Algebra Lineare e Geometria

Kieran G. O'Grady

"Sapienza"Università di Roma

12 ottobre 2023

Indice

Capitolo 0

Introduzione

Questi appunti sono stati scritti per il corso di Algebra Lineare (più o meno fino al Capitolo 5 o al Capitolo 6) e per il corso di Geometria I. Sono incompleti e vanno rivisti, questo vale soprattutto per quanto riguarda l'ultimo capitolo.

Capitolo 1

Preliminari

1.1 Insiemi

Intuitivamente un insieme è una collezione di oggetti, per esempio l'insieme I degli italiani o l'insieme A degli australiani. Gli oggetti che appartengono a un insieme sono gli *elementi* dell'insieme, per esempio Gianni Rivera è un elemento di I e non è un elemento di A, Rod Laver è un elemento di A ma non di I. La notazione

$$X := \{a, b, \dots, z\} \tag{1.1.1}$$

significa che definiamo l'insieme X come quello i cui elementi sono $a, b, \dots, z$. Per esempio potremmo porre $X := \{0, 6, 4, 2, 8, 10\}$; in parole X è l'insieme dei numeri naturali pari non maggiori di 10. Nella (1.1.1) il simbolo := sta a significare che il simbolo di sinistra denota l'espressione a destra[1], le parentesi graffe "delimitano" l'insieme.

Principio dell'estensione 1.1.1. *Un insieme è caratterizzato dagli elementi che gli appartengono ovvero, se X, Y sono insiemi, allora X è uguale a Y (in simboli $X = Y$) se e solo se X ha gli stessi elementi di Y.*

L'affermazione contenuta nel principio di estensione è ovvia (se avete capito di cosa stiamo parlando) e vi chiederete perchè mai debba essere enfatizzata; il motivo è che fa parte degli assiomi della teoria degli insiemi. Sia X un insieme e x un oggetto: la notazione $x \in X$ sigifica che x è un elemento di X e $x \notin X$ significa che x non è un elemento di X. Dato un insieme X e una proprietà P (per esempio l'insieme degli immatricolati alla Sapienza e la proprietà di essere maschi) si definisce l'insieme Y degli elementi $x \in X$ che hanno la proprietà P: in simboli

$$Y := \{x \in X \mid x \text{ ha la proprietà } P\}. \tag{1.1.2}$$

(Nell'esempio considerato Y è l'insieme dei maschi immatricolati alla Sapienza). Nella (1.1.2) la sbarra verticale | si può leggere "tale che". Noi considereremo insiemi i cui elementi sono numeri o altri oggetti matematici. Esistono notazioni standard per alcuni di questi insiemi:

1. $\mathbb{N}$ è l'insieme dei numeri naturali: i suoi elementi sono $0, 1, 2, \dots$ cioè i numeri che conoscete dall'infanzia (con l'aggiunta dello 0).

2. $\mathbb{Z}$ è l'insieme dei numeri interi: i suoi elementi sono $0, \pm 1, \pm 2, \dots$.

3. $\mathbb{Q}$ è l'insieme dei numeri razionali: un numero razionale è determinato da una coppia di interi p, q con $q \neq 0$ (il numero p/q) e si ha $p/q = p'/q'$ se e solo se $pq' - p'q = 0$.

[1] Una equaglianza del tipo $6 = 2 \cdot 3$ o $10 = 3 \cdot 3$ è un'affermazione che può essere vera (la prima) o falsa (la seconda) mentre (1.1.1) è una definizione - non ha senso chiedersi se sia vera o falsa.

4. $\mathbb{R}$ è l'insieme dei numeri reali: la costruzione dei numeri reali non è elementare, la diamo per acquisita, ci limitiamo a menzionare che un numero reale è individuato da un decimale infinito, per esempio $1,01001000100001...$, $2,39999...$ o $-3,121314151....$ (Attenzione: $2,39999...$ è uguale a $2,40000....$ che scriviamo $2,4$.)

5. Dati $a, b \in \mathbb{R}$ con $a \leqslant b$ si definiscono i seguenti sottoinsiemi di $\mathbb{R}$:

$$[a,b]:=\{x\in\mathbb{R}|a\leqslant x\leqslant b\},\quad (a,b):=\{x\in\mathbb{R}|a<x<b\},\quad [a,b):=\{x\in\mathbb{R}|a\leqslant x<b\},\quad (a,b]:=\{x\in\mathbb{R}|a<x\leqslant b\}. \tag{1.1.3}$$

   Il primo è l'*intervallo chiuso* di estremi a, b, il secondo è l'*intervallo aperto* di estremi a, b e così via. Dato $a \in \mathbb{R}$ definiamo i seguenti sottoinsiemi di $\mathbb{R}$:

$$[a,+\infty):=\{x\in\mathbb{R}|a\leqslant x\},\quad (a,+\infty):=\{x\in\mathbb{R}|a<x\},\quad (-\infty,a]:=\{x\in\mathbb{R}|x\leqslant a\},\quad (-\infty,a):=\{x\in\mathbb{R}|x<a\}. \tag{1.1.4}$$

   (Sono semirette (chiuse o aperte) di estremo a.)

6. Dato $a \in \mathbb{Z}$ (cioè a è un numero intero) definiamo

$$(a) := \{x \in \mathbb{Z} \mid x = na \text{ per un qualche } n \in \mathbb{Z}\}. \tag{1.1.5}$$

   In parole: (a) è l'insieme dei multipli (interi) di a.

Definizione 1.1.2. Un insieme X è contenuto nell'insieme Y (equivalentemente X è un *sottoinsieme* di Y) se ogni elemento di X è anche elemento di Y cioè per ogni $x \in X$ vale $x \in Y$: in simboli $X \subset Y$ (o anche $Y \supset X$). La notazione $X \not\subset Y$ (o $Y \not\supset X$) significa che X **non** è contenuto in Y cioè che esiste $x \in X$ tale che $x \notin Y$.

Esempio 1.1.3. Siccome un multiplo di 6 è anche un multiplo di 3 abbiamo $(6) \subset (3)$. D'altra parte $3 \in (3)$ ma $3 \notin (6)$ e quindi $(3) \not\subset (6)$.

Osservazione 1.1.4. Siano X, Y insiemi. Per il principio di estensione $X = Y$ se e solo se $X \subset Y$ e $Y \subset X$.

L'osservazione fatta è banale ma è utile tenerne conto quando si vuole decidere se due insiemi sono uguali: grazie all'Osservazione 1.1.4 si tratta di decidere se $X \subset Y$ *e* $Y \subset X$. Dati insiemi X, Y possiamo produrre altri insiemi a partire da X e Y.

Definizione 1.1.5. L'*unione* di X e Y è l'insieme i cui elementi sono gli x tali che $x \in X$ **o** $x \in Y$. (Attenzione: x può appartenere sia ad X che a Y.) L'unione di X e Y si denota $X \cup Y$. L'*intersezione* di X e Y è l'insieme i cui elementi sono gli x tali che $x \in X$ **e** $x \in Y$. L'intersezione di X e Y si denota $X \cap Y$.

Alcuni esempi:

$$(2) \cup \{x \in \mathbb{Z} \mid x \text{ è dispari}\} = \mathbb{Z}, \qquad (2) \cap (3) = (6), \qquad (4) \cap (6) = (12).$$

Cosa succede se consideriamo l'intersezione dell'insieme $P := (2)$ dei numeri interi pari e D l'insieme dei numeri interi dispari? Non ci sono elementi x di P e di D. Quindi se vogliamo che abbia senso l'intersezione $P \cap D$ dobbiamo accettare che ci sia un insieme che non ha elementi: questo è l'insieme vuoto, si denota $\varnothing$. Per ogni insieme X abbiamo

$$\varnothing \cup X = X, \qquad \varnothing \cap X = \varnothing.$$

L'unione e l'intersezione hanno senso anche per una famiglia arbitraria di insiemi X_i dove i è un elemento arbitrario in un insieme di indici I.

Definizione 1.1.6. L'unione $\bigcup_{i\in I} X_i$ è l'insieme i cui elementi sono gli x tali che $x \in X_i$ per **un qualche** $i \in I$, l'intersezione $\bigcap_{i\in I} X_i$ è l'insieme i cui elementi sono gli x tali che $x \in X_i$ per **tutti gli** $i \in I$.

Un esempio:

$$\bigcup_{i\in\mathbb{N}}(i) = \mathbb{Z}, \qquad \bigcap_{i\in\mathbb{N}}(i) = \{0\}.$$

Definizione 1.1.7. Siano X, Y insiemi. L'*insieme differenza* $X\backslash Y$ è

$$X\backslash Y := \{x \in X \mid x \notin Y\}.$$

Per esempio abbiamo

$$\mathbb{Z}\backslash(2) = \{x \in \mathbb{Z} \mid x \text{ è dispari}\}, \quad X\backslash\varnothing = X, \quad \varnothing\backslash X = \varnothing.$$

Definizione 1.1.8. Siano $X_1, \ldots, X_n$ insiemi. Il *prodotto cartesiano* $X_1 \times \ldots \times X_n$ è l'insieme i cui elementi sono le n-ple **ordinate** $(x_1, x_2, \ldots, x_n)$ dove $x_i \in X_i$ per $i = 1, 2, \ldots, n$. Se $X_1 = X_2 = \ldots = X_n$ denotiamo $X_1 \times \ldots \times X_n$ con X^n.

Un esempio: $\mathbb{R}^n$ è l'insieme delle n-ple ordinate di numeri reali (notazione familiare?).

1.2 Applicazioni

Siano X, Y insiemi.

Definizione 1.2.1. Un'*applicazione* (o *funzione*) da X a Y è una legge f che associa a ogni $x \in X$ un $y \in Y$ che denotiamo $f(x)$: in simboli $f\colon X \to Y$ o $X \xrightarrow{f} Y$. L'insieme X è il *dominio* della applicazione f e l'insieme Y è il suo *codominio*.

Un chiarimento riguardo la definizione di applicazione: si intende che due applicazioni $f_1\colon X_1 \to Y_1$ e $f_2\colon X_2 \to Y_2$ sono uguali se e solo se

1. $X_1 = X_2$,
2. $Y_1 = Y_2$,
3. per ogni $x \in X_1 = X_2$ si ha che $f_1(x) = f_2(x)$.

Un altro modo di vedere un'applicazione $f\colon X \to Y$ è come una procedura che a partire dall'input x produce l'output $f(x)$. Un esempio: X è l'insieme degli immatricolati alla Sapienza, Y è l'insieme dei numeri naturali e f associa a ogni immatricolato il suo anno di nascita. Esempi matematici:

$$\begin{array}{ccc} \mathbb{Z} & \xrightarrow{f} & \mathbb{Z} \\ x & \mapsto & x-5 \end{array} \qquad \begin{array}{ccc} \mathbb{R}\times\mathbb{R} & \xrightarrow{g} & \mathbb{R} \\ (a,b) & \mapsto & ab \end{array}$$

Se X è un insieme, l'applicazione *identità* da X a X è quella che associa a x se stesso; la denotiamo Id_X oppure 1_X. Quindi

$$\mathrm{Id}(x) = 1_X(x) = x \quad \forall x \in X. \tag{1.2.1}$$

(Il simbolo $\forall$ significa "per ogni".) Un'applicazione $f\colon X \to Y$ è *costante* se

$$f(x_1) = f(x_2) \qquad \forall x_1, x_2 \in X. \tag{1.2.2}$$

Dati insiemi X, Y si denota con Y^X l'insieme i cui elementi sono le applicazioni $f\colon X \to Y$ (notate l'inversione nella notazione):

$$Y^X := \{f\colon X \to Y\}. \tag{1.2.3}$$

Data un'applicazione $f\colon X \to Y$ il *grafico* di f è il sottoinsieme di Γ_f di $X \times Y$ i cui elementi sono le coppie $(x, f(x))$ per x un arbitrario elemento di X. Notate che se $X = Y = \mathbb{R}$ e associamo a ogni $(x, y) \in \mathbb{R}^2$ il punto del piano di coordinate cartesiane (x, y) (relative a un sistema di riferimento scelto) il grafico così definito corrisponde al grafico considerato a scuola. Sia $\Gamma_f \subset X \times Y$ il grafico di un'applicazione $f\colon X \to Y$; dato $x \in X$ esiste un unico elemento di Γ_f la cui prima entrata sia x (cioè uguale a $(x, *)$).

Osservazione 1.2.2. Si può dare una formulazione matematicamente precisa di applicazione $f\colon X \to Y$ evitando di fare appello al concetto di "legge che associa..." definendo un'applicazione come un sottoinsieme $\Gamma \subset X \times Y$ che ha la proprietà dei grafici appena menzionata - lasciamo i dettagli al lettore.

Supponiamo che $f\colon X \to Y$ e $g\colon Y \to Z$ siano applicazioni (notate: il codominio di f è il dominio di g). Allora possiamo definire un'applicazione da X a Z associando a $x \in X$ l'elemento $g(f(x))$ di Z: questa è la *composizione* di f e g che si denota $g \circ f$ (attenzione all'ordine - in generale $f \circ g$ non avrà senso perchè X non sarà uguale a Z). Ricapitolando

$$g \circ f(x) := g(f(x)). \tag{1.2.4}$$

Un esempio: siano $X = Y = Z$ l'insieme delle persone (viventi o morte), f l'applicazione che associa a una persona suo padre e g l'applicazione che associa a una persona sua madre. La composizione $f \circ f$ è l'applicazione che associa a una persona il nonno paterno, mentre $g \circ f$ è l'applicazione che associa a una persona la nonna paterna. Notiamo che se $f\colon X \to Y$ abbiamo

$$f \circ 1_X = 1_Y \circ f = f. \tag{1.2.5}$$

Questo giustifica la notazione 1_X per l'applicazione identità: se pensiamo alla composizione di applicazioni come analogo della moltiplicazione tra numeri vediamo che l'applicazione identità ha proprietà analoghe a quelle del numero 1. Supponiamo che $f\colon X \to Y$, $g\colon Y \to W$ e $h\colon W \to Z$ siano applicazioni: hanno senso sia $(h \circ g) \circ f$ che $h \circ (g \circ f)$ e sono entrambe applicazioni da X a Z. Abbiamo che

$$((h \circ g) \circ f)(x) = h(g(f(x))) = (h \circ (g \circ f))(x)$$

e quindi la composizione di applicazioni gode della proprietà di associatività:

$$(h \circ g) \circ f = h \circ (g \circ f). \tag{1.2.6}$$

Sia $f\colon X \to Y$ un'applicazione. Siano $A \subset X$ e $B \subset Y$. Definiamo i sottoinsiemi $f(A) \subset Y$ (l'*immagine* di A) e $f^{-1}B \subset X$ (la *controimmagine* di A) così:

$$f(A) := \{y_0 \in Y \mid \exists x_0 \in X \text{ tale che } f(x_0) = y_0\}, \qquad f^{-1}(B) := \{x_0 \in X \mid f(x_0) \in B\}. \tag{1.2.7}$$

L'*immagine* di f è $\operatorname{im} f := f(X)$. Un esempio: se $f\colon \mathbb{R} \to \mathbb{R}$ è l'applicazione quadrato, cioè $f(x) = x^2$, allora $f([1,2]) = [1,4]$, $f^{-1}([1,4]) = [1,2] \cup [-2,-1]$ e l'immagine di f è l'insieme dei reali non-negativi. Se $B = \{y_0\}$ cioè è un insieme con un solo elemento denotiamo $f^{-1}\{y_0\}$ con $f^{-1}y_0$.

Definizione 1.2.3. L'applicazione $f\colon X \to Y$ è *suriettiva* se $f(X) = Y$, è *iniettiva* se dato $y \in Y$ esiste al più un $x \in X$ tale che $f(x) = y$, è bijettiva (o biunivoca) se è iniettiva e suriettiva.

Un esempio: siano $f, g, h\colon \mathbb{R} \to \mathbb{R}$ le applicazioni definite da

$$f(x) = x^2 + 1, \qquad g(x) = x^3, \qquad h(x) = x^3 - x. \tag{1.2.8}$$

La f non è nè iniettiva nè suriettiva, la g è biunivoca, la h è suriettiva ma non iniettiva. Notate che nella definizione di applicazione dominio e codominio fanno parte dei dati che definiscono un'applicazione: quindi un'applicazione $f\colon X \to Y$ che non è suriettiva può essere "resa" suriettiva sostituendo al codominio Y il codominio $f(X)$ (il punto è che a rigor di definizione la "nuova" f non è uguale alla "vecchia" f). Nell'esempio (1.2.8) la f diventa suriettiva se la sostituiamo con l'applicazione $F\colon \mathbb{R} \to \{x \in \mathbb{R} \mid x \geqslant 1\}$ data dalla stessa formula cioè $F(x) = x^2 + 1$.

Definizione 1.2.4. Sia $f\colon X \to Y$ un'applicazione, e sia $A \subset X$ un sottoinsieme del dominio di f. La *restrizione di f ad A* è l'applicazione

$$\begin{array}{ccc} A & \xrightarrow{f_{|A}} & Y \\ a & \mapsto & f(a) \end{array}$$

Definizione 1.2.5. Siano X un insieme e $A \subset X$ un sottoinsieme. L'*inclusione* $\iota\colon A \hookrightarrow X$ è l'applicazione

$$\begin{array}{ccc} A & \xrightarrow{\iota} & X \\ a & \mapsto & a \end{array}$$

Esempio 1.2.6. Sia $f\colon X \to Y$ un'applicazione. Si può scrivere f come composizione di un'applicazione suriettiva e di una iniettiva come segue. Sia $Z := f(X)$, e sia $h\colon X \to Z$ definita da

$$\begin{array}{ccc} X & \xrightarrow{g} & Z \\ x & \mapsto & f(x) \end{array}$$

Quindi g è sostanzialmente la funzione f, però il codominio è stato ristretto. Sia $\iota\colon Z \hookrightarrow Y$ l'inclusione. L'applicazione g è suriettiva, la ι è iniettiva, e si ha $f = \iota \circ g$.

Definizione 1.2.7. Sia $f\colon X \to Y$ un'applicazione **biunivoca**. L'*applicazione inversa* $f^{-1}\colon Y \to X$ associa a $y \in Y$ l'unico $x \in X$ tale che $f(x) = y$.

Notate che la definizione di inversa di f ha senso solo se f è biunivoca. Si ha che

$$f \circ f^{-1} = f^{-1} \circ f = 1_X. \tag{1.2.9}$$

Esempio: delle tre applicazioni f, g, h definite in (1.2.8) l'unica a essere biunivoca è g quindi ha senso g^{-1} (e non hanno senso nè f^{-1} nè h^{-1}) e chiaramente $g^{-1}(y) = y^{1/3}$. Supponiamo che $f\colon X \to Y$ sia biunivoca e che $B \subset Y$: allora $f^{-1}B = f^{-1}(B)$ dove $f^{-1}B$ è dato da (1.2.7). Fate attenzione alla notazione: se f non è biunivoca l'applicazione f^{-1} non è definita, ha senso solo $f^{-1}B$ per $B \subset Y$.

Sia $f\colon X \to X$. Se m è un numero naturale positivo, si pone

$$f^m := \underbrace{f \circ f \circ \ldots \circ f}_{m}. \tag{1.2.10}$$

Se m, n sono numeri naturali positivi vale

$$f^m \circ f^n = f^{m+n}. \tag{1.2.11}$$

Si pone $f^0 := 1_X$. Con questa notazione l'uguaglianza in (1.2.11) vale per ogni $m, n \in \mathbb{N}$. Ora supponiamo che $f\colon X \to X$ sia invertibile. Allora ha senso f^m per ogni $m \in \mathbb{Z}$: infatti se m è un intero negativo si pone

$$f^m := (f^{-1} \circ f^{-1} \circ \ldots \circ f^{-1})^{-m}. \tag{1.2.12}$$

Con questa definizione (per f invertibile) l'uguaglianza in (1.2.11) vale per ogni $m, n \in \mathbb{Z}$.

Siano X, Y insiemi. Diciamo che X ha la stessa cardinalià di Y se esiste un'applicazione *biunivoca* $f\colon X \to Y$ - in simboli $X \approx Y$. Se X, Y sono insiemi finiti questo equivale a dire che X, Y hanno lo stesso numero di elementi. In questo caso il numero di elementi di X (e di Y) viene denotato $|X|$.

Se esiste un'applicazione *suriettiva* $f\colon X \to Y$ diciamo che X ha cardinalità maggiore o uguale a quella di Y - in simboli $X \succeq Y$ (o che Y ha cardinalià minore o uguale a quella di X - in simboli $Y \preceq X$). Notate che se X non è finito esistono applicazioni suriettive $f\colon X \to X$ ma non iniettive: per esempio $f\colon \mathbb{N} \to \mathbb{N}$ definita da

$$f(x) := \begin{cases} 0 & \text{se } x = 0, \\ x - 1 & \text{se } x > 0. \end{cases}$$

Supponiamo che X, Y siano insiemi finiti. Allora $X \succeq Y$ equivale alla diseguaglianza $|X| \geqslant |Y|$ (o $|Y| \leqslant |X|$), e inoltre vale $X \succeq Y$ *e* $X \preceq Y$ se e solo se $|X| = |Y|$, cioè se X e Y hanno la stessa cardinalità. Il risultato non banale enunciato sotto afferma che l'analogo vale per insiemi qualsiasi (per una dimostrazione vedete l'Appendice 2 di [3]).

Teorema 1.2.8 (Teorema di Schröder-Bernstein)**.** *Siano X, Y insiemi tali che $X \succeq Y$ e $Y \succeq X$. Allora $X \approx Y$.*

1.3 Relazioni

Sia X un insieme. Una *relazione* tra gli elementi di X (o una relazione su X) è un sottoinsieme $\mathcal{R} \subset X \times X$. Dati $x_1, x_2 \in X$ diciamo che $x_1 \mathcal{R} x_2$ se la coppia ordinata (x_1, x_2) è un elemento di $\mathcal{R}$.

Esempio 1.3.1. Sia $\mathcal{R} \subset \mathbb{R} \times \mathbb{R}$ il sottoinsieme degli (x, y) tali che $x - y \geqslant 0$. La relazione $\mathcal{R}$ è quella di "essere non più piccolo" e anzichè $x\mathcal{R}y$ scriviamo $x \geqslant y$.

Esempio 1.3.2. Sia X un insieme e sia $\mathcal{P}(X)$ l'insieme i cui elementi sono i sottoinsiemi di X. Sia $\mathcal{R} \subset \mathcal{P}(X) \times \mathcal{P}(X)$ il sottoinsieme delle coppie (A, B) tali che $A \subset B$. La relazione $\mathcal{R}$ è quella d'inclusione e anzichè $A\mathcal{R}B$ scriviamo $A \subset B$.

Esempio 1.3.3. Dato $n \in \mathbb{Z}$ sia $\mathcal{R}_n \subset \mathbb{Z} \times \mathbb{Z}$ il sottoinsieme degli (x, y) tali che $x - y \in (n)$ ovvero $(x - y)$ è un multiplo di n. Si usa scrivere $x \equiv y \pmod{n}$ anzichè $x\mathcal{R}_n y$: si legge "x è congruo a y modulo n".

Osservazione 1.3.4. Siano $x, y \in \mathbb{N}$: allora x è congruo a y modulo 10 se e solo se l'ultima cifra nello sviluppo decimale di x è uguale all'ultima cifra nello sviluppo decimale di y.

Esistono due tipi di relazione particolarmente importanti, quelle di ordine e di equivalenza.

Definizione 1.3.5. Una relazione $\mathcal{R}$ sull'insieme X è di *ordine* se

1. per ogni $x \in X$ vale $x\mathcal{R}x$ (proprietà riflessiva),
2. se $x\mathcal{R}y$ e $y\mathcal{R}x$ allora $x = y$ (antisimmetria),
3. se $x\mathcal{R}y$ e $y\mathcal{R}z$ allora $x\mathcal{R}z$ (proprietà transitiva).

È di *ordine totale* se in aggiunta per ogni $x, y \in X$ vale almeno una tra $x\mathcal{R}y$ e $y\mathcal{R}x$.

Le relazioni dell'Esempio 1.3.1 e 1.3.2 sono di ordine, la prima è di ordine totale, la seconda non è di ordine totale a meno che X sia vuoto o costituito di un singolo elemento. La relazione dell'Esempio 1.3.3 non è di ordine (quale delle tre proprietà della Definizione 1.3.5 non vale?). Notate che anche la relazione $\mathcal{R}$ su $\mathbb{R}$ definita da $x\mathcal{R}y$ se $x \leqslant y$ è di ordine.

Definizione 1.3.6. Una relazione $\mathcal{R}$ sull'insieme X è di *equivalenza* se

1. per ogni $x \in X$ vale $x\mathcal{R}x$ (proprietà riflessiva),
2. se $x\mathcal{R}y$ allora $y\mathcal{R}x$ (simmetria),
3. se $x\mathcal{R}y$ e $y\mathcal{R}z$ allora $x\mathcal{R}z$ (proprietà transitiva).

La relazione dell'Esempio 1.3.3 è di equivalenza, quella dell'Esempio 1.3.1 non lo è. Spesso una relazione di equivalenza su X si denota con "$\sim$"cioè si scrive $x_1 \sim x_2$ anzichè $x_1 \mathcal{R} x_2$. A partire dalla relazione di equivalenza $\sim$ si costruisce un insieme i cui elementi sono sottoinsiemi di X. Dato $x_0 \in X$ la *classe di* $\sim$*-equivalenza* di x_0 è

$$[x_0] := \{x \in X \mid x \sim x_0\}. \tag{1.3.1}$$

Quando non ci sono possibilità di equivoci chiamiamo $[x_0]$ la classe di equivalenza di x_0 (omettiamo il riferimento a $\sim$): si denota anche $\overline{x}_0$. Si dice che x_0 è un *rappresentante* della classe di equivalenza $[x_0]$. Un esempio: consideriamo la relazione su $\mathbb{Z}$ della congruenza modulo 2 - vedi l'Esempio1.3.3 - allora esistono due classi di equivalenza, il sottoinsieme degli interi pari e quello degli interi dispari.

Definizione 1.3.7. Sia X un insieme e $\sim$ una relazione di equivalenza su X. L'*insieme quoziente*, denotato $X/\sim$, è quello i cui elementi sono le classi di $\sim$-equivalenza. L'*applicazione quoziente* è la

$$\begin{array}{ccc} X & \xrightarrow{\pi} & X/\sim \\ x & \mapsto & [x] \end{array} \tag{1.3.2}$$

Esempio 1.3.8. Nell'esempio della congruenza modulo n - vedi l'Esempio 1.3.3 - l'insieme delle classi di equivalenza ha n elementi e cioè $[0], [1], \ldots, [n-1]$: il quoziente $\mathbb{Z}/\mathcal{R}_n$ si denota $\mathbb{Z}/(n)$.

Le classi di equivalenza di una data relazione (di equivalenza) su X hanno la proprietà di costituire una partizione di X, dove il significato di partizione è dato dalla seguente definizione.

Definizione 1.3.9. Sia X un insieme. Una *partizione* di X è una famiglia $\{X_i\}_{i\in I}$ di sottoinsiemi di X tale che

1. $\bigcup_{i\in I} X_i = X$,
2. se $i_1 \neq i_2 \in I$ allora $X_{i_1} \cap X_{i_2} = \varnothing$.

Proposizione 1.3.10. *Sia X un insieme e $\sim$ una relazione di equivalenza su X. La famiglia delle classi di $\sim$-equivalenza è una partizione di X. Viceversa data una partizione $\{X_i\}_{i\in I}$ di X esiste una e una sola relazione di equiavlenza le cui classi di equivalenza sono gli X_i.*

Dimostrazione. Verifichiamo che le classi di $\sim$-equivalenza soddisfano (1) e (2) della Definizione 1.3.9. Sia $x \in X$: siccome $x \sim x$ abbiamo $x \in [x]$ e quindi x appartiene all'unione delle classi di $\sim$-equivalenza. Questo dimostra che vale (1). Per dimostrare che vale (2) è sufficiente dimostrare che se $[x] \cap [y] \neq \varnothing$ allora $[x] = [y]$. Sia $z \in [x] \cap [y]$ e quindi $x \sim z$ e $z \sim y$. Supponiamo che $x' \in [x]$ cioè $x' \sim x$. Per la transitività di $\sim$ abbiamo che $x' \sim z$ e di nuovo per transitività si ha che $x' \sim y$: quindi $x' \in [y]$. Questo dimostra che $[x] \subset [y]$. Per dimostrare che vale $[y] \subset [x]$ si procede in modo simile. Ora supponiamo che $\{X_i\}_{i\in I}$ sia una partizione di X. Definiamo la relazione $\sim$ su X dichiarando che $x \sim x'$ se e solo se esiste $i \in I$ tale che $x, x' \in X_i$: si vede facilmente che $\sim$ è di equivalenza e che le X_i sono le sue classi di equivalenza. □

La seguente osservazione è semplice ma importante.

Osservazione 1.3.11. Sia X un insieme, $\sim$ una relazione di equivalenza su X e π l'applicazione quoziente di $\sim$. Dato un insieme Y e un'applicazione $f\colon X \to Y$ esiste una $\overline{f}\colon (X/\sim) \to Y$ tale che $f = \overline{f} \circ \pi$ se e solo se f è costante sulle classi di $\sim$-equivalenza cioè $x_1 \sim x_2$ implica che $f(x_1) = f(x_2)$. Se così è diciamo che f *discende* a $(X/\sim)$.

Un esempio: sia $f\colon \mathbb{Z} \to \{0, 1, 2, \ldots, 9\}$ l'applicazione che associa a x l'ultima cifra del suo sviluppo in base 10, quindi $f(3) = 3$, $f(15) = 5$, $f(2011) = 1$. Se x è congruo a y modulo 10 allora $f(x) = f(y)$ - vedi l'Osservazione 1.3.4 - quindi f discende a $\mathbb{Z}/(10)$ e definisce $\overline{f}\colon \mathbb{Z}/10 \to \{0, 1, 2, \ldots, 9\}$.

1.4 Induzione matematica

Consideriamo la seguente equazione:

$$1 + 2 + \ldots + n = \frac{n(n+1)}{2}. \tag{1.4.1}$$

Dimostriamo che la (1.4.1) vale per ogni n nel modo seguente. Innanzitutto osserviamo che (1.4.1) vale per $n = 1$ sostituendo 1 a n in entrambi i membri (otteniamo $1 = 1$). Ora assumiamo che la (1.4.1) valga per un certo n e dimostriamo che vale anche se sostituiamo $n + 1$ al posto di n cioè che vale

$$1 + 2 + \ldots + n + (n+1) = \frac{(n+1)(n+2)}{2}. \tag{1.4.2}$$

Per l'ipotesi che la (1.4.1) valga per n abbiamo

$$1 + 2 + \ldots + n + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2}$$

e questo dimostra che vale (1.4.2). Quindi abbiamo verificato che (1.4.1) vale per $n = 1$, e perciò anche per $n = 1+1 = 2$ e quindi anche per $n = 2+1 = 3$ etc., in definitiva abbiamo dimostrato che (1.4.1) vale per ogni naturale strettamente positivo n. Questa è una dimostrazione per *induzione* (matematica): la verifica che vale per $n = 1$ è il *primo passo*, la dimostrazione che se (1.4.1) vale per un certo n allora

vale anche sostituendo $n+1$ al posto di n è il *passo induttivo*. La (1.4.1) vale per tutti gli n una volta verificato il primo passo e dimostrato il passo induttivo perchè vale il seguente assioma (fa parte degli assiomi di Peano per l'insieme dei numeri naturali):

Assioma 1.4.1. *Sia $X \subset \mathbb{N}$ un insieme che contiene $0 \in \mathbb{N}$ e tale che valga:*

$$\text{se } X \text{ contiene } n \text{ allora contiene anche } n+1.$$

Allora $X = \mathbb{N}$.

Infatti sia $Y \subset \mathbb{N}$ il sottoinsieme degli n tali che valga (1.4.1) e $X := Y \cup \{0\}$: per quello che abbiamo dimostrato la X soddisfa le ipotesi dell'Assioma 1.4.1 e quindi $X = \mathbb{N}$. Segue che Y è l'insieme dei naturali maggiori o uguali a 1 cioè la (1.4.1) vale per ogni $n \geqslant 1$.

Osservazione 1.4.2. Il passo induttivo di una dimostrazione per induzione può anche essere formulato nel modo seguente: supponiamo che l'affermazione $\mathcal{P}_m$ (di cui vogliamo dimostrare la validità per ogni $m \in \mathbb{N}$) sia vera per ogni $m \leqslant (n-1)$ (questa è l'ipotesi induttiva), e dimostriamo che è vera per $m = n$.

1.5 L'algoritmo euclideo

Nelle scuole si impara che ogni numero naturale non nullo si decompone nel prodotto di numeri primi e che tale decomposizione è unica a meno di riordinamento. La prima affermazione è elementare, la seconda no. Qui dimostreremo la seconda affermazione. Alla base della dimostrazione c'è l'algoritmo euclideo, che permette di calcolare rapidamente il massimo comun divisore di due numeri interi.

Prima enunciamo un risultato che equivale alla divione con resto della scuola (la dimostrazione è lasciata al lettore).

Proposizione 1.5.1 (Divisione con resto)**.** *Siano $a, b \in \mathbb{Z}$, con b non nullo. Allora esistono $q \in \mathbb{Z}$ (quoziente) e $r \in \mathbb{N}$ (resto) tali che*

$$a = b \cdot q + r, \quad 0 \leqslant r < |b|. \tag{1.5.1}$$

Inoltre tali q, r sono unici.

Ora dimostriamo due importanti proprietà del massimo comun divisore di due numeri interi (non entrambi nulli).

Teorema 1.5.2 (Massimo comun divisore tra interi)**.** *Siano $a, b \in \mathbb{Z}$, non entrambi nulli. Allora esiste $m \in \mathbb{Z}$ tale che valgano le seguenti proprietà:*

1. *$m|a$ e $m|b$,*
2. *se $d \in \mathbb{Z}$ e $d|a$, $d|b$, allora $d|m$,*
3. *ed esistono $x, y \in \mathbb{Z}$ tali che $a \cdot x + b \cdot y = m$.*

Inoltre tale m è unico a meno di moltiplicazione per ± 1.

Dimostrazione. Per induzione sul minimo tra $|a|$ e $|b|$. Più precisamente per $n \geqslant 0$ sia $\mathcal{P}_n$ l'affermazione che vale la tesi del teorema per (a, b) con $\min\{|a|, |b|\} \leqslant n$. Dimostriamo per induzione su n che $\mathcal{P}_n$ vale per ogni n.

Dimostriamo che vale l'affermazione $\mathcal{P}_0$. Quindi supponiamo che $a = 0$ o $b = 0$. Possiamo assumere che $b = 0$, e quindi $a \neq 0$. Siccome 0 è un multiplo di qualsiasi intero, valgono (1), (2) e (3) se e solo se $m = \pm a$.

Dimostriamo il passo induttivo. Quindi supponiamo che $\min\{|a|, |b|\} = n + 1$. Quindi possiamo supporre che $n + 1 = |b| \leqslant |a|$. Per la Proposizione 1.5.1 esistono q, r tali che valga (1.5.1). Siccome $0 \leqslant r < |b| = (n + 1)$, per ipotesi induttiva esiste $m \in \mathbb{Z}$ tale che

(A) $m|b$ e $m|r$,

(B) se $d \in \mathbb{Z}$ e $d|b$, $d|r$, allora $d|m$,

(C) ed esistono $x', y' \in \mathbb{Z}$ tali che $b \cdot x' + r \cdot y' = m$.

Dimostriamo che valgono (1), (2) e (3). Dalla (1.5.1) segue che vale la (1) della Tesi della Proposizione. Per dimostrare che vale la (2) supponiamo che $d|a$ e $d|b$; allora $d|r$ per la (1.5.1), e quindi $d|m$ per (B). Infine abbiamo

$$m = b \cdot x' + r \cdot y' = b \cdot x' + (a - bq) \cdot y' = a \cdot y' + b \cdot (x' - qy').$$

Questo dimostra che vale anche (3). Inoltre, se m' è un altro intero con le stesse proprietà di m allora $m|m'$ e $m'|m$, quindi $m' = \pm m$. □

Siano $a, b \in \mathbb{Z}$, non entrambi nulli. Il massimo M tra i divisori comuni di a e b soddisfa (1), (2) e (3) del Teorema 1.5.2, cioè $m = \pm M$. Infatti per il Teorema $M|m$, ma essendo M il massimo tra i divisori comuni di a e b, si ha necessariamente $m = \pm M$. Per M useremo la notazione

$$\mathrm{mcd}\{a, b\} := \max\{d \in \mathbb{Z} \mid d|a, \quad d|b\}. \tag{1.5.2}$$

Per la Proposizione 1.5.1 ogni divisore comune di a e b divide $\mathrm{mcd}\{a, b\}$ - quest'affermazione non è banale. Inoltre esistono $x, y \in \mathbb{Z}$ tali che $a \cdot x + b \cdot y = \mathrm{mcd}\{a, b\}$ - di nuovo quest'affermazione non è banale. L'*algoritmo euclideo* dà un metodo efficiente per calcolare $\mathrm{mcd}\{a, b\}$ (calcolare tutti i divisori di a e b è molto laborioso). Si basa sui seguenti fatti:

(F1) $\mathrm{mcd}\{a, b\} = \mathrm{mcd}\{b, a\}$,

(F2) se $b = 0$, allora $\mathrm{mcd}\{a, b\} = |a|$,

(F3) se $b > 0$ e vale (1.5.1), allora $\mathrm{mcd}\{a, b\} = \mathrm{mcd}\{b, r\}$.

Le prime due affermazioni sono banalmente vere, l'ultima è stata dimostrata nel corso della dimostrazione della Proposizione 1.5.1 (si tratta di verificare che un intero divide sia a che b se e solo se divide sia b che r). Dati $a, b \in \mathbb{Z}$, per calcolare $\mathrm{mcd}\{a, b\}$ seguiamo i seguenti passi:

1. Scambiando a e b, se necessario, possiamo assumere, grazie a F1, che $|a| \geqslant |b|$.
2. Se $b = 0$ sappiamo che $\mathrm{mcd}\{a, b\} = |a|$ grazie a F2.
3. Se $b \neq 0$, allora, grazie a F3, sappiamo che $\mathrm{mcd}\{a, b\} = \mathrm{mcd}\{b, r\}$ dove r è il resto della divisione di a per b, e siccome $|a| \geqslant |b| > r \geqslant 0$, abbiamo $\max\{|a|, |b|\} > \max\{|b|, |r|\}$.

Quindi iterando i passi descritti produrremo $a, b, r, r_1, \ldots$ fino ad arrivare a r_n, r_{n+1} con $r_n \neq 0$ e $r_{n+1} = 0$. Per quanto detto $\mathrm{mcd}\{a, b\} = r_n$.

Esempio 1.5.3. Cacoliamo $\mathrm{mcd}\{861, 308\}$. Dividiamo 861 per 308: siccome il resto è 245, abbiamo $\mathrm{mcd}\{861, 308\} = \mathrm{mcd}\{308, 245\}$. Iterando troviamo

$$\mathrm{mcd}\{861, 308\} = \mathrm{mcd}\{308, 245\} = \mathrm{mcd}\{245, 63\} = \mathrm{mcd}\{63, 56\} = \mathrm{mcd}\{56, 7\} = \mathrm{mcd}\{7, 0\} = 7.$$

1.6 Anelli e campi

Definizione e prime proprietà

Sia A un insieme provvisto di due operazioni, la *somma*

$$\begin{array}{ccc} A \times A & \longrightarrow & A, \\ (w, z) & \mapsto & w + z \end{array} \tag{1.6.1}$$

e la *moltiplicazione*

$$\begin{array}{ccc} A \times A & \longrightarrow & A, \\ (w, z) & \mapsto & w \cdot z \end{array} \tag{1.6.2}$$

Per esempio $A = \mathbb{N}$ con le usuali operazioni di somma e prodotto, o anche $A = \mathbb{Q}$ o $A = \mathbb{R}$. Esistono moltissimi altri esempi significativi in cui l'insieme A è molto lontano dall'essere un insieme di "numeri". I simboli $+$ e $\cdot$ vengono scelti per ricordare somma e moltiplicazione di numeri, ma potrebbereo essere sostituiti da simboli qualsiasi, per esempio potremmo denotare la somma di a e b con $a \star b$, e così via.

Definizione 1.6.1. Un insieme A con operazioni (1.6.1) e (1.6.2) è un *anello* se

1. Esiste $0 \in A$ tale che $0 + z = z + 0 = z$ per ogni $z \in A$. (Esistenza di un elemento neutro per la somma.)
2. $z_1 + z_2 = z_2 + z_1$ per ogni $z_1, z_2 \in A$. (Commutatività della somma.)
3. $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$ per ogni $z_1, z_2, z_3 \in A$. (Associatività della somma.)
4. Dato $z \in A$ esiste $w \in A$ tale che $z + w = 0$ (dove 0 è come in (1)). (Esistenza dell'inverso per la somma.)
5. $(z_1 \cdot z_2) \cdot z_3 = z_1 \cdot (z_2 \cdot z_3)$ per ogni $z_1, z_2, z_3 \in A$. (Associatività del prodotto.)
6. $z_1 \cdot (z_2 + z_3) = z_1 \cdot z_2 + z_1 \cdot z_3$ per ogni $z_1, z_2, z_3 \in A$. (Distributività del prodotto rispetto alla somma.)

Gli insiemi $\mathbb{Z}$, $\mathbb{Q}$ e $\mathbb{R}$ con le usuali operazioni di somma e prodotto sono esempi di anelli. L'insieme $\mathbb{N}$ dei numeri naturali con le usuali operazioni di somma e prodotto *non* è un anello perchè non vale (4).

Definizione 1.6.2. Un anello A è *commutativo* se $a \cdot b = b \cdot a$ per ogni $a, b \in A$.

Gli insiemi $\mathbb{Z}$, $\mathbb{Q}$ e $\mathbb{R}$ con le usuali operazioni di somma e prodotto sono anelli commutativi con unità. Daremo più in là esempi significativi di anelli non commutativi, vedi l'Osservazione 3.4.9.

Proposizione 1.6.3. *Sia A un anello. Allora esiste un unico elemento $0 \in A$ tale che valga (1) della Definizione 1.6.1. Per ogni $z \in A$ si ha che $0 \cdot z = 0$. Dato $z \in A$ esiste un unico $w \in A$ tale che valga (4) della Definizione 1.6.1.*

Dimostrazione. Siano $0, 0' \in A$ tali che $0 + z = z$ e $0' + z = z$ per ogni $z \in A$. Allora $0 + 0' = 0'$, ma per la commutatività della somma $0 + 0' = 0' + 0 = 0$. Quindi $0 = 0'$: questo dimostra che esiste un unico elemento $0 \in A$ tale che valga (1) della Definizione 1.6.1.

Sia $z \in A$: dimostriamo che $0 \cdot z = 0$. Abbiamo che

$$0 \cdot z = (0 + 0) \cdot z = 0 \cdot z + 0 \cdot z \tag{1.6.3}$$

Sia w l'inverso additivo di $0 \cdot z$, cioè $0 \cdot z + w = 0$: aggiungendo w al membro di destra e di sinistra di (1.6.3) (che sono uguali) otteniamo che $0 = 0 \cdot z$.

Dimostriamo che dato $z \in A$ esiste un *unico* $w \in A$ tale che valga (4) della Definizione 1.6.1. Supponiamo che $z + w = 0 = z + w'$: la commutatività e l'associatività della somma danno

$$w' = 0 + w' = (z + w) + w' = (w + z) + w' = w + (z + w') = w + 0 = w.$$

□

Definizione 1.6.4. Sia A un anello e $z \in A$: l'unico inverso additivo di z viene denotato $-z$, e si chiama anche l'*opposto* di z.

Corollario 1.6.5. *Sia A un anello. Allora vale la* regola di cancellazione*: se $a + b = a + b'$ allora $b = b'$.*

Dimostrazione. Abbiamo

$$b = 0 + b = (-a + a) + b = -a + (a + b) = -a + (a + b') = (-a + a) + b' = 0 + b' = b'.$$

□

Definizione 1.6.6. Sia A un anello A. Un elemento $1 \in A$, non uguale a 0, tale che

$$1 \cdot a = a \cdot 1 = a \quad \forall a \in A \tag{1.6.4}$$

è un'*unità* di A. Se A ha un'unità, allora è un *anello con unità*.

Gli insiemi $\mathbb{Z}$, $\mathbb{Q}$ e $\mathbb{R}$ con le usuali operazioni di somma e prodotto sono anelli con unità. L'insieme dei numeri interi pari con le usuali operazioni di somma e prodotto è un esempio di anello senza unità.

Proposizione 1.6.7. *Se A è un anello con unità, allora esiste un unico elemento $1 \in A$ tale che valga* (1.6.4)*.*

Dimostrazione. La dimostrazione è del tutto simile alla dimostrazione dell'unicità dell'elemento neutro per la somma. Se $1, 1' \in A$ sono unità allora $1 \cdot 1' = 1'$ perchè 1 è un'unità e $1 \cdot 1' = 1$ perchè $1'$ è un'unità. Quindi $1 = 1'$. □

Noi considereremo soprattutto anelli commutativi con unità. Per questo motivo, da ora in poi per anello intendiamo *un anello commutativo con unità*, a meno di non specificare che si tratta di un anello non commutativo (o non necessariamente commutativo). Saremo particolarmente interessati ad anelli (commutativi con unità) particolari che si chiamano campi.

Definizione 1.6.8. Un anello (commutativo con unità) A è un *campo* se ogni $0 \neq z \in A$ ha un inverso moltiplicativo cioè esiste $w \in A$ tale che $w \cdot z = 1$.

Gli insiemi $\mathbb{Q}$ e $\mathbb{R}$ con le usuali operazioni sono esempi di campi, ovviamente $\mathbb{Z}$ (con le usuali operazioni) *non* è un campo. In generale denoteremo i campi con la lettera $\mathbb{K}$.

Proposizione 1.6.9. *Sia $\mathbb{K}$ un campo. Dato $0 \neq z \in \mathbb{K}$ esiste un* unico *elemento $w \in \mathbb{K}$ tale che $w \cdot z = 1$. Se $0 \neq z \in \mathbb{K}$ vale la* regola di cancellazione*: se $zw = zw'$ allora $w = w'$.*

Dimostrazione. La dimostrazione che in un campo ogni elemento non-nullo ha un unico inverso moltiplicativo è simile a quella che in un anello ogni elemento ha un unico inverso additivo. Ora supponiamo che $0 \neq z \in k$ e $zw = zw'$. Siccome $0 \neq z$ esiste z' tale che $zz' = 1$; quindi abbiamo che

$$w = 1 \cdot w = (z'z)w = z'(zw) = z'(zw') = (z'z)w' = 1 \cdot w' = w'.$$

□

Corollario 1.6.10. *Sia $\mathbb{K}$ un campo. Supponiamo che $z, w \in \mathbb{K}$ e $zw = 0$. Allora uno almeno tra z e w è uguale a 0.*

Dimostrazione. Supponiamo che $0 \neq z$. Abbiamo che $zw = 0 = z \cdot 0$ (la prima eguaglianza segue dalla Proposizione 1.6.7) e quindi $w = 0$ per la Proposizione 1.6.9. □

Definizione 1.6.11. Se $\mathbb{K}$ è un campo e $0 \neq z \in \mathbb{K}$ l'unico inverso moltiplicativo di z è denotato z^{-1}.

Esempi

Le seguenti definizioni ci servono per dare un'altra serie di esempi interessanti di campi.

Definizione 1.6.12. Sia A un anello. Un sottoinsieme $B \subset A$ è un *sottoanello* se valgono:

1. 0 e 1 sono elementi di B,
2. se $b_1, b_2 \in B$ allora sia $b_1 - b_2$ che $b_1 \cdot b_2$ sono elementi di B.

Se A è un campo un sottoanello B di A è un *sottocampo* se per ogni $b \in B$ non nullo anche b^{-1} è un elemento di B.

Sia $B \subset A$ un sottoanello. Allora la somma e la moltiplicazione di elementi di A definiscono operazioni $B \times B \to B$, e con queste operazioni B è un anello. Analogamente, la somma e la moltiplicazione di un campo $\mathbb{F}$ definiscono operazioni su un sottocampo $\mathbb{K} \subset \mathbb{F}$, e con queste operazioni $\mathbb{K}$ è un campo. Per esempio $\mathbb{Q}$ è un sottocampo di $\mathbb{R}$, e $\mathbb{Z}$ è un sottoanello di $\mathbb{R}$.

Definizione 1.6.13. Siano A un anello, $C \subset A$ un sottoanello e $\alpha \in A$. Poniamo

$$C[\alpha] := \{c_0 + c_1\alpha + \ldots + c_n\alpha^n \mid c_i \in C\}.$$

In altre parole $C[\alpha]$ è il sottoinsieme di A i cui elementi sono quelli che si possono esprimere come polinomi in α a coefficienti in C.

Un semplice ragionamento dimostra il seguente risultato.

Proposizione 1.6.14. *$C[\alpha]$ è un sottoanello di A.*

Esempio 1.6.15. Consideriamo $\mathbb{Q} \subset \mathbb{R}$, $\sqrt{2} \in \mathbb{R}$ e l'anello $\mathbb{Q}[\sqrt{2}] \subset \mathbb{R}$. Questo anello è un campo. Per vederlo notiamo che, siccome $\sqrt{2}^2 = 2 \in \mathbb{Q}$, $\sqrt{2}^3 = 2\sqrt{2}$, $\sqrt{2}^4 = 4 \in \mathbb{Q}$ etc., abbiamo

$$\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}.$$

Ora supponiamo che $(a + b\sqrt{2}) \in \mathbb{Q}[\sqrt{2}]$ sia non nullo; allora

$$(a + b\sqrt{2})(a - b\sqrt{2}) = a^2 - 2b^2 \neq 0,$$

dove la diseguaglianza vale perchè se fosse $a^2 - 2b^2 = 0$, allora, essendo a, b non entrambi nulli, sarebbero entrambi non nulli, e a/b sarebbe una radice quadrata razionale di $\sqrt{2}$, contraddizione. Quindi

$$(a + b\sqrt{2})\left(\frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\sqrt{2}\right) = 1.$$

Questo dimostra che ogni elemento non nullo ha un inverso moltiplicativo, e perciò $\mathbb{Q}[\sqrt{2}]$ è un campo.

Esempio 1.6.16. Generalizziamo l'Esempio 1.6.15 considerando $\mathbb{Q} \subset \mathbb{R}$. Sia d un numero razionale positivo. Descriviamo l'anello $\mathbb{Q}[\sqrt{d}] \subset \mathbb{R}$. Se $\sqrt{d} \in \mathbb{Q}$, allora $\mathbb{Q}[\sqrt{d}] = \mathbb{Q}$, e quindi $\mathbb{Q}[\sqrt{d}]$ è un campo. Se $\sqrt{d} \notin \mathbb{Q}$, allora $\mathbb{Q}[\sqrt{d}] \neq \mathbb{Q}$, e un ragionamento analogo a quello dell'Esempio 1.6.15 dimostra che $\mathbb{Q}[\sqrt{d}]$ è un campo.

Supponiamo che l'anello A e il sottoanello C siano campi, e denotiamoli $\mathbb{F}$ e $\mathbb{K}$ rispettivamente. Quindi $\mathbb{F} \supset \mathbb{K}$, e sia $\mathbb{F}$ che $\mathbb{K}$ sono campi. Per esempio $\mathbb{R} \supset \mathbb{Q}$. Ora sia $\alpha \in \mathbb{F}$, e chiediamoci sotto quali ipotesi $\mathbb{K}[\alpha]$ è un campo. Per esempio abbiamo visto che se $\mathbb{R} \supset \mathbb{Q}$ e $\alpha = \sqrt{d}$, allora $\mathbb{Q}[\sqrt{d}]$ è un campo. Prima di dare una risposta è opportuno introdurre una definizione.

Definizione 1.6.17. Con notazione come sopra, α è *algebrico su* $\mathbb{K}$ se esistono $c_0, \ldots, c_m \in \mathbb{K}$ con $c_0 \neq 0$ tali che

$$c_0\alpha^m + c_1\alpha^{m-1} + \ldots + c_m = 0. \tag{1.6.5}$$

(In altre parole α è soluzione di un'equazione algebrica in una incognita a coefficienti in $\mathbb{K}$.) Se α non è algebrico su $\mathbb{K}$, diciamo che è *trascendente su* $\mathbb{K}$.

Sia d un numero razionale positivo. Allora $\sqrt{d} \in \mathbb{R}$ è algebrico su $\mathbb{Q}$ perchè $(\sqrt{d})^2 - d = 0$, e $d \in \mathbb{Q}$ per ipotesi. D'altra parte un famoso Teorema di Lindemann del 1882 dimostra che π non è algebrico su $\mathbb{Q}$. Il risultato che segue risponde alla domanda fatta sopra.

Teorema 1.6.18. *Siano $\mathbb{K}$ e $\mathbb{F}$ campi, con $\mathbb{K} \subset \mathbb{F}$, e sia $\alpha \in \mathbb{F}$. Il sottoanello $\mathbb{K}[\alpha] \subset \mathbb{F}$ è un campo se e solo se α è algebrico su $\mathbb{K}$.*

Per la dimostrazione del Teorema 1.6.18 vedi la Sezione 1.7.

Esempi con un numero finito di elementi

Gli esempi dati finora di anelli e campi hanno infiniti elementi. Esistono anche anelli e campi con un numero finito di elementi. Sia $n > 1$ un numero naturale.

Lemma 1.6.19. *Siano $a, a', b, b' \in \mathbb{Z}$ tali che*

$$a \equiv a' \pmod{n}, \qquad b \equiv b' \pmod{n}.$$

Allora

$$a + b \equiv a' + b' \pmod{n}, \qquad a \cdot b \equiv a' \cdot b' \pmod{n}.$$

Dimostrazione. Per ipotesi esistono $s, t \in \mathbb{Z}$ tali che $a' = a + sn$ e $b' = b + tn$. Quindi

$$\begin{aligned} a' + b' &= a + sn + b + tn = a + b + (s + t)n, \\ a'b' &= (a + sn)(b + tn) = ab + atn + sbn + stn^2 = ab + (at + sb + stn)n. \end{aligned}$$

□

Per il Lemma 1.6.19 possiamo definire l'operazione di addizione e di moltiplicazione su $\mathbb{Z}/(n)$ ponendo

$$[a] + [b] := [a + b], \qquad [a] \cdot [b] := [a \cdot b].$$

Si verifica facilmente che $\mathbb{Z}/(n)$ è un anello. Ci chiediamo: per quali n l'anello $\mathbb{Z}/(n)$ è un campo? Se n è composto possiamo scrivere $n = ab$ dove $0 < a, b < n$ e quindi $0 \neq [a]$, $0 \neq [b]$ ma $[a] \cdot [b] = [n] = 0$. Per il Corollario 1.6.10 segue che se n è composto allora $\mathbb{Z}/(n)$ non è un campo.

Proposizione 1.6.20. *Se $p \in \mathbb{N}$ è un numero primo allora $\mathbb{Z}/(p)$ è un campo.*

Dimostrazione. Sia $[a] \in \mathbb{Z}/(p)$ non nullo. Dobbiamo dimostrare che esiste l'inverso (moltiplicativo) di $[a]$. Siccome $[a]$ è non nullo, p non divide a. Quindi $\mathrm{mcd}\{a, p\} = 1$ perchè, essendo p primo, i suoi divisori sono ± 1 e $\pm p$. Per il Teorema 1.5.2 esistono $x, y \in \mathbb{Z}$ tali che $a \cdot x + p \cdot y = 1$, e quindi $[x]$ è l'inverso (moltiplicativo) di $[a]$. □

In generale denotiamo un primo con p e poniamo

$$\mathbb{F}_p := \mathbb{Z}/(p). \tag{1.6.6}$$

Sia $\mathbb{K}$ un campo. Siccome $1 \in \mathbb{K}$, ha senso la somma

$$n \cdot 1 := \underbrace{1 + \ldots + 1}_{n}$$

che spesso denoteremo semplicemente n (e $-n$ è l'opposto di n). Quindi abbiamo associato a ogni $n \in \mathbb{Z}$ un elemento $n = n \cdot 1 \in \mathbb{K}$. Consideriamo il campo $\mathbb{F}_3$: si ha che $3 = 3 \cdot 1 = 1 + 1 + 1 = 0$, e questo dimostra che si può avere $n = 0$ (cioè $n \cdot 1 = 0$) *nel campo* $\mathbb{K}$ anche se l'intero n non è 0.

Definizione 1.6.21. La *caratteristica* di un campo $\mathbb{K}$ (notazione: $\operatorname{char}\mathbb{K}$) è 0 se per ogni intero $n \in \mathbb{Z}$ non nullo si ha $n \cdot 1 \neq 0$. Se invece esiste $n \in \mathbb{Z}$ non nullo tale che $n \cdot 1 = 0$ allora $\operatorname{char}\mathbb{K}$ è il *minimo* positivo n tale che $n \cdot 1 = 0$ (nota: se $n \cdot 1 = 0$ anche $(-n) \cdot 1 = 0$, quindi se esiste $n \in \mathbb{Z}$ non nullo tale che $n \cdot 1 = 0$ allora ne esiste uno positivo).

Per esempio $\operatorname{char}\mathbb{Q} = 0$ e $\operatorname{char}\mathbb{F}_p = p$.

Osservazione 1.6.22. Sia $\mathbb{K}$ un campo di caratteristica $p \neq 0$. Allora p è un numero primo. Infatti supponiamo che $p = ab$ con $a, b \in \mathbb{N}$, e dimostriamo che $a = p$ o $b = p$ (questo mostrerà che p è primo). Nel campo $\mathbb{K}$ si ha che

$$0 = p \cdot 1 = (ab) \cdot 1 = (a \cdot 1) \cdot (b \cdot 1)$$

e, per il Corollario 1.6.10, segue che $a \cdot 1 = 0$ o $b \cdot 1 = 0$ (in $\mathbb{K}$). Supponiamo che $a \cdot 1 = 0$: siccome $1 \leqslant a \leqslant p$ e p è il minimo intero strettamente positivo tale che $p \cdot 1 = 0$ segue che $a = p$. Analogamente se $b \cdot 1 = 0$ segue che $b = p$.

Adotteremo la seguente notazione: se A è un anello, e $a, b \in A$, allora $a|b$ (a *divide* b) significa che esiste $x \in A$ tale che $b = a \cdot x$. Per esempio $0|b$ se e solo se $b = 0$, mentre $1|b$ per ogni b.

Omomorfismi e isomorfismi

Anelli o campi che appaiono in contesti molto diversi possono essere identificabili se si considerano solo le relazioni "interne" definite dalle operazioni di addizione e moltiplicazione. La definizione seguente formalizza quest'idea (e la estende).

Definizione 1.6.23. Siano A, B anelli (commutativi con unità). Un'applicazione $f\colon A \to B$ è un *omomorfismo* se valgono

1. $f(1) = 1$,
2. se $a_1, a_2 \in A$ allora $f(a_1 + a_2) = f(a_1) + f(a_2)$ e $f(a_1 \cdot a_2) = f(a_1) \cdot f(a_2)$.

Un omomorfismo $f\colon A \to B$ è un *isomorfismo* se è biunivoco. Un isomorfismo $f\colon A \to A$ di un anello con se stesso è un *automorfismo* di A.

Se $\mathbb{K}, \mathbb{F}$ sono campi, un'applicazione $f\colon \mathbb{K} \to \mathbb{F}$ è un omomorfismo se lo è quando consideriamo $\mathbb{K}, \mathbb{F}$ come anelli, e analogamente un isomorfismo è un omomorfismo biunivoco $f\colon \mathbb{K} \to \mathbb{F}$ degli anelli $\mathbb{K}, \mathbb{F}$.

Esempio 1.6.24. L'identità di un anello è un automorfismo. Diamo un esempio di automorfismo non banale (cioè non uguale all'identità). Sia $d \in \mathbb{Q}$ un numero razionale che non è il quadrato di un numero razionale; l'applicazione $f\colon \mathbb{Q}[\sqrt{d}] \to \mathbb{Q}[\sqrt{d}]$ definita da $f(a + b\sqrt{d}) := a - b\sqrt{d}$ è un automorfismo (vedi l'Esempio 1.6.15). Un esempio di omomorfismo di anelli che non è un isomorfismo è dato dall'applicazione quoziente $\mathbb{Z} \to \mathbb{Z}/(n)$, dove $n \notin \{0, 1\}$.

Lemma 1.6.25. *Siano A, B anelli, e sia $f\colon A \to B$ un omomorfismo. Allora $f(0) = 0$ e, per $a \in A$ si ha $f(-a) = -f(a)$.*

Dimostrazione. Abbiamo

$$f(0) + 0 = f(0) = f(0 + 0) = f(0) + f(0).$$

(Attenzione, usiamo lo stesso simbolo "0" per gli elementi neutri di A e di B.) Per la legge di cancellazione, cioè il Corollario 1.6.7 otteniamo che $f(0) = 0$. Inoltre

$$f(a) + f(-a) = f(a + (-a)) = f(0) = 0.$$

Segue che $f(-a) = -f(a)$. □

Lemma 1.6.26. *Siano $\mathbb{K}, \mathbb{F}$ campi, e sia $f\colon \mathbb{K} \to \mathbb{F}$ un omomorfismo. Allora f è iniettivo e, per $x \in \mathbb{K}$ non nullo, si ha $f(x^{-1}) = f(x)^{-1}$.*

Dimostrazione. Supponiamo che f non sia iniettivo, cioè esistono $x, y \in \mathbb{K}$ tali che $x \neq y$ e $f(x) = f(y)$. Allora (usando il Lemma 1.6.25) troviamo che

$$f(x-y) = f(x) + f(-y) = f(x) - f(y) = 0.$$

Siccome $(x-y) \neq 0$ l'inverso moltiplicativo $(x-y)^{-1}$ esiste, e abbiamo

$$f(1) = f((x-y)\cdot(x-y)^{-1}) = f(x-y)\cdot f((x-y)^{-1}) = 0 \cdot f((x-y)^{-1}) = 0.$$

Questo contraddice la definizione di omomorfismo tra anelli. Ora dimostriamo che se $x \in \mathbb{K}$ è non nullo allora $f(x^{-1}) = f(x)^{-1}$. Abbiamo

$$f(x)\cdot f(x^{-1}) = f(x\cdot x^{-1}) = f(1) = 1.$$

Segue che $f(x^{-1}) = f(x)^{-1}$. □

1.7 Polinomi e funzioni razionali

Polinomi in una indeterminata

Ricordiamo la definizione di polinomio in una indeterminata[2] x a coefficienti in un campo $\mathbb{K}$. Informalmente un tale polinomio è una espressione $a_0 + a_1x + a_2x^2 + \ldots + a_dx^d$ dove $a_0, \ldots, a_d \in \mathbb{K}$. Identifichiamo due tali espressioni se sono uguali i coefficienti **non nulli** dei monomi con esponenti uguali. Siano

$$p = a_0 + a_1x + a_2x^2 + \ldots + a_dx^d, \qquad q = b_0 + b_1x + b_2x^2 + \ldots + b_ex^e \tag{1.7.1}$$

polinomi: la somma di p e q è

$$p + q := (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \ldots + (a_d + b_d)x^d, \tag{1.7.2}$$

il prodotto di p e q è

$$pq := (a_0b_0) + (a_0b_1 + a_1b_0)x + \ldots + \left(\sum_{i+j=m} a_ib_j\right)x^m + \ldots + (a_db_e)x^{d+e}. \tag{1.7.3}$$

Il lettore può avere dubbi sulla correttezza dell'uso di una lettera misteriosa "x": per spazzare via i dubbi può sostituire all'espressione $a_0 + a_1x + a_2x^2 + \ldots + a_dx^d$ una successione $(a_0, \ldots, a_i, \ldots)$ con termini nulli da un certo indice in poi. Definiamo la somma e il prodotto di due tali successioni seguendo le regole date da (1.7.2) e (1.7.3). A questo punto se chiamiamo x la successione $(0, 1, 0, \ldots, 0, \ldots)$ ci rendiamo conto che la successione $(a_0, \ldots, a_i, \ldots, a_d, 0, 0, \ldots)$ è uguale a $a_0 + a_1x + a_2x^2 + \ldots + a_dx^d$. L'insieme dei polinomi in una variabile x a coefficienti in $\mathbb{K}$ si denota $\mathbb{K}[x]$ ed è un anello (ma *non* è un campo, per esempio x non ha un inverso moltiplicativo).

Definizione 1.7.1. Un polinomio $p = a_0 + a_1x + a_2x^2 + \ldots + a_dx^d$ ha grado d se $a_d \neq 0$; in simboli $\deg p = d$. Per convenzione il polinomio nullo ha grado $-\infty$.

Notiamo che se $p, q \in \mathbb{K}[x]$ allora

$$\deg(p+q) \leqslant \max\{\deg p, \deg q\}, \qquad \deg(p\cdot q) = \deg p + \deg q. \tag{1.7.4}$$

[2]Sarebbe più preciso dire "trascendente".

Algoritmo euclideo per polinomi in una indeterminata

Esiste un analogo dei risultati della Sezione 1.5 valido per l'anello $\mathbb{K}[x]$ dei polinomi in una indeterminata a coefficienti in un campo $\mathbb{K}$. Il punto di partenza è la divisione con resto tra polinomi.

Proposizione 1.7.2 (Divisione con resto tra polinomi)**.** *Siano $a, b \in \mathbb{K}[x]$, con b non nullo. Allora esistono $q \in \mathbb{K}[x]$ (quoziente) e $r \in \mathbb{K}[x]$ (resto) tali che*

$$a = b \cdot q + r, \quad \deg r < \deg b. \tag{1.7.5}$$

Inoltre tali q, r sono unici.

Dimostrazione. Fissiamo b e dimostriamo che per ogni $a \in \mathbb{K}[x]$ esistono $q, r \in \mathbb{K}[x]$ tali che valga (1.7.5). Se $\deg b = 0$ cioè $b \in \mathbb{K}^*$ allora basta porre $q = b^{-1} \cdot a$ e $r = 0$. Ora supponiamo che $\deg b > 0$. Se $a = 0$ si pone $q = r = 0$, e perciò supponimao che $a \neq 0$, cioè $\deg a \geqslant 0$. Procediamo per induzione su $\deg a$. Se $\deg a < \deg b$ si pone $q = 0$ e $r = a$. Rimane da dimostrare il passo induttivo. Supponiamo che la divisione con resto esista per ogni $a \in \mathbb{K}[x]$ con $\deg a \leqslant (n-1)$ e dimostriamo che esiste per ogni a di grado n. Se $n < \deg b$ sappiamo che la divisone esiste, quindi possiamo supporre che $n \geqslant \deg b$. Sia $d := \deg b$ e

$$a = a_0 + a_1 x + a_2 x^2 + \ldots + a_{n+1} x^{n+1}, \qquad b = b_0 + b_1 x + b_2 x^2 + \ldots + b_d x^d.$$

Quindi $b_d \neq 0$ perchè $d := \deg b$. Poniamo

$$a' := a - b_d^{-1} a_n x^{n-d} b.$$

Allora $\deg a' < n$ e quindi per l'ipotesi induttiva possiamo scrivere $a' = bq' + r'$ dove $q', r' \in \mathbb{K}[x]$ e $\deg r' < \deg b$. Ma allora

$$a = a' + b_d^{-1} a_n x^{n-d} b = bq' + r' + b_d^{-1} a_n x^{n-d} b = (b_d^{-1} a_n x^{n-d} + q')b + r',$$

e quindi vale (1.7.5) con $q = (b_d^{-1} a_n x^{n-d} + q')$ e $r = r'$.

Ora dimostriamo l'unicità. Supponiamo che $a = b \cdot q + r = a = b \cdot q' + r'$ dove $\deg r < \deg b$ e $\deg r' < \deg b$. Allora $0 = a - a = b \cdot (q - q') + r - r'$, e quindi

$$b \cdot (q - q') = r' - r. \tag{1.7.6}$$

Se $q' \neq q$ allora $\deg(q - q') \geqslant 0$ e quindi $\deg(b \cdot (q - q')) \geqslant \deg b$ per la seconda uguaglianza in (1.7.4). Questa è una contraddizione perchè per la prima uguaglianza in (1.7.4) si ha $\deg(r' - r) < \deg b$. Segue che $q' = q$, e perciò anche $r' = r$ per (1.7.6). □

Teorema 1.7.3 (Massimo comun divisore tra polinomi)**.** *Siano $a, b \in \mathbb{K}[x]$, non entrambi nulli. Allora esiste $c \in \mathbb{K}[x]$ tale che valgano le seguenti proprietà:*

1. *$c|a$ e $c|b$,*
2. *se $d \in \mathbb{K}[x]$ e $d|a$, $d|b$, allora $d|c$,*
3. *ed esistono $s, t \in \mathbb{K}[x]$ tali che $a \cdot s + b \cdot t = c$.*

Inoltre tale c è unico a meno di moltiplicazione per elementi di $\mathbb{K}^\times$ (costanti non nulle).

Dimostrazione. Se uno dei due polinomi è nullo, possiamo assumere che sia $b = 0$, e allora (1), (2) e (3) valgono se e solo se $c = \mu a$ per $\mu \in \mathbb{K}^*$. Se né a né b è nullo, si procede in modo analogo a quanto fatto per dimostrare la Proposizione 1.5.2, l'unica differenza è che si ragiona per induzione sul minimo tra i gradi di a e di b. Siccome vale la divisione con resto tra polinomi data dalla Proposizione 1.7.2, la dimostrazione procede esattamente come nel caso di due interi. □

Siano $a, b \in \mathbb{K}[x]$, non entrambi nulli. Il *massimo comun divisore* di a e b è (l'unico) polinomio monico (cioè tale che il coefficiente del monomio di grado più alto sia 1) c tale che valgano (1), (2), (3) del Teorema 1.7.3. Usiamo la notazione $\mathrm{mcd}\{a, b\}$.

Per calcolare massimo comun divisore di due polinomi si applica l'algoritmo euclideo.

Esempio 1.7.4. Calcoliamo $\mathrm{mcd}\{x^5+x^3+2x+4,\ x^3-x^2+3x+5\}$. Il resto della divisione di x^5+x^3+2x+4 per $x^3 - x^2 + 3x + 5$ è $-9x^2 + 9$. Siccome moltiplicando uno dei polinomi per una costante non nulla non cambia il massimo comun divisore abbiamo $\mathrm{mcd}\{x^5 + x^3 + 2x + 4,\ x^3 - x^2 + 3x + 5\} = \mathrm{mcd}\{x^3 - x^2 + 3x + 5,\ x^2 - 1\}$. Iterando otteniamo

$$\begin{aligned}\mathrm{mcd}\{x^5 + x^3 + 2x + 4,\ x^3 - x^2 + 3x + 5\} &= \mathrm{mcd}\{x^3 - x^2 + 3x + 5,\ x^2 - 1\} = \\ &= \mathrm{mcd}\{x^2 - 1,\ x + 1\} = \mathrm{mcd}\{x + 1,\ 0\} = x + 1.\end{aligned}$$

Radici di polinomi in una indeterminata

Definizione 1.7.5. Sia $p \in \mathbb{K}[x]$. Una *radice* (o *zero*) di p è un $\alpha \in \mathbb{K}$ tale che $p(\alpha) = 0$.

Lemma 1.7.6 (Ruffini)**.** *Sia $p \in \mathbb{K}[x]$. Allora $\alpha \in \mathbb{K}$ è una radice di p se e solo se $(x - \alpha)$ divide p.*

Dimostrazione. Se $(x-\alpha)$ divide p allora $p = (x-\alpha)\cdot q$ dove $q \in \mathbb{K}[x]$ e quindi $p(\alpha) = (\alpha-\alpha)\cdot q(\alpha) = 0$.

Ora supponiamo che α sia una radice di p e dimostriamo che $(x-\alpha)$ divide p. Sia $p = \sum_{i=0}^{n} c_i x^i$. Se $\alpha = 0$ allora $p(\alpha) = 0$ significa che $c_0 = 0$ e quindi $p = x \cdot q$ dove $q = \sum_{i=1}^{n} c_i x^{i-1}$. Lo stesso argomento vale per un α qualsiasi perchè possiamo riscrivere p come $p = \sum_{i=0}^{n} C_i(x-\alpha)^i$ per opportuni coefficienti C_i. Infatti, ponendo $x = ((x - \alpha) + \alpha)$ ed espandendo ciascuna potenza $(x - \alpha) + \alpha)^i$ otteniamo che

$$\sum_{i=0}^{n} c_i x^i = \sum_{i=0}^{n} c_i((x - \alpha) + \alpha)^i = \sum_{i=0}^{n} C_i(x - \alpha)^i.$$

□

Il *grado* di $0 \neq p \in \mathbb{K}[x]$ è definito nel seguente modo. Per ipotesi $p = a_0 + a_1 x + \ldots + a_d x^d$: poniamo

$$\deg p := \max\{i \mid a_i \neq 0\}. \tag{1.7.7}$$

Poniamo $\deg 0 := -\infty$. Siano $p, q \in \mathbb{K}[x]$ non nulli: si verifica facilmente che

$$\deg(p + q) \leqslant \max\{\deg p, \deg q\}, \qquad \deg(p \cdot q) = \deg p + \deg q. \tag{1.7.8}$$

(Per convenzione $\max\{-\infty, n\} = n$ $-\infty + n = -\infty$ per ogni $n \in \mathbb{N}$.)

Siano $0 \neq p \in \mathbb{K}[x]$ e $\alpha \in \mathbb{K}$: osserviamo che esiste un massimo $n \in \mathbb{N}$ tali che $(x - \alpha)^n$ divide p, difatti $n \leqslant \deg p$ per (1.7.8).

Definizione 1.7.7. Siano $p \in \mathbb{K}[x]$ e $\alpha \in \mathbb{K}$. La *molteplicità di α come radice di p* è ∞ se $p = 0$ ed è uguale al massimo $n \in \mathbb{N}$ tale che $(x - \alpha)^n$ divide p se $p \neq 0$ - lo denotiamo $\mathrm{mult}_\alpha p$.

Osservazione 1.7.8. $\alpha \in \mathbb{K}$ è radice di p se e solo se la sua molteplicità come radice di p è almeno 1.

Proposizione 1.7.9. *Sia $p \in \mathbb{K}[x]$ non nullo di grado n. Allora $\mathrm{mult}_\alpha p$ è non zero per un insieme finito di $\alpha \in \mathbb{K}$ e*

$$\sum_{\alpha \in \mathbb{K}} \mathrm{mult}_\alpha p \leqslant \deg p. \tag{1.7.9}$$

Si ha eguaglianza se e solo se si può scrivere

$$p = c \cdot \prod_{i=1}^{n}(x - \alpha_i) \quad c \neq 0. \tag{1.7.10}$$

Dimostrazione. Per induzione sul grado di p. Se $n = 0$ allora $p \in \mathbb{K}$ è non nullo quindi non ha radici: perciò (1.7.9) vale banalmente e $p = c$. (Se il caso $n = 0$ appare troppo banale considerate il caso $n = 1$: allora si può scrivere $p = c \cdot (x - \alpha)$ con $c \neq 0$, p ha una radice, cioè α, di molteplicità 1 e quindi vale (1.7.9).) Ora dimostriamo il passo induttivo. Se p non ha radici non c'è nulla da dimostrare: la (1.7.9) vale banalmente. Supponiamo che p abbia una radice γ. Per il Lemma 1.7.6 esiste $q \in \mathbb{K}[x]$ tale che $p = (x - \gamma) \cdot q$: siccome $p \neq 0$ abbiamo che $q \neq 0$. La formula (1.7.8) dà che $\deg q = d - 1$. Siano $\beta_1, \ldots, \beta_\ell$ le radici distinte di q. Dalla fattorizzazione $p = (x - \gamma) \cdot q$ segue che l'insieme delle radici di p è uguale a $\{\gamma, \beta_1, \ldots, \beta_\ell\}$. Inoltre si vede subito che

$$\operatorname{mult}_\gamma p = 1 + \operatorname{mult}_\gamma q, \qquad \operatorname{mult}_{\beta_i} p = \operatorname{mult}_{\beta_i} q \quad \forall 1 \leqslant i \leqslant \ell. \tag{1.7.11}$$

Per l'ipotesi induttiva

$$\sum_{\alpha \in \mathbb{K}} \operatorname{mult}_\alpha p = 1 + \sum_{\alpha \in \mathbb{K}} \operatorname{mult}_\alpha q \leqslant 1 + \deg q = \deg p.$$

Inoltre vediamo che se si ha equaglianza deve valere $\sum_{\alpha \in \mathbb{K}} \operatorname{mult}_\alpha q = \deg q$. Per ipotesi induttiva segue che vale (1.7.10) per $p = q$: segue che vale anche per p. Il viceversa, cioè se vale (1.7.10) allora (1.7.9) è una eguaglianza, è banalmente vero. □

Corollario 1.7.10. *Sia $p \in \mathbb{K}[x]$ non nullo. Esistono al più $\deg p$ radici di p.*

Dimostrazione. Segue immediatamente dall'Osservazione 1.7.8 e da (1.7.9). □

Funzioni polinomali

Se $p = (c_0 + c_1 x + \ldots + c_d x^d) \in \mathbb{K}[x]$ e $\alpha \in \mathbb{K}$ ha senso sostituire α a x e ottenere così un elemento di $\mathbb{K}$ che si denota $p(\alpha)$:

$$p(\alpha) := c_0 + c_1 \alpha + \ldots + c_d \alpha^d.$$

Osserviamo che valgono le uguaglianze

$$(p + q)(\alpha) = p(\alpha) + q(\alpha), \qquad (p \cdot q)(\alpha) = p(\alpha) \cdot q(\alpha).$$

Quindi a $p \in \mathbb{K}[x]$ possiamo associare la *funzione polinomiale* $\mathbb{K} \to \mathbb{K}$ (che denotiamo con lo stesso simbolo p) definita da $x \mapsto p(x)$ per $x \in \mathbb{K}$.

Proposizione 1.7.11. *Sia $\mathbb{K}$ un campo. Sia $d \in \mathbb{N}$ e supponiamo che $\mathbb{K}$ abbia più di d elementi. Siano $p, q \in \mathbb{K}[x]$ di grado al più d. Le corrispondenti funzioni polinomiali $p, q \colon \mathbb{K} \to \mathbb{K}$ sono uguali se e solo se $p = q$ (cioè i coefficienti di p e q sono gli stessi). In particolare se $\mathbb{K}$ è infinito allora due funzioni polinomiali sono uguali se e solo se sono associate a polinomi uguali.*

Dimostrazione. È ovvio che se $p = q$ allora le funzioni polinomiali associate sono uguali. Ora dimostriamo che se le funzioni polinomiali sono uguali allora $p = q$. Considerando la differenza $(p - q)$ vediamo che basta dimostrare che se $p \in \mathbb{K}[x]$ ha grado al più d e la funzione polinomiale associata è uguale a 0 allora $p = 0$. Ragioniamo per assurdo. Supponiamo che $p \neq 0$. Per ipotesi esistono $\alpha_1, \ldots, \alpha_{d+1} \in \mathbb{K}$ distinti. Siccome la funzione polinomiale associata a p è uguale a 0 abbiamo che $\alpha_1, \ldots, \alpha_{d+1}$ sono radici di p: questo contraddice la Proposizione 1.7.9. □

La Proposizione 1.7.9 permette di identificare polinomi a coefficienti razionali, reali o complessi e funzioni polinomiali $\mathbb{Q} \to \mathbb{Q}$, $\mathbb{R} \to \mathbb{R}$ o $\mathbb{C} \to \mathbb{C}$ rispettivamente perchè $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$ hanno cardinalità infinita.

Esempio 1.7.12. Sia $\mathbb{K}$ un campo di cardinalità finita, per esempio $\mathbb{F}_p = \mathbb{Z}/(p)$ (vedi la Proposizione 1.6.20). Sia $d := |\mathbb{K}|$ e siano $a_1, \ldots, a_d$ gli elementi di $\mathbb{K}$. Sia $p \in \mathbb{K}[x]$ il polinomio definito da

$$p(x) := (x - a_1) \cdot \ldots \cdot (x - a_d).$$

Allora $p(x)$ non è nullo perchè il coefficiente di x^d è 1, ma la funzione polinomiale $\mathbb{K} \to \mathbb{K}$ associata a p è nulla. Per esempio, se $\mathbb{K} = \mathbb{F}_p$ il polinomio $p(x)$ è uguale a $x^p - x$. Quindi è chiaro che su un campo finito la stessa funzione polinomiale è esprimibile attraverso tanti polinomi diversi.

Dimostrazione del Teorema 1.6.18. Supponiamo che $\mathbb{K}[\alpha]$ sia un campo. Dobbiamo dimostrare che α è algebrico su $\mathbb{K}$. Possiamo assumere che $\alpha \neq 0$ perchè 0 è algebrico su $\mathbb{K}$. Siccome $\alpha \neq 0$ l'inverso α^{-1} appartiene a $\mathbb{K}[\alpha]$, quindi possiamo scrivere

$$\alpha^{-1} = d_0\alpha^m + d_1\alpha^{m-1} + \ldots + d_m$$

per opportuni $d_0, d_1, \ldots, d_m \in \mathbb{K}$, e possiamo assumere che $d_0 \neq 0$ perchè $d_0, d_1, \ldots, d_m$ non sono tutti nulli. Moltiplicando entrambi i membri per α troviamo che

$$d_0\alpha^{m+1} + d_1\alpha^m + \ldots + d_m\alpha - 1 = 0, \tag{1.7.12}$$

e quindi α è algebrico su $\mathbb{K}$.

Ora supponiamo che α sia algebrico su $\mathbb{K}$ e dimostriamo che $\mathbb{K}[\alpha]$ è un campo. Per ipotesi esiste un polinomio non nullo $p \in \mathbb{K}[x]$ tale che $p(\alpha) = 0$. Sia P un polinomio di grado minimo tra i polinomi non nulli in $\mathbb{K}[x]$ tali che si annullano in α. (Notate: considerare un polinomio di grado minimo "tra quelli che" ha senso perchè l'insieme in question non è vuoto.) Il polinomio P è *irriducibile*, cioè non si può fattorizzare P come prodotto di polinomi in $\mathbb{K}[x]$ di grado minore del grado di P (in altre parole l'unico modo di fattorizzare P è $P = \lambda \cdot (\lambda^{-1}P)$). Infatti supponiamo che $P = Q \cdot R$, dove $Q, R \in \mathbb{K}[x]$ hanno grado minore del grado di P. Siccome

$$0 = P(\alpha) = Q(\alpha) \cdot R(\alpha),$$

segue che $Q(\alpha) = 0$ o $R(\alpha) = 0$. In entrambi i casi avremmo un polinomio non nullo di grado minore del grado di P che si annulla su α, contraddizione. Ora siamo pronti a dimostrare che $\mathbb{K}[\alpha]$ è un campo. Quello che va verificato è che ogni elemento non nullo $\beta \in \mathbb{K}[\alpha]$ ha inverso appartenente a $\mathbb{K}[\alpha]$. Per ipotesi esiste $Q \in \mathbb{K}[x]$ tale che $\beta = Q(\alpha)$. Sia r il massimo comun divisore tra i polinomi Q e P. Per il Teorema 1.7.3 esistono $a, b \in \mathbb{K}[x]$ tali che

$$r(x) = a(x) \cdot P(x) + b(x) \cdot Q(x). \tag{1.7.13}$$

Sempre per il Teorema 1.7.3 sappiamo che r divide P. Ma abbiamo appena dimostrato che P è irriducibile, quindi o $r = \lambda P$ o $r = \lambda$, dove $\lambda \in \mathbb{K}^*$. Se vale la prima ipotesi allora, siccome r divide anche Q, troviamo che P divide Q, cioè $Q = P \cdot R$ dove $R \in \mathbb{K}[x]$ e sostituendo α ad x segue che $Q(\alpha) = P(\alpha) \cdot R(\alpha) = 0$, cioè $\beta = 0$, contraddizione. Quindi deve valere la seconda ipotesi, cioè $r \in \mathbb{K}^*$, perciò $r = 1$. Sostituendo α ad x nell'uguaglianza in (1.7.13) e ricordando che $\beta = Q(\alpha)$, troviamo che

$$1 = a(\alpha) \cdot P(\alpha) + b(\alpha) \cdot \beta.$$

Siccome $P(\alpha) = 0$, concludiamo che $\beta^{-1} = b(\alpha)$, e quindi l'inverso di β è in $\mathbb{K}[\alpha]$. □

Polinomi in più indeterminate

Abbiamo considerato polinomi in una indeterminata. Si definiscono in modo analogo i polinomi in n indeterminate. Se $p\colon \mathbb{N}^n \to \mathbb{K}$ e $I \in \mathbb{N}^n$, poniamo $p_I := p(I)$.

Definizione 1.7.13. $\mathbb{K}[x_1, \ldots, x_n]$ è l'insieme delle funzioni $p\colon \mathbb{N}^n \to \mathbb{K}$ che sono nulle quasi ovunque cioè tali che l'insieme degli $I \in \mathbb{N}^n$ con $p_I \neq 0$ è finito. Un *polinomio a coefficienti in* $\mathbb{K}$ *nelle indeterminate*[3] $x_1, \ldots, x_n$ è un elemento di $\mathbb{K}[x_1, \ldots, x_n]$.

[3] È più appropriato chiamarle "trascendenti".

Dato $I \in \mathbb{N}^n$ denotiamo con x^I il polinomio tale che che $p_I = 1$ e $p_J = 0$ se $J \neq I$. Se $I = (0, \dots, 0)$ denotiamo x^I con 1. Dato $p \in \mathbb{K}[x_1, \dots, x_n]$ possiamo scrivere

$$p = \sum_{I \in \mathcal{I}} p_I x^I \tag{1.7.14}$$

dove $\mathcal{I} \subset \mathbb{N}^n$ è finito. Per esempio elementi di $\mathbb{Q}[x, y]$ sono dati da

$$p := 3x^2 - xy + \frac{1}{5}y^3 - x + 30, \qquad q := 5x^{11}y^2 - \frac{3}{7}xy^5 + y^7 - 3y + 1.$$

(Quello che, a rigore, andrebbe denotato $a \cdot 1$ dove $a \in \mathbb{K}$ si denota semplicemente a.) Siano $p, q \in \mathbb{K}[x_1, \dots, x_n]$. Definiamo la *somma* $(p + q) \in \mathbb{K}[x_1, \dots, x_n]$ e il *prodotto* $p \cdot q \in \mathbb{K}[x_1, \dots, x_n]$ così:

$$(p + q)_I := p_I + q_I, \qquad (p \cdot q)_I := \sum_{J+K=I} (p_J \cdot p_K). \tag{1.7.15}$$

Notate che la sommatoria che definisce il valore di $p \cdot q$ su I ha senso perchè l'insieme delle coppie (J, K) tali che $p_J \neq 0 \neq p_K$ è finito. Inoltre anche $p \cdot q$ è una funzione nulla quasi ovunque, cioè è un polinomio. Con questa scrittura vediamo che la somma e il prodotto di polinomi corrisponde alle operazioni viste (forse) nella scuola secondaria. Il *grado* di un multiindice $I = (i_1, \dots, i_n) \in \mathbb{N}^n$ è la somma $|I| := i_1 + \ldots + i_n$.

Definizione 1.7.14. Un polinomio $p \in \mathbb{K}[x_1, \dots, x_n]$ non nullo ha grado d se nell'espressione (1.7.14) d è il massimo dei gradi dei multiindici I tale che $p_I \neq 0$; in simboli $\deg p = d$. Per convenzione il polinomio nullo ha grado $-\infty$.

Valgono le formule analoghe di quelle in (1.7.4): se $p, q \in \mathbb{K}[x_1, \dots, x_n]$ allora

$$\deg(p + q) \leqslant \max\{\deg p, \deg q\}, \qquad \deg(p \cdot q) = \deg p + \deg q. \tag{1.7.16}$$

A un polinomio $p \in \mathbb{K}[x_1, \dots, x_n]$ associamo la *funzione polinomiale*

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{p} & \mathbb{K} \\ (c_1, \dots, c_n) & \mapsto & \sum_{I \in \mathbb{N}^n} p_I c^I \end{array} \tag{1.7.17}$$

dove $c^I := c_1^{i_1} \cdot c_2^{i_2} \cdot \ldots \cdot c_n^{i_n}$. Notate che la somma, apparentemente infinita, ha senso perchè p è nulla quasi ovunque. Vale un risultato analogo alla Proposizione 1.7.11.

Proposizione 1.7.15. *Sia $\mathbb{K}$ un campo. Sia $d \in \mathbb{N}$ e supponiamo che $\mathbb{K}$ abbia più di d elementi. Siano $p, q \in \mathbb{K}[x_1, \dots, x_n]$ di grado al più d. Le corrispondenti funzioni polinomiali $p, q \colon \mathbb{K} \to \mathbb{K}$ sono uguali se e solo se $p = q$ (cioè i coefficienti di p e q sono gli stessi). In particolare se $\mathbb{K}$ è infinito allora due funzioni polinomiali sono uguali se e solo se sono associate a polinomi uguali.*

Dimostrazione. Dobbiamo dimostriamo che se le funzioni polinomiali sono uguali allora $p = q$. Considerando la differenza $(p - q)$ vediamo che basta dimostrare che se $p \in \mathbb{K}[x_1, \dots, x_n]$ ha grado al più d e la funzione polinomiale associata è uguale a 0 allora $p = 0$. La dimostrazione è per induzione su n. Se $n = 1$ l'affermazione equivale alla Proposizione 1.7.11 e quindi è vera. Ora dimostriamo il passo induttivo. Supponiamo che $n \geqslant 2$ e che il risultato valga per polinomi di grado al più d in $\mathbb{K}[x_1, \dots, x_{n-1}]$. Sia $p \in \mathbb{K}[x_1, \dots, x_n]$ tale che $p(a_1, \dots, a_n) = 0$ per ogni $(a_1, \dots, a_n) \in \mathbb{K}^n$. Siccome $\deg p \leqslant d$ possiamo scrivere

$$p = c_0 x_n^d + c_1 x_n^{d-1} + \ldots + c_d,$$

con $c_0, \dots c_d \in \mathbb{K}[x_1, \dots, x_{n-1}]$ e $\deg c_i \leqslant i$ per ogni $i \in \{0, \dots, d\}$. Abbiamo supposto che

$$p = c_0 a_n^d + c_1(a_1, \dots, a_{n-1}) a_n^{d-1} + \ldots + c_d(a_1, \dots, a_{n-1}) = 0$$

per ogni $(a_1, \dots, a_n) \in \mathbb{K}^n$. Per la Proposizione 1.7.11 segue che $c_i(a_1, \dots, a_{n-1}) = 0$ per ogni $(a_1, \dots, a_{n-1}) \in \mathbb{K}^{n-1}$. Siccome $\deg c_i \leqslant i \leqslant d$ per ogni i i polinomi $c_0, \dots, c_d$ sono nulli per l'ipotesi induttiva, e quindi p è il polinomio nullo. □

Definizione 1.7.16. Un polinomio p a coefficienti in $\mathbb{K}$ nelle indeterminate $x_1, \ldots, x_n$ è *omogeneo di grado* d se vale (1.7.14) con $i_1 + \ldots + i_n = d$ per ogni $I = (i_1, \ldots, i_n) \in \mathcal{I}$. (Nota: il polinomio 0 è omogeneo di grado d per qualsiasi d benchè il suo grado sia $-\infty$.)

Esempio 1.7.17. Sia $f \in \mathbb{F}_p[x, y]$ il polinomio omogeneo $f(x, y) = x^p y - xy^p$. Allora $f(a, b) = 0$ per ogni $(a, b) \in \mathbb{F}_p^2$, cioè la funzione polinomiale associata a f è nulla nonostante f non sia il polinomio nullo. Notate che il campo $\mathbb{F}_p$ ha p elementi e il grado di f è $p + 1$.

Funzioni razionali

Informalmente il campo delle funzioni razionali e coefficienti in $\mathbb{K}$ in una indeterminata x consiste delle "frazioni" $\frac{p}{q}$ dove $p, q \in \mathbb{K}[x]$ e $q \neq 0$. Più precisamente consideriamo l'insieme delle coppie (p, q) con $p, q \in \mathbb{K}[x]$ e $q \neq 0$, e definiamo la relazione $(p, q) \sim (a, b)$ se $p \cdot b - a \cdot q = 0$ (ottenuta formalmente da $\frac{p}{q} = \frac{a}{b}$ "eliminando" i denominatori). Si verifica facilmente che tale relazione è di equivalenza. Una *funzione razionale a coefficienti in* $\mathbb{K}$ *nell'indeterminata* x è una classe di equivalenza per la relazione appena definita. La classe di equivalenza di (p, q) si denota $\frac{p}{q}$. L'insieme i cui elementi sono tali classi di equivalenza è denotato $\mathbb{K}(x)$. Definiamo somma e moltiplicazione in $\mathbb{K}(x)$ operando come con numeri razionali. Più precisamente, dati $\frac{p}{q}, \frac{a}{b} \in \mathbb{K}(x)$ poniamo

$$\frac{p}{q} + \frac{a}{b} := \frac{pb + aq}{qb}, \qquad \frac{p}{q} \cdot \frac{a}{b} := \frac{pa}{qb}. \tag{1.7.18}$$

La definizione ha senso perchè, come si verifica facilmente, se $\frac{p'}{q'} = \frac{p}{q}$, allora

$$(pb + aq, qb) \sim (p'b + aq', q'b), \qquad (pa, qb) \sim (p'a, q'b). \tag{1.7.19}$$

Ora osserviamo che con queste operazioni $\mathbb{K}(x)$ è un campo, e che abbiamo un omomorfismo di anelli

$$\begin{array}{ccc} \mathbb{K}[x] & \hookrightarrow & \mathbb{K}(x) \\ p & \mapsto & \frac{p}{1} \end{array} \tag{1.7.20}$$

Questa inclusione realizza $\mathbb{K}[x]$ come sottoanello del campo $\mathbb{K}(x)$. Notate l'analogia tra la relazione che c'è tra $\mathbb{K}[x]$ e $\mathbb{K}(x)$ e quella che c'è tra $\mathbb{Z}$ e $\mathbb{Q}$.

Analogamente si definisce il campo $\mathbb{K}(x_1, \ldots, x_n)$ delle funzioni razionali a coefficienti in $\mathbb{K}$ e indeterminate $x_1, \ldots, x_n$, partendo da $\mathbb{K}[x_1, \ldots, x_n]$ anzichè da $\mathbb{K}[x]$. Quindi gli elementi di $\mathbb{K}(x_1, \ldots, x_n)$ sono classi di equivalenza (p, q) con $p, q \in \mathbb{K}[x_1, \ldots, x_n]$ e $q \neq 0$, dove la relazione di equivalenza $(p, q) \sim (a, b)$ è definita come sopra, cioè vale se $p \cdot b - a \cdot q = 0$. Questa inclusione realizza $\mathbb{K}[x_1, \ldots, x_n]$ come sottoanello del campo $\mathbb{K}(x_1, \ldots, x_n)$.

1.8 Fattoriali e coefficienti binomiali

Se n è un numero naturale positivo il *fattoriale di* n è

$$n! := n \cdot (n - 1) \cdot \ldots \cdot 2 \cdot 1. \tag{1.8.1}$$

Inoltre si pone $0! := 1$. Per esempio $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$.

Se n è un numero naturale positivo e $0 \leqslant i \leqslant n$ si pone

$$\binom{n}{i} := \frac{n!}{i!(n - i)!} = \frac{n \cdot (n - 1) \cdot \ldots \cdot (n - i + 1)}{i!}. \tag{1.8.2}$$

Si pone $\binom{0}{0} = 1$.

Proposizione 1.8.1. *Se A è un anello, $x, y \in A$ e $n \in \mathbb{N}$ allora*

$$(x + y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^i. \tag{1.8.3}$$

Per via della Proposizione 1.8.1 $\binom{n}{i}$ si chiama *coefficiente binomiale*. Prima di dimostrare la Proposizione 1.8.1 diamo un'interpretazione combinatoriale del fattoriale e dei coefficienti binomiali.

Proposizione 1.8.2. *Se X è un insieme con n elementi il numero degli ordinamenti totali (vedi la Definizione 1.3.5) di X è uguale a $n!$.*

Dimostrazione. Per induzione su n. Se $n = 1$ esiste un unico ordinamento, e $1! = 1$. Dimostriamo il passo induttivo. Un ordinamento totale $\preceq$ di X ha un elemento massimo cioè esiste $x_0 \in X$ tale che $x \preceq x_0$ per ogni $x \in X$ (l'ipotesi che $\preceq$ sia un ordinamento totale è essenziale). Quindi possiamo definire un'applicazione

$$\mathrm{OT}(X) \xrightarrow{\mu} X$$

dall'insieme degli ordinamenti totali di X a X associando a ogni ordinamento totale il suo elemento massimo. Se $x \in X$, la controimmagnie $\mu^{-1}(x)$ è identifcata con l'insieme degli ordinamenti totali di $X\backslash\{x\}$, e quindi ha cardinalità $(n-1)!$ per ipotesi induttiva. Questo dimostra che

$$|\,\mathrm{OT}(X)| = |X| \cdot (n-1)! = n \cdot (n-1)! = n!.$$

□

Proposizione 1.8.3. *Se X è un insieme con n elementi e $0 \leqslant i \leqslant n$, il numero dei sottoinsiemi di X di cardinalità i è uguale a $\binom{n}{i}$.*

Dimostrazione. Sia $\mathcal{P}_i(X)$ l'insieme i cui elementi sono i sottoinsiemi di X di cardinalità i, e sia $\mathrm{OT}_i(X)$ l'insieme i cui elementi sono i sottoinsiemi di X di cardinalità i con un ordinamento totale. Per capirci: se $X = \{a, b, c\}$ allora $\mathcal{P}_2(X) = \{\{a, b\}, \{a, c\}, \{b, c\}\}$, e quindi $\mathcal{P}_2(X)$ ha tre elementi, mentre $\mathrm{OT}_2(X) = \{\{a \preceq b\}, \{b \preceq a\}, \{a \preceq c\}, \{c \preceq a\}, \{b \preceq c\}, \{c \preceq b\}\}$ Imitando la dimostrazione del la Proposizione 1.8.2 si trova che

$$|\,\mathrm{OT}_i(X)| = n \cdot (n-1) \cdot \ldots \cdot (n-i+1).$$

D'altra parte c'è un'applicazione

$$\mathrm{OT}_i(X) \xrightarrow{\nu} \mathcal{P}_i(X)$$

definita associando al sottoinsieme totalmente ordinato $x_{p_1} \preceq x_{p_2} \preceq \ldots \preceq x_{p_i}$ l'insieme $\{x_{p_1}, x_{p_2}, \ldots, x_{p_i}\}$ dei suoi elementi. Se $A \in \mathcal{P}_i(X)$ la cardinalità di $\nu^{-1}(A)$ è $i!$ per la Proposizione 1.8.2, e quindi

$$\mathcal{P}_i(X) = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-i+1)}{i!} = \binom{n}{i}.$$

□

Lemma 1.8.4. *Se $n \in \mathbb{N}$ e $1 \leqslant i \leqslant n$ allora*

$$\binom{n+1}{i} = \binom{n}{i} + \binom{n-1}{i}. \tag{1.8.4}$$

Dimostrazione. Sia $X := \{1, 2, \ldots, n+1\}$. Per la Proposizione 1.8.3 il membro di sinistra di (1.8.4) è uguale alla cardinalità di $\mathcal{P}_i(X)$ (notazione come nella dimostrazione della Proposizione 1.8.3). Abbiamo la decomposizione in unione disgiunta

$$\mathcal{P}_i(X) = \{A \in \mathcal{P}_i(X) \mid (n+1) \notin A\} \sqcup \{A \in \mathcal{P}_i(X) \mid (n+1) \in A\}, \tag{1.8.5}$$

dove *unione disgiunta* significa che l'insieme di sinistra è l'unione dei due insiemi a destra, e l'intersezione degli insieme di destra è vuota. Il primo degli insiemi a destra di (1.8.5) ha cardinalità $\binom{n}{i}$ per la Proposizione 1.8.3. D'altra parte abbiamo un'applicazione biunivoca

$$\begin{array}{ccc} \{A \in \mathcal{P}_i(X) \mid (n+1) \in A\} & \longrightarrow & \mathcal{P}_{i-1}(\{1, 2, \ldots, n\}) \\ A & \mapsto & A \cap \{1, 2, \ldots, n\} \end{array}$$

e quindi il secondo degli insiemi a destra di (1.8.5) ha cardinalità $\binom{n-1}{i}$ per la Proposizione 1.8.3. Questo dimostra che vale (1.8.4). □

Dimostrazione della Proposizione 1.8.1. Per induzione su n. Se $n = 0$ la (1.8.3) vale perchè entrambi i membri sono uguali a 1. Dimostriamo il passo induttivo. Quindi supponiamo che valga (1.8.3) per $n \in \mathbb{N}$ e dimostriamo che vale

$$(x+y)^{n+1} = \sum_{i=0}^{n+1} \binom{n+1}{i} x^{n+1-i} y^i. \tag{1.8.6}$$

Moltiplicando (1.8.3) per $(x+y)$ otteniamo che

$$\begin{aligned}(x+y)^{n+1} &= \sum_{i=0}^{n} \binom{n}{i} x^{n+1-i} y^i + \sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^{i+1} = \sum_{j=0}^{n} \binom{n}{j} x^{n+1-j} y^j + \sum_{j=1}^{n+1} \binom{n}{j-1} x^{n+1-j} y^j = \\ &= \binom{n}{0} x^{n+1} + \sum_{j=1}^{n+1} \left(\binom{n}{j} + \binom{n}{j-1} \right) x^{n+1-j} y^j + \binom{n}{n} y^{n+1} = \sum_{i=0}^{n+1} \binom{n+1}{i} x^{n+1-i} y^i. \end{aligned} \tag{1.8.7}$$

Abbiamo dimostrato il passo induttivo. □

1.9 Numeri complessi

L'insieme dei *numeri complessi* $\mathbb{C}$ è definito nel modo seguente. Come insieme $\mathbb{C}$ è $\mathbb{R}^2$. La somma è quella puntuale cioè

$$(a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2). \tag{1.9.1}$$

La moltiplicazione è definita così:

$$(a_1, b_1) \cdot (a_2, b_2) := (a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1). \tag{1.9.2}$$

Il sottoinsieme di $\mathbb{C}$ dato dalle coppie $(a, 0)$ si può identificare con l'insieme dei reali nel senso che $(a_1, 0) + (a_2, 0) = (a_1 + a_2, 0)$ e $(a_1, 0) \cdot (a_2, 0) = (a_1 a_2, 0)$. Quindi da ora in poi se $a \in \mathbb{R}$ denoteremo con a il numero complesso $(a, 0)$. Poniamo

$$i := (0, 1). \tag{1.9.3}$$

Osserviamo che

$$i \cdot i = (0, 1) \cdot (0, 1) = (-1, 0) = -1. \tag{1.9.4}$$

In altre parole i è una radice di -1. Possiamo scrivere

$$(a, b) = (a, 0) + (b, 0)(0, 1) = a + bi. \tag{1.9.5}$$

Da ora in poi quando diciamo che $(a + bi)$ è un numero complesso intendiamo che $a, b \in \mathbb{R}$. (1.9.6)

Se $z = a + bi$ allora a è la *parte reale di* z e b è la *parte immaginaria di* z; in simboli

$$\Re(a + bi) = a, \qquad \Im(a + bi) = b. \tag{1.9.7}$$

Con questa scrittura le definizioni di somma e prodotto danno che

$$\begin{aligned}(a_1 + b_1 i) + (a_2 + b_2 i) &= (a_1 + a_2) + (b_1 + b_2) i, \\ (a_1 + b_1 i)(a_2 + b_2 i) &= (a_1 a_2 - b_1 b_2) + (a_1 b_2 + a_2 b_1) i. \end{aligned}$$

In particolare si verifica facilmente che $\mathbb{C}$ è un anello, di fatto $\mathbb{C}$ è un campo: l'inverso moltiplicativo di $0 \neq (a + bi)$ è dato da

$$(a + bi)^{-1} = (a^2 + b^2)^{-1} (a - bi). \tag{1.9.8}$$

(Qui $(a^2 + b^2)^{-1}$ è l'inverso del reale $(a^2 + b^2)$ in $\mathbb{R}$.)

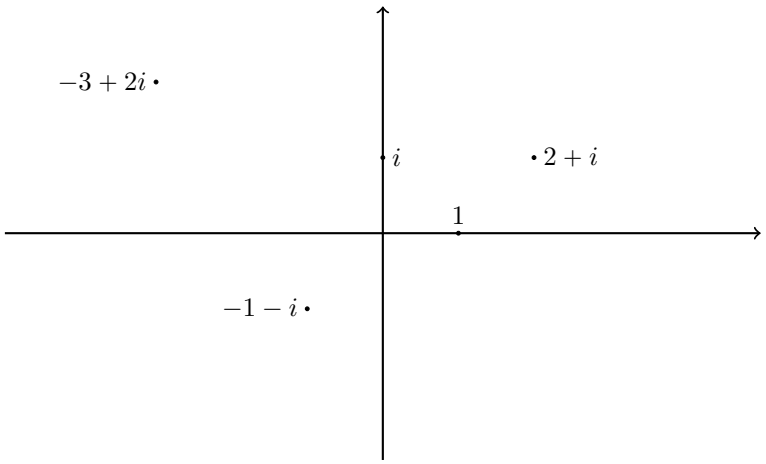

Figura 1.1: Numeri complessi e punti del piano

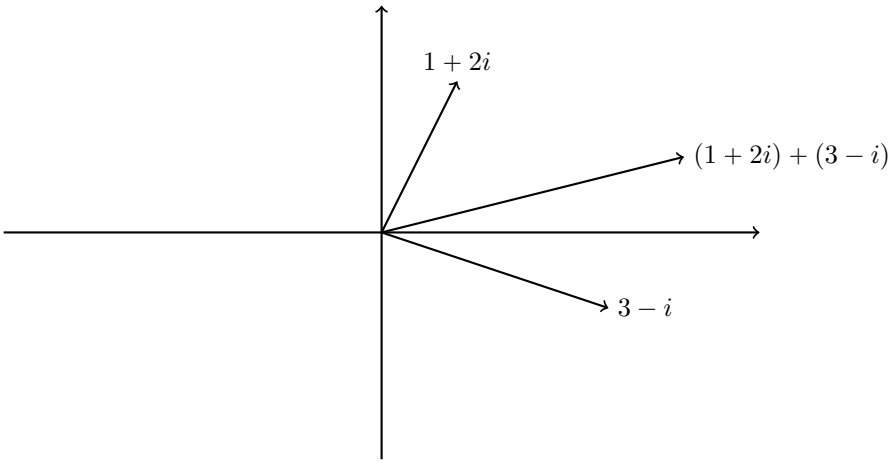

Figura 1.2: Somma di numeri complessi

Per visualizzare somma e moltiplicazione di numeri complessi scegliamo un sistema di coordinate cartesiane nel piano e associamo al numero complesso $a+bi$ il punto di coordinate (a,b). Se identifichiamo un punto P del piano con il vettore nel piano rappresentato dal segmento orientato che va dall'origine del sistema di coordinate a P, allora la somma di numeri complessi corrisponde alla "regola del parallelogramma". Per "vedere" la moltiplicazione diamo un paio di definizioni. Sia $(a+bi)\in\mathbb{C}$ (ricordate la (1.9.6)): poniamo

$$|a+bi| := (a^2+b^2)^{1/2} \tag{1.9.9}$$

e lo chiamiamo il *modulo* di $(a+bi)$. Sia $0\neq z\in\mathbb{C}$ e $w := w/|z|$. Allora $|w|=1$ cioè $w=c+di$ dove $c^2+d^2=1$ e quindi esiste $\theta\in\mathbb{R}$ tale che $w=(\cos\theta+\sin\theta i)$: il numero θ (ben determinato a meno di multipli interi di 2π) si chiama l'*argomento* di z e si indica $\mathrm{Arg}(z)$. In conclusione dato $z\in\mathbb{C}$ possiamo scrivere

$$z=\rho((\cos\theta+\sin\theta i), \qquad \rho=|z|, \quad \theta=\mathrm{Arg}(z). \tag{1.9.10}$$

(Se $z=0$ l'argomento è indeterminato: la (1.9.10) è vera con qualsiasi θ.) Ora siano $z_1,z_2\in\mathbb{C}$ e scriviamo

$$z_1=\rho_1(\cos\theta_1+\sin\theta_1 i), \quad z_2=\rho_2(\cos\theta_2+\sin\theta_2 i).$$

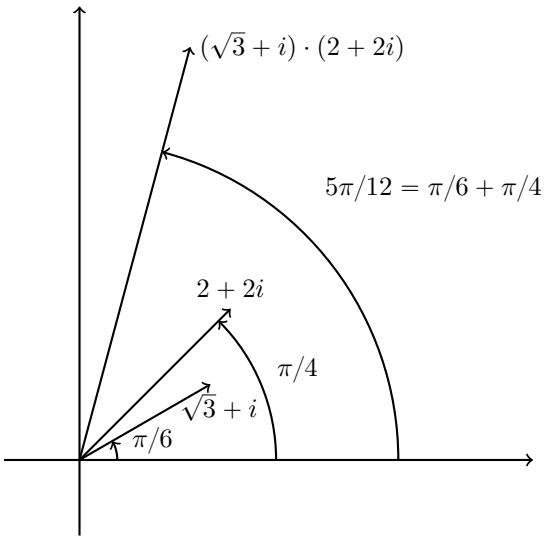

Figura 1.3: Moltiplicazione di numeri complessi

Le formule trigonometriche per il coseno e il seno della somma di angoli danno

$$\begin{aligned} z_1 z_2 = \rho_1(\cos\theta_1 + \sin\theta_1 i)\cdot\rho_2(\cos\theta_2 + \sin\theta_2 i) = \\ = \rho_1\rho_2(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 + (\cos\theta_1\sin\theta_2 + \sin\theta_1\cos\theta_2)i) = \\ = \rho_1\rho_2((\cos(\theta_1+\theta_2) + (\sin\theta_1+\theta_2)i). \end{aligned}$$

Quindi il modulo del prodotto è il prodotto dei moduli e l'argomento del prodotto è la somma degli argomenti:

$$|z_1 z_2| = |z_1|\cdot|z_2|, \qquad \mathrm{Arg}(z_1 z_2) = \mathrm{Arg}(z_1) + \mathrm{Arg}(z_2), \tag{1.9.11}$$

dove l'uguaglianza di argomenti si intende a meno di multipli interi di 2π.

L'importanza di $\mathbb{C}$ è dovuta al seguente risultato.

Teorema fondamentale dell'Algebra 1.9.1. *Sia $n > 0$ un numero naturale e $a_1 \ldots, a_n \in \mathbb{C}$. Esiste $z \in \mathbb{C}$ tale che*

$$z^n + a_1 z^{n-1} + \ldots + a_{n-1} z + a_n = 0.$$

Applicando ripetutamente il Lemma 1.7.6 segue che esistono $c_1, \ldots, c_n \in \mathbb{C}$ tali che

$$z^n + a_1 z^{n-1} + \ldots + a_{n-1} z + a_n = (z-c_1)(z-c_2)\cdot\ldots\cdot(z-c_n).. \tag{1.9.12}$$

In parole: ogni polinomio $p \in \mathbb{C}[z]$ di grado strettamente positivo è prodotto di fattori lineari (cioè polinomi di grado 1).

Illustriamo il Teorema Fondamentale dell'Algebra nel caso del polinomio $p(z) := z^n - a$. Le radici di p sono i numeri complessi w tali che $w^n = a$. Scrivendo $a = \rho(\cos\theta + \sin\theta i)$ troviamo che le n radici di p sono

$$\rho^{1/n}(\cos((\theta + s\pi)/n) + \sin((\theta + s\pi)/n)i), \quad 0 \leqslant s \leqslant (n-1). \tag{1.9.13}$$

Se rappresentiamo le radici n-esime di a con punti del piano allora otteniamo un singolo punto se $a = 0$ e i vertici di un poligono regolare con n lati se $a \neq 0$.

Definizione 1.9.2. Sia $z \in \mathbb{C}$ e scriviamo $z = a + bi$ dove $a, b \in \mathbb{R}$. Il *coniugato* di z è il numero complesso $\overline{z}$ dato da

$$\overline{z} := a - bi. \tag{1.9.14}$$

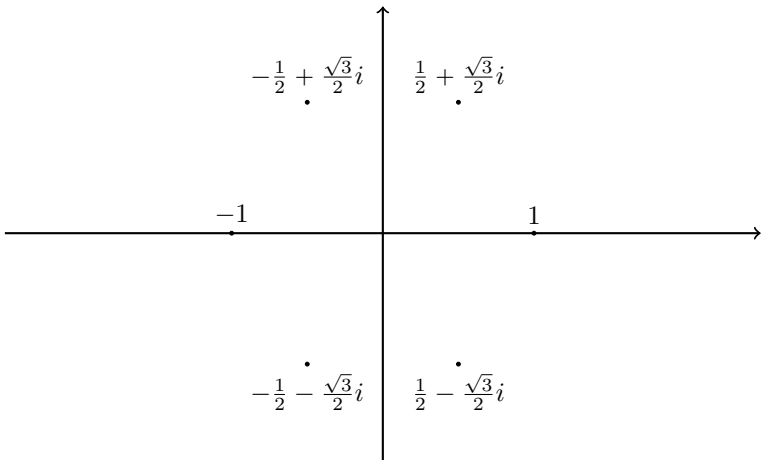

Figura 1.4: Le radici seste di 1

Un facile calcolo dà che valgono le formule

$$\overline{w+z}=\overline{w}+\overline{z},\quad \overline{wz}=\overline{w}\,\overline{z},\quad z\overline{z}=|z|^2. \tag{1.9.15}$$

In particolare, siccome $\overline{1}=1$ la coniugazione è un automorfismo del campo $\mathbb{C}$.

1.10 Gruppi

Definizione ed esempi

Definizione 1.10.1. Un *gruppo* è un insieme G dotato di un'operazione

$$\begin{array}{ccc} G\times G & \longrightarrow & G \\ (g,h) & \mapsto & g\cdot h \end{array}$$

con le seguenti proprietà:

(I) Esiste un elemento $u\in G$ tale che per ogni $g\in G$ valga $u\cdot g=g\cdot u=g$ (esistenza di un elemento unità, cioè u).

(II) Dato $g\in G$ esiste $h\in G$ tale che $g\cdot h=h\cdot g=u$, dove u è come in (1) (esistenza dell'inverso).

(III) Se $g,h,k\in G$ allora $(g\cdot h)\cdot k=g\cdot(h\cdot k)$ (associatività).

Esempio 1.10.2. Sia A un anello. L'operazione di somma dà una struttura di gruppo a A, e questo gruppo si denota $(A,+)$ (notate che in questo caso il simbolo dell'operazione è $+$). L'operazione di moltiplicazione non dà ad A una struttura di gruppo perchè come unità saremmo costretti a scegliere 1 (vedi la dimostrazione della Proposizione 1.6.7), ma allora 0 non ha inverso. Siccome 0 crea problemi, consideriamo l'operazione di moltiplicazione su

$$A^\times := A\backslash\{0\}. \tag{1.10.1}$$

Attenzione: notate che il prodotto di due elementi di $A^\times$ non è necessariamente un elemento di $A^\times$, per esempio il prodotto delle classi di congruenza modulo 6 degli elementi non nulli $[2]\in\mathbb{Z}/(6)$ e $[3]\in\mathbb{Z}/(6)$ è $[0]$. Facciamo vedere che $A^\times$ è un gruppo (questo include l'affermazione che il prodotto di due elementi di $A^\times$ è un elemento di $A^\times$) se e solo se A è un campo.

Supponiamo che $A^\times$ sia un gruppo. Allora l'unità del gruppo $A^\times$ è necessariamente 1 perchè abbiamo $u=u\cdot 1=1$, e quindi per la condizione (II) A è un campo. Ora supponiamo che A sia un campo

$\mathbb{K}$. Allora la moltiplicazione di due elementi di $\mathbb{K}^\times$ è un elemento di $\mathbb{K}^\times$ per il Corollario 1.6.10. La condizione (I) vale con unità l'unità di $\mathbb{K}$, la (II) vale perchè $\mathbb{K}$ è un campo (e non semplicemente un anello), e la (III) vale perchè vale in un qualsiasi anello.

Esempio 1.10.3. Sia X un insieme. Denotiamo con $\mathcal{S}(X)$ l'insieme delle applicazioni *biunivoche* $f\colon X \to X$ di X in se stesso. Se $f, g \in \mathcal{S}(X)$, allora la composizione $f \circ g\colon X \to X$ è essa stessa biunivoca. Quindi la composizione definisce un'operazione su $\mathcal{S}(X)$. Ora osserviamo che valgono i seguenti fatti:

(I) L'identità Id_X è in $\mathcal{S}(X)$, e per ogni $f \in \mathcal{S}(X)$ si ha $\mathrm{Id}_X \circ f = f \circ \mathrm{Id}_X = f$.

(II) Se $f, g, h \in \mathcal{S}(X)$, allora $f \circ (g \circ h) = (f \circ g) \circ h)$ (associatività).

(III) Se $f \in \mathcal{S}(X)$ allora f^{-1} è ben definita (perchè f è biunivoca) e vale $f \circ f^{-1} = f^{-1} \circ f = \mathrm{Id}_X$.

Notiamo che la composizione definisce un operazione sull'insieme $\mathcal{A}(X)$ di *tutte* le applicazioni di X in X, ma sostituendo $\mathcal{A}(X)$ a $\mathcal{S}(X)$ la proprietà (III) non vale. L'insieme $\mathcal{S}(X)$ con l'operazione di composizione è un gruppo.

Proposizione 1.10.4. *Sia G un gruppo. Allora esiste un* unico *elemento $u \in G$ tale che valga (1) della Definizione 1.10.1. Dato $g \in G$ esiste un* unico *$h \in G$ tale che valga (2) della stessa definizione, cioè $g \cdot h = h \cdot g = u$.*

Dimostrazione. La dimostrazione è del tutto simile alle dimostrazioni delle Proposizioni 1.6.7 e 1.6.9. □

Definizione 1.10.5. Un gruppo $(G, \cdot)$ è *commutativo* (o *abeliano*[4]) se per $g, h \in G$ si ha $g \cdot h = h \cdot g$, cioè se il prodotto *non* dipende dall'ordine dei fattori.

Se A è un anello allora $(A, +)$ è un gruppo commutativo e se $\mathbb{K}$ è un campo $\mathbb{K}^\times$ è un gruppo commutativo.

Notazione 1.10.6. Spesso l'operazione di un gruppo commutativo si denota $+$. In tal caso l'elemento unità si denota 0.

Sia X un insieme di cardinalità almeno 3: dimostriamo che $\mathcal{S}(X)$ *non* è commutativo. Siccome X ha cardinalità almeno 3 esistono elementi distinti $a, b, c \in X$. Siano $f, g \in \mathcal{S}(X)$ le applicazioni biunivoche definite da

$$f(a) = b, \quad f(b) = a, \quad f(x) = x \text{ se } x \notin \{a, b\}, \qquad g(b) = c, \quad g(c) = b, \quad f(x) = x \text{ se } x \notin \{b, c\}.$$

Allora $f \circ g \neq g \circ f$ perchè $f \circ g(a) = b$ e $g \circ f(a) = c$.

Definizione 1.10.7. Sia G un gruppo. L'unico elemento $u \in G$ tale che valga (1) della Definizione 1.10.1 è l'*unità* di G, e si denota 1 (si denota 0, e si chiama elemento neutro, se il gruppo è commutativo e l'operazione è denotata "+"). Dato $g \in G$ l'unico $h \in G$ tale che $g \cdot h = h \cdot g = u$ è l'*inverso* di g e si denota g^{-1} (o l'*opposto* di g se il gruppo è commutativo e l'operazione è denotata "+", e allora si denota $-g$). Inoltre vale la legge di cancellazione: se $g, h, k \in G$ e $g \cdot h = g \cdot k$ allora $h = k$.

Definizione 1.10.8. Sia G un gruppo. Un *sottogruppo di G* è un sottoinsieme $H \subset G$ che contiene l'unità di G, e tale che se $g_1, g_2 \in H$ allora anche $g_1 \cdot g_2 \in H$, e $g_1^{-1} \in H$. In altre parole H è un sottogruppo di G se è non vuoto e l'operazione di G induce una struttura di gruppo su H.

Definizione 1.10.9. $\mathbb{Z} \subset \mathbb{Q}$ è un sottogruppo, se $\mathbb{Q}$ è provvisto dell'operazione di addizione. Sia X un insieme e $Y \subset X$ un sottoinsieme. I sottoinsiemi $H, J \subset \mathcal{S}(X)$ definiti da

$$H := \{f\mathcal{S}(X) \mid f(Y) = Y\}, \qquad J := \{f\mathcal{S}(X) \mid f(y) = y \quad \forall y \in Y\}$$

sono sottogruppi di $\mathcal{S}(X)$.

[4]Da N. H. Abel (1802 - 1829) matematico norvegese.

I gruppi interessanti sono spesso gruppi di simmetrie. Diamo un esempio. Denotiamo con $\mathbb{E}^2$ "il" piano della geometria euclidea. Un'applicazione $f\colon \mathbb{E}^2 \to \mathbb{E}^2$ è una *isometria* se per oni $x, y \in \mathbb{E}^2$ i segmenti $\overline{xy}$ e $\overline{f(x)f(y)}$ sono congruenti. Per esempio sono isometrie la rotazione di un angolo fissato con centro un punto fissato, o la traslazione per un vettore fissato. Sia $\mathrm{Isom}(\mathbb{E}^2)$ l'insieme delle isometrie. Si verifica che, con l'operazione di composizione, $\mathrm{Isom}(\mathbb{E}^2)$ è un gruppo (l'unico punto non banale da dimostrare è che una isometria è suriettiva, quindi biunivoca). Questo è il gruppo delle "simmetrie" del piano euclideo.

Omomorfismi e isomorfismi

Gruppi che appaiono in contesti molto diversi possono essere identificabili se si considerano solo le relazioni "interne" definite dalle due moltiplicazioni. La definizione seguente formalizza quest'idea (e la estende).

Definizione 1.10.10. Siano G, H gruppi. Un'applicazione $f\colon G \to H$ è un *omomorfismo* se per $g_1, g_2 \in G$ vale

$$f(g_1 \cdot g_2) = f(g_1) \cdot f(g_2). \tag{1.10.2}$$

Un omomorfismo $f\colon G \to H$ è un *isomorfismo* se è biunivoco.

L'omomorfismo *banale* $f\colon G \to H$ è definito ponendo $f(g) = 1$ per ogni $g \in G$. Diamo qualche esempio significativo di omomorfismi.

Esempio 1.10.11. L'insieme dei reali positivi $\mathbb{R}_+$ con operazione il prodotto è un gruppo e $\mathbb{R}$ con operazione la somma è un gruppo (vedi l'Esempio 1.10.2). Il logaritmo naturale $\log\colon \mathbb{R}_+ \to \mathbb{R}$ è un omomorfismo tra questi gruppi. Siccome log è biunivoco è un isomorfismo. Lo stesso vale per il logaritmo in una qualsiasi base.

Siccome $\mathbb{R}$ e $\mathbb{C}$ sono campi, $\mathbb{R}^\times$ e $\mathbb{C}^\times$ sono gruppi (vedi l'Esempio 1.10.2). La prima uguaglianza di (1.9.11) afferma che l'applicazione

$$\begin{array}{ccc} \mathbb{C}^\times & \longrightarrow & \mathbb{R}^\times \\ z & \mapsto & |z| \end{array}$$

è un omomorfismo di gruppi. Notate che non è un isomorfismo perchè non è nè suriettivo nè iniettivo (se sostituiamo al codominio il gruppo $\mathbb{R}_+$ con operazione la moltiplicazione l'omomorfismo diventa suriettivo, ma non iniettivo, ovviamente).

Lemma 1.10.12. *Siano G, H gruppi, e sia $f\colon G \to H$ un omomorfismo. Allora $f(1) = 1$ e, per $g \in G$ si ha $f(g^{-1}) = f(g)^{-1}$.*

Dimostrazione. La dimostrazione è del tutto simile e a quella dei Lemmi 1.6.25 e 1.6.26. Abbiamo

$$f(1) \cdot 1 = f(1) = f(1 \cdot 1) = f(1) \cdot f(1).$$

Per la legge di cancellazione (vedi la Proposizione 1.10.4) otteniamo che $f(1) = 1$. Inoltre

$$f(g) \cdot f(g^{-1}) = f(g \cdot g^{-1}) = f(1) = f(1) = 1.$$

Moltiplicando entrambi il termine più a sinistra e quello più a destra per $f(g)^{-1}$ otteniamo che $f(g^{-1}) = f(g)^{-1}$. □

Siano G un gruppo e X un insieme. Un omomorfismo $G \to \mathcal{S}(X)$ ha un nome particolare: è un'azione di G su X.

Definizione 1.10.13. Un'*azione di G su X* è un'applicazione

$$\begin{array}{ccc} G \times X & \longrightarrow & X \\ (g, x) & \mapsto & gx \end{array} \tag{1.10.3}$$

con le seguenti proprietà:

1. $1x = x$ per ogni $x \in X$, dove 1 è l'elemento neutro di G, e
2. $(gh)x = g(hx)x$ per ogni $g, h \in G$ e $x \in X$.

Dimostriamo che un'azione di G su X si può identificare con un omomorfismo $G \to \mathcal{S}(X)$. Dato un omomorfismo $\varphi\colon G \to \mathcal{S}(X)$ definiamo

$$\begin{array}{ccc} G \times X & \longrightarrow & X \\ (g, x) & \mapsto & \varphi(g)(x) \end{array} \tag{1.10.4}$$

Vediamo che la (1) della Definizione 1.10.13 vale perchè $\varphi(1) = \mathrm{Id}_{\mathcal{S}(X)}$ (vedi il Lemma 1.10.12), e la (2) vale per definizione di omomorfismo. Viceversa, data un'azione di G su X definiamo

$$\begin{array}{ccc} G & \xrightarrow{\varphi} & \mathcal{S}(X) \\ g & \mapsto & (x \mapsto gx) \end{array} \tag{1.10.5}$$

Si verifica facilmente che φ è un omomorfismo. Inoltre le applicazioni appena definite

$$\{\text{Azioni di } G \text{ su } X\} \longrightarrow \{\text{Omomorfismi } G \to \mathcal{S}(X)\}, \quad \{\text{Omomorfismi } G \to \mathcal{S}(X)\} \longrightarrow \{\text{Azioni di } G \text{ su } X\}$$

sono l'una l'inversa dell'altra. In conclusione possiamo identificare un'azione di G su X con un omomorfismo $G \to \mathcal{S}(X)$.

Definizione 1.10.14. Siano G, H gruppi, e sia $f\colon G \to H$ un isomorfismo, con inversa f^{-1}. Allora f^{-1} è un isomorfismo.

Dimostrazione. Si tratta di dimostrare che f^{-1} è un omomorfismo, cioè che per $h_1, h_2 \in H$ si ha

$$f^{-1}(h_1 \cdot h_2) = f^{-1}(h_1) \cdot f^{-1}(h_2). \tag{1.10.6}$$

Siccome f è biunivoca è sufficiente dimostrare che le immagini per f dei termini di destra e sinistra di (1.10.6) sono uguali. Questo segue direttamente dall'ipotesi che f sia un omomorfismo. □

Esempio 1.10.15. Abbiamo visto che il logaritmo naturale $\log\colon \mathbb{R}_+ \to \mathbb{R}$ è un isomorfismo. L'inverso di log è la funzione esponenziale $\exp\colon \mathbb{R} \to \mathbb{R}_+$. Dire che è un omomorfismo equivale all'uguaglianza $e^{a+b} = e^a \cdot e^b$.

Definizione 1.10.16. I gruppi G e H sono *isomorfi* se esiste un isomorfismo $f\colon G \to H$.

Esempio 1.10.17. Il logaritmo naturale $\log\colon \mathbb{R}_+ \to \mathbb{R}$ è un isomorfismo, quindi $\mathbb{R}_+$ (con operazione il prodotto) è isomorfo a $\mathbb{R}$ con operazione la somma.

Esempio 1.10.18. Supponiamo che gli insiemi X, Y abbiano la stessa cardinalità. Allora $\mathcal{S}(X)$ e $\mathcal{S}(Y)$ sono gruppi isomorfi. Infatti sia $\varphi\colon X \to Y$ un'applicazione biunivoca, e definiamo l'applicazione

$$\begin{array}{ccc} \mathcal{S}(X) & \xrightarrow{F} & \mathcal{S}(Y) \\ g & \mapsto & \varphi \circ g \circ \varphi^{-1} \end{array}$$

Allora F è un omeomorfismo perchè

$$F(g_1 \circ g_2) = \varphi \circ (g_1 \circ g_2) \circ \varphi^{-1} = (\varphi \circ g_1) \circ (g_2 \circ \varphi^{-1}) = (\varphi \circ g_1) \circ (\varphi^{-1} \circ \varphi) \circ (g_2 \circ \varphi^{-1}) = \\ = (\varphi \circ g_1 \circ \varphi^{-1}) \circ (\varphi \circ g_2 \circ \varphi^{-1}) = F(g_1) \cdot F(g_2).$$

Inoltre F è invertibile perchè l'inversa è definita da

$$\begin{array}{ccc} \mathcal{S}(Y) & \xrightarrow{E} & \mathcal{S}(X) \\ h & \mapsto & \varphi^{-1} \circ h \circ \varphi \end{array}$$

Siccome gruppi isomorfi sono indistinguibili per quanto riguarda la struttura di gruppo, ha senso porre la seguente definizione.

Definizione 1.10.19. Sia $n \in \mathbb{N}_+$. "Il" gruppo simmetrico $\mathcal{S}_n$ è $\mathcal{S}(X)$ dove X è un arbitrario insieme di cardinalità n. (Spesso si pone $X = \{1, \dots, n\}$.)

Siano G_1, G_2 gruppi. Sul prodotto cartesiano $G_1 \times G_2$ possiamo definire l'operazione

$$\begin{array}{ccc} (G_1 \times G_2) \times (G_1 \times G_2) & \longrightarrow & (G_1 \times G_2) \\ ((g_1, g_2), (h_1, h_2)) & \mapsto & (g_1 \cdot h_1, g_2 \cdot h_2) \end{array}$$

Con questa operazione $G_1 \times G_2$ è un gruppo. Infatti l'unità è l'elemnto $(1, 1)$, l'inverso di (g_1, g_2) è (g_1^{-1}, g_2^{-1}), e l'associatività segue dall'associatività del prodotto su G_1 e G_2. Il gruppo $G_1 \times G_2$ è il *prodotto diretto* di G_1 e G_2.

Esempio 1.10.20. Il gruppo $\mathbb{Z}/(2) \times \mathbb{Z}/(2)$ ha cardinalità 4, come il gruppo $\mathbb{Z}/(4)$, ma i due gruppi non sono isomorfi. Infatti supponiamo che $f \colon \mathbb{Z}/(4) \to \mathbb{Z}/(2) \times \mathbb{Z}/(2)$ sia un isomorfismo. Siccome

$$\mathbb{Z}/(4) = \{\overline{1},\ \overline{1} + \overline{1},\ \overline{1} + \overline{1} + \overline{1},\ \overline{1} + \overline{1} + \overline{1} + \overline{1}\},$$

(ricordate che l'operazione in $\mathbb{Z}/(4)$ è denotata $+$) allora abbiamo anche che

$$\mathbb{Z}/(2) \times \mathbb{Z}/(2) = \{f(\overline{1}),\ f(\overline{1}) + f(\overline{1}),\ f(\overline{1}) + f(\overline{1}) + f(\overline{1}),\ f(\overline{1}) + f(\overline{1}) + f(\overline{1}) + f(\overline{1})\}.$$

Ma, come si verifica facilmente, non esiste $(x, y) \in \mathbb{Z}/(2) \times \mathbb{Z}/(2)$ tale che

$$\mathbb{Z}/(2) \times \mathbb{Z}/(2) = \{(x, y),\ (x, y) + (x, y),\ (x, y) + (x, y) + (x, y),\ (x, y) + (x, y) + (x, y) + (x, y)\}.$$

Invece $\mathbb{Z}/(2) \times \mathbb{Z}/(3)$ è isomorfo a $\mathbb{Z}/(6)$, dimostratelo!

Esercizi del Capitolo 1

Esercizio 1.1. *Siano*

$$X_1 := \{0, 2, 4, 6, 8\}, \quad X_2 := \{1, 2, 4, 5, 6\}, \quad X_3 := \{0, 4, 8\}.$$

Determinate $X_i \cup X_j$ e $X_i \cap X_j$ per ogni $1 \leqslant i < j \leqslant 3$.

Esercizio 1.2. *Sia $\mathbb{N}_+ \subset \mathbb{N}$ il sottoinsieme dei naturali strettamente positivi. Dimostrate che*

$$\bigcup_{n \in \mathbb{N}_+} [-\frac{(n-1)}{n}, \frac{n-1}{n}] = (-1, 1), \qquad \bigcap_{n \in \mathbb{N}_+} (-\frac{(n+1)}{n}, \frac{n+1}{n}) = [-1, 1].$$

Esercizio 1.3. *Siano X, Y insiemi. Dimostrate che*

1. *$X \cup Y = Y$ se e solo se $X \subset Y$,*
2. *$X \cap Y = Y$ se e solo se $X \supset Y$.*

Esercizio 1.4. *Siano X, Y, Z insiemi. Dimostrate che*

$$X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z).$$

Esercizio 1.5. *Se X, Y sono insiemi $X \backslash Y$ è l'insieme i cui elementi sono gli $x \in X$ che **non** sono elementi di Y. Dimostrate che*

$$X \backslash (Y \cup Z) = (X \backslash Y) \cap (X \backslash Z), \qquad X \backslash (Y \cap Z) = (X \backslash Y) \cup (X \backslash Z).$$

(Formule di de Morgan.)

Se X è un insieme finito denoteremo con $|X|$ il numero degli elementi di X.

Esercizio 1.6. *Giustificate la notazione* (1.2.3) *dimostrando che se X e Y sono finiti allora*

$$|Y^X| = |Y|^{|X|}.$$

Sia X un insieme. Denotiamo $\mathcal{P}(X)$ l'insieme i cui elementi sono i sottoinsiemi di X, per esempio $\mathcal{P}(\{1,2\}) = \{\varnothing, \{1\}, \{2\}, \{1,2\}\}$. Sia $A \subset X$ un sottoinsieme. La *funzione caratteristica* di A è la $\chi_A \colon X \to \{0,1\}$ (dovremmo denotarla $\chi_{A,X}$) definita da

$$\chi_A(x) = \begin{cases} 1 & \text{se } x \in A, \\ 0 & \text{se } x \notin A. \end{cases} \tag{1.10.7}$$

Esercizio 1.7. *Di ciascuna delle seguenti funzioni dire se è iniettiva/suriettiva/biunivoca.*

$$\begin{array}{ccc} \mathbb{Z} & \xrightarrow{f} & \mathbb{Z} \\ x & \mapsto & |x| \end{array} \qquad \begin{array}{ccc} \mathbb{Z} & \xrightarrow{g} & \mathbb{N} \\ x & \mapsto & |x| \end{array} \qquad \begin{array}{ccc} \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}} & \xrightarrow{h} & \{0,1\}^{\mathbb{N}} \\ (\{a_n\}, \{b_n\}) & \mapsto & a_0, b_0, a_1, b_1, a_2, \dots \end{array}$$

Esercizio 1.8. *Determinate quali delle seguenti applicazioni è iniettiva/suriettiva, e determinate l'inversa di quelle che sono invertibili.*

1. $\alpha \colon \mathbb{Q} \to \mathbb{Q}$ *definita da* $\alpha(x) := x^3$.
2. $\beta \colon [n] \to [n+1]$ *definita da* $\beta(i) := i + 1$.
3. $\gamma \colon [n] \to [n]$ *definita da*
$$\gamma(i) := \begin{cases} i+1 & \text{if } 1 \leqslant i \leqslant (n-1), \\ 1 & \text{if } i = n. \end{cases}$$
4. $\delta \colon \mathbb{R} \to [2, +\infty)$ *definita da* $\delta(x) := e^x + e^{-x}$.
5. $\epsilon \colon \mathbb{R} \to \mathbb{R}$ *definita da* $\epsilon(x) := x^3 + x$.
6. $\zeta \colon \mathcal{P}(X) \to \mathcal{P}(X)$ *definita da* $\zeta(A) := A^c := X \backslash A$.

Esercizio 1.9. *Sia X un insieme. Dimostrate che la funzione*

$$\begin{array}{ccc} \mathcal{P}(X) & \xrightarrow{\varphi} & \{0,1\}^X \\ A & \mapsto & \chi_A \end{array}$$

è biunivoca. Dimostrate che se X è finito allora

$$|\mathcal{P}(X)| = 2^{|X|}.$$

Esercizio 1.10. *Sia $f \colon \mathbb{R} \to \mathbb{R}$ definita da $f(x) := x^2 + x + 3$. Determinate* $\operatorname{im} f$.

Esercizio 1.11. *Siano X, Y insiemi e $f \colon X \to Y$ un'applicazione. Siano $A \subset X$ e $B \subset Y$. Verificate che*

$$A \subset f^{-1}(f(A)), \qquad f(f^{-1}(B)) \subset B. \tag{1.10.8}$$

Date esempi in cui le inclusioni di (1.10.8) *sono strette, cioè $A \neq f^{-1}(f(A))$ e $f(f^{-1}(B)) \neq B$.*

Esercizio 1.12. *Sia X un insieme, e siano $f, g, h \in X^X$, cioè f, g, h sono applicazioni da X a X. Assumiamo che*

$$f \circ g = f \circ h.$$

(a) Mostrate con opportuni esempi che si può avere $g \neq h$.

(b) Dimostrate che se f è invertibile allora $g = h$.

Esercizio 1.13. *Sia X un insieme, e sia $f \colon X \to X$. Supponiamo che esistano $m, n \in \mathbb{N}$ <u>diversi</u> tali che*

$$f^m = f^n.$$

1. *Supponiamo che f sia invertibile. Dimostrate che esiste $a \in \mathbb{N}_+$ (cioè a è un numero naturale* positivo*) tale che*
$$f^a = 1_X. \tag{1.10.9}$$
Sia $\operatorname{ord}(f)$ *il minimo $a \in \mathbb{N}_+$ tale che valga l'uguaglianza in* (1.10.9) *(*$\operatorname{ord}(f)$ *è l'*ordine *di f). Dimostrate che se $a \in \mathbb{Z}$ e $f^a = 1_X$, allora a è un multiplo di* $\operatorname{ord}(f)$.

2. *Date un esempio di f non invertibile per la quale non esiste $a \in \mathbb{N}_+$ tale che valga l'uguaglianza in* (1.10.9).

Esercizio 1.14. *Sia X un insieme, e sia $f\colon X \to X$ invertibile. Definiamo la relazione $\mathcal{R}_f$ su X ponendo*

$$x\mathcal{R}_f y \text{ se esiste } m \in \mathbb{Z} \text{ tale che } x = f^m(y).$$

Dimostrate che $\mathcal{R}_f$ è una relazione di equivalenza.

Esercizio 1.15. *Sia $f\colon \{0,1\}^{\mathbb{Z}}$ (ricordiamo che $\{0,1\}^{\mathbb{Z}}$ è l'insieme i cui elementi sono le applicazioni da $\mathbb{Z}$ a $\{0,1\}$) definita da*

$$\begin{array}{ccc} \{0,1\}^{\mathbb{Z}} & \xrightarrow{f} & \{0,1\}^{\mathbb{Z}} \\ \varphi & \mapsto & m \mapsto \varphi(m-1) \end{array}$$

1. *Dimostrate che f è invertibile, e descrivete l'inversa di f.*
2. *Sia $\mathcal{R}_f$ la relazione di equivalenza dell'Esercizio 1.14. Descrivete gli elementi $\varphi \in \{0,1\}^{\mathbb{Z}}$ la cui classe di $\mathcal{R}_f$-equivalenza è finita.*

Esercizio 1.16. *Sia $f\colon \mathbb{R} \to \mathbb{R}$ definita da $f(x) = x^2 - 6x + 10$. Descrivete*

$$f([2,5]), \quad f^{-1}([0,1]), \quad f^{-1}([0,2]).$$

Esercizio 1.17. *Dimostrate che $\mathbb{N}$, $\mathbb{Z}$ e $\mathbb{Q}$ hanno la stessa cardinalità.*

Esercizio 1.18. *Sia X un insieme e $f\colon X \to \mathcal{P}(X)$ un'applicazione. Dimostrate che f non è suriettiva. (Suggerimento: dimostrate che $A := \{x \in X \mid x \notin f(x)\}$ non è un elemento dell'immagine di f.)*

Esercizio 1.19. *Un insieme X è* numerabile *se $\mathbb{N} \geq X$ cioè se X è finito oppure ha la cardinalità di $\mathbb{N}$. Dimostrate che $\mathbb{R}$* non *è numerabile.*

Esercizio 1.20. *Ridimostrate che vale la* (1.4.1) *osservando che*

$$(1 + 2 + \ldots + n) + (n + (n-1) + \ldots + 1) = n(n+1).$$

Esercizio 1.21. *Dimostrate per induzione che*

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

Esercizio 1.22. *Dimostrate per induzione che*

$$1^3 + 2^3 + \ldots + n^3 = \frac{1}{4} n^2 (n+1)^2. \tag{1.10.10}$$

Notate che per la (1.4.1) *la formula* (1.10.10) *equivale alla formula*

$$1^3 + 2^3 + \ldots + n^3 = (1 + 2 + \ldots + n)^2.$$

Esercizio 1.23. *Calcolate i seguenti numeri del campo $\mathbb{F}_5$:*

$$3 \cdot 4^{-1}, \qquad 3^5 \cdot 2^{-2}, \qquad (1 + 2 + \ldots + 9) \cdot 3^{-10}.$$

Esercizio 1.24. *Sia p un numero primo, e sia $x \in \mathbb{F}_p$ un elemento diverso da 0. Dimostrate che*

$$x^{p-1} = 1 \tag{1.10.11}$$

seguendo i seguenti passi. (Ricordiamo che (1.10.11) *significa che se $a \in \mathbb{Z}$* non *è multiplo di p, allora $a^{p-1} - 1$ è un multiplo do p.)*

(a) *Dimostrate che esiste $1 \leqslant m \leqslant (p-1)$ tale che*

$$x^m = 1 \tag{1.10.12}$$

considerando gli elementi $x, x^2, \ldots, x^{p-1}$. (Se non esiste tale m, "contando" vediamo che esistono $1 \leqslant i < j \leqslant (p-1)$ tali che $x^i = x^j$...)

(b) *Sia m il minimo $1 \leqslant m \leqslant (p-1)$ tale che valga* (1.10.12)*: dimostrate che $m|(p-1)$.*

(c) Dal punto (b) deducete che vale (1.10.11).

Esercizio 1.25. *Calcolate*

$$(1-3i)(5+2i), \qquad (1-i)^{-1}, \qquad (3+i)\cdot(1+i)^{-1}, \qquad (1+i)^{10}$$

Esercizio 1.26. *Calcolate le radici quadrate di* $2i$ *e di* $(1+\sqrt{3}i)$.

Esercizio 1.27. *Sia* G *un gruppo finito, cioè con un numero finito di elementi. Sia* $n := |G|$ *la cardinalità di* G. *Dimostrate che se* $x \in G$ *allora*

$$x^n = 1,$$

dove 1 *è l'unità di* G. *("Copiate" la dimostrazione dell'Esercizio 1.24.) Notate che da questo risultato si riottiene il risultato dell'Esercizio 1.24 perchè l'iniseme degli elementi non nulli di* $\mathbb{F}_p$ *con operazione la moltiplicazione è un gruppo di cardinalità* $(p-1)$.

Esercizio 1.28. *Sia* G *un gruppo finito, e supponiamo che la cardinalità di* G *sia un numero primo* p. *Dimostrate che* G *è isomorfo al gruppo* $\mathbb{Z}/(p)$ *(con operazione la somma). (Usate l'Esercizio 1.27.)*

Capitolo 2

Spazi vettoriali

2.1 Gli archetipi e la definizione

Siano $\mathbb{K}$ un campo e $X = (x_1, \ldots, , x_n)$, $Y = (y_1, \ldots, , y_n)$ elementi di $\mathbb{K}^n$: definiamo la *somma* $X + Y$ come l'elemento di $\mathbb{K}^n$ dato da

$$X + Y := (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n). \tag{2.1.1}$$

Quindi abbiamo un'operazione

$$\begin{array}{ccc} \mathbb{K}^n \times \mathbb{K}^n & \longrightarrow & \mathbb{K}^n \\ (X, Y) & \mapsto & X + Y \end{array} \tag{2.1.2}$$

Si definisce anche la moltiplicazione

$$\begin{array}{ccc} \mathbb{K} \times \mathbb{K}^n & \longrightarrow & \mathbb{K}^n \\ (\lambda, X) & \mapsto & \lambda X := (\lambda x_1, \lambda x_2, \ldots, \lambda x_n) \end{array} \tag{2.1.3}$$

Si usa chiamare λ uno *scalare* e quella definita è la moltiplicazione per scalari. Uno spazio vettoriale è un insieme V, fornito di due operazioni, la somma $V \times V \to V$, e il prodotto per scalari $\mathbb{K} \times V \to V$, che hanno caratteristiche simili a quelle della somma e prodotto per scalari di $\mathbb{K}^n$.

Definizione 2.1.1. Sia $\mathbb{K}$ un campo. Uno *spazio vettoriale su* $\mathbb{K}$ è un insieme V provvisto di due operazioni, la "somma"

$$\begin{array}{ccc} V \times V & \longrightarrow & V \\ (v_1, v_2) & \mapsto & v_1 + v_2 \end{array} \tag{2.1.4}$$

e la "moltiplicazione per scalare"

$$\begin{array}{ccc} \mathbb{K} \times V & \longrightarrow & V \\ (\lambda, v) & \mapsto & \lambda v \end{array} \tag{2.1.5}$$

tali che valgano le seguenti proprietà:

1. Esiste $0 \in V$ tale che $0 + v = v$ per ogni $v \in V$.
2. Dato $v \in V$ esiste $w \in V$ tale che $v + w = 0$, dove 0 è come al punto 1.
3. Se $u, v, w \in V$, allora $(u + v) + w = u + (v + w)$.
4. Se $v, w \in V$, allora $v + w = w + v$.
5. $1v = v$ per ogni $v \in V$,
6. $(\lambda + \mu)v = \lambda v + \mu v$ per ogni $v \in V$ e $\lambda, \mu \in \mathbb{K}$ (proprietà distributiva del prodotto),
7. $\lambda(v + w) = \lambda v + \lambda w$ per ogni $v, w \in V$ e $\lambda \in \mathbb{K}$ (proprietà distributiva della somma),

8. $(\lambda\mu)v = \lambda(\mu v)$ per ogni $v \in V$ e $\lambda, \mu \in \mathbb{K}$.

Gli elementi di uno spazio vettoriale si chiamano *vettori*, gli elementi del campo $\mathbb{K}$ gli *scalari*.

Osservazione 2.1.2. Sia V uno spazio vettoriale sul campo $\mathbb{K}$. Si denota con lo stesso simbolo 0 sia l'elemento neutro del campo $\mathbb{K}$, sia un elemento (vedremo che è unico) dello spazio vettoriale per cui valgano (1) e (2) della Definizione 2.1.1: attenzione a non fare confusione!

Esempio 2.1.3. Sia $V = \mathbb{K}^n$. Si verifica facilmente che le operazioni di somma e moltiplicazione per scalari definite da (2.1.2) e (2.1.3) rispettivamente godono delle proprietà della Definizione 2.1.1, con elemento neutro $0 := (0, 0, \ldots, 0)$. Quindi $\mathbb{K}^n$ provvisto delle operazioni appena definite è uno spazio vettoriale su $\mathbb{K}$.

Osservazione 2.1.4. Valgono le proprietà da 1 a 4 della Definizione 2.1.1 se e solo se V provvisto della somma è un gruppo abeliano.

Terminologia 2.1.5. Si dice informalmente che un insieme V *è uno spazio vettoriale* sottintendendo che sono definite operazioni di somma e moltiplicazione per scalare che godono delle proprietà elencate nella Definizione 2.1.1. Gli elementi di V si dicono *vettori*.

Terminologia 2.1.6. Uno spazio vettoriale *reale* è uno spazio vettoriale su $\mathbb{R}$, uno spazio vettoriale *complesso* è uno spazio vettoriale su $\mathbb{C}$.

Esempio 2.1.7. Sia $\mathbb{E}^2$ il piano della geometria euclidea (studiato a scuola). Siano $A \neq B \in \mathbb{E}^2$: denoteremo con $\overline{AB}$ l'unica retta contenente A e B. Ricordiamo che due rette sono *parallele* se hanno intersezione vuota oppure coincidono: se $A, B, C, D \in \mathbb{E}^2$ il simbolo $\overline{AB}||\overline{CD}$ significa che o $A \neq B$, $C \neq D$ e le rette AB, CD sono parallele oppure $A = B$ o $C = D$ ($A = B$ e $C = D$ ammesso).

Un *segmento orientato* in $\mathbb{E}^2$ è una coppia ordinata (A, B) di punti di $\mathbb{E}^2$: lo indichiamo con AB - l'estremo iniziale è A, quello finale è B (quindi $AB = CD$ se e solo se $A = C$ e $B = D$). I segmenti orientati AB e CD di $\mathbb{E}^2$ sono *equipollenti* se $\overline{AB}||\overline{CD}$ *e* $\overline{AC}||\overline{BD}$. Si verifica che la relazione di equipollenza è di equivalenza (esercizio); la denotiamo $\sim$.

Un *vettore geometrico* (nel piano) è una classe di equipollenza di segmenti orientati: quindi il quoziente

$$\mathsf{V}(\mathbb{E}^2) := (\mathbb{E}^2 \times \mathbb{E}^2)/\sim \tag{2.1.6}$$

è l'insieme dei vettori geometrici. La classe di equipollenza di AB si denota $\overrightarrow{AB}$. Notiamo che dato $P \in \mathbb{E}^2$ e un vettore geometrico v esiste uno e un solo $Q \in \mathbb{E}^2$ tale che $\overrightarrow{PQ} = v$.

Si dà all'insieme $\mathsf{V}(\mathbb{E}^2)$ la struttura di spazio vettoriale nel seguente modo. Prima definiamo la somma di segmenti orientati AB e BC (cioè tali che l'estremo finale del primo è l'estremo iniziale del secondo) come il segmenti orientato AC; quindi $AB + BC := AC$. Ora siano $v, w \in \mathsf{V}(\mathbb{E}^2)$ due classi di equipollenza di segmenti orientati. Sia AB un segmento orientato che rappresenta v e sia $C \in \mathbb{E}^2$ l'unico punto tale che BC rappresenti w: quindi ha senso $AB + BC = AC$. Si dimostra che se abbiamo punti $A', B', C' \in \mathbb{E}^2$ tali che $\overrightarrow{A'B'} = v$ e $\overrightarrow{B'C'} = w$ allora $A'B' + B'C' = A'C'$ è equipollente ad AC cioè $\overrightarrow{AC} = \overrightarrow{A'C'}$. Quindi possiamo definire la somma $v + w$ come la classe di equipollenza di $AB + BC = AC$: questo definisce la somma di vettori geometrici

$$\begin{array}{ccc} \mathsf{V}(\mathbb{E}^2) \times \mathsf{V}(\mathbb{E}^2) & \longrightarrow & \mathsf{V}(\mathbb{E}^2) \\ \overrightarrow{AB} + \overrightarrow{BC} & \mapsto & \overrightarrow{AC} \end{array}$$

La moltiplicazione per scalari si definisce in modo simile. Sia $v \in \mathsf{V}(\mathbb{E}^2)$. Supponiamo che $\lambda \in \mathbb{R}$ sia non-negativo. Sia AB un segmento orientato tale che $\overrightarrow{AB} = v$. Sia r una **semiretta** con estremo A e contenente B. Sia $C \in r$ il punto tale che la distanza da A a C sia la distanza da A a B moltiplicata per λ: si dimostra che la classe di equipollenza di AC non dipende dalla scelta del rappresentante di v e quindi possiamo definire λv come la classe di equipollenza di AC. Per definire λv quando $\lambda < 0$ definiamo l'*opposto* di un vettore geometrico v così: sia AB un rappresentante di v, allora la classe di equipollenza di BA non dipende dalla scelta del rappresentante e quindi ha senso definire $-v := \overrightarrow{BA}$.

Dato $v \in \mathsf{V}(\mathbb{E}^2)$ e $\lambda \in \mathbb{R}$ negativo definiamo $\lambda v := (-\lambda)v$ - questo ha senso perchè siccome $-\lambda > 0$ il vettore $(-\lambda)v$ è stato definito in precedenza. Ora definiamo il vettore nullo $0 \in \mathsf{V}(\mathbb{E}^2)$ come la classe di equipollenza di AA.

Si verifica che $\mathsf{V}(\mathbb{E}^2)$ con le operazioni appena definite è uno spazio vettoriale reale.

Esempio 2.1.8. Siccome $\mathbb{R}$ è un sottocampo di $\mathbb{C}$ possiamo dare a $\mathbb{C}$ la struttura di spazio vettoriale su $\mathbb{R}$.

Esempio 2.1.9. Sia $\mathbb{K}$ un campo. Sull'insieme dei polinomi $\mathbb{K}[x]$ sono definite le operazioni di somma e prodotto di polinomi. Siccome $\mathbb{K} \subset \mathbb{K}[x]$ (i polinomi "costanti"), possiamo definire un prodotto scalare $\mathbb{K} \times \mathbb{K}[x] \to \mathbb{K}[x]$. Con queste operazioni $\mathbb{K}[x]$ è un $\mathbb{K}$-spazio vettoriale.

Esempio 2.1.10. Siano $\mathbb{K}$ un campo e X un insieme. Possiamo dotare l'insieme $\mathbb{K}^X$ delle funzioni $f\colon X \to \mathbb{K}$ della struttura di un $\mathbb{K}$-spazio vettoriale definendo la somma di funzioni punto per punto e analogamente il prodotto per uno scalare:

$$(f+g)(x) := f(x) + g(x), \qquad (\lambda f)(x) := \lambda f(x).$$

L'elemento neutro è la funzione identicamente nulla.

Osservazione 2.1.11. Scegliamo una unità di misura nel piano euclideo $\mathbb{E}^2$. Allora, dato un vettore v nel piano euclideo $\mathsf{V}(\mathbb{E}^2)$, ha senso considerare la lunghezza di un qualsiasi rappresentante $\overrightarrow{AB}$ di v, e siccome tale lunghezza è indipendente dal rappresentante, ha senso parlare di lunghezza di v: si chiama la *norma* di v e si denota $||v||$. Di più: possiamo definire il prodotto scalare (v, w) di due vettori $v, w \in \mathsf{V}(\mathbb{E}^2)$, procedendo come fatto a scuola. Analogamente si può definire un prodotto scalare tra vettori di $\mathbb{R}^n$. Nella definizione di spazio vettoriale, dimentichiamo tutte queste strutture, benchè siano interessanti. Il punto è che per ora concentriamo la nostra attenzione su quello che si può dedurre dal solo fatto che siano definite le operazioni di somma di vettori e prodotto per uno scalare. In questo modo si dimostrano risultati che valgono in moltissimi contesti diversi. Più in là vedremo il prodotto scalare come una struttura aggiuntiva che uno spazio vettoriale può avere, e dimostreremo risultati sui prodotti scalari (e anche altre strutture aggiuntive). Questo modo di procedere non è naturale, ma economico e redditizio.

2.2 Prime proprietà

Proposizione 2.2.1. *Sia V uno spazio vettoriale su un campo $\mathbb{K}$.*

1. *Esiste un unico elemento neutro, cioè se $0_1, 0_2 \in V$ sono tali che*

$$0_1 + v = v, \qquad 0_2 + v = v \quad \forall v \in V \tag{2.2.1}$$

   allora $0_1 = 0_2$.

2. *Dato $v \in V$ esiste un unico $w \in V$ tale che $v + w = 0$.*

3. *$0v = 0$ (lo "0" a sinistra è l'elemento neutro di $\mathbb{K}$,*

Dimostrazione. La dimostrazione è del tutto simile alla dimostrazione dell'unicità dell'elemnto neutro di un anello, e dell'opposto di un elemento di un anello. Ripetiamola velocemente. Abbiamo

$$0_1 = 0_1 + 0_2 = 0_1, \tag{2.2.2}$$

e questo dimostra (1). Per dimostrare (2) supponiamo che $v + w_1 = 0$ e $v + w_2 = 0$. Abbiamo

$$w_1 = w_1 + 0 = w_1 + (v + w_2) = (w_1 + v) + w_2 = 0 + w_2 = w_2, \tag{2.2.3}$$

e questo dimostra (2). □

Terminologia 2.2.2. Per la Proposizione 2.2.1 ha senso parlare de*ll'*elemento neutro di $0 \in V$, e, dato $v \in V$, de*ll'*unico $w \in V$ tale che $v + w = 0$, che sarà denotato $-v$ (è l'*opposto* di v).

Proposizione 2.2.3. *Sia V uno spazio vettoriale su un campo $\mathbb{K}$. Allora*

(a) $0v = 0$ *per ogni* $v \in V$*,*

(b) $\lambda 0 = 0$ *per ogni* $\lambda \in \mathbb{K}$*,*

(c) $(-1)v + v = 0$ *per ogni* $v \in V$*.*

Dimostrazione. (a): Si ha $0v = (0 + 0)v = 0v + 0v$ e aggiungendo l'opposto di $0v$ a entrambi i membri otteniamo che $0 = 0v$. (b): la dimostrazione di è del tutto simile. (c): $(-1)v + v = (-1)v + 1v = (-1 + 1)v = 0v = 0$ dà che $(-1)v + v = 0$. Quindi $(-1)v = -v$. □

2.3 Sottospazi

Definizione 2.3.1. Sia V uno spazio vettoriale su $\mathbb{K}$. Un sottoinsieme W di V è un *sottospazio* di V se è non vuoto e valgono le seguenti proprietà:

(a) dati $v_1, v_2 \in W$, la somma $(v_1 + v_2)$ è in W,

(b) dati $v \in W$ e $\lambda \in \mathbb{K}$, il prodotto scalare λv è in W.

Osservazione 2.3.2. Sia V uno spazio vettoriale su un campo $\mathbb{K}$, e sia $W \subset V$ un sottospazio. Allora $0 \in W$ e, se $w \in W$, l'opposto $-w$ è un elemento di W. Infatti, siccome W non è vuoto, esiste $w_1 \in W$, quindi $0w_1 \in W$ per (b) della Definizione 2.3.1, ma $0w_1 = 0$ per la Proposizione 2.2.3. Questo dimostra che $0 \in W$. Siccome l'opposto $-w$ è uguale a $(-1)w$, è un elemento di W per (b) della Definizione 2.3.1.

La somma di V e il prodotto per scalare di V definiscono operazioni

$$\begin{array}{ccccccc} W \times W & \longrightarrow & W & \mathbb{K} \times W & \longrightarrow & W \\ (w_1, w_2) & \mapsto & w_1 + w_2 & (\lambda, w) & \mapsto & \lambda w \end{array} \tag{2.3.1}$$

Con queste operazioni W è uno spazio vettoriale su $\mathbb{K}$, con elemento neutro l'elemento neutro di V.

Esempio 2.3.3. Siano $\mathbb{K}$ un campo e $a_1, \ldots, a_n \in \mathbb{K}$. Siano

$$U := \{(x_1, \ldots, x_n) \in \mathbb{K}^n \mid a_1x_1 + \ldots + a_nx_n = 0\}, \tag{2.3.2}$$

$$W := \{(x_1, \ldots, x_n) \in \mathbb{K}^n \mid a_1x_1 + \ldots + a_nx_n = 1\}. \tag{2.3.3}$$

Verifichiamo che U è un sottospazio di $\mathbb{K}^n$ e che W non lo è. Innanzitutto U non è vuoto perchè $(0, \ldots, 0) \in U$. Se $X = (x_1, \ldots, x_n)$ e $Y = (y_1, \ldots, y_n)$ sono vettori di U, allora $X + Y = (x_1 + y_1, \ldots, x_n + y_n)$, e

$$a_1(x_1 + y_1) + \ldots + a_n(x_n + y_2) = a_1x_1 + \ldots + a_nx_n + a_1y_1 + \ldots + a_ny_n = 0 + 0 = 0.$$

Questo dimostra che vale (a) della Definizione 2.3.1. Inoltre, se $\lambda \in \mathbb{K}$ allora

$$a_1(\lambda x_1) + \ldots + a_n(\lambda x_n) = \lambda(a_1x_1 + \ldots + a_nx_n) = \lambda \cdot 0 = 0,$$

e questo dimostra che vale (b) della Definizione 2.3.1.

Per dimostrare che W non è un sottospazio di $\mathbb{K}^n$, osserviamo che $0 = (0, \ldots, 0)$ non è un elemento di W, e quindi W non è un sottospazio di $\mathbb{K}^n$ per l'Osservazione 2.3.2.

Esempio 2.3.4. L'insieme dei polinomi $\mathbb{R}[x]$, identificato con l'insieme delle funzioni polinomiali da $\mathbb{R}$ a $\mathbb{R}$, è un sottospazio dello spazio vettoriale delle funzioni da $\mathbb{R}$ a $\mathbb{R}$ con addizione e moltiplicazione per scalari puntuali.

Esempio 2.3.5. Siano $\mathbb{K}$ un campo, e $d \in \mathbb{N}$. Sia $\mathbb{K}[x]_{\leqslant d} \subset \mathbb{K}[x]$ definito da

$$\mathbb{K}[x]_{\leqslant d} := \{a_0 + a_1 x + \ldots + a_d x^d \mid a_0, a_1, \ldots, a_d \in \mathbb{K}\}. \tag{2.3.4}$$

$\mathbb{K}[x]_{\leqslant d}$ è un sottospazio vettoriale di $\mathbb{K}[x]$.

Esempio 2.3.6. I sottospazi $W \subset \mathsf{V}(\mathbb{E}^2)$ sono i seguenti:

(I) $W = \{0\}$,

(II) il sottospazio W_r dei vettori $\overline{PQ}$ paralleli a una retta fissata r, dove $\overline{PQ}$ parallelo a r significa che $P = Q$ oppure chee $P \neq Q$ è la retta per P e Q è parallela a r,

(III) $W = \mathsf{V}(\mathbb{E}^2)$.

Infatti sia $W \subset \mathsf{V}(\mathbb{E}^2)$ un sottospazio. Sappiamo che $0 \in W$ per l'Osservazione 2.3.2. Se 0 è l'unico elemento di W, allora vale (I). Ora supponiamo che esista $0 \neq v \in W$. Quindi $v = \overrightarrow{OP}$ per opportuni $O, P \in \mathbb{E}^2$. Sia r la retta per O e P. Allora W contiene il sottospazio W_r dei vettori paralleli a r. Se $W = W_r$ allora vale (II). Infine supponiamo che $W \supsetneq W_r$. Sia $w \in (W \backslash W_r)$. Esiste (un unico) $Q \in \mathbb{E}^2$ tale che $w = \overrightarrow{OQ}$. La retta s per O e Q non è parallela a r perchè $\overrightarrow{OQ}$ non è in W_r. Dimostriamo che $W = \mathsf{V}(\mathbb{E}^2)$. Sia $u \in \mathsf{V}(\mathbb{E}^2)$, e sia $R \in \mathbb{E}^2$ l'unico punto tale che $u = \overrightarrow{OR}$. Siano r' la retta per R parallela a r e s' la retta per R parallela a s. Quindi r' non è parallela a s, e analogamente s' non è parallela a r. Quindi l'intersezione $r' \cap s$ consiste di un singolo punto Q', e l'intersezione $s' \cap r$ consiste di un singolo punto P'. Si ha

$$u = \overrightarrow{OR} = \overrightarrow{OP'} + \overrightarrow{OQ'},$$

e siccome esistono $\lambda, \mu \in \mathbb{R}$ tali che $\overrightarrow{OP'} = \lambda v$ e $\overrightarrow{OQ'} = \mu w$ segue che $u \in W$.

Proposizione 2.3.7. *Sia V uno spazio vettoriale su $\mathbb{K}$ e W_i per $i \in I$ (I è un insieme di indici) una famiglia di sottospazi vettoriali di V. L'intersezione $\bigcap_{i\in I} W_i$ è un sottospazio vettoriale di V.*

Dimostrazione. Siccome $0 \in W_i$ per ogni $i \in I$ abbiamo che $0 \in \bigcap_{i\in I} W_i$ e quindi $\bigcap_{i\in I} W_i$ non è vuoto. Siano $v_1, v_2 \in \bigcap_{i\in I} W_i$ cioè $v_1, v_2 \in W_i$ per ogni $i \in I$, e sia $\lambda \in \mathbb{K}$. Siccome W_i è un sottospazio vettoriale di V abbiamo che $(v_1 + v_2) \in W_i$ e $\lambda v_1 \in W_i$ per ogni $i \in I$ e quindi $(v_1 + v_2) \in \bigcap_{i\in I} W_i$ e $\lambda v_1 \in \bigcap_{i\in I} W_i$. □

Esempio 2.3.8. Applichiamo la Proposizione 2.3.7 all'insieme delle soluzioni di un sistema di *equazioni lineari omogenee* cioè l'insieme degli $(x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$ tali che

$$\begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots & = & 0, \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n & = & 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots & = & 0, \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & 0. \end{array} \tag{2.3.5}$$

Siccome le soluzioni di una singola equazione

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = 0$$

è un sottospazio vettoriale di $\mathbb{K}^n$ l'insieme delle soluzioni di m equazioni è l'intersezione di m sottospazi vettoriali di $\mathbb{K}^n$; per la Proposizione 2.3.7 è un sottospazio vettoriale di $\mathbb{K}^n$.

2.4 Combinazioni lineari

Definizione 2.4.1. Sia V uno spazio vettoriale su $\mathbb{K}$. Siano $v_1, v_2, \ldots, v_n \in V$. Un vettore $v \in V$ è *combinazione lineare* di $v_1, \ldots, v_n$ se esistono $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ tali che

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n. \tag{2.4.1}$$

È conveniente ammettere che n possa essere 0 cioè la collezione di vettori sia vuota: dichiariamo che solo 0 è combinazione lineare di una collezione vuota di vettori.

Esempio 2.4.2. Sia $\mathbb{K}$ un campo, e siano $e_1, e_2, \ldots, e_n \in \mathbb{K}^n$ definiti da

$$\mathbf{e}_1 := (1, 0, \ldots, 0),\ \mathbf{e}_2 := (0, 1, 0, \ldots, 0), \ldots, \mathbf{e}_n := (0, 0, \ldots, 1). \tag{2.4.2}$$

Ogni $v \in \mathbb{K}^n$ è combinazione lineare di $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Infatti

$$(x_1, x_2, \ldots, x_n) = x_1\mathbf{e}_1 : + x_2\mathbf{e}_2 : + \ldots + x_n\mathbf{e}_n. \tag{2.4.3}$$

Esempio 2.4.3. Sia $\mathbb{K}$ un campo, e sia $v = (1, 1, 1) \in \mathbb{K}^3$. Allora v non è combinazione lineare di e_1, e_2 (definiti come nell'Esempio 2.4.2). Infatti ogni combinazione lineare di e_1, e_2 ha l'ultima entrata uguale a 0.

Lemma 2.4.4. *Siano $\mathbb{K}$ un campo e V uno spazio vettoriale su $\mathbb{K}$. Un sottoinsieme $W \subset V$ è un sottospazio se e solo se contiene le combinazioni lineari di qualsiasi lista di suoi vettori.*

Dimostrazione. Supponiamo che $W \subset V$ sia un sottospazio. Siano $v_1, v_2, \ldots, v_n \in V$ e $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$. Per (b) della Definizione 2.3.1 si ha che $\lambda_i v_i$ è in W per $i \in \{1, \ldots, n\}$, e quindi $\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n$ è in W per (a) della Definizione 2.3.1. Ora supponiamo che $W \subset V$ sia un sottoinsieme tale che le combinazioni lineari di qualsiasi sua lista di vettori siano in W. Siccome 0 è combinazione lineare della lista vuota che è una lista di vettori in W, segue che $0 \in W$ e quindi W non è vuoto. Se $v_1, v_2 \in W$ allora $(1 \cdot v_1 + 1 \cdot v_2) \in W$, cioè vale (a) della Definizione 2.3.1 e, analogamente, vale (b) della Definizione 2.3.1. □

Proposizione 2.4.5. *Sia V uno spazio vettoriale su $\mathbb{K}$ e $S \subset V$ un sottoinsieme. Poniamo*

$$\langle S \rangle = \{\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n \mid v_1, v_2, \ldots, v_n \in S, \quad \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}\}. \tag{2.4.4}$$

Allora

1. *$\langle S \rangle$ è un sottospazio vettoriale di V che contiene S, e*
2. *se W è un sottospazio vettoriale di V che contiene S allora $\langle S \rangle \subset W$.*

(Informalmente $\langle S \rangle$ è il più piccolo *sottospazio vettoriale di V che contiene S).*

Dimostrazione. Siccome $0 \in \langle S \rangle$ (vedi l'ultima frase della Definizione 2.4.1) $\langle S \rangle$ non è vuoto. Se $a, b \in \langle S \rangle$, cioè

$$a = \sum_{i=1}^{m} \lambda_i v_i, \quad b = \sum_{j=1}^{n} \mu_j w_j, \qquad v_i, w_j \in S,\ \lambda_i, \mu_j \in \mathbb{K}$$

allora $a + b = \sum_{i=1}^{m} \lambda_i v_i + \sum_{j=1}^{n} \mu_j w_j$, e quindi $(a + b) \in \langle S \rangle$. Inoltre, se $\theta \in \mathbb{K}$ abbiamo che $\theta a = \sum_{i=1}^{n} (\theta\lambda_i) v_i$, e quindi $\theta a \in \langle S \rangle$. Questo dimostra che $\langle S \rangle$ è un sottospazio vettoriale di V. Se $v \in S$, allora $v = 1 \cdot v$, e quindi $v \in \langle S \rangle$. Questo dimostra che vale (1).

Se $W \subset V$ è un sottospazio che contiene S, allora $\langle S \rangle$ è contenuto in W per il Lemma 2.4.4, quindi vale (2). □

Terminologia 2.4.6. Sia V uno spazio vettoriale sul campo $\mathbb{K}$. Sia $S \subset V$ un sottoinsieme. Il sottospazio $\langle S \rangle$ della Proposizione 2.4.5 è il sottospazio vettoriale *generato da* S, e si denota anche $\mathrm{Span}(S)$. Se $v_1, \ldots, v_n \in V$ omettiamo le parentesi graffe:

$$\langle v_1, \ldots, v_n \rangle = \langle \{v_1, \ldots, v_n\} \rangle = \mathrm{Span}(v_1, \ldots, v_n).$$

Definizione 2.4.7. Sia V uno spazio vettoriale e $U, W \subset V$ sottospazi. La *somma* $U + W$ è il sottospazio di V definito da

$$U + W := \langle U \cup W \rangle = \{u + w \mid u \in U, \quad w \in W\}. \tag{2.4.5}$$

(Notate che in generale l'unione $U \cup W$ **non** è un sottospazio.)

Definizione 2.4.8. Uno spazio vettoriale su un campo $\mathbb{K}$ è *finitamente generato* se è generato da un insieme finito. Un sottospazio di uno spazio vettoriale su un campo $\mathbb{K}$ è *finitamente generato* se è finitamente generato come spazio vettoriale su $\mathbb{K}$ (vedi l'Osservazione 2.3.2).

Esempio 2.4.9. Sia $\mathbb{K}$ un campo. Lo spazio vettoriale $\mathbb{K}^n$ è finitamente generato su $\mathbb{K}$ perchè è generato dai vettori $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ (vedi l'Osservazione 2.4.2).

Esempio 2.4.10. Sia $\mathbb{K}$ un campo. Allora lo spazio vettoriale (su $\mathbb{K}$) $\mathbb{K}[x]$ *non* è finitamente generato. Infatti, siano $f_1, \ldots, f_m \in \mathbb{K}[x]$, e dimostriamo che $\langle f_1, \ldots, f_m \rangle \neq \mathbb{K}[x]$. Possiamo assumere che i polinomi $f_1, \ldots, f_m$ siano tutti non nulli perchè il sottospazio generato non cambia se scartiamo eventuali polinomi nulli. Ogni $f \in \langle f_1, \ldots, f_m \rangle$ non nullo ha grado al più uguale al massimo dei gradi degli f_j e quindi $\langle f_1, \ldots, f_m \rangle$ non è tutto $\mathbb{K}[x]$ perchè esistono polinomi di grado arbitrariamente alto.

Esempio 2.4.11. Sia $\mathbb{K}$ un campo. Allora $\mathbb{K}[x]_{\leqslant d}$ è un sottospazio finitamente generato di $\mathbb{K}[x]$, perchè è generato da $\{1, x, \ldots, x^d\}$.

Proposizione 2.4.12. *Sia V uno spazio vettoriale su $\mathbb{K}$ finitamente generato. Ogni sottospazio vettoriale di V è finitamente generato su $\mathbb{K}$.*

Dimostrazione. Sia P_n l'affermazione: *Se V ha n generatori e $W \subset V$ è un sottospazio vettoriale allora W è generato da un insieme con al più n elementi.*

Dimostriamo per induzione che P_n è vera per ogni $n \in \mathbb{N}$. Se V ha 0 generatori allora $V = \{0\}$ (per convenzione il sottospazio generato dall'insieme vuoto è $\{0\}$) e quindi $\{0\}$ è l'unico sottospazio di V: segue che P_0 è vera. Se non vi piace il caso $n = 0$ potete verificare che vale P_1: in questo caso esiste $v \in V$ tale che $V = \langle v \rangle$. Se $W = \{0\}$ allora W è generato da un insieme con 0 elementi, se $W \neq \{0\}$ esiste $av \in W$ dove $a \neq 0$, quindi $v = a^{-1}(av) \in W$ e perciò $W = V = \langle v \rangle$ e quindi W è generato dall'insieme $\{v\}$ che un elemento.

Ora dimostriamo il passo induttivo cioè supponiamo che valga P_n e dimostriamo che vale P_{n+1}. Siano $v_1, \ldots, v_{n+1}$ generatori di V cioè $V = \langle v_1, \ldots, v_{n+1} \rangle$. Sia

$$U := \langle v_1, \ldots, v_n \rangle.$$

L'intersezione $W \cap U$ è un sottospazio vettoriale di V per la Proposizione 2.3.7 e quindi è un sottospazio vettoriale di U. Siccome U è generato da n vettori e siccome per ipotesi induttiva vale P_n segue che esistono $w_1, \ldots, w_k \in W \cap U$ con $k \leqslant n$ e tali che $W \cap U = \langle w_1, \ldots, w_k \rangle$. Se $W \subset U$ allora $W \cap U = W$ e abbiamo fatto. Rimane da esaminare il caso in cui $W \not\subset U$. Dunque esiste $\widehat{w} \in (W \backslash U)$. Per ipotesi esistono $a_1, \ldots, a_{n+1} \in \mathbb{K}$ tali che

$$\widehat{w} = a_1 v_1 + a_2 v_2 + \ldots + a_{n+1} v_{n+1},$$

e siccome $w \notin U$ abbiamo che $a_{n+1} \neq 0$. Anche il vettore $\overline{w} := a_{n+1}^{-1} \widehat{w}$ appartiene a W, e abbiamo

$$\overline{w} = b_1 v_1 + b_2 v_2 + \ldots + b_n v_n + v_{n+1}.$$

Dimostriamo che

$$W = \langle w_1, w_2, \ldots, w_k, \overline{w} \rangle. \tag{2.4.6}$$

Sia $w \in W$. Per ipotesi esistono $x_1, \ldots, x_{n+1} \in \mathbb{K}$ tali che $w = x_1 v_1 + x_2 v_2 + \ldots + x_{n+1} v_{n+1}$. Il vettore

$$w - x_{n+1}\overline{w} = (x_1 - b_1 x_{n+1})v_1 + (x_2 - b_2 x_{n+1})v_2 + \ldots + \ldots (x_n - b_n x_{n+1})v_n \tag{2.4.7}$$

è in $W \cap U$. Infatti è in W perchè è combinazione lineare dei vettori w e $\overline{w}$, che appartengono al sottospazio W, e l'espressione a destra di (2.4.7) dimostra che è in U. Quindi esistono $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ tali che

$$w - x_{n+1}\overline{w} = \lambda_1 w_1 + \lambda_2 w_2 + \ldots + \ldots \lambda_k w_k.$$

Aggiungendo $x_{n+1}\overline{w}$ a entrambi i mebri vediamo che w è una combinazione lineare di $w_1, \ldots, w_k, \overline{w}$. Questo dimostra che vale (2.4.6). Siccome $k \leqslant n$, abbiamo dimostrato che W è generato da un insieme con al più $n + 1$ elementi. □

Esempio 2.4.13. Sia $\mathbb{K}$ un campo. Ogni sottospazio di $\mathbb{K}^n$ è finitamente generato perchè $\mathbb{K}^n$ è finitamente generato.

2.5 Dipendenza/indipendenza lineare

La seguente è una definizione fondamentale.

Definizione 2.5.1. Sia V uno spazio vettoriale su $\mathbb{K}$. Una relazione lineare tra $v_1, \ldots, v_n \in V$ è una uguaglianza

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0, \tag{2.5.1}$$

dove $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$. La relazione lineare (2.5.1) è *non banale* se $\lambda_1, \lambda_2, \ldots, \lambda_n$ *non* sono tutti nulli.

Se V è uno spazio vettoriale su $\mathbb{K}$, una *lista* di vettori in V è un'applicazione $f\colon I \to V$, dove I è un insieme. In generale denotiamo il vettore $f(i)$ che corrisponde a $i \in I$ con v_i, e la lista con $\{v_i\}_{i \in I}$. (Pensiamo agli elementi di I come "indici".) Se $I = \{1, \ldots, n\}$, denotiamo una lista $I \to V$ con $v_1, \ldots, v_n$.

Definizione 2.5.2. Sia V uno spazio vettoriale su $\mathbb{K}$, e sia $\{v_i\}_{i \in I}$ una lista di vettori in V. I vettori di $\{v_i\}_{i \in I}$ sono *linearmente dipendenti* se esistono indici *distinti* $i_1, \ldots, i_n \in I$ con la proprietà che esiste una relazione lineare non banale tra $v_{i_1}, \ldots, v_{i_n} \in V$. I vettori della lista $\{v_i\}_{i \in I}$ sono *linearmente indipendenti* in caso contrario (cioè se non sono lineamente dipendenti).

Esplicitiamo la definizione di vettori linearmente dipendenti/indipendenti nel caso di una lista finita $v_1, \ldots, v_n$. I vettori sono linearmente dipendenti se esiste una relazione lineare non banale tra $v_1, \ldots, v_n$, e invece sono indipendenti se

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0$$

vale solo per $0 = \lambda_1 = \lambda_2 = \ldots = \lambda_n$.

Esempio 2.5.3. I vettori $\mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathbb{K}^n$ definiti da (2.4.2) sono linearmente indipendenti, i vettori $v_1 = (2, 2)$, $v_2 = (3, 3)$ di $\mathbb{R}^2$ sono linearmente dipendenti perchè $3v_1 - 2v_2 = 0$.

Esempio 2.5.4. Nello spazio vettoriale $\mathbb{K}[x]$ i vettori $1, x, x^2, \ldots, x^n, \ldots$, cioè i vettori della lista $\{x^i\}_{i \in \mathbb{N}}$ sono lineamente indipendenti.

Esempio 2.5.5. Siano $v_1, v_2 \in \mathbb{K}^2$ dati da $v_1 = (a, b)$ e $v_2 = (c, d)$. Allora v_1, v_2 sono linearmente dipendenti se e solo se $(ad - bc) = 0$. Infatti supponiamo che

$$x_1 v_1 + x_2 v_2 = 0$$

cioè

$$ax_1 + cx_2 = 0, \qquad bx_1 + dx_2 = 0. \tag{2.5.2}$$

Moltiplicando la prima equazione per b e aggiungendogli la seconda equazione moltiplicata per $-a$ otteniamo che

$$(bc - ad)x_2 = 0. \tag{2.5.3}$$

D'altra parte moltiplicando la prima equazione di (2.5.2) per d e aggiungendogli la seconda equazione moltiplicata per $-c$ otteniamo che

$$(ad - bc)x_1 = 0. \tag{2.5.4}$$

Segue che se v_1, v_2 sono linearmente dipendenti allora $(ad - bc) = 0$: infatti esiste una soluzione non banale (x_1, x_2) di (2.5.2) e per (2.5.3) e (2.5.4) segue che $(ad - bc) = 0$. Ora dimostriamo che se $(ad - bc) = 0$ allora v_1, v_2 sono linearmente dipendenti. Se $0 = a = b = c = d$ cioè $(0, 0) = v_1 = v_2$ non c'è nulla da dire (abbiamo per esempio che $1 \cdot v_1 + 0 \cdot v_2 = 0$). Quindi possiamo supporre che $(a, c) \neq (0, 0)$ o $(b, d) \neq (0, 0)$. Nel primo caso una soluzione non banale di (2.5.2) è data da $x_1 = c$, $x_2 = -a$, nel secondo caso una soluzione non banale di (2.5.2) è data da $x_1 = d$, $x_2 = -b$.

Definizione 2.5.6. Una *matrice* 2×2 con entrate in un campo $\mathbb{K}$ è una collezione ordinata M di 4 elementi di $\mathbb{K}$, diciamo a, b, c, d. Scriviamo la matrice così:

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Le *righe* di M sono (a, b) e (c, d) rispettivamente, le sue colonne sono (a, c) e (b, d) rispettivamente. Il *determinante* di M è il numero

$$\det M := (ad - bc). \tag{2.5.5}$$

Osservazione 2.5.7. L'esempio 2.5.5 dà un caso in cui è utile disporre della nozione di matrice 2×2 e suo determinante: infatti abbiamo visto che i vettori $v_1, v_2 \in \mathbb{K}^2$ sono linearmente dipendenti se e solo se è nullo il determinante della matrice 2×2 che ha come righe (o colonne) i vettori v_1 e v_2. Nel Capitolo 3 considereremo matrici di ordine qualsiasi, e nel Capitolo 5 definiremo il determinante di matrici quadrate di ordine arbitrario.

Osservazione 2.5.8. Nella definizione di vettori linearmente dipendenti, i vettori $v_1, \ldots, v_n$ sono una *lista* di vettori, cioè un'applicazione da $\{1, 2, \ldots, n\} \to V$, e *non* un insieme. Quindi può accadere che $v_i = v_j$ per $i \neq j$, e in tal caso i vettori $v_1, \ldots, v_n$ sono linearmente dipendenti.

Proposizione 2.5.9. *Sia V uno spazio vettoriale su $\mathbb{K}$. I vettori $v_1, \ldots, v_n \in V$ sono linearmente dipendenti se e solo se esiste $1 \leqslant i \leqslant n$ tale che v_i è nel sottospazio generato dai restanti vettori, cioè*

$$v_i \in \langle v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \rangle. \tag{2.5.6}$$

Dimostrazione. Supponiamo che $v_1, \ldots, v_n$ siano linearmente dipendenti cioè esistono $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ non tutti nulli tali che valga (2.5.1). Quindi esiste $1 \leqslant i \leqslant n$ tale che $\lambda_i \neq 0$. Moltiplicando entrambi i membri di (2.5.1) per λ_i^{-1} otteniamo che

$$\lambda_i^{-1}\lambda_1 v_1 + \ldots + \lambda_i^{-1}\lambda_{i-1} v_{i-1} + v_i + \lambda_i^{-1}\lambda_{i+1} v_{i+1} + \ldots + \lambda_i^{-1}\lambda_n v_n = 0 \tag{2.5.7}$$

e dunque

$$v_i = -\lambda_i^{-1}\lambda_1 v_1 - \ldots - \lambda_i^{-1}\lambda_{i-1} v_{i-1} - \lambda_i^{-1}\lambda_{i+1} v_{i+1} - \ldots - \lambda_i^{-1}\lambda_n v_n. \tag{2.5.8}$$

Questo dimostra che $v_i \in \langle v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \rangle$. Ora supponiamo che valga (2.5.6) cioè esistono $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n \in \mathbb{K}$ tali che $v_i = a_1 v_1 + a_2 v_2 + \ldots + a_{i-1} v_{i-1} + a_{i+1} v_{i+1} + \ldots + a_n v_n$: allora abbiamo che

$$a_1 v_1 + a_2 v_2 + \ldots + a_{i-1} v_{i-1} - v_i + a_{i+1} v_{i+1} + \ldots + a_n v_n = 0 \tag{2.5.9}$$

e quindi $v_1, v_2, \ldots, v_n$ sono linearmente dipendenti. □

Osservazione 2.5.10. L'affermazione "i vettori $v_1, \ldots, v_n$ sono linearmente dipendenti" è un'affermazione sulla *lista* di vettori $v_1, \ldots, v_n$, *non* si afferma che *ciascun* vettore della lista $v_1, \ldots, v_n$ ha la proprietà di essere "linearmente indipendente". Questa è un'osservazione banale, ma esiste il pericolo di fraintendimento perchè a rigore bisognerebbe affermare che "la lista $v_1, \ldots, v_n$ è linearmente dipendente". Ovviamente, analoghe considerazione valgono per l'affermazione "i vettori $v_1, \ldots, v_n$ sono linearmente indipendenti".

2.6 Basi

La seguente è una definizione fondamentale.

Definizione 2.6.1. Sia V uno spazio vettoriale su $\mathbb{K}$. Una lista $\{v_i\}_{i\in I}$ di vettori di V è una *base* di V se i vettori della lista sono linearmente indipendenti e V è generato dai vettori della lista.

Esempio 2.6.2. I vettori $\mathbf{e}_1, \dots, \mathbf{e}_n \in \mathbb{K}^n$ definiti da (2.4.2) formano una base di $\mathbb{K}^n$: questa è la *base standard* di $\mathbb{K}^n$.

Esempio 2.6.3. Una base dello spazio vettoriale $\mathbb{K}[x]$ è data dalla lista $\{x^i\}_{i\in\mathbb{N}}$.

Proposizione 2.6.4. *Sia V uno spazio vettoriale su $\mathbb{K}$ generato dai vettori $v_1, \dots, v_n$. Allora esiste una base di V ottenuta eliminando alcuni dei v_i, cioè esistono $1 \leqslant i_1 < i_2 < \ldots < i_m \leqslant n$ tali che $\{v_{i_1}, v_{i_2}, \dots, v_{i_m}\}$ sia una base di V.*

Dimostrazione. Esiste un sottoinsieme *minimale* $\{i_1, \dots, i_m\} \subset \{1, \dots, n\}$ tale che i vettori $V = \mathrm{Span}(v_{i_1}, \dots, v_{i_m})$ siano linearmente indipendenti, dove minimale significa che per ogni

$$i \in \{i_1, \dots, i_m\} \tag{2.6.1}$$

i vettori indicizzati da $\{i_1, \dots, i_m\}\backslash\{i\}$ non generano V. Dimostriamo che $\{v_{i_1}, \dots, v_{i_m}\}$ è una base di V. Per semplificare la notazione assumiamo che $i_1 = 1, \dots, i_m = m$ (è lecito assumerlo, basta riordinare gli indici). Siccome $v_1, \dots, v_m$ generano V per definizione, rimane da dimostrare che $v_1, \dots, v_m$ sono linearmente indipendenti. Supponiamo il contrario, cioè che esista una relazione di dipendenza lineare

$$\lambda_1 v_1 + \ldots + \lambda_m v_m = 0.$$

Per la Proposizione 2.5.9 esistono $i \in \{1, \dots, m\}$ e $\mu_1, \dots, \mu_{i-1}, \mu_{i+1}, \dots, \mu_m \in \mathbb{K}$ tali che

$$v_i = \mu_1 v_1 + \ldots + \mu_{i-1} v_{i-1} + \mu_{i+1} v_{i+1} + \ldots + v_m. \tag{2.6.2}$$

Segue da questo che V è generato da $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_m$. Infatti sia $v \in V$; siccome $v_1, \dots, v_m$ generano V, esistono $\theta_1, \dots, \theta_i, \dots, \theta_m \in \mathbb{K}$ tali che

$$v = \theta_1 v_1 + \ldots + \theta_i v_i + \theta_m v_m.$$

Sostituendo v_i con la sua espressione come combinazione lineare di $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_m$ (vedi (2.6.2)), otteniamo che v è combinazione lineare di $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_m$. Abbiamo dimostrato che se $v_1, \dots, v_m$ sono linearmente dipendenti allora V è generato da $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_m$. Siccome questo contraddice la minimalità di $v_1, \dots, v_m$, segue che $v_1, \dots, v_m$ sono linearmente indipendenti. □

Dalla Proposizione 2.6.4 segue subito il seguente corollario.

Corollario 2.6.5. *Uno spazio vettoriale finitamente generato ha una base.*

Il seguente risultato è l'analogo della Proposizione 2.6.4 ottenuto sostituendo "sistema di generatori" con "lista di vettori linearmente indipendenti".

Proposizione 2.6.6. *Sia V uno spazio vettoriale su $\mathbb{K}$ finitamente generato. Siano $v_1, \dots, v_m \in V$ linearmente indipendenti: esistono $v_{m+1}, \dots, v_{m+q} \in V$ tali che $\{v_1, \dots, v_m, v_{m+1}, \dots, v_{m+q}\}$ sia una base di V. (Il caso $q = 0$ è ammesso: significa che $\{v_1, \dots, v_m\}$ è una base di V.)*

Dimostrazione. La dimostrazione è analoga a quella della Proposizione 2.6.4. Siccome V è finitamente generato esistono $u_1, \dots, u_l \in V$ tali che V sia generato da $v_1, \dots, v_m, u_1, \dots, u_l$. Siccome $v_1, \dots, v_m \in V$ sono linearmente indipendenti esiste un sottoinsieme *massimale* $\{j_1, \dots, j_q\} \subset \{1, \dots, l\}$ tale che $v_1, \dots, v_m, u_{j_1}, \dots, u_{j_q}$ sono linearmente indipendenti, cioè tale che per ogni

$$j \in \{1, \dots, l\}\backslash\{j_1, \dots, j_q\} \tag{2.6.3}$$

i vettori $v_1, \ldots, v_m, u_{j_1}, \ldots, u_{j_q}, u_j$ sono linearmente dipendenti. Dimostriamo che

$$\{v_1, \ldots, v_m, u_{j_1}, \ldots, u_{j_q}\} \tag{2.6.4}$$

è una base di V. Basta dimostrare che V è generato dai vettori di (2.6.4). Sia j tale che valga (2.6.3); per ipotesi esiste una relazione lineare non banale

$$\alpha_1 v_1 + \ldots + \alpha_m v_m + \beta_1 u_{j_1} + \ldots + \beta_q u_{j_q} + \gamma u_j = 0. \tag{2.6.5}$$

Se fosse $\gamma = 0$, allora $0 = \alpha_1 = \ldots = \alpha_m = \beta_1 = \ldots = \beta_q$ perchè $v_1, \ldots, v_m, u_{j_1}, \ldots, u_{j_q}$ sono linearmente indipendenti, ma questa è una contraddizione perchè (2.6.5) è una relazione non banale. Quindi $\gamma \neq 0$; moltiplicando (2.6.5) per γ^{-1} vediamo che u_j è nel sottospazio $\langle v_1, \ldots, v_m, u_{j_1}, \ldots, u_{j_q}\rangle$. Questo vale per ogni j tale che valga (2.6.3); siccome V è generato da $v_1, \ldots, v_m, u_1, \ldots, u_l$ segue che $V = \langle v_1, \ldots, v_m, u_{j_1}, \ldots, u_{j_q}\rangle$. □

Siano $v_1, \ldots, v_m \in V$ e $\mathcal{B} := \{v_1, \ldots, v_m, v_{m+1}, \ldots, v_{m+q}\}$ come nell'enunciato della 2.6.6: si dice che $v_1, \ldots, v_m \in V$ si *estende* alla base $\mathcal{B}$ di V. Quindi la proposizione afferma che in uno spazio vettoriale finitamente generato ogni lista di vettori linearmente indipendenti si estende a una base.

Dati $v_1, \ldots, v_n \in V$ definiamo l'applicazione

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{f} & V \\ (x_1, \ldots, x_n) & \longrightarrow & x_1 v_1 + x_2 v_2 + \ldots + x_n v_n \end{array} \tag{2.6.6}$$

Proposizione 2.6.7. *Sia V uno spazio vettoriale su $\mathbb{K}$ e $v_1, \ldots, v_n \in V$.*

(a) $v_1, \ldots, v_n$ generano V se e solo se l'applicazione f in (2.6.6) *è suriettiva, e*

(b) $v_1, \ldots, v_n$ sono linearmente indipendenti se e solo se l'applicazione f in (2.6.6) *è iniettiva.*

Dimostrazione. (a) vale per definizione. Dimostriamo che vale (b). Supponiamo che $f(x_1, x_2, \ldots, x_n) = f(y_1, y_2, \ldots, y_n)$, cioè

$$x_1 v_1 + x_2 v_2 + \ldots + x_n v_n = y_1 v_1 + y_2 v_2 + \ldots + y_n v_n.$$

Aggiungendo l'opposto del membro di destra a entrambi i membri troviamo che

$$(x_1 - y_1)v_1 + (x_2 - y_2)v_2 + \ldots + (x_n - y_n)v_n = 0,$$

Siccome $v_1, \ldots, v_n$ sono linearmente indipendenti segue che $x_i = y_i$ per ogni i. □

Corollario 2.6.8. *Sia V uno spazio vettoriale su $\mathbb{K}$ e $v_1, \ldots, v_n \in V$. Allora $\{v_1, \ldots, v_n\}$ è una base di V se e solo se l'applicazione f in* (2.6.6) *è biunivoca.*

Sia $\mathcal{B} := \{v_1, \ldots, v_n\}$ una base di V, e sia f data da (2.6.6): per il Corollario 2.6.8 la f è biunivoca e quindi è definita la sua inversa che denotiamo $X_{\mathcal{B}}$:

$$V \xrightarrow{X_{\mathcal{B}}} \mathbb{K}^n \tag{2.6.7}$$

Definizione 2.6.9. La n-pla delle *coordinate* di $v \in V$ è $X_{\mathcal{B}}(v)$.

Esempio 2.6.10. Sia $\mathcal{S} := \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ la base standard di $\mathbb{K}^n$. Le coordinate di $X = (x_1, \ldots, x_n)$ nella base $\mathcal{S}$ sono date da $(x_1, \ldots, x_n)$.

Esempio 2.6.11. Una base di $\mathbb{R}^2$ è $\mathcal{B} := \{(1,1), (1,-1)\}$ (verificatelo). Se (t_1, t_2) sono le coordinate di $X = (x_1, x_2) \in \mathbb{R}^2$ nella base $\mathcal{B}$, allora

$$(x_1, x_2) = t_1(1,1) + t_2(1,-1).$$

Quindi, per determinare (t_1, t_2) risolviamo il sistema di equazioni lineari

$$t_1 + t_2 = x_1, \qquad t_1 - t_2 = x_2.$$

Semplici calcoli danno che $t_1 = (x_1 + x_2)/2$ e $t_2 = (x_1 - x_2)/2$. Notate che le coordinate di X nella base $\mathcal{B}$ sono completamente diverse da quelle nella base standard $\mathcal{S}$.

Un fatto fondamentale è che il numero di elementi in una base di uno spazio vettoriale è indipendente dalla base. Questo risultato sarà una conseguenza del seguente lemma.

Lemma 2.6.12. *Sia V uno spazio vettoriale su $\mathbb{K}$. Supponiamo che $v_1, \dots, v_m, u \in V$ siano linearmente indipendenti (il caso $m = 0$ è ammesso) e che $v_1, \dots, v_m, w_1, \dots, w_n$ siano generatori di V. Allora $1 \leqslant n$ ed esiste $1 \leqslant i \leqslant n$ tale che V sia generato da*

$$v_1, \dots, v_m, w_1, \dots, w_{i-1}, u, w_{i+1}, \dots, w_n.$$

Dimostrazione. Siccome $V = \langle v_1, \dots, v_m, w_1, \dots, w_n \rangle$ esistono $x_1, \dots, x_m, y_1, \dots, y_n \in \mathbb{K}$ tali che

$$u = x_1 v_1 + \dots + x_m v_m + y_1 w_1 + \dots + y_n w_n. \tag{2.6.8}$$

Se fosse $0 = y_1 = y_2 = \dots = y_n$ (questa ipotesi include il caso $n = 0$) allora u sarebbe combinazione lineare di $v_1, \dots, v_m$ e quindi $v_1, \dots, v_m, u$ sarebbero linearmente dipendenti (vedi la Proposizione 2.5.9) contro l'ipotesi. Quindi $1 \leqslant n$ ed esiste $1 \leqslant i \leqslant n$ tale che $y_i \neq 0$. Moltiplicando per $-y_i^{-1}$ ambo i membri di (2.6.8) ed aggiungendo $(w_i + y_i^{-1} u)$ ai membri dell'equazione così ottenuta arriviamo all'equazione

$$w_i = -y_i^{-1} x_1 v_1 - \dots - y_i^{-1} x_m v_m - y_i^{-1} y_1 w_1 - \dots - y_i^{-1} y_{i-1} w_{i-1} + y_i^{-1} u - y_i^{-1} y_{i+1} w_{i+1} - y_i^{-1} y_n w_n.$$

Questo dimostra che $w_i \in \langle v_1, \dots, v_m, w_1, \dots, w_{i-1}, u, w_{i+1}, \dots, w_n \rangle$. Siccome V è generato da

$$v_1, \dots, v_m, w_1, \dots, w_n$$

segue che è anche generato da $v_1, \dots, v_m, w_1, \dots, w_{i-1}, u, w_{i+1}, \dots, w_n$. □

Proposizione 2.6.13. *Sia V uno spazio vettoriale su $\mathbb{K}$. Supponiamo che $r_1, \dots, r_p \in V$ siano linearmente indipendenti e che $z_1, \dots, z_q \in V$ siano generatori di V. Allora $p \leqslant q$ ed esistono $1 \leqslant j_1 < j_2 \leqslant \dots < j_p \leqslant q$ tali che V è generato da*

$$\{r_1, \dots, r_p\} \cup \{z_1, \dots, z_q\} \backslash \{z_{j_1}, z_{j_2}, \dots, z_{j_p}\}. \tag{2.6.9}$$

In altre parole: sostituendo nella lista $z_1, \dots, z_q$ ciascun z_{j_i} con r_i otteniamo un nuovo sistema di generatori.

Dimostrazione. Per induzione su p. Più precisamente sia A_p l'affermazione della proposizione: dimostriamo per induzione che è vera per ogni p. Il caso $p = 0$ è banalmente vero. (Se non vi piace il caso $p = 0$ cominciate l'induzione da $p = 1$: l'affermazione A_1 è vera per il caso $m = 0$ del 2.6.12.) Dimostriamo il passo induttivo, cioè assumiamo che A_p sia vera e dimostriamo che è vera A_{p+1}. Per l'ipotesi induttiva V è generato da (2.6.9). Applicando il Lemma 2.6.12 con $m = p$ e $u = r_{p+1}$, vediamo che vale A_{p+1}. □

Corollario 2.6.14. *Sia V uno spazio vettoriale su $\mathbb{K}$ finitamente generato, e che quindi ammette basi per il Corollario 2.6.5. Due basi di V hanno la stessa cardinalità.*

Dimostrazione. Siano $\{r_1, \dots, r_p\}$ e $\{z_1, \dots, z_q\}$ basi di V. Siccome i vettori $r_1, \dots, r_p$ sono linearmente indipendenti e i vettori $z_1, \dots, z_q$ sono generatori di V, si ha $p \leqslant q$ per la Proposizione 2.6.13. D'altra parte i vettori $z_1, \dots, z_q$ sono linearmente indipendenti e $r_1, \dots, r_p$ sono generatori di V, e quindi $q \leqslant p$ per la Proposizione 2.6.13. Concludiamo che $p = q$. □

Osservazione 2.6.15. Abbiamo dimostrato che uno spazio vettoriale finitamente generato ammette basi e che due basi hanno la stessa cardinalità. In verità le stesse affermazioni valgono per ogni spazio vettoriale, vedi [3].

Il Corollario 2.6.14 ci permette di dare la seguente definizione fondamentale.

Definizione 2.6.16. Sia V uno spazio vettoriale su $\mathbb{K}$ finitamente generato (e quindi V ammette basi per la 2.6.4). La *dimensione* di V è la cardinalità di una qualsiasi base di V.

Esempio 2.6.17. 1. Lo spazio vettoriale $\mathbb{K}^n$ ha dimensione n perchè $\mathbf{e}_1, \dots, \mathbf{e}_n$ è una base di $\mathbb{K}^n$.

2. La dimensione di $\mathbb{C}^n$ come spazio vettoriale *reale* è $2n$ perchè $\mathbf{e}_1, i\mathbf{e}_1, \mathbf{e}_2, i\mathbf{e}_2, \dots, \mathbf{e}_n, i\mathbf{e}_n$ è una base di $\mathbb{C}^n$ come spazio vettoriale reale.

3. $\mathbb{K}[x]_{\leqslant n}$ ha dimensione $(n+1)$ perchè una sua base è $\{1, x, x^2, \dots, x^n\}$.

4. Lo spazio vettoriale $\mathsf{V}(\mathbb{E}^2)$ dei vettori geometrici nel piano ha dimensione 2, perchè una sua base è data da una qualsiasi coppia di vettori *non* paralleli.

Proposizione 2.6.18. *Sia V uno spazio vettoriale su $\mathbb{K}$ finitamente generato di dimensione n.*

1. *Supponiamo che $v_1, \dots, v_m \in V$ siano linearmente indipendenti. Allora $m \leqslant n$ e se $m = n$ la lista $\{v_1, \dots, v_n\}$ è una base di V.*

2. *Supponiamo che $\langle v_1, \dots, v_m \rangle = V$. Allora $m \geqslant n$ e se $m = n$ la lista $\{v_1, \dots, v_n\}$ è una base di V.*

Dimostrazione. (1): Per la Proposizione 2.6.6 possiamo estendere $v_1, \dots, v_m$ a una base $\mathcal{B}$ di V. Siccome $\dim V = n$ la base $\mathcal{B}$ contiene n vettori e quindi $m \leqslant n$. Se $m = n$ allora $\mathcal{B} = \{v_1, \dots, v_n\}$ e quindi $W = \langle v_1, \dots, v_n \rangle = V$. (2): Per la Proposizione 2.6.4 possiamo eliminare alcuni dei v_i e ottenere una base $\mathcal{C}$ di V. Siccome $\dim V = n$ segue che $m \geqslant n$. Se $m = n$ abbiamo che $\mathcal{B} = \{v_1, \dots, v_n\}$ e quindi $W = V$. □

Esempio 2.6.19. Sia V uno uno spazio vettoriale con base $\{w_1, w_2\}$ (quindi $\dim V = 2$). Siano $v_1, v_2 \in V$ dati da

$$v_1 := aw_1 + bw_2, \qquad v_2 := cw_1 + dw_2. \tag{2.6.10}$$

Copiando gli argomenti dell'Esempio 2.5.5 si vede che v_1, v_2 sono linearmente dipendenti se e solo se $(ad - bc) = 0$ ovvero sono linearmente indipendenti se e solo se $(ad - bc) \neq 0$. Per la Proposizione 2.6.18 segue che $\{v_1, v_2\}$ è una base di V se e solo se $(ad - bc) \neq 0$.

Esempio 2.6.20. Consideriamo il sistema di equazioni lineari omogenee (2.3.5). La soluzione *banale* di (2.3.5) è quella con $x_j = 0$ per ogni $1 \leqslant j \leqslant n$. Notate che la soluzione banale esiste indipendentemente dal sistema scelto, ci interessa sapere se esiste o non esiste una soluzione non banale. Supponiamo che $n > m$ cioè che esistano più incognite che equazione: dimostriamo che esiste una soluzione non banale. Siano $v_1, \dots, v_n \in \mathbb{K}^m$ i vettori definiti da

$$v_j = (a_{1j}, a_{2j}, \dots, a_{ij}, \dots, a_{mj}).$$

Allora $(x_1, \dots, x_n)$ è soluzione di (2.3.5) se e solo se

$$x_1 v_1 + x_2 v_2 + \dots + x_n v = 0. \tag{2.6.11}$$

Siccome $\dim \mathbb{K}^m = m$ e per ipotesi $m < n$ la Proposizione 2.6.18 ci assicura che $v_1, \dots, v_n$ sono linearmente dipendenti. Quindi esistono $x_1, \dots, x_n \in \mathbb{K}$ non tutti nulli tali che valga (2.6.11) e cioè una soluzione non banale di (2.3.5).

Corollario 2.6.21. *Sia V uno spazio vettoriale su $\mathbb{K}$ finitamente generato. Se $W \subset V$ è un sottospazio allora $\dim W \leqslant \dim V$ e si ha eguaglianza se e solo se $W = V$.*

Dimostrazione. Sia $\{w_1, \dots, w_m\}$ una base di W: allora $w_1, \dots, w_m$ sono linearmente indipendenti e quindi il corollario segue dal punto (1) della Proposizione 2.6.18. □

Esempio 2.6.22. Siano $a_1, \dots, a_n \in \mathbb{K}$, e sia $W \subset \mathbb{K}^n$ il sottospazio delle soluzioni $X = (x_1, \dots, x_n)$ dell'equazione omogenea

$$a_1 x_1 + \dots + a_n x_n = 0. \tag{2.6.12}$$

Dimostriamo che

$$\dim W = \begin{cases} n & \text{se } 0 = a_1 = \ldots = a_n \\ n-1 & \text{altrimenti.} \end{cases} \tag{2.6.13}$$

Se $0 = a_1 = \ldots = a_n$, allora $W = \mathbb{K}^n$, e $\dim \mathbb{K}^n = n$ per l'Esempio 2.6.17. Ora supponiamo che esista $i \in \{1, \ldots, n\}$ tale che $a_i \neq 0$. L'equazione in (2.6.12) equivale a

$$x_i = -a_i^{-1} \cdot a_1 x_1 - a_i^{-1} \cdot a_2 x_2 - \ldots - a_i^{-1} \cdot a_n x_n. \tag{2.6.14}$$

In particolare vediamo che W non è tutto $\mathbb{K}^n$ e quindi, per il Corollario 2.6.21, la sua dimensione è strettamente più piccola della dimensione di $\mathbb{K}^n$ cioè $\dim W \leqslant (n-1)$. Inoltre dalla (2.6.14) segue che i vettori

$$\mathbf{e}_1 - a_i^{-1} \cdot a_1 \mathbf{e}_i,\ -a_i^{-1} \cdot a_2 \mathbf{e}_2,\ \ldots,\ -a_i^{-1} \cdot a_n \mathbf{e}_n \tag{2.6.15}$$

appartengono a W. Si vede facilmente che tali vettori sono linearmente indipendenti, e quindi $\dim W \geqslant (n-1)$. Siccome $\dim W \leqslant (n-1)$, segue che $\dim W = (n-1)$.

2.7 Formula di Grassmann

Consideriamo uno spazio vettoriale V su $\mathbb{K}$ e sottospazi $U, W \subset V$. Supponiamo che U e W siano finitamente generati, diciamo $U = \langle x_1, \ldots, x_p \rangle$ e $W = \langle y_1, \ldots, y_q \rangle$: allora $(U + W) = \langle x_1, \ldots, x_p, y_1, \ldots, y_q \rangle$ e quindi anche la somma $(U + W)$ è uno spazio finitamente generato. Per la 2.4.12 anche $U \cap W$ è finitamente generato. Quindi nell'ipotesi fatta le dimensioni di U, W, $(U + W)$ e $U \cap W$ sono definite. La formula di Grassmann dà una relazione tra queste dimensioni.

Proposizione 2.7.1 (Formula di Grassmann)**.** *Sia V uno spazio vettoriale su $\mathbb{K}$ e $U, W \subset V$ sottospazi finitamente generati. Allora*

$$\dim(U + W) + \dim(U \cap W) = \dim U + \dim W. \tag{2.7.1}$$

Dimostrazione. Sia $\{z_1, \ldots, z_a\}$ una base di $U \cap W$ ed estendiamola a una base $\{z_1, \ldots, z_a, u_1, \ldots, u_m\}$ di U e a una base $\{z_1, \ldots, z_a, w_1, \ldots, w_n\}$ di W. Dimostriamo che

$$\mathcal{B} := \{z_1, \ldots, z_a, u_1, \ldots, u_m, w_1, \ldots, w_n\}$$

è una base di $(U + W)$. Siccome $z_1, \ldots, z_a, u_1, \ldots, u_m$ generano U e $z_1, \ldots, z_a, w_1, \ldots, w_n$ generano W lo spazio $(U + W)$ è generato da $z_1, \ldots, z_a, u_1, \ldots, u_m, w_1, \ldots, w_n$. Rimane da dimostrare che $z_1, \ldots, z_a, u_1, \ldots, u_m, w_1, \ldots, w_n$ sono linearmente indipendenti. Supponiamo che

$$\lambda_1 z_1 + \ldots + \lambda_a z_a + \mu_1 u_1 + \ldots + \mu_m u_m + \theta_1 w_1 + \ldots + \theta_n w_n = 0. \tag{2.7.2}$$

Quindi abbiamo che

$$\lambda_1 z_1 + \ldots + \lambda_a z_a + \mu_1 u_1 + \ldots + \mu_m u_m = -(\theta_1 w_1 + \ldots + \theta_n w_n). \tag{2.7.3}$$

Il membro di sinistra di (2.7.3) è in U e il membro di destra (uguale a quello di sinistra) è in W, quindi sono entrambi in $U \cap W$. Siccome $\{z_1, \ldots, z_a\}$ è una base di $U \cap W$ esistono $\tau_1, \ldots, \tau_a \in \mathbb{K}$ tali che

$$\tau_1 z_1 + \ldots + \tau_a z_a = -(\theta_1 w_1 + \ldots + \theta_n w_n). \tag{2.7.4}$$

Ma $z_1, \ldots, z_a, w_1, \ldots, w_n$ sono linearmente indipendenti, quindi $0 = \tau_1 = \ldots = \tau_a = \theta_1 = \ldots = \theta_n$. Sostituendo $0 = \theta_1 = \ldots = \theta_n$ nella (2.7.2) otteniamo che

$$\lambda_1 z_1 + \ldots + \lambda_a z_a + \mu_1 u_1 + \ldots + \mu_m u_m = 0. \tag{2.7.5}$$

Siccome $z_1, \ldots, z_a, u_1, \ldots, u_m$ sono linearmente indipendenti segue che $0 = \lambda_1 = \ldots = \lambda_a = \mu_1 = \ldots = \mu_m$. Questo dimostra che $z_1, \ldots, z_a, u_1, \ldots, u_m, w_1, \ldots, w_n$ sono linearmente indipendenti e quindi che $\mathcal{B}$ è una base di $(U + W)$. Dunque abbiamo che

$$\dim(U + W) = a + m + n, \quad \dim U \cap W = a, \quad \dim U = a + m, \quad \dim W = a + n$$

e perciò vale (2.7.1). □

Diamo il seguente corollario della formula di Grassmann.

Corollario 2.7.2. *Sia V uno spazio vettoriale su $\mathbb{K}$ finitamente generato. Se $U, W \subset V$ sono sottospazi tali che $\dim U + \dim W > \dim V$ allora esiste un vettore non nullo in $U \cap W$.*

Dimostrazione. Per la formula di Grassman

$$\dim(U \cap W) = \dim U + \dim W - \dim(U + W).$$

Siccome $\dim(U + W) \leqslant \dim V$ segue che $\dim(U \cap W) > 0$, e quindi $U \cap W \neq \{0\}$. □

Definizione 2.7.3. Sia V uno spazio vettoriale finitamente generato e $W \subset V$ un sottospazio. La *codimensione* di W in V è

$$\operatorname{cod}(W, V) := \dim V - \dim W.$$

2.8 Costruzioni astratte di spazi vettoriali

Presenteremo due costruzioni che producono uno spazio vettoriale a partire da altri spazi vettoriali.

Somma diretta

Siano V e W spazi vettoriali sul campo $\mathbb{K}$. Definiamo la somma di elementi $(v_1, w_1), (v_2, w_2) \in V \times W$ così:

$$(v_1, w_1) + (v_2, w_2) := (v_1 + v_2, w_1 + w_2). \tag{2.8.1}$$

Dati $\lambda \in \mathbb{K}$ e $(v, w) \in V \times W$ definiamo

$$\lambda(v, w) := (\lambda v, \lambda w). \tag{2.8.2}$$

Si verifica facilmente che con queste operazioni $V \times W$ è uno spaziovettoriale su $\mathbb{K}$. L'elemento neutro è $(0_V, 0_W)$ dove 0_V e 0_W sono gli elementi neutri di V e W rispettivamente, e l'opposto di (v, w) è $(-v, -w)$. Lo spazio vettoriale $V \times W$ (con le operazioni appena definite) si denota $V \oplus W$ e si chiama la *somma diretta* di V e W.

Proposizione 2.8.1. *Siano V e W spazio vettoriali finitamente generati su $\mathbb{K}$. Allora*

$$\dim(V \oplus W) = \dim V + \dim W.$$

Dimostrazione. Siano $\{v_1, \dots, v_m\}$ e $\{w_1, \dots, w_n\}$ basi di V e W rispettivamente. Dimostreremo che

$$\{(v_1, 0_W), \dots, (v_m, 0_W), (0_V, w_1), \dots, (0_V, w_n)\} \tag{2.8.3}$$

è una base di $V \oplus W$, e la proposizione seguirà. I vettori di (2.8.3) generano $V \oplus W$ perchè dato $(v, w) \in V \oplus W$ esistono $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ e $\mu_1, \dots, \mu_n \in \mathbb{K}$ tali che $v = \sum_{i=1}^m \lambda_i v_i$ e $w = \sum_{j=1}^n \mu_j w_j$ (perchè $\{v_1, \dots, v_m\}$ genera V e $\{w_1, \dots, w_n\}$ genera W), e quindi

$$\sum_{i=1}^m \lambda_i (v_i, 0_W) + \sum_{j=1}^n \mu_j (0_V, w_j) = (\sum_{i=1}^m \lambda_i v_i, \sum_{j=1}^n \mu_j w_j) = (v, w). \tag{2.8.4}$$

Per dimostrare che i vettori di (2.8.3) sono linearmente indipendenti supponiamo che

$$\sum_{i=1}^m \lambda_i (v_i, 0_W) + \sum_{j=1}^n \mu_j (0_V, w_j) = (0_V, 0_W).$$

Guardando a (2.8.4) vediamo che necessariamente $0 = \lambda_1, \dots, \lambda_m = \mu_1, \dots, \mu_n$. □

Se $V_1, \ldots, V_m$ sono spazi vettoriali su $\mathbb{K}$ si definisce analogamente la somma di elementi di $V_1 \times \ldots \times V_m$ e la moltiplicazione di uno scalare per un elemento di $V_1 \times \ldots \times V_m$:

$$(v_1, \ldots, v_m) + (w_1, \ldots, w_m) := (v_1 + w_1, \ldots, v_m + w_m), \qquad \lambda(v_1, \ldots, v_m) := (\lambda v_1, \ldots, \lambda v_m). \quad (2.8.5)$$

Con queste operazioni $V_1 \times \ldots \times V_m$ è uno spaziovettoriale su $\mathbb{K}$ denotato $V_1 \oplus \ldots \oplus V_m$ e si chiama la *somma diretta* di $V_1, \ldots, V_m$. Supponiamo che $V_1, \ldots, V_m$ siano finitamente generati su $\mathbb{K}$; una dimostrazione analoga a quella della Proposizione 2.8.1 dà che

$$\dim(V_1 \oplus \ldots \oplus V_m) = \dim V_1 + \ldots + V_m.$$

Quoziente

Per la prossima costruzione assumiamo che V sia uno spazio vettoriale su $\mathbb{K}$ e che $W \subset V$ sia un sottospazio vettoriale. Definiamo la relazione $\overset{W}{\sim}$ su V così:

$$v_1 \overset{W}{\sim} v_2 \text{ se e solo se } (v_1 - v_2) \in W. \quad (2.8.6)$$

Si verifica facilmente che $\overset{W}{\sim}$ è una relazione di equivalenza: infatti $\overset{W}{\sim}$ è riflessiva perchè $0 \in W$, è simmetrica perchè se $w \in W$ allora l'opposto $-w \in W$ ed è transitiva perchè W è chiuso per la somma. Chiamiamo $\overset{W}{\sim}$ la *congruenza modulo* W.

Esempio 2.8.2. Se $W = V$ allora $V_1 \overset{W}{\sim} v_2$ per ogni $v_1, v_2 \in V$, cioè esiste un'unica classe di congruenza modulo W. Se $W = \{0\}$ allora $V_1 \overset{W}{\sim} v_2$ solo se $v_1 = v_2$, cioè possiamo identificare l'insieme delle classi di equivalenza con V stesso.

Esempio 2.8.3. Siano $V = \mathbb{K}^n$ e $W = \langle e_n \rangle = \{X \in \mathbb{K}^n \mid x_1 = x_2 = \ldots = x_{n-1} = 0\}$. Vettori $X, Y \in \mathbb{K}^n$ sono congruenti modulo W se solo se $x_i = y_i$ per $i \in \{1, \ldots, n-1\}$. Più in generale, sia $W \subset \mathbb{K}^n$ il sottospazio generato da $e_{j+1}, e_{j+2}, \ldots, e_n$. Allora $X, Y \in \mathbb{K}^n$ sono congruenti modulo W se solo se $x_i = y_i$ per $i \in \{1, \ldots, j\}$.

Proposizione 2.8.4. *Siano V uno spazio vettoriale su $\mathbb{K}$ e $W \subset V$ un sottospazio vettoriale. Supponiamo che $v \overset{W}{\sim} v'$ e $u \overset{W}{\sim} u'$. Allora*

$$(v + u) \overset{W}{\sim} (v' + u'), \qquad \lambda v \overset{W}{\sim} \lambda v'. \quad (2.8.7)$$

Dimostrazione. Per ipotesi $(v - v'), (u - u') \in W$; siccome W è un sottospazio è chiuso per somma,

$$(v + u) - (v' + u') = (v - v') + (u - u') \in W.$$

Questo dimostra la prima congruenza di (2.8.7). La seconda si dimostra in modo analogo. □

Per semplificare la notazione denotiamo $\overset{W}{\sim}$ con $\sim$. La Proposizione 2.8.4 permette di definire una operazione di somma su $V/\sim$. Se $[v], [u] \in V/\sim$ poniamo

$$[v] + [u] := [v + u]. \quad (2.8.8)$$

Notate che la definizione è ben posta (cioè la somma di $[v]$ e $[u]$ *non* dipende dai rappresentanti delle classi di equivalenza) grazie alla Proposizione 2.8.4. Analogamente definiamo una moltiplicazione per scalari. Se $\lambda \in \mathbb{K}$ e $[v] \in V/\sim$ poniamo

$$\lambda[v] := [\lambda v] \quad (2.8.9)$$

Di nuovo: la definizione è ben posta grazie alla Proposizione 2.8.4. Si verifica facilmente che con queste operazioni $V/\sim$ è uno spazio vettoriale su $\mathbb{K}$, con elemento neutro dato da $[0]$ (notate che $[0] = W$).

Definizione 2.8.5. Siano V uno spazio vettoriale su $\mathbb{K}$, e $W \subset V$ un sottospazio. Il *quoziente di V modulo W* è lo spazio vettoriale $V/\sim$ delle classi di congruenza modulo W con le operazioni di somma e moltiplicazione per scalari appena definiti. Lo denotiamo V/W.

Esempio 2.8.6. Se $W = V$ allora V/W è lo spazio vettoriale banale con unico elemento il vettore nullo. Se $W = \{0\}$ allora l'applicazione quoziente $V \longrightarrow V/W$ è una corrispondenza biunivoca. Se $W \subset \mathbb{K}^n$ è il sottospazio generato da $e_{j+1}, e_{j+2}, \ldots, e_n$, allora l'applicazione

$$\begin{array}{ccc} \mathbb{K}^j & \longrightarrow & \mathbb{K}^n/W \\ (t_1, \ldots, t_j) & \mapsto & (t_1, \ldots, t_j, 0, \ldots, 0) \end{array}$$

è biunivoca.

Proposizione 2.8.7. *Sia V uno spazio vettoriale finitamente generato su $\mathbb{K}$ e $W \subset V$ un sottospazio vettoriale. Allora V/W è finitamente generato e* $\dim(V/W) = \dim V - \dim W$.

Dimostrazione. Siano $v_1, \ldots, v_n$ generatori di V: allora $[v_1], \ldots, [v_n]$ sono generatori di V/W e quindi V/W è finitamente generato. Sia $\{w_1, \ldots, w_a\}$ una base di W: estendiamola a una base $\{w_1, \ldots, w_a, u_1, \ldots, u_b\}$ di V. Allora $(\dim V - \dim W) = b$ e quindi è sufficiente dimostrare che $\{[u_1], \ldots, [u_b]\}$ è una base di V/W. Prima dimostriamo che V/W è generato da $[u_1], \ldots, [u_b]$. Sia $[v] \in V/W$: siccome $w_1, \ldots, w_a, u_1, \ldots, u_b$ generano V esistono $\lambda_1, \ldots, \lambda_a, \mu_1, \ldots, \mu_b \in \mathbb{K}$ tali che

$$v = \lambda_1 w_1 + \ldots + \lambda_a w_a + \mu_1 u_1 + \ldots + \mu_b u_b \overset{W}{\sim} \mu_1 u_1 + \ldots + \mu_b u_b.$$

Quindi $[v] = \mu_1[u_1] + \ldots + \mu_b[u_b]$. Ora dimostriamo che $[u_1], \ldots, [u_b]$ sono linearmente indipendenti. Quindi supponiamo che esistano $\mu_1, \ldots, \mu_b \in \mathbb{K}$ tali che

$$\mu_1[u_1] + \ldots + \mu_b[u_b] = [0].$$

Siccome $[0] = W$ ciò significa che esistono $\lambda_1, \ldots, \lambda_a \in \mathbb{K}$ tali che

$$\mu_1 u_1 + \ldots + \mu_b u_b = \lambda_1 w_1 + \ldots + \lambda_a w_a.$$

Siccome $\{w_1, \ldots, w_a, u_1, \ldots, u_b\}$ è una base di V segue che

$$0 = \lambda_1 = \ldots = \lambda_a = \mu_1 = \ldots = \mu_b.$$

□

Esercizi del Capitolo 2

Esercizio 2.1. *Siano $X, Y, Z \in \mathbb{R}^3$ definiti da*

$$X := (1, 2, -3), \qquad Y := (3, -5, 2), \qquad Z := (1, 1, -2).$$

*Calcolate $2X - Y + Z$. Trovate $\lambda, \mu, \nu \in \mathbb{R}$ **non tutti nulli** tali che*

$$\lambda X + \mu Y + \nu Z = 0.$$

Esercizio 2.2. *Determinate quali dei seguenti sottoinsiemi $W_i \subset \mathbb{K}^3$ è un sottospazio.*

1. $\mathbb{K} = \mathbb{R}$ *e* $W_1 := \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$.
2. $\mathbb{K} = \mathbb{R}$ *e* $W_2 := \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z \leqslant 1\}$.
3. $\mathbb{K} = \mathbb{C}$ *e* $W_3 := \{(x, y, z) \in \mathbb{C}^3 \mid x^2 + y^2 + z^2 = 0\}$.
4. $\mathbb{K} = \mathbb{F}_2$ *e* $W_4 := \{(x, y, z) \in \mathbb{F}_2^3 \mid x^2 + y^2 + z^2 = 0\}$.

Esercizio 2.3. *Sia V uno spazio vettoriale e $u, v, w \in V$ tali che*

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

Dimostrate che $u = w$.

Esercizio 2.4. *Sia V un insieme. Supponiamo che sia definita un'operazione (somma)*

$$\begin{array}{ccc} V \times V & \longrightarrow & V \\ (v, w) & \mapsto & v + w \end{array}$$

e che

(a) esiste $0 \in V$ tale che $0 + v = v + 0 = v$ per ogni $v \in V$,

(b) $v + v = 0$ per ogni $v \in V$,

(c) $v + w = w + v$ per ogni $v, w \in V$, e

(d) $u + (v + w) = (u + v) + w$ per ogni $u, v, w \in V$.

Ora definite l'operazione (moltiplicazione per scalari in $\mathbb{F}_2$)

$$\begin{array}{ccc} \mathbb{F}_2 \times V & \longrightarrow & V \\ (\lambda, v) & \mapsto & \lambda v \end{array}$$

ponendo

$$\lambda v := \begin{cases} 0 & \text{se } \lambda = [0], \\ v & \text{se } \lambda = [1]. \end{cases}$$

Dimostrate che V con queste operazioni è uno spazio vettoriale su $\mathbb{F}_2$. (Ricordiamo che $\mathbb{F}_2$ è il campo delle classi di congruenza modulo 2.)

Esercizio 2.5. *Sia X un insieme, e sia $\mathcal{P}(X)$ l'insieme delle parti di X, cioè*

$$\mathcal{P}(X) = \{A \subset X\}.$$

Definiamo un'operazione (somma)

$$\begin{array}{ccc} \mathcal{P}(X) \times \mathcal{P}(X) & \longrightarrow & \mathcal{P}(X) \\ (A, B) & \mapsto & A + B \end{array}$$

ponendo

$$A + B = (A \backslash B) \cup (B \backslash A). \tag{2.8.10}$$

(Ricordiamo che $A\backslash B$ è l'insieme degli elementi di A che non sono in B.) Dimostrate che valgono (a), (b), (c) e (d) dell'esercizio 2.4, e quindi (per l'esercizio 2.4) $\mathcal{P}(X)$ ha una struttura di spazio vettoriale in cui la somma è data da (2.8.10).

Esercizio 2.6. *Nello spazio vettoriale $\mathbb{R}^2$ siano dati i vettori*

$$v_1 = (1, 2) \qquad v_2 = (4, 2) \qquad v_3 = (6, 3).$$

1. *Dire se i vettori v_1 e v_2 generano $\mathbb{R}^2$.*
2. *Dire se i vettori v_2 e v_3 generano $\mathbb{R}^2$.*

Esercizio 2.7. *Nello spazio vettoriale $\mathbb{R}^3$ siano dati i vettori*

$$v_1 = (1, 2, 1) \qquad v_2 = (1, 2, 0) \qquad v_3 = (1, 0, 1).$$

Verificare che v_1, v_2, v_3 generano $\mathbb{R}^3$.

Esercizio 2.8.

1. *Dire per quali sottospazi $W \subset \mathbb{R}^n$ il complementare $\mathbb{R}^n\backslash W$ è a sua volta un sottospazio.*
2. *Dire per quali sottospazi $W \subset \mathbb{R}^n$ l'insieme $(\mathbb{R}^n\backslash W) \cup \{0\}$ è a sua volta un sottospazio.*

Esercizio 2.9. *Sia V uno spazio vettoriale e W_1, W_2 sottospazi vettoriali di V. Dimostrare che se $W_1 \cup W_2$ è un sottospazio vettoriale di V allora $W_1 \subset W_2$ o $W_2 \subset W_1$.*

Esercizio 2.10. *Siano $v_1, v_2, v_3, v_4 \in \mathbb{R}^4$ definiti da*

$$v_1 = (1, 1, 0, -1), \quad v_2 = (1, -2, 3, 2), \quad v_3 = (1, -1, 0, 0), \quad v_4 = (0, 1, 0, 1).$$

Stabilite quali tra $\{v_1, v_2, v_3\}$, $\{v_2, v_3, v_4\}$, $\{v_3, v_4, v_1\}$ e $\{v_4, v_1, v_2\}$ sono terne di vettori linearmente dipendenti.

Esercizio 2.11. *Siano $v_1, v_2, u, w \in \mathbb{R}^2$ definiti da*

$$v_1 = (1, 1,), \quad v_2 = (1, 2), \quad u = (1, -1), \quad w = (0, 1).$$

1. *Verificate che $\mathcal{B} := \{v_1, v_2\}$ è una base di $\mathbb{R}^2$.*
2. *Calcolate le coordinate di u e w nella base $\mathcal{B}$.*

Esercizio 2.12. *Ricordiamo che $\mathbb{K}[x]_{\leqslant d}$ è il sottospazio vettoriale di $\mathbb{K}[x]$ i cui elementi sono i polinomi di grado al più d. Sia $\alpha \in \mathbb{K}$.*

1. *Dimostrate che*
$$\mathcal{B}_\alpha := \{\alpha, (x+\alpha), (x+\alpha)^2, \ldots, (x+\alpha)^d\}$$
è una base di $\mathbb{K}[x]_{\leqslant d}$.
2. *Determinate le coordinate di x^d nella base $\mathcal{B}_\alpha$. (Può essere utile ricordare la formula del binomio $(a+b)^d = \sum_{i=0}^d \binom{d}{i} a^{d-i}b^i$ dove $\binom{d}{i} := \frac{d!}{i!(d-i)!}$.)*

Esercizio 2.13. *Sia $W \subset \mathbb{K}[x]_{\leqslant d}$ definito da*

$$W := \{p \in \mathbb{K}[x]_{\leqslant d} \mid 0 = p(0) = p(-1) = p(1)\}.$$

Dimostrate che W è un sottospazio vettoriale di $\mathbb{K}[x]_d$ e calcolatene la dimensione. (Attenzione: il caso in cui $\operatorname{char}\mathbb{K} = 2$ è speciale.)

Esercizio 2.14. *Sia V uno spazio vettoriale su $\mathbb{K}$ finitamente generato e sia $\mathcal{B} := \{v_1, \ldots, v_n\}$ una base di V. Supponiamo che $v, w \in V$ e che $X_\mathcal{B}(v)$ e $X_\mathcal{B}(w)$ siano le n-ple di coordinate di V e W rispettivamente.*

1. *A cosa è uguale $X_\mathcal{B}(v + w)$?*
2. *A cosa è uguale $X_\mathcal{B}(\lambda v)$?*

Esercizio 2.15. *Sia V uno spazio vettoriale e sia $\mathcal{B} := \{v_1, \ldots, v_n\}$ una base di V. Sia*

$$u \in \langle v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \rangle.$$

(Notate che v_i "manca".) Dimostrate che $\mathcal{C} := \{v_1, \ldots, v_{i-1}, v_i + u, v_{i+1}, \ldots, v_n\}$ è una base di V.

Esercizio 2.16. *Siano $v_1, v_2, v_3 \in \mathbb{R}^3$ definiti da*

$$v_1 = (a_1, b_1, c_1), \quad v_2 = (a_2, b_2, c_2), \quad v_3 = (0, 0, 1).$$

Stabilite sotto quali condizioni $\{v_1, v_2, v_3\}$ è una base di $\mathbb{R}^3$.

Esercizio 2.17. *Sia V uno spazio vettoriale finitamente generato e siano $W_1, \ldots, W_p \subset V$ sottospazi. Si dimostri che*

$$\operatorname{cod}(W_1 \cap \ldots \cap W_p, V) \leqslant \operatorname{cod}(W_1, V) + \ldots + \operatorname{cod}(W_p, V).$$

(Suggerimento: si proceda per induzione su p e si applichi la Formula di Grassmann.)

Esercizio 2.18. *Sia $W \subset \mathbb{K}^n$ lo spazio delle soluzioni del sistema lineare omogeneo (2.3.5). Si dimostri che $\dim W \geqslant (n - m)$. (Invocate l'Esercizio 2.17 e l'Esempio 2.6.22.)*

Esercizio 2.19. *Sia $\mathbb{K}$ un campo e $\alpha_0, \alpha_1, \ldots, \alpha_n \in \mathbb{K}$ distinti. Siano $v_0, v_1, \ldots, v_n \in \mathbb{K}^{n+1}$ definiti da*

$$v_i = (\alpha_0^i, \alpha_1^i, \ldots, \alpha_n^i), \qquad 0 \leqslant i \leqslant n.$$

Dimostrate che $\{v_0, v_1, \ldots, v_n\}$ è una base di $\mathbb{K}^{n+1}$.

Esercizio 2.20. *Sia $d \in \mathbb{Q}$ e poniamo*

$$\mathbb{Q}[\sqrt{d}] := \{\alpha + \beta\sqrt{d} \mid \alpha, \beta \in \mathbb{Q}\}.$$

1. *Verificate che $\mathbb{Q}[\sqrt{d}]$ è un sottocampo di $\mathbb{C}$.*
2. *La somma e la moltiplicazione per $\mathbb{Q}$ danno a $\mathbb{Q}[\sqrt{d}]$ una struttura di spazio vettoriale su $\mathbb{Q}$: calcolatene la dimensione. (La risposta dipende dal numero d.)*

Capitolo 3

Applicazioni lineari

3.1 Definizione e prime proprietà

La definizione

Definizione 3.1.1. Siano V, W spazi vettoriali sullo stesso campo $\mathbb{K}$. Un'applicazione $f\colon V \to W$ è *lineare* se

(1) $f(v+w) = f(v) + f(w)$ per $v, w \in V$, e

(2) $f(\lambda v) = \lambda f(v)$ per $\lambda \in \mathbb{K}$ e $v \in V$.

Esempio 3.1.2. Sia

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{f} & \mathbb{K} \\ (x_1, \dots, x_n) & \mapsto & a_1x_1 + a_2x_2 + \dots a_nx_n. \end{array} \tag{3.1.1}$$

La f è lineare. Infatti se $X, Y \in \mathbb{K}^n$

$$f(X+Y) = \sum_{i=1}^n a_i(x_i + y_i) = \sum_{i=1}^n (a_ix_i + a_iy_i) = \sum_{i=1}^n a_ix_i + \sum_{i=1}^n a_iy_i = \lambda f(X) + \mu f(Y).$$

Questo mostra che vale (1) della Definizione 3.1.1. Per verificare che vale (2) della Definizione 3.1.1, siano $X \in \mathbb{K}^n$ e $\lambda \in \mathbb{K}$: si ha che

$$f(\lambda X) = \sum_{i=1}^n a_i(\lambda x_i) = \lambda \sum_{i=1}^n a_ix_i = \lambda f(X),$$

e quindi vale (2).

Esempio 3.1.3. Sia V uno spazio vettoriale finitamente generato, e sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una sua base. L'applicazione

$$V \xrightarrow{X_{\mathcal{B}}} \mathbb{K}^n$$

(vedi (2.6.7)) che associa a $v \in V$ l'n-pla delle sue coordinate nella base $\mathcal{B}$ è lineare. (Questo risponde all'Esercizio 2.14.) Per dimostrare che vale (1) della Definizione 3.1.1, cioè che

$$X_{\mathcal{B}}(v+w) = X_{\mathcal{B}}(v) + X_{\mathcal{B}}(w), \tag{3.1.2}$$

poniamo $X := X_{\mathcal{B}}(v)$ e $Y := X_{\mathcal{B}}(w)$. Quindi $v = \sum_{i=1}^n x_iv_i$ e $w = \sum_{i=1}^n y_iv_i$, e perciò

$$v + w = \sum_{i=1}^n x_iv_i + \sum_{i=1}^n y_iv_i = \sum_{i=1}^n (x_i + y_i)v_i,$$

cioè vale (3.1.2). In modo analogo si verifica che vale (2) della Definizione 3.1.1.

Osservazione 3.1.4. Siano V, W spazi vettoriali su un campo $\mathbb{K}$. Un'applicazione $f\colon V \to W$ è lineare se e solo se per $v, w \in V$ e $\lambda, \mu \in \mathbb{K}$ vale

$$f(\lambda v + \mu w) = \lambda f(v) + \mu f(w). \tag{3.1.3}$$

Infatti supponiamo che f sia lineare. Allora $f(\lambda v+\mu w) = f(\lambda v)+f(\mu w)$ per (1) della Definizione 3.1.1, e per (2) della stessa definizione abbiamo $f(\lambda v) = \lambda f(v)$ e $f(\mu w) = \mu f(w)$, quindi vale (3.1.3). Viceversa, supponiamo che valga (3.1.3) per ogni scelta di $v, w \in V$ e $\lambda, \mu \in \mathbb{K}$. Ponendo $\lambda = \mu = 1$ nella (3.1.3) otteniamo che vale (1) della Definizione 3.1.1, e ponendo $\mu = 0$ otteniamo che vale (2).

Lemma 3.1.5. *Supponiamo che $f\colon V \to W$ sia lineare. Allora*

(a) $f(0) = 0$,

(b) se $v \in V$ allora $f(-v) = -f(v)$,

(c) se $v_1, v_2, \ldots, v_n \in V$ e $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ allora

$$f(\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n) = \lambda_1 f(v_1) + \lambda_2 f(v_2) + \ldots + \lambda_n f(v_n). \tag{3.1.4}$$

Dimostrazione. (a): Basta porre $\lambda = 0$ nella (2) della Definizione 3.1.1. (b): Basta porre $\lambda = -1$ e $\mu = 0$ nella (2) della Definizione 3.1.1. (c): Supponiamo che valga (3.1.4) per ogni n e dimostriamo che f è lineare. Allora (3.1.4) vale per $n = 2$, e quindi f è lineare per l'Osservazione3.1.4. Ora supponiamo che f sia lineare e dimostriamo che vale (3.1.4) per ogni n. Si può ragionare per induzione su n. Se $n = 1$, allora (3.1.4) vale per (2) della Definizione 3.1.1. Finiamo dimostrando il passo induttivo. Supponiamo che valga (3.1.4). Allora

$$\begin{aligned} f(\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n + \lambda_{n+1} v_{n+1}) &= f(\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n) + f(\lambda_{n+1} v_{n+1}) = \\ &= \lambda_1 f(v_1) + \lambda_2 f(v_2) + \ldots + \lambda_n f(v_n) + \lambda_{n+1} f(v_{n+1}), \end{aligned}$$

cioè vale (3.1.4) con n sostituito da $n + 1$. (La prima uguaglianza vale per (1) della Definizione 3.1.1, la seconda vale per l'ipotesi induttiva e per (2) della Definizione 3.1.1.) □

Esempio 3.1.6. Mostriamo che l'Esempio 3.1.2 dà tutte le applicazione lineari da $\mathbb{K}^n$ a $\mathbb{K}$. Più precisamente, sia $f\colon \mathbb{K}^n \to \mathbb{K}$ lineare e poniamo $a_i := f(\mathbf{e}_i)$; dimostriamo che

$$f(x_1, \ldots, x_n) = a_1 x_1 + a_2 x_2 + \ldots a_n x_n. \tag{3.1.5}$$

Infatti per (c) del Lemma 3.1.5 abbiamo che

$$\begin{aligned} f(x_1, \ldots, x_n) = f(x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \ldots + x_n \mathbf{e}_n) = \\ = x_1 f(\mathbf{e}_1) + x_2 f(\mathbf{e}_2) + \ldots + x_n f(\mathbf{e}_n) = a_1 x_1 + a_2 x_2 + \ldots a_n x_n. \end{aligned}$$

Osservazione 3.1.7. L'Esempio 3.1.6 mostra che le applicazioni lineari sono applicazioni molto particolari. Inoltre, tenendo conto che in inglese "line" significa retta, spiega l'uso dell'aggettivo "lineare". Infatti vediamo che un'applicazione $f\colon \mathbb{R} \to \mathbb{R}$ è lineare se e solo se il suo grafico in $\mathbb{R}^2$ (che identifichiamo con il piano $\mathbb{E}^2$ della geometria euclidea dopo aver scelto un sistema di coordinate cartesiane) è una retta, ed analogamente un'applicazione $f\colon \mathbb{R}^2 \to \mathbb{R}$ è lineare se e solo se il suo grafico in $\mathbb{R}^3$ (che identifichiamo con lo spazio $\mathbb{E}^3$ della geometria euclidea dopo aver scelto un sistema di coordinate cartesiane) è un piano.

Esempio 3.1.8. Sia $V = \mathsf{V}(\mathbb{E}^2)$ lo spazio vettoriale reale dei vettori del piano $\mathbb{E}^2$. Sia $\rho\colon \mathbb{E}^2 \to \mathbb{E}^2$ la rotazione intorno a un punto di un angolo fissato. Se P_1Q_1, P_2Q_2 sono segmenti orientati equipollenti allora i segmenti orientati $\rho(P_1)\rho(Q_1)$ e $\rho(P_2)\rho(Q_2)$ sono equipollenti perchè ρ manda rette parallele in rette parallele . Quindi ponendo $\mathsf{V}(\overrightarrow{PQ}) := \overrightarrow{\rho(P)\rho(Q)}$ abbiamo definito un'applicazione

$$\mathsf{V}(\rho)\colon \mathsf{V}(\mathbb{E}^2) \to \mathsf{V}(\mathbb{E}^2).$$

La $\mathsf{V}(\rho)$ è lineare.

Esempio 3.1.9. Sia $c \in \mathbb{K}$. L'applicazione

$$\begin{array}{ccc} \mathbb{K}[x] & \xrightarrow{g} & \mathbb{K} \\ p & \mapsto & p(c) \end{array} \tag{3.1.6}$$

è lineare. Analogamente, sia X un insieme, e $x_0 \in X$. L'applicazione

$$\begin{array}{ccc} \mathbb{K}^X & \xrightarrow{h} & \mathbb{K} \\ \varphi & \mapsto & \varphi(x_0) \end{array} \tag{3.1.7}$$

è lineare.

Esempio 3.1.10. Siano V uno spazio vettoriale su $\mathbb{K}$, e $U \subset V$ un sottospazio. L'applicazione quoziente

$$\begin{array}{ccc} V & \xrightarrow{\pi} & V/U \\ [v] & \mapsto & [v] \end{array} \tag{3.1.8}$$

è lineare.

Primi risultati

Diamo una risposta alla domanda (un pò vaga): da cosa è determinata un'applicazione lineare? Iniziamo con un risultato di unicità.

Proposizione 3.1.11. *Siano V, W spazi vettoriali su $\mathbb{K}$. Sia $\{v_i\}_{i\in I}$ una lista di generatori di V. Supponiamo che $f, g\colon V \to W$ siano applicazioni lineari e che $f(v_i) = g(v_i)$ per tutti gli $i \in I$. Allora $f = g$.*

Dimostrazione. Dobbiamo dimostrare che per ogni $v \in V$ si ha $f(v) = g(v)$. Siccome i vettori di $\{v_i\}_{i\in I}$ generano V, esistono $i_1, \ldots, i_n \in I$ tali che

$$v = x_1 v_{i_1} + \ldots + x_n v_{i_n}.$$

Siccome f e g sono lineari,

$$\begin{aligned} f(v) = f(x_1 v_{i_1} + \ldots + x_n v_{i_n}) = x_1 f(v_{i_1}) + \ldots + x_n f(v_{i_n})) = \\ = x_1 g(v_{i_1}) + \ldots + x_n g(v_{i_n})) = g(x_1 v_{i_1} + \ldots + x_n v_{i_n}) = g(v). \end{aligned}$$

□

Il risultato seguente è in qualche senso il duale del precedente. Vale per spazi vettoriali arbitrari ma noi lo formuliamo e dimostriamo solo per spazi vettoriali finitamente generati.

Proposizione 3.1.12. *Siano V, W spazi vettoriali su $\mathbb{K}$, e supponiamo che V sia finitamente generato. Sia $\{v_1, \ldots, v_m\}$ una lista di vettori linearmente indipendenti di V, e sia $\{w_1, \ldots, w_m\}$ una lista di vettori di W. Allora esiste un'applicazione lineare $f\colon V \to W$ tale che $f(v_i) = w_i$ per $i \in \{1, \ldots, m\}$.*

Dimostrazione. Iniziamo dimostrando il risultato sotto l'ipotesi che $\{v_1, \ldots, v_m\}$ sia una base di V. Dato $v \in V$ siano $x_1, \ldots, x_m$ le coordinate di v nella base $\{v_1, \ldots, v_m\}$, cioè

$$v = x_1 v_1 + \ldots + x_m v_m.$$

Definiamo $f\colon V \to W$ ponendo

$$f(v) := x_1 w_1 + \ldots + x_m w_m.$$

Dimostriamo che f è lineare e che $f(v_i) = w_i$ per $i \in \{1, \dots, n\}$. Siano $v, w \in V$ e $\lambda, \mu \in \mathbb{K}$. Siano $x_1, \dots, x_m$ e $y_1, \dots, y_m$ rispettivamente le coordinate di v e w nella base $\{v_1, \dots, v_m\}$. Allora

$$f(\lambda v + \mu w) = f((\lambda x_1 + \mu y_1)v_1 + \dots + (\lambda x_m + \mu y_m)v_m) = (\lambda x_1 + \mu y_1)w_1 + \dots + (\lambda x_m + \mu y_m)w_m = \\ \lambda x_1 w_1 + \dots + \lambda x_m w_m + \mu y_1 w_1 + \dots + \mu y_m w_m = \lambda f(v) + \mu g(w).$$

Per l'Osservazione 3.1.4 questo dimostra che f è lineare. L'uguaglianza $f(v_i) = w_i$ vale perchè le coordinate di v_i sono $0, \dots, 0, 1, 0, \dots, 0$ con la coordinata 1 al posto i. Abbiamo dimostrato che il risultato vale con l'ipotesi aggiuntiva che $\{v_1, \dots, v_m\}$ sia una base di V. Per dimostrare il risultato in generale estendiamo $\{v_1, \dots, v_m\}$ a una base $\{v_1, \dots, v_m, v_{m+1}, \dots, v_n\}$ di V (sappiamo che è possibile perchè V è finitamente generato) e scegliamo arbitrari elementi $w_{m+1}, \dots, w_n \in V$. Per il risultato appena dimostrato esiste un'applicazione lineare $f\colon V \to W$ tale che $f(v_i)$ per $1 \in \{1, \dots, n\}$. Siccome $m \leqslant n$ abbiamo fatto. □

Osservazione 3.1.13. Nella Proposizione 3.1.12 l'ipotesi che $v_1, \dots, v_m \in V$ siano linearmente indipendenti è essenziale. Se non sono linearmente indipendenti, in generale non esiste un'applicazione lineare $f\colon V \to W$ con valori si v_i assegnati a piacere. Per esempio $f(0)$ non può essere un vettore non nullo per (a) del Lemma 3.1.5. Più in generale, se $f(v_i) = w_i$ per $i \in \{1, \dots, m\}$ e $\lambda_1 v_1 + \dots + \lambda_m v_m = 0$ è una relazione lineare non banale, per linearità di f vale la relazione lineare non banale $\lambda_1 w_1 + \dots + \lambda_m w_m = 0$. Quindi, se tale relazione non vale, la f non esiste. Analogamente, se $v_1, \dots, v_m \in V$ non generano V, allora esistono applicazioni lineari $f, g\colon V \to W$ diverse tra loro tali che $f(v_i) = g(v_i)$ per $i \in \{1, \dots, m\}$ (per esempio se $V \neq \{0\}$, allora esiste più di un'applicazione lineare $f\colon V \to W$, ma tutte le applicazioni lineari hanno valore 0 su 0.).

Il prossimo risultato segue immediatamente dalle Proposizioni 3.1.11 e 3.1.12.

Corollario 3.1.14. *Siano V, W spazi vettoriali su $\mathbb{K}$, e supponiamo che V sia finitamente generato. Sia $\{v_1, \dots, v_n\}$ una base di V, e siano $w_1, \dots, w_n \in W$. Allora esiste un'applicazione lineare $f\colon V \to W$ tale che $f(v_i) = w_i$ per $i \in \{1, \dots, n\}$, e tale applicazione lineare è unica.*

Immagine e nucleo di un'applicazione lineare

Proposizione 3.1.15. *Siano V, W spazi vettoriali su un campo $\mathbb{K}$ e $f\colon V \to W$ un'applicazione lineare. Allora $f^{-1}(0)$ è un sottospazio vettoriale di V e $\operatorname{im} f$ è un sottospazio vettoriale di W.*

Dimostrazione. Dimostriamo che $f^{-1}(0)$ è un sottospazio vettoriale di V. Siccome $f(0) = 0$ abbiamo che $f^{-1}(0)$ non è vuoto. Siano $v_1, v_2 \in f^{-1}(0)$ e $\lambda_1, \lambda_2 \in \mathbb{K}$. Per linearità di f abbiamo che

$$f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2) = \lambda_1 0 + \lambda_2 0 = 0.$$

Quindi $(\lambda_1 v_1 + \lambda_2 v_2)) \in f^{-1}(0)$: questo dimostra che $f^{-1}(0)$ è un sottospazio vettoriale di V. Ora dimostriamo che $\operatorname{im} f$ è un sottospazio vettoriale di W. Siccome V non è vuoto $\operatorname{im} f$ non è vuoto. Siano $w_1, w_2 \in \operatorname{im} f$ e $\lambda_1, \lambda_2 \in \mathbb{K}$. Quindi esistono $v_1, v_2 \in V$ tali che $f(v_i) = w_i$ e per linearità di f abbiamo che

$$\lambda_1 w_1 + \lambda_2 w_2 = \lambda_1 f(v_1) + \lambda_2 f(v_2) = f(\lambda_1 v_1 + \lambda_2 v_2) \in \operatorname{im} f.$$

□

Definizione 3.1.16. Siano V, W spazi vettoriali su un campo $\mathbb{K}$ e $f\colon V \to W$ un'applicazione lineare. Il *nucleo di f* è il sottospazio $f^{-1}(0)$, e viene denotato $\ker f$.

Esempio 3.1.17. Per $i \in \{1, \dots, m\}$ e $j \in \{1, \dots, n\}$ sia $a_{i,j} \in \mathbb{K}$. Poniamo $f_i(x_1, \dots, x_n) := a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n$. L'applicazione

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{f} & \mathbb{K}^m \\ (x_1, \dots, x_n) & \mapsto & (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)) \end{array} \tag{3.1.9}$$

è lineare. Infatti siano $X, Y \in \mathbb{K}^n$, e $\lambda, \mu \in \mathbb{K}$. Allora, siccome f_i è lineare per ogni $i \in \{1, \dots, m\}$ (vedi l'Esempio 3.1.2)

$$f(\lambda X + \mu Y) = (f_1(\lambda X + \mu Y), \dots, f_m(\lambda X + \mu Y)) = \\ = (\lambda f_1(X) + \mu f_1(Y), \dots, \lambda f_m(X) + \mu f_m(Y)) = \lambda f(X) + \mu f(Y).$$

Il nucleo di f è il sottospazio delle soluzioni del sistema di equazioni lineari omogenee (2.3.5).

Esempio 3.1.18. Sia $c \in \mathbb{K}$, e sia $g\colon \mathbb{K}[x] \to \mathbb{K}$ l'applicazione lineare definita da $g(p) = p(c)$, vedi l'Esempio 3.1.9. Il nucleo di g è il sottospazio $\{(x-c)q \mid q \in \mathbb{K}[x]\}$.

Esempio 3.1.19. Siano V uno spazio vettoriale su $\mathbb{K}$, e $U \subset V$ un sottospazio. L'applicazione quoziente $\pi\colon V \to V/U$ è lineare, vedi l'Esempio 3.1.10. Il nucleo di π è U.

Proposizione 3.1.20. *Siano V, W spazi vettoriali su un campo $\mathbb{K}$ e $f\colon V \to W$ un'appplicazione lineare. Allora f è iniettiva se e solo se* $\ker f = \{0\}$.

Dimostrazione. Supponiamo che f sia iniettiva. Siccome $f(0) = 0$ segue che $\ker f = \{0\}$. Ora supponiamo che $\ker f = \{0\}$ e dimostriamo che f è iniettiva. Supponiamo che $f(v) = f(w)$. per linearità segue che $f(v - w) = 0$ cioè $(v - w) \in \ker f$. Siccome $\ker f = \{0\}$ segue che $(v - w) = 0$ cioè $v - w$. Abbiamo dimostrato che f è iniettiva. □

Proposizione 3.1.21. *Siano V, W spazi vettoriali su un campo $\mathbb{K}$, con V finitamente generato. Sia $f\colon V \to W$ un'applicazione lineare. Allora*

$$\dim V = \dim(\ker f) + \dim(\operatorname{im} f). \tag{3.1.10}$$

(L'ipotesi che V sia finitamente generato dà che $\ker f$ *e* $\operatorname{im} F$ *sono finitamente generati e quindi le loro dimensioni sono ben definite.)*

Dimostrazione. Sia $\{v_1, \dots, v_a\}$ una base di $\ker f$ e $\{w_1, \dots, w_b\}$ una base di $\operatorname{im} f$. Siano $u_i \in V$ tali che $f(u_i) = w_i$ per $1 \leqslant i \leqslant b$. Dimostriamo che $\{v_1, \dots, v_a, u_1, \dots, u_b\}$ è una base di V. Dimostriamo che V è generato da $v_1, \dots, v_a, u_1, \dots, u_b$. Sia $v \in V$. Siccome $f(v) \in \operatorname{im} f$ e $\{w_1, \dots, w_b\}$ è una base di $\operatorname{im} f$ abbiamo che esistono $\alpha_1, \dots, \alpha_b \in \mathbb{K}$ tali che

$$f(v) = \alpha_1 w_1 + \ldots + \alpha_b w_b.$$

Per linearità di f segue che

$$f(v - \alpha_1 u_1 - \ldots + \alpha_b u_b) = f(v) - \alpha_1 w_1 - \ldots - \alpha_b w_b = 0.$$

Quindi $(v - \alpha_1 u_1 - \ldots + \alpha_b u_b) \in \ker f$: siccome $\{v_1, \dots, v_a\}$ è una base di $\ker f$ esistono $\beta_1, \dots, \beta_a \in \mathbb{K}$ tali che

$$v - \alpha_1 u_1 - \ldots - \alpha_b u_b = \beta_1 v_1 + \ldots + \beta_a v_a.$$

Segue che $v = \alpha_1 u_1 + \ldots + \alpha_b u_b + \beta_1 v_1 + \ldots + \beta_a v_a$. Ora dimostriamo che $\{v_1, \dots, v_a, u_1, \dots, u_b\}$ sono linearmente indipendenti. Supponiamo che

$$\lambda_1 v_1 + \ldots + \lambda_a v_a + \mu_1 u_1 + \ldots + \mu_b u_b = 0. \tag{3.1.11}$$

Applicando f a entrambi i membri e sfruttando la linearità di f otteniamo che

$$\mu_1 f(u_1) + \ldots + \mu_b f(u_b) = f(0) = 0.$$

Siccome $f(u_i) = w_i$ e $w_1, \dots, w_b$ sono linearmente indipendenti (costituiscono una base di $\operatorname{im} f$) segue che $0 = \mu_1 = \ldots = \mu_b$. Dalla (3.1.11) otteniamo che $0 = \lambda_1 = \ldots = \lambda_a$ ($v_1, \dots, v_a$ sono linearmente indipendenti perchè per ipotesi formano una base di $\ker f$). Questo dimostra che $\{v_1, \dots, v_a, u_1, \dots, u_b\}$ è una base di V: quindi $\dim V = a + b$ ovvero vale (3.1.10). □

Il seguente risultato segue subito dalla Proposizione 3.1.21.

Corollario 3.1.22. *Siano V, W spazi vettoriali su un campo $\mathbb{K}$. Supponiamo che V e W siano finitamente generati. Sia $f\colon V \to W$ un'applicazione lineare. Allora*

$$\dim(\ker f) \geqslant \dim V - \dim W.$$

Esempio 3.1.23. Sia $f\colon \mathbb{K}^n \to \mathbb{K}^m$ l'applicazione lineare data da (3.1.9). Applicando il Corollario 3.1.22 a f otteniamo che $\dim \ker f \geqslant (n-m)$, vedi l'Esempio 2.6.20 e l'Esercizio 2.18.

Operazioni tra applicazioni lineari

Vedremo come, a partire da applicazioni lineari date, si producano altre applicazioni lineari. Iniziamo dalla composizione di applicazioni.

Proposizione 3.1.24. *Siano U, V, W spazi vettoriali su un campo $\mathbb{K}$. Se $g\colon U \to V$ e $f\colon V \to W$ sono applicazioni lineari, allora $f \circ g$ è un'applicazione lineare.*

Dimostrazione. Abbiamo che

$$\begin{aligned} f \circ g(\lambda_1 v_1 + \lambda_2 v_2) &= f(g(\lambda_1 v_1 + \lambda_2 v_2)) = \\ &= f(\lambda_1 g(v_1) + \lambda_2 g(v_2)) = \lambda_1 f \circ g(v_1) + \lambda_2 f \circ g(v_2). \end{aligned} \tag{3.1.12}$$

Questo dimostra che $f \circ g$ è lineare. □

Ora diamo una struttura di spazio vettoriale all'insieme delle applicazioni lineari tra spazi vettoriali.

Terminologia 3.1.25. Siano V, W spazi vettoriali su un campo $\mathbb{K}$. L'insieme delle funzioni lineari $f\colon V \to W$ è denotato $\mathcal{L}(V, W)$.

Proposizione 3.1.26. *Siano V, W spazi vettoriali su un campo $\mathbb{K}$. Siano $f, g \in \mathcal{L}(V, W)$ applicazioni lineari e $\lambda \in \mathbb{K}$. Siano $(f+g)\colon V \to W$ e $\lambda f\colon V \to W$ date da*

$$(f+g)(v) := f(v) + g(v), \qquad (\lambda f)(v) := \lambda f(v). \tag{3.1.13}$$

Allora sia $(f+g)$ che λf sono applicazioni lineari.

Dimostrazione. Abbiamo che

$$\begin{aligned} (f+g)(\lambda_1 v_1 + \lambda_2 v_2) &= f(\lambda_1 v_1 + \lambda_2 v_2) + g(\lambda_1 v_1 + \lambda_2 v_2) = \\ &= \lambda_1 f(v_1) + \lambda_2 f(v_2) + \lambda_1 g(v_1) + \lambda_2 g(v_2) = \lambda_1 (f+g)(v_1) + \lambda_2 (f+g)(v_2). \end{aligned} \tag{3.1.14}$$

Questo dimostra che $(f+g)$ è lineare. Un conto simile dà che λf è lineare. □

Per la Proposizione appena dimostrata abbiamo operazioni

$$\begin{array}{ccccc} \mathcal{L}(V,W) \times \mathcal{L}(V,W) & \longrightarrow & \mathcal{L}(V,W) & \qquad \mathbb{K} \times \mathcal{L}(V,W) & \longrightarrow & \mathcal{L}(V,W) \\ (f,g) & \mapsto & f+g, & (\lambda, f) & \mapsto & \lambda f. \end{array} \tag{3.1.15}$$

La dimostrazione del seguente risultato consiste di semplici verifiche, che lasciamo al lettore.

Proposizione 3.1.27. *Siano V, W spazi vettoriali su un campo $\mathbb{K}$. Allora $\mathcal{L}(V, W)$, provvisto della somma e del prodotto per scalari in* (3.1.15) *è uno spazio vettoriale su $\mathbb{K}$.*

Notiamo che l'elemento neutro di $\mathcal{L}(V, W)$ è l'applicazione *nulla* 0, definita da $0(v) = 0$ per ogni $v \in V$.

Definizione 3.1.28. Sia V uno spazio vettoriale su un campo $\mathbb{K}$. Il *duale* di V è lo spazio vettoriale delle funzioni lineari $f\colon V \to \mathbb{K}$ (cioè $\mathcal{L}(V,\mathbb{K})$), ed è denotato $V^\vee$.

Esempio 3.1.29. Per l'Esempio 3.1.6 ogni elemento del duale di $\mathbb{K}^n$ è dato da

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{f} & \mathbb{K} \\ (x_1,\dots,x_n) & \mapsto & a_1x_1 + a_2x_2 + \dots a_nx_n. \end{array} \tag{3.1.16}$$

Inoltre la somma e prodotto scalare sul duale di $\mathbb{K}^n$ ricordano molto quelli in $\mathbb{K}^n$. Infatti, se

$$f(x_1,\dots,x_n) = a_1x_1 + a_2x_2 + \dots a_nx_n, \qquad g(x_1,\dots,x_n) = b_1x_1 + b_2x_2 + \dots b_nx_n$$

e $\lambda \in \mathbb{K}$, allora

$$(f+g)(x_1,\dots,x_n) = (a_1+b_1)x_1 + (a_2+b_2)x_2 + \dots (a_n+b_n)x_n, \tag{3.1.17}$$

$$(\lambda f)(x_1,\dots,x_n) = (\lambda a_1)x_1 + (\lambda a_2)x_2 + \dots (\lambda a_n)x_n. \tag{3.1.18}$$

3.2 Isomorfismi

Definizione 3.2.1. Siano V, W spazi vettoriali su uno stesso campo $\mathbb{K}$. Un *isomorfismo* tra V e W è un'applicazione **lineare** $f\colon V \to W$ tale che esista una $g\colon W \to V$ lineare con

$$g \circ f = \mathrm{Id}_V, \qquad f \circ g = \mathrm{Id}_W\,. \tag{3.2.1}$$

Per sottolineare che f è un isomorfismo scriviamo $f\colon V \xrightarrow{\sim} W$. Diciamo che V è *isomorfo* a W se esiste un isomorfismo $f\colon V \to W$.

Lemma 3.2.2. *Siano V, W spazi vettoriali su uno stesso campo $\mathbb{K}$. Un'applicazione **lineare** $f\colon V \to W$ è un isomorfismo se e solo se f è biunivoca.*

Dimostrazione. Se f è un isomorfismo allora è invertibile per definizione - vedi (3.2.1). Ora supponiamo che esista un'inversa g di f, cioè che valga (3.2.1), senza supporre che g sia lineare, e dimostriamo che g è lineare. Siano $w_1, w_2 \in W$ e $\lambda_1, \lambda_2 \in \mathbb{K}$. Abbiamo che

$$f(g(\lambda_1w_1 + \lambda_2w_2)) = \mathrm{Id}(\lambda_1w_1 + \lambda_2w_2) = \lambda_1w_1 + \lambda_2w_2$$

e

$$f(\lambda_1g(w_1) + \lambda_2g(w_2)) = \lambda_1f(g(w_1)) + \lambda_2f(g(w_2)) = \lambda_1w_1 + \lambda_2w_2.$$

Quindi $f(g(\lambda_1w_1 + \lambda_2w_2)) = f(\lambda_1g(w_1) + \lambda_2g(w_2))$. Siccome f è invertibile segue che

$$g(\lambda_1w_1 + \lambda_2w_2)) = \lambda_1g(w_1) + \lambda_2g(w_2)$$

e questo dimostra che g è lineare. □

Esempio 3.2.3. Sia V uno spazio vettoriale su $\mathbb{K}$, finitamente generato e sia $\mathcal{B} = \{v_1,\dots,v_n\}$ una base di V. L'applicazione

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{f} & V \\ (x_1,\dots,x_n) & \longrightarrow & x_1v_1 + x_2v_2 + \dots + x_nv_n \end{array} \tag{3.2.2}$$

è biunivoca per la Proposizione 2.6.7 e quindi f è un isomorfismo.

Osservazione 3.2.4. (1) Sia V uno spazio vettoriale: l'identità $\mathrm{Id}_V\colon V \to V$ è (banalmente) un isomorfismo.

(2) Sia $f\colon V \to W$ un isomorfismo tra spazi vettoriali su uno stesso campo $\mathbb{K}$. Per definizione anche f^{-1} è un isomorfismo.

(3) Siano U, V, W spazi vettoriali su uno stesso campo $\mathbb{K}$. Supponiamo che $f\colon U \to V$ e $g\colon V \to W$ siano isomorfismi: allora $g \circ f\colon U \to W$ è un isomorfismo (vedi la Proposizione 3.1.24).

Esempio 3.2.5. Sia V uno spazio vettoriale su $\mathbb{K}$, finitamente generato e sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V. Per l'Esempio 3.2.3 e il punto (2) dell'Osservazione 3.2.4, l'applicazione

$$V \xrightarrow{X_{\mathcal{B}}} \mathbb{K}^n, \tag{3.2.3}$$

che associa a un vettore di V il vettore delle sue coordinate, è un isomorfismo.

Esempio 3.2.6. Sia V uno spazio vettoriale su $\mathbb{K}$ finitamente generato, e sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una sua base. Per l'Esempio 3.1.6 possiamo definire un'applicazione biunivoca $\Phi\colon \mathbb{K}^n \to (\mathbb{K}^n)^\vee$ associando ad $(a_1, \dots, a_n) \in \mathbb{K}^n$ l'applicazione lineare in (3.1.16). Per le formule in (3.1.17) l'applicazione Φ è lineare, e quindi è un isomorfismo.

Supponiamo che $f\colon V \to W$ sia un isomorfismo tra spazi vettoriali sullo stesso campo $\mathbb{K}$. Per quanto concerne la struttura di spazio vettoriale possiamo identificare V e W: il risultato qui sotto dà una versione precisa di questa affermazione.

Proposizione 3.2.7. *Siano V, W spazi vettoriali su uno stesso campo $\mathbb{K}$ e supponiamo che $f\colon V \to W$ sia un isomorfismo. Siano $v_1, \dots, v_n \in V$.*

(1) $v_1, \dots, v_n$ sono linearmente dipendenti se e solo se $f(v_1), \dots, f(v_n) \in W$ sono linearmente dipendenti.

(2) $v_1, \dots, v_n$ generano V se e solo se $f(v_1), \dots, f(v_n)$ generano W.

Dimostrazione. (1): Supponiamo che $v_1, \dots, v_n$ siano linearmente dipendenti. Quindi esistono $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ non tutti nulli tali che

$$\lambda_1 v_1 + \dots + \lambda_n v_n = 0. \tag{3.2.4}$$

Applicando f a entrambi i membri di (3.2.4) e sfruttando la linearità di f otteniamo che $\lambda_1 f(v_1) + \dots + \lambda_n f(v_n) = 0$ e quindi $f(v_1), \dots, f(v_n) \in W$ sono linearmente dipendenti.

Ora supponiamo che $f(v_1), \dots, f(v_n) \in W$ siano linearmente dipendenti. Quindi esistono $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ non tutti nulli tali che

$$\lambda_1 f(v_1) + \dots + \lambda_n f(v_n) = 0. \tag{3.2.5}$$

Sia f^{-1} l'inversa di f (lineare per definizione di isomorfismo). Applicando f^{-1} a entrambi i membri di (3.2.5) e sfruttando la linearità di f^{-1} otteniamo che $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$. Questo dimostra che vale (1). Dimostriamo che vale (2). Supponiamo che V sia generato da $v_1, \dots, v_n$. Sia $w \in W$: allora esistono $\mu_1, \dots, \mu_n$ tali che

$$f^{-1}(w) = \mu_1 v_1 + \dots + \mu_n v_n. \tag{3.2.6}$$

Applicando f a entrambi i membri di (3.2.6) e sfruttando la linearità di f otteniamo che

$$w = \mu_1 f(v_1) + \dots + \mu_n f(v_n).$$

Quindi W è generato da $f(v_1), \dots, f(v_n)$ e perciò abbiamo dimostrato il "solo se". Rimane da dimostrare che se $f(v_1), \dots, f(v_n)$ generano W allora $v_1, \dots, v_n$ generano V. È sufficiente applicare quello che abbiamo appena dimostrato all'isomorfismo f^{-1} - vedi l'Osservazione 3.2.4. □

Il corollario qui sotto segue immediatamente dalla Proposizione 3.2.7.

Corollario 3.2.8. *Siano V, W spazi vettoriali su uno stesso campo $\mathbb{K}$ e supponiamo che $f\colon V \to W$ sia un isomorfismo. Assumiamo che V sia finitamente generato e sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una sua base. Allora W è finitamente generato e $\mathcal{C} = \{f(v_1), \dots, f(v_n)\}$ è una sua base. In particolare* $\dim V = \dim W$.

Per il Corollario 3.2.8 due spazi vettoriali finitamente generati isomorfi hanno la stessa dimensione. Vale il viceversa:

Proposizione 3.2.9. *Siano V, W spazi vettoriali su uno stesso campo $\mathbb{K}$. Supponiamo che V, W siano finitamente generati della stessa dimensione. Allora V è isomorfo a W.*

Dimostrazione. Sia $n := \dim V = \dim W$. Siano $\mathcal{B}$ e $\mathcal{C}$ basi di V e W rispettivamente. Allora, vedi l'Esempio 3.2.6, abbiamo isomorfismi

$$X_{\mathcal{B}} \colon V \xrightarrow{\sim} \mathbb{K}^n, \qquad X_{\mathcal{C}} \colon W \xrightarrow{\sim} \mathbb{K}^n,$$

e quindi $X_{\mathcal{C}}^{-1} \circ X_{\mathcal{B}} \colon V \to W$ è un isomorfismo - vedi l'Osservazione 3.2.4. □

Proposizione 3.2.10. *Siano V, W spazi vettoriali su uno stesso campo $\mathbb{K}$. Supponiamo che V, W siano finitamente generati e che $\dim V = \dim W$. Sia $f \colon V \to W$ lineare. Se*

(1) $\ker f = \{0\}$ *o*

(2) f è suriettiva

allora f è un isomorfismo.

Dimostrazione. (1): per la Proposizione 3.1.21 otteniamo che $\dim(\operatorname{im} f) = \dim V = \dim W$ e quindi f è suriettiva. D'altra parte parte f è iniettiva per la Proposizione 3.1.20. Per il Lemma 3.2.2 segue che f è un isomorfismo. (2): per la Proposizione 3.1.21 otteniamo che $\dim(\ker f) = \dim V - \dim W = 0$ e quindi f è iniettiva per la Proposizione 3.1.20. Per il Lemma 3.2.2 segue che f è un isomorfismo. □

Esempio 3.2.11. Sia W uno spazio vettoriale e siano $V_1, V_2 \subset W$ sottospazi. Ricordiamo che la somma diretta $V_1 \oplus V_2$ è definita nella Sezione 2.8. L'applicazione

$$\begin{array}{ccc} V_1 \oplus V_2 & \xrightarrow{f} & W \\ (v_1, v_2) & \mapsto & v_1 + v_2 \end{array}$$

è lineare (è una semplice verifica). La f è suriettiva se e solo se W è generato dai sottospazi V_1 e V_2, cioè $V_1 + V_2 = W$, ed è iniettiva se e solo se $V_1 \cap V_2 = \{0\}$ perchè

$$\ker(f) = \{(v, -v) \mid v \in V_1 \cap V_2\}.$$

Se valgono entrambe queste ipotesi, cioè f è un isomorfismo scriviamo (con un abuso di notazione)

$$W = V_1 \oplus V_2,$$

e diciamo, con un abuso della terminologia, che *W è la somma diretta di V_1 e V_2*. Analogamente, se $V_1, \ldots, V_m \subset W$ sono sottospazi vettoriali e l'applicazione lineare

$$\begin{array}{ccc} V_1 \oplus \ldots \oplus V_m & \xrightarrow{f} & W \\ (v_1, \ldots, v_m) & \mapsto & v_1 + \ldots + v_m \end{array} \tag{3.2.7}$$

è un isomorfismo scriviamo

$$W = V_1 \oplus \ldots \oplus V_m,$$

e diciamo che *W è la somma diretta di $V_1, \ldots, V_m$*. Attenzione: se $m > 2$ e la f in (3.2.7) è iniettiva allora $V_1 \cap \ldots \cap V_m = \{0\}$ ma *non vale* il viceversa.

Definizione 3.2.12. Sia V uno spazio vettoriale su un campo $\mathbb{K}$.

1. Un *automorfismo* di V è un isomorfismo $f \colon V \xrightarrow{\sim} V$.
2. $\mathrm{GL}(V)$ è l'insieme degli automorfismi $f \colon V \to V$.

Per l'Osservazione 3.2.4 valgono le seguenti proprietà:

1. se $f, g \in \mathrm{GL}(V)$ allora $f \circ g \in \mathrm{GL}(V)$,
2. $\mathrm{Id}_V \in \mathrm{GL}(V)$,
3. se $f \in \mathrm{GL}(V)$ allora $f^{-1} \in \mathrm{GL}(V)$, e
4. se $f, g, h \in \mathrm{GL}(V)$ allora $(f \circ g) \circ h = f \circ (g \circ h)$.

Quindi la composizione definisce un'operazione su $\mathrm{GL}(V)$, e con questa operazione $\mathrm{GL}(V)$ è un gruppo, che si chiama il *gruppo generale lineare*.

3.3 Il primo Teorema di isomorfismo

Siano V, W spazi vettoriali su $\mathbb{K}$, e sia $f\colon V \to W$ un'applicazione lineare. Sia

$$\pi\colon V \to V/\ker f$$

l'applicazione quoziente.

Proposizione 3.3.1. *Esiste una e una sola applicazione lineare $\overline{f}\colon V/\ker f \longrightarrow W$ tale che $\overline{f} \circ \pi = f$.*

Dimostrazione. Sia $[v] \in V/\ker f$. Definiamo $\overline{f}([v]) = f(v)$, ma dobbiamo verificare che la definizione è *ben posta*, cioè che il valore di $\overline{f}$ su una classe di equivalenza *non* dipende dal rappresentante scelto. Se $[v'] = [v]$, allora $(v' - v) \in \ker f$, e quindi

$$0 = f(v' - v) = f(v') - f(v),$$

cioè $f(v') = f(v)$. Vale $\overline{f} \circ \pi = f$ per definizione di $\overline{f}$. Una $\overline{f}$ tale che $\overline{f} \circ \pi = f$ è unica perchè l'applicazione quoziente π è suriettiva. Rimane da dimostrare che $\overline{f}$ è lineare. Se $\lambda_1, \lambda_2 \in \mathbb{K}$ e $v_1, v_2 \in V$,

$$\begin{aligned}\overline{f}(\lambda_1[v_1] + \lambda_2[v_2]) = f([\lambda_1 v_1 + \lambda_2 v_2]) = \\ = f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2) = \lambda_1 f([v_1]) + \lambda_2 f([v_2]). \quad (3.3.1)\end{aligned}$$

□

Ovviamente l'immagine di $\overline{f}$ è contenuta in $\operatorname{im} f$, e quindi definisce un'applicazione lineare $V/\ker f \longrightarrow \operatorname{im} f$ che continueremo a denotare $\overline{f}$ (abusando della notazione).

Teorema 3.3.2 (Primo Teorema di Isomorfismo)**.** *Mantenendo le ipotesi e notazioni appena introdotte, l'applicazione lineare $\overline{f}\colon V/\ker f \longrightarrow \operatorname{im} f$ è un isomorfismo.*

Dimostrazione. L'immagine di $\overline{f}$ è uguale all'immagine di f, e quindi $\overline{f}$ è suriettiva (su $\operatorname{im} f$!). Per finire basta dimostrare che $\overline{f}$ è iniettiva, cioè che se $f([v]) = 0$, allora $[v] = 0$. Ma $f([v]) = f(v)$, e quindi $v \in \ker f$, cioè $[v] = 0$. □

Osservazione 3.3.3. Mantenendo le ipotesi e notazioni appena introdotte, supponiamo che V sia finitamente generato. Allora $\dim(V/\ker f) = \dim(\operatorname{im} f)$ per il Primo Teorema di Isomorfismo, ma d'altra parte $\dim(V/\ker f) = \dim V - \dim(\ker f)$ per la Proposizione 2.8.7. Questo dimostra di nuovo che $\dim V = \dim(\ker f) + \dim(\operatorname{im} f)$, cioè la Proposizione 3.1.21.

3.4 Matrici

Le matrici sono uno strumento indispensabile per fare calcoli con le applicazioni lineari. Cominceremo definendo le operazioni tra matrici, e poi inizieremo a stabilire la relazione tra matrici e applicazioni lineari. Avremo a che fare prevalentemente con matrici le cui entrate appartengono a un campo (fissato), ma le definizioni fondamentali si formulano per matrici con entrate in un anello.

Calcolo matriciale

Sia R un anello (commutativo con unità). Una *matrice $m \times n$ a valori in R* è un'applicazione $f\colon \{1,\dots,m\} \times \{1,\dots,n\} \to R$: quindi è determinata dall'insieme dei valori $f(i,j)$ associati a (i,j) dove $1 \leqslant i \leqslant m$ e $1 \leqslant j \leqslant n$. Invece della notazione appena introdotta denotiamo una matrice con una lettera maiuscola, per esempio A, e denotiamo il valore della matrice su (i,j) con a_{ij}. Si scrive $A = (a_{ij})$. È conveniente scrivere la matrice come una tabella:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

Una matrice è *quadrata* se il numero delle sue righe è uguale al numero delle sue colonne.

Sia A una matrice $m \times n$. La *riga i-esima* della matrice $m \times n$ A è la matrice $1 \times n$ data da

$$A^i := [a_{i1}, a_{i2}, \dots, a_{i_n}]. \tag{3.4.1}$$

La *colonna j-esima* di A è la matrice $m \times 1$ data da

$$A_j := \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \tag{3.4.2}$$

Quindi a_{ij} è l'*entrata della matrice sulla riga i(-esima) e la colonna j(-esima).*

Definizione 3.4.1. Se R è un anello, $M_{m,n}(R)$ è l'insieme delle matrici $m \times n$ a valori in R. Indichiamo con $0_{m,n}$ (o con 0 quando non c'è possibiltà di confusione) la matrice $m \times n$ con entrate tutte nulle.

Esistono alcune operazioni fondamentali sulle matrici. La somma è definita da

$$\begin{array}{ccc} M_{m,n}(R) \times M_{m,n}(R) & \longrightarrow & M_{m,n}(R) \\ ((a_{ij}), (b_{ij})) & \mapsto & (a_{ij} + b_{ij}) \end{array}$$

Questo significa che l'entrata della matrice somma di A e B su riga i e colonna j è la somma $a_{ij} + b_{ij}$. Con questa operazione $M_{m,n}(R)$ è un gruppo abeliano. La moltiplicazione per scalari (in R) è definita da

$$\begin{array}{ccc} R \times M_{m,n}(R) & \longrightarrow & M_{m,n}(R) \\ (\lambda, (a_{ij})) & \mapsto & (\lambda a_{ij}) \end{array}$$

Questo significa che l'entrata della matrice A su riga i e colonna j è λa_{ij}.

Proposizione 3.4.2. *Per la somma $M_{m,n}(R) \times M_{m,n}(R) \to M_{m,n}(R)$ e il prodotto per uno scalare $R \times M_{m,n}(R) \to M_{m,n}(R)$ valgono le proprietà richieste per somma e prodotto per uno scalare di uno spazio vettoriale (vedi la Definizione 2.1.1). In particolare se $\mathbb{K}$ è un campo, allora $M_{m,n}(\mathbb{K})$ con le operazioni appena definite è uno spazio vettoriale su $\mathbb{K}$.*

Dimostrazione. Segue immediatamente dalle definizioni di somma e prodotto per uno scalare e dalle proprietà dell'anello R. L'elemento neutro di $M_{m,n}(R)$ è $0_{m,n}$. □

Osservazione 3.4.3. Sia $\mathbb{K}$ un campo. Allora l'applicazione

$$\begin{array}{ccc} M_{m,n}(\mathbb{K}) & \xrightarrow{\sim} & \mathbb{K}^{mn} \\ (a_{ij}) & \mapsto & (a_{11}, a_{12}, \dots, a_{1n}, a_{21}, \dots, a_{2n}, \dots, a_{m1}, \dots, a_{mn}) \end{array}$$

è un isomorfismo di spazi vettoriali su $\mathbb{K}$.

Esiste un'altra operazione fondamentale tra matrici.

Definizione 3.4.4. Sia R un anello. Siano $A \in M_{m,n}(R)$ e $B \in M_{n,p}(R)$. La *moltiplicazione righe per colonne* di $A \cdot B$ è la matrice $m \times p$ definita nel seguente modo. Siano $A = (a_{ij})$ e $B = (b_{jh})$. L'entrata c_{ih} (per $1 \leqslant i \leqslant 1m$ e $1 \leqslant h \leqslant p$) di $A \cdot B$ è data da

$$c_{ih} := \sum_{j=1}^{n} a_{ij} b_{jh}.$$

Consideriamo il caso in cui $m = 1 = p$: quindi

$$A = [a_{11}, \ldots, a_{1n}], \qquad B = \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix}$$

Allora

$$A \cdot B = a_{11}b_{11} + a_{12}b_{21} + \ldots + a_{1n}b_{n1}.$$

In generale

$$A \cdot B = \begin{bmatrix} A^1 \cdot B_1 & A^1 \cdot B_2 & \ldots & A^1 \cdot B_n \\ A^2 \cdot B_1 & A^2 \cdot B_2 & \ldots & A^2 \cdot B_n \\ \ldots & \ldots & \ldots & \ldots \\ A^m \cdot B_1 & A^m \cdot B_2 & \ldots & A^m \cdot B_n \end{bmatrix} \tag{3.4.3}$$

Questo giustifica il nome "moltiplicazione righe per colonne".

Esempio 3.4.5. Siano $A, B \in M_{2,2}(\mathbb{R})$ le matrici

$$A := \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}, \quad B := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{3.4.4}$$

Sia A che B sono 2×2, quindi ha senso moltiplicarle in qualsiasi ordine. Calcolando otteniamo che

$$A \cdot B := \begin{bmatrix} 0 & \lambda \\ 0 & 0 \end{bmatrix}, \quad B \cdot A := \begin{bmatrix} 0 & \mu \\ 0 & 0 \end{bmatrix}, \quad B \cdot B := \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{3.4.5}$$

I primi due prodotti di (3.4.5) fanno vedere che in generale il prodotto di matrici quadrate dello stesso ordine (e che perciò possono essere moltiplicate in qualsiasi ordine) **non** è commutativo, e il terzo prodotto di (3.4.5) dà una matrice non nulla il cui quadrato è nullo.

L'Esempio 3.4.5 dimostra che la moltiplicazione tra matrici non gode di tutte le proprietà del prodotto tra numeri reali a cui siamo abituati. Non tutto è perduto però: il prodotto tra matrici gode di alcune delle proprietà del prodotto tra numeri reali. Prima di elencare tali proprietà diamo un paio di definizioni. Siano $i, j \in \mathbb{N}$: il *simbolo di Kronecker* δ_{ij} è

$$\delta_{ij} := \begin{cases} 1 & \text{se } i = j, \\ 0 & \text{se } i \neq j. \end{cases} \tag{3.4.6}$$

Definizione 3.4.6. (1) La matrice *unità* $n \times n$ è la matrice $1_n := (\delta_{ij})$ (qui $1 \leqslant i, j \leqslant n$).

(2) Una matrice $A \in M_{n,n}(R)$ è *scalare* se esiste $\lambda \in R$ tale che $M = \lambda 1_n$.

(3) Una matrice $A \in M_{n,n}(R)$ è *diagonale* se esistono $\lambda_i \in R$ per $1 \leqslant i \leqslant n$ tali che $A = (\lambda_i \delta_{ij})$. In altre parole $A = (a_{ij})$ è diagonale se $a_{ij} = 0$ per ogni i, j con $i \neq j$.

Proposizione 3.4.7. *Siano $\lambda \in R$, $A \in M_{m,n}(R)$, $B, B' \in M_{n,p}(R)$ e $C \in M_{p,q}(R)$. Allora*

(1) $(\lambda 1_m)\cdot A = \lambda A = A\cdot(\lambda 1_n)$,

(2) $(A\cdot B)\cdot C = A\cdot(B\cdot C)$ *(proprietà associativa),*

(3) $A\cdot(B+B') = A\cdot B + A\cdot B'$ *e* $(B+B')\cdot C = B\cdot C + B'\cdot C$*(proprietà distributiva).*

Dimostrazione. (1): dimostriamo che $(\lambda 1_m)\cdot A = \lambda A$. Sia $A = (a_{ij})$ e poniamo $(\lambda 1_m)\cdot A = (b_{ih})$. Per definizione di prodotto abbiamo

$$b_{ih} = \sum_{j=1}^{m} \lambda\delta_{ij}a_{jh} = \lambda a_{ih}.$$

Questo dimostra che $(\lambda 1_m)\cdot A = \lambda A$. L'uguaglianza $A\cdot(\lambda 1_n) = \lambda A$ si dimostra con un calcolo simile. (2): sia $A = (a_{ij})$, $B = (b_{jh})$ e $C = (c_{hl})$. Poniamo $(A\cdot B)\cdot C = (s_{il})$ e $A\cdot(B\cdot C) = (t_{il})$ Per definizione di prodotto abbiamo

$$s_{il} = \sum_{h=1}^{p}(\sum_{j=1}^{n} a_{ij}b_{jh})c_{hl} = \sum_{\substack{1\leqslant j\leqslant n\\ 1\leqslant h\leqslant p}} a_{ij}b_{jh}c_{hl}$$

e

$$t_{il} = \sum_{j=1}^{n} a_{ij}(\sum_{h=1}^{p} b_{jh}c_{hl}) = \sum_{\substack{1\leqslant j\leqslant n\\ 1\leqslant h\leqslant p}} a_{ij}b_{jh}c_{hl}$$

Quindi $s_{ij} = t_{ij}$ e perciò vale (2). Dimostriamo che vale la prima eguaglianza di (3): se $m = 1 = p$ la (3) segue da un facile conto, il caso generale segue dal caso $m = 1 = p$ per la Formula (3.4.3). La seconda eguaglianza di (3) si verifica in modo simile. □

Osservazione 3.4.8. La moltiplicazione tra matrici permette di scrivere un sistema di m equazioni lineari (a coefficienti in $\mathbb{K}$) in n incognite (con valori in $\mathbb{K}$)

$$\begin{array}{ccccccc} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1 \\ \vdots \quad \vdots \quad \vdots & \vdots & \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n & = & b_i \\ \vdots \quad \vdots \quad \vdots & \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m \end{array} \tag{3.4.7}$$

in maniera compatta. Infatti siano $A \in M_{m,n}(\mathbb{K})$ la matrice con entrata a_{ij} su riga i e colonna j e $B \in M_{m,1}(\mathbb{K})$ la matrice con entrata b_i su riga i (è superfluo specificare la colonna perchè è unica). Scriviamo X per denotare la matrice colonna j e $n\times 1$ con entrata l'incognita x_j sulla riga j. Allora il sistema di equazioni in (3.4.7) equivale all'equazione $A\cdot X = B$. Questa scrittura mostra l'analogia con una semplice equazione $a\cdot x = b$, dove $a, b \in \mathbb{K}$ e x è l'incognita in $\mathbb{K}$, cioè il caso $m = n = 1$. Vedi l'Osservazione (3.5.8).

Matrici quadrate

Notiamo che il prodotto di due matrici in $M_{n,n}(R)$ è una matrice in $M_{n,n}(R)$. Viceversa se $A \in M_{m,n}(R)$, ed esiste $B \in M_{p,q}(R)$ tale che abbiano senso sia $A\cdot B$ che $B\cdot A$ allora $m = n$ (e quindi $p = q = n$).

Osservazione 3.4.9. Su $M_{n,n}(R)$ abbiamo due operazioni, la somma e il prodotto righe per colonne. Con queste operazioni $M_{n,n}(R)$ è un anello grazie alle Proposizioni 3.4.2 e 3.4.7. Inoltre, per (1) della Proposizione 3.4.7 l'elemento 1_n è un'unità, e quindi $M_{n,n}(R)$ è un anello con unità. Se $n = 1$ allora l'anello $M_{1,1}(R)$ è identificato con R e quindi è commutativo, ma se $n > 1$ allora $M_{n,n}(R)$ non è commutativo. Infatti se $n = 2$ basta considerare le matrici A, B di (3.4.4), e se $n > 2$ basta estendere le A, B inserendo entrate nulle nelle altre entrate. $M_{n,n}(R)$ è un esempio significativo di anello con unità non commutativo.

Definizione 3.4.10. Sia $A \in M_{n,n}(R)$ per un qualche n. Una matrice $B \in M_{n,n}(R)$ è un'*inversa* di A se

$$A \cdot B = 1_n = B \cdot A.$$

$A \in M_{n,n}(R)$ è *invertibile* se ha un'inversa.

Esempio 3.4.11. Siano A la matrice di (3.4.4). Allora A ha un'inversa se e solo se $\lambda \neq 0 \neq \mu$. Infatti se $\lambda \neq 0 \neq \mu$, l'inversa è data da

$$A^{-1} := \begin{bmatrix} \lambda^{-1} & 0 \\ 0 & \mu^{-1} \end{bmatrix} \tag{3.4.8}$$

D'altra parte se $\lambda = 0$, allora per qualsiasi $D \in M_{2,2}(\mathbb{K})$ la matrice $D \cdot A$ ha la prima colonna nulla e quindi non è la matrice unità, e se $\mu = 0$, allora per qualsiasi $D \in M_{2,2}(\mathbb{K})$ la matrice $A \cdot D$ ha la seconda colonna nulla e quindi non è la matrice unità. Questo mostra che, mentre in $\mathbb{K}$ ogni elemento non nullo è invertibile, in $M_{n,n}(\mathbb{K})$ per $n \geqslant 2$ esistono matrici non nulle che non sono invertibili.

Per evitare fraintendimenti, notiamo che la formula in (3.4.9) vale perchè A è una matrice diagonale, in generale l'inversa di una matrice invertibile M non ha entrate sulla diagonale principale date dagli inversi delle entrate di M. Per esempio se

$$M := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{3.4.9}$$

allora M è invertibile e $M^{-1} = M$ (ma le entrate sulla diagonale principale di M sono nulle).

Lemma 3.4.12. *Se $A, B \in M_{n,n}(R)$ sono invertibili allora $A \cdot B$ è invertibile.*

Dimostrazione. Siano A', B' inverse di A, B rispettivamente. Allora $B' \cdot A'$ è un'inversa di $A \cdot B$ (notate lo scambio nell'ordine dei fattori). Infatti

$$(B' \cdot A') \cdot (A \cdot B) = B' \cdot (A' \cdot A) \cdot B = B' \cdot 1_n \cdot B = 1_n,$$

e analogamente si verifica che $(A \cdot B) \cdot (B' \cdot A') = 1_n$. □

Poniamo

$$\mathrm{GL}_n(R) := \{A \in M_{n,n}(R) \mid A \text{ è invertibile}\}. \tag{3.4.10}$$

Per il Lemma 3.4.12 la moltiplicazione definisce un'operazione

$$\begin{array}{ccc} \mathrm{GL}_n(R) \times \mathrm{GL}_n(R) & \longrightarrow & \mathrm{GL}_n(R) \\ (A, B) & \mapsto & A \cdot B \end{array} \tag{3.4.11}$$

Con questa operazione $\mathrm{GL}_n(R)$ è un gruppo. Infatti 1_n è una unità, l'associatività vale per la Proposizione 3.4.7 e ogni elemento di $\mathrm{GL}_n(R)$ ha un'inversa per definizione. Siccome $\mathrm{GL}_n(R)$ è un gruppo, ogni suo elemento ha un'unica inversa.

Esempio 3.4.13. $\mathrm{GL}_n(\mathbb{Z}) = \{\pm 1\}$, mentre se $\mathbb{K}$ è un campo $\mathrm{GL}_n(\mathbb{K}) = \mathbb{K}^* = \mathbb{K}\backslash\{0\}$. Più in là descriveremo esplicitamente gli elementi di $\mathrm{GL}_n(\mathbb{K})$ per ogni n per $\mathbb{K}$ un campo.

Definizione 3.4.14. Sia $A \in M_{n,n}(R)$ invertibile. Denotiamo con A^{-1} l'unica inversa di A.

Osservazione 3.4.15. Sia $A \in M_{n,n}(R)$ invertibile. Se $r \in \mathbb{Z}$ poniamo

$$A^r := \begin{cases} \underbrace{A \cdot \ldots \cdot A}_{r} & \text{se } r > 0, \\ 1_n & \text{se } r = 0, \\ \underbrace{A^{-1} \cdot \ldots \cdot A^{-1}}_{-r} & \text{se } r < 0. \end{cases}$$

Con questa definizione $A^r \cdot A^s = A^{r+s}$ per ogni $r, s \in \mathbb{Z}$.

Sarà utile considerare la seguente operazione che produce una matrice $n \times m$ a partire da una matrice $m \times n$.

3.5 Da una matrice in $M_{m,n}(\mathbb{K})$ a un'applicazione lineare $\mathbb{K}^n \to \mathbb{K}^m$

Iniziamo a studiare la relazione tra matrici e applicazioni lineari, e quindi consideriamo matrici con entrate in un campo $\mathbb{K}$.

Convenzione 3.5.1. Identifichiamo $\mathbb{K}^n$ con $M_{n,1}(\mathbb{K})$ per mezzo dell'isomorfismo di spazi vettoriali

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{\sim} & M_{n,1}(\mathbb{K}) \\ (x_1, \dots, x_n) & \mapsto & \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \end{array}$$

Sia $A \in M_{m,n}(\mathbb{K})$: per la Convenzione 3.5.1 ha senso la definizione dell'applicazione

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{L_A} & \mathbb{K}^m \\ X & \mapsto & A \cdot X \end{array} \tag{3.5.12}$$

Per (1) e (3) della Proposizione 3.4.7 l'applicazione L_A è lineare. La seguente semplice osservazione è fondamentale.

Osservazione 3.5.2. Sia $A \in M_{m,n}(\mathbb{K})$, e sia $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ la base standard di $\mathbb{K}^n$. Allora

$$L_A(\mathbf{e}_j) = A_j, \qquad j \in \{1, \dots, n\}. \tag{3.5.13}$$

Proposizione 3.5.3. *Sia $f\colon \mathbb{K}^n \to \mathbb{K}^m$ un'applicazione lineare. Esiste una e una sola matrice $A \in M_{m,n}(\mathbb{K})$ tale che $f = L_A$.*

Dimostrazione. La formula (3.5.13) dà che A è univocamente determinata (se esiste) da f: infatti vediamo che f determina le colonne di A e quindi A stessa. Ora supponiamo che $f\colon \mathbb{K}^n \to \mathbb{K}^m$ sia lineare. Definiamo $A \in M_{m,n}(\mathbb{K})$ imponendo che valga (3.5.13). Dimostriamo che $L_A = f$. Per (3.5.13) abbiamo $L_A(\mathbf{e}_j) = f(\mathbf{e}_j)$ per ogni vettore $\mathbf{e}_j$ della base standard di $\mathbb{K}^n$. Siccome L_A e f sono lineari, segue dalla Proposizione 3.1.11 che $L_A = f$. □

Per la Proposizione 3.5.3 abbiamo un'applicazione biunivoca

$$\begin{array}{ccc} M_{m,n}(\mathbb{K}) & \longrightarrow & \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m) \\ A & \mapsto & L_A \end{array} \tag{3.5.14}$$

Notiamo che $M_{m,n}(\mathbb{K})$ e $\mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ sono $\mathbb{K}$-spazi vettoriali.

Proposizione 3.5.4. *L'applicazione in* (3.5.14) *è un isomorfismo di spazi vettoriali.*

Dimostrazione. L'applicazione in (3.5.14) è biunivoca per la Proposizione 3.5.3, quindi basta dimostrare che è lineare. Siano $A, B \in M_{m,n}(\mathbb{K})$. Dimostriamo che

$$L_{A+B} = L_A + L_B. \tag{3.5.15}$$

Sia $X \in M_{n,1}(\mathbb{K})$. Allora

$$L_{A+B}(X) = (A + B) \cdot X = A \cdot X + B \cdot X = L_A(X) + L_B(X),$$

(la seconda uguaglianza vale per la distributività del prodotto tra matrici), e questo dimostra che vale l'uguaglianza in (3.5.15). Si dimostra in modo simile che se $\lambda \in \mathbb{K}$ allora $L_{\lambda A} = \lambda L_A$. □

Proposizione 3.5.5. *Se $A \in M_{m,n}(\mathbb{K})$ e $B \in M_{n,p}(\mathbb{K})$, allora*

$$L_A \circ L_B = L_{A \cdot B}. \tag{3.5.16}$$

Dimostrazione. Sia $X \in \mathbb{K}^p$ (vettore colonna): per l'associatività del prodotto di matrici abbiamo che

$$(L_A \circ L_B)(X) = L_A(L_B(X)) = A \cdot (B \cdot X) = (A \cdot B) \cdot X = L_{A \cdot B}(X).$$

□

Corollario 3.5.6. *Sia $A \in M_{n,n}(\mathbb{K})$. Allora L_A è un isomorfismo se e solo se A è invertibile.*

Dimostrazione. Supponiamo che A sia invertibile. Allora

$$L_A \circ L_{A^{-1}} = L_{1_n} = \mathrm{Id}_{\mathbb{K}^n}, \quad L_{A^{-1}} \circ L_A = L_{1_n} = \mathrm{Id}_{\mathbb{K}^n},$$

e quindi L_A è un isomorfismo. Ora supponiamo che L_A sia un isomorfismo, e sia φ l'inversa di L_A. Per la Proposizione 3.5.3 esiste $B \in M_{n,n}(\mathbb{K})$ tale che $\varphi = L_B$. Quindi

$$L_{1_n} = \mathrm{Id}_n = L_A \circ L_B = L_{A \cdot B}, \quad L_{1_n} = \mathrm{Id}_n = L_B \circ L_A = L_{B \cdot A}.$$

Quindi per la Proposizione 3.5.3 abbiamo $A \cdot B = 1_n$ e $B \cdot A = 1_n$, e perciò A è invertibile. □

Osservazione 3.5.7. La Proposizione 3.5.3 e il Corollario 3.5.6 mostrano che abbiamo un'applicazione biunivoca

$$\begin{array}{ccc} \mathrm{GL}_n(\mathbb{K}) & \longrightarrow & \mathrm{GL}(\mathbb{K}^n) \\ A & \mapsto & L_A \end{array} \tag{3.5.17}$$

(Ricordiamo che $\mathrm{GL}_n(\mathbb{K})$ è il gruppo degli isomorfismi di $\mathbb{K}^n$, secondo la Definizione (3.2.12).) Inoltre l'applicazione in (3.5.17) è un isomorfismo per la Proposizione 3.5.5. Per questo motivo identifichiamo $\mathrm{GL}_n(\mathbb{K})$ e $\mathrm{GL}(\mathbb{K}^n)$.

Osservazione 3.5.8. La Proposizione 3.2.10 e il Corollario 3.5.6 danno il seguente risultato non banale: il sistema di equazioni lineari

$$\begin{array}{lllll} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1 \\ \vdots \quad \vdots \quad \vdots & \vdots & \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n & = & b_i \\ \vdots \quad \vdots \quad \vdots & \vdots & \vdots \\ a_{n1}x_1 + a_{m2}x_2 + \ldots + a_{nn}x_n & = & b_n. \end{array} \tag{3.5.18}$$

(notate che ci sono tante equazioni quante incognite) ha soluzione per **ogni** scelta di $b_1, \ldots, b_n$ se e solo se il sistema omogeneo associato, ottenuto ponendo $0 = b_1 = \ldots = b_n$, ha solo la soluzione banale.

Riscriviamo il sistema di equazioni lineari in (3.5.18) come $A \cdot X = B$, con la notazione dell'Osservazione 3.4.8. Supponiamo che A sia invertibile. Allora, moltiplicando a sinistra per A^{-1} ambo i membri dell'equazione $A \cdot X = B$, vediamo che l'unica soluzione è data da

$$X = A^{-1} \cdot B.$$

3.6 Da un'applicazione lineare a una matrice

Nella sezione 3.5 abbiamo associato a una matrice $A \in M_{m,n}(\mathbb{K})$ l'applicazione lineare $L_A \colon \mathbb{K}^n \to \mathbb{K}^m$. Questa corrispondenza biunivoca permette di interpretare le operazioni tra applicazioni lineari come operazioni tra matrici, per esempio alla somma di applicazioni lineari corrisponde la somma di matrici, alla composizione di applicazioni lineari corrisponde il prodotto righe per colonne di matrici. Se V, W sono spazi vettoriali finitamente generati su $\mathbb{K}$, esiste un'analoga corrispondenza tra applicazioni lineari $V \to W$ e matrici una volta che si siano scelte basi di V e di W. Più precisamente, siano $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V e $\mathcal{C} = \{w_1, \ldots, w_m\}$ una base di W. Per ogni $j \in \{1, \ldots, n\}$ esistono $a_{1j}, \ldots, a_{mj} \in \mathbb{K}$ tali che

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i. \tag{3.6.1}$$

Definizione 3.6.1. La matrice $M_{\mathcal{C}}^{\mathcal{B}}(f)$ associata a f è la matrice $A \in M_{m,n}(\mathbb{K})$ con entrata a_{ij} su riga i e colonna j, dove per $i \in \{1,\dots,m\}$ e $j \in \{1,\dots,n\}$ l'entrata $a_{ij} \in \mathbb{K}$ è definita dall'uguaglianza in (3.6.1).

In altre parole la colonna j-esima di $A = M_{\mathcal{C}}^{\mathcal{B}}(f)$ è la colonna delle coordinate di $f(v_j)$ nella base $\mathcal{C}$.

Esempio 3.6.2. Siano $\mathbb{K} = \mathbb{R}$, $V = W = \mathbb{R}[x]_{\leqslant 2}$ e $\mathcal{B} = \mathcal{C} = \{1, x, x^2\}$. Sia

$$\begin{array}{ccc} \mathbb{R}[x]_{\leqslant 2} & \xrightarrow{f} & \mathbb{R}[x]_{\leqslant 2} \\ p & \mapsto & p + p' \end{array}$$

La f è lineare e

$$f(1) = 1, \qquad f(x) = x + 1, \qquad f(x^2) = x^2 + 2x.$$

Quindi

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Esempio 3.6.3. Siano $\mathbb{K} = \mathbb{R}$, $V = \mathsf{V}(\mathbb{E}^2)$ e sia $\mathcal{B} = \mathcal{C} = \{\mathbf{i}, \mathbf{j}\}$ dove $\mathbf{i}, \mathbf{j}$ sono vettori di uguale lunghezza e ortogonali tra loro. Sia $\rho\colon \mathbb{E}^2 \to \mathbb{E}^2$ la rotazione di angolo θ con verso di rotazione "da $\mathbf{i}$ a $\mathbf{j}$" intorno a un punto fissato, e sia $\mathsf{V}(\rho)\colon \mathsf{V}(\mathbb{E}^2) \to \mathsf{V}(\mathbb{E}^2)$ l'associata applicazione lineare, vedi l'Esempio 3.1.8. Si ha

$$M_{\mathcal{B}}^{\mathcal{B}}(\mathsf{V}(\rho)) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{3.6.2}$$

Esempio 3.6.4. Sia $V = \mathbb{K}^n$ e $W = \mathbb{K}^m$ (quindi il campo è $\mathbb{K}$). Siano $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ e $\mathcal{C} = \{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ le basi standard di $\mathbb{K}^n$ e $\mathbb{K}^m$ rispettivamente. Sia $A \in M_{m,n}(\mathbb{K})$: allora

$$M_{\mathcal{C}}^{\mathcal{B}}(L_A) = A.$$

Questo mostra che ciò che abbiamo discusso nella sezione 3.5 è un caso particolare di quello che stiamo descrivendo in questa sezione.

Proposizione 3.6.5. *Siano V, W spazi vettoriali sullo stesso campo $\mathbb{K}$, finitamente generati. Siano $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{C} = \{w_1, \dots, w_m\}$ basi rispettivamente di V e W. Se $v \in V$, allora*

$$X_{\mathcal{C}}(f(v)) = M_{\mathcal{C}}^{\mathcal{B}}(f) \cdot X_{\mathcal{B}}(v). \tag{3.6.3}$$

Viceversa, se $A \in M_{m,n}(\mathbb{K})$ e per ogni $v \in V$ vale

$$X_{\mathcal{C}}(f(v)) = A \cdot X_{\mathcal{B}}(v), \tag{3.6.4}$$

allora $A = M_{\mathcal{C}}^{\mathcal{B}}(f)$.

Dimostrazione. Dimostriamo che vale (3.6.3) per ogni $v \in V$. Poniamo $M_{\mathcal{C}}^{\mathcal{B}}(f) = (a_{ij})$ e

$$X_{\mathcal{B}}(v) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Per linearità di f e per definizione di $M_{\mathcal{C}}^{\mathcal{B}}(f)$ abbiamo

$$f(v) = f(\sum_{j=1}^{n} x_j v_j) = \sum_{j=1}^{n} x_j f(v_j) = \sum_{j=1}^{n} x_j (\sum_{i=1}^{m} a_{ij} w_i) = \sum_{i=1}^{m} (\sum_{j=1}^{n} a_{ij} x_j) w_i.$$

Quindi la coordinata i-esima di $f(v)$ è il prodotto della riga i-esima di A per la matrice colonna $X_{\mathcal{B}}(v)$, ovvero vale (3.6.3).

Ora supponiamo che valga (3.6.4) per ogni $v \in V$. Poniamo $v = v_j$ per $j \in \{1, \dots, n\}$. Siccome $X_{\mathcal{B}}(v_j) = \mathbf{e}_j$, otteniamo che

$$X_{\mathcal{C}}(f(v_j)) = A \cdot \mathbf{e}_j = A_j,$$

cioè la colonna j di A è uguale alla colonna delle coordinate (nella base $\mathcal{C}$) di $f(v_j)$. Quindi le colonne di A sono uguali alle colonne di $M_{\mathcal{C}}^{\mathcal{B}}(f)$, e perciò $A = M_{\mathcal{C}}^{\mathcal{B}}(f)$. □

Proposizione 3.6.6. *Siano V, W spazi vettoriali sullo stesso campo $\mathbb{K}$, finitamente generati. Siano $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{C} = \{w_1, \dots, w_m\}$ basi di V e W rispettivamente. L'applicazione*

$$\begin{array}{ccc} \mathcal{L}(V,W) & \xrightarrow{M_{\mathcal{B}}^{\mathcal{C}}} & M_{m,n}(\mathbb{K}) \\ f & \mapsto & M_{\mathcal{B}}^{\mathcal{C}}(f) \end{array} \tag{3.6.5}$$

è biunivoca.

Dimostrazione. Dimostriamo che l'applicazione $M_{\mathcal{B}}^{\mathcal{C}}$ è iniettiva. Siano $f, g \in \mathcal{L}(V,W)$ tali che $M_{\mathcal{B}}^{\mathcal{C}}(f) = M_{\mathcal{B}}^{\mathcal{C}}(g)$. Siccome le colonne j di $M_{\mathcal{B}}^{\mathcal{C}}(f)$ e $M_{\mathcal{B}}^{\mathcal{C}}(g)$ sono ripettivamente $X_{\mathcal{C}}(f(v_j))$ e $X_{\mathcal{C}}(g(v_j))$, segue che $f(v_j) = g(v_j)$ per $j \in \{1, \dots, n\}$. Per la Proposizione 3.1.11 segue che $f = g$.

Ora dimostriamo che l'applicazione $M_{\mathcal{B}}^{\mathcal{C}}$ è suriettiva. Sia $A \in M_{m,n}(\mathbb{K})$. Per la Proposizione 3.1.12, esiste un'applicazione lineare $f\colon V \to W$ tale che per $j \in \{1, \dots, n\}$

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

Chiaramente $M_{\mathcal{B}}^{\mathcal{C}}(f) = A$. □

Proposizione 3.6.7. *(1) Siano V, W spazi vettoriali sullo stesso campo $\mathbb{K}$, finitamente generati. L'applicazione* (3.6.5) *è un isomorfismo di spazi vettoriali.*

(2) Siano U, V, W spazi vettoriali su $\mathbb{K}$, finitamente generati. Siano $\mathcal{B}$ una base di U, $\mathcal{C}$ una base di V e $\mathcal{D}$ una base di W. Se $g\colon U \to V$ e $f\colon V \to W$ sono applicazioni lineari, allora

$$M_{\mathcal{D}}^{\mathcal{B}}(f \circ g) = M_{\mathcal{D}}^{\mathcal{C}}(f) \cdot M_{\mathcal{C}}^{\mathcal{B}}(g). \tag{3.6.6}$$

Dimostrazione. Dimostriamo il punto (1). Siano $\lambda_1, \lambda_2 \in \mathbb{K}$ e $f_1, f_2 \in \mathcal{L}(V,W)$. Per linearità di $X_{\mathcal{C}}$ e per la Proposizione 3.6.5 abbiamo

$$\begin{aligned} X_{\mathcal{C}}((\lambda_1 f_1 + \lambda_2 f_2)(v)) &= X_{\mathcal{C}}((\lambda_1 f_1(v) + \lambda_2 f_2(v)) = \lambda_1 X_{\mathcal{C}}(f_1(v)) + \lambda_2 X_{\mathcal{C}}(f_2(v)) = \\ &= \lambda_1 M_{\mathcal{B}}^{\mathcal{C}}(f_1) X_{\mathcal{B}}(v) + \lambda_2 M_{\mathcal{B}}^{\mathcal{C}}(f_2) X_{\mathcal{B}}(v) = (\lambda_1 M_{\mathcal{B}}^{\mathcal{C}}(f_1) + \lambda_2 M_{\mathcal{B}}^{\mathcal{C}}(f_2)) X_{\mathcal{B}}(v). \end{aligned}$$

Per la Proposizione 3.6.5 concludiamo che

$$M_{\mathcal{B}}^{\mathcal{C}}(\lambda_1 f_1 + \lambda_2 f_2) = (\lambda_1 M_{\mathcal{B}}^{\mathcal{C}}(f_1) + \lambda_2 M_{\mathcal{B}}^{\mathcal{C}}(f_2))$$

cioè (3.6.5) è lineare: siccome è biunivoca è un isomorfismo per il Lemma 3.2.2.

Dimostriamo il punto (2). Abbiamo

$$\begin{aligned} X_{\mathcal{D}}((f \circ g)(v)) = X_{\mathcal{D}}((f(g(v))) &= M_{\mathcal{D}}^{\mathcal{C}}(f) \cdot X_{\mathcal{C}}(g(v)) = \\ &= M_{\mathcal{D}}^{\mathcal{C}}(f) \cdot (M_{\mathcal{C}}^{\mathcal{B}}(g) \cdot X_{\mathcal{B}}(v)) = (M_{\mathcal{D}}^{\mathcal{C}}(f) \cdot M_{\mathcal{C}}^{\mathcal{B}}(g)) \cdot X_{\mathcal{B}}(v). \end{aligned}$$

Per la Proposizione 3.6.5 concludiamo che vale (2). □

Esempio 3.6.8. Siano $\alpha, \beta \in \mathbb{R}$. Applichiamo la (2) della Proposizione 3.6.7 alla rotazione $r_{\alpha+\beta}$ dell'Esempio 3.6.3. La base $\mathcal{B}$ di $\mathcal{V}^2$ è come nell'Esempio 3.6.3. Siccome $r_{\alpha+\beta} = r_\alpha \circ r_\beta$ otteniamo che

$$\begin{bmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{bmatrix} = M_{\mathcal{B}}^{\mathcal{B}}(r_{\alpha+\beta}) = M_{\mathcal{B}}^{\mathcal{B}}(r_\alpha) \cdot M_{\mathcal{B}}^{\mathcal{B}}(r_\beta) =$$
$$= \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \cdot \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} = \begin{bmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & -\cos\alpha\sin\beta + \sin\alpha\cos\beta \\ \sin\alpha\cos\beta + \cos\alpha\sin\beta & -\sin\alpha\sin\beta + \cos\alpha\cos\beta \end{bmatrix}$$

In questo modo otteniamo le formule di addizione per sin e cos.

3.7 Operazioni elementari

Il problema

Problema 3.7.1. Dati vettori $w_1, \ldots, w_n$ di uno spazio vettoriale W finitamente generato su un campo $\mathbb{K}$, trovare una base del sottospazio vettoriale di W generato da $w_1, \ldots, w_n$.

Per dare una risposta si sceglie una base $\mathcal{C} = \{w_1, \ldots, w_m\}$ di W, e si considera la matrice $A \in M_{m,n}(\mathbb{K})$ la cui colonna j è la colonna delle coordinate di w_j. Con una serie opportuna di cosiddette operazioni elementari si trasforma A in una matrice a scala per colonne S (vedi la Definizione 3.7.5): le colonne non nulle di S sono le coordinate di una base $\operatorname{Span}(w_1, \ldots, w_m)$, e quindi questo dà una risposta al problema formulato in 3.7.1. Il punto di questo algoritmo è che è efficiente. Lo stesso algoritmo si applica per dare risposta al seguente problema.

Problema 3.7.2. Data un'applicazione lineare $f\colon V \to W$ tra spazi vettoriali finitamente generati su $\mathbb{K}$, trovare una base di $\operatorname{im} f$.

Per risolvere tale problema si scelgono basi $\mathcal{B} = \{v_1, \ldots, v_n\}$ di V e $\mathcal{C} = \{w_1, \ldots, w_m\}$ di W e si considera la matrice $A = M_{\mathcal{C}}^{\mathcal{B}}(f) \in M_{m,n}(\mathbb{K})$. Per l'equazione (3.6.3), un vettore $w \in W$ appartiene a $\operatorname{im} f$ se e solo se il vettore colonna $X_{\mathcal{C}}(w)$ è nell'immagine dell'applicazione lineare $L_A\colon \mathbb{K}^n \to \mathbb{K}^m$. Quindi risolvere il Problema 3.7.2 equivale a risolvere il seguente

Problema 3.7.3. Data una matrice $A \in M_{m,n}(\mathbb{K})$, trovare una base di $\operatorname{im} L_A$, cioè del sottospazio di $\mathbb{K}^m$ generato dalle colonne di A.

La formulazione del problema in 3.7.2 appena data mostra che il problema in 3.7.1 è un caso particolare del problema in 3.7.2.

Definizione 3.7.4. Sia $f\colon V \to W$ un'applicazione lineare tra spazi vettoriali finitamente generati su $\mathbb{K}$. Il *rango* di f è la dimensione dell'immagine di f - lo denotiamo $\operatorname{rg}(f)$. Se $A \in M_{m,n}(\mathbb{K})$ il *rango* di A è la dimensione dell'immagine di L_A - lo denotiamo $\operatorname{rg}(A)$.

Quindi un algoritmo che risolve il Problema 3.7.2 dà in particolare un algoritmo per calcolare il rango di un'applicazione lineare (tra spazi finitamente generati).

Matrici a scala per colonne

Se la matrice A ha una forma particolare si risponde facilmente al Problema 3.7.3. Definiamo quali sono le matrici "particolari" in questione. Sia $A \in M_{m,n}(\mathbb{K})$. Per $1 \leqslant j \leqslant n$ definiamo

$$p_A(j) := \begin{cases} \min\{1 \leqslant i \leqslant m \mid a_{ij} \neq 0\} & \text{se } A_j \neq 0 \\ \infty & \text{se } A_j = 0. \end{cases}$$

Quindi $p_A(j)$ misura la "profondità" della colonna j, dove la profondità è determinata dall'entrata non nulla (della colonna) più vicina alla prima riga, e profondità 1 (la "superficie dell'acqua") corrisponde alla prima riga, profondità 2 corrisponde alla seconda riga, e così via.

Definizione 3.7.5. Una matrice $A \in M_{m,n}(\mathbb{K})$ A è a *scala per colonne* se $p_A(1) < p_A(2), \ldots < p_A(n)$ (per convenzione $\infty < \infty$), in altre parole se la profondità delle sue colonne è strettamente crescente (con la convenzione che passando dalla colonna nulla alla colonna nulla la profondità aumenta). Se A_j è una colonna non nulla di A, il *pivot* di A_j è l'entrata non nulla di A_j con indice di riga più piccolo (e quindi uguale a $p_A(j)$).

Esempio 3.7.6. Le seguenti matrici reali sono a scala per colonne

$$\begin{bmatrix} 2 & 0 \\ 0 & 5 \\ 3 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 0 & 0 \\ 3 & e & 0 \end{bmatrix}, \quad \begin{bmatrix} \pi & 0 \\ 0 & \sqrt{2} \end{bmatrix}. \tag{3.7.1}$$

La prima matrice ha pivot $2, 5$, la seconda ha pivot, cioè $1/3, e$, la terza ha pivot $\pi, \sqrt{2}$. Le seguenti matrici reali non sono a scala per colonne

$$\begin{bmatrix} 2 & -1 \\ \sqrt{5} & 5 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 3 \\ -1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} \pi & \sqrt{2} \\ 0 & 3 \\ -1 & 5 \end{bmatrix}. \tag{3.7.2}$$

Proposizione 3.7.7. *Sia $S \in M_{m,n}(\mathbb{K})$ una matrice a scala per colonne. La lista delle colonne non nulle di S è una base del sottospazio di $\mathbb{K}^m$ generato dalle colonne di S (cioè* $\mathrm{im}(L_S)$*).*

Dimostrazione. Siccome $\mathrm{im}(L_S)$ è generato dalle colonne non nulle di S, basta dimostrare che le colonne non nulle di S sono linearmente indipendenti. Lo dimostriamo per induzione su r. Se $r = 1$ (o $r = 0$) l'affermazione è ovvia. Per dimostrare il passo induttivo, siano $S_1, S_2, \ldots S_r$ (con $r \geqslant 2$) le colonne non nulle di S, e supponiamo che

$$\lambda_1 S_1 + \ldots + \lambda_r S_r = 0.$$

Se $i := p_S(1)$, l'entrata al posto i della combinazione lineare $\lambda_1 S_1 + \ldots + \lambda_r S_r$ è uguale a $\lambda_1 s_{i1}$ (perchè S è a scala per colonne), e siccome $s_{i1} \neq 0$ segue che $\lambda_1 = 0$. Quindi

$$\lambda_2 S_2 + \ldots + \lambda_r S_r = 0.$$

Siccome la matrice con colonne $S_2, \ldots, S_r$ è a scala per colonne, segue per l'ipotesi induttiva che $\lambda_2 = \ldots = \lambda_r = 0$. □

Operazioni elementari su liste di vettori e sulle colonne di una matrice

Descriveremo un procedimento che permetterà di ridurci sempre al caso di una matrice a scala (per righe o per colonne) quando vogliamo risolvere il Problema 3.7.3.

Definizione 3.7.8. Sia V uno spazio vettoriale su un campo $\mathbb{K}$ e $v_1, \ldots, v_n \in V$. Le *operazioni elementari* sulla lista $v_1, \ldots, v_n \in V$ sono le seguenti:

(1) Sostituire $v_1, \ldots, v_n$ con la lista ottenuta scambiando v_i con v_j e lasciando invariati gli altri vettori.

(2) Sostituire $v_1, \ldots, v_n$ con la lista ottenuta sostituendo v_i con $v_i + \lambda v_j$ dove $i \neq j$ e lasciando invariati gli altri vettori.

(3) Sostituire $v_1, \ldots, v_n$ con la lista ottenuta moltiplicando v_i per uno scalare **non nullo** e lasciando invariati gli altri vettori.

Lemma 3.7.9. *Sia V uno spazio vettoriale su un campo $\mathbb{K}$ e $v_1, \ldots, v_n \in V$. Sia $w_1, \ldots, w_n$ una lista di vettori di V ottenuta da $v_1, \ldots, v_n$ operando con (1), (2) o (3) della Definizione 3.7.8. Allora $v_1, \ldots, v_n$ è ottenuta da $w_1, \ldots, w_n$ operando rispettivamente con (1), (2) o (3) della Definizione 3.7.8.*

Dimostrazione. Se $w_1, \dots, w_n$ è ottenuta da $v_1, \dots, v_n$ scambiando v_i con v_j allora (ri)scambiando w_i con w_j otteniamo $v_1, \dots, v_n$. Ora supponiamo che $w_1, \dots, w_n$ sia ottenuta da $v_1, \dots, v_n$ operando con (2) della Definizione 3.7.8. Allora

$$v_i = (v_i + \lambda v_j) - \lambda v_j = w_i - \lambda w_j.$$

Siccome $v_h = w_h$ per $h \neq i$ segue che $v_1, \dots, v_n$ è ottenuta da $w_1, \dots, w_n$ operando con (2) della Definizione 3.7.8, dove λ è sostituito da $-\lambda$. Se $w_1, \dots, w_n$ è ottenuta da $v_1, \dots, v_n$ moltiplicando v_i per $0 \neq \lambda$ (e lasciando invariati gli altri vettori) allora $v_1, \dots, v_n$ è ottenuta da $w_1, \dots, w_n$ moltiplicando w_i per λ^{-1} e lasciando invariati gli altri vettori. □

Proposizione 3.7.10. *Sia V uno spazio vettoriale su un campo $\mathbb{K}$ e $v_1, \dots, v_n \in V$. Sia $w_1, \dots, w_n$ una lista di vettori di V ottenuta da $v_1, \dots, v_n$ operando con una delle operazioni della Definizione 3.7.8. Allora*

$$\langle v_1, \dots, v_n \rangle = \langle w_1, \dots, w_n \rangle. \tag{3.7.3}$$

Dimostrazione. L'operazione (1) scambia l'ordine dei vettori senza cambiare l'insieme dei vettori e quindi vale (3.7.3). Ora supponiamo che $w_1, \dots, w_n$ sia ottenuta da $v_1, \dots, v_n$ operando con (2) della Definizione 3.7.8. Siccome ogni w_h è combinazione lineare di $v_1, \dots, v_n$ abbiamo che $\langle w_1, \dots, w_n \rangle \subset \langle v_1, \dots, v_n \rangle$. D'altra parte per il Lemma 3.7.9 la lista $v_1, \dots, v_n$ è ottenuta da $w_1, \dots, w_n$ operando con (2) della Definizione 3.7.8: per quanto abbiamo appena dimostrato segue che $\langle v_1, \dots, v_n \rangle \subset \langle w_1, \dots, w_n \rangle$. Quindi vale (3.7.3). Se $w_1, \dots, w_n$ è ottenuta da $v_1, \dots, v_n$ operando con (3) della Definizione 3.7.8 è chiaro che vale (3.7.3). □

Il risultato seguente è una immediata consequenza della Proposizione 3.7.10.

Corollario 3.7.11. *Sia V uno spazio vettoriale su un campo $\mathbb{K}$ e $v_1, \dots, v_n \in V$. Se $w_1, \dots, w_n$ è una lista di vettori di V ottenuta da $v_1, \dots, v_n$ operando con una serie di operazioni elementari, allora*

$$\langle v_1, \dots, v_n \rangle = \langle w_1, \dots, w_n \rangle.$$

Sia $A \in M_{m,n}(\mathbb{K})$. Le colonne di A formano una lista di vettori di $\mathbb{K}^m$. Se operiamo sulle colonne di A con una delle operazioni elementari otteniamo altri n vettori di $\mathbb{K}^m$ che sono le colonne di un'altra matrice $m \times n$. Questa è una *operazione elementare sulle colonne* di A.

Proposizione 3.7.12. *Sia $A \in M_{m,n}(\mathbb{K})$. Esiste una serie di operazioni elementari di tipo (1) e di tipo (2) sulle colonne di A il cui risultato finale è una matrice a scala per colonne S. Si ha l'uguaglianza*

$$\langle A_1, \dots, A_n \rangle = \langle S_1, \dots, S_n \rangle.$$

In particolare una base di $\mathrm{im}(L_A)$ è data dalle colonne non nulle di S e il rango di A è uguale al numero di colonne non nulle di S.

Dimostrazione. Per induzione su n, cioè il numero di colonne di A. Se A è la matrice nulla $0_{m,n}$ allora è a scala e non c'è nulla da dimostrare. Supponiamo che A non sia nulla. Sia $1 \leqslant j_0 \leqslant n$ tale che $p_A(j_0) = \min\{p_A(1), p_A(2), \dots, p_A(n)\}$. Siccome $A \neq 0_{m,n}$ abbiamo che $p_A(j_0) < \infty$. Scambiando la prima colonna con la colonna j_0 (operazione elementare sulle colonne - di tipo (1)) passiamo ad una matrice A' tale che $p_{A'}(1) = \min\{p_{A'}(1), p_{A'}(2), \dots, p_{A'}(n)\}$. Abbiamo che $p_{A'}(1) \leqslant p_{A'}(j)$ per $1 \leqslant j \leqslant n$. Ripetendo tale operazione sulle colonne successive alla prima produciamo una matrice A'' tale che $p_{A''}(1) \leqslant p_{A''}(j)$ per $1 \leqslant j \leqslant n$ e $p_{A''}(2) \leqslant p_{A''}(j)$ per $2 \leqslant j \leqslant n$. Iterando questi scambi successivi arriviamo a una matrice B tale che

$$p_B(1) \leqslant p_B(2) \leqslant p_B(3) \leqslant \cdots \leqslant p_B(n-1) \leqslant p_B(n). \tag{3.7.4}$$

Ora distinguiamo due casi. Dapprima supponiamo che

$$p_B(1) < p_B(2). \tag{3.7.5}$$

Sia C la matrice $m\times(n-1)$ che ha come colonne le colonne di B eccetto la prima, cioè $C=[B_2,\dots,B_n]$. Per ipotesi induttiva esiste una serie di operazioni elementari sulle colonne che trasforma C in una matrice $m\times(n-1)$ a scala per colonne $T=[T_2,\dots,T_n]$. Tali operazioni non modificano le entrate (nulle) di C sulle righe di indice al più $p_B(1)$, e quindi anche la matrice $m\times n$ data da $S:=[B_1,T_2,\dots,T_n]$ è a scala per colonne. Questa matrice è ottenuta da A con una serie di operazioni elementari sulle colonne.

Ora supponiamo che esista $j_0>1$ tale che

$$p_B(1)=p_B(2)=\cdots=p_B(j_0)<p_B(j_0+1). \tag{3.7.6}$$

Sia $s=p_B(1)=p_B(2)=\ldots=p_B(j_0)$. Quindi $b_{sj}\neq 0$ per $1\leqslant j\leqslant j_0$. Sostituiamo alla colonna B_2 la colonna $B_2-b_{s2}b_{s1}^{-1}B_1$: questa è una operazione elementare sulle colonne (di tipo (2)). La matrice C che otteniamo ha la proprietà che $s=p_C(1)<p_C(2)$. Se $2<j_0$ sostituiamo alla colonna C_3 la colonna $C_3-b_{s3}b_{s1}^{-1}C_1$, e così via fino a modificare la colonna j_0. Il risultato è una matrice D, ottenuta da A con una serie di operazioni elementari sulle colonne, tale che

$$p_D(1)\leqslant p_D(2)\leqslant p_D(3)\leqslant\cdots\leqslant p_D(n-1)\leqslant p_D(n), \tag{3.7.7}$$

e in più $p_D(1)<p_D(2)$. Quindi siamo nel caso analizzato in precedenza, e abbiamo fatto. □

Esempio 3.7.13. Determiniamo una base del sottospazio di $\mathbb{Q}^4$ generato dalle colonne della matrice

$$A:=\begin{bmatrix}3&2&1&5\\1&0&2&11\\2&-3&-1&-1\\6&-1&2&15\end{bmatrix}$$

Potremmo applicare l'algoritmo "abbassando" direttamente l'altezza della seconda, terza e quarta colonna, ma prima scambiamo prima e terza colonna perchè così non avremo denominatori (divideremo per 1). Indicando con il simbolo $\rightsquigarrow$ una operazione elementare, abbiamo

$$\begin{bmatrix}3&2&1&5\\1&0&2&11\\2&-3&-1&-1\\6&-1&2&15\end{bmatrix}\rightsquigarrow\begin{bmatrix}1&2&3&5\\2&0&1&11\\-1&-3&2&-1\\2&-1&6&15\end{bmatrix}\rightsquigarrow\begin{bmatrix}1&0&3&5\\2&-4&1&11\\-1&-1&2&-1\\2&-5&6&15\end{bmatrix}\rightsquigarrow$$
$$\rightsquigarrow\begin{bmatrix}1&0&0&5\\2&-4&-5&11\\-1&-1&5&-1\\2&-5&0&15\end{bmatrix}\rightsquigarrow\begin{bmatrix}1&0&0&0\\2&-4&-5&1\\-1&-1&5&4\\2&-5&0&5\end{bmatrix}\rightsquigarrow\begin{bmatrix}1&0&0&0\\2&1&-5&-4\\-1&4&5&-1\\2&5&0&-5\end{bmatrix}\rightsquigarrow$$
$$\rightsquigarrow\begin{bmatrix}1&0&0&0\\2&1&0&-4\\-1&4&25&-1\\2&5&25&-5\end{bmatrix}\rightsquigarrow\begin{bmatrix}1&0&0&0\\2&1&0&0\\-1&4&25&15\\2&5&25&15\end{bmatrix}\rightsquigarrow\begin{bmatrix}1&0&0&0\\2&1&0&0\\-1&4&25&0\\2&5&25&0\end{bmatrix}.$$

La conclusione è che una base di $\mathrm{im}(L_A)$ è data da

$$\{(1,2,-1,2),\,(0,1,4,5),\,(0,0,1,1)\}.$$

(Abbiamo moltiplicato per 25^{-1} la terza colonna della matrice a scala.)

3.8 Il procedimento di eliminazione di Gauss

Il problema

I seguenti sono problemi analoghi al Problema 3.7.2.

Problema 3.8.1. Siano V, W spazi vettoriali su un campo $\mathbb{K}$, entrambi finitamente generati. Sia $f\colon V \to W$ un'applicazione lineare. Dare un algoritmo efficiente per trovare una base di $\ker f$.

Problema 3.8.2. Siano V, W spazi vettoriali su un campo $\mathbb{K}$, entrambi finitamente generati. Sia $f\colon V \to W$ un'applicazione lineare. Dato $w \in W$, dare un algoritmo efficiente per trovare tutti gli elementi di $f^{-1}(w)$.

Il primo passo in entrambi i casi consiste nello scegliere una base $\mathcal{B} = \{v_1, \dots, v_n\}$ di V, una base $\mathcal{C} = \{w_1, \dots, w_m\}$ di W e associare a f la matrice $A = M_{\mathcal{C}}^{\mathcal{B}}(f)$. Per l'equazione (3.6.3), abbiamo

$$\ker f = \{v \in V \mid A \cdot X_{\mathcal{B}}(v) = 0\}, \qquad f^{-1}(w) = \{v \in V \mid A \cdot X_{\mathcal{B}}(v) = X_{\mathcal{C}}(w)\}.$$

Quindi risolvere i Problemi 3.8.1 e 3.8.2 equivale a risolvere il seguente

Problema 3.8.3. Data una matrice $A \in M_{m,n}(\mathbb{K})$ e $B \in M_{m,1}(\mathbb{K})$, dare un algoritmo efficiente per trovare le soluzioni (in $\mathbb{K}^n$) del sistema di equazioni lineari $A \cdot X = B$.

Matrici a scala per righe

Se la matrice A ha una forma particolare si risponde facilmente al Problema 3.8.3. Definiamo quali sono le matrici "particolari" in questione. Sia $A \in M_{m,n}(\mathbb{K})$. Per $1 \leqslant i \leqslant m$ definiamo

$$d_A(i) := \begin{cases} \min\{1 \leqslant j \leqslant n \mid a_{ij} \neq 0\} & \text{se } A^i \neq 0 \\ \infty & \text{se } A^i = 0. \end{cases}$$

Pensiamo a $d_A(i)$ come la distanza (aumentata di 1) della riga i dalla prima colonna, dove la distanza è determinata dalla colonna a cui appartiene l'entrata non nulla (della riga) con indice di colonna più piccola.

Definizione 3.8.4. Una matrice $A \in M_{m,n}(\mathbb{K})$ è a *scala per righe* se $d_A(1) < d_A(2), \dots < d_A(n)$ (per convenzione $\infty < \infty$). Se A^i è una riga non nulla, il *pivot* di A^i è l'entrata non nulla di A^i con indice di colonna più piccolo (e quindi uguale a $d_A(i)$).

Esempio 3.8.5. Le seguenti matrici reali sono a scala per righe

$$\begin{bmatrix} 2 & -1 \\ 0 & 5 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & -4 & 5 & 1 \\ 0 & 0 & 2 & 3 \end{bmatrix}, \quad \begin{bmatrix} \pi & 0 \\ 0 & \sqrt{2} \end{bmatrix}. \tag{3.8.1}$$

I pivot della prima matrice sono $2, 5$, quelli della seconda sono $1, 2$, quelli della terza sono $\pi, \sqrt{2}$. Le matrici reali

$$\begin{bmatrix} 0 & 2 \\ 3 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 & 3 \\ 0 & 2 & 1 \\ 0 & 3 & -1 \end{bmatrix} \tag{3.8.2}$$

non sono a scala per righe.

Supponiamo che $S \in M_{m,n}(\mathbb{K})$ sia a scala per righe, e di voler trovare tutte le soluzioni $X \in \mathbb{K}^n$ del sistema di equazioni lineari omogenee

$$S \cdot X = B. \tag{3.8.3}$$

Siano r le righe non nulle di S e siano

$$j_1 := d_S(1), \dots, j_r := d_S(r)$$

i corrispondenti valori di d_S, cioè gli indici di colonna dei pivot di S. Innanzitutto vediamo che se esiste un i tale che $r < i \leqslant n$ e $b_i \neq 0$, allora il sistema non ha soluzioni perchè compare l'equazione $0 = b_i$. Viceversa, risolvendo le equazioni "dal basso verso l'alto" si vede che esiste una soluzione di (3.8.3) che assume valori arbitrari di x_j per $j \in (\{1, \dots, n\} \backslash \{j_1, \dots, j_r\})$ e che, una volta assegnati questi valori, il valore di $x_{j_1}, \dots, x_{j_r}$ è univocamente determinato.

Esempio 3.8.6. Risolviamo in $\mathbb{Q}^3$ il sistema di equazioni lineari

$$\begin{aligned} x_1 - 4x_2 + 5x_3 &= b_1 \\ 2x_2 + 3x_3 &= b_2. \end{aligned}$$

L'ultima equazione dà $x_2 = -\frac{3}{2}x_3 - \frac{1}{2}b_2$, e sostituendo nella prima equazione, otteniamo $x_1 = -11x_3 + b_1 - 2b_2$. Quindi l'insieme delle soluzioni è

$$\{(-11x_3 + b_1 - 2b_2, -\frac{3}{2}x_3 - \frac{1}{2}b_2, x_3) \mid x_3 \in \mathbb{Q}\} = \{(22t + b_1 - 2b_2, 3t - \frac{1}{2}b_2, -2t) \mid t \in \mathbb{Q}\}.$$

Ricordiamo che l'insieme delle soluzioni di $A \cdot X = B$ è un sottospazio vettoriale di $\mathbb{K}^n$ se e solo se $B = 0$.

Esempio 3.8.7. Ponendo $b_1 = b_2 = 0$ nell'Esempio 3.8.6 vediamo che il sottospazio delle soluzioni ha come base $\{(22, 3, -2)\}$, e quindi ha dimensione 1.

Proposizione 3.8.8. *Sia $S \in M_{m,n}(\mathbb{K})$ a scala per righe, e siano $1 \leqslant i_1 < i_2 < \ldots < i_r \leqslant m$ gli indici delle righe non nulle di S. Allora l'applicazione lineare*

$$\begin{array}{ccc} \ker(L_S) & \xrightarrow{\varphi_S} & \mathbb{K}^{n-r} \\ X & \mapsto & (x_1, \ldots, x_{j_1-1}, x_{j_1+1}, \ldots, x_{j_r-1}, x_{j_r+1}, \ldots, x_n) \end{array} \tag{3.8.4}$$

(la φ dimentica le entrate di indici $i_1 < i_2 < \ldots < i_r$) è un isomorfismo di spazi vettoriali. In particolare

$$\dim\{X \in \mathbb{K}^n \mid S \cdot X = 0\} = n - |\{1 \leqslant i \leqslant m \mid S^i \neq 0\}|. \tag{3.8.5}$$

Dimostrazione. Per induzione su r. Se $r = 0$ il risultato è ovvio perchè in questo caso $\ker(L_S) = \mathbb{K}^n$. Dimostriamo il passo induttivo. Quindi supponiamo che $r > 0$, e che il risultato vale se sostituiamo r con $(r-1)$. Sia $T \in M_{m,n}(\mathbb{K})$ la matrice ottenuta da S sostituendo alla riga S^r la riga nulla. Per ipotesi induttiva l'applicazione lineare

$$\begin{array}{ccc} \ker(L_T) & \xrightarrow{\varphi_T} & \mathbb{K}^{n-r} \\ X & \mapsto & (x_1, \ldots, x_{j_1-1}, x_{j_1+1}, \ldots, x_{j_{r-1}-1}, x_{j_{r-1}+1}, \ldots, x_n) \end{array} \tag{3.8.6}$$

è un isomorfismo. Ovviamente

$$\ker(L_S) = \ker(L_T) \cap \{X \in \mathbb{K}^n \mid S^r \cdot X = 0\}.$$

Ora

$$\{X \in \mathbb{K}^n \mid S^r \cdot X = 0\} = \{X \in \mathbb{K}^n \mid x_{j_r} = -s_{r,j_r}^{-1}(s_{r,j_r+1}x_{j_{r+1}} + \ldots + s_{r,n}x_n)\}.$$

Segue che la restrizione di φ_T a $\ker(L_S)$ ha immagine il codominio di φ_S. Siccome φ_T è iniettiva per ipotesi induttiva, segue che φ_S è un isomorfismo. □

Operazioni elementari sulle righe di una matrice

Sia $A \in M_{m,n}(\mathbb{K})$. Le righe di A formano una lista di vettori di $\mathbb{K}^n$. Se operiamo sulle righe di A con (1), (2) o (3) della Definizione 3.7.8 otteniamo altri m vettori di $\mathbb{K}^n$ che sono le righe di un'altra matrice $m \times n$. Questa è una *operazione elementare sulle righe* di A. Osserviamo che una operazione elementare sulle righe di A corrisponde a una operazione elementare sulle colonne di un'altra matrice che si associa ad A, la sua trasposta.

Definizione 3.8.9. Sia $A \in M_{m,n}(\mathbb{K})$. La *trasposta* di A è la matrice $A^t \in M_{n,m}(\mathbb{K})$ le cui righe sono le colonne di A. Più precisamente poniamo $A = (a_{ij})$ e $A^t = (b_{ij})$. Allora $b_{ij} = a_{ji}$.

Qualche esempio di matrice e la sua trasposta:

$$A := \begin{bmatrix} 2 & 1 & 3 \\ -1 & 0 & 1 \end{bmatrix}, \quad A^t := \begin{bmatrix} 2 & -1 \\ 1 & 0 \\ 3 & 1 \end{bmatrix}, \quad B := \begin{bmatrix} 5 & 2 \\ -1 & 3 \end{bmatrix}, \quad B^t := \begin{bmatrix} 5 & -1 \\ 2 & 3 \end{bmatrix}.$$

Osservazione 3.8.10. Siano $A, B \in M_{m,n}(\mathbb{K})$, e $C \in M_{n,p}(\mathbb{K})$. Un calcolo dà le seguenti uguaglianze:

$$(A+B)^t = A^t + B^t, \qquad (B \cdot C)^t = C^t \cdot B^t.$$

Osservazione 3.8.11. Sia $A \in M_{m,n}(\mathbb{K})$, e sia $A^t \in M_{n,m}(\mathbb{K})$ la sua trasposta. Siccome le colonne di A^t sono le righe di A, è chiaro che $B \in M_{m,n}(\mathbb{K})$ è ottenuta da A con una operazione elementare sulle righe se e solo se B^t è ottenuta da A^t con una operazione elementare sulle colonne.

Per l'Osservazione appena fatta, il seguente risultato segue dalla prima parte della Proposizione 3.7.12.

Proposizione 3.8.12. *Sia $A \in M_{m,n}(\mathbb{K})$. Esiste una serie di operazioni elementari sulle righe di A il cui risultato finale è una matrice a scala per righe.*

Eliminazione di Gauss per sistemi di equazioni lineari omogenee

Sia $A \in M_{m,n}(\mathbb{K})$. Per la Proposizione 3.8.12 esiste una serie di operazioni elementari sulle righe di A il cui risultato finale è una matrice a scala per righe S. Questo fatto, unito al risultato seguente dà un algoritmo efficiente per risolvere il sistema di equazioni omogenee $A \cdot X = 0$ nelle incognite $x_1, \ldots, x_n$.

Proposizione 3.8.13. *Sia $A \in M_{m,n}(\mathbb{K})$, e sia S una matrice a scala per righe ottenuta da A con una serie di operazioni elementari sulle righe di A. Si ha l'uguaglianza*

$$\{X \in \mathbb{K}^n \mid A \cdot X = 0\} = \{X \in \mathbb{K}^n \mid S \cdot X = 0\}. \tag{3.8.7}$$

Dimostrazione. È sufficiente dimostrare che, se M è ottenuta da A con una operazione elementare sulle righe, allora

$$\{X \in \mathbb{K}^n \mid A \cdot X = 0\} = \{X \in \mathbb{K}^n \mid M \cdot X = 0\}. \tag{3.8.8}$$

L'uguaglianza è ovvia se l'operazione è di tipo (1) o (3). Ora supponiamo che M sia ottenuta da A con una operazione elementare sulle righe di tipo (2). Quindi $i \neq j \in \{1, \ldots, n\}$, $\lambda \in \mathbb{K}$ e la riga i-esima di M è $A^i + \lambda A^j$. Dimostriamo che il membro di sinistra di (3.8.8) è contenuto nel membro di destra di (3.8.8). Sia X nel membro di sinistra. Le equazioni che definiscono il membro di destra sono le stesse equazioni che definiscono il membro di sinistra eccetto quella sulla riga i che è

$$(A^i + \lambda A^j) \cdot X = 0. \tag{3.8.9}$$

Siccome X appartiene al membro di sinistra abbiamo che $A^i \cdot X = 0$ e $A^j \cdot X = 0$; segue che vale (3.8.9). Questo dimostra che il membro di sinistra di (3.8.8) è contenuto nel membro di destra di (3.8.8).

Rimane da dimostrare che il membro di destra è contenuto nel membro di sinistra. Le equazioni che definiscono il membro di sinistra sono le stesse equazioni che definiscono il membro di destra eccetto quella sulla riga i che è

$$(M^i - \lambda M^j) \cdot X = 0. \tag{3.8.10}$$

Il ragionamento appena fatto dimostra che il membro di destra di (3.8.8) è contenuto nel membro di sinistra di (3.8.8). □

L'algoritmo appena descritto si chiama *procedimento di eliminazione di Gauss*.

Corollario 3.8.14. *Sia $A \in M_{m,n}(\mathbb{K})$, e sia S una matrice a scala per righe ottenuta da A con una serie di operazioni elementari sulle righe di A. Allora*

$$\dim\{X \in \mathbb{K}^n \mid A \cdot X = 0\} = n - |\{1 \leqslant i \leqslant m \mid S^i \neq 0\}|.$$

Dimostrazione. Per la Proposizione 3.8.13 la dimensione dello spazio delle soluzioni di $A \cdot X = 0$ è uguale alla dimensione dello spazio delle soluzioni di $S \cdot X = 0$, e quest'ultima dimensione è data da (3.8.5). □

Proposizione 3.8.15. *Sia $\mathbb{K}$ un campo e $A \in M_{m,n}(\mathbb{K})$. Il rango di A^t è uguale al rango di A.*

Dimostrazione. Sia S una matrice a scala per righe ottenuta da A con operazione elementari sulle righe. Per l'Osservazione 3.8.11 abbiamo

$$\mathrm{rg}(A^t) = |\{1 \leqslant i \leqslant m \mid S^i \neq 0\}|.$$

Per il Corollario 3.8.14 otteniamo che

$$\dim\{X \in \mathbb{K}^n \mid A \cdot X = 0\} = n - |\{1 \leqslant i \leqslant m \mid S^i \neq 0\}| = n - \mathrm{rg}(A^t).$$

Ma d'altra parte

$$\dim\{X \in \mathbb{K}^n \mid A \cdot X = 0\} = n - \mathrm{rg}(A)$$

per la Proposizione 3.1.21. Le ultime due equazioni danno che $\mathrm{rg}(A) = \mathrm{rg}(A^t)$. □

Osservazione 3.8.16. La Proposizione 3.8.15 equivale alla seguente affermazione: se $A \in M_{m,n}(\mathbb{K})$ allora il sottospazio di $\mathbb{K}^n$ generato dalle *righe* di A ha la stessa dimensione del sottospazio di $\mathbb{K}^m$ generato dalle *colonne* di A. Notiamo che l'affermazione non è affatto banale. Equivalentemente: il massimo numero di righe linearmente indipendenti di A è uguale al massimo numero di colonne linearmente indipendenti di A. Possiamo anche dare la seguente versione della Proposizione 3.8.15: se, con una serie di operazioni elementari sulle *righe*, riduciamo A a una matrice a scala per righe S e, con una serie di operazioni elementari sulle *colonne*, riduciamo A a una matrice a scala per colonne T allora il numero di righe non nulle di S è uguale al numero di colonne non nulle di T (infatti il primo numero è uguale al rango di A^t, il secondo è uguale al rango di A).

Eliminazione di Gauss e dimensione

Il procedimento di eliminazione di Gauss dimostra il seguente risultato *indipendentemente dalla Proposizione 2.6.13 e dal Corollario 2.6.14.*

Proposizione 3.8.17. *Se $A \in M_{m,n}(\mathbb{K})$ e $m < n$, allora esiste una soluzione non nulla $X \in \mathbb{K}^n$ dell'equazione $A \cdot X = 0$, cioè un sistema di equazioni lineari omogenee con più incognite che equazioni ha almeno una soluzione non banale.*

A partire da questo risultato possiamo *ri*dimostrare il Corollario 2.6.14, ragionando come segue. Sia V uno spazio vettoriale su $\mathbb{K}$ finitamente generato, e siano $\mathcal{B} = \{v_1, \ldots, v_m\}$, $\mathcal{C} = \{w_1, \ldots, w_n\}$ basi di V. Supponiamo che $m < n$ e dimostriamo che i vettori $w_1, \ldots, w_n$ sono linearmente dipendenti - ovviamente questo dimostra il Corollario 2.6.14. Per $j \in \{1, \ldots, n\}$ esistono $a_{1j}, \ldots, a_{mj} \in \mathbb{K}$ tali che

$$w_j = \sum_{i=1}^{m} a_{ij} v_i.$$

Sia $A \in M_{m,n}(\mathbb{K})$ la matrice $A = (a_{ij})$. Per la Proposizione 3.8.17 esiste una soluzione non banale $\overline{X}$ dell'equazione $A \cdot X = 0$. Si ha

$$\sum_{j=1}^{n} \overline{x}_j w_j = \sum_{j=1}^{n} \overline{x}_j \left(\sum_{i=1}^{m} a_{ij} v_i \right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \overline{x}_j \right) v_i = 0,$$

e quindi i vettori $w_1, \ldots, w_n$ sono linearmente dipendenti.

Eliminazione di Gauss per sistemi di equazioni lineari arbitrarie

Il procedimento di eliminazione di Gauss si può impiegare anche per risolvere un sistema di equazione lineari arbitrario

$$A \cdot X = B \tag{3.8.11}$$

dove $A \in M_{m,n}(\mathbb{K})$, $B \in M_{m,1}(\mathbb{K})$ e X è una matrice colonna (di incognite) $n \times 1$. Si procede come segue. Sia $A = (a_{ij})$ e $B = (b_i)$. Consideriamo la matrice $m \times (n+1)$

$$[A|B] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} & b_i \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix} \tag{3.8.12}$$

Sappiamo che, con una serie di operazioni elementari sulle righe di $[A|B]$, arriviamo a una matrice a scala per righe $[S|C]$. Notiamo che anche S è una matrice a scala per righe.

Proposizione 3.8.18. *Siano $[A|B]$ e $[S|C]$ come sopra. Allora abbiamo l'uguaglianza*

$$\{X \in \mathbb{K}^n \mid A \cdot X = B\} = \{X \in \mathbb{K}^n \mid S \cdot X = C\}. \tag{3.8.13}$$

Dimostrazione. Si ragiona come nella dimostrazione della Proposizione 3.8.13. È sufficiente dimostrare che, se $[M|R]$ è ottenuta da $[A|B]$ con una operazione elementare sulle righe, allora

$$\{X \in \mathbb{K}^n \mid A \cdot X = B\} = \{X \in \mathbb{K}^n \mid M \cdot X = R\}. \tag{3.8.14}$$

Il risultato è ovvio se l'operazione è di tipo (1) o (3). Ora supponiamo l'operazione sia di tipo (2). Quindi $i \neq j \in \{1, \dots, n\}$, $\lambda \in \mathbb{K}$ e $[M|R]$ è ottenuta da $[A|B]$ sostituendo la riga i-esima (A^i, b_i) con la riga $(A^i \lambda A^j, b_i + \lambda b_j)$. Dimostriamo che il membro di sinistra di (3.8.8) è contenuto nel membro di destra di (3.8.8). Sia X nel membro di sinistra. Le equazioni che definiscono il membro di destra sono le stesse equazioni che definiscono il membro di sinistra eccetto quella sulla riga i che è

$$(A^i + \lambda A^j) \cdot X = b_i + \lambda b_j. \tag{3.8.15}$$

Siccome X appartiene al membro di sinistra abbiamo che $A^i \cdot X = b_i$ e $A^j \cdot X = b_j$; segue che vale (3.8.9). Questo dimostra che il membro di sinistra di (3.8.14) è contenuto nel membro di destra di (3.8.14).

Rimane da dimostrare che il membro di destra è contenuto nel membro di sinistra. Le equazioni che definiscono il membro di sinistra sono le stesse equazioni che definiscono il membro di destra eccetto quella sulla riga i che è

$$(M^i - \lambda M^j) \cdot X = r_i - \lambda r_j. \tag{3.8.16}$$

Il ragionamento appena fatto dimostra che il membro di destra di (3.8.14) è contenuto nel membro di sinistra di (3.8.14). □

La conclusione è che risolviamo il sistema di equazioni in (3.8.11) procedendo come segue: con una serie di operazioni elementari sulle righe "trasformiamo" $[A|B]$ in una matrice a scala per righe $[S|C]$. Per la Proposizione 3.8.18 il sistema di equazioni in (3.8.11) ha le stesse soluzioni del sistema di equazioni $S \cdot X = C$, e questo sistema si risolve "dal basso verso l'alto". Notiamo che il sistema $S \cdot X = C$ ha soluzioni (cioè il sistema $A \cdot X = B$ ha soluzioni) se e solo se per ogni $i \in \{1, \dots, m\}$ tale che $S^i = 0$ la corrispondente entrata c_i di C è nulla.

3.9 Calcolo dell'inversa di una matrice invertibile

Problema 3.9.1. Data una matrice $A \in \mathrm{GL}_n(\mathbb{K})$ (cioè $A \in M_{n,n}(\mathbb{K})$ ed è invertibile) come possiamo calcolare in modo efficiente l'inversa di A?

Descriveremo un algoritmo che produce A^{-1}. Iniziamo con il seguente risultato.

Lemma 3.9.2. *Sia $A \in \mathrm{GL}_n(\mathbb{K})$. Esiste una serie di operazioni elementari sulle righe della matrice $[A|1_n]$ (notate che è una matrice $n \times 2n$) che produce una matrice $[1_n|D]$, dove $D \in M_{n,n}(\mathbb{K})$.*

Dimostrazione. Consideriamo la matrice $n \times 2n$ data da $[A|1_n]$. Con una serie di operazioni elementari sulle righe di $[A|1_n]$ possiamo arrivare a una matrice a scala per righe $[S|C]$ dove S, C sono matrici $n \times n$. Notiamo che anche S è a scala per righe, ed è ottenuta da A con una serie di operazioni elementari sulle righe. Siccome A è invertibile $\ker(L_A) = \{0\}$, e quindi per il Corollario 3.8.14 segue che le righe di S sono tutte diverse da 0. Siccome S è quadrata questo significa che tutte le entrate di S sulla diagonale principale di S sono non nulle. Quindi moltiplicando la riga i-esima di $[S|C]$ per s_{ii}^{-1} arriviamo a $[S'|C']$ dove S' è $n \times n$ a scala per righe con entrate sulla diagonale principale uguali a 1. Ora è chiaro che con una serie di opportune operazioni elementari sulle righe di $[S'|C']$ possiamo arrivare a una matrice $[1_n|D]$. □

Proposizione 3.9.3. *Sia $A \in M_{n,n}(\mathbb{K})$ invertibile, e sia $D \in M_{n,n}(\mathbb{K})$ la matrice ottenuta a partire da A con il procedimento del Lemma 3.9.2. Allora $D = A^{-1}$.*

La Proposizione 3.9.3 dà un algoritmo che risponde al Problema 3.9.1. Prima di dimostrare la proposizione diamo un esempio di calcolo dell'inversa.

Esempio 3.9.4. Sia

$$A := \begin{bmatrix} 2 & 1 & 3 \\ -1 & 0 & 1 \\ 3 & 2 & 8 \end{bmatrix}$$

Calcoliamo A^{-1} seguendo l'algoritmo appena descritto. Dunque partiamo dalla matrice 3×6

$$\left[\begin{array}{ccc|ccc} 2 & 1 & 3 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 1 & 0 \\ 3 & 2 & 8 & 0 & 0 & 1 \end{array}\right]$$

e operiamo sulle righe in modo da trasformare la matrice a sinistra dei tratti verticali in una matrice a scala per righe. Come prima operazione moltiplichiamo la seconda riga per (-1) e poi scambiamo tra di loro le prime due righe: otteniamo

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & -1 & 0 \\ 2 & 1 & 3 & 1 & 0 & 0 \\ 3 & 2 & 8 & 0 & 0 & 1 \end{array}\right]$$

Ora moltiplichiamo la prima riga per (-2) e aggiungiamola alla seconda riga, poi moltiplichiamo la prima riga per (-3) e aggiungiamola alla terza riga: otteniamo così

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & -1 & 0 \\ 0 & 1 & 5 & 1 & 2 & 0 \\ 0 & 2 & 11 & 0 & 3 & 1 \end{array}\right]$$

Moltiplicando la seconda riga per (-2) e aggiungendola alla terza riga otteniamo

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & -1 & 0 \\ 0 & 1 & 5 & 1 & 2 & 0 \\ 0 & 0 & 1 & -2 & -1 & 1 \end{array}\right]$$

Ora la matrice a sinistra dei tratti verticali è a scala per righe, e in questo esempio le entrate sulla diagonale principale sono già uguali a 1. Rimane da operare sulle righe "dal basso" per trasformare la matrice a sinistra dei tratti verticali nella matrice 1_3. Moltiplichiamo la terza riga per (-2) e aggiungiamola alla terza riga, poi aggiungiamo la terza riga alla prima: otteniamo

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & -2 & 1 \\ 0 & 1 & 0 & 11 & 7 & -5 \\ 0 & 0 & 1 & -2 & -1 & 1 \end{array}\right]$$

Quindi

$$A^{-1} = \left[\begin{array}{ccc} -2 & -2 & 1 \\ 11 & 7 & -5 \\ -2 & -1 & 1 \end{array}\right].$$

(Provare per credere!)

Per dimostrare la Proposizione 3.9.3 dimostriamo il seguente risultato.

Lemma 3.9.5. *Siano $A, B \in M_{m,n}(\mathbb{K})$, e supponiamo che $B \in M_{m,n}(\mathbb{K})$ sia ottenuta da A con operazioni elementari sulle righe. Allora esiste $L \in \mathrm{GL}_m(\mathbb{K})$ (cioè una matrice $L \in M_{m,m}(\mathbb{K})$ invertibile) tale che $B = L \cdot A$.*

Dimostrazione. Dimostriamo che, se $B \in M_{m,n}(\mathbb{K})$ è ottenuta da A con una operazioni elementare sulle righe di tipo (1), (2) o (3), allora esiste $L \in \mathrm{GL}_m(\mathbb{K})$ tale che

$$B = L \cdot A. \tag{3.9.1}$$

Supponiamo che l'operazione elementare sia di tipo (1). Quindi esistono $k \neq h \in \{1, \ldots, m\}$ tali che $B^k = A^h$, $B^h = A^k$ e per $i \in (\{1, \ldots, m\} \backslash \{k, h\})$ si ha $A^i = B^i$. Sia $L = (l_{ij})$ data da

$$l_{ij} := \begin{cases} \delta_{ij} & \text{se } (i,j) \notin \{(k,k),(k,h),(h,k),(h,h)\}, \\ 0 & \text{se } (i,j) \in \{(k,k),(h,h)\}, \\ 1 & \text{se } (i,j) \in \{(k,h),(h,k)\}. \end{cases} \tag{3.9.2}$$

Si verifica facilmente che vale (3.9.1). Notate che $L^2 = 1_m$ e quindi L è invertibile.

Supponiamo che l'operazione elementare sia di tipo (2). Quindi esistono $k \neq h \in \{1, \ldots, m\}$ e $\lambda \in \mathbb{K}$ tali che $B^k = A^k + \lambda A^h$, e per $i \in (\{1, \ldots, m\} \backslash \{k\})$ si ha $A^i = B^i$. Sia $L = (l_{ij})$ data da

$$l_{ij} := \begin{cases} \delta_{ij} & \text{se } (i,j) \neq (k,h), \\ \lambda & \text{se } (i,j) = (k,h). \end{cases} \tag{3.9.3}$$

Si verifica facilmente che vale (3.9.1). Notate che L è invertibile perchè con una operazione elementare sulle colonne "diventa" la matrice unità.

Se l'operazione elementare è di tipo (3), cioè esistono $k \in \{1, \ldots, m\}$ e $0 \neq \mu \in \mathbb{K}$ tali che $B^k = \mu A^k$ e $A^i = B^i$ per $i \in (\{1, \ldots, m\} \backslash \{k\})$, allora vale (3.9.1) con $L = (l_{ij})$ data da

$$l_{ij} := \begin{cases} \delta_{ij} & \text{se } (i,j) \neq (k,k), \\ \mu & \text{se } (i,j) = (k,k). \end{cases} \tag{3.9.4}$$

La matrice L è invertibile perchè è diagonale con tutte le entrate sulla diagonale principale non nulle.

Ora dimostriamo il risultato in generale. Per quello che abbiamo appena dimostrato esistono $L_1, \ldots, L_r \in \mathrm{GL}_m(\mathbb{K})$ tali che $B = L_1 \cdot \ldots \cdot L_r \cdot A$. Siccome $L := L_1 \cdot \ldots \cdot L_r$ è in $\mathrm{GL}_m(\mathbb{K})$ e $B = L \cdot A$, abbiamo fatto. □

Dimostrazione della Proposizione 3.9.3. Per il Lemma 3.9.5 esiste $L \in \mathrm{GL}_n(\mathbb{K})$ tale che

$$[1_n | D] = L \cdot [A | 1_n] = [L \cdot A | L \cdot 1_n] = [L \cdot A | L]\,.$$

Quindi $L \cdot A = 1_n$, ovvero $L = A^{-1}$. Ma d'altra parte $L = D$, e perciò $D = A^{-1}$. □

3.10 Cambiamenti di base

Sia V uno spazio vettoriale su $\mathbb{K}$, finitamente generato e di dimensione n. Siano

$$\mathcal{B} = \{u_1, \dots, u_n\}, \qquad \mathcal{C} = \{w_1, \dots, w_n\}$$

basi di V. Per l'Equazione (3.6.3) vale

$$X_{\mathcal{C}}(v) = M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V) \cdot X_{\mathcal{B}}(v) \qquad \forall v \in V. \tag{3.10.1}$$

Definizione 3.10.1. $M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V)$ è la matrice del *cambiamento di base* da $\mathcal{B}$ a $\mathcal{C}$.

Quindi la matrice del cambiamento di base da $\mathcal{B}$ a $\mathcal{C}$ ci dà le $\mathcal{C}$-coordinate di un vettore a partire dalle sue $\mathcal{B}$-coordinate per mezzo della Formula (3.10.1).

Notiamo che la matrice del cambiamento di base è invertibile perchè per l'Equazione (3.6.6) si ha

$$M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}_V) \cdot M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V) = M_{\mathcal{B}}^{\mathcal{B}}(\mathrm{Id}_V) = 1_n, \qquad M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V) \cdot M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}_V) = M_{\mathcal{C}}^{\mathcal{C}}(\mathrm{Id}_V) = 1_n.$$

(Abbiamo posto $n := \dim V$.) Quindi

$$M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V)^{-1} = M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}_V). \tag{3.10.2}$$

Osservazione 3.10.2. Siano $V = \mathbb{K}^n$ e $\mathcal{S} = \{e_1, \dots, e_n\}$ la base standard. Sia $\mathcal{C}$ la base di $\mathbb{K}^n$ data da $\mathcal{C} = \{C_1, \dots, C_n\}$ dove le C_j sono matrici $n \times 1$ (matrici colonna). Allora

$$M_{\mathcal{S}}^{\mathcal{C}}(\mathrm{Id}_V) = [C_1, \dots, C_n], \qquad M_{\mathcal{C}}^{\mathcal{S}}(\mathrm{Id}_V) = [C_1, \dots, C_n]^{-1}.$$

Infatti la prima equazione vale per definizione di $M_{\mathcal{S}}^{\mathcal{C}}(\mathrm{Id}_V)$, e la seconda vale per (3.10.2).

Osservazione 3.10.3. Siano V uno spazio vettoriale finitamente generato su $\mathbb{K}$ e $\mathcal{B}, \mathcal{C}, \mathcal{D}$ sue basi. Per l'Equazione (3.6.6) si ha

$$M_{\mathcal{D}}^{\mathcal{B}}(\mathrm{Id}_V) = M_{\mathcal{D}}^{\mathcal{C}}(\mathrm{Id}_V) \cdot M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V). \tag{3.10.3}$$

Quindi possiamo esprimere la matrice del cambiamento di base tra basi arbitrarie come prodotto di matrici di cambiamento di base da una base arbitraria a una base fissata. (Analogia: se in una citta', per esempio Roma, sappiamo andare da un punto arbitrario a un punto fissato, per esempio Piazza Navona, allora sappiamo andare da un punto arbitrario a un altro punto arbitrario, per esempio passando da Piazza Navona.)

Esempio 3.10.4. Siano $\mathcal{B} = \{B_1, \dots, B_n\}$ e $\mathcal{D} = \{D_1, \dots, D_n\}$ basi di $\mathbb{K}^n$. (Come di consueto B_j e D_j sono matrici $n \times 1$.) Per le Osservazioni 3.10.2 e 3.10.3 abbiamo che

$$M_{\mathcal{D}}^{\mathcal{B}}(\mathrm{Id}_V) = M_{\mathcal{D}}^{\mathcal{S}}(\mathrm{Id}_V) \cdot M_{\mathcal{S}}^{\mathcal{B}}(\mathrm{Id}_V) = [D_1, \dots, D_n]^{-1} \cdot [B_1, \dots, B_n]. \tag{3.10.4}$$

Per illustrare questo metodo scegliamo le basi di $\mathbb{Q}^3$ date da $\mathcal{B} = \{(2,1,0), (1,0,-1), (0,-1,2)\}$ e $\mathcal{D} = \{(2,-1,3), (1,0,2), (3,1,8)\}$ (verificate che sono basi!). Allora (3.10.4) diventa

$$M_{\mathcal{D}}^{\mathcal{B}}(\mathrm{Id}_{\mathbb{K}^3}) = \begin{bmatrix} 2 & 1 & 3 \\ -1 & 0 & 1 \\ 3 & 2 & 8 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 2 \end{bmatrix}. \tag{3.10.5}$$

L'inversa che appare in (3.10.5) è stata calcolata nell'Esempio 3.9.4, e otteniamo che

$$M_{\mathcal{D}}^{\mathcal{B}}(\mathrm{Id}_{\mathbb{K}^3}) = \begin{bmatrix} -6 & -3 & 4 \\ 29 & 16 & -17 \\ -5 & -3 & 3 \end{bmatrix}.$$

Quindi, per esempio, le coordinate di $(2,1,0)$ nella base $\mathcal{D}$ sono $(-6, 29, -5)$, ovvero

$$(2,1,0) = -6(2,-1,3) + 29(1,0,2) - 5(3,1,8).$$

Abbiamo visto che una matrice di cambiamento di base è invertibile. Vale il viceversa, cioè ogni matrice invertibile è la matrice di un cambiamento di base.

Proposizione 3.10.5. *Sia V uno spazio vettoriale finitamente generato e di dimensione n su un campo $\mathbb{K}$. Sia $\mathcal{C} = \{u_1, \dots, u_n\}$ una base di V e $A \in \mathrm{GL}_n(\mathbb{K})$. Esiste una (e una sola) base $\mathcal{B}$ di V tale che $M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}_V) = A$.*

Dimostrazione. Sia $w_j \in V$ il vettore con vettore delle coordinate uguale alla j-esima colonna di A^{-1}. Esplicitamente: se $A^{-1} = (e_{ij})$ abbiamo che

$$w_j = \sum_{i=1}^{n} e_{ij} u_i.$$

Siccome le colonne di A^{-1} sono linearmente indipendenti $\mathcal{B} := \{w_1, \dots, w_n\}$ è una base di V. Abbiamo che

$$M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V) = A^{-1}.$$

Per l'equazione (3.10.2) segue che $M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}_V) = A$. □

3.11 Endomorfismi e coniugio

Definizione 3.11.1. Sia V uno spazio vettoriale. Un *endomorfismo* di V è un'applicazione lineare $f\colon V \to V$, cioè un elemento di $\mathcal{L}(V, V)$.

Poniamo

$$\mathrm{End}(V) := \mathcal{L}(V, V).$$

Sia V uno spazio vettoriale su $\mathbb{K}$, finitamente generato e di dimensione n. Sia $f\colon V \to V$ un endomorfismo di V. Scelta una base $\mathcal{C}$ di V associamo a f la matrice $M_{\mathcal{C}}^{\mathcal{C}}(f) \in M_{n,n}(\mathbb{K})$. Notate che abbiamo scelto la stessa base per V visto come dominio e come codominio: in questo modo si leggono bene le proprietà di f, per esempio f è l'identità se e solo se $M_{\mathcal{C}}^{\mathcal{C}}(f) = 1_n$. Ora chiediamoci come cambia la matrice associata a f se passiamo dalla base $\mathcal{C}$ a un'altra base $\mathcal{B}$. Per l'equazione (3.6.6) applicata a $f = \mathrm{Id}_V \circ f \circ \mathrm{Id}_V$ e l'equazione (3.10.2) abbiamo che

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = M_{\mathcal{B}}^{\mathcal{B}}(\mathrm{Id}_V \circ f \circ \mathrm{Id}_V) = M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}_V) \cdot M_{\mathcal{C}}^{\mathcal{C}}(f) \cdot M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V) = (M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V))^{-1} \cdot M_{\mathcal{C}}^{\mathcal{C}}(f) \cdot M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V). \quad (3.11.1)$$

Definizione 3.11.2. La matrice $M \in M_{n,n}(\mathbb{K})$ è *coniugata* a $N \in M_{n,n}(\mathbb{K})$ (in simboli $M \sim N$) se esiste $G \in \mathrm{GL}_n(\mathbb{K})$ tale che

$$M = G^{-1} \cdot N \cdot G. \quad (3.11.2)$$

Proposizione 3.11.3. *Sia V uno spazio vettoriale finitamente generato su $\mathbb{K}$, di dimensione n, e sia $f\colon V \to V$ un endomorfismo. Dati una base $\mathcal{C}$ di V e $M \in M_{n,n}(\mathbb{K})$, esiste una base $\mathcal{B}$ di V tale che $M = M_{\mathcal{B}}^{\mathcal{B}}(f)$ se e solo se M è coniugata a $M_{\mathcal{C}}^{\mathcal{C}}(f)$.*

Dimostrazione. Se $M = M_{\mathcal{B}}^{\mathcal{B}}(f)$ allora M è coniugata a $M_{\mathcal{C}}^{\mathcal{C}}(f)$ per l'equazione (3.11.1). Ora supponiamo che M sia coniugata a $M_{\mathcal{C}}^{\mathcal{C}}(f)$, cioè esiste $G \in \mathrm{GL}_n(\mathbb{K})$ tale che $M = G^{-1} \cdot M_{\mathcal{C}}^{\mathcal{C}}(f) \cdot G$. Per la Proposizione 3.10.5 esiste una base $\mathcal{B}$ di V tale che $M_{\mathcal{B}}^{\mathcal{C}}(f) = G^{-1}$. Per l'equazione (3.11.1) segue che $M = M_{\mathcal{B}}^{\mathcal{B}}(f)$. □

Proposizione 3.11.4. *La relazione di coniugio è di equivalenza.*

Dimostrazione. $M \sim M$ perchè $M = 1_n^{-1} \cdot N \cdot 1_n$. Supponiamo che $M \in M_{n,n}(\mathbb{K})$ sia coniugata a N e quindi che valga (3.11.2). Moltiplicando a sinistra ambo i membri di (3.11.2) per G e successivamente a destra per G^{-1} otteniamo che $G \cdot M \cdot G^{-1} = N$. Siccome $G = (G^{-1})^{-1}$ segue che N è coniugata a M. Infine supponiamo che $M \sim N$ e $N \sim P$. Quindi esistono $G, H \in \mathrm{GL}_n(\mathbb{K})$ tali che

$$M = G^{-1} \cdot N \cdot G, \qquad N = H^{-1} \cdot P \cdot H. \quad (3.11.3)$$

La matrice $H \cdot G$ è in $\mathrm{GL}_n(\mathbb{K})$ (ricordate che $\mathrm{GL}_n(\mathbb{K})$ è un gruppo) e, sostituendo l'espressione di N nella prima equazione di (3.11.3), otteniamo che

$$M = G^{-1} \cdot H^{-1} \cdot P \cdot H \cdot G = (H \cdot G)^{-1} \cdot P \cdot (H \cdot G).$$

Quindi M è coniugata a P. □

Abbiamo definito la relazione di coniugio su $M_{n,n}(\mathbb{K})$. Equivalentemente si può definire la seguente relazione su $\mathrm{End}(V)$.

Definizione 3.11.5. Sia V uno spazio vettoriale finitamente generato. Allora $f, g \in \mathrm{End}(V)$ sono *coniugati* se esiste un isomorfismo $\varphi\colon V \to V$ tale che $f = \varphi \circ f \circ \varphi^1$.

Si dimostra facilmente che la relazione appena definita su $\mathrm{End}(V)$ è di equivalenza.

3.12 Diagonalizzazione

Siano V uno spazio vettoriale finitamente generato su $\mathbb{K}$ e $f\colon V \to V$ un endomorfismo. Ci poniamo il problema di trovare una base $\mathcal{B}$ che renda la matrice $M_{\mathcal{B}}^{\mathcal{B}}(f)$ più semplice possibile. L'ideale è trovare una $\mathcal{B}$ tale che $M_{\mathcal{B}}^{\mathcal{B}}(f)$ sia una matrice diagonale.

Definizione 3.12.1. Sia $f\colon V \to V$ un endomorfismo di uno spazio vettoriale V finitamente generato su $\mathbb{K}$. La base $\mathcal{B}$ di V *diagonalizza* f se $M_{\mathcal{B}}^{\mathcal{B}}(f)$ è una matrice diagonale. Diciamo che f è *diagonalizzabile* se esiste una base che la diagonalizza. Sia $A \in M_{n,n}(\mathbb{K})$: una base $\mathcal{B}$ di $\mathbb{K}^n$ *diagonalizza* A se $M_{\mathcal{B}}^{\mathcal{B}}(L_A)$ è una matrice diagonale, e A è *diagonalizzabile* se L_A è diagonalizzabile.

Osservazione 3.12.2. La base $\mathcal{B} = \{v_1, \ldots, v_n\}$ diagonalizza f se e solo se esistano $\lambda_1, \ldots, \lambda_n$ tali che

$$f(v_i) = \lambda_i v_i \qquad i = 1, \ldots, n. \tag{3.12.1}$$

Esempio 3.12.3. Una matrice $A \in M_{n,n}(\mathbb{K})$ diagonale è diagonalizzabile perchè, se $\mathcal{S}$ è la base standard di $\mathbb{K}^n$, $M_{\mathcal{S}}^{\mathcal{S}}(L_A) = A$. Diamo un esempio meno banale. Sia $A \in M_{2,2}(\mathbb{R})$ data da

$$A := \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$$

Allora

$$L_A(1,-1) = (2,-2) = 2(1,-1), \qquad L_A(1,-2) = (3,-6) = 3(1,-2).$$

Quindi la base $\{(1,-1), (1,-2)\}$ di $\mathbb{R}^2$ diagonalizza A, e perciò A è diagonalizzabile.

Osservazione 3.12.4. Per la Proposizione 3.11.3 una matrice $A \in M_{n,n}(\mathbb{K})$ è diagonalizzabile se e solo se esiste $G \in \mathrm{GL}_n(\mathbb{K})$ tale che $\Lambda := G^{-1} \cdot A \cdot G$ sia diagonale, ovvero esiste $G \in \mathrm{GL}_n(\mathbb{K})$ tale che $A = G \cdot \Lambda \cdot G^{-1}$.

L'Osservazione 3.12.2 motiva le seguenti definizioni fondamentali. Sia $\lambda \in \mathbb{K}$: poniamo

$$V_\lambda(f) := \ker(f - \lambda \,\mathrm{Id}_V). \tag{3.12.2}$$

Definizione 3.12.5. Un $\lambda \in \mathbb{K}$ è un *autovalore* di f se $V_\lambda(f) \neq \{0\}$ cioè se esiste $0 \neq v \in V$ tale che $f(v) = \lambda v$. Un tale v si chiama *autovettore* di f. L'*autospazio* associato all'autovalore λ è $V_\lambda(f)$. Se $A \in M_{n,n}(k)$ gli autovalori, autovettori, autospazi di L_A si chiamano anche autovalori, autovettori, autospazi di A.

Esempio 3.12.6. Sia $A \in M_{2,2}(\mathbb{R})$ la matrice dell'Esempio 3.12.3. Allora 2 e 3 sono autovalori di A e gli autospazi relativi sono

$$V_2(L_A) = \langle (1,-1) \rangle, \qquad V_3(L_A) = \langle (1,-2) \rangle.$$

Esempio 3.12.7. Diamo esempi di autovettori e autovalori di un endomorfismo di uno spazio vettoriale che non è finitamente generato. Sia $V := C^\infty(\mathbb{R})$ l'insieme delle funzioni $f\colon \mathbb{R} \to \mathbb{R}$ con derivate di ogni ordine. La somma di funzioni in $C^\infty(\mathbb{R})$ è ancora in $C^\infty(\mathbb{R})$, e se $\lambda \in \mathbb{R}$, $f \in C^\infty(\mathbb{R})$ allora $\lambda \cdot f \in C^\infty(\mathbb{R})$. Con queste operazioni $C^\infty(\mathbb{R})$ è uno spazio vettoriale reale (non finitamente generato). L'applicazione $\Phi\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ definita da $\Phi(f) := f''$ è lineare. Sia $k \in \mathbb{R}$; allora le funzioni $f_k(x) = \sin kx$ e $g_k(x) = \cos kx$ sono autovettori di Φ, con autovalore associato $-k^2$.

La seguente è una riformulazione dell'Osservazione 3.12.2.

Osservazione 3.12.8. Sia V uno spazio vettoriale finitamente generato su $\mathbb{K}$. Un endomorfismo $f\colon V \to V$ è diagonalizzabile se solo se esiste una base di V i cui elementi sono autovettori di f.

Esempio 3.12.9. Dato $\theta \in \mathbb{R}$ sia $R_\theta \in M_{2,2}(\mathbb{R})$ data da

$$R_\theta := \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{3.12.3}$$

Allora R_θ è diagonalizzabile solo se θ è un multiplo intero di π, cioè $\theta = m\pi$ per un $m \in \mathbb{Z}$. Infatti $A_{m\pi} = (-1)^m 1_2$, e quindi è diagonalizzabile, mentre se θ è un multiplo intero di π non esistono autovalori di R_θ (verificatelo) e quindi per l'Osservazione 3.12.8 R_θ non è diagonalizzabile. Geometricamente: siccome R_θ è la matrice associata a una rotazione di angolo θ, vedi l'Esempio 3.6.2, è chiaro che R_θ si diagonalizza solo se θ è un multiplo intero di π.

Se invece consideriamo R_θ come matrice in $M_{2,2}(\mathbb{C})$, allora *è* diagonalizzabile. Infatti $\{(1, i), (1, -i)\}$ è una base di $\mathbb{C}^2$ di autovettroi di R_θ (con autovettori rispettivamente $\cos\theta - i\sin\theta$ e $\cos\theta + i\sin\theta$). Questo esempio dimostra che, data un'inclusione di campi $\mathbb{K} \subset \mathbb{F}$, una matrice non diagonalizzabile $A \in M_{n,n}(\mathbb{K})$ può essere diagonalizzabile vista come matrice in $M_{n,n}(\mathbb{F})$.

Esempio 3.12.10. Siano $\mathbb{K}$ un campo e $N \in M_{2,2}(\mathbb{K})$ data da

$$N := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

La matrice N non è diagonalizzabile. Infatti si verifica facilmente che gli autovettori di N sono $(t, 0)$ con $t \neq 0$ e quindi N non ha una base di autovettori. Osserviamo che, contrariamente alla matrice R_θ dell'Esempio 3.12.9, la matrice N non diventa diagonalizzabile dopo una estensione di campi $\mathbb{K} \subset \mathbb{F}$.

Spieghiamo perchè possiamo essere interessati a trovare una base che diagonalizza un endomorfismo.

Supponiamo che $A \in M_{n,n}(\mathbb{K})$ sia diagonalizzabile. Per l'Osservazione 3.12.4 esiste $G \in \mathrm{GL}_n(\mathbb{K})$ tale che $A = G \cdot \Lambda \cdot G^{-1}$. Questa uguaglianza ci permette di calcolare facilmente tutte le potenze A^r perchè

$$A^r = G \cdot \Lambda^r \cdot G^{-1}, \tag{3.12.4}$$

e Λ^r è una matrice diagonale con entrate le potenze r-esime delle entrate di Λ.

Esempio 3.12.11. Sia $A \in M_{2,2}(\mathbb{R})$ la matrice dell'Esempio 3.12.3, e sia $\mathcal{B} := \{(1, -1), (1, -2)\}$. Allora $\mathcal{B}$ è una base di $\mathbb{R}^2$ che diagonalizza A. Abbiamo

$$M_{\mathcal{B}}^{\mathcal{B}}(L_A) = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

e perciò

$$\begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix} = M_{\mathcal{S}}^{\mathcal{S}}(L_A) = M_{\mathcal{S}}^{\mathcal{B}}(\mathrm{Id}_{\mathbb{R}^2}) \cdot M_{\mathcal{B}}^{\mathcal{B}}(L_A) \cdot M_{\mathcal{B}}^{\mathcal{S}}(\mathrm{Id}_{\mathbb{R}^2}) = \\ = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}^{-1}. \tag{3.12.5}$$

Dalla (3.12.5) segue che

$$\begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}^m = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \cdot \begin{bmatrix} 2^m & 0 \\ 0 & 3^m \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 2^{m+1} - 3^m & 2^m - 3^m \\ -2^{m+1} + 2 \cdot 3^m & -2^m + 2 \cdot 3^m \end{bmatrix}, \tag{3.12.6}$$

Diamo una ulteriore motivazione. Sia $A \in M_{n,n}(\mathbb{R})$, cioè una matrice quadrata a entrate reali. L'esponenziale di A si definisce come segue. Consideriamo la somma

$$\sigma_r := 1_n + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \ldots + \frac{1}{r!}A^r. \tag{3.12.7}$$

Se $|a_{ij}| \leqslant S$ per ogni $i, j \in \{1, \ldots, n\}$ allora il valore assoluto di ciascuna entrata di A^r è al più uguale a $n^{r-1}S^r$: ne segue che le entrate di σ_r sono successioni convergenti (per $r \to \infty$). La matrice le cui entrate sono i limiti delle rispettive successioni di entrate è l'esponenziale di A, e si denota e^A. Scriviamo

$$e^A := \sum_{r=0}^{\infty} \frac{1}{r!}A^r = 1_n + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \ldots + \frac{1}{r!}A^r + \ldots \tag{3.12.8}$$

L'esponenziale è importante perchè vale il seguente analogo all'uguaglianza $d(e^t)/dt = e^t$:

$$\frac{d}{dt}e^{tA} = A \cdot e^{tA}. \tag{3.12.9}$$

Quindi una soluzione del sistema di equazioni differenziali nelle funzioni $y_1, \ldots, y_n \colon \mathbb{R} \to \mathbb{R}$ dato da

$$Y(t)' = A \cdot Y(t), , \qquad Y(0) = B, \tag{3.12.10}$$

($Y \colon \mathbb{R} \to \mathbb{R}^n$ è la funzione con entrate $y_1, \ldots, y_n$, e $B \in \mathbb{R}^n$) è dato da $Y(t) = e^{At} \cdot B$ (si dimostra che è l'unica soluzione).

Ora supponiamo che A sia diagonalizzabile e quindi che valga (3.12.4). Allora si ha che

$$e^{tA} = G \cdot (1_n + t\Lambda + \tfrac{t^2}{2}\Lambda^2 + \tfrac{t^3}{3!}A^3 + \tfrac{t^r}{r!}\Lambda^r + \ldots) \cdot G^{-1} = G \cdot \begin{bmatrix} e^{t\lambda_1} & 0 & \ldots & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & \ldots & \ldots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ldots & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & \ldots & 0 & e^{t\lambda_n} \end{bmatrix} \cdot G^{-1} \tag{3.12.11}$$

Esempio 3.12.12. Sia $A \in M_{2,2}(\mathbb{R})$ la matrice dell'Esempio 3.12.3. Dalla (3.12.6) segue che

$$e^{tA} = \begin{bmatrix} 2e^{2t} - e^{3t} & e^{2t} - e^{3t} \\ -2e^{2t} + 2e^{3t} & -e^{2t} + 2e^{3t} \end{bmatrix}.$$

3.13 Il duale di uno spazio vettoriale

Duale e biduale

Ricordiamo che il duale di uno spazio vettoriale V su $\mathbb{K}$ è lo spazio vettoriale $V^\vee = \mathcal{L}(V, \mathbb{K})$ delle applicazioni lineari $f \colon V \to \mathbb{K}$. Supponiamo che V sia finitamente generato e sia $n := \dim V$; per la Proposizione 3.6.7 abbiamo un isomorfismo $\mathcal{L}(V, \mathbb{K}) \cong M_{1,n}(\mathbb{K})$ e quindi $\dim V^\vee = n$. Questo dimostra il seguente risultato.

Proposizione 3.13.1. *Se V è uno spazio vettoriale finitamente generato su $\mathbb{K}$, allora anche $V^\vee$ è finitamente generato e $\dim V^\vee = \dim V$.*

In particolare, se valgono le ipotesi della Proposizione 3.13.1, allora V è isomorfo a $V^\vee$.

Sia $\mathcal{B} := \{v_1, \ldots, v_n\}$ una base di V. Possiamo definire una base di $V^\vee$ procedendo come segue. Sia $v_i^\vee \in V^\vee$ la funzione lineare

$$\begin{array}{ccc} V & \xrightarrow{v_i^\vee} & \mathbb{K} \\ (x_1v_1 + x_2v_2 + \ldots + x_nv_n) & \mapsto & x_i. \end{array} \tag{3.13.1}$$

In altre parole $v_i^\vee$ è l'unica applicazione lineare $V \to \mathbb{K}$ tale che

$$v_i^\vee(v_j) = \delta_{ij}, \qquad 1 \leqslant i, j \leqslant n \tag{3.13.2}$$

dove δ_{ij} è il simbolo di Kronecker, vedi (3.4.6).

Proposizione 3.13.2. *Sia V uno spazio vettoriale finitamente generato su $\mathbb{K}$ e $\mathcal{B} := \{v_1, \dots, v_n\}$ una sua base. Allora $\mathcal{B}^\vee := \{v_1^\vee, \dots, v_n^\vee\}$ è una base di $V^\vee$.*

Dimostrazione. L'applicazione

$$\begin{array}{ccc} V^\vee & \xrightarrow{\Phi} & \mathbb{K}^n \\ f & \mapsto & (f(v_1), \dots, f(v_n)) \end{array} \tag{3.13.3}$$

è lineare e biunivoca. La prima affermazione è di verifica immediata, la seconda vale per il Corollario 3.1.14. Ora notiamo che $\Phi(v_i^\vee) = e_i$, vedi l'uguaglianza in (3.13.2). Applicando il Corollario 3.2.8 all'isomorphismo $\Phi^{-1}\colon \mathbb{K}^n \to V^\vee$, segue che $\{v_1^\vee, \dots, v_n^\vee\}$ è una base di $V^\vee$. □

Terminologia 3.13.3. La base $\mathcal{B}^\vee := \{v_1^\vee, \dots, v_n^\vee\}$ è la *base duale* della base $\mathcal{B}$.

Osservazione 3.13.4. La notazione per la base duale della base $\mathcal{B}$ è *ingannevole*, perchè suggerisce che abbia senso $v_i^\vee$ indipendentemente dalla scelta della base di cui v_i fa parte. Una notazione corretta sarebbe $(v_i^{\mathcal{B}})^\vee$; per non appesantire la notazione dimentichiamo $\mathcal{B}$.

Sia V uno spazio vettoriale finitamente generato su $\mathbb{K}$, e sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una sua base. Se $\mathcal{B}^\vee = \{v_1^\vee, \dots, v_n^\vee\}$ è la base duale di $V^\vee$, l'applicazione

$$\begin{array}{ccc} V & \xrightarrow{\varphi_{\mathcal{B}}} & V^\vee \\ x_1v_1 + \ldots + x_nv_n & \mapsto & x_1v_1^\vee + \ldots + x_nv_n^\vee \end{array} \tag{3.13.4}$$

è un isomorfismo. Per definire $\varphi_{\mathcal{B}}$ abbiamo scelto la base $\mathcal{B}$, e in generale l'isomorfismo cambia se cambiamo base. Non c'è modo di definire un isomorfismo $V \xrightarrow{\sim} V^\vee$ senza fare delle scelte (a meno che V non abbia dimensione 0, oppure $\mathbb{K} = \mathbb{F}_2$ e $\dim V = 1$), per una versione precisa di quest'affermazione vedi l'Osservazione 3.13.13.

Un fatto notevole è che, se V è finitamente generato, esiste un isomorfismo naturale, cioè definito senza fare scelte, tra V e il *biduale di V*, definito come il duale del duale di V, ovvero

$$V^{\vee\vee} := (V^\vee)^\vee.$$

Per dimostrarlo consideriamo per V uno spazio vettoriale qualsiasi (non necessariamente finitamente generato) e $v \in V$ l'applicazione

$$\begin{array}{ccc} V^\vee & \xrightarrow{\mathrm{Val}(v)} & \mathbb{K} \\ f & \mapsto & f(v) \end{array} \tag{3.13.5}$$

L'applicazione $\mathrm{Val}(v)$ è lineare. Infatti siano $\lambda, \mu \in \mathbb{K}$, e $f, g \in V$; allora

$$\mathrm{Val}(v)(\lambda f + \mu g) = (\lambda f + \mu g)(v) = \lambda f(v) + \mu g(v) = \lambda\,\mathrm{Val}(v)(f) + \mu\,\mathrm{Val}(v)(g).$$

Siccome $\mathrm{Val}(v) \in (V^\vee)^\vee$ abbiamo un'applicazione

$$\begin{array}{ccc} V & \xrightarrow{\mathrm{Val}} & V^{\vee\vee} \\ v & \mapsto & \mathrm{Val}(v) \end{array} \tag{3.13.6}$$

Lemma 3.13.5. *Se V è uno spazio vettoriale (non necessariamente finitamente generato), l'applicazione* Val *in* (3.13.6) *è lineare.*

Dimostrazione. Siano $\lambda, \mu \in \mathbb{K}$, e $v, w \in V$. Allora per definizione di Val abbiamo

$$\mathrm{Val}(\lambda v + \mu w)(f) = f(\lambda v + \mu w) = \lambda f(v) + \mu f(w) = \lambda\,\mathrm{Val}(v)(f) + \mu\,\mathrm{Val}(w)(f).$$

Quindi $\mathrm{Val}(\lambda v + \mu w) = \lambda\,\mathrm{Val}(v) + \mu\,\mathrm{Val}(w)$, e perciò Val è lineare. □

Proposizione 3.13.6. *Se V è uno spazio vettoriale finitamente generato su $\mathbb{K}$, allora l'applicazione* Val *definita da* (3.13.6) *è un isomorfismo di spazi vettoriali.*

Prima di dimostrare la Proposizione 3.13.6, dimostriamo il seguente risultato.

Lemma 3.13.7. *Sia V uno spazio finitamente generato su $\mathbb{K}$, e sia $v \subset V$. Allora v è non nullo se e solo se esiste $f \in V^\vee$ tale che $f(v) \neq 0$.*

Dimostrazione. Se esiste $f \in V^\vee$ tale che $f(v) \neq 0$, allora $v \neq 0$ perchè $f(0) = 0$ per ogni $f \in V^\vee$. Ora supponiamo che $v \neq 0$ e dimostriamo che esiste $f \in V^\vee$ tale che $f(v) \neq 0$. Siccome $v \neq 0$, possiamo completare $v = v_1$ a una base $\{v_1, \dots, v_n\}$ di V. Allora $v_1^\vee(v_1) = 1 \neq 0$. □

Dimostrazione della Proposizione 3.13.6. Sappiamo che Val è lineare. Siccome V è finitamente generato, la Proposizione 3.13.1 dà che

$$\dim V = \dim(V^\vee) = \dim V^{\vee\vee}.$$

Quindi basta dimostrare che Val è iniettiva, ovvero $\ker(\mathrm{Val}) = \{0\}$. Se $0 \neq v \in V$, per il Lemma 3.13.7 esiste $f \in V^\vee$ tale che $f(v) \neq 0$, cioè $\mathrm{Val}(v)(f) \neq 0$. Questo dimostra che $\mathrm{Val}(v) \neq 0$. □

L'applicazione duale di un'applicazione lineare

Definizione 3.13.8. Siano V, W spazi vettoriali su un campo $\mathbb{K}$ e $\varphi\colon V \to W$ un'applicazione lineare. Se $f \in W^\vee$, allora la composizione $f \circ \varphi$ è lineare, e quindi ha senso porre

$$\begin{array}{ccc} W^\vee & \xrightarrow{\varphi^\vee} & V^\vee \\ f & \mapsto & f \circ \varphi \end{array}$$

La $\varphi^\vee$ è l'applicazione *duale* di φ.

Esempio 3.13.9. La duale dell'identità $\mathrm{Id}_V\colon V \to V$ è l'identità $\mathrm{Id}_{V^\vee}\colon V^\vee \to V^\vee$ perchè se $f \in V^\vee$ la composizione $f \circ \mathrm{Id}_V$ è uguale a f. Più in generale la duale della moltiplicazione per $\lambda \in \mathbb{K}$, cioè di $\lambda\,\mathrm{Id}_V$ è uguale a $\lambda\,\mathrm{Id}_{V^\vee}$.

Proposizione 3.13.10. *Siano V, W spazi vettoriali su un campo $\mathbb{K}$ e $\varphi\colon V \to W$ un'applicazione lineare. L'applicazione duale $\varphi^\vee\colon W^\vee \to V^\vee$ è lineare.*

Dimostrazione. Supponiamo che $\lambda_1, \lambda_2 \in \mathbb{K}$ e $f_1, f_2 \in W^\vee$. Allora

$$\varphi^\vee(\lambda_1 f_1 + \lambda_2 f_2)(v) = (\lambda_1 f_1 + \lambda_2 f_2)(\varphi(v)) = \lambda_1 f_1(\varphi(v)) + \lambda_2 f_2(\varphi(v)) = \lambda_1 \varphi^\vee(f_1)(v) + \lambda_2 \varphi^\vee(f_2)(v).$$

□

Supponiamo che V e W siano finitamente generati e siano

$$\mathcal{B} = \{v_1, \dots, v_n\}, \qquad \mathcal{C} = \{w_1, \dots, w_m\} \tag{3.13.7}$$

basi di V e W rispettivamente. Sia $\varphi\colon V \to W$ un'applicazione lineare: allora abbiamo la matrice associata $M_{\mathcal{C}}^{\mathcal{B}}(\varphi) \in M_{m,n}(\mathbb{K})$. Abbiamo anche le basi $\mathcal{C}^\vee$ di $W^\vee$ e $\mathcal{B}^\vee$ di $V^\vee$ e quindi la matrice associata $M_{\mathcal{B}^\vee}^{\mathcal{C}^\vee}(\varphi^\vee) \in M_{n,m}(\mathbb{K})$.

Proposizione 3.13.11. *Siano V, W spazi vettoriali finitamente generati su un campo $\mathbb{K}$ e $\varphi\colon V \to W$ un'applicazione lineare. Siano $\mathcal{B}$ e $\mathcal{C}$ basi di V e W rispettivamente. Allora*

$$M_{\mathcal{B}^\vee}^{\mathcal{C}^\vee}(\varphi^\vee) = M_{\mathcal{C}}^{\mathcal{B}}(\varphi)^t,$$

cioè $M_{\mathcal{B}^\vee}^{\mathcal{C}^\vee}(\varphi^\vee)$ è la trasposta di $M_{\mathcal{C}}^{\mathcal{B}}(\varphi)$.

Dimostrazione. Possiamo supporre che $\mathcal{B}$ e $\mathcal{C}$ siano dati da (3.13.7). Sia $M_{\mathcal{C}}^{\mathcal{B}}(\varphi) = A = (a_{ij})$. Sia $v \in V$ di coordinate $(x_1, \dots, x_n)$ nella base $\mathcal{B}$ cioè $v = \sum_{s=1}^{n} x_s v_s$. Notiamo che $v_s^\vee(v) = x_s$. Abbiamo che

$$\varphi^\vee(w_i^\vee)(v) = w_i^\vee(\varphi(v)) = \sum_{s=1}^{n} a_{is}x_s = \sum_{s=1}^{n} a_{is}v_s^\vee(v).$$

Quindi

$$\varphi^\vee(w_i^\vee) = \sum_{s=1}^{n} a_{is}v_s^\vee. \tag{3.13.8}$$

D'altra parte la colonna i-esima di $M_{\mathcal{B}^\vee}^{\mathcal{C}^\vee}(\varphi^\vee)$ è data dalle coordinate di $\varphi^\vee(w_i^\vee)$ nella base $\mathcal{B}^\vee$ e perciò la (3.13.8) dà che colonna i-esima di $M_{\mathcal{B}^\vee}^{\mathcal{C}^\vee}(\varphi^\vee)$ è la riga i-esima di $M_{\mathcal{C}}^{\mathcal{B}}(\varphi)$. □

Esempio 3.13.12. Sia $A \in M_{m,n}(\mathbb{K})$ e sia $L_A \colon \mathbb{K}^n \to \mathbb{K}^m$ l'applicazione associata ad A, cioè $L_A(X) = A \cdot X$. Se $\mathcal{B}, \mathcal{C}$ sono le matrici standard di $\mathbb{K}^n, \mathbb{K}^m$ rispettivamente, allora $M_{\mathcal{C}}^{\mathcal{B}}(L_A) = A$. Per la Proposizione 3.13.11 abbiamo

$$M_{\mathcal{B}^\vee}^{\mathcal{C}^\vee}(L_A^\vee) = A^t. \tag{3.13.9}$$

Osservazione 3.13.13. Cosa vuol dire che per ogni spazio vettoriale V finitamente generato su $\mathbb{K}$ esiste un isomorfismo $V \to V^\vee$ naturale, cioè che non dipende da scelte? Una formulazione è la seguente. Per ogni V è dato un isomorfismo $\varphi_V \colon V \xrightarrow{\sim} V^\vee$, e ogni volta che si ha un isomorfismo $f \colon V \to W$ tra spazi vettoriali V finitamente generati su $\mathbb{K}$, il diagramma

$$\begin{array}{ccc} V & \xrightarrow[\varphi_V]{} & V^\vee \\ {\scriptstyle f}\downarrow & & \uparrow{\scriptstyle f^\vee} \\ W & \xrightarrow[\varphi_W]{} & W^\vee \end{array} \tag{3.13.10}$$

commuta, cioè

$$f^\vee \circ \varphi_W \circ f = \varphi_V. \tag{3.13.11}$$

Dimostriamo che non esiste una tale collezione di isomorfismi $\varphi_V \colon V \xrightarrow{\sim} V^\vee$. Ponendo $V = W$ e $f = \lambda \operatorname{Id}_V$, dove $\lambda \in \mathbb{K}$ l'equazione in (3.13.11) diventa $\lambda^2 \varphi_V = \varphi_V$. Se $V \neq 0$ questa uguaglianza vale solo se $\lambda^2 = 1$. Se $\mathbb{K} \neq \mathbb{F}_2$ esistono $\lambda \in \mathbb{K}$ tali che $\lambda^2 \neq 1$. Questo mostra che non esiste una tale collezione di tali isomorfismi se $\mathbb{K} \neq \mathbb{F}_2$. Lasciamo al lettore il compito di dimostrare che non esistono anche se $\mathbb{K} = \mathbb{F}_2$.

Duale di un'applicazione lineare: proprietà funtoriali

Proposizione 3.13.14. *Siano U, V, W spazi vettoriali su $\mathbb{K}$ e*

$$U \xrightarrow{\psi} V \xrightarrow{\varphi} W$$

applicazioni lineari. Allora

$$(\varphi \circ \psi)^\vee = \psi^\vee \circ \varphi^\vee. \tag{3.13.12}$$

Dimostrazione. Innanzitutto notiamo che ambo i membri dell'uguaglianza in (3.13.12) sono applicazioni (lineari) da $W^\vee$ a $U^\vee$. Sia $f \in W^\vee$; allora

$$(\varphi \circ \psi)^\vee(f) = f \circ (\varphi \circ \psi) = (f \circ \varphi) \circ \psi = \psi^\vee(\varphi^\vee(f)) = (\psi^\vee \circ \varphi^\vee)(f).$$

□

Corollario 3.13.15. *Siano V, W spazi vettoriali su un campo $\mathbb{K}$ e $\varphi \colon V \to W$ un isomorfismo. Allora l'applicazione duale $\varphi^\vee \colon W^\vee \to V^\vee$ è un isomorfismo.*

Dimostrazione. Dimostriamo che $(\varphi^{-1})^\vee$ è un'inversa di $\varphi^\vee$. Per la Proposizione 3.13.14 e l'Esempio 3.13.9 abbiamo

$$\varphi^\vee \circ (\varphi^{-1})^\vee = (\varphi^{-1} \circ \varphi)^\vee = \mathrm{Id}_V^\vee = \mathrm{Id}_{V^\vee}, \quad (\varphi^{-1})^\vee \circ \varphi^\vee = (\varphi \circ \varphi^{-1})^\vee = \mathrm{Id}_W^\vee = \mathrm{Id}_{W^\vee}.$$

□

Sia $\varphi\colon V \to W$ un'applicazione lineare tra spazi vettoriali finitamente generati su $\mathbb{K}$. La duale dell'applicazione lineare $\varphi^\vee\colon W^\vee \to V^\vee$ è un'applicazione lineare $(\varphi^\vee)^\vee\colon V^{\vee\vee} \to W^{\vee\vee}$. Per la Proposizione 3.13.6 abbiamo isomorfismi $\mathrm{Val}_V\colon V \xrightarrow{\sim} V^{\vee\vee}$ e $\mathrm{Val}_W\colon W \xrightarrow{\sim} W^{\vee\vee}$, e quindi possiamo identificare $(\varphi^\vee)^\vee$ con un'applicazione $V \to W$. Vogliamo dimostrare che, con questa identificazione, $(\varphi^\vee)^\vee$ è uguale a φ. Per essere più precisi, consideriamo il diagramma di applicazioni lineari:

$$\begin{array}{ccc} V^{\vee\vee} & \xrightarrow{(\varphi^\vee)^\vee} & W^{\vee\vee} \\ {\scriptstyle \mathrm{Val}_V}\uparrow & & \uparrow{\scriptstyle \mathrm{Val}_W} \\ V & \xrightarrow[\varphi]{} & W \end{array} \tag{3.13.13}$$

Proposizione 3.13.16. *Siano V, W spazi vettoriali finitamente generati su $\mathbb{K}$, e sia $\varphi\colon V \to W$ un'applicazione lineare. Allora* $\mathrm{Val}_W \circ \varphi = (\varphi^\vee)^\vee \circ \mathrm{Val}_V$.

Dimostrazione. Dimostriamo che l'applicazione $\mathrm{Val}_W \circ \varphi\colon V \to W^{\vee\vee}$ è come segue:

$$\begin{array}{ccc} V & \xrightarrow{\mathrm{Val}_W \circ \varphi} & W^{\vee\vee} \\ v & \mapsto & (f \mapsto f(\varphi(v))) \end{array}$$

Infatti $\mathrm{Val}_W \circ \varphi(v) = \mathrm{Val}_W(\varphi(v))$, e $\mathrm{Val}_W(\varphi(v))$ è per definizione l'applicazione $W^\vee \to \mathbb{K}$ che manda $f \in W^\vee$ in $f(\varphi(v))$. Per finire basta dimostrare che l'applicazione $(\varphi^\vee)^\vee \circ \mathrm{Val}_V$ è data dalla stessa formula:

$$\begin{array}{ccc} V & \xrightarrow{(\varphi^\vee)^\vee \circ \mathrm{Val}_V} & W^{\vee\vee} \\ v & \mapsto & (f \mapsto f(\varphi(v))) \end{array}$$

Per definizione di $\mathrm{Val}_V(v)$ e del duale di un'applicazione lineare, abbiamo

$$(\varphi^\vee)^\vee \circ \mathrm{Val}_V(v) = (\varphi^\vee)^\vee(\mathrm{Val}_V(v)) = \mathrm{Val}_V(v) \circ \varphi^\vee.$$

Quindi se $f \in W^\vee$,

$$(\varphi^\vee)^\vee \circ \mathrm{Val}_V(v)(f) = \mathrm{Val}_V(v) \circ \varphi^\vee(f) = \mathrm{Val}_V(v)\,(f \circ \varphi) = f(\varphi(v)).$$

□

Corollario 3.13.17. *Siano V, W spazi vettoriali finitamente generati su $\mathbb{K}$, e sia $\varphi\colon V \to W$ un'applicazione lineare. Allora $(\varphi^\vee)^\vee$ è iniettiva se e solo se lo è φ e, analogamente, $(\varphi^\vee)^\vee$ è suriettiva se e solo se lo è φ.*

Dimostrazione. Per la Proposizione 3.13.16 abbiamo

$$(\varphi^\vee)^\vee = \mathrm{Val}_W \circ \varphi \circ \mathrm{Val}_V^{-1}, \qquad \varphi = \mathrm{Val}_W^{-1} \circ (\varphi^\vee)^\vee \circ \mathrm{Val}_V\,.$$

Il risultato segue da queste uguaglianze perchè Val_V e Val_W sono isomorfismi. □

Proposizione 3.13.18. *Siano V, W spazi vettoriali finitamente generati su un campo $\mathbb{K}$ e $\varphi\colon V \to W$ un'applicazione lineare.*

1. *φ è iniettiva se e solo se $\varphi^\vee$ è suriettiva.*

2. φ è suriettiva se e solo se $\varphi^\vee$ è iniettiva.

Prima di dimostrare la Proposizione 3.13.18, dimostriamo un caso molto particolare. Siano V uno spazio vettoriale e $U \subset V$ un sottospazio. L'inclusione

$$\begin{array}{ccc} U & \overset{\iota}{\hookrightarrow} & V \\ u & \mapsto & u \end{array}$$

è evidentemente un'applicazione lineare.

Proposizione 3.13.19. *Sia V uno spazio finitamente generato su $\mathbb{K}$, e sia $U \subset V$ un sottospazio vettoriale. La duale dell'inclusione $\iota\colon U \hookrightarrow V$ è suriettiva.*

Dimostrazione. Iniziamo notando che la duale di ι è la restrizione a U:

$$\begin{array}{ccc} V^\vee & \overset{\iota^\vee}{\longrightarrow} & U^\vee \\ f & \mapsto & f|_U. \end{array} \tag{3.13.14}$$

Sia $\mathcal{B} := \{u_1, \dots, u_m\}$ una base di U, ed estendiamola a una base $\mathcal{C} := \{u_1, \dots, u_m, w_1, \dots, w_n\}$ di V. Sia $\mathcal{C}^\vee := \{u_1^\vee, \dots, u_m^\vee, w_1^\vee, \dots, w_n^\vee\}$ la base duale di $\mathcal{C}$. Le restrizioni $u_1^\vee|_U, \dots, u_m^\vee|_U$ danno la base duale $\mathcal{B}^\vee$ di $\mathcal{B}$. Quindi l'immagine di $\iota^\vee$ contiene i vettori di una base di $U^\vee$, e siccome $\iota^\vee$ è lineare segue che è suriettiva. □

Dimostrazione della Proposizione 3.13.18. Sia $U := \operatorname{im}\varphi$. L'applicazione φ definisce un'applicazione lineare

$$\begin{array}{ccc} V & \overset{\psi}{\longrightarrow} & U \\ v & \mapsto & \varphi(v). \end{array}$$

Sia $\iota\colon U \hookrightarrow W$ l'inclusione. Allora $\varphi = \iota \circ \psi$, e quindi

$$\varphi^\vee = \psi^\vee \circ \iota^\vee \tag{3.13.15}$$

per la Proposizione 3.13.14.

Dimostriamo che vale (1). Supponiamo che φ sia iniettiva. Allora ψ è un isomorfismo, e quindi anche $\psi^\vee$ è un isomorfismo per il Corollario 3.13.15. D'altra parte $\iota^\vee$ è suriettiva per la Proposizione 3.13.19, e ne segue che $\varphi^\vee$ è suriettiva.

Ora supponiamo che φ non sia iniettiva, e sia $0 \neq v \in \ker\varphi$. Allora $f(v) = 0$ per ogni $f \in \operatorname{im}\varphi^\vee$. D'altra parte, per il Lemma 3.13.7 esiste $g \in V^\vee$ tale che $g(v) \neq 0$, e quindi $g \notin \operatorname{im}\varphi^\vee$. Questo dimostra che $\varphi^\vee$ non è suriettiva.

Per dimostare che vale (2) si può procedere direttamente come sopra, oppure si può dedurre (2) da (1) usando il Corollario 3.13.17. Seguiamo quest'ultima procedura (tipica della dualità). Supponiamo che φ sia suriettiva, e dimostriamo che $\varphi^\vee$ è iniettiva. Per il Corollario 3.13.17 $(\varphi^\vee)^\vee$ è suriettiva. Per il punto (1) (appena dimostrato) applicato a $\varphi^\vee$ segue che $\varphi^\vee$ è iniettiva.

Ora supponiamo che $\varphi^\vee$ sia iniettiva, e dimostriamo che φ è suriettiva. Per il punto (1) (appena dimostrato) applicato a $\varphi^\vee$ segue (dall'iniettività di $\varphi^\vee$) che $\varphi^{\vee\vee}$ è suriettiva, e per il Corollario 3.13.17 deduciamo che φ è suriettiva. □

Proposizione 3.13.20. *Sia $\varphi\colon V \to W$ un'applicazione lineare tra spazi vettoriali finitamente generati su un campo $\mathbb{K}$. Allora* $\operatorname{rg}(\varphi) = \operatorname{rg}(\varphi^\vee)$.

Dimostrazione. Sia $U := \operatorname{im}\varphi$, e siano $\psi\colon V \to U$ e $\iota\colon U \hookrightarrow W$ come nella dimostrazione della Proposizione 3.13.18, in particolare $\varphi^\vee = \psi^\vee \circ \iota^\vee$. Ora $\iota^\vee$ è suriettiva per la Proposizione 3.13.19, e ne segue che $\operatorname{im}(\varphi^\vee) = \operatorname{im}(\psi^\vee)$. D'altra parte $\psi^\vee$ è iniettiva per la Proposizione 3.13.18, e quindi

$$\operatorname{rg}(\varphi^\vee) = \dim\operatorname{im}(\varphi^\vee) = \dim\operatorname{im}(\psi^\vee) = \dim U^\vee = \dim U = \operatorname{rg}(\varphi).$$

□

Osservazione 3.13.21. L'Esempio 3.13.12 e la Proposizione 3.13.20 danno una dimostrazione dell'uguaglianza $\operatorname{rg}(A) = \operatorname{rg}(A^t)$ della Proposizione 3.8.15 diversa da quella data precedentemente.

Esercizi del Capitolo 3

Esercizio 3.1. *Sia $\mathbb{K}$ un campo. Quali delle seguenti applicazioni tra spazi vettoriali su $\mathbb{K}$ è lineare ?*

(1) Sia $p_0 \in \mathbb{K}[x]$ e definiamo $\Phi\colon \mathbb{K}[x] \to \mathbb{K}[x]$ ponendo $\Phi(p) := p_0 \cdot p$.

(2) Sia $\Psi\colon \mathbb{K}[x] \to \mathbb{K}[x]$ definita da $\Psi(p) := p^2$. (Attenzione: la risposta dipende dal campo $\mathbb{K}$.)

(3) Sia $\Theta\colon \mathbb{K}[x] \to \mathbb{K}[x]$ definita da $\Theta(p) := p(x^2)$.

(4) Sia $F\colon \mathbb{K}[x] \to \mathbb{K}$ definita da $F(p) := p(0) + p(1)$.

Esercizio 3.2. *Il campo dei complessi $\mathbb{C}$ è sia uno spazio vettoriale su $\mathbb{C}$ che su $\mathbb{R}$. Sia*

$$\begin{array}{ccc} \mathbb{C} & \xrightarrow{f} & \mathbb{C} \\ z & \mapsto & \overline{z} \end{array}$$

la coniugazione complessa (vedi la Definizione 1.9.2). Verificate che

*(1) f **è** un'applicazione lineare di spazi vettoriali reali.*

*(2) f **non** è un'applicazione lineare di spazi vettoriali complessi.*

Esercizio 3.3. *Ridimostrate la Formula di Grassmann, cioè la Proposizione 2.7.1, considerando l'applicazione*

$$\begin{array}{ccc} U \oplus W & \xrightarrow{F} & V \\ (u, w) & \mapsto & u - v \end{array}$$

Dimostrate che

(a) F è lineare,

(b) $\operatorname{im} F = U + W$, e

(c) $\dim(\ker F) = \dim(U \cap W)$ (definite un isomorfismo tra $U \cap W$ e $\ker F$).

Infine dimostrate la Formula di Grassmann applicando la Proposizione 3.1.21 a F.

Esercizio 3.4. *Sia $\mathbb{K}$ un campo e siano $\alpha_0, \ldots, \alpha_n \in \mathbb{K}$* distinti. *Dimostrate che l'applicazione*

$$\begin{array}{ccc} \mathbb{K}[x]_{\leqslant n} & \longrightarrow & \mathbb{K}^{n+1} \\ p & \mapsto & (p(\alpha_0), \ldots, p(\alpha_n)) \end{array}$$

è un isomorfismo.

Esercizio 3.5. *Calcolate $A \cdot B$ per le matrici*

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 1 & 3 \\ 4 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}.$$

Esercizio 3.6. *Sia*

$$A := \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

e $L_A\colon \mathbb{R}^3 \to \mathbb{R}^2$ l'applicazione lineare associata ad A. Calcolate una base di $\ker(L_A)$.

Esercizio 3.7. *Sia*

$$B := \begin{bmatrix} 1 & 2 & -1 \\ -1 & 1 & -1 \\ 0 & -3 & 2 \end{bmatrix}$$

e $L_B\colon \mathbb{R}^3 \to \mathbb{R}^3$ l'applicazione lineare associata a B. Calcolate una base di $\operatorname{im}(L_B)$.

Esercizio 3.8. *La successione di Fibonacci $\{x_n\}_{n\in\mathbb{N}}$ è definita ricorsivamente così: $1 = x_0 = x_1$ e*

$$x_{n+1} = x_n + x_{n-1}, \qquad n \geqslant 1. \tag{3.13.16}$$

Sia

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

Per $n \in \mathbb{N}$ definiamo x_n, y_n così:

$$(x_n, y_n) := L_{A^{n-1}}(1, 1).$$

Dimostrate che $\{x_n\}$ è la successione di Fibonacci.

Esercizio 3.9. *Lo scopo di questo esercizio è di dimostrare la seguente formula chiusa per i numeri di Fibonacci (vedi l'Esercizio 3.8):*

$$x_n = \frac{1}{\sqrt{5}} \cdot \left(\left(\frac{1+\sqrt{5}}{2} \right)^{n+1} - \left(\frac{1-\sqrt{5}}{2} \right)^{n+1} \right). \tag{3.13.17}$$

(a) *Sia $A \in M_{2,2}(\mathbb{R})$ data da*

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

Dimostrate che la base $\mathcal{B} := \{(2, -1+\sqrt{5}), (-2, 1+\sqrt{5})\}$ di $\mathbb{R}^2$ diagonalizza A, e più precisamente

$$L_A(2, -1+\sqrt{5}) = \left(\frac{1+\sqrt{5}}{2} \right) (2, -1+\sqrt{5}), \qquad L_A(2, -1-\sqrt{5}) = \left(\frac{1-\sqrt{5}}{2} \right) (2, -1-\sqrt{5}).$$

(b) *Deducete dal punto (a) che*

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -1+\sqrt{5} & 1+\sqrt{5} \end{bmatrix} \cdot \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} \cdot \begin{bmatrix} 2 & -2 \\ -1+\sqrt{5} & 1+\sqrt{5} \end{bmatrix}^{-1}.$$

(c) *Verificate che*

$$\begin{bmatrix} 2 & -2 \\ -1+\sqrt{5} & 1+\sqrt{5} \end{bmatrix}^{-1} = \frac{1}{4\sqrt{5}} \begin{bmatrix} 1+\sqrt{5} & 2 \\ 1-\sqrt{5} & 2 \end{bmatrix}$$

e deducetene che

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n = \frac{1}{\sqrt{5}} \begin{bmatrix} (\frac{1+\sqrt{5}}{2})^{n+1} - (\frac{1-\sqrt{5}}{2})^{n+1} & (\frac{1+\sqrt{5}}{2})^{n} - (\frac{1-\sqrt{5}}{2})^{n} \\ (\frac{1+\sqrt{5}}{2})^{n} - (\frac{1-\sqrt{5}}{2})^{n} & (\frac{1+\sqrt{5}}{2})^{n-1} - (\frac{1-\sqrt{5}}{2})^{n-1} \end{bmatrix} \tag{3.13.18}$$

(d) *Deducete la Formula (3.13.17) dal punto (c).*

Esercizio 3.10. *Sia*

$$C := \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

e $L_C \colon \mathbb{R}^2 \to \mathbb{R}^2$ l'applicazione lineare associata a C. Sia $\mathcal{B}$ la base di $\mathbb{R}^2$ data da $\mathcal{B} = \{(1,1), (1,-1)\}$.

(1) *Calcolate $M_{\mathcal{B}}^{\mathcal{B}}(L_C)$.*

(2) *Calcolate $L_{C^n}((1,-1))$.*

Esercizio 3.11. *Sia*

$$D := \begin{bmatrix} 2 & 0 & 3 \\ 1 & 1 & -2 \\ -1 & 1 & 1 \end{bmatrix}$$

e $L_D \colon \mathbb{R}^3 \to \mathbb{R}^3$ l'applicazione lineare associata a D. Sia $V \subset \mathbb{R}^3$ il sottospazio definito da

$$V := \{(x_1, x_2, x_3) \mid x_1 + x_2 + x_3 = 0\}.$$

(1) *Dimostrate che $L_D(V) \subset V$ e quindi possiamo definire un'applicazione lineare*

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ X & \mapsto & L_D(X) \end{array}$$

(2) *Sia $\mathcal{B}$ la base di V data da $\mathcal{B} = \{(1,-1,0), (0,1,-1)\}$. Calcolate $M_{\mathcal{B}}^{\mathcal{B}}(f)$.*

Esercizio 3.12. *Sia V uno spazio vettoriale finitamente generato di dimensione n. Sia $\mathcal{B}$ una base di V.*

(1) *Dimostrate che $M_{\mathcal{B}}^{\mathcal{B}}(\mathrm{Id}_V) = 1_n$.*

(2) *Sia $f \in \mathcal{L}(V,V)$. Dimostrate che f è un isomorfismo se e solo se $M_{\mathcal{B}}^{\mathcal{B}}(f)$ è invertibile e che in questo caso $M_{\mathcal{B}}^{\mathcal{B}}(f^{-1}) = M_{\mathcal{B}}^{\mathcal{B}}(f)^{-1}$.*

Esercizio 3.13. *Siano $A \in M_{m,n}(\mathbb{K})$ e $B \in M_{n,p}(\mathbb{K})$. Dimostrate che*

$$(A \cdot B)^t = B^t \cdot A^t.$$

Esercizio 3.14. *Sia V uno spazio vettoriale finitamente generato su un campo $\mathbb{K}$, di dimensione n.*

(1) Sia $f \in \mathcal{L}(V,V)$ e supponiamo che esista una base $\mathcal{B} = \{v_1, \dots, v_n\}$ di V tale che

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = (\lambda_i \delta_{ij}), \qquad \lambda_i^2 = 1. \tag{3.13.19}$$

Dimostrate che $f \circ f = \mathrm{Id}_V$.

(2) Ora supponiamo che $\operatorname{char}\mathbb{K} \neq 2$, che $f \in \mathcal{L}(V,V)$ e che $f \circ f = \mathrm{Id}_V$. Dimostrate che esiste una base $\mathcal{B}$ di V tale che valga (3.13.19). (Suggerimento: osservate che vale (3.13.19) se e solo se $f(v_i) = \lambda_i v_i$. Dato $v \in V$ calcolate $f(v \pm f(v))$.)

(3) Date un esempio di spazio vettoriale V finitamente generato su un campo $\mathbb{K}$ e $f \in \mathcal{L}(V,V)$ tale che $f \circ f = \mathrm{Id}_V$ ma non *esiste una base $\mathcal{B}$ di V tale che valga (3.13.19). (Per il punto (2) dovrà valere $\operatorname{char}\mathbb{K} = 2$.)*

Esercizio 3.15. *Siano $U, W \subset \mathbb{R}^4$ i sottospazi dati da*

$$U := \langle (1,2,3,-1), (3,5,0,2) \rangle, \qquad W := \langle (-1,0,3,2), (1,-1,1,-1), (1,-2,5,0) \rangle.$$

Date equazioni cartesiane di U e W.

Esercizio 3.16. *Sia $\mathbb{K}$ un campo e $V \subset \mathbb{K}^n$ il sottospazio*

$$V := \{X \mid x_1 + \ldots + x_n = 0\}.$$

Dare una base di $V^\vee$.

Esercizio 3.17. *Sia $\mathbb{K}$ un campo e $\Phi, \Psi \colon \mathbb{K}[x] \to \mathbb{K}[x]$ le applicazioni lineari date da*

$$\begin{array}{ccc} \mathbb{K}[x] & \xrightarrow{\Phi} & \mathbb{K}[x] \\ p & \mapsto & (x^2+3)\cdot p \end{array} \qquad \begin{array}{ccc} \mathbb{K}[x] & \xrightarrow{\Psi} & \mathbb{K}[x] \\ p(x) & \mapsto & p(-x) \end{array}$$

Siano $f, g \colon \mathbb{K}[x]^\vee$ le funzioni definite da

$$\begin{array}{ccc} \mathbb{K}[x] & \xrightarrow{f} & \mathbb{K} \\ q & \mapsto & q(0) \end{array} \qquad \begin{array}{ccc} \mathbb{K}[x] & \xrightarrow{g} & \mathbb{K} \\ q & \mapsto & q(1) \end{array}$$

Determinate

$$\Phi^\vee(f), \quad \Phi^\vee(g), \quad \Psi^\vee(f), \quad \Psi^\vee(g).$$

Esercizio 3.18. *Siano V uno spazio vettoriale su un campo $\mathbb{K}$, e $W \subset V$ un sottospazio. L'*annullatore *di W è il sottoinsieme* $\operatorname{Ann} W \subset V^\vee$ *definito da*

$$\operatorname{Ann} W := \{\varphi \in V^\vee \mid \varphi|_W = 0\}. \tag{3.13.20}$$

1. *Verificate che* $\operatorname{Ann} W$ *è un sottospazio di $V^\vee$.*
2. *Sia $\pi \colon V \to V/W$ l'applicazione quoziente. Dimostrate che*

$$(V/W)^\vee \xrightarrow{\pi^\vee} V^\vee$$

definisce un'isomorfismo tra $(V/W)^\vee$ e $\operatorname{Ann} W$.

3. *Supponiamo che V/W sia finitamente generato, e quindi anche $(V/W)^\vee$. Siano $\varphi_1, \dots, \varphi_d$ generatori di $(V/W)^\vee$. Si dimostri che*

$$W = \{v \in V \mid 0 = \varphi_1(v) = \ldots = \varphi_d(v)\}. \tag{3.13.21}$$

(Le $0 = \varphi_1(v) = \ldots = \varphi_d(v)$ si dicono equazioni cartesiane *di W.)*

Esercizio 3.19. *Sia $A \in M_{3,3}(\mathbb{R})$ data da*

$$A := \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{bmatrix}$$

1. *Verificate che A è invertibile.*

2. Calcolate A^{-1}.

Esercizio 3.20. *Sia $t \in \mathbb{R}$ e $A_t \in M_{3,3}(\mathbb{R})$ data da*

$$A_t := \begin{bmatrix} 2 & 3 & 1 \\ 3 & 5 & 0 \\ 2 & 4 & t \end{bmatrix}$$

(1) Determinare per quali t la matrice A_t è invertibile.

(2) Determinare A_t^{-1} per quei t tali che A_t è invertibile.

Esercizio 3.21. *Siano $\mathcal{B}$ e $\mathcal{C}$ le basi di $\mathbb{R}^3$ date da*

$$\mathcal{B} := \{(3,1,5),(2,1,0),(1,-1,16)\}, \quad \mathcal{C} := \{(4,5,1),(3,4,3),(2,0,-20)\}.$$

Determinate la matrice del cambiamento di base da $\mathcal{B}$ a $\mathcal{C}$.

Esercizio 3.22. *Sia $M \in M_{2,2}(\mathbb{R})$ la matrice definita da*

$$M := \begin{bmatrix} 2 & 5 \\ 1 & -2 \end{bmatrix}$$

Sia $\mathcal{B}$ la base di $\mathbb{R}^2$ data da $\mathcal{B} := \{(5,1),(1,-1)\}$.

(1) Determinare $M_{\mathcal{B}}^{\mathcal{B}}(L_M)$.

(2) Calcolare (scrivere in "forma chiusa") M^s per ogni $s \in \mathbb{N}$.

Esercizio 3.23. *Siano $A \in M_{3,3}(\mathbb{R})$ data da*

$$A := \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 6 & -11 & 6 \end{bmatrix}$$

e $\mathcal{B}$ la base di $\mathbb{R}^3$ data da

$$\mathcal{B} := \{(1,1,1),(1,2,4),(1,3,9)\}.$$

Calcolate $M_{\mathcal{B}}^{\mathcal{B}}(L_A)$.

Esercizio 3.24. *Sia V uno spazio vettoriale su $\mathbb{K}$, e sia $f\colon V \to \mathbb{K}^n$ un isomorfismo. Dimostrate che esiste una base (unica) $\mathcal{B}$ di V tale che $X_{\mathcal{B}}(v) = f(v)$ pr ogni $v \in V$.*

Esercizio 3.25. *Sia $f\colon \mathbb{Q}[x]_{\leqslant 2} \longrightarrow \mathbb{Q}^3$ l'applicazione lineare*

$$\begin{array}{ccc} \mathbb{Q}[x]_{\leqslant 2} & \xrightarrow{f} & \mathbb{Q}^3 \\ p & \mapsto & (p(1),p(2),p(3)) \end{array}$$

Siccome f è un isomorfismo (vedi l'Esercizio 3.4), esiste una base $\mathcal{B}$ di $\mathbb{Q}[x]_{\leqslant 2}$ tale che $X_{\mathcal{B}}(p) = f(p)$ per ogni $p \in \mathbb{Q}[x]_{\leqslant 2}$. D'altra parte sia $\mathcal{M} = \{1, x, x^2\}$ la base monomiale di $\mathbb{Q}[x]_{\leqslant 2}$. Calcolate la matrice dela cambiamento di base da $\mathcal{B}$ a $\mathcal{M}$.

Esercizio 3.26. *Dimostrate che le matrici*

$$A := \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \quad B := \begin{bmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix}$$

sono coniugate. Più in generale dimostrare che ogni matrice $A = (a_{ij})_{i,j=1}^n \in M_{n\times n}(\mathbb{K})$ è coniugata alla matrice $B = (b_{ij})_{i,j=1}^n$ ottenuta per "rotazione di 180°", ossia $b_{ij} = a_{n+1-i,n+1-j}$.

Capitolo 4

Spazi affini

4.1 Spazi affini

Ricordiamo che $\mathbb{E}^2$ è il piano della geometria euclidea, e che $\mathsf{V}(\mathbb{E}^2)$ è lo spazio vettoriale dei vettori geometrici, cioè l'insieme delle classi di equipollenza di segmenti orientati. Consideriamo l'applicazione

$$\begin{array}{ccc} \mathsf{V}(\mathbb{E}^2) \times \mathbb{E}^2 & \longrightarrow & \mathbb{E}^2 \\ (v, P) & \mapsto & P + v \end{array} \tag{4.1.1}$$

definita come segue. Sia $\overrightarrow{PQ}$ l'unico segmento orientato con punto iniziale P che rappresenta il vettore v: poniamo $P + v := Q$. Ora osserviamo che

(a) $P + 0 = P$ per ogni $P \in \mathbb{E}^2$.

(b) $P + (v + w) = (P + v) + w$ per ogni $P \in \mathbb{E}^2$ e $v, w \in \mathsf{V}(\mathbb{E}^2)$.

(c) dati $P, Q \in \mathbb{E}^2$ esiste un unico $v \in \mathsf{V}(\mathbb{E}^2)$ tale che $P + v = Q$.

Infatti (a) vale perchè $P + 0 = P + \overrightarrow{PP} = P$. Per dimostrare che vale (b) poniamo $v = \overrightarrow{PQ}$ e $w = \overrightarrow{QR}$. Allora

$$P + (v + w) = P + (\overrightarrow{PQ} + \overrightarrow{QR}) = P + (\overrightarrow{PR}) = R, \qquad (P + \overrightarrow{PQ}) + \overrightarrow{QR} = Q + \overrightarrow{QR} = R.$$

La (c) vale con $v = \overrightarrow{PQ}$, ed è chiaro che questo è l'unico v tale che valga $P + v = Q$.

In generale uno spazio affine è un insieme (di "punti") su cui agisce uno spazio vettoriale in modo che valgano proprietà simili ad (a), (b) e (c).

Definizione 4.1.1. Sia V uno spazio vettoriale sul campo $\mathbb{K}$. Uno *spazio affine con gruppo delle traslazioni* V è un insieme non vuoto $\mathbb{S}$ provvisto di un'applicazione

$$\begin{array}{ccc} V \times \mathbb{S} & \longrightarrow & \mathbb{S} \\ (v, P) & \mapsto & P + v \end{array} \tag{4.1.2}$$

che gode delle seguenti proprietà:

(1) $P + 0 = P$ per ogni $P \in \mathbb{S}$.

(2) $P + (v + w) = (P + v) + w$ per ogni $P \in \mathbb{S}$ e $v, w \in V$.

(3) dati $P, Q \in \mathbb{S}$ esiste un unico $v \in V$ tale che $P + v = Q$.

Gli elementi di $\mathbb{S}$ si dicono *punti*. Denotiamo con $\mathsf{V}(\mathbb{S})$ il gruppo delle traslazioni di $\mathbb{S}$, cioè $\mathsf{V}(\mathbb{S}) = V$. Diciamo che $\mathbb{S}$ è uno *spazio affine su* $\mathbb{K}$ se $\mathsf{V}(\mathbb{S})$ è uno spazio vettoriale su $\mathbb{K}$.

Esempio 4.1.2. Siano $\mathbb{K}$ un campo, $A \in M_{m,n}(\mathbb{K})$ e $B \in M_{m,1}(\mathbb{K})$. Sia

$$\mathbb{S} := \{Y \in \mathbb{K}^n \mid A \cdot Y = B\} \tag{4.1.3}$$

l'insieme delle soluzioni di un sistema di equazioni lineari (in generale non omogenee), e sia $V \subset \mathbb{K}^n$ il sottospazio vettoriale delle soluzioni dell'associato sistema di equazioni lineari omogenee, cioè

$$V := \{X \in \mathbb{K}^n \mid A \cdot X = 0\}. \tag{4.1.4}$$

Supponendo che $\mathbb{S}$ non sia vuoto, definiamo $V \times \mathbb{S} \to \mathbb{S}$ come segue. Se $X \in V$ e $Y \in \mathbb{S}$, allora $(Y + X) \in \mathbb{S}$ perchè

$$A \cdot (Y + X) = A \cdot Y + A \cdot X = B + 0 = B.$$

Quindi la somma di vettori definisce un'applicazione

$$\begin{array}{ccc} V \times \mathbb{S} & \longrightarrow & \mathbb{S} \\ (X, Y) & \mapsto & Y + X \end{array} \tag{4.1.5}$$

Valgono evidentemente (1), (2) della Definizione 4.1.1. Vale anche (3) perchè se $Y, Z \in \mathbb{S}$, allora

$$A \cdot (Z - Y) = A \cdot Z - A \cdot Y = B - B = 0,$$

quindi $(Z - Y) \in V$, e si ha $Y + (Z - Y) = Z$, ed è evidente che $Z - X$ è l'unico vettore di V con tale proprietà.

Esempio 4.1.3. Sia V uno spazio vettoriale su $\mathbb{K}$. L'applicazione

$$\begin{array}{ccc} V \times V & \longrightarrow & V \\ (u, v) & \mapsto & v + u \end{array} \tag{4.1.6}$$

dà a V una struttura di spazio affine su $\mathbb{K}$. Lo *spazio affine n-dimensionale standard su* $\mathbb{K}$ è

$$\mathbb{A}^n(\mathbb{K}) := \mathbb{K}^n \tag{4.1.7}$$

con la struttura di spazio affine appena definita.

Esempio 4.1.4. Si definisce un'applicazione

$$\begin{array}{ccc} \mathsf{V}(\mathbb{E}^3) \times \mathbb{E}^3 & \longrightarrow & \mathbb{E}^3 \\ (v, P) & \mapsto & P + v \end{array} \tag{4.1.8}$$

del tutto analoga all'applicazione in (4.1.1). Ripetiamola: se $\overrightarrow{PQ}$ è l'unico segmento orientato con punto iniziale P che rappresenta il vettore v, poniamo $P + v := Q$. Si dimostra che valgono (1), (2) e (3) della Definizione 4.1.1 come nell'analogo caso di $\mathbb{E}^2$.

Ovviamente possiamo ripetere per $\mathbb{E}^1$ (la retta della Geometria euclidea) quello che abbiamo fatto per $\mathbb{E}^2$ ed $\mathbb{E}^3$.

Osservazione 4.1.5. Sia $\mathbb{S}$ uno spazio affine con gruppo di traslazioni V. Se scegliamo un punto $P \in \mathbb{S}$ possiamo definire un'applicazione

$$\begin{array}{ccc} V & \longrightarrow & \mathbb{S} \\ v & \mapsto & P + v \end{array}$$

che è biunivoca per (1), (2) e (3) della Definizione 4.1.1. Quindi, scelto un punto di $\mathbb{S}$ possiamo identificare V con $\mathbb{S}$, ma cambiando il punto cambia l'identificazione. In un certo senso $\mathbb{S}$ è "come" V, ma senza l'elemento privilegiato 0; tutti i punti di $\mathbb{S}$ sono equivalenti.

Sia $\mathbb{S}$ uno spazio affine sullo spazio vettoriale V. Dato $v \in V$ definiamo la *traslazione*

$$\begin{array}{ccc} \mathbb{S} & \xrightarrow{\tau_v} & \mathbb{S} \\ P & \mapsto & P + v \end{array} \tag{4.1.9}$$

Osservazione 4.1.6. Le proprietà (1), (2), (3) della Definizione 4.1.1 equivalgono ripettivamente a

(I) $\tau_0 = Id_{\mathbb{S}}$,

(II) $\tau_w \circ \tau_v = \tau_{w+v}$ per ogni $v, w \in V$,

(III) dati $P, Q \in \mathbb{S}$ esiste un unico $v \in V$ tale che $\tau_v(P) = Q$.

Proposizione 4.1.7. *Sia $\mathbb{S}$ uno spazio affine sullo spazio vettoriale V. Sia $v \in V$.*

(a) L'applicazione τ_v è biunivoca.

(b) Se esiste $P \in \mathbb{S}$ tale che $\tau_v(P) = P$ allora $v = 0$. Equivalentemente: se $v \neq 0$ l'applicazione τ_v non ha punti fissi.

Dimostrazione. Per le proprietà (II) e (I) si ha $\tau_{-v} \circ \tau_v = \tau_v \circ \tau_{-v} = \tau_0 = Id_{\mathbb{S}}$ e quindi τ_v è biunivoca. Per dimostrare (2) supponiamo che $\tau_v(P) = P$. Per (I) abbiamo $\tau_0(P) = P$ e per (III) concludiamo che $v = 0$. □

In altre parole, dare una struttura di spazio affine all'insieme $\mathbb{S}$ (cioè un'applicazione (4.1.2) che soddisfa (1),(2) e (3) della Definizione 4.1.1) equivale a dare un omomorfismo

$$V \xrightarrow{\tau} \mathcal{S}(\mathbb{S}) \tag{4.1.10}$$

che in aggiunta gode della proprietà (III) dell'Osservazione 4.1.6. Inoltre

Terminologia 4.1.8. Uno spazio affine è determinato da una tripla $(\mathbb{S}, V, \tau)$ i cui elementi sono un insieme $\mathbb{S}$, uno spazio vettoriale V (il gruppo delle traslazioni) e un'applicazione τ come in (4.1.10) per cui valgono (I), (II) e (III) dell'Osservazione 4.1.6. Spesso lo denotiamo semplicemente con $\mathbb{S}$ e poniamo $\mathsf{V}(\mathbb{S}) = V$.

Ora dimostriamo un risultato che ci permette di trattare i vettori di $\mathsf{V}(\mathbb{S})$ come se fossero vettori geometrici di $\mathsf{V}(\mathbb{E}^m)$ per $m \in \{1, 2, 3\}$.

Definizione 4.1.9. Sia $\mathbb{S}$ uno spazio affine sullo spazio vettoriale V. Dati $P, Q \in \mathbb{S}$ il vettore $\overrightarrow{PQ} \in V$ è l'unico vettore tale che $P + \overrightarrow{PQ} = Q$.

Proposizione 4.1.10. *Sia $\mathbb{S}$ uno spazio affine e $P, Q, R \in \mathbb{S}$. Allora*

$$\overrightarrow{PP} = 0, \tag{4.1.11}$$

$$\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}, \tag{4.1.12}$$

$$\overrightarrow{PQ} = -\overrightarrow{QP}. \tag{4.1.13}$$

Dimostrazione. Siccome $P + 0 = P$ e $P + \overrightarrow{PP} = P$, si ha $0 = \overrightarrow{PP}$. Abbiamo

$$P + (\overrightarrow{PQ} + \overrightarrow{QR}) = (P + \overrightarrow{PQ}) + \overrightarrow{QR} = Q + \overrightarrow{QR} = R. \tag{4.1.14}$$

D'altra parte per il punto (3) della Definizione 4.1.1 esiste un unico vettore v tale che $P + v = R$ e per definizione è $\overrightarrow{PR}$; segue (4.1.12). Ora dimostriamo (4.1.13). Per (4.1.11) e (4.1.12),

$$0 = \overrightarrow{PP} = \overrightarrow{PQ} + \overrightarrow{QP}, \tag{4.1.15}$$

e quindi $\overrightarrow{PQ} = -\overrightarrow{QP}$. □

Definizione 4.1.11. La *dimensione* di uno spazio affine $\mathbb{S}$ è definita come

$$\dim \mathbb{S} := \dim \mathsf{V}(\mathbb{S}). \tag{4.1.16}$$

La definizione di dimensione di uno spazio affine è sensata: se $m \in \{1, 2, 3\}$, la dimensione dello spazio affine $\mathbb{E}^m$ è m. Una *retta affine* (o semplicemente retta) è uno spazio affine di dimensione 1, un *piano affine* (o semplicemente piano) è uno spazio affine di dimensione 2, un *solido affine* (o semplicemente solido) è uno spazio affine di dimensione 3.

Osservazione 4.1.12. La definizione di spazio affine riassume le proprietà del piano (o dello spazio) euclideo che ci permettono di parlare di rette, piani. Infatti notiamo che una retta in $\mathbb{E}^m$ è data dai punti $P_0 + sv_1$, dove $P_0 \in \mathbb{E}^m$, $v_1 \in \mathsf{V}(\mathbb{E}^m)$ è un vettore non nullo, e s è un numero reale arbitrario. Analogamente, un piano in $\mathbb{E}^m$ è dato dai punti $P_0 + sv_1 + tv_2$, dove $P_0 \in \mathbb{E}^m$, $v_1, v_2 \in \mathsf{V}(\mathbb{E}^m)$ sono vettori linearmente indipendenti, e s, t sono numeri reali arbitrari. Il punto è che analoga definizione si può dare per uno spazio affine qualsiasi (vedi la Definizione 4.3.1), e queste "rette o piani astratti" si comportano come le rette o i piani di $\mathbb{E}^m$. In uno spazio affine generale rimane vero il classico assioma di Euclide: *per due punti distinti passa una e una sola retta.* Analogamente, vale il classico assioma delle parallele: *per un punto esterno a una retta data passa una una sola retta parallela alla retta data.*

4.2 Combinazioni lineari di punti

Sia $\mathbb{S}$ uno spazio affine. Non esiste un modo sensato di definire la combinazione lineare $\lambda P + \mu Q$ di punti $P, Q \in \mathbb{S}$ se $\lambda, \mu \in \mathbb{K}$ sono arbitrari: pensate all'Esempio 4.1.2 nel caso in cui $b \neq 0$: se $X, Z \in W$ e $\lambda + \mu \neq 1$ allora $\lambda X + \mu Z \notin W$. In generale si può dare senso alle combinazioni lineari $\lambda P + \mu Q$ nel caso in cui $\lambda + \mu = 1$.

Lemma 4.2.1. *Sia $\mathbb{S}$ uno spazio affine su $\mathbb{K}$ e $P_0, \ldots, P_d, Q, R \in \mathbb{S}$. Siano $\lambda_0, \ldots, \lambda_d \in \mathbb{K}$ tali che*

$$\sum_{i=0}^{d} \lambda_i = 1. \tag{4.2.1}$$

Allora

$$Q + \sum_{i=0}^{d} \lambda_i \overrightarrow{QP_i} = R + \sum_{i=0}^{d} \lambda_i \overrightarrow{RP_i}. \tag{4.2.2}$$

Dimostrazione. Sottraendo il vettore $\sum_{i=0}^{d} \lambda_i \overrightarrow{RP_i}$ ad ambo i membri di (4.2.2) vediamo che è sufficiente verificare che

$$Q + \sum_{i=1}^{d} \lambda_i (\overrightarrow{QP_i} - \overrightarrow{RP_i}) = R. \tag{4.2.3}$$

Applicando la Proposizione 4.1.10 vediamo che (4.2.3) equivale a

$$Q + \sum_{i=0}^{d} \lambda_i \overrightarrow{QR} = R. \tag{4.2.4}$$

L'equazione (4.2.4) vale perchè per ipotesi vale (4.2.1). □

Il Lemma 4.2.1 ci permette di dare la seguente definizione.

Definizione 4.2.2. Sia $\mathbb{S}$ uno spazio affine su $\mathbb{K}$ e $P_0, \ldots, P_d \in \mathbb{S}$. Siano $\lambda_0, \ldots, \lambda_d \in \mathbb{K}$ tali che valga (4.2.1). La combinazione lineare di $P_0, \ldots, P_d \in \mathbb{S}$ con pesi $\lambda_0, \ldots, \lambda_d$ è

$$\sum_{i=0}^{d} \lambda_i P_i := Q + \sum_{i=0}^{d} \lambda_i \overrightarrow{QP_i} \tag{4.2.5}$$

dove $Q \in \mathbb{S}$ è arbitrario. (La definizione è sensata grazie al Lemma 4.2.1.)

Esempio 4.2.3. Sia $W \subset \mathbb{K}^n$ lo spazio affine dell'Esempio 4.1.2. Siano $X, Y \in W$. Dati $\lambda, \mu \in \mathbb{K}$ tali che $\lambda + \mu = 1$, la combinazione lineare di X, Y con pesi λ, μ è uguale alla combinazione lineare di vettori

$$\lambda X + \mu Y. \tag{4.2.6}$$

Notate che (4.2.6) ha senso anche se $\lambda + \mu \neq 1$, ma non apparterrà a W se $b \neq 0$,.

Esempio 4.2.4. Consideriamo lo spazio affine $\mathbb{E}^m$ per $m \in \{1, 2, 3\}$. Siano $P, Q \in \mathbb{E}^m$. Se $P \neq Q$ le combinazioni lineari di P e Q sono i punti sulla retta per P e Q. Se $P = Q$ le combinazioni lineari di P e Q sono tutte uguali a P.

4.3 Sottospazi affini

Definizione 4.3.1. Sia $\mathbb{S}$ uno spazio affine. Un sottoinsieme $\mathbb{T} \subset \mathbb{S}$ è un sottospazio affine se esistono $P_0 \in \mathbb{S}$ e un sottospazio vettoriale $W \subset \mathsf{V}(\mathbb{S})$ tali che

$$\mathbb{T} = P_0 + W := \{P_0 + w \mid w \in W\}. \tag{4.3.1}$$

Esempio 4.3.2. Ricordiamo che uno spazio vettoriale V può essere visto come spazio affine con gruppo delle traslazioni V stesso, vedi l'Esempio 4.1.3. Un sottospazio vettoriale $W \subset V$ è un sottospazio affine, ma non vale il viceversa. Infatti $\mathbb{T} \subset V$ è un sottospazio affine se è dato da $\mathbb{T} = v_0 + W$, dove $W \subset V$ è un sottospazio vettoriale, e tale $\mathbb{T}$ è un sottospazio vettoriale solo se $v_0 \in W$ (se $v_0 \in W$, allora $v_0 + W = W$, mentre se $v_0 \notin W$ allora $0 \notin (v_0 + W)$ e quindi $v_0 + W$ non è un sottospazio vettoriale). Per esempio ogni punto $v_0 \in V$ (se fossimo pedanti diremmo "ogni sottoinsieme costituito di un solo punto") è un sottospazio affine di V, ma è un sottospazio solo se $v_0 = 0$.

Proposizione 4.3.3. *Siano $\mathbb{S}$ uno spazio affine e $\mathbb{T} \subset \mathbb{S}$ un sottospazio affine dato da $\mathbb{T} = P_0 + W$, dove $W \subset \mathsf{V}(\mathbb{S})$ è un sottospazio vettoriale. Dato $P \in \mathbb{T}$, si ha*

$$\{\overrightarrow{PQ} \mid Q \in \mathbb{T}\} = W. \tag{4.3.2}$$

Dimostrazione. Esiste $v \in W$ tale che $P = P_0 + v$. Analogamente, se $Q \in \mathbb{T}$ esiste $u \in W$ tale che $Q = P_0 + u$. Siccome

$$P + (u - v) = (P_0 + v) + (u - v) = P_0 + u = Q,$$

abbiamo

$$\overrightarrow{PQ} = u - v. \tag{4.3.3}$$

Ma siccome W è un sottospazio (vettoriale) di $\mathsf{V}(\mathbb{S})$, anche $(u - v)$ appartiene a W, e perciò $\overrightarrow{PQ} \in W$. Questo dimostra che l'insieme di sinistra di (4.3.2) è contenuto nell'insieme di destra. Ora dimostriamo che l'insieme di destra di (4.3.2) è contenuto nell'insieme di sinistra. Se $w \in W$, allora $u := (v + w) \in W$ (perchè $v \in W$ e W è un sottospazio vettoriale di $\mathsf{V}(\mathbb{S})$), e quindi $Q = P_0 + u$ è un elemento di $\mathbb{T}$. Per (4.3.3) abbiamo $w = \overrightarrow{PQ}$, e quindi w è un elemento dell'insieme di sinistra. □

La Proposizione 4.3.3 mostra che abbiamo un'applicazione

$$\begin{array}{ccc} W \times \mathbb{T} & \longrightarrow & \mathbb{T} \\ (v, P) & \mapsto & P + v \end{array} \tag{4.3.4}$$

Il seguente risultato segue immediatamente da quanto già detto.

Proposizione 4.3.4. *Siano $\mathbb{S}$ uno spazio affine e $\mathbb{T} \subset \mathbb{S}$ un sottospazio affine dato da* (4.3.1)*. Allora l'applicazione* (4.3.4) *dà a $\mathbb{T}$ una struttura di spazio affine sul sottospazio vettoriale $W \subset \mathsf{V}(\mathbb{S})$.*

Il sottospazio vettoriale $W \subset \mathsf{V}(\mathbb{S})$ che appare in (4.3.1) è uguale al gruppo delle traslazioni di $\mathbb{T}$, ovvero $\mathsf{V}(\mathbb{T})$. Viene anche chiamato *giacitura di* $\mathbb{T}$.

È naturale estendere a spazi affini qualsiasi la definizione di parallelismo tra sottospazi affini di $\mathbb{E}^2$ o $\mathbb{E}^3$, nel seguente modo.

Definizione 4.3.5. Sia $\mathbb{S}$ uno spazio affine su uno spazio vettoriale V. Due sottospazi affini $\mathbb{T}_1, \mathbb{T}_2 \subset \mathbb{S}$ sono *paralleli* se $\mathsf{V}(\mathbb{T}_1) \subset \mathsf{V}(\mathbb{T}_2)$, oppure $\mathsf{V}(\mathbb{T}_1) \supset \mathsf{V}(\mathbb{T}_2)$.

Osservazione 4.3.6. Siano $\mathbb{T}_1, \mathbb{T}_2 \subset \mathbb{S}$ sottospazi affini paralleli. Se $\mathbb{T}_1 \cap \mathbb{T}_2$ non è vuoto, allora $\mathbb{T}_1 \subset \mathbb{T}_2$ oppure $\mathbb{T}_1 \supset \mathbb{T}_2$. Infatti supponiamo che $P \in \mathbb{T}_1 \cap \mathbb{T}_2$ e che $\mathsf{V}(\mathbb{T}_1) \subset \mathsf{V}(\mathbb{T}_2)$. Allora

$$\mathbb{T}_1 = P + \mathsf{V}(\mathbb{T}_1) \subset P + \mathsf{V}(\mathbb{T}_2) = \mathbb{T}_2.$$

D'altra parte $\mathbb{T}_1 \cap \mathbb{T}_2$ vuoto implica che $\mathbb{T}_1, \mathbb{T}_2$ sono paralleli solo se $\dim \mathbb{S} \leqslant 2$. Per esempio le rette $\mathbb{T}_1, \mathbb{T}_2 \subset \mathbb{A}^3(\mathbb{K})$ date da

$$\mathbb{T}_1 := \{(x, 0, 0) \mid x \in \mathbb{K}\}, \quad \mathbb{T}_2 := \{(x, 0, 1 + x) \mid x \in \mathbb{K}\}$$

non hanno punti in comune ma *non* sono parallele perchè

$$\mathsf{V}(\mathbb{T}_1) = \langle (1, 0, 0) \rangle, \quad \mathsf{V}(\mathbb{T}_2) = \langle (1, 0, 1) \rangle.$$

Dimostriamo invece che rette di un piano che non hanno punti in comune sono parallele. Equivalentemente, supponiamo che $\mathbb{S}$ sia un piano affine e che $\mathbb{T}_1, \mathbb{T}_2 \subset \mathbb{S}$ siano rette che non sono parallele, e dimostriamo che $\mathbb{T}_1 \cap \mathbb{T}_2 \neq \varnothing$. Siccome $\mathbb{T}_i$ sono rette, $\mathsf{V}(\mathbb{T}_i) = \langle v_i \rangle$ dove $v_i \neq 0$, e siccome $\mathbb{T}_1, \mathbb{T}_2$ non sono parallele, $\{v_1, v_2\}$ è una base di $\mathsf{V}(\mathbb{S})$. Siano $P_i \in \mathbb{T}_i$ per $i \in \{1, 2\}$. Siccome $\{v_1, v_2\}$ è una base di $\mathsf{V}(\mathbb{S})$, esistono $\mu_1, \mu_2 \in \mathbb{K}$ tali che

$$\overrightarrow{P_1 P_2} = \mu_1 v_1 + \mu_2 v_2.$$

Quindi

$$P_1 + \mu_1 v_1 = (P_1 + \overrightarrow{P_1 P_2}) - \mu_2 v_2 = P_2 - \mu_2 v_2.$$

Ma $(P_1 + \mu_1 v_1) \in \mathbb{T}_1$ e $(P_2 - \mu_2 v_2) \in \mathbb{T}_2$, e perciò $\mathbb{T}_1 \cap \mathbb{T}_2 \neq \varnothing$.

Proposizione 4.3.7. *Sia $\mathbb{S}$ uno spazio affine e $\mathbb{T}_i \subset \mathbb{S}$ una collezione di sottospazi affini di $\mathbb{S}$ indicizzati da un insieme I. Se l'intersezione $\bigcap_{i \in I} \mathbb{T}_i$ non è vuota, allora è un sottospazio affine di $\mathbb{S}$, con gruppo delle traslazioni l'intersezione $\bigcap_{i \in I} \mathsf{V}(\mathbb{T}_i)$.*

Dimostrazione. Sia $P_0 \in \bigcap_{i \in I} \mathbb{T}_i$ (P_0 esiste perchè per ipotesi l'intersezione $\bigcap_{i \in I} \mathbb{T}_i$ non è vuota). Dobbiamo dimostrare che

$$\bigcap_{i \in I} \mathbb{T}_i = P_0 + \bigcap_{i \in I} \mathsf{V}(\mathbb{T}_i) \tag{4.3.5}$$

È chiaro che il membro di destra di (4.3.5) è contenuto nel membro di sinistra, perchè se $v \in \bigcap_{i \in I} \mathsf{V}(\mathbb{T}_i)$, allora $(P_0 + v) \in \mathbb{T}_i$ per ogni $i \in I$. Per dimostrare che che il membro di sinistra di (4.3.5) è contenuto nel membro di destra, supponiamo che $Q \in \bigcap_{i \in I} \mathbb{T}_i$, e quindi per ogni $i \in I$ esiste $v_i \in \mathsf{V}(\mathbb{T}_i)$ tale che $Q = P_0 + v_i$. Se $i, j \in I$, allora $P_0 + v_i = Q = P_0 + v_j$ e quindi

$$P_0 + (v_i - v_j) = (P_0 + v_i) + (-v_j) = (P_0 + v_j) + (-v_j) = P_0 + (v_j - v_j) = P_0.$$

Segue che $v_i = v_j$, cioè $Q = P_0 + v$ dove $v \in \mathsf{V}(\mathbb{T}_i)$ per ogni $i \in I$ e quindi Q è un elemento del membro di destra di (4.3.5). □

Definizione 4.3.8. Siano $\mathbb{S}$ uno spazio affine e Z un sottoinsieme *non vuoto* di $\mathbb{S}$. Il *sottospazio affine di $\mathbb{S}$ generato da Z* è l'intersezione di tutti i sottospazi affini di $\mathbb{S}$ contenenti Z (notate che $\mathbb{S}$ è un sottospazio affine contenente Z, quindi stiamo intersecando gli elementi di una famiglia non vuota di sottospazi affini di $\mathbb{S}$) - lo denoteremo con $\langle Z \rangle$ o $\mathrm{Span}(Z)$.

Proposizione 4.3.9. *Siano $\mathbb{S}$ uno spazio affine e Z un sottoinsieme non vuoto di $\mathbb{S}$. Allora $\langle Z \rangle$ è un sottospazio affine contenente Z e contenuto in ogni sottospazio affine che contiene Z.*

Dimostrazione. Siccome Z è contenuto in $\langle Z \rangle$ e Z non è vuoto, $\langle Z \rangle$ non è vuoto, e per la Proposizione 4.3.7 segue che è un sottospazio affine di $\mathbb{S}$. Per costruzione è contenuto in ogni sottospazio affine che contiene Z. □

Se Z è finito, diciamo $Z = \{P_0, \dots, P_d\}$, poniamo

$$\langle P_0, \dots, P_d \rangle := \langle \{P_0, \dots, P_d\} \rangle. \tag{4.3.6}$$

Dimostriamo che

$$\langle P_0, \dots, P_d \rangle = P_0 + \langle \overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_d} \rangle. \tag{4.3.7}$$

Infatti il membro di destra di (4.3.7) è un sottospazio affine di $\mathbb{S}$ contenente $P_0, \dots, P_d$ e quindi è sufficiente dimostrare che ogni sottospazio affine $\mathbb{T} \subset \mathbb{S}$ contenente $P_0, \dots, P_d$ contiene il membro di destra di (4.3.7). Se $\mathbb{T} \subset \mathbb{S}$ è un sottospazio affine contenente $P_0, \dots, P_d$, allora per la Proposizione 4.3.3 $\mathsf{V}(\mathbb{T})$ contiene $\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_d}$, e quindi $\mathbb{T}$ contiene il membro di destra di (4.3.7).

Proposizione 4.3.10. *Siano $\mathbb{S}$ uno spazio affine e $P_0, \dots, P_d \in \mathbb{S}$. Allora*

$$\langle P_0, \dots, P_d \rangle = \left\{ \sum_{i=0}^{d} \lambda_i P_i \mid \lambda_i \in K, \quad \sum_{i=0}^{d} \lambda_i = 1 \right\}. \tag{4.3.8}$$

Dimostrazione. Per (4.3.7) abbiamo

$$\langle P_0, \dots, P_d \rangle = \left\{ P_0 + \sum_{i=1}^{d} \mu_i \overrightarrow{P_0P_i} \mid \mu_i \in K \right\} = \left\{ (1 - \sum_{i=1}^{d} \mu_i) P_0 + \sum_{i=1}^{d} \mu_i P_i \mid \mu_i \in K \right\}. \tag{4.3.9}$$

Ponendo $\lambda_0 = (1 - \sum_{i=1}^{d} \mu_i)$ e $\lambda_i = \mu_i$ per $i \in \{1, \dots, d\}$ vediamo che vale (4.3.8). □

Sia $\mathbb{T} \subset \mathbb{S}$ un sottospazio affine. Siccome $\mathbb{T}$ è uno spazio affine è ben definita la sua dimensione $\dim \mathbb{T}$; questo fatto ci permette di dare la nozione di dipendenza/indipendenza lineare di punti $P_0, \dots, P_d \in \mathbb{S}$. Osserviamo che per (4.3.7) si ha che

$$\dim \langle P_0, \dots, P_d \rangle \leqslant d. \tag{4.3.10}$$

Definizione 4.3.11. Sia $\mathbb{S}$ uno spazio affine. Una lista di punti $P_0, \dots, P_d \in \mathbb{S}$ è *linearmente dipendente* se $\dim \langle P_0, \dots, P_d \rangle < d$, è *linearmente indipendente* se $\dim \langle P_0, \dots, P_d \rangle = d$.

Come per i vettori di uno spazio vettoriale useremo l'espressione "i punti $P_0, \dots, P_d \in \mathbb{S}$ sono linearmente dipendenti/indipendenti"nonostante la dipendenza/indipendenza lineare sia una proprietà delle sequenze di punti, *non* dei singoli punti della sequenza. La (facile) dimostrazione del seguente lemma è lasciata al lettore.

Lemma 4.3.12. *Sia $\mathbb{S}$ uno spazio affine. Una lista di punti $P_0, \dots, P_d \in \mathbb{S}$ è linearmente indipendente se e solo se è iniettiva l'applicazione*

$$\begin{array}{ccc} \left\{ (\lambda_0, \dots, \lambda_d) \in K^{d+1} \mid \sum_{i=0}^{d} \lambda_i = 1 \right\} & \longrightarrow & \langle P_0, \dots, P_d \rangle \\ (\lambda_0, \dots, \lambda_d) & \mapsto & \sum_{i=0}^{d} \lambda_i P_i \end{array} \tag{4.3.11}$$

Esempio 4.3.13. Consideriamo l'Esempio 4.1.3, cioè lo spazio affine è uno spazio vettoriale V. Siano $v_0, \dots, v_d \in V$. Allora i punti $v_0, \dots, v_d$ sono linearmente indipendenti nello *spazio affine* V se e solo se i vettori $(v_1 - v_0), \dots, (v_d - v_0)$ sono linearmente indipendenti nello *spazio vettoriale* V. Segue che se i vettori $v_0, \dots, v_d$ sono linearmente indipendenti nello spazio vettoriale V allora i punti $v_0, \dots, v_d$ sono linearmente indipendenti ma *non* è vero il viceversa. Per esempio se $v \in V$ è un vettore non nullo allora i *punti* $0, v$ sono linearmente indipendenti ma i *vettori* $0, v$ non lo sono.

4.4 Applicazioni affini

Definizione, esempi e prime proprietà

Definizione 4.4.1. Un'applicazione $F\colon \mathbb{S} \to \mathbb{T}$ tra spazi affini sullo stesso campo è *affine* se esiste un'applicazione lineare $f\colon \mathsf{V}(\mathbb{S}) \to \mathsf{V}(\mathbb{T})$ tale che

$$\overrightarrow{F(P)F(Q)} = f(\overrightarrow{PQ}) \quad \forall P, Q \in \mathbb{S}. \tag{4.4.1}$$

In altre parole F manda segmenti orientati equipollenti in segmenti orientati equipollenti, e l'applicazione indotta $\mathsf{V}(\mathbb{S}) \to \mathsf{V}(\mathbb{T})$ è lineare.

Definizione 4.4.2. L'applicazione lineare f in (4.4.1) (che è univocamente determinato da F) è *associata* a F ed è denotata $\mathsf{V}(F)$.

Esempio 4.4.3. Sia $\mathbb{S}$ uno spazio affine. La traslazione $\tau_v\colon \mathbb{S} \to \mathbb{S}$ determinata da $v \in \mathsf{V}(\mathbb{S})$ è un'applicazione affine con applicazione lineare associata $\mathrm{Id}_{\mathsf{V}(\mathbb{S})}$. Infatti

$$\overrightarrow{\tau_v(P)\tau_v(Q)} = \overrightarrow{(P+v)(Q+v)},$$

e $\overrightarrow{(P+v)(Q+v)} = \overrightarrow{PQ}$ perchè

$$(P+v) + \overrightarrow{PQ} = P + (v + \overrightarrow{PQ}) = (P + \overrightarrow{PQ}) + v = Q + v.$$

Vale il viceversa: se $F\colon \mathbb{S} \to \mathbb{S}$ è un'applicazione affine tale che $\mathsf{V}(F) = \mathrm{Id}_{\mathsf{V}(\mathbb{S})}$, allora esiste $v \in \mathsf{V}(\mathbb{S})$ tale che $F = \tau_v$. Infatti sia $P_0 \in \mathbb{S}$ e poniamo $Q_0 := F(P_0)$. Dobbiamo dimostrare che $F = \tau_{\overrightarrow{P_0Q_0}}$, cioè che dato $P \in \mathbb{S}$, e posto $Q = F(P)$, si ha $\overrightarrow{PQ} = \overrightarrow{P_0Q_0}$. Questo vale perchè

$$\overrightarrow{PQ} = \overrightarrow{PP_0} + \overrightarrow{P_0Q_0} + \overrightarrow{Q_0Q} = \overrightarrow{PP_0} + \overrightarrow{P_0Q_0} + \overrightarrow{F(P_0)F(P)} = \overrightarrow{PP_0} + \overrightarrow{P_0Q_0} + \overrightarrow{P_0P} = \overrightarrow{P_0Q_0}.$$

Esempio 4.4.4. Siano $A \in M_{m,n}(\mathbb{K})$ e $B \in M_{m,1}$: l'applicazione

$$\begin{array}{ccc} \mathbb{A}^n(\mathbb{K}) & \xrightarrow{F} & \mathbb{A}^m(\mathbb{K}) \\ X & \mapsto & A \cdot X + B \end{array} \tag{4.4.2}$$

è affine, con applicazione lineare associata L_A. Infatti se $X, Y \in \mathbb{A}^n(\mathbb{K})$ si ha

$$\overrightarrow{F(X)F(Y)} = \overrightarrow{(A \cdot X + B)(A \cdot Y + B)} = A \cdot Y + B - (A \cdot X + B) = A \cdot (Y - X) = L_A(\overrightarrow{XY}).$$

(Attenzione: usiamo lo stesso simbolo per denotare punti di $\mathbb{A}^n(\mathbb{K}), \mathbb{A}^m(\mathbb{K})$ e vettori di $\mathbb{K}^n, \mathbb{K}^m$.)

Viceversa, se $F\colon \mathbb{A}^n(\mathbb{K}) \to \mathbb{A}^m(\mathbb{K})$ è un'applicazione affine, esistono $A \in M_{m,n}(\mathbb{K})$ e $B \in M_{m,1}$ tali che F sia data da (4.4.2). Infatti sia $f\colon \mathbb{K}^n \to \mathbb{K}^m$ l'applicazione lineare associata: per la Proposizione 3.5.3 esiste $A \in M_{m,n}(\mathbb{K})$ tale che $f = L_A$. Sia $B := F(0)$. Dato $X \in \mathbb{A}^n(\mathbb{K})$, l'equazione (4.4.1) con $P = 0$ e $Q = X$ dà che

$$\overrightarrow{F(0)F(X)} = L_A(\overrightarrow{0X}) = A \cdot X,$$

e quindi

$$F(X) = F(0) + \overrightarrow{F(0)F(X)} = B + A \cdot X.$$

Esempio 4.4.5. Siano $\mathbb{S}$ uno spazio affine di dimensione finita e $\mathbb{T} \subset \mathbb{S}$ un sottospazio affine. Supponiamo che $\mathsf{D} \subset \mathsf{V}(\mathbb{S})$ sia un sottospazio vettoriale e che si abbia la decomposizione in somma diretta (vedi l'Esempio 3.2.11)

$$\mathsf{V}(\mathbb{S}) = \mathsf{D} \oplus \mathsf{V}(\mathbb{T}). \tag{4.4.3}$$

Ricordiamo: ciò significa che l'applicazione

$$\begin{array}{ccc} \mathsf{D} \oplus \mathsf{V}(\mathbb{T}) & \xrightarrow{\varphi} & \mathsf{V}(\mathbb{S}) \\ (u, w) & \mapsto & u + w \end{array} \tag{4.4.4}$$

è un isomorfismo, cioè che $D \cap \mathsf{V}(\mathbb{T}) = \{0\}$ e che l'applicazione è suriettiva. Se $P \in \mathbb{S}$ il sottospazio

$$\mathbb{L}_P := P + \mathsf{D}$$

ha un unico punto d'intersezione con $\mathbb{T}$. Per dimostrarlo notiamo innanzitutto che se esistessero due punti (almeno) in $\mathbb{L}_P \cap \mathbb{T}$, diciamo Q_0 e Q_1, allora avremmo che $0 \neq \overrightarrow{Q_0Q_1} \in \mathsf{D} \cap \mathsf{V}(\mathbb{T})$, e questo contraddice l'ipotesi che vale (4.4.3). Quindi rimane da dimostrare che $\mathbb{L}_P \cap \mathbb{T}$ è non vuoto; la dimostrazione è del tutto simile a quella data nel caso in cui $\dim \mathbb{S} = 2$ e $\dim \mathbb{T} = 1$ data nell'Osservazione 4.3.6. Quindi possiamo definire la *proiezione su* $\mathbb{T}$ *parallela a* D così:

$$\begin{array}{ccc} \mathbb{S} & \xrightarrow{\pi} & \mathbb{T} \\ P & \mapsto & \mathbb{L}_P \cap \mathbb{T} \end{array}$$

Sia φ è l'isomorfismo in (4.4.4) e definiamo $f\colon \mathsf{V}(\mathbb{S}) \to \mathsf{V}(\mathbb{T})$ come la composizione di $\varphi^{-1}\colon \mathsf{V}(\mathbb{S}) \to \mathsf{D} \oplus \mathsf{V}(\mathbb{T})$ e la proiezione $\mathsf{D} \oplus \mathsf{V}(\mathbb{T}) \to \mathsf{V}(\mathbb{T})$. Chiaramente f è lineare. La proiezione su $\mathbb{T}$ parallela a D è un'applicazione affine con applicazione lineare associata f, lasciamo al lettore la verifica dettagliata di quest'affermazione.

Osservazione 4.4.6. Siano $\mathbb{S}, \mathbb{T}$ spazi affini sullo stesso campo, e sia $F\colon \mathbb{S} \to \mathbb{T}$ un'applicazione affine. Scegliamo $P_0 \in \mathbb{S}$. Allora per ogni P si ha

$$F(P) = F(P_0) + \mathsf{V}(F)(\overrightarrow{P_0P}), \tag{4.4.5}$$

perchè per definizione $\mathsf{V}(F)(\overrightarrow{P_0P}) = \overrightarrow{F(P_0)F(P)}$. Questo mostra che F è determinata da $\mathsf{V}(F)$ e dall'immagine di un punto.

La proposizione che segue mostra dà un viceversa dell'Osservazione 4.4.6.

Proposizione 4.4.7. *Siano* $\mathbb{S}, \mathbb{T}$ *spazi affini sullo stesso campo,* $P_0 \in \mathbb{S}$, $Q_0 \in \mathbb{T}$ *e* $f\colon \mathsf{V}(\mathbb{S}) \to \mathsf{V}(\mathbb{T})$ *un'applicazione lineare. L'applicazione* $F\colon \mathbb{S} \to \mathbb{T}$ *definita da*

$$F(P) = Q_0 + f(\overrightarrow{P_0P}) \tag{4.4.6}$$

è affine, con applicazione lineare associata f*.*

Dimostrazione. Siano $P_i \in \mathbb{S}$ per $i \in \{1, 2\}$ e $v_i := \overrightarrow{P_0P_i}$. Allora

$$\begin{aligned} \overrightarrow{F(P_1)F(P_2)} = \overrightarrow{(F(P_0) + f(v_1))(F(P_0) + f(v_2))} = \\ = f(v_2) - f(v_1) = f(v_2 - v_1) = f(\overrightarrow{(P_0 + v_1)(P_0 + v_2)}) = f(\overrightarrow{P_1P_2}). \end{aligned}$$

(La seconda uguaglianaza vale perchè $(F(P_0) + f(v_1)) + f(v_2) - f(v_1) = (F(P_0) + f(v_2))$. Un analogo commento si applica alla quarta uguaglianza.) □

Proposizione 4.4.8. *Siano* $\mathbb{S}$ *e* $\mathbb{T}$ *spazi affini sullo stesso campo* $\mathbb{K}$ *e* $F\colon \mathbb{S} \to \mathbb{T}$ *un'applicazione affine.*

(a) Se $\mathbb{D} \subset \mathbb{S}$ *è un sottospazio affine, allora* $F(\mathbb{D})$ *è un sottospazio affine.*

(b) Se $\mathbb{E} \subset \mathbb{T}$ *è un sottospazio affine, allora* $F^{-1}(\mathbb{E})$ *o è vuoto o è un sottospazio affine.*

(c) Siano $P_0, \dots, P_d \in \mathbb{S}$ *e* $\lambda_0, \dots, \lambda_d \in \mathbb{K}$ *tali che* $\sum_{i=0}^{d} \lambda_i = 1$*. Allora*

$$F(\sum_{i=0}^{d} \lambda_i P_i) = \sum_{i=0}^{d} \lambda_i F(P_i). \tag{4.4.7}$$

Dimostrazione. Poniamo $f := \mathsf{V}(F)$.
(a): Sia $P_0 \in \mathbb{D}$. Per l'equazione (4.4.1) abbiamo che

$$F(\mathbb{D}) = F(P_0) + f(\mathsf{V}(\mathbb{D})). \tag{4.4.8}$$

Siccome f è lineare $f(\mathsf{V}(\mathbb{D}))$ è un sottospazio vettoriale di $\mathsf{V}(\mathbb{T})$ e quindi $F(\mathbb{D})$ è un sottospazio affine.
(b): Supponiamo che $F^{-1}(\mathbb{E})$ non sia vuota e sia $P_0 \in F^{-1}(\mathbb{E})$. Per l'equazione (4.4.1) abbiamo

$$F^{-1}(\mathbb{E}) = P_0 + f^{-1}(\mathsf{V}(\mathbb{E})). \tag{4.4.9}$$

Siccome f è lineare $f^{-1}(\mathsf{V}(\mathbb{E}))$ è un sottospazio vettoriale di V e quindi $F^{-1}(\mathbb{E})$ è un sottospazio affine.
(c): Sia $Q \in \mathbb{S}$. Abbiamo

$$F(\sum_{i=0}^{d} \lambda_i P_i) = F(Q + \sum_{i=0}^{d} \lambda_i \overrightarrow{QP_i}) = F(Q) + \sum_{i=0}^{d} \lambda_i f(\overrightarrow{QP_i}) = F(Q) + \sum_{i=0}^{d} \lambda_i \overrightarrow{F(Q)F(P_i)} = \sum_{i=0}^{d} \lambda_i F(P_i).$$

□

Equazioni cartesiane di sottospazi affini

Definizione 4.4.9. Sia $\mathbb{S}$ uno spazio affine sul campo $\mathbb{K}$. Una *funzione affine* è un'applicazione affine $F\colon \mathbb{S} \to \mathbb{A}^1(\mathbb{K})$.

In altre parole esistono un'applicazione lineare $f\colon \mathsf{V}(\mathbb{S}) \to \mathbb{K}$, $P_0 \in \mathbb{S}$ e $b \in \mathbb{A}^1(\mathbb{K})$ tali che

$$F(P) = b + f(\overrightarrow{P_0P}), \qquad \forall v \in \mathsf{V}(\mathbb{S}). \tag{4.4.10}$$

Siano $\mathbb{S}$ uno spazio affine su $\mathbb{K}$ e $\mathbb{T} \subset \mathbb{S}$ un sottospazio affine. Siano $F_1, \dots, F_d\colon \mathbb{S} \to \mathbb{K}$ funzioni affini: diciamo che $0 = F_1 = \dots = F_d$ sono *equazioni cartesiane* di $\mathbb{T}$ se

$$\mathbb{T} = \{P \in \mathbb{S} \mid 0 = F_1(P) = \dots = F_d(P)\}. \tag{4.4.11}$$

Proposizione 4.4.10. *Sia $\mathbb{S}$ uno spazio affine su $\mathbb{K}$ di dimensione finita (cioè tale che $\mathsf{V}(\mathbb{S})$ sia finitamente generato). Sia $\mathbb{T} \subset \mathbb{S}$ un sottospazio affine e $m := \dim \mathbb{S} - \dim \mathbb{T}$. Esistono m funzioni affini che danno equazioni cartesiane di $\mathbb{T}$.*

Dimostrazione. Per definizione esistono $P_0 \in \mathbb{S}$ e un sottospazio vettoriale $W \subset \mathsf{V}(\mathbb{S})$ tali che

$$\mathbb{T} = \{P_0 + w \mid w \in W\}. \tag{4.4.12}$$

Inoltre $W = \mathsf{V}(\mathbb{T})$ e quindi $m = \dim \mathsf{V}(\mathbb{S}) - \dim W$. Per l'Esercizio 3.18, esistono applicazioni lineari $g_1, \dots, g_m\colon \mathsf{V}(\mathbb{S}) \to \mathbb{K}$ tali che

$$W = \{v \in \mathsf{V}(\mathbb{S}) \mid 0 = g_1(v) = \dots = g_d(v)\}. \tag{4.4.13}$$

Definiamo $F_1, \dots, F_m\colon \mathbb{S} \to \mathbb{K}$ così:

$$F_i(P) := g_i(\overrightarrow{P_0P}).$$

È chiaro che $F_i(P) = 0$ per ogni $i \in \{1, \dots, m\}$ se e solo se $P \in \mathbb{T}$. Rimane da dimostrare che ciascuna F_i è una funzione *affine*. Dimostriamo che F_i è affine con applicazione lineare associata uguale a g_i. Siano $P, Q \in \mathbb{S}$; abbiamo

$$\overrightarrow{F_i(P)F_i(Q)} = \overrightarrow{g_i(\overrightarrow{P_0P})g_i(\overrightarrow{P_0Q})} = g_i(\overrightarrow{P_0Q}) - g_i(\overrightarrow{P_0P}) = g_i(\overrightarrow{P_0Q} - \overrightarrow{P_0P}) = g_i(\overrightarrow{PQ}).$$

□

Osservazione 4.4.11. Sia $\mathbb{S}$ uno spazio affine su $\mathbb{K}$ e $F_1, \dots, F_d \colon \mathbb{S} \to \mathbb{K}$ equazioni cartesiane di un sottospazio $\mathbb{T} \subset \mathbb{S}$. Scegliamo $P_0 \in \mathbb{T}$ e per $i \in \{1, \dots, d\}$ definiamo

$$\begin{array}{ccc} \mathsf{V}(\mathbb{S}) & \xrightarrow{f_i} & \mathbb{K} \\ v & \mapsto & F(P_0 + v) \end{array} \tag{4.4.14}$$

L'applicazione f_i è lineare, infatti è identificata con $\mathsf{V}(F_i)$. Inoltre è chiaro che

$$\mathsf{V}(\mathbb{T}) = \{v \in \mathsf{V}(\mathbb{S}) \mid f_1(v) = \dots = f_d(v) = 0\}. \tag{4.4.15}$$

Esempio 4.4.12. Siano $\mathbb{S} = \mathbb{A}^4(\mathbb{Q})$ e $\mathbb{T} = P_0 + U$ dove $P_0 = (-2, -1, 1, 2)$ e $U \subset \mathbb{Q}^4$ è dato da

$$U := \langle (1, -2, 3, -4), (2, 0, 3, 1) \rangle. \tag{4.4.16}$$

Sia $f \colon \mathbb{Q}^4 \to \mathbb{Q}$ lineare, cioè

$$\begin{array}{ccc} \mathbb{R}^4 & \xrightarrow{f} & \mathbb{R} \\ (x_1, x_2, x_3, x_4) & \mapsto & \lambda_1 x_1 + \dots + \lambda_4 x_4 \end{array} \tag{4.4.17}$$

Allora $f|_U = 0$ (cioè $f \in \operatorname{Ann} U$) se e solo se

$$0 = f(1, -2, 3, -4) = f(2, 0, 3, 1)$$

cioè

$$0 = \lambda_1 - 2\lambda_2 + 3\lambda_3 - 4\lambda_4 = 2\lambda_1 + 3\lambda_3 + \lambda_4. \tag{4.4.18}$$

Risolvendo il sistema di equazioni lineari (4.4.18) troviamo che una base di $\operatorname{Ann} U$ è data da $\{2e_1^\vee + 9e_2^\vee - 4e_4^\vee, 6e_1^\vee - 3e_2^\vee - 4e_3^\vee\}$ e quindi

$$U = \{X \in \mathbb{Q}^4 \mid 0 = 2x_1 + 9x_2 - 4x_4 = 6x_1 - 3x_2 - 4x_3\}. \tag{4.4.19}$$

Le equazioni cartesiane di U sono date da (4.4.19): segue che

$$\mathbb{T} = \{X \in \mathbb{A}^4(\mathbb{Q}) \mid 0 = F_1(X) = F_2(X)\} \tag{4.4.20}$$

dove

$$F_1(X) := 2(x_1 + 2) + 9(x_2 + 1) - 4(x_4 - 2), \qquad F_2(X) := 6(x_1 + 2) - 3(x_2 + 1) - 4(x_3 - 1). \tag{4.4.21}$$

Quindi

$$\mathbb{T} = \{X \in \mathbb{A}^4(\mathbb{Q}) \mid 0 = 2x_1 + 9x_2 - 4x_4 + 21 = 6x_1 - 3x_2 - 4x_3 + 13\}.$$

Composizione di applicazioni affini

Proposizione 4.4.13. *Siano $\mathbb{S}, \mathbb{T}, \mathbb{U}$ spazi affini sullo stesso campo e $F \colon \mathbb{S} \to \mathbb{T}$, $G \colon \mathbb{T} \to \mathbb{U}$ applicazioni affini. La composizione $G \circ F \colon \mathbb{S} \to \mathbb{C}$ è affine con applicazione lineare associata $\mathsf{V}(G) \circ \mathsf{V}(F)$.*

Dimostrazione. Poniamo $f := \mathsf{V}(F)$ e $g := \mathsf{V}(G)$. Siano $P, Q \in \mathbb{S}$. Allora

$$\overrightarrow{G \circ F(P) G \circ F(Q)} = \overrightarrow{G(F(P)) G(F(Q))} = g(\overrightarrow{F(P) F(Q)} = g(f(\overrightarrow{PQ}) = f \circ g(\overrightarrow{PQ}).$$

□

Un'applicazione affine $F \colon \mathbb{S} \to \mathbb{T}$ è un *isomorfismo* se ha inversa affine $G \colon \mathbb{T} \to \mathbb{S}$.

Proposizione 4.4.14. *Siano $\mathbb{S}, \mathbb{T}$ spazi affini sullo stesso campo. Un'applicazione affine $F \colon \mathbb{S} \to \mathbb{T}$ è un* isomorfismo *se e solo se l'applicazione lineare associata è un isomorfismo di spazi vettoriali.*

Dimostrazione. Supponiamo che $F\colon \mathbb{S} \to \mathbb{T}$ sia un isomorfismo con inversa affine $G\colon \mathbb{T} \to \mathbb{S}$. Per la Proposizione 4.4.13 abbiamo

$$\mathsf{V}(G) \circ \mathsf{V}(F) = \mathsf{V}(G \circ F) = \mathsf{V}(\mathrm{Id}_{\mathbb{S}}) = \mathrm{Id}_{\mathsf{V}(\mathbb{S})}, \quad \mathsf{V}(F) \circ \mathsf{V}(G) = \mathsf{V}(F \circ G) = \mathsf{V}(\mathrm{Id}_{\mathbb{T}}) = \mathrm{Id}_{\mathsf{V}(\mathbb{T})},$$

$\mathsf{V}(F)$ è un isomorfismo. Ora supponiamo che $F\colon \mathbb{S} \to \mathbb{T}$ sia affine e che $f := \mathsf{V}(F)$ sia un isomorfismo; quindi è definita $f^{-1}\colon \mathsf{V}(\mathbb{T}) \to \mathsf{V}(\mathbb{S})$ ed è un isomorfismo di spazi vettoriali. Sia $P_0 \in \mathbb{S}$ e poniamo $Q_0 = F(P_0)$. Definiamo $G\colon \mathbb{T} \to \mathbb{S}$ così:

$$G(Q) := P_0 + f^{-1}(\overrightarrow{Q_0 Q}).$$

Si verifica subito che G è inversa di F. □

Proposizione 4.4.15. *Sia $\mathbb{S}$ uno spazio affine. Se F, G sono isomorfismo affini di $\mathbb{S}$ anche la composizione $G \circ F$ lo è.*

Dimostrazione. Per la Proposizione 4.4.13 $G \circ F$ è affine e $\mathsf{V}(G \circ F) = \mathsf{V}(G) \circ \mathsf{V}(F)$. Siccome F e G sono isomorfismi di spazi affini $\mathsf{V}(G)$ e $\mathsf{V}(F)$ sono isomorfismi di spazi vettoriali per la Proposizione 4.4.14, e quindi la composizione $\mathsf{V}(G) \circ \mathsf{V}(F)$ è un isomorfismo di spazi vettoriali. Quindi $\mathsf{V}(G \circ F)$ è un isomorfismo di spazi vettoriali, e per la Proposizione 4.4.14 segue che $G \circ F$ è un isomorfismo affine. □

Un isomorfismo affine di $\mathbb{S}$ si chiama anche *automorfismo* o *affinità* di $\mathbb{S}$. Poniamo

$$\mathrm{Aut}(\mathbb{S}) := \{F\colon \mathbb{S} \to \mathbb{S} \mid F \text{ è un automorfismo di } \mathbb{S}\}. \tag{4.4.22}$$

Per la Proposizione 4.4.15 la composizione definisce un'operazione

$$\mathrm{Aut}(\mathbb{S}) \times \mathrm{Aut}(\mathbb{S}) \quad \longrightarrow \quad \mathrm{Aut}(\mathbb{S}) \tag{4.4.23}$$

Si verifica facilmente che $\mathrm{Aut}(\mathbb{S})$ è un gruppo, è il *gruppo degli automorfismi* o *delle affinità* di $\mathbb{S}$.

Esempio 4.4.16. Per l'Esempio 4.4.4 ogni $F \in \mathrm{Aut}(\mathbb{A}^n(\mathbb{K}))$ si scrive come

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{f} & \mathbb{K}^n \\ X & \mapsto & A \cdot X + B \end{array} \tag{4.4.24}$$

dove $A \in M_{n,n}(\mathbb{K})$ e $B \in M_{n,1}$. Inoltre A deve essere invertibile per la Proposizione 4.4.14. Viceversa se A è invertibile la (4.4.24) definisce un isomorfismo affine. Quindi

$$\mathrm{Aut}(\mathbb{A}^n(\mathbb{K})) := \{F\colon \mathbb{A}^n(\mathbb{K}) \to \mathbb{A}^n(\mathbb{K}) \mid F(X) = A \cdot X + B,\ A \in \mathrm{GL}_n(\mathbb{K}),\ B \in M_{n,1} \text{ fissati}\}. \tag{4.4.25}$$

Proposizione 4.4.17. *Sia $\mathbb{S}$ uno spazio affine di dimensione finita. L'applicazione*

$$\begin{array}{ccc} \mathrm{Aut}(\mathbb{S}) & \longrightarrow & \mathrm{GL}(\mathsf{V}(\mathbb{S})) \\ F & \mapsto & \mathsf{V}(F) \end{array} \tag{4.4.26}$$

è un omomorfismo di gruppi. Inoltre $\mathsf{V}(F)$ è l'identità di $\mathsf{V}(\mathbb{S})$ se e solo se F è una traslazione di $\mathbb{S}$.

Dimostrazione. Per la Proposizione 4.4.13 si ha $\mathsf{V}(F \circ G) = \mathsf{V}(F) \circ \mathsf{V}(G)$ per ogni $F, G \in \mathrm{Aut}(\mathbb{S})$, cioè l'applicazione in (4.4.26) è un omomorfismo di gruppi. La seconda affermazione è dimostrata nell'Esempio 4.4.3. □

Coordinate affini

Sia $\mathbb{S}$ uno spazio affine su $\mathbb{K}$ di dimensione finita. Le coordinate affini su $\mathbb{S}$ sono l'analogo delle coordinate su uno spazio vettoriale di dimensione finita. L'unica differenza è che dobbiamo scegliere un'"origine", cioè un punto $O \in \mathbb{S}$ che avrà coordinate tutte nulle.

Terminologia 4.4.18. Un *sistema di riferimento affine* $RA(O;\mathcal{B})$ su uno spazio affine $\mathbb{S}$ di dimensione finita consiste di un punto $O \in \mathbb{S}$ e una base di di $\mathsf{V}(\mathbb{S})$.

Dato un sistema di riferimento affine $\mathsf{R} := RA(O;\mathcal{B})$, con $\mathcal{B} = \{v_1, \dots, v_n\}$, l'applicazione

$$\begin{array}{ccc} \mathbb{S} & \xrightarrow{X_{\mathsf{R}}} & \mathbb{A}^n(\mathbb{K}) \\ P & \mapsto & X_{\mathcal{B}}(\overrightarrow{OP}) \end{array} \tag{4.4.27}$$

è un isomorfismo affine per le Proposizioni 4.4.7 e 4.4.14.

Osservazione 4.4.19. Notiamo che (4.4.27) è caratterizzata dalle equazioni

$$X_{\mathsf{R}}(O) = 0, \quad X_{\mathsf{R}}(O + v_i) = e_i, \qquad i \in \{1, \dots, n\}, \tag{4.4.28}$$

dove $e_1, \dots, e_n \in \mathbb{A}^n(\mathbb{K})$ sono i punti dati dai vettori della base canonica.

Le *coordinate* di P nel sistema di riferimento affine $RA(O;\mathcal{B})$ sono le entrate di $X_{\mathsf{R}}(P)$, cioè le coordinate del vettore $\overrightarrow{OP}$ nella base $\mathcal{B}$. Quando $RA(O;\mathcal{B})$ è assegnato si scrive $P(a_1, \dots, a_n)$ per abbreviare l'espressione "P è il punto di $\mathbb{S}$ con coordinate $(a_1, \dots, a_n)$ nel riferimento $RA(O;\mathcal{B})$". Esplicitamente:

$$P(a_1, \dots, a_n) = O + \sum_{i=1}^{n} a_i v_i. \tag{4.4.29}$$

Esempio 4.4.20. Sia $\mathbb{S} \subset \mathbb{A}^3(\mathbb{R})$ l'insieme delle soluzioni X dell'equazione

$$x_1 + x_2 + x_3 = 3,$$

e $V \subset \mathbb{R}^3$ il sottospazio delle soluzioni dell'equazione

$$x_1 + x_2 + x_3 = 0.$$

Nell'Esempio 4.1.2 abbiamo dato una struttura di spazio affine a $\mathbb{S}$ con gruppo delle traslazioni V. Siano $O = (1,1,1) \in \mathbb{S}$ e $\mathcal{B}$ la base di V data da $\mathcal{B} := \{(1,-1,0), (1,0,-1)\}$. Sia $P = (5,-2,0) \in \mathbb{S}$: calcoliamo le coordinate di P nel riferimento $RA(O;\mathcal{B})$. Le coordinate λ, μ di P sono tali che

$$\overrightarrow{OP} = \lambda(1,-1,0) + \mu(1,0,-1).$$

Abbiamo

$$\overrightarrow{OP} = (5,-2,0) - (1,1,1) = (4,-3,-1).$$

(Notate che $(4,-3,-1) \in V$, come deve essere.) Quindi troviamo λ, μ risolvendo il sistema

$$(4,-3,-1) = \lambda(1,-1,0) + \mu(1,0,-1).$$

Troviamo che $\lambda = 3$ e $\mu = 1$, cioè P ha coordinate $(3,1)$.

Ora sia $\mathsf{R}' = RA(O';\mathcal{B}')$ un secondo sistema di riferimento affine e $X_{\mathsf{R}'} \colon \mathbb{S} \to \mathbb{A}^n(\mathbb{K})$ l'applicazione che associa a $P \in \mathbb{S}$ l'n-pla delle sue coordinate nel sistema R'. Sia $P \in \mathbb{S}$: che relazione esiste tra $X_{\mathsf{R}}(P)$ e $X_{\mathsf{R}'}(P)$? Per rispondere osserviamo che la composizione

$$X_{\mathsf{R}'} \circ X_{\mathsf{R}}^{-1} \colon \mathbb{A}^n(\mathbb{K}) \longrightarrow \mathbb{A}^n(\mathbb{K}) \tag{4.4.30}$$

è un isomorfismo di spazi affini e quindi per l'Esempio 4.4.4 esistono $A \in \mathrm{GL}_n(\mathbb{K})$ e $B \in M_{n,1}(\mathbb{K})$ tali che

$$X_{\mathsf{R}'} \circ X_{\mathsf{R}}^{-1}(Y) = A \cdot Y + B, \qquad \forall\, Y \in \mathbb{K}^n. \tag{4.4.31}$$

Sia $P = X_{\mathsf{R}}^{-1}(Y)$ cioè $Y = X_{\mathsf{R}}(P)$: possiamo riscrivere la (4.4.31) come

$$X_{\mathsf{R}'}(P) = A \cdot X_{\mathsf{R}}(P) + B, \quad \forall P \in \mathbb{S}. \qquad (A \in \mathrm{GL}_n(\mathbb{K}).) \tag{4.4.32}$$

Esempio 4.4.21. Siano $O = (1, -2) \in \mathbb{A}^2(\mathbb{K})$ e sia $\mathcal{B}$ la base di $\mathbb{K}^2$ data da $\mathcal{B} := \{(3, 2), (7, 5)\}$ (notate che è una base qualsiasi sia il campo $\mathbb{K}$). Esprimiamo le coordinate del punto $X \in \mathbb{A}^2(\mathbb{K})$ nel riferimento $RA(O; \mathcal{B})$, cioè calcoliamo $X_{\mathsf{R}}(X)$. Se $Y = X_{\mathsf{R}}(X)$ (cioè Y è la colonna delle coordinate di X nel riferimento $RA(O; \mathcal{B})$), allora

$$Y = A \cdot X + B, \quad A \in \mathrm{GL}_2(\mathbb{K}),\ B \in M_{2,1}(\mathbb{K}).$$

È più facile calcolare X a partire da Y, e poi notare che

$$X = A^{-1} \cdot Y - A^{-1} \cdot B.$$

Scriviamo

$$X = M \cdot Y + C, \qquad M = A^{-1}, \quad C = -A^{-1} \cdot B, \tag{4.4.33}$$

e calcoliamo M, C. Siccome $O = (1, -2)$ abbiamo $C = (1, -2)^t$. Per calcolare M, notiamo che $\mathsf{V}(X_{\mathsf{R}}) = L_A$ e quindi, siccome $M = A^{-1}$, L_M è l'applicazione lineare che associa alle coordinate di un vettore di $\mathbb{K}^2$ nella base $\mathcal{B}$ le sue coordinate nella base standard. Dunque

$$M = \begin{bmatrix} 3 & 7 \\ 2 & 5 \end{bmatrix}$$

Calcolando l'inversa di M, troviamo che

$$A = \begin{bmatrix} 5 & -7 \\ -2 & 3 \end{bmatrix}$$

Per (4.4.33) abbiamo

$$B = -\begin{bmatrix} 5 & -7 \\ -2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -19 \\ 8 \end{bmatrix}$$

In conclusione le coordinate di $(x_1, x_2) \in \mathbb{A}^2(\mathbb{K})$ nel riferimento $RA(O, \mathcal{B})$ sono date da

$$y_1 = 5x_1 - 7x_2 - 19, \quad y_2 = -2x_1 + 3x_2 + 8.$$

4.5 Ginnastica affine

In tutta la presente sezione $\mathbb{S}$ è uno spazio affine di dimensione finita n, e $RA(O, \mathcal{B})$ è un riferimento cartesiano su $\mathbb{S}$, dove $\mathcal{B} = \{v_1, \ldots, v_n\}$.

Equazioni parametriche di sottospazi

Sia $\mathbb{T} \subset \mathbb{S}$ un sottospazio affine. Per definizione esistono $P_0(a_1, \ldots, a_n) \in \mathbb{S}$ e un sottospazio vettoriale $W \subset \mathsf{V}(\mathbb{S})$ tali che

$$\mathbb{T} = P_0 + W := \{P_0 + w \mid w \in W\}. \tag{4.5.1}$$

Sia $\{w_1, \ldots, w_m\}$ una base di W e siano $(c_{i1}, \ldots, c_{in})$ le coordinate di w_i nella base $\mathcal{B}$, cioè

$$w_i = \sum_{j=1}^{n} c_{ij} v_j. \tag{4.5.2}$$

Segue da (4.5.1) che i punti di $\mathbb{T}$ sono dati da

$$P(t_1, \ldots, t_m) := P_0 + \sum_{i=1}^{m} t_i \left(\sum_{j=1}^{n} c_{ij} v_j \right) = P_0 + \sum_{j=1}^{n} \left(\sum_{i=1}^{m} t_i c_{ij} \right) v_j, \quad t_1, \ldots, t_m \in \mathbb{K}.$$

Per $P(t) = P(t_1, \dots, t_m) \in \mathbb{T}$ abbiamo

$$\overrightarrow{OP(t)} = \overrightarrow{OP_0} + \overrightarrow{P_0P(t)} = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} (a_j + t_i c_{ij}) \right) v_j.$$

Perciò le coordinate $(x_1, \dots, x_n)$ dei punti in $\mathbb{T}$ sono date da

$$x_j = \sum_{i=1}^{m} (a_j + t_i c_{ij}), \qquad t_1, \dots, t_m \in \mathbb{K}. \tag{4.5.3}$$

Le formule in (4.5.3) si chiamano *equazioni parametriche* di $\mathbb{T}$ (i "parametri" sono $t_1, \dots, t_r$).

Esempio 4.5.1. Se $\mathbb{T} \subset \mathbb{S}$ è una retta, cioè $\dim \mathbb{T} = 1$, allora equazioni parametriche sono

$$\mathbb{T} \begin{cases} x_1 = a_1 + l_1 t, \\ \cdots\cdots\cdots \\ x_j = a_j + l_j t, \\ \cdots\cdots\cdots \\ x_n = a_n + l_n t, \end{cases}$$

dove $a_j, l_j \in \mathbb{K}$ sono fissati con $(l_1, \dots, l_n) \neq (0, \dots, 0)$ $t \in \mathbb{K}$ (il parametro) è arbitrario. Gli scalari $l_1, \dots, l_n$, che sono le coordinate di un generatore di $\mathcal{V}(\mathbb{T})$ nella base $\mathcal{B}$, sono i *parametri direttori* di $\mathbb{T}$.

Esempio 4.5.2. Se $\mathbb{T} \subset \mathbb{S}$ è un piano, cioè $\dim \mathbb{T} = 2$, allora equazioni parametriche sono

$$\mathbb{T} \begin{cases} x_1 = a_1 + l_1 s + m_1 t, \\ \cdots\cdots\cdots \\ x_j = a_j + l_j s + m_j t, \\ \cdots\cdots\cdots \\ x_n = a_n + l_n s + m_n t, \end{cases}$$

dove $a_j, l_j, m_j \in \mathbb{K}$ sono fissati con i vettori $(l_1, \dots, l_n), (m_1, \dots, m_n) \in \mathbb{K}^n$ linearmente indipendenti, e $s, t \in \mathbb{K}$ (l parametri) sono arbitrari. Notiamo che $(l_1, \dots, l_n), (m_1, \dots, m_n) \in \mathbb{K}^n$ sono le coordinate di una base di $\mathsf{V}(\mathbb{T})$ nella base $\mathcal{B}$.

Esempio 4.5.3. Siano $P_0(a_1, \dots, a_n), P_1(b_1, \dots, b_n) \in \mathbb{S}$ distinti, cioè linearmente indipendenti. Quindi $\mathbb{T} := \langle P_0, P_1 \rangle$ è una retta. Equazioni parametriche di $\mathbb{T}$ sono

$$\mathbb{T} \begin{cases} x_1 = a_1 + (b_1 - a_1) t, \\ \cdots\cdots\cdots \\ x_j = a_j + (b_j - a_j) t, \\ \cdots\cdots\cdots \\ x_n = a_n + (b_n - a_n) t. \end{cases}$$

Infatti $\mathsf{V}(\mathbb{T})$ ha come base il vettore $\overrightarrow{P_0P_1}$, e le coordinate di questo vettore sono $(b_1 - a_1, \dots, b_n - a_n)$ perchè

$$\overrightarrow{P_0P_1} = \overrightarrow{OP_1} - \overrightarrow{OP_0} = \sum_{j=1}^{n} b_j v_j - \sum_{j=1}^{n} a_j v_j = \sum_{j=1}^{n} (b_j - a_j) v_j.$$

Esempio 4.5.4. Siano $P_0(a_1, \dots, a_n), P_1(b_1, \dots, b_n), P_2(c_1, \dots, c_n) \in \mathbb{S}$ linearmente indipendenti. Quindi $\mathbb{T} := \langle P_0, P_1, P_2 \rangle$ è un piano. Equazioni parametriche di $\mathbb{T}$ sono

$$\mathbb{T} \begin{cases} x_1 = a_1 + (b_1 - a_1) s + (c_1 - a_1) t, \\ \cdots\cdots\cdots \\ x_j = a_j + (b_j - a_j) s + (c_j - a_j) t, \\ \cdots\cdots\cdots \\ x_n = a_n + (b_n - a_n) s + (c_n - a_n) t. \end{cases}$$

Infatti $\mathsf{V}(\mathbb{T})$ ha come base $\{\overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}\}$, e le coordinate di questi vettori sono rispettivamente $(b_1 - a_1, \ldots, b_n - a_n)$ e $(c_1 - a_1, \ldots, c_n - a_n)$ perchè

$$\begin{aligned}\overrightarrow{P_0P_1} &= \overrightarrow{OP_1} - \overrightarrow{OP_0} = \sum_{j=1}^{n} b_j v_j - \sum_{j=1}^{n} a_j v_j = \sum_{j=1}^{n} (b_j - a_j) v_j,\\ \overrightarrow{P_0P_2} &= \overrightarrow{OP_2} - \overrightarrow{OP_0} = \sum_{j=1}^{n} c_j v_j - \sum_{j=1}^{n} a_j v_j = \sum_{j=1}^{n} (c_j - a_j) v_j.\end{aligned}$$

Esempio 4.5.5. Sia $\mathbb{A}$ un piano affine, e sia $RA(O, \mathcal{B})$ un riferimento affine su $\mathbb{A}$. Siano $\mathbb{L}, \mathbb{M} \subset \mathbb{A}$ le rette di equazioni parametriche

$$\mathbb{L} \begin{cases} x_1 = 3 + 2t, \\ x_2 = -1 + 3t. \end{cases} \qquad \mathbb{M} \begin{cases} x_1 = 5 + 4t, \\ x_2 = 2 + 6t. \end{cases}$$

Ci chiediamo: $\mathbb{L}$ è uguale a $\mathbb{M}$? Se sono uguali allora $\mathsf{V}(\mathbb{L}) = \mathsf{V}(\mathbb{M})$. Siccome $\mathsf{V}(\mathbb{L})$ è generato dal vettore di coordinate $(2, 3)$ e $\mathsf{V}(\mathbb{M})$ è generato dal vettore di coordinate $(4, 6)$, e $\{(2, 3), (4, 6)\}$ sono linearmente dipendenti, vediamo che $\mathsf{V}(\mathbb{L}) = \mathsf{V}(\mathbb{M})$. Questo mostra che $\mathbb{L}$ e $\mathbb{M}$ sono parallele. Siccome $\mathbb{L}$ e $\mathbb{M}$ sono parallele, sono uguali se e solo se hanno un punto in comune. Il punto $P(5, 2)$ è in $\mathbb{L}$ (valore $t = 1$), ma anche in $\mathbb{M}$ (valore $t = 0$), quindi $\mathbb{L} = \mathbb{M}$. Questo esempio illustra quanto poco canoniche siano le equazioni parametriche. Nell'esempio il punto $(5 + 4t, 2 + 6t)$ dato dal secondo gruppo di equazioni parametriche è uguale al punto che, nel primo gruppo di equazioni parametriche, corrisponde al valore del parametro dato da $1 + 2t$.

Equazioni cartesiane

Abbiamo visto che le equazioni parametriche di un sottospazio $\mathbb{T} \subset \mathbb{S}$ sono molto lontane dall'essere uniche. Invece l'equazione cartesiana, nel caso in cui $\dim \mathbb{T} = \dim \mathbb{S} - 1$, è sostanzialmente unica.

Proposizione 4.5.6. *Sia $\mathbb{T} \subset \mathbb{S}$ un sottospazio di codimensione 1, cioè $\dim \mathbb{T} = \dim \mathbb{S} - 1$. Se*

$$a_1 x_1 + \ldots + a_n x_n + b = 0 \tag{4.5.4}$$

è un'equazione cartesiana di $\mathbb{T}$, e

$$c_1 x_1 + \ldots + c_n x_n + d = 0 \tag{4.5.5}$$

è un'altra equazione cartesiana di $\mathbb{T}$, allora esiste $\lambda \in \mathbb{K}$ (non nullo) tale che

$$a_1 = \lambda c_1, \ldots, a_n = \lambda c_n, \quad b = \lambda d. \tag{4.5.6}$$

Dimostrazione. Per (4.4.15) il sottospazio vettoriale $\mathsf{V}(\mathbb{T})$ ha equazioni cartesiane

$$a_1 x_1 + \ldots + a_n x_n = 0$$

e anche l'equazione cartesiana

$$c_1 x_1 + \ldots + c_n x_n = 0.$$

In altre parole le applicazioni lineari φ, ψ (non nulle) definite da $\varphi(X) = a_1 x_1 + \ldots + a_n x_n$ e $\psi(X) = c_1 x_1 + \ldots + c_n x_n$ appartengono all'annullatore di $\mathsf{V}(\mathbb{T})$. Siccome $\dim \mathbb{T} = \dim \mathbb{S} - 1$ l'annullatore di $\mathsf{V}(\mathbb{T})$ ha dimensione 1, e quindi esiste $\lambda \in \mathbb{K}$ (non nullo) tale che $\varphi = \lambda\psi$, cioè

$$a_1 = \lambda b_1, \ldots, a_n = \lambda c_n. \tag{4.5.7}$$

Ora sia $P(x_1, \ldots, x_n) \in \mathbb{T}$, e quindi valgono (4.5.4) e (4.5.5). Moltiplicando (4.5.5) per λ otteniamo che

$$\lambda c_1 x_1 + \ldots + \lambda c_n x_n + \lambda d = 0 \tag{4.5.8}$$

Sottraendo l'ultima equazione dalla (4.5.4) segue che $b - \lambda d = 0$. Quest'ultima equazione e (4.5.7) dimostrano che vale (4.5.6). □

Vediamo come passare da equazioni parametriche a equazioni cartesiane in alcuni casi.

Esempio 4.5.7. Supponiamo che $\mathbb{S}$ sia un piano e che $\mathbb{L} \subset \mathbb{A}$ sia la retta di equazioni parametriche (con parametro t)

$$\begin{cases} x_1 = a_1 + lt, \\ x_2 = a_2 + mt. \end{cases}$$

Allora un'equazione cartesiana di $\mathbb{L}$ è

$$m(x_1 - a_1) - l(x_2 - a_2) = 0.$$

Prima di dare l'analogo esempio per un piano $\mathbb{T} \subset \mathbb{A}$, dimostriamo un risultato.

Proposizione 4.5.8. *Sia*

$$M := \begin{bmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{bmatrix} \tag{4.5.9}$$

una matrice in $M_{2,3}(\mathbb{K})$ di rango massimo, cioè 2. Allora almeno uno dei determinanti delle sue sottomatrici 2×2 è non nullo, cioè il vettore

$$(l_2m_3 - l_3m_2, -(l_1m_3 - l_3m_1), l_1m_2 - l_2m_1) \in \mathbb{K}^3 \tag{4.5.10}$$

è non nullo. Inoltre il vettore in (4.5.10) *genera lo spazio delle soluzioni del sistema di equazioni lineari omogenee associato a M.*

Dimostrazione. Dimostriamo per assurdo che il vettore in (4.5.10) è non nullo. Siccome, per ipotesi, il rango di M è 2 una delle colonne della matrice M è non nulla. Possiamo assumere che la prima colonna, cioè $(l_1, m_1)^t$ sia non nulla. Siccome è nullo il determinante della sottomatrice di M con colonne le prime due colonne, (l_2, m_2) è un multiplo di (l_1, m_1) (vedi l'Osservazione 2.5.7), e analogamente (l_3, m_3) è un multiplo di (l_1, m_1). Ma allora il rango di M è 1. Abbiamo dimostrato che il vettore in (4.5.10) è non nullo.

Un conto diretto mostra che il vettore in (4.5.10) è una soluzione del sistema di equazioni lineari omogenee associato a M. D'altra parte, siccome il rango di M è 2, lo spazio delle soluzioni di $M \cdot X = 0$ è 1, e quindi ogni soluzione è un multiplo della soluzione non nulla in (4.5.10). □

Esempio 4.5.9. Supponiamo che $\dim \mathbb{S} = 3$ e che $\mathbb{T} \subset \mathbb{A}$ sia il piano di equazioni parametriche (con parametri s e t)

$$\begin{cases} x_1 = a_1 + l_1s + m_1t, \\ x_2 = a_2 + l_2s + m_2t, \\ x_3 = a_3 + l_3s + m_3t, \end{cases}$$

Sia

$$c_1x_1 + c_2x_2 + c_3x_3 = b$$

un'equazione cartesiana di $\mathbb{T}$. Siccome $(a_1, a_2, a_3) \in \mathbb{T}$, possiamo riscriverla come

$$c_1(x_1 - a_1) + c_2(x_2 - a_2) + c_3(x_3 - a_3) = 0.$$

Siccome $(a_1 + l_1, a_2 + l_2, a_3 + l_3) \in \mathbb{T}$ e $(a_1 + m_1, a_2 + m_2, a_3 + m_3) \in \mathbb{T}$, otteniamo che

$$\begin{aligned} c_1l_1 + c_2l_2 + c_3l_3 &= 0, \\ c_1m_1 + c_2m_2 + c_3m_3 &= 0, \end{aligned}$$

cioè (c_1, c_2, c_3) è una soluzione del sistema di equazioni lineari omogenee associato alla matrice

$$M := \begin{bmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{bmatrix} \tag{4.5.11}$$

Siccome $\mathbb{T}$ è un piano i vettori $(l_1, l_2, l_3), (m_1, m_2, m_3) \in \mathbb{K}^3$ sono linearmente indipendenti, cioè M^t ha rango 2. Siccome $\mathrm{rg}(M) = \mathrm{rg}(M^t)$, segue che $\mathrm{rg}(M) = 2$, e quindi per la Proposizione 4.5.8 un'equazione cartesiana di $\mathbb{T}$ è

$$(l_2m_3 - l_3m_2)(x_1 - a_1) - (l_1m_3 - l_3m_1)(x_2 - a_2) + (l_1m_2 - l_2m_1)(x_3 - a_3) = 0.$$

Passare da equazioni cartesiane a equazioni parametriche è ancora più semplice: basta trovare tutte le soluzioni del sistema di equazioni cartesiane con il metodo di eliminazione di Gauss.

Esempio 4.5.10. Supponiamo che $\mathbb{S}$ sia un piano e che $\mathbb{L} \subset \mathbb{S}$ sia la retta di equazione cartesiana

$$3x_1 - 2x_2 = 1.$$

Per semplicità supponiamo che $\mathrm{char}\,\mathbb{K} \neq 3$. Esprimendo la x_1 in termini della x_2 otteniamo che $x_1 = \frac{2}{3}x_2 + \frac{1}{3}$. Ponendo $x_2 = t$ otteniamo le seguenti equazioni parametriche di $\mathbb{L}$

$$\begin{cases} x_1 = \frac{1}{3} + \frac{2}{3}t, \\ x_2 = t. \end{cases}$$

Esempio 4.5.11. Supponiamo che $\mathbb{S}$ sia un piano e che $\mathbb{L} \subset \mathbb{S}$ sia la retta di equazione cartesiana

$$2x_2 = 5.$$

(Quindi $\mathrm{char}\,\mathbb{K} \neq 2$.) Equazioni parametriche di $\mathbb{L}$ sono

$$\mathbb{L} \begin{cases} x_1 = t, \\ x_2 = \frac{5}{2}. \end{cases}$$

Esempio 4.5.12. Supponiamo che $\mathbb{S}$ sia un solido, cioè $\dim \mathbb{S} = 3$, e che $\mathbb{L} \subset \mathbb{S}$ sia la retta di equazioni cartesiane

$$\begin{cases} x_1 + 3x_2 - x_3 = 1 \\ 2x_1 + 7x_2 + x_3 = 5. \end{cases}$$

Il metodo di eliminazione di Gauss dà le soluzioni del sistema di equazioni:

$$x_1 = 10x_3 - 8, \quad x_2 = 3 - 3x_3.$$

Ponendo $x_3 = t$ otteniamo le seguenti equazioni parametriche di $\mathbb{L}$

$$\begin{cases} x_1 = -8 + 10t, \\ x_2 = 3 - 3t, \\ x_3 = t. \end{cases}$$

Fasci di rette e piani

Definizione 4.5.13. Sia $\mathbb{S}$ un piano affine e $P_0 \in \mathbb{S}$. Il fascio (proprio) di rette (in $\mathbb{S}$) di centro P_0 è l'insieme delle rette in $\mathbb{S}$ contenenti P_0.

Proposizione 4.5.14. *Sia $\mathbb{S}$ un piano affine e $P_0 \in \mathbb{S}$. Siano*

$$\begin{cases} a_1x_1 + a_2x_2 + b = 0 \\ c_1x_1 + c_2x_2 + d = 0 \end{cases} \tag{4.5.12}$$

equazioni cartesiane (nelle coordinate x_1, x_2) di P_0 ($\{P_0\}$ è un sottospazio affine di codimensione 2 di $\mathbb{S}$). Una retta $\mathbb{L} \subset \mathbb{S}$ contiene P_0 (cioè appartiene al fascio di centro P_0) se e solo se esistono $\lambda, \mu \in \mathbb{K}$ non entrambi nulli tali che $\mathbb{L}$ abbia equazione cartesiana

$$\lambda(a_1x_1 + a_2x_2 + b) + \mu(c_1x_1 + c_2x_2 + d) = 0. \tag{4.5.13}$$

Dimostrazione. Se $\mathbb{L}$ ha equazione data da (4.5.13) allora $P_0 \in \mathbb{L}$ perchè le coordinate di P_0 sono suluzioni del sistema in (4.5.12). Ora supponiamo che $\mathbb{L} \subset \mathbb{S}$ sia una retta che contiene P_0. Sia $Q(\overline{x}_1, \overline{x}_2) \in \mathbb{L}$ un punto diverso da P_0 (esiste perchè $\mathbb{K}$ contiene almeno 2 elementi, 0 e 1), e sia (λ_0, μ_0) una soluzione non banale dell'equazione

$$\lambda(a_1\overline{x}_1 + a_2\overline{x}_2 + b) + \mu(c_1\overline{x}_1 + c_2\overline{x}_2 + d) = 0.$$

La retta di equazione cartesiana

$$\lambda_0(a_1x_1 + a_2x_2 + b) + \mu_0(c_1x_1 + c_2x_2 + d) = 0$$

contiene P_0 e Q, quindi è uguale a $\mathbb{L}$. □

Definizione 4.5.15. Sia $\mathbb{S}$ un piano affine e $W \subset \mathsf{V}(\mathbb{S})$ un sottospazio vettoriale di dimensione 1. Il fascio (improprio) di rette (in $\mathbb{S}$) di direzione W è l'insieme delle rette in $\mathbb{S}$ con giacitura W.

Proposizione 4.5.16. *Sia $\mathbb{S}$ un piano affine e $W \subset \mathsf{V}(\mathbb{S})$ un sottospazio vettoriale di dimensione 1. Sia*

$$a_1x_1 + a_2x_2 = 0$$

un'equazione cartesiana del sottospazio $W \subset \mathsf{V}(\mathbb{S})$. Una retta $\mathbb{L} \subset \mathbb{S}$ appartiene al fascio improprio di direzione W se e solo se esiste $b \in \mathbb{K}$ tale che $\mathbb{L}$ abbia equazione cartesiana

$$a_1x_1 + a_2x_2 + b = 0. \tag{4.5.14}$$

Dimostrazione. Se $\mathbb{L}$ ha equazione come in (4.5.14) allora la sua giacitura è uguale a W, vedi (4.4.15). Ora supponiamo che la retta $\mathbb{L} \subset \mathbb{S}$ abbia giacitura W. Sia $P(\overline{x}_1, \overline{x}_2) \in \mathbb{L}$. Sia $\mathbb{J} \subset \mathbb{S}$ la retta di equazione cartesiana

$$a_1(x - \overline{x}_1) + a_2(x_2 - \overline{x}_2) = 0. \tag{4.5.15}$$

La giacitura di $\mathbb{J}$ è W e $P \in \mathbb{J}$, quindi $\mathbb{J} = \mathbb{L}$. Abbiamo fatto perchè l'equazione cartesiana in (4.5.15) è del tipo (4.5.14). □

Osservazione 4.5.17. Sia $\mathbb{L} \subset \mathbb{E}^2$ una retta. Scegliamo equazioni parametriche di $\mathbb{L}$, e sia $P(t) \in \mathbb{L}$ il punto che corrisponde al parametro $t \in \mathbb{R}$. Il fascio proprio di centro $P(t)$, per $t \to \pm\infty$ si avvicina sempre più al fascio improprio di rette parallele a $\mathbb{L}$. Analogamente, se $\mathbb{T} \subset \mathbb{E}^3$ è un piano e $\mathbb{L}(t) \subset \mathbb{T}$ è una retta che dipende da t e che "va all'infinito" per $t \to \infty$, il fascio proprio di piani di asse $\mathbb{L}(t)$ si avvicina al fascio improprio di piani paralleli a $\mathbb{T}$.

Si può dare un senso ad analoghe affermazioni per un piano o un solido affini qualsiasi.

Definizione 4.5.18. Sia $\mathbb{S}$ un solido affine, cioè $\dim \mathbb{S} = 3$ e $\mathbb{L} \subset \mathbb{S}$ una retta. Il fascio di piani (in $\mathbb{S}$) di asse $\mathbb{L}$ è l'insieme dei piani in $\mathbb{S}$ contenenti $\mathbb{L}$.

Proposizione 4.5.19. *Sia $\mathbb{S}$ un solido affine e $\mathbb{L} \subset \mathbb{S}$ una retta. Siano*

$$\begin{cases} a_1x_1 + a_2x_2 + a_3x_3 + b = 0 \\ c_1x_1 + c_2x_2 + c_3x_3 + d = 0 \end{cases}$$

equazioni cartesiane (nelle coordinate x_1, x_2, x_3) di $\mathbb{L}$. Un piano $\mathbb{T} \subset \mathbb{S}$ contiene $\mathbb{L}$ (cioè appartiene al fascio di asse $\mathbb{L}$) se e solo se esistono $\lambda, \mu \in \mathbb{K}$ non entrambi nulli tali che $\mathbb{T}$ abbia equazione cartesiana

$$\lambda(a_1x_1 + a_2x_2 + a_3x_3 + b) + \mu(c_1x_1 + c_2x_2 + c_3x_3 + d) = 0. \tag{4.5.16}$$

Dimostrazione. La dimostrazione è del tutto simile a quella della Proposizione 4.5.14. Lasciamo al lettore i dettagli.

□

Definizione 4.5.20. Sia $\mathbb{S}$ un solido affine e $W \subset \mathsf{V}(\mathbb{S})$ un sottospazio vettoriale di dimensione 2. Il fascio (improprio) di piani (in $\mathbb{S}$) di direzione W è l'insieme dei piani in $\mathbb{S}$ con giacitura W.

La dimostrazione del risultato che segue è del tutto simile a quella della Proposizione 4.5.16, e per questo viene omessa.

Proposizione 4.5.21. *Sia $\mathbb{S}$ un solido e $W \subset \mathsf{V}(\mathbb{S})$ un sottospazio vettoriale di dimensione 2. Sia*

$$a_1x_1 + a_2x_2 + a_3x_3 = 0$$

un'equazione cartesiana del sottospazio $W \subset \mathsf{V}(\mathbb{S})$. Un piano $\mathbb{T} \subset \mathbb{S}$ appartiene al fascio improprio di direzione W se e solo se esiste $b \in \mathbb{K}$ tale che $\mathbb{L}$ abbia equazione cartesiana

$$a_1x_1 + a_2x_2 + a_3x_3 + b = 0. \tag{4.5.17}$$

Esempio 4.5.22. Sia $\mathbb{S}$ un solido affine reale e $\mathbb{L} \subset \mathbb{S}$ la retta di equazioni cartesiane (nelle coordinate x_1, x_2, x_3)

$$\begin{cases} 3x_1 + 2x_2 + x_3 - 6 = 0 \\ x_1 + 2x_2 + 3x_3 - 1 = 0 \end{cases} \tag{4.5.18}$$

Sia $\mathbb{J} \subset \mathbb{S}$ la retta di equazioni parametriche

$$\begin{cases} x_1 = 1 + 2t \\ x_2 = -2 + t \\ x_3 = 3t \end{cases}$$

Verifichiamo che $\mathbb{J}$ non è parallela a $\mathbb{L}$. Le equazioni cartesiane di $\mathsf{V}(\mathbb{L})$ sono date dal sistema di equazioni omogenee associato a (4.5.18) (vedi (4.4.15)), cioè

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 0 \\ x_1 + 2x_2 + 3x_3 = 0 \end{cases} \tag{4.5.19}$$

D'altra parte un generatore di $\mathsf{V}(\mathbb{J})$ ha coordinate $(2, 1, 3)$; siccome le sue entrate non soluzioni di (4.5.19), $\mathbb{J}$ non è parallela a $\mathbb{L}$. Segue che esiste un piano $\mathbb{T}$ del fascio di asse $\mathbb{L}$ parallelo a $\mathbb{J}$, ed è unico. Calcoliamo l'equazione cartesiana di $\mathbb{T}$. Per la Proposizione 4.5.19 un'equazione cartesiana è

$$\lambda(3x_1 + 2x_2 + x_3 - 6) + \mu(x_1 + 2x_2 + 3x_3 - 1) = 0,$$

dove $\lambda, \mu \in \mathbb{R}$ non sono entrambi nulli. Quindi un'equazione cartesiana di $\mathsf{V}(\mathbb{T})$ è

$$\lambda(3x_1 + 2x_2 + x_3) + \mu(x_1 + 2x_2 + 3x_3) = 0.$$

$\mathsf{V}(\mathbb{J})$ è contenuto in $\mathsf{V}(\mathbb{T})$ (cioè $\mathbb{T}$ è parallelo a $\mathbb{J}$) se lo è un suo generatore, per esempio quello di coordinate $(2, 1, 3)$, e quindi otteniamo l'equazione in λ, μ

$$11\lambda + 13\mu = 0.$$

Una soluzione è $\lambda = 13$, $\mu = -11$, e perciò un'equazione cartesiana di $\mathbb{T}$ è

$$28x_1 + 4x_2 - 20x_3 - 67 = 0.$$

Esercizi del Capitolo 4

Esercizio 4.1. $\mathbb{S}$ *è un piano affine con riferimento affine* $RA(O;\mathcal{B})$. *Siano* $P_0(1,2)$ *e* $P_1(-1,1)$ *punti di* $\mathbb{S}$. *Scrivete equazioni parametriche e cartesiane della retta* $\langle P_0, P_1\rangle$.

Esercizio 4.2. $\mathbb{S}$ *è un piano affine reale con riferimento affine* $RA(O;\mathcal{B})$, *e* $\mathbb{L},\mathbb{J}\subset\mathbb{S}$ *sono le rette di equazioni parametriche*

$$\begin{cases} x_1 = 1+3t, \\ x_2 = -2+t \end{cases}, \qquad \begin{cases} x_1 = s, \\ x_2 = 1-s \end{cases}$$

rispettivamente. Determinate le coordinate del punto d'intersezione tra $\mathbb{L}$ *e* $\mathbb{J}$.

Esercizio 4.3. *Sia* $\mathcal{B} := \{\boldsymbol{i},\boldsymbol{j}\}$ *una base di* $\mathsf{V}(\mathbb{E}^2)$ *e* $\boldsymbol{k} := \boldsymbol{i}+2\boldsymbol{j}$, $\boldsymbol{h} := \boldsymbol{i}+\boldsymbol{j}$. *Sia* $Q\in\mathbb{E}^2$ *il punto di coordinate* $(1,-1)$ *nel riferimento affine* $RA(O;\mathcal{B})$.

(1) Verificate che $\mathcal{C} := \{\boldsymbol{k},\boldsymbol{h}\}$ *è una base di* $\mathsf{V}(\mathbb{E}^2)$.

(2) Determinate le coordinate di O *nel riferimento affine* $RA(Q;\mathcal{C})$.

Esercizio 4.4. $\mathbb{S}$ *è un solido affine reale con riferimento affine* $RA(O;\mathcal{B})$. *Siano* $P_0(1,1,-1)$, $P_1(3,0,2)$ *e* $P_2(4,2,3)$ *punti di* $\mathbb{S}$.

1. *Verificate che* P_0,P_1,P_2 non *sono allineati e quindi appartengono a un unico piano* Λ.
2. *Determinate equazioni parametriche e cartesiane di* Λ.

Esercizio 4.5. $\mathbb{S}$ *è un solido affine reale con riferimento affine* $RA(O;\mathcal{B})$, *e* $\mathbb{T}_1,\mathbb{T}_2\subset\mathbb{S}$ *sono i piani di equazioni cartesiane*

$$\mathbb{T}_1:\ 3x_1+2x_2+x_3=1, \qquad \mathbb{T}_2:\ x_1-x_2-2x_3=2.$$

Verificate che l'intersezione $\mathbb{T}_1\cap\mathbb{T}_2$ *è una retta* $\mathbb{L}$ *e determinate equazioni parametriche di* $\mathbb{L}$.

Esercizio 4.6. $\mathbb{S}$ *è un solido affine reale con riferimento affine* $RA(O;\mathcal{B})$, *e* $\mathbb{L},\mathbb{J}\subset\mathbb{S}$ *sono le rette di equazioni parametriche*

$$\begin{cases} x_1 = 1+2t, \\ x_2 = -t, \\ x_3 = 2+5t. \end{cases}, \qquad \begin{cases} x_1 = s, \\ x_2 = 1+2s \\ x_3 = 3 \end{cases}$$

rispettivamente. Determinate una equazione cartesiana del piano $\mathbb{T}$ *contenente* $\mathbb{L}$ *e parallelo a* $\mathbb{J}$.

Esercizio 4.7. $\mathbb{S}$ *è un solido affine reale con riferimento affine* $RA(O;\mathcal{B})$, *e* $\mathbb{T}_1,\mathbb{T}_2\subset\mathbb{S}$ *sono i piani di equazioni cartesiane*

$$x_1+2x_2-x_3+1=0, \qquad 2x_1+x_3-3=0.$$

rispettivamente. Determinate equazioni parametriche della retta che contiene $P(1,1,1)$ *ed è parallela a* $\mathbb{T}_1$ *e* $\mathbb{T}_2$.

Esercizio 4.8. $\mathbb{S}$ *è un piano affine reale con riferimento affine* $RA(O;\mathcal{B})$. *Siano* $P_0(1,1)$, $P_1(2,3)$, $P_2(3,5)$ *punti di* $\mathbb{S}$. *Gli studenti Anna, Marco e Lucio misurano le coordinate di* P_0, P_1 *e* P_2 *in un nuovo sistema di riferimento* $RA(Q;\mathcal{C})$ *e le loro misurazioni sono discordanti:*

(A.) $P_0(0,1)$, $P_1(-1,0)$ *e* $P_2(-2,-1)$.

(M.) $P_0(0,2)$, $P_1(1,2)$ *e* $P_2(0,3)$.

(L.) $P_0(0,0)$, $P_1(1,1)$ *e* $P_2(3,3)$.

Due tra Anna, Marco e Lucio sicuramente hanno sbagliato misurazioni: determinate chi.

Esercizio 4.9. *Sia* $\mathbb{S}$ *uno spazio affine su un campo di caratteristica zero. Siano* $P_0,\dots,P_d\in\mathbb{S}$ *linearmente indipendenti. Il* baricentro *di* $P_0,\dots,P_d$ *è il punto*

$$B(P_0,\dots,P_d) := \frac{1}{d+1}P_0+\frac{1}{d+1}P_1+\dots+\frac{1}{d+1}P_d.$$

Sia r *la retta contenente* P_d *e* $B(P_0,\dots,P_d)$. *Verificate che l'intersezione tra* r *e il sottospazio affine* $\langle P_0,P_1,\dots,P_{d-1}\rangle$ *è il baricentro* $B(P_0,\dots,P_{d-1})$.

Esercizio 4.10. *Dite se esiste/non esiste un'applicazione affine* $F\colon \mathbb{A}^2(\mathbb{R})\to\mathbb{A}^2(\mathbb{R})$ *tale che*

(1) $F(1,1) = (1,2)$, $F(3,2) = (-1,-2)$ *e* $F(2,3/2) = (0,1)$.

(2) $F(0,0) = (1,1)$, $F(1,1) = (2,1)$ *e* $F(1,-1) = (1,2)$.

Nel caso esista dare una tale F.

Esercizio 4.11. *Sia* $\mathbb{S} \subset \mathbb{A}^4(\mathbb{R})$ *il sottospazio affine* $P + U$ *dove* $P = (1,0,3,-1)$ *e* $U \subset \mathbb{R}^4$ *(qui* $\mathbb{R}^4$ *è spazio vettoriale) è il sottospazio vettoriale*

$$U = \langle (1,1,1,1), (3,2,-1,5), (4,3,0,6) \rangle.$$

Date equazioni cartesiane di $\mathbb{S}$.

Esercizio 4.12. *Siano* $\mathbb{S}$ *uno spazio affine e* $F \colon \mathbb{S} \to \mathbb{S}$ *un'affinità. Un* punto fisso *di* F *è un* $P \in \mathbb{S}$ *tale che* $F(P) = P$. *Il luogo dei punti fissi* $\mathrm{Fix}(F)$ *è l'insieme dei punti fissi di* F.

1. *Dimostrate che se* $\mathrm{Fix}(F)$ *è non vuoto, allora è un sottospazio affine di* $\mathbb{S}$.
2. *Date esempi con* $\dim \mathbb{S} = n$ *e* $\dim \mathrm{Fix}(F) = m$ *per ogni* $0 \leqslant m \leqslant n$.

Capitolo 5

Determinanti

Il determinante di una matrice quadrata $n \times n$ è una funzione polinomiale omogenea di grado n nelle entrate della matrice che gode di notevoli proprietà. In particolare per una matrice con entrate in un campo il determinante è non nullo se e solo se la matrice è invertibile. Se il campo è quello dei reali possiamo interpretare il determinante come un'area o volume (con segno).

5.1 La definizione

Sia R un anello (commutativo con unità). Sia $A \in M_{n,n}(R)$. Sia A^i_j la matrice $(n-1)\times(n-1)$ ottenuta eliminando riga i-esima e colonna j-esima di A. Definiamo la funzione

$$\mathrm{Det}_n \colon M_{n,n}(R) \longrightarrow R$$

ricorsivamente come segue:

(1) $\mathrm{Det}_1((a)) = a$.

(2) Per $n > 1$ definiamo Det_n a partire da Det_{n-1} per mezzo della formula

$$\mathrm{Det}_n(A) := \sum_{j=1}^{n} (-1)^{n+j} a_{nj} \,\mathrm{Det}_{n-1}(A^n_j). \tag{5.1.1}$$

Spieghiamo il punto (2). Assumendo di aver definito Det_{n-1}, la funzione Det_n è data da (5.1.1); siccome Det_1 è data da (1) segue che Det_2 è bene definita e quindi anche Det_3 etc. Diamo alcuni esempi. Se $n = 2$ abbiamo

$$\mathrm{Det}_2 \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = -a_{21}a_{12} + a_{22}a_{11} = a_{11}a_{22} - a_{12}a_{21}, \tag{5.1.2}$$

cioè la formula della Definizione 2.5.6. Per $n = 3$ abbiamo

$$\begin{aligned} \mathrm{Det}_3 \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \\ = a_{31}(a_{12}a_{23} - a_{13}a_{22}) - a_{32}(a_{11}a_{23} - a_{13}a_{21}) + a_{33}(a_{11}a_{22} - a_{12}a_{21}) = \\ = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}. \end{aligned} \tag{5.1.3}$$

Il *determinante* di $A \in M_{n,n}(R)$ è $\mathrm{Det}_n(A)$. Se non c'è pericolo di ambiguità scriviamo Det invece di Det_n. Si usa anche la notazione $|A|$ per $\mathrm{Det}\, A$ - in questo caso si omette di scrivere le parentesi che delimitano la matrice. Per esempio

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = \mathrm{Det} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = -2. \tag{5.1.4}$$

5.2 Applicazioni multilineari

Studieremo il determinante di matrici con entrate in un campo $\mathbb{K}$. Per capire la funzione determinante Det_n la vedremo come funzione delle n colonne di una matrice $n \times n$, cioè la vedremo come applicazione

$$\begin{array}{ccc} \underbrace{\mathbb{K}^n \times \ldots \times \mathbb{K}^n}_{n} & \xrightarrow{\mathrm{Det}_n} & \mathbb{K} \\ (A_1, \ldots, A_n) & \mapsto & \mathrm{Det}([A_1, \ldots, A_n]) \end{array} \tag{5.2.1}$$

Per questo motivo apriamo una parentesi sulle applicazioni da prodotti di spazi vettoriali su $\mathbb{K}$ a uno spazio vettoriale sullo stesso campo $\mathbb{K}$.

Definizione 5.2.1. Siano $V_1, \ldots, V_n$ e W spazi vettoriali su $\mathbb{K}$ e sia Φ un'applicazione

$$\begin{array}{ccc} V_1 \times \ldots \times V_n & \xrightarrow{\Phi} & W \\ (v_1, \ldots, v_n) & \mapsto & \Phi(v_1, \ldots, v_n) \end{array} \tag{5.2.2}$$

(1) La Φ è *lineare nell'entrata j-esima* (dove $1 \leqslant j \leqslant n$) se

$$\begin{aligned} &\Phi(v_1, \ldots, v_{j-1}, \lambda u + \mu w, v_{j+1}, \ldots, v_n) = \\ &\qquad = \lambda\Phi(v_1, \ldots, v_{j-1}, u, v_{j+1}, \ldots, v_n) + \mu\Phi(v_1, \ldots, v_{j-1}, w, v_{j+1}, \ldots, v_n) \end{aligned} \tag{5.2.3}$$

per $v_1, \ldots, v_{j-1}, u, w, v_{j+1}, \ldots, v_n \in V$ e $\lambda, \mu \in \mathbb{K}$.

(2) Φ è *multilineare* se è lineare in ciascun entrata. (Se $n = 2$ diciamo che Φ è *bi*lineare.)

Esempio 5.2.2. Consideriamo le applicazioni $\Psi_i \colon \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ definite da

$$\Psi_1(x, y) := xy^2 + x, \quad \Psi_2(x, y) := 1, \quad \Psi_3(x, y) := 5xy.$$

La Ψ_1 è lineare nella prima entrata ma non nella seconda, la Ψ_2 non è lineare in alcuna entrata, la Ψ_2 è bilineare.

Esempio 5.2.3. Sia V uno spazio vettoriale su $\mathbb{K}$. L'applicazione

$$\begin{array}{ccc} V \times V^\vee & \longrightarrow & \mathbb{K} \\ (v, f) & \mapsto & f(v) \end{array} \tag{5.2.4}$$

è bilineare.

Esempio 5.2.4. Scelta un'unità di misura, è ben definito il prodotto vettoriale di due vettori geometrici dello spazio euclideo, e quindi abbiamo l'applicazione bilineare

$$\begin{array}{ccc} \mathsf{V}(\mathbb{E}^3) \times \mathsf{V}(\mathbb{E}^3) & \longrightarrow & \mathsf{V}(\mathbb{E}^3) \\ (v, w) & \mapsto & v \wedge w \end{array} \tag{5.2.5}$$

Ricordiamo che, se $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ è una base ortonormale di $\mathsf{V}(\mathbb{E}^3)$ (ortonormale significa che ogni vettore della base ha lunghezza 1 e che coppie di vettori diversi della base sono ortogonali), allora

$$(a_1\mathbf{i} + b_1\mathbf{j} + c_1\mathbf{k}) \wedge (a_2\mathbf{i} + b_2\mathbf{j} + c_2\mathbf{k}) = (b_1c_2 - c_1b_2)\mathbf{i} - (a_1c_2 - c_1a_2)\mathbf{j} + (a_1b_2 - b_1a_2)\mathbf{k}. \tag{5.2.6}$$

Segue da questa espressione che il prodotto vettoriale è bilineare.

Terminologia 5.2.5. Un'applicazione multilineare $\Phi \colon V_1 \times \ldots \times V_n \to \mathbb{K}$ si dice *forma multilineare*.

Proposizione 5.2.6. *La funzione* Det_n *è multilineare nelle colonne.*

Dimostrazione. Per induzione su n. Il caso $n = 1$ è banalmente vero. Dimostriamo il passo induttivo. Dimostriamo che Det_n è lineare nella colonna j_0-esima, cioè che

$$\begin{aligned}\mathrm{Det}_n(A_1,\dots,A_{j_0-1},\lambda B+\mu C,A_{j_0+1},\dots,A_n)&=\\ =\lambda\,\mathrm{Det}_n(A_1,\dots,A_{j_0-1},B,A_{j_0+1},\dots,A_n)&+\mu\,\mathrm{Det}_n(A_1,\dots,A_{j_0-1},C,A_{j_0+1},\dots,A_n)\end{aligned}\tag{5.2.7}$$

dove A_j (per $j \neq j_0$), B e C sono matrici colonna $n \times 1$ - le loro entrate saranno denotate a_{ij}, b_i e c_i rispettivamente. Per $j \neq j_0$ sia $X_j \in M_{n-1,1}(\mathbb{K})$ la colonna ottenuta eliminando l'ultima entrata di A_j. Siano $Y, Z \in M_{n-1,1}(\mathbb{K})$ le colonne ottenute eliminando l'ultima entrata di B e C rispettivamente. Si ha che

$$\begin{aligned}\mathrm{Det}_n(A_1,\dots,A_{j_0-1},\lambda B+\mu C,A_{j_0+1},\dots,A_n)&=\\ =\sum_{j\neq j_0}(-1)^{n+j}a_{nj}\,\mathrm{Det}_{n-1}(X_1,\dots,X_{j-1},\widehat{X_j},X_{j+1},\dots,X_{j_0-1},\lambda Y+\mu Z,X_{j_0+1},\dots,X_n)&+\\ +(-1)^{n+j_0}(\lambda b_n+\mu c_n)\,\mathrm{Det}_{n-1}(X_1,\dots,X_{j_0-1},X_{j_0+1},\dots,X_n).&\end{aligned}\tag{5.2.8}$$

(La notazione $\widehat{X_j}$ sta per "manca la colonna X_j".) Per l'ipotesi induttiva abbiamo che

$$\begin{aligned}\mathrm{Det}_{n-1}(X_1,\dots,X_{j-1},\widehat{X_j},X_{j+1},\dots,X_{j_0-1},\lambda Y+\mu Z,X_{j_0+1},\dots,X_n)&=\\ \lambda\,\mathrm{Det}_{n-1}(X_1,\dots,X_{j-1},\widehat{X_j},X_{j+1},\dots,X_{j_0-1},Y,X_{j_0+1},\dots,X_n)&+\\ +\mu\,\mathrm{Det}_{n-1}(X_1,\dots,X_{j-1},\widehat{X_j},X_{j+1},\dots,X_{j_0-1},Z,X_{j_0+1},\dots,X_n).&\end{aligned}\tag{5.2.9}$$

Sostituendo nella (5.2.8) l'espressione della (5.2.9) otteniamo che vale (5.2.7). □

Definizione 5.2.7. Se $V_1,\dots,V_n$ e W sono spazi vettoriali su $\mathbb{K}$, $\mathrm{MultLin}(V_1\times\dots\times V_n,W)$ è l'insieme delle applicazioni multilineari $\Phi\colon V_1\times\dots\times V_n\to W$.

Se $\Phi_1,\Phi_2\in\mathrm{MultLin}(V_1\times\dots\times V_n,W)$ definiamo

$$\begin{array}{ccc}V_1\times\dots\times V_n & \xrightarrow{\Phi_1+\Phi_2} & W\\ (v_1,\dots,v_n) & \mapsto & \Phi_1(v_1,\dots,v_n)+\Phi_2(v_1,\dots,v_n)\end{array}$$

Analogamente, se $\Phi\in\mathrm{MultLin}(V_1\times\dots\times V_n,W)$ e $\lambda\in\mathbb{K}$ definiamo

$$\begin{array}{ccc}V_1\times\dots\times V_n & \xrightarrow{\lambda\Phi} & W\\ (v_1,\dots,v_n) & \mapsto & \lambda\Phi(v_1,\dots,v_n)\end{array}$$

Proposizione 5.2.8. *Siano $V_1,\dots,V_n$ e W spazi vettoriali su $\mathbb{K}$. Se $\Phi_1,\Phi_2\in\mathrm{MultLin}(V_1\times\dots\times V_n,W)$ allora $(\Phi_1+\Phi_2)\in\mathrm{MultLin}(V_1\times\dots\times V_n,W)$. Se $\Phi\in\mathrm{MultLin}(V_1\times\dots\times V_n,W)$ e $\lambda\in\mathbb{K}$, allora $\lambda\Phi\in\mathrm{MultLin}(V_1\times\dots\times V_n,W)$. Con queste operazioni $\mathrm{MultLin}(V_1\times\dots\times V_n,W)$ è uno spazio vettoriale su $\mathbb{K}$.*

Dimostrazione. La somma $(\Phi_1+\Phi_2)$ è multilineare perchè la somma di applicazioni lineari è lineare, e $\lambda\Phi$ è multilineare perchè il prodotto di uno scalare per un'applicazione lineare è lineare. La facile verifica dell'ultima affermazione è lasciata al lettore. □

Esempio 5.2.9. Siano $V_1,\dots,V_n$ spazi vettoriali su $\mathbb{K}$, e siano $\overline{v}_1,\dots,\overline{v}_n\in V$. L'applicazione

$$\begin{array}{ccc}\mathrm{MultLin}(V_1\times\dots\times V_n,\mathbb{K}) & \longrightarrow & \mathbb{K}\\ \Phi & \mapsto & \Phi(\overline{v}_1,\dots,\overline{v}_n)\end{array}$$

è lineare. Infatti la linearità segue subito dalla definizione di somma di applicazioni multilineari e moltiplicazione di uno scalare per un'applicazione multilineare.

5.3 Applicazioni multilineari alternanti

Da ora in poi ci concentriamo sul caso $V_1 = \ldots = V_n$, cioè consideremo applicazioni $\Phi\colon V^n \to W$, dove V e W sono spazi vettoriali su $\mathbb{K}$.

Definizione 5.3.1. Sia $\Phi\colon V^n \to W$.

(1) Siano $1 \leqslant j < h \leqslant n$; Φ è *alternante nelle entrate* j, h se $\Phi(v_1, \ldots, v_n) = 0$ ogni qualvolta $v_j = v_h$.

(2) Φ è *alternante* se è alternante nelle entrate j, h per ogni $1 \leqslant j < h \leqslant n$.

Esempio 5.3.2. Il prodotto vettoriale $\mathsf{V}(\mathbb{E}^3) \times \mathsf{V}(\mathbb{E}^3) \to \mathsf{V}(\mathbb{E}^3)$ (vedi l'Esempio 5.2.4) è alternante (e bilineare).

Esempio 5.3.3. Consideriamo le applicazioni $\Phi_i\colon \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ definite da

$$\Phi_1(x, y) := 3xy, \quad \Phi_2(x, y) := xy + 1, \quad \Phi_3(x, y) := x^3 - xy^2.$$

La Φ_1 è bilineare ma non alternante, la Φ_3 è alternante ma non bilineare, la Φ_2 non è nè bilineare nè alternante.

Esempio 5.3.4. Sia $A \in M_{n,n}(\mathbb{K})$. L'applicazione

$$\begin{array}{ccc} \mathbb{K}^n \times \mathbb{K}^n & \longrightarrow & \mathbb{K} \\ (X, Y) & \mapsto & X^t \cdot A \cdot Y \end{array} \tag{5.3.1}$$

è bilineare. Siccome

$$\Phi(X, X) = \sum_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} a_{ij} x_i x_j,$$

Φ è alternante se solo se $A^t = -A$ e in aggiunta $a_{ii} = 0$ per ogni $i \in \{1, \ldots, n\}$. Notate che se $\operatorname{char} \mathbb{K} \neq 2$ l'ultima condizione segue dalla prima.

Proposizione 5.3.5. *Sia $\Phi\colon V^n \to W$. Supponiamo che Φ sia multilineare e alternante.*

1. *Sia $1 \leqslant j \leqslant n$: allora*

$$\Phi(v_1, \ldots, v_{j-1}, v_j + \sum_{\substack{1 \leqslant i \leqslant n \\ i \neq j}} \mu_i v_i, v_{j+1}, \ldots, v_n) = \Phi(v_1, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n).$$

2. *Siano $1 \leqslant j < h \leqslant n$: allora*

$$\Phi(v_1, \ldots, v_{j-1}, v_h, v_{j+1}, \ldots, v_{h-1}, v_j, v_{h+1}, \ldots, v_n) = \\ = -\Phi(v_1, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_{h-1}, v_h, v_{h+1}, \ldots, v_n),$$

cioè scambiando due entrate il valore di Φ cambia segno.

Dimostrazione. Dimostriamo che vale (1). Per la multilinearità di Φ abbiamo che

$$\Phi(v_1, \ldots, v_{j-1}, v_j + \sum_{\substack{1 \leqslant i \leqslant n \\ i \neq j}} \mu_i v_i, v_{j+1}, \ldots, v_n) = \\ = \Phi(v_1, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n) + \sum_{\substack{1 \leqslant i \leqslant n \\ i \neq j}} \mu_i \Phi(v_1, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n).$$

D'altra parte ogni addendo della sommatoria è nullo perchè Φ è alternante.

Ora dimostriamo che vale (2). Supponiamo che $n = 2$ (se $n = 1$ non c'è nulla da dimostrare). Siccome Φ è alternante e multilineare abbiamo che

$$0 = \Phi(u + w, u + w) = \Phi(u, u) + \Phi(w, w) + \Phi(u, w) + \Phi(w, u) = \Phi(u, w) + \Phi(w, u).$$

Ora supponiamo che $n > 2$. Siccome la funzione

$$\begin{array}{ccc} V \times V & \longrightarrow & W \\ (u, w) & \mapsto & \Phi(v_1, \ldots, v_{j-1}, u, v_{j+1}, \ldots v_{h-1}, w, v_{h+1}, \ldots, v_n) \end{array} \tag{5.3.2}$$

è chiaramente multilineare alternante, la (2) segue dal caso $n = 2$. □

Corollario 5.3.6. *Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione n e sia $\Phi\colon V^n \to W$ un'applicazione multilineare alternante. Se $v_1, \ldots, v_n \in V$ sono linearmente dipendenti, allora $\Phi(v_1, \ldots, v_n) = 0$.*

Dimostrazione. Per la Proposizione 2.5.9 esiste $j \in \{1, \ldots, n\}$ tale che

$$v_j = \sum_{\substack{1 \leqslant i \leqslant n \\ i \neq j}} \mu_i v_i.$$

Quindi il punto (1) della Proposizione 5.3.5 dà che

$$0 = \Phi(v_1, \ldots, v_{j-1}, 0, v_{j+1}, \ldots, v_n) = \Phi(v_1, \ldots, v_{j-1}, \sum_{\substack{1 \leqslant i \leqslant n \\ i \neq j}} \mu_i v_i, v_{j+1}, \ldots, v_n) = \Phi(v_1, \ldots, v_j, \ldots, v_n).$$

□

Corollario 5.3.7. *Sia $\Phi\colon V^n \to W$. Supponiamo che Φ sia multilineare e che sia alternante nelle entrate j e $j + 1$ per ogni $1 \leqslant j < n$ (in altre parole per entrate adiacenti). Allora Φ è alternante.*

Dimostrazione. Per induzione su n. Per $n = 1$ il Lemma non dice nulla, e per $n = 2$ la tesi è uguale all'ipotesi. Dimostriamo il passo induttivo. Quindi supponiamo che la tesi valga per un $n \geqslant 2$ e dimostriamo che vale con n sostituito da $n + 1$. Dato $u \in V$ le applicazioni

$$\begin{array}{ccc} V^n & \longrightarrow & W \\ (w_1, \ldots, w_n) & \overset{\Psi(u)}{\mapsto} & \Phi(u, w_1, \ldots, w_n) \end{array} \qquad \begin{array}{ccc} V^n & \longrightarrow & W \\ (w_1, \ldots, w_n) & \overset{\Pi(u)}{\mapsto} & \Phi(w_1, \ldots, w_n, u) \end{array}$$

sono multilineari. Inoltre $\Psi(u)$ e $\Pi(u)$ sono alternanti per l'ipotesi induttiva.

Quindi $\Phi(v_1, \ldots, v_{n+1}) = 0$ se $v_i = v_j$ dove $i < j$ e $(i, j) \neq (1, n + 1)$, e rimane da dimostrare che $\Phi(v_1, \ldots, v_{n+1}) = 0$ se $v_1 = v_{n+1}$. Applicando la Proposizione 5.3.5 all'applicazione multilineare alternante $\Psi(v_1)$ vediamo che

$$\Phi(v_1, \ldots, v_{n+1}) = -\Phi(v_1, \ldots, v_{n+1}, v_n). \tag{5.3.3}$$

D'altra parte $\Pi(v_n)$ è alternante e quindi il membro di destra di (5.3.3) è nullo perchè $v_1 = v_{n+1}$. □

Terminologia 5.3.8. Un'applicazione multilineare e alternante $\Phi\colon V^n \to \mathbb{K}$ si dice *forma multilineare alternante (o antisimmetrica)* .

Ora dimostriamo che il determinante è alternante nelle colonne.

Proposizione 5.3.9. *La funzione Det_n è una forma multilineare alternante nelle colonne.*

Dimostrazione. La funzione Det_n è una forma multilineare nelle colonne per la Proposizione 5.2.6. Dobbiamo dimostrare che Det_n è alternante nelle colonne. Per induzione su n. Il caso $n = 1$ è banalmente vero. Dimostriamo il passo induttivo. Siccome Det_n è multilineare nelle colonne (per la Proposizione 5.2.6) il Corollario 5.3.7 ci dice che è sufficiente dimostrare che Det_n è alternante nelle colonne j_0 e $(j_0 + 1)$ dove $1 \leqslant j_0 < n$. Quindi supponiamo che $A_{j_0} = A_{j_0+1}$ e dimostriamo che $\mathrm{Det}_n(A) = 0$. Per $1 \leqslant j \leqslant n$ sia $X_j \in M_{n-1,1}(\mathbb{K})$ la colonna ottenuta eliminando l'ultima entrata di A_j. Si ha che

$$\begin{aligned}\mathrm{Det}_n(A_1, \ldots, A_n) = \\ = \sum_{j_0 \neq j \neq j_0+1} (-1)^{n+j} a_{nj} \,\mathrm{Det}_{n-1}(X_1, \ldots, X_{j-1}, \widehat{X_j}, X_{j+1}, \ldots, X_n) + \\ + (-1)^{n+j_0} a_{n,j_0} \,\mathrm{Det}_{n-1}(X_1, \ldots, X_{j_0-1}, \widehat{X_{j_0}}, X_{j_0+1}, \ldots, X_n) + \\ + (-1)^{n+j_0+1} a_{n,j_0+1} \,\mathrm{Det}_{n-1}(X_1, \ldots, X_{j_1-1}, \widehat{X_{j_1}}, X_{j_1+1}, \ldots, X_n). \quad (5.3.4)\end{aligned}$$

Per l'ipotesi induttiva Det_{n-1} è alternante: per ipotesi $X_{j_0} = X_{j_0+1}$ e perciò

$$\mathrm{Det}_{n-1}(X_1, \ldots, X_{j-1}, \widehat{X_j}, X_{j+1}, \ldots, X_n) = 0, \qquad 1 \leqslant j \leqslant n, \quad j_0 \neq j \neq j_0 + 1.$$

Le n-ple $(X_1, \ldots, X_{j_0-1}, \widehat{X_{j_0}}, X_{j_0+1}, \ldots, X_n)$ e $(X_1, \ldots, X_{j_0}, \widehat{X_{j_0+1}}, X_{j_0+2}, \ldots, X_n)$ sono le stesse. Siccome $((-1)^{n+j_0} a_{n,j_0} + (-1)^{n+j_0+1} a_{n,j_0+1}) = 0$ segue che è nulla anche la somma dei restanti due termini nel membro di destra di (5.3.4). □

Per capire le proprietà del determinante studieremo l'insieme di tutte le forme multilineari alternanti $V^n \to \mathbb{K}$ nel caso in cui $\dim V = n$.

Definizione 5.3.10. Se V è uno spazio vettoriale su $\mathbb{K}$, $\mathrm{Alt}(V^n, \mathbb{K})$ è l'insieme delle forme multilineari alternanti $\Phi\colon V^n \to \mathbb{K}$.

Chiaramente $\mathrm{Alt}(V^n, \mathbb{K})$ è un sottoinsieme di $\mathrm{MultLin}(V^n, \mathbb{K})$. Un attimo di riflessione mostra che vale il seguente risultato.

Proposizione 5.3.11. $\mathrm{Alt}(V^n, \mathbb{K})$ *è un sottospazio vettoriale dello spazio vettoriale (su* $\mathbb{K}$*)* $\mathrm{MultLin}(V^n, \mathbb{K})$.

Il risultato principale di questa sezione è il seguente.

Proposizione 5.3.12. *Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione n, e sia $\mathcal{B} := \{\overline{v}_1, \ldots, \overline{v}_n\}$ una base di V. L'applicazione lineare (vedi l'Esempio 5.2.9)*

$$\begin{array}{ccc} \mathrm{Alt}(V^n, \mathbb{K}) & \xrightarrow{\mathcal{E}_{\mathcal{B}}} & \mathbb{K} \\ \Phi & \mapsto & \Phi(\overline{v}_1, \ldots, \overline{v}_n) \end{array}$$

è un isomorfismo.

Prima di dimostrare la Proposizione 5.3.12 diamo un risultato preliminare.

Lemma 5.3.13. *Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione n. Se $\mathcal{B} = \{v_1, \ldots, v_n\}$ e $\mathcal{C} = \{w_1, \ldots, w_n\}$ sono basi di V, esiste una serie di operazioni elementari di tipo (1) e (2) (cioè scambi di vettori e aggiunta a un vettore di una combinazione lineare dei rimanenti vettori) che iniziano da $\{v_1, \ldots, v_n\}$ e finiscono con $\{\alpha_1 w_1, \ldots, \alpha_n w_n\}$, dove $\alpha_1, \ldots, \alpha_n$ sono scalari non nulli.*

Dimostrazione. Sia $A := M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}_V)$ la matrice del cambiamento di base da $\mathcal{B}$ a $\mathcal{C}$. Quindi, se $A = (a_{ij})$, abbiamo

$$v_j = \sum_{i=1}^{n} a_{ij} w_i,$$

cioè la colonna j-esima di A è la colonna delle coordinate di v_j nella base $\mathcal{C}$. Il lemma equivale all'affermazione che esiste una serie di operazioni elementari di tipo (1) e (2) sulle colonne che inizia con A e finisce con una matrice diagonale. Per la Proposizione 3.7.12 esiste una serie di operazioni elementari di tipo (1) e (2) sulle colonne che inizia con A e finisce con una matrice a scala per colonne

$$B := \begin{bmatrix} b_{11} & 0 & \dots & 0 & 0 \\ * & \dots & \dots & 0 & 0 \\ b_{i1} & \dots & b_{ii} & \dots & \dots \\ * & 0 & \dots & \dots & 0 \\ b_{n1} & * & \dots & * & b_{nn} \end{bmatrix}$$

La matrice A ha rango n perchè è una matrice di cambiamento di base, quindi anche B ha rango n. Segue che tutte le entrate di B sulla diagonale principale (cioè b_{ii} per $i \in \{1, \dots, n\}$)) sono diverse da 0: infatti se una è nulla, siccome B è a scala per colonne, segue che $b_{nn} = 0$, e quindi B non ha rango n, contraddizione. Ma allora esiste una serie di operazioni elementari di tipo (2) sulle colonne che inizia con B e finisce con una matrice diagonale. Infatti aggiungendo alle colonne da 1 a $(n-1)$ opportuni multipli della colonna n arriviamo a una matrice a scala per colonne (di rango n) con entrate nulle sull'ultima riga, con l'eccezione dell'entrata (n, n). Notate che la nuova matrice ha le stesse entrate di B sulla diagonale principale. Poi, aggiungendo alle colonne da 1 a $(n-2)$ opportuni multipli della colonna $n-1$ arriviamo a una matrice a scala per colonne (di rango n) con entrate nulle sulle ultime due righe, con l'eccezione delle entrate $(n-1, n-1)$ e (n, n). Notate che, come prima, la nuova matrice ha le stesse entrate di B sulla diagonale principale. Iterando arriviamo a una matrice diagonale

$$D := \begin{bmatrix} b_{11} & 0 & \dots & 0 & 0 \\ 0 & \dots & \dots & 0 & 0 \\ \dots & \dots & b_{ii} & \dots & \dots \\ 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & \dots & 0 & b_{nn} \end{bmatrix}$$

che ha le stesse entrate di B sulla diagonale principale. □

Dimostrazione della Proposizione 5.3.12. Iniziamo dimostrando che $\mathcal{E}_{\mathcal{B}}$ è iniettiva. Dobbiamo dimostrare che Φ è determinata dal singolo valore $\Phi(\overline{v}_1, \dots, \overline{v}_n)$. Siano $v_1, \dots, v_n \in V$. Se $v_1, \dots, v_n$ sono linearmente dipendenti allora $\Phi(v_1, \dots, v_n) = 0$ per il Corollario 5.3.6. Se $v_1, \dots, v_n$ sono linearmente indipendenti allora per il Lemma 5.3.13 esiste una serie di operazioni elementari di tipo (1) e (2) che iniziano da $\{v_1, \dots, v_n\}$ e finiscono con $\{\alpha_1 \overline{v}_1, \dots, \alpha_n \overline{v}_n\}$. Per la Proposizione 5.3.5 si ha

$$\Phi(v_1, \dots, v_n) = (-1)^s \alpha_1 \cdot \ldots \cdot \alpha_n \Phi(\overline{v}_1, \dots, \alpha_n \overline{v}_n),$$

dove s è il numero di operazioni elementari di tipo (1) (cioè scambio di due vettori) fatte nel passare da $\{v_1, \dots, v_n\}$ a $\{\alpha_1 \overline{v}_1, \dots, \alpha_n \overline{v}_n\}$. Quindi $\Phi(v_1, \dots, v_n)$ è determinata dal singolo valore $\Phi(\overline{v}_1, \dots, \overline{v}_n)$.

Rimane da dimostrare che $\mathcal{E}_{\mathcal{B}}$ è suriettiva. Siccome l'applicazione $X_{\mathcal{B}} \colon V \xrightarrow{\sim} \mathbb{K}^n$ è un isomorfismo e $f(v_i) = e_i$ (dove $e_1, \dots, e_n \in \mathbb{K}^n$ sono i vettori della base standard), è sufficiente dimostrare che l'applicazione

$$\begin{array}{ccc} \mathrm{Alt}((\mathbb{K}^n)^n, \mathbb{K}) & \xrightarrow{\mathcal{E}_S} & \mathbb{K} \\ \Phi & \mapsto & \Phi(e_1, \dots, e_n) \end{array}$$

non è nulla. Scegliamo $\Phi = \mathrm{Det}_n$. Un facile argomento per induzione su n dimostra che $\mathrm{Det}_n(1_n) = 1$. □

Osservazione 5.3.14. Gli argomenti appena dati forniscono il seguente algoritmo per il calcolo del determinante di una matrice $A \in M_{n,n}(\mathbb{K})$. Con una serie di operazioni elementari tipo (1) e (2) sulle colonne passiamo da A a una matrice a scala per colonne B. Se s è il numero di scambi di colonne, allora

$$\mathrm{Det}(A) = (-1)^s b_{1,1} \cdot \ldots \cdot b_{n,n}. \tag{5.3.5}$$

Proposizione 5.3.15. *Una matrice quadrata A è singolare se e solo se* $\mathrm{Det}(A) = 0$.

Dimostrazione. Con una serie di operazioni elementari tipo (1) e (2) sulle colonne passiamo da A a una matrice a scala per colonne B. Il rango di A è uguale al rango di B e quindi la matrice A è singolare se e solo se è nulla almeno una delle entrate di B sulla diagonale principale, cioè $b_{1,1} \dots b_{n,n}$ (notate che siccome B è quadrata questo accade solo se $b_{n,n} = 0$). Perciò la proposizione segue dall'uguaglianza (5.3.5). □

5.4 Determinanti e rango di una matrice

Per la Proposizione 5.3.15 una matrice quadrata $n \times n$ ha rango massimo, cioè n, se e solo se il suo determinante è non nullo. Più generalmente il rango di una matrice è determinato dai determinanti dei suoi minori.

Teorema 5.4.1 (degli orlati)**.** *Una matrice $A \in M_{m,n}(\mathbb{K})$ ha rango r se e solo se esiste una sua sottomatrice B di dimensione $r \times r$ con la proprietà che* $\mathrm{Det}\, B \neq 0$ *e che ogni sottomatrice di A di dimensione $(r+1) \times (r+1)$ contenente B abbia determinante nullo.*

Mostriamo come può essere usato il Teorema degli orlati per calcolare il rango di una matrice, poi passeremo alla dimostrazione.

Esempio 5.4.2. Calcoliamo il rango di $A \in M_{3,4}(\mathbb{Q})$ data da

$$A := \begin{bmatrix} 2 & 1 & 3 & 5 \\ 3 & 0 & 3 & 6 \\ 0 & -1 & -1 & -1 \end{bmatrix}.$$

La sottomatrice delle entrate su righe $2, 3$ e colonne $1, 2$ è

$$B := \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix},$$

e ha determinante non nullo. Esistono due sottomatrici 3×3 contenenti B, sono

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

Siccome entrambe hanno determinante nullo segue che A ha rango 2.

Dimostrazione del Teorema 5.4.1. Siano $i_1 < i_2 \dots < i_r$ e $j_1 < j_2 < \dots < j_r$ gli indici delle righe e delle colonne di B rispettivamente. Le colonne di A contenenti le colonne di B sono linearmente indipendenti. Infatti se esistesse una relazione di dipendenza lineare

$$\lambda_1 A_{j_1} + \lambda_2 A_{j_2} + \dots + \lambda_r A_{j_r} = 0$$

allora varrebbe

$$\lambda_1 B_{j_1} + \lambda_2 B_{j_2} + \dots + \lambda_r B_{j_r} = 0,$$

e questo contraddirebbe l'ipotesi che $\mathrm{Det}\, B \neq 0$. Quindi il rango di A è almeno r. Per finire la dimostrazione basta dimostrare che ogni colonna di A_j di indice diverso da $j_1, j_2, \dots, j_r$ è combinazione lineare delle colonne $A_{j_1}, A_{j_2}, \dots, A_{j_r}$, e siccome queste colonne sono linearmente indipendenti ciò equivale a dimostrare che la matrice M con colonne $A_{j_1}, A_{j_2}, \dots, A_{j_r}, A_j$ ha rango minore di $r+1$. Le righe di M con indici $i_1, i_2 \dots, i_r$ sono linearmente indipendenti perchè $\mathrm{Det}\, B \neq 0$, e quindi basta dimostrare che ogni altra riga di M è una combinazione lineare di queste righe, ovvero che ogni sottomatrice $(r+1) \times (r+1)$ di M che contiene le righe con indici $i_1, i_2 \dots, i_r$ ha determinante nullo. Ma questo vale per ipotesi. □

Una conseguenza immediata del Teorema 5.4.1 è la seguente.

Corollario 5.4.3. *Una matrice $A \in M_{m,n}(\mathbb{K})$ ha rango r se e solo se esiste una sua sottomatrice $r \times r$ di determinante non nullo e ogni sua sottomatrice $(r+1) \times (r+1)$ ha determinante nullo.*

5.5 Binet, Laplace e Cramer

La formula di Binet

Proposizione 5.5.1 (Formula di Binet)**.** *Siano $A, B \in M_{n,n}(\mathbb{K})$. Allora* $\mathrm{Det}(A \cdot B) = \mathrm{Det}(A) \cdot \mathrm{Det}(B)$.

Dimostrazione. Sia $\Phi\colon M_{n,n}(\mathbb{K}) \to \mathbb{K}$ l'applicazione definita da $\Phi(M) := \mathrm{Det}(A \cdot M)$. Dimostriamo che Φ è multilineare e alternante nelle colonne. Sia $i \in \{1, \dots, n\}$, siano $\lambda, \mu \in \mathbb{K}$ e siano $X_i, Y, Z \in M_{n,1}(\mathbb{K})$ matrici colonna, dove $i \in \{1, \dots, i-1, i+1, \dots, n\}$. Allora

$$\begin{aligned}\Phi([X_1, \dots, X_{i-1}, \lambda Y + \mu Z, X_{i+1}, \dots, X_n]) = \mathrm{Det}([A{\cdot}X_1, \dots, A{\cdot}X_{i-1}, A{\cdot}(\lambda Y + \mu Z), A{\cdot}X_{i+1}, \dots, A{\cdot}X_n]) = \\ = \lambda \,\mathrm{Det}([A{\cdot}X_1, \dots, A{\cdot}X_{i-1}, A{\cdot}Y, A{\cdot}X_{i+1}, \dots, A{\cdot}X_n]) + \mu \,\mathrm{Det}([A{\cdot}X_1, \dots, A{\cdot}X_{i-1}, A{\cdot}Z, A{\cdot}X_{i+1}, \dots, A{\cdot}X_n]) = \\ = \lambda\Phi([X_1, \dots, X_{i-1}, Y, X_{i+1}, \dots, X_n]) + \mu\Phi([X_1, \dots, X_{i-1}, Z, X_{i+1}, \dots, X_n]).\end{aligned}$$

(La seconda uguaglianza segue dalla multilinearità della funzione determinante.) Abbiamo dimostrato che Φ è multilineare nelle colonne. Si dimostra che Φ è alternante nelle colonne notando che se le colonne di indici j e k di M sono uguali, allora le colonne di indici j e k di $A \cdot M$ sono uguali.

D'altra parte anche l'applicazione $\Psi\colon M_{n,n}(\mathbb{K}) \to \mathbb{K}$ definita da $\Psi(M) := \mathrm{Det}(A) \cdot \mathrm{Det}(M)$ è multilineare e alternante nelle colonne di M perchè lo è la funzione determinante. Siccome

$$\Phi(1_n) = \mathrm{Det}(A \cdot 1_n) = \mathrm{Det}(A) = \mathrm{Det}(A) \cdot \mathrm{Det}(1_n) = \Psi(1_n),$$

segue dalla Proposizione 5.3.12 (con $V = \mathbb{K}^n$ e $\mathcal{B}$ la base standard) che $\Phi = \Psi$. □

Corollario 5.5.2. *Sia $A \in M_{n,n}(\mathbb{K})$ invertibile cioè con* $\mathrm{Det}\, A \neq 0$ *per l'Osservazione 5.3.14. Allora* $\mathrm{Det}(A^{-1}) = \mathrm{Det}(A)^{-1}$.

Dimostrazione. Per la formula di Binet abbiamo che

$$1 = \mathrm{Det}(1_n) = \mathrm{Det}(A \cdot A^{-1}) = \mathrm{Det}(A) \cdot \mathrm{Det}(A^{-1}).$$

□

Osservazione 5.5.3. La formula di Binet dà che l'applicazione

$$\begin{array}{ccc} \mathrm{GL}_n(\mathbb{K}) & \longrightarrow & \mathbb{K}^* \\ A & \mapsto & \mathrm{Det}(A) \end{array}$$

è un omomorfismo di gruppi (ricordiamo che l'operazione in $\mathbb{K}^*$ è la moltiplicazione).

La formula di Binet ha la seguente importante conseguenza.

Corollario 5.5.4. *Sia V uno spazio vettoriale su $\mathbb{K}$ finitamente generato. Sia $f\colon V \to V$ un endomorfismo. Siano $\mathcal{B}$ e $\mathcal{C}$ basi di V. Allora*

$$\mathrm{Det}(M_{\mathcal{B}}^{\mathcal{B}}(f)) = \mathrm{Det}(M_{\mathcal{C}}^{\mathcal{C}}(f)).$$

Dimostrazione. Le equazioni (3.11.1) e (3.10.2), insieme alla formula di Binet e al Corrollario 5.5.2, danno che

$$\begin{aligned}\mathrm{Det}(M_{\mathcal{B}}^{\mathcal{B}}(f)) = \mathrm{Det}\left(M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V)^{-1} \cdot M_{\mathcal{C}}^{\mathcal{C}}(f) \cdot M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V)\right) = \\ = \mathrm{Det}(M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V))^{-1} \cdot \mathrm{Det}(M_{\mathcal{C}}^{\mathcal{C}}(f)) \cdot \mathrm{Det}(M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V)) = \mathrm{Det}(M_{\mathcal{C}}^{\mathcal{C}}(f)).\end{aligned}$$

□

Definizione 5.5.5. Siano V uno spazio vettoriale su $\mathbb{K}$ finitamente generato e $f\colon V \to V$ un endomorfismo. Il *determinante di* f è

$$\mathrm{Det}(f) := \mathrm{Det}(M_{\mathcal{B}}^{\mathcal{B}}(f))$$

dove $\mathcal{B}$ è un'arbitraria base di V - la definizione ha senso grazie al Corollario 5.5.4.

Proposizione 5.5.6. *Sia V uno spazio vettoriale su $\mathbb{K}$ finitamente generato e sia $f\colon V \to V$ un endomorfismo. Allora* $\mathrm{Det}(f \circ g) = \mathrm{Det}(f) \cdot \mathrm{Det}(g)$.

Dimostrazione. Sia $\mathcal{B}$ una base di V. Per l'equazione (3.6.6) e la Formula di Binet abbiamo

$$\mathrm{Det}(f \circ g) = \mathrm{Det}(M_{\mathcal{B}}^{\mathcal{B}}(f \circ g)) = \mathrm{Det}\left(M_{\mathcal{B}}^{\mathcal{B}}(f) \cdot M_{\mathcal{B}}^{\mathcal{B}}(g)\right) = \\ = \mathrm{Det}\left(M_{\mathcal{B}}^{\mathcal{B}}(f)\right) \cdot \mathrm{Det}\left(M_{\mathcal{B}}^{\mathcal{B}}(g)\right) = \mathrm{Det}(f) \cdot \mathrm{Det}(g).$$

□

Sviluppo di Laplace

La seguente proposizione dà quello che si chiama lo *sviluppo del determinante secondo la riga i-esima*. Nel caso della riga n-esima si tratta della formula che definisce il determinante, cioè (5.1.1).

Proposizione 5.5.7. *Siano $A \in M_{n,n}(\mathbb{K})$ e $1 \leqslant i \leqslant n$. Abbiamo che*

$$\mathrm{Det}(A) := \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \,\mathrm{Det}(A_j^i). \tag{5.5.1}$$

Dimostrazione. Sia $\Phi^i\colon M_{n,n}(\mathbb{K}) \to \mathbb{K}$ la funzione definita ponendo $\Phi^i(A)$ uguale al membro di destra di (5.5.1). Allora Φ^i è multilineare e alternante nelle colonne. Infatti se $i = n$ l'affermazione segue dalle Proposizioni 5.2.6 e 5.3.9, e argomenti del tutto simili danno la dimostrazione per un i qualsiasi. Inoltre un facile argomento induttivo dà che $\Phi^i(1_n) = 1$. Per la Proposizione 5.3.12 segue che l'applicazione Φ^i è uguale all'applicazione Det_n. □

Proposizione 5.5.8. *Sia $A \in M_{n,n}(\mathbb{K})$. Allora* $\mathrm{Det}(A) = \mathrm{Det}(A^t)$.

Dimostrazione. Sia $\Phi\colon M_{n,n}(\mathbb{K}) \to \mathbb{K}$ definita da $\Phi(A) := \mathrm{Det}(A^t)$. Consideriamo la Φ come funzione delle colonne. La Proposizione 5.5.7 dà che Det_n è lineare in ciascuna riga. Siccome le colonne di A sono le righe di A^t segue che Φ è lineare in ciascuna colonna, cioè è multilineare (come funzione delle colonne). Ora dimostriamo che Φ è alterna (come funzione delle colonne). Supponiamo che due colonne di A siano uguali: allora le corrispondenti righe di A^t sono uguali. Quindi le righe di A^t sono linearmente dipendenti e perciò A^t è singolare. Per l'Osservazione 5.3.14 segue che $\mathrm{Det}_n(A^t) = 0$. Questo dimostra che Φ è alternante. D'altra parte $\Phi(1_n) = \mathrm{Det}_n(1_n^t) = \mathrm{Det}_n(1_n) = 1$. Per la Proposizione 5.3.12 segue che l'applicazione Φ è uguale all'applicazione Det_n. □

Osservazione 5.5.9. La Proposizione 5.5.8 dà che il determinante è multilineare e alternante nelle righe (oltre che nelle colonne).

Osservazione 5.5.10. Sia $A \in M_{n,n}(\mathbb{K})$ e supponiamo che $B \in M_{n,n}(\mathbb{K})$ sia ottenuta da A con una serie di operazioni elementari sulle righe di tipo (1) e (2), e che siano stati fatti s scambi di righe. Allora B^t è ottenuta da A^t con una serie di operazioni elementari sulle colonne di tipo (1) e (2), tra cui s scambi di colonne. Per l'Osservazione 5.3.14 e la Proposizione 5.5.8 abbiamo che

$$\mathrm{Det}(A) = \mathrm{Det}(A^t) = (-1)^s \,\mathrm{Det}(B^t) = (-1)^s \,\mathrm{Det}(B). \tag{5.5.2}$$

Notiamo anche che se B è a scala per righe allora il suo determinante è uguale al prodotto delle entrate sulla diagonale principale, questo segue (per esempio) dall'espansione di $\mathrm{Det}(B)$ secondo l'ultima riga. Quindi per calcolare $\mathrm{Det}(A)$ possiamo ridurre A a scala per righe o per colonne, a seconda della convenienza.

La formula seguente si chiama lo *sviluppo del determinante secondo la colonna j-esima.*

Corollario 5.5.11. *Siano $A \in M_{n,n}(\mathbb{K})$ e $1 \leqslant j \leqslant n$. Abbiamo che*

$$\operatorname{Det}(A) := \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \operatorname{Det}(A^i_j). \tag{5.5.3}$$

Dimostrazione. Per la Proposizione 5.5.8 abbiamo che $\operatorname{Det}(A) = \operatorname{Det}(A^t)$. Espandendo $\operatorname{Det}(A^t)$ secondo la riga j (che è uguale alla colonna j-esima di A), vedi (5.5.1), otteniamo (5.5.3). □

Le Formule (5.5.1) e (5.5.3) sono vantaggiose se la matrice di cui vogliamo calcolare il determinante ha molte entrate nulle.

La formula di Cramer

Sia $A \in M_{n,n}(\mathbb{K})$. Siano $1 \leqslant i, j \leqslant n$. Il *cofattore* (o complemento algebrico) di A di indici i, j è

$$A_{ij} := (-1)^{i+j} \operatorname{Det}(A^i_j). \tag{5.5.4}$$

La *matrice dei cofattori di A* (anche matrice aggiunta ma questo termine indica anche una matrice del tutto diversa), denotata A^c, è la trasposta della matrice $n \times n$ con entrate A_{ij}, cioè

$$A^c := \begin{bmatrix} A_{1,1} & A_{2,1} & \dots & \dots & \dots & \dots & A_{n,1} \\ A_{1,2} & A_{2,2} & \ddots & \ddots & \dots & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \dots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \dots & \vdots & \ddots & A_{n-1,n-1} & A_{n,n-1} \\ A_{1,n} & \dots & \dots & \dots & \dots & A_{n-1,n-1} & A_{n,n} \end{bmatrix}.$$

Esempio 5.5.12. Sia

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

Allora

$$A^c := \begin{bmatrix} -6 & -3 & 3 \\ -2 & 3 & -1 \\ 2 & 0 & -2 \end{bmatrix}.$$

Proposizione 5.5.13 (Formula di Cramer)**.** *Sia $A \in M_{n,n}(\mathbb{K})$. Allora*

$$A \cdot A^c = A^c \cdot A = (\operatorname{Det} A) 1_n. \tag{5.5.5}$$

Se A è invertibile, cioè $\operatorname{Det} A \neq 0$*, si ha che $A^{-1} = (\operatorname{Det} A)^{-1} A^c$.*

Dimostrazione. Siano $1 \leqslant i, j \leqslant n$. L'entrata al posto i, j di $A \cdot A^c$ è uguale a

$$\sum_{s=1}^{n} (-1)^{j+s} a_{is} \operatorname{Det}(A^j_s). \tag{5.5.6}$$

Sia $i = j$: lo sviluppo di $\operatorname{Det} A$ secondo la riga i-sima dà che l'entrata al posto i, i di $A \cdot A^c$ è uguale a $\operatorname{Det} A$. Ora supponiamo che $i \neq j$: la (5.5.6) è lo sviluppo secondo la riga j-esima della matrice B ottenuta dalla A sostituendo alla riga j-esima la riga i-esima di A stessa. Siccome B ha le righe i-esima e j-esima uguali è singolare e quindi $\operatorname{Det} B = 0$. Questo dimostra che le entrate di $A \cdot A^c$ che non sono sulla diagonale principale sono nulle e finisce di dimostrare (5.5.6). La formula $A^{-1} = (\operatorname{Det} A)^{-1} A^c$ segue dalla (5.5.6) moltiplicando ambo i membri della prima (o della seconda) uguaglianza per $(\operatorname{Det} A)^{-1}$. □

Esempio 5.5.14. Sia $A \in M_{3,3}(\mathbb{R})$ la matrice dell'Esempio 5.5.12. Applicando la formula di Cramer otteniamo che

$$A^{-1} = \begin{bmatrix} 1 & 1/2 & -1/2 \\ 1/3 & -1/2 & 1/6 \\ -1/3 & 0 & 1/3 \end{bmatrix}.$$

5.6 Determinante e area

Sia $\mathbb{E}^2$ il piano della Geometria euclidea. Introduciamo una unità di misura. Quindi sappiamo misurare l'area di regioni semplici (regioni poligonali, dischi, etc.). Se $\mathbf{T} \subset \mathbb{E}^2$ è una regione di cui sappiamo misurare l'area, denotiamo la sua area con $A(\mathbf{T})$. Dimostreremo che l'area di un parallelogramma è dato da un opportuno determinante. Un parallelogramma è determinato dalla scelta dei suoi vertici, che sono dati da $P_0, P_0 + v, P_0 + w, P_0 + v + w$, dove $P_0 \in \mathbb{E}^2$ e $v, w \in \mathsf{V}(\mathbb{E}^2)$. Esplicitamente tale parallelogramma è dato da

$$\Pi(P_0, v, w) := \{P_0 + sv + tw \mid (s,t) \in [0,1]^2\}. \tag{5.6.1}$$

Notate che se v e w sono linearmente dipendenti allora $\Pi(P_0, v, w)$ è un segmento (o un punto se $v = w = 0$). È conveniente considerarlo un parallelogramma "degenere". Notate che se Q_0 è un altro punto di $\mathbb{E}^2$, allora $\Pi(Q_0, v, w)$ è ottenuto da $\Pi(P_0, v, w)$, aggiungendo $\overrightarrow{P_0Q_0}$ a ciascuno dei suoi punti, cioè traslandolo del vettore $\overrightarrow{P_0Q_0}$. In particolare l'area di $\Pi(P_0, v, w)$ dipende da v, w ma non da P_0.

Proposizione 5.6.1. *Sia $\mathcal{B} = \{\boldsymbol{i}, \boldsymbol{j}\}$ una base di $\mathsf{V}(\mathbb{E}^2)$ tale che $\Pi(P_0, \boldsymbol{i}, \boldsymbol{j})$ abbia area 1. Se $v, w \in \mathsf{V}(\mathbb{E}^2)$ sono dati da*

$$v = a_{11}\boldsymbol{i} + a_{21}\boldsymbol{j}, \qquad w = a_{12}\boldsymbol{i} + a_{22}\boldsymbol{j},$$

allora l'area di $\Pi(P_0, v, w)$ è data dalla formula

$$A(\Pi(P_0, v, w)) = |\operatorname{Det} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}|.$$

Dimostrazione. Sia

$$\begin{array}{ccc} \mathsf{V}(\mathbb{E}^2) \times \mathsf{V}(\mathbb{E}^2) & \xrightarrow{\Psi} & \mathbb{R} \\ (a_{11}\mathbf{i} + a_{21}\mathbf{j}, a_{12}\mathbf{i} + a_{22}\mathbf{j}) & \mapsto & \operatorname{Det} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \end{array}$$

Ora definiamo un'applicazione $\mathsf{V}(\mathbb{E}^2) \times \mathsf{V}(\mathbb{E}^2) \to \mathbb{R}$ legata all'area. Dati $v, w \in \mathsf{V}(\mathbb{E}^2)$ sia

$$\epsilon(v,w) := \begin{cases} +1 & \text{se } \mathcal{C} := \{v, w\} \text{ è una base di } \mathsf{V}(\mathbb{E}^2) \text{ e } \operatorname{Det}(M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id})) > 0, \\ -1 & \text{se } \mathcal{C} := \{v, w\} \text{ è una base di } \mathsf{V}(\mathbb{E}^2) \text{ e } \operatorname{Det}(M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id})) < 0, \\ 0 & \text{se } v, w \text{ sono linearmente dipendenti.} \end{cases}$$

Sia

$$\begin{array}{ccc} \mathsf{V}(\mathbb{E}^2) \times \mathsf{V}(\mathbb{E}^2) & \xrightarrow{\Phi} & \mathbb{R} \\ (v, w) & \mapsto & \epsilon(v,w) \cdot A(\Pi(P_0, v, w)) \end{array}$$

Informalmente Φ associa a v, w l'area "con segno" di $\Pi(P_0, v, w)$. Vogliamo dimostrare che $\Psi = \Phi$. Siccome Ψ è bilineare, alternante e ha valore 1 sulla coppia $(\mathbf{i}, \mathbf{j})$, ci basta (per la Proposizione 5.3.12) dimostrare che anche Φ è bilineare, alternante e ha valore 1 sulla coppia $(\mathbf{i}, \mathbf{j})$. La Φ è alternante perchè $\Pi(P_0, v, v)$ è un segmento e quindi $A(\Pi(P_0, v, v)) = 0$, e $\Phi(\mathbf{i}, \mathbf{j}) = 1$ per la nostra scelta di base $\{\mathbf{i}, \mathbf{j}\}$. La dimostrazione che Φ è bilineare è leggermente più elaborata. Sia $\lambda \in \mathbb{R}$. Con semplici considerazioni geometriche vediamo che

$$\Phi(\lambda v, w) = \lambda\Phi(v, w), \qquad \Phi(v, \lambda w) = \lambda\Phi(v, w), \tag{5.6.2}$$

e che

$$\Phi(v, w + \lambda v) = \Phi(v, w) = \Phi(v + \lambda w, w). \tag{5.6.3}$$

(L'uguaglianza in (5.6.3) corrisponde al fatto che parallelogrammi con stessa base e stessa altezza hanno aree uguali.) Ora dimostriamo che per $v, w_1, w_2 \in \mathsf{V}(\mathbb{E}^2)$ si ha

$$\Phi(v, w_1 + w_2) = \Phi(v, w_1) + \Phi(v, w_2). \tag{5.6.4}$$

Se $v = 0$ tutti i termini in (5.6.4) sono nulli e quindi l'uguaglianza vale. Supponiamo che $v \neq 0$. Se w_1 e w_2 sono multipli di v di nuovo tutti i termini in (5.6.4) sono nulli. Quindi possiamo supporre che $w_1 \neq 0$, e perciò $\{v, w_1\}$ è una base di $\mathsf{V}(\mathbb{E}^2)$. Quindi esistono $\lambda, \mu \in \mathbb{R}$ tali che $w_2 = \lambda w_1 + \mu v$. Per le uguaglianze in (5.6.2) e in (5.6.3) abbiamo

$$\begin{aligned}\Phi(v, w_1 + w_2) = \Phi(v, w_1 + \lambda w_1 + \mu v) = \Phi(v, (1+\lambda)w_1) = (1+\lambda)\Phi(v, w_1) = \Phi(v, w_1) + \lambda\Phi(v, w_1) = \\ = \Phi(v, w_1) + \Phi(v, \lambda w_1) = \Phi(v, w_1) + \Phi(v, \lambda w_1 + \mu v) = \Phi(v, w_1) + \Phi(v, w_2).\end{aligned}$$

Questo dimostra che vale l'uguaglianza in (5.6.4). Quindi Φ è lineare nella seconda entrata. Siccome Φ è alternante segue che è lineare anche nella prima entrata.

□

Osservazione 5.6.2. Sia $g\colon \mathsf{V}(\mathbb{E}^2) \to \mathsf{V}(\mathbb{E}^2)$ l'endomorfismo tale che $g(\mathbf{i}) = v$ e $g(\mathbf{j}) = w$. La Proposizione 5.6.1 equivale all'affermazione che vale l'uguaglianza $A(\Pi(P_0, v, w)) = |\operatorname{Det}(g)|$.

Proposizione 5.6.3. *Sia $F\colon \mathbb{E}^2 \to \mathbb{E}^2$ un'applicazione affine e sia $f\colon \mathsf{V}(\mathbb{E}^2) \to \mathsf{V}(\mathbb{E}^2)$ l'applicazione lineare associata, cioè $f = \mathsf{V}(F)$. Sia $\boldsymbol{T} \subset \mathbb{E}^2$ un parallelogramma. Allora l'area del parallelogramma $F(\boldsymbol{T})$ è uguale all'area di $\boldsymbol{T}$ moltiplicata per $|\operatorname{Det}(f)|$.*

Dimostrazione. Il parallelogramma $\mathbf{T}$ è uguale a $\Pi(P_0, v, w)$ per un opportuno $P_0 \in \mathbb{E}^2$ e opportuni $v, w \in \mathsf{V}(\mathbb{E}^2)$, e quindi $F(\mathbf{T})$ è uguale a $\Pi(F(P_0), f(v), f(w))$. Sia $\mathcal{B} = \{\mathbf{i}, \mathbf{j}\}$ una base di $\mathsf{V}(\mathbb{E}^2)$ tale che $\Pi(P_0, \mathbf{i}, \mathbf{j})$ abbia area 1. Se $g\colon \mathsf{V}(\mathbb{E}^2) \to \mathsf{V}(\mathbb{E}^2)$ è l'endomorfismo tale che $g(\mathbf{i}) = v$ e $g(\mathbf{j}) = w$, allora

$$A(\Pi(P_0, v, w)) = |\operatorname{Det}(g)|$$

per la Proposizione 5.6.1 (vedi l'Osservazione 5.6.2). Siccome $f \circ g(\mathbf{i}) = f(v)$ e $f \circ g(\mathbf{j}) = w$, abbiamo anche che

$$A(\Pi(F(P_0), f(v), f(w))) = |\operatorname{Det}(f \circ g)| = |\operatorname{Det}(f)| \cdot |\operatorname{Det}(g)| = |\operatorname{Det}(f)| \cdot A(\Pi(P_0, v, w)).$$

(La seconda uguaglianza vale per il Teorema di Binet.) □

Risultati analoghi alle Proposizioni 5.6.1 e 5.6.3 valgono per il volume di parallelepipedi nello spazio, lasciamo al lettore il compito di darne la formulazione e la dimostrazione.

5.7 Determinante e permutazioni

Diamo la classica formula chiusa (non iterativa) per il determinante. Per semplificare la notazione poniamo

$$\mathbf{n} := \{1, \dots, n\}.$$

Definizione 5.7.1. Data $\varphi\colon \mathbf{n} \to \mathbf{n}$ sia $M_\varphi \in M_{n,n}(\mathbb{K})$ la matrice con entrate

$$m_{ij} := \begin{cases} 1 & \text{se } i = \varphi(j) \\ 0 & \text{se } j \neq \varphi(i) \end{cases}$$

Equivalentemente M_φ è la matrice la cui colonna j è il vettore $e_{\varphi(j)}$ della base standard.

Ricordiamo che una permutazione di un insieme X è un'applicazione biunivoca $\sigma\colon X \to X$, e che l'insieme delle permutazioni di X provvisto dell'operazione di composizione è un gruppo. L'insieme delle permutazioni di $\mathbf{n}$ (o di un insieme di cardinalità n) è denotato $\mathcal{S}_n$.

Osservazione 5.7.2. La matrice M_φ è non singolare se e solo se φ è biunivoca, cioè è un elemento di $\mathcal{S}_n$. Infatti se φ è biunivoca allora le colonne di M_φ sono i vettori della base standard di $\mathbb{K}^n$ (riordinari), e quindi M_φ ha rango n. D'altra parte, se φ non è biunivoca allora non è iniettiva (perchè dominio e codominio di φ hanno la stessa cardinalità n), e quindi esistono almeno due colonne di M_φ uguali, e perciò il rango di M_φ è minore di n.

Proposizione 5.7.3. *Siano $\mathbb{K}$ un campo e $A = (a_{ij}) \in M_{n,n}(\mathbb{K})$. Allora*

$$\operatorname{Det} A = \sum_{\sigma \in \mathcal{S}_n} \operatorname{Det}(M_{\sigma^{-1}}) a_{1,\sigma(1)} a_{2,\sigma(2)} \dots a_{n,\sigma(n)}. \tag{5.7.1}$$

Dimostrazione. Sia $\{e_1, \dots, e_n\}$ la base standard di $\mathbb{K}^n$. La colonna j di A è uguale a $\sum_{i=1}^n a_{ij} e_i$, e siccome Det è multilineare nelle colonne segue che

$$\begin{aligned}
\operatorname{Det} A &= \operatorname{Det}\left(\sum_{i=1}^n a_{i1} e_i, \dots, \sum_{i=1}^n a_{ij} e_i, \dots, \sum_{i=1}^n a_{in} e_i \right) = \\
&= \sum_{\mathbf{n} \xrightarrow{\varphi} \mathbf{n}} \operatorname{Det}(M_\varphi) a_{\varphi(1),1} a_{\varphi(2),2} \dots a_{\varphi(n),n} = \sum_{\tau \in \mathcal{S}_n} \operatorname{Det}(M_\tau) a_{\tau(1),1} a_{\tau(2),2} \dots a_{\tau(n),n} = \\
&= \sum_{\tau \in \mathcal{S}_n} \operatorname{Det}(M_\tau) a_{1,\tau^{-1}(1)} a_{2,\tau^1(2)} \dots a_{n,\tau^{-1}(n)} = \sum_{\sigma \in \mathcal{S}_n} \operatorname{Det}(M_{\sigma^{-1}}) a_{1,\sigma(1)} a_{2,\sigma(2)} \dots a_{n,\sigma(n)}.
\end{aligned}$$

(La seconda uguaglianza vale perchè, per l'Osservazione 5.7.2, se φ non è biunivoca allora $\operatorname{Det}(M_\varphi) = 0$.) □

Vogliamo capire come calcolare $\operatorname{Det}(M_\sigma)$ per $\sigma \in \mathcal{S}_n$. Il primo punto è il seguente risultato.

Proposizione 5.7.4. *Se σ, τ sono applicazioni da $\boldsymbol{n}$ a $\boldsymbol{n}$, allora*

$$M_{\sigma \circ \tau} = M_\sigma \cdot M_\tau. \tag{5.7.2}$$

Dimostrazione. Sia $\{e_1, \dots, e_n\}$ la base standard di $\mathbb{K}^n$. Siccome la colonna j di M_φ è il vettore $e_{\varphi(j)}$, abbiamo che

$$L_{M_\sigma \cdot M_\tau}(e_j) = L_{M_\sigma}\left(L_{M_\tau}(e_j)\right) = L_{M_\sigma}\left(e_{\tau(j)}\right) = e_{\sigma \circ \tau(j)} = L_{M_{\sigma \circ \tau}}(e_j).$$

Quindi $L_{M_\sigma \cdot M_\tau}$ e $L_{M_{\sigma \circ \tau}}$ hanno gli stessi valori sui vettori della base standard e perciò $L_{M_\sigma \cdot M_\tau} = L_{M_{\sigma \circ \tau}}$. Segue che vale (5.7.2). □

Se $\sigma \in \mathcal{S}_n$, allora segue da Binet e dalla Proposizione 5.7.4 che $\operatorname{Det}(M_\sigma) \neq 0$. Infatti

$$\operatorname{Det}(M_\sigma) \cdot \operatorname{Det}(M_{\sigma^{-1}}) = \operatorname{Det}(M_\sigma \cdot M_{\sigma^{-1}}) = \operatorname{Det}(M_{\sigma \circ \sigma^{-1}}) = \operatorname{Det}(M_{\mathrm{Id}_\mathbf{n}}) = 1.$$

Definizione 5.7.5. Poniamo

$$\begin{array}{ccc} \mathcal{S}_n & \xrightarrow{\epsilon} & \mathbb{K}^* \\ \sigma & \mapsto & \operatorname{Det}(M_\sigma) \end{array} \tag{5.7.3}$$

Proposizione 5.7.6. *L'applicazione $\epsilon \colon \mathcal{S}_n \to \mathbb{K}^*$ è un omomorfismo di gruppi. (L'operazione in $\mathbb{K}^*$ è la moltiplicazione.)*

Dimostrazione. Siano $\sigma, \tau \in \mathcal{S}_n$. Per la Proposizione 5.7.4 e la formula di Binet abbiamo

$$\epsilon(\sigma \circ \tau) = \operatorname{Det}(M_{\sigma \circ \tau}) = \operatorname{Det}(M_\sigma \cdot M_\tau) = \operatorname{Det}(M_\sigma) \cdot \operatorname{Det}(M_\tau) = \epsilon(\sigma) \cdot \epsilon(\tau).$$

□

Definizione 5.7.7. Un elemento $\sigma \in \mathcal{S}_n$ è una *trasposizione* se esistono $a \neq b \in \{1, \dots, n\}$ tali che $\sigma(a) = b$, $\sigma(b) = a$ e $\sigma(i) = i$ per $i \in (\{1, \dots, n\} \backslash \{a, b\})$.

Osservazione 5.7.8. Se $\sigma \in \mathcal{S}_n$ è una trasposizione $\epsilon(\sigma) = -1$. Infatti se $\sigma(a) = b$, $\sigma(b) = a$, allora la colonna a di M_σ è e_b, la colonna b di M_σ è e_a, e se $j \in (\mathbf{n}\backslash\{a, b\})$ la colonna j è e_j. Segue che scambiando le colonne a e b di M_σ otteniamo 1_n, e siccome $\mathrm{Det}(1_n) = 1$ otteniamo che $\mathrm{Det}(M_\sigma) = -1$.

Il seguente risultato mostra come calcolare $\epsilon(\sigma)$ per un qualsiasi $\sigma \in \mathcal{S}_n$.

Proposizione 5.7.9. *Ogni elemento di $\mathcal{S}_n$ è prodotto di trasposizioni. (Per convenzione il prodotto dell'insieme vuoto di permutazioni è l'identità di $\mathcal{S}_n$.)*

Dimostrazione. Per induzione su n. Se $n = 1$ l'affermazione è banalmente vera. (Se vi disturba iniziare dal prodotto dell'insieme vuoto di permutazioni, iniziate da $n = 2$. In questo caso l'affermazione è ancora banalmente vera.)

Dimostriamo il passo induttivo. Siano $n \geqslant 2$ e $\sigma \in \mathcal{S}_n$. Supponiamo che $\sigma(n) = n$. Allora $\sigma(i) \in \mathbf{n\text{-}1}$ per $i \in \mathbf{n\text{-}1}$, e quindi possimao definire $\tau \in \mathcal{S}_{n-1}$ ponendo $\tau(i) = \sigma(i)$ per ogni $i \in \mathbf{n\text{-}1}$. Per ipotesi induttiva $\tau = \alpha_1 \circ \ldots \circ \alpha_l$ dove ciascun α_s è una trasposizione di $\mathcal{S}_{n-1}$. Per $s \in \mathbf{l}$ definiamo $\beta_s \in \mathcal{S}_n$ ponendo

$$\beta_s(i) := \begin{cases} \alpha_s(i) & \text{se } i \in \mathbf{(n\text{-}1)}, \\ n & \text{se } i = n. \end{cases}$$

Ciascuna β_s è una trasposizione di $\mathcal{S}_n$ e $\sigma = \beta_1 \circ \ldots \circ \beta_l$. Ora supponiamo che $\sigma(n) \neq n$. Sia $m := \sigma(n)$, e definiamo $\alpha \in \mathcal{S}_n$ ponendo

$$\alpha(i) := \begin{cases} m & \text{se } i = n, \\ n & \text{se } i = m, \\ i & \text{altrimenti.} \end{cases}$$

La composizione $\alpha \circ \sigma$ manda n in n, e quindi abbiamo appena dimostrato che è un prodotto di trasposizioni:

$$\alpha \circ \sigma = \beta_1 \circ \ldots \circ \beta_l. \tag{5.7.4}$$

Siccome α è una trasposizione $\alpha \circ \alpha = \mathrm{Id}_{\mathbf{n}}$, e perciò moltiplicando ambo i membri di (5.7.4) per α otteniamo la scrittura come prodotto di trasposizioni

$$\sigma = \alpha \circ \beta_1 \circ \ldots \circ \beta_l.$$

□

Osservazione 5.7.10. Sia $\mathcal{S}_n$. Per la Proposizione 5.7.9, la Proposizione 5.7.6 e l'Osservazione 5.7.8 $\epsilon(\sigma) \in \{1, -1\}$ per ogni $\sigma \in \mathcal{S}_n$. Per questo $\epsilon(\sigma)$ si chiama il *segno* di σ. A essere precisi $\epsilon(\sigma)$ dipende dal campo $\mathbb{K}$ perchè è un elemento di $\mathbb{K}$: denotiamolo provvisoriamente $\epsilon_{\mathbb{K}}(\sigma)$. Ragionando un attimo ci si rende conto che $\epsilon_{\mathbb{K}}(\sigma)$ è noto quando è noto $\epsilon_{\mathbb{Q}}(\sigma)$. Più precisamente se $\mathrm{char}\,\mathbb{K} \neq 2$, cioè $1 \neq -1$, allora per così dire vale $\epsilon_{\mathbb{K}}(\sigma) = \epsilon_{\mathbb{Q}}(\sigma)$, mentre se $\mathrm{char}\,\mathbb{K} = 2$, cioè $1 = -1$, allora $\epsilon_{\mathbb{K}}(\sigma) = 1$ per ogni $\sigma \in \mathcal{S}_n$. Una permutazione $\sigma \in \mathcal{S}_n$ è *pari* se $\epsilon_{\mathbb{Q}}(\sigma) = 1$, ed è *dispari* se $\epsilon_{\mathbb{Q}}(\sigma) = -1$.

Esempio 5.7.11. Sia $\sigma \in \mathcal{S}_4$ definita da

$$\sigma(1) = 3, \quad \sigma(2) = 4, \quad \sigma(3) = 1, \quad \sigma(4) = 2.$$

Per calcolare la parità di σ, cioè se σ è pari o dispari, ci basta produrre una serie di scambi che partono dalla successione $\{3, 4, 1, 2\}$ e finiscono con $\{1, 2, 3, 4\}$, e contare il numero di scambi. Esplicitamente

$$\{3, 4, 1, 2\} \rightsquigarrow \{3, 2, 1, 4\} \rightsquigarrow \{1, 2, 3, 4\}.$$

Siccome il numero di scambi è 2 concludiamo che σ è una permutazione pari.

Finalmente possiamo dare la classica formula chiusa per il determinante:

Proposizione 5.7.12. *Siano $\mathbb{K}$ un campo e $A = (a_{ij}) \in M_{n,n}(\mathbb{K})$. Allora*

$$\operatorname{Det} A = \sum_{\sigma \in \mathcal{S}_n} \epsilon(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \dots a_{n,\sigma(n)}, \tag{5.7.5}$$

dove $\epsilon(\sigma) \in \{\pm 1\}$ è dato dalla Definizione 5.7.5.

Dimostrazione. Per la Proposizione 5.7.3 è sufficiente dimostrare che $\operatorname{Det}(M_{\sigma^{-1}}) = \epsilon(\sigma)$. Ma $\operatorname{Det}(M_{\sigma^{-1}}) = \epsilon(\sigma^{-1})$, e siccome l'applicazione $\epsilon\colon \mathcal{S}_n \to \mathbb{K}^*$ è un omomorfismo di gruppi (Proposizione 5.7.6), si ha che $\epsilon(\sigma^{-1}) = \epsilon(\sigma)^{-1}$. Ma $\epsilon(\sigma) \in \{\pm 1\}$ e quindi $\epsilon(\sigma)^{-1} = \epsilon(\sigma)$.

□

5.8 Determinanti con entrate in un anello

Sia R un anello (commutativo con unità). Facciamo la seguente

Ipotesi 5.8.1. R è un sottoanello di un campo $\mathbb{K}$.

Con questa ipotesi la formule di Binet, lo sviluppo di Laplace, la formula di Cramer e l'"ultima" formula per il determinante (la Proposizione 5.7.12) valgono per le matrici di $M_{n,n}(R)$ perchè

$$M_{n,n}(R) \subset M_{n,n}(\mathbb{K}).$$

In particolare valgono per $R = \mathbb{Z}$, perchè $\mathbb{Z}$ è un sottoanello di $\mathbb{Q}$, e per $R = \mathbb{K}[x]$ perchè $\mathbb{K}[x]$ è un sottoanello del campo delle funzioni razionali $\mathbb{K}(x)$.

Proposizione 5.8.2. *Sia R un anello e supponiamo che valga l'Ipotesi 5.8.1. Sia $A \in M_{n,n}(R)$. Esiste $B \in M_{n,n}(R)$ tale che $A \cdot B = B \cdot A = 1_n$ se e solo se $\operatorname{Det} A$ è invertibile in R.*

Dimostrazione. Se esiste B tale che $A \cdot B = B \cdot A = 1_n$, allora la Formula di Binet (che, come abbiamo appena osservato, vale per matrici in $M_{n,n}(R)$) dà che

$$1 = \operatorname{Det}(1_n) = \operatorname{Det}(A \cdot B) = \operatorname{Det}(A) \cdot \operatorname{Det}(B),$$

e quindi $\operatorname{Det} A$ è invertibile in R. Ora supponiamo che $\operatorname{Det} A$ sia invertibile in R, cioè $\operatorname{Det} A \neq 0$ e $\operatorname{Det} A^{-1} \in R$. Sicome la matrice dei cofattori A^c è contenuta in $M_{n,n}(R)$, lo è anche $\operatorname{Det} A^{-1} \cdot A^c \in M_{n,n}(R)$. La Fomula di Cramer (che, come abbiamo appena osservato, vale per matrici in $M_{n,n}(R)$) dà che

$$A \cdot (\operatorname{Det} A^{-1} \cdot A^c) = (\operatorname{Det} A^{-1} \cdot A^c) \cdot A = 1_n.$$

□

5.9 Polinomio caratteristico e diagonalizzazione

Polinomio caratteristico

Nella Sezione 3.12 abbiamo introdotto il problema della determinazione (se esiste) di una base che diagonalizza un endomorfismo di uno spazio vettoriale finitamente generato. Il determinante gioca un ruolo chiave nell'approccio a questo problema. Il punto di partenza è la seguente semplice osservazione.

Osservazione 5.9.1. Siano V uno spazio vettoriale finitamente generato su $\mathbb{K}$ e $f\colon V \to V$ un endomorfismo. Un $\lambda_0 \in \mathbb{K}$ è un autovalore di f se e solo se $\operatorname{Det}(\lambda_0 \operatorname{Id}_V - f) = 0$. Infatti λ_0 è un autovalore di f se e solo se esiste $0 \neq v$ tale che $f(v) = \lambda_0 v$, cioè se e solo se $\ker(\lambda_0 \operatorname{Id}_V - f) \neq \{0\}$ ovvero $\operatorname{Det}(\lambda_0 \operatorname{Id}_V - f) = 0$.

Supponiamo che $\dim V = n$, e sia $\mathcal{B}$ una base di V. Quindi $A := M_{\mathcal{B}}^{\mathcal{B}}(f) \in M_{n,n}(\mathbb{K})$. Per definizione

$$\operatorname{Det}(\lambda \operatorname{Id}_V - f) = \operatorname{Det} M_{\mathcal{B}}^{\mathcal{B}}(\lambda \operatorname{Id}_V - f) = \operatorname{Det}(\lambda 1_n - A).$$

Notiamo che $\lambda 1_n - A$ è una matrice quadrata con entrate nell'anello $\mathbb{K}[\lambda]$, e quindi il suo determinante è un ben definito elemento di $\mathbb{K}[\lambda]$.

Proposizione 5.9.2. *Sia $A \in M_{n,n}(\mathbb{K})$. Allora $\operatorname{Det}(\lambda 1_n - A)$ ha grado n ed è monico, cioè il coefficiente di λ^n è 1.*

Dimostrazione. Abbiamo

$$\lambda 1_n - A = \begin{bmatrix} \lambda - a_{11} & -a_{12} & \dots & \dots & \dots & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \ddots & \ddots & \dots & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \dots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \dots & \vdots & \ddots & \lambda - a_{n-1,n-1} & -a_{n-1,n} \\ -a_{n1} & \dots & \dots & \dots & \dots & -a_{n,n-1} & \lambda - a_{nn} \end{bmatrix} \tag{5.9.1}$$

Per la Proposizione 5.7.12, che vale per determinanti di matrici con entrate in $\mathbb{K}[\lambda]$ (vedi la Sezione 5.8) troviamo che $\operatorname{Det}(\lambda 1_n - A)$ è la somma di $n!$ addendi, ciascuno dei quali è un prodotto di polinomi in λ a coefficienti in $\mathbb{K}$, di grado al più 1, e perciò $\operatorname{Det}(\lambda 1_n - A)$ è un polinomio in λ a coefficienti in $\mathbb{K}$, di grado al più n. Inoltre l'unico addendo che dà il monomio λ^n è $(\lambda - a_{11}) \cdot \ldots \cdot (\lambda - a_{nn})$, e quindi il coefficiente di λ^n è 1. □

Definizione 5.9.3. Se $A \in M_{n,n}(\mathbb{K})$, il *polinomio caratteristico di A* è $\operatorname{Det}(\lambda 1_n - A)$, e si denota p_A.

Proposizione 5.9.4. *Se $A, B \in M_{n,n}(\mathbb{K})$ sono coniugate, allora $p_A = p_B$.*

Dimostrazione. Siccome A, B sono coniugate esiste $G \in \mathrm{GL}_n(\mathbb{K})$ tale che $A = G \cdot A \cdot G^{-1}$. Quindi

$$\begin{aligned} p_A(\lambda) &= \operatorname{Det}(\lambda 1_n - A) = \operatorname{Det}(\lambda 1_n - G \cdot B \cdot G^{-1}) = \operatorname{Det}(G \cdot \lambda 1_n \cdot G^{-1} - G \cdot B \cdot G^{-1}) = \\ &= \operatorname{Det}(G \cdot (\lambda 1_n - B) \cdot G^{-1}) = \operatorname{Det}(G) \cdot \operatorname{Det}(\lambda 1_n - B) \cdot \operatorname{Det}(G^{-1}) = \operatorname{Det}(\lambda 1_n - B) = p_B(\lambda). \end{aligned}$$

□

Definizione 5.9.5. Siano V uno spazio vettoriale finitamente generato di dimensione n, e $f\colon V \to V$ un suo endomorfismo. Allora polinomio caratteristico di $M_{\mathcal{B}}^{\mathcal{B}}(f)$, che non dipende dalla base $\mathcal{B}$ di V per la Proposizione 5.9.2, si chiama il *polinomio caratteristico di f* e si denota p_f.

Osservazione 5.9.6. Siano $A, B \in M_{n,n}(k)$. Se A e B sono coniugate, allora i loro polinomi caratteristici sono uguali. In altre parole; per ogni $i \in \{0, \dots, n\}$ il coefficiente di λ^i del determinante della matrice in (5.9.1) è un polinomio nelle $a_{1,1}, \dots, a_{n,n}$ (omogeneo di grado $(n-i)$) che è **invariante per coniugazione**. Il coefficiente di λ^n è la costante 1, che è invariante ma non dà alcuna informazione. Il coefficiente di λ^0, cioè il termine "costante" (significa che dipende da $a_{1,1}, \dots, a_{n,n}$ ma *non* da λ) è $(-1)^n \operatorname{Det}(A)$, e già sappiamo che è invariante eper coniugazione. Il coefficiente di λ^{n-1} è quella che si chiama la *traccia* di A, e ed è dato da

$$\operatorname{Tr} A := \sum_{i=1}^{n} a_{ii}.$$

Quindi se A e B sono coniugate, allora $\operatorname{Tr}(A) = \operatorname{Tr}(B)$.

Esempio 5.9.7. Siano $A, B \in M_{3,3}(\mathbb{Q})$ date da

$$A := \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, B := \begin{bmatrix} 4 & 1 & 4 \\ 0 & 2 & 3 \\ -2 & 1 & 1 \end{bmatrix}.$$

Siccome $\mathrm{Tr}(A) = 15$ e $\mathrm{Tr}(B) = 7$, A non è coniugata a B. Notate che $\mathrm{Det}(A) = 6 = \mathrm{Det}(B)$.

Diagonalizzazione

Siano V uno spazio vettoriale su $\mathbb{K}$ di dimensione n e $f\colon V \to V$ un suo endomorfismo. Supponiamo di voler capire se f è diagonalizzabile e, nel caso lo sia, determinare una base che diagonalizza f. Siccome gli autovalori di f sono le radici del polinomio caratteristico di f (per l'Osservazione 5.9.1), il primo passo da fare è calcolare il polinomio caratteristico di f. Assumiamo di essere in grado di calcolarne le radici di p_f. Il seguente risultato ci dà una condizione necessaria perchè f sia diagonalizzabile.

Proposizione 5.9.8. *Sia V uno spazio vettoriale finitamente generato su $\mathbb{K}$, di dimensione n. Se un endomorfismo $f\colon V \to V$ è diagonalizzabile, allora P_f ha n radici (contate con molteplicità) in $\mathbb{K}$, cioè esistono $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tali che $P_f = \prod_{i=1}^{n}(\lambda - \lambda_i)$.*

Dimostrazione. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V che diagonalizza f, cioè esistono $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tali che $f(v_i) = \lambda_i v_i$. Sia $A := M_{\mathcal{B}}^{\mathcal{B}}(f)$. La matrice A è diagonale, più precisamente $A = (\lambda_i \delta_{ij})$. Quindi

$$p_f = \mathrm{Det}(\lambda 1_n - A) = \mathrm{Det}(((\lambda - \lambda_i)\delta_{ij})) = \prod_{i=1}^{n}(\lambda - \lambda_i). \tag{5.9.2}$$

□

L'Esempio 3.12.10 dimostra che *non* vale il viceversa della Proposizione 5.9.8, cioè *non* è vero che se p_f è prodotto di fattori lineari allora f è diagonalizzabile; infatti il polinomio caratteristico della matrice N dell'Esempio 3.12.10 ha polinomio caratteristico λ^2, che è un prodotto di fattori lineari, ma N non è diagonalizzabile.

Proposizione 5.9.9. *Sia V uno spazio vettoriale finitamente generato su $\mathbb{K}$. Siano $f\colon V \to V$ un endomorfismo e $\lambda_0 \in \mathbb{K}$ un autovalore di f. Allora*

$$1 \leqslant \dim V_{\lambda_0}(f) \leqslant \mathrm{mult}_{\lambda_0} p_f. \tag{5.9.3}$$

Dimostrazione. La diseguaglianza $1 \leqslant \dim V_{\lambda_0}(f)$ vale per definizione di autovalore. Per dimostrare la seconda diseguaglianza poniamo $r := \dim V_{\lambda_0}(f)$. Estendiamo una base $\{v_1, \dots, v_r\}$ di $V_{\lambda_0}(f)$ a una base $\mathcal{B} = \{v_1, \dots, v_n\}$ di V. Quindi $f(v_i) = \lambda_0 v_i$ per $i \leqslant r$. Abbiamo che

$$M_{\mathcal{B}}^{\mathcal{B}}(\lambda \,\mathrm{Id}_V - f) = \begin{bmatrix} (\lambda-\lambda_0) & 0 & \dots & 0 & * & \dots & * \\ 0 & (\lambda-\lambda_0) & \dots & \vdots & \vdots & \vdots & \vdots \\ \vdots & 0 & \dots & 0 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \dots & (\lambda-\lambda_0) & \vdots & \vdots & \vdots \\ \vdots & \vdots & \dots & 0 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & * & \dots & * \end{bmatrix} \tag{5.9.4}$$

dove il numero di colonne in cui appare $(\lambda - \lambda_0)$ è uguale a r. Sviluppando il determinante secondo la prima colonna e iterando troviamo che $p_f = (\lambda - \lambda_0)^r \cdot q$ dove $q \in \mathbb{K}[\lambda]$. Segue che

$$\dim V_{\lambda_0}(f) = r \leqslant \mathrm{mult}_{\lambda_0} p_f.$$

. □

Corollario 5.9.10. *Sia V uno spazio vettoriale finitamente generato su $\mathbb{K}$ e sia $f\colon V \to V$ un endomorfismo. Allora*

$$\sum_{\lambda\in\mathbb{K}} \dim V_\lambda(f) \leqslant \dim V. \tag{5.9.5}$$

(Se $\mathbb{K}$ è un campo infinito la sommatoria sulla sinistra ha un insieme infinito di indici, ma la somma ha senso perchè $\dim V_\lambda(f) > 0$ solo se λ è un autovalore di f, e quindi per un insieme di indici di cardinalità al più $\dim V$ per l'Osservazione 5.9.1 e la Proposizione 5.9.2.)

Dimostrazione. Per la Proposizione 5.9.9, la disequazione (1.7.9) e la Proposizione 5.9.2 abbiamo che

$$\sum_{\lambda\in\mathbb{K}} \dim V_\lambda(f) \leqslant \sum_{\lambda\in\mathbb{K}} \mathrm{mult}_\lambda\, p_f \leqslant \deg p_f = \dim V. \tag{5.9.6}$$

□

Definizione 5.9.11. Sia V uno spazio vettoriale finitamente generato su $\mathbb{K}$ e sia $f\colon V \to V$ un endomorfismo. La *molteplicità geometrica* di $\lambda_0 \in \mathbb{K}$ è la dimensione di $V_{\lambda_0}(f)$.

Quindi la Proposizione 5.9.9 afferma che la molteplicità geometrica è al più uguale alla molteplicità algebrica. La proposizione che segue è il principale risultato di questa sezione.

Proposizione 5.9.12. *Sia V uno spazio vettoriale finitamente generato su $\mathbb{K}$. Un endomorfismo $f\colon V \to V$ è diagonalizzabile se e solo se vale una dellke seguenti condizioni:*

1. *Il numero di radici di p_f in $\mathbb{K}$ (contate con molteplicità) è uguale al grado di p_f, e per ogni autovalore λ di f si ha che*
$$\dim V_\lambda(f) = \mathrm{mult}_\lambda(p_f). \tag{5.9.7}$$
2. *Vale l'eguaglianza in* (5.9.5)*, cioè $\sum_{\lambda\in\mathbb{K}} \dim V_\lambda(f) = \dim V$.*

La dimostrazione della Proposizione 5.9.12 segue la dimostrazione di un risultato preliminare.

Lemma 5.9.13. *Siano V uno spazio vettoriale su $\mathbb{K}$ e $f\colon V \to V$ un endomorfismo. Se $v_1, \ldots, v_d \in V$ sono autovettori con autovalori distinti allora $v_1, \ldots, v_d$ sono linearmente indipendenti.*

Dimostrazione. Per induzione su d. Se $d = 1$ il risultato è vero perchè per definizione un autovettore è non nullo. Dimostriamo il passo induttivo. Sia $d > 1$. Supponiamo che $v_1, \ldots, v_d$ siano linearmente dipendenti. Quindi esistono $\alpha_1, \ldots, \alpha_d \in \mathbb{K}$ non tutti nulli tali che

$$0 = \alpha_1 v_1 + \ldots + \alpha_d v_d. \tag{5.9.8}$$

In verità *ciascun* α_i è non nullo perchè se un α_i si annullasse avremmo una relazione di dipendenza lineare tra una lista di autovettori con autovalori associati distinti contenente meno di d elementi, contro l'ipotesi induttiva. Siano $\lambda_1, \ldots, \lambda_d$ gli autovalori rispettivamente di $v_1, \ldots, v_d$. Applicando f otteniamo

$$0 = f(0) = f(\alpha_1 v_1 + \ldots + \alpha_d v_d) = \alpha_1\lambda_1 v_1 + \ldots + \alpha_d\lambda_d v_d. \tag{5.9.9}$$

Da questa relazione segue che nessun λ_i è nullo, perchè altrimenti avremmo una relazione di dipendenza lineare tra una lista di $(d-1)$ autovettori con autovalori associati distinti, contro l'ipotesi induttiva. Moltiplicando (5.9.9) per λ_d^{-1} otteniamo che

$$0 = \alpha_1\lambda_d^{-1}\lambda_1 v_1 + \ldots + \alpha_{d-1}\lambda_d^{-1}\lambda_{d-1} v_d + \alpha_d v_d. \tag{5.9.10}$$

Sottraendo (5.9.10) da (5.9.8) si ha che

$$\alpha_1(1 - \lambda_d^{-1}\lambda_1)v_1 + \ldots + \alpha_{d-1}(1 - \lambda_d^{-1}\lambda_{d-1})v_{d-1} = 0. \tag{5.9.11}$$

Sia $1 \leqslant i \leqslant (d-1)$. Siccome $\alpha_i \neq 0$ e, siccome $\lambda_1, \ldots, \lambda_d$ sono distinti abbiamo anche che $(1-\lambda_d^{-1}\lambda_i) \neq 0$. Quindi $\alpha_i(1 - \lambda_d^{-1}\lambda_i) \neq 0$. Per (5.9.11) segue che $v_1, \ldots, v_{d-1}$ sono linearmente dipendenti, e questo contraddice l'ipotesi induttiva. Segue che $v_1, \ldots, v_d$ sono linearmente indipendenti. □

Dimostrazione della Proposizione 5.9.12. Iniziamo osservando che le condizioni (1) e (2) sono equivalenti. Infatti se vale (1) allora vale (2) perchè

$$\sum_{\lambda\in\mathbb{K}} \mathrm{mult}_\lambda(p_f) = \deg p_f = \dim V,$$

e d'altra parte se vale (2) allora $\dim V_\lambda(f) = \mathrm{mult}_\lambda(p_f)$ per la Proposizione 5.9.9.

Ora supponiamo che f sia diagonalizzabile, e sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base che diagonalizza f. Sia λ_i l'autovalore associato a v_i, cioè $f(v_i) = \lambda_i v_i$. Il numero di radici di p_f in $\mathbb{K}$ (contate con molteplicità) è uguale alla dimensione di V per la Proposizione 5.9.8. L'espressione di p_f data da (5.9.2) mostra che

$$\mathrm{mult}_{\lambda_j}(p_f) = |\{1 \leqslant i \leqslant n \mid \lambda_i = \lambda_j\}|. \tag{5.9.12}$$

Siccome ogni v_i tale che $\lambda_i = \lambda_j$ appartiene a $V_{\lambda_j}(f)$ vediamo anche che $\dim V_{\lambda_j}(f) \geqslant \mathrm{mult}_{\lambda_j}(p_f)$, e quindi si ha equaglianza per la Proposizione 5.9.9. Abbiamo dimostrato che se f è diagonalizzabile vale (1), e quindi anche (2) perchè (1) e (2) sono equivalenti.

Ora supponiamo che valga (1) e dimostriamo che f è diagonalizzabile. Siano $\lambda_1, \ldots, \lambda_d$ gli autovalori *distinti* di f. Per $1 \leqslant i \leqslant d$ sia

$$\{v_{i,1}, \ldots, v_{i,n(i)}\}$$

una base di $V_{\lambda_i}(f)$ (quindi $n(i) = \dim V_{\lambda_i}(f)$). Dimostriamo che

$$\{v_{1,1}, \ldots, v_{1,n(1)}, \ldots, v_{i,1}, \ldots, v_{i,n(i)}, \ldots v_{d,1}, \ldots, v_{d,n(d)}\} \tag{5.9.13}$$

è una base di V. Applicando il Lemma 5.9.13 si vede che i vettori di (5.9.13) sono linearmente indipendenti, d'altra parte il loro numero è

$$n(1) + n(2) + \ldots + n(d) = \sum_{\lambda\in\mathbb{K}} \dim V_\lambda(f) = \sum_{\lambda\in\mathbb{K}} \mathrm{mult}_\lambda(p_f) = \deg p_f = \dim V.$$

(La seconda uguaglianza segue da (5.9.7), la terza dall'ipotesi che il numero di radici di p_f in $\mathbb{K}$ (contate con molteplicità) è uguale al grado di p_f.) Segue che (5.9.13) è una base di V. Siccome i vettori della base (5.9.13) sono autovettori di f la f è diagonalizzabile. □

Corollario 5.9.14. *Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione n Sia $f\colon V \to V$ un endomorfismo. Se p_f ha n radici distinte (in $\mathbb{K}$) allora f è diagonalizzabile.*

La dimostrazione della Proposizione 5.9.12 dà una procedura per trovare una base che diagonalizza un endomorfismo diagonalizzabile, *purchè si sappiano determinare le radici del polinomio caratteristico.* (Il significato dell'ultima frase meriterebbe un commento ma tralasciamo.) Illustraimo la procedura con qualche esempio.

Esempio 5.9.15. Sia $A \in M_{2,2}(\mathbb{Q})$ data da

$$A := \begin{bmatrix} 1 & 5 \\ -\frac{1}{4} & -2 \end{bmatrix}.$$

Decidiamo se A è diagonalizzabile. Il polinomio caratteristico di A è

$$p_A = \begin{vmatrix} 1 & 5 \\ -\frac{1}{4} & -2 \end{vmatrix} = \lambda^2 + \lambda - \frac{3}{4}.$$

Le radici di p_A sono $\lambda_1 = -\frac{3}{2}$ e $\lambda_2 = \frac{1}{2}$, e quindi A è diagonalizzabile per il Corollario 5.9.14. Troviamo una base che diagonalizza A. Autovettori con autovalori λ_1 e λ_2 sono dati dalle soluzioni non banali dei sistemi di equazioni lineari

$$\begin{bmatrix} 1+\frac{3}{2} & 5 \\ -\frac{1}{4} & -2+\frac{3}{2} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0, \qquad \begin{bmatrix} 1-\frac{1}{2} & 5 \\ -\frac{1}{4} & -2-\frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

Risolvendo vediamo che $\mathcal{B} := \{(2,-1),(10,-1)\}$ è una base di autovettori di A. In altre parole

$$M_{\mathcal{B}}^{\mathcal{B}}(L_A) = \begin{bmatrix} -\frac{3}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

Ora usiamo questo risultato per calcolare A^{10}. Sia $\mathcal{S}$ la base standard di $\mathbb{Q}^2$. Abbiamo

$$A = M_{\mathcal{S}}^{\mathcal{S}}(L_A) = M_{\mathcal{S}}^{\mathcal{B}}(\mathrm{Id}_{\mathbb{Q}^2}) \cdot M_{\mathcal{B}}^{\mathcal{B}}(L_A) \cdot M_{\mathcal{B}}^{\mathcal{S}}(\mathrm{Id}_{\mathbb{Q}^2}) = \begin{bmatrix} 2 & 10 \\ -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} -\frac{3}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 2 & 10 \\ -1 & -1 \end{bmatrix}^{-1}.$$

Quindi

$$A^{10} = \begin{bmatrix} 2 & 10 \\ -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} \frac{3^{10}}{2^{10}} & 0 \\ 0 & \frac{1}{2^{10}} \end{bmatrix} \cdot \begin{bmatrix} 2 & 10 \\ -1 & -1 \end{bmatrix}^{-1}.$$

Calcolando troviamo che

$$\begin{bmatrix} 2 & 10 \\ -1 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} -\frac{1}{8} & -\frac{5}{4} \\ \frac{1}{8} & \frac{1}{4} \end{bmatrix},$$

e quindi

$$A^{10} = \begin{bmatrix} 2 & 10 \\ -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} \frac{3^{10}}{2^{10}} & 0 \\ 0 & \frac{1}{2^{10}} \end{bmatrix} \cdot \begin{bmatrix} -\frac{1}{8} & -\frac{5}{4} \\ \frac{1}{8} & \frac{1}{4} \end{bmatrix} = \begin{bmatrix} \frac{5-3^{10}}{2^{12}} & \frac{5-3^{10}\cdot 5}{2^{11}} \\ \frac{1}{2^{12}} & \frac{3^{10}\cdot 5-1}{2^{12}} \end{bmatrix}.$$

(Non conviene convertire in notazione decimale.)

Esempio 5.9.16. Sia $A \in M_{3,3}(\mathbb{Q})$ data da

$$A := \begin{bmatrix} 0 & -4 & -1 \\ 3 & 13 & 3 \\ -12 & -48 & -11 \end{bmatrix}.$$

Ci chiediamo se A sia diagonalizzabile. Il polinomio caratteristico di A è $p_A = \lambda(\lambda-1)^2$, e quindi il numero delle sue radici contate con molteplicità è uguale a 3, cioè la dimensione di $\mathbb{Q}^3$. Siccome la radice 1 ha molteplicità 2, A è diagonalizzabile se e solo se $V_1(L_A)$ ha dimensione 2. L'autospazio $V_1(L_A)$ è il sottospazio delle soluzione del sistema di equazioni lineari

$$A := \begin{bmatrix} -1 & -4 & -1 \\ 3 & 12 & 3 \\ -12 & -48 & -12 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0. \tag{5.9.14}$$

Siccome la matrice 3×3 che appare in (5.9.14) ha rango 1, lo spazio delle soluzioni ha dimensione 2 e quindi A è diagonalizzabile.

5.10 Autovettori miliardari[1]

Il successo di Google è dovuta all'efficienza del suo algoritmo (PageRank) che assegna un grado di importanza ("ranking") a ciascuna pagina del web. L'algoritmo di Sergey Brin e Larry Page, i creatori di Google, riduce il calcolo dei gradi di importanza al calcolo di un autovettore di una matrice quadrata (di dimensioni enormi) di tipo particolare (stocastica). Prima descriveremo la matrice e poi discuteremo come calcolare l'autovettore (non si calcola risolvendo un sistema di milioni di equazioni lineari in milioni di incognite). Per maggiori dettagli potete consultare [2].

Siano $\{1,2,\ldots,n\}$ le pagine web. Il problema è di dare una "giusta" importanza x_i (un numero reale) alla pagina i per ogni $i \in \{1,2,\ldots,n\}$ (ovviamente l'importanza cambia giorno per giorno) sapendo, per ogni $i,j \in \{1,2,\ldots,n\}$, se esiste o non esiste un link dalla pagina j alla pagina i. Per determinare il vettore $(x_1,x_2,\ldots,x_n)$ che dà la classifica delle pagine web seguiamo due principi:

[1] Ringrazio René Schoof per avermi parlato di questo argomento, e per aver condiviso con me i suoi appunti.

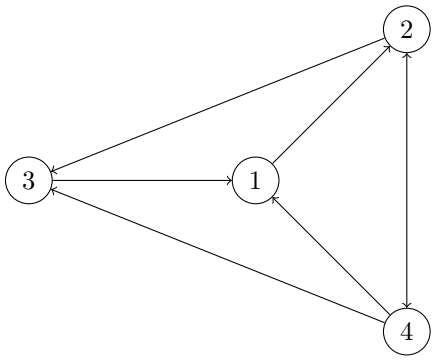

Figura 5.1: Esempio di link tra 4 pagine web

1. i è importante se riceve molti link da pagine importanti (una condizione apparentemente circolare),
2. una pagina web da cui partono molti link non conta più di una pagina web da cui partono pochi link nel formare la classifica.

Sia n_j il numero di link che partono dalla pagina web j. Una formula che ragionevolmente segue dai due principi è che deve valere l'equazione

$$x_i = \sum_{j \mapsto i} \frac{x_j}{n_j}, \tag{5.10.1}$$

dove $j \mapsto i$ significa che la pagina web j ha un link che arriva alla pagina web i. Infatti l'equazione in (5.10.1) dice che la pagina web j contribuisce al punteggio di i proporzionalmente al suo punteggio normalizzato in base al numero di link che partono dalla pagina web j. Sia $A = (a_{ij}) \in M_{n,n}(\mathbb{Q})$ la matrice definita da

$$a_{ij} := \begin{cases} \frac{1}{n_j} & \text{se } j \mapsto i, \\ 0 & \text{altrimenti} \end{cases} \tag{5.10.2}$$

L'equazione in (5.10.1) mostra che $(x_1, x_2, \ldots, x_n)$ è un autovettore di A con autovalore 1.

Esempio 5.10.1. L'esempio di link tra 4 pagine web nella Figura 5.1 dà la matrice

$$A := \begin{bmatrix} 0 & 0 & 1 & 1/3 \\ 1 & 0 & 0 & 1/3 \\ 0 & 1/2 & 0 & 1/3 \\ 0 & 1/2 & 0 & 0 \end{bmatrix}.$$

L'autospazio di autovalore 1 è generato da $(5, 6, 4, 3)$. Quindi la pagina web più importante è la 2, seguita dalla 5, la terza è la 3 e l'ultima è la 4. Notate che le prime tre hanno lo stesso numero di pagine che hanno un link verso di loro.

Ora sorgono varie domande. Data la matrice A, esiste un autovettore di autovalore 1? L'autospazio $V_1(A)$ ha dimensione 1? Supponendo la risposta sia affermativa, esiste un generatore di $V_1(A)$ con entrate non negative? (Se un autovettore $X = (x_1, \ldots, x_n)$ di autovalore 1 ha entrate negative e positive, quale "ranking" scegliamo, X o $-X$?) Infine, supponendo che le risposte alle domande precedenti siano affermative, qual'è il metodo più efficiente per calcolare un autovettore di autovalore 1? (Siccome la matrice è enorme l'eliminazione di Gauss prende troppo tempo.)

Proposizione 5.10.2. *Sia $A \in M_{n,n}(\mathbb{K})$ una matrice tale che la somma delle entrate di ciascuna colonna è uguale a 1, cioè*

$$\sum_{i=1}^{n} a_{ij} = 1 \tag{5.10.3}$$

per ogni $j \in \{1, \ldots, n\}$. Allora esiste un autovettore di A di autovalore 1.

Dimostrazione. Il vettore $(1,\dots,1)$ è un autovettore di autovalore 1 della trasposta A^t per l'uguaglianza in (5.10.3). Ma il polinomio caratteristico di A è uguale a quello di A^t perchè

$$p_{A^t}(\lambda) = \mathrm{Det}(\lambda 1_n - A^t) = \mathrm{Det}((\lambda 1_n - A^t)^t) = \mathrm{Det}(\lambda 1_n - A) = p_A(\lambda).$$

Quindi 1 è un autovalore di A. □

In generale non è vero che l'autospazio $V_1(A)$ ha dimensione 1, anche se facciamo l'ipotesi che $a_{ij} \in \mathbb{Q}$ con $a_{ij} \geqslant 0$ per ogni i, j (come nel nostro caso). Per esempio l'autospazio $V_1(A)$ ha dimensione almeno 2 se A è una matrice a blocchi

$$\begin{bmatrix} * & \dots & * & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ * & \dots & * & 0 & \dots & 0 \\ 0 & \dots & 0 & * & \dots & * \\ \vdots & \dots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & * & \dots & * \end{bmatrix}$$

Le domande formulate hanno risposte positive se la somma delle entrate di ciascuna colonna è uguale a 1 e inoltre a_{ij} è un reale positivo per ogni i, j. (Una matrice $A \in M_{n,n}(\mathbb{R})$ è *stocastica per colonne* se ha entrate non negative e la somma delle entrate di ciascuna colonna è uguale a 1.) La matrice definita da (5.10.2) non ha entrate positive, esistono molte entrate nulle. Per questo motivo si approssima A con una matrice che soddisfa tutte le proprietà appena elencate, precisamente $(1-\epsilon)A + \epsilon E_n$, dove ϵ è un reale molto piccolo ed E_n è la matrice con entrate tutte uguali a $1/n$. Fatto questo (a costo di far torto, di poco, a qualche pagina web che si troverà retrocessa) si applicano i risultati seguenti.

Proposizione 5.10.3. *Sia $A \in M_{n,n}(\mathbb{R})$ una matrice con entrate positive e tale che la somma delle entrate di ciascuna colonna sia uguale a 1. Allora l'autospazio di A di autovalore 1 ha dimensione 1 ed è generato da un vettore con tutte le entrate positive o nulle.*

Dimostrazione. Sia $X = (x_1, \dots, x_n)$ un autovettore di A di autovalore 1. Supponiamo che esistano indici $k, h \in \{1, \dots, n\}$ tali che tali che $x_k \neq 0 \neq x_h$ e x_k, x_h abbiano segni diversi (in altre parole $x_k \cdot x_h < 0$). Siccome $x_i = \sum_{j=1}^n a_{ij}x_j$ e tutti gli a_{ij} sono positivi, segue che $|x_i| < \sum_{j=1}^n a_{ij}|x_j|$. Ma allora abbiamo che

$$\sum_{i=1}^n |x_i| < \sum_{i=1}^n \sum_{j=1}^n a_{ij}|x_j| = \sum_{j=1}^n \sum_{i=1}^n a_{ij}|x_j| = \sum_{j=1}^n |x_j|,$$

e questo è assurdo. Per la Proposizione 5.10.2 sappiamo che $\dim V_1(A) \geqslant 1$. Supponiamo che $\dim V_1(A) > 1$. Siano $X, Y \in V_1(A)$ linearmente indipendenti. Siccome la matrice $2 \times n$ con righe X e Y ha rango 2, esiste una sottomatrice 2×2 che ha rango 2, diciamo che sia

$$\begin{bmatrix} x_k & x_h \\ y_k & y_h \end{bmatrix}.$$

Segue che esistono $\lambda, \mu \in \mathbb{R}$ tali che $\lambda(x_k, x_h) + \mu(y_k, y_h) = (1, -1)$. Ma allora l'autovettore di autovalore 1 dato da $\lambda X + \mu Y$ ha entrate di segni diversi ai posti k e h, e questo contraddice ciò che abbiamo dimostrato. Questo dimostra che $\dim V_1(A) = 1$. Sia X un generatore di $V_1(A)$. Per quello che abbiamo dimostrato o tutte le entrate di X sono non negative o non positive. Nel primo caso abbiamo fatto, nel secondo prendiamo come generatore $-X$. □

I prossimi due risultati danno il metodo con cui si approssima in modo efficiente un generatore di $V_1(A)$. Se $A \in M_{n,n}(\mathbb{R})$ è una matrice con entrate positive e tale che la somma delle entrate di ciascuna colonna è uguale a 1, poniamo

$$c(A) := 1 - 2\min\{a_{ij}\}. \tag{5.10.4}$$

Siccome $\sum_{i=1}^n a_{ij} = 1$ per ogni j e $a_{ij} > 0$ per ogni i, j abbiamo che $0 < \min\{a_{ij}\} \leqslant 1/n$. Quindi

$$\frac{n-2}{n} \leqslant c(A) < 1. \tag{5.10.5}$$

Lemma 5.10.4. *Sia $n \geqslant 2$ e sia $A \in M_{n,n}(\mathbb{R})$ una matrice con entrate positive e tale che la somma delle entrate di ciascuna colonna sia uguale a 1. Sia $X = (x_1, \ldots, x_n)$ un vettore di $\mathbb{R}^n$ con entrate la cui somma è nulla, e sia $Y = (y_1, \ldots, y_n) = A \cdot X$. Allora*

$$\sum_{i=1}^{n} |y_i| \leqslant c(A) \sum_{i=1}^{n} |x_i|. \tag{5.10.6}$$

Dimostrazione. Sia $f\colon \mathbb{R}^n \to \mathbb{R}$ data dalla somma delle entrate, cioè $f(X) = \sum_{i=1}^{n} x_i$. Allora $L_A(\ker(f)) \subset \ker(f)$. Infatti se $X \in \mathbb{R}^n$ abbiamo

$$f(L_A(X)) = f(A \cdot X) = \sum_{i,j \in \{1,\ldots,n\}} a_{ij} x_j = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} x_j = \sum_{j=1}^{n} x_j = f(X).$$

Quindi se $f(X) = 0$ allora $f(L_A(X)) = 0$. Ora sia X come nell'enunciato del lemma, cioè $X \in \ker(f)$, e poniamo $Y := A \cdot X$. Allora $Y \in \ker(f)$ per quello che abbiamo appena dimostrato. Se $Y = 0$ allora la diseguaglianza in (5.10.6) vale perchè $c(A) \geqslant 0$. Quindi possiamo supporre che $Y \neq 0$. Sia ϵ_i il segno di y_i (se $y_i = 0$ allora $\epsilon_i = 0$). Siccome $y_i = \sum_{i=1}^{n} a_{ij} x_j$, abbiamo $|y_i| = \epsilon_i \sum_{i=1}^{n} a_{ij} x_j$, e quindi

$$\sum_{i=1}^{n} |y_i| = \sum_{i=1}^{n} \epsilon_i \sum_{j=1}^{n} a_{ij} x_j = \sum_{j=1}^{n} \sum_{i=1}^{n} \epsilon_i a_{ij} x_j. \tag{5.10.7}$$

Siccome $Y \in \ker(f)$ esistono almeno due entrate di Y non nulle e di segni opposti, e quindi dall'uguaglianza $\sum_{i=1}^{n} a_{ij} = 1$ segue che $\sum_{i=1}^{n} \epsilon_i a_{ij} \leqslant c(A)$. Siccome $c(A) \geqslant 0$ (vedi (5.10.5)) segue che $\sum_{i=1}^{n} \epsilon_i a_{ij} x_j \leqslant c(A)|x_j|$, e allora la diseguaglianza in (5.10.6) segue dalle uguaglianze in (5.10.7). □

Corollario 5.10.5. *Sia $A \in M_{n,n}(\mathbb{R})$ una matrice con entrate positive e tale che la somma delle entrate di ciascuna colonna sia uguale a 1. Sia $X = (x_1, \ldots, x_n) \in \mathbb{R}^n$ tale che $f(X) := \sum_{i=1}^{n} x_i \neq 0$. La successione di vettori di $\mathbb{R}^n$ data da*

$$X, A \cdot X, A^2 \cdot X, \ldots, A^m \cdot X, \ldots, \tag{5.10.8}$$

tende a un generatore di $V_1(A)$.

Dimostrazione. Se $n = 1$ il risultato è banale, quindi possiamo assumere che $n \geqslant 2$. Se $Y \in V_1(A)$ è non nullo, allora $Y \notin \ker(f)$ per la Proposizione 5.10.3 (un vettore non nullo in $\ker(f)$ ha almeno due entrate di segni opposti), e perciò

$$\mathbb{R}^n = V_1(A) \oplus \ker(f).$$

Per ipotesi $X \notin \ker(f)$, quindi $X = Y + Z$ dove $Y \in V_1(A)$ è non nullo. Abbiamo

$$A^k \cdot X = A^k \cdot (Y + Z) = A^k \cdot Y + A^k \cdot Z = Y + A^k \cdot Z. \tag{5.10.9}$$

Poniamo $Z(k) = (z(k)_1, z(k)_2, \ldots, z(k)_n) := A^k \cdot Z$. Per il Lemma 5.10.4 abbiamo che

$$|z(k)_1| + \ldots + |z(k)_i| + \ldots + |z(k)_n| \leqslant c(A)^k (|z_1| + \ldots + |z_i| + \ldots + |z_n|).$$

Siccome $c(A) < 1$ (vedi (5.10.5)) segue che $A^k \cdot Z$ converge a 0 per $k \to \infty$, quindi la successione in (5.10.8) tende a Y per $k \to \infty$. Siccome $Z \notin \ker(f)$ abbiamo che $Y \neq 0$. □

5.11 Gruppo lineare reale

Sia V uno spazio vettoriale reale finitamente generato. Se $g \in \mathrm{GL}(V)$ allora $\det(g) \neq 0$ e quindi si hanno due possibilità: il determinante di g è positivo o negativo. Il determinante dell'identità è 1,

quindi positivo, e se $g, h \in \mathrm{GL}(V)$ hanno determinante positivo allora $g \cdot h$ ha determinante positivo (per Binet), e anche g^{-1} (sempre per Binet). Quindi il sottoinsieme

$$\mathrm{GL}^+(V) := \{g \in \mathrm{GL}(V) \mid \det(g) > 0\} \tag{5.11.1}$$

è un sottogruppo di $\mathrm{GL}(V)$. Dimostrermo che si può passare con continuità da ogni elemento di $\mathrm{GL}^+(V)$ a ogni altro elemento di $\mathrm{GL}^+(V)$, ma che non si può passare con continuità da un elemento di $\mathrm{GL}^+(V)$ a un elemento di $(\mathrm{GL}(V) \setminus \mathrm{GL}^+(V))$. Per dare un senso all'affermazione va definito cosa intendiamo per funzione continua da un intervallo $I \subset \mathbb{R}$ a $\mathrm{GL}(V)$.

Definizione 5.11.1. Sia $I \subset \mathbb{R}$ un intervallo. Se V è uno spazio vettoriale reale di dimensione finita n un'applicazione $\gamma\colon I \to \mathrm{GL}(V)$ è *continua* se, scelta una qualsiasi base $\mathcal{B}$ l'applicazione

$$\begin{array}{ccc} I & \longrightarrow & M_{n,n}(\mathbb{R}) \\ t & \mapsto & M^{\mathcal{B}}_{\mathcal{B}}(\gamma(t)) \end{array} \tag{5.11.2}$$

è continua (questo ha senso perchè identifichiamo $M_{n,n}(\mathbb{R})$ con $\mathbb{R}^{n^2}$).

Osservazione 5.11.2. Per assicurarsi che $\gamma\colon I \to \mathrm{GL}(V)$ sia continua è sufficiente verificare che l'applicazione in (5.11.2) sia continua per *una* base $\mathcal{B}$. Infatti se $\mathcal{C}$ è una qualsiasi base allora $M^{\mathcal{C}}_{\mathcal{C}}(\gamma(t)) = M^{\mathcal{B}}_{\mathcal{C}}(\mathrm{Id}_V) \cdot M^{\mathcal{B}}_{\mathcal{B}}(\gamma(t)) \cdot M^{\mathcal{C}}_{\mathcal{B}}(\mathrm{Id}_V)$, quindi le entrate di $M^{\mathcal{C}}_{\mathcal{C}}(\gamma(t))$ sono combinazioni lineari delle entrate di $M^{\mathcal{B}}_{\mathcal{B}}(\gamma(t))$ e perciò funzioni continue.

Definizione 5.11.3. Sia V uno spazio vettoriale reale di dimensione finita. Un automorfismo $g \in \mathrm{GL}(V)$ è *deformabile* in un automorfismo $h \in \mathrm{GL}(V)$ se esiste un'applicazione continua $\gamma\colon [a, b] \to \mathrm{GL}(V)$ tale che $\gamma(a) = g$ e $\gamma(b) = h$.

Esempio 5.11.4. Sia $A \in \mathrm{GL}_n(\mathbb{R})$ e supponiamo che $B \in \mathrm{GL}_n(\mathbb{R})$ sia ottenuta da A aggiungendo alla colonna j, cioè A_j, un multiplo di un'*altra* colonna, diciamo λA_k (dove $k \neq j$):

$$B = [A_1,\ \cdots,\ A_{j-1},\ A_j + \lambda A_k,\ A_{j+1},\ \cdots,\ A_n]\,.$$

Quindi $B \in \mathrm{GL}_n(\mathbb{R})$ e inoltre la matrice

$$\gamma(t) := [A_1,\ \cdots,\ A_{j-1},\ A_j + t\lambda A_k,\ A_{j+1},\ \cdots,\ A_n]$$

è invertibile per ogni $t \in \mathbb{R}$. Siccome l'applicazione $\gamma\colon [0, 1] \to \mathrm{GL}_n(\mathbb{R})$ è continua e $\gamma(0) = A$, $\gamma(0) = B$, questo mostra che A è deformabile in B. Considerando la trasposta vediamo anche che se $B \in \mathrm{GL}_n(\mathbb{R})$ è ottenuta da A aggiungendo a una riga una combinazione lineare delle rimanenti righe, allora A è deformabile con continuità in B.

Esempio 5.11.5. Sia $A \in \mathrm{GL}_n(\mathbb{R})$ e supponiamo che $B \in \mathrm{GL}_n(\mathbb{R})$ sia ottenuta da A scambiando tra di loro due colonne e cambiando segno a una delle due. In altre parole, se le colonne hanno indici j e k, con $j < k$, si ha

$$B = A \cdot [e_1,\ \cdots,\ e_{j-1},\ e_k,\ e_{j+1},\ \cdots,\ e_{k-1},\ -e_j,\ e_{k+1},\ \cdots,\ e_n]\,, \tag{5.11.3}$$

oppure

$$B = A \cdot [e_1,\ \cdots,\ e_{j-1},\ -e_k,\ e_{j+1},\ \cdots,\ e_{k-1},\ e_j,\ e_{k+1},\ \cdots,\ e_n]\,. \tag{5.11.4}$$

Facciamo vedere che A è deformabile in B. Basta far vedere che la matrice unità 1_n è deformabile nella matrice invertibile che è a destra del prodotto in (5.11.3) e del prodotto in (5.11.4). Sia $\alpha\colon [0, \pi/2] \to \mathrm{GL}_n(\mathbb{R})$ l'applicazione continua data da

$$\alpha(t) := [e_1,\ \cdots,\ e_{j-1},\ \cos t e_j + \sin t e_k,\ e_{j+1},\ \cdots,\ e_{k-1},\ -\sin t e_j + \cos t e_k,\ e_{k+1},\ \cdots,\ e_n]\,.$$

Allora $\alpha(0) = 1_n$ e $\alpha(\pi/2)$ è la matrice che è a destra del prodotto in (5.11.3). Analogamente si scrive un'applicazione continua $\beta\colon [0, \pi/2] \to \mathrm{GL}_n(\mathbb{R})$ tale che $\beta(0) = 1_n$ e $\beta(\pi/2)$ è la matrice che è a destra del prodotto in (5.11.4).

Osservazione 5.11.6. Se $I \subset \mathbb{R}$ è un intervallo e $\gamma\colon I \to \mathrm{GL}_n(\mathbb{R})$ è un'applicazione continua, la funzione

$$\begin{array}{ccc} I & \xrightarrow{\gamma} & \mathbb{R} \\ t & \mapsto & \det\gamma(t) \end{array}$$

è continua. Siccome γ non è mai nulla, il segno di γ è costante (Teorema di Bolzano). Ne segue che se $g \in \mathrm{GL}(V)$ è deformabile in $h \in \mathrm{GL}(V)$ allora i segni di $\det g$ e $\det h$ sono gli stessi.

Osservazione 5.11.7. La relazione tra elementi di $\mathrm{GL}(V)$ data dalla Definizione 5.11.3 è di equivalenza, lasciamo al lettore la semplice verifica di questo fatto.

Proposizione 5.11.8. *Sia V uno spazio vettoriale di dimensione finita, e siano $g, h \in \mathrm{GL}(V)$. Allora g è deformabile in h se e solo se sono entrambi elementi di $\mathrm{GL}(V)$ o di $(\mathrm{GL}(V)\setminus \mathrm{GL}^+(V))$.*

Dimostrazione. Abbiamo osservato che se $g, h \in \mathrm{GL}(V)$ hanno determinante di segni opposti, allora non sono deformabili l'uno nell'altro, vedi l'Osservazione 5.11.6.

Dimostriamo che se $g, h \in \mathrm{GL}^+(V)$, allora g è deformabile con continuità in h. Siccome $1_n \in \mathrm{GL}^+(V)$ e la deformabilità è una relazione di equivalenza, è sufficiente dimostrare che ogni $g \in \mathrm{GL}^+(V)$ è deformabile in Id_V. Questo equivale a dimostrare che ogni $A \in \mathrm{GL}^+(\mathbb{R})$ è deformabile in 1_n. La dimostrazione è per induzione su n. Il caso $n = 1$ è banalmente vero: $A = (a)$ con $a > 0$ e l'applicazione $\gamma\colon [1, 1/a] \to \mathrm{GL}_1^+(\mathbb{R})$ data da $\gamma(t) := ta$ deforma A in 1_1. Ora dimostriamo il passo induttivo. Sia $n \geqslant 2$ e sia $A \in \mathrm{GL}^+(\mathbb{R})$. Esiste una serie di operazioni elementari sulle colonne di A di tipo 2, cioè aggiunta a una colonna di un multiplo di un'altra colonna, che trasformano A in una matrice $B \in \mathrm{GL}_n(\mathbb{R})$ con una sola entrata non nulla sulla prima riga, diciamo sulla colonna j. Per l'Esempio 5.11.4 la matrice A è deformabile in B. Se $j < n$, sia $C \in \mathrm{GL}_n(\mathbb{R})$ la matrice ottenuta scambiando le colonne j e n di A e moltiplicando per -1 la "nuova" colonna n se l'entrata sulla prima riga è negativa, moltiplicando per -1 la "nuova" colonna j in caso contrario. Per l'Esempio 5.11.5 la matrice B è deformabile in C. Se moltiplichiamo la colonna n di C per $c_{n,n}^{-1}$, che è positivo, otteniamo una matrice $D \in \mathrm{GL}_n(\mathbb{R})$ con $d_{1,1} = 1$, e C si deforma in D. Ora aggiungendo opportuni multipli della prima riga alle rimanenti righe otteniamo una matrice $E \in \mathrm{GL}_n(\mathbb{R})$ con la proprietà che tutte le entrate su riga n e colonna n sono nulle eccetto per $e_{n,n}$ che è uguale a 1. La matrice D è deformabile in E per l'Esempio 5.11.4. Riassumendo, A è deformabile nella matrice

$$E = \begin{bmatrix} e_{1,1} & \dots & \dots & \dots & e_{1,n-1} & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & \dots & e_{i,j} & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ e_{n-1,1} & \dots & \dots & \dots & e_{n-1,n-1} & 0 \\ 0 & \dots & 0 & \dots & 0 & 1 \end{bmatrix}$$

La matrice $(n-1) \times (n-1)$ con entrate le $e_{i,j}$ per $i, \in \{1, \dots, n-1\}$ ha determinante uguale a $\mathrm{Det}\, E$, quindi positivo. Per l'ipotesi induttiva E è deformabile in 1_{n-1}, e segue che E è deformabile in 1_n. Questo dimostra che A si deforma in 1_n se l'unica entrata non nulla della prima riga di B (la matrice ottenuta da A al primo paso) non è sull'ultima colonna. Se invece è sull'ultima colonna, prima la scambiamo con la colonna $n-1$ (cambiando segno a un delle due colonne), ottenendo una matrice invertibile che si deforma in B, e ci ritroviamo nel caso già analizzato. Abbiamo dimostrare che due elementi di $\mathrm{GL}^+(V)$ sono deformabili l'uno nell'altro.

Ora supponiamo che $g, h \in (\mathrm{GL}(V)\setminus \mathrm{GL}^+(V))$. Allora $g^{-1} \cdot h \in \mathrm{GL}^+(V)$ (per Binet), e quindi per quanto appena dimostrato esiste un'applicazione continua $\gamma\colon [a, b] \to \mathrm{GL}(V))$ tale che $\gamma(0) = g^{-1} \cdot h$ e $\gamma(1) = \mathrm{Id}_V$. L'applicazione $\varphi\colon [a, b] \to \mathrm{GL}(V)$ definita da $\varphi(t) := g \cdot \gamma(t)$ è continua e si ha $\varphi(0) = h$, $\varphi(1) = g$. □

Esercizi del Capitolo 5

Esercizio 5.1. *Calcolate i determinanti delle seguenti matrici intere quadrate:*

$$A := \begin{bmatrix} 2 & 3 & 1 \\ 3 & 5 & 0 \\ 2 & 4 & 2 \end{bmatrix}, \qquad B := \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}.$$

Esercizio 5.2. *Calcolate le matrici dei cofattori delle A e B dell'Esercizio 5.1.*

Esercizio 5.3. *Sia $\mathbb{K}$ un campo e $x_1, x_2, \ldots, x_n \in \mathbb{K}$. Calcolate i determinanti delle seguenti matrici*

$$\begin{bmatrix} 1 & 1 \\ x_1 & x_2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1^2 & x_2^2 & x_3^2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 \\ x_1^3 & x_2^3 & x_3^3 & x_4^3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ x_1^2 & x_2^2 & \ldots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{bmatrix}.$$

Esercizio 5.4. *Sia $A \in M_{n,n}(\mathbb{Z})$. Sia p un numero primo e $\overline{A} \in M_{n,n}(\mathbb{Z}/(p))$ la matrice che si ottiene da A sostituendo all'entrata a_{ij} la classe di equivalenza di a_{ij} in $\mathbb{Z}/(p)$. Dimostrate che se* $\operatorname{rg} \overline{A} = r$ *allora* $\operatorname{Det}(A)$ *è divisibile per p^{n-r}.*

Esercizio 5.5. *I numeri* 2254*,* 4746*,* 5194 *e* 1792 *sono divisibili per* 7*. Ne* segue *che*

$$\begin{bmatrix} 2 & 2 & 5 & 4 \\ 4 & 7 & 4 & 6 \\ 5 & 1 & 9 & 4 \\ 1 & 7 & 9 & 2 \end{bmatrix} \in M_{4,4}(\mathbb{Z})$$

ha determinante divisibile per 7*. Perchè?*

Esercizio 5.6. *Per $n \geqslant 1$ sia $A(n) := (a_{ij}) \in M_{n,n}(\mathbb{Z})$ con*

$$a_{ij} = \begin{cases} 2 & \text{se } i = j, \\ -1 & \text{se } |i-j| = 1, \\ 0 & \text{altrimenti.} \end{cases} \tag{5.11.5}$$

Quindi

$$A(1) = (2), \quad A(2) = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, \quad A(3) = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}, \quad \ldots$$

Dimostrate che $\operatorname{Det} A(n) = n + 1$ *per ogni n.*

Esercizio 5.7. *Siano $A \in M_{5,1}(\mathbb{K})$ e $B \in M_{1,5}(\mathbb{K})$. Quindi $A \cdot B \in M_{5,5}(\mathbb{K})$. Qual'è il determinante di $A \cdot B$?*

Esercizio 5.8. *Sia $A \in M_{n,n}(\mathbb{K})$, e sia B la matrice ottenuta riscrivendo le colonne di A nell'ordine opposto. Quale relazione esiste tra* $\operatorname{Det} A$ *e* $\operatorname{Det} B$*?*

Esercizio 5.9. *Per quali n è vero che se $A \in \mathrm{GL}_n(\mathbb{K})$ allora $(A^c)^c = A$? (Ricordiamo che A^c è la matrice dei cofattori di A.)*

Esercizio 5.10. *Sia $\mathbb{S}$ uno spazio affine di dimensione finita n, e sia $RA(O, \mathcal{B})$ un riferimento cartesiano su $\mathbb{S}$.*

1. *Sia $d \geqslant 1$. Dimostrate che punti $P_0(a_{0,1}, \ldots, a_{0,n}), \ldots, P_d(a_{d,1}, \ldots, a_{d,n}) \in \mathbb{S}$ sono linearmente indipendenti se e solo se la matrice $d \times n$ data da*

$$\begin{bmatrix} (a_{1,1} - a_{0,1}) & \ldots & \ldots & (a_{1,n} - a_{0,n}) \\ \vdots & \vdots & \vdots & \vdots \\ (a_{d,1} - a_{0,1}) & \ldots & \ldots & (a_{d,n} - a_{0,n}) \end{bmatrix}$$

ha rango massimo cioè d.

2. *Supponiamo che* $P_0(a_{0,1},\dots,a_{0,n}),\dots,P_d(a_{d,1},\dots,a_{d,n})\in\mathbb{S}$ *siano linearmente indipendenti e sia* $\mathbb{L}\subset\mathbb{S}$ *il sottospazio lineare di dimensione* d *generato da* $P_0,\dots,P_d$. *Dimostrate che* $P(x_1,\dots,x_n)\in\mathbb{L}$ *se e solo se sono nulli tutti i determinanti dei minori* $(d+1)\times(d+1)$ *della matrice* $(d+1)\times n$ *data da*

$$\begin{bmatrix} (x_1-a_{0,1}) & \dots & \dots & (x_n-a_{0,n}) \\ (a_{1,1}-a_{0,1}) & \dots & \dots & (a_{1,n}-a_{0,n}) \\ \vdots & \vdots & \vdots & \vdots \\ (a_{d,1}-a_{0,1}) & \dots & \dots & (a_{d,n}-a_{0,n}) \end{bmatrix}$$

Esercizio 5.11. *Sia* $A\in M_{2,2}(\mathbb{Z})$ *data da*

$$A:=\begin{bmatrix} 22 & -12 \\ 35 & -19 \end{bmatrix}$$

Scrivete in forma chiusa (cioè non ricorsiva) A^m *per* $m\in\mathbb{Z}$.

Esercizio 5.12. *Sia*

$$A(x_1,\dots,x_n):=\begin{bmatrix} 1+x_1 & 1 & 1 & \dots & 1 \\ 1 & 1+x_2 & 1 & \dots & 1 \\ 1 & 1 & 1+x_3 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \dots & 1+x_n \end{bmatrix}.$$

Dimostrate che

$$\operatorname{Det} A(x_1,\dots,x_n)=x_1x_2\dots x_n+\sum_{i=1}^{n}x_1\cdot\ldots\cdot\widehat{x}_i\cdot\ldots\cdot x_n$$

dove $x_1\cdot\ldots\cdot\widehat{x}_i\cdot\ldots\cdot x_n$ *è il prodotto degli* x_s *con* $s\neq i$.

Esercizio 5.13. *Sia*

$$\Delta_n:=\operatorname{Det}\begin{bmatrix} a_0 & 1 & 0 & \dots & \dots & \dots & 0 \\ -1 & a_1 & 1 & 0 & \dots & \dots & 0 \\ 0 & -1 & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 1 & 0 \\ \vdots & \vdots & \ddots & 0 & -1 & a_{n-1} & 1 \\ 0 & \dots & \dots & \dots & 0 & -1 & a_n \end{bmatrix}$$

Dimostrate che

$$\frac{\Delta_n}{\Delta_{n-1}}=a_n+\cfrac{1}{a_{n-1}+\cfrac{1}{\dots+\cfrac{1}{a_1+\cfrac{1}{a_0}}}}.$$

Esercizio 5.14. *Sia* $A\in M_{3,3}(\mathbb{R})$ *data da*

$$A:=\begin{bmatrix} 3 & -2 & 1 \\ 0 & 1 & 1 \\ 4 & -1 & -1 \end{bmatrix}$$

1. *Calcolate autovalori e autospazi di* A.
2. *Determinate se* A *è diagonalizzabile.*

Esercizio 5.15. *Sia* $V\subset\mathbb{R}^4$ *il sottospazio*

$$V=\{X\in\mathbb{R}^4\mid x_1+x_2+x_3+x_4=0\}=\{X\in\mathbb{R}^4\mid(\boldsymbol{e}_1^\vee+\boldsymbol{e}_2^\vee+\boldsymbol{e}_3^\vee+\boldsymbol{e}_4^\vee)(X)=0\}.$$

Sia $M\in M_{4,4}(\mathbb{R})$ *definita così:*

$$M:=\begin{bmatrix} 2 & 1 & 0 & -1 \\ 1 & 0 & -1 & 2 \\ 0 & -1 & 2 & 1 \\ -1 & 2 & 1 & 0 \end{bmatrix}.$$

(a) *Verificate che*

$$L_M^\vee(\boldsymbol{e}_1^\vee + \boldsymbol{e}_2^\vee + \boldsymbol{e}_3^\vee + \boldsymbol{e}_4^\vee) = 2(\boldsymbol{e}_1^\vee + \boldsymbol{e}_2^\vee + \boldsymbol{e}_3^\vee + \boldsymbol{e}_4^\vee)$$

e ***quindi*** *(perchè ?)* $L_M(V) \subset V$. *Sia*

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ X & \mapsto & L_M(X) \end{array} \tag{5.11.6}$$

(b) *Calcolate* $\operatorname{Det} f$*, dove* f *è data da* (5.11.6)*, seguendo la definizione di* $\operatorname{Det} f$.

(c) *Notate che* $L_M(1,1,1,1) = (2,2,2,2)$. *Calcolate* $\operatorname{Det} M$ *e usate questo calcolo per (ri)determinare* $\operatorname{Det} f$. *(Suggerimento: pensate di calcolare* $\operatorname{Det} M$ *scegliendo una base il cui primo vettore è* $(1,1,1,1)$ *e gli altri formano una base di....).*

Esercizio 5.16. *Sia* $\mathbb{K}$ *un campo, e siano* $A^1, \dots, A^{n-1} \subset \mathbb{K}^n$ *vettori linearmente indipendenti, pensati come vettori-riga. Siccome i vettori sono linearmente indipendenti il sottospazio*

$$V := \langle A^1, \dots, A^{n-1} \rangle \subset \mathbb{K}^n$$

ha codimensione 1 *in* $\mathbb{K}^n$ *e quindi* $\dim \operatorname{Ann}(V) = 1$. *Sia* $A \in M_{n-1,n}(\mathbb{K})$ *la matrice le cui righe sono* $A^1, \dots, A^{n-1}$. *Dato* $1 \leqslant j \leqslant n$ *sia* $M_j \in M_{n-1,n-1}(\mathbb{K})$ *la matrice ottenuta eliminando la colonna* j*-esima da* A. *Infine sia*

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{f} & \mathbb{K} \\ X & \mapsto & \sum_{j=1}^n (-1)^j (\operatorname{Det} M_j) x_j \end{array}$$

Dimostrate che

$$\operatorname{Ann} V = \langle f \rangle.$$

Esercizio 5.17. *Sia* $\mathbb{K}$ *un campo e supponiamo che* $\operatorname{char} \mathbb{K} \neq 2$.

(1) *Sia* n *dispari e supponiamo che* $A \in M_{n,n}(\mathbb{K})$ *sia* antisimmetrica *cioè che* $A^t = -A$. . *Dimostrate che* $\operatorname{Det} A = 0$.

(2) *per ogni* n *pari date un esempio di* $A \in M_{n,n}(\mathbb{K})$ *antisimmetrica con* $\operatorname{Det} A \neq 0$.

Capitolo 6

Spazi vettoriali euclidei ed hermitiani

6.1 Motivazione

Se scegliamo una unità di misura nel piano euclideo $\mathbb{E}^2$ o nello spazio euclideo $\mathbb{E}^3$ possiamo definire la *norma* $||v||$ di un vettore $v \in \mathsf{V}(\mathbb{E}^m)$ come la lunghezza di un qualsiasi segmento orientato che rappresenta v, cioè

$$||\overrightarrow{PQ}|| = d(P,Q),$$

dove $d(P,Q)$ è la distanza tra P e Q. Possiamo fare di più, cioè definire il *prodotto scalare* di due vettori $v, w \in \mathsf{V}(\mathbb{E}^m)$ ponendo

$$v \cdot w := ||v|| \cdot ||w|| \cdot \cos\theta \tag{6.1.1}$$

dove θ è l'angolo tra segmenti orientati PQ, PR che rappresentano rispettivamente v e w. Notiamo che la norma di un vettore $v \in \mathsf{V}(\mathbb{E}^2)$ si ottiene dal prodotto scalare attraverso la formula $||v|| = (v \cdot v)^{1/2}$, ma il prodotto scalare contiene anche l'informazione sugli angoli tra rette attraverso la Formula (6.1.1). Le seguenti proprietà del prodotto scalare sono di fondamentale importanza:

1. L'applicazione $\mathsf{V}(\mathbb{E}^2) \times \mathsf{V}(\mathbb{E}^2) \to \mathbb{R}$ che manda (v, w) in $v \cdot w$ è bilineare.
2. Il prodotto scalare è simmetrico, cioè $v \cdot w = w \cdot v$ per ogni coppia di vettori $v, w \in \mathsf{V}(\mathbb{E}^2)$.
3. Se $v \in \mathsf{V}(\mathbb{E}^2)$, allora $v \cdot v \geqslant 0$, e si ha equaglianza solo se $v = 0$.

Uno spazio vettoriale euclideo è uno spazio vettoriale reale provvisto di un "prodotto scalare euclideo" che gode delle proprietà del prodotto scalare appena evidenziate - i dettagli sono nella prossima Sezione.

6.2 Spazi vettoriali euclidei

Definizione 6.2.1. Sia V uno spazio vettoriale reale. Un *prodotto scalare euclideo su* V è una forma bilineare

$$\begin{array}{ccc} V \times V & \longrightarrow & \mathbb{R} \\ (v, w) & \mapsto & \langle v, w \rangle \end{array} \tag{6.2.1}$$

simmetrica, cioè tale che $\langle v, w \rangle = \langle w, v \rangle$ per ogni coppia di vettori $v, w \in V$. Inoltre si richiede che $\langle v, v \rangle \geqslant 0$ per ogni $v \in V$, con equaglianza solo se $v = 0$.

Uno *spazio vettoriale euclideo* è una coppia $(V, \langle , \rangle)$ dove V è uno spazio vettoriale reale e $\langle , \rangle$ è prodotto scalare euclideo su V.

Il prodotto scalare su $\mathsf{V}(\mathbb{E}^m)$ della Sezione 6.1 è un prodotto scalare euclideo - lo è il *prodotto euclideo standard* su $\mathsf{V}(\mathbb{E}^m)$ (notate che è definito a meno di uno scalare perchè dipende dall'unità di misura scelta).

Esempio 6.2.2. La forma bilineare

$$\begin{array}{ccc} \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ (X, Y) & \mapsto & X^t \cdot Y = \sum_{i=1}^{n} x_i y_i \end{array} \tag{6.2.2}$$

è simmetrica, e inoltre $X^t \cdot X = \sum_{i=1}^n x_i^2$ è non negativo per ogni $X \in \mathbb{R}^n$ ed è uguale a 0 solo se $X = 0$. Quindi (6.2.2) definisce un prodotto scalare euclideo su $\mathbb{R}^n$ - è il *prodotto euclideo standard* su $\mathbb{R}^n$. Ovviamente il valore sulla coppia (X, Y) si denota $\langle X, Y\rangle$.

Esempio 6.2.3. Sia $C^0([-\pi, \pi])$ lo spazio vettoriale delle funzioni continue

$$f\colon [-\pi, \pi] \longrightarrow \mathbb{R},$$

dove l'addizione e la moltiplicazione per uno scalare sono definite punto per punto. Siano $f, g \in C^0([-\pi, \pi])$; poniamo

$$\langle f, g\rangle := \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)g(t)dt. \tag{6.2.3}$$

La $\langle , \rangle$ è evidentemente bilineare e simmetrica. Inoltre

$$\frac{1}{\pi} \int_{-\pi}^{\pi} f(t)^2 dt \geqslant 0,$$

e se l'integrale è nullo allora $f^2 = 0$ perchè la funzione f^2 è non negativa e continua, e quindi $f = 0$. Quindi (6.2.3) definisce un prodotto scalare euclideo su $C^0([-\pi, \pi])$.

Definizione 6.2.4. Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. La *norma* (associata) di un vettore $v \in V$ è data da

$$||v|| := \langle v, v\rangle^{1/2}.$$

(Quindi la norma di v è un reale non negativo.)

Esempio 6.2.5. Scegliamo un'unità di misura in $\mathbb{E}^m$ per $m \in \{2, 3\}$, e sia $\langle , \rangle$ il prodotto euclideo standard su $\mathsf{V}(\mathbb{E}^m)$. Se $v \in \mathsf{V}(\mathbb{E}^m)$ allora $||v||$ è la lunghezza di un qualsiasi segmento orientato che rappresenta il vettore v.

Esempio 6.2.6. Sia $(C^0([-\pi, \pi]), \langle , \rangle)$ lo spazio euclideo dell'Esempio 6.2.3. Se $m \in \mathbb{Z}^*$ allora

$$||\cos mt|| = ||\sin mt|| = 1, \qquad ||1|| = \sqrt{2}. \tag{6.2.4}$$

La norma in uno spazio vettoriale euclideo è definita a partire dal prodotto scalare euclideo. La proposizione che segue mostra che, viceversa, il prodotto scalare euclideo si ricostruisce a partire dalla norma.

Proposizione 6.2.7. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. Se $v, w \in V$ allora*

$$\langle v, w\rangle = \frac{1}{2}\left(||v + w||^2 - ||v||^2 - ||w||^2\right). \tag{6.2.5}$$

Dimostrazione. Segue dalle uguaglianze

$$||v + w||^2 = \langle v + w, v + w\rangle = ||v||^2 + 2\langle v, w\rangle + ||w||^2. \tag{6.2.6}$$

□

L'eguaglianza in (6.2.5) si chiama *identità di polarizzazione*.

6.3 Cauchy-Schwartz e la diseguaglianza triangolare

Teorema 6.3.1. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. Se $v, w \in V$ allora*

$$\langle v, w\rangle^2 \leqslant ||v||^2 \cdot ||w||^2, \tag{6.3.1}$$

e vale l'eguaglianza se e solo se v, w sono linearmente dipendenti.

Dimostrazione. Se $v = 0$ la tesi è banalmente vera. Ora assumiamo che $v \neq 0$. Per $x \in \mathbb{R}$ poniamo

$$p(x) := \langle xv + w, xv + w\rangle = ||v||^2x^2 + 2\langle v, w\rangle x + ||w||^2.$$

Siccome $p(x) \geqslant 0$ per ogni $x \in \mathbb{R}$, il polinomio di secondo grado p (è di secondo grado perchè $v \neq 0$) non ha radici reali oppure ha una radice reale che ha molteplicità due. Quindi il discriminante $p(x)$ è minore o uguale a 0, cioè

$$(2\langle v, w\rangle)^2 - 4||v||^2 \cdot ||w||^2 \leqslant 0.$$

Questo dimostra che vale (6.3.1). Inoltre il discriminante, cioè l'espressione a sinistra di (6.3.1), è uguale a 0 se e solo se esiste una radice dell'equazione $p(x) = 0$, cioè $x \in \mathbb{R}$ tale che $xv + w = 0$. Siccome $v \neq 0$, questo equivale alla condizione che v, w siano linearmente dipendenti. □

Corollario 6.3.2. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. Siano $v, w \in V$: si ha che*

$$||v + w|| \leqslant ||v|| + ||w||. \tag{6.3.2}$$

Inoltre vale l'eguaglianza se e solo se esiste $\lambda \geqslant 0$ tale che $v = \lambda w$ o $w = \lambda v$.

Dimostrazione. Per la diseguaglianza di Cauchy-Schwartz abbiamo che

$$||v + w||^2 = ||v||^2 + 2\langle v, w\rangle + ||w||^2 \leqslant ||v||^2 + 2||v|| \cdot ||w|| + ||w||^2 = (||v|| + ||w||)^2. \tag{6.3.3}$$

Segue che vale (6.3.2). Inoltre vediamo che in (6.3.2) vale l'eguaglianza se e solo se $\langle v, w\rangle = ||v|| \cdot ||w||$. Per il Teorema 6.3.1 se vale tale eguaglianza allora esiste $\lambda \in \mathbb{R}$ tale che $v = \lambda w$ o $w = \lambda v$, e siccome $\langle v, w\rangle = \lambda||w||^2$ nel primo caso e $\langle v, w\rangle = \lambda||v||^2$ nel secondo caso segue che $\lambda \geqslant 0$. Viceversa si verifica subito che se esiste $\lambda \geqslant 0$ tale che $v = \lambda w$ o $w = \lambda v$ allora $\langle v, w\rangle = ||v|| \cdot ||w||$. □

La diseguaglianza in (6.3.1) è la *Diseguaglianza di Cauchy-Schwarz*, quella in (6.3.2) è la *Diseguaglianaza triangolare*.

Osservazione 6.3.3. Scegliamo un'unità di misura in $\mathbb{E}^m$ per $m \in \{2, 3\}$, e sia $\langle , \rangle$ il prodotto euclideo standard su $\mathsf{V}(\mathbb{E}^m)$. La validità del Teorema 6.3.1 segue dalla Formula 6.1.1. Il Teorema 6.3.1 afferma che la diseguaglianza di Cauchy-Schwarz (e la descrizione dei casi nei quali la diseguaglianza è una uguaglianza) segue dalle proprietà che definiscono un prodotto scalare euclideo.

Se $P, Q \in \mathbb{E}^m$ denotiamo $d(P, Q)$ la distanza tra P e Q. Siano $P, Q, R \in \mathbb{E}^m$. La diseguaglianza triangolare per $v := \overrightarrow{PQ}$ e $w := \overrightarrow{QR}$ equivale (vedi l'Esempio 6.2.5) alla diseguaglianza

$$d(P, R) \leqslant d(P, Q) + d(Q, R).$$

Questo spiega il nome della diseguaglianza. (Mentre la diseguaglianza di Cauchy-Schwarz prende il nome dai matematici A. L. Cauchy[1] e H. Schwarz[2]).

Esempio 6.3.4. Per $\mathbb{R}^n$ con il prodotto euclideo standard la diseguaglianza di Cauchy-Schwarz e la diseguaglianza triangolare affermano che, se $X, Y \in \mathbb{R}^n$,

$$\left(\sum_{i=1}^n x_iy_i\right)^2 \leqslant \left(\sum_{i=1}^n x_i^2\right) \cdot \left(\sum_{j=1}^n y_j^2\right), \qquad \sum_{i=1}^n (x_i + y_i)^2 \leqslant \left(\left(\sum_{i=1}^n x_i^2\right)^{1/2} + \left(\sum_{i=1}^n y_i^2\right)^{1/2}\right)^2. \tag{6.3.4}$$

[1] Vedi https://mathshistory.st-andrews.ac.uk/Biographies/Cauchy/
[2] Vedi https://mathshistory.st-andrews.ac.uk/Biographies/Schwarz/

Ora mostriamo che si può definire l'angolo tra vettori di un qualsiasi spazio vettoriale euclideo. La diseguaglianza di Cauchy-Schwarz dà che se $v \neq 0 \neq w$ allora

$$-1 \leqslant \frac{\langle v, w\rangle}{||v|| \cdot ||w||} \leqslant 1. \tag{6.3.5}$$

Definizione 6.3.5. Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. Siano $v, w \in V$ **non nulli**. L'angolo tra v e w è l'unico $0 \leqslant \theta \leqslant \pi$ tale che

$$\cos\theta = \frac{\langle v, w\rangle}{||v|| \cdot ||w||}. \tag{6.3.6}$$

Notate che la definizione ha senso per la (6.3.5).

Notate che l'angolo tra v e w non cambia se riscaliamo v o w per un fattore positivo(quindi è definito l'angolo tra "semirette") e che non dipende dall'ordine dei vettori.

Se $\langle , \rangle$ è il prodotto euclideo standard su $\mathsf{V}(\mathbb{E}^m)$ la definizione di angolo appena data restituisce la nozione usuale di angolo - vedi (6.1.1).

Definizione 6.3.6. Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. Vettori $v, w \in V$ sono *perpendicolari* (o *ortogonali*) se $\langle v, w\rangle = 0$, cioè uno dei due è nullo oppure l'angolo tra di essi è $\pi/2$. In simboli $v \perp w$. Sottoinsiemi $S, T \subset V$ sono *ortogonali* se per ogni coppia $(v, w) \in S \times T$ si ha $v \perp w$ - in simboli $S \perp T$.

Se $S \subset V$ l'*ortogonale* di S è

$$S^\perp := \{w \in V \mid v \perp w = 0 \quad \forall v \in S\}.$$

Se $S = \{v_0\}$ (cioè è un insieme che consiste di un solo elemento) denotiamo $\{v_0\}^\perp$ con $v_0^\perp$.

Sia $\mathbb{S}$ uno spazio affine e supponiamo che $\langle , \rangle$ sia un prodotto scalare euclideo su $\mathsf{V}(\mathbb{S})$. Se $P, Q \in \mathbb{S}$ ha senso definire la distanza $d(P, Q)$ come la norma $||\overrightarrow{PQ}||$. (Vedi il Capitolo 7.) Siano $P, Q, R \in \mathbb{S}$. I vettori $\overrightarrow{PQ}, \overrightarrow{QR}$ sono ortogonali se e solo se vale il teorema di Pitagora per il triangolo di "cateti" PQ, QR, cioè se e solo se

$$||\overrightarrow{PQ}||^2 + ||\overrightarrow{QR}||^2 = ||\overrightarrow{PR}||^2.$$

Esempio 6.3.7. Sia $(C^0([-\pi, \pi]), \langle , \rangle)$ lo spazio vettoriale euclideo dell'Esempio 6.2.3. Per $m \in \mathbb{N}$ siano f_m, g_m le funzioni date da

$$f_m(t) := \cos mt, \quad g_m(t) := \sin mt. \tag{6.3.7}$$

Un calcolo[3] dà che se $(m, n) \neq (0, 0)$ allora

$$\int_{-\pi}^{\pi} f_m(t) f_n(t) dt = \pi\delta_{m,n}, \ \int_{-\pi}^{\pi} g_m(t) g_n(t) dt = \pi\delta_{m,n}, \ \int_{-\pi}^{\pi} f_m(t) g_n(t) dt = 0. \tag{6.3.8}$$

Quindi le funzioni definite da (6.3.7) per $m \in \mathbb{N}$ sono a due a due ortogonali.

6.4 Basi ortonormali

Definizione 6.4.1. Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. Una lista di vettori $\{v_1, \ldots, v_n\}$ di V è *ortonormale* (abbreviamo scrivendo che è ON) se per ogni $1 \leqslant i, j \leqslant n$ si ha che $\langle v_i, v_j\rangle = \delta_{ij}$. Una *base ortonormale* (base ON) di V è una base che è anche ON.

Esempio 6.4.2. Sia $\langle , \rangle$ il prodotto scalare standard su $\mathbb{R}^n$. La base standard di $\mathbb{R}^n$ è una base ON di $(V, \langle , \rangle)$.

Con il consueto abuso di linguaggio diciamo che vettori ..., v_i, ... sono ON se la lista ..., v_i, ... è ON.

[3] Le uguaglianze $\cos mt = (e^{imt} + e^{-imt})/2$ e $\sin mt = (e^{imt} - e^{-imt})/2i$ semplificano il calcolo degli integrali.

Esempio 6.4.3. Sia $(C^0([-\pi,\pi]),\langle,\rangle)$ lo spazio vettoriale euclideo dell'Esempio 6.2.3. Le funzioni f_m, g_m dell'Esempio 6.3.7 per $m \in \mathbb{N}^*$ formano una lista (infinita) di vettori ortonormali. Rimane una lista ON se aggiungiamo la funzione costante $1/\sqrt{2}$ (cioè $1/\sqrt{2}f_0 = 1/\sqrt{2}g_0$).

Proposizione 6.4.4. *Sia $(V,\langle,\rangle)$ uno spazio vettoriale euclideo e $\{v_1,\dots,v_n\}$ una lista di vettori ON. Allora $\{v_1,\dots,v_n\}$ sono linearmente indipendenti.*

Dimostrazione. Supponiamo che

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0. \tag{6.4.1}$$

Sia $1 \leqslant i \leqslant n$. Calcolando il prodotto scalare di v_i con ambo i membri di (6.4.1) troviamo che $\lambda_i = 0$. □

Esempio 6.4.5. Sia $(C^0([-\pi,\pi]),\langle,\rangle)$ lo spazio vettoriale euclideo dell'Esempio 6.2.3. Le funzioni

$$\frac{1}{\sqrt{2}}, \quad \cos mt, \quad \sin mt, \qquad m \in \mathbb{N}^* \tag{6.4.2}$$

formano una lista ON, vedi l'Esempio 6.4.3, e quindi sono linearmente indipendenti.

Il seguente semplice risultato mostra la convenienza di scegliere una base ON di uno spazio vettoriale euclideo.

Proposizione 6.4.6. *Sia $(V,\langle,\rangle)$ uno spazio vettoriale euclideo, e siano $v_1,\dots,v_n \in V$ vettori ortonormali. Se $x_1,y_1,\dots,x_n,y_n \in \mathbb{R}$ allora*

$$\langle \sum_{i=1}^n x_i v_i, \sum_{i=1}^n y_i v_i \rangle = \sum_{i=1}^n x_i y_i. \tag{6.4.3}$$

Dimostrazione. Segue dal calcolo

$$\langle \sum_{i=1}^n x_i v_i, \sum_{i=1}^n y_i v_i \rangle = \sum_{i,j=1}^n \langle x_i v_i, y_j v_j \rangle = \sum_{i,j=1}^n x_i y_j \langle v_i, v_j \rangle = \sum_{i,j=1}^n x_i y_j \delta_{i,j} = \sum_{i=1}^n x_i y_i.$$

□

Proposizione 6.4.7. *Sia $(V,\langle,\rangle)$ uno spazio vettoriale euclideo finitamente generato. Esiste una base ortonormale di V.*

Dimostrazione. Per induzione sulla dimensione di V. Se $\dim V = 0$ il risultato è banalmente vero. Chi è sospettoso del caso $\dim V = 0$ può iniziare l'induzione dal caso $\dim V = 1$: basta scegliere una base $\{v_1\}$ di V e sostituire a v_1 il vettore $v_1/||v_1||$ che ha norma 1, questa è una base ON di V. Ora dimostriamo il passo induttivo. Sia $v_1 \in V$ un vettore non nullo, e sia $w_1 := v_1/||v_1||$. Quindi $||w_1|| = 1$. L'applicazione

$$\begin{array}{ccc} V & \xrightarrow{\varphi} & \mathbb{R} \\ v & \mapsto & \langle v, w_1 \rangle \end{array} \tag{6.4.4}$$

è lineare perchè $\langle,\rangle$ è bilineare, e inoltre è suriettiva perchè $\langle w_1, w_1\rangle = 1$. Quindi il nucleo di φ, cioè $w_1^\perp$, ha dimensione uguale a $(\dim V - 1)$. La restrizione di $\langle,\rangle$ a $w_1^\perp$ è un prodotto scalare euclideo, e quindi esiste una base ON $\{w_2,\dots,w_n\}$ di $w_1^\perp$ per l'ipotesi induttiva (poniamo $n := \dim V$). I vettori $\{w_1, w_2, \dots, w_n\}$ sono ON e quindi sono linearmente indipendenti per la Proposizione 6.4.4, e siccome $\dim V = n$ costituiscono una base ON di V. □

6.5 Decomposizione ortogonale

Lemma 6.5.1. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo e sia $S \subset V$. L'ortogonale $S^\perp \subset V$ è un sottospazio di V.*

Dimostrazione. Se $v_0 \in V$ allora $v_0^\perp$ è il nucleo dell'applicazione lineare

$$\begin{array}{ccc} V & \longrightarrow & \mathbb{R} \\ v & \mapsto & \langle v, v_0 \rangle \end{array} \tag{6.5.1}$$

e quindi $v_0^\perp$ è un sottospazio lineare di V. Siccome

$$S^\perp = \bigcap_{v \in S} v^\perp,$$

segue che $S^\perp$ è intersezione di sottospazi lineari e perciò è un sottospazio lineare. □

Lemma 6.5.2. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. Sia $U \subset V$ un sottospazio finitamente generato e siano $u_1, \ldots, u_m$ generatori di U. Allora*

$$U^\perp = \bigcap_{i=1}^m u_i^\perp \tag{6.5.2}$$

Dimostrazione. È ovvio che il membro di sinistra di (6.5.2) è contenuto nel membro di destra. Resta da dimostrare che il membro di destra di (6.5.2) è contenuto nel membro di sinistra. Supponiamo che $v \in u_i^\perp$ per $1 \leqslant i \leqslant m$. Sia $u \in U$, e quindi $u = \sum_{i=1}^m \lambda_i u_i$. Allora

$$\langle v, u \rangle = \langle v, \sum_{i=1}^m \lambda_i u_i \rangle = \sum_{i=1}^m \lambda_i \langle v, u_i \rangle = 0.$$

□

Il prossimo risultato dà la decomposizione ortogonale del titolo della sezione.

Proposizione 6.5.3. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo e $U \subset V$ un sottospazio finitamente generato. Allora*

$$V = U \oplus U^\perp. \tag{6.5.3}$$

(Ricordiamo che l'equazione (6.5.3) *significa che l'applicazione naturale $U \oplus U^\perp \to V$ è un isomorfismo, vedete l'Esempio 3.2.11.)*

Dimostrazione. Come spiegato nell'Esempio 3.2.11 basta dimostrare che $U \cap U^\perp = \{0\}$ e che V è generato da U e $U^\perp$, cioè che

$$V = U + U^\perp. \tag{6.5.4}$$

Sia $v \in U \cap U^\perp$; allora $\langle v, v \rangle = 0$ e quindi $v = 0$. Rimane da dimostrare che vale (6.5.4). Per la Proposizione 6.4.7 esiste una base ON $\mathcal{B} := \{u_1, \ldots, u_m\}$ di U. Poniamo

$$\begin{array}{ccc} V & \xrightarrow{\pi_{\mathcal{B}}} & U \\ v & \mapsto & \sum_{i=1}^m \langle v, u_i \rangle u_i. \end{array} \tag{6.5.5}$$

L'applicazione $\pi_{\mathcal{B}}$ è lineare perchè per ciascun $i \in \{1, \ldots, m\}$ l'applicazione

$$\begin{array}{ccc} V & \longrightarrow & \mathbb{R} \\ v & \mapsto & \langle v, u_i \rangle \end{array} \tag{6.5.6}$$

è lineare. Se $v \in V$ allora

$$(v - \pi_{\mathcal{B}}(v)) \in U^\perp \tag{6.5.7}$$

perchè per ciascun $i \in \{1, \dots, m\}$ si ha $\langle v, u_i \rangle = \langle \pi_{\mathcal{B}}(v), u_i \rangle$, cioè $(v - \pi_{\mathcal{B}}(v)) \perp u_i$ e quindi (6.5.7) vale per il Lemma 6.5.2. Sia $v \in V$; siccome

$$v = \pi_{\mathcal{B}}(v) + (v - \pi_{\mathcal{B}}(v)) \tag{6.5.8}$$

e $(v - \pi_{\mathcal{B}}(v)) \in U^{\perp}$, vale (6.5.4). □

Osservazione 6.5.4. L'applicazione $\pi_{\mathcal{B}}$ definita da (6.5.5) *non* dipende dalla base ON $\mathcal{B}$ di U. Infatti se $v \in V$ il vettore $\pi_{\mathcal{B}}(v)$ è UN vettore di U tale che $(v - \pi_{\mathcal{B}}(v)) \in U^{\perp}$, e tale vettore è unico perchè vale la decomposizione in somma diretta (6.5.3).

Definizione 6.5.5. Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo e $U \subset V$ un sottospazio finitamente generato. La *proiezione ortogonale su* U è l'applicazione $\pi_{\mathcal{B}} \colon V \to U$ definita da (6.5.5), e la denotiamo π_U. (La definizione ha senso perchè $\pi_{\mathcal{B}}$ non dipende dalla base ON $\mathcal{B}$ di U.)

I risultati seguenti sono conseguenze immediate della Proposizione 6.5.3.

Corollario 6.5.6. *Siano $(V, \langle , \rangle)$ uno spazio vettoriale euclideo e $U \subset V$ un sottospazio finitamente generato. Allora*

$$\left(U^{\perp}\right)^{\perp} = U. \tag{6.5.9}$$

Corollario 6.5.7. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo finitamente generato. Sia $U \subset V$ un sottospazio. Allora*

$$\dim U^{\perp} = \dim V - \dim U. \tag{6.5.10}$$

Definizione 6.5.8. Uno spazio vettoriale euclideo $(V, \langle , \rangle)$ è *somma diretta ortogonale* dei sottospazi $W_1, \dots, W_r \subset V$ se è somma diretta di $W_1, \dots, W_r$, cioè

$$V = W_1 \oplus \dots \oplus W_r$$

e in aggiunta $W_i \perp W_j$ per ogni $i, j \in \{1, \dots, r\}$ con $i \neq j$. In tal caso adottiamo la notazione

$$V = W_1 \oplus_{\perp} \dots \oplus_{\perp} W_r$$

La seguente proposizione dà una caratterizzazione fondamentale della proiezione ortogonale su un sottospazio in termini di norma.

Proposizione 6.5.9. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. Sia $U \subset V$ un sottospazio finitamente generato e $\pi_U \colon V \to U$ la proiezione ortogonale su U. Allora*

$$||v - \pi_U(v)|| \leqslant ||v - u|| \quad \forall u \in U \tag{6.5.11}$$

e si ha eguaglianza solo se $u = \pi_U(v)$.

Dimostrazione. Abbiamo che

$$\begin{aligned} &||v - u||^2 = ||v - \pi_U(v) - (u - \pi_U(v))||^2 = \\ &= ||v - \pi_U(v)||^2 - 2\langle (v - \pi_U(v)), (u - \pi_U(v)) \rangle + ||(u - \pi_U(v))||^2 = ||v - \pi_U(v)||^2 + ||(u - \pi_U(v))||^2 \end{aligned}$$

(l'ultima uguaglianza vale perchè $(v - \pi_U(v)) \in U^{\perp}$ e $(u - \pi_U(v)) \in U$) e la proposizione segue. □

Esempio 6.5.10. Sia $(C^0([-\pi, \pi]), \langle , \rangle)$ lo spazio vettoriale euclideo dell'Esempio 6.2.3. Sia $U_n \subset C^0([-\pi, \pi])$ il sottospazio generato dalle funzioni di (6.4.2) dove $m \in \{1, \dots, n\}$. Sia $\varphi \in C^0([-\pi, \pi])$ data da $\varphi(t) = t$. La proiezione di φ sul sottospazio U_n è

$$\varphi_n(t) = 2 \sum_{m=1}^{n} \frac{(-1)^{m+1}}{m} \sin mt. \tag{6.5.12}$$

La decomposizione ortogonale

$$C^0([-\pi,\pi]) = U_n \oplus U_n^\perp$$

dà che

$$\frac{2\pi^2}{3} = \frac{1}{\pi}\int_{-\pi}^{\pi} t^2 dt = ||\varphi||^2 \geqslant \left\| 2\sum_{m=1}^{n} \frac{(-1)^{m+1}}{m} \sin mt \right\|^2 = 4\sum_{m=1}^{n} \frac{1}{m^2}. \tag{6.5.13}$$

(L'ultima eguaglianza vale perchè le funzioni $\sin mt$ per $m \in \{1,\dots,n\}$ sono ortonormali.) Per un risultato non banale (la densità delle funzioni trigonometriche nello spazio delle funzioni reali di variabile reale periodiche di periodo 2π) la differenza tra il membro di sinistra e quello di destra tende a 0 per n che tende a ∞, e segue la seguente eguaglianza scoperta da Euler:

$$\sum_{m=1}^{\infty} \frac{1}{m^2} = \frac{\pi^2}{6}.$$

Osservazione 6.5.11. Nella Proposizione 6.5.3 e nel Corollario 6.5.6 l'ipotesi che $U \subset V$ sia un sottospazio finitamente generato è necessaria perchè valga la tesi. Per dimostrarlo è sufficiente dare un esempio di un sottospazio proprio $U \subsetneq V$ tale che $U^\perp = \{0\}$. Sia ℓ^2 l'insieme delle successioni reali a quadrato sommabile:

$$\ell^2 := \left\{ \{a_n\}_{n\in\mathbb{N}} \mid a_i \in \mathbb{R} \quad \forall i \in \mathbb{N}, \quad \sum_{i=0}^{\infty} a_i^2 < \infty \right\}.$$

Siano $\{a_n\}, \{b_n\}$ successioni in ℓ^2 e $\lambda \in \mathbb{R}$; le successioni $\{a_n + b_n\}$ e $\{\lambda a_n\}$ sono in ℓ^2. Le operazioni

$$\{a_n\} + \{b_n\} := \{a_n + b_n\}, \qquad \lambda\{a_n\} := \{\lambda a_n\}$$

danno a ℓ^2 una struttura di spazio vettoriale reale. Si definisce un prodotto scalare euclideo su ℓ^2 ponendo

$$\langle \{a_n\}, \{b_n\} \rangle := \sum_{n=0}^{\infty} a_n \cdot b_n.$$

(La diseguaglianza di Cauchy-Schwartz per $\mathbb{R}^n$ con il prodotto euclideo standard - vedi (6.3.4) - dimostra che la serie è assolutamente convergente.) Ora sia

$$U := \{\{a_n\}_{n\in\mathbb{N}} \mid a_n \in \mathbb{R} \text{ per ogni } n \in \mathbb{N} \text{ e } a_n = 0 \text{ per } n \gg 0\}.$$

Chiaramente U è un sottospazio *proprio* di ℓ^2 e inoltre $U^\perp = \{0\}$.

6.6 Algoritmo di Gram-Schmidt

Dati vettori linearmente indipendenti $v_1,\dots,v_m$ di uno spazio vettoriale euclideo $(V, \langle,\rangle)$ l'*algoritmo di Gram-Schmidt* produce vettori ortogonali o, volendo, ortonormali $w_1,\dots,w_m \in V$ tali che per ogni $s \in \{1,\dots,m\}$ si abbia

$$\mathrm{Span}(w_1,\dots,w_s) = \mathrm{Span}(v_1,\dots,v_s).$$

In particolare $\{w_1,\dots,w_m\}$ è una base ortogonale del sottospazio generato da $\{v_1,\dots,v_m\}$. Per $s \in \{1,\dots,m\}$ sia

$$U_s := \mathrm{Span}(v_1,\dots,v_s),$$

e sia $\pi_s \colon V \to U_s$ la proiezione ortogonale su U_s. Poniamo $w_1 = v_1$ e per $s \in \{2,\dots,m\}$

$$w_s := v_s - \pi_{s-1}(v_s).$$

Proposizione 6.6.1. *Per $s \in \{1,\dots,m\}$ la lista $\{w_1,\dots,w_s\}$ è una base ortogonale (cioè $w_i \perp w_j$ se $i \neq j$) di U_s.*

Dimostrazione. Se $s = 1$ allora $w_1 = v_1$, quindi non c'è nulla da dimostrare. Ora dimostriamo il passo induttivo. Supponiamo che $\{w_1, \dots, w_j\}$ sia una base ortogonale di U_j, e dimostriamo che $\{w_1, \dots, w_{j+1}\}$ è una base ortogonale di U_{j+1}. Abbiamo

$$\mathrm{Span}(w_1, \dots, w_{j+1}) \subset U_{j+1} \tag{6.6.1}$$

perchè per ipotesi induttiva

$$\mathrm{Span}(w_1, \dots, w_j) = \mathrm{Span}(v_1, \dots, v_j) = U_j \subset U_{j+1},$$

e $w_{j+1} \in U_{j+1}$ perchè è una combinazione lineare di $v_{j+1} \in U_{j+1}$ e $\pi_j(v_{j+1}) \in U_j$. Inoltre w_{j+1} è ortogonale a U_j per definizione di proiezione ortogonale.

Dimostriamo che in (6.6.1) si ha eguaglianza. Per l'ipotesi induttiva

$$U_j \subset \mathrm{Span}(w_1, \dots, w_{j+1}),$$

quindi basta dimostrare che $w_{j+1} \notin U_j$. Se $w_{j+1} \in U_j$, allora $w_{j+1} = 0$ (è ortogonale a U_j, in altre parole abbiamo la decomposizione in somma diretta $V = U \oplus U^\perp$) cioè $v_{j+1} \in U_j$, e questo contraddice l'ipotesi che $v_1, \dots, v_{j+1}$ sono linearmente indipendenti. Questo dimostra che in (6.6.1) si ha eguaglianza. Siccome $\dim U_{j+1} = j + 1$ segue che $\{w_1, \dots, w_{j+1}\}$ è una base di U_{j+1}. Inoltre è ortogonale perchè per ipotesi induttiva $\{w_1, \dots, w_j\}$ è una base ortogonale di U_j e w_{j+1} è ortogonale a U_j per costruzione. □

La procedura appena descritta dà la seguente formula ricorsiva per $w_1, \dots, w_m$:

$$w_{j+1} = v_{j+1} - \sum_{i=1}^{j} \frac{\langle v_{j+1}, w_i \rangle}{||w_i||^2} w_i. \tag{6.6.2}$$

Infatti basta applicare la formula (6.5.5) per la proiezione ortogonale su U_j, dove la base ON di U_j è

$$\left\{ \frac{w_1}{||w_1||}, \dots, \frac{w_j}{||w_j||} \right\}.$$

Se, con l'algoritmo di Gram-Schmidt, vogliamo ottenere una base ON del sottospazio $\mathrm{Span}(v_1, \dots, v_m)$, prima troviamo una base ortogonale seguendo il procedimento descritto sopra, e poi normalizziamo i w_i, cioè sostituiamo a w_i il vettore

$$w_i' := \frac{w_i}{||w_i||}.$$

Esempio 6.6.2. Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo, e sia $\mathcal{B} = \{u_1, u_2, u_3\}$ una sua base ON. Siano

$$v_1 := u_1 - 2u_2 + 2u_3, \qquad v_2 := 3u_1 - 12u_2 - 9u_3.$$

Notate che v_1, v_2 sono lineamente indipendenti. Applichiamo il procedimento di Gram-Schmidt per produrre una base ON del sottospazio $\mathrm{Span}(v_1, v_2)$. Prima troviamo una base ortogonale $\{w_1, w_2\}$ e poi normalizziamo per passare a una base ON $\{w_1', w_2'\}$. Abbiamo

$$w_1 = v_1 = u_1 - 2u_2 + 2u_3.$$

Calcoliamo w_2 applicando la formula in (6.6.2):

$$w_2 = v_2 - \frac{(v_2, w_1)}{||w_1||^2} w_1 = v_2 - \frac{9}{9} w_1 = 3u_1 - 12u_2 - 9u_3 - (u_1 - 2u_2 + 2u_3) = 2u_1 - 10u_2 - 11u_3.$$

Normalizzando otteniamo

$$w_1' := \frac{1}{3}u_1 - \frac{2}{3}u_2 + \frac{2}{3}u_3, \qquad w_2' := \frac{2}{15}u_1 - \frac{2}{3}u_2 - \frac{11}{15}u_3.$$

6.7 Matrice di Gram

Definizione 6.7.1. Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. La *matrice di Gram* associata alla lista di vettori $\mathcal{B} := \{v_1, \dots, v_n\}$ di V è la matrice $n \times n$ reale $M_{\mathcal{B}}(\langle , \rangle)$ con entrata $\langle v_i, v_j \rangle$ su riga i e colonna j, cioè

$$M_{\mathcal{B}}(\langle , \rangle) := \begin{bmatrix} \langle v_1, v_1 \rangle & \dots & \langle v_1, v_n \rangle \\ \vdots & \dots & \vdots \\ \langle v_n, v_1 \rangle & \dots & \langle v_n, v_n \rangle \end{bmatrix}.$$

Esempio 6.7.2. Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo e $\mathcal{B} := \{v_1, \dots, v_n\}$ una base ON di V. Allora $M_{\mathcal{B}}(\langle , \rangle)$ è la matrice unità 1_n.

Osservazione 6.7.3. Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo e $\mathcal{B} := \{v_1, \dots, v_n\}$ una lista di vettori di V. Allora $M_{\mathcal{B}}(\langle , \rangle)$ è una matrice simmetrica perchè la sua entrata su riga i, colonna j è $\langle v_i, v_j \rangle$ che è uguale a $\langle v_j, v_i \rangle$ cioè la sua entrata su riga j, colonna i.

Prima di analizzare la matrice di Gram ci soffermiamo su una costruzione più generale.

Sia V uno spazio vettoriale finitamente generato su $\mathbb{K}$, e sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una sua base. Sia $A \in M_{n,n}(\mathbb{K})$. L'applicazione

$$\begin{array}{ccc} V \times V & \xrightarrow{\Phi_A^{\mathcal{B}}} & \mathbb{K} \\ (v, w) & \mapsto & X_{\mathcal{B}}(v)^t \cdot A \cdot X_{\mathcal{B}}(w) \end{array} \tag{6.7.1}$$

è una forma bilineare (facile verifica). Per referenza futura notiamo che

$$a_{ij} = \Phi_A^{\mathcal{B}}(v_i, v_j). \tag{6.7.2}$$

Definizione 6.7.4. Una matrice $A \in M_{n,n}(\mathbb{K})$ è *simmetrica* se $A^t = A$. Denotiamo con $M_{n,n}^+(\mathbb{K})$ il sottoinsieme di $M_{n,n}(\mathbb{K})$ i cui elementi sono le matrici simmetriche.

Osservazione 6.7.5. Si verifica facilmente che $M_{n,n}^+(\mathbb{K})$ è un sottospazio vettoriale di $M_{n,n}(\mathbb{K})$.

Definizione 6.7.6. Una forma bilineare $F \colon V \times V \to \mathbb{K}$ è *simmetrica* se $F(v, w) = F(w, v)$ per ogni $v, w \in V$. Denotiamo con $\mathrm{Bil}(V)$ l'insieme delle forme bilineari su $V \times V$ e con $\mathrm{Bil}^+(V)$ il sottoinsieme delle forme bilineari simmetriche.

Proposizione 6.7.7. *Sia V uno spazio vettoriale finitamente generato su $\mathbb{K}$ e $\mathcal{B}$ una sua base. L'applicazione*

$$\begin{array}{ccc} M_{n,n}(\mathbb{K}) & \xrightarrow{\Phi^{\mathcal{B}}} & \mathrm{Bil}(V) \\ A & \mapsto & \Phi_A^{\mathcal{B}} \end{array} \tag{6.7.3}$$

è biunivoca e l'immagine di $M_{n,n}^+(\mathbb{K})$ è $\mathrm{Bil}^+(V)$, cioè il sottoinsieme delle forme bilineari simmetriche.

Dimostrazione. Dimostriamo che $\Phi^{\mathcal{B}}$ è biunivoca. L'equazione (6.7.2) mostra che $\Phi^{\mathcal{B}}$ è iniettiva. Ora dimostriamo che $\Phi^{\mathcal{B}}$ è suriettiva. Sia $F \in \mathrm{Bil}(V)$. Per $1 \leqslant i, j \leqslant n$ sia $a_{ij} := F(v_i, v_j)$ e $A := (a_{ij})$. Si ha che

$$F\left(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j\right) = \sum_{1 \leqslant i,j \leqslant n} x_i y_j F(v_i, v_j) =$$

$$= \sum_{1 \leqslant i,j \leqslant n} a_{ij} x_i y_j = X^t \cdot A \cdot Y = \Phi_A^{\mathcal{B}}\left(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j\right).$$

Questo dimostra che $F = \Phi_A^{\mathcal{B}}$. Rimane da dimostrare che $\Phi_A^{\mathcal{B}}$ è simmetrica se e solo se A è simmetrica. Supponiamo che $\Phi_A^{\mathcal{B}}$ sia simmetrica. Poniamo $A = (a_{ij})$. Allora

$$a_{ij} = \Phi_A^{\mathcal{B}}(v_i, v_j) = \Phi_A^{\mathcal{B}}(v_j, v_i) = a_{ji} \tag{6.7.4}$$

e quindi $A^t = A$. D'altra parte se $A = A^t$ allora

$$\Phi_A^{\mathcal{B}}(v,w) = X_{\mathcal{B}}(v)^t \cdot A \cdot X_{\mathcal{B}}(w) = (X_{\mathcal{B}}(v)^t \cdot A \cdot X_{\mathcal{B}}(w)^t = X_{\mathcal{B}}(w)^t \cdot A^t \cdot X_{\mathcal{B}}(v) = X_{\mathcal{B}}(w)^t \cdot A \cdot X_{\mathcal{B}}(v) = \Phi_A^{\mathcal{B}}(w,v).$$

□

Definizione 6.7.8. Se V è uno spazio vettoriale finitamente generato su $\mathbb{K}$ e $\mathcal{B}$ una sua base,

$$\mathrm{Bil}(V) \xrightarrow{M_{\mathcal{B}}} M_{n,n}(k) \tag{6.7.5}$$

è l'inversa di $\Phi^{\mathcal{B}}$.

Osservazione 6.7.9. Se $F \in \mathrm{Bil}(V)$ allora $M_{\mathcal{B}}(F)$ è la matrice $A = (a_{ij})$ con

$$a_{ij} = F(v_i, v_j). \tag{6.7.6}$$

In particolare, la notazione appena introdotta è coerente con quella usata per la matrice di Gram di un prodotto scalare $\langle , \rangle$ associata a una base, cioè la $M_{\mathcal{B}}(\langle , \rangle)$ appena definita è uguale alla matrice di Gram della Definizione 6.7.1.

Osservazione 6.7.10. È facile verificare che $\mathrm{Bil}(V)$ e $\mathrm{Bil}^+(V)$ sono sottospazi dello spazio vettoriale delle funzioni da V^2 a $\mathbb{K}$. È ugualmente facile verificare che l'applicazione $\Phi^{\mathcal{B}}$ in (6.7.3) è lineare e (di conseguenza) anche la sua inversa $M_{\mathcal{B}}$. Quindi $\Phi^{\mathcal{B}}$ è un isomorfismo tra $M_{n,n}(\mathbb{K})$ e $\mathrm{Bil}(V)$, e perciò per la Proposizione 6.7.7 la sua restrizione al sottospazio $M_{n,n}(\mathbb{K}) \subset M_{n,n}^+(\mathbb{K})$ delle matrici simmetriche è un isomorfismo $M_{n,n}^+(\mathbb{K}) \xrightarrow{\sim} \mathrm{Bil}^+(V)$.

Proposizione 6.7.11. *Sia V uno spazio vettoriale finitamente generato su $\mathbb{K}$ e $F \in \mathrm{Bil}(V)$. Se $\mathcal{B}$ e $\mathcal{C}$ sono basi di V allora*

$$M_{\mathcal{C}}(F) = M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id})^t \cdot M_{\mathcal{B}}(F) \cdot M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}). \tag{6.7.7}$$

Dimostrazione. Per ogni $v, w \in V$ abbiamo

$$\begin{aligned} X_{\mathcal{C}}(v)^t \cdot M_{\mathcal{C}}(F) \cdot X_{\mathcal{C}}(w) &= F(v,w) = X_{\mathcal{B}}(v)^t \cdot M_{\mathcal{B}}(F) \cdot X_{\mathcal{B}}(w) = \\ &= (M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}) \cdot X_{\mathcal{C}}(v))^t \cdot M_{\mathcal{B}}(F) \cdot M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}) \cdot X_{\mathcal{C}}(w) = X_{\mathcal{C}}(v)^t \cdot (M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}))^t \cdot M_{\mathcal{B}}(F) \cdot M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}) \cdot X_{\mathcal{C}}(w) \end{aligned}$$

e segue (6.7.7). □

Osservazione 6.7.12. Riferendoci alla Proposizione 6.7.11, supponiamo che $M_{\mathcal{B}}(F)$ sia simmetrica cioè, per la Proposizione 6.7.7, che F sia una forma bilineare simmetrica. Allora

$$(M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id})^t \cdot M_{\mathcal{B}}(F) \cdot M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}))^t = M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id})^t \cdot M_{\mathcal{B}}(F)^t \cdot (M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id})^t)^t = M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id})^t \cdot M_{\mathcal{B}}(F) \cdot M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}),$$

in accordo con la Proposizione 6.7.7.

Corollario 6.7.13. *Sia V uno spazio vettoriale finitamente generato su $\mathbb{K}$ e $F \in \mathrm{Bil}(V)$. Se $\mathcal{B}$ e $\mathcal{C}$ sono basi di V allora*

$$\mathrm{Det}\, M_{\mathcal{C}}(F) = \mathrm{Det}\, M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id})^2 \cdot \mathrm{Det}\, M_{\mathcal{B}}(F). \tag{6.7.8}$$

Dimostrazione. Per la Proposizione 6.7.11 e la formula di Binet abbiamo

$$\mathrm{Det}\, M_{\mathcal{C}}(F) = \mathrm{Det}\, M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id})^t \cdot \mathrm{Det}\, M_{\mathcal{B}}(F) \cdot \mathrm{Det}\, M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}). \tag{6.7.9}$$

Siccome $\mathrm{Det}\, M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id})^t = \mathrm{Det}\, M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id})$ (per la Proposizione 5.5.8) segue (6.7.8). □

Corollario 6.7.14. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. Se $\mathcal{C}$ è una base di V allora la matrice di Gram $M_{\mathcal{C}}(\langle , \rangle)$ ha determinante (strettamente) positivo.*

Dimostrazione. Sia $\mathcal{B}$ una base ON di V (esiste per la Proposizione 6.4.7). Allora $M_{\mathcal{B}}(\langle , \rangle) = 1_n$ dove $n := \dim V$, e quindi $\operatorname{Det} M_{\mathcal{B}}(\langle , \rangle) = 1 > 0$. Per il Corollario 6.7.13 segue che $\operatorname{Det} M_{\mathcal{C}}(\langle , \rangle) > 0$. □

Osservazione 6.7.15. Il Teorema 6.3.1 segue dal Corollario 6.7.14 (e notate che nella dimostrazione del Corollario 6.7.14 non interviene il Teorema 6.3.1). Infatti siano $(V, \langle , \rangle)$ uno spazio vettoriale euclideo e $v, w \in V$. Se v, w sono linearmente indipendenti poniamo $U := \operatorname{Span}(v, w)$. La restrizione di $\langle , \rangle$ a U è un prodotto scalare euclideo, $\mathcal{B} := \{v, w\}$ è una base di U e la matrice di Gram $M_{\mathcal{B}}(\langle , \rangle_{|U})$ è data da

$$M_{\mathcal{B}}(\langle , \rangle_{|U}) = \begin{bmatrix} \langle v, v \rangle & \langle v, w \rangle \\ \langle w, v \rangle & \langle w, w \rangle \end{bmatrix}$$

Per il Corollario 6.7.14 $\operatorname{Det} M_{\mathcal{B}}(\langle , \rangle_{|U}) > 0$. Siccome $\operatorname{Det} M_{\mathcal{B}}(\langle , \rangle_{|U}) = (\langle v, v \rangle \cdot \langle w, w \rangle - \langle v, w \rangle^2)$ questo dà la diseguaglianza (stretta) di Cauchy-Schwarz per v, w sono linearmente indipendenti. Se v, w sono linearmente dipendenti si verifica subito che $\operatorname{Det} M_{\mathcal{B}}(\langle , \rangle_{|U}) = 0$.

Sia V uno spazio vettoriale reale finitamente generato, e sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una sua base. Per la Proposizione 6.7.7 ogni forma bilineare simmetrica su V è data da $\Phi^{\mathcal{B}}(A)$ per un'unica matrice simmetrica $A \in M^+_{n,n}(\mathbb{R})$. Il prossimo risultato dà un criterio che permette di determinare agevolmente se $\Phi^{\mathcal{B}}(A)$ è un prodotto scalare euclideo. Prima di dare il risultato introduciamo la seguente notazione: data $A \in M_{n,n}(\mathbb{K})$ denotiamo con $A(p)$ (per $p \in \{1, \dots, n\}$) la matrice $p \times p$ con entrata su riga i e colonna j uguale alla corrispondente entrata i, j di A. Per esempio

$$A(1) = [a_{11}], \quad A(2) := \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad A(3) := \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Proposizione 6.7.16. *Sia V uno spazio vettoriale reale finitamente generato, e sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una sua base. Sia $A \in M^+_{n,n}(\mathbb{R})$. La forma bilineare simmetrica $\Phi^{\mathcal{B}}(A)$ è un prodotto scalare euclideo se e solo se* $\operatorname{Det} A(p) > 0$ *per* $p \in \{1, \dots, n\}$.

Dimostrazione. Supponiamo che $\Phi^{\mathcal{B}}(A)$ sia un prodotto scalare euclideo. Per $p \in \{1, \dots, n\}$ sia $W_p := \langle v_1, \dots, p \rangle$ e sia $\mathcal{B}_p$ la base di W_p data da $\mathcal{B}_p := \{v_1, \dots, v_n\}$. La restrizione $\langle , \rangle_{|W_p}$ di $\langle , \rangle$ a W_p è un prodotto scalare euclideo, e $A(p)$ è la matrice di Gram $M_{\mathcal{B}_p}(\langle , \rangle_{|W_p})$. Quindi $\operatorname{Det} A(p) > 0$ per $p \in \{1, \dots, n\}$ per il Corollario 6.7.14.

Ora dimostriamo per induzione su $\dim V$ che se $\operatorname{Det} A(p) > 0$ per $p \in \{1, \dots, n\}$ allora $\Phi^{\mathcal{B}}(A)$ è un prodotto scalare euclideo. Per semplificare la notazione poniamo $F = \Phi^{\mathcal{B}}(A)$. La F è bilineare simmetrica, quindi basta dimostrare che

$$F(v, v) > 0 \quad \forall v \in (V \backslash \{0\}). \tag{6.7.10}$$

Se $\dim V = 1$ l'ipotesi $\operatorname{Det} A(1) > 0$ dà che $F(v_1, v_1) > 0$; se $v \in (V \backslash \{0\})$ allora $v = \lambda v_1$ per un $\lambda \in \mathbb{R}^*$ e quindi

$$F(v, v) = F(\lambda v_1, \lambda v_1) = \lambda^2 F(v_1, v_1) > 0.$$

Dimostriamo il passo induttivo. Supponiamo che $n = \dim V > 1$. Sia $U := \operatorname{Span}(v_1, \dots, v_{n-1})$ e sia $\mathcal{C}$ la base di U data da $\mathcal{C} := \{v_1, \dots, v_{n-1}\}$. Siccome la restrizione di F a U è la forma bilineare simmetrica $\Phi^{\mathcal{B}}(A(n-1))$, segue dall'ipotesi induttiva che la restrizione di F a U è un prodotto scalare euclideo. Per la Proposizione 6.5.3 abbiamo la decomposizione ortogonale

$$V = U \oplus U^{\perp}.$$

Per il Corollario 6.5.7 l'ortogonale $U^{\perp}$ ha dimensione 1; sia $\{w\}$ una sua base. Quindi $\mathcal{D} := \{v_1, \dots, v_{n-1}, w\}$ è una base di V. Siccome w è ortogonale a U, abbiamo

$$\operatorname{Det} M_{\mathcal{D}}(F) = \operatorname{Det} A(n-1) \cdot F(w, w). \tag{6.7.11}$$

D'altra parte $\operatorname{Det} M_{\mathcal{D}}(F)$ ha lo stesso segno di $\operatorname{Det} M_{\mathcal{B}}(F) = A(n)$ per il Corollario 6.7.13, e quindi segue che $F(w,w) > 0$. Sia $v \in V$ non nullo; per la la decomposizione ortogonale in (6.7.11) si ha $v = u + \lambda w$ con $u \in U$, $\lambda \in \mathbb{R}$ e almeno uno tra u, λ è non nullo. Quindi

$$F(v,v) = F(u+\lambda w, u+\lambda w) = F(u,u) + \lambda^2 F(w,w).$$

Se $u \neq 0$ allora $F(u,u) > 0$, e se $\lambda \neq 0$ allora $\lambda^2 F(w,w) > 0$; segue che $F(v,v) > 0$. □

Esercizio svolto 6.7.17. Sia $A \in M^+_{3,3}(\mathbb{R})$ data da

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

(a) Dimostrate che $\langle X, Y\rangle := X^t \cdot A \cdot Y$ definisce un prodotto scalare euclideo su $\mathbb{R}^3$.

(b) Diamo a $\mathbb{R}^3$ il prodotto scalare euclideo definito al punto (a). Determinare la proiezione ortogonale di $v := (1,1,1)$ su

$$U := \{(x_1,x_2,x_3) \mid x_1 + x_2 + x_3 = 0\}.$$

(a): L'applicazione $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ definita da $(X,Y) \mapsto \langle X, Y\rangle$ è bilineare e simmetrica (vedi la Proposizione 6.7.7), rimane da dimostrare che

$$\langle X, X\rangle > 0 \qquad \forall X \in (\mathbb{R}^3 \backslash \{0\}). \tag{6.7.12}$$

Applichiamo la Proposizione 6.7.16: calcolando si trova

$$\operatorname{Det} A(1) = 3, \quad \operatorname{Det} A(2) = \begin{vmatrix} 3 & 1 \\ 1 & 1 \end{vmatrix} = 2, \quad \operatorname{Det} A(3) = \begin{vmatrix} 3 & 1 & 0 \\ 1 & 1 & 2 \\ 0 & 2 & 7 \end{vmatrix} = 2.$$

Siccome i determinanti sono (strettamente) positivi vale (6.7.12).

(b): Troviamo una base ortogonale $\{w_1, w_2\}$ di U e poi applichiamo la formula in (6.6.2). Una base di U è $\{(1,-1,0),(1,0,-1)\}$. Poniamo $v_1 := (1,-1,0)$ e $v_2 := (1,0,-1)$ e applichiamo l'algoritmo di Gram-Schmidt. Si ha $w_1 = v_1$ e, siccome

$$\langle v_1, v_1\rangle = \langle e_1 - e_2, e_1 - e_2\rangle = \langle e_1, e_1\rangle - \langle e_1, e_2\rangle - \langle e_2, e_1\rangle + \langle e_2, e_2\rangle = 3 - 1 - 1 + 1 = 2$$

e

$$\langle v_1, v_2\rangle = \langle e_1 - e_2, e_1 - e_3\rangle = \langle e_1, e_1\rangle - \langle e_1, e_3\rangle - \langle e_2, e_1\rangle + \langle e_2, e_3\rangle = 3 - 0 - 1 + 2 = 4,$$

abbiamo

$$w_2 = v_2 - \frac{\langle v_2, v_1\rangle}{||v_1||^2} v_1 = (-1,2,-1).$$

Quindi la proiezione ortogonale di $v := (1,1,1)$ su U è data da

$$\pi_U(v) = \langle v, w_1\rangle \frac{w_1}{||w_1||^2} + \langle v, w_2\rangle \frac{w_2}{||w_2||^2} =$$

Abbiamo

$$\langle v, w_1\rangle = \langle v, v_1\rangle = 2$$

e

$$\langle v, w_2\rangle = [-1,2,-1] \cdot \begin{bmatrix} 3 & 1 & 0 \\ 1 & 1 & 2 \\ 0 & 2 & 7 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = -5.$$

La conclusione è che

$$\pi_U(v) = w_1 - 5w_2 =$$

6.8 Isometrie di spazi vettoriali euclidei

Definizione 6.8.1. Siano $(V, \langle , \rangle_V)$ e $(W, \langle , \rangle_W)$ spazi vettoriali euclidei. Un'applicazione lineare $f\colon V \to W$ è un'*isometria* se per ogni $v_1, v_2 \in V$ si ha

$$\langle v_1, v_2 \rangle_V = \langle f(v_1), f(v_2) \rangle_W. \tag{6.8.1}$$

Una isometria è un *isomorfismo* (di spazi vettoriali euclidei) se ha un'inversa che è un'isometria, cioè un'isometria $g\colon W \to V$ tale che $g \circ f = \mathrm{Id}_V$ e $f \circ g = \mathrm{Id}_W$. Lo spazio vettoriale euclideo $(V, \langle , \rangle_V)$ è *isomorfo* allo spazio vettoriale euclideo $(W, \langle , \rangle_W)$ se esiste un ismorfismo (di spazi vettoriali euclidei) $f\colon V \to W$.

Quando lo riteniamo necessario denoteremo un'isometria $f\colon V \to W$ di spazi vettoriali euclidei con $f\colon (V, \langle , \rangle_V) \longrightarrow (W, \langle , \rangle_W)$.

Proposizione 6.8.2. *Siano $(V, \langle , \rangle_V)$ e $(W, \langle , \rangle_W)$ spazi vettoriali euclidei, e sia $f\colon V \to W$ un'applicazione lineare. Allora f è un'isometria se e solo se per ogni $v \in V$ si ha*

$$||v||_V = ||f(v)||_W. \tag{6.8.2}$$

Dimostrazione. Se f è un'isometria allora vale (6.8.2) perchè

$$||v||_V = \langle v, v \rangle_V = \langle f(v), f(v) \rangle_W = ||f(v)||_W. \tag{6.8.3}$$

Ora supponiamo che valga (6.8.2). Se $v_1, v_2 \in V$ allora, per la Proposizione 6.2.7 cioè l'identità di polarizzazione, si ha

$$\begin{aligned}\langle v_1, v_2 \rangle_V = \frac{1}{2}\left(||v_1||_V^2 + ||v_2||_V^2 - ||v_1 + v_2||_V^2\right) = \frac{1}{2}\left(||f(v_1)||_W^2 + ||f(v_2)||_W^2 - ||f(v_1 + v_2)||_W^2\right) = \\ = \frac{1}{2}\left(||f(v_1)||_W^2 + ||f(v_2)||_W^2 - ||f(v_1) + f(v_2)||_W^2\right) = \langle f(v_1), f(v_2) \rangle_W.\end{aligned}$$

□

Corollario 6.8.3. *Se $f\colon V \to W$ è un'isometria di spazi vettoriali euclidei, allora f è iniettiva.*

Dimostrazione. Se $v \in V$ e $f(v) = 0$, allora l'equazione (6.8.2) dà che $||v|| = 0$ e quindi $v = 0$. □

Esempio 6.8.4. Sia $(V, \langle , \rangle_V)$ uno spazio vettoriale euclideo finitamente generato di dimensione n, e sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una sua base ON. L'applicazione

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{f} & V \\ X & \mapsto & \sum_{i=1}^{n} x_i v_i \end{array} \tag{6.8.4}$$

è un'isometria (il prodotto scalare euclideo su $\mathbb{R}^n$ è quello standard). Infatti se $X \in \mathbb{R}^n$ allora

$$||X||^2 = \sum_{i=1}^{n} x_i^2 = ||\sum_{i=1}^{n} x_i v_i||^2,$$

e quindi f è un'isometria per la Proposizione 6.8.2.

Proposizione 6.8.5. *Siano $(U, \langle , \rangle_U)$, $(V, \langle , \rangle_V)$ e $(W, \langle , \rangle_W)$ spazi vettoriali euclidei. Se $g\colon U \to V$ e $f\colon V \to W$ sono isometrie allora la composizione $f \circ g$ è un'isometria. Inoltre se $f\colon V \to W$ è un'isometria suriettiva, e quindi un isomorfismo di spazi vettoriali per il Corollario 6.8.3, allora l'inversa f^{-1} è un'isometria e perciò f è un isomorfismo di spazi vettoriali euclidei.*

Dimostrazione. Se $u_1, u_2 \in U$ allora

$$\langle u_1, u_2\rangle_U = \langle g(u_1), g(u_2)\rangle_V = \langle f(g(u_1)), f(g(u_2))\rangle_W = \langle f\circ g(u_1), f\circ g(u_2)\rangle_W. \tag{6.8.5}$$

Se $w_1, w_2 \in W$ allora

$$\langle f^{-1}(v_1), f^{-1}(v_2)\rangle_V = \langle f(f^{-1}(v_1)), f(f^{-1}(v_2))\rangle_W = \langle v_1, v_2\rangle_W. \tag{6.8.6}$$

□

Corollario 6.8.6. *Sia $(V, \langle,\rangle_V)$ uno spazio vettoriale euclideo. L'insieme degli isomorfismi $f\colon (V,\langle,\rangle) \to (V,\langle,\rangle)$ è un sottogruppo di* $\mathrm{GL}(V)$.

Dimostrazione. Segue dalla Proposizione 6.8.5 e dall'osservazione che l'identità Id_V è un'isomorfismo di spazi vettoriali euclidei. □

Osservazione 6.8.7. Sia $(V,\langle,\rangle)$ uno spazio vettoriale euclideo finitamente generato (cioè V è finitamente generato). Per il Corollario 6.8.3 ogni isometria $f\colon V \to V$ è suriettiva , e quindi è un isomorfismo di spazi vettoriali euclidei.

Esempio 6.8.8. Per l'Osservazione 6.8.7 l'isometria dell'Esempio 6.8.4 è un isomorfismo. Segue che spazi vettoriali euclidei finitamente generati $(V,\langle,\rangle_V)$ e $(W,\langle,\rangle_W)$ della stessa dimensione sono isomorfi (intendiamo come spazi vettoriali euclidei, già sappiamo che sono isomorfi come "semplici" spazi vettoriali). Infatti sia $n := \dim V = \dim W$. Esistono basi ON di V e di W per la Proposizione 6.4.7, e quindi esistono isomorfismi $f\colon \mathbb{R}^n \to V$ e $g\colon \mathbb{R}^n \to W$ per l'Esempio 6.8.4, e la composizione $g\circ f^1 : V \to W$ è un isomorfismo per la Proposizione 6.8.5.

Definizione 6.8.9. Sia $(V,\langle,\rangle)$ uno spazio vettoriale euclideo. Il *gruppo ortogonale di* $(V,\langle,\rangle)$ è il gruppo degli isomorfismi $f\colon (V,\langle,\rangle) \to (V,\langle,\rangle)$, e si denota $\mathrm{O}(V,\langle,\rangle)$ (o $\mathrm{O}(V)$ quando è chiaro quale sia il prodotto scalare euclideo).

Esempio 6.8.10. Il gruppo ortogonale di $\mathbb{R}^n$ con il prodotto euclideo standard si denota $\mathrm{O}_n(\mathbb{R})$. Sia $A \in \mathrm{GL}_n(\mathbb{R})$; allora $A \in \mathrm{O}_n(\mathbb{R})$ se solo se

$$A^t \cdot A = 1_n. \tag{6.8.7}$$

Infatti siano $X, Y \in \mathbb{R}^n$. Allora

$$\langle L_A(X), L_A(Y)\rangle = (A\cdot X)^t \cdot (A\cdot Y) = X^t \cdot (A^t\cdot A)\cdot Y,$$

e quindi è chiaro che se vale (6.8.7) allora $\langle L_A(X), L_A(Y)\rangle = \langle X, Y\rangle$. Vale il viceversa perchè

$$\langle L_A(e_i), L_A(e_j)\rangle = e_i^t \cdot (A^t\cdot A)\cdot e_j$$

che è uguale alla entrata di $A^t\cdot A$ su riga i e colonna j, e quindi $\langle L_A(e_i), L_A(e_j)\rangle = \langle e_i, e_j\rangle = \delta_{ij}$ dà $A^t\cdot A = (\delta_{ij}) = 1_n$.

Osservazione 6.8.11. Dall'equazione (6.8.7) segue per Binet che $\mathrm{Det}(A)^2 = 1$, e quindi $\mathrm{Det}(A) = \pm 1$. Si pone

$$\mathrm{SO}_n(\mathbb{R}) := \{A \in \mathrm{O}_n(\mathbb{R}) \mid \mathrm{Det}\, A = 1\}. \tag{6.8.8}$$

Per Binet $\mathrm{SO}_n(\mathbb{R})$ è un sottogruppo di $\mathrm{O}_n(\mathbb{R})$ - è il *gruppo ortogonale speciale.*

Più in generale sia $(V,\langle,\rangle)$ uno spazio vettoriale euclideo finitamente generato. Scegliendo una base ON vediamo che $\mathrm{Det}(g)^2 = 1$ per ogni $g \in \mathrm{O}(V,\langle,\rangle)$. Si pone

$$\mathrm{SO}(V,\langle,\rangle) := \{g \in \mathrm{O}(V,\langle,\rangle) \mid \mathrm{Det}(g) = 1\}, \tag{6.8.9}$$

e $\mathrm{SO}(V,\langle,\rangle)$ è sottogruppo di $\mathrm{O}(V,\langle,\rangle)$.

6.9 Spazi vettoriali hermitiani

Se V è uno spazio vettoriale *complesso* l'analogo di un prodotto scalare euclideo è dato dalla nozione di prodotto scalare hermitiano. Da notare che un tale analogo non è una forma bilineare ma una forma sesquilineare, vedi l'Osservazione 6.9.4.

Definizione 6.9.1. Sia V uno spazio vettoriale complesso. Un *prodotto scalare hermitiano* (da C. Hermite[4]) su V è un'applicazione

$$\begin{array}{ccc} V \times V & \longrightarrow & \mathbb{C} \\ (v, w) & \mapsto & \langle v, w\rangle \end{array} \tag{6.9.1}$$

tale che

1. dato $w \in V$ la funzione $V \to \mathbb{C}$ definita da $v \mapsto \langle v, w\rangle$ è lineare,
2. $\langle w, v\rangle = \overline{\langle v, w\rangle}$ per ogni $(v, w) \in V \times V$, e
3. $\langle v, v\rangle > 0$ se $0 \neq v \in V$. (Notate che $\langle v, v\rangle$ è reale perchè $\overline{\langle v, v\rangle} = \langle v, v\rangle$ per il punto (2).)

Uno *spazio vettoriale hermitiano* è una coppia $(V, \langle , \rangle)$ dove V è uno spazio vettoriale complesso e $\langle , \rangle$ è un prodotto scalare hermitiano su V.

Esempio 6.9.2. L'applicazione

$$\begin{array}{ccc} \mathbb{C}^n \times \mathbb{C}^n & \longrightarrow & \mathbb{C} \\ (X, Y) & \mapsto & X^t \cdot \overline{Y} = \sum_{j=1}^{n} x_j \overline{y}_j \end{array} \tag{6.9.2}$$

è un prodotto scalare hermitiano. È il *prodotto hermitiano* standard su $\mathbb{C}^n$.

Esempio 6.9.3. Sia $C^0([-\pi, \pi])_{\mathbb{C}}$ l'insieme delle funzioni $f\colon [-\pi, \pi] \to \mathbb{C}$ continue, cioè tali che $f(x) = u(x) + iv(x)$ dove $u, v\colon [-\pi, \pi] \to \mathbb{R}$ sono continue. Si verifica facilmente che $C^0([-\pi, \pi])_{\mathbb{C}}$ è un sottospazio vettoriale (complesso) di $\mathbb{C}^{[-\pi,\pi]}$. Definiamo

$$\langle f, g\rangle := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\overline{g(t)}dt, \qquad f, g \in C^0([-\pi, \pi])_{\mathbb{C}}. \tag{6.9.3}$$

L'integrale è definito calcolando la parte reale e immaginaria della funzione $f(t)\overline{g(t)}$. Esplicitamente scriviamo $f(t)\overline{g(t)} = u(t) + iv(t)$ dove $u, v\colon [-\pi, \pi] \to \mathbb{R}$ sono continue fuori da un sottoinsieme finito di $[-\pi, \pi]$: allora

$$\int_{-\pi}^{\pi} f(t)\overline{g(t)}dt := \int_{-\pi}^{\pi} u(t)dt + i \int_{-\pi}^{\pi} v(t)dt. \tag{6.9.4}$$

Si verifica facilmente che $\langle , \rangle$ è un prodotto scalare hermitiano su $C^0([-\pi, \pi])_{\mathbb{C}}$.

Osservazione 6.9.4. Sia $(V, \langle , \rangle)$ uno spazio vettoriale hermitiano. Dato $v \in V$ la funzione $V \to \mathbb{C}$ definita da $w \mapsto \langle v, w\rangle$ è *coniugato lineare*, cioè

$$\langle v, \lambda_1 w_1 + \lambda_2 w_2\rangle = \overline{\lambda}_1 \langle v, w_1\rangle + \overline{\lambda}_2 \langle v, w_2\rangle. \tag{6.9.5}$$

Infatti per i punti (1) e (2) della Definizione 6.9.1 abbiamo

$$\begin{aligned} \langle v, \lambda_1 w_1 + \lambda_2 w_2\rangle = \overline{\langle \lambda_1 w_1 + \lambda_2 w_2, v\rangle} = \overline{\lambda_1 \langle w_1, v\rangle + \lambda_2 \langle w_2, v\rangle} = \\ = \overline{\lambda}_1 \overline{\langle w_1, v\rangle} + \overline{\lambda}_2 \overline{\langle w_2, v\rangle} = \overline{\lambda}_1 \langle v, w_1\rangle + \overline{\lambda}_2 \langle v, w_2\rangle \end{aligned}$$

Un'applicazione $V \times V \to \mathbb{C}$ lineare nella prima variabile (cioè tale che valga il punto (1) della Definizione 6.9.1) e coniugato lineare nella seconda variabile (cioè tale che valga (6.9.5)) si chiama *forma sesquilineare* (vedi la Sezione 9.3). La condizione data dal punto (2) della Definizione 6.9.1 è l'equivalente, nel contesto delle forme sequilineari, della condizione di simmetria per forme bilineari reali.

[4] Vedi https://mathshistory.st-andrews.ac.uk/Biographies/Hermite/

Definizione 6.9.5. Sia $(V, \langle , \rangle)$ uno spazio vettoriale hermitiano. La *norma* di $v \in V$ è data da $||v|| := \langle v, v\rangle^{1/2}$.

Definizione 6.9.6. Sia $(V, \langle , \rangle)$ uno spazio vettoriale hermitiano. Vettori $v, w \in V$ sono *ortogonali* (o *perpendicolari*) se $\langle v, w\rangle = 0$ - in simboli $v \perp w$.

Notate che $v \perp w$ se e solo se $w \perp v$ perchè $\langle v, w\rangle = \overline{\langle w, v\rangle}$. Sottoinsiemi $S, T \subset V$ sono *ortogonali* se per ogni coppia $(v, w) \in S \times T$ si ha $v \perp w$ - in simboli $S \perp T$. Per l'ortogonalità in uno spazio vettoriale hermitiano si usa la stessa notazione e terminologia usata per gli spazi vettoriali euclidei. Una lista di vettori ortogonali o ortonormale si definisce come nel caso di spazio vettoriale euclideo.

La teoria degli spazi vettoriali hermitiani è del tutto simile a quella degli spazi vettoriali euclidei. Passiamo rapidamente in rassegna i punti salienti. I risultati delle Sezioni 6.3, 6.4, 6.5 e 6.6 e le loro dimostrazioni valgono senza modifiche anche per $(V, \langle , \rangle)$ uno spazio vettoriale hermitiano, con l'eccezione dell'analogo del Teorema 6.3.1, che si formula e si dimostra come segue.

Teorema 6.9.7. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale hermitiano. Se $v, w \in V$ allora*

$$|\langle v, w\rangle|^2 \leqslant ||v||^2 \cdot ||w||^2, \tag{6.9.6}$$

e vale l'eguaglianza se e solo se v, w sono linearmente dipendenti.

Dimostrazione. La dimostrazione è simile a quella Teorema 6.3.1, ma con un'accortezza. Se $v = 0$ la tesi è banalmente vera, e anche se $\langle v, w\rangle = 0$. Quindi possiamo assumere che $v \neq 0$ e $\langle v, w\rangle \neq 0$. Sia $\theta := \mathrm{Arg}(\langle v, w\rangle)$. Allora

$$||(\cos\theta - i\sin\theta)v|| = ||v||, \qquad \langle (\cos\theta - i\sin\theta)v, w\rangle \in \mathbb{R}.$$

Quindi la tesi vale per v, w se e solo se vale per $(\cos\theta - i\sin\theta)v, w$, e perciò è sufficiente dimostrare la tesi sotto l'ipotesi aggiuntiva che $\langle v, w\rangle \in \mathbb{R}$. Il resto della dimostrazione procede come nel caso di un prodotto scalare euclideo. Per $x \in \mathbb{R}$ poniamo

$$p(x) := \langle xv + w, xv + w\rangle = ||v||^2 x^2 + 2\langle v, w\rangle x + ||w||^2.$$

(Notate: l'eguaglianza vale perchè $\langle v, w\rangle \in \mathbb{R}$, in generale il coefficiente di x è $2\Re(\langle v, w\rangle)$.) Siccome $p(x) \geqslant 0$ per ogni $x \in \mathbb{R}$, il polinomio di secondo grado *reale* p (è di secondo grado perchè $v \neq 0$) non ha radici reali oppure ha una radice reale che ha molteplicità due. Quindi il discriminante $p(x)$ è minore o uguale a 0, cioè

$$2|\langle v, w\rangle|^2 - 4||v||^2 \cdot ||w||^2 \leqslant 0.$$

Questo dimostra che vale (6.9.6). Inoltre il discriminante, cioè l'espressione a sinistra di (6.9.6), è uguale a 0 se e solo se esiste una radice dell'equazione $p(x) = 0$, cioè $x \in \mathbb{R}$ tale che $xv + w = 0$. Siccome $v \neq 0$, questo equivale alla condizione che v, w siano linearmente dipendenti. □

Per quanto riguarda l'analogo della Sezione 6.7 vanno fatti alcuni cambiamenti, legati al fatto che un prodotto scalare hermitiano non è bilineare simmetrico ma sesquilinare e coniugato-simmetrico. La matrice di Gram associata a una lista finita di vettori si definisce come in 6.7.1. Per esempio se $\mathcal{B}$ è la base standard di $\mathbb{C}^n$ e $\langle , \rangle$ è il prodotto hermitiano standard su $\mathbb{C}^n$ allora $M_{\mathcal{B}}(\langle , \rangle) = 1_n$ (come nel caso euclideo). L'analogo dell'Osservazione 6.7.3 è come segue.

Osservazione 6.9.8. Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo e $\mathcal{B} := \{v_1, \ldots, v_n\}$ una lista di vettori di V. Allora $M_{\mathcal{B}}(\langle , \rangle)^t = \overline{M_{\mathcal{B}}(\langle , \rangle)}$ perchè l'entrata di $M_{\mathcal{B}}(\langle , \rangle)^t$ su riga i, colonna j è $\langle v_j, v_i\rangle$ che è uguale a $\overline{\langle v_i, v_j\rangle}$ cioè l'entrata di $\overline{M_{\mathcal{B}}(\langle , \rangle)}$ su riga i, colonna j.

L'Osservazione appena fatta motiva la seguente definizione.

Definizione 6.9.9. Una matrice $A \in M_{n,n}(\mathbb{C})$ è *hermitiana* se $A^t = \overline{A}$.

Sia V uno spazio vettoriale complesso finitamente generato e sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una sua base. Sia $A \in M_{n,n}(\mathbb{C})$. L'analogo dell'applicazione (6.7.1) che serve nel nostro contesto è la seguente

$$\begin{array}{ccc} V \times V & \xrightarrow{\Psi_A^{\mathcal{B}}} & \mathbb{C} \\ (v, w) & \mapsto & X_{\mathcal{B}}(v)^t \cdot A \cdot \overline{X}_{\mathcal{B}}(w). \end{array} \tag{6.9.7}$$

Il risultato sotto è l'analogo della Proposizione 6.7.7 nel nostro contesto. La dimostrazione è analoga a quella della Proposizione 6.7.7 e viene lasciata al lettore.

Proposizione 6.9.10. *Sia V uno spazio vettoriale complesso finitamente generato e sia $\mathcal{B}$ una sua base. Se $A \in M_{n,n}(\mathbb{C})$ allora $\Psi_A^{\mathcal{B}}$ è lineare nella prima variabile e coniugato-lineare nella seconda. Inoltre $\Psi_A^{\mathcal{B}}$ è* coniugato-simmetrica*, cioè per ogni $v, w \in V$*

$$\Psi_A^{\mathcal{B}}(w, v) = \overline{\Psi_A^{\mathcal{B}}(v, w)}$$

se e solo se A è una matrice hermitiana. Viceversa, se $f\colon V \times V \longrightarrow \mathbb{C}$ è un'applicazione lineare nella prima variabile e coniugato-lineare nella seconda esiste $A \in M_{n,n}(\mathbb{C})$ tale che $f = \Psi_A^{\mathcal{B}}$, e tale A è unica.

La Definizione 6.7.8 e l'Osservazione 6.7.9 rimangono sostanzialmente invariati, l'unica modifica è che l'insieme $\mathrm{Bil}(V)$ delle forme bilineari $V \times V \to \mathbb{K}$ viene sostituito dall'insieme delle forme *sesquilineari* $f\colon V \times V \longrightarrow \mathbb{C}$, cioè lineari nella prima variabile e coniugato-lineari nella seconda.

La rimanente modifica da fare è sostituire le Proposizioni 6.7.11 e il Corollario 6.7.13 con il seguente risultato, la cui dimostrazione è lasciata al lettore.

Proposizione 6.9.11. *Sia V uno spazio vettoriale complesso finitamente generato e $f\colon V \times V \longrightarrow \mathbb{C}$ una forma sesquilineari. Se $\mathcal{B}$ e $\mathcal{C}$ sono basi di V allora*

$$M_{\mathcal{C}}(f) = M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id})^t \cdot M_{\mathcal{B}}(f) \cdot \overline{M}_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}), \qquad \mathrm{Det}\, M_{\mathcal{C}}(F) = |\,\mathrm{Det}\, M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id})|^2 \cdot \mathrm{Det}\, M_{\mathcal{B}}(F). \tag{6.9.8}$$

Con questa modifica si vede facilmente che la dimostrazione della Proposizione 6.7.16 si estende a una dimostrazione del seguente criterio.

Proposizione 6.9.12. *Sia V uno spazio vettoriale complesso finitamente generato, e sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una sua base. Sia $A \in M_{n,n}(\mathbb{C})$ una matrice. La forma sesquilneare e coniugato-simmetrica $\Psi^{\mathcal{B}}(A)$ è un prodotto scalare hermitiano (cioè $\Psi^{\mathcal{B}}(A)(v, v) > 0$ per ogni $0 \neq v \in V$) se e solo se $\mathrm{Det}\, A(p) > 0$ per $p \in \{1, \dots, n\}$.*

Infine la definizione di isometria e isomorfismo tra spazi vettoriali hermitiani si definisce come nel caso euclideo, e valgono gli analoghi dei risultati della Sezione 6.8 con dimostrazioni invariate. La Definizione 6.8.9 è sostituita dalla seguente.

Definizione 6.9.13. Sia $(V, \langle , \rangle)$ uno spazio vettoriale hermitiano. Il *gruppo unitario di* $(V, \langle , \rangle)$ è il gruppo degli isomorfismi $f\colon (V, \langle , \rangle) \to (V, \langle , \rangle)$, e si denota $\mathrm{U}(V, \langle , \rangle)$ (o $\mathrm{U}(V)$ quando è chiaro quale sia il prodotto scalare hermitiano).

Il seguente esempio è analogo all'Esempio 6.8.10 e si verifica con un calcolo analogo.

Esempio 6.9.14. Il gruppo unitario di $\mathbb{C}^n$ con il prodotto hermitiano standard si denota $\mathrm{U}_n(\mathbb{C})$. Sia $A \in \mathrm{GL}_n(\mathbb{C})$; allora $A \in \mathrm{U}_n(\mathbb{C})$ se solo se

$$A^t \cdot \overline{A} = 1_n. \tag{6.9.9}$$

Dall'equazione (6.9.9) segue che $|\,\mathrm{Det}(A)| = 1$. Si pone

$$\mathrm{SU}_n(\mathbb{C}) := \{A \in \mathrm{U}_n(\mathbb{C}) \mid \mathrm{Det}\, A = 1\}. \tag{6.9.10}$$

Per Binet $\mathrm{SU}_n(\mathbb{C})$ è un sottogruppo di $\mathrm{U}_n(\mathbb{C})$ - è il *gruppo speciale unitario*.

Chiudiamo la sezione con un risultato che non ha un analogo nella Sezione 6.8 (per ora, ma vedi la Sezione 6.10))

Proposizione 6.9.15. *Sia V uno spazio vettoriale complesso finitamente generato e $\langle , \rangle$ un prodotto scalare hermitiano su V. Sia $f \in \mathrm{U}(V)$. Gli autovalori di f hanno modulo 1 ed esiste una base ON di V che diagonalizza f.*

Dimostrazione. Dimostriamo che gli autovalori di f hanno modulo 1. Sia λ un autovalore di f e v un autovettore associato. Allora

$$\langle v, v\rangle = \langle f(v), f(v)\rangle = \langle \lambda v, \lambda v\rangle = |\lambda|^2 \langle v, v\rangle. \tag{6.9.11}$$

Siccome $\langle v, v\rangle \neq 0$ (è strettamente positivo) segue che $|\lambda| = 1$. Ora dimostriamo per induzione sulla dimensione di V che esiste una base ON che diagonalizza f. Sia $n = \dim V$. Se $n = 1$ non c'è nulla da dimostrare. Dimostriamo il passo induttivo. Siccome il campo è quello dei complessi esiste un autovalore λ_n di V con autovettore v_n. Sia

$$W := v_n^{\perp} := \{w \in V \mid \langle w, v_n\rangle = 0\}. \tag{6.9.12}$$

Allora $f(W) \subset W$: infatti se $w \in W$ allora

$$0 = \langle w, v_n\rangle = \langle f(w), f(v_n)\rangle = \langle f(w), \lambda_n v_n\rangle = \overline{\lambda_n}\langle f(w), v_n\rangle = 0. \tag{6.9.13}$$

Quindi la restrizione di f a W definisce un endomorfismo $g\colon W \to W$ che è un operatore unitario per la forma hermitiana definita positiva su W data dalla restrizione di $\langle , \rangle$. Per ipotesi induttiva esiste una base ON $\{v_1, \ldots, v_{n-1}\}$ di W che diagonalizza g. Allora $\{v_1, \ldots, v_{n-1}, v_n\}$ è una base ON di V che diagonalizza f. □

Osservazione 6.9.16. Supponiamo che V e f siano come nella Proposizione 6.9.15. Sia $\{v_1, \ldots, v_n\}$ una base ON che diagonalizza f. Gli autovalori $\lambda_1, \ldots, \lambda_n$ di f hanno modulo 1 e quindi esistono $\theta_1, \ldots, \theta_n \in \mathbb{R}$ tali che $\lambda_j = e^{i\theta_j}$ per $j = 1, \ldots, n$. Quindi

$$f(v_j) = e^{i\theta_j} v_j, \qquad 1 \leqslant j \leqslant n. \tag{6.9.14}$$

6.10 Il gruppo ortogonale

Descriveremo le isometrie di uno spazio vettoriale euclideo $(V, \langle , \rangle)$ finitamente generato, cioè gli elementi di $\mathrm{O}(V, \langle , \rangle)$.

Se $\theta \in \mathbb{R}$ poniamo

$$R_\theta := \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{6.10.1}$$

Notiamo che R_θ è in $\mathrm{O}_2(\mathbb{R})$. Quindi se $(V, \langle , \rangle)$ è uno spazio vettoriale euclideo di dimensione 2 e $\mathcal{B}$ è una sua base ON allora l'applicazione lineare $f\colon V \to V$ tale che $M_{\mathcal{B}}^{\mathcal{B}}(f) = R_\theta$ è un'isometria. Notiamo anche che che $\mathcal{C}$ è un'altra base ON di V, allora $M_{\mathcal{C}}(f) = R_{\pm\theta}$, e il segno è quello del determinante della matrice (ortogonale) del cambiamento di base $M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}_V)$.

Definizione 6.10.1. Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo di dimensione 2. Un'isometria di V è una *rotazione* se, data una base ON $\mathcal{B}$ di V la matrice $M_{\mathcal{B}}^{\mathcal{B}}(f)$ è R_θ per un certo $\theta \in \mathbb{R}$ (che a meno del segno e traslazioni per multipli interi di π non dipende dalla base, vedi sopra).

In generale un'isometria $(V, \langle , \rangle) \to (V, \langle , \rangle)$ non è diagonalizzabile (su $\mathbb{R}$), per esempio se $\theta \in (\mathbb{R} \backslash \mathbb{Z}\pi)$ allora R_θ non è diagonalizzabile su $\mathbb{R}$. Nonostante ciò esiste un risultato analogo alla Proposizione 6.9.15. Spieghiamo il senso della dimostrazione che daremo con il seguente esempio.

Esempio 6.10.2. La matrice ortogonale R_θ data da (6.10.1) non è diagonalizzabile su $\mathbb{R}$ (se $\theta \notin \mathbb{Z}\pi$) però, siccome $\mathrm{O}_2(\mathbb{R}) \subset U_2(\mathbb{C})$, esiste una base ON di $\mathbb{C}^2$ (con prodotto hermitiano standard) che diagonalizza R_θ, e i corrispondenti autovalori hanno modulo 1. Infatti

$$\frac{1}{\sqrt{2}}(1,-i),\ \frac{1}{\sqrt{2}}(1,i).$$

è una base ON di autovettori di R_θ, e

$$R_\theta \cdot \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ \mp i\end{pmatrix} = (\cos\theta \pm i\sin\theta)\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ \pm i\end{pmatrix}$$

Notate che l'angolo di rotazione θ è determinato a meno del segno dagli autovalori (complessi) di R_θ (il segno di θ è legato alla scelta di un "verso" nella base ON di $\mathbb{R}^2$). Notate anche che i due autovettori della base che abbiamo dato sono l'uno il coniugato dell'altro, cioè se $v := \frac{1}{\sqrt{2}}(1,-i)$, allora la base ON di autovettori è $\{v, \overline{v}\}$, dove $\overline{v}$ è il vettore che ha come entrate i coniugati delle entrate di v. Infine la base standard, e quindi ON, di $\mathbb{R}^2$ è ottenuta a partire dalla base ON di autovettori tramite le formule

$$(1,0) = \frac{1}{\sqrt{2}}(v+\overline{v}), \qquad (0,1) = \frac{1}{\sqrt{2}}(iv - i\overline{v}). \tag{6.10.2}$$

Proposizione 6.10.3. *Siano $(V, \langle,\rangle)$ uno spazio vettoriale euclideo finitamente generato e $f \in \mathrm{O}(V)$. Esiste una decomposizione in somma diretta ortogonale (vedi la Definizione 6.5.8)*

$$V = V_1(f) \oplus_\perp V_{-1}(f) \oplus_\perp A_1 \ldots, \oplus_\perp A_c, \tag{6.10.3}$$

tale che per ogni $j \in \{1,\ldots,c\}$ il sottospazio A_j ha dimensione 2, si ha $f(A_j) = A_j$ e la restrizione di f a A_j è una rotazione.

Dimostrazione. Iniziamo dimostrando la tesi nel caso in cui V sia $\mathbb{R}^n$ con il prodotto scalare standard. È sufficiente dare una base ON $\{X_1\ldots,X_a,Y_1,\ldots,Y_b,U_1,W_1,U_2,W_2,\ldots,U_c,W_c\}$ di $\mathbb{R}^n$ e $\theta_1,\ldots,\theta_c \in \mathbb{R}$ tali che

$$f(X_p) = X_p, \quad 1 \leqslant p \leqslant a, \qquad f(Y_q) = -Y_q, \quad 1 \leqslant q \leqslant b \tag{6.10.4}$$

e

$$f(U_s) = \cos\theta_s U_s + \sin\theta_s W_s, \quad f(W_s) = -\sin\theta_s U_s + \cos\theta_s W_s, \quad 1 \leqslant s \leqslant c. \tag{6.10.5}$$

Abbiamo $f = L_A$ dove $A \in \mathrm{O}_n(\mathbb{R})$. Sia

$$\begin{array}{ccc}\mathbb{C}^n & \xrightarrow{F} & \mathbb{C}^n\\ Z & \mapsto & A\cdot Z\end{array}$$

cioè l'applicazione L_A dove A è pensata come elemento di $M_{n,n}(\mathbb{C})$. Ovviamente $F(Z) = f(Z)$ per $Z \in \mathbb{R}^n$. Siccome $A^t \cdot A$ (perchè $A \in \mathrm{O}_n(\mathbb{R})$) e $\overline{A} = A$ segue che $F \in \mathrm{U}_n(\mathbb{C})$. Per la Proposizione 6.9.15 F è diagonalizzabile e quindi

$$\sum_\lambda \dim_{\mathbb{C}} V_\lambda(F) = n. \tag{6.10.6}$$

Siano $\lambda,\mu \in \mathbb{C}$ autovalori di F e $U, W \in \mathbb{C}^n$ autovettori con autovalori λ e μ rispettivamente. Per la Proposizione 6.9.15 sappiamo che $1 = |\mu|$ cioè $\overline{\mu} = \mu^{-1}$, e quindi

$$\langle U, W\rangle = \langle F(U), F(W)\rangle = \langle \lambda U, \mu W\rangle = \lambda\overline{\mu}\langle U, W\rangle = \lambda\mu^{-1}\langle U, W\rangle. \tag{6.10.7}$$

Ne segue che se $\lambda \neq \mu$ allora $\langle U, W\rangle = 0$. In altre parole vettori appartenenti ad autospazi diversi sono ortogonali. Ora procediamo a costruire una base ON tale che valgano (6.10.4) e (6.10.5). Per ciascuno degli autovalori reali (che appartengono a $\{1,-1\}$) esiste una base ON *reale* dell'autospazio corrispondente: siano $\{X_1\ldots,X_a\}$ una base ON reale di $V_1(F)$ e $\{Y_1,\ldots,Y_b\}$ una base ON reale di $V_{-1}(F)$. Ora sia λ un autovalore *non* reale di F. Se $Z = (z_1,\ldots,z_n) \in \mathbb{C}^n$ poniamo $\overline{Z} := (\overline{z}_1,\ldots,\overline{z}_n)$.

Siccome A è reale un vettore (colonna) $Z \in \mathbb{C}^n$ è soluzione dell'equazione $(\lambda 1_n - A)\cdot Z = 0$ se e solo se $\overline{Z}$ è soluzione dell'equazione $(\overline{\lambda}1_n - A)\cdot \overline{Z} = 0$. Quindi l'applicazione $Z \mapsto \overline{Z}$ (che non è $\mathbb{C}$-lineare ma $\mathbb{R}$-lineare) definisce un isomorfismo di spazi vettoriali reali $V_\lambda(f) \xrightarrow{\sim} V_{\overline{\lambda}}(f)$. Perciò le dimensioni reali di $V_\lambda(f)$ e $V_{\overline{\lambda}}(f)$ sono uguali e, siccome le dimensioni complesse sono la metà delle dimensioni reali, abbiamo

$$\dim_{\mathbb{C}} V_{\overline{\lambda}}(f) = \dim_{\mathbb{C}} V_\lambda(f). \tag{6.10.8}$$

Per ogni coppia di autovalori complessi coniugati $\{\lambda, \overline{\lambda}\}$ scegliamo uno dei due autovalori, sia λ, e una base $\{T_1, \dots, T_d\}$ di $V_\lambda(F)$ che sia ON per il prodotto scalare hermitiano standard su $\mathbb{C}^n$. Allora $\{\overline{T}_1, \dots, \overline{T}_d\}$ è una base di $V_{\overline{\lambda}}(F)$ ON per il prodotto scalare hermitiano standard su $\mathbb{C}^n$ perchè

$$\langle \overline{T}_j, \overline{T}_k\rangle = \overline{T}_j^t \cdot \overline{T}_k = \overline{T_j^t \cdot T_k} = \overline{\delta_{jk}} = \delta_{jk}.$$

Poniamo

$$U_s := \frac{1}{\sqrt{2}}(T_s + \overline{T}_s), \qquad W_s := \frac{1}{\sqrt{2}}(iT_s - i\overline{T}_s).$$

Un facile calcolo mostra che U_s, W_s sono vettori reali, che sono ortonormali per il prodotto euclideo standard e che

$$f(U_s) = \cos\theta_s U_s + \sin\theta_s W_s, \qquad f(W_s) = -\sin\theta_s U_s + \cos\theta_s W_s. \tag{6.10.9}$$

Quindi se raccogliamo tutti i vettori vettori U_s e W_s ottenuti in tal modo e aggiungiamo i vettori $X_1 \dots, X_a, Y_1, \dots, Y_b$ abbiamo in tutto n vettori (per le equazioni in (6.10.6) e (6.10.8)) e perciò

$$X_1 \dots, X_a, Y_1, \dots, Y_b, U_1, W_1, U_2, W_2, \dots, U_c, W_c \tag{6.10.10}$$

è una base ON di V tale che valgano (6.10.4) e (6.10.5).

Ora trattiamo il caso generale. Sia $n := \dim V$. Per l'Esempio 6.8.4 esiste un isomorfismo

$$\varphi\colon (\mathbb{R}^n, \langle , \rangle_{st}) \to (V, \langle , \rangle),$$

dove $\langle , \rangle_{st}$ è il prodotto scalare standard. Siccome φ e f sono isometrie, lo è anche

$$\varphi^{-1} \circ f \circ \varphi\colon \mathbb{R}^n \to \mathbb{R}^n.$$

Quindi $\varphi^{-1} \circ f \circ \varphi$ è un elemento di $\mathrm{O}_n(\mathbb{R})$, e perciò esiste una decomposizione ortogonale data da (6.10.10) che ha le proprietà della tesi. Applicando φ a ogni addendo in (6.10.10) otteniamo una decomposizione ortogonale di V che ha le proprietà enunciate nella tesi. □

Sia V uno spazio vettoriale euclideo di dimensione finita. Ricordiamo che un'isometria di V, cioè un elemento di $\mathrm{O}(V)$, ha determinante ± 1 (vedi l'Osservazione 6.8.11), e che

$$\mathrm{SO}(V) = \{g \in \mathrm{O}(V) \mid \mathrm{Det}(g) = 1\} = \mathrm{O}(V) \cap \mathrm{GL}^+(V)$$

è un sottogruppo di $\mathrm{O}(V)$ (ricordiamo che $\mathrm{GL}^+(V)$ è il gruppo degli elementi di $\mathrm{GL}(V)$ con determinante positivo, vedi la Sezione 5.11).

Corollario 6.10.4. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo finitamente generato. Se $f \in \mathrm{SO}(V)$ esiste una decomposizione in somma diretta ortogonale*

$$V = V_1(f) \oplus_\perp B_1 \dots, \oplus_\perp B_d,$$

tale che per ogni $j \in \{1, \dots, d\}$ il sottospazio B_j ha dimensione 2, si ha $f(B_j) = B_j$ e la restrizione di f a B_j è una rotazione. Se invece $f \in (\mathrm{O}(V) \setminus \mathrm{SO}(V))$ (cioè $\det(f) = -1$) esiste una decomposizione in somma diretta ortogonale

$$V = V_1(f) \oplus_\perp W \oplus_\perp B_1 \dots, \oplus_\perp B_d,$$

tale che $\dim W = 1$, $f(v) = -v$ per ogni $v \in W$ e per ogni $j \in \{1, \dots, d\}$ il sottospazio B_j ha dimensione 2, si ha $f(B_j) = B_j$ e la restrizione di f a B_j è una rotazione.

Dimostrazione. Supponiamo che $f \in \mathrm{SO}(V)$. Per la Proposizione 6.10.3 abbiamo la decomposizione ortogonale data da (6.10.3). Siccome $1 = \det(f) = (-1)^{\dim V_{-1}(f)}$ la dimensione di $V_{-1}(f)$ è pari e quindi possiamo scegliere una decomposizione ortogonale $V_{-1}(f) = W_1 \oplus_\perp \ldots \oplus_\perp W_m$ dove $\dim W_k = 2$ per ogni $k \in \{1, \ldots, m\}$. Siccome la moltiplicazione per -1 su uno spazio vettoriale euclideo di dimensione 2 è uguale alla rotazione di angolo $\pm\pi$, basta porre $\{B_1, \ldots, B_d\} = \{A_1, \ldots, A_c, W_1, \ldots, W_m\}$. Se $f \in (\mathrm{O}(V) \setminus \mathrm{SO}(V))$ si procede in modo analogo. □

Esempio 6.10.5. Esaminiamo il contenuto del Corollario 6.10.4 nel caso in cui V abbia dimensione 2 o 3. Se $\dim V = 2$ allora gli elementi di $\mathrm{SO}(V)$ sono le rotazioni. Invece gli elementi di $(\mathrm{O}(V) \setminus \mathrm{SO}(V))$ sono le f con autovalori ± 1 e relativi autospazi ortogonali tra loro. Ora supponiamo che $\dim V = 3$. Se $f \in \mathrm{O}(V)$ allora esiste una decomposizione in somma diretta ortogonale

$$V = U \oplus_\perp B, \tag{6.10.11}$$

dove $\dim U = 1$, $\dim B = 2$, la restrizione di f a B è una rotazione e per $v \in U$ si ha $f(v) = v$ se $f \in \mathrm{SO}(V)$ e $f(v) = -v$ se $f \in (\mathrm{O}(V) \setminus \mathrm{SO}(V))$.

Dimostreremo che si può passare con continuità da ogni elemento di $\mathrm{SO}(V)$ a ogni altro elemento di $\mathrm{SO}(V)$, ma che non si può passare con continuità da un elemento di $\mathrm{SO}(V)$ a un elemento di $(\mathrm{O}(V) \setminus \mathrm{SO}(V))$. Prima diamo definizioni che sono l'analogo per $\mathrm{O}(V)$ di definizioni date nella Sezione 5.11 per $\mathrm{GL}(V)$.

Definizione 6.10.6. Sia V uno spazio vettoriale euclideo di dimensione finita, e sia $I \subset \mathbb{R}$ un intervallo. Un'applicazione $\gamma\colon I \to \mathrm{O}(V)$ è *continua* se la composizione $I \xrightarrow{\gamma} \mathrm{O}(V) \hookrightarrow \mathrm{GL}(V)$ è continua, cioè se scelta una base $\mathcal{B}$ (vedi l'Osservazione 5.11.2) l'applicazione

$$\begin{array}{ccc} I & \longrightarrow & M_{n,n}(\mathbb{R}) \\ t & \mapsto & M_{\mathcal{B}}^{\mathcal{B}}(\gamma(t)) \end{array}$$

è continua.

Esempio 6.10.7. L'applicazione

$$\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathrm{O}(2) \\ t & \mapsto & R_t = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \end{array}$$

è continua.

Definizione 6.10.8. Sia V uno spazio vettoriale euclideo di dimensione finita. Una isometria $g \in \mathrm{O}(V)$ è *deformabile in* in una isometria $h \in \mathrm{O}(V)$ se esiste un'applicazione continua $\gamma\colon [a, b] \to \mathrm{O}(V)$ tale che $\gamma(a) = g$ e $\gamma(b) = h$.

Esempio 6.10.9. Gli esempi 6.10.5 e 6.10.7 mostrano che se $g, h \in \mathrm{SO}(2)$, allora g è deformabile in h.

Per l'Osservazione 5.11.7 la relazione definita nella Definizione 6.10.8 è di equivalenza.

Proposizione 6.10.10. *Sia V uno spazio vettoriale euclideo di dimensione finita, e siano $g, h \in \mathrm{O}(V)$. Allora g è deformabile in h se e solo se $g, h \in \mathrm{SO}(V)$ oppure $g, h \in (\mathrm{O}(V) \setminus \mathrm{SO}(V))$.*

Dimostrazione. Se g e h non sono entrambi elementi di $\mathrm{SO}(V)$ o di $\mathrm{O}(V) \setminus \mathrm{SO}(V)$, allora non sono deformabili l'uno nell'altro per l'Osservazione 5.11.7.

Ora supponiamo che $g, h \in \mathrm{SO}(V)$ e dimostriamo che sono deformabili l'uno nell'altro. Siccome la relazione che stiamo considerando è di equivalenza, e siccome $\mathrm{Id}_V \in \mathrm{SO}(V)$, è sufficiente dimostrare che g è deformabile in Id_V. Questo segue immediatamente dal Corollario 6.10.4 e dall'Esempio 6.10.9.

Infine supponiamo che $f, g \in (\mathrm{O}(V) \setminus \mathrm{SO}(V))$. Si dimostra che f si deforma in g procedendo esattamente come nella dimostrazione della Proposizione 5.11.8: si ha $g^{-1} \cdot h \in \mathrm{SO}(V)$, e quindi per quanto appena dimostrato esiste un'applicazione continua $\gamma\colon [a, b] \to \mathrm{GL}_n(V))$ tale che $\gamma(0) = g^{-1} \cdot h$ e $\gamma(1) = \mathrm{Id}_V$. L'applicazione $\varphi\colon [a, b] \to \mathrm{GL}(V)$ definita da $\varphi(t) := g \cdot \gamma(t)$ è continua e si ha $\varphi(0) = h$, $\varphi(1) = g$. □

Osservazione 6.10.11. Sia V uno spazio vettoriale reale di dimensione finita. La relazione tra elementi di $\mathrm{GL}(V)$ definita da $g \sim h$ se g è deformabile con continuità in h è di equivalenza, e analogamente per la relazione definita in modo simile per elementi di $\mathrm{O}(V, \langle , \rangle)$ se $\langle , \rangle$ è un prodotto scalare euclideo.

Esercizi del Capitolo 6

Esercizio 6.1. *Siano*

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, \qquad B := \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$$

e per $X, Y \in \mathbb{R}^2$ *poniamo*

$$\langle X, Y \rangle_A := X^t \cdot A \cdot Y, \qquad \langle X, Y \rangle_B := X^t \cdot B \cdot Y.$$

Verificate che $\langle , \rangle_A$ *e* $\langle , \rangle_B$ *sono prodotti euclidei.*

Esercizio 6.2. *Sia*

$$U := \{X \in \mathbb{R}^3 \mid x_1 + 2x_2 + 3x_3 = 0\}.$$

Sia $v := (1, 1, 1)$. *Determinate la proiezione ortogonale di* v *su* U, *se* $\mathbb{R}^3$ *ha il prodotto scalare standard.*

Esercizio 6.3. *Sia* $(V, \langle , \rangle)$ *uno spazio vettoriale euclideo. Sia* $0 \neq v \in V$. *Definiamo*

$$\begin{array}{ccc} V & \xrightarrow{R_v} & V \\ w & \mapsto & w - 2\frac{\langle w, v \rangle}{||v||^2} v \end{array}$$

(1) Dimostrate che $R_v \in O(V)$.

(2) Interpretate geometricamente R_v *nel caso in cui* $V = \mathsf{V}(\mathbb{E}^m)$ *con* $m \in \{2, 3\}$.

Capitolo 7

Spazi affini euclidei

7.1 Definizione e prime proprietà

Definizione 7.1.1. Uno *spazio affine euclideo* è una coppia $(\mathbb{S}, \langle\,,\rangle)$, dove $\mathbb{S}$ è uno spazio affine reale, e $\langle\,,\rangle$ è un prodotto scalare euclideo sullo spazio vettoriale $\mathsf{V}(\mathbb{S})$ delle traslazioni di $\mathbb{S}$.

Definizione 7.1.2. Sia $(\mathbb{S}, \langle\,,\rangle)$ uno spazio affine euclideo. La *distanza* tra punti $P, Q \in \mathbb{S}$ è

$$d(P,Q) := ||\overrightarrow{PQ}||.$$

Esempio 7.1.3. Il piano euclideo $\mathbb{E}^2$ e lo spazio euclideo $\mathbb{E}^3$ con il prodotto scalare delle scuole (definito dopo aver introdotto una unità di misura) sono spazi affini euclidei. La distanza è quella definita dall'unità di misura.

Esempio 7.1.4. Uno spazio vettoriale euclideo $(V, \langle\,,\,\rangle)$ è anche uno spazio affine euclideo, se si dà a V la struttura di spazio affine in cui $T(V) = V$ e l'azione è definita dall'addizione. La distanza è data da

$$d(v,w) = ||v-w|| = \left(\langle v-w, v-w\rangle\right)^{1/2}.$$

In particolare $\mathbb{A}^n(\mathbb{R})$, cioè $\mathbb{R}^n$ visto come spazio affine (su $\mathbb{R}$), con il prodotto scalare standard, cioè

$$\langle X, Y\rangle = \sum_{i=1}^{n} x_i y_i$$

per $X, Y \in \mathsf{V}(\mathbb{A}^n_{\mathbb{R}}) = \mathbb{R}^n$ è uno spazio affine euclideo. La distanza è data da

$$d(X,Y) = \left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{1/2}.$$

Denotiamo $(\mathbb{A}^n(\mathbb{R}), \langle\,,\rangle)$ con $\mathbb{E}^n(\mathbb{R})$. Questo è lo spazio affine euclideo *standard*.

Osservazione 7.1.5. La nozione di spazio affine euclideo è modellata su quella di piano euclideo e spazio euclideo (vedi l'Esempio 7.1.3). Ci permette di trattare analoghi del piano euclideo (e dello spazio euclideo) di dimensione arbitraria.

Proposizione 7.1.6. *Sia $(\mathbb{S}, \langle\,,\rangle)$ uno spazio affine euclideo, e sia $d(\,,\,)$ la distanza associata. Allora*

1. *Per ogni $P, Q \in \mathbb{S}$ si ha $d(P,Q) \geqslant 0$, e $d(P,Q) = 0$ se e solo se $P = Q$.*
2. *$d(P,Q) = d(Q,P)$ per ogni $P, Q \in \mathbb{S}$.*
3. *Siano $P, Q, R \in \mathbb{S}$. Allora*
$$d(P,R) \leqslant d(P,Q) + d(Q,R)$$
e si ha eguaglianza se e solo se Q è un punto dell'inviluppo convesso di P e R, cioè esistono $s, t \geqslant 0$ con $s + t = 1$ tali che
$$Q = sP + tQ. \tag{7.1.1}$$

Dimostrazione. (1): $d(P,Q) = ||\overrightarrow{PQ}||$, quindi è non negativo, ed è nullo se e solo se $\overrightarrow{PQ} = 0$, cioè se e solo se $P = Q$. (2): $d(P,Q) = ||\overrightarrow{PQ}||$ e $d(Q,P)| = ||\overrightarrow{QP}||$. Siccome $\overrightarrow{QP} = -\overrightarrow{PQ}$, segue che $d(P,Q) = d(Q,P)$. (3): Per la diseguaglianza triangolare abbiamo

$$d(P,R) = ||\overrightarrow{PR}| = ||\overrightarrow{PQ} + \overrightarrow{QR}|| \leqslant ||\overrightarrow{PQ}|| + ||\overrightarrow{QR}| = d(P,Q) + d(Q,R).$$

Inoltre sappiamo che si ha eguaglianza se e solo se esiste $\lambda \geqslant 0$ tale che $\overrightarrow{PQ} = \lambda\overrightarrow{QR}$ oppure $\overrightarrow{QR} = \lambda\overrightarrow{PQ}$. Supponiamo che $\overrightarrow{PQ} = \lambda\overrightarrow{QR}$, e quindi

$$\frac{\lambda}{1+\lambda}\overrightarrow{QR} - \frac{1}{1+\lambda}\overrightarrow{PQ} = 0.$$

Allora vale (7.1.1) con $s = \frac{1}{1+\lambda}$ e $t = \frac{\lambda}{1+\lambda}$ (notate che sono non negativi, e che la somma è 1). Infatti

$$\frac{1}{1+\lambda}P + \frac{\lambda}{1+\lambda}R = Q + \frac{1}{1+\lambda}\overrightarrow{QP} + \frac{\lambda}{1+\lambda}\overrightarrow{QR} = Q - \frac{1}{1+\lambda}\overrightarrow{PQ} + \frac{\lambda}{1+\lambda}\overrightarrow{QR} = Q.$$

Analogamente si dimostra che se $\overrightarrow{QR} = \lambda\overrightarrow{PQ}$ allora Q è un punto dell'inviluppo convesso di P e R.

Il viceversa, cioè che se Q è un punto dell'inviluppo convesso di P e R allora la diseguaglainza triangolare è una eguaglianza si dimostra in modo simile. □

Se $(\mathbb{S}, \langle\,,\rangle)$ è uno spazio affine euclide, un sottoinsieme $L \subset \mathbb{S}$ è una *semiretta* se esistono un punto $P_0 \in \mathbb{S}$ e un vettore $v \in \mathsf{V}(\mathbb{S})$ tali che

$$L = \{P_0 + tv \mid t \geqslant 0\}. \tag{7.1.2}$$

Data una semiretta L, il punto P_0 è l'unico punto tale che valga l'eguaglianza in (7.1.2) (facile dimostrazione): è l'*origine* di L. Il vettore v non è univocamente determinato, infatti è determinato a meno di moltiplicarlo per un reale (strettamente) positivo. Siano $L, M \subset \mathbb{S}$ semirette, cioè

$$L = \{P_0 + tv \mid t \geqslant 0\}, \quad M = \{Q_0 + sw \mid t \geqslant 0\}.$$

L'*angolo* tra L e M è l'angolo tra v e w (vedi la Definizione 6.3.5). Notate che l'angolo è ben definito perchè v e w sono determinati a meno di riscalamento per un reale positivo.

La definizione di angolo tra rette non ha senso perchè l'angolo tra vettori non nulli v e w non è uguale all'angolo tra v e $-w$, a meno che $\langle v, w\rangle$ sia nullo. In questo caso diciamo che le rette sono *ortogonali*. Più in generale, sottospazi affini $A, B \subset \mathbb{S}$ sono *ortogonali* se i sottospazi vettoriali $\mathsf{V}(A), \mathsf{V}(B) \subset \mathsf{V}(\mathbb{S})$ sono ortogonali (vedi la Definizione 6.3.6).

Sia $(\mathbb{S}, \langle\,,\rangle)$ uno spazio affine euclideo di dimensione finita. Ricordiamo che un riferimento affine è un isomorfismo di spazi affini

$$\mathbb{S} \xrightarrow{X} \mathbb{A}^n(\mathbb{R})$$

ed è determinato dall'origine $O \in \mathbb{S}$ e da una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ di $\mathsf{V}(\mathbb{S})$. L'inversa X^{-1} è data da

$$\begin{array}{ccc} \mathbb{A}^n(\mathbb{R}) & \xrightarrow{X^{-1}} & \mathbb{S} \\ (x_1, \ldots, x_n) & \mapsto & O + \sum_{i=1}^n x_i v_i \end{array}$$

Se $P \in \mathbb{S}$ le entrate di $X(P)$ sono le coordinate di P nel riferimento affine $RA(O, \mathcal{B})$.

Definizione 7.1.7. Sia $(\mathbb{S}, \langle\,,\rangle)$ uno spazio affine euclideo di dimensione finita. Un riferimento affine di $(\mathbb{S}, \langle\,,\rangle)$ è *ortonormale* se la base $\mathcal{B}$ di $\mathsf{V}(\mathbb{S})$ è ON, e si denota $RO(O, \mathcal{B})$.

Proposizione 7.1.8. *Sia $RO(O, \mathcal{B})$ un riferimento ortonormale di $(\mathbb{S}, \langle\,,\rangle)$. Se*

$$P(x_1, \ldots, x_n), Q(y_1, \ldots, y_n) \in \mathbb{S},$$

allora

$$d(P,Q) = \left(\sum_{i=1}^n (x_i - y_i)^2\right)^{1/2}.$$

Dimostrazione. Per definizione

$$d(P,Q)) = ||\overrightarrow{PQ}|| = ||\sum_{i=1}^n (y_i - x_i)v_i|| = \left(\sum_{i=1}^n (x_i - y_i)^2\right)^{1/2}.$$

□

Convenzione 7.1.9. Da ora in poi uno spazio affine euclideo $(\mathbb{S}, \langle\,,\rangle)$ verrà denotato semplicemente con $\mathbb{S}$.

7.2 Ginnastica affine euclidea

Sia $\mathbb{S}$ uno spazio affine euclideo di dimensione finita. Passeremo in rassegna alcune costruzioni legate alla presenza della distanza tra punti di $\mathbb{S}$ e alla nozione di angolo tra vettori (non nulli) di $\mathsf{V}(\mathbb{S})$. Per svolgere gli esercizi "ginnici" sarà conveniente introdurre un riferimento ortonormale $RO(O, \mathcal{B})$, dove $\mathcal{B} = \{v_1, \ldots, v_n\}$. La prima osservazione è che le equazioni cartesiane di un sottospazio affine $\mathbb{T} \subset \mathbb{S}$

$$\begin{aligned} a_{1,1}x_1 + \ldots + a_{1,n}x_n + b_1 &= 0, && (7.2.1)\\ \ldots\ldots\ldots\ldots\ldots\ldots &= 0, && (7.2.2)\\ a_{i,1}x_1 + \ldots + a_{i,n}x_n + b_i &= 0, && (7.2.3)\\ \ldots\ldots\ldots\ldots\ldots\ldots &= 0, && (7.2.4)\\ a_{m,1}x_1 + \ldots + a_{m,n}x_n + b_m &= 0, && (7.2.5)\\ & && (7.2.6) \end{aligned}$$

si possono interpretare come

$$\mathbb{T} = \{P \in \mathbb{S} \mid \langle \overrightarrow{P_0P}, u_1 \rangle = \ldots = \langle \overrightarrow{P_0P}, u_m \rangle = 0\}, \tag{7.2.7}$$

dove

$$u_i := a_{i,1}v_1 + \ldots + a_{i,n}v_n, \tag{7.2.8}$$

e $P_0(\overline{x}_1, \ldots, \overline{x}_n) \in \mathbb{T}$ è un punto fissato. Infatti $P(x_1, \ldots, x_n) \in \mathbb{T}$ se e solo se $(x_1 - \overline{x}_1, \ldots, x_n - \overline{x}_n)$ è soluzione del sistema omogeneo associato al sistema in (7.2.1), ovvero se e solo se il vettore $\overrightarrow{P_0P} = (x_1 - \overline{x}_1)v_1 + \ldots + (x_n - \overline{x}_n)v_n$ è ortogonale al vettore u_i per $i \in \{1, \ldots, m\}$. Vediamo anche che

$$\mathsf{V}(\mathbb{T}) = \{u_1, \ldots, u_m\}^{\perp}. \tag{7.2.9}$$

Ora diamo un analogo affine della proiezione ortogonale su un sottospazio (vedi la Definizione 6.5.5).

Proposizione 7.2.1. *Siano $\mathbb{S}$ uno spazio affine euclideo di dimensione finita e $\mathbb{T} \subset \mathbb{S}$ un sottospazio affine. Dato $P \in \mathbb{S}$ esiste un punto $Q_0 \in \mathbb{T}$ che minimizza la distanza tra punti di $\mathbb{T}$ e P, cioè*

$$d(P, Q_0) \leqslant d(P, Q) \quad \forall Q \in \mathbb{T}, \tag{7.2.10}$$

e tale punto è unico, cioè si ha equaglianza solo se $Q = Q_0$.

Dimostrazione. Sia $P_0 \in \mathbb{T}$. Sia $v := \pi_{\mathsf{V}(\mathbb{T})}(\overrightarrow{P_0P})$ la proiezione ortogonale di $\overrightarrow{P_0P}$ su $\mathsf{V}(\mathbb{T})$, vedi la Definizione 6.5.5. Poniamo $Q_0 := P_0 + v$, e quindi

$$\overrightarrow{P_0P} = \overrightarrow{P_0Q_0} + \overrightarrow{Q_0P} = v + \overrightarrow{Q_0P}. \tag{7.2.11}$$

Per definizione di proiezione ortogonale il vettore $\overrightarrow{Q_0P}$ è ortogonale a ogni vettore di $\mathsf{V}(\mathbb{T})$.

Ora sia $Q \in \mathbb{T}$. Allora $\overrightarrow{PQ} = \overrightarrow{PQ_0} + \overrightarrow{Q_0Q}$, e siccome $\overrightarrow{Q_0Q} \in \mathsf{V}(\mathbb{T})$ i vettori $\overrightarrow{PQ_0}$ e $\overrightarrow{Q_0Q}$ sono ortogonali. Quindi

$$d(P, Q_0)^2 = ||\overrightarrow{PQ_0}||^2 \leqslant ||\overrightarrow{PQ_0}||^2 + ||\overrightarrow{Q_0Q}||^2 = ||\overrightarrow{PQ}||^2 = d(P, Q)^2.$$

Questo mostra che vale (7.2.10), e anche che l'uguaglianza vale solo se $Q = Q_0$. □

Definizione 7.2.2. Siano $\mathbb{S}$ uno spazio affine euclideo di dimensione finita e $\mathbb{T} \subset \mathbb{S}$ un sottospazio affine. La *proiezione ortogonale* di un punto $P \in \mathbb{S}$ su $\mathbb{T}$ è l'unico punto $Q_0 \in \mathbb{T}$ che minimizza la distanza tra punti di $\mathbb{T}$ e P, cioè tale che valga (7.2.10). La *distanza* tra P e $\mathbb{T}$ è la distanza tra P e la sua proiezione ortogonale su $\mathbb{T}$ ed è denotata $d(P, \mathbb{T})$.

Per la Proposizione 7.2.1 si ha

$$d(P, \mathbb{T}) = \min\{d(P, Q) \mid Q \in \mathbb{T}\}.$$

Esempio 7.2.3. Nelle ipotesi della Proposizione 7.2.1 calcoliamo esplicitamente la proiezione ortogonale di un punto sul sottospazio $\mathbb{T}$. Sia $RO(O, \mathcal{B})$ un riferimento ortonormale di $\mathbb{T}$. Se estendiamo $\mathcal{B}$ a una base $\widetilde{\mathcal{B}}$ di $\mathsf{V}(\mathbb{S})$ allora $RO(O, \widetilde{\mathcal{B}})$ è un riferimento ortonormale di $\mathbb{T}$. Siano m e n le dimensioni di $\mathbb{T}$ e $\mathbb{S}$ rispettivamente. Allora nelle coordinate $(x_1, \ldots, x_n)$ il sottospazio $\mathbb{T}$ ha equazioni cartesiane

$$x_{m+1} = \ldots = x_n = 0.$$

La proiezione ortogonale su $\mathbb{T}$ è data da

$$\begin{array}{ccc} \mathbb{S} & \longrightarrow & \mathbb{T} \\ P(x_1,\dots,x_n) & \mapsto & Q(x_1,\dots,x_m,0,\dots,0) \end{array} \tag{7.2.12}$$

In particolare vediamo che l'applicazione $\pi\colon \mathbb{S} \to \mathbb{T}$ che associa a P la sua proiezione ortogonale è affine. In effetti π è la proiezione su $\mathbb{T}$ parallela a $\mathsf{V}(\mathbb{T})^\perp$ definita nell'Esempio 4.4.5

Esempio 7.2.4. Di norma è laborioso cambiare coordinate affini in modo da rientrare nell'Esempio 7.2.3. Vediamo come procedere direttamente nel caso in cui $\mathbb{T} \subset \mathbb{S}$ abbia codimensione 1. Supponiamo che nelle cordinate $(x_1,\dots,x_n)$ di un riferimento ON $\mathbb{T}$ abbia equazione

$$f(x_1,\dots,x_n) = a_1x_1 + \dots + a_nx_n + b = 0,$$

a calcoliamo la proiezione ortogonale di $P(c_1,\dots,c_n)$ su $\mathbb{T}$ e la distanza $d(P,\mathbb{T})$. Sia

$$u := a_1v_1 + \dots + a_nv_n.$$

Allora (vedi (7.2.9))

$$\mathsf{V}(\mathbb{T}) = u^\perp.$$

Scegliamo $P_0(-b/a_1,0,\dots,0) \in \mathbb{T}$, supponendo che $a_1 \neq 0$. Abbiamo

$$\overrightarrow{P_0P} := \left(c_1 + \frac{b}{a_1}\right)v_1 + \sum_{i=2}^{n} c_iv_i.$$

La proiezione ortogonale Q_0 di P su $\mathbb{T}$ è caratterizzata dall'equazione

$$\overrightarrow{P_0P} = \overrightarrow{P_0Q_0} + xu,$$

dove $x \in \mathbb{R}$ è un'incognita. Notate anche che

$$d(P,\mathbb{T}) = ||xu|| = |x| \cdot ||u|| = |x| \cdot (\sum_{i=1}^{n} u_i^2)^{1/2}.$$

Quindi si tratta di determinare x. Siccome $\overrightarrow{P_0Q_0}$ è ortogonale a u abbiamo

$$\langle \overrightarrow{P_0P}, u\rangle = \langle \overrightarrow{P_0Q_0} + xu, u\rangle = x||u||^2.$$

Ma

$$\langle \overrightarrow{P_0P}, u\rangle = f(c),$$

e quindi $x = f(c_1,\dots,c_n)/||u||^2$. Segue che

$$d(P,\mathbb{T}) = \frac{|f(c_1,\dots,c_n)|}{||u||^{1/2}}$$

e la proiezione ortogonale Q_0 di P su $\mathbb{T}$ è data da

$$Q_0\left(c_1 - \frac{f(c_1,\dots,c_n)u_1}{||u||^2},\dots,c_n - \frac{f(c_1,\dots,c_n)u_n}{||u||^2}\right). \tag{7.2.13}$$

Le stesse formule valgono se $a_1 = 0$.

7.3 Applicazioni che preservano le distanze

Definizione 7.3.1. Siano $\mathbb{S}$ e $\mathbb{T}$ spazi affini euclidei. Un'applicazione $\Phi\colon \mathbb{S} \to \mathbb{T}$ *preserva le distanze* se per ogni $P, Q \in \mathbb{S}$ si ha

$$d(\Phi(P),\Phi(Q)) = d(P,Q).$$

ed è una *isometria* se preserva le distanze ed è biunivoca.

Osservazione 7.3.2. Se Φ preserva le distanze, allora è iniettiva. Quindi è biunivoca se e solo se è suriettiva. Inoltre se è una isometria, anche l'inversa preserva le distanze (e quindi è una isometria).

Esempio 7.3.3. Se $\mathbb{S}$ ha dimensione finita e $RO(O,\mathcal{B})$ è un riferimento ON, allora l'isomorfismo di spazi affini $X\colon \mathbb{S} \xrightarrow{\sim} \mathbb{E}^n(\mathbb{R})$ definito da $RO(O,\mathcal{B})$ è una isometria.

Esempio 7.3.4. Siano $\mathbb{S}$ e $\mathbb{T}$ spazi affini euclidei, e sia $\varphi\colon \mathsf{V}(\mathbb{S}) \to \mathsf{V}(\mathbb{T})$ un'applicazione lineare tale che

$$||\varphi(v)|| = ||v||$$

per ogni $v \in \mathsf{V}(\mathbb{S})$. Siano $P_0 \in \mathbb{S}$ e $Q_0 \in \mathbb{T}$. Definiamo

$$\begin{array}{ccc} \mathbb{S} & \xrightarrow{\Phi} & \mathbb{T} \\ P & \mapsto & Q_0 + \varphi(\overrightarrow{P_0P}) \end{array}$$

Allora Φ preserva le distanze. Infatti

$$d(\Phi(P),\Phi(R)) = ||\overrightarrow{\Phi(P)\Phi(R)}| = ||\varphi(\overrightarrow{PQ})|| = ||\overrightarrow{PQ}||| = d(P,Q).$$

Il prossimo risultato afferma che in dimensione finita ogni isometria è affine. (Vedi l'Esempio 7.3.4.)

Teorema 7.3.5. *Sia $\mathbb{S}$ uno spazio affine euclideo di dimensione finita. Se $\Phi\colon \mathbb{S} \to \mathbb{T}$ preserva le distanze allora esistono un'applicazione lineare $\varphi\colon \mathsf{V}(\mathbb{S}) \to \mathsf{V}(\mathbb{T})$ tale che*

$$||\varphi(v)|| = ||v|| \quad \forall v \in \mathsf{V}(\mathbb{S}), \tag{7.3.1}$$

e punti $P_0 \in \mathbb{S}$, $Q_0 \in \mathbb{T}$ tali che

$$\Phi(P) = Q_0 + \varphi(\overrightarrow{P_0P})$$

In particolare Φ è un'applicazione affine.

Per dimostrare il Teorema 7.3.5 abbiamo bisogno del seguente risultato.

Proposizione 7.3.6. *Siano $(V,\langle,\rangle)$ e $(W,\langle,\rangle)$ spazi vettoriali euclidei con $\dim V < \infty$. Sia $\varphi\colon V \to W$ un'applicazione (a priori non necessariamente lineare) tale che*

1. *$\varphi(0) = 0$ e*
2. *$||\varphi(v) - \varphi(w)|| = ||v - w||$ per ogni $v, w \in V$.*

Allora φ è un'applicazione lineare e per ogni $v \in V$ vale

$$||\varphi(v)|| = ||v||. \tag{7.3.2}$$

Dimostrazione. Iniziamo notando che per ogni $v \in V$ si ha

$$||\varphi(v)|| = ||\varphi(v) - 0||| = ||\varphi(v) - \varphi(0)|| = ||v - 0|| = ||v||, \tag{7.3.3}$$

cioè vale (7.3.2). Ora dimostriamo che per $v, w \in V$ si ha

$$\langle\varphi(v),\varphi(w)\rangle = \langle v, w\rangle. \tag{7.3.4}$$

Infatti abbiamo

$$||\varphi(v)||^2 - 2\langle\varphi(v),\varphi(w)\rangle + ||\varphi(w)||^2 = ||\varphi(v) - \varphi(w)||^2 = ||v - w||^2 = ||v||^2 - 2\langle v, w\rangle + ||w||^2,$$

e per (7.3.3) si ha $||\varphi(v)||^2 = ||v||^2$ e $||\varphi(w)||^2 = ||w||^2$, da cui segue che vale (7.3.4).

Sia $\{v_1,\dots,v_n\}$ una base ON di V. Per quello che abbiamo dimostrato finora $\varphi(v_1),\dots,\varphi(v_n)$ sono vettori ortonormali di W.

L'immagine di φ è contenuta nel sottospazio $\operatorname{Span}(\varphi(v_1),\dots,\varphi(v_n))$ (se sapessimo che φ è lineare questo sarebbe ovvio, ma stiamo appunto dimostrando che è lineare). Infatti, se non fosse così esisterebbe $v_{n+1} \in V$ tale che

$$\dim \operatorname{Span}(\varphi(v_1),\dots,\varphi(v_n),\varphi(v_{n+1})) = n + 1. \tag{7.3.5}$$

Per (7.3.4) la matrice $(n+1)\times(n+1)$ con entrate $\langle v_i, v_j\rangle$ per $i,j \in \{1,\dots,(n+1)\}$ è uguale alla matrice con entrate $\langle\varphi(v_i),\varphi(v_j)\rangle$. Questo è assurdo: la prima matrice ha rango n perchè $\dim V = n$, mentre la seconda ha rango $(n+1)$ per (7.3.5). (Se questo argomento non è chiaro, assumete che $\dim V = \dim W$, allora l'immagine di φ è contenuta nel sottospazio $\operatorname{Span}(\varphi(v_1),\dots,\varphi(v_n))$ perchè $\{\varphi(v_1),\dots,\varphi(v_n)\}$ è una base ON di W.)

Sia $v \in V$. Siccome $\{v_1,\dots,v_n\}$ è una base ON di V, sicome l'immagine di φ è contenuta nel sottospazio $\operatorname{Span}(\varphi(v_1),\dots,\varphi(v_n))$ e $\{\varphi(v_1),\dots,\varphi(v_n)\}$ è una base ON di tale sottospazio, e siccome vale (7.3.4), abbiamo

$$\varphi\left(\sum_{i=1}^n \langle v, v_i\rangle v_i\right) = \varphi(v) = \sum_{i=1}^n \langle\varphi(v),\varphi(v_i)\rangle\varphi(v_i) = \sum_{i=1}^n \langle v, v_i\rangle\varphi(v_i).$$

Questo dimostra che φ è un'applicazione lineare. □

Dimostrazione del Teorema 7.3.5. Scegliamo $P_0 \in \mathbb{S}$ e definiamo $\varphi\colon \mathsf{V}(\mathbb{S}) \to \mathsf{V}(\mathbb{T})$ ponendo

$$\varphi(\overrightarrow{P_0P}) = \overrightarrow{\Phi(P_0)\Phi(P)}.$$

Dimostriamo che φ è un'applicazione lineare e che vale (7.3.1) applicando la Proposizione 7.3.6. Dobbiamo controllare che valgano le ipotesi. È chiaro che $\varphi(0) = 0$. Per vedere che vale la seconda ipotesi, siano $P, Q \in \mathbb{S}$. Allora

$$||\varphi(\overrightarrow{P_0P}) - \varphi(\overrightarrow{P_0Q}||) = ||\overrightarrow{\Phi(P_0)\Phi(P)} - \overrightarrow{\Phi(P_0)\Phi(Q)}|| = \\ = ||\overrightarrow{\Phi(Q)\Phi(P)}|| = d(\Phi(P), \Phi(Q)) = d(P, Q) = ||\overrightarrow{PQ}|| = ||\overrightarrow{P_0P} - \overrightarrow{P_0Q}||.$$

Questo dimostra che vale anche la seconda ipotesi. Quindi per la Proposizione 7.3.6 l'applicazione φ è lineare e vale (7.3.1). Ora sia $Q_0 := \Phi(P_0)$. Allora per $P \in \mathbb{S}$ si ha

$$\Phi(P) = Q_0 + \overrightarrow{\Phi(P_0)\Phi(P)} = Q_0 + \varphi(\overrightarrow{P_0P}).$$

□

Il risultato che segue è una conseguenza immediata del Teorema 7.3.5 e dell'Esempio 7.3.4.

Corollario 7.3.7. *Sia $\mathbb{S}$ uno spazio affine euclideo di dimensione finita. Un'applicazione $\Phi\colon \mathbb{S} \to \mathbb{T}$ preserva le distanze se e solo se è affine e l'applicazione lineare $\mathsf{V}(\Phi)\colon \mathsf{V}(\mathbb{S}) \to \mathsf{V}(\mathbb{T})$ è un'isometria di spazi vettoriali euclidei.*

Esempio 7.3.8. Siano $\mathbb{S}$ uno spazio affine euclideo di dimensione finita e $\mathbb{T} \subset \mathbb{S}$ un sottospazio affine. La *riflessione nel sottospazio* $\mathbb{T}$ è l'applicazione $R_{\mathbb{T}}\colon \mathbb{S} \to \mathbb{S}$ definita come segue. Dato $P \in \mathbb{S}$, sia $Q_0 \in \mathbb{T}$ la proiezione ortogonale di P su $\mathbb{T}$ (vedi la Definizione 7.2.2); poniamo

$$R_{\mathbb{T}}(P) := Q_0 + \overrightarrow{PQ_0}.$$

Sia $RO(O, \widetilde{\mathcal{B}})$ un riferimento ON su $\mathbb{S}$ come nell'Esempio 7.2.3 (e quindi $\mathbb{T}$ ha equazioni cartesiane $x_{m+1} = \ldots = x_n = 0$). Nelle coordinate $(x_1, \ldots, x_n)$ la riflessione è data da

$$\begin{array}{ccc} \mathbb{S} & \xrightarrow{R_{\mathbb{T}}} & \mathbb{S} \\ P(x_1, \ldots, x_n) & \mapsto & Q(x_1, \ldots, x_m, -x_{m+1}, \ldots, -x_n) \end{array} \tag{7.3.6}$$

Infatti questo segue dalla formula (7.2.12) per la proiezione di $\mathbb{S}$ su $\mathbb{T}$. Come conseguenza vediamo che la riflessione è un'isometria di $\mathbb{S}$. Notiamo anche $R_{\mathbb{T}} \circ R_{\mathbb{T}} = \mathrm{Id}_{\mathbb{S}}$.

7.4 Isometrie

Considerazioni generali

Definizione 7.4.1. Siano $\mathbb{S}$ e $\mathbb{T}$ spazi affini euclidei. Un'applicazione $\Phi\colon \mathbb{S} \to \mathbb{T}$ è un'*isometria* se preserva le distanze ed è biunivoca.

Lasciamo al lettore la (semplice) dimostrazione del seguente risultato.

Proposizione 7.4.2. *Se $\mathbb{S}$ è uno spazio affine euclideo, allora l'idenità $\mathrm{Id}\colon \mathbb{S} \to \mathbb{S}$ è un'isometria. Se $\Phi\colon \mathbb{S} \to \mathbb{T}$ è un'isometria di spazi affini euclidei, allora anche l'inversa $\Phi^{-1}\colon \mathbb{T} \to \mathbb{S}$ è un'isometria. Se $\Psi\colon \mathbb{T} \to \mathbb{U}$ è un'altra isometria, allora anche la composizione $\Psi \circ \Phi\colon \mathbb{S} \to \mathbb{U}$ è un'isometria.*

Definizione 7.4.3. Spazi affini euclidei $\mathbb{S}$ e $\mathbb{T}$ sono *isometrici* se esiste un'isometria $\Phi\colon \mathbb{S} \to \mathbb{T}$. (La terminologia ha senso perchè esiste un'isometria $\Phi\colon \mathbb{S} \to \mathbb{T}$ se solo se esiste un'isometria $\Psi\colon \mathbb{T} \to \mathbb{S}$, vedi la Proposizione 7.4.2.)

Proposizione 7.4.4. *Spazi affini euclidei della stessa dimensione finita sono isometrici.*

Dimostrazione. Siano $\mathbb{S}, \mathbb{T}$ spazi affini euclidei di dimensione n. Per l'Esempio 7.3.3 esistono isometrie

$$\mathbb{S} \xrightarrow{\Phi} \mathbb{E}^n(\mathbb{R}), \qquad \mathbb{T} \xrightarrow{\Phi} \mathbb{E}^n(\mathbb{R}),$$

e la composizione $\Psi^{-1} \circ \Phi\colon \mathbb{S} \to \mathbb{T}$ è un'isometria per la Proposizione 7.4.2. □

Definizione 7.4.5. Sia $\mathbb{S}$ uno spazio affine euclideo. Per la Proposizione 7.4.2 la composizione di due isometrie di $\mathbb{S}$ (cioè isometrie da $\mathbb{S}$ a $\mathbb{S}$) è un'isometria di $\mathbb{S}$, e l'insieme delle isometrie di $\mathbb{S}$ con l'operazione di composizione è un gruppo. Lo denotiamo $\mathrm{Isom}(\mathbb{S})$.

Esempio 7.4.6. Sia $\mathbb{E}^n(\mathbb{R})$ lo spazio affine euclideo standard, vedi l'Esempio 7.1.4. Per il Teorema 7.3.5 ogni isometria di $\mathbb{E}^n(\mathbb{R})$ è un'applicazione affine la cui applicazione lineare associata $\mathbb{R}^n \to \mathbb{R}^n$ è nel gruppo ortogonale (preserva $\langle , \rangle$). Viceversa ogni tale applicazione affine è un'isometria. Quindi gli elementi di $\mathrm{Isom}(\mathbb{E}^n(\mathbb{R}))$ sono dati da

$$\begin{array}{ccc} \mathbb{E}^n(\mathbb{R}) & \longrightarrow & \mathbb{E}^n(\mathbb{R}) \\ X & \mapsto & A \cdot X + B \end{array} \tag{7.4.1}$$

dove $A \in \mathrm{O}_n(\mathbb{R})$ e $B \in M_{n,1}(\mathbb{R})$ è una matrice colonna.

Osservazione 7.4.7. Se $\mathbb{S}, \mathbb{T}$ sono spazi affini euclidei e $\Phi\colon \mathbb{S} \to \mathbb{T}$ è un'isometria, allora i gruppi $\mathrm{Isom}(\mathbb{S})$ e $\mathrm{Isom}(\mathbb{T})$ sono del tutto indistinguibili. Infatti definiamo un isomorfismo tra $\mathrm{Isom}(\mathbb{S})$ e $\mathrm{Isom}(\mathbb{T})$ così:

$$\begin{array}{ccc} \mathrm{Isom}(\mathbb{S}) & \longrightarrow & \mathrm{Isom}(\mathbb{T}) \\ f & \mapsto & \Phi \circ f \circ \Phi^{-1} \end{array} \tag{7.4.2}$$

Quindi per la Proposizione 7.4.4 il gruppo delle isometrie di un qualsiasi spazio affine euclideo di dimensione n è isomorfo al gruppo delle isometrie di $\mathbb{E}^n(\mathbb{R})$.

Osservazione 7.4.8. Una maniera equivalente di formulare la stessa affermazione è la seguente. Se $\mathbb{S}$ è uno spazio affine euclideo di dimensione finita n e $RO(O, \mathcal{B})$ è un riferimento ON di $\mathbb{S}$, allora $f\colon \mathbb{S} \to \mathbb{S}$ è un'isometria se e solo se è un'applicazione affine che nelle coordinate X di $RO(O, \mathcal{B})$ è data da (7.4.1) con $A \in \mathrm{O}_n(\mathbb{R})$ e $B \in M_{n,1}(\mathbb{R})$ una matrice colonna.

Chiudiamo con una considerazione che motiva la parte rimanente della presente sezione. In apparenza l'Esempio 7.4.6 e l'Osservazione 7.4.7 rispondono alla domanda: come si descrivono le isometrie di uno spazio affine euclideo (di dimensione finita)? In verità la descrizione esplicita delle isometrie contenuta nell'Esempio 7.4.6 non è illuminante. Se, per esempio, ci chiediamo come si scrive la stessa isometria in coordinate cartesiane diverse, la risposta non è particolarmente semplice. In altre parole, mentre la matrice ortogonale A rappresenta l'applicazione lineare $\mathsf{V}(\Phi)$, la traslazione definita da B non ha alcun significato intrinseco. Nelle prossime sottosezioni cercheremo di dare descrizioni geometriche (cioè non legate alla scelta del sistema di coordinate cartesiane) delle isometrie di uno spazio affine euclideo.

Movimenti rigidi

In questa sottosezione $\mathbb{S}$ è uno spazio affine euclideo di dimensione finita. Per il Corollario 7.3.7 abbiamo un omomorfismo *suriettivo* di gruppi

$$\begin{array}{ccc} \mathrm{Isom}(\mathbb{S}) & \xrightarrow{\mathsf{V}} & \mathrm{O}(\mathsf{V}(\mathbb{S})) \\ \Phi & \mapsto & \mathsf{V}(\Phi) \end{array} \tag{7.4.3}$$

Ricordiamo che $\mathrm{SO}(\mathsf{V}(\mathbb{S})) \subset \mathrm{O}(\mathsf{V}(\mathbb{S}))$ è il sottogruppo delle isometrie che hanno determinante 1. Poniamo

$$\mathrm{Isom}^+(\mathbb{S}) := \{\Phi \in \mathrm{Isom}(\mathbb{S}) \mid \mathsf{V}(\Phi) \in \mathrm{SO}(\mathsf{V}(\mathbb{S}))\}. \tag{7.4.4}$$

Gli elementi di $\mathrm{Isom}^+(\mathbb{S})$ si chiamano *movimenti rigidi*, la spiegazione della terminologia è nell'Esempio 7.4.13 e nella Proposizione 7.4.14.

Notiamo che $\mathrm{Isom}^+(\mathbb{S})$ è un sottogruppo di $\mathrm{Isom}(\mathbb{S})$. Infatti $\mathsf{V}(\mathrm{Id}_{\mathbb{S}}) = \mathrm{Id}_{\mathsf{V}(\mathbb{S})}$ e quindi l'identità è un elemento di $\mathrm{Isom}^+(\mathbb{S})$, e se Φ, Ψ sono elementi di $\mathrm{Isom}^+(\mathbb{S})$ allora $\mathsf{V}(\Phi \circ \Psi) = \mathsf{V}(\Phi) \circ \mathsf{V}(\Psi)$ e quindi è un elemento $\mathrm{Isom}^+(\mathbb{S})$, e anche Φ^{-1} lo è perchè $\mathsf{V}(\Phi^{-1}) = \mathsf{V}(\Phi)^{-1}$.

Esempio 7.4.9. Sia $\mathbb{E}^n(\mathbb{R})$ lo spazio affine euclideo standard. Per l'Esempio 7.4.6 gli elementi di $\mathrm{Isom}^+(\mathbb{E}^n(\mathbb{R}))$ sono dati da

$$\begin{array}{ccc} \mathbb{E}^n(\mathbb{R}) & \longrightarrow & \mathbb{E}^n(\mathbb{R}) \\ X & \mapsto & A \cdot X + B \end{array}$$

dove $A \in \mathrm{SO}_n(\mathbb{R})$ e $B \in M_{n,1}(\mathbb{R})$ è una matrice colonna.

Esiste una nozione di deformabilità di isometrie affini di uno spazio vettoriale euclideo del tutto analoga a quella di deformabilità di isometrie di spazi vettoriali euclidei. Sia $I \subset \mathbb{R}$ un intervallo, e sia $\gamma\colon I \to \mathrm{Isom}(\mathbb{S})$ un'applicazione. Scegliamo un riferimento ortonormale $RO(O, \mathcal{B})$. Per $t \in I$, siano $A(t) \in \mathrm{O}_n(\mathbb{R})$ (qui $n := \dim \mathbb{S}$) e $B(t) \in M_{n,1}(\mathbb{R})$ le matrici tali che, nelle coordinate del riferimento scelto l'isometria $\gamma(t)$ si data (vedi l'Esempio 7.4.8) da

$$X \mapsto A(t) \cdot X + B(t).$$

Definizione 7.4.10. L'applicazione $\gamma\colon I \to \mathrm{Isom}(\mathbb{S})$ è *continua* se, scelto un qualsiasi riferimento ortonormale $RO(O,\mathcal{B})$, l'applicazione $I \to \mathrm{O}_n(\mathbb{R})$ che associa a $t \in I$ la matrice $A(t)$ è continua, e anche l'applicazione $I \to M_{n,1}(\mathbb{R})$ che associa a $t \in I$ la matrice $B(t)$.

Osservazione 7.4.11. Per assicurarsi che $\gamma\colon I \to \mathrm{Isom}(\mathbb{S})$ sia continua è sufficiente verificare che le applicazioni date da $A(t)$ e $B(t)$ siano continue per *un* riferimento ortonormale $RO(O,\mathcal{B})$. Infatti siano Y sono le coordinate di un altro riferimento ortonormale. Allora essitono $M \in \mathrm{O}_n(\mathbb{R})$ e $C \in M_{n,1}(\mathbb{R})$ tali che la relazione tra le coordinate X e Y di uno stesso punto sia $Y = M \cdot X + C$. Quindi nelle coordinate Y l'applicazione $\gamma(t)$ è data da

$$Y \mapsto M \cdot A(t) \cdot M^{-1} \cdot Y - M \cdot A(t) \cdot M^{-1} \cdot C + M \cdot B(t) + C.$$

Segue che se le applicazioni $t \mapsto A(t)$ e $t \mapsto B(t)$ sono continue allora lo sono anche quelle associate alle nuove coordinate, perchè sono date da $t \mapsto M \cdot A(t) \cdot M^{-1}$ e $t \mapsto -M \cdot A(t) \cdot M^{-1} \cdot C + M \cdot B(t) + C$.

Definizione 7.4.12. Una isometria di $\Phi \in \mathrm{Isom}(\mathbb{S})$ è *deformabile in* una isometria $\Psi \in \mathrm{Isom}(\mathbb{S})$ se esiste un'applicazione continua $\gamma\colon [a,b] \to \mathrm{Isom}(\mathbb{S})$ tale che $\gamma(a) = \Phi$ e $\gamma(b) = \Psi$.

Esempio 7.4.13. Sia Φ un movimento rigido di $\mathbb{E}^n(\mathbb{R})$. Quindi $\Phi(X) = A \cdot X + B$, dove $A \in \mathrm{SO}_n(\mathbb{R})$ e $B \in M_{n,1}(\mathbb{R})$. Per la Proposizione 6.10.10 esistono un intervallo $[a,b] \subset \mathbb{R}$ e un'applicazione continua $\gamma\colon [a,b] \to \mathrm{SO}_n(\mathbb{R})$ tali che $\gamma(a) = 1_n$ e $\gamma(b) = A$. Possiamo supporre che $a < b$ perchè se $a = b$ allora $a = 1_n$, e quindi possiamo sostituire γ con un'applicazione costante. Definiamo $\mu\colon [a,b] \to \mathrm{Isom}(\mathbb{E}_n(\mathbb{R}))$ ponendo

$$\mu(t)(X) = \gamma(t) \cdot X + \frac{t-a}{b-a} B.$$

L'applicazione μ è costante e $\mu(a) = 1_n$, $\mu(b) = \Phi$. Quindi l'identità di $\mathbb{E}_n(\mathbb{R})$ è deformabile in Φ.

Come nel caso di uno spazio vettoriale reale o di uno spazio vettoriale euclideo (finitamente generati), la relazione di deformabilità è di equivalenza.

Proposizione 7.4.14. *Sia $\mathbb{S}$ uno spazio affine euclideo di dimensione finita. Due isometrie di $\mathbb{S}$ sono deformabile l'una nell'altra se e solo se appartengono entrambi a $\mathrm{Isom}^+(\mathbb{S})$ oppure entrambi a $\mathrm{Isom}(\mathbb{S}) \setminus \mathrm{Isom}^+(\mathbb{S})$.*

Dimostrazione. Se $I \subset \mathbb{R}$ è un intervallo e $\gamma\colon I \to \mathrm{Isom}(\mathbb{S})$ è un'applicazione continua, allora $\mathrm{Det}\,\mathsf{V}(\Phi(t))$ è una funzione continua di $t \in I$. Siccome $\mathrm{Det}\,\mathsf{V}(\Phi(t)) \in \{+1,-1\}$ e I è un intervallo segue che $\mathrm{Det}\,\mathsf{V}(\Phi(t))$ è costante. Quindi se due isometrie di $\mathbb{S}$ sono deformabili l'una nell'altra, allora sono entrambe in $\mathrm{Isom}^+(\mathbb{S})$ oppure entrambe in $\mathrm{Isom}(\mathbb{S}) \setminus \mathrm{Isom}^+(\mathbb{S})$.

Ora dimostriamo che se due isometrie sono entrambe in $\mathrm{Isom}^+(\mathbb{S})$ allora sono deformabili l'una nell'altra. Siccome la relazione che stiamo considerando è di equivalenza, e siccome $\mathrm{Id}_{\mathbb{S}} \in \mathrm{Isom}^+(\mathbb{S})$, è sufficiente dimostrare che ogni elemento di $\mathrm{Isom}^+(\mathbb{S})$ è deformabile in $\mathrm{Id}_{\mathbb{S}}$. Sia n la dimensione di $\mathbb{S}$. Scegliendo un riferimento ortonormale $RO(O,\mathcal{B})$ di $\mathbb{S}$ definiamo un'isometria $\Phi\colon \mathbb{S} \longrightarrow \mathbb{E}_n(\mathbb{S})$, vedi la Sottosezione 7.4. L'isometria Φ definisce l'isomorfismo $\mathrm{Isom}(\mathbb{S}) \xrightarrow{\sim} \mathrm{Isom}(\mathbb{E}_n(\mathbb{R}))$ definito in (7.4.2), e questo isomorfismo manda $\mathrm{Isom}^+(\mathbb{S})$ in $\mathrm{Isom}^+(\mathbb{E}_n(\mathbb{R}))$. Abbiamo visto nell'Esempio 7.4.13 a che ogni movimento rigido di $\mathbb{E}_n(\mathbb{R})$ è deformabile nell'identità di $\mathbb{E}_n(\mathbb{R})$ e quindi segue che ogni movimento rigido di $\mathbb{S}$ è deformabile nell'identità di $\mathbb{S}$.

Infine se $\Phi, \Psi \in (\mathrm{Isom}(\mathbb{S}) \setminus \mathrm{Isom}^+(\mathbb{S}))$ si dimostra che Φ si deforma in Ψ procedendo esattamente come nella dimostrazione della Proposizione 5.11.8 o della Proposizione 6.10.10. □

Dimensione 1

Esaminamo le isometrie di uno spazio affine euclideo $\mathbb{S}$ di dimensione 1.

È stato spiegato nella Sottosezione 7.4 che $\mathrm{Isom}(\mathbb{S})$ dipende solo dalla dimensione di $\mathbb{S}$, e quindi possiamo assumere che $\mathbb{S} = \mathbb{E}^1(\mathbb{R})$. Le isometrie sono le applicazioni affini date da

$$\begin{array}{ccc} \mathbb{E}^1(\mathbb{R}) & \xrightarrow{f} & \mathbb{E}^1(\mathbb{R}) \\ x & \mapsto & ax+b \end{array}$$

dove $a^2 = 1$. Se $a = 1$ allora f è una traslazione, non c'è altro da aggiungere. Se $a = -1$, allora $f(x) = -x + b$. Sia $RO(O,\mathcal{B})$ il riferimento ON con coordinata $y = x - \frac{b}{2}$, cioè $O = P(\frac{b}{2})$ e $\mathcal{B} = \{1\}$. In queste coordinate f è data da

$$y \mapsto -y,$$

e quindi vediamo che è la simmetria (o ribaltamento) nel punto di coordinata $x = \frac{b}{2}$. In conclusione $\mathrm{Isom}^+(\mathbb{S})$ consiste del sottogruppo delle traslazioni, $(\mathrm{Isom}(\mathbb{S}) \setminus \mathrm{Isom}^+(\mathbb{S}))$ consiste delle simmetrie di centro un punto di $\mathbb{S}$.

Dimensione 2

Esaminamo le isometrie di uno spazio affine euclideo $\mathbb{S}$ di dimensione 2.

Definizione 7.4.15. Una $\Phi \in \mathrm{Isom}(\mathbb{S})$ è una *rotazione* se esistono $P \in \mathbb{S}$ e $\theta \in \mathbb{R}$ tali che, nelle coordinate X di un riferimento ON $RO(P, \mathcal{B})$, sia data da

$$\Phi(X) = R_\theta \cdot X, \tag{7.4.5}$$

dove

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Osserviamo che se $\theta \in 2\pi\mathbb{Z}$ allora l'applicazione Φ in (7.4.5) è l'identità $\mathrm{Id}_{\mathbb{S}}$, e che se invece $\theta \notin 2\pi\mathbb{Z}$ allora Φ ha un unico punto fisso, cioè P. In quest'ultimo caso si dice che P è *il* centro di rotazione di Φ. Una *rotazione di centro* P è una rotazione che fissa P, cioè o una rotazione (non uguale all'identità) tale che P sia il suo centro di rotazione oppure l'identità.

Proposizione 7.4.16. *Sia $\mathbb{S}$ uno spazio affine euclideo di dimensione 2. I movimento rigidi di $\mathbb{S}$ sono le traslazioni e le rotazioni, cioè $\Phi \in \mathrm{Isom}^+(\mathbb{S})$ se e solo se Φ è una traslazione o una rotazione.*

Dimostrazione. Se Φ è una traslazione di $\mathbb{S}$ allora $\mathsf{V}(\Phi) = \mathrm{Id}_{\mathsf{V}(\mathbb{S})}$, e quindi Φ è un movimento rigido. Se Φ è una rotazione di $\mathbb{S}$ allora in una base ON opportuna la matrice associata a $\mathsf{V}(\Phi)$ è R_θ, e quindi $\Phi \in \mathrm{Isom}^+(\mathbb{S})$ perchè $\mathrm{Det}(R_\theta) = 1$.

Dimostriamo il viceversa. Possiamo assumere che $\mathbb{S} = \mathbb{E}^2(\mathbb{R})$. Allora $\Phi(X) = A \cdot X + B$ dove $\mathrm{Det}\, A = 1$. Per Corollario 6.10.4 abbiamo che $A = 1_2$ oppure $A = R_\theta$ dove $\theta \notin 2\pi\mathbb{Z}$. Nel primo caso Φ è una traslazione. Dimostriamo che nel secondo caso esiste un punto fisso P di Φ. Infatti, siccome 1 non è un autovalore di R_θ esiste una soluzione dell'equazione $A \cdot X + B = X$, cioè dell'equazione $(A - 1_2) \cdot X = -B$ perchè $\mathrm{Det}(A - 1_2) \neq 0$. Se Y sono le coordinate in un sistema di riferimento ON con origine nel punto fisso P, allora la Φ è data da $Y \mapsto C \cdot Y$, dove $C \in \mathrm{SO}(2)$. Per il Corollario 6.10.4 esiste $\alpha \in \mathbb{R}$ tale che $C = R_\alpha$ (di fatto $\alpha = \pm\theta$ a meno di multipli interi di 2π), e quindi Φ è una rotazione. □

Ora passiamo a descrivere le isometrie che non sono movimenti rigidi.

Definizione 7.4.17. Una $\Phi \in \mathrm{Isom}(\mathbb{S})$ è una *glissoriflessione* se è la composizione di una riflessione ortogonale in una retta $\mathbb{L} \subset \mathbb{S}$ (vedi l'Esempio 7.3.8) e una traslazione τ_v con $v \in \mathsf{V}(\mathbb{L})$.

Osservazione 7.4.18. Sia $\Phi \in \mathrm{Isom}(\mathbb{S})$ una glissoriflessione come nella Definizione 7.4.17. Allora $\Phi(\mathbb{L}) = \mathbb{L}$ perchè la riflessione in $\mathbb{L}$ fissa ogni punto di $\mathbb{L}$ e τ_v manda ogni punto $P \in \mathbb{L}$ nel punto $(P + v) \in \mathbb{L}$ (perchè $v \in \mathsf{V}(\mathbb{L})$). Sia $RO(O, \{v_1, v_2\})$ un riferimento ON con $O \in \mathbb{L}$ e $v_1 \in \mathsf{V}(\mathbb{L})$. Allora $v = av_1$ per un certo $a \in \mathbb{R}$ e $\mathbb{L}$ ha equazione cartesiana $x_2 = 0$. Nelle coordinate (x_1, x_2) l'isometria Φ è data da

$$\begin{array}{ccc} \mathbb{S} & \xrightarrow{\Phi} & \mathbb{S} \\ P(x_1, x_2) & \mapsto & Q(x_1 + a, -x_2) \end{array}$$

Da questa formula si vede che $\Phi = \tau_v \circ R_{\mathbb{L}} = R_{\mathbb{L}} \circ \tau_v$ e che una glissoriflessione non è un movimento rigido. Inoltre vediamo che $\mathbb{L}$ è l'unica retta di $\mathbb{S}$ mandata in sè stessa da Φ e tale che la restrizione alla retta sia una traslazione; $\mathbb{L}$ è l'*asse della glissoriflessione.*

Con la nostra definizione una riflessione in una retta è una glissoriflessione (con $v = 0$). Spesso si fa distinzione tra riflessioni e glissoriflessioni con $v \neq 0$, e si chiamano glissoriflessioni solo queste ultime.

Proposizione 7.4.19. *Sia $\mathbb{S}$ uno spazio affine euclideo di dimensione 2. Una isometria di $\mathbb{S}$ è in $\mathrm{Isom}(\mathbb{S}) \setminus \mathrm{Isom}^+(\mathbb{S})$, cioè non è un movimento rigido, se e solo se è una glissoriflessione.*

Dimostrazione. Abbiamo visto che le glissoriflessioni sono elementi di $\mathrm{Isom}(\mathbb{S}) \setminus \mathrm{Isom}^+(\mathbb{S})$, quindi si tratta di dimostrare che se che se $\Phi \in (\mathrm{Isom}(\mathbb{S}) \setminus \mathrm{Isom}^+(\mathbb{S}))$ allora Φ è una glissometria. Iniziamo osservando che se esiste una retta $\mathbb{L}$ tale che $\Phi(\mathbb{L}) = \mathbb{L}$, allora Φ è una glissoriflessione. Infatti siccome la restrizione di Φ a $\mathbb{L}$ è una isometria sappiamo che

1. $\Phi_{|\mathbb{L}} \colon \mathbb{L} \to \mathbb{L}$ è una traslazione, oppure
2. $\Phi_{|\mathbb{L}} \colon \mathbb{L} \to \mathbb{L}$ è la riflessione in un punto $P \in \mathbb{L}$.

Supponiamo che valga (1). Quindi $\Phi_{|\mathbb{L}}$ è la traslazione τ_v dove $v \in \mathsf{V}(\mathbb{L})$. Sia $\Psi \in \mathrm{Isom}(\mathbb{S})$ la composizione

$$\mathbb{S} \xrightarrow{\Phi} \mathbb{S} \xrightarrow{\tau_{-v}} \mathbb{S}.$$

Allora Ψ è l'identità su $\mathbb{L}$ e $\det \mathsf{V}\Psi = -1$. Segue facilmente che Ψ è la riflessione in $\mathbb{L}$ - scegliete un riferimento ON $RO(O, \{v_1, v_2\}$ con $O \in \mathbb{L}$ e $v_1 \in T(\mathbb{L})$. Quindi $\Psi = \tau_{-v} \circ \Phi$ è una riflessione, e perciò $\Phi = \tau_{-v}^{-1} \circ \Psi = \tau_v \circ \Psi$ è una glissoriflessione. Supponiamo che valga (2). Allora $\Phi(P) = P$. Scegliendo un riferimento ON $RO(P, \mathcal{B})$ vediamo che $\Phi(X) = A \cdot X$ dove $A \in (\mathrm{O}(2) \setminus \mathrm{SO}(2))$. Quindi (vedi la Proposizione 6.10.3) A ha due autovettori ortogonali con autovalori $+1$ e -1 rispettivamente, e perciò Φ è una riflessione. La conclusione di quello che abbiamo dimostrato finora è che per dimostrare la proposizione basta far vedere che esiste una retta $\mathbb{L} \subset \mathbb{S}$ tale che $\Phi(\mathbb{L}) = \mathbb{L}$.

Siccome $\mathsf{V}(\Phi) \in (\mathrm{O}(2) \setminus \mathrm{SO}(2))$ sappiamo che $\mathsf{V}(\Phi)$ ha due autovettori ortogonali v e w con autovalori $+1$ e -1 rispettivamente. Quindi se $\mathbb{L} \subset \mathbb{S}$ è una retta tale che $\mathsf{V}(\mathbb{L}) = \mathrm{Span}(v)$ oppure $\mathsf{V}(\mathbb{L}) = \mathrm{Span}(w)$, allora $\Phi(\mathbb{L})$ è una retta parallela a $\mathbb{L}$. Scegliamo una retta $\mathbb{L}$ tale che $\mathsf{V}(\mathbb{L}) = \mathrm{Span}(w)$. Definiamo l'applicazione $F \colon \mathbb{L} \to \mathbb{L}$ come segue. Dato $P \in \mathbb{L}$ sia $T_P \subset \mathbb{S}$ la retta ortogonale a $\mathbb{L}$ contenente P. Siccome $\mathsf{V}(T_P)$ è generato da w l'immagine $\Phi(R_P)$ è una retta parallela a R_P, e che perciò incontra $\mathbb{L}$ in un (unico) punto Q: poniamo $F(P) = Q$. È facile verificare che F è una isometria di $\mathbb{L}$. Inoltre $\mathsf{V}(F) = -\mathrm{Id}$ perchè $\mathsf{V}(\Phi)$ è (-1) su $\mathsf{V}(\mathbb{L})$. Quindi F è la riflessione in un punto $Q \in \mathbb{L}$. Questo non significa che $\Phi(R_Q) = R_Q$ e quindi Φ è una glissoriflessione per quanto dimostrato sopra. □

Dimensione 3

Esaminiamo i movimenti rigidi di uno spazio affine euclideo $\mathbb{S}$ di dimensione 3, cioè gli elementi di $\mathrm{Isom}^+(\mathbb{S})$.

Definizione 7.4.20. Sia $\mathbb{T} \subset \mathbb{S}$ un piano. Un'isometria $\Psi \colon \mathbb{S} \to \mathbb{S}$ è una *rotazione del piano* $\mathbb{T}$ se in un opportuno sistema di riferimento ON $RO(O, \mathcal{B})$ di coordinate (x, y, z) si ha che $\mathbb{T}$ ha equazione $z = 0$ e

$$\Psi(x, y, z) = (\cos\theta x - \sin\theta y, \sin\theta x \cos\theta y, z).$$

In altre parole Ψ manda $\mathbb{T}$ in se stesso, la restrizione di Ψ a $\mathbb{T}$ è una rotazione, e la restrizione di $\mathsf{V}(\Psi)$ all'ortogonale di $\mathsf{V}(\mathbb{T})$ è l'identità. Nella base $\mathcal{B}$ la matrice associata a $\mathsf{V}(\Psi)$ è

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{7.4.6}$$

e quindi Ψ è un movimento rigido. Notate che l'identità è una rotazione di $\mathbb{T}$ per qualsiasi piano $\mathbb{T}$. Invece una rotazione Ψ del piano $\mathbb{T}$ che non sia l'identità è una rotazione di qualsiasi piano parallelo a $\mathbb{T}$, ma di nessun altro piano. Inoltre la restrizione di Ψ a un piano parallelo a $\mathbb{T}$ ha un unico punto fisso e la retta $\mathbb{L}$ perpendicolare a $\mathbb{T}$ per il punto fisso è fissata da Ψ (cioè $\Psi(P) = P$ per ogni $P \in \mathbb{L}$).

Definizione 7.4.21. Una $\Phi \in \mathrm{Isom}(\mathbb{S})$ è un *movimento elicoidale* se esistono un piano $\mathbb{T} \subset \mathbb{S}$ e $v \in \mathsf{V}(\mathbb{T}^\perp)$ tali che $\Phi(X) = \tau_v \circ \Psi$, dove Ψ è una rotazione del piano $\mathbb{T}$.

Siccome traslazioni e rotazioni di piani sono movimenti rigidi, un movimento elicoidale è un movimento rigido. Notate che sia le traslazioni che le rotazioni di piani, secondo la nostra definizione, sono movimenti elicoidali, benchè sarebbe più appropriato parlare di movimenti elicoidali degeneri.

Osservazione 7.4.22. Sia $\Phi \in \mathrm{Isom}(\mathbb{S})$ un movimento elicoidale, dato da $\Phi(X) = \tau_v \circ \Psi$, dove Ψ è una rotazione del piano $\mathbb{T}$ e $v \in \mathsf{V}(\mathbb{T}^\perp)$. Allora si ha anche $\Phi(X) = \Psi \circ \tau_v$. Questo si può vedere scrivendo $\Psi(X)$ nelle coordinate (x, y, z) della Definizione 7.4.20, vedi la formula in (7.4.6): si ha $\tau_v(x, y, z) = (x, y, z + a)$ per un opportuno $a \in \mathbb{R}$, e segue immediatamente che $\tau_v \circ \Psi = \Psi \circ \tau_v$.

Teorema 7.4.23 (Chasles)**.** *Sia $\mathbb{S}$ uno spazio affine euclideo di dimensione 3. I movimento rigidi di $\mathbb{S}$ sono i movimenti elicoidali.*

Dimostrazione. Abbiamo già notato che i movimenti elicoidali sono movimenti rigidi, quindi si tratta di dimostrare che se Φ è un movimento rigido, allora è un movimento elicoidale. Se $\mathsf{V}(\Phi)$ è l'identità allora Φ è una traslazione, che è un movimento elicoidale. Se $\mathsf{V}(\Phi)$ non è l'identità, allora per il Corollario 6.10.4 esiste una decomposizione in somme diretta ortogonale

$$\mathsf{V}(\mathbb{S}) = U \oplus_\perp W \tag{7.4.7}$$

tale che $\mathsf{V}(\Phi)(v) = v$ per ogni $v \in U$, $\Phi(W) = W$, e la restrizione di Φ a W sia una rotazione diversa dall'identità. Sia $\mathbb{T} \subset \mathbb{S}$ un piano tale che $\mathsf{V}(\mathbb{T}) = W$. Definiamo un'applicazione $\Psi\colon \mathbb{T} \to \mathbb{T}$ come segue. Dato $P \in \mathbb{T}$ sia $\mathbb{L}_P$ la retta per P perpendicolare a $\mathbb{T}$, e quindi di giacitura U perchè la decomposizione in (7.4.7) è ortogonale. Siccome $\Phi(U) = U$ la retta immagine $\Phi(\mathbb{L}_P)$ è una retta perpendicolare a $\mathbb{T}$, in particolare esiste un unico punto di intersezione tra $\Phi(\mathbb{L}_P)$ e $\mathbb{T}$: questo è il punto $\Psi(P)$. Si verifica senza problemi che $\Psi\colon \mathbb{T} \to \mathbb{T}$ è un'isometria con $\mathsf{V}(\Psi)$ uguale alla restrizione di $\mathsf{V}(\Phi)$ a W, e quindi è una rotazione che non è l'identità. Quindi esiste un unico punto fisso dell'isometria Ψ, denotiamolo P_0. Per come è definita l'applicazione Ψ segue che la retta $\mathbb{L}_{P_0}$ per P_0 ortogonale a $\mathbb{T}$ è mandata in sè da Φ, e siccome $\mathsf{V}(\mathbb{L}_{P_0}) = U$ la restrizione di Φ a $\mathbb{L}_{P_0}$ è la traslazione di vettore un certo $v \in U$. Il movimento rigido $\Upsilon := \tau_{-v} \circ \Phi$ fissa ogni punto di $\mathbb{L}_{P_0}$, quindi manda $\mathbb{T}$ in $\mathbb{T}$. Inoltre

$$\mathsf{V}(\Upsilon) = \mathsf{V}(\tau_{-v} \circ \Phi) = \mathsf{V}(\tau_{-v}) \circ \mathsf{V}(\Phi) = \mathsf{V}(\Phi),$$

e quindi Υ è una rotazione del piano $\mathbb{T}$. Moltiplicando a sinistra per τ_v ambo i membri dell'uguaglianza $\Upsilon := \tau_{-v} \circ \Phi$ otteniamo che $\Phi = \tau_v \circ \Upsilon$ e quindi Φ è un movimento elicoidale. □

Esercizi del Capitolo 7

Esercizio 7.1. *Sia $(\mathbb{R}^n, +)$ il gruppo $\mathbb{R}^n$ con operazione la somma di vettori. Definiamo un'applicazione*

$$\mathrm{Isom}(\mathbb{E}^n(\mathbb{R})) \xrightarrow{T} (\mathbb{R}^n, +),$$

come segue. Ogni $\Phi \in \mathrm{Isom}(\mathbb{E}^n(\mathbb{R}))$ è data da

$$\begin{array}{ccc} \mathbb{E}^n(\mathbb{R}) & \xrightarrow{\Phi} & \mathbb{E}^n(\mathbb{R}) \\ X & \mapsto & A \cdot X + B \end{array}$$

Poniamo $T(\Phi) := B$. Dimostrate che T non *è un omomorfismo di gruppi (a differenza dell'applicazione $\mathrm{Isom}(\mathbb{E}^n(\mathbb{R})) \to \mathrm{O}_n(\mathbb{R})$ che manda Φ in $\mathsf{V}(\Phi)$.)*

Esercizio 7.2. *Siano $A \in M_{2,2}(\mathbb{R})$ e $B \in M_{2,1}(\mathbb{R})$ dati da*

$$A := \begin{pmatrix} 119/169 & -120/169 \\ -120/169 & -119/169 \end{pmatrix}, \quad B := \begin{pmatrix} 3 \\ 2 \end{pmatrix},$$

Definiamo

$$\begin{array}{ccc} \mathbb{E}^2 & \xrightarrow{\Phi} & \mathbb{E}^2 \\ X & \mapsto & A \cdot X + B \end{array}$$

(a) *Verificate che $\Phi \in \mathrm{Isom}^-(\mathbb{E}^2)$, e quindi Φ è una glissometria.*

(b) *Determinate l'asse della glissometria Φ.*

Esercizio 7.3. *Sia $\mathbb{A}$ un piano affine euclideo. Siano Φ_1, Φ_2 glissoriflessioni di assi $\mathbb{S}_1, \mathbb{S}_2$* ortogonali*, e vettori di traslazione v_1, v_2 rispettivamente.*

(a) *Dimostrate che $\Psi := \Phi_1 \circ \Phi_2$ è una rotazione di angolo π.*

(b) *Come si determina geometricamente il centro di Ψ ? In particolare dimostrate che se $||v_1|| = ||v_2||$ allora il centro di Ψ appartiene a una delle due bisettrici di $\mathbb{S}_1, \mathbb{S}_2$.*

Esercizio 7.4. *Sia $G \subset \mathrm{Isom}(\mathbb{E}^2(\mathbb{R}))$ il sottogruppo delle simmetrie di $\mathbb{Z}^2$, cioè*

$$G := \{\Phi \in \mathrm{Isom}(\mathbb{E}^2(\mathbb{R})) \mid \Phi(\mathbb{Z}^2) = \mathbb{Z}^2\}.$$

Descrivete gli assi delle glissometrie in G.

Capitolo 8

Forme quadratiche e forme bilineari simmetriche

8.1 Funzioni polinomiali omogenee su uno spazio vettoriale

In questo capitolo gli spazi vettoriali sono *finitamente generati*. Sia V uno spazio vettoriale su $\mathbb{K}$. Daremo senso alla nozione di funzione polinomiale omogenea $V \to \mathbb{K}$. Iniziamo con un'osservazione. Sia $P\colon V \to \mathbb{K}$ una funzione. Supponiamo che esista una base $\{v_1, \ldots, v_n\}$ di V tale che la funzione $f\colon \mathbb{K}^n \to \mathbb{K}$ definita da

$$f(x_1, \ldots, x_n) := P(x_1 v_1 + \ldots + x_n v_n) \tag{8.1.1}$$

sia polinomiale omogenea di grado d. Ciò significa (vedi la Sottosezione 1.7) che esistono $f_{d_1,\ldots,d_n} \in \mathbb{K}$ per ogni multiindice $(d_1, \ldots, d_n)$ con $d_1 + \cdots + d_n = d$ tali che

$$f(x_1, \ldots, x_n) = \sum_{d_1+\cdots+d_n=d} f_{d_1,\ldots,d_n} x_1^{d_1} x_2^{d_2} \cdots x_n^{d_n}, \tag{8.1.2}$$

per ogni $(x_1, \ldots, x_n) \in \mathbb{K}^n$. Allora se $\{w_1, \ldots, w_n\}$ è una qualsiasi altra base di V anche la funzione $r\colon \mathbb{K}^n \to \mathbb{K}$ definita da

$$r(y_1, \ldots, y_n) = P(y_1 w_1 + \ldots + y_n w_n)$$

è polinomiale omogenea di grado d. Infatti sia $A = (a_{ij}) \in M_{nn}(\mathbb{K})$ la matrice (invertibile) tale che $w_i = \sum_{j=1}^n a_{ij} v_j$ per ogni $1 \leqslant i \leqslant m$. Sostituendo questa espressione nella (8.1.1) otteniamo che

$$P(y_1 w_1 + \ldots + y_m w_m) = P\left(\left(\sum_{i=1}^n a_{i1} y_i\right) v_1 + \ldots + \left(\sum_{i=1}^n a_{in} y_j\right) v_n\right) = p\left(\left(\sum_{i=1}^n a_{i1} y_i\right), \ldots, \left(\sum_{i=1}^n a_{in} y_i\right)\right),$$

ed espandendo nelle $y_1, \ldots, y_n$ i prodotti che appaiono nell'ultima espressione vediamo che r è un polinomio omogeneo di grado d nelle $y_1, \ldots, y_n$. Per l'osservazione appena fatta ha senso porre la seguente definizione.

Definizione 8.1.1. Sia V uno spazio vettoriale su $\mathbb{K}$. Una funzione $P\colon V \to \mathbb{K}$ è *polinomiale omogenea di grado d* se, data una base $\{v_1, \ldots, v_n\}$ di V, la funzione $f\colon \mathbb{K}^n \to \mathbb{K}$ definita da (8.1.1) è un polinomio omogeneo di grado d.

Esempio 8.1.2. Una funzione polinomiale omogenea di grado 1 su V non è altro che una forma lineare su V.

Esempio 8.1.3. Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo (in particolare reale) finitamente generato. La funzione

$$\begin{array}{ccc} V & \xrightarrow{N} & \mathbb{R} \\ v & \mapsto & ||v||^2 \end{array}$$

è polinomiale omogenea di grado 2. Infatti se $\{v_1, \ldots, v_n\}$ è una base di V allora

$$N(x_1 v_1 + \ldots + x_n v_n) = \sum_{1 \leqslant i,j \leqslant n} \langle v_i, v_j \rangle x_i x_j.$$

Osservazione 8.1.4. Sia $P\colon V \to \mathbb{K}$ una funzione polinomiale omogenea di grado d. Se $t \in \mathbb{K}$ e $v \in V$, allora

$$P(tv) = t^d P(v). \tag{8.1.3}$$

Infatti sia $p\colon \mathbb{K}^n \to \mathbb{K}$ la funzione definita da (8.1.1). Siccome p è un polinomio omogeneo di grado d vale $p(tx_1, \ldots, tx_n) = t^d p(x_1, \ldots, x_n)$. Ponendo $v = x_1v_1 + \ldots + x_nv_n$ segue che $P(tv) = t^d P(v)$. L'espressione "funzione omogenea di grado d" significa esattamente che vale l'uguaglianza in (8.1.3). Infine notiamo che se $d \neq 0$ allora $P(0) = 0$.

Definizione 8.1.5. Se V è uno spazio vettoriale su $\mathbb{K}$ finitamente generato denotiamo $\mathcal{P}_d(V)$ l'insieme delle funzioni polinomiali omogenee di grado d da V in $\mathbb{K}$.

Notiamo che $\mathcal{P}_d(V)$ è un sottospazio dello spazio vettoriale su $\mathbb{K}$ delle funzioni da V in $\mathbb{K}$ perchè l'insieme dei polinomi omogenei di grado d è uno spazio vettoriale su $\mathbb{K}$. In particolare $\mathcal{P}_d(V)$ è uno spazio vettoriale su $\mathbb{K}$. Ovviamente l'elemento neutro è la funzione costante 0. Se $a \in \mathbb{K}$ denotiamo con $a\colon V \to \mathbb{K}$ la funzione costante di valore a (è polinomiale omogenea di grado 0).

Ora definiamo un'azione di $\mathrm{GL}(V)$ su $\mathcal{P}_d(V)$ e consideriamo l'associata relazione di equivalenza. Sia $P \in \mathcal{P}_d(V)$, cioè $P\colon V \to \mathbb{K}$ è una funzione polinomiale omogenee di grado d, e sia $g \in \mathrm{GL}(V)$. La funzione

$$\begin{array}{ccc} V & \xrightarrow{Q} & \mathbb{K} \\ v & \mapsto & P(g^{-1}(v)) \end{array} \tag{8.1.4}$$

è polinomiale omogenea di grado d. Infatti sia $\{v_1, \ldots, v_n\}$ una base di V, e sia $\{w_1, \ldots, w_n\}$ la base di V ottenuta ponendo $w_i := g(v_i)$. Allora

$$Q(x_1w_1 + \ldots + x_nw_n) = P(g^{-1}(x_1w_1 + \ldots + x_nw_n)) = P(x_1v_1 + \ldots + x_nv_n). \tag{8.1.5}$$

Per ipotesi l'applicazione $\mathbb{K}^n \to \mathbb{K}$ data da $(x_1, \ldots, x_n) \mapsto P(x_1v_1 + \ldots + x_nv_n)$ è polinomiale omogenea di grado d, quindi l'uguaglianza appena scritta mostra che Q è una funzione polinomiale omogenea di grado d.

Definizione 8.1.6. Siano $P \in \mathcal{P}_d(V)$ e $g \in \mathrm{GL}(V)$. La funzione polinomiale omogenea di grado d definita in (8.1.4) si denota gP, cioè $gP(v) = P(g^{-1}(v))$ per ogni $v \in V$.

Sia $P \in \mathcal{P}_d(V)$ una funzione polinomiale omogenee di grado d. Allora $\mathrm{Id}_V P = P$ e se $g, h \in \mathrm{GL}(V)$ si ha $ghP = g(hP)$ perchè

$$ghP(v) = P((gh)^{-1}(v)) = P(h^{-1} \circ g^{-1}(v)) = P(h^{-1}(g^{-1}(v))) = g(hP). \tag{8.1.6}$$

In altre parole l'applicazione

$$\begin{array}{ccc} \mathrm{GL}(V) \times \mathcal{P}_d(V) & \longrightarrow & \mathcal{P}_d(V) \\ (g, P) & \mapsto & gP \end{array} \tag{8.1.7}$$

è un'azione del gruppo $\mathrm{GL}(V)$ sullo spazio vettoriale $\mathcal{P}_d(V)$ (vedi la Definizione 1.10.13). Qui è importante notare che se fissiamo $g \in \mathrm{GL}(V)$ allora l'applicazione $\mathcal{P}_d(V) \to \mathcal{P}_d(V)$ data da $P \mapsto gP$ è lineare.

Osservazione 8.1.7. Sembrerebbe più naturale definire l'azione associando a (g, P) la funzione polinomiale omogenea $v \mapsto P(g(v))$. Con questa definizione non si ha l'uguaglianza $ghP = g(hP)$ ma piuttosto $ghP = h(gP)$.

Ora definiamo una relazione tra funzioni polinomiali omogenee di grado d dichiarando che $P \sim Q$ se esiste $g \in \mathrm{GL}(V)$ tale che $Q = gP$. Siccome l'applicazione in (8.1.7) è un'azione del gruppo $\mathrm{GL}(V)$ la relazione appena definita è di equivalenza. Infatti $P = \mathrm{Id}_V P$ mostra che la relazione è riflessiva, la relazione è simmetrica perchè se $Q = gP$ allora $P = g^{-1}Q$ (agite a sinistra su ambo i termini di $Q = gP$ con g^{-1}), e infine è transitiva per l'uguaglianza in (8.1.6).

Definizione 8.1.8. Funzioni polinomiali omogenee (di grado d) $P, Q \in \mathcal{P}_d(V)$ sono *congruenti* se $P \sim Q$, cioè se esiste $g \in \mathrm{GL}(V)$ tale che $Q = gP$.

È naturale (in molti contesti) porsi il seguente problema: classificare le classi di congruenza di funzioni polinomiali omogenee su V (di dato grado), o come si dice "le funzioni polinomiali omogenee a meno dell'azione di $\mathrm{GL}(V)$".

Esempio 8.1.9. Le classi di congruenza in $\mathcal{P}_1(V)$, cioè $V^\vee$, sono due: $\{0\}$ e $V^\vee\backslash\{0\}$. Infatti è chiaro che $\{0\}$ è la classe di congruenza di 0. D'altra parte mostriamo che se $P, Q \in (V^\vee\backslash\{0\})$ allora $P \sim Q$. Scegliamo basi $\{v_1, \dots, v_{n-1}\}$ e $\{w_1, \dots, w_{n-1}\}$ di $\ker(P)$ e $\ker(Q)$ rispettivamente, ed estendiamole a basi $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_n\}$ di V tali che $P(v_n) = 1 = Q(w_n)$. Sia $g \in \mathrm{GL}(V)$ l'unico elemento tale che $g(v_i) = w_i$ per $i \in \{1, \dots, n\}$. Mostriamo che $gP = Q$. Siccome gP e Q sono applicazione lineari è sufficiente verificare che hanno gli stessi valori sui vettori della base $\{w_1, \dots, w_n\}$. Per $i \in \{1, \dots, n\}$ abbiamo

$$gP(w_i) = P(g^{-1}(w_i)) = P(v_i) = Q(w_i).$$

(L'ultima uguaglianza vale grazie alla nostra scelta di basi $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_n\}$.)

Notiamo la seguente versione della congruenza tra funzioni polinomiali omogenee.

Osservazione 8.1.10. Due funzioni polinomiali omogenee $P, Q \in \mathcal{P}_d(V)$ sono congruenti se e solo esistono basi $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_n\}$ di V tali che le funzioni polinomiali omogenee $f, r\colon \mathbb{K}^n \to \mathbb{K}$ definite da

$$f(x_1, \dots, x_n) := P(x_1v_1 + \dots + x_nv_n), \qquad r(x_1, \dots, x_n) := Q(x_1w_1 + \dots + x_nw_n) \tag{8.1.8}$$

sono uguali. Infatti supponiamo che $Q = gP$ dove $g \in \mathrm{GL}(V)$. Scelta un'arbitraria base $\{v_1, \dots, v_n\}$ di V, poniamo $w_i := g(v_i)$ per $i \in \{1, \dots, n\}$. Siccome $g \in \mathrm{GL}(V)$ la lista $\{w_1, \dots, w_n\}$ è una base di V, e le uguaglianze in (8.1.5) mostrano che valgono le uguaglianze in (8.1.8). Viceversa se valgono le uguaglianze in (8.1.8) allora $Q = gP$ dove $g \in \mathrm{GL}(V)$ è l'automorfismo determinato dall'imporre che $w_i := g(v_i)$ per $i \in \{1, \dots, n\}$.

In altre parole classificare le funzioni polinomiali omogenee a meno di congruenza equivale a classificare i polinomi omogenei a meno di cambiamento di coordinate (attenzione però alla differenza tra polinomi e funzioni polinomiali se la cardinalità di $\mathbb{K}$ è "piccola", vedi la Proposizione 1.7.15).

La classificazione delle classi di congruenza per $d > 1$ non è nota in generale. La difficoltà della classificazione aumenta rapidamente con il crescere del grado d (eccetto nel caso banale in cui $\dim V = 1$), ed è sensibile alla scelta del campo $\mathbb{K}$. Nel presente capitolo studieremo le classi di congruenza in $\mathcal{P}_2(V)$. Anche in questo caso la classificazione può non essere semplice, la complessità del problema dipende dal campo $\mathbb{K}$. Noi daremo la classificazione (elementare) per il campo reale e quello complesso. Già per il campo $\mathbb{Q}$ la classificazione richiederebbe strumenti più sofisticati.

8.2 Forme quadratiche

Nella Sottosezione 1.7 abbiamo discusso la differenza che c'è tra polinomi e funzioni polinomiali se la cardinalità del campo è finita. Ovviamente questa differenza non esiste per funzioni polinomiali di grado 1. In questo capitolo studieremo le funzioni polinomiali omogenee di grado 2. Il risultato che segue mostra che anche per tali funzioni polinomiali omogenee non c'è differenza tra il polinomio e la funzione polinomiale associata.

Proposizione 8.2.1. *Siano $P, Q \in \mathbb{K}[x_1, \dots, x_n]$ polinomi omogenei di grado 2. Le funzioni polinomiali omogenee associate $P, Q\colon \mathbb{K}^n \to \mathbb{K}$ sono uguali se e solo se $P = Q$.*

Dimostrazione. È sufficiente dimostrare che se $P \in \mathbb{K}[x_1, \dots, x_n]$ è un polinomio omogeneo di grado 2 tale che $P(a_1, \dots, a_n) = 0$ per ogni $(a_1, \dots, a_n) \in \mathbb{K}^n$ allora $P = 0$ (vedi le dimostrazioni delle Proposizioni 1.7.11 e 1.7.15). Abbiamo

$$P = \sum_{1 \leqslant i \leqslant j \leqslant n} c_{ij} x_i x_j.$$

Dobbiamo dimostrare che $c_{ij} = 0$ per ogni i, j con $1 \leqslant i \leqslant j \leqslant n$. Se $1 \leqslant i \leqslant n$ abbiamo $c_{i,i} = P(e_i) = 0$ e se $1 \leqslant i < j \leqslant n$ abbiamo $c_{i,j} = P(e_i + e_j) = 0$ (qui e_i è l'i-esimo vettore della base standard di $\mathbb{K}^n$). □

Definizione 8.2.2. Sia V uno spazio vettoriale su $\mathbb{K}$. Una *forma quadratica su V* è una funzione polinomiale omogenea $Q\colon V \to \mathbb{K}$. Poniamo $Q(V) = \mathcal{P}_2(V)$, cioè denotiamo con $Q(V)$ lo spazio vettoriale su $\mathbb{K}$ delle forme quadratiche su V.

Ora procediamo ad associare a una forma quadratica su V un sottospazio di V.

Proposizione 8.2.3. *Sia $Q \in Q(V)$. L'insieme*

$$\{u \in V \mid Q(u + v) = Q(v) \quad \forall v \in V\} \tag{8.2.1}$$

è un sottospazio di V.

Dimostrazione. Supponiamo che per ogni $v \in V$ valga $Q(u_1 + v) = Q(v)$ e $Q(u_2 + v) = Q(v)$. Allora

$$Q((u_1 + u_2) + v) = Q(u_1 + (u_2 + v)) = Q(u_2 + v) = Q(v) \quad \forall v \in V.$$

Quindi $u_1 + u_2$ appartiene all'insieme in (8.2.1). Rimane da verificare che se u appartiene all'insieme in (8.2.1), allora tu vi appartiene per ogni $t \in K$. Se $t = 0$ questo è ovvio, quindi supponiamo $t \neq 0$. Se $v \in V$ abbiamo

$$Q(tu + v) = Q(t(u + t^{-1}v)) = t^2 Q(u + t^{-1}v) = t^2 Q(t^{-1}v) = Q(v),$$

(per la seconda uguaglianza vedi l'Osservazione 8.1.4) e quindi tu appartiene all'insieme in (8.2.1). □

Definizione 8.2.4. Sia $Q \in Q(V)$. Il *nucleo (o radicale)* di Q è il sottospazio di V dato da

$$\ker(Q) := \{u \in V \mid Q(u + v) = Q(v) \quad \forall v \in V\}.$$

e la *nullità di* Q è la dimensione di $\ker(Q)$.

Esempio 8.2.5. Sia $Q \in Q(\mathbb{K}^3)$ data da

$$Q(x, y, z) := x^2 - y^2.$$

Determiniamo $\ker(Q)$ e quindi la nullità di Q. Sia $(a, b, c) \in \mathbb{K}^3$. Allora $(a, b, c) \in \ker(Q)$ se e solo se

$$(x + a)^2 - (y + b)^2 = x^2 - y^2$$

per ogni $x, y \in \mathbb{K}$. Siccome

$$(x + a)^2 - (y + b)^2 = x^2 + 2ax + a^2 - y^2 - 2by - b^2$$

questo equivale a

$$2ax + a^2 - 2by - b^2 = 0 \quad \forall x, y \in \mathbb{K}. \tag{8.2.2}$$

Se $\operatorname{char} \mathbb{K} \neq 2$ l'uguaglianza in (8.2.2) vale se e solo se $a = b = 0$. Quindi sotto questa ipotesi

$$\ker(Q) = \{(0, 0, c) \mid c \in \mathbb{K}\},$$

cioè la nullità di Q è 1. Se invece $\operatorname{char} \mathbb{K} = 2$ allora il membro di sinistra di (8.2.2) è uguale ad $(a + b)^2$, e quindi sotto questa ipotesi

$$\ker(Q) = \{(a, a, c) \mid c \in \mathbb{K}\},$$

cioè la nullità di Q è 2.

Osservazione 8.2.6. Sia $Q \in Q(V)$. Se $u \in \ker(Q)$, allora

$$Q(u) = Q(u + 0) = Q(0) = 0,$$

ma non vale il viceversa, cioè se $Q(u) = 0$ non è necessariamente vero che $u \in \ker(Q)$. Per esempio se $Q \in Q(\mathbb{K}^3)$ è la forma quadratica dell'Esempio 8.2.5 allora $Q(1, 1, 0) = 0$, ma se $\operatorname{char} \mathbb{K} \neq 2$ il vettore $(1, 1, 0)$ non appartiene al nucleo di Q.

Definizione 8.2.7. Una forma quadratica $Q \colon V \to \mathbb{K}$ è *degenere* se $\ker(Q) \neq \{0\}$, è *non degenere* se $\ker(Q) = \{0\}$.

Proposizione 8.2.8. *Sia $Q \in Q(V)$. Se $g \in \mathrm{GL}(V)$, allora* $\ker(gQ) = g(\ker(Q))$.

Dimostrazione. Per definizione $u \in \ker(gQ)$ se e solo se $(gQ)(u + v) = gQ(v)$ per ogni $v \in V$, cioè

$$Q(g^{-1}(v)) = gQ(v) = (gQ)(u + v) = Q(g^{-1}(u) + g^{-1}(v)) \quad \forall v \in V. \tag{8.2.3}$$

Dato $w \in V$ esiste $v \in V$ tale che $w = g^{-1}(v)$, quindi (8.2.3) vale se e solo se $g^{-1}(u) \in \ker(Q)$, cioè $u \in g(\ker(Q))$. □

Una conseguenza immediata della proposizione appena dimostrata è il seguente risultato.

Corollario 8.2.9. *Se $Q, Q' \in Q(V)$ sono congruenti, allora le loro nullità sono uguali, in particolare Q è no degenere se e solo se lo è Q'.*

Un'applicazione $I\colon Q(V) \to Z$ è *invariante per congruenza* se $I(Q) = I(gQ)$ per ogni $Q \in Q(V)$ e $g \in \mathrm{GL}(V)$, cioè se è costante sugli elementi di una classe di congruenza. Sotto questa ipotesi l'applicazione I (supponendo che sia calcolabile) può permettere di stabilire che due forme quadratiche Q, Q' su V non sono congruenti: se $I(Q) \neq I(Q')$ allora Q non è congruente a Q'. Per il Corollario 8.2.9 l'applicazione $Q(V) \to \mathbb{N}$ che associa a una forma quadratica la sua nullità è invariante per congruenza.

Esempio 8.2.10. Per $r \in \{1, \dots, n\}$ sia $Q_r \in Q(\mathbb{K}^n)$ data da

$$Q_r(x_1, \dots, x_n) := x_1^2 + x_2^2 + \dots + x_r^2.$$

Supponiamo che $\operatorname{char} \mathbb{K} \neq 2$. Procedendo come nell'Esempio 8.2.5 si vede che la nullità di Q_r è $n - r$, e quindi Q_r è congruente a Q_s solo se $r = s$. Se invece $\operatorname{char} \mathbb{K} = 2$ allora le Q_r sono congruenti tra loro perchè

$$Q_r(x_1, \dots, x_n) := (x_1 + x_2 + \dots + x_r)^2,$$

e quindi Q_r è congruente a Q_1.

La classificazione delle forme quadratiche a meno di congruenza si fa nel modo seguente. Prima si definiscono invarianti per congruenza, diciamo $I_1, \dots, I_m, \dots$, e poi si dimostra che forme quadratiche Q, Q' sono congruenti se e solo se $I_1(Q) = I_1(Q'), \dots, I_m(Q) = I_m(Q'), \dots$. Si intende che gli invarianti devono essere (più o meno) calcolabili. A rigor di logica possiamo definire come invariante l'applicazione quoziente $I\colon Q(V) \to Q(V)/\!\sim$ dove $\sim$ è la relazione di congruenza, e tautologicamente Q è congruente a Q' se e solo se $I(Q) = I(Q')$, ma questa "soluzione" non è di alcun aiuto. Dimostreremo nella Sezione 8.4 che la nullità è un invariante completo per forme quadratiche complesse, cioè due forme quadratiche complesse sono congruenti se e solo se hanno la stessa nullità. In generale non basta la nullità a distinguere le forme quadratiche a meno di congruenza, per esempio nel caso di forme quadratiche reali, vedi la Sezione 8.4.

8.3 Forme quadratiche e forme bilineari simmetriche

Polarizzazione di una forma quadratica

Da ora fino alla fine del capitolo *supponiamo che la caratteristica del campo* $\mathbb{K}$ *sia diversa da* 2. Con questa ipotesi a ogni forma quadratica su V è associata una forma bilineare simmetrica su $V \times V$ (le forme bilineari e bilineari simmetriche sono state definite nella Sezione 6.7), e la teoria delle forme quadratiche equivale a quella delle forme bilineari simmetriche. Se invece il campo ha caratteristica 2 la teoria delle forme quadratiche ha delle peculiarità (come si può intuire dagli Esempi 8.2.5 e 8.2.10) che preferiamo non trattare.

Ricordiamo che una forma bilineare su V è un'applicazione bilineare

$$\begin{array}{ccc} V \times V & \xrightarrow{F} & \mathbb{K} \\ (v, w) & \mapsto & F(v, w) \end{array} \tag{8.3.1}$$

e che è simmetrica se $F(v, w) = F(w, v)$ per ogni $v, w \in V$. Denotiamo con $\mathrm{Bil}(V)$ l'insieme delle forme bilineari su V, e con $\mathrm{Bil}^+(V) \subset \mathrm{Bil}(V)$ il sottoinsieme delle forme bilineari simmetriche. Entrambi sono sottospazi vettoriali dello spazio delle applicazioni da V^2 a $\mathbb{K}$, cf. Sezione 6.7.

Data $F \in \mathrm{Bil}^+(V)$ definiamo l'applicazione $q_F\colon V \to \mathbb{K}$ così:

$$\begin{array}{ccc} V & \xrightarrow{q_F} & \mathbb{K} \\ v & \mapsto & \Phi(v, v). \end{array} \tag{8.3.2}$$

Sia $\{v_1, \dots, v_n\}$ una base di V; la bilinearità e la simmetria di F danno che

$$q_F(x_1 v_1 + \dots + x_n v_n) = \sum_{1 \leqslant i \leqslant n} F(v_i, v_i) x_i^2 + 2 \sum_{1 \leqslant i < j \leqslant n} F(v_i, v_j) x_i x_j, \tag{8.3.3}$$

e quindi q_F è una forma quadratica V.

Definizione 8.3.1. Data $F \in \mathrm{Bil}^+(V)$ la forma quadratica *associata* a F è la q_F definita da (8.3.2).

Proposizione 8.3.2. *Sia V uno spazio vettoriale su $\mathbb{K}$. L'applicazione*

$$\begin{array}{ccc} \mathrm{Bil}^+(V) & \longrightarrow & Q(V) \\ F & \mapsto & q_F \end{array} \tag{8.3.4}$$

è un isomorfismo di spazi vettoriali su $\mathbb{K}$.

Dimostrazione. Si verifica immediatamente che l'applicazione in (8.3.4) è lineare. Quindi per dimostrare che è iniettiva è sufficiente dimostrare che se $q_F = 0$ allora $F = 0$. Siano $v, w \in V$. Per definizione di q_F (e simmetria di F) si ha

$$q_F(v + w, v + w) = q_F(v, v) + 2F(v, w) + q_F(w). \tag{8.3.5}$$

Siccome $q_F = 0$ segue che $2F(v, w) = 0$ e siccome $\operatorname{char} \mathbb{K} \neq 2$ ne deduciamo che $F(v, w) = 0$. Ora dimostriamo che l'applicazione in (8.3.4) è suriettiva. Sia $f \in Q(V)$. Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V. Per definizione di forma quadratica esistono $c_{ij} \in \mathbb{K}$ per $1 \leqslant i \leqslant j \leqslant n$ tali che

$$f\left(\sum_{i=1}^{n} x_k v_k\right) = \sum_{1 \leqslant i \leqslant j \leqslant n} c_{ij} x_i x_j.$$

Per $i, j \in \{1, \ldots, n\}$ poniamo

$$a_{ij} := \begin{cases} c_{ij} & \text{se } i = j, \\ c_{ij}/2 & \text{se } i < j, \\ c_{ji}/2 & \text{se } i > j, \end{cases}$$

(Notate che 2 è invertibile perchè $\operatorname{char} \mathbb{K} \neq 2$.) Sia $A \in M_{n,n}$ la matrice simmetrica $A := (a_{ij})$, e sia F la forma bilineare $F := \Phi_A^{\mathcal{B}}$, dove $\Phi_A^{\mathcal{B}}$ (vedi (6.7.1)) è la forma bilineare simmetrica data da

$$F(v, w) := X_{\mathcal{B}}(v)^t \cdot A \cdot X_{\mathcal{B}}(w).$$

La forma bilineare F è simmetrica perchè lo è A e, come si verifica subito, $F(v, v) = f(v)$ per ogni $v \in V$, cioè $q_F = f$. □

Definizione 8.3.3. Data una forma quadratica $q \in Q(V)$ la sua controimmagine per l'applicazione (8.3.4), chiamiamola F, è la *polarizzazione* di q. Se V è finitamente generato e $\mathcal{B}$ è una sua base porremo $M_{\mathcal{B}}(q) := M_{\mathcal{B}}(F)$.

Osservazione 8.3.4. Data $q \in Q(V)$, i valori $F(v, w)$ della polarizzazione F di q sono dati dall'*identità di polarizzazione*

$$F(v, w) = 2^{-1}(f(v + w) - f(v) - f(w)). \tag{8.3.6}$$

Infatti questo segue dall'uguaglianza in (8.3.5).

Esempio 8.3.5. Siano $\mathbb{K}$ un campo e $B \in M_{n,n}(\mathbb{K})$ (non facciamo alcuna ipotesi su B). Sia $q \in Q(\mathbb{K}^n)$ data da $q(X) := X^t \cdot B \cdot X$. La polarizzazione F di q è data da

$$F(X, Y) = \frac{1}{2} X^t \cdot (B^t + B) \cdot Y. \tag{8.3.7}$$

Infatti F è una forma bilineare *simmetrica* perchè $(B^t + B)^t = (B + B^t)$ e

$$F(X, X) = \frac{1}{2} X^t \cdot (B^t + B) \cdot X = \frac{1}{2}(X^t \cdot B^t \cdot X) + \frac{1}{2}(X^t \cdot B \cdot X) = \frac{1}{2}(X^t \cdot B^t \cdot X)^t + \frac{1}{2}(X^t \cdot B \cdot X) = X^t \cdot B \cdot X = q(X).$$

Esempio 8.3.6. Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. Il prodotto euclideo $\langle , \rangle$ è la polarizzazione della norma al quadrato, cioè della forma quadratica $f \colon V \longrightarrow \mathbb{R}$ definita da $f(v) = ||v||^2$.

Osservazione 8.3.7. Supponiamo che $\mathbb{K} = \mathbb{R}$, e che $q \colon V \to \mathbb{R}$ sia una forma quadratica. Siano $v, w \in V$ e $t \in \mathbb{R}$. Per l'uguaglianza in (8.3.5) abbiamo che

$$q(v + tw) - q(v) = 2tF(v, w) + t^2 q(w),$$

e quindi

$$\lim_{t \to 0} \frac{q(v + tw) - q(v)}{t} = 2F(v, w). \tag{8.3.8}$$

In altre parole la derivata direzionale di q nel punto v e nella direzione w è uguale a $2F(v, w)$. Diamo una formulazione leggermente diversa dell'equazione in (8.3.8). Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V, e quindi

$$\varphi(x_1, \ldots, x_n) := q(x_1 v_1 + \ldots + x_n v_n)$$

è una funzione polinomiale omogenea di grado 2. Siano $a, b \in \mathbb{R}^n$. Allora l'equazione in (8.3.8) equivale all'equazione

$$\sum_{i=1}^{n} \frac{\partial \varphi(a)}{\partial x_i} b_i = 2F(a_1 v_1 + \ldots + a_n v_n, b_1 v_1 + \ldots + b_n v_n). \tag{8.3.9}$$

Infatti il termine a sinistra è uguale a alla derivata in $t = 0$ della funzione

$$\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ t & \mapsto & \varphi(a + tb), \end{array}$$

cioè la derivata direzionale della funzione φ nel punto a e nella direzione b, e per l'equazione in (8.3.8) questa derivata direzionale è uguale al membro di sinistra in (8.3.9).

Nella Sezione 8.1 abbiamo definito l'azione di $\mathrm{GL}(V)$ su $\mathcal{P}_d(V)$, lo spazio vettoriale delle funzioni polinomiali omogenee di grado d su V (se $P \in \mathcal{P}_d(V)$ e $g \in \mathrm{GL}(V)$, la funzione polinomiale gP è definita ponendo $gP(v) = P(g^{-1}(v))$). In particolare $\mathrm{GL}(V)$ agisce su $Q(V) = \mathcal{P}_2(V)$. Siccome l'applicazione che associa a $F \in \mathrm{Bil}^+(V)$ la forma quadratica q_F è un isomorfismo tra $\mathrm{Bil}^+(V)$ e $Q(V)$, l'azione di $\mathrm{GL}(V)$ su $Q(V)$ dà un'azione su $\mathrm{Bil}^+(V)$. Quest'azione si definisce direttamente così: se $g \in \mathrm{GL}(V)$ e $F \in \mathrm{Bil}^+(V)$, allora l'applicazione

$$\begin{array}{ccc} V \times V & \xrightarrow{gF} & \mathbb{K} \\ (v, w) & \mapsto & F(g^{-1}v, g^{-1}w) \end{array} \tag{8.3.10}$$

è anch'essa bilineare simmetrica. Si verifica che in questo modo abbiamo definito un'azione di $\mathrm{GL}(V)$ su $\mathrm{Bil}^+(V)$, cioè $\mathrm{Id}_V F = F$, e $g(hF) = (gh)F$. Inoltre è ovvio che quest'azione corrisponde all'azione di $\mathrm{GL}(V)$ su $Q(V)$, cioè che se $F \in \mathrm{Bil}^+(V)$ e $g \in \mathrm{GL}(V)$ allora

$$g q_F = q_{gF}.$$

In altre parole (andando nella direzione opposta, cioè da $Q(V)$ a $\mathrm{Bil}^+(V)$), si ottiene lo stesso risultato se si agisce con $g \in \mathrm{GL}(V)$ su una forma quadratica e poi la si polarizza, oppure se prima la si polarizza e poi si agisce con g sulla polarizzazione.

In accordo con la Proposizione 8.3.2 da ora in poi scriveremo spesso una forma quadratica q su V come $q(v) = X_{\mathcal{B}}(v)^t \cdot A \cdot X_{\mathcal{B}}(v)$ dove $\mathcal{B}$ è una base di V e $A \in M^+_{n,n}(\mathbb{K})$ (cioè A è una matrice $n \times n$ simmetrica). Segue che la polarizzazione di q è data da $F(v, w) = X_{\mathcal{B}}(v)^t \cdot A \cdot X_{\mathcal{B}}(w)$. Il seguente risultato si ottiene con un facile conto che lasciamo al lettore.

Proposizione 8.3.8. *Sia $A \in M^+_{n,n}(\mathbb{K})$, e sia $F \in \mathrm{Bil}^+(V)$ data da $F(v, w) = X_{\mathcal{B}}(v)^t \cdot A \cdot X_{\mathcal{B}}(w)$. Se $g \in \mathrm{GL}(V)$, allora*

$$gF(v, w) = X_{\mathcal{B}}(v)^t \cdot (g^{-1})^t \cdot A \cdot g^{-1} \cdot X_{\mathcal{B}}(w). \tag{8.3.11}$$

Notiamo che se $A \in M^+_{n,n}(\mathbb{K})$ e $h \in \mathrm{GL}(V)$, allora $h^t \cdot A \cdot h$ è simmetrica.

Definizione 8.3.9. Matrici $A, B \in M_{n,n}(\mathbb{K})$ sono *congruenti* se esiste $G \in \mathrm{GL}_n(\mathbb{K})$ tale che $A = G^t \cdot B \cdot G$.

Osservazione 8.3.10. Da quello che abbiamo detto segue che classificare le forme quadratiche su V a meno di congruenza equivale a classificare le forme bilineari simmetriche su $V \times V$ a meno di congruenza o a classificare le matrici simmetriche $n \times n$ (dove $n := \dim V$) a meno di congruenza.

La formula in (6.7.7) permette di definire un invariante per l'azione di $\mathrm{GL}(V)$ sulle forme bilineari simmetriche su V (equivalentemente sulle forme quadratiche su V). Per definire il codominio dell' invariante consideriamo la relazione $x \sim y$ su $\mathbb{K}$ definita come segue: $x \sim y$ se esiste $\mu \in K^*$ tale che $x = \mu^2 \cdot y$. La relazione è di equivalenza perchè $x = 1^2 x$, se $x = \mu^2 \cdot y$ allora $y = (\mu^{-1})^2 \cdot x$, e se $x = \mu^2 \cdot y$ e $y = \lambda^2 \cdot z$ allora $x = (\mu\lambda)^2 \cdot z$.

Definizione 8.3.11. $\mathrm{Sqf}(\mathbb{K})$ è l'insieme delle classi di equivalenza per la relazione $\sim$ su $\mathbb{K}$ definita sopra, cioè $x \sim y$ se esiste $\mu \in K^*$ tale che $x = \mu^2 \cdot y$.

Esempio 8.3.12. $\mathrm{Sqf}(\mathbb{C}) = \{[0], [1]$ e $\mathrm{Sqf}(\mathbb{R}) = \{[0], [1], [-1]\}$. Le classi di equivalenza $\mathrm{Sqf}(\mathbb{Q})$ diverse da $[0]$ sono rappresentate, senza ripetizioni, da $[p_1 \cdot p_2 \cdot \ldots \cdot p_n]$, dove $p_1 < p_2 < \ldots < p_n$ è una lista crescente di numeri primi (la classe di equivalenza $[1]$ corrisponde alla lista vuota).

Ora sia $F \in \mathrm{Bil}^+(V)$. Siano $\mathcal{B}, \mathcal{C}$ basi di V. Per la formula in (6.7.7) abbiamo la seguente relazione tra i determinanti delle matrici di Gram $M_{\mathcal{B}}(F), M_{\mathcal{C}}(F)$:

$$\mathrm{Det}\, M_{\mathcal{B}}(F) = \mathrm{Det}\, M_{\mathcal{C}}(F) \cdot \mathrm{Det}\, M^{\mathcal{C}}_{\mathcal{B}}(\mathrm{Id})^2.$$

Quindi possiamo associare a F un ben definito elemento di $\mathrm{Sqf}(\mathbb{K})$, cioè

$$\mathrm{Disc}(F) := [\mathrm{Det}\, M_{\mathcal{B}}(F)] \in \mathrm{Sqf}(\mathbb{K}),$$

dove $\mathcal{B}$ è una qualsiasi base di V, e che chiamiamo il *discriminante di* F. Inoltre per la formula in (8.3.11) se F_1 è congruente a F_2 allora $\mathrm{Disc}(F_1) = \mathrm{Disc}(F_2)$. Quindi il discriminante è un invariante per l'azione di $\mathrm{GL}(V)$ sulle forme bilineari simmetriche su V (equivalentemente sulle forme quadratiche su V).

Esempio 8.3.13. Siano f, h le forme quadratiche su $\mathbb{Q}^2$ definite da

$$f(x_1, x_2) = 2x_1^2 - 2x_1x_2 + 3x_2^2, \quad h(x_1, x_2) = 5x_1^2 + 6x_1x_2 + 3x_2^2. \tag{8.3.12}$$

Dimostriamo che f e h non sono equivalenti. Le matrici associate a f e h nella base standard $\mathcal{S}$ di $\mathbb{Q}^2$ sono

$$M_{\mathcal{S}}(f) = \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix} \qquad M_{\mathcal{S}}(h) = \begin{pmatrix} 5 & 3 \\ 3 & 3 \end{pmatrix}$$

Quindi $\mathrm{Disc}(f) = [5]$ e $\mathrm{Disc}(h) = [6]$. Siccome $5/6$ non è il quadrato di un razionale, f e h non sono equivalenti. Osserviamo che le formule in (8.3.12) definiscono anche forme quadratiche su $\mathbb{R}^2$; per non confonderci chiamiamole F e H rispettivamente. Le forme quadratiche reali F e H sono congruenti, cioè esiste $g \in \mathrm{GL}_2(\mathbb{R})$ tale che $F = gH$ (questo sarà chiaro una volta letta la Sottosezione (7.4)).

Forme bilineari e applicazioni $V \to V^\vee$

Sia $F \in \mathrm{Bil}(V)$. Se $v \in V$, allora l'applicazione da V in $\mathbb{K}$ definita da $w \mapsto F(w, v)$ è lineare perchè F è lineare a sinistra, e analogamente l'applicazione da V in $\mathbb{K}$ definite da $w \mapsto F(v, w)$ è lineare perchè F è lineare a destra. Quindi possiamo definire due applicazioni da V a $V^\vee$:

$$\begin{array}{ccc} V & \xrightarrow{D_F} & V^\vee \\ v & \mapsto & (w \mapsto F(w,v)) \end{array} \qquad \begin{array}{ccc} V & \xrightarrow{S_F} & V^\vee \\ v & \mapsto & (w \mapsto F(v,w)). \end{array} \tag{8.3.13}$$

L'applicazione D_F è lineare perchè F è lineare a sinistra, e S_F è lineare perchè F è lineare a destra. Prima di passare al prossimo risultato, consideriamo un'applicazione lineare

$$f\colon V \longrightarrow V^\vee.$$

Abbiamo l'applicazione duale di f (vedi la Sottosezione 3.13) $f^\vee\colon (V^\vee)^\vee \to V^\vee$, e siccome V è finitamente generato esiste l'isomorfismo canonico $V \xrightarrow{\sim} (V^\vee)^\vee$ che associa a $v \in V$ l'applicazione lineare $\mathrm{Val}(v)\colon V^\vee \to \mathbb{K}$ data da $\varphi \mapsto \varphi(v)$ (vedi la Proposizione 3.13.6). Ne segue che la duale di f può essere vista come applicazione lineare

$$f^\vee\colon V \to V^\vee. \tag{8.3.14}$$

(Attenzione: questo solo perchè abbiamo supposto che il codominio di f sia il duale del dominio di f.) Esplicitamente, se $v, w \in V$ allora

$$f^\vee(v)(w) = f(w)(v). \tag{8.3.15}$$

Infatti, per l'identificazione di $(V^\vee)^\vee$ con V data da Val (vedi la Proposizione 3.13.6), $f^\vee(v)$ è uguale a $f^\vee(\mathrm{Val}(v))$, cioè la composizione di $f\colon V \to V^\vee$ con $\mathrm{Val}(v)\colon V^\vee \to \mathbb{K}$.

Proposizione 8.3.14. *Sia $F \in \mathrm{Bil}(V)$, e siano $D_F, S_F\colon V \to V^\vee$ le applicazioni lineari definite in* (8.3.13).

(a) La duale di D_F è uguale a S_F e, viceversa, la duale di S_F è uguale a D_F.

(b) F è simmetrica se e solo se $D_F = S_F$.

Dimostrazione. (a): Dimostriamo che $D_F^\vee = S_F$. Siano $v, w \in V$. Per l'uguaglianza in (8.3.15) si ha

$$D_F^\vee(v)(w) = D_F(w)(v) = F(v, w) = S_F(v)(w). \tag{8.3.16}$$

Analogo ragionamento dimostra che $S_F^\vee = D_F$. (b): Ovvio. □

Definizione 8.3.15. Un'applicazione lineare $f\colon V \to V^\vee$ è *simmetrica* se $f = f^\vee$ (questo ha senso per la discussione fatta sopra, vedi (8.3.14)).

Osservazione 8.3.16. Sia $F \in \mathrm{Bil}(V)$. Per la Proposizione 8.3.14 l'applicazione D_F è simmetrica se e solo se F è simmetrica (idem per S_F).

Definizione 8.3.17. Se $F \in \mathrm{Bil}^+(V)$ denotiamo $D_F = S_F$ con $\mathcal{L}_F \colon V \to V^\vee$.

Proposizione 8.3.18. *Sia $F \in \mathrm{Bil}^+(V)$. Allora il nucleo di $\mathcal{L}_F \colon V \to V^\vee$ è uguale al nucleo della forma quadratica q_F (vedi la Definizione 8.2.4) associata a F.*

Dimostrazione. Prima dimostriamo che $\ker(\mathcal{L}_F) \subset \ker(q_F)$. Sia $v_0 \in \ker(\mathcal{L}_F)$. Se $v \in V$ si ha

$$q_F(v_0 + v) = F(v_0 + v, v_0 + v) = F(v_0, v_0) + F(v_0, v) + F(v, v_0) + F(v, v) = F(v, v) = q_F(v),$$

e quindi $v_0 \in \ker(q_F)$. Ora sia $v_0 \in \ker(q_F)$. Se $v \in V$ allora, per l'identità di polarizzazione (vedi (8.3.6)) abbiamo

$$F(v_0, v) = 2^{-1}(q_F(v_0 + v) - q_F(v_0) - q_F(v)).$$

Siccome $v_0 \in \ker(q_F)$ si ha $q_F(v_0) = 0$ (vedi l'Osservazione 8.2.6) e $q_F(v_0 + v) = q_F(v_0)$. Quindi $F(v_0, v) = 0$, cioè $v_0 \in \ker(\mathcal{L}_F)$. □

Definizione 8.3.19. Sia $F \in \mathrm{Bil}^+(V)$. Il *rango* di F è il rango di $\mathcal{L}_F \colon V \to V^\vee$, e si denota $\mathrm{rg}(F)$.

Osservazione 8.3.20. Sia $F \in \mathrm{Bil}^+(V)$. Per la Proposizione 8.3.18 la nullità di q_F (vedi la Definizione 8.2.4) è uguale a $(\dim V - \mathrm{rg}(F))$. In particolare F è non-degenere (vedi la Definizione 8.2.7) se e solo se $\mathcal{L}_F$ ha rango massimo, cioè è un isomorfismo.

Chiudiamo questa sottosezione dando l'inverso del passaggio che porta da un'applicazione bilineare $F \colon V \times V \to \mathbb{K}$ alle due applicazioni lineari $D_F, S_F \colon V \to V^\vee$. Data un'applicazione lineare $f \colon V \to V^\vee$ possiamo definire applicazioni bilineari $G, H \colon V \times V \to \mathbb{K}$ tali che $D_G = f$ e $S_H = f$ leggendo al contrario le uguaglianze in (8.3.16). Esplicitamente, poniamo

$$G(v, w) = f(v)(w), \qquad H(v, w) = f(w)(v).$$

Ne segue che G (e H) è simmetrica se e solo se lo è f. Perciò è equivalente studiare forme bilineari simmetriche su $V \times V$ o applicazioni lineari simmetriche $V \to V^\vee$ - questo include l'azione di $\mathrm{GL}(V)$ su entrambi gli spazi (l'azione su $\mathcal{L}(V, V^\vee)$ è quella indotta dall'azione simultanea su V e $V^\vee$, cioè se $g \in \mathrm{GL}(V)$ e $\varphi \in \mathcal{L}(V, V^\vee)$ si pone $g(\varphi)(v) := (g^{-1})^\vee(\varphi(g^{-1}(v)))$ per $v \in V$).

Ortogonalità

Definizione 8.3.21. Sia $F \in \mathrm{Bil}^+(V)$. I vettori $v, w \in V$ sono *F-ortogonali* (o *F-perpendicolari*) se $F(v, w) = 0$; in simboli $v \perp w$. Se $S \subset V$ l'*F-ortogonale* di S è

$$S^\perp := \{w \in V \mid F(v, w) = 0 \quad \forall v \in S\}.$$

Se $S = \{v_0\}$ (cioè consiste di un solo elemento) denotiamo $\{v_0\}^\perp$ con $v_0^\perp$.

Per definizione $v \perp w$ se $F(v, w) = 0$, ma siccome per ipotesi F è simmetrica questo equivale a $F(w, v) = 0$. In altre parole $v \perp w$ se e solo se $w \perp v$.

Spesso, in presenza di una forma bilineare simmetrica F, l'F-ortogonalità si denota semplicemente ortogonalità. Se $q \in Q(V)$ è una forma quadratica su V e $F \in \mathrm{Bil}^+(V)$ è la sua polarizzazione, q-ortogonalità (o semplicemente ortogonalità) significa F-ortogonalità.

Esempio 8.3.22. Consideriamo $\mathbb{E}^2$ (o $\mathbb{E}^3$) con il prodotto scalare standard - vedi (6.1.1). Siano $\overrightarrow{PQ}, \overrightarrow{QR} \in \mathbb{E}^2$ vettori non nulli. Allora $\overrightarrow{PQ} \perp \overrightarrow{QR}$ se e solo se la retta PQ è perpendicolare a alla retta QR. Infatti sia $\{v_1, v_2\}$ una base ortonormale di $\mathbb{E}^2$ e $\overrightarrow{PQ} = x_1 v_1 + x_2 v_2$, $\overrightarrow{QR} = y_1 v_1 + y_2 v_2$. Per il Teoremea di Pitagora la retta PQ è perpendicolare a alla retta QR se e solo se il quadrato della lunghezza di $\overrightarrow{PR}$ è uguale alla somma dei quadrati delle lunghezze di $\overrightarrow{PQ}$ e $\overrightarrow{QR}$ ovvero se e solo se

$$x_1^2 + 2x_1y_1 + y_1^2 + x_2^2 + 2x_2y_2 + y_2^2 = (x_1 + y_1)^2 + (x_2 + y_2)^2 = x_1^2 + x_2^2 + y_1^2 + y_2^2$$

cioè se e solo se $0 = x_1y_1 + x_2y_2 = \langle x_1v_1 + x_2v_2, y_1v_1 + y_2v_2 \rangle$.

Esempio 8.3.23. Sia $F \in \mathrm{Bil}^+(V)$ e sia $q_F \in Q(V)$ la forma quadratica associata. Allora $V^\perp = \ker(\mathcal{L}_F) = \ker(q_F)$.

Osservazione 8.3.24. Sia V uno spazio vettoriale finitamente generato su $\mathbb{K}$ e $F \in \mathrm{Bil}^+(V)$. Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V e $A := M_{\mathcal{B}}(F)$. Sia $w \in V$: allora

$$w^\perp = \{v \in V \mid X_{\mathcal{B}}(w)^t \cdot A \cdot X_{\mathcal{B}}(v) = 0\}. \tag{8.3.17}$$

Lemma 8.3.25. *Sia V uno spazio vettoriale su $\mathbb{K}$, e $F \in \mathrm{Bil}^+(V)$. L'ortogonale di un sottoinsieme $S \subset V$ è un sottospazio di V.*

Dimostrazione. Se $v_0 \in V$ abbiamo che $v_0^\perp = \ker(\mathcal{L}_F(v_0))$ e quindi $v_0^\perp$ è un sottospazio lineare di V. Siccome

$$S^\perp = \bigcap_{v \in S} v^\perp$$

segue che $S^\perp$ è intersezione di sottospazi lineari e perciò è un sottospazio lineare. □

Lemma 8.3.26. *Sia $F \in \mathrm{Bil}^+(V)$. Se $U \subset V$ è il sottospazio generato da $u_1, \ldots, u_m$, allora*

$$U^\perp = \bigcap_{i=1}^m u_i^\perp \tag{8.3.18}$$

Dimostrazione. È ovvio che il membro di sinistra di (8.3.18) è contenuto nel membro di destra. Resta da dimostrare che il membro di destra di (8.3.18) è contenuto nel membro di sinistra. Supponiamo che $v \in u_i^\perp$ per $1 \leqslant i \leqslant m$. Sia $u \in U$: siccome U è generato da $u_1, \ldots, u_m$ esistono $\lambda_1, \lambda_m \in \mathbb{K}$ tali che $u = \sum_{i=1}^m \lambda_i u_i$. Per linearità di F abbiamo che

$$F(v, u) = F(v, \sum_{i=1}^m \lambda_i u_i) = \sum_{i=1}^m \lambda_i F(v, u_i) = 0.$$

□

Proposizione 8.3.27. *Supponiamo che V sia finitamente generato e che $F \in \mathrm{Bil}^+(V)$ sia non degenere. Se $U \subset V$ è un sottospazio allora*

$$\dim U^\perp = \dim V - \dim U.$$

Dimostrazione. Sia $\{u_1, \ldots, u_m\}$ una base di U. Per il Lemma 8.3.26

$$U^\perp = \bigcap_{i=1}^m u_i^\perp = \bigcap_{i=1}^m \ker \mathcal{L}_F(u_i).$$

cioè $U^\perp$ è il nucleo dell'applicazione lineare

$$\begin{array}{ccc} V & \xrightarrow{\Phi} & \mathbb{K}^m \\ v & \mapsto & (\mathcal{L}_F(u_1)(v), \ldots, \mathcal{L}_F(u_m)(v)) \end{array}$$

Siccome F è non-degenere, $\mathcal{L}_F(u_1), \ldots, \mathcal{L}_F(u_m)$ sono elementi di $V^\vee$ linearmente indipendenti, e quindi Φ è suriettiva (se Φ non fosse suriettiva esisterebbe un elemnto non nullo dell'annullatore $\mathrm{Ann}(\mathrm{im}(\Phi)) \subset (\mathbb{K}^m)^\vee$, cioè una relazione di dipendenza lineare tra $\mathcal{L}_F(u_1), \ldots, \mathcal{L}_F(u_m)$). Siccome Φ è suriettiva, seegue che $U^\perp = \ker(\Phi)$ ha dimensione uguale a $(\dim V - m)$. □

Diagonalizzazione

Definizione 8.3.28. Sia V uno spazio vettoriale finitamente generato e $f \in Q(V)$. La forma quadratica f è *diagonale* nella base $\mathcal{B} = \{v_1, \ldots, v_n\}$ di V se esistono $c_1, \ldots, c_n \in \mathbb{K}$ tali che per $(x_1, \ldots, x_n) \in \mathbb{K}^n$ si abbia

$$f(x_1 v_1 + \ldots + x_n v_n) = \sum_{i=1}^n c_i x_i^2, \tag{8.3.19}$$

o in altre parole se la matrice $M_{\mathcal{B}}(f)$ è diagonale. Analogamente, $F \in \mathrm{Bil}^+(V)$ è *diagonale* nella base $\mathcal{B}$ se la matrice $M_{\mathcal{B}}(f)$ è diagonale. Nel primo caso diciamo che la base $\mathcal{B}$ *diagonalizza* f, nel secondo che la base $\mathcal{B}$ *diagonalizza* F.

Osservazione 8.3.29. Sia $F \in \mathrm{Bil}^+(V)$, e sia q_F la forma quadratica associata. Allora F è diagonale nella base $\mathcal{B}$ se e solo se lo è q_F vedi (6.7.2) e (8.3.3).

Teorema 8.3.30. *Sia V uno spazio vettoriale finitamente generato e sia $F \in \mathrm{Bil}^+(V)$. Esiste una base $\mathcal{B}$ di V che diagonalizza F.*

Dimostrazione. Per induzione sulla dimensione di V. Se $\dim V = 0$ non c'è nulla da dimostrare (se volete potete cominciare l'induzione da $\dim V = 1$, anche in questo caso non cé nulla da dimostrare). Dimostriamo il passo induttivo. Sia $n = \dim V$. Se $F = 0$ qualsiasi base diagonalizza F. Supponiamo che $F \neq 0$. Allora $q_F \neq 0$ e quindi esiste $v_0 \in V$ tale che $q_F(v_0) \neq 0$. In particolare $0 \neq v_0$ e $v_0 \notin v^\perp$. Sia $U := v_0^\perp$. La restrizione di F a $U \times U$, data da

$$\begin{array}{ccc} U \times U & \xrightarrow{G} & \mathbb{K} \\ (u_1, u_2) & \mapsto & F(u_1, u_2) \end{array} \tag{8.3.20}$$

è bilineare simmetrica. Siccome $\dim U = (n-1)$ (perchè $v_0 \neq 0$) l'ipotesi induttiva dà che esiste una base $\{w_1, \ldots, w_{n-1}\}$ di U che diagonalizza G. Ora $\mathcal{B} := \{w_1, \ldots, w_{n-1}, v_0\}$ è una base di V e l'Osservazione 6.7.9 mostra che $\mathcal{B}$ diagonalizza F. □

Corollario 8.3.31. *Sia $A \in M^+_{n,n}(\mathbb{K})$. Allora A è congruente a una matrice diagonale Λ.*

Dimostrazione. Segue dal Teorema 8.3.30 e dalla Proposizione 6.7.11. □

Corollario 8.3.32 (Lagrange)**.** *Siano V uno spazio vettoriale finitamente generato $f \in Q(V)$. Esiste una base di V che diagonalizza f.*

Dimostrazione. Segue dal Teorema 8.3.30 e dall'Osservazione 8.3.29. □

Sia $f\colon V \to \mathbb{K}$ una forma quadratica su uno spazio vettoriale finitamente generato su $\mathbb{K}$. Possiamo determinare una base che diagonalizza f senza passare per la forma bilineare associata a f procedendo come segue. Se $f = 0$ non c'è nulla da fare, supponiamo che $f \neq 0$. Quindi esiste $u_n \in V$ tale che $f(u_n) = c_n \neq 0$. Sia $\mathcal{C} = \{u_1, \ldots, u_n\}$ una base di V che estende il vettore non-nullo u_n. Sia $A = M_{\mathcal{C}}(f)$. Calcolando otteniamo che $c_n = f(u_n) = a_{nn}$. Quindi (ricordate che $c_n \neq 0$)

$$\begin{aligned} f(z_1u_1 + \ldots + z_nu_n) &= \sum_{i \leqslant i,j \leqslant (n-1)} a_{ij}x_ix_j + c_n(z_n^2 + 2c_n^{-1}\sum_{i=1}^{n-1} a_{in}z_iz_n) = \\ &= \sum_{i \leqslant i,j \leqslant (n-1)} a_{ij}x_ix_j + c_n\left(z_n + c_n^{-1}\sum_{i=1}^{n-1} a_{in}z_i\right)^2 - c_n^{-1}\left(\sum_{i=1}^{n-1} a_{in}z_i\right)^2. \end{aligned} \tag{8.3.21}$$

Sia

$$r := \sum_{i \leqslant i,j \leqslant (n-1)} a_{ij}x_ix_j - c_n^{-1}\left(\sum_{i=1}^{n-1} a_{in}z_i\right)^2.$$

Notate che $r \in \mathbb{K}[x_1, \ldots, x_{n-1}]$ è un polinomio omogeneo di grado 2. La (8.3.21) si può riscrivere così:

$$f(z_1u_1 + \ldots + z_nu_n) = r(z_1, \ldots, z_{n-1}) + c_n(z_n + c_n^{-1}\sum_{i=1}^{n-1} a_{in}z_i)^2. \tag{8.3.22}$$

Esiste una base $\mathcal{D} = \{w_1, \ldots, w_n\}$ con coordinate associate $(y_1, \ldots, y_n)$ legate alle coordinate $(z_1, \ldots, z_n)$ dalle formule

$$y_i = z_i, \quad 1 \leqslant i \leqslant (n-1), \qquad y_n = z_n + c_n^{-1}\sum_{i=1}^{n-1} a_{in}z_i. \tag{8.3.23}$$

(Questo perchè la matrice associata all'applicazione lineare definita da (8.3.23) è non-singolare). Si ha $f(y_1w_1 + \ldots + y_nw_n) = r(y_1, \ldots, y_{n-1}) + c_ny_n^2$. Sia $U \subset V$ il sottospazio generato da $w_1, \ldots, w_{n-1}$. La formula $g(y_1w_1 + \ldots + y_{n-1}w_{n-1}) := r(y_1, \ldots, y_{n-1})$ definisce una forma quadratica su U. Ora iteriamo il procedimento con V sostituito da U: arriveremo a una base che diagonalizza f. Ora supponiamo che $\mathbb{K} = \mathbb{R}$. Se $c_i \neq 0$ sostituiamo a v_i il vettore $|c_i|^{-1/2}v_i$. Se $\mathbb{K} = \mathbb{C}$ sostituiamo a ogni v_i tale che $c_i \neq 0$ il vettore $c_i^{-1/2}v_i$.

Esempio 8.3.33. Sia

$$\begin{array}{ccc} \mathbb{R}^3 & \xrightarrow{f} & \mathbb{R} \\ (x_1,x_2,x_3) & \mapsto & x_1x_2+x_2x_3+x_3x_1 \end{array}$$

Si ha che $f(0,1,1)=1\neq 0$. Siano

$$u_1=(1,0,0),\quad u_2=(0,1,0),\quad u_3=(0,1,1).$$

Si ha che

$$f(y_1u_1+y_2u_2+y_3u_3)=f(y_1,y_2+y_3,y_3)=y_1y_2+2y_1y_3+y_2y_3+y_3^2=$$
$$=y_1y_2+(y_3+y_1+\frac{1}{2}y_2)^2-(y_1+\frac{1}{2}y_2)^2=-y_1^2-\frac{1}{4}y_2^2+(y_3+y_1+\frac{1}{2}y_2)^2. \quad (8.3.24)$$

Siano (z_1,z_2,z_3) le coordinate su $\mathbb{R}^3$ date da

$$\begin{array}{rcl} z_1 &=& y_1 \\ z_2 &=& y_2 \\ z_3 &=& y_1+\frac{1}{2}y_2+y_3 \end{array}$$

Quindi

$$\begin{array}{rcl} y_1 &=& z_1 \\ y_2 &=& z_2 \\ y_3 &=& -z_1-\frac{1}{2}z_2+z_3 \end{array}$$

Perciò la base con coordinate (z_1,z_2,z_3) è $\{w_1,w_2,w_3\}$ dove

$$w_1=(1,0,-1),\quad w_2=(0,1,-1/2),\quad w_3=(0,0,1).$$

Si ha che

$$f(z_1w_1+z_2w_2+z_3w_3)=f(y_1u_1+y_2u_2+(-z_1-\frac{1}{2}z_2+z_3)u_3)=-z_1^2-\frac{1}{4}z_2^2+z_3^2.$$

Siano (t_1,t_2,t_3) le coordinate su $\mathbb{R}^3$ date da

$$\begin{array}{rcl} t_1 &=& z_1 \\ t_2 &=& z_2/2 \\ t_3 &=& z_3 \end{array}$$

La base di $\mathbb{R}^3$ che corrisponde a (t_1,t_2,t_3) è $\{r_1,r_2,r_3\}$ dove $r_1=w_1$, $r_2=w_2/2$, $r_3=w_3$. Abbiamo che

$$f(t_1r_1+t_2r_2+t_3r_3)=-t_1^2-t_2^2+t_3^2.$$

8.4 Forme quadratiche reali e complesse

Forme quadratiche complesse

La classificazione delle forme quadratiche su uno spazio vettoriale complesso è molto semplice.

Proposizione 8.4.1. *Sia V uno spazio vettoriale complesso. Due forme quadratiche su V sono congruenti se e solo se hanno lo stesso rango (o equivalentemente la stessa nullità).*

Dimostrazione. Siano $f,g\in Q(V)$. Se f è congruente a g, allora le nullità di f e g sono uguali per il Corollario 8.2.9 e quindi, per l'Osservazione 8.3.20 sono uguali anche i ranghi di f e g. Ora supponiamo che f e g abbiano lo stesso rango, e dimostriamo che sono congruenti. Per il Corollario 8.3.32 esistono basi $\mathcal{B}=\{v_1,\ldots,v_n\}$ e $\mathcal{C}=\{w_1,\ldots,w_n\}$ di V tali che $M_{\mathcal{B}}(f)$ e $M_{\mathcal{C}}(f)$ siano diagonali. Sia $r:=\mathrm{rg}(f)=\mathrm{rg}(g)$. Riordinando i vettori di $\mathcal{B}$ e di $\mathcal{C}$, se necessario, possiamo assumere che

$$f(x_1v_1+\ldots+x_nv_n)=\lambda_1x_1^2+\ldots+\lambda_rx_r^2,\qquad g(x_1w_1+\ldots+x_nw_n)=\mu_1x_1^2+\ldots+\mu_rx_r^2,$$

dove i λ_j e i μ_j sono tutti non nulli. Riscalando i vettori delle basi $\mathcal{B}$ e $\mathcal{C}$ possiamo assumere che tutti i λ_j e i μ_j siano uguali a 1. Infatti basta porre

$$v'_j:=\begin{cases}\lambda_j^{-1/2}v_j & \text{se } 1\leqslant j\leqslant r,\\ v_j & \text{se } r<j\leqslant n,\end{cases},\quad w'_j:=\begin{cases}\mu_j^{-1/2}w_j & \text{se } 1\leqslant j\leqslant r,\\ w_j & \text{se } r<j\leqslant n,\end{cases}$$

dove $\lambda_j^{-1/2}$ e $\mu_j^{-1/2}$ è una radice quadrata di λ_j^{-1} e μ_j^{-1} rispettivamente. Le matrici $M_{\mathcal{B}'}(f)$ e $M_{\mathcal{C}'}(g)$ associate a f, g nelle basi $\mathcal{B}' = \{v_1', \dots, v_n'\}$ e $\mathcal{C}' = \{w_1', \dots, w_n'\}$ sono uguali tra loro perchè

$$f(x_1v_1' + \dots + x_nv_n') = x_1^2 + \dots + x_r^2, \qquad g(x_1w_1' + \dots + x_nw_n') = x_1^2 + \dots + x_r^2,$$

e quindi f è congruente a g.

□

Forme quadratiche reali

Due forme quadratiche su uno spazio vettoriale reale che hanno lo stesso rango non sono necessariamente congruenti. Per esempio le forme quadratiche $f, h \in Q(\mathbb{R})$ definite da $f(x) = x^2$ e $h(x) = -x^2$ hanno rango 1 ma non sono congruenti perchè $f(\mathbb{R}) = [0, +\infty)$ e $h(\mathbb{R}) = (-\infty, 0]$. In questo esempio ci basta il discriminante per distinguere le due forma quadratiche giacchè il discriminante di f è $[1]$ e il discriminante di f è $[-1]$. In generale nemmeno aggiungendo il discriminante riusciamo a dare un sistema completo di invarianti per forme quadratiche reali. Per esempio le forme quadratiche $f, h \in Q(\mathbb{R}^2)$ definite da $f(x) = x_1^2 + x_2^2$ e $h(x) = -x_1^2 - x_2^2$ hanno rango 2 e discriminante $[1]$ ma non sono congruenti perchè $f(\mathbb{R}^2) = [0, +\infty)$ e $h(\mathbb{R}^2) = (-\infty, 0]$. Quindi abbiamo bisogno di introdurre nuovi specifici invarianti per forme quadratiche reali.

Definizione 8.4.2. Siano V uno spazio vettoriale *reale* e $f\colon V \to \mathbb{R}$ una forma quadratica. La f è *definita positiva* se $f(v) > 0$ per ogni $0 \neq v \in V$, è *definita negativa* se $f(v) < 0$ per ogni $0 \neq v \in V$ (ovvero se $-f$ è definita positiva). In simboli: $f > 0$ siginifica che f è definita positiva e $f < 0$ significa che f è definita negativa. Una $F \in \mathrm{Bil}^+(V)$ è *definita positiva* se q_F è definita positiva, è *definita negativa* se q_F è definita negativa. Diciamo che $f \in Q(V)$ (o $F \in \mathrm{Bil}^+(V)$) è *definita* se è definita positiva o è definita negativa.

Osservazione 8.4.3. Siano V uno spazio vettoriale *reale*. Una forma quadratica $f\colon V \to \mathbb{R}$ è definita positiva se e solo se la polarizzazione di f è un prodotto scalare euclideo. Per questo consideriamo "forma quadratica definita positiva" come sinonimo di "prodotto scalare euclideo".

Definizione 8.4.4. Una matrice $A \in M_{n,n}^+(\mathbb{R})$ è *definita positiva* se $q_A^{\mathcal{S}}$ è definita positiva, dove $\mathcal{S}$ è la base standard di $\mathbb{R}^n$, è *definita negativa* se $q_A^{\mathcal{S}}$ è definita negativa. Esplicitamente: A è definita positiva se $X^t \cdot A \cdot X > 0$ per ogni vettore colonna non nullo $X \in \mathbb{R}^n$, ed è definita negativa se $X^t \cdot A \cdot X < 0$ per ogni vettore colonna non nullo $X \in \mathbb{R}^n$. Per esempio 1_n è definita positiva e -1_n è definita negativa.

Definizione 8.4.5. Siano V uno spazio vettoriale reale e $f \in Q(V)$. La *segnatura positiva* di f (denotata $s_+(f)$) è la massima dimensione di un sottospazio vettoriale $U \subset V$ tale che $f_{|U}$ sia definita positiva, e la *segnatura negativa* di f (denotata $s_-(f)$) è la massima dimensione di un sottospazio vettoriale $U \subset V$ tale che $f_{|U}$ sia definita negativa.

Lemma 8.4.6. *Siano V uno spazio vettoriale reale e $f \in Q(V)$. Se $g \in \mathrm{GL}(V)$ allora $s_+(gf) = s_+(f)$ e $s_-(gf) = s_-(f)$.*

Dimostrazione. Abbiamo un'applicazione

$$\begin{array}{ccc} \{U \subset V \mid U \text{ è un sottospazio di } V \text{ e } (f|_U) > 0\} & \longleftrightarrow & \{W \subset V \mid W \text{ è un sottospazio di } W \text{ e } (gf|_W) > 0\} \\ U & \mapsto & g(U). \end{array} \tag{8.4.1}$$

Infatti supponiamo che $f(v) > 0$ per ogni $v \in U$ non nullo. Se $w \in g(U)$ è non nullo, allora $w = g(v)$ per un $v \in U$ non nullo, e quindi

$$gf(w) = f(g^{-1}w) = f(g^{-1}gv) = f(v) > 0.$$

Questo mostra che se $f_{|U}$ è definita positiva, allora gf è definita positiva se $g(U)$, cioè l'applicazione in (8.4.1) è ben definita. L'applicazione è chiaramente iniettiva. Siccome $f = g^{-1}(gf)$, questo mostra anche che se $gf_{|W}$ è definita positiva, allora f è definita positiva se $g^{-1}(U)$, e perciò Llapplicazione è suriettiva. Siccome l'applicazione in (8.4.1) è biunivoca segue che $s_+(gf) = s_+(f)$. Si dimostra in modo analogo che $s_-(gf) = s_-(g)$. □

In altre parole la segnatura positiva e negativa di una forma quadratica reale sono invarianti per congruenza. Dimostreremo che danno un sistema completo di invarianti.

Proposizione 8.4.7 (Sylvester)**.** *Siano V uno spazio vettoriale reale e $f \in Q(V)$. Supponiamo che $\mathcal{B} = \{v_1, \dots, v_n\}$ sia una base di V tale che*

$$f(x_1v_1 + \dots + x_nv_n) = c_1x_1^2 + \dots + c_ax_a^2 - d_{a+1}x_{a+1}^2 - \dots - d_{a+b}x_{a+b}^2. \qquad \forall\ (x_1v_1 + \dots + x_nv_n) \in V. \tag{8.4.2}$$

Supponiamo che $c_i > 0$ per ogni $1 \leqslant i \leqslant a$ e che $d_i > 0$ per ogni $a+1 \leqslant i \leqslant a+b$. Allora $s_+(f) = a$, $s_-(f) = b$.

Dimostrazione. Siano $V_+ := \langle v_1, \dots, v_a\rangle$, $V_- := \langle v_{a+1}, \dots, v_{a+b}\rangle$ e $V_0 := \langle v_{a+b+1}, \dots, v_n\rangle$. Osserviamo che

$$\dim(V_+ + V_0) = n - b, \qquad \dim(V_- + V_0) = n - a. \tag{8.4.3}$$

Siccome $f|_{V_+} > 0$ e $f|_{V_-} < 0$ abbiamo

$$s_+(f) \geqslant a, \qquad s_-(f) \geqslant b. \tag{8.4.4}$$

Supponiamo che la prima diseguaglianza sia stretta cioè $s_+(f) > a$; arriveremo a un assurdo. Per definizione esiste un sottospazio $U \subset V$ tale che $\dim U > a$ e $f|_U > 0$. Per (8.4.3) la formula di Grassmann dà che

$$\dim(U \cap (V_- + V_0)) = \dim U + \dim(V_- + V_0) - \dim(U + V_- + V_0) \geqslant \\ \geqslant \dim U + \dim(V_- + V_0) - n = \dim U - a > 0.$$

Quindi esiste $0 \neq v \in U \cap (V_- + V_0)$. Siccome $v \in (V_- + V_0)$ le sue prime a coordinate rispetto alla base $\mathcal{B}$ sono nulle; segue da (8.4.2) che $f(v) \leqslant 0$. D'altra parte $v \in U$ e per ipotesi $f|_U > 0$, quindi $f(v) > 0$. La contraddizione dimostra che non esiste un sottospazio $U \subset V$ tale che $\dim U > a$ e $f|_U > 0$; per (8.4.4) segue che $s_+(f) = a$. Si dimostra in modo analogo che non puè essere $s_-(f) > b$ e quindi $s_-(f) = b$. □

Proposizione 8.4.8. *Sia V uno spazio vettoriale reale. Forme quadratiche f, h su V sono congruenti se e solo se $s_+(f) = s_+(h)$ e $s_-(f) = s_-(h)$.*

Dimostrazione. Se f è congruente ad h allora $s_\pm(f) = s_\pm(h)$ per il Lemma 8.4.6. Ora supponiamo che $s_\pm(f) = s_\pm(h)$ e dimostriamo che f è congruente ad h. Per il Corollario 8.3.32 esistono basi $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{C} = \{w_1, \dots, w_n\}$ di V tali che $M_{\mathcal{B}}(f)$ e $M_{\mathcal{C}}(g)$ sono matrici diagonali. Siccome $s_\pm(f) = s_\pm(h)$, la Proposizione 8.4.7 dà che il numero di entrate positive sulla diagonale principale $M_{\mathcal{B}}(f)$ e $M_{\mathcal{C}}(g)$ sono uguali e analogamente il numero di entrate negative. Quindi riordinando i vettori di $\mathcal{B}$ e di $\mathcal{C}$, se necessario, possiamo assumere che

$$f(x_1v_1 + \dots + x_nv_n) = \lambda_1x_1^2 + \dots + \lambda_ax_a^2 - \lambda_{a+1}x_{a+1}^2 - \dots - \lambda_{a+b}x_{a+b}^2$$

e

$$h(x_1w_1 + \dots + x_nw_n) = \mu_1x_1^2 + \dots + \mu_ax_a^2 - \mu_{a+1}x_{a+1}^2 - \dots - \mu_{a+b}x_{a+b}^2,$$

dove i λ_j, μ_j sono tutti positivi. (Quindi $a = s_+(f) = s_+(h)$ e $b = s_-(f) = s_-(h)$.) Riscalando i vettori delle basi $\mathcal{B}$ e $\mathcal{C}$ possiamo assumere che tutti i λ_j e i μ_j siano uguali a 1 (vedi la dimostrazione della Proposizione 8.4.1). Infatti basta porre

$$v_j' := \begin{cases} \lambda_j^{-1/2}v_j & \text{se } 1 \leqslant j \leqslant a+b, \\ v_j & \text{se } a+b < j \leqslant n, \end{cases}, \qquad w_j' := \begin{cases} \mu_j^{-1/2}w_j & \text{se } 1 \leqslant j \leqslant a+b, \\ w_j & \text{se } a+b < j \leqslant n. \end{cases}$$

Le matrici $M_{\mathcal{B}'}(f)$ e $M_{\mathcal{C}'}(h)$ associate a f, h nelle basi $\mathcal{B}' = \{v_1', \dots, v_n'\}$ e $\mathcal{C}' = \{w_1', \dots, w_n'\}$ sono uguali tra loro perchè

$$f(x_1v_1' + \dots + x_nv_n') = x_1^2 + \dots + x_a^2 - x_{a+1}^2 - \dots - x_{a+b}^2, \qquad h(x_1w_1' + \dots + x_nw_n') = x_1^2 + \dots + x_a^2 - x_{a+1}^2 - \dots - x_{a+b}^2,$$

e quindi f è congruente ad h. □

8.5 Il gruppo ortogonale di una forma quadratica

Sia q una forma quadratica sullo spazio vettoriale V (sul campo $\mathbb{K}$). Consideriamo il sottoinsieme $\mathrm{O}(V,q) \subset \mathrm{GL}(V)$ definito da

$$\mathrm{O}(V,q) := \{g \in \mathrm{GL}(V) \mid gq = q\}, \tag{8.5.5}$$

cioè quello che si chiama lo *stabilizzatore di* q (relativamente all'azione di $\mathrm{GL}(V)$ su $Q(V)$ per congruenza).

Lemma 8.5.1. $\mathrm{O}(V,q)$ *è un sottogruppo di* $\mathrm{GL}(V)$.

Dimostrazione. Ricordiamo che $gq\colon V \to \mathbb{K}$ è la forma quadratica definita da $gq(v) := q(g^{-1}(v))$ per $v \in V$. Chiaramente $\mathrm{Id}_V \in \mathrm{O}(V,q)$. Ora supponiamo che $g \in \mathrm{O}(V,q)$, cioè $q(g^{-1}(v)) = q(v)$ per ogni $v \in V$. Ponendo $w = g^{-1}(v)$, possiamo riscrivere l'uguaglianza come $q(w) = q(gw)$ per ogni $v \in V$, cioè $g^{-1} \in \mathrm{O}(V,q)$. Per ultimo supponiamo che $g_1, g_2 \in \mathrm{O}(V,q)$. Segue che per ogni $v \in V$ si ha

$$(g_1g_2)q(v) = q((g_1g_2)^{-1}(v)) = q(g_2^{-1}(g_1^{-1}(v))) = q(g_1^{-1}(v)) = q(v),$$

e quindi $g_1g_2 \in \mathrm{O}(V,q)$. □

$\mathrm{O}(V,q)$ è il *gruppo ortogonale* della forma quadratica.

Osservazione 8.5.2. Sia $\mathcal{B}$ una base di V, e sia $M_{\mathcal{B}}(q)$ la matrice simmetrica associata alla forma quadratic a q, cioè $q(v) = X_{\mathcal{B}}(v)^t \cdot M_{\mathcal{B}}(q) \cdot X_{\mathcal{B}}(v)$. Un automorfismo $g \in \mathrm{GL}(V)$ è in $\mathrm{O}(V,q)$ se e solo se

$$M_{\mathcal{B}}^{\mathcal{B}}(g)^t \cdot M_{\mathcal{B}}(q) \cdot M_{\mathcal{B}}^{\mathcal{B}}(g) = M_{\mathcal{B}}(q). \tag{8.5.6}$$

Esempio 8.5.3. Poniamo $q = 0$; il gruppo $\mathrm{O}(V,0)$ è uguale a $\mathrm{GL}(V)$, e viceversa se $\mathrm{O}(V,q) = \mathrm{GL}(V)$ allora $q = 0$.

Esempio 8.5.4. Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo (di dimensione finita), e sia $q \in Q(V)$ la forma quadratica associata, cioè $q(v) = ||v||^2$. Allora $\mathrm{O}(V,q) = \mathrm{O}(V, \langle , \rangle)$.

Definizione 8.5.5. Sia V uno spazio vettoriale reale. Un *prodotto scalare* su V è una forma bilineare simmetrica non degenere $\langle , \rangle$ su V.

Sia $\langle , \rangle$ un prodotto scalare su V. Diciamo che $\langle , \rangle$ è definito positivo se la forma quadratica associata è definita positiva, cioè se $\langle v, v \rangle$ è positivo per ogni $v \in V$ non nullo. Quindi un prodotto scalare euclideo su V non è altro che un prodotto scalare definito positivo.

Esempio 8.5.6. Dato $c > 0$, consideriamo il prodotto scalare su $\mathbb{R}^4$ definito da

$$q(x_1, x_2, x_3, t) := c^2t^2 - x_1^2 - x_2^2 - x_3^2. \tag{8.5.7}$$

Il gruppo ortogonale $\mathrm{O}(\mathbb{R}^4, q)$ si chiama *gruppo di Lorentz*, ed è fondamentale nella Teoria della Relatività speciale. Sinteticamente: se (x_1, x_2, x_3, t) e (x_1', x_2', x_3', t') sono le coordinate spazio-temporali di due sistemi inerziali (con uguali unità di misura), la relazione tra coordinate di uno stesso evento è data da

$$\begin{bmatrix} x_1' \\ x_2' \\ x_3' \\ t' \end{bmatrix} = A \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ t \end{bmatrix} + B, \tag{8.5.8}$$

dove B è una matrice colonna 4×1 e $A \in \mathrm{O}(\mathbb{R}^4, q)$. Notate la differenza tra le trasformazioni di Lorentz (o meglio di Poincaré se $B \neq 0$) e quelle classiche Galileiane, in cui $t = t'$ e (x_1', x_2', x_3') si ottiene da (x_1, x_2, x_3) via un'applicazione affine ortogonale (per il prodotto euclideo standard su $\mathbb{R}^3$). Due eventi che sono simultanei nel sistema di coordinate (x_1, x_2, x_3, t) non sono necessariamente simultanei nel sistema di coordinate (x_1', x_2', x_3', t'). Ciò che rimane invariato nei due sistemi di riferimento è la velocità della luce (data da c). Esplicitamente: se (x_1, x_2, x_3, t) e (y_1, y_2, y_3, s) sono coordinate di due eventi nel primo sistema di riferimento, allora scegliendo opportunamente $A \in \mathrm{O}(\mathbb{R}^4, q)$ e $B \in M_{4,1}(\mathbb{R})$ le coordinate degli stessi eventi nel secondo sistema di riferimento possono essere date da qualsiasi coppia (x_1', x_2', x_3', t') e (y_1', y_2', y_3', s') tale che

$$c^2(t-s)^2 - (x_1-y_1)^2 - (x_2-y_2)^2 - (x_3-y_3)^2 = c^2(t'-s')^2 - (x_1'-y_1')^2 - (x_2'-y_2')^2 - (x_3'-y_3')^2.$$

Definizione 8.5.7. Siano V uno spazio vettoriale reale e $\langle , \rangle$ un prodotto scalare su V. Il *gruppo ortogonale* di $\langle , \rangle$ è il gruppo ortogonale $\mathrm{O}(V,q)$ dove $q \in Q(V)$ è la forma quadratica associata a $\langle , \rangle$, cioè $q(v) = \langle v, v \rangle$.

Il gruppo ortogonale di un prodotto scalare euclideo è stato esaminato nella Sezione 6.10. D'altra parte se $\langle v, v\rangle$ è definito negativo, cioè $\langle v, v\rangle$ è negativo per ogni $v \in V$ non nullo, allora il prodotto scalare $\langle , \rangle_0$ definito da $\langle v, w\rangle_0 := -\langle v, w\rangle$ è definito positivo, e ne segue che $\mathrm{O}(V, \langle , \rangle) = \mathrm{O}(V, \langle , \rangle_0)$, e quindi per $\mathrm{O}(V, \langle , \rangle)$ valgono i risultati della Sezione 6.10. Se il prodotto scalare non è definito (nè positivo nè negativo) il realtivo gruppo ortogonale cambia "carattere", vedi per esempio l'Esercizio 8.14.

Esercizi del Capitolo 8

Esercizio 8.1. *Sia $q \in Q(\mathbb{R}^2)$ data da $q(X) := X^t \cdot A \cdot X$ dove*

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

1. *Sia $\mathcal{B}$ la base di $\mathbb{R}^2$ data da $\mathcal{B} := \{(1,1), (1,-1)\}$. Calcolate $M_{\mathcal{B}}(q)$.*
2. *Verificate che q è definita positiva.*

Esercizio 8.2. *Sia $V \subset C^0([0,1])$ il sottospazio generato dalle funzioni*

$$f = 1, \qquad g = \cos \pi t, \qquad h := \sin \pi t.$$

e sia $F \in \mathrm{Bil}(V)$ definita da

$$\begin{array}{ccc} V \times V & \xrightarrow{F} & \mathbb{R} \\ (\phi, \psi) & \mapsto & \int_0^1 \phi\psi \end{array}$$

1. *Verificate che f, g, h sono linearmente indipendenti.*
2. *Calcolate $M_{\mathcal{B}}(F)$ dove $\mathcal{B} := \{f, g, h\}$ (è una base di V per il punto 1).*

Esercizio 8.3. *Siano $f, g \in Q(\mathbb{R}^3)$ date da*

$$f(x_1, x_2, x_3) = -7x_1^2 + 2x_1x_2 - 6x_1x_3 + 5x_2^2 - x_3^2, \quad g(x_1, x_2, x_3) = x_1^2 - 4x_1x_2 - 2x_1x_3 + 8x_2^2 + 6x_2x_3 + 2x_3^2$$

Siano $\mathcal{B}$ e $\mathcal{C}$ le basi di $\mathbb{R}^3$ date da

$$\{(1,1,3), (2,-1,0), (0,0,1)\}, \qquad \{(0,1,2), (3,0,1), (1,-1,0)\}$$

rispettivamente.

(1) Calcolate $M_{\mathcal{B}}(f)$ e $M_{\mathcal{C}}(g)$.

(2) Determinate se $(\mathbb{R}^3, f)$ è isomorfo a $(\mathbb{R}^3, g)$.

Esercizio 8.4. *Sia $f \in Q(\mathbb{R}^{2n})$ definita da*

$$f(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n) = x_1y_1 + x_2y_2 + \ldots + x_ny_n.$$

Determinate la segnatura di f.

Esercizio 8.5. *Sia $A \in M^+_{3,3}(\mathbb{R})$ definita da*

$$A := \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 3 \end{bmatrix}$$

Trovate una base che diagonalizza $q_A^{\mathcal{S}}$ ($\mathcal{S}$ è la base standard di $\mathbb{R}^3$).

Esercizio 8.6. *(a) Siano $A, B \in M_{2,2}(\mathbb{Q})$ date da*

$$A := \begin{bmatrix} 3 & 2 \\ 2 & 5 \end{bmatrix}, \qquad B := \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \tag{8.5.9}$$

e siano $f, g \in Q(\mathbb{R}^2)$ date da $f(X) := X^t \cdot A \cdot X$ e $g(X) := X^t \cdot B \cdot X$ rispettivamente. Dimostrate che f non è congruente a g.

(b) Siano $A, B \in M_{2,2}(\mathbb{R})$ date da (8.5.9) *e siano $\phi, \psi \in Q(\mathbb{R}^2)$ date da $\phi(X) := X^t \cdot A \cdot X$ e $\psi(X) := X^t \cdot B \cdot X$ rispettivamente. Dimostrate che ϕ è congruente a ψ.*

Esercizio 8.7. *Sia V uno spazio vettoriale e $f \in Q(V)$. Un sottospazio $U \subset V$ è* isotropo *per f se $f|_U$ è la forma quadratica nulla. Dimostrate che*

$$\max\{U \subset V \mid f_{|U} = 0\} = \dim \ker(f) + \min\{s_+(f),\, s_-(f)\}.$$

Esercizio 8.8. *Se $a \in \mathbb{C}$ denotiamo con $Re(a)$ e $Im(a)$ la parte reale e immaginaria di a rispettivamente. Siano $f, g\colon \mathbb{C} \times \mathbb{C} \to \mathbb{R}$ le applicazioni definite da*

$$f(w, z) := Re(w\overline{z}), \qquad g(w, z) := Im(w\overline{z}).$$

Verificate che f e g sono forme bilineari su $\mathbb{C}$ considerato come spazio vettoriale su $\mathbb{R}$. Verificate che f e g sono non-degeneri. Quale tra f e g è simmetrica ?

Esercizio 8.9. *Se $A \in M_{n,n}(\mathbb{K})$ la* traccia *di A è data da*

$$\operatorname{Tr} A := \sum_{i=1}^{n} a_{ii}.$$

Sia $\Phi\colon M_{2,2}(\mathbb{R}) \times M_{2,2}(\mathbb{R}) \to \mathbb{R}$ definita da

$$\Phi(A, B) := \operatorname{Tr}(AB).$$

Verificate che Φ è bilineare e simmetrica. Determinate una base che diagonalizza Φ.

Esercizio 8.10. *Sia $q\colon M_{n,n}(\mathbb{R}) \to \mathbb{R}$ la forma quadratica definita da $q(A) := \operatorname{Tr}(A^2)$. Determinate rango e segnatura di q. (Suggerimento: esaminate la restrizione di q al sottospazio delle matrici simmetriche/anti-simmetricha).*

Esercizio 8.11. *Sia f la forma quadratica su $\mathbb{R}^2$ data da*

$$f(x_1, x_2) = x_1^2 + 2x_1x_2 + 3x_2^2.$$

Trovate una base di $\mathbb{R}^2$, ortonormale *per il prodotto euclideo standard, che diagonalizza f.*

Esercizio 8.12. *Siano V uno spazio vettoriale finitamente generato di dimensione n, e*

$$U_1 \subset U_2 \subset \ldots \subset U_n = V$$

una catena di sottospazi vettoriali tali che $\dim U_i = i$ per $i \in \{1, \ldots, n\}$. Sia $f \in Q(V)$ una forma quadratica tale che, per ogni $1 \leqslant i \leqslant n$, la restrizione di f a U_i sia non degenere. Dimostrate che esiste una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ di V che diagonalizza f, e tale che, per ogni $1 \leqslant i \leqslant n$,

$$\langle v_1, \ldots, v_i \rangle = U_i.$$

Esercizio 8.13. *Sia $A \in M^+_{n,n}(\mathbb{R})$ una matrice reale $n \times n$ simmetrica e $f(X) = X^t \cdot A \cdot X$ la forma quadratica associata. Per $p \in \{1, \ldots, n\}$ sia $A(p)$ la matrice simmetrica $p \times p$ con entrate i, j uguale all'entrata i, j di A. Per esempio*

$$A(1) = (a_{11}), \quad A(2) := \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad A(3) := \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Supponete che $\operatorname{Det} A(p) \neq 0$ per ogni p. Dimostrate che $s_-(f)$ è uguale al numero di cambi di segno nella sequenza

$$1,\ \operatorname{Det} A(1),\ \operatorname{Det} A(2), \ldots, \operatorname{Det} A(n), \tag{8.5.10}$$

ovvero che, denotando con c il numero di cambi di segno nella sequenza (8.5.10)*, la segnatura di f è uguale a $n - 2c$. (Suggerimento: usate i risultati degli Esercizi 8.12.)*

Esercizio 8.14. *Sia $\langle , \rangle$ il prodotto scalare su $\mathbb{R}^2$ definito da*

$$\langle X, Y \rangle := X^t \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot Y$$

In altre parole $\langle X, X \rangle = 2x_1x_2$. Dimostrate che

$$\mathrm{O}(\mathbb{R}^2, \langle , \rangle) = \left\{ \begin{bmatrix} t & 0 \\ 0 & t^{-1} \end{bmatrix} \mid t > 0 \right\} \sqcup \left\{ \begin{bmatrix} t & 0 \\ 0 & t^{-1} \end{bmatrix} \mid t < 0 \right\} \sqcup \left\{ \begin{bmatrix} 0 & t \\ t^{-1} & 0 \end{bmatrix} \mid t > 0 \right\} \sqcup \left\{ \begin{bmatrix} 0 & t \\ t^{-1} & 0 \end{bmatrix} \mid t < 0 \right\}.$$

Inoltre dimostrate che ciascuno dei sottoinsiemi di $\mathrm{O}(\mathbb{R}^2, \langle , \rangle)$ elencati sopra è una classe di equivalenza per la relazione data da $A \sim B$ se A è deformabile in B (la definizione di "deformabile" si dà come per gli elementi del gruppo ortogonale di uno spazio vettoriale euclideo, vedi la Definizione 6.10.8).

Capitolo 9

Teoremi spettrali

9.1 Introduzione

Dimostreremo i seguenti risultati:

(I) Se $A \in M^+_{n,n}(\mathbb{R})$ è una matrice reale simmetrica esiste una matrice ortogonale G tale che $G^{-1} \cdot A \cdot G$ sia diagonale. In altre parole A è diagonalizzabile e in più esiste una base *ortonormale* di autovettori. Questa è una formulazione del Teorema spettrale reale, e l'aggettivo "spettrale" si riferisce all'affermazione che gli autovalori una matrice reale simmetrica sono reali (lo spettro di una matrice quadrata è la collezione dei suoi autovalori contati con molteplicità). Esistono formulazioni all'apparenza diverse ma equivalenti, vedi la Sezione 9.2. In particolare il Teorema spettrale permette di classificare le orbite (cioè le classi di equivalenza) per l'azione del gruppo ortogonale di uno spazio vettoriale euclideo $(V, \langle , \rangle)$ sullo spazio $Q(V)$ delle forme quadratiche su V (l'azione è quella che si ottiene restringendo a $\mathrm{O}(V)$ l'azione di $\mathrm{GL}(V)$).

(II) Il Teorema spettrale complesso è l'analogo complesso del Teorema spettrale reale, e afferma che se $A \in M_{n,n}(\mathbb{C})$ è una matrice complessa hermitiana allora esiste una matrice unitaria U tale che $U^{-1} \cdot A \cdot U$ sia diagonale e reale.

Prima di dimostrare il Teorema spettrale complesso discuteremo le forme hermitiane, un analogo delle forme bilineari simmetriche reali quando lo spazio vettoriale reale viene sostituito da uno spazio vettoriale complesso.

9.2 Il teorema spettrale reale

Il teorema spettrale reale: altre formulazioni

In questa sottosezione diamo formulazioni equivalenti del Teorema spettrale reale, nella Sottosezione 9.2 ne daremo la dimostrazione.

In tutto il corso di questa sezione $(V, \langle , \rangle)$ è uno spazio vettoriale euclideo.

Definizione 9.2.1. Un endomorfismo $A\colon V \to V$ è *simmetrico* se per ogni coppia di vettori $u, w \in V$ vale

$$\langle A(u), w\rangle = \langle u, A(w)\rangle. \tag{9.2.1}$$

Osservazione 9.2.2. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base ON di V. Un endomorfismo $A\colon V \to V$ è simmetrico se e solo se la matrice $M^{\mathcal{B}}_{\mathcal{B}}(A)$ è simmetrica. Per vederlo poniamo $M = (m_{ij}) := M^{\mathcal{B}}_{\mathcal{B}}(A)$. Ricordiamo che la colonna j esima di M è data dalle coordinate di $A(v_j)$, e siccome $\mathcal{B}$ è ON

$$A(v_j) = \sum_{i=1}^{n} \langle A(v_j), v_i\rangle v_i,$$

cioè $m_{ij} = \langle A(v_j), v_i\rangle$. Ora supponiamo che A sia simmetrico. Allora

$$m_{ij} = \langle A(v_j), v_i\rangle = \langle v_j, A(v_i)\rangle = \langle A(v_i), v_j\rangle = m_{ji}, \tag{9.2.2}$$

e perciò M è simmetrica. Viceversa, se M è simmetrica allora da (9.2.2) segue che per ogni coppia di vettori v_i, v_j della base $\mathcal{B}$ si ha $\langle A(v_j), v_i\rangle = \langle v_j, A(v_i)\rangle$, e per linearità di A segue che vale (9.2.1).

Poniamo

$$\mathrm{End}^+(V) := \{A \in \mathrm{End}(V) \mid A \text{ è simmetrico}\}. \tag{9.2.3}$$

Enfatizziamo che la notazione ha senso solo se $(V, \langle , \rangle)$ è uno spazio vettoriale euclideo.

Osservazione 9.2.3. Un semplice conto mostra che la somma di endomorfismi simmetrici è un endomorfismo simmetrico, e che un multiplo di un endomorfismo simmetrico è simmetrico, e quindi $\mathrm{End}^+(V)$ è un sottospazio vettoriale di $\mathrm{End}(V)$. L'applicazione

$$\begin{array}{ccc} \mathrm{End}^+(V) & \xrightarrow{M_{\mathcal{B}}^{\mathcal{B}}} & M_{n,n}^+(\mathbb{R}) \\ A & \mapsto & M_{\mathcal{B}}^{\mathcal{B}}(A), \end{array} \tag{9.2.4}$$

è un isomorfismo di spazi vettoriali, questo segue dall'Osservazione 9.2.2.

Osservazione 9.2.4. Ricordiamo che un'applicazione lineare $g\colon V \to V^\vee$ è simmetrica se è uguale all'applicazione duale $g^\vee\colon V \to V^\vee$ (vedi la Definizione 8.3.15). Se $(V, \langle , \rangle)$ è uno spazio vettoriale euclideo, gli endomorfismi $V \to V$ simmetrici corrispondono alle applicazioni lineari simmetriche $V \to V^\vee$ passando per l'isomorfismo (simmetrico) $\mathcal{L}\colon V \to V^\vee$ associato al prodotto scalare euclideo $\langle , \rangle$ (vedi la Sottosezione 8.3). Infatti un endomorfismo $A\colon V \to V$ è simmetrico se e solo se la composizione

$$V \xrightarrow{A} V \xrightarrow{\mathcal{L}} V^\vee$$

è un'applicazione simmetrica.

Proposizione 9.2.5. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo e sia $A\colon V \to V$ un endomorfismo simmetrico. Se λ, μ sono autovalori (reali) distinti di A, allora gli autospazi $V_\lambda(A)$, $V_\mu(A)$ sono ortogonali.*

Dimostrazione. Siano $u \in V_\lambda(A)$ e $w \in V_\mu(A)$. Siccome A è simmetrico abbiamo

$$\lambda\langle u, w\rangle = \langle A(u), w\rangle = \langle u, A(w)\rangle = \mu\langle u, w\rangle.$$

Siccome $\lambda \neq \mu$ segue che $\langle u, w\rangle = 0$. □

Corollario 9.2.6. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo e sia $A\colon V \to V$ un endomorfismo simmetrico. Se A è diagonalizzabile (su $\mathbb{R}$) allora esiste una base ON di V che diagonalizza A.*

Dimostrazione. Per ipotesi esiste una decomposizione in somma diretta

$$V = V_{\lambda_1}(A) \oplus \ldots \oplus V_{\lambda_m}(A).$$

Per $i \in \{1, \ldots, m\}$ scegliamo una base ON $\mathcal{B}_i = \{v_{i,1}, \ldots, i_{i,d_i}\}$ dell'autospazio $V_{\lambda_i}(A)$. La base di V data da

$$\mathcal{B} := \{v_{1,1}, \ldots, i_{1,d_1}, \ldots, v_{i,1}, \ldots, i_{i,d_i}, \ldots, v_{m,1}, \ldots, i_{m,d_m}\}$$

è costituita di autovettori di A, e per la Proposizione 9.2.5 è anche ON. □

Ora diamo una prima formulazione del teorema spettrale.

Teorema 9.2.7 (Teorema spettrale reale, prima formulazione)**.** *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo e sia $A\colon V \to V$ un endomorfismo simmetrico. Allora esiste una base ON di V che diagonalizza A.*

La seconda formulazione del Teorema spettrale reale fa intervenire le forme quadratiche. Sia $A\colon V \to V$ un endomorfismo (qualsiasi, per ora)del nostro spazio vettoriale euclideo $(V, \langle , \rangle)$. Possiamo associare ad A la forma bilineare

$$\begin{array}{ccc} V \times V & \xrightarrow{\Phi_A} & \mathbb{R} \\ (u, w) & \mapsto & \langle A(u), w\rangle \end{array}$$

Ora notiamo che A è un endomorfismo simmetrico se solo se Φ_A è forma bilineare simmetrica. Infatti siano $u, w \in V$; allora

$$\Phi_A(u, w) = \langle A(u), w\rangle, \qquad \Phi_A(w, u) = \langle A(w), u\rangle = \langle u, A(w)\rangle,$$

e quindi vale (9.2.1) se e solo se $\Phi_A(u, w) = \Phi_A(w, u)$. Quindi possiamo definire un'applicazione

$$\begin{array}{ccc} \mathrm{End}^+(V) & \xrightarrow{\Phi} & \mathrm{Bil}^+(V) \\ A & \mapsto & \Phi_A \end{array} \tag{9.2.5}$$

Ricordate che il dominio è uno spazio vettoriale (vedi l'Osservazione 9.2.3), quindi Φ è un' un'applicazione tra spazi vettoriali.

Proposizione 9.2.8. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. L'applicazione in* (9.2.5) *è un isomorphismo di spazi vettoriali.*

Dimostrazione. La Proposizione è una riformulazione dell'Osservazione 9.2.3. Più precisamente sia $\mathcal{B}$ una base ON di V. Abbiamo l'isomorphismo $M_{\mathcal{B}}^{\mathcal{B}} \colon \mathrm{End}^+(V) \xrightarrow{\sim} M_{n,n}^+(\mathbb{R})$ in (9.2.4) e l'isomorphismo

$$\begin{array}{ccc} \mathrm{Bil}^+(V) & \xrightarrow{M_{\mathcal{B}}} & M_{n,n}^+(\mathbb{R}) \\ F & \mapsto & M_{\mathcal{B}}(F). \end{array}$$

Si verifica subito che vale l'uguaglianza

$$M_{\mathcal{B}}^{\mathcal{B}}(A) = M_{\mathcal{B}}(\Phi_A). \tag{9.2.6}$$

Quindi l'applicazione lineare Φ in (9.2.5) è uguale a $M_{\mathcal{B}}^{-1} \circ M_{\mathcal{B}}^{\mathcal{B}}$, ed essendo la composizione di due isomorfismi è anch'essa un isomorfismo. □

Dato un endomorfismo $A \colon V \to V$ simmetrico, definiamo la forma quadratica q_A su V così:

$$q_A(v) := \langle A(v), v \rangle. \tag{9.2.7}$$

Componendo l'isomorfismo in (9.2.5) con l'isomorfismo $\mathrm{Bil}^+(V) \to Q(V)$ della Proposizione 8.3.2 otteniamo l'isomorfismo

$$\begin{array}{ccc} \mathrm{End}^+(V) & \xrightarrow{\sim} & Q(V) \\ A & \mapsto & q_A \end{array} \tag{9.2.8}$$

Ora notiamo che una base $\mathcal{B}$ di V diagonalizza A se e solo se diagonalizza la forma quadratica q_A: infatti questo segue subito dall'uguaglianza in (9.2.6). Questo dimostra che il Teorema spettrale reale appena enunciato, cioè il Teorema 9.2.7, è equivalente al seguente enunciato.

Teorema 9.2.9 (Teorema spettrale reale, seconda formulazione)**.** *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo e sia $f \colon V \to \mathbb{R}$ una forma quadratica. Allora esiste una base ortonormale $\mathcal{B} = \{v_1, \ldots, v_n\}$ di V che diagonalizza f, cioè tale che*

$$f(x_1 v_1 + \ldots + x_n v_n) = \sum_{i=1}^{n} \lambda_i x_i^2. \tag{9.2.9}$$

Osservazione 9.2.10. Il gruppo ortogonale $\mathrm{O}(V)$ agisce per coniugio sul dominio dell'isomorfismo in (9.2.8), cioè se $A \in \mathrm{End}^+(V)$ e $g \in \mathrm{O}(V)$ allora $g \circ A \circ g^{-1}$ è un endomorfismo simmetrico. Infatti

$$\langle g \circ A \circ g^{-1}(v), w \rangle = \langle A \circ g^{-1}(v), g^{-1}(w) \rangle = \langle g^{-1}(v), A \circ g^{-1}(w) \rangle = \langle v, g \circ A \circ g^{-1}(w) \rangle.$$

(La prima e la terza uguaglianza valgono perchè g è ortogonale, la seconda vale perchè A è simmetrica.) In alternativa possiamo scegliere una base ortonormale $\mathcal{B}$ di V e notiamo che siccome g è ortogonale si ha $M_{\mathcal{B}}^{\mathcal{B}}(g) = (M_{\mathcal{B}}^{\mathcal{B}}(g)^{-1})^t$. Quindi

$$M_{\mathcal{B}}^{\mathcal{B}}(g \circ A \circ g^{-1}) = M_{\mathcal{B}}^{\mathcal{B}}(g) \cdot M_{\mathcal{B}}^{\mathcal{B}}(A) \cdot M_{\mathcal{B}}^{\mathcal{B}}(g^{-1}) = (M_{\mathcal{B}}^{\mathcal{B}}(g)^{-1})^t \cdot M_{\mathcal{B}}^{\mathcal{B}}(A) \cdot M_{\mathcal{B}}^{\mathcal{B}}(g)^{-1}, \tag{9.2.10}$$

perciò $M_{\mathcal{B}}^{\mathcal{B}}(g \circ A \circ g^{-1})$ è simmetrica, e quindi anche A lo è. D'altra parte $\mathrm{O}(V)$ agisce per congruenza sul codominio dello stesso isomorfismo. Le due azioni sono compatibili, cioè se $A \in \mathrm{End}^+(V)$ e $g \in \mathrm{O}(V)$ allora

$$g \circ A \circ g^{-1}(v) = q_A(g^{-1}(v))$$

per ogni $v \in V$. Questo si può verificare scegliendo una base ortonormale $\mathcal{B}$ di V e procedendo come sopra (vedi le uguaglianze in (9.2.10)).

Osservazione 9.2.11. Se esaminiamo i ragionamenti svolti, ci accorgiamo che abbiamo dimostrato il seguente risultato. Un endomorfismo simmetrico $A \colon V \to V$ di uno spazio vettoriale euclideo $(V, \langle , \rangle)$ è diagonalizzato in una base ortonormale $\mathcal{B} = \{v_1, \ldots, v_n\}$ se e solo se la forma quadratica q_A è diagonale nella stessa base ortonormale $\mathcal{B}$. Inoltre, in questo caso l'autovalore corrispondente a v_i è uguale al coefficiente λ_i nella formula in (9.2.9).

Il Teorema spettrale reale: dimostrazione

Prima di procedere con Proposizioni, Lemmi etc., diamo una motivazione geometrica della dimostrazione. Quindi ci riferiamo all seconda formulazione del Teorema spettrale reale. Supponiamo che V sia $\mathbb{R}^n$ con il prodotto scalare standard, e che $f(X)$ sia una forma quadratica reale in n variabili. Assumendo la validità del Teorema spettrale reale, esiste un cambiamento ON di coordinate $X = M \cdot Y$ tale che

$$g(Y) := f(M \cdot Y) = \sum_{i=1}^{n} \lambda_i y_i^2. \tag{9.2.11}$$

Ora chiediamoci: come facciamo a determinare le nuove (buone) coordinate Y a partire da $f(X)$? Questo equivale a chiederci di determinare i vettori le cui coordinate Y sono i vettori $e_1, \dots, e_n$ (ma che avranno tutt'altre X coordinate!). La risposta la si ottiene esaminando la forma quadratica diagonale in (9.2.11). Ciascuno degli $e_1, \dots, e_n$ è un punto critico della funzione

$$\begin{array}{ccc} \mathbb{R}^n \backslash \{0\} & \xrightarrow{\rho} & \mathbb{R} \\ X & \mapsto & \frac{f(X)}{||X||^2} \end{array}$$

(E anche qualsiasi multiplo non nullo degli e_i, giachhè ρ è omogenea.) Punto critico significa che la derivata direzionale di ρ in un tale punto è nulla qualsiasi sia la direzione. Un modo geometrico di convincerci che ciascuno degli $e_1, \dots, e_n$ è un punto critico della funzione ρ consiste nel contemplare la "superficie di dimensione $(n-1)$"

$$E := \{X \in \mathbb{R}^n \mid f(X) = 1\}$$

(se $f \neq 0$ esiste $X \in \mathbb{R}^n$ tale che $f(X) \neq 0$, se $f(X) > 0$ segue che $E \neq \varnothing$, se $f(X) < 0$ sostituiamo f con $-f$). Il Teorema spettrale reale ci dice che rispetto a un sistema di coordinate ON Y (con origine in $(0, \dots, 0)$) l'equazione di E è $\sum_{i=1}^n \lambda_i y_i^2 = 1$. Quindi ciascuna delle rette generate dai vettori che hanno coordinate $e_1, \dots, e_n$ è di simmetria per E (un fatto non banale, niente affatto chiaro a priori), e da questo segue che ρ ha un punto critico in ciascuno di questi punti. Ora, siccome un massimo (se esiste) di ρ è un punto critico, vediamo che dimostrare che ρ ha un massimo ci porterà a dimostrare il Teorema spettrale reale.

Dimostreremo alcuni risultati preliminari, e alla fine della sottosezione daremo la dimostrazione del Teorema spettrale reale.

Proposizione 9.2.12. *Sia $A\colon V \to V$ un endomorfismo simmetrico. Un vettore non nullo $v \in V$ è un autovettore di A se e solo se*

$$v^\perp \subset \{w \in V \mid \langle A(v), w\rangle = 0\}, \tag{9.2.12}$$

dove $v^\perp$ è l'ortogonale di v per il prodotto scalare (notate che il membro di destra di (9.2.12) *è l'ortogonale di v per la forma bilineare simmetrica Φ_A).*

Dimostrazione. Supponiamo che v sia un autovettore di A, con autovalore $\lambda \in \mathbb{R}$. Se $w \in v^\perp$, allora

$$\langle A(v), w\rangle = \lambda \langle v, w\rangle = 0.$$

Ora supponiamo che valga l'uguaglianza in (9.2.12), e dimostriamo che v è un autovettore di A. Siccome l'ortogonale (per il prodotto scalare) di $v^\perp$ è generato da v, è sufficiente dimostrare che $A(v)$ è ortogonale a $v^\perp$. Ma questo è ciò che è scritto nell'uguaglianza di (9.2.12). □

Corollario 9.2.13. *Sia $A\colon V \to V$ un endomorfismo simmetrico. Se $v \in V$ è un autovettore di A, allora $A(v^\perp) \subset v^\perp$.*

Dimostrazione. Se $w \in v^\perp$, allora per la Proposizione 9.2.12 si ha $\langle A(v), w\rangle = 0$, e per simmetria di A segue che $\langle v, A(w)\rangle = 0$. □

Per il Corollario 9.2.13 il punto cruciale da dimostrare è che esiste un autovettore v di A. Infatti, una volta dimostrato ciò, considerando la restrizione di A a $v^\perp$ si procede per induzione sulla dimensione di V, come mostreremo. Per il Teorema fondamentale dell'Algebra esiste un autovalore complesso di A. Si può dimostrare che tale autovalore è reale, e questo finisce la dimostrazione del punto cruciale. Noi diamo una dimostrazione che non si basa sul Teorema fondamentale dell'Algebra[1], e che è di ispirazione geometrica.

[1] Probabilmente non avete mai visto una dimostrazione del T.F.A.

Proposizione 9.2.14. *Sia $A\colon V \to V$ un endomorfismo simmetrico. Esiste un vettore non nullo $v \in V$ tale che valga l'uguaglianza in* (9.2.12).

Dimostrazione. Poniamo $f := q_A$, dove $q_A \in Q(V)$ è la forma quadratica associata ad A, vedi (9.2.7). Consideriamo l'applicazione

$$\begin{array}{ccc} V\backslash\{0\} & \xrightarrow{\rho} & \mathbb{R} \\ v & \mapsto & f(v)/||v||^2 \end{array}$$

Dimostriamo che ρ ammettte massimo. Scegliendo una isometria di V con $\mathbb{R}^n$ (con prodotto euclideo standard) ci riduciamo al caso in cui V è $\mathbb{R}^n$ con il prodotto euclideo standard. Sia $F_n \subset \mathbb{R}^n$ la frontiera dell'n cubo standard, cioè

$$F_n := \{X \in \mathbb{R}^n \mid |x_i| \leqslant 1 \text{ per ogni } i \text{ e } |x_{i_0}| = 1 \text{ per un } i_0 \text{ (almeno)}\}.$$

Dimostriamo che la restrizione di ρ a F_n ha un massimo. Il caso $n = 1$ è banale perchè F_1 è un insieme con due elementi. Ora consideriamo il caso $n = 2$. La funzione $\rho(X) = f(X)/||X||^2$ è continua su $\mathbb{R}^2\backslash\{(0,0)\}$ (una funzione $\varphi\colon S \to \mathbb{R}$ con dominio un sottoinsieme $S \subset \mathbb{R}^n$ è continua se, dato un qualsiasi $\overline{X} \in S$ e un qualsiasi $\epsilon > 0$, esiste $\delta > 0$ tale che $||\varphi(X) - \varphi(\overline{X})|| < \epsilon$ per ogni $X \in S$ tale che $||X - \overline{X}|| < \delta$) e quindi anche la sua restrizione a F_2 lo è. Ma F_2 è l'unione di 4 segmenti chiusi e limitati, e quindi la restrizione di ρ a F_2 ammette massimo per il Teorema di Bolzano-Weierstrass. Un analogo ragionamento dà che la restrizione di ρ a F_n ammette massimo per ogni n - va usato l'analogo di Bolzano-Weierstrass in dimensione arbitraria: se f è una funzione continua da $[a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n] \subset \mathbb{R}^n$ a $\mathbb{R}$ allora esiste un massimo di f (potete darne una dimostrazione analoga alla classica dimostrazione di Bolzano-Weierstrass per bisezioni successive, in dimenione n si fanno 2^n-sezioni successive) Ora sia $\overline{X} \in F_n$ tale che la restrizione di ρ a F_n ha un massimo in $\overline{X}$; allora $f(\overline{X})$ è il massimo di ρ. Infatti sia $X \in (\mathbb{R}^n\backslash\{0\})$, e sia m il massimo tra $|x_1|, |x_2|, \ldots, |x_n|$. Notate che $m > 0$ perchè $X \neq 0$, e che $m^{-1}X \in F_n$ perchè ogni coordinata di $m^{-1}X$ ha valore assoluto al più 1, ed esiste $i \in \{1, \ldots, n\}$ tale che $|x_i| = m$, e perciò la i-esima coordinata di $m^{-1}X$ è uguale a ± 1. Allora

$$\rho(X) = f(X)/||X||^2 = m^2 f(m^{-1}X)/m^2||m^{-1}X||^2 = f(m^{-1}X)/||m^{-1}X||^2 = \rho(m^{-1}X) \leqslant \rho(\overline{X}).$$

Questo dimostra che ρ ammettte massimo.

Sia $v \in V\backslash\{0\}$ tale che $f(v)$ sia il massimo della funzione ρ. Dimostriamo che vale l'uguaglianza in (9.2.12). Sia $w \in v^\perp$ dove l'ortogonalità è rispetto al prodotto euclideo. Definiamo le funzioni $g, h, \xi\colon \mathbb{R} \to \mathbb{R}$ così:

$$g(s) := f(v + sw) = \langle A(v + sw), v + sw\rangle, \quad h(s) = ||v + sw||^2 = \langle v + sw, v + sw\rangle, \quad \xi(s) := \rho(v + sw) = \frac{g(s)}{h(s)}.$$

Notate che ξ è un quoziente di funzioni differenziabili (e il quoziente non è mai nullo), e quindi è differenziabile. Siccome ξ ha un massimo per $s = 0$, segue che

$$0 = \xi'(0) = \frac{g'(0)h(0) - g(0)h'(0)}{h(0)^2}. \tag{9.2.13}$$

Ora

$$g'(0) = \left.\frac{d}{ds}\right|_{s=0} (\langle A(v + sw), v + sw\rangle) = 2\langle A(v), w\rangle, \quad h'(0) = \left.\frac{d}{ds}\right|_{s=0} (\langle v + sw, v + sw\rangle) = 2\langle v, w\rangle = 0.$$

Quindi l'uguaglianza in (9.2.13) dà che $\langle A(v), w\rangle = 0$. Siccome w è un arbitrario vettore ortogonale a v, questo dimostra che vale l'uguaglianza in (9.2.12). □

Dimostrazione del Teorema spettrale reale. Per induzione sulla dimensione di V. Se $\dim V = 1$ l'enunciato è banalmente vero, ogni applicazione lineare è diagonale in qualsiasi base. Dimostriamo il passo induttivo. Per le Proposizioni 9.2.14 e 9.2.12 esiste un autovettore v di A, che possiamo assumere di norma 1 (se non lo è, lo normalizziamo). Per il Corollario 9.2.13 si ha $A(v^\perp) \subset v^\perp$, cioè A definisce un'applicazione lineare

$$\begin{array}{ccc} v^\perp & \xrightarrow{B} & v^\perp \\ w & \mapsto & A(w) \end{array}$$

Il sottospazio $v^\perp \subset V$ "eredita" il prodotto scalare di V, e cosìè anch'esso uno spazio vettoriale euclideo. È chiaro che B è un endomorfismo simmetrico di $v^\perp$. Per ipotesi induttiva esiste una base ortonormale $\mathcal{C}$ di $v^\perp$ che diagonalizza B. Aggiungendo ai vettori di $\mathcal{C}$ il vettore v, otteniamo una base ortonormale di V che diagonalizza A.

□

Implementazione del Teorema spettrale

Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. Sia A un endomorfismo simmetrico di V, o equivalentemente una forma quadratica f su V. Quello che abbiamo dimostrato in questa sezione dà il seguente algoritmo per produrree una base ON che diagonalizza A (o equivalentemente f):

1. Si sceglie una qualsiasi base ON $\mathcal{C}$ di V, e si calcola $A = M_{\mathcal{C}}(f)$.
2. Si calcolano gli autovalori di A (per il Teorema spettrale sono tutti reali), cioè le radici[2] $\lambda_1, \ldots, \lambda_m$ del polinomio caratteristico P_A.
3. Per ogni radice λ_i del polinomio caratteristico P_A, si determina una base ON dell'autospazio $V_{\lambda_i}(A)$. In questo modo si producono d_i vettori $v_{i,1}, \ldots, v_{i,d_i}$.
4. La base $\mathcal{B} := \{v_{1,1}, \ldots, v_{1,d_1}, \ldots, v_{m,1}, \ldots, v_{m,d_m}\}$ ottenuta riunendo i vettori del punto (3) è una base ON di V che diagonalizza A.

Diamo un esempio. Sia $A \in M_{3,3}(\mathbb{R})$ la matrice simmetrica definita da

$$A := \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}.$$

Per trovare una base ON che diagonalizza A, calcoliamo il polinomio caratteristico

$$p_A(\lambda); = \mathrm{Det}(\lambda 1_3 - A) = \lambda^3 - 6\lambda^2 - 3\lambda + 18.$$

Le radici di p_A sono 6, $\sqrt{3}$ e $-\sqrt{3}$. Con un calcolo semplice (ma noioso) troviamo che gli autospazi di A sono

$$\begin{aligned} V_6(A) &= \langle (1,1,1) \rangle, \\ V_{\sqrt{3}}(A) &= \langle (7 - 3\sqrt{3}, -5 - \sqrt{3}, -2 + 4\sqrt{3}) \rangle, \\ V_{-\sqrt{3}}(A) &= \langle (7 + 3\sqrt{3}, -5 + \sqrt{3}, -2 - 4\sqrt{3}) \rangle. \end{aligned}$$

Quindi una base ortogonale che diagonalizza f è

$$\mathcal{C} = \{(1,1,1),\ (7 - 3\sqrt{3}, -5 - \sqrt{3}, -2 + 4\sqrt{3}),\ (7 + 3\sqrt{3}, -5 + \sqrt{3}, -2 - 4\sqrt{3})\},$$

e una base ON che diagonalizza f si ottiene normalizzando ciascun vettore di $\mathcal{C}$.

In generale la procedura appena descritta non ha alcun valore pratico, per via del punto (2), e si usano altri metodi.

Il Teorema spettrale e classificazione di orbite

Il gruppo ortogonale $\mathrm{O}(V)$ agisce su $\mathrm{End}^+(V)$ e su $Q(V)$, vedi l'Osservazione 9.2.10, e possiamo considerare le relazioni di equivalenza su $\mathrm{End}^+(V)$ e su $Q(V)$ associate a questa azione. Esplicitamente: se $V = \mathbb{R}^n$ con il prodotto euclideo standard, allora $\mathrm{End}^+(V)$ e $Q(V)$ sono identificati con lo spazio delle matrici simmetriche $M^+_{n,n}(\mathbb{R})$ e $A, B \in M^+_{n,n}(\mathbb{R})$ sono equivalenti se esiste $G \in \mathrm{O}_n(\mathbb{R})$ tale che $G \cdot A \cdot G^{-1} = B$ (ma notate che $G^{-1} = G^t$, quindi questo equivale a $G \cdot A \cdot G^t = B$). Il Teorema spettrale reale dà che $A, B \in M^+_{n,n}(\mathbb{R})$ sono equivalenti se e solo se i loro polinomi caratteristici sono uguali.

9.3 Forme hermitiane

Per tutta la sezione V è uno spazio vettoriale complesso.

Definizione e primi risultati

Definizione 9.3.1. Una *forma sesquilineare su* V è un'applicazione

$$\begin{array}{ccc} V \times V & \xrightarrow{H} & \mathbb{C} \\ (v, w) & \mapsto & H(v, w) \end{array} \tag{9.3.1}$$

che è lineare nella prima variabile e coniugato-lineare nella seconda a variabile, cioè tale che

[2] Questo è il punto dolente: in generale possiamo solo approssimarle

1. per $w_0 \in V$ fissato, la funzione $V \to \mathbb{C}$ definita $v \mapsto H(v, w_0)$ è lineare, e
2. per $v_0 \in V$ fissato, la funzione $V \to \mathbb{C}$ definita $w \mapsto \overline{H(v_0, w)}$ è lineare.

Osservazione 9.3.2. La condizione (2) della Definizione 9.3.1 equivale alle due condizioni seguenti:

(a) Per $v_0, w_1, w_2 \in V$ si ha
$$H(v_0, w_1 + w_2) = H(v_0, w_1) + H(v_0 w_2).$$

(b) Per $v_0, w \in V$ e $\lambda \in \mathbb{C}$ si ha
$$H(v_0, \lambda w) = \overline{\lambda} H(v_0, w).$$

Infatti l'applicazione $w \mapsto \overline{H(v_0, w)}$ è lineare se e solo se

$$\overline{H(v_0, w_1 + w_2)} = \overline{H(v_0, w_1)} + \overline{H(v_0, w_2)} == \overline{H(v_0, w_1) + H(v_0, w_2)},$$

e

$$\overline{H(v_0, \lambda w)} = \lambda \overline{H(v_0, w)} = \overline{\overline{\lambda} H(v_0, w)},$$

cioè valgono rispettivamente (a) e (b).

Esempio 9.3.3. Sia $A \in M_{n,n}(\mathbb{C})$, e sia

$$\begin{array}{ccc} \mathbb{C}^n \times \mathbb{C}^n & \xrightarrow{\Psi_A} & \mathbb{C} \\ (X, Y) & \mapsto & X^t \cdot A \cdot \overline{Y} \end{array} \tag{9.3.2}$$

La Ψ_A è una forma sesquilineare. Inoltre, se H è una forma sesquilineare su $\mathbb{C}^n$, esiste $A \in M_{n,n}(\mathbb{C})$ tale che $H = \Psi_A$, e tale A è unica. Infatti, se $\{e_1, \ldots, e_n\}$ è la base standard di $\mathbb{C}^n$, allora

$$\Psi_A(e_j, e_k) = a_{jk}, \tag{9.3.3}$$

dove $A = (a_{jk})$. (Notate: quando trattiamo spazi vettoriale complessi è bene evitare di usare la lettera "i" come indice.) Questo dimostra che se $H = \Psi_A$ per un $A \in M_{n,n}(\mathbb{C})$, allora A è unico. Ma un semplice ragionamento mostra che ponendo $A = (a_{jk})$, dove gli a_{jk} sono dati da (9.3.3), vale $H = \Psi_A$.

Definizione 9.3.4. Una *forma hermitiana su* V (da Charles Hermite) è una forma sesquilineare H su V tale che
$$H(w, v) = \overline{H(v, w)} \quad \forall v, w \in V.$$

Esempio 9.3.5. Un prodotto scalare hermitiano (vedi la Definizione 6.9.1) è una forma hermitiana.

Osservazione 9.3.6. Una maniera utile di pensare alle forme sesquilineari ed hermitiane su uno spazio vettoriale complesso è come l'analogo delle forme bilineari e simmetriche su uno spazio vettoriale reale rispettivamente. Proseguendo nell'analogia, ai prodotti scalari hermitiani corrispondono i prodotti scalari euclidei.

Osservazione 9.3.7. Se H è una forma hermitiana su V, allora

$$H(v, v) \in \mathbb{R} \quad \forall v \in V.$$

Infatti, siccome H è hermitiana, $H(v, v) = \overline{H(v, v)}$, e quindi $H(v, v)$ è reale.

La teoria delle forme hermitiane su uno spazio vettoriale complesso V è del tutto simile alla teoria delle forme bilineari simmetriche su uno spazio vettoriale reale. Daremo i principali risultati. Molte dimostrazioni saranno solo accennate perchè sono del tutto simili a quelle date nel caso reale.

Definizione 9.3.8. $\mathrm{Herm}(V)$ è l'insieme delle forme hermitiane su V. Se $H_1, H_2, H \in \mathrm{Herm}(V)$ e $\lambda \in \mathbb{R}$ definiamo

$$\begin{array}{ccc} V \times V & \xrightarrow{H_1 + H_2} & \mathbb{C} \\ (v, w) & \mapsto & H_1(v, w) + H_2(v, w) \end{array}$$

e

$$\begin{array}{ccc} V \times V & \xrightarrow{\lambda H} & \mathbb{C} \\ (v, w) & \mapsto & \lambda H(v, w) \end{array} \tag{9.3.4}$$

È facile verificare che $H_1 + H_2$ e λH sono forme hermitiane. Notate che se nell'equazione (9.3.4) lo scalare λ è numero complesso non reale, allora la forma sesquilineare definita da (9.3.4) non è hermitiana. Inoltre con queste operazioni $\mathrm{Herm}(V)$ è uno spazio vettoriale reale.

Diamo l'analogo della corrispondenza che c'è tra forme bilineari simmetriche su uno spazio vettoriale e matrici simmetriche, una volta scelta una base dello spazio vettoriale. Prima va definito l'analogo di "matrice simmetrica".

Definizione 9.3.9. L'*aggiunta* di una matrice $A \in M_{n,n}(\mathbb{C})$ è la matrice

$$A^* := \overline{A}^t.$$

Inoltre $A \in M_{n,n}(\mathbb{C})$ è *hermitiana* (o *autoaggiunta*) se è uguale alla sua aggiunta.

Osservazione 9.3.10. Sia $A \in M_{n,n}(\mathbb{C})$. La forma sesquilineare Ψ_A definita in (9.3.2) è una forma hermitiana se e solo se $A = A^*$. Infatti

$$\Psi_A(Y,X) = Y^t \cdot A \cdot \overline{X} = (Y^t \cdot A \cdot \overline{X})^t = \ = \overline{X}^t \cdot A^t \cdot Y = \overline{X^t \cdot \overline{A}^t \cdot \overline{Y}} = \overline{\Psi_{A^*}(X,Y)}.$$

(La seconda uguaglianza vale perchè ogni matrice 1×1 è simmetrica.) Quindi Ψ_A è hermitiana se e solo se $\Psi_{A^*}(X,Y) = \Psi_A(X,Y)$ per ogni $X, Y \in \mathbb{C}^n$. Per (9.3.3) questo vale se e solo se $A = A^*$.

Definizione 9.3.11. $M^*_{n,n}(\mathbb{C}) \subset M_{n,n}(\mathbb{C})$ è il sottoinsieme delle matrici autoaggiunte.

Osservazione 9.3.12. $M^*_{n,n}(\mathbb{C})$ è un sottospazio vettoriale *reale* di $M_{n,n}(\mathbb{C})$. (Non è un sottospazio vettoriale complesso perchè se A è autoaggiunta, allora $a_{jk} = \overline{a_{kj}}$, e quindi iA è autoaggiunta solo se $A = 0$.) Inoltre la diagonale di una matrice autoaggiunta ha entrate reali perchè per definizione $a_{jj} = \overline{a_{jj}}$ e quindi a_{jj} è reale.

Supponiamo che V abbia dimensione finita n, e sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una sua base.

Definizione 9.3.13. Sia H una forma hermitiana su V. La matrice $M_{\mathcal{B}}(H) \in M_{n,n}(\mathbb{C})$ è la matrice con entrata su riga j e colonna k data da $H(v_j, v_k)$.

Osservazione 9.3.14. La matrice $M_{\mathcal{B}}(H) \in M_{n,n}(\mathbb{C})$ associata a una forma hermitiana H su V è autoaggiunta, cioè è un elemento di $M^*_{n,n}(\mathbb{C})$.

Esempio 9.3.15. Sia $A \in M_{n,n}(\mathbb{C})$ una matrice autoaggiunta, e $\Psi_A \in \mathrm{Herm}(\mathbb{C}^n)$ la forme hermitiana data da $\Psi_A(X,Y) = X^t \cdot A \cdot \overline{Y}$. Se $\mathcal{S}$ è la base standard di $\mathbb{C}^n$, allora $M_{\mathcal{S}}(\Psi_A) = A$.

Lasciamo al lettore la (facile) dimostrazione del seguente risultato.

Proposizione 9.3.16. *Sia V uno spazio vettoriale complesso di dimensione finita n, e sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una sua base. L'applicazione*

$$\begin{array}{ccc} \mathrm{Herm}(V) & \xrightarrow{M_{\mathcal{B}}} & M^*_{n,n}(\mathbb{C}) \\ H & \mapsto & M_{\mathcal{B}}(H) \end{array}$$

è un isomorfismo di spazi vettoriali reali.

Forme hermitiane a meno di equivalenza

Sia $H \in \mathrm{Herm}(V)$. Se $g \in \mathrm{GL}(V)$, allora l'applicazione definita da

$$\begin{array}{ccc} V \times V & \xrightarrow{gH} & \mathbb{C} \\ (v,w) & \mapsto & H(g^{-1}(v), g^{-1}(w)) \end{array}$$

è una forma hermitiana. Inoltre

1. $\mathrm{Id}_V H = H$,
2. $H = g^{-1}(gH)$, e
3. se $g_1, g_2 \in \mathrm{GL}(V)$, allora $g_1(g_2 H) = (g_1 g_2)H$.

In altre parole l'applicazione $\mathrm{GL}(V) \times \mathrm{Herm}(V) \to \mathrm{Herm}(V)$ definita assegnando gH alla coppia (g, H) è un'azione di $\mathrm{GL}(V)$ su $\mathrm{Herm}(V)$. Segue che la relazione definita sotto è di equivalenza.

Definizione 9.3.17. $H_1, H_2 \in \mathrm{Herm}(V)$ sono *equivalenti* se esiste $g \in \mathrm{GL}(V)$ tale che $H_1 = gH_2$.

Se V ha dimensione finita l'equivalenza tra forma hermitiane si traduce in una equivalenza tra matrici autoaggiunte.

Proposizione 9.3.18. *Sia $H \in \mathrm{Herm}(V)$ e $g \in \mathrm{GL}(V)$. Sia $\mathcal{B}$ una base di V. Sia $G := M_{\mathcal{B}}^{\mathcal{B}}(g^{-1})$. Allora*

$$M_{\mathcal{B}}(gH) = (G^{-1})^t \cdot M_{\mathcal{B}}(H) \cdot \overline{G^{-1}}. \tag{9.3.5}$$

Dimostrazione. Segue dalle equazioni

$$gH(v,w) = H(g^{-1}v, g^{-1}w) = X_{\mathcal{B}}(g^{-1}(v))^t \cdot M_{\mathcal{B}}(H) \cdot \overline{X_{\mathcal{B}}(g^{-1}(v))} = X_{\mathcal{B}}(v)^t \cdot (G^{-1})^t \cdot M_{\mathcal{B}}(H) \cdot \overline{G^{-1}} \cdot \overline{X_{\mathcal{B}}(w)}.$$

□

Motivati dalla proposizione precedente, notiamo che se $A \in M_{n,n}^*(\mathbb{C})$ e $G \in \mathrm{GL}_n(\mathbb{C})$, allora

$$G^t \cdot A \cdot \overline{G} \in M_{n,n}^*(\mathbb{C}).$$

(Per verificarlo notate che se $A, B \in M_{n,n}(\mathbb{C})$, allora $(A \cdot B)^* = B^* \cdot A^*$.)

Definizione 9.3.19. Definizione: $A_1, A_2 \in M_{n,n}^*(\mathbb{C})$ sono *equivalenti* se esiste $G \in \mathrm{GL}_n(\mathbb{C})$ tale che

$$A_1 = G^t \cdot A_2 \cdot \overline{G}.$$

La conclusione di ciò è che classificare forme hermitiane a meno di equivalenza equivale a classificare matrici hermitiane a meno di equivalenza.

La classificazione delle forme hermitiane a meno di equivalenza è del tutto analoga a quella delle forme bilineari simmetriche reali.

Definizione 9.3.20. Una $H \in \mathrm{Herm}(V)$ è *definita positiva* se $H(v,v) > 0$ per $v \in V$ non nullo (equivalentemente se definisce un prodotto hermitiano), ed è *definita negativa* se $H(v,v) < 0$ per $v \in V$ non nullo.

Definizione 9.3.21. Sia $H \in \mathrm{Herm}(V)$. La *segnatura positiva di H* (che denotiamo $s_+(H)$) è la massima dimensione di un sottospazio $U \subset V$ tale che la restrizione di H a U sia definita positiva, e la *segnatura negativa di H* (che denotiamo $s_-(H)$) è la massima dimensione di un sottospazio $U \subset V$ tale che la restrizione di H a U sia definita negativa.

Proposizione 9.3.22. *$H_1, H_2 \in \mathrm{Herm}(V)$ sono equivalenti se e solo se*

$$s_+(H_1) = s_+(H_2), \quad s_-(H_1) = s_-(H_2).$$

Dimostrazione. La dimostrazione è del tutto simile alla dimostrazione dell'analogo risultato per forme bilineari simmetriche su uno spazio vettoriale reale di dimensione finita.
Passo 1. Se $H \in \mathrm{Herm}(V)$ esiste una base $\mathcal{B}$ che diagonalizza H, cioè tale che $M_{\mathcal{B}}(H)$ sia diagonale. La dimostrazione è tale e quale a quella nel caso reale. Se $H = 0$ non c'è nulla da dimostrare, quindi supponimao che $H \neq 0$. Ora osserviamo che se $H \neq 0$ esiste $v \in V$ tale che $H(v,v) \neq 0$. Infatti, siccome $H \neq 0$ esistono $v_1, v_2 \in V$ tali che $H(v_1, v_2) \neq 0$. Riscalando v_1 (o v_2) possiamo assumere che $H(v_1, v_2)$ sia reale e non nullo. Segue che la funzione polinomiale

$$\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ t & \mapsto & H(v_1 + tv_2, v_1 + tv_2) \end{array}$$

non è nulla e quindi esiste $t_0 \in \mathbb{R}$ tale che $H(v_1 + t_0 v_2, v_1 + t_0 v_2) \neq 0$. Abbiamo dimostrato che esiste $v \in V$ tale che $H(v,v) \neq 0$. Per $v \in V$ definiamo l'ortogonale al solito modo:

$$v^\perp := \{w \in V \mid H(v,w) = 0\} = \{w \in V \mid H(w,v) = 0\}.$$

Ora sia $v \in V$ tale che $H(v,v) \neq 0$ (stiamo supponendo che $H \neq 0$); segue che

$$V = \mathrm{Span}(v) \oplus v^\perp. \tag{9.3.6}$$

Iterando vediamo che esiste una base $\mathcal{B}$ che diagonalizza H.
Passo 2. Supponiamo che $H \in \mathrm{Herm}(V)$ e che $M_{\mathcal{B}}(H)$ sia diagonale. Notate che le entrate di $M_{\mathcal{B}}(H)$ sono reali. Infatti fuori dalla diagonale sono nulle perchè è diagonale, e sulla diagonale sono reali per l'Osservazione 9.3.12. Si dimostra, come nel caso reale, che $s_+(H)$ è il numero di entrate positive della diagonale di $M_{\mathcal{B}}(H)$ positiva di H, e $s_-(H)$ è il numero di entrate negative della diagonale di $M_{\mathcal{B}}(H)$. Questo dimostra la Proposizione. □

9.4 Il teorema spettrale complesso

In questa sezione $(V, \langle , \rangle)$ è uno spazio vettoriale hermitiano. Prima di enunciare il teorema spettrale complesso discutiamo l'aggiunto di un operatore lineare, cioè un endomorfismo di V.

Definizione 9.4.1. Sia $(V, \langle , \rangle)$ uno spazio vettoriale hermitiano. Sia $T\colon V \to V$ un endomorfismo (o operatore lineare). Un operatore lineare $S\colon V \to V$ è *un aggiunto* di T se per ogni $v, w \in V$ vale

$$\langle T(v),\, w\rangle = \langle v,\, S(w)\rangle.$$

Lemma 9.4.2. *Sia $T\colon V \to V$ un operatore lineare, dove $(V, \langle , \rangle)$ è uno spazio vettoriale hermitiano finitamente generato. Esiste uno e un solo operatore aggiunto di T, che denotiamo T^*. Se $\mathcal{B}$ è una base ON di V, allora*

$$M_{\mathcal{B}}^{\mathcal{B}}(T^*) = M_{\mathcal{B}}^{\mathcal{B}}(T)^*. \tag{9.4.1}$$

Dimostrazione. Se $v, w \in V$, allora

$$\langle T(v), w\rangle = (M_{\mathcal{B}}^{\mathcal{B}}(T) \cdot X_{\mathcal{B}}(v))^t \cdot \overline{X}_{\mathcal{B}}(w) = X_{\mathcal{B}}(v)^t \cdot M_{\mathcal{B}}^{\mathcal{B}}(T)^t \cdot \overline{X}_{\mathcal{B}}(w) = X_{\mathcal{B}}(v)^t \cdot \overline{\left(M_{\mathcal{B}}^{\mathcal{B}}(T)^* \cdot X_{\mathcal{B}}(w)\right)}.$$

Sia $T^*\colon V \to V$ l'operatore lineare tale che valga (9.4.1). I conti appena fatti dimostrano che

$$\langle T(v),\, w\rangle = \langle v,\, T^*(w)\rangle,$$

e anche che T^* è l'unico aggiunto di T. □

Definizione 9.4.3. Un operatore lineare $T\colon V \to V$ è *autoaggiunta(o)* se $T = T^*$.

Osservazione 9.4.4. Siano $T\colon V \to V$ un operatore lineare e $\lambda \in \mathbb{C}$. Allora

$$(\lambda T)^* = \overline{\lambda} T^*.$$

Se $T_1, T_2\colon V \to V$ sono operatori lineare, allora

$$(T_1 + T_2)^* = T_1^* + T_2^*.$$

Segue che l'insieme degli operatori autoaggiunti di V è un sottospazio *reale* di $\mathrm{End}(V)$.

Osservazione 9.4.5. Consideriamo $\mathbb{C}^n$ con il prodotto hermitiano standard

$$\langle X,\, Y\rangle = X^t \cdot \overline{Y}.$$

Se $A \in M_{n,n}(\mathbb{C})$, l'aggiunto di L_A è L_{A^*}, dove $A^* = \overline{A}^t$ è l'aggiunta della matrice A, perchè la base standard di $\mathbb{C}^n$ è ON per il prodotto hermitiano standard.

In particolare il sottospazio reale degli operatori autoaggiunti è $M_{n,n}^*(\mathbb{C})$, e un operatore reale è autoaggiunto se e solo se è simmetrico.

Teorema 9.4.6 (Teorema spettrale per operatori autoaggiunti)**.** *Sia $(V, \langle\ ,\ \rangle)$ uno spazio vettoriale hermitiano finitamente generato. Sia $T\colon V \to V$ un operatore autoaggiunto, cioè tale che*

$$\langle T(v),\, w\rangle = \langle v,\, T(w)\rangle \quad \forall v, w \in V. \tag{9.4.2}$$

Allora T ha tutti gli autovalori reali, ed esiste una base ON di V che lo diagonalizza.

Dimostrazione. Iniziamo dimostrando che gli autovalori sono tutti reali. Sia λ un autovalore di T, e sia v un autovettore associato. Per (9.4.2) abbiamo

$$\lambda ||v||^2 = \langle T(v),\, v\rangle = \langle v,\, T(v)\rangle = \overline{\lambda} ||v||^2.$$

Siccome $v \neq 0$ si ha $||v|| > 0$, e quindi $\lambda = \overline{\lambda}$, cioè λ è reale.

Ora dimostriamo per induzione su $\dim V$ che esiste una base ON di V che diagonalizza T. Se $\dim V = 1$ l'affermazione è banalmente vera. Dimostriamo il passo induttivo. Procediamo come nella dimostrazione della Proposizione 6.9.15. Esiste un autovalore λ di T, sia v un autovettore associato. Normalizzando v possiamo assumere che $||v|| = 1$. Dimostriamo che $T(v^\perp) \subset v^\perp$: se $w \in v^\perp$,

$$\langle T(w),\, v\rangle = \langle w,\, T(v)\rangle = \langle w,\, \lambda v\rangle = \overline{\lambda}\langle w,\, v\rangle = 0.$$

Quindi la restrizione di T a $v^\perp$ definisce un operatore

$$\begin{array}{ccc} v^\perp & \xrightarrow{S} & v^\perp \\ w & \mapsto & T(w) \end{array}$$

che è autoaggiunto perchè lo è T. Per ipotesi induttiva esiste una base ON $\mathcal{C}$ di $v^\perp$ che diagonalizza S. Aggiungendo a $\mathcal{C}$ il vettore v otteniamo una base ON di V (perchè v è ortogonale a ogni vettore di $\mathcal{C}$, e $||v|| = 1$) che diagonalizza T. □

Osservazione 9.4.7. Nell'analogia tra spazi vettoriali euclidei e spazi vettoriali hermitiani, gli endomorfismi simmetrici corrispondono a endomorfismi (operatori) autoaggiunti. Abbiamo visto, vedi la Proposizione 9.2.8, che c'è una corrispondenza naturale tra endomorfismi simmetrici e forme quadratiche. Esiste l'analogo per spazi vettoriali hermitiani: se A è un operatore autoaggiunto su uno spazio vettoriale hermitiano (finitamente generato) $(V, \langle , \rangle)$, l'applicazione

$$\begin{array}{ccc} V \times V & \xrightarrow{H_A} & \mathbb{C} \\ (v, w) & \mapsto & \langle A(v), w\rangle \end{array}$$

è hermitiana e viceversa, data una forma hermitiana H su V, esiste un (unico) operatore autoaggiunto A tale che $H = H_A$. Questo segue dall'Osservazione 9.3.10 scegliendo una base ON di V.

Il Teorema spettrale complesso, insieme all'Osservazione 9.4.7, dà il seguente analogo del Teorema 9.2.9.

Teorema 9.4.8. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale hermitiano (finitamente generato), e sia H una forma hermitiana su V. Esiste una base ON di V che diagonalizza H.*

Esercizi del Capitolo 9

Esercizio 9.1. *Sia f la forma quadratica su $\mathbb{R}^2$ data da*

$$f(x_1, x_2) = x_1^2 + 2x_1x_2 + 3x_2^2.$$

Trovate una base di $\mathbb{R}^2$, ON per il prodotto euclideo standard, che diagonalizza f.

Esercizio 9.2. *Sia $A \in M_{2,2}(\mathbb{R})$ data da*

$$A := \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$$

1. *Verificate che $\langle X, Y\rangle := X^t \cdot A \cdot Y$ è un prodotto euclideo su $\mathbb{R}^2$.*
2. *Sia f la forma quadratica su $\mathbb{R}^2$ definita da $f(x_1, x_2) = x_1^2 + 2x_1x_2 - 3x_2^2$. Trovate una base di $\mathbb{R}^2$ che diagonalizza f e che è ON per il prodotto euclideo del punto (1)*

Esercizio 9.3. *Sia V uno spazio vettoriale complesso di dimensione finita e sia $H \in \mathrm{Herm}(V)$ una forma hermitiana su V. Dimostrate che*

$$\max\{U \subset V \mid H_{|U} = 0\} = \dim\ker(H) + \min\{s_+(H),\, s_-(H)\}.$$

(Qui $H_{|U}$ è la forma hermitian su U definita da $H(u_1, u_2)$ per $u_1, u_2 \in U$.) In altre parole, si tratta di dimostrare che vale l'analogo dell'Esercizio 8.7.

Esercizio 9.4. *Sia $A \in M^*_{n,n}(\mathbb{C})$ una matrice hermitiana. Dimostrate che A è definita positiva (cioè la forma hermitiana associata H_A è definita positiva) se e solo se*

$$\det A(1) > 0, \quad \det A(2) > 0, \ldots, \det A(n) > 0,$$

dove, come al solito

$$A(1) = (a_{11}), \quad A(2) := \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \ldots$$

(Notate che, siccome $A(m) = A(m)^*$, si ha $\det A(m) \in \mathbb{R}$.) Analogamente A è definita negativa se e solo se

$$\det A(1) < 0, \quad \det A(2) > 0, \quad \det A(3) < 0, \ldots,$$

cioè se i determinanti si alternano, iniziando da $\det A(1) < 0$.

Esercizio 9.5. *Sia $A \in M^*_{3,3}(\mathbb{C})$ data da*

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

e sia H_A la forma hermitian su $\mathbb{C}^3$ definita da A, cioè

$$H_A(X,Y) = X^t \cdot A \cdot \overline{Y}. \tag{9.4.3}$$

Determinate segnatura positiva e negativa di H_A.

Esercizio 9.6. *Dimostrate che il polinomio caratteristico di un operatore autoaggiunto $(V, \langle , \rangle)$ su uno spazio vettoriale hermitiano (finitamente generato) ha tutti i coefficienti reali.*

Esercizio 9.7. *Date una dimostrazione del Teorema spettrale complesso analoga a quella data per il Teorema spettrale reale (quindi senza usare il Teorema fondamentale dell'Algebra), seguendo i seguenti passi.*

1. *Dimostrate che l'applicazione*
$$\begin{array}{ccccc} V\backslash\{0\} & \xrightarrow{f} & \mathbb{R}v & \mapsto & \frac{\langle Av, v\rangle}{||v||^2} \end{array} \tag{9.4.4}$$
ha un massimo.
2. *Sia $0 \neq v_0$ tale che $f(v_0)$ è il massimo di f. Dimostrate che*
$$v_0^{\perp} \subset \{w \in V \mid \langle Aw, v_0\rangle = 0\}.$$
(l'ortogonalità a sinistra è ripsetto a $\langle , \rangle$).
3. *Sia v_0 come nel punto (2): dimostrate che v_0 è un autovettore di A.*

Esercizio 9.8. *Osservate che Teorema spettrale complesso implica il Teorema spettrale reale.*

Esercizio 9.9. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale hermitiano hermitiano finitamente generato. Diciamo che due forme hermitiane $H_1, H_2 \in \mathrm{Herm}(V)$ sono* equivalenti *se esiste $g \in U(V)$ (ricordate: $U(V)$ è il gruppo unitario di $(V, \langle , \rangle)$) tale che*
$$H_1(v,w) = H_2(g^{-1}(v), g^{-1}(w)), \qquad \forall v, w \in V.$$
Classificate le classi di equivalenza come segue.

(a) *Data $H \in \mathrm{Herm}(V)$ definiamo il* polinomio caratteristico *di H come il polinomio caratteristico dell'unico operatore autoaggiunto $A\colon V \to V$ tale $H = H_A$ (vedi l'Osservazione 9.4.7).*

(b) *Dimostrate che esiste $H_1, H_2 \in \mathrm{Herm}(V)$ sono equivalenti se e solo se i loro polinomi caratteristici sono uguali.*

(c) *Siano $A, B \in M^*_{n,n}(\mathbb{C})$. Dimostrate che esiste $U \in U_n(\mathbb{C})$ tale che*
$$A = U^* \cdot B \cdot U$$
se e solo se $P_A(\lambda) = P_B(\lambda)$.

Esercizio 9.10. *Sia H una forma hermitiana su uno spazio vettoriale complesso V. Definiamo*

$$\begin{array}{ccc} V \times V & \xrightarrow{g_H} & \mathbb{R} \\ (v,w) & \mapsto & \frac{1}{2}(H(v,w) + \overline{H(v,w)}) \end{array}$$

(a) *Dimostrate che g_H è una forma bilineare simmetrica sullo spazio vettoriale* reale *V.*

(b) *Dimostrate che $g_H(iv, iw) = g_H(v, w)$ per $v, w \in V$, in particolare $g_H(v, iv) = 0$ per ogni $v \in V$.*

(c) *Ora sia $g\colon V \times V \to \mathbb{R}$ una forma bilineare simmetrica sullo spazio vettoriale* reale *V, e supponiamo che*
$$g(iv, iw) = g(v, w) \quad \forall v, w \in V.$$
Dimostrate che esiste $H \in \mathrm{Herm}(V)$ (ed è unica) tale che $g = g_H$.

Capitolo 10

Coniche, quadriche

10.1 Introduzione

10.2 Ellissi, iperboli, parabole

In questa sezione $\mathbb{S}$ è un piano affine euclideo, cioè uno spazio affine euclideo di dimensione 2.

Ellissi

Definizione 10.2.1. $\mathcal{C} \subset \mathbb{S}$ è un'*ellisse* se in un opportuno riferimento ON di coordinate (x, y) ha equazione

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \tag{10.2.1}$$

cioè $P(x, y) \in \mathcal{C}$ se e solo se vale (10.2.1). Si intende sempre $a > 0$ e $b > 0$. Spesso è sottinteso che $a \geqslant b$.

Esempio 10.2.2. Un cerchio di raggio r è un'ellisse con $a = b = r$.

Osservazione 10.2.3.

1. La retta delle x (di equazione $y = 0$) e la retta delle y (di equazione $x = 0$) sono assi di simmetria di $\mathcal{C}$, cioè la riflessione di $\mathcal{C}$ nelle due retta è uguale a $\mathcal{C}$ stessa. Se $a \neq b$ allora non esistono altri assi di simmetria di $\mathcal{C}$, se $a = b$, cioè $\mathcal{C}$ è un cerchio, allora ogni retta per l'origine è un asse di simmetria (e non ce ne sono altri).
2. L'origine è il *centro* di simmetria di $\mathcal{C}$, cioè la rotazione di centro l'origine e angolo π manda $\mathcal{C}$ in se stessa. Ho scritto il centro perchè non esistono altri centri di simmetria di $\mathcal{C}$ (se $F \subset \mathbb{S}$ ha due centri di simmetria distinti allora F è *illimitato*, cioè esistono coppie di punti di F che hanno distanza arbitrariamente grande perchè il gruppo generato dalle due rotazioni di angolo π contiene una traslazione non banale).

Definizione 10.2.4. Sia $\mathcal{C}$ l'ellisse di equazione (10.2.1) ($a \geqslant b$). I *fuochi* di $\mathcal{C}$ sono i punti

$$F_1(c, 0), \quad F_2(-c, 0),$$

dove

$$c = \sqrt{a^2 - b^2}. \tag{10.2.2}$$

L'*eccentricità* di $\mathcal{C}$ è

$$e = \frac{c}{a} = \frac{\sqrt{a^2 - b^2}}{a}.$$

Il significato geometrico dei fuochi è dato dal seguente risultato.

Proposizione 10.2.5. *Sia $\mathcal{C} \subset \mathbb{S}$ l'ellisse di equazione cartesiana* (10.2.1)*, e siano F_1, F_2 i suoi fuochi. Allora*

$$\mathcal{C} = \{P \in \mathbb{S} \mid d(P, F_1) + d(P, F_2) = 2a\}. \tag{10.2.3}$$

(A parole: $\mathcal{C}$ è l'insieme dei punti tali che la somma delle distanze da P_1 e P_2 sia uguale a $2a$.) Viceversa, se $P_1, P_2 \in \mathbb{S}$ e $d > d(P_1, P_2)$, la figura definita da

$$\mathcal{D} = \{P \in \mathbb{S} \mid d(P, P_1) + d(P, P_2) = d\} \tag{10.2.4}$$

è un'ellisse.

Dimostrazione. La prima asserzione si verifica con una serie di calcoli che lascio al lettore. Ora dimostriamo il viceversa. Supponiamo che $\mathcal{C}$ sia data da (10.2.4). Siano $a > 0$ e $c > 0$ tali che

$$2c = d(P_1, P_2), \qquad 2a = d.$$

Siccome $d > d(P_1, P_2)$ abbiamo $a > c$, e quindi ha senso porre $b := \sqrt{a^2 - c^2}$. Ora sia $\mathcal{C}$ l'ellisse di equazione (10.2.1), e siano F_1, F_2 i suoi fuochi. Allora $d(F_1, F_2) = d(P_1, P_2)$, e quindi esiste una isometria $\Phi \in \mathrm{Isom}(\mathbb{S})$ che porta F_i in P_i per $i \in \{1, 2\}$. Siccome per l'ellisse di equazione cartesiana (10.2.1) vale (10.2.3), abbiamo $\Phi(\mathcal{C}) = \mathcal{D}$, e quindi $\mathcal{D}$ è un'ellisse. □

Definizione 10.2.6. Sia $\mathcal{C}$ l'ellisse di equazione (10.2.1) ($a \geqslant b$). Le *direttrici* di $\mathcal{C}$ sono le rette di equazioni

$$D_1 : x = \frac{a^2}{c}, \quad D_2 : x = -\frac{a^2}{c},$$

dove c è dato da (10.2.2).

Notate che le direttrici non sono definite se $\mathcal{C}$ è un cerchio. Il seguente risultato (la cui dimostrazione è lasciata al lettore) dà il significato geometrico delle direttrici e dell'eccentricità.

Proposizione 10.2.7. *Sia $\mathcal{C}$ l'ellisse di equazione* (10.2.1)*, siano e la sua eccentricità e D_1, D_2 le sue direttrici. Allora per $i \in \{1, 2\}$ abbiamo che*

$$\mathcal{C} = \{P \in \mathbb{S} \mid d(P, F_i) = e \cdot d(P, D_i)\}. \tag{10.2.5}$$

Notate che l'eccentricità di un'ellisse è un positivo minore di 1. La dimostrazione del seguente risultato (che è simile alla dimostrazione del "viceversa" della Proposizione 10.2.5) viene lasciata al lettore.

Proposizione 10.2.8. *Siano $P_1 \in \mathbb{S}$, $R \subset \mathbb{S}$ una retta non contenente P_1 ed $0 < e < 1$. Allora*

$$\mathcal{D} := \{P \in \mathbb{S} \mid d(P, P_1) = e \cdot d(P, R)\}$$

è un'ellisse.

Iperboli

Definizione 10.2.9. $\mathcal{C} \subset \mathbb{S}$ è un'*iperbole* se in un opportuno riferimento ON di coordinate (x, y) ha equazione

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \tag{10.2.6}$$

cioè $P(x, y) \in \mathcal{C}$ se e solo se vale (10.2.6). Si intende sempre $a > 0$ e $b > 0$.

Osservazione 10.2.10. Un'iperbole ha equazione simile a quella che definisce un'ellisse e, come faremo vedere, condivide molte delle proprietà delle ellissi. C'e' una differenza che salta agli occhi: mentre un'ellisse è una curva connessa, un'iperbole è fatta di due curve separate:

$$\mathcal{C} = \left\{ \left(a\sqrt{1 + \frac{y^2}{b^2}}, y \right) \mid y \in \mathbb{R} \right\} \sqcup \left\{ \left(-a\sqrt{1 + \frac{y^2}{b^2}}, y \right) \mid y \in \mathbb{R} \right\}.$$

Le due componenti si chiamano *rami* dell'iperbole.

Gli assi di simmetria e il centro di un'iperbole di equazione (10.2.6) sono gli stessi dell'ellisse (ma esistono sempre solo due assi di simmetria, non esistono "iperboli circolari"). Fuochi ed eccentricità si definiscono come nel caso delle ellissi, ma con un cambio di segno.

Definizione 10.2.11. Sia $\mathcal{C}$ l'iperbole di equazione (10.2.6). I *fuochi* di $\mathcal{C}$ sono i punti

$$F_1(c, 0), \quad F_2(-c, 0),$$

dove

$$c = \sqrt{a^2 + b^2}. \tag{10.2.7}$$

L'*eccentricità* di $\mathcal{C}$ è

$$e = \frac{c}{a} = \frac{\sqrt{a^2 + b^2}}{a}.$$

Il seguente risultato è l'analogo per le iperboli della Proposizione 10.2.5, e anche la sua dimostrazione è analoga, lasciamo al lettore i dettagli.

Proposizione 10.2.12. *Sia $\mathcal{C} \subset \mathbb{S}$ l'iperbole di equazione cartesiana* (10.2.6)*, e siano F_1, F_2 i suoi fuochi. Allora*

$$\mathcal{C} = \{P \in \mathbb{S} \mid |d(P, F_1) - d(P, F_2)| = 2a\}.$$

(A parole: $\mathcal{C}$ è l'insieme dei punti tali che le distanze da P_1 e P_2 differiscono di $2a$.) Viceversa, se $P_1, P_2 \in \mathbb{S}$ sono punti distinti e $0 < d < d(P_1, P_2)$, la figura definita da

$$\mathcal{D} = \{P \in \mathbb{S} \mid |d(P, P_1) - d(P, P_2)| = d\}$$

è un'iperbole.

Definizione 10.2.13. Sia $\mathcal{C}$ l'iperbole di equazione (10.2.6). Le *direttrici* di $\mathcal{C}$ sono le rette di equazioni

$$D_1 : x = \frac{a^2}{c}, \quad D_2 : x = -\frac{a^2}{c},$$

dove c è dato da (10.2.7).

Valgono risultati analoghi alle Proposizioni 10.2.7 e 10.2.8.

Proposizione 10.2.14. *Sia $\mathcal{C}$ l'iperbole di equazione* (10.2.6)*, siano e la sua eccentricità e D_1, D_2 le sue direttrici. Allora i due rami di $\mathcal{C}$ sono gli insiemi dei punti P tali che*

$$d(P, F_i) = e \cdot d(P, D_i) \tag{10.2.8}$$

per $i = 1$ e $i = 2$.

Proposizione 10.2.15. *Siano $P_1 \in \mathbb{S}$, $R \subset \mathbb{S}$ una retta non contenente P_1 ed $e > 1$. Allora*

$$\mathcal{D} := \{P \in \mathbb{S} \mid d(P, P_1) = e \cdot d(P, R)\}$$

è uno dei due rami di un'iperbole.

Parabole

Nella descrizione di ellisse e iperbole tramite fuoco e direttrice il rapporto tra le distanze è uguale all'eccenticità, che è minore di 1 nel caso di un'ellisse, e maggiore di 1 nel caso di un'iperbole. Se facciamo l'analoga costruzione con rapporto uguale a 1 otteniamo una parabola. Diamo una definizione di parabola facendo uso di coordinate, poi vedremo che la parabola si descrive anche per mezzo di fuoco e direttrice.

Definizione 10.2.16. $\mathcal{C} \subset \mathbb{S}$ è una *parabola* se in un opportuno riferimento ON di coordinate (x, y) ha equazione

$$y = \frac{x^2}{4c}, \tag{10.2.9}$$

dove $c > 0$.

Definizione 10.2.17. Sia $\mathcal{C}$ la parabola di equazione (10.2.9). Il *fuoco* di $\mathcal{C}$ è il punto $F(0, c)$ e la *direttrice* di $\mathcal{C}$ è la retta di equazione

$$D : y = -c.$$

La seguente proposizione dà la descrizione della parabola per mezzo di fuoco e direttrice, i dettagli della dimostrazione sono lasciati al lettore.

Proposizione 10.2.18. *La parabola di equazione* (10.2.9) *è l'insieme dei punti equidistanti dal fuoco e dalla direttrice, cioè tali che*

$$d(P, F) = d(P, D). \tag{10.2.10}$$

Viceversa l'insieme dei punti equidistanti da un punto F e da una retta R non contenente P è una parabola.

La conlusione è che la parabola è un caso limite (comune) di iperbole ed ellissi.

10.3 Coniche a meno di isometrie

Coniche

Sia $\mathbb{S}$ uno spazio affine euclideo di dimensione 3 (anche detto *solido* affine euclideo). Un cono (circolare retto) in $\mathbb{S}$ è descritto come segue. Sia $\mathbb{T} \subset \mathbb{S}$ un piano, e sia $\mathcal{C} \subset \mathbb{T}$ un cerchio, di centro C_0. Sia R la retta ortogonale a $\mathbb{T}$ contenente C_0, e sia $P_0 \in (R \backslash \{C_0\})$. Il *cono di vertice* P_0 *e base* $\mathcal{C}$ è la superficie $\mathcal{V}$ spazzata dalle rette $\mathrm{Span}(P_0, P)$, dove P è un arbitrario punto di $\mathcal{C}$. Scegliendo un sistema di riferimento affine ortonormale $RO(O, \{\mathbf{i}, \mathbf{j}, \mathbf{k}\})$ con $O = C_0$ e $\mathsf{V}(R) = \mathrm{Span}(\mathbf{k})$, vediamo che $\mathcal{V}$ è data da

$$\mathcal{V} = \{P(x, y, z) \mid x^2 + y^2 - \rho^2 z^2 = 0\} \tag{10.3.1}$$

per un opportuno $\rho > 0$. Quindi le coordinate dei punti di $\mathcal{V}$ sono le soluzioni di un'unica equazione (detta equazione cartesiana di $\mathcal{V}$).

Si chiama *conica* una figura piana ottenuta intersecando un cono (circolare retto) con un piano. Se $\mathbb{P}$ è il piano con cui intersechiamo, e scegliamo un riferimento affine ortonormale $RO(O', \{v, w\})$ in $\mathbb{P}$, le coordinate di $\mathcal{V} \cap \mathbb{P}$ sono le soluzione di una singola equazione polinomiale di grado 2. Infatti le coordinate (x, y, z) del punto in $\mathbb{P}$ di coordinate (s, t) (per il riferimento affine ortonormale $RO(O', \{v, w\})$) sono date da

$$x = a_1 s + b_1 t + c_1, \qquad y = a_2 s + b_2 t + c_2, \qquad z = a_3 s + b_3 t + c_3,$$

dove $a_1, \ldots, c_3 \in \mathbb{R}$ e la matrice

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}$$

ha rango 2 (perchè le righe sono le coordinate di generatori di $\mathsf{V}(\mathbb{P}) \subset \mathsf{V}(\mathbb{S})$). Quindi

$$\mathcal{V} \cap \mathbb{P} = \{P(s, t) \in \mathbb{P} \mid (a_1 s + b_1 t + c_1)^2 + (a_2 s + b_2 t + c_2)^2 - \rho^2 (a_3 s + b_3 t + c_3)^2 = 0\}.$$

Questo mostra che le coordinate dei punti di $\mathcal{V} \cap \mathbb{P}$ sono le soluzioni di un'unica equazione polinomiale di grado 2 (in generale *non* omogenea, nonostante l'equazione di $\mathcal{V}$ sia omogenea).

La discussione appena fatta motiva lo studio di figure piane i cui punti hanno coordinate che sono soluzioni di una equazione polinomiale di grado 2. Dimostreremo che tali figure sono ellissi, iperboli e parabole, oppure alcune curve "degeneri", per esempio l'unione di due rette, eventualmente coincidenti, o anche un singolo punto (osservate che una tale figura si ottiene intersecando $\mathcal{V}$ con un piano contente il vertice di $\mathcal{V}$).

Funzioni polinomiali su uno spazio affine

Se V è uno spazio vettoriale finitamente generato su $\mathbb{K}$ ha senso parlare di funzioni polinomiali $f \colon V \to \mathbb{K}$: sono le funzioni che, nelle coordinate di una base sono date da un polinomio. Tale definizione è sensata perché se f è polinomiale in una base, allora lo è in qualsiasi base. Ora sia $\mathbb{S}$ uno spazio affine su $\mathbb{K}$ (di dimensione finita). Se $f \colon \mathbb{S} \to \mathbb{K}$ è una funzione che è polinomiale quando è espressa in termini delle coordinate di un riferimento affine scelto, allora è polinomiale anche quando è espressa in termini delle coordinate di in un qualsiasi altro sistema di riferimento affine. Notiamo anche che se $f \colon \mathbb{S} \to \mathbb{K}$ ha grado d in un sistema di coordinate affini, allora ha grado d in qualsiasi altro sistema di coordinate affini. Per questo motivo la definizione che segue ha senso.

Definizione 10.3.1. Sia $\mathbb{S}$ uno spazio affine su $\mathbb{K}$ (di dimensione finita). Una funzione $f \colon \mathbb{S} \to \mathbb{K}$ è *polinomiale* se, scelto un sistema di riferimento affine $RA(O, \mathcal{B})$ con relative coordinate $\mathbb{S} \xrightarrow{X_{RA}} \mathbb{A}^n(\mathbb{R})$, la funzione

$$\begin{array}{ccc} \mathbb{A}^n(\mathbb{R}) & \xrightarrow{f_{RA}} & \mathbb{R} \\ (x_1, \ldots, x_n) & \mapsto & f(X_{RA}^{-1}(x_1, \ldots, x_n)) \end{array}$$

è polinomiale. Il *grado di* f è il grado di f_{RA}.

Osservazione 10.3.2. Se V è uno spazio vettoriale su $\mathbb{K}$ ha senso parlare di funzioni polinomiali omogenee $f \colon V \to \mathbb{K}$, ma se $\mathbb{S}$ è uno spazio affine su $\mathbb{K}$ *non* ha senso parlare di funzioni polinomiali omogenee $f \colon \mathbb{A} \to \mathbb{K}$. Questo perch'è se $f \in \mathbb{K}[x_1, \ldots, x_n]_d$ è omogenea non nulla di grado $d > 0$, e

$$X = A \cdot Y + B, \quad A \in \mathrm{GL}_n(\mathbb{K})$$

sono nuove coordinate affini, allora $f(A \cdot X + B)$ non è omogeneo (in generale). Detto in modo diverso: se V è uno spazio vettoriale, allora $f\colon V \to \mathbb{K}$ è omogeneo di grado d se

$$f(\lambda X) = \lambda^d f(X), \quad \forall \lambda \in \mathbb{K}.$$

Ma in uno spazio affine non esiste un origine privilegiata, quindi...

Definizione 10.3.3. Sia $\mathbb{S}$ uno spazio affine su $\mathbb{K}$, e sia $f\colon \mathbb{S} \to \mathbb{K}$ una funzione polinomiale. La *ipersuperficie algebrica* di equazione f è il sottoinsieme di $\mathbb{S}$ definito da

$$V(f) := \{P \in \mathbb{S} \mid f(P) = 0\}. \tag{10.3.2}$$

Se $\dim \mathbb{S} = 2$ allora $V(f)$ si dice *curva algebrica*, se $\dim \mathbb{S} = 3$ si dice *superficie algebrica*.

Osservazione 10.3.4. L'uso del nome "curva algebrica"(o "superficie algebrica") corrisponde alla nostra intuizione per molte funzioni polinomiali f, ma non per tutte. Per esempio $V(x^2 + y^2 + 1) \subset \mathbb{A}^2(\mathbb{R})$ è l'inisieme vuoto.

La classificazione delle coniche a meno di isometrie

Per il resto della sezione $\mathbb{S}$ è un piano affine euclideo. Classificheremo i sottoinsiemi ("figure") di $\mathbb{S}$ che sono definite da una singola equazione polinomiale di grado 2, cioè le curve algebriche $V(f) \subset \mathbb{S}$ dove dove $f\colon \mathbb{S} \to \mathbb{R}$ è una funzione polinomiale di grado 2. Notiamo che per definizione un'ellisse, un'iperbole o una parabola sono tali figure. Ci interessa la classificazione di tali curve a meno di isometrie di $\mathbb{S}$, cioè consideriamo equivalenti figure $\mathcal{C}_1, \mathcal{C}_2 \subset \mathbb{S}$ se esista un'isometria $\Phi\colon \mathbb{S} \to \mathbb{S}$ tale che $\Phi(\mathcal{C}_1) = \mathcal{C}_2$.

Prima di imbarcarci nella classificazione, facciamo un passo indietro, e notiamo che l'analoga classificazione per figure definite da una singola equazione di grado 1 (cioè le rette) è molto semplice: due rette qualsiasi sono equivalenti, cioè esiste un'isometria che porta una retta assegnata in un'altra retta assegnata. La classificazione delle figure definite da un'equazione polinomiale di grado 2 non è altrettanto semplice.

In realtà classificheremo le funzioni polinomiali $f\colon \mathbb{S} \to \mathbb{R}$ di grado 2 a meno di isometrie e riscalamento, anche se per alcune tali f la "curva algebrica" non è affatto una curva (vedi l'Osservazione 10.3.4).

Definizione 10.3.5. Due funzioni polinomiali

$$f, g\colon \mathbb{S} \to \mathbb{R} \tag{10.3.3}$$

sono *equivalenti per isometrie* (in simboli $f \sim g$) se esistono un'isometria $\Phi \in \mathrm{Isom}(\mathbb{S})$ e $0 \neq \lambda \in \mathbb{R}$ tali che

$$g(P) = \lambda f(\Phi(P)) \quad \forall P \in \mathbb{S}. \tag{10.3.4}$$

(Si verifica semplicemente che quella definita è una relazione di equivalenza.)

Osservazione 10.3.6. Se $f\colon \mathbb{S} \to \mathbb{R}$ è una funzione polinomiale e $0 \neq \lambda \in \mathbb{R}$, allora $V(\lambda f) = V(f)$. Questo è il motivo per cui appare il fattore non nullo λ nella Definizione 10.3.5.

Osservazione 10.3.7. Siano $f, g\colon \mathbb{S} \to \mathbb{R}$ funzioni polinomiali di grado 2. Se $f \sim g$, cioè vale (10.3.4), allora $V(g) = \Phi^{-1}(V(f))$. Non è vero il viceversa, cioè se esiste un'isometria $\Phi \in \mathrm{Isom}(\mathbb{S})$ tale che $V(g) = \Phi^{-1}(V(f))$, allora non segue che $f \sim g$. Per esempio se

$$f := (x^2 + y^2 + 1), \quad g := (3x^2 + y^2 + 1),$$

allora $V(f) = V(g) = \varnothing$, ma f non è equivalente per isometrie a g.

Osservazione 10.3.8. Funzioni polinomiali f, g come in (10.3.3) sono equivalenti per isometrie se e solo se esistono riferimenti ON $RO(O, \mathcal{B})$ e $RO(O', \mathcal{B}')$ tali che

$$f(X(P)) = \lambda \cdot g(X'(P)) \quad \forall P \in \mathbb{S},$$

dove $X(P)$ e $X'(P)$ sono le coordinate di P nei riferimenti $RO(O, \mathcal{B})$ e $RO(O', \mathcal{B}')$ rispettivamente, e $0 \neq \lambda \in \mathbb{R}$. In parole: nei due riferimenti ON f e g diventano la stessa funzione a meno di uno scalare.

Tabella 10.1: Forme canoniche di funzioni polinomiali di grado 2 in due variabili

Equazione canonica	Nome	
$E_{a,b}(x,y) := \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1, \quad a \geqslant b > 0$	ellisse	coniche non-degeneri
$C_{a,b}(x,y) := \frac{x^2}{a^2} + \frac{y^2}{b^2} + 1, \quad a \geqslant b > 0$	ellisse complessa	
$I_{a,b}(x,y) := \frac{x^2}{a^2} - \frac{y^2}{b^2} - 1, \quad a, b > 0$	iperbole	
$P_c(x,y) := \frac{x^2}{4c} - y, \quad c > 0$	parabola	
$W_{a,b}(x,y) := \frac{x^2}{a^2} + \frac{y^2}{b^2}, \quad a \geqslant b > 0$	coppia di rette complesse	coniche degeneneri
$V_{a,b}(x,y) := \frac{x^2}{a^2} - \frac{y^2}{b^2}, \quad a \geqslant b > 0$	coppia di rette incidenti	
$L_a(x,y) := \frac{x^2}{a^2} - 1, \quad a > 0$	coppia di rette parallele	
$H_a(x,y) := \frac{x^2}{a^2} + 1, \quad a > 0$	coppia di rette complesse	
$D(x,y) := x^2$	retta doppia	

Teorema 10.3.9. *Sia $f\colon \mathbb{S} \to \mathbb{R}$ una funzione polinomiale di grado 2 non nulla. Allora esiste un sistema ON di coordinate affini tale che nelle coordinate associate (x,y) la f sia data da una delle seguenti* forme canoniche *della Tabella 10.1. Inoltre siano $f, g\colon \mathbb{S} \to \mathbb{R}$ due funzioni polinomiali di grado 2. Allora f è equivalente per isometrie a g se e solo se hanno la stessa forma canonica (cioè sono entrambe $E_{*,*}$, o $C_{*,*}$, etc.), e con gli stessi parametri (cioè a, b, o a, b, c, o a, o nessun parametro nell'ultima riga).*

Ora guardiamo agli "zeri" delle funzioni che appaiono nella Tabella 10.1, cioè a $V(f)$ per f nella lista data (e spieghiamo il significato della seconda e terza colonna). Tra le coniche non-degeneri abbiamo ellissi, iperboli, parabole, e l'"ellisse complessa" (si capisce il significato geometrico di quest'ultima analizzando le soluzioni di $f(x,y) = 0$ con x, y numeri complessi). Le prime tre sono state discusse, la quarta è del tutto non interessante se ci limitiamo alle soluzioni reali perchè è l'insieme vuoto.

Tra le coniche degeneneri le coppie di rette (incidenti o parallele), e la retta con "molteplicità 2" sono geometricamente chiare, mentre per la coppia di rette complesse valgono gli stessi commenti fatti per l'ellisse complessa.

In particolare segue che una conica è necessariamente un'ellisse, o un'iperbole etc. Notate che questo è un risultato del tutto non banale, anche solo il risultato che una sezione piana di una conica abbia due assi di simmetria.

Dimostrazione del Teorema 10.3.9. Dobbiamo dimostrare che

1. funzioni date da lettere diverse (per esempio $E_{a,b}$ e $C_{a,b}$) non sono equivalenti, e funzioni date dalla stessa lettera (per esempio $E_{a,b}$ ed $E_{a',b'}$) sono equivalenti solo se le lettere che danno i sottoindici sono le stesse (per esempio $E_{a,b} \sim E_{a',b'}$ solo se $(a,b) = (a',b')$), e
2. che ogni una funzione $f\colon \mathbb{A} \to \mathbb{R}$ polinomiale di grado 2 è equivalente per isometrie a una tra $E_{a,b}, \ldots, D$.

Dimostriamo (1). Supponiamo che $f \sim g$. Per l'Osservazione 10.3.7 esiste un'isometria $\Phi \in \mathrm{Isom}(\mathbb{S})$ tale che $V(f) = \Phi(V(g))$. Guardando alla seconda colonna della Tabella 10.1 si vede facilmente che funzioni equivalenti sono date dalle stesse lettere maiuscole. Rimane da dimostrare che se $E_{a,b}$ ed $E_{a',b'}$ sono equivalenti allora $(a,b) = (a'b')$, che se $C_{a,b}$ e $C_{a',b'}$ sono equivalenti allora $(a,b) = (a'b')$, etc. È chiaro come procedere nei casi in cui gli zeri della funzione non sono un punto o l'insieme vuoto.

Dimostriamo che se $C_{a,b}$ e $C_{a',b'}$ sono equivalenti allora $(a,b) = (a',b')$. Sia $\Phi \in \mathrm{Isom}(\mathbb{S})$ tale che

$$C_{a,b}(P) = \lambda \cdot C_{a',b'}(\Phi(P)) \quad \forall P \in \mathbb{S}. \tag{10.3.5}$$

Siccome $\mathrm{im}(C_{a,b}) = [1, +\infty) = \mathrm{im}(C_{a',b'})$ e il valore minimo è raggiunto per $P = (0,0)$ in entrambi i casi, segue che $\Phi(0,0) = (0,0)$. Quindi l'uguaglianza in (10.3.5) si legge

$$\Phi(x,y) = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \tag{10.3.6}$$

dove la matrice $M = (m_{ij})$ è ortogonale, cioè $M^t = M^{-1}$. Quindi

$$\begin{pmatrix} a^{-2} & 0 \\ 0 & b^{-2} \end{pmatrix} = M^{-1} \cdot \begin{pmatrix} (a')^{-2} & 0 \\ 0 & (b')^{-2} \end{pmatrix} \cdot M, \tag{10.3.7}$$

e da questo segue facilmente che $(a,b) = (a',b')$, perchè $a \geqslant b > 0$ e $a' \geqslant b' > 0$. La dimostrazione per coppie di rette complesse è simile (per $W_{a,b}$) o più semplice (H_a).

Ora dimostriamo il punto (2). Conviene denotare le coordinate con (x,x_2). Scriviamo

$$f(x_1,x_2) = X^t \cdot M \cdot X + \mu x_1 + \nu x_2 + \theta \tag{10.3.8}$$

dove $0 \neq M \in M^+_{2,2}(\mathbb{R})$ e $\mu, \nu, \theta \in \mathbb{R}$. Dobbiamo dimostrare che esiste un cambio di coordinate ON

$$X = A \cdot Y + B, \quad A \in \mathrm{O}_2(\mathbb{R}),\ B \in M_{2,1}(\mathbb{R}) \tag{10.3.9}$$

tale che $f(A \cdot Y + B)$ sia un multiplo non nullo di una delle funzioni $E_{a,b}, C_{a,b}, \ldots, D$ della Tabella 10.1. Faremo due cambi di coordinate, la composizione sarà il cambio cercato. Se facciamo un cambio di coordinate che fissa l'origine, troviamo

$$f(A \cdot Y) = Y^t \cdot (A^t \cdot M \cdot A) \cdot Y + \mu' x_1 + \nu' x_2 + \theta',$$

dove $\mu', \nu', \theta' \in \mathbb{R}$. Per il Teorema spettrale reale esiste $A \in \mathrm{O}_2(\mathbb{R})$ tale che $A^t \cdot M \cdot A$ sia diagonale. Quindi f è equivalente

$$\alpha y_1^2 + \beta y_2^2 + \mu y_1 + \nu y_2 + \theta.$$

Ora distinguiamo i due casi:

(a) $\alpha \neq 0$ e $\beta \neq 0$,

(b) uno tra α e β è nullo.

Nel caso (a) scriviamo

$$\alpha y_1^2 + \beta y_2^2 + \mu y_1 + \nu y_2 + \theta = \alpha \left[\left(y_1 + \frac{\mu}{2\alpha} \right)^2 - \frac{\mu^2}{4\alpha^2} \right] + \beta \left[\left(y_2 + \frac{\nu}{2\beta} \right)^2 - \frac{\nu^2}{4\beta^2} \right] + \theta.$$

La traslazione

$$z_1 = y_1 + \frac{\mu}{2\alpha}, \quad z_2 = y_2 + \frac{\nu}{2\beta}$$

mostra che f è equivalente a

$$\alpha z_1^2 + \beta z_2^2 + \gamma.$$

Se $\gamma \neq 0$, dividiamo per γ, e vediamo che γ è equivalente a un $E_{a,b}$, un $C_{a,b}$ o un $I_{a,b}$. Se $\gamma = 0$ vediamo che γ è equivalente a un $W_{a,b}$, o un $V_{a,b}$.

Nel caso (b), cioè uno tra α e β è nullo, si procede in modo simile e si trova che f è equivalente a uno tra P_c, L_a, H_a e D. □

Esempio 10.3.10. Sia $\mathcal{C} \subset \mathbb{E}^2(\mathbb{R})$ data da

$$\mathcal{C}: \ 2x_1^2 + 72x_1x_2 + 23x_2^2 + 25x_1 = 0.$$

Determiniamo che tipo di "curva" è $\mathcal{C}$ (curva tra virgolette perchè potrebbe essere l'insieme vuoto, o un singolo punto).

Passo 1. Troviamo una base ON che diagonalizza la forma quadratica

$$2x_1^2 + 72x_1x_2 + 23x_2^2.$$

La matrice simmetrica corrispondente alla forma quadratica è

$$\begin{pmatrix} 2 & 36 \\ 36 & 23 \end{pmatrix}$$

Gli autovalori sono 50 e -25, con corrispondenti autovettori

$$(3,4), \quad (4,-3).$$

Normalizzando gli autovettori elencati, troviamo una base ON che diagonalizza la forma quadratica:

$$\mathcal{B} := \left\{ \left(\frac{3}{5}, \frac{4}{5}\right), \ \left(\frac{4}{5}, -\frac{3}{5}\right) \right\}.$$

Siano (y_1, y_2) le coordinate del riferimento $RO(O, \mathcal{B})$. La formula per il cambio di coordinate è

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

Nel nuovo sistema di riferimento l'equazione di $\mathcal{C}$ è

$$\mathcal{C}: \ 50y_1^2 - 25y_2^2 + 15y_1 + 20y_2 = 0.$$

Dividendo per 5 otteniamo l'equazione

$$\mathcal{C}: \ 10y_1^2 - 5y_2^2 + 3y_1 + 4y_2 = 0.$$

"Completando i quadrati"l'equazione di $\mathcal{C}$ si scrive

$$10\left[\left(y_1 + \frac{3}{20}\right)^2 - \frac{9}{400}\right] - 5\left[\left(y_2 - \frac{4}{10}\right)^2 - \frac{16}{100}\right] = 0.$$

Cambiamo riferimento ponendo

$$z_1 = y_1 + \frac{3}{20}, \quad z_2 = y_2 - \frac{4}{10},$$

e l'equazione diventa

$$\mathcal{C}: \ 10z_1^2 - 5z_2^2 + \frac{23}{40} = 0.$$

Poniamo $w_1 = z_2$ e $w_2 = z_1$. L'equazione diventa

$$\frac{200}{23}w_1^2 - \frac{400}{23}z_2^2 = 1.$$

Quindi $\mathcal{C}$ è un'iperbole, e con un pò di pazienza potremmo calcolarne fuochi, eccentricità e direttrici.

Esempio 10.3.11. Sia $\mathcal{C} \subset \mathbb{E}^2$ la parabola data da

$$\mathcal{C}: \ 25x_1^2 - 120x_1x_2 + 144x_2^2 - 494x_1 - 572x_2 + 169 = 0.$$

Determiniamo fuoco e direttrice di $\mathcal{C}$. (Non è affatto chiaro che $\mathcal{C}$ sia una parabola, lo scopriremo nel corso dei calcoli.) Come al solito il primo passo consiste nel trovare una base ON che diagonalizza la forma quadratica ottenuta dimenticando i termini di grado minore di 1, e cioè

$$f_2 := 25x_1^2 - 120x_1x_2 + 144x_2^2.$$

La matrice simmetrica corrispondente a f_2 è

$$M := \begin{pmatrix} 25 & -60 \\ -60 & 144 \end{pmatrix}$$

Gli autovalori sono

$$169, \quad 0$$

con corrispondenti autovettori di norma 1 (e ortogonali tra loro!) dati da

$$(5/13, -12/13), \quad (12/13, 5/13).$$

Quindi il primo cambio di base è dato da

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{5}{13} & \frac{12}{13} \\ -\frac{12}{13} & -\frac{5}{13} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \tag{10.3.10}$$

Nelle nuove coordinate l'equazione di $\mathcal{C}$ è

$$169y_1^2 + 338y_1 - 676y_2 + 169 = 0.$$

Tutti i coefficienti sono divisibili per 169, e quindi

$$\mathcal{C}: \ y_1^2 + 2y_1 - 4y_2 + 1 = 0.$$

Ponendo

$$z_1 = y_1 + 1, \quad z_2 = y_2,$$

abbiamo l'equazione

$$\frac{1}{4}z_1^2 - z_2 = 0.$$

Ora è chiaro che $\mathcal{C}$ è una parabola. Siano F il fuoco di $\mathcal{C}$ e D la sua direttrice: nelle coordinate (z_1, z_2) abbiamo

$$F(0,1), \quad D: \ z_2 + 1 = 0.$$

Quindi nelle coordinate (y_1, y_2) abbiamo

$$F(-1,1), \quad D: \ y_2 + 1 = 0.$$

Usando il cambio di coordinate in (10.3.10) troviamo che

$$F = \left(\frac{7}{13}, \frac{17}{13}\right), \quad D: \ 12x_1 + 5x_2 + 13 = 0.$$

Come semplificare i calcoli

I calcoli fatti negli Esempi 10.3.10 e 10.3.11 sono piuttosto lunghi. In verità esistono metodi più veloci per stabilire a quale delle funzioni $E_{a,b}$ etc. è equivalente un dato polinomio di grado 2. Qui ci limiteremo a dare un algoritmo (e a giustificarlo). Nel Capitolo 11 daremo una spiegazione geometrica di quello che stiamo per fare.

Data $f \in \mathbb{R}[x_1, x_2]$, definiamo la forma quadratica F in 3 variabili x_0, x_1, x_2 (l'*omogeneizzazione* di f) così:

$$F(x_0, x_1, x_2) := x_0^2 \cdot f\left(\frac{x_1}{x_0}, \frac{x_2}{x_0}\right). \tag{10.3.11}$$

Per esempio, se

$$f(x_1, x_2) := x_1^2 - 3x_1x_2 + 5x_2^2 - 3x_1 + 4,$$

allora

$$F(x_0, x_1, x_2) := x_1^2 - 3x_1x_2 + 5x_2^2 - 3x_0x_1 + 4x_0^2.$$

e se

$$f(x_1, x_2) := 4x_1^2 - x_1 + 4x_2 - 6,$$

allora

$$F(x_0, x_1, x_2) := 4x_1^2 - x_0x_1 + 4x_0x_2 - 6x_0^2.$$

Ora facciamo un cambiamento di coordinate:

$$X = A \cdot Y + B, \quad A \in \mathrm{O}_2(\mathbb{R}), \ B \in M_{2,1}(\mathbb{R}).$$

Poniamo

$$g(y_1, y_2) = f(a_{11}y_1 + a_{12}y_2 + b_1, a_{21}y_1 + a_{22}y_2 + b_2),$$

e sia

$$G(y_0, y_1, y_2) = y_0^2 \cdot g\left(\frac{y_1}{y_0}, \frac{y_2}{y_0}\right)$$

la forma quadratica in (y_0, y_1y_2) associata a g. L'osservazione fondamentale è che

$$G(y_0, y_1, y_2) = F(y_0, a_{11}y_1 + a_{12}y_2 + b_1y_0, a_{21}y_1 + a_{22}y_2 + b_2y_0). \tag{10.3.12}$$

Tabella 10.2: Forme canoniche omogeneizzate

Equazione canonica omogenizzata	Nome	
$\widehat{E}_{a,b}(x,y) := \frac{x^2}{a^2} + \frac{y^2}{b^2} - z^2, \quad a \geqslant b > 0$	ellisse	coniche non-degeneri
$\widehat{C}_{a,b}(x,y) := \frac{x^2}{a^2} + \frac{y^2}{b^2} + z^2, \quad a \geqslant b > 0$	ellisse complessa	
$\widehat{I}_{a,b}(x,y) := \frac{x^2}{a^2} - \frac{y^2}{b^2} - z^2, \quad a, b > 0$	iperbole	
$\widehat{P}_c(x,y) := \frac{x^2}{4c} - yz, \quad c > 0$	parabola	
$\widehat{W}_{a,b}(x,y) := \frac{x^2}{a^2} + \frac{y^2}{b^2}, \quad a \geqslant b > 0$	coppia di rette complesse	coniche degeneneri
$\widehat{V}_{a,b}(x,y) := \frac{x^2}{a^2} - \frac{y^2}{b^2}, \quad a \geqslant b > 0$	coppia di rette incidenti	
$\widehat{L}_a(x,y) := \frac{x^2}{a^2} - z^2, \quad a > 0$	coppia di rette parallele	
$\widehat{H}_a(x,y) := \frac{x^2}{a^2} + z^2, \quad a > 0$	coppia di rette complesse	
$\widehat{D}(x,y) := x^2$	retta doppia	

Lemma 10.3.12. *Se $f, g \in \mathbb{R}[x_1, x_2]$, e f è equivalente (per isometrie) a g, allora le forme quadratiche associate F e G sono equivalenti a meno di uno scalare non nullo. (Possiamo dire di più, cioè che sono equivalenti per un elemento di $\mathrm{GL}_3(\mathbb{R})$ che manda il sottospazio $\{0, *, *)\}$ in se stesso.)*

Dimostrazione. Segue dall'uguaglianza in (10.3.12). □

Osservazione 10.3.13. Non vale il viceversa del Lemma 10.3.12, cioè se F e G sono forme quadratiche equivalenti non è detto che f e g siano equivalenti.

Per dedurre conseguenze dal Lemma 10.3.12 scriviamo le forme canoniche omogeneizzate, vedi la Tabella 10.2.

Proposizione 10.3.14. *Sia $f \in \mathbb{R}[x_1, x_2]$ un polinomio di grado 2 e sia $F(x_0, x_1, x_2)$ la forma quadratica associata a f, cioè l'omogeneizzato di f.*

(I) Se F è non degenere, allora f è equivalente a una tra $E_{a,b}, C_{a,b}, I_{a,b}, P_c$, e inoltre

- *(Ia) se $F(0, x_1, x_2)$ è definita ma $F(x_0, x_1, x_2)$ non è definita (cioè ha segnatura $(2,1)$ o $(1,2)$), allora f è equivalente a un $E_{a,b}$, e quindi $f(x_1, x_2) = 0$ è un'ellisse,*
- *(Ib) se $F(0, x_1, x_2)$ è non degenere e non definita (cioè ha segnatura $(1,1)$), allora f è equivalente a un $I_{a,b}$, e quindi $f(x_1, x_2) = 0$ è un'iperbole,*
- *(Ic) se $F(x_0, x_1, x_2)$ è definita, allora f è equivalente a un $C_{a,b}$, e quindi $f(x_1, x_2) = 0$ è vuoto,*
- *(Id) se $F(0, x_1, x_2)$ è degenere, allora f è equivalente a un P_c, e quindi $f(x_1, x_2) = 0$ è una parabola.*

(II) Se F è degenere, allora f è equivalente a una tra $W_{a,b}, V_{a,b}, L_{a,b}, H_a, D$, e inoltre

- *(IIa) se $\ker F$ è generato da un vettore con $x_0 \neq 0$, e la segnatura di F è $(2,0)$ o $(0,2)$, allora f è equivalente a un $W_{a,b}$,*
- *(IIb) se $\ker F$ è generato da un vettore con $x_0 \neq 0$, e la segnatura di F è $(1,1)$, allora f è equivalente a un $V_{a,b}$,*
- *(IIc) se $\ker F$ è generato da un vettore con $x_0 = 0$, e la segnatura di F è $(1,1)$, allora f è equivalente a un L_a,*

(IId) *se* $\ker F$ *è generato da un vettore con* $x_0 = 0$*, e la segnatura di* F *è* $(2,0)$ *o* $(0,2)$*, allora* f *è equivalente a un* H_a*,*

(IIe) *se* $\dim \ker F = 2$*, allora* f *è equivalente a* D*.*

Dimostrazione. Sappiamo che f è equivalente per isometrie a una forma canonica della Tabella 10.1. Quindi la proposizione segue dal Lemma 10.3.12 e dall'osservazione che le affermazioni sono vere per le forme canoniche (esaminate le loro omogeneizzazioni nella Tabella 10.2). □

Diamo anche un algoritmo che permette di semplicifcare il calcolo del *centro* di un'ellisse o di un'iperbole, cioè l'origine di un sistema di riferimento ON nel quale $\mathcal{C}$ ha l'equazione canonica data da (10.2.1) e (10.2.6) rispettivamente. Notate che il centro di un'ellisse o di un'iperboleè *il* centro di simmetria di $\mathcal{C}$.

Proposizione 10.3.15. *Sia* $\mathbb{S}$ *un piano affine euclideo, e sia* $\mathcal{C} \subset \mathbb{S}$ *una conica a centro. Sia* $f(x_1, x_2) = 0$ *l'equazione cartesiana di* $\mathcal{C}$ *in un sistema di riferimento ON, e sia*

$$F(x_0, x_1, x_2) = x_0^2 f(x_1/x_0, x_2/x_0)$$

l'omogeneizzazione di f*, con matrice associata (nella base standard) data da* A *(cioè* $F(X) = X^t \cdot A \cdot X$*). Sia*

$$\begin{array}{ccc} \mathbb{R}^3 & \xrightarrow{L_F} & (\mathbb{R}^3)^\vee \\ X & \mapsto & A \cdot X \end{array}$$

l'applicazione simmetrica associata a F*. Infine sia sia* $(a_0, a_1, a_2) \in \mathbb{R}^3$ *l'unico vettore tale che*

$$L_F(a_0, a_1, a_2) = (1, 0, 0).$$

Allora $a_0 \neq 0$ *e le coordinate (nel sistema di riferimento ON seclto) del centro di* $\mathcal{C}$ *sono date da*

$$\left(\frac{a_1}{a_0}, \frac{a_2}{a_0}\right).$$

Dimostrazione. Un semplice calcolo mostra che l'enunciato è vero se l'equazione di $\mathcal{C}$ è canonica. Quindi l'enunciato rimane vero qualsiasi sia l'equazione di $\mathcal{C}$ perchè ogni conica è equivalente per isometrie a una conica in equazione canonica, e per via del Lemma 10.3.12. □

Esempio 10.3.16. Sia $\mathcal{C} \subset \mathbb{E}^2(\mathbb{R})$ la conica di equazione

$$\mathcal{C}: \ x_1^2 + 3x_1x_2 + 2x_2^2 + x_1 - 1 = 0.$$

(a) Verifichiamo che $\mathcal{C}$ è un'iperbole.

(b) Determiniamo l'equazione canonica di $\mathcal{C}$.

(c) Determiniamo i fuochi di $\mathcal{C}$.

(a) Verifichiamo che $\mathcal{C}$ è un'iperbole. Ricordiamo che $\mathcal{C}$ è un'iperbole se è non degenere e la forma quadratica $q(x_1, x_2)$ ottenuta dimenticando i termini di grado minore di 2 nel polinomio che definisce $\mathcal{C}$ ha segnature $(1, 1)$.

L'omogeneizzazione del polinomio che definisce $\mathcal{C}$ è

$$F(x_0, x_1, x_2) = x_1^2 + 3x_1x_2 + 2x_2^2 + x_0x_1 - x_0^2,$$

e la matrice simmetrica associata (nella base standard) è

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

Siccome $\det A = -1/4$, $\mathcal{C}$ è uno conica non degenere.

La forma quadratica in 2 variabili ottenuta dimenticando i termini di grado minore di 2 nell'equazione di $\mathcal{C}$ è

$$q(x_1, x_2) = x_1^2 + 3x_1x_2 + 2x_2^2.$$

La matrice simmetrica associata è

$$M := \begin{pmatrix} 1 & 3/2 \\ 3/2 & 2 \end{pmatrix}$$

Siccome $\det M(1) = 1$, $\det M(2) = -1/4$, le segnature sono $(1,1)$, e concludiamo che $\mathcal{C}$ è un'iperbole.

(b) Determiniamo l'equazione canonica di $\mathcal{C}$. Troviamo una base ON che diagonalizza la forma quadratica q. Per fare questo calcoliamo autovalori e corrispondenti autovettori di M. Troviamo autovettori ortogonali

$$(3, 1+\sqrt{10}), \quad (1+\sqrt{10}, -3).$$

Non sono di norma 1. Poniamo

$$\alpha := (20 + 2\sqrt{10})^{-1/2}.$$

I vettori

$$v_1 = (3\alpha, (1+\sqrt{10})\alpha), \quad v_2 = ((1+\sqrt{10})\alpha, -3\alpha)$$

costituiscono una base ON che diagonalizza la forma quadratica q. Per proseguire, calcoliamo il centro di $\mathcal{C}$ (ha senso perchè $\mathcal{C}$ è un'iperbole) applicando la Proposizione 10.3.15. Nel nostro caso troviamo che il centro di $\mathcal{C}$ è

$$Q := (4, -3).$$

Tirando le somme: sia

$$\mathcal{B} = \{(3\alpha, (1+\sqrt{10})\alpha),\ ((1+\sqrt{10})\alpha, -3\alpha)\}$$

la base ON di $\mathbb{R}^2$ determinata sopra che diagonalizza la forma quadratica q, e siano (y_1, y_2) le coordinate del riferimento ON $RO(Q, \mathcal{B})$. Allora

$$\begin{aligned} x_1 &= 3\alpha y_1 + (1+\sqrt{10})\alpha y_2 + 4 \\ x_2 &= (1+\sqrt{10})\alpha y_1 - 3\alpha y_2 - 3 \end{aligned}$$

Sostituendo nell'equazione di $\mathcal{C}$ troviamo che l'equazione di $\mathcal{C}$ nelle nuove coordinate (y_1, y_2) è

$$\frac{1+\sqrt{10}}{2} y_1^2 - \frac{13\sqrt{10}-31}{18} y_2^2 + 1 = 0.$$

Quindi l'equazione canonica è

$$\frac{13\sqrt{10}-31}{18} z_1^2 - \frac{1+\sqrt{10}}{2} z_2^2 - 1 = 0.$$

Si trova

$$\begin{aligned} a^2 &= \frac{62 + 26\sqrt{10}}{81}, \\ b^2 &= \frac{2\sqrt{10}-2}{9}. \end{aligned}$$

Quindi le coordinate (z_1, z_2) dei fuochi di $\mathcal{C}$ sono

$$\left(\pm \frac{2}{9} \sqrt{11 + 11\sqrt{10}}, 0 \right).$$

Con un pò di pazienza possiamo trovarne le coordinate nel sistema di riferimento di partenza (cioè quello standard di $\mathbb{E}^2(\mathbb{R})$).

10.4 Quadriche e iperquadriche a meno di isometrie

La classificazione

Estendiamo agli spazi affini euclidei di dimensione maggiore di 2 quello che abbiamo dimostrato per i polinomi di grado 2 su un piano affine euclideo. Più precisamente classificheremo le funzioni $f\colon \mathbb{S} \to \mathbb{R}$ polinomiali di grado 2 a meno dell'equivalenza per isometrie, dove (come nel caso $n = 2$) $f \sim g$ se esistono $\Phi \in \mathrm{Isom}(\mathbb{S})$ e $0 \neq \lambda \in \mathbb{R}$ tale che

$$f(P) = g(\Phi(P)) \quad \forall P \in \mathbb{S}$$

Come in dimensione 2 la motivazione è geometrica, cioè si vuole capire come sono fatte le (iper)superfici $\mathcal{S} \subset \mathbb{S}$ definite da un'equazione quadratica

$$\mathcal{S}:\ f(x_1, \ldots, x_n) = 0,$$

dove $(x_1, \dots, x_n)$ sono coordinate di un riferimento ON di $\mathbb{S}$, e f è un polinomio di grado 2. Queste figure si chiamano *(iper)quadriche*.

Iniziamo alcune considerazioni. Come nel caso di dimensione 2, possiamo associare a $f \in \mathbb{R}[x_1, \dots, x_n]$ (non nulla) di grado 2 una forma quadratica $F \in \mathbb{R}[x_0, \dots, x_n]_2$ (la sua omogeneizzazione) ponendo

$$F(x_0, \dots, x_n) := x_0^2 \cdot f\left(\frac{x_1}{x_0}, \dots, \frac{x_n}{x_0}\right).$$

Facciamo un cambiamento di coordinate

$$X = A \cdot Y + B, \quad A \in \mathrm{O}_n(\mathbb{R}),\ B \in M_{n,1}(\mathbb{R}),$$

e poniamo

$$g(Y) = f(A \cdot X + B).$$

L'omogeneizzazione di g è

$$G(y_0, y_1, \dots, y_n) = y_0^2 \cdot g\left(\frac{y_1}{y_0}, \dots, \frac{y_n}{y_0}\right).$$

Come nel caso $n = 2$, l'osservazione fondamentale è che

$$G(Y) = F(M \cdot Y), \tag{10.4.1}$$

dove

$$M := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ b_1 & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_n & a_{n1} & \cdots & a_{nn} \end{pmatrix} \tag{10.4.2}$$

Lemma 10.4.1. *Sia $f \in \mathbb{R}[x_1, \dots, x_n]$ un polinomio di grado 2, e sia $F \in \mathbb{R}[x_0, \dots, x_n]_2$ la sua omogeneizzazione.*

(a) Il rango di F è un invariante per equivalenza (per isometrie). Diciamo che f è non degenere *se F è non degenere.*

(b) La coppia delle segnature $(s_+(F), s_-(F))$ è un invariante per equivalenza a meno di riordinamento*, cioè se $f \sim g$ allora*

$$(s_+(F), s_-(F)) = (s_+(G), s_-(G)),$$

oppure

$$(s_+(F), s_-(F)) = (s_-(G), s_+(G))$$

(c) La coppia delle segnature di $F(0, x_1, \dots, x_n)$ a meno di riordinamento *è un invariante per equivalenza.*

Dimostrazione. (a): Il rango di F non dipende dalle coordinate scelte, e non cambia se moltiplichiamo F per uno scalare non nullo.

(b): La segnatura di F non dipende dalle coordinate scelte, e se moltiplichiamo F per un λ non nullo le segnature rimangono invariate se $\lambda > 0$ e si scambiano tra di loro se $\lambda < 0$.

(c): Vale perchè la matrice M in (10.4.2) manda il sottospazio $x_0 = 0$ in se stesso. □

Daremo la classificazione a meno di isometrie delle $f \colon \mathbb{S} \to \mathbb{R}$ di grado 2 non degeneri. Questo perchè la classificazione delle $f \colon \mathbb{S} \to \mathbb{R}$ di grado 2 degeneri si riconduce alla classificazione delle forme quadratiche modulo il gruppo ortogonale di $\mathsf{V}(\mathbb{S})$ o a quella delle $f \colon \mathbb{T} \to \mathbb{R}$ di grado 2 non degeneri dove $\dim \mathbb{T} < \dim \mathbb{S}$. Più precisamente supponiamo che $f \colon \mathbb{S} \to \mathbb{R}$ sia di grado 2 e degenere. Allora in un opportuno sistema di riferimento ON di coordinate $X \colon \mathbb{S} \xrightarrow{\sim} \mathbb{E}^n(\mathbb{R})$ vale una delle seguenti possibiltà:

(a) $f(X^{-1})$ (cioè la f scitta nelle coordinate X) è un polinomio *omogeneo* di grado 2.

1. $f(X^{-1})$ è un polinomio (in generale nonomogeneo) di grado 2 in $n-1$ (o meno) delle coordinate $x_1, \dots, x_n$.

Nel caso (a) (coni su iperquadriche) ci si riconduce alla classificazione delle forme quadratiche su $\mathsf{V}(\mathbb{S})$ modulo il gruppo ortogonale di $\mathsf{V}(\mathbb{S})$, nel caso (b) si itera la procedura.

Teorema 10.4.2. *Sia $\mathbb{S}$ uno spazio affine euclideo di dimensione n. Sia $f\colon \mathbb{S} \to \mathbb{R}$ una funzione polinomiale di grado 2 non degenere. Allora esiste un sistema di riferimento ON con coordinate $(x_1, \dots, x_n)$ tale che in queste coordinate f sia uguale a uno delle seguenti forme canoniche:*

$$E_i(x_1, \dots, x_n) := \frac{x_1^2}{a_1^2} + \dots + \frac{x_i^2}{a_i^2} - \frac{x_{i+1}^2}{a_{i+1}^2} - \dots - \frac{x_n^2}{a_n^2} + 1, \quad a_i > 0$$

dove $i \in \{0, \dots, n\}$, oppure a

$$P_j(x_1, \dots, x_n) := \frac{x_1^2}{a_1^2} + \dots + \frac{x_j^2}{a_j^2} - \frac{x_{j+1}^2}{a_{j+1}^2} - \dots - \frac{x_{n-1}^2}{a_{n-1}^2} + x_n, \quad a_j > 0,$$

dove $0 \leqslant j \leqslant \left\lfloor \frac{n-1}{2} \right\rfloor$. Inoltre siano $f, g\colon \mathbb{S} \to \mathbb{R}$ due funzioni polinomiali di grado 2 non degeneri. Allora f è equivalente per isometrie a g se e solo se hanno la stessa forma canonica (cioè sono entrambe E_i, o entrambe E_j, con gli stessi parametri a meno degli ovvi riordinamenti.

Dimostrazione. Si procede come nel caso $n = 2$.

Passo 1. Sia

$$f(x_1, \dots, x_n) = q(x_1, \dots, x_n) + l(x_1, \dots, x_n) + c,$$

dove q è una forma quadratica, l è polinomio omogeneo di grado 1 (funzione lineare), e $c \in \mathbb{R}$. L'omogeneizzazione di f è data da

$$F(x_0, \dots, x_n) = q(x_1, \dots, x_n) + x_0 \cdot l(x_1, \dots, x_n) + cx_0^2,$$

e quindi

$$q(x_1, \dots, x_n) = F(0, x_1, \dots, x_n).$$

Siano $\widetilde{q} \in \mathrm{Bil}^+(\mathbb{R}^n)$ e $\widetilde{F} \in \mathrm{Bil}^+(\mathbb{R}^{n+1})$ le polarizzazioni di q e F rispettivamente, e siano

$$\mathbb{R}^{n+1} \xrightarrow{L_{\widetilde{F}}} (\mathbb{R}^{n+1})^\vee, \quad \mathbb{R}^n \xrightarrow{L_{\widetilde{q}}} (\mathbb{R}^n)^\vee$$

le corrispondenti applicazioni lineari simmetriche definite da $\widetilde{F}$ e $\widetilde{q}$, cioè

$$L_{\widetilde{F}}(X)(Y) = \widetilde{F}(X, Y), \quad L_{\widetilde{q}}(X)(Y) = \widetilde{q}(X, Y).$$

Dimostriamo che q ha rango almeno $n - 1$, cioè

(a) q è non degenere, oppure

(b) q ha rango $n - 1$.

Siccome la forma quadratica q (e la sua polarizzazione $\widetilde{q}$) è definita sul sottospazio $x_0 = 0$, abbiamo

$$(0, x_1, \dots, x_n) \in \ker q \Leftrightarrow L_{\widetilde{F}}(0, x_1, \dots, x_n) \in \mathrm{Span}\{x_0^\vee\},$$

dove $\{x_0^\vee, \dots, x_n^\vee\}$ è la base di $(\mathbb{R}^{n+1})^\vee$ duale della base standard. Quindi restringendo $L_{\widetilde{F}}$ a $\ker q$ otteniamo un'applicazione lineare

$$\begin{array}{ccc} \ker q & \longrightarrow & \mathrm{Span}\{x_0^\vee\} \\ (0, x_1, \dots, x_n) & \mapsto & L_{\widetilde{F}}(0, x_1, \dots, x_n) \end{array}$$

Se q avesse rango minore di $n - 1$, allora $\dim \ker q \geqslant 2$, e seguirebbe che esiste $0 \neq (0, x_1, \dots, x_n)$ nel nucleo di $L_{\widetilde{F}}$. Questo contraddice l'ipotesi che F sia non degenere.

Passo 2. Dimostriamo che se vale (a), cioè q è nondegenere, allora f è equivalente per isometrie a una E_i. Per il Teorema spettrale esiste una base ON $\mathcal{B}$ che diagonalizza q. Segue che f è equivalente a

$$g(x_1, \dots, x_n) = \sum_{i=1}^n \lambda_i x_i^2 + \sum_{i=1}^n \mu_i x_i + c,$$

dove $\lambda_i \neq 0$ per ogni i. "Completando i quadrati" riscriviamo g come

$$g(x_1, \dots, x_n) = \sum_{i=1}^n \lambda_i \left(x_i + \frac{\mu_i}{2\lambda_i} \right)^2 - \sum_{i=1}^n \frac{\mu_i^2}{4\lambda_i^2} + c.$$

Quindi vediamo che g è equivalente a

$$h(x_1, \ldots, x_n) = \sum_{i=1}^{n} \lambda_i x_i^2 + d.$$

Siccome per ipotesi l'omogeneizzazione $F(x_0, \ldots, x_n)$ è non degenere, $d \neq 0$. Moltiplicando per d^{-1} e riordinando le coordinate otteniamo che f è equivalente a una E_i.

Passo 3. Dimostriamo che se vale (b), cioè q ha rango $n-1$, allora f è equivalente a una P_j con

$$0 \leqslant j \leqslant \left\lfloor \frac{n-1}{2} \right\rfloor. \tag{10.4.3}$$

La dimostrazione è simile a quella del Passo 2. Il motivo per cui possiamo assumere che valga (10.4.3) è che le forme quadratiche

$$\frac{x_1^2}{a_1^2} + \ldots + \frac{x_i^2}{a_i^2} - \frac{x_{i+1}^2}{a_{i+1}^2} - \ldots - \frac{x_{n-1}^2}{a_{n-1}^2} + x_n$$

e

$$\frac{x_1^2}{a_1^2} + \ldots + \frac{x_{n-1-i}^2}{a_{n-1-i}^2} - \frac{x_{n-i}^2}{a_{n-i}^2} - \ldots - \frac{x_{n-1}^2}{a_{n-1}^2} + x_n$$

sono equivalenti - basta moltiplicare la seconda per (-1), e poi cambiare segno alla coordinata x_n.

Passo 4. Quali delle E_i o P_j sono equivalenti? Considerando le segnature di F e q si trova che una E_i è equivalente solo ad altre E_i e così per le P_j. Infine le equivalenze tra E_i o tra P_j si ottengono solo nel modo ovvio, cioè riordinando gli a_k. □

Quadriche

Esamineremo le quadriche non degeneri, cioè le superfici definite da $V(f) \subset \mathbb{S}$ dove $\mathbb{S}$ è uno spazio affine euclideo di dimensione 3, e $f\colon \mathbb{S} \to \mathbb{R}$ è una funzione polinomiale (non nulle) di grado 2 tale che la sua omogeneizzazione (per un sistema di coordinate qualsiasi) si una forma quadratica (in 4 variabili) non degenere.

Per il Teorema 10.4.2, una quadrica non degenere è data, nelle coordinate (x, y, z) di un opportuno sistema di riferimento ON, da una delle seguenti equazioni:

$$\mathcal{Q}_1 : \ \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,$$

$$\mathcal{Q}_2 : \ \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1,$$

$$\mathcal{Q}_3 : \ \frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1,$$

$$\mathcal{Q}_4 : \ -\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1,$$

$$\mathcal{Q}_5 : \ z = \frac{x^2}{a^2} + \frac{y^2}{b^2},$$

$$\mathcal{Q}_6 : \ z = \frac{x^2}{a^2} - \frac{y^2}{b^2}.$$

$\mathcal{Q}_1$ è un *ellissoide*. In [5] trovate una costruzione dell'ellissoide simile a quella di ellissi e iperboli.

$\mathcal{Q}_2$ è un *iperboloide a una falda*, e contiene due "schiere" di rette. Per vederlo riscriviamo l'equazione di $\mathcal{Q}_2$ così:

$$\left(\frac{x}{a} + \frac{z}{c}\right) \cdot \left(\frac{x}{a} - \frac{z}{c}\right) = \left(1 + \frac{y}{b}\right) \cdot \left(1 - \frac{y}{b}\right).$$

La prima schiera è formata dalle rette di equazioni

$$\begin{aligned} \lambda \left(\frac{x}{a} + \frac{z}{c}\right) &= \mu \left(1 + \frac{y}{b}\right), \\ \mu \left(\frac{x}{a} - \frac{z}{c}\right) &= \lambda \left(1 - \frac{y}{b}\right), \end{aligned}$$

per $(\lambda, \mu) \neq (0,0)$. Notate che se moltiplichiamo λ, μ per uno stesso fattore non nullo otteniamo la stessa retta. La seconda schiera è formata dalle rette di equazioni

$$\begin{aligned} \lambda\left(\frac{x}{a} + \frac{z}{c}\right) &= \mu\left(1 - \frac{y}{b}\right), \\ \mu\left(\frac{x}{a} - \frac{z}{c}\right) &= \lambda\left(1 + \frac{y}{b}\right), \end{aligned}$$

per $(\lambda, \mu) \neq (0,0)$. L'iperboloide a una falda si chiama anche *iperboloide iperbolico*. L'aggettico "iperbolico" si riferisce alla curvatura (Gaussiana) dei suoi punti, che è negativa (whatever that means).

$\mathcal{Q}_3$ è un *iperboloide a due falde*. Dall'equazione segue che $x \leqslant -1$ o $x \geqslant 1$; per questo le due falde. È anche noto come *iperboloide ellittico*, dove "ellittico" si riferisce alla curvatura (Gaussiana) dei suoi punti, che è positiva.

$\mathcal{Q}_4$ è l'insieme vuoto.

$\mathcal{Q}_5$, che ha equazione

$$z = \frac{x^2}{a^2} + \frac{y^2}{b^2},$$

è un paraboloide ellittico. Ha la forma di una coppa infinita. Se $a = b$ è spazzata da una parabola che ruota introno al sua asse.

$\mathcal{Q}_6$, che ha equazione è un paraboloide iperbolico, e contiene due schiere di rette, come l'iperboloide iperbolico. Riscrivendo l'equazione così

$$z = \left(\frac{x}{a} + \frac{y}{b}\right) \cdot \left(\frac{x}{a} - \frac{y}{b}\right),$$

vediamo che le rette di una schiera hanno equazioni cartesiane

$$\begin{aligned} \left(\frac{x}{a} + \frac{y}{b}\right) &= \mu, \\ \mu\left(\frac{x}{a} - \frac{y}{b}\right) &= z, \end{aligned}$$

per $\mu \in \mathbb{R}$, e le rette dell'altra schiera le equazioni

$$\begin{aligned} \left(\frac{x}{a} - \frac{y}{b}\right) &= \mu, \\ \mu\left(\frac{x}{a} + \frac{y}{b}\right) &= z. \end{aligned}$$

Esempio 10.4.3. Sia $\mathcal{Q} \subset \mathbb{E}^3(\mathbb{R})$ la quadrica di equazione

$$xy - 6xz - z^2 + x + 2 = 0.$$

Cerchiamo di capire come sia fatta $\mathcal{Q}$. La forma quadratica associata al polinomio di grado 2

$$f(x,y,z) := xy - 6xz - z^2 + x + 2$$

è

$$F(w,x,y,z) := xy - 6xz - z^2 + wx + 2w^2.$$

La matrice simmetrica associata è

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

Calcolando troviamo che $\det M = 2$. Quindi F è non degenere, e perciò Q è equivalente a una delle quadriche non degeneri che abbiamo esaminato. Per decidere se $\mathcal{Q}$ è un ellissoide, o un iperboloide iperbolico, etc. calcoliamo le segnature di F e della forma quadratica $F(0,x,y,z)$.

Per $i \in \{1,2,3,4\}$ sia $M(i)$ la matrice (simmetrica) cone entrate $m_{s,t}$ per $1 \leqslant s,t \leqslant i$, dove $M = (m_{s,t})$. Abbiamo

$$\det M(1) = 2 > 0, \quad \det M(2) = -1 < 0,$$

$$\det M(3) = -2 < 0, \quad \det M(4) = \det M = 2 > 0.$$

Per un criterio che conosciamo segue che che

$$s_+(F) = 2, \quad s_-(F) = 2.$$

Nella nostra lista $\mathcal{Q}_1, \dots, \mathcal{Q}_6$ le quadriche tali che la forma quadratica associata ha segnature $(2,2)$ sono $\mathcal{Q}_2$ (iperboloide iperbolico) e $\mathcal{Q}_6$ (paraboloide iperbolico). Quindi per decidere a quale dei due sia equivalente $\mathcal{Q}$ basta determinare il rango della forma quadratica

$$F(0,x,y,z) = xy - 6xz - z^2.$$

Siccome è una forma quadratica non degenere, $\mathcal{Q}$ è un iperboloide iperbolico.

Esercizi del Capitolo 10

Esercizio 10.1. *La curva in $\mathbb{E}^2$ definita dall'equazione cartesiana*

$$\mathcal{C}: \ 73x^2 + 72xy + 52y^2 + 30x - 40y - 75 = 0$$

è un'ellisse.

1. *Scrivere la forma quadratica in 3 variabili associata al polinomio di grado 2 che definisce $\mathcal{C}$.*
2. *Dimostrare che $\mathcal{C}$ è un'ellisse calcolando 3 determinanti.*
3. *Calcolare (con maggiore fatica) i fuochi di $\mathcal{C}$, e la somma costante delle distanze dei punti di $\mathcal{C}$ dai suoi fuochi.*

Esercizio 10.2. *Per $t \in \mathbb{R}$ sia $\mathcal{C}_t \subset \mathbb{E}^2(\mathbb{R})$ la curva di equazione*

$$\mathcal{C}_t: \ x_1^2 + 6x_1x_2 + tx_2^2 - 2x_1 + 4x_2 + 3 = 0.$$

Determinate com'è fatta $\mathcal{C}_t$ al variarare di t, cioè per quali valori di t è un'ellisse, per quali un'iperbole, etc.

Esercizio 10.3. *Sia $\mathcal{C} \subset \mathbb{E}^2(\mathbb{R})$ la curva di equazione*

$$\mathcal{C}: \ x_1^2 + 2x_1x_2 + x_2^2 - 2x_1 - 6x_2 + 3 = 0.$$

(a) Dimostrare che $\mathcal{C}$ è una parabola calcolando due determinanti.

(b) Determinate fuoco e direttrice di $\mathcal{C}$.

Esercizio 10.4. *Siano $\mathcal{Q}_1, \mathcal{Q}_2, \mathcal{Q}_3 \subset \mathbb{E}^3$ le quadriche di equazioni cartesiane*

$$\mathcal{Q}_1: \ x^2 + xy + 2y^2 - z^2 + 3x - z + 25 = 0,$$

$$\mathcal{Q}_2: \ 2xy + 4xz + y^2 - 4z^2 + 2y + 4z + 3 = 0$$

e

$$\mathcal{Q}_3: \ 2xy + 4xz + y^2 - 4z^2 + 2y + 5z + 3 = 0.$$

Quale tra $\mathcal{Q}_1, \mathcal{Q}_2, \mathcal{Q}_3$ è un paraboloide?

Capitolo 11

Geometria proiettiva

11.1 Introduzione

La geometria proiettiva è legata alla prospettiva. In una riproduzione grafica di un pavimento, i punti all'"infinito" del pavimento vengono rappresentati con i punti di una retta "orizzonte", e di conseguenza rette parallele nel pavimento si intersecano in punti della retta orizzonte. Ogni volta che, guardando un'immagine ricostruiamo nella nostra mente l'oggetto riprodotto, facciamo della geometria proiettiva. Un piano proiettivo contiene un piano affine, e si può ottenere aggiungendovi "punti all'infinito". In questo modo si acquista una maggiore simmetria. Un esempio: due punti distinti di un piano affine sono incidenti (cioè contenuti in) una e una sola retta, ma non è vera l'affermazione che due rette distinte sono incidenti (cioè contengono) uno e un solo punto perchè possono essere parallele. In un piano proiettivo rimane vera la prima affermazione e in più due rette distinte sono incidenti uno e un solo punto.Per dare il sapore della geometria proiettiva prima ancora di dare definizioni, enunciamo il Teorema di Pappo: se r, r' sono rette distinte in un piano, $A, B, C \in r$ sono punti a due a due distinti (nessuno dei quali nell'intersezione $r \cap r'$) e analogamente $A', B', C' \in r'$ sono punti a due a due distinti (nessuno dei quali nell'intersezione $r \cap r'$), allora i punti di intersezione

$$BC' \cap B'C,\ AC' \cap A'C,\ AB' \cap BA'$$

sono allineati (BC' è la retta per B e C' etc.). Il teorema vale se r, r' si intersecano, ma anche se sono parallele. Analogamente, vale anche se l'intersezione $BC' \cap B'C$ è vuota, cioè le rette $BC', B'C$ sono parallele; in questo caso si intende che una retta contiene "l'intersezione" $BC' \cap B'C$ se è parallela alle due rette.

In questo capitolo gli spazi vettoriali sono sempre finitamente generati.

11.2 Spazi proiettivi

La definizione

Definizione 11.2.1. Sia V uno spazio vettoriale. Il *proiettificato* di V è l'insieme dei sottospazi vettoriali di dimensione 1 di V, e si denota $\mathbb{P}(V)$. In simboli

$$\mathbb{P}(V) := \{L \subset V \mid L \subset V \text{ sottospazio, } \dim L = 1\}.$$

La *dimensione di* $\mathbb{P}(V)$, denotata $\dim \mathbb{P}(V)$, è uguale a $(\dim V - 1)$. Uno *spazio proiettivo su* $\mathbb{K}$ è il proiettificato di uno spazio vettoriale con campo degli scalari $\mathbb{K}$.

Una *retta proiettiva* è uno spazio proiettivo di dimensione 1, e un *piano proiettivo* è uno spazio proiettivo di dimensione 2. Denotiamo con $\mathbb{P}^n(\mathbb{K})$ il proiettificato di $\mathbb{K}^{n+1}$, cioè

$$\mathbb{P}^n(\mathbb{K}) = \mathbb{P}(\mathbb{K}^{n+1}).$$

Esaminiamo $\mathbb{P}^n_{\mathbb{K}}$. Denotiamo gli elementi di $\mathbb{K}^{n+1}$ con $(x_0, x_1, \ldots, x_n)$, cioè l'indice va da 0 a n anzichè da 1 a $n+1$. Per determinare un elemento di $\mathbb{P}^n(\mathbb{K})$, cioè un sottospazio vettoriale $L \subset \mathbb{K}^{n+1}$ di dimensione 1 è sufficiente dare un vettore non nullo $X = (x_0, \ldots, x_n) \in \mathbb{K}^{n+1}$, e porre $L = \text{Span}\{X\}$. Tradizionalmente il sottospazio generato da $X = (x_0, x_1, \ldots, x_n)$ si denota con

$$[X] = [x_0, x_1, \ldots, x_n], \tag{11.2.1}$$

(in molti testi si denota $[x_0 : x_1 : \ldots : x_n]$) anzichè con $\mathrm{Span}\{X\}$, e noi adotteremo questa convenzione. Se $x_0 \neq 0$ possiamo "normalizzare" X ponendo $x_0 = 1$, e allora rimane la scelta (libera) delle entrate $(x_1, \ldots, x_n)$. Il sottoisieme $\mathbb{P}^n(\mathbb{K})$ i cui elementi sono i sottospazi generati dai vettori X con $x_0 = 0$ è lo spazio proiettivo $\mathbb{P}^{n-1}(\mathbb{K})$. In conclusione abbiamo una decomposizione

$$\mathbb{P}^n(\mathbb{K}) = \{[1, x_1, \ldots, x_n] \mid (x_1, \ldots, x_n) \in \mathbb{K}^n\} \sqcup \mathbb{P}^{n-1}(\mathbb{K}),$$

e una ovvia identificazione con $\mathbb{K}^n$ del primo sottoinsieme in cui abbiamo decomposto $\mathbb{P}^n(\mathbb{K})$. Seguendo la tradizione diciamo che $\mathbb{P}^n(\mathbb{K})$ è ottenuto dallo spazio affine $\mathbb{A}^n(\mathbb{K})$ aggiungendo i punti all'infinito, cioè i punti di $\mathbb{P}^{n-1}(\mathbb{K})$.

Diamo un'idea più geometrica dei punti all'"infinito".

$\mathbb{P}^1(\mathbb{K})$: c'è un solo punto all'infinito, cioè $[0, 1]$. Sia $\mathbb{K} = \mathbb{R}$ e immaginiamo di "camminare" sulla retta affine $\mathbb{R} \subset \mathbb{P}^1(\mathbb{R})$, per esempio secondo la formula $x_1 = lt + a$, dove t è la variabile tempo e $l, a \in \mathbb{R}$ sono costanti, con $l \neq 0$. Nella notazione di (11.2.1) il punto variabile è $P(t) = [1, lt + a]$. Se $t \to \infty$ il punto $P(t)$ non ha un limite nella retta $\mathbb{R}$, ma lo ha in $\mathbb{P}^1(\mathbb{R})$ perchè

$$P(t) = [1, lt + a] = [t^{-1}, l + at^{-1}],$$

e $[t^{-1}, l + at^{-1}]$ tende a $[0, l] = [0, 1]$ per $t \to \pm\infty$. La conclusione è che se camminate per sempre su $\mathbb{R}$ verso destra o verso sinistra, arriverete (al limite) allo stesso punto all'infinito $[0, 1]$. Questo ha senso se pensate alla riproduzione grafica di una retta su un quadro.

$\mathbb{P}^2(\mathbb{K})$: Poniamo $P(t) = [1, lt + a, mt + b]$, dove $(l, m) \neq (0, 0)$. Se $\mathbb{K} = \mathbb{R}$, e $t \to \infty$ il punto $P(t)$ non ha un limite nel piano affine $\mathbb{A}^2(\mathbb{R})$, ma lo ha in $\mathbb{P}^2(\mathbb{R})$ perchè

$$[1, lt + a, mt + b] = [t^{-1}, l + at^{-1}, m + bt^{-1}] \to [0, l, m].$$

Conclusione: $P(t)$ ha come limite il punto della retta all'infinito (che, ricordiamo, è $\mathbb{P}^1_{\mathbb{R}}$) dato da $[0, l, m]$. (Si può dare senso a quello che abbiamo detto anche per $\mathbb{P}^2(\mathbb{K})$ con $\mathbb{K}$ arbitrario.)

Spazi proiettivi e spazi affini

Diamo una versione più generale di quello che abbiamo detto sui punti all'"infinito". Sia V uno spazio vettoriale, e sia $W \subset V$ un sottospazio di codimensione 1, cioè tale che $\dim W = (\dim V - 1)$. Diamo a $\mathbb{P}(V)\backslash\mathbb{P}(W)$ una struttura di spazio affine con spazio delle traslazioni $\mathrm{Hom}(V/W, W)$, procedendo come segue. Definiamo

$$\begin{array}{ccc} (\mathbb{P}(V)\backslash\mathbb{P}(W)) \times \mathrm{Hom}(V/W, W) & \xrightarrow{T} & (\mathbb{P}(V)\backslash\mathbb{P}(W)) \\ ([v], f) & \mapsto & [v + f(\overline{v})] \end{array} \tag{11.2.2}$$

dove $\overline{v} \in (V/W)$ è la classe di equivalenza di v. Notiamo che l'applicazione in (11.2.2) è ben definita. Infatti

1. $v + f(\overline{v}) \neq 0$ perchè $f(\overline{v}) \in W$ e $v \in (V\backslash W)$, e
2. $[v + f(\overline{v})]$ non dipende dal vettore v che genera il sottospazio, perchè se v' è un altro generatore allora $v' = \mu v$ e quindi
$$[\mu v + f(\overline{\mu v})] = [\mu v + \mu f(\overline{v})] = [v + f(\overline{v})].$$
perchè f è lineare

Inoltre

1. è chiaro che $T(0) = \mathrm{Id}$,
2. si ha $T(f + g) = T(f) \circ T(g)$ perchè
$$T(f + g)[v] = [v + (f + g)(\overline{v})] = [v + g(\overline{v}) + f(\overline{v})],$$
e
$$T(f)(T(g)[v]) = T(f)([v + g(\overline{v})] = [v + g(\overline{v}) + f(\overline{v})],$$
dove l'ultima uguaglianza segue dal fatto che $g(\overline{v}) \in W$ e quindi la sua classe di equivalenza in V/W è zero,
3. $T(f)[v] = [v]$ solo se $f = 0$ perchè $T(f)[v] = [v]$ solo se $v + f(\overline{v})$ è un multiplo di v, e questo implica che $f(\overline{v}) = 0$ perchè $v \notin W$, cioè $f = 0$,

4. dato $[v'] \in (\mathbb{P}(V)\backslash\mathbb{P}(W))$ esiste $f \in \mathrm{Hom}(V/W, W)$ tale che $T(f)[v] = [v']$ perchè, siccome $\{\overline{v}\}$ è una base di V/W esiste $f \in \mathrm{Hom}(V/W, W)$ tale che $f(\overline{v}) = v' - v$ e per questa f vale $T(f)[v] = [v']$.

La conclusione è che l'applicazione in (11.2.2) dà a $(\mathbb{P}(V)\backslash\mathbb{P}(W))$ una struttura di spazio affine con spazio di traslazioni $\mathrm{Hom}(V/W, W)$. Abbiamo la decomposizione

$$\mathbb{P}(V) = (\mathbb{P}(V)\backslash\mathbb{P}(W)) \sqcup \mathbb{P}(W))$$

dove $(\mathbb{P}(V)\backslash\mathbb{P}(W))$ è "quasi tutto" $\mathbb{P}(V)$, e questo giustifica la definizione di $\dim \mathbb{P}(V)$ data in (11.2.1).

Coordinate omogenee

Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione $n+1$, e sia $\mathcal{B} = \{v_0, \ldots, v_n\}$ una sua base. Sia $\mathcal{B}^\vee = \{v_0^\vee, \ldots, v_n^\vee\}$ la base duale di $V^\vee$. Denotiamo $v_i^\vee$ anche con x_i. In altre parole, se $v \in V$ si ha

$$v = x_0(v)v_0 + \ldots + x_n(v)v_n,$$

e $x_0(v), \ldots, x_n(v)$ sono le coordinate di v nella base $\mathcal{B}$. Ora sia $[v] \in \mathbb{P}(V)$. Le *coordinate omogenee* di $[v]$ (associate alla base $\mathcal{B}$) sono $x_0(v), \ldots, x_n(v)$. Le coordinate omogenee del punto $[v] \in \mathbb{P}(V)$ sono determinate solo a meno di un fattore non nullo comune $\mu \in \mathbb{K}^*$, perchè $[\mu v] = [v]$. Per questo motivo indichiamo le coordinate omogenee di $[v]$ così:

$$[x_0(v), \ldots, x_n(v)] \in \mathbb{P}^n(\mathbb{K}).$$

In altre parole le coordinate rispetto a una base danno un'applicazione biunivoca

$$\begin{array}{ccc} \mathbb{P}(V) & \longrightarrow & \mathbb{P}^n(\mathbb{K}) \\ [v] & \mapsto & [x_0(v), \ldots, x_n(v)] \end{array}$$

Rivisitiamo, usando le coordinate omogenee, la struttura di spazio affine su $\mathbb{P}(V)\backslash\mathbb{P}(W)$ per $W \subset V$ di codimensione 1. Supponiamo che

$$W = \ker(x_0) = \left\{ \sum_{i=1}^{n} x_i v_i \right\}.$$

Se si esamina la definizione data, si vede che le coordinate affini sono date da

$$\begin{array}{ccc} \mathbb{P}(V)\backslash\mathbb{P}(\ker(x_0)) & \longrightarrow & \mathbb{K}^n \\ [v] & \mapsto & \left(\frac{x_1(v)}{x_0(v)}, \ldots, \frac{x_n(v)}{x_0(v)}\right) \end{array}$$

Se $f\colon V \to \mathbb{K}$ è un'applicazione lineare non nulla, poniamo

$$\mathbb{P}(V)_f := \mathbb{P}(V)\backslash\mathbb{P}(\ker f). \tag{11.2.3}$$

Nel caso in cui $\mathbb{P}(V) = \mathbb{P}^n(\mathbb{K})$ e non ci sono equivoci su quale sia il campo $\mathbb{K}$, denotiamo $\mathbb{P}^n(\mathbb{K})_f$ con $\mathbb{P}^n_f$.

Spesso si denota $x_i(v)/x_0(v)$ semplicemente $x_i(v)$ - possiamo pensare di normalizzare le coordinate omogenee ponendo $x_0(v) = 1$. Spesso per evitare questa potenziale fonte di malintesi si denotano con

$$[X_0(v), \ldots, X_n(v)]$$

(lettera X maiuscola) le coordinate omogenee, e quindi $x_i(v) = X_i(v)/X_0(v)$.

Sottospazi proiettivi

Definizione 11.2.2. Un sottoinsieme H di $\mathbb{P}(V)$ è un *sottospazio proiettivo* se $H = \mathbb{P}(W$, dove $W \subset V$ è un sottospazio vettoriale.

Esempio 11.2.3. Sia $W \subset V$ di codimensione 1. Allora W è dato da un'equazione cartesiana

$$W : \sum_{i=0}^{n} a_i x_i(v) = 0, \quad (a_0, \ldots, a_n) \neq (0, \ldots, 0).$$

Quella scritta sopra è anche un'*equazione cartesiana* del sottospazio proiettivo $H \subset \mathbb{P}(V)$, cioè

$$\mathbb{P}(W) = \{[v] \in \mathbb{P}(V) \mid \sum_{i=0}^{n} a_i x_i(v) = 0\}.$$

Osservazione 11.2.4. L'equazione cartesiana di $H \subset \mathbb{P}(V)$ ha senso perchè equivale a chiedere che una funzione polinomiale omogenea nelle coordinate omogenee sia nulla, e quindi se si annulla per una scelta di coordinate omogenee di un punto allora si annulla per ogni altra scelta di coordinate omogenee dello stesso punto. Non avrebbe senso se chiedessimo di annulare un polinomio non omogeneo.

Definizione 11.2.5. Sia $\mathbb{P}$ uno spazio proiettivo. Un *iperpiano di* $\mathbb{P}$ è un sottospazio proiettivo $L \subset \mathbb{P}$ di codimensione 1, cioè tale che $\dim L = (\dim \mathbb{P} - 1)$.

Esempio 11.2.6. Sia $\mathbb{P}(V)$ un piano proiettivo. Se L_1, L_2 sono rette distinte di $\mathbb{P}(V)$, allora $L_1 \cap L_2$ consiste di un singolo punto. Infatti per ipotesi $L_i = \mathbb{P}(W_i)$ dove $W_i \subset V$ ha dimensione 2, e $W_1 \neq W_2$. Quindi V, che ha dimensione 3, è generato da W_1 e W_2, e perciò Grassmann dà che $\dim(W_1 \cap W_2) = 1$. Quindi $W_1 \cap W_2$ è l'unico punto nell'intersezione $L_1 \cap L_2$.

Descriviamo $\mathbb{H} := H \cap \mathbb{P}(V)_{x_0}$. Ricordiamo che $\mathbb{P}(V)_{x_0} = \mathbb{P}(V) \backslash \mathbb{P}(\ker x_0)$ e che l'identificazione

$$\begin{array}{ccc} \mathbb{K}^n & \longrightarrow & \mathbb{P}(V)\backslash\mathbb{P}(\ker(x_0)) \\ (x_1, \ldots, x_n) & \mapsto & [v_0 + x_1 v_1 + \ldots + x_n v_n] \end{array}$$

dà coordinate affini su $\mathbb{P}(V)_{x_0}$. Ponendo $x_0 = 1$ nell'equazione cartesiana di H otteniamo un'equazione cartesiana di $\mathbb{H}$:

$$a_0 + a_1 x_1 + \ldots + a_n x_n = 0.$$

Quindi

1. Se $H \neq \mathbb{P}(\ker x_0)$, allora $H \cap \mathbb{P}(V)_{x_0}$ è un sottospazio affine di codimensione 1.
2. Se $H = \mathbb{P}(\ker x_0)$, allora $H \cap \mathbb{P}(V)_{x_0} = \varnothing$.

Più in generale, se $H \subset \mathbb{P}(V)$ è un sottospazio lineare, allora $H \cap \mathbb{P}(V)_{x_0}$ è un sottospazio affine della stessa dimensione di H, a meno che $H \subset \mathbb{P}(\ker x_0)$ (nel qual caso $H \cap \mathbb{P}(V)_{x_0} = \varnothing$).

Esiste un "viceversa". Sia $\mathbb{H} \subset \mathbb{P}(V)_{x_0}$ un sottospazio affine. Esiste uno e un solo sottospazio proiettivo $H \subset \mathbb{P}(V)$ tale che $H \cap \mathbb{P}(V)_{x_0} = \mathbb{H}$. Vediamo perchè. Se $\mathbb{H} = \mathbb{P}(V)_{x_0}$ l'affermazione è banale, si ha $H = \mathbb{P}(V)$. Supponiamo che $\dim \mathbb{H} = n - 1$, e sia

$$a_0 + a_1 x_1 + \ldots + a_n x_n = 0$$

una sua equazione cartesiana, dove $(a_1, \ldots, a_n) \neq (0, \ldots, 0)$ (altrimenti $\mathbb{H}$ sarebbe tutto $\mathbb{P}(V)_{x_0}$). L'unico sottospazio proiettivo $H \subset \mathbb{P}(V)$ con le proprietà richieste ha equazione cartesiana

$$a_0 x_0 + a_1 x_1 + \ldots + a_n x_n = 0.$$

Se $\dim \mathbb{H} < n - 1$, si procede in modo analogo, l'unica differenza è che $\mathbb{H}$ è definito da $n - \dim \mathbb{H}$ equazioni lineari, con associato sistema di equazioni lineari omogenee di rango massimo.

Definizione 11.2.7. Sia $\mathbb{H} \subset \mathbb{P}(V)_{x_0}$ un sottospazio affine. L'unico sottospazio proiettivo $H \subset \mathbb{P}(V)$ tale che

$$H \cap \mathbb{P}(V)_{x_0} = \mathbb{H}$$

è la *chiusura proiettiva di* $\mathbb{H}$.

Esempio 11.2.8. Siano $\mathbb{L}_1, \mathbb{L}_2 \subset \mathbb{P}^2_{x_0}$ le rette di equazioni affini

$$3x_1 + 2x_2 - 3 = 0, \quad 6x_1 + 4x_2 + 1 = 0,$$

e siano $L_1, L_2 \subset \mathbb{P}^2_{x_0}$ le loro chiusure proiettive. Per l'Esempio 11.2.6 sappiamo che $L_1 \cap L_2$ consiste di un singolo punto, determiniamo quale punto è. Le equazioni omogenee di L_1, L_2 sono date da

$$L_1 : \ 3x_1 + 2x_2 - 3x_0 = 0, \quad L_2 : \ 6x_1 + 4x_2 + x_0 = 0.$$

Risolvendo, troviamo che

$$L_1 \cap L_2 = \{[0, 2, -3]\}.$$

In altre parole: $\mathbb{L}_1, \mathbb{L}_2$ sono rette parallele, e quindi non "si incontrano", invece L_1, L_2 si incontrano nel punto all'infinito $[0, 2, -3]$.

Proposizione 11.2.9. *Sia $\{L_i\}_{i\in I}$ è una collezione di sottospazi proiettivi di uno spazio proiettivo $\mathbb{P}$. L'intersezione*

$$\bigcap_{i\in I} L_i$$

è un sottospazio proiettivo.

Dimostrazione. Sia $\mathbb{P} = \mathbb{P}(V)$. Per definizione $L_i = \mathbb{P}(W_i)$ dove $W_i \subset V$ è un sottospazio vettoriale. Quindi $W := \bigcap_{i\in I} W_i$ è un sottospazio vettoriale di V. Un punto $[v] \in \mathbb{P}(V)$ appartiene all'intersezione degli L_i se e solo se $v \in W$, perciò $\bigcap_{i\in I} L_i = \mathbb{P}(W)$. □

Definizione 11.2.10. Sia X un sottoinsieme di uno spazio proiettivo $\mathbb{P}$. Il sottospazio proiettivo *generato da* X è l'intersezione dei sottospazi proiettivi di $\mathbb{P}$ che contengono X. Lo denotiamo $\mathrm{Span}(X)$.

Chiaramente $\mathrm{Span}(X)$ è il più piccolo sottospazio proiettivo di $\mathbb{P}(V)$ contenente X, dove "Più piccolo" significa che è contenuto in ogni sottospazio proiettivo di $\mathbb{P}$ contenente X. Se $X = \{p_0, \dots, p_d\}$ usiamo la notazione $\mathrm{Span}(p_0, \dots, p_d)$ anzichè $\mathrm{Span}(\{p_0, \dots, p_d\})$.

Proposizione 11.2.11. *Se $[v_0], \dots, [v_d] \in \mathbb{P}(V)$ allora*

$$\mathrm{Span}([v_0], \dots, [v_d]) = \mathbb{P}(\mathrm{Span}(v_0, \dots, v_d)). \tag{11.2.4}$$

Dimostrazione. Il membro di destra di (11.2.4) contiene ciascun $[v_i]$, e se $\mathbb{P}(W) \subset \mathbb{P}(V)$ è un sottospazio proiettivo contenente $[v_0], \dots, [v_d]$ allora $v_i \in W$ per ogni i, quindi $\mathbb{P}(W)$ contiene il membro di sinistra di (11.2.4). □

Dalla Proposizione 11.2.11 segue immediatamente il seguente risultato.

Corollario 11.2.12. *Se $p_0, \dots, p_d \in \mathbb{P}(V)$ allora*

$$\dim \mathrm{Span}(p_0, \dots, p_d) \leqslant d. \tag{11.2.5}$$

Definizione 11.2.13. Punti $p_0, \dots, p_d$ di uno spazio proiettivo sono *linearmente indipendenti* se vale l'uguaglianza in (11.2.5), e sono *linearmente dipendenti* in caso contrario.

Osservazione 11.2.14. Sia $p_i = [v_i] \in \mathbb{P}(V)$ per $i \in \{0, \dots, d\}$. Per la Proposizione 11.2.11 i punti $p_0, \dots, p_d$ sono linearmente indipendenti se e solo se i vettori $v_0, \dots, v_d$ sono linearmente indipendenti.

Proposizione 11.2.15 (Grassmann proiettivo)**.** *Sia $\mathbb{P}$ uno spazio proiettivo e siano $L_1, L_2 \subset \mathbb{P}$ sottospazi proiettivi. Allora*

$$\dim(L_1 \cap L_2) = \dim L_1 + \dim L_2 - \dim \mathrm{Span}(L_1 \cup L_2).$$

Dimostrazione. Sia $\mathbb{P} = \mathbb{P}(V)$. Per $i = 1, 2$ abbiamo $L_i = \mathbb{P}(W_i)$ dove $W_i \subset V$ è un sottospazio vettoriale. Ragionando come nella dimostrazione della Proposizione vediamo che

$$\mathrm{Span}(L_1 \cup L_2) = \mathbb{P}(W_1 + W_2),$$

e d'altra parte

$$L_1 \cap L_2 = \mathbb{P}(W_1 \cap W_2).$$

Per la formula di Grassmann (per spazi vettoriali) abbiamo

$$\dim(W_1 \cap W_2) = \dim W_1 + \dim W_2 - \dim(W_1 + W_2).$$

Quindi

$$\dim(L_1 \cap L_2) = \dim(W_1 \cap W_2) - 1 = \dim W_1 - 1 + \dim W_2 - 1 - (\dim(W_1 + W_2) - 1) = \dim L_1 + \dim L_2 - \dim \mathrm{Span}(L_1 \cup L_2).$$

□

Esempio 11.2.16. Sia $\mathbb{P}$ uno spazio proiettivo. Siano $R \subset \mathbb{P}$ una retta e $H \subset \mathbb{P}$ un iperpiano. Se R non è contenuto in H allora interseca H in uno e un solo punto. Infatti siccome R non è contenuto in H il sottospazio $\mathrm{Span}(R \cup H)$ contiene strettamente H e quindi

$$\dim H < \dim \mathrm{Span}(R \cup H) \leqslant \dim \mathbb{P} = \dim H + 1.$$

Segue che $\dim \mathrm{Span}(R \cup H) = \dim \mathbb{P}$ e quindi $\mathrm{Span}(R \cup H) = \mathbb{P}$. Per la formula di Grassmann proiettiva $\dim(R \cap H) = 0$ cioè $R \cap H$ consiste di un singolo punto (uno spazio proiettivo di dimensione 0 è il proiettificato di uno spazio vettoriale di dimensione 1 e perciò consiste di un singolo punto).

11.3 Applicazioni tra spazi proiettivi

Applicazioni lineari

Supponiamo che $f\colon V \to W$ sia un'applicazione lineare iniettiva di spazi vettoriali. Allora è ben definita un'applicazione

$$\begin{array}{ccc} \mathbb{P}(V) & \xrightarrow{\mathbf{f}} & \mathbb{P}(W) \\ [v] & \mapsto & [f(v)] \end{array} \tag{11.3.1}$$

Infatti siccome f è iniettiva $f(v) \neq 0$ e quindi $f(v)$ genera un sottospazio vettoriale di W di dimensione 1. Inoltre per $\mu \in \mathbb{K}^*$ si ha $f(\mu v) = \mu f(v)$ che genera lo stesso sottospazio di $f(v)$.

Definizione 11.3.1. Un'applicazione $\mathbb{P}(V) \to \mathbb{P}(W)$ è *lineare* se è uguale a $\mathbf{f}$ per un'applicazione lineare iniettiva $f\colon V \to W$ di spazi vettoriali, ed è un *isomorfismo* se in aggiunta f è invertibile.

Osservazione 11.3.2. Se $f\colon V \to W$ è invertibile con inversa g (che, come sappiamo è lineare), allora

$$\mathbf{g}\colon \mathbb{P}(W) \to \mathbb{P}(V)$$

è inversa (destra e sinistra) di $\mathbf{f}$.

Osservazione 11.3.3. Se $\mathbf{f}\colon \mathbb{P}(V) \to \mathbb{P}(W)$ e $\mathbf{g}\colon \mathbb{P}(W) \to \mathbb{P}(U)$ sono applicazioni lineari di spazi proiettivi, la composizione

$$\mathbf{g} \circ \mathbf{f}\colon \mathbb{P}(V) \to \mathbb{P}(U)$$

è un'applicazione lineari di spazi proiettivi. In particolare la composizione di due isomorfismo di spazi proiettivi è un isomorfismo di spazi proiettivi.

Esempio 11.3.4. Siano $\mathbb{P}$ uno spazio proiettivo, $p_0 \in \mathbb{P}$ un suo punto e $H_1, H_2 \subset \mathbb{P}$ iperpiani non contenenti p_0. La *proiezione* di centro p_0

$$\pi\colon H_1 \to H_2$$

è definita come segue. Dato $q \in H_1$, siccome $p_0 \neq q$ (perchè $p_0 \notin H_1$) il sottospazio $\operatorname{Span}(p_0, q)$ generato da p_0 e q è una retta. Questa retta non è contenuta in H_2 perchè $p_0 \notin H_2$, e quindi l'intersezione $\operatorname{Span}(p_0, q) \cap H_2$ consiste di un singolo punto che è $\pi(q)$. La proiezione così definita è un isomorfismo di spazi proiettivi perchè è la composizione di due isomorfismi di spazi proiettivi. Per vederlo poniamo

$$\mathbb{P} = \mathbb{P}(V), \quad p_0 = [v_0], \quad H_i = \mathbb{P}(W_i),$$

dove $\dim W_i = \dim V - 1$ e $v_0 \notin W_i$. Abbiamo un'identificazione

$$\mathbb{P}(V/\operatorname{Span}(v_0)) = \{\mathbb{P}(U) \mid U \subset V,\ \dim U = 2,\ \ U \ni v_0\} = \{\text{rette in } \mathbb{P}(V) \text{ contenenti } [v_0]\}. \tag{11.3.2}$$

L'applicazione quoziente $V \longrightarrow V/\operatorname{Span}(v_0)$ ristretta a W_i è un isomorfismo, e quindi definisce un isomorfismo

$$\varphi_i\colon \mathbb{P}(W_i) \xrightarrow{\sim} \mathbb{P}(V/\operatorname{Span}(v_0)).$$

La proiezione $\pi\colon H_1 \to H_2$ di centro p_0 è la composizione

$$\mathbb{P}(W_1) \xrightarrow{\varphi_1} \mathbb{P}(V/\operatorname{Span}(v_0)) \xrightarrow{\varphi_2^{-1}} \mathbb{P}(W_2).$$

Infatti, tenendo conto di (11.3.2), l'applicazione φ_1 associa a $q \in H_1$ la retta $\operatorname{Span}(p_0, q)$, e l'applicazione φ_2^{-1} associa a una retta per p_0 la sua intersezione con H_2. Questo dimostra che π è un isomorfismo di spazi proiettivi.

Esempio 11.3.5. Sia $\mathcal{B} = \{v_0, \dots, v_n\}$ una base dello spazio vettoriale V. L'applicazione

$$\begin{array}{ccc} \mathbb{P}^n(\mathbb{K}) & \xrightarrow{\mathbf{f}} & \mathbb{P}(V) \\ [x_0, \dots, x_n] & \mapsto & [\sum_{i=0}^{n} x_i v_i] \end{array}$$

è un isomorfismo di spazi proiettivi. Notate che se $p \in \mathbb{P}(V)$ allora le sue coordinate omogenee (relative alla base $\mathcal{B}$) sono uguali a $\mathbf{f}^{-1}(p)$.

Esempio 11.3.6. Nel piano proiettivo $\mathbb{P}^2(\mathbb{R})$ sia $p_0 = [1,1,1]$ e siano H_1, H_2 le rette

$$H_1 : \ 3x_0 - x_1 + x_2 = 0, \quad H_2 : \ x_0 + x_1 + x_2 = 0.$$

Notiamo che p_0 non appartiene nè a H_1 nè ad H_2. Scriviamo esplicitamente la proiezione

$$\pi \colon H_1 \to H_2$$

di centro p_0. Sia $q = [t_0, t_1, t_2] \in H_1$. Allora

$$\pi(q) = [x_0, x_1, x_2]$$

è l'unico punto di H_2 allineato con p_0 e q. Quindi i vettori

$$(x_0, x_1, x_2), \ (t_0, t_1, t_2), \ (1,1,1)$$

sono linearmente dipendenti, ovvero

$$\det \begin{pmatrix} x_0 & x_1 & x_2 \\ t_0 & t_1 & t_2 \\ 1 & 1 & 1 \end{pmatrix} = 0$$

Calcolando troviamo che

$$(t_1 - t_2)x_0 + (t_2 - t_0)x_1 + (t_0 - t_1)x_2 = 0.$$

Siccome $[x_0, x_1, x_2] \in H_2$ abbiamo anche

$$x_0 + x_1 + x_2 = 0.$$

Risolvendo il sistema di due equazioni lineari omogenee (una a coefficienti polinomi in $_0, t_1, t_2$) troviamo che

$$\begin{aligned} x_0 &= 2t_0 - t_1 - t_2 \\ x_1 &= -t_0 + 2t_1 - t_2 \\ x_2 &= -t_0 - t_1 + 2t_2 \end{aligned}$$

Automorfismi

Definizione 11.3.7. Sia $\mathbb{P}$ uno spazio proiettivo. Un *automorfismo di* $\mathbb{P}$ (anche detta *proiettività*) è un isomorfismo proiettivo $\mathbf{f}\colon \mathbb{P} \to \mathbb{P}$. L'insieme degli automorfismi di $\mathbb{P}$ si denota $\mathrm{Aut}(\mathbb{P}^n)$.

Abbiamo che

1. $\mathrm{Id}_{\mathbb{P}}$ è un automorfismo di $\mathbb{P}$,
2. se $\mathbf{f}$ e $\mathbf{g}$ sono automorfismi di $\mathbb{P}$, allora anche $\mathbf{f} \circ \mathbf{g}$ è un automorfismo di $\mathbb{P}$, e
3. se $\mathbf{f}$ è un automorfismo di $\mathbb{P}$, allora anche l'inversa di $\mathbf{f}$ è un automorfismo di $\mathbb{P}$.

Quindi $\mathrm{Aut}(\mathbb{P})$ con operazione la composizione e unità $\mathrm{Id}_{\mathbb{P}}$ è un gruppo.

Sia $\mathbb{P} = \mathbb{P}(V)$. Per definizione un automorfismo di $\mathbb{P}(V)$ è dato da

$$\begin{array}{ccc} \mathbb{P}(V) & \xrightarrow{\mathbf{f}} & \mathbb{P}(V) \\ [v] & \mapsto & [f(v)] \end{array}$$

dove $f \in \mathrm{GL}(V)$. Quindi abbiamo un'applicazione suriettiva

$$\begin{array}{ccc} \mathrm{GL}(V) & \xrightarrow{T(f)} & \mathrm{Aut}(\mathbb{P}(V)) \\ f & \mapsto & \mathbf{f} \end{array}$$

L'applicazione T è un omomorfismo di gruppi, cioè

$$T(f \circ g) = \mathbf{f} \circ \mathbf{g}. \tag{11.3.3}$$

Proposizione 11.3.8. *Siano $f, g \in \mathrm{GL}(V)$. Allora gli automorfismi di $\mathbb{P}(V)$ dati da*

$$\boldsymbol{f}, \boldsymbol{g}\colon \mathbb{P}(V) \to \mathbb{P}(V)$$

sono uguali se e solo se $f \circ g^{-1}$ è un'applicazione <u>scalare</u>, cioè esiste $\lambda \in \mathbb{K}^*$ tale che

$$f(v) = \lambda g(v) \quad \forall v \in V. \tag{11.3.4}$$

Dimostrazione. Se vale (11.3.4) è chiaro che $\mathbf{f} = \mathbf{g}$. Dimostriamo il viceversa, cioè che se $\mathbf{f} = \mathbf{g}$ allora vale (11.3.4). Iniziamo dimostrando che se $h \in \mathrm{GL}(V)$ e $\mathbf{h} = \mathrm{Id}_{\mathbb{P}(V)}$ allora $h = \lambda \mathrm{Id}_V$ dove $\lambda \in \mathbb{K}^*$. Siccome $h([v]) = [v]$ per ogni $[v] \in \mathbb{P}(V)$, ogni vettore non nullo di V è autovettore di h. Questo implica che h è la moltiplicazione per uno scalare (non nullo). Infatti sia $\{v_1, \ldots, v_n\}$ una base di V. Abbiamo visto che per ogni $1 \in \{1, \ldots, n\}$ esiste $\lambda_i \in \mathbb{K}^*$ tale che

$$h(v_i) = \lambda_i v_i.$$

Se fosse $\lambda_i \neq \lambda_j$ allora il vettore $v_i + v_j$ non sarebbe autovettore di h perchè avremmo

$$h(v_i + v_j) = \lambda_i v_i + \lambda_j v_j$$

che non è multiplo di $v_i + v_j$ essendo $\lambda_i \neq \lambda_j$ e v_i, v_j linearmente indipendenti. Quindi

$$\lambda_1 = \ldots = \lambda_n, \tag{11.3.5}$$

cioè $h = \lambda \mathrm{Id}_V$.

Ora supponiamo che $\mathbf{f} = \mathbf{g}$. Quindi $\mathbf{f} \circ \mathbf{g}^{-1} = \mathrm{Id}_{\mathbb{P}(V)}$. Per (11.3.3) si ha

$$\mathbf{f} \circ \mathbf{g}^{-1} = T(f \circ g^{-1}).$$

Siccome $T(f \circ g^{-1}) = \mathrm{Id}_{\mathbb{P}(V)}$ esiste $\lambda \in \mathbb{K}^*$ tale che $f \circ g^{-1} = \lambda \mathrm{Id}_V$ (lo abbiamo appena dimostrato). □

Definizione 11.3.9. Sia $\mathbb{P}$ uno spazio proiettivo di dimensione n. I punti $p_1, \ldots, p_d \in \mathbb{P}$ sono *in posizione generale* se, dati comunque

$$1 \leqslant m \leqslant \min\{d, n+1\}$$

e m indici distinti $1 \leqslant i_1 < \ldots < i_m \leqslant d$, i punti $p_{i_1}, \ldots, p_{i_m}$ sono linearmente indipendenti.

Esempio 11.3.10. Siano $\alpha_1, \ldots, \alpha_d \in \mathbb{K}$ distinti. I punti di $\mathbb{P}^n(\mathbb{K})$ definiti da

$$[1, \alpha_1, \ldots, \alpha_1^n], \ldots, [1, \alpha_d, \ldots, \alpha_d^n]$$

sono in posizione generale.

Proposizione 11.3.11. *Siano $\mathbb{P}$ e $\mathbb{P}'$ spazi proiettivi della stessa dimensione n. Siano*

$$p_1, \ldots, p_{n+2} \in \mathbb{P}, \quad p_1', \ldots, p_{n+2}' \in \mathbb{P}'$$

punti in posizione generale in $\mathbb{P}$ e $\mathbb{P}'$ rispettivamente. Esiste uno e un solo isomorfismo di spazi proiettivi $\boldsymbol{f}\colon \mathbb{P} \to \mathbb{P}'$ tale che $\boldsymbol{f}(p_i) = p_i'$ per $i \in \{1, \ldots, n+2\}$.

Dimostrazione. Per definizione $\mathbb{P} = \mathbb{P}(V)$ e $\mathbb{P}' = \mathbb{P}(V')$ dove V, V' sono spazi vettoriali di dimensione $n+1$. Siano $p_i = [v_i]$, $p_i' = [v_i']$. Siccome $[v_1], \ldots, [v_{n+2}]$ sono in posizione generale, i vettori $v_1, \ldots, v_{n+1}$ sono linearmente indipendenti, e quindi formano una base di V. Quindi v_{n+2} è combinazione lineare di $v_1, \ldots, v_{n+1}$:

$$v_{n+2} = \mu_1 v_1 + \ldots + \mu_{n+1} v_{n+1}.$$

Nessuno dei $\mu_1, \ldots, \mu_{n+1}$ può essere nullo. Infatti se fosse $\mu_i = 0$ allora i punti $[v_j]$ con $j \neq i$ sarebbero linearmente dipendenti, e ciò contraddice l'ipotesi che i punti p_i sono in posizione generale. Analogamente i vettori $v_1', \ldots, v_{n+1}'$ sono linearmente indipendenti, e quindi formano una base di V', ed esistono $\mu_1', \ldots, \mu_{n+1}' \in \mathbb{K}$ tutti non nulli tali che

$$v_{n+2}' = \mu_1' v_1' + \ldots + \mu_{n+1}' v_{n+1}'.$$

Esiste una e una sola applicazione lineare

$$f\colon V \to V'$$

tale che

$$f(v_i) = \frac{\mu_i'}{\mu_i} v_i', \quad \forall i \in \{1, \ldots, n+1\}$$

perchè $\{v_1, \ldots, v_{n+1}\}$ è una base di V. Siccome $\{v_1', \ldots, v_{n+1}'\}$ è una base di V', f è un isomorfismo di spazi vettoriali, e perciò definisce un isomorfismo

$$\mathbf{f}\colon \mathbb{P}(V) \xrightarrow{\sim} \mathbb{P}(V').$$

Per costruzione

$$\mathbf{f}(p_i) = p_i', \quad i \in \{1, \ldots, n+2\}.$$

È anche chiaro che $\mathbf{f}$ è l'unica applicazione lineare con questa proprietà. □

Esempio 11.3.12. Per definizione un automorfismo di $\mathbb{P}^1(\mathbb{K})$ è dato da

$$\begin{array}{ccc} \mathbb{P}^1(\mathbb{K}) & \xrightarrow{\mathbf{f}} & \mathbb{P}^1(\mathbb{K}) \\ [X] & \mapsto & [M \cdot X] \end{array}$$

dove

$$M = \begin{pmatrix} m_{00} & m_{01} \\ m_{10} & m_{11} \end{pmatrix}$$

è una matrice 2×2 invertibile. Quindi

$$\mathbf{f}([X_0, X_1]) = [m_{00}X_0 + m_{01}X_1, m_{10}X_0 + m_{11}X_1].$$

Riscriviamo $\mathbf{f}$ usando la coordinata affine su $\mathbb{P}^1(\mathbb{K})_{X_0} = \mathbb{P}^1(\mathbb{K})\backslash\{[0,1]\}$ data da $x = X_1/X_0$:

$$\mathbf{f}(x) = ([1, x]) = [m_{00} + m_{01}x, m_{10} + m_{11}x] = \frac{m_{10} + m_{11}x}{m_{00} + m_{01}x}.$$

Liberandoci degli indici scriviamo

$$\mathbf{f}(x) = \frac{ax + b}{cx + d}, \qquad \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \neq 0. \tag{11.3.6}$$

Una tale $\mathbf{f}$ viene detta trasformazione *lineare fratta* o *di Möbius*. Riguardo la formula in (11.3.6): si usa porre

$$\infty = [0, 1]$$

perchè la coordinata affine di tale punto "è" uguale a ∞ (attenzione: questo ha senso in $\mathbb{P}^1(\mathbb{K})$ ma non in $\mathbb{P}^n(\mathbb{K})$ con $n > 1$). Con questa notazione

$$\mathbb{P}^1(\mathbb{K}) = \mathbb{K} \sqcup \{\infty\}.$$

Guardando (11.3.6) vediamo che

$$\mathbf{f}(\infty) = \frac{a}{c}, \quad \mathbf{f}(-d/c) = \infty.$$

11.4 Il birapporto

Siano $a_1, a_2, a_3 \in \mathbb{P}^1(\mathbb{K})$, cioè $a_1, a_2, a_3 \in \mathbb{K} \sqcup \{\infty\}$, e supponiamo che siano distinti. Allora a_1, a_2, a_3 sono in posizione generale (punti in $\mathbb{P}^1(\mathbb{K})$ sono in posizione generale se e solo se sono distinti). Per la Proposizione 11.3.11 esiste (ed è unico) $\mathbf{f} \in \mathrm{Aut}(\mathbb{P}^1(\mathbb{K}))$ tale che

$$\mathbf{f}(a_1) = 0, \quad \mathbf{f}(a_2) = \infty, \quad \mathbf{f}(a_3) = 1.$$

Esplicitamente, abbiamo

$$\mathbf{f}(x) = \frac{(a_2 - a_3)(a_1 - x)}{(a_1 - a_3)(a_2 - x)}. \tag{11.4.1}$$

Sostituendo a x un altro $a_4 \in \mathbb{K} \sqcup \{\infty\}$ otteniamo quello che si chiama il *birapporto di* $a_1, \ldots, a_4$:

$$\beta(a_1, \ldots, a_4) = \mathbf{f}(a_4) = \ = \frac{(a_2 - a_3)(a_1 - a_4)}{(a_1 - a_3)(a_2 - a_4)}. \tag{11.4.2}$$

Per esempio abbiamo

$$\beta(-2, -1, 1, 2) = \frac{8}{9}, \qquad \beta(-2, -1, 1, \infty) = \frac{(-1 - 1)(-2 - \infty)}{(-2 - 1)(-1 - \infty)} = \frac{2}{3}.$$

Notate che, siccome supponiamo che i punti siano distinti, la caratteristica del campo $\mathbb{K}$ è diversa da 3, e quindi entrambi i birapporti sono elementi di $\mathbb{K}$. Nell'ultima equazione abbiamo applicato la regola

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

Tale regola viene giustificata se scriviamo il il birapporto in coordinate proiettive. Poniamo

$$a_i = \frac{A_i}{T_i}, \tag{11.4.3}$$

cioè le coordinate proiettive del punto a_i sono $[T_i, A_i]$. Sostituendo (11.4.3) in (11.4.2) e moltiplicando numeratore e denominatore per $T_1 \cdot T_2 \cdot T_3 \cdot T_4$ si trova che

$$\beta([T_1, A_1], \dots, [T_4, A_4]) = \frac{\begin{vmatrix} T_2 & A_2 \\ T_3 & A_3 \end{vmatrix} \cdot \begin{vmatrix} T_1 & A_1 \\ T_4 & A_4 \end{vmatrix}}{\begin{vmatrix} T_1 & A_1 \\ T_3 & A_3 \end{vmatrix} \cdot \begin{vmatrix} T_2 & A_2 \\ T_4 & A_4 \end{vmatrix}}. \tag{11.4.4}$$

La formula in (11.4.2) mostra che il birapporto è un elemento di $\mathbb{K}$.

Proposizione 11.4.1. *Siano $a_1, \dots, a_4, b_1, \dots, b_4 \in \mathbb{P}^1(\mathbb{K})$ quaterne di punti distinti (cioè $a_i \neq a_j$ se $i \neq j$ e $b_i \neq b_j$ se $i \neq j$). Esiste $\boldsymbol{h} \in \mathrm{Aut}(\mathbb{P}^1(\mathbb{K}))$ tale che*

$$\boldsymbol{h}(a_i) = b_i, \quad i \in \{1, \dots, 4\} \tag{11.4.5}$$

se e solo se

$$\beta(a_1, \dots, a_4) = \beta(b_1, \dots, b_4). \tag{11.4.6}$$

Dimostrazione. Supponiamo che valga (11.4.5). Esiste $A \in M_{2,2}(\mathbb{K})$ invertibile tale che

$$\begin{array}{ccc} \mathbb{P}^1(\mathbb{K}) & \xrightarrow{\mathbf{h}} & \mathbb{P}^1(\mathbb{K}) \\ [X] & \mapsto & [A \cdot X] \end{array}$$

Guardando all'espressione (11.4.4) del birapporto vediamo che vale (11.4.6).

Ora supponiamo che valga (11.4.6) e dimostriamo che esiste un automorfismo proiettivo $\mathbf{h}$ di $\mathbb{P}^1(\mathbb{K})$ tale che valga (11.4.5). Sappiamo che esiste $\mathbf{f} \in \mathrm{Aut}(\mathbb{P}^1(\mathbb{K}))$ (ed è unica) tale che

$$\mathbf{f}(a_1) = 0, \quad \mathbf{f}(a_2) = \infty, \quad \mathbf{f}(a_3) = 1.$$

La $\mathbf{f}$ è data da (11.4.1), e

$$\mathbf{f}(a_4) = \beta(a_1, \dots, a_4).$$

Analogamente esiste $\mathbf{g} \in \mathrm{Aut}(\mathbb{P}^1(\mathbb{K}))$ tale che

$$\mathbf{g}(b_1) = 0, \quad \mathbf{g}(b_2) = \infty, \quad \mathbf{g}(b_3) = 1,$$

e

$$\mathbf{g}(b_4) = \beta(b_1, \dots, b_4).$$

Quindi

$$\mathbf{f}(a_4) = \beta(a_1, \dots, a_4) = \beta(b_1, \dots, b_4) = \mathbf{g}(b_4).$$

e vale (11.4.5) con $\mathbf{h} := \mathbf{g}^{-1} \circ \mathbf{f}$. □

Dalla Proposizione 11.4.1 segue immediatamente il seguente risultato.

Corollario 11.4.2. *Sia $\mathbb{P}$ una retta proiettiva, e siano $p_1, p_2, p_3, p_4 \in \mathbb{P}$ punti distinti. Se*

$$\boldsymbol{f}, \boldsymbol{g} \colon \mathbb{P} \xrightarrow{\sim} \mathbb{P}^1(\mathbb{K})$$

sono isomorfismi proiettivi, allora

$$\beta(\boldsymbol{f}(p_1), \boldsymbol{f}(p_2), \boldsymbol{f}(p_3), \boldsymbol{f}(p_4)) = \beta(\boldsymbol{g}(p_1), \boldsymbol{g}(p_2), \boldsymbol{g}(p_3), \boldsymbol{g}(p_4)).$$

Per la Proposizione 11.4.1 ha senso porre la seguente definizione.

Definizione 11.4.3. Sia $\mathbb{P}$ una retta proiettiva, e siano $p_1, p_2, p_3, p_4 \in \mathbb{P}$ punti distinti. Il *birapporto* di $p_1, \dots, p_4$ è dato da

$$\beta(\mathbf{f}(p_1), \mathbf{f}(p_2), \mathbf{f}(p_3), \mathbf{f}(p_4)),$$

dove $\mathbf{f} \colon \mathbb{P} \xrightarrow{\sim} \mathbb{P}^1(\mathbb{K})$ è un qualsiasi isomorfismo proiettivo.

Dalla Proposizione 11.4.1 segue anche il seguente risultato.

Proposizione 11.4.4. *Siano $\mathbb{P}, \mathbb{P}'$ rette proiettive e siano*

$$p_1, \dots, p_4 \in \mathbb{P}, \quad p'_1, \dots, p'_4 \in \mathbb{P}'$$

quaterne di punti distinti. Esiste un isomorfismo proiettivo $\boldsymbol{f} \colon \mathbb{P} \to \mathbb{P}'$ tale che

$$\boldsymbol{f}(p_i) = p'_i, \quad i \in \{1, \dots, 4\}$$

se e solo se

$$\beta(p_1, \dots, p_4) = \beta(p'_1, \dots, p'_4).$$

11.5 Dualità

La nozione di dualità per spazi proiettivi ha preceduto quella per spazi vettoriali, e ci dà una visione più geometrica della dualità per spazi vettoriali. Tutto nasce dall'osservazione che le rette di un piano proiettivo si comportano come i punti di un piano proiettivo, se decidiamo di considerare i fasci di rette come le rette di un piano. Più precisamente, introduciamo la terminologia classica.

Definizione 11.5.1. Sia $\mathbb{P}$ uno spazio proiettivo. Un punto $p \in \mathbb{P}$ e un iperpiano $L \subset \mathbb{P}$ sono *incidenti* se $p \in L$.

Ora supponiamo che $\mathbb{P}$ sia un piano proiettivo. Sappiamo che se $p_1, p_2 \in \mathbb{P}$ sono punti distinti allora esiste una e una sola retta incidente sia a p_1 che a p_2. L'osservazione è che, se scambiamo "punti" con "rette" nell'affermazione appena fatta, otteniamo l'affermazione "se R_1, R_2 sono rette distinte di $\mathbb{P}$ esiste uno e un solo punto incidente sia a R_1 che a R_2", che è anch'essa vera.

Seguendo l'approccio moderno, formuleremo la dualità per spazi proiettivi partendo dalla dualità per spazi vettoriali. Sia V uno spazio vettoriale finitamente generato.

Proposizione 11.5.2. *L'applicazione*

$$\begin{array}{ccc} \mathbb{P}(V^\vee) & \xrightarrow{\Delta} & \{H \subset \mathbb{P}(V) \mid H \text{ iperpiano}\} \\ [f] & \mapsto & \mathbb{P}(\ker f) \end{array} \tag{11.5.1}$$

è biunivoca.

Dimostrazione. Dimostriamo che Δ è suriettiva. Se $H \subset \mathbb{P}(V)$ è un iperpiano, per definizione $H = \mathbb{P}(W)$ dove $W \subset V$ è un sottospazio vettoriale di codimensione 1, cioè $\dim(V/W) = 1$. Quindi esiste un isomorfismo $V/W \xrightarrow{\sim} \mathbb{K}$. Componendo questo isomorfismo con l'applicazione quoziente $V \longrightarrow V/W$ otteniamo un'applicazione lineare suriettiva

$$f \colon V \twoheadrightarrow \mathbb{K}$$

tale che $\ker f = W$. Quindi $f \neq 0$ e perciò f determina un punto $[f] \in \mathbb{P}(V^\vee)$, e per costruzione $\Delta([f]) = \mathbb{P}(W)$.

Ora dimostriamo che Δ è iniettiva. Siano $[f], [g] \in \mathbb{P}(V^\vee)$ tali che $\Delta([f]) = \Delta([g])$, cioè

$$\ker f = \ker g. \tag{11.5.2}$$

Sia Ψ l'applicazione lineare

$$\begin{array}{ccc} V & \xrightarrow{\Psi} & \mathbb{K}^2 \\ v & \mapsto & (f(v), g(v)) \end{array}$$

Siccome vale (11.5.2) il nucleo di Ψ ha dimensione

$$\dim(\ker \Psi) = \dim(\ker f) = \dim V - 1.$$

Quindi $\dim(\operatorname{im} \Psi) = 1$, perciò esiste $(0,0) \neq (a,b) \in \mathbb{K}^2$ tale che $\operatorname{im} \Psi = \operatorname{Span}((a,b))$, e perciò

$$bf(v) - ag(v) = 0 \quad \forall v \in V.$$

Quindi f e g sono linearmente indipendenti, cioè $[f] = [g]$. □

Soffermiamoci sul significato di (11.5.1): l'insieme degli iperpiani di $\mathbb{P}(V)$, cioè i sottospazi proiettivi di codimensione 1, è in corrispondenza biunivoca con lo spazio proiettivo $\mathbb{P}(V^\vee)$, e quindi è a sua volta *uno spazio proiettivo* .

Definizione 11.5.3. Il *duale* di uno spazio proiettivo $\mathbb{P}$ è l'insieme dei suoi iperpiani, e si denota $\mathbb{P}^\vee$.

L'applicazione biunivoca Δ in (11.5.1) dà una struttura di spazio proiettivo a $\mathbb{P}^\vee$: se $\mathbb{P} = \mathbb{P}(V)$ allora $\mathbb{P}^\vee = \mathbb{P}(V^\vee)$.

Esempio 11.5.4. Come si caratterizza geometricamente una retta di iperpiani? cioè quali sono i sottoinsiemi $R \subset \mathbb{P}^\vee$ tali che $R = \mathbb{P}(U)$ dove $U \subset V^\vee$ è un sottospazio vettoriale di dimensione 2? La risposta: $R \subset \mathbb{P}^\vee$ è una retta se e solo se esiste un sottospazio proiettivo $\Lambda \subset \mathbb{P}$ di *codimensione 2* tale che

$$R = \{H \in \mathbb{P}^\vee \mid H \supset \Lambda\}. \tag{11.5.3}$$

Infatti se $R = \mathbb{P}(U)$ dove $U \subset V^\vee$ è un sottospazio vettoriale di dimensione 2, allora vale (11.5.3) con $\Lambda = \mathbb{P}(\operatorname{Ann} U)$, e il viceversa è analogo.

Esempio 11.5.5. Analogamente un sottoinsieme $S \subset \mathbb{P}^\vee$ è un sottospazio proiettivo di dimensione d se e solo se esiste un sottospazio proiettivo $\Lambda \subset \mathbb{P}$ di codimensione (d+1) tale che

$$S = \{H \in \mathbb{P}^\vee \mid H \supset \Lambda\}. \tag{11.5.4}$$

Segue che se $X \subset \mathbb{P}^\vee$ è un sottoinsieme, allora il sottospazio proiettivo di $\mathbb{P}^\vee$ generato da X è

$$\{H \in \mathbb{P}^\vee \mid H \supset \Lambda\}, \tag{11.5.5}$$

dove $\Lambda := \bigcap_{H \in X} H$.

Esempio 11.5.6. Siano $H_1, H_2, H_3 \subset \mathbb{P}^2(\mathbb{R})$ le rette di equazioni cartesiane

$$H_1 : \ 3x_0 - x_1 + x_2 = 0,$$
$$H_2 : \ 2x_0 + 5x_1 - x_2 = 0,$$
$$H_3 : \ 5x_0 + x_2 = 0.$$

Le rette sono allineate?, cioè appartengono a una retta di $\mathbb{P}^2(\mathbb{R})^\vee$? Per quello che abbiamo detto appartengono a una retta di $\mathbb{P}^2(\mathbb{R})^\vee$ se e solo se

$$H_1 \cap H_2 \cap H_3 \neq \varnothing.$$

Siccome $H_1 \cap H_2 \cap H_3 = \varnothing$ le rette non sono allinate.

Definizione 11.5.7. Un *fascio di iperpiani* in uno spazio proiettivo $\mathbb{P}$ è una retta dello spazio proiettivo duale $\mathbb{P}^\vee$.

La versione proiettiva del famoso isomorfismo naturale

$$\begin{array}{ccc} V & \xrightarrow{\sim} & (V^\vee)^\vee \\ v & \mapsto & (f \mapsto f(v)) \end{array}$$

valido per uno spazio vettoriale V di dimensione finita è l'isomorfismo

$$\begin{array}{ccc} \mathbb{P} & \xrightarrow{\sim} & (\mathbb{P}^\vee)^\vee \\ p & \mapsto & \{H \in \mathbb{P}^\vee \mid H \ni p\} \end{array}$$

La dualità permette di associare a ogni risultato di Geometria proiettiva un risultato duale (che a volte coincide con il risultato di partenza, a volte no). Infatti se un Teorema, diciamo X, vale per ogni spazio proiettivo di una certa dimensione n, allora vale anche per ogni spazio proiettivo duale di dimensione n, ma quest'ultimo l'enunciato si traduce in un enunciato (valido) per gli spazi proiettivi. Per intendersi: dove, nell'enunciato originale compare la parola punto, sostituiamo la parola iperpiano, dove compare la parola retta, sostituiamo fascio di iperpiani etc. Daremo un esempio significativo di come si può produrre un teorema apparentemente diverso a partire da un dato teorema, nel nostro caso il Teorema di Desargues.

Teorema 11.5.8. *Sia $\mathbb{P}$ un piano proiettivo, e siano $p_1, p_2, p_3, p_1', p_2', p_3' \in \mathbb{P}$ tali che*

1. *p_1, p_2, p_3 sono distinti, p_1', p_2', p_3' sono distinti e $p_i \neq p_i'$ per $i \in \{1,2,3\}$,*
2. *per $i \neq j \in \{1,2,3\}$ le rette $\mathrm{Span}(p_i, p_j)$ e $\mathrm{Span}(p_i', p_j')$ sono distinte, e*
3. *le rette $\mathrm{Span}(p_i, p_i')$ per $i \in \{1,2,3\}$ sono incidenti uno stesso punto.*

Allora i punti

$$\mathrm{Span}(p_1,p_2) \cap \mathrm{Span}(p_1',p_2'), \qquad \mathrm{Span}(p_1,p_3) \cap \mathrm{Span}(p_1',p_3'), \qquad \mathrm{Span}(p_2,p_3) \cap \mathrm{Span}(p_2',p_3')$$

sono allineati.

Prima di dare la dimostrazione del Teorema di Desargues, osserviamo che (1) e (2) sono genericamente soddisfatte mentre (3) è l'ipotesi forte del teorema. Indichiamo come dovremmo pensare all'enunciato del Teorema di Desargues. I punti p_1, p_2, p_3 sono i vertici di un triangolo (eventualmente degenere) i cui lati sono

$$\mathrm{Span}(p_1,p_2), \ \mathrm{Span}(p_1,p_3), \ \mathrm{Span}(p_2,p_3).$$

(Notate che in Geometria proiettiva non ha senso parlare del segmento tra due punti perchè ci sono *due* segmenti "tra" due punti.) Analogamente p_1', p_2', p_3' sono i vertici di un triangolo i cui lati sono

$$\mathrm{Span}(p_1',p_2'), \ \mathrm{Span}(p_1',p_3'), \ \mathrm{Span}(p_2',p_3').$$

L'ipotesi è che le rette congiungenti vertici corrispondenti sono allineate nello spazio duale, cioè sono incidenti il punto O, e la conclusione è che le intersezioni tra lati corrispondenti sono punti allineati (appartengono alla retta g).

Dimostrazione del Teorema di Desargues. Abbiamo che $\mathbb{P} = \mathbb{P}(V)$ dove V è uno spazio vettoriale di dimensione 3 su $\mathbb{K}$. Poniamo

$$p_i = [v_i], \quad p_i' = [v_i'], \quad O = [w].$$

Siccome

$$[w] \in \mathrm{Span}(p_i, p_i')$$

esistono $a_i, a_i' \in \mathbb{K}$ tali che

$$a_i v_i + a_i' v_i' = w. \tag{11.5.6}$$

Sottraendo le equazioni in (11.5.6) per $1 \leqslant i < j \leqslant 3$ otteniamo che

$$a_i v_i - a_j v_j = -a_i' v_i' + a_j' v_j'.$$

Quindi

$$\mathrm{Span}(p_1, p_2) \cap \mathrm{Span}(p_1', p_2') = [a_1 v_1 - a_2 v_2],$$
$$\mathrm{Span}(p_1, p_3) \cap \mathrm{Span}(p_1', p_3') = [a_1 v_1 - a_3 v_3],$$
$$\mathrm{Span}(p_2, p_3) \cap \mathrm{Span}(p_2', p_3') = [a_2 v_2 - a_3 v_3].$$

Siccome

$$(a_1 v_1 - a_2 v_2) - (a_1 v_1 - a_3 v_3) + (a_2 v_2 - a_3 v_3) = 0,$$

vale la tesi. □

Ora ci chiediamo: cosa afferma il Teorema duale del Teorema di Desargues? Per rispondere partiamo dal Teorema di Desargues per il duale $\mathbb{P}^\vee$ di un piano proiettivo $\mathbb{P}$, cioè l'enunciato del Teorema 11.5.8 con $\mathbb{P}^\vee$ al posto di $\mathbb{P}$, e operiamo le seguenti traduzioni:

1. a $p \in \mathbb{P}^\vee$ corrisponde una retta $R \subset \mathbb{P}$,
2. a una retta $\mathrm{Span}(q_1, q_2) \subset \mathbb{P}^\vee$ corrisponde il fascio delle rette $R \subset \mathbb{P}$ incidenti il punto $p := R_1 \cap R_2$, dove $R_i \subset \mathbb{P}$ è la retta che corrisponde a q_i (possiamo identificarlo con p purchè ci ricordiamo che significa che p significa "il fascio dlle rette contenenti p").

Quello che otteniamo è il seguente risultato.

Teorema 11.5.9 (Teorema di Desargues duale)**.** *Sia $\mathbb{P}$ un piano proiettivo. Siano $p_1, p_2, p_3, p_1', p_2', p_3' \in \mathbb{P}$ tali che*

1. *p_1, p_2, p_3 sono distinti, p_1', p_2', p_3' sono distinti, e $p_i \neq p_i'$ per $i \in \{1, 2, 3\}$,*
2. *per $i \neq j \in \{1, 2, 3\}$ le rette $\mathrm{Span}(p_i, p_j)$ e $\mathrm{Span}(p_i', p_j')$ sono distinte, e*
3. *I punti*

$$\mathrm{Span}(p_1, p_2) \cap \mathrm{Span}(p_1', p_2'), \qquad \mathrm{Span}(p_1, p_3) \cap \mathrm{Span}(p_1', p_3'), \qquad \mathrm{Span}(p_2, p_3) \cap \mathrm{Span}(p_2', p_3')$$

sono allineati.

Allora le rette $\mathrm{Span}(p_i, p_i')$ per $i \in \{1, 2, 3\}$ sono incidenti uno stesso punto.

In altre parole il duale del Teorema di Desargues ci dice che vale il "viceversa" del Teorema di Desargues. Un modo di vedere quello che abbiamo fatto: il Teorema di Desargues enunciato all'inizio e quello esposto sopra non sono lo stesso Teorema, ma se abbiamo la nozione di spazio proiettivo duale ci rendiamo conto che sono Teoremi analoghi.

Quindi un punto di vista sulla dualità è il seguente: ci permette di riconoscere come sostanzialmente uguali fenomeni all'apparenza diversi.

11.6 (Iper)quadriche proiettive

In questa sezione supponiamo che il campo $\mathbb{K}$ abbia caratteristica diversa da 2.

(Iper)quadriche in spazi proiettivi

Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione finita, e sia $F\colon V \to \mathbb{K}$ una forma quadratica non nulla. Siano $0 \neq v \in V$ e $\lambda \in \mathbb{K}^*$: allora $F(\lambda v) = \lambda^2 F(v)$ e perciò $F(v) = 0$ se e solo se $F(\lambda v) = 0$. Quindi ha senso porre

$$V(F) := \{[v] \in \mathbb{P}(V) \mid F(v) = 0\}.$$

Diciamo che $V(F)$ è una iperquadrica proiettiva in $\mathbb{P}(V)$, e che F è una sua equazione cartesiana. Se $\mathbb{P}(V)$ è un piano proiettivo diciamo che $V(F)$ è una conica proiettiva, se $\mathbb{P}(V)$ è un solido proiettivo (cioè $\dim \mathbb{P}(V) = 3$) diciamo che $V(F)$ è una quadrica proiettiva. La parola iperquadrica (o conica o quadrica) andrebbe messa tra virgolette perchè $V(F)$ può essere vuoto, per esempio $V(X_0^2 + \ldots + X_n^2) \subset \mathbb{P}^n(\mathbb{R})$, o essere ridotta a un solo punto, per esempio $V(X_1^2 + \ldots + X_n^2) \subset \mathbb{P}^n(\mathbb{R})$, o più in generale può essere un sottospazio proiettivo, per esempio $V(X_m^2 + \ldots + X_n^2) \subset \mathbb{P}^n(\mathbb{R})$ è uguale a $\{[X] \in \mathbb{P}^n(\mathbb{R}) \mid X_m = \ldots = X_n = 0\}$.

In generale però $V(F)$ merita il nome di iperquadrica (proiettiva), come si vede dal seguente esempio.

Esempio 11.6.1. Sia $F \in Q(\mathbb{R}^3)$ data da

$$F(X_0, X_1, X_2) = X_0^2 - X_1^2 + X_2^2.$$

Analizziamo l'intersezione di $V(F)$ con lo spazio affine

$$\mathbb{P}^2(\mathbb{R})_{X_0} = \{[X_0, X_1, X_2] \mid X_0 \neq 0\}.$$

Ricordiamo che le coordinate affini standard su $\mathbb{P}^n(\mathbb{R})_{X_0}$ sono date da

$$x_1 = \frac{X_1}{X_0}, x_= \frac{X_n}{X_0}.$$

Quindi un'equazione cartesiana di $Q := V(F) \cap \mathbb{P}^n(\mathbb{R})_{X_0}$ è data da

$$x_1^2 - x_2^2 = 1,$$

cioè Q è un'iperbole. Inoltre i punti di $V(F)$ che non sono in Q, cioè i punti di $V(F) \cap V(X_0)$, sono $[0, 1, 1]$ e $[0, 1, -1]$. Se proiettiamo Q su un altro piano (ne facciamo il "ritratto") non parallelo al piano contenente Q, i due punti $[0, 1, 1]$ e $[0, 1, -1]$ vengono proiettati nei punti dell'intersezione dell'"orizzonte" con la proiezione di Q.

Più in generale sia $0 \neq F \in Q(\mathbb{R}^{n+1})$ e consideriamo $V(F) \subset \mathbb{P}^n(\mathbb{R})$. Analizziamo l'intersezione di $V(F)$ con lo spazio affine

$$\mathbb{P}^n(\mathbb{R})_{X_0} = \{[X_0, \ldots, X_n] \mid X_0 \neq 0\}.$$

Le coordinate affini standard su $\mathbb{P}^n(\mathbb{R})_{X_0}$ sono

$$x_1 = \frac{X_1}{X_0}, \ldots, x_n = \frac{X_n}{X_0},$$

e perciò un'equazione cartesiana di $Q := V(F) \cap \mathbb{P}^n(\mathbb{R})_{X_0}$ è $F(1, x_1, \ldots, x_n) = 0$. Ci sono tre possibilità:

1. $F(1, x_1, \ldots, x_n)$ è un polinomio di grado 2 e quindi Q è una ellisse, o un'iperbole etc.
2. $F(1, x_1, \ldots, x_n)$ è un polinomio di grado 1, cioè $F = X_0 \cdot L(X_0, \ldots, X_n)$ dove L è un polinomio omogeneo di grado 1, e quindi Q è un iperpiano.
3. $F(1, x_1, \ldots, x_n)$ è un polinomio di grado 0 (una costante c), cioè $F = cX_0^2$.

C'è un viceversa di (1), (2) e (3). Il caso più interessante è il viceversa di (1). Sia $f \in \mathbb{R}[x_1, \ldots, x_n]$ un polinomio di grado 2 non nullo, e sia $Q \subset \mathbb{E}^n(\mathbb{R})$ la quadrica

$$Q\colon\ f(x_1, \ldots, x_n) = 0.$$

Nel Capitolo 10 abbiamo associato a f la forma quadratica $F(X_0, \ldots, X_n)$ definita da

$$F(X_0, \ldots, X_n) = X_0^2 \cdot f\left(\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}\right). \tag{11.6.1}$$

È chiaro che

$$V(F) \cap \mathbb{P}^n(\mathbb{R})_{X_0} = Q. \tag{11.6.2}$$

Quindi la forma quadratica che abbiamo associato a un'equazione cartesiana di secondo grado di $Q \subset \mathbb{A}^n_{\mathbb{R}}$ definisce una ipersuperficie quadrica $\mathbf{Q} \subset \mathbb{P}^n(\mathbb{R})$ con la proprietà che

$$\mathbf{Q} \cap \mathbb{P}^n(\mathbb{R})_{X_0} = Q. \tag{11.6.3}$$

Punti lisci e punti singolari di iperquadriche

Sia $0 \neq F \in Q(V)$ e sia $\mathbf{Q} := V(F) \subset \mathbb{P}(V)$ l'iperquadrica definita da F. La polarizzazione $\widetilde{F} \in \mathrm{Bil}_+(V)$ di F aiuta a capire la geometria di $\mathbf{Q}$. Ricordiamo che

$$\widetilde{F}(v,v) = F(v) \quad \forall v \in V.$$

Se $v \in V$ denotiamo $v^\perp$ l'ortogonale di v per la forma bilineare $\widetilde{F}$ associata a F, cioè

$$v^\perp = \{w \in V \mid \widetilde{F}(v,w) = 0\}.$$

Notiamo che se $v \in V$, allora $v^\perp$ ha codimensione 1 oppure è tutto V se $v \in \ker F$. Inoltre $[v] \in \mathbf{Q}$ se e solo se $v \in v^\perp$. In particolare

$$\mathbb{P}(\ker F) \subset V(F). \tag{11.6.4}$$

Definizione 11.6.2. Sia $0 \neq F \in Q(V)$ e $\mathbf{Q} := V(F)$. Sia $[v] \in \mathbf{Q}$.

1. $\mathbf{Q}$ è *liscia* in $[v]$ (o $[v]$ è un *punto liscio* di $\mathbf{Q}$) se $v^\perp$ ha codimensione 1.
2. $\mathbf{Q}$ è *singolare* in $[v]$ (o $[v]$ è un *punto singolare* di $\mathbf{Q}$) se $v^\perp = V$.

Un'iperquadrica è liscia se tutti i suoi punti sono lisci (cioè F è non degenere), è singolare se ha almeno un punto singolare (cioè F è degenere).

Definizione 11.6.3. Sia $F \in Q(V)$ e $\mathbf{Q} := V(F)$. Sia $[v] \in \mathbf{Q}$ un punto liscio. Lo *spazio tangente* a $\mathbf{Q}$ in $[v]$ è $\mathbf{T}_{[v]}\mathbf{Q} := \mathbb{P}(v^\perp)$.

Vogliamo giustificare la denominazione "spazio tangente". Iniziamo notando che se $[v] \in \mathbf{Q}$ e $L \subset \mathbb{P}(V)$ è una retta contenente $[v]$, allora

1. $L \cap \mathbf{Q}$ consiste di due punti distinti ($[v]$ e $[w] \neq [v]$), o
2. $L \cap \mathbf{Q}$ consiste di un solo punto, cioè $[v]$, o
3. $L \cap \mathbf{Q} = L$, cioè $L \subset \mathbf{Q}$.

Infatti $L = \mathbb{P}(U)$ dove $U \subset V$ è un sottospazio vettoriale di dimensione 2, e la restrizione $F_{|U}$ è una forma quadratica. Perciò possiamo scrivere

$$F_{|U} = \alpha \cdot \ell_1 \cdot \ell_2, \qquad \ell_1, \ell_2 \in U^\vee, \quad \alpha \in \mathbb{K}, \quad \ell_1(v) = 0.$$

Se $\alpha \neq 0$ e $\ell_2(v) \neq 0$, siamo nel caso (a). Se $\alpha \neq 0$ e $\ell_2(v) = 0$, siamo nel caso (b). Se $\alpha = 0$ siamo nel caso (c).

Lemma 11.6.4. *Sia $F \in Q(V)$ e $\boldsymbol{Q} := V(F)$. Sia $[v] \in \boldsymbol{Q}$ un punto liscio. Sia $L \subset \mathbb{P}(V)$ una retta contente $[v]$. Allora L è contenuta nello spazio tangente a $\boldsymbol{Q}$ in $[v]$ se e solo se $L \cap \boldsymbol{Q} = \{[v]\}$ oppure $L \subset \boldsymbol{Q}$.*

Dimostrazione. $L = \mathbb{P}(U)$ dove $U \subset V$ è un sottospazio vettoriale di dimensione 2. Estendiamo $\{v\}$ a una base $\{v, w\}$ di U. Siccome $F(v) = 0$ abbiamo

$$F(xv + yw) = y \cdot (2\widetilde{F}(v,w)x + F(w)y).$$

Segue che se $\widetilde{F}(v,w) \neq 0$, cioè L non è contenuta nello spazio tangente a $\mathbf{Q}$ in $[v]$, allora

$$L \cap \mathbf{Q} = \{[v], [F(w)v - 2\widetilde{F}(v,w)w]\}$$

e quindi consiste di due punti. D'altra parte, se $\widetilde{F}(v,w) = 0$, cioè L è contenuta nello spazio tangente a $\mathbf{Q}$ in $[v]$, allora $L \cap \mathbf{Q} = \{[v]\}$ se $F(w) \neq 0$ e invece $L \subset \mathbf{Q}$ se $F(w) = 0$. □

D'altra parte, il seguente risultato descrive il comportamento di una iperquadrica che ha un punto singolare.

Lemma 11.6.5. *Sia $\mathbb{P}$ uno spazio proiettivo, e sia $\boldsymbol{Q} \subset \mathbb{P}$ un'iperquadrica che ha un punto singolare x_0. Se $x \in (\boldsymbol{Q} \backslash \{x_0\})$, allora la retta contenete x_0 e x è contenuta in $\boldsymbol{Q}$.*

Dimostrazione. Per definizione $\mathbf{Q} = V(F)$, dove $0 \neq F \in Q(V)$, e $x_0 = [v_0]$ dove $v_0 \in \ker F$. Poniamo $x = [v]$. I punti di $\mathrm{Span}(x_0, x)$ sono dati da $[\alpha_0 v_0 + \alpha v]$ con $([\alpha_0, \alpha) \in \mathbb{K}^2$ non nullo. Siccome $v_0 \in \ker F$, anche $\alpha_0 v_0 \in \ker F$, e quindi

$$F(\alpha_0 v_0 + \alpha v) = F(\alpha v) = \alpha^2 F(v) = 0.$$

□

Definizione 11.6.6. Siano $\mathbb{P}$ uno spazio proiettivo, e $x_0 \in \mathbb{P}$. Un sottoinsieme $X \subset \mathbb{P}$ è un *cono di vertice* x_0 se è l'unione di rette contenenti x_0.

Il Lemma 11.6.5 afferma che una quadrica singolare è un cono di vertice un qualsiasi $[v_0]$ con $v_0 \in \ker(F)$. Il seguente risultato segue subito dal Lemma 11.6.5, e mostra che un'iperquadrica singolare in uno spazio proiettivo di dimensione n si descrive a partire un'iperquadrica in uno spazio proiettivo di dimensione $(n-1)$.

Corollario 11.6.7. *Sia $\mathbb{P}$ uno spazio proiettivo, e sia $\boldsymbol{Q} \subset \mathbb{P}$ un'iperquadrica che ha un punto singolare x_0. Se $L \subset \mathbb{P}(V)$ è un iperpiano che non contienen x_0, allora l'intersezione $Q_L := L \cap \boldsymbol{Q}$ è un'iperquadrica in L e Q è l'unione delle rette generate da x_0 e punti di Q_L.*

Ora possiamo spiegare il significato geometrico della terminologia "quadrica non degenere/degenere" riferita a una iperquadrica Q in uno spazio affine euclideo $\mathbb{S}$. Scegliendo un isomorfismo tra $\mathbb{S}$ e $\mathbb{E}^n(\mathbb{R})$ (dove $n := \dim \mathbb{S}$), possiamo assumere che $\mathbb{S}$ sia $\mathbb{E}^n(\mathbb{R})$. Inoltre identifichiamo $\mathbb{E}^n(\mathbb{R})$ con $\mathbb{P}^n(\mathbb{R})_{X_0}$. Quindi

$$Q : f(x_1, \dots, x_n) = 0,$$

dove f è un polinomio non nullo di grado 2. Come discusso nella sottosezione precedente, l'omogeneizzazione F di f definisce un'iperquadrica proiettiva $\mathbf{Q} \subset \mathbb{P}^n(\mathbb{R})$ tale che $\mathbf{Q} \cap \mathbb{E}^n(\mathbb{R}) = Q$. Quindi Q è degenere se e solo se $\mathbf{Q}$ è un cono. In questo caso il vertice di $\mathbf{Q}$ è un qualsiasi punto singolare di $\mathbf{Q}$. Se esiste un punto singolare di $\mathbf{Q}$ che non è all'infinito, cioè appartiene a $\mathbb{E}^n(\mathbb{R})$, allora Q appare come un cono nell'accezione usuale del termine, eventualmente ridotto al solo vertice, o altri casi poco visibili se ci si limita alle coordinate reali (anzichè complesse). Se invece i punti singolari di $\mathbf{Q}$ sono tutti all'infinito, Q è un cilindro.

Iperquadriche proiettive a meno di proiettività

L'azione

$$\begin{array}{ccc} \mathrm{GL}(V) \times Q(V) & \longrightarrow & Q(V) \\ (g, F) & \mapsto & gF \end{array}$$

(ricordate che $gF(v) := F(g^{-1}(v))$) ha il seguente significato geometrico: se $0 \neq F \in Q(V)$ per cui $V(F) \subset \mathbb{P}(V)$ è un'iperquadrica, allora

$$\mathbf{g}(V(F)) = V(gF), \tag{11.6.5}$$

dove $\mathbf{g}$ è la proiettività di $\mathbb{P}(V)$ definita da $\mathbf{g}([v]) = [g(v)]$. Infatti $[v] \in \mathbf{g}(V(F))$ se e solo se esiste $[w] \in V(F)$ tale che $v = g(w)$, o equivalentemente $g^{-1}(v) \in V(F)$. Perciò $[v] \in \mathbf{g}(V(F))$ se e solo se $F(g^{-1}(v)) = 0$, cioè se $gF(v) = 0$.

Definizione 11.6.8. $[F_1], [F_2] \in \mathbb{P}(Q(V))$ sono *proiettivamente equivalenti* se esistono $g \in \mathrm{GL}(V)$ e $\lambda \in \mathbb{K}^*$ tali che $F_1 = \lambda g F_2$.

Osservazione 11.6.9. Se $[F_1], [F_2] \in \mathbb{P}(Q(V))$ sono proiettivamente equivalenti, diciamo $F_1 = \lambda g F_2$ con $g \in \mathrm{GL}(V)$ e $\lambda \in \mathbb{K}^*$, allora $V(F_1) = \mathbf{g}(V(F_2))$, cioè l'iperquadrica $V(F_1)$ è l'immagine di $V(F_2)$ per la proiettività $\mathbf{g}$.

Si verifica facilmente che la relazione appena definita è di equivalenza. Classificheremo tali classi di equivalenza per $\mathbb{K} = \mathbb{R}$ e $\mathbb{K} = \mathbb{C}$. Come sempre quando cerchiamo di capire una relazione di equivalenza su un insieme, vogliamo trovare invarianti, cioè funzioni sull'insieme che assumono lo stesso valore sulgi elementi di una stessa classe di equivalenza.

Sia $[F] \in \mathbb{P}(Q(V))$: il rango di F è un invariante. Infatti se $G = \lambda g F$, dove $g \in \mathrm{GL}(V)$ e $\lambda \in \mathbb{K}^*$, allora $\mathrm{rg}(G) = \mathrm{rg}(gF)$ perchè $\lambda \in \mathbb{K}^*$ e $\mathrm{rg}(gF) = \mathrm{rg}(F)$ per il Corollario 8.2.9 e l'Osservazione 8.3.20. In particolare ha senso parlare di rango di $[F] \in \mathbb{P}(Q(V))$.

Proposizione 11.6.10. *Sia V uno spazio vettoriale complesso, e siano $[F_1], [F_2] \in \mathbb{P}(Q(V))$. Allora $[F_1]$ e $[F_2]$ sono proiettivamente equivalenti se e solo se hanno lo stesso rango.*

Dimostrazione. Segue dalla Proposizione 8.4.1. □

Ora passiamo al caso reale. Sia V uno spazio vettoriale reale. Se $F \in Q(V)$, allora sono definite le segnature positiva e negative $s_+(F)$ e $s_-(F)$. Se $g \in \mathrm{GL}(V)$, allora

$$s_+(gF) = s_+(F), \quad s_-(gF) = s_-(F).$$

Inoltre se $\lambda \in \mathbb{R}^*$ allora

$$s_\pm(\lambda F) = \begin{cases} s_\pm(F) & \text{se } \lambda > 0, \\ s_\mp(F) & \text{se } \lambda < 0. \end{cases}$$

La conclusione è che se $[F] \in \mathbb{P}(Q(V))$ l'insieme $\{s_+(F), s_-(F)\}$ i cui elementi sono le segnature di F è un invariante di $[F]$ per la relazione di equivalenza proiettiva. (Attenzione: siccome $\{s_+(F), s_-(F)\}$ è un insieme non sappiamo quale sia la segnatura positiva/negativa.)

Proposizione 11.6.11. *Sia V uno spazio vettoriale reale, e siano $[F_1], [F_2] \in \mathbb{P}(Q(V))$. Allora $[F_1]$ e $[F_2]$ sono proiettivamente equivalenti se e solo se*

$$\{s_+(F_1), s_-(F_1)\} = \{s_+(F_2), s_-(F_2)\}. \tag{11.6.6}$$

Dimostrazione. Dobbiamo dimostrare che se vale (11.6.6) allora $[F_1]$ e $[F_2]$ sono proiettivamente equivalenti. Se $s_+(F_1) = s_+(F_2)$ e quindi $s_-(F_1) = s_-(F_2)$, segue dalla Proposizione 8.4.8. Se $s_+(F_1) = s_-(F_2)$ e quindi $s_-(F_1) = s_+(F_2)$, allora

$$s_+(F_1) = s_+(-F_2), \quad s_-(F_1) = s_-(-F_2).$$

Siccome $[-F_2] = [F_2]$ basta dimostrare che $[F_1]$ è proiettivamente equivalente a $[-F_2]$, e siamo di nuovo nel caso appena trattato. □

Corollario 11.6.12. *Sia V uno spazio vettoriale reale, e siano $[F_1], [F_2] \in \mathbb{P}(Q(V))$. Allora $[F_1]$ e $[F_2]$ sono proiettivamente equivalenti se e solo se le dimensioni dei luoghi singolari sono le stesse, e la dimensione massima di un sottospazio proiettivo contenuto in $V(F_1)$ è uguale alla dimensione massima di un sottospazio proiettivo contenuto in $V(F_2)$.*

Dimostrazione. Segue dalla Proposizione 11.6.11 insieme all'Esercizio 8.7. □

Esempio 11.6.13. Se $\mathbb{P}$ è un piano proiettivo reale, due coniche non degeneri non vuote sono proiettivamente equivalenti.

Esempio 11.6.14. Se $\mathbb{P}$ è uno spazio proiettivo reale di dimensione 3, le classi di equivalenza di quadriche non degeneri sono 3, e sono rappresentate, in opportune coordinate omogenee $[X_0, \dots, X_3]$, da

$$[X_0^2 + X_1^2 + X_2^2 + X_3^2], \quad [X_0^2 + X_1^2 + X_2^2 - X_3^2], \quad [X_0^2 + X_1^2 - X_2^2 - X_3^2].$$

Notate che

$$V(X_0^2 + X_1^2 + X_2^2 + X_3^2) = \varnothing.$$

Prima di esaminare gli altri due casi, notiamo che $\mathbb{P}\backslash V(X_i)$ è uno spazio affine di dimensione 3, e diamogli una struttura di spazio affine euclideo scegliendo la forma quadratica standard nelle coordinate

$$\frac{X_0}{X_i}, \dots, \frac{X_3}{X_i}.$$

Ora notiamo che

$$V(X_0^2 + X_1^2 + X_2^2 - X_3^2) \cap (\mathbb{P}\backslash V(X_3))$$

è una sfera, ma

$$V(X_0^2 + X_1^2 + X_2^2 - X_3^2) \cap (\mathbb{P}\backslash V(X_2))$$

è un paraboloide ellittico. D'altra parte

$$V(X_0^2 + X_1^2 - X_2^2 - X_3^2) \cap (\mathbb{P}\backslash V(X_3))$$

è un'iperboloide iperbolico.

Esercizi del Capitolo 11

Esercizio 11.1. *Sia $\mathbb{P}$ una retta proiettiva e siano $p_1, p_2, p_3, p_4 \in \mathbb{P}$ punti distinti. Verificate che*

$$\beta(p_2, p_1, p_4, p_3) = \beta(p_1, p_2, p_3, p_4)$$
$$\beta(p_3, p_4, p_1, p_2) = \beta(p_1, p_2, p_3, p_4)$$
$$\beta(p_4, p_3, p_2, p_1) = \beta(p_1, p_2, p_3, p_4)$$

Deducetene che esiste una proiettività $\boldsymbol{f}\colon \mathbb{P}^1(\mathbb{K}) \to \mathbb{P}^1(\mathbb{K})$ tale che

$$\boldsymbol{f}(a_1) = a_2,\ \boldsymbol{f}(a_2) = a_1,\ \boldsymbol{f}(a_3) = a_4,\ \boldsymbol{f}(a_4) = a_3,$$

e anche una proiettività $\boldsymbol{g}\colon \mathbb{P}^1(\mathbb{K}) \to \mathbb{P}^1(\mathbb{K})$ tale che

$$\boldsymbol{g}(a_1) = a_3,\ \boldsymbol{g}(a_3) = a_1,\ \boldsymbol{g}(a_2) = a_4,\ \boldsymbol{g}(a_4) = a_2,$$

e anche...

Esercizio 11.2. *Siano $a_1, a_2, a_3, a_4 \in (\mathbb{K} \sqcup \{\infty\})$ punti distinti. Sia $\mathcal{S}_4$ il gruppo delle permutazioni di $\{1,2,3,4\}$. Sia $\sigma \in \mathcal{S}_4$. Dimostrate che la funzione razionale di $a_1, \dots, a_4$ data da*

$$\beta(a_{\sigma(1)}, a_{\sigma(2)}, a_{\sigma(3)}, a_{\sigma(4)})$$

è uguale a $\beta(a_1, a_2, a_3, a_4)$ solo se

$$\sigma \in \{\mathrm{Id},\ (12)(34),\ (13)(24),\ (14)(23)\}.$$

(Confrontate con l'Esercizio 11.1.) Quindi dalle permutazioni di a_1, a_2, a_3, a_4 otteniamo 6 funzione razionale diverse, non $4! = 24$. Dimostrate che se

$$\beta := \beta(a_1, a_2, a_3, a_4),$$

allora le espressioni diverse che otteniamo sono

$$\beta,\ \frac{1}{\beta},\ 1-\beta,\ \frac{1}{1-\beta},\ \frac{\beta-1}{\beta},\ \frac{\beta}{\beta-1}. \tag{11.6.7}$$

Esercizio 11.3. *Sia $\mathbb{P}$ una retta proiettiva. La quaterna di punti distinti $p_1, p_2, p_3, p_4 \in \mathbb{P}$ è* armonica *se*

$$\beta(a_1, a_2, a_3, a_4) = -1.$$

Per esempio la quaterna $0, \infty, 1, -1$ è armonica. Determinate il sottogruppo delle permutazioni $\sigma \in \mathcal{S}_4$ tali che

$$\beta(p_{\sigma(1)}, p_{\sigma(2)}, p_{\sigma(3)}, p_{\sigma(4)}) = \beta(p_1, p_2, p_3, p_4).$$

Esercizio 11.4. *Sia $\mathbb{P}$ uno spazio proiettivo di dimensione 3. Siano $R_1, R_2 \subset \mathbb{P}$ rette che non hanno punti in comune, e sia*

$$p \in (\mathbb{P} \backslash R_1 \backslash R_2).$$

(a) Dimostrate che esiste una e una sola retta $T \subset \mathbb{P}$ tale che

$$p \in T, \quad R_1 \cap T \neq \varnothing \neq T \cap R_2.$$

(b) Dimostrate che la corrispondente affermazione per $\mathbb{P}^\vee$, tradotta in termini di $\mathbb{P}$, è uguale ad (a), cioè (a) è autoduale.

Esercizio 11.5. *Scopo di questo esercizio è di indicare una dimostrazione più geometrica del Teorema di Desargues. Siccome le proiezioni di punti allineati sono allineati, basta dimostrare l'analogo del Teorema di Desargues in uno spazio proiettivo di dimensione 3, cioè l'enunciato del Teorema di Desargues, ma assumendo che i punti $p_1, p_2, p_3, p_1', p_2', p_3'$ appartengano a uno spazio proiettivo $\mathbb{B}$ di dimensione 3. Questo sembra più difficile, ma in verità è più semplice. Infatti sia $T \subset \mathbb{B}$ il piano che contiene il triangolo p_1, p_2, p_3 e sia $T' \subset \mathbb{B}$ il piano che contiene il triangolo p_1', p_2', p_3'. Dimostrate che ciascuna delle intersezioni*

$$\mathrm{Span}(p_1, p_2) \cap \mathrm{Span}(p_1', p_2'), \qquad \mathrm{Span}(p_1, p_3) \cap \mathrm{Span}(p_1', p_3'), \qquad \mathrm{Span}(p_2, p_3) \cap \mathrm{Span}(p_2', p_3')$$

consiste di un punto (questo non è ovvio perchè sono intersezioni tra rette di uno spazio proiettivo di dimensione 3), e che questi punti appartengono alla retta di intersezione $g = T \cap T'$.

Bibliografia

[1] M. Artin. *Algebra*, Prentice Hall, 1991.

[2] K. Bryan, T. Leise. *The $ 25,000,000,000 eigenvector: the linear algebra behind Google*, SIAM Rev. 48 (2006), pp. 56–581. https://www.math.arizona.edu/ glickenstein/math443f08/bryanleise.pdf

[3] S. Lang. *Algebra*, Springer GTM 211, 2002.

[4] E. Sernesi. *Geometria I*, Bollati Boringhieri, 1989.

[5] D. Hilbert, S. Cohn-Vossen. *Geometria intuitiva*, Biblioteca di cultura scientifica, 63 Paolo Boringhieri, Torino 1960.