

Geometria e algebra lineare

Bruno Martelli

Dipartimento di Matematica, Largo Pontecorvo 5, 56127 Pisa, Italy

E-mail address: martelli at dm dot unipi dot it

versione: 22 agosto 2018

Indice

Introduzione	1
Capitolo 1. Nozioni preliminari	3
1.1. Gli insiemi	3
1.2. Funzioni	11
1.3. Polinomi	18
1.4. Numeri complessi	22
1.5. Strutture algebriche	30
Esercizi	33
Complementi	34
Capitolo 2. Spazi vettoriali	37
2.1. Lo spazio euclideo	37
2.2. Spazi vettoriali	40
2.3. Dimensione	53
Esercizi	69
Complementi	71
Capitolo 3. Sistemi lineari	73
3.1. Algoritmi di risoluzione	73
3.2. Rango	78
3.3. Determinante	85
3.4. Algebra delle matrici	93
Esercizi	99
Complementi	101
Capitolo 4. Applicazioni lineari	103
4.1. Introduzione	103
4.2. Nucleo e immagine	110
4.3. Matrice associata	117
4.4. Endomorfismi	124
Esercizi	132
Complementi	134
Capitolo 5. Autovettori e autovalori	135
5.1. Definizioni	135

5.2. Teorema di diagonalizzabilità	147
Esercizi	153
Complementi	155
Capitolo 6. Forma di Jordan	157
6.1. Forma di Jordan	157
6.2. Teorema di Cayley – Hamilton	169
6.3. Polinomio minimo	173
Esercizi	177
Complementi	178
Capitolo 7. Prodotti scalari	179
7.1. Introduzione	179
7.2. Matrice associata	188
7.3. Sottospazio ortogonale	191
7.4. Classificazione dei prodotti scalari	195
7.5. Prodotti scalari definiti positivi	201
7.6. Isometrie	213
Capitolo 8. Lo spazio euclideo	223
8.1. Prodotto vettoriale	223
8.2. Sottospazi affini	227
8.3. Trasformazioni del piano e dello spazio	237
Capitolo 9. Teorema spettrale	245
9.1. Prodotti hermitiani	245
9.2. Endomorfismi autoaggiunti	246
9.3. Il teorema	248
Capitolo 10. Geometria proiettiva	251
10.1. Introduzione	251
10.2. Sottospazi e proiettività	254
Capitolo 11. Quadriche	259
11.1. Introduzione	259
11.2. Classificazione proiettiva	260
11.3. Classificazione affine	264
11.4. Classificazione metrica delle quadriche non degeneri	268
Indice analitico	271

Introduzione

Il testo è rilasciato con la licenza *Creative Commons-BY-NC-SA*¹. Le figure sono tutte di pubblico dominio, eccetto le seguenti, che hanno una licenza *CC-BY-SA*² e sono state scaricate da Wikimedia Commons:

- Figura 1.1 (partizione di un insieme), creata da Wshun;
- Figura 1.4 (numeri complessi), creata da Wolfkeeper;
- Figura 1.5 (numeri complessi), creata da Kmhkmh;
- Figura 1.9 creata da Cronholm144;
- Figura 3.5 (prodotto riga per colonna), creata da Bilou.
- Figura 4.3 (cubi distorti) creata da Irrons
- Figura 8.1 (regola della mano destra), creata da Acdx;
- Figure 11.1 e 11.2 (quadriche), create da Sam Derbyshire.

¹È possibile distribuire, modificare, creare opere derivate dall'originale, ma non a scopi commerciali, a condizione che venga riconosciuta la paternità dell'opera all'autore e che alla nuova opera vengano attribuite le stesse licenze dell'originale (quindi ad ogni derivato non sarà permesso l'uso commerciale).

²Come la licenza precedente, ma con anche la possibilità di uso commerciale.

CAPITOLO 1

Nozioni preliminari

In questo primo capitolo introduciamo le nozioni di base che saranno usate in tutto il libro. Supponiamo che il lettore abbia già dimestichezza con la nozione di insieme e più in generale con l'algebra che viene insegnata nelle scuole superiori.

1.1. Gli insiemi

La teoria degli insiemi è alla base di tutta la matematica moderna.

1.1.1. Gli insiemi numerici. Un insieme è generalmente indicato con due parentesi graffe $\{\}$, all'interno delle quali sono descritti tutti i suoi elementi. A volte si indicano solo alcuni elementi e si usano dei puntolini \dots per indicare il resto. Ad esempio, questo è l'insieme dei *numeri naturali*:

$$\mathbb{N} = \{0, 1, 2, 3, \dots\}$$

L'insieme \mathbb{N} contiene infiniti elementi. Se aggiungiamo i numeri negativi otteniamo l'insieme dei *numeri interi*

$$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}.$$

Se oltre ai numeri interi consideriamo tutti i numeri esprimibili come frazioni $\frac{a}{b}$, otteniamo l'insieme dei *numeri razionali*

$$\mathbb{Q} = \left\{ \dots, -\frac{2}{5}, \dots, \frac{1}{8}, \dots, \frac{9}{4}, \dots \right\}.$$

Sappiamo inoltre che in aritmetica esiste un insieme ancora più grande, chiamato \mathbb{R} , formato da tutti i *numeri reali*.

1.1.2. Dimostrazione per assurdo. La costruzione rigorosa dell'insieme \mathbb{R} dei numeri reali è complicata e ci porterebbe troppo lontano. Possiamo però intanto convincerci del fatto che molti numeri reali che ci sono familiari non sono razionali. Quella che segue è la prima proposizione del libro ed è illuminante perché contiene un esempio di *dimostrazione per assurdo*.

Proposizione 1.1.1. *Il numero $\sqrt{2}$ non è razionale.*

Dimostrazione. Supponiamo per assurdo che $\sqrt{2}$ sia razionale. Allora $\sqrt{2} = \frac{a}{b}$, dove $\frac{a}{b}$ è una frazione. Elevando al quadrato e moltiplicando per b^2 entrambi i membri otteniamo

$$a^2 = 2b^2.$$

Questa uguaglianza però non è possibile: il doppio di un quadrato non è mai un quadrato (esercizio: usare la decomposizione in fattori primi). \square

In una dimostrazione per assurdo, si nega la tesi e si dimostra che questo porta ad un assurdo. Ne deduciamo che la tesi non può essere falsa, e quindi è vera per esclusione.

1.1.3. Sottoinsiemi. Se x è un elemento dell'insieme A , scriviamo

$$x \in A$$

per dire che x appartiene ad A . Se questo non accade, scriviamo $x \notin A$. Ad esempio

$$2 \in \mathbb{Z}, \quad \sqrt{2} \notin \mathbb{Q}.$$

Se B è un sottoinsieme di A , scriviamo

$$B \subset A.$$

In questo caso tutti gli elementi di B sono anche elementi di A . Ad esempio

$$\{1, 3, 5\} \subset \{-1, 1, 3, 4, 5\}.$$

Se esiste almeno un elemento di A che non è contenuto in B , allora diciamo che A contiene B *strettamente* e possiamo usare il simbolo

$$B \subsetneq A.$$

Ad esempio gli insiemi numerici descritti precedentemente sono ciascuno contenuto strettamente nel successivo:

$$\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}.$$

D'altro canto, se valgono entrambi i contenimenti

$$A \subset B, \quad B \subset A$$

allora chiaramente $A = B$, cioè A e B sono esattamente lo stesso insieme di elementi. Questo fatto che per adesso sembra un po' inutilmente tautologico verrà usato spesso nelle dimostrazioni presenti in questi libro.

Il simbolo \emptyset indica l'*insieme vuoto*, cioè l'insieme che non contiene nessun elemento.

1.1.4. Unione, intersezione e differenza. Dati due insiemi A e B , possiamo considerare la loro *intersezione* e la loro *unione*, indicate rispettivamente:

$$A \cap B, \quad A \cup B.$$

L'intersezione consiste di tutti gli elementi che stanno *sia* in A che in B , mentre l'unione consiste di quelli che stanno in A *oppure* in B . Ad esempio, se $A = \{1, 2, 3\}$ e $B = \{3, 5\}$, allora

$$A \cap B = \{3\}, \quad A \cup B = \{1, 2, 3, 5\}.$$

La *differenza* $A \setminus B$ è l'insieme formato da tutti gli elementi di A che non stanno in B . Se B è un sottoinsieme di A , la differenza $A \setminus B$ è anche chiamata il *complementare* di B in A ed è indicata con B^c .

Esercizio 1.1.2 (Leggi di De Morgan). Siano A e B due sottoinsiemi di un insieme C . Dimostrare le seguenti uguaglianze:

$$(A \cap B)^c = A^c \cup B^c$$

$$(A \cup B)^c = A^c \cap B^c$$

Qui X^c indica il complementare di X in C , cioè $X^c = C \setminus X$.

1.1.5. Notazione insiemistica. Ricordiamo che la notazione usata in matematica per definire un insieme C è del tipo:

$$C = \{ \text{tutti quegli elementi } x \text{ tali che vale una certa proprietà } P \}.$$

Tutto ciò in linguaggio matematico si scrive più brevemente così:

$$C = \{ x \mid P \}.$$

La barra verticale $|$ è sinonimo di "tale che". Ad esempio, intersezione, unione e complemento sono definiti formalmente così:

$$A \cap B = \{x \mid x \in A \text{ e } x \in B\},$$

$$A \cup B = \{x \mid x \in A \text{ oppure } x \in B\},$$

$$A \setminus B = \{x \mid x \in A \text{ e } x \notin B\}$$

Si possono anche usare i simboli \wedge e \vee come sinonimi di *e* ed *oppure*, ed i due punti $:$ al posto della barra $|$. Ad esempio, la prima legge di De Morgan (Esercizio 1.1.2) può essere dimostrata notando che entrambi gli insiemi $(A \cap B)^c$ e $A^c \cup B^c$ possono essere descritti nel modo seguente:

$$\{x \in C \mid x \notin A \text{ oppure } x \notin B\}.$$

Questa notazione per gli insiemi è molto flessibile. Ad esempio l'insieme P dei numeri pari può essere descritto nel modo seguente:

$$P = \{2n \mid n \in \mathbb{N}\}.$$

Otteniamo ovviamente

$$P = \{0, 2, 4, \dots\}.$$

Analogamente, l'insieme dei numeri dispari è

$$D = \{2n + 1 \mid n \in \mathbb{N}\} = \{1, 3, 5, \dots\}.$$

Infine, l'insieme dei quadrati è

$$Q = \{n^2 \mid n \in \mathbb{N}\} = \{0, 1, 4, 9, 16, \dots\}.$$

Un numero pari generico si scrive come $2n$, mentre un numero dispari generico si scrive come $2n + 1$. Usando questa notazione si possono dimostrare agevolmente dei teoremi, ad esempio questo:

Proposizione 1.1.3. Ogni numero dispari è differenza di due quadrati.

Dimostrazione. Un generico numero dispari si scrive come $2n + 1$, per qualche $n \in \mathbb{N}$. Questo è effettivamente la differenza di due quadrati successivi:

$$2n + 1 = (n + 1)^2 - n^2.$$

La dimostrazione è conclusa. □

1.1.6. Quantificatori. Due simboli che giocano un ruolo fondamentale in matematica sono i *quantificatori*:

$$\forall \quad \exists$$

I simboli sono una A ed una E rovesciate ed indicano le espressioni *per ogni* (*for All*) e *esiste* (*Exists*). I quantificatori sono essenziali nella formulazione dei teoremi, e più in generale di affermazioni matematiche che possono essere vere o false. Ad esempio, l'espressione

$$\forall x \in \mathbb{R} \exists y \in \mathbb{R} : 2y = x$$

dice che qualsiasi numero reale x può essere diviso per due, ed è un'affermazione vera. I due punti ":" sono sinonimo di "tale che". D'altro canto, la stessa espressione con \mathbb{Z} al posto di \mathbb{R} :

$$\forall x \in \mathbb{Z} \exists y \in \mathbb{Z} : 2y = x$$

è falsa perché se $x = 1$ non esiste nessun $y \in \mathbb{Z}$ tale che $2y = 1$.

Il simbolo $\exists!$ indica l'espressione *esiste ed è unico*. L'affermazione seguente

$$\forall x \in \mathbb{R} \exists! y \in \mathbb{R} : 2y = x$$

continua ad essere vera: ogni numero reale è il doppio di un unico numero reale. Invece l'affermazione

$$\forall x \in \mathbb{R} : x > 0 \exists! y \in \mathbb{R} : y^2 = x$$

è falsa: ogni numero reale positivo x ha effettivamente una radice quadrata reale y , però questa non è unica perché le radici di x sono sempre due $\pm y$.

1.1.7. Dimostrazioni. Come facciamo a capire se una data affermazione matematica sia vera o falsa? Non c'è una regola generale, ma ci sono alcune indicazioni importanti: una affermazione del tipo

Per ogni x in un insieme A vale una certa proprietà P

può essere vera o falsa. Per dimostrare che è vera, serve una *dimostrazione* che mostri che la proprietà P è verificata da tutti gli elementi x nell'insieme A . Per dimostrare che è falsa, invece è sufficiente esibire un singolo $x \in A$ per cui P non sia soddisfatta.

Ad esempio, l'affermazione

$$\forall x \in \mathbb{N}, x^2 \geq 0$$

è vera, perché i quadrati sono sempre positivi. L'affermazione

$$\forall x \in \mathbb{N}, x^2 \geq 5$$

non è vera, perché non è soddisfatta ad esempio dal valore $x = 1$.

1.1.8. Prodotto cartesiano. Il *prodotto cartesiano* di due insiemi A e B è un nuovo insieme $A \times B$ i cui elementi sono tutte le coppie (a, b) dove a è un elemento qualsiasi di A e b è un elemento qualsiasi di B . Più brevemente:

$$A \times B = \{(a, b) \mid a \in A, b \in B\}.$$

Ad esempio, se $A = \{1, 2\}$ e $B = \{-1, 3, 4\}$, allora

$$A \times B = \{(1, -1), (1, 3), (1, 4), (2, -1), (2, 3), (2, 4)\}.$$

Più in generale, se A e B sono insiemi finiti con m e n elementi rispettivamente, allora il prodotto cartesiano $A \times B$ contiene mn elementi.

Il prodotto $A \times A$ è anche indicato con A^2 . Il caso $A = \mathbb{R}$ è particolarmente interessante perché ha una forte valenza geometrica. L'insieme \mathbb{R} dei numeri reali può essere interpretato come una retta. Il prodotto cartesiano $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ è l'insieme

$$\mathbb{R}^2 = \{(x, y) \mid x \in \mathbb{R}, y \in \mathbb{R}\}$$

formato da tutte le coppie (x, y) di numeri reali. In altre parole, \mathbb{R}^2 non è nient'altro che il *piano cartesiano* già studiato alle superiori: un elemento di \mathbb{R}^2 è un punto identificato dalla coppia (x, y) .

Possiamo definire in modo analogo il prodotto di un numero arbitrario di insiemi. Il prodotto cartesiano di k insiemi A_1, \dots, A_k è l'insieme

$$A_1 \times \dots \times A_k$$

i cui elementi sono le sequenze (a_1, \dots, a_k) di k elementi in cui ciascun a_i è un elemento dell'insieme A_i , per ogni $i = 1, \dots, k$. Possiamo scrivere:

$$A_1 \times \dots \times A_k = \{(a_1, \dots, a_k) \mid a_i \in A_i \forall i\}.$$

Se $k = 2$ ritroviamo il prodotto cartesiano di due insiemi già definito sopra. Anche qui, se tutti gli insiemi coincidono, il prodotto $A \times \cdots \times A$ può essere indicato semplicemente come A^k . Sempre come sopra, il caso geometricamente interessante è quello in cui $A = \mathbb{R}$ e quindi \mathbb{R}^k è l'insieme

$$\mathbb{R}^k = \{(x_1, \dots, x_k) \mid x_i \in \mathbb{R} \forall i\}.$$

Ad esempio, se $k = 3$ otteniamo

$$\mathbb{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbb{R}\}.$$

Questo insieme è l'analogo tridimensionale del piano cartesiano e può essere chiamato lo *spazio cartesiano*. Possiamo pensare ad ogni punto di \mathbb{R}^3 come ad un punto dello spazio con tre coordinate x, y, z .

1.1.9. Ragionamento per induzione. Uno degli strumenti più raffinati della matematica è il *ragionamento per induzione*, che permette di dimostrare in poche righe teoremi piuttosto complessi. Il ragionamento funziona nel modo seguente. Sia $P(n)$ una certa proposizione matematica che dipende da un numero naturale $n \geq 1$. Il nostro scopo è dimostrare che $P(n)$ è vera per ogni n . Per ottenere ciò, è sufficiente completare due passi:

- (1) Dimostrare la proposizione $P(1)$.
- (2) Per n generico, dare per buona $P(n-1)$ e dimostrare $P(n)$.

La proposizione $P(n-1)$ che viene data per buona e che quindi viene usata per dimostrare $P(n)$ è detta *ipotesi induttiva*. Il ragionamento è ben illustrato con un esempio.

Proposizione 1.1.4. *La somma dei primi n numeri naturali è*

$$1 + 2 + 3 + \cdots + n = \frac{(n+1)n}{2}.$$

Dimostrazione. Dimostriamo l'uguaglianza per induzione su n . Il passo iniziale $n = 1$ è facile: l'uguaglianza da dimostrare è semplicemente $1 = 1$. Adesso supponiamo che l'uguaglianza sia vera per $n-1$ e la dimostriamo per n . Scriviamo:

$$1 + 2 + 3 + \cdots + n = (1 + 2 + 3 + \cdots + n - 1) + n.$$

Usando l'ipotesi induttiva, sappiamo che

$$1 + 2 + 3 + \cdots + n - 1 = \frac{n(n-1)}{2}.$$

Il secondo membro dell'equazione precedente adesso diventa

$$1 + 2 + 3 + \cdots + n = \frac{n(n-1)}{2} + n = \frac{n(n-1) + 2n}{2} = \frac{(n+1)n}{2}$$

ed abbiamo concluso. □

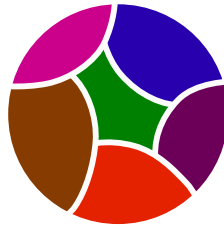


Figura 1.1. Una partizione di un insieme X in sottoinsiemi disgiunti.
Una partizione è di fatto equivalente a una relazione di equivalenza.

1.1.10. Relazione di equivalenza. Introduciamo adesso una nozione che verrà usata sporadicamente in questo libro. Sia X un insieme.

Definizione 1.1.5. Una *relazione di equivalenza* su X è una relazione \sim che occorre fra alcune coppie di elementi di X e che soddisfa queste proprietà:

- (1) $x \sim x \quad \forall x \in X$ (riflessività)
- (2) se $x \sim y$ allora $y \sim x$, $\forall x, y \in X$ (simmetria)
- (3) se $x \sim y$ e $y \sim z$, allora $x \sim z$, $\forall x, y, z \in X$ (proprietà transitiva)

Diciamo che x e y sono *in relazione fra loro* se $x \sim y$.

Esercizio 1.1.6. Fissiamo $n \in \mathbb{N}$ e definiamo una relazione sull'insieme \mathbb{Z} :

$$x \sim y \iff x - y \text{ divisibile per } n$$

Mostra che \sim è una relazione di equivalenza.

Una *partizione* di un insieme X è la suddivisione di X in sottoinsiemi disgiunti, come nella Figura 1.1. Data una partizione, possiamo definire una semplice relazione di equivalenza su X nel modo seguente: $x \sim y \iff x$ e y appartengono allo stesso sottoinsieme della partizione.

Viceversa, data una relazione di equivalenza otteniamo una partizione in sottoinsiemi formati da elementi che sono in relazione fra loro. La definizione più formale è descritta sotto.

Osservazione 1.1.7. La nozione di relazione di equivalenza è molto più concreta di quanto sembri ed è spesso lo strumento usato per partizionare un insieme in sottoinsiemi. Ad esempio, la nozione di *specie biologica* è definita (semplificando molto) nel modo seguente. Sia X l'insieme di tutti gli animali sulla terra. Diciamo che due animali $x, y \in X$ sono in relazione $x \sim y$ se questi (o alcuni loro familiari stretti) possono accoppiarsi fra loro e generare dei figli fertili: due asini sono in relazione, ma un asino ed un cavallo no. Questa ovviamente non è una definizione matematica rigorosa, ma concretamente funziona bene nella maggior parte dei casi: con una buona approssimazione, la relazione \sim è una relazione di equivalenza e da questa discende una partizione di X in sottoinsiemi come gatti, cani, ecc. Ciascun sottoinsieme è per definizione una *specie biologica*.

Formalmente, una relazione d'equivalenza \sim determina una partizione di X in questo modo: per ogni $x \in X$ definiamo $U_x \subset X$ come l'insieme di tutti gli $y \in X$ tali che $x \sim y$. Un insieme del tipo $U_x \subset X$ è detto *classe di equivalenza*. Usando la proprietà transitiva si vede facilmente che tutti gli elementi di U_x sono in relazione fra loro, ma non sono mai in relazione con nessun elemento al di fuori di U_x . Se $x, y \in X$ possono accadere due casi:

- se $x \sim y$, allora $U_x = U_y$,
- se $x \not\sim y$, allora $U_x \cap U_y = \emptyset$.

Due classi di equivalenza distinte sono anche disgiunte e l'unione di tutte le classi di equivalenza è X . Quindi le classi di equivalenza formano una partizione di X .

Ad esempio, nella \sim dell'Esercizio 1.1.6 l'insieme \mathbb{Z} è suddiviso in n classi di equivalenza U_0, U_1, \dots, U_{n-1} , dove

$$U_i = \{i + kn \mid k \in \mathbb{Z}\}.$$

Se \sim è una relazione di equivalenza su X , l'insieme quoziente X/\sim è l'insieme i cui elementi sono le classi di equivalenza di X . Ciascuna classe di equivalenza è adesso considerata un singolo elemento in X/\sim .

Ad esempio, con la \sim dell'esercizio l'insieme \mathbb{Z}/\sim è formato da n elementi:

$$\mathbb{Z}/\sim = \{U_0, U_1, \dots, U_{n-1}\}.$$

Ciascun U_i è un sottoinsieme di \mathbb{Z} .

Esempio 1.1.8. Secondo l'assiomatica moderna, i numeri razionali \mathbb{Q} sono costruiti a partire dagli interi \mathbb{Z} nel modo seguente. Sia $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$ l'insieme dei numeri interi non nulli. Consideriamo il prodotto cartesiano $\mathbb{Z} \times \mathbb{Z}^*$ formato dalle coppie (p, q) di interi con $q \neq 0$. Definiamo una relazione di equivalenza su $\mathbb{Z} \times \mathbb{Z}^*$ nel modo seguente:

$$(p, q) \sim (p', q') \iff pq' = p'q.$$

Si verifica facilmente che \sim è una relazione di equivalenza. Definiamo infine \mathbb{Q} come l'insieme quoziente:

$$\mathbb{Q} = (\mathbb{Z} \times \mathbb{Z}^*)/\sim.$$

Cosa c'entra questa definizione astratta con l'usuale insieme dei numeri razionali? Il collegamento è il seguente: interpretiamo una coppia (p, q) come una frazione $\frac{p}{q}$. La relazione di equivalenza è necessaria qui perché, come ci ricordiamo tutti, in realtà lo stesso numero razionale può essere espresso come due frazioni differenti $\frac{p}{q}$ e $\frac{p'}{q'}$, ad esempio $\frac{2}{3}$ e $\frac{4}{6}$, e questo accade precisamente quando $pq' = p'q$. Quindi alla domanda "che cos'è un numero razionale?" rispondiamo "è una classe di equivalenza di frazioni", cioè una classe di equivalenza di coppie (p, q) dove la relazione \sim è quella descritta sopra.

1.2. Funzioni

Dopo aver richiamato la teoria degli insiemi e la notazione matematica, introduciamo alcuni fra gli oggetti più usati in matematica: le funzioni.

1.2.1. Definizione. Siano A e B due insiemi. Una *funzione* da A in B è una legge f che trasforma qualsiasi punto x di A in un qualche punto y di B . Il punto y ottenuto da x tramite f è indicato come

$$y = f(x).$$

L'insieme di partenza A è detto *dominio* e l'insieme di arrivo B è detto *codominio*. Per indicare f useremo la nozione seguente:

$$f: A \longrightarrow B.$$

È importante ricordare che gli insiemi A e B sono parte integrante della funzione f , non è cioè possibile definire una funzione senza chiarire con precisione quali siano il suo dominio ed il suo codominio.

Ad esempio, possiamo prendere $A = B = \mathbb{R}$ e definire le funzioni seguenti:

$$f(x) = x^2 - 1, \quad f(x) = \text{sen } x.$$

Analogamente, possiamo definire $A = \{x \in \mathbb{R} \mid x \geq 0\}$ e $f: A \rightarrow \mathbb{R}$ come

$$f(x) = \sqrt{x}.$$

In questo caso è però fondamentale chiarire quale delle due radici di x stiamo considerando, perché $f(x)$ deve dipendere da x senza ambiguità. Generalmente si suppone che $f(x) = \sqrt{x}$ sia la radice positiva.

Osservazione 1.2.1. La nozione di funzione è molto flessibile e non si limita a considerare solo quelle funzioni che si possono scrivere esplicitamente usando le quattro operazioni $+$, $-$, \times , $:$ ed altre funzioni note come ad esempio quelle trigonometriche. Ad esempio, possiamo scegliere $A = B = \mathbb{N}$ e definire $f(n)$ come l' $(n+1)$ -esimo numero primo. Oppure scegliere $A = \mathbb{R}$, $B = \mathbb{N}$ e definire $f(x)$ come la 127-esima cifra di x nel suo sviluppo decimale. In entrambi i casi è impossibile (oppure molto difficile o inutile) scrivere f come una funzione algebrica esplicita. Dobbiamo rassegnarci al fatto che la maggior parte delle funzioni non sono scrivibili esplicitamente usando le usuali operazioni algebriche.

Due funzioni f e g sono considerate *uguali* se hanno lo stesso dominio e lo stesso codominio, e se $f(x) = g(x)$ per ogni elemento x del dominio. Quindi ad esempio le funzioni

$$f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^2, \quad g: \mathbb{N} \rightarrow \mathbb{R}, g(x) = x^2$$

non sono uguali perché hanno domini diversi, mentre le funzioni

$$f: \mathbb{N} \rightarrow \mathbb{N}, f(x) = 1, \quad g: \mathbb{N} \rightarrow \mathbb{N}, g(x) = (x-1)^2 - x^2 + 2x$$

sono invece in realtà la stessa funzione e scriviamo $f = g$.

Un modo molto semplice per modificare una funzione data consiste nel restringere il dominio ad un sottoinsieme. Se $f: A \rightarrow B$ è una funzione e $A' \subset A$ è un sottoinsieme, la *restrizione* di f a A' è la funzione $f': A' \rightarrow B$ definita esattamente come f , ponendo cioè $f'(x) = f(x)$ per ogni $x \in A'$. La restrizione f' è indicata generalmente con il simbolo $f|_{A'}$.

Ad esempio, la restrizione di $f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^2$ al sottoinsieme \mathbb{N} è indicata con $f|_{\mathbb{N}}$ ed è ovviamente la funzione

$$f|_{\mathbb{N}}: \mathbb{N} \rightarrow \mathbb{R}, \quad f|_{\mathbb{N}}(x) = x^2.$$

1.2.2. Immagine e controimmagine. Il concetto di funzione è estremamente importante in matematica e nelle scienze, ed ha quindi diritto ad un vocabolario tutto suo, che viene purtroppo raramente introdotto nelle scuole e con cui il lettore deve acquistare familiarità.

Sia $f: A \rightarrow B$ una funzione. L'*immagine* di un elemento $x \in A$ è l'elemento $f(x)$ associato a x tramite f . Più in generale, se $C \subset A$ è un qualsiasi sottoinsieme, l'*immagine* di C è l'insieme

$$f(C) = \{f(x) \mid x \in C\} \subset B.$$

L'immagine $f(C)$ è un sottoinsieme del codominio B , ed è l'unione di tutte le immagini di tutti gli elementi di C . Ad esempio, se $f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^2$ e $C = \mathbb{N}$, allora

$$f(\mathbb{N}) = \{0, 1, 4, 9, 16, \dots\}.$$

L'*immagine* della funzione f è per definizione l'immagine $f(A)$ dell'intero dominio A . Ad esempio l'immagine della funzione $f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^2$ appena descritta è la semiretta $f(\mathbb{R}) = [0, +\infty)$ formata da tutti i numeri reali positivi o nulli.

Esiste un'altra nozione che è in un certo senso opposta a quella di immagine. Se $y \in B$ è un elemento del codominio, la sua *controimmagine* è il sottoinsieme $f^{-1}(y) \subset A$ del dominio che consiste di tutti gli elementi x la cui immagine è y , in altre parole:

$$f^{-1}(y) = \{x \in A \mid f(x) = y\}.$$

Più in generale, se $C \subset B$ è un sottoinsieme del codominio, la sua *controimmagine* è il sottoinsieme $f^{-1}(C) \subset A$ del dominio che consiste di tutti gli elementi la cui immagine è contenuta in C . In altre parole:

$$f^{-1}(C) = \{x \in A \mid f(x) \in C\}.$$

Esempio 1.2.2. Considerando la stessa funzione $f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^2$ di prima, verifichiamo che:

- la controimmagine di 4 è l'insieme con due elementi $f^{-1}(4) = \{-2, 2\}$;
- la controimmagine del segmento chiuso $C = [4, 9]$ è l'unione di due segmenti chiusi $f^{-1}(C) = [-3, -2] \cup [2, 3]$.

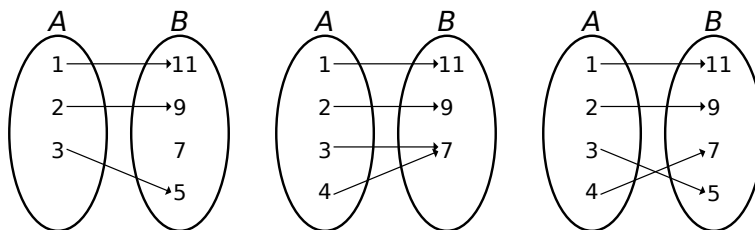


Figura 1.2. Un esempio di funzione iniettiva, suriettiva e bigettiva.

- la controimmagine del segmento chiuso $[-10, -5]$ è l'insieme vuoto $f^{-1}([-10, -5]) = \emptyset$.

Esempio 1.2.3. Consideriamo la funzione $f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = \sin x$. L'immagine di f è il segmento chiuso $[-1, 1]$. La controimmagine di \mathbb{N} è l'insieme

$$f^{-1}(\mathbb{N}) = \left\{ \dots, -\frac{3\pi}{2}, -\pi, -\frac{\pi}{2}, 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}, \dots \right\} = \left\{ \frac{n\pi}{2} \mid n \in \mathbb{Z} \right\}.$$

Esercizio 1.2.4. Dimostrare che vale sempre l'inclusione $C \subset f^{-1}(f(C))$. Costruire un esempio in cui il contenimento \subset è un'uguaglianza $=$ ed un altro esempio in cui il contenimento è stretto \subsetneq .

1.2.3. Funzioni iniettive, suriettive e bigettive. Sia $f: A \rightarrow B$ una funzione. Definiamo alcune nozioni che saranno di fondamentale importanza nei capitoli successivi.

Definizione 1.2.5. La funzione f è

- *iniettiva* se due elementi distinti $x \neq x'$ del dominio hanno sempre immagini distinte. In altre parole se

$$\forall x, x' \in A, \quad x \neq x' \implies f(x) \neq f(x');$$

- *suriettiva* se ogni elemento del codominio è immagine di almeno un elemento del dominio. In altre parole, se l'immagine di f coincide con l'intero codominio B , cioè

$$f(A) = B.$$

Si veda la Figura 1.2. Qualche esempio:

Esempio 1.2.6. La funzione $f: \mathbb{N} \rightarrow \mathbb{N}, f(x) = x^2$ è iniettiva, perché i quadrati di due numeri naturali diversi sono sempre diversi; non è suriettiva, perché esistono dei numeri naturali (ad esempio il numero 2) che non sono quadrati e quindi $f(\mathbb{N}) \subsetneq \mathbb{N}$.

Esempio 1.2.7. La funzione $f: \mathbb{R} \rightarrow [0, +\infty), f(x) = x^2$ è suriettiva perché qualsiasi numero reale $y \geq 0$ ha una radice quadrata; non è iniettiva perché questa radice non è unica, infatti $f(-x) = f(x)$.

I due esempi precedenti ci mostrano ancora quanto sia importante la scelta del dominio e del codominio nello studio delle proprietà di una funzione.

Una funzione $f: A \rightarrow B$ che è sia iniettiva che suriettiva è detta *bigezione* o *corrispondenza biunivoca*. Se f è una bigezione, ogni elemento $y \in B$ ha *esattamente* una controimmagine: ne ha almeno una perché è suriettiva, e non può averne più di una perché è iniettiva. Questo è un ragionamento cruciale che la lettrice deve assimilare bene prima di proseguire.

Se $f: A \rightarrow B$ è una bigezione, possiamo definire la *funzione inversa*

$$f^{-1}: B \rightarrow A$$

nel modo seguente: per ogni $y \in B$, l'elemento $f^{-1}(y)$ è proprio quell'unico x tale che $f(x) = y$. Nella Figura 1.2 a destra, la funzione inversa f^{-1} è ottenuta da f semplicemente invertendo il verso delle frecce. La funzione inversa f^{-1} è anch'essa una bigezione.

1.2.4. Composizione di funzioni. È utile pensare ad una funzione $f: A \rightarrow B$ come ad una trasformazione che prende come *input* un qualsiasi elemento $x \in A$ e restituisce come *output* la sua immagine $f(x) \in B$. L'elemento $f(x)$ può essere a sua volta trasformato usando un'altra funzione $g: B \rightarrow C$ e questa successione di due trasformazioni è detta *composizione di funzioni*.

Più formalmente, date due funzioni

$$f: A \rightarrow B, \quad g: B \rightarrow C$$

definiamo la loro *composizione* h come una nuova funzione

$$h: A \rightarrow C$$

nel modo seguente:

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

La funzione composta h è indicata come

$$h = g \circ f.$$

Si usa quindi il simbolo \circ per indicare l'operazione di composizione di due funzioni. Notiamo che due funzioni f e g possono essere composte solo se il codominio di f è uguale al dominio di g .

Esempio 1.2.8. Componendo le funzioni $f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^3$ e $g: \mathbb{R} \rightarrow \mathbb{R}, g(y) = \sin y$ otteniamo $h = g \circ f, h(x) = \sin x^3$. Notiamo che la composizione generalmente non è commutativa: se scambiamo l'ordine delle funzioni otteniamo un'altra funzione $j = f \circ g, j(x) = \sin^3 x$.

Se f e g sono entrambe bigezioni, allora anche $h = g \circ f$ è una bigezione e la sua inversa è

$$h^{-1} = f^{-1} \circ g^{-1}.$$

Esercizio 1.2.9. Siano $f: A \rightarrow B$ e $g: B \rightarrow C$ due funzioni.

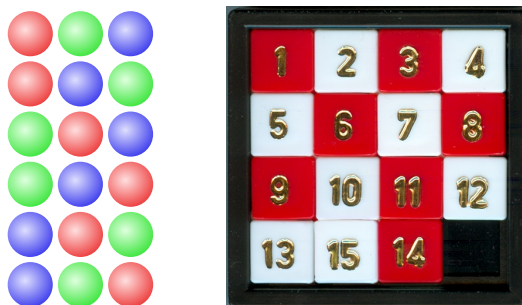


Figura 1.3. Le $3! = 6$ permutazioni di un insieme di tre palline (sinistra).
Una configurazione irrisolvibile del gioco del 15 (destra).

- (1) Se $g \circ f$ è iniettiva, allora f è iniettiva.
- (2) Se $g \circ f$ è suriettiva, allora g è suriettiva.

1.2.5. Permutazioni. Sia X un insieme di n elementi. Una *permutazione* è una bigezione

$$\sigma: X \longrightarrow X.$$

Come è noto dalle superiori, ci sono $n!$ possibili permutazioni per X , si veda la Figura 1.3-(sinistra). Notiamo che l'inversa σ^{-1} di una permutazione σ è sempre una permutazione, e la composizione $\sigma \circ \tau$ di due permutazioni σ e τ è una permutazione. La permutazione *identità* è la permutazione *id* che fissa ciascun elemento, cioè $\text{id}(x) = x \forall x \in X$.

A meno di rinominare gli elementi di X , possiamo supporre per semplicità che X sia l'insieme numerico formato dai numeri naturali da 1 a n :

$$X = \{1, \dots, n\}.$$

Vediamo come possiamo scrivere e studiare una permutazione σ . Un modo consiste nello scrivere la tabella

$$\begin{bmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{bmatrix}$$

oppure più semplicemente $[\sigma(1) \sigma(2) \dots \sigma(n)]$.

Un altro metodo consiste nello scrivere σ come prodotto di *cicli*. Se a_1, \dots, a_k sono elementi distinti di X , il *ciclo*

$$(a_1 \dots a_k)$$

indica la permutazione che trasla ciclicamente gli elementi a_1, \dots, a_k e lascia fissi tutti gli altri, cioè tale che:

$$\begin{aligned} \sigma(a_1) &= a_2, & \sigma(a_2) &= a_3, & \dots & \sigma(a_{k-1}) &= a_k, & \sigma(a_k) &= a_1, \\ \sigma(a) &= a, & \forall a &\notin \{a_1, \dots, a_k\}. \end{aligned}$$

Due cicli $(a_1 \dots a_k)$ e $(b_1 \dots b_h)$ sono *indipendenti* se $a_i \neq b_j$ per ogni i, j . Ogni permutazione si scrive come prodotto di cicli indipendenti. Ad esempio:

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 3 & 6 & 2 & 1 \end{bmatrix}$$

si scrive come prodotto di cicli

$$(1 \ 4 \ 6)(2 \ 5)(3) = (1 \ 4 \ 6)(2 \ 5).$$

Il prodotto di cicli come composizione va inteso da destra a sinistra, come le funzioni. I cicli di ordine uno possono chiaramente essere rimossi. L'ordine dei cicli indipendenti non è importante: $(1 \ 4 \ 6)(2 \ 5)$ e $(2 \ 5)(1 \ 4 \ 6)$ sono la stessa permutazione. Inoltre il ciclo $(1 \ 4 \ 6)$ può essere scritto anche come $(4 \ 6 \ 1)$ oppure $(6 \ 1 \ 4)$, ma *non* come $(1 \ 6 \ 4)$.

La notazione come prodotto di cicli ha il pregio di funzionare molto bene con le operazioni di inversione e composizione. Per scrivere l'inversa di una permutazione basta invertire i cicli. L'inversa della permutazione $(1 \ 4 \ 6)(2 \ 5)$ è $(5 \ 2)(6 \ 4 \ 1)$. Per comporre due permutazioni basta affiancare i cicli da destra a sinistra: se componiamo $(2 \ 4 \ 1)(3 \ 5 \ 6)$ e $(1 \ 4 \ 6)(2 \ 5)$ otteniamo

$$(1 \ 4 \ 6)(2 \ 5)(2 \ 4 \ 1)(3 \ 5 \ 6) = (1 \ 5)(6 \ 3 \ 2).$$

Infatti leggendo da destra a sinistra vediamo che $4 \rightarrow 1 \rightarrow 4$ sta fisso mentre $1 \rightarrow 2 \rightarrow 5$ e $5 \rightarrow 6 \rightarrow 1$.

1.2.6. Segno di una permutazione. Vediamo adesso che ciascuna permutazione σ ha un *segno* che può essere 1 oppure -1 .

Un ciclo di ordine 2 è chiamato una *trasposizione*. Una trasposizione è una permutazione che scambia due elementi e lascia fissi tutti gli altri. Si dimostra facilmente che ciascun ciclo è prodotto di trasposizioni, infatti

$$(a_1 \dots a_k) = (a_1 \ a_k)(a_1 \ a_{k-1}) \cdots (a_1 \ a_2).$$

Conseguentemente, ogni permutazione può essere ottenuta come prodotto di un certo numero n di trasposizioni. Si può fare ciò in molti modi, però la parità di n dipende solo dalla permutazione σ , in virtù della proposizione seguente.

Proposizione 1.2.10. *Se una permutazione σ si scrive in due modi diversi come prodotto di n e m trasposizioni, il numero $n - m$ è sempre pari.*

Dimostrazione. Abbiamo per ipotesi

$$\begin{aligned} \sigma &= (a_1 \ a_2)(a_3 \ a_4) \cdots (a_{2n-1} \ a_{2n}), \\ \sigma &= (b_1 \ b_2)(b_3 \ b_4) \cdots (b_{2m-1} \ b_{2m}). \end{aligned}$$

Quindi

$$\text{id} = \sigma \circ \sigma^{-1} = (a_1 \ a_2) \cdots (a_{2n-1} \ a_{2n})(b_{2m} \ b_{2m-1}) \cdots (b_2 \ b_1).$$

Dimostriamo che id non è realizzabile come prodotto di un numero dispari $h = 2k + 1$ di trasposizioni: questo implica che $m + n$ è pari e quindi $m - n$

è pari. Procediamo per induzione su $k \geq 0$. Se $k = 0$, è ovvio che id non è realizzabile come una singola trasposizione. Supponiamo il fatto dimostrato per $k - 1$ e passiamo a $k \geq 1$. Supponiamo che

$$\text{id} = (a_1 a_2) \cdots (a_{2h-1} a_{2h})$$

con $h = 2k + 1$. Siccome la permutazione è l'identità, il termine a_1 deve comparire, oltre che nella prima trasposizione a sinistra, almeno una seconda volta (letto da sinistra a destra):

$$\text{id} = (a_1 a_2) \cdots (a_1 a_i) \cdots (a_{2h-1}, a_{2h}).$$

Notiamo le seguenti uguaglianze, in cui lettere diverse indicano numeri diversi:

$$(c d)(a b) = (a b)(c d), \quad (b c)(a b) = (a c)(b c).$$

Possiamo usare queste "mosse" per modificare due trasposizioni successive senza cambiare il numero totale $h = 2k + 1$ di trasposizioni. Usando le mosse con $a = a_1$, possiamo spostare il secondo a_1 verso sinistra di un passo alla volta, finché non arriva in seconda posizione e otteniamo:

$$\text{id} = (a_1 a_2)(a_1 a_3) \cdots (a_{2h-1}, a_{2h}).$$

Se $a_2 = a_3$, abbiamo $(a_1 a_2)(a_1 a_2) = \text{id}$ e possiamo cancellare le prime due trasposizioni. Troviamo una successione di $2(k - 1) + 1$ elementi e giungiamo ad un assurdo per l'ipotesi induttiva.

Se $a_2 \neq a_3$ possiamo comunque sostituire $(a_1 a_2)(a_1 a_3)$ con $(a_1 a_3)(a_2 a_3)$. In questo modo la successione di trasposizioni contiene un a_1 in meno di prima e ripartiamo da capo. Siccome prima o poi gli a_1 finiscono, prima o poi ricadremo nel caso precedente $a_2 = a_3$. \square

Definiamo il *segno* $\text{sgn}(\sigma)$ di una permutazione σ come

$$\text{sgn}(\sigma) = (-1)^n$$

dove σ si decompone in n trasposizioni. Il segno è 1 oppure -1 ed è ben definito grazie alla Proposizione 1.2.10.

Un ciclo di ordine n ha segno $(-1)^{n-1}$. Una trasposizione ha segno -1 . L'identità ha segno 1.

Esempio 1.2.11. Usando il segno di una permutazione si può dimostrare che la configurazione mostrata nella Figura 1.3-(destra) del gioco del 15 non è risolvibile. Si procede in questo modo. La prima cosa da notare è che per trasformare la configurazione in figura in quella giusta dobbiamo fare un numero pari di mosse: questo è dovuto al fatto che, se pensiamo ad una scacchiera 4×4 , la casella vuota è inizialmente in una casella nera (in basso a destra), e ad ogni passaggio salta da una casella nera ad una bianca e viceversa. Siccome alla fine deve tornare su una nera (sempre in basso a destra), deve fare un numero pari di salti.

Ogni configurazione può essere interpretata come una permutazione σ dell'insieme $X = \{1, \dots, 15, N\}$ dove N indica la casella nera. La casella i è occupata dal tassello $\sigma(i)$. La soluzione cercata corrisponde all'identità id , in cui la casella i è occupata dal tassello $i \forall i \in X$.

La configurazione σ mostrata nella Figura 1.3 è una trasposizione (14 15) e quindi ha segno negativo $\text{sgn}(\sigma) = -1$. Ad ogni mossa che facciamo componiamo σ con una trasposizione e quindi cambiamo segno alla σ . Dopo un numero pari di mosse otterremo sempre una permutazione con segno negativo, e quindi mai l'identità. Se partiamo dalla configurazione mostrata in figura, il gioco del 15 non si può risolvere.

Il segno cambia in modo controllato per inversione e composizioni:

Proposizione 1.2.12. *Per ogni $\sigma, \tau \in S_n$ abbiamo:*

$$\text{sgn}(\sigma^{-1}) = \text{sgn}(\sigma), \quad \text{sgn}(\sigma \circ \tau) = \text{sgn}(\tau \circ \sigma).$$

Dimostrazione. Se σ e τ sono scritte come prodotto di m e n trasposizioni, possiamo scrivere σ^{-1} e $\sigma \circ \tau$ come prodotto di m e $m + n$ trasposizioni. \square

1.3. Polinomi

I polinomi sono tipi particolarmente semplici di funzioni ottenuti combinando numeri e variabili ed usando solo le operazioni $+$, $-$ e \times .

1.3.1. Definizione. Ricordiamo che un *monomio* è una espressione algebrica che ha una parte numerica (il coefficiente) ed una parte letterale; ad esempio questi sono monomi:

$$4x, \quad -2xy, \quad \sqrt{5}x^3.$$

Il *grado* di un monomio è la somma degli esponenti presenti sulle parti letterali: i tre monomi descritti sopra hanno grado 1, 2 e 3. Un monomio di grado zero è semplicemente un numero.

Un *polinomio* è una somma di monomi, ad esempio:

$$7 + 3x^2 - \sqrt{2}y^3.$$

Un polinomio è *ridotto in forma normale* se è scritto come somma di monomi con parti letterali differenti e coefficienti non nulli, oppure è il polinomio 0. Per ridurre un polinomio in forma normale è sufficiente raccogliere i monomi con la stessa parte letterale e quindi eliminare quelli con coefficiente nullo. Il *grado* di un polinomio scritto in forma normale è il massimo grado dei suoi monomi. I polinomi possono essere sommati e moltiplicati fra loro nel modo usuale.

Ci interessano particolarmente i polinomi in cui compare una sola variabile x . Un polinomio di questo tipo viene indicato con $p(x)$ oppure più semplicemente con p . Un polinomio $p(x)$ con una sola variabile può essere sempre

descritto ordinando i suoi monomi da quello di grado più alto a quello di grado più basso: otteniamo quindi una scrittura del tipo

$$p(x) = a_n x^n + \cdots + a_1 x + a_0$$

dove n è il grado di $p(x)$ e $a_n \neq 0$. Ad esempio:

$$x^3 - 2x + 5, \quad 4x^2 - 7.$$

Il coefficiente a_0 è detto *termine noto* del polinomio. Un polinomio di grado zero è semplicemente un numero a_0 .

1.3.2. Divisione con resto fra polinomi. I polinomi assomigliano ai numeri interi: possono essere sommati, moltiplicati, e si possono fare le divisioni con resto.

Prendiamo due numeri interi, ad esempio 26 e 11. Se dividiamo 26 per 11 otteniamo come quoziente 2 e come resto 4. In altre parole, otteniamo:

$$26 = 2 \cdot 11 + 4.$$

Notiamo che il resto 4 è ovviamente sempre più piccolo del divisore 11. Analogamente, dati due polinomi $p(x)$ (il *dividendo*) e $p'(x)$ (il *divisore*), esistono sempre (e sono unici) due polinomi $q(x)$ (il *quoziente*) e $r(x)$ (il *resto*) per cui

$$p(x) = q(x)p'(x) + r(x)$$

con la proprietà che il resto $r(x)$ abbia grado strettamente minore del divisore $p'(x)$. Le divisioni fra polinomi si risolvono con carta e penna esattamente con la stessa procedura usata per i numeri interi. Ad esempio, se dividiamo $p(x) = x^3 + 1$ per $p'(x) = x^2 - 1$ otteniamo

$$x^3 + 1 = x(x^2 - 1) + (x + 1)$$

e la divisione ha come quoziente x e come resto $x + 1$.

Diciamo che il numero intero 7 divide 14 ma non divide 15, perché la divisione di 14 per 7 ha resto nullo, mentre la divisione di 15 per 7 invece ha un certo resto. Usiamo la stessa terminologia per i polinomi: se la divisione fra due polinomi $p(x)$ e $p'(x)$ non ha resto, allora $p(x) = q(x)p'(x)$ per qualche quoziente $q(x)$ e diciamo che $p'(x)$ *divide* $p(x)$. Possiamo usare la barra verticale $|$ come sinonimo di "divide" e scrivere ad esempio:

$$9 \mid 18, \quad (x + 1) \mid (x^3 + 1).$$

Notiamo che effettivamente $(x^3 + 1) = (x^2 - x + 1)(x + 1)$.

1.3.3. Radici di un polinomio. Ricordiamo adesso una delle definizioni più importanti dell'algebra. Se $p(x)$ è un polinomio e a è un numero, indichiamo con $p(a)$ il numero che otteniamo sostituendo a al posto di x . Ad esempio, se $p(x) = x^2 - 3$, allora $p(-2) = 4 - 3 = 1$.

Definizione 1.3.1. Un numero a è *radice* di un polinomio $p(x)$ se $p(a) = 0$.

Ad esempio, il numero -1 è radice del polinomio $p(x) = x^3 + 1$ perché $p(-1) = 0$. La determinazione delle radici di un polinomio è uno dei problemi più classici dell'algebra. A questo scopo è utile il criterio seguente.

Proposizione 1.3.2. *Il numero a è radice di $p(x)$ se e solo se $(x - a) \mid p(x)$.*

Dimostrazione. Se dividiamo $p(x)$ per $(x - a)$, otteniamo

$$p(x) = q(x)(x - a) + r(x)$$

dove $q(x)$ è il quoziente e $r(x)$ il resto. Sappiamo che il grado di $r(x)$ è strettamente minore di quello di $(x - a)$, che è uno: quindi $r(x)$ ha grado zero, in altre parole è una costante che scriviamo semplicemente come r_0 . Quindi

$$p(x) = q(x)(x - a) + r_0.$$

Se sostituiamo a al posto di x , otteniamo

$$p(a) = r_0.$$

Quindi a è radice di $p(x)$ se e solo se $r_0 = 0$. D'altra parte $r_0 = 0$ se e solo se $(x - a)$ divide $p(x)$ e quindi concludiamo. \square

Introduciamo un'altra definizione che useremo spesso in questo libro.

Definizione 1.3.3. La *molteplicità* di una radice a di un polinomio $p(x)$ è il massimo numero k tale che $(x - a)^k$ divide $p(x)$.

Informalmente, la molteplicità di a misura "quante volte" a è radice di $p(x)$. Ad esempio, la radice 1 ha molteplicità 1 nel polinomio $x^3 - 1$ e 2 nel polinomio $x^2 - 2x + 1$.

Osserviamo un fatto semplice: se moltiplichiamo un polinomio $p(x)$ per una costante k diversa da zero, otteniamo un altro polinomio $q(x) = kp(x)$ che ha le stesse radici di $p(x)$ con le stesse molteplicità. Per questo motivo, quando studiamo le radici di un polinomio di grado n del tipo

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

possiamo dividerlo per $a_n \neq 0$; in questo modo ciascun coefficiente a_i si trasformerà in un nuovo coefficiente $b_i = a_i/a_n$ e otterremo un nuovo polinomio

$$q(x) = x^n + b_{n-1} x^{n-1} + \cdots + b_1 x + b_0$$

che ha il pregio di avere il primo coefficiente pari a 1 . Un polinomio di questo tipo è detto *monico*.

Proposizione 1.3.4. *Sia $p(x) = q_1(x)q_2(x)$. Le radici di $p(x)$ contate con molteplicità sono l'unione di quelle di $q_1(x)$ e di $q_2(x)$.*

In particolare se a è radice di $q_1(x)$ con molteplicità m_1 ed è anche radice di $q_2(x)$ con molteplicità m_2 , allora a è radice di $p(x)$ con molteplicità $m_1 + m_2$. Informalmente a è " m_1 volte" radice di $q_1(x)$ e " m_2 volte" radice di $q_2(x)$, quindi " $m_1 + m_2$ volte" radice di $p(x)$ in totale.

Dimostrazione. Se a ha molteplicità m_1 in $q_1(x)$ e m_2 in $q_2(x)$, allora $q_1(x) = (x - a)^{m_1} \bar{q}_1(x)$ e $q_2(x) = (x - a)^{m_2} \bar{q}_2(x)$ con $q_1(a) \neq 0$ e $q_2(a) \neq 0$. Allora $p(x) = (x - a)^{m_1+m_2} \bar{q}_1(x) \bar{q}_2(x)$ con $\bar{q}_1(a) \bar{q}_2(a) \neq 0$ e quindi a ha molteplicità $m_1 + m_2$ in $p(x)$. \square

Teorema 1.3.5. *Un polinomio $p(x)$ di grado $n \geq 1$ ha al più n radici, contate con molteplicità.*

Dimostrazione. Dimostriamo il teorema per induzione su n . Per quanto appena visto, a meno di dividere per il primo coefficiente possiamo supporre che il polinomio $p(x)$ sia monico. Se $n = 1$, il polinomio è del tipo $p(x) = x + a_0$ ed ha chiaramente una sola radice $-a_0$. Quindi la tesi è soddisfatta.

Supponiamo la tesi vera per $n - 1$ e la dimostriamo per n . Se $p(x)$ non ha radici, siamo a posto. Se ha almeno una radice a , allora per la Proposizione 1.3.2 possiamo dividere $p(x)$ per $x - a$ e ottenere un altro polinomio $q(x)$, cioè vale $p(x) = (x - a)q(x)$. Il polinomio $q(x)$ ha grado $n - 1$ e quindi per ipotesi induttiva ha al più $n - 1$ radici contate con molteplicità. Per la Proposizione 1.3.4, le radici di $p(x)$ contate con molteplicità sono esattamente quelle di $q(x)$ più a . Quindi $p(x)$ ha al più $n - 1 + 1 = n$ radici e abbiamo concluso. \square

Un polinomio monico di grado 1 è sempre del tipo $x - a_0$ ed ha quindi sempre una sola radice a_0 . Come sappiamo bene dalle scuole superiori, un polinomio di grado due può avere al massimo due radici reali che possono essere determinate con la nota formula di risoluzione.

Proposizione 1.3.6. *Un polinomio di secondo grado $p(x) = ax^2 + bx + c$ a coefficienti reali ha due radici reali (contate con molteplicità) se e solo se $\Delta \geq 0$, dove $\Delta = b^2 - 4ac$. Queste sono:*

$$x_{\pm} = \frac{-b \pm \sqrt{\Delta}}{2a}.$$

Dimostrazione. Se $\Delta \geq 0$, allora i numeri x_{\pm} hanno senso ed è sufficiente sostituirli in $p(x)$ per verificare che sono radici:

$$\begin{aligned} p(x_{\pm}) &= a \frac{b^2 \mp 2b\sqrt{\Delta} + \Delta}{4a^2} + b \frac{-b \pm \sqrt{\Delta}}{2a} + c \\ &= \frac{\mp b\sqrt{\Delta} + b^2 - 2ac - b^2 \pm b\sqrt{\Delta} + 2ac}{2a} = 0. \end{aligned}$$

Se $\Delta = 0$, in realtà possiamo scrivere

$$p(x) = a \left(x + \frac{b}{2a} \right)^2$$

e quindi la radice $-\frac{b}{2a}$ ha chiaramente molteplicità due. Se $\Delta < 0$, scriviamo

$$p(x) = a \left(x + \frac{b}{2a} \right)^2 + c - \frac{b^2}{4a} = a \left(x + \frac{b}{2a} \right)^2 + \frac{-\Delta}{4a}$$

e scopriamo che $p(x)$ è sempre strettamente positivo o strettamente negativo per ogni x , a seconda che a sia positivo o negativo. Quindi non può mai essere nullo, in altre parole $p(x)$ non ha radici. \square

Come sapevamo dalle superiori, esistono polinomi di secondo grado senza radici reali. Come vedremo nella prossima sezione, è possibile ovviare a questo “problema” aggiungendo all’insieme \mathbb{R} dei numeri reali altri numeri, così da formare il più vasto insieme \mathbb{C} dei *numeri complessi*.

1.4. Numeri complessi

I numeri complessi sono un ampliamento dell’insieme dei numeri reali \mathbb{R} , costruito con lo scopo di ottenere migliori qualità algebriche.

1.4.1. Definizione. Un *numero complesso* è un oggetto algebrico che si scrive nel modo seguente:

$$a + bi$$

dove a e b sono numeri reali arbitrari e i è un nuovo simbolo chiamato *unità immaginaria*. Ad esempio, questi sono numeri complessi:

$$\sqrt{7}, \quad 2 + i, \quad 23i, \quad 4 - i, \quad -1 + \pi i.$$

I numeri complessi si sommano e si moltiplicano nel modo usuale, tenendo a mente un’unica nuova relazione:

$$i^2 = -1.$$

La somma e la moltiplicazione di due numeri complessi $a + bi$ e $c + di$ si svolge quindi nel modo seguente:

$$\begin{aligned} (a + bi) + (c + di) &= a + c + (b + d)i, \\ (a + bi) \cdot (c + di) &= ac + bci + adi + bdi^2 = ac - bd + (ad + bc)i. \end{aligned}$$

Nel prodotto abbiamo usato che $i^2 = -1$. Ad esempio:

$$(7 + i) \cdot (4 - i) = 29 - 3i.$$

L’insieme dei numeri complessi è indicato con il simbolo \mathbb{C} . Aggiungendo ai numeri reali l’unità immaginaria i , abbiamo esteso la nostra sequenza di insiemi numerici:

$$\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R} \subsetneq \mathbb{C}.$$

1.4.2. Coniugio, norma e inverso. Sia $z = a + bi$ un numero complesso. I numeri a e b sono detti rispettivamente la *parte reale* e la *parte immaginaria* di z . Il numero z è reale, cioè appartiene al sottoinsieme $\mathbb{R} \subset \mathbb{C}$, se e solo se la sua parte immaginaria è nulla.

Il *coniugio* di z è il numero complesso

$$\bar{z} = a - ib$$

ottenuto da z cambiando di segno la sua parte immaginaria. Notiamo che $z = \bar{z}$ precisamente quando $b = 0$, cioè quando z è reale. Scriviamo quindi

$$z \in \mathbb{R} \iff z = \bar{z}.$$

Il *modulo* di $z = a + ib$ è il numero reale

$$|z| = \sqrt{a^2 + b^2}.$$

Il modulo $|z|$ è nullo quando $z = 0$, cioè quando $a = b = 0$, ed è strettamente positivo se $z \neq 0$. Notiamo inoltre che

$$z \cdot \bar{z} = (a + bi) \cdot (a - bi) = a^2 + b^2 = |z|^2.$$

Mostriamo adesso un fatto non banale: come nei numeri razionali e reali, ogni numero complesso $z \neq 0$ ha un *inverso* z^{-1} rispetto all'operazione di moltiplicazione, dato da

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

Infatti se moltiplichiamo z e z^{-1} otteniamo

$$z \cdot z^{-1} = \frac{z\bar{z}}{|z|^2} = 1.$$

Esempio 1.4.1. L'inverso di i è $-i$, infatti $i \cdot (-i) = 1$. L'inverso di $2 + i$ è

$$(2 + i)^{-1} = \frac{2 - i}{|2 + i|^2} = \frac{2 - i}{5}.$$

1.4.3. Il piano complesso. Mentre i numeri reali \mathbb{R} formano una retta, i numeri complessi \mathbb{C} formano un piano detto *piano complesso*. Ogni numero complesso $a + bi$ può essere identificato con il punto di coordinate (a, b) nel piano cartesiano o equivalentemente come un vettore applicato nell'origine 0 e diretto verso (a, b) , come illustrato nella Figura 1.4-(sinistra).

Gli assi delle ascisse e delle ordinate sono rispettivamente l'*asse reale* e l'*asse immaginaria* di \mathbb{C} . L'asse reale è precisamente il sottoinsieme $\mathbb{R} \subset \mathbb{C}$ formato dai numeri reali. L'asse immaginaria consiste di tutti i numeri complessi del tipo bi al variare di $b \in \mathbb{R}$.

La somma $z_1 + z_2$ di due numeri complessi z_1 e z_2 viene calcolata interpretando ciascun numero z_1 e z_2 come un vettore e sommando quindi i due vettori con l'usuale regola del parallelogramma, come mostrato nella Figura 1.5.

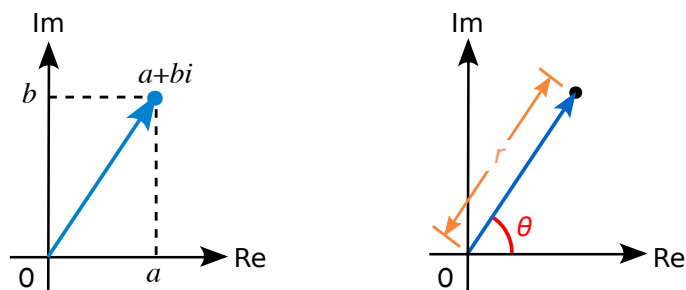


Figura 1.4. Il piano complesso (sinistra) e le coordinate polari (destra).

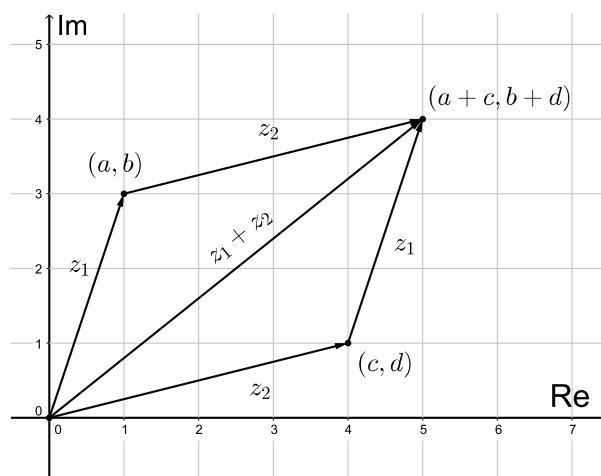


Figura 1.5. La somma $z_1 + z_2$ di due numeri complessi z_1 e z_2 può essere calcolata con la regola del parallelogramma.

Il prodotto $z_1 \cdot z_2$ di due numeri complessi è apparentemente più complicato, ma può essere visualizzato agevolmente usando le *coordinate polari*, che ora richiamiamo.

1.4.4. Coordinate polari. Come ricordato nella Figura 1.4-(destra), un punto (x, y) diverso dall'origine del piano cartesiano può essere identificato usando la lunghezza r del vettore corrispondente e l'angolo θ formato dal vettore con l'asse reale. Le *coordinate polari* del punto sono la coppia (r, θ) . Per passare dalle coordinate polari (r, θ) a quelle cartesiane (x, y) basta usare le formule

$$x = r \cos \theta, \quad y = r \sin \theta.$$

Viceversa,

$$r = \sqrt{x^2 + y^2}, \quad \theta = \arccos \frac{x}{\sqrt{x^2 + y^2}}.$$

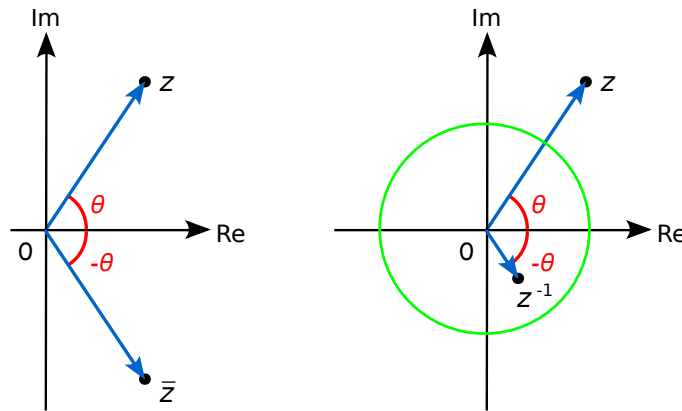


Figura 1.6. Il coniugio \bar{z} di z si ottiene specchiando z lungo l'asse reale (sinistra). L'inverso z^{-1} di z ha argomento $-\theta$ opposto a quello θ di z e modulo $|z^{-1}|$ inverso rispetto a $|z|$. La circonferenza unitaria è mostrata in figura (destra).

Tornando ai numeri complessi, un numero $z = x + yi$ può essere scritto in coordinate polari come

$$z = x + yi = r \cos \theta + (r \sin \theta)i = r(\cos \theta + i \sin \theta).$$

Notiamo che

$$|z| = \sqrt{x^2 + y^2} = r.$$

Il modulo di z è quindi descritto geometricamente come la lunghezza del vettore che descrive z . Il coniugio $\bar{z} = a - ib$ è il punto ottenuto cambiando il segno della coordinata immaginaria: geometricamente questo corrisponde a riflettere il punto rispetto all'asse reale. In coordinate polari, questo corrisponde a cambiare θ in $-\theta$ lasciando fisso r . Si veda la Figura 1.6-(sinistra).

Tornando alle coordinate polari, è comodo scrivere

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

In questo modo ogni numero complesso $z \neq 0$ si scrive come

$$z = r e^{i\theta}.$$

Il numero $r = |z|$ è il modulo di z e l'angolo θ è detto *fase* o *argomento* di z .

Il motivo profondo per cui introduciamo inaspettatamente qui la costante e di Nepero è dovuto alle rappresentazioni delle funzioni e^x , $\sin x$ e $\cos x$ come serie di potenze. Giustificare questa scelta ci porterebbe troppo lontano; per noi è sufficiente considerare questa misteriosa esponenziale complessa $e^{i\theta}$ come un simbolo che vuol dire semplicemente $\cos \theta + i \sin \theta$. Si tratta di una simbologia azzeccata, perché $e^{i\theta}$ ha le proprietà usuali dell'esponenziale:

Proposizione 1.4.2. *Vale la relazione*

$$e^{i(\theta+\varphi)} = e^{i\theta} \cdot e^{i\varphi}.$$

Dimostrazione. Usando le formule di duplicazione otteniamo

$$\begin{aligned} e^{i(\theta+\varphi)} &= \cos(\theta + \varphi) + i \operatorname{sen}(\theta + \varphi) \\ &= \cos \theta \cos \varphi - \operatorname{sen} \theta \operatorname{sen} \varphi + i(\operatorname{sen} \theta \cos \varphi + \cos \theta \operatorname{sen} \varphi) \\ &= (\cos \theta + i \operatorname{sen} \theta) \cdot (\cos \varphi + i \operatorname{sen} \varphi) \\ &= e^{i\theta} \cdot e^{i\varphi}. \end{aligned}$$

La dimostrazione è completa. \square

Adesso capiamo perché le coordinate polari sono particolarmente utili quando moltiplichiamo due numeri complessi. Se

$$z_1 = r_1 e^{i\theta_1}, \quad z_2 = r_2 e^{i\theta_2}$$

allora il loro prodotto è semplicemente

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

In altre parole:

Quando si fa il prodotto di due numeri complessi, i moduli si moltiplicano e gli argomenti si sommano.

Notiamo in particolare che se $z = r e^{i\theta} \neq 0$, il suo inverso è

$$z^{-1} = r^{-1} e^{-i\theta}.$$

L'inverso z^{-1} ha argomento $-\theta$ opposto a quello θ di z e ha modulo $|z^{-1}| = r^{-1}$ inverso rispetto a $|z| = r$, si veda la Figura 1.6-(destra).

Notiamo infine che i numeri complessi $e^{i\theta}$ al variare di θ sono precisamente i punti che stanno sulla circonferenza unitaria, determinati dall'angolo θ . In particolare per $\theta = \pi$ otteniamo la celebre *identità di Eulero*:

$$e^{i\pi} = -1.$$

Notiamo infine che due numeri complessi non nulli espressi in forma polare

$$r_0 e^{i\theta_0}, \quad r_1 e^{i\theta_1}$$

sono lo stesso numero complesso se e solo se valgono entrambi i fatti seguenti:

- $r_0 = r_1$,
- $\theta_1 = \theta_0 + 2k\pi$ per qualche $k \in \mathbb{Z}$.

1.4.5. Proprietà dei numeri complessi. I numeri complessi hanno numerose proprietà. Queste si dimostrano facilmente: in presenza di un prodotto, è spesso utile usare la rappresentazione polare. Lasciamo la dimostrazione di queste proprietà per esercizio.

Esercizio 1.4.3. Valgono i fatti seguenti per ogni $z, w \in \mathbb{C}$:

$$\begin{aligned} |z + w| &\leq |z| + |w|, & |zw| &= |z||w|, & |z|^{-1} &= \frac{1}{|z|}, \\ |z| &= |\bar{z}|, & \overline{z + w} &= \bar{z} + \bar{w}, & \overline{zw} &= \bar{z}\bar{w}. \end{aligned}$$

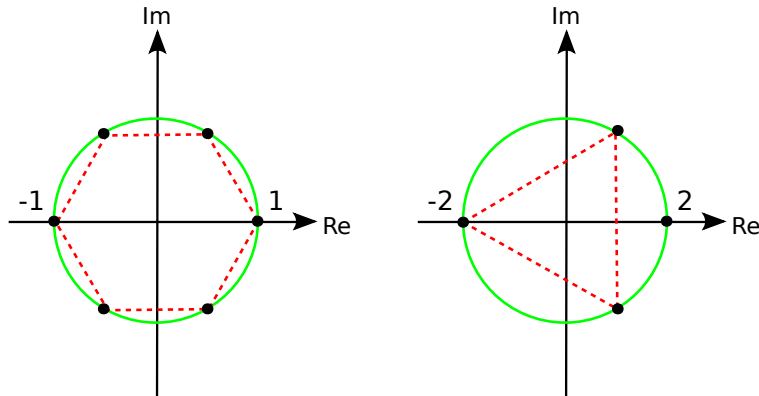


Figura 1.7. Le radici seste di 1, cioè le soluzioni di $z^6 = 1$, sono i vertici di un esagono regolare centrato nell'origine con un vertice in 1. La figura mostra la circonferenza unitaria (sinistra). Le soluzioni di $z^3 = -8$ sono i vertici di un triangolo equilatero centrato nell'origine e di raggio $\sqrt[3]{8} = 2$. La figura mostra la circonferenza di raggio 2 (destra).

1.4.6. Radici n -esime di un numero complesso. Sia $z_0 = r_0 e^{i\theta_0}$ un numero complesso fissato diverso da zero. Mostriamo come risolvere l'equazione

$$z^n = z_0$$

usando i numeri complessi. In altre parole, determiniamo tutte le radici n -esime di z_0 . Scriviamo la variabile z in forma polare come $z = r e^{i\theta}$. L'equazione adesso diventa

$$r^n e^{in\theta} = r_0 e^{i\theta_0}.$$

L'equazione è soddisfatta precisamente se valgono entrambi i fatti seguenti:

- $r = \sqrt[n]{r_0}$,
- $n\theta = \theta_0 + 2k\pi$ per qualche $k \in \mathbb{Z}$.

La seconda condizione può essere riscritta richiedendo che

$$\theta = \frac{\theta_0}{n} + \frac{2k\pi}{n}$$

per qualche $k \in \mathbb{Z}$. Otteniamo quindi precisamente n soluzioni distinte: per i valori $k = 0, 1, \dots, n-1$ otteniamo gli argomenti

$$\theta = \frac{\theta_0}{n}, \frac{\theta_0}{n} + \frac{2\pi}{n}, \dots, \frac{\theta_0}{n} + \frac{2(n-1)\pi}{n}.$$

Le n soluzioni dell'equazione $z^n = z_0$ hanno tutte lo stesso modulo $\sqrt[n]{r_0}$ e argomenti che variano in una sequenza di angoli separati da un passo costante $\frac{2\pi}{n}$. Geometricamente, questo significa che le soluzioni formano i vertici di un poligono regolare con n lati e raggio $\sqrt[n]{r_0}$. Si veda la Figura 1.7.

Esempio 1.4.4. L'equazione $z^n = 1$ ha come soluzioni i numeri complessi

$$z = e^{i\frac{2k\pi}{n}}$$

dove $k = 0, 1, \dots, n - 1$. Queste n soluzioni sono i vertici di un poligono regolare di raggio 1 con n lati, avente 1 come vertice. Questi numeri sono le *radici n -esime dell'unità*.

Esempio 1.4.5. Come mostrato nella Figura 1.7-(destra), le tre soluzioni dell'equazione $z^3 = -8$ hanno modulo $\sqrt[3]{8} = 2$ e argomento $\frac{\pi}{3}, \pi$ e $\frac{5\pi}{3}$. Si tratta dei numeri complessi

$$\begin{aligned} z_1 &= 2e^{i\frac{\pi}{3}} = 2 \left(\cos \frac{\pi}{3} + i \sin \frac{\pi}{3} \right) = 1 + \sqrt{2}i, \\ z_2 &= 2e^{i\pi} = -2, \\ z_3 &= 2e^{i\frac{5\pi}{3}} = 2 \left(\cos \frac{5\pi}{3} + i \sin \frac{5\pi}{3} \right) = 1 - \sqrt{2}i. \end{aligned}$$

Esercizio 1.4.6. Calcola le soluzioni dell'equazione $z^4 = i$.

1.4.7. Teorema fondamentale dell'algebra. Veniamo infine al vero motivo per cui abbiamo introdotto i numeri complessi.

Teorema 1.4.7 (Teorema fondamentale dell'algebra). *Ogni polinomio a coefficienti complessi ha almeno una radice complessa.*

Esistono varie dimostrazioni di questo teorema, che possono essere agevolmente trovate in rete. Quelle più accessibili usano degli strumenti analitici che si discostano da quanto trattato in questo libro e quindi le omettiamo.

Corollario 1.4.8. *Un polinomio $p(x)$ a coefficienti complessi di grado n ha esattamente n radici, contate con molteplicità.*

Dimostrazione. Dimostriamo il corollario per induzione su n . Possiamo supporre senza perdita di generalità che $p(x)$ sia monico. Se $n = 1$ allora $p(x) = x - a_0$ ha una sola radice $x = a_0$.

Dimostriamo il caso generico n dando per buono il caso $n - 1$. Sappiamo per il teorema fondamentale dell'algebra che $p(x)$ ha almeno una radice a . Quindi per la Proposizione 1.3.2 possiamo scrivere $p(x) = (x - a)q(x)$ dove $q(x)$ è un altro polinomio di grado $n - 1$. Per l'ipotesi induttiva $q(x)$ ha $n - 1$ radici contate con molteplicità, e quindi $p(x)$ ha queste $n - 1$ radici più a , quindi $p(x)$ ha esattamente n radici (sempre contate con molteplicità). \square

Per un polinomio di secondo grado

$$p(x) = ax^2 + bx + c$$

le due radici complesse si trovano usando la solita formula

$$x_{\pm} = \frac{-b \pm \sqrt{\Delta}}{2a}.$$

Adesso $\pm\sqrt{\Delta}$ indica le due radici quadrate complesse di Δ , che esistono sempre.

Esempio 1.4.9. Le radici del polinomio $x^2 + 1$ sono $\pm i$. Le radici del polinomio $x^2 + (1 - i)x - i$ sono

$$x_{\pm} = \frac{-1 + i \pm \sqrt{2i}}{2} \implies x_{\pm} = \frac{-1 + i \pm (1 + i)}{2} \implies x_+ = i, x_- = -1.$$

Scriviamo il Corollario 1.4.8 in un'altra forma:

Corollario 1.4.10. *Ogni polinomio $p(x)$ a coefficienti complessi si spezza come prodotto di polinomi di primo grado:*

$$p(x) = a_n(x - z_1) \cdots (x - z_n)$$

dove a_n è il coefficiente più grande di $p(x)$ e z_1, \dots, z_n sono le radici complesse di $p(x)$ contate con molteplicità.

1.4.8. Polinomi a coefficienti reali. Sappiamo che un polinomio $p(x)$ di grado n ha esattamente n soluzioni complesse contate con molteplicità. Se $p(x)$ ha coefficienti reali, possiamo dire qualcosa di più sulle sue radici complesse.

Proposizione 1.4.11. *Sia $p(x)$ un polinomio a coefficienti reali. Se z è una radice complessa di $p(x)$, allora \bar{z} è anch'essa radice di $p(x)$.*

Dimostrazione. Il polinomio è

$$p(x) = a_n x^n + \cdots + a_1 x + a_0$$

e per ipotesi i coefficienti a_n, \dots, a_0 sono tutti reali. Se z è radice, allora

$$p(z) = a_n z^n + \cdots + a_1 z + a_0 = 0.$$

Applicando il coniugio ad entrambi i membri e l'Esercizio 1.4.3 troviamo

$$\bar{a}_n \bar{z}^n + \cdots + \bar{a}_1 \bar{z} + \bar{a}_0 = \bar{0} = 0.$$

Siccome i coefficienti sono reali, il coniugio di a_i è sempre a_i e quindi

$$a_n \bar{z}^n + \cdots + a_1 \bar{z} + a_0 = 0.$$

In altre parole, anche \bar{z} è radice di $p(x)$. □

Possiamo dedurre un teorema di spezzamento per i polinomi a coefficienti reali che preveda una scomposizione in fattori di grado 2 e 1. I fattori di grado 2 hanno radici complesse z e \bar{z} coniugate fra loro, quelli di grado 1 hanno radici reali.

Corollario 1.4.12. *Ogni polinomio $p(x)$ a coefficienti reali si spezza come:*

$$p(x) = q_1(x) \cdots q_k(x) \cdot (x - x_1) \cdots (x - x_h)$$

dove $q_1(x), \dots, q_k(x)$ sono polinomi di grado due a coefficienti reali con $\Delta < 0$, e x_1, \dots, x_h sono le radici reali di $p(x)$ contante con molteplicità.

Dimostrazione. Ragioniamo come sempre per induzione sul grado n di $p(x)$. Se $p(x)$ ha grado 1 allora $p(x) = x - x_1$ e siamo a posto.

Se $p(x)$ ha grado n , ha qualche radice complessa z . Se z è reale, scriviamo $z = x_1$, spezziamo $p(x) = (x - x_1)q(x)$ e concludiamo per induzione su $q(x)$. Se z è complesso, allora sappiamo che anche \bar{z} è soluzione. Quindi $p(x) = (x - z)(x - \bar{z})q(x)$. Se $z = a + bi$ allora

$$q_1(x) = (x - z)(x - \bar{z}) = (x - (a + bi))(x - (a - bi)) = x^2 - 2ax + a^2 + b^2$$

è un polinomio a coefficienti reali con $\Delta = -4b^2 < 0$. Scriviamo $p(x) = q_1(x)q(x)$ e concludiamo per induzione su $q(x)$. \square

Abbiamo capito che le radici complesse non reali di un polinomio $p(x)$ a coefficienti reali sono presenti a coppie coniugate z, \bar{z} . In particolare, sono in numero pari. Ne deduciamo il fatto seguente.

Proposizione 1.4.13. *Un polinomio $p(x)$ a coefficienti reali di grado dispari ha sempre almeno una soluzione reale.*

Dimostrazione. Sappiamo che $p(x)$ ha grado n dispari e che ha n soluzioni complesse contate con molteplicità. Di queste, un numero pari non sono reali. Quindi restano un numero dispari di soluzioni reali – quindi almeno una c'è. \square

Osservazione 1.4.14. Esiste un'altra dimostrazione di questa proposizione che usa l'analisi. Siccome $p(x)$ ha grado dispari, i limiti $\lim_{x \rightarrow +\infty} p(x)$ e $\lim_{x \rightarrow -\infty} p(x)$ sono entrambi infiniti, ma con segni opposti. Quindi la funzione $p(x)$ assume valori sia positivi che negativi. Dal Teorema di esistenza degli zeri segue che la funzione $p(x)$ assume anche il valore nullo per qualche x_0 . Quindi x_0 è radice di $p(x)$.

1.5. Strutture algebriche

Quando abbiamo introdotto i numeri complessi, abbiamo notato che questi hanno diverse proprietà algebriche simili ai numeri reali. In questa sezione esplicitiamo queste proprietà e diamo un nome agli insiemi dotati di operazioni che le soddisfano.

1.5.1. Gruppi. Un *gruppo* è un insieme G dotato di una *operazione binaria*, cioè di una funzione che associa ad ogni coppia a, b di elementi in G un nuovo elemento che indichiamo con $a * b$. Il simbolo $*$ indica l'operazione binaria. L'operazione deve soddisfare i seguenti tre assiomi:

- (1) $\exists e \in G : e * a = a * e = a \forall a \in G$ (esistenza dell'*elemento neutro* e)
- (2) $a * (b * c) = (a * b) * c \forall a, b, c \in G$ (*proprietà associativa*)
- (3) $\forall a \in G, \exists a' \in G : a * a' = a' * a = e$ (esistenza dell'*inverso*)

Ad esempio, l'insieme \mathbb{Z} dei numeri interi con l'operazione di somma è un gruppo. Infatti:

- (1) esiste l'elemento neutro 0, per cui $0 + a = a + 0 = a \forall a \in \mathbb{Z}$

- (2) vale la proprietà associativa $a + (b + c) = (a + b) + c$
 (3) ogni numero intero $a \in \mathbb{Z}$ ha un inverso $a' = -a$, per cui $a + (-a) = (-a) + a = 0$.

Notiamo invece che *non* sono gruppi:

- L'insieme \mathbb{Z} con la moltiplicazione, perché non è soddisfatto (3).
- L'insieme \mathbb{N} con la somma, sempre perché non è soddisfatto (3).

Il gruppo G è *commutativo* se vale anche la *proprietà commutativa*

$$a * b = b * a \quad \forall a, b \in G.$$

I numeri interi \mathbb{Z} formano un gruppo commutativo. Formano un gruppo commutativo anche gli insiemi numerici \mathbb{Q} , \mathbb{R} e \mathbb{C} con l'operazione di somma. Un altro esempio di gruppo è l'insieme

$$S_n$$

formato da tutte le $n!$ permutazioni dell'insieme $X = \{1, \dots, n\}$, con l'operazione \circ di composizione, si veda la Sezione 1.2.5. Sappiamo infatti che:

- (1) esiste l'elemento neutro $\text{id} \in S_n$, la permutazione identità $\text{id}(i) = i$,
- (2) vale la proprietà associativa $\rho \circ (\sigma \circ \tau) = (\rho \circ \sigma) \circ \tau$, $\forall \rho, \sigma, \tau \in S_n$,
- (3) ogni permutazione σ ha una inversa σ^{-1} .

Il gruppo S_n delle permutazioni è detto *gruppo simmetrico*. A differenza dei gruppi numerici \mathbb{Z} , \mathbb{Q} , \mathbb{R} e \mathbb{C} , notiamo che S_n contiene un numero finito di elementi e non è commutativo se $n \geq 3$: è facile trovare permutazioni che non commutano, ad esempio le trasposizioni (1 2) e (2 3) non commutano:

$$(2 \ 3 \ 1) = (1 \ 2)(2 \ 3) \neq (2 \ 3)(1 \ 2) = (1 \ 3 \ 2).$$

In algebra si dimostrano vari teoremi sui gruppi. Il primo fatto da notare è che in un gruppo l'inverso di un qualsiasi elemento $a \in G$ è sempre unico.

Proposizione 1.5.1. Se $\exists a, a', a'' \in G$ tali che

$$a * a' = a' * a = e, \quad a * a'' = a'' * a = e,$$

allora $a' = a''$.

Dimostrazione. Troviamo

$$a' = a' * (a * a'') = (a' * a) * a'' = a''.$$

La dimostrazione è conclusa. \square

Un altro fatto importante è che con i gruppi si può *semplificare*. Se troviamo una espressione del tipo

$$a * b = a * c$$

con $a, b, c \in G$, allora possiamo moltiplicare per l'inverso a^{-1} a sinistra in entrambi i membri e ottenere

$$a^{-1} * a * b = a^{-1} * a * c \implies b = c.$$

Notiamo che abbiamo usato l'esistenza dell'inverso e la proprietà associativa.

1.5.2. Anelli. Negli insiemi numerici appena considerati ci sono in realtà due operazioni $+$ e \times . Introduciamo adesso una struttura algebrica che prevede la coesistenza di due operazioni binarie.

Un *anello* è un insieme A dotato di due operazioni binarie $+$ e \times che soddisfano questi assiomi:

- (1) A è un gruppo commutativo con l'operazione $+$
- (2) $a \times (b \times c) = (a \times b) \times c \quad \forall a, b, c \in A$ (*proprietà associativa del \times*)
- (3) $a \times (b + c) = a \times b + a \times c \quad \forall a, b, c \in A$ (*proprietà distributiva*)
- (4) $\exists 1 \in A : 1 * a = a * 1 = a \quad \forall a \in A$ (*elemento neutro per il \times*)

Gli interi \mathbb{Z} con le operazioni di somma e prodotto sono ancora un esempio di anello. L'elemento neutro della somma è lo 0, mentre quello del prodotto è 1. Anche gli insiemi \mathbb{Q} , \mathbb{R} e \mathbb{C} sono un anello con le operazioni di somma e prodotto. Un anello è *commutativo* se vale la proprietà commutativa per il prodotto:

$$a \times b = b \times a \quad \forall a, b \in A.$$

Gli anelli \mathbb{Z} , \mathbb{Q} , \mathbb{R} e \mathbb{C} sono tutti commutativi. In un anello A gli elementi neutri per le operazioni $+$ e \times sono generalmente indicati con 0 e 1. Dagli assiomi di anello possiamo subito dedurre qualche teorema. Ad esempio:

Proposizione 1.5.2. *Vale $0 \times a = a \times 0 = 0$ per ogni $a \in A$.*

Dimostrazione. Abbiamo

$$0 \times a = (0 + 0) \times a = 0 \times a + 0 \times a.$$

Semplificando deduciamo che $0 = 0 \times a$. Dimostrazione analoga per $a \times 0$. \square

1.5.3. Campi. Definiamo infine un'ultima struttura algebrica, che è quella più importante in questo libro.

Un *campo* è un anello commutativo K in cui vale anche il seguente assioma:

$\forall a \in K, a \neq 0 \implies \exists a' \in K : a \times a' = a' \times a = 1$ (esiste inverso per il prodotto)

Chiediamo quindi che ogni elemento a diverso da zero abbia anche un inverso rispetto al prodotto. Esempi di campi sono \mathbb{Q} , \mathbb{R} e \mathbb{C} . Notiamo invece che \mathbb{Z} non è un campo, perché ovviamente 2 non ha un inverso per il prodotto dentro l'insieme \mathbb{Z} . Riassumendo:

- \mathbb{N} non è un gruppo.
- \mathbb{Z} è un anello ma non è un campo.
- \mathbb{Q} , \mathbb{R} e \mathbb{C} sono campi.

In modo analogo alla Proposizione 1.5.1 dimostriamo che, in un qualsiasi anello A , se un elemento $a \in A$ ha un inverso per il prodotto, questo è unico. L'inverso per il prodotto di a è generalmente indicato con a^{-1} .

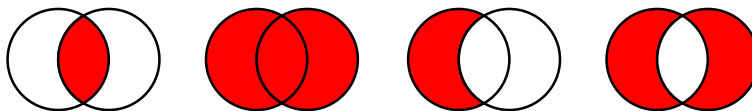


Figura 1.8. Intersezione, unione, differenza e differenza simmetrica di due insiemi.

Esercizi

Esercizio 1.1. Sia $n \in \mathbb{N}$. Dimostra che \sqrt{n} è razionale $\iff n$ è il quadrato di un numero naturale.

Esercizio 1.2. Dimostra che esistono infiniti numeri primi, nel modo seguente. Se per assurdo ne esistesse solo un numero finito p_1, \dots, p_n , allora mostra che il numero $p_1 \cdots p_n + 1$ non è divisibile per nessuno di loro.

Esercizio 1.3. La *differenza simmetrica* di due insiemi A e B è l'insieme

$$A \Delta B = (A \setminus B) \cup (B \setminus A).$$

Si veda la Figura 1.8. Mostra che Δ è commutativa e associativa:

$$A \Delta B = B \Delta A, \quad (A \Delta B) \Delta C = A \Delta (B \Delta C).$$

Esercizio 1.4. Dato un insieme X , l'*insieme delle parti* $\mathcal{P}(X)$ è l'insieme formato da tutti i sottoinsiemi di X , incluso il vuoto \emptyset e X stesso. Ad esempio, se $X = \{1, 2\}$,

$$\mathcal{P}(X) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}.$$

Mostra che se X contiene n elementi allora $\mathcal{P}(X)$ ne contiene 2^n .

Esercizio 1.5. Mostra per induzione l'uguaglianza seguente per ogni $n \geq 1$:

$$\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}.$$

Esercizio 1.6. Mostra per induzione che qualsiasi numero del tipo $n^3 + 2n$ con $n \in \mathbb{N}$ è divisibile per 3.

Esercizio 1.7. Dimostra che n rette a coppie non parallele dividono il piano in $\frac{(n+1)n}{2} + 1$ regioni differenti.

Esercizio 1.8. Determina l'unica fra le seguenti che non è una relazione di equivalenza su \mathbb{R} :

- $x \sim y \iff x - y \in \mathbb{Z}$,
- $x \sim y \iff x^3 - 4x = y^3 - 4y$,
- $x \sim y \iff x \neq y + 1$.

Esercizio 1.9. Sia $f: A \rightarrow B$ una funzione e $S, T \subset A$ due sottoinsiemi. Dimostra:

$$f(S \cup T) = f(S) \cup f(T),$$

$$f(S \cap T) \subset f(S) \cap f(T).$$

Costruisci un esempio in cui l'inclusione \subset è stretta e un altro esempio in cui è un'uguaglianza.

Esercizio 1.10. Sia $f: A \rightarrow B$ una funzione e $S, T \subset B$ due sottoinsiemi. Dimostra:

$$\begin{aligned} f^{-1}(S \cup T) &= f^{-1}(S) \cup f^{-1}(T), \\ f^{-1}(S \cap T) &= f^{-1}(S) \cap f^{-1}(T). \end{aligned}$$

Esercizio 1.11. Sia X un insieme finito. Mostra che una funzione $f: X \rightarrow X$ è iniettiva \iff è suriettiva. Mostra che questo fatto non è necessariamente vero se X è infinito.

Esercizio 1.12. Determina tutte le radici del polinomio complesso $z^4 = -16$.

Esercizio 1.13. Determina tutti i numeri complessi z tali che $z^4 = \bar{z}^3$.

Esercizio 1.14. Sia $n \in \mathbb{N}$, $n \geq 1$ fissato. Considera l'insieme

$$C_n = \{1, 2, \dots, n\}.$$

e definisci la seguente operazione binaria su C_n : per ogni $x, y \in C_n$, la loro *somma* $x + y$ è il resto della divisione di $x + y$ per n . Mostra che C_n con questa operazione è un gruppo commutativo con elemento neutro 0. Il gruppo C_n si chiama *gruppo ciclico*.

Esercizio 1.15. Sia $n \geq 2$. Sul gruppo C_n definito nell'esercizio precedente definiamo analogamente il *prodotto* $x \cdot y$ come il resto della divisione di xy per n . Mostra che C_n con questa operazione è un anello commutativo. Per quali valori di n secondo te C_n è un campo?

Esercizio 1.16. Considera l'insieme di numeri complessi

$$S = \{z \in \mathbb{C} \mid |z| = 1\}.$$

Mostra che S è un gruppo con l'operazione di prodotto. Chi è l'elemento neutro?

Esercizio 1.17. Un elemento $a \in A$ in un anello A è *invertibile* se esiste un inverso a^{-1} per il prodotto, cioè un elemento tale che $a \times a^{-1} = a^{-1} \times a = 1$. Dimostra che il prodotto ab di due elementi invertibili è anch'esso invertibile, e che $(ab)^{-1} = b^{-1}a^{-1}$.

Complementi

Infiniti numerabili e non numerabili. Mostriamo che in matematica esistono infiniti diversi fra loro, alcuni dei quali sono più grandi degli altri. Tutto parte come sempre dalla teoria degli insiemi e dalle funzioni.

Definizione 1.5.3. Un insieme infinito X è detto *numerabile* se esiste una bigezione fra \mathbb{N} e X .

Gli insiemi numerabili sono tutti in bigezione con \mathbb{N} e quindi anche in bigezione fra loro: possiamo dire che contengono "lo stesso numero di elementi", anche se questo numero è chiaramente infinito.

Concretamente, scrivere una bigezione $f: \mathbb{N} \rightarrow X$ equivale a rappresentare gli elementi di X come una successione infinita

$$X = \{f(0), f(1), f(2), \dots\}.$$

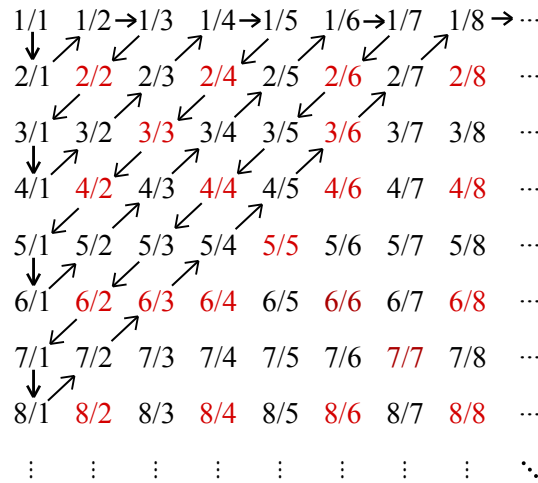


Figura 1.9. I numeri razionali \mathbb{Q} sono un insieme numerabile. Per rappresentare l'insieme dei numeri razionali positivi come una successione è sufficiente seguire il percorso qui indicato, saltando i numeri (rossi) che sono già stati incontrati precedentemente (ricordiamo che frazioni diverse possono indicare lo stesso numero). Con uno schema simile si possono inserire anche i razionali negativi o nulli e quindi ottenere una bigezione fra \mathbb{N} e \mathbb{Q} .

Ad esempio, il sottoinsieme $P \subset \mathbb{N}$ formato dai numeri pari è numerabile perché possiamo rappresentare gli elementi di P come una successione infinita:

$$P = \{0, 2, 4, 6, 8, \dots\}$$

La bigezione $f: \mathbb{N} \rightarrow P$ è definita semplicemente chiedendo che $f(n)$ sia l' n -esimo elemento della successione. In questo caso $f: \mathbb{N} \rightarrow P$ può essere scritta esplicitamente come $f(n) = 2n$.

Notiamo come, nonostante P sia contenuto strettamente in \mathbb{N} , esista una bigezione fra i due insiemi. Questa è una proprietà che possono avere solo gli insiemi infiniti: se X è finito e $Y \subset X$ è un sottoinsieme proprio, chiaramente non può esistere una bigezione fra X e Y .

Un altro esempio di insieme numerabile è l'insieme \mathbb{Z} dei numeri interi. Infatti possiamo metterli in successione nel modo seguente:

$$\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, 4, \dots\}$$

Un esempio ancora più sorprendente è quello dei razionali \mathbb{Q} . I numeri razionali possono essere messi "in fila" come indicato nella Figura 1.9. Nonostante i numeri razionali sembrano molti di più dei naturali, in realtà sono "lo stesso numero".

A questo punto può sorgere il dubbio che semplicemente tutti gli insiemi infiniti siano numerabili. Questo non è il caso:

Proposizione 1.5.4. *L'insieme \mathbb{R} dei numeri reali non è numerabile.*

Dimostrazione. Supponiamo per assurdo che sia possibile scrivere \mathbb{R} come successione infinita di numeri:

$$\mathbb{R} = \{x_0, x_1, x_2, \dots\}.$$

Scriviamo ciascun numero in forma decimale, ad esempio:

$$x_1 = 723,1291851\dots$$

$$x_2 = 12,8231452\dots$$

$$x_3 = 0,3798921\dots$$

$$x_4 = 110,0023140\dots$$

...

Ad essere precisi esistono alcuni numeri reali che possono essere rappresentati in due modi come forma decimale: quelli che ad un certo punto si stabilizzano con una successione di 9 o di 0. Ad esempio $0,18563\bar{9} = 0,18564$. In questo caso scegliamo la successione che si stabilizza con 0.

Indichiamo con n_i la i -esima cifra dopo la virgola di x_i . Definiamo adesso un nuovo numero reale x nel modo seguente:

$$x = 0, a_1 a_2 a_3 a_4 a_5 \dots$$

con questa regola: se $n_i \neq 1$ allora $a_i = 1$; se $n_i = 1$ allora $a_i = 2$.

Adesso notiamo che x non può essere uguale a nessun numero della successione x_1, \dots e quindi giungiamo ad un assurdo. \square

La tecnica usata nella dimostrazione è l'*argomento diagonale di Cantor*. Abbiamo scoperto che i numeri reali \mathbb{R} sono "di più" dei numeri naturali, interi o razionali. Sono sempre infiniti, ma di un ordine superiore.

CAPITOLO 2

Spazi vettoriali

2.1. Lo spazio euclideo

Nella geometria di Euclide, tutto ha inizio con l'introduzione di alcuni oggetti geometrici di base quali i punti, le rette e i piani, che non vengono definiti e per i quali valgono alcuni assiomi. Nella geometria analitica adottata in questo libro, partiamo invece dalla teoria degli insiemi ed in particolare dalla retta reale \mathbb{R} , e costruiamo una geometria a partire da questa.

2.1.1. Definizione. Partendo da \mathbb{R} e usando l'operazione di prodotto cartesiano fra insiemi, definiamo subito la nozione di spazio euclideo in dimensione arbitraria. Sia $n \geq 1$ un numero naturale.

Definizione 2.1.1. Lo *spazio euclideo* n -dimensionale è l'insieme \mathbb{R}^n .

Ricordiamo che \mathbb{R}^n è il prodotto cartesiano $\mathbb{R} \times \cdots \times \mathbb{R}$ di n copie di \mathbb{R} . Un elemento dell'insieme \mathbb{R}^n è una successione (x_1, \dots, x_n) di n numeri reali.

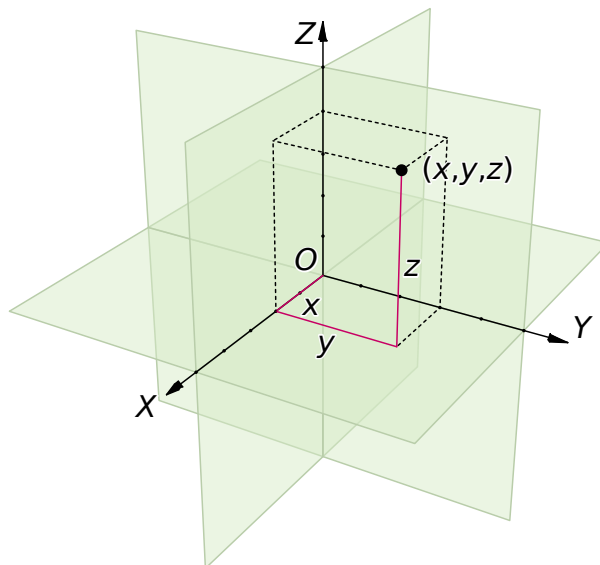


Figura 2.1. Lo spazio cartesiano \mathbb{R}^3 . Ciascun punto è una terna (x, y, z) di numeri reali. La figura mostra i tre assi e i tre piani coordinati che contengono gli assi, descritti dalle equazioni $x = 0$, $y = 0$ e $z = 0$.

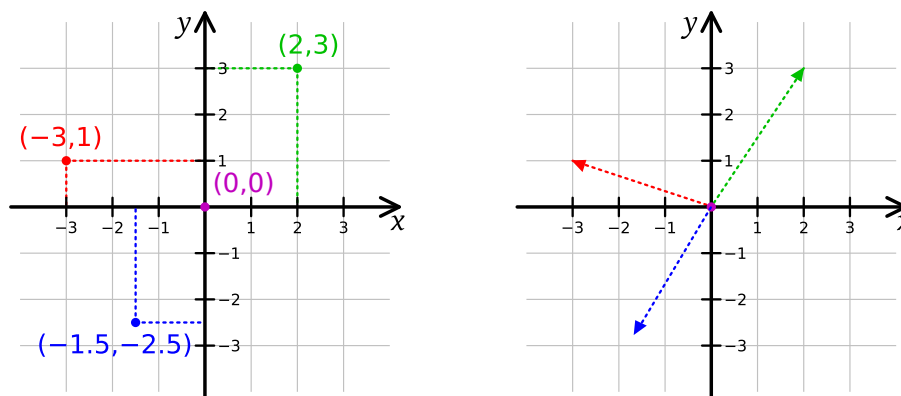


Figura 2.2. Ciascun punto $(x, y) \in \mathbb{R}^2$ può essere interpretato come un vettore applicato sull'origine e viceversa.

L'insieme \mathbb{R}^2 è il piano cartesiano, in cui ogni punto è determinato da una coppia (x, y) . Analogamente \mathbb{R}^3 è lo spazio cartesiano, in cui ogni punto è una terna (x, y, z) di numeri reali, si veda la Figura 2.1. Quando n è arbitrario (ad esempio $n = 4$) lo spazio \mathbb{R}^n è uno “spazio a n dimensioni”.

Gli spazi \mathbb{R}^2 e \mathbb{R}^3 possono essere visualizzati e studiati usando la nostra intuizione geometrica. D'altro canto, il nostro cervello non è programmato per visualizzare \mathbb{R}^n per $n \geq 4$, ma possiamo comunque studiare \mathbb{R}^n usando l'algebra, ed è esattamente quello che faremo in questo libro.

L'elemento $(0, \dots, 0)$ è l'*origine* dello spazio euclideo \mathbb{R}^n e viene indicato semplicemente con i simboli 0 oppure O .

2.1.2. Punti o vettori? Un punto $x \in \mathbb{R}^n$ è per definizione una sequenza (x_1, \dots, x_n) di numeri. Geometricamente, possiamo pensare a x come ad un punto, oppure come ad un *vettore* (cioè una freccia) che parte dall'origine 0 e arriva in x . Si veda la Figura 2.2. Entrambi i punti di vista sono ammessi e a seconda del contesto penseremo a x come ad un punto o ad un vettore. In alcuni casi, se pensiamo ad x come ad un punto possiamo usare le lettere P, Q , mentre se pensiamo ad x come vettore usiamo le lettere v, w .

Per un motivo che sarà chiaro in seguito quando introdurremo le matrici, scriveremo sempre la sequenza che identifica il punto x in verticale:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Un oggetto del genere è detto un *vettore colonna*. Diciamo che i numeri x_1, \dots, x_n sono le *coordinate* del punto (o del vettore) x .

2.1.3. Somme fra vettori. Lo spazio euclideo \mathbb{R}^n non è soltanto un insieme: ha anche una struttura algebrica che introduciamo adesso. Definiamo

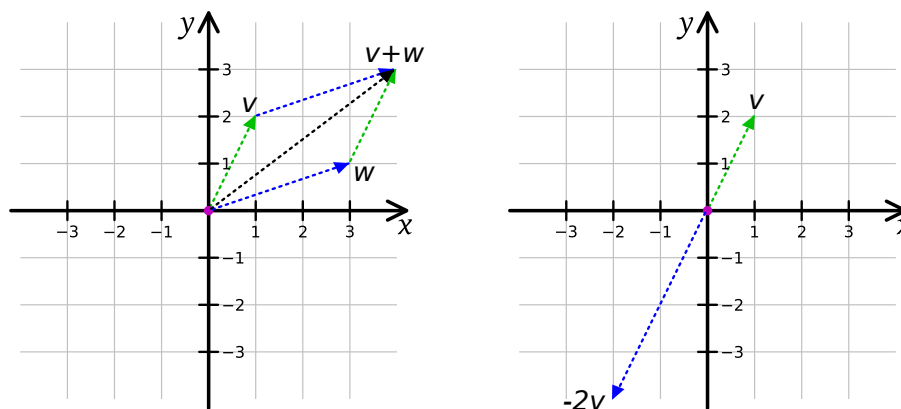


Figura 2.3. La somma $v+w$ di due vettori $v, w \in \mathbb{R}^2$ e la moltiplicazione di un vettore v per uno scalare λ (qui $\lambda = -2$).

innanzitutto una somma fra vettori. Se

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

sono due vettori di \mathbb{R}^n , allora definiamo la loro somma semplicemente facendo la somma delle singole coordinate:

$$x + y = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

Per \mathbb{R}^2 , questa corrisponde precisamente all'usuale somma di vettori con la regola del parallelogramma, come mostrato nella Figura 2.3-(sinistra).

2.1.4. Prodotto per scalare. Un'altra operazione consiste nell'allungare, accorciare o ribaltare un singolo vettore. Dato un vettore x ed un numero reale $\lambda \in \mathbb{R}$, possiamo definire un nuovo vettore

$$\lambda x = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

In \mathbb{R}^2 possiamo visualizzare λx come il vettore ottenuto da x allungandolo o accorciandolo di un fattore $|\lambda|$ e ribaltandolo se $\lambda < 0$. Si veda la Figura 2.3.

Il numero reale λ usato in questo contesto è detto *scalare* e l'operazione che consiste nel moltiplicare un vettore x per uno scalare λ è detta semplicemente *prodotto per scalare*.

2.1.5. Proprietà algebriche. Notiamo che lo spazio euclideo \mathbb{R}^n è un gruppo commutativo con l'operazione di somma appena descritta. L'elemento neutro è il vettore zero

$$0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

che corrisponde all'origine. L'inverso del vettore x è il vettore $-x$ che ha tutte le coordinate cambiate di segno.

La somma ed il prodotto per scalare inoltre soddisfano altre proprietà:

- (1) $\lambda(x + y) = \lambda x + \lambda y$
- (2) $(\lambda + \mu)x = \lambda x + \mu x$
- (3) $(\lambda\mu)x = \lambda(\mu x)$
- (4) $1x = x$

Tutte queste proprietà sono valide per ogni $x, y \in \mathbb{R}^n$ e per ogni $\lambda, \mu \in \mathbb{R}$. Sono tutte conseguenze delle usuali proprietà associative e distributive dei numeri reali.

2.2. Spazi vettoriali

Nella sezione precedente abbiamo iniziato a studiare lo spazio euclideo \mathbb{R}^n , che generalizza il piano cartesiano \mathbb{R}^2 in ogni dimensione. Abbiamo definito la somma fra vettori ed il prodotto per scalare e abbiamo verificato che queste due operazioni soddisfano alcuni assiomi.

In questa sezione daremo un nome a qualsiasi struttura algebrica che soddisfi questi assiomi: lo chiameremo *spazio vettoriale*. Il motivo di questa astrazione è che vi sono in matematica molti altri spazi che soddisfano gli stessi assiomi di \mathbb{R}^n , ed è quindi molto più conveniente considerarli tutti assieme.

2.2.1. Definizione. Fissiamo un campo \mathbb{K} . Questo è generalmente il campo $\mathbb{K} = \mathbb{R}$ dei numeri reali, ma può anche essere ad esempio il campo $\mathbb{K} = \mathbb{C}$ dei numeri complessi oppure il campo $\mathbb{K} = \mathbb{Q}$ dei numeri razionali. Gli elementi di \mathbb{K} sono detti *scalari*.

Uno *spazio vettoriale* su \mathbb{K} è un insieme V di elementi, detti *vettori*, dotato di due operazioni binarie:

- una operazione detta *somma* che associa a due vettori $v, w \in V$ un terzo vettore $v + w \in V$;
- una operazione detta *prodotto per scalare* che associa ad un vettore $v \in V$ e ad uno scalare $\lambda \in \mathbb{K}$ un vettore $\lambda v \in \mathbb{K}$.

Queste due operazioni devono soddisfare gli stessi assiomi che abbiamo verificato precedentemente per lo spazio euclideo, cioè:

- (1) L'insieme V è un gruppo commutativo con la somma $+$
- (2) $\lambda(v + w) = \lambda v + \lambda w$
- (3) $(\lambda + \mu)v = \lambda v + \mu v$

$$(4) (\lambda\mu)v = \lambda(\mu v)$$

$$(5) 1v = v$$

Le proprietà (2)-(5) devono valere per ogni $v, w \in V$ e ogni $\lambda, \mu \in \mathbb{K}$.

L'elemento neutro del gruppo V è indicato con il simbolo 0 ed è detto l'*origine* dello spazio vettoriale V . Non va confuso con lo zero 0 del campo \mathbb{K} . In questo libro il simbolo 0 può indicare varie cose differenti ed il significato preciso sarà sempre chiaro dal contesto.

A partire da questi assiomi è subito possibile dimostrare un piccolo teorema.

Proposizione 2.2.1. *Vale la relazione $0v = 0$.*

Come dicevamo, il simbolo 0 può indicare oggetti diversi: in questo enunciato, il primo 0 è l'elemento neutro di \mathbb{K} ed il secondo è l'origine di V .

Dimostrazione. Vale

$$0v = (0 + 0)v = 0v + 0v$$

e semplificando deduciamo che $0v = 0$. □

2.2.2. Lo spazio \mathbb{K}^n . L'esempio principale di spazio vettoriale su \mathbb{R} è ovviamente lo spazio euclideo \mathbb{R}^n già incontrato precedentemente. Più in generale, è possibile definire per qualsiasi campo \mathbb{K} uno spazio \mathbb{K}^n .

Sia $n \geq 1$ un numero naturale. Lo spazio \mathbb{K}^n è l'insieme delle sequenze (x_1, \dots, x_n) di numeri in \mathbb{K} . Gli elementi $v \in \mathbb{K}^n$ sono descritti generalmente come vettori colonna

$$v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n.$$

La somma e la moltiplicazione per scalare sono definiti termine a termine esattamente come visto sopra per \mathbb{R}^n , cioè:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

Facciamo alcuni esempi con dei vettori di \mathbb{C}^2 :

$$\begin{pmatrix} 1+i \\ -2 \end{pmatrix} + \begin{pmatrix} 3i \\ 1-i \end{pmatrix} = \begin{pmatrix} 1+4i \\ -1-i \end{pmatrix}, \quad (2+i) \begin{pmatrix} 3 \\ 1-i \end{pmatrix} = \begin{pmatrix} 6+3i \\ 3-i \end{pmatrix}.$$

2.2.3. Lo spazio $\mathbb{K}[x]$ dei polinomi. Come abbiamo accennato, ci sono altri oggetti matematici molto comuni che formano degli spazi vettoriali. Ad esempio, i polinomi.

Fissato un campo \mathbb{K} , indichiamo con $\mathbb{K}[x]$ l'insieme di tutti i polinomi aventi coefficienti in \mathbb{K} . Due polinomi possono essere sommati, e moltiplicando un polinomio per uno scalare otteniamo un polinomio. Ad esempio:

$$(x^3 - 2x + 1) + (4x^4 + x - 3) = 4x^4 + x^3 - x - 2, \quad 3(x^3 - 2x) = 3x^3 - 6x.$$

L'insieme $\mathbb{K}[x]$ è quindi naturalmente equipaggiato con le due operazioni di uno spazio vettoriale. Inoltre tutti gli assiomi (1) - (5) sono facilmente verificati. Quindi $\mathbb{K}[x]$ è uno spazio vettoriale su \mathbb{K} .

2.2.4. Lo spazio $F(X, \mathbb{K})$ delle funzioni. Oltre ai polinomi, anche le funzioni formano spesso uno spazio vettoriale.

Sia X un insieme qualsiasi e \mathbb{K} un campo. Consideriamo l'insieme $F(X, \mathbb{K})$ formato da tutte le funzioni $f: X \rightarrow \mathbb{K}$. Vediamo come anche $F(X, \mathbb{K})$ sia naturalmente equipaggiato con le due operazioni di uno spazio vettoriale. Date due funzioni $f, g: X \rightarrow \mathbb{K}$ possiamo definire la loro somma $f + g$ come la funzione che manda x in $f(x) + g(x)$. In altre parole:

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

A sinistra abbiamo scritto la funzione $f + g$ fra parentesi. Notiamo che anche $f + g$ è una funzione da X in \mathbb{K} . Analogamente, se $f: X \rightarrow \mathbb{K}$ è una funzione e $\lambda \in \mathbb{K}$ è uno scalare, possiamo definire una nuova funzione λf che manda x in $\lambda f(x)$. In altre parole:

$$(\lambda f)(x) = \lambda f(x).$$

Anche qui abbiamo scritto la funzione λf fra parentesi.

Esercizio 2.2.2. L'insieme $F(X, \mathbb{K})$ soddisfa gli assiomi (1)-(5) ed è quindi uno spazio vettoriale.

L'elemento neutro di $F(X, \mathbb{K})$ è la *funzione nulla* 0 , quella che fa zero su tutti gli elementi di X , cioè $0(x) = 0 \forall x \in X$.

Ad esempio, l'insieme $F(\mathbb{R}, \mathbb{R})$ di tutte le funzioni da \mathbb{R} in \mathbb{R} forma uno spazio vettoriale. Notiamo il livello di astrazione richiesto in questa sezione: una funzione, che può essere un oggetto abbastanza complicato, è interpretata come un punto oppure un vettore in uno spazio molto grande $F(X, \mathbb{R})$ che contiene tutte le possibili funzioni. Per quanto sembri inutilmente astratto, questo approccio globale è molto utile nello studio delle funzioni che capitano spesso in natura.

2.2.5. Le matrici. Dopo i polinomi e le funzioni, introduciamo un terzo oggetto matematico molto importante che rientra anch'esso nel quadro generale degli spazi vettoriali: le *matrici*.

Sia come sempre \mathbb{K} un campo fissato. Una *matrice* con m righe e n colonne a coefficienti in \mathbb{K} è una tabella rettangolare del tipo

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

in cui tutti gli mn numeri a_{11}, \dots, a_{mn} sono elementi del campo \mathbb{K} . Diciamo brevemente che A è una matrice $m \times n$. La matrice A ha effettivamente m

righe che indicheremo con il simbolo A_1, \dots, A_m , del tipo

$$A_i = (a_{i1} \quad \cdots \quad a_{in})$$

e n *colonne* che indicheremo con A^1, \dots, A^n del tipo

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

Ad esempio, le seguenti sono matrici a coefficienti in \mathbb{R} :

$$\begin{pmatrix} 1 & \sqrt{2} \\ 0 & -5 \\ 7 & \pi \end{pmatrix}, \quad (5 \quad 0 \quad \sqrt{3}).$$

Le matrici giocheranno un ruolo fondamentale in tutto il libro. Per il momento ci limitiamo a notare che due matrici A e B della stessa taglia $m \times n$

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}$$

possono essere sommate e dare quindi luogo ad una nuova matrice $A+B$ così:

$$A+B = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}.$$

Ancora una volta, si tratta semplicemente di sommare separatamente i coefficienti. Analogamente è possibile moltiplicare la matrice A per uno scalare $\lambda \in \mathbb{K}$ e ottenere una nuova matrice

$$\lambda A = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}.$$

Anche qui, abbiamo semplicemente moltiplicato per λ tutti i coefficienti della matrice. Indichiamo con $M(m, n, \mathbb{K})$ l'insieme di tutte le matrici $m \times n$ con coefficienti in \mathbb{K} , con le operazioni di somma e prodotto per scalare appena descritte. Quando il campo è sottinteso usiamo il simbolo $M(m, n)$.

Esercizio 2.2.3. L'insieme $M(m, n, \mathbb{K})$ soddisfa gli assiomi (1)-(5) ed è quindi uno spazio vettoriale.

Notiamo che una matrice $m \times 1$ è semplicemente un vettore colonna e quindi $M(m, 1, \mathbb{K}) = \mathbb{K}^m$. Un vettore colonna è una matrice formata da una sola colonna.

2.2.6. Sottospazio vettoriale. I quattro esempi di spazi vettoriali appena descritti \mathbb{K}^n , $\mathbb{K}[n]$, $F(X, \mathbb{K})$ e $M(m, n, \mathbb{K})$ sono in un certo senso i più importanti perché contengono quasi tutti gli spazi vettoriali che vedremo in questo libro. Definiamo adesso rigorosamente in che senso uno spazio vettoriale può contenerne un altro.

Sia V uno spazio vettoriale su un campo \mathbb{K} . Un *sottospazio vettoriale* di V è un sottoinsieme $W \subset V$ che soddisfa i seguenti tre assiomi:

- (1) $0 \in W$.
- (2) Se $v, v' \in W$, allora anche $v + v' \in W$.
- (3) Se $v \in W$ e $\lambda \in \mathbb{K}$, allora $\lambda v \in W$.

La prima proprietà dice che W deve contenere l'origine; la seconda può essere riassunta dicendo che W deve essere chiuso rispetto alla somma, e la terza asserendo che W deve essere chiuso rispetto al prodotto per scalare.

Il punto fondamentale qui è il seguente: un sottospazio vettoriale $W \subset V$ è esso stesso uno spazio vettoriale. Infatti, la somma ed il prodotto per scalare in W sono ben definiti grazie agli assiomi (2) e (3), e tutti gli assiomi da spazio vettoriale di V si mantengono ovviamente anche per W con la stessa origine 0 di V .

2.2.7. I sottospazi banale e totale. Ogni spazio vettoriale V ha due sottospazi molto particolari che è bene tenere a mente fin da subito:

- Il *sottospazio banale* $W = \{0\}$, formato da un punto solo, l'origine.
- Il *sottospazio totale* $W = V$, formato da tutti i vettori di V .

Tutti gli altri sottospazi W di V hanno una posizione intermedia fra quello banale e quello totale: sono sempre contenuti in V e contengono sempre $\{0\}$. In altre parole:

$$\{0\} \subset W \subset V.$$

A questo punto è relativamente facile descrivere dei sottospazi vettoriali dei quattro spazi modello \mathbb{K}^n , $\mathbb{K}[n]$, $F(X, \mathbb{K})$ e $M(m, n, \mathbb{K})$ descritti precedentemente e ottenere così molti nuovi spazi vettoriali. Questo è quello che faremo nelle prossime sezioni.

2.2.8. Sistemi lineari omogenei. Una *equazione lineare* è una equazione con n variabili x_1, \dots, x_n del tipo

$$a_1x_1 + \dots + a_nx_n = b$$

dove a_1, \dots, a_n, b sono numeri in un certo campo \mathbb{K} . L'equazione è *omogenea* se $b = 0$. Un punto x di \mathbb{K}^n è *soluzione* dell'equazione se le sue coordinate x_1, \dots, x_n la soddisfano, cioè se sostituendole nell'espressione $a_1x_1 + \dots + a_nx_n$ si ottiene proprio b .

Un *sistema lineare* è un insieme di k equazioni lineari in n variabili

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{k1}x_1 + \cdots + a_{kn}x_n = b_k \end{cases}$$

Il sistema è *omogeneo* se $b_1 = \cdots = b_k = 0$. Un punto $x \in \mathbb{K}^n$ è *soluzione* del sistema se è soluzione di ciascuna equazione. Le soluzioni del sistema formano un certo sottoinsieme $S \subset \mathbb{K}^n$.

Proposizione 2.2.4. *Le soluzioni $S \subset \mathbb{K}^n$ di un sistema lineare omogeneo formano un sottospazio vettoriale di \mathbb{K}^n .*

Dimostrazione. Consideriamo un sistema lineare omogeneo

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \vdots \\ a_{k1}x_1 + \cdots + a_{kn}x_n = 0 \end{cases}$$

Dobbiamo verificare i 3 assiomi di sottospazio per l'insieme S delle soluzioni.

(1) Il vettore 0 è soluzione, infatti per ogni equazione troviamo

$$a_{i1}0 + \cdots + a_{in}0 = 0.$$

Quindi $0 \in S$.

(2) Se due vettori x e y sono soluzioni, allora anche $x + y$ è soluzione.

Infatti per ogni equazione troviamo

$$\begin{aligned} a_{i1}(x_1 + y_1) + \cdots + a_{in}(x_n + y_n) &= \\ a_{i1}x_1 + \cdots + a_{in}x_n + a_{i1}y_1 + \cdots + a_{in}y_n &= 0 + 0 = 0. \end{aligned}$$

(3) Se x è soluzione e $\lambda \in \mathbb{K}$ è uno scalare, allora anche λx è soluzione.

Infatti per ogni equazione troviamo

$$\begin{aligned} a_{i1}(\lambda x_1) + \cdots + a_{in}(\lambda x_n) &= \\ \lambda(a_{i1}x_1 + \cdots + a_{in}x_n) &= \lambda 0 = 0. \end{aligned}$$

Quindi S è un sottospazio vettoriale di \mathbb{K}^n . □

Ad esempio, le soluzioni dell'equazione $2x - y = 0$ in \mathbb{R}^2 formano la retta mostrata nella Figura 2.4-(sinistra). Le soluzioni dell'equazione $z = 0$ in \mathbb{R}^3 formano il piano orizzontale mostrato nella Figura 2.1. Le soluzioni del sistema $\{x = 0, z = 0\}$ in \mathbb{R}^3 formano una retta, l'asse y descritto nella stessa figura. Tutti questi sono esempi di sottospazi vettoriali di \mathbb{R}^2 e \mathbb{R}^3 .

Osservazione 2.2.5. Le soluzioni S di un sistema lineare *non* omogeneo non sono un sottospazio vettoriale perché non contengono l'origine. Ad esempio l'equazione $2x - y = -2$ in \mathbb{R}^2 descrive una retta non passante per l'origine: questa non è un sottospazio vettoriale. Si veda la Figura 2.4.

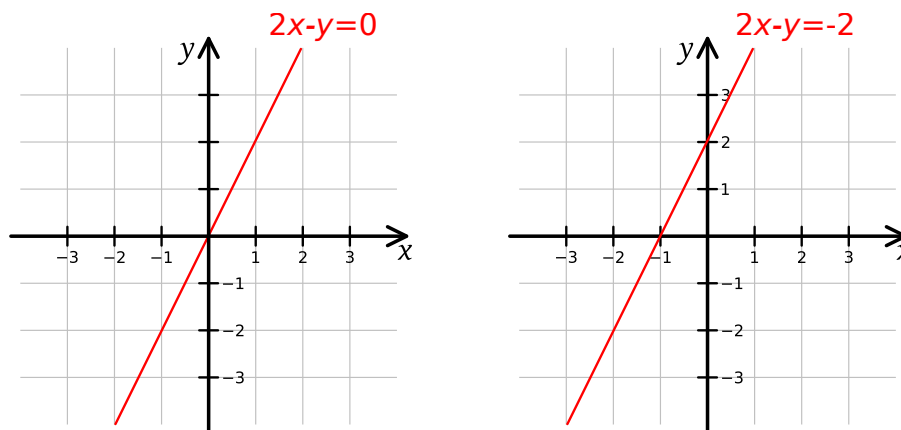


Figura 2.4. La retta $2x - y = 0$ è un sottospazio vettoriale, la retta $2x - y = -2$ no.

2.2.9. Combinazioni lineari. Sia V uno spazio vettoriale qualsiasi. Siano $v_1, \dots, v_k \in V$ dei vettori arbitrari. Una *combinazione lineare* dei vettori v_1, \dots, v_k è un qualsiasi vettore v che si ottiene come

$$v = \lambda_1 v_1 + \dots + \lambda_k v_k$$

dove $\lambda_1, \dots, \lambda_k$ sono scalari arbitrari. Ad esempio, se $V = \mathbb{R}^3$ e

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

allora una combinazione lineare arbitraria di questi due vettori è il vettore

$$\lambda_1 v_1 + \lambda_2 v_2 = \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda_2 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ 0 \end{pmatrix}.$$

Notiamo che questo vettore sta nel piano orizzontale $z = 0$. Osserviamo anche che al variare di λ_1 e λ_2 , facendo tutte le combinazioni lineari di v_1 e v_2 otteniamo precisamente tutti i punti del piano orizzontale $z = 0$. Questo ci porta alla definizione seguente.

2.2.10. Sottospazio generato. Sia V uno spazio vettoriale e $v_1, \dots, v_k \in V$ dei vettori arbitrari. Il *sottospazio generato* da v_1, \dots, v_k è il sottoinsieme di V formato da tutte le combinazioni lineari dei vettori v_1, \dots, v_k . Questo sottospazio viene indicato con il simbolo

$$\text{Span}(v_1, \dots, v_k).$$

In inglese *span* vuol dire proprio generare (in questo contesto). In simboli:

$$\text{Span}(v_1, \dots, v_k) = \{ \lambda_1 v_1 + \dots + \lambda_k v_k \mid \lambda_1, \dots, \lambda_k \in \mathbb{K} \}.$$

Proposizione 2.2.6. *Il sottoinsieme $\text{Span}(v_1, \dots, v_k)$ è un sottospazio vettoriale di V .*

Dimostrazione. Dobbiamo mostrare che $W = \text{Span}(v_1, \dots, v_k)$ soddisfa i 3 assiomi di sottospazio.

(1) $0 \in W$, infatti usando $\lambda_1 = \dots = \lambda_k = 0$ troviamo

$$0v_1 + \dots + 0v_k = 0 + \dots + 0 = 0 \in W.$$

(2) Se $v, w \in W$, allora $v + w \in W$. Infatti, per ipotesi

$$v = \lambda_1 v_1 + \dots + \lambda_k v_k$$

$$w = \mu_1 v_1 + \dots + \mu_k v_k$$

e quindi raccogliendo otteniamo

$$v + w = (\lambda_1 + \mu_1)v_1 + \dots + (\lambda_k + \mu_k)v_k.$$

Anche $v + w$ è combinazione lineare di v_1, \dots, v_k e quindi $v + w \in W$.

(3) Se $v \in W$ e $\lambda \in \mathbb{K}$, allora $\lambda v \in W$. Infatti se

$$v = \lambda_1 v_1 + \dots + \lambda_k v_k$$

allora

$$\lambda v = (\lambda \lambda_1)v_1 + \dots + (\lambda \lambda_k)v_k$$

è anch'esso combinazione lineare di v_1, \dots, v_k .

La dimostrazione è conclusa. \square

Ad esempio, se v è un singolo vettore di V , otteniamo

$$W = \text{Span}(v) = \{\lambda v \mid \lambda \in \mathbb{K}\}.$$

Per esempio se $V = \mathbb{R}^2$ e $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ troviamo che

$$W = \text{Span}(v) = \left\{ t \begin{pmatrix} 1 \\ 2 \end{pmatrix} \mid t \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} t \\ 2t \end{pmatrix} \mid t \in \mathbb{R} \right\}.$$

Notiamo che W è precisamente la retta descritta dall'equazione $y = 2x$ già mostrata nella Figura 2.4-(sinistra).

2.2.11. Forma cartesiana e forma parametrica. Un sottospazio vettoriale di \mathbb{K}^n che sia descritto come luogo di zeri di un sistema di equazioni lineari omogenee è detto in *forma cartesiana*. Un sottospazio vettoriale di \mathbb{K}^n descritto come sottospazio generato da alcuni vettori è invece detto in *forma parametrica*.

Vedremo che qualsiasi sottospazio vettoriale di \mathbb{K}^n può essere descritto in entrambi i modi. Ad esempio, il piano $W = \{z = 0\}$ in \mathbb{R}^3 può essere descritto in forma cartesiana tramite l'equazione $z = 0$, oppure in forma parametrica

come il sottospazio generato dai due vettori v_1 e v_2 indicati precedentemente. Otteniamo in questo modo:

$$W = \text{Span}(v_1, v_2) = \left\{ t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + u \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mid t, u \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} t \\ u \\ 0 \end{pmatrix} \mid t, u \in \mathbb{R} \right\}.$$

La forma parametrica è *esplicita* perché indica chiaramente come sono fatti i punti di W , al variare di alcuni parametri (in questo caso t e u). La forma cartesiana è *implicita* perché descrive W come luogo di soluzioni di una equazione (o di un sistema di equazioni). In molti casi la forma parametrica è preferibile proprio perché esplicita, ma può anche capitare che quella implicita sia più comoda, a seconda del contesto.

Vedremo nelle prossime pagine come passare da una descrizione in forma cartesiana ad una in forma parametrica, e viceversa.

2.2.12. Polinomi con restrizioni. Siano \mathbb{K} un campo e n un numero naturale qualsiasi. Indichiamo con $\mathbb{K}_n[x]$ il sottoinsieme di $\mathbb{K}[x]$ formato da tutti quei polinomi che hanno grado minore o uguale a n . In altre parole:

$$\mathbb{K}_n[x] = \{ a_n x^n + \cdots + a_1 x + a_0 \mid a_0, a_1, \dots, a_n \in \mathbb{K} \}.$$

Questa scrittura può essere interpretata in un altro modo: il sottoinsieme $\mathbb{K}_n[x]$ è il sottospazio generato dai polinomi $x^n, \dots, x, 1$, cioè:

$$\mathbb{K}_n[x] = \text{Span}(x^n, \dots, x, 1).$$

In particolare, ne deduciamo che $\mathbb{K}_n[x]$ è un sottospazio vettoriale di $\mathbb{K}[x]$.

Ci sono altri modi di imporre restrizioni sui polinomi e ottenere così dei sottospazi vettoriali di $\mathbb{K}[x]$. Ad esempio:

Proposizione 2.2.7. *Sia $a \in \mathbb{K}$ un numero fissato. I polinomi $p(x)$ tali che $p(a) = 0$ formano un sottospazio vettoriale di $\mathbb{K}[x]$.*

Dimostrazione. Sia $W \subset \mathbb{K}[x]$ il sottoinsieme formato da tutti i polinomi $p(x)$ tali che $p(a) = 0$. Come al solito, verifichiamo i 3 assiomi di sottospazio:

- (1) $0 \in W$, infatti il polinomio 0 si annulla ovunque e in particolare in a .
- (2) Se $p, q \in W$, allora $p + q \in W$. Infatti se due polinomi p e q si annullano in a , lo fa anche la loro somma:

$$(p + q)(a) = p(a) + q(a) = 0 + 0 = 0.$$

- (3) Se $p \in W$ e $\lambda \in \mathbb{K}$, allora $\lambda p \in W$. Infatti se $p(a) = 0$, allora anche $\lambda p(a) = 0$.

La dimostrazione è conclusa. □

Osservazione 2.2.8. I polinomi $p(x)$ tali che $p(a) = 1$ non formano un sottospazio perché non contengono il polinomio 0.

2.2.13. Funzioni con restrizioni. Sia X un insieme e \mathbb{K} un campo, e $F(X, \mathbb{K})$ lo spazio vettoriale formato da tutte le funzioni da X in \mathbb{K} . Come per i polinomi, è anche qui possibile costruire sottospazi di $F(X, \mathbb{K})$ imponendo delle restrizioni. Possiamo imporre che le funzioni si annullino in un qualsiasi sottoinsieme $Y \subset X$:

Esercizio 2.2.9. Sia $Y \subset X$ un sottoinsieme qualsiasi. Il sottoinsieme

$$W = \{f \in F(X, \mathbb{K}) \mid f(x) = 0 \forall x \in Y\}$$

è un sottospazio vettoriale di $F(X, \mathbb{K})$.

Traccia. Verificare i 3 assiomi. □

Possiamo richiedere che le funzioni soddisfino qualche buona proprietà, come essere continue, derivabili oppure integrabili:

Proposizione 2.2.10. Il sottoinsieme di $F(\mathbb{R}, \mathbb{R})$ formato dalle funzioni continue (oppure derivabili, oppure integrabili) è un sottospazio vettoriale.

Dimostrazione. Dobbiamo verificare i 3 assiomi. È sufficiente notare che (1) la funzione costantemente nulla è continua (e derivabile, integrabile), che (2) la somma di due funzioni continue (oppure derivabili, integrabili) è anch'essa continua (oppure derivabile, integrabile), e infine che (3) moltiplicando una funzione continua (derivabile, integrabile) per uno scalare otteniamo ancora una funzione continua (derivabile, integrabile). Queste sono tutte proprietà dimostrate durante il corso di analisi. □

Abbiamo trovato una catena di sottospazi:

$$\{\text{funzioni derivabili}\} \subset \{\text{funzioni continue}\} \subset \{\text{funz. integrabili}\} \subset F(\mathbb{R}, \mathbb{R}).$$

2.2.14. Matrici diagonali, triangolari, simmetriche e antisimmetriche.

Ricordiamo che $M(m, n, \mathbb{K})$ è lo spazio vettoriale formato dalle matrici $m \times n$ a coefficienti in \mathbb{K} . Come nei paragrafi precedenti, possiamo definire dei sottospazi di $M(m, n, \mathbb{K})$ imponendo opportune restrizioni.

Data una matrice A , indichiamo con A_{ij} oppure a_{ij} i suoi coefficienti. Introduciamo delle definizioni che ci accompagneranno in tutto il libro. Una matrice $n \times n$ è detta *quadrata*.

Definizione 2.2.11. Una matrice A quadrata $n \times n$ è:

- *diagonale* se $a_{ij} = 0 \forall i \neq j$;
- *triangolare superiore* se $a_{ij} = 0 \forall i > j$;
- *triangolare inferiore* se $a_{ij} = 0 \forall i < j$;
- *triangolare* se è triangolare inferiore o superiore;
- *simmetrica* se $a_{ij} = a_{ji} \forall i, j$;
- *antisimmetrica* se $a_{ij} = -a_{ji} \forall i, j$.

Ad esempio, le matrici seguenti sono (in ordine) diagonale, triangolare superiore, triangolare inferiore, simmetrica e antisimmetrica:

$$\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 9 \\ 0 & \sqrt{2} \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 7 & 2 \end{pmatrix}, \quad \begin{pmatrix} -1 & 2 \\ 2 & 4 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Si faccia attenzione alle definizioni: la matrice

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

è al tempo stesso diagonale, triangolare superiore, triangolare inferiore e simmetrica. La matrice nulla 0 è di tutti e 5 i tipi: è diagonale, triangolare superiore, triangolare inferiore, simmetrica e antisimmetrica.

Gli elementi $a_{11}, a_{22}, \dots, a_{nn}$ in una matrice quadrata A formano la *diagonale principale* di A . Notiamo che una matrice antisimmetrica ha necessariamente elementi nulli sulla diagonale principale, perché $a_{ii} = -a_{ii}$ implica $a_{ii} = 0$.

Indichiamo lo spazio $M(n, n, \mathbb{K})$ delle matrici quadrate $n \times n$ con $M(n, \mathbb{K})$ oppure ancora più semplicemente con $M(n)$. Indichiamo con

$$D(n), \quad T^s(n), \quad T^i(n), \quad S(n), \quad A(n)$$

rispettivamente i sottoinsiemi di $M(n)$ formati dalle matrici diagonali, triangolari superiori, triangolari inferiori, simmetriche e antisimmetriche.

Proposizione 2.2.12. *I sottoinsiemi $D(n), T^s(n), T^i(n), S(n), A(n)$ sono tutti sottospazi vettoriali di $M(n)$.*

Dimostrazione. Mostriamo l'enunciato per $S(n)$ e lasciamo gli altri casi per esercizio. I 3 assiomi del sottospazio sono verificati, perché

- (1) la matrice nulla è simmetrica,
- (2) la somma $A + B$ di due matrici simmetriche A e B è anch'essa simmetrica, infatti

$$(A + B)_{ij} = A_{ij} + B_{ij} = A_{ji} + B_{ji} = (A + B)_{ji}.$$

- (3) analogamente si verifica che se A è simmetrica e $\lambda \in \mathbb{K}$ allora λA è simmetrica.

La dimostrazione è conclusa. □

2.2.15. Intersezione di sottospazi. Sia V uno spazio vettoriale e $U, W \subset V$ due sottospazi. Da un punto di vista insiemistico è naturale considerare la loro intersezione $U \cap W$.

Proposizione 2.2.13. *L'intersezione $U \cap W$ di due sottospazi vettoriali $U, W \subset V$ è sempre un sottospazio vettoriale.*

Dimostrazione. Come sempre, verifichiamo i 3 assiomi.

- (1) $0 \in U \cap W$, perché 0 appartiene sia a U che a W .

(2) $v, w \in U \cap W \implies v + w \in U \cap W$. Infatti $v, w \in U$ implica $v + w \in U$ e $v, w \in W$ implica $v + w \in W$, quindi $v + w \in U \cap W$.

(3) $v \in U \cap W \implies \lambda v \in U \cap W \forall \lambda \in \mathbb{K}$ è analogo al precedente.

La dimostrazione è conclusa. \square

Facciamo alcuni esempi.

Esempio 2.2.14. Il piani $U = \{x = 0\}$ e $W = \{y = 0\}$ in \mathbb{R}^3 si intersecano nella retta verticale $U \cap W$, che può essere descritta in forma cartesiana come insieme delle soluzioni di questo sistema:

$$\begin{cases} x = 0 \\ y = 0 \end{cases}$$

oppure esplicitamente in forma parametrica come

$$U \cap W = \left\{ \begin{pmatrix} 0 \\ 0 \\ t \end{pmatrix} \mid t \in \mathbb{R} \right\}.$$

Più in generale, se U e W sono sottospazi di \mathbb{K}^n descritti in forma cartesiana come soluzioni di sistemi lineari omogenei, l'intersezione $U \cap W$ è descritta sempre in forma cartesiana unendo i sistemi lineari in uno solo. Questo fatto è conseguenza di un principio insiemistico più generale:

*Affiancare equazioni in un sistema
corrisponde a fare l'intersezione delle soluzioni.*

Esempio 2.2.15. Fra i sottospazi di $M(n)$ valgono le seguenti relazioni:

$$D(n) = T^s(n) \cap T^i(n), \quad \{0\} = S(n) \cap A(n).$$

La seconda relazione dice che una matrice A che è contemporaneamente simmetrica e antisimmetrica è necessariamente nulla. Infatti $a_{ij} = a_{ij}$ e $a_{ij} = -a_{ij}$ insieme implicano $a_{ij} = 0$, per ogni i, j .

2.2.16. L'unione non funziona. Dopo aver considerato l'intersezione $U \cap W$ di due sottospazi $U, W \subset V$, è naturale adesso considerare la loro unione $U \cup W$. Ci accorgiamo però immediatamente che l'unione $U \cup W$ di due sottospazi molto spesso *non* è un sottospazio, come mostra l'esempio seguente.

Esempio 2.2.16. Se $V = \mathbb{R}^2$ e $U = \{x = 0\}$ e $W = \{y = 0\}$ sono le due rette che definiscono gli assi di \mathbb{R}^2 , la loro unione $U \cup W$ non è certamente un sottospazio perché non è soddisfatto l'assioma (2): ad esempio, se prendiamo $u = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ e $w = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, vediamo che $u, w \in U \cup W$ ma la loro somma $u + w = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ non appartiene a $U + W$.

Esercizio 2.2.17. Siano $U, W \subset V$ sottospazi vettoriali. Mostra che $U \cup W$ è un sottospazio se e solo se $U \subset W$ oppure $W \subset U$.

Abbandoniamo quindi questa strada e ci accingiamo a definire un'altra operazione che sarà molto più utile in seguito.

2.2.17. Somma di sottospazi. Siano come sopra $U, W \subset V$ due sottospazi di uno spazio vettoriale V .

Definizione 2.2.18. La *somma* $U+W$ dei due sottospazi U e W è l'insieme

$$U+W = \{u+w \mid u \in U, w \in W\} \subset V.$$

La somma $U+W$ è l'insieme di tutti i vettori di V che possono essere scritti come somma $u+w$ di un vettore $u \in U$ e di un altro $w \in W$.

Proposizione 2.2.19. *L'insieme $U+W$ contiene sia U che W .*

Dimostrazione. Prendendo $u \in U$ e $0 \in W$ otteniamo $u = u+0 \in U+W$ e analogamente facciamo scrivendo $w \in W$ come $0+w \in U+W$. \square

Proposizione 2.2.20. *L'insieme $U+W$ è un sottospazio vettoriale di V .*

Dimostrazione. Come sempre, verifichiamo i 3 assiomi.

- (1) $0 \in U+W$ perché $0 \in U, 0 \in W$ e scriviamo $0 = 0+0$.
- (2) $v, v' \in U+W \implies v+v' \in U+W$. Infatti per definizione $v = u+w$ e $v' = u'+w'$ con $u, u' \in U$ e $w, w' \in W$. Ne segue che

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

con $u+u' \in U$ e $w+w' \in W$, quindi anche $v+v' \in U+W$.

- (3) $v \in U+W, \lambda \in \mathbb{K} \implies \lambda v \in U+W$. Analogo al precedente.

La dimostrazione è conclusa. \square

Prima di fornire esempi concreti, facciamo una considerazione importante che ci aiuta a capire meglio come funziona la somma fra sottospazi.

Proposizione 2.2.21. *Siano v_1, \dots, v_h e w_1, \dots, w_k vettori di uno spazio vettoriale V . Se*

$$U = \text{Span}(v_1, \dots, v_h), \quad W = \text{Span}(w_1, \dots, w_k)$$

allora

$$U+W = \text{Span}(v_1, \dots, v_h, w_1, \dots, w_k).$$

Dimostrazione. L'insieme $U+W$ è formato dagli $u+w$ dove $u \in U$ e $w \in W$. Per ipotesi due generici $u \in U$ e $w \in W$ si scrivono come

$$u = \lambda_1 v_1 + \dots + \lambda_h v_h, \quad w = \mu_1 w_1 + \dots + \mu_k w_k$$

e quindi

$$u+w = \lambda_1 v_1 + \dots + \lambda_h v_h + \mu_1 w_1 + \dots + \mu_k w_k.$$

Abbiamo dimostrato in questo modo che $u+w$ è sempre combinazione lineare dei vettori $v_1, \dots, v_h, w_1, \dots, w_k$. Abbiamo quindi mostrato l'inclusione

$$U+W \subset \text{Span}(v_1, \dots, v_h, w_1, \dots, w_k).$$

Mostriamo adesso l'inclusione opposta. Ogni vettore v che sia combinazione lineare dei vettori $v_1, \dots, v_h, w_1, \dots, w_k$ si scrive come

$$v = \lambda_1 v_1 + \dots + \lambda_h v_h + \mu_1 w_1 + \dots + \mu_k w_k.$$

Definiamo

$$u = \lambda_1 v_1 + \dots + \lambda_h v_h, \quad w = \mu_1 w_1 + \dots + \mu_k w_k$$

e troviamo che $v = u + w$. Quindi abbiamo anche

$$U + W \supset \text{Span}(v_1, \dots, v_h, w_1, \dots, w_k).$$

I due insiemi sono uguali. □

Diciamo che dei vettori $v_1, \dots, v_k \in W$ sono dei *generatori* di W se

$$W = \text{Span}(v_1, \dots, v_k).$$

Abbiamo visto precedentemente che affiancare equazioni in un sistema corrisponde a fare l'intersezione dei sottospazi. La proposizione che abbiamo appena dimostrato può essere riassunta in modo analogo.

Affiancare dei generatori corrisponde a fare la somma di sottospazi.

Esempio 2.2.22. Consideriamo i vettori in \mathbb{R}^3 :

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Gli spazi $U = \text{Span}(v_1)$ e $W = \text{Span}(v_2)$ sono due rette, corrispondenti agli assi x e y . Lo spazio somma $U + W = \text{Span}(v_1, v_2)$ è il piano orizzontale che li contiene, descritto dall'equazione $z = 0$.

Osservazione 2.2.23. Nella dimostrazione della Proposizione 2.2.21, per mostrare che due insiemi A e B sono uguali, abbiamo verificato entrambe le inclusioni $A \subset B$ e $B \subset A$. Questa è una tecnica ricorrente in matematica.

2.3. Dimensione

Sappiamo dalla scuola che un punto, una retta e un piano hanno dimensione 0, 1 e 2. In questa sezione definiamo la nozione di dimensione di uno spazio vettoriale in modo rigoroso.

2.3.1. (In)dipendenza lineare. Sia V uno spazio vettoriale su \mathbb{K} e siano $v_1, \dots, v_k \in V$ alcuni vettori. Diciamo che questi vettori sono *linearmente dipendenti* se esistono dei coefficienti $\lambda_1, \dots, \lambda_k \in \mathbb{K}$, non tutti nulli, tali che

$$\lambda_1 v_1 + \dots + \lambda_k v_k = 0.$$

Se i vettori v_1, \dots, v_k sono linearmente dipendenti, allora è possibile esprimere uno di loro in funzione degli altri. Infatti, per ipotesi esiste almeno un $\lambda_i \neq 0$ e dopo aver diviso tutto per λ_i e spostato gli addendi otteniamo

$$v_i = -\frac{\lambda_1}{\lambda_i} v_1 - \dots - \frac{\lambda_k}{\lambda_i} v_k$$

dove nel membro di destra ovviamente non compare v_i . Quindi:

Proposizione 2.3.1. *I vettori v_1, \dots, v_k sono dipendenti \iff uno di loro è esprimibile come combinazione lineare degli altri.*

I vettori v_1, \dots, v_k sono *linearmente indipendenti* se non sono linearmente dipendenti. Questa importante condizione può essere espressa nel modo seguente: dei vettori v_1, \dots, v_k sono linearmente indipendenti se e solo se

$$\lambda_1 v_1 + \dots + \lambda_k v_k = 0 \implies \lambda_1 = \dots = \lambda_k = 0.$$

In altre parole, l'unica combinazione lineare dei v_1, \dots, v_k che può dare il vettore nullo è quella banale in cui tutti i coefficienti $\lambda_1, \dots, \lambda_k$ sono nulli.

I casi $k = 1$ e $k = 2$ sono facili da studiare. Da quanto abbiamo detto, segue facilmente che:

- un vettore v_1 è dipendente $\iff v_1 = 0$;
- due vettori v_1, v_2 sono dipendenti \iff sono multipli, cioè se esiste un $k \in \mathbb{K}$ tale che $v_1 = k v_2$ oppure $v_2 = k v_1$.

I vettori $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ e $v_2 = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$ di \mathbb{R}^2 sono dipendenti; $w_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ e $w_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ sono indipendenti perché non sono multipli.

Il caso in cui ci sono tre (o più) vettori v_1, v_2, v_3 è più complesso. I vettori

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

di \mathbb{R}^3 sono dipendenti, perché $v_1 - v_2 - v_3 = 0$. Notiamo che a coppie i tre vettori sono sempre indipendenti (non sono mai multipli), ma tutti e tre non lo sono, e questo non è certamente un fatto che si vede immediatamente come nel caso $k = 2$. D'altra parte, i vettori

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

sono indipendenti: se una combinazione lineare produce il vettore nullo

$$\lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 e_3 = 0$$

allora riscriviamo entrambi i membri come vettori e scopriamo che

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Da questo deduciamo che $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Quindi l'unica combinazione lineare di e_1, e_2, e_3 che dà il vettore nullo è quella banale. Quindi i tre vettori e_1, e_2, e_3 sono indipendenti.

Osservazione 2.3.2. Nell'esempio fatto precedentemente, ciascuno dei tre vettori dipendenti v_1, v_2, v_3 può essere scritto come combinazione degli altri due, infatti $v_1 = v_2 + v_3$, oppure $v_2 = v_1 - v_3$, oppure $v_3 = v_1 - v_2$. Possono capitare anche dei vettori v_1, \dots, v_k dipendenti in cui solo alcuni dei k vettori possono essere combinazioni di altri. La costruzione di questi esempi (in \mathbb{R}^3 o in qualche altro spazio vettoriale) è lasciata per esercizio.

Proposizione 2.3.3. *Se v_1, \dots, v_k sono indipendenti, allora qualsiasi sottoinsieme di $\{v_1, \dots, v_k\}$ è anch'esso formato da vettori indipendenti.*

Dimostrazione. Supponiamo per assurdo che alcuni dei v_1, \dots, v_k siano fra loro dipendenti. Per comodità di notazione supponiamo siano i primi h , con $h < k$. Allora esiste una combinazione lineare non banale nulla:

$$\lambda_1 v_1 + \dots + \lambda_h v_h = 0.$$

Questa si estende ad una combinazione lineare non banale nulla di tutti:

$$\lambda_1 v_1 + \dots + \lambda_h v_h + 0v_{h+1} + \dots + 0v_k = 0.$$

E quindi anche i vettori v_1, \dots, v_k sarebbero dipendenti, un assurdo. \square

In particolare, ne deduciamo che se v_1, \dots, v_k sono indipendenti allora:

- i vettori v_i sono tutti diversi da zero;
- i vettori v_i sono a coppie non multipli.

Attenzione però: l'esempio sopra con $v_1, v_2, v_3 \in \mathbb{R}^3$ mostra che queste due condizioni sono necessarie ma non sufficienti affinché i vettori v_1, \dots, v_k siano indipendenti quando $k \geq 3$.

Esercizio 2.3.4. Considera i vettori seguenti di \mathbb{R}^3 :

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 5 \\ 4 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Mostra che v_1, v_2, v_3 sono dipendenti e v_1, v_2, v_4 indipendenti.

2.3.2. Basi. Introduciamo adesso una delle definizioni più importanti del libro. Sia V uno spazio vettoriale.

Definizione 2.3.5. Una sequenza $v_1, \dots, v_n \in V$ di vettori è una *base* se sono soddisfatte entrambe queste condizioni:

- (1) i vettori v_1, \dots, v_n sono indipendenti;
- (2) i vettori v_1, \dots, v_n generano V .

Ricordiamo che la seconda condizione vuol dire che $V = \text{Span}(v_1, \dots, v_n)$, cioè qualsiasi vettore di V è esprimibile come combinazione lineare dei vettori v_1, \dots, v_n . Facciamo adesso alcuni esempi fondamentali.

Proposizione 2.3.6. *Gli elementi*

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

formano una base di \mathbb{K}^n detta base canonica.

Dimostrazione. Mostriamo che i vettori e_1, \dots, e_n sono indipendenti. Supponiamo di avere una combinazione lineare nulla

$$\lambda_1 e_1 + \dots + \lambda_n e_n = 0.$$

Questa tradotta in vettori diventa

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Ne deduciamo che $\lambda_1 = \dots = \lambda_n = 0$ e quindi i v_1, \dots, v_n sono indipendenti.

Mostriamo che i vettori e_1, \dots, e_n generano \mathbb{K}^n . Un generico vettore $x \in \mathbb{K}^n$ si può effettivamente scrivere come combinazione lineare di e_1, \dots, e_n nel modo seguente:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = x_1 e_1 + x_2 e_2 + \dots + x_n e_n.$$

La dimostrazione è conclusa. \square

Ricordiamo lo spazio $\mathbb{K}_n[x]$ dei polinomi di grado minore o uguale a n .

Proposizione 2.3.7. *Gli elementi $1, x, x^2, \dots, x^n$ formano una base di $\mathbb{K}_n[x]$, detta base canonica.*

Dimostrazione. Lo spirito della dimostrazione è lo stesso di quella precedente. Dimostriamo che i vettori $1, x, \dots, x^n$ sono indipendenti: se

$$\lambda_0 \cdot 1 + \lambda_1 x + \dots + \lambda_n x^n = 0$$

allora chiaramente $\lambda_0 = \dots = \lambda_n = 0$. D'altro canto, ciascun polinomio di grado minore o uguale a n si scrive come

$$p(x) = a_n x^n + \dots + a_1 x + a_0$$

e questa scrittura è già una combinazione lineare dei vettori $1, x^1, \dots, x^n$ con coefficienti a_0, a_1, \dots, a_n . \square

Ricordiamo anche lo spazio $M(m, n, \mathbb{K})$ delle matrici $m \times n$ a coefficienti in \mathbb{K} . Per ogni $1 \leq i \leq m$ e $1 \leq j \leq n$, indichiamo con e_{ij} la matrice $m \times n$ che ha tutti zeri ovunque, tranne un 1 nella casella di riga i e colonna j . Ad esempio, nel caso delle matrici quadrate 2×2 troviamo:

$$e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad e_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad e_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Proposizione 2.3.8. *Le matrici e_{ij} con $1 \leq i \leq m$ e $1 \leq j \leq n$ formano una base di $M(m, n, \mathbb{K})$, detta base canonica*

Dimostrazione. La dimostrazione è del tutto analoga a quella per \mathbb{K}^n , solo con la complicazione notazionale che abbiamo due indici i, j invece che uno solo. Mostriamo che le matrici e_{ij} sono indipendenti. Supponiamo di avere una combinazione lineare nulla

$$\sum_{i,j} \lambda_{ij} e_{ij} = 0.$$

Qui sommiamo su $1 \leq i \leq m$ e $1 \leq j \leq n$, quindi ci sono mn addendi. Esplicitando le matrici, l'equazione diventa

$$\begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1n} \\ \vdots & \ddots & \vdots \\ \lambda_{m1} & \cdots & \lambda_{mn} \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

Da questa ricaviamo che $\lambda_{ij} = 0$ per ogni i, j . Mostriamo che le matrici e_{ij} generano: qualsiasi matrice A si scrive come

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \sum_{i,j} a_{ij} e_{ij}.$$

La dimostrazione è conclusa. \square

Abbiamo descritto delle basi *canoniche* per gli spazi \mathbb{K}^n , $\mathbb{K}_n[x]$ e $M(m, n, \mathbb{K})$. Queste però non sono certamente le uniche basi disponibili: uno spazio vettoriale ha generalmente una infinità di basi diverse e come vedremo successivamente sarà spesso importante scegliere la base giusta per risolvere i problemi che troveremo.

Esercizio 2.3.9. I vettori $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ e $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ formano una base di \mathbb{R}^2 .

2.3.3. Coordinate di un vettore rispetto ad una base. A cosa servono le basi? Servono per dare un nome a tutti i vettori dello spazio. A questo scopo useremo la proposizione seguente.

Proposizione 2.3.10. *Sia V uno spazio vettoriale e sia v_1, \dots, v_n una base di V . Ogni vettore $v \in V$ si scrive in modo unico come*

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

Dimostrazione. Sappiamo che i v_1, \dots, v_n generano, quindi sicuramente v si può scrivere come combinazione lineare dei v_1, \dots, v_n . Supponiamo che lo si possa fare in due modi diversi:

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n = \mu_1 v_1 + \dots + \mu_n v_n.$$

Spostando tutto a sinistra otteniamo

$$(\lambda_1 - \mu_1)v_1 + \dots + (\lambda_n - \mu_n)v_n = 0.$$

Siccome i vettori sono indipendenti, necessariamente $\mu_i = \lambda_i$ per ogni i . \square

I coefficienti $\lambda_1, \dots, \lambda_n$ sono le *coordinate* di v rispetto alla base v_1, \dots, v_n .

Esempio 2.3.11. Le coordinate di un vettore $x \in \mathbb{K}^n$ rispetto alla base canonica e_1, \dots, e_n di \mathbb{K}^n sono proprio i suoi coefficienti x_1, \dots, x_n , visto che

$$x = x_1 e_1 + \dots + x_n e_n.$$

Se cambiamo base questo ovviamente non è più vero. Ad esempio, se prendiamo $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ come base di \mathbb{R}^2 , il vettore $w = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ si scrive come $w = v_1 - v_2$ e quindi le sue coordinate rispetto alla base v_1, v_2 sono $1, -1$.

Esempio 2.3.12. Le coordinate di un polinomio $p(x) = a_n x^n + \dots + a_1 x + a_0$ rispetto alla base canonica $1, x, \dots, x^n$ sono a_0, a_1, \dots, a_n .

Esempio 2.3.13. Le coordinate di una matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ rispetto alla base canonica $e_{11}, e_{12}, e_{21}, e_{22}$ di $M(2, \mathbb{K})$ sono a, b, c, d .

Esempio 2.3.14. I vettori $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ sono una base di \mathbb{R}^2 per l'Esercizio 2.3.9. Descriviamo adesso un procedimento che ci permette di capire quali siano le coordinate di un vettore generico $\begin{pmatrix} x \\ y \end{pmatrix}$ rispetto a questa base. Scriviamo

$$\begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} -1 \\ 1 \end{pmatrix} + u \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

e troviamo i parametri t, u in funzione di x, y . Viene un sistema lineare

$$\begin{cases} -t + 2u = x \\ t + u = y \end{cases}$$

Nel prossimo capitolo descriveremo degli algoritmi per risolvere qualsiasi sistema lineare. Questo è abbastanza semplice e troviamo la soluzione

$$t = \frac{-x + 2y}{3}, \quad u = \frac{x + y}{3}.$$

Quindi le coordinate di $\begin{pmatrix} x \\ y \end{pmatrix}$ rispetto alla base sono $\frac{-x+2y}{3}$ e $\frac{x+y}{3}$.

2.3.4. Dimensione. Lo scopo principale di questa sezione è dimostrare questo teorema.

Teorema 2.3.15. Due basi dello stesso spazio vettoriale V contengono lo stesso numero n di elementi.

Questo teorema è fondamentale, perché ci permette di definire in modo rigoroso un concetto intuitivo, quello di dimensione.

Definizione 2.3.16. Se uno spazio vettoriale V ha una base v_1, \dots, v_n , diciamo che V ha *dimensione* n . Se V non ha una base allora diciamo che ha dimensione ∞ .

La definizione è ben posta proprio grazie al Teorema 2.3.15: lo spazio vettoriale V ha tante basi, ma hanno tutte lo stesso numero n di elementi. La dimostrazione di questo teorema si basa su questo lemma, che avrà anche altre conseguenze interessanti. Sia V uno spazio vettoriale su \mathbb{K} .

Lemma 2.3.17. Supponiamo di avere:

- dei vettori v_1, \dots, v_n che generano V , e
- dei vettori w_1, \dots, w_n indipendenti.

Allora anche i vettori w_1, \dots, w_n generano V .

Dimostrazione. Lo scopo è mostrare che $V = \text{Span}(w_1, \dots, w_n)$. Per fare ciò dimostriamo l'affermazione seguente per ogni $s = 0, 1, \dots, n$:

$$(1) \quad \exists i_{s+1} < \dots < i_n \text{ tali che } V = \text{Span}(w_1, \dots, w_s, v_{i_{s+1}}, \dots, v_{i_n}).$$

Notiamo che se $s = n$ allora abbiamo la tesi $V = \text{Span}(w_1, \dots, w_n)$ e siamo a posto. Dimostreremo questa asserzione per induzione su s . L'affermazione è vera per $s = 0$ perché $V = \text{Span}(v_1, \dots, v_n)$. Supponiamo che sia vera per $s - 1$ e la dimostriamo per s . Sappiamo per ipotesi induttiva che

$$V = \text{Span}(w_1, \dots, w_{s-1}, v_s, \dots, v_n).$$

Per semplicità notazionale, a meno di riordinare i vettori v_1, \dots, v_n scriviamo

$$V = \text{Span}(w_1, \dots, w_{s-1}, v_s, \dots, v_n).$$

In particolare possiamo ottenere w_s come combinazione lineare:

$$w_s = \lambda_1 w_1 + \dots + \lambda_{s-1} w_{s-1} + \lambda_s v_s + \dots + \lambda_n v_n.$$

Notiamo che esiste almeno un $i \geq s$ con $\lambda_i \neq 0$. Se così non fosse, questa sarebbe una relazione di dipendenza fra i soli w_1, \dots, w_s , ma questo è escluso perché sono indipendenti. A meno di riordinare i vettori supponiamo che $\lambda_s \neq 0$. Possiamo dividere tutto per λ_s e spostare gli addendi in modo da ottenere:

$$v_s = \frac{1}{\lambda_s} w_s - \frac{\lambda_1}{\lambda_s} w_1 - \dots - \frac{\lambda_{s-1}}{\lambda_s} w_{s-1} - \frac{\lambda_{s+1}}{\lambda_s} v_{s+1} - \dots - \frac{\lambda_n}{\lambda_s} v_n.$$

Quindi $v_s \in \text{Span}(w_1, \dots, w_s, v_{s+1}, \dots, v_n)$. Questo implica facilmente che

$$V = \text{Span}(w_1, \dots, w_{s-1}, v_s, \dots, v_n) \subset \text{Span}(w_1, \dots, w_s, v_{s+1}, \dots, v_n).$$

Chiaramente $\text{Span}(w_1, \dots, w_s, v_{s+1}, \dots, v_n) \subset V$ e quindi

$$V = \text{Span}(w_1, \dots, w_s, v_{s+1}, \dots, v_n).$$

La dimostrazione per induzione è conclusa. \square

Possiamo adesso dedurre facilmente il Teorema 2.3.15.

Dimostrazione. Supponiamo per assurdo che uno spazio vettoriale V contenga due basi

$$v_1, \dots, v_n, \quad w_1, \dots, w_m$$

con $n \neq m$. Sia ad esempio $n < m$.

Per ipotesi i vettori v_1, \dots, v_n generano V e i vettori w_1, \dots, w_n sono indipendenti. Dal lemma precedente deduciamo che w_1, \dots, w_n generano V . Quindi w_{n+1} può essere espresso come combinazione lineare dei w_1, \dots, w_n , e quindi i vettori w_1, \dots, w_m non sono indipendenti: assurdo. \square

Indichiamo la dimensione di uno spazio vettoriale V con il simbolo $\dim V$. Possiamo calcolare la dimensione di tutti gli spazi vettoriali dei quali abbiamo esibito una base. Troviamo:

$$\begin{aligned} \dim \mathbb{K}^n &= n, \\ \dim \mathbb{K}_n[x] &= n + 1, \\ \dim M(m, n, \mathbb{K}) &= mn. \end{aligned}$$

Ricordiamo che $\dim V = \infty$ se V non ha una base.

Proposizione 2.3.18. *Lo spazio $\mathbb{K}[x]$ dei polinomi ha dimensione infinita.*

Dimostrazione. Supponiamo per assurdo che lo spazio $\mathbb{K}[x]$ abbia una base $p_1(x), \dots, p_n(x)$. Sia N il massimo dei gradi dei polinomi $p_1(x), \dots, p_n(x)$. È chiaro che il polinomio x^{N+1} non può essere ottenuto come combinazione lineare dei $p_1(x), \dots, p_n(x)$, e questo è un assurdo. \square

Notiamo infine che per definizione uno spazio vettoriale V ha dimensione 0 se e solo se consiste solo dell'origine $V = \{0\}$.

2.3.5. Algoritmo di completamento. Abbiamo definito la dimensione di uno spazio vettoriale V come il numero di elementi in una sua base. Adesso abbiamo bisogno di algoritmi concreti per determinare una base di V in vari contesti. Iniziamo con una proposizione utile.

Proposizione 2.3.19. *Sia V uno spazio vettoriale e $v_1, \dots, v_k \in V$ dei vettori indipendenti. Sia $v_{k+1} \in V$. Allora:*

$$v_1, \dots, v_{k+1} \text{ sono indipendenti} \iff v_{k+1} \notin \text{Span}(v_1, \dots, v_k).$$

Dimostrazione. (\Rightarrow) Se per assurdo $v_{k+1} \in \text{Span}(v_1, \dots, v_k)$, allora otteniamo $v_{k+1} = \lambda_1 v_1 + \dots + \lambda_k v_k$ e quindi i vettori v_1, \dots, v_{k+1} sono dipendenti.

(\Leftarrow) Supponiamo che

$$\lambda_1 v_1 + \dots + \lambda_{k+1} v_{k+1} = 0.$$

Se $\lambda_{k+1} = 0$, otteniamo una combinazione lineare nulla dei primi v_1, \dots, v_k , quindi $\lambda_1 = \dots = \lambda_k = 0$. Se $\lambda_{k+1} \neq 0$, dividiamo tutto per λ_{k+1} ottenendo una dipendenza di v_{k+1} dai precedenti, che è assurdo. \square

Descriviamo adesso l'*algoritmo di completamento a base*. Sia V uno spazio vettoriale di dimensione n . Siano $v_1, \dots, v_k \in V$ dei vettori indipendenti qualsiasi, con $k < n$. Possiamo sempre completare v_1, \dots, v_k ad una base di V nel modo seguente. Siccome $k < n$, i vettori v_1, \dots, v_k non sono una base di V , in particolare non generano V . Allora esiste un $v_{k+1} \in V$ tale che

$$v_{k+1} \notin \text{Span}(v_1, \dots, v_k).$$

Aggiungiamo v_{k+1} alla lista, che adesso diventa v_1, \dots, v_{k+1} . La nuova lista è ancora formata da vettori indipendenti per la Proposizione 2.3.19. Continuiamo finché non otteniamo n vettori indipendenti v_1, \dots, v_n . Questi devono essere una base per il Lemma 2.3.17.

Esempio 2.3.20. Consideriamo i vettori indipendenti in \mathbb{R}^3

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

Per completare la coppia v_1, v_2 , è sufficiente aggiungere un terzo vettore qualsiasi che non sia contenuto nel piano $\text{Span}(v_1, v_2)$. Ad esempio, $v_3 = e_1$ funziona (ricordiamo che e_1, e_2, e_3 sono i vettori della base canonica di \mathbb{R}^3). Quindi v_1, v_2, e_1 è una base di \mathbb{R}^3 .

2.3.6. Algoritmo di estrazione. Descriviamo adesso l'*algoritmo di estrazione di una base*. Sia V uno spazio vettoriale e siano v_1, \dots, v_m dei generatori di V . Vogliamo estrarre da questo insieme di generatori una base per V .

L'algoritmo funziona nel modo seguente. Se i generatori v_1, \dots, v_m sono indipendenti, allora formano già una base per V e siamo a posto. Altrimenti, esiste almeno un v_i che può essere espresso come combinazione degli altri. Se lo rimuoviamo dalla lista, otteniamo una lista di $m - 1$ vettori che continuano a generare V . Dopo un numero finito di passi otteniamo una lista di vettori indipendenti, e questi sono una base di V .

Esempio 2.3.21. Consideriamo i vettori di \mathbb{R}^3

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Si verifica facilmente che questi generano \mathbb{R}^3 , ma non sono indipendenti, ad esempio $v_1 - 2v_2 + v_3 + v_4 = 0$. Ciascuno dei quattro vettori è esprimibile come combinazione degli altri tre: rimuovendo uno qualsiasi di questi quattro vettori otteniamo una base di \mathbb{R}^3 .

Concludiamo con una proposizione che risulta utile in molti casi concreti quando dobbiamo mostrare che un certo insieme di vettori è una base. In generale, per dimostrare che n vettori sono una base di V , dobbiamo dimostrare che generano e che sono indipendenti. La proposizione seguente dice che se sappiamo già che $\dim V = n$, allora una qualsiasi delle due proprietà è sufficiente, perché implica l'altra.

Proposizione 2.3.22. *Siano v_1, \dots, v_n vettori di uno spazio V di dimensione n . I fatti seguenti sono tutti equivalenti:*

- (1) v_1, \dots, v_n generano V ;
- (2) v_1, \dots, v_n sono indipendenti;
- (3) v_1, \dots, v_n sono una base per V .

Dimostrazione. (3) \Rightarrow (2) è ovvio. (2) \Rightarrow (1) è il Lemma 2.3.17.

(1) \Rightarrow (3). Se v_1, \dots, v_n non fossero indipendenti, potremmo estrarre da questi una base con meno di n vettori, un assurdo per il Teorema 2.3.15. \square

Usando questo criterio possiamo dimostrare agevolmente il fatto seguente.

Proposizione 2.3.23. *I vettori*

$$v_1 = \begin{pmatrix} a_{11} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \\ \vdots \\ 0 \end{pmatrix}, \dots, v_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ a_{3n} \\ \vdots \\ a_{nn} \end{pmatrix}$$

sono una base di $\mathbb{K}^n \iff$ i numeri a_{11}, \dots, a_{nn} sono tutti diversi da zero.

Dimostrazione. Mostriamo che i vettori v_1, \dots, v_n sono indipendenti \iff i numeri a_{11}, \dots, a_{nn} sono tutti diversi da zero. Supponiamo che

$$\lambda_1 v_1 + \dots + \lambda_n v_n = 0.$$

Questa uguaglianza fra vettori è equivalente al sistema lineare omogeneo

$$\begin{cases} a_{11}\lambda_1 + a_{12}\lambda_2 + a_{13}\lambda_3 + \dots + a_{1n}\lambda_n = 0 \\ a_{22}\lambda_2 + a_{23}\lambda_3 + \dots + a_{2n}\lambda_n = 0 \\ a_{33}\lambda_3 + \dots + a_{3n}\lambda_n = 0 \\ \vdots \\ a_{nn}\lambda_n = 0 \end{cases}$$

Mostriamo per induzione su n che un tale sistema ha come *unica* soluzione $\lambda_1 = \dots = \lambda_n = 0 \iff$ i coefficienti a_{11}, \dots, a_{nn} sono tutti diversi da zero.

Per $n = 1$ il sistema è semplicemente l'equazione $a_{11}\lambda_1 = 0$ ed è chiaro che se $a_{11} \neq 0$ allora $\lambda_1 = 0$ è l'unica soluzione, mentre se $a_{11} = 0$ allora qualsiasi numero $\lambda_1 = t \in \mathbb{K}$ è soluzione.

Supponiamo l'asserzione vera per $n-1$ e la dimostriamo per n . Se $a_{nn} \neq 0$, allora l'ultima equazione $a_{nn}\lambda_n = 0$ implica che $\lambda_n = 0$. Sostituendo $\lambda_n = 0$ in tutte le equazioni precedenti ci riconduciamo al caso $n-1$ e concludiamo per l'ipotesi induttiva. Se invece $a_{nn} = 0$, allora $\lambda_1 = 0, \dots, \lambda_{n-1} = 0, \lambda_n = t$ è soluzione per ogni $t \in \mathbb{K}$ e quindi non è vero che il sistema ha un'unica soluzione. Abbiamo concluso. \square

2.3.7. Sottospazi vettoriali. La proposizione seguente ci racconta un fatto molto intuitivo: uno spazio non può contenere un sottospazio di dimensione più grande di lui.

Proposizione 2.3.24. *Sia V uno spazio vettoriale e $U \subset V$ un sottospazio. Vale $0 \leq \dim U \leq \dim V$. Inoltre:*

- $\dim U = 0 \iff U = \{0\}$.
- $\dim U = \dim V \iff U = V$.

Dimostrazione. Sia v_1, \dots, v_k base di U . Per l'algoritmo di completamento, questa può essere completata a base v_1, \dots, v_n di V . In particolare $n \geq k$, cioè $\dim V \geq \dim U$. Inoltre $k = 0 \iff U = \{0\}$ e $n = k \iff U = V$ \square

La proposizione ci dice anche che gli unici sottospazi di V di dimensione minima 0 e massima $\dim V$ sono precisamente il sottospazio banale $\{0\}$ ed il sottospazio totale V stesso.

Corollario 2.3.25. *Uno spazio vettoriale V di dimensione 1 contiene solo due sottospazi: il sottospazio banale $\{0\}$ e il sottospazio totale V .*

Gli spazi vettoriali di dimensione 1 e 2 vengono chiamati *rette* e *piani vettoriali*. I sottospazi di \mathbb{R}^3 sono, per dimensione crescente: $\{0\}$, le rette vettoriali, i piani vettoriali, e \mathbb{R}^3 stesso.

2.3.8. Formula di Grassmann. Concludiamo questa sezione con una formula che lega le dimensioni di due sottospazi, della loro intersezione e della loro somma.

Teorema 2.3.26 (Formula di Grassmann). *Sia V uno spazio vettoriale di dimensione finita. Siano $U, W \subset V$ sottospazi. Vale la formula*

$$\dim(U + W) + \dim(U \cap W) = \dim U + \dim W.$$

Dimostrazione. Sia v_1, \dots, v_k una base di $U \cap W$. Grazie all'algoritmo di completamento, possiamo completarla a base di U :

$$v_1, \dots, v_k, u_1, \dots, u_s$$

e anche a base di W :

$$v_1, \dots, v_k, w_1, \dots, w_t.$$

Adesso vogliamo dimostrare che

$$v_1, \dots, v_k, u_1, \dots, u_s, w_1, \dots, w_t$$

è una base di $U + W$. Se ci riusciamo, abbiamo finito perché otteniamo

$$\begin{aligned} \dim(U + W) &= k + s + t = (k + s) + (k + t) - k \\ &= \dim U + \dim W - \dim U \cap W. \end{aligned}$$

Per dimostrare che quella è una base, dobbiamo verificare che i vettori generino $U + W$ e che siano indipendenti. Per mostrare che generano, consideriamo un vettore generico $u + w \in U + W$. Poiché u si scrive come combinazione dei $v_1, \dots, v_k, u_1, \dots, u_s$ e w dei $v_1, \dots, v_k, w_1, \dots, w_t$, è chiaro che $u + w$ si scrive come combinazione dell'unione $v_1, \dots, v_k, u_1, \dots, u_s, w_1, \dots, w_t$.

Per mostrare l'indipendenza, consideriamo una combinazione lineare nulla:

$$\underbrace{\lambda_1 v_1 + \dots + \lambda_k v_k}_v + \underbrace{\mu_1 u_1 + \dots + \mu_s u_s}_u + \underbrace{\eta_1 w_1 + \dots + \eta_t w_t}_w = 0.$$

Per semplicità chiamiamo v, u, w i vettori indicati, così otteniamo:

$$v + u + w = 0.$$

Da questa deduciamo che

$$w = -v - u.$$

Notiamo che $w \in W$ e $-v - u \in U$. Siccome sono uguali, entrambi w e $-v - u$ stanno in $U \cap W$. In particolare $w \in U \cap W$. Allora w può essere scritto come combinazione lineare dei v_1, \dots, v_k , del tipo

$$\eta_1 w_1 + \dots + \eta_t w_t = \alpha_1 v_1 + \dots + \alpha_k v_k.$$

Spostando tutto a sinistra otteniamo

$$-\alpha_1 v_1 - \dots - \alpha_k v_k + \eta_1 w_1 + \dots + \eta_t w_t = 0.$$

Siccome $v_1, \dots, v_k, w_1, \dots, w_t$ è una base di W , tutti i coefficienti $\alpha_1, \dots, \alpha_k, \eta_1, \dots, \eta_t$ sono nulli. Quindi $w = 0$ e allora anche $v + u = 0$, cioè

$$\lambda_1 v_1 + \dots + \lambda_k v_k + \mu_1 u_1 + \dots + \mu_s u_s = 0.$$

Analogamente, siccome $v_1, \dots, v_k, u_1, \dots, u_s$ è base di U , anche i coefficienti $\lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_s$ sono nulli. La dimostrazione è conclusa. \square

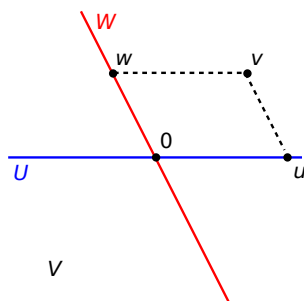


Figura 2.5. Una somma diretta $V = U \oplus W$. Ogni vettore $v \in V$ si scrive in modo unico come $v = u + w$ dove $u \in U$ e $w \in W$.

2.3.9. Somma diretta. Sia V uno spazio vettoriale. Diciamo che due sottospazi $U, W \subset V$ sono in *somma diretta* se $U \cap W = \{0\}$. In questo caso indichiamo la loro somma con $U \oplus W$. Dalla formula di Grassmann ricaviamo

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

Siamo particolarmente interessati al caso $V = U \oplus W$. Possiamo esprimere il fatto che $V = U \oplus W$ in vari modi equivalenti.

Proposizione 2.3.27. *Siano $U, W \subset V$ due sottospazi. Sono equivalenti:*

- (1) $V = U \oplus W$.
- (2) $\dim V = \dim(U + W) = \dim U + \dim W$.
- (3) *Per qualsiasi scelta di basi u_1, \dots, u_k di U e w_1, \dots, w_h di W , l'unione $u_1, \dots, u_k, w_1, \dots, w_h$ è una base di $U + W$.*
- (4) *Ogni vettore $v \in V$ si scrive in modo unico come $v = u + w$ per qualche $u \in U$ e $w \in W$. Si veda la Figura 2.5.*

Dimostrazione. (1) \Rightarrow (2) È la formula di Grassmann.

(2) \Rightarrow (3) L'unione delle basi genera $U + W$, che ha dimensione proprio $k + h$. Per la Proposizione 2.3.22, $k + h$ generatori sono una base di $U + W$.

(3) \Rightarrow (4) Ogni vettore si scrive in modo unico come combinazione di $u_1, \dots, u_k, w_1, \dots, w_h$ e quindi in modo unico come $u + w$.

(4) \Rightarrow (1) Otteniamo subito $V = U + W$. Se per assurdo ci fosse un $v \in U \cap W$ diverso da zero, questo si scriverebbe in modo non unico come $v = v + 0 = 0 + v$. \square

In una somma diretta $V = U \oplus W$, ogni vettore $v \in V$ si scrive in modo unico come somma $v = u + w$ con $u \in U$ e $w \in W$. Una rappresentazione grafica di questo fatto è mostrata nella Figura 2.5.

Notiamo ora un criterio che può essere utile per dimostrare che $V = U \oplus W$, un po' analogo alla Proposizione 2.3.22.

Proposizione 2.3.28. *Siano $U, W \subset V$ due sottospazi. Supponiamo che $\dim V = \dim U + \dim W$. I fatti seguenti sono tutti equivalenti:*

- (1) $V = U \oplus W$,
 (2) $V = U + W$,
 (3) $U + W = \{0\}$.

Dimostrazione. Segue dalla formula di Grassmann che (2) \Leftrightarrow (3). \square

Esempio 2.3.29. Due rette vettoriali in \mathbb{R}^2 distinte sono sempre in somma diretta. Un piano ed una retta vettoriali in \mathbb{R}^3 che si intersecano solo nell'origine sono sempre in somma diretta.

Esempio 2.3.30. In \mathbb{K}^n i sottospazi

$$U = \text{Span}(e_1, \dots, e_k), \quad W = \text{Span}(e_{k+1}, \dots, e_n)$$

sono in somma diretta.

Esempio 2.3.31. Consideriamo $V = \mathbb{R}^2$ e due rette

$$U = \text{Span} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad W = \text{Span} \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

Sappiamo che ogni vettore $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ si scrive in modo unico come combinazione lineare di un vettore in U e un vettore in W . Effettivamente:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \left(x + \frac{y}{2}\right) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{y}{2} \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

2.3.10. Trasposta di una matrice. Per il prossimo esempio di somma diretta abbiamo bisogno di introdurre una semplice operazione sulle matrici.

La *trasposta* di una matrice $A \in M(m, n, \mathbb{K})$ è la matrice

$${}^tA \in M(n, m, \mathbb{K})$$

definita scambiando righe e colonne, cioè:

$$({}^tA)_{ij} = A_{ji}.$$

Ad esempio:

$$A = \begin{pmatrix} 2 & 1 \\ -1 & 0 \\ 5 & 7 \end{pmatrix} \implies {}^tA = \begin{pmatrix} 2 & -1 & 5 \\ 1 & 0 & 7 \end{pmatrix}.$$

Valgono le proprietà seguenti:

$${}^t(A + B) = {}^tA + {}^tB, \quad {}^t(\lambda A) = \lambda {}^tA.$$

Se $A \in M(n)$, allora anche ${}^tA \in M(n)$. Notiamo che una matrice quadrata A è simmetrica $\Leftrightarrow {}^tA = A$ e inoltre A è antisimmetrica $\Leftrightarrow {}^tA = -A$.

2.3.11. Matrici simmetriche e antisimmetriche. Torniamo alle somme dirette di sottospazi: vogliamo descrivere un esempio con le matrici.

Esempio 2.3.32. Ricordiamo che $S(n)$ e $A(n)$ sono le matrici $n \times n$ simmetriche e antisimmetriche. Dimostriamo che

$$M(n) = S(n) \oplus A(n).$$

Dobbiamo verificare due cose:

- $M(n) = S(n) + A(n)$. Infatti ogni matrice $B \in M(n)$ si scrive come somma di una matrice simmetrica e di una antisimmetrica con il seguente trucco:

$$B = \frac{B + {}^tB}{2} + \frac{B - {}^tB}{2}.$$

Effettivamente si vede facilmente che il primo addendo è simmetrico ed il secondo antisimmetrico.

- $S(n) \cap A(n) = \{0\}$. Infatti solo la matrice nulla è contemporaneamente simmetrica e antisimmetrica.

Esercizio 2.3.33. Mostra che:

- Le matrici $e_{ij} + e_{ji}$ con $i < j$ e e_{ii} formano una base di $S(n)$,
- Le matrici $e_{ij} - e_{ji}$ con $i < j$ formano una base di $A(n)$.

Deduci che

$$\dim S(n) = 1 + \dots + n = \frac{n(n+1)}{2}, \quad \dim A(n) = 1 + \dots + n - 1 = \frac{(n-1)n}{2}.$$

Nota che effettivamente $\dim S(n) + \dim A(n) = n^2 = \dim M(n)$.

2.3.12. Somma diretta di più sottospazi. La nozione di somma diretta è molto utile e la ritroveremo varie volte in seguito. È possibile estenderla da 2 ad un numero arbitrario k di sottospazi.

Iniziamo estendendo l'operazione di somma. Sia V uno spazio vettoriale e $V_1, \dots, V_k \subset V$ sottospazi. Indichiamo con

$$V_1 + \dots + V_k$$

il sottoinsieme di V formato da tutti i vettori del tipo

$$v_1 + \dots + v_k$$

dove $v_i \in V_i$. Si verifica come nel caso $k = 2$ che $V_1 + \dots + V_k$ è un sottospazio vettoriale di V .

Diremo che V è *somma diretta* di V_1, \dots, V_k , e scriveremo

$$V = V_1 \oplus \dots \oplus V_k,$$

se valgono entrambi i fatti seguenti:

- $V = V_1 + \dots + V_k$,
- $V_i \cap (V_1 + \dots + V_{i-1} + V_{i+1} + \dots + V_k) = \{0\}$ per ogni $i = 1, \dots, k$.

La seconda condizione sembra inutilmente tecnica, ma è quella giusta, come mostra la seguente proposizione, del tutto analoga alla Proposizione 2.3.27.

Proposizione 2.3.34. *I fatti seguenti sono equivalenti:*

- (1) $V = V_1 \oplus \cdots \oplus V_k$.
- (2) $\dim V = \dim(V_1 + \cdots + V_k) = \dim V_1 + \cdots + \dim V_k$.
- (3) *Per qualsiasi scelta di basi $v_1^i, \dots, v_{h_i}^i$ di ciascun V_i , la loro unione al variare di $i = 1, \dots, k$ è una base di V .*
- (4) *Ogni vettore $v \in V$ si scrive in modo unico come*

$$v = v_1 + \cdots + v_k$$

per qualche $v_i \in V_i$.

Dimostrazione. (1) \Rightarrow (2) La prima uguaglianza è ovvia; per la seconda, mostriamo che

$$\dim(V_1 + \cdots + V_i) = \dim V_1 + \cdots + \dim V_i$$

per ogni $i = 1, \dots, k$ per induzione su i . Per $i = 1$ non c'è nulla da dimostrare, ed il caso generico è trattato nel modo seguente:

$$\begin{aligned} \dim(V_1 + \cdots + V_i) &= \dim((V_1 + \cdots + V_{i-1}) + V_i) \\ &= \dim(V_1 + \cdots + V_{i-1}) + \dim V_i \\ &= \dim V_1 + \cdots + \dim V_{i-1} + \dim V_i. \end{aligned}$$

Nella seconda uguaglianza abbiamo usato la formula di Grasmann, con

$$V_i \cap (V_1 + \cdots + V_{i-1} + V_{i+1} + \cdots + V_k) = \{0\} \implies V_i \cap (V_1 + \cdots + V_{i-1}) = \{0\}.$$

Nella terza uguaglianza abbiamo usato l'ipotesi induttiva.

(2) \Rightarrow (3) L'unione delle basi genera $V_1 + \cdots + V_k$ ed è formata esattamente da $\dim(V_1 + \cdots + V_k)$ vettori, quindi è una base di $V_1 + \cdots + V_k$.

(3) \Rightarrow (4) Ogni elemento $v \in V$ si scrive in modo unico come combinazione lineare dell'unione delle basi e quindi anche come combinazione lineare $v = v_1 + \cdots + v_k$ con $v_i \in V_i$.

(4) \Rightarrow (1) Otteniamo subito $V = V_1 + \cdots + V_k$. Se per assurdo ci fosse un $v \in V_i \cap (V_1 + \cdots + V_{i-1} + V_{i+1} + \cdots + V_k)$ diverso da zero, questo si scriverebbe in modo non unico come somma di elementi $v_j \in V_j$.

La dimostrazione è conclusa. \square

Esempio 2.3.35. Consideriamo $V = \mathbb{R}^5$ e i sottospazi $V_1 = \text{Span}(e_1)$, $V_2 = \text{Span}(e_2, e_3)$ e $V_3 = \text{Span}(e_4, e_5)$. Allora $V = V_1 \oplus V_2 \oplus V_3$.

Osservazione 2.3.36. Quando $k \geq 3$, non è sufficiente che valgano $V = V_1 + \cdots + V_k$ e $V_1 \cap \cdots \cap V_k = \{0\}$ affinché gli spazi siano in somma diretta. Ad esempio, un qualsiasi insieme di k rette vettoriali distinte in \mathbb{R}^2 soddisfa queste due ipotesi, ma per $k \geq 3$ queste rette non sono mai in somma diretta.

Esercizi

Esercizio 2.1. Sia V uno spazio vettoriale su \mathbb{K} e X un insieme. Mostra che l'insieme $F(X, V)$ ha una naturale struttura di spazio vettoriale su \mathbb{K} .

Esercizio 2.2. Siano V e W due spazi vettoriali su \mathbb{K} . Definisci su $V \times W$ una somma come $(v, w) + (v', w') = (v + v', w + w')$ e un prodotto per scalare come $\lambda(v, w) = (\lambda v, \lambda w)$. Mostra che $V \times W$ con queste operazioni è uno spazio vettoriale.

Esercizio 2.3. Siano $v_1, \dots, v_k \in V$. Mostra che $\text{Span}(v_1, \dots, v_k)$ è l'intersezione di tutti i sottospazi di V contenenti i vettori v_1, \dots, v_k .

Esercizio 2.4. Siano $v_1, \dots, v_k \in V$ indipendenti. Mostra che

$$\text{Span}(v_1, \dots, v_h) \cap \text{Span}(v_{h+1}, \dots, v_k) = \{0\}.$$

Esercizio 2.5. Sia V uno spazio vettoriale e $U, W \subset V$ sottospazi. Mostra che:

- (1) $U \cap W$ è l'intersezione di tutti i sottospazi contenuti in U e W
- (2) $U + W$ è l'intersezione di tutti i sottospazi che contengono U e W .

Esercizio 2.6. Determina delle basi dei sottospazi seguenti di \mathbb{R}^3 :

$$U = \{x + y + z = 0\}, \quad W = \{x - 2y - z = 0\}, \quad U \cap W, \quad U + W.$$

Esercizio 2.7. Sia $a \in \mathbb{K}$ un valore fissato. Sia $W \subset \mathbb{K}_n[x]$ il sottoinsieme formato dai polinomi che si annullano in a . Trova una base per W e calcolane la dimensione.

Esercizio 2.8. Trova una base di $W = \{x_1 + x_2 + \dots + x_n = 0\} \subset \mathbb{R}^n$.

Esercizio 2.9. Completa i vettori seguenti a base di \mathbb{R}^4 :

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}$$

Esercizio 2.10. Mostra che le matrici seguenti di $M(2, \mathbb{R})$ sono indipendenti e completale a base di $M(2, \mathbb{R})$.

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}.$$

Scrivi le coordinate della matrice $e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ rispetto alla base che hai trovato.

Esercizio 2.11. Sia $W \subset M(2, \mathbb{R})$ il sottospazio generato dalle tre matrici descritte nell'esercizio precedente. Per quali $t \in \mathbb{R}$ la matrice $\begin{pmatrix} 0 & 1 \\ t & 0 \end{pmatrix}$ appartiene a W ?

Esercizio 2.12. Completa i seguenti polinomi a base di $\mathbb{C}_3[x]$:

$$x + i, \quad x^2 - ix.$$

Scrivi le coordinate del polinomio x^2 rispetto alla base che hai trovato.

Esercizio 2.13. Estrai da questa lista di generatori una base di $M(2, \mathbb{C})$:

$$\begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} i & -1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ i & 0 \end{pmatrix}, \quad \begin{pmatrix} -i & 1+i \\ 0 & 1 \end{pmatrix}.$$

Esercizio 2.14. Scrivi le coordinate dei vettori

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 7 \end{pmatrix}$$

rispetto alla base $\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ di \mathbb{R}^2 .

Esercizio 2.15. Dimostra che $\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}$ è una base di $\mathbb{R}^2 \iff ad - bc \neq 0$.

Esercizio 2.16. Mostra che i vettori $v_1 = e_1, v_2 = e_1 + e_2, \dots, v_n = e_1 + \dots + e_n$ sono una base di \mathbb{R}^n .

Esercizio 2.17. Sia X un insieme finito di n elementi. Per ogni $x \in X$ definiamo la funzione $\delta_x: X \rightarrow \mathbb{K}$ in questo modo:

$$\delta_x(y) = \begin{cases} 1 & \text{se } y = x, \\ 0 & \text{se } y \neq x. \end{cases}$$

Mostra che le funzioni δ_x al variare di $x \in X$ sono una base di $F(X, \mathbb{K})$. Deduci che $\dim F(X, \mathbb{K}) = n$.

Esercizio 2.18. Siano V e W sottospazi di dimensione finita. Mostra che $\dim(V \times W) = \dim V + \dim W$ (vedi l'Esercizio 2.2).

Esercizio 2.19. Sia X un insieme con n elementi e V uno spazio vettoriale di dimensione m . Mostra che $\dim F(X, V) = mn$ (vedi l'Esercizio 2.1).

Esercizio 2.20. Sia V uno spazio vettoriale di dimensione n . Mostra che per ogni $0 \leq k \leq n$ esiste un sottospazio $W \subset V$ di dimensione k .

Esercizio 2.21. Sia V uno spazio vettoriale di dimensione finita e $U \subset V$ un sottospazio. Un *complementare algebrico* per U è un sottospazio $W \subset V$ tale che $V = U \oplus W$. Mostra che esiste sempre un complementare algebrico per U .

Esercizio 2.22. Determina 4 vettori $v_1, v_2, v_3, v_4 \in \mathbb{R}^4$ con queste proprietà:

- (1) Se elimini un qualsiasi vettore, i tre rimanenti sono sempre indipendenti.
- (2) I sottospazi $U = \text{Span}(v_1, v_2)$ e $W = \text{Span}(v_3, v_4)$ non sono in somma diretta.

Esercizio 2.23. Siano U e W sottospazi di \mathbb{R}^5 , entrambi di dimensione 3. Quali sono le possibili dimensioni di $U \cap W$? Descrivi degli esempi.

Esercizio 2.24. Considera i due sottospazi di \mathbb{R}^3 , dipendenti da una variabile $t \in \mathbb{R}$:

$$U = \text{Span} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad W_t = \text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ t \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \right).$$

Determina per quali valori di $t \in \mathbb{R}$ vale $U \oplus W = \mathbb{R}^3$.

Esercizio 2.25. Siano U una retta vettoriale e W un piano vettoriale in \mathbb{R}^3 . Dimostra che i fatti seguenti sono tutti equivalenti:

$$\mathbb{R}^3 = U \oplus W \iff U \cap W = \{0\} \iff (U \not\subseteq W \text{ e } W \not\subseteq U).$$

Esercizio 2.26. Mostra che qualsiasi spazio vettoriale V di dimensione n contiene n sottospazi V_1, \dots, V_n tutti di dimensione 1 tali che

$$V = V_1 \oplus \dots \oplus V_n.$$

Esercizio 2.27. Siano U, V, W sottospazi di \mathbb{R}^4 , tutti di dimensione 3. Siano $a = \dim(U \cap V)$, $b = \dim(V \cap W)$, $c = \dim(U \cap W)$, $d = \dim(U \cap V \cap W)$.

- (1) Dimostra che $a \in \{2, 3\}$.
- (2) Dimostra che $d > 0$.
- (3) Costruisci un esempio di U, V, W in cui $d = 2$.
- (4) Costruisci un esempio di U, V, W in cui $d = 1$.
- (5) Determina tutti i possibili valori per la quadrupla (a, b, c, d) al variare di U, V, W .

Complementi

Sistemi lineari

3.1. Algoritmi di risoluzione

Introduciamo in questa sezione un algoritmo che ci permetterà di risolvere qualsiasi sistema lineare. L'ingrediente principale consiste in alcune mosse con le quali possiamo modificare una matrice in modo controllato, note come *mosse di Gauss*.

3.1.1. Mosse di Gauss. Un *sistema lineare* è un insieme di k equazioni lineari in n variabili

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{k1}x_1 + \cdots + a_{kn}x_n = b_k \end{cases}$$

I numeri a_{ij} sono i *coefficienti* e i b_i sono i *termini noti* del sistema. Sia i coefficienti, che i termini noti, che le variabili sono in un certo campo fissato \mathbb{K} . Possiamo raggruppare i coefficienti e i termini noti efficacemente in una matrice $k \times n$ e un vettore colonna:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kn} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_k \end{pmatrix}.$$

Possiamo quindi unire tutto in un'unica matrice $k \times (n + 1)$:

$$C = (A \mid b)$$

Il nostro scopo è determinare l'insieme $S \subset \mathbb{K}^n$ delle soluzioni del sistema. A questo scopo, notiamo che esistono alcune mosse che cambiano la matrice C senza modificare l'insieme S delle soluzioni del sistema. Queste sono le seguenti e sono note come *mosse di Gauss*:

- (1) Scambiare due righe.
- (2) Moltiplicare una riga per un numero $\lambda \neq 0$.
- (3) Aggiungere ad una riga un'altra riga moltiplicata per λ qualsiasi.

Indicando con C_i la i -esima riga di C , possiamo scrivere le mosse così:

- (1) $C_i \longleftrightarrow C_j$,
- (2) $C_i \longrightarrow \lambda C_i$, $\lambda \neq 0$,
- (3) $C_i \longrightarrow C_i + \lambda C_j$.

Alcuni esempi di composizioni di mosse di tipo (1), (2) e (3):

$$\begin{pmatrix} 1 & 2 \\ 3 & 0 \\ 4 & -2 \end{pmatrix} \xrightarrow{(1)} \begin{pmatrix} 4 & -2 \\ 3 & 0 \\ 1 & 2 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 2 & -1 \\ 3 & 0 \\ 1 & 2 \end{pmatrix} \xrightarrow{(3)} \begin{pmatrix} 2 & -1 \\ 5 & 4 \\ 1 & 2 \end{pmatrix}.$$

Torniamo al sistema lineare iniziale, determinato dalla matrice $C = (A | b)$.

Proposizione 3.1.1. *Le mosse di Gauss su $C = (A | b)$ non mutano l'insieme $S \subset \mathbb{K}^n$ delle soluzioni del sistema lineare.*

Dimostrazione. La mossa (1) scambia due equazioni e questo chiaramente non tocca S . La mossa (2) moltiplica una equazione lineare

$$a_{i1}x_1 + \cdots + a_{in}x_n = b_i$$

per una costante $\lambda \neq 0$, trasformandola quindi in

$$\lambda a_{i1}x_1 + \cdots + \lambda a_{in}x_n = \lambda b_i.$$

Anche qui le soluzioni sono chiaramente le stesse di prima. L'effetto della mossa (3) sulle equazioni lineari corrispondenti alle righe i e j è il seguente: cambiamo le due equazioni

$$\begin{aligned} a_{i1}x_1 + \cdots + a_{in}x_n &= b_i \\ a_{j1}x_1 + \cdots + a_{jn}x_n &= b_j \end{aligned}$$

con

$$\begin{aligned} a_{i1}x_1 + \cdots + a_{in}x_n &= b_i \\ (a_{j1} + \lambda a_{i1})x_1 + \cdots + (a_{jn} + \lambda a_{in})x_n &= b_j + \lambda b_i \end{aligned}$$

È facile convincersi che $x \in \mathbb{K}^n$ è soluzione delle prime due se e solo se è soluzione delle seconde due. \square

3.1.2. Algoritmo di Gauss. Introduciamo un algoritmo che useremo per risolvere i sistemi lineari. L'algoritmo trasforma qualsiasi matrice C in una matrice particolare detta *a scalini*.

Sia C una matrice qualsiasi. Per ogni riga C_i di C , chiamiamo *pivot* il primo elemento non nullo della riga C_i . Una *matrice a scalini* è una matrice avente la proprietà seguente: il pivot di ogni riga è sempre strettamente più a destra del pivot della riga precedente. Ad esempio, sono matrici a scalini:

$$\begin{pmatrix} 1 & 0 & 5 \\ 0 & -1 & -1 \end{pmatrix}, \quad \begin{pmatrix} 7 & -2 & 9 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

mentre invece non sono matrici a scalini:

$$\begin{pmatrix} 1 & 2 & -1 \\ 4 & 0 & 6 \\ 0 & 7 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 8 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

L'*algoritmo di Gauss* è un algoritmo che trasforma qualsiasi matrice C in una matrice a scalini usando le 3 mosse di Gauss descritte precedentemente. Funziona così:

- (1) Se $C_{11} = 0$, scambia se puoi la prima riga con un'altra riga in modo da ottenere $C_{11} \neq 0$. Se $C_{i1} = 0$ per ogni i , passa al punto (3).
- (2) Per ogni riga C_i con $i \geq 2$ e con $C_{i1} \neq 0$ sostituisci C_i con la riga

$$C_i - \frac{C_{i1}}{C_{11}}C_1.$$

In questo modo la nuova riga C_i avrà $C_{i1} = 0$.

- (3) Adesso abbiamo ottenuto $C_{i1} = 0$ per ogni $i \geq 2$. Continua dal punto (1) lavorando sulla sottomatrice ottenuta togliendo la prima riga e la prima colonna.

Vediamo un esempio. Partiamo dalla matrice

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

Notiamo che $C_{11} = 0$. Quindi scambiamo C_1 e C_2 in modo da ottenere $C_{11} \neq 0$. Ad ogni passaggio chiamiamo sempre la nuova matrice con la lettera C per semplicità. Otteniamo

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

Ora notiamo che $C_{31} \neq 0$. Cambiamo C_3 con $C_3 + C_1$ in modo da ottenere $C_{31} = 0$. Il risultato è

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

La prima riga e la prima colonna adesso sono a posto, quindi lavoriamo sul resto. Notiamo che $C_{32} \neq 0$. Sostituiamo C_3 con $C_3 - 3C_2$ in modo da ottenere $C_{32} = 0$. Il risultato è

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

Adesso la matrice è a scalini e l'algoritmo termina.

3.1.3. Algoritmo di Gauss - Jordan. Per un motivo che sarà chiaro in seguito, dopo aver ridotto la matrice a scalini, è a volte utile fare delle ulteriori mosse di Gauss per fare in modo che tutti i numeri *sopra* i pivot siano nulli. Questo è sempre possibile usando mosse di Gauss di tipo (3), aggiungendo la riga che contiene il pivot (moltiplicata per una costante opportuna) alle righe precedenti.

Ad esempio, nella matrice C ottenuta precedentemente abbiamo $C_{12} \neq 0$. Per ottenere $C_{12} = 0$ possiamo sostituire C_1 con $C_1 - C_2$. Il risultato è una nuova matrice

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

Adesso vediamo che $C_{14} \neq 0$. Per ottenere $C_{14} = 0$ sostituiamo C_1 con $C_1 - \frac{3}{2}C_3$ e otteniamo infine

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

Come ultimo ritocco, vedremo che sarà anche utile ottenere che tutti i pivot abbiano valore 1. È sempre possibile ottenere questo semplicemente con mosse di Gauss di tipo (2).

Nel nostro esempio, i pivot C_{11} e C_{22} hanno già valore 1, però $C_{34} = -2$. Per ottenere $C_{34} = 1$ dividiamo la terza riga per -2 . Il risultato finale è

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

L'algoritmo che abbiamo appena descritto si chiama *algoritmo di Gauss - Jordan*. Consiste in due fasi:

- (1) Trasformare la matrice a scalini tramite algoritmo di Gauss.
- (2) Ottenere solo zeri sopra i pivot con mosse (3) e tutti i pivot uguali a 1 con mosse (2).

Facciamo un altro esempio di applicazione dell'algoritmo di Gauss - Jordan. Sopra ogni freccia indichiamo la mossa di Gauss corrispondente.

$$\begin{pmatrix} 1 & -1 & 3 \\ 0 & 2 & 2 \\ 1 & 0 & 4 \end{pmatrix} \xrightarrow{C_3 \rightarrow C_3 - C_1} \begin{pmatrix} 1 & -1 & 3 \\ 0 & 2 & 2 \\ 0 & 1 & 1 \end{pmatrix} \xrightarrow{C_3 \rightarrow C_3 - \frac{1}{2}C_2} \begin{pmatrix} 1 & -1 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{C_1 \rightarrow C_1 + \frac{1}{2}C_2} \begin{pmatrix} 1 & 0 & 4 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{C_2 \rightarrow \frac{1}{2}C_2} \begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

L'algoritmo di Gauss - Jordan è garantito funzionare sempre; d'altra parte, per ottenere una matrice a scalini (eventualmente con zeri sopra i pivot e pivot uguali a 1) non è necessario seguire pedissequamente l'algoritmo: qualsiasi combinazione di mosse di Gauss è lecita purché si arrivi al risultato.

3.1.4. Risoluzione di un sistema lineare. Mostriamo adesso come sia possibile descrivere esplicitamente lo spazio $S \subset \mathbb{K}^n$ delle soluzioni di un dato sistema lineare.

Il sistema è descritto da una matrice $C = (A | b)$. Usando l'algoritmo di Gauss - Jordan, possiamo trasformare C in una matrice a scalini, in cui ogni pivot ha valore 1 e tutti i numeri sopra i pivot sono nulli. La matrice C sarà indicativamente di questo tipo:

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

Adesso guardiamo le colonne che contengono i pivot, che in questo esempio sono la seconda, la quarta e la quinta. Ciascuna di queste colonne contiene un 1 al posto del pivot e 0 in tutte le altre caselle.

Ci sono due casi da considerare. Se la colonna dei termini noti b contiene un pivot, allora la matrice è indicativamente di questo tipo:

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

La riga che contiene questo pivot rappresenta l'equazione $0 = 1$ che chiaramente non può avere soluzione. In questo caso l'insieme S delle soluzioni è vuoto, cioè $S = \emptyset$.

Consideriamo adesso il caso in cui l'ultima colonna non contenga pivot. La matrice C è indicativamente di questo tipo:

$$\begin{pmatrix} 0 & 1 & a_{13} & 0 & 0 & a_{16} & b_1 \\ 0 & 0 & 0 & 1 & 0 & a_{26} & b_2 \\ 0 & 0 & 0 & 0 & 1 & a_{36} & b_3 \end{pmatrix}.$$

Ricordiamo che ciascuna colonna corrisponde ad una variabile x_1, \dots, x_n , eccetto l'ultima che contiene i termini noti. Adesso assegniamo un parametro t_1, t_2, \dots a ciascuna variabile che corrisponde ad una colonna che *non* contiene un pivot. Nel nostro caso, scriviamo $x_1 = t_1$, $x_3 = t_2$ e $x_6 = t_3$. Conseguentemente il sistema lineare è di questa forma:

$$\begin{cases} x_2 + a_{13}t_2 + a_{16}t_3 = b_1 \\ x_4 + a_{26}t_3 = b_2 \\ x_5 + a_{36}t_3 = b_3 \end{cases}$$

Spostiamo a destra dell'uguale tutti i nuovi parametri e otteniamo:

$$\begin{cases} x_2 = b_1 - a_{13}t_2 - a_{16}t_3 \\ x_4 = b_2 - a_{26}t_3 \\ x_5 = b_3 - a_{36}t_3 \end{cases}$$

Al sistema aggiungiamo anche le equazioni corrispondenti alle assegnazioni dei parametri liberi, che nel nostro caso sono $x_1 = t_1$, $x_3 = t_2$ e $x_6 = t_3$. Il

risultato è un sistema del tipo:

$$\begin{cases} x_1 = t_1 \\ x_2 = b_1 - a_{13}t_2 - a_{16}t_3 \\ x_3 = t_2 \\ x_4 = b_2 - a_{26}t_3 \\ x_5 = b_3 - a_{36}t_3 \\ x_6 = t_3 \end{cases}$$

Il sistema è risolto. I parametri t_1, t_2, \dots sono liberi e possono assumere qualsiasi valore in \mathbb{K} , e le variabili x_1, \dots, x_n dipendono da questi parametri liberi come indicato.

Notiamo che è sempre possibile usare una notazione parametrica vettoriale per esprimere le soluzioni, del tipo

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} t_1 \\ b_1 - a_{13}t_2 - a_{16}t_3 \\ t_2 \\ b_2 - a_{26}t_3 \\ b_3 - a_{36}t_3 \\ t_3 \end{pmatrix} = \begin{pmatrix} 0 \\ b_1 \\ 0 \\ b_2 \\ b_3 \\ 0 \end{pmatrix} + t_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 0 \\ -a_{13} \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t_3 \begin{pmatrix} 0 \\ -a_{16} \\ 0 \\ -a_{26} \\ -a_{36} \\ 1 \end{pmatrix}.$$

In generale, le soluzioni saranno sempre del tipo

$$x = x_0 + t_1 v_1 + \dots + t_h v_h$$

dove i vettori v_i saranno tutti della forma

$$v_i = \begin{pmatrix} ? \\ \vdots \\ ? \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

con la posizione della cifra 1 che scende al crescere di i . I vettori v_1, \dots, v_h sono sempre indipendenti per la Proposizione 2.3.23 e inoltre $h = n - r$ dove n è il numero di variabili e r è il numero di pivot.

3.2. Rango

Nella sezione precedente abbiamo visto come risolvere un sistema lineare tramite l'algoritmo di Gauss - Jordan. In questa studiamo il problema in modo più teorico, definendo e studiando il *rango* di una matrice.

3.2.1. Il sistema omogeneo associato. Consideriamo come sopra un sistema di equazioni lineari

$$(2) \quad \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{k1}x_1 + \cdots + a_{kn}x_n = b_k \end{cases}$$

Il *sistema omogeneo associato* è quello ottenuto semplicemente mettendo a zero tutti i termini noti b_i , cioè:

$$(3) \quad \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \vdots \\ a_{k1}x_1 + \cdots + a_{kn}x_n = 0 \end{cases}$$

Scriviamo la matrice $A = (a_{ij})$ dei coefficienti ed il vettore $b = (b_i)$ dei termini noti. La matrice $C = (A \mid b)$ rappresenta il sistema lineare iniziale (2), mentre $(A \mid 0)$, oppure più semplicemente A , rappresenta il sistema lineare omogeneo associato (3).

Sia $S \subset \mathbb{K}^n$ l'insieme delle soluzioni del sistema lineare iniziale e $S_0 \subset \mathbb{K}^n$ l'insieme delle soluzioni del sistema omogeneo associato. Ricordiamo che S_0 è un sottospazio vettoriale di \mathbb{K}^n per la Proposizione 2.2.4, mentre S invece non lo è perché non contiene l'origine – a meno che l'insieme lineare iniziale non sia già omogeneo e in questo caso $S = S_0$, si veda l'Osservazione 2.2.5.

I due insiemi S e S_0 sono strettamente correlati:

Proposizione 3.2.1. *Se $S \neq \emptyset$, allora S è ottenuto prendendo una qualsiasi soluzione $x \in S$ e aggiungendo a questa tutti i vettori di S_0 .*

Dimostrazione. Sia x una soluzione del sistema (2) e x' una soluzione del sistema (3). Si vede facilmente che $x + x'$ è anch'essa soluzione di (2), infatti

$$a_{i1}(x_1 + x'_1) + \cdots + a_{in}(x_n + x'_n) = b_i + 0 = b_i.$$

Se invece x'' è soluzione di (2), allora $x' = x'' - x$ è soluzione di (3) perché

$$a_{i1}(x''_1 - x_1) + \cdots + a_{in}(x''_n - x_n) = b_i - b_i = 0.$$

Quindi le soluzioni di (2) si ottengono tutte precisamente aggiungendo ad un fissato $x \in S$ le soluzioni $x' \in S_0$ di (3). \square

Esempio 3.2.2. Consideriamo il sistema in \mathbb{R}^3

$$\begin{cases} x - y + z = 1 \\ y - z = 2 \end{cases}$$

Risolvendolo troviamo che le soluzioni sono del tipo

$$\begin{pmatrix} 3 \\ t \\ t - 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ -2 \end{pmatrix} + \begin{pmatrix} 0 \\ t \\ t \end{pmatrix}$$

al variare di $t \in \mathbb{R}$. Le soluzioni del sistema omogeneo associato

$$\begin{cases} x - y + z = 0 \\ y - z = 0 \end{cases}$$

sono precisamente i vettori del tipo

$$\begin{pmatrix} 0 \\ t \\ t \end{pmatrix}.$$

3.2.2. Sottospazio affine. Cerchiamo adesso di rispondere alla seguente domanda: che tipo di luogo geometrico è l'insieme S di soluzioni di un sistema lineare? La risposta è che S è sempre un *sottospazio affine*.

Sia V uno spazio vettoriale. Un *sottospazio affine* di W è un qualsiasi sottoinsieme del tipo

$$S = \{x + v \mid v \in W\}$$

dove x è un punto fissato di V e $W \subset V$ è un sottospazio vettoriale.

Esempio 3.2.3. Abbiamo appena visto che le soluzioni S di un sistema di equazioni lineari formano sempre un sottospazio affine di \mathbb{K}^n . Infatti

$$S = \{x + v \mid v \in S_0\}$$

dove x è un punto qualsiasi di S e S_0 è l'insieme delle soluzioni del sistema lineare omogeneo associato, che è sempre un sottospazio vettoriale.

Per indicare S , usiamo semplicemente la scrittura

$$S = x + W = \{x + v \mid v \in W\}.$$

Lo stesso sottospazio affine può essere descritto in più modi diversi. Chiariamo subito quando due scritture differenti descrivono lo stesso sottospazio:

Proposizione 3.2.4. *Gli spazi affini $x + W$ e $x' + W'$ coincidono se e solo se $W = W'$ e $x - x' \in W$.*

Dimostrazione. Se $W = W'$ e $x - x' \in W$, allora chiaramente ogni vettore $x + v$ con $v \in W$ può essere scritto anche come $x' + (x - x' + v)$ con $x - x' + v \in W$ e quindi $x + W \subset x' + W$. Analogamente $x + W \supset x' + W$.

D'altro canto, se $x + W = x' + W'$ allora $W = x' - x + W'$. Questo implica che $x' - x \in W$. Analogamente $x' - x \in W'$ e quindi $W = W'$. \square

Nella descrizione di uno spazio affine S come $x + W$, lo spazio vettoriale W è determinato da S ed è detto la *giacitura* di S . Il punto x invece è un qualsiasi punto di S .

Definizione 3.2.5. La *dimensione* di S è la dimensione della giacitura W .

Un sottospazio affine di dimensione 1 o 2 è detto *retta affine* e *piano affine*. Lo spazio \mathbb{R}^3 contiene ovviamente molte rette affini e molti piani affini. Questi verranno studiati più approfonditamente nel Capitolo 8.

Tornando ai nostri sistemi lineari, abbiamo appena scoperto che lo spazio delle soluzioni S , se non è vuoto, è un sottospazio affine con una certa dimensione. La sua giacitura è lo spazio S_0 delle soluzioni del sistema omogeneo associato. Vale $\dim S = \dim S_0$ per definizione.

3.2.3. Rango. Sappiamo come risolvere un sistema lineare e che l'insieme S delle soluzioni può essere vuoto o un sottospazio affine di una certa dimensione. Cerchiamo adesso delle tecniche per capire rapidamente se un sistema lineare abbia soluzioni, e in caso affermativo la loro dimensione.

Introduciamo una nozione che comparirà spesso in questo libro. Sia A una matrice $m \times n$ a coefficienti in \mathbb{K} . Ricordiamo che indichiamo con A^1, \dots, A^n le colonne di A . Ciascun A^i è un vettore in \mathbb{K}^m .

Definizione 3.2.6. Il *rango* di A è la dimensione dello spazio

$$\text{Span}(A^1, \dots, A^n) \subset \mathbb{K}^m.$$

Il rango di A è la dimensione dello spazio generato dalle colonne. Viene comunemente indicato con $\text{rk}(A)$. Una formulazione equivalente è la seguente.

Proposizione 3.2.7. *Il rango di A è il massimo numero di colonne linearmente indipendenti di A .*

Dimostrazione. È un fatto generale che la dimensione di un sottospazio $W = \text{Span}(v_1, \dots, v_k)$ è il massimo numero di vettori indipendenti che possiamo trovare fra v_1, \dots, v_k . Segue dall'algoritmo di estrazione (esercizio). \square

Notiamo adesso che il rango è insensibile alle mosse di Gauss.

Proposizione 3.2.8. *Se modifichiamo A per mosse di Gauss sulle righe, il suo rango non cambia.*

Dimostrazione. Se esiste una relazione di dipendenza lineare fra alcune colonne, questa si conserva con le mosse di Gauss sulle righe (esercizio). Quindi il massimo numero di colonne indipendenti non cambia. \square

Corollario 3.2.9. *Il rango di A è il numero di pivot in una sua qualsiasi riduzione a scalini.*

Dimostrazione. Poiché il rango non cambia con mosse di Gauss, possiamo supporre che A sia già ridotta a scalini e inoltre usare la seconda parte dell'algoritmo di Gauss-Jordan per fare in modo che le k colonne contenenti i pivot siano precisamente i vettori e_1, \dots, e_k . Tutte le altre colonne sono combinazioni lineari di questi. Lo spazio generato dalle colonne è quindi $\text{Span}(e_1, \dots, e_k)$ ed ha dimensione k pari al numero di pivot. \square

Corollario 3.2.10. Se A è una matrice $m \times n$ allora $\text{rk}A \leq \min\{m, n\}$.

Una matrice A ha rango massimo se $\text{rk}A = \min\{m, n\}$.

3.2.4. Teorema di Rouché - Capelli. Enunciamo e dimostriamo un teorema che fornisce un algoritmo pratico ed efficace per capire se un sistema lineare ha soluzioni, e in caso affermativo la loro dimensione.

Consideriamo come al solito un sistema lineare

$$(4) \quad \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{k1}x_1 + \cdots + a_{kn}x_n = b_k \end{cases}$$

Scriviamo come sopra la matrice $A = (a_{ij})$ dei coefficienti, il vettore $b = (b_i)$ dei termini noti e la matrice completa $C = (A \mid b)$. Iniziamo con una proposizione.

Proposizione 3.2.11. Il sistema (4) ha soluzioni se e solo se

$$b \in \text{Span}(A^1, \dots, A^n).$$

Dimostrazione. Il sistema può essere riscritto nel modo seguente:

$$x_1A^1 + \cdots + x_nA^n = b.$$

A questo punto è chiaro che esistono soluzioni se e solo se b è combinazione lineare delle colonne A^1, \dots, A^n , con coefficienti x_1, \dots, x_n . \square

Passiamo al teorema più importante di questa sezione.

Teorema 3.2.12 (Rouché - Capelli). Il sistema (4) ha soluzioni se e solo se

$$\text{rk}(A \mid b) = \text{rk}(A).$$

In caso affermativo, lo spazio delle soluzioni $S \subset \mathbb{K}^n$ è un sottospazio affine di dimensione $n - \text{rk}(A)$.

Dimostrazione. Sappiamo che il sistema ha soluzioni se e solo se

$$b \in \text{Span}(A^1, \dots, A^n).$$

Se questo accade, allora

$$\text{Span}(A^1, \dots, A^n, b) = \text{Span}(A^1, \dots, A^n)$$

e quindi $\text{rk}(A \mid b) = \text{rk}(A)$. Se invece questo non accade, allora

$$\text{Span}(A^1, \dots, A^n, b) \supsetneq \text{Span}(A^1, \dots, A^n)$$

e quindi $\text{rk}(A \mid b) > \text{rk}(A)$.

Se ci sono soluzioni, sappiamo che la dimensione di S è pari alla dimensione dello spazio S_0 delle soluzioni del sistema lineare omogeneo associato. Dopo aver effettuato l'algoritmo di Gauss - Jordan, abbiamo già dimostrato nella Sezione 3.1.4 che S_0 è generato da un numero di vettori indipendenti pari

a n meno il numero di pivot, cioè $n - \text{rk}(A)$. Quindi $\dim S = \dim S_0 = n - \text{rk}(A)$. \square

Esempio 3.2.13. Consideriamo un sistema lineare dipendente da un parametro $k \in \mathbb{R}$

$$\begin{cases} x + ky = 4 - k \\ kx + 4y = 4 \end{cases}$$

Vogliamo sapere al variare di $k \in \mathbb{R}$ se ci siano soluzioni e, in caso affermativo, che dimensione abbiano. Appliciamo l'algoritmo di Gauss su $(A | b)$:

$$\begin{pmatrix} 1 & k & 4 - k \\ k & 4 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & k & 4 - k \\ 0 & 4 - k^2 & 4 - 4k + k^2 \end{pmatrix}$$

Per calcolare il rango è sufficiente fermarsi qui, non è necessario proseguire con l'algoritmo di Gauss - Jordan che è utile solo se vogliamo determinare precisamente le soluzioni. Notiamo che $4 - 4k + k^2 = (k - 2)^2$. La matrice è ridotta a scalini per tutti i possibili valori di $k \in \mathbb{R}$ e ci sono sempre due pivot, tranne per $k = 2$ in cui ce n'è uno solo. Quindi il rango di $(A | b)$ è 1 per $k = 2$ e 2 per $k \neq 2$.

La matrice dei coefficienti A con le mosse di Gauss è diventata

$$\begin{pmatrix} 1 & k \\ 0 & 4 - k^2 \end{pmatrix}$$

ed ha rango 1 per $k = \pm 2$ e rango 2 per $k \neq \pm 2$. Quindi:

- Se $k = -2$, allora $\text{rk}(A | b) \neq \text{rk}(A)$ e non ci sono soluzioni.
- Se $k = 2$, allora $\text{rk}(A | b) = \text{rk}(A) = 1$, quindi le soluzioni formano un sottospazio affine di \mathbb{R}^3 di dimensione $3-1=2$, cioè un piano affine.
- Se $k \neq \pm 2$, allora $\text{rk}(A | b) = \text{rk}(A) = 2$, quindi le soluzioni formano un sottospazio affine di \mathbb{R}^3 di dimensione $3-2=1$, cioè una retta affine.

3.2.5. Sistema lineare omogeneo. Consideriamo un sistema omogeneo

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \vdots \\ a_{k1}x_1 + \cdots + a_{kn}x_n = 0 \end{cases}$$

con matrice dei coefficienti A . Il Teorema di Rouché - Capelli per i sistemi omogenei assume una forma più semplice:

Corollario 3.2.14. *Un sistema lineare omogeneo ha sempre soluzione. Il sottospazio vettoriale $S \subset \mathbb{K}^n$ delle soluzioni ha dimensione $n - \text{rk}(A)$.*

Notiamo le due differenze principali fra i sistemi lineari omogenei e non omogenei: nei primi le soluzioni ci sono sempre e formano un sottospazio vettoriale, mentre nei secondi le soluzioni possono non esserci e se ci sono formano un sottospazio affine.

Indichiamo con $W \subset \mathbb{K}^m$ lo spazio vettoriale generato dalle colonne di A . Per definizione $\text{rk}(A) = \dim W$ e quindi per il Teorema di Rouché - Capelli

$$(5) \quad \dim S + \dim W = n.$$

Notiamo che $S \subset \mathbb{K}^n$ e $W \subset \mathbb{K}^m$ sono sottospazi vettoriali di due spazi diversi \mathbb{K}^n e \mathbb{K}^m . È utile capire cosa succeda ad entrambi S e W se cambiamo la matrice A con mosse di Gauss sulle righe e sulle colonne:

Proposizione 3.2.15. *Valgono i fatti seguenti:*

- Con mosse di Gauss sulle righe, S è immutato ma W può variare.
- Con mosse di Gauss sulle colonne, W è immutato ma S può variare.

Dimostrazione. Il caso delle righe è già stato analizzato. Per le colonne, per definizione $W = \text{Span}(A^1, \dots, A^n)$. Se scambiamo due colonne A^i, A^j o cambiamo A^i con λA^i , $\lambda \neq 0$, chiaramente W non muta. È facile verificare che non muta neppure se cambiamo A^i con $A^i + \lambda A^j$. \square

Corollario 3.2.16. *Le dimensioni di S e W non mutano né per mosse di Gauss sulle righe né sulle colonne.*

Dimostrazione. Segue da (5). \square

Abbiamo scoperto che combinando mosse di Gauss su righe e colonne i sottospazi S e W possono mutare, ma le loro dimensioni rimangono costanti.

3.2.6. Rango per righe e per colonne. Abbiamo appena dimostrato il fatto seguente. Sia A una matrice $m \times n$.

Corollario 3.2.17. *Il rango $\text{rk}A$ di una matrice A non cambia per mosse di Gauss sulle righe o sulle colonne.*

Definiamo il *rango per righe* di A come la dimensione dello spazio generato dalle righe $\text{Span}(A_1, \dots, A_m) \subset \mathbb{K}^n$. Il *rango per colonne* è l'usuale rango. In altre parole, il rango per righe di A è il rango della trasposta tA .

Proposizione 3.2.18. *Per ogni matrice A il rango per righe è uguale al rango per colonne.*

Dimostrazione. Se riduciamo A a scalini, entrambi i ranghi non cambiano per il Corollario 3.2.17. In una matrice a scalini il rango per colonne è il numero r di pivot. Quello per righe è al massimo r perché nella matrice a scalini ci sono solo r righe non nulle. Quindi il rango per righe è sempre \leq del rango per colonne. Lavorando con tA troviamo analogamente che il rango per colonne è sempre \leq di quello per righe. Quindi sono uguali. \square

Corollario 3.2.19. *Per ogni matrice A vale $\text{rk}({}^tA) = \text{rk}A$.*

Il rango di una matrice è uguale al rango della sua trasposta.

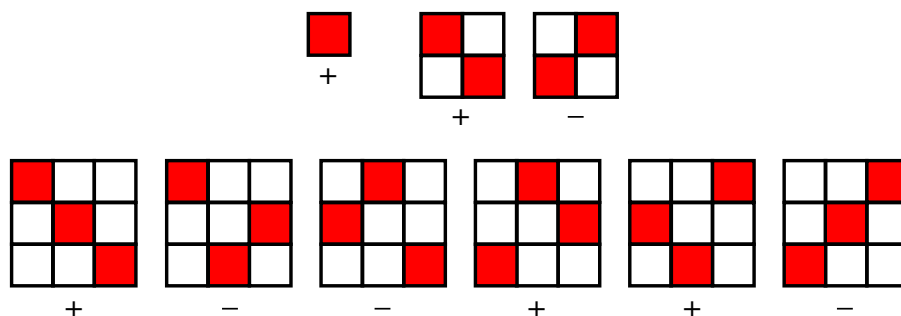


Figura 3.1. Rappresentazione grafica del determinante di una matrice $n \times n$ per $n = 1, 2$ e 3 .

3.3. Determinante

Introduciamo adesso una nozione importante che si applica solo alle matrici quadrate, quella di *determinante*.

3.3.1. Definizione. Sia A una matrice quadrata $n \times n$. Il *determinante* di A è il numero

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

Ricordiamo che S_n indica l'insieme delle $n!$ permutazioni di $\{1, \dots, n\}$, si veda la Sezione 1.2.5. Quindi questa è una sommatoria su $n!$ elementi. Il termine $\operatorname{sgn}(\sigma)$ indica il segno della permutazione σ ed è 1 oppure -1 . In seguito indichiamo la permutazione σ con il simbolo $[\sigma(1) \cdots \sigma(n)]$.

Ad esempio, se $n = 1$ l'insieme S_1 contiene solo la permutazione $\operatorname{id} = [1]$ e otteniamo semplicemente

$$\det A = a_{11}.$$

Se $n = 2$ l'insieme S_2 contiene due permutazioni $\operatorname{id} = [1 \ 2]$ e $[2 \ 1]$, e ricordiamo che le trasposizioni hanno segno -1 . Quindi

$$\det A = a_{11}a_{22} - a_{12}a_{21}.$$

Se $n = 3$ l'insieme S_3 contiene sei permutazioni:

$$\operatorname{id} = [1 \ 2 \ 3], \quad [1 \ 3 \ 2], \quad [2 \ 1 \ 3], \quad [2 \ 3 \ 1], \quad [3 \ 1 \ 2], \quad [3 \ 2 \ 1].$$

Le tre trasposizioni hanno segno -1 e le altre tre permutazioni $+1$. Quindi

$$\det A = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.$$

La Figura 3.1 mostra una rappresentazione grafica del determinante nei casi descritti $n = 1, 2, 3$. Ciascun quadrato grande corrisponde ad una permutazione ed il suo segno è indicato sotto. Ad esempio, il determinante delle

matrici seguenti

$$(3), \quad \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix}$$

è rispettivamente:

$$3, \quad 4 + 2 = 6, \quad 1 - 0 - 4 + (-4) + 0 - (-1) = -6.$$

Osservazione 3.3.1. Ispirati dalla Figura 3.1, riformuliamo la definizione di determinante con una notazione più visiva. Consideriamo una matrice quadrata A come $n \times n$ caselle con dei numeri al loro interno. Chiamiamo una *colorazione* di A la colorazione in rosso di n sue caselle, in modo che ogni riga e ogni colonna contengano esattamente una sola casella rossa.¹

C'è una chiara corrispondenza biunivoca fra permutazioni in S_n e colorazioni. Data una $\sigma \in S_n$, otteniamo una colorazione c dipingendo di rosso le caselle $1\sigma(1), \dots, n\sigma(n)$. Il *segno* della colorazione c è per definizione il segno della corrispondente permutazione σ . Il *peso* $p(c)$ della colorazione c è il prodotto dei numeri nelle caselle rosse, moltiplicato per il segno. Con questo linguaggio, il determinante di A è semplicemente la somma dei pesi di tutte le colorazioni: se indichiamo con $\text{Col}(A)$ l'insieme delle colorazioni, scriviamo

$$\det(A) = \sum_{c \in \text{Col}(A)} p(c).$$

Vediamo subito che il determinante è insensibile alla trasposizione.

Proposizione 3.3.2. *Vale* $\det({}^t A) = \det A$.

Dimostrazione. Seguiamo l'interpretazione visiva descritta nell'Osservazione 3.3.1. Data una colorazione $c \in \text{Col}(A)$, possiamo farne la trasposta e ottenere una colorazione ${}^t c \in \text{Col}({}^t A)$. Le due permutazioni corrispondenti sono una l'inversa dell'altra e quindi hanno lo stesso segno per la Proposizione 1.2.12. I due pesi sono gli stessi, cioè $p(c) = p({}^t c)$. Sommando sulle colorazioni si ottiene la tesi. \square

3.3.2. Matrici triangolari. Il calcolo del determinante di una matrice triangolare è particolarmente semplice: è il prodotto dei valori sulla diagonale principale.

Proposizione 3.3.3. *Sia* $A \in M(n)$ *una matrice triangolare superiore*

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

¹In altre parole, una colorazione è la configurazione di n torri su una scacchiera $n \times n$ che non possono mangiarsi l'una con l'altra.

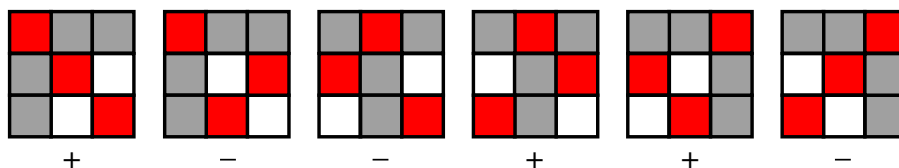


Figura 3.2. Dimostrazione dello sviluppo di Laplace.

Vale

$$\det A = a_{11} \cdots a_{nn}.$$

Dimostrazione. Usiamo l'interpretazione dell'Osservazione 3.3.1. Poiché la matrice è triangolare, l'unica colorazione che può avere peso non nullo è quella in cui le caselle della diagonale principale sono tutte rosse, che corrisponde alla permutazione identità ed ha quindi segno 1. Quindi $\det A = a_{11} \cdots a_{nn}$. \square

Lo stesso risultato vale per le matrici triangolari inferiori.

Corollario 3.3.4. *La matrice identità I_n ha determinante $\det(I_n) = 1$.*

3.3.3. Sviluppo di Laplace. Ci sono vari algoritmi che permettono di calcolare il determinante di una matrice più agevolmente che con la cruda definizione. Uno di questi è noto come lo *sviluppo di Laplace* e funziona nel modo seguente.

Sia A una matrice $n \times n$ con $n \geq 2$. Indichiamo con C_{ij} la sottomatrice $(n-1) \times (n-1)$ ottenuta da A rimuovendo la i -esima riga e la j -esima colonna.

Teorema 3.3.5 (Sviluppo di Laplace). *Per ogni i fissato vale l'uguaglianza*

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det C_{ij}.$$

Dimostrazione. Per $n = 2$ otteniamo $\det A = a_{11}a_{22} - a_{12}a_{21}$ e siamo a posto. Il caso $n = 3$ e $i = 1$ è mostrato nella Figura 3.2, che mostra l'idea chiave della dimostrazione.

Sappiamo dall'Osservazione 3.3.1 che

$$\det(A) = \sum_{c \in \text{Col}(A)} p(c).$$

Sia $\text{Col}^i(A) \subset \text{Col}(A)$ l'insieme delle colorazioni in cui la casella ij è rossa. Chiaramente abbiamo una partizione $\text{Col}(A) = \text{Col}^1(A) \cup \cdots \cup \text{Col}^n(A)$ e

$$\det(A) = \sum_{j=1}^n \sum_{c \in \text{Col}^j(A)} p(c).$$

Come suggerito dalla Figura 3.2, ogni colorazione $c \in \text{Col}^j(A)$ induce una colorazione $\pi(c)$ sulla sottomatrice C_{ij} , e la funzione $\pi: \text{Col}^j(A) \rightarrow \text{Col}(C_{ij})$ è una bigezione. Inoltre, per quello che riguarda i pesi troviamo:

$$p(c) = (-1)^{i+j} a_{ij} p(\pi(c)).$$

Il segno $(-1)^{i+j}$ è presente perché le due permutazioni corrispondenti hanno lo stesso segno se e solo se $i+j$ è pari (esercizio). Quindi possiamo scrivere:

Ok esercizio qui?

$$\begin{aligned} \det(A) &= \sum_{j=1}^n \sum_{c \in \text{Col}^j(A)} (-1)^{i+j} a_{ij} p(\pi(c)) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \sum_{c' \in \text{Col}(C_{ij})} p(c') \\ &= \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(C_{ij}). \end{aligned}$$

La dimostrazione è conclusa. \square

Vediamo un esempio. Per calcolare il determinante seguente, sviluppiamo lungo la prima riga (cioè prendiamo $i = 1$) e otteniamo

$$\begin{aligned} \det \begin{pmatrix} 1 & -1 & 0 \\ 2 & -1 & 5 \\ 1 & 1 & -1 \end{pmatrix} &= 1 \cdot \det \begin{pmatrix} -1 & 5 \\ 1 & -1 \end{pmatrix} - (-1) \cdot \det \begin{pmatrix} 2 & 5 \\ 1 & -1 \end{pmatrix} \\ &\quad + 0 \cdot \det \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} = -4 - 7 + 0 = -11. \end{aligned}$$

È chiaramente conveniente sviluppare lungo una riga che contiene degli zeri.

Notiamo inoltre che grazie al fatto che $\det({}^t A) = \det A$, è anche possibile usare lo sviluppo di Laplace sulle colonne. Ad esempio, sviluppando la matrice seguente sulla seconda colonna otteniamo:

$$\det \begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & 1 \\ \pi & 3 & \sqrt{7} \end{pmatrix} = (-1) \cdot 3 \cdot \det \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} = 3.$$

Nello sviluppo si deve sempre fare attenzione al segno $(-1)^{i+j}$ associato alla casella ij .

3.3.4. Mosse di Gauss. Veniamo adesso al vero motivo per cui abbiamo introdotto il determinante di una matrice: la sua compatibilità con le mosse di Gauss.

Proposizione 3.3.6. *Sia A una matrice $n \times n$ e sia A' ottenuta da A tramite una mossa di Gauss. Il determinante cambia nel modo seguente:*

- (1) Se A' è ottenuta da A scambiando due righe, $\det(A') = -\det(A)$.
- (2) Se A' è ottenuta moltiplicando una riga di A per uno scalare λ , allora $\det(A') = \lambda \det(A)$.

(3) Se A' è ottenuta aggiungendo ad una riga il multiplo di un'altra riga, allora $\det(A') = \det A$.

Dimostrazione. Dimostriamo la proposizione per induzione su $n \geq 2$. Se $n = 2$, tutte queste proprietà discendono facilmente dalla formula $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$. Dimostriamo il teorema per $n \geq 3$ supponendo che sia vero per $n - 1$.

Siccome $n \geq 3$, esiste sempre una riga A_i che non è implicata nella mossa di Gauss. Lo sviluppo di Laplace lungo A_i e A'_i è

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det C_{ij},$$

$$\det A' = \sum_{j=1}^n (-1)^{i+j} a'_{ij} \det C'_{ij}.$$

La matrice C'_{ij} è di taglia $(n - 1) \times (n - 1)$ ed è ottenuta da C_{ij} tramite lo stesso tipo di mossa di Gauss di A . Quindi per l'ipotesi induttiva

$$\det(C'_{ij}) = -\det(C_{ij}), \quad \lambda \det(C_{ij}) \quad \text{oppure} \quad \det(C_{ij})$$

a seconda del tipo di mossa di Gauss. Raccogliendo -1 oppure λ troviamo la stessa relazione tra $\det(A')$ e $\det A$. \square

Poiché il determinante non cambia per trasposizione, la stessa conclusione è valida anche per mosse di Gauss sulle colonne. Mostriamo ora una formula un po' più generale. Sia A una matrice $n \times n$ e supponiamo che la i -esima colonna sia del tipo

$$A^i = \lambda_1 v_1 + \dots + \lambda_h v_h$$

per qualche $v_j \in \mathbb{K}^n$ e $\lambda_j \in \mathbb{K}$. Sia B_j la matrice ottenuta da A sostituendo la colonna A^i con v_j .

Proposizione 3.3.7. Vale $\det A = \lambda_1 \det(B_1) + \dots + \lambda_h \det(B_h)$.

Dimostrazione. Stessa dimostrazione della Proposizione 3.3.6. \square

La stessa formula vale ovviamente anche con le righe al posto delle colonne.

Esempio 3.3.8. Nel calcolo del determinante di una matrice, è spesso utile combinare mosse di Gauss e sviluppo di Laplace. Ad esempio:

$$\det \begin{pmatrix} 5 & 1 & 1 & 2 & 1 \\ 1 & 4 & 1 & 1 & 1 \\ 1 & 3 & 3 & 1 & 1 \\ 1 & 1 & 1 & 6 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} = \det \begin{pmatrix} 4 & 0 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} = \det \begin{pmatrix} 4 & 0 & 0 & 1 \\ 0 & 3 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

$$= 5 \det \begin{pmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 2 & 2 \end{pmatrix} = 5 \cdot 4 \cdot 3 \cdot 2 = 120.$$

3.3.5. Matrici di rango massimo. A cosa serve il determinante? Innanzitutto a capire se una matrice quadrata ha rango massimo.

Proposizione 3.3.9. *Per ogni $A \in M(n, \mathbb{K})$ abbiamo*

$$\det A \neq 0 \iff \text{rk}A = n.$$

Dimostrazione. Usiamo l'algoritmo di Gauss con A . Ad ogni mossa di Gauss, il rango si mantiene e il determinante cambia per moltiplicazione per una costante non nulla. Quindi possiamo supporre che A sia ridotta a scalini. La matrice A è anche triangolare superiore e quindi

$$\text{rk}A = n \iff a_{11} \cdots a_{nn} \neq 0 \iff \det A \neq 0$$

per la Proposizione 3.3.3. \square

3.3.6. Calcolo del rango. Più in generale, il determinante può essere usato per calcolare il rango di una matrice qualsiasi. Ricordiamo che il determinante è definito solo per le matrici quadrate, mentre il rango è definito per tutte.

Sia A una matrice $m \times n$. Un *minore* di A di ordine k è una qualsiasi sottomatrice quadrata $k \times k$ ottenuta cancellando $m - k$ righe e $n - k$ colonne.

Proposizione 3.3.10. *Il rango di A è il massimo ordine di un suo minore B con $\det B \neq 0$.*

Dimostrazione. Con una mossa di Gauss il rango $r = \text{rk}A$ non cambia. Inoltre ciascun minore B cambia anche lui per una mossa di Gauss e anche il fatto di avere $\det B \neq 0$ quindi non cambia. Quindi ci basta dimostrare la proposizione per una matrice A ridotta a scalini.

Sia A ridotta a scalini. Il minore B che contiene i pivot è una matrice triangolare superiore $r \times r$ con numeri non nulli sulla diagonale principale, quindi $\det B \neq 0$. D'altra parte, qualsiasi minore B' di ordine $> r$ ha l'ultima riga nulla e quindi $\det B' = 0$. \square

Esempio 3.3.11. Usiamo il determinante per capire per quali $k \in \mathbb{R}$ il seguente sistema abbia soluzioni:

$$\begin{cases} y + kz = -k \\ 2x + (k-3)y + 4z = k-1 \\ x + ky - kz = k+1 \end{cases}$$

La matrice dei coefficienti è

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

e ha determinante $(k+4)(k+1)$. Quindi per $k \neq -1, -4$ sia A che $C = (A | b)$ hanno rango massimo e il sistema ha una sola soluzione $x = A^{-1}b$.

Per $k = -1$ la matrice C diventa

$$C = \left(\begin{array}{ccc|c} 0 & 1 & -1 & 1 \\ 2 & -4 & 4 & -2 \\ 1 & -1 & 1 & 0 \end{array} \right)$$

Troviamo $\text{rk}A = \text{rk}C = 2$ perché è facile trovare una relazione tra le righe di C , oppure perché tutti e 4 i minori 3×3 di C hanno determinante nullo, e il minore $\begin{pmatrix} 0 & 1 \\ 2 & -4 \end{pmatrix}$ ha determinante diverso da zero. Quindi il sistema ha soluzione, e le soluzioni formano un sottospazio affine di dimensione $3 - 2 = 1$.

Per $k = -4$ la matrice C diventa

$$C = \left(\begin{array}{ccc|c} 0 & 1 & -4 & 4 \\ 2 & -7 & 4 & -5 \\ 1 & -4 & 4 & -3 \end{array} \right)$$

Troviamo $\text{rk}A = 2$ perché A contiene il minore $\begin{pmatrix} 0 & 1 \\ 2 & -7 \end{pmatrix}$ che ha $\det \neq 0$. D'altra parte $\text{rk}C = 3$ perché C contiene un minore 3×3 con determinante non nullo:

$$\begin{aligned} \det \begin{pmatrix} 0 & -4 & 4 \\ 2 & 4 & -5 \\ 1 & 4 & -3 \end{pmatrix} &= 4 \det \begin{pmatrix} 0 & -1 & 4 \\ 2 & 1 & -5 \\ 1 & 1 & -3 \end{pmatrix} = 4 \det \begin{pmatrix} 0 & -1 & 4 \\ 0 & -1 & 1 \\ 1 & 1 & -3 \end{pmatrix} \\ &= 4 \det \begin{pmatrix} -1 & 4 \\ -1 & 1 \end{pmatrix} = 4 \cdot 3 = 12. \end{aligned}$$

Durante il calcolo abbiamo usato le mosse di Gauss per semplificare la matrice. Quindi per $k = -4$ il sistema non ha soluzioni.

3.3.7. Base di \mathbb{K}^n . Una applicazione concreta del determinante è la seguente: per capire se n vettori $v_1, \dots, v_n \in \mathbb{K}^n$ sono una base, si affiancano e si costruisce una matrice quadrata $A = (v_1 \cdots v_n)$. Quindi:

Proposizione 3.3.12. *I vettori v_1, \dots, v_n sono una base di $\mathbb{K}^n \Leftrightarrow \det A \neq 0$.*

Dimostrazione. Per la Proposizione 2.3.22, i vettori v_1, \dots, v_n sono una base \Leftrightarrow generano $\mathbb{K}^n \Leftrightarrow \text{rk}(A) = n \Leftrightarrow \det(A) \neq 0$. \square

Corollario 3.3.13. *Due vettori $\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \in \mathbb{K}^2$ sono una base $\Leftrightarrow ad - bc \neq 0$.*

Osservazione 3.3.14. Con questo criterio possiamo ridimostrare agevolmente la Proposizione 2.3.23. Se affianchiamo i vettori v_1, \dots, v_n otteniamo una matrice triangolare A con valori a_{11}, \dots, a_{nn} sulla diagonale principale, quindi con $\det A = a_{11} \cdots a_{nn}$. I vettori formano una base $\Leftrightarrow a_{11} \cdots a_{nn} \neq 0$, cioè se e solo se i valori a_{11}, \dots, a_{nn} sono tutti diversi da zero.

3.3.8. Orientazione e volume. Siano $v_1, \dots, v_n \in \mathbb{R}^n$ dei vettori, e sia $A = (v_1 \cdots v_n)$. Abbiamo appena visto che i vettori sono indipendenti $\Leftrightarrow \det A \neq 0$. Nel caso in cui siano indipendenti, quale informazione geometrica ci dà il numero reale $\det A$?

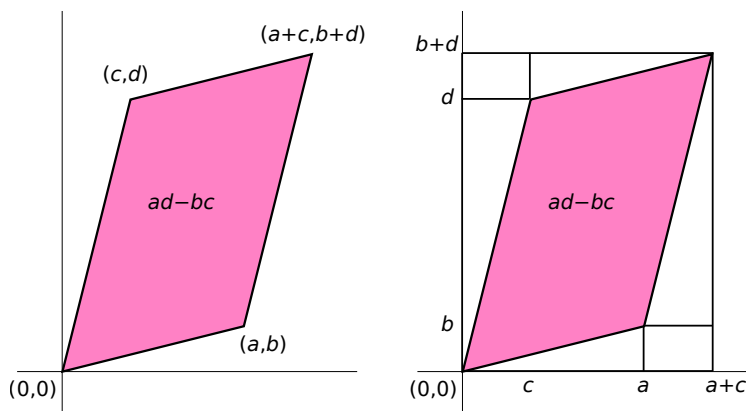


Figura 3.3. L'area del parallelogramma è il valore assoluto del determinante $|ad - bc|$.

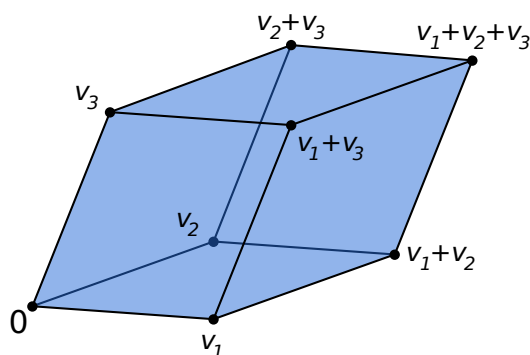


Figura 3.4. Il volume del parallelepipedo è il valore assoluto del determinante $|\det(v_1 v_2 v_3)|$.

Il numero reale $\det A$ porta con sé due informazioni. La prima è il segno: se $\det A > 0$, diciamo che la base v_1, \dots, v_n è *positiva*, e se $\det A < 0$ è *negativa*. La base canonica e_1, \dots, e_n ha $\det A = 1$ e quindi è positiva. Se scambiamo due elementi qualsiasi della base canonica, otteniamo una nuova base con $\det A = -1$ che è quindi negativa.

La seconda informazione riguarda l'area o il volume del parallelogramma o parallelepipedo generato da v_1, \dots, v_n . Per $n = 2$ il valore assoluto $|\det A| = |ad - bc|$ è l'area del parallelogramma generato da v_1 e v_2 , come mostrato nella Figura 3.3 usando

$$(a+c)(b+d) - 2bc - 2\frac{ab}{2} - 2\frac{cd}{2} = ad - bc.$$

Analogamente per $n = 3$ il valore assoluto del determinante è il volume del parallelepipedo generato da v_1, v_2 e v_3 mostrato nella Figura 3.4. Dimostreremo questa formula successivamente.

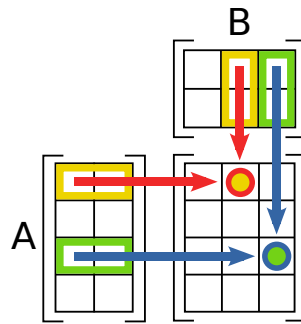


Figura 3.5. Il prodotto riga per colonna.

3.4. Algebra delle matrici

I sistemi lineari possono essere scritti in una forma più compatta usando una nuova operazione algebrica che introduciamo e studiamo in questa sezione: il *prodotto* fra matrici.

3.4.1. Prodotto fra matrici. Se A è una matrice $m \times n$ e B è una matrice $n \times p$ il *prodotto* AB è una nuova matrice $m \times p$ definita nel modo seguente: l'elemento $(AB)_{ij}$ della nuova matrice AB è

$$(AB)_{ij} = \sum_{k=1}^n A_{ik} B_{kj} = A_{i1} B_{1j} + \cdots + A_{in} B_{nj}.$$

Questo tipo di prodotto fra matrici si chiama *prodotto riga per colonna* perché l'elemento $(AB)_{ij}$ si ottiene facendo un opportuno prodotto fra la riga i -esima A_i di A e la colonna j -esima B^j di B , come suggerito dalla Figura 3.5.

Ad esempio:

$$\begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} -1 & 2 & 0 & 1 \\ 3 & 0 & 3 & 0 \end{pmatrix} = \begin{pmatrix} 5 & 1 & 6 & 1 \\ 4 & -2 & 3 & -1 \\ 9 & 0 & 9 & 0 \end{pmatrix}.$$

Ricordiamo che il prodotto AB si può fare solo se il numero di colonne di A è pari al numero di righe di B . In particolare, se A è una matrice $m \times n$ e x è una matrice $n \times 1$, cioè un vettore colonna $x \in \mathbb{K}^n$, allora il prodotto Ax è una matrice $m \times 1$, cioè un vettore colonna in \mathbb{K}^m . Ad esempio:

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

3.4.2. Notazione compatta per i sistemi lineari. Grazie al prodotto fra matrici è possibile scrivere un sistema lineare in una forma molto più compatta.

Possiamo scrivere il sistema

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{k1}x_1 + \cdots + a_{kn}x_n = b_k \end{cases}$$

semplicemente nella forma

$$Ax = b$$

dove A è la matrice $k \times n$ dei coefficienti e $b \in \mathbb{K}^n$ il vettore dei termini noti.

3.4.3. Proprietà. Studiamo le proprietà algebriche del prodotto fra matrici. Ricordiamo che il prodotto AB ha senso solo se il numero di colonne di A coincide con il numero di righe di B .

Proposizione 3.4.1. *Valgono le proprietà seguenti, per ogni matrice A, B, C per cui i prodotti abbiano senso e per ogni $\lambda \in \mathbb{K}$.*

- (1) $A(B + C) = AB + AC$ e $(A + B)C = AC + BC$ (distributività)
- (2) $A(BC) = (AB)C$ (associatività)
- (3) $\lambda(AB) = (\lambda A)B = A(\lambda B)$

Dimostrazione. Per la (1), troviamo

$$(A(B+C))_{ij} = \sum_{k=1}^n A_{ik}(B+C)_{kj} = \sum_{k=1}^n A_{ik}B_{kj} + \sum_{k=1}^n A_{ik}C_{kj} = (AB)_{ij} + (BC)_{ij}$$

e l'altra uguaglianza è analoga. Per la (2), verifichiamo che

$$\begin{aligned} (A(BC))_{ij} &= \sum_{k=1}^n A_{ik}(BC)_{kj} = \sum_{k=1}^n \sum_{h=1}^m A_{ik}B_{kh}C_{hj} \\ ((AB)C)_{ij} &= \sum_{h=1}^m (AB)_{ih}C_{hj} = \sum_{h=1}^m \sum_{k=1}^n A_{ik}B_{kh}C_{hj} \end{aligned}$$

e scambiando le sommatorie si ottiene lo stesso totale. Il punto (3) è più facile ed è lasciato per esercizio. \square

Esercizio 3.4.2. Vale la relazione

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

3.4.4. L'anello delle matrici quadrate. Consideriamo lo spazio vettoriale $M(n)$ delle matrici $n \times n$. Il prodotto fra matrici è un'operazione binaria in $M(n)$, perché il prodotto di due matrici $A, B \in M(n)$ si può sempre fare ed il risultato è una nuova matrice $AB \in M(n)$. Ricordiamo la definizione di anello data nella Sezione 1.5.2.

Proposizione 3.4.3. *L'insieme $M(n)$ delle matrici $n \times n$ è un anello con le operazioni di somma e di prodotto fra matrici. L'elemento neutro per il prodotto è la matrice*

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

i cui coefficienti sono 1 sulla diagonale principale e 0 altrove.

Dimostrazione. Sappiamo già che $M(n)$ è un gruppo commutativo con la somma. Per la Proposizione 3.4.1, il prodotto è associativo e distributivo rispetto alla somma. Resta da mostrare che I_n è un elemento neutro per il prodotto, e cioè che

$$I_n A = A I_n = A \quad \forall A \in M(n).$$

Lo verifichiamo adesso. Indichiamo I_n con I per semplicità. Troviamo:

$$(IA)_{ij} = \sum_{k=1}^n I_{ik} A_{kj} = I_{ij} A_{ij} = A_{ij}$$

perché I_{ik} è uguale a zero se $i \neq k$ e ad uno se $i = k$. L'altra uguaglianza è dimostrata in modo analogo. \square

La matrice quadrata I_n è chiamata *matrice identità* ed è spesso indicata semplicemente con I . Abbiamo quindi scoperto che lo spazio vettoriale $M(n)$ delle matrici quadrate è al tempo stesso uno spazio vettoriale e un anello.

3.4.5. Matrici (non) commutanti e (non) invertibili. Notiamo adesso due fatti importanti, che ci mostrano che l'anello $M(n)$ è ben lontano dall'essere un campo se $n \geq 2$. Ricordiamo che un campo è un anello in cui il prodotto è commutativo e ogni elemento diverso da zero ha un inverso rispetto al prodotto. Ebbene entrambe queste proprietà sono false in $M(n)$ se $n \geq 2$.

Se $n \geq 2$ l'anello $M(n)$ non è commutativo: è abbastanza raro che due matrici A e B commutino. Ad esempio, se $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ e $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, allora

$$AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = B, \quad BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

Ci sono inoltre molte matrici quadrate che non hanno un'inversa per il prodotto. Una matrice quadrata A per cui esista un'inversa per il prodotto è detta *invertibile*. In questo caso l'inversa è indicata con A^{-1} e per definizione vale $AA^{-1} = A^{-1}A = I$ dove I è la matrice identità, elemento neutro del prodotto.

Ad esempio, le due matrici A e B appena descritte non sono invertibili. Se per assurdo esistesse una inversa per A , che chiameremmo A^{-1} , dall'equazione $BA = 0$ moltiplicando entrambi i membri a destra per A^{-1} otterremmo $B = 0$, che è assurdo. Analogamente per B .

Vedremo nelle prossime pagine come determinare le matrici invertibili. Intanto vediamo che il prodotto di due matrici invertibili è sempre invertibile. Questa proprietà vale in qualsiasi anello, si veda l'Esercizio 1.17.

Proposizione 3.4.4. *Se $A, B \in M(n)$ sono invertibili, anche AB lo è e*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Dimostrazione. Verifichiamo che effettivamente

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = I,$$

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I.$$

La dimostrazione è conclusa. \square

Osservazione 3.4.5. In assenza di inversa, non si può sempre “semplificare” con il prodotto: con le matrici A e B di sopra abbiamo $AB = B$, ma non è possibile eliminare la B e ottenere $A = I_2$, infatti $A \neq I_2$. Con la somma invece si può sempre fare: se $C + D = D$ allora $C = 0$, perché aggiungiamo $-D$ ad entrambi i membri.

3.4.6. Teorema di Binet. Fra le numerose proprietà del determinante, c'è una forte compatibilità con il prodotto fra matrici: il determinante di AB è semplicemente il prodotto dei determinanti di A e B . Questo fatto molto utile è noto come *Teorema di Binet*.

Teorema 3.4.6 (Teorema di Binet). *Date due matrici $A, B \in M(n)$, vale*

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

Dimostrazione. Abbiamo

$$\det(AB) = \det \begin{pmatrix} a_{11}B_1 + \dots + a_{1n}B_n \\ \vdots \\ a_{n1}B_1 + \dots + a_{nn}B_n \end{pmatrix}$$

Se usiamo numerose volte la Proposizione 3.3.7 sulle righe possiamo scrivere questa quantità come somma di n^n addendi del tipo

$$a_{1\sigma(1)} \cdots a_{n\sigma(n)} \det \begin{pmatrix} B_{\sigma(1)} \\ \vdots \\ B_{\sigma(n)} \end{pmatrix}$$

al variare di $\sigma(1), \dots, \sigma(n) \in \{1, \dots, n\}$. Sappiamo che se due righe sono uguali allora il determinante si annulla: quindi ci possiamo limitare a considerare solo i casi in cui $\sigma(1), \dots, \sigma(n)$ sono tutti distinti, in altre parole σ è una permutazione. Quindi otteniamo

$$\det(AB) = \sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} \det \begin{pmatrix} B_{\sigma(1)} \\ \vdots \\ B_{\sigma(n)} \end{pmatrix}.$$

In ogni addendo, la permutazione σ è prodotto di un certo numero k di trasposizioni. Quindi scambiando le righe k volte si trasforma la matrice di destra in B . Ad ogni scambio il determinante cambia di segno, e ricordando che $\text{sgn}(\sigma) = (-1)^k$ otteniamo

$$\det(AB) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} \det \begin{pmatrix} B_1 \\ \vdots \\ B_n \end{pmatrix} = \det A \det B.$$

La dimostrazione è conclusa. \square

Corollario 3.4.7. Per ogni $A \in M(n, \mathbb{K})$ invertibile vale $\det A \neq 0$ e

$$\det(A^{-1}) = \frac{1}{\det A}.$$

Dimostrazione. Osserviamo che $I_n = A^{-1}A$ implica che

$$1 = \det(I_n) = \det(A^{-1}A) = \det(A^{-1}) \det A.$$

La dimostrazione è conclusa. \square

Osservazione 3.4.8. In generale $\det(A + B) \neq \det A + \det B$. Ad esempio $\det \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$ e $\det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \det \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = 1 + 1 = 2$.

3.4.7. Calcolo dell'inversa di una matrice. Introduciamo adesso un algoritmo che permette di calcolare l'inversa di una matrice A , quando esiste. Si basa, come molte altre cose in questo capitolo, sulle mosse di Gauss.

L'algoritmo funziona nel modo seguente. Si prende la matrice $n \times (2n)$

$$C = (A \mid I_n)$$

ottenuta affiancando ad A la matrice identità I_n . Trasformiamo quindi C usando l'algoritmo di Gauss - Jordan. Se $\text{rk}A = n$, alla fine dell'algoritmo tutti i pivot stanno nelle prime n colonne. In particolare abbiamo ottenuto una nuova matrice $n \times (2n)$

$$C' = (I \mid B)$$

dove B è una nuova matrice $n \times n$.

Proposizione 3.4.9. Vale $B = A^{-1}$.

Dimostrazione. Per definizione di inversa, la colonna i -esima di A^{-1} è l'unico vettore $x_0 \in \mathbb{K}^n$ tale che $Ax_0 = e_i$. Il sistema lineare $Ax = e_i$ è descritto dalla matrice $(A \mid e_i)$, con unica soluzione x_0 . Se facciamo delle mosse di Gauss, l'insieme $S = \{x_0\}$ delle soluzioni non cambia.

Siccome $\text{rk}A = n$, l'algoritmo di Gauss - Jordan trasforma $(A \mid e_i)$ in $(I_n \mid B^i)$ e siccome x_0 è ancora soluzione troviamo che adesso $I_n x_0 = B^i$, cioè $B^i = x_0$. Quindi la i -esima colonna di B è uguale all' i -esima colonna di A^{-1} , per ogni i . Quindi $B = A^{-1}$. \square

Come sempre, non è necessario seguire pedissequamente l'algoritmo di Gauss - Jordan: qualsiasi sequenza di mosse di Gauss va bene purché trasformi A in I_n . Ad esempio, determiniamo l'inversa di $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ con mosse di Gauss:

$$\left(\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array} \right) \rightarrow \left(\begin{array}{cc|cc} 1 & 0 & 1 & -1 \\ 1 & 1 & 0 & 1 \end{array} \right) \rightarrow \left(\begin{array}{cc|cc} 1 & 0 & 1 & -1 \\ 0 & 1 & -1 & 2 \end{array} \right)$$

Quindi $A^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$. Nel frattempo, abbiamo dimostrato il fatto seguente.

Proposizione 3.4.10. *Una matrice $A \in M(n)$ è invertibile $\iff \det A \neq 0$.*

Dimostrazione. Se A è invertibile, allora $\det A \neq 0$ per il Teorema di Binet. D'altro canto, se $\det A \neq 0$ allora $\text{rk} A = n$ e l'algoritmo precedente per determinare A^{-1} funziona. \square

Si può calcolare l'inversa di una matrice anche usando il determinante. Data $A \in M(n)$, la *matrice dei cofattori* è la matrice $\text{cof } A \in M(n)$ il cui elemento i, j è

$$\text{cof}(A)_{ij} = (-1)^{i+j} \det C_{ij}$$

dove C_{ij} è il minore di A ottenuto cancellando la riga i -esima e la colonna j -esima. Possiamo scrivere lo sviluppo di Laplace in questo modo:

$$\det A = \sum_{j=1}^n a_{ij} \text{cof}(A)_{ij}.$$

Usiamo adesso la matrice dei cofattori per determinare A^{-1} .

Proposizione 3.4.11. *Se A è invertibile, allora*

$$A^{-1} = \frac{1}{\det A} {}^t(\text{cof } A).$$

Dimostrazione. Dobbiamo dimostrare che

$$A \cdot {}^t(\text{cof } A) = (\det A) I_n.$$

In altre parole, il prodotto riga per colonna

$$A_i \cdot ({}^t \text{cof } A)^j = \sum_{h=1}^n A_{ih} \text{cof}(A)_{jh}$$

deve fare $\det A$ se $i = j$ e 0 se $i \neq j$. Se $i = j$, la sommatoria è semplicemente lo sviluppo di Laplace di A e quindi fa $\det A$. Se $i \neq j$, la sommatoria è lo sviluppo di Laplace della matrice A' ottenuta da A sostituendo A_i al posto di A_j , e fa $\det A' = 0$ perché A' ha due righe uguali. \square

Esempio 3.4.12. La matrice inversa di $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ con l'ipotesi $\det A \neq 0$ è

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

3.4.8. Sistema lineare con A invertibile. Notiamo adesso un fatto semplice ma rilevante. Se

$$Ax = b$$

rappresenta un sistema lineare dove A è una matrice quadrata (quindi ci sono tante variabili quante equazioni), e se A è invertibile, allora possiamo moltiplicare entrambi i membri a sinistra per A^{-1} e ottenere

$$x = A^{-1}b.$$

In questo modo scopriamo immediatamente che il sistema ha una sola soluzione e l'abbiamo determinata. Se invece A non è invertibile ovviamente questo metodo non funziona.

Proposizione 3.4.13 (Regola di Cramer). *Se $\det A \neq 0$, il sistema $Ax = b$ ha un'unica soluzione x , determinata nel modo seguente:*

$$x_i = \frac{\det B_i}{\det A} \quad \forall i = 1, \dots, n$$

dove B_i indica la matrice ottenuta sostituendo la i -esima colonna di A con b .

Dimostrazione. Sappiamo che $x = A^{-1}b$, quindi

$$x_i = (A^{-1})_i \cdot b = \frac{t((\text{cof } A)^i) \cdot b}{\det A} = \frac{\det B_i}{\det A}.$$

L'ultima uguaglianza è lo sviluppo di Laplace di B_i lungo la i -esima colonna. La dimostrazione è conclusa. \square

La regola di Cramer è molto compatta, ma la sua applicazione è spesso più dispendiosa in termini di conti dell'algoritmo di Gauss - Jordan.

Esercizi

Esercizio 3.1. Determina le soluzioni dei sistemi lineari:

$$\left\{ \begin{array}{l} 3x_1 + x_3 - x_4 = 0 \\ x_2 - 2x_3 - x_4 = -4 \\ 4x_1 + x_2 + 2x_3 - 3x_4 = 0 \\ 3x_1 - x_2 + x_4 = 1 \end{array} \right. \quad \left\{ \begin{array}{l} 2x_2 + 2x_3 + 6x_4 + x_5 = 7 \\ 2x_1 + x_2 + 7x_3 + 11x_4 = 4 \\ x_1 + 3x_3 + 4x_4 + x_5 = 4 \end{array} \right.$$

Esercizio 3.2. Determina le soluzioni del sistema lineare seguente, al variare dei parametri $\alpha, \beta \in \mathbb{R}$.

$$\left\{ \begin{array}{l} x + \alpha y + \beta z = 1 \\ 2x + 3\alpha y + 2\beta z = 2 \\ \beta x + \alpha(2 + \beta)y + (\alpha + \beta + \beta^2)z = 2\beta \end{array} \right.$$

Esercizio 3.3. Calcola il determinante delle matrici seguenti:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 5 \\ 3 & 4 & 5 & 5 & 5 \\ 4 & 5 & 5 & 5 & 5 \\ 5 & 5 & 5 & 5 & 5 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 4 & 0 & 0 & 0 \\ 2 & 1 & 5 & 9 & 1 & 0 \\ 11^{12} & 9 & 0 & 0 & 0 & 0 \\ 3 & 1 & 2 & 5 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 4 & -1 & 3 & 1 & 2 \\ 1 & 2 & 2 & 1 & 0 \\ 4 & \sqrt{3} & 4 & 1 & 1 \\ -1 & 1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 1 & 2 \end{pmatrix}.$$

Esercizio 3.4. Considera i sistemi lineari dipendenti da un parametro $c \in \mathbb{R}$:

$$\begin{cases} 2x_1 - x_2 + x_3 + 6x_4 = 13 \\ x_1 + x_2 + 5x_3 + 3x_4 = c^2 - 14 \\ x_2 + 3x_3 + x_4 = c^2 \\ x_1 + 2x_3 + 3x_4 = 7 \end{cases} \quad \begin{cases} x + y + cz = 1 \\ x + y + c^3z = 3 \\ 2x + 2y + (1+c)x = 1 \\ x + z = 0 \end{cases}$$

$$\begin{cases} x - y + 2z = 1 \\ 2x + (c-2)y + 2z = 1 \\ -x + y + (c^2 - c - 2)z = c - 1 \end{cases} \quad \begin{cases} x + cy - 2z = c \\ -x + z = 1 - c \\ -2x - 3cy + (c+4)z = -3 - c \end{cases}$$

Per ciascun sistema, determina per quali $c \in \mathbb{R}$ il sistema ha soluzione e la dimensione dello spazio delle soluzioni al variare di $c \in \mathbb{R}$.

Esercizio 3.5. Calcola l'inversa, se esiste, delle seguenti matrici:

$$\begin{pmatrix} 1 & i \\ -i & -1 \end{pmatrix}, \quad \begin{pmatrix} 3 & -1 & -5 \\ 3 & -1 & 1 \\ 2 & -2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 3 & 3 \\ 1 & 3 & 3 & 5 \end{pmatrix}.$$

Esercizio 3.6. Considera il sottospazio $W = \{x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 = 0\}$ di \mathbb{C}^5 . Determina tre sottospazi $U, V, Z \subset W$, tutti di dimensione almeno 1, tali che $W = U \oplus V \oplus Z$.

Esercizio 3.7. Sia U il sottospazio di \mathbb{R}^4 generato dai vettori

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \quad \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix}.$$

Sia V il sottospazio di \mathbb{R}^4 delle soluzioni del sistema

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_2 + 4x_3 - x_4 = 0 \\ 3x_1 + 5x_2 + 11x_3 + x_4 = 0 \end{cases}$$

Calcola le dimensioni di U , V , $U \cap V$ e $U + V$.

Esercizio 3.8. Al variare di $k \in \mathbb{R}$, sia $U_k \subset \mathbb{R}^4$ il sottospazio formato dalle soluzioni del sistema

$$\begin{cases} x + y + kz + kt = 0 \\ 2x + (2-k)y + 3kz = 0 \\ (2-k)x + 2y + 4kt = 0 \end{cases}$$

Sia inoltre

$$W = \text{Span} \left(\begin{pmatrix} 1 \\ 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 2 \\ 4 \\ 5 \end{pmatrix} \right).$$

- (1) Calcola la dimensione di U_k al variare di $k \in \mathbb{R}$.
- (2) Costruisci una base di $Z = W \cap U_0$.
- (3) Determina un sottospazio $Z' \subset \mathbb{R}^4$ tale che $Z \oplus Z' = \mathbb{R}^4$.

Esercizio 3.9. Sia $A \in M(n)$ una matrice invertibile. Dimostra che anche tA è invertibile e $({}^tA)^{-1} = {}^t(A^{-1})$.

Esercizio 3.10. Data $A \in M(n, \mathbb{K})$ e $\lambda \in \mathbb{K}$, mostra che $\det(\lambda A) = \lambda^n \det A$.

Esercizio 3.11. Dimostra che una matrice $A \in M(17, \mathbb{R})$ antisimmetrica ha determinante nullo.

Esercizio 3.12 (Matrice di Vandermonde). Siano $\alpha_1, \dots, \alpha_n \in \mathbb{K}$. Dimostra che

$$\det \begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \cdots & \alpha_n^{n-1} \end{pmatrix} = \prod_{1 \leq j < i \leq n} (\alpha_j - \alpha_i).$$

Esercizio 3.13. Sia $A \in M(n)$ una matrice a blocchi del tipo

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

dove B e D sono sottomatrici quadrate di ordine $k \times k$ e $(n - k) \times (n - k)$. La sottomatrice 0 in basso a sinistra è tutta nulla e di ordine $(n - k) \times k$. Dimostra che

$$\det A = \det B \cdot \det D.$$

Suggerimento. Dimostra l'esercizio per induzione su n ed effettuando lo sviluppo di Laplace sulla prima colonna. \square

Esercizio 3.14. Sia A una matrice a blocchi come nell'esercizio precedente:

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

Dimostra che $\text{rk}A \geq \text{rk}B + \text{rk}D$. Costruisci degli esempi in cui vale l'uguaglianza e degli esempi in cui vale la disuguaglianza stretta.

Complementi

CAPITOLO 4

Applicazioni lineari

4.1. Introduzione

Un'applicazione lineare è una funzione fra spazi vettoriali compatibile con le operazioni di somma e prodotto per scalare.

4.1.1. Definizione. Siano V e W due spazi vettoriali sullo stesso campo \mathbb{K} . Una funzione

$$f: V \longrightarrow W$$

è *lineare* se valgono gli assiomi seguenti:

- (1) $f(0) = 0$.
- (2) $f(v + w) = f(v) + f(w)$ per ogni $v, w \in V$.
- (3) $f(\lambda v) = \lambda f(v)$ per ogni $v \in V$ e $\lambda \in \mathbb{K}$.

Notiamo che in (1) il primo 0 è l'origine di V ed il secondo è l'origine di W . Quindi f deve mandare l'origine di V nell'origine di W .

4.1.2. Combinazioni lineari. Dagli assiomi di funzione lineare possiamo subito dedurre un fatto importante. Se $f: V \rightarrow W$ è lineare ed un vettore $v \in V$ è espresso come combinazione lineare di alcuni vettori di V :

$$v = \lambda_1 v_1 + \cdots + \lambda_k v_k$$

allora, poiché f è lineare, troviamo che la sua immagine è

$$\begin{aligned} f(v) &= f(\lambda_1 v_1 + \cdots + \lambda_k v_k) = f(\lambda_1 v_1) + \cdots + f(\lambda_k v_k) \\ &= \lambda_1 f(v_1) + \cdots + \lambda_k f(v_k). \end{aligned}$$

Notiamo quindi che f manda una qualsiasi combinazione lineare di vettori v_1, \dots, v_k in una combinazione lineare (con gli stessi coefficienti) delle loro immagini $f(v_1), \dots, f(v_k)$.

4.1.3. Esempi basilari. Facciamo alcuni esempi importanti.

Esempio 4.1.1 (Funzione nulla e identità). Dati due spazi vettoriali V, W su \mathbb{K} , la *funzione nulla* è la funzione $f: V \rightarrow W$ che è costantemente nulla, cioè tale che $f(v) = 0$ per ogni v . La funzione nulla è lineare: i 3 assiomi di linearità per f sono soddisfatti in modo banale.

Dato uno spazio vettoriale V , la *funzione identità* è la funzione $\text{id}_V: V \rightarrow V$ che manda ogni vettore in se stesso, cioè tale che $\text{id}_V(v) = v \forall v \in V$. La

funzione identità è lineare: anche qui i 3 assiomi sono ovviamente soddisfatti. A volte sottointendiamo V e indichiamo id_V con id .

Esempio 4.1.2. La funzione $f: \mathbb{R} \rightarrow \mathbb{R}$ data da $f(x) = 3x$ è lineare. Infatti:

- (1) $f(0) = 0$,
- (2) $f(x + x') = 3(x + x') = 3x + 3x' = f(x) + f(x') \quad \forall x, x' \in \mathbb{R}$,
- (3) $f(\lambda x) = 3\lambda x = \lambda 3x = \lambda f(x) \quad \forall x \in \mathbb{R}, \forall \lambda \in \mathbb{R}$.

Più in generale, la funzione $f(x) = kx$ è lineare per ogni $k \in \mathbb{R}$ fissato. Notiamo che per $k = 0$ otteniamo la funzione nulla $f(x) = 0$ e per $k = 1$ la funzione identità $f(x) = x$, già considerate nell'esempio precedente. Invece non sono lineari le seguenti funzioni da \mathbb{R} in \mathbb{R} :

$$f(x) = 2x + 1, \quad f(x) = x^2.$$

Per la prima non vale l'assioma (1), per la seconda non vale l'assioma (2). Effettivamente dimostreremo che le uniche funzioni lineari da \mathbb{R} in \mathbb{R} sono quelle del tipo $f(x) = kx$. Gli assiomi (1), (2) e (3) sono piuttosto restrittivi.

Esempio 4.1.3. Più in generale, sia V uno spazio vettoriale e $\lambda \in \mathbb{K}$ uno scalare fissato. La funzione $f: V \rightarrow V$ data da $f(v) = \lambda v$ è lineare. Se $\lambda = 0$ otteniamo la funzione nulla e se $\lambda = 1$ otteniamo l'identità.

Esempio 4.1.4. Generalizziamo l'Esempio 4.1.2 prendendo un campo \mathbb{K} qualsiasi e alzando la dimensione del dominio. Siano $a_1, \dots, a_n \in \mathbb{K}$ costanti fissate. La funzione

$$f: \mathbb{K}^n \rightarrow \mathbb{K}, \quad f(x) = a_1 x_1 + \dots + a_n x_n$$

è lineare. Infatti

- (1) $f(0) = a_1 0 + \dots + a_n 0 = 0$,
- (2) $f(x + x') = a_1(x_1 + x'_1) + \dots + a_n(x_n + x'_n) = a_1 x_1 + \dots + a_n x_n + a_1 x'_1 + \dots + a_n x'_n = f(x) + f(x')$.
- (3) $f(\lambda x) = a_1 \lambda x_1 + \dots + a_n \lambda x_n = \lambda(a_1 x_1 + \dots + a_n x_n) = \lambda f(x)$.

Ad esempio la funzione $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f\begin{pmatrix} x \\ y \end{pmatrix} = 2x - 3y$ è lineare.

Esempio 4.1.5. Vogliamo generalizzare ulteriormente l'esempio precedente alzando la dimensione del codominio. Prendiamo una matrice $A = (a_{ij})$ di taglia $m \times n$ e definiamo

$$L_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$$

usando il prodotto fra matrici e vettori nel modo seguente:

$$L_A(x) = Ax.$$

Nel dettaglio:

$$L_A(x) = Ax = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix}.$$

La lettera L sta per "left" e ci ricorda che L_A è l'applicazione che prende come input un vettore $x \in \mathbb{K}^n$ e lo moltiplica a sinistra per A , restituendo come output $Ax \in \mathbb{K}^m$. La linearità di L_A discende dalle proprietà delle matrici, elencate nella Proposizione 3.4.1. Dimostriamo infatti che:

- (1) $L_A(0) = A0 = 0$.
- (2) $L_A(x + x') = A(x + x') = Ax + Ax' = L_A(x) + L_A(x')$.
- (3) $L_A(\lambda x) = A(\lambda x) = \lambda Ax = \lambda L_A(x)$.

Ad esempio, se $A = \begin{pmatrix} 3 & -1 \\ 1 & 2 \end{pmatrix}$ otteniamo la funzione $L_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ data da:

$$L_A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x - y \\ x + 2y \end{pmatrix}.$$

È spesso utile scrivere l'applicazione L_A anche in una forma alternativa che usa le colonne A^1, \dots, A^n di A .

Proposizione 4.1.6. *Vale l'uguaglianza seguente:*

$$L_A(x) = x_1 A^1 + \dots + x_n A^n.$$

Dimostrazione. Si vede immediatamente che

$$\begin{aligned} L_A(x) &= \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix} = x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \\ &= x_1 A^1 + \dots + x_n A^n. \end{aligned}$$

La dimostrazione è conclusa. □

Ricordiamo che e_1, \dots, e_n è la base canonica di \mathbb{K}^n . Ne deduciamo che le colonne di A sono precisamente le immagini dei vettori e_1, \dots, e_n .

Corollario 4.1.7. *La colonna A^i è l'immagine di e_i , cioè*

$$L_A(e_i) = A^i \quad \forall i = 1, \dots, n.$$

Dimostrazione. Il vettore e_i ha coordinate $x_j = 0$ per $j \neq i$ e $x_i = 1$. □

Esaminando le colonne della matrice A si vede subito dove va la base canonica. Ad esempio, se

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

guardando le colonne capiamo subito che la applicazione L_A scambia i due vettori e_1 ed e_2 della base canonica.

Esempio 4.1.8. Ci sono due casi molto particolari che è bene capire subito:

- Se $A = 0$ è la matrice $m \times n$ nulla, allora $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$ è l'applicazione nulla.
- Se $A = I_n$ è la matrice identità $n \times n$, allora $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^n$ è l'identità.

Infatti $L_0(x) = 0x = 0$ e $L_I(x) = Ix = x$ per ogni $x \in \mathbb{K}^n$.

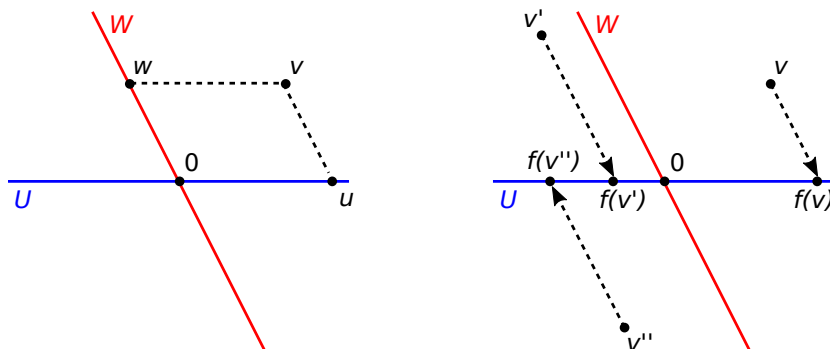


Figura 4.1. La proiezione $p_U: V \rightarrow U$ su un sottospazio determinata da una somma diretta $V = U \oplus W$.

4.1.4. Esempi più elaborati. Costruiamo adesso degli esempi di applicazioni lineari più sofisticati.

Esempio 4.1.9 (L_A sulle matrici). Sia A una matrice $m \times n$. L'applicazione

$$L_A: M(n, k) \longrightarrow M(m, k)$$

definita da

$$L_A(X) = AX$$

è una applicazione lineare. Notiamo che adesso X è una matrice $n \times k$ e $L_A(X) = AX$ è una matrice $m \times k$. La linearità segue dalle proprietà algebriche delle matrici dimostrate nella Proposizione 3.4.1:

- (1) $L_A(0) = A0 = 0$.
- (2) $L_A(X + X') = A(X + X') = AX + AX' = L_A(X) + L_A(X')$.
- (3) $L_A(\lambda X) = A(\lambda X) = \lambda AX = \lambda L_A(X)$.

Quando $k = 1$ questa applicazione lineare è la $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$ già vista.

Esempio 4.1.10 (Trasposizione di matrici). La mappa

$$f: M(m, n) \longrightarrow M(n, m)$$

$$f(A) = {}^tA$$

che prende una matrice e la traspone è una applicazione lineare, perché

$${}^t0 = 0, \quad {}^t(A + B) = {}^tA + {}^tB, \quad {}^t(\lambda A) = \lambda {}^tA.$$

Esempio 4.1.11 (Proiezione). Sia $V = U \oplus W$ uno spazio vettoriale espresso come somma diretta dei sottospazi U e W . Definiamo una applicazione lineare

$$p_U: V \longrightarrow U$$

detta *proiezione* su U nel modo seguente: per la Proposizione 2.3.27, ogni vettore $v \in V$ si scrive in modo unico come $u + w$ con $u \in U$ e $w \in W$, e noi definiamo

$$p_U(v) = u.$$

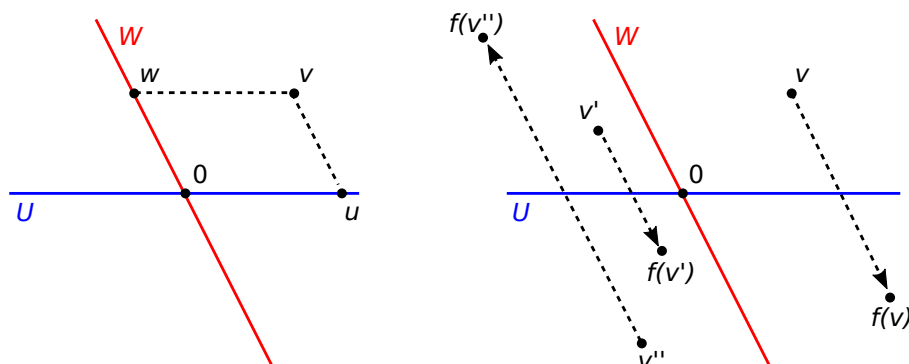


Figura 4.2. La riflessione $r_U: V \rightarrow V$ lungo un sottospazio determinata da una somma diretta $V = U \oplus W$.

Intuitivamente, lo spazio U è il luogo su cui si proiettano tutti i punti, mentre il sottospazio complementare W indica la *direzione* lungo cui i punti vengono proiettati, come mostrato nella Figura 4.1.

La proiezione è effettivamente lineare: se $v = u + w$ e $v' = u' + w'$ allora $v + v' = u + u' + w + w'$ e quindi $p_U(v + v') = u + u' = p_U(v) + p_U(v')$. Le altre due proprietà sono dimostrate in modo analogo.

Scriviamo la proiezione in un caso concreto, quello della somma diretta $\mathbb{R}^2 = U \oplus W$ già analizzata nell'Esempio 2.3.31. Abbiamo

$$U = \text{Span} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad W = \text{Span} \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

Dalla discussione dell'Esempio 2.3.31 segue che

$$p_U \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + \frac{y}{2} \\ 0 \end{pmatrix}.$$

Esempio 4.1.12 (Riflessione). Sia $V = U \oplus W$ come prima. La *riflessione* lungo U è l'applicazione lineare

$$r_U: V \rightarrow V$$

definita nel modo seguente: ogni vettore $v \in V$ si scrive in modo unico come $u + w$ con $u \in U$ e $w \in W$, e noi definiamo

$$r_U(v) = u - w.$$

Intuitivamente, lo spazio U è lo specchio lungo cui si riflettono i punti, mentre W è la direzione lungo cui riflettiamo. Si veda la Figura 4.2.

Scriviamo la riflessione con la somma diretta $\mathbb{R}^2 = U \oplus W$ dell'Esempio 2.3.31. Dai conti fatti precedentemente segue che

$$r_U \begin{pmatrix} x \\ y \end{pmatrix} = \left(x + \frac{y}{2}\right) \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{y}{2} \begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} x + y \\ -y \end{pmatrix}.$$

Esempio 4.1.13 (Valore di un polinomio in un punto). Sia $x_0 \in \mathbb{K}$ un valore fissato. Consideriamo l'applicazione

$$T: \mathbb{K}[x] \longrightarrow \mathbb{K}$$

$$T(p) = p(x_0)$$

che associa ad ogni polinomio p il suo valore $p(x_0)$ nel punto x_0 . L'applicazione T è lineare, infatti:

- (1) $T(0) = 0(x_0) = 0$,
- (2) $T(p + q) = (p + q)(x_0) = p(x_0) + q(x_0) = T(p) + T(q)$
- (3) $T(\lambda p) = \lambda p(x_0) = \lambda T(p)$.

Esempio 4.1.14 (Valori di una funzione in alcuni punti). Generalizziamo l'esempio precedente: invece che un punto solo ne prendiamo tanti, e invece che polinomi consideriamo funzioni più generali.

Sia X un insieme e $x_1, \dots, x_n \in X$ dei punti fissati. L'applicazione

$$T: F(X, \mathbb{K}) \longrightarrow \mathbb{K}^n$$

$$T(f) = \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix}.$$

associa ad una funzione $f: X \rightarrow \mathbb{K}$ il vettore formato dai suoi valori nei punti x_1, \dots, x_n . Questa funzione è lineare (esercizio).

Esempio 4.1.15 (Derivata di un polinomio). La funzione

$$T: \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$$

$$T(p) = p'$$

che associa ad un polinomio

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

la sua derivata

$$p'(x) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \dots + a_1$$

è lineare. Si può dimostrare questo fatto direttamente oppure invocare il fatto che la derivata è sempre lineare per dei teoremi di analisi: se f' è la derivata di una funzione f valgono sempre le proprietà di linearità, che sono:

$$0' = 0, \quad (f + g)' = f' + g', \quad (\lambda f)' = \lambda f'.$$

Esempio 4.1.16 (Restrizioni). Se $f: V \rightarrow W$ è una applicazione lineare e $U \subset V$ è un sottospazio, possiamo costruire una applicazione lineare

$$f|_U: U \longrightarrow W$$

semplicemente restringendo il dominio di f da V a solo U . L'applicazione ristretta è indicata con $f|_U$ ed è chiamata la *restrizione* di f a U .

Esempio 4.1.17 (Coordinate di un vettore rispetto ad una base). Sia V uno spazio vettoriale su \mathbb{K} e $\mathcal{B} = \{v_1, \dots, v_n\}$ una base¹ per V . Per ciascun $v \in V$, indichiamo con

$$[v]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{K}^n$$

il vettore colonna formato dalle coordinate $\lambda_1, \dots, \lambda_n$ di v rispetto alla base \mathcal{B} , definite nella Sezione 2.3.3. Otteniamo in questo modo una funzione

$$T: V \longrightarrow \mathbb{K}^n, \quad T(v) = [v]_{\mathcal{B}}.$$

Verifichiamo che la funzione T è lineare:

- (1) $T(0) = 0$,
- (2) $T(v + v') = T(v) + T(v')$, infatti se $v = \lambda_1 v_1 + \dots + \lambda_n v_n$ e $v' = \lambda'_1 v_1 + \dots + \lambda'_n v_n$ allora $v + v' = (\lambda_1 + \lambda'_1)v_1 + \dots + (\lambda_n + \lambda'_n)v_n$.
- (3) $T(\lambda v) = \lambda T(v)$, infatti se $v = \lambda_1 v_1 + \dots + \lambda_n v_n$ allora ovviamente $\lambda v = \lambda \lambda_1 v_1 + \dots + \lambda \lambda_n v_n$.

4.1.5. Uso delle basi. Siano V e W due spazi vettoriali su \mathbb{K} arbitrari. Cerchiamo adesso un procedimento generale per scrivere tutte le applicazioni lineari da V in W . Come in altre parti di questo libro, lo strumento chiave è la scelta di una base per V .

La proposizione seguente dice che per costruire una applicazione lineare $V \rightarrow W$ è sufficiente fissare i valori su una qualsiasi base fissata per V .

Proposizione 4.1.18. *Sia v_1, \dots, v_n una base per V . Per ogni w_1, \dots, w_n esiste ed è unica una applicazione lineare*

$$f: V \longrightarrow W$$

tale che $f(v_i) = w_i \forall i = 1, \dots, n$.

Dimostrazione. Mostriamo l'esistenza. Costruiamo f nel modo seguente: ogni $v \in V$ si scrive in modo unico come

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n$$

e noi definiamo $f: V \rightarrow W$ in modo molto naïf, sostituendo v_i con w_i :

$$f(v) = \lambda_1 w_1 + \dots + \lambda_n w_n.$$

¹Ad essere precisi, la notazione matematica $\{\}$ tra parentesi graffe indica un insieme *non ordinato* di vettori, mentre una base è in realtà un insieme *ordinato*. Con questo vogliamo dire che per noi sarà importante dire che v_1 è il primo vettore, v_2 il secondo, ecc. In particolare $\{v_1, v_2\}$ e $\{v_2, v_1\}$ sono due basi diverse. Generalmente gli insiemi ordinati vengono indicati con parentesi tonde $()$ invece che graffe $\{\}$, e quindi volendo essere rigorosi dovremmo scrivere $\mathcal{B} = (v_1, \dots, v_n)$. Preferiamo però usare comunque le graffe perché le tonde sono già usate abbondantemente per i vettori numerici e le matrici.

Si vede facilmente che questa funzione f è effettivamente lineare. D'altro canto, se $f: V \rightarrow W$ è una funzione lineare, otteniamo

$$f(v) = \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n) = \lambda_1 w_1 + \cdots + \lambda_n w_n,$$

quindi la f definita prima è l'unica applicazione lineare che rispetta le condizioni date. La dimostrazione è conclusa. \square

4.1.6. Applicazioni da \mathbb{K}^n a \mathbb{K}^m . Dimostriamo adesso un fatto a cui abbiamo già accennato: le applicazioni L_A determinate da matrici $A \in M(m, n, \mathbb{K})$ sono tutte e sole le applicazioni lineari da \mathbb{K}^n in \mathbb{K}^m .

Proposizione 4.1.19. *Per ogni $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ lineare, esiste un'unica $A \in M(m, n, \mathbb{K})$ tale che $f = L_A$.*

Dimostrazione. Sia $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ una funzione lineare. Ricordiamo che e_1, \dots, e_n è la base canonica di \mathbb{K}^n . Sia A la matrice la cui colonna A^i è uguale all'immagine $f(e_i)$ di e_i . Brevemente:

$$A = (f(e_1) \cdots f(e_n)).$$

Per il Corollario 4.1.7, per ogni $i = 1, \dots, n$ otteniamo

$$L_A(e_i) = A^i = f(e_i).$$

Poiché e_1, \dots, e_n è una base, per la Proposizione 4.1.18 abbiamo $f = L_A$.

Per quanto riguarda l'unicità, se $A \neq A'$ allora $A_{ij} \neq A'_{ij}$ per qualche i, j e quindi $L_A(e_j) \neq L_{A'}(e_j)$. Quindi A diverse danno L_A diverse. \square

4.2. Nucleo e immagine

Ad un'applicazione lineare vengono naturalmente associati due sottospazi vettoriali del dominio e del codominio, detti il *nucleo* e l'*immagine*.

4.2.1. Nucleo. Sia $f: V \rightarrow W$ un'applicazione lineare. Il *nucleo* di f è il sottoinsieme di V definito nel modo seguente:

$$\ker f = \{v \in V \mid f(v) = 0\}.$$

Il nucleo è l'insieme dei vettori che vengono mandati in zero da f . Il termine "ker" è un'abbreviazione dell'inglese *kernel*, sinonimo di nucleo.

Proposizione 4.2.1. *Il nucleo $\ker f$ è un sottospazio vettoriale di V .*

Dimostrazione. Dobbiamo verificare i 3 assiomi di sottospazio.

- (1) $0 \in \ker f$, infatti $f(0) = 0$.
- (2) $v, w \in \ker f \Rightarrow v + w \in \ker f$, infatti $f(v + w) = f(v) + f(w) = 0 + 0 = 0$.
- (3) $v \in \ker f, \lambda \in \mathbb{K} \Rightarrow \lambda v \in \ker f$, infatti $f(\lambda v) = \lambda f(v) = \lambda 0 = 0$.

La dimostrazione è conclusa. \square

Ricordiamo che una funzione $f: V \rightarrow W$ è iniettiva se $v \neq v' \Rightarrow f(v) \neq f(v')$. Capire se un'applicazione lineare è iniettiva è abbastanza facile: è sufficiente esaminare il nucleo. Sia $f: V \rightarrow W$ un'applicazione lineare.

Proposizione 4.2.2. *La funzione $f: V \rightarrow W$ è iniettiva $\iff \ker f = \{0\}$.*

Dimostrazione. (\Rightarrow) Sappiamo già che $f(0) = 0$. Siccome f è iniettiva, $v \neq 0 \Rightarrow f(v) \neq 0$. Quindi $\ker f = \{0\}$.

(\Leftarrow) Siano $v, v' \in V$ diversi. Siccome $\ker f = \{0\}$, troviamo che $v - v' \neq 0 \Rightarrow f(v) - f(v') = f(v - v') \neq 0$, e quindi $f(v) \neq f(v')$. \square

Esempio 4.2.3. Sia A una matrice $m \times n$. Il nucleo dell'applicazione $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$ è precisamente l'insieme S delle soluzioni del sistema lineare omogeneo $Ax = 0$. Quindi L_A è iniettiva precisamente se $S = \{0\}$.

Esempio 4.2.4. Consideriamo una somma diretta $V = U \oplus W$. Il nucleo della proiezione p_U definita nell'Esempio 4.1.11 è precisamente W .

4.2.2. Immagine. Sia $f: V \rightarrow W$ un'applicazione lineare. Come tutte le funzioni, la f ha una immagine che indichiamo con $\text{Im } f \subset W$.

Proposizione 4.2.5. *L'immagine $\text{Im } f$ è un sottospazio vettoriale di W .*

Dimostrazione. Dimostriamo come sempre i 3 assiomi.

(1) $0 \in \text{Im } f$ perché $f(0) = 0$.

(2) $f(v), f(v') \in \text{Im } f \Rightarrow f(v) + f(v') \in \text{Im } f$ perché $f(v) + f(v') = f(v + v')$.

(3) $f(v) \in \text{Im } f, \lambda \in \mathbb{K} \Rightarrow \lambda f(v) \in \text{Im } f$ perché $\lambda f(v) = f(\lambda v)$.

La dimostrazione è conclusa. \square

Ovviamente f è surgettiva $\iff \text{Im } f = W$. Concretamente, possiamo spesso determinare l'immagine di f usando la proposizione seguente.

Proposizione 4.2.6. *Se v_1, \dots, v_n sono generatori di V , allora*

$$\text{Im } f = \text{Span}(f(v_1), \dots, f(v_n)).$$

Dimostrazione. Ogni vettore $v \in V$ si scrive come $v = \lambda_1 v_1 + \dots + \lambda_n v_n$ per qualche $\lambda_1, \dots, \lambda_n$. Quindi

$$\begin{aligned} \text{Im } f &= \{f(\lambda_1 v_1 + \dots + \lambda_n v_n) \mid \lambda_i \in \mathbb{K}\} \\ &= \{\lambda_1 f(v_1) + \dots + \lambda_n f(v_n) \mid \lambda_i \in \mathbb{K}\} \\ &= \text{Span}(f(v_1), \dots, f(v_n)). \end{aligned}$$

La dimostrazione è conclusa. \square

Corollario 4.2.7. *Sia A una matrice $m \times n$. L'immagine dell'applicazione $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$ è precisamente lo spazio generato dalle colonne:*

$$\text{Im } L_A = \text{Span}(A^1, \dots, A^n).$$

Dimostrazione. Sappiamo che $A^i = L_A(e_j)$ e e_1, \dots, e_n generano \mathbb{K}^n . \square

Corollario 4.2.8. Per ogni matrice A vale $\text{rk}(A) = \dim \text{Im } L_A$.

4.2.3. Teorema della dimensione. Dimostriamo adesso un teorema centrale in algebra lineare.

Teorema 4.2.9 (Teorema della dimensione). Sia $f: V \rightarrow W$ una funzione lineare. Se V ha dimensione finita n , allora

$$\dim \ker f + \dim \text{Im } f = n.$$

Dimostrazione. Sia v_1, \dots, v_k una base di $\ker f$. Possiamo sempre completarla a base v_1, \dots, v_n di V . Il nostro scopo è dimostrare che $f(v_{k+1}), \dots, f(v_n)$ formano una base di $\text{Im } f$. Se ci riusciamo, abbiamo finito perché otteniamo $\dim \ker f = k$ e $\dim \text{Im } f = n - k$.

Dimostriamo che $f(v_{k+1}), \dots, f(v_n)$ generano $\text{Im } f$. Sappiamo dalla Proposizione 4.2.6 che $f(v_1), \dots, f(v_n)$ generano $\text{Im } f$. Sappiamo anche che $f(v_1) = \dots = f(v_k) = 0$, quindi possiamo rimuoverli dalla lista e ottenere che effettivamente $f(v_{k+1}), \dots, f(v_n)$ generano $\text{Im } f$.

Dimostriamo che $f(v_{k+1}), \dots, f(v_n)$ sono indipendenti. Supponiamo che

$$\lambda_{k+1}f(v_{k+1}) + \dots + \lambda_n f(v_n) = 0.$$

Dobbiamo mostrare che $\lambda_{k+1} = \dots = \lambda_n = 0$. Per linearità di f , otteniamo

$$f(\lambda_{k+1}v_{k+1} + \dots + \lambda_n v_n) = 0.$$

Quindi $\lambda_{k+1}v_{k+1} + \dots + \lambda_n v_n \in \ker f$. Siccome v_1, \dots, v_k è base di $\ker f$, possiamo scrivere

$$\lambda_{k+1}v_{k+1} + \dots + \lambda_n v_n = \alpha_1 v_1 + \dots + \alpha_k v_k$$

per qualche α_j . Spostando tutto a sinistra,

$$-\alpha_1 v_1 - \dots - \alpha_k v_k + \lambda_{k+1}v_{k+1} + \dots + \lambda_n v_n = 0.$$

I vettori v_1, \dots, v_n però sono indipendenti, quindi i coefficienti α_j, λ_j sono tutti nulli. La dimostrazione è conclusa. \square

Nel caso particolare di una applicazione lineare $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$, il teorema della dimensione era già stato dimostrato con il Teorema di Rouché - Capelli per i sistemi lineari omogenei nella Sezione 3.2.5. Abbiamo quindi due dimostrazioni molto diverse dello stesso teorema: una usa le matrici e le mosse di Gauss, l'altra direttamente i vettori ed è più generale.

Esempio 4.2.10. Nella proiezione $p_U: V \rightarrow U$ definita in 4.1.11 a partire da una somma diretta $V = U \oplus W$, troviamo

$$\ker p_U = W, \quad \text{Im } p_U = U.$$

Quindi $V = U \oplus W = \text{Im } p_U \oplus \ker p_U$ e $\dim V = \dim \text{Im } p_U + \dim \ker p_U$.

Esempio 4.2.11. Se $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, per $L_A: \mathbb{K}^2 \rightarrow \mathbb{K}^2$ troviamo:

$$\text{Im } L_A = \text{Span}(A^1, A^2) = \text{Span}(e_1),$$

$$\ker L_A = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{K}^2 \mid y = 0 \right\} = \text{Span}(e_1).$$

Quindi $\dim \ker L_A + \dim \text{Im } L_A = 1 + 1 = 2 = \dim \mathbb{K}^2$.

Nei prossimi esempi vediamo come usare il Teorema della dimensione per determinare più agevolmente il nucleo o immagine di una applicazione lineare.

Esempio 4.2.12. Sia $f: \mathbb{K}_2[x] \rightarrow \mathbb{K}$ la mappa $f(p) = p(2)$ dell'Esempio 4.1.13. Vogliamo determinare una base di $\ker f$.

Per ogni $\lambda \in \mathbb{K}$ prendiamo il polinomio costante $p(x) = \lambda \in \mathbb{K}$ e notiamo che $f(p) = \lambda$. Quindi f è suriettiva. Quindi $\dim \text{Im } f = 1$ e allora $\dim \ker f = \dim \mathbb{K}_2[x] - \dim \text{Im } f = 3 - 1 = 2$.

Sappiamo che $x-2, x^2-2x \in \ker f$. Sono indipendenti, quindi $x-2, x^2-2x$ sono una base di $\ker f$ per la Proposizione 2.3.22.

Esempio 4.2.13. Sia $T: \mathbb{R}_n[x] \rightarrow \mathbb{R}_n[x]$ la funzione derivata $T(p) = p'$ già considerata nell'Esempio 4.1.15. Vogliamo determinare l'immagine di T .

Per definizione $\ker T = \{p(x) \mid p' = 0\}$ è la retta formata dai polinomi costanti. Quindi $\dim \ker T = 1$ e allora $\dim \text{Im } T = \dim \mathbb{K}_n[x] - \dim \ker T = n+1-1 = n$. Per costruzione l'immagine è contenuta nel sottospazio $\mathbb{R}_{n-1}[x]$ formato dai polinomi di grado $\leq n-1$, che ha anche lui dimensione n . Quindi

$$\text{Im } T = \mathbb{R}_{n-1}[x].$$

Possiamo usare il Teorema della dimensione per dedurre delle caratterizzazioni dell'iniettività e della suriettività.

Corollario 4.2.14. Sia $f: V \rightarrow W$ un'applicazione lineare. Vale

$$\dim \text{Im } f \leq \dim V.$$

Inoltre:

- (1) f iniettiva $\iff \dim \text{Im } f = \dim V$.
- (2) f suriettiva $\iff \dim \text{Im } f = \dim W$.

Dimostrazione. Ricordiamo che

$$f \text{ iniettiva} \iff \ker f = \{0\} \iff \dim \ker f = 0,$$

$$f \text{ suriettiva} \iff \text{Im } f = W \iff \dim \text{Im } f = \dim W.$$

Per il Teorema della dimensione $\dim \ker f = 0 \iff \dim \text{Im } f = \dim V$. \square

Corollario 4.2.15. Sia $f: V \rightarrow W$ un'applicazione lineare e v_1, \dots, v_n una base per V . Valgono i fatti seguenti:

- f iniettiva $\iff f(v_1), \dots, f(v_n)$ sono indipendenti.
- f suriettiva $\iff f(v_1), \dots, f(v_n)$ generano W .

Esempio 4.2.16. Sia $A \in M(m, n, \mathbb{K})$. La funzione $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$ è

- iniettiva $\iff \text{rk}A = n$,
- suriettiva $\iff \text{rk}A = m$.

4.2.4. Composizione di applicazioni lineari. La composizione di due applicazioni lineari è sempre lineare.

Proposizione 4.2.17. Se $f: V \rightarrow W$ e $g: W \rightarrow Z$ sono funzioni lineari, anche la composizione $g \circ f: V \rightarrow Z$ lo è.

Dimostrazione. Valgono i 3 assiomi di linearità per $g \circ f$:

- (1) $g(f(0)) = g(0) = 0$.
- (2) $v, v' \in V \Rightarrow g(f(v + v')) = g(f(v) + f(v')) = g(f(v)) + g(f(v'))$.
- (3) $v \in V, \lambda \in \mathbb{K} \Rightarrow g(f(\lambda v)) = g(\lambda f(v)) = \lambda g(f(v))$.

La dimostrazione è conclusa. □

Per le applicazioni di tipo L_A la composizione corrisponde precisamente al prodotto fra matrici:

Proposizione 4.2.18. Siano $A \in M(k, m, \mathbb{K})$ e $B \in M(m, n, \mathbb{K})$. Consideriamo $L_A: \mathbb{K}^m \rightarrow \mathbb{K}^k$ e $L_B: \mathbb{K}^n \rightarrow \mathbb{K}^m$. Vale la relazione

$$L_A \circ L_B = L_{AB}.$$

Dimostrazione. Segue dall'associatività del prodotto fra matrici:

$$L_A(L_B(x)) = A(B(x)) = (AB)(x) = L_{AB}(x).$$

La dimostrazione è conclusa. □

4.2.5. Isomorfismi. Introduciamo una definizione.

Definizione 4.2.19. Un'applicazione lineare $f: V \rightarrow W$ è un *isomorfismo* se è bigettiva.

Ricordiamo che una funzione f è bigettiva se e solo se è contemporaneamente iniettiva e suriettiva. Sappiamo che se $f: V \rightarrow W$ è bigettiva, esiste una inversa $f^{-1}: W \rightarrow V$.

Proposizione 4.2.20. Se f è bigettiva, l'inversa f^{-1} è anch'essa lineare.

Dimostrazione. Valgono i 3 assiomi:

- (1) Chiaramente $f^{-1}(0) = 0$.
- (2) $f^{-1}(w + w') = f^{-1}(w) + f^{-1}(w')$. Infatti scriviamo $v = f^{-1}(w)$, $v' = f^{-1}(w')$ e notiamo che $f(v + v') = f(v) + f(v') = w + w'$. Quindi $f^{-1}(w + w') = v + v' = f^{-1}(w) + f^{-1}(w')$.
- (3) $f^{-1}(\lambda w) = \lambda f^{-1}(w)$. Infatti scriviamo $v = f^{-1}(w)$ e notiamo che $f(\lambda v) = \lambda f(v) = \lambda w$. Quindi $f^{-1}(\lambda w) = \lambda v = \lambda f^{-1}(w)$.

La dimostrazione è conclusa. □

Ricordiamo che una matrice quadrata $A \in M(n)$ è invertibile se ha un'inversa A^{-1} per il prodotto. C'è uno stretto legame fra matrici invertibili e isomorfismi.

Proposizione 4.2.21. *Una applicazione lineare $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^n$ è un isomorfismo $\iff A$ è invertibile.*

Dimostrazione. L'applicazione lineare L_A è un isomorfismo se e solo se esiste una L_B che la inverte, cioè tale $L_A \circ L_B = L_B \circ L_A$ sia l'identità. Tradotto in matrici vuol dire

$$AB = BA = I$$

che è esattamente la nozione di invertibilità per A . □

Dato uno spazio vettoriale V ed una base \mathcal{B} per V , ricordiamo la mappa $T: V \rightarrow \mathbb{K}^n$ che associa ad ogni $v \in V$ le sue coordinate $T(v) = [v]_{\mathcal{B}}$, definita nell'Esempio 4.1.17. La mappa T è un esempio importante di isomorfismo.

Proposizione 4.2.22. *La mappa $T: V \rightarrow \mathbb{K}^n$ è un isomorfismo.*

Dimostrazione. Se $\mathcal{B} = \{v_1, \dots, v_n\}$, l'inversa $T^{-1}: \mathbb{K}^n \rightarrow V$ è

$$T^{-1}(x) = x_1 v_1 + \dots + x_n v_n.$$

Si verifica facilmente che effettivamente $T^{-1} \circ T = \text{id}_V$ e $T \circ T^{-1} = \text{id}_{\mathbb{K}^n}$. □

Il fatto seguente segue dal Corollario 4.2.14.

Corollario 4.2.23. *Sia $f: V \rightarrow W$ un'applicazione lineare.*

- (1) *Se f è iniettiva, allora $\dim V \leq \dim W$.*
- (2) *Se f è suriettiva, allora $\dim V \geq \dim W$.*
- (3) *Se f è un isomorfismo, allora $\dim V = \dim W$.*

La proposizione seguente è analoga alle Proposizioni 2.3.22 e 2.3.28, ed è utile concretamente per dimostrare che una data f è un isomorfismo.

Proposizione 4.2.24. *Sia $f: V \rightarrow W$ un'applicazione lineare. Se $\dim V = \dim W$, allora i fatti seguenti sono equivalenti:*

- (1) *f è iniettiva,*
- (2) *f è suriettiva,*
- (3) *f è un isomorfismo.*

Dimostrazione. Per il Teorema della dimensione, se $n = \dim V = \dim W$,

$$f \text{ iniettiva} \iff \dim \ker f = 0 \iff \dim \text{Im } f = n \iff f \text{ suriettiva.}$$

La dimostrazione è conclusa. □

Osservazione 4.2.25. Notiamo delle forti analogie fra le Proposizioni 2.3.22, 2.3.28 e 4.2.24. In tutti e tre i casi abbiamo una definizione (base, somma diretta, isomorfismo) che prevede due condizioni (indipendenti e generatori, somma totale e intersezione banale, iniettiva e suriettiva), ed un criterio che dice che *se le dimensioni sono giuste* allora una qualsiasi delle due condizioni in realtà implica l'altra.

Esempio 4.2.26. La trasposizione $f: M(m, n, \mathbb{K}) \rightarrow M(n, m, \mathbb{K})$, $f(A) = {}^tA$ considerata nell'Esempio 4.1.10 è un isomorfismo. La sua inversa è ancora la trasposizione, visto che ${}^t({}^tA) = A$.

Osservazione 4.2.27. Per una matrice quadrata $A \in M(n, \mathbb{K})$, tutti i fatti seguenti sono equivalenti:

$$\begin{aligned} L_A \text{ iniettiva} &\iff L_A \text{ suriettiva} \iff L_A \text{ isomorfismo} \iff A \text{ invertibile} \\ &\iff \det A \neq 0 \iff \text{rk}A = n \iff \forall b \in \mathbb{K}^n \exists! x \in \mathbb{K}^n : Ax = b \end{aligned}$$

Tornando alle composizioni, dimostriamo una disuguaglianza che è a volte un'uguaglianza in presenza di isomorfismi.

Proposizione 4.2.28. *Siano $f: V \rightarrow W$ e $g: W \rightarrow Z$ lineari. Allora:*

- $\dim \text{Im}(g \circ f) \leq \min\{\dim \text{Im} g, \dim \text{Im} f\}$.
- *Se f è un isomorfismo, allora $\dim \text{Im}(g \circ f) = \dim \text{Im} g$.*
- *Se g è un isomorfismo, allora $\dim \text{Im}(g \circ f) = \dim \text{Im} f$.*

Dimostrazione. Vale $\text{Im}(g \circ f) \subset \text{Im} g$, quindi $\dim(g \circ f) \leq \dim \text{Im} g$. Inoltre $\text{Im}(g \circ f) = \text{Im}(g|_{\text{Im} f})$ (si veda l'Esempio 4.1.16 per la definizione di restrizione) e quindi $\dim \text{Im}(g \circ f) \leq \dim \text{Im} f$ per il Corollario 4.2.14.

Se f è un isomorfismo, allora $(g \circ f) \circ f^{-1} = g$ e applicando la disuguaglianza otteniamo anche $\dim \text{Im} g \leq \dim \text{Im}(g \circ f)$. L'altro caso è analogo. \square

4.2.6. Rango del prodotto. Usiamo le proprietà delle applicazioni lineari per mostrare agevolmente alcune proprietà del rango di una matrice.

Proposizione 4.2.29. *Siano A e B due matrici che si possono moltiplicare. Valgono i fatti seguenti:*

- $\text{rk}(AB) \leq \min\{\text{rk}A, \text{rk}B\}$.
- *Se A è invertibile, allora $\text{rk}(AB) = \text{rk}B$.*
- *Se B è invertibile, allora $\text{rk}(AB) = \text{rk}A$.*

Dimostrazione. Segue dalla Proposizione 4.2.28 applicata a L_A e L_B . \square

4.2.7. Spazi vettoriali isomorfi. Diciamo che due spazi vettoriali V e W sullo stesso campo \mathbb{K} sono *isomorfi* se esiste un isomorfismo $f: V \rightarrow W$. Per capire se due spazi vettoriali sono isomorfi è sufficiente calcolare le loro dimensioni.

Proposizione 4.2.30. *Due spazi V e W sono isomorfi $\iff \dim V = \dim W$.*

Dimostrazione. (\Rightarrow) È conseguenza del Corollario 4.2.23.

(\Leftarrow) Per ipotesi V e W hanno due basi v_1, \dots, v_n e w_1, \dots, w_m entrambe di n elementi. Definiamo un'applicazione lineare $f: V \rightarrow W$ imponendo $f(v_i) = w_i$, si veda la Proposizione 4.1.18. L'applicazione lineare f è un isomorfismo per il Corollario 4.2.15. \square

Il corollario seguente è anche una conseguenza della Proposizione 4.2.22.

Corollario 4.2.31. *Tutti gli spazi vettoriali su \mathbb{K} di dimensione n sono isomorfi a \mathbb{K}^n .*

Ad esempio $\mathbb{K}_n[k]$ è isomorfo a \mathbb{K}^{n+1} e $M(m, n, \mathbb{K})$ è isomorfo a \mathbb{K}^{mn} .

4.3. Matrice associata

Introduciamo in questa sezione delle tecniche che ci permettono di rappresentare vettori e applicazioni lineari fra spazi vettoriali arbitrari come vettori numerici e matrici. In questa rappresentazione giocano un ruolo fondamentale le basi.

4.3.1. Definizione. Se fissiamo una base di uno spazio V , possiamo rappresentare qualsiasi vettore come un vettore numerico, quello delle sue coordinate. Adesso facciamo lo stesso con le applicazioni lineari: fissando opportune basi, possiamo rappresentare qualsiasi applicazione lineare come una matrice.

Sia $f: V \rightarrow W$ un'applicazione lineare fra spazi vettoriali definiti su \mathbb{K} . Siano inoltre

$$\mathcal{B} = \{v_1, \dots, v_n\}, \quad \mathcal{C} = \{w_1, \dots, w_m\}$$

due basi rispettivamente di V e di W . Sappiamo che

$$f(v_1) = a_{11}w_1 + \dots + a_{m1}w_m$$

$$\vdots$$

$$f(v_n) = a_{1n}w_1 + \dots + a_{mn}w_m$$

per qualche insieme di coefficienti $a_{ij} \in \mathbb{K}$. Definiamo la *matrice associata* a f nelle basi \mathcal{B} e \mathcal{C} come la matrice $m \times n$

$$A = (a_{ij})$$

che raggruppa questi coefficienti, e la indichiamo con il simbolo

$$A = [f]_{\mathcal{C}}^{\mathcal{B}}$$

che ci ricorda che A dipende da f , da \mathcal{B} e da \mathcal{C} . Notiamo che la base in partenza \mathcal{B} sta in alto e quella in arrivo \mathcal{C} sta in basso. Notiamo anche che la i -esima colonna di A contiene le coordinate di $f(v_i)$ rispetto a \mathcal{C} , cioè

$$A^i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} = [f(v_i)]_{\mathcal{C}}.$$

Per prima cosa, valutiamo come sempre L_A .

Proposizione 4.3.1. *La matrice associata a L_A rispetto alle basi canoniche di \mathbb{K}^n e \mathbb{K}^m è proprio A .*

Dimostrazione. Per costruzione $f(e_i) = a_{1i}e_1 + \cdots + a_{ni}e_n$. □

4.3.2. Esempi. Ci familiarizziamo con la nozione di matrice associata con qualche esempio.

Esempio 4.3.2. Consideriamo l'applicazione lineare

$$f: \mathbb{R}_2[x] \longrightarrow \mathbb{R}^2, \quad f(p) = \begin{pmatrix} p(2) \\ p(-2) \end{pmatrix}$$

che associa ad ogni polinomio i suoi valori in 2 e in -2 . Scriviamo la matrice associata a f nelle basi canoniche $\mathcal{C}' = \{1, x, x^2\}$ di $\mathbb{R}_2[x]$ e $\mathcal{C} = \{e_1, e_2\}$ di \mathbb{R}^2 . Calcoliamo:

$$f(1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad f(x) = \begin{pmatrix} 2 \\ -2 \end{pmatrix}, \quad f(x^2) = \begin{pmatrix} 4 \\ 4 \end{pmatrix}.$$

Quindi la matrice associata è

$$[f]_{\mathcal{C}}^{\mathcal{C}'} = \begin{pmatrix} 1 & 2 & 4 \\ 1 & -2 & 4 \end{pmatrix}.$$

Se in arrivo prendiamo la base $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ invece della canonica \mathcal{C} , allora dobbiamo anche calcolare le coordinate di ciascun vettore immagine rispetto a \mathcal{B} . Risolvendo dei semplici sistemi lineari troviamo:

$$\begin{aligned} f(1) &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\ f(x) &= \begin{pmatrix} 2 \\ -2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\ f(x^2) &= \begin{pmatrix} 4 \\ 4 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ -1 \end{pmatrix} + 8 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}$$

e la matrice associata diventa quindi

$$[f]_{\mathcal{B}}^{\mathcal{C}'} = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 0 & 8 \end{pmatrix}.$$

Esempio 4.3.3. Consideriamo l'applicazione lineare

$$f: \mathbb{C}^2 \longrightarrow \mathbb{C}^3, \quad f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - y \\ 2x \\ y \end{pmatrix}.$$

La matrice associata a f rispetto alle basi canoniche è semplicemente quella dei coefficienti:

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

Se invece come basi prendiamo in partenza

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

e in arrivo

$$w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad w_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

allora per calcolare la matrice associata dobbiamo fare dei conti. Dobbiamo determinare le coordinate di $f(v_1)$ e $f(v_2)$ rispetto alla base w_1, w_2, w_3 . Risolvendo dei sistemi lineari troviamo

$$f(v_1) = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \frac{3}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

$$f(v_2) = \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix} = \frac{5}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Quindi la matrice associata a f rispetto alle basi v_1, v_2 e w_1, w_2, w_3 è

$$[f]_{w_1, w_2, w_3}^{v_1, v_2} = \begin{pmatrix} \frac{1}{2} & \frac{5}{2} \\ \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 5 \\ 3 & -1 \\ -1 & -1 \end{pmatrix}.$$

4.3.3. Proprietà. Abbiamo scoperto come caratterizzare qualsiasi applicazione lineare (fra spazi di dimensione finita) come una matrice. Adesso mostriamo le proprietà di questa caratterizzazione, che sono notevoli.

Intanto notiamo che dalla matrice associata possiamo facilmente calcolare l'immagine di qualsiasi vettore. Sia $f: V \rightarrow W$ un'applicazione lineare e siano $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{C} = \{w_1, \dots, w_m\}$ basi di V e W .

Proposizione 4.3.4. *Per ogni $v \in V$ troviamo*

$$[f(v)]_{\mathcal{C}} = [f]_{\mathcal{C}}^{\mathcal{B}} [v]_{\mathcal{B}}.$$

Dimostrazione. Dato $v = \lambda_1 v_1 + \dots + \lambda_n v_n$, troviamo

$$f(v) = \lambda_1 f(v_1) + \dots + \lambda_n f(v_n)$$

e quindi

$$[f(v)]_{\mathcal{C}} = \lambda_1 [f(v_1)]_{\mathcal{C}} + \dots + \lambda_n [f(v_n)]_{\mathcal{C}} = [f]_{\mathcal{C}}^{\mathcal{B}} [v]_{\mathcal{B}}.$$

La dimostrazione è completa. \square

Per trovare le coordinate di $f(v)$ rispetto a \mathcal{C} è sufficiente moltiplicare la matrice associata a f per le coordinate di v rispetto a \mathcal{B} .

Osservazione 4.3.5. Se scriviamo $x = [v]_{\mathcal{B}}$, $A = [f]_{\mathcal{C}}^{\mathcal{B}}$ e $y = [f(v)]_{\mathcal{C}}$, allora

$$y = Ax = L_A(x).$$

Abbiamo quindi scoperto un fatto importante: dopo aver scelto due basi per V e W , qualsiasi applicazione lineare $V \rightarrow W$ può essere interpretata in coordinate come un'applicazione del tipo $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$. È sufficiente sostituire i vettori v e $f(v)$ con le loro coordinate x e y , ed usare la matrice associata A .

Esempio 4.3.6. Nell'Esempio 4.3.2, abbiamo ottenuto la matrice associata

$$[f]_{\mathcal{C}'}^{\mathcal{C}} = \begin{pmatrix} 1 & 2 & 4 \\ 1 & -2 & 4 \end{pmatrix}$$

rispetto alle basi canoniche. Possiamo usare questa matrice per calcolare in coordinate l'immagine di qualsiasi polinomio. Il polinomio generico $p(x) = ax^2 + bx + c$ ha come coordinate rispetto a \mathcal{C}' i suoi coefficienti (in ordine inverso). Quindi la sua immagine $f(p)$ ha coordinate

$$\begin{pmatrix} 1 & 2 & 4 \\ 1 & -2 & 4 \end{pmatrix} \begin{pmatrix} c \\ b \\ a \end{pmatrix} = \begin{pmatrix} 4a + 2b + c \\ 4a - 2b + c \end{pmatrix}$$

rispetto alla base canonica \mathcal{C} di \mathbb{R}^2 . Come verifica, possiamo calcolare la f direttamente dalla definizione e ottenere effettivamente:

$$f(ax^2 + bx + c) = \begin{pmatrix} 4a + 2b + c \\ 4a - 2b + c \end{pmatrix}.$$

Troviamo un fatto analogo con la composizione di funzioni. Siano $f: U \rightarrow V$ e $g: V \rightarrow W$ due applicazioni lineari. Siano \mathcal{B}, \mathcal{C} e \mathcal{D} basi di U, V e W .

Proposizione 4.3.7. *Troviamo*

$$[g \circ f]_{\mathcal{D}}^{\mathcal{B}} = [g]_{\mathcal{D}}^{\mathcal{C}} [f]_{\mathcal{C}}^{\mathcal{B}}.$$

Dimostrazione. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$. La colonna i -esima di $[g \circ f]_{\mathcal{D}}^{\mathcal{B}}$ è

$$[g(f(v_i))]_{\mathcal{D}} = [g]_{\mathcal{D}}^{\mathcal{C}} [f(v_i)]_{\mathcal{C}}$$

per la proposizione precedente. D'altra parte $[f(v_i)]_{\mathcal{C}}$ è la colonna i -esima di $[f]_{\mathcal{C}}^{\mathcal{B}}$ e quindi concludiamo. \square

Abbiamo scoperto che nel passaggio dalle funzioni lineari alle matrici la composizione di funzioni diventa moltiplicazione fra matrici. Ora mostriamo che l'applicazione identità (che indichiamo sempre con id_V) si trasforma sempre nella matrice identità, purché si scelga la stessa base in partenza ed in arrivo.

Proposizione 4.3.8. *Sia \mathcal{B} base di uno spazio V di dimensione n . Troviamo*

$$[\text{id}_V]_{\mathcal{B}}^{\mathcal{B}} = I_n.$$

Dimostrazione. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$. Poiché $f(v_i) = v_i$, la matrice associata è I_n per costruzione. \square

Adesso mostriamo che qualsiasi matrice è matrice associata di un opportuna applicazione lineare. Siano \mathcal{B}, \mathcal{C} basi per V e W , con $n = \dim V$ e $m = \dim W$.

Proposizione 4.3.9. *Qualsiasi matrice $A \in M(m, n)$ è la matrice associata $A = [f]_{\mathcal{C}}^{\mathcal{B}}$ ad un'unica applicazione lineare $f: V \rightarrow W$.*

Dimostrazione. Siano $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{C} = \{w_1, \dots, w_m\}$. È sufficiente definire f imponendo che

$$f(v_i) = a_{1i}w_1 + \dots + a_{mi}w_m.$$

La f esiste ed è unica per la Proposizione 4.1.18. \square

Abbiamo quindi ottenuto una corrispondenza biunivoca fra applicazioni lineari e matrici (sempre dopo aver fissato delle basi). Mostriamo adesso che questa corrispondenza associa isomorfismi e matrici invertibili. Sia $f: V \rightarrow W$ una applicazione lineare, e siano \mathcal{B} e \mathcal{C} basi per V e W .

Corollario 4.3.10. *La funzione f è un isomorfismo \iff la matrice associata $[f]_{\mathcal{C}}^{\mathcal{B}}$ è invertibile, ed in questo caso la sua inversa è $[f^{-1}]_{\mathcal{B}}^{\mathcal{C}}$.*

Dimostrazione. Se f è un isomorfismo, ha una inversa $f^{-1}: W \rightarrow V$. Dalle proposizioni precedenti deduciamo che

$$I_n = [\text{id}_V]_{\mathcal{B}}^{\mathcal{B}} = [f^{-1} \circ f]_{\mathcal{B}}^{\mathcal{B}} = [f^{-1}]_{\mathcal{B}}^{\mathcal{C}} [f]_{\mathcal{C}}^{\mathcal{B}}$$

e analogamente $I_m = [\text{id}_W]_{\mathcal{C}}^{\mathcal{C}} = [f]_{\mathcal{C}}^{\mathcal{B}} [f^{-1}]_{\mathcal{B}}^{\mathcal{C}}$. Quindi effettivamente $[f^{-1}]_{\mathcal{B}}^{\mathcal{C}}$ è la matrice inversa di $[f]_{\mathcal{C}}^{\mathcal{B}}$.

D'altro canto, se $A = [f]_{\mathcal{C}}^{\mathcal{B}}$ è invertibile, grazie alla proposizione precedente esiste una $g: W \rightarrow V$ tale che $[g]_{\mathcal{B}}^{\mathcal{C}} = A^{-1}$ e deduciamo analogamente che g è l'inversa di f , quindi f è un isomorfismo. \square

4.3.4. Hom. Introduciamo adesso un altro legame forte fra applicazioni lineari e matrici. Siano V e W spazi vettoriali su \mathbb{K} di dimensione n e m . Sia

$$\text{Hom}(V, W)$$

l'insieme di tutte le applicazioni lineari $V \rightarrow W$. Il nome "hom" deriva dal fatto che in matematica il termine *omomorfismo* è usato come sinonimo di applicazione lineare.

Non dovrebbe sorprendere il fatto che $\text{Hom}(V, W)$ abbia una naturale struttura di spazio vettoriale. Date $f, g \in \text{Hom}(V, W)$ e $\lambda \in \mathbb{K}$ definiamo $f + g: V \rightarrow W$ e $\lambda f: V \rightarrow W$ nel modo usuale:

$$\begin{aligned} (f + g)(v) &= f(v) + g(v), \\ (\lambda f)(v) &= \lambda f(v). \end{aligned}$$

Lo zero in $\text{Hom}(V, W)$ è l'applicazione nulla. Siano \mathcal{B} e \mathcal{C} basi di V e W .

Proposizione 4.3.11. *La mappa*

$$T: \text{Hom}(V, W) \longrightarrow M(m, n, \mathbb{K}), \quad T(f) = [f]_{\mathcal{C}}^{\mathcal{B}}$$

è un isomorfismo di spazi vettoriali.

Dimostrazione. La linearità di T segue facilmente dalla definizione di matrice associata. Si tratta di un isomorfismo per la Proposizione 4.3.9. \square

Corollario 4.3.12. *Lo spazio $\text{Hom}(V, W)$ ha dimensione mn .*

Se $V = W$, allora $\text{Hom}(V, V)$ ha anche un'altra operazione, quella di composizione fra funzioni indicata con \circ . Si verifica facilmente che $\text{Hom}(V, V)$ è un anello con le operazioni $+$ e \circ , in cui l'elemento neutro per la somma è la funzione nulla e l'identità per la composizione è la funzione identità id_V .

La mappa $T: \text{Hom}(V, V) \rightarrow M(n)$ è quindi una funzione fra anelli. Se $\mathcal{B} = \mathcal{C}$, cioè se scegliamo la stessa base in partenza ed in arrivo, per quanto visto nella sezione precedente la mappa T è compatibile anche con le operazioni di prodotto, nel senso che:

$$T(f \circ g) = T(f)T(g), \quad T(\text{id}) = I_n, \quad T(f^{-1}) = T(f)^{-1}.$$

Riassumendo, l'operatore T che associa ad ogni funzione lineare la sua matrice associata è compatibile con gran parte delle operazioni algebriche viste finora.

4.3.5. Matrice di cambiamento di base. Abbiamo visto che la matrice associata ad un'applicazione lineare dipende in maniera essenziale dalle basi scelte. Vogliamo capire meglio come controllare questo fenomeno, e a tale scopo introduciamo uno strumento noto come *matrice di cambiamento di base*.

Sia V uno spazio vettoriale e $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{C} = \{w_1, \dots, w_n\}$ due basi di V . La *matrice di cambiamento di base* da \mathcal{B} a \mathcal{C} è la matrice

$$A = [\text{id}_V]_{\mathcal{C}}^{\mathcal{B}}.$$

La matrice A rappresenta la funzione identità nelle basi \mathcal{B} e \mathcal{C} . Quindi

$$v_i = a_{1i}w_1 + \dots + a_{ni}w_n.$$

La colonna i -esima di A contiene le coordinate di v_i rispetto a \mathcal{C} , cioè

$$A^i = [v_i]_{\mathcal{C}}.$$

La matrice di cambiamento di base A codifica quindi nelle sue colonne le coordinate di ciascun elemento di \mathcal{B} rispetto a \mathcal{C} . Chiaramente l'inversa

$$A^{-1} = [\text{id}_V]_{\mathcal{B}}^{\mathcal{C}}$$

è la matrice di cambiamento di base da \mathcal{C} a \mathcal{B} , e ha nelle sue colonne le coordinate di ciascun elemento di \mathcal{C} rispetto a \mathcal{B} .

La matrice di cambiamento di base è lo strumento che funge da dizionario tra una base e l'altra. Vediamo come. Sia $f: V \rightarrow W$ una applicazione lineare.

Siano $\mathcal{B}_1, \mathcal{B}_2$ due basi di V e $\mathcal{C}_1, \mathcal{C}_2$ due basi di W . Applicando la Proposizione 4.3.7 troviamo:

$$[f]_{\mathcal{C}_2}^{\mathcal{B}_2} = [\text{id}_W]_{\mathcal{C}_2}^{\mathcal{C}_1} \cdot [f]_{\mathcal{C}_1}^{\mathcal{B}_1} \cdot [\text{id}_V]_{\mathcal{B}_1}^{\mathcal{B}_2}.$$

Per passare da $[f]_{\mathcal{C}_1}^{\mathcal{B}_1}$ a $[f]_{\mathcal{C}_2}^{\mathcal{B}_2}$ è sufficiente moltiplicare a sinistra e a destra per le matrici di cambiamento di base.

Esempio 4.3.13. Nell'Esempio 4.3.2 abbiamo analizzato un'applicazione lineare $f: \mathbb{R}_2[x] \rightarrow \mathbb{R}^2$ e calcolato la matrice associata rispetto alle basi $\mathcal{C}' = \{1, x, x^2\}$ e $\mathcal{C} = \{e_1, e_2\}$. Il risultato era il seguente:

$$[f]_{\mathcal{C}}^{\mathcal{C}'} = \begin{pmatrix} 1 & 2 & 4 \\ 1 & -2 & 4 \end{pmatrix}.$$

Successivamente abbiamo cambiato \mathcal{C} con $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$. La matrice di cambiamento di base da \mathcal{B} a \mathcal{C} è molto semplice, perché le coordinate di un vettore rispetto a \mathcal{C} sono il vettore stesso e quindi basta affiancare i vettori di \mathcal{B} ; otteniamo

$$[\text{id}]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Adesso calcoliamo l'inversa e troviamo

$$[\text{id}]_{\mathcal{B}}^{\mathcal{C}} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Quindi la matrice associata a f nelle basi \mathcal{C}' e \mathcal{B} è

$$[f]_{\mathcal{B}}^{\mathcal{C}'} = [\text{id}]_{\mathcal{B}}^{\mathcal{C}} [f]_{\mathcal{C}}^{\mathcal{C}'} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 4 \\ 1 & -2 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 0 & 8 \end{pmatrix}$$

che coincide effettivamente con quanto trovato nell'Esempio 4.3.2.

Esempio 4.3.14. Nell'Esempio 4.3.3 abbiamo considerato un'applicazione lineare $f: \mathbb{C}^2 \rightarrow \mathbb{C}^3$ e calcolato la matrice associata rispetto alle basi canoniche \mathcal{C} e \mathcal{C}' di \mathbb{R}^2 e \mathbb{R}^3 :

$$[f]_{\mathcal{C}'}^{\mathcal{C}} = \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Successivamente abbiamo cambiato entrambe le basi con

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}, \quad \mathcal{B}' = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Calcoliamo le matrici di cambiamento di base:

$$[\text{id}_{\mathbb{C}^2}]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad [\text{id}_{\mathbb{C}^3}]_{\mathcal{C}'}^{\mathcal{B}'} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

L'inversa della seconda è

$$[\text{id}_{\mathbb{C}^3}]_{\mathcal{B}'}^{\mathcal{C}'} = \frac{1}{2} \begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}.$$

Troviamo infine

$$\begin{aligned} [f]_{\mathcal{B}'}^{\mathcal{B}} &= [\text{id}_{\mathbb{C}^3}]_{\mathcal{B}'}^{\mathcal{C}'} \cdot [f]_{\mathcal{C}'}^{\mathcal{C}} \cdot [\text{id}_{\mathbb{C}^2}]_{\mathcal{C}}^{\mathcal{B}} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 5 \\ 3 & -1 \\ -1 & -1 \end{pmatrix} \end{aligned}$$

coerentemente con quanto già visto nell'Esempio 4.3.3.

4.4. Endomorfismi

Trattiamo adesso più approfonditamente le applicazioni lineari $V \rightarrow V$ in cui dominio e codominio coincidono. Queste applicazioni vengono chiamate *endomorfismi*.

4.4.1. Definizione. Sia V uno spazio vettoriale.

Definizione 4.4.1. Un *endomorfismo* è una applicazione lineare $f: V \rightarrow V$.

Abbiamo già visto molti esempi di endomorfismi nelle pagine precedenti. Ne ricordiamo alcuni molto brevemente. In ciascuno spazio vettoriale V , la moltiplicazione per un $\lambda \in \mathbb{K}$ fissato $f(v) = \lambda v$ è un endomorfismo. Gli endomorfismi di \mathbb{K}^n sono tutti del tipo L_A dove A è una matrice quadrata $n \times n$. La derivata è un endomorfismo di $\mathbb{R}_n[x]$ e la trasposta è un endomorfismo di $M(n, \mathbb{K})$. Moltiplicando a sinistra per una matrice fissata $A \in M(n, \mathbb{K})$ si ottiene un altro endomorfismo di $M(n, \mathbb{K})$. Se $V = U \oplus W$, allora possiamo definire due endomorfismi p_U e r_U di V , la proiezione su U e la riflessione lungo U . Altri endomorfismi più geometrici di \mathbb{R}^2 e \mathbb{R}^3 verranno descritti fra poco.

4.4.2. Matrice associata. Abbiamo visto nelle pagine precedenti che se fissiamo una base \mathcal{B} per V è possibile rappresentare ogni endomorfismo f di V come una matrice $A = [f]_{\mathcal{B}}^{\mathcal{B}}$. Questa rappresentazione ha la buona proprietà di trasformare la composizione fra endomorfismi nel prodotto fra matrici:

$$[f \circ g]_{\mathcal{B}}^{\mathcal{B}} = [f]_{\mathcal{B}}^{\mathcal{B}} \cdot [g]_{\mathcal{B}}^{\mathcal{B}}.$$

Abbiamo anche visto che se \mathcal{C} è un'altra base per V e $M = [\text{id}_V]_{\mathcal{C}}^{\mathcal{B}}$ è la matrice di cambiamento di base da \mathcal{B} a \mathcal{C} , otteniamo

$$[f]_{\mathcal{B}}^{\mathcal{B}} = M^{-1} [f]_{\mathcal{C}}^{\mathcal{C}} M.$$

Ricapitolando, se A e B sono le matrici associate a f rispetto alle basi \mathcal{B} e \mathcal{C} e M è la matrice di cambiamento di base, allora

$$A = M^{-1} B M.$$

Esempio 4.4.2. Consideriamo la riflessione $r_U: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dell'Esempio 4.1.12. Avevamo espresso la funzione esplicitamente nel modo seguente:

$$r_U \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ -y \end{pmatrix}.$$

Rispetto alla base canonica $\mathcal{C} = \{e_1, e_2\}$, troviamo

$$[r_U]_{\mathcal{C}}^{\mathcal{C}} = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}.$$

Prendiamo adesso come base $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right\}$, due vettori che generano i sottospazi U e W . La matrice di cambiamento di base da \mathcal{B} a \mathcal{C} è semplicemente

$$M = [\text{id}]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$$

e la sua inversa è

$$M^{-1} = [\text{id}]_{\mathcal{B}}^{\mathcal{C}} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}.$$

Quindi la matrice associata a r_U nella base \mathcal{B} è

$$[r_U]_{\mathcal{B}}^{\mathcal{B}} = M^{-1}[r_U]_{\mathcal{C}}^{\mathcal{C}}M = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Possiamo verificare che il risultato sia giusto. Le colonne di $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ sono le coordinate delle immagini dei vettori della base \mathcal{B} rispetto alla stessa base \mathcal{B} . La prima colonna sta ad indicare che r_U manda il primo vettore di \mathcal{B} in se stesso e la seconda indica che r_U manda il secondo vettore di \mathcal{B} nel suo opposto. Quindi $r_U\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ e $r_U\left(\begin{pmatrix} -1 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$. Questo è effettivamente quello che fa la riflessione, che fissa i punti di U e ribalta quelli di W .

Esercizio 4.4.3. Calcola $[\rho_U]_{\mathcal{B}}^{\mathcal{B}}$ dove $\rho_U: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ è la proiezione dell'Esempio 4.1.11 e \mathcal{B} è la stessa base usata nell'esempio precedente.

Esempio 4.4.4. Sia V uno spazio vettoriale di dimensione n e sia $\lambda \in \mathbb{K}$. La matrice associata all'endomorfismo $f(v) = \lambda v$ è

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \lambda I_n$$

rispetto a *qualsiasi* base $\mathcal{B} = \{v_1, \dots, v_n\}$, perché $f(v_i) = \lambda v_i$ per ogni i e quindi $[f(v_i)]_{\mathcal{B}} = \lambda e_i$. Questo è coerente con il fatto che

$$M^{-1}(\lambda I_n)M = \lambda M^{-1}I_nM = \lambda M^{-1}M = \lambda I_n$$

per qualsiasi matrice invertibile M . Cambiando base la matrice associata rimane sempre λI_n .

La relazione $A = M^{-1}BM$ scoperta in questa sezione ci porta a studiare più in generale questa operazione fra matrici quadrate.

4.4.3. Similitudine fra matrici. Lavoriamo con l'anello $M(n)$ delle matrici quadrate $n \times n$. Diciamo che due matrici $A, B \in M(n)$ sono *simili* o *coniugate* se esiste una matrice invertibile $M \in M(n)$ tale che

$$A = M^{-1}BM.$$

Se A e B sono simili scriviamo $A \sim B$. Rimandiamo alla Sezione 1.1.10 per la nozione di relazione di equivalenza.

Proposizione 4.4.5. *La similitudine è una relazione di equivalenza in $M(n)$.*

Dimostrazione. Dobbiamo dimostrare tre proprietà:

- (1) $A \sim A$. Infatti prendendo $M = I_n$ otteniamo $A = I_n^{-1}AI_n$.
- (2) $A \sim B \Rightarrow B \sim A$. Infatti se $A = M^{-1}BM$, allora moltiplicando entrambi i membri a sinistra per M e a destra per M^{-1} troviamo $B = MAM^{-1}$, e definendo $N = M^{-1}$ otteniamo anche $B = N^{-1}AN$.
- (3) $A \sim B, B \sim C, \Rightarrow A \sim C$. Infatti

$$A = M^{-1}BM, B = N^{-1}CN \Rightarrow A = M^{-1}N^{-1}CNM = (NM)^{-1}C(NM).$$

La dimostrazione è completa. \square

L'insieme $M(n)$ delle matrici quadrate è quindi partizionato in sottoinsiemi disgiunti formati da matrici simili fra loro.

Esempio 4.4.6. Per ogni $\lambda \in \mathbb{K}$, la matrice λI_n è simile solo a se stessa: come abbiamo già notato, $M^{-1}(\lambda I_n)M = \lambda I_n$ per qualsiasi matrice invertibile M . In particolare le matrici nulla 0 e identità I_n sono simili solo a loro stesse.

Esempio 4.4.7. Le matrici $\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ per $k \neq 0$ sono tutte simili fra loro. Per dimostrare questo fatto usiamo la proprietà transitiva e ci limitiamo a mostrare che ciascuna $\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ con $k \neq 0$ è simile a $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Infatti se $M = \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}$ otteniamo

$$M^{-1} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} M = \begin{pmatrix} 1 & 0 \\ 0 & k^{-1} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix} = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}.$$

Le matrici $\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ con $k \neq 0$ non sono però simili alla matrice $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, perché sappiamo che la matrice identità è simile solo a se stessa.

Dimostreremo nelle prossime pagine che due matrici simili hanno molto in comune. Iniziamo con il rango e il determinante.

Proposizione 4.4.8. *Se $A \sim B$ allora $\text{rk}(A) = \text{rk}(B)$ e $\det(A) = \det(B)$. In particolare A è invertibile $\iff B$ è invertibile.*

Dimostrazione. Segue dalla Proposizione 4.2.29 e dal Teorema di Binet:

$$\det(M^{-1}BM) = \det(M^{-1}) \det B \det M = \frac{1}{\det M} \det B \det M = \det B.$$

La dimostrazione è completa. \square

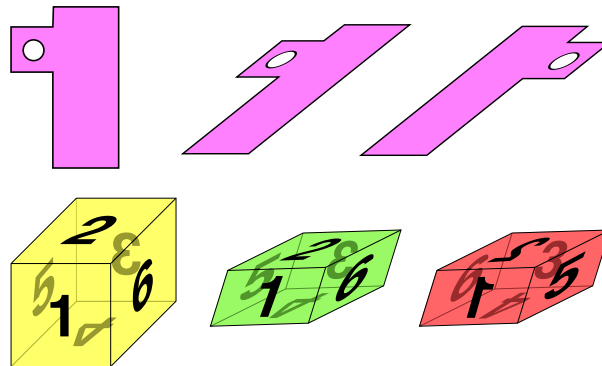


Figura 4.3. L'effetto su di un oggetto in \mathbb{R}^2 e in \mathbb{R}^3 di una trasformazione con determinante positivo e negativo.

Quindi ad esempio le matrici $\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$ e $\begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix}$ non sono simili perché hanno determinanti diversi. Il motivo per cui siamo interessati alla relazione di similitudine è il seguente, dimostrato alla fine della sezione precedente.

Proposizione 4.4.9. *Due matrici $[f]_{\mathcal{B}}^{\mathcal{B}}$ e $[f]_{\mathcal{C}}^{\mathcal{C}}$ associate allo stesso endomorfismo $f: V \rightarrow V$ ma con basi \mathcal{B}, \mathcal{C} diverse sono simili.*

Come corollario abbiamo questo fatto cruciale:

Qualsiasi oggetto associato ad una matrice quadrata A che non cambi per similitudine è ben definito sugli endomorfismi.

Chiariamo questo passaggio con un esempio importante.

4.4.4. Determinante di un endomorfismo. Definiamo il *determinante* di un endomorfismo $f: V \rightarrow V$ come il determinante della matrice associata

$$\det f = \det[f]_{\mathcal{B}}^{\mathcal{B}}$$

rispetto a *qualsiasi* base \mathcal{B} . Per quanto abbiamo appena visto infatti il determinante è invariante per similitudine: quindi il determinante della matrice $[f]_{\mathcal{B}}^{\mathcal{B}}$ non dipende da \mathcal{B} e quindi è ben definito.

Esempio 4.4.10. L'endomorfismo $f(v) = \lambda v$ in uno spazio V di dimensione n ha determinante λ^n , perché $\det(\lambda I_n) = \lambda^n$.

Sappiamo che $f: V \rightarrow V$ è un isomorfismo $\iff \det f \neq 0$. Nel caso in cui $V = \mathbb{R}^2$ oppure \mathbb{R}^3 , il determinante di f ha una chiara interpretazione geometrica, a cui abbiamo già accennato nella Sezione 3.3.8 e che studieremo fra poco.

4.4.5. Traccia di una matrice quadrata. Introduciamo adesso un'altra quantità che non cambia per similitudine, la *traccia*.

La *traccia* di una matrice quadrata $A \in M(n)$ è il numero

$$\text{tr}A = A_{11} + \cdots + A_{nn}.$$

La traccia di A è la somma dei numeri presenti sulla diagonale principale di A .

Proposizione 4.4.11. Se $A, B \in M(n)$, vale la relazione $\text{tr}(AB) = \text{tr}(BA)$.

Dimostrazione. Troviamo

$$\begin{aligned}\text{tr}(AB) &= \sum_{i=1}^n (AB)_{ii} = \sum_{i=1}^n \sum_{j=1}^n A_{ij} B_{ji} \\ \text{tr}(BA) &= \sum_{j=1}^n (BA)_{jj} = \sum_{j=1}^n \sum_{i=1}^n B_{ji} A_{ij}\end{aligned}$$

Il risultato è lo stesso. \square

Corollario 4.4.12. Due matrici simili hanno la stessa traccia.

Dimostrazione. Se $A = M^{-1}BM$, troviamo

$$\text{tr}A = \text{tr}(M^{-1}BM) = \text{tr}((M^{-1}B)M) = \text{tr}(M(M^{-1}B)) = \text{tr}B.$$

La dimostrazione è completa. \square

Quindi due matrici con tracce diverse, come ad esempio $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ e $\begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$, non sono simili. Notiamo però che esistono matrici con la stessa traccia e lo stesso determinante che non sono simili: ad esempio $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ e $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Se $f: V \rightarrow V$ è un endomorfismo e V ha dimensione finita, possiamo definire la *traccia* di f come la traccia della matrice associata $[f]_{\mathcal{B}}^{\mathcal{B}}$ rispetto a qualche base \mathcal{B} . Per il corollario precedente la traccia di f effettivamente non dipende dalla base \mathcal{B} scelta.

Esempio 4.4.13. L'endomorfismo $f(v) = \lambda v$ in uno spazio V di dimensione n ha traccia $n\lambda$, perché $\text{tr}(\lambda I_n) = n\lambda$.

4.4.6. Trasformazioni lineari del piano e dello spazio. Un endomorfismo f invertibile di \mathbb{R}^2 o \mathbb{R}^3 è detto *trasformazione lineare del piano* o *dello spazio*. Le trasformazioni hanno un ruolo importante nella geometria moderna.

Il determinante di una trasformazione lineare f di \mathbb{R}^2 o \mathbb{R}^3 ha una chiara interpretazione geometrica. Se $\det f < 0$, la trasformazione f specchia gli oggetti contenuti in \mathbb{R}^2 o \mathbb{R}^3 . La Figura 4.3 mostra l'effetto su di un oggetto di una trasformazione con determinante positivo e negativo. Entrambe possono distorcere l'oggetto, ma fra le due trasformazioni c'è un'importante differenza qualitativa: la seconda cambia l'orientazione dell'oggetto mentre la prima no. Una trasformazione con $\det < 0$ trasforma ad esempio mani destre in mani sinistre e viceversa. Tutto ciò avviene sia in \mathbb{R}^2 che in \mathbb{R}^3 .

D'altro canto, come accennato nella Sezione 3.3.8, il valore assoluto $|\det f|$ del determinante indica come vengono trasformate le aree (in \mathbb{R}^2) o i volumi (in \mathbb{R}^3) degli oggetti. Ciascun oggetto $C \subset \mathbb{R}^2$ o \mathbb{R}^3 viene trasformato in un oggetto $f(C)$ la cui area o volume è esattamente $|\det f|$ volte quella originaria di C . In particolare se $\det f = \pm 1$ allora la trasformazione f

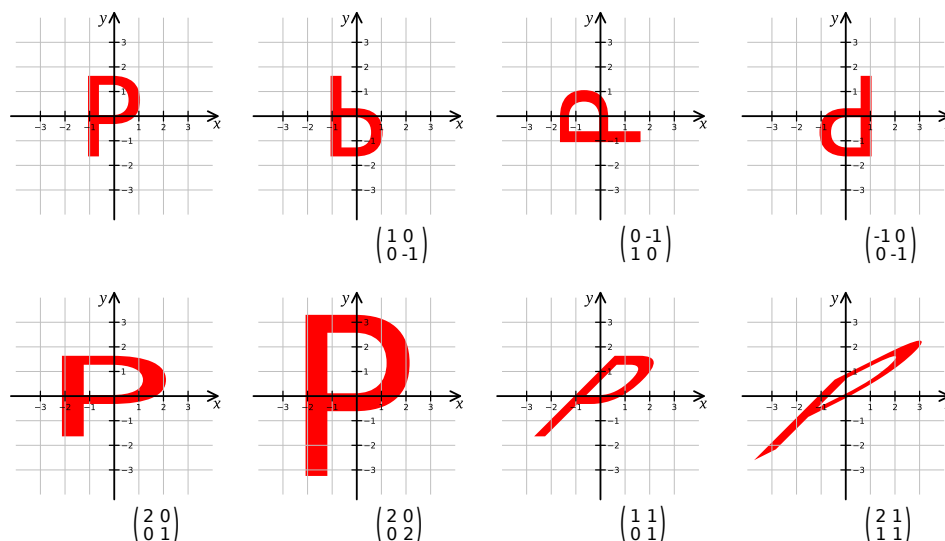


Figura 4.4. Sette trasformazioni del piano. Ciascuna trasformazione è un isomorfismo di \mathbb{R}^2 determinato dalla matrice 2×2 invertibile indicata in basso; viene mostrato il suo effetto sulla lettera P, originariamente nella posizione in alto a sinistra.

non cambia l'area o il volume degli oggetti, ma può distorcerli lo stesso come vedremo fra poco.

4.4.7. Trasformazioni del piano. Alcuni esempi di trasformazioni lineari del piano sono visualizzati in Figura 4.4. Ciascun isomorfismo è del tipo $L_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ e la matrice A corrispondente è indicata sotto la figura. La lettrice è invitata ad analizzare attentamente la matrice di ciascuna trasformazione e a convincersi geometricamente che l'effetto sulla lettera P sia quello descritto. Ricordiamo che le colonne della matrice A sono le immagini dei vettori e_1 e e_2 .

Notiamo che i primi tre esempi non distorcono la figura, e sono rispettivamente una riflessione lungo l'asse x , una rotazione antioraria di $\frac{\pi}{2}$, e una riflessione rispetto all'origine, che è la stessa cosa che una rotazione di angolo π . I quattro esempi successivi distorcono la figura: il primo la allarga orizzontalmente, il secondo è una omotetia di fattore 2, e le due trasformazioni successive sembrano distorcere la figura in modo maggiore.

I determinanti delle sette trasformazioni mostrate sono $-1, 1, 1, 2, 4, 1$ e 1 . Solo nella prima trasformazione (la riflessione) la lettera P viene specchiata. Inoltre l'area della P è preservata in tutte le trasformazioni eccetto la quarta e la quinta in cui viene moltiplicata per 2 e per 4 rispettivamente. Nelle ultime due trasformazioni c'è un forte effetto di distorsione ma l'area della P rimane la stessa.

4.4.8. Rotazioni nel piano. Introduciamo una classe importante di trasformazioni di \mathbb{R}^2 , le *rotazioni*.

Definizione 4.4.14. Una *rotazione* di angolo θ è la trasformazione $L_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ determinata dalla matrice $A = \text{Rot}_\theta$, con

$$\text{Rot}_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Proposizione 4.4.15. *La trasformazione è effettivamente una rotazione antioraria del piano di angolo θ intorno all'origine.*

Dimostrazione. Rappresentiamo un punto $\begin{pmatrix} x \\ y \end{pmatrix}$ in coordinate polari $x = \rho \cos \varphi$ e $y = \rho \sin \varphi$. Calcoliamo la sua immagine:

$$\begin{aligned} L_A \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \rho \cos \varphi \\ \rho \sin \varphi \end{pmatrix} = \begin{pmatrix} \rho(\cos \theta \cos \varphi - \sin \theta \sin \varphi) \\ \rho(\sin \theta \cos \varphi + \cos \theta \sin \varphi) \end{pmatrix} \\ &= \begin{pmatrix} \rho \cos(\theta + \varphi) \\ \rho \sin(\theta + \varphi) \end{pmatrix}. \end{aligned}$$

La trasformazione manda il punto generico con coordinate polari (ρ, φ) nel punto con coordinate polari $(\rho, \theta + \varphi)$. Effettivamente tutti i punti vengono ruotati in senso antiorario dell'angolo θ . \square

Notiamo che la matrice Rot_θ ha sempre determinante $\cos^2 \theta + \sin^2 \theta = 1$. Effettivamente le rotazioni non specchiano gli oggetti e mantengono le aree.

4.4.9. Riflessioni ortogonali nel piano. Introduciamo un'altra classe di trasformazioni del piano, le *riflessioni ortogonali*. Fissiamo un angolo θ e consideriamo la retta vettoriale r che forma un angolo $\frac{\theta}{2}$ con l'asse delle x .

Definizione 4.4.16. Una *riflessione ortogonale* rispetto alla retta r è la trasformazione $L_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ determinata dalla matrice $A = \text{Rif}_\theta$, con

$$\text{Rif}_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}.$$

Proposizione 4.4.17. *La trasformazione è effettivamente una riflessione ortogonale del piano rispetto a r .*

Dimostrazione. Rappresentiamo un punto $\begin{pmatrix} x \\ y \end{pmatrix}$ in coordinate polari $x = \rho \cos \varphi$ e $y = \rho \sin \varphi$. Calcoliamo la sua immagine:

$$\begin{aligned} L_A \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix} \begin{pmatrix} \rho \cos \varphi \\ \rho \sin \varphi \end{pmatrix} = \begin{pmatrix} \rho(\cos \theta \cos \varphi + \sin \theta \sin \varphi) \\ \rho(\sin \theta \cos \varphi - \cos \theta \sin \varphi) \end{pmatrix} \\ &= \begin{pmatrix} \rho \cos(\theta - \varphi) \\ \rho \sin(\theta - \varphi) \end{pmatrix}. \end{aligned}$$

La trasformazione manda il punto generico con coordinate polari (ρ, φ) nel punto con coordinate polari $(\rho, \theta - \varphi)$. Questa è una riflessione ortogonale rispetto a r , si veda la Figura 4.5. \square

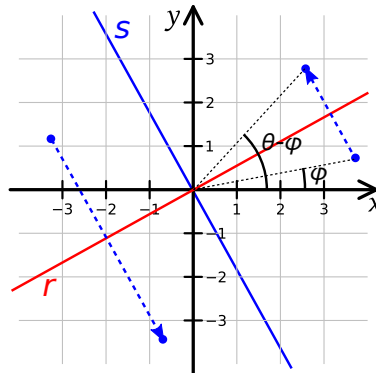


Figura 4.5. Una riflessione rispetto alla retta r che forma un angolo $\frac{\theta}{2}$ con l'asse x .

Notiamo che la matrice Rif_θ ha sempre determinante $-\cos^2 \theta - \sin^2 \theta = -1$. Le riflessioni specchiano gli oggetti ma mantengono le aree.

Osservazione 4.4.18. Sia s la retta ortogonale a r , cioè facente un angolo $\frac{\theta}{2} + \frac{\pi}{2}$ con l'asse x , come in Figura 4.5. La riflessione ortogonale descritta qui corrisponde alla riflessione determinata dalla decomposizione in somma diretta $\mathbb{R}^2 = r \oplus s$ definita nell'Esempio 4.1.12. Come già osservato nell'Esempio 4.4.2 la riflessione può essere rappresentata più agevolmente rispetto ad un'altra base $\mathcal{B} = \{v_1, v_2\}$ in cui $v_1 \in r$ e $v_2 \in s$. Poiché $f(v_1) = v_1$ e $f(v_2) = -v_2$, la matrice associata alla riflessione rispetto alla base \mathcal{B} è semplicemente

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

4.4.10. Omotetie. Una *omotetia* in \mathbb{R}^2 oppure \mathbb{R}^3 di centro l'origine è semplicemente la trasformazione

$$f(x) = \lambda x$$

dove $\lambda > 0$ è il *coefficiente* dell'omotetia. Abbiamo $f = L_A$ dove $A = \lambda I_n$, con $n = 2$ oppure 3 , cioè

$$A = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \quad \text{oppure} \quad A = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}.$$

L'omotetia ingrandisce o rimpicciolisce gli oggetti di un fattore λ . Aree e volumi vengono modificati di un fattore $\det(A) = \lambda^2$ oppure λ^3 .

4.4.11. Rotazioni intorno ad un asse nello spazio. Fra le trasformazioni dello spazio, spiccano le *rotazioni intorno ad un asse*. Studieremo queste trasformazioni nel dettaglio in seguito: per adesso ci limitiamo a definire la

rotazione di un angolo θ intorno all'asse z come la trasformazione $L_A: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ in cui la matrice A è la seguente:

$$A = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \text{Rot}_\theta & 0 \\ 0 & 1 \end{pmatrix}.$$

La seconda scrittura più compatta descrive la A come matrice a blocchi: Rot_θ è una sottomatrice 2×2 e i due "0" sono due sottomatrici 2×1 e 1×2 nulle.

Proposizione 4.4.19. *La trasformazione è effettivamente una rotazione antioraria dello spazio di angolo θ intorno all'asse z .*

Dimostrazione. Rappresentiamo un punto di \mathbb{R}^3 in *coordinate cilindriche*: queste sono semplicemente le coordinate polari per x e y con la z lasciata immutata. Passiamo cioè da (x, y, z) a (ρ, θ, z) con $x = \rho \cos \varphi$ e $y = \rho \sin \varphi$. Calcoliamo l'immagine del punto generico con coordinate cilindriche (ρ, φ, z) :

$$\begin{aligned} L_A \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \rho \cos \varphi \\ \rho \sin \varphi \\ z \end{pmatrix} = \begin{pmatrix} \rho(\cos \theta \cos \varphi - \sin \theta \sin \varphi) \\ \rho(\sin \theta \cos \varphi + \cos \theta \sin \varphi) \\ z \end{pmatrix} \\ &= \begin{pmatrix} \rho \cos(\theta + \varphi) \\ \rho \sin(\theta + \varphi) \\ z \end{pmatrix}. \end{aligned}$$

La trasformazione manda il punto generico con coordinate cilindriche (ρ, φ, z) nel punto con coordinate cilindriche $(\rho, \theta + \varphi, z)$. Effettivamente tutti i punti vengono ruotati in senso antiorario dell'angolo θ intorno all'asse z . \square

Vedremo nei prossimi capitoli come scrivere una rotazione di angolo θ intorno ad un asse arbitrario.

Esercizi

Esercizio 4.1. Sia $T: \mathbb{R}_2[x] \rightarrow \mathbb{R}_3[x]$ la funzione $T(p) = xp(x)$. Mostra che T è una funzione lineare e determina il nucleo e l'immagine di T .

Esercizio 4.2. Considera l'applicazione lineare $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ data da

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x - y + z \\ x + y + z \\ y - z \end{pmatrix}$$

Scegli una base \mathcal{B} del piano $U = \{x + y - z = 0\}$ e scrivi la matrice associata alla restrizione $f|_U: U \rightarrow \mathbb{R}^3$ rispetto a \mathcal{B} e alla base canonica di \mathbb{R}^3 .

Esercizio 4.3. Considera l'applicazione lineare $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$,

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x + y \\ x - y \\ 3x + 2y \end{pmatrix}$$

Determina il nucleo e l'immagine di f . Scrivi la matrice associata a f rispetto alle basi canoniche di \mathbb{R}^2 e \mathbb{R}^3 e rispetto alle basi seguenti:

$$\left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}, \quad \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Esercizio 4.4. Scrivi la matrice di cambiamento di base $[\text{id}]_{\mathcal{B}'}^{\mathcal{B}}$ fra le due basi seguenti di $\mathbb{R}_2[x]$:

$$\mathcal{B} = \{1 + x + x^2, x, 1 - x^2\}, \quad \mathcal{B}' = \{1 - x, x - x^2, 1\}.$$

Esercizio 4.5. Sia $f: V \rightarrow W$ un'applicazione lineare. Mostra che:

- Se $U \subset V$ è un sottospazio, allora $f(U)$ è un sottospazio di W .
- Se $Z \subset W$ è un sottospazio, allora $f^{-1}(Z)$ è un sottospazio di V .
- Se f è un isomorfismo, allora $\dim f(U) = \dim U$ per ogni sottospazio $U \subset V$.

Esercizio 4.6. Sia $T: \mathbb{R}_n[x] \rightarrow \mathbb{R}_n[x]$ la funzione $T(p) = (x+1)p'(x)$ dove $p'(x)$ è la derivata di $p(x)$. Scrivi T rispetto alla base canonica, dimostra che T è un isomorfismo e calcola la sua traccia e il suo determinante in funzione di n .

Esercizio 4.7. Sia $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Sia $L_A: M(2, \mathbb{R}) \rightarrow M(2, \mathbb{R})$ l'applicazione lineare $L_A(X) = AX$.

- (1) Determina nucleo e immagine di L_A .
- (2) Scrivi la matrice associata a L_A rispetto alla base canonica di $M(2, \mathbb{R})$.
- (3) Calcola il determinante di L_A .
- (4) Mostra che per una $A \in M(2)$ generica, l'applicazione $L_A: M(2) \rightarrow M(2)$ è un isomorfismo $\iff A$ è invertibile. Mostra inoltre che $\det L_A = (\det A)^2$.

Esercizio 4.8. Sia $A \in M(2)$ fissata e $T: M(2) \rightarrow M(2)$ l'applicazione $T(X) = AX - XA$. Mostra che T è lineare. Inoltre:

- (1) Se $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, determina il nucleo e l'immagine di T .
- (2) Dimostra che per qualsiasi A l'applicazione T non è mai un isomorfismo.
- (3) Dimostra che per qualsiasi A l'immagine di T ha dimensione ≤ 2 .

Esercizio 4.9. Determina una matrice $A \in M(3, \mathbb{R})$ tale che l'endomorfismo L_A soddisfi le proprietà seguenti: $\text{Im } L_A = \{x - 2y + z = 0\}$ e $\ker L_A = \text{Span}(e_1 + e_2 + e_3)$.

Esercizio 4.10. Siano $f: \mathbb{C}^3 \rightarrow \mathbb{C}^2$ e $g: \mathbb{C}^2 \rightarrow \mathbb{C}^3$ due applicazioni lineari. La composizione $f \circ g$ può essere un isomorfismo? E la composizione $g \circ f$? In caso affermativo fornisci un esempio, in caso negativo dimostra che non è possibile.

Esercizio 4.11. Siano $U = \{x + y = 0\}$ e $W = \text{Span}(e_2 + e_3)$ sottospazi di \mathbb{R}^3 . Dimostra che $\mathbb{R}^3 = U \oplus W$ e scrivi la matrice associata alle proiezioni p_U e p_W indotte dalla somma diretta rispetto alla base canonica di \mathbb{R}^3 .

Esercizio 4.12. Sia $f: V \rightarrow W$ un'applicazione lineare fra due spazi V e W di dimensione finita. Mostra che esistono sempre due basi \mathcal{B} e \mathcal{C} per V e W tale che

$$[f]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

dove $r = \dim \text{Im } f$ e la descrizione è a blocchi (alcuni blocchi 0 potrebbero non essere presenti).

Esercizio 4.13. Sia $f: V \rightarrow W$ un'applicazione lineare fra due spazi V e W di dimensione finita. Mostra che:

- f è iniettiva \iff esiste una $g: W \rightarrow V$ lineare tale che $g \circ f = \text{id}_V$.
- f è suriettiva \iff esiste una $g: W \rightarrow V$ lineare tale che $f \circ g = \text{id}_W$.

Esercizio 4.14. Siano $x_0, \dots, x_n \in \mathbb{K}$ punti diversi. Considera l'applicazione lineare $T: \mathbb{K}_n[x] \rightarrow \mathbb{K}^{n+1}$,

$$T(p) = \begin{pmatrix} p(x_0) \\ \vdots \\ p(x_n) \end{pmatrix}.$$

Dimostra che T è un isomorfismo. Qual è il legame con l'Esercizio 3.12?

Esercizio 4.15. Sia V uno spazio vettoriale con base $\mathcal{B} = \{v_1, \dots, v_n\}$ e $f: V \rightarrow V$ un endomorfismo. Sia \mathcal{B}' ottenuta da \mathcal{B} scambiando v_i e v_j . Mostra che $[f]_{\mathcal{B}'}$ è ottenuta da $[f]_{\mathcal{B}}$ scambiando le righe i e j e le colonne i e j .

Complementi

Spazio duale.

Autovettori e autovalori

5.1. Definizioni

Sappiamo che qualsiasi endomorfismo è rappresentabile come una matrice quadrata A dopo aver fissato una base \mathcal{B} . La domanda a cui tentiamo di rispondere in questo capitolo è la seguente: fra le infinite basi \mathcal{B} possibili, ne esistono alcune che sono più convenienti, ad esempio perché la matrice associata A risulta essere più semplice da gestire? Per rispondere a questa domanda introduciamo alcuni strumenti quali gli *autovettori*, gli *autovalori* e infine il *polinomio caratteristico*.

5.1.1. Autovettori e autovalori. Sia $T: V \rightarrow V$ un endomorfismo di uno spazio vettoriale V definito su un campo \mathbb{K} . Un *autovettore* di T è un vettore $v \neq 0$ in V per cui

$$T(v) = \lambda v$$

per qualche numero $\lambda \in \mathbb{K}$ che chiameremo *autovalore di T relativo a v* . Notiamo che λ può essere qualsiasi scalare, anche zero. D'altro canto, rimarchiamo che l'autovettore v non può essere zero per definizione. Facciamo alcuni esempi.

Esempio 5.1.1. Consideriamo l'endomorfismo $L_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ con

$$A = \begin{pmatrix} 3 & 4 \\ 0 & 2 \end{pmatrix}.$$

Poiché $L_A(e_1) = 3e_1$, il vettore e_1 è autovettore di L_A con autovalore 3. Inoltre notiamo che $L_A \begin{pmatrix} -4 \\ 1 \end{pmatrix} = \begin{pmatrix} -8 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} -4 \\ 1 \end{pmatrix}$ e quindi il vettore $\begin{pmatrix} -4 \\ 1 \end{pmatrix}$ è autovettore con autovalore 2.

Esempio 5.1.2. Sia V uno spazio vettoriale su \mathbb{K} . Fissiamo $\lambda \in \mathbb{K}$ e definiamo $f: V \rightarrow V, f(v) = \lambda v$. Tutti i vettori non nulli di V sono autovettori con lo stesso autovalore λ .

Esempio 5.1.3. Sia $L_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ con $A = \text{Rot}_\theta$ una rotazione di angolo $\theta \neq 0, \pi$. Ciascun vettore $v \in \mathbb{R}^2$ diverso da zero viene ruotato di un angolo $\theta \neq 0, \pi$ e quindi chiaramente l'immagine $L_A(v)$ non può essere un multiplo di v . L'endomorfismo L_A non ha autovettori.

Abbiamo visto tre esempi qualitativamente molto differenti: nel primo endomorfismo alcuni vettori dello spazio vettoriale sono autovettori, nel secondo tutti (eccetto lo 0) e nel terzo nessuno.

Facciamo adesso alcune osservazioni semplici ma cruciali.

Osservazione 5.1.4. Sia $f: V \rightarrow V$ un endomorfismo. Se $v \in V$ è autovettore per f con autovalore λ , allora qualsiasi multiplo $w = \mu v$ di v con $\mu \neq 0$ è anch'esso autovettore con lo stesso autovalore λ , infatti

$$f(\mu v) = \mu f(v) = \mu \lambda v = \lambda(\mu v).$$

Se $v \in V$ è autovettore, tutti i vettori non nulli nella retta $\text{Span}(v)$ sono anche loro autovettori con lo stesso autovalore λ .

Gli autovettori con autovalore 0 e 1 sono un po' speciali.

Osservazione 5.1.5. Sia $f: V \rightarrow V$ un endomorfismo. Un vettore $v \neq 0$ è autovettore per f con autovalore 0 se e solo se $f(v) = 0v = 0$, cioè se e solo se $v \in \ker f$.

Un *punto fisso* per un endomorfismo f è un vettore $v \in V$ tale che $f(v) = v$. I punti fissi sono punti che non vengono spostati da f .

Osservazione 5.1.6. Sia $f: V \rightarrow V$ un endomorfismo. Un vettore $v \neq 0$ è autovettore per f con autovalore 1 se e solo se v è un punto fisso.

Esempio 5.1.7. Nella rotazione intorno all'asse z dell'Esempio 4.4.11, tutti i vettori contenuti nell'asse z non vengono spostati da f e sono punti fissi. Tutti i vettori nell'asse z tranne l'origine sono autovettori con autovalore 1.

Riassumendo:

- Se v è autovettore con autovalore λ , ogni suo multiplo lo è.
- Gli autovettori con autovalore 0 sono i vettori non nulli del nucleo.
- Gli autovettori con autovalore 1 sono i punti fissi non nulli.

Osservazione 5.1.8. Come molte altre cose viste qui, autovettori e autovalori possono essere agevolmente studiati in coordinate rispetto ad una base. Sia $T: V \rightarrow V$ un endomorfismo di uno spazio vettoriale V . Siano \mathcal{B} una base di V e $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ la matrice associata. Grazie alla Proposizione 4.3.4, troviamo

$$T(v) = \lambda v \iff Ax = \lambda x$$

se $x = [v]_{\mathcal{B}} \in \mathbb{K}^n$. Ogni qual volta vediamo l'equazione $T(v) = \lambda v$, possiamo leggerla in coordinate come $Ax = \lambda x$.

5.1.2. Endomorfismi diagonalizzabili. Veniamo adesso al vero motivo per cui abbiamo introdotto autovettori e autovalori.

Definizione 5.1.9. Un endomorfismo $T: V \rightarrow V$ è *diagonalizzabile* se V ha una base $\mathcal{B} = \{v_1, \dots, v_n\}$ composta da autovettori per T .

Il termine “diagonalizzabile” è dovuto al fatto seguente, che è cruciale. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base qualsiasi di V .

Proposizione 5.1.10. *La matrice associata $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ è diagonale se e solo se i vettori v_1, \dots, v_n sono tutti autovettori per T .*

Dimostrazione. Il vettore v_i è un autovettore $\iff f(v_i) = \lambda_i v_i$ per qualche $\lambda_i \in \mathbb{K} \iff [f(v_i)]_{\mathcal{B}} = \lambda_i e_i$. Questo capita per ogni $i = 1, \dots, n \iff A$ è diagonale, con gli autovalori sulla diagonale principale:

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

La dimostrazione è conclusa. \square

Abbiamo scoperto che un endomorfismo T è diagonalizzabile se e solo se esiste una base \mathcal{B} tale che $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ sia una matrice diagonale. Questo accade precisamente quando \mathcal{B} è una base di autovettori, e gli elementi sulla diagonale principale di A sono precisamente i loro autovalori.

Esempio 5.1.11. L'endomorfismo $L_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ con $A = \begin{pmatrix} 3 & 4 \\ 0 & 2 \end{pmatrix}$ già considerato nell'Esempio 5.1.1 è diagonalizzabile, perché come abbiamo notato precedentemente $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ e $v_2 = \begin{pmatrix} -4 \\ 1 \end{pmatrix}$ sono entrambi autovettori e sono indipendenti, quindi formano una base. I loro autovalori sono 3 e 2. Prendendo $\mathcal{B} = \{v_1, v_2\}$ otteniamo quindi

$$[L_A]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}.$$

Esempio 5.1.12. L'endomorfismo $f(v) = \lambda v$ considerato nell'Esempio 5.1.2 è sempre diagonalizzabile, perché *qualsiasi* base di V è una base di autovettori. La matrice associata è sempre λI_n .

Esempio 5.1.13. La rotazione di angolo θ considerata nell'Esempio 5.1.3 non è diagonalizzabile per $\theta \neq 0, \pi$ perché non ha autovettori. Per $\theta = 0$ e $\theta = \pi$ la rotazione diventa rispettivamente $f(v) = v$ e $f(v) = -v$ e quindi è diagonalizzabile.

Esempio 5.1.14. L'endomorfismo $T: M(2) \rightarrow M(2)$ definito da $T(A) = {}^t A$ è diagonalizzabile: una base di autovettori è data da

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Infatti otteniamo

$$T(A_1) = A_1, \quad T(A_2) = A_2, \quad T(A_3) = A_3, \quad T(A_4) = -A_4.$$

Prendendo $\mathcal{B} = \{A_1, A_2, A_3, A_4\}$ otteniamo quindi

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

5.1.3. Matrici diagonali. Un endomorfismo $T: V \rightarrow V$ è diagonalizzabile precisamente quando esiste una base \mathcal{B} rispetto alla quale la matrice associata $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ è diagonale. Perché preferiamo le matrici diagonali? Le preferiamo perché sono molto più maneggevoli.

Notiamo innanzitutto che il prodotto Ax tra una matrice quadrata A ed un vettore x si semplifica molto se A è diagonale:

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda_1 x_1 \\ \lambda_2 x_2 \\ \vdots \\ \lambda_n x_n \end{pmatrix}.$$

Questo vuol dire che se $[T]_{\mathcal{B}}^{\mathcal{B}}$ è diagonale è molto facile calcolare le immagini dei vettori. Il determinante di una matrice diagonale A è semplicemente il prodotto degli elementi sulla diagonale principale:

$$\det A = \lambda_1 \cdots \lambda_n.$$

Anche il prodotto fra matrici diagonali è molto semplice:

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \begin{pmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \mu_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mu_n \end{pmatrix} = \begin{pmatrix} \lambda_1 \mu_1 & 0 & \dots & 0 \\ 0 & \lambda_2 \mu_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \mu_n \end{pmatrix}.$$

Questo ci dice che comporre due endomorfismi che si diagonalizzano entrambi con la stessa base di autovettori è molto semplice. In particolare, possiamo calcolare facilmente potenze come A^{100} :

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \implies A^k = \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{pmatrix}.$$

5.1.4. Matrici diagonalizzabili. Abbiamo notato che le matrici diagonali hanno ottime proprietà. Chiamiamo *diagonalizzabili* le matrici che possono essere trasformate in matrici diagonali in un senso opportuno.

Definizione 5.1.15. Una matrice $A \in M(n, \mathbb{K})$ è *diagonalizzabile* se è simile ad una matrice diagonale D .

Quindi A è diagonalizzabile \iff esiste una matrice invertibile M tale che

$$D = M^{-1}AM$$

sia diagonale. Il collegamento con la nozione di endomorfismo diagonalizzabile è chiaramente molto forte:

Proposizione 5.1.16. *Sia \mathcal{B} una base di V . Un endomorfismo $T: V \rightarrow V$ è diagonalizzabile \iff la matrice associata $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ è diagonalizzabile.*

Dimostrazione. Se T è diagonalizzabile, esiste una base \mathcal{C} di V per cui $D = [T]_{\mathcal{C}}^{\mathcal{C}}$ sia diagonale; se $M = [\text{id}]_{\mathcal{B}}^{\mathcal{C}}$ è la matrice di cambiamento di base da \mathcal{C} a \mathcal{B} otteniamo $D = M^{-1}AM$.

D'altro canto, se $D = M^{-1}AM$ è diagonale per qualche M invertibile, sia \mathcal{C} la base di V formata da vettori le cui coordinate rispetto a \mathcal{B} sono le colonne di M . Allora chiaramente $M = [\text{id}]_{\mathcal{B}}^{\mathcal{C}}$ e quindi $[T]_{\mathcal{C}}^{\mathcal{C}} = M^{-1}AM = D$ \square

Sia $A \in M(n, \mathbb{K})$ una matrice quadrata.

Corollario 5.1.17. *A è diagonalizzabile \iff L_A è diagonalizzabile.*

Le matrici diagonalizzabili sono spesso più facili da gestire. Vediamo ad esempio come calcolare la potenza di una matrice diagonalizzabile. Notiamo intanto un paio di proprietà algebriche della similitudine fra matrici:

Proposizione 5.1.18. *Valgono i fatti seguenti:*

$$\begin{aligned} A = M^{-1}BM, C = M^{-1}DM &\implies AC = M^{-1}BDM, \\ A = M^{-1}BM &\implies A^k = M^{-1}B^kM. \end{aligned}$$

per qualsiasi $k \in \mathbb{N}$.

Dimostrazione. Troviamo

$$AC = M^{-1}BMM^{-1}DM = M^{-1}BDM.$$

La seconda implicazione segue dalla prima iterata k volte. \square

Esempio 5.1.19. Prendiamo $A = \begin{pmatrix} 3 & 4 \\ 0 & 2 \end{pmatrix}$ e calcoliamo A^{100} . La matrice A non è diagonale, quindi calcolare una potenza di A direttamente è estremamente dispendioso. Sappiamo però che A è diagonalizzabile: dalla discussione svolta nell'Esempio 5.1.11 deduciamo che $M^{-1}AM = D = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$, dove

$$M = [\text{id}]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} 1 & -4 \\ 0 & 1 \end{pmatrix} \implies M^{-1} = [\text{id}]_{\mathcal{B}}^{\mathcal{C}} = \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}.$$

Qui $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ 1 \end{pmatrix} \right\}$ e \mathcal{C} è la base canonica di \mathbb{R}^2 . Quindi

$$\begin{aligned} A^{100} &= (MDM^{-1})^{100} = MD^{100}M^{-1} \\ &= \begin{pmatrix} 1 & -4 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3^{100} & 0 \\ 0 & 2^{100} \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -4 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3^{100} & 4 \cdot 3^{100} \\ 0 & 2^{100} \end{pmatrix} = \begin{pmatrix} 3^{100} & 4 \cdot 3^{100} - 2^{102} \\ 0 & 2^{100} \end{pmatrix}. \end{aligned}$$

5.1.5. Rette invarianti. Oltre che per le loro buone proprietà algebriche, gli endomorfismi diagonalizzabili sono preferibili anche per fondati motivi geometrici. Vediamo adesso che una base di autovettori fornisce un utile sistema di *rette invarianti*.

Introduciamo una nozione molto generale.

Definizione 5.1.20. Sia $T: V \rightarrow V$ un endomorfismo. Un *sottospazio T -invariante* (o più semplicemente *invariante*) è un sottospazio vettoriale $U \subset V$ tale che $T(U) \subset U$.

Un sottospazio $U \subset V$ è T -invariante se l'immagine $T(u)$ di ciascun vettore $u \in U$ è ancora un vettore di U . I sottospazi banale $\{0\}$ e totale V sono sempre T -invarianti; trovare altri sottospazi T -invarianti può essere più difficile.

Siamo qui interessati alle rette invarianti. Determinare le rette invarianti di T è di fatto equivalente a trovare gli autovettori per T , in virtù della proposizione seguente.

Proposizione 5.1.21. Sia $T: V \rightarrow V$ un endomorfismo e $v \in V$ un vettore. La retta $r = \text{Span}(v)$ è T -invariante $\iff v$ è autovettore per T .

Dimostrazione. Se r è T -invariante, $T(v) \in \text{Span}(v)$ e quindi $T(v) = \lambda v$ per qualche $\lambda \in \mathbb{K}$. D'altro canto, se v è un autovettore ciascun elemento di $\text{Span}(v)$ è anch'esso autovettore e quindi $T(\text{Span}(v)) \subset \text{Span}(v)$. \square

Per definizione, un endomorfismo $T: V \rightarrow V$ è diagonalizzabile \iff esiste una base v_1, \dots, v_n di autovettori. Per quanto abbiamo appena visto, questo è equivalente all'esistenza di un sistema $\text{Span}(v_1), \dots, \text{Span}(v_n)$ di n rette invarianti, i cui generatori sono una base per V . Questa descrizione ha una forte valenza geometrica nei casi $V = \mathbb{R}^2$ e \mathbb{R}^3 .

5.1.6. Polinomio caratteristico. Vogliamo adesso tentare di rispondere alle seguenti domande: come facciamo a capire se un dato endomorfismo T sia diagonalizzabile? Come possiamo trovare una base \mathcal{B} che diagonalizzi T , se esiste?

Negli esempi già visti precedentemente gli autovettori erano già dati e dovevamo solo verificare che funzionassero. In generale, come troviamo gli autovettori di un endomorfismo T ? Uno strumento cruciale a questo scopo è il *polinomio caratteristico*.

Definizione 5.1.22. Sia $A \in M(n, \mathbb{K})$. Il *polinomio caratteristico* di $A = (a_{ij})$ è definito nel modo seguente:

$$p_A(\lambda) = \det(A - \lambda I_n) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix}.$$

Un commento sulla notazione: normalmente un polinomio viene indicato con il simbolo $p(x)$, in cui x rappresenta la variabile; qui generalmente usiamo λ invece di x , perché il simbolo x è spesso usato per indicare un punto di \mathbb{K}^n . Il pedice A in p_A indica che il polinomio dipende dalla matrice A e a volte può essere omesso.

Da questa definizione un po' oscura di $p_A(\lambda)$ non risulta affatto chiaro che questo oggetto sia realmente un polinomio. Dimostriamo adesso che è un polinomio e identifichiamo alcuni dei suoi coefficienti.

Proposizione 5.1.23. *Il polinomio caratteristico $p_A(\lambda)$ è effettivamente un polinomio. Ha grado n e quindi è della forma*

$$p_A(\lambda) = a_n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_0.$$

Possiamo identificare alcuni dei suoi coefficienti:

- $a_n = (-1)^n$,
- $a_{n-1} = (-1)^{n-1} \text{tr} A$,
- $a_0 = \det A$.

Dimostrazione. Il caso $n = 1$ è molto facile: la matrice è $A = (a)$ e

$$p_A(\lambda) = \det(A - \lambda I) = \det(a - \lambda) = a - \lambda = -\lambda + a.$$

Otteniamo un polinomio di grado 1 con i coefficienti giusti (qui $\det A = \text{tr} A = a$). Il caso $n = 2$ è istruttivo: qui $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ e quindi

$$\begin{aligned} p_A(\lambda) &= \det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} = (a - \lambda)(d - \lambda) - bc \\ &= \lambda^2 - (a + d)\lambda + ad - bc = \lambda^2 - \text{tr} A \lambda + \det A. \end{aligned}$$

Dimostriamo adesso il caso generale. Usando l'interpretazione dell'Osservazione 3.3.1, scriviamo

$$p_A(\lambda) = \det(A - \lambda I_n) = \sum_{c \in \text{Col}(A)} p(c).$$

Qui $p(c)$ è il peso della colorazione c . Il peso $p(c)$ è il prodotto degli elementi nelle caselle rosse di $A - \lambda I_n$ moltiplicato per un segno, quindi è chiaramente

un polinomio in λ . L'unica colorazione c che può dare un polinomio di grado n è quella sulla diagonale principale, che ha peso

$$\begin{aligned} p(c) &= (a_{11} - \lambda) \cdots (a_{nn} - \lambda) \\ &= (-1)^n \lambda^n + (-1)^{n-1} (a_{11} + \dots + a_{nn}) \lambda^{n-1} + q(\lambda) \\ &= (-1)^n \lambda^n + (-1)^{n-1} \operatorname{tr} A \lambda^{n-1} + q(\lambda) \end{aligned}$$

dove $q(\lambda)$ è un polinomio di grado $\leq n-2$. Tutte le altre colorazioni c hanno come peso $p(c)$ un polinomio di grado $\leq n-2$. Quindi $p_A(\lambda)$ è un polinomio di grado n e i coefficienti a_n e a_{n-1} sono come indicato.

Resta solo da dimostrare che il termine noto a_0 è il determinante di A . Per fare ciò ricordiamo che $a_0 = p_A(0)$ e scriviamo semplicemente

$$a_0 = p_A(0) = \det(A - 0I) = \det A.$$

La dimostrazione è completa. \square

Come il determinante e la traccia, il polinomio caratteristico è invariante per similitudine:

Proposizione 5.1.24. *Se A e B sono matrici simili, allora $p_A(\lambda) = p_B(\lambda)$.*

Dimostrazione. Per ipotesi $A = M^{-1}BM$ per qualche matrice invertibile M . Notiamo che $\lambda I_n = \lambda M^{-1}M = M^{-1}(\lambda I_n)M$. Otteniamo quindi

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I_n) = \det(M^{-1}BM - M^{-1}\lambda I_n M) \\ &= \det(M^{-1}(B - \lambda I_n)M) = \det(M^{-1}) \det(B - \lambda I_n) \det M \\ &= \det(B - \lambda I_n) = p_B(\lambda) \end{aligned}$$

grazie al Teorema di Binet. \square

Sia adesso $T: V \rightarrow V$ un endomorfismo di uno spazio vettoriale V di dimensione n . Definiamo il polinomio caratteristico $p_T(\lambda)$ di T come il polinomio caratteristico $p_A(\lambda)$ della matrice associata $A = [T]_{\mathcal{B}}$ rispetto ad una qualsiasi base \mathcal{B} di V . La definizione non dipende dalla base scelta perché il polinomio caratteristico è invariante per similitudine.

Osservazione 5.1.25. Possiamo definire il polinomio caratteristico di un endomorfismo $T: V \rightarrow V$ direttamente scrivendo $p_T(\lambda) = \det(T - \lambda \operatorname{id}_V)$. Qui $T - \lambda \operatorname{id}_V$ è l'endomorfismo che manda v in $T(v) - \lambda v$.

Esempio 5.1.26. Consideriamo l'endomorfismo $T: V \rightarrow V$ dato da $f(v) = \lambda_0 v$ con $\lambda_0 \in \mathbb{K}$ fissato. La matrice associata a T rispetto a qualsiasi base è $\lambda_0 I_n$. Quindi

$$p_T(\lambda) = \det(\lambda_0 I_n - \lambda I_n) = (\lambda_0 - \lambda)^n.$$

Esempio 5.1.27. Come abbiamo già rimarcato, il polinomio caratteristico di una matrice $A \in M(2, \mathbb{K})$ è

$$p_A(\lambda) = \lambda^2 - \operatorname{tr} A \lambda + \det A.$$

Il polinomio caratteristico di una matrice $A \in M(3, \mathbb{K})$ è

$$\rho_A(\lambda) = -\lambda^3 + \operatorname{tr}A \lambda^2 + a_1 \lambda + \det A.$$

Tutti i coefficienti di $\rho_A(\lambda)$ sono invarianti per similitudine: qui a_1 è un nuovo numero invariante per similitudine, oltre alla traccia e il determinante.

Vedremo successivamente altri esempi; prima spieghiamo il motivo per cui abbiamo introdotto il polinomio caratteristico.

5.1.7. Le radici del polinomio caratteristico. Il polinomio caratteristico è uno strumento utile a trovare autovalori e autovettori di un endomorfismo. Sia V uno spazio di dimensione n e $T: V \rightarrow V$ un endomorfismo.

Proposizione 5.1.28. *Gli autovalori di T sono precisamente le radici del polinomio caratteristico $\rho_T(\lambda)$.*

Dimostrazione. Scegliamo una base \mathcal{B} e scriviamo $A = [T]_{\mathcal{B}}^{\mathcal{B}}$. Per l'Osservazione 5.1.8, uno scalare $\lambda \in \mathbb{K}$ è autovalore per T se e solo se esiste un $x \in \mathbb{K}^n$ non nullo tale che $Ax = \lambda x$. I fatti seguenti sono tutti equivalenti:

$$\begin{aligned} \exists x \neq 0 \text{ tale che } Ax = \lambda x &\iff \\ \exists x \neq 0 \text{ tale che } Ax - \lambda x &= 0 \iff \\ \exists x \neq 0 \text{ tale che } (A - \lambda I_n)x &= 0 \iff \\ \exists x \neq 0 \text{ tale che } x \in \ker(A - \lambda I_n) &\iff \\ \det(A - \lambda I_n) = 0 &\iff \rho_A(\lambda) = 0. \end{aligned}$$

La dimostrazione è completa. \square

Abbiamo scoperto un metodo per identificare gli autovalori di T : dobbiamo solo (si fa per dire...) trovare le radici del polinomio caratteristico $\rho_T(\lambda)$. Questo ovviamente non è un problema semplice da risolvere in generale: non esiste una formula per trovare le radici di un polinomio di grado qualsiasi. Per i polinomi di grado due, e quindi per le matrici 2×2 , il problema è però sempre risolvibile. Vediamo alcuni esempi 2×2 .

Esempio 5.1.29. Consideriamo l'endomorfismo $L_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ con

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

Il suo polinomio caratteristico è

$$\rho_A(\lambda) = \lambda^2 - \operatorname{tr}A \lambda + \det A = \lambda^2 - 4\lambda + 3.$$

Le sue radici sono $\lambda = 1$ e $\lambda = 3$. Quindi gli autovalori di A sono 1 e 3. Per ciascun autovalore $\lambda = 1$ e $\lambda = 3$, gli autovettori relativi sono precisamente le soluzioni non banali dei sistemi $A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$ e $A \begin{pmatrix} x \\ y \end{pmatrix} = 3 \begin{pmatrix} x \\ y \end{pmatrix}$. Questi sono:

$$\begin{cases} -x + 2y = x \\ -4x + 5y = y \end{cases}, \quad \begin{cases} -x + 2y = 3x \\ -4x + 5y = 3y \end{cases}$$

Questi sono due sistemi omogenei e le soluzioni sono rispettivamente due rette

$$\text{Span} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \text{Span} \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Due generatori qualsiasi di queste rette (ad esempio $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ e $v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$) formano una base di autovettori per L_A . La matrice A è quindi diagonalizzabile, ed è simile alla matrice $D = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$. Con $M = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ otteniamo $D = M^{-1}AM$.

Esempio 5.1.30. Consideriamo una rotazione $L_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ antioraria di un angolo $\theta \neq 0, \pi$. La rotazione è descritta dalla matrice $A = \text{Rot}_\theta$ con

$$\text{Rot}_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Ragionando geometricamente, abbiamo già dimostrato che L_A non ha autovettori. Possiamo verificarlo anche algebricamente guardando il polinomio caratteristico $p(\lambda) = \lambda^2 - 2\cos\theta + 1$ e notando che ha due radici complesse non reali, perché $\Delta = 4(\cos^2\theta - 1) < 0$. Il polinomio caratteristico non ha radici reali: quindi non ci sono autovalori, e di conseguenza neppure autovettori.

Il prossimo esempio descrive un fenomeno rilevante: la diagonalizzabilità di una matrice A può dipendere dal campo \mathbb{K} che stiamo considerando.

Esempio 5.1.31. Consideriamo l'endomorfismo $L_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ con

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

Questo è una rotazione di $\frac{\pi}{2}$ e quindi sappiamo già che non è diagonalizzabile su \mathbb{R} . Se però consideriamo l'endomorfismo $L_A: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ con la stessa matrice A le cose cambiano: il polinomio caratteristico è infatti $p(\lambda) = \lambda^2 + 1$ ed ha soluzioni $\lambda = \pm i$. Studiamo i sistemi $A \begin{pmatrix} x \\ y \end{pmatrix} = i \begin{pmatrix} x \\ y \end{pmatrix}$ e $A \begin{pmatrix} x \\ y \end{pmatrix} = -i \begin{pmatrix} x \\ y \end{pmatrix}$:

$$\begin{cases} -y = ix \\ x = iy \end{cases}, \quad \begin{cases} -y = -ix \\ x = -iy \end{cases}$$

Le soluzioni sono le rette

$$\text{Span} \begin{pmatrix} i \\ 1 \end{pmatrix}, \quad \text{Span} \begin{pmatrix} -i \\ 1 \end{pmatrix}.$$

I vettori $v_1 = \begin{pmatrix} i \\ 1 \end{pmatrix}$ e $v_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix}$ sono autovettori di $L_A: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ con autovalori i e $-i$. Quindi la matrice A non è diagonalizzabile su \mathbb{R} , ma lo è su \mathbb{C} . Prendendo $\mathcal{B} = \{v_1, v_2\}$ otteniamo

$$[L_A]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

In virtù di questo esempio, quando parliamo di diagonalizzabilità di matrici dovremo sempre specificare chiaramente con quale campo \mathbb{K} stiamo lavorando. Nel prossimo esempio mostriamo una matrice che non è diagonalizzabile né su \mathbb{R} né su \mathbb{C} :

Esempio 5.1.32. Consideriamo l'endomorfismo $L_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ con

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Il polinomio caratteristico è $p(\lambda) = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2$. Il polinomio ha una radice $\lambda = 1$ con molteplicità due. Quindi $\lambda = 1$ è l'unico autovalore. Risolvendo il sistema $A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$ si trova che lo spazio delle soluzioni è la retta $\text{Span}(e_1)$. Troviamo un autovettore per L_A , ma non ne troviamo due indipendenti: quindi L_A non è diagonalizzabile. Con lo stesso ragionamento vediamo che neppure $L_A: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ è diagonalizzabile. La matrice A non è diagonalizzabile su nessun campo, né \mathbb{R} né \mathbb{C} .

Riassumendo, abbiamo esaminato vari tipi di esempi:

- La matrice $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ è già diagonale e quindi è diagonalizzabile sia su \mathbb{C} che su \mathbb{R} . Ha un solo autovalore 1.
- La matrice $\begin{pmatrix} -1 & 2 \\ -4 & 5 \end{pmatrix}$ è diagonalizzabile sia su \mathbb{C} che su \mathbb{R} . Ha due autovalori 1 e 3.
- La matrice $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ è diagonalizzabile su \mathbb{C} ma non su \mathbb{R} . Ha due autovalori $\pm i$ su \mathbb{C} .
- La matrice $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ non è diagonalizzabile né su \mathbb{C} né su \mathbb{R} . Ha un solo autovalore 1.

Esercizio 5.1.33. Per ciascuna delle matrici seguenti, determinare autovalori e autovettori in \mathbb{R}^3 e dire se è diagonalizzabile:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

5.1.8. Matrici triangolari. Ricordiamo che una matrice quadrata A è *triangolare superiore* se è della forma

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

Il determinante di una matrice triangolare A è semplicemente $\det A = a_{11} \cdots a_{nn}$. Anche gli autovalori si determinano immediatamente:

Proposizione 5.1.34. *Gli autovalori di una matrice triangolare A sono gli elementi a_{11}, \dots, a_{nn} sulla diagonale principale.*

Dimostrazione. Vale

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ 0 & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} - \lambda \end{pmatrix} \\ &= (a_{11} - \lambda) \cdots (a_{nn} - \lambda). \end{aligned}$$

Le radici di p_A sono quindi a_{11}, \dots, a_{nn} . Il caso in cui A sia triangolare inferiore è analogo. \square

Notiamo che gli elementi a_{11}, \dots, a_{nn} possono anche ripetersi. Esaminiamo anche il caso in cui una matrice sia descritta a blocchi, con il blocco in basso a sinistra nullo.

Proposizione 5.1.35. *Sia $A \in M(n)$ del tipo*

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

con $B \in M(k)$ e $D \in M(n-k)$ per qualche k . Vale la relazione

$$p_A(\lambda) = p_B(\lambda) \cdot p_D(\lambda).$$

Dimostrazione. Scriviamo

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I) = \det \begin{pmatrix} B - \lambda I & C \\ 0 & D - \lambda I \end{pmatrix} \\ &= \det(B - \lambda I) \cdot \det(D - \lambda I) = p_B(\lambda) \cdot p_D(\lambda). \end{aligned}$$

Nella terza uguaglianza abbiamo usato l'Esercizio 3.13. \square

Quindi gli autovalori di A sono quelli di B uniti a quelli di D .

5.1.9. Esistenza di autovettori. Concludiamo questa sezione con un paio di proposizioni che mostrano l'esistenza di autovettori in alcune condizioni.

Proposizione 5.1.36. *Un endomorfismo $T: V \rightarrow V$ di uno spazio vettoriale complesso V ha sempre almeno un autovettore.*

Dimostrazione. Il polinomio caratteristico ha sempre almeno una radice (su \mathbb{C}). Quindi T ha almeno un autovalore λ , e quindi anche un autovettore. \square

Questo risultato non vale su \mathbb{R} . Ad esempio, abbiamo visto che una rotazione di un angolo $\theta \neq 0, \pi$ è una trasformazione di \mathbb{R}^2 che non ha autovettori. Possiamo comunque ottenere lo stesso risultato in dimensione dispari:

Proposizione 5.1.37. *Sia V uno spazio vettoriale reale di dimensione dispari. Un endomorfismo $T: V \rightarrow V$ ha sempre almeno un autovettore.*

Dimostrazione. Il polinomio caratteristico ha grado dispari e quindi ha sempre una radice in \mathbb{R} per la Proposizione 1.4.13. \square

Da tutte queste considerazioni algebriche deduciamo un corollario geometrico che riguarda lo spazio tridimensionale.

Corollario 5.1.38. *Un endomorfismo di \mathbb{R}^3 ha sempre una retta invariante.*

5.2. Teorema di diagonalizzabilità

Nelle pagine precedenti abbiamo descritto un metodo per trovare gli autovalori di un endomorfismo tramite il suo polinomio caratteristico. In questa sezione ci proponiamo di determinare e studiare gli autovettori. Concludiamo infine il capitolo con il *teorema di diagonalizzabilità*, che fornisce un criterio completo per capire se un endomorfismo sia diagonalizzabile.

5.2.1. Autovettori con autovalori distinti. Dimostriamo alcune proprietà algebriche degli autovettori. Sia $T: V \rightarrow V$ un endomorfismo.

Proposizione 5.2.1. *Se $v_1, \dots, v_k \in V$ sono autovettori per T con autovalori $\lambda_1, \dots, \lambda_k$ distinti, allora sono indipendenti.*

Dimostrazione. Procediamo per induzione su k . Se $k = 1$, il vettore v_1 è indipendente semplicemente perché non è nullo (per definizione, un autovettore non è mai nullo).

Diamo per buono il caso $k - 1$ e mostriamo il caso k . Supponiamo di avere una combinazione lineare nulla

$$(6) \quad \alpha_1 v_1 + \dots + \alpha_k v_k = 0.$$

Dobbiamo dimostrare che $\alpha_i = 0$ per ogni i . Applicando T otteniamo

$$(7) \quad 0 = \alpha_1 T(v_1) + \dots + \alpha_k T(v_k) = \alpha_1 \lambda_1 v_1 + \dots + \alpha_k \lambda_k v_k.$$

Moltiplicando l'equazione (6) per λ_k ricaviamo

$$\alpha_1 \lambda_k v_1 + \dots + \alpha_k \lambda_k v_k = 0$$

e sottraendo questa equazione dalla (7) deduciamo che

$$\alpha_1(\lambda_1 - \lambda_k)v_1 + \dots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} = 0.$$

Questa è una combinazione lineare nulla di $k - 1$ autovettori con autovalori distinti: per l'ipotesi induttiva tutti i coefficienti $\alpha_i(\lambda_i - \lambda_k)$ devono essere nulli. Poiché $\lambda_i \neq \lambda_k$, ne deduciamo che $\alpha_i = 0$ per ogni $i = 1, \dots, k - 1$ e usando (6) otteniamo anche $\alpha_k = 0$. La dimostrazione è completa. \square

Sia V uno spazio vettoriale su \mathbb{K} di dimensione n e $T: V \rightarrow V$ un endomorfismo. Indichiamo come sempre con $p_T(\lambda)$ il suo polinomio caratteristico.

Corollario 5.2.2. *Se il polinomio caratteristico $p_T(\lambda)$ ha n radici distinte in \mathbb{K} , l'endomorfismo T è diagonalizzabile.*

Dimostrazione. Le radici sono $\lambda_1, \dots, \lambda_n$ e ciascun λ_i è autovalore di qualche autovettore v_i . Gli autovettori v_1, \dots, v_n sono indipendenti per la Proposizione 5.2.1, quindi sono una base di V . Quindi V ha una base formata da autovettori, cioè è diagonalizzabile. \square

Rimarchiamo l'importanza del campo \mathbb{K} nell'enunciato del corollario: il criterio funziona se il polinomio caratteristico ha n radici distinte nello stesso campo \mathbb{K} su cui è definito V . Il criterio si applica anche alle matrici quadrate A , considerate sempre nello stesso campo \mathbb{K} .

Esempio 5.2.3. Come notato nell'Esempio 5.1.31, la matrice $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ non è diagonalizzabile su \mathbb{R} . Su \mathbb{C} la stessa matrice ha due autovalori distinti $\pm i$ e quindi è diagonalizzabile per il Corollario 5.2.2.

Esempio 5.2.4. Consideriamo una matrice triangolare A . Gli autovalori di A sono gli elementi a_{11}, \dots, a_{nn} sulla diagonale principale. Se gli autovalori a_{11}, \dots, a_{nn} sono tutti distinti, allora A è diagonalizzabile per il Corollario 5.2.2.

Abbiamo ottenuto un criterio di diagonalizzabilità molto economico: possiamo dimostrare in alcuni casi che una matrice è diagonalizzabile senza cercare una esplicita base di autovettori. Ad esempio, la matrice

$$A = \begin{pmatrix} 1 & 7 & -1 \\ 0 & 2 & 8 \\ 0 & 0 & 3 \end{pmatrix}$$

è diagonalizzabile perché ha tre autovalori distinti 1, 2, 3. Notiamo che trovare una base esplicita di autovettori non è immediato: per ciascun $\lambda = 1, 2, 3$ dobbiamo risolvere l'equazione $Av = \lambda v$ (il lettore è invitato a farlo).

Il teorema di diagonalizzabilità che vogliamo dimostrare sarà una generalizzazione di questo criterio. Per enunciare questo teorema dobbiamo introdurre ancora una nozione.

5.2.2. Autospazio. Sia $T: V \rightarrow V$ un endomorfismo. Per ogni autovalore λ di T definiamo l'*autospazio* $V_\lambda \subset V$ nel modo seguente:

$$V_\lambda = \{v \in V \mid T(v) = \lambda v\}.$$

L'autospazio V_λ consiste precisamente di tutti gli autovettori v con autovalore λ , più l'origine $0 \in V$ (ricordiamo che $0 \in V$ non è autovettore per definizione).

Per ogni autovalore λ di T , ha senso considerare l'endomorfismo $T - \lambda \text{id}$ di V che manda v in $T(v) - \lambda v$ e la proposizione seguente lo mette in relazione con l'autospazio V_λ .

Proposizione 5.2.5. *Per ogni autovalore λ di T abbiamo*

$$V_\lambda = \ker(T - \lambda \text{id}).$$

Dimostrazione. Troviamo

$$\begin{aligned} V_\lambda &= \{v \in V \mid T(v) = \lambda v\} = \{v \in V \mid T(v) - \lambda v = 0\} \\ &= \{v \in V \mid (T - \lambda \text{id})v = 0\} = \ker(T - \lambda \text{id}). \end{aligned}$$

La dimostrazione è conclusa. \square

Corollario 5.2.6. *L'autospazio $V_\lambda \subset V$ è un sottospazio vettoriale di V .*

Ricordiamo la somma diretta di più sottospazi definita nella Sezione 2.3.12. Sia V uno spazio vettoriale di dimensione finita. Mostriamo adesso che gli autospazi di un endomorfismo sono sempre in somma diretta.

Proposizione 5.2.7. *Sia $T: V \rightarrow V$ un endomorfismo e siano $\lambda_1, \dots, \lambda_k$ i suoi autovalori. I corrispondenti autospazi sono sempre in somma diretta:*

$$V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k}.$$

Dimostrazione. Dobbiamo mostrare che

$$V_{\lambda_i} \cap (V_{\lambda_1} + \dots + V_{\lambda_{i-1}} + V_{\lambda_{i+1}} + \dots + V_{\lambda_k}) = \{0\} \quad \forall i = 1, \dots, k.$$

Sia per assurdo $v \neq 0$ un vettore nell'intersezione. Per ipotesi $v \in V_{\lambda_i}$ e

$$v = v_1 + \dots + v_{i-1} + v_{i+1} + \dots + v_k$$

con $v_j \in V_{\lambda_j}$. Questa è una relazione di dipendenza lineare fra autovettori con autovalori distinti, un assurdo per la Proposizione 5.2.1. \square

Corollario 5.2.8. *L'endomorfismo T è diagonalizzabile se e solo se*

$$V = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k}.$$

Dimostrazione. Sappiamo già che gli autospazi sono in somma diretta, quindi dobbiamo mostrare che $V = V_{\lambda_1} + \dots + V_{\lambda_k} \iff$ esiste una base di autovettori per T .

(\Rightarrow) Unendo le basi degli V_{λ_i} otteniamo una base di autovettori per T .

(\Leftarrow) Ciascun vettore v è combinazione lineare di autovettori e quindi chiaramente $v \in V_{\lambda_1} + \dots + V_{\lambda_k}$. \square

5.2.3. Molteplicità algebrica e geometrica. Per enunciare il teorema di diagonalizzabilità abbiamo bisogno di introdurre due ultime definizioni. In tutta questa sezione V è uno spazio vettoriale di dimensione n .

Sia $T: V \rightarrow V$ un endomorfismo e sia λ un autovalore per T . La *molteplicità algebrica* $m_a(\lambda)$ è la molteplicità di λ come radice del polinomio caratteristico p_T . La *molteplicità geometrica* $m_g(\lambda)$ è la dimensione del sottospazio associato a λ , cioè

$$m_g(\lambda) = \dim V_\lambda.$$

Esempio 5.2.9. Sia $L_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ l'endomorfismo dato da

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

Il polinomio caratteristico è $p_A(\lambda) = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2$. Troviamo un solo autovalore $\lambda_1 = 1$, con molteplicità algebrica $m_a(1) = 2$. D'altra parte,

$$m_g(1) = \dim V_1 = \dim \ker(A - I) = \dim \ker \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 1.$$

Abbiamo quindi trovato in questo caso

$$m_a(1) = 2, \quad m_g(1) = 1.$$

L'esempio ci mostra che le due molteplicità possono essere diverse. In generale, la molteplicità geometrica è un numero che sta sempre tra 1 e la molteplicità algebrica:

Proposizione 5.2.10. *Sia $T: V \rightarrow V$ un endomorfismo. Per ogni autovalore λ_0 di T valgono le disuguaglianze*

$$1 \leq m_g(\lambda_0) \leq m_a(\lambda_0).$$

Dimostrazione. Se λ_0 è un autovalore, esiste almeno un autovettore $v \neq 0$ tale che $T(v) = \lambda_0 v$ e quindi l'autospazio V_{λ_0} ha dimensione almeno uno. Questo dimostra la prima disuguaglianza.

Per dimostrare la seconda, prendiamo una base $\{v_1, \dots, v_k\}$ di V_{λ_0} , formata da $k = m_g(\lambda_0) = \dim V_{\lambda_0}$ vettori, e completiamola a base $\{v_1, \dots, v_n\}$ di V . Per i primi k vettori della base abbiamo $T(v_i) = \lambda_0 v_i$. Quindi la matrice $[T]_B^B$ è della forma

$$[T]_B^B = \begin{pmatrix} \lambda_0 I_k & C \\ 0 & D \end{pmatrix}$$

dove $\lambda_0 I_k \in M(k)$ e $D \in M(n - k)$. Per il Corollario 5.1.35 il polinomio caratteristico di T è del tipo

$$p_T(\lambda) = p_{\lambda_0 I_k}(\lambda) \cdot p_D(\lambda) = (\lambda_0 - \lambda)^k \cdot p_D(\lambda).$$

Ne segue che λ_0 ha molteplicità almeno $k = m_g(\lambda_0)$ in $p_T(\lambda)$. In altre parole, vale la disuguaglianza $m_a(\lambda_0) \geq m_g(\lambda_0)$. \square

5.2.4. Matrici simili. Sia $A \in M(n, \mathbb{K})$ una matrice quadrata. Gli *autovettori* e *autovalori* di A , assieme alle loro molteplicità algebriche e geometriche, sono per definizione gli autovettori e autovalori di $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^n$.

Notiamo che le molteplicità algebrica e geometrica sono entrambe invarianti per similitudine di una matrice.

Proposizione 5.2.11. *Se A e B sono simili, hanno gli stessi autovalori con le stesse molteplicità algebriche e geometriche.*

Dimostrazione. Le matrici A e B hanno gli stessi autovalori con le stesse molteplicità algebriche perché hanno lo stesso polinomio caratteristico. Sia $m_g^X(\lambda)$ la molteplicità geometrica di λ nella matrice X . Vale $A = M^{-1}BM$ e

$$\begin{aligned} m_g^A(\lambda) &= \dim \ker(A - \lambda I_n) = \dim \ker(M^{-1}BM - M^{-1}\lambda I_n M) \\ &= \dim \ker(M^{-1}(B - \lambda I_n)M) = \dim \ker(B - \lambda I_n) = m_g^B(\lambda). \end{aligned}$$

Nella penultima uguaglianza abbiamo usato che due matrici simili hanno lo stesso rango e quindi anche la stessa dimensione del nucleo. \square

D'altro canto, due matrici simili A e B *non* hanno gli stessi autovettori.

5.2.5. Teorema di diagonalizzabilità. Possiamo finalmente enunciare il teorema seguente. Sia V uno spazio vettoriale su \mathbb{K} di dimensione n .

Teorema 5.2.12 (Teorema di diagonalizzabilità). *Un endomorfismo $T: V \rightarrow V$ è diagonalizzabile se e solo se valgono entrambi i fatti seguenti:*

- (1) $p_T(\lambda)$ ha n radici in \mathbb{K} , contate con molteplicità.
- (2) $m_a(\lambda) = m_g(\lambda)$ per ogni autovalore λ di T .

Dimostrazione. Sappiamo che gli autospazi sono in somma diretta e definiamo il sottospazio $W \subset V$ come la loro somma

$$W = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}.$$

Sappiamo che T è diagonalizzabile $\iff W = V \iff \dim W = n$. Inoltre

$$\begin{aligned} \dim W &= \dim V_{\lambda_1} + \cdots + \dim V_{\lambda_k} = m_g(\lambda_1) + \cdots + m_g(\lambda_k) \\ &\leq m_a(\lambda_1) + \cdots + m_a(\lambda_k) \leq n. \end{aligned}$$

Nella prima disuguaglianza abbiamo usato che $m_g(\lambda_i) \leq m_a(\lambda_i)$ e nella seconda che $\lambda_1, \dots, \lambda_k$ sono radici del polinomio caratteristico, che ha grado n e quindi ha al più n radici contate con molteplicità per il Teorema 1.3.5.

Chiaramente $\dim W = n$ se e solo se entrambe le disuguaglianze sono uguaglianze. La prima disuguaglianza è un uguaglianza precisamente quando $m_g(\lambda_i) = m_a(\lambda_i)$ per ogni i , e la seconda precisamente quando $p_T(\lambda)$ ha n radici in \mathbb{K} , contate con molteplicità. \square

5.2.6. Esempi. Mostriamo alcuni esempi in cui applichiamo il teorema di diagonalizzabilità.

Esempio 5.2.13. Studiamo la diagonalizzabilità su \mathbb{R} della matrice

$$A = \begin{pmatrix} 3 & 0 & 0 \\ -4 & -1 & -8 \\ 0 & 0 & 3 \end{pmatrix}.$$

Si vede facilmente che il polinomio caratteristico è

$$p_A(\lambda) = (3 - \lambda)(-1 - \lambda)(3 - \lambda),$$

quindi ha radici $\lambda_1 = 3$ con $m_a(\lambda_1) = 2$ e $\lambda_2 = -1$ con $m_a(\lambda_2) = 1$.

Tutte le radici di $p_A(\lambda)$ sono reali, quindi A è diagonalizzabile se e solo se le molteplicità algebriche e geometriche di ciascun autovalore coincidono. Per il secondo autovalore λ_2 è facile: la Proposizione 5.2.10 implica che $m_g(\lambda_2) = m_a(\lambda_2) = 1$ e quindi siamo a posto.

Dobbiamo concentrarci solo sull'autovalore λ_1 che ha $m_a(\lambda_1) = 2$. La Proposizione 5.2.10 ci dice che $m_g(\lambda_1)$ può essere 1 oppure 2: nel primo caso A non è diagonalizzabile, nel secondo sì. Facciamo i conti e troviamo

$$m_g(3) = \dim V_3 = \dim \ker(A - 3I) = 3 - \text{rk}(A - 3I).$$

Nell'ultima uguaglianza abbiamo usato il teorema della dimensione. Quindi

$$\begin{aligned} m_g(3) &= 3 - \text{rk} \begin{pmatrix} 3-3 & 0 & 0 \\ -4 & -1-3 & -8 \\ 0 & 0 & 3-3 \end{pmatrix} \\ &= 3 - \text{rk} \begin{pmatrix} 0 & 0 & 0 \\ -4 & -4 & -8 \\ 0 & 0 & 0 \end{pmatrix} = 3 - 1 = 2. \end{aligned}$$

Abbiamo scoperto che $m_a(\lambda_1) = m_g(\lambda_1) = 2$ e quindi A è diagonalizzabile.

Esempio 5.2.14. Studiamo la diagonalizzabilità su \mathbb{R} della matrice

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

al variare del parametro $t \in \mathbb{R}$. Si vede facilmente che il polinomio caratteristico è

$$p_A(\lambda) = (2 - \lambda)(\lambda^2 + t - 5).$$

Se $t > 5$, il membro di destra non ha radici reali: quindi $p_A(\lambda)$ ha una radice sola e A non è diagonalizzabile. Se $t \leq 5$, il polinomio ha tre radici reali

$$\lambda_1 = 2, \quad \lambda_2 = \sqrt{5-t}, \quad \lambda_3 = -\sqrt{5-t}.$$

Se le tre radici sono distinte, la matrice A è diagonalizzabile. Restano da considerare i casi in cui le tre radici non sono distinte, e cioè i casi $t = 1$ e $t = 5$. Questi due casi vanno analizzati separatamente con le tecniche mostrate nell'esempio precedente.

Se $t = 1$ gli autovalori sono $\lambda_1 = 2$, $\lambda_2 = 2$ e $\lambda_3 = -2$ e la matrice A è

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

Calcoliamo la molteplicità geometrica dell'autovalore 2:

$$m_g(2) = 3 - \text{rk} \begin{pmatrix} 1 & 5 & 1 \\ -1 & -5 & -1 \\ 0 & 0 & 0 \end{pmatrix} = 3 - 1 = 2.$$

Otteniamo $m_g(2) = 2 = m_a(2)$ e quindi A è diagonalizzabile.

Se $t = 5$ gli autovalori sono $\lambda_1 = 2, \lambda_2 = 0$ e $\lambda_3 = 0$ e la matrice A è

$$A = \begin{pmatrix} 3 & 9 & 1 \\ -1 & -3 & -1 \\ 0 & 0 & 2 \end{pmatrix}.$$

La molteplicità geometrica dell'autovalore 0 è $m_g(0) = 3 - \text{rk}(A) = 3 - 2 = 1 \neq 2 = m_a(0)$. Quindi A non è diagonalizzabile.

Riassumendo, la matrice A è diagonalizzabile se e solo se $t < 5$.

Esercizi

Esercizio 5.1. Quali delle matrici seguenti sono diagonalizzabili su \mathbb{R} ?

$$\begin{pmatrix} 2 & -1 & 1 \\ 6 & -3 & 4 \\ 3 & -2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ -2 & -2 & -2 \end{pmatrix}, \quad \begin{pmatrix} a & b & c \\ 0 & a & d \\ 0 & 0 & a \end{pmatrix},$$

$$\begin{pmatrix} 1 & k & 0 \\ 1 & k & 1 \\ 0 & -k & 1 \end{pmatrix}, \quad \begin{pmatrix} k & 0 & 0 \\ k+1 & -1 & 0 \\ k-1 & k & k^2 \end{pmatrix}, \quad \begin{pmatrix} k-2 & 0 & -4 & 0 \\ 0 & k & 2 & 0 \\ 0 & 0 & k+2 & 0 \\ 1 & 1 & k & 2-k \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 3 & -1 & -1 & -1 & -1 & -1 & 3 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 3 & -1 & -1 & -1 & -1 & -1 & 3 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 3 & -1 & -1 & -1 & -1 & -1 & 3 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 \end{pmatrix}.$$

Quelle che non sono diagonalizzabili su \mathbb{R} , lo sono su \mathbb{C} ? Le risposte dipendono dai parametri reali presenti nella matrice.

Esercizio 5.2. Per quali $t \in \mathbb{C}$ la matrice A studiata nell'Esempio 5.2.14 è diagonalizzabile su \mathbb{C} ?

Esercizio 5.3. Trova gli autovalori in \mathbb{C} e gli autovettori in \mathbb{C}^2 della matrice complessa $\begin{pmatrix} 1 & 2 \\ 1+i & i \end{pmatrix}$.

Esercizio 5.4. Data la matrice $A = \begin{pmatrix} 23 & -84 \\ 6 & -22 \end{pmatrix}$, determina A^{10} .

Esercizio 5.5. Sia $A \in M(2, \mathbb{R})$ una matrice fissata. Considera l'endomorfismo $L_A: M(2, \mathbb{R}) \rightarrow M(2, \mathbb{R})$ dato da $L_A(X) = AX$.

- Sia $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. Scrivi la matrice associata a L_A rispetto ad una base di $M(2, \mathbb{R})$ a tua scelta.

- Sia $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. L'endomorfismo L_A è diagonalizzabile?
- Dimostra che in generale l'endomorfismo L_A è diagonalizzabile $\iff A$ è diagonalizzabile.
- Scrivi una base di autovettori per L_A nel caso $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Esercizio 5.6. Considera l'endomorfismo $T: M(2) \rightarrow M(2)$ dato da $T(X) = {}^tX$.

- Scrivi la matrice associata a T rispetto alla base canonica di $M(2)$.
- L'endomorfismo T è diagonalizzabile?

Esercizio 5.7. Costruisci un endomorfismo di \mathbb{R}^4 senza autovettori. Più in generale, per ogni $n \geq 1$ costruisci un endomorfismo di \mathbb{R}^{2n} senza autovettori.

Esercizio 5.8. Sia $T: V \rightarrow V$ un endomorfismo invertibile. Mostra che T è diagonalizzabile se e solo se T^{-1} è diagonalizzabile.

Esercizio 5.9. È sempre vero che la composizione di due endomorfismi diagonalizzabili è diagonalizzabile? Se sì, dimostrarlo. Se no, esibisci un controesempio.

Esercizio 5.10. Mostra che se $T: V \rightarrow V$ è diagonalizzabile allora V si decompone in somma diretta come $V = \ker T \oplus \operatorname{Im} T$.

Esercizio 5.11. Mostra che una matrice $A \in M(n, \mathbb{K})$ è diagonalizzabile se e solo se lo è la sua trasposta tA .

Esercizio 5.12. Mostra che una matrice $A \in M(n, \mathbb{K})$ di rango 1 è diagonalizzabile se e solo se $\ker L_A \oplus \operatorname{Im} L_A = \mathbb{K}^n$.

Esercizio 5.13. Considera l'endomorfismo $T: \mathbb{R}_n[x] \rightarrow \mathbb{R}_n[x]$ dato da

$$T(p(x)) = (x+1)p'(x)$$

dove p' è la derivata di p . L'endomorfismo T è diagonalizzabile?

Esercizio 5.14. Sia $T: V \rightarrow V$ un endomorfismo tale che $T \circ T = T$.

- (1) Mostra che $V = \ker T \oplus \operatorname{Im} T$.
- (2) Sia $V_\lambda = \{v \in V \mid T(v) = \lambda v\}$. Mostra che $V_0 = \ker T$ e $V_1 = \operatorname{Im} T$.
- (3) Concludi che T è diagonalizzabile.

Esercizio 5.15. Sia $T: V \rightarrow V$ un endomorfismo tale che $T \circ T = \operatorname{id}$.

- (1) Sia $V_\lambda = \{v \in V \mid T(v) = \lambda v\}$. Mostra che $V = V_1 \oplus V_{-1}$. Il trucco è scrivere ogni vettore $v \in V$ come somma

$$v = \frac{v + T(v)}{2} + \frac{v - T(v)}{2}.$$

- (2) Concludi che T è diagonalizzabile.

Esercizio 5.16. Sia $T: V \rightarrow V$ un endomorfismo tale che $T \circ T = 0$.

- (1) Mostra che l'unico autovalore per T è 0.
- (2) Mostra che T è diagonalizzabile se e solo se $T = 0$.

Esercizio 5.17. Considera una matrice a blocchi

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}.$$

in cui A, B e D sono matrici quadrate. Usando l'Esercizio 3.14, mostra che per ogni autovalore λ di A vale

$$m_g^A(\lambda) \leq m_g^B(\lambda) + m_g^D(\lambda)$$

dove indichiamo con $m_g^X(\lambda)$ la molteplicità geometrica di λ nella matrice X .

Esercizio 5.18. Deduci dall'esercizio precedente che se A è diagonalizzabile allora B e D sono entrambe diagonalizzabili. È vero l'opposto?

Esercizio 5.19. Sia $T: V \rightarrow V$ un endomorfismo e $W \subset V$ un sottospazio T -invariante. Usando l'esercizio precedente, mostra che se T è diagonalizzabile allora anche la restrizione

$$T|_W: W \rightarrow W$$

è diagonalizzabile.

Complementi

CAPITOLO 6

Forma di Jordan

In questo capitolo affronteremo il problema seguente: sappiamo che un endomorfismo è più facile da studiare se è diagonalizzabile; cosa possiamo fare però nel caso in cui non lo sia? Vedremo che potremo comunque scegliere una matrice relativamente semplice, che è “quasi” diagonale, eccetto che per alcuni valori sopra la diagonale principale: questo tipo di matrice particolare e utile è detto *matrice di Jordan*.

In tutto questo capitolo il campo base \mathbb{K} sarà sempre quello dei complessi $\mathbb{K} = \mathbb{C}$. In questo modo potremo sfruttare il teorema fondamentale dell'algebra – cioè il fatto che ogni polinomio di grado n abbia n radici contate con molteplicità.

6.1. Forma di Jordan

Una *matrice di Jordan* è un tipo di matrice “quasi” diagonale in cui compaiono solo alcuni valori non nulli sopra la diagonale principale. Mostreremo che qualsiasi matrice A , anche non diagonalizzabile, è simile ad un'unica matrice di Jordan, detta *forma di Jordan* di A .

La forma di Jordan caratterizza completamente la classe di similitudine di A e può essere usata per determinare se due matrici sono simili oppure no.

6.1.1. Blocco di Jordan. In tutto questo capitolo lavoriamo sempre con il campo $\mathbb{K} = \mathbb{C}$ dei numeri complessi. Un *blocco di Jordan* è una matrice quadrata $n \times n$ del tipo

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

Qui $\lambda \in \mathbb{C}$ è uno scalare complesso fissato qualsiasi. Un blocco di Jordan $B_{\lambda,n}$ è una matrice $n \times n$ triangolare superiore in cui tutte le caselle della diagonale principale hanno lo stesso numero $\lambda \in \mathbb{C}$ e tutte le caselle sulla diagonale immediatamente superiore hanno lo stesso numero 1. Tutte le altre diagonali

hanno solo zeri. Ad esempio queste matrici sono blocchi di Jordan:

$$(3), \quad \begin{pmatrix} i & 1 \\ 0 & i \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Notiamo che una matrice 1×1 è sempre un blocco di Jordan. Osserviamo anche che un blocco di Jordan $n \times n$ è triangolare e quindi i suoi autovalori sono sulla diagonale: ne segue che la matrice ha un solo autovalore λ . Calcoliamo le sue molteplicità algebrica e geometrica.

Proposizione 6.1.1. *Le molteplicità dell'autovalore λ sono:*

$$m_a(\lambda) = n, \quad m_g(\lambda) = 1.$$

In particolare se $n \geq 2$ un blocco di Jordan non è mai diagonalizzabile.

Dimostrazione. La molteplicità algebrica è chiara perché la matrice è triangolare superiore. Per la molteplicità geometrica, notiamo che

$$m_g(\lambda) = n - \text{rk} \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix} = n - (n - 1) = 1.$$

La matrice descritta ha effettivamente rango $n - 1$ perché, dopo aver tolto la prima colonna, le colonne rimanenti sono chiaramente indipendenti. \square

6.1.2. Matrice di Jordan. Una *matrice di Jordan* è una matrice a blocchi

$$J = \begin{pmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B_k \end{pmatrix}$$

dove B_1, \dots, B_k sono blocchi di Jordan di taglia arbitraria. Notiamo che tutti i blocchi che non stanno sulla diagonale principale sono nulli. Ad esempio, queste sono matrici di Jordan:

$$\begin{pmatrix} -i & 1 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1+i \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Le prime due matrici hanno due blocchi di Jordan, la terza ne ha tre, l'ultima ne ha uno solo (è lei stessa un blocco di Jordan). Qualsiasi matrice diagonale $n \times n$ è una matrice di Jordan con n blocchi di taglia 1×1 .

Notiamo un fatto semplice ma importante: una matrice di Jordan è sempre triangolare superiore e quindi i suoi autovalori sono tutti sulla diagonale.

Proposizione 6.1.2. *Se permutiamo i blocchi di una matrice di Jordan J , otteniamo un'altra matrice di Jordan J' simile a J .*

Dimostrazione. Interpretiamo J come $J = [L_J]_{\mathcal{C}}^{\mathcal{C}}$ con $L_J: \mathbb{C}^n \rightarrow \mathbb{C}^n$ e \mathcal{C} base canonica. Usando l'Esercizio 4.15 è facile vedere che $J' = [L_J]_{\mathcal{B}}^{\mathcal{B}}$ dove \mathcal{B} è ottenuta permutando opportunamente i vettori di \mathcal{C} . Quindi $J \sim J'$. \square

6.1.3. Teorema di Jordan. Enunciamo adesso il teorema più importante di questo capitolo.

Teorema 6.1.3. *Qualsiasi matrice $A \in M(n, \mathbb{C})$ è simile ad una matrice di Jordan J . La matrice di Jordan J è unica a meno di permutare i blocchi.*

Una dimostrazione di questo teorema sarà fornita nella Sezione 6.1.6. È importante notare come la matrice di Jordan J sia unica a meno di permutare i blocchi: per questo J viene chiamata la *forma canonica di Jordan* di A .

Corollario 6.1.4. *Due matrici A e A' sono simili \iff le loro matrici di Jordan J e J' hanno gli stessi blocchi.*

Esempio 6.1.5. Consideriamo le matrici seguenti:

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Quali di queste sono simili fra loro? Le prime quattro matrici sono matrici di Jordan. La seconda e la terza sono simili, perché hanno gli stessi blocchi, solo permutati in modo diverso. La prima, la seconda e la quarta sono però tutte non simili fra loro, perché hanno blocchi diversi. L'ultima matrice non è in forma di Jordan e quindi non possiamo dire immediatamente se sia simile ad una delle precedenti: c'è del lavoro da fare.

Il Teorema 6.1.3 è molto potente, ma osserviamo che enunciato in questo modo non fornisce un metodo per costruire J a partire da A . Esiste un algoritmo per costruire J da A che descriveremo alla fine di questa sezione, ma non è semplicissimo e per adesso ci limitiamo ad usare il teorema: nei casi più semplici questo sarà sufficiente.

Esempio 6.1.6. Due matrici simili A e B hanno lo stesso determinante, la stessa traccia, lo stesso polinomio caratteristico, gli stessi autovalori, e per ciascun autovalore hanno anche le stesse molteplicità algebrica e geometrica. Usando tutte queste proprietà possiamo spesso riconoscere due matrici che non sono simili.

Ad esempio, le matrici dell'Esempio 6.1.5 hanno tutte lo stesso polinomio caratteristico (e quindi stessa traccia, determinante e autovalori). Però le molteplicità geometriche dell'autovalore 2 non sono tutte uguali: facendo un conto vediamo che queste sono rispettivamente

$$3, \quad 2, \quad 2, \quad 1, \quad 1.$$

La quinta matrice quindi può essere simile solo alla quarta. Ma lo è veramente? Usando il Teorema 6.1.3 possiamo rispondere affermativamente nel modo seguente. Ci chiediamo: quale è la forma di Jordan dell'ultima matrice? Siccome ha solo l'autovalore 2, deve essere una delle precedenti: si verifica facilmente infatti che queste sono tutte le possibili matrici di Jordan 3×3 con un solo autovalore 2. Di queste, solo la quarta può essere simile a lei, quindi sono simili.

Esempio 6.1.7. Due matrici A e B che hanno lo stesso polinomio caratteristico e le stesse molteplicità algebriche e geometriche degli autovalori, sono sempre simili? La risposta è negativa: le matrici seguenti

$$\begin{pmatrix} \boxed{1} & \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & \boxed{1} \\ 0 & 0 & 0 & \boxed{1} \end{pmatrix}, \quad \begin{pmatrix} \boxed{1} & \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & \boxed{1} \end{pmatrix}$$

hanno un solo autovalore 1 ed in entrambi i casi la molteplicità geometrica di 1 è 2, come si vede facilmente con un conto. Però si decompongono in blocchi di Jordan diversi: quella di sinistra in due blocchi di ordine 2, quella di destra in due blocchi di ordine 3 e 1. Quindi per il Teorema 6.1.3 non sono simili.

6.1.4. Matrici nilpotenti. Una matrice $A \in M(n, \mathbb{C})$ è *nilpotente* se esiste un $k \geq 0$ per cui $A^k = 0$. Ovviamente, se questo accade, allora $A^h = 0$ anche per ogni $h \geq k$. L'*indice di nilpotenza* di A è il più piccolo numero intero $k \geq 1$ per cui $A^k = 0$.

Esempio 6.1.8. Una matrice A è nilpotente con indice di nilpotenza 1 se e solo se $A = 0$. Quindi solo la matrice nulla è nilpotente con indice 1.

Ad esempio, consideriamo il blocco di Jordan $B = B_{0,2}$. Otteniamo

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

Quindi B è nilpotente con indice di nilpotenza 2. Con $B = B_{0,3}$ otteniamo

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

Questo ci dice che B è nilpotente con indice di nilpotenza 3. Con la matrice $B = B_{0,4}$ otteniamo

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

e infine $B^4 = 0$, quindi B ha indice di nilpotenza 4. Più in generale, abbiamo capito che ogni volta che moltiplichiamo $B_{0,n}$ per se stesso spostiamo i numeri 1 sulla diagonale nella diagonale immediatamente superiore. In questo modo dimostriamo il fatto seguente.

Proposizione 6.1.9. *Il blocco $B_{0,n}$ è nilpotente con indice di nilpotenza n .*

Dimostrazione. Si dimostra facilmente per induzione su i che la matrice $B_{0,n}^k$ ha zeri ovunque tranne nelle caselle $i, i+k$ in cui c'è 1. \square

Esistono molte altre matrici nilpotenti oltre alle $B_{0,n}$. Ad esempio, una qualsiasi matrice simile a $B_{0,n}$ è anch'essa nilpotente:

Esercizio 6.1.10. Siano A e B due matrici simili. Se A è nilpotente con indice k , allora anche B è nilpotente con indice k .

Esercizio 6.1.11. Le matrici 2×2 nilpotenti sono precisamente le

$$\begin{pmatrix} a & b \\ c & -a \end{pmatrix}$$

con $a, b, c \in \mathbb{C}$ tali che $bc = -a^2$. Queste sono precisamente le matrici 2×2 con traccia e determinante entrambi nulli.

La semplice caratterizzazione con traccia e determinante nulli però funziona solo per le matrici 2×2 , non per le 3×3 !

Esercizio 6.1.12. Scrivi una matrice $A \in M(3, \mathbb{C})$ con $\text{tr}A = \det A = 0$ che non sia nilpotente.

Una matrice di Jordan J con solo l'autovalore zero è nilpotente. Infatti:

$$J = \begin{pmatrix} B_{0,m_1} & 0 & \dots & 0 \\ 0 & B_{0,m_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B_{0,m_k} \end{pmatrix} \implies J^i = \begin{pmatrix} B_{0,m_1}^i & 0 & \dots & 0 \\ 0 & B_{0,m_2}^i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B_{0,m_k}^i \end{pmatrix}$$

e otteniamo $J^i = 0$ se $i = \max\{m_1, \dots, m_k\}$. Usando la forma di Jordan, possiamo fornire una caratterizzazione completa delle matrici nilpotenti:

Proposizione 6.1.13. *Una matrice $A \in M(n, \mathbb{C})$ è nilpotente $\iff p_A(\lambda) = \lambda^n \iff A$ ha solo l'autovalore zero.*

Dimostrazione. Se A ha un autovalore $\lambda \neq 0$, allora esiste un $x \in \mathbb{C}^n$ non nullo tale che $Ax = \lambda x$ e allora $A^k x = \lambda^k x \neq 0$ per ogni k . Quindi $A^k \neq 0$ per ogni k e A non è nilpotente.

Se A ha solo l'autovalore zero, la sua forma di Jordan J è fatta di blocchi di tipo $B_{0,m}$ e quindi è nilpotente. Siccome $A \sim J$, anche A è nilpotente. \square

6.1.5. Un algoritmo per determinare J . Come possiamo determinare la matrice di Jordan J associata ad una data matrice complessa A ? In molti casi è sufficiente andare per tentativi, come abbiamo visto sopra: si elencano le possibili matrici di Jordan e si usa la molteplicità geometrica per eliminarne alcune, sperando di restare alla fine con una sola. A volte però questo metodo non funziona: esponiamo ora un algoritmo più complesso che funziona sempre.

Sia λ un autovalore di A . Per ciascun numero intero $j = 1, 2, \dots$ definiamo il j -esimo autospazio generalizzato

$$V_\lambda^j = \ker(A - \lambda I_n)^j$$

e la j -esima molteplicità geometrica generalizzata

$$m_j = m_g(\lambda)_j = \dim V_\lambda^j.$$

La molteplicità m_1 è l'usuale molteplicità geometrica di λ .

Proposizione 6.1.14. *Le molteplicità generalizzate formano una successione non decrescente:*

$$m_1 \leq m_2 \leq m_3 \leq \dots$$

Dimostrazione. Scriviamo $B = A - \lambda I_n$. Per ogni $j \geq 1$, se $B^j x = 0$ per qualche $x \in \mathbb{C}^n$, allora anche $B^{j+1}x = 0$. Quindi $\ker B^j \subset \ker B^{j+1}$. \square

Proposizione 6.1.15. *Due matrici simili A e B hanno le stesse molteplicità geometriche generalizzate per ciascun autovalore.*

Dimostrazione. Se $A = M^{-1}BM$, allora

$$(A - \lambda I)^j = (M^{-1}BM - \lambda M^{-1}M)^j = (M^{-1}(B - \lambda I)M)^j = M^{-1}(B - \lambda I)^j M.$$

Quindi le matrici $(A - \lambda I)^j$ e $(B - \lambda I)^j$ sono anch'esse simili; hanno lo stesso rango, e quindi anche la stessa dimensione del nucleo. \square

Sia J la matrice di Jordan associata ad A . Poiché J e A sono simili, hanno le stesse molteplicità geometriche generalizzate. Il punto fondamentale è che queste determinano univocamente J , come mostrato nella proposizione seguente. Poniamo $m_0 = 0$.

Proposizione 6.1.16. *Per ogni $i \geq 1$, il numero $m_i - m_{i-1}$ è il numero di blocchi di J con autovalore λ e di taglia $\geq i$. In particolare la molteplicità geometrica m_1 è il numero di blocchi con autovalore λ .*

Dimostrazione. Ricordiamo che $B_{\lambda, n}$ indica un blocco di Jordan di ordine n con autovalore λ . A meno di permutare i blocchi, possiamo mettere tutti i blocchi di J con autovalore λ all'inizio, e scrivere

$$J = \begin{pmatrix} B_{\lambda, n_1} & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & B_{\lambda, n_a} & 0 \\ 0 & \dots & 0 & J' \end{pmatrix}.$$

dove n_1, \dots, n_a sono gli ordini dei blocchi con autovalore λ . La matrice di Jordan J' contiene i blocchi con gli altri autovalori. Notiamo che

$$m_1 = \dim \ker(J - \lambda I) = \dim \ker \begin{pmatrix} B_{0,n_1} & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & B_{0,n_a} & 0 \\ 0 & \dots & 0 & J' - \lambda I \end{pmatrix}.$$

La matrice $J' - \lambda I$ è invertibile (perché è triangolare superiore con elementi tutti non nulli sulla diagonale), e m_1 è il numero di colonne nulle della matrice, perché le altre colonne non nulle sono chiaramente indipendenti. C'è una colonna nulla in ogni blocco $B_{0,n}$, quindi $m_1 = a$ è il numero di tali blocchi.

Calcoliamo ora m_2 e notiamo che

$$m_2 = \dim \ker(J - \lambda I)^2 = \dim \ker \begin{pmatrix} B_{0,n_1}^2 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & B_{0,n_a}^2 & 0 \\ 0 & \dots & 0 & (J' - \lambda I)^2 \end{pmatrix}.$$

La matrice $(J' - \lambda I)^2$ è sempre invertibile, e m_2 è sempre il numero di colonne nulle (perché le colonne non nulle sono indipendenti). Le colonne nulle sono le m_1 colonne di prima, più $m_2 - m_1$ nuove colonne, una per ogni blocco $B_{0,n}$ di taglia almeno due. Quindi $m_2 - m_1$ è proprio il numero di blocchi di taglia ≥ 2 . I passi successivi della dimostrazione sono analoghi. \square

Esempio 6.1.17. L'esempio numerico seguente dovrebbe chiarire meglio cosa succede ed aiutare anche a capire la dimostrazione appena conclusa. Consideriamo

$$J = \begin{pmatrix} \boxed{2} & \boxed{1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \boxed{2} & \boxed{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{2} & \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \boxed{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{2} & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{2} \end{pmatrix}.$$

Questa ha un blocco di ordine 3, due blocchi di ordine 2 e un blocco di ordine

1. Otteniamo

$$m_1 = \dim \ker(J - 2I) = \dim \ker \begin{pmatrix} \boxed{0} & \boxed{1} & \boxed{0} & 0 & 0 & 0 & 0 & 0 \\ \boxed{0} & \boxed{0} & \boxed{1} & 0 & 0 & 0 & 0 & 0 \\ \boxed{0} & \boxed{0} & \boxed{0} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{0} & \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{0} & \boxed{0} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{0} & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{0} & \boxed{0} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{0} \end{pmatrix} = 4$$

perché sono comparse 4 colonne di zeri, una per ogni blocco: infatti la molteplicità geometrica m_1 è il numero di blocchi. Continuiamo e troviamo

$$m_2 = \dim \ker(J - 2I)^2 = \dim \ker \begin{pmatrix} \boxed{0} & \boxed{0} & \boxed{1} & 0 & 0 & 0 & 0 & 0 \\ \boxed{0} & \boxed{0} & \boxed{0} & 0 & 0 & 0 & 0 & 0 \\ \boxed{0} & \boxed{0} & \boxed{0} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{0} & \boxed{0} & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{0} & \boxed{0} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{0} & \boxed{0} & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{0} & \boxed{0} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{0} \end{pmatrix} = 7.$$

Al secondo passaggio, si sono manifestate 3 nuove colonne di zeri: una per ogni blocco di taglia ≥ 2 (i blocchi di taglia 1 erano già spariti). I blocchi di taglia ≥ 2 sono effettivamente $m_2 - m_1 = 7 - 4 = 3$.

Infine $(J - 2I)^3 = 0$ ci dice che $m_3 = 8$ e infatti i blocchi di taglia ≥ 3 sono esattamente $m_3 - m_2 = 8 - 7 = 1$.

Dalla Proposizione 6.1.16 possiamo ricavare alcune informazioni più precise sulla successione non decrescente $m_1 \leq m_2 \leq \dots$ delle molteplicità geometriche generalizzate. Sia come sopra J una matrice di Jordan e λ un autovalore per J .

Corollario 6.1.18. *Sia h la massima taglia di un blocco di Jordan in J con autovalore λ . La successione delle molteplicità geometriche generalizzate di λ è strettamente crescente fino a $m_h = m_a(\lambda)$ e poi si stabilizza:*

$$m_1 = m_g(\lambda) < m_2 < \dots < m_h = m_a(\lambda) = m_{h+1} = \dots$$

Notiamo infatti che $m_g(\lambda) = m_a(\lambda)$ precisamente quando $h = 1$, cioè quando tutti i blocchi di Jordan relativi a λ hanno taglia 1.

Esempio 6.1.19. Usiamo l'algoritmo per determinare la forma di Jordan J della matrice

$$A = \begin{pmatrix} 2 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 1 & 5 & 2 & -1 \\ 0 & -4 & 0 & 4 \end{pmatrix}.$$

Il polinomio caratteristico è $p_A(\lambda) = (\lambda - 2)^4$. Per determinare la forma di Jordan dobbiamo calcolare le molteplicità geometriche generalizzate dell'autovalore 2. Iniziamo:

$$m_1 = \dim \ker(A - 2I) = \dim \ker \begin{pmatrix} 0 & 2 & 0 & -1 \\ 0 & -2 & 0 & 1 \\ 1 & 5 & 0 & -1 \\ 0 & -4 & 0 & 2 \end{pmatrix} = 2.$$

Scopriamo che la molteplicità geometrica è 2, quindi il numero di blocchi di J è 2. Questo però non è sufficiente a determinare J , perché ci sono due casi possibili, cioè

$$\begin{pmatrix} \boxed{2} & \boxed{1} & 0 & 0 \\ 0 & \boxed{2} & 0 & 0 \\ 0 & 0 & \boxed{2} & \boxed{1} \\ 0 & 0 & 0 & \boxed{2} \end{pmatrix} \quad \text{oppure} \quad \begin{pmatrix} \boxed{2} & \boxed{1} & 0 & 0 \\ 0 & \boxed{2} & \boxed{1} & 0 \\ 0 & 0 & \boxed{2} & 0 \\ 0 & 0 & 0 & \boxed{2} \end{pmatrix}.$$

Allora continuiamo a calcolare le molteplicità geometriche generalizzate:

$$m_2 = \dim \ker(A - 2I)^2 = \dim \ker \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -4 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 3.$$

Quindi il numero di blocchi di J di taglia ≥ 2 è $m_2 - m_1 = 3 - 2 = 1$. Questo esclude il primo dei due casi possibili per J elencati sopra, quindi troviamo

$$J = \begin{pmatrix} \boxed{2} & \boxed{1} & 0 & 0 \\ 0 & \boxed{2} & \boxed{1} & 0 \\ 0 & 0 & \boxed{2} & 0 \\ 0 & 0 & 0 & \boxed{2} \end{pmatrix}.$$

6.1.6. Dimostrazione del Teorema di Jordan. In questa sezione forniamo una dimostrazione del Teorema di Jordan. Ne approfittiamo per studiare il problema dal punto di vista degli endomorfismi.

Sia V uno spazio vettoriale complesso di dimensione n e $T: V \rightarrow V$ un endomorfismo. Una *base di Jordan* per T è una base \mathcal{B} tale che la matrice associata $J = [T]_{\mathcal{B}}^{\mathcal{B}}$ sia una matrice di Jordan.

Come è fatta una base di Jordan \mathcal{B} ? A ciascun blocco di Jordan $B_{\lambda,k}$ di J corrispondono k vettori v_1, \dots, v_k di \mathcal{B} con la proprietà che

$$T(v_1) = \lambda v_1, \quad T(v_2) = \lambda v_2 + v_1, \quad \dots \quad T(v_k) = \lambda v_k + v_{k-1}.$$

È più utile scrivere le uguaglianze in questo modo:

$$(T - \lambda \text{id})(v_1) = 0, \quad (T - \lambda \text{id})(v_2) = v_1, \quad \dots \quad (T - \lambda \text{id})(v_k) = v_{k-1}.$$

Chiamiamo k vettori v_1, \dots, v_k fatti a questo modo una *catena* di vettori. Possiamo visualizzarla così:

$$0 \xleftarrow{T - \lambda \text{id}} v_1 \xleftarrow{T - \lambda \text{id}} v_2 \xleftarrow{T - \lambda \text{id}} \dots \xleftarrow{T - \lambda \text{id}} v_k.$$

Una base di Jordan per T è precisamente una base che è unione disgiunta di alcune catene di vettori (una per ogni blocco di Jordan).

Dimostriamo adesso il teorema seguente.

Teorema 6.1.20. *Ogni endomorfismo $T: V \rightarrow V$ ha una base di Jordan.*

Dimostrazione. Procediamo per induzione su $n = \dim V$. Se $n = 1$, qualsiasi generatore di V è una base di Jordan. Consideriamo il caso generico n dando per buono i casi $\leq n - 1$.

Poiché V è uno spazio vettoriale complesso, esiste sempre almeno un autovalore λ per T . Quindi l'autospazio $V_\lambda = \ker(T - \lambda \text{id})$ ha una certa dimensione $m_g(\lambda) \geq 1$. Per il Teorema della dimensione, l'immagine $\text{Im}(T - \lambda \text{id})$ ha dimensione $n - m_g(\lambda) \leq n - 1$. Ricordiamo che

$$\text{Im}(T - \lambda \text{id}) = \{T(v) - \lambda v \mid v \in V\}.$$

Il sottospazio $U = \text{Im}(T - \lambda \text{id})$ è T -invariante, infatti

$$T(T(v) - \lambda v) = T(T(v)) - \lambda T(v) = T(w) - \lambda w$$

con $w = T(v)$. Per l'ipotesi induttiva esiste una base di Jordan per l'endomorfismo ristretto $T|_U$. Questa base è unione di h catene

$$v_1^1, \dots, v_{k_1}^1, \quad \dots, \quad v_1^h, \dots, v_{k_h}^h.$$

Poiché questi vettori stanno tutti in $U = \text{Im}(T - \lambda \text{id})$, per ciascun $i = 1, \dots, h$ esiste un $v_{k_i+1}^i \in V$ tale che

$$(T - \lambda \text{id})(v_{k_i+1}^i) = v_{k_i}^i.$$

Aggiungiamo $v_{k_i+1}^i$ alla fine della i -esima catena, per ogni i . In questo modo abbiamo aggiunto h nuovi vettori alla base. Oltre a questi, ne aggiungiamo altri. Consideriamo i primi vettori delle catene v_1^1, \dots, v_1^h . Questi sono vettori indipendenti contenuti in $\ker(T - \lambda \text{id})$. Completiamoli a base di $\ker(T - \lambda \text{id})$ aggiungendo altri vettori $v_1^{h+1}, \dots, v_1^{h+s}$. Ciascuno di questi è una nuova catena formata da un solo vettore.

A questo punto dobbiamo solo dimostrare che i vettori così ottenuti

$$v_1^1, \dots, v_{k_1+1}^1, \quad \dots, \quad v_1^h, \dots, v_{k_h+1}^h, \quad v_1^{h+1}, \quad \dots, \quad v_1^{h+s}$$

formano una base di V . Mostriamo che sono linearmente indipendenti. Supponiamo di avere una combinazione lineare che li annulla:

$$\sum_{i=1}^h (\alpha_1^i v_1^i + \dots + \alpha_{k_i+1}^i v_{k_i+1}^i) + \alpha_1^{h+1} v_1^{h+1} + \dots + \alpha_1^{h+s} v_1^{h+s} = 0.$$

Applicando $T - \lambda \text{id}$ otteniamo

$$\sum_{i=1}^h (\alpha_2^i v_1^i + \dots + \alpha_{k_i+1}^i v_{k_i}^i) = 0.$$

Questi vettori sono una base di U e quindi $\alpha_j^j = 0 \forall j \geq 2$. Rimaniamo con

$$\alpha_1^1 v_1^1 + \dots + \alpha_1^{h+s} v_1^{h+s} = 0.$$

Questi vettori sono una base di $\ker(T - \lambda \text{id})$ e quindi anche questi coefficienti sono tutti zero.

Per mostrare che i vettori costruiti generano V , mostriamo che sono esattamente $n = \dim V$. Dimostriamo cioè che $k_1 + \dots + k_h + h + s = n$. Sappiamo che $\dim \text{Im}(T - \lambda \text{id}) = k_1 + \dots + k_h$ e $\dim \ker(T - \lambda \text{id}) = h + s$, quindi questo segue dal Teorema della dimensione. \square

Il Teorema 6.1.3 segue dal teorema appena dimostrato. L'esistenza di una forma di Jordan per qualsiasi matrice $A \in M(n, \mathbb{C})$ si ottiene applicando il Teorema 6.1.20 all'endomorfismo $L_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$. L'unicità a meno di permutare i blocchi segue dal fatto che il numero di blocchi di ciascuna taglia è determinato dalle molteplicità geometriche generalizzate, come avevamo già visto nella Proposizione 6.1.16.

6.1.7. Autospazi generalizzati massimali. Sia $T: V \rightarrow V$ un endomorfismo di uno spazio vettoriale complesso V di dimensione finita. Per ogni autovalore λ , sappiamo dal Corollario 6.1.18 che gli autospazi generalizzati $V_\lambda^j = \ker(T - \lambda \text{id})^j$ formano una successione del tipo

$$V_\lambda^1 \subsetneq \dots \subsetneq V_\lambda^h = V_\lambda^{h+1} = \dots$$

con $\dim V_\lambda^h = m_a(\lambda)$. Chiamiamo $V_\lambda^{\text{gen}} = V_\lambda^h$ l'*autospazio generalizzato massimale* di λ . Dalla forma di Jordan ricaviamo il fatto seguente:

Corollario 6.1.21. *Siano $\lambda_1, \dots, \lambda_k$ gli autovalori di T . Vale:*

$$V = V_{\lambda_1}^{\text{gen}} \oplus \dots \oplus V_{\lambda_k}^{\text{gen}}.$$

Ciascun $V_{\lambda_i}^{\text{gen}}$ è T -invariante. La restrizione $T|_{V_{\lambda_i}^{\text{gen}}}$ ha solo l'autovalore λ_i .

Ciascun endomorfismo su uno spazio vettoriale complesso V di dimensione finita può essere "spezzato" in una somma diretta di endomorfismi sui suoi autospazi generalizzati massimali.

6.1.8. Base di Jordan. Come possiamo trovare una base di Jordan \mathcal{B} per un dato endomorfismo $T: V \rightarrow V$? L'algoritmo non è semplicissimo: si tratta di seguire la dimostrazione del Teorema 6.1.20.

Come prima cosa determiniamo gli autovalori λ_i di T e gli autospazi generalizzati massimali $V_{\lambda_i}^{\text{gen}}$. Per il Corollario 6.1.21 l'endomorfismo T si spezza in endomorfismi su questi autospazi generalizzati che hanno ciascun un solo autovalore. Lavoriamo separatamente su ciascuna restrizione.

Ci siamo ricondotti al caso in cui T abbia un solo autovalore λ . Determiniamo la successione di sottospazi T -invarianti:

$$V \supsetneq \text{Im}(T - \lambda \text{id}) \supsetneq \dots \supsetneq \text{Im}(T - \lambda \text{id})^h = \{0\}$$

Costruiamo iterativamente una successione

$$\mathcal{B}_0 \supseteq \mathcal{B}_1 \supseteq \cdots \supseteq \mathcal{B}_{h-1} \supseteq \mathcal{B}_h = \emptyset$$

dove \mathcal{B}_j è una base di Jordan della restrizione di T al sottospazio $\text{Im}(T - \lambda \text{id})^j$. La base \mathcal{B}_0 è quindi una base di Jordan per T .

La fabbricazione delle basi funziona iterativamente da destra a sinistra nel modo seguente. Costruiamo \mathcal{B}_{j-1} aggiungendo i seguenti vettori a \mathcal{B}_j :

- (1) Per ogni vettore $v \in \mathcal{B}_j \setminus \mathcal{B}_{j+1}$, prendiamo un $w \in \text{Im}(T - \lambda \text{id})^{j-1}$ tale che $(T - \lambda \text{id})(w) = v$ e lo aggiungiamo.
- (2) Se i vettori in \mathcal{B}_{j-1} trovati finora generano $\text{Im}(T - \lambda \text{id})^{j-1}$ siamo a posto. Altrimenti li completiamo a base di $\text{Im}(T - \lambda \text{id})^{j-1}$ usando solo vettori di $\ker(T - \lambda \text{id})$. È sempre possibile fare ciò grazie al teorema della dimensione.

Questo è lo schema:

$$\begin{array}{cccccccc}
 0 & \xleftarrow{T-\lambda \text{id}} & v_1^1 & \xleftarrow{T-\lambda \text{id}} & v_2^1 & \xleftarrow{T-\lambda \text{id}} & v_3^1 & \xleftarrow{T-\lambda \text{id}} \cdots \\
 \vdots & & \vdots & & \vdots & & \vdots & \\
 0 & \xleftarrow{T-\lambda \text{id}} & v_1^a & \xleftarrow{T-\lambda \text{id}} & v_2^a & \xleftarrow{T-\lambda \text{id}} & v_3^a & \xleftarrow{T-\lambda \text{id}} \cdots \\
 & & 0 & \xleftarrow{T-\lambda \text{id}} & v_1^{a+1} & \xleftarrow{T-\lambda \text{id}} & v_2^{a+1} & \xleftarrow{T-\lambda \text{id}} \cdots \\
 & & \vdots & & \vdots & & \vdots & \\
 & & 0 & \xleftarrow{T-\lambda \text{id}} & v_1^b & \xleftarrow{T-\lambda \text{id}} & v_2^b & \xleftarrow{T-\lambda \text{id}} \cdots \\
 & & & & 0 & \xleftarrow{T-\lambda \text{id}} & v_1^{b+1} & \xleftarrow{T-\lambda \text{id}} \cdots \\
 & & & & \vdots & & \vdots & \\
 & & & & 0 & \xleftarrow{T-\lambda \text{id}} & v_1^c & \xleftarrow{T-\lambda \text{id}} \cdots \\
 & & & & & & 0 & \xleftarrow{T-\lambda \text{id}} \cdots \\
 & & & & & & \vdots & \ddots \\
 & & & & & & \vdots & \\
 & & \downarrow & & \downarrow & & \downarrow & \\
 & & \mathcal{B}^{h-1} & & \mathcal{B}^{h-2} \setminus \mathcal{B}^{h-1} & & \mathcal{B}^{h-3} \setminus \mathcal{B}^{h-2} & \cdots
 \end{array}$$

Esempio 6.1.22. Cerchiamo una base di Jordan per $L_A: \mathbb{C}^4 \rightarrow \mathbb{C}^4$ dove A è la matrice dell'Esempio 6.1.19:

$$A = \begin{pmatrix} 2 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 1 & 5 & 2 & -1 \\ 0 & -4 & 0 & 4 \end{pmatrix}.$$

La matrice ha come unico autovalore 2. Notiamo che

$$A - 2I = \begin{pmatrix} 0 & 2 & 0 & -1 \\ 0 & -2 & 0 & 1 \\ 1 & 5 & 0 & -1 \\ 0 & -4 & 0 & 2 \end{pmatrix}, \quad (A - 2I)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -4 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

e $(A - 2I)^3 = 0$. Abbiamo

$$\mathbb{C}^4 \supseteq \text{Im}(L_{A-2I}) \supseteq \text{Im}(L_{A-2I})^2 \supseteq \text{Im}(L_{A-2I})^3 = \{0\}.$$

e costruiamo partendo da destra la successione di basi:

$$\mathcal{B}_0 \supseteq \mathcal{B}_1 \supseteq \mathcal{B}_2 \supseteq \mathcal{B}_3 = \emptyset.$$

Possiamo prendere ad esempio:

$$\mathcal{B}_2 = \left\{ v_1^1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}, \quad \mathcal{B}_1 = \left\{ v_1^1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, v_2^1 = \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 2 \end{pmatrix} \right\}.$$

Notiamo che effettivamente $(A - 2I)(v_1^1) = 0$ e $(A - 2I)(v_2^1) = v_1^1$. Infine

$$\mathcal{B}_0 = \left\{ v_1^1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, v_2^1 = \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 2 \end{pmatrix}, v_3^1 = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, v_1^2 = \begin{pmatrix} -3 \\ 1 \\ 0 \\ 2 \end{pmatrix} \right\}.$$

Effettivamente $(A - 2I)(v_3^1) = v_2^1$ e $(A - 2I)(v_1^2) = 0$. Questa è una base di Jordan per L_A e quindi

$$[L_A]_{\mathcal{B}_0}^{\mathcal{B}_0} = \begin{pmatrix} \boxed{2} & 1 & 0 & 0 \\ 0 & \boxed{2} & 1 & 0 \\ 0 & 0 & \boxed{2} & 0 \\ 0 & 0 & 0 & \boxed{2} \end{pmatrix}.$$

6.2. Teorema di Cayley – Hamilton

Le matrici quadrate possono essere sommate e moltiplicate fra loro. Se A è una matrice quadrata, ha senso scrivere espressioni del tipo $A^4 - A^3 - 2A$. In altre parole, ha senso prendere un polinomio $x^4 - x^3 - 2x$ e quindi sostituire x con una matrice A . Come per le radici di un polinomio, ci possiamo chiedere quando $A^4 - A^3 - 2A = 0$.

In questa sezione studiamo come i polinomi si combinano con le matrici.

6.2.1. Polinomi di matrici. Prendiamo una matrice $A \in M(n, \mathbb{C})$ e un polinomio $p(x) \in \mathbb{C}[x]$ di grado k , entrambi a coefficienti complessi. Abbiamo

$$p(x) = a_k x^k + \dots + a_1 x + a_0.$$

Definiamo adesso una nuova matrice $p(A)$ nel modo seguente:

$$p(A) = a_k A^k + \dots + a_1 A + a_0 I_n.$$

Diciamo che il polinomio $p(x)$ *annulla* la matrice A se

$$p(A) = 0.$$

In questa equazione “0” è la matrice $n \times n$ nulla. Vogliamo studiare come sono fatti i polinomi che annullano una matrice data A . Ovviamente, fra questi

c'è sempre il *polinomio banale*, cioè la costante 0, che annulla qualsiasi cosa. Prima domanda: c'è dell'altro? La risposta è affermativa:

Proposizione 6.2.1. *Esistono sempre polinomi non banali che annullano A.*

Dimostrazione. Consideriamo le matrici

$$I_n, A, A^2, \dots, A^{n^2}.$$

Queste sono $n^2 + 1$ elementi nello spazio $M(n, \mathbb{C})$ di tutte le matrici $n \times n$. Siccome questo spazio ha dimensione n^2 , questi $n^2 + 1$ elementi non possono essere indipendenti: c'è quindi una combinazione non banale che li annulla

$$a_0 I + a_1 A + a_2 A^2 + \dots + a_{n^2} A^{n^2}$$

con $a_i \in \mathbb{C}$. In altre parole, se definiamo

$$p(x) = a_{n^2} x^{n^2} + \dots + a_1 x + a_0$$

otteniamo $p(A) = 0$. □

La dimostrazione fornisce un polinomio $p(x)$ non banale di grado al massimo n^2 che annulli A . Vogliamo adesso migliorare questo risultato, mostrando che il polinomio caratteristico di A , che ha grado solo n , annulla A . Per ottenere questo profondo risultato noto come *Teorema di Cayley – Hamilton* abbiamo bisogno di alcuni fatti preliminari.

6.2.2. Matrici a blocchi. In questo capitolo, una *matrice a blocchi* è una matrice del tipo

$$A = \begin{pmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B_k \end{pmatrix}$$

dove B_1, \dots, B_k sono matrici quadrate, di taglia arbitraria. A parte i blocchi B_1, \dots, B_k che stanno sulla diagonale, tutti gli altri blocchi della matrice sono nulli. Un esempio importante di matrice a blocchi è ovviamente una matrice di Jordan.

Proposizione 6.2.2. *Se A è una matrice a blocchi e p(x) è un polinomio, allora otteniamo*

$$p(A) = \begin{pmatrix} p(B_1) & 0 & \dots & 0 \\ 0 & p(B_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p(B_k) \end{pmatrix}.$$

In altre parole, per applicare un polinomio p su A basta farlo su ogni blocco.

Dimostrazione. Sia

$$p(x) = a_mx^m + \dots + a_1x + a_0.$$

Otteniamo

$$p(A) = a_mA^m + \dots + a_1A + a_0I_n$$

che può essere scritto come

$$a_m \begin{pmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B_k \end{pmatrix}^m + \dots + a_1 \begin{pmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B_k \end{pmatrix} + a_0 \begin{pmatrix} I & 0 & \dots & 0 \\ 0 & I & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & I \end{pmatrix}$$

dove le I nell'ultima matrice sono tutte matrici identità della taglia opportuna.

Questo è equivalente a

$$\begin{pmatrix} a_mB_1^m & 0 & \dots & 0 \\ 0 & a_mB_2^m & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_mB_k^m \end{pmatrix} + \dots + \begin{pmatrix} a_1B_1 & 0 & \dots & 0 \\ 0 & a_1B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_1B_k \end{pmatrix} + \begin{pmatrix} a_0I & 0 & \dots & 0 \\ 0 & a_0I & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_0I \end{pmatrix}$$

e sommando le matrici otteniamo

$$\begin{pmatrix} p(B_1) & 0 & \dots & 0 \\ 0 & p(B_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p(B_k) \end{pmatrix}.$$

La dimostrazione è completa. \square

6.2.3. Matrici simili. Come sappiamo, due matrici simili hanno molte cose in comune. La proposizione seguente conferma questo fatto generale.

Proposizione 6.2.3. *Se A e B sono simili e $p(x)$ è un polinomio, allora*

$$p(A) = 0 \iff p(B) = 0.$$

Dimostrazione. Sia

$$p(x) = a_mx^m + \dots + a_1x + a_0.$$

Per ipotesi $A = MBM^{-1}$ per qualche matrice invertibile M . Ricordiamo che

$$A^k = (MBM^{-1})^k = MB^kM^{-1}.$$

Otteniamo

$$\begin{aligned} p(A) &= a_m(MBM^{-1})^m + \dots + a_1MBM^{-1} + a_0I \\ &= Ma_mB^mM^{-1} + \dots + Ma_1BM^{-1} + Ma_0IM^{-1} \\ &= Mp(B)M^{-1}. \end{aligned}$$

Nell'ultima uguaglianza abbiamo raccolto M a sinistra e M^{-1} a destra. Segue che $p(B) = 0 \implies p(A) = 0$. Invertendo i ruoli di A e B si ottiene anche la freccia opposta. \square

Abbiamo scoperto che due matrici simili sono annullate precisamente dagli stessi polinomi.

6.2.4. Teorema di Cayley – Hamilton. Abbiamo infine tutti gli strumenti per enunciare e dimostrare il seguente teorema.

Teorema 6.2.4 (Cayley – Hamilton). *Il polinomio caratteristico p_A di A annulla A . In altre parole, vale*

$$p_A(A) = 0.$$

Dimostrazione. Consideriamo prima il caso in cui $A = B_{\lambda_0, n}$ sia un blocco di Jordan con autovalore λ_0 . Otteniamo $p_A = (\lambda_0 - \lambda)^n$. Quindi

$$p_A(A) = (\lambda_0 I_n - A)^n = \pm(A - \lambda_0 I_n)^n = \pm B_{0, n}^n = 0.$$

Ricordiamo infatti che $B_{\lambda_0, n} - \lambda_0 I_n = B_{0, n}$ è nilpotente di indice n . Il teorema in questo caso è dimostrato.

Passiamo ora al caso in cui $A = J$ sia una matrice di Jordan, con blocchi di Jordan B_1, \dots, B_k . Per il Corollario 5.1.35 abbiamo

$$p_J(x) = p_{B_1}(x) \cdots p_{B_k}(x).$$

Per il caso precedente, sappiamo che $p_{B_i}(B_i) = 0$ e quindi anche $p_J(B_i) = 0$, perché $p_{B_i}(x)$ divide $p_J(x)$. A questo punto la Proposizione 6.2.2 ci dà

$$p_J(J) = \begin{pmatrix} p_J(B_1) & 0 & \dots & 0 \\ 0 & p_J(B_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p_J(B_k) \end{pmatrix} = 0.$$

Il teorema è quindi verificato anche in questo caso.

Per il Teorema 6.1.3 una generica matrice A è sempre simile ad una matrice di Jordan J . Queste hanno lo stesso polinomio caratteristico $p_A(x) = p_J(x)$, e se $p_J(J) = p_A(J) = 0$ allora anche $p_A(A) = 0$ per la Proposizione 6.2.3. La dimostrazione è conclusa. \square

Esempio 6.2.5. Esaminiamo la matrice

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

Sappiamo che $p_A(\lambda) = \lambda^2 + 1$. Per il Teorema di Cayley – Hamilton abbiamo

$$p_A(A) = A^2 + I = 0.$$

Effettivamente si verifica che $A^2 = -I$.

Esempio 6.2.6. Esaminiamo la matrice

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

Sappiamo che $p_A(\lambda) = \lambda^2 - \lambda$. Per il Teorema di Cayley – Hamilton abbiamo

$$p_A(A) = A^2 - A = 0.$$

Effettivamente si verifica che $A^2 = A$.

Esempio 6.2.7. Più in generale, il polinomio caratteristico di una matrice $A \in M(2, \mathbb{C})$ qualsiasi è $\lambda^2 - \text{tr}A \lambda + \det A$. Quindi otteniamo sempre

$$A^2 = \text{tr}A \cdot A - \det A \cdot I_2.$$

6.3. Polinomio minimo

Il Teorema di Cayley – Hamilton dice che il polinomio caratteristico p_A annulla la matrice A . Se A è una matrice $n \times n$, allora p_A ha grado n . Ci poniamo adesso la domanda seguente: esistono polinomi di grado minore di n che annullano A ? Questa domanda ci porta alla definizione del *polinomio minimo* di una matrice quadrata.

6.3.1. Polinomio minimo. Ricordiamo che un polinomio $p(x)$ è *monico* se il termine di grado più alto di $p(x)$ ha coefficiente 1. In altre parole, un polinomio p di grado k è monico se è del tipo

$$p(x) = x^k + a_{k-1}x^{k-1} + \dots + a_1x + a_0.$$

Notiamo che qualsiasi polinomio non banale può essere trasformato in un polinomio monico dividendolo per il coefficiente a_k del termine di grado più alto. Sia ora $A \in M(n, \mathbb{C})$ una matrice quadrata a coefficienti complessi.

Definizione 6.3.1. Il *polinomio minimo* di A è il polinomio monico $m_A(x)$ di grado minore che annulla A .

Proposizione 6.3.2. *La definizione è ben posta.*

Dimostrazione. Nella definizione abbiamo dato per scontato che ci sia un unico polinomio $m_A(x)$ con quelle proprietà, e adesso dobbiamo dimostrarlo. Supponiamo di avere due polinomi $m_1(x)$ e $m_2(x)$ monici distinti, entrambi di grado minimo k che annullano A . Siccome sono entrambi monici, la differenza $q(x) = m_1(x) - m_2(x)$ è un polinomio non banale di grado strettamente minore di k . Otteniamo

$$q(A) = m_1(A) - m_2(A) = 0 - 0 = 0.$$

Quindi $q(x)$ è un polinomio non banale di grado $< k$ che annulla A , ma questo è assurdo. \square

Esempio 6.3.3. Se $A = 0$ è la matrice nulla, allora il suo polinomio minimo è $m_A(x) = x$. Infatti abbiamo

$$m_A(A) = A = 0$$

e quindi $m_A(x)$ annulla A . Questo è sicuramente il polinomio monico di grado minimo fra quelli che annullano A perché ha grado uno.

Se $A = I_n$ allora il suo polinomio minimo è $m_A(x) = x - 1$. Infatti abbiamo

$$m_A(A) = A - I_n = I_n - I_n = 0$$

e concludiamo come nel caso precedente. Più in generale, le uniche matrici A aventi polinomio minimo di grado uno, cioè con $m_A(x) = x - c$ per qualche $c \in \mathbb{C}$, sono le matrici del tipo $A = cI_n$.

6.3.2. Il polinomio minimo divide tutti gli altri. Sia $A \in M(n, \mathbb{C})$ una matrice quadrata $n \times n$ a coefficienti complessi. Mostriamo un'importante proprietà algebrica del polinomio minimo.

Proposizione 6.3.4. *Il polinomio minimo $m_A(x)$ divide qualsiasi polinomio in $\mathbb{C}[x]$ che annulla la matrice A .*

Dimostrazione. Sia $a(x)$ un polinomio che annulla A . Dividendo $a(x)$ per $m_A(x)$ otteniamo

$$a(x) = m_A(x)q(x) + r(x)$$

per qualche resto $r(x)$ di grado minore di quello di $m_A(x)$. Ne segue che

$$r(A) = a(A) - m_A(A)q(A) = 0 - 0q(A) = 0$$

e allora anche $r(x)$ annulla A . Poiché $r(x)$ ha grado minore di $m_A(x)$, il polinomio $r(x)$ deve essere nullo. Quindi $m_A(x)$ divide $a(x)$. \square

Corollario 6.3.5. *Il polinomio minimo $m_A(x)$ divide sempre il polinomio caratteristico $p_A(x)$.*

Dimostrazione. Per il Teorema di Cayley – Hamilton, il polinomio caratteristico $p_A(x)$ annulla A . \square

Esempio 6.3.6. Sapendo che $m_A(x)$ divide $p_A(x)$, è possibile determinare $m_A(x)$ per tentativi. Ad esempio, consideriamo

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

Vediamo che il polinomio caratteristico è $p_A(x) = (x - 1)^2$. Siccome $m_A(x)$ divide $p_A(x)$, il polinomio $m_A(x)$ è uno dei due polinomi seguenti:

$$x - 1, \quad (x - 1)^2.$$

Procediamo per tentativi, iniziando con il primo: se $q(x) = x - 1$, otteniamo $q(A) = A - I_2 \neq 0$. Quindi il primo non va bene. Allora $m_A(x) = (x - 1)^2$ è il secondo, per esclusione. La lettrice può verificare che infatti $(A - I_2)^2 = 0$.

Esempio 6.3.7. Calcoliamo il polinomio minimo della matrice

$$A = \begin{pmatrix} i & 1 & 0 \\ 0 & i & 1 \\ 0 & 0 & i \end{pmatrix}.$$

Troviamo $p_A(x) = -(x - i)^3$. Quindi il polinomio minimo è uno dei seguenti:

$$x - i, \quad (x - i)^2, \quad (x - i)^3.$$

Se $q(x) = x - i$, otteniamo $q(A) = A - iI_3 \neq 0$, quindi il primo non va bene.

Se $q(x) = (x - i)^2$, si verifica che $q(A) = (A - iI_3)^2 \neq 0$, quindi neanche il secondo va bene. Quindi $m_A(x) = (x - i)^3$.

6.3.3. Blocchi di Jordan. Calcoliamo il polinomio minimo di un blocco di Jordan. Ricordiamo che $B_{\lambda,n}$ è il blocco di Jordan $n \times n$ con autovalore λ .

Proposizione 6.3.8. *Il polinomio minimo di $B = B_{\lambda_0,n}$ è uguale al suo polinomio caratteristico:*

$$m_B(\lambda) = p_B(\lambda) = (\lambda - \lambda_0)^n.$$

Dimostrazione. Sappiamo che $m_B(\lambda)$ divide $p_B(\lambda)$ e quindi $m_B(\lambda) = (\lambda - \lambda_0)^k$ per qualche $1 \leq k \leq n$. Dobbiamo determinare il k giusto. Se prendiamo $q(\lambda) = (\lambda - \lambda_0)^k$, otteniamo

$$q(B) = (B - \lambda_0 I_n)^k = B_{0,n}^k$$

dove $B_{0,n}$ è la matrice nilpotente considerata nella Proposizione 6.1.9, che sappiamo già avere ordine di nilpotenza n . Quindi $q(B) = 0$ solo se $k = n$ e ne ricaviamo che $m_A(\lambda) = p_A(\lambda) = (\lambda - \lambda_0)^n$. \square

6.3.4. Matrici a blocchi. Consideriamo una matrice a blocchi

$$A = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}.$$

Per i polinomi caratteristici, sappiamo dal Corollario 5.1.35 che

$$p_A(x) = p_B(x)p_C(x).$$

Il polinomio minimo si comporta diversamente.

Definizione 6.3.9. Dati due polinomi monici $a(x)$ e $b(x)$, il *minimo comune multiplo* di $a(x)$ e $b(x)$ è il più piccolo polinomio monico $m(x)$ che è diviso sia da $a(x)$ che da $b(x)$.

Esercizio 6.3.10. La definizione di minimo comune multiplo è ben posta.

Esempio 6.3.11. Il minimo comune multiplo fra $(x - 1)(x - 2)$ e $(x - 1)^2$ è $(x - 1)(x - 2)^2$.

Esercizio 6.3.12. Se A è una matrice a blocchi come sopra, allora $m_A(x)$ è il minimo comune multiplo fra $m_B(x)$ e $m_C(x)$.

6.3.5. Da Jordan a polinomio minimo. Mostriamo adesso che, conoscendo già la forma di Jordan, è molto facile determinare il polinomio minimo.

Sia come sempre $A \in M(n, \mathbb{C})$ una matrice quadrata. Siano $\lambda_1, \dots, \lambda_k$ gli autovalori di A . Per il Corollario 1.4.10 il polinomio caratteristico di A è

$$p_A(\lambda) = \pm(\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_k)^{m_k}$$

dove $m_i = m_a(\lambda_i)$ è la molteplicità algebrica di λ_i . Il polinomio minimo $m_A(\lambda)$ di A divide $p_A(\lambda)$ e quindi è del tipo

$$m_A(\lambda) = \pm(\lambda - \lambda_1)^{d_1} \cdots (\lambda - \lambda_k)^{d_k}$$

dove d_1, \dots, d_k sono dei numeri da determinare. Per adesso sappiamo solo che $d_i \leq m_i$ per ogni i . Sia J la matrice di Jordan associata a A .

Proposizione 6.3.13. Il numero d_i è la massima taglia dei blocchi di Jordan in J con autovalore λ_i .

Dimostrazione. Per un singolo blocco di Jordan $B = B_{\lambda_0, n}$ otteniamo $m_B(\lambda) = (\lambda - \lambda_0)^n$ ed n è effettivamente la taglia di B , quindi ci siamo. Se J è fatto di vari blocchi di Jordan, usando più volte l'Esercizio 6.3.12 troviamo che $m_J(\lambda)$ è il minimo comune multiplo di tutti i polinomi $(\lambda - \lambda_i)^d$ al variare di tutti i blocchi di Jordan con autovalore λ_i e di taglia variabile d . Quindi otteniamo proprio

$$m_A(\lambda) = (\lambda - \lambda_1)^{d_1} \cdots (\lambda - \lambda_k)^{d_k}$$

dove effettivamente d_i è la massima taglia dei blocchi con autovalore λ_i . \square

In particolare, troviamo che $d_i \geq 1$. Quindi per ogni $i = 1, \dots, k$ abbiamo

$$1 \leq d_i \leq m_i = m_a(\lambda_i).$$

Osservazione 6.3.14. Le stesse disuguaglianze $1 \leq m_g(\lambda_i) \leq m_a(\lambda_i)$ valgono per la molteplicità geometrica $m_g(\lambda_i)$, ma attenzione a non fare confusione: qui d_i è la massima taglia dei blocchi con autovalore λ_i , mentre $m_g(\lambda_i)$ è il numero di questi blocchi. Questi due quantità non sono correlate e sono spesso differenti.

Esempio 6.3.15. Il polinomio minimo di

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

è $(x - 2)(x - 3)^3$.

Enunciamo un importante corollario.

Corollario 6.3.16. Una matrice A è diagonalizzabile se e solo se il polinomio minimo non ha radici ripetute, cioè è del tipo

$$m_A(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_n)$$

dove $\lambda_1, \dots, \lambda_n$ sono gli autovalori (distinti) di A .

Dimostrazione. Una matrice A è diagonalizzabile se e solo se la sua forma di Jordan è diagonale, e questo accade precisamente quando tutti i blocchi hanno taglia 1. \square

6.3.6. Applicazioni. Il Corollario 6.3.16 fornisce un criterio molto potente di diagonalizzabilità. Notiamo una differenza fra polinomio caratteristico e minimo: il primo si calcola subito, ma non ci dice se A sia diagonalizzabile; il secondo è più difficile da calcolare, ma ci dice subito se A sia diagonalizzabile oppure no. Mostriamo qualche applicazione di questo criterio.

Proposizione 6.3.17. Se $A \in M(n, \mathbb{C})$ è tale che $A^2 = A$, allora A è simile ad una matrice diagonale con solo valori 1 e 0 sulla diagonale.

Dimostrazione. Per ipotesi $A^2 - A = 0$ e quindi il polinomio $q(x) = x^2 - x = x(x - 1)$ annulla A . Il polinomio minimo $m_A(x)$ deve dividere $q(x)$ ed è quindi uno di questi:

$$x(x - 1), \quad x, \quad x - 1.$$

In tutti i casi $m_A(x)$ non ha radici ripetute e quindi A è diagonalizzabile. Ne segue che A è simile ad una matrice diagonale D , con gli autovalori sulla diagonale. Gli autovalori sono le radici di $m_A(x)$ e quindi possono essere solo 0 e 1. \square

L'esercizio seguente si risolve in modo analogo.

Esercizio 6.3.18. Se $A \in M(n, \mathbb{C})$ è tale che $A^2 = I$, allora A è simile ad una matrice diagonale con valori 1 e -1 sulla diagonale.

Attenzione però al caso seguente:

Esercizio 6.3.19. Se $A \in M(n, \mathbb{C})$ è tale che $A^2 = 0$, allora A è diagonalizzabile se e solo se $A = 0$.

Esercizi

Esercizio 6.1. Determina le forme di Jordan delle matrici seguenti:

$$\begin{pmatrix} 10 & 4 \\ -5 & 10 \end{pmatrix}, \quad \begin{pmatrix} 5 & -4 \\ 9 & -7 \end{pmatrix}, \quad \begin{pmatrix} 4 & 0 & 0 \\ 2 & 1 & 3 \\ 5 & 0 & 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ -1 & 1 & -1 \end{pmatrix},$$

$$\begin{pmatrix} 5 & 4 & 3 \\ -1 & 0 & -3 \\ 1 & -2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 9 & 7 & 3 \\ -9 & -7 & -4 \\ 4 & 4 & 4 \end{pmatrix}, \quad \begin{pmatrix} 2 & 2 & -1 \\ -1 & -1 & 1 \\ -1 & -2 & 2 \end{pmatrix}.$$

Esercizio 6.2. Elenca tutte le forme di Jordan che hanno polinomio caratteristico $(\lambda - 1)^2(\lambda + 2)^3$. Tra queste, determina quella che ha polinomio minimo $(\lambda - 1)(\lambda + 2)^2$.

Esercizio 6.3. Mostra che l'applicazione $f: \mathbb{C}_n[x] \rightarrow \mathbb{C}_n[x]$ che manda un polinomio $p(x)$ nella sua derivata $p'(x)$ è nilpotente di ordine $n+1$. Determina la sua forma di Jordan.

Esercizio 6.4. Mostra che le matrici complesse $n \times n$ nilpotenti di indice di nilpotenza n sono tutte simili fra loro.

Esercizio 6.5. Determina tutte le forme di Jordan 4×4 con un solo autovalore λ_0 e per ciascuna scrivi il suo polinomio minimo (sono 5.)

Esercizio 6.6. Calcola il polinomio minimo delle matrici seguenti:

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ -2 & -2 & 2 & 1 \\ 1 & 1 & -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 2 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 1 & 5 & 2 & -1 \\ 0 & -4 & 0 & -4 \end{pmatrix}, \quad \begin{pmatrix} 3 & -1 & 3 & -1 \\ 9 & -3 & 9 & -3 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 9 & -3 \end{pmatrix}$$

Esercizio 6.7. Trova due matrici $A, B \in M(2)$ che abbiano lo stesso polinomio caratteristico ma polinomi minimi diversi.

Esercizio 6.8. Costruisci due matrici complesse A e A' con lo stesso polinomio caratteristico e lo stesso polinomio minimo, che però non sono simili.

Esercizio 6.9. Costruisci due matrici complesse A e A' con lo stesso polinomio caratteristico, lo stesso polinomio minimo, e le stesse molteplicità geometriche di tutti gli autovalori, che però non sono simili.

Esercizio 6.10. Sia A una matrice complessa 7×7 tale che $(A - I)^3 = 0$ e $\text{rk}(A - I)^2 = 2$. Qual è la forma di Jordan di A ?

Esercizio 6.11. Sia A una matrice complessa quadrata tale che $A^n = I$ per qualche numero intero $n \geq 1$. Mostra che A è sempre diagonalizzabile e determina quali possano essere i suoi autovalori.

Esercizio 6.12. Sia A una matrice complessa tale che $A^2 - 2A + I = 0$. Mostra che A è diagonalizzabile se e solo se $A = I$.

Esercizio 6.13. Dimostra l'Esercizio 5.19 usando il polinomio minimo e il Corollario 6.3.16.

Esercizio 6.14. Sia $V = V_1 \oplus \dots \oplus V_k$ uno spazio vettoriale complesso e $T: V \rightarrow V$ un endomorfismo tale che $T(V_i) \subset V_i$ per ogni i . Mostra che T è diagonalizzabile se e solo se ciascuna restrizione $T|_{V_i}$ è diagonalizzabile.

Complementi

CAPITOLO 7

Prodotti scalari

In questo capitolo introduciamo per la prima volta alcuni concetti familiari della geometria euclidea: quello di *lunghezza* di un vettore, di *angolo* fra due vettori, e di *distanza* fra due punti.

Come possiamo introdurre queste nozioni? Il modo più efficiente consiste nella definizione di un *prodotto scalare* fra vettori.

7.1. Introduzione

Un endomorfismo di uno spazio vettoriale V è una funzione che prende come *input* un vettore di V e restituisce come *output* un vettore di V . Un prodotto scalare invece è una funzione che prende *due* vettori di V e restituisce uno scalare in \mathbb{R} . Questa funzione deve ovviamente soddisfare degli assiomi.

In tutto questo capitolo – a differenza dei precedenti – supporremo sempre che il campo sia $\mathbb{K} = \mathbb{R}$. In molti punti sarà importante parlare di numeri *positivi* e *negativi*, cosa che possiamo fare con \mathbb{R} e \mathbb{Q} e non con \mathbb{C} . Sarà anche importante fare la radice quadrata di un numero positivo, cosa che possiamo fare su \mathbb{R} ma non su \mathbb{Q} . Quindi supporremo sempre $\mathbb{K} = \mathbb{R}$.

7.1.1. Definizioni. Sia V uno spazio vettoriale reale.

Definizione 7.1.1. Un *prodotto scalare* su V è una applicazione

$$\begin{aligned} V \times V &\longrightarrow \mathbb{R} \\ (v, w) &\longmapsto \langle v, w \rangle \end{aligned}$$

che soddisfa i seguenti assiomi:

- (1) $\langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle$
- (2) $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$
- (3) $\langle v, w + w' \rangle = \langle v, w \rangle + \langle v, w' \rangle$
- (4) $\langle v, \lambda w \rangle = \lambda \langle v, w \rangle$
- (5) $\langle v, w \rangle = \langle w, v \rangle$

per ogni $v, v', w, w' \in V$ e ogni $\lambda \in \mathbb{R}$.

Il prodotto scalare è una applicazione che prende come *input* due vettori $v, w \in V$ e restituisce come *output* uno scalare, che indichiamo con $\langle v, w \rangle$. In alcuni aspetti il prodotto scalare si comporta come il prodotto fra numeri: gli assiomi (1) e (3) della definizione indicano che il prodotto è distributivo

rispetto alla somma, e l'assioma (5) indica che è commutativo. L'aggettivo "scalare" sta però ad indicare una differenza importante: il prodotto scalare di due vettori non è un vettore, ma "solo" uno scalare.

Un po' di terminologia: gli assiomi (1)–(4) dicono che il prodotto scalare è *bilineare*, cioè lineare "sia a sinistra che a destra"; l'assioma (5) dice che il prodotto scalare è *simmetrico*. Notiamo che gli assiomi (1), (2) e (5) implicano (3) e (4), infatti:

$$\begin{aligned}\langle v, w + w' \rangle &\stackrel{(5)}{=} \langle w + w', v \rangle \stackrel{(1)}{=} \langle w, v \rangle + \langle w', v \rangle \stackrel{(5)}{=} \langle v, w \rangle + \langle v, w' \rangle, \\ \langle v, \lambda w \rangle &\stackrel{(5)}{=} \langle \lambda w, v \rangle \stackrel{(2)}{=} \lambda \langle w, v \rangle \stackrel{(5)}{=} \lambda \langle v, w \rangle.\end{aligned}$$

Quindi in realtà gli assiomi (1), (2) e (3) sarebbero sufficienti per definire un prodotto scalare. Come prima applicazione degli assiomi, notiamo che $\langle v, 0 \rangle = 0$ per ogni $v \in V$. Infatti scrivendo 0 come $0 + 0$ troviamo che

$$\langle v, 0 \rangle = \langle v, 0 + 0 \rangle = \langle v, 0 \rangle + \langle v, 0 \rangle$$

e quindi $\langle v, 0 \rangle = 0$. In particolare otteniamo $\langle 0, 0 \rangle = 0$.

Quando vogliamo dare un nome al prodotto scalare, lo indichiamo con una lettera $g: V \times V \rightarrow \mathbb{R}$ e scriviamo $g(v, w)$ invece di $\langle v, w \rangle$.

7.1.2. Prodotto scalare degenere e definito positivo. Introduciamo ora due importanti assiomi aggiuntivi.

Definizione 7.1.2. Un prodotto scalare su V è:

- *degenere* se esiste $v \neq 0$ tale che $\langle v, w \rangle = 0$ per ogni $w \in V$;
- *definito positivo* se $\langle v, v \rangle > 0$ per ogni $v \in V$ non nullo.

I due assiomi appena introdotti sono mutualmente esclusivi:

Proposizione 7.1.3. *Un prodotto scalare definito positivo non è degenere.*

Dimostrazione. Ragioniamo per assurdo. Se fosse degenere, esisterebbe $v \in V$ non nullo tale che $\langle v, w \rangle = 0$ per ogni $w \in V$ ed in particolare avremmo $\langle v, v \rangle = 0$, contraddicendo l'ipotesi che $\langle v, v \rangle > 0$ per ogni v non nullo. \square

7.1.3. Prodotto scalare euclideo. Introduciamo un esempio importante.

Definizione 7.1.4. Il *prodotto scalare euclideo* su \mathbb{R}^n è definito come

$$\langle x, y \rangle = {}^t x \cdot y.$$

Nell'espressione ${}^t x \cdot y$ il vettore ${}^t x$ è un vettore riga e \cdot indica il prodotto fra matrici. In altre parole, se

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

allora

$$\langle x, y \rangle = (x_1, \dots, x_n) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \dots + x_n y_n.$$

Proposizione 7.1.5. *Il prodotto scalare euclideo è un prodotto scalare definito positivo su \mathbb{R}^n .*

Dimostrazione. Proviamo l'assioma (1):

$$\langle x + x', y \rangle = {}^t(x + x') \cdot y = ({}^t x + {}^t x') \cdot y = {}^t x \cdot y + {}^t x' \cdot y = \langle x, y \rangle + \langle x', y \rangle.$$

Gli altri assiomi di bilinearità si mostrano in modo simile. Per la simmetria, notiamo che l'espressione $\langle x, y \rangle = x_1 y_1 + \dots + x_n y_n$ è chiaramente simmetrica in x e y . Il prodotto scalare è definito positivo, perché se $x \neq 0$ allora

$$\langle x, x \rangle = x_1^2 + \dots + x_n^2 > 0.$$

La dimostrazione è completa. \square

Vedremo nel Capitolo 8 che il prodotto scalare euclideo su \mathbb{R}^2 e \mathbb{R}^3 coincide con l'usuale prodotto scalare fra vettori usato in fisica. In questo capitolo ci dedichiamo allo studio dei prodotti scalari nella loro generalità.

7.1.4. Matrici simmetriche. Abbiamo visto nel Capitolo 4 che una matrice quadrata $A \in M(n, \mathbb{R})$ determina un endomorfismo $L_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$, e che tutti gli endomorfismi di \mathbb{R}^n sono di questo tipo.

Analogamente, vediamo adesso che una matrice *simmetrica* $S \in M(n, \mathbb{R})$ determina un prodotto scalare g_S su \mathbb{R}^n . Dimostreremo in seguito che tutti i prodotti scalari di \mathbb{R}^n sono di questo tipo.

Ricordiamo che una matrice $S \in M(n)$ è *simmetrica* se ${}^t S = S$. Ricordiamo anche la formula

$${}^t(AB) = {}^t B {}^t A$$

che si applica per ogni prodotto di matrici AB . Più in generale, vale

$${}^t(A_1 \cdots A_k) = {}^t A_k \cdots {}^t A_1.$$

Proposizione 7.1.6. *Una matrice simmetrica S definisce un prodotto scalare g_S su \mathbb{R}^n nel modo seguente:*

$$g_S(x, y) = {}^t x \cdot S \cdot y.$$

Dimostrazione. Notiamo innanzitutto che le tre matrici ${}^t x$, S e x sono di taglia $1 \times n$, $n \times n$ e $n \times 1$, quindi il prodotto ${}^t x S y$ si può effettivamente fare e ha come risultato un numero. Quindi g_S è effettivamente una funzione $g_S: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$. L'assioma (1) di prodotto scalare segue dalla distributività del prodotto fra matrici:

$$g_S(x + x', y) = {}^t(x + x') S y = {}^t x S y + {}^t x' S y = g_S(x, y) + g_S(x', y).$$

Gli assiomi (2)-(4) si verificano in modo analogo. La simmetria (5) segue dal fatto che la matrice S è simmetrica:

$$g_S(x, y) = {}^t x S y = {}^t ({}^t x S y) = {}^t y {}^t S ({}^t x) = {}^t y S x = g_S(y, x).$$

La seconda uguaglianza è vera perché ${}^t x S y$ è un numero e quindi è uguale al suo trasposto. Nella quarta abbiamo usato che $S = {}^t S$. \square

Ogni matrice simmetrica S determina quindi un prodotto scalare g_S su \mathbb{R}^n . La notazione come prodotto fra matrici è efficiente perché è molto compatta. A volte è comunque utile sviluppare il prodotto:

Proposizione 7.1.7. *Vale l'uguaglianza*

$$g_S(x, y) = {}^t x S y = \sum_{i,j=1}^n x_i S_{ij} y_j.$$

Dimostrazione. Basta calcolare due volte il prodotto fra matrici:

$${}^t x S y = \sum_{i=1}^n x_i (S y)_i = \sum_{i=1}^n x_i \sum_{j=1}^n S_{ij} y_j = \sum_{i,j=1}^n x_i S_{ij} y_j.$$

La dimostrazione è conclusa. \square

Ricordiamo che e_1, \dots, e_n è la base canonica di \mathbb{R}^n .

Corollario 7.1.8. *Vale la relazione seguente:*

$$g_S(e_i, e_j) = {}^t e_i S e_j = S_{ij}.$$

Dimostrazione. Si applica la formula precedente, oppure si verifica direttamente che ${}^t e_i S e_j = {}^t e_i S^j = S_{ij}$ dove S^j è la j -esima colonna di S . \square

Abbiamo scoperto che l'elemento S_{ij} è precisamente il prodotto scalare $g_S(e_i, e_j)$ fra i vettori e_i ed e_j della base canonica di \mathbb{R}^n .

Esempio 7.1.9. Se come matrice simmetrica prendiamo la matrice identità, otteniamo il prodotto scalare euclideo:

$$g_{I_n}(x, y) = {}^t x I_n y = {}^t x y = x_1 y_1 + \dots + x_n y_n.$$

Ad esempio su \mathbb{R}^2 con $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ otteniamo

$$g_{I_2}(x, y) = x_1 y_1 + x_2 y_2.$$

Esempio 7.1.10. La matrice $S = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ definisce il prodotto scalare su \mathbb{R}^2 :

$$g_S(x, y) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = 2x_1^2 + x_1 y_2 + x_2 y_1 + y_2^2.$$

7.1.5. Forme quadratiche. Sia $S \in M(n, \mathbb{R})$ una matrice simmetrica. Notiamo che se poniamo $x = y$ otteniamo questa espressione:

$$g_S(x, x) = \sum_{i,j=1}^n S_{ij}x_i x_j.$$

Questo è un polinomio di secondo grado nelle variabili x_1, \dots, x_n i cui coefficienti sono proprio gli elementi S_{ij} della matrice. Un polinomio di questo tipo è chiamato *forma quadratica* e si indica generalmente con

$$q_S(x) = \sum_{i,j=1}^n S_{ij}x_i x_j.$$

Ad esempio, le matrici simmetriche

$$\begin{pmatrix} 1 & -3 \\ -3 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

definiscono le forme quadratiche

$$x_1^2 - 6x_2x_3, \quad x_1^2 + x_2^2 - x_3^2, \quad 2x_1x_2 + 2x_2x_3 + 2x_3x_1.$$

Notiamo se $i \neq j$ gli elementi $S_{ij} = S_{ji}$ contribuiscono entrambi al monomio $x_i x_j$, che ha quindi coefficiente pari a $2S_{ij}$. Di converso, la forma quadratica

$$q_S(x) = x_1^2 + 4x_1x_2 - x_2^2 + 4x_3^2$$

è determinata dalla matrice

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

Qui il monomio $4x_1x_2$ viene spezzato in due monomi $2x_1x_2 + 2x_2x_1$. La matrice identità I_n definisce la forma quadratica

$$x_1^2 + \dots + x_n^2.$$

Come gli endomorfismi sono strettamente collegati a sistemi di equazioni di primo grado, così i prodotti scalari sono strettamente collegati alle forme quadratiche. Lo studio approfondito dei prodotti scalari che faremo in questo capitolo ci servirà anche a capire le forme quadratiche successivamente nel Capitolo 11.

Notiamo che per definizione il prodotto scalare g_S è definito positivo precisamente quando la corrispondente forma quadratica è positiva:

$$q_S(x) > 0 \quad \forall x \neq 0.$$

7.1.6. Matrici diagonali. Le matrici diagonali sono matrici simmetriche molto semplici e facili da studiare. Consideriamo una matrice diagonale

$$D = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}$$

Il prodotto scalare $g_D(x, y) = {}^t x D y$ su \mathbb{R}^n è semplicemente:

$$g_D(x, y) = d_1 x_1 y_1 + \dots + d_n x_n y_n.$$

La corrispondente forma quadratica è

$$q_D(x) = d_1 x_1^2 + \dots + d_n x_n^2.$$

In particolare è facile capire quando il prodotto scalare è definito positivo:

Proposizione 7.1.11. Il prodotto scalare g_D è definito positivo $\iff d_i > 0$ per ogni $i = 1, \dots, n$.

Dimostrazione. Se $d_i > 0$ per ogni i allora otteniamo

$$q_S(x) = x_1^2 d_1 + \dots + x_n^2 d_n > 0$$

per ogni vettore non nullo x . Quindi il prodotto scalare è definito positivo. D'altra parte, se esiste un $d_i \leq 0$, allora

$$q_S(e_i) = d_i \leq 0$$

implica che il prodotto scalare non è definito positivo. \square

Analogamente è facile capire quando il prodotto scalare è non degenere:

Proposizione 7.1.12. Il prodotto scalare g_D è non degenere $\iff d_i \neq 0$ per ogni $i = 1, \dots, n$.

Dimostrazione. Notiamo che per ogni $x \in \mathbb{R}^n$ abbiamo $g_D(x, e_i) = d_i x_i$.

Se esiste un i tale che $d_i = 0$, allora $g_D(x, e_i) = 0$ per ogni $x \in \mathbb{R}^n$ e quindi il prodotto scalare è degenere.

Se $d_i \neq 0$ per ogni i , allora il prodotto scalare è non degenere: per ogni $x \in \mathbb{R}^n$ non nullo esiste un vettore $y \in \mathbb{R}^n$ tale che $g_D(x, y) \neq 0$. Infatti, siccome $x \neq 0$, deve valere $x_i \neq 0$ per qualche i , e allora $g_D(x, e_i) = d_i x_i \neq 0$. \square

Entrambe le dimostrazioni appena concluse sono molto istruttive per assimilare tutte le nozioni introdotte finora in questo capitolo.

Ad esempio, se D è la matrice diagonale

$$\begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{oppure} \quad \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}$$

il prodotto scalare g_D su \mathbb{R}^2 è degenere solo nel secondo caso ed è definito positivo solo nel terzo caso. Ricordiamo che

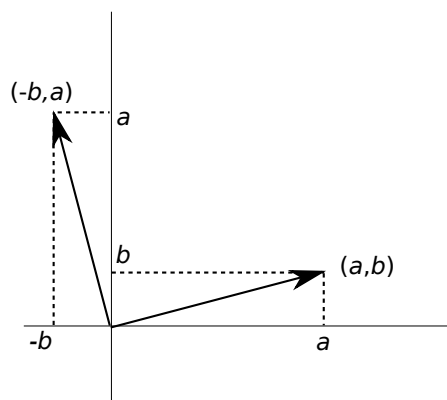


Figura 7.1. Due vettori ortogonali in \mathbb{R}^2 .

definito positivo \implies non-degenere.

Non vale però l'implicazione opposta: la prima matrice definisce un prodotto scalare non degenere che non è però definito positivo.

7.1.7. Vettori ortogonali. Introduciamo adesso alcune definizioni generali. Consideriamo un prodotto scalare su V . Due vettori $v, w \in V$ sono *ortogonali* se

$$\langle v, w \rangle = 0.$$

Esempio 7.1.13. Consideriamo il prodotto scalare euclideo su \mathbb{R}^2 . I vettori ortogonali ad un vettore dato $\begin{pmatrix} a \\ b \end{pmatrix} \neq 0$ sono quei vettori $\begin{pmatrix} x \\ y \end{pmatrix}$ tali che

$$ax + by = 0.$$

Risolvendo il sistema, ne deduciamo che i vettori ortogonali ad $\begin{pmatrix} a \\ b \end{pmatrix}$ sono precisamente i vettori che stanno nella retta

$$\text{Span} \begin{pmatrix} -b \\ a \end{pmatrix}.$$

Il vettore $\begin{pmatrix} -b \\ a \end{pmatrix}$ forma effettivamente un angolo retto con $\begin{pmatrix} a \\ b \end{pmatrix}$, come mostrato dalla Figura 7.1. Il prodotto scalare euclideo su \mathbb{R}^2 e \mathbb{R}^3 corrisponde alla geometria euclidea; se però scegliamo un altro prodotto scalare su \mathbb{R}^2 o \mathbb{R}^3 la nozione di ortogonalità può cambiare considerevolmente.

Esempio 7.1.14. Rispetto al prodotto scalare euclideo su \mathbb{R}^n , due vettori e_i ed e_j della base canonica con $i \neq j$ sono sempre ortogonali fra loro. Infatti $\langle e_i, e_j \rangle = {}^t e_i \cdot e_j = 0$.

Osservazione 7.1.15. Ricordiamo che un prodotto scalare su V è degenere se esiste un $v \in V$ non nullo tale che $\langle v, w \rangle = 0$ per ogni $w \in V$. Questa condizione può essere espressa dicendo che esiste un vettore $v \in V$ ortogonale a *tutti* i vettori di V .

7.1.8. Vettori isotropi. In un prodotto scalare, può accadere che un vettore sia ortogonale a se stesso! Sia V uno spazio vettoriale munito di un prodotto scalare. Un vettore $v \in V$ tale che $\langle v, v \rangle = 0$ è detto *isotropo*.

Se il prodotto scalare è definito positivo, come ad esempio l'usuale prodotto scalare euclideo su \mathbb{R}^2 o su \mathbb{R}^n , l'unico vettore isotropo è quello nullo. Prodotti scalari più generali possono però contenere vettori isotropi non banali.

Esempio 7.1.16. Se V è degenere, esiste un vettore $v \neq 0$ ortogonale a tutti i vettori di V . In particolare, v è ortogonale a se stesso e quindi è isotropo.

Abbiamo capito che se il prodotto scalare è degenere ci sono ovviamente dei vettori isotropi non banali. È però importante notare che, come mostra il prossimo esempio, ci possono essere dei vettori isotropi non banali anche in prodotti scalari non degeneri.

Osservazione 7.1.17. Consideriamo il prodotto scalare g_S su \mathbb{R}^n definito da una matrice simmetrica $S \in M(n, \mathbb{R})$. I vettori isotropi di g_S sono precisamente i vettori $x \in \mathbb{R}^n$ che annullano la forma quadratica q_S , cioè le soluzioni dell'equazione di secondo grado in più variabili:

$$q_S(x) = 0.$$

Esempio 7.1.18. Consideriamo la matrice simmetrica

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Per quanto visto nella sezione precedente, il prodotto scalare g_S è non degenere e non è definito positivo. La forma quadratica corrispondente è

$$q_S(x) = x_1^2 - x_2^2.$$

I vettori isotropi sono precisamente i vettori $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ tali che $x_1^2 - x_2^2 = 0$, cioè tali che $x_1 = \pm x_2$. I vettori isotropi formano quindi le due bisettrici dei quattro quadranti del piano cartesiano. Ci sono vettori isotropi non banali.

7.1.9. Radicale. Introduciamo adesso una nozione utile a gestire i prodotti scalari degeneri, quella di *radicale*.

Sia V uno spazio vettoriale munito di un prodotto scalare. Il *radicale* è l'insieme dei vettori $v \in V$ che sono ortogonali a tutti gli altri. In altre parole, un vettore v è nel radicale se $\langle v, w \rangle = 0$ per *qualsiasi* vettore $w \in V$.

Il radicale di V è generalmente indicato con V^\perp . In simboli:

$$V^\perp = \{v \in V \mid \langle v, w \rangle = 0 \forall w \in V\}.$$

Notiamo subito che se $v \in V^\perp$ allora v è anche isotropo, ma non è vero il viceversa! Un vettore v isotropo non è necessariamente nel radicale: un vettore v può essere ortogonale a se stesso senza però essere ortogonale anche a tutti gli altri. Osserviamo inoltre che

$$\text{il prodotto scalare è non degenere} \iff V^\perp = \{0\}.$$

Proposizione 7.1.19. *Il radicale V^\perp è un sottospazio vettoriale di V .*

Dimostrazione. Mostriamo che valgono i tre assiomi di sottospazio:

(1) $0 \in V^\perp$ infatti $\langle 0, w \rangle = 0 \forall w \in V$.

(2) Se $v, v' \in V^\perp$, allora per ogni $w \in V$ otteniamo

$$\langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle = 0 + 0 = 0$$

e quindi anche $v + v' \in V^\perp$.

(3) La moltiplicazione per scalare è analoga.

La dimostrazione è conclusa. \square

Come primo esempio esaminiamo i prodotti scalari g_S su \mathbb{R}^n . Sia $S \in M(n, \mathbb{R})$ una matrice simmetrica.

Proposizione 7.1.20. *Il radicale di g_S è $\ker S$. Quindi il prodotto scalare g_S è degenere $\iff \det S = 0$.*

Dimostrazione. Se $y \in \ker S$, allora per ogni $x \in \mathbb{R}^n$ abbiamo

$$g_S(x, y) = {}^t x S y = 0$$

perché $Sy = 0$. D'altra parte, se un vettore $y \in \mathbb{R}^n$ è tale che ${}^t x S y = 0$ per ogni $x \in \mathbb{R}^n$, allora scriviamo $y' = S y$ e otteniamo

$${}^t x y' = 0$$

per ogni $x \in \mathbb{R}^n$. Questo implica facilmente che $y' = 0$, cioè che $y \in \ker S$. \square

Esempio 7.1.21. Il prodotto scalare su \mathbb{R}^2 definito dalla matrice $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ considerato nell'Esempio 7.1.18 è un esempio importante da tenere a mente. Non è degenere perché $\det \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \neq 0$. Non è neppure definito positivo per la Proposizione 7.1.11. Il radicale è banale (perché è non degenere), ma ci sono comunque dei vettori isotropi (le due bisettrici dei quadranti). È quindi possibile che in un prodotto scalare un vettore sia ortogonale a se stesso ma non a tutti gli altri (cioè il vettore è isotropo ma non sta nel radicale).

7.1.10. Altri esempi. Introduciamo finalmente alcuni prodotti scalari in spazi vettoriali V diversi da \mathbb{R}^n .

Esempio 7.1.22. Un importante prodotto scalare sullo spazio $\mathbb{R}[x]$ dei polinomi è definito nel modo seguente:

$$\langle p, q \rangle = \int_0^1 p(t)q(t)dt$$

La bilinearità si verifica facilmente e la simmetria è immediata. Il prodotto scalare è definito positivo perché

$$\langle p, p \rangle = \int_0^1 p(t)^2 dt > 0$$

per ogni polinomio $p(x)$ non nullo.

È spesso utile definire un prodotto scalare anche sugli spazi di matrici:

Esercizio 7.1.23. Consideriamo il prodotto scalare su $M(n, \mathbb{R})$ dato da

$$\langle A, B \rangle = \text{tr}({}^tAB)$$

Mostra che questo è effettivamente un prodotto scalare ed è definito positivo.

Esempio 7.1.24. Consideriamo sullo spazio $\mathbb{R}_2[x]$ il prodotto scalare

$$\langle p, q \rangle = p(0)q(0) + p(1)q(1) + p(2)q(2).$$

Si verifica facilmente che questo è effettivamente un prodotto scalare. Inoltre è definito positivo: infatti

$$\langle p, p \rangle = p(0)^2 + p(1)^2 + p(2)^2 > 0$$

per qualsiasi polinomio $p(x)$ non nullo di grado ≤ 2 : infatti un tale polinomio non può annullarsi in tre valori diversi 0, 1, 2. D'altra parte, il prodotto scalare

$$\langle p, q \rangle = p(0)q(0) + p(1)q(1)$$

è degenere: se prendiamo $p(x) = x(1-x)$ otteniamo $\langle p, q \rangle = 0$ per qualsiasi polinomio $q(x) \in \mathbb{R}_2[x]$. Infine, il prodotto scalare

$$\langle p, q \rangle = p(0)q(0) + p(1)q(1) - p(2)q(2)$$

non è degenere (esercizio: da risolvere dopo aver letto la prossima sezione) ma contiene dei vettori isotropi non banali, ad esempio $p(x) = x - 1$, infatti

$$\langle p, p \rangle = (-1)(-1) + 0 \cdot 0 - 1 \cdot 1 = 0.$$

7.2. Matrice associata

Come per le applicazioni lineari, è possibile codificare i prodotti scalari in modo efficiente usando le matrici, dopo aver fissato delle basi.

7.2.1. Definizione. Sia V uno spazio vettoriale reale e

$$g: V \times V \longrightarrow \mathbb{R}$$

un prodotto scalare. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base per V .

Definizione 7.2.1. La *matrice associata* al prodotto scalare g nella base \mathcal{B} è la matrice simmetrica S il cui elemento S_{ij} è dato da

$$S_{ij} = g(v_i, v_j).$$

La matrice S è indicata a volte con $[g]_{\mathcal{B}}$.

Esempio 7.2.2. Sia $S \in M(n, \mathbb{R})$ una matrice simmetrica. La matrice associata al prodotto scalare g_S su \mathbb{R}^n rispetto alla base canonica \mathcal{C} è la matrice S stessa, cioè

$$[g_S]_{\mathcal{C}} = S.$$

Infatti per il Corollario 7.1.8 abbiamo $S_{ij} = g_S(e_i, e_j)$.

Esempio 7.2.3. Consideriamo il prodotto scalare euclideo g su \mathbb{R}^2 . Scegliamo la base

$$\mathcal{B} = \left\{ v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

e calcoliamo la matrice associata a g rispetto a \mathcal{B} :

$$S = [g]_{\mathcal{B}} = \begin{pmatrix} g(v_1, v_1) & g(v_1, v_2) \\ g(v_2, v_1) & g(v_2, v_2) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Esercizio 7.2.4. Scrivere la matrice associata al prodotto scalare su $\mathbb{R}_2[x]$

$$g(p, q) = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$$

rispetto alla base canonica $\mathcal{C} = \{1, x, x^2\}$.

Perché nella matrice associata scegliamo di mettere i numeri $S_{ij} = g(v_i, v_j)$? Il motivo principale è che questi numeri sono sufficienti per calcolare il prodotto scalare $g(v, w)$ di due vettori qualsiasi. Siano

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n, \quad w = \mu_1 v_1 + \dots + \mu_n v_n$$

due vettori arbitrari di V , espressi usando la base \mathcal{B} .

Proposizione 7.2.5. *Vale*

$$g(v, w) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j g(v_i, v_j).$$

Dimostrazione. Scriviamo

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n, \quad w = \mu_1 v_1 + \dots + \mu_n v_n.$$

Usando la bilinearità di g otteniamo

$$\begin{aligned} g(v, w) &= g(\lambda_1 v_1 + \dots + \lambda_n v_n, \mu_1 v_1 + \dots + \mu_n v_n) \\ &= \sum_{i=1}^n g(\lambda_i v_i, \mu_1 v_1 + \dots + \mu_n v_n) = \sum_{i=1}^n \sum_{j=1}^n g(\lambda_i v_i, \mu_j v_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j g(v_i, v_j). \end{aligned}$$

La dimostrazione è conclusa. □

Corollario 7.2.6. *Per ogni coppia di vettori $v, w \in V$ vale l'uguaglianza*

$$g(v, w) = {}^t[v]_{\mathcal{B}} \cdot S \cdot [w]_{\mathcal{B}}.$$

Dimostrazione. Per la Proposizione 7.2.5 abbiamo

$$g(v, w) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j S_{ij}$$

Notiamo inoltre che

$$[v]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}, \quad [w]_{\mathcal{B}} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}.$$

Operando con il prodotto riga per colonna si verifica che effettivamente

$$(\lambda_1 \ \cdots \ \lambda_n) \cdot S \cdot \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j S_{ij}.$$

Questo conclude la dimostrazione. \square

Osservazione 7.2.7. Se scriviamo $x = [v]_{\mathcal{B}}$ e $y = [w]_{\mathcal{B}}$, vediamo che

$$g(v, w) = {}^t x S y.$$

Abbiamo scoperto un fatto importante, del tutto analogo all'Osservazione 4.3.5 per le applicazioni lineari: dopo aver scelto una base, qualsiasi prodotto scalare g su qualsiasi spazio vettoriale V di dimensione n può essere interpretato come un prodotto scalare su \mathbb{R}^n del tipo $\langle x, y \rangle = {}^t x S y$. È sufficiente sostituire i vettori v e w con le loro coordinate x e y , ed usare la matrice associata S .

7.2.2. Cambiamento di base. Come abbiamo fatto per le applicazioni lineari, studiamo adesso come cambia la matrice associata ad un prodotto scalare se cambiamo base. Questo ci porterà a considerare una nuova relazione di equivalenza fra matrici simmetriche.

Siano $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{B}' = \{v'_1, \dots, v'_n\}$ due basi di V . Ricordiamo che $M = [\text{id}]_{\mathcal{B}}^{\mathcal{B}'}$ è la matrice di cambiamento di base da \mathcal{B}' a \mathcal{B} . Siano S e S' le matrici associate al prodotto scalare g nelle basi \mathcal{B} e \mathcal{B}' .

Proposizione 7.2.8. *Vale la relazione seguente:*

$$S' = {}^t M S M.$$

Dimostrazione. Per ogni vettore $v \in V$ ricordiamo che $[v]_{\mathcal{B}} = M \cdot [v]_{\mathcal{B}'}$ ed in particolare $[v'_i]_{\mathcal{B}} = M \cdot e_i = M^i$ è la i -esima colonna di M . Otteniamo

$$S'_{ij} = g(v'_i, v'_j) = {}^t [v'_i]_{\mathcal{B}} \cdot S \cdot [v'_j]_{\mathcal{B}} = {}^t M^i \cdot S \cdot M^j.$$

Questo vale per ogni i, j , quindi $S' = {}^t M S M$. \square

Esempio 7.2.9. Abbiamo già notato nell'Esempio 7.2.3 che la matrice associata al prodotto scalare euclideo di \mathbb{R}^2 rispetto alla base

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

è $S = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. La matrice di cambiamento di base è $M = [\text{id}]_{\mathcal{B}}^{\mathcal{C}} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. La Proposizione 7.2.8 ci dice che $S = {}^t M I_2 M$, e infatti verifichiamo facilmente che vale l'uguaglianza $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

7.2.3. Matrici congruenti. L'operazione di cambiamento di base per gli endomorfismi ha portato nella Sezione 4.4.3 ad una relazione di equivalenza fra matrici quadrate chiamata *similitudine*. Con i prodotti scalari, il cambiamento di base porta ad una diversa relazione di equivalenza detta *congruenza*, riservata alle matrici simmetriche.

Definizione 7.2.10. Due matrici $n \times n$ simmetriche S e S' sono *congruenti* se esiste una matrice $n \times n$ invertibile M per cui $S = {}^tMS'M$.

Esercizio 7.2.11. La congruenza fra matrici simmetriche è una relazione di equivalenza.

Si deve stare attenti a non confondere questa relazione di equivalenza con la similitudine, dove al posto della trasposta tM sta l'inversa M^{-1} .

Esempio 7.2.12. Le matrici

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

non sono simili (ad esempio, perché hanno determinanti diversi), ma sono congruenti: prendiamo $M = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{pmatrix}$ e verifichiamo facilmente che $S' = {}^tMSM$.

Due matrici simili S e S' hanno sempre lo stesso determinante, due matrici congruenti no! Possiamo comunque dimostrare quanto segue.

Proposizione 7.2.13. *Se S e S' sono congruenti, allora $\det S$ e $\det S'$ hanno lo stesso segno, cioè $\det S$ è positivo, nullo, o negativo $\iff \det S'$ è positivo, nullo, o negativo rispettivamente.*

Dimostrazione. Per ipotesi esiste una M con $\det M \neq 0$ tale che $S' = {}^tMSM$ e quindi usando il teorema di Binet troviamo

$$\det S' = \det({}^tM) \det S \det M = \det M \det S \det M = (\det M)^2 \det S.$$

La proposizione è dimostrata perché $(\det M)^2 > 0$. □

7.3. Sottospazio ortogonale

7.3.1. Sottospazio ortogonale. Sia V dotato di un prodotto scalare. Sia $W \subset V$ un sottospazio. Il *sottospazio ortogonale* W^\perp è per definizione l'insieme

$$W^\perp = \{v \in V \mid \langle v, w \rangle = 0 \ \forall w \in W\}.$$

In altre parole, l'insieme W^\perp è formato da quei vettori $v \in V$ che sono ortogonali a qualsiasi vettore di W . Notiamo che questa definizione è coerente con la notazione V^\perp già usata per definire il radicale, che come sappiamo è l'insieme dei vettori ortogonali a qualsiasi vettore di tutto lo spazio V .

Proposizione 7.3.1. *L'insieme W^\perp è effettivamente un sottospazio di V .*

Dimostrazione. Come di consueto dobbiamo mostrare tre cose:

- (1) $0 \in W^\perp$, infatti $\langle 0, w \rangle = 0$ sempre;
 (2) se $v, v' \in W^\perp$ allora $v + v' \in W^\perp$, infatti

$$\langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle = 0 + 0 = 0$$

per ogni $w \in W$;

- (3) se $v \in W^\perp$ allora $\lambda v \in W^\perp$, e questo è lasciato per esercizio.

La dimostrazione è completa. \square

Per determinare W^\perp concretamente è molto utile il fatto seguente.

Proposizione 7.3.2. Se $W = \text{Span}(w_1, \dots, w_k)$, allora

$$W^\perp = \{v \in V \mid \langle v, w_i \rangle = 0 \forall i = 1, \dots, k\}.$$

Dimostrazione. Se $v \in W^\perp$ allora $\langle v, w \rangle = 0$ per ogni $w \in W$ e quindi in particolare $\langle v, w_i \rangle = 0$ per ogni $i = 1, \dots, k$.

D'altro canto, se $\langle v, w_i \rangle = 0$ per ogni $i = 1, \dots, k$ allora per ogni $w \in W$ possiamo scrivere $w = \lambda_1 w_1 + \dots + \lambda_k w_k$ e quindi otteniamo

$$\langle v, w \rangle = \langle v, \lambda_1 w_1 + \dots + \lambda_k w_k \rangle = \lambda_1 \langle v, w_1 \rangle + \dots + \lambda_k \langle v, w_k \rangle = 0.$$

Quindi $v \in W^\perp$. \square

Concretamente, l'insieme W^\perp può essere descritto come l'insieme dei vettori v che soddisfano le k relazioni $\langle v, w_i \rangle = 0$ con $i = 1, \dots, k$.

7.3.2. Restrizioni. Sia V dotato di un prodotto scalare g . Restringendo il prodotto scalare g ai soli vettori di un sottospazio $W \subset V$, si ottiene un prodotto scalare per lo spazio W che possiamo indicare con $g|_W$. Il prodotto scalare $g|_W$ è la *restrizione* di g a W .

Proposizione 7.3.3. Notiamo che:

- (1) se g è definito positivo, anche $g|_W$ è definito positivo;
 (2) se g è non degenere, non è detto che $g|_W$ sia non degenere!

Dimostrazione. (1). Se $g(v, v) > 0$ per ogni $v \in V$ non nullo, lo stesso vale per ogni $v \in W$ non nullo.

(2) Qui dobbiamo fornire un controesempio. Prendiamo $V = \mathbb{R}^2$ con il prodotto scalare $\langle x, y \rangle = {}^t x S y$ determinato dalla matrice

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Sappiamo che S è non degenere. Definiamo quindi W come la retta $W = \text{Span}(v)$ con $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Il prodotto scalare $g|_W$ è il prodotto nullo, perché v è isotropo: quindi per ogni coppia di vettori $\lambda v, \mu v \in W$ otteniamo

$$g(\lambda v, \mu v) = \lambda \mu g(v, v) = 0.$$

In particolare la restrizione $g|_W$ è degenere. \square

Alla fine della dimostrazione abbiamo anche notato il fatto seguente.

Proposizione 7.3.4. *Se $W = \text{Span}(v)$, la restrizione $g|_W$ è degenere $\iff v$ è un vettore isotropo.*

7.3.3. Dimensione del sottospazio ortogonale. Se W è un sottospazio vettoriale, ci chiediamo adesso in che relazione siano i due sottospazi W e W^\perp . Sia $n = \dim V$ e g il prodotto scalare.

Proposizione 7.3.5. *Valgono i fatti seguenti:*

- (1) $\dim W^\perp \geq n - \dim W$;
- (2) se g è non-degenere, allora $\dim W^\perp = n - \dim W$;
- (3) se $g|_W$ è non-degenere, allora $V = W \oplus W^\perp$;
- (4) se g è definita positiva, allora $V = W \oplus W^\perp$.

Dimostrazione. Sia w_1, \dots, w_k una base di W . Un vettore $v \in V$ è ortogonale ad ogni vettore di W se (e solo se) è ortogonale a ciascun vettore della base w_1, \dots, w_k . Consideriamo l'applicazione lineare $T: V \rightarrow \mathbb{R}^k$ definita nel modo seguente:

$$T(v) = \begin{pmatrix} \langle v, w_1 \rangle \\ \vdots \\ \langle v, w_k \rangle \end{pmatrix}.$$

Lo spazio W^\perp è precisamente il nucleo di T . Per il teorema della dimensione

$$\dim W^\perp = \dim \ker T = n - \dim \text{Im } T \geq n - k$$

infatti $\dim \text{Im } T \leq k$, perché $\text{Im } T \subset \mathbb{R}^k$. Questo mostra il punto (1), quindi ora passiamo a mostrare (2). Completiamo w_1, \dots, w_k a base w_1, \dots, w_n di V e consideriamo come sopra l'applicazione lineare $T': V \rightarrow \mathbb{R}^n$ data da

$$T'(v) = \begin{pmatrix} \langle v, w_1 \rangle \\ \vdots \\ \langle v, w_n \rangle \end{pmatrix}.$$

Siccome g non è degenere, la mappa T' è iniettiva. Poiché $\dim V = \dim \mathbb{R}^n = n$, la mappa T' è anche suriettiva. Questo implica facilmente che anche T è suriettiva. Quindi $\dim \text{Im } T = k$ e $\dim W^\perp = n - k$. Il punto (2) è dimostrato.

Passiamo a dimostrare il punto (3). Se $g|_W$ è non degenere, nessun vettore non banale di W è ortogonale a tutti i vettori di W e quindi $W \cap W^\perp = \{0\}$. Per il punto (1) sappiamo che $\dim W + \dim W^\perp \geq n$, quindi per la formula di Grassmann deduciamo che $\dim(W + W^\perp) \geq \dim W + \dim W^\perp \geq n$, quindi $V = W + W^\perp$ e allora $V = W \oplus W^\perp$.

Se g è definito positivo, ogni restrizione $g|_W$ è definita positiva ed in particolare non è degenere, quindi (3) implica (4). \square

Esercizio 7.3.6. Sia V dotato di un prodotto scalare g e $U, W \subset V$ due sottospazi. Valgono i fatti seguenti:

- se $U \subset W$ allora $W^\perp \subset U^\perp$;
- $(U^\perp)^\perp \supset U$, con l'uguaglianza se g è non-degenere;
- $(U + W)^\perp = U^\perp \cap W^\perp$.

7.3.4. Esempi. Facciamo qualche esempio di calcolo di un sottospazio ortogonale. Iniziamo con il prodotto scalare euclideo su \mathbb{R}^n .

Un sottospazio $U \subset \mathbb{R}^n$ è descritto generalmente in forma parametrica, come $U = \text{Span}(v_1, \dots, v_k)$, oppure in forma cartesiana, come $U = \ker A = \{x \mid Ax = 0\}$. In entrambi i casi è facile descrivere lo spazio ortogonale U^\perp rispetto al prodotto scalare euclideo. Questo è rispettivamente

$$U^\perp = \{x \mid \langle x, v_1 \rangle = \dots = \langle x, v_k \rangle = 0\}$$

oppure

$$U^\perp = \text{Span}({}^tA_1, \dots, {}^tA_h)$$

dove A_1, \dots, A_h sono le righe di A .

Esempio 7.3.7. Se $U = \{ax + by + cz = 0\}$ è un piano in \mathbb{R}^3 , la retta ortogonale U^\perp è semplicemente

$$U^\perp = \text{Span} \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

D'altra parte, se U è la retta

$$U = \text{Span} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

allora U^\perp è il piano $U^\perp = \{ax + by + cz = 0\}$. Se

$$U = \begin{cases} ax + by + cz = 0, \\ dx + ey + fz = 0, \end{cases}$$

allora

$$U^\perp = \text{Span} \left(\begin{pmatrix} a \\ b \\ c \end{pmatrix}, \begin{pmatrix} d \\ e \\ f \end{pmatrix} \right).$$

7.3.5. Tipi di prodotti scalari. Diciamo che un prodotto scalare su V è

- *semi-definito positivo* se $\langle v, v \rangle \geq 0$ per ogni $v \in V$;
- *definito negativo* se $\langle v, v \rangle < 0$ per ogni $v \in V$ non nullo;
- *semi-definito negativo* se $\langle v, v \rangle \leq 0$ per ogni $v \in V$;
- *indefinito* se esistono v e w per cui $\langle v, v \rangle > 0$ e $\langle w, w \rangle < 0$.

Consideriamo una matrice diagonale D , con elementi d_1, \dots, d_n sulla diagonale.

Esercizio 7.3.8. Il prodotto scalare definito da D è

- semi-definito positivo $\iff d_i \geq 0$ per ogni i ;

- definito negativo $\iff d_i < 0$ per ogni i ;
- semi-definito negativo $\iff d_i \leq 0$ per ogni i ;
- indefinito \iff esistono i e j con $d_i > 0$ e $d_j < 0$.

7.4. Classificazione dei prodotti scalari

Lo scopo principale di questa sezione è ottenere una classificazione completa dei prodotti scalari.

7.4.1. Basi ortogonali e ortonormali. Sia g un prodotto scalare su uno spazio vettoriale V di dimensione n .

Definizione 7.4.1. Una base $\mathcal{B} = \{v_1, \dots, v_n\}$ è *ortogonale* se $g(v_i, v_j) = 0$ per ogni $i \neq j$. In altre parole, una base \mathcal{B} è ortogonale se due vettori distinti di \mathcal{B} sono sempre ortogonali fra loro.

Una base ortogonale è *ortonormale* se per ciascun i il numero $g(v_i, v_i)$ è 1, 0 oppure -1 .

Sia S la matrice associata a g nella base \mathcal{B} . Segue immediatamente dalla definizione che \mathcal{B} è ortogonale $\iff S$ è una matrice diagonale. Ricordiamo infatti che $S_{ij} = g(v_i, v_j)$, quindi chiedere che $g(v_i, v_j) = 0$ per ogni $i \neq j$ è come chiedere che S sia diagonale.

D'altra parte \mathcal{B} è ortonormale \iff la matrice S è diagonale, e inoltre gli elementi sulla diagonale sono soltanto di tre tipi: 1, 0 e -1 .

Esempio 7.4.2. La base canonica è una base ortonormale per il prodotto scalare euclideo. Infatti $\langle e_i, e_j \rangle = {}^t e_i e_j = 0$ per ogni $i \neq j$ e inoltre $\langle e_i, e_i \rangle = 1$ per ogni i .

Esempio 7.4.3. Consideriamo il prodotto scalare su \mathbb{R}^2 definito dalla matrice $S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Notiamo che

$$(x_1 \ y_1) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = x_2 y_1 + x_1 y_2.$$

La base $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ è una base ortogonale per il prodotto scalare, perché ${}^t v_1 S v_2 = 0$. Questa ovviamente non è l'unica base ortogonale possibile, ce ne sono molte altre: ad esempio anche $w_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $w_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ è una base ortogonale, perché ${}^t w_1 S w_2 = 0$.

7.4.2. Diagonalizzazione dei prodotti scalari. Sia V uno spazio vettoriale di dimensione n e g un prodotto scalare su V . Dimostriamo adesso il fatto seguente.

Teorema 7.4.4 (Teorema di Sylvester). *Esiste sempre una base ortogonale $\mathcal{B} = \{v_1, \dots, v_n\}$ per V .*

Dimostrazione. Lavoriamo per induzione su $n = \dim V$. Se $n = 1$, una base di V è formata da un vettore solo ed è sempre ortogonale. Supponiamo che l'enunciato sia valido per $n - 1$ e lo mostriamo per n .

Supponiamo che esista almeno un vettore $v \in V$ non isotropo. Consideriamo la retta vettoriale $W = \text{Span}(v)$. La restrizione di g a W è non-degenere perché $g(v, v) \neq 0$. Quindi la Proposizione 7.3.5-(3) ci dice che $V = W \oplus W^\perp$.

Lo spazio W^\perp ha dimensione $n - 1$ e quindi possiamo applicare l'ipotesi induttiva al prodotto scalare $g|_{W^\perp}$. Per ipotesi induttiva esiste una base ortogonale v_2, \dots, v_n per W^\perp . Aggiungendo v alla sequenza otteniamo una base $\mathcal{B} = \{v, v_2, \dots, v_n\}$ per V che è ancora ortogonale, infatti $v_i \in W^\perp$ implica $g(v, v_i) = 0$ per ogni i .

Resta da considerare il caso in cui tutti i vettori in V siano isotropi, cioè $g(v, v) = 0$ per ogni $v \in V$. La facile uguaglianza

$$g(v, w) = \frac{g(v + w, v + w) - g(v, v) - g(w, w)}{2} = \frac{0 - 0 - 0}{2} = 0$$

ci mostra che $g(v, w) = 0$ per ogni $v, w \in V$ e quindi qualsiasi base di V è ortogonale. \square

Ricordiamo che se \mathcal{B} è ortogonale allora la matrice associata a g nella base \mathcal{B} è diagonale.

Corollario 7.4.5. *Qualsiasi matrice simmetrica S è congruente ad una matrice diagonale.*

Dimostrazione. Interpretiamo S come un prodotto scalare g in \mathbb{R}^n . Per il Teorema di Sylvester esiste una base ortogonale \mathcal{B} , quindi $S' = [g]_{\mathcal{B}}$ è diagonale. Un cambiamento di base trasforma la matrice per congruenza: quindi S è congruente alla matrice diagonale S' . \square

7.4.3. Normalizzazione. Descriviamo adesso un semplice algoritmo di *normalizzazione*, che trasforma una base ortogonale in una base ortonormale.

Se $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base ortogonale per un certo prodotto scalare, possiamo sostituire ciascun vettore v_i con un nuovo vettore w_i definito nel modo seguente:

$$w_i = \begin{cases} \frac{v_i}{\sqrt{\langle v_i, v_i \rangle}} & \text{se } \langle v_i, v_i \rangle > 0, \\ v_i & \text{se } \langle v_i, v_i \rangle = 0, \\ \frac{v_i}{\sqrt{-\langle v_i, v_i \rangle}} & \text{se } \langle v_i, v_i \rangle < 0. \end{cases}$$

Il risultato è una nuova base $\{w_1, \dots, w_n\}$ in cui il numero $\langle w_i, w_i \rangle$ può essere solo 1, -1 , oppure 0, per ogni i . Il lettore è invitato a verificare questo fatto.

Notiamo inoltre che possiamo anche riordinare gli elementi di questa base in modo che prima ci siano i vettori w_i con $\langle w_i, w_i \rangle = 1$, poi quelli con $\langle w_i, w_i \rangle =$

-1 , ed infine quelli con $\langle w_i, w_i \rangle = 0$. La matrice associata $S = [g]_{\mathcal{B}}$ a questo punto diventa del tipo

$$S = \begin{pmatrix} I_k & 0 & 0 \\ 0 & -I_h & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Riassumiamo questa indagine con un enunciato. Sia come sempre V uno spazio vettoriale di dimensione n e g un prodotto scalare su V .

Proposizione 7.4.6. *Esiste sempre una base ortonormale \mathcal{B} per V tale che la matrice associata $S = [g]_{\mathcal{B}}$ sia del tipo*

$$(8) \quad S = \begin{pmatrix} I_k & 0 & 0 \\ 0 & -I_h & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Dimostrazione. Per il Teorema di Sylvester esiste una base ortogonale, che può quindi essere trasformata in ortonormale tramite la normalizzazione. \square

Definizione 7.4.7. La terna di numeri $(k, h, n - (k + h))$ è detta *segnatura* del prodotto scalare g ed è di solito indicata con i simboli (i_+, i_-, i_0) . I numeri i_+, i_-, i_0 sono chiamati *indice di positività*, *indice di negatività* e *indice di nullità* di g . La somma dei tre indici è $n = \dim V$.

Corollario 7.4.8. *Qualsiasi matrice simmetrica è congruente ad una matrice diagonale S come in (8).*

7.4.4. Classificazione. Nella sezione precedente abbiamo definito la segnatura (i_+, i_-, i_0) di un prodotto scalare, barando un po' verso la fine: dobbiamo ancora mostrare che i tre numeri dipendono solo da g e non dalla particolare base \mathcal{B} scelta per diagonalizzare il prodotto scalare. Lo facciamo con la proposizione seguente, che caratterizza i numeri i_+, i_-, i_0 più intrinsecamente e quindi fornisce un'utile informazione aggiuntiva.

Proposizione 7.4.9. *Valgono i fatti seguenti:*

- $i_+ = \max\{\dim W \mid W \subset V \text{ sottospazio con } g|_W \text{ definito positivo}\}$
- $i_- = \max\{\dim W \mid W \subset V \text{ sottospazio con } g|_W \text{ definito negativo}\}$
- $i_0 = \dim V^\perp$

Dimostrazione. Per la Proposizione 7.4.6 esiste una base ortogonale $\mathcal{B} = \{v_1, \dots, v_n\}$ per cui la matrice associata $S = [g]_{\mathcal{B}}$ è come in (8). Consideriamo i sottospazi

$$W_+ = \text{Span}(v_1, \dots, v_k), \quad W_- = \text{Span}(v_{k+1}, \dots, v_{k+h}), \\ W_0 = \text{Span}(v_{k+h+1}, \dots, v_n).$$

La Proposizione 7.1.20 ci dice che W_0 è proprio il radicale, quindi $i_0 = n - (k + h) = \dim V^\perp$.

Siano k_0 e h_0 le massime dimensioni possibili di un sottospazio W su cui la restrizione $g|_W$ sia rispettivamente definita positiva e negativa. Dobbiamo mostrare che $k_0 = k$ e $h_0 = h$.

Le restrizioni $g|_{W_+}$ e $g|_{W_-}$ sono definite positive e negative rispettivamente, quindi $k_0 \geq k$ e $h_0 \geq h$. Supponiamo per assurdo che esista un'altro sottospazio W di dimensione $k' > k$ su cui la restrizione $g|_W$ è definita positiva. Notiamo che

$$\dim W + \dim(W_- \oplus W_0) = k' + (n - k) > n$$

e quindi per la formula di Grassman l'intersezione $W \cap (W_- \oplus W_0)$ ha dimensione almeno uno. Quindi esiste un $v \neq 0$ con $v \in W$ (e quindi $\langle v, v \rangle > 0$) e $v \in W_- \oplus W_0$ (e quindi $\langle v, v \rangle \leq 0$): assurdo. Il caso definito negativo è analogo. \square

La segnatura (i_+, i_-, i_0) classifica completamente i prodotti scalari. Molte delle definizioni che abbiamo dato possono essere reinterpretate con la segnatura:

- g è definito positivo $\iff (i_+, i_-, i_0) = (n, 0, 0)$;
- g è definito negativo $\iff (i_+, i_-, i_0) = (0, n, 0)$;
- g è degenere $\iff i_0 > 0$;
- g è semi-definito positivo $\iff i_- = 0$;
- g è semi-definito negativo $\iff i_+ = 0$;
- g è indefinito $\iff i_+ > 0$ e $i_- > 0$.

Il lettore è invitato a dimostrare tutte queste facili equivalenze. Definiamo la segnatura (i_+, i_-, i_0) di una matrice simmetrica S come la segnatura del prodotto scalare $\langle x, y \rangle = {}^t x S y$ su \mathbb{R}^n . Otteniamo in particolare una classificazione completa delle matrici simmetriche a meno di congruenza:

Corollario 7.4.10. *Due matrici simmetriche sono congruenti se e solo se hanno la stessa segnatura.*

Calcolare la segnatura di una matrice diagonale è immediato.

Esercizio 7.4.11. Sia D una matrice diagonale e (i_+, i_-, i_0) la sua segnatura. I numeri i_+, i_-, i_0 sono pari al numero di elementi positivi, negativi e nulli sulla diagonale di D .

Per una matrice simmetrica S non diagonale la questione è più complessa.

7.4.5. Calcolo della segnatura. Abbiamo dimostrato che la segnatura (i_+, i_-, i_0) caratterizza un prodotto scalare, quindi è naturale adesso porsi il problema seguente: come si calcola la segnatura di un prodotto scalare dato?

Un metodo consiste nell'applicare la dimostrazione della Proposizione 7.4.4 per costruire una base ortogonale, scrivere la matrice diagonale associata, quindi contare quanti elementi sulla diagonale sono positivi, negativi, o nulli. Il problema con questo approccio è che la costruzione della base ortogonale

può essere laboriosa. Cerchiamo quindi un modo di determinare la segnatura senza dover costruire una base ortogonale.

In seguito descriveremo alcuni metodi generali per risolvere questo problema, per adesso ci accontentiamo di fare alcune considerazioni. Sia g un prodotto scalare su V e S la matrice associata a g rispetto ad una base $\mathcal{B} = \{v_1, \dots, v_n\}$ qualsiasi. Abbiamo visto nella Proposizione 7.2.13 che il segno di $\det S$ non dipende dalla base \mathcal{B} . Poiché S è congruente ad una matrice diagonale del tipo (8), deduciamo che:

- $\det S = 0 \iff i_0 > 0$;
- se $\det S > 0$, allora i_- è pari,
- se $\det S < 0$, allora i_- è dispari.

Notiamo anche un'altra cosa. Se $W \subset V$ è un sottospazio su cui la restrizione $g|_W$ è definita positiva, allora $i_+ \geq \dim W$ per la Proposizione 7.4.9. In particolare, se $S_{ii} > 0$ per qualche i , allora $\langle v_i, v_i \rangle = S_{ii} > 0$ e quindi la restrizione $g|_{\text{Span}(v_i)}$ sulla retta $\text{Span}(v_i)$ è definita positiva: ne segue che $i_+ > 0$. Riassumendo:

- se esiste i con $S_{ii} > 0$, allora $i_+ > 0$;
- se esiste i con $S_{ii} < 0$, allora $i_- > 0$.

Queste informazioni sono sufficienti per trattare il caso bidimensionale. Consideriamo il prodotto scalare su \mathbb{R}^2 definito come $\langle x, y \rangle = {}^t x S y$ per una matrice simmetrica

$$S = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

Notiamo che $\det S = ac - b^2$. Questa è la casistica:

- se $S = 0$, la segnatura è $(0, 0, 2)$;
- se $\det S = 0$ e $a > 0$ oppure $c > 0$, la segnatura è $(1, 0, 1)$;
- se $\det S = 0$ e $a < 0$ oppure $c < 0$, la segnatura è $(0, 1, 1)$;
- se $\det S > 0$ e $a > 0$ oppure $c > 0$, la segnatura è $(2, 0, 0)$;
- se $\det S > 0$ e $a < 0$ oppure $c < 0$, la segnatura è $(0, 2, 0)$;
- se $\det S < 0$, la segnatura è $(1, 1, 0)$.

Il lettore è invitato a dimostrare i vari casi usando la Proposizione 7.4.9.

Osservazione 7.4.12. Attenzione: un elemento positivo sulla diagonale di S implica $i_+ \geq 1$, ma due elementi positivi non implicano $i_+ \geq 2$. Ad esempio la segnatura di $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ è $(1, 1, 0)$.

Esempio 7.4.13. Calcoliamo la segnatura della matrice

$$S = \begin{pmatrix} 2 & 2 & 1 & 1 \\ 2 & 2 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

Notiamo che S ha rango due e quindi $i_0 = 2$. Le segnature possibili sono:

$$(2, 0, 2), \quad (1, 1, 2), \quad (0, 2, 2).$$

Ci sono elementi positivi sulla diagonale, quindi $i_+ > 0$ e la terza possibilità è esclusa. Ci resta da capire quale delle prime due segnature sia quella giusta. Se consideriamo il piano $W = \text{Span}(e_2, e_3)$, notiamo che la matrice associata alla restrizione $g|_W$ è la sottomatrice quadrata

$$\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}.$$

Questa ha determinante negativo: quindi la segnatura di $g|_W$ è $(1, 1, 0)$. Ne deduciamo che W contiene una retta su cui g è definito positivo e una su cui g è definito negativo. Quindi $i_- \geq 1$ e l'unica segnatura possibile è $(1, 1, 2)$.

7.4.6. Criterio di Jacobi. Introduciamo un criterio semplice per determinare se una matrice $n \times n$ simmetrica S è definita positiva. Per ogni $i = 1, \dots, n$ sia d_i il determinante del minore S_i di S ottenuto considerando solo le prime i righe e le prime i colonne di S .

Lemma 7.4.14 (Criterio di Jacobi). *Una matrice S è definita positiva se e solo se $d_i > 0$ per ogni i .*

Dimostrazione. La matrice S definisce un prodotto scalare $g(x, y) = {}^t x S y$ su \mathbb{R}^n e S_i rappresenta la restrizione di g sul sottospazio coordinato $V_i = \text{Span}(e_1, \dots, e_i)$. Se S è definita positiva, anche $g|_{V_i}$ lo è, ed in particolare deve valere $d_i = \det S_i > 0$ per ogni i .

Mostriamo ora l'implicazione inversa: supponiamo che $d_i > 0$ per ogni i e deduciamo che g è definito positivo. Dimostriamo il lemma per induzione su n . Se $n = 1$, la matrice è un numero positivo d_1 e siamo a posto. Supponiamo il lemma vero per $n - 1$ e lo dimostriamo per n .

Per l'ipotesi induttiva la restrizione $g|_{V_{n-1}}$ è definita positiva. Consideriamo la segnatura (i_+, i_-, i_0) di g . Siccome la restrizione $g|_{V_{n-1}}$ è definita positiva, abbiamo $i_+ \geq n - 1$ e le possibilità per la segnatura sono

$$(n, 0, 0), \quad (n - 1, 1, 0), \quad (n - 1, 0, 1).$$

Siccome $d_n > 0$, l'unica possibile è la prima. □

Corollario 7.4.15. *Una matrice S è definita negativa se e solo se $d_i < 0$ per ogni i dispari e $d_i > 0$ per ogni i pari.*

Dimostrazione. La matrice S è definita negativa se e solo se $-S$ è definita positiva. A questo punto basta notare che $\det(-A) = (-1)^k \det A$ per ogni matrice A di taglia $k \times k$. □

Con una dimostrazione analoga si dimostra anche questo criterio:

Lemma 7.4.16. Se $d_i \neq 0$ per ogni i , allora la segnatura di S è $(i_+, i_-, 0)$ dove i_+ è il numero di permanenze di segno nella successione $1, d_1, \dots, d_n$ e i_- il numero di cambiamenti di segno.

Esercizio 7.4.17. Verifica il criterio appena enunciato quando S è una matrice diagonale invertibile (cioè senza zeri sulla diagonale).

Esempio 7.4.18. Studiamo la segnatura della matrice

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

al variare di $a \in \mathbb{R}$. Otteniamo

$$d_1 = 1, \quad d_2 = 2, \quad d_3 = 6 - 2a^2.$$

Se $|a| > \sqrt{3}$ allora $d_3 < 0$ e quindi la segnatura è $(2, 1, 0)$. Se $|a| < \sqrt{3}$ allora $d_3 > 0$ e quindi la segnatura è $(3, 0, 0)$. Se $a = \pm\sqrt{3}$ il criterio di Jacobi non si applica e dobbiamo usare altri metodi. La matrice S ha rango 2 e quindi $i_0 = 3 - 2 = 1$. D'altra parte la restrizione di g al piano coordinato $W = \text{Span}(e_1, e_2)$ ha come matrice associata $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ che è definita positiva: quindi $g|_W$ è definita positiva e $i_+ \geq 2$ per la Proposizione 7.4.9. L'unico caso possibile è $(2, 0, 1)$. Riassumendo, la segnatura è

- $(3, 0, 0)$ per $|a| < \sqrt{3}$,
- $(2, 0, 1)$ per $|a| = \sqrt{3}$,
- $(2, 1, 0)$ per $|a| > \sqrt{3}$.

7.5. Prodotti scalari definiti positivi

Nella geometria euclidea che si studia alle scuole superiori, il prodotto scalare fra due vettori non è una nozione "di base" e viene generalmente definita usando angoli e lunghezze: il *prodotto scalare* di due vettori v e w che formano un angolo acuto θ come nella Figura 7.2 è il numero

$$\langle v, w \rangle = |v| \cdot |w| \cos \theta.$$

In questa formula $|v|$ e $|w|$ sono le lunghezze di v e w . La Figura 7.2 mostra una interpretazione geometrica del prodotto scalare: se proiettiamo v ortogonalmente lungo w , otteniamo un segmento s lungo $|v| \cos \theta$. Quindi il prodotto scalare fra v e w è il prodotto delle lunghezze di s e di w .

Nella geometria che si insegna in ambito universitario si preferisce scegliere la strada opposta: si introduce prima il prodotto scalare (come abbiamo fatto nella sezione precedente) e poi si introducono angoli, lunghezze e distanze.

In tutto questo capitolo V sarà sempre uno spazio vettoriale reale dotato di un prodotto scalare definito positivo.

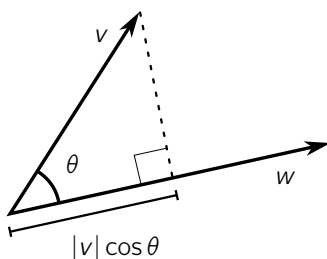


Figura 7.2. Una interpretazione geometrica del prodotto scalare fra due vettori v e w .

7.5.1. Norma. In queste pagine introduciamo i concetti di *norma* (o *lunghezza*) di un vettore, *angolo* fra vettori, e *distanza* fra punti.

Definizione 7.5.1. La *norma* di un vettore $v \in V$ è il numero reale

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

La norma di v , detta anche *modulo*, va interpretata come la *lunghezza* del vettore v . Nella definizione usiamo il fatto che il prodotto scalare è definito positivo e quindi ha senso la radice quadrata di $\langle v, v \rangle \geq 0$.

Proposizione 7.5.2. Valgono le proprietà seguenti, per ogni $v, w \in V$ e ogni $\lambda \in \mathbb{R}$:

- (1) $\|v\| > 0$ se $v \neq 0$ e $\|0\| = 0$;
- (2) $\|\lambda v\| = |\lambda| \|v\|$;
- (3) $|\langle v, w \rangle| \leq \|v\| \|w\|$;
- (4) $\|v + w\| \leq \|v\| + \|w\|$.

La disuguaglianza (3) è nota come *disuguaglianza di Cauchy-Schwarz*.

Dimostrazione. I primi due punti seguono semplicemente dalla definizione $\|v\| = \sqrt{\langle v, v \rangle}$. Il punto (1) è verificato perché il prodotto scalare è definito positivo, il (2) si ottiene così:

$$\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v \rangle} = \sqrt{\lambda^2 \langle v, v \rangle} = |\lambda| \sqrt{\langle v, v \rangle} = |\lambda| \|v\|.$$

Il punto (3) invece non è ovvio. Consideriamo due numeri $a, b \in \mathbb{R}$ e notiamo che per ogni $v, w \in V$ vale la disuguaglianza

$$\begin{aligned} 0 &\leq \|av + bw\|^2 = \langle av + bw, av + bw \rangle \\ &= a^2 \langle v, v \rangle + b^2 \langle w, w \rangle + 2ab \langle v, w \rangle = a^2 \|v\|^2 + b^2 \|w\|^2 + 2ab \langle v, w \rangle. \end{aligned}$$

Sostituendo $a = \|w\|^2$ e $b = -\langle v, w \rangle$ otteniamo

$$0 \leq \|w\|^4 \|v\|^2 + \langle v, w \rangle^2 \|w\|^2 - 2\|w\|^2 \langle v, w \rangle^2.$$

Dividendo per $\|w\|^2$ si ottiene

$$\|w\|^2 \|v\|^2 \geq \langle v, w \rangle^2$$

e la (3) è ottenuta prendendo le radici quadrate. La disuguaglianza (4) è un semplice corollario della (3):

$$\|v+w\|^2 = \|v\|^2 + \|w\|^2 + 2\langle v, w \rangle \leq \|v\|^2 + \|w\|^2 + 2\|v\|\|w\| = (\|v\| + \|w\|)^2.$$

Anche qui basta prendere le radici per concludere. \square

Esercizio 7.5.3. Per ogni $v, w \in V$ otteniamo

$$\langle v, w \rangle = \frac{1}{4}(\|v+w\|^2 - \|v-w\|^2).$$

7.5.2. Angoli. A questo punto possiamo definire la nozione di angolo fra vettori non nulli.

Definizione 7.5.4. L'*angolo* fra due vettori $v, w \in V$ non nulli è il numero $\theta \in [0, \pi]$ per cui

$$\cos \theta = \frac{\langle v, w \rangle}{\|v\|\|w\|}.$$

Effettivamente per la disuguaglianza di Cauchy-Scharz il membro di destra è un numero $x \in [-1, 1]$ e quindi esiste un unico $\theta \in [0, \pi]$ per cui $\cos \theta = x$. In altre parole, si può usare la funzione *arcocoseno* e scrivere

$$\theta = \arccos \frac{\langle v, w \rangle}{\|v\|\|w\|}.$$

Osservazione 7.5.5. Tutte le funzioni trigonometriche come $\sin(x)$, $\cos(x)$, $\tan(x)$, $\arcsin(x)$, $\arccos(x)$, ecc. possono essere definite intrinsecamente come serie di potenze.

Notiamo che θ è acuto, retto o ottuso \iff il prodotto scalare $\langle v, w \rangle$ è rispettivamente positivo, nullo o negativo.

7.5.3. Distanze. Sia come sempre V uno spazio vettoriale reale dotato di un prodotto scalare definito positivo. Possiamo infine definire la distanza fra due punti di V .

Quando interpretiamo V in modo più geometrico, può essere utile indicare i punti di V con le lettere maiuscole P, Q, R , eccetera. Possiamo anche definire il vettore

$$\overrightarrow{PQ} = Q - P.$$

Notiamo in particolare che

$$\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}.$$

Definizione 7.5.6. La *distanza* fra due punti $p, q \in V$ è il numero reale

$$d(P, Q) = \|\overrightarrow{PQ}\| = \|Q - P\|.$$

La distanza ha tre proprietà basilari.

Proposizione 7.5.7. Valgono i fatti seguenti per ogni $P, Q, R \in V$:

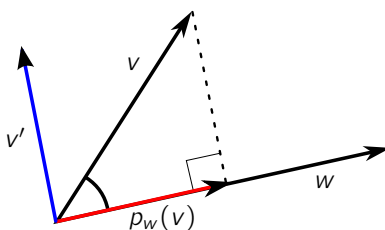


Figura 7.3. Il vettore v si decompone nella sua proiezione ortogonale $p_w(v)$ su w , che è parallela a w , e in un vettore v' che è ortogonale a w .

- (1) $d(P, Q) > 0$ se $P \neq Q$ e $d(P, P) = 0$;
- (2) $d(P, Q) = d(Q, P)$;
- (3) $d(P, R) \leq d(P, Q) + d(Q, R)$.

Dimostrazione. Le proprietà (1) e (2) sono facili conseguenze delle proprietà della norma. Per la (3) abbiamo

$$d(P, R) = \|\vec{PR}\| = \|\vec{PQ} + \vec{QR}\| \leq \|\vec{PQ}\| + \|\vec{QR}\| = d(P, Q) + d(Q, R).$$

La dimostrazione è completa. \square

La proprietà (3) è la *disuguaglianza triangolare*: nel triangolo con vertici P, Q, R , la lunghezza di ogni lato è minore della somma delle lunghezze degli altri due.

7.5.4. Proiezione ortogonale. Una nozione importante associata ad un prodotto scalare è quella di *ortogonalità*. Mostriamo adesso come sia possibile proiettare ortogonalmente vettori su altri vettori.

Sia w un vettore non nullo di V . Sia $v \in V$ un vettore qualsiasi. Definiamo la *proiezione di v su w* il vettore

$$p_w(v) = \frac{\langle v, w \rangle}{\langle w, w \rangle} w.$$

Si veda la Figura 7.3. Notiamo subito che $p_w(v)$ è ottenuto moltiplicando il vettore v per il *coefficiente di Fourier*

$$\frac{\langle v, w \rangle}{\langle w, w \rangle}.$$

Questo coefficiente comparirà in vari punti nelle pagine seguenti. La proprietà più importante della proiezione è la seguente: se indichiamo con v' il vettore

$$v' = v - p_w(v)$$

notiamo che v' è ortogonale a w come mostrato nella Figura 7.3. Infatti:

$$\begin{aligned} \langle v', w \rangle &= \langle v - p_w(v), w \rangle = \langle v - \frac{\langle v, w \rangle}{\langle w, w \rangle} w, w \rangle = \langle v, w \rangle - \frac{\langle v, w \rangle}{\langle w, w \rangle} \langle w, w \rangle \\ &= \langle v, w \rangle - \langle v, w \rangle = 0. \end{aligned}$$

Effettivamente il coefficiente di Fourier è stato introdotto proprio con questo scopo: se cerchiamo un $k \in \mathbb{R}$ tale che $v' = v - kw$ sia ortogonale a w , vediamo facilmente che l'unico k che soddisfa questa richiesta è il coefficiente di Fourier (il lettore è invitato a verificare questo fatto).

7.5.5. Componente parallela e ortogonale ad un vettore. Sia come sopra $w \in V$ un vettore non nullo fissato. Abbiamo scoperto che possiamo decomporre come nella Figura 7.3 ciascun vettore $v \in V$ in due componenti:

$$v = p_w(v) + v'$$

dove $p_w(v)$ è la componente *parallela* a w e v' è quella *ortogonale* a w .

Esempio 7.5.8. Consideriamo \mathbb{R}^2 con il prodotto scalare euclideo. Se $w = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ e $v = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ otteniamo

$$p_w(v) = \frac{\langle v, w \rangle}{\langle w, w \rangle} w = 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

e quindi decomponiamo il vettore v in $v = p_w(v) + v'$ nel modo seguente:

$$v = \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \end{pmatrix}.$$

Queste sono le componenti parallela e ortogonale a w . Facciamo un altro esempio: se $w = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ e $v = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, otteniamo

$$p_w(v) = \frac{\langle v, w \rangle}{\langle w, w \rangle} w = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

e quindi decomponiamo il vettore v in $v = p_w(v) + v'$ nel modo seguente:

$$v = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

7.5.6. Coefficienti di Fourier. Mostriamo adesso come i coefficienti di Fourier forniscano immediatamente le coordinate di un vettore rispetto ad una base ortogonale.

Sia come sempre V uno spazio vettoriale dotato di un prodotto scalare definito positivo. Sia inoltre $\mathcal{B} = \{v_1, \dots, v_n\}$ una base ortogonale.

Proposizione 7.5.9. *Per qualsiasi vettore $v \in V$ vale la relazione seguente:*

$$v = p_{v_1}(v) + \dots + p_{v_n}(v)$$

In altre parole, ciascun vettore v è somma delle sue proiezioni ortogonali sugli elementi della base. Usando i coefficienti di Fourier scriviamo:

$$(9) \quad v = \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 + \dots + \frac{\langle v, v_n \rangle}{\langle v_n, v_n \rangle} v_n.$$

Dimostrazione. Dimostriamo (9) mostrando che il vettore differenza

$$v' = v - \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \dots - \frac{\langle v, v_n \rangle}{\langle v_n, v_n \rangle} v_n$$

è nullo. Per fare ciò, calcoliamo il prodotto scalare $\langle v', v_i \rangle$ di v' con un elemento qualsiasi della base v_i e troviamo:

$$\langle v', v_i \rangle = \langle v, v_i \rangle - \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} \langle v_1, v_i \rangle - \dots - \frac{\langle v, v_n \rangle}{\langle v_n, v_n \rangle} \langle v_n, v_i \rangle.$$

Ricordiamo che i vettori v_1, \dots, v_n sono ortogonali e quindi $\langle v_i, v_j \rangle = 0$ se $i \neq j$. Quindi nel membro di destra molti addendi si annullano e resta

$$\langle v', v_i \rangle = \langle v, v_i \rangle - \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle} \langle v_i, v_i \rangle = \langle v, v_i \rangle - \langle v, v_i \rangle = 0.$$

Il vettore v' è ortogonale a tutti gli elementi v_1, \dots, v_n della base, quindi a tutti gli elementi di V . Poiché il prodotto scalare è definito positivo (ed in particolare non è degenere) otteniamo necessariamente $v' = 0$. \square

L'equazione (9) assume una forma particolarmente semplice se \mathcal{B} è una base ortonormale; in questo caso $\langle v_i, v_i \rangle = 1$ per ogni i e l'equazione diventa:

$$v = \langle v, v_1 \rangle v_1 + \dots + \langle v, v_n \rangle v_n.$$

La Proposizione 7.5.9 può essere enunciata anche nel modo seguente:

Corollario 7.5.10. *Le coordinate di un vettore v rispetto ad una base ortogonale $\mathcal{B} = \{v_1, \dots, v_n\}$ sono i coefficienti di Fourier. Scriviamo quindi:*

$$[v]_{\mathcal{B}} = \begin{pmatrix} \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} \\ \vdots \\ \frac{\langle v, v_n \rangle}{\langle v_n, v_n \rangle} \end{pmatrix}.$$

Se la base \mathcal{B} è ortonormale, scriviamo

$$[v]_{\mathcal{B}} = \begin{pmatrix} \langle v, v_1 \rangle \\ \vdots \\ \langle v, v_n \rangle \end{pmatrix}.$$

Esempio 7.5.11. Consideriamo \mathbb{R}^n con il prodotto scalare euclideo. La base canonica $\mathcal{C} = \{e_1, \dots, e_n\}$ è ortonormale. Se $x \in \mathbb{R}^n$ è un vettore qualsiasi, il coefficiente di Fourier $\langle x, e_i \rangle$ di x rispetto a \mathcal{C} è semplicemente la sua coordinata x_i . Quindi con il Corollario 7.5.10 ritroviamo semplicemente una uguaglianza che già conoscevamo:

$$[x]_{\mathcal{C}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Esempio 7.5.12. Consideriamo \mathbb{R}^2 con il prodotto scalare euclideo e prendiamo la base ortogonale

$$\mathcal{B} = \left\{ v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}.$$

Prendiamo il vettore $v = \begin{pmatrix} 7 \\ 2 \end{pmatrix}$ e verifichiamo che

$$v = \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 + \frac{\langle v, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 = \frac{9}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{5}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 7 \\ 2 \end{pmatrix}.$$

Quindi abbiamo calcolato le coordinate di $\begin{pmatrix} 7 \\ 2 \end{pmatrix}$ rispetto alla base \mathcal{B} :

$$\left[\begin{pmatrix} 7 \\ 2 \end{pmatrix} \right]_{\mathcal{B}} = \begin{pmatrix} 9/2 \\ 5/2 \end{pmatrix}.$$

L'aspetto fondamentale di questa costruzione è che le coordinate di un vettore rispetto ad una base ortogonale possono essere calcolate direttamente, senza dover impostare e risolvere un sistema lineare. Questa costruzione funziona ovviamente solo se la base v_1, \dots, v_n è ortogonale.

7.5.7. Ortogonalizzazione di Gram-Schmidt. Abbiamo appena scoperto che le basi ortogonali sono molto utili, perché le coordinate di qualsiasi vettore si calcolano facilmente tramite i coefficienti di Fourier. Ci poniamo quindi il seguente problema: come facciamo a costruire una base ortogonale?

Descriviamo adesso un algoritmo che prende come *input* dei vettori indipendenti v_1, \dots, v_k e restituisce come *output* dei vettori w_1, \dots, w_k indipendenti e ortogonali. Questo algoritmo è noto come *ortogonalizzazione di Gram-Schmidt*.

L'algoritmo è il seguente. Si costruiscono i vettori w_1, \dots, w_k induttivamente in questo modo:

$$\begin{aligned} w_1 &= v_1, \\ w_2 &= v_2 - p_{w_1}(v_2), \\ w_3 &= v_3 - p_{w_1}(v_3) - p_{w_2}(v_3), \\ &\vdots \\ w_k &= v_k - p_{w_1}(v_k) - \dots - p_{w_{k-1}}(v_k). \end{aligned}$$

In altre parole, ad ogni passo togliamo al vettore v_i le componenti parallele ai vettori w_1, \dots, w_{i-1} già trovati precedentemente, in modo che il nuovo

vettore w_i sia ortogonale a questi. Esplicitando i coefficienti di Fourier:

$$\begin{aligned} w_1 &= v_1, \\ w_2 &= v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1, \\ w_3 &= v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2, \\ &\vdots \\ w_k &= v_k - \frac{\langle v_k, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \dots - \frac{\langle v_k, w_{k-1} \rangle}{\langle w_{k-1}, w_{k-1} \rangle} w_{k-1}. \end{aligned}$$

Dimostriamo adesso che l'algoritmo funziona.

Proposizione 7.5.13. L'algoritmo di Gram-Schmidt produce effettivamente dei vettori w_1, \dots, w_k indipendenti e ortogonali. Inoltre per ogni $j = 1, \dots, k$, i primi j vettori della base generano gli stessi spazi di prima:

$$(10) \quad \text{Span}(v_1, \dots, v_j) = \text{Span}(w_1, \dots, w_j).$$

Dimostrazione. Dimostriamo tutto per induzione su j . Il caso $j = 1$ è banale, quindi esaminiamo il passo induttivo.

Il vettore w_j è costruito togliendo a v_j le componenti parallele ai vettori precedenti w_1, \dots, w_{j-1} e quindi ciò che resta è ortogonale a questi. Più formalmente, per ogni $i < j$ abbiamo

$$\begin{aligned} \langle w_i, w_j \rangle &= \langle w_i, v_j \rangle - \left\langle w_i, \frac{\langle v_j, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 \right\rangle - \dots - \left\langle w_i, \frac{\langle v_j, w_{j-1} \rangle}{\langle w_{j-1}, w_{j-1} \rangle} w_{j-1} \right\rangle \\ &= \langle w_i, v_j \rangle - \frac{\langle v_j, w_1 \rangle}{\langle w_1, w_1 \rangle} \langle w_i, w_1 \rangle - \dots - \frac{\langle v_j, w_{j-1} \rangle}{\langle w_{j-1}, w_{j-1} \rangle} \langle w_i, w_{j-1} \rangle \\ &= \langle w_i, v_j \rangle - \frac{\langle v_j, w_i \rangle}{\langle w_i, w_i \rangle} \langle w_i, w_i \rangle = 0. \end{aligned}$$

Nella terza uguaglianza abbiamo usato che $\langle w_i, w_l \rangle = 0$ se $i \neq l$. Notiamo inoltre che ciascun w_j è costruito come combinazione lineare dei v_1, \dots, v_j , quindi $\text{Span}(w_1, \dots, w_j) \subset \text{Span}(v_1, \dots, v_j)$. D'altra parte, si può anche scrivere facilmente v_j come combinazione lineare degli w_1, \dots, w_j , quindi anche $\text{Span}(w_1, \dots, w_j) \supset \text{Span}(v_1, \dots, v_j)$ e dimostriamo (10).

In particolare $\dim \text{Span}(w_1, \dots, w_j) = j$ e quindi i nuovi vettori w_1, \dots, w_j sono effettivamente indipendenti. \square

Esempio 7.5.14. Ortogonalizziamo i vettori

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

rispetto al prodotto scalare euclideo di \mathbb{R}^2 . Otteniamo

$$w_1 = v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$w_2 = v_2 - p_{w_1}(v_2) = v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{1}{1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

I vettori ottenuti w_1 e w_2 sono effettivamente ortogonali.

Esempio 7.5.15. Ortogonalizziamo i vettori

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

rispetto al prodotto scalare euclideo di \mathbb{R}^3 . Otteniamo

$$w_1 = v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

$$w_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix},$$

$$w_3 = v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{\frac{5}{2}} \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix}.$$

I vettori ottenuti w_1 , w_2 e w_3 sono effettivamente ortogonali.

7.5.8. Riscaldamento. L'operazione di *riscaldamento* di un vettore $v \neq 0$ consiste nella sostituzione di v con $v' = \lambda v$, per qualche $\lambda \neq 0$. Se w_1, \dots, w_k sono vettori ortogonali, possiamo sostituire ciascun w_i con un suo riscaldamento $w'_i = \lambda_i w_i$, ed i nuovi vettori w'_1, \dots, w'_k sono ancora ortogonali.

Nell'Esempio 7.5.15 può essere utile riscaldare i vettori nel modo seguente:

$$w'_1 = w_1, \quad w'_2 = 2w_2, \quad w'_3 = \frac{3}{2}w_3$$

ed ottenere una nuova base di vettori ortogonali, più semplice da descrivere:

$$w'_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad w'_2 = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}, \quad w'_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

La *normalizzazione* di un vettore $v \neq 0$ è il riscaldamento

$$v' = \frac{v}{\|v\|}.$$

Questa operazione è importante perché

$$\|v'\| = \frac{\|v\|}{\|v\|} = 1.$$

Il vettore normalizzato v' ha sempre norma uno. Per trasformare una base ortogonale in una ortonormale è sufficiente normalizzare i vettori. Questa operazione ha spesso l'inconveniente di introdurre delle fastidiose radici quadrate.

Esempio 7.5.16. Normalizzando la base ortogonale trovata nell'Esempio 7.5.15 otteniamo la base ortonormale seguente:

$$w'_1 = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{pmatrix}, \quad w'_2 = \begin{pmatrix} -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{3} \end{pmatrix}, \quad w'_3 = \begin{pmatrix} \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \end{pmatrix}.$$

Si verifica facilmente che hanno tutti norma uno.

7.5.9. Ortogonalità. Lavorare con vettori ortogonali ha molti vantaggi. Uno di questi è che l'indipendenza lineare è garantita a priori:

Proposizione 7.5.17. *Se $v_1, \dots, v_k \in V$ sono tutti ortogonali e non nulli, allora sono indipendenti.*

Dimostrazione. Supponiamo esista una combinazione lineare nulla

$$0 = \lambda_1 v_1 + \dots + \lambda_k v_k.$$

Facendo il prodotto scalare di entrambi i membri con v_i otteniamo

$$0 = \langle 0, v_i \rangle = \langle \lambda_1 v_1 + \dots + \lambda_k v_k, v_i \rangle = \lambda_i \langle v_i, v_i \rangle$$

perché $\langle v_j, v_i \rangle = 0$ se $j \neq i$. A questo punto $\langle v_i, v_i \rangle > 0$ implica $\lambda_i = 0$. Abbiamo scoperto che $\lambda_i = 0 \forall i$ e questo conclude la dimostrazione. \square

7.5.10. Proiezioni su sottospazi. Sia V uno spazio vettoriale con prodotto scalare definito positivo e $U \subset V$ un sottospazio. La *proiezione ortogonale*

$$p_U: V \longrightarrow V$$

è un'applicazione lineare definita nel modo seguente: poiché $V = U \oplus U^\perp$, ogni vettore $v \in V$ si scrive in modo unico come $v = u + u'$ con $u \in U$ e $u' \in U^\perp$. Definiamo quindi

$$p_U(v) = u.$$

Le componenti u e u' di v sono per definizione le proiezioni ortogonali del vettore v sui sottospazi ortogonali U e U^\perp .

Mostriamo come determinare concretamente $p_U(v)$. Entrano in gioco ancora una volta i coefficienti di Fourier.

Proposizione 7.5.18. *Sia v_1, \dots, v_k una base ortogonale per U . Allora*

$$p_U(v) = \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 + \dots + \frac{\langle v, v_k \rangle}{\langle v_k, v_k \rangle} v_k.$$

Dimostrazione. Sia v_{k+1}, \dots, v_n una base ortogonale per U^\perp . Quindi v_1, \dots, v_n per V è una base ortogonale per V . Sappiamo che

$$v = \underbrace{\frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 + \dots + \frac{\langle v, v_k \rangle}{\langle v_k, v_k \rangle} v_k}_u + \underbrace{\frac{\langle v, v_{k+1} \rangle}{\langle v_{k+1}, v_{k+1} \rangle} v_{k+1} + \dots + \frac{\langle v, v_n \rangle}{\langle v_n, v_n \rangle} v_n}_{u'}.$$

Abbiamo scritto $v = u + u'$ con $u \in U$ e $u' \in U^\perp$. Quindi per definizione di proiezione ortogonale troviamo $u = p_U(v)$. \square

Esempio 7.5.19. Consideriamo $U = \{x + y - z = 0\} \subset \mathbb{R}^3$ con il prodotto scalare euclideo. Cerchiamo una base ortogonale per U e troviamo ad esempio

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

A questo punto determiniamo la proiezione ortogonale su U di un generico vettore di \mathbb{R}^3 usando la proposizione precedente, in questo modo:

$$p_U \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{x + y + 2z}{6} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + \frac{x - y}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2x - y + z \\ -x + 2y + z \\ x + y + 2z \end{pmatrix}$$

Quindi la matrice associata a p_U nella base canonica è

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

Usando questa matrice possiamo calcolare la proiezione ortogonale di qualsiasi vettore di \mathbb{R}^3 .

Segue subito dalla definizione che la proiezione ortogonale p_U su un sottospazio $U \subset V$ manda ogni vettore di U in sé stesso, mentre manda ogni vettore di U^\perp in zero. In particolare otteniamo:

$$\text{Im } p_U = U, \quad \text{ker } p_U = U^\perp.$$

Possiamo scrivere agevolmente p_U rispetto ad una base opportuna:

Proposizione 7.5.20. Se $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di V in cui $\{v_1, \dots, v_k\}$ è base di U e $\{v_{k+1}, \dots, v_n\}$ è base di U^\perp , allora

$$[p_U]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$$

dove I_k è la matrice identità di taglia k .

Dimostrazione. Vediamo che $p_U(v_i) = v_i$ per $i = 1, \dots, k$ e $p_U(v_i) = 0$ per $i = k + 1, \dots, n$, quindi la matrice associata è del tipo richiesto. \square

In particolare abbiamo scoperto che p_U è diagonalizzabile e ha autovalori 0 e 1. Gli autospazi relativi sono precisamente U^\perp e U . Vale inoltre la relazione

$$p_U \circ p_U = p_U.$$

Abbiamo già incontrato delle trasformazioni che soddisfano questa relazione nella Proposizione 6.3.17.

Esempio 7.5.21. Sia $U = \{x + y + z = 0\}$ un piano in \mathbb{R}^3 . La base

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

soddisfa le ipotesi della Proposizione 7.5.20. Quindi p_U in questa base è

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

Nella base canonica la proiezione è rappresentata dalla matrice $B = M^{-1}AM$ con $M = [\text{id}]_{\mathcal{B}}^{\mathcal{C}}$. Facendo i conti otteniamo

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

Il lettore è invitato a controllare che usando la Proposizione 7.5.18 come abbiamo fatto nell'Esempio 7.5.19 otteniamo la stessa matrice.

Concludiamo questa sezione con una osservazione che può tornare utile. Può capitare in molti casi che sia più agevole calcolare la proiezione p_{U^\perp} sullo spazio ortogonale U^\perp invece di p_U . Si può quindi calcolare prima p_{U^\perp} e quindi ritrovare p_U scrivendo semplicemente

$$p_U(v) = v - p_{U^\perp}(v).$$

Esempio 7.5.22. Studiamo l'Esempio 7.5.21 seguendo questo suggerimento. Lo spazio ortogonale U^\perp al piano U è una retta, generata dal vettore

$$w = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Ne segue che $p_U(v) = v - p_{U^\perp}(v) = v - p_w(v)$. Quindi

$$p_U \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \frac{x+y+z}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2x - y - z \\ -x + 2y - z \\ -x - y + 2z \end{pmatrix}.$$

Abbiamo ritrovato la stessa matrice associata dell'Esempio 7.5.21.

7.6. Isometrie

Ogni volta che in matematica si definiscono delle strutture come gruppi, spazi vettoriali, prodotti scalari, esistono delle opportune trasformazioni che preservano interamente queste strutture: queste trasformazioni sono chiamate *isomorfismi* per i gruppi e per gli spazi vettoriali, e *isometrie* per i prodotti scalari.

7.6.1. Definizione. Siano V e W due spazi vettoriali dotati ciascuno di un prodotto scalare. Una *isometria* è un isomorfismo $T: V \rightarrow W$ tale che

$$(11) \quad \langle v, w \rangle = \langle T(v), T(w) \rangle \quad \forall v, w \in V.$$

Brevemente, un'isometria è un isomorfismo che preserva il prodotto scalare.

Notiamo subito che l'inversa T^{-1} di un'isometria è anch'essa un'isometria, e la composizione $T \circ T'$ di due isometrie è una isometria.

Letteralmente, il termine "isometria" indica che T preserva la geometria degli oggetti, cioè non li distorce. Quando il prodotto scalare è definito positivo, possiamo definire questa nozione in almeno tre modi equivalenti:

Proposizione 7.6.1. *Sia $T: V \rightarrow W$ un isomorfismo fra spazi dotati di un prodotto scalare definito positivo. I fatti seguenti sono equivalenti:*

- (1) T è una isometria,
- (2) T preserva la norma, cioè $\|T(v)\| = \|v\| \quad \forall v \in V$,
- (3) T preserva la distanza, cioè $d(v, w) = d(T(v), T(w)) \quad \forall v, w \in V$.

Dimostrazione. (1) \Rightarrow (2). Se T è una isometria, allora

$$\|T(v)\|^2 = \langle T(v), T(v) \rangle = \langle v, v \rangle = \|v\|^2.$$

(2) \Rightarrow (3). Vale

$$d(v, w) = \|v - w\| = \|T(v - w)\| = \|T(v) - T(w)\| = d(T(v), T(w)).$$

(3) \Rightarrow (2). Vale

$$\|v\| = d(0, v) = d(T(0), T(v)) = d(0, T(v)) = \|T(v)\|.$$

(2) \Rightarrow (1). Questo è meno facile dei precedenti:

$$\begin{aligned} \langle v, w \rangle &= \frac{\langle v + w, v + w \rangle - \langle v, v \rangle - \langle w, w \rangle}{2} = \frac{\|v + w\|^2 - \|v\|^2 - \|w\|^2}{2} \\ &= \frac{\|T(v + w)\|^2 - \|T(v)\|^2 - \|T(w)\|^2}{2} \\ &= \frac{\langle T(v + w), T(v + w) \rangle - \langle T(v), T(v) \rangle - \langle T(w), T(w) \rangle}{2} \\ &= \langle T(v), T(w) \rangle. \end{aligned}$$

La dimostrazione è completa. □

7.6.2. Con le basi. Vogliamo adesso costruire e studiare concretamente delle isometrie. Per mostrare che un isomorfismo $T: V \rightarrow W$ è una isometria dovremmo in teoria verificare la condizione (11) su ogni coppia $v, w \in V$. Come spesso accade, è sufficiente farlo per gli elementi di una base fissata. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V .

Lemma 7.6.2. *Un isomorfismo T è un'isometria se e solo se*

$$(12) \quad \langle v_i, v_j \rangle = \langle T(v_i), T(v_j) \rangle \quad \forall i, j.$$

Dimostrazione. Chiaramente (11) implica (12), mostriamo adesso il contrario. Per due generici v e w scriviamo

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n, \quad w = \mu_1 v_1 + \dots + \mu_n v_n$$

e quindi

$$\begin{aligned} \langle v, w \rangle &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j \langle v_i, v_j \rangle, \\ \langle T(v), T(w) \rangle &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j \langle T(v_i), T(v_j) \rangle \end{aligned}$$

usando il fatto che il prodotto scalare è bilineare e T è lineare. Quindi (12) implica (11). \square

Consideriamo ad esempio il caso in cui V e W siano dotati di un prodotto scalare definito positivo. Possiamo dedurre il fatto seguente.

Corollario 7.6.3. *Sia v_1, \dots, v_n una base ortonormale per V . Un isomorfismo $T: V \rightarrow W$ è una isometria \iff anche le immagini $T(v_1), \dots, T(v_n)$ formano una base ortonormale per W .*

Possiamo dire quindi che le isometrie sono precisamente gli isomorfismi che mandano basi ortonormali in basi ortonormali.

7.6.3. Con le matrici. Cerchiamo adesso di tradurre la definizione di isometria nel linguaggio concreto delle matrici. Siano V, V' spazi vettoriali dotati di prodotto scalare g, g' e di basi $\mathcal{B}, \mathcal{B}'$ rispettivamente. Consideriamo un isomorfismo $T: V \rightarrow V'$. Siano

$$S = [g]_{\mathcal{B}}, \quad S' = [g']_{\mathcal{B}'}, \quad A = [T]_{\mathcal{B}'}^{\mathcal{B}}$$

le matrici associate.

Proposizione 7.6.4. *L'isomorfismo T è una isometria se e solo se*

$$S = {}^t A S' A.$$

Dimostrazione. Sappiamo che T è una isometria se e solo se $g(v_i, v_j) = g'(T(v_i), T(v_j))$ per ogni i, j . Sappiamo anche che

$$S_{ij} = g(v_i, v_j).$$

Ricordiamo che la colonna i -esima di A è $A^i = [T(v_i)]_{\mathcal{B}'}$. Quindi

$$({}^tAS'A)_{ij} = {}^t(A^i)S'A^j = g'(T(v_i), T(v_j)).$$

Ne deduciamo che T è una isometria se e solo se $S = {}^tAS'A$. \square

In particolare, consideriamo il caso in cui V sia dotato di un prodotto scalare g e $T: V \rightarrow V$ sia un isomorfismo. Siano $S = [g]_{\mathcal{B}}$ e $A = [T]_{\mathcal{B}}$.

Corollario 7.6.5. *L'isomorfismo T è una isometria se e solo se*

$$S = {}^tASA.$$

Consideriamo adesso il caso ancora più specifico:

Corollario 7.6.6. *Sia g definito positivo e \mathcal{B} una base ortonormale. L'endomorfismo T è una isometria se e solo se*

$${}^tAA = I.$$

Dimostrazione. Con queste ipotesi $S = [g]_{\mathcal{B}} = I$, quindi la relazione $S = {}^tASA$ diventa $I = {}^tAA$. \square

7.6.4. Matrici ortogonali. Consideriamo infine il caso che ci sta più a cuore, quello di \mathbb{R}^n munito del suo prodotto scalare euclideo. Sia $A \in M(n)$ una matrice quadrata.

Corollario 7.6.7. *L'endomorfismo L_A è una isometria se e solo se ${}^tAA = I$.*

Dimostrazione. La matrice A è la matrice associata a L_A rispetto alla base canonica, che è ortonormale: quindi applichiamo il Corollario 7.6.6. \square

Una matrice $A \in M(n)$ a coefficienti reali tale che ${}^tAA = I$ è detta *ortogonale*. Studiamo adesso alcune semplici proprietà delle matrici ortogonali.

Proposizione 7.6.8. *Se $A \in M(n)$ è ortogonale, allora*

- $\det A = \pm 1$,
- *i possibili autovalori di A sono ± 1 .*

Dimostrazione. Se ${}^tAA = I$, allora per Binet

$$1 = \det I = \det({}^tA) \det A = (\det A)^2.$$

Quindi $\det A = \pm 1$. Se λ è un autovalore, esiste $v \neq 0$ tale che $Av = \lambda v$. L'isometria L_A preserva la norma e quindi

$$\|v\| = \|Av\| = \|\lambda v\| = |\lambda| \|v\|.$$

Segue che $\lambda = \pm 1$. \square

Dal fatto che $\det A = \pm 1$, deduciamo che una matrice ortogonale A è sempre invertibile; inoltre l'inversa di A è sorprendentemente semplice da determinare: dalla relazione ${}^tAA = I$ deduciamo infatti che $A^{-1} = {}^tA$. In particolare vale anche $A {}^tA = I$.

Proposizione 7.6.9. *Una matrice $A \in M(n)$ è ortogonale \iff i vettori colonna A^1, \dots, A^n formano una base ortonormale di \mathbb{R}^n .*

Dimostrazione. La relazione ${}^tAA = I$ è equivalente alla relazione $\langle A^i, A^j \rangle = \delta_{ij}$ per ogni i, j . In alternativa, questa è una conseguenza del Corollario 7.6.3, visto che i vettori colonna sono le immagini della base canonica (che è ortonormale). \square

7.6.5. Riflessioni. Introduciamo una classe importante di isometrie particolarmente semplici da descrivere. Sia V uno spazio vettoriale dotato di un prodotto scalare definito positivo e sia $U \subset V$ un sottospazio. Ricordiamo che ogni vettore $v \in V$ si decompone nelle proiezioni ortogonali su U e U^\perp :

$$v = p_U(v) + p_{U^\perp}(v).$$

Definizione 7.6.10. La *riflessione* lungo U è l'isometria

$$\text{Rif}_U(v) = v - 2p_{U^\perp}(v) = p_U(v) - p_{U^\perp}(v).$$

Se v_1, \dots, v_k e v_{k+1}, \dots, v_n sono basi ortonormali di U e U^\perp , allora $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base ortonormale di V e la matrice associata a Rif_U nella base \mathcal{B} è molto semplice:

$$[\text{Rif}_U]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{pmatrix}.$$

Questa matrice è chiaramente ortogonale: grazie alla Corollario 7.6.6 abbiamo una conferma che Rif_U sia effettivamente una isometria. Gli autospazi di Rif_U sono

$$V_1 = U, \quad V_{-1} = U^\perp.$$

Notiamo che $\det \text{Rif}_U = (-1)^{n-k} = (-1)^{n-\dim U}$. Quindi una riflessione lungo un iperpiano ha determinante -1 e *inverte* l'orientazione dello spazio. Invece una riflessione di \mathbb{R}^3 lungo una retta preserva l'orientazione.

La riflessione $\text{Rif}_{\{0\}}$ rispetto all'origine $U = \{0\}$ è semplicemente l'endomorfismo $v \mapsto -v$ che cambia di segno tutti i vettori di V . Nota che il determinante di questa applicazione è $(-1)^{\dim V}$. Se $V = \mathbb{R}^2$, la riflessione rispetto all'origine è semplicemente una rotazione di angolo π .

Esempio 7.6.11. Nel caso in cui U sia un iperpiano la riflessione Rif_U è semplice da scrivere. Sia v_0 un generatore della retta U^\perp . Troviamo

$$\text{Rif}_U(v) = v - 2 \frac{\langle v, v_0 \rangle}{\langle v_0, v_0 \rangle} v_0.$$

Nota che questa formula è molto simile a quella che rappresenta la proiezione p_U , con un coefficiente 2 al posto di 1. Ad esempio, se $U \subset \mathbb{R}^3$ è il piano

$$U = \{x + y + z = 0\}$$

otteniamo

$$\text{Rif}_U \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} - 2 \frac{x+y+z}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} x - 2y - 2z \\ -2x + y - 2z \\ -2x - 2y + 1z \end{pmatrix}.$$

Confronta con l'Esempio 7.5.22. La riflessione su U rispetto alla base canonica è rappresentata dalla matrice

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

che è effettivamente ortogonale.

7.6.6. Isometrie del piano. Vogliamo adesso classificare completamente le isometrie del piano \mathbb{R}^2 , dotato del prodotto scalare euclideo. Per quanto appena visto, questo problema è equivalente a classificare le matrici ortogonali 2×2 . Le matrici ortogonali 2×2 possono essere descritte tutte facilmente:

Proposizione 7.6.12. *Le matrici ortogonali in $M(2)$ sono le seguenti:*

$$\text{Rot}_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad \text{Rif}_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}.$$

Dimostrazione. I vettori colonna A^1 e A^2 di una matrice ortogonale A devono formare una base ortonormale. Scriviamo A^1 come un generico vettore unitario

$$A^1 = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}.$$

Il secondo vettore A^2 deve essere ortogonale ad A^1 ed unitario: ci sono solo due possibilità che differiscono solo di un segno, e otteniamo così le matrici Rot_θ e Rif_θ . \square

Notiamo che $\det \text{Rot}_\theta = 1$ e $\det \text{Rif}_\theta = -1$. Cosa rappresentano geometricamente queste isometrie? Troviamo facilmente che queste sono trasformazioni già note. Iniziamo con quelle con determinante positivo.

Proposizione 7.6.13. *L'isometria Rot_θ è una rotazione antioraria di angolo θ intorno all'origine.*

Dimostrazione. Scriviamo il generico vettore in coordinate polari:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \rho \cos \alpha \\ \rho \sin \alpha \end{pmatrix}.$$

Otteniamo quindi

$$\begin{aligned} \text{Rot}_\theta \begin{pmatrix} \rho \cos \alpha \\ \rho \sin \alpha \end{pmatrix} &= \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \rho \cos \alpha \\ \rho \sin \alpha \end{pmatrix} \\ &= \begin{pmatrix} \rho(\cos \theta \cos \alpha - \sin \theta \sin \alpha) \\ \rho(\sin \theta \cos \alpha + \cos \theta \sin \alpha) \end{pmatrix} = \begin{pmatrix} \rho \cos(\alpha + \theta) \\ \rho \sin(\alpha + \theta) \end{pmatrix}. \end{aligned}$$

Ne deduciamo che la trasformazione Rot_θ trasforma il vettore con coordinate polari (ρ, α) nel vettore con coordinate polari $(\rho, \alpha + \theta)$. In altre parole, è una rotazione antioraria di angolo θ intorno all'origine. \square

Passiamo adesso alle isometrie con determinante negativo.

Proposizione 7.6.14. *L'isometria Rif_θ è la riflessione rispetto alla retta passante per l'origine che forma un angolo $\frac{\theta}{2}$ con l'asse x .*

Dimostrazione. Per analizzare Rif_θ , scriviamo il suo polinomio caratteristico $p(\lambda) = \lambda^2 - 1$ e notiamo che Rif_θ ha autovalori ± 1 . Troviamo

$$\begin{aligned} V_1 &= \ker(\text{Rif}_\theta - I) = \ker \begin{pmatrix} \cos \theta - 1 & \sin \theta \\ \sin \theta & -\cos \theta - 1 \end{pmatrix} = \text{Span} \begin{pmatrix} \sin \theta \\ 1 - \cos \theta \end{pmatrix} \\ &= \text{Span} \begin{pmatrix} 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} \\ 1 - \cos^2 \frac{\theta}{2} + \sin^2 \frac{\theta}{2} \end{pmatrix} = \text{Span} \begin{pmatrix} 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} \\ 2 \sin^2 \frac{\theta}{2} \end{pmatrix} = \text{Span} \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix}. \end{aligned}$$

L'altro autospazio V_{-1} si determina in modo analogo e si verifica che V_1 e V_{-1} sono ortogonali. Quindi Rif_θ è una riflessione rispetto alla retta vettoriale V_1 , che forma un angolo $\frac{\theta}{2}$ con l'asse x . \square

7.6.7. Rotazioni nello spazio. Studiamo adesso alcune isometrie importanti dello spazio euclideo \mathbb{R}^3 .

Sia $r \subset \mathbb{R}^3$ una retta e $U = r^\perp$ il suo piano ortogonale. Sia

$$\mathcal{B} = \{v_1, v_2, v_3\}$$

una base ortonormale con $v_1 \in r$ e quindi $v_2, v_3 \in U$. Fissiamo un angolo θ e consideriamo l'endomorfismo T avente matrice associata

$$A = [T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}.$$

Si vede facilmente che ${}^tAA = I$, cioè la matrice A è ortogonale. Siccome \mathcal{B} è una base ortonormale, per il Corollario 7.6.6 l'endomorfismo T è una isometria.

Di che tipo di isometria si tratta? Esaminando la matrice A , si verifica facilmente che $T(r) = r$ e $T(U) = U$. Inoltre $T(v) = v$ per ogni $v \in r$, cioè T fissa tutti i punti della retta r , mentre T ruota il piano U in senso antiorario di un angolo θ . Quindi T è una *rotazione intorno all'asse r di un angolo θ* .

I casi $\theta = 0$ e $\theta = \pi$ sono un po' particolari:

- una rotazione di angolo $\theta = 0$ è l'identità;

- una rotazione di angolo $\theta = \pi$ è una riflessione rispetto alla retta r .

Ovviamente è sempre possibile scrivere T nella base canonica, cambiando la matrice per similitudine nel modo usuale. Se $M = [\text{id}]_{\mathcal{C}}^{\mathcal{B}}$ è la matrice di cambiamento di base, otteniamo

$$A' = [T]_{\mathcal{C}}^{\mathcal{C}} = MAM^{-1}.$$

La nuova matrice A' può essere però molto complicata. Poiché A' rappresenta una isometria rispetto ad una base \mathcal{C} ortonormale, la matrice A' sarà comunque ortogonale. Anche la matrice M è ortogonale, perché le sue colonne sono una base ortonormale: segue che l'inversa $M^{-1} = {}^tM$ si calcola agevolmente facendo la trasposta.

Esempio 7.6.15. Determiniamo una matrice A' che rappresenti (rispetto alla base canonica \mathcal{C}) una rotazione T intorno a

$$r = \text{Span} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

di angolo $\theta = \frac{2\pi}{3}$. Cerchiamo una base ortonormale v_1, v_2, v_3 con $v_1 \in r$. Ad esempio prendiamo questi:

$$v_1 = \frac{\sqrt{3}}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad v_3 = \frac{\sqrt{6}}{6} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

Otteniamo

$$A = [T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & \text{Rot}_{\theta} \end{pmatrix}.$$

Quindi $A' = MA {}^tM$, dove M è la matrice avente come colonne v_1, v_2, v_3 . Quindi A' è pari a

$$\frac{1}{6} \begin{pmatrix} 2\sqrt{3} & 3\sqrt{2} & \sqrt{6} \\ 2\sqrt{3} & -3\sqrt{2} & \sqrt{6} \\ 2\sqrt{3} & 0 & -2\sqrt{6} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 2\sqrt{3} & 2\sqrt{3} & 2\sqrt{3} \\ 3\sqrt{2} & -3\sqrt{2} & 0 \\ \sqrt{6} & \sqrt{6} & -2\sqrt{6} \end{pmatrix} \frac{1}{6}.$$

Svolgendo un po' di conti si trova infine la matrice

$$A' = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Il lettore può anche convincersi con metodi geometrici che effettivamente questa matrice, che permuta la base canonica, rappresenta una rotazione con asse r e angolo $\frac{2\pi}{3}$.

7.6.8. Antirotazioni. Introduciamo adesso alcune isometrie di \mathbb{R}^3 che invertono l'orientazione dello spazio.

Definizione 7.6.16. Una *antirotazione* $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ è la composizione di una rotazione intorno ad un asse r e di una riflessione rispetto al piano $U = r^\perp$.

Scriviamo la matrice associata ad una antirotazione T rispetto ad una base opportuna. Se prendiamo un vettore unitario $v_1 \in r$ e lo completiamo a base ortonormale $\mathcal{B} = \{v_1, v_2, v_3\}$, otteniamo una base ortonormale $\{v_2, v_3\}$ per U e scriviamo la composizione nel modo seguente:

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} -1 & 0 \\ 0 & I \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & \text{Rot}_\theta \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & \text{Rot}_\theta \end{pmatrix}.$$

Come per le rotazioni, i casi $\theta = 0$ e $\theta = \pi$ sono un po' particolari:

- una antirotazione di angolo $\theta = 0$ è una riflessione rispetto a U ;
- una antirotazione di angolo $\theta = \pi$ è una riflessione rispetto all'origine.

Ricapitolando, rotazioni e antirotazioni in \mathbb{R}^3 sono isometrie che, rispetto ad una base ortonormale, si rappresentano nel modo seguente:

$$\begin{pmatrix} 1 & 0 \\ 0 & \text{Rot}_\theta \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & \text{Rot}_\theta \end{pmatrix}.$$

Questo insieme di matrici contiene anche tutte le riflessioni rispetto a piani, rette, e all'origine. Le riflessioni si ottengono con i valori $\theta = 0$ oppure π .

Nella prossima sezione dimostriamo che tutte le isometrie di \mathbb{R}^3 sono rotazioni o antirotazioni.

7.6.9. Isometrie dello spazio. Abbiamo classificato precedentemente le isometrie di \mathbb{R}^2 , e passiamo adesso alle isometrie di \mathbb{R}^3 . Osserviamo innanzitutto che descrivere esplicitamente tutte le matrici ortogonali 3×3 è possibile ma molto più laborioso che in dimensione due. Ci accontenteremo quindi di dare una descrizione geometrica chiara di tutte le isometrie possibili. Ci servirà questo fatto generale:

Proposizione 7.6.17. *Sia $T: V \rightarrow V$ una isometria di uno spazio V dotato di un prodotto scalare definito positivo. Per ogni sottospazio $U \subset V$ vale*

$$T(U)^\perp = T(U^\perp).$$

Dimostrazione. Se $v \in U^\perp$, allora $\langle v, u \rangle = 0$ per ogni $u \in U$. Quindi $\langle T(v), T(u) \rangle = 0$ per ogni $u \in U$, in altre parole $T(v) \in T(U)^\perp$. Abbiamo mostrato che $T(U^\perp) \subset T(U)^\perp$. I due sottospazi in verità coincidono perché hanno entrambi la stessa dimensione $\dim V - \dim U$. \square

Corollario 7.6.18. *Se $T(U) = U$, allora $T(U^\perp) = U^\perp$.*

Il teorema seguente è un risultato importante della geometria dello spazio. La sua dimostrazione è anche interessante perché usa molti degli argomenti trattati nei capitoli precedenti: matrici associate, autovettori e prodotti scalari.

Teorema 7.6.19. *Ogni isometria di \mathbb{R}^3 è una rotazione o un'antirrotazione.*

Dimostrazione. Sia $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ una isometria. Siccome \mathbb{R}^3 ha dimensione dispari, per la Proposizione 5.1.37 esiste almeno un autovettore $v \neq 0$ per T . L'autovalore di v è ± 1 per la Proposizione 7.6.8. Quindi $T(v) = \pm v$.

Il piano $U = v^\perp$ ortogonale a v è T -invariante per il Corollario 7.6.18. Prendiamo una base ortonormale $\mathcal{B} = \{v, v_2, v_3\}$. I vettori v_2, v_3 formano una base ortonormale di U e otteniamo

$$A = [T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix}$$

dove B è una matrice 2×2 ortogonale che rappresenta la restrizione $T|_U$ nella base ortonormale $\{v_2, v_3\}$. Per la Proposizione 7.6.12, la matrice B è una matrice di rotazione o di riflessione. Se è di rotazione, otteniamo

$$A = \begin{pmatrix} \pm 1 & 0 \\ 0 & \text{Rot}_\theta \end{pmatrix}$$

e quindi A è una rotazione o antirrotazione a seconda del segno di ± 1 . Se B è una matrice di riflessione, esiste una base ortonormale $\{v'_2, v'_3\}$ di U rispetto alla quale la riflessione è la matrice $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$. Con la base $\mathcal{B}' = \{v, v'_2, v'_3\}$ otteniamo

$$A' = [T]_{\mathcal{B}'}^{\mathcal{B}'} = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Questa è una riflessione, un caso particolare di rotazione o antirrotazione. \square

Abbiamo scoperto che ogni isometria di \mathbb{R}^3 è una rotazione o una antirrotazione. Ci poniamo il problema seguente: concretamente, data una matrice ortogonale A , come faccio a determinare geometricamente di che tipo di (anti-)rotazione si tratta?

Rispondere a questa domanda è molto facile. Innanzitutto, se $\det A = 1$ si tratta di una rotazione; se invece $\det A = -1$, è una antirrotazione. L'angolo di (anti-)rotazione θ è inoltre facilmente determinato dalla traccia di A , tramite la formula seguente:

$$\text{tr}A = \pm 1 + 2 \cos \theta$$

dove $\pm 1 = \det A$ dipende se A è una rotazione o una antirrotazione. Quindi possiamo scrivere

$$\cos \theta = \frac{\text{tr}A - \det A}{2}.$$

In questo modo si determina immediatamente l'angolo θ . Se $\theta = 0$ oppure π , la matrice A rappresenta in realtà una riflessione (oppure è l'identità), che può essere descritta trovando gli autospazi V_1 e V_{-1} .

Nel caso generico $\theta \neq 0, \pi$, l'autospazio V_1 (rispettivamente, V_{-1}) è precisamente l'asse di rotazione (rispettivamente, antirrotazione).

Esempio 7.6.20. Data la matrice

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

vediamo che $\det A = 1$ e quindi è una rotazione; abbiamo $\operatorname{tr} A = 0$ e quindi $\cos \theta = -\frac{1}{2}$, che implica $\theta = \pm \frac{3\pi}{2}$. Infine, l'asse di rotazione è l'autospazio

$$V_1 = \ker(A - I) = \ker \begin{pmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} = \operatorname{Span} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Lo spazio euclideo

In questo capitolo studiamo la geometria dello spazio euclideo \mathbb{R}^n , con una particolare attenzione al piano \mathbb{R}^2 e allo spazio \mathbb{R}^3 che sono ovviamente quelli che ci riguardano più da vicino.

8.1. Prodotto vettoriale

Il prodotto scalare è una operazione che prende due vettori e restituisce uno scalare. Introduciamo qui un'altra operazione, il prodotto vettoriale, che prende due vettori e restituisce un *vettore*. Il prodotto vettoriale è definito solo su \mathbb{R}^3 .

8.1.1. Definizione. Consideriamo due vettori $v, w \in \mathbb{R}^3$:

$$v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \quad w = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}.$$

Il *prodotto vettoriale* fra v e w è il vettore

$$v \times w = \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix}.$$

Notiamo che

$$v \times w = \begin{pmatrix} d_1 \\ -d_2 \\ d_3 \end{pmatrix}$$

dove d_i è il determinante del minore 2×2 ottenuto cancellando la i -esima riga dalla matrice

$$A = \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \\ v_3 & w_3 \end{pmatrix}.$$

Un metodo per tenere a mente la definizione di prodotto vettoriale è il seguente. Si ottiene il prodotto vettoriale svolgendo formalmente il determinante di

questa "matrice" lungo l'ultima colonna:

$$\begin{aligned} v \times w &= \det \begin{pmatrix} v_1 & w_1 & e_1 \\ v_2 & w_2 & e_2 \\ v_3 & w_3 & e_3 \end{pmatrix} \\ &= \det \begin{pmatrix} v_2 & w_2 \\ v_3 & w_3 \end{pmatrix} e_1 - \det \begin{pmatrix} v_1 & w_1 \\ v_3 & w_3 \end{pmatrix} e_2 + \det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} e_3. \end{aligned}$$

Questa è solo una regola mnemonica: quella matrice in realtà non è una vera matrice perché e_1, e_2, e_3 non sono numeri ma i vettori della base canonica.

Esempio 8.1.1. Vale $e_1 \times e_2 = e_3$.

8.1.2. Proprietà. Studiamo le proprietà del prodotto vettoriale. Siano $v, w \in \mathbb{R}^3$ due vettori qualsiasi.

Proposizione 8.1.2. *Il vettore $v \times w$ è ortogonale sia a v che a w .*

Dimostrazione. Si ottiene

$$\begin{aligned} \langle v \times w, v \rangle &= \det \begin{pmatrix} v_2 & w_2 \\ v_3 & w_3 \end{pmatrix} v_1 - \det \begin{pmatrix} v_1 & w_1 \\ v_3 & w_3 \end{pmatrix} v_2 + \det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} v_3 \\ &= \det \begin{pmatrix} v_1 & w_1 & v_1 \\ v_2 & w_2 & v_2 \\ v_3 & w_3 & v_3 \end{pmatrix} = 0. \end{aligned}$$

La seconda uguaglianza è lo sviluppo di Laplace sull'ultima colonna. Il determinante è nullo perché la matrice ha due colonne uguali. Analogamente si trova $\langle v \times w, w \rangle = 0$. \square

Proposizione 8.1.3. *Il vettore $v \times w$ è nullo $\iff v$ e w sono dipendenti.*

Dimostrazione. Ricordiamo che

$$v \times w = \begin{pmatrix} d_1 \\ -d_2 \\ d_3 \end{pmatrix}$$

dove d_1, d_2, d_3 sono i determinanti dei minori 2×2 della matrice A avente come colonne v e w . Otteniamo

$$v \times w = 0 \iff \text{rk}A \leq 1 \iff v \text{ e } w \text{ sono dipendenti.}$$

La dimostrazione è conclusa. \square

Proposizione 8.1.4. *Vale l'equazione*

$$\|v \times w\|^2 + \langle v, w \rangle^2 = \|v\|^2 \|w\|^2.$$

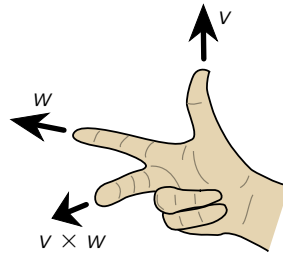


Figura 8.1. La regola della mano destra.

Dimostrazione. Troviamo

$$\begin{aligned}\|v \times w\|^2 &= (v_2 w_3 - v_3 w_2)^2 + (v_1 w_3 - v_3 w_1)^2 + (v_1 w_2 - v_2 w_1)^2 \\ &= (v_1^2 + v_2^2 + v_3^2)(w_1^2 + w_2^2 + w_3^2) - (v_1 w_1 + v_2 w_2 + v_3 w_3)^2 \\ &= \|v\|^2 \|w\|^2 - \langle v, w \rangle^2.\end{aligned}$$

La dimostrazione è completa. \square

Corollario 8.1.5. *Il modulo del prodotto vettoriale è*

$$\|v \times w\| = \|v\| \|w\| \cdot |\sin \theta| = \text{Area}(P)$$

dove θ è l'angolo formato da v e w e P è il parallelogramma con lati v e w .

Dimostrazione. Troviamo

$$\|v \times w\|^2 = \|v\|^2 \|w\|^2 - \langle v, w \rangle^2 = \|v\|^2 \|w\|^2 (1 - \cos^2 \theta) = \|v\|^2 \|w\|^2 \sin^2 \theta.$$

Quindi $\|v \times w\| = \|v\| \|w\| \cdot |\sin \theta|$ e questa è proprio l'area di P . \square

Se v e w sono dipendenti, allora $v \times w = 0$. Se sono indipendenti, allora la direzione di $v \times w$ è ortogonale al piano contenente v e w e il modulo di $v \times w$ è l'area di P . Queste due condizioni determinano il vettore $v \times w$ a meno di segno: per scegliere il "verso giusto" è sufficiente usare la regola della mano destra mostrata nella Figura 8.1 e giustificata dalla proposizione seguente.

Proposizione 8.1.6. *Se v e w sono indipendenti, la terna $v, w, v \times w$ è una base positiva di \mathbb{R}^3 .*

Dimostrazione. Dobbiamo mostrare che la matrice avente come colonne $v, w, v \times w$ ha determinante positivo. Calcoliamo sviluppando sulla terza

colonna:

$$\det \begin{pmatrix} v_1 & w_1 & \det \begin{pmatrix} v_2 & w_2 \\ v_3 & w_3 \end{pmatrix} \\ v_2 & w_2 & -\det \begin{pmatrix} v_1 & w_1 \\ v_3 & w_3 \end{pmatrix} \\ v_3 & w_3 & \det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} \end{pmatrix} = \left(\det \begin{pmatrix} v_2 & w_2 \\ v_3 & w_3 \end{pmatrix} \right)^2 + \left(\det \begin{pmatrix} v_1 & w_1 \\ v_3 & w_3 \end{pmatrix} \right)^2 + \left(\det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} \right)^2 > 0.$$

La dimostrazione è completa. \square

Possiamo quindi definire geometricamente il prodotto vettoriale $v \times w$ di due vettori indipendenti come l'unico vettore ortogonale ad entrambi, di lunghezza pari all'area del parallelogramma con lati v e w , e orientato positivamente rispetto a v e w .

Dalla definizione seguono facilmente anche i fatti seguenti:

- $v \times w = -w \times v$ per ogni $v, w \in \mathbb{R}^3$,
- il prodotto vettoriale $\times : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ è bilineare.

Come per il prodotto scalare, per bilinearità intendiamo:

$$(v + v') \times w = v \times w + v' \times w, \quad (\lambda v) \times w = \lambda(v \times w),$$

$$v \times (w + w') = v \times w + v \times w', \quad v \times (\lambda w) = \lambda(v \times w).$$

Abbiamo visto che il prodotto vettoriale ha molte proprietà. A differenza del prodotto scalare, il prodotto vettoriale non è commutativo, ma *anticommutativo*: vale $v \times w = -w \times v$.

C'è però una differenza fondamentale fra il prodotto vettoriale e altri prodotti, come quello fra numeri o fra matrici: non è verificata la proprietà associativa! Infatti ad esempio:

$$\begin{aligned} (e_1 \times e_2) \times e_2 &= e_3 \times e_2 = -e_2 \times e_3 = -e_1 \\ e_1 \times (e_2 \times e_2) &= e_1 \times 0 = 0. \end{aligned}$$

Esercizio 8.1.7 (Prodotto triplo). Siano $v_1, v_2, v_3 \in \mathbb{R}^3$. Mostra che

$$\langle v_1 \times v_2, v_3 \rangle = \langle v_1, v_2 \times v_3 \rangle = \det(v_1 | v_2 | v_3).$$

Se v_1, v_2, v_3 sono indipendenti, questo è il volume del parallelepipedo con tre lati v_1, v_2, v_3 .

8.1.3. Applicazioni. Il prodotto vettoriale è innanzitutto un metodo diretto per trovare velocemente un vettore ortogonale a due vettori indipendenti dati v_1 e v_2 .

Analogamente, si può usare il prodotto vettoriale per passare velocemente da coordinate parametriche a cartesiane per qualsiasi piano vettoriale di \mathbb{R}^3 . Se

$W = \text{Span}(v, w)$ è un piano di \mathbb{R}^3 , questo si può scrivere in forma cartesiana come

$$W = \{ax + by + cz = 0\}$$

dove a, b, c sono i coefficienti del vettore $v \times w$.

Esempio 8.1.8. Un piano

$$W = \text{Span} \left\{ \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}, \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} \right\}$$

ha come vettore ortogonale

$$v = \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} \times \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} = \begin{pmatrix} b_1 c_2 - b_2 c_1 \\ a_2 c_1 - a_1 c_2 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}$$

e quindi ha equazioni cartesiane

$$W = \{(b_1 c_2 - b_2 c_1)x + (a_2 c_1 - a_1 c_2)y + (a_1 b_2 - a_2 b_1)z = 0\}.$$

Vedremo nelle pagine seguenti altre applicazioni geometriche del prodotto vettoriale. Il prodotto vettoriale è usato in fisica, ad esempio nelle equazioni di Maxwell.

8.2. Sottospazi affini

Per definizione, un sottospazio vettoriale $U \subset V$ contiene sempre l'origine. Consideriamo adesso altri sottoinsiemi, chiamati *sottospazi affini*, che possono non contenere l'origine e sono ottenuti traslando i sottospazi vettoriali.

8.2.1. Traslazioni. Sia V uno spazio vettoriale e $v_0 \in V$ un vettore fissato. La *traslazione* di direzione v_0 è la mappa $T_{v_0}: V \rightarrow V$ definita nel modo seguente:

$$T_{v_0}(v) = v + v_0.$$

La prima considerazione da fare a proposito di questa mappa è che *non* è lineare se $v_0 \neq 0$, semplicemente perché $T_{v_0}(0) = v_0 \neq 0$.

Un'altra osservazione da fare è che la composizione di due traslazioni è sempre una traslazione:

$$T_{v_0} \circ T_{v_1} = T_{v_1} \circ T_{v_0} = T_{v_0+v_1}.$$

Notiamo anche che due traslazioni commutano sempre. Una traslazione è sempre una mappa invertibile, cioè è una corrispondenza biunivoca di V . L'inversa della traslazione T_{v_0} è la traslazione T_{-v_0} . Infatti

$$T_{-v_0} \circ T_{v_0} = T_0 = \text{id}.$$

8.2.2. Sottospazi affini. Sia V uno spazio vettoriale. Un *sottospazio affine* di V è un sottoinsieme $S \subset V$ ottenuto traslando un sottospazio vettoriale $U \subset V$ di un certo vettore v_0 , cioè

$$S = T_{v_0}(U).$$

Possiamo indicare questo sottospazio con la notazione $S = U + v_0$, che va intesa nel modo seguente:

$$S = T_{v_0}(U) = U + v_0 = \{u + v_0 \mid u \in U\}.$$

Facciamo qualche osservazione. Innanzitutto notiamo che $v_0 \in S$ perché $v_0 = 0 + v_0$. Notiamo inoltre che v_0 non è un punto speciale di S : nella notazione possiamo sostituirlo con qualsiasi altro punto $v \in S$, infatti

Proposizione 8.2.1. *Per qualsiasi $v \in S$ otteniamo*

$$S = U + v.$$

Dimostrazione. Se $v \in S$, allora $v = u + v_0$ per qualche $u \in U$ e quindi

$$U + v = (U + u) + v_0 = U + v_0.$$

Abbiamo usato che $U + u = U$ per ogni $u \in U$, un facile esercizio. \square

Esercizio 8.2.2. Se $U = \text{Span}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$ in \mathbb{R}^2 , le due rette affini

$$r_1 = U + \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad r_2 = U + \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

sono in realtà la stessa retta. Fate un disegno.

8.2.3. Supporto di un sottospazio affine. Sia $S = U + v_0$ un sottospazio affine. Il sottospazio vettoriale U è detto il *supporto* di S , e lo si ottiene traslando S all'indietro:

$$U = T_{-v_0}(S).$$

Questa operazione non dipende dal vettore $v_0 \in S$ che scegliamo. Indichiamo U con il simbolo $U = \text{supp}(S)$. Ricordiamo la notazione

$$\overrightarrow{PQ} = Q - P$$

che risulta particolarmente utile quando si studiano i sottospazi affini.

Proposizione 8.2.3. *Il supporto U di un sottospazio affine S è l'insieme*

$$U = \{\overrightarrow{PQ} \mid P, Q \in S\}$$

Dimostrazione. Sia $S = U + v_0$. Mostriamo l'inclusione \supseteq . Se $P, Q \in S$, allora $P = u_0 + v_0$ e $Q = u_1 + v_0$, quindi $\overrightarrow{PQ} = Q - P = u_1 - u_0 \in U$.

Mostriamo l'inclusione \subseteq . Se $u \in U$, allora prendendo $P = v_0$ e $Q = u + v_0$ otteniamo $\overrightarrow{PQ} = u$. \square

La *dimensione* di un sottospazio affine S è per definizione la dimensione del suo supporto $U = \text{supp}(S)$.

8.2.4. Forma parametrica e cartesiana. Siamo particolarmente interessati ai sottospazi affini dello spazio euclideo \mathbb{R}^n . Come accade in molti ambiti geometrici, un sottospazio affine può essere descritto in forma *cartesiana* (cioè *implicita*, come soluzioni di alcune equazioni) oppure *parametrica* (cioè *esplicita*, tramite alcuni parametri).

Sia $A \in M(m, n)$ e $b \in \mathbb{R}^m$. Il sistema lineare

$$S = \{x \in \mathbb{R}^n \mid Ax = b\}$$

descrive, quando ci sono soluzioni, un sottospazio affine S in *forma cartesiana*. Sappiamo per il Teorema di Rouché–Capelli che

$$S = \ker A + x_0$$

dove x_0 è una qualsiasi soluzione del sistema. Da questa scrittura vediamo subito che S è un sottospazio affine con supporto $\ker A$. Quindi

$$\dim S = \dim \ker A = n - \text{rk} A.$$

Una *forma parametrica* per S è invece una descrizione del tipo

$$S = P + U = \{P + t_1 v_1 + \dots + t_k v_k \mid t_1, \dots, t_k \in \mathbb{R}\}$$

dove $P \in S$ è un punto qualsiasi di S e $v_1, \dots, v_k \in \mathbb{R}^n$ sono una base per il supporto U . Qui ovviamente $k = \dim S$.

I sottospazi affini di \mathbb{R}^3 sono i punti, le rette, i piani, e \mathbb{R}^3 stesso.

Esempio 8.2.4. Un piano affine π in \mathbb{R}^3 è descritto da una equazione

$$\pi = \{ax + by + cz = d\}.$$

Una retta affine r in \mathbb{R}^3 è descritta da due equazioni di questo tipo, oppure più agevolmente in forma parametrica:

$$r = \left\{ \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + t \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right\}.$$

8.2.5. Esempi. Per due punti distinti in \mathbb{R}^3

$$P_0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}, \quad P_1 = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$$

passa un'unica retta r . Questa ha equazione parametrica

$$r = \{P_0 + t \overrightarrow{P_0 P_1}\}.$$

Per tre punti distinti $P_0, P_1, P_2 \in \mathbb{R}^3$ non allineati (cioè non contenuti in un'unica retta affine) passa un solo piano π . Questo ha equazione parametrica

$$\pi = \{P_0 + t \overrightarrow{P_0 P_1} + u \overrightarrow{P_0 P_2}\}.$$

Possiamo trasformare agevolmente le equazioni parametriche di un piano π in cartesiane nel modo seguente. Consideriamo un piano

$$\pi = \{P_0 + tv_1 + uv_2\}$$

dove $v_1, v_2 \in \mathbb{R}^3$ sono vettori. Calcoliamo il vettore $u = v_1 \times v_2$. Questo avrà tre coefficienti a, b, c . Il piano π può essere descritto come

$$\{ax + by + cz = d\}$$

per un certo $d \in \mathbb{R}$ che può essere trovato imponendo che $P \in \pi$.

Esempio 8.2.5. Consideriamo

$$\pi = \left\{ \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + u \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \right\}.$$

Troviamo

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \times \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

Quindi

$$\pi = \{x + 2y - z = d\}$$

per qualche $d \in \mathbb{R}$, che determiniamo imponendo

$$\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} \in \pi \implies 1 + 4 + 3 = d$$

e quindi $d = 8$. Infine otteniamo una equazione cartesiana per π :

$$\pi = \{x + 2y - z = 8\}.$$

8.2.6. Traslare l'origine. Abbiamo visto nei capitoli precedenti che per studiare un determinato problema in \mathbb{R}^n è spesso conveniente non usare la base canonica ma *cambiare base*. Quando studiamo i sottospazi affini, possiamo anche scegliere di *traslare l'origine*. Questo vuol dire fissare un punto $P \in \mathbb{R}^n$ e prendere delle nuove coordinate x' che differiscono dalle precedenti x nel modo seguente:

$$x' = x - P.$$

In altre parole:

$$\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} - \begin{pmatrix} P_1 \\ \vdots \\ P_n \end{pmatrix}.$$

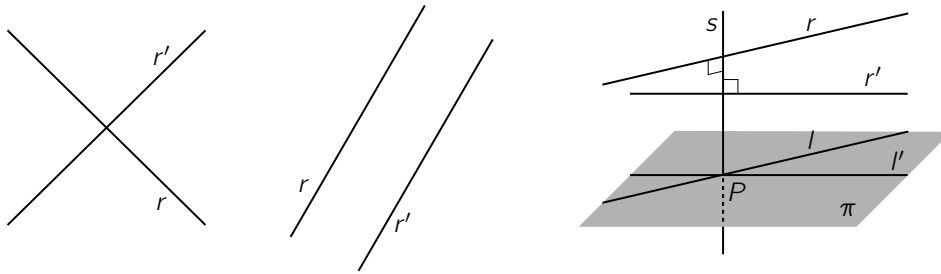


Figura 8.2. Due rette r, r' nello spazio possono essere incidenti (sinistra), parallele (centro) o sghembe (destra). Due rette sghembe hanno un'unica perpendicolare comune s (destra).

8.2.7. Intersezioni. Una peculiarità importante dei sottospazi affini, che li differenzia molto dai sottospazi vettoriali, sta nel fatto che i primi possono non intersecarsi, mentre gli ultimi devono sempre intersecarsi almeno nell'origine.

Due sottospazi affini $S, S' \subset V$ sono *incidenti* se $S \cap S' \neq \emptyset$. Due sottospazi affini incidenti si comportano un po' come i sottospazi vettoriali: se $S \cap S' \neq \emptyset$, è possibile traslare l'origine di V in modo che sia contenuta in $S \cap S'$; in questo modo S e S' diventano dei sottospazi vettoriali e possono essere studiati molto più agevolmente. Ad esempio, in questo modo verifichiamo subito che $S \cap S'$ è un nuovo sottospazio affine.

Se S e S' non sono incidenti, possono comportarsi in modo molto diverso da quanto visto per i sottospazi vettoriali, come vedremo fra poco.

8.2.8. Posizioni. Due rette affini r e r' in \mathbb{R}^3 possono essere:

- *incidenti* se $r \cap r' \neq \emptyset$;
- *parallele* se $r \cap r' = \emptyset$ e $\text{supp}(r) = \text{supp}(r')$;
- *sghembe* se $r \cap r' = \emptyset$ e $\text{supp}(r) \neq \text{supp}(r')$.

Le tre possibilità sono disegnate nella Figura 8.2.

Proposizione 8.2.6. *Date due rette sghembe r, r' , esiste un'unica retta s ortogonale ad entrambe.*

Dimostrazione. Per ipotesi $\text{supp}(r)$ e $\text{supp}(r')$ sono due rette vettoriali distinte, quindi $\pi = \text{supp}(r) + \text{supp}(r')$ è un piano vettoriale. Siano l e l' le proiezioni ortogonali di r e r' sul piano π come nella Figura 8.2-(destra). Le due rette l e l' si intersecano in un punto P . La retta s perpendicolare a π passante per P è ortogonale sia a r che a r' , si veda la Figura 8.2-(destra).

La retta s è unica: una retta ortogonale a r e r' deve essere ortogonale a π , e fra le rette ortogonali a π solo s interseca sia r che r' . \square

Un piano π ed una retta r nello spazio sono:

- *incidenti* se $\pi \cap r \neq \emptyset$,
- *paralleli* se $\pi \cap r = \emptyset$.

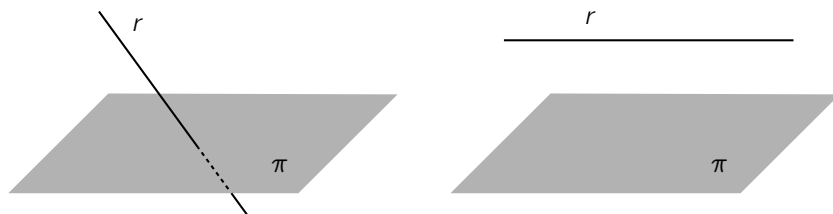


Figura 8.3. Una retta r ed un piano π possono essere incidenti (sinistra) o paralleli (destra).

I due casi sono mostrati nella Figura 8.3.

Esercizio 8.2.7. Se π e r sono paralleli, allora $\text{supp}(r) \subset \text{supp}(\pi)$.

8.2.9. Fasci. Un *fascio* di rette nel piano o di piani nello spazio è un insieme (di rette o piani) determinati da un parametro variabile.

Rette. Il fascio di rette in \mathbb{R}^2 passanti per un punto $P = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ può essere descritto prendendo le equazioni cartesiane di due rette r_0, r_1 distinte che passano per P_0 :

$$r_0 = \{a_0x + b_0y + c_0 = 0\}, \quad r_1 = \{a_1x + b_1y + c_1 = 0\}$$

e considerando tutte le loro combinazioni lineari:

$$r_{t,u} = \{t(a_0x + b_0y + c_0) + u(a_1x + b_1y + c_1) = 0\}.$$

Al variare di $t, u \in \mathbb{R}$, con la condizione che non siano entrambi nulli, otteniamo in questo modo tutte le rette passanti per P_0 . Notiamo però che c'è una certa ridondanza nel prendere due parametri t e u : per ogni $\lambda \neq 0$, i parametri (t, u) e $(\lambda t, \lambda u)$ danno in realtà la stessa retta. A meno di moltiplicare per un λ opportuno, possiamo sempre chiedere che $t + u = 1$, in altre parole possiamo prendere un parametro solo t e mettere $u = 1 - t$.

Esempio 8.2.8. Il fascio di rette passanti per $P_0 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ può essere costruito in questo modo. Se ne prendiamo due:

$$\{x - 2 = 0\}, \quad \{y - 1 = 0\}$$

e si scrive

$$r_t = \{t(x - 2) + (1 - t)(y - 1) = 0\} = \{tx + (1 - t)y - t - 1 = 0\}.$$

Il fascio di rette parallele ad una retta data r è costruito ancora più semplicemente: si prende una equazione cartesiana $\{ax + by = c\}$ di r e si sostituisce il termine c con un parametro t .

Esempio 8.2.9. Il fascio di rette parallele a $\{x + 2y = 7\}$ è $\{x + 2y = t\}$.

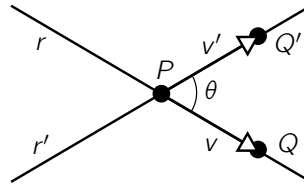


Figura 8.4. L'angolo fra due rette r e r' che si intersecano in un punto P è definito come l'angolo θ fra i vettori $v = Q - P$ e $v' = Q' - P$. Gli angoli così ottenuti sono in realtà due, θ e $\pi - \theta$, a seconda delle posizioni di Q e Q' .

Piani. Come sopra, il fascio di piani contenenti una retta r è costruito prendendo due equazioni cartesiane di due piani contenenti r e facendo una loro combinazione lineare. Il fascio di piani paralleli ad un piano con equazione $\{ax + by + cz = d\}$ si ottiene sostituendo d con un parametro t .

L'esempio seguente mostra come utilizzare agevolmente i fasci per risolvere alcuni problemi.

Esempio 8.2.10. Consideriamo la retta ed il punto

$$r = \begin{cases} x - y + 1 = 0, \\ 3x + y - z = 0, \end{cases} \quad P = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Per trovare il piano contenente r ed il punto P è sufficiente scrivere il fascio di piani contenenti r

$$\pi_t = \{t(x - y + 1) + (1 - t)(3x + y - z) = 0\}$$

ed imporre al piano generico del fascio che contenga P . In questo modo otteniamo $t + (1 - t)3 = 0$ e cioè $t = \frac{3}{2}$. Il piano cercato è $\pi_{\frac{3}{2}}$.

8.2.10. Angoli. Definiamo una nozione di angolo fra due spazi affini $A, A' \subset \mathbb{R}^3$ incidenti. Ci sono alcuni casi da considerare.

Angolo tra rette. Siano r e r' due rette che si intersecano in un punto P come nella Figura 8.4. L'angolo θ fra r e r' è definito come l'angolo formato dai vettori $v = \overrightarrow{PQ}$ e $v' = \overrightarrow{P'Q'}$ dove $Q \in r$ e $Q' \in r'$ sono diversi da P . Questa definizione si presta ad una piccola ambiguità, perché a seconda dei punti Q, Q' che scegliamo possiamo ottenere un angolo θ oppure il suo supplementare $\pi - \theta$. Generalmente si sceglie il più piccolo dei due.

Angolo tra retta e piano. Siano r una retta e π un piano che si intersecano in un punto P come nella Figura 8.5-(sinistra). Definiamo l'angolo θ fra r e π nel modo seguente: prendiamo un punto $Q \in r$ diverso da P e indichiamo con Q' la sua proiezione ortogonale su π . Se $Q' = P$, allora $r = \pi^\perp$ e poniamo $\theta = \frac{\pi}{2}$. Altrimenti l'angolo θ fra r e π è per definizione quello fra $v = \overrightarrow{PQ}$ e $v' = \overrightarrow{P'Q'}$.

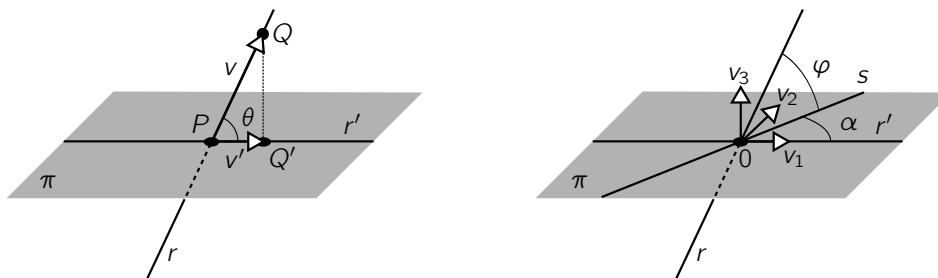


Figura 8.5. L'angolo θ fra una retta r ed un piano π che si intersecano in P è uguale all'angolo fra $v = Q - P$ e $v' = Q' - P'$ dove $Q \in r$ è un punto qualsiasi diverso da P e Q' è la sua proiezione ortogonale su π (sinistra). L'angolo θ fra r e r' è minore di quello φ fra r e s (destra).

In altre parole, proiettando r ortogonalmente su π otteniamo una retta r' e definiamo θ come l'angolo fra r e r' . Cosa succede se prendiamo un'altra retta $s \subset \pi$ al posto di r' come nella Figura 8.5-(destra)?

Proposizione 8.2.11. *Sia $s \subset \pi$ una retta che interseca r diversa da r' . L'angolo φ fra s e r è maggiore di quello θ fra r' e r .*

Dimostrazione. A meno di traslare tutto, possiamo supporre che π e r si intersechino nell'origine. Scegliamo una base ortonormale $\mathcal{B} = \{v_1, v_2, v_3\}$ come nella Figura 8.5-(destra), tale che $v_1 \in r'$ e $v_1, v_2 \in \pi$. Consideriamo adesso due vettori $v \in r$ e $u \in s$ unitari. Valgono le seguenti uguaglianze:

$$v = \cos \theta v_1 + \sin \theta v_3, \quad u = \cos \alpha v_1 + \sin \alpha v_2$$

dove $\alpha > 0$ è l'angolo tra r' e s . Otteniamo

$$\cos \varphi = \langle v, u \rangle = \cos \theta \cos \alpha < \cos \theta.$$

La funzione coseno è monotona decrescente: quindi $\varphi > \theta$. \square

Angolo tra piani. Siano π_1 e π_2 due piani che si intersecano in una retta $r = \pi_1 \cap \pi_2$. L'angolo diedrale fra π_1 e π_2 è definito nel modo seguente: si prendiamo due rette $s_1 \subset \pi_1$ e $s_2 \subset \pi_2$ incidenti ed entrambe ortogonali a r come nella Figura 8.6; l'angolo diedrale fra π_1 e π_2 è per definizione l'angolo fra s_1 e s_2 .

Esercizio 8.2.12. Siano r_1 e r_2 due rette incidenti e ortogonali a π_1 e π_2 rispettivamente. L'angolo fra π_1 e π_2 è uguale a quello fra r_1 e r_2 .

8.2.11. Distanze. Definiamo infine la *distanza* $d(A, A')$ fra due spazi affini A e A' . Se $A \cap A' \neq \emptyset$, la distanza è sempre zero. Se i sottospazi sono disgiunti la distanza è positiva ed è definita in modo diverso caso per caso.

Distanza fra punti. Come già sappiamo, la distanza $d(P, Q)$ fra due punti $P, Q \in \mathbb{R}^3$ è definita usando la norma:

$$d(P, Q) = \|\overrightarrow{PQ}\| = \|Q - P\|.$$

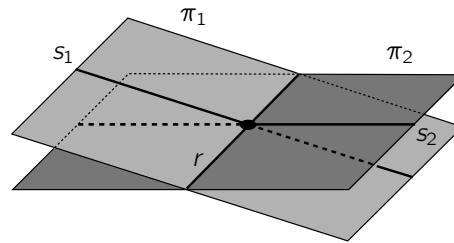


Figura 8.6. L'angolo fra due piani π_1 e π_2 che si intersecano in una retta r è definito come l'angolo fra due rette $s_1 \subset \pi_1$ e $s_2 \subset \pi_2$ incidenti entrambe ortogonali a r .

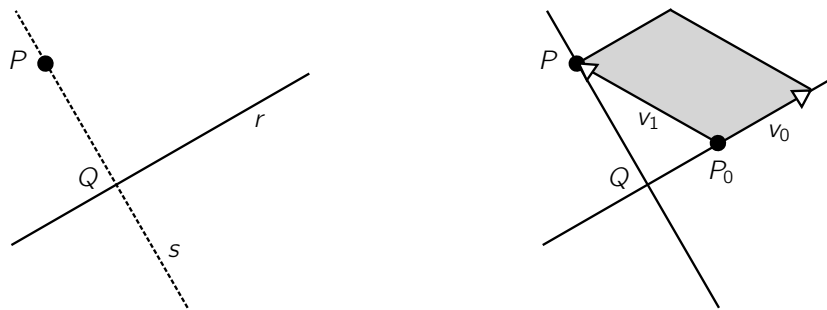


Figura 8.7. La distanza fra un punto P e una retta r è la lunghezza del segmento PQ ortogonale a r .

Distanza fra punto e retta. La distanza $d(P, r)$ fra un punto P ed una retta r nello spazio è definita nel modo seguente. Tracciamo come nella Figura 8.7 la perpendicolare s a r nel punto P e definiamo

$$d(P, r) = d(P, Q)$$

dove $Q = r \cap s$. Se la retta r è espressa in forma parametrica $r = \{P_0 + tv_0\}$ come nella Figura 8.7-(destra), la distanza si calcola facilmente usando il prodotto vettoriale.

Proposizione 8.2.13. *Vale l'uguaglianza*

$$d(P, r) = \frac{|v_0 \times v_1|}{|v_0|}$$

dove $v_1 = P - P_0$.

Dimostrazione. L'area del parallelogramma mostrato nella Figura 8.7-(destra) può essere calcolata in due modi: con il prodotto vettoriale $|v_0 \times v_1|$ e come base per altezza $|v_0| \cdot d(P, Q)$. L'uguaglianza

$$|v_0 \times v_1| = |v_0| \cdot d(P, Q)$$

implica la formula per $d(P, r) = d(P, Q)$. □

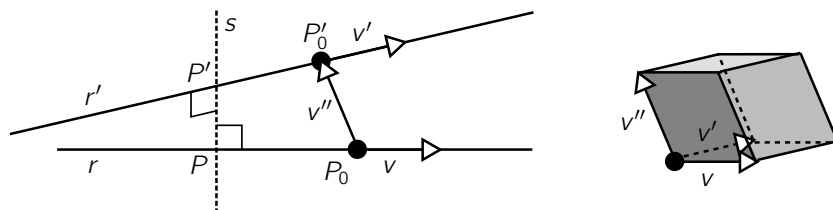


Figura 8.8. La distanza fra due rette sghembe.

Distanza fra rette. La distanza $d(r, r')$ fra due rette r ed r' disgiunte è definita nel modo seguente. Tracciamo una retta s perpendicolare ad entrambe, e definiamo

$$d(r, r') = d(P, P')$$

dove $P = r \cap s$ e $P' = r' \cap s'$. Anche in questo caso, se $r = \{P_0 + tv\}$ e $r' = \{P'_0 + tv'\}$ sono espresse in forma parametrica, la distanza si calcola facilmente (si veda la Figura 8.8).

Proposizione 8.2.14. Se r e r' sono sghembe, vale la formula

$$d(r, r') = \frac{|\det(v|v'|v'')|}{|v \times v'|}$$

dove $v'' = P'_0 - P_0$.

Dimostrazione. La dimostrazione è simile alla precedente. Il volume del parallelepipedo nella Figura 8.8-(destra) generato dai vettori v , v' e v'' può essere calcolato in due modi: come $|\det(v|v'|v'')|$ e come area di base per altezza, cioè $|v \times v'| \cdot d$. L'altezza d è proprio $d = d(P, P') = d(r, r')$ e quindi si ottiene

$$|\det(v|v'|v'')| = d(r, r') \cdot |v \times v'|$$

che implica la formula per $d(r, r')$. \square

Distanza fra punto e piano. Infine, la distanza fra un punto P_0 ed un piano π è definita in modo simile a quanto già visto: si traccia la perpendicolare s a π passante per P_0 e si definisce

$$d(P_0, \pi) = d(P_0, Q)$$

con $Q = s \cap \pi$, si veda la Figura 8.9. Se

$$\pi = \{ax + by + cz = d\}, \quad P_0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$$

allora la distanza si calcola facilmente.

Proposizione 8.2.15. Vale

$$d(P_0, \pi) = \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}$$

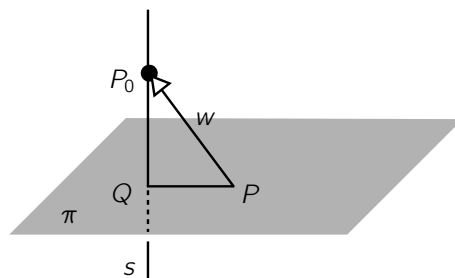


Figura 8.9. La distanza fra un punto P e un piano π .

Dimostrazione. Consideriamo

$$v = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \quad P = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

con $P \in \pi$ punto qualsiasi, come nella Figura 8.9. Il vettore v è ortogonale a π . Sia $w = P_0 - P$. La figura mostra che

$$\begin{aligned} d(P_0, Q) &= \frac{|\langle v, w \rangle|}{|v|} = \frac{|a(x_0 - x) + b(y_0 - y) + c(z_0 - z)|}{\sqrt{a^2 + b^2 + c^2}} \\ &= \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}. \end{aligned}$$

Nell'ultima uguaglianza abbiamo usato che $P \in \pi$ e quindi $ax + by + z = d$. \square

8.3. Trasformazioni del piano e dello spazio

Studiamo qui alcune trasformazioni di \mathbb{R}^2 e \mathbb{R}^3 . Nei primi paragrafi studieremo le trasformazioni vettoriali, che fissano l'origine, poi passeremo a quelle affini in cui l'origine può essere spostata a piacimento. Iniziamo con lo studio delle isometrie vettoriali del piano.

8.3.1. Affinità. Tutti gli endomorfismi di \mathbb{R}^n per definizione devono fissare l'origine. Vogliamo adesso considerare alcune trasformazioni di \mathbb{R}^n che spostano l'origine in un punto arbitrario.

Una *affinità* di \mathbb{R}^n è una mappa $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ del tipo

$$f(x) = Ax + b$$

dove A è una matrice $n \times n$ e $b \in \mathbb{R}^n$ un vettore. Chiamiamo la matrice A la *componente lineare* di f ed il vettore b la *componente traslatoria*. Notiamo che la componente traslatoria è l'immagine del vettore nullo:

$$f(0) = A0 + b = b.$$

Due esempi fondamentali di affinità sono gli endomorfismi

$$f(x) = Ax$$

in cui $b = 0$, e le *traslazioni*

$$f(x) = x + b$$

in cui $A = I$. Qualsiasi affinità $f(x) = Ax + b$ è la composizione $f = h \circ g$ di un endomorfismo $g(x) = Ax$ e di una traslazione $h(x) = x + b$.

Una affinità $f(x) = Ax + b$ è un *isomorfismo affine* se A è una matrice invertibile.

Proposizione 8.3.1. *Sia $f(x) = Ax + b$ un isomorfismo affine. Valgono i fatti seguenti:*

- (1) *la funzione $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ è biunivoca;*
- (2) *la sua inversa è $g(x) = A^{-1}x - A^{-1}b$;*
- (3) *la funzione f manda sottospazi affini in sottospazi affini;*
- (4) *gli isomorfismi affini formano un gruppo con la composizione.*

Dimostrazione. Moltiplicando entrambi i membri dell'equazione $y = Ax + b$ per A^{-1} troviamo $A^{-1}y = x + A^{-1}b$ e quindi

$$x = A^{-1}y - A^{-1}b.$$

Questo mostra in particolare che la funzione f è invertibile e quindi biunivoca. Abbiamo dimostrato (1) e (2).

Per la (3), ricordiamo che un sottospazio affine è un insieme del tipo $U + v$ con U sottospazio vettoriale. L'immagine

$$f(U + v) = A(U + v) + b = AU + Av + b$$

è un sottospazio affine $U' + v'$, con $U' = AU$ e $v' = Av + b$.

Gli isomorfismi affini formano un gruppo, perché la composizione di due isomorfismi è un isomorfismo e ogni isomorfismo ha un inversa. \square

Gli isomorfismi affini preservano anche il parallelismo: mandano rette parallele in rette parallele. Non preservano però necessariamente le distanze! A questo scopo introduciamo una nuova definizione.

8.3.2. Isometrie affini. Una *isometria affine* è un isomorfismo affine $f(x) = Ax + b$ che preserva le distanze, cioè tale che

$$d(f(P), f(Q)) = d(P, Q)$$

per ogni coppia di punti $P, Q \in \mathbb{R}^n$.

Esempio 8.3.2. Una traslazione $f(x) = x + b$ è una isometria affine, infatti $d(f(P), f(Q)) = d(P + b, Q + b) = \|(P + b) - (Q + b)\| = \|P - Q\| = d(P, Q)$.

Proposizione 8.3.3. *Le isometrie affini formano un gruppo con la composizione.*

Dimostrazione. L'inversa di una isometria è una isometria, e componendo due isometrie si ottiene una isometria. \square

Proposizione 8.3.4. *Una affinità $f(x) = Ax + b$ è una isometria affine \iff la matrice A è ortogonale.*

Dimostrazione. (\Leftarrow) Se A è ortogonale, le affinità $g(x) = Ax$ e $h(x) = x + b$ sono isometrie e quindi anche la composizione $f = h \circ g$ lo è.

(\Rightarrow) Se $f(x) = Ax + b$ è una isometria, allora

$$\|x\| = d(x, 0) = d(f(x), f(0)) = d(Ax + b, b) = \|Ax + b - b\| = \|Ax\|$$

per ogni $x \in \mathbb{R}^n$. Per la Proposizione 7.6.1 la matrice A è ortogonale. \square

8.3.3. Traslare l'origine. Quando si studiano le affinità è a volte conveniente traslare l'origine in un punto P , ottenendo quindi nuove coordinate

$$x' = x - P.$$

Se scriviamo una affinità nelle nuove coordinate x' come

$$f'(x') = Ax' + b,$$

la stessa affinità nelle coordinate originarie x sarà allora

$$f(x) = A(x - P) + b + P.$$

Possiamo anche scriverla in altri modi:

$$f(x) = Ax - AP + b + P = Ax + (I - A)P + b.$$

Notiamo che la parte lineare A della affinità è sempre la stessa, mentre la componente traslatoria è cambiata da b a $(I - A)P + b$.

Esempio 8.3.5. Sia $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ una rotazione di angolo θ intorno ad un punto $P = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$. Mostriamo adesso che f è una isometria affine e la scriviamo esplicitamente come $f(x) = Ax + b$.

Nel sistema di riferimento $x' = x - P$, la trasformazione è semplicemente $f'(x') = \text{Rot}_\theta x'$. Quindi nel sistema di riferimento iniziale x otteniamo

$$f(x) = \text{Rot}_\theta x + (I - \text{Rot}_\theta)P.$$

Esplicitamente:

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} (1 - \cos \theta)x_0 + \sin \theta y_0 \\ -\sin \theta x_0 + (1 - \cos \theta)y_0 \end{pmatrix}.$$

Come controllo, si può verificare ad esempio che $f(P) = P$.

Esercizio 8.3.6. Scrivi nella forma $f(x) = Ax + b$ la riflessione di \mathbb{R}^2 rispetto alla retta affine $x + y = 1$.

8.3.4. Punti fissi. Un *punto fisso* per una affinità $f(x) = Ax + b$ è un punto $P \in \mathbb{R}^n$ tale che $f(P) = P$.

Ad esempio, se $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ è una rotazione intorno a $P \in \mathbb{R}^2$, allora P è un punto fisso. D'altra parte, notiamo che una traslazione $f(x) = x + b$ con $b \neq 0$ non ha nessun punto fisso.

Se f ha un punto fisso P , è naturale prendere le coordinate $x' = x - P$, perché in queste nuove coordinate il punto fisso P è la nuova origine 0 e quindi l'affinità f letta nelle nuove coordinate diventa un endomorfismo:

$$f'(x') = Ax'.$$

In molti casi possiamo capire subito se c'è un punto fisso:

Proposizione 8.3.7. *Se 1 non è autovalore per A , l'affinità $f(x) = Ax + b$ ha almeno un punto fisso.*

Dimostrazione. Un punto $P \in \mathbb{R}^n$ è fisso se e solo se

$$AP + b = P \iff (A - I)P = -b.$$

Se 1 non è autovalore per A , non esiste un vettore $v \neq 0$ tale che $Av = v$, e quindi $\ker(A - I) = \{0\}$. Quindi $A - I$ è un isomorfismo ed in particolare è surgettiva: esiste un P per cui $(A - I)P = -b$. \square

8.3.5. Isometrie affini positive del piano. Classifichiamo adesso tutte le isometrie affini del piano con determinante positivo.

Proposizione 8.3.8. *Una isometria affine $f(x) = Ax + b$ con $\det A = 1$ è una traslazione oppure una rotazione intorno ad un punto $P \in \mathbb{R}^2$.*

Dimostrazione. Se f ha un punto fisso P , a meno di traslare l'origine possiamo supporre che $P = 0$, quindi l'isometria diventa vettoriale e la Proposizione 7.6.12 ci dice che f è una rotazione intorno a P .

Se f non ha punti fissi, allora A ha autovalore 1. Sempre per la Proposizione 7.6.12, deduciamo che $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ e quindi $f(x) = x + b$ è una traslazione. \square

Questa proposizione è semplice ma molto potente. Sia $f(x) = Ax + b$ una isometria affine del piano con $\det A = 1$. Sappiamo che $A = \text{Rot}_\theta$ è una matrice di rotazione; allora possiamo concludere che:

- se $\theta = 0$, cioè $A = I$, allora f è una traslazione oppure l'identità;
- se $\theta \neq 0$, allora f è una rotazione, il cui centro P può essere calcolato risolvendo l'equazione $AP + b = P$.

Da questo semplice fatto deduciamo ad esempio che:

Corollario 8.3.9. *Siano f e g due isometrie affini positive di \mathbb{R}^2 .*

- (1) *se f e g sono traslazioni, allora $f \circ g$ è una traslazione oppure l'identità;*

- (2) se f è una rotazione di angolo θ e g è una traslazione, allora $f \circ g$ è una rotazione di angolo θ ;
- (3) se f e g sono rotazioni di angolo θ e φ e $\theta + \varphi \neq 2k\pi$, allora $f \circ g$ è una rotazione di angolo $\theta + \varphi$.

Dimostrazione. Tutti i casi sono dimostrati notando che la componente rotatoria di $f \circ g$ è il prodotto delle componenti rotatorie di f e di g . \square

Esercizio 8.3.10. Scrivi due rotazioni f e g la cui composizione $f \circ g$ è una traslazione.

8.3.6. Classificazione delle isometrie affini. Vogliamo classificare adesso tutte le isometrie affini di \mathbb{R}^2 e \mathbb{R}^3 . Per questo scopo avremo bisogno di dimostrare alcuni fatti. Iniziamo con la proposizione seguente.

Proposizione 8.3.11. Se A è una matrice ortogonale, allora

$$\ker(A - I)^\perp = \text{Im}(A - I).$$

Dimostrazione. Siano $v \in \ker(A - I)$ e $w \in \mathbb{R}^n$ qualsiasi. Dobbiamo mostrare che $\langle v, (A - I)w \rangle = 0$. Per ipotesi $Av = v$ e otteniamo

$$\langle v, (A - I)w \rangle = \langle v, Aw \rangle - \langle v, w \rangle = \langle Av, Aw \rangle - \langle v, w \rangle = 0$$

perché A è una isometria vettoriale. \square

Se A è ortogonale, possiamo generalizzare la Proposizione 8.3.7 nel modo seguente.

Lemma 8.3.12. Sia $f(x) = Ax + b$ una isometria affine. A meno di traslare l'origine possiamo supporre che $b \in \ker(A - I)$.

Dimostrazione. Un cambiamento di coordinate del tipo $x = x' + d$ trasforma l'isometria affine f in

$$f'(x') = Ax' + (A - I)d + b.$$

Il nostro scopo è cercare un d per cui $b' = (A - I)d + b$ appartenga al sottospazio $\ker(A - I)$. Per la Proposizione 8.3.11 abbiamo

$$\mathbb{R}^n = \ker(A - I) \oplus \text{Im}(A - I).$$

Quindi il vettore b si scompone come $b = b_1 + b_2$ con $b_1 \in \ker(A - I)$ e $b_2 \in \text{Im}(A - I)$. Per ipotesi $b_2 = (A - I)v$ per qualche v . Basta prendere $d = -v$ e si ottiene

$$b' = (A - I)d + b = b - (A - I)v = b - b_2 = b_1 \in \ker(A - I).$$

La dimostrazione è completa. \square

Il lemma è effettivamente una generalizzazione della Proposizione 8.3.7, infatti se 1 non è autovalore per A allora $\ker(A - I) = \{0\}$ e quindi $b = 0$. Utilizziamo adesso il lemma per classificare le isometrie affini di \mathbb{R}^2 e \mathbb{R}^3 .

8.3.7. Isometrie affini del piano. Classifichiamo ora le isometrie affini di \mathbb{R}^2 . Come nel caso vettoriale, ci sono due tipi di isometrie $f(x) = Ax + b$, quelle con $\det A = 1$ che preservano l'orientazione del piano e quelle con $\det A = -1$ che la invertono.

Sappiamo già che le isometrie di \mathbb{R}^2 che preservano l'orientazione sono le traslazioni e le rotazioni intorno a qualche punto di \mathbb{R}^2 . Fra quelle che invertono l'orientazione troviamo le riflessioni rispetto a qualche retta affine r e le *glissoriflessioni*: una glissoriflessione è la composizione di una riflessione lungo una retta affine r con una traslazione non banale parallela a r . Ad esempio

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_1 \\ 0 \end{pmatrix} = \begin{pmatrix} x + b_1 \\ -y \end{pmatrix}$$

è una glissoriflessione ottenuta componendo una riflessione rispetto all'asse x con una traslazione orizzontale di passo $b_1 \in \mathbb{R}$ con $b_1 \neq 0$.

Proposizione 8.3.13. *Ogni isometria affine $f(x) = Ax + b$ di \mathbb{R}^2 è di uno dei quattro tipi seguenti:*

- se $\det A = 1$, è una traslazione o rotazione;
- se $\det A = -1$, è una riflessione o glissoriflessione.

Dimostrazione. Se f ha un punto fisso P , a meno di traslare l'origine possiamo supporre che $P = 0$ e la Proposizione 7.6.12 ci dice che f è una rotazione intorno a P o una riflessione rispetto ad una retta r contenente il punto P .

Se f non ha punti fissi, allora A ha autovalore 1. Sempre per la Proposizione 7.6.12, a meno di ruotare gli assi (cioè di prendere un'altra base ortonormale) otteniamo $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ oppure $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Nel primo caso f è una traslazione. Nel secondo, per il Lemma 8.3.12 possiamo supporre che $b \in \ker(A - I) = \ker \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix} = \text{Span}(e_1)$. Quindi $b = \begin{pmatrix} b_1 \\ 0 \end{pmatrix}$ e f è una glissoriflessione. \square

Notiamo che le isometrie affini del piano senza punti fissi sono precisamente le traslazioni e le glissoriflessioni.

8.3.8. Isometrie dello spazio. La classificazione delle isometrie affini dello spazio è analoga a quella del piano, con un paio di casi in più da considerare. Una *rototraslazione* è la composizione di una rotazione intorno ad un asse r e di una traslazione parallela a r . Ad esempio

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} b_1 \\ 0 \\ 0 \end{pmatrix}$$

è una rototraslazione ottenuta componendo una rotazione di angolo θ intorno all'asse x con una traslazione orizzontale di passo $b_1 \in \mathbb{R}$ con $b_1 \neq 0$.

Una *glissoriflessione* di \mathbb{R}^3 è la composizione di una riflessione rispetto ad un piano π e di una traslazione lungo una direzione parallela a π . Ad esempio

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ b_2 \\ b_3 \end{pmatrix}$$

è una glissoriflessione ottenuta riflettendo rispetto al piano $\pi = \{x = 0\}$ e poi trasladando di un vettore b parallelo a π .

Proposizione 8.3.14. *Ogni isometria affine $f(x) = Ax + b$ di \mathbb{R}^3 è di uno dei tipi seguenti:*

- se $\det A = 1$, è una traslazione, rotazione o rototraslazione;
- se $\det A = -1$, è una riflessione, antirotazione o glissoriflessione.

Dimostrazione. Sappiamo che A è ortogonale e quindi rappresenta l'identità, una rotazione, o una antirotazione. Se $A = I$ allora f è una traslazione. Se A è una rotazione intorno ad un asse r , allora per il Lemma 8.3.12 possiamo prendere b parallelo a r e quindi otteniamo una rototraslazione.

Se A è una antirotazione, ci sono due casi. Se l'antirotazione ha angolo zero, è in realtà una riflessione rispetto ad un piano π e per il Lemma 8.3.12 possiamo prendere b parallelo a π e ottenere una glissoriflessione. Se l'angolo è diverso da zero, allora A non ha autovalore 1 e quindi ha un punto fisso: allora f è una antirotazione. \square

Teorema spettrale

9.1. Prodotti hermitiani

In tutto il Capitolo 7 abbiamo introdotto i prodotti scalari soltanto per spazi vettoriali reali. Mostriamo adesso brevemente come sia possibile definire una nozione analoga anche per gli spazi vettoriali complessi: questa è nota con il nome di *prodotto hermitiano*.

9.1.1. Definizione. Sia V uno spazio vettoriale complesso.

Definizione 9.1.1. Un *prodotto hermitiano* su V è una applicazione

$$\begin{aligned} V \times V &\longrightarrow \mathbb{C} \\ (v, w) &\longmapsto \langle v, w \rangle \end{aligned}$$

che soddisfa i seguenti assiomi:

- (1) $\langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle$,
- (2) $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$,
- (3) $\langle v, w \rangle = \overline{\langle w, v \rangle}$

per ogni $v, v', w \in V$ e ogni $\lambda \in \mathbb{R}$.

Notiamo immediatamente che l'assioma (3) differisce da quello analogo per i prodotti scalari per la presenza di un coniugio. Il motivo di questo cambiamento sarà presto chiaro. Intanto da questi assiomi deduciamo facilmente altre proprietà:

- (4) $\langle v, w + w' \rangle = \langle v, w \rangle + \langle v, w' \rangle$,
- (5) $\langle v, \lambda w \rangle = \bar{\lambda} \langle v, w \rangle$,
- (6) $\langle 0, w \rangle = \langle v, 0 \rangle = 0$

per ogni $v, w, w' \in V$ e $\lambda \in \mathbb{R}$. Il coniugio in (3) è stato inserito per ottenere il fatto seguente:

- (7) $\langle v, v \rangle$ è un numero reale, per ogni $v \in V$.

Per dimostrare ciò, notiamo che per (3) abbiamo

$$\langle v, v \rangle = \overline{\langle v, v \rangle}$$

e quindi effettivamente $\langle v, v \rangle \in \mathbb{R}$.

9.1.2. Prodotto hermitiano definito positivo. In un prodotto hermitiano, è effettivamente cruciale che $\langle v, v \rangle$ sia un numero reale: in questo modo è perfettamente sensato chiedere che questo numero sia positivo o negativo, a seconda delle necessità. Possiamo quindi introdurre questa definizione, come abbiamo fatto nel caso reale.

Definizione 9.1.2. Un prodotto hermitiano è *definito positivo* se $\langle v, v \rangle > 0$ per ogni $v \neq 0$.

In un prodotto hermitiano definito positivo si definiscono in modo del tutto analogo al caso reale molti dei concetti già visti precedentemente: la *norma* di un vettore è

$$\|v\| = \sqrt{\langle v, v \rangle},$$

si possono quindi introdurre le nozioni di *spazio ortogonale* e di *base ortonormale* ed il procedimento di Gram – Schmidt funziona anche in questo contesto. Il *prodotto hermitiano euclideo* su \mathbb{C}^n è dato da

$$\langle x, y \rangle = {}^t x \bar{y}.$$

Come nel caso reale, il prodotto hermitiano euclideo è definito positivo.

9.1.3. Matrice associata. Come nel caso reale, se V è un prodotto hermitiano e $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di V , possiamo definire la *matrice associata* H nel modo seguente:

$$H_{ij} = \langle v_i, v_j \rangle$$

per ogni i, j . Abbiamo scelto la lettera H invece di S , perché questa matrice non è simmetrica: si tratta in realtà di una *matrice hermitiana*, cioè una matrice quadrata complessa per cui

$${}^t H = \bar{H}.$$

Qui \bar{H} è la matrice ottenuta sostituendo tutti i valori H_{ij} con i loro coniugati complessi. Ad esempio, la matrice seguente è hermitiana:

$$\begin{pmatrix} 2 & 1+i \\ 1-i & 1 \end{pmatrix}.$$

Notiamo che in una matrice hermitiana gli elementi H_{ii} sulla diagonale devono necessariamente essere reali, perché $H_{ii} = \bar{H}_{ii}$.

9.2. Endomorfismi autoaggiunti

Introduciamo in questo capitolo una importante classe di endomorfismi. In tutta questa sezione, indichiamo con V uno spazio vettoriale reale dotato di un prodotto scalare definito positivo, oppure uno spazio vettoriale complesso dotato di un prodotto hermitiano definito positivo.

9.2.1. Definizione. Un endomorfismo $T: V \rightarrow V$ è *autoaggiunto* se

$$(13) \quad \langle T(v), w \rangle = \langle v, T(w) \rangle$$

per ogni $v, w \in V$.

Notiamo subito che questa definizione ricorda un po' quella di isometria, in cui si chiede che $\langle v, w \rangle = \langle T(v), T(w) \rangle$. Ci sono effettivamente delle analogie fra le due definizioni, ma anche delle importanti differenze.

9.2.2. Con le matrici. Come per le isometrie, cerchiamo subito di tradurre la definizione di endomorfismo autoaggiunto in una condizione concreta sulle matrici associate.

Scegliamo una base ortonormale \mathcal{B} di V . Sia $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ la matrice associata ad un endomorfismo $T: V \rightarrow V$.

Proposizione 9.2.1. L'endomorfismo T è autoaggiunto se e solo se la matrice A è hermitiana.

Dimostrazione. La dimostrazione è del tutto analoga a quella della Proposizione 7.6.4. Per bilinearità, per verificare se T è autoaggiunto è sufficiente testare la condizione (13) sugli elementi v_i, v_j della base \mathcal{B} . Su questi due elementi la condizione (13) diventa

$$A_{ij} = \bar{A}_{ji}$$

per ogni i, j , in altre parole ${}^t A = \bar{A}$. □

In particolare, otteniamo il fatto seguente.

Corollario 9.2.2. Sia A una matrice $n \times n$. L'endomorfismo $L_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$ è autoaggiunto rispetto al prodotto hermitiano euclideo di \mathbb{C}^n se e solo se la matrice A è hermitiana.

Analogamente, se A è reale, l'endomorfismo $L_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ è autoaggiunto rispetto al prodotto scalare euclideo di \mathbb{R}^n se e solo se la matrice A è simmetrica (una matrice reale è hermitiana \iff è simmetrica).

Notiamo che la Proposizione 9.2.1 è valida solo se prendiamo una base \mathcal{B} ortonormale, come mostra questo esempio.

Esempio 9.2.3. Consideriamo \mathbb{R}^2 con il prodotto scalare euclideo. L'endomorfismo L_A definito dalla matrice $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ è autoaggiunto perché la matrice A è simmetrica e stiamo usando la base canonica che è ortonormale.

Se scriviamo L_A rispetto ad un'altra base ortonormale, ad esempio $\mathcal{B} = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\}$, otteniamo una nuova matrice simmetrica:

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

Se però scriviamo A rispetto ad una base *non* ortonormale, ad esempio $\mathcal{B} = \left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$, la nuova matrice associata può non essere simmetrica:

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

9.2.3. Sottospazi invarianti. Mostriamo che gli endomorfismi autoaggiunti si comportano con i sottospazi invarianti in modo simile alle isometrie.

Proposizione 9.2.4. *Sia $T: V \rightarrow V$ un endomorfismo autoaggiunto e $U \subset V$ un sottospazio. Se $T(U) \subset U$ allora $T(U^\perp) \subset U^\perp$.*

Dimostrazione. Prendiamo un vettore qualsiasi $v \in U^\perp$ e dobbiamo dimostrare che $T(v) \in U^\perp$. Per ogni vettore $u \in U$ otteniamo

$$\langle T(v), u \rangle = \langle v, T(u) \rangle = 0$$

perché $T(u) \in U$ per ipotesi. Quindi $T(v)$ è ortogonale a tutti i vettori $u \in U$, in altre parole $T(v) \in U^\perp$. \square

9.3. Il teorema

9.3.1. Enunciato. Sia come sempre V uno spazio vettoriale complesso dotato di un prodotto hermitiano definito positivo, oppure uno spazio reale dotato di un prodotto scalare definito positivo. Il seguente è uno dei più importanti teoremi di algebra lineare.

Teorema 9.3.1 (Teorema spettrale). *Un endomorfismo $T: V \rightarrow V$ è autoaggiunto \iff ha una base ortonormale di autovettori e tutti i suoi autovalori sono in \mathbb{R} .*

Dimostrazione. (\Leftarrow) Se T ha una base \mathcal{B} ortonormale di autovettori, la matrice $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ è diagonale; poiché gli autovalori sono reali, la matrice A è diagonale con numeri reali sulla diagonale e quindi è hermitiana. Per la Proposizione 9.2.1 l'endomorfismo T è autoaggiunto.

(\Rightarrow) Dimostriamo il teorema innanzitutto nel caso complesso. Iniziamo verificando che tutti gli autovalori di T sono reali: se λ è autovalore, allora c'è un autovettore v e $T(v) = \lambda v$, quindi

$$\lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle T(v), v \rangle = \langle v, T(v) \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \langle v, v \rangle.$$

Poiché $\langle v, v \rangle > 0$, deduciamo che $\lambda = \bar{\lambda}$ e quindi $\lambda \in \mathbb{R}$.

Dimostriamo adesso per induzione sulla dimensione n di V che T ha una base ortonormale di autovettori. Il caso $n = 1$ è banale: tutti gli endomorfismi di uno spazio di dimensione uno hanno una base ortonormale di autovettori, formata da un elemento solo di norma uno.

Supponiamo il fatto vero per la dimensione $n - 1$ e lo mostriamo per n . Poiché siamo sui complessi, esiste sempre un autovettore $v \in V$. La retta $\text{Span}(v)$ è chiaramente T -invariante e quindi anche lo spazio ortogonale

$U = \text{Span}(v)^\perp$ è T -invariante, per la Proposizione 9.2.4, cioè vale $T(U) \subset U$. La restrizione

$$T|_U: U \rightarrow U$$

è un endomorfismo autoaggiunto di uno spazio vettoriale U di dimensione $n-1$. Per l'ipotesi induttiva, esiste una base v_2, \dots, v_n ortonormale formata da autovettori per $T|_U$. Aggiungendo il vettore iniziale v , rinormalizzato in modo che abbia norma uno, otteniamo una base $\mathcal{B} = \{v, v_2, \dots, v_n\}$ ortonormale formata da autovettori di T .

Resta infine da mostrare il caso reale. □

Possiamo riscrivere il teorema nel linguaggio delle matrici. Consideriamo \mathbb{R}^n con l'usuale prodotto scalare euclideo.

Corollario 9.3.2. *Sia A una matrice $n \times n$ reale. Sono fatti equivalenti:*

- (1) A è simmetrica;
- (2) L_A ha una base ortonormale di autovettori;
- (3) esiste una matrice ortogonale M tale che

$${}^tMAM = M^{-1}AM = D$$

sia una matrice diagonale.

Dimostrazione. L'equivalenza (1) \Leftrightarrow (2) è il teorema spettrale. L'esistenza di una base \mathcal{B} di autovettori per A è equivalente all'esistenza di una matrice M tale che $M^{-1}AM = D$ sia diagonale. Inoltre \mathcal{B} è ortonormale se e solo se M è ortogonale, quindi (2) \Leftrightarrow (3). Notiamo che $M^{-1} = {}^tM$ in questo caso. □

Riassumendo, abbiamo scoperto due fatti non banali: il primo è che ogni matrice simmetrica reale è diagonalizzabile, ed il secondo è che sia possibile scegliere una base di autovettori che è anche ortonormale. Solo le matrici simmetriche hanno entrambe queste proprietà.

9.3.2. Segnatura. Il teorema spettrale ha varie conseguenze inaspettate, ed una di queste è un metodo definitivo per calcolare la segnatura (i_+, i_-, i_0) di una matrice simmetrica S .

Proposizione 9.3.3. *Sia S una matrice simmetrica e (i_+, i_-, i_0) la sua segnatura. I numeri i_+ , i_- e i_0 sono pari al numero di autovalori positivi, negativi e nulli di S .*

Dimostrazione. Sappiamo dalla Proposizione 9.3.2 che esiste una matrice ortogonale M per cui

$${}^tMAM = M^{-1}AM = D$$

sia diagonale. Le matrici A e D sono simultaneamente simili e congruenti! La matrice D ha la stessa segnatura di A perché sono congruenti, e gli stessi autovalori di A perché sono simili. Quindi è sufficiente mostrare la proposizione per D , ma per le diagonali la proposizione è vera per l'Esercizio 7.4.11. □

Abbiamo apparentemente trovato un metodo infallibile per determinare la segnatura di una matrice simmetrica S : è sufficiente determinare i suoi autovalori. Ma come troviamo gli autovalori di S ? Questi sono le radici del polinomio caratteristico $p_S(x)$, ma sfortunatamente non esistono formule risolutive per trovare le radici di un polinomio quando questo ha grado ≥ 5 , e di fatto l'unica formula gestibile è quella in grado 2 che si impara alle superiori.

Per fortuna non è necessario determinare precisamente gli autovalori, è sufficiente il loro *segno*. A questo scopo possiamo usare il *criterio di Cartesio*, che non dimostriamo:

Teorema 9.3.4 (Criterio di Cartesio). *Sia $p(x) = a_n x^n + \dots + a_m x^m$ un polinomio con n radici reali, scritto in modo che il primo e l'ultimo coefficiente a_n e a_m siano entrambi non nulli. Valgono i fatti seguenti:*

- (1) *il numero di radici nulle di $p(x)$ è m ;*
- (2) *il numero di radici positive di $p(x)$ è pari al numero di cambiamenti di segno nella sequenza a_n, \dots, a_m dei coefficienti, da cui siano stati rimossi i coefficienti nulli.*

Ad esempio, il polinomio $p(x) = x^4 - x^2$ ha due radici nulle (perché $m = 2$) e una radice positiva perché la sequenza dei coefficienti è $1, -1$ e contiene una variazione. Effettivamente sappiamo che le radici sono $0, 0, 1, -1$.

Il numero di radici negative è n meno il numero di radici nulle e positive (attenzione: questo non è necessariamente il numero di permanenze di segno nella successione dei coefficienti!)

Con il criterio di Cartesio, è sempre possibile determinare la segnatura di una matrice esaminando i segni dei coefficienti del suo polinomio caratteristico. Notiamo però che il calcolo esplicito del polinomio caratteristico di una matrice $n \times n$ può essere dispendioso: quando funziona, il criterio di Jacobi spiegato nella Sezione 7.4.6 è più immediato.

Geometria proiettiva

10.1. Introduzione

Introduciamo in questo capitolo una nuova geometria, chiamata *geometria proiettiva*. La geometria proiettiva è un arricchimento della geometria affine, realizzato aggiungendo allo spazio euclideo \mathbb{R}^n alcuni “punti all’infinito”. Questo arricchimento presenta numerosi vantaggi, sia algebrici che geometrici. Renderemo in particolare rigoroso l’assunto informale che due rette parallele “si incontrano all’infinito”.

10.1.1. Lo spazio proiettivo. Possiamo descrivere lo *spazio proiettivo n -dimensionale* \mathbb{P}^n nel modo seguente. Ogni punto $P \in \mathbb{P}^n$ è descritto con un simbolo di questo tipo:

$$[x_1, \dots, x_{n+1}]$$

e ci sono due regole fondamentali:

- (1) i valori x_i non sono tutti nulli, cioè il simbolo $[0, \dots, 0]$ è vietato;
- (2) due simboli $[x_1, \dots, x_{n+1}]$ e $[\lambda x_1, \dots, \lambda x_{n+1}]$ indicano lo stesso punto, per ogni $\lambda \neq 0$.

I valori x_1, \dots, x_{n+1} sono le *coordinate omogenee* di P . Ad esempio, nel piano proiettivo \mathbb{P}^2 ogni punto ha tre coordinate omogenee $[x_1, x_2, x_3]$, però le coordinate $[1, -3, 2]$ e $[5, -15, 10]$ indicano in realtà lo stesso punto: ogni punto ha molti nomi. Il punto $[0, 0, 0]$ però non è ammesso, non esiste.

10.1.2. I punti al finito e all’infinito. Consideriamo adesso un punto $[x_1, \dots, x_{n+1}]$ nello spazio proiettivo \mathbb{P}^n . Si possono presentare due casi:

- Se $x_{n+1} \neq 0$, allora possiamo usare la regola (1) e dividere ogni coefficiente per x_{n+1} . Quindi il punto ha anche coordinate omogenee

$$\left[\frac{x_1}{x_{n+1}}, \dots, \frac{x_n}{x_{n+1}}, 1 \right].$$

In questo caso diciamo che il punto è *al finito*.

- Se $x_{n+1} = 0$, diciamo che il punto è *all’infinito*.

In altre parole, usando la regola (1) possiamo sempre supporre che l’ultima coordinata x_{n+1} sia 1 oppure 0, una specie di variabile booleana che ci dice se il punto è al finito o all’infinito.

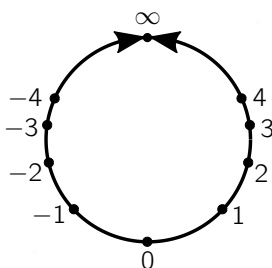


Figura 10.1. La retta proiettiva \mathbb{P}^1 è ottenuta aggiungendo un punto all'infinito alla retta reale \mathbb{R} .

Notiamo adesso che i punti al finito corrispondono precisamente ad \mathbb{R}^n . Possiamo infatti interpretare un punto di \mathbb{R}^n con coordinate x_1, \dots, x_n come il punto al finito

$$[x_1, \dots, x_n, 1].$$

La filosofia che sta dietro alla costruzione di \mathbb{P}^n è adesso chiara: tramite l'introduzione di una nuova coordinata x_{n+1} , che di fatto possiamo supporre che sia solo 1 o 0, abbiamo aggiunto ad \mathbb{R}^n dei punti all'infinito, che sono della forma

$$[x_1, \dots, x_n, 0].$$

Cerchiamo adesso di studiare i punti all'infinito nei casi più semplici.

10.1.3. La retta proiettiva. Studiamo la retta proiettiva \mathbb{P}^1 . La retta proiettiva \mathbb{P}^1 è formata da punti del tipo $[x_1, x_2]$. Come abbiamo visto, a meno di moltiplicare entrambe le coordinate per x_2^{-1} possiamo supporre che x_2 sia 1 oppure 0. Nel primo caso, otteniamo i punti al finito $[x, 1]$ che corrispondono ai punti $x \in \mathbb{R}$ della retta reale \mathbb{R} .

Nel secondo caso, otteniamo i punti all'infinito $[x_1, 0]$. Ricordiamo adesso che il punto $[0, 0]$ non è ammesso: quindi $x_1 \neq 0$, e a meno di dividere per x_1 possiamo supporre che il punto abbia coordinate $[1, 0]$. Quindi abbiamo scoperto che in \mathbb{P}^1 esiste un solo punto all'infinito $[1, 0]$, che chiamiamo ∞ .

Riassumendo, la retta proiettiva \mathbb{P}^1 è ottenuta aggiungendo ad \mathbb{R} un punto solo $\infty = [1, 0]$. La retta proiettiva è in realtà una specie di circonferenza come nella Figura 10.1.

10.1.4. Il piano proiettivo. Consideriamo adesso il piano proiettivo \mathbb{P}^2 . I punti al finito sono del tipo $[x_1, x_2, 1]$ e corrispondono ai punti (x_1, x_2) del piano \mathbb{R}^2 . I punti all'infinito sono invece del tipo $[x_1, x_2, 0]$.

A differenza di \mathbb{P}^1 , qui non abbiamo un solo punto all'infinito, ma ne abbiamo infiniti. Ad esempio, i punti seguenti

$$[1, 0, 0], \quad [1, 1, 0], \quad [1, 2, 0], \quad [0, 1, 0]$$

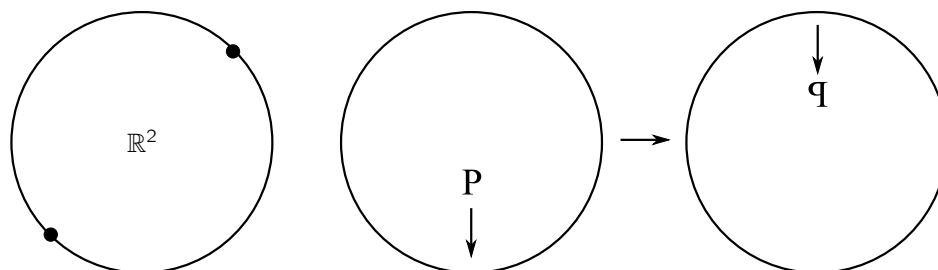


Figura 10.2. Il piano proiettivo \mathbb{P}^2 è ottenuto aggiungendo i “punti all’infinito” a \mathbb{R}^2 . Possiamo pensare a \mathbb{R}^2 come ad un disco senza bordo, cui abbiamo aggiunto dei punti di bordo: due punti opposti sono però lo stesso punto (sinistra). È possibile per un oggetto attraversare l’infinito e trovarsi quindi nella parte opposta del disco: attenzione però che l’attraversamento causa uno specchiamento dell’oggetto (destra).

sono tutti all’infinito e sono diversi fra loro. Qual è il loro significato geometrico? Geometricamente, il punto $[x_1, x_2, 0]$ è un punto all’infinito che troviamo nella direzione indicata dal vettore

$$v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Dobbiamo cioè pensare che se prendiamo una qualsiasi retta affine del tipo

$$r = P + tv$$

e spostiamo il punto sulla retta mandando t a $\pm\infty$, troviamo all’infinito proprio il punto $[x_1, x_2, 0]$. Notiamo, come nel caso di \mathbb{P}^1 , che troviamo lo stesso punto $[x_1, x_2, 0]$ se percorriamo la retta affine in entrambe le direzioni, cioè con $t \rightarrow +\infty$ e $t \rightarrow -\infty$. Notiamo anche che due rette affini puntano sullo stesso punto all’infinito se e solo se sono parallele: questo ci permetterà di dire nella prossima sezione che due rette parallele si intersecano all’infinito proprio in questo punto.

Abbiamo disegnato \mathbb{P}^1 come una circonferenza nella Figura 10.1. Qual è la forma di \mathbb{P}^2 ? La risposta è più complicata e cerchiamo di fornire un’idea nella Figura 10.2: pensiamo ad \mathbb{R}^2 come ad un disco senza bordo, cui aggiungiamo i punti di bordo, con l’accortezza però che due punti opposti rappresentano lo stesso punto. In questo modo l’infinito può essere attraversato da qualsiasi oggetto, come mostrato nella Figura 10.2: il lettore può concentrarsi e convincersi però che dopo l’attraversamento l’oggetto risulta specchiato, come in figura.

Si può rappresentare la forma di \mathbb{P}^3 alla stessa maniera, disegnando \mathbb{R}^3 come una sfera solida tridimensionale senza bordo e aggiungendo i punti di bordo, in cui però due punti antipodali risultano identificati. Il lettore può verificare che in questo caso l’attraversamento dell’infinito *non* causa uno specchiamento dell’oggetto!

10.2. Sottospazi e proiettività

Ogni spazio che si rispetti ha i suoi sottospazi e le sue mappe, e lo spazio proiettivo \mathbb{P}^n non fa eccezione.

10.2.1. Sottospazi proiettivi. Un *sottospazio proiettivo* di \mathbb{P}^n è un luogo di punti $S \subset \mathbb{P}^n$ descritto da un sistema di k equazioni lineari omogenee:

$$S = \{[x_1, \dots, x_{n+1}] \in \mathbb{P}^n \mid Ax = 0\}.$$

dove A è una matrice $k \times (n+1)$ e x ha coordinate x_1, \dots, x_{n+1} . Ad esempio, una retta proiettiva r in \mathbb{P}^2 è descritta da una equazione

$$a_1x_1 + a_2x_2 + a_3x_3 = 0.$$

Osservazione 10.2.1. Ricordiamo una peculiarità dello spazio proiettivo: ciascun punto ha più nomi, quindi quando scriviamo delle equazioni dobbiamo stare attenti. Le equazioni lineari omogenee (cioè senza termine noto) del tipo $a_1x_1 + \dots + a_nx_n = 0$ in \mathbb{P}^n hanno senso, perché se sono soddisfatte da $[x_1, \dots, x_n]$ allora sono soddisfatte anche da $[\lambda x_1, \dots, \lambda x_n]$. Un'equazione affine non omogenea (cioè con un termine noto) del tipo $x_1 + x_2 - x_3 = 3$ invece *non* ha senso nello spazio proiettivo! Il punto $[1, 1, -1]$ la soddisfa e $[2, 2, -2]$ no, però indicano lo stesso punto: l'equazione non ha senso.

Definiamo la *dimensione* del sottospazio proiettivo $S \subset \mathbb{P}^n$ come

$$\dim S = n - \text{rk}A.$$

Esempio 10.2.2. Nel piano proiettivo, i punti

$$P_1 = [1, 0, 0], \quad P_2 = [0, 1, 0], \quad P_3 = [0, 0, 1]$$

e le rette

$$r_1 = \{x_1 = 0\}, \quad r_2 = \{x_2 = 0\}, \quad r_3 = \{x_3 = 0\}$$

sono disegnati nella Figura 10.3-(sinistra). Le tre rette tagliano il piano proiettivo in quattro triangoli, indicati con colori differenti nella Figura 10.3-(destra). Quando si fanno i disegni, a volte non cerchiamo di distinguere troppo i punti al finito e all'infinito, perché nella geometria proiettiva si comportano tutti allo stesso modo: ad esempio, in questo caso i punti in r_3 in realtà starebbero tutti all'infinito, ma li abbiamo disegnati come se fossero al finito.

Come nelle altre geometrie che abbiamo visto, possiamo chiamare la *codimensione* di S il numero $n - \dim S$ e *iperpiano* un sottospazio proiettivo di codimensione uno (ad esempio, una retta di \mathbb{P}^2 oppure un piano in \mathbb{P}^3).

Come già notato nell'esempio, l'equazione $x_{n+1} = 0$ in \mathbb{P}^n descrive precisamente tutti i punti all'infinito. Abbiamo quindi scoperto che i punti all'infinito formano un iperpiano, detto *iperpiano all'infinito*. Nel piano proiettivo \mathbb{P}^2 i punti all'infinito formano la *retta all'infinito* di \mathbb{P}^2 , di equazione $x_3 = 0$.

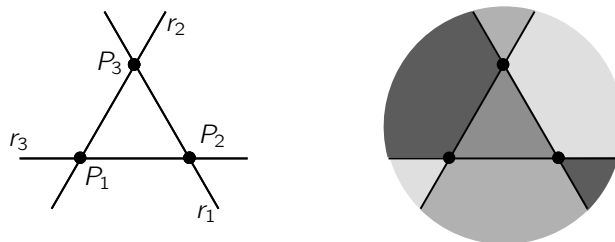


Figura 10.3. Tre punti e tre rette in \mathbb{P}^2 (sinistra). Le rette tagliano il piano proiettivo in quattro triangoli (destra).

10.2.2. Forma parametrica. Come in geometria affine, è possibile descrivere un sottospazio proiettivo anche in forma parametrica. Ad esempio, la retta passante per due punti

$$P = [x_1, \dots, x_{n+1}], \quad [y_1, \dots, y_{n+1}]$$

può essere scritta in forma parametrica come l'insieme di punti del tipo

$$[tx_1 + uy_1, \dots, tx_{n+1} + uy_{n+1}]$$

dove t, u sono variabili reali non entrambe nulle. Questa forma si estende agevolmente anche a sottospazi di dimensione maggiore.

Esempio 10.2.3. La retta r in \mathbb{P}^2 passante per $P = [1, 1, 0]$ e $Q = [0, 1, -1]$ è data dai punti

$$r = \{[t, t + u, -u]\}.$$

Si passa facilmente ad una scrittura cartesiana notando che

$$r = \{-x_1 + x_2 + x_3 = 0\}.$$

Il piano π in \mathbb{P}^3 passante per i tre punti non allineati $P = [1, -1, 0, 0]$, $Q = [1, 0, 1, 0]$ e $R = [0, 0, 0, 1]$ in forma parametrica è dato da tutti i punti del tipo

$$\pi = \{[t + u, -t, u, s]\}$$

al variare dei parametri t, u, s non tutti nulli. In forma cartesiana diventa

$$\pi = \{x_1 + x_2 - x_3 = 0\}.$$

10.2.3. Completamento di un sottospazio affine. Mostriamo adesso che i sottospazi affini di \mathbb{R}^n diventano sottospazi proiettivi di \mathbb{P}^n tramite un procedimento di *completamento*, che consiste nell'aggiunta dei loro punti all'infinito.

Consideriamo uno spazio affine $S \subset \mathbb{R}^n$ descritto da un sistema di equazioni $Ax = b$ di \mathbb{R}^n . Come sottolineato nell'Osservazione 10.2.1, queste equazioni non sono omogenee e quindi non hanno senso in \mathbb{P}^n . Rimediamo facilmente, trasformandole in equazioni omogenee nel modo seguente.

Il *completamento* di S in \mathbb{P}^n è il sottospazio proiettivo \bar{S} di \mathbb{P}^n ottenuto sostituendo ciascuna equazione $a_{i1}x_1 + \dots + a_{in}x_n = b_i$ che definisce S con la sua versione omogenea

$$a_{i1}x_1 + \dots + a_{in}x_n = b_i x_{n+1}$$

in cui il termine noto b_i diventa $b_i x_{n+1}$. Questo processo è chiamato *omogeneizzazione* delle equazioni. Il completamento \bar{S} è il luogo di zeri del sistema omogeneo che otteniamo.

Il completamento \bar{S} contiene S , perché se (x_1, \dots, x_n) è soluzione del primo sistema, allora $[x_1, \dots, x_n, 1]$ è soluzione del secondo. Il completamento \bar{S} aggiunge ad S i suoi "punti all'infinito", cioè le direzioni di tutte le rette contenute in S .

Esempio 10.2.4. Sia $r \subset \mathbb{R}^2$ la retta affine descritta dall'equazione non omogenea $x_1 + x_2 - 2 = 0$. Il suo completamento $\bar{r} \subset \mathbb{P}^2$ è la retta proiettiva

$$x_1 + x_2 - 2x_3 = 0$$

ottenuta sostituendo il termine noto 2 con $2x_3$. La retta r è formata da tutti i punti del tipo $\begin{pmatrix} t \\ 2-t \end{pmatrix}$. Il completamento \bar{r} contiene questi punti, nella forma $[t, 2-t, 1]$, ed in più anche il punto all'infinito $[1, -1, 0]$. Quindi

$$\bar{r} = r \cup \{[1, -1, 0]\}.$$

Il punto all'infinito $[1, -1, 0]$ è quello che corrisponde alla direzione di r .

Esempio 10.2.5. Il piano $x_1 = 1$ in \mathbb{R}^3 si completa al piano $x_1 = x_4$ in \mathbb{P}^3 . I punti all'infinito che abbiamo aggiunto sono quelli che soddisfano questa equazione e anche $x_4 = 0$. Questi sono precisamente tutti i punti del tipo

$$[0, t, u, 0].$$

Questa è una retta contenuta nel piano all'infinito. In generale, completando un piano si aggiunge la sua retta di punti all'infinito.

Nel piano proiettivo non ci sono rette parallele: due rette distinte si intersecano sempre in un punto.

Esempio 10.2.6. Due rette affini parallele

$$a_1x_1 + a_2x_2 = b_1, \quad a_1x_1 + a_2x_2 = b_2$$

con $b_1 \neq b_2$ non si intersecano in \mathbb{R}^2 , ma i loro completamenti

$$a_1x_1 + a_2x_2 = b_1x_3, \quad a_1x_1 + a_2x_2 = b_2x_3$$

si intersecano nel piano proiettivo \mathbb{P}^2 nel punto all'infinito $[a_2, -a_1, 0]$.

10.2.4. Proiettività. Le affinità sono trasformazioni particolari dello spazio affine. Analogamente, le *proiettività* sono trasformazioni particolari dello spazio proiettivo.

Sia A una matrice $(n+1) \times (n+1)$. La *proiettività* definita da A è la mappa $f: \mathbb{P}^n \rightarrow \mathbb{P}^n$ che manda un punto di coordinate proiettive x nel punto di coordinate proiettive Ax . Se A è invertibile, allora f è invertibile e viene detta *isomorfismo proiettivo*.

Esempio 10.2.7. La matrice

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

descrive la proiettività

$$[x_1, x_2, x_3] \mapsto [x_1 + 2x_2, x_2, -2x_3]$$

del piano proiettivo \mathbb{P}^2 .

Le proiettività di \mathbb{P}^n formano un gruppo rispetto alla composizione. Notiamo che due matrici A e λA con $\lambda \in \mathbb{R}$ non nullo descrivono la stessa proiettività.

Esercizio 10.2.8. Scrivi una proiettività $T: \mathbb{P}^2 \rightarrow \mathbb{P}^2$ tale che $T(P_1) = P_2$, $T(P_2) = P_3$ e $T(P_3) = P_1$ dove P_1, P_2, P_3 sono i punti dell'Esempio 10.2.2.

10.2.5. Punti fissi. Notiamo un fatto interessante.

Proposizione 10.2.9. *Un punto $P = [x_1, \dots, x_{n+1}]$ è punto fisso per una proiettività descritta da $A \in M(n+1, \mathbb{R})$ se e soltanto se*

$$v = \begin{pmatrix} x_1 \\ \vdots \\ x_{n+1} \end{pmatrix}$$

è autovettore per A .

Dimostrazione. Il punto P è fisso se e solo se $[v] = [Av]$, cioè se e solo se Av è multiplo di v . \square

Esempio 10.2.10. La matrice A descritta nell'Esempio 10.2.7 ha due autovalori 1 e -2 , con autospazi

$$V_1 = \text{Span} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad V_{-2} = \text{Span} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Quindi i punti fissi in \mathbb{P}^2 della proiettività descritta da A sono due: $[1, 0, 0]$ e $[0, 0, 1]$. D'altra parte, se

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

allora gli autospazi di B sono

$$V_1 = \{x_3 = 0\}, \quad V_{-2} = \text{Span} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

I punti fissi in \mathbb{P}^2 della proiettività descritta da B sono la retta $r = \{x_3 = 0\}$ ed il punto $[0, 0, 1]$.

Dalla teoria degli autovettori che abbiamo sviluppato nei capitoli precedenti possiamo dedurre facilmente un fatto non banale.

Proposizione 10.2.11. *Ogni proiettività di \mathbb{P}^2 ha almeno un punto fisso.*

Dimostrazione. Ogni matrice 3×3 ha almeno un autovettore. \square

La proposizione è abbastanza sorprendente perché non è valida in geometria affine: le traslazioni $x \mapsto x + b$ sono affinità di \mathbb{R}^2 senza punti fissi. Il fatto è che le traslazioni fissano tutti i punti all'infinito. Esaminiamo meglio questo fenomeno.

10.2.6. Da affinità a proiettività. Nell'ultimo paragrafo abbiamo suggerito che una traslazione in \mathbb{R}^n possa essere considerata una proiettività. Si tratta di un fenomeno generale: qualsiasi trasformazione affine $f(x) = Ax + b$ di \mathbb{R}^n può essere *completata* ad una trasformazione proiettiva, usando la matrice $(n+1) \times (n+1)$ data da

$$A' = \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}.$$

Questa proiettività effettivamente completa l'affinità originale, perché i punti finiti del tipo $[x, 1]$, vanno effettivamente in $[Ax + b, 1]$. I punti all'infinito, di tipo $[x, 0]$, invece vanno in $[Ax, 0]$. In particolare, se f è una traslazione, allora $A = I$ e quindi i punti all'infinito sono tutti punti fissi.

Se A è invece una rotazione in \mathbb{R}^2 intorno ad un punto, i punti all'infinito vengono tutti ruotati. Se A è una rototraslazione in \mathbb{R}^3 , il suo completamento ha un solo punto fisso all'infinito: il punto che corrisponde alla direzione dell'asse.

Esercizio 10.2.12. Scrivi una proiettività di \mathbb{P}^3 senza punti fissi.

Quadriche

11.1. Introduzione

Nei capitoli precedenti abbiamo studiato molti oggetti *lineari*, cioè definiti usando equazioni lineari, ovvero di primo grado; in questo capitolo iniziamo lo studio di oggetti non lineari, partendo da quelli più semplici: quelli *quadratici*, definiti usando equazioni di secondo grado. Questi oggetti comprendiamo le *coniche* nel piano, quali ellissi, iperboli e parabole.

11.1.1. Polinomi. Conosciamo bene i polinomi $p(x)$ in una variabile x , ed introduciamo adesso i polinomi a più variabili.

Siano x_1, \dots, x_n delle variabili. Un *monomio* è un prodotto

$$\lambda x_{i_1} \cdots x_{i_k}$$

dove $\lambda \in \mathbb{R}$ è un numero non nullo e i_1, \dots, i_k sono indici qualsiasi tra 1 e n . Gli indici possono ripetersi e k è il *grado* del monomio. Ad esempio

$$27, \quad 7x_2, \quad 2x_1x_3, \quad \pi x_1^2x_4, \quad -x_2^2x_3^2$$

sono monomi di grado 0, 1, 2, 3, 4, rispettivamente. Un *polinomio* $p(x)$ è una somma di monomi: lo indichiamo sempre con $p(x)$ per semplicità, ma ci ricordiamo che in realtà x è un vettore di variabili x_1, \dots, x_n , cioè un punto generico di \mathbb{R}^n . Il *grado* di $p(x)$ è il massimo grado dei monomi che lo compongono. Ad esempio

$$p(x) = 3x_1x_3^2 - x_1x_2 + 4x_3^2 - x_1 + 7$$

ha grado 3. Un polinomio è *omogeneo* se tutti i suoi monomi hanno lo stesso grado. Il polinomio $p(x)$ appena descritto non è omogeneo, invece

$$q(x) = 2x_1x_2 - x_3^2 + x_1x_3$$

è omogeneo.

11.1.2. Varietà affini. Se $p(x)$ è un polinomio in una sola variabile x , chiamiamo *radice* di p un punto $x \in \mathbb{R}$ per cui $p(x) = 0$. Dalla teoria dei polinomi sappiamo che, se p non è il polinomio nullo, allora ha al più d radici, dove d è il grado di p . Se $p(x)$ ha grado uno o due, sappiamo anche determinare le radici tramite formule esplicite.

Se $p(x)$ è un polinomio a più variabili x_1, \dots, x_n , la situazione ovviamente si complica. Il *luogo degli zeri* di $p(x)$, cioè l'insieme dei punti $x \in \mathbb{R}^n$ in cui

$p(x)$ si annulla, è chiamato *varietà affine*: questo insieme è spesso infinito e difficile da studiare.

Se $p(x)$ ha grado uno, allora $p(x) = a_1x_1 + \dots + a_nx_n + b$ ed il luogo degli zeri di $p(x)$ è per definizione un sottospazio affine: otteniamo un oggetto che abbiamo già studiato e che sappiamo trattare senza problemi.

Se $p(x)$ ha grado due, otteniamo precisamente l'oggetto di studio di questo capitolo.

Definizione 11.1.1. Una *quadrica affine* è il luogo di zeri in \mathbb{R}^n di un polinomio $p(x)$ di grado due.

Nelle prossime pagine useremo gli strumenti dell'algebra lineare per studiare e classificare le quadriche affini. La nostra impostazione sarà la seguente: studiamo prima il problema analogo nello spazio proiettivo \mathbb{P}^n , dove lo studio è *più facile* che in \mathbb{R}^n . L'aggiunta di punti all'infinito infatti semplificherà molto la discussione. Poi passeremo al caso affine.

11.2. Classificazione proiettiva

Introduciamo in questa sezione le quadriche proiettive e le classifichiamo.

11.2.1. Varietà proiettive. Se $p(x)$ è un polinomio qualsiasi nelle variabili x_1, \dots, x_{n+1} , possiamo cercare di interpretarlo nello spazio proiettivo \mathbb{P}^n dove ogni punto ha coordinate omogenee $[x_1, \dots, x_{n+1}]$. Non è però detto che l'equazione $p(x) = 0$ abbia senso in questo contesto, perché ricordiamo che in \mathbb{P}^n ogni punto ha molti nomi e tutto deve essere coerente. Ad esempio, il polinomio

$$p(x_1, x_2, x_3) = x_1x_2 - x_3$$

sembra annullarsi nel punto $[1, 1, 1]$, però non si annulla nel punto $[2, 2, 2]$: questo è un problema, perché i due simboli rappresentano in realtà lo stesso punto... il luogo degli zeri di $p(x)$ in \mathbb{P}^2 non è ben definito.

Per fortuna ci vengono in aiuto i polinomi omogenei.

Proposizione 11.2.1. Se $p(x)$ ha grado d ed è omogeneo, allora $p(\lambda x) = \lambda^d p(x)$ per ogni $\lambda \in \mathbb{R}$.

Dimostrazione. Tutti i monomi di $p(x)$ hanno lo stesso grado d . Ogni monomio è della forma

$$\alpha x_{i_1} \cdots x_{i_d}$$

e se sostituiamo λx al posto di x , cioè mettiamo λx_i al posto di x_i per ogni i , questo monomio si trasforma in

$$\alpha \lambda^d x_{i_1} \cdots x_{i_d}$$

e quindi viene moltiplicato per λ^d . Mettendo in evidenza il fattore λ^d che spunta da ciascun monomio otteniamo $p(\lambda x) = \lambda^d p(x)$. \square

Corollario 11.2.2. Se $p(x)$ è omogeneo e $p(x) = 0$, allora $p(\lambda x) = 0$ per ogni $\lambda \in \mathbb{R}$.

Se $p(x)$ è un polinomio omogeneo nelle variabili x_1, \dots, x_{n+1} , il luogo di zeri di $p(x)$ nello spazio proiettivo \mathbb{P}^n è ben definito ed è chiamato *varietà proiettiva*. Avevamo già notato questo fatto nel caso in cui p abbia grado uno: in questo caso $p(x) = a_1x_1 + \dots + a_{n+1}x_{n+1}$ e il luogo di zeri di $p(x)$ è un iperpiano proiettivo.

Passiamo al grado due.

11.2.2. Quadriche proiettive. Una *quadrica proiettiva* è il luogo di zeri in \mathbb{P}^n di un polinomio $p(x)$ omogeneo di grado due. Il polinomio $p(x)$ può essere scritto come

Consideriamo una proiettività $x = Mx'$ di \mathbb{P}^n determinata da una matrice invertibile $M \in M(n+1)$. La proiettività trasforma il polinomio $p(x) = {}^t x A x$ in un nuovo polinomio

$$p'(x') = p(Mx') = {}^t (Mx') A (Mx') = {}^t x' ({}^t M A M) x'.$$

Quindi la matrice A che descrive la conica cambia per congruenza in una nuova matrice $A' = {}^t M A M$.

11.2.3. Classificazione delle quadriche proiettive. Diciamo che due quadriche sono *proiettivamente equivalenti* se esiste una proiettività che trasforma una nell'altra.

Teorema 11.2.3 (Classificazione proiettiva delle quadriche). *Ogni quadrica in \mathbb{P}^n è proiettivamente equivalente ad una del tipo*

$$(14) \quad x_1^2 + \dots + x_k^2 - x_{k+1}^2 - \dots - x_{k+h}^2 = 0$$

per qualche $k, h \geq 0$ con $1 \leq k + h \leq n + 1$.

Dimostrazione. Sia $p(x) = {}^t x A x$ una quadrica. Per il Corollario 7.4.8 esiste una M invertibile per cui $D = {}^t M A M$ è una matrice diagonale con valori $+1, -1, 0$ sulla diagonale. La matrice M descrive una proiettività che trasforma la quadrica nella forma cercata. \square

L'equazione (14) è la *forma canonica* della quadrica ${}^t x A x = 0$ ed è determinata direttamente da A nel modo seguente: la terna $(k, h, n - (k + h))$ è la segnatura della matrice simmetrica A .

Notiamo che se moltiplichiamo tutta l'equazione (14) per -1 otteniamo un'altra forma canonica che descrive la stessa quadrica (il luogo di zeri non cambia) con k e h invertiti.

La quadrica è *non degenera* se $k + h = n + 1$, in altre parole se l'indice di nullità di A è zero. Ricordiamo che questo è equivalente a $\det A \neq 0$.

Passiamo ora a studiare più da vicino le quadriche in $\mathbb{P}^1, \mathbb{P}^2$ e \mathbb{P}^3 .

11.2.4. Quadriche in \mathbb{P}^1 . Studiare una quadrica ${}^t xAx = 0$ in \mathbb{P}^1 è molto simile a studiare le soluzioni di un polinomio di secondo grado in una variabile: praticamente è la stessa cosa, l'unica differenza è che qui fra le possibili soluzioni c'è anche il punto all'infinito. Otteniamo 0, 1 oppure 2 punti, a seconda del determinante di A che gioca lo stesso ruolo del Δ in una equazione di secondo grado.

Corollario 11.2.4. *Ogni quadrica proiettiva in \mathbb{P}^1 è equivalente a una di queste:*

- (1) *l'insieme vuoto $x_1^2 + x_2^2 = 0$ se $\det A > 0$;*
- (2) *due punti $x_1^2 - x_2^2 = 0$ se $\det A < 0$;*
- (3) *un punto $x_1^2 = 0$ se $\det A = 0$.*

Dimostrazione. Le possibili segnature di A sono

$$(2, 0, 0), \quad (1, 1, 0), \quad (1, 0, 1).$$

Ricordiamo che le segnature $(k, h, 2-(k+h))$ e $(h, k, 2-(k+h))$ rappresentano la stessa quadrica. Otteniamo quindi le tre forme canoniche elencate. Nella prima non ci sono soluzioni, nella seconda $x_1^2 = x_2^2$ ce ne sono due e sono $[1, 1]$ e $[1, -1]$, nella terza $x_1^2 = 0$ otteniamo solo $[0, 1]$. \square

11.2.5. Coniche proiettive. Una *conica proiettiva* è una quadrica proiettiva nel piano \mathbb{P}^2 . Queste sono classificate grazie al Teorema 11.2.3.

Corollario 11.2.5. *Ogni conica proiettiva è equivalente a una di queste:*

- (1) *l'insieme vuoto $x_1^2 + x_2^2 + x_3^2 = 0$,*
- (2) *la circonferenza $x_1^2 + x_2^2 - x_3^2 = 0$,*
- (3) *il punto $x_1^2 + x_2^2 = 0$,*
- (4) *le due rette incidenti $x_1^2 - x_2^2 = 0$,*
- (5) *la retta doppia $x_1^2 = 0$.*

Dimostrazione. Le possibili segnature di A sono

$$(3, 0, 0), \quad (2, 1, 0), \quad (2, 0, 1), \quad (1, 1, 1), \quad (1, 0, 2).$$

Ricordiamo che le segnature $(k, h, 3-(k+h))$ e $(h, k, 3-(k+h))$ rappresentano la stessa conica. Otteniamo quindi le cinque forme canoniche elencate.

Studiamo i cinque casi individualmente:

- (1) l'insieme è vuoto perché l'equazione non ha soluzioni in \mathbb{P}^2 ;
- (2) l'equazione implica $x_3 \neq 0$ e quindi possiamo porre $x_3 = 1$; la conica è contenuta interamente in $\mathbb{R}^2 \subset \mathbb{P}^2$, nella forma

$$x_1^2 + x_2^2 = 1$$

che è ovviamente una circonferenza;

- (3) l'unica soluzione in \mathbb{P}^2 è il punto $[0, 0, 1]$;
- (4) l'equazione si spezza in $(x_1 + x_2)(x_1 - x_2)$ e otteniamo l'unione di due rette $x_1 = -x_2$ e $x_1 = x_2$, che si intersecano nel punto $[0, 0, 1]$;

(5) otteniamo la retta $x_1 = 0$, ma con “molteplicità due”, come se fosse contata due volte.

La dimostrazione è completa. \square

In particolare una conica proiettiva non degenera è vuota oppure una circonferenza.

Esempio 11.2.6. Consideriamo la famiglia di coniche proiettive

$$C_k = \{kx_0^2 + x_1^2 + 5x_2^2 + 4x_0x_1 - 4x_0x_2 - 2x_1x_2 = 0\}$$

dipendenti da un parametro $k \in \mathbb{R}$. La matrice dei coefficienti è

$$A = \begin{pmatrix} k & 2 & -2 \\ 2 & 1 & -1 \\ -2 & -1 & 5 \end{pmatrix}.$$

Il determinante è $\det A = 4k - 16$.

- se $k > 4$ la segnatura è $(3, 0, 0)$ e quindi $C_k = \emptyset$;
- se $k = 4$ la segnatura è $(2, 0, 1)$ e quindi C_4 è un punto;
- se $k < 4$ la segnatura è $(2, 1, 0)$ e quindi C_k è una circonferenza.

Passiamo adesso dal piano allo spazio.

11.2.6. Quadriche proiettive in \mathbb{P}^3 . Usiamo il Teorema 11.2.3 per classificare le quadriche in \mathbb{P}^3 .

Corollario 11.2.7. *Ogni quadrica proiettiva in \mathbb{P}^3 è equivalente a una di queste:*

- (1) *l'insieme vuoto* $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 0$,
- (2) *la sfera* $x_1^2 + x_2^2 + x_3^2 - x_4^2 = 0$,
- (3) *il toro* $x_1^2 + x_2^2 - x_3^2 - x_4^2 = 0$,
- (4) *il punto* $x_1^2 + x_2^2 + x_3^2 = 0$,
- (5) *il doppio cono* $x_1^2 + x_2^2 - x_3^2 = 0$,
- (6) *i due piani incidenti* $x_1^2 - x_2^2 = 0$,
- (7) *la retta* $x_1^2 + x_2^2 = 0$,
- (8) *il piano doppio* $x_1^2 = 0$.

Dimostrazione. Le possibili segnature di A sono

$(4, 0, 0)$, $(3, 1, 0)$, $(2, 2, 0)$, $(3, 0, 1)$, $(2, 1, 1)$, $(1, 1, 2)$, $(2, 0, 2)$, $(1, 0, 3)$.

Discutiamo brevemente caso per caso:

- (1) l'equazione non ha soluzione in \mathbb{P}^3 ;
- (2) l'equazione implica $x_4 \neq 0$, poniamo $x_4 = 1$ e otteniamo la sfera

$$x_1^2 + x_2^2 + x_3^2 = 1$$

nello spazio \mathbb{R}^3 ;

- (3) riscriviamo l'equazione come $x_1^2 + x_2^2 = x_3^2 + x_4^2$ e a meno di riscalamento possiamo scrivere $x_1^2 + x_2^2 = x_3^2 + x_4^2 = 1$; Possiamo scrivere equazioni parametriche

$$x_1 = \cos \theta, \quad x_2 = \sin \theta, \quad x_3 = \cos \phi, \quad x_4 = \sin \phi$$

che dipendono da due angoli indipendenti θ e ϕ ; uno spazio parametrizzato in questo modo è effettivamente un toro;

- (4) l'unica soluzione è $[0, 0, 0, 1]$;
 (5) il luogo di zeri dell'equazione in \mathbb{R}^3 è effettivamente un doppio cono;
 (6) si spezza in $(x_1 + x_2)(x_1 - x_2) = 0$, quindi otteniamo due piani $x_1 = -x_2$ e $x_1 = x_2$ che si intersecano nella retta $x_1 = x_2 = 0$;
 (7) otteniamo la retta $x_1 = x_2 = 0$;
 (8) otteniamo il piano $x_1 = 0$ con "molteplicità due".

La dimostrazione è completa. \square

11.3. Classificazione affine

Classifichiamo in questa sezione le quadriche affini.

11.3.1. Omogeneizzazione. Consideriamo un polinomio $p(x)$ nelle variabili x_1, \dots, x_n avente grado d . Il processo di *omogeneizzazione* trasforma $p(x)$ in un polinomio omogeneo $q(x)$ nelle variabili x_1, \dots, x_{n+1} nel modo seguente: si moltiplica ciascun monomio

$$\lambda x_{i_1} \cdots x_{i_k}$$

di $p(x)$ avente un certo grado k per l'elemento x_{n+1}^{d-k} , in modo da ottenere così un nuovo monomio

$$\lambda x_{i_1} \cdots x_{i_k} x_{n+1}^{d-k}$$

di grado d . I nuovi monomi hanno tutti lo stesso grado d ed il nuovo polinomio $q(x)$ è effettivamente omogeneo.

Esempio 11.3.1. Consideriamo il caso $n = 2$. Se omogeneizziamo il polinomio $p(x) = x_1 x_2 + x_1^2 - x_2 + 7$ otteniamo

$$q(x) = x_1 x_2 + x_1^2 - x_2 x_3 + 7 x_3^2.$$

11.3.2. Completamento. Nella Sezione 10.2.3 abbiamo definito la nozione di completamento di uno spazio affine, e la estendiamo adesso alle quadriche.

Sia $Q \subset \mathbb{R}^n$ una quadrica, luogo di zeri di un polinomio $p(x)$ di secondo grado in x_1, \dots, x_n . Sia $q(x)$ il polinomio omogeneo di secondo grado in x_1, \dots, x_{n+1} ottenuto omogeneizzando $p(x)$. Definiamo il *completamento* $\bar{Q} \subset \mathbb{P}^n$ di Q come la quadrica proiettiva luogo di zeri di $q(x)$.

Proposizione 11.3.2. *Il completamento \bar{Q} contiene Q .*

Dimostrazione. Se (x_1, \dots, x_n) è soluzione di $p(x)$, allora $[x_1, \dots, x_n, 1]$ è soluzione di $q(x)$. Quindi $Q \subset \bar{Q}$. \square

Come nel caso dei sottospazi affini, la quadrica proiettiva \bar{Q} è ottenuta aggiungendo alla quadrica affine Q i suoi "punti all'infinito".

Esempio 11.3.3. Se $Q = \{x_1^2 + x_2^2 = 1\}$ è una circonferenza in \mathbb{R}^2 , il completamento $\bar{Q} = \{x_1^2 + x_2^2 = x_3^2\}$ è ancora una circonferenza: nessun punto all'infinito è stato aggiunto.

Se $Q = \{x_2 = x_1^2\}$ è una parabola in \mathbb{R}^2 , il completamento $\bar{Q} = \{x_2x_3 = x_1^2\} = \{x_1^2 - x_2x_3 = 0\}$ è una conica proiettiva descritta dalla matrice

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

La matrice ha segnatura $(2, 1, 0)$ e quindi \bar{Q} è una circonferenza. È stato aggiunto il punto all'infinito $[0, 1, 0]$.

Se $Q = \{x_1x_2 = 1\}$ è una iperbole in \mathbb{R}^2 , il completamento $\bar{Q} = \{x_1x_2 = x_3^2\}$ è ancora una circonferenza in \mathbb{P}^2 . Sono stati aggiunti due punti all'infinito $[1, 0, 0]$ e $[0, 1, 0]$.

11.3.3. Classificazione affine delle coniche. Una conica affine è il luogo di zeri di un polinomio

$$p(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1 + 2b_2x_2 + c.$$

Il polinomio omogeneizzato è

$$q(x_1, x_2, x_3) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1x_3 + 2b_2x_2x_3 + cx_3^2$$

e la sua matrice di coefficienti è

$$\bar{A} = \begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{12} & a_{22} & b_2 \\ b_1 & b_2 & c \end{pmatrix}.$$

Definiamo anche

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Sia C la conica definita da $p(x)$ e \bar{C} il suo completamento, definito da $q(x)$. Supponiamo che \bar{C} sia non-degenere, cioè che $\det \bar{A} \neq 0$. Sappiamo allora che \bar{C} è l'insieme vuoto oppure una circonferenza, a seconda della segnatura di \bar{A} . Nel primo caso anche C è vuoto e abbiamo concluso. Nel secondo, si presentano tre possibilità: possiamo interpretare il seguente teorema come definizione di ellisse, parabola e iperbole.

Teorema 11.3.4. *La circonferenza \bar{C} interseca la retta all'infinito $\{x_3 = 0\}$ in $k = 0, 1$ o 2 punti. Inoltre:*

- $k = 0 \iff \det A > 0$ e in questo caso C è una ellisse,
- $k = 1 \iff \det A = 0$ e in questo caso C è una parabola,
- $k = 2 \iff \det A < 0$ e in questo caso C è una iperbole.

Dimostrazione. La circonferenza \bar{C} è il luogo di zeri

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1x_3 + 2b_2x_2x_3 + cx_3^2 = 0.$$

Se intersechiamo con la retta $\{x_3 = 0\}$ otteniamo

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = 0.$$

Abbiamo già visto che le soluzioni in \mathbb{P}^1 di questo polinomio omogeneo di secondo grado sono 0, 1, 2 a seconda che $\det A$ sia positivo, nullo o negativo. La dimostrazione è conclusa. \square

Esercizio 11.3.5. Consideriamo la famiglia di coniche

$$C_t = \{(2t - 1)x^2 + 6txy + ty^2 + 2x = 0\}$$

dipendenti da un parametro $t \in \mathbb{R}$. Cerchiamo di capire che tipo di conica è C_t al variare di t . Scriviamo la matrice

$$\bar{A} = \begin{pmatrix} 2t - 1 & 3t & 1 \\ 3t & t & 0 \\ 2 & 0 & 0 \end{pmatrix}.$$

La conica è degenere solo per $t = 0$: per questo valore $C_0 = \{-x^2 + 2x = 0\}$ è l'unione di due rette parallele $\{x = 0\}$ e $\{x = 2\}$.

Se $t \neq 0$ la conica è non degenere. Per il Lemma 7.4.14 e il Corollario 7.4.15, la matrice \bar{A} non è definita (né positiva né negativa) e quindi è indefinita per ogni t . Segue che C_t non è mai vuota.

Calcoliamo $\det A = t(2t - 1) - 9t^2 = -7t^2 - t$ e deduciamo che C_t è una ellisse per $t \in (-\frac{1}{7}, 0)$, una parabola per $t = -\frac{1}{7}$ e una iperbole per $t \in (-\infty, -\frac{1}{7}) \cup (0, +\infty)$.

11.3.4. Classificazione metrica delle coniche non-degeneri. Sia C una conica affine descritta da un'equazione $p(x_1, x_2) = 0$ come nella sezione precedente. Possiamo scrivere il polinomio brevemente così:

$$p(x) = {}^t x A x + 2 {}^t b x + c.$$

Un *centro* della conica è un punto $P \in \mathbb{R}^2$ tale che $AP + b = 0$. Se P è un centro, è utile traslare il sistema di riferimento inserendo delle nuove coordinate $x' = x - P$ in cui il centro risulti l'origine.

Proposizione 11.3.6. *Nelle coordinate x' il polinomio $p(x)$ diventa*

$$p'(x') = {}^t x' A x' + c'$$

per qualche $c' \in \mathbb{R}$.

Dimostrazione. Il polinomio $p(x)$ si trasforma in

$$\begin{aligned} p'(x') &= p(x' + P) = {}^t(x' + P)A(x' + P) + 2 {}^t b(x' + P) + c \\ &= {}^t x' A x' + {}^t P A x' + {}^t x' A P + {}^t P A P + 2 {}^t b x' + 2 {}^t b P + c \\ &= {}^t x' A x' + 2({}^t P A + b)x' + c' \\ &= {}^t x' A x' + c' \end{aligned}$$

dove $c' = {}^t P A P + 2 {}^t b P + c$. \square

Il termine “centro” indica che P è un centro di simmetria per C : infatti sostituendo x' con $-x'$ l'espressione di $p'(x')$ resta invariata. Quindi se $r_P: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ è la riflessione rispetto al centro P otteniamo

$$r_P(C) = C.$$

Notiamo innanzitutto che se C è una ellisse o una iperbole allora $\det A \neq 0$ e quindi esiste un unico P tale che $AP = -b$, in altre parole la conica C ha un unico centro.

Possiamo a questo punto agevolmente classificare le coniche non degeneri a meno di isometrie affini di \mathbb{R}^2 .

Proposizione 11.3.7. *Ogni conica non degenera è isometrica a una di queste:*

- $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ (ellisse);
- $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ (iperbole);
- $y = ax^2$ (parabola)

per qualche $a, b > 0$.

Dimostrazione. Se la conica C è una ellisse o una iperbole, dopo aver traslato il centro nell'origine diventa della forma ${}^t x A x - c = 0$. Necessariamente $c \neq 0$ perché la conica è non degenera e quindi $\det \bar{A} \neq 0$. Possiamo dividere per c e ottenere ${}^t x A x = 1$.

A questo punto usiamo il teorema spettrale: la matrice A è simmetrica e quindi ha una base ortonormale di autovettori. Se usiamo questa base come nuovo sistema di riferimento, la matrice A è diagonale ed abbiamo concluso.

Se C è una parabola, possiamo comunque diagonalizzare A con il teorema spettrale ed ottenere

$$\bar{A} = \begin{pmatrix} a_{11} & 0 & b_1 \\ 0 & 0 & b_2 \\ b_1 & b_2 & c \end{pmatrix}.$$

Qui $a_{11}, b_2 \neq 0$ perché $\det \bar{A} \neq 0$. Allora $p(x)$ è della forma

$$p(x) = a_{11}x^2 + 2b_1x + 2b_2y + c.$$

Quindi l'equazione della parabola diventa

$$y = ax^2 + bx + c.$$

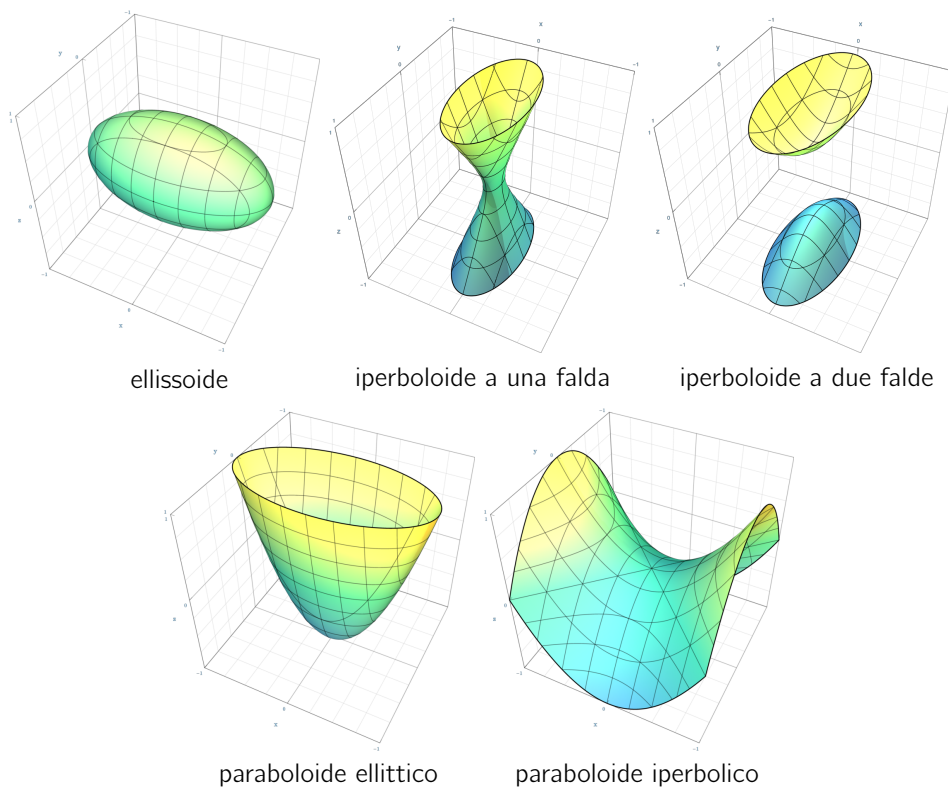


Figura 11.1. Le quadriche non degeneri.

Una traslazione opportuna la trasforma nella forma cercata. □

11.4. Classificazione metrica delle quadriche non degeneri

Con tecniche simili a quelle sviluppate nelle pagine precedenti possiamo dimostrare anche questo fatto.

Proposizione 11.4.1. *Ogni quadrica non degenera non vuota in \mathbb{R}^3 è isometrica a una di queste:*

- (1) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ (ellissoide);
- (2) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$ (iperboloide ad una falda);
- (3) $\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$ (iperboloide a due falde);
- (4) $z = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ (paraboloide ellittico);
- (5) $z = \frac{x^2}{a^2} - \frac{y^2}{b^2}$ (paraboloide iperbolico);

per qualche $a, b, c > 0$.

Le superfici sono mostrate nella Figura 11.1. Analogamente possiamo classificare le quadriche degeneri:

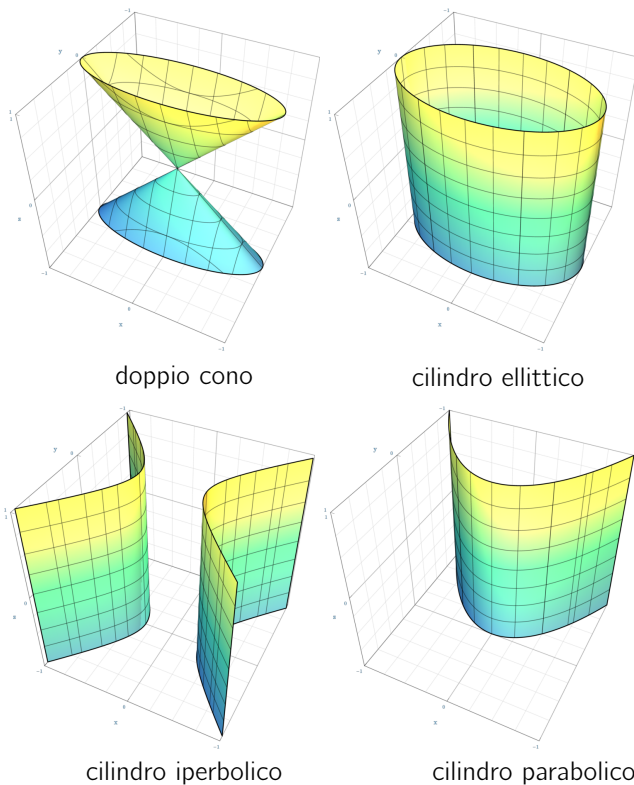


Figura 11.2. Alcune quadriche degeneri.

Proposizione 11.4.2. *Ogni quadrica degenera in \mathbb{R}^3 è isometrica a una di queste:*

- (1) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$ (doppio cono);
- (2) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ (cilindro ellittico);
- (3) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ (cilindro iperbolico);
- (4) $y = ax^2$ (cilindro parabolico);
- (5) due piani (incidenti, paralleli, oppure un piano doppio);
- (6) una retta;
- (7) un punto.

Alcune di queste sono mostrate in Figure 11.2.

Indice analitico

- Hom(V, W), 121
- ker f , 110
- algoritmo
 - di completamento a base, 60
 - di estrazione di una base, 61
 - di Gauss, 74
 - di Gauss - Jordan, 75
- anello, 32
- applicazione lineare, 103
- asse
 - immaginaria, 23
 - reale, 23
- autospazio, 148
 - generalizzato, 162
- autovalore, 135
- autovettore, 135
- base, 55
 - di Jordan, 165
- base canonica
 - di \mathbb{K}^n , 56
 - di $\mathbb{K}_n[x]$, 56
 - di $M(m, n, \mathbb{K})$, 57
- bigezione, 14
- blocco di Jordan, 157
- campo, 32
- combinazione lineare, 46
- coniugio di un numero complesso, 23
- controimmagine di una funzione, 12
- coordinate
 - cilindriche, 132
 - di un vettore rispetto ad una base, 58
 - polari, 24
- corrispondenza biunivoca, 14
- dimensione
 - di un sottospazio affine, 80
 - di uno spazio vettoriale, 59
- dimostrazione
 - per assurdo, 3
 - per induzione, 8
- divisione fra polinomi, 19
- endomorfismo
 - diagonalizzabile, 136
- forma cartesiana, 47
- forma parametrica, 47
- forma quadratica, 183
- formula di Grassmann, 63
- funzione, 11
 - bigettiva, 13
 - composta, 14
 - iniettiva, 13
 - inversa, 14
 - lineare, 103
 - suriettiva, 13
- gruppo, 30
 - simmetrico, 31
- immagine di una funzione, 12
- indipendenza lineare, 53
- insieme
 - numerabile, 34
- insieme vuoto, 4
- isomorfismo, 114
- leggi di De Morgan, 5
- matrice, 42
 - a blocchi, 170
 - antisimmetrica, 49, 67
 - associata
 - ad un prodotto scalare, 188
 - ad una applicazione lineare, 117, 124
 - dei cofattori, 98
 - di cambiamento di base, 122

- di Jordan, 158
- di Vandermonde, 101
- diagonale, 49, 138
- diagonalizzabile, 138
- identità, 95
- nilpotente, 160
- quadrata, 49
- simmetrica, 49, 67
- trasposta, 66
- triangolare, 49
- matrici
 - congruenti, 191
 - simili, 126
- modulo
 - di un numero complesso, 23
- molteplicità di una radice, 20
- molteplicità geometrica di un autovalore, 149
- mossa di Gauss, 73
- nucleo, 110
- numero
 - complesso, 22
 - intero, 3
 - naturale, 3
 - razionale, 3
 - reale, 3
- permutazione, 15
- piano complesso, 23
- pivot, 74
- polinomio
 - caratteristico, 140
 - monico, 20
- prodotto
 - cartesiano, 7
 - fra matrici, 93
- prodotto scalare, 179
 - definito positivo, 180
 - degenere, 180
 - euclideo, 180
- proiezione, 106
- quantificatore, 6
- radicale, 186
- radicale di un prodotto scalare, 186
- radice di un polinomio, 19
- rango, 81
- regola
 - di Cramer, 99
- relazione di equivalenza, 9
- restrizione di una funzione, 108
- retta
 - invariante, 140
- riflessione, 107
- rotazione, 130
- segno di una permutazione, 17
- similitudine fra matrici, 126
- sistema omogeneo associato, 79
- somma
 - di sottospazi, 52
 - diretta, 65
- somma diretta, 67
- sottospazio
 - affine, 80
 - generato, 46
 - invariante, 140
 - vettoriale, 44
- spazio
 - cartesiano, 8
 - euclideo, 37
 - vettoriale, 40
- sviluppo di Laplace, 87
- teorema
 - della dimensione, 112
 - di Binet, 96
 - di Cayley - Hamilton, 172
 - di Jordan, 159
 - di Rouché - Capelli, 82
 - fondamentale dell'algebra, 28
- traccia di una matrice, 127
- trasformazione del piano, 129
- trasformazione lineare, 128
- trasposizione, 16
- vettore
 - isotropo, 186
- vettori ortogonali, 185