

Geometria e algebra lineare (II parte)

Bruno Martelli

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

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

versione: 5 aprile 2018

Indice

Introduzione	1
Capitolo 1. Autovettori e autovalori	3
1.1. Introduzione	3
1.2. Autovettori, autovalori e polinomio caratteristico	4
1.3. Teorema di diagonalizzabilità	10
Capitolo 2. Forma di Jordan	17
2.1. Forma di Jordan	17
2.2. Teorema di Cayley – Hamilton	24
2.3. Polinomio minimo	27
Capitolo 3. Prodotti scalari	33
3.1. Introduzione	33
3.2. Classificazione dei prodotti scalari	44
3.3. Prodotti scalari definiti positivi	51
3.4. Isometrie	62
Capitolo 4. Lo spazio euclideo	73
4.1. Prodotto vettoriale	73
4.2. Sottospazi affini	77
4.3. Trasformazioni del piano e dello spazio	87
Capitolo 5. Teorema spettrale	95
5.1. Prodotti hermitiani	95
5.2. Endomorfismi autoaggiunti	96
5.3. Il teorema	98
Capitolo 6. Geometria proiettiva	101
6.1. Introduzione	101
6.2. Sottospazi e proiettività	104
Capitolo 7. Quadriche	109
7.1. Introduzione	109
7.2. Classificazione proiettiva	110
7.3. Classificazione affine	114
7.4. Classificazione metrica delle quadriche non degeneri	118

Introduzione

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

- Figura 4.1 (regola della mano destra), creata da Acdx;
- Figure 7.1 e 7.2 (quadrice), create da Sam Derbyshire.

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

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

Autovettori e autovalori

1.1. Introduzione

Sia $T: V \rightarrow V$ un endomorfismo di uno spazio vettoriale V di dimensione n . Fissata una base $\mathcal{B} = \{v_1, \dots, v_n\}$ di V , possiamo descrivere T concretamente come una matrice quadrata $A = [T]_{\mathcal{B}}^{\mathcal{B}}$. In questo capitolo studiamo questo problema: esiste una base \mathcal{B} migliore delle altre, che produca una matrice A particolarmente semplice?

Ad esempio, se A fosse una matrice diagonale

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

otterremmo una descrizione soddisfacente di T . Se A fosse diagonale, il prodotto riga per colonna si semplificherebbe molto:

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda_1 x_1 \\ \lambda_2 x_2 \\ \vdots \\ \lambda_n x_n \end{pmatrix}.$$

L'osservazione geometrica fondamentale è che spiegheremo fra poche righe è la seguente: se A è diagonale, allora la trasformazione T manda ogni vettore v_i della base \mathcal{B} in un multiplo $\lambda_i v_i$ di se stesso. La retta $r_i = \text{Span}(v_i)$ è *invariante*, cioè

$$T(r_i) \subset r_i.$$

Le rette r_1, \dots, r_n formano quindi degli "assi" per T . L'endomorfismo T agisce su ciascuna retta r_i ampliandola (se $\lambda_i > 1$), lasciando fissi tutti i punti (se $\lambda_i = 1$), contraendola (se $0 < \lambda_i < 1$), collassandola tutta nell'origine (se $\lambda_i = 0$), o ribaltandola (se $\lambda_i < 0$). L'endomorfismo T può agire in modi diversi su assi diversi.

Ci sono altri motivi per preferire le matrici diagonali come la A descritta sopra. Intanto il determinante di A si calcola semplicemente come

$$\det A = \lambda_1 \cdots \lambda_n.$$

Inoltre il prodotto di due matrici diagonali è semplicissimo da calcolare:

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \begin{pmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \mu_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mu_n \end{pmatrix} = \begin{pmatrix} \lambda_1\mu_1 & 0 & \dots & 0 \\ 0 & \lambda_2\mu_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n\mu_n \end{pmatrix}.$$

Possiamo ad esempio calcolare facilmente potenze come A^{100} che indicano geometricamente che la trasformazione T viene iterata 100 volte.

Diciamo che un endomorfismo T è *diagonalizzabile* se esiste una base \mathcal{B} per cui $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ sia diagonale. Lo scopo di questo capitolo è rispondere alle seguenti domande: come facciamo a capire se T è diagonalizzabile? Come possiamo trovare una base \mathcal{B} che diagonalizzi T , se esiste?

Per rispondere dovremo introdurre un po' di definizioni.

1.2. Autovettori, autovalori e polinomio caratteristico

1.2.1. Autovettori e autovalori. Sia come sempre $T: V \rightarrow V$ un endomorfismo di uno spazio vettoriale V definito su un campo \mathbb{K} . Un *autovettore* è un vettore $v \neq 0$ in V per cui

$$f(v) = \lambda v$$

per qualche numero $\lambda \in \mathbb{K}$ che chiameremo *autovalore*. Notiamo che λ può essere qualsiasi scalare, anche zero. Introduciamo adesso una definizione importante.

Definizione 1.2.1. Un endomorfismo $T: V \rightarrow V$ è *diagonalizzabile* se V ha una base v_1, \dots, v_n composta da autovettori per T .

Questa definizione è effettivamente equivalente a quella data nell'introduzione, in virtù del fatto seguente. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V .

Proposizione 1.2.2. La matrice $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ è diagonale se e solo se i vettori v_1, \dots, v_n sono tutti autovettori.

Dimostrazione. Il vettore v_i è un autovettore $\iff f(v_i) = \lambda_i v_i$. Le coordinate di $\lambda_i v_i$ rispetto a \mathcal{B} sono il vettore colonna

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

Quindi v_i è un autovettore se e solo se la i -esima colonna di A è un vettore di questo tipo. La matrice A è diagonale precisamente quando questo accade per ogni i . \square

Abbiamo quindi scoperto che un endomorfismo T è diagonalizzabile se e solo se esiste una base \mathcal{B} tale che $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ sia una matrice diagonale, come accennato nell'introduzione a questo capitolo.

Esempio 1.2.3. L'endomorfismo $L_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ con

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

è diagonalizzabile: una base di autovettori è data da

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

Infatti otteniamo $Av_1 = 2v_1$ e $Av_2 = -2v_2$, quindi v_1 e v_2 sono autovettori, con autovalori rispettivamente 2 e -2 . Prendendo $\mathcal{B} = \{v_1, v_2\}$ otteniamo

$$[L_A]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Esempio 1.2.4. Sia $M(2)$ lo spazio vettoriale delle matrici quadrate reali 2×2 . L'endomorfismo $T: M(2) \rightarrow M(2)$ definito da $T(A) = {}^tA$ è diagonalizzabile: una base di autovettori è data da

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Infatti otteniamo

$$T(A_1) = A_1, \quad T(A_2) = A_2, \quad T(A_3) = A_3, \quad T(A_4) = -A_4.$$

Prendendo $\mathcal{B} = \{A_1, A_2, A_3, A_4\}$ otteniamo

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Negli esempi precedenti, gli autovettori sono già dati e dobbiamo solo verificare che funzionino. In generale, come troviamo gli autovettori di un endomorfismo T ? Introduciamo uno strumento che sarà cruciale per rispondere a questa domanda.

1.2.2. Polinomio caratteristico. Il *polinomio caratteristico* di una matrice quadrata A è definito nel modo seguente:

$$p_A(\lambda) = \det(A - \lambda I).$$

Un commento sulla notazione: normalmente un polinomio viene indicato con il simbolo $p(x)$, in cui x rappresenta la variabile; qui a volte usiamo λ invece di x , per un motivo che sarà chiaro in seguito. Il pedice A in p_A indica semplicemente che il polinomio dipende dalla matrice A .

Da questa definizione un po' oscura di $p_A(\lambda)$ non risulta affatto chiaro che questo oggetto sia realmente un polinomio. Dimostriamo adesso che è un polinomio e identifichiamo alcuni dei suoi coefficienti. Sia n come di consueto la dimensione di V e quindi la taglia di A .

Proposizione 1.2.5. *Il polinomio caratteristico $p_A(\lambda)$ è effettivamente un polinomio. Ha grado n e quindi è della forma*

$$p_A(\lambda) = a_n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_0.$$

Possiamo identificare alcuni dei suoi coefficienti:

- $a_n = (-1)^n$,
- $a_{n-1} = (-1)^{n-1} \text{tr} A$,
- $a_0 = \det A$.

Dimostrazione. Dimostreremo la proposizione per induzione su n . Il caso $n = 1$ è ovviamente molto facile: la matrice è $A = (a)$ e scriviamo

$$p_A(\lambda) = \det(A - \lambda I) = \det(a - \lambda) = a - \lambda = -\lambda + a.$$

Otteniamo un polinomio di grado 1 con i coefficienti giusti (qui $\det A = \text{tr} A = a$). Anche se non è necessario per il passaggio induttivo, è comunque istruttivo esaminare anche il caso $n = 2$, in cui $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ e quindi

$$\begin{aligned} p_A(\lambda) &= \det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} = (a - \lambda)(d - \lambda) - bc \\ &= \lambda^2 - (a + d)\lambda + ad - bc = \lambda^2 - \text{tr} A \lambda + \det A. \end{aligned}$$

Dimostriamo adesso il passaggio induttivo: supponiamo la proposizione vera per $n - 1$ e la verifichiamo per n . Scriviamo A come (a_{ij}) e calcoliamo

$$p_A(\lambda) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix}.$$

Con lo sviluppo di Laplace lungo la prima colonna otteniamo

$$p_A(\lambda) = (a_{11} - \lambda) \det \begin{pmatrix} a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix} + q(\lambda).$$

Il termine $q(\lambda)$ è ottenuto continuando lo sviluppo di Laplace: il termine j -esimo dello sviluppo è calcolato cancellando la prima colonna e la j -esima riga di A ; quando $j \geq 2$ cancelliamo due caselle diverse contenenti un λ , quindi le caselle rimanenti che contengono λ sono $n - 2$. Iterando lo sviluppo si deduce facilmente che $q(\lambda)$ è un polinomio di grado $\leq n - 2$.

Possiamo adesso usare l'ipotesi induttiva sulla matrice di taglia $n - 1$ presente nel primo termine dello sviluppo di Laplace e dedurre che

$$\begin{aligned} p_A(t) &= (a_{11} - \lambda)((-1)^{n-1}\lambda^{n-1} + (-1)^{n-2}(a_{22} + \dots + a_{nn})\lambda^{n-2} + \dots) + \dots \\ &= (-1)^n \lambda^n + (-1)^{n-1}(a_{11} + a_{22} + \dots + a_{nn})\lambda^{n-1} + \dots \end{aligned}$$

Resta solo da dimostrare che il termine noto a_0 è il determinante di A . Per fare ciò ricordiamo che $a_0 = p_A(0)$ e scriviamo semplicemente

$$a_0 = p_A(0) = \det(A - 0I) = \det A.$$

La dimostrazione è completa. \square

Il polinomio caratteristico ci sarà utile per il motivo seguente. Gli autovalori ed autovettori di una matrice quadrata A sono per definizione quelli dell'endomorfismo $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^n$.

Proposizione 1.2.6. Gli autovalori di A sono precisamente le radici del polinomio caratteristico $p_A(\lambda)$.

Dimostrazione. Uno scalare $\lambda \in \mathbb{K}$ è autovalore per A se e solo se esiste un vettore $v \in \mathbb{K}^n$ non nullo tale che $Av = \lambda v$. I fatti seguenti sono tutti equivalenti:

$$\begin{aligned} \exists v \neq 0 \text{ tale che } Av &= \lambda v \iff \\ \exists v \neq 0 \text{ tale che } Av - \lambda v &= 0 \iff \\ \exists v \neq 0 \text{ tale che } (A - \lambda I)v &= 0 \iff \\ \exists v \neq 0 \text{ tale che } v \in \ker(A - \lambda I) &\iff \\ \det(A - \lambda I) = 0 &\iff p_A(\lambda) = 0. \end{aligned}$$

La dimostrazione è completa. \square

Abbiamo scoperto come possiamo identificare gli autovalori di A : dobbiamo solo (si fa per dire...) trovare le radici del polinomio caratteristico $p_A(\lambda)$. Questo ovviamente non è un problema semplice in generale: purtroppo non esiste una formula per trovare le radici di un polinomio di grado qualsiasi.

Esempio 1.2.7. Consideriamo la matrice

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

Il suo polinomio caratteristico è

$$p_A(\lambda) = \lambda^2 - \operatorname{tr}A\lambda + \det A = \lambda^2 - 4\lambda + 3.$$

Le sue radici sono $\lambda = 1$ e $\lambda = 3$. Quindi gli autovalori di A sono 1 e 3. Per ciascun autovalore $\lambda = 1$ e $\lambda = 3$, i suoi autovettori sono precisamente le soluzioni non banali dei sistemi rispettivamente $Av = v$ e $Av = 3v$. Risolvendo i sistemi, troviamo ad esempio due autovettori

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Effettivamente $Av_1 = v_1$ e $Av_2 = 3v_2$. I due vettori sono indipendenti e quindi formano una base di \mathbb{R}^2 . La matrice A è quindi diagonalizzabile, ed è simile alla matrice $D = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$.

Esempio 1.2.8. Consideriamo una rotazione $\text{Rot}_\theta: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ antioraria di un angolo $\theta \neq 0, \pi$. La rotazione è descritta dalla matrice

$$\text{Rot}_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Consideriamo il caso $\theta \neq 0, \pi$. In questo caso, è chiaro geometricamente che Rot_θ non ha autovettori: ciascun vettore non banale v è ruotato di un angolo $\theta \neq 0, \pi$ e quindi non può diventare un multiplo λv di se stesso.

Possiamo verificare che non ci sono autovettori anche guardando il polinomio caratteristico $p(\lambda) = \lambda^2 - 2\cos\theta + 1$ e notando che ha due radici complesse non reali, perché $\Delta = \cos^2\theta - 1 < 0$. Il polinomio caratteristico non ha radici reali: quindi non ci sono autovalori, e di conseguenza neppure autovettori.

Il prossimo esempio mostra come la diagonalizzabilità di una matrice A dipenda dal campo \mathbb{K} che stiamo considerando!

Esempio 1.2.9. La matrice

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

rappresenta una rotazione di $\frac{\pi}{2}$ e quindi sappiamo già che non è diagonalizzabile su \mathbb{R} . Se però la consideriamo come matrice su \mathbb{C} , le cose cambiano: il polinomio caratteristico è infatti $p(\lambda) = \lambda^2 + 1$ ed ha soluzioni $\lambda = \pm i$. Studiando i sistemi $Av = iv$ e $Av = -iv$ troviamo che i vettori seguenti

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

sono autovettori con autovalori i e $-i$ rispettivamente. Quindi la matrice A non è diagonalizzabile su \mathbb{R} , ma lo è su \mathbb{C} . Prendendo $\mathcal{B} = \{v_1, v_2\}$ otteniamo

$$[L_A]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Esempio 1.2.10. La matrice

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

ha polinomio caratteristico $p(\lambda) = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2$. Il polinomio ha una radice $\lambda = 1$ con molteplicità due. Quindi $\lambda = 1$ è l'unico autovalore. Risolvendo il sistema $Av = v$ si trova facilmente che gli autovettori sono i multipli di $v = (1, 0)$. La matrice A non è diagonalizzabile perché questi vettori non sono sufficienti per generare \mathbb{R}^2 . Troviamo un autovettore, ma non ne troviamo due indipendenti. Lo stesso ragionamento mostra che questa matrice non è diagonalizzabile neppure su \mathbb{C} .

Esercizio 1.2.11. Per ciascuna delle matrici seguenti, determinare autovalori e autovettori in \mathbb{R}^3 e dire se è diagonalizzabile:

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

Se $T: V \rightarrow V$ è un endomorfismo, possiamo comunque definire il polinomio caratteristico $p_T(\lambda)$ come il polinomio caratteristico $p_A(\lambda)$ di una qualsiasi matrice associata $A = [T]_{\mathcal{B}}$. Questa definizione ha senso, perché due matrici associate diverse sono sempre simili, e la proposizione seguente ci dice che due matrici simili hanno lo stesso polinomio caratteristico:

Proposizione 1.2.12. *Se A e B sono matrici simili, allora $p_A(\lambda) = p_B(\lambda)$.*

Dimostrazione. Per ipotesi $A = M^{-1}BM$ per qualche matrice invertibile M . Otteniamo quindi

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I) = \det(M^{-1}BM - M^{-1}\lambda I M) \\ &= \det(M^{-1}(B - \lambda I)M) = \det(M^{-1}) \det(B - \lambda I) \det M \\ &= \det(B - \lambda I) = p_B(\lambda) \end{aligned}$$

grazie al teorema di Binet. □

1.2.3. Matrici triangolari. Ricordiamo che una matrice quadrata A è *triangolare* se è della forma

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

In altre parole, vale $a_{ij} = 0$ per ogni $i > j$. Il determinante di una matrice triangolare A è particolarmente semplice da calcolare: usando lo sviluppo di Laplace si verifica che

$$\det A = a_{11} \cdots a_{nn}.$$

Anche gli autovalori si determinano immediatamente:

Proposizione 1.2.13. *Gli autovalori di una matrice triangolare A sono gli elementi a_{11}, \dots, a_{nn} sulla diagonale.*

Dimostrazione. Vale

$$\begin{aligned} p_A(\lambda) = \det(A - \lambda I) &= \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ 0 & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} - \lambda \end{pmatrix} \\ &= (a_{11} - \lambda) \cdots (a_{nn} - \lambda). \end{aligned}$$

Le radici di p_A sono quindi a_{11}, \dots, a_{nn} . □

Notiamo che gli elementi a_{11}, \dots, a_{nn} possono anche ripetersi.

1.2.4. Matrici a blocchi. Ricordiamo il seguente esercizio. Indichiamo con $M(n)$ lo spazio delle matrici quadrate $n \times n$, sul campo base \mathbb{K} .

Esercizio 1.2.14. Sia $A \in M(n)$ del tipo

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

con $B \in M(k)$ e $D \in M(n - k)$ per qualche k . Valgono le relazioni

$$\begin{aligned} \det A &= \det B \cdot \det D, \\ \text{rk} A &\geq \text{rk} B + \text{rk} D. \end{aligned}$$

Ne ricaviamo una relazione analoga per il polinomio caratteristico.

Corollario 1.2.15. *Sia $A \in M(n)$ del tipo*

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

con $B \in M(k)$ e $D \in M(n - k)$ per qualche k . Vale la relazione

$$p_A(\lambda) = p_B(\lambda) \cdot p_D(\lambda).$$

Dimostrazione. Scriviamo

$$\begin{aligned} p_A(\lambda) = \det(A - \lambda I) &= \det \begin{pmatrix} B - \lambda I & C \\ 0 & D - \lambda I \end{pmatrix} \\ &= \det(B - \lambda I) \cdot \det(D - \lambda I) = p_B(\lambda) \cdot p_D(\lambda). \end{aligned}$$

La dimostrazione è completa. □

1.3. Teorema di diagonalizzabilità

Lo scopo di questa sezione è la definizione di un criterio di diagonalizzabilità per endomorfismi e matrici quadrate. Iniziamo mostrando una proprietà algebrica degli autovettori.

1.3.1. Autovettori con autovalori distinti. Sia $T: V \rightarrow V$ un endomorfismo qualsiasi.

Proposizione 1.3.1. *Se $v_1, \dots, v_k \in V$ sono autovettori per T con autovalori $\lambda_1, \dots, \lambda_k$ distinti, allora sono indipendenti.*

Dimostrazione. Lo dimostriamo per induzione su k . Se $k = 1$, il vettore v_1 è indipendente semplicemente perché non è nullo (per definizione, un autovettore non è mai nullo).

Diamo per buono il caso $k - 1$ e mostriamo il caso k . Supponiamo di avere una combinazione lineare nulla

$$(1) \quad \alpha_1 v_1 + \dots + \alpha_k v_k = 0.$$

Dobbiamo dimostrare che $\alpha_i = 0$ per ogni i . Applicando T otteniamo

$$(2) \quad 0 = \alpha_1 T(v_1) + \dots + \alpha_k T(v_k) = \alpha_1 \lambda_1 v_1 + \dots + \alpha_k \lambda_k v_k.$$

Moltiplicando l'equazione (1) per λ_k ricaviamo

$$\alpha_1 \lambda_k v_1 + \dots + \alpha_k \lambda_k v_k = 0$$

e sottraendo questa equazione dalla (2) deduciamo che

$$\alpha_1(\lambda_k - \lambda_1)v_1 + \dots + \alpha_{k-1}(\lambda_k - \lambda_{k-1})v_{k-1} = 0.$$

Questa è una combinazione lineare nulla di $k - 1$ autovettori con autovalori distinti: per l'ipotesi induttiva tutti i coefficienti $\alpha_i(\lambda_k - \lambda_i)$ devono essere nulli. Poiché $\lambda_i \neq \lambda_k$, ne deduciamo che $\alpha_i = 0$ per ogni $i = 1, \dots, k - 1$ e usando (1) otteniamo anche $\alpha_k = 0$. La dimostrazione è completa. \square

Corollario 1.3.2. *Se $p_T(\lambda)$ ha n radici reali distinte, l'endomorfismo T è diagonalizzabile.*

Dimostrazione. Le radici sono $\lambda_1, \dots, \lambda_n$ e ciascun λ_i è autovalore di qualche autovettore v_i . Gli autovettori v_1, \dots, v_n sono indipendenti per la Proposizione 1.3.1, quindi sono una base di V . Quindi V ha una base formata da autovettori, cioè è diagonalizzabile. \square

Ad esempio, consideriamo una matrice triangolare A . Gli autovalori di A sono gli elementi a_{11}, \dots, a_{nn} sulla diagonale. Abbiamo visto nell'Esempio 1.2.10 che una matrice diagonale può non essere diagonalizzabile in generale: se però gli autovalori a_{11}, \dots, a_{nn} sono tutti distinti, allora A è diagonalizzabile per il Corollario 1.3.2.

Abbiamo ottenuto un criterio di diagonalizzabilità molto economico: possiamo dimostrare in alcuni casi che una matrice è diagonalizzabile senza cercare una esplicita base di autovettori. Ad esempio, la matrice

$$A = \begin{pmatrix} 1 & 7 & -1 \\ 0 & 2 & 8 \\ 0 & 0 & 3 \end{pmatrix}$$

è diagonalizzabile perché ha tre autovalori distinti 1, 2, 3. Notiamo che trovare una base esplicita di autovettori non è immediato: per ciascun $\lambda = 1, 2, 3$ dobbiamo risolvere l'equazione $Av = \lambda v$ (il lettore è invitato a farlo).

Il teorema di diagonalizzabilità che vogliamo dimostrare sarà una generalizzazione di questo criterio. Per enunciare questo teorema dobbiamo introdurre ancora una nozione.

1.3.2. Autospatio. Sia $T: V \rightarrow V$ un endomorfismo. Per ogni autovalore λ di T definiamo l'*autospatio* $V_\lambda \subset V$ nel modo seguente:

$$V_\lambda = \{v \in V \mid T(v) = \lambda v\}.$$

In altre parole, V_λ consiste precisamente di tutti gli autovettori v con autovalore λ , più l'origine $0 \in V$ (ricordiamo che $0 \in V$ non è mai autovettore per definizione).

Ricordiamo l'endomorfismo identità $\text{id}: V \rightarrow V$, definito semplicemente da $\text{id}(v) = v$. Per ogni autovalore λ di T , ha senso considerare l'endomorfismo $T - \lambda \text{id}$ e la proposizione seguente lo mette in relazione con V_λ .

Proposizione 1.3.3. *Vale*

$$V_\lambda = \ker(T - \lambda \text{id}).$$

Dimostrazione. Otteniamo

$$\begin{aligned} V_\lambda &= \{v \in V \mid T(v) = \lambda v\} = \{v \in V \mid T(v) - \lambda v = 0\} \\ &= \{v \in V \mid (T - \lambda \text{id})v = 0\} = \ker(T - \lambda \text{id}). \end{aligned}$$

La dimostrazione è conclusa. \square

Ne deduciamo in particolare che l'autospatio $V_\lambda \subset V$ è un sottospazio vettoriale di V .

1.3.3. Molteplicità algebrica e geometrica. Per enunciare il teorema di diagonalizzabilità abbiamo bisogno di introdurre ancora due definizioni.

Sia $T: V \rightarrow V$ un endomorfismo e λ sia un autovalore per T . La *molteplicità algebrica* $m_a(\lambda)$ è la molteplicità di λ come radice del polinomio caratteristico p_T . La *molteplicità geometrica* $m_g(\lambda)$ è la dimensione del sottospazio associato a λ , cioè

$$m_g(\lambda) = \dim V_\lambda.$$

Esempio 1.3.4. Sia $L_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ l'endomorfismo dato da

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

Il polinomio caratteristico è $p_A(\lambda) = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2$. Troviamo un solo autovalore $\lambda_1 = 1$, con molteplicità algebrica $m_a(1) = 2$. D'altra parte,

$$m_g(1) = \dim V_1 = \dim \ker(A - I) = \dim \ker \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 1.$$

Abbiamo quindi trovato in questo caso

$$m_a(1) = 2, \quad m_g(1) = 1.$$

In particolare vediamo che le due molteplicità non sono sempre uguali.

1.3.4. Teorema di diagonalizzabilità. Possiamo finalmente enunciare e dimostrare il teorema seguente.

Teorema 1.3.5 (Teorema di diagonalizzabilità). *Un endomorfismo $T: V \rightarrow V$ è diagonalizzabile se e solo se valgono entrambi i fatti seguenti:*

- (1) *il polinomio caratteristico $p_T(\lambda)$ ha tutte le radici nel campo \mathbb{K} ;*
- (2) *$m_a(\lambda) = m_g(\lambda)$ per ogni autovalore λ .*

Dimostrazione. Se T è diagonalizzabile, esiste una base $\mathcal{B} = \{v_1, \dots, v_n\}$ formata da autovettori, con autovalori $\lambda_1, \dots, \lambda_n$. Prendiamo la matrice associata $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ e notiamo che è diagonale: da questo fatto segue che $p_A(\lambda) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$ e quindi il polinomio caratteristico ha tutte le radici in \mathbb{K} . Usando la matrice A , si verifica anche che $m_a(\lambda) = m_g(\lambda)$, e questo fatto è lasciato come esercizio. Quindi otteniamo sia (1) che (2).

D'altra parte, supponiamo che valgano (1) e (2) e dimostriamo adesso che T è diagonalizzabile. Siano $\lambda_1, \dots, \lambda_k$ gli autovalori. Sia $n = \dim V$. Da (1) segue che

$$m_a(\lambda_1) + \dots + m_a(\lambda_k) = n.$$

Da (2) deduciamo anche che

$$m_g(\lambda_1) + \dots + m_g(\lambda_k) = n.$$

Sia \mathcal{B}^i una base dell'autospazio V_{λ_i} , per ogni $i = 1, \dots, k$. Vogliamo dimostrare adesso che

$$\mathcal{B} = \mathcal{B}^1 \cup \dots \cup \mathcal{B}^k$$

è una base per V . Questo conclude la dimostrazione, perché \mathcal{B} è fatta di autovettori: l'endomorfismo T ha una base \mathcal{B} di autovettori e quindi è diagonalizzabile per definizione.

Indichiamo per semplicità $m_i = m_g(\lambda_i)$. Notiamo che la base \mathcal{B}^i contiene precisamente $m_i = \dim V_{\lambda_i}$ elementi. Ne segue che \mathcal{B} contiene precisamente $m_1 + \dots + m_k = n = \dim V$ elementi. Quindi per dimostrare che \mathcal{B} è una base, è sufficiente verificare che gli n elementi in \mathcal{B} sono indipendenti. Scriviamo

$$\mathcal{B}^i = \{v_1^i, \dots, v_{m_i}^i\}$$

per ogni $i = 1, \dots, k$. Supponiamo di avere una combinazione lineare nulla fra i vettori di \mathcal{B} , del tipo

$$(3) \quad \alpha_1^1 v_1^1 + \dots + \alpha_{m_1}^1 v_{m_1}^1 + \dots + \alpha_1^k v_1^k + \dots + \alpha_{m_k}^k v_{m_k}^k = 0.$$

Il vettore

$$w_i = \alpha_1^i v_1^i + \dots + \alpha_{m_i}^i v_{m_i}^i$$

appartiene a V_{λ_i} . Possiamo riscrivere l'equazione (3) nel modo seguente:

$$w_1 + \dots + w_k = 0.$$

Per ciascun w_i , ci sono due possibilità: o $w_i = 0$, oppure w_i è un autovettore con autovalore λ_i . Il secondo caso non può però accadere: altrimenti l'equazione $w_1 + \dots + w_k = 0$ fornirebbe una combinazione lineare nulla fra autovettori con autovalori distinti, ma questo è proibito per la Proposizione 1.3.1. Otteniamo quindi

$$w_1 = \dots = w_k = 0.$$

Allora per ogni $i = 1, \dots, k$ otteniamo

$$\alpha_1^i v_1^i + \dots + \alpha_{n_i}^i v_{n_i}^i = 0.$$

I vettori $v_1^i, \dots, v_{n_i}^i$ formano la base \mathcal{B}_i e quindi sono indipendenti: allora necessariamente $\alpha_j^i = 0$ per ogni i, j . Quindi i coefficienti della combinazione lineare originale (3) sono tutti nulli e abbiamo dimostrato (con un po' di fatica...) che i vettori di \mathcal{B} sono indipendenti. \square

Prima di applicare il teorema, mostriamo un paio di disuguaglianze che saranno utili per semplificare alcuni argomenti.

Proposizione 1.3.6. *Per ogni autovalore λ_0 valgono le disuguaglianze*

$$1 \leq m_g(\lambda_0) \leq m_a(\lambda_0).$$

Dimostrazione. Se λ_0 è un autovalore, esiste almeno un autovettore $v \neq 0$ tale che $T(v) = \lambda_0 v$ e quindi l'autospazio V_{λ_0} ha dimensione almeno uno. Questo dimostra la prima disuguaglianza.

Per dimostrare la seconda, prendiamo una base $\{v_1, \dots, v_k\}$ di V_{λ_0} , formata da $k = m_g(\lambda_0) = \dim V_{\lambda_0}$ vettori, e completiamola a base $\{v_1, \dots, v_n\}$ di V . Per i primi k vettori della base abbiamo $T(v_i) = \lambda_0 v_i$. Quindi la matrice $[T]_{\mathcal{B}}^{\mathcal{B}}$ è della forma

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} \lambda_0 I & C \\ 0 & D \end{pmatrix}$$

dove $\lambda_0 I \in M(k)$ e $D \in M(n - k)$. Per il Corollario 1.2.15 il polinomio caratteristico di T è del tipo

$$p_T(\lambda) = p_{\lambda_0 I}(\lambda) \cdot p_D(\lambda) = (\lambda_0 - \lambda)^k \cdot p_D(\lambda).$$

Ne segue che λ_0 ha molteplicità almeno $k = m_g(\lambda_0)$ in $p_T(\lambda)$. In altre parole, vale la disuguaglianza $m_a(\lambda_0) \geq m_g(\lambda_0)$. \square

1.3.5. Esempi. Mostriamo alcuni esempi in cui applichiamo il teorema di diagonalizzabilità.

Esempio 1.3.7. Studiamo la diagonalizzabilità su \mathbb{R} della matrice

$$A = \begin{pmatrix} 3 & 0 & 0 \\ -4 & -1 & -8 \\ 0 & 0 & 3 \end{pmatrix}.$$

Si vede facilmente che il polinomio caratteristico è

$$p_A(\lambda) = (3 - \lambda)(-1 - \lambda)(3 - \lambda),$$

quindi ha radici $\lambda_1 = 3$ con $m_a(\lambda_1) = 2$ e $\lambda_2 = -1$ con $m_a(\lambda_2) = 1$.

Tutte le radici di $p_A(\lambda)$ sono reali, quindi A è diagonalizzabile se e solo se le molteplicità algebriche e geometriche di ciascun autovalore coincidono. Per il secondo autovalore λ_2 è facile: la Proposizione 1.3.6 implica che $m_g(\lambda_2) = m_a(\lambda_2) = 1$ e quindi siamo a posto.

Dobbiamo concentrarci solo sull'autovalore λ_1 che ha $m_a(\lambda_1) = 2$. La Proposizione 1.3.6 ci dice che $m_g(\lambda_1)$ può essere 1 oppure 2: nel primo caso A non è diagonalizzabile, nel secondo sì. Facciamo i conti e troviamo

$$m_g(3) = \dim V_3 = \dim \ker(A - 3I) = 3 - \text{rk}(A - 3I).$$

Nell'ultima uguaglianza abbiamo usato il teorema della dimensione. Quindi

$$\begin{aligned} m_g(3) &= 3 - \text{rk} \begin{pmatrix} 3-3 & 0 & 0 \\ -4 & -1-3 & -8 \\ 0 & 0 & 3-3 \end{pmatrix} \\ &= 3 - \text{rk} \begin{pmatrix} 0 & 0 & 0 \\ -4 & -4 & -8 \\ 0 & 0 & 0 \end{pmatrix} = 3 - 1 = 2. \end{aligned}$$

Abbiamo scoperto che $m_a(\lambda_1) = m_g(\lambda_1) = 2$ e quindi A è diagonalizzabile.

Esempio 1.3.8. Studiamo la diagonalizzabilità su \mathbb{R} della matrice

$$A = \begin{pmatrix} 3 & t+4 & 5 \\ -1 & -3 & -1 \\ 0 & 0 & 2 \end{pmatrix}.$$

al variare del parametro $t \in \mathbb{R}$. Si vede facilmente che il polinomio caratteristico è

$$p_A(\lambda) = (2 - \lambda)(\lambda^2 + t - 5).$$

Se $t > 5$, il membro di destra non ha radici reali: quindi $p_A(\lambda)$ ha una radice sola e A non è diagonalizzabile. Se $t \leq 5$, il polinomio ha tre radici reali

$$\lambda_1 = 2, \quad \lambda_2 = \sqrt{5-t}, \quad \lambda_3 = -\sqrt{5-t}.$$

Se le tre radici sono distinte, la matrice A è diagonalizzabile. Restano da considerare i casi in cui le tre radici non sono distinte, e cioè i casi $t = 1$

e $t = 5$. Questi due casi vanno analizzati separatamente con le tecniche mostrate nell'esempio precedente.

Esercizio 1.3.9. Quali delle matrici seguenti sono diagonalizzabili su \mathbb{R} ?

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ -2 & -2 & -2 \end{pmatrix}, \quad \begin{pmatrix} a & b & c \\ 0 & a & d \\ 0 & 0 & a \end{pmatrix}, \quad \begin{pmatrix} 1 & k & 0 \\ 1 & k & 1 \\ 0 & -k & 1 \end{pmatrix}$$

La risposta dipende dai parametri reali presenti nella matrice.

1.3.6. Esistenza di autovettori. Usiamo le proprietà algebriche dei polinomi per dimostrare un paio di proposizioni sull'esistenza di autovettori in generale.

Proposizione 1.3.10. *Un endomorfismo $T: V \rightarrow V$ di uno spazio vettoriale complesso V ha sempre almeno un autovettore.*

Dimostrazione. Il polinomio caratteristico ha sempre almeno una radice (su \mathbb{C}). Quindi T ha almeno un autovalore λ , e quindi anche un autovettore. \square

Questo risultato non è vero su \mathbb{R} . Ad esempio, abbiamo visto che una rotazione di un angolo $\theta \neq 0, \pi$ è una trasformazione di \mathbb{R}^2 che non ha autovettori. Possiamo comunque ottenere lo stesso risultato in dimensione dispari:

Proposizione 1.3.11. *Sia V uno spazio vettoriale reale di dimensione dispari. Un endomorfismo $T: V \rightarrow V$ ha sempre un autovettore.*

Dimostrazione. Il polinomio caratteristico ha grado dispari e quindi ha sempre una radice in \mathbb{R} . \square

CAPITOLO 2

Forma di Jordan

In questo capitolo affronteremo il problema seguente: sappiamo che un endomorfismo è più facile da studiare se è diagonalizzabile; cosa possiamo fare però nel caso in cui non lo sia? Vedremo che potremo comunque scegliere una matrice relativamente semplice, che è “quasi” diagonale, eccetto che per alcuni valori sopra la diagonale: questo tipo di matrice particolare e utile è detto *matrice di Jordan*.

In tutto questo capitolo il campo base \mathbb{K} sarà sempre quello dei complessi $\mathbb{K} = \mathbb{C}$. In questo modo potremo sfruttare il teorema fondamentale dell'algebra – cioè il fatto che ogni polinomio di grado n abbia n radici contate con molteplicità.

2.1. Forma di Jordan

2.1.1. Blocco di Jordan. Un *blocco di Jordan* è una matrice quadrata $n \times n$ del tipo

$$B_{\lambda,n} = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}.$$

Qui $\lambda \in \mathbb{C}$ è uno scalare complesso fissato qualsiasi. In altre parole un blocco di Jordan $B_{\lambda,n}$ è una matrice $n \times n$ triangolare superiore in cui tutte le caselle della diagonale principale hanno lo stesso numero $\lambda \in \mathbb{C}$ e tutte le caselle sulla diagonale successiva hanno lo stesso numero 1. Tutte le altre diagonaloni hanno solo zeri. Ad esempio queste sono blocchi di Jordan:

$$(3), \quad \begin{pmatrix} i & 1 \\ 0 & i \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Notiamo che una matrice 1×1 è sempre di Jordan. Osserviamo anche che una matrice di Jordan $n \times n$ ha un solo autovalore λ .

Proposizione 2.1.1. *Le molteplicità dell'autovalore λ sono:*

$$m_a(\lambda) = n, \quad m_g(\lambda) = 1.$$

In particolare se $n \geq 2$ un blocco di Jordan non è mai diagonalizzabile.

Dimostrazione. La molteplicità algebrica è chiara perché la matrice è triangolare superiore. Per la molteplicità geometrica, notiamo che

$$m_g(\lambda) = n - \text{rk} \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix} = n - (n - 1) = 1.$$

La matrice descritta ha effettivamente rango $n - 1$ perché, dopo aver tolto la prima colonna, le colonne rimanenti sono chiaramente indipendenti. \square

2.1.2. Matrice di Jordan. Una *matrice di Jordan* è una matrice a blocchi

$$M = \begin{pmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B_k \end{pmatrix}$$

dove B_1, \dots, B_k sono blocchi di Jordan di taglia arbitraria. Notiamo che tutti i blocchi che non stanno sulla diagonale principale sono nulli. Ad esempio, queste sono matrici di Jordan:

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

Le prime due matrici hanno due blocchi di Jordan, la terza ne ha tre, l'ultima ne ha uno solo (è lei stessa un blocco di Jordan). Qualsiasi matrice diagonale $n \times n$ è una matrice di Jordan con n blocchi di taglia 1×1 .

Notiamo un fatto semplice ma importante: una matrice di Jordan è sempre triangolare superiore e quindi i suoi autovalori sono tutti sulla diagonale.

2.1.3. Teorema di Jordan. Enunciamo adesso il teorema più importante di questa sezione.

Teorema 2.1.2. *Qualsiasi matrice $n \times n$ a coefficienti complessi A è simile ad una matrice di Jordan J . La matrice di Jordan J è unica a meno di permutare i blocchi.*

Una dimostrazione di questo teorema sarà fornita più avanti in questo capitolo. È importante notare come la matrice di Jordan J sia unica (a meno di permutazione di blocchi): per questo J viene chiamata la *forma canonica di Jordan* di M .

Corollario 2.1.3. *Due matrici M e M' sono simili \iff le loro matrici di Jordan J e J' hanno gli stessi blocchi.*

Esempio 2.1.4. Consideriamo le matrici seguenti:

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

Quali di queste sono simili fra loro? La seconda e la terza sono simili, perché hanno gli stessi blocchi, solo permutati in modo diverso. La prima, la seconda e la quarta sono però tutte non simili fra loro, perché hanno blocchi diversi. L'ultima matrice non è in forma di Jordan e quindi non possiamo dire immediatamente se sia simile ad una delle precedenti: c'è del lavoro da fare.

Il Teorema 2.1.2 è molto potente, ma osserviamo che enunciato in questo modo non fornisce un metodo per costruire J a partire da M . Esiste un algoritmo per costruire J da M che descriveremo alla fine di questa sezione, ma non è semplicissimo e per adesso ci limitiamo ad usare il teorema: nei casi più semplici questo sarà sufficiente.

Esempio 2.1.5. Due matrici simili A e B hanno lo stesso determinante, la stessa traccia, lo stesso polinomio caratteristico, gli stessi autovalori, e per ciascun autovalore hanno anche le stesse molteplicità algebrica e geometrica. Usando tutte queste proprietà possiamo spesso riconoscere due matrici che non sono simili.

Ad esempio, le matrici dell'Esempio 2.1.4 hanno tutte lo stesso polinomio caratteristico (e quindi stessa traccia, determinante e autovalori). Però le molteplicità geometriche dell'autovalore 2 non sono tutte uguali: facendo un conto vediamo che queste sono rispettivamente

$$3, \quad 2, \quad 2, \quad 1, \quad 1.$$

La quinta matrice quindi può essere simile solo alla quarta. Ma lo è veramente? Usando il Teorema 2.1.2 possiamo rispondere affermativamente nel modo seguente. Ci chiediamo: quale è la forma di Jordan dell'ultima matrice? Siccome ha solo l'autovalore 2, deve essere una delle precedenti: si verifica facilmente infatti che queste sono infatti tutte le possibili matrici di Jordan 3×3 con un solo autovalore 2. Di queste, solo la quarta può essere simile a lei, quindi sono simili.

Esempio 2.1.6. Due matrici A e B che hanno lo stesso polinomio caratteristico e le stesse molteplicità algebriche e geometriche degli autovalori, sono sempre simili? La risposta è negativa: le matrici seguenti

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

hanno un solo autovalore 1 ed in entrambi i casi la molteplicità geometrica di 1 è 2, come si vede facilmente con un conto. Però si decompongono in blocchi di Jordan diversi: quella di sinistra in due blocchi di ordine 2, quella di destra in due blocchi di ordine 3 e 1. Quindi per il Teorema 2.1.2 non sono simili.

2.1.4. Matrici nilpotenti. Una matrice quadrata A è *nilpotente* se esiste un $k \geq 0$ per cui $A^k = 0$. Ovviamente, se questo accade, allora $A^h = 0$ anche per ogni $h \geq k$. L'*indice di nilpotenza* di A è il più piccolo numero intero $k \geq 1$ per cui $A^k = 0$.

Esempio 2.1.7. Una matrice A è nilpotente con indice di nilpotenza 1 se e solo se $A = 0$. Quindi solo la matrice nulla è nilpotente con indice 1.

Ad esempio, consideriamo il blocco di Jordan $B = B_{0,2}$. Otteniamo

$$B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Quindi B è nilpotente con indice di nilpotenza 2. Con $B = B_{0,3}$ otteniamo

$$B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad B^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Questo ci dice che B è nilpotente con indice di nilpotenza 3. Con la matrice $B = B_{0,4}$ otteniamo

$$B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

e infine $B^4 = 0$, quindi B ha indice di nilpotenza 4. Più in generale, abbiamo capito che ogni volta che moltiplichiamo $B_{0,n}$ per se stesso spostiamo i numeri 1 sulla diagonale nella diagonale immediatamente superiore. In questo modo dimostriamo il fatto seguente.

Proposizione 2.1.8. *Il blocco $B_{0,n}$ è nilpotente con indice di nilpotenza n .*

Esistono molte altre matrici nilpotenti oltre alle $B_{0,n}$. Notiamo infatti che qualsiasi matrice simile a $B_{0,n}$ è anch'essa nilpotente, con lo stesso indice di nilpotenza n .

Esercizio 2.1.9. Determinare tutte le matrici complesse 2×2 nilpotenti.

2.1.5. Algoritmo per determinare J . Come possiamo determinare la matrice di Jordan J associata ad una data matrice complessa A ? In molti casi è sufficiente andare per tentativi, come abbiamo visto sopra: si elencano le possibili matrici di Jordan e si usa la molteplicità geometrica per eliminarne alcune, sperando di restare alla fine con una sola. A volte però questo non funziona: esponiamo adesso un algoritmo più complesso che funziona sempre.

Sia λ un autovalore di A . Per ciascun numero intero $j = 1, 2, \dots$ definiamo l'*autospazio generalizzato*

$$V_\lambda^j = \ker(A - \lambda I)^j$$

e la *molteplicità geometrica generalizzata*

$$m_j = m_g(\lambda)_j = \dim V_\lambda^j.$$

La molteplicità m_1 è l'usuale molteplicità geometrica di λ , mentre per le successive otteniamo sempre

$$m_1 \leq m_2 \leq m_3 \leq \dots$$

Proposizione 2.1.10. Due matrici simili A e B hanno le stesse molteplicità geometriche generalizzate per ciascun autovalore.

Dimostrazione. Se $A = M^{-1}BM$, allora

$$(A - \lambda I)^j = (M^{-1}BM - \lambda M^{-1}M)^j = (M^{-1}(B - \lambda I)M)^j = M^{-1}(B - \lambda I)^j M.$$

Quindi le matrici $(A - \lambda I)^j$ e $(B - \lambda I)^j$ sono anch'esse simili; hanno lo stesso rango, e quindi anche la stessa dimensione del nucleo. \square

Sia J la matrice di Jordan associata ad A . Poiché J e A sono simili, hanno le stesse molteplicità geometriche generalizzate. Il punto fondamentale è che queste determinano univocamente J , come mostrato nella proposizione seguente. Poniamo $m_0 = 0$.

Proposizione 2.1.11. Per ogni $i \geq 1$, il numero $m_i - m_{i-1}$ è il numero di blocchi di J con autovalore λ e di taglia $\geq i$. In particolare la molteplicità geometrica m_1 è il numero di blocchi con autovalore λ .

Dimostrazione. Ricordiamo che $B_{\lambda,n}$ indica un blocco di Jordan di ordine n con autovalore λ . A meno di permutare i blocchi, possiamo mettere tutti i blocchi di J con autovalore λ all'inizio, e scrivere

$$J = \begin{pmatrix} B_{\lambda,n_1} & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & B_{\lambda,n_a} & 0 \\ 0 & \dots & 0 & J' \end{pmatrix}.$$

dove n_1, \dots, n_a sono gli ordini dei blocchi con autovalore λ . La matrice di Jordan J' contiene i blocchi con gli altri autovalori. Notiamo che

$$m_1 = \dim \ker(J - \lambda I) = \dim \ker \begin{pmatrix} B_{0,n_1} & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & B_{0,n_a} & 0 \\ 0 & \dots & 0 & J' - \lambda I \end{pmatrix}.$$

La matrice $J' - \lambda I$ è invertibile (perché è triangolare superiore con elementi tutti non nulli sulla diagonale), e si vede subito che m_1 è il numero di colonne

nulle della matrice. C'è una colonna nulla in ogni blocco $B_{0,n}$, quindi $m_1 = a$ è il numero di tali blocchi.

Calcoliamo ora m_2 e notiamo che

$$m_2 = \dim \ker(J - \lambda I)^2 = \dim \ker \begin{pmatrix} B_{0,n_1}^2 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & B_{0,n_a}^2 & 0 \\ 0 & \dots & 0 & (J' - \lambda I)^2 \end{pmatrix}.$$

La matrice $(J' - \lambda I)^2$ è sempre invertibile, e m_2 è sempre il numero di colonne nulle. Le colonne nulle sono le m_1 colonne di prima, più $m_2 - m_1$ nuove colonne, una per ogni blocco $B_{0,n}$ di taglia almeno due. Quindi $m_2 - m_1$ è proprio il numero di blocchi di taglia ≥ 2 . I passi successivi della dimostrazione sono analoghi. \square

Esempio 2.1.12. L'esempio numerico seguente dovrebbe chiarire meglio cosa succede ed aiutare anche a capire la dimostrazione appena conclusa. Consideriamo

$$J = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Questa ha un blocco di ordine 3, due blocchi di ordine 2 e un blocco di ordine 1. Otteniamo

$$m_1 = \dim \ker(J - 2I) = \dim \ker \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = 4$$

perché sono comparse 4 colonne di zeri, una per ogni blocco: infatti la molteplicità geometrica m_1 è il numero di blocchi. Continuiamo e troviamo

$$m_2 = \dim \ker(J - 2I)^2 = \dim \ker \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = 7.$$

Al secondo passaggio, si sono manifestate 3 nuove colonne di zeri: una per ogni blocco di taglia ≥ 2 (i blocchi di taglia 1 erano già spariti). I blocchi di taglia ≥ 2 sono effettivamente $m_2 - m_1 = 7 - 4 = 3$.

Infine $(J - 2I)^3 = 0$ ci dice che $m_3 = 8$ e infatti i blocchi di taglia ≥ 3 sono esattamente $m_3 - m_2 = 8 - 7 = 1$.

Esempio 2.1.13. Usiamo l'algoritmo per determinare la forma di Jordan J della matrice

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

Verifichiamo che il polinomio caratteristico è $p_A(\lambda) = (\lambda - 2)^4$. Per determinare la forma di Jordan dobbiamo calcolare le molteplicità geometriche generalizzate dell'autovalore 2. Iniziamo:

$$m_1 = \dim \ker(A - 2I) = \dim \ker \begin{pmatrix} 0 & 2 & 0 & -1 \\ 0 & -2 & 0 & 1 \\ 1 & 5 & 0 & -1 \\ 0 & -4 & 0 & 2 \end{pmatrix} = 4 - 2 = 2.$$

Scopriamo che la molteplicità geometrica è 2, quindi il numero di blocchi di J è 2. Questo però non è sufficiente a determinare J , perché ci sono due casi possibili, cioè

$$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad \text{oppure} \quad \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Allora continuiamo a calcolare le molteplicità geometriche generalizzate:

$$m_2 = \dim \ker(A - 2I)^2 = \dim \ker \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -6 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 4 - 1 = 3.$$

Quindi il numero di blocchi di J di taglia ≥ 2 è $m_2 - m_1 = 3 - 2 = 1$. Questo esclude il primo dei due casi possibili per J elencati sopra, quindi troviamo

$$J = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

2.2. Teorema di Cayley – Hamilton

In questa sezione vogliamo iniziare ad abbinare polinomi e matrici – più precisamente, dato un polinomio $p(x)$ e una matrice A , diamo un senso all'espressione $p(A)$ e ci chiediamo quando questa espressione si annulla.

2.2.1. Polinomi di matrici. Sia A una matrice $n \times n$ e $p(x)$ un polinomio di grado k . Il polinomio è del tipo

$$p(x) = a_k x^k + \dots + a_1 x + a_0.$$

Definiamo adesso una nuova matrice $p(A)$ nel modo seguente:

$$p(A) = a_k A^k + \dots + a_1 A + a_0 I_k.$$

Diciamo che il polinomio $p(x)$ *annulla* la matrice A se

$$p(A) = 0.$$

In questa equazione "0" è la matrice $n \times n$ nulla. Vogliamo studiare come sono fatti i polinomi che annullano una matrice data A . Ovviamente, fra questi c'è sempre il *polinomio banale*, cioè la costante 0, che annulla qualsiasi cosa. Prima domanda: c'è dell'altro? La risposta è affermativa:

Proposizione 2.2.1. *Esistono sempre polinomi non banali che annullano A .*

Dimostrazione. Consideriamo le matrici

$$I, A, A^2, \dots, A^{n^2}.$$

Queste sono $n^2 + 1$ elementi nello spazio $M(n)$ di tutte le matrici $n \times n$. Siccome questo spazio ha dimensione n^2 , questi $n^2 + 1$ elementi non possono essere indipendenti: c'è quindi una combinazione non banale che li annulla

$$a_0 I + a_1 A + a_2 A^2 + \dots + a_{n^2} A^{n^2}.$$

In altre parole, se definiamo

$$p(x) = a_{n^2} x^{n^2} + \dots + a_1 x + a_0$$

otteniamo $p(A) = 0$. □

La dimostrazione fornisce un polinomio $p(x)$ di grado al massimo n^2 che annulla A . Vogliamo adesso migliorare questo risultato, mostrando che il polinomio caratteristico di A , che ha grado solo n , annulla A . Per ottenere questo profondo risultato noto come *Teorema di Cayley – Hamilton* abbiamo bisogno di alcuni fatti preliminari.

2.2.2. Matrici a blocchi. In questo capitolo, una *matrice a blocchi* è una matrice del tipo

$$M = \begin{pmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B_k \end{pmatrix}$$

dove B_1, \dots, B_k sono matrici quadrate, di taglia arbitraria. A parte i blocchi B_1, \dots, B_k che stanno sulla diagonale, tutti gli altri blocchi della matrice sono nulli. Un esempio importante di matrice a blocchi è ovviamente una matrice di Jordan.

Proposizione 2.2.2. *Se M è una matrice a blocchi e $p(x)$ è un polinomio, allora otteniamo*

$$p(M) = \begin{pmatrix} p(B_1) & 0 & \dots & 0 \\ 0 & p(B_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p(B_k) \end{pmatrix}.$$

In altre parole, per applicare un polinomio p su M basta farlo su ogni blocco.

Dimostrazione. Sia

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

Otteniamo

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

che può essere scritto come

$$a_n \begin{pmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B_k \end{pmatrix}^n + \dots + a_1 \begin{pmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B_k \end{pmatrix} + a_0 \begin{pmatrix} I & 0 & \dots & 0 \\ 0 & I & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & I \end{pmatrix}$$

dove le I nell'ultima matrice sono tutte matrici identità della taglia opportuna.

Questo è equivalente a

$$\begin{pmatrix} a_n B_1^n & 0 & \dots & 0 \\ 0 & a_n B_2^n & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_n B_k^n \end{pmatrix} + \dots + \begin{pmatrix} a_1 B_1 & 0 & \dots & 0 \\ 0 & a_1 B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_1 B_k \end{pmatrix} + \begin{pmatrix} a_0 I & 0 & \dots & 0 \\ 0 & a_0 I & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_0 I \end{pmatrix}.$$

Sommando le matrici otteniamo

$$\begin{pmatrix} p(B_1) & 0 & \dots & 0 \\ 0 & p(B_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p(B_k) \end{pmatrix}.$$

La dimostrazione è completa. \square

2.2.3. Matrici simili. Come sappiamo, due matrici simili hanno molte cose in comune. La proposizione seguente conferma questo fatto generale.

Proposizione 2.2.3. *Se A e B sono simili e $p(x)$ è un polinomio, allora*

$$p(A) = 0 \iff p(B) = 0.$$

Dimostrazione. Sia

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

Per ipotesi $A = MBM^{-1}$ per qualche matrice invertibile M . Ricordiamo che

$$A^k = (MBM^{-1})^k = MB^k M^{-1}.$$

Otteniamo

$$\begin{aligned} p(A) &= a_n (MBM^{-1})^n + \dots + a_1 MBM^{-1} + a_0 I \\ &= M a_n B^n M^{-1} + \dots + M a_1 B M^{-1} + M a_0 I M^{-1} \\ &= M p(B) M^{-1}. \end{aligned}$$

Nell'ultima uguaglianza abbiamo raccolto M a sinistra e M^{-1} a destra. Segue che $p(B) = 0 \implies p(A) = 0$. Invertendo i ruoli di A e B si ottiene anche la freccia opposta. \square

Abbiamo scoperto che due matrici simili sono annullate precisamente dagli stessi polinomi.

2.2.4. Teorema di Cayley – Hamilton. Abbiamo infine tutti gli strumenti per enunciare e dimostrare il fatto seguente.

Teorema 2.2.4 (Cayley – Hamilton). *Il polinomio caratteristico p_A di A annulla A . In altre parole, vale*

$$p_A(A) = 0.$$

Dimostrazione. Consideriamo prima il caso in cui $A = B_{\lambda_0, n}$ sia un blocco di Jordan con autovalore λ_0 . Otteniamo $p_A = (\lambda_0 - \lambda)^n$. Quindi

$$p_A(A) = (\lambda_0 I - A)^n = \pm (A - \lambda_0 I)^n = \pm B_{0, n}^n = 0.$$

Ricordiamo infatti che $B_{\lambda_0, n} - \lambda_0 I = B_{0, n}$ è nilpotente di indice n . Il teorema in questo caso è dimostrato.

Passiamo ora al caso in cui $A = J$ sia una matrice di Jordan, con blocchi di Jordan B_1, \dots, B_k . Notiamo che

$$p_J = p_{B_1} \cdots p_{B_k}.$$

Per il caso precedente, sappiamo che $p_{B_i}(B_i) = 0$ e quindi anche $p_J(B_i) = 0$, perché p_{B_i} è un fattore di p_J . A questo punto la Proposizione 2.2.2 ci dà

$$p_J(J) = \begin{pmatrix} p_J(B_1) & 0 & \dots & 0 \\ 0 & p_J(B_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p_J(B_k) \end{pmatrix} = 0.$$

Il teorema è quindi verificato anche in questo caso.

Una generica matrice A è sempre simile ad una matrice di Jordan J . Queste hanno lo stesso polinomio caratteristico $p_A = p_J$, e se $p_J(J) = 0$ allora anche $p_A(A) = 0$ per la Proposizione 2.2.3. La dimostrazione è conclusa. \square

Esempio 2.2.5. La matrice

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

rappresenta su \mathbb{R}^2 una rotazione antioraria di \mathbb{R}^2 di angolo $\frac{\pi}{2}$. Sappiamo che $p_A(\lambda) = \lambda^2 + 1$. Per il Teorema di Cayley – Hamilton abbiamo

$$p_A(A) = A^2 + I = 0.$$

Effettivamente si verifica che $A^2 = -I$.

2.3. Polinomio minimo

Il Teorema di Cayley – Hamilton dice che il polinomio caratteristico p_A annulla la matrice A . Se A è una matrice $n \times n$, allora p_A ha grado n . Ci poniamo adesso la domanda seguente: esistono polinomi di grado minore di n che annullano A ? Questa domanda ci porta alla definizione del *polinomio minimo* di una matrice quadrata.

2.3.1. Polinomio minimo. Ricordiamo che un polinomio $p(x)$ è *monico* se il termine di grado più alto di $p(x)$ ha coefficiente 1. In altre parole, un polinomio p di grado k è monico se è del tipo

$$p(x) = x^k + a_{k-1}x^{k-1} + \dots + a_1x + a_0.$$

Notiamo che qualsiasi polinomio non banale può essere trasformato in un polinomio monico dividendolo per il coefficiente a_k del termine di grado più alto. Sia ora A una matrice complessa $n \times n$.

Definizione 2.3.1. Il *polinomio minimo* di A è il polinomio monico $m_A(x)$ di grado minore che annulla A .

Proposizione 2.3.2. *La definizione è ben posta.*

Dimostrazione. Nella definizione abbiamo dato per scontato che ci sia un unico polinomio m_A con quelle proprietà, e adesso dobbiamo dimostrarlo. Supponiamo di avere due polinomi $m_1(x)$ e $m_2(x)$ monici distinti, entrambi di grado minimo k che annullano A . Siccome sono entrambi monici, la differenza $q(x) = m_1(x) - m_2(x)$ è un polinomio non banale di grado strettamente minore di k . Otteniamo

$$q(A) = m_1(A) - m_2(A) = 0 - 0 = 0.$$

Quindi $q(x)$ è un polinomio non banale di grado $< k$ che annulla A , ma questo è assurdo. \square

Esempio 2.3.3. Se $A = 0$ è la matrice nulla, allora il suo polinomio minimo è $m_A(x) = x$. Infatti abbiamo

$$m_A(A) = A = 0$$

e quindi m_A annulla A . Questo è sicuramente il polinomio monico di grado minimo fra quelli che annullano A perché ha grado uno.

Se $A = I$ allora il suo polinomio minimo è $m_A(x) = x - 1$. Infatti abbiamo

$$m_A(A) = A - I = I - I = 0$$

e concludiamo come nel caso precedente. Più in generale, le uniche matrici A aventi polinomio minimo di grado uno, cioè con $m_A(x) = x - c$ per qualche $c \in \mathbb{C}$, sono le matrici del tipo $A = cI$.

2.3.2. Il polinomio minimo divide tutti gli altri. Siano $a(x)$ e $b(x)$ due polinomi. Ricordiamo che possiamo sempre dividere il primo per il secondo e ottenere

$$a(x) = b(x)q(x) + r(x)$$

dove $q(x)$ è il quoziente e $r(x)$ è il resto, che ha grado strettamente più piccolo di $b(x)$. Nel caso in cui $r(x)$ sia il polinomio nullo, otteniamo $a(x) = b(x)q(x)$ e diciamo che $b(x)$ divide $a(x)$.

Sia A una matrice quadrata $n \times n$.

Proposizione 2.3.4. *Il polinomio minimo m_A divide qualsiasi polinomio che annulla A .*

Dimostrazione. Sia $a(x)$ un polinomio che annulla A . Dividendo $a(x)$ per $m_A(x)$ otteniamo

$$a(x) = m_A(x)q(x) + r(x)$$

per qualche r di grado minore di quello di m_A . Ne segue che

$$r(A) = a(A) - m_A(A)q(A) = 0 - 0q(A) = 0$$

e allora anche r annulla A . Poiché r ha grado minore di m_A , il polinomio r deve essere nullo. Quindi m_A divide a . \square

Corollario 2.3.5. *Il polinomio minimo m_A divide sempre il polinomio caratteristico p_A .*

Dimostrazione. Per il Teorema di Cayley – Hamilton, il polinomio caratteristico p_A annulla A . \square

Esempio 2.3.6. Sapendo che m_A divide p_A , è possibile determinare m_A per tentativi. Ad esempio, consideriamo

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

Vediamo che il polinomio caratteristico è $p_A(x) = (x-1)^2$. Siccome m_A divide p_A , il polinomio m_A è uno dei due polinomi seguenti:

$$x - 1, \quad (x - 1)^2.$$

Procediamo per tentativi, iniziando con il primo: se $q(x) = x - 1$, otteniamo $q(A) = A - I \neq 0$. Quindi il primo non va bene. Allora $m_A(x) = (x - 1)^2$ è il secondo, per esclusione. Il lettore può verificare che infatti $(A - I)^2 = 0$.

Esempio 2.3.7. Calcoliamo il polinomio minimo della matrice

$$A = \begin{pmatrix} i & 1 & 0 \\ 0 & i & 1 \\ 0 & 0 & i \end{pmatrix}.$$

Troviamo $p_A(x) = -(x - i)^3$. Quindi il polinomio minimo è uno dei seguenti:

$$x - i, \quad (x - i)^2, \quad (x - i)^3.$$

Se $q(x) = x - i$, otteniamo $q(A) = A - iI \neq 0$, quindi il primo non va bene. Se $q(x) = (x - i)^2$, si verifica che $q(A) = (A - iI)^2 \neq 0$, quindi neanche il secondo va bene. Quindi $m_A(x) = (x - i)^3$.

2.3.3. Blocchi di Jordan. Calcoliamo il polinomio minimo di un blocco di Jordan. Ricordiamo che $B_{\lambda,n}$ è il blocco di Jordan $n \times n$ con autovalore λ .

Proposizione 2.3.8. *Il polinomio minimo di $B = B_{\lambda_0,n}$ è uguale al polinomio caratteristico:*

$$m_B(\lambda) = p_B(\lambda) = (\lambda - \lambda_0)^n.$$

Dimostrazione. Sappiamo che m_B divide p_B e quindi $m_B(\lambda) = (\lambda - \lambda_0)^k$ per qualche $1 \leq k \leq n$. Dobbiamo determinare il k giusto. Se prendiamo $q(\lambda) = (\lambda - \lambda_0)^k$, otteniamo

$$q(B) = (B - \lambda_0 I)^k = B_{0,n}^k$$

dove $B_{0,n}$ è la matrice nilpotente considerata nella Proposizione 2.1.8, che sappiamo già avere ordine di nilpotenza n . Quindi $q(B) = 0$ solo se $k = n$ e ne ricaviamo che $m_A = p_A = (\lambda - \lambda_0)^n$. \square

2.3.4. Matrici a blocchi. Consideriamo una matrice a blocchi

$$A = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}.$$

Per i polinomi caratteristici, sappiamo che

$$p_A(x) = p_B(x)p_C(x).$$

Il polinomio minimo si comporta diversamente.

Definizione 2.3.9. Dati due polinomi monici $a(x)$ e $b(x)$, il *minimo comune multiplo* di a e b è il più piccolo polinomio monico $m(x)$ che è divisibile sia da a che da b .

Come per i numeri interi, il minimo comune multiplo $m(x)$ fra $a(x)$ e $b(x)$ è un polinomio che divide il prodotto $a(x)b(x)$, ma potrebbe avere grado minore.

Esempio 2.3.10. Il minimo comune multiplo fra $(x-1)(x-2)$ e $(x-1)^2$ è $(x-1)(x-2)^2$.

Esercizio 2.3.11. Se A è una matrice a blocchi come sopra, allora $m_A(x)$ è il minimo comune multiplo fra $m_B(x)$ e $m_C(x)$.

2.3.5. Da Jordan a polinomio minimo. Mostriamo adesso che se conosciamo già la forma di Jordan, è molto facile determinare il polinomio minimo.

Sia come sempre A una matrice complessa $n \times n$. Siano $\lambda_1, \dots, \lambda_k$ gli autovalori di A . Il polinomio caratteristico di A è

$$p_A(\lambda) = \pm(\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_k)^{m_k}$$

dove $m_i = m_a(\lambda_i)$ è la molteplicità algebrica di λ_i . Il polinomio minimo m_A di A divide p_A e quindi è del tipo

$$m_A(\lambda) = \pm(\lambda - \lambda_1)^{d_1} \cdots (\lambda - \lambda_k)^{d_k}$$

dove d_1, \dots, d_k sono dei numeri da determinare. Per adesso sappiamo solo che $d_i \leq m_i$ per ogni i . Sia J la matrice di Jordan associata a A .

Proposizione 2.3.12. Il numero d_i è la massima taglia dei blocchi di Jordan in J con autovalore λ_i .

Dimostrazione. Per un singolo blocco di Jordan $B = B_{\lambda_0, n}$ otteniamo $m_B(\lambda) = (\lambda - \lambda_0)^n$ ed n è effettivamente la taglia di B , quindi ci siamo. Se J è fatto di vari blocchi di Jordan, usando più volte l'Esercizio 2.3.11 troviamo che p_J è il minimo comune multiplo di tutti i polinomi $(\lambda - \lambda_i)^d$ al variare di tutti i blocchi di Jordan con autovalore λ_i e di taglia variabile d . Quindi otteniamo proprio

$$m_A(\lambda) = (\lambda - \lambda_1)^{d_1} \cdots (\lambda - \lambda_k)^{d_k}$$

dove effettivamente d_i è la massima taglia dei blocchi con autovalore λ_i . \square

In particolare, troviamo che $d_i \geq 1$. Quindi per ogni $i = 1, \dots, k$ abbiamo

$$1 \leq d_i \leq m_i = m_a(\lambda_i).$$

Osservazione 2.3.13. Le stesse disuguaglianze valgono per la molteplicità geometrica, ma attenzione a non fare confusione: il numero d_i *non* è uguale a $m_g(\lambda_i)$ in generale! Infatti d_i è la massima taglia dei blocchi con autovalore λ_i , mentre $m_g(\lambda_i)$ è il numero di questi blocchi. Questi due quantità non sono correlate e sono spesso differenti.

Esempio 2.3.14. Il polinomio minimo di

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

è $(x - 2)(x - 3)^3$.

Enunciamo un importante corollario.

Corollario 2.3.15. *Una matrice A è diagonalizzabile se e solo se il polinomio minimo non ha radici ripetute, cioè è del tipo*

$$m_A(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_n)$$

dove $\lambda_1, \dots, \lambda_n$ sono gli autovalori (distinti) di A .

Dimostrazione. Una matrice A è diagonalizzabile se e solo se la sua forma di Jordan è diagonale, e questo accade precisamente quando tutti i blocchi hanno taglia 1. \square

2.3.6. Applicazioni. Il Corollario 2.3.15 fornisce un criterio molto potente di diagonalizzabilità. Notiamo una differenza fra polinomio caratteristico e minimo: il primo si calcola subito, ma non ci dice se A è diagonalizzabile; il secondo è più difficile da calcolare, ma ci dice subito se A è diagonalizzabile o no. Mostriamo qualche applicazione di questo criterio.

Proposizione 2.3.16. *Se A è una matrice complessa tale che $A^2 = A$, allora A è simile ad una matrice diagonale con valori 1 e 0 sulla diagonale.*

Dimostrazione. Per ipotesi il polinomio $q(x) = x^2 - x = x(x - 1)$ annulla A . Allora il polinomio minimo $m_A(x)$ deve dividere $q(x)$ ed è quindi uno di questi:

$$x(x - 1), \quad x, \quad x - 1.$$

In tutti i casi $m_A(x)$ non ha radici ripetute e quindi A è diagonalizzabile. Ne segue che A è simile ad una matrice diagonale D , con gli autovalori sulla diagonale. Gli autovalori sono le radici di $m_A(x)$ e quindi possono essere solo 0 e 1. \square

L'esercizio seguente si risolve in modo analogo.

Esercizio 2.3.17. Se A è una matrice complessa tale che $A^2 = I$, allora A è simile ad una matrice diagonale con valori 1 e -1 sulla diagonale.

Attenzione però al caso seguente:

Esercizio 2.3.18. Se A è una matrice complessa tale che $A^2 = 0$, allora A è diagonalizzabile se e solo se $A = 0$.

CAPITOLO 3

Prodotti scalari

In questo capitolo introduciamo per la prima volta alcuni concetti familiari della geometria euclidea: quello di *lunghezza* di un vettore, di *angolo* fra due vettori, e di *distanza* fra due punti.

Come possiamo introdurre queste nozioni? Il modo più efficiente consiste nella definizione di un *prodotto scalare* fra vettori.

3.1. Introduzione

3.1.1. Definizioni. Sia V uno spazio vettoriale reale. In tutto questo capitolo – a differenza dei precedenti – supporremo sempre che il campo sia \mathbb{R} e non \mathbb{C} : in molti punti sarà importante parlare di numeri *positivi* e *negativi*, cosa che possiamo fare solo in \mathbb{R} e non in \mathbb{C} .

Definizione 3.1.1. Un *prodotto scalare* su V è una applicazione

$$\begin{aligned} V \times V &\longrightarrow \mathbb{R} \\ (v, w) &\longmapsto \langle v, w \rangle \end{aligned}$$

che soddisfa i seguenti assiomi:

- (1) $\langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle$
- (2) $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$
- (3) $\langle v, w + w' \rangle = \langle v, w \rangle + \langle v, w' \rangle$
- (4) $\langle v, \lambda w \rangle = \lambda \langle v, w \rangle$
- (5) $\langle v, w \rangle = \langle w, v \rangle$

per ogni $v, v', w, w' \in V$ e ogni $\lambda \in \mathbb{R}$.

Il prodotto scalare è una applicazione che prende come *input* due vettori $v, w \in V$ e restituisce come *output* uno scalare, che indichiamo con $\langle v, w \rangle$. In alcuni aspetti il prodotto scalare si comporta come il prodotto fra numeri: gli assiomi (1) e (3) della definizione indicano che il prodotto è distributivo rispetto alla somma, e l'assioma (5) indica che è commutativo. L'aggettivo "scalare" sta però ad indicare una differenza importante: il prodotto scalare di due vettori non è un vettore, ma "solo" uno scalare.

Un po' di terminologia: gli assiomi (1)–(4) dicono che il prodotto scalare è *bilineare*, cioè lineare "sia a sinistra che a destra"; l'assioma (5) dice che il prodotto scalare è *simmetrico*. Notiamo che gli assiomi (1), (2) e (5) implicano (3) e (4).

Come prima applicazione degli assiomi, notiamo che $\langle v, 0 \rangle = 0$ per ogni $v \in V$. Infatti

$$\langle v, 0 \rangle = \langle v, 0 + 0 \rangle = \langle v, 0 \rangle + \langle v, 0 \rangle$$

implica $\langle v, 0 \rangle = 0$. In particolare otteniamo $\langle 0, 0 \rangle = 0$. Introduciamo ora degli assiomi aggiuntivi.

Definizione 3.1.2. Un prodotto scalare è:

- *degenere* se esiste $v \neq 0$ tale che $\langle v, w \rangle = 0$ per ogni $w \in V$;
- *definito positivo* se $\langle v, v \rangle > 0$ per ogni $v \in V$ non nullo.

Proposizione 3.1.3. *Un prodotto scalare definito positivo non è degenere.*

Dimostrazione. Ragioniamo per assurdo. Se fosse degenere, esisterebbe $v \in V$ non nullo tale che $\langle v, w \rangle = 0$ per ogni $w \in V$ ed in particolare avremmo $\langle v, v \rangle = 0$, contraddicendo l'ipotesi che $\langle v, v \rangle > 0$ per ogni v non nullo. \square

3.1.2. Prodotto scalare euclideo. Introduciamo un esempio importante.

Definizione 3.1.4. Il *prodotto scalare euclideo* su \mathbb{R}^n è definito come

$$\langle x, y \rangle = {}^t x \cdot y.$$

Nell'espressione ${}^t x \cdot y$ il vettore ${}^t x$ è un vettore riga e \cdot indica il prodotto fra matrici. In altre parole, se

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

allora

$$\langle x, y \rangle = (x_1, \dots, x_n) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \dots + x_n y_n.$$

Proposizione 3.1.5. *Il prodotto scalare euclideo è un prodotto scalare definito positivo su \mathbb{R}^n .*

Dimostrazione. Proviamo l'assioma (1):

$$\langle x + x', y \rangle = {}^t(x + x') \cdot y = ({}^t x + {}^t x') \cdot y = {}^t x \cdot y + {}^t x' \cdot y = \langle x, y \rangle + \langle x', y \rangle.$$

Gli altri assiomi di bilinearità si mostrano in modo simile. Per la simmetria, notiamo che l'espressione $\langle x, y \rangle = x_1 y_1 + \dots + x_n y_n$ è chiaramente simmetrica in x e y . Il prodotto scalare è definito positivo, perché se $x \neq 0$ allora

$$\langle x, x \rangle = x_1^2 + \dots + x_n^2 > 0.$$

La dimostrazione è completa. \square

3.1.3. Matrici simmetriche. Generalizziamo adesso l'esempio precedente. Ricordiamo che una matrice $S \in M(n)$ è *simmetrica* se ${}^tS = S$. Ricordiamo anche la formula

$${}^t(AB) = {}^tB {}^tA$$

che si applica per ogni prodotto di matrici AB . Più in generale, vale

$${}^t(A_1 \cdots A_k) = {}^tA_k \cdots {}^tA_1.$$

Proposizione 3.1.6. *Una matrice simmetrica S definisce un prodotto scalare su \mathbb{R}^n nel modo seguente:*

$$\langle x, y \rangle = {}^t_x S y$$

Dimostrazione. Notiamo innanzitutto che il prodotto di tre matrici ${}^t_x S y$ è ben definito e ha come risultato un numero. Verifichiamo l'assioma (1):

$$\langle x + x', y \rangle = {}^t(x + x') S y = {}^t_x S y + {}^t_{x'} S y = \langle x, y \rangle + \langle x', y \rangle.$$

Gli assiomi (2)-(4) si verificano in modo analogo. Per la simmetria:

$$\langle x, y \rangle = {}^t_x S y = {}^t({}^t_x S y) = {}^t_y {}^tS {}^t({}^t_x) = {}^t_y S x = \langle y, x \rangle.$$

La seconda uguaglianza è vera perché ${}^t_x S y$ è un numero e quindi è uguale al suo trasposto. Nella quarta abbiamo usato che $S = {}^tS$. \square

Ogni matrice simmetrica S definisce quindi un prodotto scalare su \mathbb{R}^n . Il prodotto scalare euclideo corrisponde alla matrice identità I , per cui

$$\langle x, y \rangle = {}^t_x I y = {}^t_x y.$$

3.1.4. Matrici diagonali. Le matrici diagonali sono matrici simmetriche molto semplici e facili da studiare. Consideriamo una matrice diagonale generica

$$D = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}$$

Il prodotto scalare $\langle x, y \rangle = {}^t_x D y$ può essere scritto semplicemente nel modo seguente

$$\langle x, y \rangle = d_1 x_1 y_1 + \dots + d_n x_n y_n.$$

In particolare è facile capire quando il prodotto scalare è definito positivo:

Proposizione 3.1.7. *Il prodotto scalare definito da D è definito positivo $\iff d_i > 0$ per ogni $i = 1, \dots, n$.*

Dimostrazione. Se $d_i > 0$ per ogni i allora otteniamo

$$\langle x, x \rangle = x_1^2 d_1 + \dots + x_n^2 d_n > 0$$

per ogni vettore non nullo x . Quindi il prodotto scalare è definito positivo. D'altra parte, se esiste un $d_i \leq 0$, allora

$$\langle e_i, e_i \rangle = d_i \leq 0$$

implica che il prodotto scalare non è definito positivo. \square

3.1.5. Vettori ortogonali e isotropi. Sia g un prodotto scalare su V . Due vettori $v, w \in V$ sono *ortogonali* se

$$\langle v, w \rangle = 0.$$

Può accadere che un vettore sia ortogonale a se stesso! Un vettore $v \in V$ tale che $\langle v, v \rangle = 0$ è detto *isotropo*.

Esempio 3.1.8. Consideriamo il prodotto scalare su \mathbb{R}^2 dato dalla matrice simmetrica

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

Vediamo subito che

$$\langle x, y \rangle = x_1y_1 - x_2y_2.$$

Quindi i vettori isotropi sono precisamente i vettori $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ tali che $x_1^2 - x_2^2 = 0$, cioè tali che $x_1 = \pm x_2$. I vettori isotropi formano quindi le due bisettrici dei quattro quadranti del piano cartesiano. In particolare notiamo che i vettori isotropi *non* formano un sottospazio vettoriale.

3.1.6. Radicale. Sia g un prodotto scalare su V . Il *radicale* di g è l'insieme dei vettori $v \in V$ che sono ortogonali a tutti gli altri. In altre parole, un vettore v è nel radicale se $\langle v, w \rangle = 0$ per *qualsiasi* vettore $w \in V$.

Il radicale di V è generalmente indicato con V^\perp . In simboli:

$$V^\perp = \{v \in V \mid \langle v, w \rangle = 0 \forall w \in V\}.$$

Notiamo subito che se $v \in V^\perp$ allora v è anche isotropo, ma non è vero il viceversa! Un vettore v isotropo non è necessariamente nel radicale: un vettore v può essere ortogonale a se stesso senza però essere ortogonale anche a tutti gli altri.

Proposizione 3.1.9. *Il radicale V^\perp è un sottospazio vettoriale di V .*

Dimostrazione. Se $v, v' \in V^\perp$, allora per ogni $w \in V$ otteniamo

$$\langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle = 0 + 0 = 0$$

e quindi anche $v + v' \in V^\perp$. La moltiplicazione per scalare è analoga. \square

Osservazione 3.1.10. Per definizione il prodotto scalare g è non degenere se e solo se $V^\perp = \{0\}$.

Consideriamo il caso in cui S sia una matrice simmetrica, che definisce quindi un prodotto scalare su \mathbb{R}^n tramite la formula

$$\langle x, y \rangle = {}^t x S y.$$

In questo caso è facile determinare il radicale.

Proposizione 3.1.11. *Il radicale di g è $\ker S$. Quindi il prodotto scalare è degenere $\iff \det S = 0$.*

Dimostrazione. Se $y \in \ker S$, allora per ogni $x \in \mathbb{R}^n$ abbiamo

$$\langle x, y \rangle = {}^t x S y = 0$$

perché $Sy = 0$. D'altra parte, se un vettore $y \in \mathbb{R}^n$ è tale che ${}^t x S y = 0$ per ogni $x \in \mathbb{R}^n$, allora scriviamo $y' = S y$ e otteniamo

$${}^t x y' = 0$$

per ogni $x \in \mathbb{R}^n$. Questo implica facilmente che $y' = 0$, cioè che $y \in \ker S$. \square

Esempio 3.1.12. Il prodotto scalare su \mathbb{R}^2 definito dalla matrice $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ considerato nell'Esempio 3.1.8 è un esempio importante da tenere a mente. Non è degenere perché $\det \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \neq 0$. Non è neppure definito positivo per la Proposizione 3.1.7. Il radicale è banale (perché è non degenere), ma ci sono comunque dei vettori isotropi (le due bisettrici dei quadranti). È quindi possibile che in un prodotto scalare un vettore sia ortogonale a se stesso ma non a tutti gli altri (cioè è isotropo ma non sta nel radicale).

3.1.7. Altri esempi. Introduciamo alcuni prodotti scalari in spazi vettoriali V diversi da \mathbb{R}^n .

Esempio 3.1.13. Un importante prodotto scalare sullo spazio $\mathbb{R}[x]$ dei polinomi è definito nel modo seguente:

$$\langle p, q \rangle = \int_0^1 p(t)q(t) dt$$

La bilinearità si verifica facilmente e la simmetria è immediata. Il prodotto scalare è definito positivo perché

$$\langle p, p \rangle = \int_0^1 p(t)^2 dt > 0$$

per ogni polinomio $p(x)$ non nullo.

È spesso utile definire un prodotto scalare anche sugli spazi di matrici:

Esercizio 3.1.14. Il prodotto scalare su $M(n)$ definito da

$$\langle A, B \rangle = \text{tr}(B {}^t A)$$

è definito positivo.

Esempio 3.1.15. Sia $\mathbb{R}_2[x]$ lo spazio dei polinomi di grado ≤ 2 . Scriviamo

$$\langle p, q \rangle = p(0)q(0) + p(1)q(1) + p(2)q(2).$$

Si verifica facilmente che anche questo è un prodotto scalare su $\mathbb{R}_2[x]$. Il prodotto scalare è definito positivo: infatti

$$\langle p, p \rangle = p(0)^2 + p(1)^2 + p(2)^2 > 0$$

per qualsiasi polinomio $p(x)$ non nullo di grado ≤ 2 : infatti un tale polinomio non può annullarsi in tre valori diversi 0, 1, 2. D'altra parte, il prodotto scalare

$$\langle p, q \rangle = p(0)q(0) + p(1)q(1)$$

non è definito positivo: se prendiamo $p(x) = x(1-x)$ otteniamo $\langle p, p \rangle = 0$. In particolare $p(x)$ è isotropo.

3.1.8. Matrice associata. Come per le applicazioni lineari, è possibile codificare i prodotti scalari in modo efficiente usando le matrici, dopo aver fissato delle basi.

Sia V uno spazio vettoriale reale e

$$g: V \times V \longrightarrow \mathbb{R}$$

un prodotto scalare: scriviamo $g(v, w)$ invece di $\langle v, w \rangle$. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base per V .

Definizione 3.1.16. La *matrice associata* a g nella base \mathcal{B} è la matrice simmetrica S il cui elemento S_{ij} è dato da

$$S_{ij} = g(v_i, v_j).$$

La matrice S è indicata a volte con $[g]_{\mathcal{B}}$.

Esempio 3.1.17. Se S è una matrice simmetrica che induce il prodotto scalare $g(x, y) = {}^t x S y$ su \mathbb{R}^n , allora la matrice associata a g nella base canonica di \mathbb{R}^n è proprio S stessa. Infatti $S_{ij} = {}^t e_i S e_j = g(e_i, e_j)$.

Esempio 3.1.18. Consideriamo il prodotto scalare euclideo g su \mathbb{R}^2 . Se scegliamo un'altra base

$$\mathcal{B} = \left\{ v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

verifichiamo allora che la matrice associata a g rispetto a \mathcal{B} è

$$S = [g]_{\mathcal{B}} = \begin{pmatrix} g(v_1, v_1) & g(v_1, v_2) \\ g(v_2, v_1) & g(v_2, v_2) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Esercizio 3.1.19. Scrivere la matrice associata al prodotto scalare su $\mathbb{R}_2[x]$

$$g(p, q) = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$$

rispetto alla base canonica $\mathcal{C} = \{1, x, x^2\}$.

Perché nella matrice associata scegliamo di mettere i numeri $S_{ij} = g(v_i, v_j)$? Il motivo principale è che questi numeri sono sufficienti per calcolare il prodotto scalare $g(v, w)$ di due vettori qualsiasi. Siano

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n, \quad w = \mu_1 v_1 + \dots + \mu_n v_n$$

due vettori arbitrari.

Proposizione 3.1.20. *Vale*

$$g(v, w) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j g(v_i, v_j).$$

Dimostrazione. Scriviamo

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

Usando la bilinearità di g otteniamo

$$\begin{aligned} g(v, w) &= g(\lambda_1 v_1 + \dots + \lambda_n v_n, \mu_1 v_1 + \dots + \mu_n v_n) \\ &= \sum_{i=1}^n g(\lambda_i v_i, \mu_1 v_1 + \dots + \mu_n v_n) = \sum_{i=1}^n \sum_{j=1}^n g(\lambda_i v_i, \mu_j v_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j g(v_i, v_j). \end{aligned}$$

La dimostrazione è conclusa. □

Modifichiamo leggermente l'enunciato della proposizione precedente:

Proposizione 3.1.21. *Per ogni coppia di vettori $v, w \in V$ vale l'uguaglianza*

$$g(v, w) = {}^t[v]_{\mathcal{B}} \cdot S \cdot [w]_{\mathcal{B}}.$$

Dimostrazione. Per la Proposizione 3.1.20 abbiamo

$$g(v, w) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j S_{ij}$$

Notiamo inoltre che

$$[v]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}, \quad [w]_{\mathcal{B}} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}.$$

Operando con il prodotto riga per colonna si verifica che effettivamente

$$(\lambda_1 \quad \dots \quad \lambda_n) \cdot S \cdot \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j S_{ij}.$$

Questo conclude la dimostrazione. □

In altre parole, se scriviamo $x = [v]_{\mathcal{B}}$ e $y = [w]_{\mathcal{B}}$, vediamo che

$$g(v, w) = {}^t x S y.$$

Abbiamo quindi scoperto un fatto importante: dopo aver scelto una base, qualsiasi prodotto scalare g su qualsiasi spazio vettoriale V di dimensione n può essere interpretato come un prodotto scalare su \mathbb{R}^n del tipo $\langle x, y \rangle = {}^t x S y$. È sufficiente sostituire i vettori v e w con le loro coordinate x e y , ed usare la matrice associata S .

3.1.9. Cambiamento di base. Studiamo adesso come cambia la matrice associata ad un prodotto scalare se sostituiamo la base $\mathcal{B} = \{v_1, \dots, v_n\}$ con una nuova base $\mathcal{B}' = \{v'_1, \dots, v'_n\}$. Ricordiamo che $M = [\text{id}]_{\mathcal{B}'}^{\mathcal{B}}$ è la matrice di cambiamento di base. Siano S e S' le matrici associate al prodotto scalare g nelle basi \mathcal{B} e \mathcal{B}' .

Proposizione 3.1.22. *Vale la relazione seguente:*

$$S' = {}^t M S M.$$

Dimostrazione. Per ogni vettore $v \in V$ ricordiamo che $[v]_{\mathcal{B}} = M \cdot [v]_{\mathcal{B}'}$ ed in particolare $[v'_i]_{\mathcal{B}} = M \cdot e_i$ è la i -esima colonna M^i di M . Otteniamo

$$S'_{ij} = g(v'_i, v'_j) = {}^t [v'_i]_{\mathcal{B}} \cdot S \cdot [v'_j]_{\mathcal{B}} = {}^t M^i \cdot S \cdot M^j.$$

Questo vale per ogni i, j , quindi $S' = {}^t M S M$. \square

Esempio 3.1.23. Abbiamo già notato che la matrice associata al prodotto scalare euclideo di \mathbb{R}^2 rispetto alla base

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

è $S = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. La matrice di cambiamento di base con la base canonica $\mathcal{C} = \{e_1, e_2\}$ è $M = [\text{id}]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. La Proposizione 3.1.22 ci dice che $S = {}^t M I M$, e infatti verifichiamo che $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Definizione 3.1.24. Due matrici simmetriche S e S' sono *congruenti* se esiste una matrice invertibile M per cui $S = {}^t M S' M$.

Esercizio 3.1.25. La congruenza fra matrici simmetriche è una relazione di equivalenza.

Si deve stare attenti a non confondere questa relazione di equivalenza con la similitudine, dove al posto della trasposta ${}^t M$ sta l'inversa M^{-1} .

Esempio 3.1.26. Le matrici

$$S = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad S' = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

non sono simili (ad esempio, perché hanno determinanti diversi), ma sono congruenti: prendiamo $M = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{pmatrix}$ e verifichiamo facilmente che $S' = {}^t M S M$.

Abbiamo implicitamente dimostrato questo fatto.

Proposizione 3.1.27. *Sue matrici sono congruenti \iff rappresentano lo stesso prodotto scalare con basi diverse.*

Due matrici simili S e S' hanno lo stesso determinante, due matrici congruenti no! Possiamo comunque dimostrare quanto segue.

Proposizione 3.1.28. *Se S e S' sono congruenti, allora $\det S$ e $\det S'$ hanno lo stesso segno, cioè $\det S$ è positivo, nullo, o negativo $\iff \det S'$ è positivo, nullo, o negativo.*

Dimostrazione. Per ipotesi esiste una M con $\det M \neq 0$ tale che $S' = {}^tMSM$ e quindi usando il teorema di Binet troviamo

$$\det S' = \det({}^tM) \det S \det M = \det M \det S \det M = (\det M)^2 \det S.$$

La proposizione è dimostrata perché $(\det M)^2 > 0$. □

3.1.10. Restrizioni. Sia V dotato di un prodotto scalare g . Restringendo il prodotto scalare g ai soli vettori di un sottospazio $W \subset V$, si ottiene un prodotto scalare per lo spazio W che possiamo indicare con $g|_W$. Il prodotto scalare $g|_W$ è la *restrizione* di g a W .

Proposizione 3.1.29. *Notiamo che:*

- (1) *se g è definito positivo, anche $g|_W$ è definito positivo;*
- (2) *se g è non degenere, non è detto che $g|_W$ sia non degenere!*

Dimostrazione. (1). Se $g(v, v) > 0$ per ogni $v \in V$ non nullo, lo stesso vale per ogni $v \in W$ non nullo.

(2) Qui dobbiamo fornire un controesempio. Prendiamo $V = \mathbb{R}^2$ con il prodotto scalare $\langle x, y \rangle = {}^txSy$ determinato dalla matrice

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

Sappiamo che S è non degenere. Definiamo quindi W come la retta $W = \text{Span}(v)$ con $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Il prodotto scalare $g|_W$ è il prodotto nullo, perché v è isotropo: quindi per ogni coppia di vettori $\lambda v, \mu v \in W$ otteniamo

$$g(\lambda v, \mu v) = \lambda \mu g(v, v) = 0.$$

In particolare la restrizione $g|_W$ è degenere. □

Alla fine della dimostrazione abbiamo anche notato il fatto seguente.

Proposizione 3.1.30. *Se $W = \text{Span}(v)$, la restrizione $g|_W$ è degenere $\iff v$ è un vettore isotropo.*

3.1.11. Sottospazio ortogonale. Sia V dotato di un prodotto scalare. Sia $W \subset V$ un sottospazio. Il *sottospazio ortogonale* W^\perp è per definizione l'insieme

$$W^\perp = \{v \in V \mid \langle v, w \rangle = 0 \forall w \in W\}.$$

In altre parole, l'insieme W^\perp è formato da quei vettori $v \in V$ che sono ortogonali a qualsiasi vettore di W . Notiamo che questa definizione è coerente con la notazione V^\perp già usata per definire il radicale, che come sappiamo è l'insieme dei vettori ortogonali a qualsiasi vettore di tutto lo spazio V .

Proposizione 3.1.31. *L'insieme W^\perp è effettivamente un sottospazio di V .*

Dimostrazione. Come di consueto dobbiamo mostrare tre cose:

- (1) $0 \in W^\perp$, infatti $\langle 0, w \rangle = 0$ sempre;
- (2) se $v, v' \in W^\perp$ allora $v + v' \in W^\perp$, infatti

$$\langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle = 0 + 0 = 0$$

per ogni $w \in W$;

- (3) se $v \in W^\perp$ allora $\lambda v \in W^\perp$, e questo è lasciato per esercizio.

La dimostrazione è completa. \square

Per determinare W^\perp concretamente è molto utile il fatto seguente.

Proposizione 3.1.32. *Se $W = \text{Span}(w_1, \dots, w_k)$, allora*

$$W^\perp = \{v \in V \mid \langle v, w_i \rangle = 0 \forall i = 1, \dots, k\}.$$

Dimostrazione. Se $v \in W^\perp$ allora $\langle v, w \rangle = 0$ per ogni $w \in W$ e quindi in particolare $\langle v, w_i \rangle = 0$ per ogni $i = 1, \dots, k$.

D'altro canto, se $\langle v, w_i \rangle = 0$ per ogni $i = 1, \dots, k$ allora per ogni $w \in W$ possiamo scrivere $w = \lambda_1 w_1 + \dots + \lambda_k w_k$ e quindi otteniamo

$$\langle v, w \rangle = \langle v, \lambda_1 w_1 + \dots + \lambda_k w_k \rangle = \lambda_1 \langle v, w_1 \rangle + \dots + \lambda_k \langle v, w_k \rangle = 0.$$

Quindi $v \in W^\perp$. \square

Concretamente, l'insieme W^\perp può essere descritto come l'insieme dei vettori v che soddisfano le k relazioni $\langle v, w_i \rangle = 0$ con $i = 1, \dots, k$.

3.1.12. Dimensione del sottospazio ortogonale. Se W è un sottospazio vettoriale, ci chiediamo adesso in che relazione siano i due sottospazi W e W^\perp . Sia $n = \dim V$ e g il prodotto scalare.

Proposizione 3.1.33. *Valgono i fatti seguenti:*

- (1) $\dim W^\perp \geq n - \dim W$;
- (2) se g è non-degenere, allora $\dim W^\perp = n - \dim W$;
- (3) se $g|_W$ è non-degenere, allora $V = W \oplus W^\perp$;
- (4) se g è definita positiva, allora $V = W \oplus W^\perp$.

Dimostrazione. Sia w_1, \dots, w_k una base di W . Un vettore $v \in V$ è ortogonale ad ogni vettore di W se (e solo se) è ortogonale a ciascun vettore della base w_1, \dots, w_k . Consideriamo l'applicazione lineare $T: V \rightarrow \mathbb{R}^k$ definita nel modo seguente:

$$T(v) = \begin{pmatrix} \langle v, w_1 \rangle \\ \vdots \\ \langle v, w_k \rangle \end{pmatrix}.$$

Lo spazio W^\perp è precisamente il nucleo di T . Per il teorema della dimensione

$$\dim W^\perp = \dim \ker T = n - \dim \operatorname{Im} T \geq n - k$$

infatti $\dim \operatorname{Im} T \leq k$, perché $\operatorname{Im} T \subset \mathbb{R}^k$. Questo mostra il punto (1), quindi ora passiamo a mostrare (2). Completiamo w_1, \dots, w_k a base w_1, \dots, w_n di V e consideriamo come sopra l'applicazione lineare $T': V \rightarrow \mathbb{R}^n$ data da

$$T'(v) = \begin{pmatrix} \langle v, w_1 \rangle \\ \vdots \\ \langle v, w_n \rangle \end{pmatrix}.$$

Siccome g non è degenere, la mappa T' è iniettiva. Poiché $\dim V = \dim \mathbb{R}^n = n$, la mappa T' è anche suriettiva. Questo implica facilmente che anche T è suriettiva. Quindi $\dim \operatorname{Im} T = k$ e $\dim W^\perp = n - k$. Il punto (2) è dimostrato.

Passiamo a dimostrare il punto (3). Se $g|_W$ è non degenere, nessun vettore non banale di W è ortogonale a tutti i vettori di W e quindi $W \cap W^\perp = \{0\}$. Per il punto (1) sappiamo che $\dim W + \dim W^\perp \geq n$, quindi per la formula di Grassmann deduciamo che $\dim(W + W^\perp) \geq \dim W + \dim W^\perp \geq n$, quindi $V = W + W^\perp$ e allora $V = W \oplus W^\perp$.

Se g è definito positivo, ogni restrizione $g|_W$ è definita positiva ed in particolare non è degenere, quindi (3) implica (4). \square

Esercizio 3.1.34. Sia V dotato di un prodotto scalare g e $U, W \subset V$ due sottospazi. Valgono i fatti seguenti:

- se $U \subset W$ allora $W^\perp \subset U^\perp$;
- $(U^\perp)^\perp \supset U$, con l'uguaglianza se g è non-degenere;
- $(U + W)^\perp = U^\perp \cap W^\perp$.

3.1.13. Esempi. Facciamo qualche esempio di calcolo di un sottospazio ortogonale. Iniziamo con il prodotto scalare euclideo su \mathbb{R}^n .

Un sottospazio $U \subset \mathbb{R}^n$ è descritto generalmente in forma parametrica, come $U = \operatorname{Span}(v_1, \dots, v_k)$, oppure in forma cartesiana, come $U = \ker A = \{x \mid Ax = 0\}$. In entrambi i casi è facile descrivere lo spazio ortogonale U^\perp rispetto al prodotto scalare euclideo. Questo è rispettivamente

$$U^\perp = \{x \mid \langle x, v_1 \rangle = \dots = \langle x, v_k \rangle = 0\}$$

oppure

$$U^\perp = \operatorname{Span}({}^tA_1, \dots, {}^tA_h)$$

dove A_1, \dots, A_h sono le righe di A .

Esempio 3.1.35. Se $U = \{ax + by + cz = 0\}$ è un piano in \mathbb{R}^3 , la retta ortogonale U^\perp è semplicemente

$$U^\perp = \text{Span} \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

D'altra parte, se U è la retta

$$U = \text{Span} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

allora U^\perp è il piano $U^\perp = \{ax + by + cz = 0\}$. Se

$$U = \begin{cases} ax + by + cz = 0, \\ dx + ey + fz = 0, \end{cases}$$

allora

$$U^\perp = \text{Span} \left(\begin{pmatrix} a \\ b \\ c \end{pmatrix}, \begin{pmatrix} d \\ e \\ f \end{pmatrix} \right).$$

3.1.14. Tipi di prodotti scalari. Diciamo che un prodotto scalare su V è

- *semi-definito positivo* se $\langle v, v \rangle \geq 0$ per ogni $v \in V$;
- *definito negativo* se $\langle v, v \rangle < 0$ per ogni $v \in V$ non nullo;
- *semi-definito negativo* se $\langle v, v \rangle \leq 0$ per ogni $v \in V$;
- *indefinito* se esistono v e w per cui $\langle v, v \rangle > 0$ e $\langle w, w \rangle < 0$.

Consideriamo una matrice diagonale D , con elementi d_1, \dots, d_n sulla diagonale.

Esercizio 3.1.36. Il prodotto scalare definito da D è

- semi-definito positivo $\iff d_i \geq 0$ per ogni i ;
- definito negativo $\iff d_i < 0$ per ogni i ;
- semi-definito negativo $\iff d_i \leq 0$ per ogni i ;
- indefinito \iff esistono i e j con $d_i > 0$ e $d_j < 0$.

3.2. Classificazione dei prodotti scalari

Lo scopo principale di questa sezione è ottenere una classificazione completa dei prodotti scalari.

3.2.1. Basi ortogonali e ortonormali. Sia g un prodotto scalare su uno spazio vettoriale V di dimensione n .

Definizione 3.2.1. Una base $\mathcal{B} = \{v_1, \dots, v_n\}$ è *ortogonale* se $g(v_i, v_j) = 0$ per ogni $i \neq j$. In altre parole, una base \mathcal{B} è ortogonale se due vettori distinti di \mathcal{B} sono sempre ortogonali fra loro.

Una base ortogonale è *ortonormale* se per ciascun i il numero $g(v_i, v_i)$ è 1, 0 oppure -1 .

Sia S la matrice associata a g nella base \mathcal{B} . Segue immediatamente dalla definizione che \mathcal{B} è ortogonale $\iff S$ è una matrice diagonale. Ricordiamo infatti che $S_{ij} = g(v_i, v_j)$, quindi chiedere che $g(v_i, v_j) = 0$ per ogni $i \neq j$ è come chiedere che S sia diagonale.

D'altra parte \mathcal{B} è ortonormale \iff la matrice S è diagonale, e inoltre gli elementi sulla diagonale sono soltanto di tre tipi: 1, 0 e -1 .

Esempio 3.2.2. La base canonica è una base ortonormale per il prodotto scalare euclideo. Infatti $\langle e_i, e_j \rangle = {}^t e_i e_j = 0$ per ogni $i \neq j$ e inoltre $\langle e_i, e_i \rangle = 1$ per ogni i .

Esempio 3.2.3. Consideriamo il prodotto scalare su \mathbb{R}^2 definito dalla matrice $S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Notiamo che

$$(x_1 \ y_1) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = x_2 y_1 + x_1 y_2.$$

La base $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ è una base ortogonale per il prodotto scalare, perché ${}^t v_1 S v_2 = 0$. Questa ovviamente non è l'unica base ortogonale possibile, ce ne sono molte altre: ad esempio anche $w_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $w_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ è una base ortogonale, perché ${}^t w_1 S w_2 = 0$.

3.2.2. Diagonalizzazione dei prodotti scalari. Sia V uno spazio vettoriale di dimensione n e g un prodotto scalare su V . Dimostriamo adesso il fatto seguente.

Teorema 3.2.4 (Teorema di Sylvester). *Esiste sempre una base ortogonale $\mathcal{B} = \{v_1, \dots, v_n\}$ per V .*

Dimostrazione. Lavoriamo per induzione su $n = \dim V$. Se $n = 1$, una base di V è formata da un vettore solo ed è sempre ortogonale. Supponiamo che l'enunciato sia valido per $n - 1$ e lo mostriamo per n .

Supponiamo che esista almeno un vettore $v \in V$ non isotropo. Consideriamo la retta vettoriale $W = \text{Span}(v)$. La restrizione di g a W è non-degenere perché $g(v, v) \neq 0$. Quindi la Proposizione 3.1.33-(3) ci dice che $V = W \oplus W^\perp$.

Lo spazio W^\perp ha dimensione $n - 1$ e quindi possiamo applicare l'ipotesi induttiva al prodotto scalare $g|_{W^\perp}$. Per ipotesi induttiva esiste una base ortogonale v_2, \dots, v_n per W^\perp . Aggiungendo v alla sequenza otteniamo una base

$\mathcal{B} = \{v, v_2, \dots, v_n\}$ per V che è ancora ortogonale, infatti $v_i \in W^\perp$ implica $g(v, v_i) = 0$ per ogni i .

Resta da considerare il caso in cui tutti i vettori in V siano isotropi, cioè $g(v, v) = 0$ per ogni $v \in V$. La facile uguaglianza

$$g(v, w) = \frac{g(v+w, v+w) - g(v, v) - g(w, w)}{2} = \frac{0 - 0 - 0}{2} = 0$$

ci mostra che $g(v, w) = 0$ per ogni $v, w \in V$ e quindi qualsiasi base di V è ortogonale. \square

Ricordiamo che se \mathcal{B} è ortogonale allora la matrice associata a g nella base \mathcal{B} è diagonale.

Corollario 3.2.5. *Qualsiasi matrice simmetrica S è congruente ad una matrice diagonale.*

Dimostrazione. Interpretiamo S come un prodotto scalare g in \mathbb{R}^n . Per il Teorema di Sylvester esiste una base ortogonale \mathcal{B} , quindi $S' = [g]_{\mathcal{B}}$ è diagonale. Un cambiamento di base trasforma la matrice per congruenza: quindi S è congruente alla matrice diagonale S' . \square

3.2.3. Normalizzazione. Descriviamo adesso un semplice algoritmo di *normalizzazione*, che trasforma una base ortogonale in una base ortonormale.

Se $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base ortogonale per un certo prodotto scalare, possiamo sostituire ciascun vettore v_i con un nuovo vettore w_i definito nel modo seguente:

$$w_i = \begin{cases} \frac{v_i}{\sqrt{\langle v_i, v_i \rangle}} & \text{se } \langle v_i, v_i \rangle > 0, \\ v_i & \text{se } \langle v_i, v_i \rangle = 0, \\ \frac{v_i}{\sqrt{-\langle v_i, v_i \rangle}} & \text{se } \langle v_i, v_i \rangle < 0. \end{cases}$$

Il risultato è una nuova base $\{w_1, \dots, w_n\}$ in cui il numero $\langle w_i, w_i \rangle$ può essere solo 1, -1 , oppure 0, per ogni i . Il lettore è invitato a verificare questo fatto.

Notiamo inoltre che possiamo anche riordinare gli elementi di questa base in modo che prima ci siano i vettori w_i con $\langle w_i, w_i \rangle = 1$, poi quelli con $\langle w_i, w_i \rangle = -1$, ed infine quelli con $\langle w_i, w_i \rangle = 0$. La matrice associata $S = [g]_{\mathcal{B}}$ a questo punto diventa del tipo

$$S = \begin{pmatrix} I_k & 0 & 0 \\ 0 & -I_h & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Riassumiamo questa indagine con un enunciato. Sia come sempre V uno spazio vettoriale di dimensione n e g un prodotto scalare su V .

Proposizione 3.2.6. *Esiste sempre una base ortonormale \mathcal{B} per V tale che la matrice associata $S = [g]_{\mathcal{B}}$ sia del tipo*

$$(4) \quad S = \begin{pmatrix} I_k & 0 & 0 \\ 0 & -I_h & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Dimostrazione. Per il Teorema di Sylvester esiste una base ortogonale, che può quindi essere trasformata in ortonormale tramite la normalizzazione. \square

Definizione 3.2.7. La terna di numeri $(k, h, n - (k + h))$ è detta *segnatura* del prodotto scalare g ed è di solito indicata con i simboli (i_+, i_-, i_0) . I numeri i_+, i_-, i_0 sono chiamati *indice di positività*, *indice di negatività* e *indice di nullità* di g . La somma dei tre indici è $n = \dim V$.

Corollario 3.2.8. *Qualsiasi matrice simmetrica è congruente ad una matrice diagonale S come in (4).*

3.2.4. Classificazione. Nella sezione precedente abbiamo definito la segnatura (i_+, i_-, i_0) di un prodotto scalare, barando un po' verso la fine: dobbiamo ancora mostrare che i tre numeri dipendono solo da g e non dalla particolare base \mathcal{B} scelta per diagonalizzare il prodotto scalare. Lo facciamo con la proposizione seguente, che caratterizza i numeri i_+, i_-, i_0 più intrinsecamente e quindi fornisce un'utile informazione aggiuntiva.

Proposizione 3.2.9. *Valgono i fatti seguenti:*

- $i_+ = \max\{\dim W \mid W \subset V \text{ sottospazio con } g|_W \text{ definito positivo}\}$
- $i_- = \max\{\dim W \mid W \subset V \text{ sottospazio con } g|_W \text{ definito negativo}\}$
- $i_0 = \dim V^\perp$

Dimostrazione. Per la Proposizione 3.2.6 esiste una base ortogonale $\mathcal{B} = \{v_1, \dots, v_n\}$ per cui la matrice associata $S = [g]_{\mathcal{B}}$ è come in (4). Consideriamo i sottospazi

$$W_+ = \text{Span}(v_1, \dots, v_k), \quad W_- = \text{Span}(v_{k+1}, \dots, v_{k+h}), \\ W_0 = \text{Span}(v_{k+h+1}, \dots, v_n).$$

La Proposizione 3.1.11 ci dice che W_0 è proprio il radicale, quindi $i_0 = n - (k + h) = \dim V^\perp$.

Siano k_0 e h_0 le massime dimensioni possibili di un sottospazio W su cui la restrizione $g|_W$ sia rispettivamente definita positiva e negativa. Dobbiamo mostrare che $k_0 = k$ e $h_0 = h$.

Le restrizioni $g|_{W_+}$ e $g|_{W_-}$ sono definite positive e negative rispettivamente, quindi $k_0 \geq k$ e $h_0 \geq h$. Supponiamo per assurdo che esista un'altro sottospazio W di dimensione $k' > k$ su cui la restrizione $g|_W$ è definita positiva. Notiamo che

$$\dim W + \dim(W_- \oplus W_0) = k' + (n - k) > n$$

e quindi per la formula di Grassman l'intersezione $W \cap (W_- \oplus W_0)$ ha dimensione almeno uno. Quindi esiste un $v \neq 0$ con $v \in W$ (e quindi $\langle v, v \rangle > 0$) e $v \in W_- \oplus W_0$ (e quindi $\langle v, v \rangle \leq 0$): assurdo. Il caso definito negativo è analogo. \square

La segnatura (i_+, i_-, i_0) classifica completamente i prodotti scalari. Molte delle definizioni che abbiamo dato possono essere reinterpretate con la segnatura:

- g è definito positivo $\iff (i_+, i_-, i_0) = (n, 0, 0)$;
- g è definito negativo $\iff (i_+, i_-, i_0) = (0, n, 0)$;
- g è degenere $\iff i_0 > 0$;
- g è semi-definito positivo $\iff i_- = 0$;
- g è semi-definito negativo $\iff i_+ = 0$;
- g è indefinito $\iff i_+ > 0$ e $i_- > 0$.

Il lettore è invitato a dimostrare tutte queste facili equivalenze. Definiamo la segnatura (i_+, i_-, i_0) di una matrice simmetrica S come la segnatura del prodotto scalare $\langle x, y \rangle = {}^t x S y$ su \mathbb{R}^n . Otteniamo in particolare una classificazione completa delle matrici simmetriche a meno di congruenza:

Corollario 3.2.10. Due matrici simmetriche sono congruenti se e solo se hanno la stessa segnatura.

Calcolare la segnatura di una matrice diagonale è immediato.

Esercizio 3.2.11. Sia D una matrice diagonale e (i_+, i_-, i_0) la sua segnatura. I numeri i_+, i_-, i_0 sono pari al numero di elementi positivi, negativi e nulli sulla diagonale di D .

Per una matrice simmetrica S non diagonale la questione è più complessa.

3.2.5. Calcolo della segnatura. Abbiamo dimostrato che la segnatura (i_+, i_-, i_0) caratterizza un prodotto scalare, quindi è naturale adesso porsi il problema seguente: come si calcola la segnatura di un prodotto scalare dato?

Un metodo consiste nell'applicare la dimostrazione della Proposizione 3.2.4 per costruire una base ortogonale, scrivere la matrice diagonale associata, quindi contare quanti elementi sulla diagonale sono positivi, negativi, o nulli. Il problema con questo approccio è che la costruzione della base ortogonale può essere laboriosa. Cerchiamo quindi un modo di determinare la segnatura senza dover costruire una base ortogonale.

In seguito descriveremo alcuni metodi generali per risolvere questo problema, per adesso ci accontentiamo di fare alcune considerazioni. Sia g un prodotto scalare su V e S la matrice associata a g rispetto ad una base $\mathcal{B} = \{v_1, \dots, v_n\}$ qualsiasi. Abbiamo visto nella Proposizione 3.1.28 che il segno di $\det S$ non dipende dalla base \mathcal{B} . Poiché S è congruente ad una matrice diagonale del tipo (4), deduciamo che:

- $\det S = 0 \iff i_0 > 0$;
- se $\det S > 0$, allora i_- è pari,
- se $\det S < 0$, allora i_- è dispari.

Notiamo anche un'altra cosa. Se $W \subset V$ è un sottospazio su cui la restrizione $g|_W$ è definita positiva, allora $i_+ \geq \dim W$ per la Proposizione 3.2.9. In particolare, se $S_{ii} > 0$ per qualche i , allora $\langle v_i, v_i \rangle = S_{ii} > 0$ e quindi la restrizione $g|_{\text{Span}(v_i)}$ sulla retta $\text{Span}(v_i)$ è definita positiva: ne segue che $i_+ > 0$. Riassumendo:

- se esiste i con $S_{ii} > 0$, allora $i_+ > 0$;
- se esiste i con $S_{ii} < 0$, allora $i_- > 0$.

Queste informazioni sono sufficienti per trattare il caso bidimensionale. Consideriamo il prodotto scalare su \mathbb{R}^2 definito come $\langle x, y \rangle = {}^t x S y$ per una matrice simmetrica

$$S = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

Notiamo che $\det S = ac - b^2$. Questa è la casistica:

- se $S = 0$, la segnatura è $(0, 0, 2)$;
- se $\det S = 0$ e $a > 0$ oppure $c > 0$, la segnatura è $(1, 0, 1)$;
- se $\det S = 0$ e $a < 0$ oppure $c < 0$, la segnatura è $(0, 1, 1)$;
- se $\det S > 0$ e $a > 0$ oppure $c > 0$, la segnatura è $(2, 0, 0)$;
- se $\det S > 0$ e $a < 0$ oppure $c < 0$, la segnatura è $(0, 2, 0)$;
- se $\det S < 0$, la segnatura è $(1, 1, 0)$.

Il lettore è invitato a dimostrare i vari casi usando la Proposizione 3.2.9.

Osservazione 3.2.12. Attenzione: un elemento positivo sulla diagonale di S implica $i_+ \geq 1$, ma due elementi positivi non implicano $i_+ \geq 2$. Ad esempio la segnatura di $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ è $(1, 1, 0)$.

Esempio 3.2.13. Calcoliamo la segnatura della matrice

$$S = \begin{pmatrix} 2 & 2 & 1 & 1 \\ 2 & 2 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

Notiamo che S ha rango due e quindi $i_0 = 2$. Le signature possibili sono:

$$(2, 0, 2), \quad (1, 1, 2), \quad (0, 2, 2).$$

Ci sono elementi positivi sulla diagonale, quindi $i_+ > 0$ e la terza possibilità è esclusa. Ci resta da capire quale delle prime due signature sia quella giusta. Se consideriamo il piano $W = \text{Span}(e_2, e_3)$, notiamo che la matrice associata alla restrizione $g|_W$ è la sottomatrice quadrata

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

Questa ha determinante negativo: quindi la segnatura di $g|_W$ è $(1, 1, 0)$. Ne deduciamo che W contiene una retta su cui g è definito positivo e una su cui g è definito negativo. Quindi $i_- \geq 1$ e l'unica segnatura possibile è $(1, 1, 2)$.

3.2.6. Criterio di Jacobi. Introduciamo un criterio semplice per determinare se una matrice $n \times n$ simmetrica S è definita positiva. Per ogni $i = 1, \dots, n$ sia d_i il determinante del minore S_i di S ottenuto considerando solo le prime i righe e le prime i colonne di S .

Lemma 3.2.14 (Criterio di Jacobi). *Una matrice S è definita positiva se e solo se $d_i > 0$ per ogni i .*

Dimostrazione. La matrice S definisce un prodotto scalare $g(x, y) = {}^t x S y$ su \mathbb{R}^n e S_i rappresenta la restrizione di g sul sottospazio coordinato $V_i = \text{Span}(e_1, \dots, e_i)$. Se S è definita positiva, anche $g|_{V_i}$ lo è, ed in particolare deve valere $d_i = \det S_i > 0$ per ogni i .

Mostriamo ora l'implicazione inversa: supponiamo che $d_i > 0$ per ogni i e deduciamo che g è definito positivo. Dimostriamo il lemma per induzione su n . Se $n = 1$, la matrice è un numero positivo d_1 e siamo a posto. Supponiamo il lemma vero per $n - 1$ e lo dimostriamo per n .

Per l'ipotesi induttiva la restrizione $g|_{V_{n-1}}$ è definita positiva. Consideriamo la segnatura (i_+, i_-, i_0) di g . Siccome la restrizione $g|_{V_{n-1}}$ è definita positiva, abbiamo $i_+ \geq n - 1$ e le possibilità per la segnatura sono

$$(n, 0, 0), \quad (n - 1, 1, 0), \quad (n - 1, 0, 1).$$

Siccome $d_n > 0$, l'unica possibile è la prima. □

Corollario 3.2.15. *Una matrice S è definita negativa se e solo se $d_i < 0$ per ogni i dispari e $d_i > 0$ per ogni i pari.*

Dimostrazione. La matrice S è definita negativa se e solo se $-S$ è definita positiva. A questo punto basta notare che $\det(-A) = (-1)^k \det A$ per ogni matrice A di taglia $k \times k$. □

Con una dimostrazione analoga si dimostra anche questo criterio:

Lemma 3.2.16. *Se $d_i \neq 0$ per ogni i , allora la segnatura di S è $(i_+, i_-, 0)$ dove i_+ è il numero di permanenze di segno nella successione $1, d_1, \dots, d_n$ e i_- il numero di cambiamenti di segno.*

Esercizio 3.2.17. Verifica il criterio appena enunciato quando S è una matrice diagonale invertibile (cioè senza zeri sulla diagonale).

Esempio 3.2.18. Studiamo la segnatura della matrice

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

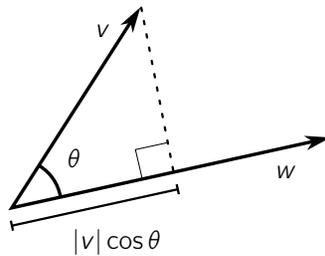


Figura 3.1. Una interpretazione geometrica del prodotto scalare fra due vettori v e w .

al variare di $a \in \mathbb{R}$. Otteniamo

$$d_1 = 1, \quad d_2 = 2, \quad d_3 = 6 - 2a^2.$$

Se $|a| > \sqrt{3}$ allora $d_3 < 0$ e quindi la segnatura è $(2, 1, 0)$. Se $|a| < \sqrt{3}$ allora $d_3 > 0$ e quindi la segnatura è $(3, 0, 0)$. Se $a = \pm\sqrt{3}$ il criterio di Jacobi non si applica e dobbiamo usare altri metodi. La matrice S ha rango 2 e quindi $i_0 = 3 - 2 = 1$. D'altra parte la restrizione di g al piano coordinato $W = \text{Span}(e_1, e_2)$ ha come matrice associata $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ che è definita positiva: quindi $g|_W$ è definita positiva e $i_+ \geq 2$ per la Proposizione 3.2.9. L'unico caso possibile è $(2, 0, 1)$. Riassumendo, la segnatura è

- $(3, 0, 0)$ per $|a| < \sqrt{3}$,
- $(2, 0, 1)$ per $|a| = \sqrt{3}$,
- $(2, 1, 0)$ per $|a| > \sqrt{3}$.

3.3. Prodotti scalari definiti positivi

Nella geometria euclidea che si studia alle scuole superiori, il prodotto scalare fra due vettori non è una nozione "di base" e viene generalmente definita usando angoli e lunghezze: il *prodotto scalare* di due vettori v e w che formano un angolo acuto θ come in Figura 3.1 è il numero

$$\langle v, w \rangle = |v| \cdot |w| \cos \theta.$$

In questa formula $|v|$ e $|w|$ sono le lunghezze di v e w . La Figura 3.1 mostra una interpretazione geometrica del prodotto scalare: se proiettiamo v ortogonalmente lungo w , otteniamo un segmento s lungo $|v| \cos \theta$. Quindi il prodotto scalare fra v e w è il prodotto delle lunghezze di s e di w .

Nella geometria che si insegna in ambito universitario si preferisce scegliere la strada opposta: si introduce prima il prodotto scalare (come abbiamo fatto nella sezione precedente) e poi si introducono angoli, lunghezze e distanze.

In tutto questo capitolo V sarà sempre uno spazio vettoriale reale dotato di un prodotto scalare definito positivo.

3.3.1. Norma. In queste pagine introduciamo i concetti di *norma* (o *lunghezza*) di un vettore, *angolo* fra vettori, e *distanza* fra punti.

Definizione 3.3.1. La *norma* di un vettore $v \in V$ è il numero reale

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

La norma di v , detta anche *modulo*, va interpretata come la *lunghezza* del vettore v . Nella definizione usiamo il fatto che il prodotto scalare è definito positivo e quindi ha senso la radice quadrata di $\langle v, v \rangle \geq 0$.

Proposizione 3.3.2. *Valgono le proprietà seguenti, per ogni $v, w \in V$ e ogni $\lambda \in \mathbb{R}$:*

- (1) $\|v\| > 0$ se $v \neq 0$ e $\|0\| = 0$;
- (2) $\|\lambda v\| = |\lambda| \|v\|$;
- (3) $|\langle v, w \rangle| \leq \|v\| \|w\|$;
- (4) $\|v + w\| \leq \|v\| + \|w\|$.

La disuguaglianza (3) è nota come *disuguaglianza di Cauchy-Schwarz*.

Dimostrazione. I primi due punti seguono semplicemente dalla definizione $\|v\| = \sqrt{\langle v, v \rangle}$. Il punto (1) è verificato perché il prodotto scalare è definito positivo, il (2) si ottiene così:

$$\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v \rangle} = \sqrt{\lambda^2 \langle v, v \rangle} = |\lambda| \sqrt{\langle v, v \rangle} = |\lambda| \|v\|.$$

Il punto (3) invece non è ovvio. Consideriamo due numeri $a, b \in \mathbb{R}$ e notiamo che per ogni $v, w \in V$ vale la disuguaglianza

$$\begin{aligned} 0 &\leq \|av + bw\|^2 = \langle av + bw, av + bw \rangle \\ &= a^2 \langle v, v \rangle + b^2 \langle w, w \rangle + 2ab \langle v, w \rangle = a^2 \|v\|^2 + b^2 \|w\|^2 + 2ab \langle v, w \rangle. \end{aligned}$$

Sostituendo $a = \|w\|^2$ e $b = -\langle v, w \rangle$ otteniamo

$$0 \leq \|w\|^4 \|v\|^2 + \langle v, w \rangle^2 \|w\|^2 - 2\|w\|^2 \langle v, w \rangle^2.$$

Dividendo per $\|w\|^2$ si ottiene

$$\|w\|^2 \|v\|^2 \geq \langle v, w \rangle^2$$

e la (3) è ottenuta prendendo le radici quadrate. La disuguaglianza (4) è un semplice corollario della (3):

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2\langle v, w \rangle \leq \|v\|^2 + \|w\|^2 + 2\|v\| \|w\| = (\|v\| + \|w\|)^2.$$

Anche qui basta prendere le radici per concludere. \square

Esercizio 3.3.3. Per ogni $v, w \in V$ otteniamo

$$\langle v, w \rangle = \frac{1}{4} (\|v + w\|^2 - \|v - w\|^2).$$

3.3.2. Angoli. A questo punto possiamo definire la nozione di angolo fra vettori non nulli.

Definizione 3.3.4. L'*angolo* fra due vettori $v, w \in V$ non nulli è il numero $\theta \in [0, \pi]$ per cui

$$\cos \theta = \frac{\langle v, w \rangle}{\|v\| \|w\|}.$$

Effettivamente per la disuguaglianza di Cauchy-Scharz il membro di destra è un numero $x \in [-1, 1]$ e quindi esiste un unico $\theta \in [0, \pi]$ per cui $\cos \theta = x$. In altre parole, si può usare la funzione *arcocoseno* e scrivere

$$\theta = \arccos \frac{\langle v, w \rangle}{\|v\| \|w\|}.$$

Osservazione 3.3.5. Tutte le funzioni trigonometriche come $\sin(x)$, $\cos(x)$, $\tan(x)$, $\arcsin(x)$, $\arccos(x)$, ecc. possono essere definite intrinsecamente come serie di potenze.

Notiamo che θ è acuto, retto o ottuso \iff il prodotto scalare $\langle v, w \rangle$ è rispettivamente positivo, nullo o negativo.

3.3.3. Distanze. Sia come sempre V uno spazio vettoriale reale dotato di un prodotto scalare definito positivo. Possiamo infine definire la distanza fra due punti di V .

Quando interpretiamo V in modo più geometrico, può essere utile indicare i punti di V con le lettere maiuscole P, Q, R , eccetera. Possiamo anche definire il vettore

$$\overrightarrow{PQ} = Q - P.$$

Notiamo in particolare che

$$\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}.$$

Definizione 3.3.6. La *distanza* fra due punti $p, q \in V$ è il numero reale

$$d(P, Q) = \|\overrightarrow{PQ}\| = \|Q - P\|.$$

La distanza ha tre proprietà basilari.

Proposizione 3.3.7. *Valgono i fatti seguenti per ogni $P, Q, R \in V$:*

- (1) $d(P, Q) > 0$ se $P \neq Q$ e $d(P, P) = 0$;
- (2) $d(P, Q) = d(Q, P)$;
- (3) $d(P, R) \leq d(P, Q) + d(Q, R)$.

Dimostrazione. Le proprietà (1) e (2) sono facili conseguenze delle proprietà della norma. Per la (3) abbiamo

$$d(P, R) = \|\overrightarrow{PR}\| = \|\overrightarrow{PQ} + \overrightarrow{QR}\| \leq \|\overrightarrow{PQ}\| + \|\overrightarrow{QR}\| = d(P, Q) + d(Q, R).$$

La dimostrazione è completa. \square

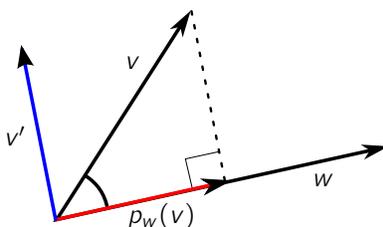


Figura 3.2. Il vettore v si decompone nella sua proiezione ortogonale $p_w(v)$ su w , che è parallela a w , e in un vettore v' che è ortogonale a w .

La proprietà (3) è la *disuguaglianza triangolare*: nel triangolo con vertici P, Q, R , la lunghezza di ogni lato è minore della somma delle lunghezze degli altri due.

3.3.4. Proiezione ortogonale. Una nozione importante associata ad un prodotto scalare è quella di *ortogonalità*. Mostriamo adesso come sia possibile proiettare ortogonalmente vettori su altri vettori.

Sia w un vettore non nullo di V . Sia $v \in V$ un vettore qualsiasi. Definiamo la *proiezione di v su w* il vettore

$$p_w(v) = \frac{\langle v, w \rangle}{\langle w, w \rangle} w.$$

Si veda la Figura 3.2. Notiamo subito che $p_w(v)$ è ottenuto moltiplicando il vettore v per il *coefficiente di Fourier*

$$\frac{\langle v, w \rangle}{\langle w, w \rangle}.$$

Questo coefficiente compare in vari punti nelle pagine seguenti. La proprietà più importante della proiezione è la seguente: se indichiamo con v' il vettore

$$v' = v - p_w(v)$$

notiamo che v' è ortogonale a w come mostrato in Figura 3.2. Infatti:

$$\begin{aligned} \langle v', w \rangle &= \langle v - p_w(v), w \rangle = \langle v - \frac{\langle v, w \rangle}{\langle w, w \rangle} w, w \rangle = \langle v, w \rangle - \frac{\langle v, w \rangle}{\langle w, w \rangle} \langle w, w \rangle \\ &= \langle v, w \rangle - \langle v, w \rangle = 0. \end{aligned}$$

Effettivamente il coefficiente di Fourier è stato introdotto proprio con questo scopo: se cerchiamo un $k \in \mathbb{R}$ tale che $v' = v - kw$ sia ortogonale a w , vediamo facilmente che l'unico k che soddisfa questa richiesta è il coefficiente di Fourier (il lettore è invitato a verificare questo fatto).

3.3.5. Componente parallela e ortogonale ad un vettore. Sia come sopra $w \in V$ un vettore non nullo fissato. Abbiamo scoperto che possiamo decomporre come in Figura 3.2 ciascun vettore $v \in V$ in due componenti:

$$v = p_w(v) + v'$$

dove $p_w(v)$ è la componente *parallela* a w e v' è quella *ortogonale* a w .

Esempio 3.3.8. Consideriamo \mathbb{R}^2 con il prodotto scalare euclideo. Se $w = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ e $v = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ otteniamo

$$p_w(v) = \frac{\langle v, w \rangle}{\langle w, w \rangle} w = 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

e quindi decomponiamo il vettore v in $v = p_w(v) + v'$ nel modo seguente:

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

Queste sono le componenti parallela e ortogonale a w . Facciamo un altro esempio: se $w = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ e $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, otteniamo

$$p_w(v) = \frac{\langle v, w \rangle}{\langle w, w \rangle} w = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

e quindi decomponiamo il vettore v in $v = p_w(v) + v'$ nel modo seguente:

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

3.3.6. Coefficienti di Fourier. Mostriamo adesso come i coefficienti di Fourier forniscano immediatamente le coordinate di un vettore rispetto ad una base ortogonale.

Sia come sempre V uno spazio vettoriale dotato di un prodotto scalare definito positivo. Sia inoltre $\mathcal{B} = \{v_1, \dots, v_n\}$ una base ortogonale.

Proposizione 3.3.9. *Per qualsiasi vettore $v \in V$ vale la relazione seguente:*

$$v = p_{v_1}(v) + \dots + p_{v_n}(v)$$

In altre parole, ciascun vettore v è somma delle sue proiezioni ortogonali sugli elementi della base. Usando i coefficienti di Fourier scriviamo:

$$(5) \quad v = \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 + \dots + \frac{\langle v, v_n \rangle}{\langle v_n, v_n \rangle} v_n.$$

Dimostrazione. Dimostriamo (5) mostrando che il vettore differenza

$$v' = v - \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \dots - \frac{\langle v, v_n \rangle}{\langle v_n, v_n \rangle} v_n$$

è nullo. Per fare ciò, calcoliamo il prodotto scalare $\langle v', v_i \rangle$ di v' con un elemento qualsiasi della base v_i e troviamo:

$$\langle v', v_i \rangle = \langle v, v_i \rangle - \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} \langle v_1, v_i \rangle - \dots - \frac{\langle v, v_n \rangle}{\langle v_n, v_n \rangle} \langle v_n, v_i \rangle.$$

Ricordiamo che i vettori v_1, \dots, v_n sono ortogonali e quindi $\langle v_i, v_j \rangle = 0$ se $i \neq j$. Quindi nel membro di destra molti addendi si annullano e resta

$$\langle v', v_i \rangle = \langle v, v_i \rangle - \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle} \langle v_i, v_i \rangle = \langle v, v_i \rangle - \langle v, v_i \rangle = 0.$$

Il vettore v' è ortogonale a tutti gli elementi v_1, \dots, v_n della base, quindi a tutti gli elementi di V . Poiché il prodotto scalare è definito positivo (ed in particolare non è degenere) otteniamo necessariamente $v' = 0$. \square

L'equazione (5) assume una forma particolarmente semplice se \mathcal{B} è una base ortonormale; in questo caso $\langle v_i, v_i \rangle = 1$ per ogni i e l'equazione diventa:

$$v = \langle v, v_1 \rangle v_1 + \dots + \langle v, v_n \rangle v_n.$$

La Proposizione 3.3.9 può essere enunciata anche nel modo seguente:

Corollario 3.3.10. *Le coordinate di un vettore v rispetto ad una base ortogonale $\mathcal{B} = \{v_1, \dots, v_n\}$ sono i coefficienti di Fourier. Scriviamo quindi:*

$$[v]_{\mathcal{B}} = \begin{pmatrix} \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} \\ \vdots \\ \frac{\langle v, v_n \rangle}{\langle v_n, v_n \rangle} \end{pmatrix}.$$

Se la base \mathcal{B} è ortonormale, scriviamo

$$[v]_{\mathcal{B}} = \begin{pmatrix} \langle v, v_1 \rangle \\ \vdots \\ \langle v, v_n \rangle \end{pmatrix}.$$

Esempio 3.3.11. Consideriamo \mathbb{R}^n con il prodotto scalare euclideo. La base canonica $\mathcal{C} = \{e_1, \dots, e_n\}$ è ortonormale. Se $x \in \mathbb{R}^n$ è un vettore qualsiasi, il coefficiente di Fourier $\langle x, e_i \rangle$ di x rispetto a \mathcal{C} è semplicemente la sua coordinata x_i . Quindi con il Corollario 3.3.10 ritroviamo semplicemente una uguaglianza che già conoscevamo:

$$[x]_{\mathcal{C}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Esempio 3.3.12. Consideriamo \mathbb{R}^2 con il prodotto scalare euclideo e prendiamo la base ortogonale

$$\mathcal{B} = \left\{ v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}.$$

Prendiamo il vettore $v = \begin{pmatrix} 7 \\ 2 \end{pmatrix}$ e verifichiamo che

$$v = \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 + \frac{\langle v, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 = \frac{9}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{5}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 7 \\ 2 \end{pmatrix}.$$

Quindi abbiamo calcolato le coordinate di $\begin{pmatrix} 7 \\ 2 \end{pmatrix}$ rispetto alla base \mathcal{B} :

$$\left[\begin{pmatrix} 7 \\ 2 \end{pmatrix} \right]_{\mathcal{B}} = \begin{pmatrix} 9/2 \\ 5/2 \end{pmatrix}.$$

L'aspetto fondamentale di questa costruzione è che le coordinate di un vettore rispetto ad una base ortogonale possono essere calcolate direttamente, senza dover impostare e risolvere un sistema lineare. Questa costruzione funziona ovviamente solo se la base v_1, \dots, v_n è ortogonale.

3.3.7. Ortogonalizzazione di Gram-Schmidt. Abbiamo appena scoperto che le basi ortogonali sono molto utili, perché le coordinate di qualsiasi vettore si calcolano facilmente tramite i coefficienti di Fourier. Ci poniamo quindi il seguente problema: come facciamo a costruire una base ortogonale?

Descriviamo adesso un algoritmo che prende come *input* dei vettori indipendenti v_1, \dots, v_k e restituisce come *output* dei vettori w_1, \dots, w_k indipendenti e ortogonali. Questo algoritmo è noto come *ortogonalizzazione di Gram-Schmidt*.

L'algoritmo è il seguente. Si costruiscono i vettori w_1, \dots, w_k induttivamente in questo modo:

$$\begin{aligned} w_1 &= v_1, \\ w_2 &= v_2 - p_{w_1}(v_2), \\ w_3 &= v_3 - p_{w_1}(v_3) - p_{w_2}(v_3), \\ &\vdots \\ w_k &= v_k - p_{w_1}(v_k) - \dots - p_{w_{k-1}}(v_k). \end{aligned}$$

In altre parole, ad ogni passo togliamo al vettore v_i le componenti parallele ai vettori w_1, \dots, w_{i-1} già trovati precedentemente, in modo che il nuovo vettore w_i sia ortogonale a questi. Esplicitando i coefficienti di Fourier:

$$\begin{aligned} w_1 &= v_1, \\ w_2 &= v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1, \\ w_3 &= v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2, \\ &\vdots \\ w_k &= v_k - \frac{\langle v_k, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \dots - \frac{\langle v_k, w_{k-1} \rangle}{\langle w_{k-1}, w_{k-1} \rangle} w_{k-1}. \end{aligned}$$

Dimostriamo adesso che l'algoritmo funziona.

Proposizione 3.3.13. *L'algoritmo di Gram-Schmidt produce effettivamente dei vettori w_1, \dots, w_k indipendenti e ortogonali. Inoltre per ogni $j = 1, \dots, k$, i primi j vettori della base generano gli stessi spazi di prima:*

$$(6) \quad \text{Span}(v_1, \dots, v_j) = \text{Span}(w_1, \dots, w_j).$$

Dimostrazione. Dimostriamo tutto per induzione su j . Il caso $j = 1$ è banale, quindi esaminiamo il passo induttivo.

Il vettore w_j è costruito togliendo a v_j le componenti parallele ai vettori precedenti w_1, \dots, w_{j-1} e quindi ciò che resta è ortogonale a questi. Più formalmente, per ogni $i < j$ abbiamo

$$\begin{aligned} \langle w_i, w_j \rangle &= \langle w_i, v_j \rangle - \left\langle w_i, \frac{\langle v_j, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 \right\rangle - \dots - \left\langle w_i, \frac{\langle v_j, w_{j-1} \rangle}{\langle w_{j-1}, w_{j-1} \rangle} w_{j-1} \right\rangle \\ &= \langle w_i, v_j \rangle - \frac{\langle v_j, w_1 \rangle}{\langle w_1, w_1 \rangle} \langle w_i, w_1 \rangle - \dots - \frac{\langle v_j, w_{j-1} \rangle}{\langle w_{j-1}, w_{j-1} \rangle} \langle w_i, w_{j-1} \rangle \\ &= \langle w_i, v_j \rangle - \frac{\langle v_j, w_i \rangle}{\langle w_i, w_i \rangle} \langle w_i, w_i \rangle = 0. \end{aligned}$$

Nella terza uguaglianza abbiamo usato che $\langle w_i, w_l \rangle = 0$ se $i \neq l$. Notiamo inoltre che ciascun w_j è costruito come combinazione lineare dei v_1, \dots, v_j , quindi $\text{Span}(w_1, \dots, w_j) \subset \text{Span}(v_1, \dots, v_j)$. D'altra parte, si può anche scrivere facilmente v_j come combinazione lineare degli w_1, \dots, w_j , quindi anche $\text{Span}(w_1, \dots, w_j) \supset \text{Span}(v_1, \dots, v_j)$ e dimostriamo (6).

In particolare $\dim \text{Span}(w_1, \dots, w_j) = j$ e quindi i nuovi vettori w_1, \dots, w_j sono effettivamente indipendenti. \square

Esempio 3.3.14. Ortogonalizziamo i vettori

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

rispetto al prodotto scalare euclideo di \mathbb{R}^2 . Otteniamo

$$\begin{aligned} w_1 &= v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ w_2 &= v_2 - p_{w_1}(v_2) = v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{1}{1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned}$$

I vettori ottenuti w_1 e w_2 sono effettivamente ortogonali.

Esempio 3.3.15. Ortogonalizziamo i vettori

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

rispetto al prodotto scalare euclideo di \mathbb{R}^3 . Otteniamo

$$w_1 = v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

$$w_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix},$$

$$w_3 = v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix}.$$

I vettori ottenuti w_1 , w_2 e w_3 sono effettivamente ortogonali.

3.3.8. Riscaldamento. L'operazione di *riscaldamento* di un vettore $v \neq 0$ consiste nella sostituzione di v con $v' = \lambda v$, per qualche $\lambda \neq 0$. Se w_1, \dots, w_k sono vettori ortogonali, possiamo sostituire ciascun w_i con un suo riscaldamento $w'_i = \lambda_i w_i$, ed i nuovi vettori w'_1, \dots, w'_k sono ancora ortogonali.

Nell'Esempio 3.3.15 può essere utile riscaldare i vettori nel modo seguente:

$$w'_1 = w_1, \quad w'_2 = 2w_2, \quad w'_3 = \frac{3}{2}w_3$$

ed ottenere una nuova base di vettori ortogonali, più semplice da descrivere:

$$w'_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad w'_2 = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}, \quad w'_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

La *normalizzazione* di un vettore $v \neq 0$ è il riscaldamento

$$v' = \frac{v}{\|v\|}.$$

Questa operazione è importante perché

$$\|v'\| = \frac{\|v\|}{\|v\|} = 1.$$

Il vettore normalizzato v' ha sempre norma uno. Per trasformare una base ortogonale in una ortonormale è sufficiente normalizzare i vettori. Questa operazione ha spesso l'inconveniente di introdurre delle fastidiose radici quadrate.

Esempio 3.3.16. Normalizzando la base ortogonale trovata nell'Esempio 3.3.15 otteniamo la base ortonormale seguente:

$$w'_1 = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{pmatrix}, \quad w'_2 = \begin{pmatrix} -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{3} \end{pmatrix}, \quad w'_3 = \begin{pmatrix} \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \end{pmatrix}.$$

Si verifica facilmente che hanno tutti norma uno.

3.3.9. Ortogonalità. Lavorare con vettori ortogonali ha molti vantaggi. Uno di questi è che l'indipendenza lineare è garantita a priori:

Proposizione 3.3.17. *Se $v_1, \dots, v_k \in V$ sono tutti ortogonali e non nulli, allora sono indipendenti.*

Dimostrazione. Supponiamo esista una combinazione lineare nulla

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

Facendo il prodotto scalare di entrambi i membri con v_i otteniamo

$$0 = \langle 0, v_i \rangle = \langle \lambda_1 v_1 + \dots + \lambda_k v_k, v_i \rangle = \lambda_i \langle v_i, v_i \rangle$$

perché $\langle v_j, v_i \rangle = 0$ se $j \neq i$. A questo punto $\langle v_i, v_i \rangle > 0$ implica $\lambda_i = 0$. Abbiamo scoperto che $\lambda_i = 0 \forall i$ e questo conclude la dimostrazione. \square

3.3.10. Proiezioni su sottospazi. Sia V uno spazio vettoriale con prodotto scalare definito positivo e $U \subset V$ un sottospazio. La *proiezione ortogonale*

$$p_U: V \longrightarrow V$$

è un'applicazione lineare definita nel modo seguente: poiché $V = U \oplus U^\perp$, ogni vettore $v \in V$ si scrive in modo unico come $v = u + u'$ con $u \in U$ e $u' \in U^\perp$. Definiamo quindi

$$p_U(v) = u.$$

Le componenti u e u' di v sono per definizione le proiezioni ortogonali del vettore v sui sottospazi ortogonali U e U^\perp .

Mostriamo come determinare concretamente $p_U(v)$. Entrano in gioco ancora una volta i coefficienti di Fourier.

Proposizione 3.3.18. *Sia v_1, \dots, v_k una base ortogonale per U . Allora*

$$p_U(v) = \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 + \dots + \frac{\langle v, v_k \rangle}{\langle v_k, v_k \rangle} v_k.$$

Dimostrazione. Sia v_{k+1}, \dots, v_n una base ortogonale per U^\perp . Quindi v_1, \dots, v_n per V è una base ortogonale per V . Sappiamo che

$$v = \underbrace{\frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 + \dots + \frac{\langle v, v_k \rangle}{\langle v_k, v_k \rangle} v_k}_u + \underbrace{\frac{\langle v, v_{k+1} \rangle}{\langle v_{k+1}, v_{k+1} \rangle} v_{k+1} + \dots + \frac{\langle v, v_n \rangle}{\langle v_n, v_n \rangle} v_n}_{u'}.$$

Abbiamo scritto $v = u + u'$ con $u \in U$ e $u' \in U^\perp$. Quindi per definizione di proiezione ortogonale troviamo $u = p_U(v)$. \square

Esempio 3.3.19. Consideriamo $U = \{x + y - z = 0\} \subset \mathbb{R}^3$ con il prodotto scalare euclideo. Cerchiamo una base ortogonale per U e troviamo ad esempio

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

A questo punto determiniamo la proiezione ortogonale su U di un generico vettore di \mathbb{R}^3 usando la proposizione precedente, in questo modo:

$$p_U \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{x+y+2z}{6} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + \frac{x-y}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2x-y+z \\ -x+2y+z \\ x+y+2z \end{pmatrix}$$

Quindi la matrice associata a p_U nella base canonica è

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

Usando questa matrice possiamo calcolare la proiezione ortogonale di qualsiasi vettore di \mathbb{R}^3 .

Segue subito dalla definizione che la proiezione ortogonale p_U su un sottospazio $U \subset V$ manda ogni vettore di U in sé stesso, mentre manda ogni vettore di U^\perp in zero. In particolare otteniamo:

$$\text{Im } p_U = U, \quad \text{ker } p_U = U^\perp.$$

Possiamo scrivere agevolmente p_U rispetto ad una base opportuna:

Proposizione 3.3.20. *Se $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di V in cui $\{v_1, \dots, v_k\}$ è base di U e $\{v_{k+1}, \dots, v_n\}$ è base di U^\perp , allora*

$$[p_U]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$$

dove I_k è la matrice identità di taglia k .

Dimostrazione. Vediamo che $p_U(v_i) = v_i$ per $i = 1, \dots, k$ e $p_U(v_i) = 0$ per $i = k+1, \dots, n$, quindi la matrice associata è del tipo richiesto. \square

In particolare abbiamo scoperto che p_U è diagonalizzabile e ha autovalori 0 e 1. Gli autospazi relativi sono precisamente U^\perp e U . Vale inoltre la relazione

$$p_U \circ p_U = p_U.$$

Abbiamo già incontrato delle trasformazioni che soddisfano questa relazione nella Proposizione 2.3.16.

Esempio 3.3.21. Sia $U = \{x + y + z = 0\}$ un piano in \mathbb{R}^3 . La base

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

soddisfa le ipotesi della Proposizione 3.3.20. Quindi p_U in questa base è

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

Nella base canonica la proiezione è rappresentata dalla matrice $B = M^{-1}AM$ con $M = [\text{id}]_{\mathcal{B}}^{\mathcal{C}}$. Facendo i conti otteniamo

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

Il lettore è invitato a controllare che usando la Proposizione 3.3.18 come abbiamo fatto nell'Esempio 3.3.19 otteniamo la stessa matrice.

Concludiamo questa sezione con una osservazione che può tornare utile. Può capitare in molti casi che sia più agevole calcolare la proiezione p_{U^\perp} sullo spazio ortogonale U^\perp invece di p_U . Si può quindi calcolare prima p_{U^\perp} e quindi ritrovare p_U scrivendo semplicemente

$$p_U(v) = v - p_{U^\perp}(v).$$

Esempio 3.3.22. Studiamo l'Esempio 3.3.21 seguendo questo suggerimento. Lo spazio ortogonale U^\perp al piano U è una retta, generata dal vettore

$$w = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Ne segue che $p_U(v) = v - p_{U^\perp}(v) = v - p_w(v)$. Quindi

$$p_U \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \frac{x+y+z}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2x - y - z \\ -x + 2y - z \\ -x - y + 2z \end{pmatrix}.$$

Abbiamo ritrovato la stessa matrice associata dell'Esempio 3.3.21.

3.4. Isometrie

Ogni volta che in matematica si definiscono delle strutture come gruppi, spazi vettoriali, prodotti scalari, esistono delle opportune trasformazioni che preservano interamente queste strutture: queste trasformazioni sono chiamate *isomorfismi* per i gruppi e per gli spazi vettoriali, e *isometrie* per i prodotti scalari.

3.4.1. Definizione. Siano V e W due spazi vettoriali dotati ciascuno di un prodotto scalare. Una *isometria* è un isomorfismo $T: V \rightarrow W$ tale che

$$(7) \quad \langle v, w \rangle = \langle T(v), T(w) \rangle \quad \forall v, w \in V.$$

Brevemente, un'isometria è un isomorfismo che preserva il prodotto scalare.

Notiamo subito che l'inversa T^{-1} di un'isometria è anch'essa un'isometria, e la composizione $T \circ T'$ di due isometrie è una isometria.

Letteralmente, il termine "isometria" indica che T preserva la geometria degli oggetti, cioè non li distorce. Quando il prodotto scalare è definito positivo, possiamo definire questa nozione in almeno tre modi equivalenti:

Proposizione 3.4.1. Sia $T: V \rightarrow W$ un isomorfismo fra spazi dotati di un prodotto scalare definito positivo. I fatti seguenti sono equivalenti:

- (1) T è una isometria,
- (2) T preserva la norma, cioè $\|T(v)\| = \|v\| \quad \forall v \in V$,
- (3) T preserva la distanza, cioè $d(v, w) = d(T(v), T(w)) \quad \forall v, w \in V$.

Dimostrazione. (1) \Rightarrow (2). Se T è una isometria, allora

$$\|T(v)\|^2 = \langle T(v), T(v) \rangle = \langle v, v \rangle = \|v\|^2.$$

(2) \Rightarrow (3). Vale

$$d(v, w) = \|v - w\| = \|T(v - w)\| = \|T(v) - T(w)\| = d(T(v), T(w)).$$

(3) \Rightarrow (2). Vale

$$\|v\| = d(0, v) = d(T(0), T(v)) = d(0, T(v)) = \|T(v)\|.$$

(2) \Rightarrow (1). Questo è meno facile dei precedenti:

$$\begin{aligned} \langle v, w \rangle &= \frac{\langle v+w, v+w \rangle - \langle v, v \rangle - \langle w, w \rangle}{2} = \frac{\|v+w\|^2 - \|v\|^2 - \|w\|^2}{2} \\ &= \frac{\|T(v+w)\|^2 - \|T(v)\|^2 - \|T(w)\|^2}{2} \\ &= \frac{\langle T(v+w), T(v+w) \rangle - \langle T(v), T(v) \rangle - \langle T(w), T(w) \rangle}{2} \\ &= \langle T(v), T(w) \rangle. \end{aligned}$$

La dimostrazione è completa. \square

3.4.2. Con le basi. Vogliamo adesso costruire e studiare concretamente delle isometrie. Per mostrare che un isomorfismo $T: V \rightarrow W$ è una isometria dovremmo in teoria verificare la condizione (7) su ogni coppia $v, w \in V$. Come spesso accade, è sufficiente farlo per gli elementi di una base fissata. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V .

Lemma 3.4.2. Un isomorfismo T è un'isometria se e solo se

$$(8) \quad \langle v_i, v_j \rangle = \langle T(v_i), T(v_j) \rangle \quad \forall i, j.$$

Dimostrazione. Chiaramente (7) implica (8), mostriamo adesso il contrario. Per due generici v e w scriviamo

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n, \quad w = \mu_1 v_1 + \dots + \mu_n v_n$$

e quindi

$$\begin{aligned} \langle v, w \rangle &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j \langle v_i, v_j \rangle, \\ \langle T(v), T(w) \rangle &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j \langle T(v_i), T(v_j) \rangle \end{aligned}$$

usando il fatto che il prodotto scalare è bilineare e T è lineare. Quindi (8) implica (7). \square

Consideriamo ad esempio il caso in cui V e W siano dotati di un prodotto scalare definito positivo. Possiamo dedurre il fatto seguente.

Corollario 3.4.3. *Sia v_1, \dots, v_n una base ortonormale per V . Un isomorfismo $T: V \rightarrow W$ è una isometria \iff anche le immagini $T(v_1), \dots, T(v_n)$ formano una base ortonormale per W .*

Possiamo dire quindi che le isometrie sono precisamente gli isomorfismi che mandano basi ortonormali in basi ortonormali.

3.4.3. Con le matrici. Cerchiamo adesso di tradurre la definizione di isometria nel linguaggio concreto delle matrici. Siano V, V' spazi vettoriali dotati di prodotto scalare g, g' e di basi $\mathcal{B}, \mathcal{B}'$ rispettivamente. Consideriamo un isomorfismo $T: V \rightarrow V'$. Siano

$$S = [g]_{\mathcal{B}}, \quad S' = [g']_{\mathcal{B}'}, \quad A = [T]_{\mathcal{B}'}$$

le matrici associate.

Proposizione 3.4.4. *L'isomorfismo T è una isometria se e solo se*

$$S = {}^t A S' A.$$

Dimostrazione. Sappiamo che T è una isometria se e solo se $g(v_i, v_j) = g'(T(v_i), T(v_j))$ per ogni i, j . Sappiamo anche che

$$S_{ij} = g(v_i, v_j).$$

Ricordiamo che la colonna i -esima di A è $A^i = [T(v_i)]_{\mathcal{B}'}$. Quindi

$$({}^t A S' A)_{ij} = {}^t (A^i) S' A^j = g'(T(v_i), T(v_j)).$$

Ne deduciamo che T è una isometria se e solo se $S = {}^t A S' A$. \square

In particolare, consideriamo il caso in cui V sia dotato di un prodotto scalare g e $T: V \rightarrow V$ sia un isomorfismo. Siano $S = [g]_{\mathcal{B}}$ e $A = [T]_{\mathcal{B}}$.

Corollario 3.4.5. *L'isomorfismo T è una isometria se e solo se*

$$S = {}^t A S A.$$

Consideriamo adesso il caso ancora più specifico:

Corollario 3.4.6. *Sia g definito positivo e \mathcal{B} una base ortonormale. L'endomorfismo T è una isometria se e solo se*

$${}^t A A = I.$$

Dimostrazione. Con queste ipotesi $S = [g]_{\mathcal{B}} = I$, quindi la relazione $S = {}^t A S A$ diventa $I = {}^t A A$. \square

3.4.4. Matrici ortogonali. Consideriamo infine il caso che ci sta più a cuore, quello di \mathbb{R}^n munito del suo prodotto scalare euclideo. Sia $A \in M(n)$ una matrice quadrata.

Corollario 3.4.7. *L'endomorfismo L_A è una isometria se e solo se ${}^tAA = I$.*

Dimostrazione. La matrice A è la matrice associata a L_A rispetto alla base canonica, che è ortonormale: quindi applichiamo il Corollario 3.4.6. \square

Una matrice $A \in M(n)$ a coefficienti reali tale che ${}^tAA = I$ è detta *ortogonale*. Studiamo adesso alcune semplici proprietà delle matrici ortogonali.

Proposizione 3.4.8. *Se $A \in M(n)$ è ortogonale, allora*

- $\det A = \pm 1$,
- *i possibili autovalori di A sono ± 1 .*

Dimostrazione. Se ${}^tAA = I$, allora per Binet

$$1 = \det I = \det({}^tA) \det A = (\det A)^2.$$

Quindi $\det A = \pm 1$. Se λ è un autovalore, esiste $v \neq 0$ tale che $Av = \lambda v$. L'isometria L_A preserva la norma e quindi

$$\|v\| = \|Av\| = \|\lambda v\| = |\lambda| \|v\|.$$

Segue che $\lambda = \pm 1$. \square

Dal fatto che $\det A = \pm 1$, deduciamo che una matrice ortogonale A è sempre invertibile; inoltre l'inversa di A è sorprendentemente semplice da determinare: dalla relazione ${}^tAA = I$ deduciamo infatti che $A^{-1} = {}^tA$. In particolare vale anche $A {}^tA = I$.

Proposizione 3.4.9. *Una matrice $A \in M(n)$ è ortogonale \iff i vettori colonna A^1, \dots, A^n formano una base ortonormale di \mathbb{R}^n .*

Dimostrazione. La relazione ${}^tAA = I$ è equivalente alla relazione $\langle A^i, A^j \rangle = \delta_{ij}$ per ogni i, j . In alternativa, questa è una conseguenza del Corollario 3.4.3, visto che i vettori colonna sono le immagini della base canonica (che è ortonormale). \square

3.4.5. Riflessioni. Introduciamo una classe importante di isometrie particolarmente semplici da descrivere. Sia V uno spazio vettoriale dotato di un prodotto scalare definito positivo e sia $U \subset V$ un sottospazio. Ricordiamo che ogni vettore $v \in V$ si decompone nelle proiezioni ortogonali su U e U^\perp :

$$v = p_U(v) + p_{U^\perp}(v).$$

Definizione 3.4.10. La *riflessione* lungo U è l'isometria

$$\text{Rif}_U(v) = v - 2p_{U^\perp}(v) = p_U(v) - p_{U^\perp}(v).$$

Se v_1, \dots, v_k e v_{k+1}, \dots, v_n sono basi ortonormali di U e U^\perp , allora $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base ortonormale di V e la matrice associata a Rif_U nella base \mathcal{B} è molto semplice:

$$[\text{Rif}_U]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{pmatrix}.$$

Questa matrice è chiaramente ortogonale: grazie alla Corollario 3.4.6 abbiamo una conferma che Rif_U sia effettivamente una isometria. Gli autospazi di Rif_U sono

$$V_1 = U, \quad V_{-1} = U^\perp.$$

Notiamo che $\det \text{Rif}_U = (-1)^{n-k} = (-1)^{n-\dim U}$. Quindi una riflessione lungo un iperpiano ha determinante -1 e *inverte* l'orientazione dello spazio. Invece una riflessione di \mathbb{R}^3 lungo una retta preserva l'orientazione.

La riflessione $\text{Rif}_{\{0\}}$ rispetto all'origine $U = \{0\}$ è semplicemente l'endomorfismo $v \mapsto -v$ che cambia di segno tutti i vettori di V . Nota che il determinante di questa applicazione è $(-1)^{\dim V}$. Se $V = \mathbb{R}^2$, la riflessione rispetto all'origine è semplicemente una rotazione di angolo π .

Esempio 3.4.11. Nel caso in cui U sia un iperpiano la riflessione Rif_U è semplice da scrivere. Sia v_0 un generatore della retta U^\perp . Troviamo

$$\text{Rif}_U(v) = v - 2 \frac{\langle v, v_0 \rangle}{\langle v_0, v_0 \rangle} v_0.$$

Nota che questa formula è molto simile a quella che rappresenta la proiezione p_U , con un coefficiente 2 al posto di 1. Ad esempio, se $U \subset \mathbb{R}^3$ è il piano

$$U = \{x + y + z = 0\}$$

otteniamo

$$\text{Rif}_U \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} - 2 \frac{x+y+z}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} x - 2y - 2z \\ -2x + y - 2z \\ -2x - 2y + 1z \end{pmatrix}.$$

Confronta con l'Esempio 3.3.22. La riflessione su U rispetto alla base canonica è rappresentata dalla matrice

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

che è effettivamente ortogonale.

3.4.6. Isometrie del piano. Vogliamo adesso classificare completamente le isometrie del piano \mathbb{R}^2 , dotato del prodotto scalare euclideo. Per quanto appena visto, questo problema è equivalente a classificare le matrici ortogonali 2×2 . Le matrici ortogonali 2×2 possono essere descritte tutte facilmente:

Proposizione 3.4.12. *Le matrici ortogonali in $M(2)$ sono le seguenti:*

$$\text{Rot}_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad \text{Rif}_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}.$$

Dimostrazione. I vettori colonna A^1 e A^2 di una matrice ortogonale A devono formare una base ortonormale. Scriviamo A^1 come un generico vettore unitario

$$A^1 = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}.$$

Il secondo vettore A^2 deve essere ortogonale ad A^1 ed unitario: ci sono solo due possibilità che differiscono solo di un segno, e otteniamo così le matrici Rot_θ e Rif_θ . \square

Notiamo che $\det \text{Rot}_\theta = 1$ e $\det \text{Rif}_\theta = -1$. Cosa rappresentano geometricamente queste isometrie? Troviamo facilmente che queste sono trasformazioni già note. Iniziamo con quelle con determinante positivo.

Proposizione 3.4.13. *L'isometria Rot_θ è una rotazione antioraria di angolo θ intorno all'origine.*

Dimostrazione. Scriviamo il generico vettore in coordinate polari:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \rho \cos \alpha \\ \rho \sin \alpha \end{pmatrix}.$$

Otteniamo quindi

$$\begin{aligned} \text{Rot}_\theta \begin{pmatrix} \rho \cos \alpha \\ \rho \sin \alpha \end{pmatrix} &= \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \rho \cos \alpha \\ \rho \sin \alpha \end{pmatrix} \\ &= \begin{pmatrix} \rho(\cos \theta \cos \alpha - \sin \theta \sin \alpha) \\ \rho(\sin \theta \cos \alpha + \cos \theta \sin \alpha) \end{pmatrix} = \begin{pmatrix} \rho \cos(\alpha + \theta) \\ \rho \sin(\alpha + \theta) \end{pmatrix}. \end{aligned}$$

Ne deduciamo che la trasformazione Rot_θ trasforma il vettore con coordinate polari (ρ, α) nel vettore con coordinate polari $(\rho, \alpha + \theta)$. In altre parole, è una rotazione antioraria di angolo θ intorno all'origine. \square

Passiamo adesso alle isometrie con determinante negativo.

Proposizione 3.4.14. *L'isometria Rif_θ è la riflessione rispetto alla retta passante per l'origine che forma un angolo $\frac{\theta}{2}$ con l'asse x .*

Dimostrazione. Per analizzare Rif_θ , scriviamo il suo polinomio caratteristico $p(\lambda) = \lambda^2 - 1$ e notiamo che Rif_θ ha autovalori ± 1 . Troviamo

$$\begin{aligned} V_1 &= \ker(\text{Rif}_\theta - I) = \ker \begin{pmatrix} \cos \theta - 1 & \sin \theta \\ \sin \theta & -\cos \theta - 1 \end{pmatrix} = \text{Span} \begin{pmatrix} \sin \theta \\ 1 - \cos \theta \end{pmatrix} \\ &= \text{Span} \begin{pmatrix} 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} \\ 1 - \cos^2 \frac{\theta}{2} + \sin^2 \frac{\theta}{2} \end{pmatrix} = \text{Span} \begin{pmatrix} 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} \\ 2 \sin^2 \frac{\theta}{2} \end{pmatrix} = \text{Span} \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix}. \end{aligned}$$

L'altro autospazio V_{-1} si determina in modo analogo e si verifica che V_1 e V_{-1} sono ortogonali. Quindi Rif_θ è una riflessione rispetto alla retta vettoriale V_1 , che forma un angolo $\frac{\theta}{2}$ con l'asse x . \square

3.4.7. Rotazioni nello spazio. Studiamo adesso alcune isometrie importanti dello spazio euclideo \mathbb{R}^3 .

Sia $r \subset \mathbb{R}^3$ una retta e $U = r^\perp$ il suo piano ortogonale. Sia

$$\mathcal{B} = \{v_1, v_2, v_3\}$$

una base ortonormale con $v_1 \in r$ e quindi $v_2, v_3 \in U$. Fissiamo un angolo θ e consideriamo l'endomorfismo T avente matrice associata

$$A = [T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}.$$

Si vede facilmente che ${}^tAA = I$, cioè la matrice A è ortogonale. Siccome \mathcal{B} è una base ortonormale, per il Corollario 3.4.6 l'endomorfismo T è una isometria.

Di che tipo di isometria si tratta? Esaminando la matrice A , si verifica facilmente che $T(r) = r$ e $T(U) = U$. Inoltre $T(v) = v$ per ogni $v \in r$, cioè T fissa tutti i punti della retta r , mentre T ruota il piano U in senso antiorario di un angolo θ . Quindi T è una *rotazione intorno all'asse r di un angolo θ* .

I casi $\theta = 0$ e $\theta = \pi$ sono un po' particolari:

- una rotazione di angolo $\theta = 0$ è l'identità;
- una rotazione di angolo $\theta = \pi$ è una riflessione rispetto alla retta r .

Ovviamente è sempre possibile scrivere T nella base canonica, cambiando la matrice per similitudine nel modo usuale. Se $M = [\text{id}]_{\mathcal{C}}^{\mathcal{B}}$ è la matrice di cambiamento di base, otteniamo

$$A' = [T]_{\mathcal{C}}^{\mathcal{C}} = MAM^{-1}.$$

La nuova matrice A' può essere però molto complicata. Poiché A' rappresenta una isometria rispetto ad una base \mathcal{C} ortonormale, la matrice A' sarà comunque ortogonale. Anche la matrice M è ortogonale, perché le sue colonne sono una base ortonormale: segue che l'inversa $M^{-1} = {}^tM$ si calcola agevolmente facendo la trasposta.

Esempio 3.4.15. Determiniamo una matrice A' che rappresenti (rispetto alla base canonica \mathcal{C}) una rotazione T intorno a

$$r = \text{Span} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

di angolo $\theta = \frac{2\pi}{3}$. Cerchiamo una base ortonormale v_1, v_2, v_3 con $v_1 \in r$. Ad esempio prendiamo questi:

$$v_1 = \frac{\sqrt{3}}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad v_3 = \frac{\sqrt{6}}{6} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

Otteniamo

$$A = [T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & \text{Rot}_{\theta} \end{pmatrix}.$$

Quindi $A' = MA^tM$, dove M è la matrice avente come colonne v_1, v_2, v_3 . Quindi A' è pari a

$$\frac{1}{6} \begin{pmatrix} 2\sqrt{3} & 3\sqrt{2} & \sqrt{6} \\ 2\sqrt{3} & -3\sqrt{2} & \sqrt{6} \\ 2\sqrt{3} & 0 & -2\sqrt{6} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 2\sqrt{3} & 2\sqrt{3} & 2\sqrt{3} \\ 3\sqrt{2} & -3\sqrt{2} & 0 \\ \sqrt{6} & \sqrt{6} & -2\sqrt{6} \end{pmatrix} \frac{1}{6}.$$

Svolgendo un po' di conti si trova infine la matrice

$$A' = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Il lettore può anche convincersi con metodi geometrici che effettivamente questa matrice, che permuta la base canonica, rappresenta una rotazione con asse r e angolo $\frac{2\pi}{3}$.

3.4.8. Antirotazioni. Introduciamo adesso alcune isometrie di \mathbb{R}^3 che invertono l'orientazione dello spazio.

Definizione 3.4.16. Una *antirotazione* $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ è la composizione di una rotazione intorno ad un asse r e di una riflessione rispetto al piano $U = r^\perp$.

Scriviamo la matrice associata ad una antirotazione T rispetto ad una base opportuna. Se prendiamo un vettore unitario $v_1 \in r$ e lo completiamo a base ortonormale $\mathcal{B} = \{v_1, v_2, v_3\}$, otteniamo una base ortonormale $\{v_2, v_3\}$ per U e scriviamo la composizione nel modo seguente:

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} -1 & 0 \\ 0 & I \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & \text{Rot}_{\theta} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & \text{Rot}_{\theta} \end{pmatrix}.$$

Come per le rotazioni, i casi $\theta = 0$ e $\theta = \pi$ sono un po' particolari:

- una antirotazione di angolo $\theta = 0$ è una riflessione rispetto a U ;
- una antirotazione di angolo $\theta = \pi$ è una riflessione rispetto all'origine.

Ricapitolando, rotazioni e antirotazioni in \mathbb{R}^3 sono isometrie che, rispetto ad una base ortonormale, si rappresentano nel modo seguente:

$$\begin{pmatrix} 1 & 0 \\ 0 & \text{Rot}_{\theta} \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & \text{Rot}_{\theta} \end{pmatrix}.$$

Questo insieme di matrici contiene anche tutte le riflessioni rispetto a piani, rette, e all'origine. Le riflessioni si ottengono con i valori $\theta = 0$ oppure π .

Nella prossima sezione dimostriamo che tutte le isometrie di \mathbb{R}^3 sono rotazioni o antirotazioni.

3.4.9. Isometrie dello spazio. Abbiamo classificato precedentemente le isometrie di \mathbb{R}^2 , e passiamo adesso alle isometrie di \mathbb{R}^3 . Osserviamo innanzitutto che descrivere esplicitamente tutte le matrici ortogonali 3×3 è possibile ma molto più laborioso che in dimensione due. Ci accontenteremo quindi di dare una descrizione geometrica chiara di tutte le isometrie possibili. Ci servirà questo fatto generale:

Proposizione 3.4.17. *Sia $T: V \rightarrow V$ una isometria di uno spazio V dotato di un prodotto scalare definito positivo. Per ogni sottospazio $U \subset V$ vale*

$$T(U)^\perp = T(U^\perp).$$

Dimostrazione. Se $v \in U^\perp$, allora $\langle v, u \rangle = 0$ per ogni $u \in U$. Quindi $\langle T(v), T(u) \rangle = 0$ per ogni $u \in U$, in altre parole $T(v) \in T(U)^\perp$. Abbiamo mostrato che $T(U^\perp) \subset T(U)^\perp$. I due sottospazi in verità coincidono perché hanno entrambi la stessa dimensione $\dim V - \dim U$. \square

Corollario 3.4.18. *Se $T(U) = U$, allora $T(U^\perp) = U^\perp$.*

Il teorema seguente è un risultato importante della geometria dello spazio. La sua dimostrazione è anche interessante perché usa molti degli argomenti trattati nei capitoli precedenti: matrici associate, autovettori e prodotti scalari.

Teorema 3.4.19. *Ogni isometria di \mathbb{R}^3 è una rotazione o un'antirotazione.*

Dimostrazione. Sia $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ una isometria. Siccome \mathbb{R}^3 ha dimensione dispari, per la Proposizione 1.3.11 esiste almeno un autovettore $v \neq 0$ per T . L'autovalore di v è ± 1 per la Proposizione 3.4.8. Quindi $T(v) = \pm v$.

Il piano $U = v^\perp$ ortogonale a v è T -invariante per il Corollario 3.4.18. Prendiamo una base ortonormale $\mathcal{B} = \{v, v_2, v_3\}$. I vettori v_2, v_3 formano una base ortonormale di U e otteniamo

$$A = [T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix}$$

dove B è una matrice 2×2 ortogonale che rappresenta la restrizione $T|_U$ nella base ortonormale $\{v_2, v_3\}$. Per la Proposizione 3.4.12, la matrice B è una matrice di rotazione o di riflessione. Se è di rotazione, otteniamo

$$A = \begin{pmatrix} \pm 1 & 0 \\ 0 & \text{Rot}_\theta \end{pmatrix}$$

e quindi A è una rotazione o antirotazione a seconda del segno di ± 1 . Se B è una matrice di riflessione, esiste una base ortonormale $\{v'_2, v'_3\}$ di U rispetto alla

quale la riflessione è la matrice $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$. Con la base $\mathcal{B}' = \{v, v'_2, v'_3\}$ otteniamo

$$A' = [T]_{\mathcal{B}'}^{\mathcal{B}'} = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Questa è una riflessione, un caso particolare di rotazione o antirotazione. \square

Abbiamo scoperto che ogni isometria di \mathbb{R}^3 è una rotazione o una antirotazione. Ci poniamo il problema seguente: concretamente, data una matrice ortogonale A , come faccio a determinare geometricamente di che tipo di (anti-)rotazione si tratta?

Rispondere a questa domanda è molto facile. Innanzitutto, se $\det A = 1$ si tratta di una rotazione; se invece $\det A = -1$, è una antirotazione. L'angolo di (anti-)rotazione θ è inoltre facilmente determinato dalla traccia di A , tramite la formula seguente:

$$\operatorname{tr} A = \pm 1 + 2 \cos \theta$$

dove $\pm 1 = \det A$ dipende se A è una rotazione o una antirotazione. Quindi possiamo scrivere

$$\cos \theta = \frac{\operatorname{tr} A - \det A}{2}.$$

In questo modo si determina immediatamente l'angolo θ . Se $\theta = 0$ oppure π , la matrice A rappresenta in realtà una riflessione (oppure è l'identità), che può essere descritta trovando gli autospazi V_1 e V_{-1} .

Nel caso generico $\theta \neq 0, \pi$, l'autospazio V_1 (rispettivamente, V_{-1}) è precisamente l'asse di rotazione (rispettivamente, antirotazione).

Esempio 3.4.20. Data la matrice

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

vediamo che $\det A = 1$ e quindi è una rotazione; abbiamo $\operatorname{tr} A = 0$ e quindi $\cos \theta = -\frac{1}{2}$, che implica $\theta = \pm \frac{3\pi}{2}$. Infine, l'asse di rotazione è l'autospazio

$$V_1 = \ker(A - I) = \ker \begin{pmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} = \operatorname{Span} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Lo spazio euclideo

In questo capitolo studiamo la geometria dello spazio euclideo \mathbb{R}^n , con una particolare attenzione al piano \mathbb{R}^2 e allo spazio \mathbb{R}^3 che sono ovviamente quelli che ci riguardano più da vicino.

4.1. Prodotto vettoriale

Il prodotto scalare è una operazione che prende due vettori e restituisce uno scalare. Introduciamo qui un'altra operazione, il prodotto vettoriale, che prende due vettori e restituisce un *vettore*. Il prodotto vettoriale è definito solo su \mathbb{R}^3 .

4.1.1. Definizione. Consideriamo due vettori $v, w \in \mathbb{R}^3$:

$$v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \quad w = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}.$$

Il *prodotto vettoriale* fra v e w è il vettore

$$v \times w = \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix}.$$

Notiamo che

$$v \times w = \begin{pmatrix} d_1 \\ -d_2 \\ d_3 \end{pmatrix}$$

dove d_i è il determinante del minore 2×2 ottenuto cancellando la i -esima riga dalla matrice

$$A = \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \\ v_3 & w_3 \end{pmatrix}.$$

Un metodo per tenere a mente la definizione di prodotto vettoriale è il seguente. Si ottiene il prodotto vettoriale svolgendo formalmente il determinante di

questa "matrice" lungo l'ultima colonna:

$$\begin{aligned} v \times w &= \det \begin{pmatrix} v_1 & w_1 & e_1 \\ v_2 & w_2 & e_2 \\ v_3 & w_3 & e_3 \end{pmatrix} \\ &= \det \begin{pmatrix} v_2 & w_2 \\ v_3 & w_3 \end{pmatrix} e_1 - \det \begin{pmatrix} v_1 & w_1 \\ v_3 & w_3 \end{pmatrix} e_2 + \det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} e_3. \end{aligned}$$

Questa è solo una regola mnemonica: quella matrice in realtà non è una vera matrice perché e_1, e_2, e_3 non sono numeri ma i vettori della base canonica.

Esempio 4.1.1. Vale $e_1 \times e_2 = e_3$.

4.1.2. Proprietà. Studiamo le proprietà del prodotto vettoriale. Siano $v, w \in \mathbb{R}^3$ due vettori qualsiasi.

Proposizione 4.1.2. *Il vettore $v \times w$ è ortogonale sia a v che a w .*

Dimostrazione. Si ottiene

$$\begin{aligned} \langle v \times w, v \rangle &= \det \begin{pmatrix} v_2 & w_2 \\ v_3 & w_3 \end{pmatrix} v_1 - \det \begin{pmatrix} v_1 & w_1 \\ v_3 & w_3 \end{pmatrix} v_2 + \det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} v_3 \\ &= \det \begin{pmatrix} v_1 & w_1 & v_1 \\ v_2 & w_2 & v_2 \\ v_3 & w_3 & v_3 \end{pmatrix} = 0. \end{aligned}$$

La seconda uguaglianza è lo sviluppo di Laplace sull'ultima colonna. Il determinante è nullo perché la matrice ha due colonne uguali. Analogamente si trova $\langle v \times w, w \rangle = 0$. \square

Proposizione 4.1.3. *Il vettore $v \times w$ è nullo $\iff v$ e w sono dipendenti.*

Dimostrazione. Ricordiamo che

$$v \times w = \begin{pmatrix} d_1 \\ -d_2 \\ d_3 \end{pmatrix}$$

dove d_1, d_2, d_3 sono i determinanti dei minori 2×2 della matrice A avente come colonne v e w . Otteniamo

$$v \times w = 0 \iff \text{rk}A \leq 1 \iff v \text{ e } w \text{ sono dipendenti.}$$

La dimostrazione è conclusa. \square

Proposizione 4.1.4. *Vale l'equazione*

$$\|v \times w\|^2 + \langle v, w \rangle^2 = \|v\|^2 \|w\|^2.$$

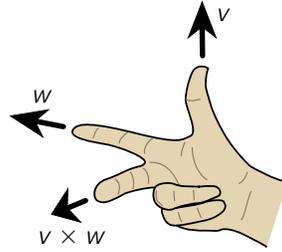


Figura 4.1. La regola della mano destra.

Dimostrazione. Troviamo

$$\begin{aligned}\|v \times w\|^2 &= (v_2 w_3 - v_3 w_2)^2 + (v_1 w_3 - v_3 w_1)^2 + (v_1 w_2 - v_2 w_1)^2 \\ &= (v_1^2 + v_2^2 + v_3^2)(w_1^2 + w_2^2 + w_3^2) - (v_1 w_1 + v_2 w_2 + v_3 w_3)^2 \\ &= \|v\|^2 \|w\|^2 - \langle v, w \rangle^2.\end{aligned}$$

La dimostrazione è completa. \square

Corollario 4.1.5. *Il modulo del prodotto vettoriale è*

$$\|v \times w\| = \|v\| \|w\| \cdot |\sin \theta| = \text{Area}(P)$$

dove θ è l'angolo formato da v e w e P è il parallelogramma con lati v e w .

Dimostrazione. Troviamo

$$\|v \times w\|^2 = \|v\|^2 \|w\|^2 - \langle v, w \rangle^2 = \|v\|^2 \|w\|^2 (1 - \cos^2 \theta) = \|v\|^2 \|w\|^2 \sin^2 \theta.$$

Quindi $\|v \times w\| = \|v\| \|w\| \cdot |\sin \theta|$ e questa è proprio l'area di P . \square

Se v e w sono dipendenti, allora $v \times w = 0$. Se sono indipendenti, allora la direzione di $v \times w$ è ortogonale al piano contenente v e w e il modulo di $v \times w$ è l'area di P . Queste due condizioni determinano il vettore $v \times w$ a meno di segno: per scegliere il "verso giusto" è sufficiente usare la regola della mano destra mostrata in Figura 4.1 e giustificata dalla proposizione seguente.

Proposizione 4.1.6. *Se v e w sono indipendenti, la terna $v, w, v \times w$ è una base positiva di \mathbb{R}^3 .*

Dimostrazione. Dobbiamo mostrare che la matrice avente come colonne $v, w, v \times w$ ha determinante positivo. Calcoliamo sviluppando sulla terza

colonna:

$$\det \begin{pmatrix} v_1 & w_1 & \det \begin{pmatrix} v_2 & w_2 \\ v_3 & w_3 \end{pmatrix} \\ v_2 & w_2 & -\det \begin{pmatrix} v_1 & w_1 \\ v_3 & w_3 \end{pmatrix} \\ v_3 & w_3 & \det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} \end{pmatrix} = \left(\det \begin{pmatrix} v_2 & w_2 \\ v_3 & w_3 \end{pmatrix} \right)^2 + \left(\det \begin{pmatrix} v_1 & w_1 \\ v_3 & w_3 \end{pmatrix} \right)^2 + \left(\det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} \right)^2 > 0.$$

La dimostrazione è completa. \square

Possiamo quindi definire geometricamente il prodotto vettoriale $v \times w$ di due vettori indipendenti come l'unico vettore ortogonale ad entrambi, di lunghezza pari all'area del parallelogramma con lati v e w , e orientato positivamente rispetto a v e w .

Dalla definizione seguono facilmente anche i fatti seguenti:

- $v \times w = -w \times v$ per ogni $v, w \in \mathbb{R}^3$,
- il prodotto vettoriale $\times : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ è bilineare.

Come per il prodotto scalare, per bilinearità intendiamo:

$$(v + v') \times w = v \times w + v' \times w, \quad (\lambda v) \times w = \lambda(v \times w),$$

$$v \times (w + w') = v \times w + v \times w', \quad v \times (\lambda w) = \lambda(v \times w).$$

Abbiamo visto che il prodotto vettoriale ha molte proprietà. A differenza del prodotto scalare, il prodotto vettoriale non è commutativo, ma *anticommutativo*: vale $v \times w = -w \times v$.

C'è però una differenza fondamentale fra il prodotto vettoriale e altri prodotti, come quello fra numeri o fra matrici: non è verificata la proprietà associativa! Infatti ad esempio:

$$\begin{aligned} (e_1 \times e_2) \times e_2 &= e_3 \times e_2 = -e_2 \times e_3 = -e_1 \\ e_1 \times (e_2 \times e_2) &= e_1 \times 0 = 0. \end{aligned}$$

Esercizio 4.1.7 (Prodotto triplo). Siano $v_1, v_2, v_3 \in \mathbb{R}^3$. Mostra che

$$\langle v_1 \times v_2, v_3 \rangle = \langle v_1, v_2 \times v_3 \rangle = \det(v_1 | v_2 | v_3).$$

Se v_1, v_2, v_3 sono indipendenti, questo è il volume del parallelepipedo con tre lati v_1, v_2, v_3 .

4.1.3. Applicazioni. Il prodotto vettoriale è innanzitutto un metodo diretto per trovare velocemente un vettore ortogonale a due vettori indipendenti dati v_1 e v_2 .

Analogamente, si può usare il prodotto vettoriale per passare velocemente da coordinate parametriche a cartesiane per qualsiasi piano vettoriale di \mathbb{R}^3 . Se

$W = \text{Span}(v, w)$ è un piano di \mathbb{R}^3 , questo si può scrivere in forma cartesiana come

$$W = \{ax + by + cz = 0\}$$

dove a, b, c sono i coefficienti del vettore $v \times w$.

Esempio 4.1.8. Un piano

$$W = \text{Span} \left\{ \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}, \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} \right\}$$

ha come vettore ortogonale

$$v = \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} \times \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} = \begin{pmatrix} b_1 c_2 - b_2 c_1 \\ a_2 c_1 - a_1 c_2 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}$$

e quindi ha equazioni cartesiane

$$W = \{(b_1 c_2 - b_2 c_1)x + (a_2 c_1 - a_1 c_2)y + (a_1 b_2 - a_2 b_1)z = 0\}.$$

Vedremo nelle pagine seguenti altre applicazioni geometriche del prodotto vettoriale. Il prodotto vettoriale è usato in fisica, ad esempio nelle equazioni di Maxwell.

4.2. Sottospazi affini

Per definizione, un sottospazio vettoriale $U \subset V$ contiene sempre l'origine. Consideriamo adesso altri sottoinsiemi, chiamati *sottospazi affini*, che possono non contenere l'origine e sono ottenuti traslando i sottospazi vettoriali.

4.2.1. Traslazioni. Sia V uno spazio vettoriale e $v_0 \in V$ un vettore fissato. La *traslazione* di direzione v_0 è la mappa $T_{v_0}: V \rightarrow V$ definita nel modo seguente:

$$T_{v_0}(v) = v + v_0.$$

La prima considerazione da fare a proposito di questa mappa è che *non* è lineare se $v_0 \neq 0$, semplicemente perché $T_{v_0}(0) = v_0 \neq 0$.

Un'altra osservazione da fare è che la composizione di due traslazioni è sempre una traslazione:

$$T_{v_0} \circ T_{v_1} = T_{v_1} \circ T_{v_0} = T_{v_0+v_1}.$$

Notiamo anche che due traslazioni commutano sempre. Una traslazione è sempre una mappa invertibile, cioè è una corrispondenza biunivoca di V . L'inversa della traslazione T_{v_0} è la traslazione T_{-v_0} . Infatti

$$T_{-v_0} \circ T_{v_0} = T_0 = \text{id}.$$

4.2.2. Sottospazi affini. Sia V uno spazio vettoriale. Un *sottospazio affine* di V è un sottoinsieme $S \subset V$ ottenuto traslando un sottospazio vettoriale $U \subset V$ di un certo vettore v_0 , cioè

$$S = T_{v_0}(U).$$

Possiamo indicare questo sottospazio con la notazione $S = U + v_0$, che va intesa nel modo seguente:

$$S = T_{v_0}(U) = U + v_0 = \{u + v_0 \mid u \in U\}.$$

Facciamo qualche osservazione. Innanzitutto notiamo che $v_0 \in S$ perché $v_0 = 0 + v_0$. Notiamo inoltre che v_0 non è un punto speciale di S : nella notazione possiamo sostituirlo con qualsiasi altro punto $v \in S$, infatti

Proposizione 4.2.1. *Per qualsiasi $v \in S$ otteniamo*

$$S = U + v.$$

Dimostrazione. Se $v \in S$, allora $v = u + v_0$ per qualche $u \in U$ e quindi

$$U + v = (U + u) + v_0 = U + v_0.$$

Abbiamo usato che $U + u = U$ per ogni $u \in U$, un facile esercizio. \square

Esercizio 4.2.2. Se $U = \text{Span}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$ in \mathbb{R}^2 , le due rette affini

$$r_1 = U + \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad r_2 = U + \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

sono in realtà la stessa retta. Fate un disegno.

4.2.3. Supporto di un sottospazio affine. Sia $S = U + v_0$ un sottospazio affine. Il sottospazio vettoriale U è detto il *supporto* di S , e lo si ottiene traslando S all'indietro:

$$U = T_{-v_0}(S).$$

Questa operazione non dipende dal vettore $v_0 \in S$ che scegliamo. Indichiamo U con il simbolo $U = \text{supp}(S)$. Ricordiamo la notazione

$$\overrightarrow{PQ} = Q - P$$

che risulta particolarmente utile quando si studiano i sottospazi affini.

Proposizione 4.2.3. *Il supporto U di un sottospazio affine S è l'insieme*

$$U = \{\overrightarrow{PQ} \mid P, Q \in S\}$$

Dimostrazione. Sia $S = U + v_0$. Mostriamo l'inclusione \supseteq . Se $P, Q \in S$, allora $P = u_0 + v_0$ e $Q = u_1 + v_0$, quindi $\overrightarrow{PQ} = Q - P = u_1 - u_0 \in U$.

Mostriamo l'inclusione \subseteq . Se $u \in U$, allora prendendo $P = u + v_0$ e $Q = v_0$ otteniamo $\overrightarrow{PQ} = u$. \square

La *dimensione* di un sottospazio affine S è per definizione la dimensione del suo supporto $U = \text{supp}(S)$.

4.2.4. Forma parametrica e cartesiana. Siamo particolarmente interessati ai sottospazi affini dello spazio euclideo \mathbb{R}^n . Come accade in molti ambiti geometrici, un sottospazio affine può essere descritto in forma *cartesiana* (cioè *implicita*, come soluzioni di alcune equazioni) oppure *parametrica* (cioè *esplicita*, tramite alcuni parametri).

Sia $A \in M(m, n)$ e $b \in \mathbb{R}^m$. Il sistema lineare

$$S = \{x \in \mathbb{R}^n \mid Ax = b\}$$

descrive, quando ci sono soluzioni, un sottospazio affine S in *forma cartesiana*. Sappiamo per il Teorema di Rouché–Capelli che

$$S = \ker A + x_0$$

dove x_0 è una qualsiasi soluzione del sistema. Da questa scrittura vediamo subito che S è un sottospazio affine con supporto $\ker A$. Quindi

$$\dim S = \dim \ker A = n - \text{rk} A.$$

Una *forma parametrica* per S è invece una descrizione del tipo

$$S = P + U = \{P + t_1 v_1 + \dots + t_k v_k \mid t_1, \dots, t_k \in \mathbb{R}\}$$

dove $P \in S$ è un punto qualsiasi di S e $v_1, \dots, v_k \in \mathbb{R}^n$ sono una base per il supporto U . Qui ovviamente $k = \dim S$.

I sottospazi affini di \mathbb{R}^3 sono i punti, le rette, i piani, e \mathbb{R}^3 stesso.

Esempio 4.2.4. Un piano affine π in \mathbb{R}^3 è descritto da una equazione

$$\pi = \{ax + by + cz = d\}.$$

Una retta affine r in \mathbb{R}^3 è descritta da due equazioni di questo tipo, oppure più agevolmente in forma parametrica:

$$r = \left\{ \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + t \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right\}.$$

4.2.5. Esempi. Per due punti distinti in \mathbb{R}^3

$$P_0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}, \quad P_1 = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$$

passa un'unica retta r . Questa ha equazione parametrica

$$r = \{P_0 + t \overrightarrow{P_0 P_1}\}.$$

Per tre punti distinti $P_0, P_1, P_2 \in \mathbb{R}^3$ non allineati (cioè non contenuti in un'unica retta affine) passa un solo piano π . Questo ha equazione parametrica

$$\pi = \{P_0 + t \overrightarrow{P_0 P_1} + u \overrightarrow{P_0 P_2}\}.$$

Possiamo trasformare agevolmente le equazioni parametriche di un piano π in cartesiane nel modo seguente. Consideriamo un piano

$$\pi = \{P_0 + tv_1 + uv_2\}$$

dove $v_1, v_2 \in \mathbb{R}^3$ sono vettori. Calcoliamo il vettore $u = v_1 \times v_2$. Questo avrà tre coefficienti a, b, c . Il piano π può essere descritto come

$$\{ax + by + cz = d\}$$

per un certo $d \in \mathbb{R}$ che può essere trovato imponendo che $P \in \pi$.

Esempio 4.2.5. Consideriamo

$$\pi = \left\{ \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + u \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \right\}.$$

Troviamo

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \times \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

Quindi

$$\pi = \{x + 2y - z = d\}$$

per qualche $d \in \mathbb{R}$, che determiniamo imponendo

$$\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} \in \pi \implies 1 + 4 + 3 = d$$

e quindi $d = 8$. Infine otteniamo una equazione cartesiana per π :

$$\pi = \{x + 2y - z = 8\}.$$

4.2.6. Traslare l'origine. Abbiamo visto nei capitoli precedenti che per studiare un determinato problema in \mathbb{R}^n è spesso conveniente non usare la base canonica ma *cambiare base*. Quando studiamo i sottospazi affini, possiamo anche scegliere di *traslare l'origine*. Questo vuol dire fissare un punto $P \in \mathbb{R}^n$ e prendere delle nuove coordinate x' che differiscono dalle precedenti x nel modo seguente:

$$x' = x - P.$$

In altre parole:

$$\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} - \begin{pmatrix} P_1 \\ \vdots \\ P_n \end{pmatrix}.$$

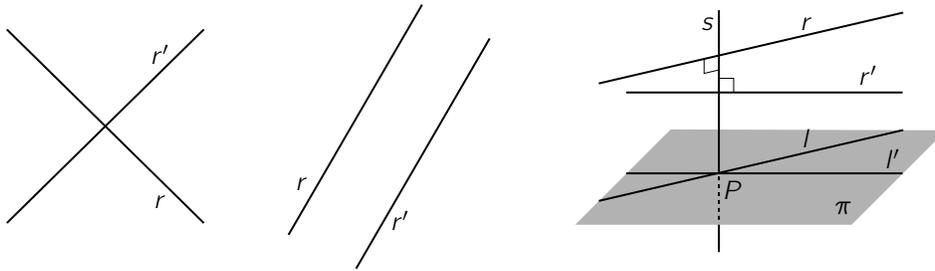


Figura 4.2. Due rette r, r' nello spazio possono essere incidenti (sinistra), parallele (centro) o sghembe (destra). Due rette sghembe hanno un'unica perpendicolare comune s (destra).

4.2.7. Intersezioni. Una peculiarità importante dei sottospazi affini, che li differenzia molto dai sottospazi vettoriali, sta nel fatto che i primi possono non intersecarsi, mentre gli ultimi devono sempre intersecarsi almeno nell'origine.

Due sottospazi affini $S, S' \subset V$ sono *incidenti* se $S \cap S' \neq \emptyset$. Due sottospazi affini incidenti si comportano un po' come i sottospazi vettoriali: se $S \cap S' \neq \emptyset$, è possibile traslare l'origine di V in modo che sia contenuta in $S \cap S'$; in questo modo S e S' diventano dei sottospazi vettoriali e possono essere studiati molto più agevolmente. Ad esempio, in questo modo verificiamo subito che $S \cap S'$ è un nuovo sottospazio affine.

Se S e S' non sono incidenti, possono comportarsi in modo molto diverso da quanto visto per i sottospazi vettoriali, come vedremo fra poco.

4.2.8. Posizioni. Due rette affini r e r' in \mathbb{R}^3 possono essere:

- *incidenti* se $r \cap r' \neq \emptyset$;
- *parallele* se $r \cap r' = \emptyset$ e $\text{supp}(r) = \text{supp}(r')$;
- *sghembe* se $r \cap r' = \emptyset$ e $\text{supp}(r) \neq \text{supp}(r')$.

Le tre possibilità sono disegnate in Figura 4.2.

Proposizione 4.2.6. *Date due rette sghembe r, r' , esiste un'unica retta s ortogonale ad entrambe.*

Dimostrazione. Per ipotesi $\text{supp}(r)$ e $\text{supp}(r')$ sono due rette vettoriali distinte, quindi $\pi = \text{supp}(r) + \text{supp}(r')$ è un piano vettoriale. Siano l e l' le proiezioni ortogonali di r e r' sul piano π come in Figura 4.2-(destra). Le due rette l e l' si intersecano in un punto P . La retta s perpendicolare a π passante per P è ortogonale sia a r che a r' , si veda la Figura 4.2-(destra).

La retta s è unica: una retta ortogonale a r e r' deve essere ortogonale a π , e fra le rette ortogonali a π solo s interseca sia r che r' . \square

Un piano π ed una retta r nello spazio sono:

- *incidenti* se $\pi \cap r \neq \emptyset$,
- *paralleli* se $\pi \cap r = \emptyset$.

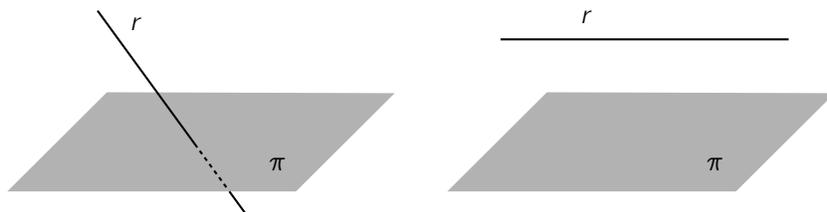


Figura 4.3. Una retta r ed un piano π possono essere incidenti (sinistra) o paralleli (destra).

I due casi sono mostrati in Figura 4.3.

Esercizio 4.2.7. Se π e r sono paralleli, allora $\text{supp}(r) \subset \text{supp}(\pi)$.

4.2.9. Fasci. Un *fascio* di rette nel piano o di piani nello spazio è un insieme (di rette o piani) determinati da un parametro variabile.

Rette. Il fascio di rette in \mathbb{R}^2 passanti per un punto $P = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ può essere descritto prendendo le equazioni cartesiane di due rette r_0, r_1 distinte che passano per P_0 :

$$r_0 = \{a_0x + b_0y + c_0 = 0\}, \quad r_1 = \{a_1x + b_1y + c_1 = 0\}$$

e considerando tutte le loro combinazioni lineari:

$$r_{t,u} = \{t(a_0x + b_0y + c_0) + u(a_1x + b_1y + c_1) = 0\}.$$

Al variare di $t, u \in \mathbb{R}$, con la condizione che non siano entrambi nulli, otteniamo in questo modo tutte le rette passanti per P_0 . Notiamo però che c'è una certa ridondanza nel prendere due parametri t e u : per ogni $\lambda \neq 0$, i parametri (t, u) e $(\lambda t, \lambda u)$ danno in realtà la stessa retta. A meno di moltiplicare per un λ opportuno, possiamo sempre chiedere che $t + u = 1$, in altre parole possiamo prendere un parametro solo t e mettere $u = 1 - t$.

Esempio 4.2.8. Il fascio di rette passanti per $P_0 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ può essere costruito in questo modo. Se ne prendono due:

$$\{x - 2 = 0\}, \quad \{y - 1 = 0\}$$

e si scrive

$$r_t = \{t(x - 2) + (1 - t)(y - 1) = 0\} = \{tx + (1 - t)y - t - 1 = 0\}.$$

Il fascio di rette parallele ad una retta data r è costruito ancora più semplicemente: si prende una equazione cartesiana $\{ax + by = c\}$ di r e si sostituisce il termine c con un parametro t .

Esempio 4.2.9. Il fascio di rette parallele a $\{x + 2y = 7\}$ è $\{x + 2y = t\}$.

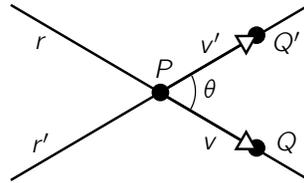


Figura 4.4. L'angolo fra due rette r e r' che si intersecano in un punto P è definito come l'angolo θ fra i vettori $v = Q - P$ e $v' = Q' - P$. Gli angoli così ottenuti sono in realtà due, θ e $\pi - \theta$, a seconda delle posizioni di Q e Q' .

Piani. Come sopra, il fascio di piani contenenti una retta r è costruito prendendo due equazioni cartesiane di due piani contenenti r e facendo una loro combinazione lineare. Il fascio di piani paralleli ad un piano con equazione $\{ax + by + cz = d\}$ si ottiene sostituendo d con un parametro t .

L'esempio seguente mostra come utilizzare agevolmente i fasci per risolvere alcuni problemi.

Esempio 4.2.10. Consideriamo la retta ed il punto

$$r = \begin{cases} x - y + 1 = 0, \\ 3x + y - z = 0, \end{cases} \quad P = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Per trovare il piano contenente r ed il punto P è sufficiente scrivere il fascio di piani contenenti r

$$\pi_t = \{t(x - y + 1) + (1 - t)(3x + y - z) = 0\}$$

ed imporre al piano generico del fascio che contenga P . In questo modo otteniamo $t + (1 - t)3 = 0$ e cioè $t = \frac{3}{2}$. Il piano cercato è $\pi_{\frac{3}{2}}$.

4.2.10. Angoli. Definiamo una nozione di angolo fra due spazi affini $A, A' \subset \mathbb{R}^3$ incidenti. Ci sono alcuni casi da considerare.

Angolo tra rette. Siano r e r' due rette che si intersecano in un punto P come in Figura 4.4. L'angolo θ fra r e r' è definito come l'angolo formato dai vettori $v = \overrightarrow{PQ}$ e $v' = \overrightarrow{P'Q'}$ dove $Q \in r$ e $Q' \in r'$ sono diversi da P . Questa definizione si presta ad una piccola ambiguità, perché a seconda dei punti Q, Q' che scegliamo possiamo ottenere un angolo θ oppure il suo supplementare $\pi - \theta$. Generalmente si sceglie il più piccolo dei due.

Angolo tra retta e piano. Siano r una retta e π un piano che si intersecano in un punto P come in Figura 4.5-(sinistra). Definiamo l'angolo θ fra r e π nel modo seguente: prendiamo un punto $Q \in r$ diverso da P e indichiamo con Q' la sua proiezione ortogonale su π . Se $Q' = P$, allora $r = \pi^\perp$ e poniamo $\theta = \frac{\pi}{2}$. Altrimenti l'angolo θ fra r e π è per definizione quello fra $v = \overrightarrow{PQ}$ e $v' = \overrightarrow{P'Q'}$.

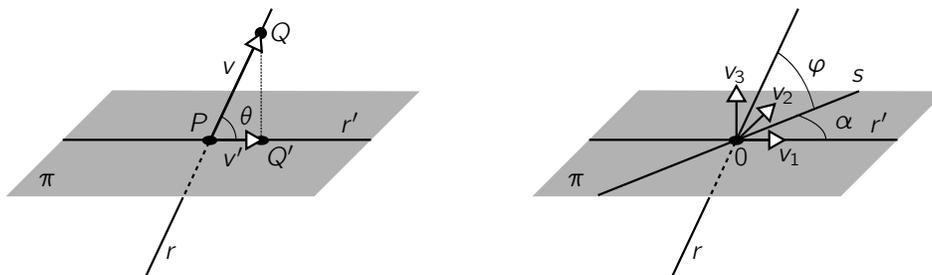


Figura 4.5. L'angolo θ fra una retta r ed un piano π che si intersecano in P è uguale all'angolo fra $v = Q - P$ e $v' = Q' - P'$ dove $Q \in r$ è un punto qualsiasi diverso da P e Q' è la sua proiezione ortogonale su π (sinistra). L'angolo θ fra r e r' è minore di quello φ fra r e s (destra).

In altre parole, proiettando r ortogonalmente su π otteniamo una retta r' e definiamo θ come l'angolo fra r e r' . Cosa succede se prendiamo un'altra retta $s \subset \pi$ al posto di r' come in Figura 4.5-(destra)?

Proposizione 4.2.11. *Sia $s \subset \pi$ una retta che interseca r diversa da r' . L'angolo φ fra s e r è maggiore di quello θ fra r' e r .*

Dimostrazione. A meno di traslare tutto, possiamo supporre che π e r si intersechino nell'origine. Scegliamo una base ortonormale $\mathcal{B} = \{v_1, v_2, v_3\}$ come nella Figura 4.5-(destra), tale che $v_1 \in r'$ e $v_1, v_2 \in \pi$. Consideriamo adesso due vettori $v \in r$ e $u \in s$ unitari. Valgono le seguenti uguaglianze:

$$v = \cos \theta v_1 + \sin \theta v_3, \quad u = \cos \alpha v_1 + \sin \alpha v_2$$

dove $\alpha > 0$ è l'angolo tra r' e s . Otteniamo

$$\cos \varphi = \langle v, u \rangle = \cos \theta \cos \alpha < \cos \theta.$$

La funzione coseno è monotona decrescente: quindi $\varphi > \theta$. \square

Angolo tra piani. Siano π_1 e π_2 due piani che si intersecano in una retta $r = \pi_1 \cap \pi_2$. L'angolo diedrale fra π_1 e π_2 è definito nel modo seguente: si prendono due rette $s_1 \subset \pi_1$ e $s_2 \subset \pi_2$ incidenti ed entrambe ortogonali a r come in Figura 4.6; l'angolo diedrale fra π_1 e π_2 è per definizione l'angolo fra s_1 e s_2 .

Esercizio 4.2.12. Siano r_1 e r_2 due rette incidenti e ortogonali a π_1 e π_2 rispettivamente. L'angolo fra π_1 e π_2 è uguale a quello fra r_1 e r_2 .

4.2.11. Distanze. Definiamo infine la *distanza* $d(A, A')$ fra due spazi affini A e A' . Se $A \cap A' \neq \emptyset$, la distanza è sempre zero. Se i sottospazi sono disgiunti la distanza è positiva ed è definita in modo diverso caso per caso.

Distanza fra punti. Come già sappiamo, la distanza $d(P, Q)$ fra due punti $P, Q \in \mathbb{R}^3$ è definita usando la norma:

$$d(P, Q) = \|\overrightarrow{PQ}\| = \|Q - P\|.$$

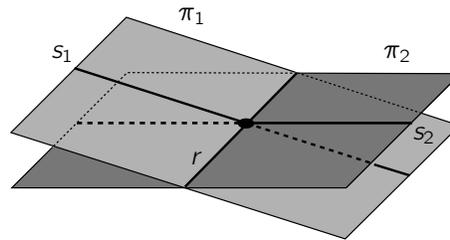


Figura 4.6. L'angolo fra due piani π_1 e π_2 che si intersecano in una retta r è definito come l'angolo fra due rette $s_1 \subset \pi_1$ e $s_2 \subset \pi_2$ incidenti entrambe ortogonali a r .

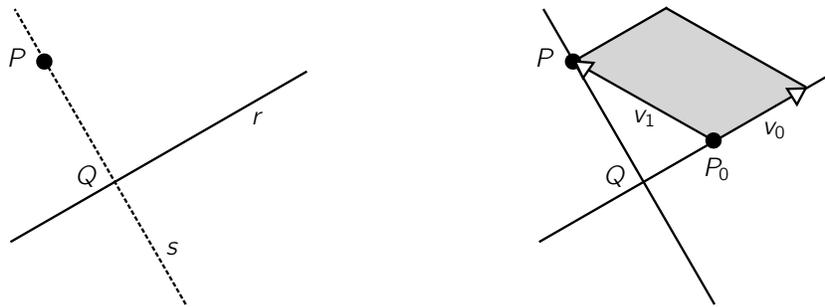


Figura 4.7. La distanza fra un punto P e una retta r è la lunghezza del segmento PQ ortogonale a r .

Distanza fra punto e retta. La distanza $d(P, r)$ fra un punto P ed una retta r nello spazio è definita nel modo seguente. Tracciamo come in Figura 4.7 la perpendicolare s a r nel punto P e definiamo

$$d(P, r) = d(P, Q)$$

dove $Q = r \cap s$. Se la retta r è espressa in forma parametrica $r = \{P_0 + tv_0\}$ come in Figura 4.7-(destra), la distanza si calcola facilmente usando il prodotto vettoriale.

Proposizione 4.2.13. *Vale l'uguaglianza*

$$d(P, r) = \frac{|v_0 \times v_1|}{|v_0|}$$

dove $v_1 = P - P_0$.

Dimostrazione. L'area del parallelogramma mostrato in Figura 4.7-(destra) può essere calcolata in due modi: con il prodotto vettoriale $|v_0 \times v_1|$ e come base per altezza $|v_0| \cdot d(P, Q)$. L'uguaglianza

$$|v_0 \times v_1| = |v_0| \cdot d(P, Q)$$

implica la formula per $d(P, r) = d(P, Q)$. □

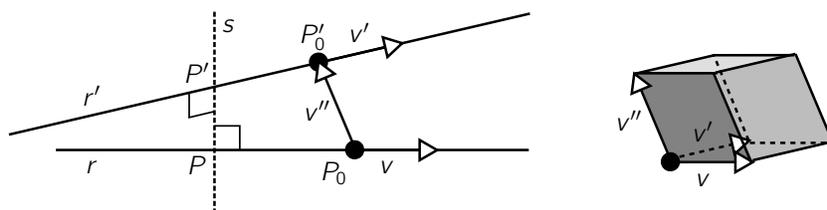


Figura 4.8. La distanza fra due rette sghembe.

Distanza fra rette. La distanza $d(r, r')$ fra due rette r ed r' disgiunte è definita nel modo seguente. Tracciamo una retta s perpendicolare ad entrambe, e definiamo

$$d(r, r') = d(P, P')$$

dove $P = r \cap s$ e $P' = r' \cap s'$. Anche in questo caso, se $r = \{P_0 + tv\}$ e $r' = \{P'_0 + tv'\}$ sono espresse in forma parametrica, la distanza si calcola facilmente (si veda la Figura 4.8).

Proposizione 4.2.14. Se r e r' sono sghembe, vale la formula

$$d(r, r') = \frac{|\det(v|v'|v'')|}{|v \times v'|}$$

dove $v'' = P'_0 - P_0$.

Dimostrazione. La dimostrazione è simile alla precedente. Il volume del parallelepipedo in Figura 4.8-(destra) generato dai vettori v , v' e v'' può essere calcolato in due modi: come $|\det(v|v'|v'')|$ e come area di base per altezza, cioè $|v \times v'| \cdot d$. L'altezza d è proprio $d = d(P, P') = d(r, r')$ e quindi si ottiene

$$|\det(v|v'|v'')| = d(r, r') \cdot |v \times v'|$$

che implica la formula per $d(r, r')$. \square

Distanza fra punto e piano. Infine, la distanza fra un punto P_0 ed un piano π è definita in modo simile a quanto già visto: si traccia la perpendicolare s a π passante per P_0 e si definisce

$$d(P_0, \pi) = d(P_0, Q)$$

con $Q = s \cap \pi$, si veda la Figura 4.9. Se

$$\pi = \{ax + by + cz = d\}, \quad P_0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$$

allora la distanza si calcola facilmente.

Proposizione 4.2.15. Vale

$$d(P_0, \pi) = \frac{ax_0 + by_0 + cz_0 - d}{\sqrt{a^2 + b^2 + c^2}}$$

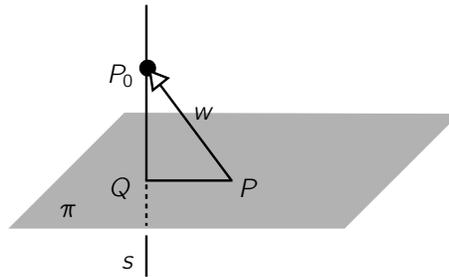


Figura 4.9. La distanza fra un punto P e un piano π .

Dimostrazione. Consideriamo

$$v = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \quad P = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

con $P \in \pi$ punto qualsiasi, come in Figura 4.9. Il vettore v è ortogonale a π . Sia $w = P_0 - P$. La figura mostra che

$$\begin{aligned} d(P_0, Q) &= \frac{|\langle v, w \rangle|}{|v|} = \frac{|a(x_0 - x) + b(y_0 - y) + c(z_0 - z)|}{\sqrt{a^2 + b^2 + c^2}} \\ &= \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}. \end{aligned}$$

Nell'ultima uguaglianza abbiamo usato che $P \in \pi$ e quindi $ax + by + z = d$. \square

4.3. Trasformazioni del piano e dello spazio

Studiamo qui alcune trasformazioni di \mathbb{R}^2 e \mathbb{R}^3 . Nei primi paragrafi studieremo le trasformazioni vettoriali, che fissano l'origine, poi passeremo a quelle affini in cui l'origine può essere spostata a piacimento. Iniziamo con lo studio delle isometrie vettoriali del piano.

4.3.1. Affinità. Tutti gli endomorfismi di \mathbb{R}^n per definizione devono fissare l'origine. Vogliamo adesso considerare alcune trasformazioni di \mathbb{R}^n che spostano l'origine in un punto arbitrario.

Una *affinità* di \mathbb{R}^n è una mappa $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ del tipo

$$f(x) = Ax + b$$

dove A è una matrice $n \times n$ e $b \in \mathbb{R}^n$ un vettore. Chiamiamo la matrice A la *componente lineare* di f ed il vettore b la *componente traslatoria*. Notiamo che la componente traslatoria è l'immagine del vettore nullo:

$$f(0) = A0 + b = b.$$

Due esempi fondamentali di affinità sono gli endomorfismi

$$f(x) = Ax$$

in cui $b = 0$, e le *traslazioni*

$$f(x) = x + b$$

in cui $A = I$. Qualsiasi affinità $f(x) = Ax + b$ è la composizione $f = h \circ g$ di un endomorfismo $g(x) = Ax$ e di una traslazione $h(x) = x + b$.

Una affinità $f(x) = Ax + b$ è un *isomorfismo affine* se A è una matrice invertibile.

Proposizione 4.3.1. *Sia $f(x) = Ax + b$ un isomorfismo affine. Valgono i fatti seguenti:*

- (1) *la funzione $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ è biunivoca;*
- (2) *la sua inversa è $g(x) = A^{-1}x - A^{-1}b$;*
- (3) *la funzione f manda sottospazi affini in sottospazi affini;*
- (4) *gli isomorfismi affini formano un gruppo con la composizione.*

Dimostrazione. Moltiplicando entrambi i membri dell'equazione $y = Ax + b$ per A^{-1} troviamo $A^{-1}y = x + A^{-1}b$ e quindi

$$x = A^{-1}y - A^{-1}b.$$

Questo mostra in particolare che la funzione f è invertibile e quindi biunivoca. Abbiamo dimostrato (1) e (2).

Per la (3), ricordiamo che un sottospazio affine è un insieme del tipo $U + v$ con U sottospazio vettoriale. L'immagine

$$f(U + v) = A(U + v) + b = AU + Av + b$$

è un sottospazio affine $U' + v'$, con $U' = AU$ e $v' = Av + b$.

Gli isomorfismi affini formano un gruppo, perché la composizione di due isomorfismi è un isomorfismo e ogni isomorfismo ha un inversa. \square

Gli isomorfismi affini preservano anche il parallelismo: mandano rette parallele in rette parallele. Non preservano però necessariamente le distanze! A questo scopo introduciamo una nuova definizione.

4.3.2. Isometrie affini. Una *isometria affine* è un isomorfismo affine $f(x) = Ax + b$ che preserva le distanze, cioè tale che

$$d(f(P), f(Q)) = d(P, Q)$$

per ogni coppia di punti $P, Q \in \mathbb{R}^n$.

Esempio 4.3.2. Una traslazione $f(x) = x + b$ è una isometria affine, infatti $d(f(P), f(Q)) = d(P + b, Q + b) = \|(P + b) - (Q + b)\| = \|P - Q\| = d(P, Q)$.

Proposizione 4.3.3. *Le isometrie affini formano un gruppo con la composizione.*

Dimostrazione. L'inversa di una isometria è una isometria, e componendo due isometrie si ottiene una isometria. \square

Proposizione 4.3.4. *Una affinità $f(x) = Ax + b$ è una isometria affine \iff la matrice A è ortogonale.*

Dimostrazione. (\Leftarrow) Se A è ortogonale, le affinità $g(x) = Ax$ e $h(x) = x + b$ sono isometrie e quindi anche la composizione $f = h \circ g$ lo è.

(\Rightarrow) Se $f(x) = Ax + b$ è una isometria, allora

$$\|x\| = d(x, 0) = d(f(x), f(0)) = d(Ax + b, b) = \|Ax + b - b\| = \|Ax\|$$

per ogni $x \in \mathbb{R}^n$. Per la Proposizione 3.4.1 la matrice A è ortogonale. \square

4.3.3. Traslare l'origine. Quando si studiano le affinità è a volte conveniente traslare l'origine in un punto P , ottenendo quindi nuove coordinate

$$x' = x - P.$$

Se scriviamo una affinità nelle nuove coordinate x' come

$$f'(x') = Ax' + b,$$

la stessa affinità nelle coordinate originarie x sarà allora

$$f(x) = A(x - P) + b + P.$$

Possiamo anche scriverla in altri modi:

$$f(x) = Ax - AP + b + P = Ax + (I - A)P + b.$$

Notiamo che la parte lineare A della affinità è sempre la stessa, mentre la componente traslatoria è cambiata da b a $(I - A)P + b$.

Esempio 4.3.5. Sia $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ una rotazione di angolo θ intorno ad un punto $P = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$. Mostriamo adesso che f è una isometria affine e la scriviamo esplicitamente come $f(x) = Ax + b$.

Nel sistema di riferimento $x' = x - P$, la trasformazione è semplicemente $f'(x') = \text{Rot}_\theta x'$. Quindi nel sistema di riferimento iniziale x otteniamo

$$f(x) = \text{Rot}_\theta x + (I - \text{Rot}_\theta)P.$$

Esplicitamente:

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} (1 - \cos \theta)x_0 + \sin \theta y_0 \\ -\sin \theta x_0 + (1 - \cos \theta)y_0 \end{pmatrix}.$$

Come controllo, si può verificare ad esempio che $f(P) = P$.

Esercizio 4.3.6. Scrivi nella forma $f(x) = Ax + b$ la riflessione di \mathbb{R}^2 rispetto alla retta affine $x + y = 1$.

4.3.4. Punti fissi. Un punto fisso per una affinità $f(x) = Ax + b$ è un punto $P \in \mathbb{R}^n$ tale che $f(P) = P$.

Ad esempio, se $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ è una rotazione intorno a $P \in \mathbb{R}^2$, allora P è un punto fisso. D'altra parte, notiamo che una traslazione $f(x) = x + b$ con $b \neq 0$ non ha nessun punto fisso.

Se f ha un punto fisso P , è naturale prendere le coordinate $x' = x - P$, perché in queste nuove coordinate il punto fisso P è la nuova origine 0 e quindi l'affinità f letta nelle nuove coordinate diventa un endomorfismo:

$$f'(x') = Ax'.$$

In molti casi possiamo capire subito se c'è un punto fisso:

Proposizione 4.3.7. *Se 1 non è autovalore per A , l'affinità $f(x) = Ax + b$ ha almeno un punto fisso.*

Dimostrazione. Un punto $P \in \mathbb{R}^n$ è fisso se e solo se

$$AP + b = P \iff (A - I)P = -b.$$

Se 1 non è autovalore per A , non esiste un vettore $v \neq 0$ tale che $Av = v$, e quindi $\ker(A - I) = \{0\}$. Quindi $A - I$ è un isomorfismo ed in particolare è surgettiva: esiste un P per cui $(A - I)P = -b$. \square

4.3.5. Isometrie affini positive del piano. Classifichiamo adesso tutte le isometrie affini del piano con determinante positivo.

Proposizione 4.3.8. *Una isometria affine $f(x) = Ax + b$ con $\det A = 1$ è una traslazione oppure una rotazione intorno ad un punto $P \in \mathbb{R}^2$.*

Dimostrazione. Se f ha un punto fisso P , a meno di traslare l'origine possiamo supporre che $P = 0$, quindi l'isometria diventa vettoriale e la Proposizione 3.4.12 ci dice che f è una rotazione intorno a P .

Se f non ha punti fissi, allora A ha autovalore 1. Sempre per la Proposizione 3.4.12, deduciamo che $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ e quindi $f(x) = x + b$ è una traslazione. \square

Questa proposizione è semplice ma molto potente. Sia $f(x) = Ax + b$ una isometria affine del piano con $\det A = 1$. Sappiamo che $A = \text{Rot}_\theta$ è una matrice di rotazione; allora possiamo concludere che:

- se $\theta = 0$, cioè $A = I$, allora f è una traslazione oppure l'identità;
- se $\theta \neq 0$, allora f è una rotazione, il cui centro P può essere calcolato risolvendo l'equazione $AP + b = P$.

Da questo semplice fatto deduciamo ad esempio che:

Corollario 4.3.9. *Siano f e g due isometrie affini positive di \mathbb{R}^2 .*

- (1) *se f e g sono traslazioni, allora $f \circ g$ è una traslazione oppure l'identità;*

- (2) se f è una rotazione di angolo θ e g è una traslazione, allora $f \circ g$ è una rotazione di angolo θ ;
- (3) se f e g sono rotazioni di angolo θ e φ e $\theta + \varphi \neq 2k\pi$, allora $f \circ g$ è una rotazione di angolo $\theta + \varphi$.

Dimostrazione. Tutti i casi sono dimostrati notando che la componente rotatoria di $f \circ g$ è il prodotto delle componenti rotatorie di f e di g . \square

Esercizio 4.3.10. Scrivi due rotazioni f e g la cui composizione $f \circ g$ è una traslazione.

4.3.6. Classificazione delle isometrie affini. Vogliamo classificare adesso tutte le isometrie affini di \mathbb{R}^2 e \mathbb{R}^3 . Per questo scopo avremo bisogno di dimostrare alcuni fatti. Iniziamo con la proposizione seguente.

Proposizione 4.3.11. Se A è una matrice ortogonale, allora

$$\ker(A - I)^\perp = \text{Im}(A - I).$$

Dimostrazione. Siano $v \in \ker(A - I)$ e $w \in \mathbb{R}^n$ qualsiasi. Dobbiamo mostrare che $\langle v, (A - I)w \rangle = 0$. Per ipotesi $Av = v$ e otteniamo

$$\langle v, (A - I)w \rangle = \langle v, Aw \rangle - \langle v, w \rangle = \langle Av, Aw \rangle - \langle v, w \rangle = 0$$

perché A è una isometria vettoriale. \square

Se A è ortogonale, possiamo generalizzare la Proposizione 4.3.7 nel modo seguente.

Lemma 4.3.12. Sia $f(x) = Ax + b$ una isometria affine. A meno di traslare l'origine possiamo supporre che $b \in \ker(A - I)$.

Dimostrazione. Un cambiamento di coordinate del tipo $x = x' + d$ trasforma l'isometria affine f in

$$f'(x') = Ax' + (A - I)d + b.$$

Il nostro scopo è cercare un d per cui $b' = (A - I)d + b$ appartenga al sottospazio $\ker(A - I)$. Per la Proposizione 4.3.11 abbiamo

$$\mathbb{R}^n = \ker(A - I) \oplus \text{Im}(A - I).$$

Quindi il vettore b si scompone come $b = b_1 + b_2$ con $b_1 \in \ker(A - I)$ e $b_2 \in \text{Im}(A - I)$. Per ipotesi $b_2 = (A - I)v$ per qualche v . Basta prendere $d = -v$ e si ottiene

$$b' = (A - I)d + b = b - (A - I)v = b - b_2 = b_1 \in \ker(A - I).$$

La dimostrazione è completa. \square

Il lemma è effettivamente una generalizzazione della Proposizione 4.3.7, infatti se 1 non è autovalore per A allora $\ker(A - I) = \{0\}$ e quindi $b = 0$. Utilizziamo adesso il lemma per classificare le isometrie affini di \mathbb{R}^2 e \mathbb{R}^3 .

4.3.7. Isometrie affini del piano. Classifichiamo ora le isometrie affini di \mathbb{R}^2 . Come nel caso vettoriale, ci sono due tipi di isometrie $f(x) = Ax + b$, quelle con $\det A = 1$ che preservano l'orientazione del piano e quelle con $\det A = -1$ che la invertono.

Sappiamo già che le isometrie di \mathbb{R}^2 che preservano l'orientazione sono le traslazioni e le rotazioni intorno a qualche punto di \mathbb{R}^2 . Fra quelle che invertono l'orientazione troviamo le riflessioni rispetto a qualche retta affine r e le *glissoriflessioni*: una glissoriflessione è la composizione di una riflessione lungo una retta affine r con una traslazione non banale parallela a r . Ad esempio

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_1 \\ 0 \end{pmatrix} = \begin{pmatrix} x + b_1 \\ -y \end{pmatrix}$$

è una glissoriflessione ottenuta componendo una riflessione rispetto all'asse x con una traslazione orizzontale di passo $b_1 \in \mathbb{R}$ con $b_1 \neq 0$.

Proposizione 4.3.13. *Ogni isometria affine $f(x) = Ax + b$ di \mathbb{R}^2 è di uno dei quattro tipi seguenti:*

- se $\det A = 1$, è una traslazione o rotazione;
- se $\det A = -1$, è una riflessione o glissoriflessione.

Dimostrazione. Se f ha un punto fisso P , a meno di traslare l'origine possiamo supporre che $P = 0$ e la Proposizione 3.4.12 ci dice che f è una rotazione intorno a P o una riflessione rispetto ad una retta r contenente il punto P .

Se f non ha punti fissi, allora A ha autovalore 1. Sempre per la Proposizione 3.4.12, a meno di ruotare gli assi (cioè di prendere un'altra base ortonormale) otteniamo $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ oppure $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Nel primo caso f è una traslazione. Nel secondo, per il Lemma 4.3.12 possiamo supporre che $b \in \ker(A - I) = \ker \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix} = \text{Span}(e_1)$. Quindi $b = \begin{pmatrix} b_1 \\ 0 \end{pmatrix}$ e f è una glissoriflessione. \square

Notiamo che le isometrie affini del piano senza punti fissi sono precisamente le traslazioni e le glissoriflessioni.

4.3.8. Isometrie dello spazio. La classificazione delle isometrie affini dello spazio è analoga a quella del piano, con un paio di casi in più da considerare. Una *rototraslazione* è la composizione di una rotazione intorno ad un asse r e di una traslazione parallela a r . Ad esempio

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} b_1 \\ 0 \\ 0 \end{pmatrix}$$

è una rototraslazione ottenuta componendo una rotazione di angolo θ intorno all'asse x con una traslazione orizzontale di passo $b_1 \in \mathbb{R}$ con $b_1 \neq 0$.

Una *glissoriflessione* di \mathbb{R}^3 è la composizione di una riflessione rispetto ad un piano π e di una traslazione lungo una direzione parallela a π . Ad esempio

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ b_2 \\ b_3 \end{pmatrix}$$

è una glissoriflessione ottenuta riflettendo rispetto al piano $\pi = \{x = 0\}$ e poi trasladando di un vettore b parallelo a π .

Proposizione 4.3.14. *Ogni isometria affine $f(x) = Ax + b$ di \mathbb{R}^3 è di uno dei tipi seguenti:*

- se $\det A = 1$, è una traslazione, rotazione o rototraslazione;
- se $\det A = -1$, è una riflessione, antirotazione o glissoriflessione.

Dimostrazione. Sappiamo che A è ortogonale e quindi rappresenta l'identità, una rotazione, o una antirotazione. Se $A = I$ allora f è una traslazione. Se A è una rotazione intorno ad un asse r , allora per il Lemma 4.3.12 possiamo prendere b parallelo a r e quindi otteniamo una rototraslazione.

Se A è una antirotazione, ci sono due casi. Se l'antirotazione ha angolo zero, è in realtà una riflessione rispetto ad un piano π e per il Lemma 4.3.12 possiamo prendere b parallelo a π e ottenere una glissoriflessione. Se l'angolo è diverso da zero, allora A non ha autovalore 1 e quindi ha un punto fisso: allora f è una antirotazione. \square

Teorema spettrale

5.1. Prodotti hermitiani

In tutto il Capitolo 3 abbiamo introdotto i prodotti scalari soltanto per spazi vettoriali reali. Mostriamo adesso brevemente come sia possibile definire una nozione analoga anche per gli spazi vettoriali complessi: questa è nota con il nome di *prodotto hermitiano*.

5.1.1. Definizione. Sia V uno spazio vettoriale complesso.

Definizione 5.1.1. Un *prodotto hermitiano* su V è una applicazione

$$\begin{aligned} V \times V &\longrightarrow \mathbb{C} \\ (v, w) &\longmapsto \langle v, w \rangle \end{aligned}$$

che soddisfa i seguenti assiomi:

- (1) $\langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle$,
- (2) $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$,
- (3) $\langle v, w \rangle = \overline{\langle w, v \rangle}$

per ogni $v, v', w \in V$ e ogni $\lambda \in \mathbb{R}$.

Notiamo immediatamente che l'assioma (3) differisce da quello analogo per i prodotti scalari per la presenza di un coniugio. Il motivo di questo cambiamento sarà presto chiaro. Intanto da questi assiomi deduciamo facilmente altre proprietà:

- (4) $\langle v, w + w' \rangle = \langle v, w \rangle + \langle v, w' \rangle$,
- (5) $\langle v, \lambda w \rangle = \bar{\lambda} \langle v, w \rangle$,
- (6) $\langle 0, w \rangle = \langle v, 0 \rangle = 0$

per ogni $v, w, w' \in V$ e $\lambda \in \mathbb{R}$. Il coniugio in (3) è stato inserito per ottenere il fatto seguente:

- (7) $\langle v, v \rangle$ è un numero reale, per ogni $v \in V$.

Per dimostrare ciò, notiamo che per (3) abbiamo

$$\langle v, v \rangle = \overline{\langle v, v \rangle}$$

e quindi effettivamente $\langle v, v \rangle \in \mathbb{R}$.

5.1.2. Prodotto hermitiano definito positivo. In un prodotto hermitiano, è effettivamente cruciale che $\langle v, v \rangle$ sia un numero reale: in questo modo è perfettamente sensato chiedere che questo numero sia positivo o negativo, a seconda delle necessità. Possiamo quindi introdurre questa definizione, come abbiamo fatto nel caso reale.

Definizione 5.1.2. Un prodotto hermitiano è *definito positivo* se $\langle v, v \rangle > 0$ per ogni $v \neq 0$.

In un prodotto hermitiano definito positivo si definiscono in modo del tutto analogo al caso reale molti dei concetti già visti precedentemente: la *norma* di un vettore è

$$\|v\| = \sqrt{\langle v, v \rangle},$$

si possono quindi introdurre le nozioni di *spazio ortogonale* e di *base ortonormale* ed il procedimento di Gram – Schmidt funziona anche in questo contesto. Il *prodotto hermitiano euclideo* su \mathbb{C}^n è dato da

$$\langle x, y \rangle = {}^t x \bar{y}.$$

Come nel caso reale, il prodotto hermitiano euclideo è definito positivo.

5.1.3. Matrice associata. Come nel caso reale, se V è un prodotto hermitiano e $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di V , possiamo definire la *matrice associata* H nel modo seguente:

$$H_{ij} = \langle v_i, v_j \rangle$$

per ogni i, j . Abbiamo scelto la lettera H invece di S , perché questa matrice non è simmetrica: si tratta in realtà di una *matrice hermitiana*, cioè una matrice quadrata complessa per cui

$${}^t H = \bar{H}.$$

Qui \bar{H} è la matrice ottenuta sostituendo tutti i valori H_{ij} con i loro coniugati complessi. Ad esempio, la matrice seguente è hermitiana:

$$\begin{pmatrix} 2 & 1+i \\ 1-i & 1 \end{pmatrix}.$$

Notiamo che in una matrice hermitiana gli elementi H_{ii} sulla diagonale devono necessariamente essere reali, perché $H_{ii} = \bar{H}_{ii}$.

5.2. Endomorfismi autoaggiunti

Introduciamo in questo capitolo una importante classe di endomorfismi. In tutta questa sezione, indichiamo con V uno spazio vettoriale reale dotato di un prodotto scalare definito positivo, oppure uno spazio vettoriale complesso dotato di un prodotto hermitiano definito positivo.

5.2.1. Definizione. Un endomorfismo $T: V \rightarrow V$ è *autoaggiunto* se

$$(9) \quad \langle T(v), w \rangle = \langle v, T(w) \rangle$$

per ogni $v, w \in V$.

Notiamo subito che questa definizione ricorda un po' quella di isometria, in cui si chiede che $\langle v, w \rangle = \langle T(v), T(w) \rangle$. Ci sono effettivamente delle analogie fra le due definizioni, ma anche delle importanti differenze.

5.2.2. Con le matrici. Come per le isometrie, cerchiamo subito di tradurre la definizione di endomorfismo autoaggiunto in una condizione concreta sulle matrici associate.

Scegliamo una base ortonormale \mathcal{B} di V . Sia $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ la matrice associata ad un endomorfismo $T: V \rightarrow V$.

Proposizione 5.2.1. L'endomorfismo T è autoaggiunto se e solo se la matrice A è hermitiana.

Dimostrazione. La dimostrazione è del tutto analoga a quella della Proposizione 3.4.4. Per bilinearità, per verificare se T è autoaggiunto è sufficiente testare la condizione (9) sugli elementi v_i, v_j della base \mathcal{B} . Su questi due elementi la condizione (9) diventa

$$A_{ij} = \bar{A}_{ji}$$

per ogni i, j , in altre parole ${}^t A = \bar{A}$. □

In particolare, otteniamo il fatto seguente.

Corollario 5.2.2. Sia A una matrice $n \times n$. L'endomorfismo $L_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$ è autoaggiunto rispetto al prodotto hermitiano euclideo di \mathbb{C}^n se e solo se la matrice A è hermitiana.

Analogamente, se A è reale, l'endomorfismo $L_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ è autoaggiunto rispetto al prodotto scalare euclideo di \mathbb{R}^n se e solo se la matrice A è simmetrica (una matrice reale è hermitiana \iff è simmetrica).

Notiamo che la Proposizione 5.2.1 è valida solo se prendiamo una base \mathcal{B} ortonormale, come mostra questo esempio.

Esempio 5.2.3. Consideriamo \mathbb{R}^2 con il prodotto scalare euclideo. L'endomorfismo L_A definito dalla matrice $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ è autoaggiunto perché la matrice A è simmetrica e stiamo usando la base canonica che è ortonormale.

Se scriviamo L_A rispetto ad un'altra base ortonormale, ad esempio $\mathcal{B} = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\}$, otteniamo una nuova matrice simmetrica:

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

Se però scriviamo A rispetto ad una base *non* ortonormale, ad esempio $\mathcal{B} = \left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$, la nuova matrice associata può non essere simmetrica:

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

5.2.3. Sottospazi invarianti. Mostriamo che gli endomorfismi autoaggiunti si comportano con i sottospazi invarianti in modo simile alle isometrie.

Proposizione 5.2.4. *Sia $T: V \rightarrow V$ un endomorfismo autoaggiunto e $U \subset V$ un sottospazio. Se $T(U) \subset U$ allora $T(U^\perp) \subset U^\perp$.*

Dimostrazione. Prendiamo un vettore qualsiasi $v \in U^\perp$ e dobbiamo dimostrare che $T(v) \in U^\perp$. Per ogni vettore $u \in U$ otteniamo

$$\langle T(v), u \rangle = \langle v, T(u) \rangle = 0$$

perché $T(u) \in U$ per ipotesi. Quindi $T(v)$ è ortogonale a tutti i vettori $u \in U$, in altre parole $T(v) \in U^\perp$. \square

5.3. Il teorema

5.3.1. Enunciato. Sia come sempre V uno spazio vettoriale complesso dotato di un prodotto hermitiano definito positivo, oppure uno spazio reale dotato di un prodotto scalare definito positivo. Il seguente è uno dei più importanti teoremi di algebra lineare.

Teorema 5.3.1 (Teorema spettrale). *Un endomorfismo $T: V \rightarrow V$ è autoaggiunto \iff ha una base ortonormale di autovettori e tutti i suoi autovalori sono in \mathbb{R} .*

Dimostrazione. (\Leftarrow) Se T ha una base \mathcal{B} ortonormale di autovettori, la matrice $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ è diagonale; poiché gli autovalori sono reali, la matrice A è diagonale con numeri reali sulla diagonale e quindi è hermitiana. Per la Proposizione 5.2.1 l'endomorfismo T è autoaggiunto.

(\Rightarrow) Dimostriamo il teorema innanzitutto nel caso complesso. Iniziamo verificando che tutti gli autovalori di T sono reali: se λ è autovalore, allora c'è un autovettore v e $T(v) = \lambda v$, quindi

$$\lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle T(v), v \rangle = \langle v, T(v) \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \langle v, v \rangle.$$

Poiché $\langle v, v \rangle > 0$, deduciamo che $\lambda = \bar{\lambda}$ e quindi $\lambda \in \mathbb{R}$.

Dimostriamo adesso per induzione sulla dimensione n di V che T ha una base ortonormale di autovettori. Il caso $n = 1$ è banale: tutti gli endomorfismi di uno spazio di dimensione uno hanno una base ortonormale di autovettori, formata da un elemento solo di norma uno.

Supponiamo il fatto vero per la dimensione $n - 1$ e lo mostriamo per n . Poiché siamo sui complessi, esiste sempre un autovettore $v \in V$. La retta $\text{Span}(v)$ è chiaramente T -invariante e quindi anche lo spazio ortogonale

$U = \text{Span}(v)^\perp$ è T -invariante, per la Proposizione 5.2.4, cioè vale $T(U) \subset U$. La restrizione

$$T|_U: U \rightarrow U$$

è un endomorfismo autoaggiunto di uno spazio vettoriale U di dimensione $n-1$. Per l'ipotesi induttiva, esiste una base v_2, \dots, v_n ortonormale formata da autovettori per $T|_U$. Aggiungendo il vettore iniziale v , rinormalizzato in modo che abbia norma uno, otteniamo una base $\mathcal{B} = \{v, v_2, \dots, v_n\}$ ortonormale formata da autovettori di T .

Resta infine da mostrare il caso reale. □

Possiamo riscrivere il teorema nel linguaggio delle matrici. Consideriamo \mathbb{R}^n con l'usuale prodotto scalare euclideo.

Corollario 5.3.2. *Sia A una matrice $n \times n$ reale. Sono fatti equivalenti:*

- (1) A è simmetrica;
- (2) L_A ha una base ortonormale di autovettori;
- (3) esiste una matrice ortogonale M tale che

$${}^tMAM = M^{-1}AM = D$$

sia una matrice diagonale.

Dimostrazione. L'equivalenza (1) \Leftrightarrow (2) è il teorema spettrale. L'esistenza di una base \mathcal{B} di autovettori per A è equivalente all'esistenza di una matrice M tale che $M^{-1}AM = D$ sia diagonale. Inoltre \mathcal{B} è ortonormale se e solo se M è ortogonale, quindi (2) \Leftrightarrow (3). Notiamo che $M^{-1} = {}^tM$ in questo caso. □

Riassumendo, abbiamo scoperto due fatti non banali: il primo è che ogni matrice simmetrica reale è diagonalizzabile, ed il secondo è che sia possibile scegliere una base di autovettori che è anche ortonormale. Solo le matrici simmetriche hanno entrambe queste proprietà.

5.3.2. Segnatura. Il teorema spettrale ha varie conseguenze inaspettate, ed una di queste è un metodo definitivo per calcolare la segnatura (i_+, i_-, i_0) di una matrice simmetrica S .

Proposizione 5.3.3. *Sia S una matrice simmetrica e (i_+, i_-, i_0) la sua segnatura. I numeri i_+ , i_- e i_0 sono pari al numero di autovalori positivi, negativi e nulli di S .*

Dimostrazione. Sappiamo dalla Proposizione 5.3.2 che esiste una matrice ortogonale M per cui

$${}^tMAM = M^{-1}AM = D$$

sia diagonale. Le matrici A e D sono simultaneamente simili e congruenti! La matrice D ha la stessa segnatura di A perché sono congruenti, e gli stessi autovalori di A perché sono simili. Quindi è sufficiente mostrare la proposizione per D , ma per le diagonali la proposizione è vera per l'Esercizio 3.2.11. □

Abbiamo apparentemente trovato un metodo infallibile per determinare la segnatura di una matrice simmetrica S : è sufficiente determinare i suoi autovalori. Ma come troviamo gli autovalori di S ? Questi sono le radici del polinomio caratteristico $p_S(x)$, ma sfortunatamente non esistono formule risolutive per trovare le radici di un polinomio quando questo ha grado ≥ 5 , e di fatto l'unica formula gestibile è quella in grado 2 che si impara alle superiori.

Per fortuna non è necessario determinare precisamente gli autovalori, è sufficiente il loro *segno*. A questo scopo possiamo usare il *criterio di Cartesio*, che non dimostriamo:

Teorema 5.3.4 (Criterio di Cartesio). *Sia $p(x) = a_n x^n + \dots + a_m x^m$ un polinomio con n radici reali, scritto in modo che il primo e l'ultimo coefficiente a_n e a_m siano entrambi non nulli. Valgono i fatti seguenti:*

- (1) *il numero di radici nulle di $p(x)$ è m ;*
- (2) *il numero di radici positive di $p(x)$ è pari al numero di cambiamenti di segno nella sequenza a_n, \dots, a_m dei coefficienti, da cui siano stati rimossi i coefficienti nulli.*

Ad esempio, il polinomio $p(x) = x^4 - x^2$ ha due radici nulle (perché $m = 2$) e una radice positiva perché la sequenza dei coefficienti è $1, -1$ e contiene una variazione. Effettivamente sappiamo che le radici sono $0, 0, 1, -1$.

Il numero di radici negative è n meno il numero di radici nulle e positive (attenzione: questo non è necessariamente il numero di permanenze di segno nella successione dei coefficienti!)

Con il criterio di Cartesio, è sempre possibile determinare la segnatura di una matrice esaminando i segni dei coefficienti del suo polinomio caratteristico. Notiamo però che il calcolo esplicito del polinomio caratteristico di una matrice $n \times n$ può essere dispendioso: quando funziona, il criterio di Jacobi spiegato nella Sezione 3.2.6 è più immediato.

Geometria proiettiva

6.1. Introduzione

Introduciamo in questo capitolo una nuova geometria, chiamata *geometria proiettiva*. La geometria proiettiva è un arricchimento della geometria affine, realizzato aggiungendo allo spazio euclideo \mathbb{R}^n alcuni “punti all’infinito”. Questo arricchimento presenta numerosi vantaggi, sia algebrici che geometrici. Renderemo in particolare rigoroso l’assunto informale che due rette parallele “si incontrano all’infinito”.

6.1.1. Lo spazio proiettivo. Possiamo descrivere lo *spazio proiettivo n -dimensionale* \mathbb{P}^n nel modo seguente. Ogni punto $P \in \mathbb{P}^n$ è descritto con un simbolo di questo tipo:

$$[x_1, \dots, x_{n+1}]$$

e ci sono due regole fondamentali:

- (1) i valori x_i non sono tutti nulli, cioè il simbolo $[0, \dots, 0]$ è vietato;
- (2) due simboli $[x_1, \dots, x_{n+1}]$ e $[\lambda x_1, \dots, \lambda x_{n+1}]$ indicano lo stesso punto, per ogni $\lambda \neq 0$.

I valori x_1, \dots, x_{n+1} sono le *coordinate omogenee* di P . Ad esempio, nel piano proiettivo \mathbb{P}^2 ogni punto ha tre coordinate omogenee $[x_1, x_2, x_3]$, però le coordinate $[1, -3, 2]$ e $[5, -15, 10]$ indicano in realtà lo stesso punto: ogni punto ha molti nomi. Il punto $[0, 0, 0]$ però non è ammesso, non esiste.

6.1.2. I punti al finito e all’infinito. Consideriamo adesso un punto $[x_1, \dots, x_{n+1}]$ nello spazio proiettivo \mathbb{P}^n . Si possono presentare due casi:

- Se $x_{n+1} \neq 0$, allora possiamo usare la regola (1) e dividere ogni coefficiente per x_{n+1} . Quindi il punto ha anche coordinate omogenee

$$\left[\frac{x_1}{x_{n+1}}, \dots, \frac{x_n}{x_{n+1}}, 1 \right].$$

In questo caso diciamo che il punto è *al finito*.

- Se $x_{n+1} = 0$, diciamo che il punto è *all’infinito*.

In altre parole, usando la regola (1) possiamo sempre supporre che l’ultima coordinata x_{n+1} sia 1 oppure 0, una specie di variabile booleana che ci dice se il punto è al finito o all’infinito.

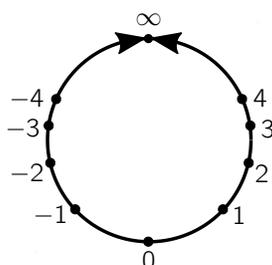


Figura 6.1. La retta proiettiva \mathbb{P}^1 è ottenuta aggiungendo un punto all'infinito alla retta reale \mathbb{R} .

Notiamo adesso che i punti al finito corrispondono precisamente ad \mathbb{R}^n . Possiamo infatti interpretare un punto di \mathbb{R}^n con coordinate x_1, \dots, x_n come il punto al finito

$$[x_1, \dots, x_n, 1].$$

La filosofia che sta dietro alla costruzione di \mathbb{P}^n è adesso chiara: tramite l'introduzione di una nuova coordinata x_{n+1} , che di fatto possiamo supporre che sia solo 1 o 0, abbiamo aggiunto ad \mathbb{R}^n dei punti all'infinito, che sono della forma

$$[x_1, \dots, x_n, 0].$$

Cerchiamo adesso di studiare i punti all'infinito nei casi più semplici.

6.1.3. La retta proiettiva. Studiamo la retta proiettiva \mathbb{P}^1 . La retta proiettiva \mathbb{P}^1 è formata da punti del tipo $[x_1, x_2]$. Come abbiamo visto, a meno di moltiplicare entrambe le coordinate per x_2^{-1} possiamo supporre che x_2 sia 1 oppure 0. Nel primo caso, otteniamo i punti al finito $[x, 1]$ che corrispondono ai punti $x \in \mathbb{R}$ della retta reale \mathbb{R} .

Nel secondo caso, otteniamo i punti all'infinito $[x_1, 0]$. Ricordiamo adesso che il punto $[0, 0]$ non è ammesso: quindi $x_1 \neq 0$, e a meno di dividere per x_1 possiamo supporre che il punto abbia coordinate $[1, 0]$. Quindi abbiamo scoperto che in \mathbb{P}^1 esiste un solo punto all'infinito $[1, 0]$, che chiamiamo ∞ .

Riassumendo, la retta proiettiva \mathbb{P}^1 è ottenuta aggiungendo ad \mathbb{R} un punto solo $\infty = [1, 0]$. La retta proiettiva è in realtà una specie di circonferenza come in Figura 6.1.

6.1.4. Il piano proiettivo. Consideriamo adesso il piano proiettivo \mathbb{P}^2 . I punti al finito sono del tipo $[x_1, x_2, 1]$ e corrispondono ai punti (x_1, x_2) del piano \mathbb{R}^2 . I punti all'infinito sono invece del tipo $[x_1, x_2, 0]$.

A differenza di \mathbb{P}^1 , qui non abbiamo un solo punto all'infinito, ma ne abbiamo infiniti. Ad esempio, i punti seguenti

$$[1, 0, 0], \quad [1, 1, 0], \quad [1, 2, 0], \quad [0, 1, 0]$$

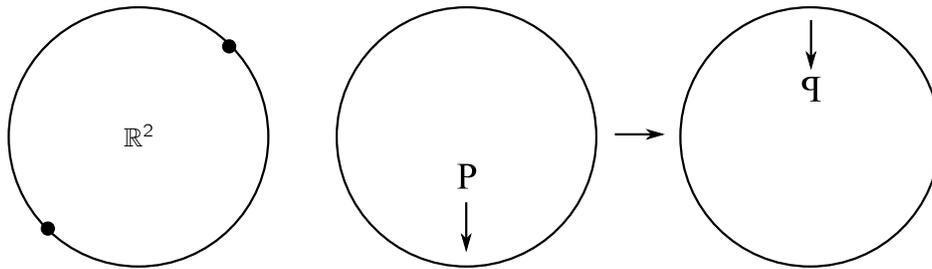


Figura 6.2. Il piano proiettivo \mathbb{P}^2 è ottenuto aggiungendo i “punti all’infinito” a \mathbb{R}^2 . Possiamo pensare a \mathbb{R}^2 come ad un disco senza bordo, cui abbiamo aggiunto dei punti di bordo: due punti opposti sono però lo stesso punto (sinistra). È possibile per un oggetto attraversare l’infinito e trovarsi quindi nella parte opposta del disco: attenzione però che l’attraversamento causa uno specchiamento dell’oggetto (destra).

sono tutti all’infinito e sono diversi fra loro. Qual è il loro significato geometrico? Geometricamente, il punto $[x_1, x_2, 0]$ è un punto all’infinito che troviamo nella direzione indicata dal vettore

$$v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Dobbiamo cioè pensare che se prendiamo una qualsiasi retta affine del tipo

$$r = P + tv$$

e spostiamo il punto sulla retta mandando t a $\pm\infty$, troviamo all’infinito proprio il punto $[x_1, x_2, 0]$. Notiamo, come nel caso di \mathbb{P}^1 , che troviamo lo stesso punto $[x_1, x_2, 0]$ se percorriamo la retta affine in entrambe le direzioni, cioè con $t \rightarrow +\infty$ e $t \rightarrow -\infty$. Notiamo anche che due rette affini puntano sullo stesso punto all’infinito se e solo se sono parallele: questo ci permetterà di dire nella prossima sezione che due rette parallele si intersecano all’infinito proprio in questo punto.

Abbiamo disegnato \mathbb{P}^1 come una circonferenza nella Figura 6.1. Qual è la forma di \mathbb{P}^2 ? La risposta è più complicata e cerchiamo di fornire un’idea nella Figura 6.2: pensiamo ad \mathbb{R}^2 come ad un disco senza bordo, cui aggiungiamo i punti di bordo, con l’accortezza però che due punti opposti rappresentano lo stesso punto. In questo modo l’infinito può essere attraversato da qualsiasi oggetto, come mostrato in Figura 6.2: il lettore può concentrarsi e convincersi però che dopo l’attraversamento l’oggetto risulta specchiato, come in figura.

Si può rappresentare la forma di \mathbb{P}^3 alla stessa maniera, disegnando \mathbb{R}^3 come una sfera solida tridimensionale senza bordo e aggiungendo i punti di bordo, in cui però due punti antipodali risultano identificati. Il lettore può verificare che in questo caso l’attraversamento dell’infinito *non* causa uno specchiamento dell’oggetto!

6.2. Sottospazi e proiettività

Ogni spazio che si rispetti ha i suoi sottospazi e le sue mappe, e lo spazio proiettivo \mathbb{P}^n non fa eccezione.

6.2.1. Sottospazi proiettivi. Un *sottospazio proiettivo* di \mathbb{P}^n è un luogo di punti $S \subset \mathbb{P}^n$ descritto da un sistema di k equazioni lineari omogenee:

$$S = \{[x_1, \dots, x_{n+1}] \in \mathbb{P}^n \mid Ax = 0\}.$$

dove A è una matrice $k \times (n+1)$ e x ha coordinate x_1, \dots, x_{n+1} . Ad esempio, una retta proiettiva r in \mathbb{P}^2 è descritta da una equazione

$$a_1x_1 + a_2x_2 + a_3x_3 = 0.$$

Osservazione 6.2.1. Ricordiamo una peculiarità dello spazio proiettivo: ciascun punto ha più nomi, quindi quando scriviamo delle equazioni dobbiamo stare attenti. Le equazioni lineari omogenee (cioè senza termine noto) del tipo $a_1x_1 + \dots + a_nx_n = 0$ in \mathbb{P}^n hanno senso, perché se sono soddisfatte da $[x_1, \dots, x_n]$ allora sono soddisfatte anche da $[\lambda x_1, \dots, \lambda x_n]$. Un'equazione affine non omogenea (cioè con un termine noto) del tipo $x_1 + x_2 - x_3 = 3$ invece *non* ha senso nello spazio proiettivo! Il punto $[1, 1, -1]$ la soddisfa e $[2, 2, -2]$ no, però indicano lo stesso punto: l'equazione non ha senso.

Definiamo la *dimensione* del sottospazio proiettivo $S \subset \mathbb{P}^n$ come

$$\dim S = n - \text{rk}A.$$

Esempio 6.2.2. Nel piano proiettivo, i punti

$$P_1 = [1, 0, 0], \quad P_2 = [0, 1, 0], \quad P_3 = [0, 0, 1]$$

e le rette

$$r_1 = \{x_1 = 0\}, \quad r_2 = \{x_2 = 0\}, \quad r_3 = \{x_3 = 0\}$$

sono disegnati in Figura 6.3-(sinistra). Le tre rette tagliano il piano proiettivo in quattro triangoli, indicati con colori differenti in Figura 6.3-(destra). Quando si fanno i disegni, a volte non cerchiamo di distinguere troppo i punti al finito e all'infinito, perché nella geometria proiettiva si comportano tutti allo stesso modo: ad esempio, in questo caso i punti in r_3 in realtà starebbero tutti all'infinito, ma li abbiamo disegnati come se fossero al finito.

Come nelle altre geometrie che abbiamo visto, possiamo chiamare la *codimensione* di S il numero $n - \dim S$ e *iperpiano* un sottospazio proiettivo di codimensione uno (ad esempio, una retta di \mathbb{P}^2 oppure un piano in \mathbb{P}^3).

Come già notato nell'esempio, l'equazione $x_{n+1} = 0$ in \mathbb{P}^n descrive precisamente tutti i punti all'infinito. Abbiamo quindi scoperto che i punti all'infinito formano un iperpiano, detto *iperpiano all'infinito*. Nel piano proiettivo \mathbb{P}^2 i punti all'infinito formano la *retta all'infinito* di \mathbb{P}^2 , di equazione $x_3 = 0$.

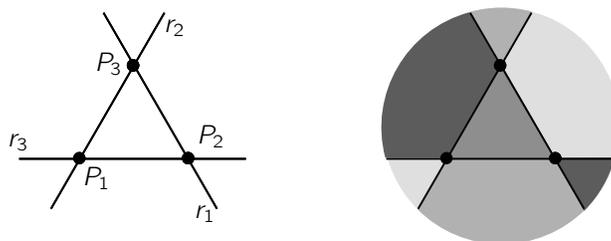


Figura 6.3. Tre punti e tre rette in \mathbb{P}^2 (sinistra). Le rette tagliano il piano proiettivo in quattro triangoli (destra).

6.2.2. Forma parametrica. Come in geometria affine, è possibile descrivere un sottospazio proiettivo anche in forma parametrica. Ad esempio, la retta passante per due punti

$$P = [x_1, \dots, x_{n+1}], \quad [y_1, \dots, y_{n+1}]$$

può essere scritta in forma parametrica come l'insieme di punti del tipo

$$[tx_1 + uy_1, \dots, tx_{n+1} + uy_{n+1}]$$

dove t, u sono variabili reali non entrambe nulle. Questa forma si estende agevolmente anche a sottospazi di dimensione maggiore.

Esempio 6.2.3. La retta r in \mathbb{P}^2 passante per $P = [1, 1, 0]$ e $Q = [0, 1, -1]$ è data dai punti

$$r = \{[t, t + u, -u]\}.$$

Si passa facilmente ad una scrittura cartesiana notando che

$$r = \{-x_1 + x_2 + x_3 = 0\}.$$

Il piano π in \mathbb{P}^3 passante per i tre punti non allineati $P = [1, -1, 0, 0]$, $Q = [1, 0, 1, 0]$ e $R = [0, 0, 0, 1]$ in forma parametrica è dato da tutti i punti del tipo

$$\pi = \{[t + u, -t, u, s]\}$$

al variare dei parametri t, u, s non tutti nulli. In forma cartesiana diventa

$$\pi = \{x_1 + x_2 - x_3 = 0\}.$$

6.2.3. Completamento di un sottospazio affine. Mostriamo adesso che i sottospazi affini di \mathbb{R}^n diventano sottospazi proiettivi di \mathbb{P}^n tramite un procedimento di *completamento*, che consiste nell'aggiunta dei loro punti all'infinito.

Consideriamo uno spazio affine $S \subset \mathbb{R}^n$ descritto da un sistema di equazioni $Ax = b$ di \mathbb{R}^n . Come sottolineato nell'Osservazione 6.2.1, queste equazioni non sono omogenee e quindi non hanno senso in \mathbb{P}^n . Rimediamo facilmente, trasformandole in equazioni omogenee nel modo seguente.

Il *completamento* di S in \mathbb{P}^n è il sottospazio proiettivo \bar{S} di \mathbb{P}^n ottenuto sostituendo ciascuna equazione $a_{i1}x_1 + \dots + a_{in}x_n = b_i$ che definisce S con la sua versione omogenea

$$a_{i1}x_1 + \dots + a_{in}x_n = b_i x_{n+1}$$

in cui il termine noto b_i diventa $b_i x_{n+1}$. Questo processo è chiamato *omogeneizzazione* delle equazioni. Il completamento \bar{S} è il luogo di zeri del sistema omogeneo che otteniamo.

Il completamento \bar{S} contiene S , perché se (x_1, \dots, x_n) è soluzione del primo sistema, allora $[x_1, \dots, x_n, 1]$ è soluzione del secondo. Il completamento \bar{S} aggiunge ad S i suoi "punti all'infinito", cioè le direzioni di tutte le rette contenute in S .

Esempio 6.2.4. Sia $r \subset \mathbb{R}^2$ la retta affine descritta dall'equazione non omogenea $x_1 + x_2 - 2 = 0$. Il suo completamento $\bar{r} \subset \mathbb{P}^2$ è la retta proiettiva

$$x_1 + x_2 - 2x_3 = 0$$

ottenuta sostituendo il termine noto 2 con $2x_3$. La retta r è formata da tutti i punti del tipo $\begin{pmatrix} t \\ 2-t \end{pmatrix}$. Il completamento \bar{r} contiene questi punti, nella forma $[t, 2-t, 1]$, ed in più anche il punto all'infinito $[1, -1, 0]$. Quindi

$$\bar{r} = r \cup \{[1, -1, 0]\}.$$

Il punto all'infinito $[1, -1, 0]$ è quello che corrisponde alla direzione di r .

Esempio 6.2.5. Il piano $x_1 = 1$ in \mathbb{R}^3 si completa al piano $x_1 = x_4$ in \mathbb{P}^3 . I punti all'infinito che abbiamo aggiunto sono quelli che soddisfano questa equazione e anche $x_4 = 0$. Questi sono precisamente tutti i punti del tipo

$$[0, t, u, 0].$$

Questa è una retta contenuta nel piano all'infinito. In generale, completando un piano si aggiunge la sua retta di punti all'infinito.

Nel piano proiettivo non ci sono rette parallele: due rette distinte si intersecano sempre in un punto.

Esempio 6.2.6. Due rette affini parallele

$$a_1x_1 + a_2x_2 = b_1, \quad a_1x_1 + a_2x_2 = b_2$$

con $b_1 \neq b_2$ non si intersecano in \mathbb{R}^2 , ma i loro completamenti

$$a_1x_1 + a_2x_2 = b_1x_3, \quad a_1x_1 + a_2x_2 = b_2x_3$$

si intersecano nel piano proiettivo \mathbb{P}^2 nel punto all'infinito $[a_2, -a_1, 0]$.

6.2.4. Proiettività. Le affinità sono trasformazioni particolari dello spazio affine. Analogamente, le *proiettività* sono trasformazioni particolari dello spazio proiettivo.

Sia A una matrice $(n+1) \times (n+1)$. La *proiettività* definita da A è la mappa $f: \mathbb{P}^n \rightarrow \mathbb{P}^n$ che manda un punto di coordinate proiettive x nel punto di coordinate proiettive Ax . Se A è invertibile, allora f è invertibile e viene detta *isomorfismo proiettivo*.

Esempio 6.2.7. La matrice

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

descrive la proiettività

$$[x_1, x_2, x_3] \mapsto [x_1 + 2x_2, x_2, -2x_3]$$

del piano proiettivo \mathbb{P}^2 .

Le proiettività di \mathbb{P}^n formano un gruppo rispetto alla composizione. Notiamo che due matrici A e λA con $\lambda \in \mathbb{R}$ non nullo descrivono la stessa proiettività.

Esercizio 6.2.8. Scrivi una proiettività $T: \mathbb{P}^2 \rightarrow \mathbb{P}^2$ tale che $T(P_1) = P_2$, $T(P_2) = P_3$ e $T(P_3) = P_1$ dove P_1, P_2, P_3 sono i punti dell'Esempio 6.2.2.

6.2.5. Punti fissi. Notiamo un fatto interessante.

Proposizione 6.2.9. *Un punto $P = [x_1, \dots, x_{n+1}]$ è punto fisso per una proiettività descritta da $A \in M(n+1, \mathbb{R})$ se e soltanto se*

$$v = \begin{pmatrix} x_1 \\ \vdots \\ x_{n+1} \end{pmatrix}$$

è autovettore per A .

Dimostrazione. Il punto P è fisso se e solo se $[v] = [Av]$, cioè se e solo se Av è multiplo di v . \square

Esempio 6.2.10. La matrice A descritta nell'Esempio 6.2.7 ha due autovalori 1 e -2 , con autospazi

$$V_1 = \text{Span} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad V_{-2} = \text{Span} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Quindi i punti fissi in \mathbb{P}^2 della proiettività descritta da A sono due: $[1, 0, 0]$ e $[0, 0, 1]$. D'altra parte, se

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

allora gli autospazi di B sono

$$V_1 = \{x_3 = 0\}, \quad V_{-2} = \text{Span} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

I punti fissi in \mathbb{P}^2 della proiettività descritta da B sono la retta $r = \{x_3 = 0\}$ ed il punto $[0, 0, 1]$.

Dalla teoria degli autovettori che abbiamo sviluppato nei capitoli precedenti possiamo dedurre facilmente un fatto non banale.

Proposizione 6.2.11. *Ogni proiettività di \mathbb{P}^2 ha almeno un punto fisso.*

Dimostrazione. Ogni matrice 3×3 ha almeno un autovettore. \square

La proposizione è abbastanza sorprendente perché non è valida in geometria affine: le traslazioni $x \mapsto x + b$ sono affinità di \mathbb{R}^2 senza punti fissi. Il fatto è che le traslazioni fissano tutti i punti all'infinito. Esaminiamo meglio questo fenomeno.

6.2.6. Da affinità a proiettività. Nell'ultimo paragrafo abbiamo suggerito che una traslazione in \mathbb{R}^n possa essere considerata una proiettività. Si tratta di un fenomeno generale: qualsiasi trasformazione affine $f(x) = Ax + b$ di \mathbb{R}^n può essere *completata* ad una trasformazione proiettiva, usando la matrice $(n+1) \times (n+1)$ data da

$$A' = \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}.$$

Questa proiettività effettivamente completa l'affinità originale, perché i punti finiti del tipo $[x, 1]$, vanno effettivamente in $[Ax + b, 1]$. I punti all'infinito, di tipo $[x, 0]$, invece vanno in $[Ax, 0]$. In particolare, se f è una traslazione, allora $A = I$ e quindi i punti all'infinito sono tutti punti fissi.

Se A è invece una rotazione in \mathbb{R}^2 intorno ad un punto, i punti all'infinito vengono tutti ruotati. Se A è una rototraslazione in \mathbb{R}^3 , il suo completamento ha un solo punto fisso all'infinito: il punto che corrisponde alla direzione dell'asse.

Esercizio 6.2.12. Scrivi una proiettività di \mathbb{P}^3 senza punti fissi.

CAPITOLO 7

Quadriche

7.1. Introduzione

Nei capitoli precedenti abbiamo studiato molti oggetti *lineari*, cioè definiti usando equazioni lineari, ovvero di primo grado; in questo capitolo iniziamo lo studio di oggetti non lineari, partendo da quelli più semplici: quelli *quadratici*, definiti usando equazioni di secondo grado. Questi oggetti comprendono le *coniche* nel piano, quali ellissi, iperboli e parabole.

7.1.1. Polinomi. Conosciamo bene i polinomi $p(x)$ in una variabile x , ed introduciamo adesso i polinomi a più variabili.

Siano x_1, \dots, x_n delle variabili. Un *monomio* è un prodotto

$$\lambda x_{i_1} \cdots x_{i_k}$$

dove $\lambda \in \mathbb{R}$ è un numero non nullo e i_1, \dots, i_k sono indici qualsiasi tra 1 e n . Gli indici possono ripetersi e k è il *grado* del monomio. Ad esempio

$$27, \quad 7x_2, \quad 2x_1x_3, \quad \pi x_1^2x_4, \quad -x_2^2x_3^2$$

sono monomi di grado 0, 1, 2, 3, 4, rispettivamente. Un *polinomio* $p(x)$ è una somma di monomi: lo indichiamo sempre con $p(x)$ per semplicità, ma ci ricordiamo che in realtà x è un vettore di variabili x_1, \dots, x_n , cioè un punto generico di \mathbb{R}^n . Il *grado* di $p(x)$ è il massimo grado dei monomi che lo compongono. Ad esempio

$$p(x) = 3x_1x_3^2 - x_1x_2 + 4x_3^2 - x_1 + 7$$

ha grado 3. Un polinomio è *omogeneo* se tutti i suoi monomi hanno lo stesso grado. Il polinomio $p(x)$ appena descritto non è omogeneo, invece

$$q(x) = 2x_1x_2 - x_3^2 + x_1x_3$$

è omogeneo.

7.1.2. Varietà affini. Se $p(x)$ è un polinomio in una sola variabile x , chiamiamo *radice* di p un punto $x \in \mathbb{R}$ per cui $p(x) = 0$. Dalla teoria dei polinomi sappiamo che, se p non è il polinomio nullo, allora ha al più d radici, dove d è il grado di p . Se $p(x)$ ha grado uno o due, sappiamo anche determinare le radici tramite formule esplicite.

Se $p(x)$ è un polinomio a più variabili x_1, \dots, x_n , la situazione ovviamente si complica. Il *luogo degli zeri* di $p(x)$, cioè l'insieme dei punti $x \in \mathbb{R}^n$ in cui

$p(x)$ si annulla, è chiamato *varietà affine*: questo insieme è spesso infinito e difficile da studiare.

Se $p(x)$ ha grado uno, allora $p(x) = a_1x_1 + \dots + a_nx_n + b$ ed il luogo degli zeri di $p(x)$ è per definizione un sottospazio affine: otteniamo un oggetto che abbiamo già studiato e che sappiamo trattare senza problemi.

Se $p(x)$ ha grado due, otteniamo precisamente l'oggetto di studio di questo capitolo.

Definizione 7.1.1. Una *quadrica affine* è il luogo di zeri in \mathbb{R}^n di un polinomio $p(x)$ di grado due.

Nelle prossime pagine useremo gli strumenti dell'algebra lineare per studiare e classificare le quadriche affini. La nostra impostazione sarà la seguente: studiamo prima il problema analogo nello spazio proiettivo \mathbb{P}^n , dove lo studio è *più facile* che in \mathbb{R}^n . L'aggiunta di punti all'infinito infatti semplificherà molto la discussione. Poi passeremo al caso affine.

7.2. Classificazione proiettiva

Introduciamo in questa sezione le quadriche proiettive e le classifichiamo.

7.2.1. Varietà proiettive. Se $p(x)$ è un polinomio qualsiasi nelle variabili x_1, \dots, x_{n+1} , possiamo cercare di interpretarlo nello spazio proiettivo \mathbb{P}^n dove ogni punto ha coordinate omogenee $[x_1, \dots, x_{n+1}]$. Non è però detto che l'equazione $p(x) = 0$ abbia senso in questo contesto, perché ricordiamo che in \mathbb{P}^n ogni punto ha molti nomi e tutto deve essere coerente. Ad esempio, il polinomio

$$p(x_1, x_2, x_3) = x_1x_2 - x_3$$

sembra annullarsi nel punto $[1, 1, 1]$, però non si annulla nel punto $[2, 2, 2]$: questo è un problema, perché i due simboli rappresentano in realtà lo stesso punto... il luogo degli zeri di $p(x)$ in \mathbb{P}^2 non è ben definito.

Per fortuna ci vengono in aiuto i polinomi omogenei.

Proposizione 7.2.1. Se $p(x)$ ha grado d ed è omogeneo, allora $p(\lambda x) = \lambda^d p(x)$ per ogni $\lambda \in \mathbb{R}$.

Dimostrazione. Tutti i monomi di $p(x)$ hanno lo stesso grado d . Ogni monomio è della forma

$$\alpha x_{i_1} \cdots x_{i_d}$$

e se sostituiamo λx al posto di x , cioè mettiamo λx_i al posto di x_i per ogni i , questo monomio si trasforma in

$$\alpha \lambda^d x_{i_1} \cdots x_{i_d}$$

e quindi viene moltiplicato per λ^d . Mettendo in evidenza il fattore λ^d che spunta da ciascun monomio otteniamo $p(\lambda x) = \lambda^d p(x)$. \square

Corollario 7.2.2. Se $p(x)$ è omogeneo e $p(x) = 0$, allora $p(\lambda x) = 0$ per ogni $\lambda \in \mathbb{R}$.

Se $p(x)$ è un polinomio omogeneo nelle variabili x_1, \dots, x_{n+1} , il luogo di zeri di $p(x)$ nello spazio proiettivo \mathbb{P}^n è ben definito ed è chiamato *varietà proiettiva*. Avevamo già notato questo fatto nel caso in cui p abbia grado uno: in questo caso $p(x) = a_1x_1 + \dots + a_{n+1}x_{n+1}$ e il luogo di zeri di $p(x)$ è un iperpiano proiettivo.

Passiamo al grado due.

7.2.2. Quadriche proiettive. Una *quadrica proiettiva* è il luogo di zeri in \mathbb{P}^n di un polinomio $p(x)$ omogeneo di grado due. Il polinomio $p(x)$ può essere scritto come

$$(10) \quad p(x) = \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} \alpha_{ij} x_i x_j.$$

Questa scrittura è un po' ridondante, perché se $i \neq j$ il termine $x_i x_j$ compare due volte nella sommatoria, come $\alpha_{ij} x_i x_j$ e come $\alpha_{ji} x_j x_i$. Supponiamo per comodità che $\alpha_{ij} = \alpha_{ji}$.

Sia $A \in M(n+1)$ la matrice con $A_{ij} = \alpha_{ij}$. Possiamo riscrivere l'espressione (10) semplicemente come

$$(11) \quad p(x) = {}^t x A x.$$

Poiché $\alpha_{ij} = \alpha_{ji}$, la matrice A è simmetrica.

Esempio 7.2.3. Il polinomio omogeneo

$$p(x) = x_1^2 + 4x_1x_2 - x_2^2 + 4x_3^2$$

può essere scritto come

$$p(x) = (x_1, x_2, x_3) \begin{pmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Si noti come il monomio $4x_1x_2$ viene spezzato in due monomi $2x_1x_2 + 2x_2x_1$ per ottenere una matrice simmetrica.

La forma (11) è particolarmente comoda. Consideriamo una proiettività $x = Mx'$ di \mathbb{P}^n determinata da una matrice invertibile $M \in M(n+1)$. La proiettività trasforma il polinomio $p(x) = {}^t x A x$ in un nuovo polinomio

$$p'(x') = p(Mx') = {}^t (Mx') A (Mx') = {}^t x' ({}^t M A M) x'.$$

Quindi la matrice A che descrive la conica cambia per congruenza in una nuova matrice $A' = {}^t M A M$.

7.2.3. Classificazione delle quadriche proiettive. Diciamo che due quadriche sono *proiettivamente equivalenti* se esiste una proiettività che trasforma una nell'altra.

Teorema 7.2.4 (Classificazione proiettiva delle quadriche). *Ogni quadrica in \mathbb{P}^n è proiettivamente equivalente ad una del tipo*

$$(12) \quad x_1^2 + \dots + x_k^2 - x_{k+1}^2 - \dots - x_{k+h}^2 = 0$$

per qualche $k, h \geq 0$ con $1 \leq k + h \leq n + 1$.

Dimostrazione. Sia $p(x) = {}^t xAx$ una quadrica. Per il Corollario 3.2.8 esiste una M invertibile per cui $D = {}^t MAM$ è una matrice diagonale con valori $+1, -1, 0$ sulla diagonale. La matrice M descrive una proiettività che trasforma la quadrica nella forma cercata. \square

L'equazione (12) è la *forma canonica* della quadrica ${}^t xAx = 0$ ed è determinata direttamente da A nel modo seguente: la terna $(k, h, n - (k + h))$ è la segnatura della matrice simmetrica A .

Notiamo che se moltiplichiamo tutta l'equazione (12) per -1 otteniamo un'altra forma canonica che descrive la stessa quadrica (il luogo di zeri non cambia) con k e h invertiti.

La quadrica è *non degenera* se $k + h = n + 1$, in altre parole se l'indice di nullità di A è zero. Ricordiamo che questo è equivalente a $\det A \neq 0$.

Passiamo ora a studiare più da vicino le quadriche in $\mathbb{P}^1, \mathbb{P}^2$ e \mathbb{P}^3 .

7.2.4. Quadriche in \mathbb{P}^1 . Studiare una quadrica ${}^t xAx = 0$ in \mathbb{P}^1 è molto simile a studiare le soluzioni di un polinomio di secondo grado in una variabile: praticamente è la stessa cosa, l'unica differenza è che qui fra le possibili soluzioni c'è anche il punto all'infinito. Otteniamo 0, 1 oppure 2 punti, a seconda del determinante di A che gioca lo stesso ruolo del Δ in una equazione di secondo grado.

Corollario 7.2.5. *Ogni quadrica proiettiva in \mathbb{P}^1 è equivalente a una di queste:*

- (1) *l'insieme vuoto* $x_1^2 + x_2^2 = 0$ se $\det A > 0$;
- (2) *due punti* $x_1^2 - x_2^2 = 0$ se $\det A < 0$;
- (3) *un punto* $x_1^2 = 0$ se $\det A = 0$.

Dimostrazione. Le possibili segnature di A sono

$$(2, 0, 0), \quad (1, 1, 0), \quad (1, 0, 1).$$

Ricordiamo che le segnature $(k, h, 2 - (k + h))$ e $(h, k, 2 - (k + h))$ rappresentano la stessa quadrica. Otteniamo quindi le tre forme canoniche elencate. Nella prima non ci sono soluzioni, nella seconda $x_1^2 = x_2^2$ ce ne sono due e sono $[1, 1]$ e $[1, -1]$, nella terza $x_1^2 = 0$ otteniamo solo $[0, 1]$. \square

7.2.5. Coniche proiettive. Una *conica proiettiva* è una quadrica proiettiva nel piano \mathbb{P}^2 . Queste sono classificate grazie al Teorema 7.2.4.

Corollario 7.2.6. *Ogni conica proiettiva è equivalente a una di queste:*

- (1) l'insieme vuoto $x_1^2 + x_2^2 + x_3^2 = 0$,
- (2) la circonferenza $x_1^2 + x_2^2 - x_3^2 = 0$,
- (3) il punto $x_1^2 + x_2^2 = 0$,
- (4) le due rette incidenti $x_1^2 - x_2^2 = 0$,
- (5) la retta doppia $x_1^2 = 0$.

Dimostrazione. Le possibili segnature di A sono

$$(3, 0, 0), \quad (2, 1, 0), \quad (2, 0, 1), \quad (1, 1, 1), \quad (1, 0, 2).$$

Ricordiamo che le segnature $(k, h, 3-(k+h))$ e $(h, k, 3-(k+h))$ rappresentano la stessa conica. Otteniamo quindi le cinque forme canoniche elencate.

Studiamo i cinque casi individualmente:

- (1) l'insieme è vuoto perché l'equazione non ha soluzioni in \mathbb{P}^2 ;
- (2) l'equazione implica $x_3 \neq 0$ e quindi possiamo porre $x_3 = 1$; la conica è contenuta interamente in $\mathbb{R}^2 \subset \mathbb{P}^2$, nella forma

$$x_1^2 + x_2^2 = 1$$

che è ovviamente una circonferenza;

- (3) l'unica soluzione in \mathbb{P}^2 è il punto $[0, 0, 1]$;
- (4) l'equazione si spezza in $(x_1 + x_2)(x_1 - x_2)$ e otteniamo l'unione di due rette $x_1 = -x_2$ e $x_1 = x_2$, che si intersecano nel punto $[0, 0, 1]$;
- (5) otteniamo la retta $x_1 = 0$, ma con "molteplicità due", come se fosse contata due volte.

La dimostrazione è completa. \square

In particolare una conica proiettiva non degenera è vuota oppure una circonferenza.

Esempio 7.2.7. Consideriamo la famiglia di coniche proiettive

$$C_k = \{kx_0^2 + x_1^2 + 5x_2^2 + 4x_0x_1 - 4x_0x_2 - 2x_1x_2 = 0\}$$

dipendenti da un parametro $k \in \mathbb{R}$. La matrice dei coefficienti è

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

Il determinante è $\det A = 4k - 16$.

- se $k > 4$ la segnatura è $(3, 0, 0)$ e quindi $C_k = \emptyset$;
- se $k = 4$ la segnatura è $(2, 0, 1)$ e quindi C_4 è un punto;
- se $k < 4$ la segnatura è $(2, 1, 0)$ e quindi C_k è una circonferenza.

Passiamo adesso dal piano allo spazio.

7.2.6. Quadriche proiettive in \mathbb{P}^3 . Usiamo il Teorema 7.2.4 per classificare le quadriche in \mathbb{P}^3 .

Corollario 7.2.8. *Ogni quadrica proiettiva in \mathbb{P}^3 è equivalente a una di queste:*

- (1) l'insieme vuoto $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 0$,
- (2) la sfera $x_1^2 + x_2^2 + x_3^2 - x_4^2 = 0$,
- (3) il toro $x_1^2 + x_2^2 - x_3^2 - x_4^2 = 0$,
- (4) il punto $x_1^2 + x_2^2 + x_3^2 = 0$,
- (5) il doppio cono $x_1^2 + x_2^2 - x_3^2 = 0$,
- (6) i due piani incidenti $x_1^2 - x_2^2 = 0$,
- (7) la retta $x_1^2 + x_2^2 = 0$,
- (8) il piano doppio $x_1^2 = 0$.

Dimostrazione. Le possibili segnature di A sono

$$(4, 0, 0), \quad (3, 1, 0), \quad (2, 2, 0), \quad (3, 0, 1), \quad (2, 1, 1), \quad (1, 1, 2), \quad (1, 0, 3).$$

Discutiamo brevemente caso per caso:

- (1) l'equazione non ha soluzione in \mathbb{P}^3 ;
- (2) l'equazione implica $x_4 \neq 0$, poniamo $x_4 = 1$ e otteniamo la sfera

$$x_1^2 + x_2^2 + x_3^2 = 1$$

nello spazio \mathbb{R}^3 ;

- (3) riscriviamo l'equazione come $x_1^2 + x_2^2 = x_3^2 + x_4^2$ e a meno di riscalamento possiamo scrivere $x_1^2 + x_2^2 = x_3^2 + x_4^2 = 1$; Possiamo scrivere equazioni parametriche

$$x_1 = \cos \theta, \quad x_2 = \sin \theta, \quad x_3 = \cos \phi, \quad x_4 = \sin \phi$$

che dipendono da due angoli indipendenti θ e ϕ ; uno spazio parametrizzato in questo modo è effettivamente un toro;

- (4) l'unica soluzione è $[0, 0, 0, 1]$;
- (5) il luogo di zeri dell'equazione in \mathbb{R}^3 è effettivamente un doppio cono;
- (6) si spezza in $(x_1 + x_2)(x_1 - x_2) = 0$, quindi otteniamo due piani $x_1 = -x_2$ e $x_1 = x_2$ che si intersecano nella retta $x_1 = x_2 = 0$;
- (7) otteniamo il piano $x_1 = 0$ con "molteplicità due".

La dimostrazione è completa. □

7.3. Classificazione affine

Classifichiamo in questa sezione le quadriche affini.

7.3.1. Omogeneizzazione. Consideriamo un polinomio $p(x)$ nelle variabili x_1, \dots, x_n avente grado d . Il processo di *omogeneizzazione* trasforma $p(x)$ in un polinomio omogeneo $q(x)$ nelle variabili x_1, \dots, x_{n+1} nel modo seguente: si moltiplica ciascun monomio

$$\lambda x_{i_1} \cdots x_{i_k}$$

di $p(x)$ avente un certo grado k per l'elemento x_{n+1}^{d-k} , in modo da ottenere così un nuovo monomio

$$\lambda x_{i_1} \cdots x_{i_k} x_{n+1}^{d-k}$$

di grado d . I nuovi monomi hanno tutti lo stesso grado d ed il nuovo polinomio $q(x)$ è effettivamente omogeneo.

Esempio 7.3.1. Consideriamo il caso $n = 2$. Se omogeneizziamo il polinomio $p(x) = x_1 x_2 + x_1^2 - x_2 + 7$ otteniamo

$$q(x) = x_1 x_2 + x_1^2 - x_2 x_3 + 7 x_3^2.$$

7.3.2. Completamento. Nella Sezione 6.2.3 abbiamo definito la nozione di completamento di uno spazio affine, e la estendiamo adesso alle quadriche.

Sia $Q \subset \mathbb{R}^n$ una quadrica, luogo di zeri di un polinomio $p(x)$ di secondo grado in x_1, \dots, x_n . Sia $q(x)$ il polinomio omogeneo di secondo grado in x_1, \dots, x_{n+1} ottenuto omogeneizzando $p(x)$. Definiamo il *completamento* $\bar{Q} \subset \mathbb{P}^n$ di Q come la quadrica proiettiva luogo di zeri di $q(x)$.

Proposizione 7.3.2. *Il completamento \bar{Q} contiene Q .*

Dimostrazione. Se (x_1, \dots, x_n) è soluzione di $p(x)$, allora $[x_1, \dots, x_n, 1]$ è soluzione di $q(x)$. Quindi $Q \subset \bar{Q}$. \square

Come nel caso dei sottospazi affini, la quadrica proiettiva \bar{Q} è ottenuta aggiungendo alla quadrica affine Q i suoi "punti all'infinito".

Esempio 7.3.3. Se $Q = \{x_1^2 + x_2^2 = 1\}$ è una circonferenza in \mathbb{R}^2 , il completamento $\bar{Q} = \{x_1^2 + x_2^2 = x_3^2\}$ è ancora una circonferenza: nessun punto all'infinito è stato aggiunto.

Se $Q = \{x_2 = x_1^2\}$ è una parabola in \mathbb{R}^2 , il completamento $\bar{Q} = \{x_2 x_3 = x_1^2\} = \{x_1^2 - x_2 x_3 = 0\}$ è una conica proiettiva descritta dalla matrice

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

La matrice ha segnatura $(2, 1, 0)$ e quindi \bar{Q} è una circonferenza. È stato aggiunto il punto all'infinito $[0, 1, 0]$.

Se $Q = \{x_1 x_2 = 1\}$ è una iperbole in \mathbb{R}^2 , il completamento $\bar{Q} = \{x_1 x_2 = x_3^2\}$ è ancora una circonferenza in \mathbb{P}^2 . Sono stati aggiunti due punti all'infinito $[1, 0, 0]$ e $[0, 1, 0]$.

7.3.3. Classificazione affine delle coniche. Una conica affine è il luogo di zeri di un polinomio

$$p(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1 + 2b_2x_2 + c.$$

Il polinomio omogeneizzato è

$$q(x_1, x_2, x_3) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1x_3 + 2b_2x_2x_3 + cx_3^2$$

e la sua matrice di coefficienti è

$$\bar{A} = \begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{12} & a_{22} & b_2 \\ b_1 & b_2 & c \end{pmatrix}.$$

Definiamo anche

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Sia C la conica definita da $p(x)$ e \bar{C} il suo completamento, definito da $q(x)$. Supponiamo che \bar{C} sia non-degenere, cioè che $\det \bar{A} \neq 0$. Sappiamo allora che \bar{C} è l'insieme vuoto oppure una circonferenza, a seconda della segnatura di \bar{A} . Nel primo caso anche C è vuoto e abbiamo concluso. Nel secondo, si presentano tre possibilità: possiamo interpretare il seguente teorema come definizione di ellisse, parabola e iperbole.

Teorema 7.3.4. *La circonferenza \bar{C} interseca la retta all'infinito $\{x_3 = 0\}$ in $k = 0, 1$ o 2 punti. Inoltre:*

- $k = 0 \iff \det A > 0$ e in questo caso C è una ellisse,
- $k = 1 \iff \det A = 0$ e in questo caso C è una parabola,
- $k = 2 \iff \det A < 0$ e in questo caso C è una iperbole.

Dimostrazione. La circonferenza \bar{C} è il luogo di zeri

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1x_3 + 2b_2x_2x_3 + cx_3^2 = 0.$$

Se intersechiamo con la retta $\{x_3 = 0\}$ otteniamo

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = 0.$$

Abbiamo già visto che le soluzioni in \mathbb{P}^1 di questo polinomio omogeneo di secondo grado sono 0, 1, 2 a seconda che $\det A$ sia positivo, nullo o negativo. La dimostrazione è conclusa. \square

Esercizio 7.3.5. Consideriamo la famiglia di coniche

$$C_t = \{(2t - 1)x^2 + 6txy + ty^2 + 2x = 0\}$$

dipendenti da un parametro $t \in \mathbb{R}$. Cerchiamo di capire che tipo di conica è C_t al variare di t . Scriviamo la matrice

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

La conica è degenera solo per $t = 0$: per questo valore $C_0 = \{-x^2 + 2x = 0\}$ è l'unione di due rette parallele $\{x = 0\}$ e $\{x = 2\}$.

Se $t \neq 0$ la conica è non degenera. Per il Lemma 3.2.14 e il Corollario 3.2.15, la matrice \bar{A} non è definita (né positiva né negativa) e quindi è indefinita per ogni t . Segue che C_t non è mai vuota.

Calcoliamo $\det A = t(2t - 1) - 9t^2 = -7t^2 - t$ e deduciamo che C_t è una ellisse per $t \in (-\frac{1}{7}, 0)$, una parabola per $t = -\frac{1}{7}$ e una iperbole per $t \in (-\infty, -\frac{1}{7}) \cup (0, +\infty)$.

7.3.4. Classificazione metrica delle coniche non-degeneri. Sia C una conica affine descritta da un'equazione $p(x_1, x_2) = 0$ come nella sezione precedente. Possiamo scrivere il polinomio brevemente così:

$$p(x) = {}^t x A x + 2 {}^t b x + c.$$

Un *centro* della conica è un punto $P \in \mathbb{R}^2$ tale che $AP + b = 0$. Se P è un centro, è utile traslare il sistema di riferimento inserendo delle nuove coordinate $x' = x - P$ in cui il centro risulti l'origine.

Proposizione 7.3.6. *Nelle coordinate x' il polinomio $p(x)$ diventa*

$$p'(x') = {}^t x' A x' + c'$$

per qualche $c' \in \mathbb{R}$.

Dimostrazione. Il polinomio $p(x)$ si trasforma in

$$\begin{aligned} p'(x') &= p(x' + P) = {}^t (x' + P) A (x' + P) + 2 {}^t b (x' + P) + c \\ &= {}^t x' A x' + {}^t P A x' + {}^t x' A P + {}^t P A P + 2 {}^t b x' + 2 {}^t b P + c \\ &= {}^t x' A x' + 2({}^t P A + b)x' + c' \\ &= {}^t x' A x' + c' \end{aligned}$$

dove $c' = {}^t P A P + 2 {}^t b P + c$. □

Il termine "centro" indica che P è un centro di simmetria per C : infatti sostituendo x' con $-x'$ l'espressione di $p'(x')$ resta invariata. Quindi se $r_P: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ è la riflessione rispetto al centro P otteniamo

$$r_P(C) = C.$$

Notiamo innanzitutto che se C è una ellissi o una iperbole allora $\det A \neq 0$ e quindi esiste un unico P tale che $AP = -b$, in altre parole la conica C ha un unico centro.

Possiamo a questo punto agevolmente classificare le coniche non degeneri a meno di isometrie affini di \mathbb{R}^2 .

Proposizione 7.3.7. *Ogni conica non degenera è isometrica a una di queste:*

- $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ (ellisse);
- $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ (iperbole);

- $y = ax^2$ (parabola)

per qualche $a, b > 0$.

Dimostrazione. Se la conica C è una ellisse o una iperbole, dopo aver traslato il centro nell'origine diventa della forma ${}^t xAx - c = 0$. Necessariamente $c \neq 0$ perché la conica è non degenere e quindi $\det \bar{A} \neq 0$. Possiamo dividere per c e ottenere ${}^t xAx = 1$.

A questo punto usiamo il teorema spettrale: la matrice A è simmetrica e quindi ha una base ortonormale di autovettori. Se usiamo questa base come nuovo sistema di riferimento, la matrice A è diagonale ed abbiamo concluso.

Se C è una parabola, possiamo comunque diagonalizzare A con il teorema spettrale ed ottenere

$$\bar{A} = \begin{pmatrix} a_{11} & 0 & b_1 \\ 0 & 0 & b_2 \\ b_1 & b_2 & c \end{pmatrix}.$$

Qui $a_{11}, b_2 \neq 0$ perché $\det \bar{A} \neq 0$. Allora $p(x)$ è della forma

$$p(x) = a_{11}x^2 + 2b_1x + 2b_2y + c.$$

Quindi l'equazione della parabola diventa

$$y = ax^2 + bx + c.$$

Una traslazione opportuna la trasforma nella forma cercata. \square

7.4. Classificazione metrica delle quadriche non degeneri

Con tecniche simili a quelle sviluppate nelle pagine precedenti possiamo dimostrare anche questo fatto.

Proposizione 7.4.1. *Ogni quadrica non degenere non vuota in \mathbb{R}^3 è isometrica a una di queste:*

- (1) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ (ellissoide);
- (2) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$ (iperboloide ad una falda);
- (3) $\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$ (iperboloide a due falde);
- (4) $z = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ (paraboloide ellittico);
- (5) $z = \frac{x^2}{a^2} - \frac{y^2}{b^2}$ (paraboloide iperbolico);

per qualche $a, b, c > 0$.

Le superfici sono mostrate in Figura 7.1. Analogamente possiamo classificare le quadriche degeneri:

Proposizione 7.4.2. *Ogni quadrica degenere in \mathbb{R}^3 è isometrica a una di queste:*

- (1) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$ (doppio cono);
- (2) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ (cilindro ellittico);

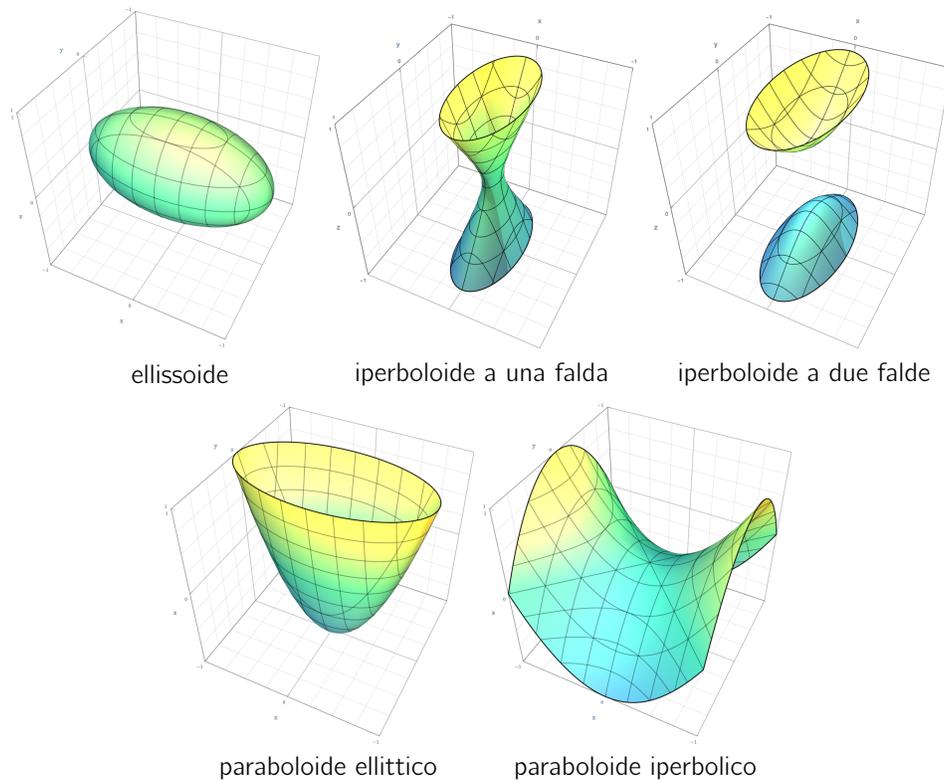
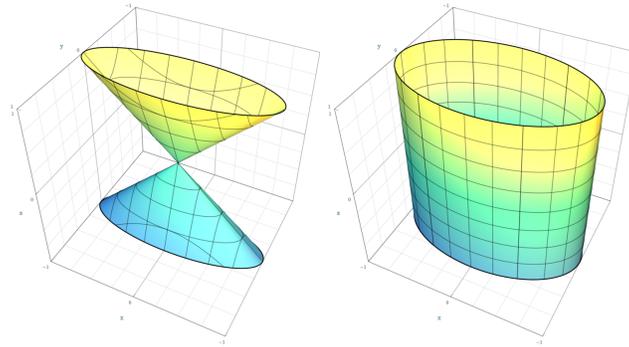


Figura 7.1. Le quadriche non degeneri.

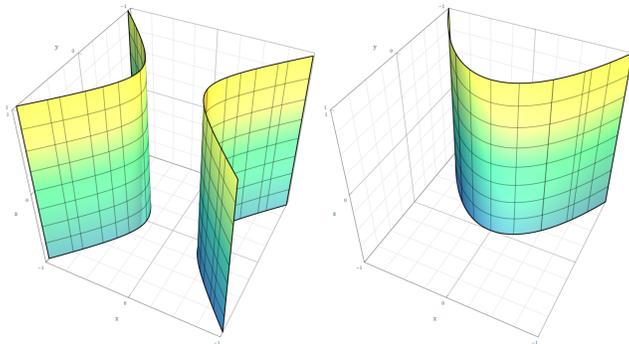
- (3) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ (*cilindro iperbolico*);
- (4) $y = ax^2$ (*cilindro parabolico*);
- (5) *due piani (incidenti, paralleli, oppure un piano doppio)*;
- (6) *una retta*;
- (7) *un punto*.

Alcune di queste sono mostrate in Figure 7.2.



doppio cono

cilindro ellittico



cilindro iperbolico

cilindro parabolico

Figura 7.2. Alcune quadriche degeneri.