Algebra Lineare e Geometria

Kieran G. O'Grady

"Sapienza" Università di Roma

Aggiornamento 3/11/2016

Indice

0 Introduzione **5**

1 Preliminari **7**
1.1 Insiemi . . . 7
1.2 Funzioni . . . 9
1.3 Relazioni . . . 11
1.4 Induzione matematica . . . 13
1.5 Anelli e campi . . . 14
1.6 Polinomi . . . 16
1.7 L'algoritmo euclideo . . . 19
1.8 Numeri complessi . . . 19

2 Spazi vettoriali **25**
2.1 Gli archetipi e la definizione . . . 25
2.2 Prime proprietà . . . 27
2.3 Sottospazi . . . 28
2.4 Dipendenza/indipendenza lineare . . . 31
2.5 Basi . . . 33
2.6 Formula di Grassmann . . . 38
2.7 Costruzioni astratte di spazi vettoriali . . . 39

3 Geometria affine, I **45**
3.1 Coordinate affini nel piano . . . 45
3.2 Coordinate affini nello spazio . . . 46
3.3 Giacitura e parallelismo . . . 48
3.4 Spazi affini . . . 49
3.5 Combinazioni lineari di punti . . . 51
3.6 Sottospazi affini . . . 52

4 Applicazioni lineari e matrici **57**
4.1 Applicazioni lineari: definizione e prime proprietà . . . 57
4.2 Isomorfismi . . . 61
4.3 Il primo Teorema di isomorfismo . . . 64
4.4 Matrici . . . 65
4.5 La matrice associata ad un'applicazione lineare . . . 70
4.6 Operazioni elementari sulle matrici, I . . . 72
4.7 Il duale di uno spazio vettoriale . . . 77

4.8 Operazioni elementari sulle matrici, II 80
4.9 Cambiamenti di base e coniugio 83

5 Geometria affine, II **89**
5.1 Applicazioni affini 89
5.2 Composizione di applicazioni affini 91
5.3 Cambiamenti di coordinate affini 92
5.4 Equazioni cartesiane 92

6 Determinanti **95**
6.1 La definizione 95
6.2 Applicazioni multilineari 96
6.3 Proprietà del determinante 98
6.4 La Formula di Binet 100
6.5 Sviluppo di Laplace 101
6.6 Permutazioni e determinante 102
6.7 La formula di Cramer 103
6.8 Determinante e area 104

7 Forme quadratiche e bilineari simmetriche **109**
7.1 Forme quadratiche su k^n 109
7.2 Funzioni polinomiali su uno spazio vettoriale 110
7.3 Forme bilineari simmetriche e forme quadratiche 112
7.4 Ortogonalità 117
7.5 Diagonalizzazione 119
7.6 Spazi vettoriali quadratici 121
7.7 Spazi vettoriali euclidei 124
7.8 Il teorema spettrale 128

8 Coniche e quadriche **137**
8.1 Coniche e quadriche affini 137
8.2 Spazi affini euclidei 139
8.3 Forma canonica euclidea di coniche e quadriche 140

9 Endomorfismi **143**
9.1 Motivazione 143
9.2 Autovalori, autospazi 144
9.3 Molteplicità algebrica e geometrica di autovalori 146
9.4 Forme Hermitiane 149
9.5 Teorema spettrale per operatori autoaggiunti 153
9.6 Il Teorema spettrale per operatori simmetrici 155
9.7 La forma canonica di Jordan 157

Capitolo 0

Introduzione

Queste note sono una introduzione all'algebra lineare e alla trattazione della geometria elementare per mezzo dell'algebra lineare. Un tipico problema di algebra lineare: descrivere le soluzioni di un sistema di equazioni di grado 1, per esempio

$$\begin{aligned} 3x + 2y - z &= 1, \\ x + y + z &= 3, \\ 2x - y + 2z &= 2. \end{aligned}$$

(L'unica soluzione è $x = 0$, $y = 4/3$, $z = 5/3$.) Le equazioni si dicono *lineari* perché, se (x, y, z) sono le coordinate di un punto dello spazio relativamente a un sistema di coordinate cartesiane, allora le soluzioni di una singola equazione sono i punti di un piano. (Quindi stiamo intersecando 3 piani nello spazio, ci aspettiamo che ci sia un solo punto d'intersezione, o una retta in comune, o un piano in comune o nessun punto d'intersezione.) Qui vediamo il collegamento con la geometria (e il motivo per l'uso dell'aggettivo "lineare"). In verità è conveniente iniziare con lo studio dell'algebra lineare e successivamente formulare il concetto di spazio affine o euclideo a partire dal concetto di spazio vettoriale.
Cosa ci aspettiamo che lo studente impari durante il corso? Innanzitutto alcuni semplici algoritmi che permettono di risolvere problemi "concreti" di algebra lineare, per esempio risolvere un sistema di equazioni lineari. Inoltre dovrà imparare a ragionare in astratto, cioè senza scegliere coordinate: se non si sa fare questo si rischia di saper calcolare senza sapere cosa si sta calcolando. Infine si dovrà capire il dizionario "algebra lineare" - "geometria".

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 perché 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 $\mid$ 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 $0, \pm 1, \pm 2, \dots$.

[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.

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 $pq' - p'q = 0$.

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 $a \leq b$ si definiscono i seguenti sottoinsiemi di $\mathbb{R}$:

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

   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 $\mathbb{R}$:

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

   (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 $\emptyset$. Per ogni insieme X abbiamo

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

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 $X_1, \ldots, X_n$ 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 Funzioni

Siano X, Y insiemi.

Definizione 1.2.1. Una *funzione* (o *applicazione*) 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 funzione f e l'insieme Y è il suo *codominio*.

Un chiarimento riguardo la definizione di funzione: si intende che due funzioni $f_1\colon X_1 \to Y_1$ 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 una funzione $f\colon X \to Y$ 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:

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

Se X è un insieme la funzione *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".) Una funzione $f\colon X \to Y$ 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 una funzione $f\colon X \to Y$ 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 una funzione $f\colon X \to Y$; 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 funzione $f\colon X \to Y$ evitando di fare appello al concetto di "legge che associa..." definendo una funzione 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 funzioni (notate: il codominio di f è il dominio di g). Allora possiamo definire una funzione 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 perché 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 degli persone (viventi o morte), f la funzione che associa a una persona suo padre e g la funzione che associa a una persona sua madre. La composizione $f \circ g$ è la funzione che associa a una persona il nonno materno, mentre $g \circ f$ è la funzione 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 la funzione identità: se pensiamo alla composizione di funzioni come analogo della moltiplicazione tra numeri vediamo che la funzione 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 funzioni: hanno senso sia $(h \circ g) \circ f$ che $h \circ (g \circ f)$ e sono entrambe funzioni 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 funzioni gode della proprietà di associatività:

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

Sia $f\colon X \to Y$ una funzione. 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 B, anche detta *immagine inversa*) così:

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

L'*immagine* di f è $\operatorname{im} f := f(X)$. Un esempio: se $f\colon \mathbb{R} \to \mathbb{R}$ è la funzione 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 denotimao $f^{-1}\{y_0\}$ con $f^{-1}y_0$.

Definizione 1.2.3. La funzione $f\colon X \to Y$ 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 $f, g, h, q\colon \mathbb{R} \to \mathbb{R}$ le funzioni definite da

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

La f non è né iniettiva né suriettiva; la g è biunivoca; la h è suriettiva ma non iniettiva; la q è iniettiva ma non suriettiva. Notate che nella definizione di funzione dominio e codominio fanno parte dei dati che definiscono una funzione: quindi una funzione $f\colon X \to Y$ che non è suriettiva può essere "resa"suriettiva sostituendo al codominio Y il codominio $f(Y)$ (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 la funzione $F\colon \mathbb{R} \to \{x \in \mathbb{R} \mid x \geq 1\}$ data dalla stessa formula cioè $F(x) = x^2 + 1$.

Definizione 1.2.4. Sia $f\colon X \to Y$ una funzione **biunivoca**. La *funzione 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 quattro funzioni f, g, h, q definite in (1.2.8) l'unica a essere biunivoca è g quindi ha senso g^{-1} (e non hanno senso né f^{-1} né h^{-1} né q^{-1}) e chiaramente $g^{-1}(y) = y^{1/3}$. Supponiamo che $f\colon X \to Y$ sia biunivoca e sia $B \subset Y$: 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 f^{-1} non ha senso, ha senso solo se è seguito da un sottoinsieme del codominio. Ora supponiamo che $f\colon X \to X$ sia invertibile. Allora ha senso f^m per ogni $m \in \mathbb{Z}$: infatti si pone

$$f^m = \begin{cases} \underbrace{f \circ f \circ \ldots \circ f}_{m} & \text{if } m > 0, \\ 1_X & \text{if } m = 0, \\ \underbrace{f^{-1} \circ f^{-1} \circ \ldots \circ f^{-1}}_{-m} & \text{if } m < 0. \end{cases} \tag{1.2.10}$$

Notiamo che con questa definizione abbiamo che

$$f^m \circ f^n = f^{m+n} \quad \forall m, n \in \mathbb{Z}. \tag{1.2.11}$$

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 $x - y \geq 0$. La relazione $\mathcal{R}$ è quella di "essere non più piccolo" e anziché $x\mathcal{R}y$ scriviamo $x \geq y$.

Esempio 1.3.2. 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 $x \equiv y \pmod{n}$ anziché $x\mathcal{R}_n y$: si legge "x è congruo a y modulo n".

Osservazione 1.3.3. Siano $x, y \in \mathbb{Z}$: 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 1.3.4. 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).

La relazione dell'**Esempio 1.3.1** è di ordine, quella dell'**Esempio 1.3.2** non lo è (quale delle tre proprietà della **Definizione 1.3.4** non vale?). Notate che anche la relazione $\mathcal{R}$ su $\mathbb{R}$ definita da $x\mathcal{R}y$ se $x \leq y$ è di ordine.

Definizione 1.3.5. 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 1.3.2** 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 $x_1 \sim x_2$ anziché $x_1\mathcal{R}x_2$. 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'**Esempio 1.3.2** - allora esistono due classi di equivalenza, il sottoinsieme degli interi pari e quello degli interi dispari.

Definizione 1.3.6. 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 1.3.7. Nell'esempio della congruenza modulo n - vedi l'**Esempio 1.3.2** - 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 1.3.8. 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 $X_{i_1} \cap X_{i_2} = \emptyset$.

Proposizione 1.3.9. *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 1.3.8**. 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 $[x] \cap [y] \neq \emptyset$ 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 1.3.10. Sia X un insieme, $\sim$ una relazione di equivalenza su X e π l'applicazione quoziente di $\sim$. Dato un insieme Y e una funzione $f\colon X \to Y$ 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 $f\colon \mathbb{Z} \to \{0, 1, 2, \ldots, 9\}$ la funzione 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 1.3.3** - 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 perché 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:*

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

Allora $X = \mathbb{N}$.

Dall'assiome segue, sommando tutti gli element di X per un numero naturale N, che

$$\left.\begin{array}{l} N \in X \subset \mathbb{N} \\ \text{se } X \text{ contiene } n, \text{ allora } X \text{ contiene anche } n+1 \end{array}\right\} \implies X \supset \{n \in \mathbb{N} \,|\, n \geq N\}$$

Infatti sia $X \subset \mathbb{N}$ il sottoinsieme degli n tali che valga (1.4.1): per quello che abbiamo dimostrato, la X contiene 1 e, se contiene n, contiene anche $n+1$. Segue che X contiene l'insieme dei naturali maggiori o uguali a 1 cioè la (1.4.1) vale per ogni $n \geq 1$.

1.5 Anelli e campi

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

$$\begin{array}{ccc} A \times A & \longrightarrow & A, \\ (w, z) & \mapsto & w + z \end{array} \tag{1.5.1}$$

e la *moltiplicazione*

$$\begin{array}{ccc} A \times A & \longrightarrow & A, \\ (w, z) & \mapsto & w \cdot z \end{array} \tag{1.5.2}$$

Definizione 1.5.1. Un insieme A con operazioni (1.5.1) e (1.5.2) è un *anello* se

1. Esiste $0 \in A$ tale che $0 + z = z$ 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. Esiste $0 \neq 1 \in A$ tale che $1 \cdot z = z$ per ogni $z \in A$. (Esistenza di una unità per il prodotto.)
6. $z_1 \cdot z_2 = z_2 \cdot z_1$ per ogni $z_1, z_2 \in A$. (Commutatività del prodotto.)
7. $(z_1 \cdot z_2) \cdot z_3 = z_1 \cdot (z_2 \cdot z_3)$ per ogni $z_1, z_2, z_3 \in A$. (Associatività del prodotto.)
8. $z_1 \cdot (z_2 + z_3) = z_1 \cdot z_2 + z_1 \cdot z_3$ 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: non vale (4). Noi saremo interessati soprattutto ad anelli particolare che si chiamano campi.

Definizione 1.5.2. Un anello A è un *campo* se ogni $0 \neq z \in A$ ha un inverso moltiplicativo cioè esiste $w \in A$ tale che $w \cdot z = 1$ (dove 1 è come in (5) della **Definizione 1.5.1**).

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 k.

Proposizione 1.5.3. *Sia A un anello. Allora esiste un* unico *elemento $0 \in A$ tale che valga (1) della* **Definizione 1.5.1** *ed esiste un* unico *elemento $1 \in A$ tale che valga (5) della* **Definizione 1.5.1**. *Dato $z \in A$ esiste un* unico *$w \in A$ tale che valga (4) della* **Definizione 1.5.1**. *Per ogni $z \in A$ si ha che $0 \cdot z = 0$.*

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) di **Definizione 1.5.1**. La dimostrazione che esiste un *unico* elemento $1 \in A$ tale che valga (5) di **Definizione 1.5.1** è del tutto simile. Dimostriamo che dato $z \in A$ esiste un *unico* $w \in A$ tale che valga (4) di **Definizione 1.5.1**. Supponiamo che $z + w = 0 = z + w'$: la commutatività e l'associatività della somma danno

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

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

$$0 \cdot z = (0 + 0) \cdot z = 0 \cdot z + 0 \cdot z \tag{1.5.3}$$

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 (1.5.3) (che sono uguali) otteniamo che $0 = 0 \cdot z$. □

Proposizione 1.5.4. *Sia k un campo. Dato $0 \neq z \in K$ esiste un* unico *elemento $w \in k$ tale che $w \cdot z = 1$. Se $0 \neq z \in k$ 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 1.5.5. *Sia k un campo. Supponiamo che $z, w \in k$ 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 1.5.3**) e quindi $w = 0$ per la **Proposizione 1.5.4**. □

Definizione 1.5.6. Sia A un anello e $z \in A$: l'unico inverso additivo di z viene denotato $-z$. Se k è un campo e $0 \neq z \in k$ l'unico inverso moltiplicativo di z viene denotato z^{-1}.

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 1.5.7. *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

$$a'+b'=a+sn+b+tn=a+b+(s+t)n, \qquad a'b'=(a+sn)(b+tn)=ab+atn+sbn+stn^2=ab+(at+sb+stn)n.$$

□

Per il **Lemma 1.5.7** 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 1.5.5** segue che se n è composto allora $\mathbb{Z}/(n)$ non è un campo. Un risultato non-banale (vedi per esempio Teorema (2.5) di [1]) dà che se n è un numero primo allora $\mathbb{Z}/(n)$ *è* un campo. In generale denotiamo un primo con p e poniamo

$$\mathbb{F}_p := \mathbb{Z}/(p). \tag{1.5.4}$$

Sia k un campo. Siccome $1 \in k$ ha senso la somma

$$\underbrace{1 + \ldots + 1}_{n}$$

che denoteremo n (e $-n$ sarà l'opposto (inverso additivo) di n). Quindi abbiamo associato a ogni $n \in \mathbb{Z}$ un elemnto $n \in k$. Consideriamo il campo $\mathbb{F}_3$: si ha che $3 = 1+1+1 = 0$, e questo dimostra che si puo' avere $n = 0$ *nel campo* k anche se l'intero n non è 0.

Definizione 1.5.8. La *caratteristica* di un campo k (notazione: $\operatorname{char} k$)è 0 se $n \neq 0$ (nel campo k) per ogni intero $n \in \mathbb{Z}$ non nullo ed è il *minimo* $p \in \mathbb{N}$ *non nullo* tale che $p = 0$ *in* k.

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

Osservazione 1.5.9. Sia k un campo di caratteristica $p \neq 0$. Allora p è un numero primo. Infatti supponiamo che $p = ab$ con $a, b \in \mathbb{N}$. Allora nel campo k si ha che $0 = p = a \cdot b$ e per il **Corollario 1.5.5** segue che $a = 0$ o $b = 0$ (in k). Supponiamo che $a = 0$ (in k): siccome $a \leq p$ e p è il minimo intero strettamente positivo tale che $p = 0$ segue che $a = p$ (come numeri naturali). Analogamente se $b = 0$ (in k) segue che $b = p$ (in $\mathbb{N}$).

1.6 Polinomi

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

$$p = a_0 + a_1x + a_2x^2 + \ldots + a_dx^d, \qquad q = b_0 + b_1x + b_2x^2 + \ldots + b_ex^e \tag{1.6.1}$$

polinomi: la somma di p e q è

$$p + q := (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \ldots + (a_d + b_d)x^d, \tag{1.6.2}$$

[2]Sarebbe più preciso dire "trascendente".

il prodotto di p e q è

$$pq := (a_0b_0) + (a_0b_1 + a_1b_0)x + \ldots + (\sum_{i+j=m} a_ib_j)x^m + \ldots + (a_db_e)x^{d+e}. \tag{1.6.3}$$

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 (1.6.2) e (1.6.3). A questo punto se chiamiamo x la successione $(0, 1, 0, \ldots, 0, \ldots)$ ci renderemo 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 k si denota $k[x]$ ed è un anello (ma *non* è un campo, per esempio x non ha un inverso moltiplicativo). Il *grado* di $0 \neq p \in k[x]$ è definito nel seguente modo. Per ipotesi $p = a_0 + a_1x + \ldots + a_dx^d$: poniamo

$$\deg p := \max\{i \mid a_i \neq 0\}. \tag{1.6.4}$$

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

$$\deg(p+q) \leq \max\{\deg p, \deg q\}, \qquad \deg(p \cdot q) = \deg p + \deg q. \tag{1.6.5}$$

(Per convenzione $\max\{-\infty, n\} = n$ $-\infty + n = -\infty$ per ogni $n \in \mathbb{N}$.) Sia $p \in k[x]$: una *radice* (o *zero*) di p è un $\alpha \in k$ tale che $p(\alpha) = 0$.

Definizione 1.6.1. Siano $p, q \in k[x]$. Allora q *divide* p se esiste $m \in k[x]$ tale che $p = m \cdot q$.

Lemma 1.6.2 (Ruffini)**.** *Siano $p \in k[x]$ e $\alpha \in k$. Allora α è una radice di p se e solo se $(x - \alpha)$ divide p.*

Dimostrazione. Se $p = (x - \alpha) \cdot q$ è ovvio che α è una radice di p. Ora supponiamo che α sia una radice di $p = \sum_{i=0}^n c_ix^i$. Scrivendo $x = ((x - \alpha) + \alpha)$ e sostituendo nel polinomio p otteniamo che

$$p = \sum_{i=0}^n c_i((x - \alpha) + \alpha)^i = (x - \alpha) \cdot q + p(\alpha), \qquad q \in k[x].$$

Siccome $p(\alpha) = 0$ segue il risultato. □

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

Definizione 1.6.3. Siano $p \in k[x]$ e $\alpha \in k$. 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 1.6.4. $\alpha \in k$ è radice di p se e solo se la sua molteplicità come radice di p è almeno 1.

Notate che la molteplicità di α è ∞ se e solo se $p = 0$.

Proposizione 1.6.5. *Sia $p \in k[x]$ non nullo di grado n. Allora $\mathrm{mult}_\alpha p$ è non zero per un insieme finito di $\alpha \in k$ e*

$$\sum_{\alpha \in k} \mathrm{mult}_\alpha p \leq \deg p. \tag{1.6.6}$$

Si ha eguaglianza se e solo se si può scrivere

$$p = c \cdot \prod_{i=1}^n (x - \alpha_i) \quad c \neq 0. \tag{1.6.7}$$

Dimostrazione. Per induzione sul grado di p. Se $n = 0$ allora $p \in k$ è non nullo quindi non ha radici: perciò (1.6.6) 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 (1.6.6).) Ora dimostriamo il passo induttivo. Se p non ha radici non c'è nulla da dimostrare: la (1.6.6) vale banalmente. Supponiamo che p abbia una radice γ. Per il **Lemma 1.6.2** esiste $q \in k[x]$ tale che $p = (x - \gamma) \cdot q$: siccome $p \neq 0$ abbiamo che $q \neq 0$. La formula (1.6.5) 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

$$\operatorname{mult}_\gamma p = 1 + \operatorname{mult}_\gamma q, \qquad \operatorname{mult}_{\beta_i} p = \operatorname{mult}_{\beta_i} q \quad \forall 1 \leq i \leq \ell. \tag{1.6.8}$$

Per l'ipotesi induttiva

$$\sum_{\alpha \in k} \operatorname{mult}_\alpha p = 1 + \sum_{\alpha \in k} \operatorname{mult}_\alpha q \leq 1 + \deg q = \deg p.$$

Inoltre vediamo che se si ha equaglianza deve valere $\sum_{\alpha \in k} \operatorname{mult}_\alpha q = \deg q$. Per ipotesi induttiva segue che vale (1.6.7) per $p = q$: segue che vale anche per p. Il viceversa, cioè se vale (1.6.7) allora (1.6.6) è una eguaglianza, è banalmente vero. □

Corollario 1.6.6. *Sia $p \in k[x]$ non nullo. Esistono al più $\deg p$ radici di p.*

Dimostrazione. Segue immediatamente dall'**Osservazione 1.6.4** e da (1.6.6). □

A un polinomio $p = (a_0 + a_1 x + \ldots + a_d x^d) \in k[x]$ possiamo associare la *funzione polinomiale* $k \to k$ (che denotiamo con lo stesso simbolo p) definita da

$$p(x) = a_0 + a_1 x + \ldots + a_d x^d, \qquad x \in k. \tag{1.6.9}$$

Corollario 1.6.7. *Sia k un campo. Sia $d \in \mathbb{N}$ e supponiamo che k abbia più di d elementi. Siano $p, q \in k[x]$ di grado al più d. Le corrispondenti funzioni polinomiali $p, q \colon k \to k$ sono uguali se e solo se $p = q$ (cioè i coefficienti di p e q sono gli stessi). In particolare se 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 $p \in k[x]$ 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 $\alpha_1, \ldots, \alpha_{d+1} \in k$ 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 1.6.5**. □

Il **Corollario 1.6.7** permette di identificare polinomi a coefficienti reali e funzioni polinomiali $\mathbb{R} \to \mathbb{R}$.

Abbiamo considerato polinomi in una indeterminata. Definiremo in modo analogo i polinomi in n indeterminate. Sia $p \colon \mathbb{N}^n \to k$: se $I \in \mathbb{N}^n$ poniamo $p_I := p(I)$.

Definizione 1.6.8. $k[x_1, \ldots, x_n]$ è l'insieme delle funzioni $p \colon \mathbb{N}^n \to k$ 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 k nelle indeterminate*[3] $x_1, \ldots, x_n$ è un elemento di $k[x_1, \ldots, x_n]$.

[3] È più appropriato chiamarle "trascendenti".

Dato $I \in \mathbb{N}^n$ denotiamo con x^I il polinomio che manda I in 1 e $J \neq I$ in 0. Se $I = (0, \ldots, 0)$ denotiamo x^I con 1. Siano $p, q \in k[x_1, \ldots, x_n]$. Definiamo la *somma* $(p+q) \in k[x_1, \ldots, x_n]$ e il *prodotto* $p \cdot q \in k[x_1, \ldots, x_n]$ così:

$$(p+q)_I := p_I + q_I, \qquad (p \cdot q)_I := \sum_{J+K=I} (p_J \cdot p_K). \tag{1.6.10}$$

Notate che la sommatoria che definisce il valore di $p \cdot q$ su I ha senso perché l'insieme delle coppie (J, K) tali che $p_J \neq 0 \neq p_K$ è finito. Inoltre siccome $p, q \in k[x_1, \ldots, x_n]$ anche $p \cdot q$ è una funzione nulla quasi ovunque, cioè è un polinomio. Dato $a \in k$ gli associamo $p_a \in k[x_1, \ldots, x_n]$ con $p_a(0, \ldots, 0) = a$ e $p_a(I) = 0$ per $I \neq (0, \ldots, 0)$. In questo modo abbiamo una inclusione $k \hookrightarrow k[x_1, \ldots, x_n]$. Dato $p \in k[x_1, \ldots, x_n]$ possiamo scrivere

$$p = \sum_{I \in \mathcal{I}} a_I x^I \tag{1.6.11}$$

dove $\mathcal{I} \subset \mathbb{N}^n$ è finito. Con questa scrittura vediamo che la somma e il prodotto di polinomi corrisponde alle operazioni a cui siete abituati dalla scuola media. A un polinomio $p \in k[x_1, \ldots, x_n]$ associamo la *funzione polinomiale*

$$\begin{array}{ccc} k^n & \xrightarrow{p} & k \\ (c_1, \ldots, c_n) & \mapsto & \sum_{I \in \mathbb{N}^n} p_I c^I \end{array} \tag{1.6.12}$$

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 perché p è nulla quasi ovunque.

Definizione 1.6.9. Un polinomio P a coefficienti in k nelle indeterminate $x_1, \ldots, x_n$ è *omogeneo di grado* d se

$$P = \sum_{I \in \mathcal{I}} a_I x^I \tag{1.6.13}$$

dove per ogni $I = (i_1, \ldots, i_n) \in \mathcal{I}$ si ha che $i_1 + \ldots + i_n = d$.

1.7 L'algoritmo euclideo

1.8 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

$$(a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2). \tag{1.8.1}$$

La moltiplicazione è definita così:

$$(a_1, b_1) \cdot (a_2, b_2) := (a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1). \tag{1.8.2}$$

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

$$i := (0, 1). \tag{1.8.3}$$

Osserviamo che

$$i \cdot i = (0,1) \cdot (0,1) = (-1,0) = -1. \tag{1.8.4}$$

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

$$(a,b) = (a,0) + (b,0)(0,1) = a + bi. \tag{1.8.5}$$

Da ora in poi quando diciamo che $(a+bi)$ è un numero complesso intendiamo che $a, b \in \mathbb{R}$. (1.8.6)

Con questa scrittura le definizioni di somma e prodotto danno che

$$(a_1+b_1i)+(a_2+b_2i)=(a_1+a_2)+(b_1+b_2)i, \qquad (a_1+b_1i)(a_2+b_2i)=(a_1a_2-b_1b_2)+(a_1b_2+a_2b_1)i. \tag{1.8.7}$$

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

$$(a+bi)^{-1} = (a^2+b^2)^{-1}(a-bi). \tag{1.8.8}$$

(Qui $(a^2+b^2)^{-1}$ è l'inverso del reale (a^2+b^2) in $\mathbb{R}$.) 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). In questo modo 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 (1.8.6)): poniamo

$$|a+bi| := (a^2+b^2)^{1/2} \tag{1.8.9}$$

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

$$z = \rho((\cos\theta + \sin\theta i), \qquad \rho = |z|, \quad \theta = \mathrm{Arg}(z). \tag{1.8.10}$$

(Se $z = 0$ l'argomento è indeterminato: la (1.8.10) è vera con qualsiasi θ.) Ora siano $z_1, z_2 \in \mathbb{C}$ e scriviamo

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

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

$$z_1z_2=\rho_1\rho_2(\cos\theta_1\cos\theta_2-\sin\theta_1\sin\theta_2+(\cos\theta_1\sin\theta_2+\sin\theta_1\cos\theta_2)i)=\rho_1\rho_2((\cos(\theta_1+\theta_2)+(\sin\theta_1+\theta_2)i). \tag{1.8.11}$$

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

$$|z_1z_2| = |z_1| \cdot |z_2|, \qquad \mathrm{Arg}(z_1z_2) = \mathrm{Arg}(z_1) + \mathrm{Arg}(z_2), \tag{1.8.12}$$

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 1.8.1. *Sia $n > 0$ un numero naturale e $a_1 \ldots, a_n \in \mathbb{C}$. Esiste $z \in \mathbb{C}$ tale che*

$$z^n + a_1z^{n-1} + \ldots + a_{n-1}z + a_n = 0.$$

Applicando ripetutamente il **Lemma 1.6.2** segue che esistono $c_1, \ldots, c_n \in \mathbb{C}$ tali che

$$z^n + a_1 z^{n-1} + \ldots + a_{n-1} z + a_n = (z - c_1)(z - c_2) \cdot \ldots \cdot (z - c_n).. \tag{1.8.13}$$

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

$$\rho^{1/n}(\cos((\theta + s \cdot 2\pi)/n) + \sin((\theta + s \cdot 2\pi)/n)i), \quad 0 \leq s \leq (n-1). \tag{1.8.14}$$

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 1.8.2. 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

$$\overline{z} := a - bi. \tag{1.8.15}$$

Un facile calcolo dà che valgono le formule

$$\overline{w + z} = \overline{w} + \overline{z}, \quad \overline{wz} = \overline{w}\,\overline{z}, \quad z\overline{z} = |z|^2. \tag{1.8.16}$$

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 $1 \leq i < j \leq 3$.

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 $X \setminus Y$ è l'insieme i cui elementi sono gli $x \in X$ che* **non** *sono elementi di Y. Dimostrate che*

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

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

$$\chi_A(x) = \begin{cases} 1 & \text{se } x \in A, \\ 0 & \text{se } x \notin A. \end{cases} \tag{1.8.17}$$

Esercizio 1.7. *Sia X un insieme. Dimostrate che la funzione*

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

è biunivoca. Dimostrate che se X è finito allora

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

Esercizio 1.8. *Di ciascuna delle seguenti funzioni dire se è iniettiva/suriettiva/biunivoca.*

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

Esercizio 1.9. *Sia $f \colon \mathbb{R} \to \mathbb{R}$ definita da $f(x) := x^2 + x + 3$. Determinate* $\operatorname{im} f$.

Esercizio 1.10. *Siano X, Y insiemi e $f \colon X \to Y$ un'applicazione. Siano $A \subset X$ e $B \subset Y$. Verificate che*

$$A \subset f^{-1}(f(A)), \qquad f(f^{-1}(B)) \subset B. \tag{1.8.18}$$

Date esempi in cui le inclusioni di (1.8.18) *sono strette, cioè $A \neq f^{-1}(f(A))$ e $f(f^{-1}(B)) \neq B$.*

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 esiste un'applicazione *suriettiva* $f \colon X \to Y$ diciamo che X ha cardinalià maggiore o uguale a quella di Y - in simboli $X \succeq Y$ (o che Y ha cardinalià minore o uguale a quella di X - n simboli $Y \preceq X$). Se X e Y sono insieme finiti allora X ha la stessa cardinalià di Y se e solo se $|X| = |Y|$ e X ha cardinalià maggiore o uguale a quella di Y se e solo se $|X| \geq |Y|$. (Potreste chiedervi se è vero in generale, come nel caso finito, che $X \succeq Y$ e $Y \succeq X$ implica che $X \approx Y$: la risposta è SÌ, è il contenuto del Teorema di Schröder-Bernstein - vedi l'Appendice 2 di [2]).

Esercizio 1.11. *Dimostrate che $\mathbb{N}$, $\mathbb{Z}$ e $\mathbb{Q}$ hanno la stessa cardinalità.*

Esercizio 1.12. *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 1.13. *Un insieme X è* numerabile *se $\mathbb{N} \succeq X$ cioè se X è finito oppure ha la cardinalità di $\mathbb{N}$. Dimostrate che $\mathbb{R}$* non *è numerabile.*

Esercizio 1.14. *Ridimostrate che vale la* (1.4.1) *osservando che*

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

Esercizio 1.15. *Dimostrate per induzione che*

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

Esercizio 1.16. *Dimostrate per induzione che*

$$1^3 + 2^3 + \ldots + n^3 = \frac{1}{4} n^2 (n+1)^2. \tag{1.8.19}$$

Notate che per la (1.4.1) *la formula* (1.8.19) *equivale alla formula*

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

Esercizio 1.17. *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 1.18. *Calcolate*

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

Esercizio 1.19. *Calcolate le radici quadrate di* $2i$ *e di* $(1 + \sqrt{3}i)$*.*

Esercizio 1.20. *Usando il teorema fondamental dell'algebra, dimostrare che un polinomio a coefficienti reali si può scrivere come prodotto di polinomi reali di grado* 2 *senza radici reali e di polinomi reali di grado* 1*.*

Capitolo 2

Spazi vettoriali

2.1 Gli archetipi e la definizione

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

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

Quindi abbiamo un'operazione

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

Si definisce anche la moltiplicazione

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

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 $k \times V \to V$, che hanno caratteristiche simili a quelle della somma e prodotto per scalari di k^n.

Definizione 2.1.1. Sia k un campo. Uno *spazio vettoriale su* k è un insieme V provvisto di un elemento $0 \in V$ e 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

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

tali che valgano le seguenti proprietà:

1. $0 + v = v$ per ogni $v \in V$ (si dice che 0 è un *elemento neutro*),
2. $(u + v) + w = u + (v + w)$ per ogni $u, v, w \in V$ (proprietà associativa della somma),
3. $u + v = v + u$ per ogni $u, v \in V$ (la somma è commutativa),
4. per ogni $v \in V$ esiste $w \in V$ tale che $v + w = 0$ (esistenza di un opposto),

5. $1v = v$ per ogni $v \in V$,
6. $(\lambda + \mu)v = \lambda v + \mu v$ per ogni $v \in V$ e $\lambda, \mu \in k$ (proprietà distributiva del prodotto),
7. $\lambda(v + w) = \lambda v + \lambda w$ per ogni $v, w \in V$ e $\lambda \in k$ (proprietà distributiva della somma),
8. $(\lambda\mu)v = \lambda(\mu v)$ per ogni $v \in V$ e $\lambda, \mu \in k$.

Gli elementi di uno spazio vettoriale si chiamano *vettori*.

Esempio 2.1.2. Sia $V = k^n$. 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 (1)-(8) della **Definizione 2.1.1**, con elemento neutro $\mathbf{0} := (0, 0, \ldots, 0)$. Quindi k^n provvisto delle operazioni appena definite è uno spazio vettoriale su k.

Terminologia 2.1.3. Uno spazio vettoriale *reale* è uno spazio vettoriale su $\mathbb{R}$, uno spazio vettoriale *complesso* è uno spazio vettoriale su $\mathbb{C}$.

È spesso più utile pensare ai vettori di $\mathbb{R}^n$ non come a qualcosa che ci descriva la posizione di un oggetto nello spazio n-dimensionale, quanto piuttosto ad un'azione di spostamento dentro lo spazio n-dimensionale. Infatti, mentre ha senso effettuare prima uno spostamento e poi un altro (ciò che corrisponde alla somma di vettori), non ha senso pensare ad una somma di posizioni di oggetti. Chiaramente un'azione di spostamento può essere applicata a oggetti che si trovino in posizioni differenti nello spazio n-dimensionale: a tal riguardo, vedete l'Esempio 2.1.4 seguente.

Esempio 2.1.4. Sia $\mathcal{A}^2$ il piano della geometria euclidea (studiato a scuola). Siano $A \neq B \in \mathcal{A}^2$: denoteremo con $\overline{AB}$ l'unica retta contenente A e B. Ricordiamo che due rette sono *parallele* se hanno intersezione vuota oppure coincidono: se $A, B, C, D \in \mathcal{A}^2$ 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 $\mathcal{A}^2$ è una coppia ordinata (A, B) di punti di $\mathcal{A}^2$: 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 $\mathcal{A}^2$ 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 $\mathcal{V}^2 := \mathcal{A}^2/\sim$ è l'insieme dei vettori geometrici. La classe di equipollenza di AB si denota $\overrightarrow{AB}$. Notiamo che dato $P \in \mathcal{A}^2$ e un vettore geometrico v esiste uno e un solo $Q \in \mathcal{A}^2$ tale che $\overrightarrow{PQ} = v$.

Si dà all'insieme $\mathcal{V}^2$ 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 $v, w \in \mathcal{V}^2$ due classi di equipollenza di segmenti orientati. Sia AB un segmento orientato che rappresenta v e sia $C \in \mathcal{A}^2$ l'unico punto tale che BC rappresenti w: quindi ha senso $AB + BC = AC$. Si dimostra che se abbiamo punti $A', B', C' \in \mathcal{A}^2$ 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

$$\begin{array}{ccc} \mathcal{V}^2 \times \mathcal{V}^2 & \longrightarrow & \mathcal{V}^2 \\ \overrightarrow{AB} + \overrightarrow{BC} & \mapsto & \overrightarrow{AC} \end{array}$$

La moltiplicazione per scalari si definisce in modo simile. Sia $v \in \mathcal{V}^2$. 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 $v \in \mathcal{V}^2$ e $\lambda \in \mathbb{R}$ negativo definiamo $\lambda v := (-\lambda)v$ - questo ha senso perché siccome $-\lambda > 0$ il vettore $(-\lambda)v$ è stato definito in precedenza. Ora definiamo il vettore nullo $\mathbf{0} \in \mathcal{V}^2$ come la classe di equipollenza di AA.

Si verifica che $\mathcal{V}^2$ con le operazioni appena definite è uno spazio vettoriale reale.

Esempio 2.1.5. Siccome $\mathbb{R}$ è un sottocampo di $\mathbb{C}$ possiamo dare a $\mathbb{C}$ la struttura di spazio vettoriale su $\mathbb{R}$.

Esempio 2.1.6. Sia k un campo. Sull'insieme dei polinomi $k[x]$ sono definite le operazioni di somma e prodotto di polinomi. Siccome $k \subset k[x]$ (i polinomi "costanti"), possiamo definire un prodotto scalare $k \times k[x] \to k[x]$. Con queste operazioni $k[x]$ è un k-spazio vettoriale.

Esempio 2.1.7. Sia k un campo e sia $d \in \mathbb{N}$. Sia $k[x]_{\leq d}$ l'insieme che contiene i polinomi a coefficienti in k di grado al più d e il polinomio nullo. Come per $k[x]$, possiamo analogamente definire le operazioni di somma tra due polinomi e di prodotto di un polinomio per scalare su $k[x]_{\leq d}$. Con queste operazioni $k[x]_{\leq d}$ è un k-spazio vettoriale.

Esempio 2.1.8. Siano k un campo e X un insieme. Possiamo dotare l'insieme k^X delle funzioni $f\colon X \to k$ della struttura di un 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 2.1.9. Scegliamo una unità di misura nel piano euclideo $\mathcal{A}^2$. Allora, dato un vettore v nel piano euclideo $\mathcal{V}^2$, 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 $v, w \in \mathcal{V}^2$, 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 sul campo k. 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 $0_1 = 0_2$ - *questo è* l'elemento neutro *di* V *e sarà denotato* 0. *Analogamente, dato* $v \in V$ *esiste un unico* $w \in V$ *tale che* $v + w = 0$.

Dimostrazione. Applichiamo le equazioni di (2.2.1) con $v = 0_2$ e $v = 0_1$, e ricordiamo che la somma di vettori è commutativa; otteniamo che

$$0_2 = 0_1 + 0_2 = 0_2 + 0_1 = 0_1.$$

Segue che $0_1 = 0_2$. Ora supponiamo che $v + w_1 = 0 = v + w_2$. Per la proprietà associativa della somma,

$$w_1 = 0 + w_1 = (w_2 + v) + w_1 = w_2 + (v + w_1) = w_2 + 0 = w_2,$$

e quindi $w_1 = w_2$. □

Osserviamo anche che, se V è uno spazio vettoriale sul campo k, valgono le seguenti uguaglianze:

$$0v = 0, \quad \lambda 0 = 0, \quad (-1)v + v = 0, \qquad \forall v \in V, \lambda \in k. \tag{2.2.2}$$

Infatti $0v = (0 + 0)v = 0v + 0v$ e aggiungendo l'opposto di $0v$ a entrambi i membri otteniamo che $0 = 0v$. La dimostrazione della seconda equaglianza è del tutto simile. Infine $(-1)v + v = (-1)v + 1v = (-1 + 1)v = 0v = 0$ dà che $(-1)v + v = 0$.

Terminologia 2.2.2. Sia V uno spazio vettoriale sul campo k. Dato $v \in V$ l'unico $w \in V$ tale che $v + w = 0$ è l'*opposto* di v e sarà denotato $-v$ (in accordo con la terza eguaglianza di (2.2.2)).

Osservazione 2.2.3. Sia V uno spazio vettoriale sul campo k. Si denota con lo stesso simbolo sia l'elemento neutro del campo k, che l'elemento neutro dello spazio vettoriale: attenzione a non confondere i due elementi neutri!

2.3 Sottospazi

Definizione 2.3.1. Sia V uno spazio vettoriale su k. Un sottoinsieme W di V è un *sottospazio* di V se è non vuoto e se dati $v_1, v_2 \in W$, $\lambda_1, \lambda_2 \in k$ si ha che $(\lambda_1 v_1 + \lambda_2 v_2) \in W$.

Esempio 2.3.2. Siano k un campo e $a_1, \ldots, a_n \in k$. Siano

$$W_1 := \{(x_1,\ldots,x_n) \in k^n \mid a_1x_1 + a_2x_2 + \ldots + a_nx_n = 0\}, \quad W_2 := \{(x_1,\ldots,x_n) \in k^n \mid a_1x_1 + a_2x_2 + \ldots + a_nx_n = 1\}. \tag{2.3.1}$$

Si verifica facilmente che W_1 è un sottospazio di V e che W_2 *non* è un sottospazio di V.

Esempio 2.3.3. 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 2.3.4. L'insieme dei polinomi $\mathbb{R}[x]_{\leq d}$ di grado al più d è un $\mathbb{R}$-sottospazio vettoriale dello spazio vettoriale $\mathbb{R}[x]$ dei polinomi reali di grado arbitrario.

Osservazione 2.3.5. Siano k un campo e V uno spazio vettoriale su k. Un sottoinsieme $W \subset V$ è un sottospazio se e solo se W è chiuso per l'operazione di somma e per la moltiplicazione per scalari di V e, provvisto di queste operazioni è uno spazio vettoriale su k. In particolare un sottospazio contiene l'elemento neutro di V.

Lemma 2.3.6. *Sia V uno spazio vettoriale su 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 $\lambda \in k$. 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 2.3.7. Applichiamo il **Lemma 2.3.6** all'insieme delle soluzioni di un sistema di *equazioni lineari omogenee* cioè l'insieme degli $(x_1, x_2, \ldots, x_n) \in k^n$ tali che

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots & = & 0, \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n & = & 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots & = & 0, \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & 0. \end{array} \tag{2.3.2}$$

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 k^n l'insieme delle soluzioni di m equazioni è l'intersezione di m sottospazi vettoriali di k^n; per il **Lemma 2.3.6** è un sottospazio vettoriale di k^n.

Proposizione 2.3.8. *Sia V uno spazio vettoriale su k e $S \subset V$ un sottoinsieme. Esiste un unico sottospazio vettoriale $U \subset V$ che ha le seguenti proprietà:*

1. *S è contenuto in U.*
2. *Se W è un sottospazio vettoriale di V che contiene S allora $U \subset W$.*

(Informalmente U è il più piccolo *sottospazio vettoriale di V che contiene S).*

Dimostrazione. Sia $\mathcal{F}$ la famiglia dei sottospazio vettoriali di V che contengono S - notate che $\mathcal{F}$ non è vuota perché $S \subset V$. Sia U l'intersezione dei sottospazi in $\mathcal{F}$: è un sottospazio di V per il **Lemma 2.3.6**. Chiaramente U contiene S, inoltre se W è un sottospazio vettoriale di V che contiene S allora $W \in \mathcal{F}$ e quindi $U \subset W$. Quindi U soddisfa sia (1) che (2). Ora supponiamo che esista un sottospazio vettoriale $U' \subset V$ tale che valgano (1) e (2) con U sostituito da U': siccome $S \subset U$ segue che $U' \subset U$. D'altra parte siccome valgono (1) e (2) e $S \subset U'$ abbiamo che $U \subset U'$: quindi $U = U'$. □

Definizione 2.3.9. Sia V uno spazio vettoriale su k. Sia $S \subset V$ un sottoinsieme: l'unico sottospazio $U \subset V$ tale che valgano (1) e (2) della **Proposizione 2.3.8** è chiamato il sottospazio vettoriale *generato* da S e si denota $\langle S \rangle$ (in alcuni libri, tale sottospazio è anche denotato come "span(S)"). Se $v_1, \ldots, v_n \in V$ denotiamo $\langle \{v_1, \ldots, v_n\} \rangle$ con $\langle v_1, \ldots, v_n \rangle$.

Esiste una descrizione più esplicita del sottospazio generato da un sottoinsieme S di uno spazio vettoriale V. Prima diamo una definizione fondamentale.

Definizione 2.3.10. Sia V uno spazio vettoriale su 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 $\lambda_1, \lambda_2, \ldots, \lambda_n \in k$ tali che

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n. \tag{2.3.3}$$

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.

Proposizione 2.3.11. *Siano V uno spazio vettoriale su k e $S \subset V$ un sottoinsieme. Allora $\langle S \rangle$ è uguale all'insieme i cui elementi sono le combinazioni lineari di arbitrarie collezioni finite di vettori in S:*

$$\langle S \rangle = \{\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n \mid v_1, v_2, \ldots, v_n \in V, \quad \lambda_1, \lambda_2, \ldots, \lambda_n \in k\}. \tag{2.3.4}$$

Dimostrazione. Sia U il membro di destra di (2.3.4). Dobbiamo dimostrare che $\langle S \rangle = U$. È sufficiente dimostrare che U è un sottospazio di V e che valgono (1) e (2) della **Proposizione 2.3.8**. Verifichiamo che U è un sottospazio di V: non è vuoto perché $0 \in U$ (vedi l'ultima frase della **Definizione 2.3.10**) e se $u, u' \in U$ cioè

$$u = \sum_{i=1}^{n} \lambda_i v_i, \quad u' = \sum_{j=1}^{n} \mu_j w_j, \qquad v_i, w_j \in S,\ \lambda_i, \mu_j \in k$$

allora $u + u' = \sum_{i=1}^{n} \lambda_i v_i + \sum_{j=1}^{n} \mu_j w_j$ e quindi appartiene a U, inoltre se $\alpha \in k$ abbiamo che $\alpha u = \sum_{i=1}^{n} (\alpha \lambda_i) v_i$ e quindi appartiene a U. Ora dimostriamo che $S \subset U$. Sia $v \in S$: siccome $v = 1v$ abbiamo che $v \in U$. Rimane da dimostrare che se $W \subset V$ è un sottospazio contenente S allora $U \subset W$. Sia $v \in U$ cioè vale (2.3.3) con $v_i \in S$, allora $v_1, v_2, \ldots, v_n \in W$ e siccome W è un sottospazio segue che $v \in W$. □

Il seguente corollario è semplice ma utile.

Corollario 2.3.12. *Sia V uno spazio vettoriale su k e siano $S, R \subset V$ sottoinsiemi. Allora $R \subset \langle S \rangle$ se e solo se $\langle S \cup R \rangle = \langle S \rangle$.*

Dimostrazione. Supponiamo dapprima $\langle S \cup R \rangle = \langle S \rangle$. Allora $R \subset \langle S \cup R \rangle = \langle S \rangle$.

Supponiamo invece ora $R \subset \langle S \rangle$. Chiaramente $\langle S \rangle \subset \langle S \cup R \rangle$. Vogliamo dimostrare che $\langle S \cup R \rangle \subset \langle S \rangle$.
Sia $v \in \langle S \cup R \rangle$. Esistono elementi $v_1, \ldots, v_n \in S$ e $w_1, \ldots, w_m \in R$ e coefficienti $a_1, \ldots, a_n, b_1, \ldots, b_m \in k$ tali che

$$v = a_1 v_1 + \cdots + a_n v_n + b_1 w_1 + \cdots + b_m w_m$$

Poiché $w_i \in R \subset \langle S \rangle$, possiamo anche scrivere

$$w_i = c_{i,1} u_{i,1} + \cdots + c_{i,r} u_{i,r}$$

dove $u_{i,j} \in S$ e $c_{i,j} \in k$. Sostituendo, otteniamo

$$v = \sum_{h=1}^{n} a_h v_h + \sum_{i=1}^{m} \sum_{j=1}^{r} b_i c_{i,j} u_{i,j}$$

e dunque v è combinazione lineare di elementi in S. Ne segue che $v \in \langle S \rangle$. □

Definizione 2.3.13. Sia V uno spazio vettoriale su k. Un sottospazio $W \subset V$ è *finitamente generato* se è generato da un insieme finito.

Esempio 2.3.14. Lo spazio vettoriale k^n è finitamente generato su k perché è generato dai vettori

$$\mathbf{e}_1 := (1, 0, \ldots, 0),\ \mathbf{e}_2 := (0, 1, 0, \ldots, 0), \ldots, \mathbf{e}_n := (0, 0, \ldots, 1). \tag{2.3.5}$$

Esempio 2.3.15. Sia

$$\mathbb{R}[x]_{\leq d} := \{p \in \mathbb{R}[x] \mid p = 0 \text{ o } p \neq 0 \text{ e } \deg p \leq d\}.$$

Allora $\mathbb{R}[x]_{\leq d}$ è un sottospazio finitamente generato di $\mathbb{R}[x]$, perché è generato da $\{1, x, \ldots, x^d\}$.

Esempio 2.3.16. Lo spazio vettoriale $k[x]$ (con somma e moltiplicazioni puntuali) *non* è finitamente generato su k. Infatti, siano $f_1, \ldots, f_m \in k[x]$, e dimostriamo che $\langle f_1, \ldots, f_m \rangle \neq k[x]$. Possiamo assumere che i polinomi $f_1, \ldots, f_m$ siano tutti non nulli perché il sottopsazio 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 $k[x]$ perché esistono polinomi di grado arbitrariamente alto.

Definizione 2.3.17. Sia V uno spazio vettoriale e $U, W \subset B$ sottospazi. La *somma* $U + W$ è il sottospazio di V definito da

$$U + W := \langle U \cup W \rangle = \{u + w \mid u \in U, \quad w \in W\}. \tag{2.3.6}$$

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

2.4 Dipendenza/indipendenza lineare

La seguente è una definizione fondamentale.

Definizione 2.4.1. Sia V uno spazio vettoriale su k. Siano $v_1, \ldots, v_n \in V$. Una relazione lineare tra $v_1, \ldots, v_n$ è una uguaglianza

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0, \tag{2.4.1}$$

dove $\lambda_1, \lambda_2, \ldots, \lambda_n \in k$. La relazione lineare (2.4.1) è *non banale* se $\lambda_1, \lambda_2, \ldots, \lambda_n$ *non* sono tutti nulli. Diciamo che $v_1, \ldots, v_n$ sono *linearmente dipendenti* se esiste una relazione lineare non banale tra $v_1, \ldots, v_n \in V$, e che sono *linearmente indipendenti* in caso contrario.

Esplicitiamo la definizione di vettori $v_1, \ldots, v_n$ linearmente indipendenti: vuol dire che

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

solo se $0 = \lambda_1 = \lambda_2 = \ldots = \lambda_n$.

Esempio 2.4.2. I vettori $\mathbf{e}_1, \ldots, \mathbf{e}_n \in k^n$ definiti da (2.3.5) sono linearmente indipendenti, i vettori $v_1 = (2, 2)$, $v_2 = (3, 3)$ di $\mathbb{R}^2$ sono linearmente dipendenti perché $3v_1 - 2v_2 = \mathbf{0}$.

Esempio 2.4.3. Siano $v_1, v_2 \in k^2$ 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

$$x_1 v_1 + x_2 v_2 = \mathbf{0}$$

cioè

$$ax_1 + cx_2 = 0, \qquad bx_1 + dx_2 = 0. \tag{2.4.2}$$

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

$$(bc - ad)x_2 = 0. \tag{2.4.3}$$

D'altra parte moltiplicando la prima equazione di (2.4.2) per d e aggiungendogli la seconda equazione moltiplicata per $-c$ otteniamo che

$$(ad - bc)x_1 = 0. \tag{2.4.4}$$

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 (2.4.2) e per (2.4.3) e (2.4.4) 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 $1 \cdot v_1 + 0 \cdot v_2 = \mathbf{0}$). Quindi possiamo supporre che $(a, c) \neq (0, 0)$ o $(b, d) \neq (0, 0)$. Nel primo caso una soluzione non banale di (2.4.2) è data da $x_1 = c$, $x_2 = -a$, nel secondo caso una soluzione non banale di (2.4.2) è data da $x_1 = d$, $x_2 = -b$.

Definizione 2.4.4. Una *matrice* 2×2 con entrate in un campo k è una collezione ordinata M di 4 elementi di 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

$$\det M := (ad - bc). \tag{2.4.5}$$

L'**Esempio 2.4.3** dà un caso in cui è utile disporre della nozione di matrice 2×2 e suo determinante: infatti abbiamo visto che i vettori $v_1, v_2 \in k^2$ sono linearmente dipendenti se e solo se è nullo il determinante della matrice 2×2 che ha come righe i vettori v_1 e v_2. Nel **Capitolo 4** considereremo matrici di ordine qualsiasi, e nel **Capitolo 6** definiremo il determinante di matrici quadrate di ordine arbitrario.

Osservazione 2.4.5. 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. Tuttavia, è chiaro che la dipendenza/indipendenza lineare non varia se cambiamo l'ordine dei vettori nella lista. Dunque ha senso dire che un insieme $S \subset V$ di vettori è linearmente indipendente se, comunque ordinati i vettori in S, sono linearmente indipendenti.

Lemma 2.4.6. *Sia W uno spazio vettoriale su k non finitamente generato. Allora, per ogni $m \geq 1$, esiste una m-upla $w_1, \ldots, w_m \in W$ di vettori linearmente indipendenti.*

Dimostrazione. Per induzione su $m \geq 1$.
Nel caso $m = 1$, notiamo che $W \neq \{0\}$, in quanto lo spazio vettoriale $\{0\}$ è finitamente generato. Sia dunque $w_1 \in W$ con $w_1 \neq 0$. Chiaramente w_1 è linearmente indipendente.
Supponiamo ora vera l'asserzione per m e quindi supponiamo di aver costruito una m-upla $w_1, \ldots, w_m \in W$ linearmente indipendente. L'insieme $\{w_1, \ldots, w_m\}$ non può generare W, in quanto W non è finitamente generato, e dunque esiste un vettore $w_{m+1} \in W \setminus \langle w_1, \ldots, w_m \rangle$. Asseriamo che $w_1, \ldots, w_{m+1}$ siano linearmente indipendenti. Infatti, se $a_1 w_1 + \cdots + a_{m+1} w_{m+1} = 0$ per opportuni $a_i \in k$, dovrebbe aversi $a_{m+1} = 0$ (perché $w_{m+1} \notin \langle w_1, \ldots, w_m \rangle$) e di conseguenza $a_1 = \cdots = a_m = 0$ (perché $w_1, \ldots, w_m$ sono linearmente indipendenti). Questo completa il passo induttivo. □

Osservazione 2.4.7. L'affermazione "i vettori $v_1, \dots, v_n$ sono linearmente dipendenti" è un'affermazione sulla *lista* di vettori $v_1, \dots, v_n$, *non* si afferma che *ciascun* vettore della lista $v_1, \dots, v_n$ ha la proprietà di essere "linearmente indipendente". Questa è un'osservazione banale, ma esiste il pericolo di fraintendimento perché a rigore bisognerebbe affermare che "la lista $v_1, \dots, v_n$ è linearmente dipendente". Ovviamente, analoghe considerazione valgono per l'affermazione "i vettori $v_1, \dots, v_n$ sono linearmente indipendenti".

Osservazione 2.4.8. È possibile definire relazioni lineari e dipendenza/indipendenza lineare per un insieme qualunque (anche infinito!) di vettori $S \subset V$. Una relazione lineare fra vettori della lista S è una uguaglianza

$$a_1 v_1 + \cdots + a_h v_h = 0$$

dove $v_1, \dots, v_h$ sono vettori distinti in S e $a_1, \dots, a_h \in k$, e tale relazione lineare si dice *banale* se $a_1 = \cdots = a_h = 0$. Sottolineiamo che una relazione lineare coinvolge un numero finito di addendi!
Come nel caso di una lista finita, diciamo che S è *linearmente indipendente* se non ci sono relazioni lineari non banali in S, e *linearmente dipendente* altrimenti.

2.5 Basi

La seguente è una definizione fondamentale.

Definizione 2.5.1. Sia V uno spazio vettoriale su k. Siano $v_1, \dots, v_n \in V$. Diciamo che $(v_1, \dots, v_n)$ è una *base (finita)* di V se $v_1, \dots, v_n$ generano V e sono linearmente indipendenti.

Un esempio: i vettori $\mathbf{e}_1, \dots, \mathbf{e}_n \in k^n$ definiti da (2.3.5) formano una base di k^n: questa è la *base standard* di k^n. Notate che, se V ha una base finita, allora V è finitamente generato, quindi per esempio lo spazio vettoriale $k[x]$ non ha una base finita. Esiste una nozione più generale di base che ammette il caso di basi con infiniti elementi: ogni spazio vettoriale ammette una base secondo la definizione più generale, vedi [2].

Come nel caso di una lista linearmente indipendente di vettori, diciamo che un insieme finito $\{v_1, \dots, v_n\}$ di vettori in V è una *base (finita)* se, comunque si ordinino tali $v_1, \dots, v_n$, essi formino una base finita.

Proposizione 2.5.2. *Sia V uno spazio vettoriale su k. Sia $S \subset V$ un sottoinsieme linearmente indipendente in V e siano $w_1, \dots, w_m \in V$ tali che $S \cup \{w_1, \dots, w_m\}$ generi V.*
Allora vale una ed una sola delle due seguenti asserzioni:

(a) la lista $S \cup \{w_1, \dots, w_m\}$ è una base di V;

(b) esiste $1 \leq i \leq m$ tale che $S \cup \{w_1, \dots, w_{i-1}, w_{i+1}, \dots, w_m\}$ generi V.

Dimostrazione. Se $S \cup \{w_1, \dots, w_{m+1}\}$ è linearmente indipendente, allora segue la tesi nel caso (a).
Supponiamo dunque che $S \cup \{w_1, \dots, w_{m+1}\}$ non sia linearmente indipendente, e quindi che esistano $v_1, \dots, v_h \in S$ e coefficienti $a_1, \dots, a_h, b_1, \dots, b_{m+1} \in k$ non tutti nulli tali che

$$a_1 v_1 + \cdots + a_h v_h + b_1 w_1 + \cdots + b_{m+1} w_{m+1} = 0$$

Poiché S è linearmente indipendente, non si può avere $b_1 = b_2 = \cdots = b_{m+1} = 0$ e dunque esiste un $1 \leq i \leq m+1$ tale che $b_i \neq 0$. Allora, sommando ambo i membri per $-b_i w_i$ e

dividendo per $-b_i$, otteniamo

$$w_i = -\frac{a_1}{b_i}v_1 - \cdots - \frac{a_h}{b_i}v_h - \frac{b_1}{b_i}w_1 - \cdots - \frac{b_{i-1}}{b_i}w_{i-1} - \frac{b_{i+1}}{b_i}w_{i+1} - \cdots - \frac{b_{m+1}}{b_i}w_{m+1}$$

Dunque $w_i \in \langle S \cup \{w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_{m+1}\}$. Per il **Corollario 2.3.12**, si ha

$$\langle S \cup \{w_1, \ldots, w_{m+1}\}\rangle = \langle S \cup \{w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_{m+1}\}\rangle$$

e quindi $S \cup \{w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_{m+1}\}$ è un insieme di generatori per V. □

Corollario 2.5.3. *Sia V uno spazio vettoriale su k. Sia $S \subset V$ un sottoinsieme di vettori linearmente indipendenti e siano $w_1, \ldots, w_m \in V$ tali che $S \cup \{w_1, \ldots, w_m\}$ generi V.*
Allora esistono h e $1 \leq j_1 < \cdots < j_h \leq m$ tali che $S \cup \{w_{j_1}, \ldots, w_{j_h}\}$ sia una base di V.

Dimostrazione. Procediamo per induzione su $m \geq 0$.
Per $m = 0$, l'enunciato segue dalle assunzioni. Supponiamo ora l'enunciato vero per m e dimostriamolo per $m + 1$.
Consideriamo $S \cup \{w_1, \ldots, w_{m+1}\}$. Per la proposizione precedente, si hanno due casi. Nel caso (a), $S \cup \{w_1, \ldots, w_{m+1}\}$ è linearmente indipendente e quindi otteniamo la tesi con $k = m + 1$ e $j_h = h$. Nel caso (b), abbiamo che $S \cup \{w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_{m+1}\}$ è un insieme di generatori. Per ipotesi induttiva, otteniamo una base $S \cup \{w_{j_1}, \ldots, w_{j_h}\}$ di V per qualche h, con $j_1, \ldots, j_h \neq i$. In ogni caso, abbiamo ottenuto la tesi. □

Ne discendono alcune utili conseguenze.

Corollario 2.5.4. *Sia V uno spazio vettoriale su k generato dai vettori $v_1, \ldots, v_n$. Allora esiste una base di V ottenuta eliminando alcuni dei v_i, cioè esistono $1 \leq i_1 < i_2 < \ldots < i_h \leq n$ tali che $\{v_{i_1}, v_{i_2}, \ldots, v_{i_h}\}$ sia una base di V.*

Dimostrazione. Segue dal **Corollario 2.5.3**, ponendo $S = \emptyset$ e usando l'insieme di generatori $\{v_1, \ldots, v_n\}$. □

Corollario 2.5.5. *Sia V uno spazio vettoriale su k finitamente generato. Siano $v_1, \ldots, v_n \in V$ linearmente indipendenti: esistono $v_{n+1}, \ldots, v_{n+h} \in V$ tali che $\{v_1, \ldots, v_n, v_{n+1}, \ldots, v_{n+h}\}$ sia una base di V. (Il caso $h = 0$ è ammesso: significa che $\{v_1, \ldots, v_n\}$ è una base di V.)*

Dimostrazione. Sia $\{w_1, \ldots, w_m\}$ un insieme di generatori di V. Ponendo $S = \{v_1, \ldots, v_n\}$, abbiamo che S è linearmente indipendente e chiaramente $S \cup \{w_1, \ldots, w_m\}$ genera V. Applicando il **Corollario 2.5.3**, otteniamo una base $S \cup \{w_{i_1}, \ldots, w_{i_h}\}$ e dunque la tesi segue ponendo $v_{n+j} := w_{i_j}$. □

Corollario 2.5.6. *Sia V uno spazio vettoriale su k. Supponiamo che $v_1, \ldots, v_r, u \in V$ siano linearmente indipendenti (il caso $r = 0$ è ammesso) e che $v_1, \ldots, v_r, w_1, \ldots, w_s$ siano generatori di V. Allora esiste $1 \leq i \leq s$ tale che V sia generato da*

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

Dimostrazione. Prendendo $S = \{v_1, \ldots, v_r, u\}$, si ha che $S \cup \{w_1, \ldots, w_s\}$ genera V. Possiamo quindi applicare la **Proposizione 2.5.2**.

Poiché anche $S \cup \{w_1, \ldots, w_s\} \setminus \{u\}$ genera V, ne segue che $u \in \langle S \cup \{w_1, \ldots, w_s\}\rangle$ e dunque u si scrive come combinazione lineare dei v_j e dei w_j. Quindi $S \cup \{w_1, \ldots, w_s\}$ non è linearmente indipendente e non siamo nel caso (a) della **Proposizione 2.5.2**: siamo nel caso (b), che è esattamente la tesi voluta. □

Siano $v_1, \dots, v_n \in V$ e $\mathcal{B} := \{v_1, \dots, v_n, v_{n+1}, \dots, v_{n+h}\}$ come nell'enunciato del **Corollario 2.5.5**: si dice che $v_1, \dots, v_n \in V$ 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.

Proposizione 2.5.7. *Sia V uno spazio vettoriale su k di dimensione n e sia $v_1, \dots, v_m \in V$ una lista (ordinata) di vettori. Sia infine dato un vettore $v \in V$.*

(i) Supponiamo che $v_1, \dots, v_m$ generino V. Allora esistono $a_1, \dots, a_m \in k$ tali che $v = a_1v_1 + a_2v_2 + \cdots + a_mv_m$.

(ii) Supponiamo che $v_1, \dots, v_m$ siano linearmente indipendenti. Se v si può scrivere come $v = a_1v_1 + a_2v_2 + \cdots + a_mv_m$ con $a_1, \dots, a_m \in k$, allora tale m-upla $(a_1, \dots, a_m) \in k^m$ è unica.

(iii) Supponiamo che $v_1, \dots, v_m$ sia una base di V (e quindi $n = m$). Allora $\exists!\, a_1, \dots, a_m \in k$ tali che $v = a_1v_1 + a_2v_2 + \cdots + a_mv_m$.

Dimostrazione. La (i) segue dalla definizione di insieme di generatori e dalla **Proposizione 2.3.11**.
Per quanto riguarda la (ii), supponiamo che $v \in V$ si scriva come combinazione lineare dei $v_1, \dots, v_m$ in due modi, ossia che esistano $a_1, \dots, a_m, b_1, \dots, b_m \in k$ tali che

$$a_1v_1 + a_2v_2 + \cdots + a_mv_m = v = b_1v_1 + b_2v_2 + \cdots + b_mv_m$$

Ne segue che $(a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \cdots + (a_m - b_m)v_m = 0$ e dunque $a_1 - b_1 = a_2 - b_2 = \cdots = a_m - b_m = 0$ perché $v_1, \dots, v_m$ sono linearmente indipendenti. Dunque $a_i = b_i$ per ogni $1 \leq i \leq m$ e le due scritture coincidono.
Infine, la (iii) si ottiene combinando (i) e (ii). □

Come immediata conseguenza della **Proposizione 2.5.7**(iii) otteniamo che i vettori di uno spazio vettoriale su k possono essere messi in corrispondenza con quelli di k^n, quando si sia scelta una base (ordinata) di n vettori.

Corollario 2.5.8. *Sia V uno spazio vettoriale su k finitamente generato e $\mathcal{B} = (v_1, \dots, v_n)$ una sua base (ordinata). Allora l'applicazione*

$$\begin{array}{ccc} k^n & \xrightarrow{f} & V \\ (a_1, \dots, a_n) & \longrightarrow & a_1v_1 + a_2v_2 + \ldots + a_nv_n \end{array} \tag{2.5.1}$$

è biunivoca.

Sia f data da (2.5.1): per il **Corollario 2.5.8** la f è biunivoca e quindi è definita la sua inversa che denotiamo $X_{\mathcal{B}}$:

$$V \xrightarrow{X_{\mathcal{B}}} k^n \tag{2.5.2}$$

La n-pla $X_{\mathcal{B}}(v)$ associata a v è l'n-pla delle sue *coordinate* relative alla base $\mathcal{B}$.

Esempio 2.5.9. Sia $\mathcal{S} := \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ la base standard di k^n. Le coordinate di $X = (x_1, \dots, x_n)$ nella base $\mathcal{S}$ sono date da $(x_1, \dots, x_n)$.

Esempio 2.5.10. 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. La seguente proposizione rappresenta, nel caso di uno spazio vettoriale finitamente generato, una versione più precisa del **Corollario 2.5.3**. La dimostrazione procede similmente per induzione, ma i ruoli sono invertiti.

Proposizione 2.5.11. *Sia V uno spazio vettoriale su k. Supponiamo che $v_1, \dots, v_n \in V$ siano linearmente indipendenti e che $w_1, \dots, w_m \in V$ siano generatori di V. Allora $n \leq m$ ed esistono $1 \leq j_1 < j_2 \leq \dots < j_h \leq m$ tali che V è generato da*

$$\{v_1, \dots, v_n\} \cup \{w_1, \dots, w_m\} \setminus \{w_{j_1}, w_{j_2}, \dots, w_{j_h}\}. \tag{2.5.3}$$

(In altre parole: sostituendo nella lista $w_1, \dots, w_h$ ciascun w_{j_i} con v_i otteniamo un nuovo sistema di generatori.)

Dimostrazione. Per induzione su $n \geq 0$. Più precisamente sia A_n l'affermazione della proposizione: dimostriamo per induzione che è vera per ogni n.
Il caso $n = 0$ è banalmente vero. Dimostriamo il passo induttivo, cioè assumiamo che A_n sia vera e dimostriamo che è vera A_{n+1}.
Per l'ipotesi induttiva V è generato da (2.5.3). Applicando il **Corollario 2.5.6** con $r = n$ e $u = v_{n+1}$, ossia considerando l'insieme $\{v_1, \dots, v_n, u = v_{n+1}\}$ di vettori linearmenti indipendenti tale che $\{v_1, \dots, v_n\} \cup \Big(\{w_1, \dots, w_m\} \setminus \{w_{j_1}, \dots, w_{j_h}\}\Big)$ genera, vediamo che vale A_{n+1}. □

Corollario 2.5.12. *Sia V uno spazio vettoriale su k, finitamente generato. Supponiamo che $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_m\}$ siano basi di V. Allora $n = m$.*

Dimostrazione. I vettori $v_1, \dots, v_n$ sono linearmente indipendenti e $w_1, \dots, w_m$ sono generatori di V: per la **Proposizione 2.5.11** abbiamo che $n \leq m$. D'altra parte i vettori $w_1, \dots, w_m$ sono linearmente indipendenti e $v_1, \dots, v_n$ sono generatori di V: per la **Proposizione 2.5.11** abbiamo che $m \leq n$: segue che $n = m$. □

Un altro semplice corollario è il seguente.

Corollario 2.5.13. *Sia V uno spazio vettoriale su k finitamente generato e sia $W \subset V$ un sottospazio. Allora W è finitamente generato.*

Dimostrazione. Per il **Corollario 2.5.4**, V ammette una base $\{v_1, \dots, v_n\}$. Per il **Lemma 2.4.6**, esiste un sottoinsieme linearmente indipendente $\{w_1, \dots, w_{n+1}\}$ di $W \subset V$ composto da $n + 1$ vettori. Questo contraddice la **Proposizione 2.5.11**. □

Il **Corollario 2.5.12** ci permette di dare la seguente definizione fondamentale.

Definizione 2.5.14. Sia V uno spazio vettoriale su k finitamente generato (e quindi V ammette basi per il **Corollario 2.5.4**). La *dimensione* di V è la cardinalità di una qualsiasi base di V.

Esempio 2.5.15.

1. Lo spazio vettoriale k^n ha dimensione n perché $\mathbf{e}_1, \dots, \mathbf{e}_n$ è una base di k^n.
2. Se $k' \subset k$ è un sottocampo, allora restringendo gli scalari a k', diamo a k^n una struttura di spazio vettoriale su k', e la dimensione di k^n come k'-spazio vettoriale sarà diversa da quella come k-spazio vettoriale se $k' \neq k$. Un esempio: la dimensione di $\mathbb{C}^n$ come spazio vettoriale *reale* è $2n$, perché $\mathbf{e}_1, i\mathbf{e}_1, \mathbf{e}_2, i\mathbf{e}_2, \dots, \mathbf{e}_n, i\mathbf{e}_n$ è una base di $\mathbb{C}^n$ come spazio vettoriale reale.

3. $k[x]_{\leq n}$ ha dimensione $(n+1)$ perché una sua base è $\{1, x, x^2, \ldots, x^n\}$.

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

Proposizione 2.5.16. *Sia V uno spazio vettoriale su k finitamente generato di dimensione n.*

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

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

Dimostrazione. (1): Per il **Corollario 2.5.5** possiamo estendere $v_1, \ldots, v_m$ a una base $\mathcal{B}$ di V. Siccome $\dim V = n$ la base $\mathcal{B}$ contiene n vettori e quindi $m \leq n$. Se $m = n$ allora $\mathcal{B} = \{v_1, \ldots, v_n\}$ e quindi $W = \langle v_1, \ldots, v_n \rangle = V$. (2): Per il **Corollario 2.5.4** possiamo eliminare alcuni dei v_i e ottenere una base $\mathcal{C}$ di V. Siccome $\dim V = n$ segue che $m \geq n$. Se $m = n$ abbiamo che $\mathcal{B} = \{v_1, \ldots, v_n\}$ e quindi $W = V$. □

Esempio 2.5.17. 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

$$v_1 := aw_1 + bw_2, \qquad v_2 := cw_1 + dw_2. \tag{2.5.4}$$

Copiando gli argomenti dell'**Esempio 2.4.3** 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 2.5.16** segue che $\{v_1, v_2\}$ è una base di V se e solo se $(ad - bc) \neq 0$.

Esempio 2.5.18. Consideriamo il sistema di equazioni lineari omogenee (2.3.2). La soluzione *banale* di (2.3.2) è quella con $x_j = 0$ per ogni $1 \leq j \leq n$. 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 $v_1, \ldots, v_n \in k^m$ i vettori definiti da

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

Allora $(x_1, \ldots, x_n)$ è soluzione di (2.3.2) se e solo se

$$x_1 v_1 + x_2 v_2 + \ldots + x_n v = 0. \tag{2.5.5}$$

Siccome $\dim k^m = m$ e per ipotesi $m < n$ la **Proposizione 2.5.16** ci assicura che $v_1, \ldots, v_n$ sono linearmente dipendenti. Quindi esistono $x_1, \ldots, x_n \in k$ non tutti nulli tali che valga (2.5.5) e cioè una soluzione non banale di (2.3.2).

Corollario 2.5.19. *Sia V uno spazio vettoriale su k finitamente generato. Se $W \subset V$ è un sottospazio allora $\dim W \leq \dim V$ e si ha eguaglianza se e solo se $W = V$.*

Dimostrazione. Sia $\{w_1, \ldots, w_m\}$ una base di W: allora $w_1, \ldots, w_m$ sono linearmente indipendenti e quindi il corollario segue dal punto (1) della **Proposizione 2.5.16**. □

2.6 Formula di Grassmann

Consideriamo uno spazio vettoriale V su 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 il **Corollario 2.5.13** 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 2.6.1 (Formula di Grassmann)**.** *Sia V uno spazio vettoriale su k e $U, W \subset V$ sottospazi finitamente generati. Allora*

$$\dim(U + W) + \dim(U \cap W) = \dim U + \dim W. \tag{2.6.1}$$

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, allora $U \cup W \subset \langle z_1, \ldots, z_a, u_1, \ldots, u_m, w_1, \ldots, w_n \rangle$. Dunque $U + W = \langle U \cup W \rangle$ è generato da $\mathcal{B}$.
Rimane da dimostrare che $z_1, \ldots, z_a, u_1, \ldots, u_m, w_1, \ldots, w_n$ sono linearmente indipendenti. Supponiamo che

$$\lambda_1 z_1 + \ldots + \lambda_a z_a + \mu_1 u_1 + \ldots + \mu_m u_m + \theta_1 w_1 + \ldots + \theta_n w_n = 0. \tag{2.6.2}$$

Sia $v := \lambda_1 z_1 + \ldots + \lambda_a z_a + \mu_1 u_1 + \ldots + \mu_m u_m$, da cui

$$\lambda_1 z_1 + \ldots + \lambda_a z_a + \mu_1 u_1 + \ldots + \mu_m u_m = v = -(\theta_1 w_1 + \ldots + \theta_n w_n). \tag{2.6.3}$$

Il membro di sinistra di (2.6.3) è in U e il membro di destra è in W, quindi $v \in U \cap W$. Siccome $(z_1, \ldots, z_a)$ è una base di $U \cap W$, per la **Proposizione 2.5.7**(iii) il vettore v può scriversi in modo unico come combinazione lineare di $z_1, \ldots, z_a$ ossia $v = \tau_1 z_1 + \cdots + \tau_a z_a$, che può anche essere letta come $v = \tau_1 z_1 + \cdots + \tau_a z_a + 0u_1 + \cdots + 0u_m$. D'altra parte, $z_1, \ldots, z_a, u_1, \ldots, u_m$ è una base di U e dunque per la **Proposizione 2.5.7**(iii) $v = \lambda_1 z_1 + \ldots + \lambda_a z_a + \mu_1 u_1 + \ldots + \mu_m u_m$ è l'unico modo di scrivere v come combinazione lineare di $z_1, \ldots, z_a, u_1, \ldots, u_m$. Ne segue che $\mu_1 = \cdots = \mu_m = 0$ e $\lambda_i = \tau_i$ per $1 \leq i \leq a$. Ma allora la relazione lineare (2.6.2) si riduce a

$$\lambda_1 z_1 + \ldots + \lambda_a z_a + \theta_1 w_1 + \ldots + \theta_n w_n = 0$$

e quindi $\lambda_1 = \cdots = \lambda_a = \theta_1 = \cdots = \theta_n = 0$ perché $z_1, \ldots, z_a, w_1, \ldots, w_n$ sono linearmente indipendenti.
Questo dimostra che $\mathcal{B}$ è una base di $(U + W)$ e dunque

$$\dim(U + W) = a + m + n, \quad \dim U \cap W = a, \quad \dim U = a + m, \quad \dim W = a + n$$

e perciò vale (2.6.1). □

Esempio 2.6.2. Siano $a_1, \ldots, a_n \in k$, e sia $W \subset k^n$ il sottospazio delle soluzioni $X = (x_1, \ldots, x_n)$ dell'equazione omogenea

$$a_1x_1 + \ldots + a_nx_n = 0. \tag{2.6.4}$$

Dimostriamo che

$$\dim W = \begin{cases} n & \text{se } 0 = a_1 = \ldots = a_n \\ n-1 & \text{altrimenti.} \end{cases} \tag{2.6.5}$$

Sia $f\colon k^n \to k$ l'applicazione $f(x_1, \ldots, x_n) := a_1x_1 + \ldots + a_nx_n$. Quindi

$$W = \{X \in k^n \mid f(X) = 0\}. \tag{2.6.6}$$

Se $f(X) = 0$ per ogni X, cioè $0 = a_1 = \ldots = a_n$, allora $W = k^n$, e quindi $\dim W = \dim k^n = n$. Ora supponiamo che esista $\overline{X} \in k^n$ tale che $f(\overline{X}) \neq 0$. Sia $U := \langle \overline{X} \rangle$: allora $\dim U = 1$ perché $\overline{X} \neq \mathbf{0}$. Siccome $U \cap W = \{0\}$, la formula di Grassmann (2.6.1) dà che

$$\dim(U + W) = \dim U + \dim W - \dim U \cap W = \dim U + \dim W = 1 + \dim W. \tag{2.6.7}$$

Quindi $\dim W = \dim(U + W) - 1$ e perciò sarà sufficiente dimostrare che

$$(U + W) = k^n. \tag{2.6.8}$$

A questo scopo, notiamo che, se $\lambda \in k$ e $X, Z \in k^n$,

$$f(\lambda X) = \lambda f(X), \qquad f(X + Z) = f(X) + f(Z). \tag{2.6.9}$$

Dato $X \in k^n$, definiamo $Y \in k^n$ così:

$$Y := X - \frac{f(X)}{f(\overline{X})}\overline{X}.$$

Ricordando (2.6.9), troviamo che

$$f(Y) = f\left(X - \frac{f(X)}{f(\overline{X})}\overline{X}\right) = f(X) - f\left(\frac{f(X)}{f(\overline{X})}\overline{X}\right) = f(X) - \frac{f(X)}{f(\overline{X})}f(\overline{X}) = f(X) - f(X) = 0,$$

cioè $Y \in W$. Siccome

$$X := Y + \frac{f(X)}{f(\overline{X})}\overline{X},$$

vediamo che X è somma di un vettore di W e un vettore di U. Questo dimostra (2.6.8).

Definizione 2.6.3. Sia V uno spazio vettoriale finitamente generato e $W \subset V$ un sottospazio. La *codimensione* di W in V è

$$\operatorname{cod}(W, V) := \dim V - \dim W.$$

Nell'**Esempio 2.6.2** la codimensione di W in k^n è 0 se $0 = a_1 = \ldots = a_n$ ed è 1 altrimenti.

2.7 Costruzioni astratte di spazi vettoriali

Presenteremo due costruzioni che producono uno spazio vettoriale a partire da altri spazi vettoriali.

2.7.1 Somma diretta

Siano V e W spazi vettoriali sul campo k. Definiamo la somma di elementi $(v_1, w_1), (v_2, w_2) \in V \times W$ così:

$$(v_1, w_1) + (v_2, w_2) := (v_1 + v_2, w_1 + w_2). \tag{2.7.1}$$

Dati $\lambda \in k$ e $(v, w) \in V \times W$ definiamo

$$\lambda(v, w) := (\lambda v, \lambda w). \tag{2.7.2}$$

Si verifica facilmente che con queste operazioni $V \times W$ è uno spaziovettoriale (su 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 2.7.1. *Siano V e W spazio vettoriali finitamente generati su k. Allora*

$$\dim(V \oplus W) = \dim V + \dim W.$$

Dimostrazione. Siano $\{v_1, \ldots, v_m\}$ e $\{w_1, \ldots, w_n\}$ basi di V e W rispettivamente. Dimostreremo che

$$\{(v_1, 0_W), \ldots, (v_m, 0_W), (0_V, w_1), \ldots, (0_V, w_n)\} \tag{2.7.3}$$

è una base di $V \oplus W$, e la proposizione seguirà. I vettori di (2.7.3) generano $V \oplus W$ perché dato $(v, w) \in V \oplus W$ esistono $\lambda_1, \ldots, \lambda_m \in k$ e $\mu_1, \ldots, \mu_n \in k$ tali che $v = \sum_{i=1}^m \lambda_i v_i$ e $w = \sum_{j=1}^n \mu_j w_j$ (perché $\{v_1, \ldots, v_m\}$ genera V e $\{w_1, \ldots, w_n\}$ genera W), e quindi

$$\sum_{i=1}^m \lambda_i (v_i, 0_W) + \sum_{j=1}^n \mu_j (0_V, w_j) = (\sum_{i=1}^m \lambda_i v_i, \sum_{j=1}^n \mu_j w_j) = (v, w). \tag{2.7.4}$$

Per dimostrare che i vettori di (2.7.3) 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 (2.7.4) vediamo che necessariamente $0 = \lambda_1, \ldots, \lambda_m = \mu_1, \ldots, \mu_n$. □

2.7.2 Quoziente

Per la prossima costruzione assumiamo che V sia uno spazio vettoriale su k e che $W \subset V$ sia un sottospazio vettoriale. Definiamo la relazione $\overset{W}{\sim}$ su V così:

$$v_1 \overset{W}{\sim} v_2 \text{ se e solo se } (v_1 - v_2) \in W. \tag{2.7.5}$$

Si verifica facilmente che $\overset{W}{\sim}$ è una relazione di equivalenza: infatti $\overset{W}{\sim}$ è riflessiva perché $0 \in W$, è simmetrica perché se $w \in W$ allora l'opposto $-w \in W$ ed è transitiva perché W è chiuso per la somma. Chiamiamo $\overset{W}{\sim}$ la *congruenza modulo W*.

Proposizione 2.7.2. *Siano V uno spazio vettoriale su k e $W \subset V$ un sottospazio vettoriale. Supponiamo che $v \overset{W}{\sim} v'$ e $u \overset{W}{\sim} u'$. Allora*

$$(v + u) \overset{W}{\sim} (v' + u'), \qquad \lambda v \overset{W}{\sim} \lambda v'. \tag{2.7.6}$$

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 (2.7.6). La seconda si dimostra in modo analogo. □

La **Proposizione 2.7.4** permette di definire una operazione di somma su $V/\overset{W}{\sim}$. Siano $[v], [u] \in V/\overset{W}{\sim}$: poniamo

$$[v] + [u] := [v + u]. \tag{2.7.7}$$

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 2.7.4**. Analogamente definiamo una moltiplicazione per scalari. Siano $\lambda \in k$ e $[v] \in V/\overset{W}{\sim}$: poniamo

$$\lambda[v] := [\lambda v] \tag{2.7.8}$$

Di nuovo: la definizione è ben posta grazie alla **Proposizione 2.7.4**. Si verifica facilmente che con queste operazioni $V/\overset{W}{\sim}$ è uno spazio vettoriale su k, con elemento neutro dato da $[0]$ (notate che $[0] = W$).

Definizione 2.7.3. Siano V uno spazio vettoriale su k, e $W \subset V$ un sottospazio. Il *quoziente di V modulo W* è lo spazio vettoriale $V/\overset{W}{\sim}$ con le operazioni di somma e moltiplicazione per scalari appena definiti; lo si denota V/W.

Proposizione 2.7.4. *Sia V uno spazio vettoriale finitamente generato su 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 $\lambda_1, \ldots, \lambda_a, \mu_1, \ldots, \mu_b \in k$ 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 $\mu_1, \ldots, \mu_b \in k$ tali che

$$\mu_1[u_1] + \ldots + \mu_b[u_b] = [0].$$

Siccome $[0] = W$ ciò significa che esistono $\lambda_1, \ldots, \lambda_a \in k$ 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*

$$\lambda X + \mu Y + \nu Z = \mathbf{0}.$$

Esercizio 2.2. *Determinate quali dei seguenti sottoinsiemi $W_i \subset k^3$ è un sottospazio.*

1. *$k = \mathbb{R}$ e $W_1 := \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$.*
2. *$k = \mathbb{R}$ e $W_2 := \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z \leq 1\}$.*
3. *$k = \mathbb{C}$ e $W_3 := \{(x, y, z) \in \mathbb{C}^3 \mid x^2 + y^2 + z^2 = 0\}$.*
4. *$k = \mathbb{F}_2$ 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. *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 2.5. *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 2.6.

1. *Dire per quali sottospazi $W \subset \mathbb{R}^n$ il complementare $\mathbb{R}^n \setminus W$ è a sua volta un sottospazio.*
2. *Dire per quali sottospazi $W \subset \mathbb{R}^n$ l'insieme $(\mathbb{R}^n \setminus W) \cup \{0\}$ è a sua volta un sottospazio.*

Esercizio 2.7. *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 2.8. *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 2.9. *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 2.10. *Osserviamo che l'insieme $k[x]_{\leq d}$ (vedi* **Esempio 2.3.15***) i cui elementi sono i polinomi di grado al più d oppure uguali a zero è un sottospazio vettoriale di $k[x]$.*

1. *Dimostrate che $\mathcal{B} := \{1, 1+x, (1+x)^2\}$ è una base di $k[x]_{\leq 2}$.*
2. *Trovate le coordinate di 1 di x e di x^2 bella base $\mathcal{B}$.*

Esercizio 2.11. *Sia $W \subset k[x]_{\leq d}$ definito da*

$$W := \{p \in k[x]_{\leq d} \mid 0 = p(0) = p(-1) = p(1)\}.$$

Dimostrate che W è un sottospazio vettoriale di $k[x]_d$ e calcolatene la dimensione. (Attenzione: il caso in cui $\operatorname{char} k = 2$ è speciale.)

Esercizio 2.12. *Sia V uno spazio vettoriale su 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 2.13. *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 2.14. *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 2.15. *Sia V uno spazio vettoriale finitamente generato e siano $W_1, \ldots, W_p \subset V$ sottospazi. Si dimostri che*

$$\operatorname{cod}(W_1 \cap \ldots \cap W_p, V) \leq \operatorname{cod}(W_1, V) + \ldots + \operatorname{cod}(W_p, V).$$

(Suggerimento: si proceda per induzione su p e si applichi la Formula di Grassmann.)

Esercizio 2.16. *Sia $W \subset k^n$ lo spazio delle soluzioni del sistema lineare omogeneo* (2.3.2)*. Si dimostri che $\dim W \geq (n-m)$. (Invocate l'***Esercizio 2.15** *e l'* **Esempio 2.6.2***.)*

Esercizio 2.17. *Sia k un campo e $\alpha_0, \alpha_1, \ldots, \alpha_n \in k$ distinti. Siano $v_0, v_1, \ldots, v_n \in k^{n+1}$ definiti da*

$$v_i = (\alpha_0^i, \alpha_1^i, \ldots, \alpha_n^i), \qquad 0 \leq i \leq n.$$

Dimostrate che $\{v_0, v_1, \ldots, v_n\}$ è una base di k^{n+1}.

Esercizio 2.18. *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

Geometria affine, I

La nozione di base e relative coordinate di uno spazio vettoriale permette di introdurre coordinate affini per i punti del piano o dello spazio: le coordinate cartesiane introdotte nella scuola sono casi particolari (corrispondono a basi cosiddette ortonormali). Introdurremo le coordinate affini e daremo le equazioni di rette e piani in coordinate affini. Daremo anche la definizione generale di spazio affine. A questo proposito è importante osservare che è conveniente *definire* cosa è uno spazio affine a partire dalla nozione di spazio vettoriale. In particolare da un punto di vista formale la nozione di spazio vettoriale *precede* quella di spazio affine anche se viene motivata con argomenti di geometria affine (classi di equipollenza di segmenti orientati nel piano e nello spazio) che possiamo considerare di natura intuitiva oppure fondati sugli assiomi di Hilbert (rielaborazione degli assiomi di Euclide) per punti, rette, piani dello spazio.

3.1 Coordinate affini nel piano

Sia $\mathcal{A}^2$ il piano della geometria euclidea. Scegliamo un punto $O \in \mathcal{A}^2$ e una base $\{\mathbf{i}, \mathbf{j}\}$ di $\mathcal{V}^2$. Dato $P \in \mathcal{A}^2$ esistono $x, y \in \mathbb{R}$ (e sono unici) tali che

$$\overrightarrow{OP} = x\mathbf{i} + y\mathbf{j}. \tag{3.1.1}$$

Associamo a P la coppia (x, y) e diciamo che x, y sono le *coordinate* di P nel *riferimento affine* determinato dalla scelta di O e della base $\{\mathbf{i}, \mathbf{j}\}$, che indicheremo con $RA(O; \mathbf{i}, \mathbf{j})$. La x è l'*ascissa* di P e la y è l'*ordinata* di P. Spesso scriveremo $P(x_0, y_0)$ per dire "P è il punto con coordinate (x_0, y_0)". Notate che le coordinate del punto O sono $(0, 0)$; il punto O è l'*origine* del sistema $RA(O; \mathbf{i}, \mathbf{j})$. Si ottengono le coordinate cartesiane viste a scuola quando i vettori $\mathbf{i}$, $\mathbf{j}$ sono di uguale lunghezza e perpendicolari (si dice che la base $\{\mathbf{i}, \mathbf{j}\}$ è *ortonormale*). Dati $P_0, P_1 \in \mathcal{A}^2$ di coordinate (x_0, y_0) e (x_1, y_1) rispettivamente le coordinate del vettore $\overrightarrow{P_0P_1}$ nella base $\{\mathbf{i}, \mathbf{j}\}$ si ottengono così:

$$\overrightarrow{P_0P_1} = \overrightarrow{OP_2} - \overrightarrow{OP_1} = x_1\mathbf{i} + y_1\mathbf{j} - (x_0\mathbf{i} + y_0\mathbf{j}) = (x_1 - x_0)\mathbf{i} + (y_1 - y_0)\mathbf{j}. \tag{3.1.2}$$

Ora supponiamo che $P_0 \neq P_1$ e quindi esiste una unica retta $\overline{P_0P_1}$ contenente P_0 e P_1. Un punto $P \in \mathcal{A}^2$ appartiene a $\overline{P_0P_1}$ se e solo se il vettore $\overrightarrow{P_0P}$ è un multiplo del vettore $\overrightarrow{P_0P_1}$ ovvero esiste $t \in \mathbb{R}$ tale che

$$\overrightarrow{P_0P} = t\overrightarrow{P_0P_1}. \tag{3.1.3}$$

Siano (x, y) le coordinate di P: l'equazione (3.1.3) equivale alle due equazioni

$$\begin{aligned} x &= x_0 + t(x_1 - x_0), \\ y &= y_0 + t(y_1 - y_0). \end{aligned}$$

Queste sono equazioni *parametriche* della retta $\overline{P_0P_1}$ - il parametro è t. Spesso si pone

$$l = (x_1 - x_0), \qquad m = (y_1 - y_0)$$

e l, m sono chiamati *coefficienti direttori*. Quindi abbiamo

$$\begin{aligned} x &= x_0 + lt, \\ y &= y_0 + mt. \end{aligned} \tag{3.1.4}$$

Notate che i coefficienti direttori sono definiti a meno di una costante moltiplicativa. Le coordinate (x, y) dei punti di r soddisfano l'equazione

$$ax + by + c = 0, \tag{3.1.5}$$

dove $a = m$, $b = -l$ e $c = -mx_0 + ly_0$. Questa è una equazione *cartesiana* della retta r. Viceversa se $(a, b) \neq (0, 0)$ le soluzioni di (3.1.5) sono le coordinate dei punti di una retta.

3.2 Coordinate affini nello spazio

Sia $\mathcal{A}^3$ lo spazio della geometria euclidea. Scegliamo un punto $O \in \mathcal{A}^3$ e una base $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ di $\mathcal{V}^3$. Dato $P \in \mathcal{A}^3$ esistono $x, y, z \in \mathbb{R}$ (e sono unici) tali che

$$\overrightarrow{OP} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}. \tag{3.2.1}$$

Associamo a P la terna (x, y, z) e diciamo che x, y, z sono le *coordinate* di P nel *riferimento affine* determinato dalla scelta di O e della base $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$, che indicheremo con $RA(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$. Il punto O è l'*origine* del sistema $RA(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$. Dati $P_0, P_1 \in \mathcal{A}^3$ di coordinate (x_0, y_0, z_0) e (x_1, y_1, z_1) rispettivamente le coordinate del vettore $\overrightarrow{P_0P_1}$ nella base $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ si ottengono così:

$$\overrightarrow{P_0P_1} = \overrightarrow{OP_1} - \overrightarrow{OP_0} = x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k} - (x_0\mathbf{i} + y_0\mathbf{j} + z_0\mathbf{k}) = (x_1 - x_0)\mathbf{i} + (y_1 - y_0)\mathbf{j} + (z_1 - z_0)\mathbf{k}. \tag{3.2.2}$$

Ora supponiamo che $P_0 \neq P_1$. Ragionando come per le rette nel piano vediamo che un punto P appartiene alla retta $\overline{P_0P_1}$ se e solo se le sue coordinate (x, y, z) sono date da

$$\begin{aligned} x &= x_0 + t(x_1 - x_0), \\ y &= y_0 + t(y_1 - y_0), \\ z &= z_0 + t(z_1 - z_0) \end{aligned}$$

per un qualche $t \in \mathbb{R}$. Queste sono equazioni *parametriche* della retta $\overline{P_0P_1}$. Spesso si pone $l = (x_1 - x_0)$, $m = (y_1 - y_0)$, $n = (z_1 - z_0)$. Quindi abbiamo

$$\begin{aligned} x &= x_0 + lt, \\ y &= y_0 + mt, \\ z &= z_0 + nt. \end{aligned} \tag{3.2.3}$$

Diciamo che l, m, n sono *coefficienti direttori* di r; sono definiti a meno di una costante moltiplicativa.

Ora siano $P_0, P_1, P_2 \in \mathcal{A}^3$ tre punti non allineati. Esiste un'unico piano Λ contenente P_0, P_1, P_2. Scriviamo equazioni parametriche di Λ. Poniamo

$$\overrightarrow{P_0P_1} = l_1\mathbf{i} + m_1\mathbf{j} + n_1\mathbf{k}, \qquad \overrightarrow{P_0P_2} = l_2\mathbf{i} + m_2\mathbf{j} + n_2\mathbf{k}.$$

Allora $P_0, P_1, P_2 \in \mathcal{A}^3$ non sono allineati se e solo se $\overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}$ sono linearmente indipendenti.

Osservazione 3.2.1. Sia V uno spazio vettoriale con base $\{v_1, v_2, v_3\}$. Siano

$$u = l_1v_1 + m_1v_2 + n_1v_3, \qquad w = l_2v_1 + m_2v_2 + n_2v_3.$$

Allora u, w sono linearmente dipendenti se e solo se

$$(m_1n_2 - n_1m_2, -(l_1n_2 - n_1l_2), l_1m_2 - m_1l_2) = (0, 0, 0). \tag{3.2.4}$$

Infatti u, w sono linearmente dipendenti se e solo se esitono $x, y \in k$, non entrambi nulli, tali che

$$0 = xu + yw = (xl_1 + yl_2)v_1 + (xm_1 + ym_2)v_2 + (xn_1 + yn_2)v_3. \tag{3.2.5}$$

Siccome $\{v_1, v_2, v_3\}$ è una base di V, vale (3.2.5) se e solo se

$$0 = (xl_1 + yl_2) = (xm_1 + ym_2) = (xn_1 + yn_2). \tag{3.2.6}$$

Ora supponiamo che u, w siano linearmente dipendenti. Allora $(l_1v_1 + m_1v_2), (l_2v_1 + m_2v_2)$ sono linearmente dipendenti, e quindi $(l_1m_2 - m_1l_2) = 0$ per l'**Esempio 2.5.17**, inoltre $(l_1v_1 + n_1v_3), (l_2v_1 + n_2v_3)$ sono linearmente dipendenti, e quindi $(l_1n_2 - n_1l_2) = 0$, e, analogamente otteniamo che $(m_1n_2 - n_1m_2) = 0$. Questo dimostra che se u, w sono linearmente dipendenti, allora vale (3.2.4). Ora supponiamo che valga (3.2.4), e dimostriamo che u, w sono linearmente dipendenti. Se $n_1 = n_2 = 0$, allora $u = l_1v_1 + m_1v_2$ e $w = l_2v_1 + m_2v_2$, quindi in questo caso u, w sono linearmente dipendenti per l'**Esempio 2.5.17**. Se invece n_1, n_2 non sono entrambi nulli, notiamo che, per (3.2.4), si ha

$$n_2u - n_1w = 0,$$

e quindi u, w sono linearmente dipendenti.

Siano (x_0, y_0, z_0) le coordinate di P_0: allora $P \in \Lambda$ se e solo se $\overrightarrow{P_0P} \in \langle \overrightarrow{P_0P_1}, \overrightarrow{P_0P_2} \rangle$ ovvero $\overrightarrow{P_0P} = t\overrightarrow{P_0P_1} + u\overrightarrow{P_0P_2}$ per qualche $t, u \in \mathbb{R}$. In altre parole equazioni parametriche di Λ sono

$$\begin{aligned} x &= x_0 + l_1t + l_2u, \\ y &= y_0 + m_1t + m_2u, \\ z &= z_0 + n_1t + n_2u. \end{aligned} \tag{3.2.7}$$

Ora siano (a, b, c) le entrate del vettore di $\mathbb{R}^3$ a destra di (3.2.4). Un semplice calcolo dà che

$$0 = al_1 + bm_1 + cn_1 = al_2 + bm_2 + cn_2$$

e quindi se $P \in \Lambda$ le sue coordinate (x, y, z) soddsfano l'equazione *cartesiana*

$$ax + by + cz + d = 0. \tag{3.2.8}$$

dove $d = -ax_0 - by_0 - cz_0$. Viceversa se le coordinate (x, y, z) di P soddisfano l'equazione (3.2.8) allora $P \in \Lambda$. D'altra parte se scegliamo $a, b, c, d \in \mathbb{R}$ con $(a, b, c) \neq (0, 0, 0)$ allora le soluzioni di (3.2.8) sono le coordinate dei punti appartenenti a un piano.

3.3 Giacitura e parallelismo

Siano $r \subset \mathcal{A}^3$ e $\Lambda \subset \mathcal{A}^3$ una retta e un piano rispettivamente. Definiamo

$$\mathcal{G}(r) := \{\overrightarrow{PQ} \mid P, Q \in r\}, \tag{3.3.1}$$
$$\mathcal{G}(\Lambda) := \{\overrightarrow{PQ} \mid P, Q \in \Lambda\}. \tag{3.3.2}$$

Si verifica facilmente che $\mathcal{G}(r)$ e $\mathcal{G}(\Lambda)$ sono sottospazi vettoriali di $\mathcal{V}^3$, di dimensioni 1 e 2 rispettivamente: si chiamano le *giaciture* di r e Λ rispettivamente. Si definisce in modo analogo la giacitura di una retta in $\mathcal{A}^2$.

Esempio 3.3.1.

1. Sia $r \subset \mathcal{A}^2$ la retta di equazione (3.1.4). La giacitura $\mathcal{G}(r)$ è il sottospazio di $\mathcal{V}^2$ generato dal vettore $(l\mathbf{i} + m\mathbf{j})$.

2. Sia $r \subset \mathcal{A}^2$ la retta di equazione (3.1.5). La giacitura $\mathcal{G}(r)$ è il sottospazio di $\mathcal{V}^2$ dato da

$$\mathcal{G}(r) = \{x\mathbf{i} + y\mathbf{j} \mid ax + by = 0\}. \tag{3.3.3}$$

   In altre parole le coordinate dei vettori di $\mathcal{G}(r)$ sono le soluzione dell'equazione lineare omogenea associata all'equazione cartesiana di r.

3. La giacitura della retta r di equazioni parametriche (3.2.3) è il sottospazio di $\mathcal{V}^3$ generato dal vettore $(l\mathbf{i} + m\mathbf{j} + n\mathbf{k})$.

4. La giacitura del piano Λ di equazioni parametriche (3.2.7) è il sottospazio di $\mathcal{V}^3$ generato dai vettori $(l_1\mathbf{i} + m_1\mathbf{j} + n_1\mathbf{k})$ e $(l_2\mathbf{i} + m_2\mathbf{j} + n_2\mathbf{k})$.

5. Sia $\Lambda \subset \mathcal{A}^3$ il piano di equazione (3.2.8). La giacitura $\mathcal{G}(\Lambda)$ è il sottospazio di $\mathcal{V}^3$ dato da

$$\mathcal{G}(\Lambda) = \{x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \mid ax + by + cz = 0\}. \tag{3.3.4}$$

   In altre parole le coordinate dei vettori di $\mathcal{G}(\Lambda)$ sono le soluzione dell'equazione lineare omogenea associata all'equazione cartesiana di Λ.

La seguente definizione è in accordo con la nozione intuitiva di parallelismo (attenzione: è conveniente stabilire che una retta è parallela a se stessa e parallela a ogni piano che la contiene).

Definizione 3.3.2. Siano $r, r' \subset \mathcal{A}^3$ e $\Lambda, \Lambda' \subset \mathcal{A}^3$ due rette e due piani, rispettivamente. Allora

1. r e r' sono *parallele* (in simboli $r||r'$) se $\mathcal{G}(r) = \mathcal{G}(r')$,

2. r è *parallela* a Λ (in simboli $r||\Lambda$) se $\mathcal{G}(r) \subset \mathcal{G}(\Lambda)$,

3. Λ, Λ' sono *parallele* se $\mathcal{G}(\Lambda) = \mathcal{G}(\Lambda')$.

Esempio 3.3.3. La retta r di equazioni parametriche (3.2.3) è parallela al piano Λ di equazione cartesiana (3.2.8) se e solo se

$$al + bm + cn = 0. \tag{3.3.5}$$

Osservazione 3.3.4. Siano $r, r' \subset \mathcal{A}^2$ rette *non* parallele: allora r, r' si intersecano in un punto. È sufficiente dimostrare che $r \cap r' \neq \emptyset$ (se avessero almeno *due* punti in comune avrebbero la stessa giacitura). Infatti siano $0 \neq v \in \mathcal{G}(r)$ e $0 \neq v' \in \mathcal{G}(r')$. Per l'ipotesi che r, r' non sono parallele i vettori v, v' sono linearmente indipendenti e quindi $\{v, v'\}$ è una base di $\mathcal{V}^2$ giacchè $\dim \mathcal{V}^2 = 2$. Siano $P \in r$ e $P' \in r'$: esistono $\lambda, \lambda' \in \mathbb{R}$ tali che

$$\overrightarrow{PP'} = \lambda v + \lambda' v'. \tag{3.3.6}$$

Siccome $\mathcal{G}(r) = \langle v \rangle$ esiste $Q \in r$ tale che $\overrightarrow{PQ} = \lambda v$. Allora $Q \in r'$ (e quindi $Q \in r \cap r'$). Infatti $\overrightarrow{PQ} + \overrightarrow{QP'} = \overrightarrow{PP'}$ e l'equazione (3.3.6) danno che $\overrightarrow{QP'} = \lambda' v'$ e quindi $\overrightarrow{QP'} \in \mathcal{G}(r')$: siccome $P' \in r'$ segue che $Q \in r'$. Analogamente si dimostra che se $r \subset \mathcal{A}^3$ è una retta *non* parallela al piano $\Lambda \subset \mathcal{A}^3$ allora r, Λ si intersecano in un punto. Non vale nulla di analogo per rette *non* parallele in $\mathcal{A}^3$: in generale non avranno punti in comune (e si diranno *sghembe*).

3.4 Spazi affini

Abbiamo definito lo spazio vettoriale dei vettori geometrici a partire dal piano euclideo e dallo spazio euclideo. In realtà è più conveniente da un punto di vista logico iniziare con la nozione di spazio vettoriale, a partire da questa si definisce cosa è uno spazio affine su un dato spazio vettoriale. Quindi si può definire il piano euclideo come uno spazio affine su $\mathbb{R}^2$, e lo spazio euclideo come uno spazio affine su $\mathbb{R}^3$.

Definizione 3.4.1. Sia V uno spazio vettoriale (su un campo k). Uno spazio affine con spazio vettoriale associato V è un insieme non vuoto $\mathbb{A}$ provvisto di una "azione"(di traslazione)

$$\begin{array}{ccc} \mathbb{A} \times V & \longrightarrow & \mathbb{A} \\ (P, v) & \mapsto & P + v \end{array} \tag{3.4.1}$$

che gode delle seguenti proprietà:

- $P + 0 = P$ per ogni $P \in \mathbb{A}$.
- $P + (v + w) = (P + v) + w$ per ogni $P \in \mathbb{A}$ e $v, w \in V$.
- dati $P, Q \in \mathbb{A}$ esiste un unico $v \in V$ tale che $P + v = Q$.

I punti di uno spazio affine sono i suoi elementi. Uno *spazio affine su* k è uno spazio affine con spazio vettoriale associato che ha k come campo degli scalari.

Esempio 3.4.2. Il piano euclideo $\mathcal{A}^2$ con spazio vettoriale associato $\mathcal{V}^2$. Dato $P \in \mathcal{A}^2$ e $v \in \mathcal{V}^2$ definiamo $P + v$ come l'unico $Q \in \mathcal{A}^2$ tale che il segmento orientato $\overrightarrow{PQ}$ rappresenti il vettore v.

Esempio 3.4.3. Sia k un campo e $W \subset k^n$ l'insieme delle soluzioni dell'equazione lineare

$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = b, \tag{3.4.2}$$

dove $a_i, b \in k$. Sia $V \subset k^n$ il sottospazio vettoriale delle soluzioni dell'equazione lineare omogenea ottenuto da (3.4.2) sostituendo b con 0:

$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = 0. \tag{3.4.3}$$

Notate che W è un sottospazio vettoriale di k^n se e solo se $b = 0$ cioè se è uguale a V (se $b \neq 0$ allora $\mathbf{0} \notin W$ e quindi W non è un sottospazio vettoriale di k^n). Definiamo un'azione di V su W così

$$\begin{array}{ccc} W \times V & \longrightarrow & W \\ (X, Y) & \mapsto & X + Y \end{array} \tag{3.4.4}$$

Si verifica facilmente che le proprietà (a), (b) e (c) della **Definizione 3.4.1** sono soddisfatte.

Esempio 3.4.4. Uno spazio vettoriale V è uno spazio affine su se stesso: l'azione è data dalla somma di vettori

$$\begin{array}{ccc} V \times V & \longrightarrow & V \\ (u, v) & \mapsto & u + v \end{array} \tag{3.4.5}$$

Lo *spazio affine n-dimensionale standard su k* è

$$\mathbb{A}^n_k := k^n \tag{3.4.6}$$

con la struttura di spazio affine appena definita.

Sia $\mathbb{A}$ uno spazio affine sullo spazio vettoriale V. Dato $v \in V$ definiamo

$$\begin{array}{ccc} \mathbb{A} & \xrightarrow{T_v} & \mathbb{A} \\ P & \mapsto & P + v \end{array} \tag{3.4.7}$$

Osservazione 3.4.5. Le proprietà (a), (b), (c) della **Definizione 3.4.1** equivalgono ripettivamente a

(a') $T_0 = Id_{\mathbb{A}}$,

(b') $T_v \circ T_w = T_{v+w}$ per ogni $v, w \in V$,

(c') dati $P, Q \in \mathbb{A}$ esiste un unico $v \in V$ tale che $T_v(P) = Q$.

Lemma 3.4.6. *Sia $\mathbb{A}$ uno spazio affine sullo spazio vettoriale V. Sia $v \in V$.*

(1) L'applicazione T_v è biunivoca.

(2) Se esiste $P \in \mathbb{A}$ tale che $T_v(P) = P$ allora $v = 0$. Equivalentemente: se $v \neq 0$ l'applicazione T_v non ha punti fissi.

Dimostrazione. Per la proprietà (b') si ha $T_{-v} \circ T_v = T_v \circ T_{-v} = T_0$ e per la proprietà (a') concludiamo che T_v è biunivoca. Questo dimostra (1). Per dimostrare (2) supponiamo che $T_v(P) = P$. Per (a') abbiamo $T_0(P) = P$ e per (c') concludiamo che $v = 0$. □

Definizione 3.4.7. Sia $\mathbb{A}$ uno spazio affine sullo spazio vettoriale V. Dati $P, Q \in \mathbb{A}$ il vettore $\overrightarrow{PQ} \in V$ è l'unico vettore tale che $P + \overrightarrow{PQ} = Q$.

Osserviamo che, fissato un punto P in $\mathbb{A}$, l'applicazione

$$c_P : \begin{array}{ccc} \mathbb{A} & \longrightarrow & V \\ Q & \mapsto & \overrightarrow{PQ} \end{array} \tag{3.4.8}$$

è biunivoca, con inversa

$$c_P^{-1} : \begin{array}{ccc} V & \longrightarrow & \mathbb{A} \\ v & \mapsto & P + v \end{array} \tag{3.4.9}$$

Dunque i punti dello spazio affine $\mathbb{A}$ possono essere messi in corrispondenza biunivoca con i vettori dello spazio vettoriale V ma tale corrispondenza *non è "canonica"*, ossia dipende dalla scelta in un punto P in $\mathbb{A}$.

Osservazione 3.4.8. Consideriamo l' **Esempio 3.4.3**: dati $X = (x_i)$ e $Z = (z_i)$ in W abbiamo

$$\overrightarrow{XZ} = (z_1 - x_1, z_2 - x_2, \ldots, z_n - x_n) = Z - X. \tag{3.4.10}$$

Lemma 3.4.9. *Sia $\mathbb{A}$ uno spazio affine e $P, Q, R \in \mathbb{A}$. Allora*

$$\overrightarrow{PP} = 0, \tag{3.4.11}$$
$$\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}, \tag{3.4.12}$$
$$\overrightarrow{PQ} = -\overrightarrow{QP}. \tag{3.4.13}$$

Dimostrazione. Siccome $P + 0 = P$ e $P + \overrightarrow{PP} = P$, si ha $0 = \overrightarrow{PP}$. Abbiamo

$$P + (\overrightarrow{PQ} + \overrightarrow{QR}) = (P + \overrightarrow{PQ}) + \overrightarrow{QR} = Q + \overrightarrow{QR} = R. \tag{3.4.14}$$

D'altra parte per la proprietà (c) della **Definizione 3.4.1** esiste un unico vettore v tale che $P + v = R$ e per definizione è $\overrightarrow{PR}$; segue (3.4.12). Ora dimostriamo (3.4.13). Per (3.4.11) e (3.4.12),

$$0 = \overrightarrow{PP} = \overrightarrow{PQ} + \overrightarrow{QP}, \tag{3.4.15}$$

e quindi $\overrightarrow{PQ} = -\overrightarrow{QP}$. □

Definizione 3.4.10. La *dimensione* di uno spazio affine $\mathbb{A}$ sullo spazio vettoriale V è definita come

$$\dim \mathbb{A} := \dim V. \tag{3.4.16}$$

La definizione di dimensione di uno spazio affine è sensata: basti pensare al caso di un piano o dello spazio ordinario. Una *retta* è uno spazio affine di dimensione 1, un *piano* è uno spazio affine di dimensione 2.

3.5 Combinazioni lineari di punti

Sia $\mathbb{A}$ uno spazio affine. Non esiste un modo sensato di definire la combinazione lineare $\lambda P + \mu Q$ di punti $P, Q \in \mathbb{A}$ se $\lambda, \mu \in k$ sono arbitrari: pensate all' **Esempio 3.4.3** 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 3.5.1. *Sia $\mathbb{A}$ uno spazio affine su k e $P_0, \ldots, P_d, Q, R \in \mathbb{A}$. Siano $\lambda_0, \ldots, \lambda_d \in k$ tali che*

$$\sum_{i=0}^{d} \lambda_i = 1. \tag{3.5.1}$$

Allora

$$Q + \sum_{i=0}^{d} \lambda_i \overrightarrow{QP_i} = R + \sum_{i=0}^{d} \lambda_i \overrightarrow{RP_i}. \tag{3.5.2}$$

Dimostrazione. Sottraendo il vettore $\sum_{i=0}^{d} \lambda_i \overrightarrow{RP_i}$ ad ambo i membri di (3.5.2) vediamo che è sufficiente verificare che

$$Q + \sum_{i=1}^{d} \lambda_i (\overrightarrow{QP_i} - \overrightarrow{RP_i}) = R. \tag{3.5.3}$$

Applicando il **Lemma 3.4.9** vediamo che (3.5.3) equivale a

$$Q + \sum_{i=0}^{d} \lambda_i \overrightarrow{QR} = R. \tag{3.5.4}$$

L'equazione (3.5.4) vale perché per ipotesi vale (3.5.1). □

Il **Lemma 3.5.1** ci permette di dare la seguente definizione.

Definizione 3.5.2. Sia $\mathbb{A}$ uno spazio affine su k e $P_0, \ldots, P_d \in \mathbb{A}$. Siano $\lambda_0, \ldots, \lambda_d \in k$ tali che valga (3.5.1). La combinazione lineare di $P_0, \ldots, P_d \in \mathbb{A}$ con pesi $\lambda_0, \ldots, \lambda_d$ è

$$\sum_{i=0}^{d} \lambda_i P_i := Q + \sum_{i=0}^{d} \lambda_i \overrightarrow{QP_i} \tag{3.5.5}$$

dove $Q \in \mathbb{A}$ è arbitrario. (La definizione è sensata grazie al **Lemma 3.5.1**.)

Esempio 3.5.3. Sia $W \subset k^n$ lo spazio affine dell' **Esempio 3.4.3**. Siano $X, Y \in W$. Dati $\lambda, \mu \in k$ tali che $\lambda + \mu = 1$, la combinazione lineare di X, Y con pesi λ, μ è uguale alla combinazione lineare di vettori

$$\lambda X + \mu Y. \tag{3.5.6}$$

Notate che (3.5.6) ha senso anche se $\lambda + \mu \neq 1$, ma non apparterrà a W se $b \neq 0$,.

Esempio 3.5.4. Consideriamo l' **Esempio 3.4.2**. Siano $P, Q \in \mathbb{A}$. 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.

3.6 Sottospazi affini

Definizione 3.6.1. Sia $\mathbb{A}$ uno spazio affine su uno spazio vettoriale V. Un sottoinsieme non vuoto $\mathbb{B} \subset \mathbb{A}$ è un sottospazio affine se esistono $P \in \mathbb{A}$ e un sottospazio vettoriale $W \subset V$ tali che

$$\mathbb{B} = P + W := \{p + w \mid w \in W\}.$$

Osservazione 3.6.2. Siano $\mathbb{A}$ uno spazio affine su uno spazio vettoriale V, e $\mathbb{B} \subset \mathbb{A}$ un sottospazio affine. Supponiamo di avere $\mathbb{B} = P + W = Q + U$ dove $P, Q \in \mathbb{A}$ e $W, U \subset V$ sono sottospazi vettoriali. Allora $U = W$; infatti, siccome $Q \in (P + W)$, esiste $w_0 \in W$ tale che $Q = P + w_0$, e quindi $P + W = Q + U$ si traduce nell'uguaglianza $W = w_0 + U$ (dove $w_0 + U := \{w_0 + u \mid u \in U\}$), da cui segue $W = U$. Quindi c'è un unico sottospazio vettoriale di V associato a $\mathbb{B}$: è la *giacitura* di $\mathbb{B}$, denotata $\mathcal{G}(\mathbb{B})$. Notiamo anche che $\mathbb{B}$ è in modo naturale uno spazio affine con spazio vettoriale associato $\mathcal{G}(\mathbb{B})$.

È naturale estendere a spazii affini qualsiasi la definizione di parallelismo tra sottospazi affine di $\mathcal{A}^2$ o $\mathcal{A}^3$, nel seguente modo.

Definizione 3.6.3. Sia $\mathbb{A}$ uno spazio affine su uno spazio vettoriale V. Due sottospazi affini $\mathbb{B}_1, \mathbb{B}_2 \subset \mathbb{A}$ sono *paralleli* se $\mathcal{G}(\mathbb{B}_1) \subset \mathcal{G}(\mathbb{B}_2)$, oppure $\mathcal{G}(\mathbb{B}_1) \supset \mathcal{G}(\mathbb{B}_2)$.

Proposizione 3.6.4. *Sia $\mathbb{A}$ uno spazio affine e $\mathbb{B}_i$ per $i \in I$ una collezione di sottospazi affini di $\mathbb{A}$. Se l'intersezione $\bigcap_{i \in I} \mathbb{B}_i$ è non vuota allora è un sottospazio affine di $\mathbb{A}$, con giacitura l'intersezione delle giaciture $\mathcal{G}(\mathbb{B}_i)$ per $i \in I$.*

Dimostrazione. Sia V lo spazio vettoriale associato ad $\mathbb{A}$. Sia $P \in \bigcap_{i\in I} \mathbb{B}_i$ e sia $W := \bigcap_{i\in I} \mathcal{G}(\mathbb{B}_i)$. Quindi W è un sottospazio vettoriale di V. Abbiamo che $(P + W) = \bigcap_{i\in I} \mathbb{B}_i$ e quindi $\bigcap_{i\in I} \mathbb{B}_i$ è un sottospazio affine di $\mathbb{A}$. □

Esempio 3.6.5. Siano $b_1, \ldots, b_m \in k$, e $a_{ij} \in k$ per $1 \leq i \leq m$ e $1 \leq j \leq n$. Il sottoinsieme delle soluzioni del sistema di equazioni

$$\begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots & = & * \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n & = & b_i \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots & = & * \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m. \end{array} \tag{3.6.1}$$

o è vuoto, oppure è un sottospazio affine di k^n, in quanto intersezione di sottospazio affini, e la sua giacitura è il sottospazio di k^n delle soluzioni del sistema lineare omogeneo associato a (3.6.1), cioè ottenuto sostituendo 0 a ciascun b_i.

Dato un sottoinsieme $Z \subset \mathbb{A}$ esiste un minimo sottospazio affine $\langle Z\rangle \subset \mathbb{A}$ contenente Z per la **Proposizione 3.6.4**.

Definizione 3.6.6. Siano $\mathbb{A}$ uno spazio affine e $Z \subset \mathbb{A}$. Il sottospazio affine di $\mathbb{A}$ generato da Z è l'intersezione di tutti i sotospazi affini di $\mathbb{A}$ contenenti Z - lo denoteremo con $\langle Z\rangle$.

Quindi $\langle Z\rangle$ è un sottospazio affine contenente Z e contenuto in ogni sottospazio affine che contiene Z. Se Z è finito $Z = \{P_0, \ldots, P_d\}$ poniamo

$$\langle P_0, \ldots, P_d\rangle := \langle\{P_0, \ldots, P_d\}\rangle. \tag{3.6.2}$$

Esplicitamente

$$\langle P_0, \ldots, P_d\rangle = P_0 + \langle\overrightarrow{P_0P_1}, \ldots, \overrightarrow{P_0P_d}\rangle. \tag{3.6.3}$$

Infatti il membro di destra di (3.6.3) è un sottospazio affine di $\mathbb{A}$ contenente $P_0, \ldots, P_d$ e quindi è sufficiente dimostrare che ogni sottospazio affine $\mathbb{B} \subset \mathbb{A}$ contenente $P_0, \ldots, P_d$ contiene il membro di destra di (3.6.3); questo è chiaro. Da (3.6.3) segue che abbiamo anche

$$\langle P_0, \ldots, P_d\rangle = \left\{ \sum_{i=0}^{d} \lambda_i P_i \;\middle|\; \lambda_i \in k, \quad \sum_{i=0}^{d} \lambda_i = 1 \right\}. \tag{3.6.4}$$

Infatti

$$\langle P_0, \ldots, P_d\rangle = \left\{ P_0 + \sum_{i=1}^{d} \lambda_i \overrightarrow{P_0P_i} \;\middle|\; \lambda_1, \ldots, \lambda_d \in k \right\} = \tag{3.6.5}$$

$$= \left\{ (1 - \sum_{i=1}^{d} \lambda_i)P_0 + \sum_{i=1}^{d} \lambda_i P_i \;\middle|\; \lambda_1, \ldots, \lambda_d \in k \right\} \tag{3.6.6}$$

e la conclusione segue osservando che $\lambda_0 = 1 - \sum_{i=1}^{d} \lambda_i$.

Sia $\mathbb{B} \subset \mathbb{A}$ un sottospazio affine. Siccome $\mathbb{B}$ è uno spazio affine è ben definita la sua dimensione $\dim \mathbb{B}$; questo fatto ci permette di dare la nozione di dipendenza/indipendenza lineare di punti $P_0, \ldots, P_d \in \mathbb{A}$. Osserviamo che per (3.6.3) si ha che

$$\dim\langle P_0, \ldots, P_d\rangle \leq d. \tag{3.6.7}$$

Definizione 3.6.7. Sia $\mathbb{A}$ uno spazio affine. Una sequenza di punti $P_0, \dots, P_d \in \mathbb{A}$ 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 $P_0, \dots, P_d \in \mathbb{A}$ 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 3.6.8. *Sia $\mathbb{A}$ uno spazio affine. Una sequenza di punti $P_0, \dots, P_d \in \mathbb{A}$ è linearmente indipendente se e solo se è iniettiva l'applicazione*

$$\begin{array}{ccc} \left\{(\lambda_0, \dots, \lambda_d) \in k^{d+1} \;\middle|\; \sum_{i=0}^d \lambda_i = 1\right\} & \longrightarrow & \langle P_0, \dots, P_d\rangle \\ (\lambda_0, \dots, \lambda_d) & \mapsto & \sum_{i=0}^d \lambda_i P_i \end{array} \tag{3.6.8}$$

Esempio 3.6.9. Consideriamo l' **Esempio 3.4.4**. Siano $v_0, \dots, v_d \in V = \mathbb{A}$. 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. Se $v \neq 0$ allora i punti $0, v$ sono linearmente indipendenti ma ovviamente i vettori $0, v$ non lo sono.

Esercizi del Capitolo 3

Esercizio 3.1. *Nel $RA(O; \mathbf{i}, \mathbf{j})$ siano $P_0(1, 2)$ e $P_1(-1, 1)$. Scrivere equazioni parametriche e cartesiane della retta $r := \overline{P_0P_1}$.*

Esercizio 3.2. *Nel $RA(O; \mathbf{i}, \mathbf{j})$ siano r, r' le rette di equazioni parametriche*

$$\begin{array}{rcl} x & = & 1 + 3t, \\ y & = & -2 + t. \end{array}$$

e

$$\begin{array}{rcl} x & = & s, \\ y & = & 1 - s. \end{array}$$

rispettivamente. Determinate le coordinate del punto d'intersezione tra r e r'.

Esercizio 3.3. *Sia $\{\mathbf{i}, \mathbf{j}\}$ una base di $\mathcal{V}^2$ e $\mathbf{k} := \mathbf{i} + 2\mathbf{j}$, $\mathbf{h} := \mathbf{i} + \mathbf{j}$. Sia $Q \in \mathcal{A}^2$ il punto di coordinate $(1, -1)$ nel $RA(O; \mathbf{i}, \mathbf{j})$.*

(1) Verificate che $\{\mathbf{k}, \mathbf{h}\}$ è una base di $\mathcal{V}^2$.

(2) Determinate le coordinate di O nel $RA(Q; \mathbf{k}, \mathbf{h})$.

Esercizio 3.4. *Nel $RA(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ siano $P_0(1, 1, -1)$, $P_1(3, 0, 2)$ e $P_2(4, 2, 3)$.*

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 3.5. *Nel $RA(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ siano Λ_1 e Λ_2 i piani di equazioni cartesiane*

$$\Lambda_1 : \; 3x + 2y + z = 1, \qquad \Lambda_2 : \; x - y - 2z = 2.$$

Verificate che l'intersezione $\Lambda_1 \cap \Lambda_2$ è una retta r e determinate equazioni parametriche di r.

Esercizio 3.6. *Nel $RA(O;\mathbf{i},\mathbf{j},\mathbf{k})$ siano r, r' le rette di equazioni parametriche*

$$\begin{aligned} x &= 1+2t, \\ y &= -t, \\ z &= 2+5t. \end{aligned}$$

e

$$\begin{aligned} x &= s, \\ y &= 1+2s, \\ z &= 3 \end{aligned}$$

rispettivamente. Determinate una equazione cartesiana del piano Λ contenente r e parallelo a r'.

Esercizio 3.7. *Nel $RA(O;\mathbf{i},\mathbf{j},\mathbf{k})$ siano Λ, Λ' i piani di equazioni cartesiane*

$$x+2y-z+1=0$$

e

$$2x+z-3=0.$$

rispettivamente. Determinate equazioni parametriche della retta r parallela a Λ e Λ' e passante per il punto $P(1,1,1)$.

Esercizio 3.8. *Nel $RA(O;\mathbf{i},\mathbf{j})$ siano $P_0(1,1)$, $P_1(2,3)$, $P_2(3,5)$. Gli studenti Anna, Marco e Lucio misurano le coordinate di P_0, P_1 e P_2 in un nuovo sistema di riferimento $RA(Q;\mathbf{k},\mathbf{h})$ e le loro misurazioni sono discordanti:*

(Anna) *$P_0(0,1)$, $P_1(-1,0)$ e $P_2(-2,-1)$.*

(Marco) *$P_0(0,2)$, $P_1(1,2)$ e $P_2(0,3)$.*

(Lucio) *$P_0(0,0)$, $P_1(1,1)$ e $P_2(3,3)$.*

Due tra Anna, Marco e Lucio sicuramente ha sbagliato misurazioni: determinate chi.

Esercizio 3.9. *Sia $\mathbb{A}$ uno spazio affine. Siano $P_0,\dots,P_d \in \mathbb{A}$ 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 $B(P_0,\dots,P_{d-1})$.

Capitolo 4

Applicazioni lineari e matrici

Siano V, W spazi vettoriali su uno stesso campo k: un'applicazione $f\colon V \to W$ è lineare se "commuta" con le operazioni di somma e di moltiplicazione per scalari. Gli oggetti fondamentali dell'algebra lineare sono gli spazi vettoriali e le applicazioni lineari tra di essi - sono l'esempio più semplice di morfismi di strutture algebriche. Sono nozioni fondamentali perché in generale si cerca di ridurre qualsiasi problema matematico a un problema riguardante un'applicazione lineare tra spazi vettoriali: un esempio è il differenziale di una funzione in un punto, è un'applicazione lineare che ci dice molto sul comportamento locale della funzione.

4.1 Applicazioni lineari: definizione e prime proprietà

4.1.1 La definizione

Definizione 4.1.1. Siano V, W spazi vettoriali su un campo k. Un'applicazione $f\colon V \to W$ è *lineare* se dati $v_1, v_2 \in V$ e $\lambda_1, \lambda_2 \in k$ vale

$$f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2). \tag{4.1.1}$$

Osservazione 4.1.2. Siano V, W spazi vettoriali su un campo k. Un'applicazione $f\colon V \to W$ è lineare se e solo se:

(1) $f(v_1 + v_2) = f(v_1) + f(v_2)$ per $v_1, v_2 \in V$, e

(2) $f(\lambda v) = \lambda f(v)$ per $\lambda \in k$ e $v \in V$.

Infatti supponiamo che f sia lineare. Ponendo $\lambda_1 = \lambda_2 = 1$ nella (4.1.2) otteniamo che vale (1) e ponendo $\lambda_1 = 1$, $\lambda_2 = 0$ otteniamo che vale (2). Viceversa supponiamo che valgano (1) e (2). Dati $v_1, v_2 \in V$ e $\lambda_1, \lambda_2 \in k$ abbiamo che

$$f(\lambda_1 v_1 + \lambda_2 v_2) = f(\lambda_1 v_1) + f(\lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2). \tag{4.1.2}$$

Osservazione 4.1.3. Supponiamo che $f\colon V \to W$ sia lineare. Dalla (4.1.2) segue che:

(1) $f(0) = 0$, basta porre $0 = \lambda_1 = \lambda_2$.

(2) Se $v \in V$ allora $f(-v) = -f(v)$, basta porre $v_1 = v$ $v_2 = 0$, $\lambda_1 = -1$.

(3) Se $v_1, v_2, \ldots, v_n \in V$ e $\lambda_1, \lambda_2, \ldots, \lambda_n \in k$ vale

$$f(\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n) = \lambda_1 f(v_1) + \lambda_2 f(v_2) + \ldots + \lambda_n f(v_n). \tag{4.1.3}$$

(Applicate (4.1.2) $(n-1)$ volte.)

Esempio 4.1.4. Sia

$$\begin{array}{ccc} k^n & \xrightarrow{f} & k \\ (x_1,\dots,x_n) & \mapsto & a_1x_1+a_2x_2+\dots a_nx_n. \end{array} \tag{4.1.4}$$

La f è lineare. Infatti siano $X,Y\in k^n$ e $\lambda,\mu\in k$: si ha che

$$f(\lambda X+\mu Y)=\sum_{i=1}^n a_i(\lambda x_i+\mu y_i)=\sum_{i=1}^n a_i\lambda x_i+\sum_{i=1}^n a_i\mu y_i=\lambda\sum_{i=1}^n a_ix_i+\mu\sum_{i=1}^n a_iy_i=\lambda f(X)+\mu f(Y).$$

Viceversa supponiamo che $f\colon k^n\to k$ sia lineare. Sia $a_i:=f(\mathbf{e}_i)$: dimostriamo che f è data da (4.1.4). Infatti per linearità (vedi (4.1.3)) abbiamo che

$$f(x_1\mathbf{e}_1+x_2\mathbf{e}_2+\dots+x_n\mathbf{e}_n)=x_1f(\mathbf{e}_1)+x_2f(\mathbf{e}_2)+\dots+x_nf(\mathbf{e}_n)=a_1x_1+a_2x_2+\dots a_nx_n.$$

Esempio 4.1.5. Sia $c\in k$. L'applicazione

$$\begin{array}{rccc} \mathrm{val}_c: & k[x] & \longrightarrow & k \\ & p & \mapsto & p(c) \end{array} \tag{4.1.5}$$

è lineare. Analogamente, sia X un insieme e sia $x_0\in X$. L'applicazione

$$\begin{array}{rccc} \mathrm{val}_{x_0}: & k^X & \longrightarrow & k \\ & \varphi & \mapsto & \varphi(x_0) \end{array} \tag{4.1.6}$$

è lineare, dove ricordiamo che $k^X:=\{\varphi: X\to k\}$.

Esempio 4.1.6. Siano V uno spazio vettoriale su k e sia $U\subset V$ un sottospazio. L'applicazione quoziente

$$\begin{array}{ccc} V & \xrightarrow{\pi} & V/U \\ v & \mapsto & [v] \end{array} \tag{4.1.7}$$

è lineare.

Proposizione 4.1.7. *Siano V,W spazi vettoriali su uno stesso campo k e $f\colon V\to W$ un'applicazione lineare. Se $v_1,\dots,v_n\in V$ sono linearmente dipendenti allora $f(v_1),\dots,f(v_n)\in W$ sono linearmente dipendenti.*

Dimostrazione. Per ipotesi esistono $\lambda_1,\dots,\lambda_n\in k$ non tutti nulli tali che

$$\lambda_1v_1+\dots+\lambda_nv_n=0. \tag{4.1.8}$$

Applicando f a entrambi i membri di (4.1.8) e sfruttando la linearità di f otteniamo che $\lambda_1f(v_1)+\dots+\lambda_nf(v_n)=0$ e quindi $f(v_1),\dots,f(v_n)\in W$ sono linearmente dipendenti. □

4.1.2 Immagine e nucleo di un'applicazione lineare

Proposizione 4.1.8. *Siano V,W spazi vettoriali su un campo 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 e contiene $0\in V$. Siano $v_1,v_2\in f^{-1}(0)$ e $\lambda_1,\lambda_2\in k$. Per linearità di f abbiamo che

$$f(\lambda_1v_1+\lambda_2v_2)=\lambda_1f(v_1)+\lambda_2f(v_2)=\lambda_10+\lambda_20=0.$$

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.

Certamente $\operatorname{im} f$ contiene $0 \in W$, in quanto $f(0) = 0$. Ora, siano $w_1, w_2 \in \operatorname{im} f$ e $\lambda_1, \lambda_2 \in k$. Quindi esistono $v_1, v_2 \in V$ tali che $f(v_i) = w_i$, da cui

$$\lambda_1 w_1 + \lambda_2 w_2 = \lambda_1 f(v_1) + \lambda_2 f(v_2) = f(\lambda_1 v_1 + \lambda_2 v_2) \in \operatorname{im} f$$

per linearità di f. □

Definizione 4.1.9. Siano V, W spazi vettoriali su un campo k e $f\colon V \to W$ un'applicazione lineare. Il *nucleo di* f è il sottospazio $f^{-1}(0)$, lo si denota $\ker f$.

Esempio 4.1.10. Siano $a_{11}, a_{12}, \ldots, a_{ij}, \ldots, a_{mn}$ elementi di k, dove $1 \le i \le m$ e $1 \le j \le n$. Si verifica facilmente che l'applicazione

$$\begin{array}{ccc} k^n & \xrightarrow{f} & k^m \\ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \mapsto & \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix} \end{array} \tag{4.1.9}$$

è lineare. Il nucleo di f è il sottospazio delle soluzioni del sistema di equazioni lineari omogenee (2.3.2).

Esempio 4.1.11. Sia $c \in k$, e sia $\operatorname{val}_c :: k[x] \to k$ l'applicazione lineare definita da $\operatorname{val}_c(p) = p(c)$, vedi l'**Esempio 4.1.5**. Il nucleo di val_c è il sottospazio $\{(x-c)q \in k[x] \mid q \in k[x]\}$ di $k[x]$.

Esempio 4.1.12. Sia X un insieme e $x_0 \in X$ un suo elemento. Sia $\operatorname{val}_{x_0}\colon k^X \to k$ l'applicazione lineare definita da $\operatorname{val}_{x_0}(\varphi) = \varphi(x_0)$, vedi l'**Esempio 4.1.5**. Il nucleo di val_{x_0} è il sottospazio $\{\varphi \in k^X \mid \varphi(x_0) = 0\}$ di k^X.

Esempio 4.1.13. Siano V uno spazio vettoriale su k e sia $U \subset V$ un sottospazio. L'applicazione quoziente $\pi\colon V \to V/U$ è lineare, vedi l'**Esempio 4.1.6**. Il nucleo di π è U.

Proposizione 4.1.14. *Siano V, W spazi vettoriali su un campo k e $f\colon V \to W$ un'applicazione 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. □

Lemma 4.1.15. *Siano V, W spazi vettoriali sul campo k e sia $S \subset V$ un insieme di generatori per V. Allora l'immagine* $\operatorname{im} f$ *dell'applicazione lineare $f\colon V \to W$ è generata dall'insieme $f(S) := \{f(v) \in W \mid v \in S\}$.*

Dimostrazione. Chiaramente $f(S) \subset \operatorname{im} f$ e dunque $\langle f(S)\rangle \subset \operatorname{im} f$. Per dimostrare che $\operatorname{im} f \subset \langle f(S)\rangle$, consideriamo un vettore $w \in \operatorname{im} f$. Esiste $v \in V$ tale che $f(v) = w$. Poiché S genera V, esistono vettori $v_1, \ldots, v_m \in S$ e $a_1, \ldots, a_m \in k$ tali che $v = a_1 v_1 + \cdots + a_m v_m$. Ne segue che $w = f(v) = f(a_1 v_1 + \cdots + a_m v_m) = a_1 f(v_1) + \cdots + a_m f(v_m) \in \langle f(S)\rangle$. □

Corollario 4.1.16. *Sia $f\colon V \to W$ un'applicazione lineare di spazi vettoriali sul campo k e supponiamo V finitamente generato. Allora* $\ker f$ *e* $\operatorname{im} f$ *sono finitamente generati.*

Dimostrazione. Essendo $\ker f$ un sottospazio di V, esso è finitamente generato per il **Corollario 2.5.13**. D'altra parte, se $S = \{v_1, \ldots, v_n\}$ è un insieme finito di generatori di V, allora $f(S) = \{f(v_1), \ldots, f(v_n)\}$ è un insieme finito di generatori di $\operatorname{im} f$ per il **Lemma 4.1.15**. □

Proposizione 4.1.17. *Siano V, W spazi vettoriali su un campo k, con V finitamente generato. Sia $f\colon V \to W$ un'applicazione lineare. Allora*

$$\dim V = \dim(\ker f) + \dim(\operatorname{im} f). \tag{4.1.10}$$

(L'ipotesi che V sia finitamente generato dà che $\ker f$ *e* $\operatorname{im} f$ *sono finitamente generati per il* **Corollario 4.1.16** *e quindi le loro dimensioni sono ben definite.)*

Dimostrazione. Sia $\{v_1, \ldots, v_a\}$ una base di $\ker f$ e sia $\{v_1, \ldots, v_a, u_1, \ldots, u_b\}$ il suo completamento ad una base di V. Consideriamo i vettori $w_i := f(u_i) \in W$ per $1 \leq i \leq b$. La tesi segue se dimostriamo che $\{w_1, \ldots, w_b\}$ è una base di $\operatorname{im} f$, da cui seguirà che

$$\dim(\operatorname{im} f) = b = (a+b) - a = \dim(V) - \dim(\ker f). \tag{4.1.11}$$

Certamente $\{w_1, \ldots, w_b\}$ genera $\operatorname{im} f$ per il **Lemma 4.1.15**. Per dimostrare che $w_1, \ldots, w_b$ sono linearmente indipendenti, consideriamo una relazione lineare

$$\lambda_1 w_1 + \cdots + \lambda_b w_b = 0 \tag{4.1.12}$$

con $\lambda_1, \ldots, \lambda_b \in k$: vogliamo dimostrare che $\lambda_1 = \cdots = \lambda_b = 0$. Per la linearità di f abbiamo

$$f(\lambda_1 u_1 + \cdots + \lambda_b u_b) = \lambda_1 w_1 + \cdots + \lambda_b w_b = 0 \tag{4.1.13}$$

e dunque il vettore $v := \lambda_1 u_1 + \cdots + \lambda_b u_b$ appartiene a $\ker f$. D'altra parte, v può essere scritto come $v = \mu_1 v_1 + \cdots + \mu_a v_a$ per opportuni $\mu_1, \ldots, \mu_a \in k$, in quanto $\{v_1, \ldots, v_a\}$ è una base di $\ker f$. Dalla relazione lineare

$$(\lambda_1 u_1 + \cdots + \lambda_b u_b) - (\mu_1 v_1 + \cdots + \mu_a v_a) = v - v = 0 \tag{4.1.14}$$

concludiamo che $\mu_1 = \cdots = \mu_a = \lambda_1 = \cdots = \lambda_b = 0$, in quanto $\{v_1, \ldots, v_a, u_1, \ldots, u_b\}$ sono una base di V. □

Il seguente risultato segue subito dalla **Proposizione 4.1.17**.

Corollario 4.1.18. *Siano V, W spazi vettoriali su un campo k. Supponiamo che V e W siano finitamente generati. Sia $f\colon V \to W$ un'applicazione lineare. Allora*

$$\dim(\ker f) \geq \dim V - \dim W.$$

Sia $f\colon k^n \to k^m$ l'applicazione lineare data da (4.1.9). Applicando il **Corollario 4.1.18** a f otteniamo che $\dim \ker f \geq (n-m)$, vedi l' **Esempio 2.5.18** e l'**Esercizio 2.16**.

4.1.3 Operazioni tra applicazioni lineari

Proposizione 4.1.19. *Siano V, W spazi vettoriali su un campo k. Siano $f, g\colon V \to W$ applicazioni lineari e $\lambda \in k$. Siano $(f+g)\colon V \to W$ e $\lambda f\colon V \to W$ date da*

$$(f+g)(v) := f(v) + g(v), \qquad (\lambda f)(v) := \lambda f(v). \tag{4.1.15}$$

Allora sia $(f+g)$ che λf sono applicazioni lineari.

Dimostrazione. Abbiamo che

$$\begin{aligned}(f+g)(\lambda_1 v_1 + \lambda_2 v_2) &= f(\lambda_1 v_1 + \lambda_2 v_2) + g(\lambda_1 v_1 + \lambda_2 v_2) = \\ &= \lambda_1 f(v_1) + \lambda_2 f(v_2) + \lambda_1 g(v_1) + \lambda_2 g(v_2) = \lambda_1 (f+g)(v_1) + \lambda_2 (f+g)(v_2).\end{aligned} \tag{4.1.16}$$

Questo dimostra che $(f+g)$ è lineare. Un conto simile dà che λf è lineare. □

Terminologia 4.1.20. Siano V, W spazi vettoriali su un campo k. L'insieme delle funzioni lineari $f\colon V \to W$ è denotato $\mathcal{L}(V, W)$.

Osservazione 4.1.21. Sia $0 \in \mathcal{L}(V, W)$ l'applicazione *nulla* definita da $0(v) = 0$ per ogni $v \in V$. Allora $\mathcal{L}(V, W)$, provvisto dell'applicazione nulla, e della somma e prodotto per scalari di (4.1.15) è uno spazio vettoriale su k - lasciamo la verifica al lettore.

Lemma 4.1.22. *Siano U, V, W spazi vettoriali su un campo 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

$$\begin{aligned}f \circ g(\lambda_1 v_1 + \lambda_2 v_2) &= f(g(\lambda_1 v_1 + \lambda_2 v_2)) = \\ &= f(\lambda_1 g(v_1) + \lambda_2 g(v_2)) = \lambda_1 f \circ g(v_1) + \lambda_2 f \circ g(v_2).\end{aligned} \tag{4.1.17}$$

Questo dimostra che $f \circ g$ è lineare. □

Definizione 4.1.23. Sia V uno spazio vettoriale su un campo k. Il *duale* di V è lo spazio vettoriale delle funzioni lineari $f\colon V \to k$ (cioè $\mathcal{L}(V, k)$), ed è denotato V^*.

4.2 Isomorfismi

Definizione 4.2.1. Siano V, W spazi vettoriali su uno stesso campo 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

$$g \circ f = \mathrm{Id}_V, \qquad f \circ g = \mathrm{Id}_W\,. \tag{4.2.1}$$

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 4.2.2. *Siano V, W spazi vettoriali su uno stesso campo 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 (4.2.1). Ora supponiamo che esista un'inversa g di f, cioè che valga (4.2.1), senza supporre che g sia lineare, e dimostriamo che g è lineare. Siano $w_1, w_2 \in W$ e $\lambda_1, \lambda_2 \in k$. Abbiamo che

$$f(g(\lambda_1 w_1 + \lambda_2 w_2)) = \mathrm{Id}(\lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 w_1 + \lambda_2 w_2$$

e

$$f(\lambda_1 g(w_1) + \lambda_2 g(w_2)) = \lambda_1 f(g(w_1)) + \lambda_2 f(g(w_2)) = \lambda_1 w_1 + \lambda_2 w_2.$$

Quindi $f(g(\lambda_1 w_1 + \lambda_2 w_2)) = f(\lambda_1 g(w_1) + \lambda_2 g(w_2))$. Siccome f è invertibile segue che

$$g(\lambda_1 w_1 + \lambda_2 w_2)) = \lambda_1 g(w_1) + \lambda_2 g(w_2)$$

e questo dimostra che g è lineare. □

Esempio 4.2.3. Sia V uno spazio vettoriale su k, finitamente generato e sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V. L'applicazione

$$\begin{array}{ccc} k^n & \xrightarrow{f} & V \\ (x_1, \ldots, x_n) & \longrightarrow & x_1 v_1 + x_2 v_2 + \ldots + x_n v_n \end{array} \tag{4.2.2}$$

è biunivoca per il **Corollario 2.5.8** e quindi f è un isomorfismo.

Osservazione 4.2.4. (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 k. Per definizione anche f^{-1} è un isomorfismo.

(3) Siano U, V, W spazi vettoriali su uno stesso campo 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 **Lemma 4.1.22**).

Segue che la relazione di isomorfismo tra spazi vettoriali è di equivalenza.

Esempio 4.2.5. Sia V uno spazio vettoriale su k, finitamente generato e sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V. Per l'**Esempio 4.2.3** e il punto (2) dell'**Osservazione 4.2.4**, l'applicazione

$$V \xrightarrow{X_{\mathcal{B}}} k^n, \tag{4.2.3}$$

che associa a un vettore di V il vettore delle sue coordinate, è un isomorfismo.

Supponiamo che $f \colon V \to W$ sia un isomorfismo tra spazi vettoriali sullo stesso campo 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 4.2.6. *Siano V, W spazi vettoriali su uno stesso campo k e supponiamo che $f \colon V \to W$ sia un isomorfismo. Siano $v_1, \ldots, v_n \in V$.*

(1) $v_1, \ldots, v_n$ sono linearmente dipendenti se e solo se $f(v_1), \ldots, f(v_n) \in W$ sono linearmente dipendenti.

(2) $v_1, \ldots, v_n$ generano V se e solo se $f(v_1), \ldots, f(v_n)$ generano W.

Dimostrazione. Ricordiamo che f è biunivoca e f^{-1} è lineare, e quindi $\operatorname{im} f = W$ e $\operatorname{im} f^{-1} = V$.

La (1) è equivalente a dimostrare che $v_1, \dots, v_n$ sono linearmente dipendenti se e solo se $f(v_1), \dots, f(v_n)$ sono linearmente dipendenti.

Se $v_1, \dots, v_n$ sono linearmente dipendenti allora $f(v_1), \dots, f(v_n) \in W$ sono linearmente dipendenti per la **Proposizione 4.1.7**. Viceversa, se $f(v_1), \dots, f(v_n) \in W$ sono linearmente dipendenti, allora $v_1 = f^{-1}(f(v_1)), \dots, v_n = f^{-1}(f(v_n)) \in V$ sono linearmente dipendenti sempre per la **Proposizione 4.1.7** applicata a f^{-1}.

Per la (2), se $v_1, \dots, v_n$ generano V, allora $f(v_1), \dots, f(v_n)$ generano $\operatorname{im} f = W$ per il **Lemma 4.1.15**. Viceversa, se $f(v_1), \dots, f(v_n)$ generano W, allora $v_1 = f^{-1}(f(v_1)), \dots, v_n = f^{-1}(f(v_n))$ generano $\operatorname{im} f^{-1} = V$ sempre per il **Lemma 4.1.15** applicato a f^{-1}. □

Il corollario qui sotto segue immediatamente dalla **Proposizione 4.2.6**.

Corollario 4.2.7. *Siano V, W spazi vettoriali su uno stesso campo 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 4.2.7** due spazi vettoriali finitamente generati isomorfi hanno la stessa dimensione. Vale il viceversa:

Proposizione 4.2.8. *Siano V, W spazi vettoriali su uno stesso campo 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 4.2.5**, abbiamo isomorfismi

$$X_{\mathcal{B}}\colon V \xrightarrow{\sim} k^n, \qquad X_{\mathcal{C}}\colon W \xrightarrow{\sim} k^n,$$

e quindi $X_{\mathcal{C}}^{-1} \circ X_{\mathcal{B}}\colon V \to W$ è un isomorfismo - vedi **Osservazione 4.2.4**. □

Proposizione 4.2.9. *Siano V, W spazi vettoriali finitamente generati su uno stesso campo k e tali che* $\dim V = \dim W$. *Sia $f\colon V \to W$ lineare e supponiamo che almeno una delle seguenti due ipotesi sia soddisfatta:*

(1) $\ker f = \{0\}$*;*

(2) f è suriettiva.

Allora f è un isomorfismo.

Dimostrazione. (1): per la **Proposizione 4.1.17** 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 4.1.14**. Per il **Lemma 4.2.2** segue che f è un isomorfismo. (2): per la **Proposizione 4.1.17** otteniamo che $\dim(\ker f) = \dim V - \dim W = 0$ e quindi f è iniettiva per la **Proposizione 4.1.14**. . Per il **Lemma 4.2.2** segue che f è un isomorfismo. □

Definizione 4.2.10. Sia V uno spazio vettoriale su un campo k.

1. Un *automorfismo* di V è un isomorfismo $f\colon V \xrightarrow{\sim} V$.
2. Il *gruppo lineare generale* $\mathrm{GL}(V)$ è l'insieme degli automorfismi $f\colon V \xrightarrow{\sim} V$.

3. Sia k un campo: si pone
$$\mathrm{GL}_n(k) := \mathrm{GL}(k^n). \tag{4.2.4}$$

La definizione di $\mathrm{GL}(V)$ e l'**Osservazione 4.2.4** danno che:

1. $\mathrm{Id}_V \in \mathrm{GL}(V)$.
2. Se $f, g \in \mathrm{GL}(V)$, allora $f \circ g \in \mathrm{GL}(V)$.
3. Se $f \in \mathrm{GL}(V)$, allora $f^{-1} \in \mathrm{GL}(V)$.
4. Se $f, g, h \in \mathrm{GL}(V)$, allora $(f \circ g) \circ h = f \circ (g \circ h)$.

L'insieme $\mathrm{GL}(V)$ degli automorfismi di V munito dell'operazione $\circ$ di composizione si dice *gruppo* perché valgono le 4 proprietà elencate sopra.

4.3 Il primo Teorema di isomorfismo

Siano V, W spazi vettoriali su k, e sia $f\colon V \to W$ un'applicazione lineare. Sia
$$\pi\colon V \to V/\ker f$$
l'applicazione quoziente.

Proposizione 4.3.1. *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 perché l'applicazione quoziente π è suriettiva. Rimane da dimostrare che $\overline{f}$ è lineare. Se $\lambda_1, \lambda_2 \in k$ e $v_1, v_2 \in V$,
$$\overline{f}(\lambda_1[v_1]+\lambda_2[v_2]) = f([\lambda_1 v_1 + \lambda_2 v_2]) = f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2) = \lambda_1 f([v_1]) + \lambda_2 f([v_2]).$$
□

Ovviamente l'immagine di $\overline{f}$ è contenuta in $\mathrm{im}\, f$, e quindi definisce un'applicazione lineare $V/\ker f \longrightarrow \mathrm{im}\, f$ che continueremo a denotare $\overline{f}$ (abusando della notazione).

Teorema 4.3.2 (Primo Teorema di Isomorfismo)**.** *Mantenendo le ipotesi e notazioni appena introdotte, l'applicazione lineare $\overline{f}\colon V/\ker f \longrightarrow \mathrm{im}\, f$ è un isomorfismo.*

Dimostrazione. L'immagine di $\overline{f}$ è uguale all'immagine di f, e quindi $\overline{f}$ è suriettiva (su $\mathrm{im}\, f$!). 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 4.3.3. Mantenendo le ipotesi e notazioni appena introdotte, supponiamo che V sia finitamente generato. Allora $\dim(V/\ker f) = \dim \mathrm{im}\, f$ per il Primo Teorema di Isomorfismo, ma d'altra parte $\dim(V/\ker f) = \dim V - \dim \ker f$ per la **Proposizione 2.7.4**. Questo dimostra di nuovo che $\dim V = \dim \ker f + \dim \mathrm{im}\, f$, cioè la **Proposizione 4.1.17**.

4.4 Matrici

Le matrici sono uno strumento indispensabile per fare conti con applicazioni lineari. Cominceremo definendo le operazioni tra matrici, e poi inizieremo a stabilire la relazione tra matrici e applicazioni lineari.

4.4.1 Calcolo matriciale

Una *matrice* $m \times n$ *a valori in* k è un'applicazione $\{1, \ldots, m\} \times \{1, \ldots, n\} \to k$: quindi è determinata dall'insieme dei valori $a_{ij} \in k$ associati a (i, j) dove $1 \leq i \leq m$ e $1 \leq j \leq n$. È conveniente scrivere la matrice come una tabella:

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ a_{31} & a_{32} & \ldots & a_{3n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

Denotiamo la matrice A con (a_{ij}). La *riga* i-esima di A è

$$A^i := (a_{i1}, a_{i2}, \ldots, a_{i_n}) \in k^n. \tag{4.4.1}$$

La *colonna* j-esima di A è

$$A_j := \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \in k^m \tag{4.4.2}$$

Definizione 4.4.1. $M_{m,n}(k)$ è l'insieme delle matrici $m \times n$ a valori in k.

Esistono alcune operazioni fondamentali sulle matrici. La somma è definita da

$$\begin{array}{ccc} M_{m,n}(k) \times M_{m,n}(k) & \longrightarrow & M_{m,n}(k) \\ ((a_{ij}), (b_{ij})) & \mapsto & (a_{ij} + b_{ij}) \end{array}$$

Possiamo identificare in modo ovvio $M_{m,n}(k)$ con k^{mn} e con questa identificazione la somma corrisponde alla somma in k^{mn}. La moltiplicazione di vettori di k^{mn} per scalari (in k) corrisponde alla moltiplicazione

$$\begin{array}{ccc} k \times M_{m,n}(k) & \longrightarrow & M_{m,n}(k) \\ (\lambda, (a_{ij})) & \mapsto & (\lambda a_{ij}) \end{array}$$

Con queste operazioni $M_{m,n}(k)$ è uno spazio vettoriale su k e abbiamo un isomorfismo

$$k^{mn} \xrightarrow{\sim} M_{m,n}(k).$$

Terminologia 4.4.2. Indicheremo con $0_{m,n}$ la matrice nulla $m \times n$ ovvero l'elemento neutro di $M_{m,n}(k)$.

Esiste un'altra operazione fondamentale sulle matrici.

Definizione 4.4.3. Siano $A \in M_{m,n}(k)$ e $B \in M_{n,p}(k)$. 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 $1 \leq i \leq 1m$ e $1 \leq h \leq p$) 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

$$A = (a_1, \ldots, a_n), \qquad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Allora

$$A \cdot B = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n.$$

In generale

$$A \cdot B = \begin{bmatrix} A^1 \cdot B_1 & A^1 \cdot B_2 & \ldots & A^1 \cdot B_n \\ A^2 \cdot B_1 & A^2 \cdot B_2 & \ldots & A^2 \cdot B_n \\ \ldots & \ldots & \ldots & \ldots \\ A^m \cdot B_1 & A^m \cdot B_2 & \ldots & A^m \cdot B_n \end{bmatrix} \tag{4.4.3}$$

Questo giustifica il nome "moltiplicazione righe per colonne".

Esempio 4.4.4. Siano $A, B, C \in M_{2,2}(\mathbb{R})$ le matrici

$$A := \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix}, \quad B := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad C := \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}. \tag{4.4.4}$$

Tutte le matrici di (4.4.4) sono 2×2, quindi ha senso moltiplicare due qualsiasi tali matrici. Calcolando otteniamo che

$$A \cdot B := \begin{bmatrix} 0 & a_1 \\ 0 & 0 \end{bmatrix}, \quad B \cdot A := \begin{bmatrix} 0 & a_2 \\ 0 & 0 \end{bmatrix}, \quad B \cdot B := \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad C \cdot C := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{4.4.5}$$

Qundi vediamo che la moltiplicazione di matrici non ha le stesse proprietà algebriche della moltiplicazione di numeri reali (o più in genrale di elementi di un campo). I primi due prodotti di (4.4.5) fanno vedere che in generale il prodotto **non** è commutativo. Il terzo prodotto di (4.4.5) dà una matrice non nulla il cui quadrato è nullo, il quarto prodotto dimostra che esistono infinite "radici quadrate" di una stessa matrice.

L'**Esempio 4.4.4** 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

$$\delta_{ij} := \begin{cases} 1 & \text{se } i = j, \\ 0 & \text{se } i \neq j. \end{cases} \tag{4.4.6}$$

Definizione 4.4.5. (1) La matrice *unità* $n \times n$ è la matrice $1_n := (\delta_{ij})$ (qui $1 \leq i, j \leq n$).

(2) Una matrice $A \in M_{n,n}(k)$ è *scalare* se esiste $\lambda \in k$ tale che $M = \lambda 1_n$.

(3) Una matrice $A \in M_{n,n}(k)$ è *diagonale* se esistono $\lambda_i \in k$ per $1 \leq i \leq n$ 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 4.4.6. *Siano $\lambda \in k$, $A \in M_{m,n}(k)$, $B, B' \in M_{n,p}(k)$ e $C \in M_{p,q}(k)$. 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

$$s_{il} = \sum_{h=1}^{p} (\sum_{j=1}^{n} a_{ij} b_{jh}) c_{hl} = \sum_{\substack{1 \leq j \leq n \\ 1 \leq h \leq p}} a_{ij} b_{jh} c_{hl}$$

e

$$t_{il} = \sum_{j=1}^{n} a_{ij} (\sum_{h=1}^{p} b_{jh} c_{hl}) = \sum_{\substack{1 \leq j \leq n \\ 1 \leq h \leq p}} a_{ij} b_{jh} c_{hl}$$

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 (4.4.3). La seconda eguaglianza di (3) si verifica in modo simile. □

Osservazione 4.4.7. Sia A una matrice *quadrata* cioè $A \in M_{n,n}(k)$ per un qualche n. Quindi ogni prodotto che coinvolge solo fattori uguali ad A (per esmpio $A \cdot ((A \cdot A) \cdot A)$) ha senso. Se cambiamo la disposizione delle parentesi il prodotto non cambia perché il prodotto è l'associativo. Si pone $A^0 := 1_n$. Quindi ha senso A^r per un qualsiasi $r \in \mathbb{N}$. Vale

$$A^r \cdot A^s = A^{r+s}, \qquad r, s \in \mathbb{N}.$$

Definizione 4.4.8. Sia A una matrice *quadrata* cioè $A \in M_{n,n}(k)$ per un qualche n. Una matrice $B \in M_{n,n}(k)$ è una *inversa* di A se

$$A \cdot B = 1_n = B \cdot A.$$

Esempio 4.4.9. Siano A, B, C le matrici di (4.4.4). A ha un'inversa se e solo se $a_1 \neq 0 \neq a_2$, B non ha inversa, C ha inversa (uguale a C).

Lemma 4.4.10. *Se $A \in M_{n,n}(k)$ ha un'inversa allora ha un'unica inversa.*

Dimostrazione. Siano B, B' inverse di A. Allora

$$B' = B' \cdot 1_n = B' \cdot (A \cdot B) = (B' \cdot A) \cdot B = 1_n \cdot B = B$$

□

Definizione 4.4.11. Sia $A \in M_{n,n}(k)$ invertibile. Denotiamo con A^{-1} l'unica inversa di A.

Osservazione 4.4.12. Sia $A \in M_{n,n}(k)$ invertibile. Se $r \in \mathbb{Z}$ è negativo poniamo

$$A^r := (A^{-1})^{-r}.$$

Quindi ha senso A^r per ogni $r \in \mathbb{Z}$. Vale

$$A^r \cdot A^s = A^{r+s}, \qquad r, s \in \mathbb{Z}.$$

Sarà utile considerare la seguente operazione che produce una matrice $n \times m$ a partire da una matrice $m \times n$.

Definizione 4.4.13. Sia $A \in M_{m,n}(k)$. La *trasposta* di A è la matrice $A^t \in M_{n,m}(k)$ le cui righe sono le colonne di A. Più precisamente poniamo $A = (a_{ij})$ e $A^t = (b_{ij})$. Allora $b_{ij} = a_{ji}$.

Osservazione 4.4.14. Siano $A, B \in M_{m,n}(k)$, e $C \in M_{n,p}(k)$. Un calcolo dà le seguenti uguaglianze:

$$(A + B)^t = A^t + B^t, \qquad (B \cdot C)^t = C^t \cdot B^t.$$

4.4.2 Matrici e applicazioni lineari $k^n \to k^m$.

Ora iniziamo a studiare la relazione tra matrici e applicazioni lineari. Sia $A \in M_{m,n}(k)$: da (1) e (3) della **Proposizione 4.4.6** segue che l'applicazione

$$\begin{array}{ccc} k^n & \xrightarrow{L_A} & k^m \\ X & \mapsto & A \cdot X \end{array} \tag{4.4.7}$$

è lineare - qui l'elemento $X \in k^n$ è visto come matrice $n \times 1$ cioè come *vettore colonna*. Sia $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ la base standard di k^n. Allora

$$L_A(\mathbf{e}_j) = A_j. \tag{4.4.8}$$

Proposizione 4.4.15. *Sia $f\colon k^n \to k^m$ un'applicazione lineare. Esiste una e una sola matrice $A \in M_{m,n}(k)$ tale che $f = L_A$.*

Dimostrazione. La formula (4.4.8) 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 $f\colon k^n \to k^m$ sia lineare. Definiamo $A \in M_{m,n}(k)$ imponendo che valga (4.4.8). Dimostriamo che $L_A = f$. Sia $v \in k^n$: allora esistono $x_1, \ldots, x_n \in k$ tali che $v = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n$. Per linearità di f e L_A abbiamo che

$$\begin{aligned} f(v) = f(x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n) = x_1 f(\mathbf{e}_1) + \ldots + x_n f(\mathbf{e}_n) = x_1 A_1 + \ldots + x_n A_n = \\ = x_1 L_A(\mathbf{e}_1) + \ldots + x_n L_A(\mathbf{e}_n) = L_A(x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n) = \varphi_A(v). \end{aligned} \tag{4.4.9}$$

Questo dimostra che $L_A = f$. □

Per la **Proposizione 4.4.15** abbiamo un'applicazione biunivoca

$$\begin{array}{ccc} M_{m,n}(k) & \longrightarrow & \mathcal{L}(k^n, k^m) \\ A & \mapsto & L_A \end{array} \tag{4.4.10}$$

Sia $M_{m,n}(k)$ che $\mathcal{L}(k^n, k^m)$ sono k-spazi vettoriali. Si verifica facilmente che (4.4.10) è un'applicazione lineare: siccome è anche biunivoca segue che è un isomorfismo di spazi vettoriali per il **Lemma 4.2.2**.

Proposizione 4.4.16. *Se $A \in M_{m,n}(k)$ e $B \in M_{n,p}(k)$, allora*

$$L_A \circ L_B = L_{A \cdot B}. \tag{4.4.11}$$

Dimostrazione. Sia $X \in k^p$ (vettore colonna): per l'associatività del prodotto di matrici abbiamo che

$$(L_A \circ L_B)(X) = L_A(L_B(X)) = A \cdot (B \cdot X) = (A \cdot B) \cdot X = L_{A \cdot B}(X).$$

□

Esempio 4.4.17. Sia

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

Definiamo $x_n, y_n \in \mathbb{N}$ così:

$$(x_n, y_n) := L_{A^{n-1}}(1, 1).$$

Allora la successione $\{x_n\}$ è la successione di Fibonacci[1]. Infatti

$$(x_{n+1}, y_{n+1}) := L_{A^n}(1, 1) = L_A \circ L_{A^{n-1}}(1, 1) = L_A((x_n, y_n)) = (x_n + y_n, x_n). \tag{4.4.12}$$

Sostituendo $(n-1)$ a n nella (4.4.12) abbiamo che $y_n = x_{n-1}$ e quindi (4.4.12) dà che $x_{n+1} = (x_n + x_{n-1})$ cioè $\{x_n\}$ soddisfa la formula ricorsiva che definisce la successione di Fibonacci. Siccome $x_1 = 1$ e $x_2 = 2$ segue $\{x_n\}$ è la successione di Fibonacci. Per la (4.4.12) segue anche che $\{y_{n+1}\}$ è la successione di Fibonacci.

Osservazione 4.4.18. Sia $A \in M_{n,n}(k)$. Per (4.4.11) la matrice A è invertibile se e solo se L_A è invertibile cioè è un isomorfismo. Per la **Proposizione 4.2.9** otteniamo il seguente risultato (non banale): il sistema di equazioni lineari

$$\begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots & = & *, \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n & = & b_i, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots & = & *, \\ a_{n1}x_1 + a_{m2}x_2 + \ldots + a_{nn}x_n & = & b_n. \end{array} \tag{4.4.13}$$

(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.

L'isomorfismo (4.4.10) dà l'identificazione

$$\mathrm{GL}_n(k) = \{A \in M_{n,n}(k) \mid A \text{ è invertibile}\}.. \tag{4.4.14}$$

(Ricordiamo che $\mathrm{GL}_n(k)$ è il gruppo degli isomorfismi di k^n, vedi (4.2.4).) Inoltre l'uguaglianza (4.4.11) dà che la composizione di elementi di $\mathrm{GL}_n(k)$ è identificata con il prodotto righe per colonne di matrici.

[1]Per definizione $1 = x_0 = x_1$ e $x_n = x_{n-1} + x_{n-2}$ per $n \geq 2$.

4.5 La matrice associata ad un'applicazione lineare

Siano V, W spazi vettoriali sullo stesso campo k e $f\colon V \to W$ un'applicazione lineare. Supponiamo che V, W siano finitamente generati. Scegliamo una base $\mathcal{B} = \{v_1, \dots, v_n\}$ di V e una base $\mathcal{C} = \{w_1, \dots, w_m\}$ di W. Associamo a f la matrice $A = (a_{ij}) \in M_{m,n}(k)$ definita così:

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i. \tag{4.5.1}$$

In altre parole la colonna j-esima di A è la colonna delle coordinate di $f(v_j)$ nella base $\mathcal{C}$.

Definizione 4.5.1. La matrice $M_{\mathcal{C}}^{\mathcal{B}}(f)$ è la matrice data da (4.5.1).

Esempio 4.5.2. Siano $k = \mathbb{R}$, $V = W = \mathbb{R}[x]_{\leq 2}$ e $\mathcal{B} = \mathcal{C} = \{1, x, x^2\}$. Sia

$$\begin{array}{ccc} \mathbb{R}[x]_{\leq 2} & \xrightarrow{f} & \mathbb{R}[x]_{\leq 2} \\ p & \mapsto & p + p' \end{array}$$

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 4.5.3. Siano $k = \mathbb{R}$, $V = W = \mathcal{V}^2$ e sia $\mathcal{B} = \mathcal{C} = \{\mathbf{i}, \mathbf{j}\}$ dove $\mathbf{i}, \mathbf{j}$ sono vettori di uguale lunghezza e ortogonali tra loro. Siano $P \in \mathcal{A}^2$ e $\theta \in \mathbb{R}$: sia $R_\theta\colon \mathcal{A}^2 \to \mathcal{A}^2$ la rotazione di centro P e angolo θ con verso di rotazione "da $\mathbf{i}$ a $\mathbf{j}$". Se P_1Q_1 e P_2Q_2 sono segmenti orientati equipollenti anche $R_\theta(P_1)R_\theta(Q_1)$ e $R_\theta(P_2)R_\theta(Q_2)$ sono equipollenti: quindi R_θ induce un'applicazione

$$r_\theta\colon \mathcal{V}^2 \to \mathcal{V}^2.$$

La r_θ è lineare (verificatelo). Abbiamo

$$M_{\mathcal{B}}^{\mathcal{B}}(r_\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{4.5.2}$$

Esempio 4.5.4. Sia $V = k^n$ e $W = k^m$ (quindi il campo è 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 k^n e k^m rispettivamente. Sia $A \in M_{m,n}(k)$: allora

$$M_{\mathcal{C}}^{\mathcal{B}}(L_A) = A.$$

Proposizione 4.5.5. *Siano V, W spazi vettoriali sullo stesso campo k. Supponiamo che V, W siano finitamente generati. Siano $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V e $\mathcal{C} = \{w_1, \dots, w_m\}$ una base di W. Siano $X_{\mathcal{B}}(v)$ e $X_{\mathcal{C}}(f(v))$ le matrici colonna delle coordinate di v e $f(v)$ nelle basi $\mathcal{B}$ e $\mathcal{C}$ rispettivamente. Abbiamo*

$$X_{\mathcal{C}}(f(v)) = M_{\mathcal{C}}^{\mathcal{B}}(f) \cdot X_{\mathcal{B}}(v). \tag{4.5.3}$$

Sia $M \in M_{m,n}$ tale che valga (4.5.3) *con $M_{\mathcal{C}}^{\mathcal{B}}(f)$ sostituita da M. Allora $M = M_{\mathcal{C}}^{\mathcal{B}}(f)$.*

Dimostrazione. Poniamo $A := M_{\mathcal{C}}^{\mathcal{B}}(f)$. Per linearità di f e per definizione di $M_{\mathcal{C}}^{\mathcal{B}}(f)$ abbiamo

$$f(\sum_{j=1}^{n} x_j v_j) = \sum_{j=1}^{n} x_j f(v_j) = \sum_{j=1}^{n} x_j (\sum_{i=1}^{m} a_{ij} w_i) = \sum_{i=1}^{m} (\sum_{j=1}^{n} a_{ij} x_j) w_i.$$

Segue che la coordinata i-esima di $f(v)$ è il prodotto della riga i-esima di A per la matrice colonna $X_{\mathcal{B}}(v)$: questo dimostra che vale (4.5.3). Ora supponiamo che valga (4.5.3) con $M_{\mathcal{C}}^{\mathcal{B}}(f)$ sostituita da M. Allora $(M - M_{\mathcal{C}}^{\mathcal{B}}(f)) \cdot X_{\mathcal{B}}(\mathbf{e}_j) = 0$ per $1 \leq j \leq n$. Segue che la colonna j-esima di $(M - M_{\mathcal{C}}^{\mathcal{B}}(f))$ è nulla per $1 \leq j \leq n$, cioè $(M - M_{\mathcal{C}}^{\mathcal{B}}(f))$ è la matrice nulla ovvero $M = M_{\mathcal{C}}^{\mathcal{B}}(f)$. □

Poniamo $A := M_{\mathcal{C}}^{\mathcal{B}}(f)$: è conveniente sintetizzare la (4.5.3) con il seguente diagramma

(4.5.4)

$$\begin{array}{ccc} V & \xrightarrow{\quad f \quad} & W \\ \downarrow{\scriptstyle X_{\mathcal{B}}} & & \downarrow{\scriptstyle X_{\mathcal{C}}} \\ k^n & \xrightarrow{\quad L_A \quad} & k^m \end{array}$$

e osservando che la (4.5.3) equivale ad affermare che partendo da $v \in V$ e arrivando a un vettore di k^m seguendo le due strade possibili arriveremo comunque allo stesso vettore.

Proposizione 4.5.6. *Siano V, W spazi vettoriali sullo stesso campo k. Supponiamo che V, W siano finitamente generati. Siano $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V e $\mathcal{C} = \{w_1, \ldots, w_m\}$ una base di W. L'applicazione*

$$\begin{array}{ccc} \mathcal{L}(V,W) & \longrightarrow & M_{m,n}(k) \\ f & \mapsto & M_{\mathcal{B}}^{\mathcal{C}}(f) \end{array} \tag{4.5.5}$$

è biunivoca.

Dimostrazione. L'iniettività segue da (4.5.3). Per dimostrare la suriettività consideriamo $A \in M_{m,n}(k)$. Poniamo $f := X_{\mathcal{C}}^{-1} \circ L_A \circ X_{\mathcal{B}} \colon V \to W$ - guardate (7.2.1). La f è una composizione di applicazioni lineari e quindi è lineare. Verifichiamo che $M_{\mathcal{B}}^{\mathcal{C}}(f) = A$. Abbiamo che

$$f(\mathbf{e}_j) = X_{\mathcal{C}}^{-1} \circ L_A(X_{\mathcal{B}}(v_j)) = X_{\mathcal{C}}^{-1} \circ L_A(\mathbf{e}_j) = X_{\mathcal{C}}^{-1}(\sum_{i=1}^{m} a_{ij} \mathbf{e}'_i) = \sum_{i=1}^{m} a_{ij} w_i.$$

Questo dimostra che $M_{\mathcal{B}}^{\mathcal{C}}(f) = A$. □

Proposizione 4.5.7. *(1) L'applicazione* (4.5.5) *è un isomorfismo di spazi vettoriali.*

(2) Siano U, V, W spazi vettoriali su k e siano $g \colon U \to V$, $f \colon V \to W$ applicazioni lineari. Siano $\mathcal{B}$ una base di U, $\mathcal{C}$ una base di V e $\mathcal{D}$ una base di W. Allora

$$M_{\mathcal{D}}^{\mathcal{B}}(f \circ g) = M_{\mathcal{D}}^{\mathcal{C}}(f) \cdot M_{\mathcal{C}}^{\mathcal{B}}(g). \tag{4.5.6}$$

Dimostrazione. (1). Siano $\lambda_1, \lambda_2 \in k$ e $f_1, f_2 \in \mathcal{L}(V,W)$. Per linearità di $X_{\mathcal{C}}$ e per la **Proposizione 4.5.5** 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 4.5.5** 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 (4.5.5) è lineare: siccome è biunivoca è un isomorfismo per il **Lemma 4.2.2**. (2). Abbiamo

$$X_{\mathcal{D}}((f \circ g)(v)) = X_{\mathcal{D}}((f(g(v))) = M_{\mathcal{D}}^{\mathcal{C}}(f) \cdot X_{\mathcal{C}}(g(v)) = \\ = M_{\mathcal{D}}^{\mathcal{C}}(f) \cdot (M_{\mathcal{C}}^{\mathcal{B}}(g) \cdot X_{\mathcal{B}}(v)) = (M_{\mathcal{D}}^{\mathcal{C}}(f) \cdot M_{\mathcal{C}}^{\mathcal{B}}(g)) \cdot X_{\mathcal{B}}(v).$$

Per la **Proposizione 4.5.5** concludiamo che vale (2). □

Esempio 4.5.8. Siano $\alpha, \beta \in \mathbb{R}$. Applichiamo la (2) della **Proposizione 4.5.7** alla rotazione $r_{\alpha+\beta}$ dell'**Esempio 4.5.3**. La base $\mathcal{B}$ di $\mathcal{V}^2$ è come nell'**Esempio 4.5.3**. 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.

4.6 Operazioni elementari sulle matrici, I

4.6.1 Il problema

Siano V, W spazi vettoriali su un campo k. Supponiamo che V, W siano finitamente generati. Sia $f\colon V \to W$ un'applicazione lineare.

Problema 4.6.1. Dare algoritmi efficienti per

(1) trovare una base di $\operatorname{im} f$,

(2) trovare una base di $\ker f$.

Definizione 4.6.2. Sia $f\colon V \to W$ un'applicazione lineare tra spazi vettoriali finitamente generati su un campo k. Il *rango* di f è la dimensione dell'immagine di f - lo denotiamo $\operatorname{rk} f$. Se $A \in M_{m,n}(k)$ il *rango* di A è la dimensione dell'immagine di L_A - lo denotiamo $\operatorname{rk} A$.

Quindi uno dei problemi che vogliamo risolvere è quello di calcolare il rango di un'applicazione lineare (tra spazi finitamente generati). Il primo passo consiste nello scegliere una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ di V, una base $\mathcal{C} = \{w_1, \ldots, w_m\}$ di W e associare a f la matrice

$$A = M_{\mathcal{C}}^{\mathcal{B}}(f) \in M_{m,n}(k). \tag{4.6.1}$$

Un vettore $w \in W$ appartiene a $\operatorname{im} f$ se e solo se il vettore colonna $X_{\mathcal{C}}(w)$ è nel sottospazio di k^m generato dalle colonne di A, e, analogamente, $v \in V$ è in $\ker f$ se e solo se $A \cdot X_{\mathcal{B}}(v) = 0$. Ne segue che risolvere **Problema 4.6.1** equivale a risolvere il seguente

Problema 4.6.3. Data una matrice $A \in M_{m,n}(k)$, dare algoritmi efficienti per

(1) trovare una base del sottospazio di k^m generato dalle colonne di A (cioè una base di $\operatorname{im} L_A$),

(2) trovare una base dello spazio delle soluzioni (in k^n) del sistema di equazioni omogenee $A \cdot X = 0$ associato ad A (cioè una base di $\ker L_A$).

4.6.2 Matrici a scala

Se la matrice A ha una forma particolare si risponde facilmente al primo punto del **Problema 4.6.3**, e, analogamente per il punto (2) del **Problema 4.6.3**. Definiamo quali sono le matrici "particolari" in questione. Sia $A \in M_{m,n}(k)$. Per $1 \leq i \leq m$ definiamo

$$p_A(i) := \begin{cases} \min\{1 \leq j \leq n \mid a_{ij} \neq 0\} & \text{se } A^i \neq \mathbf{0} \\ \infty & \text{se } A^i = \mathbf{0}. \end{cases}$$

Per $1 \leq j \leq n$ definiamo

$$q_A(j) := \begin{cases} \min\{1 \leq i \leq m \mid a_{ij} \neq 0\} & \text{se } A_j \neq \mathbf{0} \\ \infty & \text{se } A_j = \mathbf{0}. \end{cases}$$

Definizione 4.6.4. Sia $A \in M_{m,n}(k)$.

(1) A è a *scala per colonne* se $q_A(1) < q_A(2), \ldots < q_A(n)$ (per convenzione $\infty < \infty$).

(2) A è a *scala per righe* se $p_A(1) < p_A(2), \ldots < p_A(m)$.

Esempio 4.6.5. Le seguenti matrici sono a scala per colonne

$$\begin{bmatrix} 2 & 0 \\ 0 & 5 \\ 3 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 0 & 0 \\ 3 & e & 0 \end{bmatrix}, \quad \begin{bmatrix} \pi & 0 \\ 0 & \sqrt{2} \end{bmatrix}, \tag{4.6.2}$$

e le seguenti sono a scala per righe

$$\begin{bmatrix} 2 & -1 \\ 0 & 5 \end{bmatrix}, \quad \begin{bmatrix} 1 & -4 & 5 \\ 0 & 2 & 3 \end{bmatrix}, \quad \begin{bmatrix} \pi & 0 \\ 0 & \sqrt{2} \end{bmatrix}. \tag{4.6.3}$$

D'altra parte, ciascuna delle seguenti matrici non è a scala per colonne, né a scala per righe.

$$\begin{bmatrix} 2 & -1 \\ \sqrt{5} & 5 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 3 \\ -1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} \pi & \sqrt{2} \\ 0 & 3 \\ -1 & 5 \end{bmatrix}. \tag{4.6.4}$$

Osservazione 4.6.6. Una matrice A è a scala per righe se e solo se la sua trasposta A^t è a scala per colonne.

Ora osserviamo che, se A è a scala per colonne, si trova immediatamente una base di $\operatorname{im} L_A$, e che se A è a scala per righe, si trova facilmente una base di $\ker L_A$.

Osservazione 4.6.7. Sia $S \in M_{m,n}(k)$, e supponiamo che S sia a scala per colonne. Il sottospazio di k^m generato dalle colonne di S (cioè $\operatorname{im} L_S$) ha per base l'insieme delle colonne non nulle di S, in particolare la sua dimensione è uguale al numero di colonne non nulle di S. Infatti $\operatorname{im} L_S$ delle colonne non nulle di S, e si vede facilmente che le colonne non nulle di S sono linearmente indipendenti: se $S_1, S_2, \ldots S_r$ sono le colonne non nulle di S, e

$$\lambda_1 S_1 + \ldots \lambda_r S_r = 0,$$

allora $\lambda_1 = 0$ perché $q_S(1) > q_S(2) > \ldots q_S(r) > 0$, e, analogamente $\lambda_2 = \ldots = \lambda_r = 0$.

Osservazione 4.6.8. Sia $S \in M_{m,n}(k)$, e supponiamo che S sia a scala per righe. L'insieme delle soluzioni di $S \cdot X = \mathbf{0}$ si ottiene facilmente cominciando a risolvere le equazioni cominciando "dal basso". Per esempio, se S è la seconda matrice di (4.6.3), dobbiamo risolvere il sistema di equazioni lineari omogenee

$$\begin{aligned} x_1 - 4x_2 + 5x_3 &= 0 \\ 2x_2 + 3x_3 &= 0. \end{aligned}$$

L'ultima equazione dà $x_2 = -3x_3/2$, e sostituendo nella prima equazione, otteniamo $x_1 = -11x_3$. Quindi il sottospazio delle soluzioni è

$$\{(22t, 3t, -2t)\}.$$

Questo procedimento dà la seguente formula:

$$\dim\{X \in k^n \mid S \cdot X = \mathbf{0}\} = n - |\{1 \leq i \leq m \mid S^i \neq \mathbf{0}\}|. \tag{4.6.5}$$

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 4.6.3**.

4.6.3 Operazioni elementari su liste di vettori e sulle colonne di una matrice

Definizione 4.6.9. Sia V uno spazio vettoriale su un campo 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 4.6.10. *Sia V uno spazio vettoriale su un campo 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 4.6.9**. *Allora $v_1, \ldots, v_n$ è ottenuta da $w_1, \ldots, w_n$ operando rispettivamente con (1), (2) o (3) della* **Definizione 4.6.9**.

Dimostrazione. Se $w_1, \ldots, w_n$ è ottenuta da $v_1, \ldots, v_n$ scambiando v_i con v_j allora (ri)scambiando w_i con w_j otteniamo $v_1, \ldots, v_n$. Ora supponiamo che $w_1, \ldots, w_n$ sia ottenuta da $v_1, \ldots, v_n$ operando con (2) della **Definizione 4.6.9**. 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 $v_1, \ldots, v_n$ è ottenuta da $w_1, \ldots, w_n$ operando con (2) della **Definizione 4.6.9**, dove λ è sostituito da $-\lambda$. Se $w_1, \ldots, w_n$ è ottenuta da $v_1, \ldots, v_n$ moltiplicando v_i per $0 \neq \lambda$ (e lasciando invariati gli altri vettori) allora $v_1, \ldots, v_n$ è ottenuta da $w_1, \ldots, w_n$ moltiplicando w_i per λ^{-1} e lasciando invariati gli altri vettori. □

Proposizione 4.6.11. *Sia V uno spazio vettoriale su un campo 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 4.6.9**. *Allora*

$$\langle v_1, \dots, v_n \rangle = \langle w_1, \dots, w_n \rangle. \tag{4.6.6}$$

Dimostrazione. L'operazione (1) scambia l'ordine dei vettori senza cambiare l'insieme dei vettori e quindi vale (4.6.6). Ora supponiamo che $w_1, \dots, w_n$ sia ottenuta da $v_1, \dots, v_n$ operando con (2) della **Definizione 4.6.9**. 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 4.6.10** la lista $v_1, \dots, v_n$ è ottenuta da $w_1, \dots, w_n$ operando con (2) della **Definizione 4.6.9**: per quanto abbiamo appena osservato segue che $\langle v_1, \dots, v_n \rangle \subset \langle w_1, \dots, w_n \rangle$. Quindi vale (4.6.6). Se $w_1, \dots, w_n$ è ottenuta da $v_1, \dots, v_n$ operando con (3) della **Definizione 4.6.9** è chiaro che vale (4.6.6). □

Sia $A \in M_{m,n}(k)$. Le colonne di A formano una lista di vettori di k^m. Se operiamo sulle colonne di A con una delle operazioni della**Definizione 4.6.9** otteniamo altri n vettori di k^m che sono le colonne di un'altra matrice $m \times n$. Questa è una *operazione elementare sulle colonne* di A.

Proposizione 4.6.12. *Sia $A \in M_{m,n}(k)$. 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 e quindi esiste $1 \leq j \leq n$ tale che $q_A(j) \neq \infty$. Sia $1 \leq j_0 \leq n$ tale che $q_A(j_0) = \min\{q_A(1), q_A(2), \dots, q_A(n)\}$. Siccome $A \neq 0_{m,n}$ abbiamo che $q_A(j_0) < \infty$. Scambiando la prima colonna con la colonna j_0 (operazione elementare sulle colonne - di tipo (1)) passiamo ad una matrice A_1 tale che $q_{A_1}(1) = \min\{q_{A_1}(1), q_{A_1}(2), \dots, q_{A_1}(n)\}$. Sia $s = q_{A_1}(1)$: quindi $a_{s,1} \neq 0$. Abbiamo che $q_{A_1}(1) \leq q_{A_1}(j)$ per $1 \leq j \leq n$. Supponiamo che $s = q_{A_1}(1) = q_{A_1}(j_0)$ per $1 < j_0 \leq n$. Sostituiamo alla colonna A_{j_0} la colonna $A_{j_0} - a_{s,j_0} a_{s,1}^{-1} A_1$: questa è una operazione elementare sulle colonne (di tipo (2)). La matrice A_2 che otteniamo ha la proprietà che $s = q_{A_2}(1) < q_{A_2}(j_0)$ e siccome le colonne con indice diverso da j_0 non sono cambiate $q_{A_2}(1) \leq q_{A_2}(j)$ per $1 \leq j \leq n$. Procedendo in modo simile con operazioni elementari sulle colonne (di tipo (2)) arriviamo a una matrice B tale che $s = q_B(1)$ e $q_B(1) < q_B(j)$ per ogni $1 < j \leq n$. Sia C la matrice $m \times (n-1)$ ottenuta eliminando la prima colonna di B. Per ipotesi induttiva esiste una serie di operazioni elementari sulle colonne di C il cui risultato è una matrice ($m \times (n-1)$ a scala per colonne: vedendole come operazioni elementari su B otteniamo la desiderata matrice a scala per colonne. L'affermazione riguardante la base di $\mathrm{im}(L_A)$ e il rango di A segue dall'**Osservazione 4.6.7** e dalla **Proposizione 4.6.11**. □

Osservazione 4.6.13. Sia V uno spazio vettoriale finitamente generato (su un campo k) di dimensione m. Siano $v_1, \dots, v_n \in V$. Sia $\mathcal{B}$ una base di V e sia A la matrice $m \times n$ la cui colonna j-esima è la colonna delle coordinate di v_j nella base $\mathcal{B}$. Riduciamo A a scala per colonne: sia S la matrice a scala ottenuta. Siano $w_1, \dots, w_r \in V$ i vettori le cui coordinate

sono le colonne **non nulle** di A. Per la **Proposizione 4.6.11** abbiamo che $\{w_1, \ldots, w_r\}$ è una base di $\langle v_1, \ldots, v_n \rangle$. In altre parole per trovare una base di $\langle v_1, \ldots, v_n \rangle$ consideriamo l'applicazione lineare $k^n \to V$ definita da $f(X) := \sum_{j=1}^n x_j v_j$ e troviamo una base di $\operatorname{im} f$ con la riduzione a scala per colonne di $M_{\mathcal{B}}^{\mathcal{S}}(f)$ dove $\mathcal{S}$ è la base standard di k^n.

4.6.4 Operazioni elementari su liste di funzioni lineari e sulle righe di una matrice

Sia V uno spazio vettoriale su un campo. Ricordiamo che $V^* := \mathcal{L}(V, k)$, vedi **Definizione 4.1.23**. Se V è finitamente generato, allora V^* è finitamente generato, di dimensione uguale a quella di V.

Proposizione 4.6.14. *Sia V uno spazio vettoriale su un campo k e $f_1, \ldots, f_m \in V^*$. Supponiamo che $g_1, \ldots, g_m \in V^*$ siano ottenuti da $f_1, \ldots, f_m$ operando con una delle operazioni della* **Definizione 4.6.9**. *Allora*

$$\{v \in V \mid 0 = f_1(v) = \ldots = f_m(v)\} = \{v \in V \mid 0 = g_1(v) = \ldots = g_m(v)\}. \tag{4.6.7}$$

Dimostrazione. Osserviamo che

$$\{v \in V \mid 0 = f_1(v) = \ldots = f_m(v)\} = \{v \in V \mid f(v) = 0 \quad \forall f \in \langle f_1, \ldots, f_m \rangle\}. \tag{4.6.8}$$

Infatti, supponiamo che v appartenga al membro di sinistra di (4.6.8); se $f \in \langle f_1, \ldots, f_m \rangle$, allora esistono $\lambda_1, \ldots, \lambda_m \in k$ tali che $f = \lambda_1 f_1 + \ldots + \lambda_m f_m$, e quindi

$$f(v) = \lambda_1 f_1(v) + \ldots + \lambda_m f_m(v) = 0.$$

Questo dimostra che il membro di sinistra di (4.6.8) è contenuto nel membro di destra di (4.6.8). D'altra parte, il membro di destra di (4.6.8) è contenuto nel membro di sinistra di (4.6.8) perché $f_i \in \langle f_1, \ldots, f_m \rangle$ per ogni $i \in \{1, \ldots, m\}$. La proposizione segue dall'uguaglianza (4.6.8), e dalla **Proposizione 4.6.11** applicata al sottospazio di V^* generato da $f_1, \ldots, f_m$. □

Sia $A \in M_{m,n}(k)$. Le righe di A formano una lista di vettori di k^n. Se operiamo sulle righe di A con (1) o (2) della **Definizione 4.6.9** otteniamo altri m vettori di k^n che sono le righe di un'altra matrice $m \times n$. Questa è una *operazione elementare sulle righe* di A.

Proposizione 4.6.15. *Sia $A \in M_{m,n}(k)$. Esiste una serie di operazioni elementari di tipo (1) e di tipo (2) sulle righe di A il cui risultato finale è una matrice a scala per righe S. Si ha l'uguaglianza*

$$\{X \in k^n \mid A \cdot X = \mathbf{0}\} = \{X \in k^n \mid S \cdot X = \mathbf{0}\}, \tag{4.6.9}$$

e la dimensione del nucleo di L_A è uguale alla differenza tra n e il numero di righe non nulle di S.

Dimostrazione. La prima parte segue dalla **Proposizione 4.6.12** applicata alla trasposta A^t. Per dimostrare (4.6.9) definiamo, per $i \in \{1, \ldots, m\}$, le applicazioni $f_i, g_i \in k^n \to k$ così:

$$\begin{array}{ccc} k^n & \xrightarrow{f_i} & k \\ X & \mapsto & A^i \cdot X \end{array} \qquad \begin{array}{ccc} k^n & \xrightarrow{g_i} & k \\ X & \mapsto & S^i \cdot X, \end{array}$$

dove A^i, S^i sono le righe i-esime di A e S rispettivamente, e $X \in k^n$ è visto come matrice (colonna) $n \times 1$. Per ipotesi, la lista $g_1, \ldots, g_m \in V^*$ è ottenuta da $f_1, \ldots, f_m$ operando con una serie di operazioni della **Definizione 4.6.9**, e quindi (4.6.9) segue dalla **Proposizione 4.6.14**. L'ultima affermazione segue da (4.6.9) e da (4.6.5). □

Terminologia 4.6.16. L'algoritmo descritto nella **Proposizione 4.6.15** va sotto il nome di *procedimento di eliminazione di Gauss*.

4.7 Il duale di uno spazio vettoriale

4.7.1 Duale e biduale

Ricordiamo che il duale di uno spazio vettoriale V su k è lo spazio vettoriale $\mathcal{L}(V,k)$ delle applicazioni lineari $f\colon V \to k$. Supponiamo che V sia finitamente generato e sia $n := \dim V$; per la **Proposizione 4.5.7** abbiamo un isomorfismo $\mathcal{L}(V,k) \cong M_{1,n}(k)$ e quindi $\dim V^* = n = \dim V$. Sia $\mathcal{B} := \{v_1, \ldots, v_n\}$ una base di V. Possiamo definire una base di V^* procedendo come segue. Sia $v_i^* \in V^*$ la funzione lineare

$$\begin{array}{ccc} V & \xrightarrow{v_i^*} & k \\ (x_1v_1 + x_2v_2 + \ldots + x_nv_n) & \mapsto & x_i. \end{array} \tag{4.7.1}$$

In altre parole v_i^* è l'unica applicazione lineare $V \to k$ tale che

$$v_i^*(v_j) = \delta_{ij}, \qquad 1 \leq i,j \leq n \tag{4.7.2}$$

dove δ_{ij} è il simbolo di Kronecker, vedi (4.4.6).

Proposizione 4.7.1. *Sia V uno spazio vettoriale finitamente generato su k e $\mathcal{B} := \{v_1, \ldots, v_n\}$ una sua base. Allora $\mathcal{B}^* := \{v_1^*, \ldots, v_n^*\}$ è una base di V^*.*

Dimostrazione. Per la **Proposizione 4.5.7** abbiamo un isomorfismo

$$\begin{array}{ccc} V^* & \xrightarrow{\Phi} & M_{1,n}(k) \\ f & \mapsto & M_{\mathcal{S}}^{\mathcal{B}}(f) \end{array} \tag{4.7.3}$$

dove $\mathcal{S} = \{1\}$ è la base standard di k. Sia $\{e_1, \ldots, e_n\}$ la base standard di $k^n = M_{1,n}(k)$; si ha che $\Phi^{-1}(e_i) = v_i^*$. Per il **Corollario 4.2.7** segue che $\{v_1^*, \ldots, v_n^*\}$ è una base di V^*. □

Terminologia 4.7.2. La base $\mathcal{B}^* := \{v_1^*, \ldots, v_n^*\}$ è la *base duale* della base $\mathcal{B}$.

Osservazione 4.7.3. La notazione per la base duale della base $\mathcal{B}$ è *ingannevole*, perché suggerisce che abbia senso v_i^* indipendentemente dalla scelta della base di cui v_i fa parte. Una notazione corretta sarebbe $(v_i^{\mathcal{B}})^*$; per non appesantire la notazione dimentichiamo $\mathcal{B}$.

Se V è uno spazio vettoriale finitamente generato su k, allora $V \cong V^*$ (perché hanno la stessa dimensione), ma non esiste un modo canonico di dare un isomorfismo tra V e V^*. D'altra parte, se V^{**} denota il biduale di V, cioè

$$V^{**} := (V^*)^*,$$

esiste un isomorfismo canonico tra V e V^{**}. Per vederlo, osserviamo che, se $v \in V$, allora l'applicazione

$$\begin{array}{ccc} V^* & \xrightarrow{\Phi(v)} & k \\ f & \mapsto & f(v) \end{array} \tag{4.7.4}$$

è lineare. Quindi possiamo definire un'applicazione

$$\begin{array}{ccc} V & \xrightarrow{\Phi} & V^{**} \\ v & \mapsto & \Phi(v) \end{array} \tag{4.7.5}$$

Ricapitolando, se $v \in V$, e $f \in V^*$,

$$\Phi(v)(f) = f(v). \tag{4.7.6}$$

Proposizione 4.7.4. *Se V è uno spazio vettoriale finitamente generato su k, allora l'applicazione definita da* (4.7.5) *è un isomorfismo di spazi vettoriali.*

Dimostrazione. Siano $\lambda_1, \lambda_2 \in k$, e $v_1, v_2 \in V$. Allora (ricordate (4.7.6))

$$\Phi(\lambda_1 v_1 + \lambda_2 v_2)(f) = f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2) = \lambda_1 \Phi(v_1)(f) + \lambda_2 \Phi(v_2)(f).$$

Quindi $\Phi(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 \Phi(v_1) + \lambda_2 \Phi(v_2)$, e perciò Φ è lineare. Per dimostrare che Φ è un isomorfismo, è sufficiente dimostrare che Φ è iniettiva, giacchè V e V^{**} hanno la stessa dimensione. Sia $0 \neq v \in V$, e dimostriamo che $\Phi(v) \neq 0$, cioè che esiste $f \in V^*$ 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 $v_1^*(v_1) = 1 \neq 0$. □

4.7.2 Applicazione duale di un'applicazione lineare

Definizione 4.7.5. Siano V, W spazi vettoriali su un campo k e $\phi\colon V \to W$ un'applicazione lineare. Se $f \in W^*$, allora la composizione $f \circ \phi$ è lineare, e quindi ha senso porre

$$\begin{array}{ccc} W^* & \xrightarrow{\phi^*} & V^* \\ f & \mapsto & f \circ \phi \end{array}$$

La ϕ^* è l'applicazione *duale* di ϕ.

Proposizione 4.7.6. *Siano V, W spazi vettoriali su un campo k e $\phi\colon V \to W$ un'applicazione lineare. L'applicazione duale $\phi^*\colon W^* \to V^*$ è lineare.*

Dimostrazione. Supponiamo che $\lambda_1, \lambda_2 \in k$ e $f_1, f_2 \in W^*$. Allora

$$\phi^*(\lambda_1 f_1 + \lambda_2 f_2)(v) = (\lambda_1 f_1 + \lambda_2 f_2)(\phi(v)) = \lambda_1 f_1(\phi(v)) + \lambda_2 f_2(\phi(v)) = \lambda_1 \phi^*(f_1)(v) + \lambda_2 \phi^*(f_2)(v).$$

□

Esempio 4.7.7. Sia $U \subset V$ un sottospazio vettoriale, e $\iota\colon U \hookrightarrow V$ l'inclusione. Ovviamente ι è lineare. La trasposta di ι è l'applicazione

$$\begin{array}{ccc} V^* & \xrightarrow{\iota^*} & U^* \\ f & \mapsto & f|_U. \end{array} \tag{4.7.7}$$

Verifichiamo che ι^* è suriettiva. Infatti sia $\mathcal{B} := \{u_1, \dots, u_a\}$ una base di U, ed estendiamola a una base $\mathcal{C} := \{u_1, \dots, u_a, w_1, \dots, w_b\}$ di V. Sia $\mathcal{C}^* := \{u_1^*, \dots, u_a^*, w_1^*, \dots, w_b^*\}$ la base duale di $\mathcal{C}$. Le restrizioni $u_1^*|_U, \dots, u_a^*|_U$ danno la base duale $\mathcal{B}^*$ di $\mathcal{B}$, e quindi ι^* è suriettiva. In particolare, se $0 \neq v \in V$, esiste $f \in V^*$ tale che $f(v) \neq 0$: basta considerare $U = \langle v \rangle$.

Proposizione 4.7.8. *L'applicazione duale dell'identità $\mathrm{Id}_V\colon V \to V$ è l'identità $\mathrm{Id}_{V^*}\colon V^* \to V^*$. Supponiamo che U, V, W siano spazi vettoriali su k, e che*

$$U \xrightarrow{\psi} V \xrightarrow{\phi} W$$

siano applicazioni lineari. Allora

$$(\phi \circ \psi)^* = \psi^* \circ \phi^*. \tag{4.7.8}$$

Dimostrazione. La prima affermazione è banalmente vera. Ora dimostriamo (4.7.8). Sia $f \in W^*$; allora

$$(\phi \circ \psi)^*(f) = f \circ (\phi \circ \psi) = (f \circ \phi) \circ \psi = \psi^*(\phi^*(f)) = (\psi^* \circ \phi^*)(f).$$

□

Corollario 4.7.9. *Siano V, W spazi vettoriali su un campo k e $\phi\colon V \to W$ un isomorfismo. Allora l'applicazione duale $\phi^*\colon W^* \to V^*$ è un isomorfismo.*

Dimostrazione. Dimostriamo che $(\phi^{-1})^*$ è un'inversa di ϕ^*. Per la **Proposizione 4.7.8**,

$$\phi^* \circ (\phi^{-1})^* = (\phi^{-1} \circ \phi)^* = \mathrm{Id}_V^* = \mathrm{Id}_{V^*}, \qquad (\phi^{-1})^* \circ \phi^* = (\phi \circ \phi^{-1})^* = \mathrm{Id}_W^* = \mathrm{Id}_{W^*}.$$

□

Proposizione 4.7.10. *Siano V, W spazi vettoriali finitamente generati su un campo k e $\phi\colon V \to W$ un'applicazione lineare.*

1. *La ϕ è iniettiva se e solo se ϕ^* è suriettiva.*

2. *La ϕ è suriettiva se e solo se ϕ^* è iniettiva.*

Dimostrazione. Sia $U := \operatorname{im}\phi$. L'applicazione ϕ definisce un'applicazione lineare $\psi\colon V \to U$ (data da $\psi(v) := \phi(v)$). Sia $\iota\colon U \hookrightarrow W$ l'inclusione. Allora $\phi = \iota \circ \psi$, e quindi

$$\phi^* = \psi^* \circ \iota^* \tag{4.7.9}$$

per la **Proposizione 4.7.8**. Ora supponiamo che ϕ sia iniettiva. Allora ψ è un isomorfismo, e quindi anche ψ^* è un isomorfismo per il **Corollario 4.7.9**. D'altra parte ι^* è suriettiva per l'**Esempio 4.7.7**, e ne segue che ϕ^* è suriettiva. Ora supponiamo che ϕ non sia iniettiva, e sia $0 \neq v \in \ker\phi$. Allora $f(v) = 0$ per ogni $f \in \operatorname{im}\phi^*$. D'altra parte, per l'**Esempio 4.7.7**, esiste $f \in V^*$ tale che $f(v) \neq 0$, e quindi ϕ^* non è suriettiva. Abbiamo dimostrato (1). La dimostrazione di (2) è simile, lasciamo i dettagli al lettore. □

Il seguente risultato sull'applicazione duale di un'applicazione lineare darà una conseguenza non banale sulle matrici (vedi **Corollario 4.7.13**).

Proposizione 4.7.11. *Sia $\phi\colon V \to W$ un'applicazione lineare tra spazi vettoriali finitamente generati su un campo k. Allora* $\operatorname{rk}\phi = \operatorname{rk}\phi^*$.

Dimostrazione. Sia $U := \operatorname{im}\phi$, e siano $\psi\colon V \to U$ e $\iota\colon U \hookrightarrow W$ come nella dimostrazione della **Proposizione 4.7.10**, in particolare vale (4.7.9). Ora ι^* è suriettiva per l'**Esempio 4.7.7**, e ne segue che $\operatorname{im}(\phi^*) = \operatorname{im}(\psi^*)$. D'altra parte ψ^* è iniettiva per la **Proposizione 4.7.10**, e quindi

$$\operatorname{rk}\phi^* = \dim\operatorname{im}(\phi^*) = \dim\operatorname{im}(\psi^*) = \dim U^* = \dim U = \operatorname{rk}\phi.$$

□

4.7.3 Rango di una matrice e della sua trasposta

Supponiamo che V e W siano finitamente generati e siano

$$\mathcal{B} = \{v_1, \dots, v_n\}, \qquad \mathcal{C} = \{w_1, \dots, w_m\} \tag{4.7.10}$$

basi di V e W rispettivamente. Sia $\phi\colon V \to W$ un'applicazione lineare: allora abbiamo la matrice associata $M_{\mathcal{C}}^{\mathcal{B}}(\phi) \in M_{m,n}(k)$. Abbiamo anche le basi $\mathcal{C}^*$ di W^* e $\mathcal{B}^*$ di V^* e quindi la matrice associata $M_{\mathcal{B}^*}^{\mathcal{C}^*}(\phi^*) \in M_{n,m}(k)$.

Proposizione 4.7.12. *Siano V, W spazi vettoriali finitamente generati su un campo k e $\phi\colon V \to W$ un'applicazione lineare. Siano $\mathcal{B}$ e $\mathcal{C}$ basi di V e W rispettivamente. Allora*

$$M_{\mathcal{B}^*}^{\mathcal{C}^*}(\phi^*) = M_{\mathcal{C}}^{\mathcal{B}}(\phi)^t,$$

cioè $M_{\mathcal{B}^*}^{\mathcal{C}^*}(\phi^*)$ è la trasposta di $M_{\mathcal{C}}^{\mathcal{B}}(\phi)$.

Dimostrazione. Possiamo supporre che $\mathcal{B}$ e $\mathcal{C}$ siano dati da (4.7.10). Sia $M_{\mathcal{C}}^{\mathcal{B}}(\phi) = A = (a_{ij})$. 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 $v_s^*(v) = x_s$. Abbiamo che

$$\phi^*(w_i^*)(v) = w_i^*(\phi(v)) = \sum_{s=1}^n a_{is} x_s = \sum_{s=1}^n a_{is} v_s^*(v).$$

Quindi

$$\phi^*(w_i^*) = \sum_{s=1}^n a_{is} v_s^*. \tag{4.7.11}$$

D'altra parte la colonna i-esima di $M_{\mathcal{B}^*}^{\mathcal{C}^*}(\phi^*)$ è data dalle coordinate di $\phi^*(w_i^*)$ nella base $\mathcal{B}^*$ e perciò la (4.7.11) dà che colonna i-esima di $M_{\mathcal{B}^*}^{\mathcal{C}^*}(\phi^*)$ è la riga i-esima di $M_{\mathcal{C}}^{\mathcal{B}}(\phi)$. □

Corollario 4.7.13. *Sia k un campo e $A \in M_{m,n}(k)$. Il rango di A^t è uguale al rango di A.*

Dimostrazione. Sia $\phi := L_A$. Siano $\mathcal{B}$ e $\mathcal{C}$ le basi standard di k^n e k^m rispettivamente. Per la **Proposizione 4.7.12** abbiamo che $A^t = M_{\mathcal{B}^*}^{\mathcal{C}^*}(L_A^t)$. Per la **Proposizione 4.7.11** segue che $\operatorname{rk} A^t = \operatorname{rk} A$. □

Il **Corollario 4.7.13** equivale alla seguente affermazione: se $A \in M_{m,n}(k)$ allora il sottospazio di k^n generato dalle *righe* di A ha la stessa dimensione del sottospazio di k^m generato dalle *colonne* di A. Notiamo che l'affermazione non è affatto banale. Possiamo anche dare la seguente versione del **Corollario 4.7.13**: 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).

4.8 Operazioni elementari sulle matrici, II

Siano V, W spazi vettoriali su un campo k. Supponiamo che V, W siano finitamente generati. Sia $f\colon V \to W$ un'applicazione lineare.

Problema 4.8.1. Dare un algoritmo efficiente per descrivere la controimmagine di $w \in W$, in particolare descrivere esplicitamente f^{-1} se f è un isomorfismo.

Scegliamo una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ di V, una base $\mathcal{C} = \{w_1, \ldots, w_m\}$ di W e associamo a f la matrice

$$A := M_{\mathcal{C}}^{\mathcal{B}}(f) \in M_{m,n}(k). \tag{4.8.1}$$

Siano $w \in W$ e $v \in V$: per (4.5.3) abbiamo che $v \in f^{-1}(w)$ (ovvero $f(v) = w$) se e solo

$$A \cdot X_{\mathcal{B}}(v) = X_{\mathcal{C}}(w).$$

Quindi il **Problema 4.8.1** equivale a risolvere il sistema di equazione lineari

$$A \cdot X = B \tag{4.8.2}$$

dove X è una matrice colonna (di incognite) $n \times 1$ e B è una matrice colonna $m \times 1$ (di termini noti). Notate che se $B = \mathbf{0}$ il sistema (4.8.2) si risolve seguendo l'algoritmo descritto nella **Proposizione 4.6.15**; in generale si può procedere come segue. Sia $A = (a_{ij})$ e $B = (b_i)$. Consideriamo la matrice $m \times (n+1)$

$$[A|B] = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \ldots & a_{in} & b_i \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{bmatrix} \tag{4.8.3}$$

Con una serie di operazioni elementari sulle righe di $[A|B]$ possiamo arrivare a una matrice $[S|C]$ dove S è una matrice a **scala per righe** $m \times n$ e C è una matrice $m \times 1$ cioè una colonna di lunghezza m.

Proposizione 4.8.2. *Siano $[A|B]$ e $[S|C]$ come sopra. Allora abbiamo l'uguaglianza di sottospazi di k^n (scriviamo i vettori di k^n come matrici $n \times 1$)*

$$\{X \in k^n \mid A \cdot X = B\} = \{X \in k^n \mid S \cdot X = C\}. \tag{4.8.4}$$

Dimostrazione. Ragioniamo come nella dimostrazione della **Proposizione 4.6.14**. Supponiamo che con una operazione elementare sulle righe di $[A|B]$ abbiamo otteniamo $[M|R]$ dove M è una matrice $m \times n$ e R è una matrice $m \times 1$: dimostreremo che

$$\{X \in k^n \mid A \cdot X = B\} = \{X \in k^n \mid M \cdot X = R\}. \tag{4.8.5}$$

La proposizione seguirà da questo risultato. Il risultato è del tutto ovvio se l'operazione è di tipo (1) o (3). Ora supponiamo che $[M|R]$ sia ottenuta da $[A|B]$ sostituendo la riga i-esima $(a_{i1}, \ldots, a_{in}, b_i)$ con la riga $(a_{i1} + \lambda a_{j1}, \ldots, a_{in} + \lambda a_{jn}, b_i + \lambda b_j)$ dove $j \neq i$. Dimostriamo che il membro di sinistra di (4.8.5) è contenuto nel membro di destra di (4.8.5). 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

$$(A^i + \lambda A^j) \cdot X = (b_i + \lambda b_j). \tag{4.8.6}$$

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 (4.8.6). Questo dimostra che il membro di sinistra di (4.8.5) è contenuto nel membro di destra di (4.8.5). Rimane da dimostrare che il membro di destra è contenuto nel membro di sinistra. Per il **Lemma 4.6.10** sappiamo che sia $[A|B]$ è ottenuta da $[M|R]$ con una operazione elemntare di tipo (2) e quindi per quello che abbiamo dimostrato il membro di destra è contenuto nel membro di sinistra. □

Ora supponiamo che S sia una matrice a **scala per righe** $m \times n$ e C una matrice $m \times 1$: le soluzioni del sistema di equazioni lineari $S \cdot X = C$ si trovano facilmente cominciando a risolvere le equazioni "dal basso". In particolare abbiamo la seguente

Osservazione 4.8.3. Sia S sia una matrice a **scala per righe** $m \times n$ e C una matrice $m \times 1$: il sistema di equazioni lineari $S \cdot X = C$ ha una soluzione se e solo se $c_i = 0$ per ogni indice $1 \leq i \leq m$ tale che la riga S^i sia nulla.

Terminologia 4.8.4. L'algoritmo appena descritto si chiama *procedimento di eliminazione di Gauss.*

Avendo risolto il **Problema 4.8.1** ci poniamo un altra domanda.

Problema 4.8.5. Data una matrice quadrata *invertibile* $A \in M_{n,n}(k)$ come possiamo calcolare in modo efficiente l'inversa di A ?

Descriviamo un algoritmo che produce A^{-1}, poi spiegheremo perché è corretto. 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 $[S|C]$ dove S è $n \times n$ a scala per righe e C è una matrice $n \times n$. Per ipotesi Aè invertibile quindi $\ker(L_S) = \{\mathbf{0}\}$ ovvero tutte le righe di S sono non nulle: segue che le entrate di S sulla diagonale principale di S sono tutte 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 operiamo sulle righe di $[S'|C']$ cominciando "dal basso"e arriveremo a una matrice $[1_n|D]$.

Proposizione 4.8.6. *Sia $A \in M_{n,n})(k)$ invertibile. Sia $D \in M_{n,n}(k)$ la matrice ottenuta a partire da A con il procedimento descritto sopra. Allora $D = A^{-1}$.*

Dimostrazione. Sia $Y \in k^n$ vista come matrice $n \times 1$. Sia $X = A^{-1} \cdot Y$: allora

$$A \cdot X = A \cdot (A^{-1} \cdot Y) = (A \cdot (A^{-1}) \cdot Y = 1_n \cdot Y = Y.$$

Quindi X è una soluzione del sistema di equazioni lineari

$$A \cdot X = Y. \tag{4.8.7}$$

Notiamo che (4.8.7) ha una unica soluzione perché L_A è biunivoca. Applichiamo l'algoritmo per risolvere un sistema di equazioni lineari a (4.8.7) con colonna Y. Allora avremo che la soluzione di (4.8.7) è la soluzione di $S \cdot X = C \cdot Y$ dove S e C sono le matrici descritte sopra (pensate che la riga i-esima di 1_n cioè l'i-esimo vettore della base standard rappresenti y_i). Con la moltiplicazione per gli inversi delle entrate sulla diagonal principale di S troviamo che X è la soluzione di $S' \cdot X = C' \cdot Y$. Infine con le operazioni sulle righe "dal basso"troviamo che X è la soluzione di $1_n \cdot X = D \cdot Y$ cioè $X = D \cdot Y$ - quindi $A^{-1} = D$. □

Esempio 4.8.7. 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 !)

4.9 Cambiamenti di base e coniugio

Sia V uno spazio vettoriale su k, finitamente generato e di dimensione n. Siano

$$\mathcal{B} = \{u_1, \ldots, u_n\}, \qquad \mathcal{C} = \{w_1, \ldots, w_n\}$$

basi di V. Ci interessa sapere quale relazione esiste tra le coordinate di un vettore nella base $\mathcal{B}$ e nella base $\mathcal{C}$. Abbiamo gli isomorfismi $X_{\mathcal{B}} \colon V \xrightarrow{\sim} k^n$ e $X_{\mathcal{C}} \colon V \xrightarrow{\sim} k^n$ che associano a $v \in V$ il vettore delle coordinate di v nelle basi $\mathcal{B}$ e $\mathcal{C}$ rispettivalmente. Componendo otteniamo l'isomorfismo $X_{\mathcal{C}} \circ X_{\mathcal{B}}^{-1} \colon k^n \xrightarrow{\sim} k^n$. Per la **Proposizione 4.4.15** esiste $A \in \mathrm{GL}_n(k)$ tale che $X_{\mathcal{C}} \circ X_{\mathcal{B}}^{-1} = L_A$. Questa A è la matrice del *cambiamento di base* da $\mathcal{B}$ a $\mathcal{C}$; si ha che

$$X_{\mathcal{C}}(v) = A \cdot X_{\mathcal{B}}(v) \qquad \forall v \in V. \tag{4.9.1}$$

Poniamo $v = u_j$ nella (4.9.1): otteniamo che la colonna delle delle coordinate di u_j nella base $\mathcal{C}$ è uguale alla colonna j-esima di A (cioè A_j). In altre parole

$$\text{la matrice del cambiamento di base da } \mathcal{B} \text{ a } \mathcal{C} \text{ è } M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V). \tag{4.9.2}$$

Ora notiamo che A è invertibile perché $L_A = X_{\mathcal{C}} \circ X_{\mathcal{B}}^{-1}$ e $X_{\mathcal{C}} \circ X_{\mathcal{B}}^{-1}$ è invertibile in quanto composizione di isomorfismi. Moltiplicando ambo i membri di (4.9.1) per A^{-1} vediamo che l'inversa di A è la matrice del cambiamento di base da $\mathcal{C}$ a $\mathcal{B}$:

$$M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}_V) = M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V)^{-1}. \tag{4.9.3}$$

Osservazione 4.9.1. Sia $V = k^n$ e $\mathcal{S} = \{e_1, \ldots, e_n\}$ la base standard. Se $\mathcal{C} = \{C_1, \ldots, C_n\}$ dove le C_j sono colonne allora la colonna j-esima di $M_{\mathcal{S}}^{\mathcal{C}}(\mathrm{Id}_V)$ è uguale a C_j. Quindi

$$M_{\mathcal{C}}^{\mathcal{S}}(\mathrm{Id}_V) = [C_1, \ldots, C_n]^{-1}.$$

Osservazione 4.9.2. Sia V uno spazio vettoriale finitamente generato su k e $\mathcal{B}, \mathcal{C}, \mathcal{D}$ sue basi. Allora

$$M_{\mathcal{D}}^{\mathcal{B}}(\mathrm{Id}_V) = M_{\mathcal{D}}^{\mathcal{C}}(\mathrm{Id}_V) \cdot M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V). \tag{4.9.4}$$

Qundi 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, per esempio la base standard se $V = k^n$. In concreto: supponiamo che $V = k^n$ e $\mathcal{B} = \{B_1, \ldots, B_n\}$ e $\mathcal{D} = \{D_1, \ldots, D_n\}$ dove le B_j e D_j sono matrici $n \times 1$: abbiamo che

$$M_{\mathcal{D}}^{\mathcal{B}}(\mathrm{Id}_V) = M_{\mathcal{D}}^{\mathcal{S}}(\mathrm{Id}_V) \cdot M_{\mathcal{S}}^{\mathcal{B}}(\mathrm{Id}_V) = [D_1, \ldots, D_n]^{-1} \cdot [B_1, \ldots, B_n]. \tag{4.9.5}$$

L'equazione (4.9.3) dimostra 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 4.9.3. *Sia V uno spazio vettoriale finitamente generato e di dimensione n su un campo k. Sia $\mathcal{C} = \{u_1, \ldots, u_n\}$ una base di V e $A \in \mathrm{GL}_n(k)$. 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

$$w_j = \sum_{i=1}^{n} e_{ij} u_j.$$

Siccome le colonne di A^{-1} sono linearmente indipendenti $\mathcal{B} := \{w_1, \ldots, w_n\}$ è una base di V. Abbiamo che

$$M_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id}_V) = A^{-1}.$$

Per l'equazione (4.9.3) segue che $M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}_V) = A$. □

Definizione 4.9.4. Sia V uno spazio vettoriale.

1. Un *endomorfismo* di V è un'applicazione lineare $f\colon V \to V$.
2. $\mathrm{End}(V)$ è l'insieme degli endomorfismi di V (cioè $\mathcal{L}(V, V)$).

Sia V uno spazio vettoriale su un campo k: supponiamo che V is finitamente generato e di dimensione n. Sia $f\colon V \to V$ un endomorfismo di V. Scelta una base $\mathcal{C}$ di V possiamo associare a f la matrice $M^{\mathcal{C}}_{\mathcal{C}}(f) \in M_{n,n}(k)$. Notate che siccome dominio e codominio di f coincidono abbiamo scelto la stessa base per V visto come dominio e visto 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 dall base $\mathcal{C}$ a un'altra base $\mathcal{B}$. Per la (4.5.6) abbiamo che

$$M^{\mathcal{B}}_{\mathcal{B}}(f) = M^{\mathcal{B}}_{\mathcal{B}}(\mathrm{Id}_V \circ f \circ \mathrm{Id}_V) = M^{\mathcal{C}}_{\mathcal{B}}(\mathrm{Id}_V) \cdot M^{\mathcal{C}}_{\mathcal{C}}(f) \cdot M^{\mathcal{B}}_{\mathcal{C}}(\mathrm{Id}_V). \tag{4.9.6}$$

Definizione 4.9.5. La matrice $M \in M_{n,n}(k)$ è *coniugata* a $N \in M_{n,n}(k)$ (in simboli $M \sim N$) se esiste una $A \in \mathrm{GL}_n(k)$ tale che

$$M = A^{-1} \cdot N \cdot A. \tag{4.9.7}$$

Proposizione 4.9.6. *Sia V uno spazio vettoriale finitamente generato e di dimensione n su un campo k. Sia $f\colon V \to V$ un endomorfismo e $\mathcal{C} = \{u_1, \ldots, u_n\}$ una base di V. Data $M \in M_{n,n}(k)$ 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 (4.9.6). Ora supponiamo che M sia coniugata a $M^{\mathcal{C}}_{\mathcal{C}}(f)$ e quindi esiste $A \in M_{n,n}(k)$ invertibile tale che $M = A^{-1} \cdot M^{\mathcal{C}}_{\mathcal{C}}(f) \cdot A$. Per la **Proposizione 4.9.3** esiste una base $\mathcal{B}$ di V tale che $M^{\mathcal{C}}_{\mathcal{B}}(f) = A^{-1}$. Per l'equazione (4.9.6) segue che $M = M^{\mathcal{B}}_{\mathcal{B}}(f)$. □

Proposizione 4.9.7. *La relazione di coniugio è di equivalenza.*

Dimostrazione. $M \sim M$ perché $M = 1_n^{-1} \cdot N \cdot 1_n$. Supponiamo che $M \in M_{n,n}(k)$ sia coniugata a N e quindi che valga (4.9.7). Moltiplicando a sinistra ambo i membri di (4.9.7) per A e successivamente a destra per A^{-1} otteniamo che $A \cdot M \cdot A^{-1} = N$. Siccome $A = (A^{-1})^{-1}$ segue che N è coniugata a M. Infine supponiamo che $M \sim N$ e $N \sim P$. Quindi esistono $A, B \in M_{n,n}(k)$ invertibili tali che

$$M = A^{-1} \cdot N \cdot A, \qquad N = B^{-1} \cdot P \cdot B. \tag{4.9.8}$$

La matrice $B \cdot A$ è invertibile perché $L_{B \cdot A} = L_B \circ L_A$ è composizione di isomorfismi e quindi è un isomorfismo. Sostituendo l'espressione di N nella prima equazione di (4.9.8) otteniamo che

$$M = A^{-1} \cdot B^{-1} \cdot P \cdot B \cdot A = (B \cdot A)^{-1} \cdot P \cdot (B \cdot A).$$

Questo dimostra che M è coniugata a P. □

Esercizi del Capitolo 4

Esercizio 4.1. *Sia k un campo. Quali delle seguenti applicazioni tra spazi vettoriali su k è lineare ?*

(1) Sia $p \in k[x]$ e $\Phi\colon k[x] \to k[x]$ definita da

$$\Phi(q) := p \cdot q.$$

(2) Sia $\Psi\colon k[x] \to k[x]$ definita da

$$\Psi(p) := p^2.$$

(Attenzione: la risposta dipende dal campo k.)

(3) Sia $\Theta\colon k[x] \to k[x]$ *definita così: dato* $p \in k[x]$ *poniamo*

$$\Theta(p) := p(x^2).$$

(4) Sia $F\colon k[x] \to k$ *definita da*

$$F(p) := p(0) + p(1).$$

Esercizio 4.2. *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 **Definizione 1.8.2***). Verificate che*

(1) f **è** *un'applicazione lineare di spazi vettoriali reali.*

(2) f **non** *è un'applicazione lineare di spazi vettoriali complessi.*

Esercizio 4.3. *Sia* k *un campo e* $\alpha_0, \ldots, \alpha_n \in k$ distinti*. Dimostrate che l'applicazione*

$$\begin{array}{ccc} k[x]_{\leq n} & \longrightarrow & k^{n+1} \\ p & \mapsto & (p(\alpha_0), \ldots, p(\alpha_n)) \end{array}$$

è un isomorfismo.

Esercizio 4.4. *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 4.5. *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 4.6. *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* $\mathrm{im}(L_B)$*.*

Esercizio 4.7. *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 4.8. *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 4.9. *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 4.10. *Siano* $A \in M_{m,n}(k)$ *e* $B \in M_{n,p}(k)$. *Dimostrate che*

$$(A \cdot B)^t = B^t \cdot A^t.$$

Esercizio 4.11. *Sia* V *uno spazio vettoriale su un campo* k *in cui* $2 \neq 0$ *(e quindi esiste* 2^{-1}*). Supponiamo che* V *sia finitamente generato di dimensione* n. *Sia* $f \in \mathcal{L}(V,V)$.

(1) *Supponiamo che esista una base* $\mathcal{B} = \{v_1,\dots,v_n\}$ *di* V *tale che*

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = (\lambda_i \delta_{ij}), \qquad \lambda_i^2 = 1. \tag{4.9.9}$$

Dimostrate che $f \circ f = \mathrm{Id}_V$.

(2) *Ora supponiamo che* $f \circ f = \mathrm{Id}_V$. *Dimostrate che esiste una base* $\mathcal{B}$ *di* V *tale che valga* (4.9.9). *(Suggerimento: osservate che vale* (4.9.9) *se e solo se* $f(v_i) = \lambda_i v_i$. *Dato* $v \in V$ *calcolate* $f(v \pm f(v))$.*)*

Esercizio 4.12. *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 4.13. *Sia* k *un campo e* $V \subset k^n$ *il sottospazio*

$$V := \{X \mid x_1 + \ldots + x_n = 0\}.$$

Dare una base di V^*.

Esercizio 4.14. *Sia* k *un campo e* $\Phi, \Psi \colon k[x] \to k[x]$ *le applicazioni lineari date da*

$$\begin{array}{ccc} k[x] & \xrightarrow{\Phi} & k[x] \\ p & \mapsto & (x^2+3)\cdot p \end{array} \qquad \begin{array}{ccc} k[x] & \xrightarrow{\Psi} & k[x] \\ p(x) & \mapsto & p(-x) \end{array}$$

Siano $f, g \colon \mathbb{R}[x]^*$ *le funzioni definite da*

$$\begin{array}{ccc} k[x] & \xrightarrow{f} & k \\ q & \mapsto & q(0) \end{array} \qquad \begin{array}{ccc} k[x] & \xrightarrow{g} & k \\ q & \mapsto & q(1) \end{array}$$

Determinate

$$\Phi^* f, \quad \Phi^* g, \quad \Psi^* f, \quad \Psi^* g.$$

Esercizio 4.15. *Siano* V *uno spazio vettoriale su un campo* k, *e* $W \subset V$ *un sottospazio. L'*annullatore *di* W *è il sottoinsieme* $\operatorname{Ann} W \subset V^*$ *definito da*

$$\operatorname{Ann} W := \{\varphi \in V^* \mid \varphi|_W = 0\}. \tag{4.9.10}$$

1. *Verificate che* $\operatorname{Ann} W$ *è un sottospazio di* V^*.
2. *Sia* $\pi\colon V \to V/W$ *l'applicazione quoziente. Dimostrate che*

$$(V/W)^* \xrightarrow{\pi^*} V^*$$

definisce un'isomorfismo tra $(V/W)^*$ *e* $\operatorname{Ann} W$.

3. *Supponiamo che* V/W *sia finitamente generato, e quindi anche* $(V/W)^*$. *Siano* $\varphi_1, \dots, \varphi_d$ *generatori di* $(V/W)^*$. *Si dimostri che*

$$W = \{v \in V \mid 0 = \varphi_1(v) = \dots = \varphi_d(v)\}. \tag{4.9.11}$$

(Le $0 = \varphi_1(v) = \dots = \varphi_d(v)$ *si dicono* equazioni cartesiane *di* W.*)*

Esercizio 4.16. *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 4.17. *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 4.18. *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 4.19. *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 4.20. *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 $M_{\mathcal{B}}^{\mathcal{B}}(L_A)$.

Capitolo 5

Geometria affine, II

5.1 Applicazioni affini

Sia k un campo e $\mathbb{A}$, $\mathbb{B}$ spazi affini su k e siano rispettivamente V e W gli spazi vettoriali associati.

Definizione 5.1.1. Un'applicazione $F\colon \mathbb{A} \to \mathbb{B}$ è *affine* se esiste un'applicazione lineare $f\colon V \to W$ tale che

$$\overrightarrow{F(P)F(Q)} = f(\overrightarrow{PQ}) \quad \forall P, Q \in \mathbb{A}. \tag{5.1.1}$$

In altre parole F manda segmenti orientati equipollenti in segemnti orientati equipollenti. Dalla (5.1.1) segue che la f è univocamente determinata da F: è l'applicazione lineare *associata* a F.

Proposizione 5.1.2. *Sia $F\colon \mathbb{A} \to \mathbb{B}$ un'applicazione affine e $f\colon V \to W$ l'applicazione lineare associata. Sia $P \in \mathbb{A}$. Allora*

$$F(P + v) = F(P) + f(v), \qquad \forall P \in \mathbb{A}, \quad \forall v \in V. \tag{5.1.2}$$

Viceversa sia $F\colon \mathbb{A} \to \mathbb{B}$ un'applicazione e supponiamo che esistano un'applicazione lineare $f\colon V \to W$ e $P \in \mathbb{A}$ tali che valga (5.1.2)*: allora F è affine con applicazione lineare associata $f\colon V \to W$.*

Dimostrazione. Supponiamo che $F\colon \mathbb{A} \to \mathbb{B}$ sia affine con applicazione lineare associata $f\colon V \to W$. Poniamo $Q = P + v$, cioè $v = \overrightarrow{PQ}$: allora

$$F(Q) = F(P) + \overrightarrow{F(P)F(Q)} = F(P) + f(\overrightarrow{PQ}) = F(P) + f(v).$$

Ora supponiamo che valga (5.1.2) con f lineare. Siano $Q_0, Q_1 \in \mathbb{A}$ e $v_i \in V$ tali che $Q_i = P + v_i$. Allora

$$\overrightarrow{F(Q_0)F(Q_1)} = \overrightarrow{(F(P) + f(v_0))(F(P) + f(v_1))} = \\ = f(v_1) - f(v_0) = f(v_1 - v_0) = f(\overrightarrow{(P + v_0)(P + v_1)}) = f(\overrightarrow{Q_0Q_1}).$$

□

Sia $F\colon \mathbb{A} \to \mathbb{B}$ un'applicazione affine, con applicazione lineare associata f. La (5.1.2) mostra che F è univocamente determinata dal suo valore in un punto e dall'applicazione

lineare associata f. Viceversa dati un'applicazione lineare $f\colon V \to W$, $P \in \mathbb{A}$ e $Q \in \mathbb{B}$ l'applicazione

$$\begin{array}{ccc} \mathbb{A} & \longrightarrow & \mathbb{B} \\ P+v & \mapsto & Q+f(v) \end{array} \tag{5.1.3}$$

è affine.

Esempio 5.1.3. Sia $w \in V$ e $\tau_w\colon \mathbb{A} \to \mathbb{A}$ definita da

$$\tau_w(P) := P + w. \tag{5.1.4}$$

La τ_w è affine con applicazione lineare associata Id_V. La τ_w è una *traslazione*.

Esempio 5.1.4. I k-spazi vettoriali k^n e k^m sono anche spazi affini (su k). Sia $A \in M_{m,n}(k)$ e $B \in M_{m,1}$: l'applicazione

$$\begin{array}{ccc} k^n & \longrightarrow & k^m \\ X & \mapsto & A \cdot X + B \end{array} \tag{5.1.5}$$

è affine. L'applicazione lineare associata è L_A. Viceversa ogni applicazione affine $F\colon k^n \to k^m$ è di questo tipo. Infatti sia $f\colon k^n \to k^m$ l'applicazione lineare associata: per la **Proposizione 4.4.15** esiste $A \in M_{m,n}(k)$ tale che $f = L_A$. Sia $B := F(\mathbf{0})$. Dato $X \in k^n$ abbiamo che

$$F(X) = F(\mathbf{0} + X) = F(\mathbf{0}) + f(X) = B + A \cdot X. \tag{5.1.6}$$

(Notate che nell'equazione (5.1.6) l'elemento neutro $\mathbf{0}$ è visto come punto dello spazio affine k^n mentre X è visto come vetttore di k^n.)

Esempio 5.1.5. Sia $\mathbb{A}$ uno spazio affine sul campo k. Una *funzione affine* $F\colon \mathbb{A} \to k$ è un'applicazione affine dove k è uno spazio affine in quanto spazio vettoriale (su k): quindi esistono un'applicazione lineare $f\colon V \to k$, $P \in \mathbb{A}$ e $b \in k$ tali che

$$F(P+v) = b + f(v), \qquad \forall v \in V. \tag{5.1.7}$$

Osservazione 5.1.6. Siano $\mathbb{A}$ e $\mathbb{B}$ spazi affini sul campo k e $F\colon \mathbb{A} \to \mathbb{B}$ un'applicazione affine. Siano V e W gli spazi vettoriali associati a $\mathbb{A}$ e $\mathbb{B}$ rispettivamente e $f\colon V \to W$ l'applicazione lineara associata a F.

1. Sia $\mathbb{D} \subset \mathbb{A}$ un sottospazio affine. Allora $F(\mathbb{D})$ è un sottospazio affine. Infatti siano $P \in \mathbb{D}$ e $\mathcal{G}(\mathbb{D})$ la giacitura di $\mathbb{D}$, vedi **Osservazione 3.6.2**. Per la **Proposizione 5.1.2** abbiamo che
$$F(\mathbb{D}) = F(P) + f(\mathcal{G}(\mathbb{D})) \tag{5.1.8}$$
e siccome $f(\mathcal{G}(\mathbb{D}))$ è un sottospazio vettoriale di W (perché f è lineare) segue che $F(\mathbb{D})$ è un sottospazio affine di $\mathbb{B}$.

2. Sia $\mathbb{E} \subset \mathbb{B}$ un sottospazio affine. La controimmagine $F^{-1}(\mathbb{E})$ o è vuota o è un sottospazio affine. Infatti supponiamo che $F^{-1}(\mathbb{E})$ non sia vuota e sia $P \in F^{-1}(\mathbb{E})$. Sia $\mathcal{G}(\mathbb{E})$ la giacitura di $\mathbb{E}$. Per la **Proposizione 5.1.2** abbiamo che
$$F^{-1}(\mathbb{E}) = P + f^{-1}(\mathcal{G}(\mathbb{E})) \tag{5.1.9}$$
e siccome $f^{-1}(\mathcal{G}(\mathbb{E}))$ è un sottospazio vettoriale di V (perché f è lineare) segue che $F^{-1}(\mathbb{E})$ è un sottospazio affine di $\mathbb{A}$.

Proposizione 5.1.7. *Siano $\mathbb{A}, \mathbb{B}$ spazi affini su un campo k e $F\colon \mathbb{A} \to \mathbb{B}$ un'applicazione affine. Siano $P_0, \dots, P_d \in \mathbb{A}$ e $\lambda_0, \dots, \lambda_d \in k$ tali che $\sum_{i=0}^{d} \lambda_i = 1$. Allora*

$$F(\sum_{i=0}^{d} \lambda_i P_i) = \sum_{i=0}^{d} \lambda_i F(P_i). \tag{5.1.10}$$

Dimostrazione. Siano $Q \in \mathbb{A}$ e f l'applicazione lineare associata a F. Abbiamo

$$F(\sum_{i=0}^{d} \lambda_i P_i) = F(Q + \sum_{i=0}^{d} \lambda_i \overrightarrow{QP_i}) = F(Q) + \sum_{i=0}^{d} \lambda_i f(\overrightarrow{QP_i}) = F(Q) + \sum_{i=0}^{d} \lambda_i \overrightarrow{F(Q)F(P_i)} = \sum_{i=0}^{d} \lambda_i F(P_i).$$

□

5.2 Composizione di applicazioni affini

Proposizione 5.2.1. *Siano $F\colon \mathbb{A} \to \mathbb{B}$ e $G\colon \mathbb{B} \to \mathbb{C}$ applicazioni affini, con applicazioni lineari associate $f\colon U \to V$ $g\colon V \to W$. Allora $G \circ F\colon \mathbb{A} \to \mathbb{C}$ è affine con applicazione lineare associata $g \circ f$.*

Dimostrazione. Siano $P \in \mathbb{A}$ e $u \in U$: allora

$$G \circ F(P + u) = G(F(P + u)) = G(F(P) + f(u))) = G(F(P)) + g(f(u)) = G(F(P)) + g \circ f(u).$$

Quindi la proposizione segue dalla **Proposizione 5.1.2**. □

Un'applicazione affine $F\colon \mathbb{A} \to \mathbb{B}$ è un *isomorfismo* se ha inversa affine $g\colon \mathbb{B} \to \mathbb{A}$.

Proposizione 5.2.2. *Un'applicazione affine $F\colon \mathbb{A} \to \mathbb{B}$ è un isomorfismo se e solo se l'applicazione lineare associata è un isomorfismo di spazi vettoriali.*

Dimostrazione. Supponiamo che $F\colon \mathbb{A} \to \mathbb{B}$ sia isomorfismo con inversa affine $g\colon \mathbb{B} \to \mathbb{A}$. Siano $f\colon V \to W$ e $g\colon W \to V$ le applicazioni lineari associate. Allora $g \circ f\colon V \to V$ e $f \circ g\colon W \to W$ sono associate a $G \circ F = \mathrm{Id}_{\mathbb{A}}$ e $F \circ G = \mathrm{Id}_{\mathbb{B}}$ rispettivamente; segue che sono entramebe l'identità e quindi f è un isomorfismo. Ora supponiamo che $F\colon \mathbb{A} \to \mathbb{B}$ sia affine con applicazioni lineari associata $f\colon V \to W$ che è un isomorfismo. Sia $F(P) = Q$. Definiamo $G\colon W \to V$ così:

$$G(Q + w) := P + f^{-1}(w).$$

Si verifica subito che G è inversa di F. □

Definizione 5.2.3. Sia $\mathbb{A}$ uno spazio affine su un campo k. Il *gruppo delle affinità* di $\mathbb{A}$ è l'insieme degli isomorfismi (affini) $F\colon \mathbb{A} \to \mathbb{A}$.

Da quanto abbiamo già osservato seguono i seguenti fatti:

1. L'identità $\mathrm{Id}\colon \mathbb{A} \to \mathbb{A}$ è un'affinità di $\mathbb{A}$.
2. Se F, G sono affinità di $\mathbb{A}$ anche la composizione $F \circ G$ loè.
3. Se F, G, H sono affinità di $\mathbb{A}$ allora $F \circ (G \circ H) = (F \circ G) \circ H$.
4. Se F è un'affinità di $\mathbb{A}$ anche l'inversa F^{-1} loè.

L'insieme degli isomorfismi (affini) $F\colon \mathbb{A} \to \mathbb{A}$ si chiama *gruppo* delle affinità perché valgono 1, 2, 3 e 4.

Esempio 5.2.4. Per l'**Esempio 5.1.4** ogni isomorfismo affine $f\colon k^n \to k^n$ si scrive come

$$\begin{array}{ccc} k^n & \xrightarrow{f} & k^n \\ X & \mapsto & A \cdot X + B \end{array} \tag{5.2.1}$$

dove $A \in M_{n,n}(k)$ e $B \in M_{n,1}$. Inoltre A deve essere invertibile per la **Proposizione 5.2.2**. Viceversa se A è invertibile la (5.2.1) definisce un isomorfismo affine. Quindi abbiamo descritto il gruppo delle affinità di k^n (e perciò anche di ogni spazio affine di dimensione finita).

5.3 Cambiamenti di coordinate affini

Sia $\mathbb{A}$ uno spazio affine con spazio vettoriale associato V, spazio vettoriale su k. Supponiamo che V sia finitamente generato e che $\mathcal{B} = \{v_1, \ldots, v_n\}$ sia una base di V. Scegliamo on'origine $O \in \mathbb{A}$. Allora abbiamo il sistema di riferimento affine $\mathsf{R} = RA(O; \mathcal{B})$. Sia

$$X_{\mathsf{R}}\colon \mathbb{A} \to k^n \tag{5.3.1}$$

l'applicazione che associa a $P \in \mathbb{A}$ l'n-pla delle sue coordinate nel sistema R. Si verifica facilmente che X_{R} è un isomorfismo di spazi affini. Ora sia $\mathsf{R}' = RA(O'; \mathcal{B}')$ un secondo sistema di riferimento affine e $X_{\mathsf{R}'}\colon \mathbb{A} \to k^n$ l'applicazione che associa a $P \in \mathbb{A}$ l'n-pla delle sue coordinate nel sistema R'. Sia $P \in \mathbb{A}$: che relazione esiste tra $X_{\mathsf{R}}(P)$ e $X_{\mathsf{R}'}(P)$? La composizione

$$X_{\mathsf{R}'} \circ X_{\mathsf{R}}^{-1}\colon k^n \longrightarrow k^n \tag{5.3.2}$$

è un isomorfismo di spazi affini e quindi per l'**Esempio 5.1.4** esistono $A \in M_{n,n}(k)$ invertibile e $B \in M_{n,1}(k)$ tali che

$$X_{\mathsf{R}'} \circ X_{\mathsf{R}}^{-1}(Y) = A \cdot Y + B, \qquad \forall\, Y \in k^n. \tag{5.3.3}$$

Sia $P = X_{\mathsf{R}}^{-1}(Y)$ cioè $Y = X_{\mathsf{R}}(P)$: possiamo riscrivere la (5.3.3) come

$$X_{\mathsf{R}'}(P) = A \cdot X_{\mathsf{R}}(P) + B, \qquad \forall P \in \mathbb{A}. \tag{5.3.4}$$

5.4 Equazioni cartesiane

Siano $\mathbb{A}$ uno spazio affine su k e $\mathbb{B} \subset \mathbb{A}$ un sottospazio affine, vedi **Definizione 3.6.1**. Siano $F_1, \ldots, F_d\colon \mathbb{A} \to k$ funzioni affini: diciamo che $0 = F_1 = \ldots = F_d$ sono *equazioni cartesiane* di $\mathbb{B}$ se

$$\mathbb{B} = \{P \in \mathbb{A} \mid 0 = F_1(P) = \ldots = F_d(P)\}. \tag{5.4.1}$$

Proposizione 5.4.1. *Sia $\mathbb{A}$ uno spazio affine su k con spazio vettoriale associato V finitamente generato. Se $\mathbb{B} \subset \mathbb{A}$ è un sottospazio affine allora esistono equazioni cartesiane di $\mathbb{B}$.*

Dimostrazione. Per definizione esistono $P \in \mathbb{A}$ e un sottospazio vettoriale $W \subset V$ tali che

$$\mathbb{B} = \{P + w \mid w \in W\}. \tag{5.4.2}$$

Per l'**Esercizio 4.15**, esistono applicazioni lineari $g_1, \dots, g_d \colon V \to k$ tali che

$$W = \{v \in V \mid 0 = g_1(v) = \dots = g_d(v)\}. \tag{5.4.3}$$

Siano $F_1, \dots, F_d \colon \mathbb{A} \to k$ definite così:

$$F_i(P + v) := g_i(v). \tag{5.4.4}$$

□

Osserviamo che la (5.4.4) si può scrivere

$$F_i(Q) := g_i(\overline{PQ}).$$

Esempio 5.4.2. Siano $\mathbb{A} = \mathbb{R}^4$ e $\mathbb{B} = P + U$ dove $P = (-2, -1, 1, 2)$ e $U \subset \mathbb{R}^4$ (qui $\mathbb{R}^4$ è lo spazio vettoriale $\mathbb{R}^4$) è dato da

$$U := \langle (1, -2, 3, -4), (2, 0, 3, 1) \rangle. \tag{5.4.5}$$

Sia $f \in (\mathbb{R}^4)^*$ cioè

$$\begin{array}{ccc} \mathbb{R}^4 & \xrightarrow{f} & \mathbb{R} \\ (x_1, x_2, x_3, x_4) & \mapsto & \lambda_1 x_1 + \dots + \lambda_4 x_4 \end{array} \tag{5.4.6}$$

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

$$0 = \lambda_1 - 2\lambda_2 + 3\lambda_3 - 4\lambda_4 = 2\lambda_1 + 3\lambda_3 + \lambda_4. \tag{5.4.7}$$

Risolvendo il sistema di equazioni lineari (5.4.7) troviamo che una base di $\operatorname{Ann} U$ è data da $\{2e_1^* + 9e_2^* - 4e_4^*, 6e_1^* - 3e_2^* - 4e_3^*\}$ e quindi

$$U = \{X \in \mathbb{R}^4 \mid 0 = 2x_1 + 9x_2 - 4x_4 = 6x_1 - 3x_2 - 4x_3\}. \tag{5.4.8}$$

Le equazioni cartesiane di U sono date da (5.4.8): segue che

$$\mathbb{B} = \{X \in \mathbb{R}^4 \mid 0 = F_1(X) = F_2(X)\} \tag{5.4.9}$$

dove

$$F_1(X) := 2(x_1+2)+9(x_2+1)-4(x_4-2), \qquad F_2(X) := 6(x_1+2)-3(x_2+1)-4(x_3-1). \tag{5.4.10}$$

Quindi

$$\mathbb{B} = \{X \in \mathbb{R}^4 \mid 0 = 2x_1 + 9x_2 - 4x_4 + 21 = 6x_1 - 3x_2 - 4x_3 + 13\}.$$

Esercizi del Capitolo 5

Esercizio 5.1. *Dite se esiste/non esiste un'applicazione affine $F \colon \mathbb{R}^2 \to \mathbb{R}^2$ 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 5.2. *Sia* $\mathbb{A} \subset \mathbb{R}^4$ *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 $\mathbb{A}$*.*

Esercizio 5.3. *Siano* $\mathbb{A}$ *un piano affine (spazio affine di dimensione* 2*) e* $\mathbb{L}_1, \mathbb{L}_2 \subset \mathbb{A}$ *due rette (sottospazi affini di dimensione* 1*). Sia* $\mathbb{L} \subset \mathbb{A}$ *una retta che* non *è parallela a* $\mathbb{L}_1$ *né a* $\mathbb{L}_2$*.*

1. *Dimostrate che dato* $P \in \mathbb{L}_1$ *l'unica retta parallela a* $\mathbb{L}$ *contenente* P *incontra* $\mathbb{L}_2$ *in un unico punto.*
2. *Per il punto 1 possiamo definire* $F \colon \mathbb{L}_1 \to \mathbb{L}_2$ *associando a* $P \in \mathbb{L}_1$ *l'intersezione di* $\mathbb{L}_2$ *con l'unica retta parallela a* $\mathbb{L}$ *contenente* P*. Dimostrate che* F *è un isomorfismo affine.*

Esercizio 5.4. *Siano* $\mathbb{A}$ *uno spazio affine e* $F \colon \mathbb{A} \to \mathbb{A}$ *un'affinità. Un* punto fisso *di* F *è un* $P \in \mathbb{A}$ *tale che* $F(P) = P$*. Il luogo dei punti fissi* $\mathrm{Fix}(F)$ *è l'insieme dei punti fissi di* F*.*

1. *Dimostrate che* $\mathrm{Fix}(F)$ *è un sottospazio affine di* $\mathbb{A}$*.*
2. *Date esempi con* $\dim \mathbb{A} = n$ *e* $\dim \mathrm{Fix}(F) = m$ *per ogni* $0 \leq m \leq n$*.*

Capitolo 6

Determinanti

Sia k un campo. Il determinante dà una funzione polinomiale $M_{n,n}(k) \to k$ che vale 0 su A se e solo se A *non* è invertibile. Nel caso in cui $k = \mathbb{R}$ e $n = 2$ o $n = 3$ possiamo interpretare il determinante come un'area o, rispettivamente, un volume (con segno). Il determinante è uno strumento teorico importante.

6.1 La definizione

Sia k un campo. Sia $A \in M_{n,n}(k)$. Sia A^i_j la matrice $(n-1) \times (n-1)$ ottenuta eliminando riga i-esima e colonna j-esima di A. Definiamo una funzione

$$\mathrm{Det}_n \colon M_{n,n}(k) \longrightarrow k$$

così:

(1) $\mathrm{Det}_1((a)) = a$.

(2) Per $n > 1$ definiamo $\det_n$ ricorsivamente:

$$\mathrm{Det}_n(A) := \sum_{j=1}^{n} (-1)^{n+j} a_{nj} \,\mathrm{Det}_{n-1}(A^n_j). \tag{6.1.1}$$

Spieghiamo il punto (2). Assumendo di aver definito Det_{n-1}, la funzione Det_n è data da (6.1.1); siccome Det_1 è data da (1) segue che Det_2 è bene definita e quindi anche Det_3 etc. Diamo alcuni esempi.

$$\mathrm{Det}_2 \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = -a_{21}a_{12} + a_{22}a_{11} = a_{11}a_{22} - a_{12}a_{21} \tag{6.1.2}$$

cioè la formula imparata nelle scuole medie. Abbiamo anche

$$\begin{aligned} \mathrm{Det}_3 \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \\ = a_{31}(a_{12}a_{23} - a_{13}a_{22}) - a_{32}(a_{11}a_{23} - a_{13}a_{21}) + a_{33}(a_{11}a_{22} - a_{12}a_{21}) = \\ = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}. \end{aligned} \tag{6.1.3}$$

La funzione determinante ha varie notevoli proprietà. Il Teorema che segue dà una di queste proprietà.

Teorema 6.1.1. *Una $A \in M_{n,n}(k)$ è invertibile se e solo se* $\mathrm{Det}(A) \neq 0$.

Il **Teorema 6.1.1** verrà dimostrato in seguito, vedi il **Corollario 6.3.5**.

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

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = \mathrm{Det} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = -2. \tag{6.1.4}$$

6.2 Applicazioni multilineari

Una matrice $A \in M_{n,n}(k)$ è univocamente determinata dalle sue colonne $A_1, \ldots, A_n \in k^n$ e viceversa n vettori colonna $A_1, \ldots, A_n \in k^n$ individuano un'unica matrice (di cui sono rispettivamente prima, seconda,..., ultima colonna), quindi possiamo considerare Det_n come una funzione di n vettori colonna. Prima di studiare le proprietà del determinante (visto come funzione delle colonne) ci soffermeremo a studiare la seguente classe di funzioni. Sia V uno spazio vettoriale su k e sia

$$\begin{array}{ccc} \underbrace{V \times \ldots \times V}_{n} & \xrightarrow{\Phi} & k \\ (v_1, \ldots, v_n) & \mapsto & \Phi(v_1, \ldots, v_n) \end{array} \tag{6.2.1}$$

La funzione Det_n è un esempio di tale funzione; in questo caso $V = k^n$ e perciò $\dim V = n$, in generale $\dim V$ e n sono diversi (e V non è necessariamente finitamente generato).

Definizione 6.2.1. Sia $\Phi\colon V^n \to k$.

(1) Sia $1 \leq j \leq n$; Φ è *lineare nell'entrata j-esima* se

$$\begin{aligned} &\Phi(v_1, \ldots, v_{j-1}, \lambda u + \mu w, v_{j+1}, \ldots, v_n) = \\ &\qquad = \lambda \Phi(v_1, \ldots, v_{j-1}, u, v_{j+1}, \ldots, v_n) + \mu \Phi(v_1, \ldots, v_{j-1}, w, v_{j+1}, \ldots, v_n) \end{aligned} \tag{6.2.2}$$

per $v_1, \ldots, v_{j-1}, u, w, v_{j+1}, \ldots, v_n \in V$ e $\lambda, \mu \in k$.

(2) Φ è *multilineare* (o n-lineare) se è lineare in ciascun entrata.

Esempio 6.2.2. Consideriamo le applicazioni $\Psi_i\colon k \times k \to k$ definite da

$$\Psi_1(x, y) := 3x^2 y, \quad \Psi_2(x, y) := 1, \quad \Psi_3(x, y) := 5xy.$$

La Ψ_1 è lineare nella prima entrata ma non nella seconda (a meno che $k = \mathbb{F}_2$), la Ψ_2 non è lineare in alcuna entrata, la Ψ_2 è bilineare.

Definizione 6.2.3. Sia $\Phi\colon V^n \to k$.

(1) Siano $1 \leq j < h \leq n$; 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 $1 \leq j < h \leq n$.

Esempio 6.2.4. Consideriamo le applicazioni $\Phi_i \colon k \times k \to k$ 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 né bilineare né alternante.

Osservazione 6.2.5. Sia $\Phi \colon V^n \to k$.

(a) Φ è lineare nell'entrata j-esima se e solo se per ogni $v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n \in V$ la funzione

$$\begin{array}{ccc} V & \longrightarrow & k \\ u & \mapsto & \Phi(v_1, \dots, v_{j-1}, u, v_{j+1}, \dots, v_n) \end{array} \tag{6.2.3}$$

è lineare.

(b) Φ è alternante nelle entrate j, h se e solo se la funzione

$$\begin{array}{ccc} V \times V & \longrightarrow & k \\ (u,w) & \mapsto & \Phi(v_1, \dots, v_{j-1}, u, v_{j+1}, \dots v_{h-1}, w, v_{h+1}, \dots, v_n) \end{array} \tag{6.2.4}$$

è alternante.

Lemma 6.2.6. *Sia $\Phi \colon V^n \to k$. Supponiamo che Φ sia multilineare e alternante. Siano $1 \le j < h \le n$: allora*

$$\Phi(v_1, \dots, v_{j-1}, v_h, v_{j+1}, \dots, v_{h-1}, v_j, v_{h+1}, \dots, v_n) = \\ = -\Phi(v_1, \dots, v_{j-1}, v_j, v_{j+1}, \dots, v_{h-1}, v_h, v_{h+1}, \dots, v_n),$$

cioè scambiando due entrate il valore di Φ cambia segno.

Dimostrazione. 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 $n > 2$: la funzione (6.2.4) è alternante per l'**Osservazione 6.2.5** e quindi la proposizione segue dal caso $n = 2$. □

Osservazione 6.2.7. Notate che entrambe le ipotesi del **Lemma 6.2.6** sono necessarie perché valga la tesi. Infatti siano Φ_1, Φ_3 le applicazioni dell'**Esempio 6.2.4**: Φ_1 è multilineare ma non alternante, Φ_3 è alternante ma non multilineare e abbiamo

$$\Phi_1(2,1) \neq -\Phi_1(1,2), \qquad \Phi_3(1,0) \neq -\Phi_1(0,1).$$

Lemma 6.2.8. *Sia $\Phi \colon V^n \to k$. Supponiamo che Φ sia multilineare e che sia alternante nelle entrate j e $j+1$ per ogni $1 \le j < n$ (in altre parole per entrate adiacenti). Allora Φ è alternante.*

Dimostrazione. Siano $1 \le j < h \le n$. Dimostriamo che Φ è alternante nelle entrate j e h. Se $(h-j) = 1$ è vero per ipotesi. Quindi possiamo assumere che Φ è alternante nelle entrate j_0 e h_0 (con $1 \le j_0 < h_0 \le n$) ogni qualvolta $(h_0 - j_0) < (h - j)$. Poniamo $v_j = v_h = w$; per il **Lemma 6.2.6** abbiamo che

$$\Phi(v_1, \dots, v_{j-1}, w, v_{j+1}, \dots, v_{h-1}, w, v_{h+1}, \dots, v_n) = \\ = -\Phi(v_1, \dots, v_{j-1}, w, v_{j+1}, \dots, v_{h-2}, w, v_{h-1}, v_{h+1}, \dots, v_n) \tag{6.2.5}$$

e siccome Φ è alternante nelle entrate $j_0 = j$ e $h_0 = h - 1$ segue che il membro di destra di (6.2.5) è nullo. □

6.3 Proprietà del determinante

Dimostreremo che la funzione Det_n gode di alcune proprietà algebriche che la rendono interessante. Queste proprietà danno un algoritmo efficiente per calcolare il determinante. Dati $A_1, \ldots, A_n \in k^n$, visti come matrici colonna, denoteremo con $\mathrm{Det}_n(A_1, \ldots, A_n)$ il determinante della matrice $n \times n$ le cui colonne sono $A_1, \ldots, A_n$.

Proposizione 6.3.1. *La funzione* Det_n *gode delle seguenti proprietà:*

(1) $\mathrm{Det}_n(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$.

(2) È multilineare.

(3) È alternante.

Dimostrazione. Per induzione su n. Il caso $n = 1$ è banalmente vero. Dimostriamo il passo induttivo. Il punto (1) segue immediatamente da (6.1.1). Verifichiamo il passo induttivo per il punto (2). Dimostriamo che Det_n è lineare nella colonna j_0-esima, cioè che

$$\begin{aligned}&\mathrm{Det}_n(A_1, \ldots, A_{j_0-1}, \lambda B + \mu C, A_{j_0+1}, \ldots, A_n) = \\ &= \lambda \,\mathrm{Det}_n(A_1, \ldots, A_{j_0-1}, B, A_{j_0+1}, \ldots, A_n) + \mu \,\mathrm{Det}_n(A_1, \ldots, A_{j_0-1}, C, A_{j_0+1}, \ldots, A_n) \quad (6.3.1)\end{aligned}$$

dove vettori 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 $X_j \in M_{n-1,1}(k)$ la colonna ottenuta eliminando l'ultima entrata di A_j. Siano $Y, Z \in M_{n-1,1}(k)$ le colonne ottenute eliminando l'ultima entrata di B e C rispettivamente. Si ha che

$$\begin{aligned}&\mathrm{Det}_n(A_1, \ldots, A_{j_0-1}, \lambda B + \mu C, A_{j_0+1}, \ldots, A_n) = \\ &= \sum_{j \neq j_0} (-1)^{n+j} a_{nj} \,\mathrm{Det}_{n-1}(X_1, \ldots, X_{j-1}, \widehat{X_j}, X_{j+1}, \ldots, X_{j_0-1}, \lambda Y + \mu Z, X_{j_0+1}, \ldots, X_n) + \\ &\qquad + (-1)^{n+j_0} (\lambda b_n + \mu c_n) \,\mathrm{Det}_{n-1}(X_1, \ldots, X_{j_0-1}, X_{j_0+1}, \ldots, X_n). \quad (6.3.2)\end{aligned}$$

(La notazione $\widehat{X_j}$ sta per "manca la colonna X_j".) Per l'ipotesi induttiva abbiamo che

$$\begin{aligned}&\mathrm{Det}_{n-1}(X_1, \ldots, X_{j-1}, \widehat{X_j}, X_{j+1}, \ldots, X_{j_0-1}, \lambda Y + \mu Z, X_{j_0+1}, \ldots, X_n) = \\ &\qquad \lambda \,\mathrm{Det}_{n-1}(X_1, \ldots, X_{j-1}, \widehat{X_j}, X_{j+1}, \ldots, X_{j_0-1}, Y, X_{j_0+1}, \ldots, X_n) + \\ &\qquad\quad + \mu \,\mathrm{Det}_{n-1}(X_1, \ldots, X_{j-1}, \widehat{X_j}, X_{j+1}, \ldots, X_{j_0-1}, Z, X_{j_0+1}, \ldots, X_n). \quad (6.3.3)\end{aligned}$$

Sostituendo nella (6.3.2) l'espressione della (6.3.3) otteniamo che vale (6.3.1). Ora dimostriamo che vale il punto (3). Per il **Lemma 6.2.8** è sufficiente dimostrare che Det_n è alternante nelle colonne j_0 e $(j_0 + 1)$ dove $1 \leq j_0 < n$. Quindi supponiamo che $A_{j_0} = A_{j_0+1}$ e dimostriamo che $\mathrm{Det}_n(A) = 0$. Per $1 \leq j \leq n$ sia $X_j \in M_{n-1,1}(k)$ la colonna ottenuta eliminando l'ultima entrata di A_j. Si ha che

$$\begin{aligned}&\mathrm{Det}_n(A_1, \ldots, A_n) = \\ &\qquad = \sum_{j_0 \neq j \neq j_0+1} (-1)^{n+j} a_{nj} \,\mathrm{Det}_{n-1}(X_1, \ldots, X_{j-1}, \widehat{X_j}, X_{j+1}, \ldots, X_n) + \\ &\qquad\quad + (-1)^{n+j_0} a_{n,j_0} \,\mathrm{Det}_{n-1}(X_1, \ldots, X_{j_0-1}, \widehat{X_{j_0}}, X_{j_0+1}, \ldots, X_n) + \\ &\qquad\qquad + (-1)^{n+j_0+1} a_{n,j_0+1} \,\mathrm{Det}_{n-1}(X_1, \ldots, X_{j_1-1}, \widehat{X_{j_1}}, X_{j_1+1}, \ldots, X_n). \quad (6.3.4)\end{aligned}$$

Per l'ipotesi induttiva Det_{n-1} è alternante: per ipotesi $X_{j_0} = X_{j_0+1}$ e perciò

$$\mathrm{Det}_{n-1}(X_1, \dots, X_{j-1}, \widehat{X_j}, X_{j+1}, \dots, X_n) = 0, \qquad 1 \leq j \leq n, \quad j_0 \neq j \neq j_0 + 1.$$

Le n-ple $(X_1, \dots, X_{j_0-1}, \widehat{X_{j_0}}, X_{j_0+1}, \dots, X_n)$ e $(X_1, \dots, X_{j_0}, \widehat{X_{j_0+1}}, X_{j_0+2}, \dots, X_n)$ 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 (6.3.4). □

Consideriamo una funzione $\Phi\colon (k^n)^n \to k$, cioè come in (6.2.1) con $V = k^n$. Identificando una matrice $n \times n$ con la n-pla ordinata delle sue colonne possiamo pensare Φ come una funzione $\Phi\colon M_{n,n}(k) \to k$.

Proposizione 6.3.2. *Sia $\Phi\colon M_{n,n}(k) \to k$. Supponiamo che Φ, vista come funzione delle colonne, sia multilineare e alternante.*

(1) Se $A \in M_{n,n}(k)$ è singolare allora $\Phi(A) = 0$.

(2) Siano $A, B \in M_{n,n}(k)$ e supponiamo che B sia ottenuta da A con una serie di operazioni elementari sulle colonne di tipo (1) e (2) (vedi la **Definizione 4.6.9***), e che il numero di scambi di colonne sia s. Allora $\Phi(A) = (-1)^s \Phi(B)$.*

(3) Sia A a scala per colonne. Allora

$$\Phi(A) = a_{1,1} \cdot a_{2,2} \cdot \ldots \cdot a_{n,n} \Phi(1_n). \tag{6.3.5}$$

Dimostrazione. (1): Siccome A è singolare esiste una colonna di A, diciamo A_{j_0}, che è combinazione lineare delle rimanenti colonne:

$$A_{j_0} = \sum_{j \neq j_0} \lambda_j A_j.$$

La linearità di Φ nella colonna j_0-esima insieme alla proprietà di essere alternante dà che

$$\Phi(A) = \Phi(A_1, \dots, A_{j_0-1}, \sum_{j \neq j_0} \lambda_j A_j, A_{j_0+1}, \dots, A_n) = \\ = \sum_{j \neq j_0} \lambda_j \Phi(A_1, \dots, A_{j_0-1}, A_j, A_{j_0+1}, \dots, A_n) = 0.$$

(2): Una operazione di tipo (1) cambia segno al determinante per il **Lemma 6.2.6**, d'altra parte il calcolo appena fatto dimostra che una operazione di tipo (2) non cambia il valore del determinante. Questo finisce la dimostrazione del punto (2). (3): Supponiamo che esista $1 \leq i_0 \leq n$ tale che $a_{i_0,i_0} = 0$. Siccome A è a scala per colonne segue che $a_{ii} = 0$ per ogni $i_0 \leq i \leq n$. In particolare l'ultima colonna di A è nulla, e quindi A è singolare. Per il punto (1) segue che $\Phi(A) = 0$: siccome il membro di destra di (6.3.5) è nullo abbiamo dimostrato che se uno degli a_{ii} è nullo allora vale (6.3.5). Ora supponiamo che ciascun a_{ii} sia diverso da 0. Con questa ipotesi esiste una serie di operazioni elementari di tipo (2) che trasforma A in una matrice diagonale con le stesse entrate sulla diagonale principale: per il punto (2) segue che

$$\Phi(A) = \Phi\left(\begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{11} & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & a_{ii} & \dots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}\right) = a_{1,1} \cdot a_{2,2} \cdot \ldots \cdot a_{n,n} \Phi(1_n).$$

(La seconda uguaglianza segue dalla multilinearità di Φ.) □

Corollario 6.3.3. *Sia* $\Phi\colon M_{n,n}(k) \to k$. *Supponiamo che* Φ, *vista come funzione delle colonne, sia multilineare e alternante.*

(1) Se $A \in M_{n,n}(k)$ *allora*

$$\Phi(A) = \Phi(1_n)\,\mathrm{Det}_n(A). \tag{6.3.6}$$

(2) Supponiamo che $\Phi(1_n) \neq 0$. *Allora* $\Phi(A) = 0$ *se e solo se* A *è singolare.*

Dimostrazione. (1): Esiste una serie di operazioni elementari di tipo (1) e (2) sulle colonne di A che trasformano A in una matrice a scala per colonne B - vedi **Proposizione 4.6.12**. Sia s il numero di scambi di colonne. La **Proposizione 6.3.2** applicata a Φ e Det_n dà che

$$\begin{array}{rclcl} \Phi(A) & = & (-1)^s\Phi(B) & = & (-1)^s b_{11}\cdot b_{22}\cdot\ldots\cdot b_{nn}\Phi(1_n), \\ \mathrm{Det}_n(A) & = & (-1)^s\,\mathrm{Det}_n(B) & = & (-1)^s b_{11}\cdot b_{22}\cdot\ldots\cdot b_{nn}. \end{array} \tag{6.3.7}$$

Sostituendo nella prima equazione l'espressione per $(-1)^s b_{11}\cdot b_{22}\cdot\ldots\cdot b_{nn}$ data dalla seconda equazione otteniamo (6.3.6). (2): Se A è singolare allora $\Phi(A) = 0$ per il punto (1) della **Proposizione 6.3.2**. Ora supponiamo che $\Phi(A) \neq 0$. Sia B come nella dimostrazione del punto (1): si ha che $b_{11}\cdot b_{22}\cdot\ldots\cdot b_{nn} \neq 0$ perché A non è singolare. Per la prima equazione di (6.3.7) e l'ipotesi $\Phi(1_n) \neq 0$ vediamo che $\Phi(A) \neq 0$. □

Corollario 6.3.4. *Sia* $\Phi\colon M_{n,n}(k) \to k$. *Supponiamo che* Φ, *vista come funzione delle colonne, sia multilineare e alternante, e che* $\Phi(1_n) = 1$. *Allora* $\Phi = \mathrm{Det}_n$.

Dimostrazione. Segue immediatamente dal **Corollario 6.3.3**. □

Da ora in poi la notazione Det_n sarà sostituita quasi sempre da Det. Il seguente corollario dimostra che vale il **Teorema 6.1.1**.

Corollario 6.3.5. *Una* $A \in M_{n,n}(k)$ *è invertibile se e solo se* $\mathrm{Det}(A) \neq 0$.

Dimostrazione. Segue dal **Corollario 6.3.3** e dal fatto che $\mathrm{Det}(1_n) = 1 \neq 0$. □

Osservazione 6.3.6. La **Proposizione 6.3.2** dà un metodo efficiente per calcolare il determinante di una matrice quadrata A di ordine grande. Con operazioni elementari di tipo (1) e (2) trasformiamo A in una matrice B a scala per colonne: il determinante di A si calcola applicando i punti (2) e (3) della **Proposizione 6.3.2**.

6.4 La Formula di Binet

Proposizione 6.4.1 (Formula di Binet)**.** *Siano* $A, B \in M_{n,n}(k)$. *Allora* $\mathrm{Det}(A\cdot B) = \mathrm{Det}(A)\cdot \mathrm{Det}(B)$.

Dimostrazione. Sia $\Phi\colon M_{n,n}(k) \to k$ l'applicazione definita da $\Phi(M) := \mathrm{Det}(A\cdot M)$. Si verifica facilmente che Φ, come funzione delle colonne è multilineare e alternante. Per il **Corollario 6.3.3** segue che

$$\mathrm{Det}(A\cdot B) = \Phi(B) = \Phi(1_n)\,\mathrm{Det}(B) = \mathrm{Det}(A)\cdot\mathrm{Det}(B).$$

□

Corollario 6.4.2. *Sia* $A \in M_{n,n}(k)$ *invertibile cioè con* $\mathrm{Det}\,A \neq 0$ *per il* **Corollario 6.3.5**. *Allora* $\mathrm{Det}(A^{-1}) = \mathrm{Det}(A)^{-1}$.

Dimostrazione. Per la formula di Binet abbiamo che

$$1 = \mathrm{Det}(1_n) = \mathrm{Det}(A \cdot A^{-1}) = \mathrm{Det}(A) \cdot \mathrm{Det}(A^{-1}).$$

□

La formula di Binet ha la seguente importante conseguenza.

Corollario 6.4.3. *Sia V uno spazio vettoriale su 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. Segue dalle equazioni (4.9.6) e (4.9.3) insieme al **Corollario 6.4.2**. □

Definizione 6.4.4. Siano V uno spazio vettoriale su 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 6.4.3**.

La Formula di Binet dà che dat $f, g \in \mathrm{End}(V)$ vale la seguente equazione:

$$\mathrm{Det}(f \circ g) = \mathrm{Det}(f) \cdot \mathrm{Det}(g). \tag{6.4.1}$$

6.5 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 è la Formula (6.1.1) che definisce il determinante.

Proposizione 6.5.1. *Siano $A \in M_{n,n}(k)$ e $1 \leq i \leq n$. Abbiamo che*

$$\mathrm{Det}(A) := \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \,\mathrm{Det}(A_j^i). \tag{6.5.1}$$

Dimostrazione. Sia $\Phi^i\colon M_{n,n}(k) \to k$ la funzione definita ponendo $\Phi^i(A)$ uguale al membrod di destra di (6.5.1). Allora Φ^i, vista come funzione delle (n) colonne, è multilineare e alternante: per la dimostrazione nel caso $i = n$ vedi la **Proposizione 6.3.1**, la dimostrazione per un i qualsiasi è del tutto analoga. Per il **Corollario 6.3.3** segue che

$$\Phi^i(A) = \Phi^i(1_n)\,\mathrm{Det}(A) \qquad \forall A \in M_{n.n}(k).$$

Un facile calcolo dà che $\Phi^i(1_n) = 1$. □

Proposizione 6.5.2. *Sia $A \in M_{n,n}(k)$. Allora* $\mathrm{Det}(A) = \mathrm{Det}(A^t)$.

Dimostrazione. Sia $\Phi\colon M_{n,n}(k) \to k$ definita da $\Phi(A) := \mathrm{Det}(A^t)$. Consideriamo la Φ come funzione delle colonne. La **Proposizione 6.5.1** 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 la **Proposizione 6.3.2** 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 il **Corollario 6.3.4** segue che $\mathrm{Det}_n(A^t) = \Phi(A) = \mathrm{Det}_n(A)$. □

Osservazione 6.5.3. Sia $A \in M_{n,n}(k)$ e supponiamo che $B \in M_{n,n}(k)$ 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 la **Proposizione 6.3.2** e la **Proposizione 6.5.2** abbiamo che

$$\mathrm{Det}(A) = \mathrm{Det}(A^t) = (-1)^s \mathrm{Det}(B^t) = (-1)^s \mathrm{Det}(B). \tag{6.5.2}$$

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 6.5.4. *Siano $A \in M_{n,n}(k)$ e $1 \leq j \leq n$. Abbiamo che*

$$\mathrm{Det}(A) := \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \mathrm{Det}(A^i_j). \tag{6.5.3}$$

Dimostrazione. Per la **Proposizione 6.5.2** abbiamo che $\mathrm{Det}(A) = \mathrm{Det}(A^t)$. Espandendo $\mathrm{Det}(A^t)$ secondo la riga j (che è uguale alla colonna j-esima di A), vedi (6.5.1), otteniamo (6.5.3). □

Le Formule (6.5.1) e (6.5.3) sono casi particolare dell'espansione di Laplace di un determinante: sono vantaggiose se la matrice di cui vogliamo calcolare il determinante ha molte entrate nulle.

6.6 Permutazioni e determinante

Definizione 6.6.1. Una *permutazione* di un insieme X è un'applicazione biunivoca $\sigma\colon X \to X$. L'insieme delle permutazioni di $\{1, \ldots, n\}$ (o di uni insieme di cardinalità n) si denota $\mathcal{S}_n$.

Osservazione 6.6.2. L'identità di X è una permutazione, l'inversa di una permutazione di X è una permutazione, e la composizione di due permutazioni di X è una permutazione.

Definizione 6.6.3. Un elemnto $\sigma \in \mathcal{S}_n$ è una *trasposizione* se scambia tra di loro $1 \leq a < b \leq n$ e lascia invariati gli altri elementi, cioè $\sigma(a) = b$, $\sigma(b) = a$ e $\sigma(i) = i$ per $i \in (\{1, \ldots, n\} \setminus \{a, b\})$.

Definiremo un'applicazione (denominata *segno*)

$$\mathcal{S}_n \xrightarrow{\epsilon} \{1, -1\} \tag{6.6.1}$$

con le seguenti proprietà:

1. $\epsilon(\sigma \circ \tau) = \epsilon(\sigma) \cdot \epsilon(\tau)$ per ogni $\sigma, \tau \in \mathcal{S}_n$,
2. $\epsilon(\sigma) = -1$ per ogni trasposizione σ.

Definizione 6.6.4. Data $\sigma \in \mathcal{S}_n$ la matrice $M_\sigma \in M_{n,n}(\mathbb{Q})$ ha entrate m_{ij} definite così:

$$m_{ij} := \begin{cases} 1 & \text{se } i = \sigma(j) \\ 0 & \text{se } i \neq \sigma(j) \end{cases}$$

Siano $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ la base standard di $\mathbb{Q}^n$ e $1 \leq j \leq n$. Si verifica immediatamente che

$$L_{M_\sigma}(\mathbf{e}_j) = \mathbf{e}_{\sigma(j)}. \tag{6.6.2}$$

Applicando l'espansione del determinante secondo una riga si verifica facilmente che Det $L_{M_\sigma} = \pm 1$. Definiamo l'applicazione (6.6.1) così:

$$\epsilon(\sigma) := \mathrm{Det}(L_{M_\sigma}). \tag{6.6.3}$$

Siano $\sigma, \tau \in \mathcal{S}_n$: la Formula di Binet dà che

$$\epsilon(\sigma \circ \tau) = \epsilon(\sigma)\epsilon(\tau). \tag{6.6.4}$$

Inoltre se $\sigma \in \mathcal{S}_n$ è una trasposizione un calcolo diretto dà che $\epsilon(\sigma) = -1$. Abbiamo dimostrato che la funzione segno ha le due proprietà promesse.

Osservazione 6.6.5. Sia $\sigma \in \mathcal{S}_n$. Allora σ si può scrivere (non in modo unico) come composizione di trasposizioni (esercizio). Supponiamo che $\sigma = \tau_1 \circ \ldots \circ \tau_r = \theta_1 \circ \ldots \circ \theta_s$ dove $\tau_1, \ldots, \tau_r, \theta_1, \ldots, \theta_s$ sono trasposizioni. Le proprietà della funzione segno danno che $(r - s)$ è pari.

Proposizione 6.6.6. *Siano k un campo e $A = (a_{ij}) \in M_{n,n}(k)$. Allora*

$$\mathrm{Det}\, A = \sum_{\sigma \in \mathcal{S}_n} \epsilon(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \ldots a_{n,\sigma(n)}. \tag{6.6.5}$$

Dimostrazione. Sia $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ la base standard di k^n. La funzione Det è multilineare e alternante (come funzione delle righe), quindi

$$\begin{aligned}\mathrm{Det}\, A = \mathrm{Det}\Big(\textstyle\sum_{j=1}^n a_{1j}\mathbf{e}_j, \ldots, \sum_{j=1}^n a_{ij}\mathbf{e}_j, \ldots, \sum_{j=1}^n a_{nj}\mathbf{e}_j\Big) = \\ \textstyle\sum_{\sigma \in \mathcal{S}_n} \mathrm{Det}(M_\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \ldots a_{n,\sigma(n)} = \sum_{\sigma \in \mathcal{S}_n} \epsilon(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \ldots a_{n,\sigma(n)}.\end{aligned} \tag{6.6.6}$$

□

6.7 La formula di Cramer

Sia $A \in M_{n,n}(k)$. Siano $1 \leq i, j \leq n$. Il *cofattore* (o complemento algebrico) di A di indici i, j è

$$A^{ij} := (-1)^{i+j} \mathrm{Det}(A^i_j). \tag{6.7.1}$$

La *matrice dei cofattori di A* (anche matrice aggiunta ma questo termine indica anche una matrice del tutto diversa) è la trasposta della matrice $n \times n$ con entrate A_{ij}:

$$A^c := (A_{ji}).$$

Esempio 6.7.1. 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 6.7.2 (Formula di Cramer)**.** *Sia $A \in M_{n,n}(k)$. Allora*

$$A \cdot A^c = A^c \cdot A = (\operatorname{Det} A)1_n. \tag{6.7.2}$$

Se A è invertibile, cioè $\operatorname{Det} A \neq 0$, si ha che $A^{-1} = (\operatorname{Det} A)^{-1}A^c$.

Dimostrazione. Siano $1 \leq i, j \leq n$. L'entrata al posto i, j di $A \cdot A^c$ è uguale a

$$\sum_{s=1}^{n}(-1)^{j+s}a_{is}\operatorname{Det}(A_s^j). \tag{6.7.3}$$

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 $\operatorname{Det} A$. Oraa supponiamo che $i \neq j$: la (6.7.3) 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 (6.7.3). La formula $A^{-1} = (\operatorname{Det} A)^{-1}A^c$ segue dalla (6.7.3) moltiplicando ambo i membri della prima (o della seconda) uguaglianza per $(\operatorname{Det} A)^{-1}$. □

Esempio 6.7.3. Sia $A \in M_{3,3}(\mathbb{R})$ la matrice dell'**Esempio 6.7.1**. Applicando la formula di Cramer otteniamo che

$$A^{-1} = \begin{bmatrix} 1 & 1/2 & -1/2 \\ 1/3 & -1/2 & 1/6 \\ -1/3 & 0 & 1/3 \end{bmatrix}.$$

6.8 Determinante e area

Sia $\mathcal{V}^2$ lo spazio vettoriale dei vettori nel piano $\mathcal{A}^2$. Scegliamo una base $\mathcal{B} = \{\mathbf{i}, \mathbf{j}\}$ di $\mathcal{V}^2$ *ortonormale* cioè tale che $\mathbf{i}$ e $\mathbf{j}$ abbiano lunghezza 1 e siano ortogonali.

Proposizione 6.8.1. *Siano $v, w \in \mathcal{V}^2$ e $O \in \mathcal{A}^2$. Sia $\mathbf{T}$ il parallelogramma[1] di vertici*

$$O, O + v, O + w, O + v + w. \tag{6.8.1}$$

Per $1 \leq i, j \leq 2$ siano $a_{ij} \in \mathbb{R}$ tali che

$$v = a_{11}\mathbf{i} + a_{21}\mathbf{j}, \qquad w = a_{21}\mathbf{i} + a_{22}\mathbf{j},$$

cioè le coordinate di v e w nella base $\mathcal{B}$. Sia $A \in M_{2,2}(\mathbb{R})$ la matrice con entrate a_{ij}. Allora l'area di $\mathbf{T}$ è uguale a $|\operatorname{Det} A|$.

Dimostrazione. Supponiamo che $a_{21} = 0$. Allora l'altezza del parallelogramma $\mathbf{Q}$ rispetto al lato $\overline{O, O + v}$ è uguale ad $|a_{22}|$ e quindi l'area di $\mathbf{Q}$ è uguale a

$$|a_{11}| \cdot |a_{22}| = |\operatorname{Det} A|. \tag{6.8.2}$$

Abbiamo dimostrato che l'area di $\mathbf{Q}$ è uguale a $|\operatorname{Det} A|$ se $a_{21} = 0$. In generale sia $\theta \in \mathbb{R}$ tale che

$$a_{11}\sin\theta + a_{21}\cos\theta = 0 \tag{6.8.3}$$

[1] Se v e w sono linearmente dipendenti $\mathbf{T}$ è un parallelogramma degenere cioè contenuto in una retta.

e poniamo

$$M_\theta := \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{6.8.4}$$

Siano (x_1, x_2) le coordinate affine del sistema di riferimento $RA(O; \mathbf{i}, \mathbf{j})$ e identifichiamo $\mathcal{A}^2$ per mezzo delle coordinate (x_1, x_2). Sia $F\colon \mathcal{A}^2 \to \mathcal{A}^2$ la rotazione di angolo θ definita da

$$\begin{array}{ccc} \mathcal{A}^2 & \longrightarrow & \mathcal{A}^2 \\ X & \mapsto & M_\theta \cdot X. \end{array} \tag{6.8.5}$$

Siano $P := F(O+v)$ e $Q := F(O+W)$. La rotazione F porta $\mathbf{T}$ nel parallelogramma $\mathbf{S}$ di lati (contingui) OP e OQ, quindi l'area di $\mathbf{T}$ è uguale all'area di $\mathbf{S}$. Sia $B \in M_{22}(\mathbb{R})$ la matrice con prima colonna le coordinate di P e seconda colonna le coordinate di Q. Per (6.8.3) abbiamo che $b_{21} = 0$. Per quanto visto sopra l'area di $\mathbf{S}$ è uguale a $\operatorname{Det} B$. D'altra parte la formula di Binet dà che

$$\operatorname{Det}(A) = \operatorname{Det}(M_\theta \cdot B) = \operatorname{Det}(M_\theta) \cdot \operatorname{Det}(B) = \operatorname{Det}(B). \tag{6.8.6}$$

Siccome l'area di $\mathbf{T}$ è uguale all'area di $\mathbf{S}$ questo dimostra che vale la proposizione. □

Un simile risultato vale per il volume di parallelopipedi nello spazio - verrà dimostrato in seguito.

Proposizione 6.8.2. *Sia $F\colon \mathcal{A}^2 \to \mathcal{A}^2$ un'affinità e $f\colon \mathcal{V}^2 \to \mathcal{V}^2$ l'applicazione lineare associata. Sia $\mathbf{T} \subset \mathcal{A}^2$ un parallelogramma. Allora l'area del parallelogramma $F(\mathbf{T})$ è uguale all'area di $\mathbf{T}$ moltiplicata per $|\operatorname{Det}(f)|$.*

Dimostrazione. Sia $\mathcal{B} = \{\mathbf{i}, \mathbf{j}\}$ una base ortonormale di $\mathcal{V}^2$. Sia $f\colon \mathcal{V}^2 \to \mathcal{V}^2$ l'applicazione lineare associata a F e $M := M_{\mathcal{B}}^{\mathcal{B}}(f)$. Supponiamo che $\mathbf{T}$ sia il parallelogramma di vertici dati da (6.8.1). Sia $A = (a_{ij})$ la matrice con colonne le coordinate di v e w, come nell'enunciato della **Proposizione 6.8.1**, e quindi l'area di $\mathbf{T}$ è uguale a $|\operatorname{Det} A|$. Il parallelogramma $F(\mathbf{T})$ ha vertici $F(O)$, $F(O)+f(v)$, $F(O)+f(w)$, $F(O)+f(v)+f(w)$, e quindi per la **Proposizione 6.8.1**, la sua area è uguale a

$$|\operatorname{Det}(M \cdot A)| = |\operatorname{Det}(f)| \cdot |\operatorname{Det}(A)|. \tag{6.8.7}$$

□

Esercizi del Capitolo 6

Esercizio 6.1. *Calcolate i determinanti delle seguenti matrici reali 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 6.2. *Calcolate le matrici dei cofattori delle A e B dell'***Esercizio 6.1**.

Esercizio 6.3. *Sia k un campo e $x_1, x_2, \ldots, x_n \in k$. 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 6.4. *Sia $A \in M_{n,n}(\mathbb{Q})$. Supponiamo che le entrate di A siano in $\mathbb{Z}$ (cioè sono numeri* interi*). Dimostrate che esiste $B \in M_{n,n}(\mathbb{Q})$ con entrate in $\mathbb{Z}$ tale che $A \cdot B = 1_n$ se e solo se* $\operatorname{Det} A = \pm 1$.

Esercizio 6.5. *Sia $A \in M_{n,n}(\mathbb{Q})$ con entrate in $\mathbb{Z}$. 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* $\operatorname{rk} \overline{A} = r$ *allora* $\operatorname{Det}(A)$ *è divisibile per* p^{n-r}.

Esercizio 6.6. *Per $n \geq 1$ sia $A_n := (a_{ij})$ dove*

$$a_{ij} = \begin{cases} 2 & \text{se } i = j, \\ -1 & \text{se } |i-j| = 1, \\ 0 & \text{altrimenti.} \end{cases} \tag{6.8.8}$$

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 6.7. *Sia k un campo e $x_1, \ldots, x_n \in k$. Sia $A(x_1, \ldots, x_n) \in M_{n,n}(k)$ definita così:*

$$A(x_1, \ldots, x_n) := \begin{bmatrix} 1+x_1 & 1 & 1 & \ldots & 1 \\ 1 & 1+x_2 & 1 & \ldots & 1 \\ 1 & 1 & 1+x_3 & \ldots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \ldots & 1+x_n \end{bmatrix}.$$

Dimostrate che

$$\operatorname{Det} A(x_1, \ldots, x_n) = x_1 x_2 \ldots x_n + \sum_{i=1}^{n} x_1 \cdot \ldots \cdot \widehat{x}_i \cdot \ldots \cdot x_n$$

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 6.8. *Sia $V \subset \mathbb{R}^4$ il sottospazio*

$$V = \{X \in \mathbb{R}^4 \mid x_1 + x_2 + x_3 + x_4 = 0\} = \{X \in \mathbb{R}^4 \mid (\mathbf{e}_1^* + \mathbf{e}_2^* + \mathbf{e}_3^* + \mathbf{e}_4^*)(X) = 0\}.$$

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

$$L_M^*(\mathbf{e}_1^* + \mathbf{e}_2^* + \mathbf{e}_3^* + \mathbf{e}_4^*) = 2(\mathbf{e}_1^* + \mathbf{e}_2^* + \mathbf{e}_3^* + \mathbf{e}_4^*)$$

e **quindi** *(perché ?) $L_M(V) \subset V$. Sia*

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ X & \mapsto & L_M(X) \end{array} \tag{6.8.9}$$

(b) Calcolate $\operatorname{Det} f$*, dove f è data da* (6.8.9)*, 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 6.9. *Sia k un campo. Siano $A^1, \dots, A^{n-1} \subset k^n$ vettori linearmente indipendenti, pensati come vettori-riga. Siccome i vettori sono linearmente indipendenti il sottospazio*

$$V := \langle A^1, \dots, A^{n-1} \rangle \subset k^n$$

ha codimensione 1 in k^n e quindi $\dim \mathrm{Ann}(V) = 1$. Sia $A \in M_{n-1,n}(k)$ la matrice le cui righe sono $A^1, \dots, A^{n-1}$. Dato $1 \leq j \leq n$ sia $M_j \in M_{n-1,n-1}(k)$ la matrice ottenuta eleiminando la colonna j-esima da A. Sia $c_j := (-1)^j \operatorname{Det} M_j$. Sia

$$\begin{array}{ccc} k^n & \xrightarrow{f} & k \\ X & \mapsto & \sum_{j=1}^n c_j x_j \end{array}$$

Dimostrate che

$$\mathrm{Ann}\, V = \langle f \rangle.$$

Esercizio 6.10. *Sia k un campo e supponiamo che $\mathrm{char}\, k \neq 2$.*

(1) Sia n dispari e supponiamo che $A \in M_{n,n}(k)$ sia antisimmetrica cioè che $A^t = -A$. . Dimostrate che $\operatorname{Det} A = 0$.

(2) per ogni n pari date un esempio di $A \in M_{n,n}(k)$ antisimmetrica con $\operatorname{Det} A \neq 0$.

Capitolo 7

Forme quadratiche e bilineari simmetriche

Le funzioni $f\colon k^n \to k$ più semplici sono quelle polinomiali e tra le funzioni polinomiali quelle di grado al più 1 cioè data da $f(X) = a_1x_1 + \dots a_nx_n + b$ dove $a_1, \dots, a_n, b \in k$. Se $b = 0$ allora f è una funzione lineare, in generale è una funzione affine - sono state studiate nei capitoli precedenti e la loro importanza dovrebbe essere evidente. La f è una funzione polinomiale di grado (al più) 2 se è data da $f(X) = \sum_{1\le i\le j\le n} a_{ij}x_ix_j + b_1x_1 + \dots b_nx_n + c$ dove $a_{ij}, b_1, \dots, b_n, c \in k$. Perchè studiarle ? Una motivazione geometrica: l'insieme dei punti del piano le cui coordinate sono le soluzioni di $f(x_1, x_2) = 0$ con f di grado 2 è una figura geometrica ben nota: una conica (in generale, può anche essere un oggetto "degenere"). Un'altra motivazione (qui $k = \mathbb{R}$): come le funzioni di grado al più 1 approssimano un'arbitraria funzione $f\colon \mathbb{R}^n \to \mathbb{R}$ nelle vicinanze di un dato $a \in \mathbb{R}^n$ così le funzioni di grado 2 (omogenee) approssimano un'arbitraria funzione $f\colon \mathbb{R}^n \to \mathbb{R}$ nelle vicinanze di un dato $a \in \mathbb{R}^n$ dove tutte le derivate parziali di f si annullano (un punto critico di f).

Studieremo soprattutto le funzioni polinomiali di grado 2 omogenee. I polinomi omogenei sono, a dispetto delle apparenze, degli oggetti lineari. Infatti associeremo a una tale funzione polinomiale un'applicazione bilineare simmetrica cioè un oggetto "lineare". L'esempio standard di una applicazione bilineare simmetrica è il prodotto scalare sullo spazio dei vettori del piano o dello spazio - in questo caso il polinomio di grado 2 è la funzione che associa a un vettore il quadrato della lunghezza.

In questo capitolo *assumeremo sempre che il campo k ha caratteristica diversa da* 2, e gli *spazi vettoriali saranno finitamente generati se non specifichiamo altrimenti.*

7.1 Forme quadratiche su k^n

Forma di grado d in n variabili è sinonimo di polinomio omogeneo di grado d in n variabili. Una *forma quadratica* è una forma di grado 2, cioè un polinomio

$$f(x_1, \dots, x_n) = \sum_{1\le i\le j\le n} b_{ij}x_ix_j, \tag{7.1.1}$$

dove $b_{ij} \in k$ per ogni coppia di indici $1 \leq i \leq j \leq n$. Si associa a f la matrice $n \times n$ simmetrica $A = (a_{ij})$, dove

$$a_{ij} := \begin{cases} b_{ij} & \text{se } i = j, \\ b_{ij}/2 & \text{se } i < j, \\ b_{ji}/2 & \text{se } i > j. \end{cases}$$

Con questa definizione vale l'uguaglianza

$$f(X) = X^t \cdot A \cdot X \qquad \forall X \in k^n.$$

Ora consideriamo l'applicazione bilineare

$$\begin{array}{ccc} k^n \times k^n & \xrightarrow{F} & k \\ (X, Y) & \mapsto & X^t \cdot A \cdot Y. \end{array}$$

La F dà le derivate direzionali della funzione f (si può dare senso alla derivata per un campo qualsiasi, qui ci accontenteremo del caso $k = \mathbb{R}$). Infatti, per $s \in k$ e $X, Y \in k^n$ si ha

$$f(X + sY) = (X + sY)^t \cdot A \cdot (X + sY) = f(X) + (X^t \cdot A \cdot Y + Y^t \cdot A \cdot X)s + f(Y)s^2, \quad (7.1.2)$$

e $X^t \cdot A \cdot Y = Y^t \cdot A \cdot X$ perché

$$X^t \cdot A \cdot Y = (X^t \cdot A \cdot Y)^t = Y^t \cdot A^t \cdot X = Y^t \cdot A \cdot X.$$

(La prima uguaglianza vale perché $X^t \cdot A \cdot Y$ è una matrice 1×1, l'ultima vale perché per costruzione A è una matrice simmetrica.) Quindi (7.1.2) si può riscrivere

$$f(X + sY) = f(X) + 2X^t \cdot A \cdot Y s + f(Y)s^2.$$

Segue che, se $k = \mathbb{R}$,

$$\left.\frac{df(X + sY)}{ds}\right|_{s=0} = 2F(X, Y).$$

Nelle prime due sezioni studieremo le nozioni di forma quadratica e forma bilineare simmetrica su un arbitrario spazio vettoriale.

7.2 Funzioni polinomiali su uno spazio vettoriale

Sia V uno spazio vettoriale su k. Daremo senso alla nozione di funzione polinomiale $V \to k$. Sia $f\colon V \to k$ una funzione. Siano $v_1, \ldots, v_n \in V$, e supponiamo che esista un polinomio $p \in k[x_1, \ldots, x_n]$ tale che

$$f(x_1v_1 + \ldots + x_nv_n) = p(x_1, \ldots, x_n) \qquad \forall\ (x_1v_1 + \ldots + x_nv_n) \in V. \quad (7.2.1)$$

Proposizione 7.2.1. *Con notazione come sopra supponiamo che $w_1, \ldots, w_m \in V$ e che*

$$\langle w_1, \ldots, w_m \rangle = \langle v_1, \ldots, v_n \rangle. \quad (7.2.2)$$

Allora esiste $q \in k[y_1, \ldots, y_m]$ tale che

$$f(y_1w_1 + \ldots + y_mw_m) = q(y_1, \ldots, y_m) \qquad \forall\ (y_1w_1 + \ldots + y_mw_m) \in V.$$

Inoltre se p è omogeneo di grado d anche q lo è.

Dimostrazione. Infatti per (7.2.2) esiste una matrice $A = (a_{ij}) \in M_{mn}(k)$ tale che valga $w_i = \sum_{j=1}^{n} a_{ij} v_j$ per ogni $1 \leq i \leq m$. Sostituendo questa espressione nella (7.2.1) otteniamo che

$$f(y_1w_1+\ldots+y_mw_m)=f\big(\big(\textstyle\sum_{j=1}^{n} a_{1j}y_j\big)v_1+\ldots+\big(\sum_{j=1}^{n} a_{nj}y_j\big)v_n\big)=p\big(\big(\textstyle\sum_{j=1}^{n} a_{1j}y_j\big),\ldots,\big(\sum_{j=1}^{n} a_{nj}y_j\big)v_n\big)\in k[y_1,\ldots,y_n],$$

e la proposizione segue immediatamente. □

Per la **Proposizione 7.2.1** ha senso porre la seguente definizione.

Definizione 7.2.2. Sia V uno spazio vettoriale su k. Una funzione $f\colon V \to k$ è *polinomiale* se gode della seguente proprietà. Dati $\{v_1, \ldots, v_n\}$ di V vale (7.2.1) con p polinomio; se p è omogeneo di grado d allora f è una *forma* di grado d. Una *forma quadratica* (su V) è una forma di grado 2.

Esempio 7.2.3. Siano $f, g, h\colon \mathbb{R}^2 \to \mathbb{R}$ definite da

$$f(x_1, x_2) := x_1 + 3x_2 + 1, \qquad g(x_1, x_2) := x_1^2 - x_1x_2 + x_2^2, \qquad h(x_1, x_2) := \sin x_1 + \sin x_2.$$

Sia f che g sono funzioni polinomiali. La g è una forma quadratica. La h non è una funzione polinomiale.

Esempio 7.2.4. Una forma di grado 1 su V non è altro che una funzione lineare $V \to k$.

Osservazione 7.2.5. Siano V uno spazio vettoriale e $\mathcal{B} = \{v_1, \ldots, v_n\}$ una sua base. Supponiamo che valga (7.2.1). Allora f è una funzione polinomiale (e se p è omogenea di grado d allora f è una forma di grado d). Infatti siano $w_1, \ldots, w_m \in V$. Estendiamo $w_1, \ldots, w_m \in V$ a generatori $w_1, \ldots, w_m, u_1, \ldots, u_a$ di V. Per la **Proposizione 7.2.1** esiste $q \in k[y_1, \ldots, y_m, z_1, \ldots, z_a]$ tale che

$$f(y_1w_1+\ldots+y_mw_m+z_1u_1+\ldots+z_au_a)=q(y_1,\ldots,y_m,z_1,\ldots,z_a) \qquad \forall\ (y_1w_1+\ldots+y_mw_m+z_1u_1+\ldots+z_au_a)\in V$$

e quindi $f(y_1w_1 + \ldots + y_mw_m) = q(y_1, \ldots, y_m, 0, \ldots, 0)$ per ogni $(y_1w_1 + \ldots + y_mw_m)$: siccome $q(y_1, \ldots, y_m, 0, \ldots, 0)$ è un polinomio questo dimostra che f è una funzione polinomiale (omogenea di grado d se p è omogeneo di grado d).

Esempio 7.2.6. Sia V uno spazio vettoriale su k e $\mathcal{B} = \{v_1, \ldots, v_n\}$ una sua base. Sia $A \in M_{n,n}(k)$. Definiamo $q_A^{\mathcal{B}}\colon V \to k$ così:

$$q_A^{\mathcal{B}}(x_1v_1 + \ldots + x_nv_n) = X^t \cdot A \cdot X \qquad X := \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{7.2.3}$$

Siccome

$$q_A^{\mathcal{B}}(x_1v_1 + \ldots + x_nv_n) = \sum_{1 \leq i,j \leq n} a_{ij}x_ix_j, \tag{7.2.4}$$

$q_A^{\mathcal{B}}$ è una forma quadratica.

Definizione 7.2.7. $Q(V)$ è l'insieme delle forme quadratiche su V.

Proposizione 7.2.8. *$Q(V)$ è un sottospazio dello spazio vettoriale delle funzioni da V a k.*

Dimostrazione. Si tratta di dimostrare che, date $f, g \in Q(V)$ e $\lambda, \mu \in k$, l'applicazione

$$\begin{array}{ccc} V & \overset{(\lambda f+\mu g)}{\longrightarrow} & k \\ v & \mapsto & \lambda f(v) + \mu g(v) \end{array} \tag{7.2.5}$$

è una forma quadratica. Questo segue dal fatto che l'insieme dei polinomi omogenei di grado 2 in n variabili (incluso il polinomio nullo) è un sossospazio vettoriale dello spazio vettoriale dei polinomi. □

Un problema che affronteremo è il seguente.

Problema 7.2.9. Sia V uno spazio vettoriale su un campo k e $f\colon V \to k$ una forma quadratica. Esiste una base $\{v_1, \ldots, v_n\}$ di V tale che $f(x_1v_1 + \ldots + x_nv_n)$ sia un polinomio particolarmente semplice? per esempio del tipo

$$c_1x_1^2 + c_2x_2^2 + \ldots + c_nx_n^2, \qquad c_i \in k, \tag{7.2.6}$$

eventualmente con coefficienti c_i che appartengono a un sottoinsieme assegnato di k (per esempio $\{0, 1\}$, o $\{0, \pm 1\}$).

7.3 Forme bilineari simmetriche e forme quadratiche

7.3.1 Forme bilineari

Definizione 7.3.1. Sia V uno spazio vettoriale su un campo k. Una *forma bilineare* su V è una funzione bilineare

$$\begin{array}{ccc} V \times V & \overset{F}{\longrightarrow} & k \\ (v, w) & \mapsto & F(v, w) \end{array} \tag{7.3.1}$$

La forma bilineare F è *simmetrica* se $F(v, w) = F(w, v)$ per ogni $v, w \in V$.

Esempio 7.3.2. Scegliamo una unità di misura nel piano o nello spazio, cosicché sia ben definita la lunghezza $||v||$ di un vettore $v \in \mathcal{V}^2$ o $v \in \mathcal{V}^3$, dove $\mathcal{V}^2$, $\mathcal{V}^3$ sono rispettivamente lo spazio vettoriale reale dei vettori del piano, e lo spazio vettoriale reale dei vettori dello spazio. Su $\mathcal{V}^2$ e $\mathcal{V}^3$ si definisce il prodotto scalare ponendo

$$\langle v, w \rangle = ||v|| \cdot ||w|| \cdot \cos\theta, \tag{7.3.2}$$

dove θ è l'angolo tra v e w. Si può verificare con semplici argomenti geometrici che

$$\begin{array}{ccc} \mathcal{V}^2 \times \mathcal{V}^2 & \longrightarrow & \mathbb{R} \\ (v, w) & \mapsto & \langle v, w \rangle \end{array} \tag{7.3.3}$$

è una forma bilineare simmetrica, e analogamente per il prodotto scalare $\mathcal{V}^3 \times \mathcal{V}^3 \to \mathbb{R}$.

Esempio 7.3.3. L'applicazione

$$\begin{array}{ccc} k^2 \times k^2 & \overset{F}{\longrightarrow} & k \\ (A_1, A_2) & \mapsto & \operatorname{Det}[A_1, A_2] \end{array}$$

è una forma bilineare, ma *non* è simmetrica. (Si dice che F è *antisimmetrica* perché $\operatorname{Det}[A_1, A_2] = -\operatorname{Det}[A_2, A_1]$.)

Definizione 7.3.4. $\mathrm{Bil}(V)$ è l'insieme delle forme bilineari su V, e $\mathrm{Bil}_+(V) \subset \mathrm{Bil}(V)$ è il sottoinsieme delle forme bilineari simmetriche su V.

Proposizione 7.3.5. *Sia* $\mathrm{Bil}(V)$ *che* $\mathrm{Bil}_+(V)$ *sono sottospazi dello spazio vettoriale delle funzioni da* V^2 *a* k.

Dimostrazione. Siano $F, G \in \mathrm{Bil}(V)$ e $\lambda, \mu \in k$. L'applicazione

$$\begin{array}{ccc} V \times V & \stackrel{(\lambda F + \mu G)}{\longrightarrow} & k \\ (v, w) & \mapsto & \lambda F(v,w) + \mu G(v,w) \end{array} \tag{7.3.4}$$

è lineare in ciascuna delle variabili perché V^* è un sottospazio dello spazio vettoriale delle applicazioni da V in k. È evidente che se F e G sono simmetriche anche $\lambda F(v,w) + \mu G(v,w)$ è simmetrica. □

Siano V uno spazio vettoriale su k e $F \in \mathrm{Bil}(V)$. Sia $v \in V$: per bilinearità le applicazioni da V in k definite da $w \mapsto F(w,v)$ e da $w \mapsto F(v,w)$ sono lineari. Quindi possiamo definire due applicazioni da V a V^*:

$$\begin{array}{ccc} V & \stackrel{D_F}{\longrightarrow} & V^* \\ v & \mapsto & (w \mapsto F(w,v)) \end{array} \qquad \begin{array}{ccc} V & \stackrel{S_F}{\longrightarrow} & V^* \\ v & \mapsto & (w \mapsto F(v,w)). \end{array} \tag{7.3.5}$$

Proposizione 7.3.6. *Siano* V *uno spazio vettoriale su* k *e* $F \in \mathrm{Bil}(V)$. *Allora*

1. *sia* D_F *che* S_F *sono applicazioni lineari,*
2. *la duale di* D_F *è uguale a* S_F *e, viceversa, la duale di* S_F *è uguale a* D_F,
3. *la* F *è simmetrica se e solo se* $D_F = S_F$.

Dimostrazione. (1): D_F è un'applicazione lineare perché F è lineare a destra, e S_F è un'applicazione lineare perché F è lineare a sinistra. (2): Iniziamo spiegando il senso dell'affermazione che la duale di D_F è uguale a S_F. La duale $D_F^* \colon (V^*)^* \to V^*$ è identificata con un'applicazione (lineare) $V \to V^*$ grazie all'isomorfismo naturale $V \cong (V^*)^*$, e quindi ha senso affermare che $D_F^* = S_F$. Analogamente possiamo identificare S_F^* con un'applicazione lineare $V \to V^*$, e quindi ha senso affermare che $S_F^* = D_F$. Dimostriamo che $D_F^* = S_F$. Il valore di $D_F^*(u) \in V^*$ su z è uguale a $F(u,z)$, e d'altra parte il valore di $S_F(u) \in V^*$ su z è uguale a $F(u,z)$, cioè sono uguali. Analogoragionamento dimostra che $S_F^* = D_F$. (3): Ovvio. □

Definizione 7.3.7. Siano V uno spazio vettoriale su k e $F \in \mathrm{Bil}_+(V)$. Denotiamo $D_F = S_F$ con $\mathcal{L}_F \colon V \to V^\vee$.

7.3.2 Forme bilineari e matrici

Sia V uno spazio vettoriale finitamente generato su k, e $\mathcal{B} = \{v_1, \ldots, v_n\}$ una sua base. Sia $A \in M_{n,n}(k)$. L'applicazione

$$\begin{array}{ccc} V \times V & \stackrel{\Phi_A^{\mathcal{B}}}{\longrightarrow} & k \\ (v, w) & \mapsto & X_{\mathcal{B}}(v)^t \cdot A \cdot X_{\mathcal{B}}(w) \end{array} \tag{7.3.6}$$

è bilineare (facile verifica). Per referenza futura notiamo che

$$a_{ij} = \Phi_A^{\mathcal{B}}(v_i, v_j). \tag{7.3.7}$$

Siano $A, C \in M_{n,n}(k)$ e $\lambda \in k$. Un facile calcolo dà che

$$\Phi_A^{\mathcal{B}} + \Phi_C^{\mathcal{B}} = \Phi_{A+C}^{\mathcal{B}}, \qquad \lambda \Phi_A^{\mathcal{B}} = \Phi_{\lambda A}^{\mathcal{B}}$$

In altre parole l'applicazione

$$\begin{array}{ccc} M_{n,n}(k) & \xrightarrow{\Phi^{\mathcal{B}}} & \mathrm{Bil}(V) \\ A & \mapsto & \Phi_A^{\mathcal{B}} \end{array} \tag{7.3.8}$$

è lineare.

Definizione 7.3.8. $M_{n,n}^+(k) \subset M_{n,n}(k)$ è il sottospazio delle matrici simmetriche.

Proposizione 7.3.9. *Sia V uno spazio vettoriale finitamente generato su k e $\mathcal{B}$ una sua base. L'applicazione*

$$\begin{array}{ccc} M_{n,n}(k) & \xrightarrow{\Phi^{\mathcal{B}}} & \mathrm{Bil}(V) \\ A & \mapsto & \Phi_A^{\mathcal{B}} \end{array} \tag{7.3.9}$$

è un isomorfismo di spazi vettoriali. L'immagine di $M_{n,n}^+(k)$ è il sottospazio $\mathrm{Bil}_+(V) \subset \mathrm{Bil}(V)$ delle forme bilineari simmetriche.

Dimostrazione. Dimostriamo che $\Phi^{\mathcal{B}}$ è biunivoca. Supponiamo che $\Phi_A^{\mathcal{B}} = 0$. Per la (7.3.10) segue che $a_{ij} = 0$ per ogni $1 \leq i, j \leq n$ e quindi $A = 0$. Questo dimostra che $\Phi^{\mathcal{B}}$ è iniettiva. Ora dimostriamo che $\Phi^{\mathcal{B}}$ è suriettiva. Sia $F \in \mathrm{Bil}(V)$. Per $1 \leq i, j \leq n$ sia $a_{ij} := F(v_i, v_j)$. Si ha che

$$F\left(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j\right) = \sum_{1 \leq i,j \leq n} x_i y_j F(v_i, v_j) =$$

$$= \sum_{1 \leq i,j \leq n} a_{ij} x_i y_j = X^t \cdot A \cdot Y = \Phi_A^{\mathcal{B}}\left(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j\right).$$

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

$$a_{ij} = \Phi_A^{\mathcal{B}}(v_i, v_j) = \Phi_A^{\mathcal{B}}(v_j, v_i) = a_{ji} \tag{7.3.10}$$

e quindi $A^t = A$. D'altra parte se $A = A^t$ allora

$$\Phi_A^{\mathcal{B}}(v,w) = X_{\mathcal{B}}(v)^t \cdot A \cdot X_{\mathcal{B}}(w) = (X_{\mathcal{B}}(v)^t \cdot A \cdot X_{\mathcal{B}}(w)^t = X_{\mathcal{B}}(w)^t \cdot A^t \cdot X_{\mathcal{B}}(v) = X_{\mathcal{B}}(w)^t \cdot A \cdot X_{\mathcal{B}}(v) = \Phi_A^{\mathcal{B}}(w,v). \tag{7.3.11}$$

□

Sia V uno spazio vettoriale finitamente generato su k e $\mathcal{B}$ una sua base. Denoteremo con $M_{\mathcal{B}}$ l'inversa di $\Phi^{\mathcal{B}}$:

$$M_{\mathcal{B}} \colon \mathrm{Bil}(V) \xrightarrow{\sim} M_{n,n}(k). \tag{7.3.12}$$

Osservazione 7.3.10. Sia V uno spazio vettoriale finitamente generato su k e $\mathcal{B}$ una sua base. Se $F \in \mathrm{Bil}(V)$ allora $A = M_{\mathcal{B}}(F)$ è data da

$$a_{ij} = F(v_i, v_j). \tag{7.3.13}$$

Proposizione 7.3.11. *Sia V uno spazio vettoriale finitamente generato su k e $F \in \mathrm{Bil}(V)$. Siano $\mathcal{B}$ e $\mathcal{C}$ basi di V. Allora*

$$M_{\mathcal{C}}(F) = (M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}))^t \cdot M_{\mathcal{B}}(F) \cdot M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}). \tag{7.3.14}$$

Dimostrazione. Per ogni $v, w \in V$ abbiamo

$$X_{\mathcal{C}}(v)^t \cdot M_{\mathcal{C}}(F) \cdot X_{\mathcal{C}}(w) = F(v,w) = X_{\mathcal{B}}(v)^t \cdot M_{\mathcal{B}}(F) \cdot X_{\mathcal{B}}(w) = (M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}) \cdot X_{\mathcal{C}}(v))^t \cdot M_{\mathcal{B}}(F) \cdot M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}) \cdot X_{\mathcal{C}}(w) =$$
$$X_{\mathcal{C}}(v)^t \cdot (M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}))^t \cdot M_{\mathcal{B}}(F) \cdot M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{Id}) \cdot X_{\mathcal{C}}(w) \tag{7.3.15}$$

e la proposizione segue. □

La seguente definizione è motivata dalla **Proposizione 7.3.11**.

Definizione 7.3.12. Matrici $A, B \in M_{n,n}(k)$ sono *congruenti* se esiste $G \in \mathrm{GL}_n(k)$ tale che $A = G^t \cdot B \cdot G$.

Osservazione 7.3.13. Riferendoci alla **Proposizione 7.3.9** e alla **Proposizione 7.3.11**, notiamo che se $A \in M_{n,n}^+(k)$ allora ogni matrice congruente ad A è simmetrica.

7.3.3 Polarizzazione

Associeremo a una $F \in \mathrm{Bil}_+(V)$ una forma quadratica. Definiamo $q_F \colon V \to k$ così:

$$q_F(v) := \Phi(v, v). \tag{7.3.16}$$

Osservazione 7.3.14. Siano $v_1, \ldots, v_n \in V$; la bilinearità e la simmetria di F danno che

$$q_F(x_1 v_1 + \ldots + x_n v_n) = \sum_{1 \le i \le n} F(v_i, v_i) x_i^2 + 2 \sum_{1 \le i < j \le n} F(v_i, v_j) x_i x_j. \tag{7.3.17}$$

Quindi q_F è una forma quadratica.

Definizione 7.3.15. Siano V uno spazio vettoriale su k e $F \in \mathrm{Bil}_+(V)$: la forma quadratica *associata* a F è q_F.

Proposizione 7.3.16. *Sia V uno spazio vettoriale su k. L'applicazione*

$$\begin{array}{ccc} \mathrm{Bil}_+(V) & \longrightarrow & Q(V) \\ F & \mapsto & q_F \end{array} \tag{7.3.18}$$

è un isomorfismo di spazi vettoriali.

Dimostrazione. Segue dalla **Proposizione 7.3.9** e dalla **Sezione 7.1**. Osserviamo che, data $f \in Q(V)$, l'unica $F \in \mathrm{Bil}_+(V)$ tale che $q_F = f$ è data dalla formula

$$F(v, w) = 2^{-1}(f(v + w) - f(v) - f(w)). \tag{7.3.19}$$

Infatti per una tale F vale

$$f(v+w) = F(v+w, v+w) = F(v,v) + 2F(v,w) + F(w,w) = f(v) + 2F(v,w) + f(w). \tag{7.3.20}$$

□

In parole: se V è uno spazio vettoriale su k possiamo identificare forme bilineari **simmetriche** su V con forme quadratiche su V.

Definizione 7.3.17. Sia V uno spazio vettoriale su k. Data una forma quadratica $q \in Q(V)$ la sua controimmagine per l'applicazione (7.3.18), chiamamola F, è la forma bilineare simmetrica *associata* a q - diciamo anche che F è la *polarizzazione* di q. Se V è finitamente generato e $\mathcal{B}$ è una sua base porremo $M_{\mathcal{B}}(q) := M_{\mathcal{B}}(F)$.

Esempio 7.3.18. Siano k un campo e $B \in M_{n,n}(k)$ (non facciamo alcuna ipotesi su B). Sia $q \in Q(k^n)$ data da $q(X) := X^t \cdot B \cdot X$. La forma bilineare simmetrica associata a q è la F definita da

$$F(X,Y) = \frac{1}{2} X^t \cdot (B^t + B) \cdot Y. \tag{7.3.21}$$

Infatti F è una forma bilineare *simmetrica* perché $(B^t + B)^t = (B + B^t)$ e

$$F(X,X)=\tfrac{1}{2}X^t\cdot(B^t+B)\cdot X=\tfrac{1}{2}(X^t\cdot B^t\cdot X)+\tfrac{1}{2}(X^t\cdot B\cdot X)=\tfrac{1}{2}(X^t\cdot B^t\cdot X)^t+\tfrac{1}{2}(X^t\cdot B\cdot X)=X^t\cdot B\cdot X=q(X). \tag{7.3.22}$$

Osservazione 7.3.19. Applicando la **Proposizione 7.3.16** a $V = k^n$ otteniamo che se $A \in M^+_{n,n}(k)$ allora $X^t \cdot A \cdot X = 0$ per ogni $X \in k^n$ se e solo se $A = 0$. Infatti sia $\mathcal{S}$ la base standard di k^n e F la polarizzazione della forma quadratica $q^{\mathcal{S}}_A$ (data da $q_A(X) := X^t \cdot A \cdot X$). Allora $A = M_{\mathcal{S}}(F)$ e quindi per la **Proposizione 7.3.16** abbiamo che $q^{\mathcal{S}}_A = 0$ se e solo se $A = 0$. Notate che se $A \in M_{n,n}(k)$ non è simmetrica può accadere che $X^t \cdot A \cdot X = 0$ per ogni $X \in k^n$ senza che A sia nulla. Più precisamente l'**Esempio 7.3.18** e la **Proposizione 7.3.16** danno che $X^t \cdot A \cdot X = 0$ per ogni $X \in k^n$ se e solo se A è *anti*simmetrica.

Esempio 7.3.20. Scegliamo una unità di misura nel piano o nello spazio, cosicché sia ben definita la lunghezza $||v||$ di un vettore $v \in \mathcal{V}^2$ o $v \in \mathcal{V}^3$. Allora

$$\begin{array}{ccc} \mathcal{V}^2 & \xrightarrow{q} & \mathbb{R} \\ v & \mapsto & ||v||^2 \end{array} \tag{7.3.23}$$

è una forma quadratica. Infatti se $\{\mathbf{i}_1, \mathbf{i}_2\}$ è una base ortonormale di $\mathcal{V}^2$, cioè tale che $\mathbf{i}_1, \mathbf{i}_2$ hanno lunghezza 1 e sono perpendicolari, allora $q(x_1\mathbf{i}_1 + x_2\mathbf{i}_2) = x_1^2 + x_2^2$ per il Teorema di Pitagora. Analogamente, $q\colon \mathcal{V}^3 \to \mathbb{R}$ definita da $q(v) := ||v||^2$ è una forma quadratica. In entrambi i casi la polarizzazione di q è uguale al prodotto scalare richiamato nell'**Esempio 7.3.2**.

Osservazione 7.3.21. Sia $f \in Q(V)$ e sia $F \in \mathrm{Bil}_+(V)$ la forma bilineare simmetrica associata a f. La F dà le derivate direzionali di f. Infatti sia $t \in k$: per la (7.3.20) abbiamo che

$$f(v + tw) - f(v) = 2tF(v,w) + t^2 f(w). \tag{7.3.24}$$

Ora per semplicità supponiamo che $k = \mathbb{R}$. Allora la (7.3.24) dà che la derivata di f nel punto v e nella direzione w è uguale a $2F(v,w)$.

7.3.4 Prodotti scalari

Definizione 7.3.22. Sia V uno spazio vettoriale su k. Una $F \in \mathrm{Bil}(V)$ è *non-degenere* se $D_F\colon V \to V^\vee$ e $S_F\colon V \to V^\vee$ sono iniettive.

Osservazione 7.3.23. Sia V uno spazio vettoriale su k. Una $F \in \mathrm{Bil}(V)$ è *non-degenere* se e solo se $\mathrm{Det}(M_{\mathcal{B}}(F)) \neq 0$.

Definizione 7.3.24. Sia V uno spazio vettoriale *reale*. Un *prodotto scalare* su V è una $F \in \mathrm{Bil}_+(V)$ non-degenere. Siano $v, w \in V$: si pone $\langle v, w\rangle := F(v, w)$.

Definizione 7.3.25. Il *prodotto scalare standard* su $\mathbb{R}^n$ è quello definito da (7.3.21). Analogamente il *prodotto scalare standard* su $\mathcal{V}^2$ o $\mathcal{V}^3$ è quello dell' **Esempio 7.3.2** (è definito a meno di uno scalare perché dobbiamo scegliere una unità di misura).

Definizione 7.3.26. Siano V uno spazio vettoriale *reale* e $q\colon V \to \mathbb{R}$ una forma quadratica. La q è *definita positiva* se $q(v) > 0$ per ogni $0 \neq v \in V$, è *definita negativa* se $q(v) < 0$ per ogni $0 \neq v \in V$ (ovvero se $-f$ è definita positiva). In simboli: $q > 0$ siginifica che q è definita positiva e $q < 0$ siginifica che q 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 $q \in Q(V)$ (o $F \in \mathrm{Bil}_+(V)$) è *definita* se è definita positiva o è definita negativa.

Osservazione 7.3.27. Siano V uno spazio vettoriale *reale* e $q\colon V \to \mathbb{R}$ una forma quadratica definita. Allora la polarizzazione F di q è una forma bilineare non-degenere, e come di consueto si pone $\langle v, w\rangle = F(v, w)$. Se q è definita positiva la *norma* di $v \in V$ è definita da

$$||v|| := q(v)^{1/2} = \langle v, v\rangle^{1/2} = F(v, w)^{1/2}. \tag{7.3.25}$$

Il prodotto scalare standard su $\mathbb{R}^n$ è uno degli archetipi di prodotto scalare, l'altro è quello dell'**Esempio 7.3.2**.

Definizione 7.3.28. 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[1].

Osservazione 7.3.29. $A \in M^+_{n,n}(\mathbb{R})$ è definita positiva se e solo se $X^t \cdot A \cdot X > 0$ per ogni vettore colonna $X \in M_{n,1}(\mathbb{R})$ non nullo, analogamente $A \in M^+_{n,n}(\mathbb{R})$ è definita negativa se e solo se $X^t \cdot A \cdot X < 0$ per ogni vettore colonna $X \in M_{n,1}(\mathbb{R})$ non nullo.

7.4 Ortogonalità

Definizione 7.4.1. Sia V uno spazio vettoriale su k, . Sia $F \in \mathrm{Bil}_+(V)$. I vettori $v, w \in V$ sono *perpendicolari* se $F(v, w) = 0$ - in simboli $v \perp w$. Se $S \subset V$ l'*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 $F(w, v) = 0$. Se F fosse una forma bilineare arbitraria dovremmo considerare l'ortogonale destro e l'ortogonale sinistro.

Osservazione 7.4.2. Consideriamo $\mathcal{V}^2$ (o $\mathcal{V}^3$) con il prodotto scalare standard - vedi **Definizione 7.3.25**. Siano $\overrightarrow{PQ}, \overrightarrow{QR} \in \mathcal{V}^2$ vettori non nulli. Allora $\overrightarrow{PQ} \perp \overrightarrow{QR}$ se e solo se la retta PQ è perpendicolare a alla retta QR. Infatti sia $\{\mathbf{i}_1, \mathbf{i}_2\}$ una base ortonormale di $\mathcal{V}^2$ e $\overrightarrow{PQ} = x_1\mathbf{i}_1 + x_2\mathbf{i}_2$, $\overrightarrow{QR} = y_1\mathbf{i}_1 + y_2\mathbf{i}_2$. 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 $0 = x_1y_1 + x_2y_2 = \langle x_1\mathbf{i}_1 + x_2\mathbf{i}_2, y_1\mathbf{i}_1 + y_2\mathbf{i}_2\rangle$.

[1]Notate che se $A \in M^+_{n,n}(\mathbb{R})$ è definita positiva allora $q_A^{\mathcal{B}}$ è definita positiva qualsiasi sia la base $\mathcal{B}$, e analogamente per $A \in M^+_{n,n}(\mathbb{R})$ definita negativa.

Esempio 7.4.3. Sia V uno spazio vettoriale su k e $F \in \mathrm{Bil}_+(V)$. Allora $V^\perp = \ker \mathcal{L}_F$: lo denoteremo $\ker F$.

Lemma 7.4.4. *Sia V uno spazio vettoriale su 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

$$S^\perp = \bigcap_{v \in S} v^\perp \tag{7.4.1}$$

segue che $S^\perp$ è intersezione di sottospazi lineari e perciò è un sottospazio lineare. □

Lemma 7.4.5. *Sia V uno spazio vettoriale su k e $F \in \mathrm{Bil}_+(V)$. Sia $U \subset V$ un sottospazio finitamente generato e siano $u_1, \dots, u_m$ generatori di U. Allora*

$$U^\perp = \bigcap_{i=1}^{m} u_i^\perp \tag{7.4.2}$$

Dimostrazione. È ovvio che il membro di sinistra di (7.4.2) è contenuto nel membro di destra. Resta da dimostrare che il membro di destra di (7.4.2) è contenuto nel membro di sinistra. Supponiamo che $v \in u_i^\perp$ per $1 \leq i \leq m$. Sia $u \in U$: siccome U è generato da $u_1, \dots, u_m$ esistono $\lambda_1, \lambda_m \in k$ tali che $u = \sum_{i=1}^m \lambda_i u_i$. Per linearità di F abbiamo che

$$F(v, u) = F(v, \sum_{i=1}^{m} \lambda_i u_i) = \sum_{i=1}^{m} \lambda_i F(v, u_i) = 0.$$

□

Osservazione 7.4.6. Sia V uno spazio vettoriale finitamente generato su k e $F \in \mathrm{Bil}_+(V)$. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V e $A := M_{\mathcal{B}}(F)$. Sia $w \in V$: allora

$$w^\perp = \{v \in V \mid X_{\mathcal{B}}(w)^t \cdot A \cdot X_{\mathcal{B}}(v) = 0\}. \tag{7.4.3}$$

Proposizione 7.4.7. *Sia V uno spazio vettoriale finitamente generato su k e $F \in \mathrm{Bil}_+(V)$ non degenere. Sia $U \subset V$ un sottospazio. Allora*

$$\dim U^\perp = \dim V - \dim U. \tag{7.4.4}$$

Dimostrazione. Sia $\{u_1, \dots, u_m\}$ una base di U. Per il **Lemma 7.4.5**

$$U^\perp = \bigcap_{i=1}^{m} u_i^\perp = \bigcap_{i=1}^{m} \ker \mathcal{L}_F(u_i). \tag{7.4.5}$$

Siccome F è non-degenere, $\mathcal{L}_F(u_1), \dots, \mathcal{L}_F(u_m)$ sono elementi di V^* lindamente indipendenti, e quindi il membro di destra di (7.4.5) ha dimensione uguale a $\dim V - m$. □

7.5 Diagonalizzazione

Data una forma quadratica f su uno spazio vettoriale V ci chiediamo se esiste una base $\mathcal{B}$ di V tale che valga (7.2.6).

Definizione 7.5.1. Sia V uno spazio vettoriale finitamente generato su k e $\mathcal{B} = \{v_1, \ldots, v_n\}$ una sua base. Una forma quadratica $f\colon V \to k$ è *diagonale* nella base $\mathcal{B}$ se esistono $c_1, \ldots, c_n \in k$ tali che

$$f(x_1v_1 + \ldots + x_nv_n) = \sum_{i=1}^{n} c_i x_i^2 \qquad \forall\ (x_1v_1 + \ldots + x_nv_n) \in V. \tag{7.5.1}$$

Una forma bilineare $F\colon V \times V \to k$ 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 che la base $\mathcal{B}$ *diagonalizza* F.

Osservazione 7.5.2. Sia V uno spazio vettoriale finitamente generato su k e $\mathcal{B} = \{v_1, \ldots, v_n\}$ una sua base. Sia $F\colon V \times V \to k$ una forma bilineare *simmetrica* e q_F la forma quadratica associata. Allora F è diagonale nella base $\mathcal{B}$ se e solo se lo è q_F, è sufficiente guardare a (7.3.17) e (**??**).

Osservazione 7.5.3. Se $F\colon V \times V \to k$ è una forma bilineare arbitraria e definiamo $f\colon V \to k$ come in (7.3.16) la f può essere diagonale in una base $\mathcal{B}$ anche se la F non lo è. Per esempio sia F *anti-simmetrica* cioè $F(v, w) = -F(w, v)$ per ogni $v, w \in V$: allora $f(v) = 0$ per ogni $v \in V$ e quindi f è diagonale in qualsiasi base, mentre F sarà diagonale in una base solo se è nulla.

Teorema 7.5.4. *Sia V uno spazio vettoriale finitamente generato su k. Sia $F\colon V \times V \to k$ una forma bilineare simmetrica. Esiste una base $\mathcal{B}$ di V che diagonalizza F. Se $k = \mathbb{R}$ possiamo assumere che le entrate di $M_{\mathcal{B}}(F)$ appartengano a $\{-1, 0, 1\}$. Se $k = \mathbb{C}$ possiamo assumere che le entrate di $M_{\mathcal{B}}(F)$ appartengano a $\{0, 1\}$.*

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

$$\begin{array}{ccc} U \times U & \xrightarrow{G} & k \\ (u_1, u_2) & \mapsto & F(u_1, u_2) \end{array} \tag{7.5.2}$$

è bilineare simmetrica. Siccome $\dim U = (n-1)$ (perché $v_0 \neq 0$) 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 7.3.10** mostra che $\mathcal{B}$ diagonalizza F. Abbiamo dimostrato la prima affermazione del teorema. Se $k = \mathbb{R}$ o $k = \mathbb{C}$ procediamo di nuovo per induzione su n. Se $\dim V = 0$ non c'è nulla da dimostrare. Per il passo induttivo procediamo come sopra, in aggiunta notiamo che siccome ogni reale positivo ha una radice quadrata e ogni numero complesso ha una radice quadrata possiamo riscalare v_0 nel caso reale in modo tale che $q_F(v_0) \in \{0, \pm 1\}$ e nel caso complesso in modo che $q_F(v_0) \in \{0, 1\}$. □

Corollario 7.5.5. *Sia $A \in M^{+}_{n,n}(k)$. Allora A è congruente a una matrice diagonale Λ. Se $k = \mathbb{C}$ possiamo assumere che le entrate di Λ appartengano a $\{0, 1\}$, $k = \mathbb{R}$ che appartengano a $\{0, \pm 1\}$.*

Dimostrazione. Segue dal **Teorema 7.5.4** e dalla **Proposizione 7.3.11**. □

Corollario 7.5.6 (Lagrange)**.** *Sia V uno spazio vettoriale finitamente generato su k. Sia $f\colon V \to k$ una forma quadratica. Esiste una base di V che diagonalizza f. Se $k = \mathbb{R}$ possiamo assumere che i c_i che appaiono in* (7.5.1) *appartengano a $\{-1, 0, 1\}$. Se $k = \mathbb{C}$ possiamo assumere che i c_i che appaiono in* (7.5.1) *appartengano a $\{0, 1\}$.*

Dimostrazione. Segue dal **Teorema 7.5.4** e dall'**Osservazione 7.5.2**. □

Sia $f\colon V \to k$ una forma quadratica su uno spazio vettoriale finitamente generato su 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 $c_n \neq 0$)

$$f(z_1u_1 + \ldots + z_nu_n) = \sum_{i\leq i,j\leq(n-1)} a_{ij}x_ix_j + c_n(z_n^2 + 2c_n^{-1}\sum_{i=1}^{n-1} a_{in}z_iz_n) =$$

$$= \sum_{i\leq i,j\leq(n-1)} a_{ij}x_ix_j + c_n\left(z_n + c_n^{-1}\sum_{i=1}^{n-1} a_{in}z_i\right)^2 - c_n^{-1}\left(\sum_{i=1}^{n-1} a_{in}z_i\right)^2. \quad (7.5.3)$$

Sia

$$r := \sum_{i\leq i,j\leq(n-1)} a_{ij}x_ix_j - c_n^{-1}\left(\sum_{i=1}^{n-1} a_{in}z_i\right)^2.$$

Notate che $r \in k[x_1, \ldots, x_{n-1}]$ è un polinomio omogeneo di grado 2. La (7.5.3) si può riscrivere così:

$$f(z_1u_1 + \ldots + z_nu_n) = r(z_1, \ldots, z_{n-1}) + c_n(z_n + c_n^{-1}\sum_{i=1}^{n-1} a_{in}z_i)^2. \quad (7.5.4)$$

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

$$y_i = z_i, \quad 1 \leq i \leq (n-1), \qquad y_n = z_n + c_n^{-1}\sum_{i=1}^{n-1} a_{in}z_i. \quad (7.5.5)$$

(Questo perché la matrice associata all'applicazione lineare definita da (7.5.5) 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 $k = \mathbb{R}$. Se $c_i \neq 0$ sostituiamo a v_i il vettore $|c_i|^{-1/2}v_i$. Se $k = \mathbb{C}$ sostituiamo a ogni v_i tale che $c_i \neq 0$ il vettore $c_i^{-1/2}v_i$.

Esempio 7.5.7. 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

$$f(y_1u_1 + y_2u_2 + y_3u_3) = f(y_1, y_2 + y_3, y_3) = y_1y_2 + 2y_1y_3 + y_2y_3 + y_3^2 = \\ y_1y_2 + (y_3 + y_1 + \frac{1}{2}y_2)^2 - (y_1 + \frac{1}{2}y_2)^2 = -y_1^2 - \frac{1}{4}y_2^2 + (y_3 + y_1 + \frac{1}{2}y_2)^2. \quad (7.5.6)$$

Siano (z_1, z_2, z_3) le coordinate su $\mathbb{R}^3$ date da

$$\begin{aligned} z_1 &= y_1 \\ z_2 &= y_2 \\ z_3 &= y_1 + \tfrac{1}{2}y_2 + y_3 \end{aligned}$$

Quindi

$$\begin{aligned} y_1 &= z_1 \\ y_2 &= z_2 \\ y_3 &= -z_1 - \tfrac{1}{2}z_2 + z_3 \end{aligned}$$

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

$$\begin{aligned} t_1 &= z_1 \\ t_2 &= z_2/2 \\ t_3 &= z_3 \end{aligned}$$

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

$$f(t_1r_1 + t_2r_2 + t_3r_3) = -t_1^2 - t_2^2 + t_3^2.$$

7.6 Spazi vettoriali quadratici

Definizione 7.6.1. Uno spazio (vettoriale) *quadratico* su k è una coppia (V, q) dove V è uno spazio vettoriale su k e $q \in Q(V)$. Siano (V, q) e (W, r) spazi quadratici su k. Un *isomorfismo* $(V, q) \xrightarrow{\sim} (W, r)$ è un isomorfismo $\phi\colon V \to W$ di spazi vettoriali tale che

$$r(\phi(v)) = q(v) \quad \forall v \in V. \quad (7.6.1)$$

Un *automorfismo* di (V, q) è un isomorfismo $(V, q) \xrightarrow{\sim} (V, q)$. Diciamo che (V, q) è *isomorfo* a (W, r) se esiste un isomorfismo $(V, q) \xrightarrow{\sim} (W, r)$.

Esempio 7.6.2. Siano $f, g, h \in Q(\mathbb{R}^2)$ date da

$$f(x, y) = x^2 - y^2, \quad g(x, y) = xy, \quad h(x, y) := x^2 + y^2. \quad (7.6.2)$$

Sia

$$\begin{array}{ccc} \mathbb{R}^2 & \xrightarrow{\phi} & \mathbb{R}^2 \\ (x, y) & \mapsto & (x + y, x - y) \end{array}$$

Allora ϕ è un isomorfismo $(\mathbb{R}^2, f) \xrightarrow{\sim} (\mathbb{R}^2, g)$. Siccome ϕ è un isomorfismo segue che g e h sono congruenti. D'altra parte f e h non sono isomorfi. Infatti $f(1,1) = 0$. Se f e h fossero congruenti esisterebbe un isoomorfismo $\phi\colon \mathbb{R}^2 \to \mathbb{R}^2$ tale che $h(\phi(x,y)) = f(x,y)$ e quindi $h(\phi(1,1)) = f(1,1) = 0$. Siccome ϕ è un isomorfismo $\phi(1,1) \neq (0,0)$ e quindi $h(\phi(1,1)) > 0$: contraddizione.

Esempio 7.6.3. Sia $\mathcal{V}^2$ con la forma quadratica q dell'**Esempio 7.3.20** (quadrato della lunghezza di un vettore). Un isomorfismo $\phi\colon \mathcal{V}^2 \xrightarrow{\sim} \mathcal{V}^2$ è un automorfismo di $(\mathcal{V}^2, q)$ se e solo se conserva le lunghezze dei vettori. Quindi una rotazione è un automorfismo di $(\mathcal{V}^2, q)$.

Osservazione 7.6.4. Si verifica facilmente che la relazione di isomorfismo tra spazi quadratici è di equivalenza.

Osservazione 7.6.5. Siano (V, q) e (W, r) spazi quadratici su k e siano $F \in \mathrm{Bil}_+(V)$ e $G \in \mathrm{Bil}_+(W)$ le polarizzazioni di q e r rispettivamente.

1. Supponiamo che $\phi\colon V \to W$ sia un isomorfismo di spazi quadratici. Allora
$$G(\phi(v), \phi(w)) = F(v, w) \quad \forall (v, w) \in V \times V. \tag{7.6.3}$$
2. Supponiamo che $\phi\colon V \to W$ sia un isomorfismo di spazi vettoriali tale che valga (7.6.3). Allora ϕ è un isomorfismo di spazi quadratici.

Osservazione 7.6.6. Sia V uno spazio vettoriale e $F \in \mathrm{Bil}_+(V)$. Denoteremo con (V, F) lo spazio quadratico (V, q_F). Siano W uno spazio vettoriale e $G \in \mathrm{Bil}_+(W)$. Per l'**Osservazione 7.6.5** un isomorfismo $\phi\colon V \to W$ definisce un isomorfismo di spazi vettoriali quadratici $(V, F) \xrightarrow{\sim} (W, G)$ se e solo se vale (7.6.3).

Proposizione 7.6.7. *Siano V e W spazi vettoriale finitamente generati della stessa dimensione n. Siano $q \in Q(V)$ e $r \in Q(W)$. Siano $\mathcal{B}$ e $\mathcal{C}$ basi di V e W rispettivamente. Sia $\phi\colon V \xrightarrow{\sim} W$ un isomorfismo di spazi vettoriali. Allora ϕ è un isomorfismo di spazi quadratici se e solo se*
$$M_{\mathcal{B}}(q) = (M_{\mathcal{C}}^{\mathcal{B}}(\phi))^t \cdot M_{\mathcal{C}}(r) \cdot M_{\mathcal{C}}^{\mathcal{B}}(\phi). \tag{7.6.4}$$

Dimostrazione. Supponiamo che ϕ sia un isomorfismo di spazi quadratici. Allora
$$X_{\mathcal{B}}(v)^t \cdot M_{\mathcal{B}}(q) \cdot X_{\mathcal{B}}(v) = (M_{\mathcal{C}}^{\mathcal{B}}(\phi) \cdot X_{\mathcal{B}}(v))^t \cdot M_{\mathcal{C}}(r) \cdot M_{\mathcal{C}}^{\mathcal{B}}(\phi) \cdot X_{\mathcal{B}}(v) \tag{7.6.5}$$
e quindi
$$X^t \cdot \left(M_{\mathcal{B}}(q) - M_{\mathcal{C}}^{\mathcal{B}}(\phi)^t \cdot M_{\mathcal{C}}(r) \cdot M_{\mathcal{C}}^{\mathcal{B}}(\phi)\right) \cdot X = 0 \quad \forall X \in k^n. \tag{7.6.6}$$
Per l'**Osservazione 7.3.19** segue che vale (7.6.4). Lo stesso calcolo dà che vale il viceversa. □

Esempio 7.6.8. Sia (V, q) uno spazio quadratico su k. Il *gruppo ortogonale* di (V, q) è
$$\mathrm{O}(V, q) := \{\phi\colon V \xrightarrow{\sim} V \mid \phi \text{ è un automorfismo di } (V, q)\}. \tag{7.6.7}$$
Supponiamo che $V = \mathbb{R}^n$ e che q sia la forma quadratica $q(X) := X^t \cdot X$ (con polarizzazione il prodotto scalare standard, vedi **Definizione 7.3.25**) - in questo caso si pone $\mathrm{O}_n(\mathbb{R}) := \mathrm{O}(\mathbb{R}^n, q)$. Per la **Proposizione 7.6.7** abbiamo che
$$\mathrm{O}_n(\mathbb{R}) = \{A \in \mathrm{GL}_n(\mathbb{R}) \mid A^t \cdot A = 1_n\}. \tag{7.6.8}$$
In altre parole $A \in \mathrm{O}_n(\mathbb{R})$ se e solo se le colonne di A formano una base ortonormale di $\mathbb{R}^n$ (rispetto al prodotto scalare standard), cioè $(A_i, A_j) = \delta_{ij}$ per $1 \leq i, j \leq n$.

Problema 7.6.9. Siano (V,q), (W,r) spazi quadratici. Come facciamo a decidere se (V,q) è isomorfo a (W,r) ? La difficoltà del problema dipende dal campo k. Daremo una soluzione nel caso in cui k sia $\mathbb{R}$ o $\mathbb{C}$. Cominciamo con il definire un invariante delle forme quadratiche.

Definizione 7.6.10. Sia V uno spazio vettoriale finitamente generato. Sia $q \in Q(V)$ e $F \in \mathrm{Bil}_+(V)$ la forma bilineare simmetrica associata a q. Il *rango* di q è il rango di $\mathcal{L}_F$: lo denotiamo $r(q)$.

Osservazione 7.6.11. Siano (V,f) e (W,g) spazi quadratici finitamente generati. Se (V,f) è isomorfo a (W,g) allora $r(f) = r(g)$. (Equivalentemente: se $r(f) \neq r(g)$ allora (V,f) non è isomorfo a (W,g).) Infatti siano $\mathcal{B}$ e $\mathcal{C}$ basi di V e W rispettivamente. Allora $r(f) = \mathrm{rk}\, M_{\mathcal{B}}(f)$ e $r(g) = \mathrm{rk}\, M_{\mathcal{C}}(g)$. D'altra parte $M_{\mathcal{B}}(f)$ e $M_{\mathcal{C}}(g)$ sono congruenti per la **Proposizione 7.6.7** e si verifica facilmente che matricic congruenti hanno lo stesso rango.

Proposizione 7.6.12. *Siano (V,f) e (W,g) spazio quadratici* complessi. *Allora (V,f) è isomorfo a (W,g) se e solo se $\dim V = \dim W$ e $r(f) = r(g)$.*

Dimostrazione. Se (V,f) è isomorfo a (W,g) allora $\dim V = \dim W$ e $r(f) = r(g)$ - vedi **Osservazione 7.6.11**. Ora supponiamo che $\dim V = \dim W$ e $r(f) = r(g)$. Dimostriamo che (V,f) è isomorfo a (W,g). Per il **Corollario 7.5.6** esistono basi $\mathcal{B}$ e $\mathcal{C}$ di V tali che $M_{\mathcal{B}}(f)$ e $M_{\mathcal{C}}(g)$ sono matrici diagonali con entrate non nulle uguali a 1. Le matrici $M_{\mathcal{B}}(f), M_{\mathcal{C}}(g)$ hanno lo stesso numero di righe/colonne perché $\dim V = \dim W$ e, siccome $r(f) = r(g)$, hanno lo stesso numero di entrate non nulle (il rango di una matrice diagonale è uguale al numero di entrate non nulle). Segue che, riordinando i vettori delle basi se necessario, si ha $M_{\mathcal{B}}(f) = M_{\mathcal{C}}(g)$ e quindi (V,f) è isomorfo a (W,g) per la **Proposizione 7.6.7**. □

Ora affrontiamo il **Problema 7.6.9** nel caso in cui $k = \mathbb{R}$. Non è vero l'analogo della **Proposizione 7.6.12**. Per esempio le forme quadratiche reali f, h su $\mathbb{R}^2$ della (7.6.2) hanno rango 2 ma $(\mathbb{R}^2, f)$ *non* è isomorfo a $(\mathbb{R}^2, h)$.

Definizione 7.6.13. Sia V uno spazio vettoriale **reale** . Denotiamo con $s_+(f)$ la massima dimensione di un sottospazio $U \subset V$ tale che $f|_U$ è definita positiva, e con $s_-(f)$ la massima dimensione di un sottospazio $U \subset V$ tale che $f|_U$ è definita negativa. La *segnatura* di f è

$$s(f) := s_+(f) - s_-(f).$$

Osservazione 7.6.14. Siano (V,f) e (W,g) uno spazi quadratici *reali*. Se (V,f) è isomorfo a (W,g) allora $s_+(f) = s_+(g)$, $s_-(f) = s_-(g)$ e $s(f) = s(g)$. Infatti sia $\phi\colon V \xrightarrow{\sim} W$ un isomorfismo di spazi quadratici. L'uguaglianza $s_+(f) = s_+(g)$ segue dall'osservazione che abbiamo una corrispondenza biunivoca

$$\begin{array}{ccc} \{U \subset V | U \text{ è un sottospazio di } V \text{ e } (f|_U) > 0\} & \xrightarrow{\sim} & \{P \subset W | P \text{ è un sottospazio di } W \text{ e } (g|_P) > 0\} \\ U & \mapsto & \phi(U). \end{array} \tag{7.6.9}$$

Si dimostra in modo analogo che $s_-(f) = s_-(g)$. L'uguaglianza $s(f) = s(g)$ segue immediatamente dalle uguaglianze $s_+(f) = s_+(g)$ e $s_-(f) = s_-(g)$.

Proposizione 7.6.15 (Sylvester)**.** *Sia V uno spazio vettoriale reale e $q\colon V \to \mathbb{R}$ una forma quadratica. Supponiamo che $\mathcal{B} = \{v_1, \ldots, v_n\}$ sia una base di V tale che*

$$q(x_1v_1 + \ldots + x_nv_n) = c_1x_1^2 + \ldots + c_ax_a^2 - d_{a+1}x_{a+1}^2 - \ldots - d_{a+b}x_{a+b}^2. \qquad \forall\, (x_1v_1 + \ldots + x_nv_n) \in V. \tag{7.6.10}$$

Supponiamo che $c_i > 0$ per ogni $1 \leq i \leq a$ e che $d_i > 0$ per ogni $a+1 \leq i \leq a+b$. Allora $s_+(q) = a$, $s_-(q) = b$ e quindi $s(q) = a - b$.

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

$$\dim(V_+ + V_0) = n - b, \qquad \dim(V_- + V_0) = n - a. \tag{7.6.11}$$

Siccome $q|_{V_+} > 0$ e $q|_{V_-} < 0$ abbiamo

$$s_+(q) \geq a, \qquad s_-(q) \geq b. \tag{7.6.12}$$

Supponiamo che la prima diseguaglianza sia stretta cioè $s_+(q) > a$; arriveremo a un assurdo. Per definizione esiste un sottospazio $U \subset V$ tale che $\dim U > a$ e $q|_U > 0$. Per (7.6.11) la formula di Grassmann dà che

$$\begin{aligned}\dim(U \cap (V_- + V_0)) = \dim U + \dim(V_- + V_0) - \dim(U + V_- + V_0) \geq \\ \geq \dim U + \dim(V_- + V_0) - n = \dim U - a > 0.\end{aligned}$$

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 (7.6.10) che $q(v) \leq 0$. D'altra parte $v \in U$ e per ipotesi $q|_U > 0$, quindi $q(v) > 0$. La contraddizione dimostra che non esiste un sottospazio $U \subset V$ tale che $\dim U > a$ e $q|_U > 0$; per (7.6.12) segue che $s_+(q) = a$. Si dimostra in modo analogo che non puè essere $s_-(q) > b$ e quindi $s_-(q) = b$. □

Corollario 7.6.16. *Sia V uno spazio vettoriale* **reale** *e $f \in Q(V)$. Sia $\mathcal{B}$ una base di V che diagonalizza f, cioè vale* (7.5.1)*, e supponiamo che $c_i \in \{\pm 1, 0\}$ - vedi il* **Corollario 7.5.6**. *Allora*

$$|\{i \mid c_i = 1\}| = (r(q) + s(q))/2, \qquad |\{j \mid c_j = -1\}| = (r(q) - s(q))/2. \tag{7.6.13}$$

Dimostrazione. Siano $a := |\{i \mid c_i = 1\}|$ e $b := |\{i \mid c_i = -1\}|$. Allora $r(q) = a + b$ e per la **Proposizione 7.6.15** $s(q) = a - b$. Sommando e sottraendo le due uguaglianze si ottiene il corollario. □

Proposizione 7.6.17. *Siano (V, f) e (W, g) spazi quadratici* reali. *Allora (V, f) è isomorfo a (W, g) se e solo se $\dim V = \dim W$, $r(f) = r(g)$ e $s(f) = s(g)$.*

Dimostrazione. Se (V, f) è isomorfo a (W, g) allora $\dim V = \dim W$, $r(f) = r(g)$ e $s(f) = s(g)$ - vedi l' **Osservazione 7.6.11** e l' **Osservazione 7.6.14**. Ora supponiamo che $\dim V = \dim W$, $r(f) = r(g)$ e $s(f) = s(g)$. Dimostriamo che (V, f) è isomorfo a (W, g). Per il **Corollario 7.5.6** esistono basi $\mathcal{B}$ e $\mathcal{C}$ di V tali che $M_{\mathcal{B}}(f)$ e $M_{\mathcal{C}}(g)$ sono matrici diagonali con entrate non nulle uguali a ± 1. Le matrici $M_{\mathcal{B}}(f), M_{\mathcal{C}}(g)$ hanno lo stesso numero di righe/colonne perché $\dim V = \dim W$. Dall'ipotesi che $r(f) = r(g)$ e $s(f) = s(g)$ e dal **Corollario 7.6.16** segue che il numero di entrate uguali a 1 di $M_{\mathcal{B}}(f)$ è uguale al numero di entrate uguali a 1 di $M_{\mathcal{C}}(g)$ e che il numero di entrate uguali a -1 di $M_{\mathcal{B}}(f)$ è uguale al numero di entrate uguali a -1 di $M_{\mathcal{C}}(g)$. Quindi, riordinando i vettori delle basi $\mathcal{B}$ e $\mathcal{C}$, segue che $M_{\mathcal{B}}(f) = M_{\mathcal{C}}(g)$ e per la **Proposizione 7.6.7** otteniamo che (V, f) è isomorfo a (W, g). □

7.7 Spazi vettoriali euclidei

Uno *spazio vettoriale euclideo* è uno spazio quadratico *reale* (V, q) tale che q è definita positiva. Denoteremo

$$\begin{array}{ccc} V \times V & \longrightarrow & \mathbb{R} \\ (v, w) & \mapsto & \langle v, w \rangle \end{array} \tag{7.7.1}$$

la polarizzazione di q. Quindi $\langle , \rangle$ è un prodotto scalare definito positivo. Spesso definiremo q via $\langle , \rangle$, quindi denoteremo (V, q) con $(V, \langle , \rangle)$. Esempi da tenere in mente: il prodotto scalare standard su $\mathbb{R}^n$ - vedi (7.3.21) - e il prodotto scalare "classico" tra vettori geometrici - vedi l' **Esempio 7.3.2**. Un altro esempio interessante è il seguente.

Esempio 7.7.1. L'**Esempio ??** definisce un prodotto scalare definito positivo su $C^0([-\pi, \pi])$.

Ricordiamo che in uno spazio vettoriale euclideo $(V, \langle , \rangle)$ la norma di $v \in V$ è definita da $||v|| = \langle v, v\rangle^{1/2}$ - vedi (7.3.25).

Teorema 7.7.2 (Diseguaglianza di Cauchy-Schwarz)**.** *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. Siano $v, w \in V$: si ha che*

$$\langle v, w\rangle^2 \leq ||v||^2 \cdot ||w||^2. \tag{7.7.2}$$

Dimostrazione. Sia $x \in \mathbb{R}$: siccome $\langle , \rangle$ è definito positivo abbiamo che

$$p(x) := ||v||^2 x^2 + 2\langle v, w\rangle x + ||w||^2 = \langle xv + w, xv + w\rangle \geq 0.$$

Segue che il polinomio p ha al più una radice reale e perciò

$$(2\langle v, w\rangle)^2 - 4||v||^2 \cdot ||w||^2 \leq 0.$$

Segue che vale la (7.7.2). □

Corollario 7.7.3 (Diseguaglianaza triangolare)**.** *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. Siano $v, w \in V$: si ha che*

$$||v + w|| \leq ||v|| + ||w||. \tag{7.7.3}$$

Dimostrazione. Per la diseguaglianza di Cauchy-Schwartz abbiamo che

$$||v + w||^2 = ||v||^2 + 2\langle v, w\rangle + ||w||^2 \leq ||v||^2 + 2||v|| \cdot ||w|| + ||w||^2 = (||v|| + ||w||)^2. \tag{7.7.4}$$

Segue il corollario. □

Mostreremo che si può definire l'angolo tra vettori di un qualsiasi spazio vettoriale con prodotto scalare definito positivo. La diseguaglianza di Cauchy-Schwarz dà che se $v \neq 0 \neq w$ allora

$$-1 \leq \frac{\langle v, w\rangle}{||v|| \cdot ||w||} \leq 1. \tag{7.7.5}$$

Definizione 7.7.4. Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo. Siano $v, w \in V$ **non nulli**. L'angolo tra v e w è l'unico $0 \leq \theta \leq \pi$ tale che

$$\cos\theta = \frac{\langle v, w\rangle}{||v|| \cdot ||w||}. \tag{7.7.6}$$

Notate che la definizione ha senso per la (7.7.5).

Notate che l'angolo tra v e w non cambia se riscaliamo v o w (quindi è definito l'angolo tra "semirette") e che non dipende dall'ordine dei vettori. Se $\langle , \rangle$ è il prodotto scalare su $\mathcal{V}^2$ o $\mathcal{V}^3$ dell'**Esempio 7.3.2** allora la definizione di angolo appena data coincide con la nozione usuale di angolo - vedi (7.3.2).

Esempio 7.7.5. Nello spazio vettoriale euclideo $C^0([-\pi,\pi])$ - vedi l'**Esempio 7.7.1** - siano f_m e g_n le funzioni definite da

$$f_m(t) := cosmt, \quad g_n(t) := \sin nt, \qquad m,n \in \mathbb{N}_+. \tag{7.7.7}$$

Un calcolo[2] dà che

$$\pi\delta_{ab}=\int_{-\pi}^{\pi} f_a(t)f_b(t)dt=\langle [f_a],[f_b]\rangle,\ \pi\delta_{ab}=\int_{-\pi}^{\pi} g_a(t)g_b(t)dt=\langle [g_a],[g_b]\rangle,\ 0=\int_{-\pi}^{\pi} f_a(t)g_b(t)dt=\langle [f_a],[g_b]\rangle, \qquad a,b\in\mathbb{N}_+. \tag{7.7.8}$$

Quindi le funzioni definite da (7.7.7) sono a due a due ortogonali.

Terminologia 7.7.6. Siano V, W spazi vettoriali reali euclidei con prodotti scalari $\langle,\rangle_V$ e $\langle,\rangle_W$ rispettivamente. Una *isometria* $f\colon V \to W$ è un isomorfismo $(V,\langle,\rangle_V) \xrightarrow{\sim} (W,\langle,\rangle_W)$ di spazi quadratici, cioè un isomorfismo di spazi vettoriali tale che per ogni $v_1, v_2 \in V$ si ha

$$\langle v_1, v_2\rangle_V = \langle f(v_1), f(v_2)\rangle_W. \tag{7.7.9}$$

Definizione 7.7.7. 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 $1 \le i,j \le n$ si ha che $\langle v_i, v_j\rangle = \delta_{ij}$.

Lemma 7.7.8. *Sia $(V,\langle,\rangle)$ uno spazio vettoriale euclideo e $\{v_1,\ldots,v_n\}$ una lista di vettori ON. Allora $\{v_1,\ldots,v_n\}$ sono linearmente indipendenti.*

Dimostrazione. Supponiamo che

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0. \tag{7.7.10}$$

Sia $1 \le i \le n$. Calcolando il prodotto scalare di v_i con ambo i membri di (7.7.10) troviamo che $\lambda_i = 0$. □

Esempio 7.7.9. Sia $C^0([-\pi,\pi])$ lo spazio euclideo dell'**Esempio 7.7.1**. Le funzioni

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{1}{\sqrt{\pi}}\cos at, \quad \frac{1}{\sqrt{\pi}}\sin at, \qquad 1 \le a \le n \tag{7.7.11}$$

formano una lista ON: questo segue dai calcoli dell'**Esempio 7.7.9** e dalla formula

$$0 = \int_{-\pi}^{\pi} \cos atdt = \int_{-\pi}^{\pi} \sin atdt, \qquad a \in \mathbb{N}_+. \tag{7.7.12}$$

Quindi le funzioni di (7.7.11) sono linearmente indipendenti.

Proposizione 7.7.10. *Sia $(V,\langle,\rangle)$ uno spazio vettoriale euclideo finitamente generato. Esiste una base ortonormale di V.*

Dimostrazione. Il **Teorema 7.5.4** dà che esiste una base $\mathcal{B}$ di V tale che $M_{\mathcal{B}}(\langle,\rangle)$ è diagonale con ciascuna entrata in $\{0,1\}$. Siccome $\langle,\rangle$ è definita positiva è non-degenere e quindi tutte le entrate sulla diagonale principale di $M_{\mathcal{B}}(\langle,\rangle)$ sono uguali a 1. Quindi $M_{\mathcal{B}}(\langle,\rangle) = 1_n$ e perciò $\mathcal{B}$ è una base ON. □

[2]Le uguaglianze $\cos mt = (e^{imt} + e^{-imt})/2$ e $\sin mt = (e^{imt} - e^{-imt})/2i$ rendono semplice il calcolo degli integrali.

Una tipica costruzione che si fa in uno spazio spazio vettoriale euclideo è la proiezione (ortogonale) su un sottospazio: generalizza la proiezione ortogonale su una retta del piano o dello spazio o su un piano dello spazio.

Proposizione 7.7.11. *Sia $(V, \langle , \rangle)$ uno spazio vettoriale euclideo e $U \subset V$ un sottospazio finitamente generato. Esiste un'applicazione lineare $\pi\colon V \to U$ tale che*

$$(v - \pi(v)) \in U^{\perp} \quad \forall v \in V \tag{7.7.13}$$

ed è unica.

Dimostrazione. Per la **Proposizione 7.7.10** esiste una base ON di U, sia $\mathcal{B} = \{u_1, \dots, u_m\}$. Poniamo

$$\pi(v) := \sum_{i=1}^{m} (v, u_i) u_i. \tag{7.7.14}$$

Allora $\pi\colon V \to U$ è lineare e vale (7.7.13) perché $\mathcal{B}$ è ON. Sia $v \in V$ e supponiamo che $w \in U$ è tale che $(v - w) \in U^{\perp}$: allora

$$\pi(v) - w = ((v - w) - (v - \pi(v))) \in U^{\perp}. \tag{7.7.15}$$

Siccome $(\pi(v) - w) \in U$ e la restrizione di $\langle , \rangle$ a U è non-degenere (lo è su tutto V per ipotesi) segue che $(\pi(v) - w) = 0$. Questo dimostra l'unicità di π. □

Siano V, U e π come nella **Proposizione 7.7.11**. Quindi abbiamo l'uguaglianza

$$v = \pi(v) + (v - \pi(v)), \quad \pi(v) \in U, \quad (v - \pi(v)) \in U^{\perp}. \tag{7.7.16}$$

In altre parole abbiamo decomposto v nella somma di un vettore in U e di un vettore perpendicolare a U. La seguente proposizione dà una caratterizzazione della proiezione ortogonale $\pi(v)$ in termini di distanza.

Proposizione 7.7.12. *Siano V, U e π come nella* **Proposizione 7.7.11**. *Sia $u \in U$: allora*

$$||v - \pi(v)|| \leq ||v - u|| \quad \forall u \in U \tag{7.7.17}$$

e si ha eguaglianza solo se $u = \pi(v)$.

Dimostrazione. Abbiamo che

$$||v-u||^2 = ||v-\pi(v)-(u-\pi(v))||^2 = ||v-\pi(v)||^2 - 2\langle (v-\pi(v)), (u-\pi(v))\rangle + ||(u-\pi(v))||^2 = \\ = ||v-\pi(v)||^2 + ||(u-\pi(v))||^2 \tag{7.7.18}$$

(la seconda uguaglianza vale perché $(u - \pi(v)) \in U^{\perp}$) e la proposizione segue. □

Esempio 7.7.13. Sia $\mathcal{R}(\mathbb{T})$ lo spazio euclideo dell'**Esempio 7.7.1**. Sia $U_n \subset \mathcal{R}(\mathbb{T})$ il sottospazio generato dalle (classi di equivalenza delle) funzioni di (7.7.11). Sia $\phi \in \mathcal{R}(\mathbb{T})$ la (classi di equivalenza della) funzione ϕ tale che

$$\phi(t) := \begin{cases} t & \text{se } -\pi < t < \pi, \\ 0 & \text{se } t \in \pi\mathbb{Z}. \end{cases} \tag{7.7.19}$$

La proiezione di $[\phi]$ sul sottospazio U_n, è la classe di equivalenza della funzione ϕ_n data da

$$\phi_n(t) = 2 \sum_{a=1}^{n} \frac{(-1)^{a+1}}{a} \sin at. \tag{7.7.20}$$

7.8 Il teorema spettrale

Supponiamo che $(V,(,))$ sia uno spazio vettoriale euclideo e che $f \in Q(V)$ sia una forma quadratica. Sappiamo che esiste una base $\mathcal{B}$ in cui f è diagonale. Esiste una tale $\mathcal{B}$ ortonormale? La risposta è affermativa e va sotto il nome di Teorema spettrale (spiegheremo più in là l'origine del nome). Dimostreremo il Teorema spettrale, e discuteremo il problema di dare un algoritmo per trovare una base ortonormale che diagonalizza f.

7.8.1 Dimostrazione del Teorema spettrale

Teorema 7.8.1 (Teorema spettrale, 1° enunciato)**.** *Sia $(V,(,))$ uno spazio vettoriale euclideo di dimensione n e $f\colon V \to \mathbb{R}$ una forma quadratica. Esiste una base ortonormale $\mathcal{B} = \{v_1,\dots,v_n\}$ di V che diagonalizza f, cioè tale che*

$$f(x_1v_1 + \ldots + x_nv_n) = \sum_{i=1}^{n} \lambda_i x_i^2. \tag{7.8.1}$$

Dimostrazione. Per motivare la dimostrazione, supponiamo che valga (7.8.1). Sia F la forma bilineare simmetrica associata a f. Allora

$$F(\sum_{i=1} x_iv_i, \sum_{j=1} x_jv_j) = \sum_{i=1}^{n} \lambda_i x_i y_i.$$

Ne segue che, *indicando con $v_i^\perp$ l'ortogonale di v_i per il prodotto euclideo*,

$$v_i^\perp = \langle v_1,\dots,v_{i-1},v_{i+1},\dots,v_n\rangle \subset \{w \in V \mid F(w,v_i) = 0\}. \tag{7.8.2}$$

Più precisamente, si ha eguaglianza se $\lambda_i \neq 0$, e inclusione stretta se $\lambda_i = 0$. Viceversa, *supponiamo* che esista $v_n \in V$ tale che valga (7.8.2) (con $i = n$), e poniamo $\lambda_n := f(v_n)$. Siano $\{u_1,\dots,u_{n-1}\}$ una base di $v_n^\perp$, (ortonormale o no, non è rilevante), e $\mathcal{C} = \{u_1,\dots,u_{n-1},v_n\}$ la base di V ottenuta aggiungendo v_n. Allora la matrice $M_{\mathcal{C}}(f)$ è del tipo

$$M_{\mathcal{C}}(f) = \begin{bmatrix} * & \dots & * & 0 \\ * & \dots & * & \vdots \\ * & \dots & * & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}.$$

A questo punto si può iterare il procedimento, cioè cercare $v_{n-1} \in v_n^\perp$ tale che $v_{n-1}^\perp \subset \{w \in V \mid F(w,v_{n-1}) = 0\}$, e così via. Il risultato (ammettendo che esistano v_n, v_{n-1} etc.) sarà una base ortogonale che diagonalizza f. Per ottenere una base ortonormale basterà normalizzare la base ortogonale (dividere ogni vettore della base ortogonale per la sua norma). Il ragionamento appena fatto dimostra che è sufficiente far vedere che esiste $0 \neq v_n \in V$ tale che valga

$$v_n^\perp \subset \{w \in V \mid F(w,v_n) = 0\}. \tag{7.8.3}$$

Infatti, se vale (7.8.3) per qualsiasi spazio vettoriale euclideo $(V,(,))$ di dimensione n, e forma quadratica $f \in Q(V)$, allora vale anche per lo spazio vettoriale euclideo $v_n^\perp$ con prodotto euclideo la restrizione di $(,)$, e forma quadratica la restrizione di f. Quindi esiste $v_{n-1} \in v_n^\perp$ tale che valga (7.8.3) con v_{n-1} al posto di v_n e $F_{v_n^\perp \times v_n^\perp}$ al posto di F; iterando $n-1$ volte si arriva a una base ortogonale che diagonalizza f.

Ora dimostriamo che esiste $0 \neq v_n \in V$ tale che valga (7.8.3). Consideriamo l'applicazione

$$\begin{array}{ccc} V \setminus \{\mathbf{0}\} & \xrightarrow{\rho} & \mathbb{R} \\ v & \mapsto & f(v)/||v||^2 \end{array}$$

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

$$F_n := \{X \in \mathbb{R}^n \mid |x_i| \leq 1 \text{ per ogni } i \text{ e } |x_{i_0}| = 1 \text{ per un } i_0 \text{ (almeno)}\}.$$

Dimostriamo che la restrizione di ρ a F_n ha un massimo. Il caso $n = 1$ è banale perché F_1 è un insieme con due elementi. Ora consideriamo il caso $n = 2$. La F_2 è unione di 4 segmenti chiusi e limitati, e siccome la funzione $\rho(X) = f(X)/||X||^2$ è continua su $\mathbb{R}^2 \setminus \{\mathbf{0}\}$ segue per il Teorema di Bolzano-Weierstrass che la restrizione di ρ a F_2 ammette massimo. 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 $\mathbb{R}$ [3] allora esiste un massimo di f. 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 $X \in (\mathbb{R}^n \setminus \{0\})$, e sia m il massimo tra $|x_1|, |x_2|, \ldots, |x_n|$. Notate che $m > 0$ perché $X \neq 0$, e che $m^{-1}X \in F_n$ perché 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

$$\rho(X)=f(X)/||X||^2=m^2 f(m^{-1}X)/m^2||m^{-1}X||^2=f(m^{-1}X)/||m^{-1}X||^2=\rho(m^{-1}X)\leq\rho(\overline{X}).$$

Questo dimostra che ρ ammettte massimo. Ora torniamo a considerare V (non è necessario identificarlo con $\mathbb{R}^n$ in quello che segue). Sia $v_n \in V \setminus \{\mathbf{0}\}$ tale che $f(v_n)$ sia il massimo della funzione ρ. Sia $v \in v_n^{\perp}$ dove l'ortogonalità è rispetto al prodotto euclideo. La funzione

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{\xi} & \mathbb{R} \\ s & \mapsto & \rho(v_n + sv) \end{array} \tag{7.8.4}$$

è 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

$$\xi'(0) = 0. \tag{7.8.5}$$

Scriviamo $\xi = g/h$, dove $g, h \colon \mathbb{R} \to \mathbb{R}$ sono le funzioni definite da

$$g(s) := f(v_n + sv) = F(v_n + sv, v_n + sv), \qquad h(s) = ||v_n + sv||^2 = (v_n + sv, v_n + sv).$$

Allora

$$\xi'(0) = \frac{g'(0)h(0) - g(0)h'(0)}{h(0)^2}.$$

Ora $h'(0) = 0$ perche $v_n \perp v$, e quindi

$$\xi'(0) = \frac{g'(0)}{h(0)} = \frac{g'(0)}{||v_n||^2} = \frac{2F(v_n, v)}{||v_n||^2}.$$

Dalla (7.8.5) segue che $F(v_n, v) = 0$. Siccome v è un arbitrario vettore ortogonale a v_n, questo dimostra che vale (7.8.3). □

[3] Una funzione da un sottoinsieme $S \subset \mathbb{R}^n$ a $\mathbb{R}$ è continua dati $\overline{X} \in S$ e $\epsilon > 0$, esiste $\delta > 0$ tale che $||f(X) - f(\overline{X})|| < \epsilon$ per ogni $X \in S$ tale che $||X - \overline{X}|| < \delta$.

7.8.2 Polinomio caratteristico di una matrice quadrata

Dato uno spazio vettoriale euclideo $(V,(,))$, e una forma quadratica $f \in Q(V)$, come possiamo procedere per trovare una base ON che diagonalizza f? Dobbiamo trovare un vettore $0 \neq v_n \in V$ tale che valga (7.8.3). Per trovare v_n, scegliamo una qualsiasi base *ortonormale* $\mathcal{B}$ di V. Sia $A = M_{\mathcal{B}}(f)$ la matrice simmetrica associata a f nella base $\mathcal{B}$. Sia $v = \sum_{i=1}^n x_i v_i$ un vettore di V. Allora

$$\begin{aligned} v^\perp &= \left\{ w = \sum_{i=1}^n y_i v_i \mid Y^t \cdot X = 0 \right\}, \\ \{w \in V \mid F(w,v) = 0\} &= \left\{ w = \sum_{i=1}^n y_i v_i \mid Y^t \cdot A \cdot X = 0 \right\} \end{aligned}$$

Segue che

$$v^\perp \subset \{w \in V \mid F(w,v) = 0\} \tag{7.8.6}$$

se e solo se esiste $\lambda \in \mathbb{R}$ tale che

$$A \cdot X = \lambda X. \tag{7.8.7}$$

Infatti vale il seguente risultato.

Proposizione 7.8.2. *Sia V uno spazio vettoriale sul campo k, e siano $f, g \in V^*$. Allora $\ker f \subset \ker g$ se e solo se esiste $\lambda \in k$ tale che $g = \lambda f$.*

Dimostrazione. Se esiste $\lambda \in k$ tale che $g = \lambda f$, chiaramente $\ker f \subset \ker g$. Ora supponiamo che $\ker f \subset \ker g$. L'applicazione

$$\begin{array}{ccc} V & \xrightarrow{\varphi} & k^2 \\ v & \mapsto & (f(v), g(v)) \end{array} \tag{7.8.8}$$

è lineare, e non è suriettiva perché $(0,1)$ non è nella sua immagine. Quindi l'immagine di φ è un sottospazio proprio di k^2, e perciò esistono $\alpha, \beta \in k$, non entrambi nulli, tali che $\alpha f(v) + \beta g(v) = 0$ per ogni $v \in V$. Se $\beta \neq 0$, allora $g = (-\alpha/\beta) f$. Se $\beta = 0$ allora $\alpha \neq 0$, e quindi $f = 0$; ma allora $\ker f = V$, e quindi $\ker g = V$, cioè $g = f$. □

L'affermazione che $v^\perp \subset \{w \in V \mid F(w,v) = 0\}$ equivale all'esistenza di $\lambda \in \mathbb{R}$ tale che valga (7.8.7) segue dalla **Proposizione 7.8.2** per $V = \mathbb{R}^n$ e $f, g \colon \mathbb{R}^n \to \mathbb{R}$ le funzioni lineari definite da $f(Y) := Y^t \cdot X$, $g(Y) := Y^t \cdot A \cdot X$. Ora supponiamo che esista $0 \neq X \in \mathbb{R}^n$ tale che valga (7.8.7). Allora la matrice $(\lambda 1_n - A)$ è singolare perché $0 \neq X$ è nel suo nucleo, e quindi

$$\mathrm{Det}(\lambda 1_n - A) = 0. \tag{7.8.9}$$

Viceversa, se vale (7.8.9), allora esiste X tale che valga (7.8.7). Questo ci porta a studiare la funzione di λ definita dal membro di sinistra di (7.8.9). Tale funzione ha senso per una qualsiasi matrice quadrata.

Proposizione 7.8.3. *Siano k un campo e $A \in M_{n,n}(k)$. Allora $\mathrm{Det}(\lambda 1_n - A)$ è un polinomio in λ a coefficienti in k, monico di grado n.*

Dimostrazione.

$$\mathrm{Det}(\lambda 1_n - A) = \det \begin{bmatrix} \lambda - a_{11} & -a_{12} & \dots & \dots & \dots & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \ddots & \ddots & \dots & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \dots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \dots & \vdots & \ddots & \lambda - a_{n-1,n-1} & -a_{n-1,n} \\ -a_{n1} & \dots & \dots & \dots & \dots & -a_{n,n-1} & \lambda - a_{nn} \end{bmatrix} \tag{7.8.10}$$

Espandiamo il determinante secondo la **Proposizione 6.6.6**; ogni addendo a destra di (6.6.5) è un prodotto di polinomi in λ a coefficienti in k, di grado al più 1, e questro dimostra che $\mathrm{Det}(\lambda 1_n - A)$ è un polinomio in λ a coefficienti in 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 vediamo che il coefficiente di λ^n è 1, cioè il polinomio è monico. □

Definizione 7.8.4. Sia $A \in M_{n,n}(k)$. Il *polinomio caratteristico* di A è

$$P_A = \det(\lambda 1_n - A). \tag{7.8.11}$$

Esempio 7.8.5. Siano

$$A := \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}, \qquad B := \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}.$$

Un facile calcolo dà che

$$P_A(\lambda) = \lambda^2 - 4\lambda + 5, \qquad P_B(\lambda) = \lambda^2 - 4\lambda - 1.$$

Esempio 7.8.6. Sia $A \in M_{n,n}(k)$ un matrice diagonale. Allora

$$P_A(\lambda) = (\lambda - a_{11}) \cdot \ldots \cdot (\lambda - a_{nn}),$$

e quindi le radici di P_A sono le entrate della diagonale principale di A.

Proposizione 7.8.7. *Se $A, B \in M_{n,n}(k)$ sono matrice coniugate, allora $P_A = P_B$, cioè i loro polinomi caratteristici sono uguali.*

Dimostrazione. Esiste $G \in \mathrm{GL}_n(k)$ tale che $A = G^{-1} \cdot B \cdot G$, e quindi

$$P_A(\lambda) = \mathrm{Det}(\lambda 1_n - G^{-1} \cdot B \cdot G) = \mathrm{Det}(G^{-1} \cdot (\lambda 1_n - B) \cdot G) = \mathrm{Det}(G)^{-1} \cdot \mathrm{Det}(\lambda 1_n - B) \cdot \mathrm{Det}(G) = P_B(\lambda).$$

□

Ora torniamo al problema di trovare una base ON di uno spazio vettoriale euclideo V che diagonalizza una $f \in Q(V)$. Come abbiamo visto, dobbiamo trovare una radice di (7.8.9), cioè una radice del polinomio caratteristico P_A. Il Teorema fondamentale dell'algebra assicura che esiste una soluzione *complessa* di (7.8.9). Il contenuto fondamentale del Teorema spettrale è che tutte le radici di (7.8.9) sono reali. Notiamo che se $A \in M_{n,n}(\mathbb{R})$ è una matrice qualsiasi, può ben succedere che le radici di P_A *non* siano tutte reali, riandate all'**Esempio 7.8.5**.

Proposizione 7.8.8. *Se $A \in M_{n,n}(\mathbb{R})$ è una matrice simmetrica, allora tutte le radici del suo polinomio caratteristico sono reali. Più precisamente esistono $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ tali che*

$$P_A(\lambda) = (\lambda - \alpha_1) \cdot \ldots \cdot (\lambda - \alpha_n).$$

Dimostrazione. Sia $f \in Q(\mathbb{R}^n)$ definita da $f(X) = X^t \cdot A \cdot X$. Sia $(,)$ il prodotto euclideo standard su $\mathbb{R}^n$. Per il Teorema spettrale, esiste una base ON (per il prodotto euclideo standard) $\mathcal{B}$ che diagonalizza f. Quindi $M_{\mathcal{B}}(f) = \Delta$, dove Δ è una matrice diagonale. D'altra parte $M_{\mathcal{B}}(f) = G^t \cdot A \cdot G$, dove $G = M_{\mathcal{S}}^{\mathcal{B}}(\mathrm{Id}_{\mathbb{R}^n})$ è la matrice di cambiamento di coordinate dalla base $\mathcal{B}$ alla base standard $\mathcal{S}$. Siccome $\mathcal{B}$ è una base ON, la $G \in \mathrm{O}_n(\mathbb{R})$, cioè $G^t \cdot G = 1_n$, ovvero $G^t = G^{-1}$. Quindi A è coniugata (oltre che congruente) a Δ, e per la **Proposizione 7.8.7**, segue che $P_A = P_\Delta$. Ma $P_\Delta = (\lambda - \alpha_1) \cdot \ldots \cdot (\lambda - \alpha_1)$, dove $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ sono le entrate della diagonale principale di Δ. □

7.8.3 Versione "algoritmica" del Teorema spettrale

Proposizione 7.8.9. *Siano $(V, (,))$ uno spazio vettoriale euclideo e $f \in Q(V)$. Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ON che diagonalizza f. Sia $v \in V$. Allora vale* (7.8.6) *se e solo se v è combinazione lineare di $v_{i_1}, \ldots, v_{i_m}$ con la proprietà che*

$$f(v_{i_1}) = \ldots = f(v_{i_m}).$$

Dimostrazione. Facile esercizio. □

Siano $(V, (,))$ uno spazio vettoriale euclideo e $f \in Q(V)$. Dalla **Proposizione 7.8.9**, e dai risultati della **Sottosezione 7.8.2**, segue che il seguente algoritmo produce una base ON che diagonalizza f.

1. Si sceglie una qualsiasi base ON $\mathcal{C}$ di V, e si calcola $A = M_{\mathcal{C}}(f)$.

2. Si calcolano le radici del polinomio caratteristico P_A (per il Teorema spettrale sono tutte reali).

3. Per ogni radice λ_i del polinomio caratteristico P_A, si determina una base ON del sottospazio di $\mathbb{R}^n$ delle soluzioni X di $(\lambda_i 1_n - A) \cdot X = 0$. In questo modo si producono n vettori $X(1), \ldots, X(n)$ di $\mathbb{R}^n$.

4. Siano $v_1, \ldots, v_n \in V$ i vettori le cui coordinate sono date dai vettori $X(1), \ldots, X(n)$ rispettivamente. Allora $\{v_1, \ldots, v_n\}$ è una base ON di V che diagonalizza f.

Esempio 7.8.10. Sia $A \in M_{3,3}(\mathbb{R})$ la matrice simmetrica definita da

$$A := \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix},$$

e sia $f \in Q(\mathbb{R}^3)$ la forma quadratica associata: $f(X) = X^t \cdot A \cdot X$. Per trovare una base ON che diagonalizza f, calcoliamo

$$\mathrm{Det}(\lambda 1_3 - A) = \lambda^3 - 6\lambda^2 - 3\lambda + 18. \tag{7.8.12}$$

Le radici di (7.8.12) sono 6, $\sqrt{3}$ e $-\sqrt{3}$. Con un clacolo semplice (ma noioso) troviamo che

$$\begin{aligned}\{X \in \mathbb{R}^3 \mid (6 \cdot 1_3 - A) \cdot X = 0\} &= \langle (1,1,1) \rangle,\\ \{X \in \mathbb{R}^3 \mid (\sqrt{3} \cdot 1_3 - A) \cdot X = 0\} &= \langle (7-3\sqrt{3}, -5-\sqrt{3}, -2+4\sqrt{3}) \rangle,\\ \{X \in \mathbb{R}^3 \mid (-\sqrt{3} \cdot 1_3 - A) \cdot X = 0\} &= \langle (7+3\sqrt{3}, -5+\sqrt{3}, -2-4\sqrt{3}) \rangle.\end{aligned}$$

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 $\mathcal{C}$.

Esercizi del Capitolo 7

Esercizio 7.1. *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 7.2. *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 7.3. *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 7.4. *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 7.5. *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 7.6. *Siano (V, f) e (W, g) spazi vettoriali quadratici.*

(a) *Dimostrate che se (V, f) è isomorfo a (W, g) allora, scelte una base $\mathcal{B}$ di V e una base $\mathcal{C}$ di W, esiste $m \in k^*$ $(= (k \setminus \{0\}))$ tale che*

$$\operatorname{Det} M_{\mathcal{B}}(f) = m^2 \cdot \operatorname{Det} M_{\mathcal{C}}(g). \tag{7.8.13}$$

(b) *Siano $A, B \in M_{2,2}(\mathbb{Q})$ date da*

$$A := \begin{bmatrix} 3 & 2 \\ 2 & 5 \end{bmatrix}, \qquad B := \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \tag{7.8.14}$$

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 $(\mathbb{Q}^2, f)$ non *è isomorfo a $(\mathbb{Q}^2, g)$. (Suggerimento: invocate il punto (a).)*

(c) *Siano $A, B \in M_{2,2}(\mathbb{R})$ date da (7.8.14) 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 $(\mathbb{R}^2, f)$* è *isomorfo a $(\mathbb{R}^2, \psi)$.*

Esercizio 7.7. *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. Supponiamo che f sia non-degenere e che $U \subset V$ sia isotropo per f. Dimostrate che* $\dim U \leq \dim V/2$.

Esercizio 7.8. *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 7.9. *Se $A \in M_{n,n}(k)$ 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 7.10. *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 7.11. *Sia*

$$U := \{X \in \mathbb{R}^3 \mid x_1 + 2x_2 + 3x_3 = 0\}. \tag{7.8.15}$$

Sia $v := (1, 1, 1)$. Determinate la proiezione ortogonale di v su U, se $\mathbb{R}^3$ ha il prodotto scalare standard.

Esercizio 7.12. *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 è una isometria di V.

(2) Interpretate geometricamente R_v nel caso in cui $V = \mathcal{V}^2$ o $V = \mathcal{V}^3$.

Esercizio 7.13. *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 7.14. *Sia $A \in M_{2,2}(\mathbb{R})$ data da*

$$A := \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$$

1. *Si verifichi 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 7.15. *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 $1 \leq i \leq n$, 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 $1 \leq i \leq n$,

$$\langle v_1, \ldots, v_i\rangle = U_i.$$

Esercizio 7.16. *Sia V uno spazio vettoriale reale finitamente generato, e f una forma quadratica su V. Siano $\mathcal{B}$ e $\mathcal{C}$ basi di V. Si dimostri che $\operatorname{Det} M_{\mathcal{B}}(f)$ e $\operatorname{Det} M_{\mathcal{C}}(f)$ sono entrambi nulli, o hanno lo stesso segno.*

Esercizio 7.17. *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

$$1,\ \operatorname{Det} A(1),\ \operatorname{Det} A(2), \ldots, \operatorname{Det} A(n), \tag{7.8.16}$$

ovvero che, denotando con c il numero di cambi di segno nella sequenza (7.8.16)*, la segnatura di f è uguale a $n - 2c$. (Suggerimento: usate i risultati dell'***Esercizio 7.15** *e dell'***Esercizio 7.16***.)*

Esercizio 7.18. *Sia $A \in M_{n,n}(k)$. Dimostrate che*

$$P_A(\lambda) = \lambda^n - (\operatorname{Tr} A)\lambda^{n-1} + \ldots + \det A, \tag{7.8.17}$$

*dove la traccia $\operatorname{Tr} A = \sum_{i=1}^n a_{ii}$ è come nell'***Esercizio 7.9***. In particolare, segue che $\operatorname{Tr}(G^1 \cdot A \cdot G) = \operatorname{Tr} A$; deducte questo risultato dalll'***Esercizio 7.9***.*

Capitolo 8

Coniche e quadriche

8.1 Coniche e quadriche affini

Sia $\mathbb{A}$ uno spazio affine **reale** di dimensione 2 - per esempio $\mathcal{A}^2$. Una *conica* in $\mathbb{A}$ è l'insieme dei punti le cui coordinate rispetto a un sistema di coordinate affini sono le soluzioni reali di un polinomio reale $f(x_1, x_2)$ di grado 2

$$C = \{p \mid f(x_1(p), x_2(p)) = 0\}. \tag{8.1.1}$$

La $f(x_1, x_2) = 0$ si dice *equazione cartesiana* di C. La definizione ha senso perché se (y_1, y_2) è un nuovo sistema di coordinate allora esistono una matrice invertibile $A \in M_{2,2}(\mathbb{R})$ e un vettore colonna $B \in M_{2,1}(\mathbb{R})$ tali che la relazione tra vecchie e nuove coordinate è

$$X = A \cdot Y + B. \tag{8.1.2}$$

Sostituendo l'espressione delle (x_1, x_2) data sopra nella $f(x_1, x_2)$ abbiamo che

$$C = \{p \mid f(a_{11}y_1(p) + a_{12}y_2(p) + b_1, a_{21}y_1(p) + a_{22}y_2(p) + b_2) = 0\}. \tag{8.1.3}$$

Siccome $f(a_{11}y_1 + a_{12}y_2 + b_1, a_{21}y_1 + a_{22}y_2 + b_2)$ è un polinomio di grado 2 nelle (y_1, y_2) vediamo che la definizione di conica è ben posta. I risultati ottenuti sulle forme quadratiche daranno *forme canoniche affini* per le coniche.

Proposizione 8.1.1. *Sia C una conica nel piano. Esiste un sistema di riferimento affine $RA(O, x_1, x_2)$ tale che C abbia per equazione cartesiana una delle equazioni in forma canonica della Tabella* (8.1)*. L'equazione cartesiana canonica di C è (appunto) unica.*

Dimostrazione. Siano (x_1, x_2) coordinate cartesiane nel piano. Supponiamo che C sia data da (8.1.1) e scriviamo

$$f(x_1, x_2) = q(x_1, x_2) + \mu x_1 + \nu x_2 + \theta \tag{8.1.4}$$

dove q è una forma quadratica non nulla. Quindi esiste una matrice simmetrica $M \in M_{2,2}(\mathbb{R})$ non nulla tale che

$$q(x_1, x_2) = X^t \cdot M \cdot X. \tag{8.1.5}$$

Ora siano (y_1, y_2) coordinate rispetto a un nuovo sistema di coordinate cartesiane; esistono una matrice invertibile $A \in M_{2,2}(\mathbb{R})$ e un vettore colonna $B \in M_{2,1}(\mathbb{R})$ tali che la relazione tra vecchie e nuove coordinate di uno stesso punto è data da (8.1.2). Sia

$$g(y_1, y_2) = f(a_{11}y_1 + a_{12}y_2 + b_1, a_{21}y_1 + a_{22}y_2 + b_2).$$

Nel nuovo sistema di riferimento la conica C ha equazione cartesiana $g(y_1, y_2) = 0$. Scriviamo

$$g(y_1, y_2) = q'(y_1, y_2) + \mu' y_1 + \nu' y_2 + \theta' \tag{8.1.6}$$

Tabella 8.1: Equazione canonica delle coniche in $\mathbb{A}^2_{\mathbb{R}}$

Equazione canonica	Nome	
$x_1^2 + x_2^2 - 1 = 0$	ellisse	coniche non-degeneri
$x_1^2 + x_2^2 + 1 = 0$	ellisse complessa	
$x_1^2 - x_2^2 - 1 = 0$	iperbole	
$x_1^2 - x_2 = 0$	parabola	
$x_1^2 + x_2^2 = 0$	coppia di rette complesse coniugate	coniche degeneneri
$x_1^2 - x_2^2 = 0$	coppia di rette incidenti	
$x_1^2 - 1 = 0$	coppia di rette parallele	
$x_1^2 = 0$	retta doppia	

dove q' è una forma quadratica non nulla. Un facile conto dà che

$$g(y_1, y_2) = Y^t \cdot A^t \cdot M \cdot A \cdot Y. \tag{8.1.7}$$

Per il **Corollario 7.5.6** esiste A invertibile tale che $A^t \cdot M \cdot A$ sia diagonale con entrate in $\{0, \pm 1\}$. Quindi in un opportuno sistema di coordinate cartesiane (y_1, y_2) abbiamo che C ha equazione

$$\pm y_1^2 \pm y_2^2 + \mu' y_1 + \nu' y_2 + \theta' \tag{8.1.8}$$

oppurre

$$\pm y_1^2 + \mu' y_1 + \nu' y_2 + \theta'. \tag{8.1.9}$$

Se $\alpha \in \mathbb{R}$ è non nullo gli zeri di $f(y_1, y_2)$ sono gli stessi zeri di $\alpha f(y_1, y_2)$, quindi possiamo assumere che nelle equzioni (8.1.8) e (8.1.9) il coefficient di y_1^2 sia 1. Supponiamo che C abbia equazione $y_1^2 + \epsilon y_2^2 + \mu' y_1 + \nu' y_2 + \theta'$ dove $\epsilon = \pm 1$. Si ha

$$y_1^2 + \epsilon y_2^2 + \mu' y_1 + \nu' y_2 + \theta' = (y_1 + \mu'/2)^2 + \epsilon(y_2 + \epsilon\nu'/2)^2 + \theta' - (\mu'/2)^2 - \epsilon(\nu'/2)^2.$$

Passando al sistema di coordinate (z_1, z_2) tali che $z_1 = (y_1 + \mu'/2)$ e $z_2 = (y_2 + \epsilon\nu'/2)$ l'equazione cartesiana di C diventa $z_1^2 + \epsilon z_2^2 + d = 0$ dove $d := (\theta' - (\mu'/2)^2 - \epsilon(\nu'/2)^2)$. Se $d = 0$ abbiamo una forma canonica e C è una coppia di rette complesse coniugate oppure una coppia di rette (reali) incidenti. Se $d < 0$ una equazione cartesiana di C è

$$\left(\frac{z_1}{\sqrt{|d|}}\right)^2 + \epsilon \left(\frac{z_2}{\sqrt{|d|}}\right)^2 - 1 = 0.$$

Passando a coordinate cartesiane (w_1, w_2) date da $w_i := z_i/\sqrt{|d|}$ otteniamo una equazione in forma canonica e vediamo che C o è una ellisse o una iperbole. Se $d > 0$ otteniamo in modo simile un'equazione canonica di una iperbole oppure di una conica liscia complessa. Se C ha equazione $\pm y_1^2 + \mu' y_1 + \nu' y_2 + \theta'$ si procede in modo simile. Per dimostrare che l'equazione canonica è unica si dimostra per prima cosa che se $C_1, C_2 \subset \mathbb{A}$ sono coniche non vuote,

$$C_1 = \{p \mid f_1(x_1(p), x_2(p)) = 0\}, \qquad C_2 = \{p \mid f_2(x_1(p), x_2(p)) = 0\},$$

allora esiste $0 \neq \alpha \in \mathbb{R}$ tale che $f_1 = \alpha f_2$. Da questo l'unicità della forma canonica segue subito. □

Tabella 8.2: Equazione canonica delle quadriche non-degeneri in $\mathbb{A}^3_{\mathbb{R}}$

Equazione canonica	Nome	Tipo
$x_1^2 + x_2^2 + x_3^2 - 1 = 0$	ellissoide	
$x_1^2 - x_2^2 + x_3^2 + 1 = 0$	ellissoide complesso	
$x_1^2 + x_2^2 - x_3^2 - 1 = 0$	iperboloide iperbolico (o a una falda)	
$x_1^2 + x_2^2 - x_3^2 + 1 = 0$	iperboloide ellittico (o a due falde)	
$x_1^2 + x_2^2 - x_3 = 0$	paraboloide ellittico	
$x_1^2 - x_2^2 - x_3 = 0$	paraboloide iperbolico	

Diamo una interpretazione alternativa della **Proposizione 8.1.1**. Sia $f\colon \mathbb{A} \to \mathbb{A}$ un automorfismo (come spaio affine), si chiama anche *affinità*. Se X è il vettore colonna di coordinate affini su $\mathbb{A}$ esistono $A \in M_{2,2}(\mathbb{R})$ invertibile e $B \in M_{2,1}(\mathbb{R})$ vettore colonna tali che $X(f(p)) = A \cdot X(p) + B$ per ogni $p \in \mathbb{A}$. Le affinità includono i movimenti rigidi ma anche le dilatazioni e altre trasformazioni che modificano gli angoli. Un esempio tipico di trasformazione affine f tra piani diversi Π_1 e Π_2 nello spazio è la proiezione "da un punto all'infinito"cioè $f(p)$ è l'intersezione di Π_2 con l'unica retta R passante per p e parallela a una retta fissata R_0 (non parallela a Π_1 né a Π_2). La **Proposizione 8.1.1** afferma che modulo le affinità tutte le ellissi sono equivalenti, e così le iperboli, le parabole, etc. In altre parole se un'ellisse C_2 è ottenuta da una curva C_1 applicando una affinità (per esempio una proiezione dall'infinito) sappiamo che C_1 è anch'essa un'ellisse, ma nulla di più.

Ora sia $\mathbb{A}$ uno spazio affine **reale** di dimensione 3. Una *quadrica* in $\mathbb{A}$ è l'insieme dei punti le cui coordinate rispetto a un sistema di coordinate affini sono le soluzioni reali di un polinomio reale $f(x_1, x_2, x_3)$ di grado 2

$$Q = \{p \mid f(x_1(p), x_2(p), x_3(p)) = 0\}. \tag{8.1.10}$$

La $f(x_1, x_2, x_3) = 0$ si dice *equazione cartesiana* di Q. Ragionando come nel caso delle coniche si vede che la definizione ha senso, cioè se vale (8.1.10) in un sistema di coordinate allora in qualsiasi sistema di coordinate affini Q è l'insieme dei punti le cui coordinate sono gli "zeri"di un polinomio di grado 2 (che dipende dal sistema di coordinate !). La dimostrazione della **Proposizione 8.1.1** si può adattare per dare equazioni canoniche delle quadriche e per dimostrare che l'equazione canonica è unica. Nella Tabella (8.2) abbiamo elencato le equazioni canoniche delle quadriche cosidette *non-degeneri*, le altre (quelle cosidette degeneri) sono coni, cilindri, coppie di piani, piani doppi, rette, punti o l'insieme vuoto.

8.2 Spazi affini euclidei

Sia $\mathbb{A}$ uno spazio affine reale con spazio vettoriale associato V. Supponiamo che V sia uno spazio vettoriale euclideo cioè che sia provvisto di un prodotto scalare definito positivo $\langle , \rangle$: diciamo che $\mathbb{A}$ è uno *spazio affine euclideo*. Dati $P, Q \in \mathbb{A}$ definiamo la distanza tra P e Q come

$$d(P, Q) := ||\overrightarrow{PQ}||. \tag{8.2.1}$$

La distanza ha le seguenti proprietà:

(1) $d(P, Q) \geq 0$ e si ha equaglianza solo se $P = Q$.

(2) $d(P,Q) = d(Q,P)$.

(3) Per il **Corollario 7.7.3** vale la *diseguaglianza triangolare*

$$d(P,Q) \leq d(P,R) + d(R,Q). \tag{8.2.2}$$

Il primo esempio che viene in mente è lo spazio affine $\mathcal{A}^2$ con la distanza determinata dalla scelta di una unità di misura. In uno spazio affine euclideo $\mathbb{A}$ di dimensione n ha senso la nozione di *sistema di riferimento ortonormale*: è un sistema di riferimento affine $RA(O;\mathbf{i}_1,\ldots,\mathbf{i}_n)$ tale che la base di V è ortonormale. Denoteremo $RA(O;\mathbf{i}_1,\ldots,\mathbf{i}_n)$ con $RC(O;\mathbf{i}_1,\ldots,\mathbf{i}_n)$ per sottolineare che il riferimento è ortonormale. Supponiamo che $P,Q \in \mathbb{A}$ abbiano coordinate $(x_1,\ldots,x_n)$ e $(y_1,\ldots,y_n)$ nel riferimento $RC(O;\mathbf{i}_1,\ldots,\mathbf{i}_n)$: allora

$$d(P,Q) = \sum_{i=1}^{n}(x_i - y_i)^2. \tag{8.2.3}$$

Siano X e Y le coordinate relative a due sistemi di riferimento ortonormali: la relazione che lega le coordinate dello stesso punto nei due sistemi di coordinate è $X = A \cdot Y + B$ dove $A \in M_{n,n}(\mathbb{R})$ è ortogonale e $B \in M_{n,1}(\mathbb{R})$.

Definizione 8.2.1. Sia $\mathbb{A}$ uno spazio affine euclideo. Un'applicazione $f\colon \mathbb{A} \to \mathbb{A}$ è una isometria se conserva le distanze cioè se per ogni $P,Q \in \mathbb{A}$ si ha che $d(P,Q) = d(f(P),f(Q))$.

Proposizione 8.2.2. *Sia $\mathbb{A}$ uno spazio affine euclideo con spazio vettoriale associato V. Sia $f\colon \mathbb{A} \to \mathbb{A}$. Allora f è una isometria se e solo se è un'affinità e l'applicazione lineare associata $F\colon V \to V$ è una isometria $F \in O(V)$.*

Dimostrazione. Supponiamo che f sia una isometria. Siccome f preserva le distanze manda parallelogrammi in parallelogrammi e quindi è un'affinità. $F\colon V \to V$ è una isometria perché lo è f. Il viceversa (se f è un'affinità con applicazione lineare associata $F \in O(V)$ allora f è una isometria) è immediato. □

8.3 Forma canonica euclidea di coniche e quadriche

Sia $\mathbb{A}$ uno spazio affine euclideo di dimensione 2, con spazio vettoriale associato V e prodotto scalare $\langle , \rangle$. Sia $C \subset \mathbb{A}$ una conica. Sappiamo (vedi la **Sezione 8.1**) che esiste un sistema di coordinate affini nelle quali C è data da una delle equazioni cartesiane elencate in (8.1). Siccome $\mathbb{A}$ è uno spazio affine euclideo ha senso considerare sistemi di riferimento cartesiani *ortonormali* $RA(O;\mathbf{i},\mathbf{j})$ ciè tali che $\{\mathbf{i},\mathbf{j}\}$ sia una base ortonormale di V. Se X e Y sono coordinate relative a due tali riferimenti la relazione tra le coordinate di uno stesso punto p è data da

$$X(p) = A \cdot Y(p) + B, \qquad A \in \mathcal{O}_2(\mathbb{R}). \tag{8.3.1}$$

Il Teorema spettrale per operatori simmetrici, nella versione data dal **Teorema 9.6.8**, permette di dimostrare che esiste un sistema di riferimento cartesiano ortonormale $RA(O;\mathbf{i},\mathbf{j})$ nel quale C è data da una delle equazioni cartesiane elencate in (8.1). La dimostrazione è un semplice adattamento della dimostrazione del **Proposizione 8.1.1** - lasciamo i dettagli al lettore. Un risultato analogo si dimostra per le quadriche in uno spazio affine euclideo di dimostrazione 3.

Osservazione 8.3.1. Sia $\mathbb{A}$ uno piano affine euclideo a $L \subset \mathbb{A}$ una retta. La *riflessione* nella retta L è l'unica isometria $\rho_L\colon \mathbb{A} \to \mathbb{A}$ tale che ρ_L ristretta a L è l'identità, $\rho_L \circ \rho_L = \mathrm{Id}_{\mathbb{A}}$ e $\rho_L \neq \mathrm{Id}_{\mathbb{A}}$. Se (x_1,x_2) sono coordinate di un sistema di riferimento ortonormale e L è l'asse delle x_1 allora ρ_L porta il punto di coordinate (x_1,x_2) nel punto di coordinate $(x_1,-x_2)$. Sia $F \subset \mathbb{A}$; diciamo che L è un *asse di simmetria* di F se $\rho_L(F) = F$. Le forme canoniche euclidee delle coniche mostrano che una conica ha almeno un asse di simmetria. Per esempio un ellisse ha due assi di simmetria (ortogonali tra loro), una parabola ne ha uno.

Tabella 8.3: Equazione canonica euclidea delle coniche in $\mathbb{A}^2_{\mathbb{R}}$

Equazione canonica	Nome	
$\frac{x_1^2}{a_1^2} + \frac{x_x^2}{a_2^2} - 1 = 0,\ a_1 \geq a_2 > 0$	ellisse	coniche non-degeneri
$\frac{x_1^2}{a_1^2} + \frac{x_x^2}{a_2^2} + 1 = 0\ a_1 \geq a_2 > 0$	ellisse complessa	
$\frac{x_1^2}{a_1^2} - \frac{x_x^2}{a_2^2} - 1 = 0\ a_1 > 0,\ a_2 > 0$	iperbole	
$\frac{x_1^2}{a^2} - x_2 = 0,\ a > 0$	parabola	
$\frac{x_1^2}{a_1^2} + \frac{x_x^2}{a_2^2} = 0,\ a_1 \geq a_2 > 0$	coppia di rette complesse coniugate	coniche degeneneri
$\frac{x_1^2}{a_1^2} - \frac{x_x^2}{a_2^2} = 0,\ a_1 \geq a_2 > 0$	coppia di rette incidenti	
$\frac{x_1^2}{a^2} - 1 = 0,\ a > 0$	coppia di rette parallele	
$x_1^2 = 0$	retta doppia	

Esercizi del Capitolo 8

Esercizio 8.1. *Siano (x, y) coordinate affini su $\mathcal{A}^2$. Per ciascuna delle seguenti coniche determinate se è degenere/non-degenere e se è non-degenere dite se è un'ellisse, un'iperbole...*

$C_1 : 2x^2 - 10xy - y^2 - 2x - 4y = 0, \quad C_2 : xy - 3x + y - 3 = 0, \quad C_3 : 9x^2 - 6xy + y^2 - 7x + y - 1 = 0.$

Esercizio 8.2. *Sia $\mathbb{A}$ un piano affine euclideo e (x_1, x_2) coordinate di un sistema di riferimento ortonormale. Sia $C \subset \mathbb{A}$ l'ellisse di equazione cartesiana*

$$3x_1^2 + 2x_1x_2 + 2x_2^2 + x_2 = 0. \tag{8.3.2}$$

Date equazioni cartesiane degli assi di simmetria di C.

Capitolo 9

Endomorfismi

Il problema di cui ci occuperemo prevalentemente sarà il seguente: dato un endomorfismo f di uno spazio vettoriale finitamente generato dare una base tale che la matrice associata a f sia semplice, per esempio diagonale.

9.1 Motivazione

Sia V uno spazio vettoriale finitamente generato su k e $f\colon V \to V$ un suo endomorfismo. Ci porremo il problema di trovare una base $\mathcal{B}$ che renda facile analizzare f attraverso la matrice $M_{\mathcal{B}}^{\mathcal{B}}(f)$. L'ideale è trovare una $\mathcal{B}$ tale che $M_{\mathcal{B}}^{\mathcal{B}}(f)$ sia una matrice diagonale. Se la base $\mathcal{B} = \{v_1, \ldots, v_n\}$ è tale che $M_{\mathcal{B}}^{\mathcal{B}}(f)$ è diagonale si dice che *diagonalizza* f: equivale a richiedere che esistano $\lambda_1, \ldots, \lambda_n$ (qui $n := \dim V$) tali che

$$f(v_i) = \lambda_i v_i \qquad i = 1, \ldots, n. \tag{9.1.1}$$

Attenzione: esistono f per cui *non* esiste una base che diagonalizza f, vedi l'**Esempio 9.3.1** e l'**Esempio 9.3.2**, se esiste una tale base diciamo che f è *diagonalizzabile*. Notate la somiglianza con il problema di diagonalizzare una forma quadratica (o bilineare simmetrica). In questa sezione daremo qualche motivazione per questo problema. Sia $A \in M_{n,n}(k)$. Per la **Proposizione 4.9.6** trovare una base che diagonalizza L_A equivale a trovare $G \in \mathrm{GL}_n(k)$ tale che $\Lambda = G^{-1} \cdot A \cdot G$ sia diagonale - esplicitamente le colonne di G sono i vettori della base $\mathcal{B}$ che diagonalizza. Quindi se abbiamo trovato un tale G abbiamo che $A = G \cdot \Lambda \cdot G^{-1}$. Questa uguaglianza ci permette di calcolare facilmente tutte le potenze A^r perchè

$$A^r = G \cdot \Lambda^r \cdot G^{-1} \tag{9.1.2}$$

e Λ^r è una matrice diagonale con entrate le potenze r-esime delle entrate di Λ. Sia A a entrate reali; le potenze di A intervengono nell'esponenziale di A, che si definisce come segue. Consideriamo la somma

$$\sigma_r := 1_n + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \ldots + \frac{1}{r!}A^r. \tag{9.1.3}$$

Se $|a_{ij}| \leq S$ per ogni $1 \leq i, j \leq n$ 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, la denotiamo e^A, e scriviamo

$$e^A := \sum_{r=0}^{\infty} \frac{1}{r!}A^r = 1_n + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \ldots + \frac{1}{r!}A^r + \ldots \tag{9.1.4}$$

L'esponenziale è importante perché vale una relazione analoga all'uguaglianza $d(e^t)/dt = e^t$:

$$\frac{d}{dt}e^{tA} = A \cdot e^{tA}. \tag{9.1.5}$$

Quindi una soluzione del sistema di equazioni differenziali nelle funzioni $y_1, \ldots, y_n\colon \mathbb{R} \to \mathbb{R}$ dato da

$$Y(t)' = A \cdot Y(t),, \qquad Y(0) = B, \tag{9.1.6}$$

($Y\colon \mathbb{R} \to \mathbb{R}^n$ è la funzione con entrate $y_1, \dots, y_n$, 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 (9.1.2). Allora si ha che

$$e^{tA} = G \cdot (1_n + t\Lambda + \tfrac{t^2}{2}\Lambda^2 + \tfrac{t^3}{3!}A^3 + \tfrac{t^r}{r!}\Lambda^r + \dots) \cdot G^{-1} = G \cdot \begin{bmatrix} e^{t\lambda_1} & 0 & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & \dots & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \dots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & e^{t\lambda_n} \end{bmatrix} \cdot G^{-1} \tag{9.1.7}$$

9.2 Autovalori, autospazi

Sia V uno spazio vettoriale su k e $f\colon V \to V$ un endomorfismo. Sia $\lambda \in k$: poniamo

$$V_\lambda(f) := \ker(f - \lambda \operatorname{Id}_V). \tag{9.2.1}$$

Definizione 9.2.1. Un $\lambda \in k$ 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.

La seguente osservazione è già stata fatta informalmente, e giustifica la **Definizione 9.2.1**.

Osservazione 9.2.2. Sia V uno spazio vettoriale finitamente generato su 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 9.2.3. Sia

$$A := \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$$

e $L_A\colon \mathbb{R}^2 \to \mathbb{R}^2$ l'applicazione lineare associata. Allora 2 e 3 sono autovalori di A e gli autospazi relativi sono

$$V_2(L_A) = \{(t, -t) \mid t \in \mathbb{R}\}, \qquad V_3(L_A) = \{(t, -2t) \mid t \in \mathbb{R}\}.$$

Esempio 9.2.4. Sia $V := C^\infty(\mathbb{R})$ e $\Phi\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ definita da $\Phi(f) := f''$. 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 $-k^2$.

Esempio 9.2.5. Sia $A \in M_{2,2}(\mathbb{R})$ la matrice dell'**Esempio 9.2.3**. Allora $\mathcal{B} := \{(1, -1), (1, -2)\}$ è una base di $\mathbb{R}^2$ i cui elementi sono autovettori di A, quindi A è diagonalizzabile. Esplicitamente

$$M_{\mathcal{B}}^{\mathcal{B}}(L_A) = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

e perciò

$$\begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix} = M_{\mathcal{S}}^{\mathcal{S}}(L_A) = M_{\mathcal{S}}^{\mathcal{B}}(\operatorname{Id}_{\mathbb{R}^2}) \cdot M_{\mathcal{B}}^{\mathcal{B}}(L_A) \cdot M_{\mathcal{B}}^{\mathcal{S}}(\operatorname{Id}_{\mathbb{R}^2}) = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}^{-1}. \tag{9.2.2}$$

Notiamo che da (9.2.2) segue che

$$\begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}^m = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \cdot \begin{bmatrix} 2^m & 0 \\ 0 & 3^m \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 2^{m+1} - 3^m & 2^m - 3^m \\ -2^{m+1} + 2 \cdot 3^m & -2^m + 2 \cdot 3^m \end{bmatrix},$$

e che

$$e^A = \begin{bmatrix} 2e^2 - e^3 & e^2 - e^3 \\ -2e^2 + 2e^3 & -e^2 + 2e^3 \end{bmatrix}.$$

Proposizione 9.2.6. *Siano V uno spazio vettoriale finitamente generato di dimensione n, e $f\colon V \to V$ un suo endomorfismo. Allora $\mathrm{Det}(\lambda \mathrm{Id}_V - f)$ è un polinomio in λ a coefficienti in k, monico di grado n (monico significa che il coefficiente di λ^n è uguale a 1).*

Dimostrazione. Conseguenza immediata della **Proposizione 7.8.3**. □

Definizione 9.2.7. Sia V uno spazio vettoriale finitamente generato su k e $f\colon V \to V$ un endomorfismo. Il *polinomio caratteristico* di f è

$$P_f := \mathrm{Det}(\lambda \mathrm{Id}_V - f) \in k[\lambda]. \tag{9.2.3}$$

Notiamo che, se $A \in M_{n,n}(k)$, il polinomio caratteristico di L_A è uguale al polinomio caratteristico di A:

$$P_{L_A} = P_A. \tag{9.2.4}$$

Osservazione 9.2.8. Sia V uno spazio vettoriale finitamente generato su k e $f\colon V \to V$ un endomorfismo. Un $\lambda_0 \in k$ è un autovalore di f se e solo se è una radice di P_f. Infatti λ_0 è un autovalore di f se e solo se $\ker(\lambda_0 \mathrm{Id}_V - f) \neq \{0\}$ ovvero se e solo se esiste $0 \neq v$ tale che $f(v) = \lambda_0 v$. Vediamo anche che se $\lambda_0 \in k$ è un autovalore di f allora gli autovettori con autovalore λ_0 sono gli elementi non-nulli di $\ker(\lambda_0 \mathrm{Id}_V - f)$.

Esempio 9.2.9. Sia $A \in M_{2,2}(\mathbb{R})$ data da

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

Allora

$$P_A = \det \begin{bmatrix} \lambda - 1 & -1 \\ -1 & \lambda \end{bmatrix} = \lambda^2 - \lambda - 1.$$

Quindi gli autovalori di A sono

$$\frac{1+\sqrt{5}}{2}, \quad \frac{1-\sqrt{5}}{2}$$

e si trova che

$$L_A(2,-1+\sqrt{5}) = (1+\sqrt{5},2) = (\tfrac{1+\sqrt{5}}{2})(2,-1+\sqrt{5}), \qquad L_A(2,-1-\sqrt{5}) = (1-\sqrt{5},2) = (\tfrac{1-\sqrt{5}}{2})(2,-1-\sqrt{5}).$$

Perciò la base $\mathcal{B} := \{(2,-1+\sqrt{5}),(-2,1+\sqrt{5})\}$ diagonalizza A. Segue che

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = M_{\mathcal{S}}^{\mathcal{S}}(L_A) = M_{\mathcal{S}}^{\mathcal{B}}(\mathrm{Id}_{\mathbb{R}^2}) \cdot M_{\mathcal{B}}^{\mathcal{B}}(L_A) \cdot M_{\mathcal{B}}^{\mathcal{S}}(\mathrm{Id}_{\mathbb{R}^2}) = \begin{bmatrix} 2 & -2 \\ -1+\sqrt{5} & 1+\sqrt{5} \end{bmatrix} \cdot \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} \cdot \begin{bmatrix} 2 & -2 \\ -1+\sqrt{5} & 1+\sqrt{5} \end{bmatrix}^{-1}.$$

Un facile calcolo dà che

$$\begin{bmatrix} 2 & -2 \\ -1+\sqrt{5} & 1+\sqrt{5} \end{bmatrix}^{-1} = \frac{1}{4\sqrt{5}} \begin{bmatrix} 1+\sqrt{5} & 2 \\ 1-\sqrt{5} & 2 \end{bmatrix}$$

Segue (con una serie di calcoli) che

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^m = \frac{1}{\sqrt{5}} \begin{bmatrix} (\frac{1+\sqrt{5}}{2})^{m+1} - (\frac{1-\sqrt{5}}{2})^{m+1} & (\frac{1+\sqrt{5}}{2})^m - (\frac{1-\sqrt{5}}{2})^m \\ (\frac{1+\sqrt{5}}{2})^m - (\frac{1-\sqrt{5}}{2})^m & (\frac{1+\sqrt{5}}{2})^{m-1} - (\frac{1-\sqrt{5}}{2})^{m-1} \end{bmatrix} \tag{9.2.5}$$

Questa uguaglianza dà una formula chiusa per i numeri di Fibonacci (vedi l'**Esempio 4.4.17**). Ricordiamo che la successione di Fibonacci $\{x_m\}_{m\geq 0}$ è definita ricorsivamente, ponendo $1 = x_0 = x_1$ e $x_m = x_{m-1} + x_{m-2}$ per $m \geq 2$. I numeri di Fibonacci sono i termini della successione di Fibonacci. Dimostriamo che

$$x_m = \frac{1}{\sqrt{5}} \cdot \left(\left(\frac{1+\sqrt{5}}{2} \right)^{m+1} - \left(\frac{1-\sqrt{5}}{2} \right)^{m+1} \right). \tag{9.2.6}$$

Infatti, come visto nell'**Esempio 4.4.17**,

$$\begin{bmatrix} x_{m+1} \\ x_m \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^m \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{9.2.7}$$

Applicando (9.2.5), si trova la formula (9.2.6), se si tiene conto dell'uguaglianza $\lambda^2 = \lambda + 1$ per $\lambda = (1 \pm \sqrt{5})/2$.

Esempio 9.2.10. Se $A \in M_{n,n}(\mathbb{R})$ è simmetrica, allora A è diagonalizzabile per il Teorema spettrale (cioè il **Teorema 7.8.1**). Infatti, per il Teorema spettrale esiste una base ON $\mathcal{B} = \{v_1, \ldots, v_n\}$ di $\mathbb{R}^n$ (con prodotto euclideo standard) tale che la forma quadratica $f \in Q(\mathbb{R}^n)$ definita $f(X) := X^t \cdot A \cdot X$ è diagonale nella base $\mathcal{B}$, cioè

$$f\left(\sum_{i=1}^{n} x_i v_i\right) = \sum \lambda_i x_i^2.$$

Sia $G := M_{\mathcal{B}}^{\mathcal{S}}(\mathrm{Id}_{\mathbb{R}^n})$ la matrice del cambiamento di base, dalla base standard $\mathcal{S} = \{e_1, \ldots, e_n\}$ alla base $\mathcal{B}$; allora

$$A = G^t \cdot \Lambda \cdot G, \tag{9.2.8}$$

dove Λ è la matrice diagonale cone entrata $\lambda_i \delta_{ij}$ su riga i e colonna j. D'altra parte $G^t \cdot G = 1_n$ perché la base $\mathcal{B}$ è ON per il prodotto euclideo standard, cioè $G^t = G^{-1}$. Quindi (9.2.8) dà che A è anche coniugata a Λ, e perciò è diagonalizzabile.

9.3 Molteplicità algebrica e geometrica di autovalori

I seguenti esempi dimostrano che esistono endomorfismi che *non* sono diagonalizzabili.

Esempio 9.3.1. Sia k un campo e $A \in M_{2,2}(k)$ data da

$$A := \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}.$$

Allora

$$P_A = \det \begin{bmatrix} \lambda - 3 & -1 \\ 0 & \lambda - 3 \end{bmatrix} = (\lambda - 3)^2.$$

Quindi 3 è l'unico autovalore di A. Un facile calcolo dà che gli autovettori sono i vettori $(a, 0)$ dove $a \neq 0$. Segue che A *non* è diagonalizzabile.

Esempio 9.3.2. Sia $A \in M_{2,2}(\mathbb{R})$ data da

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

Allora

$$P_A = \det \begin{bmatrix} \lambda & 1 \\ -1 & \lambda \end{bmatrix} = \lambda^2 + 1.$$

Quindi A *non* ha autovalori e in particolare non è diagonalizzabile. Ora consideriamo A come elemento di $M_{2,2}(\mathbb{C})$: allora P_A ha le due radici i e $-i$ e quindi A ha due autovalori. Un facile conto dà che $(1, -i)$ e $(1, i)$ sono autovettori di A:

$$L_A(1, -i) = i(1, -i), \qquad L_A(1, i) = -i(1, i).$$

Perciò la base $\mathcal{B} := \{(1, -i), (1, i)\}$ diagonalizza A.

Osservazione 9.3.3. L'**Esempio 9.3.1** e l'**Esempio 9.3.2** sono sostanzialmente diversi. Nel primo esempio la matrice A non è diagonalizzabile qualsiasi sia il campo che si considera (i numeri 1 e 3 hanno senso in qualsiasi campo). Nel secondo esempio A è diagonalizzabile se il campo contiene una radice di -1, cioè l'eventuale non diagonalizzabilità è imputabile a inadeguatezza del campo.

Lemma 9.3.4. *Sia V uno spazio vettoriale finitamente generato, di dimensione n. Se un endomorfismo $f\colon V \to V$ è diagonalizzabile, allora P_f ha n radici (contate con molteplicità) in k, cioè esistono $\lambda_1, \dots, \lambda_n \in k$ tali che $P_f = \prod_{i=1}^{n}(\lambda - \lambda_i)$.*

Dimostrazione. Sia $\mathcal{B}$ una base di V che diagonalizzi f e quindi esistono $\lambda_1, \dots, \lambda_n \in k$ tali che $M_{\mathcal{B}}^{\mathcal{B}}(f) = (\lambda_i \delta_{ij})$. Allora $P_f = \prod_{i=1}^{n}(\lambda - \lambda_i)$. □

L'**Esempio 9.3.1** dimostra che *non* vale il viceversa del **Lemma 9.3.4**, cioè *non* è vero che se P_f è prodotto di fattori lineari allora f è diagonalizzabile. Vedremo quale altra ipotesi occorre fare per garantire che un endomorfismo sia diagonalizzabile.

Proposizione 9.3.5. *Sia V uno spazio vettoriale finitamente generato su k. Siano $f\colon V \to V$ un endomorfismo e $\lambda_0 \in k$ un autovalore di f. Allora*

$$1 \leq \dim V_{\lambda_0}(f) \leq \operatorname{mult}_{\lambda_0} P_f. \tag{9.3.1}$$

Dimostrazione. La diseguaglianza di sinistra di (9.3.1) vale per l'**Osservazione 9.2.8**. Rimane da dimostrare che vale la diseguaglianza di destra di (9.3.1). Sia $\{v_1, \dots, v_{\dim V_{\lambda_0}}\}$ una base di $V_{\lambda_0}(f)$ ed estendiamola a una base $\mathcal{B} = \{v_1, \dots, v_n\}$ di V. Quindi $f(v_i) = \lambda_0 v_i$ per $1 \leq i \leq \dim V_{\lambda_0}(f)$. Abbiamo che

$$M_{\mathcal{B}}^{\mathcal{B}}(\lambda \operatorname{Id}_V - f) = \begin{bmatrix} (\lambda-\lambda_0) & 0 & \dots & 0 & * & \dots & * \\ 0 & (\lambda-\lambda_0) & \dots & \vdots & \vdots & \vdots & \vdots \\ \vdots & 0 & \dots & 0 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \dots & (\lambda-\lambda_0) & \vdots & \vdots & \vdots \\ \vdots & \vdots & \dots & 0 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & * & \dots & * \end{bmatrix} \tag{9.3.2}$$

dove il numero di colonne in cui appare $(\lambda - \lambda_0)$ è uguale a $\dim V_{\lambda_0}$. Sviluppando il determinante secondo la prima colonna e iterando troviamo che $P_f = (\lambda - \lambda_0)^r \cdot q$ dove $r := \dim V_{\lambda_0}(f)$ e $q \in k[\lambda]$. Quindi $\operatorname{mult}_{\lambda_0} P_f \geq r = \dim V_{\lambda_0}(f)$. □

Corollario 9.3.6. *Sia V uno spazio vettoriale su k e $f\colon V \to V$ un endomorfismo. Allora*

$$\sum_{\lambda \in k} \dim V_\lambda(f) \leq \dim V. \tag{9.3.3}$$

Dimostrazione. Per la **Proposizione 9.3.5** e la disequazione (1.6.6) abbiamo che

$$\sum_{\lambda \in k} \dim V_\lambda(f) \leq \sum_{\lambda \in k} \operatorname{mult}_\lambda P_f \leq \deg P_f. \tag{9.3.4}$$

Siccome $\deg P_f = \dim V$ - vedi il **Lemma 9.3.4** - segue il risultato. □

Esempio 9.3.7. Sia

$$A = \begin{bmatrix} \lambda_0 & 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \dots & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \dots & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & \dots & \vdots & \ddots & \ddots & 1 \\ 0 & \dots & \dots & \dots & \dots & 0 & \lambda_0 \end{bmatrix} \tag{9.3.5}$$

(Le entrate sulla diagonale principale sono uguali a λ_0, quelle immediatamente sopra sono uguali a 1, le altre sono 0.) Abbiamo che

$$P_A = (\lambda - \lambda_0)^n$$

quindi l'unico autovalore di A è λ_0 e $\mathrm{mult}_{\lambda_0} P_A = n$. Inoltre $\dim V_{\lambda_0}(L_A) = 1$. Quindi in questo caso la (9.3.3) è una diseguaglianza stretta a meno che $n = 1$. Vediamo anche che se $n > 1$ allora A *non* è diagonalizzabile.

Lemma 9.3.8. *Sia V uno spazio vettoriale su k e $f\colon V \to V$ un endomorfismo. Siano $v_1, \ldots, v_d \in V$ autovettori con autovalori $\lambda_1, \ldots, \lambda_d$ a due a due distinti. Allora $v_1, \ldots, v_d$ sono linearmente indipendenti.*

Dimostrazione. Per induzione su d. Se $d = 1$ il risultato è vero perché 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 $\alpha_1, \ldots, \alpha_d \in k$ non tutti nulli tali che

$$0 = \alpha_1 v_1 + \ldots + \alpha_d v_d. \tag{9.3.6}$$

Di fatto

$$\alpha_i \neq 0 \quad \forall 1 \leq i \leq d. \tag{9.3.7}$$

Infatti 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. Ora applichiamo f:

$$0 = f(0) = f(\alpha_1 v_1 + \ldots + \alpha_d v_d) = \alpha_1 \lambda_1 v_1 + \ldots + \alpha_d \lambda_d v_d. \tag{9.3.8}$$

Siccome $d > 1$ e gli autovalori sono distinti esiste un λ_i non nullo: riordinando possiamo assumere che sia λ_d. Moltiplicando (9.3.8) per λ_d^{-1} otteniamo che

$$0 = \alpha_1 \lambda_d^{-1} \lambda_1 v_1 + \ldots + \alpha_{d-1} \lambda_d^{-1} \lambda_{d-1} v_d + \alpha_d v_d. \tag{9.3.9}$$

Sottraendo (9.3.9) da (9.3.6) si ha che

$$\alpha_1 (1 - \lambda_d^{-1} \lambda_1) v_1 + \ldots + \alpha_{d-1} (1 - \lambda_d^{-1} \lambda_{d-1}) v_{d-1} = 0. \tag{9.3.10}$$

Sia $1 \leq i \leq (d-1)$. Per (9.3.7) sappiamo che $\alpha_i \neq 0$ 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 (9.3.10) 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. □

Proposizione 9.3.9. *Sia V uno spazio vettoriale finitamente generato su k e $f\colon V \to V$ un endomorfismo. Allora f è diagonalizzabile se e solo se il numero di radici di P_f in k (contate con molteplicità) è uguale al grado di P_f, e per ogni autovalore λ di f si ha che*

$$\dim V_\lambda(f) = \mathrm{mult}_\lambda(P_f). \tag{9.3.11}$$

Dimostrazione. Supponiamo che esista una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ che diagonalizza f. Il numero di radici di P_f in k (contate con molteplicità) è uguale alla dimensione di V per il **Lemma 9.3.4**. Per l'**Osservazione 9.2.2** ciascun v_i è un autovettore di f: sia λ_i l'autovalore associato. Sia λ_j un autovalore di f; l'espressione di P_f data nel **Lemma 9.3.4** mostra che

$$\mathrm{mult}_{\lambda_j}(P_f) = |\{1 \leq i \leq n \mid \lambda_i = \lambda_j\}|. \tag{9.3.12}$$

Siccome ogni v_i tale che $\lambda_i = \lambda_j$ appartiene a $V_{\lambda_j}(f)$ vediamo anche che $\dim V_{\lambda_j}(f) = \mathrm{mult}_{\lambda_j}(P_f)$, e quindi si ha equaglianza per la **Proposizione 9.3.5**. Ora supponiamo che il numero di radici di P_f in k (contate con molteplicità) è uguale alla dimensione di V, e che valga (9.3.11) per ogni autovalore λ di f: dimostriamo che f è diagonalizzabile. Siano $\lambda_1, \ldots, \lambda_d$ gli autovalori *distinti* di f. Per $1 \leq i \leq d$ 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

$$\{v_{1,1}, \dots, v_{1,n(1)}, \dots, v_{i,1}, \dots, v_{i,n(i)}, \dots v_{d,1}, \dots, v_{d,n(d)}\} \tag{9.3.13}$$

è una base di V. Applicando il **Lemma 9.3.8** si vede che i vettori di (9.3.13) sono linearmente indipendenti, d'altra parte il loro numero è

$$n(1) + n(2) + \ldots + n(d) = \sum_{\lambda \in k} \dim V_\lambda(f) = \sum_{\lambda \in k} \mathrm{mult}_\lambda(P_f) = \deg P_f = \dim V.$$

(La seconda uguaglianza segue da (9.3.11), la terza dall'ipotesi che il numero di radici di P_f in k (contate con molteplicità) è uguale al grado di P_f.) Segue che (9.3.13) è una base di V. Siccome i vettori della base (9.3.13) sono autovettori di f la f è diagonalizzabile - vedi l'**Osservazione 9.2.2**. □

Corollario 9.3.10. *Sia V uno spazio vettoriale su k di dimensione n Sia $f\colon V \to V$ un endomorfismo. Se P_f ha n radici distinte (in k) allora f è diagonalizzabile.*

9.4 Forme Hermitiane

9.4.1 Definizione di forma Hermitiana

Per spazi vettoriali reali esiste la nozione di prodotto scalare definito positivo e quindi di spazio vettoriale euclideo. Se V è uno spazio vettoriale *complesso* la nozione di forma bilineare simmetrica definita positiva non ha senso. Per poter definire l'analogo complesso di spazio vettoriale euclideo introduciamo le forme hermitiane.

Definizione 9.4.1. Sia V uno spazio vettoriale complesso. Una *forma hermitiana* su V è una funzione

$$\begin{array}{ccc} V \times V & \xrightarrow{H} & \mathbb{C} \\ (v, w) & \mapsto & H(v, w) \end{array} \tag{9.4.1}$$

che gode delle seguenti proprietà.

1. Dato $w_0 \in V$ la funzione $V \to \mathbb{C}$ definita da $v \mapsto H(v, w)$ è lineare. (Cioè $H(,)$ è lineare a sinistra.)
2. $H(w, v) = \overline{H(v, w)}$.

Esempio 9.4.2. Sia $A \in M_{n,n}(\mathbb{C})$ e definiamo $H\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ ponendo

$$H(X, Y) := X^t \cdot A \cdot \overline{Y}, \qquad X, Y \in \mathbb{C}^n. \tag{9.4.2}$$

Allora H è un prodotto hermitiano su $\mathbb{C}^n$ se e solo se

$$A^t = \overline{A}. \tag{9.4.3}$$

Infatti $H(,)$ è lineare a sinistra qualsiasi sia A, e inoltre

$$\overline{H(Y, X)} = \overline{Y^t \cdot A \cdot \overline{X}} = \overline{(Y^t \cdot A \cdot \overline{X})^t} = \overline{\overline{X}^t \cdot A^t \cdot Y} = X^t \cdot \overline{A}^t \cdot \overline{Y}. \tag{9.4.4}$$

Segue che $H(Y, X) = \overline{H(X, Y)}$ per ogni $X, Y \in \mathbb{C}^n$ se e solo se $\overline{A}^t = A$, ovvero vale (9.4.3). Se $H = 1_n$ otteniamo la forma hermitiana *standard* su $\mathbb{C}^n$

$$\langle X, Y \rangle = X^t \cdot \overline{Y}. \tag{9.4.5}$$

(In questo caso denotiamo $H(X, Y)$ con $\langle X, Y \rangle$, vedi la **Sottosezione 9.4.2**.)

Esempio 9.4.3. 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

$$H(f,g):=\int_{-\pi}^{\pi} f(x)\overline{g(x)}dx,\qquad f,g\in C^0([-\pi,\pi])_{\mathbb{C}}. \tag{9.4.6}$$

L'integrale di Riemann è definito calcolando la parte reale e immaginaria della funzione $f(x)\overline{g(x)}$. Esplicitamente scriviamo $f(x)\overline{g(x)}=u(x)+iv(x)$ dove $u,v\colon[-\pi,\pi]\to\mathbb{R}$ sono continue fuori da un sottoinsieme finito di $[-\pi,\pi]$: allora

$$\int_{-\pi}^{\pi} f(x)\overline{g(x)}dx:=\int_{-\pi}^{\pi}u(x)dx+i\int_{-\pi}^{\pi}v(x)dx. \tag{9.4.7}$$

Si verifica facilmente che $H(,)$ è un prodotto scalare hermitiano su $C^0([-\pi,\pi])_{\mathbb{C}}$.

Osservazione 9.4.4. Sia V uno spazio vettoriale complesso e H una forma hermitiana su V.

1. Siano $v,w_1,w_2\in V$ e $\lambda_1,\lambda_2\in\mathbb{C}$. Si ha che
$$H(v,\lambda_1w_1+\lambda_2w_2)=\overline{H(\lambda_1w_1+\lambda_2w_2,v)}=\overline{\lambda}_1H(w_1,v)+\overline{\lambda}_2H(w_2,v). \tag{9.4.8}$$
2. Sia $v\in V$; siccome $H(v,v)=\overline{H(v,v)}$ vediamo che $H(v,v)$ è un numero reale.
3. Vettori $v,w\in V$ sono *ortogonali* se $H(v,w)=0$, in simboli $v\perp w$. Notate che i $v\perp w$ se e solo se $w\perp v$. Sia $S\subset V$ un sottoinsieme. L'*ortogonale* di S è
$$S^{\perp}:=\{w\in V\mid v\perp w\quad\forall v\in S\}. \tag{9.4.9}$$
Se $S=\{v\}$ denotiamo $S^{\perp}$ con $v^{\perp}$. Notiamo che $v^{\perp}$ è il nucleo dell'applicazione lineare
$$\begin{array}{ccc} V & \longrightarrow & \mathbb{C}\\ w & \mapsto & H(w,v)\end{array} \tag{9.4.10}$$
e quindi è un sottospazio lineare di V. Siccome $S^{\perp}$ è l'intersezione dei sottospazi $v^{\perp}$ per $v\in S$ ne segue che $S^{\perp}$ è un sottospazio di V.

9.4.2 Forme Hermitiane definite positive

Definizione 9.4.5. Sia V uno spazio vettoriale complesso. Una forma hermitiana H su V è *definita positiva* se per ogni $0\neq v\in V$ si ha che
$$H(v,v)>0. \tag{9.4.11}$$
Una forma hermitiana su V definita positiva verrà denotata spesso con $\langle,\rangle$, cioè porremo $\langle v,w\rangle=H(v,w)$.

Esempio 9.4.6. La forma hermitiana standard, vedi (9.4.5), è definita positiva perché
$$\langle X,X\rangle=X^t\cdot\overline{X}=\sum_{j=1}^{n}|x_j|^2. \tag{9.4.12}$$
La forma hermitiana dell'**Esempio 9.4.3** è definita positiva.

Definizione 9.4.7. Sia V uno spazio vettoriale complesso e $\langle,\rangle$ una forma hermitiana definita positiva su V.

1. La *norma* di $v\in V$ è data da $||v||:=\langle v,v\rangle^{1/2}$.
2. Il *gruppo unitario* di $(V,\langle,\rangle)$ è l'insieme $U(V)$ delle applicazioni lineari invertibili $f\colon V\to V$ tali che
$$\langle f(v),f(w)\rangle=\langle v,w\rangle\quad\forall v,w\in V. \tag{9.4.13}$$
(Chiamiamo f un *operatore unitario*.) Il gruppo unitario di $\mathbb{C}^n$ con il prodotto hermitiano standard è denotato $\mathrm{U}(n)$.

3. Supponiamo che V sia finitamente generato. Una base $\{v_1, \dots, v_n\}$ di V è *ortonormale* (abbreviamo con ON) se
$$\langle v_i, v_j \rangle = \delta_{ij} \quad \forall 1 \leq i, j \leq n. \tag{9.4.14}$$

Osservazione 9.4.8. Supponiamo che V sia finitamente generato. Sia $f\colon V \to V$ un'applicazione lineare tale che vale (9.4.13). Allora f è invertibile e perciò è un operatore unitario. Infatti sia $0 \neq v \in V$: allora
$$\langle f(v), f(v) \rangle = \langle v, v \rangle > 0 \tag{9.4.15}$$
e quindi $f(v) \neq 0$. Quindi f è iniettiva e siccome V è finitamente generato segue che f è invertibile.

Osservazione 9.4.9. Supponiamo che V sia finitamente generato. Allora esiste una base ON di V. Infatti si può ragionare per induzione sulla dimensione di V. Se $\dim V = 1$ sia $0 \neq v \in V$. Allora $||v|| \neq 0$ e $\{v/||v||$ è una base ON di V. Per dimostrare il passo induttivo supponiamo che $\dim V = n$ e sia $0 \neq v \in V$. Allora $v^\perp$ ha dimensione $(n-1)$ e per l'ipotesi induttiva (applicata alla restrizione di $\langle , \rangle$ a $v^\perp$) esiste una base ON $\{v_1, \dots, v_{n-1}\}$ di $v^\perp$. Sia $v_n := v/||v||$; allora $\{v_1, \dots, v_n\}$ è una base ON di V.

Osservazione 9.4.10. Supponiamo che V sia finitamente generato e che $\mathcal{B} = \{v_1, \dots, v_n\}$ sia una base ON di V. Allora
$$\langle u, w \rangle = X_{\mathcal{B}}(u)^t \cdot \overline{X}_{\mathcal{B}}(w). \tag{9.4.16}$$
Infatti siano $X := X_{\mathcal{B}}(u)$ e $Y := X_{\mathcal{B}}(w)$. Quindi $u = \sum_{j=1}^n x_j v_j$ e $w = \sum_{k=1}^n y_k v_k$. Siccome $\mathcal{B}$ è una base ON
$$\langle u,w\rangle = \langle \textstyle\sum_{j=1}^n x_j v_j, \sum_{k=1}^n y_k v_k \rangle = \sum_{1 \leq j,k \leq n} \langle x_j v_j, y_k v_k \rangle = \sum_{1 \leq j,k \leq n} x_j \overline{y}_k \langle v_j, v_k \rangle = \sum_{1 \leq j,k \leq n} x_j \overline{y}_k \delta_{jk} = \sum_{j=1}^n x_j \overline{y}_j. \tag{9.4.17}$$

Osservazione 9.4.11. Il gruppo unitario $\mathrm{U}(n)$ è descritto come segue. Sia $A \in M_{n,n}(\mathbb{C})$ allora $L_A \in \mathrm{U}(n)$ se e solo se
$$A^t \cdot \overline{A} = 1_n. \tag{9.4.18}$$
Più in generale supponiamo che V sia finitamente generato, con base ON $\mathcal{B} = \{v_1, \dots, v_n\}$. Sia $f \in \mathrm{End}(V)$; allora $f \in \mathrm{U}(V)$ se e solo se
$$M_{\mathcal{B}}^{\mathcal{B}}(f)^t \cdot \overline{M_{\mathcal{B}}^{\mathcal{B}}(f)} = 1_n. \tag{9.4.19}$$

Proposizione 9.4.12. *Sia V uno spazio vettoriale complesso finitamente generato e $\langle , \rangle$ una forma hermitiana definita positiva 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
$$\langle v, v \rangle = \langle f(v), f(v) \rangle = \langle \lambda v, \lambda v \rangle = |\lambda|^2 \langle v, v \rangle. \tag{9.4.20}$$
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
$$W := v_n^\perp := \{w \in V \mid \langle w, v_n \rangle = 0\}. \tag{9.4.21}$$
Allora $f(W) \subset W$: infatti se $w \in W$ allora
$$0 = \langle w, v_n \rangle = \langle f(w), f(v_n) \rangle = \langle f(w), \lambda_n v_n \rangle = \overline{\lambda_n} \langle f(w), v_n \rangle = 0. \tag{9.4.22}$$
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 $\{v_1, \dots, v_{n-1}\}$ di W che diagonalizza g. Allora $\{v_1, \dots, v_{n-1}, v_n\}$ è una base ON di V che diagonalizza f. □

Osservazione 9.4.13. Supponiamo che V e f siano come nella **Proposizione 9.4.12**. Sia $\{v_1, \dots, v_n\}$ una base ON che diagonalizza f. Gli autovalori $\lambda_1, \dots, \lambda_n$ di f hanno modulo 1 e quindi esistono $\theta_1, \dots, \theta_n \in \mathbb{R}$ tali che $\lambda_j = e^{i\theta_j}$ per $j = 1, \dots, n$. Quindi

$$f(v_j) = e^{i\theta_j} v_j, \qquad 1 \leq j \leq n. \tag{9.4.23}$$

9.4.3 Applicazioni ortogonali di uno spazio vettoriale euclideo

È naturale chiedersi se esiste un risultato analogo della **Proposizione 9.4.12**, quando sostituiamo allo spazio vettoriale complesso con forma hermitiana definita positiva uno spazio vettoriale euclideo. Se pensiamo a una rotazione del piano vediamo che in generale una isometria di uno spazio vettoriale euclideo non è diagonalizzabile; il seguente risultato dà la generalizzazione corretta.

Proposizione 9.4.14. *Sia $(V, (,))$ uno spazio vettoriale euclideo finitamente generato. Sia $f \in \mathrm{O}(V)$. Esistono una base ON $\{v_1 \dots, v_a, z_1, \dots, z_b, u_1, w_1, u_2, w_2, \dots, u_c, w_c\}$ di V e $\theta_1, \dots, \theta_c \in \mathbb{R}$ tali che*

$$f(v_p) = v_p, \quad 1 \leq p \leq a, \qquad f(z_q) = -z_q, \quad 1 \leq q \leq b \tag{9.4.24}$$

e

$$f(u_s) = \cos\theta_s u_s + \sin\theta_s w_s, \quad f(w_s) = -\sin\theta_s u_s + \cos\theta_s w_s, \quad 1 \leq s \leq c. \tag{9.4.25}$$

Dimostrazione. Scegliamo una base ON $\mathcal{B} = \{v_1, \dots, v_n\}$ di V. Allora

$$(v, w) = X_{\mathcal{B}}(v)^t \cdot X_{\mathcal{B}}(w) \qquad \forall v, w \in V. \tag{9.4.26}$$

Ne segue che possiamo assumere che V sia $\mathbb{R}^n$ con il prodotto scalare standard. Quindi $f = L_A$ dove $A \in \mathrm{O}(n)$, cioè $A^t \cdot A = 1_n$. Denotiamo con $F\colon \mathbb{C}^n \to \mathbb{C}^n$ l'applicazione data da L_A dove A è pensata come elemento di $M_{n,n}(\mathbb{C})$: quindi $F(X) = f(X)$ per ogni $X \in \mathbb{R}^n$. Siccome $\overline{A} = A$ segue che $F \in \mathrm{U}(n)$. Per la **Proposizione 9.4.12** sappiamo che F è diagonalizzabile e quindi

$$\sum_{\lambda} \dim V_\lambda(F) = n, \qquad \dim V_\lambda(F) = \mathrm{mult}_\lambda(P_F) \quad \forall\ \lambda \in \mathbb{C}. \tag{9.4.27}$$

Siano $\lambda, \mu \in \mathbb{C}$ autovalori di F e $v, w \in \mathbb{C}^n$ autovettori con autovalori λ e μ rispettivamente. Allora

$$\langle v, w\rangle = \langle F(v), F(w)\rangle = \langle \lambda v, \mu w\rangle = \lambda\overline{\mu}\langle v, w\rangle. \tag{9.4.28}$$

Per la **Proposizione 9.4.12** sappiamo che $1 = |\mu|$ cioè $\overline{\mu} = \mu^{-1}$: ne segue che se $\lambda \neq \mu$ allora $\langle v, w\rangle = 0$. In altre parole vettori appartenenti ad autospazi diversi sono ortogonali. Ora cominciamo a costruire una base ON che diagonalizza f. Per ciascuno degli autovettori reali (che appartengono a $\{1, -1\}$) esiste una base ON *reale* dell'autospazio corrispondente: siano $\{v_1 \dots, v_a\}$ una base ON reale di $V_1(F)$ e $\{z_1, \dots, z_b\}$ una base ON reale di $V_{-1}(F)$. Ora sia λ un autovalore *non* reale di F. Siccome A è una matrice reale anche $\overline{\lambda}$ è un autovalore di f e

$$\dim V_{\overline{\lambda}}(f) = \mathrm{mult}_{\overline{\lambda}}(P_f) = \mathrm{mult}_\lambda(P_f) = \dim V_\lambda(f). \tag{9.4.29}$$

Inoltre se $u \in V_\lambda(F)$ allora $\overline{u} \in V_{\overline{\lambda}}(F)$. Per ogni coppia di autovalori complessi coniugati $\{\lambda, \overline{\lambda}\}$ scegliamo uno dei due autovalori, sia λ, e una base ON $\{t_1, \dots, t_d\}$ di $V_\lambda(F)$. Allora $\{\overline{t}_1, \dots, \overline{t}_d\}$ è una base ON di $V_{\overline{\lambda}}(F)$. Poniamo

$$u_s := \frac{1}{\sqrt{2}}(t_s + \overline{t}_s), \tag{9.4.30}$$

$$w_s := \frac{1}{\sqrt{2}}(it_s - i\overline{t}_s). \tag{9.4.31}$$

Un facile clacolo dà che $u_s, w_s \in \mathbb{R}^n$, che sono ortonormali e che

$$f(u_s) = \cos\theta_s u_s + \sin\theta_s w_s, \qquad f(w_s) = -\sin\theta_s u_s + \cos\theta_s w_s. \tag{9.4.32}$$

Quindi se raccogliamo tutti i vettori vettori u_s e w_s ottenuti in tal modo e aggiungiamo i vettori $v_1 \dots, v_a, z_1, \dots, z_b$ abbiamo in tutto n vettori (per l'equazione (9.4.27)) e perciò

$$v_1 \dots, v_a, z_1, \dots, z_b, u_1, w_1, u_2, w_2, \dots, u_c, w_c \tag{9.4.33}$$

è una base ON di V tale che valgono (9.4.24) e (9.4.25). □

Osservazione 9.4.15. Il contenuto della **Proposizione 9.4.14** nel caso dello spazio $\mathcal{V}^2$ dei vettori geometrici del piano è il seguente: se $f \in \mathrm{O}(\mathcal{V}^2)$ allora f è l'identità, o la riflessione in un sottospazio di dimensione 1 (il caso di due autovalori reali, uno 1 e l'altro -1) o una rotazione. Analogamente il contenuto della **Proposizione 9.4.14** nel caso dello spazio $\mathcal{V}^3$ dei vettori geometrici dello spazio è il seguente: se $f \in \mathrm{O}(\mathcal{V}^3)$ allora f è l'identità, o la riflessione in un sottospazio di dimensione 2 (il caso di autovalori reali, 1 con molteplicità 2 e -1 con molteplicità 1) o una rotazione o una rotazione seguita dalla riflessione nel piano della rotazione.

9.5 Teorema spettrale per operatori autoaggiunti

Siano V uno spazio vettoriale complesso e $\langle , \rangle$ una forma hermitiana definita positiva su V. Dimostreremo che certi endomorfismi di V sono diagonalizzabili. Cominciamo con i preliminari necessari per definire la classe di endomorfismi che considereremo. Un *operatore* di V non è altro che un endomorfismo $T\colon V \to V$. Ci prenderemo la libertà di denotare $T(v)$ con Tv quando non ciò non genererà equivoci.

Lemma 9.5.1. *Siano V uno spazio vettoriale complesso e $\langle , \rangle$ una forma hermitiana definita positiva su V. Sia $T\colon V \to V$ un operatore. Supponiamo che $A, B\colon V \to V$ siano operatori tali che*

$$\langle Tv, w\rangle = \langle v, Aw\rangle = \langle v, Bw\rangle \qquad \forall v, w \in V. \tag{9.5.1}$$

Allora $A = B$.

Dimostrazione. Dalla (9.5.3) segue che $\langle v, (A-B)w\rangle = 0$ per ogni $v, w \in V$. Ponendo $v = (A-B)w$ otteniamo che

$$0 = \langle (A-B)w, (A-B)w\rangle = ||(A-B)w||^2 \qquad \forall w \in V \tag{9.5.2}$$

e quindi $(A-B)w = 0$ per ogni $w \in V$. Perciò $Aw = Bw$ per ogni $w \in V$, ovvero $A = B$. □

Definizione 9.5.2. Siano V uno spazio vettoriale complesso e $\langle , \rangle$ una forma hermitiana definita positiva su V. Sia $T\colon V \to V$ un operatore. Se esiste un operatore $A\colon V \to V$ tale che

$$\langle Tv, w\rangle = \langle v, Aw\rangle \qquad \forall v, w \in V \tag{9.5.3}$$

(unico per il **Lemma 9.5.1**) diciamo che A è l'*aggiunto* di T e lo denotiamo T^*. Quindi si ha che

$$\langle Tv, w\rangle = \langle v, T^*w\rangle \qquad \forall v, w \in V. \tag{9.5.4}$$

Diciamo che T è *autoaggiunto* se $T = T^*$.

Proposizione 9.5.3. *Siano V uno spazio vettoriale complesso* finitamente generato *e $\langle , \rangle$ una forma hermitiana definita positiva su V. Sia $T\colon V \to V$ un operatore. Esiste l'aggiunto di T.*

Dimostrazione. Sia $\mathcal{B}$ una base ON di V e $A := M_{\mathcal{B}}^{\mathcal{B}}(T)$. Siano $v, w \in V$; per l'**Osservazione 9.4.10** abbiamo che

$$\langle Tv, w\rangle = (A \cdot X_{\mathcal{B}}(v))^t \overline{X}_{\mathcal{B}}(w) = X_{\mathcal{B}}(v)^t \cdot A^t \cdot \overline{X}_{\mathcal{B}}(w) = X_{\mathcal{B}}(v)^t \cdot \overline{\overline{A}^t} \cdot X_{\mathcal{B}}(w). \tag{9.5.5}$$

Quindi $T^*\colon V \to V$ è l'unico operatore tale che $M_{\mathcal{B}}^{\mathcal{B}}(T^*) = \overline{A}^t$ vale (9.5.4). □

Definizione 9.5.4. L'*aggiunta* di una matrice $A \in M_{n,n}(\mathbb{C})$ è la matrice $A^* \in M_{n,n}(\mathbb{C})$ tale che $L_{A^*} = L_A^*$, dove $\mathbb{C}^n$ è provvisto del prodotto hermitiano standard; esplicitamente $A^* = \overline{A}^t$. Diciamo che $A \in M_{n,n}(\mathbb{C})$ è *autoaggiunta* se $A = A^*$, cioè se $A = \overline{A}^t$.

Proposizione 9.5.5. *Siano V uno spazio vettoriale complesso finitamente generato e $\langle , \rangle$ una forma hermitiana definita positiva su V.*

1. *Sia $T\colon V \to V$ un operatore. Allora*

$$\begin{array}{ccc} V \times V & \xrightarrow{F} & \mathbb{C} \\ (v,w) & \mapsto & \langle Tv, w\rangle \end{array} \tag{9.5.6}$$

   è una forma hermitiana su V se e solo se T è autoaggiunto.

2. *Data una forma hermitiana H su V esiste un unico operatore autoaggiunto $T\colon V \to V$ tale che $H(v,w) = \langle Tv, w\rangle$ per ogni $v, w \in V$.*

Dimostrazione. (1): F è lineare a sinistra perché T è lineare (sia che T sia autoggiunto, sia che non lo sia). Rimane da dimostrare che $F(v,w) = \overline{F(w,v)}$ se e solo se T è autoaggiunto. Ora,

$$F(v,w) = \langle Tv, w\rangle, \qquad \overline{F(w,v)} = \overline{\langle Tw, v\rangle} = \overline{\langle w, T^*v\rangle} = \langle T^*v, w\rangle.$$

Segue il punto (1). Dimostriamo (2): Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base ON di V, e, in analogia con quanto fatto per forme bilineari simmetriche, poniamo $M_{\mathcal{B}}(H) = A$, dove

$$a_{ij} = H(v_i, v_j).$$

Allora $A^t = \overline{A}$, cioè A è autoaggiunta. Sia $T\colon V \to V$ l'operatore tale che $M_{\mathcal{B}}^{\mathcal{B}}(T) = A$; allora $H(v,w) = \langle Tv, w\rangle$ per ogni $v, w \in V$, e T è autoaggiunto perché A lo è. □

Teorema 9.5.6 (Teorema spettrale per operatori autoaggiunti)**.** *Siano V uno spazio vettoriale complesso finitamente generato, $\langle , \rangle$ una forma hermitiana definita positiva su V e $T\colon V \to V$ un operatore autoaggiunto (per $\langle , \rangle$). Gli autovalori di T sono reali ed esiste una base ortonormale di V che diagonalizza T.*

Dimostrazione. La dimostrazione è per induzione sulla dimensione di V. Se $\dim V = 1$ ogni endomorfismo di V è la moltiplicazione per uno scalare e l'affermazione è banalmente vera. Dimostriamo il passo induttivo; poniamo $n := \dim V > 1$. Il polinomio caratteristico P_T ha almeno una radice λ: dimostriamo che λ è reale. Sia v un autovettore con autovalore λ; siccome $T = T^*$ abbiamo che

$$\lambda ||v||^2 = \lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Tv, v\rangle = \langle v, Tv\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle = \overline{\lambda}||v||^2. \tag{9.5.7}$$

Ma $||v||^2 \neq 0$ perché $v \neq 0$ e quindi $\lambda = \overline{\lambda}$ cioè λ è reale. Abbiamo dimostrato che gli autovalori di T sono reali. Sia $v_n \in V$ un autovettore con autovalore associato λ_n e poniamo

$$W := v_n^{\perp} = \{w \in V \mid \langle w, v_n\rangle = 0\}. \tag{9.5.8}$$

Allora W è un sottospazio di dimensione $(n-1)$. Dimostriamo che $T(W) \subset W$. Sia $w \in W$: allora

$$\langle Tw, v_n\rangle = \langle w, Tv_n\rangle = \langle w, \lambda_n v_n\rangle = \overline{\lambda}_n\langle w, v_n\rangle = 0$$

e perciò $T(w) \in W$. La restrizione di $\langle , \rangle$ a W è un prodotto hermitiano definito positivo. Abbiamo dimostrato che $T(W) \subset W$ e quindi la restrizione di T a W è un operatore autoaggiunto di W (relativamente alla restrizione di $\langle , \rangle$ a W). Per l'ipotesi induttiva esiste una base ortonormale $\{v_1, \dots, v_{n-1}\}$ di W che diagonalizza la restrizione di T a W. La base $\{v_1, \dots, v_{n-1}, v_n\}$ di V è ortonormale e diagonalizza T. □

Corollario 9.5.7. *Siano V uno spazio vettoriale complesso finitamente generato, $\langle , \rangle$ una forma hermitiana definita positiva su V e H una forma hermitian su V. Esiset una base ON $\mathcal{B} = \{v_1, \dots, v_n\}$ di V che diagonalizza H, cioè tale che*

$$H\left(\sum_{i=1}^{n} x_i v_i, \sum_{j=1}^{n} y_j v_j\right) = \sum_{i=1}^{n} \lambda_i x_i \overline{y}_i.$$

Dimostrazione. Segue dal Teorema spettrale per operatori autoaggiunti, e dal punto (2) della **Proposizione 9.5.5**. □

9.6 Il Teorema spettrale per operatori simmetrici

Come abbiamo visto un endomorfismo di uno spazio vettoriale (finitamente generato) non è necessariamente diagonalizzabile. In questa sezione dimostreremo che se un endomorfismo di uno spazio vettoriale euclideo gode di una certa proprietà allora è diagonalizzabile. Faremo vedere che questo risultato equivale ad affermare che, data una forma quadratica su uno spazio vettoriale euclideo, esiste una base *ortonormale* che diagonalizza la forma quadratica.

9.6.1 Il Teorema spettrale

Definizione 9.6.1. Sia V uno spazio vettoriale euclideo, con prodotto scalare $\langle , \rangle$. Un operatore $T\colon V \to V$ è *simmetrico*[1] se per ogni $v, w \in V$

$$\langle Tv, w\rangle = \langle v, Tw\rangle. \tag{9.6.1}$$

Osservazione 9.6.2. Supponiamo che V sia finitamente generato. Sia $\mathcal{B}$ una base ortonormale di V e quindi $\langle v, w\rangle = X_{\mathcal{B}}(v)^t \cdot X_{\mathcal{B}}(w)$ per ogni $v, w \in V$. Allora $T \in \mathrm{End}(V)$ è simmetrico se e solo se $A := M_{\mathcal{B}}^{\mathcal{B}}(T)$ è una matrice simmetrica. Infatti $T \in \mathrm{End}(V)$ è simmetrico se e solo se per ogni $v, w \in V$

$$(A \cdot X_{\mathcal{B}}(v))^t \cdot X_{\mathcal{B}}(w) = \langle Tv, w\rangle = \langle v, Tw\rangle = X_{\mathcal{B}}(v)^t \cdot (A \cdot X_{\mathcal{B}}(w)) \tag{9.6.2}$$

e questo vale se e solo se $A^t = A$.

Teorema 9.6.3 (Teorema spettrale per operatori simmetrici)**.** *Siano $(V, \langle , \rangle)$ uno spazio vettoriale euclideo finitamente generato e $T\colon V \to V$ un operatore simmetrico. Esiste una base ortonormale che diagonalizza T.*

Dimostrazione. Per induzione sulla dimensione di V. Se $\dim V = 1$ ogni endomorfismo di V è la moltiplicazione per uno scalare e l'affermazione è banalmente vera. Dimostriamo il passo induttivo; poniamo $n := \dim V > 1$. Il polinomio caratteristico P_T ha almeno una radice complessa λ: dimostriamo che λ è reale. Sia $\mathcal{B}$ una base ortonormale di V: allora $A := M_{\mathcal{B}}^{\mathcal{B}}$ è una matrice simmetrica e λ è un suo autovalore quando A è vista come matrice *complessa* (perché radice del suo polinomio caratteristico), cioè esiste $\mathbf{0} \neq X \in \mathbb{C}^n$ tale che $A \cdot X = \lambda X$. Se $M \in M_{r,s}(\mathbb{C})$ denotiamo con $\overline{M}$ la matrice con entrate i coniugati delle entrate di M. Siccome $A^t = A$ e $\overline{A} = A$ abbiamo che

$$\lambda X^t \cdot \overline{X} = (A \cdot X)^t \cdot \overline{X} = X^t \cdot (A \cdot \overline{X}) = X^t \cdot \overline{(A \cdot X)} = X^t \cdot \overline{\lambda X} = \overline{\lambda} X^t \cdot \overline{X}. \tag{9.6.3}$$

Ora osserviamo che $X^t \cdot \overline{X} = \sum_{i=1}^{n} |x_i|^2$ è un numero reale strettamente positivo perché $X \neq \mathbf{0}$. Dalla (9.6.8) segue che $\lambda = \overline{\lambda}$ e quindi λ è reale. Siccome λ è un autovalore (reale) di T esiste una autovettore $v_n \in V$ con autovalore associato λ. Sia

$$W := v_n^{\perp} = \{w \in V \mid \langle w, v_n\rangle = 0\}. \tag{9.6.4}$$

[1] Il termine simmetrico è giustificato dalla **Proposizione 9.6.5**.

Siccome $\langle , \rangle$ è non-degenere W è un sottospazio di dimensione $(n-1)$. Dimostriamo che $T(W) \subset W$. Sia $w \in W$: allora

$$\langle Tw, v_n \rangle = \langle w, Tv_n \rangle = \langle w, \lambda v_n \rangle = \lambda \langle w, v_n \rangle = 0$$

e perciò $T(w) \in W$. La restrizione di $\langle , \rangle$ a W è un prodotto scalare definito positivo e quindi dà a W una struttura di spazio vettoriale euclideo di dimensione $(n-1)$. Abbiamo dimostrato che $T(W) \subset W$ e quindi la restrizione di T a W è un operatore simmetrico su W. Per l'ipotesi induttiva esiste una base ortonormale $\{v_1, \ldots, v_{n-1}\}$ di W che diagonalizza la restrizione di T a W. La base $\{v_1, \ldots, v_{n-1}, v_n\}$ di V è ortonormale e diagonalizza T. □

Osservazione 9.6.4.

1. Il Teorema spettrale dà che se $A \in M_{n,n}(\mathbb{R})^+$ è una matrice simmetrica il suo polinomio caratterstico ha tutte le radici reali (cioè è prodotto di polinomi reali di grado 1).
2. Lo *spettro* di un operatore è l'insieme dei suoi autovalori. Il contenuto essenziale del Teorema spettrale è che gli autovalori di un operatore simmetrico sono tutti reali, ovvero il suo spettro è reale - questo spiega il nome del teorema.

9.6.2 Forme quadratiche su spazi vettoriali euclidei

Sia V uno spazio vettoriale euclideo, con prodotto scalare $\langle , \rangle$, e $T\colon V \to V$ un operatore. L'applicazione

$$\begin{array}{ccc} V \times V & \longrightarrow & \mathbb{R} \\ (v, w) & \mapsto & \langle Tv, w \rangle \end{array} \tag{9.6.5}$$

è bilineare.

Proposizione 9.6.5. *L'applicazione bilineare* (9.6.5) *è simmetrica se e solo se T è un operatore simmetrico.*

Dimostrazione. L'applicazione bilineare (9.6.5) è simmetrica se e solo se

$$\langle Tv, w \rangle = \langle Tw, v \rangle \tag{9.6.6}$$

per ogni $v, w \in V$. La simmetria di $\langle , \rangle$ dà che $\langle Tw, v \rangle = \langle v, Tw \rangle$ e perciò (9.6.6) equivale a $\langle Tv, w \rangle = \langle v, Tw \rangle$. □

Proposizione 9.6.6. *Sia V uno spazio vettoriale euclideo finitamente generato, con prodotto scalare $\langle , \rangle$. Sia $F \in \mathrm{Bil}_+(V)$ una forma bilineare simmetrica. Esiste uno e un solo operatore simmetrico $T\colon V \to V$ tale che*

$$F(v, w) = \langle Tv, w \rangle \qquad \forall v, w \in V. \tag{9.6.7}$$

Dimostrazione. Sia $\mathcal{B}$ una base ortonormale di V. La matrice $M_{\mathcal{B}}(F)$ è simmetrica. Sia $T\colon V \to V$ l'operatore tale che $M_{\mathcal{B}}^{\mathcal{B}}(T) = M_{\mathcal{B}}(F)$. Allora T è simmetrico per l'**Osservazione 9.6.2**. Si ha che

$$F(v, w) = X_{\mathcal{B}}(v)^t \cdot M_{\mathcal{B}}(F) \cdot X_{\mathcal{B}}(w) = ((M_{\mathcal{B}}^{\mathcal{B}}(T) \cdot X_{\mathcal{B}}(v))^t \cdot X_{\mathcal{B}}(w) = \langle Tv, w \rangle. \tag{9.6.8}$$

Questo conto dimostra anche che se vale (9.6.7) allora $M_{\mathcal{B}}^{\mathcal{B}}(T) = M_{\mathcal{B}}(F)$ e quindi T è unico. □

Definizione 9.6.7. Siano V, $\langle , \rangle$, $F \in \mathrm{Bil}_+(V)$ e $T\colon V \to V$ come nell'enuciato della **Proposizione 9.6.6**: gli *autovalori* di F sono gli autovalori di T.

Teorema 9.6.8. *Siano $(V, \langle , \rangle)$ uno spazio vettoriale euclideo finitamente generato e q una forma quadratica su V. Esiste una base ortonormale che diagonalizza q.*

Dimostrazione. Sia F la polarizzazione di q: dunque F è una forma bilineare simmetrica e basta trovare una base ON che diagonalizza F. Per la **Proposizione 9.6.6** esiste un operatore simmetrico $T\colon V \to V$ tale che valga (9.6.7). Per il teorema spettrale esiste una base ON $\{v_1, \ldots, v_n\}$ che diagonalizza T: quindi esistono $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tali che $Tv_i = \lambda_i v_i$. Si ha che

$$F(v_i, v_j) = \langle Tv_i, v_j \rangle = \lambda_i \langle v_i, v_j \rangle = \lambda_i \delta_{ij} \tag{9.6.9}$$

e perciò la base ON $\{v_1, \ldots, v_n\}$ diagonalizza F. □

Osservazione 9.6.9. Sia f una forma quadratica su $\mathbb{R}^n$, quindi esiste $A \in M^+_{n,n}(\mathbb{R})$ tale che $f(X) = X^t \cdot A \cdot X$. Consideriamo il prodotto scalare standard $\langle , \rangle$ su $\mathbb{R}^n$, cioè $\langle X, Y\rangle = X^t \cdot Y$. Gli autovalori di f sono gli autovalori di A perché l'operatore simmetrico T tale che valga (9.6.7) è L_A. Sia $X = G \cdot Z$ un cambiamento di coordinate, quindi $G \in \mathrm{GL}_n(\mathbb{R})$. La matrice di f nel nuovo sistema di coordinate è $B := G^t \cdot A \cdot G$. Se $G \in \mathrm{O}_n(\mathbb{R})$ allora $G^t = G^{-1}$, quindi B è coniugata di A e perciò i suoi autovalori sono gli stessi di A - questo è in accordo col fatto che la **Definizione 9.6.7** è sensata. Se G è arbitraria non c'è alcun motivo per cui gli autovalori di B debbano esser gli stessi di A, però il **Teorema 9.6.8** e la **Proposizione 7.6.15** danno che il numero di autovalori positivi di A e di B sono eguali, e così per il numero di autovalori negativi e nulli.

9.7 La forma canonica di Jordan

Esercizi del Capitolo 9

Esercizio 9.1. *Siano $A, B \in M_{n,n}(k)$.*

1. *Dimostrate che se A, B commutano (cioè $A \cdot B = B \cdot A$) allora $e^{A+B} = e^A \cdot e^B$.*
2. *Date esempi di A, B tali che $e^{A+B} \neq e^A \cdot e^B$.*

Esercizio 9.2. *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 9.3. *Sia*

$$A := \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}$$

Esprimete e^{tA} come matrice con entrate combinazioni lineari di esponenziali.

Esercizio 9.4. *Sia $A \in M_{3,3}(\mathbb{R})$ data da*

$$A := \begin{bmatrix} 5 & -3 & -1 \\ -3 & 2 & 2 \\ -1 & 2 & 0 \end{bmatrix}$$

Trovate una base di $\mathbb{R}^3$, ortonormale *per il prodotto euclideo standard, che diagonalizza A.*

Esercizio 9.5. *Sia $A \in M_{3,3}(\mathbb{R})$ data da*

$$A := \begin{bmatrix} 0 & 3 & -6 \\ 3 & 1 & 4 \\ -6 & 4 & -5 \end{bmatrix}$$

Determinate una base ortogonale[2] *che diagonalizza A.*

Esercizio 9.6. *Siano $H_1, H_2, H_3 \colon \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ definite da*

$$\begin{array}{ccc} \mathbb{C}^2 \times \mathbb{C}^2 & \xrightarrow{H_1} & \mathbb{C} \\ (W, Z) & \mapsto & w_1\overline{z}_2 + 3w_2\overline{z}_1 \end{array} \tag{9.7.1}$$

[2]Non necessariamente ortonormale.

$$\begin{array}{ccc} \mathbb{C}^2 \times \mathbb{C}^2 & \xrightarrow{H_2} & \mathbb{C} \\ (W, Z) & \mapsto & w_1\overline{z}_1 + (1+i)w_2\overline{z}_1 + (1-i)w_1\overline{z}_2 \end{array} \tag{9.7.2}$$

$$\begin{array}{ccc} \mathbb{C}^2 \times \mathbb{C}^2 & \xrightarrow{H_3} & \mathbb{C} \\ (W, Z) & \mapsto & w_1\overline{z}_1 + w_2\overline{z}_2 \end{array} \tag{9.7.3}$$

Quali delle H_i sono forme hermitiane ?

Esercizio 9.7. *Per $a \in \mathbb{R}$ sia $H_a \colon \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ la forma hermitiana definita da*

$$\begin{array}{ccc} \mathbb{C}^2 \times \mathbb{C}^2 & \xrightarrow{H_a} & \mathbb{C} \\ (W, Z) & \mapsto & w_1\overline{z}_1 + w_1\overline{z}_1 + aiw_2\overline{z}_1 - aiw_1\overline{z}_2 \end{array} \tag{9.7.4}$$

Determinate per quali valori di a la H_a è definita positiva.

Bibliografia

[1] M. Artin. *Algebra*, Prentice Hall, 1991.

[2] S. Lang. *Algebra*, Springer GTM 211, 2002.