Dispensa del corso di Geometria e Algebra Lineare[1]

Ingegneria dell'Automazione, Elettrica, Elettronica, Informatica e Telecomunicazioni

Anno Accademico 2017/2018

Marco Compagnoni

Dipartimento di Matematica, Politecnico di Milano
E-mail address, M. Compagnoni: `marco.compagnoni@polimi.it`

Indice

CAPITOLO 1

Introduzione

La geometria nacque per esigenze pratiche (ad esempio in agricoltura, nella costruzione di edifici e nell'astronomia) come scienza sperimentale e come tale fu probabilmente una delle prime della storia dell'umanità. I più antichi risultati empirici riguardanti lunghezze, angoli, aree e volumi risalgono alle civiltà egizie, babilonesi, cinesi e dell'antica India. Attraverso procedimenti induttivi, molte nozioni, tecniche e concetti geometrici si accumularono nel corso dei secoli, al punto da rendere necessario una sistematizzazione della disciplina.

Così come Democrito può essere considerato il padre della fisica, grazie alla sua idea di atomo che è a fondamento dell'approccio riduzionistico di tale scienza, allo stesso modo Euclide deve essere considerato il fondatore della matematica come oggi viene propriamente intesa. La sua opera più famosa sono naturalmente gli Elementi. Composti da 13 volumi, in essi Euclide raccoglie i risultati a lui noti riguardanti la geometria piana, la geometria solida e l'aritmetica dei numeri razionali ed irrazionali. Di fatto, essi costituiscono il più antico trattato organico di matematica a noi pervenutoci, uno dei prodotti più significativi della civiltà greca antica. L'aspetto più importante di questa opera è sicuramente la metodologia in essa sviluppata. Euclide cercò di sistematizzare la geometria inventando il cosiddetto approccio assiomatico. La sua idea fu di individuare un nucleo ristretto di concetti fondamentali, che fossero veri in se senza bisogno di prova e che non fossero deducibili l'uno dall'altro.

Ad esempio, nel caso della geometria piana Euclide definisce alcune entità elementari come i punti, i segmenti, le rette, i piani e gli angoli. Successivamente vengono asseriti cinque postulati:

i) un segmento di linea retta può essere disegnato tra ogni coppia di punti;
ii) qualsiasi segmento può essere esteso ad una linea retta;
iii) dato un qualsiasi segmento, è possibile disegnare un cerchio avente tale segmento come raggio ed un suo punto estremo come centro;
iv) tutti gli angoli retti sono congruenti;
v) date una linea retta ed un punto non appartenente a tale linea, esiste una ed una sola linea contenente il punto e non intersecante la retta.[1]

Su tali fondamenti Euclide tenta poi di sviluppare in modo puramente deduttivo tutti gli altri risultati, come ad esempio i teoremi di congruenza dei triangoli oppure il teorema di Pitagora, la cui validità viene verificata attraverso una dimostrazione.

Benché in un'ottica moderna gli Elementi non soddisfino tutti i caratteri di rigore e completezza necessari in una teoria, l'impostazione assiomatica scelta da Euclide è alla base di tutta la matematica successivamente sviluppatasi fino ai nostri giorni. Inoltre, gli Elementi hanno costituito un punto di riferimento ed ispirazione per gli studiosi di geometria per più di due millenni. Di fatto la geometria euclidea è stata per secoli l'unica geometria conosciuta e studiata.

É interessante osservare che nel passaggio dal metodo induttivo a quello deduttivo, la geometria perde il suo carattere di scienza sperimentale. Questo implica che non è necessariamente vero che essa

[1] L'ultimo postulato è stato qui enunciato in maniera differente rispetto alla formulazione originaria di Euclide. In questa forma equivalente, esso è noto come postulato delle rette parallele.

debba fornire una adeguata descrizione del mondo fisico. In effetti la fisica moderna, ed in particolare la teoria della relatività, ha dimostrato che non è possibile descrivere la realtà attraverso modelli teorici basati sulla geometria euclidea, come invece accade per la fisica newtoniana. D'altra parte, l'utilizzo del metodo assiomatico ha anche l'effetto di svincolare la geometria dalla sua funzione originaria, ovvero la descrizione del mondo a noi sensibile. Infatti, seguendo tale approccio è assolutamente legittimo modificare i postulati di partenza per poi studiarne le conseguenze utilizzando esclusivamente i principi della logica. Seguendo questo ragionamento vennero formulate nel diciannovesimo secolo le cosiddette geometrie non euclidee, sviluppate tra gli altri da Bolyai, Lobacevskij e Riemann. Analogamente, è lecito cercare di superare la limitazione concernente spazi ed oggetti geometrici al più tridimensionali.

Nella matematica moderna esistono diverse riformulazioni rigorose della geometria di Euclide, la più nota delle quali è dovuta ad Hilbert, così come diverse sue generalizzazioni. Le finalità del corso di Geometria e Algebra Lineare sono:

- costruire un modello per la geometria euclidea. Ciò significa che vogliamo ottenere un insieme in cui sia possibile definire i concetti fondamentali di punti, rette, angoli e distanze e per i quali sia possibile dimostrare rigorosamente la validità dei postulati di Euclide e quindi dei teoremi da questi conseguenti;
- fare in modo che il precedente modello non sia limitato al mondo tridimensionale, ma che sia valido in dimensione qualunque, possibilmente anche infinita;
- dotarci di strumenti teorici e computazionali grazie ai quali sia possibile studiare in modo sistematico ed efficiente le proprietà degli oggetti geometrici che introdurremo.

La realizzazione di questi obiettivi passerà attraverso lo studio approfondito dell'algebra lineare. Insieme all'analisi matematica, questa disciplina costituisce uno dei pilastri della matematica contemporanea, il cui campo di applicazione all'interno della matematica stessa, delle scienze e dell'ingegneria è letteralmente sconfinato.

Questa dispensa è così strutturata. Nel Capitolo 2 introdurremo della terminologia ed alcuni concetti di base che utilizzeremo durante tutto il resto del corso. Il Capitolo 3 è dedicato alla presentazione delle matrici, che dal punto di vista computazionale sono lo strumento fondamentale con cui lavoreremo. Nel Capitolo 4 discuteremo degli spazi vettoriali, che costituiscono la tipologia di insiemi più importante dell'algebra lineare, mentre nei Capitoli 5 e 6 studieremo le funzioni definite tra tali insiemi. Con gli strumenti fino a qui sviluppati, nel Capitolo 7 potremo iniziare a studiare la geometria affine, che può essere pensata come la geometria euclidea senza le nozioni di distanze ed angoli. Per poter discutere di questi ultimi concetti sarà necessario studiare le proprietà dei cosiddetti spazi vettoriali euclidei e delle loro funzioni, nei Capitoli 8 e 9. Quindi applicheremo questo studio a livello geometrico nel Capitolo 10. Nel Capitolo 11 concluderemo il nostro cammino utilizzando tutto quello che abbiamo sviluppato per lo studio delle ipersuperfici quadriche, che sono la naturale generalizzazione in dimensione arbitraria delle curve coniche studiate dagli antichi greci.

CAPITOLO 2

Cenni di Algebra Astratta

2.1. Insiemi, relazioni e funzioni

Il concetto di insieme è fondamentale per tutta la matematica. La trattazione assiomatica rigorosa della teoria degli insiemi esula dagli scopi del corso. In questo note sarà invece utilizzata la cosiddetta teoria ingenua degli insiemi, la cui conoscenza costituisce un prerequisito del corso. Pertanto, un insieme A è intuitivamente definito come una collezione di oggetti distinti $a \in A$, detti elementi dell'insieme. La cardinalità di un insieme, indicata con $|A|$, è il numero di elementi che lo costituiscono.

Esempio 2.1. L'insieme privo di elementi viene detto insieme vuoto ed indicato con $\emptyset = \{\ \}$. La sua cardinalità è $|\emptyset| = 0$. Un insieme $A = \{a_i, 1 \leq i \leq n\}$ costituito da n elementi distinti ha cardinalità $|A| = n$. Gli insiemi numerici sono esempi fondamentali di insiemi di cardinalità infinita. La loro costruzione è argomento del corso di Analisi I, pertanto qui noi ci limitiamo a richiamarli brevemente:

- l'insieme dei numeri naturali $\mathbb{N} = \{0, 1, 2, \dots, n-1, n, n+1, \dots\}$;
- l'insieme dei numeri interi $\mathbb{Z} = \{\dots, -n-1, -n, -n+1, \dots, -2, -1, 0, 1, 2, \dots, n-1, n, n+1, \dots\}$;
- l'insieme dei numeri razionali $\mathbb{Q} = \left\{q = \frac{m}{n} \,|\, m, n \in \mathbb{Z},\ n \neq 0\right\}$;
- l'insieme dei numeri reali $\mathbb{R}$, per il quale esistono diverse definizioni equivalenti (ad esempio, $\mathbb{R}$ può essere definito come l'insieme di tutte le successioni di Cauchy di numeri razionali). Informalmente, $\mathbb{R}$ può essere pensato come l'insieme di tutti i numeri con rappresentazione decimale finita od infinita, periodica o non periodica (ad esempio, sono reali i numeri 1.1, $1.\bar{1}$, $\pi = 3.14\dots$);
- l'insieme dei numeri complessi $\mathbb{C} = \{z = x + iy \,|\, x, y \in \mathbb{R},\ i^2 = -1\}$.

Definizione 2.2. *Dati n insiemi $A_1 = \{a_{11}, \dots, a_{1m_1}\}, \dots, A_n = \{a_{n1}, \dots, a_{nm_n}\}$, il loro prodotto cartesiano è definito come l'insieme delle n-uple ordinate costituite da un elemento di ciascun insieme:*

$$A_1 \times \dots \times A_n = \{(a_{1i_1}, \dots, a_{ni_n}) | 1 \leq i_1 \leq m_1, \dots, 1 \leq i_n \leq m_n\}.$$

Se tutti gli insiemi di partenza sono uguali $A_1 = \dots = A_n = A$, il loro prodotto cartesiano si denota con A^n. Una relazione n-aria R tra $A_1, \dots, A_n$ è un sottoinsieme di $A_1 \times \dots \times A_n$.

In questo corso tratteremo esclusivamente relazioni binarie tra due insiemi A, B. In tal caso, una relazione R è un sottoinsieme di $A \times B$ formato da coppie (a_i, b_j) composte da un elemento $a_i \in A$ ed un elemento $b_j \in B$. Se $(a_i, b_j) \in R$ si indica $a_i \sim_R b_j$.

Esempio 2.3. Consideriamo due insiemi $A = \{a_1, a_2\}$ e $B = \{b_1, b_2\}$ di cardinalità $m = n = 2$. Allora $A \times B = \{(a_1, b_1), (a_1, b_2), (a_2, b_1), (a_2, b_2)\}$ ha cardinalità $mn = 4$. L'insieme delle parti di $A \times B$ è composto da $2^{mn} = 16$ elementi, che sono esattamente tutti i possibili sottoinsiemi di $A \times B$, ovvero tutte le possibili relazioni tra A e B. Le elenchiamo tutte:

$$\begin{aligned}
&R_0 = \emptyset \\
&R_1 = \{(a_1, b_1)\},\ R_2 = \{(a_1, b_2)\},\ R_3 = \{(a_2, b_1)\},\ R_4 = \{(a_2, b_2)\}, \\
&R_5 = \{(a_1, b_1), (a_1, b_2)\},\ R_6 = \{(a_1, b_1), (a_2, b_1)\},\ R_7 = \{(a_1, b_1), (a_2, b_2)\}, \\
&R_8 = \{(a_1, b_2), (a_2, b_1)\},\ R_9 = \{(a_1, b_2), (a_2, b_2)\},\ R_{10} = \{(a_2, b_1), (a_2, b_2)\}, \\
&R_{11} = \{(a_1, b_1), (a_1, b_2), (a_2, b_1)\},\ R_{12} = \{(a_1, b_1), (a_1, b_2), (a_2, b_2)\}, \\
&R_{13} = \{(a_1, b_1), (a_2, b_1), (a_2, b_2)\},\ R_{14} = \{(a_1, b_2), (a_2, b_1), (a_2, b_2)\}, \\
&R_{15} = \{(a_1, b_1), (a_1, b_2), (a_2, b_1), (a_2, b_2)\} = A \times B.
\end{aligned}$$

Una classe speciale di relazioni sono le funzioni, che informalmente possono essere pensate come delle regole che mettono in relazione ogni elemento di A ad un unico elemento di B.

Definizione 2.4. *Dati gli insiemi A e B, una funzione f da A in B è una relazione tra i due insiemi tale che ogni elemento di A compare come componente di un unico elemento di f.*

Esempio 2.5. Dall'insieme $A = \{a_1, a_2\}$ all'insieme $B = \{b_1, b_2\}$ esistono esattamente $2^2 = 4$ funzioni. Le notazioni standard per le funzioni sono:

$$f_1 = R_6 \qquad \text{che esprimiamo come} \qquad \begin{array}{rccc} f_1: & A & \to & B \\ & a_1 & \mapsto & b_1 \\ & a_2 & \mapsto & b_1 \end{array} \qquad \text{oppure} \qquad f_1(a_1) = b_1,\ f_1(a_2) = b_1;$$

$$f_2 = R_7 \qquad \text{che esprimiamo come} \qquad \begin{array}{rccc} f_2: & A & \to & B \\ & a_1 & \mapsto & b_1 \\ & a_2 & \mapsto & b_2 \end{array} \qquad \text{oppure} \qquad f_2(a_1) = b_1,\ f_2(a_2) = b_2;$$

$$f_3 = R_8 \qquad \text{che esprimiamo come} \qquad \begin{array}{rccc} f_3: & A & \to & B \\ & a_1 & \mapsto & b_2 \\ & a_2 & \mapsto & b_1 \end{array} \qquad \text{oppure} \qquad f_3(a_1) = b_2,\ f_3(a_2) = b_1;$$

$$f_4 = R_9 \qquad \text{che esprimiamo come} \qquad \begin{array}{rccc} f_4: & A & \to & B \\ & a_1 & \mapsto & b_2 \\ & a_2 & \mapsto & b_2 \end{array} \qquad \text{oppure} \qquad f_4(a_1) = b_2,\ f_4(a_2) = b_2;$$

In generale si chiama B^A l'insieme di tutte le funzioni da A in B. Se i due insiemi hanno cardinalità finita, si può dimostrare che $|B^A| = |B|^{|A|}$.

Definizione 2.6. *Data $f : A \to B$, allora:*

i) *A è detto dominio e B è detto codominio di f;*

ii) *dato $a \in A$, l'elemento $b = f(a) \in B$ è detto immagine di a tramite la funzione f;*

iii) *l'insieme immagine di f è $Im(f) = \{b \in B \mid b = f(a)$ per un qualche $a \in A\} \subseteq B$;*

iv) *dato $b \in Im(f)$, la sua controimmagine è $f^{-1}(b) = \{a \in A \mid f(a) = b\} \subseteq A$.*

Definizione 2.7. *Data $f : A \to B$, allora:*

i) *f è iniettiva se $|f^{-1}(b)| = 1$ per ogni $b \in Im(f)$;*

ii) *f è suriettiva se $Im(f) = B$;*

iii) *f è biietttiva, o biunivoca, se è iniettiva e suriettiva.*

Esempio 2.8. Studiamo le proprietà delle funzioni dell'Esempio 2.5.

- $\mathrm{Im}(f_1) = \{b_1\} \subset B,\ f_1^{-1}(b_1) = A,\ f_1^{-1}(b_2) = \emptyset$. Quindi f_1 non è né invertibile né suriettiva.
- $\mathrm{Im}(f_2) = B,\ f_2^{-1}(b_1) = \{a_1\},\ f_2^{-1}(b_2) = \{a_2\}$. Quindi f_2 è biunivoca.
- $\mathrm{Im}(f_3) = B,\ f_3^{-1}(b_1) = \{a_2\},\ f_3^{-1}(b_2) = \{a_1\}$. Quindi f_3 è biunivoca.
- $\mathrm{Im}(f_4) = \{b_2\},\ f_4^{-1}(b_2) = \emptyset,\ f_4^{-1}(b_2) = A$. Quindi f_4 non è né invertibile né suriettiva.

Concludiamo la sezione richiamando le ultime importanti nozioni sulle funzioni. Date $f : A \to B$ e $g : B \to C$, la composizione delle due funzioni è la funzione $g \circ f : A \to C$ che ad ogni elemento $a \in A$ associa $(g \circ f)(a) = g(f(a)) \in C$. Un caso particolare di composizione si ha con la funzione inversa. Essa è la funzione $f^{-1} : B \to A$ tale che $f^{-1}(f(a)) = a$ e $f(f^{-1}(b)) = b$ per ogni $a \in A$ e $b \in B$. Si può dimostrare che f^{-1} esiste, ed in tal caso è unica, se e solo se f è biunivoca. In questa situazione, f viene anche detta invertibile.

Infine, la restrizione di f ad $U \subseteq A$ è la funzione

$$\begin{array}{rccc} f_U : & U & \to & B \\ & u & \mapsto & f(u) \end{array}.$$

2.2. Operazioni, strutture algebriche ed omomorfismi

Le operazioni sono particolari tipi di funzioni.

Definizione 2.9. *Dati $n+1$ insiemi $A_1, \dots, A_{n+1}$ non vuoti, un'operazione n-aria $*$ è una funzione*

$$\begin{array}{rccc} * : & A_1 \times \cdots \times A_n & \to & A_{n+1} \\ & (a_1, \dots, a_n) & \mapsto & *(a_1, \dots, a_n) \end{array}.$$

Se $A_1 = \cdots = A_{n+1}$ l'operazione viene detta interna, altrimenti si dice esterna. Se $n = 2$ l'operazione si dice binaria e si indica $*(a_1, a_2) = a_1 * a_2$.

Esempio 2.10. Sull'insieme dei numeri naturali $\mathbb{N}$ è definita la somma. Questa è un'operazione binaria interna. Infatti:

$$\begin{array}{rccc} + : & \mathbb{N} \times \mathbb{N} & \to & \mathbb{N} \\ & (m_1, m_2) & \mapsto & m_1 + m_2 \end{array}.$$

La differenza di due numeri naturali è anch'essa un'operazione binaria, ma non è interna in quanto la differenza di due naturali in generale è un intero:

$$\begin{array}{rccc} - : & \mathbb{N} \times \mathbb{N} & \to & \mathbb{Z} \\ & (m_1, m_2) & \mapsto & m_1 - m_2 \end{array}.$$

Una funzione tra due insiemi può essere interpretata come una operazione unaria. È interna se il dominio ed il codominio coincidono. Un esempio notevole è il seguente.

Definizione 2.11. *Dato un insieme A, chiamamo insieme diagonale di A*

$$\Delta_A = \{(a, a) \,|\, a \in A\} \subseteq A \times A.$$

Come relazione, Δ_A definisce una funzione detta identità. Essa associa ad ogni elemento di A l'elemento stesso:

$$\begin{array}{rccc} \mathrm{Id}_A : & A & \to & A \\ & a & \mapsto & a \end{array}.$$

Esempio 2.12. Sia $A = \{a_1, a_2\}$. Esistono esattamente $2^{2^2} = 16$ differenti operazioni binarie interne ad A. Rappresentiamo le operazioni attraverso delle tabelle, in cui nel posto di riga i e colonna j inseriamo il valore di $a_i * a_j$:

$*_1$	a_1	a_2
a_1	a_1	a_1
a_2	a_1	a_1

$*_2$	a_1	a_2
a_1	a_2	a_1
a_2	a_1	a_1

$*_3$	a_1	a_2
a_1	a_1	a_2
a_2	a_1	a_1

$*_4$	a_1	a_2
a_1	a_1	a_1
a_2	a_2	a_1

$*_5$	a_1	a_2
a_1	a_1	a_1
a_2	a_1	a_2

$*_6$	a_1	a_2
a_1	a_2	a_2
a_2	a_1	a_1

$*_7$	a_1	a_2
a_1	a_2	a_1
a_2	a_2	a_1

$*_8$	a_1	a_2
a_1	a_2	a_1
a_2	a_1	a_2

$*_9$	a_1	a_2
a_1	a_1	a_2
a_2	a_2	a_1

$*_{10}$	a_1	a_2
a_1	a_1	a_2
a_2	a_1	a_2

$*_{11}$	a_1	a_2
a_1	a_1	a_1
a_2	a_2	a_2

$*_{12}$	a_1	a_2
a_1	a_2	a_2
a_2	a_2	a_1

$*_{13}$	a_1	a_2
a_1	a_2	a_2
a_2	a_1	a_2

$*_{14}$	a_1	a_2
a_1	a_2	a_1
a_2	a_2	a_2

$*_{15}$	a_1	a_2
a_1	a_1	a_2
a_2	a_2	a_2

$*_{16}$	a_1	a_2
a_1	a_2	a_2
a_2	a_2	a_2

Definizione 2.13. *Una struttura algebrica è costituita da una collezione di insiemi* $A_1, \dots, A_n$*, detti supporto della struttura, ed una collezione di operazioni* $*_1, \dots, *_m$ *su questi insiemi. Si indica con* $(A_1, \dots, A_n, *_1, \dots, *_m)$*. Un omomorfismo è una funzione* f *tra due strutture algebriche che preserva le loro operazioni. Se tale funzione è biunivoca e la funzione inversa* f^{-1} *è a sua volta un omomorfismo, allora* f *si dice isomorfismo e le due strutture si dicono isomorfe.*

Lo studio sistematico delle strutture algebriche è l'argomento principale dell'Algebra. In particolare, l'Algebra Lineare è la disciplina che si occupa di investigare le proprietà delle strutture note con il nome di spazi vettoriali. Per poter arrivare a definirli è però necessario presentare preliminarmente un paio di altre strutture algebriche, ovvero i gruppi ed i campi, che ricorreranno spesso nella nostra trattazione.

Definizione 2.14. *Sia* G *un insieme su cui è definita un'operazione interna* $*$ *soddisfacente le seguenti tre proprietà:*

i) esiste un elemento neutro $e \in G$ *tale che* $e * a = a * e = a$ *per ogni* $a \in A$;

ii) per ogni $a \in G$ *esiste l'elemento inverso* $a^{-1} \in G$ *tale che* $a * a^{-1} = a^{-1} * a = e$;

iii) vale la proprietà associativa $(a * b) * c = a * (b * c)$ *per ogni* $a, b, c \in G$.

Allora la struttura $(G, *)$ *si dice gruppo. Se inoltre*

iv) vale la proprietà commutativa $a * b = b * a$ *per ogni* $a, b \in G$,

allora $(G, *)$ *è detto gruppo commutativo o abeliano.*

Esempio 2.15. Consideriamo l'insieme $A = \{a_1, a_2\}$ e verifichiamo quali delle proprietà di gruppo sono soddisfatte dalle operazioni dell'Esempio 2.12:

- $*_3, *_4, *_6, *_7, *_{13}, *_{14}$ non soddisfano alcuna proprietà;
- $*_{10}, *_{11}$ soddisfano solo la proprietà iii;
- $*_2, *_{12}$ soddisfano solo la proprietà iv;
- $*_1, *_{16}$ soddisfano solo le proprietà iii e iv;
- $*_5$ soddisfa la proprietà i, con elemento neutro $e = a_2$, e le proprietà iii e iv, ma non esiste l'inverso di a_1;
- $*_{15}$ soddisfa la proprietà i, con elemento neutro $e = a_1$, e le proprietà iii e iv, ma non esiste l'inverso di a_2;
- $*_8$ soddisfa tutte le proprietà, con elemento neutro $e = a_2$ ed inverso $a_1^{-1} = a_1$;
- $*_9$ soddisfa tutte le proprietà, con elemento neutro $e = a_1$ ed inverso $a_2^{-1} = a_2$.

Quindi $(A, *_8)$ e $(A, *_9)$ sono gruppi abeliani. Osserviamo che questi due gruppi sono identici se scambiamo i nomi di a_1 ed a_2.

Esempio 2.16. Per gli insiemi numerici abbiamo:

- $(\mathbb{N}, +)$ e $(\mathbb{N}, \cdot)$ non sono gruppi, perché in entrambi i casi non vale la proprietà ii;
- $(\mathbb{Z}, +)$ è un gruppo abeliano, $(\mathbb{Z}, \cdot)$ non è un gruppo perché non vale ii;
- $(\mathbb{Q}, +)$ e $(\mathbb{Q} \setminus \{0\}, \cdot)$ sono gruppi abeliani (analogo per $\mathbb{R}$ e $\mathbb{C}$).

Definizione 2.17. *Dati due gruppi $(A, *_A)$ e $(B, *_B)$, un omomorfismo di gruppi è una funzione $\phi : A \to B$ tale che $\phi(a_1 *_A a_2) = \phi(a_1) *_B \phi(a_2)$ per ogni $a_1, a_2 \in A$.*

Esempio 2.18.

- Dato $A = \{a_1, a_2\}$, la funzione $\phi : A \to A$ definita come $\phi(a_1) = a_2$ e $\phi(a_2) = a_1$ è un isomorfismo di gruppi da $(A, *_8)$ ad $(A, *_9)$ e viceversa. I due gruppi si dicono isomorfi. Per quanto visto nell'Esempio 2.15, possiamo affermare che, a meno di isomorfismi, esiste un unico gruppo di cardinalità 2.
- La funzione
$$\begin{array}{rccc} f: & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & e^x \end{array}$$
soddisfa la proprietà
$$f(x+y) = e^{x+y} = e^x \cdot e^y = f(x) \cdot f(y).$$
Quindi f è un omomorfismo di gruppi tra $(\mathbb{R}, +)$ ed $(\mathbb{R} \setminus \{0\}, \cdot)$. Non è un isomorfismo in quanto f non è suriettiva. Per avere un isomorfismo, dobbiamo restringere il codominio di f alla sua immagine. Ovvero, definito $\mathbb{R}^+ = \{x \in \mathbb{R} \,|\, x > 0\}$, la funzione
$$\begin{array}{rccc} \bar{f}: & \mathbb{R} & \to & \mathbb{R}^+ \\ & x & \mapsto & e^x \end{array}$$
è un isomorfismo di gruppi tra $(\mathbb{R}, +)$ ed $(\mathbb{R}^+, \cdot)$.

Definizione 2.19. *Sia $\mathbb{K}$ un insieme su cui sono definite due operazioni interne $*, \circ$ tali che:*

i) $(\mathbb{K}, *)$ *è un gruppo abeliano con elemento neutro e;*
ii) definito $\mathbb{K}^* = \mathbb{K} \setminus \{e\}$, allora $(\mathbb{K}^*, \circ)$ è un gruppo abeliano;
iii) vale la proprietà distributiva $a \circ (b * c) = (a \circ b) * (a \circ c)$ per ogni $a, b, c \in \mathbb{K}$.

Allora la struttura $(\mathbb{K}, *, \circ)$ si dice campo.

Convenzionalmente si è soliti chiamare somma la prima operazione in $\mathbb{K}$, indicata con $* = +$ ed il cui elemento neutro è 0, mentre la seconda operazione si dice moltiplicazione, indicata con $\circ = \cdot$ ed il cui elemento neutro è 1. Si dice che $(\mathbb{K}, +)$ e $(\mathbb{K}^*, \cdot)$ costituiscono rispettivamente la struttura additiva e la struttura moltiplicativa del campo.

Esempio 2.20. $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$ sono campi rispetto alle usuali operazioni. L'insieme $A = \{a_1, a_2\}$ con le operazioni $*, \circ$ indicate nelle tabelle è un esempio di campo composto da un numero finito di elementi. Esso viene denotato con $\mathbb{Z}_2$.

$*$	a_1	a_2
a_1	a_1	a_2
a_2	a_2	a_1

$\circ$	a_1	a_2
a_1	a_1	a_1
a_2	a_1	a_2

Definizione 2.21. *Dati due campi* $(A, *_A, \circ_A)$ *e* $(B, *_B, \circ_B)$, *un omomorfismo di campi è una funzione* $\phi : A \to B$ *tale che* $\phi(a_1 *_A a_2) = \phi(a_1) *_B \phi(a_2)$ *e* $\phi(a_1 \circ_A a_2) = \phi(a_1) \circ_B \phi(a_2)$ *per ogni* $a_1, a_2 \in A$.

Esempio 2.22. Fissato $a \in \mathbb{K}$, consideriamo la funzione

$$\begin{array}{rccc} f_a : & \mathbb{K} & \to & \mathbb{K} \\ & x & \mapsto & ax \end{array}.$$

Allora

$$f_a(x+y) = a(x+y) = ax + ay = f_a(x) + f_a(y)$$

per ogni $a \in \mathbb{K}$, mentre

$$f_a(x \cdot y) = a(x \cdot y) = ax \cdot ay = f_a(x) \cdot f_a(y)$$

è vera se e solo se $a = 0, 1$. Quindi f_a è sempre un omomorfismo di gruppi da $(\mathbb{K}, +)$ in se stesso, ma è un omomorfismo di campi da $(\mathbb{K}, +, \cdot)$ in se stesso se e solo se $a = 0, 1$. Chiaramente, $f_1 = \mathrm{Id}_{\mathbb{K}}$ è un isomorfismo di campi. Invece f_0 è detto funzione nulla oppure omomorfismo nullo, ed ovviamente non è invertibile.

2.3. Polinomi

Vista la loro importanza nell'Algebra, in questa sezione introduciamo i polinomi a coefficienti in un campo $\mathbb{K}$ e riassumiamo brevemente le loro proprietà fondamentali.

Definizione 2.23. *Un polinomio nella variabile* x *e con coefficienti* $(a_0, \ldots, a_n) \in \mathbb{K}^{n+1}$ *ed* $n \in \mathbb{N}$, *è l'espressione*

$$P(x) = \sum_{i=0}^{n} a_i x^i.$$

L'insieme di tutti i polinomi nella variabile x *si indica con* $\mathbb{K}[x]$.

Un polinomio nelle m *variabili* $x_1, \ldots, x_m$ *è definito induttivamente come l'espressione*

$$P(x_1, \ldots, x_m) = \sum_{i=0}^{n} Q_i(x_1, \ldots, x_{m-1})\, x_m^i,$$

dove $Q_1, \ldots, Q_n$ *sono polinomi nelle prime* $m-1$ *variabili. L'insieme di tutti i polinomi di questo tipo si indica con* $\mathbb{K}[x_1, \ldots, x_m]$.

Ogni polinomio è scrivibile in modo unico come somma di monomi, ovvero termini del tipo $a_{i_1 \ldots i_m} x_1^{i_1} \ldots x_m^{i_m}$. Il grado di tale monomio è definito come $i_1 + \cdots + i_m$. Il grado del polinomio è il massimo grado dei suoi monomi aventi coefficiente $a_{i_1 \ldots i_m} \neq 0$. Il polinomio nullo è quello i cui coefficienti sono tutti zero, e per definizione ha grado indeterminato. Esiste una stretta relazione tra polinomio ed associata funzione polinomiale. Quest'ultimo oggetto si ottiene con la valutazione del polinomio sugli elementi di un insieme. Ad esempio, la funzione polinomiale $f \in \mathbb{K}^{\mathbb{K}}$ associata a $P \in \mathbb{K}[x]$ è

$$\begin{array}{rccc} f : & \mathbb{K} & \to & \mathbb{K} \\ & a & \mapsto & P(a) \end{array}.$$

Commettendo un abuso di notazione, è usuale identificare il polinomio con la funzione polinomiale associata sul corrispondente campo. A meno di ambiguità, in questo note adotteremo anche noi tale pratica.

Ora concentriamoci sui polinomi in una variabile. In tal caso esiste un unico monomio per ogni grado. Il polinomio si dice monico se il monomio di grado massimo ha coefficiente 1. È considerato prerequisito del corso la conoscenza dell'algoritmo euclideo per la divisione di polinomi in una variabile. Il seguente risultato è conseguenza di tale algoritmo.

Lemma 2.24. *Consideriamo una coppia* $(A, B) \in \mathbb{K}[x] \times \mathbb{K}[x]$ *soddisfacente* $B \neq 0$. *Allora esiste un'unica coppia* $(Q, R) \in \mathbb{K}[x] \times \mathbb{K}[x]$ *tale per cui* $A = BQ + R$, *soddisfacente la condizione* $\text{grado}(R) < \text{grado}(B)$ *oppure* $R = 0$.

I polinomi Q ed R vengono detti rispettivamente quoziente e resto della divisione di A per B. Se $R = 0$ si dice che B divide A.

Esempio 2.25. Definiamo $A(x) = 1 - x^2$ e scegliamo $B_0(x) = 1,\ B_1(x) = 1 - x$ e $B_2(x) = 1 + x^3$. Studiamo la divisione $A = B_iQ_i + R_i$, per $0 \leq i \leq 2$. Allora:

- $1 - x^2 = 1\,(1 - x^2)$ e quindi $Q_0(x) = A(x)$ ed $R_0(x) = 0$;
- $1 - x^2 = (1 - x)(1 + x)$ e quindi $Q_1(x) = 1 + x$ ed $R_1(x) = 0$;
- $1 - x^2 = (1 + x^3)\,0 + (1 - x^2)$ e quindi $Q_2(x) = 0$ ed $R_2(x) = A(x)$.

Lemma 2.26. *Dati* $P \in \mathbb{K}[x]$ *ed* $r \in \mathbb{K}$, *esiste un numero naturale massimo* $m \leq \text{grado}(P)$ *tale che* $(x - r)^m$ *divide* P.

Dimostrazione. Sicuramente $(x - r)^0 = 1$ divide il polinomio, pertanto $m \in \mathbb{N}$ esiste. Inoltre, se $(x - r)^m$ divide P significa che $P(x) = (x - r)^m Q(x)$. Allora vale $\text{grado}(P) = m + \text{grado}(Q)$, da cui $m \leq \text{grado}(P)$. □

Utilizzando le notazioni dell'ultimo lemma, diamo la seguente definizione.

Definizione 2.27. *Il numero* m *viene detto molteplicità algebrica di* r *rispetto a* P. *Se* $m \geq 1$, *allora* r *si dice radice di* P. *La radice si dice semplice se* $m = 1$.

Proposizione 2.28. *L'elemento* $r \in \mathbb{K}$ *è radice di* $P \in \mathbb{K}[x]$ *se e solo se è uno zero della funzione polinomiale associata* $f \in \mathbb{K}^{\mathbb{K}}$.

Dimostrazione. Per definizione, r è zero di f se $f(r) = P(r) = 0$. Dobbiamo provare che questo è possibile se e solo se $(x - r)$ divide il polinomio P. Osserviamo che il Lemma 2.24 garantisce l'esistenza di $Q(x) \in \mathbb{K}[x]$ ed $R(x) = a_0$ tali che $P(x) = (x - r)Q(x) + a_0$. Allora $f(r) = (r - r)Q(r) + a_0 = 0$ se e solo se $a_0 = 0$, il che completa la dimostrazione. □

Esempio 2.29. Dato $P(x) = 1 - x^2$, vale $P(x) = -(x-1)^1(x+1)^1$. Pertanto esistono solo due radici del polinomio, ovvero $r_1 = 1$ ed $r_2 = -1$, entrambe semplici. È semplice verificare che $P(1) = P(-1) = 0$.

Osserviamo che molti autori invertono i ruoli della Definizione 2.27 e della Proposizione 2.28, usando quest'ultima come definizione di radice di un polinomio. Naturalmente, a causa del contenuto della Proposizione, le due cose sono del tutto equivalenti.

Concludiamo la nostra veloce presentazione dei polinomi con due importanti risultati, che ci torneranno utili in particolare nei Capitoli 6 e 9. Le loro dimostrazioni sono più complicate rispetto a quelle dei teoremi visti finora ed esulano dagli scopi del corso.

Teorema 2.30. *Se esistono, le radici* $r_1, \ldots, r_k \in \mathbb{K}$ *di un polinomio* $P \in \mathbb{K}[x]$ *di grado* n *soddisfano* $m_1 + \cdots + m_k \leq n$, *dove* m_i *è la molteplicità algebrica di* r_i *per* $i = 1, \ldots, k$. *Per ogni campo* $\mathbb{K}$ *esiste un altro campo* $\overline{\mathbb{K}}$ *che lo contiene, tale che ogni polinomio* $\overline{P} \in \overline{\mathbb{K}}[x]$ *di grado* n *ha le radici soddisfacenti* $m_1 + \cdots + m_k = n$.

La dimostrazione di questo teorema si può trovare in un buon libro di Algebra. Si dice che nel campo $\overline{\mathbb{K}}$ esistono esattamente n radici di $\overline{P}$ contate con la loro molteplicità. In particolare, il teorema implica che ogni $P \in \mathbb{K}[x] \subseteq \overline{\mathbb{K}}[x]$ ha esattamente n radici nel campo $\overline{\mathbb{K}}$. Si può inoltre dimostrare che esiste un campo $\overline{\mathbb{K}}$ minimo rispetto all'inclusione, e che tale campo è unico a meno di isomorfismi di campi. Usando queste notazioni, possiamo dare la seguente definizione.

Definizione 2.31. *Il campo $\overline{\mathbb{K}}$ si dice chiusura algebrica di $\mathbb{K}$. Un campo che coincida con la sua chiusura algebrica si dice algebricamente chiuso.*

Infine, il seguente risultato, noto come Teorema fondamentale dell'Algebra, è dovuto a Gauss e la sua dimostrazione richiede tecniche di Analisi Complessa.

Teorema 2.32. *Il campo $\mathbb{C}$ dei numeri complessi è algebricamente chiuso. Inoltre, $\mathbb{C}$ è la chiusura algebrica di $\mathbb{R}$ e contiene la chiusura algebrica di $\mathbb{Q}$.*

Esempio 2.33. Consideriamo il polinomio $P(x) = 1 + x^2 \in \mathbb{Q}[x]$. Ovviamente non esistono radici di P in $\mathbb{Q}$ od $\mathbb{R}$. Ma se pensiamo al polinomio come ad un elemento di $\mathbb{C}[x]$, allora esistono $r_1 = i$ ed $r_2 = -i$. Osserviamo che il polinomio si può quindi fattorizzare nel prodotto di due binomi di primo grado, cioè $P(x) = (x - i)(x + i)$.

L'ultima osservazione dell'esempio precedente è del tutto generalizzabile, essendo una conseguenza diretta della definizione di radice di un polinomio e di campo algebricamente chiuso.

Corollario 2.34. *Sia $P \in \mathbb{K}[x]$ dove $\mathbb{K}$ è un campo algebricamente chiuso. Siano inoltre $r_1, \ldots, r_k$ le sue radici, aventi molteplicità $m_1, \ldots, m_k$, ed a_n il coefficiente del monomio di grado massimo di P. Allora $P(x) = a_n(x - r_1)^{m_1} \ldots (x - r_k)^{m_k}$.*

2.4. Relazioni di equivalenza ed insieme quoziente

Una seconda classe importante di relazioni sono le cosiddette relazioni di equivalenza.

Definizione 2.35. *Una relazione R da A in A è detta di equivalenza se:*

i) è riflessiva, ovvero $(a, a) \in R$ per ogni $a \in A$;
ii) è simmetrica, ovvero se $(a_1, a_2) \in R$ allora $(a_2, a_1) \in R$;
iii) è transitiva, ovvero se $(a_1, a_2), (a_2, a_3) \in R$ allora $(a_1, a_3) \in R$.

Esempio 2.36. Consideriamo l'insieme $A = \{a_1, a_2\}$ e le relazioni da A in A:

$$
\begin{aligned}
&R_0 = \emptyset \\
&R_1 = \{(a_1, a_1)\},\ R_2 = \{(a_1, a_2)\},\ R_3 = \{(a_2, a_1)\},\ R_4 = \{(a_2, a_2)\}, \\
&R_5 = \{(a_1, a_1), (a_1, a_2)\},\ R_6 = \{(a_1, a_1), (a_2, a_1)\},\ R_7 = \{(a_1, a_1), (a_2, a_2)\} = \Delta_A, \\
&R_8 = \{(a_1, a_2), (a_2, a_1)\},\ R_9 = \{(a_1, a_2), (a_2, a_2)\},\ R_{10} = \{(a_2, a_1), (a_2, a_2)\}, \\
&R_{11} = \{(a_1, a_1), (a_1, a_2), (a_2, a_1)\},\ R_{12} = \{(a_1, a_1), (a_1, a_2), (a_2, a_2)\}, \\
&R_{13} = \{(a_1, a_1), (a_2, a_1), (a_2, a_2)\},\ R_{14} = \{(a_1, a_2), (a_2, a_1), (a_2, a_2)\}, \\
&R_{15} = \{(a_1, a_1), (a_1, a_2), (a_2, a_1), (a_2, a_2)\} = A \times A.
\end{aligned}
$$

Lasciamo come esercizio al lettore la verifica delle seguenti affermazioni:

- R_8, R_{11}, R_{14} soddisfano solo la proprietà ii;
- $R_2, R_3, R_5, R_6, R_9, R_{10}$ soddisfano solo la proprietà iii;
- R_{12}, R_{13} soddisfano solo le proprietà i e iii;

- R_0, R_1, R_4 soddisfano solo le proprietà ii e iii;
- R_7, R_{15} soddisfano tutte le proprietà e quindi sono relazioni di equivalenza su A. Osserviamo che $a_1 \nsim_{R_7} a_2$ mentre $a_1 \sim_{R_{15}} a_2$.

È facile dimostrare che per qualsiasi insieme A, le relazioni Δ_A e $A \times A$ sono sempre di equivalenza e vengono dette relazioni di equivalenza banali. Δ_A coincide con la relazione di uguaglianza, ovvero $a_1 \sim_{\Delta_A} a_2$ se e solo se $a_1 = a_2$. Invece, secondo la relazione $A \times A$, tutti gli elementi di A sono tra loro equivalenti. Una generica relazione di equivalenza R soddisfa $\Delta_A \subseteq R \subseteq A \times A$, cioè si trova a metà strada tra i due casi banali e può essere quindi considerata come una generalizzazione del concetto di uguaglianza.

Definizione 2.37. *Siano A un insieme ed R una relazione di equivalenza su A. La classe di equivalenza di $a \in A$ è il sottoinsieme di A definito come*

$$[a]_R = \{b \in A \,|\, b \sim_R a\}.$$

Teorema 2.38. *Sia A un insieme ed R una relazione di equivalenza su A. Allora ogni elemento $a \in A$ appartiene ad un'unica classe di equivalenza.*

Dimostrazione. L'esistenza della classe contenente a è conseguenza della riflessività di R, che implica $a \in [a]_R$. Per l'unicità, se ipotizziamo che $a \in [b]_R$ per un qualche $b \in A$ diverso da a, allora è necessario dimostrare che $[b]_R = [a]_R$. Sia c un generico elemento di $[a]_R$. Sfruttando la transitività di R abbiamo:

$$c \sim_R a \quad \text{ed} \quad a \sim_R b \quad \Rightarrow \quad c \sim_R b \quad \Rightarrow \quad c \in [b]_R$$

e di conseguenza $[a]_R \subseteq [b]_R$. Il viceversa si dimostra in maniera analoga, sfruttando anche la simmetria di R. In dettaglio, supponiamo $c \in [b]_R$. Allora:

$$c \sim_R b \quad \text{ed} \quad a \sim_R b \quad \Rightarrow \quad c \sim_R b \quad \text{e} \quad b \sim_R a \quad \Rightarrow \quad c \sim_R a \quad \Rightarrow \quad c \in [a]_R$$

e di conseguenza $[b]_R \subseteq [a]_R$. La doppia inclusione implica la tesi. □

Un corollario immediato del teorema è che l'insieme A è l'unione di tutte le sue classi di equivalenza. Poiché l'intersezione tra due classi distinte è l'insieme vuoto, si dice che la loro unione è un'unione disgiunta.

Definizione 2.39. *Sia A un insieme ed R una relazione di equivalenza su A. L'insieme delle classi di equivalenza rispetto ad R contenute in A è detto insieme quoziente A/R. Inoltre, si chiama proiezione naturale la funzione*

$$\begin{array}{rccc} \pi_R : & A & \to & A/R \\ & a & \mapsto & [a]_R \end{array}.$$

Esempio 2.40. Nell'Esempio 2.36 abbiamo due possibili relazioni di equivalenza.

- $[a_1]_{R_7} = \{a_1\}$, $[a_2]_{R_7} = \{a_2\}$. Di conseguenza $[a_1]_{R_7} \cap [a_2]_{R_7} = \emptyset$ e $[a_1]_{R_7} \cup [a_2]_{R_7} = A$. Infine, $A/R_7 = \{[a_1]_{R_7}, [a_2]_{R_7}\}$.
- $[a_1]_{R_{15}} = [a_2]_{R_{15}} = \{a_1, a_2\} = A$. Quindi, $A/R_{15} = \{[a_1]_{R_{15}}\}$.

È importante osservare che gli elementi dell'insieme quoziente sono a loro volta degli insiemi.

Normalmente, se non vi sono ambiguità sulla relazione di equivalenza da considerare, si è soliti omettere il pedice R nei simboli $\sim_R$, $[\ ]_R$ e π_R.

2.5. Relazione d'ordine

L'ultimo fondamentale tipo di relazione che consideriamo in queste note è la relazione d'ordine.

Definizione 2.41. *Una relazione R da A in A è detta d'ordine se:*

i) è riflessiva, ovvero $(a, a) \in R$ per ogni $a \in A$;

ii) è antisimmetrica, ovvero se $(a_1, a_2), (a_2, a_1) \in R$ allora $a_1 = a_2$;

iii) è transitiva, ovvero se $(a_1, a_2), (a_2, a_3) \in R$ allora $(a_1, a_3) \in R$.

Esempio 2.42. In $\mathbb{N}$, consideriamo la relazione

$$R = \{(m, n) \in \mathbb{N}^2 \mid (n - m) \in \mathbb{N}\}.$$

Essa è una relazione d'ordine. Infatti

- $(m, m) \in R$, poiché $m - m = 0 \in \mathbb{N}$, quindi R è riflessiva;
- se $(m, n), (n, m) \in R$, allora $n - m$ e $m - n$ sono entrambi numeri naturali, ma questo è possibile se e solo se $m = n$. Pertanto, R è antisimmetrica;
- se $(m, n), (n, p) \in R$, allora $p - m = (p - n) + (n - m) \in \mathbb{N}$. Quindi, $(m, p) \in R$ e perciò la relazione è transitiva.

La relazione R appena definita si chiama relazione di minore od uguale, che si rappresenta con il simbolo $\leq$. Cioè, se vale $m \sim_R n$ allora si scrive $m \leq n$. Si può definire in maniera analoga una relazione di minore od uguale su $\mathbb{Z}$, $\mathbb{Q}$ ed $\mathbb{R}$, ma non su $\mathbb{C}$.

Esempio 2.43. In $\mathbb{N}^*$, consideriamo la relazione

$$R = \{(m, n) \in \mathbb{N}^{*2} \mid n = mp \text{ con } p \in \mathbb{N}^*\}.$$

Anche questa è una relazione d'ordine. Infatti

- $(m, m) \in R$ poiché $m = m1$, quindi R è riflessiva;
- se $(m, n), (n, m) \in R$, allora $n = mp$ ed $m = nq$. Ricaviamo quindi $n = npq$, da cui segue $pq = 1$. L'unica soluzione di questa equazione in $\mathbb{N}^*$ è $p = q = 1$, che implica $m = n$. Pertanto, R è antisimmetrica;
- se $(m, n), (n, p) \in R$, allora $p = nq = mrq \in \mathbb{N}^*$. Quindi, $(m, p) \in R$ e perciò la relazione è transitiva.

La relazione R appena definita si chiama relazione di divisibilità, che si rappresenta con il simbolo $|$. Ovvero, se vale $m \sim_R n$ allora si scrive $m|n$.

Esempio 2.44. Tra le relazioni individuate nell'Esempio 2.36, lasciamo al lettore verificare che solo R_7, R_{12} ed R_{13} sono relazioni d'ordine.

Osserviamo che nell'Esempio 2.42, per ogni $m, n \in \mathbb{N}$ vale $(m, n) \in R$ oppure $(n, m) \in R$. Al contrario, nell'Esempio 2.43 esistono $m, n \in \mathbb{N}$ tali che $(m, n), (n, m) \notin R$. Ad esempio, basta scegliere m ed n coprimi tra loro.

Definizione 2.45. *Siano A un insieme ed R una relazione d'ordine su A. R si dice relazione d'ordine totale, e la coppia (A, R) si dice insieme totalmente ordinato, se per ogni $a_1, a_2 \in A$ vale $(a_1, a_2) \in R$ oppure $(a_2, a_1) \in R$. In caso contrario si dice che R è una relazione d'ordine parziale e la coppia (A, R) si dice insieme parzialmente ordinato.*

Osservazione 2.46. Quando si definisce un insieme con la seguente notazione

$$A = \{a_1, a_2, \dots, a_{n-1}, a_n, a_{n+1}, \dots\},$$

si intende che la collezione di oggetti indicati sia priva di un ordine specifico. Questo significa che potremmo anche scrivere

$$A = \{a_2, a_1, \dots, a_{n-1}, a_n, a_{n+1}, \dots\}.$$

Naturalmente, in tali situazioni è semplice assegnare un ordine canonico sull'insieme A. Infatti, data la funzione

$$\begin{array}{rccc} \phi: & A & \to & \mathbb{N} \\ & a_n & \mapsto & n \end{array},$$

si può definire

$$R_\phi = \{(a_m, a_n) \in A^2 \,|\, \phi(a_m) \leq \phi(a_n)\} = \{(a_m, a_n) \in A^2 \,|\, m \leq n\}.$$

Lasciamo al lettore la facile verifica che R_ϕ è una relazione d'ordine totale su A. Nel proseguio di questa dispensa, quando parleremo di insieme A ordinato, intenderemo la coppia (A, R_ϕ). In tal caso, cambiare l'ordine degli elementi dell'insieme produce un'insieme ordinato differente.

CAPITOLO 3

Matrici

3.1. Definizione, operazioni e proprietà elementari

Definizione 3.1. *Dati* $M = \{1, \ldots, m\}$ *e* $N = \{1, \ldots, n\}$, *una matrice di ordine* (m, n) *ad elementi nel campo* $\mathbb{K}$ *è una funzione*

$$\begin{array}{rccc} A: & M \times N & \longrightarrow & \mathbb{K} \\ & (i,j) & \longmapsto & a_{ij} \end{array}.$$

L'insieme delle matrici di tipo (m, n) *su campo* $\mathbb{K}$ *viene chiamato* $\mathrm{Mat}(m, n; \mathbb{K})$.

Esempio 3.2. Prendiamo $M = N = \{1, 2\}$ e $\mathbb{K} = \mathbb{Z}_2$. Allora esistono esattamente $2^{mn} = 16$ matrici. La matrice uno è la funzione che vale identicamente 1 su ogni elemento di $M \times N$:

$$A(1, 1) = a_{11} = 1, \quad A(1, 2) = a_{12} = 1, \quad A(2, 1) = a_{21} = 1, \quad A(2, 2) = a_{22} = 1.$$

Normalmente si è soliti rappresentare una matrice incasellando le immagini di A in una tabella con m righe e n colonne:

$$A = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix}.$$

La matrice

$$A_{R(i)} = [a_{i1} \ \ldots \ a_{in}] \in \mathrm{Mat}(1, n; \mathbb{K})$$

è detta riga i–esima di A, mentre la matrice

$$A_{C(j)} = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \in \mathrm{Mat}(m, 1; \mathbb{K})$$

è detta colonna j–esima di A. La matrice A può essere allora pensata come composta da m matrici riga di tipo $(1, n)$, oppure da n matrici colonne di tipo $(m, 1)$:

$$A = \begin{bmatrix} A_{R(1)} \\ \hline \vdots \\ \hline A_{R(m)} \end{bmatrix} = \left[\begin{array}{c|c|c} A_{C(1)} & \ldots & A_{C(n)} \end{array} \right].$$

Se applichiamo la rappresentazione con la tabella all'esempio precedente Esempio 3.2, abbiamo

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad A_{R(1)} = A_{R(2)} = \begin{bmatrix} 1 & 1 \end{bmatrix}, \qquad A_{C(1)} = A_{C(2)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Osservazione 3.3. Sebbene una n-upla in $\mathbb{K}^n$ ed una matrice formata da una sola riga siano due oggetti ben distinti, definiti in maniera differente, esiste una funzione biunivoca naturale tra $\mathbb{K}^n$ e

$\mathrm{Mat}(1, n; \mathbb{K})$:

(1)
$$f: \begin{array}{ccc} \mathbb{K}^n & \longrightarrow & \mathrm{Mat}(1, n; \mathbb{K}) \\ (a_1, \dots, a_n) & \longmapsto & [a_1 \ \dots \ a_n] \end{array}.$$

Analogamente, esiste una funzione biunivoca naturale tra $\mathbb{K}^m$ e $\mathrm{Mat}(m, 1; \mathbb{K})$:

(2)
$$g: \begin{array}{ccc} \mathbb{K}^m & \longrightarrow & \mathrm{Mat}(m, 1; \mathbb{K}) \\ (a_1, \dots, a_m) & \longmapsto & \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \end{array}.$$

In particolare, f e g definiscono un'ovvia corrispondenza biunivoca tra $\mathbb{K}$ e $\mathrm{Mat}(1, 1; \mathbb{K})$. Anche se il lettore deve sempre tenere conto della distinzione tra gli insiemi e gli oggetti precedenti, in queste note a volte identificheremo n-uple e corrispondenti matrici righe o colonne utilizzando implicitamente le funzioni f oppure g, al fine di semplificare la nostra esposizione.

Definizione 3.4. *Sull'insieme $M \times N$ definiamo:*

- *la funzione nulla*

$$0_{mn}: \begin{array}{ccc} M \times N & \longrightarrow & \mathbb{K} \\ (i, j) & \longmapsto & 0 \end{array};$$

- *la funzione delta di Kronecker*

$$\mathbb{I}_{mn}: \begin{array}{ccc} M \times N & \longrightarrow & \mathbb{K} \\ (i, j) & \longmapsto & \delta_{ij} \end{array},$$

  dove il simbolo δ_{ij} vale 1 se $i = j$ e 0 altrimenti.

Le due funzioni precedenti sono delle matrici e svolgono un ruolo speciale all'interno dell'algebra matriciale. Nella terminologia delle matrici, esse vengono rispettivamente chiamate matrice nulla e la matrice identità. Ad esempio,

$$0_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \text{e} \qquad \mathbb{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Osserviamo che abbiamo utilizzato la convenzione di scrivere un solo pedice nel caso $m = n$.

Un'altra comoda notazione per rappresentare una matrice è $A = [a_{ij}]$. Utilizzandola, possiamo agevolmente definire le prime due operazioni che coinvolgono matrici.

Definizione 3.5. *In* $\mathrm{Mat}(m, n; \mathbb{K})$ *si definiscono:*

i) l'operazione interna di somma

$$+: \begin{array}{ccc} \mathrm{Mat}(m, n; \mathbb{K}) \times \mathrm{Mat}(m, n; \mathbb{K}) & \longrightarrow & \mathrm{Mat}(m, n; \mathbb{K}) \\ ([a_{ij}], [b_{ij}]) & \longmapsto & [a_{ij} + b_{ij}] \end{array};$$

ii) l'operazione esterna di moltiplicazione per uno scalare

$$\cdot: \begin{array}{ccc} \mathbb{K} \times \mathrm{Mat}(m, n; \mathbb{K}) & \longrightarrow & \mathrm{Mat}(m, n; \mathbb{K}) \\ (t, [a_{ij}]) & \longmapsto & [t\, a_{ij}] \end{array}.$$

Esempio 3.6. Consideriamo $\mathrm{Mat}(2,2;\mathbb{K})$, allora

$$\begin{aligned}\left(t_1\cdot\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}\right)+\left(t_2\cdot\begin{bmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{bmatrix}\right) &= \begin{bmatrix}t_1\,a_{11} & t_1\,a_{12}\\ t_1\,a_{21} & t_1\,a_{22}\end{bmatrix}+\begin{bmatrix}t_2\,b_{11} & t_2\,b_{12}\\ t_2\,b_{21} & t_2\,b_{22}\end{bmatrix}\\ &= \begin{bmatrix}t_1\,a_{11}+t_2\,b_{11} & t_1\,a_{12}+t_2\,a_{12}\\ t_1\,a_{21}+t_2\,a_{21} & t_1\,a_{22}+t_2\,a_{22}\end{bmatrix}.\end{aligned}$$

Essendo definite elemento per elemento, è elementare verificare che le due operazioni soddisfano molte delle usuali proprietà di somma e prodotto in un campo, tra cui quelle contenute nella seguente proposizione.

Proposizione 3.7. *Siano $A, B \in \mathrm{Mat}(m,n;\mathbb{K})$ e $t_1, t_2 \in \mathbb{K}$. Allora:*

i) $(\mathrm{Mat}(m,n;\mathbb{K}),+)$ *è un gruppo abeliano, con elemento neutro 0_{mn} ed inverso additivo di A la matrice $-1\cdot A$;*

ii) la moltiplicazione per uno scalare è distributiva rispetto alla somma di matrici, cioè $t\cdot(A+B)=(t\cdot A)+(t\cdot B)$;

iii) la moltiplicazione per uno scalare è distributiva rispetto alla somma in $\mathbb{K}$, cioè $(t_1+t_2)\cdot A=(t_1\cdot A)+(t_2\cdot A)$;

iv) la moltiplicazione per uno scalare è omogenea rispetto alla moltiplicazione in $\mathbb{K}$, cioè $t_1\cdot(t_2\cdot A)=(t_1t_2)\cdot A$;

v) la moltiplicazione per uno scalare soddisfa la proprietà di normalizzazione, ovvero $1\cdot A=A$.

Tra matrici è definita anche una terza operazione, detta prodotto di matrici oppure prodotto righe per colonne.

Definizione 3.8. *Il prodotto righe per colonne è l'operazione*

$$\begin{array}{rccc} *: & \mathrm{Mat}(m,p;\mathbb{K})\times\mathrm{Mat}(p,n;\mathbb{K}) & \longrightarrow & \mathrm{Mat}(m,n;\mathbb{K})\\ & ([a_{ij}],[b_{ij}]) & \longmapsto & \left[\sum_{k=1}^{p} a_{ik}b_{kj}\right]\end{array}.$$

Per prima cosa, osserviamo che il prodotto di due matrici $A*B$ è definito se e solo se il numero di colonne di A è uguale al numero di righe di B. Analizziamo il caso $m=n=1$:

$$[a_{11}\ \dots\ a_{1p}] * \begin{bmatrix}b_{11}\\ \vdots\\ b_{p1}\end{bmatrix} = [\,a_{11}b_{11}+\cdots+a_{1p}b_{p1}\,] \in \mathrm{Mat}(1,1;\mathbb{K}).$$

Utilizzando l'identificazione naturale tra $\mathrm{Mat}(1,1;\mathbb{K})$ e $\mathbb{K}$, è facile osservare che l'elemento di posto (i,j) della matrice $A*B$ è uguale al prodotto $A_{R(i)}*B_{C(j)}$. Ad esempio, se $A\in\mathrm{Mat}(2,2;\mathbb{K})$ e $B\in\mathrm{Mat}(2,3;\mathbb{K})$ abbiamo

$$\begin{aligned}\left[\begin{array}{cc}a_{11} & a_{12}\\ \hline a_{21} & a_{22}\end{array}\right] * \left[\begin{array}{c|c|c}b_{11} & b_{12} & b_{13}\\ b_{21} & b_{22} & b_{23}\end{array}\right] &= \begin{bmatrix}A_{R(1)}*B_{C(1)} & A_{R(1)}*B_{C(2)} & A_{R(1)}*B_{C(3)}\\ A_{R(2)}*B_{C(1)} & A_{R(2)}*B_{C(2)} & A_{R(2)}*B_{C(3)}\end{bmatrix}\\ &= \begin{bmatrix}a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} & a_{11}b_{13}+a_{12}b_{23}\\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} & a_{21}b_{13}+a_{22}b_{23}\end{bmatrix}.\end{aligned}$$

Altri modi convenienti di scrivere il prodotto tra A e B sono contenuti nel seguente lemma, la cui verifica è lasciata al lettore.

Lemma 3.9. *Date $A\in\mathrm{Mat}(m,p;\mathbb{K})$ e $B\in\mathrm{Mat}(p,n;\mathbb{K})$ vale* *i)* $A * B = \left[\begin{array}{c} a_{11} \cdot B_{R(1)} + \cdots + a_{1p} \cdot B_{R(p)} \\ \hline \vdots \\ \hline a_{m1} \cdot B_{R(1)} + \cdots + a_{mp} \cdot B_{R(p)} \end{array}\right] = \left[\begin{array}{c} A_{R(1)} * B \\ \hline \vdots \\ \hline A_{R(m)} * B \end{array}\right];$

ii) $A * B = \left[\begin{array}{c|c|c} b_{11} \cdot A_{C(1)} + \cdots + b_{p1} \cdot A_{C(n)} & \ldots & b_{1n} \cdot A_{C(1)} + \cdots + b_{pn} \cdot A_{C(n)} \end{array}\right]$

$\qquad = \left[\begin{array}{c|c|c} A * B_{C(1)} & \ldots & A * B_{C(n)} \end{array}\right].$

Come osservato precedentemente, l'operazione $*$ soddisfa solo alcune delle proprietà abituali delle operazioni nei campi.

Proposizione 3.10. *Il prodotto riga per colonna:*

i) è associativo, cioè $(A * B) * C = A * (B * C)$ *per ogni* $A \in \mathrm{Mat}(m, n; \mathbb{K})$, $B \in \mathrm{Mat}(n, p; \mathbb{K})$, $C \in \mathrm{Mat}(p, q; \mathbb{K})$;

ii) è distributivo rispetto alla somma di matrici, cioè $A * (B + C) = (A * B) + (A * C)$ *per ogni* $A \in \mathrm{Mat}(m, n; \mathbb{K})$ *e* $B, C \in \mathrm{Mat}(n, p; \mathbb{K})$, *ed analogamente* $(A + B) * C = (A * C) + (B * C)$ *per ogni* $A, B \in \mathrm{Mat}(m, n; \mathbb{K})$ *e* $C \in \mathrm{Mat}(n, p; \mathbb{K})$;

iii) è omogeneo rispetto al prodotto per uno scalare, cioè $t \cdot (A * B) = (t \cdot A) * B = A * (t \cdot B)$ *per ogni* $t \in \mathbb{K}$, $A \in \mathrm{Mat}(m, n; \mathbb{K})$ *e* $B \in \mathrm{Mat}(n, p; \mathbb{K})$;

iv) ha $\mathbb{I}_n$ *come elemento neutro, nel senso che date* $A \in \mathrm{Mat}(m, n; \mathbb{K})$ *e* $B \in \mathrm{Mat}(n, p; \mathbb{K})$, *vale* $A * \mathbb{I}_n = A$ *ed* $\mathbb{I}_n * B = B$.

Dimostrazione. La dimostrazione si basa sul verificare le uguaglianza per il singolo elemento di posto (i, j) delle matrici coivolte, utilizzando le proprietà di somma e moltiplicazione in un campo.

i) $((A * B) * C)_{ij} = \sum_{l=1}^{p} \left(\sum_{k=1}^{n} a_{ik} b_{kl} \right) c_{lj} = \sum_{k=1}^{n} a_{ik} \left(\sum_{l=1}^{p} b_{kl} c_{lj} \right) = (A * (B * C))_{ij}$.

ii) Verifichiamo la prima proprietà, in quanto la seconda è analoga:

$$(A*(B+C))_{ij} = \sum_{k=1}^{n} a_{ik} \, (B + C)_{kj} = \sum_{k=1}^{n} a_{ik} \, (b_{kj} + c_{kj}) = \sum_{k=1}^{n} a_{ik} b_{kj} + \sum_{k=1}^{n} a_{ik} c_{kj} = (A*B)_{ij} + (A*C)_{ij}.$$

iii) Come per il punto precedente, verifichiamo solo la prima proprietà:

$$(t \cdot (A * B))_{ij} = t \left(\sum_{k=1}^{n} a_{ik} b_{kj} \right) = \sum_{k=1}^{n} (t \, a_{ik}) \, b_{kj} = ((t \cdot A) * B)_{ij} \, .$$

iv) $(A * \mathbb{I}_n)_{ij} = \sum_{k=1}^{n} a_{ik} \, \delta_{kj} = a_{ij}$ e analogamente l'altra identità. □

D'altra parte, in generale il prodotto di matrici non soddisfa la proprietà commutativa, di annullamento del prodotto e di esistenza dell'inverso. Verifichiamolo su dei controesempi. Siano

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \qquad \Rightarrow \qquad A * B = \begin{bmatrix} b_{21} & b_{22} \\ 0 & 0 \end{bmatrix}, \quad B * A = \begin{bmatrix} 0 & b_{11} \\ 0 & b_{21} \end{bmatrix}.$$

Allora:

- $A * B = B * A$ se e solo se $b_{11} = b_{22}$ e $b_{21} = 0$;
- $A * B = 0_{22}$ se e solo se $b_{21} = b_{22} = 0$, ovvero il prodotto di due matrici non nulle può essere la matrice nulla (ad esempio scegliendo $b_{11} \neq 0$);
- non esiste alcuna matrice B tale per cui $A * B = \mathbb{I}_2$, cioè la matrice A non è invertibile.

Lo studio dell'esistenza ed il calcolo della matrice inversa di una matrice data è un argomento molto importante dell'algebra matriciale, a cui dedicheremo l'intera Sezione 3.5.

Completiamo infine la sezione dando la definizione di trasposizione.

Definizione 3.11. *Data $A = [a_{ij}] \in \mathrm{Mat}(m, n; \mathbb{K})$, si definisce la matrice trasposta $A^T \in \mathrm{Mat}(n, m; \mathbb{K})$ la funzione*

$$\begin{array}{rccc} A^T: & N \times M & \longrightarrow & \mathrm{Mat}(n, m; \mathbb{K}) \\ & (j, i) & \longmapsto & a_{ij} \end{array}.$$

Nella usuale rappresentazione delle matrici abbiamo

$$A = \left[\begin{array}{ccc} a_{11} & \dots & a_{1n} \\ \hline \vdots & & \vdots \\ \hline a_{m1} & \dots & a_{mn} \end{array}\right] \qquad \Rightarrow \qquad A^T = \left[\begin{array}{c|c|c} a_{11} & \dots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{mn} \end{array}\right].$$

Dalla definizione segue immediatamente che $A^T_{R(j)} = A_{C(j)}$ e $A^T_{C(i)} = A_{R(i)}$.

Proposizione 3.12. *La trasposizione:*

i) *è un'involuzione, cioè $(A^T)^T = A$ per ogni $A \in \mathrm{Mat}(m, n; \mathbb{K})$;*
ii) *soddisfa la proprietà di linearità $(t_1 \cdot A + t_2 \cdot B)^T = t_1 \cdot A^T + t_2 \cdot B^T$ per ogni $A, B \in \mathrm{Mat}(m, n; \mathbb{K})$ e $t_1, t_2 \in \mathbb{K}$;*
iii) *soddisfa $(A * B)^T = B^T * A^T$ per ogni $A \in \mathrm{Mat}(m, p; \mathbb{K}), B \in \mathrm{Mat}(p, n; \mathbb{K})$.*

Dimostrazione. Le prime due proprietà sono elementari. Verifichiamo la terza. L'elemento di posto (i, j) della matrice a sinistra dell'uguaglianza è $((A * B)^T)_{ij} = (A * B)_{ji} = \sum_{k=1}^{p} a_{jk} b_{ki}$. Questo coincide con il corrispondente elemento della matrice a destra $(B^T * A^T)_{ij} = \sum_{k=1}^{p} b_{ki} a_{jk}$. □

Sfruttando l'associatività del prodotto di matrici, l'ultima proprietà si generalizza facilmente:

$$(3) \qquad (A_1 * \cdots * A_k)^T = ((A_1 * \cdots * A_{k-1}) * A_k)^T = A_k^T * (A_1 * \cdots * A_{k-1})^T = \cdots = A_k^T * \cdots * A_1^T.$$

3.2. Metodo di eliminazione di Gauss e rango

Definizione 3.13. *Data una matrice $A \in \mathrm{Mat}(m, n; \mathbb{K})$, si definisce pivot p_i il primo elemento diverso da zero della riga $A_{R(i)}$. La matrice A si dice a scala di rango r se:*

i) *ciascuna delle prime r righe di A contiene un pivot e le successive sono nulle;*
ii) *la posizione (i, j) del pivot p_i è più a sinistra della posizione $(i + 1, k)$ del pivot successivo p_{i+1}, ovvero $j < k$, per ogni $i \in \{1, \dots, r - 1\}$.*

Esempio 3.14. Siano

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \ B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \quad \text{e} \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

La prima matrice ha due pivot $p_1 = a_{11} = 1$ e $p_2 = a_{22} = 1$ ed è una matrice a scala di rango $r = 2$. La seconda matrice ha tre pivot $p_1 = a_{11} = 1$, $p_2 = a_{22} = 1$ e $p_3 = a_{32} = -1$, non è a scala in quanto il secondo pivot si trova nella stessa colonna del terzo. Infine, la terza matrice ha due pivot $p_1 = a_{11} = 1$ e $p_3 = a_{33} = 1$, non è a scala in quanto la seconda riga è nulla mentre la terza non lo è. Osserviamo come casi particolari che la matrice nulla e la matrice identità sono entrambe a scala.

Definizione 3.15. *Data* $A \in \mathrm{Mat}(m,n;\mathbb{K})$, *chiamiamo operazioni elementari sulle righe le seguenti tre trasformazioni della matrice:*

i) la permutazione delle righe $A_{R(i)}$ *ed* $A_{R(j)}$;

ii) la moltiplicazione della riga $A_{R(i)}$ *per uno scalare non nullo* $t \in \mathbb{K}^*$;

iii) la sostituzione delle riga $A_{R(i)}$ *con la somma della riga stessa ed un multiplo di una seconda riga* $A_{R(i)} + t \cdot A_{R(j)}$, *dove* $t \in \mathbb{K}$.

Teorema 3.16. *È sempre possibile trasformare una matrice* $A \in \mathrm{Mat}(m,n;\mathbb{K})$ *in una matrice a scala* $S \in \mathrm{Mat}(m,n;\mathbb{K})$ *attraverso una sequenza finita di operazioni elementari del primo e terzo tipo.* S *viene detta riduzione a scala di* A.

Dimostrazione. La dimostrazione è costruttiva in quanto definiamo un algoritmo che consente il calcolo di S a partire dalla matrice generica A.

Algoritmo 1 Metodo di Eliminazione di Gauss (MEG)

Input: $A \in \mathrm{Mat}(m,n;\mathbb{K})$.

Output: $S \in \mathrm{Mat}(m,n;\mathbb{K})$ ridotta a scala.

1: Se A è a scala, allora $S = A$ e l'algoritmo termina. Altrimenti definiamo $l = 1$.

2: Sia j_{l+k} l'indice della colonna di A contenente il pivot p_{l+k}, per ogni $k \in \{0,\dots,m-l\}$. Calcoliamo il numero naturale minimo $\bar{k}$ tale per cui $j_{l+\bar{k}} \le j_{l+k}$ per ogni $k \in \{0,\dots,m-l\}$.

3: Se $\bar{k} = 0$, andiamo alla linea 4. Se $\bar{k} \neq 0$, ridefiniamo A permutando le righe $A_{R(l)}$ ed $A_{R(l+\bar{k})}$, e ritorniamo alla linea 2.

4: Ridefiniamo A sostituendo la riga $A_{R(l+k)}$ con la riga $A_{R(l+k)} - \dfrac{a_{l+k\, j_{l+k}}}{p_{j_l}} A_{R(l)}$, per ogni $k \in \{1,\dots,m-l\}$.

5: Se A è a scala, allora $S = A$ e l'algoritmo termina. Altrimenti, ridefiniamo $l = l + 1$ ed iteriamo la procedura a partire dalla linea 2.

È semplice constatare che, con l'esecuzione delle linee 2 e 3 dell'algoritmo, il pivot p_l della matrice A viene a trovarsi a sinistra rispetto a tutti i pivot successivi. In seguito, con l'applicazione del comando nella linea 4, si pongono a zero tutti gli elementi situati sotto p_l. Con queste due operazioni si è quindi creato un gradino della matrice a scala S. Successivamente, attraverso l'esecuzione della linea 5, il numero l aumenta di 1 e si ripete l'algoritmo a partire dalla riga successiva della matrice. Poiché $l \le m < \infty$, concludiamo che l'algoritmo termina in un numero finito di iterazioni. $\square$

Esempio 3.17. Sia

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \ (m,n) = (3,4) \xrightarrow{1} \begin{array}{c} A \text{ non è a scala,} \\ l = 1. \end{array} \xrightarrow{2}$$

$$\text{Prima iterazione: } \begin{cases} j_1 = 1, \ p_1 = 1 \\ j_2 = 1, \ p_2 = 1 \\ j_3 = 2, \ p_3 = 1 \end{cases} \Rightarrow \bar{k} = 0 \xrightarrow{3,4} A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \xrightarrow{5} \begin{array}{c} A \text{ non è a scala,} \\ l = 2. \end{array}$$

$$\text{Seconda iterazione: } \begin{cases} j_2 = 3, \ p_2 = -1 \\ j_3 = 2, \ p_3 = 1 \end{cases} \Rightarrow \bar{k} = 1 \xrightarrow{3} A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \end{bmatrix} \xrightarrow{2}$$

$$\begin{cases} j_2 = 2, \ p_2 = 1 \\ j_3 = 3, \ p_3 = -1 \end{cases} \Rightarrow \bar{k} = 0 \xrightarrow{3,4} A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \end{bmatrix} \xrightarrow{5} \begin{array}{c} A \text{ è a scala,} \\ S = A. \end{array}$$

L'algoritmo di eliminazione definito nella dimostrazione del Teorema 3.16 è deterministico, nel senso che associa ad una data matrice A un'unica matrice ridotta S. Attraverso quest'unica matrice possiamo definire il rango della matrice di partenza.

Definizione 3.18. *Il rango di $A \in \mathrm{Mat}(m, n; \mathbb{K})$ è la funzione che associa ad A il rango della sua riduzione a scala S, ottenuta attraverso il Metodo di Eliminazione di Gauss.*

D'altra parte è possibile ottenere riduzioni a scala differenti della matrice A, applicando in maniera diversa le operazioni elementari. Questa è una pratica usuale che è spesso conveniente per semplificare e velocizzare i conti da fare.

Esempio 3.19.

$$A = \begin{bmatrix} 2 & 0 & 2 & 0 \\ 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \end{bmatrix} \xrightarrow[R_3 - 2R_2 \to R_3]{\frac{1}{2}R_1 \to R_1} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_2 - R_1 \to R_2} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} = S$$

$$\Rightarrow \quad r(A) = r(S) = 2.$$

Nell'esempio, la notazione $R_3 - 2R_2 \to R_3$ significa che la terza riga della matrice trasformata corrisponde alla differenza tra la terza e due volte la seconda riga della matrice precedente. È importante osservare che cambiando l'ordine delle operazioni la matrice S ottenuta risulterà differente, ma di rango uguale a quella che si otterrebbe dall'esatta applicazione dell'Algoritmo 1. In generale, il rango di A potrebbe essere definito come il rango di una sua qualsiasi riduzione a scala. Questo suggerisce che il rango sia una quantità intrinseca della matrice A. Ritorneremo su queste affermazioni dimostrandole quando parleremo di spazi vettoriali delle righe e delle colonne di una matrice. Dalla definizione di rango segue immediatamente una sua proprietà elementare ma molto utile.

Lemma 3.20. *Data $A \in \mathrm{Mat}(m, n, \mathbb{K})$, si ha $0 \leq r(A) \leq \min(m, n)$.*

Dimostrazione. Nella matrice ridotta a scala esiste al più un unico pivot per ogni riga ed ogni colonna. $\square$

3.3. Sistemi lineari

Una applicazione fondamentale delle matrici, ed in particolare del metodo di eliminazione di Gauss, è la risoluzione dei sistemi lineari.

Definizione 3.21. *Consideriamo il seguente sistema di m equazioni lineari con coefficienti a_{ij} e termini noti b_i nel campo $\mathbb{K}$ ed n incognite x_j, dove $1 \leq i \leq m$ e $1 \leq j \leq n$:*

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

Allora la matrice dei coefficienti e le colonne dei termini noti e delle incognite sono:

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \quad \text{ed} \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

La matrice completa (od orlata) del sistema lineare è $[A|B]$*, ottenuta aggiungendo alla destra di* A *la colonna* B*. Infine, il sistema lineare definito dai coefficienti della matrice completa* $[A|0_{m1}]$ *si chiama sistema omogeneo associato al sistema* $[A|B]$*.*

Osserviamo che un sistema lineare può essere riscritto in forma matriciale utilizzando il prodotto righe per colonne. Infatti, trovare la n-upla $(x_1, \dots, x_n) \in \mathbb{K}^n$ soluzione del sistema lineare è equivalente a trovare la matrice colonna X tale che $A * X = B$. Possiamo identificare i due oggetti con la funzione (2) definita nell'Osservazione 3.3. Ad esempio:

$$\begin{cases} x_1 + x_2 & = & 1 \\ x_1 - x_2 & = & 0 \end{cases} \quad \Rightarrow \quad A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \ B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ [A|B] = \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & -1 & 0 \end{array}\right] \text{ ed } X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Visto che

$$A * \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = B,$$

allora la soluzione del sistema è $(x_1, x_2) = \left(\frac{1}{2}, \frac{1}{2}\right)$.

Proposizione 3.22. *Sia* $[A|B]$ *la matrice completa di un sistema lineare e sia* $[S|B']$ *una sua riduzione a scala. Allora il sistema associato alla matrice ridotta ha lo stesso insieme delle soluzioni del sistema di partenza.*

Dimostrazione. A seguito di un'operazione elementare sulle righe di $[A|B]$ otteniamo una nuova matrice $[\tilde{A}|\tilde{B}]$ e quindi passiamo dal sistema lineare originario ad un sistema trasformato. Dobbiamo dimostrare che le soluzioni del sistema di partenza non cambiano a seguito dell'applicazione di tali operazioni elementari. Questo è evidente nel caso delle permutazioni di righe e nella moltiplicazione di una riga per uno scalare non nullo. Studiamo invece con più attenzione il terzo tipo di trasformazioni. Supponiamo che la riga $[A_{R(i)}|b_i]$ venga sostituita da $[\tilde{A}_{R(i)}|\tilde{b}_i] = [A_{R(i)} + tA_{R(j)}|b_i + tb_j]$. Concentriamoci sulle equazioni i e j dei due sistemi lineari associati, visto che le altre rimangono inalterate. Nel sistema di partenza abbiamo

$$(4) \qquad \begin{cases} a_{i1}x_1 + \dots + a_{in}x_n - b_i & = & 0 \\ a_{j1}x_1 + \dots + a_{jn}x_n - b_j & = & 0 \end{cases},$$

mentre le equazioni del sistema trasformato possono essere scritte come

$$(5) \qquad \begin{cases} (a_{i1}x_1 + \dots + a_{in}x_n - b_i) + t(a_{j1}x_1 + \dots + a_{jn}x_n - b_j) & = & 0 \\ a_{j1}x_1 + \dots + a_{jn}x_n - b_j & = & 0 \end{cases}.$$

Dovrebbe essere evidente che se $(x_1, \dots, x_n)$ è soluzione di 4, allora risolve anche 5 per ogni $t \in \mathbb{K}$. Viceversa, se $(x_1, \dots, x_n)$ risolve la seconda equazione del sistema 5, allora la prima equazione si può semplificare ad

$$a_{i1}x_1 + \dots + a_{in}x_n - b_i = 0.$$

Ma questo significa che $(x_1, \dots, x_n)$ soluzione del sistema trasformato deve risolvere anche il sistema lineare di partenza. $\square$

La proposizione precedente è fondamentale per poter definire un metodo efficiente per la risoluzione di sistemi lineari. Infatti ci consente di risolvere un sistema lineare complicato attraverso la soluzione di un sistema più semplice, quello associato alla matrice ridotta a scala.

Esempio 3.23. Risolviamo il sistema lineare

$$\begin{cases} x+y-z &= 1 \\ x-y+z &= 1 \\ 2x+2y &= 0 \end{cases} \Rightarrow [A|B] = \left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 2 & 2 & 0 & 0 \end{array}\right] \xrightarrow[R_3-2R_1\to R_3]{R_2-R_1\to R_2} \left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & -2 & 2 & 0 \\ 0 & 0 & 2 & -2 \end{array}\right] = [S|B'].$$

Quindi il sistema di partenza è equivalente al sistema ridotto

$$\begin{cases} x+y-z &= 1 \\ -2y+2z &= 0 \\ 2z &= -2 \end{cases}.$$

La soluzione la possiamo ottenere partendo dall'ultima equazione e risolvendo a ritroso quelle precedenti.

$$\begin{cases} x = 1-y+z = 1 \\ y = z = -1 \\ z = -1 \end{cases} \Rightarrow \quad X = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

Questo metodo è del tutto generale e si chiama algoritmo di Gauss per la risoluzione dei sistemi lineari. Come detto precedentemente, questo è uno dei metodi più efficienti dal punto di vista computazionale per la risoluzione di un generico sistema lineare.

Grazie al metodo di eliminazione possiamo enunciare e dimostrare i due teoremi fondamentali sui sistemi lineari. Il primo è il teorema di Rouché–Capelli, che ci permette di studiare il numero di soluzioni di un sistema senza necessariamente risolverlo. Nella sua dimostrazione daremo anche esplicitamente la formula risolutiva di un sistema lineare, basata sul metodo di eliminazione. Il secondo risultato è il cosiddetto teorema di struttura delle soluzioni.

Teorema 3.24. *Consideriamo il sistema lineare associato alla matrice completa* $[A|B] \in \mathrm{Mat}(m, n+1; \mathbb{K})$. *Allora:*

i) la soluzione non esiste se e solo se $r([A|B]) > r(A)$;

ii) la soluzione esiste ed è unica se e solo se $r([A|B]) = r(A) = n$;

iii) esistono infinite soluzioni dipendenti da $n - r(A)$ *parametri se e solo se* $r([A|B]) = r(A) < n$.

Dimostrazione. Grazie alla Proposizione 3.22 è possibile verificare il teorema assumendo che la matrice completa sia ridotta a scala.

i) La condizione $r([A|B]) > r(A)$ è verificata se e solo se la riga r–esima contenente l'ultimo pivot è $[0 \ \dots \ 0 \,|\, p_r]$. Per il sistema lineare associato, questa riga corrisponde all'equazione $0 = p_r$, che ovviamente non ha soluzione.

ii) $r([A|B]) = r(A) = n$ significa che ogni colonna della matrice A contiene un pivot. Pertanto il sistema lineare associato è

$$\begin{cases} p_1x_1 + a_{12}x_2 + \cdots + a_{1\,n-1}x_{n-1} + a_{1n}x_n &= b_1 \\ p_2x_2 + \cdots + a_{2\,n-1}x_{n-1} + a_{2n}x_n &= b_2 \\ \vdots & \vdots \\ p_{n-1}x_{n-1} + a_{n-1\,n}x_n &= b_{n-1} \\ p_nx_n &= b_n \end{cases},$$

più eventualmente un certo numero di righe nulle corrispondenti alle identità $0 = 0$. Definiamo

$$R_i(x_{i+1}, \dots, x_n) = \sum_{j=1}^{n-i} a_{i\,i+j}\, x_{i+j} \qquad \text{per} \qquad 1 \le i \le n-1.$$

Allora il sistema precedente si può riscrivere come

$$\begin{cases} p_1 x_1 + R_1(x_2, \dots, x_n) &= b_1 \\ p_2 x_2 + R_2(x_3, \dots, x_n) &= b_2 \\ \vdots & \vdots \\ p_{n-1} x_{n-1} + R_{n-1}(x_{n-1}) &= b_{n-1} \\ p_n x_n &= b_n \end{cases} \quad \Rightarrow \quad \begin{cases} x_1 &= \frac{b_1 - R_1(x_2,\dots,x_n)}{p_1} \\ \vdots & \vdots \\ x_{n-1} &= \frac{b_{n-1} - R_{n-1}(x_n)}{p_{n-1}} \\ x_n &= \frac{b_n}{p_n} \end{cases}.$$

Leggendo le espressioni precedenti dal basso verso l'alto, appare chiaro che la soluzione esiste ed è unica.

iii) La condizione $r([A|B]) = r(A) < n$ è verificata se e solo se $n - r(A)$ colonne di A, ovvero $n - r(A)$ incognite del sistema lineare, non sono associate a pivot. Dopo aver eventualmente rinominato le variabili in modo che tali $n - r(A)$ incognite siano $x_{r+1}, \dots, x_n$, le $r(A)$ equazioni non banali del sistema lineare sono

$$\begin{cases} p_1 x_1 + a_{12} x_2 + \dots + a_{1\,r-1} x_{r-1} + a_{1r} x_r + a_{1\,r+1} x_{r+1} + \dots + a_{1n} x_n &= b_1 \\ p_2 x_2 + \dots + a_{1\,r-1} x_{r-1} + a_{2r} x_r + a_{2\,r+1} x_{r+1} + \dots + a_{2n} x_n &= b_2 \\ \vdots & \vdots \\ p_{r-1} x_{r-1} + a_{r-1\,r} x_r + a_{r-1\,r+1} x_{r+1} + \dots + a_{r-1\,n} x_n &= b_{r-1} \\ p_r x_r + a_{r\,r+1} x_{r+1} + \dots + a_{rn} x_n &= b_r \end{cases}.$$

Ponendo $x_{r+1} = t_1, \dots, x_n = t_{n-r}$ dove $t_i \in \mathbb{K}$ sono parametri liberi, le restanti r incognite $x_1, \dots, x_r$ sono univocamente determinate dalle equazioni del sistema come funzioni dei parametri:

$$\begin{cases} x_1 &= \frac{b_1 - R_1(x_2,\dots,x_n)}{p_1} = \frac{b_1 - R_1(x_2,\dots,x_r,t_1,\dots,t_{n-r})}{p_1} \\ \vdots & \vdots \\ x_r &= \frac{b_r - R_r(x_{r+1},\dots,x_n)}{p_r} = \frac{b_r - R_r(t_1,\dots,t_{n-r})}{p_r} \end{cases}.$$

Pertanto esistono infinite soluzioni dipendenti da $n - r(A)$ parametri.

Infine, osserviamo che un sistema lineare ridotto a scala può assumere solo una delle tre forme precedentemente elencate, da cui deduciamo l'implicazione inversa del teorema. □

Corollario 3.25. *Un sistema lineare omogeneo ammette sempre soluzione. In particolare, esiste sempre la soluzione banale $X = 0_{n1}$.*

Dimostrazione. L'enunciato segue da $r([A|0_{m1}]) = r(A)$ ed $A * 0_{n1} = 0_{m1}$. □

Definizione 3.26. *Chiamiamo nucleo o kernel di una matrice $A \in \mathrm{Mat}(m, n; \mathbb{K})$ l'insieme $\ker(A)$ delle soluzioni del sistema omogeneo $[A|0_{m1}]$.*

Esempio 3.27. Dato il parametro $h \in \mathbb{K}$, discutere la risolubilità del sistema lineare

$$\begin{cases} x + y - hz &= 4 \\ y + hz &= h + 1 \\ hx + (h-1)y - 2z &= h + 1 \end{cases}.$$

Effettuiamo la riduzione a scala della matrice associata:

$$[A|B] = \left[\begin{array}{ccc|c} 1 & 1 & -h & 4 \\ 0 & 1 & h & h+1 \\ h & h-1 & -2 & h+1 \end{array}\right] \xrightarrow{R_3 - hR_1 \to R_3} \left[\begin{array}{ccc|c} 1 & 1 & -h & 4 \\ 0 & 1 & h & h+1 \\ 0 & -1 & -2+h^2 & -3h+1 \end{array}\right]$$

$$\xrightarrow{R_3 + R_2 \to R_3} \left[\begin{array}{ccc|c} 1 & 1 & -h & 4 \\ 0 & 1 & h & h+1 \\ 0 & 0 & -2+h+h^2 & -2h+2 \end{array}\right] = [S|B'].$$

La risolubilità del sistema dipende dal valore di h. Infatti:

$$r([A]) = \begin{cases} 3 & \text{se} \quad h \neq -2, 1 \\ 2 & \text{se} \quad h = -2, 1 \end{cases}, \qquad r([A|B]) = \begin{cases} 3 & \text{se} \quad h \neq 1 \\ 2 & \text{se} \quad h = 1 \end{cases}.$$

Pertanto il sistema non ha soluzione se $h = 2$, ha un'unica soluzione se $h \neq -2, 1$ ed ha infinite soluzioni dipendenti da un parametro se $h = 1$. Andiamo a risolvere il sistema nei diversi casi.

- $h \neq -2, 1$: l'unica soluzione è

$$\begin{cases} x + y - hz &= 4 \\ y + hz &= h+1 \\ z &= -\frac{1}{h+2} \end{cases} \quad \Rightarrow \quad X = \begin{bmatrix} \frac{-h^2-3h+6}{h+2} \\ \frac{h^2+5h+2}{h+2} \\ -\frac{2}{h+2} \end{bmatrix}.$$

- $h = 1$: esistono infinite soluzione dipendenti dal parametero t

$$\begin{cases} x + y - z &= 4 \\ y + z &= 2 \\ 0 &= 0 \end{cases} \quad \Rightarrow \quad X(t) = \begin{bmatrix} 2+2t \\ 2-t \\ t \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} + t \cdot \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}.$$

Se $h = 1$, osserviamo che

$$A * \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} = B, \qquad A * \left(t \cdot \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \right) = 0_{31}.$$

Questo non è un caso, bensì è il contenuto del teorema di struttura delle soluzioni di un sistema lineare.

Teorema 3.28. *Sia $[A|B] \in \mathrm{Mat}(m, n+1; \mathbb{K})$ la matrice completa di un sistema lineare risolvibile. Allora la soluzione generale del sistema ha la forma $X = X_p + X_0$, dove X_p è una qualunque soluzione particolare del sistema ed X_0 è la soluzione generale del sistema omogeneo associato.*

Dimostrazione. L'ipotesi del teorema è che il sistema sia risolvibile, pertanto esiste la sua soluzione generale $X(t_1, \ldots, t_{n-r})$ dipendente da $n - r(A)$ parametri. Per ogni scelta dei parametri abbiamo una soluzione particolare X_p. Dopo aver definito $X_0(t_1, \ldots, t_{n-r}) = X(t_1, \ldots, t_{n-r}) - X_p$, dobbiamo verificare che questa sia la soluzione generale del sistema omogeneo $[A|0_{m1}]$. Ma ciò è vero in quanto:

- $A * X_0(t_1, \ldots, t_{n-r}) = A * (X(t_1, \ldots, t_{n-r}) - X_p) = A * X(t_1, \ldots, t_{n-r}) - A * X_p = B - B = 0_{m1}$;
- $X_0(t_1, \ldots, t_{n-r})$ è funzione di $n - r(A)$ parametri liberi, tanti quanti sono quelli previsti dal teorema di Rouché-Capelli per la soluzione generale del sistema $[A|0_{m1}]$. $\square$

3.4. Matrici quadrate, triangolari, diagonali, simmetriche ed antisimmetriche

Concentriamoci ora su alcune classi importanti di matrici.

Definizione 3.29. *Una matrice $A \in \mathrm{Mat}(n,n;\mathbb{K})$ si dice quadrata di ordine n. Inoltre:*

i) se $a_{ij} = 0$ per ogni $i > j$ (rispettivamente $i \geq j$), allora la matrice si dice triangolare alta (strettamente alta). Indichiamo con $\mathbb{T}_a(n;\mathbb{K})$ (rispettivamente $\mathbb{T}_{sa}(n;\mathbb{K})$) l'insieme delle matrici triangolari alte (strettamente alte);

ii) se $a_{ij} = 0$ per ogni $i < j$ (rispettivamente $i \leq j$), allora la matrice si dice triangolare bassa (strettamente bassa). Indichiamo con $\mathbb{T}_b(n;\mathbb{K})$ (rispettivamente $\mathbb{T}_{sb}(n;\mathbb{K})$) l'insieme delle matrici triangolari basse (strettamente basse);

iii) se $a_{ij} = 0$ per ogni $i \neq j$, allora la matrice si dice diagonale. Indichiamo con $\mathbb{D}(n;\mathbb{K})$ l'insieme delle matrici diagonali.

Esempio 3.30. Dati $a, b, c, d, e \in \mathbb{K}$, sia

$$A = \begin{bmatrix} a & b & d \\ e & a & c \\ e & e & a \end{bmatrix}.$$

Allora

- $A \in \mathbb{T}_a(3;\mathbb{K})$ per ogni $a, b, c, d \in \mathbb{K}$ ed $e = 0$;
- $A \in \mathbb{T}_{sa}(3;\mathbb{K})$ per $a = e = 0$ ed ogni $b, c, d \in \mathbb{K}$;
- $A \in \mathbb{T}_b(3;\mathbb{K})$ per ogni $a, e \in \mathbb{K}$ e $b = c = d = 0$;
- $A \in \mathbb{T}_{sb}(3;\mathbb{K})$ per $a = b = c = d = 0$ ed ogni $e \in \mathbb{K}$;
- $A \in \mathbb{D}(3;\mathbb{K})$ per ogni $a \in \mathbb{K}$ e $b = c = d = e = 0$.

Osserviamo che una matrice quadrata a scala è necessariamente triangolare alta, ma in generale non è vero il viceversa. Ad esempio, se $a = b = e = 0$ e $c, d \neq 0$, allora A è triangolare alta ma non è a scala.

Proposizione 3.31. *Gli insiemi $\mathrm{Mat}(n,n;\mathbb{K})$, $\mathbb{T}_a(n;\mathbb{K})$, $\mathbb{T}_{sa}(n;\mathbb{K})$, $\mathbb{T}_b(n;\mathbb{K})$, $\mathbb{T}_{sb}(n;\mathbb{K})$, $\mathbb{D}(n;\mathbb{K})$ sono chiusi rispetto alla somma, al prodotto per uno scalare ed al prodotto righe per colonne. Questo significa che, date due matrici A, B appartenenti ad uno di questi insiemi e $t \in \mathbb{K}$, abbiamo $A + B$, $t \cdot A$ ed $A * B$ appartenenti ancora all'insieme dato.*

Dimostrazione. La verifica della chiusura rispetto alle operazioni $+$ e $\cdot$ è elementare, così come quella di $\mathrm{Mat}(n,n;\mathbb{K})$ rispetto al prodotto di matrici. Dimostriamo la chiusura di $\mathbb{T}_a(n;\mathbb{K})$) rispetto a $*$. Date due matrici A, B triangolari alte, calcoliamo l'elemento di posto (i,j) del loro prodotto, dove $i > j$:

$$(A * B)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{i-1} a_{ik} b_{kj} + \sum_{k=i}^{n} a_{ik} b_{kj} = \sum_{k=1}^{i-1} 0 \cdot b_{kj} + \sum_{k=i}^{n} a_{ik} \cdot 0 = 0.$$

Le altre verifiche sulle matrici triangolari sono analoghe. Per le matrici diagonali è sufficiente osservare che $\mathbb{D}(n;\mathbb{K}) = \mathbb{T}_a(n;\mathbb{K}) \cap \mathbb{T}_b(n;\mathbb{K})$. Quindi la chiusura di $\mathbb{D}(n;\mathbb{K})$ rispetto a $*$ segue dalla chiusura dei due insiemi delle matrici triangolari rispetto alla stessa operazione. $\square$

La chiusura di $\mathrm{Mat}(n,n;\mathbb{K})$ rispetto ad $*$ permette di definire l'elevazione ad una potenza naturale di una matrice e quindi il concetto di polinomio di matrici. Infatti chiamiamo

$$A^k = \begin{cases} \mathbb{I}_n & \text{se } k = 0, \\ A * \ldots * A & k \text{ volte se } k \in \mathbb{N}^*. \end{cases}$$

Allora un polinomio di A a coefficienti $c_i \in \mathbb{K},\ 0 \leq i \leq k$, è un'espressione del tipo

$$P(A) = \sum_{i=0}^{k} c_i \cdot A^i = c_0 \cdot \mathbb{I}_n + c_1 \cdot A + c_2 \cdot A^2 + \cdots + c_k \cdot A^k.$$

Un'altra classe importante di matrici quadrate sono contenute nella seguente definizione.

Definizione 3.32. *Una matrice* $A \in \mathrm{Mat}(n, n; \mathbb{K})$ *si dice:*

i) simmetrica se $a_{ij} = a_{ji}$, *per ogni* i, j. *Denotiamo con* $\mathbb{S}(n; \mathbb{K})$ *l'insieme delle matrici simmetriche;*

ii) antisimmetrica se $a_{ij} = -a_{ji}$, *per ogni* i, j. *Denotiamo con* $\mathbb{A}(n; \mathbb{K})$ *l'insieme delle matrici antisimmetriche.*

Esempio 3.33. Dati $a, b, c \in \mathbb{K}$, sia

$$A = \begin{bmatrix} a & b+c & 0 \\ b-c & a & b+c \\ 0 & b-c & a \end{bmatrix}.$$

Allora

- $A \in \mathbb{S}(3; \mathbb{K})$ per ogni $a, b \in \mathbb{K}$ e $c = 0$;
- $A \in \mathbb{A}(3; \mathbb{K})$ per $a = b = 0$ ed ogni $c \in \mathbb{K}$.

Osserviamo in particolare che gli elementi lungo la diagonale di una matrice antisimmetrica sono necessariamente uguali a zero.

È immediato verificare la seguente proprietà.

Proposizione 3.34. *Gli insiemi* $\mathbb{S}(n; \mathbb{K})$ *e* $\mathbb{A}(n; \mathbb{K})$ *sono chiusi rispetto alla somma ed al prodotto per uno scalare.*

D'altra parte, i due insiemi non sono chiusi rispetto al prodotto righe per colonne, come si può osservare nei seguenti controesempi.

Esempio 3.35. Siano

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix},\ B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \in \mathbb{S}(2; \mathbb{K}) \quad \Rightarrow \quad A * B = \begin{bmatrix} 1 & 1 \\ 0 & 1. \end{bmatrix} \notin \mathbb{S}(2; \mathbb{K}).$$

Questo dimostra che $\mathbb{S}(2; \mathbb{K})$ non è chiuso rispetto al prodotto. Analogamente, siano

$$A = B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in \mathbb{A}(2; \mathbb{K}) \quad \Rightarrow \quad A * B = \begin{bmatrix} -1 & 0 \\ 0 & -1. \end{bmatrix} \notin \mathbb{A}(2; \mathbb{K}).$$

Questo dimostra che $\mathbb{A}(2; \mathbb{K})$ non è chiuso rispetto al prodotto.

Proposizione 3.36. *Siano* $A, B \in \mathbb{S}(n; \mathbb{K})$ *e* $C, D \in \mathbb{A}(n; \mathbb{K})$. *Allora:*

i) $A * B \in \mathbb{S}(n; \mathbb{K})$ *se e solo se le due matrici commutano, ovvero* $A * B = B * A$;

ii) $C * D \in \mathbb{A}(n; \mathbb{K})$ *se e solo se le due matrici anticommutano, ovvero* $C * D = -D * C$.

Dimostrazione. Dimostriamo il primo punto, essendo il secondo analogo. Per l'ipotesi $A, B \in \mathbb{S}(n; \mathbb{K})$, abbiamo $(A * B)^T = B^T * A^T = B * A$. Supponiamo ora $A * B \in \mathbb{S}(n; \mathbb{K})$. Allora $A * B = (A * B)^T = B * A$. Viceversa, siano le due matrici commutanti. Allora $A * B = B * A = (A * B)^T$. $\square$

Enunciamo e dimostriamo due utili proprietà delle matrici simmetriche ed antisimmetriche che utilizzeremo in seguito.

Proposizione 3.37. *Sia* $A \in \mathrm{Mat}(m,n;\mathbb{K})$*. Allora:*

i) $A * A^T \in \mathbb{S}(m;\mathbb{K})$*;*
ii) $A^T * A \in \mathbb{S}(n;\mathbb{K})$*.*

DIMOSTRAZIONE. $(A*A^T)^T = (A^T)^T * A^T = A*A^T$ da cui la prima tesi. La seconda è analoga. □

Proposizione 3.38. *Data* $A \in \mathrm{Mat}(n,n;\mathbb{K})$*, definiamo le due matrici* $A_S = \frac{1}{2}(A + A^T)$ *ed* $A_A = \frac{1}{2}(A - A^T)$*. Allora:*

i) $A_S \in \mathbb{S}(n,\mathbb{K})$ *ed* $A_A \in \mathbb{A}(n;\mathbb{K})$*;*
ii) (A_S, A_A) *è l'unica coppia in* $\mathbb{S}(n;\mathbb{K}) \times \mathbb{A}(n;\mathbb{K})$ *tale che* $A = A_S + A_A$*.*

DIMOSTRAZIONE. La verifica del primo punto è immediata. Per il secondo punto, per prima cosa osserviamo che

$$A_S + A_A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T) = A.$$

Per dimostrare l'unicità della decomposizione, supponiamo che $(B,C) \in \mathbb{S}(n;\mathbb{K}) \times \mathbb{A}(n;\mathbb{K})$ sia una seconda coppia di matrici tali che $A = B + C$. Allora

$$A_S = \frac{1}{2}(A + A^T) = \frac{1}{2}((B + C) + (B + C)^T) = \frac{1}{2}(B + C + B - C) = B$$

ed analogamente

$$A_A = \frac{1}{2}(A - A^T) = \frac{1}{2}((B + C) - (B + C)^T) = \frac{1}{2}(B + C - B + C) = C.$$ □

Esempio 3.39. Sia

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{Mat}(2,2;\mathbb{K}) \quad \Rightarrow \quad A_S = \begin{bmatrix} a & \frac{b+c}{2} \\ \frac{b+c}{2} & d \end{bmatrix} \in \mathbb{S}(2;\mathbb{K}),\ A_A = \begin{bmatrix} 0 & \frac{b-c}{2} \\ \frac{-b+c}{2} & 0 \end{bmatrix} \in \mathbb{A}(2;\mathbb{K}).$$

3.5. Matrici elementari ed inversione di matrici quadrate

In questa sezione daremo la definizione della matrice inversa di una matrice quadrata, studieremo le condizioni di invertibilità di una matrice ed infine presenteremo un algoritmo per il calcolo della matrice inversa.

Definizione 3.40. *Date le matrici* $A, B, C \in \mathrm{Mat}(n,n;\mathbb{K})$*, si dice che*

i) B *è inversa sinistra di* A *se* $B * A = \mathbb{I}_n$*;*
ii) C *è inversa destra di* A *se* $A * C = \mathbb{I}_n$*.*

A si dice regolare o invertibile o non singolare se esistono un'inversa sinistra B ed un'inversa destra C uguali tra loro. In tal caso, $B = C$ si dice inversa di A.

Esempio 3.41. Consideriamo la matrice

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \in \mathrm{Mat}(2,2;\mathbb{Q}).$$

Calcolare l'inversa destra C di A significa risolvere l'equazione

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} x & z \\ y & w \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

nelle incognite x, y, z, w. Questa è equivalente al sistema lineare

$$\begin{cases} x+y=1 \\ z+w=0 \\ y=0 \\ w=1 \end{cases} \quad \Rightarrow \quad \begin{cases} x=1 \\ y=0 \\ z=-1 \\ w=1 \end{cases} \quad \Rightarrow \quad C = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}.$$

D'altra parte, è semplice verificare che $C * A = \mathbb{I}_2$, da cui otteniamo che C è anche l'inversa sinistra B di A. Di conseguenza la matrice A è invertibile, con unica matrice inversa $B = C$.

Il teorema fondamentale caratterizzante l'insieme delle matrici invertibili è il seguente.

Teorema 3.42. *$A \in \mathrm{Mat}(n, n; \mathbb{K})$ è invertibile se e soltanto se $r(A) = n$. In tal caso l'inversa è unica.*

La dimostrazione del teorema è abbastanza articolata e richiede la verifica di diversi risultati intermedi che presenteremo nelle prossime sottosezioni.

3.5.1. Unicità dell'inversa e proprietà elementari della matrice inversa. In questa sottosezione ci concentreremo sullo studio di alcune importanti proprietà delle matrici inverse, nell'ipotesi che queste esistano. In primo luogo, verifichiamo l'unicità della matrice inversa di una matrice data.

Proposizione 3.43. *Sia $A \in \mathrm{Mat}(n, n; \mathbb{K})$ una matrice che ammette inversa sinistra B e destra C. Allora:*

i) $B = C$ e sono entrambe uniche;
ii) la matrice A è invertibile e l'inversa è unica.

Dimostrazione. Verifichiamo che $B = C$ attraverso la seguente catena di uguaglianze:

$$B = B * \mathbb{I}_n = B * (A * C) = (B * A) * C = \mathbb{I}_n * C = C.$$

Naturalmente, se esistesse un'altra inversa destra $\tilde{C}$ (rispettivamente inversa sinistra $\tilde{B}$), per lo stesso ragionamento dovrebbe valere $\tilde{C} = B = C$ (rispettivamente $\tilde{B} = C = B$). Quindi B e C sono uniche. Per definizione, esse coincidono con la matrice inversa che a sua volta è unica. □

Come conseguenza di questo risultato, possiamo affermare che, quando esiste, è lecito riferirsi al singolare alla matrice inversa di A. Essa viene indicata con A^{-1}.

La matrice inversa gode di alcune proprietà elementari.

Proposizione 3.44. *Siano $A, B \in \mathrm{Mat}(n, n; \mathbb{K})$ invertibili. Allora l'operazione di inversione di una matrice:*

i) è un'involuzione, cioè $(A^{-1})^{-1} = A$;
ii) soddisfa $(A * B)^{-1} = B^{-1} * A^{-1}$;
iii) commuta con la trasposizione, ovvero $(A^{-1})^T = (A^T)^{-1}$.

Dimostrazione. Per ciascun punto dell'enunciato utilizziamo la proprietà di unicità dell'inversa:

i) segue direttamente dalla definizione di A^{-1};
ii) la seguente catena di uguaglianze

$$(B^{-1} * A^{-1}) * (A * B) = B^{-1} * (A^{-1} * A) * B = B^{-1} * \mathbb{I}_n * B = B^{-1} * B = \mathbb{I}_n$$

prova che $B^{-1} * A^{-1}$ è l'inversa sinistra di $A * B$. Analogamente si dimostra che essa è anche inversa destra e quindi coincide con l'inversa del prodotto delle due matrici;

iii) l'ultimo punto segue da $\mathbb{I}_n = \mathbb{I}_n^T = (A * A^{-1})^T = (A^{-1})^T * A^T$. □

Grazie alla proprietà associativa ed all'identità (3), è facile verificare che valgono anche le seguenti generalizzazioni delle proprietà precedenti:

(6) $(A_1 * \cdots * A_k)^{-1} = ((A_1 * \cdots * A_{k-1}) * A_k)^{-1} = A_k^{-1} * (A_1 * \cdots * A_{k-1})^{-1} = \cdots = A_k^{-1} * \cdots * A_1^{-1}$

e

(7) $((A_1 * \cdots * A_k)^{-1})^T = (A_k^{-1} * \cdots * A_1^{-1})^T = (A_1^{-1})^T * \cdots * (A_k^{-1})^T = (A_1^T)^{-1} * \cdots * (A_k^T)^{-1}$,

per ogni insieme $\{A_1, \ldots, A_k\}$ di matrici invertibili.

3.5.2. Matrici elementari e riduzione a scala. Per studiare le condizioni di esistenza della matrice inversa, è necessario introdurre le cosiddette matrici elementari.

Definizione 3.45. *Nell'insieme* $\mathrm{Mat}(n, n; \mathbb{K})$, *attraverso le tre operazioni elementari effettuate sulla matrice identità* $\mathbb{I}_n$, *si costruiscono le seguenti matrici chiamate elementari:*

i) la matrice di permutazione $P(i, j)$ *ottenuta scambiando le righe* $R(i)$ *e* $R(j)$ *di* $\mathbb{I}_n$;

ii) la matrice $T(i; t)$ *ottenuta moltiplicando la riga* $R(i)$ *di* $\mathbb{I}_n$ *per lo scalare* $t \in \mathbb{K}^*$;

iii) la matrice $T(i, j; t)$ *ottenuta sostituendo alla riga* $R(i)$ *di* $\mathbb{I}_n$ *la somma* $R(i) + t \cdot R(j)$, *con* $t \in \mathbb{K}$.

Esempio 3.46. In $\mathrm{Mat}(3, 3, \mathbb{K})$ abbiamo

$$P(1,2) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad T(2;3) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad T(2,1;-1) = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Proposizione 3.47. *Data* $A \in \mathrm{Mat}(m, n; \mathbb{K})$, *le operazioni elementari sulle righe sono equivalenti alla moltiplicazione a sinistra della corrispondente matrice elementare, ovvero:*

i) la permutazione delle righe $A_{R(i)}$ *e* $A_{R(j)}$ *è equivalente a* $P(i, j) * A$;

ii) la moltiplicazione della riga $A_{R(i)}$ *per uno scalare non nullo* $t \in \mathbb{K}^*$ *è equivalente a* $T(i; t) * A$;

iii) la sostituzione delle riga $A_{R(i)}$ *con* $A_{R(i)} + t \cdot A_{R(j)}$ *è equivalente a* $T(i, j; t) * A$.

Dimostrazione. La proposizione segue dall'applicazione del Lemma 3.9 ai singoli casi. Verifichiamo esplicitamente il primo, essendo gli altri analoghi.

$$(P(i,j) * A)_{R(k)} = \sum_{l=1}^{n} p(i,j)_{kl} \cdot A_{R(l)} = \begin{cases} \displaystyle\sum_{l=1}^{n} \delta_{jl} \cdot A_{R(l)} = A_{R(j)} & \text{se} \quad k = i \\ \displaystyle\sum_{l=1}^{n} \delta_{il} \cdot A_{R(l)} = A_{R(i)} & \text{se} \quad k = j \\ \displaystyle\sum_{l=1}^{n} \delta_{kl} \cdot A_{R(l)} = A_{R(k)} & \text{se} \quad k \neq i, j \end{cases} \qquad \square$$

Corollario 3.48. *È sempre possibile trasformare una matrice* $A \in \mathrm{Mat}(m, n; \mathbb{K})$ *in una matrice a scala* $S \in \mathrm{Mat}(m, n; \mathbb{K})$ *attraverso il prodotto a sinistra di una sequenza finita di matrici elementari.*

Dimostrazione. L'enunciato segue facilmente dal Teorema 3.16 e dalla Proposizione 3.47. $\square$

Grazie a questa rappresentazione della riduzione a scala, è semplice verificare che l'algoritmo di eliminazione è invertibile, nel senso che è sempre possibile partire dalla matrice ridotta S e tornare alla matrice A attraverso operazioni elementari. A tal fine, è necessario enunciare il seguente lemma, la cui verifica è lasciata al lettore come esercizio.

Lemma 3.49. *Le matrici elementari sono invertibili e le loro inverse sono ancora matrici elementari. In particolare:*

i) $P(i,j)^{-1} = P(i,j)$;

ii) $T(i;t)^{-1} = T(i;t^{-1})$;

iii) $T(i,j;t)^{-1} = T(i,j;-t)$.

Proposizione 3.50. *Sia* $S = E_1 * \cdots * E_k * A$ *una riduzione a scala della matrice* A *attraverso le matrici elementari* E_i, $i = 1, \ldots, k$. *Allora si può ottenere* A *attraverso operazioni elementari applicate ad* S *come* $A = E_k^{-1} * \cdots * E_1^{-1} * S$.

Dimostrazione. Grazie al Lemma 3.49, sappiamo che le matrici elementari sono invertibili, ovvero che $E_1^{-1}, \ldots, E_k^{-1}$ esistono. Dall'identità (6) segue inoltre che

$$(E_1 * \cdots * E_k)^{-1} = E_k^{-1} * \cdots * E_1^{-1}.$$

Di conseguenza:

$$E_k^{-1} * \cdots * E_1^{-1} * S = (E_1 * \cdots * E_k)^{-1} * (E_1 * \cdots * E_k) * A = A. \qquad \square$$

La Proposizione 3.50 può essere riformulata dicendo che per ogni matrice esiste sempre una decomposizione della stessa come prodotto di k matrici elementari ed una matrice a scala.

Esempio 3.51. Riprendiamo la riduzione a scala dell'Esempio 3.19. Dalla Proposizione 3.47 abbiamo

$$A = \begin{bmatrix} 2 & 0 & 2 & 0 \\ 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \end{bmatrix}, \quad S = T(2,1;-1) * T(3,2;-2) * T(1;\frac{1}{2}) * A.$$

Infatti

$$\begin{aligned} S &= \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} * \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} * A = \\ &= \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} * \begin{bmatrix} 2 & 0 & 2 & 0 \\ 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

A questo punto verifichiamo la Proposizione 3.50:

$$T(1;2) * T(3,2;2) * T(2,1;1) * S = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 2 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 2 & 0 \\ 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \end{bmatrix} = A.$$

3.5.3. Condizione sufficiente di esistenza dell'inversa ed algoritmo di Gauss-Jordan.

In questa sottosezione utilizzeremo i risultati precedenti sulle matrici elementari, al fine di provare un verso del Teorema 3.42, ovvero la condizione sufficiente per l'invertibilità di una matrice.

Lemma 3.52. *È possibile ridurre una matrice* $A \in \mathrm{Mat}(n,n;\mathbb{K})$ *alla matrice identità* $\mathbb{I}_n$ *attraverso operazioni elementari sulle righe se e soltanto se* $r(A) = n$.

Dimostrazione. Se $r(A) = n$, allora la riduzione a scala S di A ottenuta utilizzando l'Algoritmo 1 avrà un pivot per ogni elemento della diagonale. Naturalmente, possiamo proseguire oltre nell'applicazione di operazioni elementari alla matrice S. In particolare, possiamo dividere ogni sua riga per il valore del pivot corrispondente, così da ottenere 1 in ogni elemento della diagonale. Infine, attraverso operazioni del terzo tipo, possiamo porre a 0 tutti gli elementi posti sopra ciascun pivot, così da ottenere la matrice $\mathbb{I}_n$. Chiaramente, queste operazioni non sono possibili se $r(A) < n$. $\square$

Corollario 3.53. *Una matrice $A \in \mathrm{Mat}(n, n; \mathbb{K})$ è il prodotto di matrici elementari se e solo se il suo rango è n.*

DIMOSTRAZIONE. Sia $A = E_1 * \cdots * E_k$, dove E_i sono matrici elementari per $i = 1, \ldots, k$. Essendo il prodotto di matrici invertibili, dall'identità (6) segue che anche A è invertibile, con matrice inversa $A^{-1} = (E_1 * \cdots * E_k)^{-1} = E_k^{-1} * \cdots * E_1^{-1}$. Questo significa che $E_k^{-1} * \cdots * E_1^{-1} * A = \mathbb{I}_n$, ovvero la matrice A è riducibile alla matrice identità. Per il lemma precedente, segue quindi che $r(A) = n$. Partendo dal fondo e rileggendo al contrario il ragionamento precedente, si dimostra anche l'implicazione inversa. $\square$

Corollario 3.54. *Una matrice $A \in \mathrm{Mat}(n, n; \mathbb{K})$ di rango n è invertibile.*

DIMOSTRAZIONE. L'enunciato segue direttamente dalla dimostrazione del Corollario 3.53. $\square$

Esempio 3.55.

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \\ 0 & 1 & 4 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 2 \\ 0 & 1 & 4 \end{bmatrix} \xrightarrow{R_3 - R_2 \to R_3} \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix} \xrightarrow[R_1 - R_2 \to R_1]{\frac{1}{2} R_3 \to R_3} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow[R_1 - R_3 \to R_1]{R_2 - 2R_3 \to R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbb{I}_3.$$

Di conseguenza

$$\mathbb{I}_3 = T(1, 3; -1) * T(2, 3; -2) * T(1, 2; -1) * T\left(3; \frac{1}{2}\right) * T(3, 2; -1) * P(1, 2) * A$$

e quindi

$$A = P(1, 2) * T(3, 2; 1) * T(3; 2) * T(1, 2; 1) * T(2, 3; 2) * T(1, 3; 1).$$

Per il Corollario 3.54, la matrice inversa di A è

$$A^{-1} = T(1, 3; -1) * T(2, 3; -2) * T(1, 2; -1) * T\left(3; \frac{1}{2}\right) * T(3, 2; -1) * P(1, 2) = \begin{bmatrix} -1/2 & 1 & -1/2 \\ 2 & 0 & -1 \\ -1/2 & 0 & 1/2 \end{bmatrix}.$$

Osservazione 3.56. Come abbiamo visto, il Corollario 3.54 fornisce una tecnica per il calcolo della matrice inversa. Nella pratica non si utilizza quasi mai direttamente tale approccio, quanto piuttosto il seguente basato sulle operazioni elementari associate ad $A^{-1} = E_1^{-1} * \cdots * E_k^{-1}$. Più precisamente, consideriamo la matrice $[A|\mathbb{I}_n]$. Grazie al Lemma 3.9 abbiamo

$$A^{-1} * [A|\mathbb{I}_n] = [A^{-1} * A|A^{-1} * \mathbb{I}_n] = [\mathbb{I}_n|A^{-1}]$$

e di conseguenza

$$[\mathbb{I}_n|A^{-1}] = E_1^{-1} * \cdots * E_k^{-1} * [A|\mathbb{I}_n].$$

Quindi è possibile ottenere la matrice inversa di A compiendo la riduzione $[A|\mathbb{I}_n] \to [\mathbb{I}_n|A^{-1}]$, applicando cioè all'intera matrice composta $[A|\mathbb{I}_n]$ le operazioni elementari che riducono A all'identità. In questa forma, questo metodo di inversione è noto come algoritmo di Gauss-Jordan.

Esempio 3.57.

$$[A|\mathbb{I}_2] = \left[\begin{array}{cc|cc} -1 & 2 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & -1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \end{array}\right] = [\mathbb{I}_2|A^{-1}].$$

Tramite calcolo esplicito possiamo verificare che $A * A^{-1} = A^{-1} * A = \mathbb{I}_2$.

3.5.4. Condizione necessaria di esistenza dell'inversa. Per completare la dimostrazione del Teorema 3.42 rimane da verificare che $r(A) = n$ è anche condizione necessaria per l'invertibilità della matrice A. A tal fine andiamo a studiare meglio il legame tra matrici invertibili e sistemi lineari.

Lemma 3.58. *Sia $A \in \mathrm{Mat}(n, n; \mathbb{K})$ con $r(A) < n$. Allora esiste $B \in \mathrm{Mat}(n, 1; \mathbb{K})$ tale che il sistema lineare associato ad $[A|B]$ non ha soluzione.*

Dimostrazione. Se $r(A) < n$, allora a seguito della riduzione a scala della matrice esiste una riga $A_{R(i)}$ che viene trasforma nella riga nulla 0_{1n}. Scegliamo la matrice colonna B tale che $(B)_{j1} = \delta_{ij}$ per $1 \leq j \leq n$. Nella riduzione a scala di $[A|B]$, la i-esima riga viene trasformata in $[0_{1n}|c]$ con $c \neq 0$. Di conseguenza $r([A|B]) > r(A)$ e quindi il sistema non ha soluzione. □

Proposizione 3.59. *Sia $A \in \mathrm{Mat}(n, n; \mathbb{K})$ con $r(A) < n$. Allora non esistono le matrici inversa destra ed inversa sinistra e quindi A non è invertibile.*

Dimostrazione. Supponiamo per assurdo che esista C inversa sinistra di A. Consideriamo il sistema lineare omogeneo $[A|0_{n1}]$ e sia X una sua soluzione, che sappiamo esistere grazie al Corollario 3.25. Allora

$$X = \mathbb{I}_n * X = (C * A) * X = C * (A * X) = C * 0_{n1} = 0_{n1}.$$

Questo significa che l'unica soluzione possibile è quella banale, ma ciò è impossibile in quanto contraddice il teorema di Rouchè-Capelli.

Ancora per assurdo, supponiamo che esista D inversa sinistra di A. Consideriamo un sistema $[A|B]$ privo di soluzione, che sappiamo esistere grazie al Lemma 3.58. Allora

$$A * (D * B) = (A * D) * B = \mathbb{I}_n * B = B,$$

cioè $D * B$ è soluzione del sistema. Ancora una volta questo è impossibile, in quanto ciò contraddice la nostra ipotesi. □

3.5.5. Osservazioni finali e teorema di Cramer. La seguente proposizione, che riassume i risultati della sezione.

Proposizione 3.60. *Data $A \in \mathrm{Mat}(n, n; \mathbb{K})$, le seguenti affermazioni sono equivalenti:*

i) esiste l'inversa sinistra di A;
ii) esiste l'inversa destra di A;
iii) la matrice A è invertibile;
iv) il rango di A è massimo.

Dimostrazione. Il Teorema 3.42 garantisce l'equivalenza degli enunciati (iii) e (iv). Se (iii) è vero, inoltre per definizione sono veri (i) e (ii). Infine, la Proposizione 3.59 ha come conseguenza che se (i) oppure (ii) sono veri allora deve valere anche (iv). Componendo opportunamente queste implicazioni, abbiamo quindi che la validità di ogni punto dell'enunciato è sufficiente a dimostrare tutte le altre. □

Concludiamo la sezione presentando una prima applicazione dell'esistenza della matrice inversa, utile per lo studio dei sistemi lineari. Infatti, adattando al caso in cui $r(A) = n$ i ragionamenti utilizzati nella dimostrazione della Proposizione 3.59, si arriva facilmente a provare il seguente risultato, noto come teorema di Cramer.

Corollario 3.61. *Consideriamo il sistema lineare definito dalla matrice completa* $[A|B] \in \mathrm{Mat}(n+1, n; \mathbb{K})$. *Se A è invertibile, allora il sistema ammette l'unica soluzione* $X = A^{-1} * B$.

Esempio 3.62. Sia

$$[A|B] = \left[\begin{array}{cc|c} -1 & 2 & 1 \\ 1 & -1 & 1 \end{array}\right].$$

Sfruttando i risultati contenuti nell'Esempio 3.57, otteniamo che il sistema lineare $A * X = B$ ha unica soluzione

$$X = A^{-1} * B = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} * \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

3.6. Traccia di una matrice

Le due più importanti funzioni definite sull'insieme delle matrici quadrate di ordine n a valori in $\mathbb{K}$ sono la traccia ed il determinante. Per ogni matrice esse forniscono degli scalari contenenti informazioni significative sulla matrice stessa. In questa sezione ci dedichiamo allo studio della traccia.

Definizione 3.63. *La traccia è la funzione*

$$\begin{array}{rccc} \mathrm{Tr}: & \mathrm{Mat}(n, n; \mathbb{K}) & \longrightarrow & \mathbb{K} \\ & A & \longmapsto & \displaystyle\sum_{i=1}^{n} a_{ii} \end{array}.$$

Ad esempio

$$\mathrm{Tr}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a + d.$$

Proposizione 3.64. *Per ogni $A, B \in \mathrm{Mat}(n, n; \mathbb{K})$ ed $t_1, t_2 \in \mathbb{K}$, la traccia soddisfa le proprietà di:*

i) linearità $\mathrm{Tr}(t_1 \cdot A + t_2 \cdot B) = t_1 \,\mathrm{Tr}(A) + t_2 \,\mathrm{Tr}(B)$;
ii) invarianza rispetto alla trasposizione $\mathrm{Tr}(A^T) = \mathrm{Tr}(A)$;
iii) ciclicità $\mathrm{Tr}(A * B) = \mathrm{Tr}(B * A)$.

Dimostrazione. Le prime due proprietà sono immediate. Verifichiamo la terza:

$$\mathrm{Tr}(A * B) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} b_{ji} \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} b_{ji} a_{ij} \right) = \mathrm{Tr}(B * A).$$

□

La proprietà di ciclicità è così chiamata poiché, insieme alla proprietà associativa del prodotto di matrici, implica che $\mathrm{Tr}(A_1 * \cdots * A_{k-1} * A_k) = \mathrm{Tr}(A_k * A_1 * \cdots * A_{k-1})$. È però importante osservare che l'uguaglianza non è vera per una generica permutazione dei fattori. Ad esempio, in generale $\mathrm{Tr}(A_1 * \cdots * A_{k-1} * A_k) \neq \mathrm{Tr}(A_1 * \cdots * A_k * A_{k-1})$.

3.7. Determinante: definizione e relazione con il rango

La definizione e lo studio delle proprietà del determinante richiedono un impegno decisamente maggiore rispetto alla traccia.

Definizione 3.65. *Data* $A \in \mathrm{Mat}(m, n; \mathbb{K})$, *la sottomatrice* $A_{\widehat{i_1 \ldots i_k}\,\widehat{j_1 \ldots j_l}} \in \mathrm{Mat}(m-k, n-l; \mathbb{K})$ *è la matrice con gli stessi elementi di* A *meno le righe* $A_{R(i_1)}, \ldots, A_{R(i_k)}$ *e le colonne* $A_{C(j_1)}, \ldots, A_{C(j_l)}$. *Nel caso non venga eliminata alcuna riga o colonna, indichiamo rispettivamente* $A_{\widehat{0}\,\widehat{j_1 \ldots j_l}}$ *e* $A_{\widehat{i_1 \ldots i_k}\,\widehat{0}}$.

Esempio 3.66.

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \\ 0 & 1 & 4 \end{bmatrix} = A_{\widehat{0}\widehat{0}} \quad \Rightarrow \quad A_{\widehat{1}\widehat{1}} = \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix}, \; A_{\widehat{1}\widehat{2}} = \begin{bmatrix} 1 & 3 \\ 0 & 4 \end{bmatrix}, \; A_{\widehat{1}\widehat{3}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \; A_{\widehat{1\,2}\,\widehat{1\,2}} = [4].$$

Definizione 3.67. *Data* $A \in \mathrm{Mat}(n, n; \mathbb{K})$, *il determinante di* A *è la funzione*

$$\begin{array}{rccc} \det: & \mathrm{Mat}(n, n; \mathbb{K}) & \longrightarrow & \mathbb{K} \\ & A & \longmapsto & \det(A) = |A| \end{array}$$

definita iterativamente su n *nel seguente modo:*

- $n = 1: \ \det([a_{11}]) = a_{11}$;
- $n > 1: \ \det(A) = \sum_{i=1}^{n} a_{1i} C_{1i}$, *dove* $C_{ij} = (-1)^{i+j} \det(A_{\widehat{i}\widehat{j}})$ *è detto complemento algebrico di* a_{ij}.

Esempio 3.68.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \Rightarrow \quad \det(A) = a_{11}C_{11} + a_{12}C_{12} = a \cdot (-1)^{1+1} \det([d]) + b \cdot (-1)^{1+2} \det([c]) = ad - bc.$$

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \\ 0 & 1 & 4 \end{bmatrix} \quad \Rightarrow \quad \det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} = 0 \cdot 1 + 1 \cdot (-4) + 2 \cdot 1 = -2.$$

$$C_{11} = (-1)^{1+1} \begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} = 1, \quad C_{12} = (-1)^{1+2} \begin{vmatrix} 1 & 3 \\ 0 & 4 \end{vmatrix} = -4, \quad C_{13} = (-1)^{1+3} \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1.$$

L'importanza del determinante è dovuta al seguente fondamentale teorema.

Teorema 3.69. *Sia* $A \in \mathrm{Mat}(n, n; \mathbb{K})$, *allora* $\det(A) = 0$ *se e solo se* $r(A) < n$.

Questo teorema ha diverse notevoli conseguenze, ad iniziare da quelle di seguito elencate.

Corollario 3.70. *Sia* $A \in \mathrm{Mat}(n, n; \mathbb{K})$, *allora:*

- *il sistema lineare associato ad* $[A|B] \in \mathrm{Mat}(n, n+1; \mathbb{K})$ *ha unica soluzione se e solo se* $\det(A) \neq 0$;
- A *è invertibile se e solo se* $\det(A) \neq 0$.

Dimostrazione. Il primo ed il secondo punto seguono dal Teorema 3.69 e rispettivamente dal teorema di Rouché-Capelli e dal Teorema 3.42. □

Per poter dimostrare il Teorema 3.69 dovremo enunciare e verificare diverse proprietà del determinante, al cui studio sarà riservato l'intero capitolo. In particolare, andremo a studiare il comportamento del determinante a seguito dell'esecuzione di operazioni elementari sulle righe di una matrice. Come avremo modo di vedere, durante tale investigazione otterremo anche dei metodi alternativi per il calcolo del determinante stesso, che saranno più efficienti dal punto di vista del costo computazionale, rispetto al calcolo tramite la definizione.

3.7.1. Determinante ed operazioni elementari sulle righe. Per prima cosa investighiamo l'operazione elementare di permutazione di due righe.

Lemma 3.71. *Data* $A \in \mathrm{Mat}(n, n; \mathbb{K})$, *sia* B *la matrice ottenuta permutando le prime due righe di* A. *Allora* $\det(B) = -\det(A)$.

DIMOSTRAZIONE. Per definizione abbiamo

$$\det(A) = \sum_{i=1}^{n} (-1)^{1+i}\, a_{1i} \det\left(A_{\widehat{1}\,\widehat{i}}\right). \tag{8}$$

A sua volta

$$\det\left(A_{\widehat{1}\,\widehat{i}}\right) = \begin{cases} \displaystyle\sum_{j=2}^{n} (-1)^{j}\, a_{2j} \det\left(A_{\widehat{1}\,\widehat{2}\,\widehat{1}\,\widehat{j}}\right) & \text{se} \quad i = 1, \\ \displaystyle\sum_{j=1}^{i-1} (-1)^{1+j}\, a_{2j} \det\left(A_{\widehat{1}\,\widehat{2}\,\widehat{j}\,\widehat{i}}\right) + \sum_{j=i+1}^{n} (-1)^{j}\, a_{2j} \det\left(A_{\widehat{1}\,\widehat{2}\,\widehat{i}\,\widehat{j}}\right) & \text{se} \quad 1 < i < n, \\ \displaystyle\sum_{j=1}^{n-1} (-1)^{1+j}\, a_{2j} \det\left(A_{\widehat{1}\,\widehat{2}\,\widehat{j}\,\widehat{n}}\right) & \text{se} \quad i = n. \end{cases} \tag{9}$$

Dopo aver sostituito (9) in (8), raccogliamo i coefficienti che moltiplicano $\det\left(A_{\widehat{1}\,\widehat{2}\,\widehat{i}\,\widehat{j}}\right)$ nell'espressione risultante di $\det(A)$:

$$\det(A) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (-1)^{1+i+j} (a_{1i}a_{2j} - a_{1j}a_{2i}) \det\left(A_{\widehat{1}\,\widehat{2}\,\widehat{i}\,\widehat{j}}\right).$$

Analogamente

$$\begin{aligned} \det(B) &= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (-1)^{1+i+j} (b_{1i}b_{2j} - b_{1j}b_{2i}) \det\left(B_{\widehat{1}\,\widehat{2}\,\widehat{i}\,\widehat{j}}\right) = \\ &= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (-1)^{1+i+j} (a_{2i}a_{1j} - a_{2j}a_{1i}) \det\left(A_{\widehat{1}\,\widehat{2}\,\widehat{i}\,\widehat{j}}\right), \end{aligned}$$

dove nella seconda uguaglianza abbiamo usate le proprietà della matrice B. Confrontando le due formule segue la tesi. □

Proposizione 3.72. *Data* $A \in \mathrm{Mat}(n, n; \mathbb{K})$, *sia* B *la matrice ottenuta permutando una qualsiasi coppia di righe di* A. *Allora* $\det(B) = -\det(A)$.

DIMOSTRAZIONE. La dimostrazione viene compiuta per induzione su n. Se $n = 2$ abbiamo un caso particolare del Lemma 3.71. Supponiamo ora vera la proposizione fino ad $n - 1$, e sia B ottenuta dallo scambio di $A_{R(i)}$ ed $A_{R(j)}$, $1 \le i < j \le n$. Esistono tre casi differenti:

- se $i = 1, j = 2$ allora l'enunciato è equivalente al Lemma 3.71;
- se $j > i > 1$ allora

$$\det(B) = \sum_{k=1}^{n} (-1)^{1+k}\, b_{1k} \det\left(B_{\widehat{1}\,\widehat{k}}\right) = \sum_{k=1}^{n} (-1)^{1+k}\, a_{1k} \left(-\det\left(A_{\widehat{1}\,\widehat{k}}\right)\right) = -\det(A),$$

  dove nella seconda uguaglianza abbiamo utilizzato l'ipotesi induttiva;
- se $i = 1, j > 2$ allora lo scambio delle due righe si può ottenere scambiando in successione $A_{R(2)}$ ed $A_{R(j)}$, quindi $A_{R(1)}$ ed $A_{R(j)}$ ed infine $A_{R(1)}$ ed $A_{R(2)}$. Ogni scambio ricade in uno dei due casi precedenti e pertanto comporta un cambio di segno del determinante. Componendo i tre scambi si ha $\det(B) = -\det(A)$. □

Corollario 3.73. *Se una matrice $A \in \mathrm{Mat}(n, n; \mathbb{K})$ ha due righe uguali allora* $\det(A) = 0$.

Dimostrazione. L'enunciato è conseguenza della Proposizione 3.72, in quanto permutando le due righe uguali otteniamo $\det(A) = -\det(A)$. $\square$

Concentriamoci ora sulle due rimanenti operazioni elementari, di secondo e terzo tipo. A tal fine, abbiamo preliminarmente bisogno del seguente risultato, noto come sviluppo di Laplace per righe del determinante. Questo teorema ci permette anche di eliminare il ruolo particolare assegnato dalla Definizione 3.67 alla prima riga di una matrice.

Proposizione 3.74. *Data $A \in \mathrm{Mat}(n, n; \mathbb{K})$, per ogni $i \in \{1, \ldots, n\}$ si ha*

$$\det(A) = \sum_{j=1}^{n} a_{ij} C_{ij}.$$

Dimostrazione. La matrice

$$B = \begin{bmatrix} A_{R(i)} \\ A_{R(1)} \\ \vdots \\ A_{R(n)} \end{bmatrix} \in \mathrm{Mat}(n, n; \mathbb{K})$$

si può ottenere dalla matrice A permutando la riga $A_{R(i)}$ con le precedenti $i-1$ righe. Di conseguenza

$$\det(B) = \sum_{j=1}^{n} a_{ij} (-1)^{1+j} \det\left(A_{\widehat{i}\,\widehat{j}}\right) = (-1)^{i-1} \det(A)$$

da cui

$$\det(A) = \sum_{j=1}^{n} a_{ij} (-1)^{i+j} \det\left(A_{\widehat{i}\,\widehat{j}}\right) = \sum_{j=1}^{n} a_{ij} C_{ij}.$$ $\square$

Esempio 3.75.

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \\ 0 & 1 & 4 \end{bmatrix} \quad \Rightarrow \quad \det(A) = a_{21}C_{21} + a_{22}C_{22} + a_{23}C_{23} = 1 \cdot (-2) + 1 \cdot 0 + 3 \cdot 0 = -2.$$

$$C_{21} = (-1)^{2+1} \begin{vmatrix} 1 & 2 \\ 1 & 4 \end{vmatrix} = -2, \quad C_{22} = (-1)^{2+2} \begin{vmatrix} 0 & 2 \\ 0 & 4 \end{vmatrix} = 0, \quad C_{23} = (-1)^{2+3} \begin{vmatrix} 0 & 1 \\ 0 & 1 \end{vmatrix} = 0.$$

Lemma 3.76. *Date $A \in \mathrm{Mat}(n, n; \mathbb{K})$, $t \in \mathbb{K}$ e $\tilde{A} \in \mathrm{Mat}(1, n; \mathbb{K})$, siano B e C le matrici ottenute da A sostituendo alla riga $A_{R(i)}$ rispettivamente le righe $\tilde{A}$ e $t \cdot A_{R(i)} + \tilde{A}$. Allora*

$$\det(C) = t \det(A) + \det(B).$$

Dimostrazione. Utilizzando lo sviluppo di Laplace, abbiamo

$$\det(C) = \sum_{k=1}^{n} c_{ik} (-1)^{i+k} \det\left(C_{\widehat{i}\,\widehat{k}}\right) = \sum_{k=1}^{n} (t\, a_{ik} + \tilde{a}_{1k}) (-1)^{i+k} \det\left(A_{\widehat{i}\,\widehat{k}}\right) = t \det(A) + \det(B).$$ $\square$

Corollario 3.77. *Data $A \in \mathrm{Mat}(n, n; \mathbb{K})$, sia B una matrice ottenuta a seguito di operazioni elementari del secondo e terzo tipo. Allora:*

i) se $A_{R(i)}$ è sostituita da $t \cdot A_{R(i)}$ con $t \in \mathbb{K}^*$, vale $\det(B) = t \det(A)$;

ii) se $A_{R(i)}$ è sostituita da $A_{R(i)} + t \cdot A_{R(j)}$ con $t \in \mathbb{K}$, $j \neq i$, vale $\det(B) = \det(A)$.

Dimostrazione. La tesi segue dall'applicazione del Lemma 3.76 e, per il secondo punto, del Corollario 3.73. $\square$

3.7.2. Determinante e riduzione a scala. Grazie all'analisi compiuta nella precedente sottosezione, segue immediatamente il seguente importante risultato.

Proposizione 3.78. *Data* $A \in \mathrm{Mat}(n,n;\mathbb{K})$, *sia* $S \in \mathrm{Mat}(n,n;\mathbb{K})$ *una sua riduzione a scala ottenuta attraverso* k *permutazioni di righe ed* r *moltiplicazioni di righe di* A *per gli scalari* $t_i \in \mathbb{K}^*$, $i = 1,\dots,r$. *Allora*

$$\det(S) = (-1)^k\, t_1 \dots t_r \det(A).$$

Dal punto di vista applicativo, la Proposizione 3.78 fornisce un metodo per calcolare il determinante di una matrice A attraverso il calcolo di una sua matrice ridotta S. Si può dimostrare che tale tecnica è computazionalmente più efficiente rispetto al calcolo basato sugli sviluppi di Laplace, anche in considerazione del fatto che il determinante di S è ottenibile in modo molto semplice. Infatti, poiché una matrice quadrata a scala è necessariamente triangolare alta, su S possiamo applicare il seguente risultato.

Proposizione 3.79. *Sia* $A \in \mathrm{Mat}(n,n;\mathbb{K})$ *una matrice triangolare, allora* $\det(A) = \prod_{i=1}^n a_{ii}$.

Dimostrazione. La dimostrazione si svolge per induzione su n. L'enunciato è ovviamente vero se $n = 1$. Supponiamo ora che sia vero fino ad $n-1$. Abbiamo due casi:

- se $A \in \mathbb{T}_a(n;\mathbb{K})$ allora
$$\det(A) = \sum_{j=1}^n a_{nj}\,(-1)^{n+j} \det\left(A_{\widehat{n}\,\widehat{j}}\right) = a_{nn} \det\left(A_{\widehat{n}\,\widehat{n}}\right) = a_{nn}\left(\prod_{i=1}^{n-1} a_{ii}\right) = \prod_{i=1}^n a_{ii};$$
- se $A \in \mathbb{T}_b(n;\mathbb{K})$ allora
$$\det(A) = \sum_{j=1}^n a_{1j}\,(-1)^{1+j} \det\left(A_{\widehat{1}\,\widehat{j}}\right) = a_{11} \det\left(A_{\widehat{1}\,\widehat{1}}\right) = a_{11}\left(\prod_{i=2}^{n} a_{ii}\right) = \prod_{i=1}^n a_{ii}.$$

In entrambe le situazioni abbiamo utilizzato l'ipotesi induttiva nella penultima uguaglianza. □

Corollario 3.80. $\det(\mathbb{I}_n) = 1$.

Esempio 3.81. Nell'Esempio 3.55 abbiamo

$$\mathbb{I}_3 = T(1,3;-1) * T(2,3;-2) * T(1,2;-1) * T\left(3;\frac{1}{2}\right) * T(3,2;-1) * P(1,2) * A,$$

quindi $k = 1$ e $t_1 = \frac{1}{2}$. Di conseguenza

$$1 = \det(\mathbb{I}_3) = (-1)^1 \cdot \frac{1}{2} \cdot \det(A) \quad \Rightarrow \quad \det(A) = -2.$$

Grazie ai risultati contenuti in questa sottosezione, possiamo finalmente completare la dimostrazione del Teorema 3.69.

Dimostrazione del Teorema 3.69. Date $A \in \mathrm{Mat}(n,n;\mathbb{K})$ ed S una sua riduzione a scala, dalla Proposizione 3.78 segue che $\det(A) = 0$ se e solo se $\det(S) = 0$. Ma dalla Proposizione 3.79 questo è possibile se e solo se $r(S) < n$, da cui segue la tesi. □

Osservazione 3.82. Per completezza, osserviamo che è possibile dimostrare che il determinante è l'unica funzione che rispetto alle operazioni elementari si comporti nel modo descritto nella Sottosezione 3.7.1 e che in aggiunta soddisfi il Corollario 3.80. In alcuni testi di algebra lineare di taglio più matematico, questa proprietà viene infatti presa come definizione stessa di determinante.

3.8. Determinante ed algebra delle matrici

In Algebra è sempre molto importante studiare il comportamento delle funzioni definite su un insieme rispetto alle operazioni definenti la struttura associata a tale insieme. Normalmente, le funzioni più significative sono quelle che presentano un qualche comportamento peculiare. In questa sezione analizzeremo per prima cosa il comportamento del determinante rispetto alle operazione definite su $\mathrm{Mat}(n,n;\mathbb{K})$ e rispetto alla trasposizione. Successivamente investigheremo la relazione tra determinante ed inversione di matrici.

3.8.1. Determinante, operazioni tra matrici e trasposizione. In primo luogo, è facile verificare attraverso controesempi che il determinante non soddisfa la linearità. In particolare, il determinante non è una funzione additiva, ovvero $\det(A+B) \neq \det(A)+\det(B)$. Ad esempio, se $A, B \in \mathrm{Mat}(2,2;\mathbb{K})$:

$$\begin{aligned}
\det(A+B) &= \begin{vmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{vmatrix} \\
&= (a_{11}a_{22}-a_{12}a_{21})+(b_{11}b_{22}-b_{12}b_{21})+(a_{11}b_{22}+b_{11}a_{22}-a_{12}b_{21}-b_{12}a_{21}) \\
&= \det(A)+\det(B)+\begin{vmatrix} a_{11} & a_{12} \\ b_{21} & b_{22} \end{vmatrix}+\begin{vmatrix} b_{11} & b_{12} \\ a_{21} & a_{22} \end{vmatrix}.
\end{aligned}$$

Inoltre, data $A \in \mathrm{Mat}(n,n;\mathbb{K})$ e lo scalare $t \in \mathbb{K}$, utilizzando il Corollario 3.77 si può dimostrare facilmente che vale $\det(t\cdot A) = t^n \det(A)$.

Il determinante ha un comportamento decisamente migliore rispetto al prodotto righe per colonne, come enunciato nel seguente fondamentale teorema dovuto a Binét.

Teorema 3.83. *Il determinante è una funzione moltiplicativa, ovvero date* $A, B \in \mathrm{Mat}(n,n;\mathbb{K})$ *si ha* $\det(A*B) = \det(A)\det(B)$.

Dimostrazione. Supponiamo che valga $r(A) = r(B) = n$. Per il Lemma 3.52, esistono $E_1, \ldots, E_p$ ed $\tilde{E}_1, \ldots, \tilde{E}_q$ matrici elementari tali che

$$\mathbb{I}_n = E_1 * \cdots * E_p * A \qquad \text{ed} \qquad \mathbb{I}_n = \tilde{E}_1 * \cdots * \tilde{E}_q * B.$$

Per la Proposizione 3.78, abbiamo

$$1 = \det(\mathbb{I}_n) = (-1)^k\, t_1 \ldots t_r \det(A) \qquad \text{ed} \qquad 1 = \det(\mathbb{I}_n) = (-1)^l\, \tilde{t}_1 \ldots \tilde{t}_s \det(B),$$

dove k è il numero di permutazioni e $t_1, \ldots, t_r$ sono i coefficienti delle operazioni elementari del secondo tipo associate ad $E_1, \ldots, E_p$, ed analogamente per l e $\tilde{t}_1, \ldots, \tilde{t}_s$ rispetto ad $\tilde{E}_1, \ldots, \tilde{E}_q$. Quindi,

$$\det(A)\det(B) = (-1)^{k+l}\left(t_1 \ldots t_r\, \tilde{t}_1 \ldots \tilde{t}_s\right)^{-1}.$$

D'altra parte, abbiamo

$$A = E_p^{-1} * \cdots * E_1^{-1},\ B = \tilde{E}_q^{-1} * \cdots * \tilde{E}_1^{-1} \qquad \Rightarrow \qquad A*B = E_p^{-1} * \cdots * E_1^{-1} * \tilde{E}_q^{-1} * \cdots * \tilde{E}_1^{-1}$$

e quindi

$$\mathbb{I}_n = \tilde{E}_1 * \cdots * \tilde{E}_q * E_1 * \cdots * E_p * (A*B).$$

Questo significa che $\tilde{E}_1, \ldots, \tilde{E}_q, E_1, \ldots, E_p$ sono associate alla riduzione della matrice prodotto $A*B$ alla matrice $\mathbb{I}_n$. Di conseguenza

$$1 = \det(\mathbb{I}_n) = (-1)^{k+l}\, t_1 \ldots t_r\, \tilde{t}_1 \ldots \tilde{t}_s\, \det(A*B).$$

Dal confronto tra le due espressioni ottenute segue la tesi $\det(A * B) = \det(A)\det(B)$.

Se invece la matrice A, oppure B, non ha rango massimo, allora $\det(A)\det(B) = 0$. In questo caso il teorema vale se e solo se anche $\det(A * B) = 0$, cioè se e solo se $r(A * B) < n$. Supponiamo che $r(A) < n$ e sia $S = E_1 * \cdots * E_p * A$ una sua riduzione a scala. Allora S ha almeno l'ultima riga nulla, per cui anche il prodotto $S * B$ avrà l'ultima riga nulla. Questo significa che $r(S * B) < n$. Sia $\tilde{S}$ una riduzione a scala di $S * B$, allora

$$\tilde{S} = \tilde{E}_1 * \cdots * \tilde{E}_q * (S * B) = \tilde{E}_1 * \cdots * \tilde{E}_q * E_1 * \cdots * E_p * (A * B).$$

Di conseguenza, $\tilde{S}$ è anche una riduzione a scala di $A * B$. Poiché $r(\tilde{S}) = r(S * B) < n$, segue che $r(A * B) < n$. Supponiamo invece $r(B) < n$. Allora per il Teorema di Rouché-Capelli esiste $C \in \mathrm{Mat}(n, 1; \mathbb{K})$ non nulla tale che $B * C = 0_{n1}$. Di conseguenza $(A * B) * C = A * (B * C) = A * 0_{n1} = 0_{n1}$, pertanto $r(A * B) < n$. $\square$

Grazie al teorema di Binét possiamo anche studiare il comportamento del determinante rispetto all'operazione di trasposizione.

Proposizione 3.84. *Sia* $A \in \mathrm{Mat}(n, n; \mathbb{K})$, *allora* $\det(A) = \det(A^T)$.

Dimostrazione. A causa della Proposizione 3.79, l'enunciato è vero per tutte le matrici triangolari e quindi in particolare per le matrici elementari $T(i; t)$ e $T(i, j; t)$. D'altra parte, la proposizione è ovviamente vera per le matrici di permutazione, dato che $P(i, j)^T = P(i, j)$. Sia ora A una matrice generica ed $S = E_1 * \cdots * E_p * A$ una sua riduzione a scala. Grazie al teorema di Binét abbiamo

$$\det(S) = \det(E_1) \dots \det(E_p) \det(A).$$

Trasponendo il prodotto di matrici, otteniamo $S^T = A^T * E_p^T * \cdots * E_1^T$ e quindi

$$\det(S^T) = \det(A^T) \det(E_p^T) \dots \det(E_1^T) = \det(E_1) \dots \det(E_p) \det(A^T).$$

Dato che S è una matrice triangolare, possiamo uguagliare le due espressioni precedenti:

$$\det(E_1) \dots \det(E_n) \det(A) = \det(E_1) \dots \det(E_n) \det(A^T).$$

Osserviamo infine che le matrici elementari hanno rango massimo per costruzione, quindi il loro determinante è diverso da 0. Di conseguenza $\det(A) = \det(A^T)$. $\square$

Dalla proposizione precedente seguono infine due corollari molto utili. In particolare, il risultato contenuto nel Corollario 3.86, noto come sviluppo di Laplace per colonne, stabilisce la totale simmetria tra righe e colonne nel calcolo del determinante di una matrice.

Corollario 3.85. $\det(A) = 0$ *se:*

i) una riga o una colonna di A è identicamente nulla;

ii) due righe o due colonne di A sono uguali.

Dimostrazione. In entrambi i casi il rango della matrice A o della sua trasposta è minore di n. Pertanto il determinante è nullo per il Teorema 3.69. $\square$

Corollario 3.86 (Sviluppo di Laplace per colonne)**.** *Data* $A \in \mathrm{Mat}(n, n; \mathbb{K})$, *per ogni* $j \in \{1, \dots, n\}$ *vale*

$$\det(A) = \sum_{i=1}^{n} a_{ij} C_{ij}.$$

Dimostrazione. La tesi segue dallo sviluppo di Laplace per righe e dalla Proposizione 3.84. $\square$

Esempio 3.87.

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \\ 0 & 1 & 4 \end{bmatrix} \quad \Rightarrow \quad \det(A) = a_{11}C_{11} + a_{21}C_{21} + a_{31}C_{31} = 0 \cdot 1 + 1 \cdot (-2) + 0 \cdot 1 = -2.$$

$$C_{11} = (-1)^{1+1} \begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} = 1, \quad C_{21} = (-1)^{2+1} \begin{vmatrix} 1 & 2 \\ 1 & 4 \end{vmatrix} = -2, \quad C_{31} = (-1)^{3+1} \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} = 1.$$

3.8.2. Determinante ed inversione di matrici. Come sappiamo dal Corollario 3.70, l'invertibilità o meno di una matrice dipende dal valore del suo determinante. In realtà, la relazione tra matrice inversa e determinante è ancora più precisa. Iniziamo ad investigarla partendo dal cosiddetto teorema di Laplace sulle matrici quadrate, che contiene e generalizza i risultati relativi agli sviluppi di Laplace per righe e per colonne.

Definizione 3.88. *Sia $A \in \mathrm{Mat}(n, n; \mathbb{K})$ e siano C_{ij} i suoi complementi algebrici, con $i, j = 1, \ldots, n$. Si dice aggiunta di A la matrice $A^* \in \mathrm{Mat}(n, n; \mathbb{K})$ i cui elementi sono $(A^*)_{ij} = C_{ji}$.*

Teorema 3.89. *Sia $A \in \mathrm{Mat}(n, n; \mathbb{K})$ e sia A^* la sua matrice aggiunta, allora vale*

$$A * A^* = A^* * A = \det(A) \cdot \mathbb{I}_n.$$

Dimostrazione. Gli elementi del primo prodotto sono

$$(A * A^*)_{ij} = \sum_{k=1}^{n} a_{ik}(A^*)_{kj} = \sum_{k=1}^{n} a_{ik}C_{jk}.$$

Se $j = i$ l'espressione precedente coincide esattamente con lo sviluppo di Laplace del determinante di A rispetto alla riga i, enunciato nel Proposizione 3.74. Se $j \neq i$, l'espressione è ancora interpretabile come un determinante, questa volta della matrice ottenuta attraverso la sostituzione della riga $A_{R(j)}$ in A con la riga $A_{R(i)}$. Di conseguenza, per il Corollario 3.73 tale determinante è nullo. Quindi

$$(A * A^*)_{ij} = \begin{cases} \det(A) & \text{se} \quad i = j, \\ 0 & \text{se} \quad i \neq j, \end{cases},$$

o equivalentemente $A * A^* = \det(A) \cdot \mathbb{I}_n$. La seconda identità si dimostra in maniera analoga utilizzando lo sviluppo di Laplace per colonne contenuto nel Corollario 3.86. □

La seguente formula di inversione di una matrice è conseguenza diretta del Teorema di Laplace.

Corollario 3.90. *Sia $A \in \mathrm{Mat}(n, n; \mathbb{K})$ invertibile, allora $A^{-1} = \frac{1}{\det(A)} \cdot A^*$.*

Esempio 3.91.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \Rightarrow \quad A^* = \begin{bmatrix} C_{11} & C_{21} \\ C_{12} & C_{22} \end{bmatrix} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad \Rightarrow \quad A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

L'utilità del teorema di inversione attraverso la matrice aggiunta è principalmente teorica, in quanto permette di scrivere una forma chiusa per la matrice inversa. D'altra parte, dal punto di vista computazionale questo metodo è molto più costoso dell'algoritmo di Gauss–Jordan, in quanto richiede il calcolo del determinante di n^2 matrici di tipo $(n-1, n-1)$. Di fatto questo procedimento è realizzabile solo per n non troppo grande.

Un secondo corollario del Teorema 3.89 è la cosiddetta regola di Cramer per la risoluzione dei sistemi lineari determinati.

Proposizione 3.92. *Data* $[A|B] \in \mathrm{Mat}(n+1, n; \mathbb{K})$ *la matrice completa di un sistema lineare con* $r(A) = n$, *sia* A_i *la matrice ottenuta da* A *sostituendo alla colonna* $A_{C(i)}$ *la colonna* B. *Se* X *è la colonna delle soluzioni, allora vale*

$$x_{i1} = \frac{\det(A_i)}{\det(A)}.$$

Dimostrazione. Poiché $r(A) = n$, vale il teorema di Cramer 3.61 e quindi

$$X = A^{-1} * B = \frac{1}{\det(A)} \cdot (A^* * B),$$

da cui

$$x_{i1} = \frac{1}{\det(A)} \sum_{j=1}^{n} C_{ji} b_{j1}.$$

Come abbiamo già osservato nella dimostrazione del Teorema 3.89, l'ultima sommatoria è interpretabile come il determinante della matrice ottenuta da A attraverso la sostituzione della colonna $A_{C(i)}$ con la colonna dei termini noti B. □

Analogamente al corollario precedente, il metodo di Cramer ha notevole utilità teorica nella dimostrazione di teoremi, ma quasi nessuna rilevanza pratica, poiché è estremamente pesante dal punto di visto del costo computazionale.

Concludiamo la sezione, ed il capitolo, con un teorema caratterizzante le proprietà algebriche dell'insieme delle matrici invertibili e del determinante.

Teorema 3.93. *L'insieme delle matrici* $A \in \mathrm{Mat}(n, n; \mathbb{K})$ *invertibili è un gruppo (non abeliano se* $n > 1$*) rispetto alla moltiplicazione. Esso viene indicato con* $GL(n; \mathbb{K})$ *ed è detto gruppo lineare generale di ordine* n *sul campo* $\mathbb{K}$. *Il determinante è un omomorfismo suriettivo di gruppi da* $GL(n; \mathbb{K})$ *in* $(\mathbb{K}^*, \cdot)$.

Dimostrazione. Bisogna in primo luogo dimostrare che l'insieme $GL(n; \mathbb{K})$ è chiuso rispetto alla moltiplicazione. A tal fine osserviamo che $A, B \in GL(n; \mathbb{K})$ se, e soltanto se, $\det(A)$ e $\det(B)$ non sono nulli. Ma allora, per il Teorema di Binét, abbiamo $\det(A * B) = \det(A) \det(B) \neq 0$ e quindi $A * B \in GL(n; \mathbb{K})$.

A questo punto verifichiamo le proprietà gruppali. Ovviamente $GL(n; \mathbb{K})$ contiene l'elemento neutro della moltiplicazione, cioè la matrice identità $\mathbb{I}_n$. Inoltre, se $A \in GL(n; \mathbb{K})$ allora per definizione esiste la sua inversa A^{-1}, che è a sua volta invertibile con inversa A e quindi $A^{-1} \in GL(n; \mathbb{K})$. Infine, la proprietà associativa è ereditata dall'associatività del prodotto tra tutte le matrici in $\mathrm{Mat}(n, n; \mathbb{K})$.

Relativamente all'ultimo punto dell'enunciato, dalla definizione di $GL(n; \mathbb{K})$ sappiamo che la sua immagine attraverso la funzione determinante è un sottoinsieme di $\mathbb{K}^*$. D'altra parte, ogni $t \in \mathbb{K}^*$ ha almeno una controimmagine nella matrice elementare $T(1; t) \in GL(n; \mathbb{K})$, quindi la funzione è suriettiva. Inoltre, il Teorema di Binét afferma che il determinante rispetta la struttura moltiplicativa dei due gruppi, e quindi è un omomorfismo. □

Corollario 3.94. *Sia* $A \in GL(n; \mathbb{K})$, *allora vale* $\det(A^{-1}) = (\det(A))^{-1}$.

Dimostrazione. La tesi si dimostra sfruttando la proprietà di omomorfismo del determinante:

$$1 = \det(\mathbb{I}_n) = \det(A * A^{-1}) = \det(A) \det(A^{-1}),$$

da cui segue $\det(A^{-1}) = \det(A)^{-1}$. □

CAPITOLO 4

Spazi vettoriali

In questo capitolo definiremo e studieremo in astratto la struttura algebrica di spazio vettoriale, che è la struttura fondamentale dell'algebra lineare. L'esempio più importante di spazio vettoriale è l'insieme delle matrici definito nel Capitolo 3.

4.1. Definizione, esempi, proprietà elementari ed omomorfismi

Definizione 4.1. *Sia V un insieme e $\mathbb{K}$ un campo, su cui sono definite un'operazione interna detta somma vettoriale*

$$\begin{array}{rccc} +: & V \times V & \longrightarrow & V \\ & (\mathbf{v_1}, \mathbf{v_2}) & \longmapsto & \mathbf{v_1} + \mathbf{v_2} \end{array}$$

ed un'operazione esterna detta prodotto per uno scalare

$$\begin{array}{rccc} \cdot: & \mathbb{K} \times V & \longrightarrow & V \\ & (t, \mathbf{v}) & \longmapsto & t \cdot \mathbf{v} \end{array}.$$

$(V, \mathbb{K}, +, \cdot)$ *si dice spazio vettoriale se:*

i) $(V, +)$ *è un gruppo abeliano, con elemento neutro ed inverso di* $\mathbf{v}$ *denotati rispettivamente* $\mathbf{0}$ *e* $-\mathbf{v}$*;*
ii) la moltiplicazione per uno scalare è distributiva rispetto alla somma vettoriale, cioè $t \cdot (\mathbf{v_1} + \mathbf{v_2}) = (t \cdot \mathbf{v_1}) + (t \cdot \mathbf{v_2})$ *per ogni* $t \in \mathbb{K}$ *e* $\mathbf{v_1}, \mathbf{v_2} \in V$*;*
iii) la moltiplicazione per uno scalare è distributiva rispetto alla somma in $\mathbb{K}$*, cioè* $(t_1 + t_2) \cdot \mathbf{v} = (t_1 \cdot \mathbf{v}) + (t_2 \cdot \mathbf{v})$ *per ogni* $t_1, t_2 \in \mathbb{K}$ *e* $\mathbf{v} \in V$*;*
iv) la moltiplicazione per uno scalare è omogenea rispetto alla moltiplicazione in $\mathbb{K}$*, cioè* $t_1 \cdot (t_2 \cdot \mathbf{v}) = (t_1 t_2) \cdot \mathbf{v}$ *per ogni* $t_1, t_2 \in \mathbb{K}$ *e* $\mathbf{v} \in V$*;*
v) la moltiplicazione per uno scalare soddisfa la proprietà di normalizzazione, ovvero $1 \cdot \mathbf{v} = \mathbf{v}$ *per ogni* $\mathbf{v} \in V$*.*

Gli elementi $\mathbf{v}$ *di* V *si dicono vettori.*

Esempio 4.2. Esistono alcuni esempi fondamentali di spazi vettoriali che utilizzeremo in tutto il corso.

- A seguito della Proposizione 3.7, la struttura $(\mathrm{Mat}(m, n; \mathbb{K}), \mathbb{K}, +, \cdot)$ è uno spazio vettoriale.
- Dato il campo $\mathbb{K}$, sull'insieme $\mathbb{K}^n$ delle n-uple di scalari sono naturalmente definite le due operazioni di somma e prodotto:

$$\begin{array}{rccc} +: & \mathbb{K}^n \times \mathbb{K}^n & \longrightarrow & \mathbb{K}^n \\ & ((a_1, \ldots, a_n), (b_1, \ldots, b_n)) & \longmapsto & (a_1 + b_1, \ldots, a_n + b_n) \end{array},$$

$$\begin{array}{rccc} \cdot: & \mathbb{K} \times \mathbb{K}^n & \longrightarrow & \mathbb{K}^n \\ & (t, (a_1, \ldots, a_n)) & \longmapsto & (ta_1, \ldots, ta_n) \end{array}.$$

  Le proprietà di spazio vettoriale sono conseguenza diretta delle proprietà di somma e prodotto in $\mathbb{K}$. In particolare $(\mathbb{K}, \mathbb{K}, +, \cdot)$ è uno spazio vettoriale in cui le due operazioni coincidono esattamente con quelle del campo.

- Consideriamo l'insieme dei polinomi in una variabile x ed a coefficienti nel campo $\mathbb{K}$
$$\mathbb{K}[x] = \{P(x) = \sum_{i=0}^{n} a_i x^i \mid a_i \in \mathbb{K}, n \in \mathbb{N}\}.$$
Definiamo le operazioni di somma tra polinomi e prodotto per uno scalare nel modo usuale. Sia $n > m$, allora
$$\begin{array}{rccc} + : & \mathbb{K}[x] \times \mathbb{K}[x] & \longrightarrow & \mathbb{K}[x] \\ & \left(\sum_{i=0}^{m} a_i x^i \,,\, \sum_{i=0}^{n} b_i x^i\right) & \longmapsto & \sum_{i=0}^{n} (a_i + b_i) x^i \end{array},$$
dove nella somma assumiamo $a_{m+1} = \cdots = a_n = 0$. Inoltre
$$\begin{array}{rccc} \cdot : & \mathbb{K} \times V & \longrightarrow & V \\ & \left(t \,,\, \sum_{i=0}^{n} a_i x^i\right) & \longmapsto & \sum_{i=0}^{n} (t\, a_i)\, x^i \end{array}.$$
È semplice verificare tutte le proprietà richiesta nella Definizione 4.1, di conseguenza $(\mathbb{K}[x], \mathbb{K}, +, \cdot)$ è uno spazio vettoriale.
- Dati un insieme A ed un campo $\mathbb{K}$, sia $\mathbb{K}^A$ l'insieme delle funzioni da A in $\mathbb{K}$. In $\mathbb{K}^A$ definiamo l'operazione di somma e di prodotto elemento per elemento:
$$\begin{array}{rccc} + : & \mathbb{K}^A \times \mathbb{K}^A & \longrightarrow & \mathbb{K}^A \\ & (f_1, f_2) & \longmapsto & f_1 + f_2 \end{array},$$
dove $(f_1 + f_2)(a) = f_1(a) + f_2(a)$ per ogni $a \in A$;
$$\begin{array}{rccc} \cdot : & \mathbb{K} \times \mathbb{K}^A & \longrightarrow & \mathbb{K}^A \\ & (t\,,\, f) & \longmapsto & t \cdot f \end{array},$$
dove $(t \cdot f)(a) = t\, f(a)$ per ogni $a \in A$. Allora $(\mathbb{K}^A, \mathbb{K}, +, \cdot)$ è uno spazio vettoriale. Lo spazio delle matrici è un caso particolare di spazio di funzioni. Infatti $(\mathrm{Mat}(m, n; \mathbb{K}), \mathbb{K}, +, \cdot) = (\mathbb{K}^{M \times N}, \mathbb{K}, +, \cdot)$ dove M, N sono gli insiemi degli indici di riga e colonna della matrice. Lasciamo al lettore verificare che questo esempio è in realtà un caso particolare dello spazio vettoriale V^A, costituito da tutte le funzioni da un insieme A ad uno spazio vettoriale V.
- Naturalmente, esistono infiniti esempi di strutture algebriche dotate di due operazioni che non siano spazi vettoriali. Ad esempio, $(\mathrm{Mat}(n, n; \mathbb{K}), \mathbb{K}, *, \cdot)$ non è uno spazio vettoriale, anche se l'insieme è chiuso rispetto ad entrambe le operazioni. Infatti, $(\mathrm{Mat}(n, n; \mathbb{K}), *, \cdot)$ non è un gruppo abeliano poiché in generale non esiste l'inversa di una matrice ed inoltre l'operazione non è commutativa. Infine, non valgono le proprietà distributive richieste nella Definizione 4.1.

Gli spazi vettoriali godono di alcune *proprietà elementari*, la cui dimostrazione richiede però l'attento utilizzo degli assiomi contenuti nella Definizione 4.1. Questo è un esercizio molto istruttivo sul modo di ragionare assiomatico e totalmente astratto tipico della matematica, ed in particolare dell'Algebra.

Proposizione 4.3. *Sia* $(V, \mathbb{K}, +, \cdot)$ *uno spazio vettoriale e siano* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ *e* $t \in \mathbb{K}$. *Allora:*

i) vale la proprietà di cancellazione della somma, ovvero $\mathbf{u} + \mathbf{v} = \mathbf{u} + \mathbf{w}$ *se e solo se* $\mathbf{v} = \mathbf{w}$;

ii) vale la proprietà di annullamento del prodotto, ovvero $t \cdot \mathbf{v} = \mathbf{0}$ *se e solo se* $t = 0$ *oppure* $\mathbf{v} = \mathbf{0}$;

iii) l'elemento neutro $\mathbf{0}$ *e l'inverso additivo di* $\mathbf{v}$ *sono unici. In particolare, l'inverso è uguale a* $-1 \cdot \mathbf{v}$.

Dimostrazione. Dobbiamo utilizzare in sequenza le proprietà caratterizzanti gli spazi vettoriali. A scopo didattico, eseguiamo esplicitamente ogni singolo passaggio della dimostrazione.

i) Supponiamo che $\mathbf{u}+\mathbf{v}=\mathbf{u}+\mathbf{w}$. Dato che $(V,+)$ è un gruppo abeliano, allora esiste l'inverso additivo $-\mathbf{u}$ di $\mathbf{u}$. Poiché la somma è un'operazione ben definita, abbiamo che

$$-\mathbf{u}+(\mathbf{u}+\mathbf{v})=-\mathbf{u}+(\mathbf{u}+\mathbf{w}).$$

Per la proprietà associativa della somma, vale

$$(-\mathbf{u}+\mathbf{u})+\mathbf{v}=(-\mathbf{u}+\mathbf{u})+\mathbf{w}.$$

Dalla definizione di inverso ed elemento neutro, segue che

$$\mathbf{0}+\mathbf{v}=\mathbf{0}+\mathbf{w} \qquad \text{e quindi} \qquad \mathbf{v}=\mathbf{w}.$$

Viceversa, se $\mathbf{v}=\mathbf{w}$ allora l'enunciato è ovvio.

ii) Sia $t=0$, per la proprietà distributiva abbiamo

$$0\cdot\mathbf{v}=(0+0)\cdot\mathbf{v}=(0\cdot\mathbf{v})+(0\cdot\mathbf{v}).$$

D'altra parte, per le proprietà dell'elemento neutro della somma vettoriale, abbiamo $0\cdot\mathbf{v}=(0\cdot\mathbf{v})+\mathbf{0}$ e quindi

$$(0\cdot\mathbf{v})+\mathbf{0}=(0\cdot\mathbf{v})+(0\cdot\mathbf{v}).$$

A seguito della proprietà di cancellazione della somma, allora vale $0\cdot\mathbf{v}=\mathbf{0}$. Sia invece $\mathbf{v}=\mathbf{0}$. Analogamente a prima, abbiamo

$$t\cdot\mathbf{0}=t\cdot(\mathbf{0}+\mathbf{0})=(t\cdot\mathbf{0})+(t\cdot\mathbf{0}).$$

Di conseguenza,

$$(t\cdot\mathbf{0})+\mathbf{0}=(t\cdot\mathbf{0})+(t\cdot\mathbf{0})$$

e quindi $t\cdot\mathbf{0}=\mathbf{0}$.

Viceversa, sia $t\cdot\mathbf{v}=\mathbf{0}$. Se $t=0$, l'enunciato è dimostrato. Se $t\neq 0$, allora esiste t^{-1} e quindi

$$t^{-1}\cdot(t\cdot\mathbf{v})=(t^{-1}t)\cdot\mathbf{v}=1\cdot\mathbf{v}=\mathbf{v}=t^{-1}\cdot\mathbf{0}=\mathbf{0}.$$

Nell'equazione abbiamo utilizzato le proprietà di omogeneità e normalizzazione del prodotto per uno scalare e, nell'ultima uguaglianza, l'altro verso della proprietà di annullamento del prodotto che abbiamo precedentemente dimostrato.

iii) Sia $\mathbf{e}$ un elemento neutro rispetto alla somma. Allora $\mathbf{v}+\mathbf{e}=\mathbf{v}=\mathbf{v}+\mathbf{0}$. Di conseguenza, per la proprietà di cancellazione della somma si ha $\mathbf{e}=\mathbf{0}$. In modo del tutto analogo si dimostra che se $\mathbf{v},\mathbf{w}$ sono inversi additivi di $\mathbf{u}$, allora $\mathbf{v}=\mathbf{w}$.

Per le proprietà di annullamento del prodotto, distributiva e di normalizzazione, abbiamo invece

$$\mathbf{0}=0\cdot\mathbf{v}=(1+(-1))\cdot\mathbf{v}=(1\cdot\mathbf{v})+(-1\cdot\mathbf{v})=\mathbf{v}+(-1\cdot\mathbf{v})$$

da cui la tesi. □

Come abbiamo visto nel Capitolo 2, ad ogni struttura algebrica sono associati i relativi omomorfismi, ovvero le funzioni che rispettano la struttura algebrica stessa. Concludiamo questa sezione declinando il concetto di omomorfismo nel caso degli spazi vettoriali.

Definizione 4.4. *Siano $(V,\mathbb{K},+_V,\cdot_V)$ e $(W,\mathbb{K},+_W,\cdot_W)$ due spazi vettoriali sul medesimo campo. Allora la funzione $f:V\to W$ è un omomorfismo di spazi vettoriali, o applicazione lineare, se per ogni $\mathbf{v},\tilde{\mathbf{v}}\in V$ e $t\in\mathbb{K}$, vale*

i) $f(\mathbf{v}+_V\tilde{\mathbf{v}})=f(\mathbf{v})+_W f(\tilde{\mathbf{v}})$;

ii) $f(t \cdot_V \mathbf{v}) = t \cdot_W f(\mathbf{v})$.

L'insieme delle applicazioni lineari da V in W viene indicato con $\mathrm{Hom}(V, W)$.

Normalmente è prassi non utilizzare i pedici per distinguere le operazioni in V e W, ma il lettore deve sempre ricordarsi che queste sono operazioni differenti.

Lo studio dettagliato delle applicazioni lineari è l'argomento del Capitolo 5. In questa sezione ci limiteremo a fornire alcune informazioni indispensabili per lo studio degli spazi vettoriali. Per prima cosa osserviamo che le due richieste della Definizione 4.4 possono essere riassunte in una sola.

Lemma 4.5. *Una funzione $f : V \to W$ tra due spazi vettoriali è un'applicazione lineare se e solo se $f(t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2}) = t_1 \cdot f(\mathbf{v_1}) + t_2 \cdot f(\mathbf{v_2})$ per ogni $\mathbf{v_1}, \mathbf{v_2} \in V$ e $t_1, t_2 \in \mathbb{K}$.*

Dimostrazione. Se f è un'applicazione lineare, allora

$$f(t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2}) = f(t_1 \cdot \mathbf{v_1}) + f(t_2 \cdot \mathbf{v_2}) = t_1 \cdot f(\mathbf{v_1}) + t_2 \cdot f(\mathbf{v_2}).$$

Al contrario, supponiamo vera la seconda parte dell'enunciato. Allora

$$f(\mathbf{v_1} + \mathbf{v_2}) = f(1 \cdot \mathbf{v_1} + 1 \cdot \mathbf{v_2}) = 1 \cdot f(\mathbf{v_1}) + 1 \cdot f(\mathbf{v_2}) = f(\mathbf{v_1}) + f(\mathbf{v_2}),$$

$$f(t \cdot \mathbf{v_1}) = f(t \cdot \mathbf{v_1} + 0 \cdot \mathbf{v_2}) = t \cdot f(\mathbf{v_1}) + 0 \cdot f(\mathbf{v_2}) = t \cdot f(\mathbf{v_1}).$$ □

Due esempi banali di applicazioni lineari sono l'applicazione nulla

$$\begin{array}{rccc} 0: & V & \longrightarrow & V \\ & \mathbf{v} & \longmapsto & \mathbf{0} \end{array}$$

e l'applicazione identità

$$\begin{array}{rccc} \mathrm{Id}_V: & V & \longrightarrow & V \\ & \mathbf{v} & \longmapsto & \mathbf{v} \end{array}.$$

Sappiamo che un isomorfismo di strutture algebriche è un omomorfismo invertibile, la cui funzione inversa è ancora un omomorfismo. L'identità è un esempio ovvio di isomorfismo di spazi vettoriali. In generale, per le applicazioni lineari vale la seguente proprietà.

Proposizione 4.6. *Data $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ invertibile, allora f^{-1} è un'applicazione lineare e quindi f è un isomorfismo.*

Dimostrazione. Siano $\mathbf{w_1}, \mathbf{w_2} \in W$, essendo f invertibile allora esistono $\mathbf{v_1}, \mathbf{v_2} \in V$ tali che $f(\mathbf{v_1}) = \mathbf{w_1}$ e $f(\mathbf{v_2}) = \mathbf{w_2}$. Di conseguenza

$$\begin{aligned} f^{-1}(t_1 \cdot \mathbf{w_1} + t_2 \cdot \mathbf{w_2}) &= f^{-1}(t_1 \cdot f(\mathbf{v_1}) + t_2 \cdot f(\mathbf{v_2})) = f^{-1}(f(t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2})) = \\ &= t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2} = t_1 \cdot f^{-1}(\mathbf{w_1}) + t_2 \cdot f^{-1}(\mathbf{w_2}). \end{aligned}$$

Nella seconda uguaglianza abbiamo utilizzato la linearità di f. □

Esempio 4.7. Consideriamo le funzioni

$$\begin{array}{rccc} f: & \mathbb{K}^n & \longrightarrow & \mathrm{Mat}(1, n; \mathbb{K}) \\ & (a_1, \ldots, a_n) & \longmapsto & [a_1 \ \ldots \ a_n] \end{array}.$$

e

$$\begin{array}{rccc} g: & \mathbb{K}^m & \longrightarrow & \mathrm{Mat}(m, 1; \mathbb{K}) \\ & (a_1, \ldots, a_m) & \longmapsto & \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \end{array}.$$

introdotte nella Sezione 3.1. E' evidente che f e g sono applicazioni lineari biunivoche e quindi isomorfismi. Esse permettono l'identificazione degli spazi vettoriali composti da righe e colonne di matrici con quelli composti da n-uple di numeri.

Osservazione 4.8. Lasciamo al lettore la dimostrazione del risultato analogo a quello della Proposizione 4.6 nel caso di isomorfismi di gruppi e campi. La verifica è basata su un ragionamento del tutto simile a quello seguito nella dimostrazione precedente.

4.2. Sottospazi

In algebra astratta, data una struttura algebrica su un insieme di supporto A, si parla di sottostruttura quando si ha un sottoinsieme $B \subseteq A$ su cui è possibile definire la stessa tipologia di struttura definita su A. Sviluppiamo questo concetto più concretamente nel caso degli spazi vettoriali.

4.2.1. Definizione ed esempi.

Definizione 4.9. *Sia $(V, \mathbb{K}, +, \cdot)$ uno spazio vettoriale ed U un sottoinsieme di V. Allora $(U, \mathbb{K}, +, \cdot)$ si dice sottospazio di $(V, \mathbb{K}, +, \cdot)$ se a sua volta è uno spazio vettoriale. $(U, \mathbb{K}, +, \cdot)$ si dice sottospazio proprio se $U \neq V$.*

Osserviamo che le operazioni in U devono essere le stesse operazioni definite in V, o per essere più precisi le loro restrizioni all'insieme U.

Esempio 4.10. Elenchiamo alcuni dei più rilevanti sottospazi vettoriali con cui lavoreremo.

- Ogni spazio vettoriale $(V, \mathbb{K}, +, \cdot)$ ha due sottospazi banali, lo spazio vettoriale stesso $(V, \mathbb{K}, +, \cdot)$ ed il sottospazio $(\{\mathbf{0}\}, \mathbb{K}, +, \cdot)$.
- Consideriamo lo spazio vettoriale delle matrici $(\mathrm{Mat}(m, n; \mathbb{K}), \mathbb{K}, +, \cdot)$. A seguito delle Proposizioni 3.31 e 3.34, ciascun sottoinsieme $\mathbb{T}_a(n, \mathbb{K})$, $\mathbb{T}_{sa}(n, \mathbb{K})$, $\mathbb{T}_b(n, \mathbb{K})$, $\mathbb{T}_{sb}(n, \mathbb{K})$, $\mathbb{D}(n, \mathbb{K})$, $\mathbb{S}(n, \mathbb{K})$ e $\mathbb{A}(n, \mathbb{K})$ definisce un sottospazio vettoriale di $(\mathrm{Mat}(m, n; \mathbb{K}), \mathbb{K}, +, \cdot)$. L'insieme $GL(n, \mathbb{K})$ delle matrici invertibili non definisce un sottospazio vettoriale, perché non contiene l'elemento neutro 0_n della somma di matrici.
- Lo spazio dei polinomi pensati come funzioni da $\mathbb{K}$ in $\mathbb{K}$ è un sottospazio di $(\mathbb{K}^{\mathbb{K}}, \mathbb{K}, +, \cdot)$. Se $\mathbb{K} = \mathbb{R}$, l'insieme delle funzioni continue, oppure l'insieme delle funzioni differenziabili, da $\mathbb{R}$ in $\mathbb{R}$ definiscono sottospazi di $(\mathbb{R}^{\mathbb{R}}, \mathbb{R}, +, \cdot)$.
- L'insieme dei polinomi di grado minore od uguale ad n definisce il sottospazio $(\mathbb{K}[x]_n, \mathbb{K}, +, \cdot)$ di $(\mathbb{K}[x], \mathbb{K}, +, \cdot)$. È infatti facile verificare la chiusura di $\mathbb{K}[x]_n$ rispetto alle usuali operazioni di somma tra polinomi e prodotto di un polinomio per uno scalare.

Osserviamo che, a meno che vi siano ambiguità sul campo o sulle operazioni da utilizzare, è prassi compiere un abuso di notazione indicando uno spazio vettoriale $(V, \mathbb{K}, +, \cdot)$ semplicemente con il suo insieme di supporto V.

Proposizione 4.11. *Sia $(V, \mathbb{K}, +, \cdot)$ uno spazio vettoriale. Allora, $U \subseteq V$ è un sottospazio se e solo se:*

i) U è chiuso rispetto alla somma vettoriale $+$;
ii) U è chiuso rispetto al prodotto per uno scalare $\cdot$;
iii) U contiene l'elemento neutro $\mathbf{0}$.

Dimostrazione. Se U è un sottospazio di V, allora per definizione devono valere le tre proprietà elencate. Verifichiamo il viceversa. Se U è chiuso rispetto a $+, \cdot$ allora su U sono definite due operazioni che godono automaticamente di quasi tutte le proprietà della Definizione 4.1, poichè queste valgono per ogni elemento di V (si dice che le proprietà vengono ereditate da V). Dobbiamo solo verificare esplicitamente che U contenga l'elemento neutro $\mathbf{0}$ e l'inverso $-\mathbf{u}$, per ogni vettore $\mathbf{u} \in U$. Ma questo segue dalle proprietà elementari degli spazi vettoriali e dalla chiusura di U, infatti $\mathbf{0} = 0 \cdot \mathbf{u} \in U$ e $-\mathbf{u} = -1 \cdot \mathbf{u} \in U$. Osserviamo che il ragionamento appena compiuto garantisce l'appartenenza di $\mathbf{0}$ ad U solo nell'ipotesi che esista almeno un elemento $\mathbf{u} \in U$. La terza richiesta dell'enunciato è pertanto necessaria proprio per garantire che $U \neq \emptyset$. □

Proposizione 4.12. *Sia* $A \in \mathrm{Mat}(m, n; \mathbb{K})$, *allora* $\ker(A)$ *è un sottospazio di* $\mathrm{Mat}(n, 1; \mathbb{K})$.

Dimostrazione. Verifichiamo la chiusura di $\ker(A)$ rispetto alle due operazioni. Siano $B, C \in \ker(A)$ e $t \in \mathbb{K}$. Allora

$$A * (B + C) = (A * B) + (A * C) = 0_{m1} + 0_{m1} = 0_{m1} \quad \text{e} \quad A * (t \cdot B) = t \cdot (A * B) = t \cdot 0_{m1} = 0_{m1},$$

quindi $B + C, t \cdot B \in \ker(A)$. L'appartenenza dell'elemento neutro 0_{n1} a $\ker(A)$ è garantita dal Corollario 3.25. □

Esempio 4.13. Sia

$$A = \begin{bmatrix} 1 & 2 \\ 4 & 8 \end{bmatrix}.$$

Riducendo a scala la matrice, abbiamo

$$\ker(A) = \ker\left(\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}\right).$$

Esistono due modi di rappresentare il nucleo della matrice. Una rappresentazione algebrica o cartesiana di $\ker(A)$ è semplicemente il sistema lineare associato ad una riduzione a scala di A :

$$\ker(A) : \quad \left\{ \; x + 2y = 0 \right. .$$

Una rappresentazione parametrica di $\ker(A)$ è la scrittura esplicita delle soluzioni del sistema dipendenti da $n - r(A)$ parametri:

$$\ker(A) : \quad X(t) = t \cdot \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \quad t \in \mathbb{K}.$$

Indagheremo in modo più sistematico le rappresentazioni algebriche e parametriche di un generico sottospazio nella Proposizione 4.68, al termine del capitolo.

Osserviamo invece che l'insieme delle soluzioni di un sistema non omogeneo $[A|B]$ non è un sottospazio di $\mathrm{Mat}(n, 1; \mathbb{K})$. In primo luogo, esso non contiene l'elemento neutro, poiché $A * 0_{n1} = 0_{m1} \neq B$. Inoltre, l'insieme non è chiuso rispetto alla somma ed al prodotto. Infatti, siano C, D soluzioni di $[A|B]$ e sia $t \in \mathbb{K}$, allora

$$A * (C + D) = (A * C) + (A * D) = B + B = 2 \cdot B \neq B$$

e

$$A * (t \cdot C) = t \cdot (A * C) = t \cdot B \neq B \quad \text{se} \quad t \neq 1.$$

Come per le applicazioni lineari, anche per i sottospazi le richieste di chiusura rispetto alle due operazioni possono essere sintetizzate in una sola.

Lemma 4.14. *Un sottoinsieme U non vuoto di uno spazio vettoriale V è un sottospazio se e solo se per ogni $\mathbf{u_1}, \mathbf{u_2} \in U$ e $t_1, t_2 \in \mathbb{K}$ vale $t_1 \cdot \mathbf{u_1} + t_2 \cdot \mathbf{u_2} \in U$.*

Dimostrazione. Se U è un sottospazio chiaramente deve valere $t_1 \cdot \mathbf{u_1} + t_2 \cdot \mathbf{u_2} \in U$. Supponiamo valga il viceversa. Allora basta porre rispettivamente $t_1 = t_2 = 1$ e $t_2 = 0$ per ottenere le richieste di chiusura rispetto alla somma ed alla moltiplicazione della Proposizione 4.11. Di conseguenza U è un sottospazio. □

4.2.2. Operazioni sui sottospazi. Nella teoria degli insiemi, sappiamo che tra i sottoinsiemi di un insieme dato (o più precisamente nell'insieme delle parti di tale insieme) sono definite le operazioni di intersezione ed unione. Andiamo a vedere come si comporta la struttura di sottospazio rispetto a queste due operazioni.

Definizione 4.15. *Dati lo spazio vettoriale V e due sottospazi U_1, U_2, l'insieme somma dei sottospazi è*

$$U_1 + U_2 = \{\mathbf{v} \in V \mid \mathbf{v} = \mathbf{u_1} + \mathbf{u_2},\ \mathbf{u_1} \in U_1, \mathbf{u_2} \in U_2\}.$$

Proposizione 4.16. *Sia V uno spazio vettoriale ed U_1, U_2 due sottospazi. Allora:*

i) $U_1 \cap U_2$ è un sottospazio di V;
ii) $U_1 \cup U_2$ è un sottospazio di V se e solo se $U_1 \subseteq U_2$ oppure $U_2 \subseteq U_1$;
iii) $U_1 + U_2$ è il più piccolo sottospazio di V contenente $U_1 \cup U_2$.

Dimostrazione. Verifichiamo separatamente ogni singola affermazione.

i) Dobbiamo dimostrare la chiusura di $U_1 \cap U_2$ rispetto alla somma ed al prodotto. Per definizione di intersezione, se $\mathbf{u}, \tilde{\mathbf{u}} \in U_1 \cap U_2$ allora i due vettori appartengono a ciascuno dei singoli sottospazi U_1, U_2. A causa della chiusura dei sottospazi rispetto alla somma, segue che $\mathbf{u} + \tilde{\mathbf{u}} \in U_1, U_2$ e di conseguenza $\mathbf{u} + \tilde{\mathbf{u}} \in U_1 \cap U_2$. Analogamente si verifica la chiusura rispetto al prodotto. Infine, poiché U_1, U_2 sono sottospazi, si ha $\mathbf{0} \in U_1, U_2$ e quindi $\mathbf{0} \in U_1 \cap U_2$.

ii) Se $U_1 \subseteq U_2$, allora $U_1 \cup U_2 = U_2$ e quindi l'unione è un sottospazio per ipotesi (analogamente se $U_2 \subseteq U_1$). Dimostriamo l'implicazione inversa per contrapposizione. Supponiamo che $U_1 \not\subseteq U_2$ e $U_2 \not\subseteq U_1$. Di conseguenza, esistono $\mathbf{u_1} \in U_1$ ed $\mathbf{u_2} \in U_2$ tali che $\mathbf{u_1} \notin U_2$ ed $\mathbf{u_2} \notin U_1$. Allora $\mathbf{u_1} + \mathbf{u_2}$ non può appartenere ad U_1. Infatti, se per assurdo lo fosse avremmo $(\mathbf{u_1} + \mathbf{u_2}) - \mathbf{u_1} = \mathbf{u_2} \in U_1$, dove abbiamo sfruttato la chiusura di U_1 rispetto alla somma. Ma questo è in contraddizione con l'ipotesi. Analogamente, $\mathbf{u_1} + \mathbf{u_2}$ non può appartenere ad U_2. Pertanto, abbiamo due vettori $\mathbf{u_1}, \mathbf{u_2} \in U_1 \cup U_2$ tali che $\mathbf{u_1} + \mathbf{u_2} \notin U_1 \cup U_2$. Quindi $U_1 \cup U_2$ non è un sottospazio, da cui segue la tesi.

iii) Per prima cosa, verifichiamo che $U_1 + U_2$ è un sottospazio di V. Siano $\mathbf{v}, \tilde{\mathbf{v}} \in U_1 + U_2$ e $t \in \mathbb{K}$, allora $\mathbf{v} = \mathbf{u_1} + \mathbf{u_2}$ e $\tilde{\mathbf{v}} = \tilde{\mathbf{u}}_1 + \tilde{\mathbf{u}}_2$ dove $\mathbf{u_1}, \tilde{\mathbf{u}}_1 \in U_1$ e $\mathbf{u_2}, \tilde{\mathbf{u}}_2 \in U_2$. Di conseguenza

$$\mathbf{v} + \tilde{\mathbf{v}} = (\mathbf{u_1} + \mathbf{u_2}) + (\tilde{\mathbf{u}}_1 + \tilde{\mathbf{u}}_2) = (\mathbf{u_1} + \tilde{\mathbf{u}}_1) + (\mathbf{u_2} + \tilde{\mathbf{u}}_2),$$

$$t \cdot \mathbf{v} = t \cdot (\mathbf{u_1} + \mathbf{u_2}) = (t \cdot \mathbf{u_1}) + (t \cdot \mathbf{u_2}).$$

Poiché U_1, U_2 sono sottospazi, si ha $\mathbf{u_1} + \tilde{\mathbf{u}}_1$, $t \cdot \mathbf{u_1} \in U_1$ e $\mathbf{u_2} + \tilde{\mathbf{u}}_2$, $t \cdot \mathbf{u_2} \in U_2$. Segue che $\mathbf{v} + \tilde{\mathbf{v}}$, $t \cdot \mathbf{v}$ appartengono a $U_1 + U_2$. Infine, $\mathbf{0} = \mathbf{0} + \mathbf{0} \in U_1 \cup U_2$.

È evidente che $U_1, U_2 \subseteq U_1 + U_2$ e quindi $U_1 \cup U_2 \subseteq U_1 + U_2$. Infatti, sia $\mathbf{u_1} \in U_1$, allora $\mathbf{u_1} = \mathbf{u_1} + \mathbf{0} \in U_1 + U_2$ (analogo per $\mathbf{u_2} \in U_2$). Rimane da dimostrare che se W è un qualsiasi sottospazio di V contenente $U_1 \cup U_2$, allora $U_1 + U_2 \subseteq W$. Siano $\mathbf{u_1} \in U_1$ ed $\mathbf{u_2} \in U_2$, allora $\mathbf{u_1}, \mathbf{u_2} \in W$ e per le proprietà di sottospazio $\mathbf{u_1} + \mathbf{u_2} \in W$. Di conseguenza $U_1 + U_2 \subseteq W$. □

Le operazioni di somma ed intersezione godono della proprietà associativa, commutativa e di esistenza dell'elemento neutro:

$$(U_1 + U_2) + U_3 = U_1 + (U_2 + U_3), \qquad U_1 + U_2 = U_2 + U_1, \qquad U + \{\mathbf{0}\} = U,$$
$$(U_1 \cap U_2) \cap U_3 = U_1 \cap (U_2 \cap U_3), \qquad U_1 \cap U_2 = U_2 \cap U_1, \qquad U \cap V = U.$$

Entrambe le operazioni non soddisfano invece la proprietà di esistenza dell'inverso. A questo punto, è semplice generalizzare il concetto di somma ed intersezione per un numero finito arbitrario di sottospazi $U_1, \ldots, U_n \subseteq V$:

$$U_1 + \cdots + U_n = \sum_{i=1}^{n} U_i = \{\mathbf{v} \in V \mid \mathbf{v} = \mathbf{u_1} + \cdots + \mathbf{u_n},\ \mathbf{u_1} \in U_1, \ldots, \mathbf{u_n} \in U_n\},$$
$$U_1 \cap \cdots \cap U_n = \bigcap_{i=1}^{n} U_i = \{\mathbf{v} \in V \mid \mathbf{v} \in U_1, \ldots, \mathbf{v} \in U_n\}.$$

Definizione 4.17. *Siano V uno spazio vettoriale ed $U_1, \ldots, U_n \subseteq V$ sottospazi. Allora V si dice somma diretta dei sottospazi se per ogni $\mathbf{v} \in V$ esiste un'unica n-upla $(\mathbf{u_1}, \ldots, \mathbf{u_n}) \in U_1 \times \cdots \times U_n$ tale che $\mathbf{v} = \mathbf{u_1} + \cdots + \mathbf{u_n}$. In tal caso si usa la notazione $V = U_1 \oplus \cdots \oplus U_n = \bigoplus_{i=1}^{n} U_i$.*

Anche la somma diretta soddisfa la proprietà associativa, commutativa ed esistenza dell'elemento neutro:

$$(U_1 \oplus U_2) \oplus U_3 = U_1 \oplus (U_2 \oplus U_3), \qquad U_1 \oplus U_2 = U_2 \oplus U_1, \qquad U \oplus \{\mathbf{0}\} = U.$$

La seguente proposizione offre una caratterizzazione della somma diretta, che da alcuni autori viene usata come definizione.

Proposizione 4.18. *Siano V uno spazio vettoriale, $U_1, \ldots, U_m \subseteq V$ con $m \geq 2$ sottospazi di V ed infine*

$$U_{\widehat{k}} = \sum_{\substack{i=1 \\ i \neq k}}^{m} U_i.$$

Allora $V = \bigoplus_{i=1}^{m} U_i$ se e solo se:

i) $V = \sum_{i=1}^{m} U_i$;

ii) $U_k \cap U_{\widehat{k}} = \{\mathbf{0}\}$ *per ogni* $1 \leq k \leq m$.

Dimostrazione. L'ipotesi (i) è condizione necessaria e sufficiente a garantire l'esistenza della decomposizione di ogni $\mathbf{v} \in V$ come somma di vettori in $U_1, \ldots, U_m$. Dobbiamo dimostrare che l'ipotesi (ii) è equivalente alla richiesta di unicità di tale decomposizione.

Supponiamo che V sia somma diretta ed $\mathbf{u_k} \in U_k \cap U_{\widehat{k}}$. Segue che per ogni $t \in \mathbb{K}$ la coppia $(t \cdot \mathbf{u_k}, -t \cdot \mathbf{u_k}) \in U_k \times U_{\widehat{k}}$ fornisce una decomposizioni del vettore nullo: $\mathbf{0} = (t \cdot \mathbf{u_k}) + (-t \cdot \mathbf{u_k})$. Per costruzione, $\mathbf{u_k}$ può essere a sua volta decomposto come somma di vettori $\mathbf{u_i} \in U_i$, $1 \leq i \leq m$, $i \neq k$. Quindi, per ogni t esiste la decomposizione

$$\mathbf{0} = t \cdot \mathbf{u_k} - \sum_{\substack{i=1 \\ i \neq k}}^{m} t \cdot \mathbf{u_i}.$$

Tali decomposizioni sono tutte distinte se $\mathbf{u_k} \neq \mathbf{0}$, pertanto l'unicità richiesta dalla definizione di somma diretta impone che $U_k \cap U_{\widehat{k}} = \{\mathbf{0}\}$.

Viceversa, supponiamo che per ogni $k = 1, \dots, m$ valga $U_k \cap U_{\widehat{k}} = \{\mathbf{0}\}$. Siano

$$\mathbf{v} = \sum_{i=1}^{m} \mathbf{u_i} = \sum_{i=1}^{m} \tilde{\mathbf{u}}_{\mathbf{i}}$$

due decomposizioni di $\mathbf{v} \in V$ (se è vera l'ipotesi (i), sicuramente ne esiste almeno una), e chiamiamo

$$\mathbf{u}_{\widehat{\mathbf{k}}} = \sum_{\substack{i=1\\ i\neq k}}^{m} \mathbf{u_i}, \qquad \tilde{\mathbf{u}}_{\widehat{\mathbf{k}}} = \sum_{\substack{i=1\\ i\neq k}}^{m} \tilde{\mathbf{u}}_{\mathbf{i}}.$$

Allora, $\mathbf{v} = \mathbf{u_k} + \mathbf{u}_{\widehat{\mathbf{k}}} = \tilde{\mathbf{u}}_{\mathbf{k}} + \tilde{\mathbf{u}}_{\widehat{\mathbf{k}}}$ e di conseguenza $\mathbf{u_k} - \tilde{\mathbf{u}}_{\mathbf{k}} = \tilde{\mathbf{u}}_{\widehat{\mathbf{k}}} - \mathbf{u}_{\widehat{\mathbf{k}}} \in U_k \cap V_{\widehat{k}}$. Per ipotesi, deve valere $\mathbf{u_k} - \tilde{\mathbf{u}}_{\mathbf{k}} = \mathbf{0}$ e quindi $\mathbf{u_k} = \tilde{\mathbf{u}}_{\mathbf{k}}$. Poiché questo è vero per ogni k, segue l'unicità della decomposizione e quindi la tesi. □

Esempio 4.19. Consideriamo lo spazio vettoriale delle matrici quadrate $\mathrm{Mat}(n, n; \mathbb{K})$. Allora:

- per la Proposizione 3.38 abbiamo $\mathrm{Mat}(n, n; \mathbb{K}) = \mathbb{S}(n, \mathbb{K}) + \mathbb{A}(n, \mathbb{K})$. Inoltre, l'unica matrice sia simmetrica che antisimmetrica è la matrice nulla, cioè $\mathbb{S}(n, \mathbb{K}) \cap \mathbb{A}(n, \mathbb{K}) = \{0_n\}$. In conclusione, $\mathrm{Mat}(n, n; \mathbb{K}) = \mathbb{S}(n, \mathbb{K}) \oplus \mathbb{A}(n, \mathbb{K})$. Questo è coerente con il fatto che per ogni matrice A esiste un'unica coppia di matrici simmetriche ed antisimmetriche (A_S, A_A) tali che $A = A_S + A_A$;
- ogni matrice può essere scritta come somma di una matrice triangolare alta ed una triangolare bassa, quindi $\mathrm{Mat}(n, n; \mathbb{K}) = \mathbb{T}_a(n, \mathbb{K}) + \mathbb{T}_b(n, \mathbb{K})$. Dato che $\mathbb{T}_a(n, \mathbb{K}) \cap \mathbb{T}_b(n, \mathbb{K}) = \mathbb{D}(n, \mathbb{K})$, in questo caso $\mathrm{Mat}(n, n; \mathbb{K}) \neq \mathbb{T}_a(n, \mathbb{K}) \oplus \mathbb{T}_b(n, \mathbb{K})$. Infatti, esistono infiniti modi di scrivere una generica matrice come somma di una matrice triangolare alta ed una bassa.

Esempio 4.20. Siano $V = \mathrm{Mat}(3, 1; \mathbb{Q})$ ed $U_1 = \ker(A_1)$, $U_2 = \ker(A_2)$ dove

$$A_1 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 2 & 4 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Allora,

$$U_1 = \left\{ t_1 \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix},\ t_1 \in \mathbb{Q} \right\}, \qquad U_2 = \left\{ t_2 \cdot \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix},\ t_2 \in \mathbb{Q} \right\}.$$

La somma vettoriale dei due nuclei si ottiene in maniera semplice partendo dalla loro rappresentazione parametrica:

$$U_1 + U_2 = \left\{ t_1 \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + t_2 \cdot \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix},\ t_1, t_2 \in \mathbb{Q} \right\}.$$

Al contrario, l'intersezione dei due nuclei è legata in maniera naturale alla loro rappresentazione algebrica:

$$U_1 \cap U_2 = \ker\left(\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \right) = \ker\left(\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 2 & 4 & 2 \\ 0 & 0 & 1 \end{bmatrix} \right) = \ker\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \right) = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}.$$

Definiamo il sottospazio

$$U_3 = \left\{ t_3 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\ t_3 \in \mathbb{Q} \right\}$$

e verifichiamo che $V = U_1 \oplus U_2 \oplus U_3$. Dobbiamo dimostrare che esiste un unico modo di scrivere ogni vettore di V come somma di tre vettori in U_1, U_2, U_3. Per trovare i tre vettori dobbiamo risolvere un sistema lineare:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = t_1 \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + t_2 \cdot \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} + t_3 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \begin{cases} a = t_1 + 2t_2 + t_3 \\ b = -t_2 \\ c = -t_1 \end{cases} \quad \Rightarrow \quad \begin{cases} t_1 = -a \\ t_2 = -b \\ t_3 = a + 2b + c \end{cases}.$$

Poiché la soluzione del sistema esite ed è unica per ogni $a, b, c \in \mathbb{K}$, segue la tesi.

4.3. Combinazioni lineari, basi di uno spazio vettoriale e mappa delle componenti

Una proprietà fondamentale degli spazi vettoriali è la possibilità di descrivere tutte le loro proprietà utilizzando sottoinsiemi particolari di vettori, che nei casi più fortunati sono di cardinalità finita. Stiamo parlando delle basi di uno spazio vettoriale, che saranno l'argomento principale delle prossime tre sezioni di questo capitolo.

Definizione 4.21. *Sia $(V, \mathbb{K}, +, \cdot)$ uno spazio vettoriale con $U = \{\mathbf{u_1}, \dots, \mathbf{u_m}\} \subseteq V$ e $t_1, \dots, t_n \in \mathbb{K}$. Allora il vettore*

$$\mathbf{u} = \sum_{i=1}^{m} t_i \cdot \mathbf{u_i}$$

è detto combinazione lineare dei vettori in U.

Definizione 4.22. *L'insieme ordinato $\mathcal{B} = \{\mathbf{v_1}, \dots, \mathbf{v_n}\} \subseteq V$ si dice base dello spazio vettoriale V se ogni vettore $\mathbf{v} \in V$ è uguale ad un'unica combinazione lineare dei vettori di $\mathcal{B}$.*

Esempio 4.23. In ogni spazio vettoriale che abbiamo studiato, esistono delle basi particolari con cui è semplice lavorare. Queste vengono dette basi canoniche o standard.

- In $\mathbb{K}^n$ l'insieme $\mathcal{B}_n = \{\mathbf{e_1} = (1, 0, \dots, 0), \mathbf{e_2} = (0, 1, \dots, 0), \dots, \mathbf{e_n} = (0, 0, \dots, 1)\}$ è una base. Infatti, ogni vettore è scrivibile in modo unico come

$$(a_1, \dots, a_n) = \sum_{i=1}^{n} a_i \cdot \mathbf{e_i}.$$

- Nello spazio $\mathrm{Mat}(m, n; \mathbb{K})$ consideriamo le matrici

$$\begin{array}{rccc} E_{ij}: & M \times N & \longrightarrow & \mathbb{K} \\ & (k, l) & \longmapsto & \delta_{ik}\, \delta_{jl} \end{array}.$$

  Ad esempio, se $m = n = 2$, abbiamo:

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

  Allora, $\mathcal{B}_{mn} = \{E_{ij} \mid (i, j) \in M \times N\}$ è una base di $\mathrm{Mat}(m, n; \mathbb{K})$, detta base canonica delle matrici di tipo (m, n). Infatti, ogni matrice può essere scritta in modo unico come combinazione lineare degli elementi della base canonica:

$$A = [a_{ij}] = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \cdot E_{ij}.$$

  Ad esempio,

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a_{12} \cdot \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + a_{21} \cdot \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + a_{22} \cdot \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Prendendo degli opportuni sottoinsiemi di $\mathcal{B}_{nn}$, si ottengono in modo naturale le basi canoniche di $\mathbb{T}_a(n;\mathbb{K})$, $\mathbb{T}_{sa}(n;\mathbb{K})$, $\mathbb{T}_b(n;\mathbb{K})$, $\mathbb{T}_{sb}(n;\mathbb{K})$ e $\mathbb{D}(n;\mathbb{K})$. Invece, le basi canoniche di $\mathbb{S}(n;\mathbb{K})$ e $\mathbb{A}(n;\mathbb{K})$ sono rispettivamente $\mathcal{B}_{\mathbb{S}(n;\mathbb{K})} = \{S_{ij} = \frac{1}{2}\cdot(E_{ij}+E_{ji}) \,|\, 1 \leq i \leq j \leq n\}$ e $\mathcal{B}_{\mathbb{A}(n;\mathbb{K})} = \{A_{ij} = E_{ij} - E_{ji} \,|\, 1 \leq i < j \leq n\}$. Ad esempio, se $m = n = 2$, abbiamo:

$$S_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad S_{12} = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}, \quad S_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

- L'insieme dei monomi $\mathcal{B} = \{x^i \,|\, i \geq 0\}$ è la base canonica di $\mathbb{K}[x]$. La base canonica del sottospazio $\mathbb{K}[x]_n$ dei polinomi di grado minore od uguale ad n è $\mathcal{B}_n = \{x^i \,|\, n \geq i \geq 0\}$.
- Non esiste una base dello spazio vettoriale banale $V = \{\mathbf{0}\}$. Infatti, l'unico sottoinsieme non vuoto di V è l'insieme V stesso, quindi l'unica possibile base dello spazio sarebbe $\mathcal{B} = \{\mathbf{0}\}$. Ma tale insieme non è una base, perché il vettore $\mathbf{0} \in V$ può essere scritto in infiniti modi distinti rispetto a $\mathcal{B}$:

$$\mathbf{0} = t \cdot \mathbf{0} \qquad \text{per ogni } t \in \mathbb{K}.$$

Un primo risultato importante legato all'esistenza delle basi è la possibilità di associare in modo naturale ad ogni vettore una n-upla di numeri in $\mathbb{K}$.

Definizione 4.24. *Sia V uno spazio vettoriale e $\mathcal{B} = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ una sua base. Per ogni vettore $\mathbf{v} \in V$, la combinazione lineare $\mathbf{v} = t_1 \cdot \mathbf{v_1} + \cdots + t_n \cdot \mathbf{v_n}$ è detta decomposizione di $\mathbf{v}$ rispetto alla base $\mathcal{B}$. I coefficienti $t_1, \dots, t_n$ della decomposizione si dicono componenti di $\mathbf{v}$ rispetto a $\mathcal{B}$. Infine, la funzione*

$$\begin{array}{rccc} \phi_{\mathcal{B}}: & V & \longrightarrow & \mathrm{Mat}(n,1;\mathbb{K}) \\ & \mathbf{v} & \longmapsto & \mathbf{v}|_{\mathcal{B}} = \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix} \end{array}.$$

è detta mappa delle componenti.

Esempio 4.25. Consideriamo le basi canoniche dell'esempio precedente.

- In $\mathbb{K}^n$ consideriamo la base canonica $\mathcal{B}_n = \{\mathbf{e_1}, \dots, \mathbf{e_n}\}$. La decomposizione di $\mathbf{v} = (a_1, \dots, a_n)$ rispetto a tale base è

$$\mathbf{v} = \sum_{i=1}^{n} a_n \cdot \mathbf{e_i}.$$

  Quindi le componenti di $\mathbf{v}$ sono

$$\phi_{\mathcal{B}_n}(\mathbf{v}) = \mathbf{v}|_{\mathcal{B}_n} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

  Osserviamo che la mappa $\phi_{\mathcal{B}_n}$ coincide con la funzione g definita nell'Osservazione 3.3.
- Sia

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

  La matrice può essere pensata sia come un elemento di $\mathrm{Mat}(2,2;\mathbb{K})$, sia di $\mathbb{S}(2,\mathbb{K})$. Le decomposizioni di A relativamente alle rispettive basi canoniche sono

$$A = a \cdot E_{11} + b \cdot E_{12} + b \cdot E_{21} + c \cdot E_{22} \qquad \text{e} \qquad A = a \cdot S_{11} + 2b \cdot S_{12} + c \cdot S_{22}.$$

Di conseguenza, le componenti di A sono:

$$\phi_{\mathcal{B}_{22}}(A) = A|_{\mathcal{B}_{22}} = \begin{bmatrix} a \\ b \\ b \\ c \end{bmatrix} \qquad \text{e} \qquad \phi_{\mathcal{B}_{\mathbb{S}(2;\mathbb{K})}}(A) = A|_{\mathbb{S}(2;\mathbb{K})} = \begin{bmatrix} a \\ 2b \\ c \end{bmatrix}.$$

- Sia invece $P(x) = a + bx + cx^2$. Anche in questo caso, il polinomio può essere pensato come appartenente a spazi vettoriali diversi, ad esempio $\mathbb{K}[x]_2$ e $\mathbb{K}[x]_3$. Le decomposizioni di $P(x)$ relativamente alle basi canoniche sono:

$$P(x) = a \cdot x^0 + b \cdot x^1 + c \cdot x^2 \qquad \text{e} \qquad P(x) = a \cdot x^0 + b \cdot x^1 + c \cdot x^2 + 0 \cdot x^3.$$

Di conseguenza, le componenti di $P(x)$ sono:

$$\phi_{\mathcal{B}_2}(P(x)) = P(x)|_{\mathcal{B}_2} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \qquad \text{e} \qquad \phi_{\mathcal{B}_3}(P(x)) = P(x)|_{\mathcal{B}_3} = \begin{bmatrix} a \\ b \\ c \\ 0 \end{bmatrix}.$$

Teorema 4.26. *Dato uno spazio vettoriale V ed una sua base $\mathcal{B} = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$, la mappa delle componenti $\phi_{\mathcal{B}}$ è un isomorfismo di spazi vettoriali tra V e $\mathrm{Mat}(n, 1; \mathbb{K})$.*

Dimostrazione. È evidente che la funzione $\phi_{\mathcal{B}}$ è invertibile. Infatti, per ogni $[\,t_1 \dots t_n\,]^T \in \mathrm{Mat}(n, 1; \mathbb{K})$ esiste un unico vettore $\mathbf{v} \in V$ avente tali componenti, ovvero

$$\mathbf{v} = \sum_{i=1}^{n} t_i \cdot \mathbf{v_i}.$$

Grazie alla Proposizione 4.6, resta solo da dimostrare che $\phi_{\mathcal{B}}$ è lineare. Siano $\mathbf{v}, \tilde{\mathbf{v}} \in V$ con decomposizioni

$$\mathbf{v} = \sum_{i=1}^{n} t_i \cdot \mathbf{v_i}, \qquad \tilde{\mathbf{v}} = \sum_{i=1}^{n} \tilde{t}_i \cdot \mathbf{v_i}$$

e siano $\alpha, \beta \in \mathbb{K}$. Allora

$$\phi_{\mathcal{B}}(\alpha \cdot \mathbf{v} + \beta \cdot \tilde{\mathbf{v}}) = \phi_{\mathcal{B}}\left(\alpha \cdot \sum_{i=1}^{n} t_i \cdot \mathbf{v_i} + \beta \cdot \sum_{i=1}^{n} \tilde{t}_i \cdot \mathbf{v_i}\right) = \phi_{\mathcal{B}}\left(\sum_{i=1}^{n} (\alpha\, t_i + \beta \tilde{t}_i) \cdot \mathbf{v_i}\right) = \begin{bmatrix} \alpha\, t_1 + \beta\, \tilde{t}_1 \\ \vdots \\ \alpha\, t_n + \beta\, \tilde{t}_n \end{bmatrix}.$$

Ma l'ultima matrice colonna la possiamo riscrivere come

$$\begin{bmatrix} \alpha\, t_1 + \beta\, \tilde{t}_1 \\ \vdots \\ \alpha\, t_n + \beta\, \tilde{t}_n \end{bmatrix} = \alpha \cdot \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix} + \beta \cdot \begin{bmatrix} \tilde{t}_1 \\ \vdots \\ \tilde{t}_n \end{bmatrix} = \alpha \cdot \phi_{\mathcal{B}}(\mathbf{v}) + \beta \cdot \phi_{\mathcal{B}}(\tilde{\mathbf{v}}),$$

da cui segue la tesi. □

Il teorema precedente non solo fissa una corrispondenza biunivoca tra vettori e componenti, ma ci permette anche di eseguire ogni possibile operazione tra vettori traducendole in operazioni sulle componenti, in quanto il calcolo delle coordinate e le operazioni sui vettori commutano tra loro. Da questo punto di vista si spiega l'affermazione secondo cui lo spazio delle matrici è l'esempio più importante di spazio vettoriale, poiché qualsiasi altro spazio (finitamente generato, vedi Sezione 4.5) è identificabile con quest'ultimo attraverso la mappa delle componenti. Vedremo nel Capitolo 5 che questo isomorfismo ammette un analogo anche per le applicazioni lineari tra spazi vettoriali.

4.4. Vettori generatori, indipendenza lineare e basi

In questa sezione caratterizziamo le basi di uno spazio vettoriale V. Iniziamo dando la definizione di insiemi di vettori generatori ed insiemi linearmente indipendenti.

Definizione 4.27. *Dato $U = \{\mathbf{u_1}, \dots, \mathbf{u_m}\} \subseteq V$, l'insieme di tutte le combinazioni lineari di tali vettori si dice spazio generato o span di U, e viene indicato con*

$$\mathcal{L}(\mathbf{u_1}, \dots, \mathbf{u_m}) = \mathrm{Span}(\mathbf{u_1}, \dots, \mathbf{u_m}) = \left\{ \mathbf{v} \in V \,\middle|\, \mathbf{v} = \sum_{i=1}^{m} t_i \cdot \mathbf{u_i},\ t_1, \dots, t_m \in \mathbb{K} \right\}.$$

I vettori in U si dicono generatori di $\mathcal{L}(\mathbf{u_1}, \dots, \mathbf{u_m})$.

Proposizione 4.28. *Lo spazio generato dai vettori $\mathbf{u_1}, \dots, \mathbf{u_m} \in V$ è un sottospazio vettoriale di V.*

Dimostrazione. Siano $\mathbf{v}, \tilde{\mathbf{v}} \in \mathcal{L}(\mathbf{u_1}, \dots, \mathbf{u_n})$ e $t \in \mathbb{K}$. Allora:

$$\mathbf{v} + \tilde{\mathbf{v}} = \left(\sum_{i=1}^{m} t_i \cdot \mathbf{u_i} \right) + \left(\sum_{i=1}^{n} \tilde{t}_i \cdot \mathbf{u_i} \right) = \sum_{i=1}^{m} (t_i + \tilde{t}_i) \cdot \mathbf{u_i} \in \mathcal{L}(\mathbf{u_1}, \dots, \mathbf{u_n});$$

$$t \cdot \mathbf{v} = t \cdot \left(\sum_{i=1}^{m} t_i \cdot \mathbf{u_i} \right) = \sum_{i=1}^{m} (t\, t_i) \cdot \mathbf{u_i} \in \mathcal{L}(\mathbf{u_1}, \dots, \mathbf{u_n}).$$

Nelle precedenti uguaglianze si sono utilizzate le proprietà di associatività, commutatività e distributività delle operazioni negli spazi vettoriali. Infine,

$$\mathbf{0} = \sum_{i=1}^{m} 0 \cdot \mathbf{u_i} \in \mathcal{L}(\mathbf{u_1}, \dots, \mathbf{u_n}).$$

□

Definizione 4.29. *L'insieme $U = \{\mathbf{u_1}, \dots, \mathbf{u_m}\} \subseteq V$ è linearmente indipendente se l'unica combinazione lineare di risultato $\mathbf{0}$ è quella banale, ovvero*

$$\mathbf{0} = \sum_{i=1}^{m} 0 \cdot \mathbf{u_i}.$$

Se un insieme non è indipendente, allora si dice linearmente dipendente.

Proposizione 4.30. *L'insieme $U = \{\mathbf{u_1}, \dots, \mathbf{u_m}\} \subseteq V$, dove $m \geq 2$, è linearmente dipendente se e solo se esiste almeno un vettore di U combinazione lineare dei rimanenti $m-1$ vettori.*

Dimostrazione. Per definizione, U è linearmente dipendente se esiste una combinazione lineare

$$\sum_{i=1}^{m} t_i \cdot \mathbf{u_i} = \mathbf{0}$$

con almeno un coefficiente $t_1, \dots, t_m$ diverso da 0. Senza perdere in generalità, supponiamo $t_m \neq 0$. Allora vale

$$\mathbf{v_m} = \sum_{i=1}^{n-1} -\frac{t_i}{t_m} \cdot \mathbf{v_i}.$$

Il viceversa si dimostra facilmente in maniera analoga. □

Lemma 4.31. *Un insieme di vettori U contenente il vettore $\mathbf{0}$ è linearmente dipendente.*

Dimostrazione. Poiché $\mathbf{0} = t \cdot \mathbf{0}$ per ogni $t \in \mathbb{K}$, l'insieme U è dipendente. □

Benché il concetto di dipendenza ed indipendenza lineare sia riferito ad un insieme di vettori, è comune abusare della notazione dicendo che gli m vettori di U sono linearmente dipendenti od indipendenti.

A questo punto possiamo dare la prima caratterizzazione di base di uno spazio vettoriale.

Teorema 4.32. *Un insieme ordinato di vettori $\mathcal{B} = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ è una base di V se e solo se:*

i) $\mathcal{B}$ è un insieme di generatori di V;

ii) $\mathcal{B}$ è linearmente indipendente.

Dimostrazione. Se $\mathcal{B}$ è una base, le due affermazione dell'enunciato seguono direttamente dalla Definizione 4.22. Infatti, affinché ogni vettore $v \in V$ sia esprimibile come combinazione lineare dei vettori $\mathcal{B}$, tale insieme deve generare l'intero spazio vettoriale. Inoltre, la richiesta di unicità delle decomposizioni implica in particolare che debba esistere un'unica decomposizione del vettore $\mathbf{0}$, che necessariamente è quella banale e quindi l'insieme $\mathcal{B}$ è indipendente.

Viceversa, se vale la prima proprietà elencata allora ogni vettore $v \in V$ è combinazione lineare dei vettori in $\mathcal{B}$. Dobbiamo dimostrare che se è vera anche la seconda proprietà, allora la decomposizione di $\mathbf{v}$ è unica. Supponiamo esistano due decomposizioni

$$\mathbf{v} = \sum_{i=1}^{n} t_i \cdot \mathbf{v_i} = \sum_{i=1}^{n} \tilde{t}_i \cdot \mathbf{v_i}\,.$$

Allora

$$\mathbf{0} = \mathbf{v} - \mathbf{v} = \left(\sum_{i=1}^{n} t_i \cdot \mathbf{v_i}\right) - \left(\sum_{i=1}^{n} \tilde{t}_i \cdot \mathbf{v_i}\right) = \sum_{i=1}^{n} (t_i - \tilde{t}_i) \cdot \mathbf{v_i}.$$

La richiesta di indipendenza implica che $t_i = \tilde{t}_i$ per ogni $i = 1, \ldots, n$, da cui la tesi. □

Esempio 4.33. Nello spazio $\mathrm{Mat}(2, 1; \mathbb{Q})$, consideriamo l'insieme U i cui elementi sono i vettori

$$\mathbf{u_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \mathbf{u_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \text{e} \qquad \mathbf{u_3} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}.$$

Allora:

- U è un 'inseme di generatori, in quanto ogni vettore in $\mathrm{Mat}(2, 1; \mathbb{Q})$ può essere ottenuto come
$$\begin{bmatrix} a \\ b \end{bmatrix} = a \cdot \mathbf{u_1} + b \cdot \mathbf{u_2} + 0 \cdot \mathbf{u_3}.$$
- U non è linearmente indipendente, poiché $\mathbf{u_1} + \mathbf{u_2} + \mathbf{u_3} = \mathbf{0}$.
- ogni coppia di vettori in U è una base di $\mathrm{Mat}(2, 1; \mathbb{Q})$:
$$\begin{bmatrix} a \\ b \end{bmatrix} = a \cdot \mathbf{u_1} + b \cdot \mathbf{u_2} = (a - b) \cdot \mathbf{u_1} - b \cdot \mathbf{u_3} = (a - b) \cdot \mathbf{u_2} - b \cdot \mathbf{u_3}.$$

Concludiamo la sezione con una seconda caratterizzazione delle basi, formulata utilizzando il linguaggio dei sottospazi vettoriali e delle operazioni coinvolgenti sottospazi.

Teorema 4.34. *Dato $U = \{\mathbf{u_1}, \ldots, \mathbf{u_m}\} \subseteq V$, per ogni $1 \leq k \leq m$ chiamiamo*

$$U_k = \mathcal{L}(\mathbf{u_k}) \qquad e \qquad U_{\widehat{k}} = \sum_{\substack{i=1 \\ i \neq k}}^{m} U_i.$$

Allora:

i) U è un insieme di generatori di V se e solo se $V = \sum_{k=1}^{m} U_k$;

ii) se $\mathbf{0} \notin U$ *ed* $m \geq 2$*, l'insieme* U *è linearmente indipendente se e solo se* $U_k \cap U_{\widehat{k}} = \{\mathbf{0}\}$ *per ogni* $1 \leq k \leq m$*;*

iii) se $\mathbf{0} \notin U$*, l'insieme ordinato* U *è una base di* V *se e solo se* $V = \bigoplus_{k=1}^{m} U_k$*.*

Dimostrazione. Per prima cosa osserviamo che

$$U_i = \mathcal{L}(\mathbf{u_i}) = \{\mathbf{v} \in V \,|\, \mathbf{v} = t_i \cdot \mathbf{u_i},\ t_i \in \mathbb{K}\}.$$

Per definizione di somma di sottospazi, $V = \sum_{k=1}^{m} U_k$ se e solo se ogni $\mathbf{v} \in V$ è esprimibile come combinazione lineare di vettori appartenenti ai singoli U_i, ovvero

$$\mathbf{v} = \sum_{i=1}^{m} t_i \cdot \mathbf{u_i}.$$

Ma questo è vero se e solo se U è un insieme di generatori di V. D'altra parte, per definizione di somma diretta, $V = \bigoplus_{k=1}^{m} U_k$ se e solo se la precedente combinazione lineare è unica, e quindi se e solo se U è una base di V. Infine, grazie alla Proposizione 4.30 sappiamo che U è dipendente se e solo se almeno un vettore $\mathbf{u_k}$ è combinazione lineare degli altri. Ma questo è vero se e solo se $\mathbf{u_k} \in U_{\widehat{k}}$, da cui la tesi. □

4.5. Esistenza delle basi e dimensione

Finora abbiamo discusso delle proprietà delle basi, ma a meno di esempi specifici, non abbiamo dimostrato la loro esistenza nel caso di un generico spazio vettoriale. Colmiamo questa lacuna nel caso degli spazi finitamente generati, su cui porremo le nostre attenzioni da ora in poi.

Definizione 4.35. *Uno spazio vettoriale* V *si dice finitamente generato se esiste un insieme finito di generatori, ovvero* $V = \mathcal{L}(\mathbf{v_1}, \ldots, \mathbf{v_m})$ *con* $m \in \mathbb{N}$*.*

Esempio 4.36. Lo spazio $\mathrm{Mat}(m, n; \mathbb{K})$, con tutti i suoi sottospazi, e lo spazio $\mathbb{K}[x]_n$ con $n \in \mathbb{N}$ sono spazi finitamente generati, in quanto le loro basi canoniche sono costituite da un numero finito di elementi. Lo spazio $\mathbb{K}[x]$ non è uno spazio finitamente generato. In generale, gli spazi di funzioni sono gli esempi più rilevanti di spazi vettoriali non finitamente generati. Il loro studio è uno degli argomenti principali dell'Analisi Funzionale.

Teorema 4.37. *Sia* $V \neq \{\mathbf{0}\}$ *uno spazio vettoriale finitamente generato, con* $U = \{\mathbf{v_1}, \ldots, \mathbf{v_m}\}$ *un insieme di generatori. Allora esiste una base di* V *composta da* $n \leq m$ *vettori di* U*.*

Dimostrazione. La dimostrazione prevede la definizione di un algoritmo, detto di eliminazione, per individuare una base all'interno di U. Se U è linearmente indipendente, allora è una base ed $n = m$. Se U non è indipendente, allora per la Proposizione 4.30 esiste un vettore $\mathbf{u_k} \in U$ che è combinazione lineare degli altri generatori:

$$\mathbf{u_k} = \sum_{i=1}^{k-1} t_i \cdot \mathbf{u_i} + \sum_{i=k+1}^{m} t_i \cdot \mathbf{u_i}.$$

Definiamo $U_{\widehat{k}} = U \setminus \{\mathbf{u_k}\}$, che è ancora un insieme di generatori di V. Infatti, per ogni $\mathbf{v} \in V$ abbiamo:

$$\mathbf{v} = \sum_{i=1}^{m} \tilde{t}_i \cdot \mathbf{u_i} = \sum_{i=1}^{k-1} (\tilde{t}_i + \tilde{t}_k t_i) \cdot \mathbf{u_i} + \sum_{i=k+1}^{m} (\tilde{t}_i + \tilde{t}_k t_i) \cdot \mathbf{u_i}.$$

Ripetiamo la procedura partendo dall'insieme $U_{\widehat{k}}$. Poiché lo spazio è finitamente generato ed è diverso da $\{\mathbf{0}\}$, l'algoritmo termina in un numero finito di iterazioni, restituendo un'insieme di generatori linearmente indipendente e quindi una base di V. $\square$

Esempio 4.38. Riprendiamo i vettori dell'Esempio 4.33. Nello spazio $\mathrm{Mat}(2,1;\mathbb{Q})$, l'insieme U è formato da

$$\mathbf{u_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \mathbf{u_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \text{e} \qquad \mathbf{u_3} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}.$$

U non è linearmente indipendente, infatti $\mathbf{u_3} = -\mathbf{u_1} - \mathbf{u_2}$. Definiamo quindi $U_{\widehat{3}} = \{\mathbf{u_1}, \mathbf{u_2}\}$, che è una base di $\mathrm{Mat}(2,1;\mathbb{Q})$.

Grazie al teorema precedente possiamo definire il concetto fondamentale di dimensione di uno spazio vettoriale. A tal fine è pero necessario dimostrare preliminarmente un paio di risultati. Il primo è noto come lemma di Steinitz (o teorema dello scambio), il secondo come teorema della dimensione.

Lemma 4.39. *Sia $V = \mathcal{L}(\mathbf{v_1}, \dots, \mathbf{v_m})$, allora ogni insieme $U = \{\mathbf{u_1}, \dots, \mathbf{u_q}\} \subseteq V$ con $q > m$ è linearmente dipendente.*

Dimostrazione. Se $V = \{\mathbf{0}\}$, allora ogni insieme $U \subseteq V$ è linearmente dipendente. Se $V \neq \{\mathbf{0}\}$, grazie all'algoritmo di eliminazione esiste una base di V che, senza perdere in generalità, possiamo assumere essere $\mathcal{B} = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$, dove $n \leq m$. Di conseguenza i vettori di U possono essere decomposti come

$$\mathbf{u_j} = \sum_{i=1}^{n} a_{ij} \cdot \mathbf{v_i},$$

per ogni $1 \leq j \leq q$. Consideriamo ora la combinazione lineare

$$\mathbf{0} = \sum_{j=1}^{q} t_j \cdot \mathbf{u_j} = \sum_{j=1}^{q} t_j \cdot \left(\sum_{i=1}^{n} a_{ij} \cdot \mathbf{v_i} \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{q} t_j \, a_{ij} \right) \cdot \mathbf{v_i}.$$

Dato che $\mathcal{B}$ è indipendente, l'unica combinazione dei suoi vettori con risultato $\mathbf{0}$ è quella banale. Di conseguenza, deve valere

$$\sum_{j=1}^{q} a_{ij} \, t_j = 0$$

per ogni $1 \leq i \leq n$. Abbiamo pertanto un sistema lineare omogeneo nelle incognite $t_1, \dots, t_q$ la cui matrice associata è $A = [a_{ij}] \in \mathrm{Mat}(n, q; \mathbb{K})$. Grazie al Lemma 3.20, sappiamo che $r(A) \leq \min(n, q) = n < q$. Quindi, per il Teorema di Rouché-Capelli, esistono infinite soluzioni dipendenti da $q - r(A) > 0$ parametri. Questo implica infine che esistono combinazioni lineari non banali dei vettori di U che restituiscono il vettore nullo e quindi U è per definizione linearmente dipendente. $\square$

Teorema 4.40. *Sia V uno spazio vettoriale finitamente generato, allora ogni base di V è composta dallo stesso numero di vettori.*

Dimostrazione. Siano $\mathcal{B}_1$ e $\mathcal{B}_2$ due basi dello spazio V, con $n_1 = |\mathcal{B}_1|$ e $n_2 = |\mathcal{B}_2|$. Poiché $\mathcal{B}_1$ é un insieme di generatori e $\mathcal{B}_2$ è linearmente indipendente, per il Lemma 4.39 deve valere $n_2 \leq n_1$. Ovviamente, scambiando i ruoli di $\mathcal{B}_1$ e $\mathcal{B}_2$ deve valere anche $n_1 \leq n_2$ e quindi $n_1 = n_2$. $\square$

Definizione 4.41. *La dimensione di uno spazio vettoriale V è*

$$\dim(V) = \begin{cases} 0 & \textit{se} \quad V = \{\mathbf{0}\}, \\ |\mathcal{B}| & \textit{se} \quad V \neq \{\mathbf{0}\} \textit{ è finitamente generato}, \\ +\infty & \textit{se} \quad V \textit{ non è finitamente generato}, \end{cases}$$

dove nel secondo caso $\mathcal{B}$ è una base di V.

Esempio 4.42. Grazie alle basi canoniche che abbiamo introdotto nell'Esempio 4.23, abbiamo:

- $\dim(\mathbb{K}^n) = n$;
- $\dim(\mathrm{Mat}(m, n; \mathbb{K})) = mn$;
- $\dim(\mathbb{T}_a(n; \mathbb{K})) = \dim(\mathbb{T}_b(n; \mathbb{K})) = \dim(\mathbb{S}(n; \mathbb{K})) = \frac{n(n+1)}{2}$;
- $\dim(\mathbb{T}_{sa}(n; \mathbb{K})) = \dim(\mathbb{T}_{sb}(n; \mathbb{K})) = \dim(\mathbb{A}(n; \mathbb{K})) = \frac{n(n-1)}{2}$;
- $\dim(\mathbb{D}(n; \mathbb{K})) = n$;
- $\dim(\mathbb{K}[x]_n) = n + 1$.

Dal Teorema della dimensione seguono alcune conseguenze immediate molto utili nelle applicazioni e nella risoluzione di esercizi, in quanto semplificano la verifica che un insieme di vettori costituisca una base, qualora si conosca la dimensione dello spazio vettoriale in cui si sta lavorando.

Corollario 4.43. *Sia V uno spazio vettoriale con $\dim(V) = n > 0$ e sia U un sottoinsieme di V avente cardinalità $|U| = m$. Allora:*

i) se U è linearmente indipendente, allora $m \leq n$. Se inoltre $m = n$, allora U è una base di V;
ii) se U è un insieme di generatori, allora $m \geq n$. Se inoltre $m = n$, allora U è una base di V.

Dimostrazione. Verifichiamo singolarmente i singoli enunciati.

i) Per ipotesi esiste un insieme di generatori di V di cardinalità n, pertanto dal Lemma 4.39 abbiamo che ogni insieme composto da $m > n$ vettori è linearmente dipendente. Sia ora $U = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ indipendente. Per quanto abbiamo appena dimostrato, per ogni $\mathbf{v} \in V$ l'insieme $U \cup \{\mathbf{v}\}$ è linearmente dipendente. Questo significa che esiste una combinazione lineare non banale del tipo
$$\mathbf{0} = t \cdot \mathbf{v} + \sum_{i=1}^{n} t_i \cdot \mathbf{v_i}.$$
Necessariamente $t \neq 0$, altrimenti l'insieme U sarebbe dipendente. Di conseguenza vale
$$\mathbf{v} = \sum_{i=1}^{n} -\frac{t_i}{t} \cdot \mathbf{v_i},$$
il che significa che U è un insieme di generatori di V e quindi una base;

ii) per l'algoritmo di eliminazione, ogni insieme U di generatori contiene una base di V, di conseguenza la sua cardinalità deve essere maggiore od uguale ad n. Se $|U| = n$, ovviamente la base contenuta deve coincidere con l'insieme stesso. $\square$

Concludiamo infine la sezione con un secondo algoritmo utile alla costruzione di una base di uno spazio vettoriale finitamente generato.

Proposizione 4.44. *Sia V uno spazio vettoriale di dimensione $\dim(V) = n > 0$. Allora ogni insieme $U = \{\mathbf{u_1}, \ldots, \mathbf{u_m}\} \subset V$ indipendente può essere esteso ad una base di V.*

Dimostrazione. La dimostrazione si basa sul cosiddetto algoritmo di completamento. Se $m = n$, allora per il Corollario 4.43 l'insieme U è una base di V. Se invece $m < n$, allora U non è un insieme di generatori di V e quindi esiste un vettore $\mathbf{v} \in V \backslash \mathcal{L}(\mathbf{u_1}, \ldots, \mathbf{u_m})$. Verifichiamo che l'insieme $\tilde{U} = U \cup \{\mathbf{v}\}$, di cardinalità $m + 1$, è linearmente indipendente. Consideriamo la combinazione lineare
$$\mathbf{0} = t \cdot \mathbf{v} + \sum_{i=1}^{m} t_i \cdot \mathbf{u_i}.$$

Se $t = 0$ anche tutti gli altri coefficienti devono essere uguali a zero, poiché U è indipendente. Se $t \neq 0$, come nella dimostrazione del Corollario 4.43 arriviamo a

$$\mathbf{v} = \sum_{i=1}^{m} -\frac{t_i}{t} \cdot \mathbf{u_i}.$$

Ma questo è impossibile, dato che $\mathbf{v} \notin \mathcal{L}(\mathbf{u_1}, \ldots, \mathbf{u_m})$. A questo punto ripetiamo la procedura sull'insieme $\tilde{U}$. Dato che lo spazio V è finitamente generato, l'algoritmo termina in un numero finito di iterazioni. $\square$

Esempio 4.45. Sia $V = \mathbb{R}[x]_2$ ed $U = \{1 + x + x^2, 1 - x + x^2\}$. L'insieme è linearmente indipendente, infatti

$$t_1 \cdot (1 + x + x^2) + t_2 \cdot (1 - x + x^2) = (t_1 + t_2) \cdot 1 + (t_1 - t_2) \cdot x + (t_1 + t_2) \cdot x^2 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2.$$

se e solo se $t_1 = t_2 = 0$. Pertanto possiamo utilizzare l'algoritmo di completamento per ottenere una base di V. Ad esempio, è semplice verificare che $1 \notin \mathcal{L}(1 + x + x^2, 1 - x + x^2)$ e quindi $\tilde{U} = U \cup \{1\}$ è linearmente indipendente. Poiché $\dim(V) = 3$, l'insieme $\tilde{U}$ é una base di V.

4.6. Dimensione di sottospazi e formula di Grassmann

In questa sezione utilizziamo i concetti di base e dimensione per studiare le proprietà dei sottospazi. Iniziamo con un risultato intuitivo sulla dimensione dei sottospazi, che però come sempre deve essere rigorosamente dimostrato.

Proposizione 4.46. *Sia V uno spazio vettoriale finitamente generato ed $U \subseteq V$ un sottospazio. Allora U è finitamente generato ed in particolare $\dim(U) \leq \dim(V)$. L'uguaglianza vale se e soltanto se $U = V$.*

Dimostrazione. Il caso $U = \{\mathbf{0}\}$ è ovvio. Supponiamo $U \neq \{\mathbf{0}\}$. L'algoritmo di completamento ci permette di costruire una base $\mathcal{B} = \{\mathbf{u_1}, \ldots, \mathbf{u_m}\}$ di U a condizione che l'algoritmo termini in un numero finito di iterazioni. Nel nostro caso questo è garantito dal fatto che la cardinalità massima di un sottoinsieme indipendente di U è pari a $\dim(V)$, per il Corollario 4.43. Di conseguenza U è finitamente generato e $\dim(U) \leq \dim(V)$. Se $\dim(U) = \dim(V)$ allora $\mathcal{B}$ è anche una base di V, da cui segue che $U = \mathcal{L}(\mathbf{u_1}, \ldots, \mathbf{u_m}) = V$. Al contrario, se $U = V$ l'enunciato è ovvio. $\square$

Esempio 4.47. Data $A \in \mathrm{Mat}(m, n; \mathbb{K})$, sappiamo che il suo nucleo è un sottospazio di $\mathrm{Mat}(n, 1; \mathbb{K})$. Ogni soluzione del sistema omogeneo associato ad A può essere scritta in modo unico come combinazione lineare di $n - r(A)$ colonne, ottenibili attraverso l'applicazione del metodo di eliminazione di Gauss. Di conseguenza, queste colonne costituiscono una base di $\ker(A)$. Se chiamiamo nullità $k(A)$ la dimensione del nucleo della matrice, allora vale $k(A) = \dim(\ker(A)) = n - r(A) \leq n$.

È interessante osservare come la proposizione precedente implica che per verificare l'uguaglianza di due spazi vettoriali è sufficiente dimostrare che uno dei due è contenuto nell'altro e che essi hanno la stessa dimensione.

Il nostro obbiettivo è ora studiare come si comporta la dimensione rispetto alle operazioni tra sottospazi. A tal fine enunciamo preliminarmente il seguente lemma.

Lemma 4.48. *Siano V uno spazio vettoriale ed $U, W \subseteq V$ due sottospazi. Allora:*

i) $\dim(U \cap W) \leq \min(\dim(U), \dim(W))$;

ii) l'unione di una base di U ed una di W costituisce un insieme di generatori di $U + W$ ed inoltre $\max(\dim(U), \dim(W)) \leq \dim(U + W) \leq \dim(U) + \dim(W)$.

Dimostrazione. Il punto (i) e la prima disuguaglianza del punto (ii) sono conseguenza diretta della Proposizione precedente. Per le restanti affermazioni, siano $\mathcal{B}_U = \{\mathbf{u_1}, \ldots, \mathbf{u_h}\}$ e $\mathcal{B}_W = \{\mathbf{w_1}, \ldots, \mathbf{w_k}\}$ rispettivamente basi di U e W. Ogni vettore $\mathbf{v} \in U + W$ può essere scritto come somma di una coppia $(\mathbf{u}, \mathbf{w}) \in U \times W$. Di conseguenza

$$\mathbf{v} = \mathbf{u} + \mathbf{w} = \sum_{i=1}^{h} \alpha_i \cdot \mathbf{u_i} + \sum_{j=1}^{k} \beta_j \cdot \mathbf{w_j}.$$

Questo significa che $\mathcal{B}_U \cup \mathcal{B}_W$ è un insieme di generatori di $U + W$ e quindi

$$\dim(U + W) \leq |\mathcal{B}_U \cup \mathcal{B}_W| \leq |\mathcal{B}_U| + |\mathcal{B}_W| = \dim(U) + \dim(W). \qquad \square$$

Possiamo a questo punto presentare il seguente importante teorema dovuto a Grassmann.

Teorema 4.49. *Siano V uno spazio vettoriale ed $U, W \subseteq V$ due sottospazi finitamente generati. Allora:*

$$\dim(U + W) + \dim(U \cap W) = \dim(U) + \dim(W).$$

Dimostrazione. Chiamiamo $\dim(U) = h$, $\dim(W) = k$ e $\dim(U \cap W) = l$. Per il Lemma precedente, vale $l \leq \min(h, k)$. Sia $\mathcal{B}_{U\cap W} = \{\mathbf{v_1}, \ldots, \mathbf{v_l}\}$ una base dell'intersezione, allora grazie all'algoritmo di completamento possiamo estenderla rispettivamente ad una base di U ed una di W :

$$\mathcal{B}_U = \{\mathbf{v_1}, \ldots, \mathbf{v_l}, \mathbf{u_1}, \ldots, \mathbf{u_{h-l}}\} \qquad \text{e} \qquad \mathcal{B}_W = \{\mathbf{v_1}, \ldots, \mathbf{v_l}, \mathbf{w_1}, \ldots, \mathbf{w_{k-l}}\}.$$

Se potessimo dimostrare che $\mathcal{B}_U \cup \mathcal{B}_W = \{\mathbf{v_1}, \ldots, \mathbf{v_l}, \mathbf{u_1}, \ldots, \mathbf{u_{h-l}}, \mathbf{w_1}, \ldots, \mathbf{w_{k-l}}\}$ è una base di $U + W$, il teorema seguirebbe facilmente, in quanto varrebbe

$$\dim(U + W) = l + (h - l) + (k - l) = h + k - l = \dim(U) + \dim(W) - \dim(U \cap W).$$

Verifichiamo quindi la validità delle ipotesi del Teorema 4.32 per l'insieme $\mathcal{B}_U \cup \mathcal{B}_W$:

- $\mathcal{B}_U \cup \mathcal{B}_W$ è un insieme di generatori di $U + W$, in quanto è l'unione di una base di U ed una di W;
- consideriamo la combinazione lineare

$$\mathbf{0} = \sum_{s=1}^{l} t_s \cdot \mathbf{v_s} + \sum_{i=1}^{h-l} \alpha_i \cdot \mathbf{u_i} + \sum_{j=1}^{k-l} \beta_j \cdot \mathbf{w_j}.$$

  Chiamiamo

$$\mathbf{u} = \sum_{s=1}^{l} t_s \cdot \mathbf{v_s} + \sum_{i=1}^{h-l} \alpha_i \cdot \mathbf{u_i} \in U$$

  e

$$\mathbf{w} = \sum_{j=1}^{k-l} \beta_j \cdot \mathbf{w_j} \in W.$$

  Da $\mathbf{u} + \mathbf{w} = \mathbf{0}$ segue che $\mathbf{w} = -\mathbf{u} \in U$ e quindi $\mathbf{w} \in U \cap W$. Di conseguenza, esiste la decomposizione

$$\mathbf{w} = \sum_{i=1}^{l} \gamma_i \cdot \mathbf{v_i}$$

  rispetto alla base $\mathcal{B}_{U\cap W}$. Allora vale

$$\mathbf{0} = \mathbf{w} - \mathbf{w} = \sum_{i=1}^{l} \gamma_i \cdot \mathbf{v_i} - \sum_{j=1}^{k-l} \beta_j \cdot \mathbf{w_j}.$$

  Poiché i vettori di questa combinazione lineare costituiscono la base B_W, necessariamente tutti i coefficienti devono essere uguali a zero: $\gamma_1 = \cdots = \gamma_l = \beta_1 = \cdots = \beta_{k-l} = 0$. Allora $\mathbf{w} = \mathbf{0} = \mathbf{u}$ e quindi $t_1 = \cdots = t_l = \alpha_1 = \cdots = \alpha_{h-l} = 0$, da cui otteniamo l'indipendenza di $\mathcal{B}_U \cup \mathcal{B}_W$. $\square$

Esempio 4.50. Siano $V = \mathrm{Mat}(3,1;\mathbb{Q})$ ed $U = \ker(A)$, $W = \ker(B)$ dove

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Allora,

$$U = \left\{ s_1 \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} + s_2 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \\ -2 \end{bmatrix}, \ s_1, s_2 \in \mathbb{Q} \right\}, \quad W = \left\{ t_1 \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \ t_1, t_2 \in \mathbb{Q} \right\}.$$

I due sottospazi hanno $\dim(U) = 2$ e $\dim(W) = 2$. L'intersezione dei due sottospazi è

$$\begin{aligned} U_1 \cap U_2 &= \ker\left(\begin{bmatrix} A \\ B \end{bmatrix} \right) = \ker\left(\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \right) = \\ &= \ker\left(\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right) = \left\{ \alpha \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \ \alpha \in \mathbb{Q} \right\}. \end{aligned}$$

Di conseguenza $\dim(U+W) = \dim(U) + \dim(W) - \dim(U \cap W) = 2 + 2 - 1 = 3$. Dall'analisi delle basi dei sottospazi e della loro intersezione, e dalla dimostrazione del Teorema di Grassmann, abbiamo che

$$\mathcal{B}_{U+W} = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} \right\}.$$

Completiamo la sezione generalizzando i precedenti risultati per un numero arbitrario di sottospazi. In particolare, nel Corollario 4.52 presenteremo un risultato significativo caratterizzante le somme dirette, che ci verrà utile quando parleremo di applicazioni lineari diagonalizzabili nel Capitolo 6.

Proposizione 4.51. *Siano V uno spazio vettoriale finitamente generato ed $U_1, \ldots, U_m$ sottospazi di V. Nell'ipotesi $m \geq 2$, definiamo*

$$W = \sum_{i=1}^{m} U_i \qquad ed \qquad U_{\widehat{k}} = \sum_{\substack{i=1 \\ i \neq k}}^{m} U_i \,,$$

per ogni $k = 1, \ldots, m$. Allora:

i) per ogni k vale $\dim(W) = \dim(U_k) + \dim(U_{\widehat{k}}) - \dim(U_k \cap U_{\widehat{k}})$;

ii) se $\dim(U_k \cap U_{\widehat{k}}) = 0$ *per ogni k, allora* $\dim(W) = \sum_{i=1}^{m} \dim(U_i)$.

Dimostrazione. La prima formula è un'applicazione del Teorema 4.49, se teniamo conto della decomposizione $W = U_k + U_{\widehat{k}}$. Il secondo risultato lo possiamo verificare attraverso l'induzione su m. Il caso $m = 2$ coincide con la formula di Grassmann. Ipotizziamo ora che il lemma sia vero fino ad $m - 1$. Allora

$$\dim(W) = \dim(U_m) + \dim(U_{\widehat{m}}) - \dim(U_m \cap U_{\widehat{m}}) = \dim(U_m) + \sum_{i=1}^{m-1} \dim(U_i) - 0 = \sum_{i=1}^{m} \dim(U_i),$$

dove nella seconda uguaglianza abbiamo utilizzato l'ipotesi induttiva. □

Corollario 4.52. *Lo spazio vettoriale V finitamente generato è somma diretta di $U_1, \ldots, U_m$ se e solo se*

i) $\dim(V) = \sum_{i=1}^{m} \dim(U_i)$;

ii) $\dim(U_k \cap U_{\widehat{k}}) = 0$ *per ogni* $k = 1, \ldots, m$.

Dimostrazione. Dato $W = \sum_{i=1}^{m} U_i$, per la Proposizione 4.18, lo spazio V è somma diretta se e solo se

$$V = W \qquad \text{ed} \qquad U_k \cap U_{\widehat{k}} = \{\mathbf{0}\} \text{ per ogni } k = 1, \ldots, m.$$

La seconda condizione è ovviamente equivalente a $\dim(U_k \cap U_{\widehat{k}}) = 0$. Assumendo che sia vera tale ipotesi, dimostriamo che la richiesta (i) dell'enunciato è equivalente alla condizione $V = W$. Da un lato, se V coincide con la somma dei sottospazi allora la condizione sulla dimensione è uno dei contenuti della Proposizione 4.51. Viceversa, se è vera l'ipotesi (i) allora $\dim(W) = \sum_{i=1}^{m} \dim(U_i) = \dim(V)$. Dato che per costruzione $W \subseteq V$, l'uguaglianza delle dimensioni implica $W = V$. $\square$

4.7. Spazi delle righe e delle colonne di una matrice

In questa sezione andremo ad indagare le proprietà degli spazi generati dalle righe e dalle colonne di una matrice. Questo studio è importante dal punto di vista sia teorico, sia applicativo. Per quanto riguarda il primo aspetto, grazie a questa analisi potremo in particolare fornire una caratterizzazione intrinseca del rango di una matrice. Dal punto di vista delle applicazioni, otterremo invece degli strumenti che, in cooperazione con la mappa delle componenti, risulteranno molto efficaci per studiare le proprietà di un generico spazio vettoriale, dei suoi sottospazi e delle loro basi. Ora ci concentreremo sull'enunciazione e dimostrazione dei risultati principali riguardanti gli spazi righe e colonne, lasciando la discussione dei risvolti applicativi alla Sezione 4.8.

Definizione 4.53. *Data $A \in \mathrm{Mat}(m, n; \mathbb{K})$, gli spazi delle righe e delle colonne della matrice sono rispettivamente $R(A) = \mathcal{L}(A_{R(1)}, \ldots, A_{R(m)}) \subseteq \mathrm{Mat}(1, n; \mathbb{K})$ e $C(A) = \mathcal{L}(A_{C(1)}, \ldots, A_{C(n)}) \subseteq \mathrm{Mat}(m, 1; \mathbb{K})$.*

Ad esempio, sia

$$A = \begin{bmatrix} 1 & e & \pi \\ 0 & 1 & \imath \end{bmatrix} \in \mathrm{Mat}(2, 3; \mathbb{C}).$$

Allora

$$R(A) = \mathcal{L}\left(\begin{bmatrix} 1 & e & \pi \end{bmatrix}, \begin{bmatrix} 0 & 1 & \imath \end{bmatrix}\right) \subset \mathrm{Mat}(1, 3; \mathbb{C})$$

e

$$C(A) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} e \\ 1 \end{bmatrix}, \begin{bmatrix} \pi \\ \imath \end{bmatrix}\right) = \mathrm{Mat}(2, 1; \mathbb{C}).$$

Il teorema fondamentale che vogliamo dimostrare è il seguente.

Teorema 4.54. *Data $A \in \mathrm{Mat}(m, n; \mathbb{K})$ vale $r(A) = \dim(R(A)) = \dim(C(A))$.*

La dimostrazione di questo risultato è piuttosto articolata, pertanto la spezziamo in più parti. Per prima cosa verifichiamo le affermazioni riguardanti lo spazio delle righe.

Proposizione 4.55. *Sia $A \in \mathrm{Mat}(m, n; \mathbb{K})$ ed S una sua riduzione a scala. Allora:*

i) $R(A) = R(S)$;

ii) $\dim(R(A)) = \dim(R(S)) = r(A)$ *e le righe non nulle di S costituiscono una base di $R(A)$.*

Dimostrazione. Come sappiamo, S è ottenuta attraverso una sequenza di operazioni elementari eseguite sulla matrice A. Nell'applicazione di ogni singola operazione elementare sostituiamo ad una data riga di A una combinazione lineare delle righe stesse. Questo significa che lo spazio delle righe della matrice trasformata è contenuto in $R(A)$ e quindi vale l'inclusione $R(S) \subseteq R(A)$. D'altra parte sappiamo che le operazioni elementari sono invertibili, cioè è possibile ottenere A attraverso operazioni elementari sulle righe di S. Di conseguenza vale $R(A) \subseteq R(S)$ e quindi $R(A) = R(S)$, da cui ovviamente segue anche $\dim(R(A)) = \dim(R(S))$. La dimostrazione si completa provando che le $r(A)$ righe non nulle di S costituiscono una base di $R(S)$. Per costruzione abbiamo

$$S = \begin{bmatrix} 0 & \dots & p_1 & \dots & s_{1q_2} & \dots & s_{1q_r} & \dots & s_{1n} \\ 0 & \dots & 0 & \dots & p_2 & \dots & s_{2q_r} & \dots & s_{2n} \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & p_r & \dots & s_{rn} \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 \end{bmatrix} \in \mathrm{Mat}(m, n; \mathbb{K}),$$

dove p_i e q_i sono rispettivamente i pivot e gli indici delle colonne contenenti i pivot di S, $1 \leq i \leq r(A)$. È evidente che le prime $r(A)$ righe di S sono un insieme di generatori di $R(S)$. Per verificare l'indipendenza, consideriamo la combinazione lineare

$$0_{1n} = \sum_{i=1}^{r(A)} t_i \cdot S_{R(i)} = \begin{bmatrix} 0 & \dots & t_1 p_1 & \dots & t_2 p_2 + t_1 s_{1q_2} & \dots & t_r p_r + \sum_{i=1}^{r-1} t_i s_{iq_r} & \dots & \sum_{i=1}^{r} t_i s_{in} \end{bmatrix}.$$

Dovrebbe essere chiaro che questa equazione matriciale nelle incognite $t_1, \dots, t_r$ ha come unica soluzione quella banale. □

La relazione tra la dimensione dello spazio delle colonne di una matrice ed il suo rango è meno visibile. Questo fatto è ragionevole, in quanto abbiamo definito la riduzione a scala di una matrice attraverso operazioni elementari sulle sue righe, il cui effetto sulle relative colonne non è immediatamente comprensibile. Per completare la dimostrazione del Teorema 4.54 dovremo quindi dimostrare un altro significativo teorema, dovuto a Kronecker, che stabilisce un legame profondo tra il rango di una matrice ed il determinante delle sue sottomatrici quadrate.

Lemma 4.56. *Sia* $A \in \mathrm{Mat}(m, n; \mathbb{K})$.

i) Se $m = n$, *l'insieme delle righe e l'insieme delle colonne di* A *sono linearmente indipendenti se e solo se* $\det(A) \neq 0$;

ii) se $m > n$, *l'insieme delle colonne di* A *è linearmente indipendente se e solo se esiste una sottomatrice* $B \in \mathrm{Mat}(n, n; \mathbb{K})$ *di* A *tale che* $\det(B) \neq 0$;

iii) se $n > m$, *l'insieme delle righe di* A *è linearmente indipendente se e solo se esiste una sottomatrice* $B \in \mathrm{Mat}(m, m; \mathbb{K})$ *di* A *tale che* $\det(B) \neq 0$.

Dimostrazione. Verifichiamo singolarmente ogni punto del lemma.

i) Per il Teorema 3.69 sappiamo che $\det(A) \neq 0$ se e solo se $r(A) = n$ e quindi, per la Proposizione 4.55, se e solo se le n righe di A sono indipendenti. Grazie alla Proposizione 3.84 sappiamo che $\det(A^T) = \det(A)$, quindi è possibile ripetere lo stesso ragionamento sulle righe di A^T, che ovviamente corrispondono alle colonne di A.

ii) Consideriamo la combinazione lineare
$$0_{m1} = \sum_{i=1}^{n} t_i \cdot A_{C(i)}.$$
Grazie al Lemma 3.9, l'equazione è equivalente ad
$$A * \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix} = 0_{m1}.$$
Per il Teorema di Rouché-Capelli, questo sistema omogeneo ha unica soluzione banale se e solo se $r(A) = n$. Per la Proposizione 4.55, questo è vero se e solo se esistono n righe di A linearmente indipendenti. Considerando la matrice B formata da tali righe, la tesi segue dall'applicazione del primo punto del lemma.

iii) L'insieme delle righe di A corrisponde all'insieme delle colonne di A^T. Pertanto l'ultimo enunciato si dimostra facilmente applicando il punto ii del lemma alla matrice A^T, in cooperazione con la Proposizione 3.84. □

Teorema 4.57. *Sia $A \in \mathrm{Mat}(m, n; \mathbb{K})$, allora $r(A) = p$ se e solo se:*

i) esiste una sottomatrice $B \in \mathrm{Mat}(p, p; \mathbb{K})$ di A tale che $\det(B) \neq 0$;

ii) ogni sottomatrice $C \in \mathrm{Mat}(p+1, p+1; \mathbb{K})$ di A ottenuta orlando la matrice B soddisfa $\det(C) = 0$.

Dimostrazione. Supponiamo che $r(A) = p$, allora il numero massimo di righe di A costituenti un insieme linearmente indipendente è p. Consideriamo la matrice $\tilde{B} \in \mathrm{Mat}(p, n; \mathbb{K})$ formata da uno di questi insiemi di righe, allora grazie al Lemma 4.56 sappiamo che esiste una sottomatrice $B \in \mathrm{Mat}(p, p; \mathbb{K})$ di $\tilde{B}$ (e quindi di A) tale che $\det(B) \neq 0$. D'altra parte, ragionando in maniera analoga si dimostra che ogni sottomatrice $C \in \mathrm{Mat}(p + 1, p + 1; \mathbb{K})$ di A deve soddisfare $\det(C) = 0$, dato che non esistono insiemi linearmente indipendenti composti da $p + 1$ righe di A.

Viceversa, sia $B \in \mathrm{Mat}(p, p; \mathbb{K})$ sottomatrice di A tale che $\det(B) \neq 0$. Sempre per il Lemma 4.56, segue che le righe della matrice $\tilde{B} \in \mathrm{Mat}(p, n; \mathbb{K})$ contenente B formano un insieme linearmente indipendente. Infine, se per ogni matrice $C \in \mathrm{Mat}(p+1, p+1; \mathbb{K})$ ottenuta orlando B vale $\det(C) = 0$, questo significa che le righe di ogni matrice $\tilde{C} \in \mathrm{Mat}(p + 1, n; \mathbb{K})$ ottenuta orlando $\tilde{B}$ formano un insieme linearmente dipendente. Di conseguenza, per la Proposizione 4.30 si deduce che l'ulteriore riga di $\tilde{C}$ è combinazione lineare delle p righe di $\tilde{B}$. Pertanto le p righe di $\tilde{B}$ formano un insieme di generatori di $R(A)$ e quindi una base di tale spazio, da cui $r(A) = p$. □

Il Teorema di Kronecker viene anche detto Teorema dei minori orlati, concetto che ora ci apprestiamo a definire.

Definizione 4.58. *Sia $A \in \mathrm{Mat}(m, n; \mathbb{K})$ ed $B \in \mathrm{Mat}(p, p; \mathbb{K})$ una sottomatrice di A. Allora il determinante di B viene detto minore di ordine p di A.*

Il Teorema 4.57 afferma che il rango di una matrice coincide con l'ordine massimo dei suoi minori diversi da zero. A questo punto possiamo finalmente completare la dimostrazione del Teorema 4.54.

Corollario 4.59. *Sia $A \in \mathrm{Mat}(m, n; \mathbb{K})$, allora:*

i) $r(A) = r(A^T)$;

ii) $\dim(C(A)) = r(A)$.

Dimostrazione. Poiché il determinante non cambia a seguito di trasposizione, si ha che i minori di A sono uguali ai minori di A^T e quindi per il Teorema di Kronecker vale $r(A) = r(A^T)$. Di conseguenza $\dim(C(A)) = \dim(R(A^T)) = r(A^T) = r(A)$. $\square$

Esempio 4.60. Sia

$$A = \begin{bmatrix} 0 & 1 & 0 & 3 \\ 1 & -1 & -1 & 2 \\ 1 & 0 & -1 & 5 \end{bmatrix} \in \mathrm{Mat}(3,4;\mathbb{Q}).$$

Allora consideriamo la sottomatrice quadrata

$$B = A_{\widehat{3}\widehat{4}\,\widehat{3}} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \in \mathrm{Mat}(2,2;\mathbb{Q}) \qquad \Rightarrow \qquad \det(B) = -1 \neq 0 \qquad \Rightarrow \qquad r(A) \geq 2.$$

Esistono due possibili orlature di B :

$$C_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -1 & -1 \\ 1 & 0 & -1 \end{bmatrix} \in \mathrm{Mat}(3,3;\mathbb{Q}) \qquad \Rightarrow \qquad \det(C_1) = 0,$$

$$C_2 = \begin{bmatrix} 0 & 1 & 3 \\ 1 & -1 & 2 \\ 1 & 0 & 5 \end{bmatrix} \in \mathrm{Mat}(3,3;\mathbb{Q}) \qquad \Rightarrow \qquad \det(C_2) = 0,$$

da cui $r(A) = 2$. Nella dimostrazione del teorema di Kronecker e del Corollario 4.59, è stato provato anche che le righe e le colonne di una matrice A che compongono una sottomatrice quadrata B di ordine massimo avente $\det(B) \neq 0$, costituiscono rispettivamente una base di $R(A)$ e di $C(A)$. Nel caso in esempio, questo implica che $\mathcal{B}_{R(A)} = \left\{A_{R(1)}, A_{R(2)}\right\}$ e $\mathcal{B}_{C(A)} = \left\{A_{C(1)}, A_{C(2)}\right\}$.

Dalla Proposizione 4.55 sappiamo che le righe della matrice ridotta a scala costituiscono una base dello spazio delle righe della matrice di partenza. Concludiamo la sezione riequilibrando la situazione rispetto alle colonne.

Proposizione 4.61. *Sia $A \in \mathrm{Mat}(m,n;\mathbb{K})$ avente $r(A) = r$ e sia S una sua riduzione a scala, con $q_1, \ldots, q_r$ gli indici delle colonne di S contenenti i pivot. Allora $\{A_{C(q_1)}, \ldots, A_{C(q_r)}\}$ è una base di $C(A)$.*

Dimostrazione. Dal Teorema 4.54 sappiamo che $\dim(C(A)) = r(A) = r$, per cui è sufficiente dimostrare che l'insieme $\{A_{C(q_1)}, \ldots, A_{C(q_r)}\}$ è linearmente indipendente. Consideriamo la matrice $\tilde{A} \in \mathrm{Mat}(m,r;\mathbb{K})$ composta da tali colonne di A. È semplice osservare che se eseguiamo su $\tilde{A}$ le stesse operazioni elementari che servono a ridurre la matrice A ad S, otteniamo la matrice $\tilde{S} \in \mathrm{Mat}(m,r;\mathbb{K})$ composta solo dalle colonne $\{S_{C(q_1)}, \ldots, S_{C(q_r)}\}$ di S contenenti i pivot. Dovrebbe essere evidente che $r(\tilde{A}) = r(\tilde{S}) = r$, da cui segue la tesi. $\square$

È importante osservare che $\{S_{C(q_1)}, \ldots, S_{C(q_r)}\}$ è una base di $C(S)$ ma non di $C(A)$. Infatti in generale $C(A) \neq C(S)$.

Esempio 4.62. Sia

$$A = \begin{bmatrix} 0 & 1 & 0 & 3 \\ 1 & -1 & -1 & 2 \\ 1 & 0 & -1 & 5 \end{bmatrix} \xrightarrow[R_3 \leftrightarrow R_2]{R_3 \leftrightarrow R_1} \begin{bmatrix} 1 & 0 & -1 & 5 \\ 0 & 1 & 0 & 3 \\ 1 & -1 & -1 & 2 \end{bmatrix} \xrightarrow{R_3 - R_1 + R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 0 & -1 & 5 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} = S.$$

Quindi $\mathcal{B}_{R(A)} = \mathcal{B}_{R(S)} = \left\{S_{R(1)}, S_{R(2)}\right\}$, $\mathcal{B}_{C(A)} = \left\{A_{C(1)}, A_{C(2)}\right\}$ e $\mathcal{B}_{C(S)} = \left\{S_{C(1)}, S_{C(2)}\right\}$. Osserviamo che in questo caso $C(A) \neq C(S)$.

4.8. Mappa delle componenti e spazi delle righe e delle colonne

Concludiamo il capitolo sugli spazi vettoriali presentando alcune conseguenze significative dei risultati dimostrati nella sezione precedente, che possano essere utili in particolare nella risoluzione degli esercizi sugli spazi vettoriali. Chiaramente, la prima importante ed immediata implicazione del Teorema 4.54 è che il rango di una matrice é una quantità intrinseca associata alla matrice stessa, che non dipende quindi dalla sua particolare riduzione a scala. Un'altra implicazione diretta del teorema riguarda la teoria dei sistemi lineari.

Corollario 4.63. *Il sistema lineare associato alla matrice completa $[A|B]$ ha soluzione se e solo se $B \in C(A)$.*

Dimostrazione. Per il teorema di Rouché-Capelli, la soluzione esiste se e solo se $r(A) = r([A|B])$. Per il Teorema 4.54, questo è possibile se e solo se $\dim(C(A)) = \dim(C([A|B]))$. Poiché $C(A) \subseteq C([A|B])$, le due dimensioni coincidono se e solo se $C(A) = C([A|B])$, da cui la tesi. □

Le conseguenze più significative dei risultati contenuti nella Sezione 4.7 si hanno nello studio degli spazi vettoriali. Infatti, l'utilizzo congiunto della mappa delle componenti e delle proprietà degli spazi delle righe e delle colonne di una matrice, consentono l'esecuzione sistematica ed efficiente di ogni tipo di calcolo da compiere su un generico spazio. Al fine di illustrare le potenzialità di tali tecniche, anticipiamo alcuni risultati sugli isomorfismi di spazi vettoriali che dimostreremo nel Capitolo 5.

Proposizione 4.64. *Dati V e W spazi vettoriali sul medesimo campo $\mathbb{K}$, sia $f : V \to W$ un isomorfismo. Allora:*

i) *$U \subseteq V$ è un sottospazio se e solo se $f(U) \subseteq W$ è un sottospazio;*
ii) *$\{\mathbf{u_1}, \dots, \mathbf{u_m}\}$ è un insieme di generatori di $U \subseteq V$ se e solo se $\{f(\mathbf{u_1}), \dots, f(\mathbf{u_m})\}$ è un insieme di generatori di $f(U) \subseteq W$;*
iii) *$\{\mathbf{u_1}, \dots, \mathbf{u_m}\}$ è linearmente indipendente se e solo se $\{f(\mathbf{u_1}), \dots, f(\mathbf{u_m})\}$ è linearmente indipendente;*
iv) *$\{\mathbf{u_1}, \dots, \mathbf{u_m}\}$ è una base di $U \subseteq V$ se e solo se $\{f(\mathbf{u_1}), \dots, f(\mathbf{u_m})\}$ è una base di $f(U) \subseteq W$.*

La proposizione precedente è valida in particolare per la mappa delle componenti $\phi_{\mathcal{B}}$. Infatti, dato lo spazio vettoriale V di dimensione n su campo $\mathbb{K}$ e fissata una sua base $\mathcal{B}$, la funzione $\phi_{\mathcal{B}}$ è un isomorfismo tra V ed il corrispondente insieme delle componenti $\mathrm{Im}(\phi_{\mathcal{B}}) = \mathrm{Mat}(n, 1; \mathbb{K})$. Sfruttando queste proprietà ed i risultati della Sezione 4.7, si arriva facilmente alla seguente proposizione.

Proposizione 4.65. *Consideriamo uno spazio vettoriale V di dimensione $n < \infty$ su campo $\mathbb{K}$, una base $\mathcal{B}$ di V ed un sottospazio $U = \mathcal{L}\{\mathbf{u_1}, \dots, \mathbf{u_m}\} \subseteq V$. Sia A la matrice le cui colonne sono le componenti di $\mathbf{u_1}, \dots, \mathbf{u_m}$ rispetto a $\mathcal{B}$:*

$$A = \left[\ \mathbf{u_1}|_{\mathcal{B}}\ \middle|\ \dots\ \middle|\ \mathbf{u_m}|_{\mathcal{B}}\ \right] \in \mathrm{Mat}(n, m; \mathbb{K}).$$

i) *Se S è una riduzione a scala di A, con $q_1, \dots, q_r$ gli indici delle colonne di S contenenti i pivot, allora $\dim(U) = r(A)$ e $\{\mathbf{u_{q_1}}, \dots, \mathbf{u_{q_r}}\}$ è una base di U.*
ii) *Se $\tilde{S}$ è una riduzione a scala di A^T, allora $\dim(U) = r(A^T)$ e le trasposizioni delle righe non nulle di $\tilde{S}$ costituiscono una base di $\phi_{\mathcal{B}}(U)$.*

Esempio 4.66. Vediamo in un paio di esempio come la Proposizione 4.65 costituisca uno strumento molto utile per individuare le basi di spazi e sottospazi vettoriali.

- Siano $V = \mathbb{R}[x]_2$ ed $\mathbf{u_1} = 1 + x,\ \mathbf{u_2} = x + x^2,\ \mathbf{u_3} = 1 + x^2$. Supponiamo di voler verificare se i tre vettori costituiscono una base di V. Rispetto alla base canonica $\mathcal{B}_2 = \{1, x, x^2\}$ di V abbiamo

$$A = \begin{bmatrix} 1+x|_{\mathcal{B}_2} & x+x^2|_{\mathcal{B}_2} & 1+x^2|_{\mathcal{B}_2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \xrightarrow[R_2-R_1-R_3 \to R_3]{R_3 \to R_2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \end{bmatrix} = S.$$

  Allora $\dim(\mathcal{L}(\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3})) = \dim(C(A)) = r(A) = 3 = \dim(V)$, quindi $\mathcal{L}(\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}) = V$ ed $\{\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}\}$ è una base di V.
- Siano $V = \mathbb{R}[x]_3$ ed $\mathbf{u_1} = 1+x,\ \mathbf{u_2} = 1-x,\ \mathbf{u_3} = x+x^2+x^3,\ \mathbf{u_4} = 1+x^2+x^3$. Vogliamo trovare una base dello spazio generato dai quattro vettori. Rispetto alla base canonica $\mathcal{B}_3 = \{1, x, x^2, x^3\}$ di V abbiamo

$$A = \begin{bmatrix} 1+x|_{\mathcal{B}_3} & 1-x|_{\mathcal{B}_3} & x+x^2+x^3|_{\mathcal{B}_3} & 1+x^2+x^3|_{\mathcal{B}_3} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad \Rightarrow$$

$$A^T = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \xrightarrow[R_4-R_1 \leftrightarrow R_4]{-\frac{1}{2}(R_2-R_1) \leftrightarrow R_2} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & -1 & 1 & 1 \end{bmatrix} \xrightarrow[R_4+2R_2-R_3 \to R_4]{R_3-R_2 \to R_3} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \tilde{S}.$$

  Allora $\dim(\mathcal{L}(\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}, \mathbf{u_4})) = r(A^T) = 3$ e la trasposizione delle prime tre righe di $\tilde{S}$ costituiscono una base di $\phi_{\mathcal{B}_3}(\mathcal{L}(\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}, \mathbf{u_4}))$. Questo significa che una base di $\mathcal{L}(\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}, \mathbf{u_4})$ è $\{1 + x, x, x^2 + x^3\}$.

Esempio 4.67. Siano $V = \mathbb{R}[x]_3$, i sottospazi $U = \mathcal{L}(1 + x, 1 - x^2)$ e $W = \{P(x) \in V \mid P(1) = 0\}$ e $\mathcal{B}_3 = \{1, x, x^2, x^3\}$ la base canonica di V. Vogliamo calcolare le dimensioni e le basi di $U, W, U+W, U\cap W$. Consideriamo la solita matrice

$$A = \begin{bmatrix} 1+x|_{\mathcal{B}_3} & 1-x^2|_{\mathcal{B}_3} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & -1 \\ 0 & 0 \end{bmatrix} \xrightarrow[R_3+R_1-R_2 \to R_3]{R_2-R_1 \to R_2} \begin{bmatrix} 1 & 1 \\ 0 & -1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = S.$$

Questo implica $\dim(U) = 2$ e quindi $\{1 + x, 1 - x^2\}$ è una base di U. Per il secondo sottospazio vale

$$W = \{P(x) = a \cdot 1 + b \cdot x + c \cdot x^2 + d \cdot x^3 \mid a + b + c + d = 0\},$$

da cui $d = -a - b - c$ e quindi

$$W = \{P(x) = a \cdot (1 - x^3) + b \cdot (x - x^3) + c \cdot (x^2 - x^3)\}.$$

Di conseguenza $\dim(W) = 3$ ed $\{1 - x^3, x - x^3, x^2 - x^3\}$ è una base di W.

Concentriamoci su $U + W$ e $U \cap W$. Sappiamo che

$$U + W = \mathcal{L}(1 + x, 1 - x^2, 1 - x^3, x - x^3, x^2 - x^3),$$

quindi per calcolare la dimensione ed una base di $U+W$ studiamo la matrice associata ai generatori

$$B = \begin{bmatrix} 1+x|_{\mathcal{B}_3} & 1-x^2|_{\mathcal{B}_3} & 1-x^3|_{\mathcal{B}_3} & x-x^3|_{\mathcal{B}_3} & x^2-x^3|_{\mathcal{B}_3} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & -1 & -1 & -1 \end{bmatrix}.$$

Applichiamo il Teorema di Kronecker, considerando la sottomatrice composta dalle prime quattro colonne di B:

$$\det\left(\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & -1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & -1 \end{bmatrix}\right) - \det\left(\begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & -1 & -1 \end{bmatrix}\right) = 1+1 = 2 \neq 0.$$

Di conseguenza, le prime quattro colonne di B ed i corrispondenti vettori generatori sono indipendenti. Inoltre, $r(B) = 4 = \dim(U+W)$ e quindi $U+W = V$, una cui possibile base è $\mathcal{B}_3$. Dal Teorema di Grassmann segue che $\dim(U \cap W) = \dim(U) + \dim(W) - \dim(U+W) = 2+3-4 = 1$. Poiché $1-x^2 = 1\cdot(1-x^3) - 1\cdot(x^2-x^3)$, abbiamo $1-x^2 \in U \cap W$ e quindi questo binomio costituisce una base del sottospazio intersezione.

Concludiamo infine il capitolo con una discussione sistematica sulla rappresentazione parametrica ed algebrica di sottospazi, che abbiamo introdotto nell'Esempio 4.13.

Proposizione 4.68. *Sia V uno spazio vettoriale su campo $\mathbb{K}$, di dimensione n e con base $\mathcal{B}$. Inoltre, sia U un sottospazio di V di dimensione m e base $\mathcal{B}_U = \{\mathbf{u_1}, \dots, \mathbf{u_m}\}$. Chiamiamo*

$$A = \left[\ \mathbf{u_1}|_{\mathcal{B}}\ \middle|\ \dots\ \middle|\ \mathbf{u_m}|_{\mathcal{B}}\ \right] \in \mathrm{Mat}(n,m;\mathbb{K})$$

ed

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

la colonna delle componenti del generico vettore $\mathbf{v} \in V$ rispetto a $\mathcal{B}$. Allora:

i) esiste la decomposizione

$$X = \sum_{i=1}^{m} t_i \cdot A_{C(i)}$$

se e solo se $\mathbf{v} \in U$. In tal caso, questa viene detta rappresentazione parametrica del sottospazio U rispetto alla base $\mathcal{B}$;

ii) se riduciamo a scala la matrice A contenuta all'interno della matrice orlata $[A|X]$, otteniamo una matrice $[S|X']$ le cui ultime $n-m$ righe costituiscono una rappresentazione algebrica di U rispetto alla base $\mathcal{B}$.

Dimostrazione. Il primo punto dell'enunciato è immediato, in quanto per la Proposizione 4.64 le colonne di A sono una base di $\phi_{\mathcal{B}}(U)$ e quindi la decomposizione esiste se e solo se $\phi_{\mathcal{B}}(\mathbf{v}) = X$ appartiene a $\phi_{\mathcal{B}}(U)$, da cui la tesi.

Per il secondo punto osserviamo che la matrice A ha rango m, di conseguenza riducendola a scala si ottiene la matrice S avente $S_{R(m+1)} = \dots = S_{R(n)} = 0_{1m}$. Eseguendo le stesse operazioni elementari necessarie a compiere tale riduzione anche sulla colonna X, le risultanti righe $X'_{R(m+1)}, \dots, X'_{R(n)}$ saranno delle combinazioni lineari non nulle di $x_1, \dots, x_n$. Di conseguenza il sistema $[A|X]$ ha soluzione, od equivalentemente $\mathbf{v}$ appartiene ad U, se e solo se $X'_{R(m+1)} = \cdots = X'_{R(n)} = 0$. Queste ultime condizioni su $x_1, \ldots, x_n$ costituiscono un sistema di $n-m$ equazioni lineari in n incognite, che per ipotesi ha infinite soluzioni dipendenti da m parametri. Pertanto, dal Teorema di Rouché-Capelli segue che le equazioni $X'_{R(m+1)} = \cdots = X'_{R(n)} = 0$ sono indipendenti e definiscono quindi una rappresentazione algebrica di U. □

Esempio 4.69. Sia $V = \mathbb{A}(3,\mathbb{Q})$ lo spazio delle matrici antisimmetriche di ordine 3, la cui base canonica $\mathcal{B}_{\mathbb{A}(3,\mathbb{Q})}$ è costituita dalle matrici

$$A_{12} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad A_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \qquad A_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}.$$

Consideriamo il sottospazio U di V generato da

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

Poiché le componenti di B rispetto alla base canonica sono

$$B|_{\mathcal{B}_{\mathbb{A}(3,\mathbb{Q})}} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

una rappresentazione parametrica di U è

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Osserviamo che la matrice corrispondente alle componenti x, y, z è

$$\phi^{-1}_{\mathcal{B}_{\mathbb{A}(3,\mathbb{Q})}}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = x \cdot A_{12} + y \cdot A_{13} + z \cdot A_{23} = \begin{bmatrix} 0 & x & y \\ -x & 0 & z \\ -y & -z & 0 \end{bmatrix} = \begin{bmatrix} 0 & t & t \\ -t & 0 & t \\ -t & -t & 0 \end{bmatrix}.$$

Per trovare una rappresentazione algebrica di U, riduciamo la matrice

$$\left[\begin{array}{c|c} 1 & x \\ 1 & y \\ 1 & z \end{array}\right] \xrightarrow[R_3 - R_1 \to R_3]{R_2 - R_1 \to R_2} \left[\begin{array}{c|c} 1 & x \\ 0 & -x+y \\ 0 & -x+z \end{array}\right].$$

Quindi U è l'insieme delle matrici antisimmetriche le cui componenti rispetto a $\mathcal{B}_{\mathbb{A}(3,\mathbb{Q})}$ soddisfano

$$\begin{cases} -x + y = 0 \\ -x + z = 0 \end{cases} \qquad \Rightarrow \qquad x = y = z\ .$$

Grazie alla Proposizione 4.68, abbiamo la seguente caratterizzazione dei sottospazi di uno spazio vettoriale.

Corollario 4.70. *Sia V uno spazio vettoriale di dimensione n su campo $\mathbb{K}$, con base $\mathcal{B}$. Allora il sottoinsieme U di V è un sottospazio di dimensione m se e solo se esiste una matrice $M \in \mathrm{Mat}(n-m, n; \mathbb{K})$ di rango massimo tale che $\phi_{\mathcal{B}}(U) = \ker(M)$. In tal caso, $\phi_{\mathcal{B}}$ è un isomorfismo tra U e $\ker(M)$.*

CAPITOLO 5

Applicazioni lineari

5.1. Definizione, nucleo ed immagine

Come abbiamo visto nel capitolo precedente, le applicazioni lineari sono le funzioni naturali tra due spazi vettoriali, che commutano con le operazioni ivi definite. In questa sezione inizieremo a studiare alcune proprietà generali di tali funzioni. Data la loro importanza, partiamo riproponendo la loro definizione.

Definizione 4.4. *Siano* $(V, \mathbb{K}, +_V, \cdot_V)$ *e* $(W, \mathbb{K}, +_W, \cdot_W)$ *due spazi vettoriali sul medesimo campo. Allora la funzione* $f : V \to W$ *è un omomorfismo di spazi vettoriali, o applicazione lineare, se per ogni* $\mathbf{v}, \tilde{\mathbf{v}} \in V$ *e* $t \in \mathbb{K}$, *vale*

i) $f(\mathbf{v} +_V \tilde{\mathbf{v}}) = f(\mathbf{v}) +_W f(\tilde{\mathbf{v}})$;

ii) $f(t \cdot_V \mathbf{v}) = t \cdot_W f(\mathbf{v})$.

L'insieme delle applicazioni lineari da V *in* W *viene indicato con* $\mathrm{Hom}(V, W)$.

Esempio 5.1. Alcuni esempi fondamentali di applicazioni lineari sono:

- la mappa delle componenti, che è un'applicazione lineare da un qualsiasi spazio V di dimensione n allo spazio delle matrici $\mathrm{Mat}(n, 1; \mathbb{K})$;
- la traccia di una matrice da $\mathrm{Mat}(n, n; \mathbb{K})$ a $\mathbb{K}$;
- l'applicazione naturale associata alla matrice $A \in \mathrm{Mat}(m, n; \mathbb{K})$, definita come

$$\begin{array}{rccc} f_A : & \mathrm{Mat}(n, 1; \mathbb{K}) & \longrightarrow & \mathrm{Mat}(m, 1; \mathbb{K}) \\ & X & \longmapsto & A * X \end{array}.$$

La linearità di f_A è conseguenza immediata delle proprietà del prodotto tra matrici;

- in $\mathbb{K}[x]$ siano

$$P(x) = \sum_{i=0}^{n} a_i\, x^i, \qquad P'(x) = \sum_{i=0}^{n-1} (i+1)\, a_{i+1}\, x^i \qquad \text{e} \qquad Q(x) = \sum_{i=1}^{n+1} \frac{a_{i-1}}{i}\, x^i.$$

La derivata algebrica di un polinomio è la funzione

$$\begin{array}{rccc} \dfrac{d}{dx} : & \mathbb{K}[x] & \longrightarrow & \mathbb{K}[x] \\ & P(x) & \longmapsto & P'(x) \end{array},$$

mentre, per ogni $y_0 \in \mathbb{K}$ fissato, l'integrale definito algebrico è la funzione

$$\begin{array}{rccc} \displaystyle\int_{y_0}^{y} dx : & \mathbb{K}[x] & \longrightarrow & \mathbb{K}[y] \\ & P(x) & \longmapsto & Q(y) - Q(y_0) \end{array}.$$

Osserviamo che nella loro definizioni non abbiamo utilizzato il concetto di limite, che è strettamente attinente all'Analisi ed assume senso solo se $\mathbb{K} = \mathbb{R}$. Però ovviamente le definizioni di derivata ed integrale algebrici coincidono con quelle analitiche nel caso di polinomi reali. Verifichiamo la linearità della derivata e dell'integrale. Dati i polinomi

$$P_1(x) = \sum_{i=0}^{n} a_i\, x^i \qquad \text{e} \qquad P_2(x) = \sum_{i=0}^{n} b_i\, x^i,$$

consideriamo la combinazione lineare

$$t_1 \cdot P_1(x) + t_2 \cdot P_2(x) = \sum_{i=0}^{n} (t_1 a_i + t_2 b_i)\, x^i.$$

Allora

$$\begin{aligned} \frac{d}{dx}(t_1 \cdot P_1(x) + t_2 \cdot P_2(x)) &= \sum_{i=0}^{n-1} (i+1)(t_1 a_{i+1} + t_2 b_{i+1})\, x^i \\ &= t_1 \sum_{i=0}^{n-1} (i+1) a_{i+1}\, x^i + t_2 \sum_{i=0}^{n-1} (i+1) b_{i+1}\, x^i \\ &= t_1 \cdot \frac{d}{dx} P_1(x) + t_2 \cdot \frac{d}{dx} P_2(x)\,. \end{aligned}$$

Analogamente,

$$\begin{aligned} \int_{y_0}^{y} (t_1 \cdot P_1(x) + t_2 \cdot P_2(x)) &= \sum_{i=1}^{n+1} \frac{t_1 a_{i-1} + t_2 b_{i-1}}{i}\, y^i - \sum_{i=1}^{n+1} \frac{t_1 a_{i-1} + t_2 b_{i-1}}{i}\, y_0^i \\ &= t_1 \left(\sum_{i=1}^{n+1} \frac{a_{i-1}}{i}\, y^i - \sum_{i=1}^{n+1} \frac{a_{i-1}}{i}\, y_0^i \right) + t_2 \left(\sum_{i=1}^{n+1} \frac{b_{i-1}}{i}\, y^i - \sum_{i=1}^{n+1} \frac{b_{i-1}}{i}\, y_0^i \right) \\ &= t_1 \cdot \int_{y_0}^{y} P_1(x) + t_2 \cdot \int_{y_0}^{y} P_2(x)\,. \end{aligned}$$

- Siano V uno spazio vettoriale ed U un sottospazio. Allora la mappa di inclusione è la funzione

$$i_{U;V}: \begin{array}{ccc} U & \longrightarrow & V \\ \mathbf{u} & \longmapsto & \mathbf{u} \end{array}.$$

  Lasciamo al vettore la verifica della linearità di $i_{U;V}$. Osserviamo che se $U = V$, allora $i_{V;V} = \mathrm{Id}_V$.

- Siano V uno spazio vettoriale ed U, W due sottospazi tali che $V = U \oplus W$. Per definizione, per ogni vettore $\mathbf{v} \in V$ esiste un'unica coppia $(\mathbf{u}, \mathbf{w}) \in U \times W$ tale che $\mathbf{v} = \mathbf{u} + \mathbf{w}$. Di conseguenza è naturalmente definita la funzione

$$\mathcal{P}_{U;W}: \begin{array}{ccc} V & \longrightarrow & U \\ \mathbf{v} & \longmapsto & \mathbf{u} \end{array},$$

  detta proiezione su U parallela a W. Questa funzione è un'applicazione lineare. Infatti, supponiamo che i due vettori $\mathbf{v_1}, \mathbf{v_2} \in V$ abbiano decomposizioni $\mathbf{v_1} = \mathbf{u_1} + \mathbf{w_1}$ e $\mathbf{v_2} = \mathbf{u_2} + \mathbf{w_2}$. Allora

$$\begin{aligned} \mathcal{P}_{U;W}(t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2}) &= \mathcal{P}_{U;W}(t_1 \cdot (\mathbf{u_1} + \mathbf{w_1}) + t_2 \cdot (\mathbf{u_2} + \mathbf{w_2})) \\ &= \mathcal{P}_{U;W}((t_1 \cdot \mathbf{u_1} + t_2 \cdot \mathbf{u_2}) + (t_1 \cdot \mathbf{w_1} + t_2 \cdot \mathbf{w_2})) \\ &= t_1 \cdot \mathbf{u_1} + t_2 \cdot \mathbf{u_2} \\ &= t_1 \cdot \mathcal{P}_{U;W}(\mathbf{v_1}) + t_2 \cdot \mathcal{P}_{U;W}(\mathbf{v_2}). \end{aligned}$$

- Siano ancora $V = U \oplus W$ e $\mathbf{v} = \mathbf{u} + \mathbf{w}$. La riflessione di $\mathbf{v}$ rispetto ad U e parallela a W è la funzione

$$\mathcal{R}_{U;W}: \begin{array}{ccc} V & \longrightarrow & V \\ \mathbf{v} & \longmapsto & \mathbf{u} - \mathbf{w} \end{array},$$

  La dimostrazione della linearità di $\mathcal{R}_{U;W}$ è del tutto analoga a quella della proiezione.

- Concludiamo con un esempio di applicazione non lineare tra due spazi vettoriali. La funzione $f(x) = x^2$ da $\mathbb{K}$ in $\mathbb{K}$ non è lineare. Infatti

$$f(t_1 \cdot x_1 + t_2 \cdot x_2) = (t_1x_1 + t_2x_2)^2 = t_1^2x_1^2 + 2t_1t_2x_1x_2 + t_2^2x_2^2 \neq t_1x_1^2 + t_2x_2^2 = t_1 \cdot f(x_1) + t_2 \cdot f(x_2).$$

Ad ogni applicazione lineare sono associati due insiemi.

Definizione 5.2. *Siano V, W spazi vettoriali su campo $\mathbb{K}$ ed $f \in \mathrm{Hom}(V, W)$. Allora:*

i) il nucleo dell'applicazione è l'insieme

$$\ker(f) = \{\mathbf{v} \in V \mid f(\mathbf{v}) = \mathbf{0}\} \subseteq V;$$

ii) l'immagine dell'applicazione è l'insieme

$$\mathrm{Im}(f) = \{\mathbf{w} \in W \mid \mathbf{w} = f(\mathbf{v}),\ \textit{per qualche } \mathbf{v} \in V\} \subseteq W.$$

Questi insiemi contengono informazioni importanti riguardanti l'applicazione f. Iniziamo lo studio delle loro proprietà.

Lemma 5.3. *Per ogni applicazione lineare $f \in \mathrm{Hom}(V, W)$ vale $f(\mathbf{0_V}) = \mathbf{0_W}$.*

Dimostrazione. $f(\mathbf{0_V}) = f(\mathbf{v} - \mathbf{v}) = f(\mathbf{v}) - f(\mathbf{v}) = \mathbf{0_W}$. □

Proposizione 5.4. *Siano V, W spazi vettoriali su campo $\mathbb{K}$ ed $f \in \mathrm{Hom}(V, W)$. Allora:*

i) $\ker(f)$ *è un sottospazio di* V;

ii) $\mathrm{Im}(f)$ *è un sottospazio di* W.

Dimostrazione. Le verifiche delle due proprietà di sottospazio sono simili ed entrambe sfruttano la Proposizione 4.11 ed il Lemma 4.14:

i) siano $\mathbf{v_1}, \mathbf{v_2} \in \ker(f)$ e $t_1, t_2 \in \mathbb{K}$. Allora

$$f(t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2}) = t_1 \cdot f(\mathbf{v_1}) + t_2 \cdot f(\mathbf{v_2}) = t_1 \cdot \mathbf{0_W} + t_2 \cdot \mathbf{0_W} = \mathbf{0_W}$$

e di conseguenza $t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2} \in \ker(f)$. Inoltre, per il Lemma 5.3 sappiamo che $\mathbf{0_V} \in \ker(f)$;

ii) siano $\mathbf{w_1}, \mathbf{w_2} \in \mathrm{Im}(f)$ e $t_1, t_2 \in \mathbb{K}$. Allora $\mathbf{w_1} = f(\mathbf{v_1})$ e $\mathbf{w_2} = f(\mathbf{v_2})$, e di conseguenza

$$t_1 \cdot \mathbf{w_1} + t_2 \cdot \mathbf{w_2} = t_1 \cdot f(\mathbf{v_1}) + t_2 \cdot f(\mathbf{v_2}) = f(t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2}).$$

Pertanto $t_1 \cdot \mathbf{w_1} + t_2 \cdot \mathbf{w_2} \in \mathrm{Im}(f)$. Inoltre, sempre per il Lemma 5.3 sappiamo che $\mathbf{0_W} = f(\mathbf{0_V}) \in \mathrm{Im}(f)$.

□

Esempio 5.5. Studiamo nucleo ed immagine di ogni applicazione lineare dell'Esempio 5.1.

- La mappa delle componenti fissa una corrispondenza biunivoca tra vettori e matrici colonne di lunghezza pari alla dimensione dello spazio vettoriale. Questo ha come conseguenza $\ker(\phi_{\mathcal{B}}) = \{\mathbf{0_V}\}$ e $\mathrm{Im}(\phi_{\mathcal{B}}) = \mathrm{Mat}(n, 1; \mathbb{K})$.
- La traccia è chiaramente un'applicazione suriettiva, quindi $\mathrm{Im}(\mathrm{Tr}) = \mathbb{K}$. La richiesta di traccia nulla impone invece un'equazione lineare omogenea sulle entrate di una matrice:

$$\sum_{i_1}^{n} a_{ii} = 0.$$

Questa equazione definisce un sottospazio di $\mathrm{Mat}(n, n; \mathbb{K})$ di dimensione $n^2 - 1$.

- Data $A \in \text{Mat}(m, n; \mathbb{K})$, il nucleo dell'applicazione f_A è l'insieme delle soluzioni di $A * X = 0_{m1}$, ovvero $\ker(f_A) = \ker(A)$ (più correttamente, ma anche più pedantemente, i due nuclei sono in realtà solo isomorfi tra loro, attraverso l'isomorfismo (1) definito nell'Osservazione 3.3). L'immagine di f_A è invece l'insieme di tutte le matrici $B \in \text{Mat}(m, 1; \mathbb{K})$ che possono essere scritte come
$$B = A * \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix} = \sum_{i=1}^{n} t_i \cdot A_{C(i)},$$
per ogni $t_1, \ldots, t_n \in \mathbb{K}$. Questo significa che $\text{Im}(f_A) = C(A)$, lo spazio delle colonne di A.
- La derivata algebrica è suriettiva, mentre il suo nucleo è costituito dai polinomi di grado zero. Per ogni y_0 fissato, l'integrale algebrico ha come immagine un sottospazio di $\mathbb{K}[x]$ isomorfo allo spazio generato dai polinomi di grado maggiore di zero, mentre il suo nucleo è il solo polinomio nullo.
- Il nucleo della mappa d'inclusione è il sottospazio banale $\{\mathbf{0}\}$, mentre la sua immagine è U.
- Il nucleo della proiezione $\mathcal{P}_{U;W}$ è il sottospazio W. Infatti, i vettori $\mathbf{v} \in \ker(\mathcal{P}_{U;W})$ sono tutti ed i soli vettori aventi come decomposizione una coppia del tipo $(\mathbf{0}, \mathbf{w}) \in U \times W$, da cui $\mathbf{v} = \mathbf{0} + \mathbf{w} = \mathbf{w} \in W$. Invece, l'immagine di $\mathcal{P}_{U;W}$ è U. Infatti, ogni vettore $\mathbf{u} \in U$ è immagine di $\mathbf{v} = \mathbf{u} + \mathbf{0} \in V$, da cui $U \subseteq \text{Im}(\mathcal{P}_{U;W})$. Ma essendo U il codominio della proiezione, ne segue l'uguaglianza dei due insiemi e quindi anche la suriettività dell'applicazione.
- Chiaramente la riflessione $\mathcal{R}_{U;W}$ è un isomorfismo, quindi $\ker(\mathcal{R}_{U;W}) = \{\mathbf{0}\}$ ed $\text{Im}(\mathcal{R}_{U;W}) = V$.

Come abbiamo visto nel terzo esempio, esiste uno stretto legame tra i sistemi lineari associati alla matrice A e l'applicazione lineare f_A. In particolare X è soluzione del sistema associato ad $[A|B]$ se e solo se $f_A(X) = B$, cioè se e solo se $X \in f_A^{-1}(B)$. Questo fatto sta alla base della forte analogia tra il Teorema 3.28 di struttura delle soluzioni di un sistema lineare ed il seguente teorema di struttura delle controimmagini di un'applicazione lineare, essendo essenzialmente il primo un caso particolare del secondo.

Teorema 5.6. *Siano V, W spazi vettoriali su campo $\mathbb{K}$ ed $f \in \text{Hom}(V, W)$. Dato $\mathbf{w} \in \text{Im}(f)$ e $\mathbf{v_p} \in V$ una sua controimmagine particolare, l'insieme delle controimmagini di $\mathbf{w}$ è*
$$f^{-1}(\mathbf{w}) = \{\mathbf{v} \in V \mid \mathbf{v} = \mathbf{v_p} + \mathbf{v_0}\,,\ \text{dove } \mathbf{v_0} \in \ker(f)\}.$$

Dimostrazione. Il vettore $\mathbf{v}$ appartiene a $f^{-1}(\mathbf{w})$ se e solo se $f(\mathbf{v} - \mathbf{v_p}) = f(\mathbf{v}) - f(\mathbf{v_p}) = \mathbf{w} - \mathbf{w} = \mathbf{0_W}$, ovvero se e solo se $\mathbf{v_0} = \mathbf{v} - \mathbf{v_p} \in \ker(f)$. □

Corollario 5.7. *L'applicazione $f \in \text{Hom}(V, W)$ è iniettiva se e solo se $\ker(f) = \{\mathbf{0_V}\}$.*

Dimostrazione. Se f è iniettiva, allora la controimmagine di $\mathbf{0_W}$ è un solo vettore e quindi necessariamente $\mathbf{0_V}$. Viceversa, se $\ker(f) = \{\mathbf{0_V}\}$ il teorema di struttura afferma che $f^{-1}(\mathbf{w})$ è costituito da un solo vettore per ogni $\mathbf{w} \in W$ e quindi f è iniettiva. □

Esempio 5.8. Studiamo nucleo ed immagine dell'applicazione $f_h \in \text{Hom}(\text{Mat}(2, 1; \mathbb{K}), \text{Mat}(2, 1; \mathbb{K}))$ associata ad
$$A_h = \begin{bmatrix} h & 1 \\ 0 & 1 \end{bmatrix} \in \text{Mat}(2, 2; \mathbb{K}),$$
dove $h \in \mathbb{K}$. Se $h \neq 0$ l'applicazione è iniettiva e suriettiva. Infatti $r(A_h) = 2$ e quindi $\dim(\ker(f_h)) = \dim(\ker(A_h)) = n - r(A_h) = 0$, mentre $\dim(\text{Im}(f_h)) = \dim(C(A_h)) = r(A_h) = 2$. Dalla prima equazione si deduce $\ker(f_h) = \{0_{21}\}$ e dalla seconda $\text{Im}(f_h) = \text{Mat}(2, 1; \mathbb{K})$.

Se $h = 0$ abbiamo $r(A_h) = 1$, quindi $\dim(\ker(f_h)) = 2 - 1 = 1$ e $\dim(\mathrm{Im}(f_h)) = 1$. In tal caso le basi dei due sottospazi sono rispettivamente

$$\mathcal{B}_{\ker(f_h)} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\} \qquad \text{e} \qquad \mathcal{B}_{\mathrm{Im}(f_h)} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}.$$

La controimmagine di un vettore si ottiene risolvendo il sistema lineare associato. Quindi, se $h \neq 0$ abbiamo

$$f^{-1}\left(\begin{bmatrix} a \\ b \end{bmatrix} \right) = \left\{ \begin{bmatrix} \frac{a-b}{h} \\ b \end{bmatrix} \right\}.$$

Se $h = 0$ la controimmagine è l'insieme vuoto se $a \neq b$, mentre

$$f^{-1}\left(\begin{bmatrix} a \\ a \end{bmatrix} \right) = \left\{ \begin{bmatrix} 0 \\ a \end{bmatrix} + t \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ t \in \mathbb{K} \right\}.$$

5.2. Algebra delle applicazioni lineari

Lo scopo di questa sezione è indagare le operazioni naturali che possono essere definite sull'insieme delle applicazioni lineari. Nell'Esempio 4.2 contenuto all'inizio del Capitolo 4 abbiamo accennato al fatto che l'insieme delle funzioni W^A definite sull'insieme A a valori nello spazio vettoriale W è a sua volta uno spazio vettoriale, rispetto alle operazioni di somma e prodotto elemento per elemento. Ora studieremo in dettaglio questa struttura nel caso degli omomorfismi tra due spazi vettoriali, in cui cioè anche l'insieme di partenza A è uno spazio vettoriale V.

Proposizione 5.9. *Siano V e W due spazi vettoriali sul medesimo campo $\mathbb{K}$. Nell'insieme* $\mathrm{Hom}(V, W)$ *sono definite le operazioni di somma e prodotto elemento per elemento:*

$$\begin{array}{rccc} + : & \mathrm{Hom}(V,W) \times \mathrm{Hom}(V,W) & \longrightarrow & \mathrm{Hom}(V,W) \\ & (f_1, f_2) & \longmapsto & f_1 + f_2 \end{array},$$

dove $(f_1 + f_2)(\mathbf{v}) = f_1(\mathbf{v}) + f_2(\mathbf{v})$ per ogni $\mathbf{v} \in V$;

$$\begin{array}{rccc} \cdot : & \mathbb{K} \times \mathrm{Hom}(V,W) & \longrightarrow & \mathrm{Hom}(V,W) \\ & (t\,,\, f) & \longmapsto & t \cdot f \end{array},$$

dove $(t \cdot f)(\mathbf{v}) = t\, f(\mathbf{v})$ per ogni $\mathbf{v} \in V$. Allora $(\mathrm{Hom}(V,W), \mathbb{K}, +, \cdot)$ *è uno spazio vettoriale.*

Dimostrazione. Per prima cosa verifichiamo che le due operazioni siano ben definite, cioè che il loro risultato sia ancora un'applicazione lineare. Dati $\mathbf{v_1}, \mathbf{v_2} \in V$ e $t_1, t_2 \in \mathbb{K}$ abbiamo

$$\begin{aligned} (f_1 + f_2)(t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2}) &= f_1(t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2}) + f_2(t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2}) = \\ &= (t_1 \cdot f_1(\mathbf{v_1}) + t_2 \cdot f_1(\mathbf{v_2})) + (t_1 \cdot f_2(\mathbf{v_1}) + t_2 \cdot f_2(\mathbf{v_2})) = \\ &= t_1 \cdot (f_1(\mathbf{v_1}) + f_2(\mathbf{v_1})) + t_2 \cdot (f_1(\mathbf{v_2} + f_2(\mathbf{v_2})) = \\ &= t_1 \cdot (f_1 + f_2)(\mathbf{v_1}) + t_2 \cdot (f_1 + f_2)(\mathbf{v_2}), \end{aligned}$$

per ogni $(f_1, f_2) \in \mathrm{Hom}(V, W) \times \mathrm{Hom}(V, W)$. Analogamente,

$$\begin{aligned} (t \cdot f)(t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2}) &= t \cdot f(t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2}) = \\ &= t \cdot (t_1 \cdot f(\mathbf{v_1}) + t_2 \cdot f(\mathbf{v_2})) = \\ &= t_1 \cdot (t \cdot f(\mathbf{v_1})) + t_2 \cdot (t \cdot f(\mathbf{v_2})), \end{aligned}$$

per ogni $(t, f) \in \mathbb{K} \times \mathrm{Hom}(V, W)$.

Le proprietà richieste per la somma ed il prodotto di applicazioni lineari seguono facilmente dalla loro definizione elemento per elemento. Ad esempio, verifichiamo l'associatività della somma:

$$\begin{aligned}((f_1+f_2)+f_3)(\mathbf{v}) &= (f_1+f_2)(\mathbf{v})+f_3(\mathbf{v}) = (f_1(\mathbf{v})+f_2(\mathbf{v}))+f_3(\mathbf{v}) = \\ &= f_1(\mathbf{v})+(f_2(\mathbf{v})+f_3(\mathbf{v})) = f_1(\mathbf{v})+(f_2+f_3)(\mathbf{v}) = (f_1+(f_2+f_3))(\mathbf{v}).\end{aligned}$$

Nella terza uguaglianza abbiamo utilizzato l'associatività della somma in W. Le altre proprietà si dimostrano in modo simile. Osserviamo in particolare che l'elemento neutro della somma è l'applicazione nulla

$$\begin{array}{rccc} 0_{VW}: & V & \longrightarrow & W \\ & \mathbf{v} & \longmapsto & \mathbf{0_W} \end{array},$$

mentre l'inversa additiva di $f \in \mathrm{Hom}(V,W)$ è

$$\begin{array}{rccc} -f: & V & \longrightarrow & W \\ & \mathbf{v} & \longmapsto & -f(\mathbf{v}) \end{array}. \qquad \square$$

Tra funzioni è definita anche una terza operazione, la composizione $\circ$. Andiamo ad investigare il comportamento di $\circ$ rispetto alla struttura di $\mathrm{Hom}(V,W)$.

Proposizione 5.10. *Siano U,V,W,Z spazi vettoriali ed $f \in \mathrm{Hom}(U,V)$, $g \in \mathrm{Hom}(V,W)$. Allora:*

$$\begin{array}{rccc} \circ: & \mathrm{Hom}(U,V)\times\mathrm{Hom}(V,W) & \longrightarrow & \mathrm{Hom}(U,W) \\ & (f,g) & \longmapsto & g\circ f \end{array}.$$

Inoltre, la composizione di applicazioni lineari:

i) è associativa, cioè $(h\circ g)\circ f = h\circ(g\circ f)$ per ogni $f \in \mathrm{Hom}(U,V)$, $g \in \mathrm{Hom}(V,W)$ ed $h \in \mathrm{Hom}(W,Z)$;

ii) è distributiva rispetto alla somma di funzioni, cioè $h\circ(f+g) = (h\circ f)+(h\circ g)$ per ogni $f,g \in \mathrm{Hom}(U,V)$ ed $h \in \mathrm{Hom}(V,W)$, ed analogamente $(g+h)\circ f = (g\circ f)+(h\circ f)$ per ogni $f \in \mathrm{Hom}(U,V)$ ed $g,h \in \mathrm{Hom}(V,W)$;

iii) è omogenea rispetto al prodotto per uno scalare, cioè $t\cdot(g\circ f) = (t\cdot g)\circ f = g\circ(t\cdot f)$ per ogni $t \in \mathbb{K}$, $f \in \mathrm{Hom}(U,V)$ e $g \in \mathrm{Hom}(V,W)$;

iv) ha Id_V come elemento neutro, nel senso che date $f \in \mathrm{Hom}(U,V)$ e $g \in \mathrm{Hom}(V,W)$, vale $\mathrm{Id}_V\circ f = f$ e $g\circ\mathrm{Id}_V = g$.

Dimostrazione. Per prima cosa è necessario verificare che la composizione di due applicazioni lineari è ancora un'applicazione lineare:

$$\begin{aligned}(g\circ f)(t_1\cdot\mathbf{v_1}+t_2\cdot\mathbf{v_2}) &= g(f(t_1\cdot\mathbf{v_1}+t_2\cdot\mathbf{v_2})) \\ &= g(t_1\cdot f(\mathbf{v_1})+t_2\cdot f(\mathbf{v_2})) \\ &= t_1\cdot g(f(\mathbf{v_1}))+t_2\cdot g(f(\mathbf{v_2})) \\ &= t_1\cdot(g\circ f)(\mathbf{v_1})+t_2\cdot(g\circ f)(\mathbf{v_2}).\end{aligned}$$

Dimostriamo ora le singole proprietà della composizione. Queste sono vere se e solo se valgono per ogni $\mathbf{u} \in U$:

i) $((h\circ g)\circ f)(\mathbf{u}) = (h\circ g)(f(\mathbf{u})) = h(g(f(\mathbf{u}))) = h((g\circ f)(\mathbf{u})) = (h\circ(g\circ f))(\mathbf{u})$;

ii) verifichiamo la prima proprietà, in quanto la seconda è analoga:

$$(h\circ(f+g))(\mathbf{u}) = h((f+g)(\mathbf{u})) = h(f(\mathbf{u})+g(\mathbf{u})) = h(f(\mathbf{u}))+h(g(\mathbf{u})) =$$

$$= (h\circ f)(\mathbf{u})+(h\circ g)(\mathbf{u}) = ((h\circ f)+(h\circ g))(\mathbf{u});$$

iii) come per il punto precedente, verifichiamo solo la prima proprietà:

$$(t \cdot (g \circ f))(\mathbf{u}) = t \cdot (g \circ f)(\mathbf{u}) = t \cdot g(f(\mathbf{u})) = (t \cdot g)(f(\mathbf{u})) = ((t \cdot g) \circ f)(\mathbf{u});$$

iv) $(\mathrm{Id}_V \circ f)(\mathbf{u}) = \mathrm{Id}_V(f(\mathbf{u})) = f(\mathbf{u})$ ed analogamente l'altra identità. $\square$

In generale, la composizione di funzioni non soddisfa la proprietà commutativa, di annullamento del prodotto e di esistenza dell'inverso.

Esempio 5.11. Consideriamo gli spazi vettoriali delle matrici colonna $U = \mathrm{Mat(n, 1; \mathbb{K})}$, $V = \mathrm{Mat(m, 1; \mathbb{K})}$ e $W = \mathrm{Mat(p, 1; \mathbb{K})}$. Le matrici $A, B \in \mathrm{Mat(m, n; \mathbb{K})}$ e $C \in \mathrm{Mat(n, p; \mathbb{K})}$ definiscono rispettivamente le applicazioni lineari $f_A, f_B \in \mathrm{Hom}(U, V)$ ed $f_C \in \mathrm{Hom}(V, W)$. Allora per ogni $X \in U$ e $t \in \mathbb{K}$ valgono:

- $(f_A + f_B)(X) = f_A(X) + f_B(X) = (A * X) + (B * X) = (A + B) * X = f_{A+B}(X)$;
- $(t \cdot f_A)(X) = t \cdot f_A(X) = t \cdot (A * X) = (t \cdot A) * X = f_{t \cdot A}(X)$;
- l'applicazione nulla è associata alla matrice zero, cioè $0_{UV} = f_{0_{mn}}$;
- l'inversa additiva di f_A è ovviamente f_{-A}.

Queste osservazioni implicano che la funzione

$$\begin{array}{rccc} \Upsilon : & \mathrm{Mat}(m, n) & \longrightarrow & \mathrm{Hom}(U, V) \\ & A & \longmapsto & f_A \end{array}$$

è un omomorfismo di spazi vettoriali. Inoltre valgono:

- $(f_C \circ f_A)(X) = f_C(f_A(X)) = f_C(A * X) = C * (A * X) = (C * A) * X = f_{C*A}(X)$;
- l'applicazione identica è $\mathrm{Id_U} = f_{\mathbb{I}_n}$.

Ciò significa che Υ rispetta anche le strutture moltiplicative date rispettivamente dal prodotto righe per colonne nello spazio delle matrici e dalla composizione nello spazio delle applicazioni lineari. Da questa proprietà deduciamo che la composizione di due applicazioni lineari f_A ed f_B, qualora esista, gode della proprietà commutativa se e solo se commutano le due matrici A e B. Inoltre, f_A è invertibile se e solo se la matrice A è invertibile, ed in tal caso ovviamente vale $(f_A)^{-1} = f_{A^{-1}}$. Nel linguaggio dell'Algebra Astratta, si dice che Υ è un omomorfismo tra l'algebra delle matrici e l'algebra delle applicazioni lineari.

Esempio 5.12. Consideriamo la derivata e l'integrale algebrici. Partendo dalla loro definizione nell'Esempio 5.1, non è difficile constatare che

$$\frac{d}{dx} \circ \int_{x_0}^{x} dx = \mathrm{Id}_{\mathbb{K}[x]}.$$

Al contrario

$$\int_{x_0}^{x} dx \circ \frac{d}{dx} \neq \mathrm{Id}_{\mathbb{K}[x]},$$

in quanto nell'operazione di derivazione si perde il termine costante di un polinomio, che non può più essere recuperato attraverso l'integrazione. Quindi la derivata è l'inversa sinistra dell'integrale, ma non l'inversa destra. Questo fenomeno può accadere esclusivamente in spazi di dimensione infinita, come appunto ad esempio $\mathbb{K}[x]$.

5.3. Isomorfismi e teorema di isomorfismo

Per definizione, gli isomorfismi di spazi vettoriali sono applicazioni lineari per le quali esiste l'applicazione inversa, la quale a sua volta è lineare. Dalla Proposizione 4.6 sappiamo che $f \in \mathrm{Hom}(U, V)$ è un isomorfismo se e solo se è invertibile. Utilizzando il linguaggio sviluppato nella sezione precedente, un'applicazione lineare è un isomorfismo se esiste la sua inversa rispetto all'operazione di composizione.

È semplice verificare che l'insieme degli isomorfismi tra due spazi vettoriali U e V non è chiuso rispetto al prodotto per uno scalare e rispetto alla somma di funzioni. Basta infatti osservare che $0 \cdot f = 0_{UV}$ ed $f + (-f) = 0_{UV}$, ed ovviamente 0_{UV} non è un'applicazione invertibile. Questo implica che l'insieme degli isomorfismi non è un sottospazio vettoriale di $\mathrm{Hom}(U, V)$. Il comportamento rispetto alla composizione è invece migliore. Infatti se $f \in \mathrm{Hom}(U, V)$ e $g \in \mathrm{Hom}(V, W)$ sono isomorfismi, allora $g \circ f : U \to W$ è un'applicazione lineare biunivoca, in quanto composizione di applicazioni biunivoche. Quindi $g \circ f$ è ancora un isomorfismo.

Attraverso il concetto di isomorfismo possiamo definire una relazione binaria sull'insieme di tutti gli spazi vettoriali a supporto sul medesimo campo $\mathbb{K}$. Diciamo che la coppia ordinata di spazi vettoriali (V, W) è in relazione se esiste un isomorfismo $f : V \to W$.

Proposizione 5.13. *La relazione di isomorfismo è di equivalenza, pertanto viene indicata con il simbolo* $\sim$.

Dimostrazione. Ricordiamo che una relazione è di equivalenza se soddisfa le proprietà di riflessitività, simmetria e transitività. La prima proprietà è vera se ogni spazio vettoriale V è isomorfo a se stesso, cioè $V \sim V$. Ma questo è ovvio, in quanto $\mathrm{Id}_V : V \to V$ è un isomorfismo. Per la simmetria osserviamo invece che se $f : V \to W$ è un isomorfismo tra due dati spazi, allora esiste l'isomorfismo inverso $f^{-1} : W \to V$ e quindi $V \sim W$ implica $W \sim V$. Infine, consideriamo tre spazi U, V, W e due isomorfismi $f : U \to V$ e $g : V \to W$. Per quanto osservato precedentemente, anche $g \circ f$ è un isomorfismo. Questo significa che $U \sim V$ e $V \sim W$ implicano $U \sim W$. □

Una classe di equivalenza rispetto all'isomorfismo è costituita da tutti gli spazi vettoriali, definiti sul medesimo campo, che sono isomorfi tra loro.

Una importante caratterizzazione degli isomorfismi è data in termini di basi.

Proposizione 5.14. *Siano V, W spazi vettoriali sul medesimo campo $\mathbb{K}$. Allora l'applicazione $f \in \mathrm{Hom}(V, W)$ è:*

i) suriettiva se e solo se l'immagine di un insieme di generatori di V è un insieme di generatori di W;
ii) iniettiva se e solo se l'immagine di ogni insieme linearmente indipendente in V è linearmente indipendente in W;
iii) un isomorfismo se e solo se l'immagine di una base di V è una base di W.

Dimostrazione. Sia $U = \{\mathbf{u_1}, \dots, \mathbf{u_n}\}$ un sottoinsieme di V ed $f(U) = \{f(\mathbf{u_1}), \dots, f(\mathbf{u_n})\}$ la sua immagine.

i) Supponiamo che U sia un insieme di generatori, ovvero che sia possibile ottenere ogni $\mathbf{v} \in V$ come

$$\mathbf{v} = \sum_{i=1}^{n} t_i \cdot \mathbf{u_i}.$$

Per definizione, l'applicazione f è suriettiva se ogni $\mathbf{w} \in W$ è immagine di un vettore $\mathbf{v} \in V$, e quindi

$$\mathbf{w} = f(\mathbf{v}) = f\left(\sum_{i=1}^{n} t_i \cdot \mathbf{u_i}\right) = \sum_{i=1}^{n} t_i \cdot f(\mathbf{u_i}).$$

Ma questo è possibile se e solo se $f(U)$ è un insieme di generatori di W.

ii) Supponiamo che U sia linearmente indipendente. Consideriamo la combinazione lineare

$$\mathbf{0_W} = \sum_{i=1}^{n} t_i \cdot f(\mathbf{u_i}) = f\left(\sum_{i=1}^{n} t_i \cdot \mathbf{u_i}\right). \tag{10}$$

Se f è iniettiva, l'unica controimmagine di $\mathbf{0_W}$ è $\mathbf{0_V}$, quindi

$$\sum_{i=1}^{n} t_i \cdot \mathbf{u_i} = \mathbf{0_V}.$$

Dato che U è indipendente, allora $t_1 = \cdots = t_n = 0$ e quindi anche $f(U)$ è indipendente. Viceversa, sia $f(U)$ indipendente per ogni U indipendente. Allora, in particolare $f(U)$ è indipendente se U è composto da un unico vettore $\mathbf{u} \in V \setminus \{\mathbf{0_V}\}$. Di conseguenza deve valere $f(\mathbf{u}) \neq \mathbf{0_W}$. Questo implica che $\ker(f) = \{\mathbf{0_V}\}$ e quindi per il Corollario 5.7 segue che f è iniettiva.

iii) Se $f \in \mathrm{Hom}(V, W)$ è un isomorfismo, allora è un'applicazione iniettiva e suriettiva. In tal caso, se U è una base allora dai punti (i) e (ii) otteniamo che $f(U)$ è una base. Viceversa, supponiamo che data una base U di V, la sua immagine $f(U)$ sia una base di W. Poiché U ed $f(U)$ generano rispettivamente V e W, dal punto (i) segue che f è suriettiva. Infine, riprendiamo ancora la combinazione lineare 10. Se $f(U)$ è una base, l'unica soluzione è quella banale con $t_1 = \cdots = t_n = 0$, di conseguenza f è anche iniettiva e pertanto è un isomorfismo. □

Nel caso di spazi finitamente generati, la proposizione precedente ci permette di caratterizzare le classi di isomorfismo in modo molto semplice, attraverso il seguente importante risultato noto come Teorema di isomorfismo.

Teorema 5.15. *Due spazi vettoriali V e W finitamente generati sul medesimo campo $\mathbb{K}$ sono isomorfi se e solo se* $\dim(V) = \dim(W)$.

Dimostrazione. Il caso $V = \{\mathbf{0}\}$ è ovvio, poiché uno spazio vettoriale W in corrispondenza biunivoca con V deve contenere un solo vettore, e quindi $W = V = \{\mathbf{0}\}$. Supponiamo invece che V e W non siano $\{\mathbf{0}\}$. Se $f \in \mathrm{Hom}(V, W)$ è un isomorfismo, allora la Proposizione 5.14 implica che l'immagine di ogni base di V attraverso f è una base di W. Di conseguenza $\dim(V) = \dim(W)$. Viceversa, se $\dim(V) = \dim(W) = n$ allora le rispettive mappe delle componenti rispetto a due generiche basi dei due spazi definiscono due isomorfismi $\phi_{\mathcal{B}_V} : V \to \mathrm{Mat(n, 1; \mathbb{K})}$ e $\phi_{\mathcal{B}_W} : W \to \mathrm{Mat(n, 1; \mathbb{K})}$. Quindi $V \sim \mathrm{Mat(n, 1; \mathbb{K})}$ e $W \sim \mathrm{Mat(n, 1; \mathbb{K})}$, da cui per le proprietà di simmetria e transitività otteniamo $V \sim W$. □

Le classi di isomorfismo sono quindi gli insiemi di tutti gli spazi vettoriali definiti sul medesimo campo e di uguale dimensione. Di conseguenza esistono infinite classi di isomorfismo. Segnaliamo al lettore che il precedente teorema non è vero nel caso di spazi non finitamente generati. Più precisamente, uno spazio di dimensione finita ed uno di dimensione infinita ovviamente non sono isomorfi, ma esistono esempi di spazi vettoriali di dimensione infinita, sul medesimo campo, che non sono tra loro isomorfi.

Esempio 5.16. Siano $V = \mathbb{R}[x]_2$ e $W = \mathbb{S}(2; \mathbb{R})$. Dato che $\dim(V) = \dim(W) = 3$, i due spazi sono isomorfi. Costruiamo esplicitamente un isomorfismo. Prendiamo le basi canoniche dei due spazi:

$$\mathcal{B}_2 = \{1, x, x^2\} \qquad \text{e} \qquad \mathcal{B}_{\mathbb{S}(2;\mathbb{R})} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

Allora, dato $P = a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2$ abbiamo

$$\begin{aligned}\left(\phi^{-1}_{\mathcal{B}_{\mathbb{S}(2;\mathbb{R})}} \circ \phi_{\mathcal{B}_3}\right)(P) &= \phi^{-1}_{\mathcal{B}_{\mathbb{S}(2;\mathbb{R})}}\left(\phi_{\mathcal{B}_3}(P)\right) = \phi^{-1}_{\mathcal{B}_{\mathbb{S}(2;\mathbb{R})}}\left(\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}\right) = \\ &= a_0 \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a_1 \cdot \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix} + a_2 \cdot \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a_0 & \frac{a_1}{2} \\ \frac{a_1}{2} & a_2 \end{bmatrix}.\end{aligned}$$

In generale, fissate due basi $\mathcal{B}_V = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ e $\mathcal{B}_W = \{\mathbf{w_1}, \dots, \mathbf{w_n}\}$ di due spazi vettoriali V e W di uguale dimensione, un isomorfismo naturale si costruisce come

$$\begin{aligned}\phi^{-1}_{\mathcal{B}_W} \circ \phi_{\mathcal{B}_V} : \quad V &\longrightarrow W \\ \mathbf{v_i} &\longmapsto \mathbf{w_i}\end{aligned}$$

per ogni $1 \leq i \leq n$.

Osservazione 5.17. É fondamentale prestare attenzione alla richiesta che i due spazi vettoriali V e W siano definiti sullo stesso campo. Ad esempio $V = (\mathbb{C}, \mathbb{C}, +_{\mathbb{C}}, \cdot_{\mathbb{C}})$ e $W = (\mathbb{R}, \mathbb{R}, +_{\mathbb{R}}, \cdot_{\mathbb{R}})$ sono entrambi spazi vettoriali unidimensionali, i cui vettori sono rispettivamente i numeri complessi z ed i numeri reali x. Ma V e W non sono isomorfi, in quanto non esistono nemmeno applicazioni lineari da V in W. Per dimostrarlo, supponiamo per assurdo che f sia una tale applicazione, che associa al vettore $z \in V$ il vettore $x = f(z) \in W$. Ma se $t \in \mathbb{C}$ è un numero complesso con parte immaginaria non nulla, l'uguaglianza $f(t \cdot_{\mathbb{C}} z) = t \cdot_{\mathbb{R}} f(z)$ non ha senso, perché l'operazione a secondo membro non è definita.

Considerando invece $\tilde{V} = (\mathbb{C}, \mathbb{R}, +_{\mathbb{C}}, \cdot_{\mathbb{C}})$, abbiamo a che fare con uno spazio vettoriale reale bidimensionale. Infatti, ogni numero complesso z può essere scritto in modo unico come $z = x \cdot_{\mathbb{C}} 1 + y \cdot_{\mathbb{C}} i$ con $x, y \in \mathbb{R}$, da cui deduciamo che $\{1, i\}$ è una base di $\tilde{V}$. In questo senso, è corretto affermare che sul sostegno $\mathbb{C}^n$ è possibile definire sia la sua struttura naturale di spazio vettoriale complesso di dimensione n, sia quella di spazio vettoriale reale di dimensione $2n$. Quest'ultima è isomorfa alla struttura vettoriale naturale di $\mathbb{R}^{2n}$.

Concludiamo questa sezione dimostrando finalmente la Proposizione 4.64, che riproponiamo.

Proposizione 4.64. *Dati V e W spazi vettoriali sul medesimo campo $\mathbb{K}$, sia $f : V \to W$ un isomorfismo. Allora:*

i) $U \subseteq V$ *è un sottospazio se e solo se* $f(U) \subseteq W$ *è un sottospazio;*

ii) $\{\mathbf{u_1}, \dots, \mathbf{u_m}\}$ *è un insieme di generatori di* $U \subseteq V$ *se e solo se* $\{f(\mathbf{u_1}), \dots, f(\mathbf{u_m})\}$ *è un insieme di generatori di* $f(U) \subseteq W$;

iii) $\{\mathbf{u_1}, \dots, \mathbf{u_m}\}$ *è linearmente indipendente se e solo se* $\{f(\mathbf{u_1}), \dots, f(\mathbf{u_m})\}$ *è linearmente indipendente;*

iv) $\{\mathbf{u_1}, \dots, \mathbf{u_m}\}$ *è una base di* $U \subseteq V$ *se e solo se* $\{f(\mathbf{u_1}), \dots, f(\mathbf{u_m})\}$ *è una base di* $f(U) \subseteq W$.

Dimostrazione. Per ogni affermazione è sufficiente verificare il primo verso dell'implicazione. Infatti, se questo è vero, a causa dell'esistenza dell'isomorfismo inverso f^{-1} vale anche l'implicazione contraria. Ma allora, il primo punto dell'enunciato è conseguenza della Proposizione 5.4, mentre i rimanenti seguono dalla Proposizione 5.14. $\square$

5.4. Teorema di rappresentazione

Il lettore più attento dovrebbe aver notato una forte somiglianza tra la struttura algebrica definita sull'insieme delle matrici e quella sull'insieme delle applicazioni lineari tra due spazi vettoriali.

In particolare, queste due strutture sono entrambe degli spazi vettoriali ed inoltre esiste una totale interscambiabilità tra le Proposizioni 3.10 e 5.10 rispetto alle proprietà della operazioni $*$ e $\circ$. Infine, abbiamo verificato nell'Esempio 5.11 che esiste un omomorfismo naturale Υ dall'algebra delle matrici nell'algebra di opportune applicazioni lineari. In questa sezione vogliamo addentrarci ulteriormente nell'investigazione di questo legame. In primo luogo ci serve il seguente risultato preliminare, noto come Teorema di interpolazione.

Proposizione 5.18. *Dati V, W spazi vettoriali sul medesimo campo $\mathbb{K}$, con $n = \dim(V) < \infty$, siano $\mathcal{B} = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ una base di V e $\{\mathbf{w_1}, \ldots, \mathbf{w_n}\}$ un generico sottoinsieme di W. Allora esiste un'unica applicazione lineare $f \in \mathrm{Hom}(V, W)$ tale che $f(\mathbf{v_i}) = \mathbf{w_i}$ per ogni $1 \leq i \leq n$.*

Dimostrazione. Per prima cosa dimostriamo l'esistenza di f. Dato $\mathbf{v} \in V$, esiste unica la decomposizione

$$\mathbf{v} = \sum_{i=1}^{n} t_i \cdot \mathbf{v_i}.$$

Definiamo allora la cosiddetta funzione interpolante $f : V \to W$ come

$$f(\mathbf{v}) = \sum_{i=1}^{n} t_i \cdot \mathbf{w_i}.$$

Osserviamo che la richiesta $n < \infty$ è necessaria per garantire la convergenza della sommatoria. Per prima cosa, l'esistenza ed unicità della decomposizione garantisce che f sia ben definita. Inoltre è immediato verificare che $f(\mathbf{v_i}) = \mathbf{w_i}$ per ogni $1 \leq i \leq n$. Infine, proviamo la linearità della funzione. Siano $\mathbf{v}, \tilde{\mathbf{v}} \in V$ e $t, \tilde{t} \in \mathbb{K}$, con

$$\tilde{\mathbf{v}} = \sum_{i=1}^{n} \tilde{t}_i \cdot \mathbf{v_i}.$$

Allora

$$\begin{aligned} f(t \cdot \mathbf{v} + \tilde{t} \cdot \tilde{\mathbf{v}}) &= f\left(t \cdot \left(\sum_{i=1}^{n} t_i \cdot \mathbf{v_i}\right) + \tilde{t} \cdot \left(\sum_{i=1}^{n} \tilde{t}_i \cdot \mathbf{v_i}\right)\right) = f\left(\sum_{i=1}^{n} (tt_i + \tilde{t}\tilde{t}_i) \cdot \mathbf{v_i}\right) = \\ &= \sum_{i=1}^{n} (tt_i + \tilde{t}\tilde{t}_i) \cdot \mathbf{w_i} = t \cdot \left(\sum_{i=1}^{n} t_i \cdot \mathbf{w_i}\right) + \tilde{t}\left(\cdot \sum_{i=1}^{n} \tilde{t}_i \cdot \mathbf{w_i}\right) = t \cdot f(\mathbf{v}) + \tilde{t} \cdot f(\tilde{\mathbf{v}}). \end{aligned}$$

Dimostriamo ora l'unicità della funzione. Sia $\tilde{f} \in \mathrm{Hom}(V, W)$ un'applicazione tale che $\tilde{f}(\mathbf{v_i}) = \mathbf{w_i}$ per ogni $1 \leq i \leq n$. Allora per ogni $\mathbf{v} \in V$ vale

$$\tilde{f}(\mathbf{v}) = \tilde{f}\left(\sum_{i=1}^{n} t_i \cdot \mathbf{v_i}\right) = \sum_{i=1}^{n} t_i \cdot \tilde{f}(\mathbf{v_i}) = \sum_{i=1}^{n} t_i \cdot \mathbf{w_i} = f(\mathbf{v}),$$

da cui deduciamo $\tilde{f} = f$. □

La proposizione precedente afferma che una applicazione lineare è univocamente determinata dalla sua azione sugli elementi di una qualunque base dello spazio di partenza.

Esempio 5.19. Determiniamo tutte le possibili applicazioni lineari in $\mathrm{Hom}(\mathbb{K}, \mathbb{K})$. La base canonica di $\mathbb{K}$ è $\{1\}$. Per la Proposizione 5.18, ogni applicazione è univocamente determinata se fissiamo $f(1) = a$, dove a è un dato elemento di $\mathbb{K}$. Infatti, sul generico elemento t dello spazio di partenza la funzione vale $f(t) = f(t \cdot 1) = t \cdot f(1) = ta$. In conclusione, esiste una corrispondenza biunivoca tra $\mathrm{Hom}(\mathbb{K}, \mathbb{K})$ e $\mathbb{K}$.

Siamo ora pronti ad enunciare e dimostrare un risultato fondamentale sulle applicazioni lineari, il cosiddetto Teorema di rappresentazione. Cominciamo con il definire una funzione che associa ad ogni applicazione lineare in $\mathrm{Hom}(V, W)$ una matrice, per ogni fissata scelta delle basi di V e W.

Definizione 5.20. *Siano V, W spazi vettoriali finitamente generati sul medesimo campo $\mathbb{K}$, con basi rispettivamente $B_V = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ e $B_W = \{\mathbf{w_1}, \ldots, \mathbf{w_m}\}$. Allora la mappa delle componenti per le applicazioni lineari è la funzione*

$$\begin{array}{rccc} \phi_{\mathcal{B}_V \mathcal{B}_W} : & \mathrm{Hom}(V, W) & \longrightarrow & \mathrm{Mat}(m, n; \mathbb{K}) \\ & f & \longmapsto & F|_{\mathcal{B}_V \mathcal{B}_W} \end{array},$$

dove

$$F|_{\mathcal{B}_V \mathcal{B}_W} = \left[\begin{array}{c|c|c} f(\mathbf{v_1})|_{\mathcal{B}_W} & \ldots & f(\mathbf{v_n})|_{\mathcal{B}_W} \end{array} \right].$$

Si dice che $F|_{\mathcal{B}_V \mathcal{B}_W}$ rappresenta l'applicazione f rispetto alle due basi scelte.

Esempio 5.21.

- Per ogni V, W finitamente generati ed ogni scelta delle basi, le applicazioni 0_{VW} ed Id_V sono rappresentate rispettivamente da

$$\begin{aligned} 0_{VW}|_{\mathcal{B}_V \mathcal{B}_W} &= \left[\begin{array}{c|c|c} 0_{VW}(\mathbf{v_1})|_{\mathcal{B}_W} & \ldots & 0_{VW}(\mathbf{v_n})|_{\mathcal{B}_W} \end{array} \right] = \\ &= \left[\begin{array}{c|c|c} \mathbf{0_W}|_{\mathcal{B}_W} & \ldots & \mathbf{0_W}|_{\mathcal{B}_W} \end{array} \right] = \\ &= \left[\begin{array}{c|c|c} 0_{m1} & \ldots & 0_{m1} \end{array} \right] = 0_{mn} \end{aligned}$$

  e

$$\begin{aligned} \mathrm{Id}_V|_{\mathcal{B}_V \mathcal{B}_V} &= \left[\begin{array}{c|c|c} \mathrm{Id}_V(\mathbf{v_1})|_{\mathcal{B}_V} & \ldots & \mathrm{Id}_V(\mathbf{v_n})|_{\mathcal{B}_V} \end{array} \right] = \\ &= \left[\begin{array}{c|c|c} \mathbf{v_1}|_{\mathcal{B}_V} & \ldots & \mathbf{v_n}|_{\mathcal{B}_V} \end{array} \right] = \mathbb{I}_n \ . \end{aligned}$$

- Siano $V = \mathrm{Mat}(n, 1; \mathbb{K}), W = \mathrm{Mat}(m, 1; \mathbb{K})$ ed $A \in \mathrm{Mat}(m, n; \mathbb{K})$. Rispetto alle basi canoniche di V e W, l'applicazione naturale f_A associata alla matrice A è rappresentata da

$$\begin{aligned} F_A|_{\mathcal{B}_{n1} \mathcal{B}_{m1}} &= \left[\begin{array}{c|c|c} f_A(E_{11})|_{\mathcal{B}_{m1}} & \ldots & f_A(E_{n1})|_{\mathcal{B}_{m1}} \end{array} \right] = \\ &= \left[\begin{array}{c|c|c} (A * E_{11})|_{\mathcal{B}_{m1}} & \ldots & (A * E_{n1})|_{\mathcal{B}_{m1}} \end{array} \right] = \\ &= \left[\begin{array}{c|c|c} A_{C(1)}|_{\mathcal{B}_{m1}} & \ldots & A_{C(n)}|_{\mathcal{B}_{m1}} \end{array} \right] = \\ &= \left[\begin{array}{c|c|c} A_{C(1)} & \ldots & A_{C(n)} \end{array} \right] = A \ . \end{aligned}$$

- In $V = W = \mathbb{K}[x]_2$, consideriamo la base $\mathcal{B}_2 = \{1, x, x^2\}$ di V e la base $\mathcal{B}' = \{x^2, x, 1\}$ di W. Allora la derivata algebrica è rappresentata da

$$\begin{aligned} \tfrac{d}{dx}|_{\mathcal{B}_2 \mathcal{B}'} &= \left[\begin{array}{c|c|c} \tfrac{d}{dx}(1)|_{\mathcal{B}'} & \tfrac{d}{dx}(x)|_{\mathcal{B}'} & \tfrac{d}{dx}(x^2)|_{\mathcal{B}'} \end{array} \right] = \\ &= \left[\begin{array}{c|c|c} 0|_{\mathcal{B}'} & 1|_{\mathcal{B}'} & 2x|_{\mathcal{B}'} \end{array} \right] = \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}. \end{aligned}$$

- Sia $V = U \oplus W$, allora una base $\mathcal{B}_V$ di V è data dall'unione di una base $\mathcal{B}_U = \{\mathbf{u_1}, \dots, \mathbf{u_n}\}$ di U ed una base $\mathcal{B}_W = \{\mathbf{w_1}, \dots, \mathbf{w_m}\}$ di W. La matrice che rappresenta l'inclusione $i_{U;V}$ è
$$\begin{aligned}\phi_{\mathcal{B}_U\mathcal{B}_V}(i_{U;V}) &= \left[\ i_{U;V}(\mathbf{u_1})|_{\mathcal{B}_V} \ \middle|\ \dots \ \middle|\ i_{U;V}(\mathbf{u_n})|_{\mathcal{B}_V}\ \right] = \\ &= \left[\ \mathbf{u_1}|_{\mathcal{B}_V} \ \middle|\ \dots \ \middle|\ \mathbf{u_n}|_{\mathcal{B}_V}\ \right] = \left[\begin{array}{c} \mathbb{I}_n \\ \hline 0_{mn} \end{array}\right].\end{aligned}$$
La matrice che rappresenta la proiezione $\mathcal{P}_{U;W}$ è invece
$$\begin{aligned}\phi_{\mathcal{B}_V\mathcal{B}_U}(\mathcal{P}_{U;W}) &= \left[\ \mathcal{P}_{U;W}(\mathbf{u_1})|_{\mathcal{B}_U} \ \middle|\ \dots \ \middle|\ \mathcal{P}_{U;W}(\mathbf{u_n})|_{\mathcal{B}_U} \ \middle|\ \mathcal{P}_{U;W}(\mathbf{w_1})|_{\mathcal{B}_U} \ \middle|\ \dots \ \middle|\ \mathcal{P}_{U;W}(\mathbf{w_m})|_{\mathcal{B}_U}\ \right] = \\ &= \left[\ \mathbf{u_1}|_{\mathcal{B}_U} \ \middle|\ \dots \ \middle|\ \mathbf{u_n}|_{\mathcal{B}_U} \ \middle|\ \mathbf{0}|_{\mathcal{B}_U} \ \middle|\ \dots \ \middle|\ \mathbf{0}|_{\mathcal{B}_U}\ \right] = [\,\mathbb{I}_n \mid 0_{nm}\,]\end{aligned}$$
ed infine la matrice che rappresenta la riflessione $\mathcal{R}_{U;W}$ è
$$\begin{aligned}\phi_{\mathcal{B}_V\mathcal{B}_U}(\mathcal{R}_{U;W}) &= \left[\ \mathcal{R}_{U;W}(\mathbf{u_1})|_{\mathcal{B}_V} \ \middle|\ \dots \ \middle|\ \mathcal{R}_{U;W}(\mathbf{u_n})|_{\mathcal{B}_V} \ \middle|\ \mathcal{R}_{U;W}(\mathbf{w_1})|_{\mathcal{B}_V} \ \middle|\ \dots \ \middle|\ \mathcal{R}_{U;W}(\mathbf{w_m})|_{\mathcal{B}_V}\ \right] = \\ &= \left[\ \mathbf{u_1}|_{\mathcal{B}_V} \ \middle|\ \dots \ \middle|\ \mathbf{u_n}|_{\mathcal{B}_V} \ \middle|\ -\mathbf{w_1}|_{\mathcal{B}_V} \ \middle|\ \dots \ \middle|\ -\mathbf{w_m}|_{\mathcal{B}_V}\ \right] = \left[\begin{array}{c|c} \mathbb{I}_n & 0_{nm} \\ \hline 0_{mn} & -\mathbb{I}_m \end{array}\right].\end{aligned}$$

Teorema 5.22. *Siano V, W spazi vettoriali finitamente generati sul medesimo campo $\mathbb{K}$, con basi rispettivamente B_V, B_W, e sia $f \in \mathrm{Hom}(V, W)$. Allora $F|_{\mathcal{B}_V\mathcal{B}_W}$ è l'unica matrice per cui vale*
$$f(\mathbf{v})|_{\mathcal{B}_W} = F|_{\mathcal{B}_V\mathcal{B}_W} * \mathbf{v}|_{\mathcal{B}_V}, \tag{11}$$
per ogni $\mathbf{v} \in V$. Inoltre la mappa delle componenti $\phi_{\mathcal{B}_V\mathcal{B}_W}$:

i) è un isomorfismo di spazi vettoriali;

ii) rispetta la struttura moltiplicativa, ovvero se U è uno spazio vettoriale su campo $\mathbb{K}$ con base $\mathcal{B}_U$ e $g \in \mathrm{Hom}(U, V)$, allora
$$\phi_{\mathcal{B}_U\mathcal{B}_W}(f \circ g) = \phi_{\mathcal{B}_V\mathcal{B}_W}(f) * \phi_{\mathcal{B}_U\mathcal{B}_V}(g).$$

Dimostrazione. Per prima cosa verifichiamo che valga la formula (11). Se $\mathcal{B}_V = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$, allora per ogni $\mathbf{v} \in V$ esiste unica la decomposizione
$$\mathbf{v} = \sum_{i=1}^{n} t_i \cdot \mathbf{v_i}.$$
Utilizzando la linearità di f, abbiamo
$$f(\mathbf{v}) = \sum_{i=1}^{n} t_i \cdot f(\mathbf{v_i})$$
e quindi
$$f(\mathbf{v})|_{\mathcal{B}_W} = \sum_{i=1}^{n} t_i \cdot f(\mathbf{v_i})|_{\mathcal{B}_W} = \sum_{i=1}^{n} t_i \cdot F|_{\mathcal{B}_V\mathcal{B}_W, C(i)} = F|_{\mathcal{B}_V\mathcal{B}_W} * \mathbf{v}|_{\mathcal{B}_V},$$
dove nell'ultima uguaglianza abbiamo utilizzato il Lemma 3.9. Per dimostrare l'unicità della matrice, supponiamo che $A \in \mathrm{Mat}(m, n; \mathbb{K})$ soddisfi le richieste del teorema, ovvero
$$f(\mathbf{v})|_{\mathcal{B}_W} = A * \mathbf{v}|_{\mathcal{B}_V},$$
per ogni $\mathbf{v} \in V$. Questo è vero in particolare per i vettori di $\mathcal{B}_V$. Quindi, per $1 \le i \le n$ abbiamo
$$f(\mathbf{v_i})|_{\mathcal{B}_W} = A * E_{i1} = F|_{\mathcal{B}_V\mathcal{B}_W} * E_{i1}$$
e di conseguenza, ancora grazie al Lemma 3.9, vale $A_{R(i)} = F|_{\mathcal{B}_V\mathcal{B}_W, R(i)}$. Naturalmente questo implica che $A = F|_{\mathcal{B}_V\mathcal{B}_W}$.

Concentriamoci ora sulle proprietà algebriche della mappa delle componenti.

i) Verifichiamo la linearità della funzione:

$$\begin{aligned}
\phi_{\mathcal{B}_V\mathcal{B}_W}(t_1 \cdot f_1 + t_2 \cdot f_2) &= \left[\ (t_1 \cdot f_1 + t_2 \cdot f_2)(\mathbf{v_1})|_{\mathcal{B}_W}\ \middle|\ \dots\ \middle|\ (t_1 \cdot f_1 + t_2 \cdot f_2)(\mathbf{v_n})|_{\mathcal{B}_W}\ \right] \\
&= \left[\ (t_1 \cdot f_1(\mathbf{v_1}) + t_2 \cdot f_2(\mathbf{v_1}))|_{\mathcal{B}_W}\ \middle|\ \dots\ \middle|\ (t_1 \cdot f_1(\mathbf{v_n}) + t_2 \cdot f_2(\mathbf{v_n}))|_{\mathcal{B}_W}\ \right] \\
&= \left[\ t_1 \cdot f_1(\mathbf{v_1})|_{\mathcal{B}_W} + t_2 \cdot f_2(\mathbf{v_1})|_{\mathcal{B}_W}\ \middle|\ \dots\ \middle|\ t_1 \cdot f_1(\mathbf{v_n})|_{\mathcal{B}_W} + t_2 \cdot f_2(\mathbf{v_n})|_{\mathcal{B}_W}\ \right] \\
&= t_1 \cdot \left[\ f_1(\mathbf{v_1})|_{\mathcal{B}_W}\ \middle|\ \dots\ \middle|\ f_1(\mathbf{v_n})|_{\mathcal{B}_W}\ \right] + t_2 \cdot \left[\ f_2(\mathbf{v_1})|_{\mathcal{B}_W}\ \middle|\ \dots\ \middle|\ f_2(\mathbf{v_n})|_{\mathcal{B}_W}\ \right] \\
&= t_1 \cdot \phi_{\mathcal{B}_V\mathcal{B}_W}(f_1) + t_2 \cdot \phi_{\mathcal{B}_V\mathcal{B}_W}(f_2).
\end{aligned}$$

Rimane da analizzare la biunivocità dell'applicazione. Per definizione, questa proprietà è realizzata se e solo se per ogni matrice

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \in \mathrm{Mat}(m, n; \mathbb{K})$$

esiste un'unica applicazione $f \in \mathrm{Hom}(V, W)$ tale che

$$\phi_{\mathcal{B}_V\mathcal{B}_W}(f) = A.$$

Se tale f esiste, in particolare vale

$$f(\mathbf{v_i})|_{\mathcal{B}_W} = \phi_{\mathcal{B}_V\mathcal{B}_W}(f)_{C(i)} = A_{C(i)},$$

per ogni $1 \leq i \leq n$. Data $\mathcal{B}_W = \{\mathbf{w_1}, \dots, \mathbf{w_m}\}$, questo significa che

$$f(\mathbf{v_i}) = \sum_{j=1}^{m} a_{ji} \cdot \mathbf{w_j}.$$

La Proposizione 5.18 garantisce che esiste esattamente un'unica funzione f di questo tipo.

ii) Sia $\mathcal{B}_U = \{\mathbf{u_1}, \dots, \mathbf{u_p}\}$, allora

$$\begin{aligned}
\phi_{\mathcal{B}_U\mathcal{B}_W}(f \circ g) &= \left[\ (f \circ g)(\mathbf{u_1})|_{\mathcal{B}_W}\ \middle|\ \dots\ \middle|\ (f \circ g)(\mathbf{u_p})|_{\mathcal{B}_W}\ \right] \\
&= \left[\ f(g(\mathbf{u_1}))|_{\mathcal{B}_W}\ \middle|\ \dots\ \middle|\ f(g(\mathbf{u_p}))|_{\mathcal{B}_W}\ \right] \\
&= \left[\ \phi_{\mathcal{B}_V\mathcal{B}_W}(f) * g(\mathbf{u_1})|_{\mathcal{B}_V}\ \middle|\ \dots\ \middle|\ \phi_{\mathcal{B}_V\mathcal{B}_W}(f) * g(\mathbf{u_p})|_{\mathcal{B}_V}\ \right] \\
&= \phi_{\mathcal{B}_V\mathcal{B}_W}(f) * \left[\ g(\mathbf{u_1})|_{\mathcal{B}_V}\ \middle|\ \dots\ \middle|\ g(\mathbf{u_p})|_{\mathcal{B}_V}\ \right] \\
&= \phi_{\mathcal{B}_V\mathcal{B}_W}(f) * \phi_{\mathcal{B}_U\mathcal{B}_V}(g)\,.
\end{aligned}$$

□

Esempio 5.23. Riconsideriamo l'Esempio 5.12, restringendo le applicazioni derivata ed integrale su uno spazio finitamente generato. In questo esempio consideriamo la derivata come la funzione $f_1 = \frac{d}{dx} \in \mathrm{Hom}(\mathbb{K}[x]_2, \mathbb{K}[x]_1)$ e l'integrale è la funzione $f_2 = \int_0^x dx \in \mathrm{Hom}(\mathbb{K}[x]_1, \mathbb{K}[x]_2)$. Rispetto alle basi canoniche, le applicazioni sono rappresentate da:

$$F_1|_{\mathcal{B}_2\mathcal{B}_1} = \begin{bmatrix} f_1(1)|_{\mathcal{B}_1} & f_1(x)|_{\mathcal{B}_1} & f_1(x^2)|_{\mathcal{B}_1} \end{bmatrix} = \begin{bmatrix} 0|_{\mathcal{B}_1} & 1|_{\mathcal{B}_1} & 2x|_{\mathcal{B}_1} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

$$F_2|_{\mathcal{B}_1\mathcal{B}_2} = \begin{bmatrix} f_2(1)|_{\mathcal{B}_2} & f_2(x)|_{\mathcal{B}_2} \end{bmatrix} = \begin{bmatrix} x|_{\mathcal{B}_2} & \frac{x^2}{2}|_{\mathcal{B}_2} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

Allora

$$\begin{aligned}(f_2 \circ f_1)(a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2)|_{\mathcal{B}_2} &= F_2|_{\mathcal{B}_1\mathcal{B}_2} * F_1|_{\mathcal{B}_2\mathcal{B}_1} * (a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2)|_{\mathcal{B}_2} = \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ a_1 \\ a_2 \end{bmatrix}\end{aligned}$$

e di conseguenza $(f_2 \circ f_1)(a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2) = a_1 \cdot x + a_2 \cdot x^2$, come si può facilmente verificare utilizzando direttamente la definizione delle due funzioni. Analogamente

$$(f_1 \circ f_2)(a_0 \cdot 1 + a_1 \cdot x)|_{\mathcal{B}_1} = F_1|_{\mathcal{B}_2\mathcal{B}_1} * F_1|_{\mathcal{B}_1\mathcal{B}_2} * (a_0 \cdot 1 + a_1 \cdot x)|_{\mathcal{B}_1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$$

e di conseguenza $(f_2 \circ f_1)(a_0 \cdot 1 + a_1 \cdot x) = a_0 \cdot 1 + a_1 \cdot x$.

Esempio 5.24. Dato $V = U \oplus W$, è possibile eseguire entrambe le composizioni $i_{U;V} \circ \mathcal{P}_{U;W} \in \mathrm{Hom}(V, V)$ e $\mathcal{P}_{U;W} \circ i_{U;V} \in \mathrm{Hom}(U, U)$. Considerate le basi $\mathcal{B}_U = \{\mathbf{u_1}, \ldots, \mathbf{u_n}\}$, $\mathcal{B}_W = \{\mathbf{w_1}, \ldots, \mathbf{w_m}\}$ e $\mathcal{B}_V = \mathcal{B}_U \cup \mathcal{B}_W$, le matrici che rappresentano le due composizioni sono rispettivamente

$$\phi_{\mathcal{B}_V\mathcal{B}_V}(i_{U;V} \circ \mathcal{P}_{U;W}) = \left[\begin{array}{c} \mathbb{I}_n \\ \hline 0_{mn} \end{array}\right] * [\mathbb{I}_n | 0_{nm}] = \left[\begin{array}{c|c} \mathbb{I}_n & 0_{nm} \\ \hline 0_{mn} & 0_m \end{array}\right]$$

e

$$\phi_{\mathcal{B}_U\mathcal{B}_U}(\mathcal{P}_{U;W} \circ i_{U;V}) = [\mathbb{I}_n | 0_{nm}] * \left[\begin{array}{c} \mathbb{I}_n \\ \hline 0_{mn} \end{array}\right] = \mathbb{I}_n.$$

Il primo caso corrisponde alla proiezione da V su U, dove però la funzione di proiezione ha codominio V. In alcuni testi, questa viene presa come definizione stessa di proiezione. Nel proseguo del nostro studio, noi tenderemo ad usare entrambe le definizioni a seconda della particolare convenienza, specificando in ogni caso la nostra scelta. Nel secondo caso, la composizione delle due applicazioni coincide con l'identità su U.

Corollario 5.25. *Siano V, W spazi vettoriali finitamente generati sul medesimo campo $\mathbb{K}$, con $\dim(V) = n$ e $\dim(W) = m$. Allora $\dim(\mathrm{Hom}(V, W)) = mn$.*

Dimostrazione. L'enunciato segue dai Teoremi 5.15 e 5.22 e da $\dim(\mathrm{Mat}(m, n; \mathbb{K})) = mn$. □

Corollario 5.26. *Siano V, W spazi vettoriali finitamente generati sul medesimo campo $\mathbb{K}$, con $\dim(V) = \dim(W) = n$. Allora $f \in \mathrm{Hom}(V, W)$ è un isomorfismo se e solo se $F|_{\mathcal{B}_V\mathcal{B}_W}$ è invertibile, ed in tal caso*

$$\phi_{\mathcal{B}_V\mathcal{B}_W}(f^{-1}) = F|^{-1}_{\mathcal{B}_V\mathcal{B}_W}.$$

Dimostrazione. Per la biunivocità e per la moltiplicatività della mappa delle componenti, f^{-1} esiste se e solo se esiste $F|^{-1}_{\mathcal{B}_V\mathcal{B}_W}$. Infatti vale:

$$\mathbb{I}_n = \phi_{\mathcal{B}_V\mathcal{B}_V}(\mathrm{Id}_V) = \phi_{\mathcal{B}_V\mathcal{B}_V}(f^{-1} \circ f) = \phi_{\mathcal{B}_W\mathcal{B}_V}(f^{-1}) * \phi_{\mathcal{B}_V\mathcal{B}_W}(f).$$ □

5.5. Applicazioni del teorema di rappresentazione

Il Teorema di rappresentazione ci permette di analizzare le proprietà di una data applicazione lineare attraverso lo studio di una qualsiasi matrice che la rappresenta. Dedichiamo questa sezione conclusiva allo studio di alcune importanti conseguenze di questo teorema.

Proposizione 5.27. *Siano V, W spazi vettoriali finitamente generati sul medesimo campo $\mathbb{K}$, con basi rispettivamente $\mathcal{B}_V$ e $\mathcal{B}_W$, e sia $f \in \mathrm{Hom}(V, W)$. Allora:*

i) $\phi_{\mathcal{B}_V}$ *è un isomorfismo tra* $\ker(f)$ *e* $\ker(F|_{\mathcal{B}_V\mathcal{B}_W})$;

ii) $\phi_{\mathcal{B}_W}$ *è un isomorfismo tra* $\mathrm{Im}(f)$ *e* $C(F|_{\mathcal{B}_V\mathcal{B}_W})$.

DIMOSTRAZIONE. Dato che le mappe delle componenti $\phi_{\mathcal{B}_V}$ e $\phi_{\mathcal{B}_W}$ sono omomorfismi iniettivi definiti sugli spazi V e W, lo sono anche le loro restrizioni rispettivamente a $\ker(f)$ e $\mathrm{Im}(f)$. Pertanto l'unica cosa rimanente da verificare è che $\phi_{\mathcal{B}_V}(\ker(f)) = \ker(F|_{\mathcal{B}_V\mathcal{B}_W})$ e $\phi_{\mathcal{B}_W}(\mathrm{Im}(f)) = C(F|_{\mathcal{B}_V\mathcal{B}_W})$. Le dimostrazioni delle due proprietà sono molto simili.

i) Il vettore $\mathbf{v} \in V$ appartiene a $\ker(f)$ se e solo se $f(\mathbf{v}) = \mathbf{0_W}$. A causa del teorema di rappresentazione, questo è vero se e solo se $F|_{\mathcal{B}_V\mathcal{B}_W} * \mathbf{v}|_{\mathcal{B}_V} = \mathbf{0_W}|_{\mathcal{B}_W} = 0_{m1}$, ovvero se e solo se $\phi_{\mathcal{B}_V}(\mathbf{v}) \in \ker(F|_{\mathcal{B}_V\mathcal{B}_W})$.

ii) Il vettore $\mathbf{w} \in W$ appartiene a $\mathrm{Im}(f)$ se e solo se esiste $\mathbf{v} \in V$ tale che $\mathbf{w} = f(\mathbf{v})$. Sempre per il teorema di rappresentazione, questo è vero se e solo se $\mathbf{w}|_{\mathcal{B}_W} = F|_{\mathcal{B}_V\mathcal{B}_W} * \mathbf{v}|_{\mathcal{B}_V}$. Quest'ultima equazione matriciale rappresenta un sistema lineare, le cui incognite sono le componenti di $\mathbf{v}$. Grazie al Corollario 4.63, sappiamo che la soluzione del sistema esiste se e solo se $\mathbf{w}|_{\mathcal{B}_W} \in C(F|_{\mathcal{B}_V\mathcal{B}_W})$. $\square$

Esempio 5.28. Dati $V = \mathbb{R}[x]_1$ e $W = \mathbb{S}(2;\mathbb{R})$, consideriamo la funzione $f : V \to W$ definita come

$$f(P(x)) = \begin{bmatrix} P(1) & P(0) \\ P(0) & P(-1) \end{bmatrix}.$$

Lasciamo al lettore la verifica della linearità di f. Utilizziamo invece la Proposizione 5.27 per studiare il suo nucleo e la sua immagine. Scegliamo di rappresentare l'applicazione rispetto alle basi canoniche dei due spazi:

$$F|_{\mathcal{B}_1\mathcal{B}_{\mathbb{S}(2,\mathbb{R})}} = \begin{bmatrix} f(1)|_{\mathcal{B}_{\mathbb{S}(2,\mathbb{R})}} & f(x)|_{\mathcal{B}_{\mathbb{S}(2,\mathbb{R})}} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}|_{\mathcal{B}_{\mathbb{S}(2,\mathbb{R})}} & \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}|_{\mathcal{B}_{\mathbb{S}(2,\mathbb{R})}} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 1 & -1 \end{bmatrix}.$$

Riduciamo a scala la matrice:

$$F|_{\mathcal{B}_1\mathcal{B}_{\mathbb{S}(2,\mathbb{R})}} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 1 & -1 \end{bmatrix} \xrightarrow[R_3-R_1\to R_3]{R_2-2R_1\to R_2} \begin{bmatrix} 1 & 1 \\ 0 & -2 \\ 0 & -2 \end{bmatrix} \xrightarrow[R_3-R_2\to R_3]{-\frac{1}{2}R_2\to R_2} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Di conseguenza $r(F|_{\mathcal{B}_1\mathcal{B}_{\mathbb{S}(2,\mathbb{R})}}) = 2 = \dim(C(F|_{\mathcal{B}_1\mathcal{B}_{\mathbb{S}(2,\mathbb{R})}})) = \dim(\mathrm{Im}(f))$. Una base di $C(A)$ è costituita dalle due colonne della matrice. Esse corrispondono alle componenti dei vettori della base di $\mathrm{Im}(f)$ rispetto a $\mathcal{B}_{\mathbb{S}(2,\mathbb{R})}$. Per costruzione abbiamo quindi che una base di $\mathrm{Im}(f)$ sono le due matrici immagini dei polinomi 1 e x, cioè

$$\mathcal{B}_{\mathrm{Im}(f)} = \left\{ \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right\}.$$

Il nucleo di f ha dimensione $\dim(\ker(f)) = \dim(\ker(F|_{\mathcal{B}_1\mathcal{B}_{\mathbb{S}(2,\mathbb{R})}})) = 2 - r(F|_{\mathcal{B}_1\mathcal{B}_{\mathbb{S}(2,\mathbb{R})}}) = 0$. Quindi il nucleo è il sottospazio banale, $\ker(f) = \{0\}$, da cui deduciamo che f è inettiva.

Ovviamente la matrice $F_{\mathcal{B}_V\mathcal{B}_W}$ rappresentante l'applicazione f dipende dalla scelta delle basi di V e W, di conseguenza altrettanto vale per il suo nucleo e lo spazio delle colonne. La Proposizione 5.27 afferma però che tutti questi spazi vettoriali sono isomorfi e quindi, per il Teorema di isomorfismo, devono avere la stessa dimensione.

Definizione 5.29. *Sia* $f \in \mathrm{Hom}(V, W)$, *allora:*

i) il rango dell'applicazione è la dimensione dello spazio immagine $r(f) = \dim(\mathrm{Im}(f))$;

ii) la nullità dell'applicazione è la dimensione del nucleo $k(f) = \dim(\ker(f))$.

Per quanto detto precedentemente, per ogni matrice rappresentativa dell'applicazione valgono

$$r(f) = \dim(C(F|_{\mathcal{B}_V\mathcal{B}_W})) = r(F|_{\mathcal{B}_V\mathcal{B}_W}) \quad \text{e} \quad k(f) = \dim(\ker(F|_{\mathcal{B}_V\mathcal{B}_W})) = k(F|_{\mathcal{B}_V\mathcal{B}_W}).$$

Il seguente risultato, conosciuto come Teorema di nullità più rango, fornisce una relazione tra queste due quantità.

Teorema 5.30. *Sia $f \in \mathrm{Hom}(V, W)$, dove lo spazio V è finitamente generato. Allora*

$$\dim(V) = r(f) + k(f).$$

Dimostrazione. Per prima cosa, consideriamo la restrizione dell'applicazione

$$\begin{array}{rccc} \tilde{f}: & V & \longrightarrow & \mathrm{Im}(f) \\ & \mathbf{v} & \longmapsto & f(\mathbf{v}) \end{array}.$$

Il dominio di $\tilde{f}$ ha dimensione finita per ipotesi, ma anche il codominio di $\tilde{f}$ soddisfa questa proprietà. Infatti, osserviamo che se $\mathcal{B}_V = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ è base di V, allora la Proposizione 5.4 afferma che $\{f(\mathbf{v_1}), \dots, f(\mathbf{v_n})\}$ è un insieme di generatori di $\mathrm{Im}(f)$, quindi $\dim(\mathrm{Im}(f)) \leq \dim(V) = n < \infty$. Di conseguenza, è possibile rappresentare l'applicazione $\tilde{f}$ attraverso il Teorema di rappresentazione.

Se scegliamo una base $\mathcal{B}_{\mathrm{Im}(f)}$ del codominio di $\tilde{f}$, questa applicazione è rappresentata da una matrice $F|_{\mathcal{B}_V\mathcal{B}_{\mathrm{Im}(f)}} \in \mathrm{Mat}(m, n; \mathbb{K})$, dove $m = \dim(\mathrm{Im}(f))$ ed $n = \dim(V)$. Poiché $\mathrm{Im}(f) = \mathrm{Im}(\tilde{f})$ e $\ker(f) = \ker(\tilde{f})$, segue che $r(f) = r(\tilde{f}) = r(F|_{\mathcal{B}_V\mathcal{B}_{\mathrm{Im}(f)}})$ e $k(f) = k(\tilde{f}) = k(F|_{\mathcal{B}_V\mathcal{B}_{\mathrm{Im}(f)}})$. Il teorema di Rouché-Capelli ci dice che

$$k(F|_{\mathcal{B}_V\mathcal{B}_{\mathrm{Im}(f)}})) = n - r(F|_{\mathcal{B}_V\mathcal{B}_{\mathrm{Im}(f)}}))$$

da cui

$$k(f) = \dim(V) - r(f)$$

e quindi la tesi. $\square$

Corollario 5.31. *Siano V, W spazi vettoriali finitamente generati sul medesimo campo $\mathbb{K}$ e sia $f \in \mathrm{Hom}(V, W)$. Allora:*

i) f è iniettiva se e solo se $r(f) = \dim(V)$;
ii) f è suriettiva se e solo se $r(f) = \dim(W)$;
iii) f è biettiva se e solo se $r(f) = \dim(V) = \dim(W)$.

Dimostrazione. Il primo punto segue dal Corollario 5.7 e dal Teorema 5.30. Per il secondo punto, osserviamo che f è suriettiva se e solo se $\mathrm{Im}(f) = W$. Poiché $\mathrm{Im}(f)$ è un sottospazio di W, dalla Proposizione 4.46 deduciamo che f è suriettiva se e solo se $r(f) = \dim(W)$. Infine, l'ultimo punto segue dai due precedenti. $\square$

Concludiamo il capitolo sulle applicazioni lineari con lo studio del cambiamento di base. Come abbiamo già accennato, la matrice rappresentativa di una applicazione dipende dalla scelta delle basi del dominio e del codominio. È però possibile passare da una rappresentazione all'altra attraverso una semplice regola, che qui di seguito illustriamo.

Definizione 5.32. *Dato uno spazio vettoriale V finitamente generato, siano $\mathcal{B} = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ e $\mathcal{B}' = \{\mathbf{v'_1}, \dots, \mathbf{v'_n}\}$ due basi. Allora*

$$M_{\mathcal{B}\mathcal{B}'} = \left[\begin{array}{c|c|c} \mathbf{v_1}|_{\mathcal{B}'} & \dots & \mathbf{v_n}|_{\mathcal{B}'} \end{array}\right] \in \mathrm{Mat}(n, n; \mathbb{K})$$

è detta matrice di cambio base.

Proposizione 5.33. *Siano V, W spazi vettoriali finitamente generati sul medesimo campo $\mathbb{K}$, rispettivamente con basi $\mathcal{B}_V, \mathcal{B}'_V$ e $\mathcal{B}_W, \mathcal{B}'_W$. Allora:*

i) per ogni $\mathbf{v} \in V$ vale $\mathbf{v}|_{\mathcal{B}'_V} = M_{\mathcal{B}_V\mathcal{B}'_V} * \mathbf{v}|_{\mathcal{B}_V}$;

ii) per ogni $f \in \mathrm{Hom}(V, W)$ vale $F|_{\mathcal{B}'_V\mathcal{B}'_W} = M_{\mathcal{B}_W\mathcal{B}'_W} * F|_{\mathcal{B}_V\mathcal{B}_W} * M_{\mathcal{B}'_V\mathcal{B}_V}$.

Dimostrazione. Consideriamo l'applicazione identità Id_V e la sua matrice rappresentativa rispetto alle basi $\mathcal{B}_V = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ e $\mathcal{B}'_V = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$:

$$\phi_{\mathcal{B}_V\mathcal{B}'_V}(\mathrm{Id}_V) = \left[\begin{array}{c|c|c} \mathrm{Id}_V(\mathbf{v_1})|_{\mathcal{B}'_V} & \dots & \mathrm{Id}_V(\mathbf{v_n})|_{\mathcal{B}'_V} \end{array}\right] = \left[\begin{array}{c|c|c} \mathbf{v_1}|_{\mathcal{B}'_V} & \dots & \mathbf{v_n}_{\mathcal{B}'_V} \end{array}\right] = M_{\mathcal{B}_V\mathcal{B}'_V}.$$

Allora:

i) per ogni $\mathbf{v} \in V$ vale

$$\mathbf{v}|_{\mathcal{B}'_V} = \mathrm{Id}_V(\mathbf{v})|_{\mathcal{B}'_V} = \phi_{\mathcal{B}_V\mathcal{B}'_V}(\mathrm{Id}_V) * \mathbf{v}|_{\mathcal{B}_V} = M_{\mathcal{B}_V\mathcal{B}'_V} * \mathbf{v}|_{\mathcal{B}_V},$$

dove nella seconda uguaglianza abbiamo applicato il Teorema di rappresentazione;

ii) utilizzando la regola di cambio base per le componenti di un vettore dimostrata nel punto precedente ed il Teorema di rappresentazione, otteniamo

$$\begin{aligned} f(\mathbf{v})|_{\mathcal{B}'_W} &= M_{\mathcal{B}_W\mathcal{B}'_W} * f(\mathbf{v})|_{\mathcal{B}_W} = \\ &= M_{\mathcal{B}_W\mathcal{B}'_W} * (F|_{\mathcal{B}_V\mathcal{B}_W} * \mathbf{v}|_{\mathcal{B}_V}) = \\ &= M_{\mathcal{B}_W\mathcal{B}'_W} * (F|_{\mathcal{B}_V\mathcal{B}_W} * (M_{\mathcal{B}'_V\mathcal{B}_V} * \mathbf{v}|_{\mathcal{B}'_V})) \\ &= (M_{\mathcal{B}_W\mathcal{B}'_W} * F|_{\mathcal{B}_V\mathcal{B}_W} * M_{\mathcal{B}'_V\mathcal{B}_V}) * \mathbf{v}|_{\mathcal{B}'_V}. \end{aligned}$$

Nell'ultima uguaglianza abbiamo utilizzato l'associatività del prodotto di matrici. La dimostrazione si completa sfruttando l'unicità della matrice che rappresenta f rispetto alla coppia di basi $\mathcal{B}'_V, \mathcal{B}'_W$. □

Esempio 5.34. Dato $V = W = \mathbb{K}[x]_1$, consideriamo l'applicazione derivata $\frac{d}{dx} \in \mathrm{Hom}(V, W)$ e rappresentiamola rispetto alle coppie di basi $\mathcal{B}_V = \mathcal{B}_W = \{1, x\} = \mathcal{B}$ e $\mathcal{B}'_V = \mathcal{B}'_W = \{1+x, 1-x\} = \mathcal{B}'$. Utilizzando la definizione abbiamo:

$$F|_{\mathcal{B}\mathcal{B}} = \begin{bmatrix} \frac{d}{dx}1|_{\mathcal{B}} & \frac{d}{dx}x|_{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} 0|_{\mathcal{B}} & -1|_{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

$$F|_{\mathcal{B}'\mathcal{B}'} = \begin{bmatrix} \frac{d}{dx}(1+x)|_{\mathcal{B}'} & \frac{d}{dx}(1-x)|_{\mathcal{B}'} \end{bmatrix} = \begin{bmatrix} 1|_{\mathcal{B}'} & 1|_{\mathcal{B}'} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

La due matrici di cambio base sono:

$$M_{\mathcal{B}\mathcal{B}'} = \begin{bmatrix} 1|_{\mathcal{B}'} & x|_{\mathcal{B}'} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}, \qquad M_{\mathcal{B}'\mathcal{B}} = \begin{bmatrix} 1+x|_{\mathcal{B}'} & 1-x|_{\mathcal{B}'} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Allora è facile verificare che vale la regola del cambio base:

$$F|_{\mathcal{B}'\mathcal{B}'} = M_{\mathcal{B}\mathcal{B}'} * F|_{\mathcal{B}\mathcal{B}} * M_{\mathcal{B}'\mathcal{B}}.$$

Notiamo anche che $M_{\mathcal{B}\mathcal{B}'} * M_{\mathcal{B}'\mathcal{B}} = M_{\mathcal{B}'\mathcal{B}} * M_{\mathcal{B}\mathcal{B}'} = \mathbb{I}_2$. Questo è un risultato generale.

Lemma 5.35. *Dato uno spazio vettoriale V finitamente generato con $\mathcal{B}$ e $\mathcal{B}'$ due basi, allora vale $M_{\mathcal{B}'\mathcal{B}} = M^{-1}_{\mathcal{B}\mathcal{B}'}$.*

Dimostrazione. Utilizziamo il Teorema di rappresentazione sulla matrice identità:

$$\mathbb{I}_n = \mathrm{Id}_V|_{\mathcal{B}\mathcal{B}} = M_{\mathcal{B}'\mathcal{B}} * \mathrm{Id}_V|_{\mathcal{B}'\mathcal{B}'} * M_{\mathcal{B}\mathcal{B}'} = M_{\mathcal{B}'\mathcal{B}} * \mathbb{I}_n * M_{\mathcal{B}\mathcal{B}'} = M_{\mathcal{B}'\mathcal{B}} * M_{\mathcal{B}\mathcal{B}'},$$

da cui segue la tesi. □

Osservazione 5.36. Dall'ultimo lemma deduciamo che le matrici di cambio base sono invertibili. Vale anche il viceversa, ovvero che ogni matrice invertibile corrisponde ad una matrice di cambio base. Sia infatti $\mathcal{B} = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ una base dello spazio vettoriale V finitamente generato su campo $\mathbb{K}$, e sia $S = [s_{ij}] \in GL(n; \mathbb{K})$. Definiamo i vettori

$$\mathbf{v_i'} = \sum_{j=1}^{n} s_{ji} \cdot \mathbf{v_i},$$

che formano l'insieme $\mathcal{B}' = \{\mathbf{v_1'}, \ldots, \mathbf{v_n'}\}$. Segue che

$$\left[\ \mathbf{v_1'}|_{\mathcal{B}}\ \middle|\ \ldots\ \middle|\ \mathbf{v_n'}|_{\mathcal{B}}\ \right] = S.$$

Poiché $\det(S) \neq 0$, questo significa che $\mathcal{B}'$ è un insieme indipendente formato da n vettori e quindi è una base di V, mentre S è la matrice del cambiamento di base da $\mathcal{B}$ a $\mathcal{B}'$. Grazie a questo ragionamento possiamo in conclusione affermare che esiste una corrispondenza biunivoca tra l'insieme delle basi di V ed il gruppo delle matrici invertibili $GL(n; \mathbb{K})$.

CAPITOLO 6

Endomorfismi e diagonalizzazione

6.1. Endomorfismi, autovalori, autovettori ed autospazi

In questo capitolo ci concentreremo sull'analisi di una classe particolare di applicazioni lineari, quelle per le quali dominio e codominio coincidono.

Definizione 6.1. *Dato V uno spazio vettoriale, un'applicazione lineare $f \in \mathrm{Hom}(V, V)$ si dice endomorfismo di V. L'insieme degli endomorfismi di V si indica con* $\mathrm{End}(V)$*. Un endomorfismo che sia anche isomorfismo si dice automorfismo di V ed indichiamo con $GL(V)$ l'insieme di tutti gli automorfismi di V.*

Nello spazio $\mathrm{End}(V)$ abbiamo una seconda operazione interna oltre alla somma, ovvero la composizione di due applicazioni lineari. Questo ci permette di definire l'elevazione ad una potenza naturale di una applicazione e quindi il concetto di polinomio di applicazioni. Chiamiamo

$$f^k = \begin{cases} \mathrm{Id}_V & \text{se } k = 0, \\ f \circ \ldots \circ f & k \text{ volte se } k \in \mathbb{N}^*. \end{cases}$$

Allora un polinomio di f a coefficienti $c_i \in \mathbb{K},\ 0 \leq i \leq k$, è un'espressione del tipo

$$P(f) = \sum_{i=0}^{k} c_i \cdot f^i = c_0 \cdot \mathrm{Id}_V + c_1 \cdot f + c_2 \cdot f^2 + \cdots + c_k \cdot f^k.$$

In uno spazio V finitamente generato, scelta una coppia di basi $\mathcal{B}, \mathcal{B}'$ un endomorfismo f è rappresentato da una matrice quadrata $F|_{\mathcal{B}\mathcal{B}'}$. Benché non indispensabile, di solito è conveniente scegliere uguali le basi $\mathcal{B}$ e $\mathcal{B}'$ di V necessarie a rappresentare f. In tal caso, si utilizza la notazione semplificata $F|_{\mathcal{B}}$ per indicare la sua matrice rappresentativa. Allora, per la linearità e la moltiplicatività della mappa delle componenti, abbiamo

$$\phi_{\mathcal{B}}(P(f)) = P(F|_{\mathcal{B}}) = \sum_{i=0}^{k} c_i \cdot F|_{\mathcal{B}}{}^i.$$

Nel caso degli automorfismi, l'operazione di composizione ammette l'esistenza dell'applicazione inversa e quindi la struttura $(GL(V), \circ)$ è un gruppo. Inoltre, le proprietà di biiettività e moltiplicatività della mappa delle componenti ci permettono di affermare che $(GL(V), \circ)$ è isomorfo a $GL(n; \mathbb{K})$ tramite $\phi_{\mathcal{B}}$, dove ovviamente $n = \dim(V)$.

Gli endomorfismi presentano alcune proprietà molto speciali, che andiamo ad illustrare nel prossimo esempio.

Esempio 6.2. In $V = \mathrm{Mat}(2, 1; RR)$ consideriamo l'applicazione $f_A \in \mathrm{End}(V)$ associata alla matrice

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

Le due colonne della matrice A corrispondono alle immagini dei due vettori della base canonica $\mathcal{B}_{21} = \{E_{11}, E_{21}\}$. Quindi

$$f_A(E_{11}) = 2 \cdot E_{11} - 1 \cdot E_{21} \qquad \text{e} \qquad f_A(E_{21}) = 3 \cdot E_{11} - 2 \cdot E_{21}.$$

Benché il concetto di rotazione verrà definito con precisione nel Capitolo 8 quando parleremo di angoli tra vettori, possiamo qui affermare che un endomorfismo f agisce su un vettore $\mathbf{v}$ (anche) attraverso una rotazione se l'immagine $f(\mathbf{v})$ non è proporzionale al vettore di partenza. In caso contrario, diciamo che l'applicazione f effettua un puro riscalamento del vettore. Dal punto di vista geometrico, possiamo quindi interpretare la precedente funzione f_A dicendo che essa è una trasformazione dei vettori di V che in generale agisce attraverso un riscalamento ed una rotazione. Esistono però due vettori speciali, su cui f_A si comporta in maniera più semplice. Definiamo

$$\mathbf{v_1} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad \text{e} \qquad \mathbf{v_2} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

Allora

$$\begin{aligned} f_A(\mathbf{v_1}) &= \begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix} * \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -1 \cdot \mathbf{v_1}, \\ f_A(\mathbf{v_2}) &= \begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix} * \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix} = 1 \cdot \mathbf{v_2}. \end{aligned}$$

I due vettori vengono trasformati in multipli di loro stessi. Geometricamente questo significa che $\mathbf{v_1}, \mathbf{v_2}$ vengono esclusivamente riscalati da f_A.

Analizziamo l'importanza di questa osservazione per lo studio dell'applicazione. Dato che $\mathbf{v_1}, \mathbf{v_2}$ sono linearmente indipendenti, essi costituiscono una base $\mathcal{B}$ di V. Rispetto ad essa, la matrice rappresentativa di f_A è

$$F|_{\mathcal{B}} = \begin{bmatrix} f_A(\mathbf{v_1})|_{\mathcal{B}} & f_A(\mathbf{v_2})|_{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Come si può osservare, $F|_{\mathcal{B}}$ è estremamente semplice, essendo una matrice diagonale le cui entrate non nulle sono i coefficienti per i quali vengono riscalati i vettori $\mathbf{v_1}, \mathbf{v_2}$. Rispetto a questa base è immediato verificare alcune proprietà di f_A. Ad esempio, è chiaro che f_A è un'applicazione biettiva. É altrettanto evidente che f_A è un'involuzione, ovvero $f_A^2 = \mathrm{Id}_V$. Questo è meno apparente dall'analisi della matrice di partenza A, in quanto è necessario eseguire esplicitamente i conti per verificare che $A^2 = \mathbb{I}_2$. Infine, dal comportamento su $\mathbf{v_1}$ e $\mathbf{v_2}$ si capisce facilmente che f_A è la riflessione rispetto al sottospazio $\mathcal{L}(\mathbf{v_2})$ e parallela a $\mathcal{L}(\mathbf{v_1})$.

Lo studio della diagonalizzabilità di un endomorfismo consiste nell'indagare l'esistenza di una base di vettori avente le proprietà speciali che abbiamo visto nell'esempio, rispetto alla quale l'applicazione è rappresentata appunto da una matrice diagonale.

Definizione 6.3. *Sia V uno spazio vettoriale su campo $\mathbb{K}$ e sia $f \in \mathrm{End}(V)$. Allora il vettore $\mathbf{v} \in V$ si dice autovettore di f se:*

i) $\mathbf{v} \neq \mathbf{0}$;

ii) esiste $\lambda \in \mathbb{K}$ tale che $f(\mathbf{v}) = \lambda \cdot \mathbf{v}$.

Lo scalare λ è detto autovalore di f associato all'autovettore $\mathbf{v}$.

Definizione 6.4. *Sia V uno spazio vettoriale su campo $\mathbb{K}$ e sia $f \in \mathrm{End}(V)$. L'applicazione è detta diagonalizzabile se esiste una base di $\mathcal{B}$ di V composta da autovettori di f.*

Possiamo ora facilmente ricondurci alle considerazioni sulla matrice rappresentativa di una applicazione diagonalizzabile, attraverso il seguente teorema, noto anche come primo criterio di diagonalizzabilità.

Teorema 6.5. *Sia V uno spazio vettoriale finitamente generato su campo $\mathbb{K}$ e sia $f \in \operatorname{End}(V)$. Allora l'applicazione è diagonalizzabile se e solo se esiste una base esiste una base di $\mathcal{B}$ di V tale che la matrice rappresentativa $F|_{\mathcal{B}}$ di f è diagonale.*

Dimostrazione. Sia f diagonalizzabile, con $\mathcal{B} = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ base di V composta da autovettori associati agli autovalori $\lambda_1 \dots, \lambda_n$. Allora

$$F|_{\mathcal{B}} = \begin{bmatrix} f(\mathbf{v_1})|_{\mathcal{B}} & \dots & f(\mathbf{v_n})|_{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} \lambda_1 \cdot \mathbf{v_1}|_{\mathcal{B}} & \dots & \lambda_n \cdot \mathbf{v_n}|_{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix} \in \mathbb{D}(n; \mathbb{K}).$$

Viceversa, sia $\mathcal{B} = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ una base rispetto a cui $F|_{\mathcal{B}}$ sia diagonale. Se denotiamo con a_{ij} gli elementi della matrice, allora

$$f(\mathbf{v_i}) = \sum_{j=1}^{n} a_{ij} \cdot \mathbf{v_j} = a_{ii} \cdot \mathbf{v_i}\,.$$

Di conseguenza $\mathbf{v_i}$ è un autovettore di f con autovalore associato a_{ii}, per ogni $1 \leq i \leq n$. □

Esempio 6.6.

- Nello spazio V di dimensione n, l'applicazione Id_V è diagonalizzabile. Infatti, rispetto a qualsiasi base $\mathcal{B}$ la matrice rappresentativa è
$$\operatorname{Id}_V|_{\mathcal{B}} = \mathbb{I}_n\,,$$
che ovviamente è diagonale. In questo caso tutti i vettori di V sono autovettori di Id_V, tutti associati al medesimo autovalore 1. Analogamente, anche l'applicazione nulla 0_V è diagonalizzabile in quanto rappresentata dalla matrice $0_n \in \mathbb{D}(n, \mathbb{K})$, ed ogni vettore di V è autovettore associato all'autovalore 0.
- Non tutte le applicazioni sono diagonalizzabili. Ad esempio, consideriamo l'applicazione f_A definita dalla matrice
$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$
che agisce sullo spazio $V = \operatorname{Mat}(2, 1; \mathbb{R})$. La sua azione sugli elementi della base canonica è
$$f_A(E_{11}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} * \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \text{ed} \qquad f_A(E_{21}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} * \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$
Come vedremo meglio quando parleremo di spazi euclidei, dal punto di vista geometrico questa applicazione è interpretabile come la pura rotazione dei vettori della base in senso antiorario di un angolo pari a $\frac{\pi}{2}$. Per linearità, l'applicazione agisce nello stesso modo su tutti i vettori colonna. Pertanto, non possono esistere autovettori di f_A.

Osservazione 6.7. Benché il concetto di diagonalizzabilità è definito relativamente ad una applicazione, è prassi comune parlare di autovettori ed autovalori della matrice A quando ci si riferisce agli autovalori ed autovettori dell'endomorfismo associato f_A.

Concludiamo la sezione studiando l'insieme degli autovettori di un'applicazione data. Come abbiamo appena visto, esistono endomorfismi che non possiedono autovettori. D'altra parte, qualora esistano, allora ve ne sono infiniti associati a ciascun autovalore.

Definizione 6.8. *Sia V uno spazio vettoriale su campo $\mathbb{K}$ e sia $f \in \mathrm{End}(V)$. Se λ è un autovalore di f, allora l'insieme*

$$V_\lambda = \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda \cdot \mathbf{v}\}$$

si chiama autospazio di f associato a λ.

Osserviamo che un autospazio contiene tutti gli autovettori associati al medesimo autovalore λ. Inoltre, vale $f(\mathbf{0}) = \mathbf{0} = \lambda \cdot \mathbf{0}$ per ogni $\lambda \in \mathbb{K}$, quindi anche $\mathbf{0}$ appartiene all'autospazio, benché per definizione tale vettore non sia un autovettore. Per gli autospazi vale il seguente teorema.

Proposizione 6.9. *L'insieme V_λ è un sottospazio di V avente* $\dim(V) \geq 1$.

Dimostrazione. Abbiamo appena dimostrato che $\mathbf{0} \in V_\lambda$, rimane pertanto da verificare la chiusura dell'insieme rispetto alle combinazioni lineari. Se $\mathbf{v_1}, \mathbf{v_2} \in V_\lambda$ e $t_1, t_2 \in \mathbb{K}$, allora

$$f(t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2}) = t_1 \cdot f(\mathbf{v_1}) + t_2 \cdot f(\mathbf{v_2}) = t_1 \cdot (\lambda \cdot \mathbf{v_1}) + t_2 \cdot (\lambda \cdot \mathbf{v_2}) = \lambda \cdot (t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2}),$$

da cui deduciamo che $t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2} \in V_\lambda$. Infine, $V_\lambda \neq \{\mathbf{0}\}$ altrimenti per definizione λ non sarebbe un autovalore. Di conseguenza $\dim(V_\lambda) \geq 1$. □

6.2. Relazione di similitudine

La matrice rappresentativa di un'applicazione $f \in \mathrm{Hom}(V, W)$ dipende dalla scelta delle basi $\mathcal{B}_V$ di V. D'altra parte, tali matrici sono legate tra loro dalla regola del cambio di base enunciata nella Proposizione 5.33. Analizziamo più a fondo tale relazione nel caso degli endomorfismi.

Definizione 6.10. *Date due matrici $A, B \in \mathrm{Mat}(n, n; \mathbb{K})$, B si dice simile ad A, o in relazione di similitudine con A, se esiste $S \in \mathrm{Mat}(n, n; \mathbb{K})$ invertibile tale che $B = S^{-1} * A * S$.*

Proposizione 6.11. *La similitudine è una relazione di equivalenza nell'insieme* $\mathrm{Mat}(n, n; \mathbb{K})$.

Dimostrazione. In primo luogo vale la proprietà riflessiva, poiché $A = \mathbb{I}_n^{-1} * A * \mathbb{I}_n$. Supponiamo ora che B sia simile ad A, allora $B = S^{-1} * A * S$. Di conseguenza vale $A = S * B * S^{-1}$, ovvero A è simile a B e quindi vale la proprietà di simmetria. Infine, siano B simile ad A e C simile a B. Allora $C = \tilde{S}^{-1} * B * \tilde{S} = \tilde{S}^{-1} * (S^{-1} * A * S) * \tilde{S} = (S * \tilde{S})^{-1} * A * (S * \tilde{S})$, ovvero C è simile ad A e quindi vale anche la proprietà transitiva. □

Denotiamo con R la relazione di similitudine. Ricordiamo che la classe di similitudine $[A]$ è l'insieme di tutte le matrici simili ad A, l'insieme quoziente $\mathrm{Mat}(n, n; \mathbb{K})/R$ è l'insieme i cui elementi sono tali classi ed infine la proiezione naturale π è la funzione che associa ad ogni matrice la sua classe di similitudine. Una proprietà importante della relazione di similitudine è che le funzioni naturali sull'insieme $\mathrm{Mat}(n, n; \mathbb{K})$ quali rango, determinante e traccia dipendono solo dalla classe di similitudine di una matrice.

Proposizione 6.12. *Le seguenti funzioni sull'insieme quoziente $\mathrm{Mat}(n, n; \mathbb{K})/R$ sono ben definite:*

i) il rango di una classe di similitudine

$$\begin{array}{rccc} r: & \mathrm{Mat}(n, n; \mathbb{K})/R & \longrightarrow & \mathbb{N} \\ & [A] & \longmapsto & r(A) \end{array};$$

ii) la traccia di una classe di similitudine

$$\begin{array}{rccc} \mathrm{Tr}: & \mathrm{Mat}(n, n; \mathbb{K})/R & \longrightarrow & \mathbb{K} \\ & [A] & \longmapsto & \mathrm{Tr}(A) \end{array};$$

iii) il determinante di una classe di similitudine

$$\begin{array}{rccc} \det: & \mathrm{Mat}(n,n;\mathbb{K})/R & \longrightarrow & \mathbb{K} \\ & [A] & \longmapsto & \det(A) \end{array}.$$

Dimostrazione. Dobbiamo verificare che rango, traccia e determinante di due matrici simili A e $B = S^{-1} * A * S$ sono uguali.

i) Poiché S è invertibile, per il Corollario 3.53 essa è il prodotto di matrici elementari. La stessa cosa vale naturalmente anche per S^{-1}. Di conseguenza la matrice $S^{-1} * A$ è il risultato di una sequenza finita di operazioni elementari applicate alle righe di A e quindi $r(S^{-1} * A) = r(A)$. Analogamente, $(S^{-1} * A) * S$ si ottiene operando una sequenza finita di operazioni elementari sulle colonne di $S^{-1} * A$. Come conseguenza dei risultati contenuti nel Sezione 4.7, sappiamo che esiste una totale simmetria tra le righe e le colonne di una matrice relativamente alle proprietà del rango, di conseguenza $r((S^{-1} * A) * S) = r(S^{-1} * A) = r(A)$.

ii) Grazie alla proprietà di ciclicità, abbiamo

$$\mathrm{Tr}(B) = \mathrm{Tr}(S^{-1} * A * S) = \mathrm{Tr}(S * S^{-1} * A)) = \mathrm{Tr}(\mathbb{I}_n * A) = \mathrm{Tr}(A).$$

iii) Grazie al teorema di Binét ed alle proprietà del determinante, abbiamo

$$\det(B) = \det(S^{-1} * A * S) = \det(S^{-1}) \det(A) \det(S) = \det(S)^{-1} \det(S) \det(A) = \det(A). \qquad \square$$

Rango, determinante e traccia vengono detti invarianti per similitudine. La proposizione precedente è particolarmente utile per certificare la non similitudine tra due matrici. Infatti, se due matrici hanno differente anche un singolo invariante, allora sicuramente non possono essere simili. Il viceversa non è vero. Ad esempio, consideriamo le matrici

$$A = \mathbb{I}_2 \quad \text{e} \quad B = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \qquad \Rightarrow \qquad r(A) = r(B) = 2,\ \mathrm{Tr}(A) = \mathrm{Tr}(B) = 2,\ \det(A) = \det(B) = 1,$$

per ogni elemento $a \in \mathbb{K}$. D'altra parte, per definizione le matrici A e B sono simili se e solo se esiste una matrice invertibile $S \in \mathrm{Mat}(2,2;\mathbb{K})$ tale che

$$S^{-1} * A * S = S^{-1} * \mathbb{I}_2 * S = S^{-1} * S = \mathbb{I}_2 = B,$$

ovvero se e solo se $a = 0$.

Esistono altri invarianti per similitudine di una matrice A. I più significativi si ricavano attraverso il calcolo del cosiddetto polinomio caratteristico di A. Per introdurlo, abbiamo bisogno del seguente risultato preliminare.

Lemma 6.13. *Il determinante della matrice $A \in \mathrm{Mat}(n,n;\mathbb{K})$ è un polinomio multivariato di grado n in $\mathbb{K}[a_{11},\dots,a_{nn}]$, dove le variabili a_{ij} sono tutte le entrate della matrice, per $1 \le i,j \le n$. Il polinomio è ottenibile come combinazione lineare con coefficienti ± 1 di tutti i possibili monomi del tipo $a_{1j_1} \dots a_{nj_n}$, dove $j_1,\dots,j_n$ sono distinti tra loro.*

Dimostrazione. Verifichiamo l'enunciato per induzione su n. Il caso $n = 1$ è ovvio. Supponiamo quindi che la tesi sia vera fino ad $n-1$. Allora

$$\det(A) = \sum_{j=1}^{n} a_{1j} C_{1j} = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det(A_{\widehat{1}\widehat{j}})$$

si può facilmente scrivere nella forma prevista dal lemma, poiché per costruzione $A_{\widehat{1}\widehat{j}}$ ha ordine $n-1$ e non contiene elementi della riga 1 e della colonna j. $\square$

La verifica esplicita del lemma è già stata fatta per il caso $n = 2$ nell'Esempio 3.68. Vediamo il caso delle matrici di ordine 3 :

$$\begin{aligned}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} = \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.\end{aligned}$$

Proposizione 6.14. *Data $A \in \mathrm{Mat}(n, n; \mathbb{K})$, l'espressione $P_A(\lambda) = \det(A - \lambda \cdot \mathbb{I}_n)$ è un polinomio in $\mathbb{K}[\lambda]_n$, detto polinomio caratteristico di A. In particolare, i coefficienti di grado 0, $n-1$ ed n sono rispettivamente $c_0 = \det(A)$, $c_{n-1} = (-1)^{n-1}\mathrm{Tr}(A)$ e $c_n = (-1)^n$. Tutti i coefficienti e le radici di P_A sono invarianti per similitudine.*

Dimostrazione. Per prima cosa, riscriviamo esplicitamente il polinomio:

$$P_A(\lambda) = \det(A - \lambda \cdot \mathbb{I}_n) = \begin{vmatrix} a_{11} - \lambda & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} - \lambda \end{vmatrix}.$$

Grazie al Lemma 6.13, sappiamo che $\det(A - \lambda \cdot \mathbb{I}_n)$ è un polinomio i cui monomi sono tutti ed i soli prodotti di n elementi di $A - \lambda \cdot \mathbb{I}_n$, uno per ogni riga ed ogni colonna della matrice.

Se interpretiamo P_A come un polinomio nella sola variabile λ, allora i coefficienti sono per costruzione elementi di $\mathbb{K}$. In tal caso, il monomio di grado massimo di P_A lo otteniamo dal prodotto degli elementi lungo la diagonale di $A - \lambda \cdot \mathbb{I}_n$, ed ovviamente è $(-1)^n\lambda^n$. I monomi di $\det(A - \lambda \cdot \mathbb{I}_n)$ che contengono un elemento a_{ij} fuori dalla diagonale, hanno grado rispetto a λ minore di $n-1$. Infatti, per il Lemma 6.13 essi non possono contenere gli elementi $a_{ii} - \lambda$ e $a_{jj} - \lambda$, che stanno rispettivamente nella stessa riga e colonna di a_{ij}. Di conseguenza, anche il monomio di grado $n-1$ rispetto a λ si ottiene attraverso il prodotto degli elementi lungo la diagonale. Dovrebbe essere chiaro che esso è

$$(-\lambda)^{n-1}(a_{11} + \dots + a_{nn}) = (-1)^{n-1}\mathrm{Tr}(A)\lambda^{n-1}.$$

Il coefficiente di grado 0 rispetto a λ è invece uguale a $c_0 = P_A(0) = \det(A - 0 \cdot \mathbb{I}_n) = \det(A)$.

Infine, poiché il determinante di una matrice dipende solo dalla classe di similitudine, anche $P_A(\lambda)$ soddisfa tale proprietà. Infatti, dato $B = S^{-1} * A * S$, vale

$$\begin{aligned}P_B(\lambda) &= \det(B - \lambda \cdot \mathbb{I}_n) = \det(S^{-1} * A * S - \lambda \cdot S^{-1} * \mathbb{I}_n * S) = \\ &= \det(S^{-1} * (A - \lambda \cdot \mathbb{I}_n) * S) = \det(A - \lambda \cdot \mathbb{I}_n) = P_A(\lambda).\end{aligned}$$

Ovviamente questo implica che anche tutti i coefficienti e le radici del polinomio sono a loro volta invarianti. □

Esempio 6.15. Calcoliamo il polinomio caratteristico della generica matrice $A \in \mathrm{Mat}(n, n; \mathbb{K})$ per $n \leq 3$.

- $n = 1$: la matrice è $A = [a_{11}]$, di conseguenza $c_0 = \det(A) = \mathrm{Tr}(A) = a_{11}$, mentre $c_1 = -1$. Quindi
$$P_A(\lambda) = a_{11} - \lambda.$$
- $n = 2$: il polinomio caratteristico ha grado 2 ed i suoi coefficienti sono $c_0 = \det(A)$, $c_1 = -\mathrm{Tr}(A)$ e $c_2 = 1$. Quindi
$$P_A(\lambda) = \det(A) - \mathrm{Tr}(A)\lambda + \lambda^2.$$

- $n = 3$: il polinomio caratteristico ha grado 3 ed i suoi coefficienti sono $c_0 = \det(A),\ c_2 = \mathrm{Tr}(A)$ e $c_3 = -1$. Data la matrice
$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$
definiamo il cosiddetto invariante secondo
$$I_2(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}.$$
Lasciamo come esercizio al lettore verificare che $c_1 = -I_2(A)$. Quindi
$$P_A(\lambda) = \det(A) - I_2(A)\lambda + \mathrm{Tr}(A)\lambda^2 - \lambda^3.$$

6.3. Polinomio caratteristico e calcolo di autovalori ed autovettori

L'esempio fondamentale di matrici simili è quello riguardante le matrici rappresentative del medesimo endomorfismo. Infatti, dati $f \in \mathrm{End}(V)$ e due basi $\mathcal{B}, \mathcal{B}'$ di V, le matrici rappresentative rispetto alle due basi sono legate da
$$F|_{\mathcal{B}'} = M_{\mathcal{B}\mathcal{B}'} * F|_{\mathcal{B}} * M_{\mathcal{B}'\mathcal{B}} = M^{-1}_{\mathcal{B}'\mathcal{B}} * F|_{\mathcal{B}} * M_{\mathcal{B}'\mathcal{B}}.$$
Questo fatto e la Proposizione 6.12 ci consentono di definire la traccia, il determinante ed il polinomio caratteristico di un endomorfismo.

Definizione 6.16. *Sia V uno spazio vettoriale finitamente generato su campo $\mathbb{K}$ e sia $f \in \mathrm{End}(V)$. Sia inoltre $\mathcal{B}$ una base di V ed $F|_{\mathcal{B}}$ la matrice rappresentativa di f rispetto a tale base. Allora:*

i) la traccia dell'endomorfismo è la funzione
$$\begin{array}{rccc} \mathrm{Tr}: & \mathrm{End}(V) & \longrightarrow & \mathbb{K} \\ & f & \longmapsto & \mathrm{Tr}(F|_{\mathcal{B}}) \end{array};$$
ii) il determinante dell'endomorfismo è la funzione
$$\begin{array}{rccc} \det: & \mathrm{End}(V) & \longrightarrow & \mathbb{K} \\ & f & \longmapsto & \det(F|_{\mathcal{B}}) \end{array}.$$
iii) il polinomio caratteristico dell'endomorfismo è l'espressione
$$P_f(\lambda) = \det(f - \lambda \cdot \mathrm{Id}_V).$$

Ovviamente, molti dei risultati validi per le matrici quadrate sono a questo punto riformulabili in termini di endomorfismi. Ad esempio, un endomorfismo f è invertibile, e pertanto è un automorfismo, se e solo se $\det(f) = 0$. Se $F|_{\mathcal{B}}$ è la matrice rappresentativa di f rispetto alla base $\mathcal{B}$, il polinomio caratteristico dell'endomorfismo si calcola come
$$P_f(\lambda) = \det(F|_{\mathcal{B}} - \lambda \cdot \mathbb{I}_n).$$
Dalla Proposizione 6.14 segue immediatamente che $P_f(\lambda) \in \mathbb{K}[\lambda]_n$, ed i suoi coefficienti di grado 0, $n-1$ ed n sono rispettivamente $c_0 = \det(f),\ c_{n-1} = (-1)^{n-1}\mathrm{Tr}(f)$ e $c_n = (-1)^n$.

Torniamo ora al problema centrale di questo capitolo ed analizziamo il ruolo degli strumenti sviluppati finora nello studio della diagonalizzabilità di un endomorfismo. In primo luogo, da quanto osservato in precedenza sappiamo che due matrici rappresentanti la medesima applicazione sono simili. Vale anche il viceversa. Infatti, come abbiamo spiegato al termine della Sezione 5.5, ogni matrice invertibile S definisce un cambiamento di base. Questo significa che due matrici simili rappresentano il medesimo endomorfismo rispetto a due opportune basi. Questo ragionamento dimostra il seguente criterio di diagonalizzabilità di un'applicazione lineare.

Proposizione 6.17. *Siano V uno spazio vettoriale finitamente generato e $\mathcal{B}$ una sua base. Allora $f \in \mathrm{End}(V)$ è diagonalizzabile se e solo se la matrice rappresentativa $F|_{\mathcal{B}}$ è simile ad una matrice diagonale D, cioè se e solo se è possibile scrivere $D = S^{-1} * F|_{\mathcal{B}} * S$. In tal caso, S si dice matrice diagonalizzante di $F|_{\mathcal{B}}$.*

Finora è rimasta ancora aperta la questione di come calcolare gli autovalori e gli autovettori di un'applicazione data. Questo problema viene affrontato e risolto dal seguente teorema.

Teorema 6.18. *Sia V uno spazio vettoriale di dimensione finita su campo $\mathbb{K}$ e sia $f \in \mathrm{End}(V)$. Allora:*

i) $\lambda \in \mathbb{K}$ *è un autovalore di f se e solo se è radice del polinomio caratteristico, ovvero $P_f(\lambda) = 0$;*

ii) $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ *è un autovettore associato a λ se e solo se $\mathbf{v} \in \ker(f - \lambda \cdot \mathrm{Id}_V)$.*

Dimostrazione. Un vettore $\mathbf{v}$ non nullo è autovettore di f se e solo se $f(\mathbf{v}) = \lambda \cdot \mathbf{v} = \lambda \cdot \mathrm{Id}_V(\mathbf{v})$, ovvero $(f - \lambda \cdot \mathrm{Id}_V)(\mathbf{v}) = \mathbf{0}$. Questo è equivalente a richiedere $\mathbf{v} \in \ker(f - \lambda \cdot \mathrm{Id}_V) \setminus \{\mathbf{0}\}$. D'altra parte, questo ultimo insieme non è vuoto se e solo se l'applicazione $f - \lambda \cdot \mathrm{Id}_V$ non è iniettiva, ovvero se e solo se $f - \lambda \cdot \mathrm{Id}_V$ non è un isomorfismo e quindi

$$\det(f - \lambda \cdot \mathrm{Id}_V) = 0,$$

da cui la tesi. □

Esempio 6.19. Riprendiamo l'applicazione dell'Esempio 6.2:

$$A = \begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix} \qquad \Rightarrow \qquad f_A(\mathbf{v}) = A * \mathbf{v}.$$

Il polinomio caratteristico è

$$P_{f_A}(\lambda) = \det(f_A - \lambda \cdot \mathrm{Id}_V) = \det(f_A) - \mathrm{Tr}(f_A)\lambda + \lambda^2 = -1 + \lambda^2.$$

Di conseguenza, esistono due autovalori $\lambda_1 = 1$ e $\lambda_2 = -1$. Gli autospazi associati si ottengono calcolando il nucleo delle corrispondenti applicazioni:

$$V_1 = \ker(f_A - \lambda_1 \cdot \mathrm{Id}_V) = \ker(A - \mathbb{I}_2) = \ker\left(\begin{bmatrix} 1 & 3 \\ -1 & -3 \end{bmatrix}\right) = \ker\left(\begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}\right) = \mathcal{L}\left(\begin{bmatrix} 3 \\ -1 \end{bmatrix}\right),$$

$$V_{-1} = \ker(f_A - \lambda_2 \cdot \mathrm{Id}_V) = \ker(A + \mathbb{I}_2) = \ker\left(\begin{bmatrix} 3 & 3 \\ -1 & -1 \end{bmatrix}\right) = \ker\left(\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\right) = \mathcal{L}\left(\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right).$$

6.4. Criteri di diagonalizzabilità

In questa sezione forniremo dei criteri utili a decidere la diagonalizzabilità di un endomorfismo f, che a volte consentono di evitare il calcolo esplicito degli autospazi dell'applicazione. Preliminarmente, consigliamo al lettore il ripasso degli argomenti della Sezione 2.3.

Definizione 6.20. *Dato V uno spazio vettoriale finitamente generato sul campo $\mathbb{K}$, siano f un endomorfismo di V e λ un autovalore di f.*

i) La molteplicità algebrica $\mathrm{ma}(\lambda)$ *è la molteplicità dell'autovalore come radice del polinomio caratteristico.*

ii) La molteplicità geometrica $\mathrm{mg}(\lambda)$ *è la dimensione dell'autospazio V_λ.*

Definizione 6.21. *Dato V uno spazio vettoriale finitamente generato sul campo $\mathbb{K}$, siano f un endomorfismo di V e λ un autovalore di f.*

i) λ si dice regolare se $\mathrm{ma}(\lambda) = \mathrm{mg}(\lambda)$.

ii) λ si dice semplice se $\mathrm{ma}(\lambda) = 1$.

I due risultati principali che vogliamo dimostrare in questa sezione sono i seguenti. Il primo è noto come secondo criterio di diagonalizzabilità.

Teorema 6.22. *Sia V uno spazio vettoriale di dimensione $n < \infty$ sul campo $\mathbb{K}$ e sia $f \in \mathrm{End}(V)$. L'endomorfismo è diagonalizzabile se e solo se:*

i) il polinomio caratteristico $P_f(\lambda)$ ha esattamente n radici in $\mathbb{K}$, contate con molteplicità;

ii) ogni autovalore di f è regolare.

Ovviamente la prima condizione è automaticamente soddisfatta se $\mathbb{K}$ è algebricamente chiuso. Il secondo risultato è invece conosciuto come condizione sufficiente di diagonalizzabilità.

Teorema 6.23. *Sia V uno spazio vettoriale di dimensione $n < \infty$ sul campo $\mathbb{K}$ e sia $f \in \mathrm{End}(V)$. L'endomorfismo è diagonalizzabile se:*

i) il polinomio caratteristico $P_f(\lambda)$ ha esattamente n radici in $\mathbb{K}$, contate con molteplicità;

ii) ogni autovalore di f è semplice.

Esempio 6.24. La dimostrazione dei due teoremi precedenti è articolata e prevede la verifica di diversi risultati intermedi. Prima di intraprendere tale cammino, mostriamo l'utilità dei due enunciati nello studio della diagonalizzabilità del generico endomorfismo $f \in \mathrm{End}(V)$, dove V è uno spazio vettoriale reale o complesso di dimensione 2. In tal caso il polinomio caratteristico e le sue radici sono

$$P_f(\lambda) = \det(f) - \mathrm{Tr}(f)\,\lambda + \lambda^2 \qquad \Rightarrow \qquad \lambda_\pm = \frac{\mathrm{Tr}(f) \pm \sqrt{\mathrm{Tr}(f)^2 - 4\det(f)}}{2}.$$

Per il Teorema 6.23, abbiamo che

- su campo reale, f è diagonalizzabile se $\mathrm{Tr}(f)^2 > 4\det(f)$;
- su campo complesso, f è diagonalizzabile se $\mathrm{Tr}(f)^2 \neq 4\det(f)$.

Per il Teorema 6.22, f non è diagonalizzabile su campo reale se $\mathrm{Tr}(f)^2 < 4\det(f)$. La situazione con $\mathrm{Tr}(f)^2 = 4\det(f)$ è la più delicata. Infatti, in tal caso esiste un'unica radice $\bar{\lambda} = \frac{\mathrm{Tr}(f)}{2}$ del polinomio caratteristico, che sicuramente appartiene al campo su cui si sta lavorando e pertanto è un autovalore di f. D'altra parte, poiché $\mathrm{ma}(\bar{\lambda}) = 2$ l'autovalore non é semplice e potrebbe non essere nemmeno regolare. Per poter risolvere questa situazione abbiamo bisogno del seguente risultato.

Lemma 6.25. *Dato lo spazio vettoriale V di dimensione n su campo $\mathbb{K}$, sia $f \in \mathrm{End}(V)$ avente un unico autovalore λ di molteplicità algebrica n. Allora l'endomorfismo è diagonalizzabile se e solo se $f = \lambda \cdot \mathrm{Id}_V$.*

Dimostrazione. Per il secondo criterio di diagonalizzabilità, f è diagonalizzabile se e solo se $\mathrm{mg}(\lambda) = \mathrm{ma}(\lambda) = n = \dim(V)$, ovvero se e solo se $V_\lambda = V$. In tal caso, ogni vettore $\mathbf{v} \in V$ è un autovettore associato all'autovalore λ. Quindi $f(\mathbf{v}) = \lambda \cdot \mathbf{v}$ per ogni $\mathbf{v} \in V$, da cui segue la tesi. □

Possiamo quindi completare lo studio degli endomorfismi su spazi bidimensionali, affermando che se $\mathrm{Tr}(f)^2 = 4\det(f)$ l'endomorfismo è diagonalizzabile se e solo se $f = \bar{\lambda} \cdot \mathrm{Id}_V$.

Esempio 6.26. Applichiamo la classificazione precedente ad alcuni esempi di endomorfismi f_A associati alla matrice $A \in \mathrm{Mat}(2, 2; \mathbb{K})$.

- Consideriamo ancora una volta l'Esempio 6.2:
$$A = \begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix} \qquad \Rightarrow \qquad \mathrm{Tr}(A)^2 = 0 > -4 = 4\det(A)$$
da cui otteniamo che f_A è diagonalizzabile per entrambi $\mathbb{K} = \mathbb{R}, \mathbb{C}$.
- Prendiamo la matrice di rotazione dell'esempio Esempio 6.2:
$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad \Rightarrow \qquad \mathrm{Tr}(A)^2 = 0 < 4 = 4\det(A)$$
da cui otteniamo che f_A non è diagonalizzabile su campo reale ma lo è su campo complesso. Infatti, dal calcolo degli autospazi si può vedere che una base di $\mathrm{Mat}(2,1;\mathbb{C})$ composta da autovettori di $f_A \in \mathrm{End}(\mathrm{Mat}(2,1;\mathbb{C}))$ è
$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -i \end{bmatrix}, \begin{bmatrix} 1 \\ i \end{bmatrix} \right\}.$$
Gli autovalori associati sono $\lambda_\pm = \pm i$, entrambi semplici.
- Consideriamo infine l'applicazione definita dalla matrice
$$A = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \qquad \Rightarrow \qquad \mathrm{Tr}(A)^2 = 4 = 4\det(A).$$
Esiste un unico autovalore $\bar{\lambda} = 1$, di molteplicità algebrica 2. Pertanto, f_A è diagonalizzabile se e solo se $A = \bar{\lambda} \cdot \mathbb{I}_2 = \mathbb{I}_2$ e quindi se e solo se $a = 0$. Infatti, se $a \neq 0$ dallo studio dell'autospazio V_1 otteniamo:
$$V_1 = \ker(A - \mathbb{I}_2) = \ker\left(\begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix} \right) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \right).$$
Quindi la molteplicità geometrica di $\bar{\lambda}$ è 1 e non esiste una base di autovettori, indipendentemente dal campo in cui stiamo operando.

Torniamo ora alla dimostrazione dei Teoremi 6.22 e 6.23.

Lemma 6.27. *Siano V uno spazio finitamente generato ed f un endomorfismo di V. Se $\bar{\lambda}$ è autovalore di f, allora* $1 \leq \mathrm{mg}(\bar{\lambda}) \leq \mathrm{ma}(\bar{\lambda})$.

Dimostrazione. Dato $V_{\bar{\lambda}}$ l'autospazio associato all'autovalore, per prima cosa osserviamo che $\mathrm{mg}(\bar{\lambda}) = \dim(V_{\bar{\lambda}}) \geq 1$. Per dimostrare l'altra disuguaglianza, consideriamo una base $\mathcal{B}_{\bar{\lambda}} = \{\mathbf{v_1}, \dots, \mathbf{v_g}\}$ dell'autospazio, dove $g = \dim(V_{\bar{\lambda}})$. Utilizzando l'algoritmo di completamento, possiamo ottenere una base $\mathcal{B} = \mathcal{B}_{\bar{\lambda}} \cup \{\mathbf{v_{g+1}}, \dots, \mathbf{v_n}\}$ dell'intero spazio V. Rispetto a $\mathcal{B}$, la matrice rappresentativa di f è
$$\begin{aligned} F|_{\mathcal{B}} &= \begin{bmatrix} f(\mathbf{v_1})|_{\mathcal{B}} & \dots & f(\mathbf{v_g})|_{\mathcal{B}} & f(\mathbf{v_{g+1}})|_{\mathcal{B}} & \dots & f(\mathbf{v_n})|_{\mathcal{B}} \end{bmatrix} = \\ &= \begin{bmatrix} \bar{\lambda} \cdot \mathbf{v_1}|_{\mathcal{B}} & \dots & \bar{\lambda} \cdot \mathbf{v_g}|_{\mathcal{B}} & f(\mathbf{v_{g+1}})|_{\mathcal{B}} & \dots & f(\mathbf{v_n})|_{\mathcal{B}} \end{bmatrix} = \\ &= \left[\begin{array}{c|c} \bar{\lambda} \cdot \mathbb{I}_g & A \\ \hline 0_{n-g\,g} & B \end{array}\right]. \end{aligned}$$
Quindi il polinomio caratteristico è
$$P_f(\lambda) = \det(F|_{\mathcal{B}} - \lambda \cdot \mathbb{I}_n) = \left|\begin{array}{c|c} (\bar{\lambda} - \lambda) \cdot \mathbb{I}_g & A \\ \hline 0_{n-g\,g} & B - \lambda \cdot \mathbb{I}_{n-g} \end{array}\right| = (\bar{\lambda} - \lambda)^g \det(B - \lambda \cdot \mathbb{I}_{n-g}).$$
Di conseguenza $\bar{\lambda}$ è una radice di $P_f(\lambda)$ avente molteplicità maggiore od uguale a g, quindi $\mathrm{ma}(\bar{\lambda}) \geq g = \mathrm{mg}(\bar{\lambda})$. $\square$

Corollario 6.28. *Un autovalore semplice è regolare.*

Dimostrazione. Per definizione, un autovalore λ è semplice se $\mathrm{ma}(\lambda) = 1$. In tal caso, il Lemma 6.27 implica $\mathrm{mg}(\lambda) = 1$. $\square$

Grazie al corollario precedente deduciamo che la condizione sufficiente di diagonalizzabilità è conseguenza diretta del secondo criterio di diagonalizzabilità.

Rimane quindi da verificare la validità del Teorema 6.22.

Lemma 6.29. *Sia V uno spazio vettoriale ed f un endomorfismo di V. Un insieme $U_r = \{\mathbf{v_1}, \dots, \mathbf{v_r}\}$ di autovettori di f, associati ad r autovalori distinti, è linearmente indipendente.*

Dimostrazione. Verifichiamo l'enunciato per induzione su r. Il caso $r = 1$ è immediato, poiché per definizione $\mathbf{v_1} \neq \mathbf{0}$ e quindi U_1 è indipendente. Come ipotesi induttiva, supponiamo che il lemma sia vero per insiemi di autovettori di cardinalità minore ad r. Per assurdo, assumiamo ora che U_r sia linearmente dipendente. Allora esiste un vettore in U_r che può essere scritto come combinazione lineare degli altri. Senza perdere in generalità, supponiamo

$$\mathbf{v_r} = \sum_{i=1}^{r-1} t_i \cdot \mathbf{v_i}\,, \qquad t_i \in \mathbb{K}.$$

Se $\lambda_1, \dots, \lambda_r$ sono gli autovalori associati agli autovettori, vale quindi

$$f(\mathbf{v_r}) = \sum_{i=1}^{r-1} t_i \cdot f(\mathbf{v_i}) = \sum_{i=1}^{r-1} t_i \cdot (\lambda_i \cdot \mathbf{v_i}).$$

D'altra parte,

$$f(\mathbf{v_r}) = \lambda_r \cdot \mathbf{v_r} = \lambda_r \cdot \left(\sum_{i=1}^{r-1} t_i \cdot \mathbf{v_i} \right).$$

Sottraendo le due espressioni, otteniamo

$$\mathbf{0} = \sum_{i=1}^{r-1} t_i \cdot (\lambda_i \cdot \mathbf{v_i}) - \lambda_r \cdot \left(\sum_{i=1}^{r-1} t_i \cdot \mathbf{v_i} \right) = \sum_{i=1}^{r-1} t_i(\lambda_i - \lambda_r) \cdot \mathbf{v_i}.$$

A causa dell'ipotesi induttiva, deve valere $t_i(\lambda_i - \lambda_r) = 0$ per $1 \leq i < r$. Dato che gli autovalori sono distinti, segue che $t_i = 0$ per $1 \leq i < r$, ma questo è impossibile perché $\mathbf{v_r} \neq \mathbf{0}$. $\square$

Proposizione 6.30. *Dato la spazio vettoriale V finitamente generato sul campo $\mathbb{K}$, sia f un endomorfismo di V avente r autovalori distinti $\lambda_1, \dots, \lambda_r$ con autospazi associati $V_{\lambda_1}, \dots, V_{\lambda_r}$. Allora f è diagonalizzabile se e solo se $V = \bigoplus_{i=1}^{r} V_{\lambda_i}$.*

Dimostrazione. Se $r = 1$, l'enunciato è ovvio, pertanto supponiamo $r > 1$. Per il primo verso dell'implicazione, ipotizziamo che f sia diagonalizzabile e dimostriamo che lo spazio vettoriale è somma diretta degli autospazi, utilizzando la Proposizione 4.18. Chiaramente, per costruzione abbiamo

$$\sum_{i=1}^{r} V_{\lambda_i} \subseteq V.$$

Inoltre, per ipotesi esiste una base $\mathcal{B} = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ di V composta da autovettori di f. Essendo $\mathcal{B}$ un insieme di generatori di V, abbiamo

$$V = \mathcal{L}(\mathbf{v_1}, \dots, \mathbf{v_n}) \subseteq \sum_{i=1}^{r} V_{\lambda_i}\,.$$

La doppia inclusione implica $V = \sum_{i=1}^{r} V_{\lambda_r}$. Definiamo ora

$$V_{\widehat{\lambda_k}} = \sum_{\substack{i=1 \\ i \neq k}}^{r} V_{\lambda_i}, \qquad \text{per} \qquad 1 \leq k \leq r,$$

e consideriamo un vettore $\mathbf{u_k} \in V_{\lambda_k} \cap V_{\widehat{\lambda_k}}$. Allora esiste la decomposizione

$$\mathbf{u_k} = \sum_{\substack{i=1 \\ i \neq k}}^{r} \mathbf{u_i}\,,$$

dove $\mathbf{u_i} \in V_{\lambda_i}$. Di conseguenza, l'insieme $\{\mathbf{u_1}, \ldots, \mathbf{u_r}\}$ è linearmente dipendente. Affinché ciò sia compatibile con il Lemma 6.29, necessariamente deve valere $\mathbf{u_i} = \mathbf{0}$ per ogni $1 \leq i \leq r$, il che completa la prima parte della dimostrazione.

L'implicazione inversa è più semplice da verificare. Supponiamo che V sia somma diretta degli autospazi e siano $\mathcal{B}_1, \ldots, \mathcal{B}_r$ rispettivamente basi di $V_{\lambda_1}, \ldots, V_{\lambda_r}$. Grazie all'ipotesi, l'unione di tali basi è una base di V, che per costruzione è composta da autovettori di f. Quindi l'endomorfismo è diagonalizzabile. □

Corollario 6.31. *Su uno spazio V finitamente generato, l'endomorfismo f è diagonalizzabile se e solo se* $\sum_{i=1}^{r} \mathrm{mg}(\lambda_i) = \dim(V)$.

Dimostrazione. Nella dimostrazione della proposizione precedente abbiamo verificato che $V_{\lambda_k} \cap V_{\widehat{\lambda_k}} = \{\mathbf{0}\}$. Di conseguenza, l'enunciato è conseguenza del Corollario 4.52. □

Grazie ai risultati fino a qui ottenuti, possiamo finalmente completare la dimostrazione del secondo criterio di diagonalizzabilità.

Dimostrazione del Teorema 6.22. Siano $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ le radici distinte di $P_f(\lambda)$. Per il Lemma 6.27, vale $\mathrm{mg}(\lambda_i) \leq \mathrm{ma}(\lambda_i)$ per ogni $1 \leq i \leq r$, di conseguenza

$$\sum_{i=1}^{r} \mathrm{mg}(\lambda_i) \leq \sum_{i=1}^{r} \mathrm{ma}(\lambda_i) \leq n = \dim(V).$$

Per il Corollario 6.31, l'applicazione f è diagonalizzabile se e solo se vale l'uguaglianza. A sua volta, questo è possibile se e solo se:

i) $\sum_{i=1}^{r} \mathrm{ma}(\lambda_i) = n$, cioè il polinomio caratteristico ha esattamente n radici in $\mathbb{K}$, contate con moltiplicità;
ii) $\mathrm{mg}(\lambda_i) = \mathrm{ma}(\lambda_i)$ per ogni $1 \leq i \leq r$, ovvero se e solo se tutti gli autovalori di f sono regolari. □

Osservazione 6.32. La Proposizione 6.30 comporta anche altre interessanti conseguenze. In primo luogo, se f è un endomorfismo diagonalizzabile di V, allora per ogni vettore $\mathbf{v} \in V$ esiste un'unica decomposizione composta da autovettori relativi agli autovalori distinti $\lambda_1, \ldots, \lambda_r$:

$$\mathbf{v} = \sum_{i=1}^{r} \mathbf{v_i} \qquad \text{con } \mathbf{v_i} \in V_{\lambda_i}.$$

Questa decomposizione di $\mathbf{v}$ chiaramente dipende dall'applicazione f che stiamo considerando. D'altra parte, l'endomorfismo f ammette invece una decomposizione caratteristica legata ai suoi autovalori, che prende il nome di *decomposizione spettrale.* Infatti, definiamo le proiezioni

$$\begin{array}{rccc} \mathcal{P}_i : & V & \longrightarrow & V \\ & \mathbf{v} & \longmapsto & \mathbf{v_i} \end{array},$$

dove $\mathbf{v_i} \in V_{\lambda_i}$ è l'autovettore associato a λ_i che compare nella decomposizione di $\mathbf{v}$. Allora vale

$$f = \sum_{i=1}^{r} \lambda_i \cdot \mathcal{P}_i \,.$$

Per dimostrarlo, basta eseguire la verifica diretta:

$$f(\mathbf{v}) = f\left(\sum_{i=1}^{r} \mathbf{v_i}\right) = \sum_{i=1}^{r} f(\mathbf{v_i}) = \sum_{i=1}^{r} \lambda_i \cdot \mathbf{v_i} = \sum_{i=1}^{r} \lambda_i \cdot \mathcal{P}_i(\mathbf{v}).$$

Infine, lasciamo come esercizio al lettore la verifica dell'implicazione opposta a quella appena dimostrata, ovvero che la decomposizione spettrale di un endomorfismo esiste solo se l'endomorfismo è diagonalizzabile.

6.5. Complementi

Includiamo in questa sezione conclusiva alcuni risultati interessanti sugli argomenti affrontati in questo capitolo.

6.5.1. Algebra degli autospazi. Supponiamo che V sia uno spazio vettoriale ed f un endomorfismo di V.

Lemma 6.33. *Se V è finitamente generato, allora f è un automorfismo se e solo se è suriettivo, oppure se e solo se è iniettivo.*

Dimostrazione. L'enunciato segue facilmente dal Teorema di nullità più rango e dal Corollario 5.31. $\square$

Proposizione 6.34. *Sia λ autovalore di f con autospazio associato V_λ. Allora:*

i) $\lambda = 0$ *se e solo se* $\ker(f) \neq \{\mathbf{0}\}$, *ed in tal caso* $V_0 = \ker(f)$;

ii) se V è finitamente generato, f è invertibile se e solo se ogni autovalore soddisfa $\lambda \neq 0$. In tal caso λ^{-1} è autovalore di f^{-1} con autospazio associato $V_{\lambda^{-1}} = V_\lambda$. Inoltre, f è diagonalizzabile se e solo se f^{-1} è diagonalizzabile.

Dimostrazione. Per definizione, $\mathbf{v} \in V_0$ se e solo se $f(\mathbf{v}) = 0 \cdot \mathbf{v} = \mathbf{0}$, ovvero se e solo se $\mathbf{v} \in \ker(f)$, da cui segue il primo punto dell'enunciato. Concentriamoci sul secondo punto. Grazie al Lemma 6.33, possiamo osservare che f è invertibile se e solo se è iniettiva. Per il punto (i), questo è vero se e solo se tutti gli autovalori λ di f sono diversi da 0. In tal caso esistono $f^{-1} \in \mathrm{End}(V)$ e $\lambda^{-1} \in \mathbb{K}^*$. Supponiamo ora che $\mathbf{v} \in V_\lambda$. Allora

$$f^{-1}(\mathbf{v}) = f^{-1}\left(\frac{1}{\lambda} \cdot f(\mathbf{v})\right) = \frac{1}{\lambda} \cdot f^{-1} \circ f(\mathbf{v}) = \frac{1}{\lambda} \cdot \mathbf{v},$$

da cui deduciamo che λ^{-1} è autovalore di f^{-1} e $V_\lambda \subseteq V_{\lambda^{-1}}$. Le implicazioni e l'inclusione inversa si ottengono osservando che nel ragionamento precedente possiamo scambiare il ruolo di f ed f^{-1}. Infine, per la Proposizione 6.30 sappiamo che f è diagonalizzabile se e solo se lo spazio V è somma diretta degli autospazi di f. Ma per quanto appena dimostrato, questo è vero se e solo se V è somma diretta degli autospazi di f^{-1}. $\square$

Proposizione 6.35. *Sia λ autovalore di f con autospazio associato V_λ. Allora:*

i) *l'applicazione* $t \cdot f \in \mathrm{End}(V)$ *ammette* $t\lambda$ *tra i suoi autovalori. Inoltre, se* $t \neq 0$, *l'autospazio associato è* $V_{t\lambda} = V_\lambda$;

ii) *l'applicazione* $f^m \in \mathrm{End}(V)$ *ammette* λ^m *tra i suoi autovalori e l'autospazio associato* V_{λ^m} *contiene* V_λ;

iii) *l'applicazione* $Q(f) \in \mathrm{End}(V)$ *ammette* $Q(\lambda)$ *tra i suoi autovalori e l'autospazio associato* $V_{Q(\lambda)}$ *contiene* V_λ.

Infine, se f *è diagonalizzabile allora lo sono anche tutte le applicazioni introdotte nei punti precedenti.*

Dimostrazione.

i) Se $t = 0$, l'enunciato è ovvio ed in tal caso $V_0 = V$. Se $t \neq 0$, allora $\mathbf{v}$ è un autovettore di f associato a λ se e solo se

$$(t \cdot f)(\mathbf{v}) = t \cdot f(\mathbf{v}) = t \cdot (\lambda \cdot \mathbf{v}) = t\lambda \cdot \mathbf{v},$$

ovvero se e solo $\mathbf{v}$ è autovettore di $t \cdot f$ associato a $t\lambda$.

ii) Utilizziamo l'induzione su m. Il caso $m = 1$ è vero per ipotesi. Supponiamo che l'enunciato valga fino ad $m - 1$. Allora se $\mathbf{v} \in V_\lambda$, segue che

$$f^m(\mathbf{v}) = f^{m-1}(f(\mathbf{v})) = f^{m-1}(\lambda \cdot \mathbf{v}) = \lambda \cdot f^{m-1}(\mathbf{v}) = \lambda \cdot (\lambda^{m-1} \cdot \mathbf{v}) = \lambda^m \cdot \mathbf{v},$$

dove nella penultima uguaglianza abbiamo utilizzato l'ipotesi induttiva. Quindi λ^m è autovalore di f^m e $V_\lambda \subseteq V_{\lambda^m}$.

iii) Il terzo punto si dimostra in maniera analoga al precedente.

Infine, se f è diagonalizzabile allora V è somma degli autospazi di f. Per le inclusioni sopra dimostrate, allora V è anche somma degli autospazi di ogni singola applicazione $t \cdot f$, f^m e $Q(f)$. Grazie al Lemma 6.29 ed alla Proposizione 6.30, questo è sufficiente a dimostrare la diagonalizzabilità di tali endomorfismi. $\square$

Esempio 6.36. Prendiamo la matrice

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

e l'applicazione associata f_A. Allora $P_A(\lambda) = \lambda^2$, da cui deduciamo che esiste un unico autovalore $\lambda = 0$. L'autospazio V_0 coincide con $\ker(f_A)$ ed è uguale a

$$V_0 = \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right).$$

È semplice verificare che f_A^2 è l'applicazione nulla, da cui $V_0 = V$. Come si può osservare, in questo caso vale l'inclusione stretta $V_\lambda \subset V_{\lambda^2}$.

6.5.2. Criterio di similitudine. Sappiamo che due matrici simili hanno uguali invarianti di similitudine. Nel verso opposto vale il seguente risultato.

Proposizione 6.37. *Siano* $A, B \in \mathrm{Mat}(n, n; \mathbb{K})$, *con* A *diagonalizzabile. Allora le due matrici sono simili se e solo se anche* B *è diagonalizzabile e gli autovalori delle due matrici, comprese le relative moltiplicità, sono uguali.*

Dimostrazione. Per ipotesi, A è simile ad una matrice diagonale D, sulla cui diagonale compaiono gli autovalori $\lambda_1, \dots, \lambda_r$ di A in numero pari alle rispettive molteplicità. Poiché la similitudine è una relazione di equivalenza, allora B è simile ad A se e solo se B è simile a D. Questo è vero se e solo se vale la condizione richiesta nell'enunciato. $\square$

Esempio 6.38. Studiamo la similitudine delle due matrici

$$A_h = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 1 & h^2 \\ -1 & 1 & 1 \end{bmatrix} \qquad \text{e} \qquad B_h = \begin{bmatrix} 1+h & 0 & 0 \\ 0 & 1-h & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

dipendenti dal parametro $h \in \mathbb{R}$. La matrice A_h ha polinomio caratteristico

$$P_h(\lambda) = \begin{vmatrix} 2-\lambda & 0 & 0 \\ 1 & 1-\lambda & h^2 \\ -1 & 1 & 1-\lambda \end{vmatrix} = (2-\lambda)((1-\lambda)^2 - h^2) = (2-h)(1-\lambda-h)(1-\lambda+h).$$

Gli autovalori sono $\lambda_1 = 2$, $\lambda_2 = 1-h$ e $\lambda_3 = 1+h$. La matrice B_h è diagonale con autovalori $\mu_1 = 1+h$, $\mu_2 = 1-h$ e $\mu_3 = 2$. Quindi $\lambda_1 = \mu_3$, $\lambda_2 = \mu_2$ e $\lambda_3 = \mu_1$ per ogni valore di h. Pertanto le due matrici sono simili per tutti ed i soli valori di h per cui A_h risulti diagonalizzabile.

- Se $h \neq 0, \pm 1$, i tre autovalori sono semplici, quindi A_h è diagonalizzabile.
- $h = -1$ abbiamo $\lambda_1 = \lambda_2 = 2$, mentre se $h = 1$ vale $\lambda_1 = \lambda_3 = 2$. In entrambi i casi, l'autospazio associato all'autovalore 2 è

$$V_2 = \ker \left(\begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix} \right).$$

  Dallo studio delle righe si deduce che la matrice ha rango 1, di conseguenza la molteplicità geometrica dell'autovalore è 2 e perciò l'autovalore è regolare. Il rimanente autovalore 0 è semplice e di conseguenza regolare. Quindi le matrici $A_{\pm 1}$ sono diagonalizzabili.
- Infine, se $h = 0$ vale $\lambda_2 = \lambda_3 = 1$. L'autospazio associato a tale autovalore è

$$V_1 = \ker \left(\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ -1 & 1 & 0 \end{bmatrix} \right).$$

  Dallo studio delle spazio delle colonne si deduce facilmente che la matrice ha rango 2, di conseguenza la molteplicità geometrica dell'autovalore è 1 e quindi l'autovalore non è regolare. Pertanto la matrice A_0 non è diagonalizzabile.

6.5.3. Teorema di Cayley–Hamilton. Enunciamo e dimostriamo il teorema di Cayley–Hamilton per le matrici quadrate.

Teorema 6.39. *Ogni matrice* $A \in \mathrm{Mat}(n, n; \mathbb{K})$ *è radice del suo polinomio caratteristico, ovvero* $P_A(A) = 0_n$.

Dimostrazione. Il polinomio caratteristico di A è

$$P_A(\lambda) = \det(A - \lambda \cdot \mathbb{I}_n) = \sum_{i=0}^{n} c_i \, \lambda^i.$$

Per il teorema di Laplace, abbiamo

$$\det(A - \lambda \cdot \mathbb{I}_n) \cdot \mathbb{I}_n = (A - \lambda \cdot \mathbb{I}_n) * (A - \lambda \cdot \mathbb{I}_n)^*.$$

Grazie alla definizione di matrice aggiunta ed al Lemma 6.13, sappiamo che ogni elemento di $(A - \lambda \cdot \mathbb{I}_n)^*$ è un polinomio in λ di grado al più $n-1$. Pertanto esiste la decomposizione

$$(A - \lambda \cdot \mathbb{I}_n)^* = \sum_{i=0}^{n-1} \lambda^i \cdot B_i,$$

per opportune matrice $B_i \in \mathrm{Mat}(n,n;\mathbb{K})$. Quindi vale

$$\begin{aligned}(A-\lambda\cdot\mathbb{I}_n)*(A-\lambda\cdot\mathbb{I}_n)^* &= (A-\lambda\cdot\mathbb{I}_n)*\sum_{i=0}^{n-1}\lambda^i\cdot B_i\\ &= A*B_0+\sum_{i=1}^{n-1}\lambda^i\cdot(A*B_i-B_{i-1})-\lambda^n\cdot B_{n-1}\\ &= \det(A-\lambda\cdot\mathbb{I}_n)\cdot\mathbb{I}_n=\sum_{i=0}^{n}c_i\,\lambda^i\cdot\mathbb{I}_n.\end{aligned}$$

Da questa identità possiamo ricavare i coefficienti del polinomio caratteristico:

$$\begin{cases} c_0\cdot\mathbb{I}_n = A*B_0,\\ c_i\cdot\mathbb{I}_n = A*B_i-B_{i-1} \quad \text{se} \quad 1\le i\le n-1,\\ c_n\cdot\mathbb{I}_n=-B_{n-1}.\end{cases}$$

Pertanto

$$P_A(A)=\sum_{i=0}^{n}c_i\cdot A^i=\sum_{i=0}^{n}A^i*(c_i\cdot\mathbb{I}_n)=A*B_0+\sum_{i=1}^{n-1}(A^{i+1}*B_i-A^i*B_{i-1})-A^n*B_{n-1}=0_n,$$

dove nell'ultima uguaglianza abbiamo utilizzato le proprietà delle sommatorie telescopiche. □

Corollario 6.40. *Data* $A\in\mathrm{Mat}(n,n;\mathbb{K})$ *con polinomio caratteristico* $P_A(\lambda)=\sum_{i=0}^{n}c_i\,\lambda^i$, *allora:*

i) l'insieme $\{\mathbb{I}_n, A,\dots,A^n\}$ *è linearmente dipendente;*

ii) se A *è invertibile, allora*

$$A^{-1}=-\frac{1}{\det(A)}\cdot\left(\sum_{i=1}^{n}c_i\cdot A^{i-1}\right).$$

Dimostrazione. Il primo punto è conseguenza diretta del teorema di Cayley-Hamilton. Per il secondo, osserviamo che

$$P_A(A)=\sum_{i=0}^{n}c_i\cdot A^i=c_0\cdot\mathbb{I}_n+A*\sum_{i=1}^{n}c_i\cdot A^{i-1}=0.$$

Tenendo conto che $c_0=\det(A)\neq 0$, dall'equazione precedente otteniamo

$$-\det(A)\cdot\mathbb{I}_n=A*\sum_{i=1}^{n}c_i\cdot A^{i-1},$$

da cui la tesi. □

Esempio 6.41. Calcoliamo esplicitamente l'inversa di una matrice quadrata A di ordine 2 utilizzando il Corollario 6.40. Data

$$A=\begin{bmatrix} a & b\\ c & d\end{bmatrix}\in\mathrm{Mat}(2,2;\mathbb{K}),$$

allora

$$A^{-1}=\frac{1}{\det(A)}\cdot(\mathrm{Tr}(A)\cdot\mathbb{I}_n-A)=\frac{1}{ad+bc}\cdot\left(\begin{bmatrix} a+d & 0\\ 0 & a+d\end{bmatrix}-\begin{bmatrix} a & b\\ c & d\end{bmatrix}\right)=\frac{1}{ad+bc}\cdot\begin{bmatrix} d & -b\\ -c & a\end{bmatrix}.$$

Ovviamente, sfruttando il teorema di rappresentazione e tutte le proprietà relative al polinomio caratteristico di un'applicazione lineare, si dimostra facilmente l'analogo del Teorema 6.39 per gli endomorfismi.

Corollario 6.42. *Siano* V *uno spazio finitamente generato ed* f *un endomorfismo. Allora* $P_f(f)\in\mathrm{End}(V)$ *è l'applicazione nulla.*

CAPITOLO 7

Geometria affine

In questo capitolo utilizzeremo gli strumenti algebrici introdotti nei capitoli precedenti per affrontare in maniera rigorosa lo studio di problemi di natura geometrica. Ciò che faremo sarà introdurre il concetto di spazio affine, per poi descriverne le sue proprietà. Anche in questa situazione continueremo a seguire l'approccio assiomatico fino ad ora utilizzato, ma cercando di fornire un'interpretazione geometrica dei ragionamenti che svilupperemo. In questo modo otterremo gli strumenti necessari per investigare la geometria di oggetti in dimensione qualsiasi, mantenendo però il contatto con una loro descrizione intuitiva. Al termine del capitolo, faremo infine una breve discussione sul legame tra l'assiomatizzazione della geometria che qui presentiamo e gli assiomi classici della geometria euclidea.

7.1. Spazi affini e sistemi di riferimento

Prima di affrontare questo corso, il lettore era probabilmente abituato ad un'idea differente di vettore rispetto a quella assiomatica sviluppata nel Capitolo 4. In fisica, ed in particolare nello studio della cinematica, i vettori sono entità geometriche che servono a descrivere la posizione dei corpi in funzione del tempo. In tale contesto, un vettore è intuitivamente definito come un oggetto rappresentante lo spostamento da un punto di partenza P ad un punto di arrivo Q. Attraverso la definizione di spazio affine vogliamo recuperare questa interpretazione, basandola su un approccio assiomatico rigoroso. Esistono diverse definizioni equivalenti di spazio affine, noi adotteremo quella basata sui cosiddetti assiomi di Weyl.

Definizione 7.1. *Consideriamo un insieme non vuoto $\mathcal{A}$, uno spazio vettoriale V ed una funzione*

$$\begin{array}{rccc} \psi: & \mathcal{A}\times\mathcal{A} & \longrightarrow & V \\ & (P,Q) & \longmapsto & \overrightarrow{PQ} \end{array}.$$

La struttura $\mathbb{A}=(\mathcal{A},V,\psi)$ è uno spazio affine se:

i) per ogni $P\in\mathcal{A}$ fissato, la funzione

$$\begin{array}{rccc} \psi_P: & \{P\}\times\mathcal{A} & \longrightarrow & V \\ & (P,Q) & \longmapsto & \overrightarrow{PQ} \end{array}$$

è biunivoca;

ii) vale la regola del parallelogramma, o di Chasles

$$\psi(P,Q)+\psi(Q,R)=\psi(P,R) \quad \text{per ogni } P,Q,R\in\mathcal{A}.$$

Definizione 7.2. *Dato $\mathbb{A}=(\mathcal{A},V,\psi)$, gli elementi $P\in\mathcal{A}$ si dicono punti, la coppia $(P,Q)\in\mathcal{A}\times\mathcal{A}$ si dice segmento orientato da P a Q ed il vettore $\overrightarrow{PQ}\in V$ si dice vettore geometrico da P a Q. L'insieme $\mathcal{A}$ e lo spazio vettoriale V sono detti rispettivamente sostegno e giacitura di $\mathbb{A}$. La dimensione dello spazio affine è definita come* $\dim(\mathbb{A})=\dim(V)$*. Lo spazio si dice retta affine se* $\dim(\mathbb{A})=1$ *e piano affine se* $\dim(\mathbb{A})=2$.

Analogamente agli spazi vettoriali, in assenza di ambiguità si è soliti identificare uno spazio affine $\mathbb{A}$ con il suo insieme di supporto $\mathcal{A}$. Il lettore deve però prestare attenzione al fatto che questo è un abuso di notazione, in quanto sul medesimo insieme è possibile definire differenti strutture affini.

Prima di fornire alcuni esempi di spazi affini, cerchiamo di interpretare il significato della Definizione 7.1. La funzione ψ intuitivamente associa ad ogni coppia di punti P, Q lo spostamento o traslazione tra gli stessi, rappresentato dal vettore $\overrightarrow{PQ}$. In questa interpretazione, la prima richiesta riguardante ψ è equivalente ad affermare che è possibile partire da un punto P e raggiungere tutti gli altri punto dello spazio attraverso tutte le possibili traslazioni, e viceversa ogni possibile traslazione è univocamente determinata dal punto di partenza P e dal punto di arrivo Q dello spostamento. Infine, la seconda richiesta combina in modo naturale le traslazioni: se ci muoviamo prima da P a Q e successivamente da Q ad R, questo è equivalente ad andare direttamente da P ad R.

Esempio 7.3.

- Sia V uno spazio vettoriale e ψ la funzione

  $$\begin{array}{rccc} \psi : & V \times V & \longrightarrow & V \\ & (\mathbf{v_1}, \mathbf{v_2}) & \longmapsto & \mathbf{v_2} - \mathbf{v_1} \end{array}.$$

  È evidente che $\psi_{\mathbf{v_1}}$ è biunivoca. Infatti, per ogni $\mathbf{w} \in V$ esiste un unico vettore $\mathbf{v_2} \in V$ tale che $\mathbf{w} = \psi_{\mathbf{v_1}}(\mathbf{v_2}) = \mathbf{v_2} - \mathbf{v_1}$, ovvero $\mathbf{v_2} = \mathbf{v_1} + \mathbf{w}$. Inoltre

  $$\psi(\mathbf{v_1}, \mathbf{v_2}) + \psi(\mathbf{v_2}, \mathbf{v_3}) = (\mathbf{v_2} - \mathbf{v_1}) + (\mathbf{v_3} - \mathbf{v_2}) = \mathbf{v_3} - \mathbf{v_1} = \psi(\mathbf{v_1}, \mathbf{v_3})$$

  per ogni $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3} \in V$. Quindi (V, V, ψ) è uno spazio affine ed è indicato con $\mathbb{A}_V$. Lo spazio affine associato allo spazio vettoriale $(\mathbb{K}^n, \mathbb{K}, +, \cdot)$ viene indicato con $\mathbb{A}^n_{\mathbb{K}}$.
- Consideriamo l'insieme

  $$\mathcal{A} = \{(x, y) \in \mathbb{K}^2 \mid x^2 - y = 0\}.$$

  Geometricamente, questa è una parabola di vertice $(0, 0)$, i cui punti hanno coordinate (x, x^2). La funzione

  $$\begin{array}{rccc} \psi : & \mathcal{A} \times \mathcal{A} & \longrightarrow & \mathbb{K} \\ & ((x_1, x_1^2), (x_2, x_2^2)) & \longmapsto & x_2 - x_1 \end{array}$$

  definisce una struttura di spazio affine $\mathbb{A} = (\mathcal{A}, \mathbb{K}, \psi)$. Infatti, fissati $P = (x_1, x_1^2) \in \mathbb{A}$ e $t \in \mathbb{K}$, vale

  $$\psi_P^{-1}(t) = (x_1 + t, (x_1 + t)^2).$$

  Inoltre è facile verificare la regola del parallelogramma. La dimensione di $\mathbb{A}$ è uno, quindi rispetto alla struttura definita da ψ abbiamo una retta affine. Questo esempio può essere generalizzato in modo naturale per ogni insieme $\mathcal{A}$ che possa essere visto come il grafico di una funzione f definita su uno spazio vettoriale V a valori nell'insieme A, cioè

  $$\mathcal{A} = \{(\mathbf{x}, y) \in V \times A \mid y = f(\mathbf{x})\}.$$

Investighiamo ora alcune importanti proprietà degli spazi affini, che seguono immediatamente dalla loro definizione. Iniziamo elencando un paio di proprietà elementari.

Proposizione 7.4. *Sia $\mathbb{A} = (\mathcal{A}, V, \psi)$ uno spazio affine. Allora:*

i) per ogni punto $P \in \mathcal{A}$ vale $\overrightarrow{PP} = \mathbf{0}$;
ii) per ogni segmento orientato $(P, Q) \in \mathcal{A} \times \mathcal{A}$ vale $-\overrightarrow{PQ} = \overrightarrow{QP}$.

Dimostrazione. La prima proprietà segue dalla regola del parallelogramma e dall'unicità dell'elemento neutro di V. Infatti $\overrightarrow{PP} + \overrightarrow{PQ} = \overrightarrow{PQ}$, da cui $\overrightarrow{PP} = \mathbf{0}$. A questo punto è facile provare la seconda affermazione, poiché $\overrightarrow{PQ} + \overrightarrow{QP} = \overrightarrow{PP} = \mathbf{0}$, da cui la tesi. □

Scelto un punto arbitrario $O \in \mathcal{A}$, la richiesta che la funzione ψ_O sia una corrispondenza biunivoca tra i punti di $\mathcal{A}$ ed i vettori di V è di fondamentale importanza per lo studio di $\mathbb{A}$. Infatti, una volta scelta una base $\mathcal{B}$ di V, possiamo comporre ψ_O con la mappa delle componenti $\phi_{\mathcal{B}}$, ottenendo quindi una corrispondenza biunivoca tra $\mathcal{A}$ e lo spazio delle matrici colonna $\mathrm{Mat}(n, 1; \mathbb{K})$, dove $n = \dim(\mathbb{A})$ e $\mathbb{K}$ è il campo su cui è definito V. In tal modo possiamo assegnare ad ogni punto di $\mathcal{A}$ un'unica n–upla in $\mathbb{K}^n$, e viceversa. Arriviamo quindi alla definizione di sistema di riferimento e di mappa delle coordinate.

Definizione 7.5. *Un sistema di riferimento di uno spazio affine $\mathbb{A} = (\mathcal{A}, V, \psi)$ è un sottoinsieme ordinato $\mathcal{B}_O$ di $\mathcal{A} \cup V$ composto da un punto $O \in \mathcal{A}$ e dai vettori di una base $\mathcal{B}$ di V. La funzione*

$$\begin{array}{rccc} \phi_{\mathcal{B}_O}: & \mathcal{A} & \longrightarrow & \mathrm{Mat}(n, 1; \mathbb{K}) \\ & P & \longmapsto & P|_{\mathcal{B}_O} = (\phi_{\mathcal{B}} \circ \psi_O)(P) \end{array}$$

è detta mappa delle coordinate e gli elementi della matrice $(\phi_{\mathcal{B}} \circ \psi_O)(P)$ si chiamano coordinate di P rispetto a $\mathcal{B}_O$. Il punto O è detto origine del sistema di riferimento.

Esempio 7.6. Dato $V = \mathbb{K}[x]_2$, consideriamo lo spazio affine $\mathbb{A}_V$ ed il sistema di riferimento $\mathcal{B}_{Q(x)}$ in cui l'origine è il polinomio $Q(x) = a_0 + a_1 x + a_2 x^2$ e la base di V è quella canonica $\mathcal{B} = \{1, x, x^2\}$. Allora le coordinate di $P(x) = b_0 + b_1 x + b_2 x^2 \in \mathbb{A}_V$ sono

$$\begin{aligned} P(x)|_{\mathcal{B}_{Q(x)}} &= \phi_{\mathcal{B}}(\psi_{Q(x)}(P(x))) = \phi_{\mathcal{B}}(P(x) - Q(x)) = \phi_{\mathcal{B}}(P(x)) - \phi_{\mathcal{B}}(Q(x)) = \\ &= \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} - \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} b_0 - a_0 \\ b_1 - a_1 \\ b_2 - a_2 \end{bmatrix}. \end{aligned}$$

Osserviamo che le coordinate di $P(x)$ rispetto a $\mathcal{B}_{Q(x)}$ coincidono con le componenti del polinomio rispetto a $\mathcal{B}$ se e solo se $Q(x)$ è il polinomio nullo.

Osservazione 7.7. I risultati dell'esempio precedente possono essere generalizzati su qualunque spazio affine. Sia $\mathcal{B}_O$ un sistema di riferimento in $\mathbb{A}$, allora le coordinate del generico punto $P \in \mathbb{A}$ soddisfano la seguente relazione fondamentale:

$$P|_{\mathcal{B}_O} = \phi_{\mathcal{B}}(\psi_O(P)) = \phi_{\mathcal{B}}\left(\overrightarrow{OP}\right) = \overrightarrow{OP}|_{\mathcal{B}}.$$

Da questa possiamo ricavare che:

$$\overrightarrow{PQ}|_{\mathcal{B}} = \left(\overrightarrow{OQ} - \overrightarrow{OP}\right)|_{\mathcal{B}} = \overrightarrow{OQ}|_{\mathcal{B}} - \overrightarrow{OP}|_{\mathcal{B}} = Q|_{\mathcal{B}_O} - P|_{\mathcal{B}_O}.$$

Nel caso di uno spazio affine $\mathbb{A}_V$, se la giacitura V ha una base canonica $\mathcal{B}$, allora chiamiamo sistema di riferimento canonico quello composto da $\mathcal{B}_{\mathbf{0}} = \{\mathbf{0_V}\} \cup \mathcal{B}$. Rispetto a questo sistema, per ogni $\mathbf{v} \in \mathbb{A}_V$ vale

$$\mathbf{v}|_{\mathcal{B}_{\mathbf{0}}} = (\mathbf{v} - \mathbf{0})|_{\mathcal{B}} = \mathbf{v}.$$

Per i sistemi di riferimento affini esiste una caratterizzazione simile a quella relativa alle basi di spazi vettoriali che è stata fornita nel Teorema 4.32.

Definizione 7.8. *Sia $S = \{P_0, \ldots, P_n\}$ un insieme di $n+1$ punti dello spazio affine $\mathbb{A}$ ed $U = \{\overrightarrow{P_0P_1}, \ldots, \overrightarrow{P_0P_n}\}$ il corrispondente sottoinsieme ordinato di vettori appartenenti alla giacitura V. Allora:*

i) S è affinamente indipendente se l'insieme U è linearmente indipendente;

ii) S è un insieme di generatori di $\mathbb{A}$ se U è un insieme generatore di V. In tal caso si scrive $\mathbb{A} = Span(P_0, \ldots, P_n) = \mathcal{L}(P_0, \ldots, P_n)$.

Anche se potrebbe apparire il contrario, il punto P_0 non è privilegiato all'interno dell'insieme S. Infatti, la definizione precedente non dipende dalla scelta del punto P_0. Ad esempio, consideriamo l'insieme $U' = \{\overrightarrow{P_1P_0}, \overrightarrow{P_1P_2}, \ldots, \overrightarrow{P_1P_n}\}$, in cui è stato scelto il punto P_1 a svolgere il ruolo che P_0 ha in U. Ogni vettore di U è scrivibile come $\overrightarrow{P_0P_i} = -\overrightarrow{P_1P_0} + \overrightarrow{P_1P_i},\ i = 1, \ldots, n$. Allora non è difficile verificare che U è indipendente, oppure è un insieme di generatori, se e solo se anche U' lo è. La seguente proposizione è conseguenza immediata della definizione di sistema di riferimento e del Teorema 4.32.

Proposizione 7.9. *Dato il punto O nello spazio affine $\mathbb{A}$ ed i vettori $\mathbf{v_1}, \ldots, \mathbf{v_n}$ nella giacitura V, siano $P_1, \ldots, P_n$ gli unici punti di $\mathbb{A}$ tali che $\mathbf{v_i} = \overrightarrow{OP_i}$, per ogni $i = 1, \ldots, n$. Allora l'insieme ordinato $\mathcal{B}_O = \{O, \mathbf{v_1}, \ldots, \mathbf{v_n}\}$ è un sistema di riferimento se e solo se $\{O, P_1, \ldots, P_n\}$ è un'insieme affinamente indipendente composto da generatori di $\mathbb{A}$.*

Osserviamo infine che per i sistemi di riferimento valgono le proposizione analoghe al Teorema 4.40 ed al Corollario 4.43.

Esempio 7.10. Nel piano affine $\mathbb{A}^2_{\mathbb{R}}$, consideriamo i punti $P_0 = (0,0),\ P_1 = (1,0),\ P_2 = (0,1)$ ed infine $Q = (1,1)$. Seguendo la procedura della Proposizione 7.9, i punti P_0, P_1, P_2 determinano sei possibili sistemi di riferimento distinti di $\mathbb{A}^2_{\mathbb{R}}$, rispetto a cui il punto Q avrà in generale coordinate differenti.

- Scegliamo il sistema di riferimente avente come origine il punto P_0 e come vettori costituenti la base $\mathcal{B}_0 = \{\overrightarrow{P_0P_1} = P_1 - P_0 = (1,0), \overrightarrow{P_0P_2} = P_2 - P_0 = (0,1)\}$. Osserviamo che le sottrazioni necessarie a calcolare i vettori di $\mathcal{B}_0$ sono operazioni ben definite, se teniamo conto che P_0, P_1, P_2 sono per definizione elementi dello spazio vettoriale $\mathbb{R}^2$. Questo sistema di riferimento è quello canonico per $\mathbb{A}^2_{\mathbb{R}}$. Rispetto ad esso, le coordinate di Q sono:

$$Q|_{\mathcal{B}_{P_0}} = \overrightarrow{P_0Q}|_{\mathcal{B}_0} = (Q - P_0)|_{\mathcal{B}_0} = (1,1)|_{\mathcal{B}_0} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

  Un secondo sistema di riferimento si può ottenere mantenendo la stessa origine e scambiando l'ordine dei vettori, ovvero $\mathcal{B}'_{P_0} = \{P_0, \overrightarrow{P_0P_2}, \overrightarrow{P_0P_1}\}$. In questo caso:

$$Q|_{\mathcal{B}'_{P_0}} = \overrightarrow{P_0Q}|_{\mathcal{B}'_0} = (Q - P_0)|_{\mathcal{B}'_0} = (1,1)|_{\mathcal{B}'_0} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

- Scegliamo l'origine P_1 e la base $\mathcal{B}_1 = \{\overrightarrow{P_1P_0} = P_0 - P_1 = (-1,0), \overrightarrow{P_1P_2} = P_2 - P_1 = (-1,1)\}$. Rispetto al sistema di riferimento $\mathcal{B}_{P_1} = \{P_1, \overrightarrow{P_1P_0}, \overrightarrow{P_1P_2}\}$, le coordinate di Q sono:

$$Q|_{\mathcal{B}_{P_1}} = \overrightarrow{P_1Q}|_{\mathcal{B}_1} = (Q - P_1)|_{\mathcal{B}_1} = (0,1)|_{\mathcal{B}_1} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

  Come prima, possiamo mantenere la stessa origine e scambiare l'ordine dei vettori ottenendo il sistema di riferimento $\mathcal{B}'_{P_1} = \{P_1, \overrightarrow{P_1P_2}, \overrightarrow{P_1P_0}\}$. In questo caso:

$$Q|_{\mathcal{B}'_{P_1}} = \overrightarrow{P_1Q}|_{\mathcal{B}'_1} = (Q - P_1)|_{\mathcal{B}'_1} = (0,1)|_{\mathcal{B}'_1} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

- Infine, scegliamo l'origine P_2 e la base $\mathcal{B}_2 = \{\overrightarrow{P_2P_0} = P_0 - P_2 = (0,-1), \overrightarrow{P_2P_1} = P_1 - P_2 = (1,-1)\}$. Rispetto al sistema di riferimento $\mathcal{B}_{P_2} = \{P_2, \overrightarrow{P_2P_0}, \overrightarrow{P_2P_1}\}$, le coordinate di Q sono:

$$Q|_{\mathcal{B}_{P_2}} = \overrightarrow{P_2Q}|_{\mathcal{B}_2} = (Q - P_2)|_{\mathcal{B}_2} = (1,0)|_{\mathcal{B}_2} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

  Per l'ultima volta, mantenendo la stessa origine e scambiando l'ordine dei vettori otteniamo il sistema di riferimento $\mathcal{B}'_{P_2} = \{P_2, \overrightarrow{P_2P_1}, \overrightarrow{P_2P_0}\}$. In questo caso:

$$Q|_{\mathcal{B}'_{P_2}} = \overrightarrow{P_2Q}|_{\mathcal{B}'_2} = (Q - P_2)|_{\mathcal{B}'_2} = (0,1)|_{\mathcal{B}'_2} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

7.2. Sottospazi affini

La definizione di sottospazio affine segue l'usuale schema valido per ogni tipo di sottostruttura.

Definizione 7.11. *Sia $\mathbb{A} = (\mathcal{A}, V, \psi)$ uno spazio affine ed S un sottoinsieme di $\mathcal{A}$. Denotiamo $\psi|_{S\times S}$ la restrizione di ψ. Allora S si dice sottospazio affine di $\mathbb{A}$ se $\mathrm{Im}(\psi|_{S\times S})$ è sottospazio vettoriale di V ed $(S, \mathrm{Im}(\psi|_{S\times S}), \psi|_{S\times S})$ è a sua volta uno spazio affine. Se $\dim(S) = \dim(\mathbb{A}) - 1$, il sottospazio si dice iperpiano di $\mathbb{A}$.*

D'altra parte, i sottospazi affini sono facilmente caratterizzabili.

Proposizione 7.12. *Sia $\mathbb{A} = (\mathcal{A}, V, \psi)$ uno spazio affine, S un sottoinsieme di $\mathcal{A}$ e P un punto di S. Dato l'insieme $U = \{\overrightarrow{PQ} \in V \mid Q \in S\}$, allora S è un sottospazio affine di $\mathbb{A}$ se e solo se U è un sottospazio vettoriale di V. In tal caso, U è la giacitura di S.*

Dimostrazione. Preliminarmente osserviamo che $U = \mathrm{Im}((\psi|_{S\times S})_P) \subseteq \mathrm{Im}(\psi|_{S\times S})$. La dimostrazione di entrambi i versi della proposizione si basa sul verificare che vale anche l'inclusione opposta, e quindi $U = \mathrm{Im}(\psi|_{S\times S})$. Questo prova automaticamente anche l'ultima affermazione dell'enunciato. Per prima cosa, supponiamo che S sia un sottospazio affine. Allora per definizione $(\psi|_{S\times S})_P$ è suriettiva e quindi $\mathrm{Im}((\psi|_{S\times S})_P) = \mathrm{Im}(\psi|_{S\times S})$. Pertanto U è un sottospazio vettoriale di V.

Viceversa, sia U un sottospazio vettoriale. Ogni elemento di $\mathrm{Im}(\psi|_{S\times S})$ è un vettore $\overrightarrow{Q_1Q_2}$ immagine di un segmento orientato $(Q_1, Q_2) \in S \times S$. Per la regola del parallelogramma, abbiamo

$$\overrightarrow{Q_1Q_2} = \overrightarrow{Q_1P} + \overrightarrow{PQ_2} = -\overrightarrow{PQ_1} + \overrightarrow{PQ_2},$$

che appartiene ad U per le proprietà di sottospazio vettoriale. Quindi $\mathrm{Im}(\psi|_{S\times S}) \subseteq U$, da cui $\mathrm{Im}(\psi|_{S\times S}) = U$. Pertanto $\mathrm{Im}(\psi|_{S\times S})$ è un sottospazio di V. Rimangono da verificare le proprietà di $\psi|_{S\times S}$. Poiché ψ_P è iniettiva per ogni $P \in \mathbb{A}$, lo sarà anche la sua restrizione $(\psi|_{S\times S})_P$ per ogni $P \in S$. Inoltre abbiamo appena provato che $\mathrm{Im}((\psi|_{S\times S})_P) = \mathrm{Im}(\psi|_{S\times S})$, di conseguenza $(\psi|_{S\times S})_P$ è suriettiva. Infine, la proprietà di Chasles è ereditata da ψ. Quindi S è un sottospazio affine. □

La Proposizione 7.12 può essere interpretata dicendo che un sottospazio affine S è completamente determinato se conosciamo un suo punto P e la sua giacitura U, poiché il punto $Q \in \mathbb{A}$ appartiene ad S se e solo se $\overrightarrow{PQ}$ appartiene ad U. Quindi, una retta affine è univocamente determinata da un suo punto P e da un vettore non nullo appartenente alla sua giacitura. Analogamente, un piano affine è univocamente determinato da un suo punto P e da due vettori indipendenti costituenti una base della sua giacitura.

Due tipologie importanti di sottospazi affini sono presentati nelle seguenti proposizioni.

Proposizione 7.13. *Siano $P_0, \dots, P_n$ punti dello spazio affine $\mathbb{A}$ sul campo $\mathbb{K}$. Il sottoinsieme*

$$S = \mathcal{L}(P_0, \dots, P_n) = \left\{ Q \in \mathbb{A} \;\middle|\; \overrightarrow{P_0Q} = \sum_{i=1}^{n} t_i \cdot \overrightarrow{P_0P_i},\ t_i \in \mathbb{K} \right\}$$

generato da tali punti è un sottospazio affine di $\mathbb{A}$.

Dimostrazione. Chiaramente S è il sottospazio contenente il punto P_0 e di giacitura $U = \mathcal{L}(\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_n})$. □

Come già osservato nella sezione precedente, il punto P_0 non svolge un ruolo privilegiato tra i generatori. Infatti è facile verificare che

$$S = \left\{ Q \in \mathbb{A} \;\middle|\; \overrightarrow{P_kQ} = \sum_{i=0}^{n} t_i \cdot \overrightarrow{P_kP_i},\ t_i \in \mathbb{K} \right\}$$

per ogni $0 \leq k \leq n$.

Proposizione 7.14. *Dati gli spazi vettoriali V e W, consideriamo l'applicazione $f \in \mathrm{Hom}(V, W)$ e lo spazio affine $\mathbb{A}_V = (V, V, \psi)$. Allora $f^{-1}(\mathbf{w})$ è un sottospazio affine di $\mathbb{A}_V$ per ogni $\mathbf{w} \in \mathrm{Im}(f) \subseteq W$.*

Dimostrazione. Verifichiamo la tesi utilizzando la Proposizione 7.12. Ricordiamo che per il Teorema 5.6, data una controimmagine particolare $\mathbf{v_p}$ di $\mathbf{w}$ vale

$$f^{-1}(\mathbf{w}) = \{\mathbf{v} \in V \mid \mathbf{v} = \mathbf{v_p} + \mathbf{v_0}\,,\ \text{dove } \mathbf{v_0} \in \ker(f)\}.$$

Quindi, per i punti $\mathbf{v_p}$ e $\mathbf{v}$ di $f^{-1}(\mathbf{w})$, abbiamo

$$\overrightarrow{\mathbf{v_p}\mathbf{v}} = \psi(\mathbf{v_p}, \mathbf{v}) = \mathbf{v} - \mathbf{v_p} = (\mathbf{v_p} + \mathbf{v_0}) - \mathbf{v_p} = \mathbf{v_0}$$

e pertanto $U = \{\overrightarrow{\mathbf{v_p}\mathbf{v}} \in V \mid \mathbf{v} \in f^{-1}(\mathbf{w})\} = \ker(f)$, che è un sottospazio di V. □

È possibile dimostrare anche l'implicazione inversa della proposizione precedente, ovvero che per ogni sottospazio affine S di $\mathbb{A}_V$ esiste un'applicazione lineare f ed un vettore $\mathbf{w} \in \mathrm{Im}(f) \subseteq W$ tale che $S = f^{-1}(\mathbf{w})$.

Esempio 7.15. Consideriamo l'insieme delle soluzioni di un sistema lineare di m equazioni a coefficienti in $\mathbb{K}$ ed n incognite. In forma matriciale, il sistema corrisponde all'equazione $A * X = B$, ovvero $f_A(X) = B$. Pertanto X è soluzione se e solo se appartiene ad $f_A^{-1}(B)$. Dalla proposizione precedente, segue quindi che l'insieme delle soluzioni, se non è vuoto, è un sottospazio affine di $\mathbb{A}^n_{\mathbb{K}}$. Ad esempio, risolviamo il sistema lineare

$$\begin{cases} x + y + kz = k \\ x + y - kz = k \end{cases}$$

dipendente dal paramentro $k \in \mathbb{R}$. In forma matriciale abbiamo

$$[A|B] = \left[\begin{array}{ccc|c} 1 & 1 & k & k \\ 1 & 1 & -k & k \end{array}\right] \xrightarrow{R_2 - R_1 \to R_2} \left[\begin{array}{ccc|c} 1 & 1 & k & k \\ 0 & 0 & -2k & 0 \end{array}\right] = [S|B'].$$

Se $k \neq 0$, l'insieme delle soluzioni è la retta affine

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} k \\ 0 \\ 0 \end{bmatrix} + t \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

Se $k = 0$, l'insieme delle soluzioni è il piano affine

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t_1 \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Il legame tra sottospazi affini e sistemi lineari è molto stretto. Infatti, valgono risultati simili a quelli contenuti nella Proposizione 4.68 e nel Corollario 4.70, le cui dimostrazioni sono del tutto analoghe ed i cui dettagli sono lasciati al lettore.

Proposizione 7.16. *Sia $\mathbb{A}$ uno spazio affine di dimensione $n < \infty$ su campo $\mathbb{K}$ e sia $\mathcal{B}_O$ un sistema di riferimento. Inoltre, sia S un sottospazio affine di dimensione m e sia $\mathcal{B}_Q = \{Q, \mathbf{u_1}, \ldots, \mathbf{u_m}\}$ un sistema di riferimento di S. Chiamiamo*

$$A = \left[\ \mathbf{u_1}|_{\mathcal{B}}\ \middle|\ \ldots\ \middle|\ \mathbf{u_m}|_{\mathcal{B}}\ \right] \in \mathrm{Mat}(n, m; \mathbb{K}), \qquad B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = Q|_{\mathcal{B}_O}$$

ed

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = P|_{\mathcal{B}_O}$$

la colonna delle coordinate del generico punto $P \in \mathbb{A}$ rispetto a $\mathcal{B}_O$. Allora:

i) esiste la decomposizione

$$X = B + \sum_{i=1}^{m} t_i \cdot A_{C(i)}$$

se e solo se $P \in S$. In tal caso, questa viene detta rappresentazione parametrica del sottospazio S rispetto al sistema di riferimento $\mathcal{B}_O$;

ii) se riduciamo a scala la matrice A contenuta all'interno della matrice orlata $[A|X - B]$, otteniamo una matrice $[S|X']$ le cui ultime $n - m$ righe costituiscono una rappresentazione algebrica di S rispetto al sistema di riferimento $\mathcal{B}_O$.

Corollario 7.17. *Sia $\mathbb{A}$ uno spazio affine di dimensione $n < \infty$ su campo $\mathbb{K}$ e sia $\mathcal{B}_O$ un sistema di riferimento. Allora il sottoinsieme non vuoto S di $\mathbb{A}$ è un sottospazio di dimensione m se e solo se esiste una matrice $[A|B] \in \mathrm{Mat}(n - m, n + 1; \mathbb{K})$ di rango massimo tale che $\phi_{\mathcal{B}_O}(S)$ è l'insieme delle soluzioni del sistema associato ad $[A|B]$. In tal caso, $\phi_{\mathcal{B}}$ è un isomorfismo tra la giacitura U di S e* $\ker(A)$.

Esempio 7.18. Nello spazio affine $\mathbb{A}^3_{\mathbb{R}}$, cerchiamo il piano Π contenente i punti $P_1 = (1, 0, 0), P_2 = (0, 1, 0), P_3 = (0, 0, 1)$. Sappiamo che il piano è univocamente determinato da un punto, ad esempio P_1, e da due vettori costituenti una base della giacitura. Consideriamo ad esempio i vettori

$$\overrightarrow{P_1P_2} = P_2 - P_1 = (-1, 1, 0), \qquad \overrightarrow{P_1P_3} = P_3 - P_1 = (-1, 0, 1).$$

I due vettori sono chiaramente indipendenti, quindi il piano è l'insieme di tutti i punti di $Q \in \mathbb{A}^3_{\mathbb{R}}$ tali che

$$\overrightarrow{P_1Q} = t_1 \cdot \overrightarrow{P_1P_2} + t_2 \cdot \overrightarrow{P_1P_3} \qquad \text{per ogni } t_1, t_2 \in \mathbb{R},$$

mentre $\mathcal{B}_{P_1} = \{P_1, \overrightarrow{P_1P_2}, \overrightarrow{P_1P_3}\}$ è un sistema di riferimento per Π.

Il sistema di riferimento canonico di $\mathbb{A}^3_{\mathbb{R}}$ è $\mathcal{B}_{P_0} = \{P_0, \overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, \overrightarrow{P_0P_3}\}$, dove l'origine è il punto $P_0 = (0,0,0)$ e la base associata della giacitura è $\mathcal{B}_0 = \{(1,0,0),(0,1,0),(0,0,1)\}$. Utilizzando le notazioni della Proposizione 7.16, abbiamo

$$A = \begin{bmatrix} \overrightarrow{P_1P_2}|_{\mathcal{B}_0} & \overrightarrow{P_1P_3}|_{\mathcal{B}_0} \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = P_1|_{\mathcal{B}_{P_0}}.$$

Quindi, una rappresentazione parametrica di Π è

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t_1 \cdot \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Una rappresentazione algebrica del piano si ottiene riducendo la matrice

$$\left[\begin{array}{cc|c} -1 & -1 & x-1 \\ 1 & 0 & y \\ 0 & 1 & z \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 0 & y \\ 0 & 1 & z \\ 0 & 0 & x+y+z-1 \end{array}\right],$$

da cui si ricava l'equazione $x + y + z - 1 = 0$.

Esempio 7.19. Sia $V = \mathrm{Mat}(2,2;\mathbb{Q})$, una cui base $\mathcal{B}$ è costituita dalle matrici

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \qquad D = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Scegliamo la matrice D come origine dello spazio affine $\mathbb{A}_V$, ottenendo così il sistema di riferimento $\mathcal{B}_D$. Consideriamo ora il sottoinsieme S di $\mathbb{A}_V$ i cui elementi sono le matrici

$$M = t_1 \cdot A + t_2 \cdot B + C + D,$$

dove $t_1, t_2 \in \mathbb{Q}$. Questo è un sottospazio affine di $\mathbb{A}_V$, caratterizzato dal punto $C + D \in S$ e la cui giacitura è $U = \mathcal{L}(A, B) \subset V$. Sia M la generica matrice di S avente coordinate

$$M|_{\mathcal{B}_D} = (M - D)|_{\mathcal{B}} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}.$$

Questo significa che

$$M = \phi^{-1}_{\mathcal{B}_D}(M|_{\mathcal{B}_D}) = D + x \cdot A + y \cdot B + z \cdot C + w \cdot D = \begin{bmatrix} 1+x+y+z+w & 1+y+z+w \\ 1+z+w & 1+w \end{bmatrix}.$$

Una rappresentazione parametrica di S rispetto a $\mathcal{B}$ è

$$\begin{aligned} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} &= (C+D)|_{\mathcal{B}_D} + t_1 \cdot A|_{\mathcal{B}} + t_2 \cdot B|_{\mathcal{B}} = C|_{\mathcal{B}} + t_1 \cdot A|_{\mathcal{B}} + t_2 \cdot B|_{\mathcal{B}} = \\ &= \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + t_1 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}. \end{aligned}$$

Per trovare una rappresentazione algebrica di S, consideriamo la matrice

$$\left[\begin{array}{cc|c} A|_{\mathcal{B}} & B|_{\mathcal{B}} & M|_{\mathcal{B}_D} - (C+D)|_{\mathcal{B}_D} \end{array}\right] = \left[\begin{array}{cc|c} 1 & 0 & x \\ 0 & 1 & y \\ 0 & 0 & z-1 \\ 0 & 0 & w \end{array}\right].$$

Quindi S è l'insieme delle matrici le cui coordinate rispetto a $\mathcal{B}_D$ soddisfano

$$\begin{cases} z = 1 \\ w = 0 \end{cases}.$$

Osservazione 7.20. Dato uno spazio affine $\mathbb{A}$, sia S un sottoinsieme di $\mathbb{A}$ che ammetta a sua volta una struttura affine. Allora non necessariamente S è un sottospazio affine di $\mathbb{A}$. Nell'Esempio 7.3 abbiamo visto che la parabola $S \subset \mathbb{A}^2_{\mathbb{K}}$, luogo delle soluzioni dell'equazione $y = x^2$, ammette una struttura affine. Ma tale struttura è diversa da quella dello spazio $\mathbb{A}^2_{\mathbb{K}}$ che contiene l'insieme. Di conseguenza S non è un sottospazio di $\mathbb{A}^2_{\mathbb{K}}$. Infatti, per il Corollario 7.17, ogni sottospazio affine di $\mathbb{A}^2_{\mathbb{K}}$ deve essere scrivibile come insieme degli zeri di un sistema di equazioni lineari nelle coordinate x, y dello spazio, cosa ovviamente non vera per la parabola.

7.3. Mutua posizione di sottospazi affini

Dal punto di vista geometrico, gli spazi affini sono caratterizzabili come gli insiemi in cui è possibile studiare i cosiddetti problemi di parallelismo ed incidenza.

Definizione 7.21. *Dato lo spazio affine $\mathbb{A}$, siano S e T due sottospazi di dimensione maggiore di zero ed aventi rispettivamente giaciture U e W. Allora, S e T si dicono:*

i) paralleli se $U \subseteq W$ oppure $W \subseteq U$;
ii) incidenti se non sono paralleli ed hanno almeno un punto in comune;
iii) sghembi se non sono paralleli od incidenti.

Osserviamo che le tre condizioni sono mutuamente esclusive e ricoprono tutte i possibili casi di mutua posizioni tra sottospazi affini. É immediato verificare che la relazione di parallelismo gode della proprietà riflessiva e di simmetria, ma non di quella transitiva. Infatti, nello spazio affine tridimensionale due rette parallele al medesimo piano possono essere parallele tra loro, incidenti o sghembe. Notiamo che due sottospazi sono paralleli anche se uno è contenuto nell'altro. In questo caso vale la seguente proposizione.

Proposizione 7.22. *Siano S, T due sottospazi affini paralleli di $\mathbb{A}$, con $\dim(S) \leq \dim(T) < \infty$. Supponiamo inoltre che $S \cap T \neq \emptyset$:*

i) se $\dim(S) < \dim(T)$ allora $S \subset T$;
ii) se $\dim(S) = \dim(T)$ allora $S = T$.

DIMOSTRAZIONE. Chiamiamo rispettivamente U e W le giaciture di S e T, che per la Proposizione 7.12 sono i sottospazi vettoriali

$$U = \{\overrightarrow{PQ} \mid Q \in S\} \qquad \text{e} \qquad W = \{\overrightarrow{PR} \mid R \in T\},$$

dove $P \in S \cap T$. La relazione $\dim(S) \leq \dim(T)$ significa che $\dim(U) \leq \dim(W)$. Assieme all'ipotesi di parallelismo, questo implica che $U \subseteq W$. Di conseguenza tutti i punti Q di S appartengono anche a T, cioè $S \subseteq T$. Analogamente, l'implicazione inversa vale se e solo se $\dim(T) \leq \dim(S)$, da cui segue la tesi. □

Conseguenza diretta della Definizione 7.21 e della Proposizione 7.22 è il cosiddetto postulato delle rette parallele della geometria euclidea, o per meglio dire la sua generalizzazione in dimensione qualunque.

Corollario 7.23. *Dati S un sottospazio di dimensione finita e P un punto dello spazio affine $\mathbb{A}$, esiste un unico sottospazio T parallelo ad S, contenente P e soddisfacente* $\dim(T) = \dim(S)$.

DIMOSTRAZIONE. Come abbiamo detto a seguito della Proposizione 7.12, ogni sottospazio è univocamente determinato da un suo punto e dalla sua giacitura. Se U è la giacitura di S, definiamo T come il sottospazio contenente P e di giacitura U. Per costruzione S e T sono paralleli e vale $\dim(S) = \dim(T)$. Supponiamo che T' sia un secondo sottospazio contenente P, parallelo ad S e soddisfacente $\dim(T') = \dim(T)$. Allora $T' = T$ per il secondo punto della Proposizione 7.22. □

Esempio 7.24. In $\mathbb{A}^2_{\mathbb{R}}$ fissiamo il sistema di riferimento $\mathcal{B}_O = \{O, \mathbf{v_1}, \mathbf{v_2}\}$. Data la retta S avente rappresentazione parametrica

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + t \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

cerchiamo una rappresentazione parametrica ed una algebrica della retta T parallela ad S e passante per il punto P di coordinate

$$P|_{\mathcal{B}_O} = \begin{bmatrix} 1 \\ k \end{bmatrix},$$

dove $k \in \mathbb{R}$. Per il Corollario 7.23 sappiamo che T esiste ed è unica. In particolare, la sua giacitura deve coincidere con quella di S. Pertanto una sua rappresentazione parametrica è

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ k \end{bmatrix} + t \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Ovviamente per $k = 0$ risulta $T = S$. Una rappresentazione algebrica di T si ottiene eliminando il parametro t dalla descrizione precedente:

$$\begin{cases} x = 1 + t \\ y = k + t \end{cases} \quad \Rightarrow \quad \begin{cases} t = x - 1 \\ y = k + x - 1 \end{cases} \quad \Rightarrow \quad x - y + k - 1 = 0.$$

Di conseguenza, una rappresentazione algebrica di S è $x - y - 1 = 0$. Come possiamo osservare, le equazioni delle due rette parallele differiscono solo per il valore del termine costante.

Le considerazioni dell'esempio precedente possono essere generalizzate a qualsiasi spazio affine. Per farlo, studiamo preliminarmente le proprietà dell'intersezione di due sottospazi affini. Negli enunciati dei prossimi teoremi, assumiamo che $\mathbb{A} = (\mathcal{A}, V, \psi)$ sia uno spazio affine di dimensione $n < \infty$ mentre S e T sono sottospazi rispettivamente di dimensione p e q, con $p \geq q$. Fissato un sistema di riferimento $\mathcal{B}_O$ di $\mathbb{A}$, supponiamo che i sistemi lineari associati ad $[A|B] \in \mathrm{Mat}(n-p, n+1)$ ed $[A'|B'] \in \mathrm{Mat}(n-q, n+1)$ siano rappresentazioni algebriche rispettivamente di S e T. Infine, definiamo

$$[\tilde{A}|\tilde{B}] = \left[\begin{array}{c|c} A & B \\ A' & B' \end{array}\right] \in \mathrm{Mat}(2n - p - q, n + 1; \mathbb{K}).$$

Lemma 7.25. *Se $S \cap T$ non è vuota, allora è un sottospazio affine di $\mathbb{A}$. In tal caso, una rappresentazione algebrica di $S \cap T$ è il sistema associato alla riduzione a scala della matrice $[\tilde{A}|\tilde{B}]$.*

DIMOSTRAZIONE. Le coordinate dei punti di $S \cap T$ devono essere soluzioni comuni di $[A|B]$ e $[A'|B']$, pertanto costituiscono l'insieme delle soluzioni di $[\tilde{A}|\tilde{B}]$. Per il Corollario 7.17, allora $S \cap T$ è un sottospazio affine di $\mathbb{A}$. □

Proposizione 7.26. *S e T sono:*

i) paralleli con $T \subseteq S$ se e solo se $r([\tilde{A}|\tilde{B}]) = r(\tilde{A}) = n - q$;

ii) paralleli disgiunti se e solo se $r([\tilde{A}|\tilde{B}]) > r(\tilde{A}) = n - q$;

iii) incidenti se e solo se $r([\tilde{A}|\tilde{B}]) = r(\tilde{A}) > n - q$;

iv) sghembi se e solo se $r([\tilde{A}|\tilde{B}]) > r(\tilde{A}) > n - q$;

Dimostrazione. Per il Lemma 7.25, l'intersezione $S \cap T$ ha come rappresentazione algebrica il sistema associato alla riduzione a scala di $[\tilde{A}|\tilde{B}]$. Studiamo le proprietà di tale matrice, utilizzando il Teorema di Rouché-Capelli. Dato che $\dim(S) = p$ e $\dim(T) = q$, le due matrici $[A|B] \in \mathrm{Mat}(n-p, n)$ ed $[A'|B'] \in \mathrm{Mat}(n-q, n)$ hanno rispettivamente rango $n - p$ ed $n - q$. Chiamiamo inoltre U e W le giaciture di S e T, che sono rispettivamente isomorfe a $\ker(A)$ e $\ker(A')$, attraverso la mappa delle componenti $\phi_{\mathcal{B}}$. Andiamo ora a verificare i vari casi dell'enunciato.

Per prima cosa indaghiamo la condizione di incidenza. Abbiamo $S \cap T \neq \emptyset$ se e solo se il sistema completo $[\tilde{A}|\tilde{B}]$ rappresentante l'intersezione ammette soluzione, e quindi se e solo se $r([\tilde{A}|\tilde{B}]) = r(\tilde{A})$. In tal caso, la giacitura di $S \cap T$ è $U \cap W$, che è isomorfa a $\ker(\tilde{A})$.

Concentriamoci invece sulla condizione di parallelismo. Dato che $p \geq q$, abbiamo che $U \not\subset W$. Di conseguenza, S e T sono paralleli se e solo se $W \subseteq U$. A sua volta questo è possibile se e solo se $U \cap W = W$, ovvero $\dim(U \cap W) = \dim(W)$. A livello di rappresentazione algebrica, questo è vero se e solo se $r(\tilde{A}) = r(A') = n - q$.

Infine, le quattro possibili combinazioni delle condizioni di incidenza e parallelismo corrispondono ai quattro casi della proposizione. □

Corollario 7.27. *Se S è un iperpiano di $\mathbb{A}$, allora S e T sono paralleli oppure incidenti.*

Dimostrazione. S è descritto da una sola equazione lineare, pertanto $r([A'|B'])+1 \geq r([\tilde{A}|\tilde{B}]) \geq r([A'|B']) = n - q$. Quindi la doppia disuguaglianza stretta $r([\tilde{A}|\tilde{B}]) > r(\tilde{A}) > n - q$ è impossibile. □

Corollario 7.28. *Siano U e W le giaciture di S e T, tali che $V = U + W$. Allora l'intersezione dei due sottospazi non è vuota. In particolare, l'intersezione è un unico punto se e solo se $V = U \oplus W$.*

Dimostrazione. Poiché $U \cap W$ è isomorfo a $\ker([\tilde{A}|\tilde{B}])$, segue che $\dim(U \cap W) = n - r([\tilde{A}|\tilde{B}])$. Da questa relazione e dalla formula di Grassmann otteniamo:

$$r([\tilde{A}|\tilde{B}]) = n - \dim(U \cap W) = n - (\dim(U) + \dim(W) - \dim(U + W)) = 2n - p - q = r(\tilde{A}),$$

da cui segue la prima parte della tesi. Per l'ultimo punto osserviamo che il sistema $[\tilde{A}|\tilde{B}]$ ha un unica soluzione se e solo se $r([\tilde{A}|\tilde{B}]) = n$, ovvero se e solo se $\dim(U \cap W) = 0$ e quindi $V = U \oplus W$. □

Esempio 7.29. Fissato un sistema di riferimento $\mathcal{B}_O$ di $\mathbb{A}^3_{\mathbb{R}}$, consideriamo il piano Π di equazione $x - y + hz - h = 0$ e la retta r avente rappresentazione algebrica $hx + (h+1)y + z = x + y - z - 1 = 0$, dove h è un parametro reale. Π è un iperpiano di $\mathbb{A}^3_{\mathbb{R}}$, di conseguenza la retta r è parallela oppure incidente a Π, come garantito dal Corollario 7.27. Applichiamo la Proposizione 7.26, tenendo conto che la dimensione dello spazio è $n = 3$, la dimensione del piano è $p = 2$ e della retta è $q = 1$:

$$[\tilde{A}|\tilde{B}] = \left(\begin{array}{ccc|c} 1 & -1 & h & h \\ h & h+1 & 1 & 0 \\ 1 & 1 & -1 & 1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & 1 & 1+h & -h \\ 0 & -2 & 1+h & -1+h \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & 1 & 1+h & -h \\ 0 & 0 & 3(1+h) & -1-h \end{array}\right).$$

Quindi, se $h \neq -1$ allora $r([\tilde{A}|\tilde{B}]) = r([\tilde{A}]) = 3 > n - q$ e di conseguenza il piano e la retta sono incidenti. Esiste un unico punto di intersezione, avente coordinate

$$x = \frac{1+2h}{3}, \qquad y = \frac{1-2h}{3}, \qquad z = -\frac{1}{3}.$$

Le giaciture U, W di piano e retta si ricavano risolvendo le equazioni omogenee associate a tali sottospazi affini:

$$U|_{\mathcal{B}} = \ker\left(\begin{bmatrix} 1 & -1 & h \end{bmatrix}\right) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ h \\ 1 \end{bmatrix}\right),$$

$$W|_{\mathcal{B}} = \ker\left(\begin{bmatrix} h & h+1 & 1 \\ 1 & 1 & -1 \end{bmatrix}\right) = \ker\left(\begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1+h \end{bmatrix}\right) = \mathcal{L}\left(\begin{bmatrix} 2+h \\ -1-h \\ 1 \end{bmatrix}\right).$$

Lasciamo al lettore verificare che se $h \neq 1$, esse soddisfano $\mathbb{R}^3 = U \oplus W$. Infine, se $h = -1$ allora $r([\tilde{A}|\tilde{B}]) = r([\tilde{A}]) = 2 = n - q$ e pertanto la retta risulta contenuta nel piano.

Esempio 7.30. Consideriamo un esempio meno intuitivo, che ci permette di apprezzare la generalità degli strumenti che abbiamo sviluppato in questo capitolo. Nei cosiddetti problemi di interpolazione si cercano tutte le funzioni, all'interno di una opportuna classe, che in determinati punti $x_1, \dots, x_n$ del dominio assumono valori fissati $y_1, \dots, y_n$. Per semplicità consideriamo l'insieme delle funzione $\mathbb{K}^{\mathbb{K}}$, che sappiamo essere uno spazio vettoriale di dimensione infinita. Le funzioni interpolanti formano un sottoinsieme di $\mathbb{K}^{\mathbb{K}}$, che è un sottospazio affine dello spazio affine associato a $\mathbb{K}^{\mathbb{K}}$. Per dimostrarlo consideriamo la cosiddetta mappa di valutazione di una funzione nel punto $x \in \mathbb{K}$

$$\begin{array}{rccc} \chi_x : & \mathbb{K}^{\mathbb{K}} & \longrightarrow & \mathbb{K} \\ & f & \longmapsto & f(x) \end{array}.$$

Non è difficile verificare che χ_x è un'applicazione lineare in $\mathrm{Hom}(\mathbb{K}^{\mathbb{K}}, \mathbb{K})$. Infatti

$$\chi_x(t_1 \cdot f_1 + t_2 \cdot f_2) = (t_1 \cdot f_1 + t_2 \cdot f_2)(x) = t_1 f_1(x) + t_2 f_2(x) = t_1 \cdot \chi_x(f_1) + t_2 \cdot \chi_x(f_2).$$

Osserviamo che $\chi_x^{-1}(y) = \{f \in \mathbb{K}^{\mathbb{K}} \,|\, f(x) = y\}$, che per la Proposizione 7.14 è un sottospazio affine di $\mathbb{K}^{\mathbb{K}}$. Poiché l'insieme delle funzioni interpolanti è $\bigcap_{i=1}^{n} \chi_{x_i}^{-1}(y_i)$, dalla Proposizione 7.22 segue la tesi. Queste considerazioni sono valide in qualsiasi spazio di funzioni.

Ad esempio, consideriamo gli insiemi

$$S = \{P(x) \in \mathbb{R}[x]_2 \,|\, P(0) = P'(0) = k,\ k \in \mathbb{R}\} \qquad \text{e} \qquad T = \{Q(x) \in \mathbb{R}[x]_2 \,|\, Q(1) = Q'(1) = 0\}.$$

Entrambi sono sottospazi affini di $\mathbb{R}[x]_2$. Per studiarli, scegliamo il sistema di riferimento canonico di tale spazio: $\mathcal{B}_O = \{0(x), 1, x, x^2\}$. Le coordinate dei generici polinomi $P(x) = p_0 \cdot 1 + p_1 \cdot x + p_2 \cdot x^2 \in S$ e $Q(x) = q_0 \cdot 1 + q_1 \cdot x + q_2 \cdot x^2 \in T$ sono quindi

$$P(x)|_{\mathcal{B}_O} = \begin{bmatrix} p_0 \\ p_1 \\ p_2 \end{bmatrix} \qquad \text{e} \qquad Q(x)|_{\mathcal{B}_O} = \begin{bmatrix} q_0 \\ q_1 \\ q_2 \end{bmatrix}.$$

Una coppia di rappresentazioni algebriche di S e T sono i sistemi lineari

$$\begin{cases} p_0 = k \\ p_1 = k \end{cases} \qquad \text{e} \qquad \begin{cases} q_0 + q_1 + q_2 = 0 \\ q_1 + 2q_2 = 0 \end{cases}.$$

Di conseguenza, i polinomi interpolanti in S hanno la forma $P(x) = k(1 + x) + p_2 x^2$ mentre quelli in T sono $Q(x) = q_2(1 - 2x + x^2)$, dove $p_2, q_2 \in \mathbb{R}$.

I due sottospazi sono quindi delle rette affini in $\mathbb{R}[x]_2$. Possiamo studiarne la mutua posizione. Le matrici associate ai due sistemi lineari sono

$$[A|B] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & k \\ 0 & 1 & 0 & k \end{array}\right] \qquad \text{ed} \qquad [A'|B'] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right],$$

da cui

$$[\tilde{A}|\tilde{B}] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & k \\ 0 & 1 & 0 & k \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right] \xrightarrow[R_4 - R_2 \to R_4]{R_3 - R_1 - R_2 \to R_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & k \\ 0 & 1 & 0 & k \\ 0 & 0 & 1 & -2k \\ 0 & 0 & 2 & -k \end{array}\right] \xrightarrow{R_4 - 2R_3 \to R_4} \left[\begin{array}{ccc|c} 1 & 0 & 0 & k \\ 0 & 1 & 0 & k \\ 0 & 0 & 1 & -2k \\ 0 & 0 & 0 & 3k \end{array}\right].$$

Deduciamo che S e T sono sghembe per $k \neq 0$, mentre sono incidenti per $k = 0$. In quest'ultimo caso, S e T sono entrambi sottospazi vettoriali di $\mathbb{R}[x]_2$, che si intersecano nella sola origine. Quindi il polinomio nullo è l'unico polinomio interpolante (di secondo grado) che soddisfa le condizioni di appartenenza ad entrambi gli insiemi.

7.4. Fasci di iperpiani

Definizione 7.31. *Sia $\mathbb{A}$ uno spazio affine di dimensione $n \geq 2$. Allora:*

i) se S è un sottospazio affine di dimensione $n - 2$, l'insieme di tutti gli iperpiani contenenti S è detto fascio proprio di iperpiani con sostegno S;

ii) se T è un iperpiano, l'insieme di tutti gli iperpiani paralleli a T è detto fascio improprio di iperpiani paralleli a T.

Fissato un sistema di riferimento $\mathcal{B}_O$ in $\mathbb{A}$, sappiamo che una qualsiasi rappresentazione algebrica di un iperpiano T è costituita da una sola equazione lineare $[A|B] \in \mathrm{Mat}(1, n + 1; \mathbb{K})$. L'intersezione di due generici iperpiani distinti T_1, T_2 del fascio, di equazioni $[A_1|B_1], [A_2|B_2] \in \mathrm{Mat}(1, n + 1; \mathbb{K})$, è rappresentata dal sistema

$$[\tilde{A}|\tilde{B}] = \left[\begin{array}{c|c} A_1 & B_1 \\ A_2 & B_2 \end{array}\right] \in \mathrm{Mat}(2, n + 1).$$

Se il fascio è proprio, allora $T_1 \cap T_2 \neq \emptyset$ e di conseguenza la soluzione del sistema esiste e, per il teorema di Rouché-Capelli, deve avere dimensione $n - 2$. Questo significa che l'intersezione coincide esattamente con S. Pertanto, T_1, T_2 sono sufficienti a definire completamente il fascio. Utilizzando le precedenti notazioni, vale la seguente caratterizzazione dei fasci.

Proposizione 7.32. *T appartiene al fascio contenente i due iperpiani distinti T_1, T_2 se e solo se esistono $t_1, t_2 \in \mathbb{K}$ tali che l'equazione associata a $[t_1 \cdot A_1 + t_2 \cdot A_2 | t_1 \cdot B_1 + t_2 \cdot B_2] \in \mathrm{Mat}(1, n; \mathbb{K})$ è una rappresentazione algebrica di T.*

Dimostrazione. Le coordinate dei punti nell'intersezione $T_1 \cap T_2 \cap T$ sono le soluzioni del sistema lineare

$$[A'|B'] = \left[\begin{array}{c|c} A_1 & B_1 \\ A_2 & B_2 \\ A & B \end{array}\right] \in \mathrm{Mat}(3, n + 1).$$

Nel caso proprio, le prime due righe di A' sono indipendenti e quindi le prime due equazioni descrivono il sostegno S. Di conseguenza, l'intersezione dei tre piani è ancora uguale ad S se e solo se se l'ultima riga $[A|B]$ del sistema lineare è combinazione delle due righe precedenti. Nel caso improprio, A_1 ed A_2 sono proporzionali, ovvero $A_2 = t \cdot A_1$, ma con $B_2 \neq t \cdot B_1$. L'iperpiano T appartiene al fascio se e solo se A è a sua volta proporzionale alle righe A_1, A_2. Quindi $r(A') = 1$ e di conseguenza $r(A'|B') = 2$. Questo significa che $[A|B]$ è combinazione lineare di $[A_1|B_1]$ ed $[A_2|B_2]$. □

Naturalmente, nel caso di spazi affini di dimensione due si parla di fasci di rette e in dimensione tre di fasci di piani.

Esempio 7.33. Consideriamo il fascio delle rette in $\mathbb{A}^2_{\mathbb{K}}$ contenenti il punto P, supponendo che le sue coordinate rispetto ad un sistema di riferimento siano x_P, y_P. L'equazione della generica retta del fascio si può ottenere come combinazione lineare delle due rette parallele agli assi coordinati e contenenti P:

$$t_1(x - x_P) + t_2(y - y_P) = 0.$$

In questa situazione, il supporto P del fascio viene anche detto punto base del fascio stesso. È ora immediato trovare la retta passante per P ed un secondo punto Q di coordinate x_Q, y_Q. Infatti è sufficiente imporre che una retta del fascio contenga Q:

$$t_1(x_Q - x_P) + t_2(y_Q - y_P) = 0 \qquad \Rightarrow \qquad \frac{t_1}{t_2} = -\frac{y_Q - y_P}{x_Q - x_P},$$

da cui

$$(y_Q - y_P)(x - x_P) - (x_Q - x_P)(y - y_P) = 0.$$

Esempio 7.34. Fissato un sistema di riferimento $\mathcal{B}_O$ nello spazio affine $\mathbb{A}^3_{\mathbb{R}}$, cerchiamo il piano contenente il punto P di coordinate

$$P|_{\mathcal{B}_O} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

e parallelo al piano Π di equazione $x+y+z = 0$. Per prima cosa, sappiamo che due piani sono paralleli se e solo se le loro rappresentazioni algebriche differiscono per il solo termine costante, in modo che la loro intersezione sia vuota. Nel nostro caso, il generico piano del fascio dei piani paralleli a Π ha equazione $x + y + z = k$, con $k \in \mathbb{R}$. Imponendo il passaggio per P otteniamo:

$$1 + 1 + 1 = k \qquad \Rightarrow \qquad x + y + z = 3.$$

7.5. Trasformazioni affini

Seguendo l'usuale approccio algebrico che abbiamo imparato a conoscere, una volta definita una struttura e studiate le sue proprietà principali, successivamente ci si concentra sulle funzioni naturali tra tali oggetti algebrici. Vediamo brevemente ciò nel caso degli spazi affini.

Definizione 7.35. *Siano* $\mathbb{A} = (\mathcal{A}, V, \psi_{\mathbb{A}})$ *e* $\mathbb{B} = (\mathcal{B}, W, \psi_{\mathbb{B}})$ *due spazi affini sul medesimo campo* $\mathbb{K}$. *Allora la funzione* $f : \mathcal{A} \to \mathcal{B}$ *è un omomorfismo di spazi affini, o trasformazione affine, se esiste un'applicazione lineare* $\bar{f} : V \to W$ *tale che per ogni segmento orientato* $(P, Q) \in \mathbb{A}$ *vale*

$$\psi_{\mathbb{B}}(f(P), f(Q)) = \bar{f}(\psi_{\mathbb{A}}(P, Q)).$$

Chiamiamo $\mathrm{Hom}(\mathbb{A}, \mathbb{B})$ *l'insieme di tutte le trasformazioni affini da* $\mathbb{A}$ *a* $\mathbb{B}$.

La precedente definizione richiede che ogni coppia di punti (P, Q) venga mandata in $(f(P), f(Q))$ tale che $\overrightarrow{f(P)f(Q)} = \bar{f}(\overrightarrow{PQ})$. Visto la stretto legame tra applicazioni lineari e trasformazioni affini, molti risultati contenuti nel Capitolo 5 ammettono una naturale estensione nel caso affine. Ad esempio, l'analogo del teorema di interpolazione afferma che una trasformazione affine è univocamente determinata dalla sua azione su un sistema di riferimento dello spazio di partenza. Qui ci limitiamo ad enunciare e dimostrare esplicitamente il teorema di rappresentazione per le trasformazioni affini.

Definizione 7.36. *Siano $\mathbb{A} = (\mathcal{A}, V, \psi_{\mathbb{A}})$ e $\mathbb{B} = (\mathcal{B}, W, \psi_{\mathbb{B}})$ spazi affini di dimensione finita sul medesimo campo $\mathbb{K}$, con sistemi di riferimento rispettivamente $\mathcal{B}_{\mathbb{A}} = \{O_{\mathbb{A}}\} \cup \mathcal{B}_V$ e $\mathcal{B}_{\mathbb{B}} = \{O_{\mathbb{B}}\} \cup \mathcal{B}_W$. Allora la mappa delle coordinate per le trasformazioni affini è la funzione*

$$\begin{array}{rccc} \phi_{\mathcal{B}_{\mathbb{A}}\mathcal{B}_{\mathbb{B}}}: & \mathrm{Hom}(\mathbb{A}, \mathbb{B}) & \longrightarrow & \mathrm{Mat}(m, n+1; \mathbb{K}) \\ & f & \longmapsto & [F|_{\mathcal{B}_V \mathcal{B}_W} \mid f(O_{\mathbb{A}})|_{\mathcal{B}_{\mathbb{B}}}] \end{array},$$

dove $F|_{\mathcal{B}_V \mathcal{B}_W}$ è la matrice rappresentante $\bar{f} \in \mathrm{Hom}(V, W)$.

Teorema 7.37. *Siano $\mathbb{A} = (\mathcal{A}, V, \psi_{\mathbb{A}})$ e $\mathbb{B} = (\mathcal{B}, W, \psi_{\mathbb{B}})$ spazi affini rispettivamente di dimensione n e m, sul medesimo campo $\mathbb{K}$, con sistemi di riferimento $\mathcal{B}_{\mathbb{A}} = \{O_{\mathbb{A}}\} \cup \mathcal{B}_V$ e $\mathcal{B}_{\mathbb{B}} = \{O_{\mathbb{B}}\} \cup \mathcal{B}_W$. Data $f \in \mathrm{Hom}(\mathbb{A}, \mathbb{B})$, allora $(F|_{\mathcal{B}_V \mathcal{B}_W} \mid f(O_{\mathbb{A}})|_{\mathcal{B}_{\mathbb{B}}})$ è l'unica matrice per cui vale*

$$f(P)|_{\mathcal{B}_{\mathbb{B}}} = (F|_{\mathcal{B}_V \mathcal{B}_W} \mid f(O_{\mathbb{A}})|_{\mathcal{B}_{\mathbb{B}}}) * \left(\frac{P|_{\mathcal{B}_{\mathbb{A}}}}{1} \right), \tag{12}$$

per ogni $P \in \mathbb{A}$.

Dimostrazione. Verifichiamo la validità della formula (12), attraverso la seguente catena di uguaglianza. In primo luogo utilizziamo l'Osservazione 7.7 e successivamente la regola di Chasles:

$$f(P)|_{\mathcal{B}_{\mathbb{B}}} = \overrightarrow{O_{\mathbb{B}} f(P)}|_{\mathcal{B}_W} = \left(\overrightarrow{O_{\mathbb{B}} f(O_{\mathbb{A}})} + \overrightarrow{f(O_{\mathbb{A}}) f(P)} \right)|_{\mathcal{B}_W}.$$

Sfruttiamo quindi la linearità della mappa delle coordinate e la definizione di trasformazione affine:

$$f(P)|_{\mathcal{B}_{\mathbb{B}}} = \overrightarrow{O_{\mathbb{B}} f(O_{\mathbb{A}})}|_{\mathcal{B}_W} + \overrightarrow{f(O_{\mathbb{A}}) f(P)}|_{\mathcal{B}_W} = \overrightarrow{O_{\mathbb{B}} f(O_{\mathbb{A}})}|_{\mathcal{B}_W} + \bar{f}\left(\overrightarrow{O_{\mathbb{A}} P} \right)|_{\mathcal{B}_W}.$$

Possiamo ora usare il teorema di rappresentazione delle applicazioni lineari ed ancora l'Osservazione 7.7:

$$f(P)|_{\mathcal{B}_{\mathbb{B}}} = \overrightarrow{O_{\mathbb{B}} f(O_{\mathbb{A}})}|_{\mathcal{B}_W} + F|_{\mathcal{B}_V \mathcal{B}_W} * \overrightarrow{O_{\mathbb{A}} P}|_{\mathcal{B}_V} = F|_{\mathcal{B}_V \mathcal{B}_W} * P|_{\mathcal{B}_{\mathbb{A}}} + f(O_{\mathbb{A}})|_{\mathcal{B}_{\mathbb{B}}}.$$

Infine, applicando il Lemma 3.9, è facile verificare che l'ultima espressione è equivalente alla formula (12). L'unicità della matrice è conseguenza dell'unicità della matrice rappresentante $\bar{f}$. □

Esempio 7.38. Definiamo alcune trasformazioni affini significative dal punto di vista geometrico.

- In $\mathbb{A} = (\mathcal{A}, V, \psi)$ la funzione identità

$$\begin{array}{rccc} \mathrm{Id}_{\mathcal{A}}: & \mathcal{A} & \longrightarrow & \mathcal{A} \\ & P & \longmapsto & P \end{array}$$

  è ovviamente una trasformazione affine, la cui applicazione lineare associata è Id_V. Come per le applicazioni lineare, attraverso la rappresentazione di questa trasformazione otteniamo la regola del cambio di coordinate tra due sistemi di riferimento $\mathcal{B}_{\mathbb{A}} = \{O\} \cup \mathcal{B}_V$ e $\mathcal{B}'_{\mathbb{A}} = \{O'\} \cup \mathcal{B}'_V$:

$$P|_{\mathcal{B}'_{\mathbb{A}}} = \mathrm{Id}_{\mathbb{A}}(P)|_{\mathcal{B}'_{\mathbb{A}}} = \mathrm{Id}_V|_{\mathcal{B}_V \mathcal{B}'_V} * P|_{\mathcal{B}_{\mathbb{A}}} + \mathrm{Id}_{\mathbb{A}}(O)|_{\mathcal{B}'_{\mathbb{A}}} = M_{\mathcal{B}_V \mathcal{B}'_V} * P|_{\mathcal{B}_{\mathbb{A}}} + O|_{\mathcal{B}'_{\mathbb{A}}}.$$

Come esempio esplicito, consideriamo i punti $P_0 = (0,0), P_1 = (1,0), P_2 = (0,1) \in \mathbb{A}^2_{\mathbb{R}}$. Studiamo il cambio di coordinate tra i due sistemi di riferimento $\mathcal{B}_{P_0} = \left\{P_0, \overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}\right\}$ e $\mathcal{B}_{P_1} = \left\{P_1, \overrightarrow{P_1P_0}, \overrightarrow{P_1P_2}\right\}$. Per prima cosa, le basi della giacitura associate ai due sistemi sono

$$\mathcal{B}_0 = \left\{\overrightarrow{P_0P_1} = (1,0), \overrightarrow{P_0P_2} = (0,1)\right\} \qquad \text{e} \qquad \mathcal{B}_1 = \left\{\overrightarrow{P_1P_0} = (-1,0), \overrightarrow{P_1P_2} = (-1,1)\right\}.$$

Di conseguenza, la matrice del cambiamento di base è

$$M|_{\mathcal{B}_0\mathcal{B}_1} = \left[\begin{array}{cc} \overrightarrow{P_0P_1}|_{\mathcal{B}_1} & \overrightarrow{P_0P_2}|_{\mathcal{B}_1} \end{array}\right] = \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}.$$

Invece, il termine additivo nel cambiamento di coordinate è

$$P_0|_{\mathcal{B}_{P_1}} = (P_0 - P_1)|_{\mathcal{B}_1} = (-1,0)|_{\mathcal{B}_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Quindi, le coordinante di un generico punto P nei due sistemi di riferimento sono legate da

$$P|_{\mathcal{B}_{P_1}} = \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix} * P|_{\mathcal{B}_{P_0}} + \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

- In $\mathbb{A} = (\mathcal{A}, V, \psi)$ scegliamo il vettore $\mathbf{v} \in V$. Allora la traslazione definita da $\mathbf{v}$ è la funzione

$$\begin{array}{rccc} t_{\mathbf{v}}: & \mathcal{A} & \longrightarrow & \mathcal{A} \\ & P & \longmapsto & \psi_P^{-1}(\mathbf{v}) \end{array},$$

ovvero $t_{\mathbf{v}}$ associa ad ogni punto P il punto $t_{\mathbf{v}}(P)$ tale che $\overrightarrow{Pt_{\mathbf{v}}(P)} = \mathbf{v}$. Per prima cosa osserviamo che $t_{\mathbf{v}}$ è ben definita per ogni $\mathbf{v} \in V$, in quanto la funzione ψ_P è biunivoca per ogni $P \in \mathcal{A}$. Secondariamente, verifichiamo che $t_{\mathbf{v}}$ sia una trasformazione affine. Per ogni segmento orientato (P,Q) abbiamo

$$\overrightarrow{t_{\mathbf{v}}(P)t_{\mathbf{v}}(Q)} = \overrightarrow{t_{\mathbf{v}}(P)P} + \overrightarrow{PQ} + \overrightarrow{Qt_{\mathbf{v}}(Q)} = -\overrightarrow{Pt_{\mathbf{v}}(P)} + \overrightarrow{PQ} + \overrightarrow{Qt_{\mathbf{v}}(Q)} = -\mathbf{v} + \overrightarrow{PQ} + \mathbf{v} = \overrightarrow{PQ},$$

da cui l'applicazione lineare associata alla traslazione è ancora una volta Id_V. A livello di coordinate, questo implica che per ogni $P \in \mathbb{A}$ vale

$$t_{\mathbf{v}}(P)|_{\mathcal{B}_O} = \mathbb{I}_n * P|_{\mathcal{B}_O} + t_{\mathbf{v}}(O)|_{\mathcal{B}_O} = P|_{\mathcal{B}_O} + \overrightarrow{Ot_{\mathbf{v}}(O)}|_{\mathcal{B}} = P|_{\mathcal{B}_O} + \mathbf{v}|_{\mathcal{B}}.$$

- Dati $\mathbb{A} = (\mathcal{A}, V, \psi_{\mathbb{A}})$ e $\mathbb{B} = (\mathcal{B}, W, \psi_{\mathbb{B}})$ e scelti due punti $A \in \mathcal{A}$ e $B \in \mathcal{B}$, ad ogni applicazione lineare $f \in \mathrm{Hom}(V,W)$ è associata la funzione

$$\begin{array}{rccc} f_{AB}: & \mathcal{A} & \longrightarrow & \mathcal{B} \\ & P & \longmapsto & \psi_{\mathbb{B},B}^{-1}(f(\psi_{\mathbb{A},A}(P))) \end{array},$$

ovvero f_{AB} associa ad un punto $P \in \mathcal{A}$ il punto $Q \in \mathcal{B}$ tale che $\overrightarrow{BQ} = f(\overrightarrow{AP})$. Questa funzione è una trasformazione affine e valgono $f_{AB}(A) = B$ ed $\overline{f_{AB}} = f$. Quindi, se $F|_{\mathcal{B}_V\mathcal{B}_W}$ è la matrice rappresentante f, abbiamo

$$f_{AB}(P)|_{\mathcal{B}_B} = F|_{\mathcal{B}_V\mathcal{B}_W} * P|_{\mathcal{B}_A}.$$

- Siano S, T sottospazi di $\mathbb{A} = (\mathcal{A}, V, \psi)$ tali che le loro giaciture U, W soddisfino $V = U \oplus W$. Per il Corollario 7.28 esiste un unico punto di intersezione $O \in S \cap T$. La proiezione su S parallela a T è la funzione

$$\begin{array}{rccc} \mathcal{P}_{S;T}: & \mathcal{A} & \longrightarrow & S \\ & P & \longmapsto & \psi_O^{-1}(\mathcal{P}_{U;W}(\psi_O(P))) \end{array},$$

ovvero $\mathcal{P}_{S;T}$ associa ad un punto $P \in \mathcal{A}$ il punto $Q \in S$ tale che $\overrightarrow{OQ} = \mathcal{P}_{U;W}(\overrightarrow{OP})$. Analogamente, la riflessione rispetto ad S e parallela a T è la funzione

$$\begin{array}{rccc} \mathcal{R}_{S;T}: & \mathcal{A} & \longrightarrow & \mathcal{A} \\ & P & \longmapsto & \psi_O^{-1}(\mathcal{R}_{U;W}(\psi_O(P))) \end{array}.$$

Anche la proiezione e la riflessione sono trasformazioni affini, con $\overline{\mathcal{P}_{S;T}} = \mathcal{P}_{U;W}$ ed $\overline{\mathcal{R}_{S;T}} = \mathcal{R}_{U;W}$.

Ad esempio, nello spazio affine $\mathbb{A}^3_{\mathbb{R}}$ consideriamo il piano e la retta affini

$$S = \{(x, y, z) \in \mathbb{A}^3_{\mathbb{R}} \mid x + y + z = 1\},$$
$$T = \{(x, y, z) \in \mathbb{A}^3_{\mathbb{R}} \mid (x, y, z) = (1, 0, 0) + t(1, 2, 1),\ t \in \mathbb{R}\}.$$

Le giaciture dei due sottospazi sono rispettivamente

$$U = \mathcal{L}((1, -1, 0), (1, 0, -1)), \qquad W = \mathcal{L}((1, 2, 1)).$$

Vale $\mathbb{R}^3 = U \oplus W$ e l'unico punto di intersezione dei due sottospazi è $O = (1, 0, 0)$. Dato $P = (1, 1, 1) \in \mathbb{A}^3_{\mathbb{R}}$, abbiamo

$$\overrightarrow{OP} = (0, 1, 1) = -\frac{1}{2} \cdot (1, 0, -1) + \frac{1}{2} \cdot (1, 2, 1) \qquad \Rightarrow \qquad \overrightarrow{OQ} = \mathcal{P}_{U;W}(\overrightarrow{OP}) = -\frac{1}{2} \cdot (1, 0, -1)$$

e quindi $\mathcal{P}_{S;T}(P) = Q = \left(\frac{1}{2}, 0, \frac{1}{2}\right)$. La riflessione $R = \mathcal{R}_{S:T}(P)$ si ricava invece da

$$\overrightarrow{OR} = \mathcal{R}_{U;W}(\overrightarrow{OP}) = -\frac{1}{2} \cdot (1, 0, -1) - \frac{1}{2} \cdot (1, 2, 1) = (-1, -1, 0),$$

per cui $R = (0, -1, 0)$.

- Nel caso in cui S è un iperpiano, si definisce anche la proiezione su S di centro $O \in \mathcal{A}$. Iniziamo osservando che per ogni punto P dello spazio esiste un'unica retta r contente O e P. Sia T l'unico iperpiano parallelo ad S e contenente O. Se P non appartiene a T, allora r non è parallela ad S ed interseca l'iperpiano in un unico punto Q. Utilizzando questa notazione, definiamo

$$\begin{array}{rccc} \mathcal{P}_{S;O}: & \mathcal{A} \setminus T & \longrightarrow & S \\ & P & \longmapsto & Q \end{array}.$$

Ad esempio, in $\mathbb{A}^2_{\mathbb{R}}$ consideriamo la proiezione sulla retta $S = \{(x, y) \in \mathbb{A}^2_{\mathbb{R}} \mid x = 1\}$ e di centro $O = (0, 0)$. La retta contenente O ed il punto (a, b) ha equazione $bx - ay = 0$ ed interseca S nel punto $\left(1, \frac{b}{a}\right)$, se $a \neq 0$. Quindi

$$\mathcal{P}_{S;O}((a, b)) = \left(1, \frac{b}{a}\right) \qquad \text{se } a \neq 0.$$

La funzione $\mathcal{P}_{S;O}$ non è una trasformazione affine, come si può osservare esplicitamente nell'esempio precedente. Per ogni $t \in \mathbb{R}^*$, i punti $P_t = (t, -1)$ e $Q_t = (t, 1)$ definiscono il medesimo vettore $\overrightarrow{P_tQ_t} = (0, 2) \in \mathbb{R}^2$. Di conseguenza, se la funzione $\mathcal{P}_{S;O}$ fosse affine, allora il vettore $\overrightarrow{\mathcal{P}_{S;O}(P_t)\mathcal{P}_{S;O}(Q_t)}$ non dovrebbe dipendere da t. Ma questo non è vero in quanto

$$\overrightarrow{\mathcal{P}_{S;O}(P_t)\mathcal{P}_{S;O}(Q_t)} = \mathcal{P}_{S;O}(Q_t) - \mathcal{P}_{S;O}(P_t) = \left(1, \frac{1}{t}\right) - \left(1, -\frac{1}{t}\right) = \left(0, \frac{2}{t}\right).$$

Concludiamo la sezione analizzando brevemente le trasformazioni affini invertibili.

Definizione 7.39. *Dati due spazi affini $\mathbb{A}$ e $\mathbb{B}$, una trasformazione affine $f \in \operatorname{Hom}(\mathbb{A}, \mathbb{B})$ la cui applicazione lineare associata $\bar{f}$ è un isomorfismo si dice affinità. In tal caso $\mathbb{A}$ e $\mathbb{B}$ si dicono affinamente equivalenti.*

Le traslazioni e le riflessioni sono delle affinità. Un altro esempio di affinità sono le omotetie, ovvero le trasformazioni affini associate alle trasformazioni lineari del tipo $f = c \cdot \mathrm{Id}_V$. Dal punto di vista geometrico, le affinità si possono caratterizzare come le trasformazioni che non alterano le relazioni di incidenza e parallelismo. Ovvero l'immagine di sottospazi paralleli attraverso affinità sono ancora sottospazi paralleli della stessa dimensione di quelli di partenza, ed analogamente per tutti gli altri casi di mutua posizione. Per le affinità valgono molte proprietà analoghe a quelle riguardanti gli isomorfismi di spazi vettoriali. In particolare, lasciamo al lettore la verifica del seguente risultato.

Proposizione 7.40. *Siano $\mathbb{A}, \mathbb{B}$ due spazi affini. Allora:*

i) *$\mathbb{A}, \mathbb{B}$ sono affinamente equivalenti se e solo se hanno la stessa dimensione;*

ii) *una trasformazione affine $f \in \mathrm{Hom}(\mathbb{A}, \mathbb{B})$ è un'affinità se e solo se l'immagine di un sistema di riferimento di $\mathbb{A}$ è un sistema di riferimento di $\mathbb{B}$.*

7.6. La geometria affine e gli Elementi di Euclide

Ricordiamo i cinque postulati della geometria euclidea piana:

i) un segmento di linea retta può essere disegnato tra ogni coppia di punti;

ii) qualsiasi segmento può essere esteso ad una linea retta;

iii) dato un qualsiasi segmento, è possibile disegnare un cerchio avente tale segmento come raggio ed un suo punto estremo come centro;

iv) tutti gli angoli retti sono congruenti;

v) date una linea retta ed un punto non appartenente a tale linea, esiste una ed una sola linea contenente il punto e non intersecante la retta.

In questo capitolo, all'interno del contesto assiomatico degli spazi affini, noi abbiamo definito i concetti di punto, segmento, retta e piano. Per costruzione, ogni coppia di punti definisce un unico segmento ed esiste un'unica retta contenente tale segmento. Inoltre, abbiamo introdotto il concetto di parallelismo e dimostrato il quinto postulato di Euclide nel Corollario 7.23[1] Abbiamo pertanto costruito un modello per i postulati (i), (ii) e (v). Sono rimasti però completamente estranei alla nostra trattazione i postulati (iii) e (iv). Infatti, al momento non abbiamo nemmeno l'idea di come definire un angolo ed una circonferenza di un dato raggio. Per poter parlare di questi concetti abbiamo bisogno di arricchire gli strumenti di cui disponiamo, e lo faremo nel nostro solito modo, cioè definendo una nuova struttura. Nel Capitolo 8 introdurremo i cosiddetti spazi vettoriali euclidei, che costituiranno la base algebrica per poter completare il percorso che abbiamo tracciato nel Capitolo 1, cosa che poi faremo nel Capitolo 10 riguardante per l'appunto la geometria euclidea.

---

[1] Il lettore non deve rimanere confuso da questo fatto. Infatti, questo non significa che il quinto postulato sia logicamente derivabile dai primi due, in quanto il nostro risultato è stato ottenuto in un contesto completamente differente, basato sugli assiomi degli spazi affini che la formulazione classica della geometria euclidea non prevede affatto.

CAPITOLO 8

Spazi vettoriali euclidei

Come abbiamo già discusso nella Sezione 7.6, gli strumenti algebrici presentati finora ci permettono di riottenere in forma rigorosa e generalizzata solo parte della geometria classica studiata a partire dagli antichi greci. In particolare, non siamo ancora in grado di analizzare le cosiddette proprietà metriche degli oggetti geometrici, ovvero le loro lunghezze e gli angoli da essi formati. Colmeremo tale vuoto in questo capitolo, grazie alla definizione della struttura di spazio vettoriale euclideo. Precisiamo subito che, diversamente da quanto fatto finora, in questo contesto non è possibile lavorare in completa generalità relativamente al campo di supporto $\mathbb{K}$, ma è invece necessario utilizzare il campo dei numeri reali $\mathbb{R}$. In particolare, esula dagli scopi di queste dispense la trattazione del corrispondente complesso degli spazi vettoriali euclidei, ovvero gli spazi vettoriali hermitiani.

8.1. Definizione, esempi e proprietà elementari

Definizione 8.1. *Sia V uno spazio vettoriale reale. Un prodotto scalare su V è un'operazione binaria esterna a valori reali, denotata con*

$$\begin{array}{rccc} \langle\ ,\ \rangle : & V \times V & \longrightarrow & \mathbb{R} \\ & (\mathbf{v_1}, \mathbf{v_2}) & \longmapsto & \langle \mathbf{v_1}, \mathbf{v_2} \rangle \end{array}$$

e soddisfacente le seguenti proprietà:

i) è lineare rispetto al primo argomento, ovvero $\langle t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2}, \mathbf{v_3} \rangle = t_1 \langle \mathbf{v_1}, \mathbf{v_3} \rangle + t_2 \langle \mathbf{v_2}, \mathbf{v_3} \rangle$ per ogni $t_1, t_2 \in \mathbb{K}$ e $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3} \in V$;

ii) è simmetrica, ovvero $\langle \mathbf{v_1}, \mathbf{v_2} \rangle = \langle \mathbf{v_2}, \mathbf{v_1} \rangle$ per ogni $\mathbf{v_1}, \mathbf{v_2} \in V$;

iii) è definita positiva, cioè $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ per ogni $\mathbf{v} \in V$, e vale esattamente 0 solo se $\mathbf{v} = \mathbf{0}$.

Definizione 8.2. *Uno spazio vettoriale euclideo è la struttura algebrica $(V, \langle\ ,\ \rangle)$ composta da uno spazio vettoriale reale ed un prodotto scalare.*

Come al solito, è possibile definire sul medesimo spazio vettoriale differenti prodotti scalari e quindi differenti strutture euclidee. D'altra parte, in mancanza di ambiguità identificheremo normalmente uno spazio euclideo con lo spazio vettoriale di supporto.

Definizione 8.3. *In uno spazio vettoriale euclideo V, la norma indotta dal prodotto scalare è la funzione*

$$\begin{array}{rccc} \|\ \| : & V & \longrightarrow & \mathbb{R} \\ & \mathbf{v} & \longmapsto & \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} \end{array} .$$

Grazie all'assioma di positività del prodotto scalare, la norma è una funzione ben definita e per costruzione essa associa ad ogni vettore un numero non negativo. Intuitivamente, questo numero deve essere pensato come la lunghezza del vettore. Osserviamo quindi che un vettore appartenente ad uno spazio vettoriale non ha una sua lunghezza intrinseca, bensì questa dipende da come è definito il prodotto scalare, cioè dalla struttura euclidea dello spazio.

Esempio 8.4. Per ogni spazio vettoriale presentato nell'Esempio 4.2, dopo aver posto $\mathbb{K} = \mathbb{R}$, studiamo i cosiddetti prodotti scalari canonici o standard, che permettono di definire le corrispondenti strutture euclidee.

- Nello spazio $\mathbb{R}^n$, scegliamo due generici vettori $\mathbf{u} = (u_1, \ldots, u_n)$ e $\mathbf{v} = (v_1, \ldots, v_n)$. Il prodotto scalare canonico è

$$\langle \mathbf{u}, \mathbf{v} \rangle_E = \sum_{i=1}^{n} u_i v_i.$$

È immediato verificare la validità delle prime due proprietà dei prodotti scalari, infatti

$$\langle \alpha \cdot \mathbf{u} + \beta \cdot \mathbf{v}, \mathbf{w} \rangle_E = \sum_{i=1}^{n} (\alpha u_i + \beta v_i) w_i = \alpha \sum_{i=1}^{n} u_i w_i + \beta \sum_{i=1}^{n} v_i w_i = \alpha \langle \mathbf{u}, \mathbf{w} \rangle_E + \beta \langle \mathbf{v}, \mathbf{w} \rangle_E ,$$

$$\langle \mathbf{u}, \mathbf{v} \rangle_E = \sum_{i=1}^{n} u_i v_i = \sum_{i=1}^{n} v_i u_i = \langle \mathbf{v}, \mathbf{u} \rangle_E .$$

Inoltre,

$$\langle \mathbf{v}, \mathbf{v} \rangle_E = \sum_{i=1}^{n} v_i^2 \geq 0$$

in quanto l'espressione è somma di quadrati di numeri reali. Naturalmente tale somma è uguale a zero se e solo se tutti gli elementi v_i sono nulli, per $1 \leq i \leq n$. La norma indotta dal prodotto scalare è

$$\|\mathbf{v}\|_E = \sqrt{\sum_{i=1}^{n} v_i^2}.$$

Queste due funzioni vengono anche dette prodotto scalare e norma euclidei.

- Siano $A, B \in \mathrm{Mat}(m, n; \mathbb{R})$, allora il prodotto scalare canonico o di Frobenius è

$$\langle A, B \rangle_F = \mathrm{Tr}(A^T * B).$$

Per verificarne le proprietà, andiamo a sviluppare esplicitamente questa operazione rispetto alle entrate delle due matrici:

$$\mathrm{Tr}(A^T * B) = \sum_{i=1}^{n} (A^T * B)_{ii} = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ji} b_{ji}.$$

Il lettore noti la stretta analogia tra il prodotto scalare di Frobenius e quello euclideo visto nell'esempio precedente. In questa forma, la dimostrazione della linearità, simmetria e positività del prodotto è del tutto simile al caso del prodotto euclideo. Infine, la norma di Frobenius di una matrice è

$$\|A\|_F = \sqrt{\mathrm{Tr}(A^T * A)} = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ji}^2} .$$

- Dato $I \subseteq \mathbb{R}$, l'insieme delle funzioni quadrato sommabili su I secondo Riemann[1] è

$$L_2(I; \mathbb{R}) = \left\{ f : I \to \mathbb{R} \;\middle|\; \int_I f(x)^2 dx \text{ esiste finito} \right\}.$$

Sfruttando le proprietà dell'integrale di Riemann studiate nel corso di Analisi, è possibile verificare che $L_2(I; \mathbb{R})$ è uno spazio vettoriale reale, rispetto alle solite operazioni di somma e prodotto negli

[1] Per pignoleria, osserviamo che nella definizione dello spazio $L_2(I; \mathbb{R})$ che si può trovare in un qualunque testo di Analisi Funzionale, l'integrazione è fatta secondo Lebesgue. Questo è necessario per avere la cosiddetta proprietà di completezza dello spazio vettoriale.

spazi di funzioni. Su questo spazio possiamo definire il seguente prodotto scalare canonico, detto anche prodotto L_2:

$$\langle f, g\rangle_I = \int_I f(x)g(x)dx\,.$$

Grazie alla definizione di $L_2(I;\mathbb{R})$, l'integrale esiste finito per ogni coppia di funzioni in questo spazio ed inoltre soddisfa tutte le proprietà dei prodotti scalari. La norma L_2 di una funzione è

$$\|f\|_I = \sqrt{\int_I f(x)^2 dx}\,.$$

In particolare, le funzioni polinomiali sono tutte quadrato sommabili su intervalli limitati di $\mathbb{R}$. Sfruttando il Teorema fondamentale del calcolo, il prodotto L_2 di due polinomi si può calcolare utilizzando l'integrale algebrico. Infatti, supponiamo per semplicità che $I = [a, b]$ ed $P, Q \in \mathbb{K}[x]$. Allora

$$\langle P, Q\rangle_I = \int_a^b P(x)Q(x)dx\,.$$

Ad esempio, se $I = [0, 1]$ ed i due polinomi sono $P(x) = 1 + x$ e $Q(x) = x$, si ha

$$\langle P, Q\rangle_I = \int_0^1 (1 + x)x\,dx = \int_0^1 x + x^2\,dx = \frac{x^2}{2} + \frac{x^3}{3}\bigg|_0^1 = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}\,.$$

La norma L_2 dei due polinomi è

$$\|P\| = \sqrt{\int_0^1 (1 + x)^2\,dx} = \sqrt{\int_0^1 1 + 2x + x^2\,dx} = \sqrt{x + x^2 + \frac{x^3}{3}\bigg|_0^1} = \sqrt{1 + 1 + \frac{1}{3}} = \sqrt{\frac{7}{3}}\,,$$

$$\|Q\| = \sqrt{\int_0^1 x^2\,dx} = \sqrt{\frac{x^3}{3}\bigg|_0^1} = \sqrt{\frac{1}{3}}\,.$$

Esempio 8.5. Nello spazio vettoriale $V = \mathrm{Mat}(n, 1;\mathbb{R})$, consideriamo una matrice simmetrica $G \in \mathbb{S}(n;\mathbb{R})$ e l'operazione che ad ogni $X, Y \in V$ associa

$$\langle X, Y\rangle_G = X^T * G * Y.$$

Sfruttando la naturale identificazione tra $\mathrm{Mat}(1, 1;\mathbb{R})$ ed $\mathbb{R}$, possiamo pensare al risultato dell'operazione come ad un numero reale. E' facile verificare che $\langle X, Y\rangle$ è lineare nel primo argomento e che è simmetrica, infatti:

$$\langle t_1 \cdot X_1 + t_2 \cdot X_2, Y\rangle_G = (t_1{\cdot}X_1{+}t_2{\cdot}X_2)^T{*}G{*}Y = t_1{\cdot}X_1^T{*}G{*}Y{+}t_2{\cdot}X_2^T{*}G{*}Y = t_1\langle X_1, Y\rangle_G{+}t_2\langle X_2, Y\rangle_G\,,$$

$$\langle Y, X\rangle_G = Y^T * G * X = (Y^T * G * X)^T = X^T * G^T * (Y^T)^T = X^T * G * Y = \langle X, Y\rangle_G\,.$$

Se richiediamo che la matrice G sia definita positiva, ovvero che

$$\langle X, X\rangle_G = X^T * G * X \geq 0$$

per ogni $X \in V$ e sia esattamente zero se e solo se $X = 0_{n1}$, allora l'operazione $\langle\ ,\ \rangle_G$ è un prodotto scalare su V. Vedremo in seguito che una condizione necessaria e sufficiente per la positività di G è che tutti i suoi autovalori siano numeri reali strettamente positivi.

Ad esempio, in $\mathrm{Mat}(2, 1;\mathbb{R})$ andiamo a studiare tutti i possibili prodotti scalari di questo tipo:

$$G = \begin{bmatrix} a & b \\ b & c \end{bmatrix},\ X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},\ Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \qquad \Rightarrow \qquad X^T * G * Y = ax_1y_1 + b(x_1y_2 + x_2y_1) + cx_2y_2.$$

Gli autovalori di G sono:

$$\lambda_\pm = \frac{\mathrm{Tr}(G) \pm \sqrt{\mathrm{Tr}(G)^2 - 4\det(G)}}{2} = \frac{\mathrm{Tr}(G)}{2}\left(1 \pm \sqrt{1 - \frac{4\det(G)}{\mathrm{Tr}(G)^2}}\right).$$

La matrice è definita positiva se e solo se gli autovalori sono reali e strettamente positivi, ovvero se e solo se

$$\mathrm{Tr}(G) > 0 \qquad \text{e} \qquad 0 < \frac{4\det(G)}{\mathrm{Tr}(G)^2} \leq 1.$$

L'ultima disuguaglianza è ridondante, infatti

$$4\det(G) \leq \mathrm{Tr}(G)^2 \qquad \Leftrightarrow \qquad 4ac - 4b^2 \leq a^2 + 2ac + c^2 \qquad \Leftrightarrow \qquad -4b^2 \leq a^2 - 2ac + c^2 = (a-c)^2,$$

che ovviamente è sempre vera per qualunque $a, b, c \in \mathbb{R}$. Quindi le condizioni su G affinché tale matrice definisca un prodotto scalare sono $\mathrm{Tr}(G) > 0$ e $\det(G) > 0$, ovvero $a + c > 0$ e $ac > b^2$.

Presentiamo ora alcune proprietà elementari del prodotto scalare e della norma indotta dal prodotto.

Proposizione 8.6. *Sia V uno spazio vettoriale euclideo. Allora:*

i) il prodotto scalare è bilineare, ovvero per ogni $\alpha_i, \beta_j \in \mathbb{R}$ e $\mathbf{u_i}, \mathbf{v_j} \in V$, con $1 \leq i \leq m$ e $1 \leq j \leq n$, vale

$$\left\langle \sum_{i=1}^{m} \alpha_i \cdot \mathbf{u_i}\,, \sum_{j=1}^{n} \beta_j \cdot \mathbf{v_j} \right\rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_i \beta_j \left\langle \mathbf{u_i}, \mathbf{v_j} \right\rangle;$$

ii) il prodotto scalare è non degenere, ovvero $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ per ogni $\mathbf{v} \in V$ se e solo se $\mathbf{u} = \mathbf{0}$;

iii) vale la legge di annullamento per la norma, ovvero $\|\mathbf{v}\| = 0$ se e solo se $\mathbf{v} = \mathbf{0}$;

iv) vale la proprietà di omogeneità della norma, cioè per ogni $t \in \mathbb{R}$ e $\mathbf{v} \in V$ abbiamo $\|t \cdot \mathbf{v}\| = |t| \, \|\mathbf{v}\|$.

Dimostrazione. Verifichiamo singolarmente ogni punto dell'enunciato.

i) Per definizione, il prodotto scalare è lineare nel primo argomento, pertanto

$$\left\langle \sum_{i=1}^{m} \alpha_i \cdot \mathbf{u_i}\,, \sum_{j=1}^{n} \beta_j \cdot \mathbf{v_j} \right\rangle = \sum_{i=1}^{m} \alpha_i \left\langle \mathbf{u_i}\,, \sum_{j=1}^{n} \beta_j \cdot \mathbf{v_j} \right\rangle.$$

Utilizziamo ora la simmetria del prodotto scalare:

$$\sum_{i=1}^{m} \alpha_i \left\langle \mathbf{u_i}\,, \sum_{j=1}^{n} \beta_j \cdot \mathbf{v_j} \right\rangle = \sum_{i=1}^{m} \alpha_i \left\langle \sum_{j=1}^{n} \beta_j \cdot \mathbf{v_j}\,, \mathbf{u_i} \right\rangle.$$

Grazie nuovamente alla linearità del primo argomento, otteniamo

$$\sum_{i=1}^{m} \alpha_i \left\langle \sum_{j=1}^{n} \beta_j \cdot \mathbf{v_j}\,, \mathbf{u_i} \right\rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_i \beta_j \left\langle \mathbf{v_j}, \mathbf{u_i} \right\rangle,$$

da cui ricaviamo infine la tesi utilizzando ancora una volta la simmetria del prodotto.

ii) Sia $\mathbf{u} = \mathbf{0}$. Sfruttando la linearità del prodotto scalare si ricava

$$\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{w} - \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle - \langle \mathbf{w}, \mathbf{v} \rangle = 0.$$

D'altra parte, per la positività del prodotto scalare, $\mathbf{0}$ è l'unico vettore il cui prodotto con se stesso è nullo.

iii) La legge di annullamento segue direttamente dalla proprietà di positività del prodotto scalare.

iv) Infine, grazie ancora alla bilinearità del prodotto scalare, abbiamo

$$\|t \cdot \mathbf{v}\| = \sqrt{\langle t \cdot \mathbf{v}, t \cdot \mathbf{v} \rangle} = \sqrt{t^2 \langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{t^2 \|\mathbf{v}\|^2} = |\, t \, \|\mathbf{v}\| \,| = |t| \, \|\mathbf{v}\|. \qquad \square$$

Proposizione 8.7. *Sia V uno spazio vettoriale euclideo, allora per ogni $\mathbf{u}, \mathbf{v} \in V$ valgono la formula di Carnot*

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle$$

e la formula di polarizzazione

$$\langle \mathbf{u}, \mathbf{v} \rangle = \frac{\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2}{2}.$$

DIMOSTRAZIONE. Chiaramente le due formule sono equivalenti, in quanto legate da una trasformazione algebrica elementare. Verifichiamo la prima:

$$\|\mathbf{u}+\mathbf{v}\|^2 = \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle,$$

dove nelle uguaglianze abbiamo sfruttato la bilinearità e la simmetria del prodotto scalare. □

L'ultima proposizione è particolarmente interessante, perché ci permette di scambiare i ruoli del prodotto scalare e della norma. Infatti, grazie alla formula di polarizzazione possiamo dire che il prodotto scalare, e quindi la struttura euclidea, è completamente determinata se conosciamo le norme di tutti i vettori appartenenti allo spazio vettoriale di supporto.

Concludiamo infine la sezione fornendo la definizione di versore e di normalizzazione.

Definizione 8.8. *Siano V uno spazio vettoriale euclideo e $\mathbf{v} \in V$ un vettore. Allora:*

i) se $\|\mathbf{v}\| = 1$ il vettore è detto versore;
ii) se $\|\mathbf{v}\| \neq 0$, il versore $\tilde{\mathbf{v}} = \dfrac{\mathbf{v}}{\|\mathbf{v}\|}$ è detto normalizzazione di $\mathbf{v}$ oppure versore associato a $\mathbf{v}$.

Un sottoinsieme U di V si dice normalizzato se è composto da versori.

Esempio 8.9. In $\mathbb{R}^n$ la base canonica $\mathcal{B}_n = \{\mathbf{e_1}, \ldots, \mathbf{e_n}\}$ è composta da versori, rispetto al prodotto scalare euclideo. Infatti:

$$\langle \mathbf{e_i}, \mathbf{e_j} \rangle_E = \sum_{k=1}^{n} \delta_{ik}\delta_{jk} = \delta_{ij} \qquad \Rightarrow \qquad \|\mathbf{e_i}\|_E = 1$$

per ogni $1 \leq i, j \leq n$. La stessa cosa si può dimostrare per la base canonica $\mathcal{B}_{mn}$ dello spazio $\mathrm{Mat}(m, n; \mathbb{K})$ relativamente al prodotto scalare di Frobenius. Invece, la base canonica $\mathcal{B}_n = \{1, \ldots, x^n\}$ di $\mathbb{R}[x]_n$ non è normalizzata rispetto al prodotto scalare L_2. Ad esempio, se $I = [0, 1]$ allora vale

$$\|x^i\|_I = \sqrt{\int_0^1 x^{2i}\, dx} = \sqrt{\left.\frac{x^{2i+1}}{2i+1}\right|_0^1} = \frac{1}{\sqrt{2i+1}}.$$

La base composta dai vettori normalizzati è dunque $\tilde{\mathcal{B}}_n = \{\sqrt{2i+1}\, x^i,\ 1 \leq i \leq n\}$.

8.2. Matrice di Gram e teorema di rappresentazione

Esattamente come per le applicazioni lineari, anche i prodotti scalari definibili su uno spazio vettoriale reale si possono rappresentare attraverso opportune matrici, una volta che è stata fissata una base.

Definizione 8.10. *Siano V uno spazio vettoriale euclideo ed $U = \{\mathbf{u_1}, \ldots, \mathbf{u_m}\}$ un sottoinsieme finito di V. Allora la matrice di Gram associata ad U è*

$$G|_U = [\langle \mathbf{u_i}, \mathbf{u_j} \rangle] \in \mathrm{Mat}(m, m; \mathbb{R}).$$

Il determinate di $G|_U$ si dice gramiano di U e lo indichiamo con $\det(G|_U) = G(U) = G(\mathbf{u_1}, \ldots, \mathbf{u_m})$.

Esempio 8.11. Calcoliamo le matrice di Gram di alcune basi degli spazi euclidei presentati negli Esempi 8.4 e 8.5.

- In $\mathbb{R}^n$ prendiamo la base canonica $\mathcal{B}_n = \{\mathbf{e_1}, \dots, \mathbf{e_n}\}$. La matrice di Gram rispetto al prodotto scalare euclideo è
$$G|_{\mathcal{B}_n} = [\langle \mathbf{e_i}, \mathbf{e_j} \rangle] = [\delta_{ij}] = \mathbb{I}_n.$$
Osserviamo che se cambiamo base, la matrice di Gram cambia anch'essa. Ad esempio, in $\mathbb{R}^3$ scegliamo la base $\mathcal{B} = \{\mathbf{v_1} = (1,0,0), \mathbf{v_2} = (1,1,0), \mathbf{v_3} = (1,1,1)\}$. Allora
$$G|_{\mathcal{B}} = \begin{bmatrix} \langle \mathbf{v_1}, \mathbf{v_1} \rangle_E & \langle \mathbf{v_1}, \mathbf{v_2} \rangle_E & \langle \mathbf{v_1}, \mathbf{v_3} \rangle_E \\ \langle \mathbf{v_2}, \mathbf{v_1} \rangle_E & \langle \mathbf{v_2}, \mathbf{v_2} \rangle_E & \langle \mathbf{v_2}, \mathbf{v_3} \rangle_E \\ \langle \mathbf{v_3}, \mathbf{v_1} \rangle_E & \langle \mathbf{v_3}, \mathbf{v_2} \rangle_E & \langle \mathbf{v_3}, \mathbf{v_3} \rangle_E \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$
- In $\mathrm{Mat}(n,n;\mathbb{R})$ prendiamo la base canonica $\mathcal{B}_{nn} = \{E_{11}, \dots, E_{nn}\}$. Anche in questo caso, la matrice di Gram rispetto al prodotto scalare di Frobenius è l'identità, ovvero $G|_{\mathcal{B}_{nn}} = \mathbb{I}_{n^2}$. Ad esempio, nel caso $n = 2$ abbiamo
$$G|_{\mathcal{B}} = \begin{bmatrix} \langle E_{11}, E_{11} \rangle_F & \langle E_{11}, E_{12} \rangle_F & \langle E_{11}, E_{21} \rangle_F & \langle E_{11}, E_{22} \rangle_F \\ \langle E_{12}, E_{11} \rangle_F & \langle E_{12}, E_{12} \rangle_F & \langle E_{12}, E_{21} \rangle_F & \langle E_{12}, E_{22} \rangle_F \\ \langle E_{21}, E_{11} \rangle_F & \langle E_{21}, E_{12} \rangle_F & \langle E_{21}, E_{21} \rangle_F & \langle E_{21}, E_{22} \rangle_F \\ \langle E_{22}, E_{11} \rangle_F & \langle E_{22}, E_{12} \rangle_F & \langle E_{22}, E_{21} \rangle_F & \langle E_{22}, E_{22} \rangle_F \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \mathbb{I}_4.$$
- Nello spazio $\mathbb{K}[x]_n$ scegliamo il prodotto L_2 sull'intervallo $I = [0,1]$. La matrice di Gram associata alla base canonica $\mathcal{B}_n = \{1, \dots, x^n\}$ è
$$G|_{\mathcal{B}_n} = \left[\left\langle x^i, x^j \right\rangle_I\right] = \left[\int_0^1 x^{i+j}\,dx\right] = \left[\left.\frac{x^{i+j+1}}{i+j+1}\right|_0^1\right] = \left[\frac{1}{i+j+1}\right].$$
Ad esempio, se $n = 2$ abbiamo
$$G|_{\mathcal{B}_2} = \begin{bmatrix} \langle 1, 1 \rangle_I & \langle 1, x \rangle_I & \left\langle 1, x^2 \right\rangle_I \\ \langle x, 1 \rangle_I & \langle x, x \rangle_I & \left\langle x, x^2 \right\rangle_I \\ \left\langle x^2, 1 \right\rangle_I & \left\langle x^2, x \right\rangle_I & \left\langle x^2, x^2 \right\rangle_I \end{bmatrix} = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix}.$$
$G|_{\mathcal{B}_n}$ è nota come matrice di Hilbert ed è un caso particolare di matrice di Hankel e di matrice di Cauchy.
- In $\mathrm{Mat}(n,1;\mathbb{R})$ consideriamo il prodotto scalare $\langle\ ,\ \rangle_G$ definito da una matrice $G \in \mathcal{S}(n;\mathbb{R})$. Rispetto alla base canonica $\mathcal{B}_{n1} = \{E_{11}, \dots, E_{n1}\}$, la matrice di Gram è
$$G|_{\mathcal{B}_{n1}} = \left[\langle E_{i1}, E_{j1} \rangle_G\right] = \left[E_{1i}^T * G * E_{j1}\right] = [(G)_{ij}],$$
per cui $G|_{\mathcal{B}_{n1}} = G$.

Proposizione 8.12. *Siano V uno spazio vettoriale euclideo finitamente generato e $\mathcal{B} = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ una base. Allora:*

i) l'unica matrice $A \in \mathrm{Mat}(n,n;\mathbb{R})$ tale che $\langle \mathbf{u}, \mathbf{w} \rangle = \mathbf{u}|_{\mathcal{B}}^T * A * \mathbf{w}|_{\mathcal{B}}$ per ogni $\mathbf{u}, \mathbf{w} \in V$ è la matrice di Gram $G|_{\mathcal{B}}$;

ii) la matrice di Gram associata alla base $\mathcal{B}'$ è $G|_{\mathcal{B}'} = M_{\mathcal{B}'\mathcal{B}}^T * G|_{\mathcal{B}} * M_{\mathcal{B}'\mathcal{B}}$.

Dimostrazione. Consideriamo lo sviluppo dei vettori $\mathbf{u}, \mathbf{w}$ rispetto alla base $\mathcal{B}$:
$$\mathbf{u} = \sum_{i=1}^n \alpha_i \mathbf{v_i} \qquad \text{e} \qquad \mathbf{v} = \sum_{j=1}^n \beta_j \mathbf{v_j}.$$

Sostituendo queste espressioni nel prodotto scalare otteniamo

$$\langle \mathbf{u},\mathbf{w}\rangle = \left\langle \sum_{i=1}^{n} \alpha_i \mathbf{v_i}, \sum_{j=1}^{n} \beta_j \mathbf{v_j} \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\beta_j \langle \mathbf{v_i},\mathbf{v_j}\rangle = \mathbf{u}|_{\mathcal{B}}^T * G|_{\mathcal{B}} * \mathbf{w}|_{\mathcal{B}}\,.$$

D'altra parte, se A soddisfa la formula $\langle \mathbf{u},\mathbf{w}\rangle = \mathbf{u}|_{\mathcal{B}}^T * A * \mathbf{w}|_{\mathcal{B}}$, allora:

$$(G|_{\mathcal{B}})_{ij} = \langle \mathbf{v_i},\mathbf{v_j}\rangle = \mathbf{v_i}|_{\mathcal{B}}^T * A * \mathbf{v_j}|_{\mathcal{B}} = E_{i1}^T * A * E_{j1} = a_{ij},$$

da cui ricaviamo $A = G$. Concentriamo ora sul punto (ii). Dal punto precedente e dalla regola del cambio base per le componenti di un vettore, sappiamo che

$$\langle \mathbf{u},\mathbf{w}\rangle = \mathbf{u}|_{\mathcal{B}}^T * G|_{\mathcal{B}} * \mathbf{w}|_{\mathcal{B}} = (M_{\mathcal{B}'\mathcal{B}} * \mathbf{u}|_{\mathcal{B}'})^T * G|_{\mathcal{B}} * (M_{\mathcal{B}'\mathcal{B}} * \mathbf{w}|_{\mathcal{B}'}) = \mathbf{u}|_{\mathcal{B}'}^T * (M_{\mathcal{B}'\mathcal{B}}^T * G|_{\mathcal{B}} * M_{\mathcal{B}'\mathcal{B}}) * \mathbf{w}|_{\mathcal{B}'}.$$

La tesi segue quindi dall'unicità della matrice rappresentativa del prodotto scalare, che abbiamo prima dimostrato. □

Grazie alla proposizione precedente possiamo affermare che un prodotto scalare è completamente determinato se conosciamo la matrice di Gram rispetto ad una qualunque base dello spazio vettoriale, ovvero se conosciamo il prodotto scalare tra tutte le coppie di elementi contenuti in tale base.

Esempio 8.13. In $\mathbb{K}[x]_2$ calcoliamo il prodotto scalare L_2 su $I = [0,1]$ tra due generici polinomi $P(x) = a_0 + a_1x + a_2x^2$ e $Q(x) = b_0 + b_1x + b_2x^2$:

$$\begin{aligned}
\langle P,Q\rangle &= \int_0^1 P(x)Q(x)\,dx = P|_{\mathcal{B}_2}^T * G|_{\mathcal{B}_2} * Q|_{\mathcal{B}_2} = \\
&= \begin{bmatrix} a_0 & a_1 & a_2 \end{bmatrix} * \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix} * \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} a_0 & a_1 & a_2 \end{bmatrix} * \begin{bmatrix} b_0 + \frac{b_1}{2} + \frac{b_2}{3} \\ \frac{b_0}{2} + \frac{b_1}{3} + \frac{b_2}{4} \\ \frac{b_0}{3} + \frac{b_1}{4} + \frac{b_2}{5} \end{bmatrix} = \\
&= a_0b_0 + \frac{a_0b_1}{2} + \frac{a_0b_2}{3} + \frac{a_1b_0}{2} + \frac{a_1b_1}{3} + \frac{a_1b_2}{4} + \frac{a_2b_0}{3} + \frac{a_2b_1}{4} + \frac{a_2b_2}{5}\,.
\end{aligned}$$

Due matrici quadrate A e B, legate da una trasformazione del tipo $B = P^T * G * P$ dove P è invertibile, si dicono congruenti. In generale si può dimostrare che due matrici sono congruenti se e solo se rappresentano il medesimo prodotto scalare rispetto a basi differenti.

Infine, un'importante applicazione della matrice di Gram è la verifica dell'indipendenza lineare di un insieme di vettori.

Proposizione 8.14. *Siano V uno spazio vettoriale euclideo ed $U = \{\mathbf{u_1},\dots,\mathbf{u_n}\}$ un sottoinsieme finito di vettori. Allora $G(U)$ è diverso da zero se e solo se U è un insieme linearmente indipendente.*

Dimostrazione. Sappiamo che il determinante di una matrice è diverso da zero se e solo se le colonne della matrice stessa formano un insieme linearmente indipendente. Studiamo allora la combinazione lineare delle colonne di $G|_U$:

$$\sum_{i=1}^{n} t_i \cdot \begin{bmatrix} \langle \mathbf{u_1},\mathbf{u_i}\rangle \\ \vdots \\ \langle \mathbf{u_n},\mathbf{u_i}\rangle \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n t_i \langle \mathbf{u_1},\mathbf{u_i}\rangle \\ \vdots \\ \sum_{i=1}^n t_i \langle \mathbf{u_n},\mathbf{u_i}\rangle \end{bmatrix} = \begin{bmatrix} \langle \mathbf{u_1}, \sum_{i=1}^n t_i \cdot \mathbf{u_i}\rangle \\ \vdots \\ \langle \mathbf{u_n}, \sum_{i=1}^n t_i \cdot \mathbf{u_i}\rangle \end{bmatrix}.$$

Se una tale combinazione esiste, allora segue che

$$\left\| \sum_{i=1}^{n} t_i \cdot \mathbf{u_i} \right\|^2 = \left\langle \sum_{i=1}^{n} t_i \cdot \mathbf{u_i}, \sum_{i=1}^{n} t_i \cdot \mathbf{u_i} \right\rangle = \sum_{j=1}^{n} t_j \left\langle \mathbf{u_j}, \sum_{i=1}^{n} t_i \cdot \mathbf{u_i} \right\rangle = \sum_{i=1}^{n} t_j\, 0 = 0.$$

Per la legge di annullamento, questo implica

$$\sum_{i=1}^{n} t_i \cdot \mathbf{u_i} = \mathbf{0}.$$

Viceversa, ipotizzando che esista una combinazione lineare dei vettori con risultato nullo, è facile osservare che è nulla anche la corrispondente combinazione lineare delle colonne. Di conseguenza, l'insieme U è linearmente indipendente se e solo se lo è anche l'insieme delle colonne di $G|_U$, cioè se e solo se $\det(G|_U) \neq 0$. $\square$

Esempio 8.15. Utilizziamo l'ultima proposizione per verificare l'indipendenza lineare dell'insieme $U = \{\mathbf{u_1} = (1,-1,0,0), \mathbf{u_2} = (0,1,-1,0), \mathbf{u_3} = (1,0,-1,0)\} \subset \mathbb{R}^4$. Rispetto al prodotto scalare euclideo, il gramiano di U è

$$G(U) = \begin{vmatrix} \langle \mathbf{u_1}, \mathbf{u_1} \rangle_E & \langle \mathbf{u_1}, \mathbf{u_2} \rangle_E & \langle \mathbf{u_1}, \mathbf{u_3} \rangle_E \\ \langle \mathbf{u_2}, \mathbf{u_1} \rangle_E & \langle \mathbf{u_2}, \mathbf{u_2} \rangle_E & \langle \mathbf{u_2}, \mathbf{u_3} \rangle_E \\ \langle \mathbf{u_3}, \mathbf{u_1} \rangle_E & \langle \mathbf{u_3}, \mathbf{u_2} \rangle_E & \langle \mathbf{u_3}, \mathbf{u_3} \rangle_E \end{vmatrix} = \begin{vmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{vmatrix} = 0.$$

Di conseguenza, l'insieme U è linearmente dipendente, come d'altra parte segue facilmente dalla relazione $\mathbf{u_1} + \mathbf{u_2} = \mathbf{u_3}$.

8.3. Ortogonalità, complemento ortogonale e proiezione ortogonale

Il terzo criterio di congruenza della geometria euclidea piana afferma che due triangoli aventi i lati della stessa lunghezza sono congruenti. Questo significa che tali triangoli possono essere sovrapposti attraverso movimenti rigidi e di conseguenza essi devono avere anche gli angoli uguali. In altri termini, si può dire che gli angoli di un triangolo sono completamente determinati dalle lunghezze dei suoi lati. In questa e nelle prossime due sezioni vogliamo portare queste osservazioni all'interno del formalismo che stiamo sviluppando, per arrivare a definire l'angolo tra due vettori come una funzione delle loro norme, o equivalentemente, dei loro prodotti scalari. Iniziamo con il presentare il concetto di ortogonalità.

Definizione 8.16. *Sia V uno spazio vettoriale euclideo. Due vettori $\mathbf{v_1}, \mathbf{v_2}$ si dicono ortogonali se $\langle \mathbf{v_1}, \mathbf{v_2} \rangle = 0$. In tal caso si scrive $\mathbf{v_1} \perp \mathbf{v_2}$. Un sottoinsieme U di V si dice ortogonale se ogni suo elemento è ortogonale a tutti gli altri ed inoltre si dice ortonormale se i vettori sono anche normalizzati.*

Come per la norma, anche l'ortogonalità di U non è un concetto intrinsecamente definito in uno spazio vettoriale, ma dipende invece dalla struttura euclidea dello spazio. Osserviamo anche che $\mathbf{0}$ è l'unico vettore ortogonale a tutti gli altri vettori ed in particolare è l'unico ortogonale a se stesso. Quest'ultima proprietà è invece indipendente dal prodotto scalare scelto.

Esempio 8.17. Riprendendo l'Esempio 8.9, sappiamo che i vettori delle basi canoniche di $\mathbb{R}^n$ e di $\mathrm{Mat}(m,n;\mathbb{K})$ sono ortogonali, rispetto al prodotto scalare euclideo e di Frobenius. Quindi queste si dicono basi ortonormali. Invece, la base canonica di $\mathbb{R}[x]_n$, oltre a non essere normalizzata, non è nemmeno ortogonale rispetto al prodotto L_2. Sempre ponendo $I = [0,1]$, abbiamo infatti

$$\left\langle x^i, x^j \right\rangle = \int_0^1 x^{i+j}\, dx = \left. \frac{x^{i+j+1}}{i+j+1} \right|_0^1 = \frac{1}{i+j+1} \neq 0$$

per ogni $1 \leq i, j \leq n$.

Le osservazioni precedenti si possono ottenere anche studiando le matrici di Gram associate alle basi. Infatti, è evidente che un insieme U di n vettori è

i) normalizzato se e solo se tutti gli elementi lungo la diagonale di $G|_U$ sono uguali ad uno;

ii) ortogonale se e solo se $G|_U \in \mathbb{D}(n; \mathbb{R})$;

iii) ortonormale se e solo se $G|_U = \mathbb{I}_n$.

Vogliamo ora mostrare che la definizione precedente di ortogonalità permette di riottenere alcuni importanti risultati della geometria euclidea. In primo luogo, il cosiddetto Teorema di Pitagora generalizzato.

Proposizione 8.18. *In uno spazio euclideo V, sia $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ un insieme ortogonale. Allora:*

$$\left\| \sum_{i=1}^{n} \mathbf{v_i} \right\|^2 = \sum_{i=1}^{n} \|\mathbf{v_i}\|^2 .$$

Dimostrazione. Verifichiamolo per induzione su n. Il caso $n = 2$ è un corollario della formula di Carnot. Consideriamo il caso generale, supponendo vera la tesi fino ad $n-1$. Sempre per la formula di Carnot, abbiamo

$$\left\| \sum_{i=1}^{n} \mathbf{v_i} \right\|^2 = \left\| \sum_{i=1}^{n-1} \mathbf{v_i} + \mathbf{v_n} \right\|^2 = \left\| \sum_{i=1}^{n-1} \mathbf{v_i} \right\|^2 + \|\mathbf{v_n}\|^2 + 2 \left\langle \sum_{i=1}^{n-1} \mathbf{v_i} \, , \mathbf{v_n} \right\rangle .$$

Grazie all'ipotesi induttiva ed alla linearità del prodotto scalare, abbiamo

$$\left\| \sum_{i=1}^{n-1} \mathbf{v_i} \right\|^2 + \|\mathbf{v_n}\|^2 + 2 \left\langle \sum_{i=1}^{n-1} \mathbf{v_i} \, , \mathbf{v_n} \right\rangle = \sum_{i=1}^{n-1} \|\mathbf{v_i}\|^2 + \|\mathbf{v_n}\|^2 + 2 \sum_{i=1}^{n-1} \langle \mathbf{v_i}, \mathbf{v_n} \rangle .$$

Dato che tutti i vettori sono ortogonali tra loro, segue la tesi. □

Per definire il concetto di proiezione ortogonale, è necessario prima introdurre il complemento ortogonale di un insieme di vettori.

Definizione 8.19. *Siano V uno spazio euclideo ed $U \subseteq V$. Allora il complemento ortogonale $U^\perp$ di U è l'insieme dei vettori ortogonali a tutti i vettori di U, ovvero*

$$U^\perp = \{\mathbf{v} \in V \mid \langle \mathbf{v}, \mathbf{u} \rangle = 0 \text{ per ogni } \mathbf{u} \in U\}.$$

In uno spazio euclideo V finitamente generato, individuare il complemento ortogonale di un insieme di vettori è equivalente a risolvere un sistema lineare. Infatti, i vettori $\mathbf{v}$ ortogonali ad $U = \{\mathbf{u_1}, \ldots, \mathbf{u_m}\}$ sono quelli per cui vale

$$\begin{cases} \langle \mathbf{u_1}, \mathbf{v} \rangle & = & 0, \\ \vdots & & \vdots \\ \langle \mathbf{u_m}, \mathbf{v} \rangle & = & 0. \end{cases}$$

Fissando una base $\mathcal{B} = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ di V il precedente sistema risulta equivalente a

$$\begin{cases} \mathbf{u_1}|_{\mathcal{B}}^T * G|_{\mathcal{B}} * \mathbf{v}|_{\mathcal{B}} & = & 0, \\ \vdots & & \vdots \\ \mathbf{u_m}|_{\mathcal{B}}^T * G|_{\mathcal{B}} * \mathbf{v}|_{\mathcal{B}} & = & 0 \end{cases} \qquad \Rightarrow \qquad A^T * G|_{\mathcal{B}} * \mathbf{v}|_{\mathcal{B}} = 0_{m1},$$

dove

$$A = \left[\begin{array}{c|c|c} \mathbf{u_1}|_{\mathcal{B}} & \ldots & \mathbf{u_m}|_{\mathcal{B}} \end{array} \right].$$

Quindi il computo di $U^\perp$ si risolve individuando le soluzioni del sistema omogeneo $\left[A^T * G|_{\mathcal{B}} \mid 0_{m1} \right]$.

Esempio 8.20.

- In $\mathbb{R}^3$ cerchiamo il complemento ortogonale di $U = \{(1,1,0),(1,0,1)\}$ rispetto al prodotto euclideo. Scegliendo la base canonica $\mathcal{B}_3 = \{\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}\}$, abbiamo

$$A^T * G|_{\mathcal{B}_3} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} * \mathbb{I}_n = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \xrightarrow{R_2 - R_1 \to R_2} \begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}.$$

Quindi $U^\perp = \{(-t,t,t)\}$, dove $t \in \mathbb{R}$.
- Cerchiamo il complemento ortogonale dell'insieme $U = \{1, x^2\} \subset \mathbb{R}[x]_2$, rispetto al prodotto L_2 con $I = [0,1]$. Rispetto alla base canonica $\mathcal{B}_3 = \{1, x, x^2\}$, abbiamo

$$A^T * G|_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix} = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/3 & 1/4 & 1/5 \end{bmatrix} \xrightarrow{R_2 - \frac{1}{3}R_1 \to R_2} \begin{bmatrix} 1 & 1/2 & 1/3 \\ 0 & 1/12 & 4/45 \end{bmatrix}.$$

Quindi le componenti dei vettori ortogonali sono

$$\mathbf{v}|_{\mathcal{B}_3} = t \cdot \begin{bmatrix} 3 \\ -16 \\ 15 \end{bmatrix},$$

da cui $U^\perp = \{t \cdot (3 - 16x + 15x^2)\}$, dove $t \in \mathbb{R}$.

Proposizione 8.21. *Siano V uno spazio euclideo ed $U \subseteq V$. Chiamiamo $\bar{U}$ il sottospazio di V generato da U. Allora:*

i) *$U^\perp$ è un sottospazio di V;*
ii) *$U^\perp = \bar{U}^\perp$;*
iii) *se V è finitamente generato vale $V = \bar{U} \oplus U^\perp$.*

DIMOSTRAZIONE.

i) Proviamo l'enunciato utilizzando il Lemma 4.14. Per prima cosa $U^\perp$ non è vuoto perché contiene il vettore $\mathbf{0}$, che come sappiamo è ortogonale ad ogni $\mathbf{v} \in V$. Inoltre, siano $t_1, t_2 \in \mathbb{R}$ e $\mathbf{v_1}, \mathbf{v_2} \in U^\perp$. Allora per ogni $\mathbf{u} \in U$ vale

$$\langle t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2}, \mathbf{u} \rangle = t_1 \langle \mathbf{v_1}, \mathbf{u} \rangle + t_2 \langle \mathbf{v_2}, \mathbf{u} \rangle = 0 + 0 = 0.$$

Quindi $t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2} \in U^\perp$, da cui la tesi.

ii) Per definizione, i vettori in $U^\perp$ sono perpendicolari a tutti i vettori $\mathbf{u} \in U$. Di conseguenza sono perpendicolari anche a tutte le combinazioni lineari di questi vettori, per via della linearità del prodotto scalare. Questo significa che $U^\perp \subseteq \bar{U}^\perp$. Viceversa, i vettori in $\bar{U}^\perp$ sono perpendicolari a tutti i vettori in $\bar{U}$ e quindi anche ai vettori in U, dato che $U \subseteq \bar{U}$. Di conseguenza $\bar{U}^\perp \subseteq U^\perp$, e quindi $U^\perp = \bar{U}^\perp$.

iii) Per prima cosa, osserviamo che un vettore $\mathbf{v} \in \bar{U} \cap U^\perp$ deve essere perpendicolare a se stesso, ma questo implica che $\mathbf{v} = \mathbf{0}$. Di conseguenza $\bar{U} \cap U^\perp = \{\mathbf{0}\}$. Rimane da dimostrare che $V = \bar{U} + U^\perp$, e per farlo sfruttiamo il fatto che V è uno spazio avente $\dim(V) = n < \infty$. In tal caso, esistono una base $\mathcal{B}$ di V ed una base $\mathcal{B}_{\bar{U}} = \{\mathbf{u_1}, \ldots, \mathbf{u_m}\}$ di $\bar{U}$, con $m \leq n$. Definiamo

$$A = \left[\begin{array}{c|c|c} \mathbf{u_1}|_{\mathcal{B}} & \ldots & \mathbf{u_m}|_{\mathcal{B}} \end{array} \right] \in \mathrm{Mat}(n, m; \mathbb{R})$$

e $G|_{\mathcal{B}}$ la matrice di Gram. Per quanto precedentemente illustrato, sappiamo che la mappa delle coordinate $\phi_{\mathcal{B}}$ è un isomorfismo tra $U^\perp$ e $\ker(A^T * G|_{\mathcal{B}})$. Inoltre, la Proposizione 8.14 garantisce che la matrice di Gram associata alla base $\mathcal{B}$ è invertibile, quindi a sua volta $\ker(A^T * G|_{\mathcal{B}})$ è isomorfo a $\ker(A^T)$. In conclusione, abbiamo

$$\dim(U^\perp) = \dim(\ker(A^T)) = n - r(A^T) = n - m = \dim(V) - \dim(\bar{U}),$$

dove abbiamo utilizzato il fatto che $\mathcal{B}_{\bar{U}}$ è una base e quindi $r(A^T) = r(A) = m = \dim(\bar{U})$. L'uguaglianza precedente implica

$$\dim(V) = \dim(\bar{U}) + \dim(U^\perp).$$

Tale relazione è sufficiente a dimostrare la tesi, per via del Corollario 4.52. □

Corollario 8.22. *Siano V uno spazio euclideo finitamente generato ed $U \subseteq V$. Se $\bar{U}$ è il sottospazio di V generato da U, allora $(U^\perp)^\perp = \bar{U}$.*

Dimostrazione. Usiamo i risultati della proposizione precedente. Dobbiamo dimostrare che $(\bar{U}^\perp)^\perp = \bar{U}$. Per definizione, i vettori in $\bar{U}$ sono ortogonali ai vettori in $\bar{U}^\perp$, quindi $\bar{U} \subseteq (\bar{U}^\perp)^\perp$. Inoltre, sappiamo che $V = \bar{U} \oplus \bar{U}^\perp = \bar{U}^\perp \oplus (\bar{U}^\perp)^\perp$ e quindi

$$\dim(V) = \dim(\bar{U}) + \dim(\bar{U}^\perp) = \dim(\bar{U}^\perp) + \dim((\bar{U}^\perp)^\perp)\,,$$

da cui $\dim(\bar{U}) = \dim((\bar{U}^\perp)^\perp)$. A questo punto la tesi segue applicando la Proposizione 4.46. □

Ad ogni decomposizione di uno spazio vettoriale come somma diretta di suoi sottospazi è associata in modo naturale una proiezione. La proiezione corrispondente a $V = U \oplus U^\perp$ gode di proprietà speciali, che andremo ora a presentare.

Definizione 8.23. *Siano V uno spazio euclideo finitamente generato ed $U \subseteq V$ un sottospazio. Allora:*

- *la proiezione ortogonale su U è la funzione $\mathcal{P}_{U;U^\perp}$, che indichiamo semplicemente con $\mathcal{P}_U$;*
- *la riflessione ortogonale rispetto ad U è la funzione $\mathcal{R}_{U;U^\perp}$, che indichiamo con $\mathcal{R}_U$.*

Per costruzione, per ogni $\mathbf{v} \in V$ abbiamo che $\mathcal{P}_U(\mathbf{v})$ è un vettore appartenente ad U mentre $\mathbf{v} - \mathcal{P}_U(\mathbf{v})$ appartiene al complemento ortogonale $U^\perp$, quindi in particolare abbiamo $\mathbf{v} - \mathcal{P}_U(\mathbf{v}) \perp \mathcal{P}_U(\mathbf{v})$. Infine, la seguente proprietà metrica caratterizza la proiezione ortogonale.

Proposizione 8.24. *Siano V uno spazio euclideo finitamente generato ed $U \subseteq V$ un sottospazio. Per ogni $\mathbf{v} \in V$ vale*

$$\mathcal{P}_U(\mathbf{v}) = \operatorname*{argmin}_{\mathbf{u} \in U} \|\mathbf{v} - \mathbf{u}\|\,.$$

Dimostrazione. Dobbiamo dimostrare che vale

$$\|\mathbf{v} - \mathbf{u}\| > \|\mathbf{v} - \mathcal{P}_U(\mathbf{v})\|$$

per ogni $\mathbf{u} \in U$ diverso da $\mathcal{P}_U(\mathbf{v})$. Consideriamo la funzione

$$\|\mathbf{v} - \mathbf{u}\|^2 = \|(\mathbf{v} - \mathcal{P}_U(\mathbf{v})) + (\mathcal{P}_U(\mathbf{v}) - \mathbf{u})\|^2\,,$$

il cui argomento di minimo è ovviamente uguale a quello della norma. A destra dell'uguale abbiamo $\mathbf{v} - \mathcal{P}_U(\mathbf{v}) \in U^\perp$ e $\mathcal{P}_U(\mathbf{v}) - \mathbf{u} \in U$. Di conseguenza i due vettori sono ortogonali e possiamo quindi utilizzare il Teorema di Pitagora:

$$\|\mathbf{v} - \mathbf{u}\|^2 = \|\mathbf{v} - \mathcal{P}_U(\mathbf{v})\|^2 + \|\mathcal{P}_U(\mathbf{v}) - \mathbf{u}\|^2 > \|\mathbf{v} - \mathcal{P}_U(\mathbf{v})\|^2$$

se $\mathbf{u} \neq \mathcal{P}_U(\mathbf{v})$. □

Esempio 8.25. Per ogni $h \in \mathbb{R}$, calcoliamo la proiezione ortogonale di $\mathbf{v} = (h, 1, 1)$ sul sottospazio $U = \{(x, y, z) \in \mathbb{R}^3 \,|\, x + y + z = 0\}$, rispetto al prodotto scalare euclideo. Per prima cosa dobbiamo individuare una base di U ed una di $U^\perp$. Il sottospazio U è definito attraverso una rappresentazione algebrica, per trovare una base calcoliamo una sua rappresentazione parametrica:

$$x + y + z = 0 \qquad \Rightarrow \qquad (x, y, z) = t_1 \cdot (-1, 1, 0) + t_2(-1, 0, 1).$$

Di conseguenza $\mathcal{B}_U = \{(-1, 1, 0), (-1, 0, 1)\}$. Il complemento ortogonale $U^\perp$ si calcola risolvendo il sistema lineare

$$\begin{cases} -x + y &= 0 \\ -x + z &= 0 \end{cases} \qquad \Rightarrow \qquad (x, y, z) = t \cdot (1, 1, 1).$$

Quindi $\mathcal{B}_{U^\perp} = \{(1, 1, 1)\}$. Per le proprietà di somma diretta, una base di V è $\mathcal{B} = \mathcal{B}_U \cup \mathcal{B}_{U^\perp} = \{(-1, 1, 0), (-1, 0, 1), (1, 1, 1)\}$. La proiezione ortogonale di $\mathbf{v}$ si ottiene calcolando la sua decomposizione rispetto alla base $\mathcal{B}$. Utilizzando le solite tecniche, si ricava

$$\mathbf{v} = \frac{1-h}{3} \cdot (-1, 1, 0) + \frac{1-h}{3} \cdot (-1, 0, 1) + \frac{h+2}{3} \cdot (1, 1, 1).$$

Per definizione, $\mathcal{P}_U(\mathbf{v}) = \dfrac{1-h}{3} \cdot (-2, 1, 1)$. In particolare:

- se $h = 1$ allora $\mathcal{P}_U(\mathbf{v}) = \mathbf{0}$, quindi il vettore $\mathbf{v}$ appartiene ad $U^\perp$;
- se $h = -2$ allora $\mathcal{P}_U(\mathbf{v}) = \mathbf{v}$, pertanto il vettore $\mathbf{v}$ appartiene ad U.

8.4. Basi ortonormali, proiezione ortogonale ed algoritmo di Gram-Schmidt

Come sappiamo, in uno spazio vettoriale possono esistere basi più comode di altre dal punto di vista computazionale, ma dal lato teorico esse sono tutte da considerarsi perfettamente equivalenti tra loro. Al contrario, negli spazi euclidei esiste una classe naturalmente privilegiata di basi, ovvero quelle formate da vettori ortonormali. Esse godono infatti di alcune proprietà speciali.

Proposizione 8.26. *In uno spazio vettoriale euclideo V finitamente generato, siano $U \subseteq V$ un sottospazio e $\mathcal{B} = \{\mathbf{u_1}, \ldots, \mathbf{u_m}\}$ una sua base ortonormale. Allora la proiezione ortogonale di un vettore $\mathbf{v} \in V$ su U è uguale a*

$$\mathcal{P}_U(\mathbf{v}) = \sum_{i=1}^{m} \langle \mathbf{v}, \mathbf{u_i} \rangle \cdot \mathbf{u_i}\,. \tag{13}$$

Dimostrazione. Siano $U^\perp$ il complemento ortogonale di U e $\mathcal{B}^\perp = \{\tilde{\mathbf{u}}_\mathbf{1}, \ldots, \tilde{\mathbf{u}}_\mathbf{n}\}$ una sua base. Allora $\mathcal{B} \cup \mathcal{B}^\perp$ è una base di V. Per il teorema di interpolazione, ogni applicazione lineare è univocamente determinata dalla sua azione su una base del dominio. Per definizione di proiezione ortogonale, valgono

$$\begin{cases} \mathcal{P}_U(\mathbf{u_j}) = \mathbf{u_j} & \text{per} \quad 1 \le j \le m, \\ \mathcal{P}_U(\tilde{\mathbf{u}}_\mathbf{j}) = \mathbf{0} & \text{per} \quad 1 \le j \le n. \end{cases}$$

D'altra parte, abbiamo

$$\begin{cases} \displaystyle\sum_{i=1}^{m} \langle \mathbf{u_j}, \mathbf{u_i} \rangle \cdot \mathbf{u_i} = \sum_{i=1}^{m} \delta_{ji} \cdot \mathbf{u_i} = \mathbf{u_j} & \text{per} \quad 1 \le j \le m, \\ \displaystyle\sum_{i=1}^{m} \langle \tilde{\mathbf{u}}_\mathbf{j}, \mathbf{u_i} \rangle \cdot \mathbf{u_i} = \sum_{i=1}^{m} 0 \cdot \mathbf{u_i} = \mathbf{0} & \text{per} \quad 1 \le i \le n, \end{cases}$$

da cui segue la tesi. □

Corollario 8.27. *Siano V uno spazio vettoriale euclideo finitamente generato e $\mathcal{B} = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ una base ortonormale. Allora per ogni $\mathbf{v} \in V$ vale la decomposizione*

$$\mathbf{v} = \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{v_i} \rangle \cdot \mathbf{v_i} \, .$$

Dimostrazione. La proiezione di un vettore $\mathbf{v} \in V$ sullo spazio V stesso coincide con la funzione identità. Di conseguenza la tesi segue dalla Proposizione 8.26. □

Il Corollario 8.27 permette di calcolare le componenti di un vettore rispetto ad una base ortonormale attraverso il prodotto scalare del vettore con gli elementi della base stessa. Specialmente nella letteratura riguardante l'Analisi Funzionale, tali componenti sono anche note come Coefficienti di Fourier del vettore rispetto alla base ortonormale scelta.

Esempio 8.28. In $V = \mathrm{Mat}(2,1;\mathbb{R})$ consideriamo il prodotto scalare $\langle \ , \ \rangle_G$ definito dalla matrice

$$G = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}.$$

Una base ortonormale di V è

$$\mathcal{B} = \left\{ A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, B = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}.$$

Infatti valgono

$$\langle A, A \rangle_G = A^T * G * A = 1, \qquad \langle B, B \rangle_G = B^T * G * B = 1, \qquad \langle A, B \rangle_G = A^T * G * B = 0.$$

Quindi la decomposizione della generica colonna

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathrm{Mat}(2,1;\mathbb{R})$$

rispetto alla base $\mathcal{B}$ è

$$X = \langle X, A \rangle_G \cdot A + \langle X, B \rangle_G \cdot B = (X^T * G * A) \cdot A + (X^T * G * B) \cdot B = \frac{x_1 + x_2}{2} \cdot A + \frac{x_1 - x_2}{2} \cdot B \, .$$

Finora abbiamo verificato l'esistenza delle basi ortonormali in alcuni esempi specifici. La dimostrazione della loro esistenza per qualsiasi spazio euclideo è costruttiva, in quanto è basata sul cosiddetto Algoritmo di Gram-Schmidt, che permette di costruire una base ortonormale a partire da una base generica.

Lemma 8.29. *Siano V uno spazio vettoriale euclideo ed $U = \{\mathbf{u_1}, \dots, \mathbf{u_m}\} \subseteq V$ un insieme ortogonale non contenente il vettore nullo. Allora U è linearmente indipendente.*

Dimostrazione. La matrice di Gram $G|_U$ è diagonale, con tutti gli elementi lungo la diagonale diversi da zero. Pertanto il suo determinante è diverso da zero e quindi l'insieme U è indipendente, grazie alla Proposizione 8.14. □

Teorema 8.30. *Siano V uno spazio vettoriale euclideo finitamente generato ed $U = \{\mathbf{u_1}, \dots, \mathbf{u_m}\} \subseteq V$ un insieme linearmente indipendente. Chiamiamo $U_i = \mathcal{L}(\mathbf{u_1}, \dots, \mathbf{u_i})$, per $1 \leq i \leq m$. Allora l'insieme $\tilde{U} = \{\tilde{\mathbf{u}}_\mathbf{1}, \dots, \tilde{\mathbf{u}}_\mathbf{m}\}$ formato dai vettori*

$$\begin{aligned} \tilde{\mathbf{u}}_\mathbf{1} &= \mathbf{u_1} \\ \tilde{\mathbf{u}}_\mathbf{i} &= \mathbf{u_i} - \mathcal{P}_{U_{i-1}}(\mathbf{u_i}) \qquad \textit{per} \qquad 2 \leq i \leq m \end{aligned}$$

è linearmente indipendente ed ortogonale.

Dimostrazione. Per prima cosa, osserviamo che per costruzione abbiamo:

i) $U_j \subset U_i$ per ogni $1 \leq j < i \leq m$, poiché ogni U_i è indipendente;

ii) $\mathbf{\tilde{u}_i}$ è ortogonale ad U_{i-1} per ogni $2 \leq i \leq m$;

iii) $\mathbf{\tilde{u}_i}$ è contenuto in U_i, per ogni $1 \leq i \leq m$.

Dopo aver definito $\tilde{U}_i = \{\mathbf{\tilde{u}_1}, \ldots, \mathbf{\tilde{u}_i}\}$ per $1 \leq i \leq m$, la dimostrazione del teorema si compie provando per induzione su i che $\tilde{U}_i$ è indipendente ed ortogonale.

Il caso $i = 1$ è ovvio, in quanto $\mathbf{\tilde{u}_1} = \mathbf{u_1} \neq \mathbf{0}$ e quindi $\tilde{U}_1$ è indipendente. Inoltre, è anche ortogonale poiché non vi sono altri vettori in $\tilde{U}_1$ tali per cui il prodotto scalare di questi con $\mathbf{\tilde{u}_1}$ sia diverso da zero.

Supponendo ora che la tesi sia vera fino ad $i-1$, verifichiamola per i. Per l'ipotesi induttiva, i primi $i-1$ vettori di $\tilde{U}_i$ sono indipendenti ed ortogonali, quindi rimane da studiare il comportamento del solo vettore $\mathbf{\tilde{u}_i}$. Per le osservazioni iniziali (i) e (iii), sappiamo che

$$\tilde{U}_{i-1} = \{\mathbf{\tilde{u}_1}, \ldots, \mathbf{\tilde{u}_{i-1}}\} \subseteq \bigcup_{j=1}^{i-1} U_j = U_{i-1}.$$

Utilizzando l'osservazione (ii), segue quindi che $\mathbf{\tilde{u}_i}$ è ortogonale ad $\tilde{U}_{i-1}$ e perciò $\tilde{U}_i$ è un insieme ortogonale. Grazie al Lemma 8.29, per dimostrare l'indipendenza di $\tilde{U}_i$ è quindi sufficiente verificare che $\mathbf{\tilde{u}_i} \neq \mathbf{0}$. Se per assurdo assumessimo $\mathbf{\tilde{u}_i} = \mathbf{0}$, questo significherebbe che $\mathbf{u_i} = \mathcal{P}_{U_{i-1}}(\mathbf{u_i}) \in U_{i-1}$, ovvero che $\mathbf{u_i}$ è una combinazione lineare di $\mathbf{u_1}, \ldots, \mathbf{u_{i-1}}$. Ma questo è impossibile, poiché per ipotesi U è un insieme linearmente indipendente. □

Seguendo le notazioni del teorema precedente, è possibile ora enunciare il teorema di esistenza delle basi ortonormali.

Corollario 8.31. *Siano V uno spazio vettoriale euclideo finitamente generato e $\mathcal{B} = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\} \subseteq V$ una sua base. Allora l'insieme*

$$\tilde{\mathcal{B}} = \left\{ \frac{\mathbf{\tilde{v}_1}}{\|\mathbf{\tilde{v}_1}\|}, \ldots, \frac{\mathbf{\tilde{v}_n}}{\|\mathbf{\tilde{v}_n}\|} \right\}$$

è una base ortonormale di V.

Dimostrazione. Per ipotesi, la dimensione di V è n. Utilizzando l'Algoritmo di Gram-Schmidt, otteniamo che l'insieme $\{\mathbf{\tilde{v}_1}, \ldots, \mathbf{\tilde{v}_n}\}$ è ortogonale e linearmente indipendente, quindi è una base ortogonale di V. Normalizzandola, otteniamo la base ortonormale $\tilde{\mathcal{B}}$. □

Utilizzando la Proposizione 8.26, possiamo dare la formulazione classica dell'Algoritmo di Gram-Schmidt, che è quella effettivamente implementabile a livello computazionale. Data la base di partenza $\mathcal{B} = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$, la base ortonormale finale si ottiene calcolando per prima cosa la base ortogonale composta da

$$\begin{aligned} \mathbf{\tilde{v}_1} &= \mathbf{v_1}\,, \\ \mathbf{\tilde{v}_i} &= \mathbf{v_i} - \sum_{j=1}^{i-1} \frac{\langle \mathbf{v_i}, \mathbf{\tilde{v}_j} \rangle}{\|\mathbf{\tilde{v}_j}\|^2} \cdot \mathbf{\tilde{v}_j} \qquad \text{per} \qquad 2 \leq i \leq n\,. \end{aligned}$$

Infine, si ricava la base $\tilde{\mathcal{B}}$ normalizzando ogni singolo vettore $\mathbf{\tilde{v}_i}$, per $1 \leq i \leq n$.

Esempio 8.32. Riprendiamo l'Esempio 8.25 e risolviamolo utilizzando le tecniche che abbiamo imparato in questa sezione. Per calcolare la proiezione di $\mathbf{v} = (h, 1, 1)$ sul sottospazio $U = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$ è necessario conoscere una base ortonormale di U. Risolvendo il sistema lineare, sappiamo che una base del sottospazio è $\mathcal{B}_U = \{\mathbf{u_1} = (-1, 1, 0), \mathbf{u_2} = (-1, 0, 1)\}$, ma è facile verificare che essa non è ortonormale rispetto al prodotto euclideo. Possiamo però applicare a $\mathcal{B}_U$ l'Algoritmo di Gram-Schmidt:

$$\begin{aligned}
\tilde{\mathbf{u}}_\mathbf{1} &= \mathbf{u}_\mathbf{1} = (-1,1,0),\\
\tilde{\mathbf{u}}_\mathbf{2} &= \mathbf{u}_\mathbf{2} - \frac{\langle \mathbf{u}_\mathbf{2}, \tilde{\mathbf{u}}_\mathbf{1}\rangle}{\|\tilde{\mathbf{u}}_\mathbf{1}\|^2}\cdot \tilde{\mathbf{u}}_\mathbf{1} = (-1,0,1) - \frac{1}{2}\cdot(-1,1,0) = \left(-\frac{1}{2}, -\frac{1}{2}, 1\right).
\end{aligned}$$

Dopo aver normalizzato i due vettori otteniamo:

$$\tilde{\mathcal{B}} = \left\{ \mathbf{u}'_\mathbf{1} = \frac{\tilde{\mathbf{u}}_\mathbf{1}}{\|\tilde{\mathbf{u}}_\mathbf{1}\|}, \mathbf{u}'_\mathbf{2} = \frac{\tilde{\mathbf{u}}_\mathbf{2}}{\|\tilde{\mathbf{u}}_\mathbf{2}\|} \right\} = \left\{ \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right) \right\}.$$

Infine, la proiezione ortogonale di $\mathbf{v}$ su U è

$$\mathcal{P}_U(\mathbf{v}) = \langle \mathbf{v}, \mathbf{u}'_\mathbf{1}\rangle \cdot \mathbf{u}'_\mathbf{1} + \langle \mathbf{v}, \mathbf{u}'_\mathbf{2}\rangle \cdot \mathbf{u}'_\mathbf{2} = \frac{1-h}{2}\cdot(-1,1,0) + \frac{1-h}{6}\cdot(-1,-1,2) = \frac{1-h}{3}\cdot(-2,1,1).$$

Osservazione 8.33. Benché la nostra definizione della proiezione ortogonale $\mathcal{P}_U$ sia valida solo all'interno di uno spazio euclideo V finitamente generato, grazie alla Proposizione 8.26 essa può essere estesa anche al caso in cui U è un sottospazio finitamente generato di uno spazio V di dimensione infinita. Infatti, in questa situazione esiste una base ortonormale $\mathcal{B} = \{\mathbf{u}_\mathbf{1}, \dots, \mathbf{u}_\mathbf{m}\}$ di U, per cui possiamo utilizzare la formula (13) come definizione di $\mathcal{P}_U$. Non è difficile dimostrare che con questa scelta continuano a valere tutte le proprietà della proiezione che abbiamo verificato nel caso di dimensione finita. Quindi d'ora in poi non considereremo più la restrizione agli spazi euclidei finitamente generati per asserire l'esistenza della proiezione ortogonale su un sottospazio di dimensione finita.

8.5. Angolo tra vettori

Grazie ai risultati delle precedenti sezioni, possiamo dimostrare due importanti relazioni coinvolgenti prodotti scalari e norme di vettori, note in letteratura come disuguaglianza di Schwarz e disuguaglianza triangolare.

Proposizione 8.34. *In uno spazio vettoriale euclideo V, siano $\mathbf{u}, \mathbf{v} \in V$. Allora:*

i) $|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|$ *e l'uguaglianza vale se e solo se* $\mathbf{v} = t\cdot\mathbf{u}$ *con* $t \in \mathbb{R}$, *oppure* $\mathbf{u} = \mathbf{0}$;

ii) $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ *e l'uguaglianza vale se e solo se* $\mathbf{v} = t\cdot\mathbf{u}$ *con* $t \in [0, +\infty]$, *oppure* $\mathbf{u} = \mathbf{0}$.

Dimostrazione. Se $\mathbf{u} = \mathbf{0}$, a causa della Proposizione 8.6 vale l'uguaglianza per entrambe le relazioni. In caso contrario, chiamiamo $U = \mathcal{L}(\mathbf{u})$. Allora

$$\begin{aligned}
\|\mathbf{v}\|^2 &= \|\mathcal{P}_U(\mathbf{v}) + (\mathbf{v} - \mathcal{P}_U(\mathbf{v}))\|^2 = \|\mathcal{P}_U(\mathbf{v})\|^2 + \|\mathbf{v} - \mathcal{P}_U(\mathbf{v})\|^2 \geq \|\mathcal{P}_U(\mathbf{v})\|^2 =\\
&= \left\| \frac{\langle \mathbf{v}, \mathbf{u}\rangle}{\|\mathbf{u}\|^2}\cdot\mathbf{u} \right\|^2 = \frac{\langle \mathbf{v}, \mathbf{u}\rangle^2}{\|\mathbf{u}\|^4}\,\|\mathbf{u}\|^2 = \frac{\langle \mathbf{v}, \mathbf{u}\rangle^2}{\|\mathbf{u}\|^2},
\end{aligned}$$

da cui segue $\|\mathbf{u}\|^2\,\|\mathbf{v}\|^2 \geq \langle \mathbf{v}, \mathbf{u}\rangle^2$ e quindi la tesi. L'uguaglianza è vera se e solo se $\mathbf{v} - \mathcal{P}_U(\mathbf{v}) = \mathbf{0}$, cioè se e solo se $\mathbf{v} \in U$ o equivalentemente $\mathbf{v} = t\cdot\mathbf{u}$.

Grazie all'appena dimostrata disuguaglianza di Schwarz, possiamo verificare anche la disuguaglianza triangolare:

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle &\leq \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\,|\langle \mathbf{u}, \mathbf{v}\rangle|\\
&\leq \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\,\|\mathbf{u}\|\,\|\mathbf{v}\| = (\|\mathbf{u}\| + \|\mathbf{v}\|)^2.
\end{aligned}$$

L'uguaglianza vale se e solo se valgono le condizioni per l'uguaglianza nella relazione di Schwarz ed in aggiunta $\langle \mathbf{u}, \mathbf{v}\rangle \geq 0$. La combinazione di queste condizioni impone che

$$\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, t\cdot\mathbf{u}\rangle = t\,\|\mathbf{u}\|^2 \geq 0,$$

da cui $t \geq 0$. $\square$

Se $\mathbf{u}$ e $\mathbf{v}$ sono entrambi non nulli, la disuguaglianza di Schwarz può essere espressa anche nella seguente forma:

$$-1 \leq \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \, \|\mathbf{v}\|} \leq 1.$$

Grazie a questa osservazione, siamo ora pronti a definire l'angolo compreso tra due vettori.

Definizione 8.35. *In uno spazio vettoriale euclideo V, siano $\mathbf{u}, \mathbf{v} \in V$ entrambi non nulli. Allora, l'angolo tra i due vettori è*

$$\widehat{\mathbf{uv}} = \arccos\left(\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \, \|\mathbf{v}\|} \right).$$

L'angolo $\widehat{\mathbf{uv}}$ è ben definito ed è una quantità compresa nell'intervallo $[0, \pi]$. Naturalmente esso dipende dalla struttura euclidea associata allo spazio vettoriale. In ogni caso abbiamo che:

- $\widehat{\mathbf{uv}} = 0$ se e solo se $\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\| \, \|\mathbf{v}\|$, cioè se e solo se $\mathbf{v} = t \cdot \mathbf{u}$ con $t > 0$;
- $\widehat{\mathbf{uv}} = \pi$ se e solo se $\langle \mathbf{u}, \mathbf{v} \rangle = -\|\mathbf{u}\| \, \|\mathbf{v}\|$, cioè se e solo se $\mathbf{v} = t \cdot \mathbf{u}$ con $t < 0$;
- $\widehat{\mathbf{uv}} = \frac{\pi}{2}$ se e solo se $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, cioè se e solo se i due vettori sono ortogonali.

Osserviamo infine che, se sono note le norme di due vettori e l'angolo tra loro compreso, allora il loro prodotto scalare si può calcolare come

$$\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos \widehat{\mathbf{uv}}.$$

Esempio 8.36.

- In $\mathbb{R}^2$ calcoliamo l'angolo tra i vettori $\mathbf{e_1} = (1, 0)$ ed $\mathbf{x} = (x, y)$, rispetto al prodotto scalare euclideo:
$$\widehat{\mathbf{e_1 x}}_E = \arccos\left(\frac{x}{\sqrt{x^2 + y^2}} \right).$$
- In $\mathbb{R}^3$ calcoliamo l'angolo tra i vettori $\mathbf{u} = (0, k, 1)$ e $\mathbf{v} = (1, 1, 0)$ rispetto al prodotto scalare euclideo:
$$\widehat{\mathbf{uv}}_E = \arccos\left(\frac{k}{\sqrt{2(1 + k^2)}} \right).$$
In particolare, i due vettori sono ortogonali se e solo se $k = 0$, mentre $\widehat{\mathbf{uv}} = \frac{\pi}{4}$ per $k = 1$ e $\widehat{\mathbf{uv}} = \frac{3\pi}{4}$ per $k = -1$. Osserviamo che l'angolo dipende dalla struttura euclidea. Ad esempio, consideriamo il prodotto scalare definito dalla matrice di Gram
$$G|_{\mathcal{B}_3} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$
Allora
$$\widehat{\mathbf{uv}}_G = \arccos\left(\frac{1 + k}{\sqrt{3(1 + k^2)}} \right).$$
In questo caso i due vettori sono ortogonali se e solo se $k = -1$.
- In $V = \mathrm{Mat}(2, 2; \mathbb{R})$ calcoliamo l'angolo compreso tra le due matrici
$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad \text{e} \qquad B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

rispetto al prodotto scalare di Frobenius:

$$\widehat{AB}_F = \arccos\left(\frac{0}{\sqrt{20}}\right) = \frac{\pi}{2}.$$

In generale, non è difficile dimostrare che $\mathbb{S}(n;\mathbb{R})^\perp = \mathbb{A}(n;\mathbb{R})$.

- In $V = \mathbb{R}[x]_1$ calcoliamo l'angolo tra $P(x) = 1$ e $Q(x) = x$, rispetto al prodotto scalare L_2 con $I = [0,1]$:

$$\widehat{PQ}_I = \arccos\left(\frac{1/2}{1/\sqrt{3}}\right) = \frac{\pi}{6}.$$

Questo dimostra che la base canonica di V non è ortogonale. Lo diventa se calcoliamo l'integrale su intervalli simmetrici $I = [-a,a]$.

- In $V = \mathrm{Mat}(2,1;\mathbb{R})$, consideriamo l'endomorfismo f_A associato alla matrice

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

Nell'Esempio 6.6 abbiamo osservato che f_A agisce ruotando tutti i vettori del medesimo angolo. Chiariamo meglio tale affermazione. Dato un generico vettore $B = a \cdot E_{11} + b \cdot E_{21}$, abbiamo $f_A(B) = -b \cdot E_{11} + a \cdot E_{21}$. Rispetto al prodotto scalare di Frobenius, allora valgono

$$\begin{aligned} \|f_A(B)\|_F &= \sqrt{b^2 + a^2} = \|B\|_F \ , \\ \widehat{Bf_A(B)}_F &= \arccos\left(\frac{-ab + ab}{a^2 + b^2}\right) = \arccos(0) = \frac{\pi}{2} \ . \end{aligned}$$

8.6. Spazi vettoriali orientati ed angolo orientato

Come abbiamo visto nella sezione precedente, l'angolo θ tra due vettori $\mathbf{u}, \mathbf{v}$ è per definizione una quantità nell'intervallo $[0,\pi]$. Si dice che questo è l'angolo convesso compreso tra i due vettori. Questo non sempre è completamente soddisfacente. Ad esempio, fissato un versore di riferimento $\mathbf{u}$ in uno spazio bidimensionale V, è intuitivo pensare che ogni altro versore $\mathbf{v} \in V$ sia completamente determinato dall'angolo che esso forma con $\mathbf{u}$. Ma questo non è vero se utilizziamo la nozione di angolo convesso. Infatti, dato $U = \mathcal{L}(\mathbf{u})$, non è difficile verificare che

$$\langle \mathcal{R}_U(\mathbf{v}), \mathbf{u}\rangle = \langle \mathbf{v}, \mathbf{u}\rangle$$

e quindi

$$\widehat{\mathcal{R}_U(\mathbf{v})\mathbf{u}} = \widehat{\mathbf{v}\mathbf{u}}.$$

Utilizzando l'angolo convesso rispetto ad $\mathbf{u}$, non possiamo quindi distinguere un versore $\mathbf{v}$ dal suo riflesso $\mathcal{R}_U(\mathbf{v})$. Questo è equivalente a dire che utilizzando le nozioni fino ad ora introdotte, non possiamo definire il concetto di senso di rotazione oppure di destra e sinistra.

Per superare tali limitazioni è necessario introdurre la struttura di spazio vettoriale orientato.

Definizione 8.37. *Dato uno spazio vettoriale reale finitamente generato, la base $\mathcal{B}$ è orientata come $\mathcal{B}'$, se* $\det(M_{\mathcal{B}\mathcal{B}'}) > 0$.

La richiesta che lo spazio vettoriale sia reale garantisce che la disuguaglianza abbia senso, in quanto $\det(M_{\mathcal{B}\mathcal{B}'})$ è necessariamente un numero reale. Naturalmente è possibile generalizzare questa ipotesi, richiedendo semplicemente che il campo di supporto dello spazio vettoriale sia ordinato, come ad esempio il campo dei numeri razionali $\mathbb{Q}$.

Esempio 8.38.

- In ogni spazio vettoriale reale, due basi $\mathcal{B}$ e $\mathcal{B}'$ che differiscono solo per l'ordine di una coppia di vettori, non sono orientate nello stesso modo. Ad esempio, date $\mathcal{B} = \{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ e $\mathcal{B}' = \{\mathbf{v_2}, \mathbf{v_1}, \dots, \mathbf{v_n}\}$ abbiamo
$$M_{\mathcal{B}\mathcal{B}'} = \left[\begin{array}{cccc} \mathbf{v_1}|_{\mathcal{B}'} & \mathbf{v_2}|_{\mathcal{B}'} & \dots & \mathbf{v_n}|_{\mathcal{B}'} \end{array}\right] = \begin{bmatrix} 0 & 1 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = P(1,2),$$
da cui $\det(M_{\mathcal{B}\mathcal{B}'}) = \det(P(1,2)) = -1 < 0$.
- In $V = \mathrm{Mat}(n, 1; \mathbb{R})$, consideriamo l'automorfismo f_A associato ad una matrice invertibile $A \in \mathrm{Mat}(n, n; \mathbb{R})$. L'immagine rispetto ad f_A della base canonica $\mathcal{B}_{n1} = \{E_{11}, \dots, E_{n1}\}$ di V è la base $\mathcal{B}'_{n1} = \{f_A(E_{11}), \dots, f_A(E_{n1})\}$. È facile verificare che $M_{\mathcal{B}_n\mathcal{B}'_n} = A$. Ad esempio, per $n = 2$ siano
$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \qquad \text{e} \qquad B = \begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix}.$$
Rispetto al prodotto di Frobenius, sappiamo che f_A effettua una rotazione di $\frac{\pi}{2}$ di tutti i vettori di V. Il suo determinante è $\det(f_A) = \det(A) = 1 = \det(M_{\mathcal{B}_n\mathcal{B}'_n})$, di conseguenza $\mathcal{B}$ è orientata come $\mathcal{B}'_n$. Invece, f_B è una riflessione. Il suo determinante è $\det(f_B) = \det(B) = -1 = \det(M_{\mathcal{B}_n\mathcal{B}'_n})$ e di conseguenza $\mathcal{B}$ non è orientata come $\mathcal{B}'$.

La Definizione 8.37 stabilisce una relazione nell'insieme di tutte le possibili basi di spazio vettoriale reale, che chiamiamo relazione di orientazione.

Proposizione 8.39. *L'orientazione è una relazione di equivalenza sull'insieme delle basi di uno spazio vettoriale reale. Inoltre, per ogni spazio esistono esattamente due classi di orientazione.*

Dimostrazione. Verifichiamo le tre proprietà delle relazioni di equivalenza.

- Riflessività: una base $\mathcal{B}$ è orientata come se stessa, poiché in questo caso la matrice del cambiamento di base è $M_{\mathcal{B}\mathcal{B}} = \mathbb{I}_n$, il cui determinante è $\det(M_{\mathcal{B}\mathcal{B}}) = 1 > 0$.
- Simmetria: supponiamo che la base $\mathcal{B}$ sia orientata come $\mathcal{B}'$, ovvero $\det(M_{\mathcal{B}\mathcal{B}'}) > 0$. Allora vale $\det(M_{\mathcal{B}'\mathcal{B}}) = \det(M_{\mathcal{B}\mathcal{B}'}^{-1}) = \det(M_{\mathcal{B}\mathcal{B}'})^{-1} > 0$, pertanto anche $\mathcal{B}'$ è orientata come $\mathcal{B}$.
- Transitività: sia $\mathcal{B}$ orientata come $\mathcal{B}'$, che a sua volta supponiamo orientata come $\mathcal{B}''$. Allora, $\det(M_{\mathcal{B}\mathcal{B}''}) = \det(M_{\mathcal{B}'\mathcal{B}''} * M_{\mathcal{B}\mathcal{B}'}) = \det(M_{\mathcal{B}'\mathcal{B}''})\det(M_{\mathcal{B}\mathcal{B}'}) > 0$ e di conseguenza $\mathcal{B}$ è orientata come $\mathcal{B}''$.

Per la seconda proprietà enunciata, osserviamo che la classe di orientazione di una base $\mathcal{B}$ dello spazio V è per definizione l'insieme
$$[\,\mathcal{B}\,] = \{\tilde{\mathcal{B}} \text{ basi di } V \mid \det(M_{\mathcal{B}\tilde{\mathcal{B}}}) > 0\}.$$
Se $\mathcal{B}'$ non appartiene a $[\,\mathcal{B}\,]$, allora necessariamente $\det(M_{\mathcal{B}\mathcal{B}'}) < 0$, poiché la matrice del cambiamento di base è invertibile e quindi $\det(M_{\mathcal{B}\mathcal{B}'}) \neq 0$. La classe di orientazione di $\mathcal{B}'$ è
$$[\,\mathcal{B}'\,] = \{\tilde{\mathcal{B}} \text{ basi di } V \mid \det(M_{\mathcal{B}'\tilde{\mathcal{B}}}) > 0\}.$$
D'altra parte vale
$$\det(M_{\mathcal{B}'\tilde{\mathcal{B}}}) = \det(M_{\mathcal{B}\tilde{\mathcal{B}}} * M_{\mathcal{B}'\mathcal{B}}) = \det(M_{\mathcal{B}\tilde{\mathcal{B}}})\det(M_{\mathcal{B}'\mathcal{B}}),$$
pertanto
$$[\,\mathcal{B}'\,] = \{\tilde{\mathcal{B}} \text{ basi di } V \mid \det(M_{\mathcal{B}\tilde{\mathcal{B}}})\det(M_{\mathcal{B}'\mathcal{B}}) > 0\} = \{\tilde{\mathcal{B}} \text{ basi di } V \mid \det(M_{\mathcal{B}\tilde{\mathcal{B}}}) < 0\}.$$

Dovrebbe essere chiaro che $[\mathcal{B}] \cup [\mathcal{B}']$ costituisce l'insieme di tutte le basi di V, da cui segue la tesi. $\square$

Definizione 8.40. *In uno spazio vettoriale reale V finitamente generato, sia $\mathbb{B}$ l'insieme quoziente avente come elementi le due classi di orientazione di V. Allora:*

i) un'orientazione di V è una funzione biunivoca $\xi : \mathbb{B} \to \{-1, +1\}$;
ii) la struttura (V, ξ) si dice spazio vettoriale reale orientato;
iii) una base appartenente a $\xi^{-1}(1)$ si dice orientata positivamente;
iv) una base appartenente a $\xi^{-1}(-1)$ si dice orientata negativamente.

Un'orientazione corrisponde quindi all'assegnazione arbitraria degli aggettivi positiva e negativa alle due classi di orientazione delle basi. L'orientazione canonica di uno spazio vettoriale è quella per cui la base canonica di tale spazio è orientata positivamente. Come al solito, a meno di ambiguità, uno spazio orientato (V, ξ) viene indicato con il solo spazio di supporto V.

A questo punto possiamo dare la definizione di angolo orientato in uno spazio bidimensionale.

Definizione 8.41. *Siano $\mathbf{u}, \mathbf{v}$ due vettori non nulli di uno spazio vettoriale euclideo orientato V di dimensione 2. Allora:*

i) $\mathbf{v}$ si dice a sinistra (destra) di $\mathbf{u}$ se l'insieme $\{\mathbf{u}, \mathbf{v}\}$ è una base di V orientata positivamente (negativamente);
ii) l'angolo orientato tra $\mathbf{u}$ ed $\mathbf{v}$ è

$$\widehat{\widehat{\mathbf{u}\mathbf{v}}} = \begin{cases} \widehat{\mathbf{u}\mathbf{v}} & \textit{se } \mathbf{u} \textit{ è a sinistra di } \mathbf{v}, \textit{ oppure } \mathbf{v} = t \cdot \mathbf{u} \textit{ con } t \in \mathbb{R}^*; \\ -\widehat{\mathbf{u}\mathbf{v}} & \textit{se } \mathbf{u} \textit{ è a destra di } \mathbf{v}. \end{cases}$$

L'angolo orientato tra due vettori è quindi un numero reale θ nell'intervallo $(-\pi, \pi]$.

Esempio 8.42. Sia $V = \mathbb{R}^2$ lo spazio vettoriale con struttura euclidea ed orientazione canoniche. Calcoliamo l'angolo orientato tra i versori $\mathbf{e_1} = (1, 0)$ e $\mathbf{v} = (\cos\theta, \sin\theta)$, dove $\theta \in (-\pi, \pi]$. La matrice del cambiamento di base da $\mathcal{B}' = \{\mathbf{e_1}, \mathbf{v}\}$ a $\mathcal{B}_2 = \{\mathbf{e_1}, \mathbf{e_2}\}$ è

$$M_{\mathcal{B}'\mathcal{B}_2} = \begin{bmatrix} \mathbf{e_1}|_{\mathcal{B}_2} & \mathbf{v}|_{\mathcal{B}_2} \end{bmatrix} = \begin{bmatrix} 1 & \cos\theta \\ 0 & \sin\theta \end{bmatrix} \qquad \text{e quindi} \qquad \det(M_{\mathcal{B}'\mathcal{B}_2}) = \sin\theta.$$

Questo significa che $\mathbf{v}$ è rispettivamente a sinistra di $\mathbf{e_1}$ per $\theta \in (0, \pi)$ ed è a destra per $\theta \in (-\pi, 0)$. Invece, l'angolo tra i due vettori è

$$\widehat{\mathbf{e_1}\mathbf{v}}_E = \arccos\left(\frac{\cos\theta}{\sqrt{\cos^2\theta + \sin^2\theta}}\right) = \arccos(\cos\theta) = \begin{cases} \theta & \text{se} \quad \theta \in [0, \pi], \\ -\theta & \text{se} \quad \theta \in (-\pi, 0). \end{cases}$$

Combinando le due cose abbiamo

$$\widehat{\widehat{\mathbf{e_1}\mathbf{v}}}_E = \theta.$$

Come si può vedere, abbiamo stabilito una corrispondenza biunivoca tra i versori di V ed i corrispondenti angoli orientati. L'orientazione canonica di $\mathbb{R}^2$ è detta anche orientazione antioraria e l'angolo orientato viene anche detto angolo misurato in senso antiorario. Viceversa, l'orientazione opposta si dice orientazione oraria.

Non è possibile estendere la definizione di angolo orientato tra due vettori $\mathbf{u}, \mathbf{v}$ appartenenti ad uno spazio vettoriale orientato V di dimensione generica. Il problema è che l'orientazione di V non definisce automaticamente una orientazione canonica nel sottospazio $\mathcal{L}(\mathbf{u}, \mathbf{v})$. Ad esempio, sia $V = \mathbb{R}^3$ con struttura euclidea ed orientazione canoniche. Questa è anche nota come orientazione destrorsa di $\mathbb{R}^3$ ed una qualunque base orientata positivamente si dice terna destrorsa. Si potrebbe pensare di dichiarare

$\mathbf{v}$ a destra di $\mathbf{u}$ se esistesse un terzo vettore $\mathbf{w}$ tale che $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ sia una terna destrorsa. Consideriamo ad esempio i vettori $\mathbf{e_1}$ ed $\mathbf{e_2}$. La base canonica $\{\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}\}$ è orientata positivamente, quindi secondo la precedente definizione $\mathbf{e_2}$ si troverebbe a sinistra di $\mathbf{e_1}$. D'altra parte, anche $\{\mathbf{e_2}, \mathbf{e_1}, -\mathbf{e_3}\}$ è orientata positivamente, quindi in questo caso sarebbe il vettore $\mathbf{e_1}$ a trovarsi a sinistra di $\mathbf{e_2}$. Questo ci porta a concludere che non esiste un'orientazione naturale del sottospazio $\mathcal{L}(\mathbf{e_1}, \mathbf{e_2})$ che sia ereditata da quella di V.

8.7. Duale di Hodge e prodotto vettoriale

Negli spazi euclidei orientati è possibile definire un'ulteriore operazione tra vettori, che è molto importante in alcune applicazioni. A tal fine, abbiamo bisogno di alcuni risultati preliminari. Iniziamo fornendo la seguente caratterizzazione dei vettori.

Proposizione 8.43. *Sia V uno spazio vettoriale euclideo. Un vettore $\mathbf{u} \in V$ è completamente determinato da:*

i) il suo modulo $\|\mathbf{u}\|$;

ii) se $\mathbf{u} \neq \mathbf{0}$, la sua direzione, ovvero il sottospazio unidimensionale U a cui appartiene;

iii) se $\mathbf{u} \neq \mathbf{0}$, il suo verso, ovvero l'orientazione dell'insieme $\{\mathbf{u}\}$ rispetto ad una scelta dell'orientazione nel sottospazio U.

Dimostrazione. Se $\|\mathbf{u}\| = 0$ allora $\mathbf{u} = \mathbf{0}$. Altrimenti, $U = \mathcal{L}(\mathbf{u})$ è l'unico sottospazio unidimensionale di V che lo contiene. Fissiamo un'orientazione in U e scegliamo una sua base $\mathcal{B}_U = \{\tilde{\mathbf{u}}\}$ ortonormale e positivamente orientata. Il vettore $\mathbf{u}$ è univocamente individuato dalla sua decomposizione rispetto a $\mathcal{B}_U$, cioè da $t \in \mathbb{R}$ tale che $\mathbf{u} = t \cdot \tilde{\mathbf{u}}$. Grazie alla conoscenza della norma di $\mathbf{u}$ possiamo calcolare il modulo della componente t:

$$\|\mathbf{u}\| = \|t \cdot \tilde{\mathbf{u}}\| = |t| \, \|\tilde{\mathbf{u}}\| = |t|.$$

Infine, la conoscenza del verso di $\mathbf{u}$ ci permette di calcolare il segno della componente. Infatti, l'orientazione di $\{\mathbf{u}\}$ è data dal segno di

$$\det([\,\mathbf{u}|_{\mathcal{B}_U}\,]) = \det([\,t\,]) = t. \qquad \square$$

Il lettore che ha già incontrato i vettori in altri corsi, ad esempio nello studio della fisica, probabilmente riconoscerà nella precedente proposizione la definizione più familiare di tali oggetti. Notiamo però come essa sia tutt'altro che semplice da fornire seguendo un approccio assiomatico. Infatti, noi l'abbiamo potuta presentare solo dopo aver introdotto la struttura tutt'altro che banale di spazio vettoriale euclideo orientato. Dal punto di vista delle applicazioni, questa definizione di vettore risulta naturale in quanto è implicitamente assunto che lo spazio fisico tridimensionale presenti una struttura di questo tipo, perlomeno nell'ambito della fisica non relativistica.

In secondo luogo, è necessario il seguente risultato riguardante il gramiano di un insieme di vettori, che abbiamo introdotto nella Sezione 8.2.

Lemma 8.44. *Siano V uno spazio euclideo orientato ed $U = \{\mathbf{u_1}, \ldots, \mathbf{u_m}\} \subseteq V$. Allora il gramiano dell'insieme U è un numero reale non negativo, ovvero $G(U) \geq 0$. L'uguaglianza vale se e solo se U è linearmente dipendente.*

Dimostrazione. Per la Proposizione 8.14, vale $G(U) = 0$ se e solo se U è linearmente dipendente. Supponiamo invece che U sia indipendente e che quindi costituisca una base dello spazio $\mathcal{L}(U)$. Per costruzione questo è uno spazio vettoriale finitamente generato, pertanto esiste una sua base ortonormale

$\mathcal{B} = \{\mathbf{v_1}, \ldots, \mathbf{v_m}\}$. Definiamo quindi la matrice

$$A = \begin{bmatrix} \mathbf{u_1}|_{\mathcal{B}} & \ldots & \mathbf{u_m}|_{\mathcal{B}} \end{bmatrix} \in \mathrm{Mat}(m, m; \mathbb{R}).$$

Per le proprietà delle basi ortonormali, è facile verificare che

$$G|_U = A^T * A \qquad \Rightarrow \qquad \det(G|_U) = \det(A^T)\det(A) = \det(A)^2 > 0. \qquad \square$$

Definizione 8.45. *Sia V uno spazio vettoriale euclideo orientato di dimensione $2 \le n < \infty$. L'operazione duale di Hodge è l'operazione interna $(n-1)$–aria*

$$\begin{array}{rccc} * : & V^{n-1} & \longrightarrow & V \\ & (\mathbf{u_1}, \ldots, \mathbf{u_{n-1}}) & \longmapsto & \mathbf{u} \end{array}$$

che associa ad $U = \{\mathbf{u_1}, \ldots, \mathbf{u_{n-1}}\}$ il vettore $\mathbf{u}$ caratterizzato da:

i) modulo $\|\mathbf{u}\| = \sqrt{G(U)}$;

ii) se $\|\mathbf{u}\| \neq 0$, direzione $U^\perp$;

iii) se $\|\mathbf{u}\| \neq 0$, verso tale che l'insieme $\{\mathbf{u_1}, \ldots, \mathbf{u_{n-1}}, \mathbf{u}\}$ è una base positivamente orientata di V.

Verifichiamo subito che l'operazione precedente è ben definita. In primo luogo, grazie al Lemma 8.44 abbiamo $G(U) \ge 0$. Inoltre, se $G(U) \neq 0$ allora U è un insieme linearmente indipendente, nel cui caso $\dim(U^\perp) = 1$ e quindi $U^\perp$ definisce una direzione di V. Infine, $\{\mathbf{u_1}, \ldots, \mathbf{u_{n-1}}, \mathbf{u}\}$ è effettivamente una base di V per ogni $\mathbf{u} \in U^\perp \setminus \{\mathbf{0}\}$.

Esempio 8.46.

- Se $n = 2$, il duale di Hodge di $\mathbf{u}$ è il vettore $*\mathbf{u}$ ortogonale a $\mathbf{u}$, posto alla sua sinistra e di modulo
$$\|*\mathbf{u}\| = \sqrt{\det([\langle \mathbf{u}, \mathbf{u}\rangle])} = \|\mathbf{u}\|.$$
- Se $n = 3$ il duale di Hodge è un'operazione binaria, meglio conosciuta come prodotto vettoriale, la cui notazione usuale è
$$\mathbf{u} = *(\mathbf{u_1}, \mathbf{u_2}) = \mathbf{u_1} \wedge \mathbf{u_2}.$$
Attraverso la cosiddetta identità di Lagrange, si può calcolare il modulo del prodotto vettoriale in funzione dell'angolo convesso tra $\mathbf{u_1}$ e $\mathbf{u_2}$:
$$\begin{aligned} \|\mathbf{u_1} \wedge \mathbf{u_2}\|^2 &= \begin{vmatrix} \langle \mathbf{u_1}, \mathbf{u_1}\rangle & \langle \mathbf{u_1}, \mathbf{u_2}\rangle \\ \langle \mathbf{u_2}, \mathbf{u_1}\rangle & \langle \mathbf{u_2}, \mathbf{u_2}\rangle \end{vmatrix} = \|\mathbf{u_1}\|^2 \|\mathbf{u_2}\|^2 - \langle \mathbf{u_1}, \mathbf{u_2}\rangle^2 = \\ &= \|\mathbf{u_1}\|^2 \|\mathbf{u_2}\|^2 (1 - \cos^2 \widehat{\mathbf{u_1}\mathbf{u_2}}) = \|\mathbf{u_1}\|^2 \|\mathbf{u_2}\|^2 \sin^2 \widehat{\mathbf{u_1}\mathbf{u_2}}. \end{aligned}$$
Dato che $\widehat{\mathbf{u_1}\mathbf{u_2}} \in [0, \pi]$, segue che $\sin \widehat{\mathbf{u_1}\mathbf{u_2}} \ge 0$ e quindi
$$\|\mathbf{u_1} \wedge \mathbf{u_2}\| = \|\mathbf{u_1}\| \, \|\mathbf{u_2}\| \sin \widehat{\mathbf{u_1}\mathbf{u_2}},$$

La definizione che abbiamo dato di duale di Hodge è significativa dal punto di vista geometrico, ma è poco pratica dal lato operativo. Fortunatamente esiste una caratterizzazione di questa operazione che ci permette di calcolarla in maniera efficiente. Per proseguire in tale direzione, è necessario fornire preliminarmente la nozione di determinante formale. Ricordiamo che una matrice $A \in \mathrm{Mat}(m, n; \mathbb{K})$ è per definizione una funzione

$$\begin{array}{rccc} A: & M \times N & \longrightarrow & \mathbb{K} \\ & (i, j) & \longmapsto & a_{ij} \end{array}$$

dove $M = \{1, \ldots, m\}$ e $N = \{1, \ldots, n\}$. Il determinante è a sua volta una funzione che combina in maniera opportuna le entrate di A in modo da ottenere un unico elemento di $\mathbb{K}$. E' possibile pensare di generalizzare il concetto di matrice sostituendo al campo $\mathbb{K}$ un inseme d'arrivo generico. A questo punto, se gli elementi della matrice generalizzata A possono ancora essere sommati e moltiplicati tra loro, si può calcolare il determinante della matrice stessa, che chiamiamo determinante formale di A. Ad esempio, calcoliamo il determinante formale della matrice generalizzata

$$A = \begin{bmatrix} 1 & x \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \end{bmatrix} \qquad \Rightarrow \qquad \det(A) = 1 \cdot \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} - x \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -x & -1 \\ 1 & -x \end{bmatrix}.$$

Proposizione 8.47. *Siano V uno spazio vettoriale euclideo orientato di dimensione $2 \leq n < \infty$ e $\mathcal{B} = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ una sua base ortonormale positivamente orientata. Sia inoltre $U = \{\mathbf{u_1}, \dots, \mathbf{u_{n-1}}\} \subseteq V$, i cui vettori hanno componenti*

$$\mathbf{u_1}|_{\mathcal{B}} = \begin{bmatrix} u_{11} \\ \vdots \\ u_{n1} \end{bmatrix}, \dots, \mathbf{u_{n-1}}|_{\mathcal{B}} = \begin{bmatrix} u_{1\,n-1} \\ \vdots \\ u_{n\,n-1} \end{bmatrix}.$$

Il vettore ottenuto attraverso il seguente determinante formale

$$\mathbf{u} = \begin{vmatrix} u_{11} & \dots & u_{1\,n-1} & \mathbf{v_1} \\ \vdots & & \vdots & \vdots \\ u_{n1} & \dots & u_{n\,n-1} & \mathbf{v_n} \end{vmatrix}$$

è uguale al duale di Hodge $*(\mathbf{u_1}, \dots, \mathbf{u_{n-1}})$.

Dimostrazione. Se U è linearmente dipendente allora anche le prime $n-1$ colonne della matrice generalizzata sono dipendenti, quindi per le proprietà del determinante abbiamo $\mathbf{u} = \mathbf{0} = *(\mathbf{u_1}, \dots, \mathbf{u_{n-1}})$.

Supponiamo invece che U sia indipendente e di conseguenza $\mathbf{u} \neq \mathbf{0}$. In tal caso la decomposizione di $\mathbf{u}$ rispetto alla base $\mathcal{B}$ è

$$\mathbf{u} = \sum_{i=1}^{n} C_{in} \cdot \mathbf{v_i},$$

dove C_{in} sono i complementi algebrici della matrice generalizzata rispetto all'ultima colonna. Dato che $\mathcal{B}$ è ortonormale, segue che per ogni vettore $\mathbf{v} \in V$ avente decomposizione $\mathbf{v} = \sum_{i=1}^{n} t_i \cdot \mathbf{v_i}$ vale

$$\langle \mathbf{v}, \mathbf{u} \rangle = \sum_{i=1}^{n} t_i \, C_{in} = \begin{vmatrix} u_{11} & \dots & u_{1\,n-1} & t_1 \\ \vdots & & \vdots & \vdots \\ u_{n1} & \dots & u_{n\,n-1} & t_n \end{vmatrix}.$$

Grazie alla formula precedente ed alle proprietà del determinante, abbiamo quindi che $\langle \mathbf{u_i}, \mathbf{u} \rangle = 0$ per ogni $1 \leq i \leq n-1$. Di conseguenza $\mathbf{u} \in U^{\perp}$, che per costruzione ha dimensione 1. A sua volta questo implica che l'insieme $\mathcal{B}' = \{\mathbf{u_1}, \dots, \mathbf{u_{n-1}}, \mathbf{u}\}$ è una base di V. Per studiarne l'orientazione, consideriamo la matrice del cambiamento di base da $\mathcal{B}'$ a $\mathcal{B}$:

$$M_{\mathcal{B}'\mathcal{B}} = \begin{bmatrix} u_{11} & \dots & u_{1\,n-1} & C_{1n} \\ \vdots & & \vdots & \vdots \\ u_{n1} & \dots & u_{n\,n-1} & C_{nn} \end{bmatrix}.$$

Sviluppando il determinante di $M_{\mathcal{B}\mathcal{B}'}$ rispetto all'ultima colonna, abbiamo

$$\det(M_{\mathcal{B}'\mathcal{B}}) = \sum_{i=1}^{n} {C_{in}}^2 = \|\mathbf{u}\|^2 > 0,$$

quindi $\mathcal{B}'$ è orientata positivamente.

Resta da verificare che $\|\mathbf{u}\| = \sqrt{G(U)}$. Grazie all'ortonormalità di $\mathcal{B}$, al Teorema 3.89 ed alla formula precedente, vale

$$M_{\mathcal{B}'\mathcal{B}}^T * M_{\mathcal{B}'\mathcal{B}} = \left[\begin{array}{c|c} G|_U & 0_{n-1\,1} \\ \hline 0_{1\,n-1} & \|\mathbf{u}\|^2 \end{array}\right],$$

da cui otteniamo

$$\|\mathbf{u}\|^4 = \det(M_{\mathcal{B}'\mathcal{B}})^2 = \det(M_{\mathcal{B}'\mathcal{B}}^T)\det(M_{\mathcal{B}'\mathcal{B}}) = \det(M_{\mathcal{B}'\mathcal{B}}^T * M_{\mathcal{B}'\mathcal{B}}) = \|\mathbf{u}\|^2 \det(G|_U)$$

e quindi la tesi. □

Esempio 8.48. Nello spazio $V = \mathbb{R}^4$ con prodotto scalare ed orientamento canonici, calcoliamo il complemento ortogonale di $U = \mathcal{L}(\mathbf{u_1} = (1, 0, -1, 0), \mathbf{u_2} = (0, 1, 0, -1), \mathbf{u_3} = (0, 0, 1, 0)$ utilizzando il duale di Hodge:

$$\mathbf{u} = \begin{vmatrix} u_{11} & u_{12} & u_{13} & u_{14} & \mathbf{e_1} \\ u_{21} & u_{22} & u_{23} & u_{24} & \mathbf{e_2} \\ u_{31} & u_{32} & u_{33} & u_{34} & \mathbf{e_3} \\ u_{41} & u_{42} & u_{43} & u_{44} & \mathbf{e_4} \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & \mathbf{e_1} \\ 0 & 1 & 0 & \mathbf{e_2} \\ -1 & 0 & 1 & \mathbf{e_3} \\ 0 & -1 & 0 & \mathbf{e_4} \end{vmatrix} = \mathbf{e_2} + \mathbf{e_4} = (0, 1, 0, 1)$$

e quindi $U^\perp = \mathcal{L}((0, 1, 0, 1))$. L'insieme $\{\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}, \mathbf{u}\}$ è una base positivamente orientata di V.

Osservazione 8.49. Il prodotto scalare $\langle \mathbf{v}, *(\mathbf{u_1}, \dots, \mathbf{u_{n-1}}) \rangle$ che abbiamo utilizzato nella dimostrazione del teorema precedente è anche detto prodotto misto. A seguito di quanto abbiamo visto nelle dimostrazioni del Lemma 8.44 e della Proposizione 8.47, vale

$$|\langle \mathbf{v}, *(\mathbf{u_1}, \dots, \mathbf{u_{n-1}} \rangle| = \sqrt{G(\mathbf{u_1}, \dots, \mathbf{u_{n-1}}, \mathbf{v})}.$$

Di conseguenza, tale prodotto soddisfa l'importante proprietà di essere diverso da zero se e solo se l'insieme $\{\mathbf{u_1}, \dots, \mathbf{u_{n-1}}, \mathbf{v}\}$ è una base dello spazio vettoriale.

Un utilizzo importante del duale di Hodge si ha nelle applicazioni fisiche. Ad esempio, nella meccanica classica il momento angolare di un punto materiale si definisce utilizzando questa operazione. Per questo motivo, concludiamo la sezione focalizzandoci su spazi vettoriali euclidei orientati di dimensione $n = 3$. Presenteremo alcune utili proprietà del prodotto vettoriale, partendo da quelle elementari.

Proposizione 8.50. *Sia V uno spazio vettoriale euclideo orientato di dimensione 3. Allora il prodotto vettoriale soddisfa le seguenti proprietà:*

i) è bilineare, ovvero

$$(t_1 \cdot \mathbf{u_1} + t_2 \cdot \mathbf{u_2}) \wedge \mathbf{u_3} = t_1 \cdot \mathbf{u_1} \wedge \mathbf{u_3} + t_2 \cdot \mathbf{u_2} \wedge \mathbf{u_3}$$

$$\mathbf{u_1} \wedge (t_2 \cdot \mathbf{u_2} + t_3 \cdot \mathbf{u_3}) = t_2 \cdot \mathbf{u_1} \wedge \mathbf{u_2} + t_3 \cdot \mathbf{u_1} \wedge \mathbf{u_3}$$

per ogni $t_1, t_2, t_3 \in \mathbb{K}$ ed $\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3} \in V$;

ii) è antisimmetrico, ovvero $\mathbf{u_1} \wedge \mathbf{u_2} = -\mathbf{u_2} \wedge \mathbf{u_1}$ per ogni $\mathbf{u_1}, \mathbf{u_2} \in V$.

Dimostrazione. La bilinearità segue dalla Proposizione 8.47 e dalla linearità del determinante rispetto alle combinazioni lineari nelle colonne di una matrice. Analogamente, l'antisimmetria segue dal cambio di segno del determinante a seguito di uno scambio di colonne. □

Le seguenti proprietà richiedono invece una verifica più attenta.

Proposizione 8.51. *Sia V uno spazio vettoriale euclideo orientato di dimensione 3. Allora, per ogni $\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3} \in V$ valgono le seguenti identità:*

i) la ciclicità del prodotto misto $\langle \mathbf{u_1}, \mathbf{u_2} \wedge \mathbf{u_3} \rangle = \langle \mathbf{u_2}, \mathbf{u_3} \wedge \mathbf{u_1} \rangle = \langle \mathbf{u_3}, \mathbf{u_1} \wedge \mathbf{u_2} \rangle$;

ii) la formula del prodotto triplo $\mathbf{u_1} \wedge (\mathbf{u_2} \wedge \mathbf{u_3}) = \langle \mathbf{u_1}, \mathbf{u_3} \rangle \cdot \mathbf{u_2} - \langle \mathbf{u_1}, \mathbf{u_2} \rangle \cdot \mathbf{u_3}$;

iii) l'identità di Jacobi $\mathbf{u_1} \wedge (\mathbf{u_2} \wedge \mathbf{u_3}) + \mathbf{u_2} \wedge (\mathbf{u_3} \wedge \mathbf{u_1}) + \mathbf{u_3} \wedge (\mathbf{u_1} \wedge \mathbf{u_2}) = \mathbf{0}$.

Dimostrazione. Sia $\mathcal{B} = \{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$ una base ortonormale positivamente orientata di V, rispetto a cui i tre vettori hanno componenti

$$\mathbf{u_1}|_{\mathcal{B}} = \begin{bmatrix} u_{11} \\ u_{21} \\ u_{31} \end{bmatrix}, \quad \mathbf{u_2}|_{\mathcal{B}} = \begin{bmatrix} u_{12} \\ u_{22} \\ u_{32} \end{bmatrix}, \quad \mathbf{u_3}|_{\mathcal{B}} = \begin{bmatrix} u_{13} \\ u_{23} \\ u_{33} \end{bmatrix}.$$

La prima identità segue dalla seguente proprietà del determinante:

$$\begin{vmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{vmatrix} = \begin{vmatrix} u_{12} & u_{13} & u_{11} \\ u_{22} & u_{23} & u_{21} \\ u_{32} & u_{33} & u_{31} \end{vmatrix} = \begin{vmatrix} u_{13} & u_{11} & u_{12} \\ u_{23} & u_{21} & u_{22} \\ u_{33} & u_{31} & u_{32} \end{vmatrix}.$$

Verifichiamo la seconda identità relativamente alla prima componente rispetto alla base $\mathcal{B}$, essendo le altre due componenti analoghe. Grazie alla ciclicità del prodotto misto, a sinistra dell'uguale abbiamo:

$$\langle \mathbf{v_1}, \mathbf{u_1} \wedge (\mathbf{u_2} \wedge \mathbf{u_3}) \rangle = \langle \mathbf{u_2} \wedge \mathbf{u_3}, \mathbf{v_1} \wedge \mathbf{u_1} \rangle = \begin{vmatrix} u_{12} & u_{13} & 0 \\ u_{22} & u_{23} & -u_{31} \\ u_{32} & u_{33} & u_{21} \end{vmatrix} =$$

$$= u_{21}u_{12}u_{23} - u_{21}u_{22}u_{13} + u_{31}u_{12}u_{33} - u_{31}u_{32}u_{13}.$$

Sviluppiamo esplicitamente anche l'espressione a destra dell'uguale:

$$\begin{aligned} \langle \mathbf{u_1}, \mathbf{u_3} \rangle \langle \mathbf{v_1}, \mathbf{u_2} \rangle - \langle \mathbf{u_1}, \mathbf{u_2} \rangle \langle \mathbf{v_1}, \mathbf{u_3} \rangle &= (u_{11}u_{13} + u_{21}u_{23} + u_{31}u_{33})u_{12} - (u_{11}u_{12} + u_{21}u_{22} + u_{31}u_{32})u_{13} \\ &= u_{21}u_{23}u_{12} + u_{31}u_{33}u_{12} - u_{21}u_{22}u_{13} - u_{31}u_{32}u_{13}. \end{aligned}$$

Dal confronto delle due formule segue la tesi.

Infine, utilizzando la formula del prodotto triplo appena dimostrata, abbiamo:

$$\mathbf{u_1} \wedge (\mathbf{u_2} \wedge \mathbf{u_3}) + \mathbf{u_2} \wedge (\mathbf{u_3} \wedge \mathbf{u_1}) + \mathbf{u_3} \wedge (\mathbf{u_1} \wedge \mathbf{u_2}) =$$

$$= \langle \mathbf{u_1}, \mathbf{u_3} \rangle \cdot \mathbf{u_2} - \langle \mathbf{u_1}, \mathbf{u_2} \rangle \cdot \mathbf{u_3} + \langle \mathbf{u_2}, \mathbf{u_1} \rangle \cdot \mathbf{u_3} - \langle \mathbf{u_2}, \mathbf{u_3} \rangle \cdot \mathbf{u_1} + \langle \mathbf{u_3}, \mathbf{u_2} \rangle \cdot \mathbf{u_1} - \langle \mathbf{u_3}, \mathbf{u_1} \rangle \cdot \mathbf{u_2} = \mathbf{0}. \quad \square$$

Una conseguenza interessante dell'identità di Jacobi è che in generale vale

$$\mathbf{u_1} \wedge (\mathbf{u_2} \wedge \mathbf{u_3}) - (\mathbf{u_1} \wedge \mathbf{u_2}) \wedge \mathbf{u_3} = -\mathbf{u_2} \wedge (\mathbf{u_3} \wedge \mathbf{u_1}) \neq \mathbf{0},$$

ovvero il prodotto vettoriale non soddisfa la proprietà associativa, contrariamente a tutte le operazioni che abbiamo incontrato finora. È quindi fondamentale indicare l'ordine quando si compongono più prodotti vettoriali, o equivalentemente si devono sempre mantenere tutte le parentesi nelle espressioni coinvolgenti tale operazione.

CAPITOLO 9

Isometrie ed applicazioni aggiunte

Questo capitolo è dedicato allo studio delle applicazioni lineari che hanno un comportamento speciale rispetto alla struttura di spazio vettoriale euclideo. Stiamo ovviamente parlando degli omomorfismi di tale struttura, che vengono chiamate isometrie, ma anche delle cosiddette applicazioni lineari aggiunte ed in particolare autoaggiunte. Queste ultime hanno un comportamento particolare rispetto al problema della diagonalizzazione e ricoprono un ruolo molto importante nell'algebra lineare, dal punto di vista sia teorico sia applicativo.

9.1. Isometrie: definizione, esempi e proprietà elementari

Ricordiamo che un omomorfismo è una funzione che rispetta le strutture algebriche degli insiemi tra cui è definita. Nel caso degli spazi vettoriali euclidei, dobbiamo quindi parlare di applicazioni lineari che commutano con l'operazione prodotto scalare.

Definizione 9.1. *Siano* $(V, \langle\ ,\ \rangle_V)$ *e* $(W, \langle\ ,\ \rangle_W)$ *due spazi vettoriali euclidei. Allora la funzione* $f : V \to W$ *è un omomorfismo di spazi euclidei, o isometria, se è un'applicazione lineare tale che per ogni* $\mathbf{v_1}, \mathbf{v_2} \in V$ *vale*

$$\langle f(\mathbf{v_1}), f(\mathbf{v_2})\rangle_W = \langle \mathbf{v_1}, \mathbf{v_2}\rangle_V .$$

A causa della corrispondenza biunivoca tra prodotti scalari e norme definita dalla formula di polarizzazione, chiaramente le isometrie possono essere caratterizzate anche come le applicazioni lineari $f \in \mathrm{Hom}(V, W)$ che per ogni $\mathbf{v} \in V$ soddisfano $\|f(\mathbf{v})\|_W = \|\mathbf{v}\|_V$. In termini geometrici, le isometrie sono le trasformazioni lineare che non alterano le lunghezze dei vettori e gli angoli convessi tra loro definiti.

Esempio 9.2.

- Siano V uno spazio vettoriale euclideo ed $U \subseteq V$ un sottospazio. Naturalmente, U è a sua volta uno spazio euclideo rispetto alla restrizione del prodotto scalare definito su V. Allora l'inclusione $i_{U;V}$ è un'isometria, infatti per ogni $\mathbf{u_1}, \mathbf{u_2} \in U$ vale

  $$\langle i_{U;V}(\mathbf{u_1}), i_{U;V}(\mathbf{u_2})\rangle_V = \langle \mathbf{u_1}, \mathbf{u_2}\rangle_V = \langle \mathbf{u_1}, \mathbf{u_2}\rangle_U .$$

  In particolare, l'identità Id_V è un'isometria.

  Anche la riflessione ortogonale $\mathcal{R}_U$ è un isometria. Infatti, dato $\mathbf{v} \in V$ consideriamo la decomposizione $\mathbf{v} = \mathbf{u} + \tilde{\mathbf{u}}$, dove $(\mathbf{u}, \tilde{\mathbf{u}}) \in U \times U^\perp$. Allora vale $\mathcal{R}_U(\mathbf{v}) = \mathbf{u} - \tilde{\mathbf{u}}$ e quindi

  $$\|\mathcal{R}_U(\mathbf{v})\|^2 = \|\mathbf{u} - \tilde{\mathbf{u}}\|^2 = \|\mathbf{u}\|^2 + \|\tilde{\mathbf{u}}\|^2 = \|\mathbf{u} + \tilde{\mathbf{u}}\|^2 = \|\mathbf{v}\| ,$$

  da cui $\|\mathcal{R}_U(\mathbf{v})\| = \|\mathbf{v}\|$. Al contrario, la proiezione ortogonale $\mathcal{P}_U$ sul sottospazio U non è un'isometria. Infatti l'immagine di ogni vettore $\tilde{\mathbf{u}} \in U^\perp$ è il vettore $\mathbf{0} \in U$, quindi in generale $\|\mathcal{P}_U(\tilde{\mathbf{u}})\| = 0 \neq \|\tilde{\mathbf{u}}\|$.

- In $V = \mathrm{Mat}(2,1;\mathbb{R})$ consideriamo l'endomorfismo f_A associato alla matrice

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

  Come sappiamo, rispetto al prodotto scalare canonico, l'applicazione f_A effettua la rotazione di tutti i vettori del medesimo angolo. Pertanto, f_A è un'isometria, in quanto non altera le lunghezze dei vettori. Verifichiamo questo fatto anche per controllo diretto attraverso la definizione. Dati due vettori $B, C \in V$, consideriamo le loro decomposizioni

$$B = a \cdot E_{11} + b \cdot E_{21} \qquad \text{e} \qquad C = c \cdot E_{11} + c \cdot E_{21}.$$

  Allora $f_A(B) = -b \cdot E_{11} + a \cdot E_{21}$ ed $f_A(C) = -d \cdot E_{11} + d \cdot E_{21}$, da cui otteniamo

$$\langle f_A(B), f_A(C) \rangle_F = bd + ac = \langle B, C \rangle_F .$$

- In $V = \mathrm{Mat}(2,1;\mathbb{R})$ consideriamo l'endomorfismo f_A associato alla matrice

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

  Nell'Esempio 6.6 abbiamo visto che f_A è diagonalizzabile, con autovalori $\lambda_1 = 1, \lambda_2 = -1$ e con una base di autovettori associati

$$\mathcal{B} = \left\{ \mathbf{v_1} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \mathbf{v_2} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}.$$

  In generale, data una base $\mathcal{B} = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ di uno spazio vettoriale reale V, esiste sempre un prodotto scalare rispetto al quale tale base è ortonormale. Infatti, dati due generici vettori $\mathbf{u}, \mathbf{w}$ dello spazio V, aventi decomposizione

$$\mathbf{u} = \sum_{i=1}^{n} \alpha_i \cdot \mathbf{v_i} \qquad \text{e} \qquad \mathbf{w} = \sum_{i=1}^{n} \beta_i \cdot \mathbf{v_i},$$

  definiamo l'operazione

$$\langle \mathbf{u}, \mathbf{w} \rangle_{\mathcal{B}} = \sum_{i=1}^{n} \alpha_i \beta_i.$$

  Lasciamo come esercizio al lettore verificare che $\langle\ ,\ \rangle_{\mathcal{B}}$ è un prodotto scalare su V e che $\mathcal{B}$ è ortonormale rispetto a tale prodotto.

  Nel nostro esempio, proviamo a studiare il comportamento della funzione f_A rispetto al prodotto scalare associato alla base degli autovettori $\mathcal{B}$. Dati due vettori $\mathbf{u}, \mathbf{w} \in V$, consideriamo le loro decomposizioni

$$\mathbf{u} = \alpha_1 \cdot \mathbf{v_1} + \alpha_2 \cdot \mathbf{v_2} \qquad \text{e} \qquad \mathbf{w} = \beta_1 \cdot \mathbf{v_1} + \beta_2 \cdot \mathbf{v_2}.$$

  Ricordando che $\lambda_1 = 1$ e $\lambda_2 = -1$, si ha

$$\begin{aligned} \langle f_A(\mathbf{u}), f_A(\mathbf{w}) \rangle_{\mathcal{B}} &= \langle \alpha_1\lambda_1 \cdot \mathbf{v_1} + \alpha_2\lambda_2 \cdot \mathbf{v_2}, \beta_1\lambda_1 \cdot \mathbf{v_1} + \beta_2\lambda_2 \cdot \mathbf{v_2} \rangle = \\ &= \alpha_1\beta_1\lambda_1^2 + \alpha_2\beta_2\lambda_2^2 = \alpha_1\beta_1 + \alpha_2\beta_2 = \\ &= \langle \mathbf{u}, \mathbf{w} \rangle_{\mathcal{B}} . \end{aligned}$$

  Quindi f_A è un isometria dello spazio euclideo $(V, \langle\ ,\ \rangle_{\mathcal{B}})$. Osserviamo però che f_A non è una rotazione, perché ad esempio l'autovettore $\mathbf{v_1}$ rimane uguale a se stesso a seguito dell'applicazione di f_A. In questo caso stiamo invece parlando della riflessione ortogonale rispetto all'autospazio V_1.

Concludiamo la sezione presentando alcune proprietà elementari delle isometrie.

Proposizione 9.3. *Siano* $(V, \langle\ ,\ \rangle_V), (W, \langle\ ,\ \rangle_W)$ *spazi vettoriali euclidei ed* $f \in \mathrm{Hom}(V, W)$:

i) *se* f *è un'isometria allora è iniettiva;*

ii) *supponendo* $\dim(V) = \dim(W) < +\infty$, *se* f *è un'isometria allora è un isomorfismo;*

iii) *supponendo* $\dim(V) = \dim(W) < +\infty$, *allora* f *è un'isometria se e solo se l'immagine di una base ortonormale di* V *è una base ortonormale di* W.

Dimostrazione. Sia $\mathbf{v}$ un vettore non nullo di V. Se f è un'isometria, allora

$$\|f(\mathbf{v})\|_W = \|\mathbf{v}\|_V \neq 0$$

e quindi $f(\mathbf{v}) \neq \mathbf{0}$. Questo significa che $\ker(f) = \{\mathbf{0}\}$ e quindi il Corollario 5.7 garantisce che f è iniettiva. Se vale inoltre $\dim(V) = \dim(W) < +\infty$, per il Corollario 5.31 abbiamo che f è un isomorfismo.

Per quanto concerne l'ultimo punto, grazie a quanto appena dimostrato sappiamo che se f è un'isometria allora è un isomorfismo e quindi l'immagine di una base $\mathcal{B}_V \subseteq V$ è una base $\mathcal{B}_W \subseteq W$. Inoltre, le isometrie non alterano distanze ed angoli, quindi se $\mathcal{B}_V$ è ortonormale allora lo è anche $\mathcal{B}_W$. Viceversa, siano $\mathcal{B}_V = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ e $\mathcal{B}_W = \{f(\mathbf{v_1}), \dots, f(\mathbf{v_n})\}$ rispettivamente basi ortonormali di V e W. Per ogni $\mathbf{u}, \mathbf{v} \in V$ esistono le decomposizioni

$$\mathbf{u} = \sum_{i=1}^{n} \alpha_i \cdot \mathbf{v_i} \qquad \text{e} \qquad \mathbf{w} = \sum_{j=1}^{n} \beta_j \cdot \mathbf{v_j},$$

attraverso cui ricaviamo

$$\langle f(\mathbf{u}), f(\mathbf{v})\rangle_W = \left\langle \sum_{i=1}^{n} \alpha_i \cdot f(\mathbf{v_i}), \sum_{j=1}^{n} \beta_j \cdot f(\mathbf{v_j}) \right\rangle_W = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\beta_j \langle f(\mathbf{v_i}), f(\mathbf{v_j})\rangle_W = \sum_{i=1}^{n} \alpha_i\beta_i = \langle \mathbf{u}, \mathbf{v}\rangle_V .$$

Quindi, f è un'isometria. □

Corollario 9.4. *Siano* V, W *spazi vettoriali euclidei soddisfacenti* $\dim(W) < \dim(V)$. *Allora non esistono isometrie in* $\mathrm{Hom}(V, W)$.

Dimostrazione. Se gli spazi hanno dimensione finita, per il teorema di nullità più rango non possono esistere omomorfismi iniettivi da V in W e quindi non possono esistere isometrie. S V non è finitamente generato, basta prendere un suo sottospazio finitamente generato avente dimensione maggiore di $\dim(W)$ e ripetere il ragionamento precedente. □

9.2. Endomorfismi isometrici e matrici ortogonali

In questa sezione ci focalizzeremo sullo studio delle isometrie da uno spazio vettoriale V in se stesso. Se V è finitamente generato, la Proposizione 9.3 garantisce che esse sono degli automorfismi, pertanto formano un sottoinsieme di $GL(V)$ e le loro matrici rappresentative formano un sottoinsieme di $GL(n; \mathbb{R})$, dove $n = \dim(V)$. Approfondiamo meglio questi concetti.

Definizione 9.5. *In uno spazio vettoriale euclideo* V, *un endomorfismo isometrico è un'applicazione* $f \in \mathrm{End}(V)$ *che sia anche un'isometria. L'insieme di tutti gli endomorfismi isometrici su* V *si indica con* $O(V)$.

Proposizione 9.6. *Dato uno spazio vettoriale euclideo* V, *l'insieme* $O(V)$ *è un sottogruppo di* $GL(V)$.

Dimostrazione. Per prima cosa verifichiamo che $O(V)$ è chiuso rispetto alla composizione di funzioni. Date $f_1, f_2 \in O(V)$, per ogni vettore $\mathbf{v} \in V$ vale

$$\|(f_1 \circ f_2)(\mathbf{v})\| = \|f_1(f_2(\mathbf{v}))\| = \|f_2(\mathbf{v})\| = \|\mathbf{v}\|,$$

quindi $f_1 \circ f_2 \in O(V)$. Rimane da verificare che la funzione Id_V e l'inversa f^{-1} di un isometria f appartengono ad $O(V)$. La prima proprietà l'abbiamo già osservata nel capitolo precedente. Per la seconda, notiamo in primo luogo che f^{-1} esiste in quanto $f \in GL(V)$. Quindi, per ogni $\mathbf{v} \in V$ vale

$$\|\mathbf{v}\| = \|\mathrm{Id}_V(\mathbf{v})\| = \left\|(f \circ f^{-1})(\mathbf{v})\right\| = \left\|f(f^{-1}(\mathbf{v}))\right\| = \left\|f^{-1}(\mathbf{v})\right\|,$$

dove nell'ultima uguaglianza abbiamo sfruttato l'ipotesi che f sia un'isometria. Di conseguenza otteniamo che anche $f^{-1} \in O(V)$. □

Definizione 9.7. *Una matrice $Q \in \mathrm{Mat}(n, n; \mathbb{R})$ si dice ortogonale se soddisfa*

$$Q * Q^T = \mathbb{I}_n.$$

L'insieme delle matrici ortogonali si indica con $O(n; \mathbb{R})$.

Esempio 9.8.

- Ovviamente la matrice $\mathbb{I}_n$ è ortogonale, in quanto $\mathbb{I}_n^T = \mathbb{I}_n$ e quindi la verifica di ortogonalità diventa $\mathbb{I}_n^2 = \mathbb{I}_n$.
- Sia
$$Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix},$$
allora è facile verificare che $Q * Q^T = \mathbb{I}_2$.

Le matrici ortogonali soddisfano alcune proprietà elementari.

Proposizione 9.9. *Data $Q \in \mathrm{Mat}(n, n; \mathbb{R})$, allora*

i) $Q \in O(n; \mathbb{R})$ se e solo se anche $Q^T \in O(n; \mathbb{R})$, ed in tal caso $Q^{-1} = Q^T$;
ii) $Q \in O(n; \mathbb{R})$ se e solo se le sue righe e le sue colonne formano rispettivamente una base ortonormale di $\mathrm{Mat}(1, n; \mathbb{R})$ e di $\mathrm{Mat}(n, 1; \mathbb{R})$, rispetto al prodotto scalare di Frobenius;
iii) se $Q \in O(n; \mathbb{R})$ allora $\det(Q) = \pm 1$.

Dimostrazione. Verifichiamo ogni singolo punto della proposizione.

i) Per definizione, la matrice Q è ortogonale se Q^T è la sua inversa destra. Ma per le proprietà delle matrici invertibili, questo significa che Q è invertibile e Q^T è anche la sua inversa sinistra, da cui otteniamo $Q^T * Q = \mathbb{I}_n$. Quindi vale
$$\mathbb{I}_n = \mathbb{I}_n^T = (Q^T * Q)^T = Q^T * (Q^T)^T,$$
da cui segue che $Q^T \in O(n; \mathbb{R})$. Il viceversa è analogo.

ii) La seguente identità
$$(Q * Q^T)_{ij} = \sum_{k=1}^{n} q_{ik}\, q_{jk} = \left\langle Q_{R(i)}, Q_{R(j)} \right\rangle_F = \delta_{ij}$$
è soddisfatta se e solo se le righe di Q formano una base ortonormale di $\mathrm{Mat}(1, n; \mathbb{R})$. L'affermazione analoga sulle colonne si ottiene considerando la relazione $Q^T * Q = \mathbb{I}_n$, che abbiamo dimostrato nel punto precedente essere conseguenza dell'ortogonalità di Q.

iii) Partendo dalla definizione di matrice ortogonale e sfrutando il Teorema di Binét, abbiamo

$$1 = \det(\mathbb{I}_n) = \det(Q * Q^T) = \det(Q)\det(Q^T) = \det(Q)^2 \qquad \Rightarrow \qquad \det(Q) = \pm 1. \qquad \square$$

Proposizione 9.10. *L'insieme $O(n;\mathbb{R})$ è un sottogruppo di $GL(n;\mathbb{R})$, detto gruppo ortogonale.*

DIMOSTRAZIONE. Dobbiamo dimostrare che $O(n;\mathbb{R})$ è chiuso rispetto al prodotto di matrici. Date $Q, R \in O(n;\mathbb{R})$, abbiamo

$$(Q * R) * (Q * R)^T = Q * R * R^T * Q^T = Q * \mathbb{I}_n * Q^T = Q * Q^T = \mathbb{I}_n,$$

quindi $Q * R \in O(n;\mathbb{R})$. Ovviamente $\mathbb{I}_n$ appartiene ad $O(n;\mathbb{R})$ ed infine, grazie alla Proposizione 9.9, sappiamo che $Q^{-1} = Q^T \in O(n;\mathbb{R})$. $\square$

Possiamo a questo punto enunciare e dimostrare il teorema che lega endomorfismi isometrici e matrici ortogonali.

Teorema 9.11. *Siano V uno spazio vettoriale euclideo finitamente generato e $\mathcal{B}$ una sua base ortonormale. Allora la mappa delle componenti*

$$\begin{array}{rccc} \phi_{\mathcal{B}}: & O(V) & \longrightarrow & O(n;\mathbb{R}) \\ & f & \longmapsto & F|_{\mathcal{B}} \end{array}$$

è un isomorfismo di gruppi.

DIMOSTRAZIONE. Sappiamo che $\phi_{\mathcal{B}}$ è un isomorfismo di gruppi da $GL(V)$ a $GL(n;\mathbb{R})$. Di conseguenza, dobbiamo solo verificare che l'immagine di $O(V)$ tramite $\phi_{\mathcal{B}}$ corrisponda al gruppo $O(n;\mathbb{R})$. Sia $f \in GL(V)$ e $\mathcal{B} = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$. Sfruttando il fatto che $\mathcal{B}$ è ortonormale e quindi $G|_{\mathcal{B}} = \mathbb{I}_n$, abbiamo

$$\delta_{ij} = \langle f(\mathbf{v_i}), f(\mathbf{v_j})\rangle = f(\mathbf{v_i})|_{\mathcal{B}}^T * G|_{\mathcal{B}} * f(\mathbf{v_j})_{\mathcal{B}} = f(\mathbf{v_i})|_{\mathcal{B}}^T * f(\mathbf{v_j})_{\mathcal{B}} = \left\langle F|_{\mathcal{B},C(i)}, F|_{\mathcal{B},C(j)} \right\rangle_F,$$

per ogni $1 \le i, j \le n$. Questo significa che le colonne di $F|_{\mathcal{B}}$ formano una base ortonormale di $\mathrm{Mat}(n,1;\mathbb{R})$. Grazie alla Proposizione 9.9, segue che $F|_{\mathcal{B}} \in O(n;\mathbb{R})$. Viceversa, se $Q \in O(n;\mathbb{R})$ allora per il teorema di rappresentazione esiste $f \in \mathrm{End}(V)$ tale che $F|_{\mathcal{B}} = Q$. Sfruttando la sequenza di uguaglianze precedente, abbiamo

$$\langle f(\mathbf{v_i}), f(\mathbf{v_j})\rangle = \left\langle F|_{\mathcal{B},C(i)}, F|_{\mathcal{B},C(j)} \right\rangle_F = \left\langle Q_{C(i)}, Q_{C(j)} \right\rangle_F = \delta_{ij}$$

e quindi f è un'isometria. $\square$

È importante osservare che il teorema precedente richiede che $\mathcal{B}$ sia una base ortonormale. Se questo requisito non è soddisfatto, non è necessariamente vero che la matrice rappresentante un'isometria sia ortogonale, o viceversa che una matrice ortogonale rappresenti un'isometria.

Corollario 9.12. *Siano V uno spazio vettoriale euclideo finitamente generato ed $f \in O(V)$, allora* $\det(f) = \pm 1$.

DIMOSTRAZIONE. La tesi è conseguenza del Teorema 9.11 e della Proposizione 9.9. $\square$

9.3. Rotazioni

Il Corollario 9.12 permette di distinguere due classi differenti di isometrie, a seconda del valore del loro determinante. Per inquadrare meglio la questione, è necessario lavorare in uno spazio vettoriale euclideo orientato.

Definizione 9.13. *Sia (V, ξ) uno spazio vettoriale euclideo orientato. Una rotazione di (V, ξ) è un'endomorfismo isometrico f dello spazio che rispetta l'orientazione, ovvero tale che per ogni base $\mathcal{B}$ di V vale $\xi([\, f(\mathcal{B})\,]) = \xi([\,\mathcal{B}\,])$. L'insieme di tutte le rotazioni di (V, ξ) è indicato con $SO(V)$.*

In altri termini, una rotazione è un'applicazione lineare che manda ogni base ortonormale positivamente (negativamente) orientata in un'altra base ortonormale positivamente (negativamente) orientata. Le rotazioni si possono facilmente caratterizzare in base al segno del loro determinante.

Lemma 9.14. *L'insieme $SL(n;\mathbb{K}) = \{A \in GL(n;\mathbb{K}) \mid \det(A) = 1\}$ è un sottogruppo di $GL(n;\mathbb{K})$, detto gruppo speciale lineare. A sua volta, l'insieme $SO(n;\mathbb{R}) = O(n;\mathbb{R}) \cap SL(n;\mathbb{R})$ è un sottogruppo di $O(n;\mathbb{R})$, detto gruppo speciale ortogonale.*

Dimostrazione. Date $A, B \in SL(n;\mathbb{K})$, abbiamo

$$\det(A * B) = \det(A)\det(B) = 1,$$

quindi $A * B \in SL(n;\mathbb{K})$, cioé l'insieme è chiuso rispetto al prodotto. Inoltre, per ogni $A \in SL(n;\mathbb{K})$ esiste la matrice inversa e $\det(A^{-1}) = \det(A)^{-1} = 1$, per cui anche $A^{-1} \in SL(n;\mathbb{K})$. Infine, $\mathbb{I}_n \in SL(n;\mathbb{K})$. Quindi $SL(n;\mathbb{K})$ è un gruppo. L'ultimo punto del lemma si verifica in modo analogo, oppure segue in maniera più elegante dalla proprietà generale che l'intersezione di due gruppi è un gruppo, la cui verifica è lasciata al lettore. □

Teorema 9.15. *Sia (V, ξ) uno spazio vettoriale euclideo orientato. Allora:*

i) le rotazioni di (V, ξ) sono le isometrie $f \in O(V)$ che soddisfano $\det(f) = 1$;
ii) l'insieme $SO(V)$ è un sottogruppo di $O(V)$;
iii) data $\mathcal{B}$ una base ortonormale di V, la mappa delle componenti

$$\begin{array}{rccc} \phi_{\mathcal{B}}: & SO(V) & \longrightarrow & SO(n;\mathbb{R}) \\ & f & \longmapsto & F|_{\mathcal{B}} \end{array}$$

è un isomorfismo di gruppi.

Dimostrazione. Sia $\mathcal{B} = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ una base di V e $\mathcal{B}' = f(\mathcal{B}) = \{f(\mathbf{v_1}), \ldots, f(\mathbf{v_n})\}$. Se f è un isometria allora $\mathcal{B}'$ è una base di V e quindi possiamo rappresentare f rispetto alla coppia di basi $\mathcal{B}, \mathcal{B}'$:

$$F|_{\mathcal{B}\mathcal{B}'} = \begin{bmatrix} f(\mathbf{v_1})|_{\mathcal{B}'} & \ldots & f(\mathbf{v_n})|_{\mathcal{B}'} \end{bmatrix} = \mathbb{I}_n.$$

Per la regola del cambiamento di base abbiamo

$$\mathbb{I}_n = F|_{\mathcal{B}\mathcal{B}'} = M_{\mathcal{B}\mathcal{B}'} * F|_{\mathcal{B}}$$

e quindi

$$\det(f) = \det(F|_{\mathcal{B}}) = \det(M_{\mathcal{B}\mathcal{B}'})^{-1}.$$

Di conseguenza vale $\xi([\, f(\mathcal{B})\,]) = \xi([\,\mathcal{B}\,])$ se e solo se $\det(f) > 0$, da cui segue il primo punto del teorema. Sfruttando questa proprietà, la verifica del secondo punto è del tutto analoga a quella svolta nel Lemma 9.14 ed è lasciata al lettore. Infine, utilizzando ancora il primo punto del teorema, segue che la funzione $\phi_{\mathcal{B}}$ è una corrispondenza biunivoca tra $SO(V)$ ed $SO(n;\mathbb{R})$, da cui segue l'ultima affermazione contenuta nell'enunciato. $\square$

A seguito dell'ultimo teorema, possiamo osservare che l'insieme delle rotazioni non dipende dalla scelta dell'orientazione ξ dello spazio, pertanto si parla semplicemente di rotazioni dello spazio euclideo V.

Esempio 9.16. Nell'Esempio 8.38 abbiamo visto due isometrie di $\mathrm{Mat}(2,1;\mathbb{R})$, definite dalle matrici

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

L'applicazione f_A è una rotazione rispetto al prodotto scalare di Frobenius, poiché $A \in SO(2;R)$. Invece l'applicazione f_B è un'isometria rispetto al prodotto scalare che rende ortonormali due suoi autovettori, ma non è una rotazione, come abbiamo già spiegato nell'Esempio 9.2. Non deve sorprendere che $B \notin O(2;\mathbb{R})$, dato che la base canonica non è ortonormale rispetto a tale prodotto scalare. D'altra parte vale $\det(B) = \det(f_B) = -1$, poiché il determinante di un'applicazione non dipende dalla scelta della base rappresentativa e quindi dal fatto che questa sia o meno ortonormale.

Per definizione, le rotazioni sono gli automorfismi degli spazi vettoriali euclidei orientati. È naturale pertanto chiedersi come esse si comportino rispetto all'operazione duale di Hodge.

Proposizione 9.17. *Sia V uno spazio vettoriale euclideo orientato di dimensione $2 \leq n < \infty$ e sia $f \in SO(V)$. Allora f commuta con il duale di Hodge, ovvero per ogni $\mathbf{u_1}, \dots, \mathbf{u_{n-1}} \in V$ vale*

$$*(f(\mathbf{u_1}), \dots, f(\mathbf{u_{n-1}})) = f(*(\mathbf{u_1}, \dots, \mathbf{u_{n-1}})).$$

Dimostrazione. Dobbiamo verificare l'uguaglianza della norma e, nell'ipotesi che questa non sia nulla, della direzione e del verso dei due vettori.

- Norma: sfruttando le proprietà delle isometrie, abbiamo

$$\|f(*(\mathbf{u_1}, \dots, \mathbf{u_{n-1}}))\|^2 = \|*(\mathbf{u_1}, \dots, \mathbf{u_{n-1}})\|^2 = \begin{vmatrix} \langle \mathbf{u_1}, \mathbf{u_1} \rangle & \dots & \langle \mathbf{u_1}, \mathbf{u_{n-1}} \rangle \\ \vdots & & \vdots \\ \langle \mathbf{u_{n-1}}, \mathbf{u_{n-1}} \rangle & \dots & \langle \mathbf{u_{n-1}}, \mathbf{u_{n-1}} \rangle \end{vmatrix} =$$

$$= \begin{vmatrix} \langle f(\mathbf{u_1}), f(\mathbf{u_1}) \rangle & \dots & \langle f(\mathbf{u_1}), f(\mathbf{u_{n-1}}) \rangle \\ \vdots & & \vdots \\ \langle f(\mathbf{u_{n-1}}), f(\mathbf{u_{n-1}}) \rangle & \dots & \langle f(\mathbf{u_{n-1}}), f(\mathbf{u_{n-1}}) \rangle \end{vmatrix} = \|*(f(\mathbf{u_1}), \dots, f(\mathbf{u_{n-1}}))\|^2.$$

- Direzione: ancora per le proprietà delle isometrie, per ogni $1 \leq i \leq n-1$ vale

$$\langle f(*(\mathbf{u_1}, \dots, \mathbf{u_{n-1}})), f(\mathbf{u_i}) \rangle = \langle *(\mathbf{u_1}, \dots, \mathbf{u_{n-1}}), \mathbf{u_i} \rangle = 0,$$

  quindi $f(*(\mathbf{u_1}, \dots, \mathbf{u_{n-1}}))$ appartiene allo spazio perpendicolare ad $f(\mathbf{u_1}), \dots, f(\mathbf{u_{n-1}})$, ovvero a $\mathcal{L}(*(f(\mathbf{u_1}), \dots, f(\mathbf{u_{n-1}})))$.
- Verso: per definizione di duale di Hodge, la base

$$\{f(\mathbf{u_1}), \dots, f(\mathbf{u_{n-1}}), *(f(\mathbf{u_1}), \dots, f(\mathbf{u_{n-1}}))\}$$

  è orientata positivamente. D'altra parte, per definizione di rotazione, l'orientazione della base

$$\{f(\mathbf{u_1}), \dots, f(\mathbf{u_{n-1}}), f(*(\mathbf{u_1}, \dots, \mathbf{u_{n-1}}))\}$$

  è uguale all'orientazione di

$$\{\mathbf{u_1}, \dots, \mathbf{u_{n-1}}, *(\mathbf{u_1}, \dots, \mathbf{u_{n-1}})\},$$

che ancora una volta è positiva per definizione di duale di Hodge. Quindi i due vettori hanno la medesima orientazione. $\square$

9.4. Isometrie in dimensione bassa

Sfruttando gli strumenti sviluppati nella sezione precedente, indaghiamo in maniera esaustiva gli endomorfismi isometrici nel caso di spazi vettoriali euclidei di dimensione 2 e 3. Preliminarmente, abbiamo bisogno del seguente risultato.

Lemma 9.18. *Sia V uno spazio vettoriale euclideo avente* $\dim(V) = n < \infty$ *e sia $f \in O(V)$. Allora:*

i) gli unici possibili autovalori di f sono $\lambda = \pm 1$;

ii) se n è dispari e $\det(f) = 1$ *(rispettivamente* $\det(f) = -1$*), sicuramente* 1 *(rispettivamente* -1*) è autovalore di f.*

iii) se n è pari e $\det(f) = -1$*, sicuramente* 1 *e* -1 *sono autovalori di f.*

Dimostrazione. Per definizione, $\lambda \in \mathbb{R}$ è autovalore se esiste un vettore $\mathbf{v} \in V$ non nullo tale che $f(\mathbf{v}) = \lambda \cdot \mathbf{v}$. In tal caso vale

$$\|\mathbf{v}\| = \|f(\mathbf{v})\| = \|\lambda \cdot \mathbf{v}\| = |\lambda| \, \|\mathbf{v}\| \, .$$

Semplificando $\|\mathbf{v}\|$ otteniamo $|\lambda| = 1$, da cui segue il primo punto dell'enunciato.

Le dimostrazioni dei punti (ii) e (iii) sono simili. In primo luogo, ricordiamo che λ è autovalore di f se e solo se è una radice reale del polinomio caratteristico $P_f(\lambda)$. Per quanto studiato nella Sezione 6.3, abbiamo

$$P_f(\lambda) = \det(f) + \cdots + (-1)^n \lambda^n.$$

Supponiamo che n sia dispari e che valga $\det(f) = 1$. La funzione polinomiale $P_f(\lambda)$ è continua da $\mathbb{R}$ in $\mathbb{R}$, pertanto possiamo qui sfruttare alcuni risultati dell'Analisi. Per prima cosa valgono:

$$P_f(0) = 1 \qquad \text{mentre} \qquad \lim_{\lambda \to +\infty} P_f(\lambda) = \lim_{\lambda \to +\infty} -\lambda^n = -\infty.$$

Dalla definizione di limite all'infinito e dal teorema di esistenza degli zeri, sappiamo che esiste una radice di P_f nell'intervallo $[0, +\infty]$. Per quanto provato nel punto precedente, questa deve essere necessariamente $\lambda = +1$. Il caso in cui $\det(f) = -1$ è analogo, l'unica differenza consiste nel considerare il limite di $P_f(\lambda)$ per $\lambda \to -\infty$. Infine, se n è pari e $\det(f) = -1$, attraverso lo stesso ragionamento si prova l'esistenza di entrambe le radici reali ± 1. $\square$

In quanto segue, indicheremo con V_+ e V_- gli autospazi associati rispettivamente agli autovalori $+1$ e -1, qualora questi esistano. Iniziamo con la classificazione degli endomorfismi isometrici in dimensione 2.

Teorema 9.19. *Sia V uno spazio euclideo di dimensione* 2 *e sia $f \in O(V)$. Allora:*

i) se $\det(f) = 1$ *l'isometria f ruota ogni vettore $\mathbf{v}$ di un angolo convesso θ soddisfacente* $\mathrm{Tr}(f) = 2\cos\theta$;

ii) se $\det(f) = -1$ *l'isometria f è la riflessione ortogonale rispetto all'autospazio V_+.*

Dimostrazione. Sia $\mathcal{B} = \{\mathbf{v_1}, \mathbf{v_2}\}$ una base ortonormale di V. L'isometria f è rappresentata da una matrice ortogonale

$$F|_{\mathcal{B}} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in O(2; \mathbb{R}).$$

Questo significa che le righe di $F|_{\mathcal{B}}$ costituiscono un insieme ortonormale, ovvero

$$\begin{cases} a^2 + b^2 = 1 \\ c^2 + d^2 = 1 \\ ac + bd = 0 \end{cases}.$$

Il sistema può essere risolto utilizzando le funzioni trigonometriche:

$$\begin{cases} a = \cos\alpha \qquad b = \sin\alpha \\ c = \cos\beta \qquad d = \sin\beta \\ \cos\alpha\cos\beta + \sin\alpha\sin\beta = 0 \end{cases}$$

dove $\alpha, \beta \in (-\pi, \pi]$. L'ultima equazione è riscrivibile come

$$\cos(\alpha - \beta) = 0,$$

che ha soluzioni $\beta = \alpha \pm \frac{\pi}{2}$. A seconda del segno abbiamo due situazioni differenti.

Se $\beta = \alpha + \frac{\pi}{2}$, allora

$$F|_{\mathcal{B}} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix},$$

quindi $\det(f) = \cos^2\alpha + \sin^2\alpha = 1$ e $\mathrm{Tr}(f) = 2\cos\alpha$. Dato $\mathbf{v} \in V$ avente decomposizione $\mathbf{v} = t_1 \cdot \mathbf{v_1} + t_2 \cdot \mathbf{v_2}$, il prodotto scalare tra $\mathbf{v} \in V$ e la sua immagine $f(\mathbf{v})$ è

$$\langle \mathbf{v}, f(\mathbf{v}) \rangle = \begin{bmatrix} t_1 & t_2 \end{bmatrix} * \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} * \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = (t_1^2 + t_2^2)\cos\alpha$$

e quindi l'angolo convesso da loro definito è

$$\theta = \widehat{\mathbf{v}f(\mathbf{v})} = \arccos\left(\frac{(t_1^2 + t_2^2)\cos\alpha}{(t_1^2 + t_2^2)}\right) = \arccos(\cos\alpha) = \begin{cases} \alpha & \text{se} \quad \alpha \in [0, \pi], \\ 2\pi - \alpha & \text{se} \quad \alpha \in [\pi, 2\pi). \end{cases}$$

Tale angolo è costante per ogni vettore $\mathbf{v}$ e soddisfa $\mathrm{Tr}(f) = 2\cos\theta$.

Se $\beta = \alpha - \frac{\pi}{2}$, allora

$$F|_{\mathcal{B}} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix},$$

quindi $\det(f) = -\cos^2\alpha - \sin^2\alpha = -1$. Per il Lemma 9.18, esistono entrambi gli autovalori ± 1 e di conseguenza f è diagonalizzabile. Sia $\mathcal{B} = \{\mathbf{v}_+, \mathbf{v}_-\}$ una base di V costituita da autovettori di f. Su tali vettori abbiamo $f(\mathbf{v}_\pm) = \pm\mathbf{v}_\pm$, ovvero f è la riflessione rispetto a V_+ e parallela a V_-. Inoltre

$$\langle \mathbf{v}_+, \mathbf{v}_- \rangle = \langle f(\mathbf{v}_+), f(\mathbf{v}_-) \rangle = \langle \mathbf{v}_+, -\mathbf{v}_- \rangle = -\langle \mathbf{v}_+, \mathbf{v}_- \rangle,$$

per cui $\langle \mathbf{v}_+, \mathbf{v}_- \rangle = 0$. Quindi $V_+^\perp = V_-$ ed f coincide con la riflessione ortogonale $\mathcal{R}_{V_+}$. □

Il seguente teorema classifica invece le isometrie in dimensione 3, ed è dovuto ad Eulero.

Teorema 9.20. *Sia V uno spazio euclideo di dimensione 3 e sia f un endomorfismo isometrico. Allora esistono quattro differenti situazioni.*

i) $f = \mathrm{Id}_V$ e quindi V coincide con V_+.

ii) $f = -\mathrm{Id}_V$ e quindi V coincide con V_-. L'isometria riflette ogni vettore rispetto a $\mathbf{0}$.

iii) *Se* $\det(f) = 1$ *ed* $f \neq \mathrm{Id}_V$, *allora l'autospazio* V_+ *ha dimensione* 1. *Ogni vettore* $\mathbf{v} \in V$ *può essere decomposto come*

$$\mathbf{v} = \mathbf{v}_+ + \mathbf{v}_+^\perp, \qquad \textit{dove} \qquad (\mathbf{v}_+, \mathbf{v}_+^\perp) \in V_+ \times V_+^\perp.$$

L'isometria f *lascia inalterata la componente* $\mathbf{v}_+$ *e ruota la componente* $\mathbf{v}_+^\perp$ *all'interno di* $V_+^\perp$, *di un angolo convesso* θ *soddisfacente* $\mathrm{Tr}(f) = 1 + 2\cos\theta$.

iv) *Se* $\det(f) = -1$ *ed* $f \neq -\mathrm{Id}_V$, *allora l'autospazio* V_- *ha dimensione* 1. *Ogni vettore* $\mathbf{v} \in V$ *può essere decomposto come*

$$\mathbf{v} = \mathbf{v}_- + \mathbf{v}_-^\perp, \qquad \textit{dove} \qquad (\mathbf{v}_-, \mathbf{v}_-^\perp) \in V_- \times V_-^\perp.$$

L'isometria f *riflette ortogonalmente la componente* $\mathbf{v}_-$ *rispetto a* $V_-^\perp$ *e ruota la componente* $\mathbf{v}_-^\perp$ *all'interno di* $V_-^\perp$, *di un angolo* θ *soddisfacente* $\mathrm{Tr}(f) = -1 + 2\cos\theta$.

Dimostrazione. Per prima cosa osserviamo che i quattro casi esauriscono tutte le possibilità. Inoltre, le proprietà elencate in (i) e (ii) sono evidenti. Dimostriamo invece il caso (iii), essendo poi il caso (iv) completamente analogo. L'ipotesi $\det(f) = 1$ ed il Lemma 9.18 garantiscono l'esistenza dell'autovalore 1 e quindi $\dim(V_+) \geq 1$. Dato un versore $\mathbf{v}_+ \in V_+$, è possibile costruire una base ortonormale $\mathcal{B} = \{\mathbf{v}_+, \mathbf{u_1}, \mathbf{u_2}\}$ di V. La matrice rappresentativa di f rispetto a $\mathcal{B}$ è

$$F|_\mathcal{B} = \begin{bmatrix} f(\mathbf{v}_+)|_\mathcal{B} & f(\mathbf{u_1})|_\mathcal{B} & f(\mathbf{u_2})|_\mathcal{B} \end{bmatrix} = \begin{bmatrix} 1 & a & b \\ 0 & c & d \\ 0 & e & g \end{bmatrix} \in SO(3;\mathbb{R}).$$

A questo punto dimostriamo che $V_+ = \mathcal{L}(\mathbf{v}_+)$ e che il sottospazio $U = \mathcal{L}(\mathbf{u_1}, \mathbf{u_2})$ è uguale a $V_+^\perp$. Essendo $F|_\mathcal{B}$ ortogonale, deve valere $1 + a^2 + b^2 = 1$, che ha come unica soluzione reale $a = b = 0$. Questo significa che $f(\mathbf{u_1}), f(\mathbf{u_2}) \in U$ e quindi la restrizione $f|_U$ è un endomorfismo isometrico di U. La sua matrice rappresentativa rispetto a $\mathcal{B}_U = \{\mathbf{u_1}, \mathbf{u_2}\}$ è

$$A = \begin{bmatrix} c & d \\ e & g \end{bmatrix},$$

che ovviamente deve essere ortogonale. Inoltre vale

$$1 = \det(f) = \det(F|_\mathcal{B}) = 1 \det(A) = \det(A),$$

da cui ricaviamo che $A \in SO(2;\mathbb{R})$. Questo significa che $f|_U$ è una rotazione e per ipotesi $f|_U \neq \mathrm{Id}_U$. Pertanto non possono esistere vettori $\mathbf{u} \in U$ tali che $f(\mathbf{u}) = \mathbf{u}$, ovvero $V_+ \cap U = \{\mathbf{0}\}$. D'altra parte, per costruzione vale $V = V_+ + U$ e quindi $V = V_+ \oplus U$. Di conseguenza

$$\dim(V_+) = \dim(V) - \dim(U) = 3 - 2 = 1 \qquad \Rightarrow \qquad V_+ = \mathcal{L}(\mathbf{v}_+).$$

Analogamente vale

$$\dim(V_+^\perp) = \dim(V) - \dim(V_+) = 3 - 1 = 2 \qquad \Rightarrow \qquad U = V_+^\perp.$$

Chiaramente $f|_{V_+} = \mathrm{Id}_{V_+}$, mentre il comportamento di $f|_{V_+^\perp}$ è determinato dal Teorema 9.19. Questo significa che f ruota tutti i vettori di $V_+^\perp$ all'interno di tale sottospazio, di un angolo θ soddisfacente $2\cos\theta = \mathrm{Tr}(A) = \mathrm{Tr}(F|_\mathcal{B}) - 1$, da cui otteniamo $\mathrm{Tr}(f) = 1 + 2\cos\theta$. $\square$

Esempio 9.21. Nello spazio euclideo $V = \mathrm{Mat}(3,1;\mathbb{R})$ con prodotto scalare di Frobenius, studiamo l'endomorfismo f_A definito dalla matrice

$$A = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix}.$$

Per prima cosa, osserviamo che $A * A^T = \mathbb{I}_3$ e che $\det(A) = 1$, quindi $A \in SO(3;\mathbb{R})$. Dato che la base canonica di V è ortonormale, questo implica che $f_A \in SO(V)$, ovvero l'endomorfismo è una rotazione. Chiamiamo asse di rotazione l'autospazio V_+ e piano di rotazione il complemento ortogonale $V_+^\perp$. Abbiamo:

$$V_+ = \ker(A - \mathbb{I}_3) = \ker\left(\begin{bmatrix} 1/\sqrt{3} - 1 & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1 & -2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} - 1 \end{bmatrix}\right) =$$

$$= \mathcal{L}\left(\begin{bmatrix} 2 + \sqrt{14 - 4\sqrt{6}} \\ 2 - \sqrt{6} \\ 2 + \sqrt{2} \end{bmatrix}\right) \approx \mathcal{L}\left(\begin{bmatrix} 4.05 \\ -0.45 \\ 3.41 \end{bmatrix}\right),$$

$$V_+^\perp = \mathcal{L}\left(\begin{bmatrix} 1/\sqrt{3} \\ -1 \\ -2/\sqrt{6} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{2} \\ 1/\sqrt{6} - 1 \end{bmatrix}\right) \approx \mathcal{L}\left(\begin{bmatrix} 0.58 \\ -1 \\ -0.82 \end{bmatrix}, \begin{bmatrix} 0.58 \\ 0.71 \\ -0.59 \end{bmatrix}\right).$$

Infine, l'angolo di rotazione si ricava da

$$\cos\theta = \frac{\mathrm{Tr}(A) - 1}{2} = \frac{1/\sqrt{3} + 1/\sqrt{6} - 1}{2} \qquad \Rightarrow \qquad \theta \approx 90^\circ\, 25'.$$

9.5. Omomorfismo aggiunto ed endomorfismi autoaggiunti

La seconda parte di questo capitolo è dedicata a studiare l'impatto della struttura euclidea nei problemi di diagonalizzazione. In particolare, arriveremo ad enunciare e dimostrare il fondamentale Teorema spettrale ed a definire la cosiddetta Decomposizione ai Valori Singolari di un'applicazione lineare (Singular Value Decomposition o SVD). In questa sezione, iniziamo a definire e studiare un oggetto centrale della nostra analisi.

Definizione 9.22. *Siano $(V, \langle\ ,\ \rangle_V), (W, \langle\ ,\ \rangle_W)$ spazi vettoriali euclidei ed $f \in \mathrm{Hom}(V,W)$. Allora un omomorfismo $g \in \mathrm{Hom}(W,V)$ tale che*

$$\langle f(\mathbf{v}), \mathbf{w}\rangle_W = \langle \mathbf{v}, g(\mathbf{w})\rangle_V\,,$$

per ogni $(\mathbf{v},\mathbf{w}) \in V \times W$ si dice omomorfismo aggiunto di f.

Seguendo le notazioni della definizione precedente, abbiamo il seguente risultato riguardante l'esistenza e l'unicità dell'applicazione aggiunta.

Proposizione 9.23. *Per ogni $f \in \mathrm{Hom}(V,W)$, se g esiste allora è unico ed è denotato con f^*. Inoltre, se V, W sono finitamente generati allora f^* esiste.*

Dimostrazione. Siano $g, h \in \mathrm{Hom}(W,V)$ omomorfismi aggiunti di f. Allora

$$\langle f(\mathbf{v}), \mathbf{w}\rangle_W = \langle \mathbf{v}, g(\mathbf{w})\rangle_V = \langle \mathbf{v}, h(\mathbf{w})\rangle_V$$

e quindi

$$\langle \mathbf{v}, g(\mathbf{w}) - h(\mathbf{w})\rangle_V = 0,$$

per ogni $(\mathbf{v}, \mathbf{w}) \in V \times W$. Per la proprietà di non degenerazione del prodotto scalare, deve valere $g(\mathbf{w}) - h(\mathbf{w}) = \mathbf{0}$ e quindi $g = h$.

Supponiamo ora che i due spazi V e W siano finitamente generati e che $\mathcal{B}_V = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\} \subset V$ e $\mathcal{B}_W = \{\mathbf{w_1}, \ldots, \mathbf{w_m}\} \subset V$ siano due loro basi ortonormali. Definiamo la funzione $g : V \to W$ come

$$g(\mathbf{w}) = \sum_{i=1}^{n} \langle f(\mathbf{v_i}), \mathbf{w} \rangle_W \cdot \mathbf{v_i}.$$

Grazie alla linearità del prodotto scalare, è facile verificare che g è un omomorfismo. Inoltre, per ogni $\mathbf{v} \in V$ esiste la decomposizione $\mathbf{v} = \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{v_i} \rangle_V \cdot \mathbf{v_i}$, da cui ricaviamo

$$\begin{aligned}
\langle f(\mathbf{v}), \mathbf{w} \rangle_W &= \left\langle f\left(\sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{v_i} \rangle_V \cdot \mathbf{v_i}\right), \mathbf{w} \right\rangle_W = \left\langle \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{v_i} \rangle_V \cdot f(\mathbf{v_i}), \mathbf{w} \right\rangle_W = \\
&= \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{v_i} \rangle_V \langle f(\mathbf{v_i}), \mathbf{w} \rangle_W = \sum_{i=1}^{n} \langle \mathbf{v}, \langle f(\mathbf{v_i}), \mathbf{w} \rangle_W \cdot \mathbf{v_i} \rangle_V = \\
&= \left\langle \mathbf{v}, \sum_{i=1}^{n} \langle f(\mathbf{v_i}), \mathbf{w} \rangle_W \cdot \mathbf{v_i} \right\rangle_V = \langle \mathbf{v}, g(\mathbf{w}) \rangle_V .
\end{aligned}$$

Quindi $g = f^*$. □

Presentiamo alcune proprietà elementari dell'applicazione aggiunta.

Proposizione 9.24. *Siano* $(U, \langle\ ,\ \rangle_U), (V, \langle\ ,\ \rangle_V)$ *e* $(W, \langle\ ,\ \rangle_W)$ *spazi vettoriali euclidei ed* $f, g \in \mathrm{Hom}(U, V), h \in \mathrm{Hom}(V, W)$ *applicazioni lineari dotate di aggiunta. Allora l'operazione di calcolo dell'aggiunta:*

i) è un'involuzione, cioè $(f^*)^* = f$;
ii) soddisfa la proprietà di linearità $(\alpha \cdot f + \beta \cdot g)^* = \alpha \cdot f^* + \beta \cdot g^*$, *per ogni* $\alpha, \beta \in \mathbb{R}$;
iii) soddisfa la proprietà $(h \circ f)^* = f^* \circ h^*$;
iv) se esiste f^{-1} *allora* $(f^{-1})^* = (f^*)^{-1}$.

Dimostrazione. Verifichiamo i singoli punti dell'enunciato.

i) Per definizione e per la proprietà di simmetria del prodotto scalare, per ogni $(\mathbf{v}, \mathbf{u}) \in V \times U$ vale

$$\langle f^*(\mathbf{v}), \mathbf{u} \rangle_U = \langle \mathbf{v}, f(\mathbf{u}) \rangle_V .$$

Data l'unicità dell'aggiunta, questo implica che $f = (f^*)^*$.

ii) Per ogni $(\mathbf{u}, \mathbf{v}) \in U \times V$ abbiamo

$$\begin{aligned}
\langle (\alpha \cdot f + \beta \cdot g)(\mathbf{u}), \mathbf{v} \rangle_V &= \alpha \langle f(\mathbf{u}), \mathbf{v} \rangle_V + \beta \langle g(\mathbf{u}), \mathbf{v} \rangle_V = \alpha \langle \mathbf{u}, f^*(\mathbf{v}) \rangle_U + \beta \langle \mathbf{u}, g^*(\mathbf{v}) \rangle_U = \\
&= \langle \mathbf{u}, (\alpha \cdot f^* + \beta \cdot g^*)(\mathbf{v}) \rangle_U .
\end{aligned}$$

iii) Per ogni $(\mathbf{u}, \mathbf{w}) \in U \times W$ abbiamo

$$\langle (h \circ f)(\mathbf{v}), \mathbf{w} \rangle_W = \langle h(f(\mathbf{v})), \mathbf{w} \rangle_W = \langle f(\mathbf{v}), h^*(\mathbf{w}) \rangle_W = \langle \mathbf{v}, f^*(h^*(\mathbf{w})) \rangle_W = \langle \mathbf{v}, (f^* \circ h^*)(\mathbf{w}) \rangle_W .$$

iv) Dalla proprietà precedente abbiamo che $\mathrm{Id}_U{}^* = (f^{-1} \circ f)^* = f^* \circ (f^{-1})^*$. Poiché $\mathrm{Id}_U{}^* = \mathrm{Id}_U$, segue la tesi. □

In spazi di dimensione finita, esiste una caratterizzazione importante dell'aggiunta di un'applicazione lineare in termini della matrice rappresentativa.

Proposizione 9.25. *Siano* $(V, \langle\ ,\ \rangle_V)$ *e* $(W, \langle\ ,\ \rangle_W)$ *due spazi vettoriali euclidei finitamente generati e siano* $\mathcal{B}_V = \{\mathbf{v_1}, \dots, \mathbf{v_n}\} \subset V$ *e* $\mathcal{B}_W = \{\mathbf{w_1}, \dots, \mathbf{w_m}\} \subset W$ *due loro basi ortonormali. Data* $f \in \operatorname{Hom}(V, W)$*, l'applicazione* $g \in \operatorname{Hom}(W, V)$ *è l'aggiunta di* f *se e solo se*

$$\phi_{\mathcal{B}_W\mathcal{B}_V}(g) = \phi_{\mathcal{B}_V\mathcal{B}_W}(f)^T.$$

Dimostrazione. In primo luogo ricordiamo che f^* esiste ed è unica. La sua matrice rappresentativa è

$$F^*|_{\mathcal{B}_W\mathcal{B}_V} = \begin{bmatrix} f^*(\mathbf{w_1})|_{\mathcal{B}_V} & \dots & f^*(\mathbf{w_m})|_{\mathcal{B}_V} \end{bmatrix} = \begin{bmatrix} \langle \mathbf{v_1}, f^*(\mathbf{w_1})\rangle & \dots & \langle \mathbf{v_1}, f^*(\mathbf{w_m})\rangle \\ \vdots & & \vdots \\ \langle \mathbf{v_n}, f^*(\mathbf{w_1})\rangle & \dots & \langle \mathbf{v_n}, f^*(\mathbf{w_m})\rangle \end{bmatrix} =$$

$$= \begin{bmatrix} \langle f(\mathbf{v_1}), \mathbf{w_1}\rangle & \dots & \langle f(\mathbf{v_1}), \mathbf{w_m}\rangle \\ \vdots & & \vdots \\ \langle f(\mathbf{v_n}), \mathbf{w_1}\rangle & \dots & \langle f(\mathbf{v_n}), \mathbf{w_m}\rangle \end{bmatrix} = \begin{bmatrix} f(\mathbf{v_1})|_{\mathcal{B}_V}^T \\ \vdots \\ f(\mathbf{v_n})|_{\mathcal{B}_V}^T \end{bmatrix} = F|_{\mathcal{B}_V\mathcal{B}_W}^T.$$

Viceversa, se esiste un'applicazione g tale che $\phi_{\mathcal{B}_W\mathcal{B}_V}(g) = \phi_{\mathcal{B}_V\mathcal{B}_W}(f)^T$, allora per quanto appena dimostrato vale $\phi_{\mathcal{B}_W\mathcal{B}_V}(g) = \phi_{\mathcal{B}_W\mathcal{B}_V}(f^*)$. Dato che $\phi_{\mathcal{B}_W\mathcal{B}_V}$ è un isomorfismo, allora $g = f^*$. □

La Proposizione 9.25 permette di calcolare facilmente l'applicazione aggiunta di un'applicazione data f, una volta che sia nota una sua rappresentazione A rispetto ad una coppia di basi ortonormali. Infatti, f^* è semplicemente l'applicazione rappresentata dalla matrice trasposta A^T.

Corollario 9.26. *Siano* $(V, \langle\ ,\ \rangle_V)$ *e* $(W, \langle\ ,\ \rangle_W)$ *due spazi vettoriali euclidei finitamente generati e sia* $f \in \operatorname{Hom}(V, W)$*. Allora:*

$$\operatorname{Im}(f) = \ker(f^*)^\perp \qquad \textit{ed} \qquad \operatorname{Im}(f^*) = \ker(f)^\perp.$$

Dimostrazione. Siano $\mathbf{w} \in \operatorname{Im}(f)$ e $\tilde{\mathbf{w}} \in \ker(f^*)$. Allora esiste $\mathbf{v} \in V$ tale che $\mathbf{w} = f(\mathbf{v})$, quindi

$$\langle \mathbf{w}, \tilde{\mathbf{w}}\rangle = \langle f(\mathbf{v}), \tilde{\mathbf{w}}\rangle = \langle \mathbf{v}, f^*(\tilde{\mathbf{w}})\rangle = \langle \mathbf{v}, \mathbf{0_V}\rangle = 0.$$

Questo significa che $\mathbf{w} \perp \tilde{\mathbf{w}}$ e di conseguenza $\operatorname{Im}(f) \subseteq \ker(f)^\perp$. Per dimostrare l'uguaglianza, confrontiamo la dimensione dei due sottospazi. Date due basi $\mathcal{B}_V, \mathcal{B}_W$ degli spazi vettoriali, abbiamo

$$\dim(\operatorname{Im}(f)) = r(F|_{\mathcal{B}_V\mathcal{B}_W}) = r(F|_{\mathcal{B}_V\mathcal{B}_W}^T) = r(F|_{\mathcal{B}_W\mathcal{B}_V}^*) = \dim(\operatorname{Im}(f^*)) = \dim(W) - \dim(\ker(f^*)),$$

dove nell'ultima uguaglianza è stato utilizzato il Teorema di nullità più rango applicato ad f^*. Dato che $W = \ker(f^*) \oplus \ker(f^*)^\perp$, segue che

$$\dim(\operatorname{Im}(f)) = \dim(W) - \dim(\ker(f^*)) = \dim(\ker(f^*)^\perp)$$

e quindi $\operatorname{Im}(f) = \ker(f^*)^\perp$. L'uguaglianza $\operatorname{Im}(f^*) = \ker(f)^\perp$ si dimostra in maniera analoga scambiando i ruoli di f e f^*. □

Se consideriamo gli endomorfismi di uno spazio vettoriale, un ruolo particolare è svolto dalle applicazioni che soddisfano $f^* = f$, che nel caso di dimensione finita sono rappresentate da matrici simmetriche.

Definizione 9.27. *Siano* V *uno spazio vettoriale euclideo ed* $f \in \operatorname{End}(V)$*. Allora l'applicazione* f *si dice autoaggiunta se* $f^* = f$*, o equivalentemente se*

$$\langle f(\mathbf{v_1}), \mathbf{v_2}\rangle = \langle \mathbf{v_1}, f(\mathbf{v_2})\rangle$$

per ogni $\mathbf{v_1}, \mathbf{v_2} \in V$*. L'insieme di tutti gli endomorfismi autoaggiunti è indicato con* $\mathbb{S}(V)$*.*

La motivazione della scelta della notazione sarà chiara nel prossimo teorema. Preliminarmente, ricordiamo che $\mathbb{S}(n;\mathbb{R})$ è l'insieme delle matrici quadrate simmetriche di ordine n ad elementi reali e che esso costituisce un sottospazio di $\mathrm{Mat}(n,n;\mathbb{R})$.

Proposizione 9.28. *Sia V uno spazio vettoriale euclideo di dimensione $n<\infty$. Allora:*

i) l'insieme $\mathbb{S}(V)$ è un sottospazio vettoriale di $\mathrm{End}(V)$;

ii) data una base ortonormale $\mathcal{B}$ di V, la mappa delle componenti

$$\begin{array}{rccc} \phi_{\mathcal{B}}: & \mathbb{S}(V) & \longrightarrow & \mathbb{S}(n;\mathbb{R}) \\ & f & \longmapsto & F|_{\mathcal{B}} \end{array}$$

è un isomorfismo di spazi vettoriali.

Dimostrazione. L'applicazione 0_V è autoaggiunta. Infatti $\langle 0_V(\mathbf{v_1}),\mathbf{v_2}\rangle = \langle \mathbf{v_1}, 0_V(\mathbf{v_2})\rangle = 0$ per ogni $\mathbf{v_1},\mathbf{v_2}\in V$. Inoltre, per ogni $f_1,f_2\in\mathbb{S}(V)$ e $t_1,t_2\in\mathbb{R}$ vale

$$(t_1\cdot f_1 + t_2\cdot f_2)^* = t_1\cdot f_1^* + t_2\cdot f_2^* = t_1\cdot f_1 + t_2\cdot f_2,$$

quindi $\mathbb{S}(V)$ è chiuso rispetto alla somma ed al prodotto per uno scalare e di conseguenza $\mathbb{S}(V)$ è un sottospazio di $\mathrm{End}(V)$. Per il secondo punto è sufficiente dimostrare che $\phi_{\mathcal{B}}(\mathbb{S}(V)) = \mathbb{S}(n;\mathbb{R})$. Ma questo segue facilmente dalla Proposizione 9.25. Infatti, f è autoaggiunta se e solo se $F|_{\mathcal{B}} = F^*|_{\mathcal{B}} = F|_{\mathcal{B}}^T$ e quindi se e solo se $F|_{\mathcal{B}}\in\mathbb{S}(n;\mathbb{R})$. □

Il teorema precedente caratterizza quindi le applicazioni autoaggiunte in spazi finitamente generati come quelle che sono rappresentabili da matrici simmetriche rispetto ad una base ortonormale.

Esempio 9.29. Alcune importanti tipologie di endomorfismi di uno spazio vettoriale euclideo V che abbiamo già incontrato sono autoaggiunti. Oltre al caso banale $f=\mathrm{Id}_V$, tra questi vi sono la proiezione ortogonale $\mathcal{P}_U$ e la riflessione $\mathcal{R}_U$ rispetto ad un sottospazio U. Infatti, supponiamo che $\dim(U) = n$, $\dim(U^\perp) = m$ e siano $\mathcal{B}_U,\mathcal{B}_{U^\perp}$ basi ortonormali dei due sottospazi $U,U^\perp$. Allora rispetto a $\mathcal{B}=\mathcal{B}_U\cup\mathcal{B}_{U^\perp}$ le due applicazioni sono rappresentate da

$$\phi_{\mathcal{B}}(\mathcal{P}_U) = \left[\begin{array}{c|c} \mathbb{I}_n & 0_{nm} \\ \hline 0_{mn} & 0_{mm}\end{array}\right] \qquad \text{e} \qquad \phi_{\mathcal{B}}(\mathcal{R}_U) = \left[\begin{array}{c|c} \mathbb{I}_n & 0_{nm} \\ \hline 0_{mn} & -\mathbb{I}_m\end{array}\right],$$

che sono chiaramente matrici simmetriche.

Consideriamo invece una rotazione bidimensionale f di un angolo α. La matrice rappresentativa rispetto ad una base ortonormale è

$$F|_{\mathcal{B}} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}.$$

Non essendo simmetrica, l'applicazione non è autoaggiunta. L'applicazione f^* è rappresentata da

$$F^*|_{\mathcal{B}} = F|_{\mathcal{B}}^T = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} = \begin{bmatrix} \cos(-\alpha) & \sin(-\alpha) \\ -\sin(-\alpha) & \cos(-\alpha) \end{bmatrix},$$

ovvero f^* è ancora una rotazione, dello stesso angolo convesso α, ma con senso di rotazione opposto rispetto ad f.

9.6. Teorema spettrale

In uno spazio vettoriale euclideo sappiamo che le basi ortonormali svolgono un ruolo privilegiato rispetto alla altre basi. Questo è vero anche nei problemi di diagonalizzazione.

Definizione 9.30. *Siano V uno spazio vettoriale euclideo ed $f \in \mathrm{End}(V)$. Allora l'applicazione f si dice ortogonalmente diagonalizzabile se esiste una base ortonormale di V composta da autovettori di f.*

Le applicazioni ortogonalmente diagonalizzabili in spazi finitamente generati possono essere caratterizzate in termini delle loro matrici rappresentative.

Lemma 9.31. *Siano V uno spazio vettoriale euclideo finitamente generato e $\mathcal{B}$ una sua base ortonormale. La base $\mathcal{B}'$ è anch'essa ortonormale se e solo se le matrici del cambiamento di base $M_{\mathcal{B}\mathcal{B}'}$ e $M_{\mathcal{B}'\mathcal{B}}$ sono ortogonali.*

Dimostrazione. Sia $\mathcal{B}' = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$, allora

$$M_{\mathcal{B}'\mathcal{B}} = \begin{bmatrix} \mathbf{v_1}|_{\mathcal{B}} & \dots & \mathbf{v_n}|_{\mathcal{B}} \end{bmatrix}$$

è ortogonale se e solo se $\mathbf{v_i}|_{\mathcal{B}}^T * \mathbf{v_j}|_{\mathcal{B}} = \langle \mathbf{v_i}, \mathbf{v_j} \rangle = \delta_{ij}$, per ogni $1 \leq i, j \leq n$. Questo è vero se e solo se $\mathcal{B}'$ è ortonormale. Infine, poiché $M_{\mathcal{B}\mathcal{B}'} = M_{\mathcal{B}'\mathcal{B}}^{-1}$, una matrice è ortogonale se e solo se anche l'altra lo è. □

Proposizione 9.32. *Siano V uno spazio vettoriale euclideo soddisfacente $\dim(V) = n < \infty$, $\mathcal{B}$ una base ortonormale di V ed $f \in \mathrm{End}(V)$. Allora f è ortogonalmente diagonalizzabile se e solo se esiste una matrice $Q \in O(n; \mathbb{R})$ tale che*

$$Q^T * F|_{\mathcal{B}} * Q = D,$$

dove D è una matrice diagonale.

Dimostrazione. Per il criterio enunciato nella Proposizione 6.17, f è diagonalizzabile se e solo se esiste $S \in GL(n; R)$ tale che $S^{-1} * F|_{\mathcal{B}} * S = D$. In tal caso, la matrice S è la matrice del cambiamento di base $M_{\mathcal{B}'\mathcal{B}}$ da una base di autovettori $\mathcal{B}'$ alla base iniziale $\mathcal{B}$. Per il Lemma 9.31, la base $\mathcal{B}'$ è ortonormale se e solo se $M_{\mathcal{B}'\mathcal{B}}$ è una matrice ortogonale Q, la cui inversa è $Q^{-1} = Q^T$, da cui segue la tesi. □

Analizzando le rispettive definizioni e proprietà, non sembra esistere alcun legame apparente tra i concetti studiati nelle ultime due sezioni, ovvero gli endomorfismi autoaggiunti e quelli ortogonalmente diagonalizzabili. Per questa ragione, il Teorema spettrale è uno dei risultati più sorprendenti dell'algebra lineare.

Teorema 9.33. *Sia V uno spazio vettoriale euclideo finitamente generato. Allora $f \in \mathrm{End}(V)$ è autoaggiunto se e solo se è ortogonalmente diagonalizzabile.*

Esempio 9.34. Come abbiamo visto nell'Esempio 9.29, le proiezioni e le riflessioni ortogonali rispetto ad un sottospazio U di uno spazio vettoriale V sono endomorfismi autoaggiunti. Il Teorema spettrale garantisce quindi che esse siano ortogonalmente diagonalizzabili. Infatti, l'insieme costituito dall'unione di una base ortonormale di U ed una base ortonormale di $U^{\perp}$ è una base ortonormale di V, composta da autovettori per entrambe le applicazioni. Al contrario, consideriamo una rotazione di un angolo $\alpha \neq 0$ in uno spazio bidimensionale. Nell'Esempio 9.29 abbiamo mostrato che tale applicazione non è autoaggiunta. Questo è confermato dal fatto che essa non è diagonalizzabile e quindi tanto meno può essere ortogonalmente diagonalizzabile.

La dimostrazione del Teorema spettrale richiede la verifica di alcuni risultati intermedi. Il primo riguarda gli autovalori di una matrice simmetrica e per dimostrarlo avremo bisogno di lavorare nello spazio $\mathrm{Mat}(n, 1; \mathbb{C})$. A tal fine, osserviamo preliminarmente che ogni elemento $Z \in \mathrm{Mat}(n, 1; \mathbb{C})$ può essere decomposto come

$$Z = \begin{bmatrix} z_{11} \\ \vdots \\ z_{n,1} \end{bmatrix} = \begin{bmatrix} x_{11} + i\, y_{11} \\ \vdots \\ x_{n,1} + i\, y_{n,1} \end{bmatrix} = \begin{bmatrix} x_{11} \\ \vdots \\ x_{n,1} \end{bmatrix} + i \cdot \begin{bmatrix} y_{11} \\ \vdots \\ y_{n,1} \end{bmatrix} = X + i \cdot Y,$$

dove $X, Y \in \mathrm{Mat}(n, 1; \mathbb{R})$ si dicono rispettivamente parte reale ed immaginaria di Z. Definiamo quindi il complesso coniugato di Z come $\bar{Z} = X - i \cdot Y$. Ovviamente $Z \in \mathrm{Mat}(n, 1; \mathbb{R})$ se e solo se $\bar{Z} = Z$.

Lemma 9.35. *Il polinomio caratteristico di una matrice $A \in \mathbb{S}(n; \mathbb{R})$ ha n radici reali contate con molteplicità.*

Dimostrazione. Dato che $\mathbb{R} \subset \mathbb{C}$, la matrice A è anche un elemento di $\mathbb{S}(n; \mathbb{C})$ e quindi definisce un'applicazione associata $f_A \in \mathrm{End}(\mathrm{Mat}(n, 1; \mathbb{C}))$. Dalla Proposizione 6.14, sappiamo che il polinomio caratteristico P_A ha grado n. Essendo $\mathbb{C}$ un campo algebricamente chiuso, P_A ha esattamente n radici nel campo complesso, contate con molteplicità. Ogni radice λ corrisponde ad un autovalore di f_A, ovvero esiste un autovettore $Z \in \mathrm{Mat}(n, 1; \mathbb{C})$ associato a λ tale che $A * Z = \lambda \cdot Z$. Moltiplicando entrambi i membri dell'identità a sinistra per $\bar{Z}^T$, otteniamo

$$\bar{Z}^T * A * Z = \bar{Z}^T * (\lambda \cdot Z) = \lambda \cdot (\bar{Z}^T * Z).$$

Se assumiamo che $\bar{Z}^T * Z \neq 0_{11}$, potremmo sfruttare l'identificazione naturale tra $\mathrm{Mat}(1, 1; \mathbb{C})$ e $\mathbb{C}$ per ottenere

$$\lambda = \frac{\bar{Z}^T * A * Z}{\bar{Z}^T * Z}.$$

Tale relazione ci servirà per dimostrare che $\lambda \in \mathbb{R}$. Analizziamo singolarmente il numeratore ed il denominatore della frazione.

- Per le proprietà del complesso coniugato, abbiamo

  $$\overline{\bar{Z}^T * A * Z} = Z^T * \bar{A} * \bar{Z}.$$

  Dato che questa quantità è una matrice di ordine 1, il suo trasposto coincide con la matrice stessa, quindi

  $$Z^T * \bar{A} * \bar{Z} = (Z^T * \bar{A} * \bar{Z})^T = \bar{Z}^T * \bar{A}^T * Z.$$

  Infine, poiché A è una matrice reale simmetrica, vale

  $$\bar{Z}^T * \bar{A}^T * Z = \bar{Z}^T * A * Z.$$

  Queste uguaglianze implicano che il numeratore della precedente frazione è reale.
- Grazie alla decomposizione in parte reale ed immaginaria di Z, possiamo scrivere

  $$\bar{Z}^T * Z = (X - i \cdot Y)^T (X + i \cdot Y) = (X^T * X + Y^T * Y) + i \cdot (X^T * Y - Y^T * X).$$

  Poiché $X, Y \in \mathrm{Mat}(n, 1; \mathbb{R})$, la componente reale del prodotto può essere identificata con la somma delle norme di Frobenius al quadrato delle due matrici, mentre la componente complessa è uguale alla differenza di due prodotti scalari aventi lo stesso valore:

  $$\begin{aligned}(X^T * X + Y^T * Y) + i \cdot (X^T * Y - Y^T * X) &= (\|X\|_F^2 + \|Y\|_F^2) + i\,(\langle X, Y\rangle_F - \langle Y, X\rangle_F) = \\ &= \|X\|_F^2 + \|Y\|_F^2 .\end{aligned}$$

Dato che $Z \neq 0_{n1}$ in quanto è un autovettore di f_A, segue che X, Y non possono essere entrambi nulli e quindi la somma delle loro norme è un numero reale strettamente positivo.

In conclusione, essendo λ il quoziente di due quantità reali, a sua volta deve essere un numero reale. □

Il secondo risultato preliminare riguarda il comportamento di un'applicazione autoaggiunta su un suo autospazio e sul relativo complemento ortogonale.

Lemma 9.36. *Siano V uno spazio vettoriale euclideo ed f un endomorfismo autoaggiunto con autovalore λ. Allora i sottospazi V_λ e $V_\lambda^\perp$ sono invarianti rispetto ad f, ovvero $f(V_\lambda) \subseteq V_\lambda$ ed $f(V_\lambda^\perp) \subseteq V_\lambda^\perp$.*

Dimostrazione. L'invarianza di V_λ segue dalle definizioni di autovettore e di autospazio. Infatti, se $\mathbf{v} \in V_\lambda$ allora $f(\mathbf{v}) = \lambda \cdot \mathbf{v}$. Dato che V_λ è un sottospazio, esso è chiuso rispetto alla moltiplicazione per uno scalare e quindi $f(\mathbf{v}) \in V_\lambda$, che è equivalente al primo punto della tesi. Sia invece $\mathbf{v}' \in V_\lambda^\perp$. Allora vale

$$\langle \mathbf{v}, f(\mathbf{v}') \rangle = \langle f(\mathbf{v}), \mathbf{v}' \rangle = \langle \lambda \cdot \mathbf{v}, \mathbf{v}' \rangle = \lambda \langle \mathbf{v}, \mathbf{v}' \rangle = 0$$

per ogni $\mathbf{v} \in V$. Quindi $f(\mathbf{v}') \in V_\lambda^\perp$, che completa la dimostrazione. □

Dall'ultimo lemma segue che le restrizioni di un'endomorfismo autoaggiunto f ad ogni suo autospazio ed al relativo complemento ortogonale sono ancora endomorfismi. In particolare, essi saranno ancora autoaggiunti, ovvero $f|_{V_\lambda} \in \mathbb{S}(V_\lambda)$ ed $f|_{V_\lambda^\perp} \in \mathbb{S}(V_\lambda^\perp)$, per ogni autovalore λ di f.

Possiamo a questo punto affrontare la dimostrazione del Teorema spettrale.

Dimostrazione del Teorema 9.33. Verifichiamo il primo verso dell'implicazione. Per prima cosa, osserviamo che se f è un endomorfismo autoaggiunto allora la sua matrice rappresentativa rispetto ad una qualunque base ortonormale di V è simmetrica. A causa del Lemma 9.35, esistono quindi n autovalori di f, contati tenendo conto della loro molteplicità algebrica. Nel caso particolare in cui esista un autovalore λ di molteplicità geometrica $g = n$, allora $V_\lambda = V$ e quindi ogni vettore di V è un autovettore di f. Pertanto, una qualunque base ortonormale di V è anche una base di autovettori dell'endomorfismo e di conseguenza f è ortogonalmente diagonalizzabile.

Dimostriamo il teorema nel caso generale attraverso l'induzione sulla dimensione n dello spazio V. Se $n = 1$ rientriamo nel caso speciale precedente, poiché necessariamente esiste un solo autovalore la cui molteplicità geometrica è $g = 1 = n$.

Come ipotesi induttiva, supponiamo ora che il teorema spettrale sia vero in tutti gli spazi vettoriali di dimensione minore di n. Dobbiamo provare che questo implica la sua validità anche su uno spazio V di dimensione n. Sia λ un autovalore di molteplicità geometrica $1 \leq g < n$. Il complemento ortogonale $V_\lambda^\perp$ ha dimensione $1 \leq n - g < n$. Per il Lemma 9.36, abbiamo quindi due endomorfismi autoaggiunti $f|_{V_\lambda}$ e $f|_{V_\lambda^\perp}$ definiti su spazi vettoriali non banali di dimensione minore di n, per i quali possiamo utilizzare l'ipotesi induttiva. Di conseguenza, esistono due basi ortonormali $\mathcal{B}_{V_\lambda}$ e $\mathcal{B}_{V_\lambda^\perp}$ dei rispettivi sottospazi, composte da autovettori delle due restrizioni e quindi anche di f. Poiché $V = V_\lambda \oplus V_{\lambda^\perp}$, per le proprietà di somma diretta l'insieme $\mathcal{B} = \mathcal{B}_{V_\lambda} \cup \mathcal{B}_{V_\lambda^\perp}$ è una base di V, che per costruzione è ortonormale e composta da autovettori di f. Questo dimostra che f è ortogonalmente diagonalizzabile.

La dimostrazione dell'implicazione inversa è molto più semplice. Infatti, supponiamo che f sia ortogonalmente diagonalizzabile. Per ipotesi esiste una base ortonormale $\mathcal{B}$ di V composta da autovettori di f. Rispetto a tale base, la matrice rappresentativa $F|_{\mathcal{B}}$ è ovviamente diagonale e quindi simmetrica. Di conseguenza, dalla Proposizione 9.28 segue che l'applicazione f è autoaggiunta. □

Corollario 9.37. *Siano V uno spazio vettoriale euclideo ed $f \in \mathbb{S}(V)$ con λ_1, λ_2 due autovalori distinti. Allora i relativi autospazi soddisfano $V_{\lambda_1} \perp V_{\lambda_2}$.*

DIMOSTRAZIONE. Dalla dimostrazione del teorema spettrale sappiamo che

$$V = V_{\lambda_1} \oplus V_{\lambda_1}^{\perp}$$

e che a sua volta $V_{\lambda_1}^{\perp}$ è somma diretta di autospazi di f, ovviamente distinti da V_{λ_1} :

$$V_{\lambda_1}^{\perp} = \bigoplus_{i=2}^{n} V_{\lambda_i}.$$

Di conseguenza

$$V = \bigoplus_{i=1}^{n} V_{\lambda_i}.$$

Ricordiamo che la decomposizione di V come somma diretta di autospazi di un'applicazione è unica a meno di riordinamenti. Questo implica che $V_{\lambda_1}^{\perp}$ è somma diretta di tutti gli autospazi di f diversi da V_{λ_1}, che sono quindi tutti ortogonali V_{λ_1} stesso. □

Nei libri di algebra lineare centrati sul calcolo matriciale, il teorema spettrale viene spesso presentato in termini di decomposizioni di matrici simmetriche. Per legare tale versione del teorema alla nostra trattazione, ricordiamo i seguenti fatti:

- ogni matrice $A \in \mathrm{Mat}(n, n; \mathbb{R})$ definisce un endomorfismo $f_A \in \mathrm{End}(V)$ su $V = \mathrm{Mat}(n, 1; \mathbb{R})$;
- la matrice che rappresenta f_A rispetto alla base canonica $\mathcal{B}_{n1}$ di V è A stessa;
- rispetto alla struttura euclidea canonica su V, definita dal prodotto di Frobenius, la base $\mathcal{B}_{n1}$ è ortonormale.

Allora il teorema spettrale implica la seguente proposizione.

Proposizione 9.38. *Una matrice $A \in \mathrm{Mat}(n, n; \mathbb{R})$ può essere decomposta come il prodotto*

$$A = Q * D * Q^T$$

con $Q \in O(n; \mathbb{R})$ e $D \in \mathbb{D}(n; \mathbb{R})$ se e solo se A è simmetrica. In tal caso, le entrate sulla diagonale di D sono gli autovalori di A e le colonne di Q sono i corrispondenti autovettori.

DIMOSTRAZIONE. Per la Proposizione 9.32, f_A è ortogonalmente diagonalizzabile in V se e solo se vale

$$D = Q^T * A * Q$$

e quindi

$$A = Q * D * Q^T.$$

Per il Teorema spettrale, questo può succedere se e solo se f_A è autoaggiunta e quindi, per la Proposizione 9.28, se e solo se A è simmetrica. L'ultima affermazione segue dalle proprietà generali delle applicazioni diagonalizzabili. □

Esempio 9.39. Consideriamo la matrice

$$A = \begin{bmatrix} 0 & 2a-1 & 1 \\ a & 0 & b \\ b & a & 0 \end{bmatrix},$$

dove $a, b \in \mathbb{R}$. La matrice è simmetrica se e solo se valgono

$$\begin{cases} 2a - 1 = a \\ 1 = b \\ b = a \end{cases} \qquad \Rightarrow \qquad \begin{cases} a = 1 \\ b = 1 \end{cases}.$$

In tal caso, il polinomio caratteristico di A è

$$P_A(\lambda) = \det(A - \lambda \cdot \mathbb{I}_3) = \begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = -\lambda^3 + 3\lambda + 2.$$

Osserviamo che -1 è una radice del polinomio, in quanto

$$V_{-1} = \ker(A + \mathbb{I}_3) = \ker\left(\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}\right) = \mathcal{L}\left(\mathbf{v_1} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{v_2} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}\right).$$

Dato che A è diagonalizzabile, la molteplicità algebrica di -1 è uguale a 2, poiché essa coincide con quella geometrica. Quindi esiste un secondo autovalore $\bar{\lambda}$, tale che

$$P_A(\lambda) = -\lambda^3 + 3\lambda + 2 = -(\lambda + 1)^2(\lambda - \bar{\lambda}).$$

Uguagliando i termini costanti delle due espressioni, otteniamo $\bar{\lambda} = 2$. L'autospazio associato è

$$V_2 = \ker(A + \mathbb{I}_3) = \left(\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}\right) = \left(\begin{bmatrix} 1 & 1 & -2 \\ 0 & -3 & 3 \\ 0 & 3 & -3 \end{bmatrix}\right) = \mathcal{L}\left(\mathbf{v_3} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right).$$

Come si può osservare, $V_{-1} \perp V_2$ ma la base di V_{-1} non è ortonormale. Dato che $V_{-1} = V_2^{\perp}$, al fine di ottenerne un autovettore associato a -1 ed ortogonale a $\mathbf{v_1}$, possiamo calcolare il prodotto vettoriale tra $\mathbf{v_1}$ e $\mathbf{v_3}$:

$$\tilde{\mathbf{v}}_\mathbf{2} = \mathbf{v_1} \wedge \mathbf{v_3} = \begin{vmatrix} 1 & 1 & E_{11} \\ -1 & 1 & E_{21} \\ 0 & 1 & E_{31} \end{vmatrix} = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}.$$

Dopo aver normalizzato gli autovettori, otteniamo una base ortonormale di V :

$$\mathcal{B} = \left\{ \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \right\}.$$

Quindi, la matrice A può essere decomposta come

$$A = Q * D * Q^T = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix} * \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} * \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ -1/\sqrt{6} & -1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}.$$

Concludiamo la sezione presentando un'applicazione del teorema spettrale. Nell'Esempio 8.5 abbiamo visto che una matrice $G \in \mathbb{S}(n; \mathbb{R})$ definisce un prodotto scalare nello spazio $\mathrm{Mat}(n, 1; \mathbb{R})$ se e solo se è definita positiva, ovvero se vale $X^T * G * X > 0$ per ogni $X \neq 0_{n1}$. A tal proposito, vale il seguente criterio.

Proposizione 9.40. *La matrice $G \in \mathbb{S}(n; \mathbb{R})$ è definita positiva se e solo se tutti i suoi autovalori sono strettamente positivi.*

Dimostrazione. Il Teorema 9.33 garantisce che esiste una base ortonormale $\mathcal{B} = \{X_1, \ldots, X_n\}$ di $\mathrm{Mat}(n, 1; \mathbb{R})$ composta da autovettori di A, rispetto al prodotto scalare di Frobenius. Allora, per ogni $X \in \mathrm{Mat}(n, 1; \mathbb{R})$ esiste la decomposizione

$$X = \sum_{i=1}^{n} t_i \cdot X_i.$$

Dati $\lambda_1, \ldots, \lambda_n$ gli autovalori associati agli autovettori nella base $\mathcal{B}$, vale

$$\begin{aligned} X^T * G * X &= \left(\sum_{i=1}^{n} t_i \cdot X_i \right)^T * G * \left(\sum_{j=1}^{n} t_j \cdot X_j \right) = \\ &= \left(\sum_{i=1}^{n} t_i \cdot X_i \right)^T * \left(\sum_{j=1}^{n} t_j \cdot (G * X_j) \right) = \\ &= \left(\sum_{i=1}^{n} t_i \cdot X_i \right)^T * \left(\sum_{j=1}^{n} t_j \cdot (\lambda_j \cdot X_j) \right) = \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} t_i t_j \lambda_j \cdot (X_i^T * X_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} t_i t_j \lambda_j \delta_{ij} = \sum_{i=1}^{n} t_i^2 \lambda_i. \end{aligned}$$

Chiaramente il prodotto è positivo per ogni $X \neq 0_{n1}$ se e solo se gli autovalori sono tutti strettamente positivi. $\square$

Esempio 9.41. Studiamo quali delle matrici

$$G_h = \begin{bmatrix} 3 & 1 & h \\ 1 & 1 & 1 \\ h & 1 & 2 \end{bmatrix}$$

definiscono un prodotto scalare, al variare di $h \in \mathbb{R}$. Il polinomio caratteristico è

$$P_h(\lambda) = \det(G_h) - II(G_h)\lambda + \mathrm{Tr}(G_h)\lambda^2 - \lambda^3 = (1 + 2h - h^2) + (-9 + h^2)\lambda + 6\lambda^2 - \lambda^3.$$

Grazie al Teorema spettrale sappiamo che le radici di P_h sono tutte reali. Il loro segno è determinato dalla regola di Cartesio, ovvero il numero di radici strettamente positive è uguale al numero di variazioni del segno tra i coefficienti ordinati del polinomio, trascurando eventuali coefficienti nulli. Il segno dei coefficienti di grado 2 e 3 è fissato, dobbiamo invece studiare il segno dei coefficienti di grado 0 e 1 in funzione di h :

- $c_0 = 1 + 2h - h^2$ ha radici $\bar{h}_\pm = 1 \pm \sqrt{2}$, è positivo in $(\bar{h}_-, \bar{h}_+)$ ed è negativo in $(-\infty, \bar{h}_-) \cup (\bar{h}_+, +\infty)$;
- $c_1 = -9 + h^2$ ha radici $\tilde{h}_\pm = \pm 3$, è positivo in $(-\infty, \tilde{h}_-) \cup (\tilde{h}_+, +\infty)$ ed è negativo in $(\tilde{h}_-, \tilde{h}_+)$.

I segni dei coefficienti sono alternati se e solo se $h \in (\bar{h}_-, \bar{h}_+)$. In tal caso le radici del polinomio sono tutte positive, di conseguenza G_h è definita positiva e determina un prodotto scalare in $\mathrm{Mat}(3, 1; \mathbb{R})$.

9.7. Decomposizione ai valori singolari

Il Teorema spettrale, e più in generale la teoria degli autovalori ed autovettori, sono strumenti molto potenti per lo studio delle applicazioni lineari e delle matrici, però essi possono essere pienamente sfruttati solo nel caso di endomorfismi diagonalizzabili, possibilmente simmetrici. È lecito chiedersi se sia possibile generalizzare tali mezzi di indagine per lo studio di omomorfismi generici, tra spazi vettoriali potenzialmente differenti, e di matrici non necessariamente quadrate. Tale quesito trova risposta positiva nella teoria della decomposizione ai valori singolari di una applicazione lineare o di una matrice, tecnica che svolge un ruolo di primaria importanza in un'enorme quantità di applicazioni scientifiche ed ingegneristiche[1].

In ciò che segue assumeremo che $(V, \langle\ ,\ \rangle_V)$ e $(W, \langle\ ,\ \rangle_W)$ siano spazi vettoriali euclidei finitamente generati, con $\dim(V) = n$ e $\dim(W) = m$, e che f sia un omomorfismo in $\mathrm{Hom}(V, W)$.

Lemma 9.42. *$f^* \circ f \in \mathbb{S}(V)$ ed $f \circ f^* \in \mathbb{S}(W)$ ed entrambi hanno autovalori non negativi.*

Dimostrazione. Verifichiamo l'enunciato per l'applicazione $f^* \circ f$, essendo quello su $f \circ f^*$ analogo. Per prima cosa, vale $(f^* \circ f)^* = f^* \circ (f^*)^* = f^* \circ f$ e quindi $f^* \circ f$ è autoaggiunto. Per il Teorema spettrale, esiste una base ortonormale $\mathcal{B}_V = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ di V composta da autovettori di $f^* \circ f$, per i quali vale

$$(f^* \circ f)(\mathbf{v_i}) = \lambda_i \cdot \mathbf{v_i}$$

per ogni $1 \leq i \leq n$. Di conseguenza abbiamo

$$\langle \mathbf{v_i}, f^* \circ f(\mathbf{v_i}) \rangle_V = \langle \mathbf{v_i}, \lambda_i \cdot \mathbf{v_i} \rangle_V = \lambda_i \left\| \mathbf{v_i} \right\|_V^2 .$$

D'altra parte, vale anche

$$\langle \mathbf{v_i}, f^* \circ f(\mathbf{v_i}) \rangle_V = \langle f(\mathbf{v_i}), f(\mathbf{v_i}) \rangle_W = \left\| f(\mathbf{v_i}) \right\|_W^2$$

e quindi $\lambda_i = \dfrac{\left\| f(\mathbf{v_i}) \right\|_W^2}{\left\| \mathbf{v_i} \right\|_V^2} \geq 0$ per ogni $1 \leq i \leq n$. □

Definizione 9.43. *Sia $\dim(V) = n$, $\dim(W) = m$. Tenendo conto delle loro molteplicità, ordiniamo rispettivamente gli autovalori di $f^* \circ f$ e di $f \circ f^*$ in modo tale che $\lambda_1 \geq \cdots \geq \lambda_n$ e $\lambda_1^* \geq \cdots \geq \lambda_m^*$. Allora:*

i) un autovettore di $f^* \circ f$ è detto vettore singolare destro di f, mentre $\sigma_i = \sqrt{\lambda_i}$ è detto valore singolare destro di f, per $1 \leq i \leq n$;

ii) un autovettore di $f \circ f^*$ è detto vettore singolare sinistro di f, mentre $\sigma_i^* = \sqrt{\lambda_i^*}$ è detto valore singolare sinistro di f, per $1 \leq i \leq m$.

Osserviamo che se f è un endomorfismo autoaggiunto allora $f^* \circ f = f \circ f^* = f^2$ e quindi i valori singolari a destra e sinistra coincidono e sono uguali ai valori assoluti degli autovalori di f. Analogamente, i vettori singolari sono uguali agli autovettori di f.

Esempio 9.44. Dati $V = \mathrm{Mat}(2, 1; \mathbb{R})$ e $W = \mathrm{Mat}(3, 1; \mathbb{R})$, siano $f_A \in \mathrm{Hom}(V, W)$ ed $f_A^* \in \mathrm{Hom}(W, V)$ le applicazioni definite rispettivamente dalle matrici

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \qquad \text{e} \qquad A^T = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Allora $f_A^* \circ f_A$ ed $f_A \circ f_A^*$ sono rispettivamente associate alle matrici

$$A^T * A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \qquad \text{e} \qquad A * A^T = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & 2 \end{bmatrix}.$$

Gli autovalori di $f_A^* \circ f_A$ sono $\lambda_1 = 3$ e $\lambda_2 = 1$, di conseguenza i valori singolari a destra di f_A sono $\sigma_1 = \sqrt{3}$ e $\sigma_2 = 1$. Gli autovalori di $f_A \circ f_A^*$ sono invece $\lambda_1^* = 3 = \lambda_1$, $\lambda_2^* = 1 = \lambda_2$ e $\lambda_3^* = 0$, quindi i valori singolari a sinistra di f_A sono $\sigma_1^* = \sqrt{3} = \sigma_1$, $\sigma_2^* = 1 = \sigma_2$ e $\sigma_3^* = 0$.

[1]Per un'introduzione al problema e ad alcuni suoi risvolti applicativi, suggeriamo la lettura dell'articolo divulgativo We Recommend a Singular Value Decomposition, sul sito dell'American Mathematical Society.

La coincidenza tra i valori singolari non nulli a destra e a sinistra non è un incidente, come è dimostrato nella seguente proposizione.

Proposizione 9.45. *Sia σ un numero reale strettamente positivo. Allora le seguenti affermazioni sono equivalenti:*

i) esiste una coppia $(\mathbf{v}, \mathbf{w}) \in V \times W$ tale che

$$f(\mathbf{v}) = \sigma \cdot \mathbf{w} \qquad \textit{ed} \qquad f^*(\mathbf{w}) = \sigma \cdot \mathbf{v};$$

ii) σ è valore singolare destro di f;

iii) σ è valore singolare sinistro di f.

In tal caso, $\mathbf{v}$ è un vettore singolare destro e $\mathbf{w}$ è un vettore singolare sinistro di f.

Dimostrazione. Supponiamo che valga (i). Allora

$$(f^* \circ f)(\mathbf{v}) = f^*(\sigma \cdot \mathbf{w}) = \sigma^2 \cdot \mathbf{v} \qquad \text{ed} \qquad (f \circ f^*)(\mathbf{w}) = f(\sigma \cdot \mathbf{v}) = \sigma^2 \cdot \mathbf{w},$$

da cui seguono i punti (ii) e (iii) e l'ultima affermazione dell'enunciato.

Supponiamo invece che valga (ii). Allora esiste $\mathbf{v} \in V$ tale che $(f^* \circ f)(\mathbf{v}) = \sigma^2 \cdot \mathbf{v}$. In W definiamo il vettore $\mathbf{w} = \frac{1}{\sigma} \cdot f(\mathbf{v})$. Allora ovviamente vale $f(\mathbf{v}) = \sigma \cdot \mathbf{w}$ ed inoltre

$$f^*(\mathbf{w}) = f^*\left(\frac{1}{\sigma} \cdot f(\mathbf{v})\right) = \frac{1}{\sigma} \cdot (f^* \circ f)(\mathbf{v}) = \sigma \cdot \mathbf{v},$$

ovvero abbiamo provato la validità del punto (i). Analogamente si verifica anche che (iii) implica (i). La dimostrazione si completa componendo le varie implicazioni. □

Definizione 9.46. *I valori singolari non nulli a destra ed a sinistra sono detti valori singolari di f.*

Possiamo ora enunciare e dimostrare il risultato principale della sezione riguardante le applicazioni lineari, che chiamiamo Teorema della decomposizione ai valori singolari.

Teorema 9.47. *Siano $\sigma_1 \geq \cdots \geq \sigma_r > 0$ i valori singolari di f. Allora esistono due basi ortonormali $\mathcal{B}_V = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ di V e $\mathcal{B}_W = \{\mathbf{w_1}, \ldots, \mathbf{w_m}\}$ di W, rispettivamente composte da vettori singolari a destra ed a sinistra di f, tali che*

$$\begin{cases} f(\mathbf{v_i}) = \sigma_i \cdot \mathbf{w_i} & \textit{se} \quad 1 \leq i \leq r, \\ f(\mathbf{v_i}) = \mathbf{0_W} & \textit{se} \quad r < i \leq n \end{cases} \qquad \textit{ed} \qquad \begin{cases} f^*(\mathbf{w_i}) = \sigma_i \cdot \mathbf{v_i} & \textit{se} \quad 1 \leq i \leq r, \\ f^*(\mathbf{w_i}) = \mathbf{0_V} & \textit{se} \quad r < i \leq m. \end{cases}$$

Inoltre, per ogni $\mathbf{v} \in V$ vale la decomposizione

$$f(\mathbf{v}) = \sum_{i=1}^{r} \sigma_i \langle \mathbf{v_i}, \mathbf{v} \rangle_V \cdot \mathbf{w_i}.$$

Dimostrazione. L'esistenza di una base ortonormale $\mathcal{B}_V = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ composta da vettori singolari a destra di f è garantita dal Lemma 9.42 e dal Teorema spettrale. Senza perdere in generalità, supponiamo che tali vettori siano ordinati in maniera coerente con i valori singolari di f. In tal caso, osserviamo subito che $f(\mathbf{v_i}) = \mathbf{0_W}$ per ogni $r < i \leq n$. Infatti

$$\|f(\mathbf{v_i})\|_W^2 = \langle f(\mathbf{v_i}), f(\mathbf{v_i}) \rangle_W = \langle \mathbf{v_i}, (f^* \circ f)(\mathbf{v_i}) \rangle_V = \langle \mathbf{v_i}, \lambda_i \cdot \mathbf{v_I} \rangle_V = \langle \mathbf{v_i}, \mathbf{0_V} \rangle_V = 0,$$

poiché per ipotesi $\lambda_i = 0$ se $r < i \leq m$. Ne segue anche che $\operatorname{Im}(f) = \mathcal{L}(f(\mathbf{v_1}), \ldots, f(\mathbf{v_r}))$.

Costruiamo invece $\mathcal{B}_W$ nel seguente modo:

- definiamo $\mathbf{w_i} = \frac{1}{\sigma_i} \cdot f(\mathbf{v_i})$ per ogni $1 \leq i \leq r$. Tali vettori costituiscono una base ortonormale di $\mathrm{Im}(f)$, poiché vale

$$\langle \mathbf{w_i}, \mathbf{w_j} \rangle_W = \left\langle \frac{1}{\sigma_i} \cdot f(\mathbf{v_i}), \frac{1}{\sigma_j} \cdot f(\mathbf{v_j}) \right\rangle_W = \frac{1}{\sigma_i \sigma_j} \langle f(\mathbf{v_i}), f(\mathbf{v_j}) \rangle_W = \frac{1}{\sigma_i \sigma_j} \langle \mathbf{v_i}, (f^* \circ f)(\mathbf{v_j}) \rangle_V =$$
$$= \frac{1}{\sigma_i \sigma_j} \langle \mathbf{v_i}, \lambda_j \cdot \mathbf{v_j} \rangle_V = \frac{\lambda_j}{\sigma_i \sigma_j} \langle \mathbf{v_i}, \mathbf{v_j} \rangle_V = \frac{\lambda_j}{\sigma_i \sigma_j} \delta_{ij} = \frac{\lambda_i}{\sigma_i^2} \delta_{ij} = \delta_{ij}\,;$$

- dato che $\dim(\mathrm{Im}(f)^\perp) = \dim(V) - \dim(\mathrm{Im}(f)) = m - r$, scegliamo i vettori $\mathbf{w_i}$ per $r < i \leq m$ in modo che essi costituiscano una base ortonormale di $\mathrm{Im}(f)^\perp$.

Per le proprietà del complemento ortogonale, l'insieme $\mathcal{B}_W = \{\mathbf{w_1}, \ldots, \mathbf{w_m}\}$ è una base ortonormale di W.

Per costruzione e per quanto visto nella dimostrazione della Proposizione 9.45, i vettori $\mathbf{v_i}$ e $\mathbf{w_i}$ per $1 \leq i \leq r$ soddisfano le proprietà dell'enunciato. Precedentemente abbiamo anche dimostrato che $f(\mathbf{v_i}) = \mathbf{0_W}$ per ogni $r < i \leq n$. Rimangono da provare le proprietà dei vettori $\mathbf{w_i}$ per $r < i \leq n$. Essi verificano $f^*(\mathbf{w_i}) = \mathbf{0_V}$, poiché $\mathrm{Im}(f)^\perp = \ker(f^*)$ grazie al Corollario 9.26. Inoltre, essi sono vettori normali a sinistra di f, poiché vale $(f \circ f^*)(\mathbf{w_i}) = f(f^*(\mathbf{w_i})) = f(\mathbf{0_V}) = \mathbf{0_W} = 0 \cdot \mathbf{w_i}$.

Infine, la validità della decomposizione di $f(\mathbf{v})$ si può verificare facilmente sui vettori della base $\mathcal{B}_V$. Grazie al teorema di interpolazione delle applicazioni lineare, segue poi la validità della formula per ogni $\mathbf{v} \in V$. □

Dalla dimostrazione del teorema si deduce che ad ogni base $\mathcal{B}_V$ di vettori singolari a destra di f si associa una base $\mathcal{B}_W$ di vettori singolari a sinistra, che soddisfa le proprietà enunciate. $\mathcal{B}_W$ è unica a meno di isometrie nello spazio $\mathrm{Im}(f)^\perp$. Ragionando analogamente sull'applicazione aggiunta f^*, si può dimostrare anche l'implicazione speculare. Si dice che $\mathcal{B}_V$ e $\mathcal{B}_W$ sono basi duali associate ai valori singolari di f. Dall'enunciato del teorema precedente e dalla sua dimostrazione, si ricavano facilmente alcune proprietà interessanti della decomposizione ai valori singolari di una applicazione. Le raccogliamo nel seguente corollario, la cui verifica è lasciata come esercizio al lettore.

Corollario 9.48. *Siano $\sigma_1 \geq \cdots \geq \sigma_r > 0$ i valori singolari di f e siano $\mathcal{B}_V = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ e $\mathcal{B}_W = \{\mathbf{w_1}, \ldots, \mathbf{w_m}\}$ due basi ortonormali duali associate a tali valori. Allora:*

i) $\mathcal{B}_W$ e $\mathcal{B}_V$ sono basi duali di f^* associate agli stessi valori singolari;

ii) f ed f^* hanno rango r;

iii) $\{\mathbf{w_1}, \ldots \mathbf{w_r}\}$ è una base ortonormale di $\mathrm{Im}(f) = \ker(f^*)^\perp$, mentre $\{\mathbf{v_1}, \ldots \mathbf{v_r}\}$ è una base ortonormale di $\mathrm{Im}(f^*) = \ker(f)^\perp$;

iv) $\{\mathbf{v_{r+1}}, \ldots, \mathbf{v_n}\}$ è una base ortonormale di $\ker(f) = \mathrm{Im}(f^*)^\perp$, mentre $\{\mathbf{w_{r+1}}, \ldots, \mathbf{w_m}\}$ è una base ortonormale di $\ker(f^*) = \mathrm{Im}(f)^\perp$.

Il corollario precedente motiva l'importanza della decomposizione ai valori singolari. Infatti, se uno conosce una decomposizione di un'applicazione allora conosce le informazioni più importanti che caratterizzano l'applicazione stessa.

Esempio 9.49. Nell'Esempio 9.29 abbiamo visto che una rotazione f di un angolo α nello spazio bidimensionale V non è un'applicazione autoaggiunta. La sua matrice rappresentativa rispetto ad una base ortonormale è

$$F|_{\mathcal{B}} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}.$$

Le applicazioni $f^* \circ f$ ed $f \circ f^*$ sono rappresentate da

$$F|_{\mathcal{B}}^{T} * F|_{\mathcal{B}} = F|_{\mathcal{B}} * F|_{\mathcal{B}}^{T} = \mathbb{I}_n,$$

per cui i valori singolari di f sono $\sigma_1 = \sigma_2 = 1$ e qualunque vettore $\mathbf{v} \in V$ è vettore singolare destro e sinistro. Si deduce che la rotazione f ha rango 2, quindi è biettiva.

Così come il Teorema spettrale per le matrici simmetriche, anche la decomposizione ai valori singolari può essere formulata direttamente nel linguaggio delle matrici. Anche in questo caso, ricordiamo preliminarmente i seguenti fatti:

- ogni matrice $A \in \mathrm{Mat}(m, n; \mathbb{R})$ definisce un omomorfismo $f_A \in \mathrm{Hom}(V, W)$ da $V = \mathrm{Mat}(n, 1; \mathbb{R})$ in $W = \mathrm{Mat}(m, 1; \mathbb{R})$;
- la matrice che rappresenta f_A rispetto alle basi canoniche $\mathcal{B}_{n1}$ di V e $\mathcal{B}_{m1}$ di W è A stessa;
- rispetto alle strutture euclidee canoniche su V e W, definite dal prodotto di Frobenius, le basi $\mathcal{B}_{n1}, \mathcal{B}_{m1}$ sono ortonormali.

Chiamiamo allora valori singolari e vettori singolari di A i corrispondenti valori e vettori di f_A.

Proposizione 9.50. *Ogni matrice $A \in \mathrm{Mat}(m, n; \mathbb{R})$ può essere decomposta come il prodotto*

$$A = P * \Sigma * Q^T,$$

dove:

- $P \in O(m; \mathbb{R})$, *le cui colonne costituiscono una base ortonormale di vettori singolari a sinistra di A;*
- $Q \in O(n; \mathbb{R})$, *le cui colonne costituiscono una base ortonormale duale di vettori singolari a destra di A;*
- $\Sigma \in \mathrm{Mat}(m, n; \mathbb{R})$, *i cui unici elementi non nulli sono $(\Sigma)_{ii} = \sigma_i$, i valori singolari di A.*

Dimostrazione. Le matrici P, Q e Σ esistono perché la decomposizione ai valori singolari di f_A esiste per ogni A. Allora vale

$$P^T * A * Q = P^T * \left[\begin{array}{ccc|c} \sigma_1 \cdot P_{C(1)} & \dots & \sigma_r \cdot P_{C(r)} & 0_{m\, n-r} \end{array}\right] = \left[\begin{array}{ccc|c} \sigma_1 & \dots & 0 & \\ \vdots & & \vdots & 0_{r\, n-r} \\ 0 & \dots & \sigma_r & \\ \hline & 0_{m-r\, r} & & 0_{m-r\, n-r} \end{array}\right] = \Sigma.$$

La tesi segue sfruttando l'ortogonalità di P e Q. □

Prima di presentare un esempio finale esplicativo dei precedenti risultati, concludiamo l'analisi teorica studiando alcune interessanti proprietà metriche legate alla decomposizione ai valori singolari, per gli omomorfismi e per le matrici, molto utili dal punto di vista delle applicazioni.

Proposizione 9.51. *Siano $\sigma_1 \geq \dots \geq \sigma_r$ i valori singolari di f. Dato un vettore $\mathbf{v} \in V$, chiamiamo $\mathbf{u}$ la sua proiezione ortogonale su $\ker(f)^{\perp}$. Allora valgono*

$$\sigma_1 \|\mathbf{u}\| \geq \|f(\mathbf{v})\| \geq \sigma_r \|\mathbf{u}\|.$$

La prima oppure la seconda uguaglianza è verificata se e solo se $\mathbf{u}$ è multiplo rispettivamente del vettore singolare destro $\mathbf{v_1}$ oppure del vettore singolare destro $\mathbf{v_r}$.

Dimostrazione. Siano $\mathcal{B}_V = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ e $\mathcal{B}_W = \{\mathbf{w_1}, \dots, \mathbf{w_m}\}$ due basi ortonormali duali associate ai valori singolari. Ogni vettore $\mathbf{v}$ e la sua proiezione $\mathbf{u}$ possono essere decomposti come

$$\mathbf{v} = \sum_{i=1}^{n} t_i \cdot \mathbf{v_i} \qquad \text{e} \qquad \mathbf{u} = \sum_{i=1}^{r} t_i \cdot \mathbf{v_i}.$$

L'immagine di $\mathbf{v}$ è

$$f(\mathbf{v}) = \sum_{i=1}^{n} t_i \cdot f(\mathbf{v_i}) = \sum_{i=1}^{r} t_i \, \sigma_i \cdot \mathbf{w_i}$$

e quindi vale

$$\|f(\mathbf{v})\|^2 = \left\| \sum_{i=1}^{r} t_i \, \sigma_i \cdot \mathbf{w_i} \right\|^2 = \sum_{i=1}^{r} t_i^2 \, \sigma_i^2 \, \|\mathbf{w_i}\|^2 = \sum_{i=1}^{r} t_i^2 \, \sigma_i^2 \, .$$

Allora

$$\|f(\mathbf{v})\|^2 \le \sum_{i=1}^{r} t_i^2 \, \sigma_1^2 = \sigma_1^2 \sum_{i=1}^{r} t_i^2 = \sigma_1^2 \, \|\mathbf{u}\|^2 \, ,$$

da cui otteniamo $\sigma_1 \|\mathbf{u}\| \ge \|f(\mathbf{v})\|$. Se $\mathbf{u} = \mathbf{0}$ l'uguaglianza è soddisfatta. Se $\mathbf{u} \ne \mathbf{0}$, l'uguaglianza è vera se e solo se $t_2 = \dots = t_r = 0$. In entrambi i casi deve valere $\mathbf{u} = t_1 \cdot \mathbf{v_1}$. In maniera analoga si dimostrano le proprietà della minorazione di $\|f(\mathbf{v})\|$. □

Nei prossimi enunciati useremo la notazione $\sigma_i(f)$ per indicare i valori singolari di $f \in \mathrm{Hom}(V, W)$, definendo $\sigma_i(f) = 0$ per $i > r$.

Corollario 9.52. *Siano $f_1, f_2 \in \mathrm{Hom}(V, W)$, allora $\sigma_1(f_1) + \sigma_1(f_2) \ge \sigma_1(f_1 + f_2)$.*

Dimostrazione. Se $f_1 + f_2 = 0_{V,W}$ la disuguaglianza è banalmente soddisfatta. Altrimenti abbiamo $\sigma_1(f_1 + f_2) > 0$. Sia $\mathbf{v}$ un corrispondente vettore singolare a destra. Ricordiamo che la norma di ogni vettore è sempre maggiore od uguale alla norma della sua proiezione ortogonale su un sottospazio. Di conseguenza, grazie alla Proposizione 9.51 ed alla disuguaglianza triangolare, valgono

$$\sigma_1(f_1) + \sigma_1(f_2) \ge \frac{\|f_1(\mathbf{v})\| + \|f_2(\mathbf{v})\|}{\|\mathbf{v}\|} \ge \frac{\|f_1(\mathbf{v}) + f_2(\mathbf{v})\|}{\|\mathbf{v}\|} = \frac{\|(f_1 + f_2)(\mathbf{v})\|}{\|\mathbf{v}\|} = \sigma_1(f_1 + f_2). \qquad \square$$

L'equivalente del corollario precedente per gli autovalori di una applicazione autoaggiunta è noto fin dal diciannovesimo secolo ed è stato generalizzato da Weyl nel 1912. L'omonimo teorema implica la seguente disuguaglianza per i valori singolari di un'applicazione, che noi non dimostreremo:

$$\sigma_i(f_1) + \sigma_j(f_2) \ge \sigma_{i+j-1}(f_1 + f_2). \tag{14}$$

Grazie a questa disuguaglianza possiamo infine dimostrare il Teorema di Eckart-Young, riguardante l'approssimazione di basso rango di una matrice.

Teorema 9.53. *Sia $A \in \mathrm{Mat}(m, n; \mathbb{R})$ con valori singolari $\sigma_1(A) \ge \dots \ge \sigma_r(A)$ ed avente la decomposizione $A = P * \Sigma * Q^T$. Definiamo $\Sigma_k \in \mathrm{Mat}(m, n; \mathbb{R})$ come la matrice i cui unici elementi diversi da zero sono $(\Sigma_k)_{ii} = \sigma_i(A)$ per $1 \le i \le k$, e di conseguenza definiamo $A_k = P * \Sigma_k * Q^T$. Allora A_k è soluzione del problema di minimizzazione della norma $\|A - X\|_F$, per $X \in \mathrm{Mat}(m, n; \mathbb{R})$ e soddisfacente $r(X) = k$.*

Dimostrazione. Per costruzione $r(A_k) = k$. Osserviamo che

$$\|A\|_F^2 = \left\|P * \Sigma * Q^T\right\|_F^2 = \mathrm{Tr}((P * \Sigma * Q^T)^T * (P * \Sigma * Q^T)) =$$

$$= \mathrm{Tr}(Q * \Sigma^T * P^T * P * \Sigma * Q^T) = \mathrm{Tr}(\Sigma^T * \Sigma) = \|\Sigma\|_F^2 = \sum_{i=1}^{n} \sigma_i(A)^2.$$

Analogamente,

$$\|A - A_k\|_F^2 = \left\|P * (\Sigma - \Sigma_k) * Q^T\right\|_F^2 = \|\Sigma - \Sigma_k\|_F^2 = \sum_{i=k+1}^{n} \sigma_i(A)^2.$$

Sia $X \in \mathrm{Mat}(m, n; \mathbb{R})$ di rango k, di conseguenza $\sigma_{k+1}(X) = 0$. Applicando la disuguaglianza (14), abbiamo

$$\sigma_i(A - X) = \sigma_i(A - X) + \sigma_{k+1}(X) \geq \sigma_{i+k}(A).$$

per $1 \leq i \leq n$. Allora segue che

$$\|A - X\|_F^2 = \sum_{i=1}^{n} \sigma_i(A - X)^2 \geq \sum_{i=k+1}^{n} \sigma_i(A)^2 = \|A - A_k\|_F^2. \qquad \square$$

Esempio 9.54. Abbiamo visto nell'Esempio 9.44 che la matrice

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}$$

ha valori singolari $\sigma_1 = \sqrt{3}$ e $\sigma_2 = 1$, per cui la matrice ha rango 2. Le basi ortonormali composte da vettori singolari si ricavano dalla diagonalizzazione di

$$A^T * A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \qquad \text{e} \qquad A * A^T = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & 2 \end{bmatrix}.$$

La base ortonormale di vettori singolari a destra è composta da autovettori di $A^T * A$:

$$\mathbf{v_1} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \mathbf{v_2} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \qquad \Rightarrow \qquad Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

I vettori $\mathbf{v_1}^T, \mathbf{v_2}^T$ formano una base di $R(A)$. La base ortonormale duale si ricava come

$$\mathbf{w_1} = \frac{1}{\sigma_1} \cdot A * \mathbf{v_1} = \begin{bmatrix} 1/\sqrt{6} \\ -1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}, \qquad \mathbf{w_2} = \frac{1}{\sigma_2} \cdot A * \mathbf{v_2} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix},$$

$$\mathbf{w_3} = \mathbf{w_1} \wedge \mathbf{w_2} = \begin{vmatrix} 1/\sqrt{6} & 1/\sqrt{2} & E_{11} \\ -1/\sqrt{6} & 1/\sqrt{2} & E_{21} \\ -2/\sqrt{6} & 0 & E_{31} \end{vmatrix} = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \qquad \Rightarrow \qquad P = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ -1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ -2/\sqrt{6} & 0 & 1/\sqrt{3} \end{bmatrix}.$$

Lasciamo al lettore verificare che $\mathbf{w_1}, \mathbf{w_2}, \mathbf{w_3}$ sono autovettori di $A * A^T$ e che formano una base di $C(A)$. La decomposizione ai valori singolari di A è

$$A = P * \Sigma * Q^T = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ -1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ -2/\sqrt{6} & 0 & 1/\sqrt{3} \end{bmatrix} * \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} * \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

La migliore approssimazione di rango 1 di A è la matrice

$$A_1 = P * \Sigma_1 * Q^T = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ -1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ -2/\sqrt{6} & 0 & 1/\sqrt{3} \end{bmatrix} * \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} * \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \\ -1 & 1 \end{bmatrix}.$$

É facile verificare che $\|A - A_1\|_F = 1 = \sqrt{\sigma_2^2}$.

CAPITOLO 10

Geometria euclidea

In questo capitolo riprendiamo lo studio della geometria, concentrandoci sugli aspetti metrici della stessa. Questo è possibile grazie agli strumenti che abbiamo sviluppato nei Capitoli 8 e 9.

10.1. Spazi euclidei, sistemi di riferimento cartesiani ed isometrie

Definizione 10.1. *Uno spazio euclideo* $\mathbb{E} = (\mathcal{A}, V, \langle\ ,\ \rangle, \psi)$ *è uno spazio affine* $\mathbb{A} = (\mathcal{A}, V, \psi)$ *per il quale* $(V, \langle\ ,\ \rangle)$ *è uno spazio vettoriale euclideo. I sottospazi euclidei* S *di* $\mathbb{E}$ *sono i sottospazi affini di* $\mathbb{A}$ *con struttura euclidea ereditata da* $\mathbb{E}$*. Infine, la dimensione di uno spazio euclideo è uguale alla sua dimensione affine.*

Indicheremo con $\mathbb{E}_V$ lo spazio euclideo associato ad $\mathbb{A}_V = (V, V, \psi)$, in cui la struttura euclidea è quella canonica (ammesso che questa esista). In particolare, lo spazio euclideo canonico n-dimensionale è $\mathbb{E}^n = (\mathbb{R}^n, (\mathbb{R}^n, \mathbb{R}, +, \cdot), \langle\ ,\ \rangle_E, \psi)$, avente quindi lo spazio affine $\mathbb{A}^n_{\mathbb{R}}$ come supporto e struttura euclidea fissata dal prodotto scalare euclideo.

Noi abbiamo intrapreso nel Capitolo 8 lo studio degli spazi vettoriali euclidei con l'obiettivo di definire lunghezze ed angoli. Questo è infatti ciò che distingue la geometria euclidea da quella affine. Iniziamo con lo studio di tali proprietà per i segmenti orientati.

Definizione 10.2. *Siano* P_1, Q_1, P_2, Q_2 *punti di uno spazio euclideo* $\mathbb{E}$*. Allora:*

i) la distanza tra due punti è la funzione

$$\begin{array}{cccc} d: & \mathcal{A} \times \mathcal{A} & \longrightarrow & \mathbb{R} \\ & (P_1, Q_1) & \longmapsto & \left\| \overrightarrow{P_1Q_1} \right\| \end{array}.$$

$d(P_1, Q_1)$ è anche detta lunghezza del segmento orientato (P_1, Q_1);

ii) l'angolo tra i segmenti orientati (P_1, Q_1) e (P_2, Q_2) è

$$\widehat{(P_1, Q_1)(P_2, Q_2)} = \widehat{\overrightarrow{P_1Q_1}\overrightarrow{P_2Q_2}}.$$

Presentiamo alcune proprietà elementari della distanza, la cui dimostrazione è diretta conseguenza delle proprietà della norma.

Proposizione 10.3. *Per ogni punto* P, Q, R *nello spazio euclideo* $\mathbb{E}$*, la distanza soddisfa le proprietà:*

i) di positività, ovvero $d(P, Q) \geq 0$ *ed inoltre* $d(P, Q) = 0$ *se e solo se* $P = Q$*;*

ii) di simmetria, ovvero $d(P, Q) = d(Q, P)$*;*

iii) triangolare, ovvero $d(P, Q) + d(Q, R) \geq d(P, R)$*.*

Grazie alla formula di polarizzazione, la struttura euclidea di $\mathbb{E}$ si può ricavare dalla conoscenza della lunghezza di ogni suo segmento orientato. Osserviamo però che non tutte le funzioni che soddisfano le proprietà della Proposizione 10.3 sono distanze indotte da un prodotto scalare. In generale, un insieme su cui è definita una funzione distanza che soddisfa tali proprietà si dice spazio metrico. Gli spazi euclidei sono esempi particolari di spazi metrici.

La definizione di sistemi di riferimento che abbiamo dato nel Capitolo 7 si specializza in maniera naturale nel caso degli spazi euclidei, in modo da tenere conto della struttura metrica aggiuntiva esistente su questi ultimi.

Definizione 10.4. *Un sistema di riferimento cartesiano di* $\mathbb{E} = (\mathcal{A}, V, \langle\ ,\ \rangle, \psi)$ *è un sistema di riferimento* $\mathcal{B}_O$ *dello spazio affine associato* $\mathbb{A}$*, formato da un'origine* $O \in \mathcal{A}$ *ed una base ortonormale* $\mathcal{B}$ *di* V*. Le coordinate associate a* $\mathcal{B}_O$ *si dicono coordinate cartesiane.*

Parleremo di spazi euclidei orientati qualora lo spazio vettoriale di supporto sia euclideo orientato. In tal caso, distingueremo tra sistemi di riferimento cartesiani orientati positivamente e negativamente, a seconda dell'orientazione delle basi associate. Anche il concetto di angolo orientato si può naturalmente mutuare dalla corrispondente definizione negli spazi vettoriali euclidei.

Esempio 10.5. In $\mathbb{E}^n$ calcoliamo le lunghezze e gli angoli tra i segmenti (O, P_i) per $1 \leq i \leq n$, dove

$$(O)_k = 1 \quad \text{per} \quad 1 \leq k \leq n \qquad \text{e} \qquad (P_i)_k = \begin{cases} 1 & \text{per} \quad 1 \leq k \leq n, \ k \neq i, \\ 0 & \text{per} \qquad\quad k = i. \end{cases}$$

Abbiamo

$$\overrightarrow{OP_i} = P_i - O = -\mathbf{e_i} \qquad \Rightarrow \qquad \left\| \overrightarrow{OP_i} \right\| = 1, \quad \widehat{\overrightarrow{OP}_i\overrightarrow{OP}_j} = \frac{\pi}{2},$$

per $1 \leq i, j \leq n$ ed $i \neq j$. Quindi l'insieme $\mathcal{B}_O = \{O, \overrightarrow{OP_1}, \ldots, \overrightarrow{OP_n}\}$ è un sistema di riferimento cartesiano di $\mathbb{E}^n$. Se scegliamo l'orientazione canonica di $\mathbb{E}_n$, rispetto a cui $\mathcal{B}_n = \{\mathbf{e_1}, \ldots, \mathbf{e_n}\}$ è orientata positivamente, otteniamo che l'orientazione di $\mathcal{B}_O$ dipende dalla dimensione dello spazio. Infatti

$$M_{\mathcal{B}_O \mathcal{B}_n} = \begin{bmatrix} -\mathbf{e_1}|_{\mathcal{B}_n} & \ldots & -\mathbf{e_n}|_{\mathcal{B}_n} \end{bmatrix} = -\mathbb{I}_n \qquad \Rightarrow \qquad \det(M_{\mathcal{B}_O \mathcal{B}_n}) = (-1)^n.$$

Infine, completiamo la sezione studiando le funzioni naturali tra spazi euclidei.

Definizione 10.6. *Siano* $\mathbb{E}_1 = (\mathcal{A}_1, V_1, \langle\ ,\ \rangle_1, \psi_1)$ *ed* $\mathbb{E}_2 = (\mathcal{A}_2, V_2, \langle\ ,\ \rangle_2, \psi_2)$ *due spazi euclidei ed* $f \in \operatorname{Hom}(\mathbb{A}_1, \mathbb{A}_2)$ *una trasformazione affine tra i due rispettivi spazi affini di supporto. Allora* f *si dice isometria se l'applicazione lineare associata* $\bar{f}$ *è un'isometria da* $(V_1, \langle\ ,\ \rangle_1)$ *in* $(V_2, \langle\ ,\ \rangle_2)$.

Esempio 10.7. Naturalmente l'identità e le traslazioni sono isometrie, in quanto sono entrambe associate all'applicazione lineare identica. Sia invece S un sottospazio affine di $\mathbb{E} = (\mathcal{A}, V, \psi)$ avente giacitura U e sia O un punto di S. Per ogni punto $P \in \mathcal{A}$ definiamo la proiezione e la riflessione ortogonale rispettivamente come

$$\mathcal{P}_S(P) = \psi_O^{-1}(\mathcal{P}_U(\psi_O(P))) \qquad \text{e} \qquad \mathcal{R}_S(P) = \psi_O^{-1}(\mathcal{R}_U(\psi_O(P))).$$

Questo significa che $\mathcal{P}_S(P)$ è il punto tale per cui $\overrightarrow{O\mathcal{P}_S(P)} = \mathcal{P}_U(\overrightarrow{OP})$, mentre per $\mathcal{R}_S(P)$ vale $\overrightarrow{O\mathcal{R}_S(P)} = \mathcal{R}_U(\overrightarrow{OP})$. Dobbiamo dimostrare che queste due funzioni non dipendono dalla scelta di O. Essendo verifiche analoghe, sviluppiamo il ragionamento solo per la proiezione ortogonale. Siano $O' \in S$ e $Q' = \psi_{O'}^{-1}(\mathcal{P}_U(\psi_{O'}(P)))$. Sfruttando la regola di Chasles, abbiamo

$$\overrightarrow{O\mathcal{P}_S(P)} = \mathcal{P}_U(\overrightarrow{OP}) = \mathcal{P}_U(\overrightarrow{OO'} + \overrightarrow{O'P}) = \mathcal{P}_U(\overrightarrow{OO'}) + \mathcal{P}_U(\overrightarrow{O'P}) = \overrightarrow{OO'} + \overrightarrow{O'Q'} = \overrightarrow{OQ'},$$

da cui si ricava $Q' = \mathcal{P}_S(P)$. Osserviamo in particolare che $\mathbf{0} = \overrightarrow{\mathcal{P}_S(P)\mathcal{P}_S(P)} = \mathcal{P}_U(\overrightarrow{\mathcal{P}_S(P)P})$, ovvero la proiezione ortogonale di un punto soddisfa la naturale proprietà $\overrightarrow{\mathcal{P}_S(P)P} \in U^\perp$. Infine, abbiamo chiaramente che le applicazione lineari associate a $\mathcal{P}_S$ ed $\mathcal{R}_S$ sono rispettivamente la proiezione ortogonale $\mathcal{P}_U$ e la riflessione ortogonale $\mathcal{R}_U$. Per quanto detto nell'Esempio 9.2, segue che $\mathcal{P}_S$ non è un'isometria mentre $\mathcal{R}_S$ lo è.

Ad esempio, nello spazio euclideo $\mathbb{E}^3$ consideriamo il piano

$$S = \{(x, y, z) \in \mathbb{A}^3_{\mathbb{R}} \,|\, x + y + z = 1\}.$$

Una base ortonormale della giacitura U di S è

$$\mathcal{B}_U = \left\{\mathbf{u_1} = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right), \mathbf{u_2} = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)\right\}.$$

Vogliamo calcolare la proiezione e la riflessione ortogonale del punto $P = (1, 1, 1)$ rispetto ad S. Il punto $O = (1, 0, 0)$ appartenente ad S definisce il vettore $\overrightarrow{OP} = (0, 1, 1)$, da cui ricaviamo

$$\overrightarrow{O\mathcal{P}_S(P)} = \mathcal{P}_U(\overrightarrow{OP}) = \left\langle \overrightarrow{OP}, \mathbf{u_1} \right\rangle \cdot \mathbf{u_1} + \left\langle \overrightarrow{OP}, \mathbf{u_2} \right\rangle \cdot \mathbf{u_2} = \left(-\frac{1}{2}, \frac{1}{2}, 0\right) + \left(-\frac{1}{6}, -\frac{1}{6}, \frac{2}{6}\right) = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

e quindi $\mathcal{P}_S(P) = (1/3, 1/3, 1/3)$. La riflessione si ricava invece da

$$\overrightarrow{O\mathcal{R}_S(P)} = \mathcal{R}_U(\overrightarrow{OP}) = \overrightarrow{OP} + 2 \cdot \overrightarrow{P\mathcal{P}_S(P)} = (0, 1, 1) + 2 \cdot \left(\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) - (1, 1, 1)\right) = -\left(\frac{4}{3}, \frac{1}{3}, \frac{1}{3}\right),$$

da cui $\mathcal{R}_S(P) = -(1/3, 1/3, 1/3)$.

Chiaramente, se due spazi euclidei hanno la stessa dimensione, allora f è un'isometria se e solo se l'immagine di un sistema di riferimento cartesiano è ancora un sistema cartesiano. Naturalmente in tal caso le isometrie si possono caratterizzare geometricamente come le affinità che lasciano inalterate gli angoli e le distanze tra gli oggetti geometrici. Strettamente legato al concetto di isometria vi è quello di congruenza.

Definizione 10.8. *Siano $\mathbb{E}_1$ ed $\mathbb{E}_2$ due spazi euclidei. Due loro sottoinsiemi A_1 ed A_2 si dicono congruenti se esiste un isometria tale che $f(A_1) = A_2$.*

Lasciamo al lettore verificare che la congruenza è una relazione di equivalenza. Per definizione, due insiemi sono congruenti se e solo se i punti che li costituiscono hanno uguali angoli e distanze. Questo significa che negli spazi euclidei da noi definiti è valido il (iv) postulato di Euclide, ovvero due angoli retti sono sempre congruenti.

10.2. Distanza di un punto da un sottospazio

In uno spazio metrico, possiamo dare la definizione generale di distanza tra un punto ed un sottoinsieme di punti.

Definizione 10.9. *Sia S un sottoinsieme non vuoto di uno spazio metrico e P un punto dello spazio. Allora*

$$d(S, P) = \inf_{Q \in S} d(Q, P).$$

La distanza tra due punti è per definizione una funzione a valori reali inferiormente limitata. In generale, non è detto che esista un minimo di $d(Q, P)$ per $Q \in S$, ma sicuramente esiste il suo estremo inferiore e quindi $d(S, T)$ è ben definita. In questa sezione studieremo tale distanza nel caso in cui S sia un sottospazio affine di uno spazio euclideo $\mathbb{E}$. Per prima cosa vedremo che in questa situazione particolare esiste esattamente un punto $Q \in S$ tale che $d(S, T) = d(P, Q)$, grazie al quale potremo calcolare tale quantità.

Proposizione 10.10. *Siano $\mathbb{E}$ uno spazio euclideo, P un punto ed S un sottospazio di dimensione finita. Allora:*

i) $d(S, P) = \left\| \overrightarrow{\mathcal{P}_S(P)P} \right\|$;

ii) se $Q \in S$ e $\mathcal{B} = \{\mathbf{u_1}, \dots, \mathbf{u_n}\}$ è una base della giacitura U di S, vale

$$d(S,P) = \sqrt{\frac{G\left(\mathbf{u_1}, \dots, \mathbf{u_n}, \overrightarrow{QP}\right)}{G\left(\mathbf{u_1}, \dots, \mathbf{u_n}\right)}}.$$

Dimostrazione. Sia O un punto fissato in S. Allora

$$d(S,P) = \inf_{Q \in S} d(Q,P) = \inf_{Q \in S} \left\|\overrightarrow{QP}\right\| = \inf_{Q \in S} \left\|\overrightarrow{OP} - \overrightarrow{OQ}\right\|.$$

Al variare di Q in S, i vettori geometrici $\overrightarrow{OQ}$ descrivono tutti i vettori della giacitura U di S. Quindi

$$d(S,P) = \inf_{\mathbf{u} \in U} \left\|\overrightarrow{OP} - \mathbf{u}\right\|.$$

Dalla Proposizione 8.24 sappiamo che

$$\operatorname*{argmin}_{\mathbf{u} \in U} \left\|\overrightarrow{OP} - \mathbf{u}\right\| = \mathcal{P}_U(\overrightarrow{OP})$$

e quindi

$$d(S,P) = \left\|\overrightarrow{OP} - \mathcal{P}_U(\overrightarrow{OP})\right\| = \left\|\overrightarrow{OP} - \overrightarrow{O\mathcal{P}_S(P)}\right\| = \left\|\overrightarrow{\mathcal{P}_S(P)P}\right\|.$$

Per il secondo punto dell'enunciato, osserviamo che se $P \in S$ allora $\overrightarrow{QP} \in U$, di conseguenza $G\left(\mathbf{u_1}, \dots, \mathbf{u_n}, \overrightarrow{QP}\right) = 0$ e quindi correttamente otteniamo $d(S,P) = 0$. Altrimenti, definiamo lo spazio vettoriale $W = U + \mathcal{L}(\overrightarrow{QP})$, che ha dimensione $n+1$. Il vettore geometrico $\overrightarrow{\mathcal{P}_S(P)P}$ appartiene a W, poiché

$$\overrightarrow{\mathcal{P}_S(P)P} = \overrightarrow{\mathcal{P}_S(P)Q} + \overrightarrow{QP}.$$

Inoltre, per costruzione $\overrightarrow{\mathcal{P}_S(P)P}$ appartiene al complemento ortogonale $U^\perp$ di U in W. Dato che anche $\mathbf{u} = *(\mathbf{u_1}, \dots, \mathbf{u_n}) \in U^\perp$, deve valere $\overrightarrow{\mathcal{P}_S(P)P} = t \cdot \mathbf{u}$ e quindi

$$d(S,P) = |t|\,\|\mathbf{u}\| = \frac{\left|\left\langle \overrightarrow{\mathcal{P}_S(P)P}, \mathbf{u}\right\rangle\right|}{\|\mathbf{u}\|} = \frac{\left|\left\langle \overrightarrow{\mathcal{P}_S(P)Q} + \overrightarrow{QP}, \mathbf{u}\right\rangle\right|}{\|\mathbf{u}\|} = \frac{\left|\left\langle \overrightarrow{QP}, \mathbf{u}\right\rangle\right|}{\|\mathbf{u}\|}.$$

Per definizione $\|\mathbf{u}\| = \sqrt{G(\mathbf{u_1}, \dots, \mathbf{u_n})}$, mentre grazie all'Osservazione 8.49 vale

$$\left|\left\langle \overrightarrow{QP}, \mathbf{u}\right\rangle\right| = \sqrt{G\left(\mathbf{u_1}, \dots, \mathbf{u_n}, \overrightarrow{QP}\right)},$$

da cui segue la tesi. □

Esempio 10.11. In $\mathbb{E}^3$ calcoliamo la distanza tra il punto $P = (1,1,1)$ ed il piano

$$S = \{(x,y,z) \in \mathbb{A}^3_{\mathbb{R}} \mid x+y+z = 1\}.$$

Nell'Esempio 10.7 abbiamo calcolato la proiezione $\mathcal{P}_S(P) = (1/3, 1/3, 1/3)$, da cui segue

$$d(S,P) = \left\|\overrightarrow{\mathcal{P}_S(P)P}\right\| = \left\|\left(\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right)\right\| = \frac{2}{\sqrt{3}}.$$

Alternativamente, sappiamo anche che una base ortonormale della giacitura U di S è

$$\mathcal{B}_U = \left\{\mathbf{u_1} = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right), \mathbf{u_2} = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)\right\},$$

a cui è ovviamente associato il gramiano $G(\mathbf{u_1}, \mathbf{u_2}) = 1$. Il punto $O = (1,0,0)$ appartiene ad S, da cui $\overrightarrow{OP} = (0,1,1)$. Pertanto vale

$$G\left(\mathbf{u_1}, \mathbf{u_2}, \overrightarrow{OP}\right) = \begin{vmatrix} 1 & 0 & -1/\sqrt{2} \\ 0 & 1 & -1/\sqrt{6} \\ -1/\sqrt{2} & -1/\sqrt{6} & 2 \end{vmatrix} = \frac{4}{3}$$

e quindi

$$d(S,P)=\sqrt{\frac{G\left(\mathbf{u_1},\mathbf{u_2},\overrightarrow{OP}\right)}{G(\mathbf{u_1},\mathbf{u_2})}}=\frac{2}{\sqrt{3}}.$$

Ogni sottoinsieme S di $\mathbb{E}$ può essere visto come il luogo dei punti P dello spazio soddisfacenti $d(S,P)=0$. Questo ci permette di ricavare una formula esplicita per la rappresentazione algebrica di un iperpiano.

Corollario 10.12. *Siano $\mathbb{E}$ uno spazio euclideo di dimensione n e Π un iperpiano. Sia inoltre $\mathcal{B}_O$ un sistema di riferimento cartesiano con $O\in\Pi$. Allora:*

i) data la giacitura $U=\mathcal{L}(\mathbf{u_1},\ldots,\mathbf{u_{n-1}})$ dell'iperpiano ed il punto P dello spazio, una rappresentazione algebrica di Π è

$$\det\left(\begin{bmatrix}\mathbf{u_1}|_{\mathcal{B}} & \ldots & \mathbf{u_{n-1}}|_{\mathcal{B}} & P|_{\mathcal{B}_O}\end{bmatrix}\right)=0;$$

ii) se Π ha rappresentazione algebrica

$$\sum_{i=1}^{n}a_ix_i=0,$$

il vettore di componenti $a_1\ldots,a_n$ è ortogonale all'iperpiano ed inoltre la distanza di Π dal punto P di coordinate $\bar{x}_1,\ldots,\bar{x}_n$ è

$$d(\Pi,P)=\frac{|\sum_{i=1}^{n}a_i\bar{x}_i|}{\sqrt{\sum_{i=1}^{n}a_i^2}}.$$

Dimostrazione. I punti dell'iperpiano sono tutti ed i soli punti P tali che $d(\Pi,P)=0$. Questa condizione è verificata se e solo se $G\left(\mathbf{u_1},\ldots,\mathbf{u_n},\overrightarrow{OP}\right)=0$. Dalla dimostrazione del Lemma 8.44 sappiamo che a sua volta questa equazione è equivalente a

$$\det\left(\begin{bmatrix}\mathbf{u_1}|_{\mathcal{B}} & \ldots & \mathbf{u_{n-1}}|_{\mathcal{B}} & P|_{\mathcal{B}_O}\end{bmatrix}\right)=0.$$

L'ultima espressione è un'equazione lineare nelle coordinate di P e quindi è una rappresentazione algebrica di Π :

$$\sum_{i=1}^{n}C_{in}x_i=0,$$

dove C_{in} sono i complementi algebrici rispetto all'ultima colonna della precedente matrice.

Di conseguenza, tutte le rappresentazioni algebriche del piano Π sono del tipo

$$\alpha\left(\sum_{i=1}^{n}C_{in}\,x_i\right)=\sum_{i=1}^{n}\alpha\,C_{in}\,x_i=0,$$

con $\alpha\in\mathbb{R}^*$. Per la Proposizione 8.47, i complementi C_{in} sono le componenti rispetto a $\mathcal{B}$ del vettore $*(\mathbf{u_1},\ldots,\mathbf{u_{n-1}})$, che per definizione è ortogonale ad U. Infine, per quanto visto nelle dimostrazioni delle Proposizioni 8.47 e 10.10, vale

$$\frac{\left|\sum_{i=1}^{n}\alpha\,C_{in}\,\bar{x}_i\right|}{\sqrt{\sum_{i=1}^{n}\alpha^2\,C_{in}^2}}=\frac{\left|\sum_{i=1}^{n}C_{in}\,\bar{x}_i\right|}{\sqrt{\sum_{i=1}^{n}C_{in}^2}}=\frac{\left|\left\langle *(\mathbf{u_1},\ldots,\mathbf{u_{n-1}}),\overrightarrow{OP}\right\rangle\right|}{\|*(\mathbf{u_1},\ldots,\mathbf{u_{n-1}})\|}=\sqrt{\frac{G\left(\mathbf{u_1},\ldots,\mathbf{u_{n-1}},\overrightarrow{OP}\right)}{G\left(\mathbf{u_1},\ldots,\mathbf{u_{n-1}}\right)}}=d(\Pi,P).\square$$

Esempio 10.13. In $\mathbb{E}^4$ sia $\mathcal{B} = \{\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}, \mathbf{e_4}\}$ la base canonica di $\mathbb{R}^4$, che per definizione è ortonormale. Ogni iperpiano Π è individuato da 4 punti. Siano $P_1 = (1, 0, 0, 0)$, $P_2 = (1, 1, 0, 0)$, $P_3 = (1, 1, 1, 0)$ e $P_4 = (1, 1, 1, 1)$. I vettori geometrici $\overrightarrow{P_1P_i}$ per $2 \leq i \leq 4$ costituiscono una base della giacitura U di Π. Dato $P = (x, y, z, w)$ un generico punto di $\mathbb{E}^4$, una rappresentazione algebrica di Π è

$$\det\left(\left[\ \overrightarrow{P_1P_2}|_{\mathcal{B}} \quad \overrightarrow{P_1P_3}|_{\mathcal{B}} \quad \overrightarrow{P_1P_4}|_{\mathcal{B}} \quad P|_{\mathcal{B}_{P_1}}\ \right]\right) = \begin{vmatrix} 0 & 0 & 0 & x-1 \\ 1 & 1 & 1 & y \\ 0 & 1 & 1 & z \\ 0 & 0 & 1 & w \end{vmatrix} = -x + 1 = 0.$$

La distanza di P dal piano è

$$d(\Pi, P) = \frac{|-x+1|}{\sqrt{(-1)^2}} = |x - 1|.$$

Osserviamo che benché ogni sottospazio affine S possa essere descritto come il luogo dei punti P tali che $d(S, P) = 0$, questo non ci consente in generale di scrivere una rappresentazione cartesiana di S. Ad esempio, consideriamo l'asse coordinato delle x in $\mathbb{E}^3$. Esso contiene $O = (0, 0, 0)$ ed ha giacitura $U = \mathcal{L}(\mathbf{e_1})$. Se $P = (x, y, z)$, l'asse è il luogo delle soluzioni di

$$G(\mathbf{e_1}, \overrightarrow{OP}) = 0 \qquad \Rightarrow \qquad \begin{vmatrix} 1 & x \\ x & x^2+y^2+z^2 \end{vmatrix} = y^2 + z^2 = 0.$$

D'altra parte, la rappresentazione cartesiana dell'asse coordinato è $y = z = 0$, ovvero un sistema composto da due equazioni lineari. La precedente equazione di secondo grado rappresenta invece più correttamente una superficie quadrica degenere. Ritorneremo più approfonditamente su tale questione nel Capitolo 11.

Chiudiamo questa sezione con una discussione informale relativa alla misura dei volumi negli spazi euclidei. In $\mathbb{E}^2$ un parallelogramma è la regione geometrica avente come vertici tre punti indipendenti P_1, P_2, P_3 ed il punto P_4 associato a $\overrightarrow{P_1P_4} = \overrightarrow{P_1P_2} + \overrightarrow{P_1P_3}$. Dato il sottospazio $\mathcal{L}(P_1, P_2)$ coincidente con la retta contenente il segmento (P_1, P_2), allora l'area del parallelogrammo è per definizione la lunghezza della base (P_1, P_2) per la relativa altezza, ovvero

$$A(P_1, P_2, P_3) = d(P_1, P_2)\, d(\mathcal{L}(P_1, P_2), P_3) = \sqrt{G\left(\overrightarrow{P_1P_2}\right)}\sqrt{\frac{G\left(\overrightarrow{P_1P_2}, \overrightarrow{P_1P_3}\right)}{G\left(\overrightarrow{P_1P_2}\right)}} = \sqrt{G\left(\overrightarrow{P_1P_2}, \overrightarrow{P_1P_3}\right)}.$$

In $\mathbb{E}^3$ un parallelepipedo è la regione geometrica avente come vertici quattro punti indipendenti P_1, P_2, P_3, P_4 ed i punti P_5, P_6, P_7, P_8 associati a

$$\overrightarrow{P_1P_5} = \overrightarrow{P_1P_2} + \overrightarrow{P_1P_3}, \qquad \overrightarrow{P_1P_6} = \overrightarrow{P_1P_2} + \overrightarrow{P_1P_4}, \qquad \overrightarrow{P_1P_7} = \overrightarrow{P_1P_3} + \overrightarrow{P_1P_4},$$
$$\overrightarrow{P_1P_8} = \overrightarrow{P_1P_2} + \overrightarrow{P_1P_3} + \overrightarrow{P_1P_4}.$$

Il suo volume è per definizione l'area della base, costituita dal parallelogrammo di vertici P_1, P_2, P_3, P_5, per la relativa altezza, ovvero

$$\begin{aligned} V(P_1, P_2, P_3, P_4) &= A(P_1, P_2, P_3)\, d(\mathcal{L}(P_1, P_2, P_3), P_4) = \\ &= \sqrt{G\left(\overrightarrow{P_1P_2}, \overrightarrow{P_1P_3}\right)}\sqrt{\frac{G\left(\overrightarrow{P_1P_2}, \overrightarrow{P_1P_3}, \overrightarrow{P_1P_4}\right)}{G\left(\overrightarrow{P_1P_2}, \overrightarrow{P_1P_3}\right)}} = \\ &= \sqrt{G\left(\overrightarrow{P_1P_2}, \overrightarrow{P_1P_3}, \overrightarrow{P_1P_4}\right)}. \end{aligned}$$

Iterando questo ragionamento, si può definire il volume del *parallelotopo* in $\mathbb{E}^n$ costruito sui vertici $P_1, \dots, P_{n+1}$:

$$V(P_1, \dots, P_{n+1}) = \sqrt{G\left(\overrightarrow{P_1P_2}, \dots, \overrightarrow{P_1P_{n+1}}\right)} = \left| \det \left(\left[\begin{array}{ccc} \overrightarrow{P_1P_2}|_{\mathcal{B}} & \dots & \overrightarrow{P_1P_{n+1}}|_{\mathcal{B}} \end{array} \right] \right) \right|,$$

dove nell'ultima espressione $\mathcal{B}$ è una base ortonormale dello spazio vettoriale $\mathbb{R}^n$ a supporto di $\mathbb{E}^n$. Ad esempio, in $\mathbb{E}^4$ scegliamo i punti $P_1 = (1,0,0,0)$, $P_2 = (1,1,0,0)$, $P_3 = (1,1,1,0)$, $P_4 = (1,1,1,1)$ e $P = (x,y,z,w)$ dell'Esempio 10.13. Allora

$$V(P_1,P_2,P_3,P_4,P) = \left| \det \left(\left[\begin{array}{cccc} \overrightarrow{P_1P_2}|_{\mathcal{B}} & \overrightarrow{P_1P_3}|_{\mathcal{B}} & \overrightarrow{P_1P_4}|_{\mathcal{B}} & \overrightarrow{P_1P}|_{\mathcal{B}} \end{array} \right] \right) \right| = |x-1|.$$

10.3. Distanza tra sottospazi

La distanza tra due generici sottoinsiemi di punti in uno spazio metrico si definisce generalizzando in modo ovvio la Definizione 10.9.

Definizione 10.14. *Siano S, T sottoinsiemi non vuoti di uno spazio metrico. La distanza tra i due sottoinsiemi è*

$$d(S,T) = \inf_{(P,Q) \in S \times T} d(P,Q).$$

Ancora una volta ci focalizzeremo sullo studio della distanza tra due sottospazi affini S e T di uno spazio euclideo $\mathbb{E}$. Vedremo che il modo di calcolare tale quantità dipende dalla mutua posizione dei sottospazi. A tale scopo abbiamo bisogno del seguente risultato preliminare.

Lemma 10.15. *Siano $\mathbb{E}$ uno spazio euclideo ed S, T due sottospazi sghembi di dimensione finita. Allora esiste un unico sottospazio affine T' di dimensione minima, contenente T e parallelo ad S.*

Dimostrazione. Chiamiamo U e W rispettivamente le giaciture di S e T. Definiamo T' il sottospazio affine contenente un generico punto P di T e di giacitura $U+W$. Per costruzione $U+W$ è il più piccolo sottospazio vettoriale contenente U e W. Dalla Proposizione 7.22 segue che $T \subseteq T'$ mentre per definizione S è parallelo a T'. Ovviamente qualsiasi altro sottospazio affine che soddisfa tali proprietà deve contenere T', da cui seguono le proprietà di dimensione minima e di unicità di T'. □

Proposizione 10.16. *Siano $\mathbb{E}$ uno spazio euclideo ed S, T due sottospazi di dimensione finita e giacitura U, W. Se i due sottospazi sono:*

i) incidenti vale $d(S,T) = 0$;

ii) paralleli con $W \subseteq U$, vale $d(S,T) = d(S,P)$ dove P è un generico punto di T;

iii) sghembi vale $d(S,T) = d(S,T')$, dove T' è il più piccolo sottospazio affine contenente T e parallelo ad S.

Dimostrazione. Il primo caso è ovvio. Nel caso di sottospazi paralleli, iniziamo osservando che

$$d(S,T) = \inf_{(P,Q) \in S \times T} d(P,Q) = \inf_{Q \in T} d(S,Q).$$

Dobbiamo quindi dimostrare che $d(S,Q_1) = d(S,Q_2)$ per ogni coppia di punti $Q_1, Q_2 \in T$. Grazie alla Proposizione 10.10 ed alla regola di Chasles, abbiamo

$$d(S,Q_1) = \left\| \overrightarrow{\mathcal{P}_S(Q_1)Q_1} \right\| = \left\| \overrightarrow{\mathcal{P}_S(Q_1)\mathcal{P}_S(Q_2)} + \overrightarrow{\mathcal{P}_S(Q_2)Q_2} + \overrightarrow{Q_2Q_1} \right\|.$$

Dato che $\mathcal{P}_S(Q_1) \in S$, per definizione di proiezione ortogonale vale

$$\overrightarrow{\mathcal{P}_S(Q_1)\mathcal{P}_S(Q_2)} = \mathcal{P}_U(\overrightarrow{\mathcal{P}_S(Q_1)Q_2}).$$

D'altra parte, il vettore $\overrightarrow{Q_2Q_1}$ appartiene a $W \subseteq U$ e quindi

$$\overrightarrow{Q_2Q_1} = \mathcal{P}_U(\overrightarrow{Q_2Q_1}).$$

Di conseguenza abbiamo

$$\overrightarrow{\mathcal{P}_S(Q_1)\mathcal{P}_S(Q_2)} + \overrightarrow{Q_2Q_1} = \mathcal{P}_U(\overrightarrow{\mathcal{P}_S(Q_1)Q_2}) + \mathcal{P}_U(\overrightarrow{Q_2Q_1}) = \mathcal{P}_U(\overrightarrow{\mathcal{P}_S(Q_1)Q_2} + \overrightarrow{Q_2Q_1}) = \mathcal{P}_U(\overrightarrow{\mathcal{P}_S(Q_1)Q_1}) = \mathbf{0},$$

da cui si ricava

$$d(S, Q_1) = \left\| \overrightarrow{\mathcal{P}_S(Q_2)Q_2} \right\| = d(S, Q_2).$$

Infine, consideriamo lo scenario con sottospazi sghembi. Grazie al Lemma 10.15 sappiamo che T' esiste ed ha giacitura $U + W$. Poiché $T \subseteq T'$, vale

$$d(S, T') = \inf_{(P,Q) \in S \times T'} d(P, Q) \leq \inf_{(P,Q) \in S \times T} d(P, Q) = d(S, T).$$

Dal punto precedente, sappiamo inoltre che per ogni $P \in S$ vale $d(S, T') = \left\| \overrightarrow{\mathcal{P}_{T'}(P)P} \right\|$. L'uguaglianza $d(S, T') = d(S, T)$ è vera se tra i punti di S ne esiste almeno uno per cui $\mathcal{P}_{T'}(P) \in T$. Per dimostrarne l'esistenza, consideriamo la proiezione ortogonale $\mathcal{P}_{T'}(S)$ del sottospazio S su T'. $\mathcal{P}_{T'}(S)$ è un sottoinsieme affine di T' con giacitura $\mathcal{P}_{U+W}(U) = U$. Dunque i due sottospazi $\mathcal{P}_{T'}(S)$ e T dello spazio affine T' soddisfano i requisiti del Corollario 7.28, pertanto la loro intersezione non è vuota. Scegliendo $P \in \mathcal{P}_{T'}^{-1}(\mathcal{P}_{T'}(S) \cap T)$ completiamo la dimostrazione. □

Grazie alla proposizione precedente possiamo finalmente dare la formula generale per il calcolo della distanza tra due sottospazi affini.

Teorema 10.17. *Siano $\mathbb{E}$ uno spazio euclideo ed S, T due sottospazi affini di dimensione finita e giacitura U, W. Dati $(P, Q) \in S \times T$ e $\mathcal{B} = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ una base di $U + W$, vale*

$$d(S, T) = \sqrt{\frac{G\left(\mathbf{v_1}, \dots, \mathbf{v_n}, \overrightarrow{PQ}\right)}{G\left(\mathbf{v_1}, \dots, \mathbf{v_n}\right)}}.$$

Dimostrazione. Supponiamo che S e T siano incidenti e sia $O \in S \cap T$. Allora vale

$$\overrightarrow{PQ} = \overrightarrow{PO} + \overrightarrow{OQ} \in U + W,$$

di conseguenza $G\left(\mathbf{v_1}, \dots, \mathbf{v_n}, \overrightarrow{PQ}\right) = 0$ e quindi otteniamo correttamente $d(S, T) = 0$.

Supponiamo invece che S e T siano paralleli con $W \subseteq U$. Di conseguenza $U + W = U$ e quindi $\mathcal{B}$ è una base della giacitura U di S. Grazie alla Proposizione 10.16 sappiamo che $d(S, T) = d(S, Q)$, che a sua volta è uguale a

$$d(S, Q) = \sqrt{\frac{G\left(\mathbf{v_1}, \dots, \mathbf{v_n}, \overrightarrow{PQ}\right)}{G\left(\mathbf{v_1}, \dots, \mathbf{v_n}\right)}},$$

per la Proposizione 10.10.

Ancora grazie alla Proposizione 10.16, nell'ultimo caso di sottospazi sghembi vale $d(S, T) = d(S, T')$, dove T' contiene T, è parallelo ad S ed ha giacitura $U + W$. Dato che $P \in T'$, in particolare vale $d(S, T) = d(P, T')$. Osserviamo inoltre che $\mathcal{B}$ è una base della giacitura di T', il punto Q appartiene a T' ed infine $U \subseteq U + W$. Applicando di nuovo la Proposizione 10.10 e sfruttando la proprietà $G\left(\mathbf{v_1}, \dots, \mathbf{v_n}, \overrightarrow{QP}\right) = G\left(\mathbf{v_1}, \dots, \mathbf{v_n}, \overrightarrow{PQ}\right)$, segue la tesi. □

Esempio 10.18. In $\mathbb{E}^3$, sia S l'insieme i cui punti $P=(x,y,z)$ soddisfano le equazioni

$$\begin{cases} x+ky-1=0 \\ z-k(k-1)=0 \end{cases},$$

che dipendono dal parametro reale k. Sia invece T l'asse coordinato delle y, ovvero il sottospazio

$$\begin{cases} x=0 \\ z=0 \end{cases}.$$

S e T sono entrambi delle rette, hanno quindi dimensione $p=q=1$ in uno spazio di dimensione $n=3$. Le loro giaciture sono rispettivamente

$$U=\mathcal{L}(\,(k,-1,0)\,) \qquad \text{e} \qquad W=\mathcal{L}(\,(0,1,0)\,).$$

Studiamo la loro mutua posizione e la reciproca distanza. Il sistema lineare che descrive la loro intersezione è

$$[\tilde{A}|\tilde{B}]=\left[\begin{array}{ccc|c} 1 & k & 0 & 1 \\ 0 & 0 & 1 & k(k-1) \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & k & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & k(k-1) \end{array}\right].$$

Applichiamo le Proposizioni 7.26 e 10.16.

- Se $k=0$ abbiamo $r([\tilde{A}|\tilde{B}])=3$ mentre $r([\tilde{A}])=2=n-q$, quindi le rette sono parallele non coincidenti. Entrambe sono contenute nel piano $z=0$. Il punto $Q=(1,0,0)$ appartiene ad S, mentre l'origine O appartiene a T, quindi

  $$d(S,T)=d(Q,T)=\left\|\overrightarrow{\mathcal{P}_T(Q)Q}\right\|.$$

  La proiezione di Q su T si ricava da

  $$\overrightarrow{O\mathcal{P}_T(Q)}=\mathcal{P}_W(\overrightarrow{OQ})=\langle(1,0,0),\mathbf{e_2}\rangle\cdot\mathbf{e_2}=\mathbf{0},$$

  da cui $\mathcal{P}_T(Q)=O$. In conclusione

  $$d(S,T)=\left\|\overrightarrow{OQ}\right\|=\|(1,0,0)\|=1.$$

- Se $k=1$ abbiamo $r([\tilde{A}|\tilde{B}])=r([\tilde{A}])=3>n-q$, quindi le rette sono incidenti. In questo caso S e T appartengono al medesimo piano $z=0$ e l'unico punto in $S\cap T$ é $P=(0,1,0)$. Ovviamente vale $d(S,T)=0$.
- Se $k\neq 0,1$ abbiamo infine $r([\tilde{A}|\tilde{B}])=4$ mentre $r([\tilde{A}])=3>n-q$, quindi le rette sono sghembe. Il piano T' contenente T e parallelo ad S è descritto da:

  $$(x,y,z)=(0,0,0)+t_1\cdot(0,1,0)+t_2\cdot(k,-1,0),$$

  ovvero $z=0$. Il punto $Q=(1,0,k(k-1))$ appartiene ad S, mentre l'origine O appartiene a T'. Una base ortonormale di $U+W$ è $\mathcal{B}_{U+W}=\{\mathbf{e_1},\mathbf{e_2}\}$. La proiezione di Q su T' si ricava da

  $$\overrightarrow{O\mathcal{P}_{T'}(Q)}=\mathcal{P}_{U+W}(\overrightarrow{OQ})=\langle(1,0,k(k-1)),\mathbf{e_1}\rangle\cdot\mathbf{e_1}+\langle(1,0,k(k-1)),\mathbf{e_2}\rangle\cdot\mathbf{e_2}=\mathbf{e_1}$$

  e quindi $\mathcal{P}_{T'}(Q)=(1,0,0)$. In conclusione

  $$d(S,T)=d(S,T')=\left\|\overrightarrow{\mathcal{P}_{T'}(Q)Q}\right\|=\|(0,0,k(k-1))\|=|k(k-1)|.$$

Verifichiamo i risultati precedenti utilizzando il Teorema 10.17. Abbiamo $Q=(1,0,k(k-1))\in S$ ed $O=(0,0,0)\in T$ per ogni $k\in\mathbb{R}$, quindi $\overrightarrow{OQ}=(1,0,k(k-1))$.

- Se $k = 0$, una base di $U + W$ è $\mathcal{B}_{U+W} = \{\mathbf{e_2}\}$, quindi

$$G\left(\mathbf{e_2}, \overrightarrow{OQ}\right) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

  da cui

$$d(S,T) = \sqrt{\frac{G\left(\mathbf{e_2}, \overrightarrow{OQ}\right)}{G\left(\mathbf{e_2}\right)}} = 1.$$

- Se $k \neq 0$, una base di $U + W$ è $\mathcal{B}_{U+W} = \{\mathbf{e_1}, \mathbf{e_2}\}$, quindi

$$G\left(\mathbf{e_1}, \mathbf{e_2}, \overrightarrow{OQ}\right) = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 + (k(k-1))^2 \end{vmatrix} = (k(k-1))^2$$

  da cui

$$d(S,T) = \sqrt{\frac{G\left(\mathbf{e_1}, \mathbf{e_2}, \overrightarrow{OQ}\right)}{G\left(\mathbf{e_1}, \mathbf{e_2}\right)}} = |k(k-1)|.$$

10.4. Angoli tra sottospazi

Per definire l'angolo tra sottospazi affini è prima necessario introdurre la nozione di angolo tra due sottospazi vettoriali. In quanto segue U e W saranno due sottospazi di uno spazio vettoriale euclideo V, rispettivamente di dimensione $\dim(U) = m$ e $\dim(W) = n$ con $1 \leq m, n < \infty$. Sarà inoltre importante distinguere con cura tra le seguenti funzioni:

- la proiezione ortogonale $\mathcal{P}_W : V \to W$;
- la sua restrizione $\mathcal{P}_{UW} : U \to W$;
- la sua composizione con l'immersione, ovvero $\widetilde{\mathcal{P}}_W = i_{W,V} \circ \mathcal{P}_W : V \to V$.

Esiste una definizione naturale di angolo nel caso in cui uno dei due sottospazi sia monodimensionale.

Definizione 10.19. *Se $m = 1$, sia $\mathcal{B}_U = \{\mathbf{u}\}$ è una base di U. Allora l'angolo tra i due sottospazi è*

$$\widehat{UW} = \begin{cases} \widehat{\mathbf{u}\,\widetilde{\mathcal{P}}_W(\mathbf{u})} & \textit{se} \quad \widetilde{\mathcal{P}}_W(\mathbf{u}) \neq \mathbf{0}, \\ \dfrac{\pi}{2} & \textit{se} \quad \widetilde{\mathcal{P}}_W(\mathbf{u}) = \mathbf{0}. \end{cases}$$

Osserviamo che $\widehat{UW} \in [0, \frac{\pi}{2}]$. Per dimostrarlo sfruttiamo le proprietà di idempotenza ed autoaggiuntezza della proiezione, grazie a cui otteniamo

$$\frac{\left\langle \mathbf{u}, \widetilde{\mathcal{P}}_W(\mathbf{u}) \right\rangle}{\|\mathbf{u}\| \left\|\widetilde{\mathcal{P}}_W(\mathbf{u})\right\|} = \frac{\left\langle \mathbf{u}, \widetilde{\mathcal{P}}_W^2(\mathbf{u}) \right\rangle}{\|\mathbf{u}\| \left\|\widetilde{\mathcal{P}}_W(\mathbf{u})\right\|} = \frac{\left\langle \widetilde{\mathcal{P}}_W(\mathbf{u}), \widetilde{\mathcal{P}}_W(\mathbf{u}) \right\rangle}{\|\mathbf{u}\| \left\|\widetilde{\mathcal{P}}_W(\mathbf{u})\right\|} = \frac{\left\|\widetilde{\mathcal{P}}_W(\mathbf{u})\right\|}{\|\mathbf{u}\|} > 0$$

se $\mathbf{u} \neq \mathbf{0}$. Di conseguenza

$$\widehat{\mathbf{u}\,\widetilde{\mathcal{P}}_W(\mathbf{u})} = \arccos\left(\frac{\left\|\widetilde{\mathcal{P}}_W(\mathbf{u})\right\|}{\|\mathbf{u}\|}\right) \in \left[0, \frac{\pi}{2}\right).$$

Questo ci permette anche di verificare che l'angolo tra i due sottospazi non dipende dal particolare vettore $\mathbf{u}$. Infatti, ogni altro vettore in $U \setminus \{\mathbf{0}\}$ si può scrivere come $t \cdot \mathbf{u}$ con $t \in \mathbb{R}^*$, per il quale vale

$$\frac{\left\|\widetilde{\mathcal{P}}_W(t \cdot \mathbf{u})\right\|}{\|t \cdot \mathbf{u}\|} = \frac{|t| \left\|\widetilde{\mathcal{P}}_W(\mathbf{u})\right\|}{|t| \, \|\mathbf{u}\|} = \frac{\left\|\widetilde{\mathcal{P}}_W(\mathbf{u})\right\|}{\|\mathbf{u}\|}.$$

Esempio 10.20. Nello spazio vettoriale $V = \mathbb{R}^3$ con prodotto scalare euclideo, consideriamo i sottospazi

$$U = \{(x,y,z) \in \mathbb{R}^3 \,|\, y = z = 0\} \qquad \text{e} \qquad W = \{(x,y,z) \in \mathbb{R}^3 \,|\, ax + bz = 0\}.$$

per $a, b \in \mathbb{R}$ non entrambi nulli. Due basi ortonormali dei sottospazi sono

$$\mathcal{B}_U = \{\mathbf{e_1}\} \qquad \text{e} \qquad \mathcal{B}_W = \left\{ \mathbf{v} = \frac{1}{\sqrt{a^2+b^2}} \cdot (-b, 0, a), \mathbf{e_2} \right\}.$$

Segue che

$$\widetilde{\mathcal{P}}_W(\mathbf{e_1}) = \langle \mathbf{e_1}, \mathbf{v} \rangle \cdot \mathbf{v} + \langle \mathbf{e_1}, \mathbf{e_2} \rangle \cdot \mathbf{e_2} = \frac{b}{a^2+b^2} \cdot (b, 0, -a)$$

e quindi

$$\widehat{UW} = \arccos\left(\frac{|b|}{\sqrt{a^2+b^2}} \right).$$

In particolare, i due sottospazi sono ortogonali se $b = 0$, mentre $U \subset W$ se $a = 0$.

La generalizzazione della precedente definizione di angolo per sottospazi U e W di dimensione arbitraria non è affatto banale. Infatti, ad ogni sottospazio unidimensionale di U è associato un suo specifico angolo rispetto a W. Questo problema è stato affrontato per la prima volta da Jordan alla fine del diciannovesimo secolo ed è stato risolto definendo i cosiddetti angoli canonici o principali. Qui noi seguiremo un approccio moderno alla questione, basato sulla decomposizione ai valori singolari che abbiamo sviluppato nella Sezione 9.7.

Se $\mathcal{B} = \{\mathbf{u_1}, \dots, \mathbf{u_m}\}$ è una base di U, prendendo ispirazione dal caso monodimensionale risulterebbe naturale definire gli angoli tra U e W come $\widehat{UW}_i = \widehat{\mathbf{u_i}\, \widetilde{\mathcal{P}}_W(\mathbf{u_i})}$, per $1 \le i \le m$. Ovviamente il problema di tale definizione è che questi angoli dipendono dalla scelta della base $\mathcal{B}$. Inoltre, in questo modo risulta poco chiara la relazione tra gli angoli $\widehat{UW}_i$ e $\widehat{WU}_i$, che invece dovrebbero essere legati tra loro in modo naturale. É possibile superare entrambe queste difficoltà se consideriamo due basi associate in modo canonico alla proiezione ortogonale $\mathcal{P}_{UW}$, ovvero una base $\mathcal{B}_U$ composta da vettori singolari a destra della proiezione e la base duale $\mathcal{B}_W$ composta da vettori singolari a sinistra. Approfondiamo con ordine questa idea. Precisiamo che in quanto segue compiremo sistematicamente un abuso di notazione utilizzando lo stesso simbolo di vettore per $\mathbf{u} \in U$ ed $i_{UV}(\mathbf{u}) \in V$ ed analogamente per $\mathbf{w} \in W$ ed $i_{WV}(\mathbf{w}) \in V$.

Definizione 10.21. *Siano $\mathcal{B}_U = \{\mathbf{u_1}, \dots, \mathbf{u_m}\}$ e $\mathcal{B}_W = \{\mathbf{w_1}, \dots, \mathbf{w_n}\}$ due basi di U e W. Allora la matrice di Gram mista delle due basi è*

$$G|_{\mathcal{B}_U \mathcal{B}_W} = [\langle \mathbf{u_i}, \mathbf{w_j} \rangle] \in \mathrm{Mat}(m, n; \mathbb{R}).$$

Lemma 10.22. *Le proiezioni ortogonali $\mathcal{P}_{UW}$ e $\mathcal{P}_{WU}$ sono una l'aggiunta dell'altra, pertanto hanno gli stessi valori singolari e le stesse basi ortonormali duali composte da vettori singolari.*

Dimostrazione. Siano $\mathcal{B}_U = \{\mathbf{u_1}, \dots, \mathbf{u_m}\}$ e $\mathcal{B}_W = \{\mathbf{w_1}, \dots, \mathbf{w_n}\}$ due basi ortonormali di U e W. La matrice rappresentativa di $\mathcal{P}_{UW}$ è

$$\phi_{\mathcal{B}_U \mathcal{B}_W}(\mathcal{P}_{UW}) = \begin{bmatrix} \mathcal{P}_{UW}(\mathbf{u_1})|_{\mathcal{B}_W} & \dots & \mathcal{P}_{UW}(\mathbf{u_1})|_{\mathcal{B}_W} \end{bmatrix} = \begin{bmatrix} \langle \mathbf{u_1}, \mathbf{w_1} \rangle & \dots & \langle \mathbf{u_1}, \mathbf{w_n} \rangle \\ \vdots & & \vdots \\ \langle \mathbf{u_m}, \mathbf{w_1} \rangle & \dots & \langle \mathbf{u_m}, \mathbf{w_n} \rangle \end{bmatrix} = G|_{\mathcal{B}_U \mathcal{B}_W}.$$

Analogamente vale $\phi_{\mathcal{B}_W \mathcal{B}_U}(\mathcal{P}_{WU}) = G|_{\mathcal{B}_W \mathcal{B}_U} = G|^T_{\mathcal{B}_U \mathcal{B}_W}$, da cui segue la tesi. $\square$

Proposizione 10.23. *Siano* $\sigma_1 \geq \cdots \geq \sigma_r > 0$ *i valori singolari della proiezione ortogonale* $\mathcal{P}_{UW}$ *e siano* $\mathcal{B}_U = \{\mathbf{u_1}, \ldots, \mathbf{u_m}\}$ *e* $\mathcal{B}_W = \{\mathbf{w_1}, \ldots, \mathbf{w_n}\}$ *due basi ortonormali duali composte da vettori singolari. Allora*

$$\widehat{\mathbf{u_i}\,\widetilde{\mathcal{P}}_W(\mathbf{u_i})} = \widehat{\mathbf{w_i}\,\widetilde{\mathcal{P}}_U(\mathbf{w_i})} = \arccos(\sigma_i)$$

per ogni $1 \leq i \leq r$. *Inoltre,* $\mathbf{u_{r+1}}, \ldots, \mathbf{u_m} \in W^{\perp}$ *e* $\mathbf{w_{r+1}}, \ldots, \mathbf{w_n} \in U^{\perp}$.

Dimostrazione. Per definizione vale $\mathcal{P}_{UW}(\mathbf{u_i}) = \sigma_i \cdot \mathbf{w_i}$ per ogni $1 \leq i \leq r$. Di conseguenza

$$\left\langle \mathbf{u_i}, \widetilde{\mathcal{P}}_W(\mathbf{u_i}) \right\rangle = \left\langle \mathbf{u_i}, \widetilde{\mathcal{P}}^2_W(\mathbf{u_i}) \right\rangle = \left\langle \widetilde{\mathcal{P}}_W(\mathbf{u_i}), \widetilde{\mathcal{P}}_W(\mathbf{u_i}) \right\rangle = \langle \sigma_i \cdot \mathbf{w_i}, \sigma_i \cdot \mathbf{w_i} \rangle = \sigma_i^2 \left\| \mathbf{w_i} \right\|^2 = \sigma_i^2.$$

Inoltre vale

$$\left\| \widetilde{\mathcal{P}}_W(\mathbf{u_i}) \right\| = \left\| \sigma_i \cdot \mathbf{w_i} \right\| = |\sigma_i| \left\| \mathbf{w_i} \right\| = \sigma_i.$$

Di conseguenza

$$\widehat{\mathbf{u_i}\,\widetilde{\mathcal{P}}_W(\mathbf{u_i})} = \arccos\left(\frac{\left\langle \mathbf{u_i}, \widetilde{\mathcal{P}}_W(\mathbf{u_i}) \right\rangle}{\left\| \mathbf{u_i} \right\| \left\| \widetilde{\mathcal{P}}_W(\mathbf{u_i}) \right\|} \right) = \arccos\left(\frac{\sigma_i^2}{\sigma_i} \right) = \arccos\left(\sigma_i \right)$$

per ogni $1 \leq i \leq r$. La seconda uguaglianza segue dal Lemma 10.22 e dalle identità analoghe alle precedenti valide per l'applicazione aggiunta. Infine, l'ultima affermazione segue direttamente da

$$\mathcal{P}_{UW}(\mathbf{u_{r+1}}) = \cdots = \mathcal{P}_{UW}(\mathbf{u_m}) = \mathbf{0_W} \qquad \text{e} \qquad \mathcal{P}_{WU}(\mathbf{w_{r+1}}) = \cdots = \mathcal{P}_{WU}(\mathbf{w_n}) = \mathbf{0_U}. \qquad \square$$

É importante sottolineare che la precedente proposizione è valida per ogni coppia di basi duali. Quindi gli angoli tra i vettori singolari e le relative proiezioni sono delle quantità intrinseche associate ai due sottospazi, che non dipendono dalla particolare scelta delle relative basi.

Definizione 10.24. *Gli angoli principali o canonici tra* U *e* W *sono*

$$\begin{cases} \widehat{UW}_i = \widehat{WU}_i = \arccos(\sigma_i) & \textit{se} \quad 1 \leq i \leq r, \\ \widehat{UW}_i = \dfrac{\pi}{2} & \textit{se} \quad r < i \leq m, \\ \widehat{WU}_i = \dfrac{\pi}{2} & \textit{se} \quad r < i \leq v, \end{cases}$$

dove $\sigma_1 \geq \cdots \geq \sigma_r > 0$ *sono i valori singolari delle proiezioni ortogonali* P_{UW} *e* P_{WU}.

Prima di presentare un esempio conclusivo, facciamo alcune importanti osservazioni:

- gli angoli principali sono compresi nell'intervallo $[0, \frac{\pi}{2}]$, in quanto i valori singolari sono per definizione quantità positive. Naturalmente, vale anche $0 < \sigma_i \leq 1$ per ogni $1 \leq i \leq r$;
- i valori singolari di P_{UW} e P_{WU} sono uguali ai valori singolari delle relative matrici rappresentative, ovvero delle matrici $G|_{\mathcal{B}_U\mathcal{B}_W}$ e $G|^T_{\mathcal{B}_U\mathcal{B}_W}$ per ogni scelta delle basi $\mathcal{B}_U$ e $\mathcal{B}_W$ dei due sottospazi;
- se $m = 1$, riotteniamo la Definizione 10.19 per l'unico angolo $\widehat{UW}$. Inoltre $\widehat{WU}_1 = \widehat{UW}$ e $\widehat{WU}_i = \frac{\pi}{2}$ per ogni $2 \leq i \leq n$.

Infine, possiamo finalmente definire in maniera naturale gli angoli tra sottospazi affini.

Definizione 10.25. *Gli angoli principali tra due sottospazi affini* S *e* T *di uno spazio euclideo* $\mathbb{E}$ *sono uguali agli angoli principali tra le loro rispettive giaciture.*

Esempio 10.26. In $\mathbb{E}^3$, i sottospazi affini

$$S = \{(x, y, z) \in \mathbb{R}^3 \mid y = z = 1\} \qquad \text{e} \qquad T = \{(x, y, z) \in \mathbb{R}^3 \mid ax + bz = -1\}.$$

hanno come giaciture i sottospazi vettoriali U e W dell'Esempio 10.20. Alle basi ortonormali

$$\mathcal{B}_U = \{\mathbf{e_1}\} \qquad \text{e} \qquad \mathcal{B}_W = \left\{ \mathbf{v} = \frac{1}{\sqrt{a^2+b^2}} \cdot (-b, 0, a), \mathbf{e_2} \right\}.$$

di U e W corrisponde la matrice di Gram mista

$$G|_{\mathcal{B}_U \mathcal{B}_W} = \left[\begin{array}{cc} \langle \mathbf{e_1}, \mathbf{v} \rangle & \langle \mathbf{e_1}, \mathbf{e_2} \rangle \end{array} \right] = \left[\begin{array}{cc} \frac{-b}{\sqrt{a^2+b^2}} & 0 \end{array} \right].$$

Dato che $G|_{\mathcal{B}_U \mathcal{B}_W} * G|^T_{\mathcal{B}_U \mathcal{B}_W} = \left[\frac{b^2}{a^2+b^2} \right]$, ne segue che l'unico valore singolare è $\sigma_1 = \frac{|b|}{\sqrt{a^2+b^2}}$. Due basi ortonormali duali di vettori singolari sono $\tilde{\mathcal{B}}_U = \{\mathbf{e_1}\}$ e $\tilde{\mathcal{B}}_W = \{\tilde{\mathbf{v}} = -\mathrm{sgn}(b) \cdot \mathbf{v}, \mathbf{e_2}\}$. Infatti valgono

$$\mathcal{P}_{UW}(\mathbf{e_1}) = \langle \mathbf{e_1}, \mathbf{v} \rangle \cdot \mathbf{v} + \langle \mathbf{e_1}, \mathbf{e_2} \rangle \cdot \mathbf{e_2} = \frac{-b}{\sqrt{a^2+b^2}} \cdot \mathbf{v} = \sigma_1 \cdot \tilde{\mathbf{v}},$$

$$\mathcal{P}_{WU}(\tilde{\mathbf{v}}) = \langle \tilde{\mathbf{v}}, \mathbf{e_1} \rangle \cdot \mathbf{e_1} = \frac{-b}{\sqrt{a^2+b^2}} \cdot \mathbf{v} = \sigma_1 \cdot \tilde{\mathbf{v}} \qquad \text{e} \qquad \mathcal{P}_{WU}(\mathbf{e_2}) = \langle \mathbf{e_2}, \mathbf{e_1} \rangle \cdot \mathbf{e_1} = \mathbf{0_U},$$

da cui ricaviamo

$$\widehat{ST}_1 = \widehat{\mathbf{e_1}\mathcal{P}_{UW}(\mathbf{e_1})} = \widehat{TS}_1 = \widehat{\tilde{\mathbf{v}}\mathcal{P}_{WU}(\tilde{\mathbf{v}})} = \arccos\left(\frac{|b|}{\sqrt{a^2+b^2}}\right) \qquad \text{e} \qquad \widehat{TS}_2 = \frac{\pi}{2}.$$

$\widehat{ST}_1$ coincide con l'angolo tra la retta S e la sua proiezione ortogonale $\mathcal{P}_T(S)$ che abbiamo calcolato nell'Esempio 10.20.

Consideriamo invece il piano coordinato $\Pi = \{(x, y, z) \in \mathbb{R}^3 \mid y = 0\}$, la cui giacitura Z ha base $\mathcal{B}_Z = \{\mathbf{e_1}, \mathbf{e_3}\}$. Calcoliamo gli angoli tra i due piani T e Π. La matrice di Gram mista è

$$G|_{\mathcal{B}_W \mathcal{B}_Z} = \left[\begin{array}{cc} \langle \mathbf{v}, \mathbf{e_1} \rangle & \langle \mathbf{v}, \mathbf{e_3} \rangle \\ \langle \mathbf{e_2}, \mathbf{e_1} \rangle & \langle \mathbf{e_2}, \mathbf{e_3} \rangle \end{array} \right] = \left[\begin{array}{cc} \frac{-b}{\sqrt{a^2+b^2}} & \frac{a}{\sqrt{a^2+b^2}} \\ 0 & 0 \end{array} \right] \qquad \Rightarrow \qquad G|_{\mathcal{B}_W \mathcal{B}_Z} * G|^T_{\mathcal{B}_W \mathcal{B}_Z} = \left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array} \right].$$

Anche in questo caso esiste un unico valore singolare $\sigma_1 = 1$. Due basi ortonormali duali di vettori singolari sono $\tilde{\mathcal{B}}_W = \{\mathbf{v}, \mathbf{e_2}\}$ e $\tilde{\mathcal{B}}_Z = \{\mathbf{v}, \mathbf{v}' = \frac{1}{\sqrt{a^2+b^2}} \cdot (a, 0, b)\}$. Infatti valgono

$$\mathcal{P}_{WZ}(\mathbf{v}) = \langle \mathbf{v}, \mathbf{v} \rangle \cdot \mathbf{v} + \left\langle \mathbf{v}, \mathbf{v}' \right\rangle \cdot \mathbf{v}' = \mathbf{v} = \sigma_1 \cdot \tilde{\mathbf{v}} \qquad \text{e} \qquad \mathcal{P}_{WZ}(\mathbf{e_2}) = \langle \mathbf{e_2}, \mathbf{v} \rangle \cdot \mathbf{v} + \left\langle \mathbf{e_2}, \mathbf{v}' \right\rangle \cdot \mathbf{v}' = \mathbf{0_Z},$$

$$\mathcal{P}_{ZW}(\mathbf{v}) = \langle \mathbf{v}, \mathbf{v} \rangle \cdot \mathbf{v} + \langle \mathbf{v}, \mathbf{e_2} \rangle \cdot \mathbf{e_2} = \mathbf{v} = \sigma_1 \cdot \tilde{\mathbf{v}} \qquad \text{e} \qquad \mathcal{P}_{ZW}(\mathbf{v}') = \left\langle \mathbf{v}', \mathbf{v} \right\rangle \cdot \mathbf{v} + \left\langle \mathbf{v}', \mathbf{e_2} \right\rangle \cdot \mathbf{e_2} = \mathbf{0_W},$$

da cui ricaviamo

$$\widehat{T\Pi}_1 = \widehat{\mathbf{v}\mathcal{P}_{WZ}(\mathbf{v})} = \widehat{\Pi T}_1 = \widehat{\tilde{\mathbf{v}}\mathcal{P}_{ZW}(\mathbf{v})} = \arccos(1) = 0 \qquad \text{e} \qquad \widehat{T\Pi}_2 = \widehat{\Pi T}_2 = \frac{\pi}{2}.$$

Gli angoli $\widehat{T\Pi}_1 = \widehat{\Pi T}_1 = 0$ sono conseguenza del fatto che T e Π si intersecano lungo una retta di giacitura $\mathcal{L}(\mathbf{v})$, che ovviamente è parallela ad entrambi i piani. La condizione $\widehat{T\Pi}_2 = \widehat{\Pi T}_2 = \frac{\pi}{2}$ ci dice invece che i due piani si intersecano perpendicolarmente lungo tale retta.

CAPITOLO 11

Ipersuperfici quadriche

Uno dei principali successi degli antichi geometrici greci è stato lo studio, all'interno della geometria euclidea, delle cosiddette sezioni coniche, ovvero le curve ottenibili come l'intersezione di un cono con un piano. Concludiamo il corso con la presentazione di tali oggetti geometrici all'interno del formalismo che abbiamo fino ad ora sviluppato. In tutta questa sezione assumiamo di lavorare in uno spazio euclideo orientato $\mathbb{E}$ di dimensione n.

11.1. Quadriche

Iniziamo definendo due sottoinsiemi naturali di $\mathbb{E}$.

Definizione 11.1. *L'ipersfera e l'iperdisco di centro C e raggio r sono rispettivamente i sottoinsiemi di punti*

$$\mathscr{S}_{C,r} = \{P \in \mathbb{E} \mid d(C,P) = r\} \qquad e \qquad \mathscr{D}_{C,r} = \{P \in \mathbb{E} \mid d(C,P) \leq r\}.$$

Se $\dim(\mathbb{E}) = 2$ *si parla rispettivamente di circonferenza e cerchio, mentre se* $\dim(\mathbb{E}) = 3$ *di sfera e disco.*

Scegliamo un sistema di riferimento cartesiano positivamente orientato $\mathcal{B}_O$ di $\mathbb{E}$, rispetto a cui il generico punto P ed il centro C hanno coordinate

$$P|_{\mathcal{B}_O} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad \text{e} \qquad C|_{\mathcal{B}_O} = \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_n \end{bmatrix}.$$

Allora vale

$$\overrightarrow{CP}|_{\mathcal{B}} = P|_{\mathcal{B}_O} - C|_{\mathcal{B}_O} = \begin{bmatrix} x_1 - \bar{x}_1 \\ \vdots \\ x_n - \bar{x}_n \end{bmatrix}$$

e quindi

$$d(C,P) = \left\| \overrightarrow{CP} \right\| = \sqrt{\overrightarrow{CP}|_{\mathcal{B}}^T * \overrightarrow{CP}|_{\mathcal{B}}} = \sqrt{\sum_{i=1}^{n} (x_i - \bar{x}_i)^2}\,.$$

Pertanto le rappresentazioni algebriche, o cartesiane, dell'ipersfera e dell'iperdisco sono rispettivamente

$$\sum_{i=1}^{n} (x_i - \bar{x}_i)^2 = r^2 \qquad \text{e} \qquad \sum_{i=1}^{n} (x_i - \bar{x}_i)^2 \leq r^2.$$

Dato che queste equazioni ammettono soluzione in $\mathbb{R}^n$, con la precedente definizione di ipersfera si soddisfa il (iii) postulato di Euclide e si completa quindi la costruzione del modello per la geometria euclidea, che ci eravamo posti come obiettivo principale del nostro studio nell'Introduzione di queste dispense. All'interno di tale modello è possibile riprodurre e studiare in modo efficiente tutti i risultati noti delle geometria euclidea. Come detto sopra, qui forniremo un importante esempio di questo fatto, ovvero lo studio delle curve coniche e della loro generalizzazione in dimensione maggiore di 2.

Definizione 11.2. *Sia $\mathcal{B}_O$ un sistema di riferimento cartesiano positivamente orientato di $\mathbb{E}$. Una ipersuperficie quadrica $\mathcal{Q}$ è il luogo dei punti le cui coordinate $x_1, \ldots, x_n$ sono soluzioni di un'equazione di secondo grado del tipo*

$$\sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i=1}^{n}\sum_{j>i}^{n} a_{ij}x_ix_j + 2\sum_{i=1}^{n} b_ix_i + c = 0,$$

dove a_{ij}, b_i e c sono coefficienti reali per ogni $1 \leq i \leq j \leq n$. Se $\dim(\mathbb{E}) = 2$ le ipersuperfici quadriche si dicono curve coniche, mentre se $\dim(\mathbb{E}) = 3$ si parla di superfici quadriche.

Naturalmente l'ipersfera è un esempio di ipersuperficie quadrica.

L'utilità del fattore 2 che compare nella precedente equazione si palesa utilizzando la cosiddetta rappresentazione matriciale delle quadriche, che ora presentiamo seguendo le notazioni della Definizione 11.2.

Definizione 11.3. *La matrice dei coefficienti di secondo grado di $\mathcal{Q}$ è $A \in \mathbb{S}(n;\mathbb{R})$, i cui elementi sono*

$$(A)_{ij} = \begin{cases} a_{ij} & \textit{se } 1 \leq i \leq j \leq n, \\ a_{ji} & \textit{altrimenti.} \end{cases}$$

La colonna dei coefficienti di primo grado di $\mathcal{Q}$ è $B \in \mathrm{Mat}(n,1;\mathbb{R})$, i cui elementi sono $(B)_{i1} = b_i$ per $1 \leq i \leq n$. Infine, la matrice completa di $\mathcal{Q}$ è

$$C = \left[\begin{array}{c|c} A & B \\ \hline B^T & c \end{array}\right] \in \mathbb{S}(n+1;\mathbb{R}).$$

Proposizione 11.4. *Data le colonne*

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathrm{Mat}(n,1;\mathbb{R}) \qquad e \qquad Y = \left[\begin{array}{c} X \\ \hline 1 \end{array}\right] \in \mathrm{Mat}(n+1,1;\mathbb{R}),$$

l'equazione dell'ipersuperficie quadrica $\mathcal{Q}$ è uguale a

$$X^T * A * X + 2 \cdot B^T * X + c = 0 \qquad \textit{oppure} \qquad Y^T * C * Y = 0.$$

La dimostrazione della proposizione è una semplice questione di calcolo che lasciamo al lettore.

Esempio 11.5. In $\mathbb{E}^2$ l'equazione generale di una curva conica è

$$a_{11}x_1^2 + a_{22}x_2^2 + 2a_{12}x_1x_2 + 2b_1x_1 + 2b_2x_2 + c = 0.$$

Le matrici associate alla conica sono

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \in \mathbb{S}(2;\mathbb{R}), \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \in \mathrm{Mat}(2,1;\mathbb{R}) \quad \text{e} \quad C = \begin{bmatrix} a_{11} & a_{12} & b_1 \\ a_{12} & a_{22} & b_2 \\ b_1 & b_2 & c \end{bmatrix} \in \mathbb{S}(3;\mathbb{R}).$$

In $\mathbb{E}^3$ l'equazione generale di una superficie quadrica è

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3 + 2b_1x_1 + 2b_2x_2 + 2b_3x_3 + c = 0.$$

Le matrici associate alla quadrica sono

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \in \mathbb{S}(3;\mathbb{R}), \quad B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \in \mathrm{Mat}(3,1;\mathbb{R}) \quad \text{e}$$

$$C = \begin{bmatrix} a_{11} & a_{12} & a_{13} & b_1 \\ a_{12} & a_{22} & a_{23} & b_2 \\ a_{13} & a_{23} & a_{33} & b_3 \\ b_1 & b_2 & b_3 & c \end{bmatrix} \in \mathbb{S}(4;\mathbb{R}).$$

11.2. Riduzione in forma canonica delle quadriche

La Definizione 11.2 di ipersuperficie quadrica $\mathcal{Q}$ richiede preliminarmente la scelta arbitraria di un sistema di riferimento cartesiano dello spazio euclideo. Come sempre in queste situazioni, è necessario chiedersi cosa succeda qualora noi scegliessimo due sistemi di riferimento differenti. In particolare:

- $\mathcal{Q}$ sarà descritta in entrambi i casi da un'equazione di secondo grado? Come saranno legate le due equazioni?
- esiste un sistema di riferimento rispetto a cui l'equazione di $\mathcal{Q}$ assume una forma particolarmente semplice?

I due punti precedenti ci introducono al problema dell'individuazione del sistema di riferimento canonico di $\mathcal{Q}$. Preliminarmente, ricordiamo la regola del cambiamento di coordinate tra sistemi di riferimento, che abbiamo ottenuto nell'Esempio 7.38. Siano $\mathcal{B}_O$ e $\tilde{\mathcal{B}}_O$ due sistemi di riferimento cartesiani positivamente orientati dello spazio euclideo $\mathbb{E}$ di dimensione n. Allora le coordinate di un punto P rispetto ai due sistemi sono legate da

$$P|_{\mathcal{B}_O} = M_{\tilde{\mathcal{B}}\mathcal{B}} * P|_{\tilde{\mathcal{B}}_O} + \tilde{O}|_{\mathcal{B}_O},$$

dove $M_{\tilde{\mathcal{B}}\mathcal{B}} \in SO(n;\mathbb{R})$. Per alleggerire la notazione, chiamiamo

$$X = P|_{\mathcal{B}_O}, \quad \tilde{X} = P|_{\tilde{\mathcal{B}}_O}, \quad T = \tilde{O}|_{\mathcal{B}_O} \quad \text{e} \quad Q = M_{\tilde{\mathcal{B}}\mathcal{B}}.$$

Di conseguenza

$$X = Q * \tilde{X} + T.$$

Seguendo queste notazioni, investighiamo il cambiamento nella rappresentazione di $\mathcal{Q}$ a seguito della modifica del sistema di riferimento cartesiano.

Proposizione 11.6. *Sia $\mathcal{Q}$ una quadrica che rispetto a $\mathcal{B}_O$ abbia matrici associate $A \in \mathbb{S}(n;\mathbb{R})$, $B \in \mathrm{Mat}(n,1;\mathbb{R})$, $C \in \mathbb{S}(n+1;\mathbb{R})$ e termine noto $c \in \mathbb{R}$. Allora rispetto a $\tilde{\mathcal{B}}_O$ l'ipersuperficie $\mathcal{Q}$ è ancora descritta da un'equazione di secondo grado, le cui matrici associate sono:*

$$\tilde{A} = Q^T * A * Q \in \mathbb{S}(n;\mathbb{R}), \quad \tilde{B} = Q^T * (A * T + B) \in \mathrm{Mat}(n,1;\mathbb{R}) \quad e \quad \tilde{C} = F^T * C * F,$$

dove

$$F = \left[\begin{array}{c|c} Q & T \\ \hline 0_{1n} & 1 \end{array}\right] \in \mathrm{Mat}(n+1,n+1;\mathbb{R}).$$

Infine, il termine noto di $\mathcal{Q}$ rispetto a $\tilde{\mathcal{B}}_O$ è

$$\tilde{c} = T^T * A * T + 2 \cdot B^T * T + c.$$

DIMOSTRAZIONE. L'equazione della quadrica è

$$\begin{aligned} 0 &= X^T * A * X + 2 \cdot B^T * X + c = \\ &= (Q * \tilde{X} + T)^T * A * (Q * \tilde{X} + T) + 2 \cdot B^T * (Q * \tilde{X} + T) + c = \\ &= \tilde{X}^T * Q^T * A * Q * \tilde{X} + \tilde{X}^T * Q^T * A * T + T^T * A * Q * \tilde{X} + 2 \cdot B^T * Q * \tilde{X} + \\ &\quad + T^T * A * T + 2 \cdot B^T * T + c\,. \end{aligned}$$

Dalle proprietà della trasposizione e dalla simmetria di A, segue che

$$\tilde{X}^T * Q^T * A * T = (\tilde{X}^T * Q^T * A * T)^T = T^T * A * Q * \tilde{X}.$$

Di conseguenza, possiamo riscrivere l'equazione di $\mathcal{Q}$ come

$$\tilde{X}^T * (Q^T * A * Q) * \tilde{X} + 2 \cdot (T^T * A + B^T) * Q * \tilde{X} + (T^T * A * T + 2 \cdot B^T * T + c) = 0\,,$$

da cui seguono immediatamente le affermazioni dell'enunciato riguardanti $\tilde{A}$, $\tilde{B}$ e $\tilde{c}$. Per la proprietà della matrice completa della quadrica, possiamo effettuare la verifica diretta:

$$F^T * C * F = \left[\begin{array}{c|c} Q^T & 0_{n1} \\ \hline T^T & 1 \end{array}\right] * \left[\begin{array}{c|c} A & B \\ \hline B^T & c \end{array}\right] * \left[\begin{array}{c|c} Q & T \\ \hline 0_{1n} & 1 \end{array}\right] =$$

$$= \left[\begin{array}{c|c} Q^T * A * Q & Q^T * (A * T + B) \\ \hline (T^T * A + B^T) * Q & T^T * A * T + 2 \cdot B^T * T + c \end{array}\right] = \left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline \tilde{B}^T & \tilde{c} \end{array}\right] = \tilde{C}. \qquad \square$$

Corollario 11.7. *A seguito del cambio di coordinate $X = Q * \tilde{X} + T$ valgono le seguenti identità:*

i) $r(A) = r(\tilde{A})$ *e* $r(C) = r(\tilde{C})$;

ii) $P_A(\lambda) = P_{\tilde{A}}(\lambda)$;

iii) $\det(C) = \det(\tilde{C})$.

Dimostrazione. La matrice Q è ortogonale speciale e quindi è invertibile. Di conseguenza $r(A)$ e $P_A(\lambda)$ non cambiano a seguito della trasformazione $\tilde{A} = Q^T * A * Q$. Inoltre $\det(F) = 1\ \det(Q) = 1$, quindi anche F è invertibile. Di conseguenza $r(C)$ non cambia a seguito della trasformazione ed inoltre $\det(\tilde{C}) = \det(F^T)\det(C)\det(F) = \det(F)^2\det(C) = \det(C)$. $\square$

Definizione 11.8. *Data una rappresentazione della quadrica $\mathcal{Q}$ attraverso le matrici A e C, gli invarianti metrici sono:*

- $I_i = (-1)^{n-i} c_{n-i}$, *dove c_i sono i coefficienti del polinomio caratteristico $P_A(\lambda)$ per $1 \leq i \leq n$;*
- $I_{n+1} = \det(C)$.

Esempio 11.9. In $\mathbb{E}^2$ gli invarianti metrici di una curva conica sono

$$I_1 = -c_1 = \mathrm{Tr}(A), \qquad I_2 = c_0 = \det(A) \qquad \text{ed} \qquad I_3 = \det(C).$$

In $\mathbb{E}^3$ gli invarianti metrici di una superficie quadrica sono

$$I_1 = c_2 = \mathrm{Tr}(A), \qquad I_2 = -c_1 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{13} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{23} & a_{33} \end{vmatrix},$$

$$I_3 = c_0 = \det(A) \qquad \text{ed} \qquad I_4 = \det(C).$$

Grazie al Corollario 11.7 sappiamo che gli invarianti metrici rimangono uguali a seguito del cambiamento del sistema di riferimento cartesiano. Bisogna però prestare attenzione al fattore moltiplicativo globale dell'equazione della quadrica. Infatti, se l'equazione di $\mathcal{Q}$ viene moltiplicata per $t \in \mathbb{R}^*$, otteniamo ancora una rappresentazione della quadrica:

$$t\left(\sum_{i=1}^n a_{ii}x_i^2 + 2\sum_{i=1}^n\sum_{j>i}^n a_{ij}x_ix_j + 2\sum_{i=1}^n b_ix_i + c\right) = \sum_{i=1}^n ta_{ii}x_i^2 + 2\sum_{i=1}^n\sum_{j>i}^n ta_{ij}x_ix_j + 2\sum_{i=1}^n tb_ix_i + tc = 0.$$

É semplice verificare che a seguito di questa trasformazione dell'equazione, le matrici associate a $\mathcal{Q}$ diventano $\tilde{A} = t \cdot A$ e $\tilde{C} = t \cdot C$ e quindi gli invarianti trasformati sono $\tilde{I}_i = t^i I_i$, per ogni $1 \leq i \leq n+1$.

Possiamo ora enunciare e dimostrare il teorema di riduzione in forma canonica delle quadriche.

Teorema 11.10. *Supponiamo che le matrici A e C associate alla quadrica $\mathcal{Q}$ abbiano rispettivamente rango r e q. Inoltre, siano $\lambda_1, \ldots, \lambda_r$ gli autovalori non nulli di A, eventualmente coincidenti. Allora esiste un sistema di riferimento cartesiano positivamente orientato $\tilde{\mathcal{B}}$ di $\mathbb{E}$ che è canonico per $\mathcal{Q}$, rispetto al quale l'equazione della quadrica è:*

i) $\lambda_1\tilde{x}_1^2 + \cdots + \lambda_r\tilde{x}_r^2 = 0$ *se* $q = r$;

ii) $\lambda_1\tilde{x}_1^2 + \cdots + \lambda_r\tilde{x}_r^2 + \tilde{c} = 0$, *con* $\tilde{c} \neq 0$, *se* $q = r + 1$;

iii) $\lambda_1\tilde{x}_1^2 + \cdots + \lambda_r\tilde{x}_r^2 + 2p\tilde{x}_{r+1} = 0$, *con* $p > 0$, *se* $q = r + 2$.

Dimostrazione. Dobbiamo costruire il cambiamento di coordinate $X = Q * \tilde{X} + T$ necessario a porre l'equazione della quadrica in una delle forme dell'enunciato. Per prima cosa osserviamo che la matrice A è simmetrica e quindi è ortogonalmente diagonalizzabile. Questo significa che possiamo scegliere come Q una matrice ortogonale speciale che diagonalizza A, ovvero tale per cui $\tilde{A} = Q^T * A * Q \in \mathbb{D}(n; \mathbb{R})$. Se la matrice A ha rango r, allora esistono r autovalori non nulli, eventualmente coincidenti. Ordinando opportunamente le variabili, il termine di secondo grado del polinomio associato a $\mathcal{Q}$ nel nuovo sistema di riferimento è $\lambda_1\tilde{x}_1^2 + \cdots + \lambda_r\tilde{x}_r^2$. Rimane quindi da individuare la matrice colonna T.

Se $r = r([A|B])$, allora per il Teorema di Rouchè-Capelli esiste una colonna T tale che $A*T+B = 0_{n1}$. Di conseguenza, a seguito del cambio di coordinate $X = Q * \tilde{X} + T$, si ottiene la rappresentazione

$$\lambda_1\tilde{x}_1^2 + \cdots + \lambda_r\tilde{x}_r^2 + \tilde{c} = 0$$

di $\mathcal{Q}$. Dato che

$$\tilde{C} = \left[\begin{array}{c|c} \tilde{A} & 0_{n1} \\ \hline 0_{1n} & \tilde{c} \end{array}\right] \qquad \text{segue che} \qquad q = \begin{cases} r & \text{se } \tilde{c} = 0, \\ r+1 & \text{se } \tilde{c} \neq 0. \end{cases}$$

Se $r \neq r([A|B])$ la situazione è più complicata e bisogna comporre più cambiamenti di coordinate per ottenere la forma canonica di $\mathcal{Q}$. Elenchiamoli separatamente con ordine:

i) eseguiamo come prima la rotazione $X = Q_1 * X_1$, a seguito della quale $A_1 = Q_1^T * A * Q_1$ è diagonale, $B_1 = Q_1^T * B$ e $c_1 = c$;

ii) definiamo la matrice colonna T_2 i cui elementi sono

$$(T_2)_{i1} = \begin{cases} -\dfrac{(B_1)_{i1}}{\lambda_i} & \text{se} \quad 1 \leq i \leq r, \\ 0 & \text{se} \quad r+1 \leq i \leq n. \end{cases}$$

A seguito della trasformazione $X_1 = X_2 + T_2$, otteniamo $A_2 = A_1$. La colonna $B_2 = A_1 * T_2 + B_1$ ha i primi r elementi nulli, mentre $c_2 = -\sum_{i=1}^{r} \frac{(B_1)_{i1}^2}{\lambda_i} + c_1$. Osserviamo che la condizione sui ranghi della matrice impone che $B_2 \neq 0_{n1}$;

iii) B_2 appartiene al sottospazio vettoriale

$$U = \{M \in \mathrm{Mat}(n, 1; \mathbb{R}) \,|\, (M)_{i1} = 0 \text{ per } 1 \leq i \leq r\},$$

la cui dimensione è $n-r$. Utilizzando gli algoritmi di completamento e di Gram-Schmidt, l'insieme formato dalla sola colonna B_2 può essere esteso ad una base ortonormale positivamente orientata $\left\{\frac{B_2}{\|B_2\|_F}, M_1, \ldots, M_{n-r-1}\right\}$ di U. Definiamo quindi la matrice ortogonale speciale

$$Q_3 = \left[\begin{array}{c|cccc} \mathbb{I}_{nr} & \frac{B_2}{\|B_2\|_F} & M_1 & \ldots & M_{n-r-1} \end{array}\right] \in SO(n; \mathbb{R}).$$

A seguito della trasformazione $X_2 = Q_3 * X_3$ si ottiene $A_3 = Q_3^T * A_2 * Q_3 = A_2$, $B_3 = Q_3^T * B_2$, il cui unico elemento non nullo si trova al posto $(1, r+1)$ ed è uguale a $p = \|B_2\|_F > 0$, ed infine $c_3 = c_2$;

iv) definiamo la colonna T_4 il cui unico elemento non nullo si trova al posto $(1, r+1)$ ed è uguale a $-\frac{c_3}{2p}$. Dalla trasformazione $X_3 = X_4 + T_4$ otteniamo $A_4 = A_3$, $B_4 = B_3$ e $c_4 = 0$.

Componendo tutte le trasformazioni abbiamo

$$X = Q_1 * X_1 = Q_1 * (X_2 + T_2) = Q_1 * (Q_3 * X_3) + Q_1 * T_2 = Q_1 * Q_3 * X_4 + Q_1 * (Q_3 * T_4 + T_2) = Q * \tilde{X} + T,$$

da cui $Q = Q_1 * Q_3$ e $T = Q_1 * (Q_3 * T_4 + T_2)$. A seguito di questa trasformazione di coordinate si ottiene la matrice completa

$$\tilde{C} = \left[\begin{array}{cccccc|c} \lambda_1 & \dots & 0 & 0 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \dots & \lambda_r & 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 & p \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ \hline 0 & \dots & 0 & p & \dots & 0 & 0 \end{array}\right].$$

La rappresentazione associata di $\mathcal{Q}$ è

$$\lambda_1 \tilde{x}_1^2 + \cdots + \lambda_r \tilde{x}_r^2 + 2p\tilde{x}_{r+1} = 0.$$

Dato che gli elementi non nulli di $\tilde{C}$ appartengono ad $r+2$ righe e colonne differenti, segue immediatamente che $q = r(\tilde{C}) = r(C) = r+2$. $\square$

Esempio 11.11. In $\mathbb{E}^3$ consideriamo la superficie quadrica $\mathcal{Q}$, che rispetto al sistema di riferimento $\mathcal{B}$ ha equazione

$$x^2 + y^2 + z^2 + 2xy + 2xz + 2yz + 2y + 2z + 1 = 0.$$

Le matrici associate a $\mathcal{Q}$ sono

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \qquad \text{e} \qquad C = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}.$$

É facile verificare che $r(A) = 1$ e $r(C) = 3$. Gli autovalori di A sono $\lambda_1 = 3$ e $\lambda_2 = \lambda_3 = 0$, corrispondenti alla seguente base ortonormale di autovettori

$$\mathcal{B} = \left\{ \mathbf{v_1} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \mathbf{v_2} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}, \mathbf{v_3} = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix} \right\}.$$

La matrice Q_1 che diagonalizza A è

$$Q_1 = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \end{bmatrix} \qquad \Rightarrow$$

$$A_1 = Q_1^T * A * Q_1 = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad B_1 = Q_1^T * B = \begin{bmatrix} 2/\sqrt{3} \\ -1/\sqrt{2} \\ -1/\sqrt{6} \end{bmatrix}, \qquad c_1 = c = 1.$$

La matrice colonna T_2 è

$$T_2 = \begin{bmatrix} -2/3\sqrt{3} \\ 0 \\ 0 \end{bmatrix} \qquad \Rightarrow$$

$$A_2 = A_1 = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad B_2 = A_1 * T_2 + B_1 = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ -1/\sqrt{6} \end{bmatrix}, \qquad c_2 = -\frac{(B_1)_{11}^2}{\lambda_1} + c_1 = \frac{5}{9}.$$

La matrice ortogonale Q_3 è

$$Q_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\sqrt{3}/2 & 1/2 \\ 0 & -1/2 & -\sqrt{3}/2 \end{bmatrix} \qquad \Rightarrow$$

$$A_3 = Q_3^T * A_2 * Q_3 = A_2 = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad B_3 = Q_3^T * B_2 = \begin{bmatrix} 0 \\ 2/\sqrt{6} \\ 0 \end{bmatrix}, \qquad c_3 = c_2 = \frac{5}{9}.$$

La colonna T_4 è

$$T_4 = \begin{bmatrix} 0 \\ -5/6\sqrt{6} \\ 0 \end{bmatrix} \qquad \Rightarrow$$

$$A_4 = A_3 = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad B_4 = A_3 * T_4 + B_3 = B_3 = \begin{bmatrix} 0 \\ 2/\sqrt{6} \\ 0 \end{bmatrix}, \qquad c_4 = 2 \cdot B_3^T * T_4 + c_3 = 0.$$

La trasformazione di coordinate completa è generata da

$$Q = Q_1 * Q_3 = \begin{bmatrix} 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \end{bmatrix} \qquad \text{e} \qquad T = Q_1 * (Q_3 * T_4 + T_2) = \frac{1}{36} \cdot \begin{bmatrix} 2 \\ -13 \\ -13 \end{bmatrix},$$

ovvero

$$\begin{cases} x = \frac{1}{\sqrt{3}}\tilde{x} - \frac{2}{\sqrt{6}}\tilde{y} + \frac{1}{18} \\ y = \frac{1}{\sqrt{3}}\tilde{x} + \frac{1}{\sqrt{6}}\tilde{y} - \frac{1}{\sqrt{2}}\tilde{z} - \frac{13}{36} \\ z = \frac{1}{\sqrt{3}}\tilde{x} + \frac{1}{\sqrt{6}}\tilde{y} + \frac{1}{\sqrt{2}}\tilde{z} - \frac{13}{36} \end{cases}.$$

L'equazione canonica della quadrica nel sistema di riferimento $\tilde{\mathcal{B}}_O$ è

$$3\tilde{x}^2 + \frac{4}{\sqrt{6}}\tilde{y} = 0.$$

Osservazione 11.12. Esistono due modi differenti di interpretare le trasformazioni che abbiamo descritto in questa sottosezione. Quello che noi abbiamo fatto è stato pensare ad esse come ad un cambiamento del sistema di coordinate da $\mathcal{B}$ a $\tilde{\mathcal{B}}$, che comporta una modifica delle equazioni che descrivono la quadrica fissata $\mathcal{Q}$ (*interpretazione passiva* della trasformazione). Alternativamente, si può pensare di lasciare fisso il sistema di riferimento $\mathcal{B}$ e trasformare invece $\mathcal{Q}$ in una nuova quadrica $\tilde{\mathcal{Q}}$ (*interpretazione attiva* della trasformazione). In questo caso le due ipersuperfici sono legate dalla rotazione rappresentata dalla matrice Q e dalla traslazione rappresentata dalla matrice T. La loro composizione è un'isometria di $\mathbb{E}$ e quindi le due quadriche sono congruenti. In tale approccio, la riduzione in forma canonica di $\mathcal{Q}$ corrisponde all'individuazione della quadrica più semplice $\tilde{\mathcal{Q}}$ appartenente alla classe di congruenza $[\mathcal{Q}]$. Tenendo conto anche del fattore moltiplicativo globale dell'equazione, ogni classe di congruenza è univocamente determinata dal valore di alcuni parametri, che vengono detti moduli dell'ipersuperficie.

11.3. Classificazione delle curve coniche

Classifichiamo tutte le classi di congruenza di curve coniche nel piano euclideo $\mathbb{E}^2$, utilizzando gli invarianti metrici ed il Teorema 11.10.

Teorema 11.13. *In $\mathbb{E}^2$ l'equazione di ogni curva conica $\mathcal{Q}$ può essere ridotta ad una ed una sola delle forme canoniche elencate nella seguente tabella:*

$r(C)$	$r(A)$	I_3	I_2	$I_1 I_3$	*Forma canonica*	*Moduli*	*Nome*
3	2	$\neq 0$	> 0	< 0	$\frac{\tilde{x}^2}{\alpha^2} + \frac{\tilde{y}^2}{\beta^2} = 1$	$\alpha, \beta > 0$	*Ellisse con punti reali*
3	2	$\neq 0$	> 0	> 0	$\frac{\tilde{x}^2}{\alpha^2} + \frac{\tilde{y}^2}{\beta^2} = -1$	$\alpha, \beta > 0$	*Ellisse privo di punti reali*
3	2	$\neq 0$	< 0		$\frac{\tilde{x}^2}{\alpha^2} - \frac{\tilde{y}^2}{\beta^2} = 1$	$\alpha, \beta > 0$	*Iperbole*
3	1	$\neq 0$	$= 0$		$\tilde{x}^2 - 2\rho\tilde{y} = 0$	$\rho > 0$	*Parabola*
2	2	$= 0$	> 0	$= 0$	$\frac{\tilde{x}^2}{\alpha^2} + \tilde{y}^2 = 0$	$\alpha > 0$	*Coppia di rette incidenti con un unico punto reale*
2	2	$= 0$	< 0	$= 0$	$\frac{\tilde{x}^2}{\alpha^2} - \tilde{y}^2 = 0$	$\alpha > 0$	*Coppia di rette incidenti con infiniti punti reali*
2	1	$= 0$	$= 0$	$= 0$	$\frac{\tilde{x}^2}{\alpha^2} = 1$	$\alpha > 0$	*Coppia di rette parallele con infiniti punti reali*
2	1	$= 0$	$= 0$	$= 0$	$\frac{\tilde{x}^2}{\alpha^2} = -1$	$\alpha > 0$	*Coppia di rette parallele prive di punti reali*
1	1	$= 0$	$= 0$	$= 0$	$\tilde{x}^2 = 0$		*Retta doppia*

DIMOSTRAZIONE. Il teorema segue direttamente dal Teorema 11.10 e dal calcolo degli invarianti eseguito sulla rappresentazione canonica di Q. Ad esempio, studiamo il primo caso elencato nella tabella. Se $r(C) = 3$ e $r(A) = 2$ sappiamo che l'equazione della conica si può ridurre a

$$\lambda_1 \tilde{x}^2 + \lambda_2 \tilde{y}^2 + \tilde{c} = 0.$$

Segue immediatamente che $I_3 = \lambda_1 \lambda_2 \tilde{c} \neq 0$ ed $I_2 = \lambda_1 \lambda_2 \neq 0$.

Supponiamo inoltre che $I_2 > 0$. Allora i due autovalori di A hanno lo stesso segno. Di conseguenza la disuguaglianza $I_1 I_3 = (\lambda_1 + \lambda_2)\lambda_1 \lambda_2 \tilde{c} < 0$ implica che il segno di $\tilde{c}$ è discorde da quello degli autovalori. Definiamo le quantità strettamente positive

$$\alpha = \sqrt{-\frac{\tilde{c}}{\lambda_1}} \qquad \text{e} \qquad \beta = \sqrt{-\frac{\tilde{c}}{\lambda_2}}\,.$$

Moltiplicando l'equazione della conica per $-\tilde{c}^{-1}$ si ottiene

$$\frac{\tilde{x}^2}{\alpha^2} + \frac{\tilde{y}^2}{\beta^2} = 1.$$

Gli altri casi si ricavano in maniera analoga. □

Presentiamo le proprietà delle principali tipologie di curve coniche attraverso alcuni esempi.

Esempio 11.14. Consideriamo la curva conica $\mathcal{Q}$ di equazione

$$x^2 + 3y^2 - 2y - 1 = 0$$

rispetto ad un sistema di riferimento cartesiano $\mathcal{B}_O$. Le matrici associate a $\mathcal{Q}$ sono

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \qquad \text{e} \qquad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & -1 & -1 \end{bmatrix}.$$

Gli invarianti sono $I_1 = 4,\ I_2 = 3,\ I_3 = -4$. Grazie al Teorema 11.13 deduciamo che $\mathcal{Q}$ è un'ellisse con punti reali. Gli autovalori di A sono $\lambda_1 = 1$ e $\lambda_2 = 3$, mentre per il termine noto osserviamo che

$$\tilde{c} = \frac{I_3}{I_2} = -\frac{4}{3}.$$

Quindi una rappresentazione canonica di $\mathcal{Q}$ è

$$\tilde{x}^2 + 3\tilde{y}^2 - \frac{4}{3} = 0.$$

Dal calcolo dei moduli otteniamo

$$\alpha = \frac{2}{\sqrt{3}}, \quad \beta = \frac{1}{3} \qquad \Rightarrow \qquad \frac{\tilde{x}^2}{\left(\frac{2}{\sqrt{3}}\right)^2} + \frac{\tilde{y}^2}{\left(\frac{1}{3}\right)^2} = 1.$$

Calcoliamo il cambiamento di coordinate $X = Q * \tilde{X} + T$ necessario a ridurre $\mathcal{Q}$ in forma canonica. Dato che A è già una matrice diagonale, abbiamo $Q = \mathbb{I}_2$. La matrice T è la soluzione di $A * T = -B$, quindi:

$$[A| - B] = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 3 & 1 \end{array}\right] \qquad \Rightarrow \qquad \begin{cases} t_{11} = 0 \\ t_{21} = \frac{1}{3} \end{cases} \qquad \Rightarrow \qquad \begin{cases} x = \tilde{x} \\ y = \tilde{y} + \frac{1}{3} \end{cases}.$$

Elenchiamo alcune proprietà geometriche dell'ellisse $\mathcal{Q}$.

- L'ellisse è una curva con due assi di simmetria r_1, r_2 ed un centro di simmetria C. Nel sistema di riferimento canonico, r_1, r_2 coincidono con gli assi coordinati $\tilde{x}, \tilde{y}$ mentre C è l'origine $\tilde{O}$. Nel sistema di riferimento di partenza il centro ha coordinate

$$C|_{\mathcal{B}_O} = Q * C|_{\tilde{\mathcal{B}}_O} + T = T = \begin{bmatrix} 0 \\ \frac{1}{3} \end{bmatrix}.$$

  Dati $\mathbf{v_1}, \mathbf{v_2}$ gli autovettori di A, gli assi di simmetria hanno rappresentazione parametrica

$$r_1|_{\mathcal{B}_O} :\ P|_{\mathcal{B}_O} = C|_{\mathcal{B}_O} + t \cdot \mathbf{v_1}|_{\mathcal{B}} = \begin{bmatrix} 0 \\ \frac{1}{3} \end{bmatrix} + t \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

$$r_2|_{\mathcal{B}_O} :\ P|_{\mathcal{B}_O} = C|_{\mathcal{B}_O} + t \cdot \mathbf{v_2}|_{\mathcal{B}} = \begin{bmatrix} 0 \\ \frac{1}{3} \end{bmatrix} + t \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

- Esistono altri punti di $\mathbb{E}^2$ che sono speciali rispetto all'ellisse, ovvero i vertici ed i fuochi. I vertici sono i punti di intersezione dell'ellisse con gli assi di simmetria. Nel sistema di riferimento canonico hanno coordinate

$$V_1|_{\tilde{\mathcal{B}}_O} = \begin{bmatrix} \alpha \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{3}} \\ 0 \end{bmatrix}, \qquad V_2|_{\tilde{\mathcal{B}}_O} = \begin{bmatrix} -\alpha \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{2}{\sqrt{3}} \\ 0 \end{bmatrix},$$

$$V_3|_{\tilde{\mathcal{B}}_O} = \begin{bmatrix} 0 \\ \beta \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{3} \end{bmatrix}, \qquad V_4|_{\tilde{\mathcal{B}}_O} = \begin{bmatrix} 0 \\ -\beta \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{3} \end{bmatrix}.$$

L'ellisse è interamente contenuta all'interno della regione limitata $[-\alpha, \alpha] \times [-\beta, \beta] \subset \mathbb{R}^2$. Se supponiamo che $\alpha \geq \beta$ (come nel nostro esempio), definiamo inoltre $\varphi = \sqrt{\alpha^2 - \beta^2} = \sqrt{11}/3$. I fuochi dell'ellisse sono i punti di coordinate

$$F_1|_{\tilde{\mathcal{B}}_O} = \begin{bmatrix} \varphi \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{11}}{3} \\ 0 \end{bmatrix}, \qquad F_2|_{\tilde{\mathcal{B}}_O} = \begin{bmatrix} -\varphi \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{\sqrt{11}}{3} \\ 0 \end{bmatrix}.$$

Attraverso calcoli algebrici elementari si può facilmente verificare che l'ellisse è caratterizzabile come il luogo dei punti P del piano euclideo che soddisfano la seguente equazione:

$$d(P, F_1) + d(P, F_2) = 2\alpha.$$

Nel caso particolare $\alpha = \beta$ allora $F_1 = F_2 = C$, quindi $\mathcal{Q}$ è una circonferenza di raggio α.

Esempio 11.15. Consideriamo la curva conica $\mathcal{Q}$ di equazione

$$3x^2 - 3y^2 - 8xy + 2x + 4y - 2 = 0$$

rispetto ad un sistema di riferimento cartesiano $\mathcal{B}_O$. Le matrici associate a $\mathcal{Q}$ sono

$$A = \begin{bmatrix} 3 & -4 \\ -4 & -3 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \text{e} \qquad C = \begin{bmatrix} 3 & -4 & 1 \\ -4 & -3 & 2 \\ 1 & 2 & -2 \end{bmatrix}.$$

Gli invarianti sono $I_1 = 0,\ I_2 = -25,\ I_3 = 25$. Grazie al Teorema 11.13 deduciamo che $\mathcal{Q}$ è un'iperbole. Dato che $I_1 = \lambda_1 + \lambda_2$ e $I_2 = \lambda_1\lambda_2$, segue che gli autovalori di A sono $\lambda_1 = 5$ e $\lambda_2 = -5$. Il termine noto è invece uguale a

$$\tilde{c} = \frac{I_3}{I_2} = -1.$$

Quindi una rappresentazione canonica di $\mathcal{Q}$ è

$$5\tilde{x}^2 - 5\tilde{y}^2 - 1 = 0.$$

In questo caso i moduli sono definiti come

$$\alpha = \sqrt{-\frac{\tilde{c}}{\lambda_1}} = \frac{1}{\sqrt{5}}, \quad \beta = \sqrt{\frac{\tilde{c}}{\lambda_2}} = \frac{1}{\sqrt{5}} \qquad \Rightarrow \qquad \frac{\tilde{x}^2}{\left(\frac{1}{\sqrt{5}}\right)^2} - \frac{\tilde{y}^2}{\left(\frac{1}{\sqrt{5}}\right)^2} = 1.$$

Calcoliamo il cambiamento di coordinate $X = Q * \tilde{X} + T$ necessario a ridurre $\mathcal{Q}$ in forma canonica. Gli autospazi di A sono

$$V_5 = \ker(A - 5 \cdot \mathbb{I}_2) = \ker\left(\begin{bmatrix} -2 & -4 \\ -4 & -8 \end{bmatrix}\right) = \ker\left(\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}\right) = \mathcal{L}\left(\begin{bmatrix} -2 \\ 1 \end{bmatrix}\right)$$

e

$$V_{-5} = \ker(A + 5 \cdot \mathbb{I}_2) = \ker\left(\begin{bmatrix} 8 & -4 \\ -4 & 2 \end{bmatrix}\right) = \ker\left(\begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix}\right) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right).$$

Di conseguenza la matrice che diagonalizza A è

$$Q = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \in SO(2; \mathbb{R}).$$

La matrice T è la soluzione di $A * T = -B$, quindi:

$$[A|-B] = \left[\begin{array}{cc|c} 3 & -4 & -1 \\ -4 & -3 & -2 \end{array}\right] \xrightarrow{3R_2 + 4R_1 \to R_2} \left[\begin{array}{cc|c} 3 & -4 & -1 \\ 0 & -25 & -10 \end{array}\right] \qquad \Rightarrow \qquad \begin{cases} t_{11} = \frac{1}{5} \\ t_{21} = \frac{2}{5} \end{cases}$$

$$\Rightarrow \qquad \begin{cases} x = \frac{2}{\sqrt{5}}\,\tilde{x} + \frac{1}{\sqrt{5}}\,\tilde{y} + \frac{1}{5} \\ y = -\frac{1}{\sqrt{5}}\,\tilde{x} + \frac{2}{\sqrt{5}}\,\tilde{y} + \frac{2}{5} \end{cases} .$$

Elenchiamo alcune proprietà geometriche dell'iperbole $\mathcal{Q}$.

- Come per l'ellisse, anche l'iperbole è una curva con due assi di simmetria r_1, r_2 ed un centro di simmetria C. Nel sistema di riferimento canonico, r_1, r_2 coincidono con gli assi coordinati $\tilde{x}, \tilde{y}$ mentre C è l'origine $\tilde{O}$. Nel sistema di riferimento di partenza il centro ha coordinate

$$C|_{\mathcal{B}_O} = Q * C|_{\tilde{\mathcal{B}}_O} + T = T = \begin{bmatrix} \frac{1}{5} \\ \frac{2}{5} \end{bmatrix},$$

  mentre gli assi di simmetria hanno rappresentazione parametrica

$$r_1|_{\mathcal{B}_O} : \; P|_{\mathcal{B}_O} = C|_{\mathcal{B}_O} + t \cdot \mathbf{v_1}|_{\mathcal{B}} = \begin{bmatrix} \frac{1}{5} \\ \frac{2}{5} \end{bmatrix} + t \cdot \begin{bmatrix} -2 \\ 1 \end{bmatrix},$$

$$r_2|_{\mathcal{B}_O} : \; P|_{\mathcal{B}_O} = C|_{\mathcal{B}_O} + t \cdot \mathbf{v_2}|_{\mathcal{B}} = \begin{bmatrix} \frac{1}{5} \\ \frac{2}{5} \end{bmatrix} + t \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

- L'iperbole ha due vertici e due fuochi. I vertici sono i punti di intersezione dell'ellisse con gli assi di simmetria. Nel sistema di riferimento canonico hanno coordinate

$$V_1|_{\tilde{\mathcal{B}}_O} = \begin{bmatrix} \alpha \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ 0 \end{bmatrix}, \qquad V_2|_{\tilde{\mathcal{B}}_O} = \begin{bmatrix} -\alpha \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{5}} \\ 0 \end{bmatrix}.$$

  L'iperbole è contenuta nell'unione di due strisce non limitate del piano $([-\infty, -\alpha] \cup [\alpha, \infty]) \times \mathbb{R} \subset \mathbb{R}^2$. Definiamo inoltre $\varphi = \sqrt{\alpha^2 + \beta^2} = \sqrt{2/5}$. I fuochi dell'iperbole sono i punti di coordinate

$$F_1|_{\tilde{\mathcal{B}}_O} = \begin{bmatrix} \varphi \\ 0 \end{bmatrix} = \begin{bmatrix} \sqrt{\frac{2}{5}} \\ 0 \end{bmatrix}, \qquad F_2|_{\tilde{\mathcal{B}}_O} = \begin{bmatrix} -\varphi \\ 0 \end{bmatrix} = \begin{bmatrix} -\sqrt{\frac{2}{5}} \\ 0 \end{bmatrix}.$$

  L'iperbole è caratterizzabile come il luogo dei punti P del piano euclideo che soddisfano la seguente equazione:

$$|d(P, F_1) - d(P, F_2)| = 2\alpha.$$

- Esistono altre due rette speciali $r_\pm$ per l'iperbole, dette asintoti. Nel sistema di riferimento canonico questi hanno equazione

$$r_\pm|_{\tilde{\mathcal{B}}_O} : \; \beta\tilde{x} \pm \alpha\tilde{y} = 0 \qquad \Rightarrow \qquad \tilde{x} \pm \tilde{y} = 0.$$

  In geometria proiettiva, gli asintoti sono caratterizzati come le due rette che intersecano tangenzialmente $\mathcal{Q}$ nei due punti all'infinito della curva. Osserviamo che nel nostro esempio i due asintoti sono le bisetrici dei quadranti del sistema di riferimento canonico. In questo caso si dice che l'iperbole è equilatera. Ciò accade se e solo se $\alpha = \beta$, ovvero se e solo se $\lambda_1 = -\lambda_2$ e quindi $I_1 = 0$.

Esempio 11.16. Consideriamo la curva conica $\mathcal{Q}$ di equazione

$$x^2 + y^2 + 2xy + 4\sqrt{2}x + 2 = 0$$

rispetto ad un sistema di riferimento cartesiano $\mathcal{B}_O$. Le matrici associate a $\mathcal{Q}$ sono

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix} \qquad \text{e} \qquad C = \begin{bmatrix} 1 & 1 & 2\sqrt{2} \\ 1 & 1 & 0 \\ 2\sqrt{2} & 0 & 2 \end{bmatrix}.$$

Gli invarianti sono $I_1 = 2,\ I_2 = 0,\ I_3 = -8$. Grazie al Teorema 11.13 deduciamo che $\mathcal{Q}$ è una parabola. Dato che $I_1 = \lambda_1 + \lambda_2$ ed $I_2 = \lambda_1\lambda_2$, segue che gli autovalori di A sono $\lambda_1 = 2$ e $\lambda_2 = 0$. Il coefficiente di primo grado è invece

$$p = \sqrt{-\frac{I_3}{I_1}} = 2.$$

Quindi una rappresentazione canonica di $\mathcal{Q}$ è

$$2\tilde{x}^2 + 4\tilde{y} = 0.$$

In questo caso l'unico modulo è

$$\rho = \frac{p}{|\lambda_1|} = 1 \qquad \Rightarrow \qquad \tilde{x}^2 - 2\tilde{y} = 0.$$

Notiamo che il segno del termine lineare può sempre essere scelto negativo, poiché abbiamo la libertà di eventualmente ruotare di un angolo π gli assi $\tilde{x}, \tilde{y}$, quando definiamo $\tilde{\mathcal{B}}_O$.

Calcoliamo il cambiamento di coordinate $X = Q * \tilde{X} + T$ necessario a ridurre $\mathcal{Q}$ in forma canonica, seguendo il procedimento utilizzato nella dimostrazione del Teorema 11.13. Gli autospazi di A sono

$$V_2 = \ker(A - 2 \cdot \mathbb{I}_2) = \ker\left(\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\right) = \ker\left(\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}\right) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$$

e

$$V_0 = \ker(A) = \ker\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right) = \ker\left(\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\right) = \mathcal{L}\left(\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right).$$

Di conseguenza la matrice Q_1 che diagonalizza A è

$$Q_1 = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \in SO(2;\mathbb{R}) \qquad \Rightarrow$$

$$A_1 = Q_1^T * A * Q_1 = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \qquad B_1 = Q_1^T * B = \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \qquad c_1 = c = 2.$$

La matrice colonna T_2 è

$$T_2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \qquad \Rightarrow$$

$$A_2 = A_1 = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \qquad B_2 = A_1 * T_2 + B_1 = \begin{bmatrix} 0 \\ -2 \end{bmatrix}, \qquad c_2 = -\frac{(B_1)_{11}^2}{\lambda_1} + c_1 = 0.$$

Essendo già arrivati alla forma canonica, la procedura si interrompe qui. La trasformazione completa di coordinate è quindi generata da

$$Q = Q_1 = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \qquad \text{e} \qquad T = Q_1 * T_2 = \begin{bmatrix} -1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix},$$

ovvero

$$\begin{cases} x = \frac{1}{\sqrt{2}}\tilde{x} - \frac{1}{\sqrt{2}}\tilde{y} - \frac{1}{\sqrt{2}} \\ y = \frac{1}{\sqrt{2}}\tilde{x} + \frac{1}{\sqrt{2}}\tilde{y} - \frac{1}{\sqrt{2}} \end{cases}.$$

Elenchiamo alcune proprietà geometriche della parabola $\mathcal{Q}$.

- Diversamente dall'ellisse e dall'iperbole, la parabola ha un solo asse di simmetria r. Nel sistema di riferimento canonico, r è l'asse coordinato $\tilde{y}$. Inoltre, non esiste un centro di simmetria di $\mathcal{Q}$. In questo caso l'origine $\tilde{O}$ del sistema di riferimento canonico viene posta in corrispondenza del vertice V della parabola. Nel sistema di riferimento di partenza, il vertice ha coordinate

$$V|_{\mathcal{B}_O} = Q * V|_{\tilde{\mathcal{B}}_O} + T = T = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix},$$

mentre l'asse di simmetria ha rappresentazione parametrica

$$r|_{\mathcal{B}_O}: \; P|_{\mathcal{B}_O} = V|_{\mathcal{B}_O} + t \cdot \mathbf{v_2}|_{\mathcal{B}} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} + t \cdot \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Osserviamo che la parabola è interamente contenuta nel semipiano $\tilde{y} \geq 0$.
- La parabola ha un solo fuoco F ed un asse direttore s. Nel sistema di riferimento canonico sono descritti da

$$F|_{\tilde{\mathcal{B}}_O} = \begin{bmatrix} 0 \\ \frac{\rho}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{2} \end{bmatrix}, \qquad s|_{\tilde{\mathcal{B}}_O}: \; \tilde{y} + \frac{\rho}{2} = 0.$$

La parabola è caratterizzabile come il luogo dei punti P del piano euclideo che soddisfano la seguente equazione:

$$d(P, F) = d(P, s).$$

In generale, dati un punto F ed una retta s in $\mathbb{E}^2$, consideriamo il luogo $\mathcal{Q}$ dei punti P che soddisfano l'equazione

$$d(P, F) = \varepsilon \, d(P, s),$$

dove $0 < \varepsilon < +\infty$. É possibile dimostrare che:

- se $0 < \varepsilon < 1$ allora $\mathcal{Q}$ è un'ellisse, e tutte le ellissi possono essere definite in tal modo;
- se $\varepsilon = 1$ allora $\mathcal{Q}$ è una parabola e tutte le parabole possono essere definite in tal modo;
- se $1 < \varepsilon < +\infty$ allora $\mathcal{Q}$ è un'iperbole, e tutte le iperboli possono essere definite in tal modo.

Il parametro ε viene detto eccentricità della conica $\mathcal{Q}$.

Esempio 11.17. Consideriamo infine la curva conica $\mathcal{Q}$ di equazione

$$2x^2 - y^2 - xy + 2x + y = 0$$

rispetto ad un sistema di riferimento cartesiano $\mathcal{B}_O$. Le matrici associate a $\mathcal{Q}$ sono

$$A = \begin{bmatrix} 2 & -1/2 \\ -1/2 & -1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 \\ 1/2 \end{bmatrix} \qquad \text{e} \qquad C = \begin{bmatrix} 2 & -1/2 & 1 \\ -1/2 & -1 & 1/2 \\ 1 & 1/2 & 0 \end{bmatrix}.$$

Gli invarianti sono $I_1 = 1$, $I_2 = -9/4$, $I_3 = 0$. Quindi $r(C) \leq 3$, nel qual caso la conica si dice degenere. Grazie al Teorema 11.13 deduciamo che $\mathcal{Q}$ è una coppia di rette incidenti con infiniti punti reali. Lasciamo come esercizio al lettore la costruzione dell'equazione e del sistema di riferimento canonico. Osserviamo qui invece che il polinomio quadratico che definisce $\mathcal{Q}$ può essere fattorizzato in due polinomi lineari:

$$2x^2 - y^2 - xy + 2x + y = (x - y + 1)(2x + y).$$

Di conseguenza l'insieme $\mathcal{Q}$ è l'unione delle due rette

$$r|_{\mathcal{B}_O}: \; x - y + 1 = 0 \qquad \text{e} \qquad s|_{\mathcal{B}_O}: \; 2x + y = 0,$$

che hanno intersezione nel punto

$$C|_{\mathcal{B}_O} = \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix}.$$

Il punto C è centro di simmetria della conica e corrisponde all'origine del sistema di riferimento canonico. Gli assi di simmetria corrispondono alle rette bisetrici degli angoli formati da r ed s, e corrispondono agli assi coordinati nel sistema di riferimento canonico.

11.4. Classificazione delle superfici quadriche

Analogamente alle curve coniche, classifichiamo tutte le classi di congruenza di superfici quadriche in $\mathbb{E}^3$. Anche il seguente teorema è una diretta conseguenza del Teorema 11.10 e della definizione degli invarianti metrici.

Teorema 11.18. *In $\mathbb{E}^3$ l'equazione di ogni superficie quadrica $\mathcal{Q}$ può essere ridotta ad una ed una sola delle forme canoniche elencate nella seguente tabella:*

$r(C)$	$r(A)$	I_4	I_3	I_2	$I_1 I_3$	*Forma canonica*	*Moduli*	*Nome*
4	3	<0	$\neq 0$	>0	>0	$\frac{\tilde{x}^2}{\alpha^2}+\frac{\tilde{y}^2}{\beta^2}+\frac{\tilde{z}^2}{\gamma^2}=1$	$\alpha,\beta,\gamma>0$	*Ellissoide con punti reali*
4	3	>0	$\neq 0$	>0	>0	$\frac{\tilde{x}^2}{\alpha^2}+\frac{\tilde{y}^2}{\beta^2}+\frac{\tilde{z}^2}{\gamma^2}=-1$	$\alpha,\beta,\gamma>0$	*Ellissoide privo di punti reali*
4	3	>0	$\neq 0$	≤ 0		$\frac{\tilde{x}^2}{\alpha^2}+\frac{\tilde{y}^2}{\beta^2}-\frac{\tilde{z}^2}{\gamma^2}=1$	$\alpha,\beta,\gamma>0$	*Iperboloide ad una falda od iperbolico*
4	3	>0	$\neq 0$		≤ 0			
4	3	<0	$\neq 0$	≤ 0		$\frac{\tilde{x}^2}{\alpha^2}-\frac{\tilde{y}^2}{\beta^2}-\frac{\tilde{z}^2}{\gamma^2}=1$	$\alpha,\beta,\gamma>0$	*Iperboloide a due falde od ellittico*
4	3	<0	$\neq 0$		≤ 0			
4	2	<0	$=0$	>0	$=0$	$\frac{\tilde{x}^2}{\alpha^2}+\tilde{y}^2-2\rho\tilde{z}=0$	$\alpha,\rho>0$	*Paraboloide ellittico*
4	2	>0	$=0$	<0	$=0$	$\frac{\tilde{x}^2}{\alpha^2}-\tilde{y}^2-2\rho\tilde{z}=0$	$\alpha,\rho>0$	*Paraboloide iperbolico*
3	3	$=0$	$\neq 0$	>0	>0	$\frac{\tilde{x}^2}{\alpha^2}+\frac{\tilde{y}^2}{\beta^2}+\tilde{z}^2=0$	$\alpha,\beta>0$	*Cono con un unico punto reale*
3	3	$=0$	$\neq 0$	≤ 0		$\frac{\tilde{x}^2}{\alpha^2}+\frac{\tilde{y}^2}{\beta^2}-\tilde{z}^2=0$	$\alpha,\beta>0$	*Cono con infiniti punti reali*
3	3	$=0$	$\neq 0$		≤ 0			
3	2	$=0$	$=0$	>0	$=0$	$\frac{\tilde{x}^2}{\alpha^2}+\frac{\tilde{y}^2}{\beta^2}=1$	$\alpha,\beta>0$	*Cilindro ellittico con punti reali*
3	2	$=0$	$=0$	>0	$=0$	$\frac{\tilde{x}^2}{\alpha^2}+\frac{\tilde{y}^2}{\beta^2}=-1$	$\alpha,\beta>0$	*Cilindro ellittico privo di punti reali*
3	2	$=0$	$=0$	<0	$=0$	$\frac{\tilde{x}^2}{\alpha^2}-\frac{\tilde{y}^2}{\beta^2}=1$	$\alpha,\beta>0$	*Cilindro iperbolico*
3	1	$=0$	$=0$	$=0$	$=0$	$\tilde{x}^2-2\rho\tilde{y}=0$	$\rho>0$	*Cilindro parabolico*
2	2	$=0$	$=0$	>0	$=0$	$\frac{\tilde{x}^2}{\alpha^2}+\tilde{y}^2=0$	$\alpha>0$	*Coppia di piani incidenti con una retta di punti reali*
2	2	$=0$	$=0$	<0	$=0$	$\frac{\tilde{x}^2}{\alpha^2}-\tilde{y}^2=0$	$\alpha>0$	*Coppia di piani incidenti con infiniti punti reali*
2	1	$=0$	$=0$	$=0$	$=0$	$\frac{\tilde{x}^2}{\alpha^2}=1$	$\alpha>0$	*Coppia di piani paralleli con infiniti punti reali*
2	1	$=0$	$=0$	$=0$	$=0$	$\frac{\tilde{x}^2}{\alpha^2}=-1$	$\alpha>0$	*Coppia di piani paralleli privi di punti reali*
1	1	$=0$	$=0$	$=0$	$=0$	$\tilde{x}^2=0$		*Piano doppio*

Esempio 11.19. Consideriamo la superficie quadrica $\mathcal{Q}$ di equazione

$$3x^2 + 2y^2 + 3z^2 + 2xz - 4 = 0$$

rispetto ad un sistema di riferimento cartesiano $\mathcal{B}_O$. Le matrici associate a $\mathcal{Q}$ sono

$$A = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad \text{e} \qquad C = \begin{bmatrix} 3 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 1 & 0 & 3 & 0 \\ 0 & 0 & 0 & -4 \end{bmatrix}.$$

Gli invarianti sono $I_1 = 8,\ I_2 = 20,\ I_3 = 16,\ I_4 = -64$. Grazie al Teorema 11.18 deduciamo che $\mathcal{Q}$ è un'ellissoide con punti reali. Gli autovalori di A sono le radici di

$$P_A(\lambda) = (2-\lambda)((3-\lambda)^2 - 1) = (2-\lambda)^2(4-\lambda),$$

pertanto $\lambda_1 = \lambda_2 = 2$ e $\lambda_3 = 4$. Per il termine noto osserviamo che

$$\tilde{c} = \frac{I_4}{I_3} = -4.$$

Quindi una rappresentazione canonica di $\mathcal{Q}$ è

$$2\tilde{x}^2 + 2\tilde{y}^2 + 4\tilde{z}^2 - 4 = 0.$$

Dal calcolo dei moduli otteniamo

$$\alpha = \sqrt{-\frac{\tilde{c}}{\lambda_1}} = \sqrt{2}, \quad \beta = \sqrt{-\frac{\tilde{c}}{\lambda_2}} = \sqrt{2}, \quad \gamma = \sqrt{-\frac{\tilde{c}}{\lambda_3}} = 1 \qquad \Rightarrow \qquad \frac{\tilde{x}^2}{\left(\sqrt{2}\right)^2} + \frac{\tilde{y}^2}{\left(\sqrt{2}\right)^2} + \tilde{z}^2 = 1.$$

Calcoliamo il cambiamento di coordinate $X = Q * \tilde{X} + T$ necessario a ridurre $\mathcal{Q}$ in forma canonica. Gli autospazi di A sono:

$$V_2 = \ker(A - 2\cdot\mathbb{I}_3) = \ker\left(\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}\right) = \ker\left(\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right),$$

$$V_4 = \ker(A - 4\cdot\mathbb{I}_3) = \ker\left(\begin{bmatrix} -1 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & -1 \end{bmatrix}\right) = \ker\left(\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right).$$

Di conseguenza la matrice che diagonalizza A è

$$Q = \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} \in SO(3;\mathbb{R}).$$

La matrice T è la soluzione di $A * T = -B$, quindi:

$$[A|-B] = \left[\begin{array}{ccc|c} 3 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 1 & 0 & 3 & 0 \end{array}\right] \xrightarrow[R_1 - 3R_3 \to R_3]{R_3 \to R_1} \left[\begin{array}{ccc|c} 1 & 0 & 3 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -8 & 0 \end{array}\right] \qquad \Rightarrow \qquad \begin{cases} t_{11} = 0 \\ t_{21} = 0 \\ t_{31} = 0 \end{cases}.$$

Questo risultato è coerente con il fatto che i termini lineari dell'equazione di $\mathcal{Q}$ sono già nulli nel sistema di riferimento di partenza. Quindi otteniamo:

$$\begin{cases} x = \frac{1}{\sqrt{2}}\,\tilde{x} + \frac{1}{\sqrt{2}}\,\tilde{z} \\ y = \tilde{y} \\ z = -\frac{1}{\sqrt{2}}\,\tilde{x} + \frac{1}{\sqrt{2}}\,\tilde{z} \end{cases}.$$

Elenchiamo alcune proprietà geometriche dell'ellissoide $\mathcal{Q}$.

- Analogamente al caso bidimensionale dell'ellisse, in generale un'ellissoide è una superficie con tre assi di simmetria r_1, r_2, r_3 ed un centro di simmetria C. Ogni coppia di assi genera un piano di simmetria. Nel sistema di riferimento canonico, r_1, r_2, r_3 coincidono con gli assi coordinati $\tilde{x}, \tilde{y}, \tilde{z}$, i piani di simmetria sono i piani coordinati mentre C è l'origine $\tilde{O}$. Nel sistema di riferimento di partenza il centro ha coordinate

$$C|_{\mathcal{B}_O} = Q * C|_{\tilde{\mathcal{B}}_O} + T = T = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

  Gli assi di simmetria contengono il centro C e sono paralleli agli autovettori $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$ di A. L'ellissoide ha 8 vertici definiti come i punti di intersezione della superficie con gli assi di simmetria. L'ellissoide è interamente contenuto all'interno della regione limitata $[-a, a] \times [-b, b] \times [-c, c] \subset \mathbb{R}^3$. Nel nostro caso l'autospazio V_2 ha dimensione 2. Geometricamente questo significa che tutte le rette contenenti il centro e parallele a V_2 sono assi di simmetria dell'ellissoide. In questa situazione si dice che la superficie è una quadrica di rotazione, concetto che approfondiremo nella prossima sezione.

- Il risultato dell'intersezione tra un piano π e una superficie quadrica $\mathcal{Q}$ è una curva conica $\mathcal{Q}'$. Questo significa che se scegliamo un sistema di riferimento cartesiano $\mathcal{B}'_O$ nel piano π, l'insieme $\mathcal{Q}'$ è descritto da un'equazione quadratica. Ad esempio, scegliamo come π il piano coordinato xy. Allora

$$\mathcal{Q}'|_{\mathcal{B}_O} : \begin{cases} 3x^2 + 2y^2 + 3z^2 + 2xz - 4 = 0 \\ z = 0 \end{cases}$$

  Definiamo il sistema di riferimento cartesiano $\mathcal{B}'_O = \{O, \mathbf{e_1}, \mathbf{e_2}\}$ di π. Per ogni P in π abbiamo

$$P|_{\mathcal{B}_O} = O|_{\mathcal{B}_O} + t_1 \cdot \mathbf{e_1}|_{\mathcal{B}} + t_2 \cdot \mathbf{e_2}|_{\mathcal{B}} \qquad \Rightarrow \qquad \begin{cases} x = t_1 \\ y = t_2 \\ z = 0 \end{cases},$$

  da cui segue

$$\mathcal{Q}'|_{\mathcal{B}'_O} : \quad 3t_1^2 + 2t_2^2 - 4 = 0$$

  che evidentemente è un'ellisse. Prendiamo invece il piano π di giacitura V_2 e contenente il centro C. Se definiamo $\mathcal{B}'_O = \{C, \mathbf{v_1}, \mathbf{v_2}\}$, per ogni P in π abbiamo

$$P|_{\mathcal{B}'_O} = C|_{\mathcal{B}'_O} + t_1 \cdot \mathbf{v_1}|_{\mathcal{B}} + t_2 \cdot \mathbf{v_2}|_{\mathcal{B}} \qquad \Rightarrow \qquad \begin{cases} x = \frac{1}{\sqrt{2}} t_1 \\ y = t_2 \\ z = -\frac{1}{\sqrt{2}} t_1 \end{cases}.$$

  Allora

$$\mathcal{Q}'|_{\mathcal{B}'_O} : \quad 2t_1^2 + 2t_2^2 - 4 = 0,$$

  che è l'equazione di una circonferenza. Questo accade per ogni conica $\mathcal{Q}'$ ottenuta dall'intersezione di $\mathcal{Q}$ con un piano di giacitura V_2 ed è conseguenza del fatto che $\mathcal{Q}$ è una superficie di rotazione.

Esempio 11.20. Consideriamo la superficie quadrica $\mathcal{Q}$ di equazione

$$2x^2 - 2y^2 - 2z^2 - 2yz + 4x - 1 = 0$$

rispetto ad un sistema di riferimento cartesiano $\mathcal{B}_O$. Le matrici associate a $\mathcal{Q}$ sono

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & -1 \\ 0 & -1 & -2 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} \qquad \text{e} \qquad C = \begin{bmatrix} 2 & 0 & 0 & 2 \\ 0 & -2 & -1 & 0 \\ 0 & -1 & -2 & 0 \\ 2 & 0 & 0 & -1 \end{bmatrix}.$$

Gli invarianti sono $I_1 = -2,\ I_2 = -5,\ I_3 = 6,\ I_4 = -18$. Grazie al Teorema 11.18 deduciamo che $\mathcal{Q}$ è un iperboloide a due falde. Gli autovalori di A sono le radici di

$$P_A(\lambda) = (2-\lambda)((-2-\lambda)^2 - 1) = (2-\lambda)(-1-\lambda)(-3-\lambda),$$

pertanto $\lambda_1 = 2,\ \lambda_2 = -1$ e $\lambda_3 = -3$. Per il termine noto osserviamo che

$$\tilde{c} = \frac{I_4}{I_3} = -3.$$

Quindi una rappresentazione canonica di $\mathcal{Q}$ è

$$2\tilde{x}^2 - \tilde{y}^2 - 3\tilde{z}^2 - 3 = 0.$$

Dal calcolo dei moduli otteniamo

$$\alpha = \sqrt{-\frac{\tilde{c}}{\lambda_1}} = \sqrt{\frac{3}{2}}, \quad \beta = \sqrt{\frac{\tilde{c}}{\lambda_2}} = \sqrt{3}, \quad \gamma = \sqrt{\frac{\tilde{c}}{\lambda_3}} = 1 \qquad \Rightarrow \qquad \frac{\tilde{x}^2}{\left(\sqrt{\frac{3}{2}}\right)^2} - \frac{\tilde{y}^2}{\left(\sqrt{3}\right)^2} - \tilde{z}^2 = 1.$$

Calcoliamo il cambiamento di coordinate $X = Q * \tilde{X} + T$ necessario a ridurre $\mathcal{Q}$ in forma canonica. Gli autospazi di A sono:

$$V_2 = \ker(A - 2\cdot \mathbb{I}_3) = \ker\left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & -4 & -1 \\ 0 & -1 & -4 \end{bmatrix}\right) = \ker\left(\begin{bmatrix} 0 & 1 & 4 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right),$$

$$V_{-1} = \ker(A + \cdot \mathbb{I}_3) = \ker\left(\begin{bmatrix} 3 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{bmatrix}\right) = \ker\left(\begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \mathcal{L}\left(\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}\right),$$

$$V_{-3} = \ker(A + 3\cdot \mathbb{I}_3) = \ker\left(\begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}\right) = \ker\left(\begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \mathcal{L}\left(\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right).$$

Di conseguenza la matrice che diagonalizza A è

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \in SO(3;\mathbb{R}).$$

La matrice T è la soluzione di $A * T = -B$, quindi:

$$[A| - B] = \left[\begin{array}{ccc|c} 2 & 0 & 0 & -2 \\ 0 & -2 & -1 & 0 \\ 0 & -1 & -2 & 0 \end{array}\right] \xrightarrow[-R_3 \to R_2]{R_2 - 2R_3 \to R_3} \left[\begin{array}{ccc|c} 2 & 0 & 0 & -2 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & 0 \end{array}\right] \qquad \Rightarrow \qquad \begin{cases} t_{11} = -1 \\ t_{21} = 0 \\ t_{31} = 0 \end{cases}.$$

Quindi otteniamo:

$$\begin{cases} x = \tilde{x} - 1 \\ y = \frac{1}{\sqrt{2}}\,\tilde{y} + \frac{1}{\sqrt{2}}\,\tilde{z} \\ z = -\frac{1}{\sqrt{2}}\,\tilde{y} + \frac{1}{\sqrt{2}}\,\tilde{z} \end{cases}.$$

Tutti gli iperboloidi hanno tre assi di simmetria r_1, r_2, r_3, tre piani di simmetria ed un centro di simmetria C. Gli assi di simmetria sono paralleli agli autovettori $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$ di A. Nel sistema di riferimento canonico, r_1, r_2, r_3 coincidono con gli assi coordinati $\tilde{x}, \tilde{y}, \tilde{z}$, mentre C è l'origine $\tilde{O}$. L'iperboloide a due falde è composto da due parti disconnesse, simmetriche rispetto al piano $\tilde{x}\tilde{y}$ e contenute nei due semispazi $(-\infty, -\alpha] \times \mathbb{R}^2$ ed $[\alpha, +\infty) \times \mathbb{R}^2$. L'iperboloide ha 2 vertici definiti come i punti di intersezione dell'iperboloide con l'asse coordinato $\tilde{z}$.

Esempio 11.21. Infine consideriamo la superficie quadrica $\mathcal{Q}$ di equazione

$$x^2 + y^2 + z^2 + 2xy + 2xz + 2yz + 2y + 2z + 1 = 0.$$

La sua riduzione in forma canonica è già stata studiata nell'Esempio 11.11:

$$3\tilde{x}^2 + \frac{4}{\sqrt{6}}\tilde{y} = 0.$$

La superficie è un cilindro parabolico, che può essere pensato come l'unione delle infinite parabole date dall'intersezione di $\mathcal{Q}$ con i piani di equazione $\tilde{z} = k$, per ciascun $k \in \mathbb{R}$. Inoltre, per ogni punto di coordinate $(\tilde{x}_0, \tilde{y}_0, \tilde{z}_0)$ della superficie $\mathcal{Q}$ passa una retta interamente contenuta nella quadrica, avente rappresentazione algebrica $\tilde{x} = \tilde{x}_0,\ \tilde{y} = \tilde{y}_0$. Questa è una proprietà comune a tutte le quadriche degeneri caratterizzate da $r(C) = 3$ e dall'invariante $I_4 = 0$. Le proprietà di simmetria di $\mathcal{Q}$ si ricavano da quelle delle parabole di cui è costituito. In particolare, esiste un intero piano di simmetria di equazione $\tilde{x} = 0$, mentre la retta $\tilde{x} = \tilde{z} = 0$ è costituita dai vertici di tutte tali parabole.

11.5. Superfici di rotazione, coni e cilindri generalizzati

Concludiamo il capitolo con lo studio di alcune superfici particolari nello spazio euclideo tridimensionale.

Definizione 11.22. *In $\mathbb{E}^3$ consideriamo una retta r, un punto V ed una direzione U. Allora:*

i) una superficie di rotazione di asse r è un famiglia continua ad un parametro $\mathcal{S}$ di circonferenze, tutte giacenti su piani ortogonali all'asse ed i cui centri appartengono ad r;

ii) un cono generalizzato di vertice V è una famiglia continua ad un parametro $\mathcal{S}$ di rette tutte contenenti V;

iii) un cilindro generalizzato di direzione U è una famiglia continua ad un parametro $\mathcal{S}$ di rette tutte di giacitura U.

Noi qui ci presentiamo esempi di tali superfici che siano anche quadriche. Osserviamo che le superfici di rotazioni prendono il nome dal fatto che esse possono essere costruite attraverso la rotazione di una curva rispetto all'asse di rotazione.

Esempio 11.23. In $\mathbb{E}^3$ consideriamo le due rette

$$r : \begin{cases} x = 0 \\ y - z = 0 \end{cases}, \qquad \text{e} \qquad s_h : \begin{cases} x + y - 2h = 0 \\ y - h = 0 \end{cases},$$

dove h è un parametro reale. Costruiamo la superficie $\mathcal{S}_h$ ottenuta dalla rotazione di s_h rispetto ad r. A tal fine, ci serve una rappresentazione parametrica di r, da cui otteniamo un punto $Q \in r$ ed un vettore $\mathbf{v}$ parallelo alla retta:

$$r : \begin{cases} x = 0 \\ y = t \\ z = t \end{cases} \qquad \Rightarrow \qquad Q = (0, 0, 0), \quad \mathbf{v} = (0, 1, 1).$$

Per definizione, $\mathcal{S}_h$ è l'unione di tutte le circonferenze caratterizzate da:

- avere come centro un punto di r;
- giacere su un piano perpendicolare a $\mathbf{v}$;
- intersecare s_h.

Dato $P_0 = (x_0, y_0, z_0)$ un generico punto di s_h, le coordinate del generico punto $P = (x, y, z)$ di $\mathcal{S}_h$ devono soddisfare il seguente sistema di equazioni:

$$\begin{cases} d(P,Q) = d(P_0,Q) \\ \langle \mathbf{v}, \overrightarrow{P_0P} \rangle = 0 \\ x_0 + y_0 - 2h = 0 \\ y_0 - h = 0 \end{cases} \quad \Rightarrow \quad \begin{cases} x^2 + y^2 + z^2 = x_0^2 + y_0^2 + z_0^2 \\ y - y_0 + z - z_0 = 0 \\ x_0 + y_0 - 2h = 0 \\ y_0 - h = 0 \end{cases}.$$

Le prime due equazioni del sistema definiscono la circonferenza contenuta in $\mathcal{S}_h$ e passante per il punto P_0, ottenuta come l'intersezione della sfera di centro Q e contenente P_0 e del piano ortogonale a $\mathbf{v}$ ed anch'esso contenente P_0. Esprimendo x_0, y_0, z_0 in funzione di x, y, z otteniamo:

$$\begin{cases} x_0 = h \\ y_0 = h \\ z_0 = y + z - h \\ x^2 + y^2 + z^2 = 3h^2 + y^2 + z^2 + 2yz - 2hy - 2hz \end{cases}.$$

Di conseguenza la superficie $\mathcal{S}_h$ ha equazione

$$x^2 - 2yz + 2hy + 2hz - 3h^2 = 0,$$

da cui si deduce che $\mathcal{S}_h$ è una quadrica. Le matrici associate sono

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ h \\ h \end{bmatrix} \qquad \text{e} \qquad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & h \\ 0 & -1 & 0 & h \\ 0 & h & h & -3h^2 \end{bmatrix}.$$

Gli invarianti sono $I_1 = 1,\ I_2 = -1,\ I_3 = 1,\ I_4 = h^2$. Grazie al Teorema 11.18 deduciamo che $\mathcal{S}_h$ è un iperboloide ad una falda se $h \neq 0$, mentre è un cono reale se $h = 0$.

Analizziamo alcune proprietà di queste superfici.

- L'equazione canonica di un iperboloide ad una falda è

$$\frac{\tilde{x}^2}{\alpha^2} + \frac{\tilde{y}^2}{\beta^2} - \frac{\tilde{z}^2}{\gamma^2} = 1.$$

  Diversamente dall'iperboloide a due falde, esso è formato da un'unica componente connessa ed ha 4 vertici, definiti come i punti di intersezione dell'iperboloide con gli assi $\tilde{x}$ e $\tilde{y}$. Le superfici $\mathcal{S}_h$ con $h \neq 0$ sono iperboloidi di rotazione, per i quali $\alpha = \beta$. Questo significa che la quadrica è simmetrica rispetto allo scambio delle coordinate canoniche $\tilde{x}$ e $\tilde{y}$ e di conseguenza ogni asse nel piano $\tilde{x}\tilde{y}$ è di simmetria. Nel caso $h = 0$ abbiamo invece un cono. Questo è un altro esempio di quadrica degenere. Il vertice del cono coincide con il centro di simmetria della quadrica e con l'origine del sistema di riferimento canonico.

- La classificazione di $\mathcal{S}_h$ è strettamente legata alla mutua posizione delle due rette. Infatti, analizziamo il sistema lineare associato all'intersezione $r \cap s_h$, che è descritto dalle matrici

$$[A|B] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 1 & 0 & 2h \\ 0 & 1 & 0 & h \end{array}\right] \xrightarrow[R_4-R_2 \to R_4]{R_3-R_1-R_2 \to R_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 2h \\ 0 & 0 & 1 & h \end{array}\right] \xrightarrow{R_4-R_3 \to R_4} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 2h \\ 0 & 0 & 0 & -h \end{array}\right].$$

Segue che:
  - $r(A) = 3 < 4 = r([A|B])$ se $h \neq 0$, nel qual caso le rette sono sghembe;
  - $r(A) = 3 = r([A|B])$ se $h = 0$, nel qual caso le rette sono incidenti. Il vertice del cono coincide con il punto di intersezione di $r \cap s_0$.
- Per costruzione, l'iperboloide è costituito da infinite rette. Lasciamo al lettore verificare che $\mathcal{S}_h$ si può ottenere anche dalla rotazione della retta

$$\tilde{s}_h : \begin{cases} x + y - 2h = 0 \\ y - 2z - h = 0 \end{cases}$$

rispetto ad r. Quindi per ogni punto di $\mathcal{S}_h$ passano due rette distinte. Questa proprietà caratterizza tutti gli iperboloidi ed i paraboloidi iperbolici, che vengono detti superfici quadriche doppiamente rigate. Se $h = 0$ le due famiglie di rette coincidono. Infatti il cono è l'unione di un'unica famiglia di rette, tutte contenenti il vertice. In questo caso si parla di quadrica unicamente rigata. Tutte le quadriche degeneri con $r(C) = 3$ sono unicamente rigate.
- Il *piano polare* di una quadrica $\mathcal{Q}$ rispetto al punto $Q = (\tilde{x}, \tilde{y}, \tilde{z})$ di $\mathbb{E}^3$ è definito come il piano di equazione

$$\begin{bmatrix} \tilde{x} & \tilde{y} & \tilde{z} & 1 \end{bmatrix} * C * \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = 0.$$

Se Q appartiene alla superficie quadrica, allora il piano polare coincide con il *piano tangente* a $\mathcal{Q}$ in Q, ovvero è l'unico piano che interseca la superficie $\mathcal{Q}$ lungo una conica degenere contenente il punto Q. Ad esempio, consideriamo il punto $Q_h = (h, h, 0)$, che appartiene a $\mathcal{S}_h$ per ogni $h \in \mathbb{R}$. Il piano tangente Π_h ad $\mathcal{S}_h$ nel punto Q_h ha equazione

$$\begin{bmatrix} h & h & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & h \\ 0 & -1 & 0 & h \\ 0 & h & h & -3h^2 \end{bmatrix} * \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = h(x + y - 2h) = 0.$$

L'intersezione tra Π_h ed $\mathcal{S}_h$ è descritta dal sistema

$$\begin{cases} x^2 - 2yz + 2hy + 2hz - 3h^2 = 0 \\ x + y - 2h = 0 \end{cases} \Rightarrow \begin{cases} (y - h)(y - 2z - h) = 0 \\ x + y - 2h = 0 \end{cases},$$

che ha come soluzione la conica degenere costituita dall'unione delle due rette

$$r_1 : \begin{cases} x + y - 2h = 0 \\ y - h = 0 \end{cases} \quad \text{e} \quad r_2 : \begin{cases} x + y - 2h = 0 \\ y - 2z - h = 0 \end{cases}.$$

Osserviamo infine che per $h = 0$ il punto Q_0 coincide con il vertice del cono. Come è intuitivo aspettarsi, non esiste il piano tangente al cono nel suo vertice (il vertice si dice *punto singolare* della superficie.)

Esempio 11.24. Consideriamo la superficie quadrica $\mathcal{Q}$ di equazione

$$x_0^2 + y_0^2 + 4x_0y_0 - 2x_0z_0 - 2y_0 - 1 = 0.$$

Il punto $Q = (0, 0, 0)$ non appartiene alla superficie. Il piano polare Π di $\mathcal{Q}$ rispetto a Q ha equazione

$$\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 2 & -1 & 0 \\ 2 & 1 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 \end{bmatrix} * \begin{bmatrix} x_0 \\ y_0 \\ z_0 \\ 1 \end{bmatrix} = -y_0 - 1 = 0\ .$$

È possibile dimostrare che tutte le rette tangenti a $\mathcal{Q}$ e contenenti Q passano per i punti dell'intersezione $\mathcal{Q} \cap \Pi$. L'unione di tali rette tangenti forma il *cono tangente* $\mathcal{S}$ alla quadrica $\mathcal{Q}$ e di vertice Q. Dato $P_0 = (x_0, y_0, z_0)$ un generico punto di $\mathcal{Q} \cap \Pi$, le coordinate del generico punto $P = (x, y, z)$ del cono tangente devono soddisfare il seguente sistema di equazioni:

$$\begin{cases} \overrightarrow{QP} = t \cdot \overrightarrow{QP_0} \\ x_0^2 + y_0^2 + 4x_0y_0 - 2x_0z_0 - 2y_0 - 1 = 0 \\ y_0 + 1 = 0 \end{cases} \Rightarrow \begin{cases} x = tx_0 \\ y = ty_0 \\ z = tz_0 \\ y_0 = -1 \\ x_0^2 + y_0^2 + 4x_0y_0 - 2x_0z_0 - 2y_0 - 1 = 0 \end{cases},$$

da cui otteniamo

$$\begin{cases} t = y \\ x_0 = \dfrac{x}{y} \\ y_0 = 1 \\ z_0 = \dfrac{z}{y} \\ x^2 + 2y^2 + 4xy - 2xz = 0 \end{cases}.$$

Quindi il cono tangente $\mathcal{S}$ ha equazione

$$x^2 + 2y^2 + 4xy - 2xz = 0,$$

che lasciamo verificare al lettore essere un cono quadrico reale.

Esempio 11.25. Consideriamo la curva conica $\mathcal{Q}$ di equazioni

$$\begin{cases} x_0^2 - y_0^2 - 2x_0y_0 + 2x_0 - 4z_0 = 0 \\ x_0 - y_0 - z_0 \end{cases}$$

ed il piano $\Pi :\ x - 1 = 0$. Il vettore perpendicolare a Π è $\mathbf{n} = (1, 0, 0)$. La proiezione ortogonale di $\mathcal{Q}$ su Π si può ottenere come l'intersezione di tale piano con il cilindro $\mathcal{S}$ costituito dalle rette di $\mathbb{E}^3$ intersecanti $\mathcal{Q}$ e parallele ad $\mathbf{n}$. Dato $P_0 = (x_0, y_0, z_0)$ un generico punto di $\mathcal{Q}$, le coordinate del generico punto $P = (x, y, z)$ di $\mathcal{S}$ devono soddisfare il seguente sistema di equazioni:

$$\begin{cases} \overrightarrow{P_0P} = t \cdot \mathbf{n} \\ x_0^2 - y_0^2 - 2x_0y_0 + 2x_0 - 4z_0 = 0 \\ x_0 - y_0 - z_0 = 0 \end{cases} \Rightarrow \begin{cases} x - x_0 = t \\ y - y_0 = 0 \\ z - z_0 = 0 \\ x_0 - y_0 - z_0 = 0 \\ x_0^2 - y_0^2 - 2x_0y_0 + 2x_0 - 4z_0 = 0 \end{cases},$$

da cui otteniamo

$$\begin{cases} x_0 = x - t = y + z \\ y_0 = y \\ z_0 = z \\ t = x - y - z \\ -2y^2 + z^2 + 2y - 2z = 0 \end{cases} .$$

Quindi la proiezione di $\mathcal{Q}$ su Π è la curva conica descritta dal sistema

$$\begin{cases} 2y^2 - z^2 - 2y + 2z = 0 \\ x - 1 = 0 \end{cases} .$$