrt{D}\sqrt{D} = D$ (basta prendere le radici quadrate degli elementi sulla diagonale di D). Ma allora se si pone $U = \sqrt{D}P^{-1}$ si ha

$$A = {}^t\!P^{-1}DP^{-1} = {}^t\!P^{-1}\sqrt{D}\sqrt{D}P^{-1} = {}^t\!P^{-1}{}^t\!\sqrt{D}\sqrt{D}P^{-1} = {}^t\!UU.$$

Dato che P è unitriangolare superiore, anche P^{-1} lo è. Un piccolo programmino che esegue questi semplici calcoli è il seguente.

```
from matrice import *
from lagrange import *
from math import sqrt

def cholesky(A):
 n=A.num_righe
 P=tBAB(A)
 D=P.t() * A * P
 for i in range(n):
  D[i,i] = sqrt(D[i,i])
 return (D * (P^(-1)))
```

66 Sia $A = QR$ la decomposizione QR. Allora se poniamo $\boldsymbol{y} = R\boldsymbol{x}$, si ha $A\boldsymbol{x} = Q\boldsymbol{y}$ e la formula (8) del Teorema (IV.45) dice che $\mathrm{proj}_W \boldsymbol{b} = \sum_{i=1}^{n}(\boldsymbol{b}\cdot\boldsymbol{q}_i)\boldsymbol{q}_i$, dove i $\boldsymbol{q}_i$ sono le colonne di Q, e dunque $y_i = \boldsymbol{b}\cdot\boldsymbol{q}_i$ per ogni i. In forma matriciale si scrive $\boldsymbol{y} = {}^t\!Q\boldsymbol{b} \implies \boldsymbol{x} = R^{-1}{}^t\!Q\boldsymbol{b}$. Ecco la brevissima implementazione.

```
from matrice import *
from QRdec import *

def minimi_quadrati(A,b):
 """A Matrice, b n-upla """
 Q,R=QRdec(A)
 return (R^(-1)) * Q.t() * Matrice([b]).t()
```

Riferimenti bibliografici

[Abate(1996)] M. Abate. *Geometria*. McGraw–Hill, Milano, 1996.

[Adams(2003)] R. A. Adams. *Calcolo Differenziale, vol.1 e vol.2*. CEA, Milano, 2003.

[Aigner e Ziegler(2004)] M. Aigner e G. M. Ziegler. *Proofs from The Book*. Springer-Verlag, Berlin, 2004.

[Ambrosetti e Prodi(1995)] A. Ambrosetti e G. Prodi. *A primer of nonlinear analysis*, volume 34. Cambridge University Press, Cambridge, 1995.

[Andronov et al.(1973)] A. A. Andronov, E. A. Leontovich, I. I. Gordon, e A. G. Maĭer. *Qualitative theory of second-order dynamic systems*. Halsted Press, 1973.

[Apostol(1967)] T. M. Apostol. *Calculus. Vol. I: One-variable calculus, with an introduction to linear algebra*. Blaisdell Publishing Co., 1967.

[Apostol(1969)] ———. *Calculus. Vol. II: Multi-variable calculus and linear algebra, with applications to differential equations and probability*. Blaisdell Publishing Co., 1969.

[Apostol(1977)] ———. *Calcolo – volume secondo, geometria*. Bollati Boringhieri, Torino, 1977.

[Apostol(1978)] ———. *Calcolo – volume terzo, Analisi 2*. Bollati Boringhieri, Torino, 1978.

[Arnold(2006)] V. I. Arnold. *Ordinary differential equations*. Springer-Verlag, Berlin, 2006.

[Axler(1997)] S. Axler. *Linear algebra done right*. Springer, Paris, 1997.

[Banach(1922)] S. Banach. Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales,. *Fund. Math.*, 3:133–181, 1922.

[Berger(1987)] M. Berger. *Geometry I, II*. Springer, Berlin–Heidelberg–New York, 1987.

[Birkhoff e Rota(1989)] G. Birkhoff e G.-C. Rota. *Ordinary differential equations*. John Wiley & Sons Inc., New York, 1989.

[Boyce e DiPrima(2001)] W. E. Boyce e R. C. DiPrima. *Elementary differential equations and boundary value problems*. John Wiley & Sons Inc., New York, 2001.

[Boyer(2004)] C. B. Boyer. *History of Analytic Geometry*. Dover, New York, 2004.

[Bramanti et al.(2000)] M. Bramanti, C. Pagani, e S. Salsa. *Matematica. Calcolo infinitesimale e algebra lineare*. Zanichelli, Bologna, 2000.

[Caccioppoli(1930)] R. Caccioppoli. Un teorema generale sull esistenza di elementi uniti in una trasformazione funzionale. *Rend. Accad. dei Lincei*, 11:794–799, 1930.

[Carrier e Pearson(1991)] G. F. Carrier e C. E. Pearson. *Ordinary differential equations*, volume 6 di *Classics in Applied Mathematics*. SIAM, 1991.

[Coddington(1961)] E. A. Coddington. *An introduction to ordinary differential equations*. Prentice-Hall Inc., 1961.

[Conti(1993)] F. Conti. *Calcolo. Teoria e applicazioni*. McGraw–Hill, Milano, 1993.

[Courant e Hilbert(1953)] R. Courant e D. Hilbert. *Methods of mathematical physics. Vol. I*. Interscience Publishers, Inc., New York, N.Y., 1953.

[Crowe(1967)] M. J. Crowe. *History of Vector Analysis*. University of Notre Dame Press, London, 1967.

[De Marco(1992)] G. De Marco. *Analisi due: secondo corso di analisi matematica per l'università*. Zanichelli (Decibel), Bologna Padova, 1992.

[Devaney(2003)] R. L. Devaney. *An introduction to chaotic dynamical systems*. Westview Press, Boulder, CO, 2003.

[Dolcher(1978)] M. Dolcher. *Algebra lineare*. Zanichelli, Bologna, 1978.

[Fox(1987)] C. Fox. *An introduction to the calculus of variations*. Dover Publications Inc., New York, 1987.

[Guggenheimer(1967)] H. Guggenheimer. *Plane Geometry and its Groups*. Holden Day, San Francisco, 1967.

[Hilbert e Cohn-Vossen(1972)] D. Hilbert e S. Cohn-Vossen. *Geometria intuitiva*. Bollati Boringhieri, Torino, 1972.

[Hirsch e Smale(1974)] M. W. Hirsch e S. Smale. *Differential equations, dynamical systems, and linear algebra*. Academic Press, 1974.

[Hirsch et al.(2004)] M. W. Hirsch, S. Smale, e R. L. Devaney. *Differential equations, dynamical systems, and an introduction to chaos*, volume 60. Elsevier/Academic Press, Amsterdam, 2004.

[Klein(1956)] F. Klein. *Lectures on the Icosahedron*. Dover, New York, 1956.

[Kolmogorov e Fomīn(1975)] A. N. Kolmogorov e S. V. Fomīn. *Introductory real analysis*. Dover Publications Inc., New York, 1975. Translated from the second Russian edition and edited by Richard A. Silverman, Corrected reprinting.

[Lacroix(1797)] S. F. Lacroix. *Traité du Calcul Différentiel et du Calcul Intégral*. Paris, 1797.

[Lang(1970)] S. Lang. *Algebra lineare*. Boringhieri, Torino, 1970. Ottava edizione.

[Lyndon(1985)] R. Lyndon. *Groups and Geometry*. Cambridge University Press, Cambridge, 1985.

[Maderna e Soardi(1987)] C. Maderna e P. M. Soardi. *Lezioni di analisi matematica*. CLUP, Milano, 1987.

[Marcellini e Sbordone(2004)] P. Marcellini e C. Sbordone. *Elementi di calcolo : [versione semplificata per i nuovi corsi di laurea]*. Liguori, Napoli, 2004.

[Marsden e Tromba(1976)] J. E. Marsden e A. J. Tromba. *Vector Calculus*. W.H. Freeman and Company, New York, 1976.

[Messer(1994)] R. Messer. *Linear algebra, Gateway to Mathematics*. Harper Collins, New York, 1994.

[Pagani e Salsa(1990)] C. D. Pagani e S. Salsa. *Analisi matematica - Vol I*. Zanichelli, Bologna, 1990.

[Pagani e Salsa(1991)] ———. *Analisi matematica - Vol II*. Zanichelli, Bologna, 1991.

[Peano(1888)] G. Peano. *Calcolo Geometrico, secondo l'Ausdehnungslehre di H. Grassmann, preceduto dalle Operazioni della Logica Deduttiva*. F.lli Bocca Editori, Torino, 1888.

[Piccinini et al.(1978)] L. Piccinini, G. Stampacchia, e G. Vidossich. *Equazioni differenziali ordinarie in $\mathbb{R}^n$*. Liguori Editore, Napoli, 1978.

[Pontryagin(1962)] L. S. Pontryagin. *Ordinary differential equations*. Addison-Wesley, 1962.

[Riccardi(2004)] G. Riccardi. *Calcolo differenziale ed integrale*. Springer Italia, Milano, 2004.

[Robbiano(2007)] L. Robbiano. *Algebra lineare per tutti*. Springer, Milano, 2007.

[Royden(1988)] H. L. Royden. *Real analysis*. Macmillan Publishing Company, New York, 1988.

[Rudin(1976)] W. Rudin. *Principles of mathematical analysis*. McGraw-Hill Book Co., New York, 1976. International Series in Pure and Applied Mathematics.

[Rudin(1987)] ———. *Real and complex analysis*. McGraw-Hill Book Co., New York, 1987.

[Salsa e Squellati(2004)] S. Salsa e A. Squellati. *Calcolo infinitesimale e algebra lineare*. Zanichelli, Bologna, 2004.

[Sernesi(1991,1994,2000)] E. Sernesi. *Geometria, vol. I–II*. Bollati Boringhieri, Torino, 1991,1994,2000.

[Shahriari(2006)] S. Shahriari. *Approximately calculus*. American Mathematical Society, Providence, RI, 2006.

[Simmons(1972)] G. F. Simmons. *Differential equations with applications and historical notes*. McGraw-Hill, New York, 1972.

[Stewart(2002)] J. Stewart. *Calcolo – Funzioni di più variabili*. Apogeo, Milano, 2002.

[Tricomi(1965)] F. G. Tricomi. *Lezioni di analisi matematica*. CEDAM, Padova, 1965.

[Weyl(1952)] H. Weyl. *Symmetry*. Princeton University Press, Princeton, 1952.

APOGEO SPICCHI

M. Conti, D. L. Ferrario, S. Terracini, G. Verzini, V. Barutello

ANALISI MATEMATICA

VOL. 2 con elementi di geometria e calcolo vettoriale

Gli altri capitoli