

Università di Roma Tor Vergata
Corso di Laurea in
Scienza dei Media e della Comunicazione

Appunti di
Algebra Lineare

Indice

| | |
|---|----|
| Capitolo 1. Introduzione | 1 |
| 1.1. Numeri interi, razionali e reali | 1 |
| 1.2. Alcune notazioni | 4 |
| 1.3. Appendice: i numeri naturali, gli interi, i razionali ed i reali | 5 |
| Capitolo 2. Spazi vettoriali | 15 |
| 2.1. Definizione di spazio vettoriale | 16 |
| 2.2. Proprietà della somma | 17 |
| 2.3. Altre proprietà delle operazioni sui vettori | 19 |
| 2.4. Combinazioni lineare di vettori | 21 |
| 2.5. Sistemi di vettori linearmente dipendenti e indipendenti | 22 |
| 2.6. Sottospazi vettoriali ed insiemi di generatori | 25 |
| 2.7. Base di uno spazio vettoriale | 26 |
| Capitolo 3. Applicazioni lineari e matrici | 32 |
| 3.1. Definizione di applicazione lineare | 32 |
| 3.2. Immagine e nucleo di una applicazione lineare | 33 |
| 3.3. Spazi vettoriali di polinomi ed applicazioni lineari | 36 |
| 3.4. Esercizi sulle applicazioni lineari su spazi di polinomi | 37 |
| 3.5. Applicazioni lineari fra \mathbb{R}^n e \mathbb{R}^m e calcolo con matrici | 38 |
| 3.6. Spazi vettoriali di matrici | 40 |
| 3.7. Rango di una matrice | 43 |
| 3.8. Applicazioni lineari e matrici invertibili | 45 |
| 3.9. Il calcolo della matrice inversa | 46 |
| 3.10. Minori ed orli di una matrice ed invertibilità | 50 |
| 3.11. Esercizi | 54 |
| 3.12. Isomorfismi fra spazi vettoriali | 60 |
| Capitolo 4. Cambiamento di base | 61 |
| 4.1. Trasformazione di coordinate sotto cambiamento di base | 61 |
| 4.2. Matrice di una applicazione lineare e cambiamento di basi | 65 |
| 4.3. Cenni introduttivi sulla diagonalizzazione | 71 |
| 4.4. Esercizi | 72 |
| Capitolo 5. Determinante di matrici | 77 |

| | |
|---|-----|
| Capitolo 6. Prodotto scalare e ortogonalità | 80 |
| 6.1. Introduzione: prodotto scalare euclideo nel piano ed in \mathbb{R}^n | 80 |
| 6.2. Spazi vettoriali su \mathbb{C} | 81 |
| 6.3. * La definizione generale di prodotto scalare | 82 |
| 6.4. Matrici complesse, matrici autoaggiunte e matrici simmetriche | 83 |
| 6.5. * Norma e prodotti scalari definiti positivi | 84 |
| 6.6. Ortogonalità | 85 |
| 6.7. Procedimento di ortogonalizzazione di Gram-Schmidt | 88 |
| 6.8. Matrici ortogonali e matrici unitarie | 90 |
| 6.9. * Matrice associata ad un prodotto scalare | 92 |
| Capitolo 7. Funzionali lineari e dualità | 96 |
| Capitolo 8. Autovalori, autovettori e diagonalizzabilità di matrici | 97 |
| 8.1. Triangolarizzazione e diagonalizzazione | 97 |
| 8.2. Autovalori, autovettori e diagonalizzazione | 98 |
| 8.3. Ulteriori esercizi sulla diagonalizzazione sul campo \mathbb{R} | 105 |
| 8.4. Autovalori complessi e diagonalizzazione in $M_{\mathbb{C}}^{nn}$ | 113 |
| 8.5. Diagonalizzabilità di matrici simmetriche o autoaggiunte | 114 |
| 8.6. Esercizi sulla diagonalizzazione di matrici simmetriche | 116 |
| 8.7. Triangolarizzazione e forma canonica di Jordan | 116 |
| 8.8. Una applicazione: dinamica della popolazione | 116 |
| Capitolo 9. Somma diretta di spazi vettoriali | 117 |
| Capitolo 10. Esempi di geometria analitica nel piano | 118 |
| Capitolo 11. Esempi di geometria analitica in \mathbb{R}^3 | 119 |
| Bibliografia | 120 |

CAPITOLO 1

Introduzione

Questo capitolo preliminare stabilisce la terminologia sui tipi di numeri ed operazioni aritmetiche che usiamo in seguito. Per questi concetti è sufficiente fare appello all'intuizione del lettore, che, nel caso ritenga utile una maggior precisione, può leggere l'Appendice di questo capitolo oppure far riferimento a [1] (un riferimento bibliografico utile per tutto il presente libro).

1.1. Numeri interi, razionali e reali

Con il termine *algebra* si intende il calcolo, e metodi di calcolo, di numeri naturali, interi, razionali, reali e complessi.

Il primo concetto basilare è quello di *insieme*, che intuitivamente è una collezione di elementi descritti da una qualche proprietà. Per esempio l'insieme dei miei vestiti blu è formato dai vestiti che possiedo nel mio armadio e che sono di colore blu.

Osserviamo che il numero di elementi di un insieme può essere finito (come i miei vestiti blu) o infinito (come gli insiemi di numeri su cui lavoriamo).

L'insieme dei numeri *naturali*, che si indica con \mathbb{N} , è formato dai numeri che si possono contare:

$$0, 1, 2, 3, 4, 5, \dots$$

In questo corso elementare si può forse procedere contando che i lettori abbiano già una intuizione precisa dei numeri naturali (con l'operazione di somma), dei numeri interi, in cui c'è anche l'operazione di differenza, e dei numeri razionali, nei quali si introducono anche le operazioni di prodotto e di quoziente. In ogni caso, presentiamo un cenno delle definizioni e costruzioni rigorose nell'Appendice 1.3.

L'insieme dei numeri *naturali*, che si indica con \mathbb{N} , è formato da numeri che si possono ordinare consecutivamente:

$$0, 1, 2, 3, 4, 5, \dots$$

I numeri *interi*, il cui insieme è denotato con \mathbb{Z} , sono i numeri naturali e i loro *opposti*:

$$0, 1, -1, 2, -2, 3, -3, 4, -4, \dots$$

I numeri naturali e interi si possono sommare e moltiplicare. Si può fare la differenza di due numeri interi, ma non in generale di due numeri naturali: per esempio $1 - 2$ è l'intero -1 che non è un numero naturale.

Lo stesso problema si presenta con i numeri interi per la divisione: il numero

$$1 : 2 = \frac{1}{2}$$

non è intero. Si introducono allora i numeri *razionali*, il cui insieme si denota con \mathbb{Q} , che è formato dai quozienti di due numeri interi (con denominatore diverso da zero).

Prima di scrivere la definizione rigorosa di numero razionale, ricordiamo che in insiemistica si usano di solito i simboli di appartenenza \in e di sottoinsieme \subset . Per esempio il fatto che l'insieme dei numeri naturali è sottoinsieme dei numeri interi (cioè ogni numero naturale è anche un intero), che a sua volta è sottoinsieme dei numeri razionali, si scrive:

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}.$$

Il fatto che $1/2$ non è un numero intero, ma è razionale, si scrive:

$$\frac{1}{2} \notin \mathbb{Z}, \quad \frac{1}{2} \in \mathbb{Q}.$$

DEFINIZIONE 1.1.1. L'insieme dei numeri razionali è

$$\mathbb{Q} = \left\{ \frac{p}{q} \mid p, q \in \mathbb{Z}, q > 0 \right\},$$

dove possiamo supporre p e q *primi tra loro*, cioè senza divisori comuni. Si può supporre $q > 0$, perché se fosse $q < 0$, allora si potrebbe moltiplicare numeratore e denominatore per -1 , ottenendo così una frazione con denominatore positivo.

Per esempio:

$$\frac{2}{-5} = \frac{(-1) \cdot 2}{(-1) \cdot (-5)} = \frac{-2}{5} = -\frac{2}{5}.$$

Con i numeri razionali si può fare la divisione (per un numero diverso da zero):

$$\frac{p}{q} : \frac{m}{n} = \frac{p}{q} \cdot \frac{n}{m} = \frac{pn}{qm}$$

dove il denominatore qm è diverso da zero, perché sia q che m sono diversi da zero.

È conveniente rappresentare i numeri naturali, interi e razionali (e come vedremo anche quelli reali) su una retta:

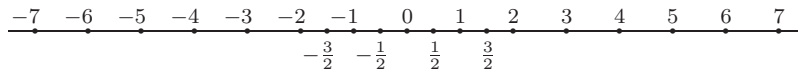


FIGURA 1. Rappresentazione grafica dei numeri naturali, interi e razionali

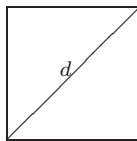
Ma anche i numeri razionali non sono sufficienti per misurare gli oggetti che troviamo in natura. Consideriamo per esempio la diagonale di un quadrato, come in figura 2.

Se il lato è lungo 1 cm, allora per il teorema di Pitagora la diagonale è:

$$d = \sqrt{2} = 0,4142 \dots \text{ cm},$$

cioè d è un numero tale che $d^2 = 2$.

PROPOSIZIONE 1.1.2. Il numero $d = \sqrt{2}$ non è razionale.

FIGURA 2. La diagonale d di un quadrato

DIMOSTRAZIONE. Supponiamo per assurdo che $\sqrt{2} = p/q$, con p e q numeri interi primi tra loro. Elevando entrambi i membri al quadrato si ottiene:

$$2 = \frac{p^2}{q^2}, \quad \text{cioè} \quad p^2 = 2q^2.$$

Ne segue che p deve essere un numero pari, quindi $p = 2m$, per un certo intero m . Sostituendo p con $2m$ nella formula precedente, si trova che:

$$p^2 = 4m^2 = 2q^2, \quad \text{perciò} \quad 2m^2 = q^2.$$

Ma allora anche q deve essere un numero pari, divisibile per 2, e questo contraddice l'ipotesi che p e q non abbiano fattori in comune. Abbiamo trovato così una contraddizione, quindi l'ipotesi che $\sqrt{2}$ fosse un numero razionale non può essere vera. \square

Si considerano allora anche i numeri *reali*, il cui insieme si indica con \mathbb{R} , che si possono scrivere come numeri interi seguiti, dopo la virgola, da infinite cifre decimali. I numeri reali si possono rappresentare sulla stessa retta di figura 1 e possiamo immaginare che ogni punto della retta corrisponde ad un numero reale.

Possiamo pensare anche i punti di un *piano* come elementi di un insieme su cui poter fare operazioni, come per esempio la somma. Infatti possiamo associare ad ogni punto del piano, determinato da una coppia di numeri reali, un *vettore*, di cui i numeri reali sono le coordinate. I vettori si possono sommare e moltiplicare per uno scalare e formano così uno *spazio vettoriale*, che definiremo dettagliatamente nel prossimo capitolo 2.

Più avanti vedremo anche che ad ogni coppia di numeri reali (ovvero, ad ogni punto del piano) possono essere associati i cosiddetti numeri *complessi*, il cui insieme si denota con \mathbb{C} . Il numero complesso associato alla coppia (a, b) si denota con

$$a + bi$$

dove i è l'unità *immaginaria*, che è un numero complesso tale che

$$i^2 = -1, \quad \text{ovvero} \quad i = \sqrt{-1}. \quad (1.1.1)$$

Noi però ci occuperemo principalmente dei numeri reali e non di quelli complessi. Fra gli insiemi di numeri che abbiamo introdotto valgono le seguenti relazione di inclusione:

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

All'interno del corso però i numeri complessi avranno un ruolo abbastanza marginale e quasi sempre avremo a che fare solo con i numeri reali.

Abbiamo detto che con il termine “algebra” si intende il calcolo di operazioni quali la somma e il prodotto di numeri.

Con il termine “algebra lineare”, che è il contenuto di questo corso, si intende lo studio e la risoluzione dei *sistemi di equazioni lineari*, come per esempio:

$$\begin{cases} 2x + 3y = 1 \\ -x + 5y = -2 \end{cases} \quad (1.1.2)$$

cioè di un numero finito di equazioni in cui compaiono *variabili lineari*, ovvero le incognite compaiono nelle espressioni solo con *grado uno* (il grado è l’esponente dell’incognita, che, essendo sempre e solo 1, di solito si tralascia).

Lo strumento per risolvere tali sistemi saranno i *vettori* e le *matrici*. Per esempio il sistema lineare di due equazioni (1.1.2) verrà scritto nel modo seguente:

$$\begin{pmatrix} 2 & 3 \\ -1 & 5 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix},$$

dove il secondo membro è il vettore dei termini noti e al primo membro

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

è la matrice quadrata di ordine 2×2 dei coefficienti delle incognite. Per risolvere tale sistema lineare si useranno allora le proprietà dei vettori e delle matrici che vedremo nel seguito.

1.2. Alcune notazioni

Gli intervalli di numeri reali li indicheremo nel modo seguente:

$$[1, 2] = \{ x \in \mathbb{R} \mid 1 \leq x \leq 2 \},$$

$$[1, 2) = \{ x \in \mathbb{R} \mid 1 \leq x < 2 \},$$

$$(1, +\infty) = \{ x \in \mathbb{R} \mid 1 < x \}.$$

Ricordiamo che possiamo descrivere un insieme elencando tutti gli elementi o indicando una proprietà. Per esempio indichiamo l’insieme dei numeri naturali dispari così:

$$\{ n \in \mathbb{N} \mid n \text{ dispari} \},$$

oppure nel modo seguente:

$$\{ n \in \mathbb{N} \mid n = 2m + 1, \text{ con } m \in \mathbb{N} \}.$$

Una costruzione che useremo spesso è la seguente:

DEFINIZIONE 1.2.1. Consideriamo due insiemi S e T . Il *prodotto cartesiano* di S e T è:

$$S \times T = \{ (s, t) \mid s \in S, t \in T \}.$$

ESEMPIO 1.2.2. Il prodotto cartesiano che considereremo molto spesso è $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$, detto *piano reale*, che è l'insieme formato dalle coppie (a, b) di numeri reali. Vedremo più avanti che è conveniente scrivere queste coppie in verticale, cioè

$$\begin{pmatrix} a \\ b \end{pmatrix}$$

invece che in orizzontale. Come si può fare la somma di due numeri reali, così si possono sommare le coppie di numeri reali, facendo la somma componente per componente: se $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ e $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ sono due coppie qualsiasi di numeri reali, allora si definisce

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \end{pmatrix}.$$

□

ESEMPIO 1.2.3. Consideriamo i punti $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ e $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ di \mathbb{R}^2 . Allora la loro somma è:

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

□

Come già sapete, si è soliti rappresentare gli elementi di \mathbb{R}^2 usando gli assi cartesiani. Nel prossimo capitolo 2 torneremo subito su questo esempio.

Più in generale si può considerare il prodotto cartesiano di \mathbb{R} per se stesso un numero finito n di volte:

$$\underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ fattori}} = \mathbb{R}^n$$

che è un insieme i cui elementi sono le n -uple di numeri reali. Per $n = 3$, si ottiene \mathbb{R}^3 che è chiamato *spazio reale*. Per ogni n , si potrà associare ad una n -pla di numeri reali un *vettore*, ottenendo così uno spazio vettoriale di *dimensione* n . In questi spazi vettoriali si potranno fare le stesse operazioni di somma e moltiplicazioni per scalari come nel piano.

1.3. Appendice: i numeri naturali, gli interi, i razionali ed i reali

Partiamo da un solo concetto *primitivo*, cioè non definito, quello di insieme. Un sottoinsieme E_- di un insieme E_+ è un insieme i cui elementi sono tutti contenuti in E_+ ; se E_- non coincide con E_+ si dice che è un *sottoinsieme proprio*.

Definizione assiomatica dei numeri naturali: assiomi di Peano. I numeri naturali sono un insieme \mathbb{N} che verifica i cinque assiomi seguenti, introdotti da Peano:

P1. Esiste un elemento in \mathbb{N} che denotiamo con 1.

P2. Per ogni elemento n di \mathbb{N} esiste un altro elemento n' di \mathbb{N} che designiamo come il *successivo* di n .

P3. 1 non è il successivo di nessun altro numero naturale.

P4. Se $n \neq m$ allora $n' \neq m'$.

P5. (**Assioma di induzione.**) Se una proprietà è verificata dal numero 1, ed è tale che, se è verificata da n allora è verificata dal suo successivo n' , allora essa è verificata da tutti gli interi.

Grazie all'assioma di induzione, possiamo definire due operazioni su \mathbb{N} , come segue.

DEFINIZIONE 1.3.1. (**L'operazione di somma su \mathbb{N} .**) Definiamo $n + 1 = n'$, e poi induttivamente $n + m' = (n + m)'$. Per l'assioma di induzione, questo definisce la somma per ogni coppia di naturali.

DEFINIZIONE 1.3.2. (**L'operazione di moltiplicazione su \mathbb{N} .**) Definiamo $n \cdot 1 = n$, e poi induttivamente $n \cdot m' = (n \cdot m) + 1$ (in altre parole, $n \cdot (m + 1) = (n \cdot m) + 1$). Per l'assioma di induzione, questo definisce la moltiplicazione per ogni coppia di naturali.

È facile verificare che valgono le proprietà seguenti.

A1. (**Proprietà associativa della somma:** $m + (n + k) = (m + n) + k$ per ogni $m, n, k \in \mathbb{N}$.)

A2. (**Proprietà commutativa della somma:** $m + n = n + m$ per ogni $m, n \in \mathbb{N}$.)

M1. (**Proprietà associativa della moltiplicazione:** $m \cdot (n \cdot k) = (m \cdot n) \cdot k$ per ogni $m, n, k \in \mathbb{N}$.)

M2. (**Proprietà commutativa della moltiplicazione:** $m \cdot n = n \cdot m$ per ogni $m, n \in \mathbb{N}$.)

M3. (**Esistenza dell'elemento neutro per la moltiplicazione.**) $n \cdot 1 = n$ per ogni $n \in \mathbb{N}$.

D. (**Proprietà distributiva della somma rispetto alla moltiplicazione:** $n \cdot (m + k) = (n \cdot m) + (n \cdot k)$ per ogni $m, n, k \in \mathbb{N}$.)

Vogliamo costruire in maniera concreta una copia degli interi. Per questo scopo dobbiamo esporre vari argomenti preliminari.

DEFINIZIONE 1.3.3. (**Prodotto cartesiano.**) Dati due insiemi A e B , il *prodotto cartesiano* $A \times B$ è l'insieme delle coppie ordinate $(a, b) : a \in A, b \in B$.

Si noti che il prodotto cartesiano $\mathbb{R} \times \mathbb{R}$ viene in tal modo visualizzato geometricamente come un piano: ciascun punto del piano corrisponde infatti, in modo unico, ad una coppia di coordinate. Analogamente, $\mathbb{Z} \times \mathbb{Z}$ è il reticolo dei punti con entrambe le coordinate intere. Prendendo il prodotto cartesiano di tre copie di \mathbb{R} si ottiene lo spazio tridimensionale, ed analogamente per il reticolo tridimensionale degli interi.

NOTA 1.3.4. È ovvio dalla definizione che il prodotto cartesiano è associativo, ma non commutativo. \square

DEFINIZIONE 1.3.5. (**Relazione di equivalenza.**) Dato un insieme E , un sottoinsieme $R \in E \times E$ si dice una *relazione di equivalenza* se valgono le seguenti tre proprietà:

(i) **riflessiva:** $(a, a) \in R$ per ogni $a \in E$

(ii) **simmetrica:** $(a, b) \in R$ se $(b, a) \in R$

(iii) **transitiva:** se $(a, b) \in R$ e $(b, c) \in R$ allora anche $(a, c) \in R$

Se $(a, b) \in R$ scriviamo $a \sim b$. Con questa notazione le tre proprietà sopra elencate diventano:

(i) $a \sim a \quad \forall a \in E$

(ii) **simmetrica:** $a \sim b \Rightarrow b \sim a$

(iii) **transitiva:** $a \sim b$ e $b \sim c \Rightarrow a \sim c$

I sottoinsiemi di E consistenti di elementi mutuamente equivalenti (questa è una definizione ben posta grazie alla proprietà transitiva!) si chiamano le *classi di equivalenza* di R . Gli elementi di una classe di equivalenza si chiamano *rappresentanti* della classe.

DEFINIZIONE 1.3.6. (**Funzioni.**) Una funzione (o *mappa*) f da un insieme A ad un insieme B è un sottoinsieme di $A \times B$ tale che non ci siano in f due coppie con lo stesso primo elemento.

Invece di scrivere $(a, b) \in f$ è consuetudine scrivere $f = f(a)$. Questa notazione ispira l'interpretazione abituale, che identifica la funzione f con una *legge* che ad ogni *variabile* $a \in A$ associa un *valore* $b \in B$ (ed uno solo: questo è cio

‘o che vuol dire che non ci sono due coppie diverse con lo stesso primo elemento). In questa interpretazione, e rammentando la visualizzazione data poco sopra del prodotto cartesiano $A \times B$ come un *piano* generato dagli *assi* A e B , la nozione da noi introdotta di funzione si identifica con quella del suo grafico. Il fatto che non ci siano due coppie con lo stesso primo elemento significa allora che per ogni variabile la funzione ha al più un solo valore, cioè che il grafico non interseca più di una volta nessuna *retta verticale*.

ESEMPIO 1.3.7. (**Equipotenza.**) Introduciamo sulla famiglia di tutti gli insiemi una relazione di equivalenza, che chiamiamo *equipotenza*, nel modo seguente: due insiemi A e B sono equipotenti se esiste una funzione da A a B biunivoca, cioè tale che per ogni $a \in A$ esiste un $b \in B$ (e necessariamente uno solo) tale che $f(a) = b$, e per ogni $b \in B$ esiste un $a \in A$ (ed uno solo) tale che $b = f(a)$.

Lasciamo al lettore la facile verifica del fatto che valgono le proprietà della Definizione 1.3.5. □

NOTA 1.3.8. Il lettore che ha già sviluppato un'intuizione sicura dell'insieme dei numeri interi, che fra bpoco definiremo in maniera rigorosa, può osservare che gli interi sono equipotenti ad alcuni loro sottoinsiemi propri! In effetti, chiamiamo \mathbb{N} l'insieme degli interi e con \mathbb{P} il sottoinsieme proprio degli interi pari: allora la mappa $n \mapsto 2n$ è una equipotenza fra \mathbb{N} e \mathbb{P} . Questo fatto motiva la prossima definizione. □

DEFINIZIONE 1.3.9. (**Insiemi finiti.**) Un insieme si dice *finito* se non è equipotente ad alcun suo sottoinsieme proprio.

Ora introduciamo la **costruzione dei naturali**.

DEFINIZIONE 1.3.10. I numeri naturali sono le classi di equivalenza degli insiemi finiti rispetto alla relazione di equipotenza. L'insieme dei numeri naturali si indica con \mathbb{N} . Il numero naturale n dato dalla classe di equivalenza di un insieme A si chiama la *cardinalità* di A .

Questa definizione chiarisce cosa sia l'insieme degli interi, ma non ci dice come costruire un singolo intero. Essa dice che i numeri naturali possono essere considerati come la proprietà che hanno in comune tutti gli insiemi con lo stesso numero di elementi, ma non avrebbe senso dire che un dato numero naturale n può essere considerato come la proprietà che hanno in comune tutti gli insiemi con lo stesso numero n di elementi: questa è una tautologia, perché per spigare cosa sia n fa riferimento a n stesso.

La costruzione rigorosa è la seguente. Il numero 0 è la classe di equivalenza dell'insieme vuoto (quello che non contiene elementi: si noti che questa classe contiene il solo insieme vuoto). Un insieme si chiama un *singleton* se l'unico suo sottoinsieme proprio è l'insieme vuoto. Tutti i singleton sono nella stessa classe di equivalenza (esercizio!), e tale classe è il numero 1.

L'operazione di somma su \mathbb{N} . Consideriamo ora un insieme finito A e sia n la sua cardinalità: se B è un singleton disgiunto da A , diciamo che la cardinalità di $A \cup B$ è $n + 1$. Questo definisce la somma fra un intero ed il numero 1. C'è una naturale estensione di questo concetto a tutte le coppie di interi (cioè una funzione da $\mathbb{N} \times \mathbb{N} \mapsto \mathbb{N}$), ed è l'unica definizione di somma che rispetta la proprietà associativa: $n + 2 = n + (1 + 1) = (n + 1) + 1$, e così via. È facile ed intuitivo vedere che la somma verifica anche la proprietà commutativa. Per una esposizione più rigorosa di questa e delle successive costruzioni dovremmo utilizzare l'assioma di induzione: lasciamo la dimostrazione rigorosa come esercizio.

L'operazione di moltiplicazione su \mathbb{N} . La moltiplicazione è una funzione $\mathbb{N} \times \mathbb{N} \mapsto \mathbb{N}$ che verifica la proprietà distributiva del prodotto rispetto alla somma. la sua costruzione induttiva è la seguente: $2n = (1 + 1)n = n + n = n \cdot 2$, $3n = (1 + 1 + 1)n = n + n + n = n \cdot 3$, e così via. È facile vedere che la moltiplicazione è commutativa, associativa e distributiva rispetto alla somma.

DEFINIZIONE 1.3.11. (**L'ordinamento di \mathbb{N} .**) Diciamo che due numeri naturali n , m verificano la disuguaglianza $n < m$ se esiste $k \in \mathbb{N}$ tale che $m = n + k$.

È facile vedere che la somma ed il prodotto rispettano l'ordinamento su \mathbb{N} :

O1. Se $n < m$ allora $n + k < m + k$ per ogni $k \in \mathbb{N}$.

O2. Se $n < m$ allora $n \cdot k < m \cdot k$ per ogni $k \in \mathbb{N}$.

Più in generale introduciamo la seguente definizione.

DEFINIZIONE 1.3.12. (**Relazione d'ordine.**) Dato un insieme E , un sottoinsieme $R \subseteq E \times E$ si dice una *relazione d'ordine*, o un *ordinamento*, se valgono le seguenti tre proprietà:

- (i) **riflessiva:** $(a, a) \in R$ per ogni $a \in E$
- (ii) **antisimmetrica:** se $(a, b) \in R$ e $(b, a) \in R$ allora $a = b$

(iii) **transitiva:** se $(a, b) \in R$ e $(b, c) \in R$ allora anche $(a, c) \in R$

Se $(a, b) \in R$ scriviamo $a \preceq b$. Con questa notazione le tre proprietà sopra elencate diventano:

- (i) $a \preceq a \ \forall a \in E$
- (ii) $a \preceq b$ e $b \preceq a \Rightarrow a = b$
- (iii) $a \preceq b$ e $b \preceq c \Rightarrow a \preceq c$

Ovviamente, l'ordinamento dei numeri naturali (ed in seguito degli interi, o dei razionali, o dei reali) è un esempio.

Abbiamo definito la somma di due numeri naturali. Non possiamo invece definire la differenza fra due naturali m e n altro che se $n \leq m$. Sotto questa ipotesi, segue direttamente dalla definizione di ordinamento che esiste un naturale k (ed uno solo) tale che $n + k = m$: il numero k si chiama la differenza fra m e n , e si scrive $k = m - n$. È anche facile verificare che la differenza, nei casi in cui esiste, verifica le proprietà associativa, commutativa e distributiva rispetto alla moltiplicazione. Ma come osservato, essa non esiste per tutte le coppie in \mathbb{N} .

Introduciamo una estensione \mathbb{Z} di \mathbb{N} a cui sia possibile estendere l'operazione di differenza in modo che si applichi a tutte le coppie: \mathbb{Z} si chiama l'insieme degli interi (o interi con segno, o interi relativi).

Visualizzazione geometrica dell'operazione di differenza. Per prima cosa diamo una visualizzazione geometrica della differenza, in termini del prodotto cartesiano $\mathbb{N} \times \mathbb{N}$, che come osservato è un reticolo. Fissato $k \in \mathbb{N}$, per quali coppie di naturali (n, m) , cioè per quali punti del reticolo, si ha che $n - m = k$? Esattamente per i punti che giacciono sulla (semi)retta di pendenza 1 che interseca l'asse orizzontale al punto k . Anzi, richiedere che la differenza $n - m$ sia la stessa per diverse coppie di naturali è una relazione di equivalenza, le cui classi di equivalenza sono i tali rette. Allora è immediato immergere \mathbb{N} in questo insieme di rette: ogni k corrisponde alla retta a pendenza 1 che passa per $(0, k)$. L'operazione di somma si trasporta a queste classi di equivalenza. Infatti possiamo definire la somma di due rette come la retta che si ottiene sommando come vettori i loro punti, cioè l'insieme di tutte le coppie $(n_1 + n_2, m_1 + m_2)$ dove (n_1, m_1) appartiene alla prima retta e (n_2, m_2) alla seconda. Quest'insieme è ancora una retta perché, se $m_1 - n_1 = k_1$ e $m_2 - n_2 = k_2$, allora $(m_1 + m_2) - (n_1 + n_2) = k_1 + k_2$ non dipende dalla scelta del singolo punto sulla prima o sulla seconda retta, ma solo da k_1 e k_2 . In altre parole, la differenza (ed analogamente la somma) sono invarianti sulle classi di equivalenza date dalle rette, cioè indipendenti dai rappresentanti della classe di equivalenza (diciamo che sono *ben definite* sulle classi).

NOTA 1.3.13. Una presentazione più elegante, ma meno trasparente, avrebbe evitato l'uso del segno meno che abbiamo utilizzato nello scrivere la relazione di equivalenza $(n_1, m_1) \sim (n_2, m_2) \Leftrightarrow m_1 - n_1 = m_2 - n_2$. Avremmo infatti potuto, equivalentemente, scrivere $(n_1, m_1) \sim (n_2, m_2) \Leftrightarrow m_1 + n_2 = n_1 + m_2$. Le classi di equivalenza della relazione si scrivono anch'esse senza ricorrere al segno meno, in questo modo: la classe di equivalenza del numero naturale k è la semiretta $\{k + n, n\}$. □

Estensione dai naturali \mathbb{N} agli interi \mathbb{Z} . A questo punto è facile estendere \mathbb{N} ad un insieme \mathbb{Z} dove la differenza sia definita per ogni coppia di elementi. Abbiamo visto che \mathbb{N} è realizzabile come l'insieme delle (semi)rette del reticolo $V \times \mathbb{N}$ a pendenza 1 che tagliano le ascisse in punti non a sinistra di 0, cioè delle semirette che giacciono nel primo quadrante. Consideriamo l'insieme, più grande, di tutte le rette, o semirette, semirette a pendenza 1. e operazioni di somma e di differenza si estendono come prima, e come prima dipendono solo dalle rette ma non dalla scelta dei singoli punti in esse. Questo insieme di rette lo chiamiamo \mathbb{Z} . Se una semiretta non giace nel primo quadrante, essa interseca l'asse orizzontale in un punto a sinistra dell'origine. Il primo di tali punti lo indichiamo con -1 , il secondo con -2 e così via: osserviamo che i punti (n, m) della retta che passa per $-k$ sono quelli per cui $n - m = k$. Osserviamo anche che le rette passanti per $-k$ e per k hanno per somma la retta passante per 0, perché se $n_1 - m_1 = k$ e $m_2 - n_2 = k$ allora $(m_1 + m_2) - (n_1 + n_2) = 0$. In questo stesso modo vediamo che tutte le proprietà della somma si estendono a \mathbb{Z} .

Anche la moltiplicazione è ben definita sulle classi di equivalenza date dalla semirette: osserviamo che, se $m_1 - n_1 = k_1$ e $m_2 - n_2 = k_2$, allora $k_1 k_2 = (m_1 - n_1)(m_2 - n_2) = m_1 m_2 + n_1 n_2 - m_1 n_2 - n_1 m_2$. Quindi definiamo la moltiplicazione delle due rette di pendenza 1 che contengono rispettivamente i punti (n_1, m_1) e (n_2, m_2) come la retta che contiene il punto $(m_1 m_2 + n_1 n_2, m_1 n_2 + n_1 m_2)$. Si verifica facilmente che questa definizione di moltiplicazione è ben posta: essa dipende solo dalle classi di equivalenza, le rette, e non dalla scelta dei rappresentanti usati per formularla (esercizio).

Ora è facile, anche se lungo, verificare che le proprietà: associativa e commutativa della somma e della moltiplicazione, l'esistenza dell'elemento neutro della moltiplicazione, la proprietà distributiva della somma rispetto alla moltiplicazione e la prima proprietà di ordinamento **O1** valgono in \mathbb{Z} ; invece la seconda proprietà di ordinamento **O2** si estende in questo modo:

O2. Se $n < m$ allora $n \cdot k < m \cdot k$ per ogni $k > 0 \in \mathbb{Z}$, $m \cdot k < n \cdot k$ per ogni $k < 0 \in \mathbb{N}$, e $n \cdot 0 = 0$ per ogni $n \in \mathbb{Z}$.

In particolare, segue dalla proprietà **O2** che 0 *non ha divisori*: se $nm = 0$ allora uno fra n e m deve essere zero, perché altrimenti il loro prodotto sarebbe o strettamente positivo o strettamente negativo.

Inoltre, è facile vedere che valgono altre due proprietà della somma:

A3. (Esistenza dell'elemento neutro per la somma.) Esiste un numero $0 \in \mathbb{Z}$ tale che $n + 0 = n$ per ogni $n \in \mathbb{Z}$.

A4. (Esistenza dell'opposto.) Per ogni $n \in \mathbb{Z}$ esiste un $m \in \mathbb{Z}$ tale che $n + m = 0$ (si scrive $m = -n$).

Infatti, l'elemento 0 corrisponde alla classe di equivalenza della retta bisettrice: $\{(n, n), n \in \mathbb{Z}\}$, perché se (k, l) sta in un'altra retta, allora $(k+n, l+n) \sim (k, l)$ perché $(k+n)+l = (l+n)+k$ per le proprietà associative e commutative della somma. Questo prova la proprietà **A3**. Per provare **A4**, data una coppia (n, m) corrispondente a qualche $k \in \mathbb{Z}$, una coppia appartenente

alla retta corrispondente a $-k$ è (m, n) , dal momento che $(n, m) + (m, n) = (n + m, n + m)$ e l'ultima coppia appartiene alla semiretta associata a 0.

NOTA 1.3.14. L'elemento 0 è unico. Infatti, se ce ne fossero due (chiamiamoli 0_1 e 0_2), allora si avrebbe $0_1 = 0_1 + 0_2 = 0_2 + 0_1 = 0_2$, per la proprietà commutativa della somma. \square

Analogamente, l'opposto è unico: più in generale, vale la legge di cancellazione seguente.

COROLLARIO 1.3.15. *Per ogni $n, m \in \mathbb{Z}$ esiste un unico $k \in \mathbb{Z}$ tale che $n + k = m$.*

DIMOSTRAZIONE. È facile verificare che k esiste: basta porre $k = m - n$, la verifica è immediata a partire dalle quattro proprietà della somma. Se ci fossero due diversi k_1 e k_2 con la proprietà dell'enunciato, avremmo $n + k_1 = m = n + k_2$, e sommando $-n$ ad entrambi i membri si trova, di nuovo dalle proprietà associative e di esistenza degli opposti, che $k_1 = k_2$. \square

Un insieme con una operazione di somma che verifica le proprietà **A1**, **A2**, **A3** ed **A4** si chiama un *gruppo commutativo*. La regola di cancellazione vale, per questo argomento, in tutti i gruppi commutativi.

Usando le proprietà che abbiamo introdotto si possono provare i risultati cruciali dell'aritmetica: ad esempio l'esistenza e l'unicità della scomposizione in fattori primi (*Teorema fondamentale dell'aritmetica*), le proprietà del massimo comun divisore e del minimo comune multiplo, il teorema della divisione con resto (per ogni $m, n \in \mathbb{N}$ esistono $q \in \mathbb{N}$ e $0 \leq r < m$ tali che $n = qm + r$, e sono unici), l'algoritmo euclideo per la determinazione del massimo comun divisore tramite applicazioni iterate della divisione con resto, e le proprietà delle congruenze modulo un naturale N .

L'operazione di divisione ed i numeri razionali. Abbiamo introdotto in \mathbb{Z} una moltiplicazione, ma non l'operazione inversa, la divisione. È chiaro che la struttura ordinata e sequenziale degli interi, derivante dall'assioma di Peano P5, impedisce di trovare in \mathbb{Z} il reciproco di ogni elemento, e quindi di definire la divisione: infatti i reciproci dei numeri maggiori di 1 dovrebbero essere compresi fra 0 e 1, ma abbiamo visto che non esistono interi con questa proprietà.

Per definire la divisione dobbiamo estendere gli interi ad un insieme più grande, quello dei razionali \mathbb{Q} . Vogliamo definire il reciproco di qualunque numero intero non nullo, cioè vogliamo definire le frazioni $\frac{m}{n}$ se $n \neq 0$. A questo fine definiamo una nuova relazione di equivalenza sulle coppie (n, m) con $m \neq 0$: due tali coppie sono equivalenti *per dilatazione* se esiste $k \in \mathbb{Z}$, $k \neq 0$, tali che $(n_2, m_2) = (kn_1, km_1)$. Preferiamo riformulare questa relazione nel seguente modo equivalente: $(n_1, m_1) \approx (n_2, m_2)$ se $n_1 m_2 = m_1 n_2$. È facile verificare che si tratta di una relazione di equivalenza. Le classi di equivalenza sono i punti del reticolo $\mathbb{Z} \times \mathbb{Z}$ allineati radialmente rispetto all'origine: in altre parole, una classe di equivalenza corrisponde ai punti del reticolo sulla stessa retta passante per l'origine. Osserviamo che in ogni classe di equivalenza esiste una coppia (n, m) di distanza minima dall'origine (questa coppia è unica

se si sceglie il suo secondo elemento $m > 0$): essa è caratterizzata dalla proprietà che n e m sono relativamente primi, cioè senza fattori comuni. Quindi, quando si scrivono i razionali come frazioni, questa rappresentazione senza fattori primi è unica (assumendo il denominatore positivo); essa si chiama *rappresentazione ridotta*.

Le operazioni di somma e di moltiplicazione vengono definite, come sempre, sui rappresentanti delle classi di equivalenza. Definiamo queste operazioni in modo che ricalchino le proprietà consuete che vogliamo avere sulle frazioni:

- **Somma in $\mathbb{Q} \equiv \mathbb{N} \times \mathbb{N} \setminus \{0\}$:** $(n_1, m_1) \pm (n_2, m_2) \approx (n_1 m_2 \pm m_1 n_2, n_1 m_2)$
- **Moltiplicazione in \mathbb{Q} :** $(n_1, m_1) \cdot (n_2, m_2) \approx (n_1 n_2, m_1 m_2)$

È elementare verificare che i rappresentanti del numero 1 sono le coppie (n, n) con $n \neq 0$ (cioè la bisettrice), e quelli del numero 0 sono le coppie $0, n$ con $n \neq 0$ (cioè l'asse orizzontale). È altrettanto immediato verificare che tutte le proprietà aritmetiche di \mathbb{Z} si estendono a \mathbb{Q} . Inoltre, ogni razionale non nullo ha un reciproco:

M4. **Esistenza del reciproco in \mathbb{Q} .**) Per ogni $n, m \in \mathbb{Q}$ con $m \neq 0$ esiste uno ed un solo $k \in \mathbb{Q}$ tale che $n = mk$.

Per dimostrare M4 basta verificare che, se n è rappresentato da (j_1, l_1) e m da (j_2, l_2) , allora il razionale k rappresentato da $(j_1 l_2, j_2 l_1)$ verifica $n = mk$. L'unicità è ovvia: se $mk_1 = n = mk_2$ allora $m \cdot (k_1 - k_2) = 0$, ma l'unico divisore di 0 è 0 e quindi $k_1 - k_2 = 0$.

Un insieme dotato delle operazioni di somma e moltiplicazione con le proprietà elencate più sopra e tale che ogni elemento non nullo ha un reciproco si chiama un *campo*.

Immersione isomorfa degli interi nei razionali. Gli interi sono stati costruiti a partire dal reticolo $\mathbb{N} \times \mathbb{N}$ dei naturali come classi di equivalenza della relazione di equivalenza delle rette a pendenza 1, mentre i razionali sono stati costruiti a partire dal reticolo $\mathbb{Z} \times \mathbb{Z}$ degli interi (che estende il precedente) come classi di equivalenza della relazione di equivalenza della dilatazione. Le due relazioni sono diverse, e quindi non è ovvio che \mathbb{Z} si immerga in \mathbb{Q} preservando le operazioni aritmetiche: se questo avviene, l'immersione ϕ si chiama un *isomorfismo* di \mathbb{Z} su $\phi(\mathbb{Z})$.

Esiste una naturale immersione isomorfa. Questa immersione è data da $\phi(n) = (n, 1)$, $n \in \mathbb{Z}$. È chiaro che ϕ è iniettiva, e si verifica immediatamente che $\phi(n_1 \pm n_2) = \phi(n_1) \pm \phi(n_2)$.

L'immersione preserva l'ordinamento. La relazione d'ordine di \mathbb{Z} si estende a \mathbb{Q} . L'ordinamento su \mathbb{Q} si definisce così: se $r \in \mathbb{Q}$ è rappresentato dalla coppia (n, m) ($m \neq 0$), allora si dice che $r > 0$ se n e m sono entrambi positivi o entrambi negativi, e invece $r < 0$ se i segni di n e m sono opposti (ovviamente c'è solo un caso residuo, quello in cui $n = 0$, nel qual caso si ha $r = 0$). È facile vedere che due numeri interi verificano la relazione d'ordine $n < m$ rispetto all'ordinamento di \mathbb{Z} se e solo se le loro immagini in \mathbb{Q} verificano la stessa relazione rispetto all'ordinamento di \mathbb{Q} : $\phi(n) < \phi(m)$. Quindi l'immersione di \mathbb{Z} in \mathbb{Q} preserva l'ordinamento.

Risoluzione di equazioni quadratiche e sezioni di Dedekind in \mathbb{Q} . Abbiamo dimostrato nella Proposizione 1.1.2 che l'equazione $x^2 = 2$ non ammette soluzioni in \mathbb{Q} .

Per risolverla dobbiamo estendere il campo \mathbb{Q} dei razionali ad un campo più grande, quello dei numeri reali, che si denota con \mathbb{R} . Vediamo come. Poiché la funzione x^2 è strettamente crescente per $x \geq 0$, e $1^2 = 1 < 2 < 2^2 = 4$, in qualsiasi estensione di \mathbb{Q} che preservi l'ordinamento deve valere $1 < \sqrt{2} < 2$. Ora consideriamo le frazioni $\frac{11}{10}, \frac{12}{10}, \dots, \frac{19}{10}$. Poiché $11^2 < 12^2 < 13^2 < 14^2 = 196 < 200 < 15^2 = 225$, deve valere $1.4 = \frac{14}{10} < 2 < \frac{15}{10} = 1.5$. Continuando così abbiamo $1.41 < 2 < 1.42$, $1.414 < 2 < 1.415$, e così via. In questo modo, in ogni campo che estende \mathbb{Q} preservandone l'ordinamento e nel quale esiste la radice quadrata di 2, costruiamo una successione crescente di razionali minori di $\sqrt{2}$ ed una decrescente di razionali maggiori di $\sqrt{2}$ in cui la differenza fra il maggiorante ed il minorante di indice n è inferiore a 10^{-n} . Ora dobbiamo fare una ulteriore ipotesi, che a questo punto è chiaramente equivalente alla risolubilità dell'equazione $x^2 = 2$: nell'estensione di \mathbb{Q} non esiste alcun numero positivo minore di 10^{-n} simultaneamente per tutti gli n (*proprietà archimedeica*.) Oppure, equivalentemente, riformuliamo l'ipotesi nel modo seguente:

Assioma delle sezioni di Dedekind: Chiamiamo *classe maggiorante* in \mathbb{Q} un sottoinsieme proprio J_+ di \mathbb{Q} tale che, per ogni $r \in J_+$, tutti i $q \in \mathbb{Q}$ con $q > r$ appartengono a J_+ . In maniera simmetrica si definiscono le classi minoranti. Il complementare di J_+ è una classe minorante che indichiamo con J_- . Si osservi che abbiamo appena visto che, se $J_+ = \{r \in \mathbb{Q} : r^2 > 2\}$, allora J_+ non ha minimo in \mathbb{Q} e J_- non ha massimo, ma naturalmente ogni elemento di J_- è minore di tutti gli elementi di J_+ ed ogni elemento di J_+ maggiora tutti gli elementi di J_- . Una coppia di insiemi J_+, J_- in cui ciascun elemento del primo insieme maggiora ogni elemento del secondo e viceversa si chiama una *sezione di Dedekind*. Chiamiamo *insieme \mathbb{R} dei numeri reali* l'insieme delle sezioni di Dedekind dei numeri razionali. Estendiamo a \mathbb{R} l'ordinamento di \mathbb{Q} nel modo seguente: il numero reale corrispondente ad una sezione di Dedekind J_+, J_- è minore o uguale di ogni razionale in J_- e maggiore o uguale di ogni razionale in J_+ (esso si chiama *l'elemento separatore* della sezione). È chiaro che l'immersione di \mathbb{Q} in \mathbb{R} rispetta l'ordinamento. Data una sezione di Dedekind J_+, J_- , un numero reale minore o uguale di ogni $r \in J_+$ e maggiore o uguale di ogni $s \in J_-$ si chiama un *elemento separatore* della sezione. L'assioma di Dedekind dice che per ogni sezione esiste un unico elemento separatore in \mathbb{R} . (Per esercizio si dimostri che, in particolare, J_+ ha minimo in \mathbb{R} oppure J_- ha massimo). Quindi i numeri reali sono gli elementi separatori delle sezioni dei razionali.

Resta solo da estendere a \mathbb{R} le operazioni di somma e di moltiplicazione. La somma di due reali x_1 (associato alla sezione J_+, J_-) e x_2 (associato alla sezione I_+, I_-) è l'elemento separatore della sezione $J_+ + I_+, J_- + I_-$ (qui l'insieme somma $J_+ + I_+$ è definito come $\{x + y : x \in J_+, y \in I_+\}$). Abbiamo visto, dall'esempio di $\sqrt{2}$, che una sezione di Dedekind (cioè un numero reale) corrisponde a due successioni di approssimanti razionali, in cui tutti i numeri della prima sono minoranti di tutti quelli della seconda e tutti quelli della seconda sono maggioranti di tutti quelli della prima, e con l'ulteriore proprietà che i numeri minoranti si avvicinano a quelli maggioranti a meno di una precisione arbitrariamente piccola. Allora è naturale definire la somma trasportandola da questi approssimanti razionali, i quali, avendo

un numero finito di cifre decimali oppure uno sviluppo periodico, ci permettono di calcolare la somma colonna per colonna (con eventuale riporto). Quindi la somma di due numeri reali si approssima numericamente troncando i due reali ad approssimanti razionali con lo stesso numero di cifre decimali, sommando tali razionali e poi facendo migliorare l'approssimazione col ripetere il calcolo con via via più cifre decimali. È facile verificare che la somma così definita su \mathbb{R} ha tutte le proprietà precedentemente dimostrate su \mathbb{Q} .

La moltiplicazione si estende in modo analogo, prendendo la sezione prodotto $J_+ \cdot I_+, J_- \cdot I_-$. Qui però c'è una difficoltà tecnica che ci accenniamo. Se due numeri reali r_1 e r_2 sono entrambi non negativi, allora il prodotto $J_+ \cdot I_+$ delle loro classi maggioranti è la classe maggiorante del prodotto, e tutto procede come nel caso della somma. Ma se $r_1 < 0 < r_2$ allora la classe maggiorante è $J_+ \cdot I_-$, e se infine $r_1, r_2 < 0$ allora la classe maggiorante è $J_- \cdot I_-$. Nei vari casi la moltiplicazione si esegue sugli approssimanti razionali e poi si trasporta per approssimazioni successive, ma quali siano gli approssimanti dal di sotto e quali quelli dal di sopra dipende dalla casistica che abbiamo elencato.

CAPITOLO 2

Spazi vettoriali

Ad ogni punto $P = (x, y)$ del piano reale \mathbb{R}^2 possiamo associare un *vettore*, cioè un segmento orientato che parte dall'origine $O = (0, 0)$ del piano e arriva al punto fissato P , come in figura 1. Denotiamo questo vettore con \vec{OP} .

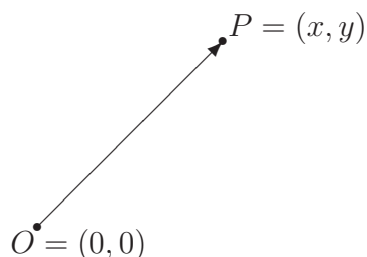


FIGURA 1. Vettore con punto iniziale O e punto finale P .

Ricordiamo che un vettore è determinato dalla sua *lunghezza* (o *modulo*), dalla sua *direzione* e dal suo *verso*. Come è ben noto, si possono sommare due vettori con la cosiddetta *regola del parallelogramma*: infatti si possono pensare i due vettori v, w come lati di un parallelogramma e la loro somma $v + w$ corrisponde alla diagonale del parallelogramma, come mostrato in figura 2.

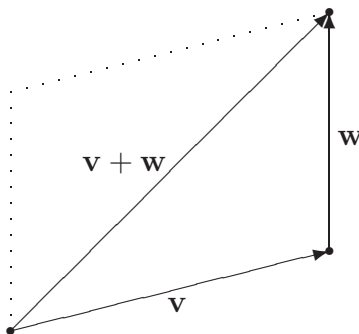


FIGURA 2. Somma di due vettori con la regola del parallelogramma.

Dato un vettore, si può considerare il suo *opposto*, che è il vettore che ha lo stesso modulo, la stessa direzione di quello dato, ma verso opposto. Inoltre si può moltiplicare un vettore per un numero reale $k > 0$, che è il vettore con la stessa direzione e lo stesso verso del vettore dato, ma con la lunghezza moltiplicata per k .

È conveniente identificare un vettore \vec{OP} con il punto del piano P , cioè con la coppia di numeri reali (x, y) che sono le coordinate di P nel piano; quindi consideriamo le proprietà dei vettori in base alle loro *coordinate*, e non direttamente in base alla lunghezza, direzione e verso del vettore.

2.1. Definizione di spazio vettoriale

DEFINIZIONE 2.1.1. Uno **spazio vettoriale** è un insieme X provvisto di due operazioni:

- la **somma**, che associa ad ogni coppia di elementi di X un terzo elemento di X

$$x, y \in X \mapsto x + y \in X,$$

chiamato *somma* di x e y ;

- la **moltiplicazione per scalari**, che ad ogni elemento di X e ad ogni numero reale associa un altro elemento di X

$$x \in X, \lambda \in \mathbb{R} \mapsto \lambda \cdot x \in X,$$

detto *moltiplicazione di x per lo scalare λ* ;

che devono soddisfare alcune proprietà che vedremo in dettaglio fra poco. Talvolta si dice spazio *lineare*, al posto di spazio *vettoriale*.

ESEMPIO 2.1.2. Nel prodotto cartesiano $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$, detto *piano reale*, formato dalle coppie (a, b) di numeri reali, possiamo definire le due operazioni di somma e moltiplicazione per uno scalare $\lambda \in \mathbb{R}$ nel seguente modo:

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} := \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \end{pmatrix}, \quad (2.1.1)$$

$$\lambda \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} := \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \end{pmatrix}, \quad (2.1.2)$$

cioè componente per componente.

Associamo ad ogni elemento $P = (a, b)$ di \mathbb{R}^2 il vettore *applicato* \vec{OP} , cioè il vettore con punto iniziale l'origine O del piano e con punto finale P , come appena visto a pag. 15. Si verifica subito che la somma definita dalla formula (2.1.1) corrisponde proprio alla somma di due vettori con la regola del parallelogramma e che la moltiplicazione per lo scalare $\lambda \in \mathbb{R}$ definita dalla formula (2.1.2) corrisponde alla moltiplicazione della lunghezza del vettore per λ , se λ è positivo, oppure alla moltiplicazione della lunghezza del vettore opposto per $-\lambda$, se λ è negativo. \square

Questa corrispondenza giustifica il termine di *spazio vettoriale* per X e la seguente:

DEFINIZIONE 2.1.3. Gli elementi di uno spazio vettoriale X sono detti **vettori**.

2.2. Proprietà della somma

DEFINIZIONE 2.2.1. La somma di due elementi di uno spazio vettoriale X deve verificare le seguenti quattro proprietà:

- (1) la **commutatività**, cioè per ogni $x, y \in X$ deve valere

$$x + y = y + x;$$

- (2) l'**associatività**, cioè per ogni $x, y, z \in X$ deve valere

$$(x + y) + z = x + (y + z);$$

- (3) esistenza dell'elemento **neutro**;

- (4) esistenza dell'**opposto**.

Queste proprietà sono molto naturali perché sono valide per tutti gli insiemi di numeri che già conoscete: interi, razionali e reali. Vedremo in futuro, però, esempi di operazioni su insiemi che non sono commutative, come per esempio la moltiplicazione di due matrici.

La proprietà commutativa ci dice semplicemente che cambiando l'ordine degli addendi, la somma non cambia.

La proprietà associativa invece ci dice che possiamo scrivere la somma di tre vettori senza parentesi, cioè possiamo scrivere:

$$x + y + z$$

senza alcuna ambiguità, invece di scrivere $(x + y) + z$, oppure $x + (y + z)$. Usando la proprietà associativa si può dimostrare che possiamo scrivere anche la somma di n vettori senza indicare alcuna parentesi:

$$x_1 + x_2 + x_3 + \cdots + x_n.$$

Per esempio, la somma di $n = 4$ vettori può essere fatta in molti modi:

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= (x_1 + x_2) + (x_3 + x_4) = ((x_1 + x_2) + x_3) + x_4 = \\ &= (x_1 + (x_2 + x_3)) + x_4 = x_1 + ((x_2 + x_3) + x_4) = \\ &= x_1 + (x_2 + (x_3 + x_4)), \end{aligned}$$

ma il risultato è sempre lo stesso, quindi è inutile distinguere con le parentesi in quale ordine eseguire l'addizione. Ricordiamo che per la proprietà commutativa possiamo anche scambiare di posto i vettori.

Quando si considerano un certo numero di vettori qualsiasi, diciamo cinque vettori, si è soliti scrivere:

$$x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5$$

dove la lettera x indica che i vettori sono indeterminati e i numeri 1, 2, 3, 4 e 5 sono detti *indici*. Per esempio 4 è l'indice dell'elemento x_4 .

Per indicare la somma dei cinque vettori si può scrivere

$$x_1 + x_2 + x_3 + x_4 + x_5$$

oppure, in maniera più compatta, si può scrivere

$$\sum_{i=1}^5 x_i,$$

dove la lettera greca \sum (sigma maiuscola) indica proprio la somma (spesso viene detta anche *sommatoria*) e i viene detto *indice* della sommatoria. Un altro modo per indicare la medesima somma è:

$$\sum_{1 \leq j \leq 5} x_j.$$

Si noti che non ha importanza quale lettera viene usata per denotare l'indice della sommatoria, anche se spesso viene usata la lettera i .

Se invece volessimo indicare, fra dodici vettori x_1, \dots, x_{12} dati, l'addizione dei vettori con indice pari, potremmo scrivere così:

$$\sum_{\substack{1 \leq i \leq 12 \\ i \text{ pari}}} x_i.$$

Torniamo alle proprietà della somma in uno spazio vettoriale e vediamo cosa vogliono dire la terza e la quarta proprietà della definizione 2.2.1.

L'esistenza dell'elemento neutro significa che esiste un elemento, che indicheremo con 0_X , tale che:

$$0_X + x = x, \quad \text{per ogni } x \in X. \quad (2.2.1)$$

Si dice allora che 0_X è l'elemento **neutro** di X . L'esistenza dell'opposto significa che per ogni elemento x di X esiste un elemento y di X tale che

$$x + y = 0_X. \quad (2.2.2)$$

Si dice che y è l'**opposto di** x e si denota di solito con $-x$.

Naturalmente, se consideriamo l'insieme dei vettori nel piano, l'elemento neutro è il vettore di lunghezza nulla, mentre l'opposto di un vettore dato, come abbiamo già accennato all'inizio del capitolo, è il vettore con stessa lunghezza, stessa direzione e verso opposto del vettore fissato.

ESEMPIO 2.2.2. Nel piano reale \mathbb{R}^2 , considerato con le operazioni definite con le formule (2.1.1) e (2.1.2), l'elemento neutro è $(0, 0) \in \mathbb{R}^2$, mentre l'opposto di (a_1, a_2) è

$$(-a_1, -a_2),$$

perché $(a_1 + a_2) + (-a_1, -a_2) = (a_1 - a_1, a_2 - a_2) = (0, 0) = 0_X$. □

ESERCIZIO 2.2.3. Dimostrare che l'elemento neutro della somma di uno spazio vettoriale è *unico*. □

SOLUZIONE. Supponiamo che y e z siano due elementi dello spazio vettoriale X che soddisfino la definizione di elemento neutro (2.2.1). Allora si ha che $y = y + z$, considerando y come elemento neutro, e che $z = z + y$, considerando invece z come elemento neutro. Ne segue allora che

$$y = y + z = z + y = z$$

come volevasi dimostrare. \square

ESERCIZIO 2.2.4. Sia x un elemento qualsiasi di uno spazio vettoriale X . Dimostrare che l'opposto di x è *unico*. \square

SOLUZIONE. Supponiamo che y e z siano due opposti di x . Per definizione di 0_X si ha che $y = y + 0_X$. Per ipotesi z è opposto di x , quindi possiamo scrivere $0_X = x + z$. Allora si ha che:

$$y = y + 0_X = y + (x + z) = (y + x) + z = 0_X + z = z$$

dove la terza uguaglianza (da sinistra) segue dalla proprietà associativa, la quarta uguaglianza dall'ipotesi che y è opposto di x e infine la quinta e ultima segue dalla definizione di 0_X .

\square

OSSERVAZIONE 2.2.5. Si noti che vale $0_X + 0_X = 0_X$, per definizione dell'elemento neutro 0_X applicata ad 0_X stesso, quindi

$$-0_X = 0_X,$$

per definizione di opposto (2.2.2).

2.3. Altre proprietà delle operazioni sui vettori

DEFINIZIONE 2.3.1. La moltiplicazione di un vettore per uno scalare in uno spazio vettoriale X deve verificare le seguenti due proprietà:

(1) l'*associatività*, cioè:

$$\beta \cdot (\alpha \cdot x) = (\beta \cdot \alpha) \cdot x, \quad \text{per ogni } \alpha, \beta \in \mathbb{R}, x \in X;$$

(2) esistenza dell'elemento neutro, che denotiamo con 1:

$$1 \cdot x = x, \quad \text{per ogni } x \in X.$$

Inoltre devono essere verificate altre due proprietà della moltiplicazione per scalari rispetto alla somma, le quali sono dette *proprietà distributive*:

- $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$, per ogni $\alpha \in \mathbb{R}$ e $x, y \in X$;
- $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$, per ogni $\alpha, \beta \in \mathbb{R}$ e $x \in X$.

La prima è detta proprietà distributiva della moltiplicazione per scalari rispetto alla somma, mentre la seconda è la proprietà distributiva della somma rispetto alla moltiplicazione per scalari.

Anche se la formulazione delle proprietà può apparire non immediatamente chiara, in verità traduce in formule proprietà che intuitivamente sono evidenti. Per esempio, se $\alpha = 2$, la proprietà distributiva della moltiplicazione per scalari rispetto alla somma ci dice semplicemente che $2(x + y) = 2x + 2y$, come è ragionevole immaginare.

ESEMPIO 2.3.2. Consideriamo il piano reale \mathbb{R}^2 come negli esempi 2.1.2 e 2.2.2. Allora è facile verificare che le operazioni somma e moltiplicazione per scalari definite dalle formule (2.1.1) e (2.1.2) soddisfano tutte le proprietà che abbiamo elencato per gli spazi vettoriali. Si può affermare quindi che \mathbb{R}^2 è uno spazio vettoriale con tali operazioni. \square

Nei prossimi tre esercizi mostriamo alcune proprietà riguardanti le operazioni di somma e prodotto per scalari che seguono facilmente dalle definizioni e dalle altre proprietà già viste, come sarà chiaro dalle dimostrazioni.

ESERCIZIO 2.3.3. Dimostrare che, per ogni x, y, z elementi di uno spazio vettoriale X , se

$$x + y = x + z \quad (2.3.1)$$

allora $y = z$. \square

SOLUZIONE. Ricordiamo che per ogni $x \in X$ esiste l'opposto di x , che indichiamo con $-x$. Sommando membro a membro $-x$, la formula (2.3.1) è equivalente alla seguente:

$$-x + x + y = -x + x + z. \quad (2.3.2)$$

Il primo membro diventa:

$$-x + (x + y) = (-x + x) + y = 0_X + y = y,$$

perché la prima uguaglianza segue dalla proprietà associativa dell'addizione, la seconda uguaglianza segue dalla definizione di opposto, mentre l'ultima uguaglianza segue dalla definizione di elemento neutro 0_X .

Similmente il secondo membro di (2.3.2) diventa:

$$-x + (x + z) = (-x + x) + z = 0_X + z = z.$$

Concludiamo quindi che $y = z$, come richiesto. \square

ESERCIZIO 2.3.4. Dimostrare che $0 \cdot x = 0_X$, per ogni $x \in X$. \square

SOLUZIONE. Ricordiamo che possiamo scrivere il numero reale 0 come $0 = 0 + 0$, quindi

$$0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x \quad (2.3.3)$$

dove la seconda uguaglianza segue dalla proprietà distributiva dell'addizione rispetto alla moltiplicazione per scalari. Possiamo riscrivere il primo membro così:

$$0 \cdot x = 0_X + 0 \cdot x,$$

per definizione di elemento neutro, così la formula (2.3.3) diventa:

$$0_X + 0 \cdot x = 0 \cdot x + 0 \cdot x$$

e si conclude applicando l'esercizio 2.3.3. □

ESERCIZIO 2.3.5. Dimostrare che $(-1) \cdot x = -x$, per ogni $x \in X$. □

SOLUZIONE. Ricordando che $x = 1 \cdot x$ e la proprietà distributiva, si ha che:

$$x + (-1) \cdot x = 1 \cdot x + (-1) \cdot x = (1 - 1) \cdot x = 0 \cdot x = 0_X$$

dove l'ultima uguaglianza segue dall'esercizio 2.3.4. Si conclude allora per l'unicità dell'opposto di x . □

2.4. Combinazioni lineari di vettori

DEFINIZIONE 2.4.1. Nello spazio vettoriale X consideriamo due vettori x_1 e x_2 . Dati due scalari $\alpha_1, \alpha_2 \in \mathbb{R}$, si dice che:

$$\alpha_1 x_1 + \alpha_2 x_2,$$

che si può scrivere anche così:

$$\sum_{i=1}^2 \alpha_i x_i,$$

è **combinazione lineare** dei due vettori dati. Gli scalari α_1 e α_2 sono detti **coefficienti** di x_1 e x_2 , rispettivamente.

Per esempio se $\alpha_1 = 2$ e $\alpha_2 = 3$, allora

$$2x_1 + 3x_2$$

è una combinazione lineare di x_1 e x_2 .

ESEMPIO 2.4.2. Siano $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$ e $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ elementi del piano reale \mathbb{R}^2 . Calcoliamo la combinazione lineare di questi due vettori con coefficienti rispettivamente -1 e -2 :

$$-1 \begin{pmatrix} 2 \\ -1 \end{pmatrix} - 2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \cdot 2 - 2 \cdot 0 \\ -1 \cdot (-1) - 2 \cdot 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \end{pmatrix}.$$

□

Le combinazioni lineari si possono fare anche per tre o più vettori:

DEFINIZIONE 2.4.3. Se x_1, \dots, x_n sono n vettori dati e $\alpha_1, \dots, \alpha_n$ sono n scalari (tanti quanti i vettori), allora

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$$

è la **combinazione lineare** di x_1, \dots, x_n con coefficienti $\alpha_1, \dots, \alpha_n$, rispettivamente, che possiamo scrivere anche:

$$\sum_{i=1}^n \alpha_i x_i.$$

Se consideriamo un solo vettore x e un qualsiasi scalare α , allora il *multiplo* αx di x è anch'esso detto combinazione lineare di x .

ESEMPIO 2.4.4. Consideriamo $\begin{pmatrix} 1 \\ -1 \end{pmatrix} \in \mathbb{R}^2$. Allora:

$$\alpha \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} \alpha \\ -\alpha \end{pmatrix}$$

è un multiplo di $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ per ogni $\alpha \in \mathbb{R}$. □

ESEMPIO 2.4.5. Consideriamo i seguenti tre elementi di \mathbb{R}^2 :

$$\begin{pmatrix} -3 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

La combinazione lineare di questi tre elementi con coefficienti rispettivamente 0, 1 e -1 è

$$0 \begin{pmatrix} -3 \\ 0 \end{pmatrix} + \begin{pmatrix} -2 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 - 1 \\ 1 - 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \end{pmatrix}.$$

□

2.5. Sistemi di vettori linearmente dipendenti e indipendenti

Consideriamo un vettore non nullo $x \in X$. Nella definizione 2.4.3 abbiamo visto che i multipli αx di x sono combinazioni lineari di x .

DEFINIZIONE 2.5.1. Fissato un vettore non nullo $x \in X$, un vettore $y \in X$ che non si può scrivere nella forma $y = \alpha x$, per nessun $\alpha \in \mathbb{R}$, è detto **linearmente indipendente** da x .

Per esempio il vettore $y = \begin{pmatrix} -3 \\ 0 \end{pmatrix} \in \mathbb{R}^2$ non è linearmente indipendente dal vettore $x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, perché $y = -3x$ (cfr. esempio 2.4.4).

Il vettore $z = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, invece, è linearmente indipendente da x perché non si può scrivere come $z = \alpha x$ per nessun $\alpha \in \mathbb{R}$. Infatti, supponiamo per assurdo che esista $\alpha \in \mathbb{R}$ tale che $z = \alpha x$. Allora dovrebbe essere:

$$\begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha \\ -\alpha \end{pmatrix},$$

cioè $\alpha = 2$, per l'uguaglianza della prima coordinata, e $-\alpha = 0$, per l'uguaglianza della seconda, e troviamo così una contraddizione (cfr. ancora esempio 2.4.4).

OSSERVAZIONE 2.5.2. Se rappresentiamo i vettori nel piano reale come usuale, i multipli di un vettore dato x stanno sulla retta con la stessa direzione del vettore x . Quindi un vettore y linearmente indipendente da x è un vettore che non sta su tale retta.

DEFINIZIONE 2.5.3. Siano x e y due vettori di uno spazio vettoriale X . Si dice che x e y sono **linearmente dipendenti** se non sono linearmente indipendenti, cioè se x è un multiplo di y , o viceversa y è multiplo di x .

Se y è il vettore nullo 0_X , allora per ogni $x \in X$ i due vettori x e y sono linearmente dipendenti, perché possiamo scrivere $y = 0_X = 0x$. In altri termini il vettore nullo è multiplo di qualsiasi altro vettore.

OSSERVAZIONE 2.5.4. Se x e y sono vettori non nulli, allora x è multiplo di y se e solo se y è multiplo di x . Infatti, se x è multiplo di y , allora esiste $\alpha \in \mathbb{R}$ tale che $x = \alpha y$, con $\alpha \neq 0$, perché $x \neq 0$. Allora anche y è multiplo di x , perché dividendo per α troviamo che $y = \frac{1}{\alpha}x$.

DEFINIZIONE 2.5.5. Siano x e y due vettori dello spazio vettoriale X . Si dice che un vettore z è **linearmente indipendente** da x e y se non esistono $\alpha, \beta \in \mathbb{R}$ tali che

$$z = \alpha x + \beta y. \quad (2.5.1)$$

In caso contrario, cioè se esistono siffatti α e β , si dice che z è **linearmente dipendente** da x e y .

Osserviamo che l'equazione (2.5.1) si può riscrivere nel modo seguente:

$$\alpha x + \beta y - z = 0_X,$$

cioè abbiamo trovato una combinazione lineare di x , y e z , con coefficienti non tutti nulli, la cui somma è il vettore nullo. Viceversa, se esiste una combinazione lineare di x , y e z

$$\delta x + \xi y + \gamma z = 0, \quad \text{tale che } \gamma \neq 1$$

allora, dividendo per γ , si trova che

$$\frac{\delta}{\gamma} x + \frac{\xi}{\gamma} y + z = 0$$

che possiamo riscrivere come:

$$z = \alpha x + \beta y, \quad \text{dove } \alpha = -\frac{\delta}{\gamma}, \quad \beta = -\frac{\xi}{\gamma}$$

cioè z è combinazione lineare di x e y e quindi è linearmente dipendenti da essi.

ESEMPIO 2.5.6. Siano $x = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ e $y = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ due elementi di \mathbb{R}^2 . Il vettore $z = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ è linearmente dipendente da x e y perché $z = x + y$. \square

Più in generale, la nozione di dipendenza lineare si può dare anche per più di due vettori:

DEFINIZIONE 2.5.7. Consideriamo n vettori x_1, \dots, x_n di uno spazio vettoriale X . Si dice che x_1, \dots, x_n sono **linearmente indipendenti** se non esiste una combinazione lineare di x_1, \dots, x_n , a coefficienti non tutti nulli, tale che la loro somma sia il vettore nullo 0_X . In altri termini, x_1, \dots, x_n sono linearmente indipendenti se

$$\lambda_1 x_1 + \dots + \lambda_n x_n = 0_X \implies \lambda_1 = \dots = \lambda_n = 0.$$

Con la prossima proposizione verifichiamo che la definizione 2.5.7 è concorde con le precedenti definizioni 2.5.1 e 2.5.5.

PROPOSIZIONE 2.5.8. *Siano x_1, \dots, x_n vettori di uno spazio vettoriale X . Allora x_1, \dots, x_n sono linearmente indipendenti se e solo se non esiste alcun x_k , con $1 \leq k \leq n$, che sia combinazione lineare dei rimanenti.*

DIMOSTRAZIONE. Supponiamo che esista x_k che sia combinazione lineare dei rimanenti. Per semplicità, poniamo che sia $k = n$. Allora esistono $\alpha_1, \dots, \alpha_{n-1} \in \mathbb{R}$ tali che

$$x_n = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_{n-1} x_{n-1}.$$

Riscrivendo la formula precedente nel modo seguente:

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_{n-1} x_{n-1} - x_n = 0,$$

cioè abbiamo trovato una combinazione lineare con coefficienti non tutti nulli che dà 0_X , quindi x_1, \dots, x_n sono linearmente dipendenti secondo la definizione (2.5.7).

Viceversa, se x_1, \dots, x_n sono linearmente dipendenti, esiste almeno una combinazione lineare

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_{n-1} x_{n-1} + \alpha_n x_n = 0,$$

con almeno un coefficiente non nullo, che supponiamo per semplicità essere α_n . Allora possiamo dividere tutto per α_n e portare x_n dall'altro membro:

$$x_n = -\frac{\alpha_1}{\alpha_n} x_1 - \frac{\alpha_2}{\alpha_n} x_2 + \dots - \frac{\alpha_{n-1}}{\alpha_n} x_{n-1},$$

cioè abbiamo scritto x_n come combinazione lineare dei rimanenti. □

Riordinando i vettori se necessario possiamo riscrivere l'enunciato precedente in questo modo:

COROLLARIO 2.5.9. *Siano x_1, \dots, x_n vettori di uno spazio vettoriale X . Allora x_1, \dots, x_n sono linearmente indipendenti se e solo se non esiste alcun x_k , con $1 \leq k \leq n$, che sia combinazione lineare dei vettori precedenti x_1, \dots, x_{k-1} .*

ESERCIZIO 2.5.10. Dimostrare che i vettori $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ e $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ di \mathbb{R}^2 sono linearmente indipendenti usando la definizione 2.5.7.

Svolgimento. Consideriamo una combinazione lineare dei due vettori fissati che dia 0_X :

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

Allora deve essere

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

quindi l'uguaglianza della prima coordinata implica che $\lambda_1 = 0$ e quella della seconda che $0 = -\lambda_1 + \lambda_2 = \lambda_2$. Abbiamo così dimostrato che l'unica combinazione lineare dei due vettori

fissati che dà 0_X è quella con tutti i coefficienti nulli, quindi i vettori dati sono linearmente indipendenti per la definizione 2.5.7. \square

2.6. Sottospazi vettoriali ed insiemi di generatori

DEFINIZIONE 2.6.1. Sia X uno spazio vettoriale. Un sottoinsieme Y di X si dice un **sottospazio vettoriale** di X se valgono le seguenti due condizioni:

- (1) per ogni $y, z \in Y$, si ha $y + z \in Y$;
- (2) per ogni $y \in Y$ e $\lambda \in \mathbb{R}$, si ha $\lambda y \in Y$.

In particolare dalla seconda condizione segue che $0_X \in Y$.

In altri termini, un sottospazio vettoriale Y di X è un sottoinsieme di X in cui si possono fare le operazioni di somma e moltiplicazione per scalari definite in X senza uscire da Y .

ESEMPIO 2.6.2. Consideriamo il vettore $\begin{pmatrix} 1 \\ -1 \end{pmatrix} \in \mathbb{R}^2$ e tutti i suoi multipli $\alpha \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} \alpha \\ -\alpha \end{pmatrix}$, per $\alpha \in \mathbb{R}$, come nell'esempio 2.4.4. Allora l'insieme

$$\left\{ \begin{pmatrix} \alpha \\ -\alpha \end{pmatrix} : \alpha \in \mathbb{R} \right\} \quad (2.6.1)$$

è un sottospazio vettoriale di \mathbb{R}^2 .

Infatti, sommando due elementi dell'insieme (2.6.1) troviamo:

$$\begin{pmatrix} \alpha \\ -\alpha \end{pmatrix} + \begin{pmatrix} \beta \\ -\beta \end{pmatrix} = \begin{pmatrix} \alpha + \beta \\ -\alpha - \beta \end{pmatrix}$$

che è ancora un elemento dell'insieme. In modo simile si verifica anche la proprietà (2) della definizione 2.6.1. \square

DEFINIZIONE 2.6.3. Si dice che l'insieme (2.6.1) è **generato** dal vettore $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Sia X uno spazio vettoriale e consideriamo dei vettori x_1, \dots, x_n di X , per un certo $n \geq 1$.

PROPOSIZIONE 2.6.4. L'insieme di vettori che sono combinazioni lineari di x_1, \dots, x_n , cioè

$$\{x = \lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n : \lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}\}, \quad (2.6.2)$$

è un sottospazio lineare di X .

DEFINIZIONE 2.6.5. Si dice che (2.6.2) è il sottospazio vettoriale di X **generato da** x_1, \dots, x_n .

Osserviamo che 0_X appartiene al sottospazio generato da x_1, \dots, x_n perché

$$0_X = 0 \cdot x_1 + 0 \cdot x_2 + \dots + 0 \cdot x_n.$$

DEFINIZIONE 2.6.6. Diciamo che $\{y_1, \dots, y_k\}$ è un **sistema di generatori** di uno spazio vettoriale X se ogni vettore in X è combinazione lineare di y_1, \dots, y_k , cioè se per ogni $x \in X$

esistono scalari $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ tali che

$$x = \sum_{i=1}^k \alpha_i y_i.$$

ESEMPIO 2.6.7. Consideriamo lo spazio reale a tre dimensioni \mathbb{R}^3 . I tre vettori

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

formano un sistema di generatori, infatti ogni vettore $(\alpha_1, \alpha_2, \alpha_3)$ di \mathbb{R}^3 si può scrivere come combinazione lineare di questi tre vettori:

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \alpha_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \alpha_3 \end{pmatrix}.$$

□

2.7. Base di uno spazio vettoriale

Il nostro principale obiettivo ora è provare che il numero di vettori linearmente indipendenti non può essere maggiore del numero di elementi di un sistema di generatori. Questo fatto ci servirà per dimostrare che ogni *base* di uno spazio vettoriale, cioè ogni sistema di generatori linearmente indipendenti, ha lo stesso numero di elementi. Proviamo prima un enunciato ausiliario.

PROPOSIZIONE 2.7.1. *Sia X uno spazio vettoriale. Supponiamo di avere due sistemi linearmente indipendenti, $S = \{x_1, \dots, x_n\}$ e $T = \{y_1, \dots, y_k\}$. Se gli n vettori x_1, \dots, x_n formano un sistema di generatori per X , allora $k \leq n$.*

DIMOSTRAZIONE. Consideriamo la famiglia T_1 di vettori $\{y_1, x_1, x_2, \dots, x_n\}$. Poiché i vettori x_1, \dots, x_n formano un sistema di generatori, y_1 è combinazione lineare di essi e quindi la famiglia T_1 è linearmente dipendente. Per il Corollario 2.5.9, esiste almeno un x_j che dipende linearmente dai precedenti vettori in T_1 : scartiamo questo vettore x_j e ci riduciamo alla famiglia $S_1 = \{y_1, x_1, x_2, \dots, x_{j-1}, x_{j+1}, \dots, x_n\}$, che continua a generare tutto X perché il vettore che abbiamo scartato è combinazione lineare di quelli in S_1 . Aggiungiamo ora il secondo vettore di T , cioè formiamo la famiglia $T_2 = \{y_2, y_1, x_1, x_2, \dots, x_{j-1}, x_{j+1}, \dots, x_n\}$. Come prima, questa famiglia è linearmente dipendente, ed uno dei suoi vettori dipende linearmente dai precedenti. Questo vettore non può essere y_1 , perché y_1, \dots, y_k sono indipendenti, quindi deve essere qualche x_i con $i \neq j$. Scartiamo anche x_i e continuiamo in questo modo. Ad ogni passo otteniamo una famiglia linearmente dipendente che genera tutto X , da cui scartare un vettore (uno degli x , non degli y) senza che la sottofamiglia risultante smetta di essere un sistema di generatori. Se i vettori in T fossero meno di quelli in S , alla fine ci ritroveremmo

con una famiglia linearmente dipendente di vettori tutti contenuti in $T = \{y_1, \dots, y_k\}$. Ma allora T non è linearmente indipendente, ed abbiamo una contraddizione. \square

ESERCIZIO 2.7.2. Con un argomento simmetrico rispetto a quello della Proposizione 2.7.1, provare l'enunciato seguente:

Sia X uno spazio vettoriale. Supponiamo di avere due sistemi di generatori,

$$S = \{y_1, \dots, y_k\} \text{ e } T = \{x_1, \dots, x_n\}.$$

Se i k vettori y_1, \dots, y_n formano un sistema linearmente indipendente, allora $k \leq n$.

Svolgimento. La dimostrazione è simmetrica rispetto a quella precedente. Consideriamo la famiglia T_1 di vettori $\{y_1, x_1, x_2, \dots, x_n\}$. Esattamente come prima, y_1 è combinazione lineare di x_1, \dots, x_n e quindi la famiglia T_1 è linearmente dipendente. Per il Corollario 2.5.9, esiste almeno un x_j che dipende linearmente dai precedenti vettori in T_1 : scartiamo questo vettore x_j e ci riduciamo alla famiglia $S_1 = \{y_1, x_1, x_2, \dots, x_{j-1}, x_{j+1}, \dots, x_n\}$, che come prima continua a generare tutto X . Aggiungiamo allora il secondo vettore di S per formare la famiglia $T_2 = \{y_2, y_1, x_1, x_2, \dots, x_{j-1}, x_{j+1}, \dots, x_n\}$. Di nuovo, questa famiglia è un sistema di generatori linearmente dipendente, e possiamo scartare uno dei suoi vettori senza che questo scarto le faccia perdere la proprietà di generare tutto X . Questo vettore che possiamo scartare non è però y_1 perché S è un sistema linearmente indipendente. Continuando così, se il numero dei vettori in S fosse maggiore di quelli di T , ci ridurremmo ad una sottofamiglia linearmente dipendente di vettori tutti in S : questo contraddirebbe l'ipotesi che S sia linearmente indipendente. \square

Dalla Proposizione precedenti segue facilmente la conclusione desiderata:

TEOREMA 2.7.3. *Sia X uno spazio vettoriale. Supponiamo di avere un sistema T di n vettori non nulli linearmente indipendenti x_1, \dots, x_n , ed un sistema S di k generatori non nulli $\{y_1, \dots, y_k\}$. Allora $n \leq k$.*

DIMOSTRAZIONE. Se il sistema S di generatori è anche linearmente indipendente, l'enunciato coincide con la Proposizione 2.7.1. Assumiamo quindi che non lo sia, e scartiamo il primo vettore che è combinazione lineare dei precedenti, riducendolo quindi ad una sottofamiglia S_1 consistente di $k - 1$ vettori la quale, come prima, è ancora un sistema di generatori. Se adesso S_1 risulta essere linearmente indipendente, allora, sempre per la Proposizione 2.7.1, si deve avere $n \leq k - 1 < k$ ed il teorema è dimostrato. Se invece S_1 è linearmente dipendente, continuiamo a scartare vettori finché non lo diventa (prima o poi deve diventarlo, perché una famiglia consistente di un solo vettore non nullo è linearmente indipendente).

In questo modo abbiamo provato che si può estrarre da S una sottofamiglia S' di generatori linearmente indipendenti, ed il numero di vettori di S' è non inferiore a quello di T . \square

DEFINIZIONE 2.7.4. Una **base** di uno spazio vettoriale è un sistema di generatori linearmente indipendenti.

Sia X uno spazio vettoriale e $\{x_1, \dots, x_n\}$ una sua base. Siccome x_1, \dots, x_n è un sistema di generatori di X , ogni elemento x di X si può scrivere come combinazione lineare di x_1, \dots, x_n . Il fatto che x_1, \dots, x_n siano anche linearmente indipendenti, implicano che la scrittura di x come combinazione lineare di x_1, \dots, x_n è *unica*, come mostra la seguente:

PROPOSIZIONE 2.7.5. *Sia $\{x_1, \dots, x_n\}$ una base di uno spazio vettoriale X . Sia x un vettore qualsiasi di X . Allora esistono e sono univocamente determinati $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tali che*

$$x = \lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n. \quad (2.7.1)$$

DIMOSTRAZIONE. Come abbiamo già osservato prima dell'enunciato, x_1, \dots, x_n è un sistema di generatori. Quindi, per ogni $x \in X$, esiste sicuramente la combinazione lineare (2.7.1). Supponiamo che esistano anche $\lambda'_1, \dots, \lambda'_n$ tali che $x = \sum_{i=1}^n \lambda'_i x_i$. Allora avremmo che:

$$\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n = \lambda'_1 x_1 + \lambda'_2 x_2 + \dots + \lambda'_n x_n$$

e quindi, portando tutto al primo membro, che:

$$(\lambda_1 - \lambda'_1)x_1 + (\lambda_2 - \lambda'_2)x_2 + \dots + (\lambda_n - \lambda'_n)x_n = 0_X.$$

Ma x_1, \dots, x_n sono linearmente indipendenti e l'equazione precedente è una combinazione lineare che dà 0_X . Perciò tale combinazione lineare deve avere tutti i coefficienti nulli, cioè deve essere:

$$\lambda_1 = \lambda'_1, \quad \lambda_2 = \lambda'_2, \quad \dots \quad \lambda_n = \lambda'_n,$$

come volevasi dimostrare. □

Il prossimo teorema ci mostra una proprietà fondamentale che hanno tutte le basi di uno stesso spazio vettoriale.

TEOREMA 2.7.6. *Due basi di uno spazio vettoriale hanno lo stesso numero di elementi.*

DIMOSTRAZIONE. Siano $\{x_1, \dots, x_n\}$ e $\{y_1, \dots, y_m\}$ due basi dello stesso spazio vettoriale X . In particolare x_1, \dots, x_n sono linearmente indipendenti e $\{y_1, \dots, y_m\}$ è un sistema di generatori, quindi il teorema 2.7.3 implica che $n \leq m$. D'altra parte, è vero che anche y_1, \dots, y_m sono linearmente indipendenti e $\{x_1, \dots, x_n\}$ è un sistema di generatori, perciò ancora lo stesso teorema 2.7.3 implica pure che $m \leq n$. Ne segue allora che $n = m$, cioè due basi qualsiasi hanno lo stesso numero di elementi, come volevasi dimostrare. □

Possiamo quindi dare la seguente:

DEFINIZIONE 2.7.7. La **dimensione** di uno spazio vettoriale è il numero di elementi di una base, che per il teorema precedente non dipende dalla base scelta. Se n è la dimensione dello spazio vettoriale X , allora scriviamo:

$$\dim X = n.$$

ESEMPIO 2.7.8. Consideriamo i seguenti n vettori di \mathbb{R}^n :

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Come per \mathbb{R}^2 o \mathbb{R}^3 , è facile verificare che questi n vettori sono linearmente indipendenti e che generano tutto \mathbb{R}^n . \square

Concludiamo questa sezione e questo capitolo mostrando come costruire delle basi di spazio vettoriale.

Supponiamo di avere un sistema di generatori $\{x_1, \dots, x_5\}$ di X . Se questi elementi fossero linearmente indipendenti, sarebbero una base. Ma se non sono linearmente indipendenti, allora bisogna trovare una base.

Per esempio, supponiamo che x_5 sia una combinazione lineare degli altri, cioè di x_1, x_2, x_3 e x_4 :

$$x_5 = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4 \quad (2.7.2)$$

con $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{R}$. Allora x_1, x_2, x_3 e x_4 generano X , perché sappiamo che ogni vettore $x \in X$ è combinazione lineare di x_1, \dots, x_5 , quindi esistono $\lambda_1, \dots, \lambda_5 \in \mathbb{R}$ tali che

$$\begin{aligned} x &= \lambda_1 x_1 + \dots + \lambda_4 x_4 + \lambda_5 x_5 = \\ &= \lambda_1 x_1 + \dots + \lambda_4 x_4 + \lambda_5 (\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4) \end{aligned}$$

dove l'ultima uguaglianza segue da (2.7.2), perciò x è combinazione lineare di x_1, \dots, x_4 .

Generalizzando il ragionamento precedente ad un sistema di n generatori $\{x_1, \dots, x_n\}$ di uno spazio vettoriale X si dimostra la seguente:

PROPOSIZIONE 2.7.9. *Sia $\{x_1, \dots, x_n\}$ un sistema di generatori di uno spazio vettoriale X . Allora esiste un sottoinsieme $\{x_{i_1}, \dots, x_{i_k}\}$ del sistema di generatori che è una base di X .*

DIMOSTRAZIONE. Se x_1, \dots, x_n sono linearmente indipendenti, allora abbiamo già una base di X con $k = n$ e $i_1 = 1, i_2 = 2, \dots, i_k = n$. Se invece x_1, \dots, x_n sono linearmente dipendenti, allora esiste una combinazione lineare:

$$\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n = 0_X$$

dove i coefficienti $\lambda_1, \dots, \lambda_n$ non sono tutti nulli. Allora esiste i tale che $\lambda_i \neq 0$ e possiamo scrivere x_i in funzione dei rimanenti:

$$x_i = -\frac{\lambda_1}{\lambda_i} x_1 - \dots - \frac{\lambda_{i-1}}{\lambda_i} x_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} x_{i+1} - \dots - \frac{\lambda_n}{\lambda_i} x_n.$$

Ne segue che $\{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n\}$ è ancora un sistema di generatori di X . Se ora $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ sono linearmente indipendenti, abbiamo trovato una base di X e

abbiamo finito. Se invece sono linearmente dipendenti, allora esiste una combinazione lineare con coefficienti non tutti nulli che è 0_X , da cui possiamo ricavare uno dei vettori x_j in funzione dei rimanenti. Ripetendo lo stesso ragionamento a questi vettori rimasti, dopo un numero finito di volte troveremo un sistema di generatori di X formato da vettori linearmente indipendenti, che quindi è una base di X . Per costruzione, i vettori della base saranno appartenenti al sistema di generatori dato in partenza. \square

Vediamo ora un altro metodo per trovare una base di uno spazio vettoriale X . Sia x_1 un vettore non nullo di X . Se x_1 genera X , allora x_1 è una base. Altrimenti, se x_1 non genera X , esiste un altro elemento x_2 linearmente indipendente da x_1 , cioè che non è multiplo di x_1 . Se x_1 e x_2 generano X , allora sono una base, perché sono linearmente indipendenti. Altrimenti esiste un vettore x_3 che non è combinazione lineare di x_1 e x_2 . Si procede allo stesso modo finché non si trova un sistema di generatori che sono linearmente indipendenti per costruzione, che quindi formano una base.

Si può formalizzare questa idea con la seguente proposizione:

PROPOSIZIONE 2.7.10. *Siano x_1, \dots, x_m vettori linearmente indipendenti di uno spazio vettoriale X (può essere $m = 1$, nel qual caso x_1 è un vettore non nullo qualsiasi). Se la dimensione di X è $n > m$, allora esistono dei vettori x_{m+1}, \dots, x_n tali che $\{x_1, \dots, x_n\}$ è una base di X .*

DIMOSTRAZIONE. Consideriamo il sottospazio vettoriale generato da x_1, \dots, x_m . Scegliamo un vettore qualsiasi x_{m+1} non appartenente a questo sottospazio. Allora x_1, \dots, x_{m+1} sono linearmente indipendenti. Infatti, se non lo fossero, esisterebbero $\lambda_1, \dots, \lambda_{m+1}$ non tutti nulli tali che:

$$\lambda_1 x_1 + \dots + \lambda_m x_m + \lambda_{m+1} x_{m+1} = 0_X;$$

ora ci sono due possibilità: o $\lambda_{m+1} = 0$, ma allora x_1, \dots, x_m sarebbero linearmente dipendenti, in contraddizione con l'ipotesi; oppure $\lambda_{m+1} \neq 0$, ma allora potremmo scrivere x_{m+1} come combinazione lineare di x_1, \dots, x_m , contraddicendo l'ipotesi di aver scelto x_m non appartenente al sottospazio generato da x_1, \dots, x_n .

A questo punto possiamo ripetere il ragionamento ai vettori linearmente indipendenti x_1, \dots, x_{m+1} e scegliere un vettore qualsiasi x_{m+2} non appartenente al sottospazio vettoriale da essi generato. Con la stessa dimostrazione appena fatta, si vede che x_1, \dots, x_{m+2} sono linearmente indipendenti.

Continuando così, si costruisce una base x_1, \dots, x_n come volevasi dimostrare. \square

ESEMPIO 2.7.11. Consideriamo il sistema di generatori $\{x_1, x_2, x_3\}$, dello spazio vettoriale $X = \mathbb{R}^2$, dove:

$$x_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Vediamo come trovare una base di X seguendo la proposizione [2.7.9](#).

I tre vettori x_1 , x_2 e x_3 sono linearmente dipendenti, infatti

$$x_1 + 3x_2 - 2x_3 = \begin{pmatrix} 2 \\ -1 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 + 0 \cdot 3 - 2 \cdot 1 \\ -1 + 3 \cdot 1 - 2 \cdot 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Allora si ha che:

$$x_1 = 2x_3 - 3x_2$$

e quindi $\{x_2, x_3\}$ è ancora un sistema di generatori di X . Siccome $X = \mathbb{R}^2$ ha dimensione 2, allora $\{x_2, x_3\}$ è una base di X . \square

ESEMPIO 2.7.12. Consideriamo il vettore

$$x_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

dello spazio vettoriale $X = \mathbb{R}^2$. Vediamo come costruire una base di X seguendo la proposizione 2.7.10.

Il vettore x_1 genera il sottospazio vettoriale di X formato dai vettori

$$\begin{pmatrix} \alpha \\ \alpha \end{pmatrix}$$

dove α è un numero reale qualsiasi. Scegliamo un vettore x_2 non appartenente a tale sottospazio, per esempio

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

Allora x_1 e x_2 sono sicuramente linearmente indipendenti e quindi $\{x_1, x_2\}$ è una base di X perché $\dim X = 2$. \square

Si noti che nell'esempio precedente avremmo potuto scegliere un altro vettore x_2 , come

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

o anche

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

o in altri infiniti modi, perché ci sono infiniti vettori non appartenenti al sottospazio vettoriale generato da x_1 .

CAPITOLO 3

Applicazioni lineari e matrici

3.1. Definizione di applicazione lineare

DEFINIZIONE 3.1.1. Siano \mathbf{X} e \mathbf{Y} spazi vettoriali. Una applicazione $T : \mathbf{X} \longrightarrow \mathbf{Y}$ si chiama *lineare* se rispetta le operazioni di spazio vettoriale di \mathbf{X} e \mathbf{Y} , cioè

$$\begin{aligned} T(\mathbf{x}_1 + \mathbf{x}_2) &= T(\mathbf{x}_1) + T(\mathbf{x}_2), & \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbf{X}, \\ T(\lambda \mathbf{x}) &= \lambda T(\mathbf{x}), & \forall \mathbf{x} \in \mathbf{X}, \forall \lambda \in \mathbb{R}. \end{aligned}$$

NOTA 3.1.2. L'immagine tramite T di una combinazione lineare di vettori in \mathbf{X} è la combinazione lineare degli immagini di questi vettori con i coefficienti corrispondenti :

$$T(\lambda_1 \mathbf{x}_1 + \dots + \lambda_k \mathbf{x}_k) = \lambda_1 T(\mathbf{x}_1) + \dots + \lambda_k T(\mathbf{x}_k).$$

□

ESEMPIO 3.1.3. Siano $X = Y = \mathbb{R}^2$ e consideriamo le applicazioni $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Allora l'applicazione:

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x - 3y \\ -x + y \end{pmatrix}$$

è lineare, infatti soddisfa le due proprietà della definizione 3.1.1, come si può verificare direttamente. □

Ma vediamo anche un esempio di applicazione non lineare.

ESEMPIO 3.1.4. Consideriamo ancora le applicazioni $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. L'applicazione:

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^2 + y \\ y \end{pmatrix}$$

non è lineare, perché compare un esponente 2 in una delle incognite. Da ciò segue per esempio che:

$$T \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 9 \\ 0 \end{pmatrix} \neq \begin{pmatrix} 5 \\ 0 \end{pmatrix} = T \begin{pmatrix} 1 \\ 0 \end{pmatrix} + T \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

□

Si verifica facilmente che, se $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ sono spazi vettoriali e $T : \mathbf{X} \longrightarrow \mathbf{Y}$, $S : \mathbf{Y} \longrightarrow \mathbf{Z}$ sono applicazioni lineari, allora la *composizione*

$$S \circ T : \mathbf{X} \ni \mathbf{x} \longmapsto S(T(\mathbf{x})) \in \mathbf{Z}$$

è pure una applicazione lineare.

3.2. Immagine e nucleo di una applicazione lineare

Sia $T : \mathbf{X} \longrightarrow \mathbf{Y}$ una applicazione lineare. Allora

$$\text{Ker}(T) := \{\mathbf{x} \in \mathbf{X}; T(\mathbf{x}) = \mathbf{0}_{\mathbf{Y}}\} \subset \mathbf{X},$$

chiamato il *nucleo* di T , è un sottospazio lineare di \mathbf{X} . Infatti, $\text{Ker}(T)$ contiene $\mathbf{0}_{\mathbf{X}}$ perché $T(\mathbf{0}_{\mathbf{X}}) = \mathbf{0}_{\mathbf{Y}}$,

$$\begin{aligned} \mathbf{x}_1, \mathbf{x}_2 \in \text{Ker}(T) &\implies T(\mathbf{x}_1 + \mathbf{x}_2) = T(\mathbf{x}_1) + T(\mathbf{x}_2) = \mathbf{0}_{\mathbf{Y}} \\ &\implies \mathbf{x}_1 + \mathbf{x}_2 \in \text{Ker}(T) \end{aligned}$$

e

$$\begin{aligned} \mathbf{x} \in \text{Ker}(T), \lambda \in \mathbb{R} &\implies T(\lambda \mathbf{x}) = \lambda T(\mathbf{x}) = \mathbf{0}_{\mathbf{Y}} \\ &\implies \lambda \mathbf{x} \in \text{Ker}(T). \end{aligned}$$

Poi,

$$T(\mathbf{X}) := \{T(\mathbf{x}); \mathbf{x} \in \mathbf{X}\} \subset \mathbf{Y},$$

chiamato l'*immagine* di T , è un sottospazio lineare di \mathbf{Y} . Infatti, $T(\mathbf{X})$ contiene $\mathbf{0}_{\mathbf{Y}} = T(\mathbf{0}_{\mathbf{X}})$,

$$T(\mathbf{x}_1) + T(\mathbf{x}_2) = T(\mathbf{x}_1 + \mathbf{x}_2) \in T(\mathbf{X})$$

e

$$\lambda T(\mathbf{x}) = T(\lambda \mathbf{x}) \in T(\mathbf{X}).$$

Esiste una relazione notevole fra la dimensione del nucleo e la dimensione dell'immagine di una applicazione lineare:

TEOREMA 3.2.1. (Teorema della dimensione.) *Se \mathbf{X} e \mathbf{Y} sono spazi vettoriali di dimensione finita e $T : \mathbf{X} \longrightarrow \mathbf{Y}$ è una applicazione lineare, allora*

$$\dim(\text{Ker}(T)) + \dim(T(\mathbf{X})) = \dim(\mathbf{X}).$$

Dimostrazione. (facoltativa.) Sia $\mathbf{v}_1, \dots, \mathbf{v}_p$ una base di $\text{Ker}(T)$ e $\mathbf{f}_1, \dots, \mathbf{f}_q$ una base di $T(\mathbf{X})$. Allora esistono $\mathbf{x}_1, \dots, \mathbf{x}_q \in \mathbf{X}$ tale che

$$\mathbf{f}_1 = T(\mathbf{x}_1), \dots, \mathbf{f}_q = T(\mathbf{x}_q).$$

Verifichiamo che i vettori $\mathbf{v}_1, \dots, \mathbf{v}_p, \mathbf{x}_1, \dots, \mathbf{x}_q \in \mathbf{X}$ sono linearmente indipendenti. Infatti, se

$$\alpha_1 \mathbf{v}_1 + \dots + \alpha_p \mathbf{v}_p + \beta_1 \mathbf{x}_1 + \dots + \beta_q \mathbf{x}_q = \mathbf{0}_{\mathbf{X}}, \quad (3.2.1)$$

allora

$$\begin{aligned}
 \beta_1 \mathbf{f}_1 + \dots + \beta_q \mathbf{f}_q &= \beta_1 T(\mathbf{x}_1) + \dots + \beta_q T(\mathbf{x}_q) \\
 &= T(\beta_1 \mathbf{x}_1 + \dots + \beta_q \mathbf{x}_q) \\
 &= T(\underbrace{\alpha_1 \mathbf{v}_1 + \dots + \alpha_p \mathbf{v}_p}_{\in \text{Ker}(T)} + \beta_1 \mathbf{x}_1 + \dots + \beta_q \mathbf{x}_q) \\
 &= T(\mathbf{0}_\mathbf{X}) = \mathbf{0}_\mathbf{Y}
 \end{aligned}$$

e dall'indipendenza lineare dei vettori $\mathbf{f}_1, \dots, \mathbf{f}_q$ risulta che tutti i coefficienti β_1, \dots, β_q si annullano. Ora da (3.2.1) risulta che $\alpha_1 \mathbf{v}_1 + \dots + \alpha_p \mathbf{v}_p = \mathbf{0}_\mathbf{X}$ e l'indipendenza lineare del sistema $\mathbf{v}_1, \dots, \mathbf{v}_p$ implica che anche i coefficienti $\alpha_1, \dots, \alpha_p$ si annullano.

Verifichiamo anche che $\mathbf{v}_1, \dots, \mathbf{v}_p, \mathbf{x}_1, \dots, \mathbf{x}_q \in \mathbf{X}$ generano lo spazio vettoriale \mathbf{X} . Sia $\mathbf{x} \in \mathbf{X}$ arbitrario. Poiché $\mathbf{f}_1, \dots, \mathbf{f}_q$ generano $T(\mathbf{X})$, esistono dei scalari μ_1, \dots, μ_q tale che

$$T(\mathbf{x}) = \mu_1 \mathbf{f}_1 + \dots + \mu_q \mathbf{f}_q = T(\mu_1 \mathbf{x}_1 + \dots + \mu_q \mathbf{x}_q),$$

cioè

$$\mathbf{x} - (\mu_1 \mathbf{x}_1 + \dots + \mu_q \mathbf{x}_q) \in \text{Ker}(T).$$

Ma $\mathbf{v}_1, \dots, \mathbf{v}_p$ generano $\text{Ker}(T)$, perciò esistono sei scalari $\lambda_1, \dots, \lambda_p$ tale che

$$\mathbf{x} - (\mu_1 \mathbf{x}_1 + \dots + \mu_q \mathbf{x}_q) = \lambda_1 \mathbf{v}_1 + \dots + \lambda_p \mathbf{v}_p$$

e così

$$\mathbf{x} = \lambda_1 \mathbf{v}_1 + \dots + \lambda_p \mathbf{v}_p + \mu_1 \mathbf{x}_1 + \dots + \mu_q \mathbf{x}_q.$$

Concludiamo che $\mathbf{v}_1, \dots, \mathbf{v}_p, \mathbf{x}_1, \dots, \mathbf{x}_q$ è una base di \mathbf{X} e così $\dim(\mathbf{X})$ è uguale a $p + q = \dim(\text{Ker}(T)) + \dim(T(\mathbf{X}))$.

□

NOTA 3.2.2. Una applicazione lineare $T : \mathbf{X} \longrightarrow \mathbf{Y}$ è *iniettiva*, cioè applica vettori diversi in vettori diversi, ovvero

$$\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{X}, T(\mathbf{x}_1) = T(\mathbf{x}_2) \implies \mathbf{x}_1 = \mathbf{x}_2, \quad (3.2.2)$$

se e solo se $\text{Ker}(T) = \{\mathbf{0}_\mathbf{X}\}$, ovvero

$$\mathbf{x} \in \mathbf{X}, T(\mathbf{x}) = \mathbf{0}_\mathbf{Y} \implies \mathbf{x} = \mathbf{0}_\mathbf{X}. \quad (3.2.3)$$

Infatti, (3.2.2) implica ovviamente (3.2.3), perciò se T è iniettiva, allora $\text{Ker}(T) = \{\mathbf{0}_\mathbf{X}\}$. Viceversa, se $\text{Ker}(T) = \{\mathbf{0}_\mathbf{X}\}$ e $\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{X}$ sono tali che $T(\mathbf{x}_1) = T(\mathbf{x}_2)$, allora $T(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}_\mathbf{Y}$ e, grazie a (3.2.3), risulta che $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}_\mathbf{X}$, cioè $\mathbf{x}_1 = \mathbf{x}_2$. □

Se \mathbf{X} e \mathbf{Y} sono di dimensione finita, allora, per il teorema della dimensione 3.2.1,

$$T \text{ iniettiva} \iff \dim(\text{Ker}(T)) = 0 \iff \dim(T(\mathbf{X})) = \dim(\mathbf{X}).$$

Poiché $T(\mathbf{X})$ è un sottospazio lineare di \mathbf{Y} e così $\dim(T(\mathbf{X})) \leq \dim(\mathbf{Y})$, per l'injectività di T è necessario che $\dim(\mathbf{X}) \leq \dim(\mathbf{Y})$.

Ricordiamo che una applicazione lineare $T : \mathbf{X} \longrightarrow \mathbf{Y}$ si chiama *surgettiva* se $T(\mathbf{X}) = \mathbf{Y}$. Se \mathbf{X} e \mathbf{Y} sono di dimensione finita, allora, tenendo conto del teorema della dimensione 3.2.1,

$$\begin{aligned} T \text{ surgettiva} &\iff \dim(T(\mathbf{X})) = \dim(\mathbf{Y}) \\ &\iff \dim(\text{Ker}(T)) = \dim(\mathbf{X}) - \dim(\mathbf{Y}). \end{aligned}$$

In particolare, per la surgettività di T è necessario che $\dim(\mathbf{X}) - \dim(\mathbf{Y})$ sia ≥ 0 , cioè che $\dim(\mathbf{X}) \geq \dim(\mathbf{Y})$.

Se l'applicazione lineare $T : \mathbf{X} \longrightarrow \mathbf{Y}$ è *bigettiva*, cioè simultaneamente iniettiva e surgettiva, allora dobbiamo avere $\dim(\mathbf{X}) = \dim(\mathbf{Y})$. Se \mathbf{X} e \mathbf{Y} sono di dimensione finita uguale, allora dal teorema della dimensione 3.2.1 risulta subito che

$$\dim(\text{Ker}(T)) = 0 \iff \dim(T(\mathbf{X})) = \dim(\mathbf{Y}).$$

Quindi, in questo caso,

$$\begin{aligned} T \text{ bigettiva} &\iff \dim(\mathbf{X}) = \dim(\mathbf{Y}) \ \& \ T \text{ iniettiva} \\ &\iff \dim(\mathbf{X}) = \dim(\mathbf{Y}) \ \& \ T \text{ surgettiva.} \end{aligned} \tag{3.2.4}$$

Se T è bigettiva, allora, associando ad ogni $\mathbf{y} \in \mathbf{Y}$ l'unico $\mathbf{x} \in \mathbf{X}$ che soddisfa $\mathbf{y} = T(\mathbf{x})$, otteniamo l'applicazione inversa $T^{-1} : \mathbf{Y} \longrightarrow \mathbf{X}$ di T , per cui abbiamo

$$T^{-1} \circ T = I_{\mathbf{X}}, \quad T \circ T^{-1} = I_{\mathbf{Y}},$$

ove $I_{\mathbf{X}} : \mathbf{X} \ni \mathbf{x} \longmapsto \mathbf{x} \in \mathbf{X}$ e $I_{\mathbf{Y}} : \mathbf{Y} \ni \mathbf{y} \longmapsto \mathbf{y} \in \mathbf{Y}$ indicano le rispettive applicazioni identiche.

Si verifica facilmente che, se $T : \mathbf{X} \longrightarrow \mathbf{Y}$, $S : \mathbf{Y} \longrightarrow \mathbf{Z}$ sono applicazioni lineari, allora

$$T, S \text{ iniettive} \implies S \circ T \text{ iniettiva} \implies T \text{ iniettiva}, \tag{3.2.5}$$

$$T, S \text{ surgettive} \implies S \circ T \text{ surgettiva} \implies S \text{ surgettiva}. \tag{3.2.6}$$

Tenendo conto di (3.2.4), risulta che se \mathbf{X} , \mathbf{Y} hanno la stessa dimensione finita, allora

$$S \circ T \text{ iniettiva} \implies T \text{ bigettiva},$$

mentre se \mathbf{Y} , \mathbf{Z} hanno la stessa dimensione finita, allora

$$S \circ T \text{ surgettiva} \implies S \text{ bigettiva}.$$

3.3. Spazi vettoriali di polinomi ed applicazioni lineari

DEFINIZIONE 3.3.1. (**Polinomi.**) Un *polinomio* (con coefficienti reali) è una espressione $P(x)$ della forma

$$a_0 + a_1 x + \dots + a_n x^n \text{ ossia } \sum_{k=0}^n a_k x^k,$$

ove a_0, a_1, \dots, a_n sono *coefficienti* reali, $n \geq 0$ è un intero ed x è una *variabile*. Due polinomi

$$\sum_{k=0}^n a_k x^k, \quad \sum_{k=0}^m b_k x^k$$

con $m \geq n$ sono considerati uguali se $a_k = b_k$ per $k \leq n$ e $b_k = 0$ per $k > n$. L'insieme \mathcal{P} di tutti i polinomi è uno spazio vettoriale rispetto all'addizione

$$\sum_{k=0}^n a_k x^k + \sum_{k=0}^m b_k x^k := \sum_{k=0}^n (a_k + b_k) x^k$$

ed alla moltiplicazione con scalari

$$\lambda \sum_{k=0}^n a_k x^k := \sum_{k=0}^n (\lambda a_k) x^k.$$

Il *grado* di un polinomio non zero $P(x) = \sum_{k=0}^n a_k x^k$ è il più grande intero $0 \leq d \leq n$ per quale $a_d \neq 0$, $a_k = 0$ per $k > d$

e si indica con $\deg(P)$. Il grado del polinomio zero è per definizione $-\infty$. Definiamo anche il prodotto di due polinomi

$$P(x) = \sum_{k=0}^n a_k x^k, \quad Q(x) = \sum_{k=0}^m b_k x^k$$

come segue:

$$(PQ)(x) := \sum_{j=0}^{n+m} \left(\sum_{\substack{0 \leq k_1 \leq n \\ 0 \leq k_2 \leq m \\ k_1 + k_2 = j}} a_{k_1} b_{k_2} \right) x^j,$$

ove $a_{k_1} = 0$ per $k_1 > n$ e $b_{k_2} = 0$ per $k_2 > m$. Il grado di PQ è uguale alla somma $\deg(P) + \deg(Q)$.

La moltiplicazione dei polinomi è associativa e commutativa. Inoltre vale la proprietà di distributività:

$$(P + Q)R = PR + QR.$$

Se $P(x) = \sum_{k=0}^n a_k x^k$ è un polinomio ed r è un numero reale, allora $P(r)$ è il numero che

si ottiene sostituendo $x = r$ in $P(x)$, cioè $\sum_{k=0}^n a_k r^k$.

È noto che per qualsiasi polinomi P e D , $D \neq 0$, possiamo dividere P con D con resto: esistono unici polinomi Q e R tale che

$$P = DQ + R, \quad \deg(R) < \deg(D).$$

Q si chiama il *quoziente* della divisione e verifica $\deg(D) + \deg(Q) = \deg(P)$. R si chiama il *resto* della divisione.

Se $D(x) = d_0 + d_1 x$, allora R è la costante $P\left(-\frac{d_0}{d_1}\right)$.

Indichiamo per ogni intero $n \geq 0$ con \mathcal{P}_n l'insieme di tutti i polinomi di grado $\leq n$. Allora \mathcal{P}_n è un sottospazio lineare di \mathcal{P} , di dimensione $n + 1$, avendo come base

$$1, x, \dots, x^n.$$

La derivata $P'(x)$ di un polinomio $P(x) = \sum_{k=0}^n a_k x^k$ è

$$a_1 x + 2a_2 x + \dots + n a_n x^{n-1} = \sum_{k=1}^n k a_k x^{k-1}.$$

Per il grado della derivata di un polinomio P non costante abbiamo

$$\deg(P') = \deg(P) - 1.$$

La derivata di un polinomio costante è 0.

Regole di calcolo per la derivazione:

$$(P + Q)' = P' + Q', \quad (\lambda P)' = \lambda P', \quad (PQ)' = P'Q + PQ'.$$

3.4. Esercizi sulle applicazioni lineari su spazi di polinomi

ESERCIZIO 3.4.1. Si trovi il nucleo e l'immagine dell'applicazione lineare $T : \mathcal{P}_3 \longrightarrow \mathcal{P}_4$ che assegna ad ogni $P \in \mathcal{P}_3$ il polinomio $(x + 2)P(x)$.

Svolgimento. $\text{Ker}(T) = \{0\}$, cioè T è iniettiva. Infatti, ogni $0 \neq P \in \mathcal{P}_3$ ha un grado ≥ 0 ed allora il grado di $T(P)$, cioè di $(x + 2)P(x)$, è uguale ad $1 + \deg(P) \geq 1$.

$T(\mathcal{P}_3)$ consiste da tutti i polinomi di grado ≤ 4 che sono divisibili con $x - 2$, cioè quelli si annullano in -2 . In particolare, T non è surgettiva. \square

ESERCIZIO 3.4.2. Si trovi il nucleo e l'immagine dell'applicazione lineare $T : \mathcal{P}_5 \longrightarrow \mathcal{P}_1$ che assegna ad ogni $P \in \mathcal{P}_5$ il resto nella divisione di P con $x^2 + 1$.

Svolgimento. $\text{Ker}(T)$ consiste da tutti i polinomi di grado ≤ 5 che sono divisibili con $x^2 + 1$. In particolare, T non è iniettiva.

L'immagine di T è uguale a \mathcal{P}_1 , cioè T è surgettiva. Infatti, il resto della divisione con $x^2 + 1$ di un qualsiasi $P \in \mathcal{P}_1$ è uguale a P stesso. \square

ESERCIZIO 3.4.3. Si trovi il nucleo e l'immagine dell'applicazione lineare $T : \mathcal{P}_4 \longrightarrow \mathcal{P}_3$ che assegna ad ogni $P \in \mathcal{P}_4$ la sua derivata.

Svolgimento. $\text{Ker}(T)$ consiste di tutti i polinomi costanti. In particolare, T non è iniettiva.

D'altro canto, T è surgettiva. Infatti, il polinomio generico

$$P(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3,$$

è la derivata di

$$a_0 x + \frac{a_1}{2} x^2 + \frac{a_2}{3} x^3 + \frac{a_3}{4} x^4.$$

□

ESERCIZIO 3.4.4. Si trovino il nucleo e l'immagine dell'applicazione lineare $T : \mathcal{P}_3 \longrightarrow \mathcal{P}_3$ che assegna ad ogni $P \in \mathcal{P}_3$ la derivata di $(x-3)P(x)$.

Svolgimento. T è iniettiva. Infatti, se $P \in \mathcal{P}_3$ è tale che $T(P) = 0$, allora il polinomio $(x-3)P(x)$ è costante, che è possibile solo se $P = 0$.

Poiché T è una applicazione lineare fra spazi vettoriali della stessa dimensione finita, dall'iniettività di T risulta la sua surgettività. □

3.5. Applicazioni lineari fra \mathbb{R}^n e \mathbb{R}^m e calcolo con matrici

Abbiamo visto che l'immagine tramite T di una combinazione lineare di vettori in \mathbb{R}^n è la combinazione lineare degli immagini di questi vettori con i coefficienti corrispondenti :

$$T(\lambda_1 \mathbf{x}_1 + \dots + \lambda_k \mathbf{x}_k) = \lambda_1 T(\mathbf{x}_1) + \dots + \lambda_k T(\mathbf{x}_k).$$

Quindi una applicazione lineare $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ è completamente determinata dai valori assunti negli elementi della base naturale

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{v}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

di \mathbb{R}^n . Infatti, per ogni vettore in \mathbb{R}^n

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \sum_{j=1}^n x_j \mathbf{v}_j$$

abbiamo

$$T(\mathbf{x}) = T\left(\sum_{j=1}^n x_j \mathbf{v}_j\right) = \sum_{j=1}^n x_j T(\mathbf{v}_j). \quad (3.5.1)$$

Pertanto, se i vettori $T(\mathbf{v}_1), \dots, T(\mathbf{v}_n) \in \mathbb{R}^m$, scritti esplicitamente, sono

$$T(\mathbf{v}_1) = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \\ \vdots \\ \alpha_{m1} \end{pmatrix}, T(\mathbf{v}_2) = \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \alpha_{32} \\ \vdots \\ \alpha_{m2} \end{pmatrix}, \dots, T(\mathbf{v}_n) = \begin{pmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \alpha_{3n} \\ \vdots \\ \alpha_{mn} \end{pmatrix},$$

possiamo riscrivere (3.5.1) come segue :

$$\begin{aligned} T(\mathbf{x}) &= x_1 \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \\ \vdots \\ \alpha_{m1} \end{pmatrix} + x_2 \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \alpha_{32} \\ \vdots \\ \alpha_{m2} \end{pmatrix} + \dots + x_n \begin{pmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \alpha_{3n} \\ \vdots \\ \alpha_{mn} \end{pmatrix} \\ &\stackrel{!}{=} \underbrace{\begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \alpha_{31} & \alpha_{32} & \dots & \alpha_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix}}_{=: A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}}_{= \mathbf{x}}, \end{aligned}$$

ove il *prodotto* della matrice A di m righe ed n colonne, detta matrice $m \times n$, con il vettore \mathbf{x} di n componenti è il vettore di m componenti

$$\begin{pmatrix} \alpha_{11} x_1 + \alpha_{12} x_2 + \alpha_{13} x_3 + \dots + \alpha_{1n} x_n \\ \alpha_{21} x_1 + \alpha_{22} x_2 + \alpha_{23} x_3 + \dots + \alpha_{2n} x_n \\ \alpha_{31} x_1 + \alpha_{32} x_2 + \alpha_{33} x_3 + \dots + \alpha_{3n} x_n \\ \vdots \\ \alpha_{m1} x_1 + \alpha_{m2} x_2 + \alpha_{m3} x_3 + \dots + \alpha_{mn} x_n \end{pmatrix}.$$

La matrice $m \times n$ ottenuta si chiama *la matrice associata* a T (rispetto alle basi naturali di \mathbb{R}^n e \mathbb{R}^m).

Viceversa, se A è una matrice $m \times n$, allora l'applicazione

$$T_A : \mathbb{R}^n \ni \mathbf{x} \longmapsto A\mathbf{x} \in \mathbb{R}^m,$$

ove il prodotto $A\mathbf{x}$ è definito come di cui sopra, è una applicazione lineare e la matrice associata a T_A è A . Perciò

$$A \longmapsto T_A$$

è una corrispondenza biunivoca tra le matrici $m \times n$ e le applicazioni lineari $\mathbb{R}^n \rightarrow \mathbb{R}^m$. ricordiamo che la colonna j -esima di A è uguale all'immagine dell'elemento j -esimo della base naturale di \mathbb{R}^n tramite T_A .

3.6. Spazi vettoriali di matrici

Cominciamo con due notazioni. La matrice dell'applicazione identicamente zero $\mathbb{R}^n \ni \mathbf{x} \mapsto \mathbf{0}_m \in \mathbb{R}^m$ è

$$\mathbf{0}_{mn} := \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \in M_{mn} \text{ (}\mathbf{0}_n \text{ se } m = n\text{)},$$

mentre la matrice dell'applicazione identica $\mathbb{R}^n \ni \mathbf{x} \mapsto \mathbf{x} \in \mathbb{R}^n$ è la *matrice unità* d'ordine n :

$$I_n := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \in M_n.$$

NOTAZIONE 3.6.1. (Applicazioni lineari da \mathbb{R}^n a \mathbb{R}^m .) Indichiamo con $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ l'insieme di tutte le applicazioni lineari $\mathbb{R}^n \rightarrow \mathbb{R}^m$, e con M_{mn} l'insieme di tutte le matrici $m \times n$. Poniamo anche $\mathcal{L}(\mathbb{R}^n)$ per $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ ed M_n per M_{nn} .

NOTA 3.6.2. L'insieme $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ è uno spazio vettoriale rispetto alle operazioni seguenti:

- la somma $T_1 + T_2$ di $T_1, T_2 \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ è definita da $(T_1 + T_2)(\mathbf{x}) := T_1(\mathbf{x}) + T_2(\mathbf{x})$,
- il prodotto λT di $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ è definita da $(\lambda T)(\mathbf{x}) := \lambda T(\mathbf{x})$.

Infatti, se a $T_1, T_2 \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ corrispondono le matrici

$$A_1 = \left(\alpha_{jk} \right)_{\substack{1 \leq j \leq m \\ 1 \leq k \leq n}}, \quad A_2 = \left(\beta_{jk} \right)_{\substack{1 \leq j \leq m \\ 1 \leq k \leq n}},$$

allora a $T_1 + T_2$ corrisponde

$$A_1 + A_2 := \left(\alpha_{jk} + \beta_{jk} \right)_{\substack{1 \leq j \leq m \\ 1 \leq k \leq n}},$$

mentre se a $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ corrisponde

$$A = \left(\gamma_{jk} \right)_{\substack{1 \leq j \leq m \\ 1 \leq k \leq n}},$$

allora a λT corrisponde

$$\lambda A := \left(\lambda \gamma_{jk} \right)_{\substack{1 \leq j \leq m \\ 1 \leq k \leq n}}.$$

□

Così abbiamo definito la somma di due matrici dello stesso tipo, e la moltiplicazione di una matrice per un scalare reale. Ora definiamo il prodotto di due matrici (un'operazione in più rispetto a quelle necessarie per definire gli spazi vettoriali).

DEFINIZIONE 3.6.3. (*Prodotto di matrici.*) Siano $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $S : \mathbb{R}^m \rightarrow \mathbb{R}^p$ applicazioni lineari tali che il dominio di S sia uguale allo spazio vettoriale immagine di T . Allora possiamo considerare la composizione $S \circ T : \mathbb{R}^n \rightarrow \mathbb{R}^p$, che è anch'essa lineare. Se $A = (\alpha_{jk})_{j,k} \in M_{mn}$ e $B = (\beta_{hl})_{h,l} \in M_{pm}$ sono le matrici associate a T ed S , allora la *matrice prodotto* BA associata ad $S \circ T$ è la matrice di tipo $p \times n$ data dalla formula di *moltiplicazione righe per colonne* seguente:

$$BA := \left((\beta_{h,1} \ \dots \ \beta_{h,m}) \cdot \begin{pmatrix} \alpha_{1,k} \\ \vdots \\ \alpha_{m,k} \end{pmatrix} \right)_{\substack{1 \leq h \leq p \\ 1 \leq k \leq n}} = \left(\sum_{j=1}^m \beta_{h,j} \alpha_{j,k} \right)_{\substack{1 \leq h \leq p \\ 1 \leq k \leq n}}.$$

Perciò il prodotto BA di due matrici B ed A esiste solo se B ha tante colonne quante sono le righe di A . In questo caso BA ha tante righe quante ne ha B , e tante colonne quante ne ha A . Se la h -esima riga di B e la k -esima colonna di A sono

$$\text{--- } \mathbf{r}_h \text{ ---} \quad \text{rispettivamente} \quad \begin{array}{c} | \\ \mathbf{c}_k \\ | \end{array},$$

allora \mathbf{r}_h e \mathbf{c}_k hanno lo stesso numero di componenti e l'elemento di BA all'incrocio della h -esima riga e della k -esima colonna è

$$\begin{aligned} & (\text{primo componente di } \mathbf{r}_h) \cdot (\text{primo componente di } \mathbf{c}_k) \\ & + (\text{secondo componente di } \mathbf{r}_h) \cdot (\text{secondo componente di } \mathbf{c}_k) \\ & \quad \vdots \\ & + (\text{l'ultimo componente di } \mathbf{r}_h) \cdot (\text{l'ultimo componente di } \mathbf{c}_k). \end{aligned}$$

Osserviamo che, per $A \in M_{mn}$ e $\mathbf{x} \in \mathbb{R}^n$, il vettore $A\mathbf{x}$, considerato come matrice $m \times 1$, è il prodotto della matrice $m \times n$ A e la matrice $n \times 1$ \mathbf{x} .

Se $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ è una applicazione lineare, allora

$$\begin{aligned} \text{Ker}(T) &:= \{\mathbf{x} \in \mathbb{R}^n ; T(\mathbf{x}) = \mathbf{0}_m\} \subset \mathbb{R}^n \text{ (il nucleo di } T) \\ T(\mathbb{R}^n) &:= \{T(\mathbf{x}) ; \mathbf{x} \in \mathbb{R}^n\} \subset \mathbb{R}^m \text{ (l'immagine di } T) \end{aligned}$$

sono sottospazi lineari. Ricordiamo cosa dice il teorema della dimensione 3.2.1 in questo caso:

$$\dim(\text{Ker}(T)) + \dim(T(\mathbb{R}^n)) = \dim(\mathbb{R}^n) = n.$$

Ricordiamo (Nota 3.2.2) che T è *iniettiva* se e solo se $\text{Ker}(T) = \{\mathbf{0}_n\}$.

Di conseguenza, usando anche il teorema della dimensione 3.2.1, risulta che T è iniettiva se e solo se $\dim(T(\mathbb{R}^n)) = n$. Poiché $\dim(T(\mathbb{R}^n)) \leq \dim(\mathbb{R}^m) = m$, per l'iniettività di T è necessario che $n \leq m$.

D'altro canto, dal Teorema della dimensione 3.2.1 risulta che T è *surgettiva*, cioè $T(\mathbb{R}^n) = \mathbb{R}^m$, se e solo se $\dim(\text{Ker}(T)) = n - m$. Quindi per la surgettività di T è necessario che $n \geq m$.

Sia ora

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix}$$

una matrice $m \times n$. Allora $\text{Ker}(T_A) = \{\mathbf{x} \in \mathbb{R}^n; A\mathbf{x} = \mathbf{0}_m\}$ è l'insieme di tutte le soluzioni

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

del sistema omogeneo $A\mathbf{x} = \mathbf{0}_m$, ovvero

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1n}x_n = 0 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \dots + \alpha_{2n}x_n = 0 \\ \vdots \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \dots + \alpha_{mn}x_n = 0 \end{cases}.$$

Perciò T_A è iniettiva esattamente quando questo sistema ha la sola soluzione $\mathbf{x} = \mathbf{0}_n$. Poiché il sistema precedente si scrive !)

$$A\mathbf{x} = \sum_{j=1}^n x_j \mathbf{c}_j \quad \text{ove } \mathbf{c}_1, \dots, \mathbf{c}_n \text{ sono le colonne di } A, \quad (3.6.1)$$

l'iniettività di T_A è equivalente all'indipendenza lineare delle sue colonne.

ESEMPIO 3.6.4. (Il prodotto di matrici non è commutativo.)

Consideriamo le matrici quadrate di ordine $n \times n$. Siano A e B due matrici di tale ordine. Allora possiamo sempre fare il prodotto $A \cdot B$ e $B \cdot A$, ma in generale otteniamo due risultati diversi. Ecco un esempio.

Siano A e B le seguenti matrici quadrate di ordine 2×2 :

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

Allora il prodotto $A \cdot B$ è

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

che è diverso dal prodotto $B \cdot A$

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

Ciò dimostra che il prodotto di matrici *non* è commutativo, cioè in generale $AB \neq BA$.

□

3.7. Rango di una matrice

Osserviamo che, a causa di (3.6.1), l'immagine di T_A è il sottospazio lineare di \mathbb{R}^m generato dalle colonne di A . Perciò il *rango* di A , definito come la dimensione dell'immagine dell'applicazione lineare $T_A : \mathbb{R}^n \ni \mathbf{x} \mapsto A\mathbf{x} \in \mathbb{R}^m$, è uguale alla dimensione del sottospazio lineare di \mathbb{R}^m generato dalle colonne di A . Poiché ogni sistema di generatori contiene una base, possiamo dire che il rango di A è il numero massimo di colonne linearmente indipendenti. Indichiamo il rango di A con $\text{rg}(A)$ o $\text{rg}A$. Poiché A ha n colonne, $\text{rg}(A) \leq n$. D'altro canto, poiché le colonne di A appartengono ad \mathbb{R}^m e la dimensione di ogni sottospazio lineare di \mathbb{R}^m è minore o uguale alla dimensione m di \mathbb{R}^m , abbiamo anche $\text{rg}(A) \leq m$. Quindi:

$$\text{rg}(A) \leq \max\{m, n\}, \quad A \in M_{mn}.$$

NOTA 3.7.1. (**Invertibilità di una matrice.**) L'iniettività di T_A è equivalente a $\text{rg}(A) = n$, mentre la surgettività di T_A è equivalente a $\text{rg}(A) = m$. Perciò T_A è bigettiva se e solo se A è quadrata e $\text{rg}(A) = m = n$. In questo caso A si chiama *invertibile*. \square

DEFINIZIONE 3.7.2. (**Matrice trasposta.**) La *trasposta* di una matrice $A \in M_{mn}$ è la matrice $A^\top \in M_{nm}$ che si ottiene da A cambiando le colonne in righe (ed analogamente le righe in colonne):

$$\left(\begin{array}{cccc} | & | & & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_n \\ | & | & & | \end{array} \right)^\top = \left(\begin{array}{ccc} - & \mathbf{c}_1 & - \\ - & \mathbf{c}_2 & - \\ & \vdots & \\ - & \mathbf{c}_n & - \end{array} \right),$$

cioè

$$\left(\begin{array}{cccc} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{array} \right)^\top = \left(\begin{array}{cccc} \alpha_{11} & \alpha_{21} & \dots & \alpha_{m1} \\ \alpha_{12} & \alpha_{22} & \dots & \alpha_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1n} & \alpha_{2n} & \dots & \alpha_{mn} \end{array} \right).$$

NOTA 3.7.3. (**Regole di calcolo per la matrice trasposta.**)

$$\begin{aligned} (A+B)^\top &= A^\top + B^\top \quad \text{per } A, B \in M_{mn}, \\ (\lambda A)^\top &= \lambda A^\top, \\ (BA)^\top &= A^\top B^\top \quad \text{per } A \in M_{mn}, B \in M_{pm}, \\ (A^\top)^\top &= A. \end{aligned}$$

\square

PROPOSIZIONE 3.7.4.

$$\text{rg}(A) = \text{rg}(A^\top), \quad A \in M_{mn}. \quad (3.7.1)$$

Dimostrazione. (facoltativa.) Osserviamo che la trasposta \mathbf{x}^\top di ogni $\mathbf{x} \in \mathbb{R}^n$, considerata come matrice $n \times 1$, è una matrice $1 \times n$, cioè un vettore riga. Usando le regole di calcolo di cui sopra, otteniamo una uguaglianza importante:

$$\mathbf{y}^\top A \mathbf{x} = (A^\top \mathbf{y})^\top \mathbf{x}, \quad A \in M_{mn}, \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m \quad (3.7.2)$$

(tutti i prodotti di matrici in (3.7.2) hanno senso perché il numero di colonne di ciascuna corrisponde al numero di righe della successiva). Ne segue che

$$\{\mathbf{x} \in \mathbb{R}^n; A \mathbf{x} = \mathbf{0}_m\} \cap A^\top \mathbb{R}^m = \{\mathbf{0}_n\}, \quad A \in M_{mn}. \quad (3.7.3)$$

Infatti, se

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

è tale che nello stesso tempo $A \mathbf{x} = \mathbf{0}_m$ e $\mathbf{x} = A^\top \mathbf{y}$ per almeno un $\mathbf{y} \in \mathbb{R}^m$, allora (3.7.2) implica che

$$\mathbf{x}^\top \mathbf{x} = (A^\top \mathbf{y})^\top \mathbf{x} = \mathbf{y}^\top A \mathbf{x} = \mathbf{y}^\top \mathbf{0}_m = (0).$$

Ma l'unico elemento di $\mathbf{x}^\top \mathbf{x}$ è uguale a $x_1^2 + \dots + x_n^2$, perciò dobbiamo avere $x_1 = \dots = x_n = 0$, cioè $\mathbf{x} = \mathbf{0}_n$.

(Vedremo più avanti che l'applicazione

$$\mathbb{R}^n \times \mathbb{R}^n \ni (\mathbf{x}, \mathbf{y}) \longmapsto \mathbf{y}^\top \mathbf{x} \in \mathbb{R},$$

chiamata il *prodotto scalare (naturale)* su \mathbb{R}^n , è di importanza fondamentale anche per altri propositi.)

Ora siamo pronti per dimostrare che

$$\text{rg}(A^\top) \leq \text{rg}(A) \quad \text{per ogni } A \in M_{mn}. \quad (3.7.4)$$

Indichiamo con r il rango di $\text{rg}(A^\top)$. Allora esistono r colonne di $\text{rg}(A^\top)$, diciamo $\mathbf{x}_1, \dots, \mathbf{x}_r \in \mathbb{R}^n$, linearmente indipendenti. Verifichiamo che i vettori $A \mathbf{x}_1, \dots, A \mathbf{x}_r$ sono ancora linearmente indipendenti:

Se $\sum_{j=1}^r \lambda_j A \mathbf{x}_j = \mathbf{0}_m$, cioè $A \left(\sum_{j=1}^r \lambda_j \mathbf{x}_j \right) = \mathbf{0}_m$, allora $\sum_{j=1}^r \lambda_j \mathbf{x}_j$ appartiene al nucleo di $\mathbb{R}^n \ni \mathbf{x} \rightarrow A \mathbf{x} \in \mathbb{R}^m$. Ma questa somma si trova anche in $A^\top \mathbb{R}^m$, perciò (3.7.3) implica che $\sum_{j=1}^r \lambda_j \mathbf{x}_j = \mathbf{0}_n$. Siccome i vettori $\mathbf{x}_1, \dots, \mathbf{x}_r$ sono linearmente indipendenti, concludiamo che tutti i coefficienti $\lambda_1, \dots, \lambda_r$ si annullano.

Ne segue che la dimensione dell'immagine di $\mathbb{R}^n \ni \mathbf{x} \rightarrow A \mathbf{x} \in \mathbb{R}^m$, cioè il rango di A , è maggiore o uguale di r . Questo prova (3.7.4).

Infine, applicando la disuguaglianza dimostrata per A^\top , otteniamo anche la disuguaglianza inversa: $\text{rg}(A) = \text{rg}((A^\top)^\top) \leq \text{rg}(A^\top)$. \square

NOTA 3.7.5. Per (3.7.1) il rango di una matrice è uguale anche alla dimensione dello spazio lineare generato dalle sue righe, ovvero al numero massimo di righe linearmente indipendenti. \square

COROLLARIO 3.7.6. (Calcolo del rango.) *Per calcolare il rango di una matrice A , basta applicare l'eliminazione di Gauss ad A e contare il numero delle righe non zero nella tabella ottenuto.*

Dimostrazione. I passaggi dell'eliminazione di Gauss non cambiano lo spazio lineare generato dalle righe e le righe non zero nella tabella finale costituiscono una base di questo spazio lineare. Possiamo calcolare il rango di A anche applicando l'eliminazione di Gauss alle righe della matrice trasposta, cioè alle colonne di A . \square

ESERCIZIO 3.7.7. Si determini il rango della matrice

$$\begin{pmatrix} 3 & 1 & 1 & 4 \\ 0 & 4 & 10 & 1 \\ 1 & 7 & 17 & 3 \\ 2 & 2 & 4 & 3 \end{pmatrix}.$$

Per semplificare i calcoli fin dall'inizio, sottraiamo prima l'ultima riga dalla prima. Poi svolgiamo l'eliminazione di Gauss usuale, ottenendo

$$\begin{aligned} \operatorname{rg} \begin{pmatrix} 3 & 1 & 1 & 4 \\ 0 & 4 & 10 & 1 \\ 1 & 7 & 17 & 3 \\ 2 & 2 & 4 & 3 \end{pmatrix} &= \operatorname{rg} \begin{pmatrix} 1 & -1 & -3 & 1 \\ 0 & 4 & 10 & 1 \\ 1 & 7 & 17 & 3 \\ 2 & 2 & 4 & 3 \end{pmatrix} \\ &= \operatorname{rg} \begin{pmatrix} 1 & -1 & -3 & 1 \\ 0 & 4 & 10 & 1 \\ 0 & 8 & 20 & 2 \\ 0 & 4 & 10 & 1 \end{pmatrix} = \operatorname{rg} \begin{pmatrix} 1 & -1 & -3 & 1 \\ 0 & 4 & 10 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\ &= 2. \end{aligned}$$

\square

3.8. Applicazioni lineari e matrici invertibili

Abbiamo introdotto la nozione di matrice invertibile nella Nota 3.7.1. Perché una matrice A sia invertibile, per prima cosa A dev'essere quadrata. In tal caso, $A \in M_n$ è invertibile se e solo se $\operatorname{rg}(A)$ è uguale ad n , cioè assume il valore massimo possibile.

Sia $A \in M_n$. Se A è invertibile, allora l'applicazione lineare $T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ ha una inversa $T_A^{-1} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, che è anch'essa lineare (verificare per esercizio). La matrice associata a

T_A^{-1} si chiama *l'inversa* di A ed è indicata con A^{-1} . Poiché $T_A^{-1} \circ T_A$ e $T_A \circ T_A^{-1}$ sono uguali all'applicazione identica su \mathbb{R}^n , per le matrici associate abbiamo

$$A A^{-1} = A^{-1} A = I_n.$$

PROPOSIZIONE 3.8.1. Se $A, B \in M_n$

$$A B = I_n \iff B A = I_n \implies A \text{ è invertibile ed } A^{-1} = B. \quad (\text{I})$$

Dimostrazione. (facoltativa.) Supponiamo dapprima che $A B = I_n$. Allora $T_A \circ T_B = T_{AB} = T_{I_n} = I_{\mathbb{R}^n}$. Per (3.2.6) T_A è surgettiva e poi (3.2.4) implica che T_A è di fatto bigettiva. Perciò A è invertibile e

$$B = I_n B = (A^{-1} A) B = A^{-1} (A B) = A^{-1} I_n = A^{-1}.$$

Analogamente, se $B A = I_n$, allora $T_B \circ T_A = T_{BA} = T_{I_n} = I_{\mathbb{R}^n}$ implica l'iniettività, e poi la bigettività di T_A . Di conseguenza, A è anche in questo caso invertibile e

$$B = B I_n = B (A A^{-1}) = (B A) A^{-1} = I_n A^{-1} = A^{-1}.$$

□

3.9. Il calcolo della matrice inversa

Usiamo l'uguaglianza (I) per calcolare la matrice inversa.

Sia infatti

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix}$$

una matrice quadrata di ordine n e poniamoci il problema di verificare l'invertibilità di A e, nel caso di invertibilità, di calcolare la matrice inversa A^{-1} .

Grazie ad (I), il problema posto è equivalente a verificare l'esistenza di una matrice

$$B = \begin{pmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \dots & \beta_{nn} \end{pmatrix} = \begin{pmatrix} | & | & & | \\ \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_n \\ | & | & & | \end{pmatrix}$$

con $A B = I_n$ e, nel caso di esistenza, a calcolarla. Ma $A B = I_n$ significa

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix} \begin{pmatrix} | & | & & | \\ \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_n \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \\ | & | & & | \end{pmatrix},$$

ove $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ è la base naturale di \mathbb{R}^n , cioè i n sistemi di equazioni lineari

$$\begin{pmatrix} & \\ & A \\ & \end{pmatrix} \begin{pmatrix} | \\ \mathbf{b}_1 \\ | \end{pmatrix} = \begin{pmatrix} | \\ \mathbf{v}_1 \\ | \end{pmatrix},$$

.....

(3.9.1)

$$\begin{pmatrix} & \\ & A \\ & \end{pmatrix} \begin{pmatrix} | \\ \mathbf{b}_n \\ | \end{pmatrix} = \begin{pmatrix} | \\ \mathbf{v}_n \\ | \end{pmatrix}.$$

Per la soluzione di questi n sistemi usiamo il metodo di eliminazione di Gauss. Poiché gli coefficienti sono i stessi, possiamo svolgere l'eliminazione di Gauss *simultaneamente per tutti i sistemi*.

Partiamo quindi con la tabella

$$\begin{array}{cccc|cccc} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} & 1 & 0 & \dots & 0 \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} & 0 & 0 & \dots & 1 \end{array}.$$

Tramite il processo di eliminazione di Gauss otteniamo una tabella con la matrice dei coefficienti triangolare superiore, cioè n sistemi triangolari superiori:

$$\begin{array}{cccc|cccc} \alpha'_{11} & \alpha'_{12} & \dots & \alpha'_{1n} & \gamma'_{11} & \gamma'_{12} & \dots & \gamma'_{1n} \\ 0 & \alpha'_{22} & \dots & \alpha'_{2n} & \gamma'_{21} & \gamma'_{22} & \dots & \gamma'_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha'_{nn} & \gamma'_{n1} & \gamma'_{n2} & \dots & \gamma'_{nn} \end{array}.$$
(3.9.2)

Se uno degli elementi diagonali $\alpha'_{11}, \alpha'_{22}, \dots, \alpha'_{nn}$ si annulla, allora il rango di A è $< n$ e quindi A non è invertibile. Infatti, se per esempio α'_{33} è il primo elemento zero sulla diagonale, allora la terza colonna di coefficienti

$$\begin{pmatrix} \gamma'_{13} \\ \gamma'_{23} \\ \gamma'_{33} \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma'_{13} \\ \gamma'_{23} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

appartiene al sottospazio lineare bidimensionale

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} ; x_1, x_2 \in \mathbb{R} \right\}$$

di \mathbb{R}^n , generato dai vettori linearmente indipendenti

$$\begin{pmatrix} \gamma'_{11} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \gamma'_{12} \\ \gamma'_{22} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \gamma'_{11}, \gamma'_{22} \neq 0.$$

Se invece ogni elemento sulla diagonale è non zero, allora A è invertibile. Per trovare la matrice inversa, cioè risolvere i n sistemi (3.9.1) che sono equivalenti ai n sistemi (3.9.2), applichiamo alla tabella (3.9.2) il processo di eliminazione di Gauss, partendo dal coefficiente *pivot* γ'_{nn} nell'angolo destro inferiore e procedendo verso alto ed a sinistra:

Sommando alla $(n-1)$ -esima riga il multiplo con $-\frac{\alpha'_{n-1,n}}{\alpha'_{nn}}$ dell'ultima riga, azzeriamo

l'elemento sopra α'_{nn} . Poi, sommando alla $(n-2)$ -esima riga il multiplo con $-\frac{\alpha'_{n-2,n}}{\alpha'_{nn}}$ dell'ultima riga, azzeriamo anche il secondo elemento sopra α'_{nn} . Dopo $n-1$ passi la matrice dei coefficienti diventa

$$\begin{array}{ccccc} \alpha'_{11} & \alpha''_{12} & \dots & \alpha''_{1,n-1} & 0 \\ 0 & \alpha'_{22} & \dots & \alpha''_{2,n-1} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \alpha'_{n-1,n-1} & 0 \\ 0 & 0 & \dots & 0 & \alpha'_{nn} \end{array}$$

Ora, sommando successivamente multipli scalari appropriati della $(n-1)$ -esima riga a tutte le righe precedenti, azzeriamo tutti gli elementi della $(n-1)$ -esima colonna che si trovano sopra α'_{nn} . Procediamo in questo modo fino all'azzeramento del coefficiente sopra α'_{22} .

Il risultato è una tabella che rappresenta n sistemi diagonali equivalenti a (3.9.1) :

$$\begin{array}{ccccc|cccc} \alpha'_{11} & 0 & \dots & 0 & 0 & \gamma''_{11} & \gamma''_{12} & \dots & \gamma''_{1n} \\ 0 & \alpha'_{22} & \dots & 0 & 0 & \gamma''_{21} & \gamma''_{22} & \dots & \gamma''_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha'_{n-1,n-1} & 0 & \gamma''_{n-1,1} & \gamma''_{n-1,2} & \dots & \gamma''_{n-1,n-1} \\ 0 & 0 & \dots & 0 & \alpha'_{nn} & \gamma''_{n1} & \gamma''_{n2} & \dots & \gamma''_{nn} \end{array}$$

Poiché tutti gli elementi sulla diagonale sono non zero, possiamo dividere ogni k -esima riga con l'elemento diagonale α'_{kk} su di essa, ottenendo

$$\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & 0 & \gamma''_{11} & \gamma''_{12} & \dots & \gamma''_{1n} \\ 0 & 1 & \dots & 0 & 0 & \gamma''_{21} & \gamma''_{22} & \dots & \gamma''_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & \gamma''_{n-1,1} & \gamma''_{n-1,2} & \dots & \gamma''_{n-1,n-1} \\ 0 & 0 & \dots & 0 & 1 & \gamma''_{n1} & \gamma''_{n2} & \dots & \gamma''_{nn} \end{array} \quad (3.9.3)$$

Ora da (3.9.3) leggiamo che la soluzione del primo sistema in (3.9.1), cioè la prima colonna della matrice B inversa ad A , è

$$\begin{pmatrix} | \\ \mathbf{b}_1 \\ | \end{pmatrix} = \begin{pmatrix} \gamma'_{11} \\ \vdots \\ \gamma'_{n1} \end{pmatrix}.$$

Poi, la soluzione del secondo sistema in (3.9.1), cioè la seconda colonna della matrice B inversa ad A , è

$$\begin{pmatrix} | \\ \mathbf{b}_2 \\ | \end{pmatrix} = \begin{pmatrix} \gamma'_{12} \\ \vdots \\ \gamma'_{n2} \end{pmatrix}$$

e così via fino all'uguaglianza

$$\begin{pmatrix} | \\ \mathbf{b}_n \\ | \end{pmatrix} = \begin{pmatrix} \gamma'_{1n} \\ \vdots \\ \gamma'_{nn} \end{pmatrix}.$$

Perciò la matrice inversa di A è la matrice dei termini noti in (3.9.3):

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix}^{-1} = \begin{pmatrix} \gamma''_{11} & \gamma''_{12} & \dots & \gamma''_{1n} \\ \gamma''_{21} & \gamma''_{22} & \dots & \gamma''_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma''_{n1} & \gamma''_{n2} & \dots & \gamma''_{nn} \end{pmatrix}.$$

□

ESEMPIO 3.9.1. Vediamo se la matrice

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

è invertibile o no e nel caso di invertibilità calcoliamone l'inversa.

Svolgimento. Tramite l'eliminazione di Gauss riduciamo A a forma triangolare superiore:

$$\begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 & 1 & -2 & 1 & 0 & 0 & 1 \\ -1 & 3 & 0 & 0 & 1 & 0 & 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 1 & -1 & 3 & 0 & 0 & 1 & 0 \end{array},$$

$$\begin{array}{ccc|ccc} 1 & -2 & 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \end{array} , \quad \begin{array}{ccc|ccc} 1 & -2 & 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 & 1 & 1 \end{array} .$$

Poiché tutti gli elementi sulla diagonale sono non zero, A è invertibile. Per calcolare la matrice inversa, tramite operazioni elementari sulle righe azzeriamo anche i coefficienti sopra la diagonale:

$$\begin{array}{ccc|ccc} 1 & -2 & 0 & -1 & 1 & 2 \\ 0 & 1 & 0 & -1 & 2 & 2 \\ 0 & 0 & -1 & -1 & 1 & 1 \end{array} , \quad \begin{array}{ccc|ccc} 1 & 0 & 0 & -3 & 5 & 6 \\ 0 & 1 & 0 & -1 & 2 & 2 \\ 0 & 0 & -1 & -1 & 1 & 1 \end{array} .$$

Ora normalizziamo ogni riga dividendone tutti gli elementi per il suo elemento diagonale: in tal modo otteniamo come matrice di coefficienti la matrice unità:

$$\begin{array}{ccc|ccc} 1 & 0 & 0 & -3 & 5 & 6 \\ 0 & 1 & 0 & -1 & 2 & 2 \\ 0 & 0 & 1 & 1 & -1 & -1 \end{array} .$$

Concludiamo che

$$A^{-1} = \begin{pmatrix} -3 & 5 & 6 \\ -1 & 2 & 2 \\ 1 & -1 & -1 \end{pmatrix} .$$

□

3.10. Minori ed orli di una matrice ed invertibilità

DEFINIZIONE 3.10.1. (**Minori, ovvero sottomatrici.**) Sia

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix}$$

una matrice $m \times n$. Una *sottomatrice* (detta anche un *minore*) di A è una matrice che si ottiene cancellando alcune righe e colonne di A :

$$B = \begin{pmatrix} \alpha_{j_1 k_1} & \alpha_{j_1 k_2} & \dots & \alpha_{j_1 k_q} \\ \alpha_{j_2 k_1} & \alpha_{j_2 k_2} & \dots & \alpha_{j_2 k_q} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{j_p k_1} & \alpha_{j_p k_2} & \dots & \alpha_{j_p k_q} \end{pmatrix}$$

è la sottomatrice ottenuta tramite la cancellazione di tutte le righe di A tranne quelle con indici $j_1 < j_2 < \dots < j_p$, e di tutte le colonne tranne quelle con indici $k_1 < k_2 < \dots < k_q$, ove $1 \leq p \leq n$ e $1 \leq q \leq m$.

NOTA 3.10.2. Se B è una sottomatrice di A , allora $\text{rg}(B) \leq \text{rg}(A)$.

□

DEFINIZIONE 3.10.3. (**Orli di una matrice.**) Sia B una sottomatrice di A . Una sottomatrice B' di A si ottiene *orlando* B se B' si ottiene cancellando una riga ed una colonna di meno di quante si cancellano per ottenere B .

Per esempio, togliendo le righe di indici 1 e 2, nonché le colonne di indici 2, 3 e 5 della matrice

$$A = \begin{pmatrix} 7 & -1 & 0 & 2 & 3 \\ -1 & 2 & 3 & 1 & -1 \\ -2 & 5 & -4 & 7 & 0 \\ 0 & -1 & 1 & 0 & 1 \end{pmatrix},$$

si ottiene la sottomatrice

$$B = \begin{pmatrix} -2 & 7 \\ 0 & 0 \end{pmatrix}.$$

Una sottomatrice di A che si ottiene orlando B è, per esempio,

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

che si ottiene cancellando solo la prima riga e le colonne di indici 2 e 3 di A (abbiamo orlato B con la seconda riga e la quinta colonna di A). Un altro esempio si ottiene orlando B con la prima riga e la terza colonna di A :

$$\begin{pmatrix} 7 & 0 & 2 \\ -2 & -4 & 7 \\ 0 & 1 & 0 \end{pmatrix}.$$

Il rango di una matrice qualsiasi può essere espresso per mezzo di sottomatrici quadrate:

TEOREMA 3.10.4. (**Teorema degli orli.**) Sia A una matrice $m \times n$. Allora $\text{rg}(A) = r$ se e solo se esiste una sottomatrice quadrata B di ordine r di A , tale che

- B è invertibile e
- ogni sottomatrice quadrata di ordine $r + 1$ di A , che si ottiene orlando B , è non invertibile.

Dimostrazione. (facoltativa.) Sia $\text{rg}(A) = r$. Allora esistono r colonne linearmente indipendenti, diciamo quelli con indici

$$k_1 < k_2 < \dots < k_r,$$

di A . Se cancelliamo tutte le altre colonne, otteniamo una sottomatrice A' di A con m righe ed r colonne, tale che $\text{rg}(A') = r$. Ora, per (3.7.1), esistono r righe linearmente indipendenti, diciamo quelle con indici

$$j_1 < j_2 < \dots < j_r,$$

di A' . Se cancelliamo tutte le altre righe di A' otteniamo una sottomatrice quadrata B di A' di ordine r con tutte le righe linearmente indipendenti, quindi B è invertibile. Ovviamente, B è una sottomatrice anche di A e si verifica facilmente che ogni sottomatrice quadrata di ordine $r + 1$ di A che si ottiene orlando B è non invertibile.

Supponiamo adesso che esista una sottomatrice quadrata B di ordine r di A tale che le due condizioni nell'enunciato siano soddisfatte. Siano

$$j_1 < j_2 < \dots < j_r,$$

gli indici delle righe di A e

$$k_1 < k_2 < \dots < k_r,$$

gli indici delle colonne di A che non si cancellano per ottenere B . Allora le colonne di A con indici k_1, k_2, \dots, k_r sono linearmente indipendenti. Verifichiamo che tutte le altre colonne di A sono combinazioni lineari di queste, dimostrando quindi che il rango di A è uguale ad r .

Supponiamo che una colonna di A , il cui indice indichiamo con k' , non sia combinazione lineare delle colonne con indici k_1, k_2, \dots, k_r . Allora le colonne con indici k', k_1, k_2, \dots, k_r sono linearmente indipendenti. Pertanto il rango della sottomatrice A'' di A che si ottiene cancellando tutte le colonne con indici diversi da k', k_1, k_2, \dots, k_r è uguale a $r + 1$.

Per (3.7.1) sappiamo che la dimensione del sottospazio lineare di \mathbb{R}^m generato dalle righe di A'' è uguale ad $r + 1$. Ma le righe con indici j_1, j_2, \dots, j_r , cioè le righe di B , sono linearmente indipendenti. Se tutte le altre righe di A'' fossero combinazioni lineari di queste, allora il rango di A'' sarebbe r , perciò esiste almeno una riga di A'' , diciamo quella con indice j' , che non è combinazione lineare delle righe con indici j_1, j_2, \dots, j_r . Allora le righe di A'' con indici $j'', j_1, j_2, \dots, j_r$ sono linearmente indipendenti e quindi la sottomatrice quadrata B'' di ordine $r + 1$ di A'' che si ottiene cancellando tutte le righe con indice diverso da $j'', j_1, j_2, \dots, j_r$ è invertibile. Ma B'' è una sottomatrice di A ottenuta orlando B e quindi otteniamo una contraddizione all'ipotesi fatta su B . \square

Il rango interviene nel seguente criterio per la risolubilità di un sistema di equazioni lineari:

TEOREMA 3.10.5. (Rouché-Capelli.) *Siano*

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix}$$

una matrice $m \times n$ e

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m.$$

Allora il sistema $A\mathbf{x} = \mathbf{b}$ di m equazioni lineari in n incognite ammette (almeno) una soluzione se e solo se

$$\operatorname{rg} \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix} = \operatorname{rg} \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} & b_1 \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \underbrace{\alpha_{m1} \ \alpha_{m2} \ \dots \ \alpha_{mn}}_{=A} & \underbrace{b_m}_{=\mathbf{b}} \end{pmatrix}.$$

Dimostrazione. (facoltativa.) Siano $\mathbf{c}_1, \dots, \mathbf{c}_n$ le colonne di A e supponiamo che il sistema $A\mathbf{x} = \mathbf{b}$ ammetta una soluzione

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n.$$

Allora dalla formula (3.6.1) segue

$$\mathbf{b} = A\mathbf{x} = \sum_{j=1}^n x_j \mathbf{c}_j.$$

Perciò i vettori $\mathbf{c}_1, \dots, \mathbf{c}_n, \mathbf{b}$ generano lo stesso sottospazio lineare di \mathbb{R}^m di $\mathbf{c}_1, \dots, \mathbf{c}_n$, la cui dimensione è uguale a

$$\operatorname{rg} \begin{pmatrix} | & | & & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_n \\ | & | & & | \end{pmatrix} = \operatorname{rg} \begin{pmatrix} | & | & & | & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_n & \mathbf{b} \\ | & | & & | & | \end{pmatrix}. \quad (3.10.1)$$

$\underbrace{\hspace{10em}}_{=A}$

Supponiamo viceversa che valga (3.10.1). Se \mathbf{X} ed $\mathbf{X}_{\mathbf{b}}$ indicano i sottospazi lineari di \mathbb{R}^m generati rispettivamente

- (1) dalle colonne di A e
- (2) dalle colonne di A e dal termine noto \mathbf{b} ,

allora (3.10.1) significa che $\dim \mathbf{X} = \dim \mathbf{X}_{\mathbf{b}}$ e quindi implica l'uguaglianza $\mathbf{X} = \mathbf{X}_{\mathbf{b}}$. Ora, se il rango di A , è r , esistono r colonne $\mathbf{c}_{k_1}, \dots, \mathbf{c}_{k_r}$ di A che costituiscono una base per \mathbf{X} . Poiché $\mathbf{b} \in \mathbf{X}$, segue che \mathbf{b} è combinazione lineare dei vettori $\mathbf{c}_{k_1}, \dots, \mathbf{c}_{k_r}$, e quindi, a maggior ragione, di tutte le colonne di A . Ora, per (3.6.1), questo significa che il sistema $A\mathbf{x} = \mathbf{b}$ ammette soluzioni. \square

ESEMPIO 3.10.6. Troviamo tutti i valori del parametro reale a per i quali il sistema di equazioni lineari

$$\begin{cases} 2x_1 - x_2 - x_3 = 1 \\ -2x_1 + a^2x_2 + (1-a)x_3 = -1 \\ x_1 + 3x_2 + x_3 = 2 \end{cases}$$

non ha soluzione.

Svolgimento. Per il teorema di Rouché-Capelli 3.10.5 il sistema ha soluzione se e soltanto se

$$\operatorname{rg} \begin{pmatrix} 2 & -1 & -1 \\ -2 & a^2 & 1-a \\ 1 & 3 & 1 \end{pmatrix} = \operatorname{rg} \begin{pmatrix} 2 & -1 & -1 & 1 \\ -2 & a^2 & 1-a & -1 \\ 1 & 3 & 1 & 2 \end{pmatrix}.$$

Tramite operazioni elementari sulle righe si ottiene

$$\begin{aligned} \operatorname{rg} \begin{pmatrix} 2 & -1 & -1 \\ -2 & a^2 & 1-a \\ 1 & 3 & 1 \end{pmatrix} &= \operatorname{rg} \begin{pmatrix} 1 & 3 & 1 \\ 2 & -1 & -1 \\ -2 & a^2 & 1-a \end{pmatrix} = \\ &= \operatorname{rg} \begin{pmatrix} 1 & 3 & 1 \\ 0 & -7 & -3 \\ 0 & 6+a^2 & 3-a \end{pmatrix} = \operatorname{rg} \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & \frac{3}{7} \\ 0 & 6+a^2 & 3-a \end{pmatrix} = \\ &= \operatorname{rg} \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & \frac{3}{7} \\ 0 & 0 & \frac{1}{7}(3-7a-3a^2) \end{pmatrix} = \begin{cases} 3 & \text{se } 3-7a-3a^2 \neq 0, \\ 2 & \text{se } 3-7a-3a^2 = 0. \end{cases} \end{aligned}$$

Ne segue che nel caso $3-7a-3a^2 \neq 0$ il sistema ha soluzione. Se invece $3-7a-3a^2 = 0$, allora abbiamo

$$\begin{aligned} \operatorname{rg} \begin{pmatrix} 2 & -1 & -1 & 1 \\ -2 & a^2 & 1-a & -1 \\ 1 & 3 & 1 & 2 \end{pmatrix} &= \operatorname{rg} \begin{pmatrix} 1 & 3 & 1 & 2 \\ 2 & -1 & -1 & 1 \\ -2 & a^2 & 1-a & -1 \end{pmatrix} = \\ &= \operatorname{rg} \begin{pmatrix} 1 & 3 & 1 & 2 \\ 0 & -7 & -3 & -3 \\ 0 & 6+a^2 & 3-a & 3 \end{pmatrix} = \operatorname{rg} \begin{pmatrix} 1 & 3 & 1 & 2 \\ 0 & 1 & \frac{3}{7} & \frac{3}{7} \\ 0 & 6+a^2 & 3-a & 3 \end{pmatrix} = \\ &= \operatorname{rg} \begin{pmatrix} 1 & 3 & 1 & 2 \\ 0 & 1 & \frac{3}{7} & \frac{3}{7} \\ 0 & 0 & \frac{1}{7}(3-7a-3a^2) & \frac{1}{7}(3-3a^2) \end{pmatrix} = \\ &= \operatorname{rg} \begin{pmatrix} 1 & 3 & 1 & 2 \\ 0 & 1 & \frac{3}{7} & \frac{3}{7} \\ 0 & 0 & 0 & a \end{pmatrix} = 3 \end{aligned}$$

perché $3-7a-3a^2 = 0 \Rightarrow a \neq 0$. Perciò il sistema non ha soluzione esattamente quando $3-7a-3a^2 = 0$, ossia $a = \frac{-7 \pm \sqrt{85}}{6}$. \square

3.11. Esercizi

ESERCIZIO 3.11.1. Si verifichi che l'applicazione

$$\mathbb{R}^3 \ni \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto \begin{pmatrix} x_1 + x_2 \\ x_2 + x_3 \end{pmatrix} \in \mathbb{R}^2$$

è lineare e si trovi la matrice associata.

Soluzione. Gli elementi della base naturale di \mathbb{R}^* sono mandati in

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

Pertanto la matrice associata è

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

e l'applicazione può essere scritta nella forma

$$\mathbb{R}^3 \ni \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^2.$$

□

ESERCIZIO 3.11.2. Si trovi una matrice A di tipo 4×3 tale che, per ogni $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$, la combinazione lineare

$$\lambda_1 \begin{pmatrix} 2 \\ -3 \\ 0 \\ -1 \end{pmatrix} + \lambda_2 \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 5 \\ 2 \\ -3 \end{pmatrix}$$

sia uguale al prodotto

$$A \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix}.$$

Soluzione. Si usa l'uguaglianza (3.6.1) :

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

□

ESERCIZIO 3.11.3. Si trovi la matrice dell'applicazione lineare $T : \mathbb{R}_3 \longrightarrow \mathbb{R}^3$ che manda

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 3 \\ 4 \\ -2 \end{pmatrix}.$$

Soluzione. Le colonne della matrice A associata a T sono le immagini tramite T degli elementi della base naturale di \mathbb{R}^3 , perciò

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

□

ESERCIZIO 3.11.4. Esiste una applicazione lineare $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ che manda

$$\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}?$$

Se esistono tali applicazioni, se ne trovi una.

Soluzione. Se T esistesse, sarebbe surgettiva, quindi, per (I), bigettiva. Inoltre T^{-1} manderebbe

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix},$$

perciò, come nell'Esercizio 3.11.2, la matrice associata a T^{-1} dovrebbe essere

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

Ma questa matrice è invertibile, avendo l'inversa

$$\begin{pmatrix} 0 & 1 & 2 \\ -1 & 3 & 0 \\ 1 & -2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -3 & 5 & 6 \\ -1 & 2 & 2 \\ 1 & -1 & -1 \end{pmatrix}$$

(si veda l'Esempio 3.9.1). Ne concludiamo che

$$T : \mathbb{R}^3 \ni \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto \begin{pmatrix} -3 & 5 & 6 \\ -1 & 2 & 2 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$$

è l'unica applicazione lineare con le proprietà desiderate.

□

ESERCIZIO 3.11.5. Si trovi il nucleo e l'immagine dell'applicazione lineare

$$T : \mathbb{R}^3 \ni \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} 2 & -1 & 1 \\ 1 & 1 & 5 \\ 2 & 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 .$$

Soluzione. Il nucleo di T consiste da tutte le soluzioni del sistema omogeneo

$$\begin{cases} 2x_1 - x_2 + x_3 = 0 \\ x_1 + x_2 + 5x_3 = 0 \\ 2x_1 + 4x_3 = 0 \end{cases}$$

che ora risolviamo.

$$\begin{array}{ccc|ccc|ccc|ccc|ccc|} 2 & -1 & 1 & 0 & 1 & 1 & 5 & 0 & 1 & 1 & 5 & 0 & 1 & 1 & 5 & 0 \\ 1 & 1 & 5 & 0 & 2 & -1 & 1 & 0 & 0 & -3 & -9 & 0 & 0 & 1 & 3 & 0 \\ 2 & 0 & 4 & 0 & 2 & 0 & 4 & 0 & 0 & -2 & -6 & 0 & 0 & 0 & 0 & 0 \end{array} ,$$

$x_3 =$ arbitrario

$$\begin{aligned} x_2 &= -3x_3 \\ x_1 &= -x_2 - 5x_3 = -2x_3 \end{aligned} , \text{ cioè } \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2x_3 \\ -3x_3 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} -2 \\ -3 \\ 1 \end{pmatrix} .$$

Concludiamo che $\text{Ker}(T)$ consiste da tutti i multipli scalari di $\begin{pmatrix} -2 \\ -3 \\ 1 \end{pmatrix}$:

$$\text{Ker}(T) = \left\{ \lambda \begin{pmatrix} -2 \\ -3 \\ 1 \end{pmatrix} ; \lambda \in \mathbb{R} \right\} .$$

D'altra parte, l'immagine di T è il sottospazio lineare di \mathbb{R}^3 generato dalle colonne della matrice associata, cioè dai vettori

$$\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} , \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} , \begin{pmatrix} 1 \\ 5 \\ 4 \end{pmatrix} .$$

Volendo, possiamo trovare una base per questo sottospazio, scrivendo i vettori di cui sopra come le righe di una matrice, svolgendo l'eliminazione di Gauss e poi considerando le righe non zero della matrice risultante:

$$\begin{array}{cccc} 2 & 1 & 2 & -1 & 1 & 0 & -1 & 1 & 0 & -1 & 1 & 0 \\ -1 & 1 & 0 & 2 & 1 & 2 & 0 & 3 & 2 & 0 & 3 & 2 \\ 1 & 5 & 4 & 1 & 5 & 4 & 0 & 6 & 4 & 0 & 0 & 0 \end{array}$$

Così otteniamo la base $\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix}$ per $T(\mathbb{R}^3)$:

$$T(\mathbb{R}^3) = \left\{ \lambda_1 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} ; \lambda_1, \lambda_2 \in \mathbb{R} \right\}.$$

□

ESERCIZIO 3.11.6. Si determini il rango della matrice

$$\begin{pmatrix} 2 & 3 & 1 & 1 \\ 1 & -1 & 5 & 3 \\ 5 & 5 & 3 & 5 \\ 3 & 7 & -7 & -1 \end{pmatrix}.$$

Soluzione. Tramite operazioni elementari sulle righe si ottiene

$$\begin{aligned} \text{rg} \begin{pmatrix} 2 & 3 & 1 & 1 \\ 1 & -1 & 5 & 3 \\ 5 & 5 & 3 & 5 \\ 3 & 7 & -7 & -1 \end{pmatrix} &= \text{rg} \begin{pmatrix} 1 & -1 & 5 & 3 \\ 2 & 3 & 1 & 1 \\ 2 & -2 & 10 & 6 \\ 3 & 7 & -7 & -1 \end{pmatrix} = \\ &= \text{rg} \begin{pmatrix} 1 & -1 & 5 & 3 \\ 0 & 5 & -9 & -5 \\ 0 & 0 & 0 & 0 \\ 0 & 10 & -22 & -10 \end{pmatrix} = \text{rg} \begin{pmatrix} 1 & -1 & 5 & 3 \\ 0 & 5 & -9 & -5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -4 & 0 \end{pmatrix} = \\ &= 3. \end{aligned}$$

□

ESERCIZIO 3.11.7. Si dica se la matrice

$$A = \begin{pmatrix} 2 & 7 & 3 \\ 3 & 9 & 4 \\ 1 & 5 & 3 \end{pmatrix}$$

è invertibile o no. Nel caso di invertibilità si calcoli la sua matrice inversa.

Soluzione. Tramite operazioni elementari sulle righe si azzerano i coefficienti sotto la diagonale:

$$\begin{array}{ccc|ccc} 2 & 7 & 3 & 1 & 0 & 0 & 1 & 5 & 3 & 0 & 0 & 1 \\ 3 & 9 & 4 & 0 & 1 & 0 & 2 & 7 & 3 & 1 & 0 & 0 \\ 1 & 5 & 3 & 0 & 0 & 1 & 3 & 9 & 4 & 0 & 1 & 0 \end{array},$$

$$\begin{array}{ccc|ccc} 1 & 5 & 3 & 0 & 0 & 1 \\ 0 & -3 & -3 & 1 & 0 & -2 \\ 0 & -6 & -5 & 0 & 1 & -3 \end{array} , \quad \begin{array}{ccc|ccc} 1 & 5 & 3 & 0 & 0 & 1 \\ 0 & 1 & 1 & -1/3 & 0 & 2/3 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array}$$

Poiché tutti gli elementi del diagonale sono non zero, la matrice A è diagonalizzabile. Usando di nuovo operazioni elementari sulle righe, si azzerano anche i coefficienti sopra la diagonale:

$$\begin{array}{ccc|ccc} 1 & 5 & 0 & 6 & -3 & -2 \\ 0 & 1 & 0 & 5/3 & -1 & -1/3 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array} , \quad \begin{array}{ccc|ccc} 1 & 0 & 0 & -7/3 & 2 & -1/3 \\ 0 & 1 & 0 & 5/3 & -1 & -1/3 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array} .$$

Così

$$A^{-1} = \begin{pmatrix} -7/3 & 2 & -1/3 \\ 5/3 & -1 & -1/3 \\ -2 & 1 & 1 \end{pmatrix} .$$

□

ESERCIZIO 3.11.8. Si dica se la matrice

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

è invertibile o no. Nel caso di invertibilità si calcoli la sua matrice inversa.

Soluzione. Tramite operazioni elementari sulle righe si azzerano i coefficienti sotto la diagonale:

$$\begin{array}{cccc|cccc} 2 & 1 & -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 1 & 0 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & -2 & 0 & 0 & 0 & 1 \end{array} \quad \begin{array}{cccc|cccc} -1 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & -2 & 0 & 1 & 0 & -2 & 0 & 1 \\ -3 & 0 & 1 & 0 & -3 & 0 & 1 & 0 \\ 0 & 2 & 1 & -2 & 0 & 2 & 1 & -2 \end{array} \quad \begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{cccc|cccc} -1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & -2 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & -3 & 1 & 0 & -3 & 0 & -2 & 0 \\ 0 & 2 & 1 & -2 & 0 & 0 & 0 & 1 \end{array} \quad \begin{array}{cccc|cccc} -1 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & -2 & 0 & 1 & 0 & -2 & 0 & 1 \\ 0 & -3 & 1 & 0 & 0 & -1 & 1 & -1 \\ 0 & 2 & 1 & -2 & 0 & 0 & 1 & -1 \end{array} \quad \begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ -3 & -1 & -2 & 0 & -3 & -1 & -2 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{array}$$

$$\begin{array}{cccc|cccc} -1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 & 3 & 1 & 2 & 0 \\ 0 & -2 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 & 0 & 1 \end{array} \quad \begin{array}{cccc|cccc} -1 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 & 1 & -1 & 1 \\ 0 & 0 & -2 & 3 & 0 & 0 & -2 & 3 \\ 0 & 0 & 1 & -1 & 0 & 0 & 1 & -1 \end{array} \quad \begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 3 & 1 & 2 & 0 & 3 & 1 & 2 & 0 \\ 6 & 3 & 4 & 0 & 6 & 3 & 4 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{array}$$

$$\begin{array}{cccc|cccc|cccc}
-1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & -1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\
0 & 1 & -1 & 1 & 3 & 1 & 2 & 0 & 0 & 1 & -1 & 1 & 3 & 1 & 2 & 0 \\
0 & 0 & 1 & -1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & -1 & 0 & 1 & 0 & 1 \\
0 & 0 & -2 & 3 & 6 & 3 & 4 & 0 & 0 & 0 & 0 & 1 & 6 & 5 & 4 & 2
\end{array} .$$

Poiché tutti gli elementi della diagonale sono non nulli, la matrice A è diagonalizzabile. Usando di nuovo operazioni elementari sulle righe, si azzeriamo anche i coefficienti sopra la diagonale:

$$\begin{array}{cccc|cccc|cccc|cccc}
1 & -1 & 0 & 0 & -1 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 2 & 2 & 1 & 1 \\
0 & 1 & 0 & 0 & 3 & 2 & 2 & 1 & 0 & 1 & 0 & 0 & 3 & 2 & 2 & 1 \\
0 & 0 & 1 & -1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 6 & 6 & 4 & 3 \\
0 & 0 & 0 & 1 & 6 & 5 & 4 & 2 & 0 & 0 & 0 & 1 & 6 & 5 & 4 & 2
\end{array} .$$

Pertanto

$$A^{-1} = \begin{pmatrix} 2 & 2 & 1 & 1 \\ 3 & 2 & 2 & 1 \\ 6 & 6 & 4 & 3 \\ 6 & 5 & 4 & 2 \end{pmatrix} .$$

□

3.12. Isomorfismi fra spazi vettoriali

CAPITOLO 4

Cambiamento di base

4.1. Trasformazione di coordinate sotto cambiamento di base

Sia \mathbf{X} uno spazio vettoriale ed $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è una base di \mathbf{X} . Allora

$$\mathbb{R}^n \ni \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \mapsto \lambda_1 \mathbf{v}_1 + \dots + \lambda_n \mathbf{v}_n \in \mathbf{X}$$

è una applicazione lineare bigettiva e la sua applicazione inversa è

$$F_{\mathcal{V}} : \mathbf{X} \ni \mathbf{x} \mapsto \begin{pmatrix} \lambda_1^{(\mathcal{V})}(\mathbf{x}) \\ \vdots \\ \lambda_n^{(\mathcal{V})}(\mathbf{x}) \end{pmatrix} \in \mathbb{R}^n,$$

ove $\lambda_1^{(\mathcal{V})}(\mathbf{x}), \dots, \lambda_n^{(\mathcal{V})}(\mathbf{x})$ sono i coefficienti di $\mathbf{v}_1, \dots, \mathbf{v}_n$ nello sviluppo di \mathbf{x} quale combinazione lineare degli elementi della base \mathcal{V} :

$$\mathbf{x} = \lambda_1^{(\mathcal{V})}(\mathbf{x}) \mathbf{v}_1 + \dots + \lambda_n^{(\mathcal{V})}(\mathbf{x}) \mathbf{v}_n.$$

$F_{\mathcal{V}}$ applica \mathbf{v}_k nel k -esimo elemento della base naturale di \mathbb{R}^n , quindi \mathcal{V} nella base naturale di \mathbb{R}^n . Perciò $F_{\mathcal{V}}$ ci offre una identificazione dello spazio vettoriale \mathbf{X} con \mathbb{R}^n , tale che ad \mathcal{V} corrisponda la base naturale di \mathbb{R}^n .

NOTAZIONE 4.1.1. La mappa $F_{\mathcal{V}}$ introdotto qui sopra è la mappa delle coordinate, che per ogni vettore restituisce le sue coordinate rispetto alla base scelta. Facciamo un esempio.

ESEMPIO 4.1.2. Siano $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ la base canonica

$$\mathcal{E} = \mathcal{E}_3 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

di \mathbb{R}^3 e $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ la base seguente: $\mathbf{v}_1 = \mathbf{e}_1 + \mathbf{e}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \mathbf{e}_1 - \mathbf{e}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$,

$\mathbf{v}_3 = \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Sia $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$: quindi x_i è la i -esima coordinata (o componente) di \mathbf{x} nella base canonica. Troviamo le componenti y_i dello stesso vettore \mathbf{x} nella base \mathcal{V} .

Svolgimento. Per prima cosa scriviamo la matrice M (nella base canonica!) della applicazione lineare U che manda la base canonica nella base \mathcal{V} (nell'ordine: $U(\mathbf{e}_i) = (v)_i$, $i = 1, 2, 3$). Questo è facile: la i -sima colonna di M_U è il vettore \mathbf{v}_i , cioè

$$M = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Un facile calcolo basato sull'eliminazione di Gauss, che lasciamo per esercizio, mostra che la matrice inversa M^{-1} , che indichiamo con $C = C_{\mathcal{V}, \mathcal{E}}$ è

$$C = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

(si arriva allo stesso risultato risolvendo per sostituzione il sistema lineare

$$A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

il che è facile perché la terza equazione del sistema è $x_3 = y_3$, e quindi in pratica il sistema lineare è bidimensionale invece che tridimensionale).

Abbiamo $\mathbf{x} = \sum_{i=1}^3 x_i \mathbf{e}_i$ e vogliamo trovare i numeri y_i tali che $\mathbf{x} = \sum_{j=1}^3 y_j \mathbf{v}_j$. D'altra parte, $U^{-1} \mathbf{v}_i = \mathbf{e}_i$, quindi

$$\mathbf{x} = \sum_{i=1}^3 x_i \mathbf{e}_i = \sum_{i=1}^3 x_i U^{-1} \mathbf{v}_i = \sum_{i=1}^3 x_i \sum_{j=1}^3 A_{ij} \mathbf{v}_j .$$

Scambiamo ora l'ordine delle due somme:

$$\sum_{j=1}^3 y_j \mathbf{v}_j = \sum_{j=1}^3 \sum_{i=1}^3 x_i A_{ij} \mathbf{v}_j .$$

Poiché i vettori \mathbf{v}_j , $j = 1, 2, 3$ sono una base i coefficienti dell'ultima uguaglianza devono essere uguali a sinistra e a destra: quindi per $j = 1, 2, 3$

$$y_j = \sum_{i=1}^3 A_{ij} .$$

L'ultima uguaglianza si può scrivere come prodotto righe per colonne

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = (x_1, x_2, x_3) A ,$$

dove però il vettore \mathbf{x} ora è il fattore di sinistra nel prodotto, ed è quindi scritto come vettore riga anziché colonna (altrimenti il prodotto sarebbe stato colonne per righe!).

In questo esempio, svolgendo il prodotto, si trova la seguente regola di trasformazione di coordinate:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}x_1 + \frac{1}{2}x_2 \\ \frac{1}{2}x_1 - \frac{1}{2}x_2 \\ x_3 \end{pmatrix}.$$

□

NOTA 4.1.3. Si osservi che, sebbene la matrice che trasforma i vettori \mathbf{e}_i negli \mathbf{v}_i sia M_U , nella trasformazione delle coordinate dei vettori sotto questo cambiamento di base interviene invece la matrice inversa $C = M_U^{-1}$. Si dice che le coordinate si trasformano in maniera controgradiente rispetto ai vettori di base. □

Riassumendo quanto esposto nell'Esempio 4.1.2, abbiamo mostrato:

PROPOSIZIONE 4.1.4. (Regola di trasformazione di coordinate dalla base canonica ad un'altra base.) *Se indichiamo con \mathbf{e}_i i vettori della base canonica e con \mathbf{v}_i ($i = 1, \dots, n$) quelli di un'altra base in \mathbb{R}^n , e con U la trasformazione lineare che verifica $U\mathbf{e}_i = \mathbf{v}_i$, allora la trasformazione delle coordinate dello stesso vettore $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i = \sum_{i=1}^n y_i \mathbf{v}_i$ sotto questo cambiamento di base è data da*

$$y_j = \sum_{i=1}^n x_i A_{ij}, \quad (4.1.1)$$

dove $A := M_U^{-1}$ è l'inversa della matrice M_U che esprime la trasformazione lineare U nella base canonica.

NOTAZIONE 4.1.5. (Matrice di passaggio dalla base \mathcal{F} alla base \mathcal{V} .) *A causa del suo ruolo nella regola di trasformazione di coordinate, la matrice $C = C_{\mathcal{V}, \mathcal{F}} := M_U^{-1}$ si chiama la matrice di passaggio (o del cambiamento di base) dalla base canonica*

$$\mathcal{E}_n = \left\{ \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix} \right\}$$

alla base \mathcal{V} . Essa è quindi la matrice che realizza la seguente trasformazione $F_{\mathcal{V}} \circ F_{\mathcal{E}}^{-1}$

$$\mathbb{R}^n \xrightarrow{F_{\mathcal{E}}^{-1}} (X) \xrightarrow{F_{\mathcal{V}}} \mathbb{R}^n$$

Rammentiamo che la colonna k -sima di tale matrice consiste dei coefficienti dello sviluppo rispetto alla base \mathcal{V} del k -simo vettore della base canonica (per maggiori dettagli si veda la successiva Nota 4.1.7).

Più in generale, nel passare da una base \mathcal{F} alla base \mathcal{V} la matrice $C = C_{\mathcal{V}, \mathcal{F}}$ che realizza la corrispondente trasformazione di coordinate $F_{\mathcal{V}} \circ \mathcal{F}^{-1}$

$$\mathbb{R}^n \xrightarrow{F_{\mathcal{F}}^{-1}} (X) \xrightarrow{F_{\mathcal{V}}} \mathbb{R}^n$$

è quella associata alla trasformazione lineare $F_{\mathcal{V}} \circ F_{\mathcal{F}}^{-1}$

NOTA 4.1.6. È immediato che $C_{\mathcal{V},\mathcal{F}} = C_{\mathcal{V},\mathcal{E}} C_{\mathcal{E},\mathcal{F}}$, perché $F_{\mathcal{V}} \circ \mathcal{F}^{-1} = F_{\mathcal{V}} \circ \mathcal{E}^{-1} \circ F_{\mathcal{E}} \circ \mathcal{F}^{-1}$. Quindi tutte le matrici di passaggio si esprimono in termini di matrici di passaggio dalla base canonica e loro inverse. \square

NOTA 4.1.7. (**Forma della matrice di cambiamento di base.**) Estendendo l'osservazione fatta per il passaggio dalla base canonica alla base \mathcal{V} , osserviamo ora che la colonna k -esima di $C_{\mathcal{F},\mathcal{E}}$ consiste dei coefficienti dello sviluppo di \mathbf{e}_k rispetto alla base \mathcal{F} : se

$$\mathbf{e}_k = \sum_{j=1}^n c_{jk} \mathbf{f}_j,$$

allora la colonna k -esima di $C_{\mathcal{F},\mathcal{E}}$ è

$$\begin{pmatrix} c_{1k} \\ c_{2k} \\ \vdots \\ c_{nk} \end{pmatrix},$$

quindi

$$C_{\mathcal{F},\mathcal{E}} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{pmatrix}. \quad (4.1.2)$$

Nello spazio vettoriale \mathbb{R}^n c'è un modo semplice per scrivere la matrice di passaggio da una base ad un'altra.

Cominciamo con la matrice di passaggio da una base $\mathcal{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ di \mathbb{R}^n alla base canonica

$$\mathcal{E}_n = \left\{ \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix} \right\}$$

e viceversa. Poiché i coefficienti dello sviluppo di \mathbf{f}_k rispetto ad \mathcal{E}_n sono le componenti del vettore \mathbf{f}_k , la matrice di passaggio da \mathcal{F} a \mathcal{E}_n è

$$C_{\mathcal{E}_n,\mathcal{F}} = \begin{pmatrix} | & & | \\ \mathbf{f}_1 & \dots & \mathbf{f}_n \\ | & & | \end{pmatrix}.$$

Ora possiamo calcolare anche la matrice di passaggio da \mathcal{E}_n a \mathcal{F} :

$$C_{\mathcal{F},\mathcal{E}_n} = C_{\mathcal{E}_n,\mathcal{F}}^{-1} = \begin{pmatrix} | & & | \\ \mathbf{f}_1 & \dots & \mathbf{f}_n \\ | & & | \end{pmatrix}^{-1}.$$

Se $\mathcal{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ e $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ sono due basi qualsiasi di \mathbb{R}^n , allora calcoliamo la matrice di passaggio da \mathcal{F} a \mathcal{V} quale il prodotto della matrice di passaggio da \mathcal{F} a \mathcal{E}_n con la matrice di passaggio da \mathcal{E}_n a \mathcal{V} :

$$C_{\mathcal{V}, \mathcal{F}} = C_{\mathcal{V}, \mathcal{E}_n} C_{\mathcal{E}_n, \mathcal{F}} = \begin{pmatrix} | & & | \\ \mathbf{v}_1 & \dots & \mathbf{v}_n \\ | & & | \end{pmatrix}^{-1} \begin{pmatrix} | & & | \\ \mathbf{v}_1 & \dots & \mathbf{v}_n \\ | & & | \end{pmatrix}.$$

□

4.2. Matrice di una applicazione lineare e cambiamento di basi

DEFINIZIONE 4.2.1. (*Matrice associata ad una trasformazione lineare in una data coppia di basi.*) Siano \mathbf{X}, \mathbf{Y} spazi vettoriali e $T : \mathbf{X} \longrightarrow \mathbf{Y}$ una applicazione lineare. Indichiamo con $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base di \mathbf{X} e con $\mathcal{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ una base di \mathbf{Y} . Identificando \mathbf{X} con \mathbb{R}^n ed \mathbf{Y} con \mathbb{R}^m tramite $F_{\mathcal{V}}$ ed $F_{\mathcal{F}}$, T diventa una applicazione lineare $\mathbb{R}^n \longrightarrow \mathbb{R}^m$. Più precisamente, a $T : \mathbf{X} \longrightarrow \mathbf{Y}$ corrisponde univocamente (una volta scelte le basi) $T_{\mathcal{F}, \mathcal{V}} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ definita da $T_{\mathcal{F}, \mathcal{V}} := F_{\mathcal{F}} \circ T \circ F_{\mathcal{V}}^{-1}$:

$$\begin{array}{ccc} \mathbf{X} & \xrightarrow{T} & \mathbf{Y} \\ F_{\mathcal{V}}^{-1} \uparrow & & \downarrow F_{\mathcal{F}} \\ \mathbb{R}^n & \xrightarrow{T_{\mathcal{F}, \mathcal{V}}} & \mathbb{R}^m \end{array}$$

La matrice $A_{\mathcal{F}, \mathcal{V}} = A_{\mathcal{F}, \mathcal{V}}(T) \in M_{mn}$ associata a $T_{\mathcal{F}, \mathcal{V}}$ si chiama la *matrice associata a T rispetto alle basi \mathcal{V} e \mathcal{F}* . Se $\mathbf{X} = \mathbf{Y}$ e $\mathcal{V} = \mathcal{F}$, allora questa matrice, associata alla trasformazione $F_{\mathcal{V}} \circ T \circ F_{\mathcal{V}}^{-1}$, viene chiamata la *matrice associata a T rispetto alla base \mathcal{V}* e la si indica semplicemente con $A_{\mathcal{V}}(T)$.

NOTA 4.2.2. La colonna k -sima di $A_{\mathcal{F}, \mathcal{V}}$ è

$$\begin{aligned} & T_{\mathcal{F}, \mathcal{V}}(\text{l'elemento } k\text{-simo della base naturale di } \mathbb{R}^n) \\ &= T_{\mathcal{F}, \mathcal{V}}(F_{\mathcal{V}}(\mathbf{v}_k)) = F_{\mathcal{F}}(T(\mathbf{v}_k)), \end{aligned}$$

cioè i coefficienti dello sviluppo di $T(\mathbf{v}_k)$ rispetto alla base \mathcal{F} di \mathbf{Y} considerati come un vettore in \mathbb{R}^m . Allora, per la definizione di $A_{\mathcal{F}, \mathcal{V}}$, l'immagine di

$$\mathbf{x} = \sum_{k=1}^n \lambda_k \mathbf{v}_k \in \mathbf{X}$$

tramite T è

$$T(\mathbf{x}) = \sum_{j=1}^m \mu_k \mathbf{f}_j \in \mathbf{Y},$$

ove

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix} = A_{\mathcal{F}, \mathcal{V}} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

□

ESEMPIO 4.2.3. Si calcoli la matrice dell'applicazione lineare T data da

$$\mathbb{R}^2 \ni \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$$

rispetto alla base

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

di \mathbb{R}^2 .

Svolgimento. Sia

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

la base canonica in \mathbb{R}^2 . Osserviamo che

$$\mathbf{v}_1 = \mathbf{e}_1 + \mathbf{e}_2 \quad \text{e} \quad \mathbf{v}_2 = \mathbf{e}_1 - \mathbf{e}_2, \quad (4.2.1)$$

da cui

$$\mathbf{e}_1 = \frac{1}{2}(\mathbf{v}_1 + \mathbf{v}_2) \quad (4.2.2)$$

$$\mathbf{e}_2 = \frac{1}{2}(\mathbf{v}_1 - \mathbf{v}_2) \quad (4.2.3)$$

(si noti che (4.2.2) è il calcolo dell'inversa della trasformazione di cambio di base (4.2.1). Pertanto,

$$T(\mathbf{e}_1) = \mathbf{e}_1 + \mathbf{e}_2 = \mathbf{v}_1 \quad (4.2.4)$$

$$T(\mathbf{e}_2) = \mathbf{e}_1 + 3\mathbf{e}_2 = \mathbf{v}_1 + 2\mathbf{e}_2 = \frac{3}{2}\mathbf{v}_1 - \mathbf{v}_2. \quad (4.2.5)$$

Quindi, in seguito alla Nota 4.2.2, nelle basi \mathcal{E} e \mathcal{V} la matrice di T , $A_{\mathcal{V}, \mathcal{E}}$, è

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

Ma la domanda che ci siamo posti è di trovare la matrice $A_{\mathcal{V}} = A_{\mathcal{V}, \mathcal{V}}$ di T nella base \mathcal{V} . A questo scopo dobbiamo esprimere i vettori \mathbf{e}_1 ed \mathbf{e}_2 nel primo membro di (4.2.4) in termini

della base $\mathbf{v}_1, \mathbf{v}_2$, facendo uso della trasformazione inversa del cambio di base calcolata in (4.2.2). Si ottiene:

$$\begin{aligned}\frac{1}{2}T(\mathbf{v}_1) + \frac{1}{2}T(\mathbf{v}_2) &= \mathbf{v}_1 \\ \frac{1}{2}T(\mathbf{v}_1) - \frac{1}{2}T(\mathbf{v}_2) &= \frac{3}{2}\mathbf{v}_1 - \mathbf{v}_2,\end{aligned}$$

da cui, risolvendo,

$$\begin{aligned}T(\mathbf{v}_1) &= \frac{5}{4}\mathbf{v}_1 - \frac{1}{2}\mathbf{v}_2 \\ T(\mathbf{v}_2) &= -\frac{1}{4}\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2.\end{aligned}$$

Pertanto

$$A_{\mathcal{V}} = \begin{pmatrix} \frac{5}{4} & -\frac{1}{4} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

□

ESEMPIO 4.2.4. Si consideri l'applicazione lineare T che manda i vettori della base canonica di \mathbb{R}^3 , $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, rispettivamente nei vettori $\mathbf{v}_1 := \mathbf{e}_1 + 2\mathbf{e}_2$, $\mathbf{v}_2 = \mathbf{e}_1$, $\mathbf{v}_3 = \mathbf{e}_3$. Si calcoli la matrice $A_{\mathcal{V}}$ di T nella base $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

Svolgimento. Esprimiamo l'immagine sotto T dei vettori \mathbf{v}_i in termini ancora dei vettori \mathbf{v}_i :

$$\begin{aligned}T(\mathbf{v}_1) &= T(\mathbf{e}_1) + 2T(\mathbf{e}_2) = 2\mathbf{e}_1 + 2\mathbf{e}_2 = \mathbf{v}_1 = \mathbf{v}_2 \\ T(\mathbf{v}_2) &= T(\mathbf{e}_1) = \mathbf{v}_1 \\ T(\mathbf{v}_3) &= T(\mathbf{e}_3) = \mathbf{v}_3.\end{aligned}$$

Si noti che nell'ultimo passaggio abbiamo espresso i vettori \mathbf{e}_i in termini dei \mathbf{v}_i , quindi anche in questo caso abbiamo invertito la matrice del cambiamento di base che manda gli \mathbf{e}_i nei \mathbf{v}_i , $i = 1, 2, 3$.) Pertanto la matrice che cerchiamo è

$$A_{\mathcal{V}} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

□

Il ruolo della matrice inversa del cambiamento di base in questi due esempi lascia intravedere la forma generale della soluzione del problema del cambiamento della rappresentazione matriciale di un operatore lineare al cambiare delle basi:

TEOREMA 4.2.5. Siano \mathbf{X}, \mathbf{Y} spazi vettoriali di dimensione rispettivamente n e m , e $T : \mathbf{X} \longrightarrow \mathbf{Y}$ una applicazione lineare. Indichiamo con $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ e $\mathcal{V}' = \{\mathbf{v}'_1, \dots, \mathbf{v}'_n\}$ due basi di \mathbf{X} e con $\mathcal{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ e $\mathcal{F}' = \{\mathbf{f}'_1, \dots, \mathbf{f}'_m\}$ due basi di \mathbf{Y} .

Siano $C_V = C_{\mathcal{V}', \mathcal{V}}$ la matrice (4.1.2) del cambiamento di base dalla base \mathcal{V} alla base \mathcal{V}' , e $C_F = C_{\mathcal{F}', \mathcal{F}}$ la matrice (4.1.2) del cambiamento di base dalla base \mathcal{F} alla base \mathcal{F}' . Allora, con la notazione stabilita nella Definizione 4.2.1, si ha

$$A_{\mathcal{F}', \mathcal{V}'} = C_V^{-1} A_{\mathcal{F}, \mathcal{V}} C_F .$$

Dimostrazione. Dalla Definizione 4.2.1 sappiamo che la matrice $A_{\mathcal{F}, \mathcal{V}}$ rappresenta (nelle basi canoniche di \mathbb{R}^n e \mathbb{R}^m) la trasformazione lineare $T_{\mathcal{F}, \mathcal{V}} := F_{\mathcal{F}} \circ T \circ F_{\mathcal{V}}^{-1}$. Perciò, per ogni vettore \mathbf{x} in \mathbf{X} ,

$$\begin{aligned} C_V^{-1} A_{\mathcal{F}, \mathcal{V}} C_F \mathbf{x} &= (F_{\mathcal{F}'} \circ F_{\mathcal{F}}^{-1})^{-1} \circ (F_{\mathcal{F}} \circ T \circ F_{\mathcal{V}}^{-1}) \circ (F_{\mathcal{V}} \circ F_{\mathcal{V}'}^{-1}) \mathbf{x} \\ &= F_{\mathcal{F}'} \circ F_{\mathcal{F}}^{-1} \circ F_{\mathcal{F}} \circ T \circ F_{\mathcal{V}}^{-1} \circ F_{\mathcal{V}} \circ F_{\mathcal{V}'}^{-1} \mathbf{x} \\ &= F_{\mathcal{F}'} \circ T \circ F_{\mathcal{V}'}^{-1} \mathbf{x} = A_{\mathcal{F}', \mathcal{V}'} \mathbf{x} . \end{aligned}$$

□

NOTA 4.2.6. Riassumendo,

$$A_{\mathcal{V}', \mathcal{F}'}(T) = C_{\mathcal{V}', \mathcal{V}} A_{\mathcal{V}, \mathcal{F}}(T) C_{\mathcal{F}, \mathcal{F}'} = C_{\mathcal{V}, \mathcal{V}'}^{-1} A_{\mathcal{V}, \mathcal{F}}(T) C_{\mathcal{F}, \mathcal{F}'} .$$

In particolare, nel caso di $A \in M_{mn}$, una base $\mathcal{F} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ di \mathbb{R}^n ed una base $\mathcal{V} = \{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ di \mathbb{R}^m , la matrice dell'applicazione lineare

$$T_A : \mathbb{R}^n \ni \mathbf{x} \longmapsto A \mathbf{x} \in \mathbb{R}^m$$

rispetto alle basi \mathcal{F} ed \mathcal{V} è uguale ad

$$\begin{aligned} A_{\mathcal{V}, \mathcal{F}}(T_A) &= C_{\mathcal{N}_m, \mathcal{V}}^{-1} A_{\mathcal{N}_m, \mathcal{N}_n}(T_A) C_{\mathcal{N}_n, \mathcal{F}} = C_{\mathcal{N}_m, \mathcal{V}}^{-1} A C_{\mathcal{N}_n, \mathcal{F}} \\ &= \begin{pmatrix} | & & | \\ \mathbf{f}_1 & \dots & \mathbf{f}_m \\ | & & | \end{pmatrix}^{-1} A \begin{pmatrix} | & & | \\ \mathbf{e}_1 & \dots & \mathbf{e}_n \\ | & & | \end{pmatrix} . \end{aligned}$$

Chiaramente,

$$\begin{aligned} T \text{ iniettiva} &\iff T_{\mathcal{F}, \mathcal{V}} \text{ iniettiva,} \\ T \text{ surgettiva} &\iff T_{\mathcal{F}, \mathcal{V}} \text{ surgettiva,} \end{aligned}$$

e quindi

$$T \text{ bigettiva} \iff T_{\mathcal{F}, \mathcal{V}} \text{ bigettiva} \iff A_{\mathcal{F}, \mathcal{V}} \text{ invertibile.}$$

La matrice della applicazione lineare identicamente zero è la matrice zero rispetto a qualsiasi basi. La matrice della applicazione identica $I_{\mathbf{X}}$ di uno spazio vettoriale \mathbf{X} rispetto ad una base di \mathbf{X} è chiaramente la matrice unità. Se però $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ e $\mathcal{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ sono due basi diverse di \mathbf{X} , allora la matrice $C_{\mathcal{F}, \mathcal{V}} := A_{\mathcal{F}, \mathcal{V}}(I_{\mathbf{X}})$ di $I_{\mathbf{X}}$ rispetto alle basi \mathcal{V} ed \mathcal{F} trasforma le coordinate di un vettore $\mathbf{x} \in \mathbf{X}$ rispetto ad \mathcal{V} nelle coordinate dello stesso \mathbf{x} rispetto a \mathcal{F} , e quindi è esattamente la *matrice di passaggio* dalla base \mathcal{V} alla base \mathcal{F} definita nella Notazione 4.1.5. □

ESEMPIO 4.2.7. Sia \mathcal{P}_2 lo spazio vettoriale di tutti i polinomi (con coefficienti reali) di grado ≤ 2 , cioè tutti i polinomi della forma

$$a_0 + a_1 x + a_2 x^2, \quad a_0, a_1, a_2 \in \mathbb{R}.$$

Chiamiamo *base canonica* di \mathcal{P}_2 la base $1, x, x^2$, e consideriamo l'applicazione lineare $T : \mathcal{P}_2 \longrightarrow \mathcal{P}_2$ che assegna ad ogni $P \in \mathcal{P}_2$ la derivata di $(x-1)P(x)$. Calcoliamo la matrice associata a T rispetto alla base naturale di \mathcal{P}_2 .

Svolgimento. Poiché

$$\begin{aligned} T(1) &= (x-1)' = 1 = 1 + 0x + 0x^2, \\ T(x) &= (x^2 - x)' = 2x - 1 = -1 + 2x + 0x^2, \\ T(x^2) &= (x^3 - x^2)' = 3x^2 - 2x = 0 - 2x + 3x^2, \end{aligned}$$

la matrice A associata a T rispetto alla base naturale di \mathcal{P}_2 ha le colonne

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

e perciò

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

Di conseguenza, la derivata del polinomio $(x-1)(a_0 + a_1 x + a_2 x^2)$ è uguale al polinomio $b_0 + b_1 x + b_2 x^2$, ove

$$\begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}.$$

La matrice A è invertibile, perciò T è bigettiva e perciò ad ogni $Q \in \mathcal{P}_2$ corrisponde un unico $P \in \mathcal{P}_2$ tale che Q sia la derivata di $(x-1)P(x)$. \square

I seguenti fatti sono ormai ovvi:

PROPOSIZIONE 4.2.8. Siano $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ spazi vettoriali, $T : \mathbf{X} \longrightarrow \mathbf{Y}$, $S : \mathbf{Y} \longrightarrow \mathbf{Z}$ applicazioni lineari, ed $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, $\mathcal{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_m\}$, $\mathcal{G} = \{\mathbf{g}_1, \dots, \mathbf{g}_p\}$ basi di $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$, rispettivamente. Allora la matrice associata alla composizione $S \circ T$ rispetto ad \mathcal{V} e \mathcal{G} è il prodotto $A_{\mathcal{G}, \mathcal{F}}(S) A_{\mathcal{F}, \mathcal{V}}(T) \in M_{np}$.

Se T è bigettiva, allora $A_{\mathcal{V}, \mathcal{F}}(T^{-1}) A_{\mathcal{F}, \mathcal{V}}(T) = A_{\mathcal{V}, \mathcal{V}}(T^{-1} \circ T) = A_{\mathcal{V}, \mathcal{V}}(I_{\mathbf{X}})$ è uguale alla matrice unità. Risulta che $A_{\mathcal{V}, \mathcal{F}}(T^{-1}) = A_{\mathcal{F}, \mathcal{V}}(T)^{-1}$.

In particolare, la matrice di passaggio da una base di \mathbf{X} ad un'altra base è l'inversa della matrice di passaggio dalla seconda base alla prima.

ESEMPIO 4.2.9. Consideriamo di nuovo lo spazio vettoriale \mathcal{P}_2 di tutti i polinomi di grado ≤ 2 . Troviamo le matrici di passaggio:

- dalla base naturale di \mathcal{P}_2 , cioè $1, x, x^2$, alla base $1, x-1, (x-1)^2$,
- dalla base $1, x-1, (x-1)^2$ alla base naturale di \mathcal{P}_2 .

Dopo di che, troviamo la matrice associata all'applicazione lineare $T : \mathcal{P}_2 \longrightarrow \mathcal{P}_2$, che assegna ad ogni $P \in \mathcal{P}_2$ la derivata di $(x-1)P(x)$ ed è stata considerata qui sopra, rispetto alla base $1, x-1, (x-1)^2$. Che connessione esiste fra la matrice trovata e quella calcolata nell'esempio precedente?

Svolgimento. Calcoliamo prima la matrice di passaggio dalla base $1, x-1, (x-1)^2$ alla base naturale di \mathcal{P}_2 : poiché

$$\begin{aligned} 1 &= 1 + 0x + 0x^2, \\ x-1 &= -1 + 1x + 0x^2, \\ (x-1)^2 &= 1 - 2x + 1x^2, \end{aligned}$$

essa è

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

Troviamo la matrice di passaggio nel senso inverso invertendo la matrice C , col metodo di eliminazione di Gauss:

$$\begin{array}{ccc|ccc|ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 & 1 & -1 & 0 & 1 & 0 & -1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & -2 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 2 & 0 & 1 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}$$

$$C^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix};$$

ma avremmo potuto trovare il risultato anche direttamente: poiché

$$\begin{aligned} 1 &= 1 + 0(x-1) + 0(x-1)^2, \\ x &= 1 + 1(x-1) + 0(x-1)^2, \\ x^2 &= 1 + 2(x-1) + 1(x-1)^2, \end{aligned}$$

la matrice di passaggio dalla base naturale di \mathcal{P}_2 alla base $1, x-1, (x-1)^2$ è quella con le colonne

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Ora possiamo trovare la matrice della trasformazione T , sia in maniera diretta analoga a quanto fatto nell'Esempio 4.2.7, sia applicando il Teorema 4.2.5. I due metodi sono equivalenti: illustriamo il secondo. Poiché

$$T(1) = (x-1)' = 1,$$

$$\begin{aligned} T(x-1) &= ((x-1)^2)' = 2(x-1), \\ T((x-1)^2) &= ((x-1)^3)' = 3(x-1)^2, \end{aligned}$$

la matrice di T rispetto alla base $1, x-1, (x-1)^2$ è la matrice diagonale

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

Abbiamo calcolato le matrici di passaggio C e C^{-1} , e grazie al Teorema 4.2.5 ne ricaviamo di nuovo la matrice di T rispetto alla base naturale di \mathcal{P}_2 , già calcolata nell'esempio precedente:

$$\begin{aligned} C \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} C^{-1} &= \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{pmatrix}. \end{aligned}$$

□

4.3. Cenni introduttivi sulla diagonalizzazione

Anticipiamo qui, come breve cenno, un argomento che sarà trattato in maggior dettaglio in seguito nel Capitolo 8. Siano \mathbf{X}, \mathbf{Y} spazi vettoriali della stessa dimensione n , $T : \mathbf{X} \rightarrow \mathbf{Y}$ una applicazione lineare, e $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, $\mathcal{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ basi di \mathbf{X} e \mathbf{Y} , rispettivamente. Allora la matrice associata a T rispetto alle basi \mathcal{V} e \mathcal{F} è *diagonale*, cioè ha tutti gli elementi fuori della diagonale uguali a zero, se e solo se

$$T(\mathbf{v}_k) \text{ è un multiplo scalare } \lambda_k \mathbf{f}_k \text{ di } \mathbf{f}_k, \quad 1 \leq k \leq n. \quad (4.3.1)$$

In tal caso si ha:

$$A_{\mathcal{F}, \mathcal{V}} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

Se $\mathbf{X} = \mathbf{Y}$ e $\mathcal{V} = \mathcal{F}$, allora la condizione (4.3.1) si scrive :

$$T(\mathbf{v}_k) = \lambda_k \mathbf{v}_k \text{ per un opportuno } \lambda_k \in \mathbb{R}, \quad 1 \leq k \leq n.$$

DEFINIZIONE 4.3.1. In generale, se T è una applicazione lineare di uno spazio vettoriale \mathbf{X} in se stesso, un scalare λ si chiama *autovalore* di T se esiste $\mathbf{0}_{\mathbf{X}} \neq \mathbf{x} \in \mathbf{X}$ tale che

$$T(\mathbf{x}) = \lambda \mathbf{x}. \quad (4.3.2)$$

Tutti i vettori *non nulli* $\mathbf{x} \in \mathbf{X}$, che soddisfano (4.3.2), si chiamano *autovettori* di T corrispondenti all'autovalore λ ed il sottospazio lineare

$$\text{Ker}(T - \lambda I_{\mathbf{X}}) = \{\mathbf{x} \in \mathbf{X}; T(\mathbf{x}) = \lambda \mathbf{x}\}$$

di \mathbf{X} si chiama *l'autospazio* di T corrispondente a λ .

Se A è una matrice quadrata d'ordine n , allora gli autovalori, autovettori ed autospazi di $T_A : \mathbb{R}^n \ni \mathbf{x} \mapsto A\mathbf{x} \in \mathbb{R}^n$ si chiamano rispettivamente autovalori, autovettori ed autospazi della matrice A .

4.4. Esercizi

ESERCIZIO 4.4.1. Sia A la matrice 3×3 seguente:

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

Con un calcolo diretto analogo a quello dell'Esempio 4.2.3, si calcoli come cambia la matrice dell'applicazione lineare indotta da A (nella base canonica), cioè

$$\mathbb{R}^3 \ni \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$$

quando si passa alla base

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

di \mathbb{R}^3 .

Soluzione. Poiché

$$\begin{aligned} T(\mathbf{v}_1) &= \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} = 3\mathbf{v}_1 + 0\mathbf{v}_2 + 0\mathbf{v}_3, \\ T(\mathbf{v}_2) &= \mathbf{0}_3 = 0\mathbf{v}_1 + 0\mathbf{v}_2 + 0\mathbf{v}_3, \\ T(\mathbf{v}_3) &= \mathbf{0}_3 = 0\mathbf{v}_1 + 0\mathbf{v}_2 + 0\mathbf{v}_3, \end{aligned}$$

la matrice desiderata è

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Osserviamo che la matrice così ottenuta è diagonale. Quindi il cambiamento di base ha diagonalizzato la rappresentazione matriciale della applicazione lineare data: cioè, la nuova base è una base di autovettori. Daremo maggiori dettagli nel Capitolo 8. \square

ESERCIZIO 4.4.2. Si calcoli la matrice di passaggio dalla base naturale di \mathbb{R}^3 alla base

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

Soluzione. La matrice richiesta è

$$B = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & -1 \\ 1 & 0 & 1 \end{pmatrix}^{-1}.$$

La calcoliamo:

$$\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & -1 & -1 & 0 & 1 & 0 & 0 & -2 & -1 & -1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & -1 & 1 & -1 & 0 & 1 \\ \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 & -1 & 0 & 1 & -1 & 1 & 0 & -1 \\ 0 & -2 & -1 & -1 & 1 & 0 & 0 & 0 & -3 & 1 & 1 & -2 \\ \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 & -1 & 0 & 1 & 0 & 2/3 & -1/3 & -1/3 \\ 0 & 0 & 1 & -1/3 & -1/3 & 2/3 & 0 & 0 & 1 & -1/3 & -1/3 & 2/3 \\ \\ & & & 1 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ & & & 0 & 1 & 0 & 2/3 & -1/3 & -1/3 \\ & & & 0 & 0 & 1 & -1/3 & -1/3 & 2/3 \end{array}.$$

Concludiamo che

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

□

ESERCIZIO 4.4.3. Si risolva l'Esercizio 4.4.1 non con il calcolo diretto, bensì applicando il Teorema di cambiamento di base 4.2.5.

Soluzione. Per prima cosa scriviamo la matrice $C = C_{\mathcal{V}, \mathcal{E}}$ di passaggio dalla base canonica \mathcal{E} alla base \mathcal{V} . Chiaramente C è la matrice che ha per colonne i vettori di \mathcal{V} :

$$C = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & -1 \\ 1 & 0 & 1 \end{pmatrix}.$$

L'inversa $B := C^{-1} = C_{\mathcal{E}, \mathcal{V}}$ è stata calcolata nel precedente Esercizio 4.4.2: essa vale

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

Ora, per il Teorema 4.2.5, la matrice $A_{\mathcal{V}}$ associata alla trasformazione lineare T nella base \mathcal{V} è

$$A_{\mathcal{V}} = B^{-1}AB,$$

dove, come già nell'Esercizio 4.4.2, abbiamo indicato con $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ la matrice $A_{\mathcal{E}} =$ che rappresenta l'operatore nella base canonica. Svolgendo le moltiplicazioni fra matrici si trova che

$$A_{\mathcal{V}} = B^{-1}AB = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (4.4.1)$$

Si osservi che è opportuno verificare l'identità (4.4.1), ma farlo è equivalente a verificare che $A = BA_{\mathcal{V}}B^{-1}$, una identità che richiede calcoli molto più semplici perché la matrice $A_{\mathcal{V}} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ ha tutti i coefficienti nulli tranne il primo. L'elementare verifica viene lasciata al lettore.

Osserviamo che *il cambiamento di basi, e la corrispondente operazione matriciale $A \longrightarrow B^{-1}AB$, ha diagonalizzato la matrice A di partenza.* Si veda il Capitolo 8 nel seguito.

□

ESERCIZIO 4.4.4. Si trovi la matrice quadrata A d'ordine 3 tale che l'applicazione lineare $\mathbb{R}^3 \ni \mathbf{x} \longmapsto A\mathbf{x} \in \mathbb{R}^3$ trasformi

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \mapsto 3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \mapsto \frac{1}{2} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Soluzione. Indichiamo

$$\mathbf{v}_1 := \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 := \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 := \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Allora $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ è una base di \mathbb{R}^3 e l'applicazione lineare T di cui sopra è stata definita tramite la sua matrice

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

rispetto a questa base. Poiché la matrice del passaggio dalla base $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ alla base canonica di \mathbb{R}^3 è

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix},$$

la matrice del passaggio dalla base canonica di \mathbb{R}^3 alla base $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ sarà C^{-1} . La calcoliamo:

$$\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & -1 & 0 & 1 & 0 & 0 & 1 & -1 & -1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & -1 & 0 & 1 \\ \hline 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 & 0 & 1 & 0 & -1 & 1/2 & 1/2 \\ 0 & 0 & 2 & 0 & -1 & 1 & 0 & 0 & 1 & 0 & -1/2 & 1/2 \end{array},$$

e perciò

$$C^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1/2 & 1/2 \\ 0 & -1/2 & 1/2 \end{pmatrix},$$

la matrice associata a T rispetto alla base canonica di \mathbb{R}^3 , cioè la matrice A per cui $T = T_A$, è

$$\begin{aligned} & C \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{pmatrix} C^{-1} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1/2 & 1/2 \\ 0 & -1/2 & 1/2 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 0 & 0 \\ 2 & 3/4 & 1/4 \\ 2 & 1/4 & 3/4 \end{pmatrix}. \end{aligned}$$

□

ESERCIZIO 4.4.5. Si scriva la matrice dell'applicazione lineare

$$\mathbb{R}^2 \ni \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} 2 & 1 \\ -1 & 3 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^3$$

rispetto alle basi

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ e } \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Soluzione.

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 2 & 1 \\ -1 & 3 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ -3/2 & -5/2 \\ -1/2 & 5/2 \end{pmatrix}.$$

□

CAPITOLO 5

Determinante di matrici

Consideriamo le matrici *quadrate*, cioè le matrici di ordine $n \times n$, per un certo $n \geq 1$. Ad ogni matrice quadrata si può associare un numero reale che in qualche modo ne *determina* alcune proprietà fondamentali.

DEFINIZIONE 5.0.6. Il ***determinante*** è una funzione che associa ad ogni matrice A di ordine $n \times n$ un numero reale, che indichiamo con $\det(A)$, che soddisfa le seguenti proprietà:

- $\det(I) = 1$;
- se A ha due righe uguali, allora $\det(A) = 0$;
- $\det(A)$ è una funzione lineare sulle righe di A .

L'ultima proprietà significa che se moltiplichiamo una riga di A per uno scalare λ , allora anche il determinante viene moltiplicato per λ . Allo stesso modo se sommiamo ad una riga di A un vettore, allora il determinante della matrice ottenuta è la somma di $\det(A)$ e del determinante della matrice con il vettore al posto della riga di A .

Dalle proprietà della definizione 5.0.6, si possono dimostrare altre proprietà della funzione determinante:

PROPOSIZIONE 5.0.7. *Consideriamo le matrici quadrate di ordine $n \times n$ e la funzione determinante su di esse. Allora:*

- se A ha una riga nulla, allora $\det(A) = 0$;
- scambiando due righe qualunque di A , allora $\det(A)$ cambia di segno;
- se le righe di A sono linearmente dipendenti, allora $\det(A) = 0$.

Si può dimostrare che esiste davvero, ed è unica, la funzione determinante che soddisfa la definizione 5.0.6 e le proprietà indicate.

Nel caso di matrici 2×2 , il determinante è facile da calcolare:

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12},$$

dove

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

ESEMPIO 5.0.8. Consideriamo la matrice A quadrata di ordine 2×2 :

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

allora il determinante di A è:

$$\det(A) = \begin{vmatrix} 2 & 1 \\ -1 & 0 \end{vmatrix} = 2 \cdot 0 - (-1) \cdot 1 = 1.$$

□

OSSERVAZIONE 5.0.9. Il determinante è definito solo per le matrici *quadrate*, cioè di ordine $n \times n$, per un certo intero $n \geq 1$. Se invece una matrice A non è quadrata, cioè è di ordine $n \times m$ con $n \neq m$, allora il determinante *non è definito*.

Il determinante di una matrice di ordine 3×3

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

si può calcolare con la formula di Sarrus:

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

ESEMPIO 5.0.10. Consideriamo la matrice

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

Allora il determinante di A è:

$$\det(A) = 2 + (-1) + 0 - 0 - (-2) - (-4) = 7.$$

□

In generale, se consideriamo una matrice quadrata di ordine $n \times n$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \quad (5.0.2)$$

allora si può calcolare il determinante sviluppandolo lungo una riga o una colonna con il metodo di Laplace. La dimostrazione è per induzione sull'ordine n delle matrici (questa dimostrazione si basa sull'assioma di induzione, assioma **P5** dei numeri interi, (1.3)). Il risultato è il seguente:

PROPOSIZIONE 5.0.11. *Sia A una matrice quadrata di ordine $n \times n$ come nella formula (5.0.2). Scegliendo di sviluppare la i -esima riga, il determinante di A è:*

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}),$$

dove A_{ij} è la matrice quadrata di ordine $(n-1) \times (n-1)$ ottenuta da A eliminando la i -esima riga e la j -esima colonna.

Nella proposizione 5.0.7 abbiamo visto che se le righe della matrice A sono linearmente dipendenti, allora il determinante di A è nullo. Vale anche il viceversa, come afferma la seguente:

PROPOSIZIONE 5.0.12. *Il determinante $\det(A)$ di una matrice A è zero se e solo se le righe di A sono linearmente indipendenti.*

Si può dimostrare anche che una matrice A è invertibile se e solo se il determinante di A è diverso da zero. In tal caso, l'inversa di A è

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} A_{11} & -A_{21} & \cdots & (-1)^{n-1}A_{n1} \\ -A_{12} & A_{22} & \cdots & (-1)^n A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ (-1)^{n-1}A_{1n} & (-1)^n A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

dove A_{ij} è il determinante della matrice di ordine $(n-1) \times (n-1)$ ottenuta da A eliminando la i -esima riga e la j -esima colonna.

Un'altra proprietà importante del determinante è data dal seguente:

TEOREMA 5.0.13 (Binet). *Siano A e B sono due matrici quadrate di ordine $n \times n$. Allora*

$$\det(A \cdot B) = \det(A) \cdot \det(B). \quad (5.0.3)$$

ESERCIZIO 5.0.14. Consideriamo due matrici quadrate A e B di ordine 2×2 o 3×3 . Verificate che vale la formula (5.0.3). \square

CENNO DI SOLUZIONE. Vediamo esplicitamente il caso delle matrici di ordine 2×2 . Lasciamo al lettore il caso di quelle di ordine 3×3 . Siano A e B matrici di ordine 2×2 :

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Allora il prodotto AB è la seguente matrice di ordine 2×2 :

$$AB = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix},$$

che ha determinante

$$\begin{aligned} \det(AB) &= (a_{11}b_{11} + a_{12}b_{21})(a_{21}b_{12} + a_{22}b_{22}) - (a_{21}b_{11} + a_{22}b_{21})(a_{11}b_{12} + a_{12}b_{22}) = \\ &= a_{12}b_{21}a_{21}b_{12} + a_{11}b_{11}a_{22}b_{22} - a_{21}b_{11}a_{12}b_{22} - a_{22}b_{21}a_{11}b_{12} = \\ &= (a_{11}a_{22} - a_{21}a_{12})(b_{11}b_{22} - b_{21}b_{12}) = \det(A) \cdot \det(B), \end{aligned}$$

che è proprio ciò che volevamo dimostrare. \square

CAPITOLO 6

Prodotto scalare e ortogonalità

6.1. Introduzione: prodotto scalare euclideo nel piano ed in \mathbb{R}^n

Questa sezione ha carattere introduttivo: tutte le definizioni e proprietà qui illustrate verranno riprese nel seguito in veste più generale e rigorosa.

DEFINIZIONE 6.1.1. Il *prodotto scalare naturale*, o *euclideo* di due vettori \mathbf{x} e \mathbf{y} in \mathbb{R}^2 è la lunghezza della proiezione ortogonale di \mathbf{y} su \mathbf{x} :

$$\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta$$

dove $\|\cdot\|$ denota la lunghezza di un vettore e θ è l'angolo formato dai vettori \mathbf{x} e \mathbf{y} .

Si noti che il prodotto scalare di \mathbf{x} e \mathbf{y} è un numero reale, cioè appunto uno scalare, il che spiega la terminologia.

Vorremmo trovare un metodo per calcolare il prodotto scalare in termini delle coordinate dei vettori, senza dover usare la trigonometria per ricavare il coseno dell'angolo da essi formato: indubbiamente tale metodo è vantaggioso, anche se a prima vista sembra assoggettare il risultato alla scelta di una base.

Consideriamo allora due vettori \mathbf{x} e \mathbf{y} nel piano \mathbb{R}^2 , diciamo $\mathbf{x} = (x_1, x_2)$ e $\mathbf{y} = (y_1, y_2)$, dove x_1, x_2 e y_1, y_2 sono le coordinate rispetto alla base canonica $e_1 = (1, 0)$ e $e_2 = (0, 1)$ di \mathbb{R}^2 .

Siano θ_1 e θ_2 gli angoli formati rispettivamente dai vettori \mathbf{x} e \mathbf{y} con la semiretta positiva delle ascisse. Allora l'angolo formato dai vettori \mathbf{x} e \mathbf{y} è

$$\theta = \theta_2 - \theta_1.$$

Ricordiamo le *formule di addizione e sottrazione* dei coseni:

$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$$

che ci serviranno fra poco e osserviamo che per definizione si ha che:

$$\cos \theta_1 = \frac{x_1}{\|\mathbf{x}\|}, \quad \sin \theta_1 = \frac{x_2}{\|\mathbf{x}\|}, \quad \cos \theta_2 = \frac{y_1}{\|\mathbf{y}\|}, \quad \sin \theta_2 = \frac{y_2}{\|\mathbf{y}\|}.$$

Dalle due equazioni precedenti segue che:

$$\cos \theta = \cos(\theta_2 - \theta_1) = \cos \theta_1 \cos \theta_2 + \sin \theta_1 \sin \theta_2 = \frac{x_1 y_1}{\|\mathbf{x}\| \|\mathbf{y}\|} + \frac{x_2 y_2}{\|\mathbf{x}\| \|\mathbf{y}\|} = \frac{x_1 y_1 + x_2 y_2}{\|\mathbf{x}\| \|\mathbf{y}\|}$$

Ma allora il prodotto scalare di \mathbf{x} e \mathbf{y} è

$$\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta = x_1 y_1 + x_2 y_2. \quad (6.1.1)$$

Scrivendo i vettori \mathbf{x} e \mathbf{y} come colonne:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

il prodotto scalare si può scrivere anche così:

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \mathbf{x}^t \mathbf{y}.$$

OSSERVAZIONE 6.1.2. Il prodotto scalare è commutativo: scambiando di posto i due vettori, il prodotto scalare non cambia:

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x},$$

come è evidente dalla formula (6.1.1) e dalla definizione intuitiva data all'inizio.

Un prodotto scalare naturale si può definire anche in \mathbb{R}^n analogamente al caso \mathbb{R}^2 .

DEFINIZIONE 6.1.3. Dati due vettori di \mathbb{R}^n

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

definiamo il *prodotto scalare euclideo* di \mathbf{x} e \mathbf{y} come

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^t \mathbf{y} = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

ESEMPIO 6.1.4. Consideriamo i seguenti due vettori in \mathbb{R}^3 :

$$\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Allora il loro prodotto scalare è

$$2 \cdot 1 + 1 \cdot 0 - 1 \cdot 1 = 1.$$

□

6.2. Spazi vettoriali su \mathbb{C}

La definizione di spazio vettoriale sul campo complesso \mathbb{C} differisce da quella degli spazi vettoriali su \mathbb{R} , Definizione 2.1.1, solo perché al posto di scalari reali si usano ora scalari complessi.

DEFINIZIONE 6.2.1. Dato uno spazio vettoriale reale $X_{\mathbb{R}}$, la sua *complessificazione* $X_{\mathbb{C}}$ è lo spazio vettoriale di tutte le combinazioni lineari *a coefficienti complessi* dei vettori di una qualsiasi base $\mathbf{x}_1, \dots, \mathbf{x}_n$ di $X_{\mathbb{R}}$. È chiaro che lo spazio $X_{\mathbb{C}}$ non dipende dalla scelta della base in $X_{\mathbb{R}}$ usata per costruirlo, e che quella base è anche una base in $X_{\mathbb{C}}$: quindi $X_{\mathbb{C}}$ ha la stessa dimensione su \mathbb{C} che $X_{\mathbb{R}}$ ha su \mathbb{R} (cautela: poiché il campo complesso \mathbb{C} è uno spazio vettoriale di dimensione due sul campo di scalari \mathbb{R} , ogni spazio vettoriale complesso $X_{\mathbb{C}}$ di dimensione n si può anche considerare come uno spazio vettoriale su $X_{\mathbb{R}}$, ed in tal caso la sua dimensione su \mathbb{R} è $2n$: una base è $\mathbf{x}_1, i\mathbf{x}_1, \mathbf{x}_2, i\mathbf{x}_2, \dots, \mathbf{x}_n, i\mathbf{x}_n$. Se uno spazio vettoriale $X_{\mathbb{C}}$ su \mathbb{C} è la complessificazione di uno spazio vettoriale $X_{\mathbb{R}}$ su \mathbb{R} , allora $X_{\mathbb{R}}$ si chiama una *forma reale* di \mathbb{C} .

6.3. * La definizione generale di prodotto scalare

Estendiamo la nozione di prodotto scalare con le seguenti definizioni.

DEFINIZIONE 6.3.1. (**Forme bilineari e forme hermitiane.**)

- (i) Dato uno spazio vettoriale \mathbf{X} su un campo K (in questo libro $K = \mathbb{R}$ o \mathbb{C} , i numeri reali o rispettivamente i numeri complessi), una applicazione $\phi : \mathbf{X} \times \mathbf{X} \longrightarrow K$ si dice una *forma bilineare* se è lineare rispetto ad entrambe le variabili, cioè se
 - (a₁) $\phi(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) = \phi(\mathbf{x}_1, \mathbf{y}) + \phi(\mathbf{x}_2, \mathbf{y})$
 - (a₂) $\phi(\mathbf{x}, \mathbf{y}_1 + \mathbf{y}_2) = \phi(\mathbf{x}, \mathbf{y}_1) + \phi(\mathbf{x}, \mathbf{y}_2)$
 - (b) $\phi(\lambda \mathbf{x}, \mathbf{y}) = \phi(\mathbf{x}, \lambda \mathbf{y}) = \lambda \phi(\mathbf{x}, \mathbf{y})$ per ogni scalare λ .
- (ii) Sia \mathbf{X} uno spazio vettoriale sul campo \mathbb{C} . Una applicazione $\phi : \mathbf{X} \times \mathbf{X} \longrightarrow \mathbb{C}$ si dice una *forma hermitiana* se è lineare rispetto alla prima variabile ed *antilineare* rispetto alla seconda, cioè se
 - (a₁) $\phi(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) = \phi(\mathbf{x}_1, \mathbf{y}) + \phi(\mathbf{x}_2, \mathbf{y})$
 - (a₂) $\phi(\mathbf{x}, \mathbf{y}_1 + \mathbf{y}_2) = \phi(\mathbf{x}, \mathbf{y}_1) + \phi(\mathbf{x}, \mathbf{y}_2)$
 - (b₁) $\phi(\lambda \mathbf{x}, \mathbf{y}) = \lambda \phi(\mathbf{x}, \mathbf{y})$ per ogni scalare λ
 - (b₂) $\phi(\mathbf{x}, \lambda \mathbf{y}) = \bar{\lambda} \phi(\mathbf{x}, \mathbf{y})$ per ogni scalare λ (qui, come sempre, $\bar{\lambda}$ denota il complesso coniugato di λ)

DEFINIZIONE 6.3.2. (**Prodotto scalare.**)

- (i) Un *prodotto scalare sul campo reale* su uno spazio vettoriale reale \mathbf{X} è una forma bilineare $\phi : \mathbf{X} \times \mathbf{X} \longrightarrow \mathbb{R}$ *simmetrica*, nel senso che

$$\phi(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{y}, \mathbf{x})$$

per ogni \mathbf{x}, \mathbf{y} in \mathbf{X} .

- (ii) Più in generale, un *prodotto scalare sul campo complesso* (detto anche *prodotto hermitiano*) su uno spazio vettoriale complesso \mathbf{X} è una forma bilineare $\phi : \mathbf{X} \times \mathbf{X} \longrightarrow \mathbb{C}$ *simmetrica coniugata*, nel senso che

$$\phi(\mathbf{x}, \mathbf{y}) = \overline{\phi(\mathbf{y}, \mathbf{x})}$$

per ogni \mathbf{x}, \mathbf{y} in \mathbf{X} .

ESEMPIO 6.3.3. Il prodotto scalare euclideo in \mathbb{R}^n , definito nella Sezione 6.1 mediante la formula

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n x_i y_i \quad (6.3.1)$$

dove i numeri x_i e y_i sono le coordinate dei vettori \mathbf{x} e \mathbf{y} rispetto alla base canonica (oppure rispetto ad una qualsiasi base prefissata) è un esempio di prodotto scalare reale. Si noti che questa definizione dipende dalla scelta di base.

L'esempio corrispondente di prodotto scalare complesso su \mathbb{C}^n è l'analogo prodotto scalare euclideo

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n x_i \overline{y_i}, \quad (6.3.2)$$

che dipende anch'esso dalla base scelta. \square

6.4. Matrici complesse, matrici autoaggiunte e matrici simmetriche

Una volta scelte le basi in due spazi vettoriali complessi di dimensioni rispettivamente n e m , ogni applicazione lineare fra i due spazi è rappresentata da una matrice $m \times n$ come nella sezione 3.5, ma questa volta le matrici sono a coefficienti complessi.

NOTAZIONE 6.4.1. *L'insieme della matrici $m \times n$ a coefficienti complessi è uno spazio vettoriale su \mathbb{C} che si indica con $M_m^{\mathbb{C}} n$. Per chiarezza, quando ci sia adito a dubbio, indicheremo lo spazio delle matrici reali con $M_m^{\mathbb{R}} n$.*

Rammentiamo dal Capitolo 3 la regola con cui si associa una matrice ad una applicazione lineare, riformulandola ora in termini di prodotti scalari:

PROPOSIZIONE 6.4.2. (**Matrice associata ad una applicazione lineare.**)

- (i) Denotiamo con $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ la base canonica in \mathbb{C}^n . Per ogni applicazione lineare $T : \mathbb{C}^n \longrightarrow \mathbb{C}^n$, la matrice $A = A_T$ associata a T nella base canonica ha per coefficienti

$$a_{ij} = T\mathbf{e}_i \cdot \mathbf{e}_j.$$

- (ii) Sia \mathbf{X} uno spazio vettoriale con base $\mathcal{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ e sia $T : \mathbf{X} \longrightarrow \mathbf{Y}$ una applicazione lineare. La matrice $A = A_{\mathcal{X}, \mathcal{X}}(T)$ associata a T nella base \mathcal{X} ha per coefficienti

$$a_{ij} = \langle T\mathbf{e}_i, \mathbf{f}_j \rangle.$$

Da queste espressioni segue immediatamente:

COROLLARIO 6.4.3. Sia $A \in M_{nn}^{\mathbb{R}}$. Allora, per ogni \mathbf{x} e $\mathbf{y} \in \mathbb{R}^n$,

$$A\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot A^T \mathbf{y}.$$

DEFINIZIONE 6.4.4. (**Agiunto.**) Sia $A \in M_{mn}^{\mathbb{C}}$. Si chiama *matrice aggiunta* (o semplicemente *aggiunto*) di A la matrice $A^* \in M_{nm}^{\mathbb{C}}$ data da

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix}^* = \begin{pmatrix} \overline{\alpha_{11}} & \overline{\alpha_{21}} & \dots & \overline{\alpha_{m1}} \\ \overline{\alpha_{12}} & \overline{\alpha_{22}} & \dots & \overline{\alpha_{m2}} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{\alpha_{1n}} & \overline{\alpha_{2n}} & \dots & \overline{\alpha_{mn}} \end{pmatrix}$$

(di nuovo, la barra indica il complesso coniugato). In altre parole, $A^* = \overline{A}^{\top} = \overline{A}^{\top}$.

NOTA 6.4.5. Per la matrice A^* valgono regole di calcolo analoghe a quelle della Nota 3.7.3; inoltre, se A è una matrice quadrata, fra aggiunto e prodotto scalare complesso c'è lo stesso legame che abbiamo già illustrato nel Corollario 6.4.3 fra trasposta e prodotto scalare reale:

$$\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A^*\mathbf{y} \rangle, \quad (6.4.1)$$

ovvero, con la terminologia alternativa introdotta in (6.3.2) per il prodotto scalare euclideo,

$$A\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot A^*\mathbf{y}. \quad (6.4.2)$$

Per linearità basta provare (??) per i vettori della base canonica: in tal caso, siccome $A\mathbf{e}_i \cdot \mathbf{e}_j = \alpha_{ij}$, (??) è precisamente la definizione di aggiunto 6.4.4. \square

DEFINIZIONE 6.4.6. (**Matrici simmetriche e matrici autoaggiunte.**) Una matrice $A \in M_{nn}^{\mathbb{R}}$ si dice *simmetrica* se $A = A^{\top}$. Più in generale, una matrice $A \in M_{nn}^{\mathbb{C}}$ si dice *autoaggiunta* se $A = A^*$. Analogamente, se una applicazione lineare verifica $\langle A\mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{x}, A\mathbf{x} \rangle$, A si dice *simmetrica* se agisce su uno spazio vettoriale reale, ed *autoaggiunta* se agisce su uno spazio vettoriale complesso.

6.5. * Norma e prodotti scalari definiti positivi

DEFINIZIONE 6.5.1. (**Norma.**) Una *norma* su uno spazio vettoriale \mathbf{X} è una funzione $N : \mathbf{X} \rightarrow \mathbb{R}$ con le seguenti proprietà:

- (i) $N(\mathbf{x}) \geq 0$ per ogni vettore \mathbf{x}
- (ii) $N(\mathbf{x}) = 0$ se e solo se $\mathbf{x} = \mathbf{0}$
- (iii) (**disuguaglianza triangolare.**) $N(\mathbf{x} + \mathbf{y}) \leq N(\mathbf{x}) + N(\mathbf{y})$ per tutti i vettori \mathbf{x} e \mathbf{y} .

NOTA 6.5.2. (**Norma euclidea.**) La *lunghezza* di un vettore $\mathbf{x} = (x_1, x_2, \dots, x_n)$ in \mathbb{R}^n , data da

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^n |x_i|^2}$$

è una norma, che si indica con *norma euclidea*; è da questo esempio di norma che proviene il nome della disuguaglianza triangolare (ogni lato di un triangolo ha lunghezza non superiore

alla somma degli altri due). Analogamente, la lunghezza di un vettore nello spazio vettoriale \mathbb{C}^n sul campo complesso è la norma definita esattamente come sopra ma interpretando il modulo nel senso del modulo dei numeri complessi. \square

DEFINIZIONE 6.5.3. (*Prodotto scalare definito positivo.*) Si dice che un prodotto scalare è *definito positivo* se $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ e $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ se e solo se $\mathbf{x} = \mathbf{0}$.

In virtù delle Definizioni 6.5.3 e 6.5.1 il seguente Corollario è evidente:

COROLLARIO 6.5.4. (Norma indotta da un prodotto scalare definito positivo.) *Se un prodotto scalare sullo spazio vettoriale \mathbf{X} è definito positivo, l'espressione $\langle \mathbf{x}, \mathbf{x} \rangle$ è una norma su X .*

ESEMPIO 6.5.5. Alla luce della Nota 6.5.2, la norma euclidea in \mathbb{R}^n (rispettivamente \mathbb{C}^n) è indotta dal rispettivo prodotto scalare euclideo. Quindi il prodotto scalare euclideo è definito positivo. \square

DEFINIZIONE 6.5.6. (*Prodotto scalare non degenero.*) Si dice che un prodotto scalare è *non degenero* se per ciascun vettore \mathbf{x} la condizione $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ per ogni \mathbf{y} implica $\mathbf{x} = \mathbf{0}$, cioè se non esistono vettori non nulli ortogonali a tutti i vettori.

Da queste definizioni risulta ovvio il Corollario seguente:

COROLLARIO 6.5.7. *Ogni prodotto scalare definito positivo è non degenero.*

6.6. Ortogonalità

LEMMAG 6.6.1. (Polarizzazione di forme bilineari.)

- (i) *Sia \mathbf{X} uno spazio vettoriale sul campo reale e $\phi : \mathbf{X} \times \mathbf{X} \longrightarrow \mathbb{R}$ una forma bilineare. Allora $\phi(\mathbf{x}, \mathbf{x}) = 0$ per ogni vettore \mathbf{x} se e solo se $\phi(\mathbf{x}, \mathbf{y}) = 0$ per ogni coppia di vettori \mathbf{x} e \mathbf{y} .*
- (i) *Più in generale, lo stesso risultato vale se \mathbf{X} è uno spazio vettoriale sui complessi e $\phi : \mathbf{X} \times \mathbf{X} \longrightarrow \mathbb{C}$ una forma hermitiana.*

Dimostrazione. L'implicazione inversa è ovvia. Per l'implicazione diretta, consideriamo dapprima il caso reale (parte (i)). Per ogni \mathbf{x}, \mathbf{y} si formi il vettore $\mathbf{v}_+ = \mathbf{x} + \mathbf{y}$. Si ha

$$0 = \phi(\mathbf{v}_+, \mathbf{v}_+) = \phi(\mathbf{x}, \mathbf{x}) + \phi(\mathbf{x}, \mathbf{y}) + \phi(\mathbf{y}, \mathbf{x}) + \phi(\mathbf{y}, \mathbf{y}) = 2\phi(\mathbf{x}, \mathbf{y})$$

per la proprietà di simmetria del prodotto scalare (Definizione 6.3.2 (i)). Questo prova (i).

La dimostrazione di (ii) (il caso complesso) è simile. Ora il prodotto scalare è quello complesso. Si considerino il vettore \mathbf{v}_1 di prima ed il vettore $\mathbf{v}_2 = \mathbf{x} + i\mathbf{y}$. L'argomento di prima, applicato di nuovo al vettore \mathbf{v}_1 , questa volta dà

$$0 = \phi(\mathbf{v}_1, \mathbf{v}_1) = \phi(\mathbf{x}, \mathbf{x}) + \phi(\mathbf{x}, i\mathbf{y}) + \phi(i\mathbf{y}, \mathbf{x}) + \phi(i\mathbf{y}, i\mathbf{y}) = 2\operatorname{Re}(\phi(\mathbf{x}, \mathbf{y})) ,$$

ed applicando lo stesso argomento al vettore \mathbf{v}_2 troviamo

$$0 = \phi(\mathbf{v}_2, \mathbf{v}_2) = \phi(\mathbf{x}, i\mathbf{y}) + \phi(i\mathbf{y}, \mathbf{x}) = 2 \operatorname{Re}(-i\phi(\mathbf{x}, \mathbf{y})) = 2 \operatorname{Im}(\phi(\mathbf{x}, \mathbf{y})) .$$

Concludiamo che $\phi(\mathbf{x}, \mathbf{y}) = 0$. Questo prova (ii).

DEFINIZIONE 6.6.2. (*Ortogonalità.*) Due vettori x e y di uno spazio vettoriale X si dicono **ortogonali**, o *perpendicolari*, se

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0.$$

DEFINIZIONE 6.6.3. (*Sottospazio ortogonale.*) Scelto un prodotto scalare in \mathbf{X} , per ogni sottoinsieme $\mathbf{E} \subset \mathbf{X}$ l'insieme

$$\mathbf{V}^\perp := \{\mathbf{x} : \langle \mathbf{x}, \mathbf{v} \rangle = 0 \quad \text{per ogni } \mathbf{v} \in \mathbf{E}\}$$

si chiama *ortogonale* di \mathbf{E} .

PROPOSIZIONE 6.6.4. (*Complemento ortogonale e proiezione ortogonale.*) In ogni spazio vettoriale \mathbf{X} munito di prodotto scalare si ha:

- (i) L'ortogonale \mathbf{E}^\perp di ogni sottoinsieme (non solo sottospazio!) $\mathbf{E} \subset \mathbf{X}$ è un sottospazio vettoriale di \mathbf{X} , che d'ora in avanti chiameremo sottospazio ortogonale di \mathbf{E} .
- (ii) $\mathbf{0}^\perp = \mathbf{X}$.
- (iii) Se il prodotto scalare è non degenere, allora $\mathbf{X}^\perp = \mathbf{0}$
- (iv) Il prodotto scalare è non degenere se e solo se per ogni vettore \mathbf{v} il sottospazio unidimensionale \mathbf{V} che esso genera è strettamente contenuto in \mathbf{X} , cioè $\dim \mathbf{V} < \dim \mathbf{X}$.
- (v) Se il prodotto scalare è definito positivo, allora per ogni sottospazio \mathbf{V} si ha $\mathbf{V} \cap \mathbf{V}^\perp = \emptyset$ ed ogni $\mathbf{x} \in \mathbf{X}$ si decompone in modo unico come $\mathbf{x} = \mathbf{v} + \mathbf{w}$ con $\mathbf{v} \in \mathbf{V}$ e $\mathbf{w} \in \mathbf{V}^\perp$. Il vettore \mathbf{v} si chiama la proiezione ortogonale di \mathbf{x} su \mathbf{V} . In tal caso per ogni sottospazio vettoriale $\mathbf{V} \subset \mathbf{X}$ si ha $\dim \mathbf{V} + \dim \mathbf{V}^\perp = \dim \mathbf{X}$.

Dimostrazione. Le parti (i), (ii) e (iii) sono ovvie: ne lasciamo la dimostrazione al lettore per esercizio. Per provare (iv), basta ricordare (Definizione 6.5.6) che il prodotto scalare è non degenere se e solo se nessun vettore è ortogonale a tutto lo spazio. Per provare (v) osserviamo che, se $\mathbf{x} \in \mathbf{V} \cap \mathbf{V}^\perp$, allora $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ e quindi $\mathbf{x} = \mathbf{0}$ perché il prodotto scalare è definito positivo. Rimane solo da dimostrare l'enunciato sulle dimensioni. Per ogni $\mathbf{x} \in \mathbf{X}$ scriviamo $\mathbf{x} = \mathbf{v} + \mathbf{w}$ come sopra e definiamo l'operatore di proiezione su \mathbf{V} come $P\mathbf{x} = \mathbf{v}$. È evidente che P è una applicazione lineare, $P : \mathbf{X} \rightarrow \mathbf{X}$, l'immagine di P è il sottospazio \mathbf{V} e $\ker P = \mathbf{V}^\perp$. Ora l'identità $\dim \mathbf{V} + \dim \mathbf{V}^\perp = \dim \mathbf{X}$ non è altro che il teorema della dimensione, Teorema 3.2.1.

NOTA 6.6.5. La dimostrazione della parte (v) della Proposizione 6.6.4, cioè dell'esistenza della proiezione ortogonale, qui è formulata senza far ricorso alla nozione di continuità della norma. Il nostro approccio vale pertanto in dimensione finita, ma l'enunciato è più generale, e vale per una classe di spazi vettoriali a dimensione infinita muniti di prodotto

scalare (gli *spazi di Hilbert*). Per la dimostrazione nel caso generale rinviando il lettore ad un [libro su questo argomento](#). \square

NOTA 6.6.6. La proprietà della Proposizione ?? non vale se il prodotto scalare è soltanto non degenere. Infatti, sotto tali ipotesi, possono esistere vettori di norma zero, cioè ortogonali a sé stessi. Ad esempio, il prodotto scalare dell'Esempio 6.9.2 (iii), diciamo associato alla matrice diagonale

$$D = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 0 & \\ & & & \ddots \\ & & & & 0 \end{pmatrix}$$

si espande come

$$\mathbf{x} \cdot \mathbf{y} = x_1 \overline{y_1} - x_2 \overline{y_2}$$

e quindi il vettore $(1, -1, 0, \dots, 0)$ ha norma nulla, cioè è ortogonale a sé stesso, ma il prodotto scalare è non degenere come osservato nell'Esempio 6.9.2 (iii) e 6.9.2 (iv). \square

LEMMAG 6.6.7. *Se i vettori $\mathbf{x}_1, \dots, \mathbf{x}_k$ sono a due a due ortogonali rispetto ad un prodotto scalare non degenere, allora essi sono linearmente indipendenti.*

Dimostrazione. Siano λ_i scalari tali che $\sum_{i=1}^k \lambda_i \mathbf{x}_i = \mathbf{0}$: dobbiamo mostrare che $\lambda_j = 0$ per ogni j con $1 \leq j \leq k$. Ma per ogni j si ha

$$0 = \left\langle \sum_{i=1}^k \lambda_i \mathbf{x}_i, \mathbf{x}_j \right\rangle = \lambda_j \langle \mathbf{x}_j, \mathbf{x}_j \rangle.$$

Poiché il prodotto scalare è non degenere, $\langle \mathbf{x}_j, \mathbf{x}_j \rangle \neq 0$ per ogni j : ne segue che $\lambda_j = 0$.

NOTAZIONE 6.6.8. *Dato un prodotto scalare, un insieme di vettori si dice un sistema ortogonale rispetto a quel prodotto scalare se essi sono a due a due ortogonali. Se inoltre tutti i vettori sono di norma 1 il sistema si chiama un empsistema ortonormale.*

TEOREMA 6.6.9. **(Esistenza di basi ortogonali ed ortonormali.)**

- (i) *Sia \mathbf{X} uno spazio vettoriale con un prodotto scalare: allora \mathbf{X} ha una base ortogonale (cautela: non è detto che esista una base ortonormale, perché i vettori di base potrebbero avere lunghezza zero rispetto a questo prodotto scalare, e quindi non essere normalizzabili).*
- (ii) *Se il prodotto scalare è definito positivo, allora \mathbf{X} ha una base ortonormale.*

Dimostrazione.

- (i) Procediamo per induzione su $n = \dim \mathbf{X}$ (la validità di questo metodo dimostrativo segue dall'assioma di induzione, assioma **P5** dei numeri interi, (1.3)). Se $n = 1$ le basi hanno un solo elemento: basta allora scegliere un qualsiasi vettore non nullo.

Supponiamo $n > 1$. Ci sono due possibilità:

- o $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ per ogni \mathbf{x} , oppure
- esiste un vettore \mathbf{x}_1 tale che $\langle \mathbf{x}_1, \mathbf{x}_1 \rangle \neq 0$.

Nel primo caso, il Lemma di polarizzazione 6.6.1 ci dice che il prodotto scalare è identicamente zero, cioè $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ per ogni \mathbf{x}, \mathbf{y} . In questo caso banale ogni base è ortogonale (non ortonormale, perché tutti i vettori hanno norma zero!), e l'enunciato vale.

Nel secondo caso, indichiamo con \mathbf{X}_1 il sottospazio generato da \mathbf{x}_1 . Sappiamo che la norma $\langle \mathbf{x}_1, \mathbf{x}_1 \rangle$ di \mathbf{x}_1 è non nulla, e questo equivale a dire che \mathbf{X}_1 non è contenuto nel suo ortogonale \mathbf{X}_1^\perp . Quindi, anche se ora, a differenza della Proposizione ?? (v), non sappiamo se $\dim \mathbf{X}_1 + \dim \mathbf{X}_1^\perp = n$, certamente sappiamo che $\dim \mathbf{X}_1^\perp < n$.

Ogni vettore $\mathbf{x} \in \mathbf{X}$ è la somma di un vettore in \mathbf{X}_1 ed un altro vettore nel sottospazio ortogonale \mathbf{X}_1^\perp : non abbiamo l'unicità della decomposizione come nella Proposizione ?? (v), perché ora non stiamo supponendo che il prodotto scalare sia definito positivo, ma l'esistenza sì, perché

$$\mathbf{x} = \left(\mathbf{x} - \frac{\langle \mathbf{x}, \mathbf{x}_1 \rangle}{\langle \mathbf{x}_1, \mathbf{x}_1 \rangle} \mathbf{x}_1 \right) + \frac{\langle \mathbf{x}, \mathbf{x}_1 \rangle}{\langle \mathbf{x}_1, \mathbf{x}_1 \rangle} \mathbf{x}_1. \quad (6.6.1)$$

Poiché $\dim \mathbf{X}_1^\perp < n$, per ipotesi di induzione possiamo assumere che esista una base ortogonale $\{\mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n\}$ in \mathbf{X}_1^\perp . Allora tutti questi vettori sono ortogonali a \mathbf{x}_1 , e quindi $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n\}$ è una base ortogonale in \mathbf{X} . Questo prova (i).

- (ii) Anche in questa parte della dimostrazione, se la dimensione $n = 1$ il risultato è ovvio: si prende un vettore non nullo e si osserva che, poiché il prodotto scalare è definito positivo, la norma di quel vettore è positiva, quindi normalizzandolo si ottiene una base ortonormale (consistente in un solo vettore, in questo caso).

Se invece $n > 1$ consideriamo un sottospazio \mathbf{X}_1 di dimensione 1, generato da un vettore \mathbf{x}_1 che come sopra possiamo assumere di norma 1. Il sottospazio ortogonale \mathbf{X}_1^\perp ha dimensione $n - 1$ (Proposizione ?? (v)), e quindi, per ipotesi di induzione, ha una base ortonormale che indichiamo con $\{\mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n\}$. Allora, analogamente alla parte (i) della dimostrazione, $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n\}$ è una base ortonormale in \mathbf{X} .

6.7. Procedimento di ortogonalizzazione di Gram-Schmidt

In questa sezione ci proponiamo di fornire una dimostrazione costruttiva dell'esistenza di basi ortonormali in spazi vettoriali muniti di un prodotto scalare definito positivo (Teorema 6.6.9 (ii)), basandoci sulla proiezione ortogonale (Proposizione ?? (v)) ed applicando iterativamente l'identità (6.6.1).

Consideriamo una base $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ di uno spazio vettoriale \mathbf{X} . Scegliamo in \mathbf{X} un prodotto scalare definito positivo, in modo che il prodotto scalare di ogni vettore non nullo con sé stesso sia positivo (una norma). Vogliamo costruire esplicitamente una base *ortonormale* $\{\mathbf{x}'_1, \dots, \mathbf{x}'_n\}$. Il metodo che stiamo per sviluppare si chiama il **procedimento di ortogonalizzazione di Gram-Schmidt**.

Partiamo da $x'_1 = x_1$.

Definiamo x'_2 nel modo seguente:

$$x'_2 = x_2 - \frac{x_2 \cdot x'_1}{x'_1 \cdot x'_1} x'_1.$$

Poi definiamo x'_3 :

$$x'_3 = x_3 - \frac{x_3 \cdot x'_2}{x'_2 \cdot x'_2} x'_2 - \frac{x_3 \cdot x'_1}{x'_1 \cdot x'_1} x'_1.$$

e andiamo avanti così, cioè:

$$x'_i = x_i - \frac{x_i \cdot x'_{i-1}}{x'_{i-1} \cdot x'_{i-1}} x'_{i-1} - \dots - \frac{x_i \cdot x'_1}{x'_1 \cdot x'_1} x'_1.$$

PROPOSIZIONE 6.7.1. *L'insieme $\{x'_1, \dots, x'_n\}$ così costruito è una base ortogonale di X .*

DIMOSTRAZIONE. Si verifica direttamente con i calcoli che $x'_i \cdot x'_j = 0$. □

Si noti che, per costruzione, il sottospazio vettoriale generato da x'_1, \dots, x'_i , per ogni i , coincide con il sottospazio generato da x_1, \dots, x_i .

ESEMPIO 6.7.2. Consideriamo i seguenti tre vettori in \mathbb{R}^4 :

$$x_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Vogliamo trovare una base ortogonale del sottospazio di \mathbb{R}^4 generato da x_1, x_2, x_3 con il procedimento di ortogonalizzazione di Gram-Schmidt.

Innanzitutto poniamo

$$x'_1 = x_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}.$$

Dopo si definisce:

$$x'_2 = x_2 - \frac{x_2 \cdot x'_1}{x'_1 \cdot x'_1} x'_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Infine si definisce:

$$x'_3 = x_3 - \frac{x_3 \cdot x'_2}{x'_2 \cdot x'_2} x'_2 - \frac{x_3 \cdot x'_1}{x'_1 \cdot x'_1} x'_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{4} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}$$

Abbiamo trovato così la seguente base ortogonale:

$$x'_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \quad x'_2 = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad x'_3 = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}.$$

□

6.8. Matrici ortogonali e matrici unitarie

LEMMA 6.8.1. (Preservazione di norme e di angoli.) *Siano \mathbf{x} e \mathbf{y} due vettori in uno spazio vettoriale complesso $X_{\mathbb{C}}$, ed A una applicazione lineare di $X_{\mathbb{C}}$ in sé. Allora $\langle A\mathbf{x}, A\mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle$ per ogni vettore \mathbf{x} se e solo se $\langle A\mathbf{x}, A\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$ per ogni coppia di vettori \mathbf{x}, \mathbf{y} .*

Dimostrazione. Costruiamo la forma hermitiana

$$\phi(\mathbf{x}, \mathbf{y}) = \langle A\mathbf{x}, A\mathbf{y} \rangle - \langle \mathbf{x}, \mathbf{y} \rangle.$$

L'enunciato segue immediatamente applicando a questa forma hermitiana il Lemma 6.6.1.

DEFINIZIONE 6.8.2. (Matrici ortogonali e matrici unitarie.) Una matrice quadrata $A \in M_n^{\mathbb{R}}$ si dice *ortogonale* se preserva la norma dei vettori (che abbiamo anche chiamato lunghezza), cioè se $\|A\mathbf{x}\| = \langle A\mathbf{x}, A\mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2$ per ogni vettore \mathbf{x} . Analogamente, una matrice quadrata $A \in M_n^{\mathbb{C}}$ si dice *unitaria* se preserva la norma in \mathbb{C}^n .

COROLLARIO 6.8.3. *Il prodotto righe per colonne di matrici unitarie (rispettivamente, ortogonali) è una matrice unitaria (rispettivamente, ortogonale).*

Dimostrazione. Ogni matrice definisce una applicazione lineare, ed il prodotto righe per colonne corrisponde alla composizione delle applicazioni (Definizione 3.6.3). Se due applicazioni preservano la norma, anche il prodotto la preserva.

NOTA 6.8.4. Se A è una matrice ortogonale o unitaria, allora, grazie al lemma 6.6.1, per ogni coppia di vettori \mathbf{x}, \mathbf{y} in \mathbb{R}^n o \mathbb{C}^n rispettivamente si ha $A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$. In particolare, ricordando che se due vettori in \mathbb{R}^n formano un angolo θ allora $\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta$, ed osservando che $\|A\mathbf{x}\| = \|\mathbf{x}\|$ per la Definizione 6.8.2, concludiamo che l'azione di una matrice ortogonale A su \mathbb{R}^n preserva gli angoli fra i vettori. □

ESEMPIO 6.8.5. (**Matrici di rotazione.**) La Nota 6.8.4 mostra che le matrici ortogonali corrispondono agli operatori di rotazione su \mathbb{R}^n : sono quelle che ruotano una base ortonormale in un'altra base ortonormale, quindi le loro colonne sono vettori ortonormali. Ad esempio, in \mathbb{R}^2 , le loro colonne sono i vettori del tipo $(\cos \theta, \sin \theta)$, se chiamiamo θ l'angolo di rotazione (si veda più in generale il successivo Teorema 6.8.7). In particolare, la più generale matrice reale ortogonale a dimensione 2 è

$$\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

che manda il vettore $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ in $\begin{pmatrix} \cos \theta \\ -\sin \theta \end{pmatrix}$, e quindi è una rotazione di un angolo θ in senso antiorario. \square

COROLLARIO 6.8.6. (i) Una matrice $A \in M_{nn}^{\mathbb{R}}$ è ortogonale se e solo se $A^{-1} = A^{\top}$
(ii) Una matrice $A \in M_{nn}^{\mathbb{C}}$ è unitaria se e solo se $A^{-1} = A^*$

Dimostrazione. Basta provare la proprietà (ii), perché essa implica la (i).

Per la precedente Nota 6.8.4, A è unitaria se e solo se, per ogni \mathbf{x}, \mathbf{y} in \mathbb{C}^n , si ha $\mathbf{x} \cdot A^* A \mathbf{y} = A \mathbf{x} \cdot A \mathbf{y} = \mathbf{x} \cdot \mathbf{y}$. Quindi $\mathbf{x} \cdot (A^* A - \mathbb{I}) \mathbf{y} = 0$. Poiché il prodotto scalare è non degenere, questo equivale a dire che $A^* A = \mathbb{I}$, cioè $A^* = A^{-1}$. Si noti che, a sua volta, quest'ultima asserzione equivale a dire che $AA^* = \mathbb{I}$.)

TEOREMA 6.8.7. Le seguenti condizioni sono equivalenti:

- (i) una matrice $A \in M_{nn}^{\mathbb{C}}$ è unitaria;
- (ii) le colonne di A sono vettori ortonormali rispetto al prodotto scalare euclideo in \mathbb{C}^n ,
- (iii) le righe di A sono vettori ortonormali rispetto al prodotto scalare euclideo in \mathbb{C}^n ,
- (iv) A^* è unitaria,
- (v) A manda una base ortonormale in una base ortonormale (rispetto al prodotto scalare euclideo).

In particolare, una matrice reale è ortogonale se e solo se le sue colonne sono ortonormali rispetto al prodotto scalare in \mathbb{R}^n , e se e solo se lo sono le sue righe, e se e solo se lo è A^{\top} , e se e solo se manda una base ortonormale di \mathbb{R}^n in una base ortonormale di \mathbb{R}^n .

Dimostrazione. Sia $A \in M_{nn}^{\mathbb{C}}$ unitaria. Indichiamo come sempre con $\{\mathbf{e}_i\}$ i vettori della base canonica in \mathbb{C}^n . Utilizziamo la seguente notazione abituale: il simbolo δ_{ij} , detto *simbolo di Kronecker*, vale 1 se $i = j$ e 0 altrimenti. Segue dalla Nota 6.8.4 che

$$A \mathbf{e}_i \cdot A \mathbf{e}_j = \mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}. \quad (6.8.1)$$

Pertanto le colonne di A , cioè i vettori $A \mathbf{e}_i$, sono una famiglia ortonormale, e quindi (i) implica (ii). Viceversa, se le colonne sono ortonormali, cioè se vale (6.8.1), allora, scrivendo $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$, otteniamo per linearità che $A \mathbf{x} \cdot A \mathbf{x} = \sum_{i,j=1}^n x_i \overline{x_j} \delta_{ij} = \sum_{i=1}^n |x_i|^2 = \|\mathbf{x}\|^2$, e quindi A è unitaria. Pertanto (ii) implica (i).

Ora, se A è unitaria, grazie a (ii) essa è invertibile, quindi lo è anche A^* , perché $\det A^* = \det \overline{A^\top} = \overline{\det A^\top} = \overline{\det A}$. Perciò ogni vettore $\mathbf{y} \in \mathbb{C}^n$ si può scrivere (in modo unico) come $\mathbf{y} = A^* \mathbf{x}$ per qualche vettore \mathbf{x} . Pertanto, per definizione di aggiunto (Definizione 6.4.4),

$$A^* \mathbf{x} \cdot A^* \mathbf{x} = \mathbf{y} \cdot \mathbf{y} = A\mathbf{y} \cdot A\mathbf{y} = A^{*-1} \mathbf{y} \cdot A^{*-1} \mathbf{y} = \mathbf{x} \cdot \mathbf{x}$$

quindi A^* è unitaria e (i) implica (iv). Poiché $A^{**} = A$, anche (iv) implica (i). Allora, in base a (ii), il fatto che A^* sia unitaria equivale ad asserire che le sue colonne formino una famiglia ortonormale di vettori: ma queste colonne sono il complesso coniugato delle righe di A (Definizione 6.4.4), e quindi le righe di A sono una famiglia ortonormale, e (iv) equivale a (iii). Infine, il fatto che le colonne di A siano un sistema ortonormale chiaramente equivale a dire che A manda la base canonica in una base ortonormale. Pertanto la matrice che manda una base ortonormale nella base canonica è l'inversa A^{-1} di una matrice unitaria: ma per una matrice unitaria si ha $A^{-1} = A^*$, e A^* è ancora unitaria grazie alla parte (iv) del teorema. Da questo segue che una matrice che manda una base ortonormale \mathcal{F} in una base ortonormale \mathcal{V} è il prodotto di due matrici unitarie (quella che manda \mathcal{F} nella base canonica e quella che manda la base canonica in \mathcal{V}). Per il Corollario 6.8.3 questo prodotto è ancora una matrice unitaria. Questo prova la parte (v).

6.9. * Matrice associata ad un prodotto scalare

L'Esempio 6.3.3 sottolinea che, quando il prodotto scalare è espresso in termini di coordinate, esso evidentemente dipende dalla scelta della base. Chiariamo ora questa dipendenza.

PROPOSIZIONE 6.9.1. (Matrice associata ad una forma bilineare o hermitiana.) *Esiste una corrispondenza biunivoca fra le forme bilineari o hermitiane ϕ su \mathbb{R}^n (rispettivamente \mathbb{C}^n) e le matrici M $n \times n$, che verifica la regola seguente: la matrice M dà origine alla forma ϕ_M definita da*

$$\phi_M(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot M\mathbf{y}$$

ovvero, facendo uso del prodotto righe per colonne,

$$\phi_M(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top M\mathbf{y}. \quad (6.9.1)$$

In altre parole, in \mathbb{R}^n

$$\phi_M(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^n m_{ij} x_i y_j$$

(dove i numeri x_i e y_i sono le coordinate nella base canonica) e più in generale in \mathbb{C}^n

$$\phi_M(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^n m_{ij} x_i \overline{y_j}.$$

Inoltre:

- ϕ_M è un prodotto scalare reale se e solo se M è simmetrica (nel senso della Definizione 6.4.6, $M = M^\top$);

- ϕ_M è un prodotto scalare complesso (ovvero prodotto hermitiano) se e solo se M è autoaggiunta (nel senso della Definizione 6.4.6, $M = M^* := \overline{M}^T$);
- se ϕ_M è un prodotto scalare, allora esso è non degenere se e solo se $\det M \neq 0$;
- se ϕ_M è un prodotto scalare, allora esso è definito positivo se e solo se M è definita positiva, nel senso che $\sum_{i,j=1}^n m_{ij}x_i\overline{x_j} \geq 0$ per ogni n -pla di numeri complessi $\{x_1, \dots, x_n\}$.

Dimostrazione. Data una matrice M a dimensione n , è chiaro che ϕ_M è una forma bilineare se si usa il prodotto scalare reale su \mathbb{R}^n , o hermitiana si usa il prodotto scalare complesso su \mathbb{C}^n , nel senso dell'Esempio 6.3.3. Viceversa, data una forma bilineare (rispettivamente hermitiana) ϕ , consideriamo la matrice M i cui coefficienti sono

$$m_{ij} = \phi(\mathbf{e}_i, \mathbf{e}_j)$$

dove i vettori \mathbf{e}_i sono i vettori della base canonica. Grazie alla bilinearità si ha

$$\phi(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^n m_{ij}x_i\overline{y_j}$$

dove i numeri x_i e y_i sono le coordinate nella base canonica. Ne segue che M è la matrice associata alla forma ϕ , cioè che $\phi = \phi_M$, e quindi la corrispondenza è biunivoca.

Delle restanti proprietà, le prime due sono conseguenze ovvie delle formule della Proposizione 6.4.2. Proviamo la terza per un prodotto scalare reale. La forma bilineare ϕ_M dà luogo ad un prodotto scalare degenere se e solo se esiste un vettore $\mathbf{y} \neq \mathbf{0}$ tale che $\phi_M(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot M\mathbf{y} = 0$ per ogni \mathbf{x} . In particolare, $M\mathbf{y} \cdot M\mathbf{y} = 0$. Ma il prodotto scalare euclideo è definito positivo (Nota 6.5.5), quindi $M\mathbf{y} = \mathbf{0}$ e M non è iniettiva. Viceversa, se $\ker M \neq \mathbf{0}$, sia \mathbf{y} un vettore non nullo in $\ker M$: allora $\phi_M(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot M\mathbf{y} = 0$ per ogni \mathbf{x} ed il prodotto scalare ϕ_M è degenere. Infine, dire che ϕ_M è un prodotto scalare definito positivo equivale a dire che $\sum_{i,j=1}^n m_{ij}x_i\overline{y_j} = \phi_M(\mathbf{x}, \mathbf{x}) \geq 0$ per ogni \mathbf{x} .

ESEMPIO 6.9.2. (*Prodotti scalari associati a matrici diagonali.*)

- (i) Il prodotto scalare euclideo è associato alla matrice identità: $\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot \mathbb{I}\mathbf{y}$.
- (ii) Un prodotto scalare ϕ_D associato ad una matrice diagonale D è definito positivo se e solo se tutti i termini diagonali $\{d_i, \quad i = 1, \dots, n\}$ della matrice sono positivi. Infatti un tale prodotto scalare si espande come $\phi_D(\mathbf{x}, \mathbf{x}) = \sum_{i=1}^n d_i|x_i|^2$. Perciò, se i d_i sono tutti positivi, allora $\phi_D(\mathbf{x}, \mathbf{x}) = \sum_{i=1}^n d_i|x_i|^2 \geq 0$, invece se esiste un termine negativo, diciamo d_1 , allora $\phi_D(\mathbf{e}_1, \mathbf{e}_1) = d_1 < 0$.
- (iii) Un prodotto scalare ϕ_D associato ad una matrice diagonale D che possiede due termini diagonali di segno opposto, diciamo $d_1 < 0 < d_2$ non è quindi definito positivo, ed in particolare l'espressione $N(\mathbf{x}) = \phi_D(\mathbf{x}, \mathbf{x})$ non è una norma perché esistono vettori non nulli con $N(\mathbf{x}) = 0$. Ad esempio, consideriamo i vettori $\mathbf{x} = (x_1, x_2, 0, 0, \dots, 0)$ con tutte le componenti nulle dopo la seconda: per questi vettori si ha $\phi_D(\mathbf{x}, \mathbf{x}) = d_1|x_1|^2 - d_2|x_2|^2$, ed il secondo membro si annulla per opportuni

x_1 e x_2 non nulli. Nonostante questo, il prodotto \mathbf{x} tale che $\phi_D(\mathbf{x}, \mathbf{y}) = 0$ per ogni vettore \mathbf{y} , potremmo prendere al posto di \mathbf{y} lo i -simo scalare e non degenerare a meno che qualcuno dei termini diagonali sia nullo. Infatti, se esistesse un vettore non nullo vettore della base canonica, \mathbf{e}_i , ed otterremmo $d_i x_i = \phi_D(\mathbf{x}, \mathbf{e}_i) = 0$, da cui $x_i = 0$ visto che $d_i \neq 0$. Ma allora tutte le componenti x_i di \mathbf{x} sono nulle, il che contraddice l'ipotesi $\mathbf{x} \neq \mathbf{0}$.

- (iv) Se invece la matrice diagonale D ha un termine diagonale nullo, diciamo $d_1 = 0$, allora ϕ_D è degenerare, perché $\phi_D(\mathbf{e}_1, \mathbf{y}) = d_1 |y_1|^2 = 0$ per ogni vettore \mathbf{y} .

□

NOTAZIONE 6.9.3. (Prodotti scalari in forma diagonale.) Quando un prodotto scalare è associato ad una matrice diagonale, come nel precedente Esempio 6.9.2, diciamo che esso è espresso in forma diagonale, o anche senza termini misti. Se il prodotto scalare non è espresso in forma diagonale ma lo diventa in una base diversa, allora diciamo che esso è riducibile a forma diagonale tramite un cambiamento di base.

Ora finalmente possiamo discutere cosa succede quando si fissa un prodotto scalare, ad esempio il prodotto scalare euclideo nella base canonica, e poi si cambia base.

PROPOSIZIONE 6.9.4. (Prodotti scalari e cambio di base.)

- (i) Sia M la matrice associata ad un prodotto scalare in una data base \mathcal{F} nel senso della Proposizione 6.9.1, e sia \mathcal{F}' una nuova base. Sia $C = C_{\mathcal{F}, \mathcal{F}'}$ la matrice di passaggio dalla base \mathcal{F}' alla base \mathcal{F} , come nella Notazione 4.1.5. Allora la matrice associata al prodotto scalare nella base \mathcal{F}' è $C^\top M C$.
- (ii) Se le basi \mathcal{F} e \mathcal{F}' sono basi ortonormali, allora la matrice associata al prodotto scalare nella nuova base \mathcal{F}' è la matrice $C^{-1} M C$.
- (iii) La base canonica è ortogonale per un prodotto scalare se e solo se la matrice associata al prodotto scalare nella base canonica è diagonale; analogamente, una base è ortogonale per un prodotto scalare se e solo se la matrice ad esso associata in quella base è una matrice diagonale.
- (iv) Ogni prodotto scalare è riducibile a forma diagonale tramite un cambiamento di base. Esiste quindi una matrice invertibile C tale che $C^\top M C$ è una matrice diagonale.
- (v) Se il prodotto scalare è definito positivo, la matrice di cambiamento di base che realizza la diagonalizzazione è unitaria nel caso complesso, e ortogonale nel caso reale (come definite nella Definizione 6.8.2). Data la matrice M associata al prodotto scalare, esiste quindi una matrice unitaria (od ortogonale) C tale che $C^\top M C = C^{-1} M C$ è una matrice diagonale.

Dimostrazione. Abbiamo visto (6.9.1) che il legame fra il prodotto scalare la matrice M è il seguente: $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top M \mathbf{y}$. Per la bilinearità (proprietà (i) – (a₁) e (i) – (a₂) della Definizione

6.3.1, e analoghe nel caso hermitiano), il prodotto scalare è univocamente identificato dai coefficienti di matrice $m_{ij} = \mathbf{f}_j^\top M \mathbf{f}_i$ (nella base \mathcal{F}). D'altra parte, abbiamo osservato nella Nota **4.1.3** che, se $A = C_{\mathcal{F}', \mathcal{F}}$, allora $A^{-1} \mathbf{f}_i = \mathbf{f}'_i$ per ogni $i = 1, \dots, n$; poiché $A^{-1} = C_{\mathcal{F}, \mathcal{F}'} := C$ questo equivale a dire $C^{-1} \mathbf{f}_i = \mathbf{f}'_i$, cioè $C \mathbf{f}'_i = \mathbf{f}_i$. Allora nella base \mathcal{F}' il prodotto scalare è associato alla matrice $m'_{ij} := \mathbf{f}'_j^\top M \mathbf{f}'_i$, cioè alla matrice $M' = C^\top M C$. Questo prova (i).

Per provare (ii), osserviamo che, se le due basi sono ortonormali, allora la matrice C di cambiamento di base è una matrice unitaria (o, nel caso reale, ortogonale) per il Teorema **6.8.7**. Pertanto $C^\top = C^{-1}$, e la parte (ii) segue dalla parte (i).

La base canonica è ortogonale rispetto al prodotto scalare associato in essa ad una matrice M se e solo se $m_{ij} = \langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0$ per $i \neq j$, cioè se e solo se la matrice M è diagonale. Identico ragionamento vale in qualunque altra base. Questo prova (iii).

Proviamo ora la parte (iv). Sia M la matrice associata al prodotto scalare in una data base. Sappiamo che M è simmetrica nel caso reale, e più in generale autoaggiunta nel caso complesso. In entrambi i casi, per la parte (iii) di questo teorema, M è diagonale se e solo se la base è ortogonale rispetto al prodotto scalare. D'altro canto, grazie al Teorema **6.6.9**, una base ortogonale esiste. Quindi, in questa base, M si riduce a forma diagonale, ed il prodotto scalare diventa espresso in forma diagonale.

In maggior dettaglio, supponiamo dapprima che lo spazio vettoriale sia \mathbb{R}^n e che le matrici siano reali. Sia C la matrice che implementa il cambiamento di base: se indichiamo con \mathbf{e}_i i vettori della base canonica e con \mathbf{f}_i quelli della base ortogonale, abbiamo $C \mathbf{e}_i = \mathbf{f}_i$. Consideriamo un nuovo prodotto scalare

$$\langle \mathbf{x}, \mathbf{y} \rangle' := \langle C \mathbf{x}, C \mathbf{y} \rangle = (C \mathbf{x})^\top M C \mathbf{y} = \mathbf{x}^\top C^\top M C \mathbf{x}.$$

Questo nuovo prodotto scalare è quello ottenuto passando alla nuova base, e cioè quello trovato nella parte (i) della dimostrazione, la cui matrice associata è $C^\top M C$. Osserviamo che la base canonica è ortogonale rispetto a questo nuovo prodotto scalare:

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle' = \langle \mathbf{f}_i, \mathbf{f}_j \rangle = 0 \quad \text{se } i \neq j$$

perché la base dei vettori \mathbf{f}_i è ortogonale rispetto al prodotto scalare originale. Quindi la matrice associata al nuovo prodotto scalare, cioè $C^\top M C$, è diagonale grazie alla parte (iii). Nel caso dello spazio complesso \mathbb{C}^n l'argomento è analogo. Questo prova (iv).

La parte (v) equivale a dimostrare che per ogni matrice A autoaggiunta (rispettivamente, simmetrica) esiste una matrice U unitaria (rispettivamente, ortogonale) tale che $U^\top A U = U^{-1} A U$ è diagonale. Questo risultato verrà dimostrato in seguito nel Corollario **8.5.4**.

CAPITOLO 7

Funzionali lineari e dualità

CAPITOLO 8

Autovalori, autovettori e diagonalizzabilità di matrici

In questo capitolo analizziamo in maggiore profondità la nozione di *diagonalizzazione* accennata nella Sezione 4.3 ed elaborata nell'Esercizio 4.4.3.

8.1. Triangolarizzazione e diagonalizzazione

Siano \mathbf{X}, \mathbf{Y} due spazi vettoriali e $T : \mathbf{X} \longrightarrow \mathbf{Y}$ una applicazione lineare. Siano anche $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ e $\mathcal{V}' = \{\mathbf{v}'_1, \dots, \mathbf{v}'_n\}$ due basi di \mathbf{X} , mentre $\mathcal{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ e $\mathcal{F}' = \{\mathbf{f}'_1, \dots, \mathbf{f}'_m\}$ sono due basi di \mathbf{Y} . Il Teorema 4.2.5 esprime la matrice di T rispetto ad \mathcal{V}' ed \mathcal{F}' in termini della matrice di T rispetto ad \mathcal{V} ed \mathcal{F} e le matrici di cambiamenti di base nel modo seguente:

$$A_{\mathcal{F}', \mathcal{V}'}(T) = C_{\mathcal{F}', \mathcal{F}} A_{\mathcal{F}, \mathcal{V}}(T) C_{\mathcal{V}, \mathcal{V}'} = C_{\mathcal{F}, \mathcal{F}'}^{-1} A_{\mathcal{F}, \mathcal{V}}(T) C_{\mathcal{V}, \mathcal{V}'}.$$

In particolare, nel caso di $A \in M_{mn}$, una base $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ di \mathbb{R}^n ed una base $\mathcal{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ di \mathbb{R}^m , la matrice dell'applicazione lineare

$$T_A : \mathbb{R}^n \ni \mathbf{x} \longmapsto A \mathbf{x} \in \mathbb{R}^m$$

rispetto alle basi \mathcal{V} ed \mathcal{F} è uguale ad

$$\begin{aligned} A_{\mathcal{F}, \mathcal{V}}(T_A) &= C_{\mathcal{N}_m, \mathcal{F}}^{-1} A_{\mathcal{N}_m, \mathcal{N}_n}(T_A) C_{\mathcal{N}_n, \mathcal{V}} = C_{\mathcal{N}_m, \mathcal{F}}^{-1} A C_{\mathcal{N}_n, \mathcal{V}} \\ &= \begin{pmatrix} | & & | \\ \mathbf{f}_1 & \dots & \mathbf{f}_m \\ | & & | \end{pmatrix}^{-1} A \begin{pmatrix} | & & | \\ \mathbf{v}_1 & \dots & \mathbf{v}_n \\ | & & | \end{pmatrix}. \end{aligned}$$

Pertanto, se possiamo trovare delle basi \mathcal{V} ed \mathcal{F} tale che $A_{\mathcal{F}, \mathcal{V}}(T_A)$ abbia una forma particolare, per esempio

- **triangolare** (*superiore*), cioè con tutti gli elementi sotto la diagonale uguali a zero,
- o
- **diagonale**, cioè con tutti gli elementi fuori della diagonale uguali a zero,

allora A si esprime tramite una matrice di questo tipo particolare mediante la formula

$$A = \begin{pmatrix} | & & | \\ \mathbf{f}_1 & \dots & \mathbf{f}_m \\ | & & | \end{pmatrix} A_{\mathcal{F}, \mathcal{V}}(T_A) \begin{pmatrix} | & & | \\ \mathbf{v}_1 & \dots & \mathbf{v}_n \\ | & & | \end{pmatrix}^{-1}. \quad (8.1.1)$$

Particolarmente interessante è il caso in cui $n = m$ e $\mathcal{V} = \mathcal{F}$. Allora la matrice in (8.1.1) è quadrata A d'ordine n :

$$A = \begin{pmatrix} | & & | \\ \mathbf{v}_1 & \dots & \mathbf{v}_n \\ | & & | \end{pmatrix} A_{\mathcal{V},\mathcal{V}}(T_A) \begin{pmatrix} | & & | \\ \mathbf{v}_1 & \dots & \mathbf{v}_n \\ | & & | \end{pmatrix}^{-1}. \quad (8.1.2)$$

Se $A_{\mathcal{V},\mathcal{V}}(T_A)$ è triangolare, diciamo che A è stata *triangularizzata*, mentre se $A_{\mathcal{V},\mathcal{V}}(T_A)$ è diagonale, diciamo che A è stata *diagonalizzata* (la stessa terminologia è stata introdotta nella Sezione 4.3). Nel seguito rivolgiamo l'attenzione al problema di trovare come diagonalizzare una matrice quadrata. Un esempio è stato già presentato nell'Esercizio 4.4.3.

8.2. Autovalori, autovettori e diagonalizzazione

Siano \mathbf{X}, \mathbf{Y} spazi vettoriali della stessa dimensione n , $T : \mathbf{X} \rightarrow \mathbf{Y}$ una applicazione lineare, ed $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, $\mathcal{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ basi di \mathbf{X} e \mathbf{Y} , rispettivamente. Allora la matrice associata a T rispetto ad \mathcal{V} e \mathcal{F} è diagonale se e solo se

$$T(\mathbf{v}_k) \text{ è un multiplo scalare } \lambda_k \mathbf{f}_k \text{ di } \mathbf{f}_k, \quad 1 \leq k \leq n, \quad (8.2.1)$$

ed in tal caso abbiamo:

$$A_{\mathcal{F},\mathcal{V}} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

Se $\mathbf{X} = \mathbf{Y}$ e $\mathcal{V} = \mathcal{F}$, allora la condizione (8.2.1) si scrive :

$$T(\mathbf{v}_k) = \lambda_k \mathbf{v}_k \text{ per un opportuno } \lambda_k \in \mathbb{R}, \quad 1 \leq k \leq n.$$

DEFINIZIONE 8.2.1. (*Autovalori, autovettori ed autospazi.*) Se T è una applicazione lineare di uno spazio vettoriale \mathbf{X} in se stesso, un scalare λ si chiama *autovalore* di T se esiste $\mathbf{0}_{\mathbf{X}} \neq \mathbf{x} \in \mathbf{X}$ tale che

$$T(\mathbf{x}) = \lambda \mathbf{x}. \quad (8.2.2)$$

Tutti i vettori non nulli $\mathbf{x} \in \mathbf{X}$ che soddisfano (8.2.2) si chiamano *autovettori* di T corrispondenti all'autovalore λ ed il sottospazio lineare

$$\text{Ker}(T - \lambda I_{\mathbf{X}}) = \{\mathbf{x} \in \mathbf{X}; T(\mathbf{x}) = \lambda \mathbf{x}\}$$

di \mathbf{X} si chiama l'*autospazio* di T corrispondente a λ . La dimensione di un'autospazio $\text{Ker}(T - \lambda I_{\mathbf{X}})$ si chiama la *molteplicità geometrica* di λ .

NOTA 8.2.2. Osserviamo che il multiplo di un autovettore per un scalare non nullo è ancora un autovettore corrispondente allo stesso autovalore:

$$T(\mathbf{x}) = \lambda \mathbf{x} \implies T(\alpha \mathbf{x}) = \lambda (\alpha \mathbf{x}).$$

□

Con queste definizioni possiamo dire che la matrice di una applicazione lineare $T : \mathbf{X} \longrightarrow \mathbf{X}$ rispetto ad una base $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ di \mathbf{X} è diagonale se e solo se tutti i vettori \mathbf{v}_k della base sono autovettori di T .

Il prossimo risultato dimostra la proprietà cruciale che autovettori corrispondenti ad autovalori diversi sono linearmente indipendenti:

TEOREMA 8.2.3. (Indipendenza lineare degli autovettori.) *Siano \mathbf{X} uno spazio vettoriale e $T : \mathbf{X} \longrightarrow \mathbf{X}$ una applicazione lineare. Se*

$\lambda_1, \dots, \lambda_k$ sono autovalori diversi di T e

\mathbf{x}_j è un autovettore di T corrispondente a λ_j , $1 \leq j \leq k$,

allora i vettori $\mathbf{x}_1, \dots, \mathbf{x}_k$ sono linearmente indipendenti.

Dimostrazione. (Facoltativa.)

Dobbiamo verificare l'implicazione

$$\left. \begin{array}{l} \lambda_1, \dots, \lambda_k \text{ autovalori diversi di } T \\ \mathbf{0}_{\mathbf{X}} \neq \mathbf{x}_j \in \text{Ker}(T - \lambda_j I_{\mathbf{X}}), 1 \leq j \leq k \\ \sum_{j=1}^k \alpha_j \mathbf{x}_j = \mathbf{0}_{\mathbf{X}} \end{array} \right\} \implies \alpha_1 = \dots = \alpha_k = 0. \quad (I_k)$$

L'implicazione (I_1) è chiara.

Supponiamo adesso che l'implicazione (I_k) non sia sempre vera e sia $k \geq 2$ il più piccolo numero naturale per il quale essa non vale. Allora esistono

autovalori diversi $\lambda_1, \dots, \lambda_k$ di T ,

autovettori $\mathbf{x}_1 \in \text{Ker}(T - \lambda_1 I_{\mathbf{X}}), \dots, \mathbf{x}_k \in \text{Ker}(T - \lambda_k I_{\mathbf{X}})$,

scalari $\alpha_1, \dots, \alpha_k$ non tutti nulli

tali che

$$\sum_{j=1}^k \alpha_j \mathbf{x}_j = \mathbf{0}_{\mathbf{X}} \quad \text{e quindi anche} \quad \sum_{j=1}^k \alpha_j \lambda_j \mathbf{x}_j = T\left(\sum_{j=1}^k \alpha_j \mathbf{x}_j\right) = \mathbf{0}_{\mathbf{X}}.$$

Allora

$$\sum_{j=1}^{k-1} \alpha_j (\lambda_k - \lambda_j) \mathbf{x}_j = \lambda_k \sum_{j=1}^k \alpha_j \mathbf{x}_j - \sum_{j=1}^k \alpha_j \lambda_j \mathbf{x}_j = \mathbf{0}_{\mathbf{X}}.$$

Per il modo in cui è stato scelto k , i $k-1$ autovettori $\mathbf{x}_1, \dots, \mathbf{x}_{k-1}$ sono linearmente indipendenti: perciò dobbiamo avere $\alpha_j \underbrace{(\lambda_k - \lambda_j)}_{\neq 0} = 0$, e così $\alpha_j = 0$ per ogni $1 \leq j \leq k-1$.

Ma allora l'uguaglianza $\sum_{j=1}^k \alpha_j \mathbf{x}_j = \mathbf{0}_{\mathbf{X}}$ diventa $\alpha_k \mathbf{x}_k = \mathbf{0}_{\mathbf{X}}$ e pertanto anche $\alpha_k = 0$, in contraddizione con l'ipotesi che non tutti i scalari $\alpha_1, \dots, \alpha_k$ siano nulli. \square

Grazie a questo Teorema 8.2.3, se $\lambda_1, \dots, \lambda_k$ sono autovalori diversi di una applicazione lineare $T : \mathbf{X} \longrightarrow \mathbf{X}$ e, per ogni $1 \leq j \leq k$, $\mathbf{v}_1^{(j)}, \dots, \mathbf{v}_{d_j}^{(j)}$ è una base dell'autospazio $\text{Ker}(T - \lambda_j I_{\mathbf{X}})$, allora

$$\underbrace{\mathbf{v}_1^{(1)}, \dots, \mathbf{v}_{d_1}^{(1)}}_{d_1}, \dots, \underbrace{\mathbf{v}_1^{(k)}, \dots, \mathbf{v}_{d_k}^{(k)}}_{d_k}$$

è un sistema linearmente indipendente e pertanto è una base del sottospazio lineare di \mathbf{X} generato da tutti gli autovettori di T .

Infatti, se $\sum_{p=1}^{d_1} \alpha_p^{(1)} \mathbf{v}_p^{(1)} + \dots + \sum_{p=1}^{d_k} \alpha_p^{(k)} \mathbf{v}_p^{(k)} = \mathbf{0}_{\mathbf{X}}$, allora per il teorema sull'indipendenza lineare degli autovettori $\mathbf{x}_1 = \dots = \mathbf{x}_k = \mathbf{0}_{\mathbf{X}}$. Siccome, per ogni $1 \leq j \leq k$, $\mathbf{v}_1^{(j)}, \dots, \mathbf{v}_{d_j}^{(j)}$ sono linearmente indipendenti, risulta che $\alpha_1^{(j)} = \dots = \alpha_{d_j}^{(j)} = 0$.

Abbiamo così dimostrato il risultato seguente:

PROPOSIZIONE 8.2.4. *Esiste una base di \mathbf{X} consistente di soli autovettori di una applicazione lineare $T : \mathbf{X} \longrightarrow \mathbf{X}$, ossia una base rispetto a quale la matrice di T è diagonale, se e solo se la somma delle molteplicità geometriche di tutti gli autovalori di T è uguale alla dimensione di \mathbf{X} . In particolare, se T ha tanti autovalori diversi quanta è la dimensione di \mathbf{X} , allora otteniamo una tale base scegliendo per ogni autovalore un corrispondente autovettore.*

Se A è una matrice quadrata d'ordine n , allora gli autovalori, autovettori ed autospazi di $T_A : \mathbb{R}^n \ni \mathbf{x} \longmapsto A\mathbf{x} \in \mathbb{R}^n$ si chiamano rispettivamente autovalori, autovettori ed autospazi della matrice A . Per diagonalizzare A dobbiamo trovare una base di \mathbb{R}^n consistente da soli autovettori di A . Questo è possibile se e solo se la somma delle molteplicità geometriche di tutti gli autovalori di A è uguale a n .

NOTA 8.2.5. (Procedimento per la diagonalizzazione.) Per diagonalizzare una matrice quadrata A di ordine n , dobbiamo prima calcolare tutti gli autovalori di A , cioè

- tutti i $\lambda \in \mathbb{R}$ per i quali il sistema omogeneo $(A - \lambda I_n)\mathbf{x} = \mathbf{0}_n$ ha almeno una soluzione non nulla.

Se $\lambda_1, \dots, \lambda_k$ sono tutti gli autovalori diversi di A , allora per ogni $1 \leq j \leq k$ dobbiamo trovare una base per il relativo autospazio, cioè

- una base $\mathbf{v}_1^{(j)}, \dots, \mathbf{v}_{d_j}^{(j)}$ per lo spazio vettoriale di tutte le soluzioni del sistema omogeneo $(A - \lambda_j I_n)\mathbf{x} = \mathbf{0}_n$.

Se $d_1 + \dots + d_k = n$, allora A è diagonalizzabile,

$$\underbrace{\mathbf{v}_1^{(1)}, \dots, \mathbf{v}_{d_1}^{(1)}}_{d_1}, \dots, \underbrace{\mathbf{v}_1^{(k)}, \dots, \mathbf{v}_{d_k}^{(k)}}_{d_k}$$

è una base di \mathbb{R}^n consistente di soli autovettori di A e la formula (8.1.2) di diagonalizzazione diventa:

$$A = \left(\underbrace{\begin{pmatrix} | & & | \\ \mathbf{v}_1^{(1)} & \dots & \mathbf{v}_{d_1}^{(1)} \\ | & & | \end{pmatrix}}_{d_1} \dots \underbrace{\begin{pmatrix} | & & | \\ \mathbf{v}_1^{(k)} & \dots & \mathbf{v}_{d_k}^{(k)} \\ | & & | \end{pmatrix}}_{d_k} \right) \cdot \begin{pmatrix} \lambda_1 & \dots & 0 & & & \\ \vdots & \ddots & \vdots & & & 0 \\ 0 & \dots & \lambda_1 & & & \\ & & & \ddots & & \\ & & & & \lambda_k & \dots & 0 \\ & 0 & & & \vdots & \ddots & \vdots \\ & & & & 0 & \dots & \lambda_k \end{pmatrix} \cdot \left(\underbrace{\begin{pmatrix} | & & | \\ \mathbf{v}_1^{(1)} & \dots & \mathbf{v}_{d_1}^{(1)} \\ | & & | \end{pmatrix}}_{d_1} \dots \underbrace{\begin{pmatrix} | & & | \\ \mathbf{v}_1^{(k)} & \dots & \mathbf{v}_{d_k}^{(k)} \\ | & & | \end{pmatrix}}_{d_k} \right)^{-1}.$$

□

Il passo più difficile è trovare tutti gli autovalori di A . Questo si fa risolvendo trovando i valori di λ per i quali $A - \lambda \mathbb{I}$ non è invertibile. Diamo due esempi nei quale determiniamo la perdita dell'invertibilità dapprima in maniera artigianale mediante il metodo di eliminazione di Gauss, e poi con il calcolo del determinante.

ESEMPIO 8.2.6. Diagonalizziamo le matrici

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

Svolgimento. Gli autovalori di A sono i numeri reali λ per quali il sistema omogeneo

$$\begin{pmatrix} 2 - \lambda & -1 \\ -1 & 2 - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

cioè

$$\begin{cases} (2 - \lambda)x_1 - x_2 = 0 \\ -x_1 + (2 - \lambda)x_2 = 0 \end{cases},$$

ammette una soluzione non banale. Risolviamo questo sistema dapprima usando il metodo di eliminazione di Gauss: dai calcoli

$$\begin{array}{ccc|c} 2 - \lambda & -1 & 0 \\ -1 & 2 - \lambda & 0 \end{array}, \quad \begin{array}{ccc|c} -1 & 2 - \lambda & 0 \\ 2 - \lambda & -1 & 0 \end{array}, \quad \begin{array}{ccc|c} -1 & 2 - \lambda & 0 \\ 0 & (2 - \lambda)^2 - 1 & 0 \end{array}$$

risulta che il sistema ha soluzione non banale se e solo se $(2 - \lambda)^2 - 1 = 0$, cioè $2 - \lambda = \pm 1$, $\lambda = 2 \mp 1$. Pertanto gli autovalori di A sono $\lambda_1 = 1$ e $\lambda_2 = 3$.

A questo punto sappiamo già che A è diagonalizzabile: è una matrice 2×2 che ha due autovalori distinti.

Troviamo un autovettore corrispondente a $\lambda_1 = 1$: il sistema

$$\begin{array}{cc|c} -1 & 1 & 0 \\ 0 & 0 & 0 \end{array}$$

ha la soluzione non banale $x_1 = 1$, $x_2 = 1$. Poi troviamo un autovettore corrispondente anche a $\lambda_2 = 3$:

$$\begin{array}{cc|c} -1 & -1 & 0 \\ 0 & 0 & 0 \end{array}$$

ha la soluzione non banale $x_1 = 1$, $x_2 = -1$. Perciò gli autovettori

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ (per } \lambda_1 = 1 \text{)}, \quad \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ (per } \lambda_2 = 3 \text{)}$$

costituiscono una base di \mathbb{R}^2 ed una diagonalizzazione di A è

$$\begin{aligned} A &= \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}. \end{aligned}$$

Per trovare gli autovalori di A

abbiamo calcolato per quali λ la matrice $A - \lambda \mathbb{I}$ non è invertibile. Questo passaggio si poteva svolgere in un colpo solo per tutti gli autovalori, semplicemente imponendo che $\det(A - \lambda \mathbb{I}) = 0$ (come visto nel Capitolo 5). La funzione delle variabili λ data da $\det(A - \lambda \mathbb{I})$ si chiama *il polinomio caratteristico* della matrice A , ed è un polinomio dello stesso grado della dimensione di A , nel caso presente di grado 2, quindi risolvibile per radicali. Le sue radici sono tutti gli autovalori. (Quando $n > 2$ non sempre si sanno trovare le radici dell'equazione, e quindi il metodo del determinante porta alla soluzione solo se si riescono a trovare per ispezione diretta abbastanza radici).

Nel caso presente si ottiene:

$$P(\lambda) = \begin{vmatrix} 2 - \lambda & -1 \\ -1 & 2 - \lambda \end{vmatrix} = (2 - \lambda)^2 - 1 = \lambda^2 - 4\lambda + 3$$

le cui radici sono appunto i due autovalori $\lambda = 1$ e $\lambda = 3$ precedentemente trovati.

In maniera analoga si procede con la matrice B . Si trova che il solo autovalore di B è 0 e che l'autospazio di B corrispondente a questo autovalore consiste dai multipli scalari del vettore $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Perciò non esiste una base di \mathbb{R}^2 consistente da soli autovettori di B , ossia B non è diagonalizzabile. \square

ESERCIZIO 8.2.7. Si diagonalizzi la matrice

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

Soluzione. Calcoliamo anche questa volta gli autovalori dapprima in maniera diretta, con il metodo di eliminazione di Gauss, e poi in maniera più rapida tramite le radici del *polinomio caratteristico* introdotto nel precedente Esempio 8.2.6.

Gli autovalori di A sono gli scalari λ per quali il sistema omogeneo

$$\begin{pmatrix} -1-\lambda & 1 & 0 \\ 4 & -\lambda & 4 \\ 0 & 1 & 1-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

cioè

$$\begin{cases} (-1-\lambda)x_1 + x_2 = 0 \\ 4x_1 - \lambda x_2 + 4x_3 = 0 \\ x_2 + (1-\lambda)x_3 = 0 \end{cases},$$

ammette soluzione non zero. Applichiamo il metodo di eliminazione di Gauss:

$$\begin{array}{ccc|ccc|c} -1-\lambda & 1 & 0 & 0 & 1 & -\frac{\lambda}{4} & 1 & 0 \\ 4 & -\lambda & 4 & 0 & 0 & 1 & 1-\lambda & 0 \\ 0 & 1 & 1-\lambda & 0 & -1-\lambda & 1 & 0 & 0 \\ \hline 1 & -\frac{\lambda}{4} & 1 & 0 & 1 & -\frac{\lambda}{4} & 1 & 0 \\ 0 & 1 & 1-\lambda & 0 & 0 & 1 & 1-\lambda & 0 \\ 0 & 1-\frac{\lambda}{4}-\frac{\lambda^2}{4} & 1+\lambda & 0 & 0 & 0 & \frac{9\lambda}{4}-\frac{\lambda^3}{4} & 0 \end{array}$$

mostra che il sistema ha soluzione non zero se e solo se $\frac{9\lambda}{4} - \frac{\lambda^3}{4} = 0$. Perciò gli autovalori di A sono $\lambda_1 = 0$, $\lambda_2 = 3$, $\lambda_3 = -3$.

Avremmo trovato più velocemente questi tre autovalori se avessimo calcolato le radici del polinomio caratteristico

$$\begin{vmatrix} -1-\lambda & 1 & 0 \\ 4 & -\lambda & 4 \\ 0 & 1 & 1-\lambda \end{vmatrix},$$

che, dopo varie semplificazioni, si trova essere un multiplo di $9\lambda - \lambda^3$. Si noti che questo calcolo, persino nel presente caso di dimensione piccola ($n = 3$), è laborioso. In effetti, il determinante (o meglio un suo multiplo) si calcola più facilmente proprio se si esegue l'eliminazione di Gauss! Si osservi infatti che il determinante non cambia sotto le operazioni di riga del metodo di Gauss tranne per la moltiplicazione per costanti (-1 quando si scambiano due righe adiacenti, e α quando si moltiplica una riga per la costante α): quindi, già in questo esempio a dimensione soltanto 3, il modo più agevole di calcolare il determinante è di svolgere l'eliminazione di gauss come abbiamo fatto sopra e poi calcolare il determinante dell'ultima riduzione, quella

in cui la matrice è diventata triangolare, e quindi ha per determinante il prodotto dei termini diagonali.

Abbiamo trovato i tre autovalori di A . Poiché A è una matrice 3×3 ed ha tre autovalori distinti, essa è diagonalizzabile. Per trovare una diagonalizzazione di A , troviamo un autovettore ad ogni autovalore:

- Autovettore corrispondente a $\lambda_1 = 0$:

$$\begin{array}{ccc|c} -1 & 1 & 0 & 0 \\ 4 & 0 & 4 & 0 \\ 0 & 1 & 1 & 0 \end{array}, \quad \begin{array}{ccc|c} -1 & 1 & 0 & 0 \\ 0 & 4 & 4 & 0 \\ 0 & 1 & 1 & 0 \end{array}, \quad \begin{array}{ccc|c} -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array},$$

$$\mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

- Autovettore corrispondente a $\lambda_2 = 3$:

$$\begin{array}{ccc|c} -4 & 1 & 0 & 0 \\ 4 & -3 & 4 & 0 \\ 0 & 1 & -2 & 0 \end{array}, \quad \begin{array}{ccc|c} -4 & 1 & 0 & 0 \\ 0 & -2 & 4 & 0 \\ 0 & 1 & -2 & 0 \end{array}, \quad \begin{array}{ccc|c} -4 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array},$$

$$\mathbf{y} = \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}.$$

- Autovettore corrispondente a $\lambda_3 = -3$:

$$\begin{array}{ccc|c} 2 & 1 & 0 & 0 \\ 4 & 3 & 4 & 0 \\ 0 & 1 & 4 & 0 \end{array}, \quad \begin{array}{ccc|c} 2 & 1 & 0 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 1 & 4 & 0 \end{array}, \quad \begin{array}{ccc|c} 2 & 1 & 0 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{array},$$

$$\mathbf{z} = \begin{pmatrix} 2 \\ -4 \\ 1 \end{pmatrix}.$$

Abbiamo così ottenuto la diagonalizzazione

$$\begin{pmatrix} -1 & 1 & 0 \\ 4 & 0 & 4 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 4 & -4 \\ -1 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 2 \\ 1 & 4 & -4 \\ -1 & 2 & 1 \end{pmatrix}^{-1}.$$

□

8.3. Ulteriori esercizi sulla diagonalizzazione sul campo \mathbb{R}

ESERCIZIO 8.3.1. Si dica se la matrice

$$\begin{pmatrix} 7 & 6 \\ 6 & 2 \end{pmatrix}$$

è diagonalizzabile o no e se lo è si trovi una sua diagonalizzazione.

Soluzione. Il polinomio caratteristico della matrice è

$$P(\lambda) = \begin{vmatrix} 7-\lambda & 6 \\ 6 & 2-\lambda \end{vmatrix} = (7-\lambda)(2-\lambda) - 36 = \lambda^2 - 9\lambda - 22$$

e risolvendo l'equazione $P(\lambda) = 0$ risultano gli autovalori $\lambda_1 = 11$ e $\lambda_2 = -2$. Poiché $\lambda_1 \neq \lambda_2$, la matrice è diagonalizzabile.

- Autovettore corrispondente a $\lambda_1 = 11$:

$$\begin{array}{ccc} \begin{vmatrix} -4 & 6 \\ 6 & -9 \end{vmatrix} \begin{array}{c} 0 \\ 0 \end{array} & , & \begin{vmatrix} -2 & 3 \\ 2 & -3 \end{vmatrix} \begin{array}{c} 0 \\ 0 \end{array} & , & \begin{vmatrix} -2 & 3 \\ 0 & 0 \end{vmatrix} \begin{array}{c} 0 \\ 0 \end{array} \end{array}$$

$$\mathbf{x} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

- Autovettore corrispondente a $\lambda_2 = -2$:

$$\begin{array}{ccc} \begin{vmatrix} 9 & 6 \\ 6 & 4 \end{vmatrix} \begin{array}{c} 0 \\ 0 \end{array} & , & \begin{vmatrix} 3 & 2 \\ 3 & 2 \end{vmatrix} \begin{array}{c} 0 \\ 0 \end{array} & , & \begin{vmatrix} 3 & 2 \\ 0 & 0 \end{vmatrix} \begin{array}{c} 0 \\ 0 \end{array} \end{array}$$

$$\mathbf{y} = \begin{pmatrix} -2 \\ 3 \end{pmatrix}.$$

Si ottiene la diagonalizzazione

$$\begin{aligned} \begin{pmatrix} 7 & 6 \\ 6 & 2 \end{pmatrix} &= \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 11 & 0 \\ 0 & -2 \end{pmatrix} \cdot \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 11 & 0 \\ 0 & -2 \end{pmatrix} \cdot \begin{pmatrix} \frac{3}{13} & \frac{2}{13} \\ -\frac{2}{13} & \frac{3}{13} \end{pmatrix}. \end{aligned}$$

□

ESERCIZIO 8.3.2. Si discuti la diagonalizzazione della matrice

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

Soluzione. Il polinomio caratteristico della matrice è

$$\begin{vmatrix} -1-\lambda & 1 & 3 \\ 0 & 1-\lambda & 0 \\ 3 & 1 & -1-\lambda \end{vmatrix},$$

che si calcola usando sviluppo rispetto alla seconda riga:

$$(1-\lambda) \begin{vmatrix} -1-\lambda & 3 \\ 3 & -1-\lambda \end{vmatrix} = (1-\lambda)(\lambda^2 + 2\lambda - 8).$$

Risultano tre autovalori diversi: $1, -4, 2$, quindi la matrice è diagonalizzabile. Gli autovettori corrispondenti all'autovalore λ sono le soluzioni non nulle del sistema omogeneo

$$\begin{pmatrix} -1-\lambda & 1 & 3 \\ 0 & 1-\lambda & 0 \\ 3 & 1 & -1-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Si ottiene

$$\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \text{ per } 1, \quad \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \text{ per } -4, \quad \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \text{ per } 2.$$

Si ha quindi la seguente diagonalizzazione della matrice:

$$\begin{aligned} & \begin{pmatrix} -1 & 1 & 3 \\ 0 & 1 & 0 \\ 3 & 1 & -1 \end{pmatrix} = \\ & = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix}^{-1} = \\ & = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 & 0 \\ 1/2 & 0 & -1/2 \\ 1/2 & 1 & 1/2 \end{pmatrix}. \end{aligned}$$

□

ESERCIZIO 8.3.3. Si discuta la diagonalizzazione della matrice

$$\begin{pmatrix} 5 & -7 & 8 \\ -7 & 5 & 8 \\ 8 & 8 & -10 \end{pmatrix}.$$

Soluzione. Il polinomio caratteristico della matrice è

$$P(\lambda) = \begin{vmatrix} 5-\lambda & -7 & 8 \\ -7 & 5-\lambda & 8 \\ 8 & 8 & -10-\lambda \end{vmatrix},$$

che si calcola sottraendo la seconda riga alla prima ed osservando che dopo questa operazione la prima riga diventa divisibile per $12 - \lambda$:

$$\begin{aligned}
 P(\lambda) &= \begin{vmatrix} 12 - \lambda & \lambda - 12 & 0 \\ -7 & 5 - \lambda & 8 \\ 8 & 8 & -10 - \lambda \end{vmatrix} = \\
 &= (12 - \lambda) \begin{vmatrix} 1 & -1 & 0 \\ -7 & 5 - \lambda & 8 \\ 8 & 8 & -10 - \lambda \end{vmatrix} = \\
 &= (12 - \lambda) \begin{vmatrix} 1 & 0 & 0 \\ -7 & -2 - \lambda & 8 \\ 8 & 16 & -10 - \lambda \end{vmatrix} = \\
 &= (12 - \lambda) \begin{vmatrix} -2 - \lambda & 8 \\ 16 & -10 - \lambda \end{vmatrix} = \\
 &= (12 - \lambda) (\lambda^2 + 12\lambda - 108).
 \end{aligned}$$

Si ottengono quindi gli autovalori $\lambda_1 = 12$ e

$$\lambda_{2,3} = -6 \pm \sqrt{36 + 108} = -6 \pm 12, \quad \text{cioè } \lambda_2 = 6, \lambda_3 = -18,$$

- Autovettore corrispondente a $\lambda_1 = 12$:

$$\begin{array}{ccc|c} -7 & -7 & 8 & 0 \\ -7 & -7 & 8 & 0 \\ 8 & 8 & -22 & 0 \end{array}, \quad \begin{array}{ccc|c} -7 & -7 & 8 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -90 & 0 \end{array},$$

$$\mathbf{x} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

- Autovettore corrispondente a $\lambda_2 = 6$:

$$\begin{array}{ccc|c} -1 & -7 & 8 & 0 \\ -7 & -1 & 8 & 0 \\ 8 & 8 & -16 & 0 \end{array}, \quad \begin{array}{ccc|c} -1 & -7 & 8 & 0 \\ -7 & -1 & 8 & 0 \\ 1 & 1 & -2 & 0 \end{array},$$

$$\begin{array}{ccc|c} -1 & -7 & 8 & 0 \\ 0 & 48 & -48 & 0 \\ 0 & -6 & 6 & 0 \end{array}, \quad \begin{array}{ccc|c} -1 & -7 & 8 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array},$$

$$\mathbf{y} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

- Autovettore corrispondente a $\lambda_3 = -18$:

$$\begin{array}{ccc|c} 23 & -7 & 8 & 0 \\ -7 & 23 & 8 & 0 \\ 8 & 8 & 8 & 0 \end{array} \quad , \quad \begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 23 & -7 & 8 & 0 \\ -7 & 23 & 8 & 0 \end{array}$$

$$\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -30 & -15 & 0 \\ 0 & 30 & 15 & 0 \end{array} \quad , \quad \begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

$$\mathbf{z} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

Si ottiene la diagonalizzazione

$$\begin{aligned} & \begin{pmatrix} 5 & -7 & 8 \\ -7 & 5 & 8 \\ 8 & 8 & -10 \end{pmatrix} = \\ & = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & 1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 12 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -18 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & 1 & -2 \end{pmatrix}^{-1} = \\ & = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & 1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 12 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -18 \end{pmatrix} \cdot \begin{pmatrix} 1/2 & -1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 1/6 & 1/6 & -1/3 \end{pmatrix}. \end{aligned}$$

□

ESERCIZIO 8.3.4. Si discuti la diagonalizzazione della matrice

$$\begin{pmatrix} 16 & -12 & 0 \\ -12 & 9 & 0 \\ 0 & 0 & 25 \end{pmatrix}.$$

Soluzione. Anzitutto, la matrice è simmetrica e perciò certamente diagonalizzabile, come dimostreremo in un prossimo teorema

Il polinomio caratteristico della matrice è

$$\begin{aligned} P(\lambda) &= \begin{vmatrix} 16 - \lambda & -12 & 0 \\ -12 & 9 - \lambda & 0 \\ 0 & 0 & 25 - \lambda \end{vmatrix} = (25 - \lambda) \begin{vmatrix} 16 - \lambda & -12 \\ -12 & 9 - \lambda \end{vmatrix} = \\ &= (25 - \lambda) ((16 - \lambda)(9 - \lambda) - 144) = -\lambda(\lambda - 25)^2, \end{aligned}$$

perciò gli autovalori della matrice sono $\lambda_1 = 0$ e $\lambda_2 = 25$.

- Autovettori corrispondenti a $\lambda_1 = 0$:

$$\begin{array}{ccc|c}, & \begin{array}{ccc|c} 4 & -3 & 0 & 0 \\ -4 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}, & \begin{array}{ccc|c} 4 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}, \end{array}$$

$$\mathbf{x} = \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}.$$

- Autovettori corrispondenti a $\lambda_2 = 25$: dai calcoli

$$\begin{array}{ccc|c}, & \begin{array}{ccc|c} -3 & -4 & 0 & 0 \\ -3 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}, & \begin{array}{ccc|c} 3 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}, \end{array}$$

risulta che la forma generale di questi autovettori è

$$\begin{pmatrix} -\frac{4}{3}x_2 \\ x_2 \\ x_3 \end{pmatrix} = \frac{1}{3}x_2 \begin{pmatrix} -4 \\ 3 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

perciò abbiamo due autovettori linearmente indipendenti corrispondenti a $\lambda_2 = 25$:

$$\mathbf{y} = \begin{pmatrix} -4 \\ 3 \\ 0 \end{pmatrix}, \mathbf{z} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Si ottiene la diagonalizzazione:

$$\begin{aligned} & \begin{pmatrix} 16 & -12 & 0 \\ -12 & 9 & 0 \\ 0 & 0 & 25 \end{pmatrix} \\ &= \begin{pmatrix} 3 & -4 & 0 \\ 4 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 \\ 0 & 25 & 0 \\ 0 & 0 & 25 \end{pmatrix} \cdot \begin{pmatrix} 3 & -4 & 0 \\ 4 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \\ &= \begin{pmatrix} 3 & -4 & 0 \\ 4 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 \\ 0 & 25 & 0 \\ 0 & 0 & 25 \end{pmatrix} \cdot \begin{pmatrix} 3/25 & 4/25 & 0 \\ -4/25 & 3/25 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

□

ESERCIZIO 8.3.5. Si discuta la diagonalizzabilità della matrice

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

Soluzione. Il polinomio caratteristico della matrice è

$$P(\lambda) = \begin{vmatrix} 3-\lambda & 1 & 2 \\ -2 & -\lambda & -2 \\ -1 & 0 & -1-\lambda \end{vmatrix},$$

che si calcola sommando la seconda riga alla prima ed osservando che dopo questa operazione la prima riga diventa divisibile per $1-\lambda$:

$$\begin{aligned} P(\lambda) &= \begin{vmatrix} 1-\lambda & 1-\lambda & 0 \\ -2 & -\lambda & -2 \\ -1 & 0 & -1-\lambda \end{vmatrix} = (1-\lambda) \begin{vmatrix} 1 & 1 & 0 \\ -2 & -\lambda & -2 \\ -1 & 0 & -1-\lambda \end{vmatrix} = \\ &= (1-\lambda) \begin{vmatrix} 1 & 0 & 0 \\ -2 & 2-\lambda & -2 \\ -1 & 1 & -1-\lambda \end{vmatrix} = (1-\lambda) \begin{vmatrix} 2-\lambda & -2 \\ 1 & -1-\lambda \end{vmatrix} = \\ &= (1-\lambda)(\lambda^2 - \lambda) = -\lambda(\lambda - 1)^2. \end{aligned}$$

Pertanto si trovano gli autovalori $\lambda_1 = 0$, $\lambda_2 = 1$.

Risolvendo il sistema omogeneo

$$\begin{pmatrix} 3-\lambda & 1 & 2 \\ -2 & -\lambda & -2 \\ -1 & 0 & -1-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

per $\lambda = 0$, si trovano come soluzioni i multipli scalari del vettore

$$\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix},$$

mentre risolvendolo per $\lambda = 1$, si trovano i multipli scalari del vettore

$$\begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}.$$

Di conseguenza il sottospazio lineare di \mathbb{R}^3 generato da tutti gli autovettori della matrice è il sottospazio generato dai due vettori di cui sopra e pertanto non è uguale a \mathbb{R}^3 . Concludiamo che la matrice non è diagonalizzabile. \square

ESERCIZIO 8.3.6. Si trovino gli autovalori della matrice

$$\begin{pmatrix} 5 & -2 & 0 & 1 \\ -2 & 5 & 1 & 0 \\ 0 & 1 & 5 & -2 \\ 1 & 0 & -2 & 5 \end{pmatrix}.$$

Soluzione. Il polinomio caratteristico della matrice è

$$P(\lambda) = \begin{vmatrix} 5-\lambda & -2 & 0 & 1 \\ -2 & 5-\lambda & 1 & 0 \\ 0 & 1 & 5-\lambda & -2 \\ 1 & 0 & -2 & 5-\lambda \end{vmatrix},$$

che si comincia a calcolare sommando la seconda, terza e quarta riga alla prima ed osservando che dopo questa operazione la prima riga diventa divisibile per $4 - \lambda$:

$$\begin{aligned} P(\lambda) &= \begin{vmatrix} 4-\lambda & 4-\lambda & 4-\lambda & 4-\lambda \\ -2 & 5-\lambda & 1 & 0 \\ 0 & 1 & 5-\lambda & -2 \\ 1 & 0 & -2 & 5-\lambda \end{vmatrix} = \\ &= (4-\lambda) \begin{vmatrix} 1 & 1 & 1 & 1 \\ -2 & 5-\lambda & 1 & 0 \\ 0 & 1 & 5-\lambda & -2 \\ 1 & 0 & -2 & 5-\lambda \end{vmatrix} = \\ &= (4-\lambda) \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 7-\lambda & 3 & 2 \\ 0 & 1 & 5-\lambda & -2 \\ 0 & -1 & -3 & 4-\lambda \end{vmatrix} = \\ &= (4-\lambda) \begin{vmatrix} 7-\lambda & 3 & 2 \\ 1 & 5-\lambda & -2 \\ -1 & -3 & 4-\lambda \end{vmatrix}. \end{aligned}$$

Ora si sottraggono alla prima colonna la seconda e la terza e si osserva che dopo questa operazione la prima colonna diventa divisibile per $2 - \lambda$:

$$\begin{aligned} P(\lambda) &= (4-\lambda) \begin{vmatrix} 2-\lambda & 3 & 2 \\ -2+\lambda & 5-\lambda & -2 \\ -2+\lambda & -3 & 4-\lambda \end{vmatrix} = \\ &= (4-\lambda)(2-\lambda) \begin{vmatrix} 1 & 3 & 2 \\ -1 & 5-\lambda & -2 \\ -1 & -3 & 4-\lambda \end{vmatrix} = \\ &= (4-\lambda)(2-\lambda) \begin{vmatrix} 1 & 3 & 2 \\ 0 & 8-\lambda & 0 \\ 0 & 0 & 6-\lambda \end{vmatrix} = \\ &= (4-\lambda)(2-\lambda)(8-\lambda)(6-\lambda). \end{aligned}$$

Risultano gli autovalori $\lambda_1 = 2$, $\lambda_2 = 4$, $\lambda_3 = 6$, $\lambda_4 = 8$.

□

ESERCIZIO 8.3.7. Consideriamo la matrice

$$A = \begin{pmatrix} 2 & 6 & -9 \\ -1 & -5 & 9 \\ 0 & p & 8 \end{pmatrix},$$

dove p è un parametro reale.

- (i) Si calcoli il polinomio caratteristico di A .
- (ii) Per che valori di p tutti gli autovalori di A sono reali?
- (iii) Per che valori di p tutti gli autovalori di A sono reali e distinti?
- (iv) Per che valori di p la matrice A è diagonalizzabile?

Soluzione.

- (i) Il polinomio caratteristico di A è

$$\begin{aligned} \begin{vmatrix} 2-\lambda & 6 & -9 \\ -1 & -5-\lambda & 9 \\ 0 & p & 8-\lambda \end{vmatrix} &= \begin{vmatrix} 1-\lambda & 1-\lambda & 0 \\ -1 & -5-\lambda & 9 \\ 0 & p & 8-\lambda \end{vmatrix} = \\ &= (1-\lambda) \begin{vmatrix} 1 & 1 & 0 \\ -1 & -5-\lambda & 9 \\ 0 & p & 8-\lambda \end{vmatrix} = \\ &= (1-\lambda) \begin{vmatrix} 1 & 1 & 0 \\ 0 & -4-\lambda & 9 \\ 0 & p & 8-\lambda \end{vmatrix} = \\ &= (1-\lambda) (\lambda^2 - 4\lambda - 32 - 9p). \end{aligned}$$

- (ii) Gli autovalori di A sono $\lambda_1 = 1$ e $\lambda_{2,3} = 2 \pm 3\sqrt{4+p}$, quindi sono tutti reali se e soltanto se $p \geq -4$.
- (iii) Risulta

$$\lambda_2 = \lambda_3 \quad \text{per} \quad p = -4$$

e

$$\lambda_1 = \lambda_3 \quad \text{per} \quad p = -4 + \frac{1}{9} = -\frac{35}{9},$$

perciò A risulta avere tutti gli autovalori reali e distinti esattamente per $-\frac{35}{9} \neq p > -4$.

- (iv) Calcolando tutti gli autovettori esplicitamente nei casi $p = -4$ (autovalori 1 e 2) e $p = -\frac{35}{9}$ (autovalori 1 e 2), si trova che essi non generano \mathbb{R}^3 in nessuno di questi casi. Sapendo poi che A è diagonalizzabile quando tutti gli autovalori sono reali e distinti, risulta che A è diagonalizzabile esattamente in questo caso, cioè quando $-\frac{35}{9} \neq p > -4$.

□

8.4. Autovalori complessi e diagonalizzazione in $M_{(nn)}^{\mathbb{C}}$

Il polinomio caratteristico di ogni matrice $A \in M_{nn}^{\mathbb{C}}$ è un polinomio di grado n a coefficienti complessi. Un celebre risultato, il Teorema Fondamentale dell'Algebra, la cui dimostrazione qui viene omessa, afferma che ogni polinomio complesso di grado n ha sempre n radici complesse, se le si conta con la loro molteplicità. Quindi, a differenza del caso reale, ogni matrice complessa di grado n ha sempre n autovalori complessi, ma alcuni possono essere ripetuti. Se gli autovalori sono distinti, anche se magari non reali, i corrispondenti autovettori sono indipendenti (il Teorema di indipendenza 8.2.3 continua a valere nel caso complesso, con la stessa dimostrazione). Se invece un autovalore è ripetuto, diciamo con molteplicità m , allora non è detto che il suo autospazio contenga m autovettori linearmente indipendenti, e se questo non accade non si può avere una base di autovettori e la matrice non è diagonalizzabile. (Per avere esempi, si veda in seguito la Sezione ??).

Si osservi che, se la matrice A è reale, nondimeno il suo polinomio caratteristico, che è a coefficienti reali, può avere qualche radice λ complessa. In tal caso questi autovalori complessi devono avere autovettori con qualche componente complessa, altrimenti l'equazione $A\mathbf{x} = \lambda\mathbf{x}$ non potrebbe valere per $\mathbf{x} \neq \mathbf{0}$. Pertanto A non è diagonalizzabile in $M_{nn}^{\mathbb{R}}$ (si dice che A non è *diagonalizzabile sui reali*). Però potrebbe esserlo in $M_{nn}^{\mathbb{C}}$. Ecco un esempio:

ESEMPIO 8.4.1. (**Diagonalizzazione sui complessi di matrici di rotazione.**) Come nell'Esempio 6.8.5, sia A_θ la matrice reale di rotazione 2×2 data da

$$A_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.$$

È ovvio che, a meno che $\theta = 0$ o π (nei quali casi $A_\theta = \pm \mathbb{I}$ lascia invariante la direzione di ogni vettore nel piano), A_θ non può avere autovettori reali, perché ogni vettore non nullo in \mathbb{R}^2 viene ruotato di un angolo diverso da 0 e da π e quindi non viene mandato in un multiplo di se stesso. Vedremo infatti che la matrice ha autovettori complessi invece che reali, e quindi non è diagonalizzabile sui reali, ma anche che i due autovettori, complessi coniugati, sono diversi (certo! se no sarebbero reali!), e quindi A_θ è *diagonalizzabile sui complessi*.

In effetti, si ha

$$\det(A_\theta - \lambda \mathbb{I}) = \begin{vmatrix} \cos \theta - \lambda & \sin \theta \\ -\sin \theta & \cos \theta - \lambda \end{vmatrix} = (\cos \theta - \lambda)^2 + \sin^2 \theta = \lambda^2 - 2\lambda \cos \theta + 1,$$

le cui radici sono $\lambda_{\pm} = \cos \theta \pm i \sin \theta = e^{\pm i\theta}$. La forma diagonale di A_θ , su \mathbb{C} , è quindi

$$\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}.$$

Gli autovettori complessi \mathbf{x}_{\pm} sono i vettori nel nucleo di $A - \lambda_{\pm} \mathbb{I}$: il metodo di eliminazione di Gauss ci permette di trovarli. Lasciamo i calcoli al lettore.

8.5. Diagonalizzabilità di matrici simmetriche o autoaggiunte

La definizione di matrici simmetriche o autoaggiunte è stata data in 6.4.6.

TEOREMA 8.5.1. (Autovalori ed autovettori di matrici simmetriche o autoaggiunte.)

Sia A una matrice simmetrica (o più in generale autoaggiunta). Allora

- (i) *Tutti gli autovalori di A sono numeri reali.*
- (ii) *Autovettori di A corrispondenti ad autovalori diversi sono ortogonali.*
- (iii) *Esiste almeno un autovalore reale di A .*

Dimostrazione.

- (i) Se la matrice A è autoaggiunta essa agisce sullo spazio complesso \mathbb{C}^n ; se è simmetrica, agisce sullo spazio reale \mathbb{R}^3 ma più in generale si può considerare la sua azione sullo spazio complesso \mathbb{C}^3 . In entrambi i casi i suoi autovalori, a priori, potrebbero essere complessi (ovviamente, quando A è una matrice a coefficienti reali, se essa ha un autovalore complesso il corrispondente autovettore deve essere a coefficienti complessi, e quindi non esiste in \mathbb{R}^3 , ma - al più - in \mathbb{C}^3). Nella dimostrazione consideriamo il caso più generale dell'azione della matrice autoaggiunta A sullo spazio complesso \mathbb{C}^n , munito quindi del prodotto scalare sui complessi

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i \overline{y_i}$$

introdotto in (6.3.2).

Sia \mathbf{x} un autovettore di A con autovalore λ . Poiché A è a coefficienti reali abbiamo $A^* = A^\top = A$, e quindi

$$A\mathbf{x} \cdot \mathbf{x} = \mathbf{x} \cdot A\mathbf{x}$$

grazie a (6.4.2).

Pertanto

$$\lambda \mathbf{x} \cdot \mathbf{x} = A\mathbf{x} \cdot \mathbf{x} = \mathbf{x} \cdot A\mathbf{x} = \overline{\lambda} \mathbf{x} \cdot \mathbf{x}.$$

Dal momento che l'autovettore \mathbf{x} deve essere non nullo, la sua norma $\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$ è diversa da zero, e quindi l'identità precedente implica $\lambda \in \mathbb{R}$.

- (ii) Siano $\mathbf{x}_1, \mathbf{x}_2$ autovettori di A con rispettivi autovalori $\lambda_1 \neq \lambda_2$. Allora

$$\lambda_1 \mathbf{x}_1 \cdot \mathbf{x}_2 = A\mathbf{x}_1 \cdot \mathbf{x}_2 = \mathbf{x}_1 \cdot A\mathbf{x}_2 = \lambda_2 \mathbf{x}_1 \cdot \mathbf{x}_2$$

(nell'ultima uguaglianza non abbiamo scritto $\overline{\lambda_2}$ perchè l'autovalore è reale grazie alla parte (i) del teorema). Poiché $\lambda_1 \neq \lambda_2$ ne segue $\mathbf{x}_1 \cdot \mathbf{x}_2 = 0$.

- (iii) Per il Teorema Fondamentale dell'algebra, il polinomio caratteristico di A ha almeno una radice, e quindi A ha almeno un autovalore; per la parte (i) questo autovalore è reale.

Quando si sa che due autovettori sono ortogonali, li si può scegliere ortonormali: basta normalizzarli, cioè dividerli per la loro norma (lunghezza). Quindi il precedente Teorema 8.5.1 porta a questa conseguenza:

COROLLARIO 8.5.2. *Se tutti gli autovalori di una matrice autoaggiunta (in particolare simmetrica) sono distinti, allora la matrice è diagonalizzabile, ed esiste una base ortonormale di autovettori. Equivalentemente, se tutti gli autovalori di un operatore autoaggiunto (in particolare simmetrico) sono distinti, allora esiste una base ortonormale di autovettori.*

TEOREMA 8.5.3. (Diagonalizzabilità di operatori autoaggiunti.)

- (i) *Ogni matrice autoaggiunta è diagonalizzabile sui complessi, e la matrice di cambiamento di base che la diagonalizza è una matrice unitaria. Equivalentemente, ogni applicazione lineare autoaggiunta (noi diremo ora operatore autoaggiunto) ammette una base ortonormale di autovettori.*
- (ii) *Ogni matrice simmetrica è diagonalizzabile sui reali, e la matrice di cambiamento di base che la diagonalizza è una matrice ortogonale. Equivalentemente, ogni operatore simmetrico su uno spazio vettoriale reale ammette una base ortonormale di autovettori.*

Dimostrazione. Dimostriamo solo la parte (i): la dimostrazione della parte (ii) è identica. Gli autovettori con autovalori distinti si possono scegliere ortonormali per il Corollario 8.5.2. Quindi basta provare che per ogni autovalore si possono trovare tanti autovettori quanta è la molteplicità dell'autovalore (se non sono fra loro ortogonali li si può ortogonalizzare all'interno dell'autospazio con il procedimento di Gram-Schmidt illustrato nella Sezione 6.7, che rimpiazza i vettori con loro combinazioni lineari opportune, le quali quindi restano autovettori).

Equivalentemente, ora mostriamo la seguente asserzione: *ogni operatore autoaggiunto A ammette una base ortonormale di autovettori.* La dimostrazione ricalca da vicino quella del Teorema 6.6.9 (cioè proprio del teorema di proiezione ortogonale che dà origine alla ortogonalizzazione di Gram-Schmidt: in effetti le due dimostrazioni sono essenzialmente equivalenti, visto che ora stiamo usando il prodotto scalare euclideo, che è definito positivo).

Procediamo di nuovo per induzione sulla dimensione n dello spazio vettoriale su cui A opera (ancora una volta stiamo usando l'assioma di induzione degli interi, (1.3)). Se $n = 1$ l'asserzione è ovvia, perché qualsiasi vettore non nullo forma una base, e si può normalizzare. Assumiamo ora $n > 1$. Per il Teorema 8.5.1 (iii) l'operatore A ha un autovalore reale λ ; indichiamo con \mathbf{x}_1 un autovettore corrispondente, e sia \mathbf{X}_1 lo spazio vettoriale che esso genera. Mostriamo che il complemento ortogonale \mathbf{X}_1^\perp è *invariante* sotto l'operatore A , cioè che $A\mathbf{X}_1^\perp \subset \mathbf{X}_1^\perp$. Questo è vero perché, se $\mathbf{x} \in \mathbf{X}_1^\perp$, cioè se $\mathbf{x} \cdot \mathbf{x}_1 = 0$, allora

$$A\mathbf{x} \cdot \mathbf{x}_1 = \mathbf{x} \cdot A\mathbf{x}_1 = 0 = \lambda \mathbf{x} \cdot A\mathbf{x}_1 = 0.$$

D'altra parte, sempre perché il prodotto scalare euclideo è definito positivo, si ha $\dim \mathbf{X}_1^\perp = n - 1$ (Proposizione ?? (v); qui basterebbe usare il fatto che il prodotto scalare euclideo è

non degenerare e ricorrere alla Proposizione ?? (iv)). Perciò, per ipotesi di induzione, \mathbf{X}_1^\perp ha una base ortonormale: aggiungendo a questa base il vettore \mathbf{x}_1 , come nella dimostrazione del Teorema 6.6.9, si ottiene una base ortonormale per \mathbf{X} . L'asserzione è dimostrata.

La matrice che porta dalla base canonica alla base ortonormale, cioè la matrice del cambiamento di base che diagonalizza A , è unitaria (nel caso reale, ortogonale) per il Teorema 6.8.7, perchè manda la base canonica, ovviamente ortonormale rispetto al prodotto scalare euclideo, in una base ortonormale.

COROLLARIO 8.5.4. *Per ogni matrice A autoaggiunta (rispettivamente, simmetrica) esiste una matrice U unitaria (rispettivamente, ortogonale) tale che $U^\top AU = U^{-1}AU$ è diagonale.*

8.6. Esercizi sulla diagonalizzazione di matrici simmetriche

8.7. Triangolarizzazione e forma canonica di Jordan

8.8. Una applicazione: dinamica della popolazione

CAPITOLO 9

Somma diretta di spazi vettoriali

CAPITOLO 10

Esempi di geometria analitica nel piano

CAPITOLO 11

Esempi di geometria analitica in \mathbb{R}^3

Bibliografia

- [1] M. Abate, *Algebra Lineare*, McGraw-Hill Italia, 2000. [1](#)