Capitolo 13

Prodotto vettoriale (o esterno) in $\mathbb{R}^3$

Definizione 13.0.1. Il prodotto esterno o vettoriale (o prodotto wedge) di due vettori di $\mathbb{R}^3$ è l'applicazione $\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ tale che $(x, y) \to x \times y$ o $x \wedge y$ definito da

$$x \times y = \left(\det \begin{pmatrix} x_2 & x_3 \\ y_2 & y_3 \end{pmatrix}, -\det \begin{pmatrix} x_1 & x_3 \\ y_1 & y_3 \end{pmatrix}, \det \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}, \right)$$

che si usa anche scrivere simbolicamente come determinante

$$\det \begin{pmatrix} e_1 & e_2 & e_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix},$$

dove (e_1, e_2, e_3) come al solito è la base canonica di $\mathbb{R}^3$. Il determinante va pensato sviluppato con la regola di Laplace secondo la prima riga.

Il prodotto vettoriale gode delle seguenti proprietà:

1. è bilineare;
2. è antisimmetrico;
3. $x \times y = 0$ se e solo se x, y sono linearmente dipendenti;
4. $x \times y$ è ortogonale sia a x sia a y, dunque è ortogonale al piano generato dai due vettori.

Dimostrazione. Le 1,2,3 seguono dal fatto che le coordinate di $x \times y$ sono i minori di ordine 2 della matrice $\begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}$; la 4. segue dallo sviluppo di Laplace dei determinanti nulli

$$\det \begin{pmatrix} x_1 & x_2 & x_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}, \det \begin{pmatrix} y_1 & y_2 & y_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}.$$

$\square$

Osserviamo che si ha anche $\langle x \times y, z\rangle = \det \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix}$, ed è nullo se e solo se x, y, z sono linearmente dipendenti: è detto **prodotto misto** dei tre vettori x, y, z.

Calcoliamo la norma di $x \times y$:

$$\begin{aligned} \|x \times y\|^2 &= (x_2y_3 - x_3y_2)^2 + (x_1y_3 - x_3y_1)^2 + (x_1y_2 - x_2y_1)^2 = \\ &= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1y_1 + x_2y_2 + x_3y_3)^2 = \\ &= \|x\|^2\|y\|^2 - \langle x, y\rangle^2 = \\ &= \|x\|^2\|y\|^2 - \|x\|^2\|y\|^2 \cos^2\alpha = \quad \alpha \text{ l'angolo convesso di } x \text{ e } y \\ &= \|x\|^2\|y\|^2(1 - \cos^2\alpha) = \\ &= \|x\|^2\|y\|^2 \sin^2\alpha. \end{aligned}$$

Perciò $\|x \times y\| = \|x\|\|y\| \sin\alpha$, perché α è convesso, e coincide con l'area del parallelogramma di lati i vettori x e y.

Capitolo 14

Forme bilineari simmetriche e cambi di base

Sia V uno spazio vettoriale reale di dimensione finita, $b : V \times V \to \mathbb{R}$ una forma bilineare simmetrica. Siano $\mathcal{B}, \mathcal{B}'$ basi di V. Vogliamo vedere quale relazione intercorre fra le matrici $A = M_{\mathcal{B}}(b)$ e $A' = M_{\mathcal{B}'}(b)$.

Proposizione 14.0.1. *Detta $S = M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}_V)$ la matrice del cambio di base, si ha $A' = {}^tSAS$.*

Dimostrazione. Siano $v, w \in V$ due vettori e denotiamo con x, risp. x' le colonne delle coordinate di v rispetto a $\mathcal{B}$, risp. $\mathcal{B}'$, e con y, risp. y', quelle di w. Allora $x = Sx'$, $y = Sy'$. Le espressioni analitiche di b nelle coordinate rispetto alle due basi sono $b(v, w) = {}^txAy = {}^tx'A'y'$. D'altra parte ${}^txAy = {}^t(Sx')A(Sy') = {}^tx'({}^tSAS)y'$, e quindi $A' = {}^tSAS$ (per l'Osservazione 44). $\square$

Definizione 14.0.2. Due matrici quadrate $n \times n$ $A, A' \in M(n \times n, \mathbb{K})$ sono dette **congruenti** se esiste una matrice invertibile S tale che $A' =^t SAS$.

Osservazione 50. La Proposizione 14.0.1 afferma che due matrici che rappresentano la stessa forma bilineare simmetrica rispetto a basi diverse sono congruenti. È facile verificare che la relazione di congruenza fra matrici quadrate dello stesso ordine è una relazione d'equivalenza. Tale relazione è diversa dalla relazione di similitudine, che intercorre invece fra matrici che rappresentano lo stesso endomorfismo rispetto a basi diverse.

Tuttavia, il Teorema spettrale per le matrici (Corollario 12.3.2) afferma che per ogni matrice simmetrica reale A, esiste una matrice **ortogonale** S che diagonalizza A, ovvero ${}^tSAS = S^{-1}AS = D$, matrice diagonale; infatti ${}^tS = S^{-1}$. Quindi tramite S, che è ortogonale, A è sia simile sia congruente a una matrice diagonale, e precisamente alla matrice diagonale che ha sulla diagonale principale gli autovalori $\lambda_1, \ldots, \lambda_n$ di A, presi ciascuno tante volte quant'è la sua molteplicità algebrica. In questo modo S diagonalizza sia la forma bilineare simmetrica sia l'endomorfismo L_A di $\mathbb{R}^n$ aventi A come matrice rispetto alla base canonica.

Se $\mathbb{R}^n$ è dotato della struttura di spazio euclideo con il prodotto scalare canonico, la matrice ortogonale S si può interpretare come matrice di passaggio da $\mathcal{C}$ a una base $\mathcal{B}$ ortonormale. Quindi come corollario del Teorema spettrale, data una matrice simmetrica reale A, esiste una base ortonormale $\mathcal{B}$ di $\mathbb{R}^n$ rispetto a cui la forma bilineare simmetrica b su $\mathbb{R}^n$ associata ad A è rappresentata da una matrice diagonale.

Nelle coordinate x, y rispetto a tale base $\mathcal{B}$, $b(v, w) = \lambda_1 x_1 y_1 + \lambda_2 x_2 y_2 + \cdots + \lambda_n x_n y_n$. Il problema della diagonalizzazione per forme bilineari simmetriche reali ha dunque sempre risposta positiva. Ciò viene ulteriormente precisato e migliorato nel seguente teorema.

Teorema 14.0.3 (Teorema di trasformazione ad assi principali)**.** *Consideriamo $\mathbb{R}^n$ spazio euclideo con il prodotto scalare canonico. Sia $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ una forma bilineare simmetrica. Sia $A = M_{\mathcal{C}}(b)$ la matrice simmetrica $n \times n$ che la rappresenta rispetto alla base canonica. Allora:*

1. *esiste una base ortonormale $\mathcal{B} = (v_1, \ldots, v_n)$ formata da autovettori di A; ovvero esiste $S = M^{\mathcal{B}}_{\mathcal{C}}(\mathrm{id}_{\mathbb{R}^n})$, matrice ortogonale, tale che*

$$M_{\mathcal{B}}(b) = {}^tSAS = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \tag{14.1}$$

diagonale, con $\lambda_1, \ldots, \lambda_n$ gli autovalori di A, tutti reali.

2. *esiste una base ortogonale $\mathcal{B}'$ di autovettori di A, tale che*

$$M_{\mathcal{B}'}(b) = {}^tTAT = \left(\begin{array}{ccc} \begin{matrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & 1 \end{matrix} & \text{\Large 0} & \text{\Large 0} \\ \text{\Large 0} & \begin{matrix} -1 & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & -1 \end{matrix} & \text{\Large 0} \\ \text{\Large 0} & \text{\Large 0} & \text{\Large 0} \end{array}\right) : \tag{14.2}$$

è una matrice diagonale con soli 1, -1, 0 sulla diagonale principale, e precisamente tanti 1 quanti sono gli autovalori positivi, tanti -1 quanti gli autovalori negativi.

Dimostrazione. La 1. è ancora una riformulazione del Teorema spettrale.

Per dimostrare la 2., eventualmente riordinando i vettori di $\mathcal{B}$, povveromo supporre che gli autovalori $\lambda_1, \dots, \lambda_r$ siano > 0, gli autovalori $\lambda_{r+1}, \dots, \lambda_{r+s}$ siano < 0, e che $\lambda_{r+s+1} = \cdots = \lambda_n = 0$. Definiamo la base $\mathcal{B}' = (v_1', \dots, v_n')$ ponendo

$$v_i' = \begin{cases} \frac{v_i}{\sqrt{\lambda_i}} & \text{se } i \leq r, \text{ ovvero } \lambda_i > 0, \\ \frac{v_i}{\sqrt{-\lambda_i}} & \text{se } r < i \leq r+s, \text{ ovvero } \lambda_i < 0, \\ v_i & \text{se } r+s < i \leq n, \text{ ovvero } \lambda_i = 0. \end{cases}$$

Allora $b(v_i', v_j') = 0$ se $i \neq j$, mentre

$$b(v_i', v_i') = \begin{cases} b(\frac{v_i}{\sqrt{\lambda_i}}, \frac{v_i}{\sqrt{\lambda_i}}) = (\frac{1}{\sqrt{\lambda_i}})^2 b(v_i, v_i) = 1 & \text{se } i \leq r, \\ b(\frac{v_i}{\sqrt{-\lambda_i}}, \frac{v_i}{\sqrt{-\lambda_i}}) = (\frac{1}{\sqrt{-\lambda_i}})^2 b(v_i, v_i) = -1 & \text{se } r < i \leq r+s, \\ b(v_i, v_i) = 0 & \text{se } r+s < i \leq n. \end{cases}$$

□

Con la notazione precedente, $M_{\mathcal{B}}(b)$ ha r autovalori positivi, s autovalori negativi, e i rimanenti $n-r-s$ autovalori uguali a 0. Nelle coordinate x', y' rispetto alla nuova base $\mathcal{B}'$, $b(v,w) = x_1'y_1' + x_2'y_2' + \cdots + x_r'y_r' - x_{r+1}'y_{r+1}' - \cdots - x_{r+s}'y_{r+s}'$. Osserviamo che la matrice $T = M_{\mathcal{C}}^{\mathcal{B}'}(\mathrm{id}_{\mathbb{R}^n})$ in generale non è ortogonale. Le rette generate dai vettori v_i, che coincidono con le rette generate dai vettori v_i', sono dette **assi principali** per b. Il cambio di base effettuato per passare da $\mathcal{B}$ a $\mathcal{B}'$ non cambia le direzione dei tre assi coordinati, ma consiste semplicemente nel riscalare i vettori.

Definizione 14.0.4. Data una forma bilineare simmetrica $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, l'applicazione $q : \mathbb{R}^n \to \mathbb{R}$ definita da $q(x) = b(x,x)$ è detta **forma quadratica** associata a b.

Detta A la matrice simmetrica $M_{\mathcal{C}}(b)$ di b rispetto alla base canonica, $q(x) = {}^t x A x = \sum_{ij} a_{ij} x_i x_j$ è rappresentata da un polinomio omogeneo di secondo grado in $x_1, \dots, x_n$. Per la forma quadratica, il Teorema 14.0.3 si può riformulare così:

Teorema 14.0.5. *Sia $q : \mathbb{R}^n \to \mathbb{R}$ la forma quadratica su $\mathbb{R}^n$ (dotato della struttura euclidea standard) associata a una forma bilineare simmetrica. Allora:*

1. *esiste una base ortonormale di $\mathbb{R}^n$ tale che, in coordinate rispetto a questa base, q si scrive $q(x_1, \dots, x_n) = \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2$;*

2. *esiste una base ortogonale di $\mathbb{R}^n$ tale che, in coordinate rispetto a questa base, q si scrive $q(x_1', \dots, x_n') = x_1'^2 + \cdots + x_r'^2 - x_{r+1}'^2 - \cdots - x_{r+s}'^2$.*

Teorema 14.0.6 (Teorema di Sylvester - Legge d'inerzia)**.** *1. Sia V uno spazio vettoriale reale di dimensione finita n, sia $b : V \times V \to \mathbb{R}$ una forma bilineare simmetrica. Il numero di autovalori positivi, negativi e nulli di $M_{\mathcal{B}}(b)$ dipende solo da b ed è indipendente dalla base fissata.*

2. Ogni matrice simmetrica reale è congruente a una e una sola matrice diagonale della forma (14.2).

Dimostrazione. Sia $\mathcal{B} = (v_1, \dots, v_n)$ e supponiamo che $M_{\mathcal{B}}(b)$ sia della forma (14.1) con $\lambda_1, \dots, \lambda_r > 0$, $\lambda_{r+1}, \dots, \lambda_{r+s} < 0$, e $\lambda_i = 0$ per $i > r + s$. Sia $W_+ = \langle v_1, \dots, v_r \rangle$, $W_- = \langle v_{r+1}, \dots, v_{r+s} \rangle$, $W_0 = \langle w_{r+s+1}, \dots, v_n \rangle$. Allora:

- $b(v, v)$ è sempre > 0 su W_+, < 0 su W_-, 0 su W_0;
- $V = W_+ \oplus W_- \oplus W_0$;
- r è la massima dimensione dei sottospazi $U \subset V$ tali che $b(v, v) > 0$ per ogni $v \in U$ $v \neq 0$, ovvero b è definita positiva se ristretta a U. Infatti se b è definita positiva su U, allora $U \cap (W_+ \cup W_0) = (0)$, quindi la somma $U \oplus W_- \oplus W_0$ è diretta. Allora $\dim(U \oplus W_- \oplus W_0) = \dim U + \dim W_+ \dim W_0 = \dim U + s + (n - r - s) = \dim U + n - r \leq n = \dim V$. Quindi $\dim U \leq r$.

Questa caratterizzazione di r non dipende dalla base scelta. Per s vale una caratterizzazione analoga, come massima dimensione dei sottospazi dove $b(v, v) < 0$ per ogni $v \neq 0$. Questo conclude la dimostrazione. □

Definizione 14.0.7. Con la notazione precedente, la **segnatura** di una forma bilineare simmetrica reale b, o della forma quadratica q a essa associata, o della matrice che le rappresenta rispetto a una base fissata, è la coppia (r, s), o la terna (r, s, n).

La forma bilineare simmetrica b, o la forma quadratica q, è detta:

1. **definita positiva** se ha segnatura $(n, 0)$, o $(n, 0, n)$;
2. **definita negativa** se ha segnatura $(0, n)$, o $(0, n, n)$;
3. **semidefinita positiva** se ha segnatura $(m, 0)$ o $(m, 0, n)$ con $m < n$;
4. **semidefinita negativa** se ha segnatura $(0, m)$ o $(0, m, n)$ con $m < n$;
5. **indefinita** in tutti gli altri casi.

Due matrici simmetriche reali sono dunque congruenti se e solo se hanno la stessa segnatura.

Notiamo che il segno assunto da q nei 5 casi è come segue:

1. $q(v) \geq 0$ per ogni v e $q(v) = 0$ se e solo se $v = 0$; ovvero b è un prodotto scalare;

2. $q(v) \leq 0$ per ogni v e $q(v) = 0$ se e solo se $v = 0$;

3. $q(v) \geq 0$ per ogni v ma esiste $v \neq 0$ tale che $q(v) = 0$;

4. $q(v) \leq 0$ per ogni v ma esiste $v \neq 0$ tale che $q(v) = 0$;

5. esistono vettori $v, w \in V$ con $q(v) > 0, q(w) < 0$.

Concludiamo enunciando senza dimostrazione un utile criterio che permette di stabilire la segnatura di una matrice simmetrica reale solo guardando il suo polinomio caratteristico.

Teorema 14.0.8 (Criterio di Cartesio). *Sia $p(x) \in \mathbb{R}[x]$ un polinomio di grado n a coefficienti reali:*

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n;\ n > 0,\ a_n \neq 0.$$

Sia m il minimo intero tale che $a_m \neq 0$. **Supponiamo che tutte le radici di p siano reali.** *Allora:*

1. *0 è radice di $p(x)$ se e solo se $m > 0$ e in tal caso m è la molteplicità algebrica di 0 come radice di $p(x)$;*

2. *il numero di radici positive di $p(x)$ (contate con la relativa molteplicità algebrica) è uguale al numero di variazioni di segno nella sequenza dei coefficienti* **non nulli** *di $p(x)$;*

3. *il numero di radici negative di $p(x)$ è pari a $n -$ (numero di radici positive) $- m$.*

Il Criterio di Cartesio si applica al polinomio caratteristico di ogni matrice simmetrica reale, in quanto sappiamo che le sue radici sono tutte reali.

Corollario 14.0.9. *La forma bilineare simmetrica b associata a una matrice A di ordine n è un prodotto scalare se e solo se la sequenza dei coefficienti del polinomio caratteristico $p_A(x)$ presenta n variazioni di segno; in particolare nessun coefficiente di $p_A(x)$ può annullarsi, e si ha $m = 0$ ovvero 0 non è una sua radice.*

Dimostrazione. Infatti dire che b è un prodotto scalare significa che gli autovalori di A sono tutti positivi, ovvero che il polinomio caratteristico di A ha radici tutte positive. In particolare 0 non può essere una sua radice, e la sequenza dei coefficienti deve avere segni alterni. □

Esempi 14.0.10. Consideriamo la matrice

$$A_c = \begin{pmatrix} 4 & 1 & -1 \\ 1 & 1 & 0 \\ -4 & 0 & c \end{pmatrix}$$

dove c è un parametro reale. Vogliamo determinare per quali valori di c A_c definisce un prodotto scalare. Il polinomio caratteristico è

$$p_{A_c}(x) = \det \begin{pmatrix} 4-x & 1 & -1 \\ 1 & 1-x & 0 \\ -4 & 0 & c-x \end{pmatrix} = -x^3 + (c+5)x^2 + (13-5c)x + (3c-16).$$

Il primo segno è negativo, dunque, per avere un prodotto scalare, dovranno valere le condizioni $c+5 > 0$, $13-5c < 0$ e $3c-16 > 0$; queste tre condizioni equivalgono all'unica condizione $c > \frac{16}{3}$. Osserviamo che soltanto per $c = \frac{16}{3}$ il termine noto si annulla, dunque 0 è un autovalore; in tal caso ci sono due variazioni di segno, dunque due radici positive e una nulla, e la forma bilineare corrispondente è semidefinita positiva. Per tutti gli altri valori di c vi sono due variazioni di segno e 0 non è radice, dunque la terza radice è negativa, e la forma risulta indefinita.

14.1 Esempio di trasformazione ad assi principali

Sia A la seguente matrice simmetrica reale:

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

A è la matrice dell'endomorfismo autoaggiunto L_A rispetto alla base canonica di $\mathbb{R}^3$, con prodotto scalare standard. Ad A corrisponde anche la forma bilineare simmetrica b su $\mathbb{R}^3$ definita da: $b(x,y) = {}^t xAy = 2x_1y_1 - 3x_1y_3 - 3x_3y_1 - x_2y_2 + 2x_3y_3$. Si ha $A = M_{\mathcal{C}}(b)$. La forma quadratica $q : \mathbb{R}^3 \to \mathbb{R}$ è definita da $q(x) = b(x,x) = 2x_1^2 - 6x_1x_3 - x_2^2 + 2x_3^2$. Applichiamo il Teorema spettrale per diagonalizzare A mediante una matrice ortogonale. Calcoliamo il polinomio caratteristico di A per trovare i suoi autovalori:

$$p_A(x) = \det \begin{pmatrix} 2-x & 0 & -3 \\ 0 & -1-x & 0 \\ -3 & 0 & 2-x \end{pmatrix} = -(1+x)[(2-x)^2-9] = -(1+x)(x^2-4x-5) = -(x+1)^2(x-5).$$

Gli autovalori sono -1 e 5, con $m_a(-1) = 2$ e $m_a(5) = 1$. Allora il Teorema spettrale ci garantisce che esiste una matrice ortogonale S tale che

$$S^{-1}AS = {}^tSAS = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{pmatrix}, \tag{14.3}$$

matrice diagonale con gli autovalori di A sulla diagonale. S è la matrice di un cambio di base: $S = M^{\mathcal{B}}_{\mathcal{C}}(\mathrm{id}_{\mathbb{R}^3})$, dove $\mathcal{B} = (v_1, v_2, v_3)$ è fatta così: v_1, v_2 formano una base ortonormale di $\mathrm{Aut}(-1)$, e v_3 è un vettore di norma 1 che genera $\mathrm{Aut}(5)$.

Per trovare $\mathrm{Aut}(-1)$, calcoliamo $A + \mathbb{I}_3 = \begin{pmatrix} 3 & 0 & -3 \\ 0 & 0 & 0 \\ -3 & 0 & 3 \end{pmatrix}$ di rango 1, quindi $x_1 - x_3 = 0$ è l'equazione di $\mathrm{Aut}(-1)$. Una base è formata da $(1, 0, 1), (0, 1, 0)$, è già ortogonale; la normalizziamo e troviamo $v_1 = \frac{1}{\sqrt{2}}(1, 0, 1)$, $v_2 = e_2$. Per trovare $\mathrm{Aut}(5)$, calcoliamo $A - 5\mathbb{I}_3 = \begin{pmatrix} -3 & 0 & -3 \\ 0 & -6 & 0 \\ -3 & 0 & -3 \end{pmatrix}$ di rango 2, quindi le equazioni sono $\begin{cases} x_1 + x_3 = 0 \\ x_2 = 0. \end{cases}$

Una base di $\mathrm{Aut}(5)$ è $(1, 0, -1)$, lo normalizziamo e otteniamo $v_3 = \frac{1}{\sqrt{2}}(1, 0, -1)$. Abbiamo così ottenuto la base ortonormale diagonalizzante $\mathcal{B} = (v_1, v_2, v_3)$. Allora per questa scelta della base $\mathcal{B}$ otteniamo

$$S = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{pmatrix}:$$

matrice ortogonale. Come abbiamo visto nella Sezione **??**, le equazioni del cambio di coordinate sono $x = Sy$, cioè per esteso $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = S \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$, dove x_1, x_2, x_3 sono coordinate rispetto a $\mathcal{C}$ e y_1, y_2, y_3 rispetto a $\mathcal{B}$. Quindi equivalentemente $y = S^{-1}x = {}^tSx$, perché S è ortogonale. Esplicitamente abbiamo:

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

Gli assi principali sono gli assi y_1, y_2, y_3, cioè le tre rette generate dai tre vettori di $\mathcal{B}$: l'asse y_1 è la retta $\langle v_1 \rangle$ di equazioni $y_2 = y_3 = 0$, ovvero $x_2 = x_1 - x_3 = 0$, l'asse y_2 è la retta $\langle v_2 \rangle$ di equazioni $y_1 = y_3 = 0$ ovvero $x_1 + x_3 = x_1 - x_3 = 0$; l'asse y_3 è la retta $\langle v_3 \rangle$ di equazioni

$y_1 = y_2 = 0$, ovvero $x_1 + x_3 = x_2 = 0$. La forma quadratica q allora si può scrivere:

$$q(x) = b(x,x) = {}^t x A x = {}^t(Sy)A(Sy) = {}^t y({}^t SAS)y \overset{\text{per la (14.3)}}{=} -y_1^2 - y_2^2 + 5y_3^2 = q'(y),$$

quindi q è stata diagonalizzata: la sua espressione rispetto agli assi principali non contiene i termini "misti" $y_i y_j$ con $i \neq j$.

Per il Teorema 14.0.3, povveromo fare un ulteriore cambio di base, senza cambiare i tre assi, ma solo riscalando, in modo che la forma quadratica sia rappresentata da una matrice diagonale con soli $1, -1$ sulla diagonale. Essendo $\lambda_1 = \lambda_2 = -1$, $\lambda_3 = 5$ avremo la base ortogonale $\mathcal{B}'$ con $v_1' = v_1$, $v_2' = v_2$, $v_3' = \frac{1}{\sqrt{5}} v_3$. Quindi $M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_{\mathbb{R}^3}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\sqrt{5}} \end{pmatrix}$. Le corrispondenti coordinate z_1, z_2, z_3 sono date da

$$y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\sqrt{5}} \end{pmatrix} z, \quad \text{quindi} \begin{cases} y_1 &= z_1 \\ y_2 &= z_2 \\ y_3 &= \frac{1}{\sqrt{5}} z_3 \end{cases}.$$

La forma quadratica allora si trasforma: $-y_1^2 - y_2^2 + 5y_3^2 = -z_1^2 - z_2^2 + z_3^2$. e corrisponde alla matrice diagonale $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Osserviamo che la segnatura è la coppia $(1,2)$, la forma è indefinita.

Povveromo dare un'interpretazione geometrica considerando l'insieme

$$X = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid q(x) = 2x_1^2 - 6x_1x_3 - x_2^2 + 2x_3^2 = 1\}.$$

È una **superficie quadrica** di cui vogliamo capire la geometria. Se passiamo alle coordinate (y_1, y_2, y_3) l'equazione della superficie diventa $-y_1^2 - y_2^2 + 5y_3^2 = 1$. Ora se tagliamo X con un piano affine di equazione $y_3 = k$, troviamo $y_1^2 + y_2^2 = 5k^2 - 1$: è l'equazione di

- una circonferenza se $5k^2 - 1 > 0$, cioè $k > \frac{\sqrt{5}}{5}$ o $k < -\frac{\sqrt{5}}{5}$;
- un punto se $k = \pm\frac{\sqrt{5}}{5}$
- vuoto se $-\frac{\sqrt{5}}{5} < k < \frac{\sqrt{5}}{5}$

Se invece tagliamo X con piani del tipo $y_1 = k$ o $y_2 = k$ troviamo delle iperboli, per esempio $5y_3^2 - y_1^2 = 1 + k^2$.

Nella trasformazione $x = Sy$, o $y = S^{-1}x = {}^tSx$ la superficie non viene "deformata" in quanto, essendo S ortogonale, si conservano le lunghezze. Quindi passare agli assi principali ci permette di capire la "forma" di X. Nel passare invece alle coordinate z_1, z_2, z_3 le lunghezze non si conservano e la superficie viene deformata.

Capitolo 15

Triangolarizzazione e forma canonica di Jordan

15.1 Matrici di Jordan

Definizione 15.1.1. Una matrice di Jordan, o un blocco di Jordan, è una matrice J $n \times n$ della forma

$$\begin{pmatrix} \lambda & 0 & \cdots & 0 & 0 \\ 1 & \lambda & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 1 & \lambda & 0 \\ 0 & 0 & \cdots & 1 & \lambda \end{pmatrix} \quad \text{oppure} \quad \begin{pmatrix} \lambda & 1 & \cdots & 0 & 0 \\ 0 & \lambda & \ddots & 0 & 0 \\ \vdots & & \ddots & \ddots & 0 \\ 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{pmatrix},$$

dove λ è un qualunque scalare. Nel primo caso J è una matrice triangolare inferiore con tutti λ sulla diagonale principale e 1 sulla diagonale immediatamente sotto, mentre nel secondo caso J è una matrice triangolare superiore con tutti λ sulla diagonale principale e tutti 1 sulla diagonale immediatamente sopra.

Questa classe di matrici generalizza il precedente esempio 9.5.3. Il nome viene dal matematico francese Camille Jordan (1838-1922).

Il polinomio caratteristico di J, in entrambi i casi, è $P_J(x) = (\lambda - x)^n$; $p_J(x)$ si fattorizza in n fattori lineari uguali, J ha un unico autovalore λ con molteplicità algebrica n. L'autospazio

$\mathrm{Aut}(\lambda)$ ha dimensione pari a $n - \mathrm{rg}(A - \lambda\mathbb{I}_n)$, dove (nel primo caso)

$$A - \lambda\mathbb{I}_n = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ & \ddots & \ddots & 0 \\ 0 & \cdots & 1 & 0 \end{pmatrix},$$

dunque $\mathrm{rg}(A - \lambda\mathbb{I}_n) = n - 1$ e $m_g(\lambda) = 1$. Analogamente nel secondo caso. Inoltre nel primo caso $\mathrm{Aut}(\lambda) =< e_n >$ perché ha equazioni $x_1 = \cdots = x_{n-1} = 0$, mentre nel secondo caso $\mathrm{Aut}(\lambda) =< e_1 >$ perché ha equazioni $x_2 = \cdots = x_n = 0$. Dunque J non è diagonalizzabile, qualunque sia il campo base, ma è una matrice triangolare.

15.2 Triangolarizzazione

Definizione 15.2.1. 1. Sia $f : V \to V$ un endomorfismo di un $\mathbb{K}$-spazio vettoriale, $\dim V = n$. f è trangolarizzabile (o triangolabile) se esiste una base $\mathcal{B}$ di V tale che la matrice $M_{\mathcal{B}}(f)$ sia triangolare (superiore o inferiore).

2. Una matrice $A \in M(n \times n, \mathbb{K})$ è triangolarizzabile se è simile a una matrice triangolare.

Teorema 15.2.2. *Un endomorfismo $f : V \to V$ è triangolarizzabile se e solo se il suo polinomio caratteristico $p_f(x) \in \mathbb{K}[x]$ è fattorizzabile in fattori lineari, cioè ha tutte le sue radici in $\mathbb{K}$.*

Dimostrazione. Un'implicazione è facile: se f è triangolarizzabile, esiste $\mathcal{B}$, base di V, tale che

$$M_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & a_{12} & \cdots & a_{1n} \\ 0 & \lambda_2 & \cdots & a_{2n} \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Perciò $p_f(x) = (\lambda_1 - x) \cdots (\lambda_n - x)$. Analogamente se la matrice è diagonale inferiore.

Viceversa procediamo per induzione su $n = \dim V$. Per $n = 1$ è ovvio perché ogni matrice 1×1 è triangolare. Supponiamo vero il teorema per $n - 1$ e lo dimostriamo per n. Sia dunque f un endomorfismo di V, spazio vettoriale di dimensione n, e supponiamo che $p_f(x)$ sia prodotto di fattori lineari. Allora $p_f(x)$ ha almeno una radice $\lambda \in \mathbb{K}$ che è dunque un autovalore di f, sia $v_1 \neq 0$ un relativo autovettore: $f(v_1) = \lambda v_1$. Prolunghiamo v_1 a una base di V: $\tilde{\mathcal{B}} = (v_1, w_2, \ldots, w_n)$. Allora $V =< v_1 > \oplus W$, dove $W =< w_2, \ldots, w_n >$, W ha dimensione $n-1$ e $\mathcal{B}' = (w_2, \dots, w_n)$ è una sua base. Consideriamo

$$A = M_{\tilde{\mathcal{B}}}(f) = \begin{pmatrix} \lambda & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdots & & & \\ 0 & a_{n2} & \cdots & a_{nn} \end{pmatrix};$$

povveromo interpretare la sottomatrice $A' = \begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \cdots & & \\ a_{n2} & \cdots & a_{nn} \end{pmatrix}$ come matrice rispetto a $\mathcal{B}'$ di un endomorfismo g di W, il cui polinomio caratteristico è $|A' - x\mathbb{I}_{n-1}|$. Ma $p_f(x) = (\lambda - x)|A' - x\mathbb{I}_{n-1}|$, perché A è triangolare a blocchi; per ipotesi $p_f(x)$ è prodotto di fattori lineari, quindi anche $p_g(x)$ è prodotto di fattori lineari. Povveromo quindi applicare l'ipotesi induttiva: g è triangolarizzabile, quindi esiste una base $(v_2, \dots, v_n)$ di W rispetto a cui la matrice di g è triangolare superiore. Chiamiamo allora $\mathcal{B}$ la base di V ottenuta come unione di v_1 con $\mathcal{B}'$; la matrice $M_{\mathcal{B}}(f)$ risulta triangolare superiore.

In effetti: ogni vettore v di V ha un'unica espressione come somma di un multiplo di v_1 e di un vettore di W (perché $V = < v_1 > \oplus W$); detta π l'applicazione $\pi : V \to W$ che manda v nel suo addendo che sta in W, g risulta uguale a $\pi \circ f$ ristretta a W. □

Usando il Teorema 15.2.2 e il Teorema fondamentale dell'Algebra 9.6.1, si ottiene il seguente Corollario.

Corollario 15.2.3. *Ogni endomorfismo di un $\mathbb{C}$-spazio vettoriale di dimensione finita è triangolarizzabile. Similmente, ogni matrice quadrata di $M(n \times n, \mathbb{C})$ è simile a una matrice traingolare.*

15.3 Potenze di un endomorfismo e autospazi generalizzati

Sia $g : V \to V$ un endomorfismo. Le **potenze di** g sono gli endomorfismi $g^2 := g \circ g$, $g^3 := g \circ g \circ g$, e così via. Valgono le seguenti semplici proprietà:

Proposizione 15.3.1.
1. $\ker g \subset \ker g^2 \subset \ker g^3 \subset \dots$ *(*)*
2. se $\ker g^k = \ker g^{k+1}$ *per un certo* $k \geq 1$*, da quel momento in poi nella catena (*) sono tutte uguaglianze;*
3. Se $v \in \ker g^k$*, allora* $g(v) \in \ker g^{k-1}$.

Dimostrazione. 1. Se $g^k(v) = 0$ allora $g^{k+1}(v) = g(g^k(v)) = 0$. Dunque $\ker g^k \subset \ker g^{k+1}$ per ogni k.

2. Supponiamo $\ker g^k = \ker g^{k+1}$, e sia $v \in \ker g^{k+2}$. Allora $g^{k+2}(v) = 0 = g^{k+1}g(v)$, e perciò $g(v) \in \ker g^{k+1} = \ker g^k$. Dunque $g^k(g(v)) = 0$ ovvero $g^{k+1}(v) = 0$. Si conclude che $\ker g^{k+2} = \ker g^{k+1}$.

3. Se $g^k(v) = 0$, allora $g^{k-1}(g(v)) = 0$, e quindi $g(v) \in \ker g^{k-1}$. □

Applicheremo questa proposizione nel caso in cui è dato un endomorfismo $f : V \to V$, di cui ci interessa trovare una forma "canonica", $\lambda \in \mathbb{K}$ è un autovalore di f, e $g = f - \lambda\,\mathrm{id}_V$. In questo caso $\ker g = \mathrm{Aut}(\lambda)$ e ha dimensione $m_g(\lambda)$.

Vale la seguente proposizione di cui omettiamo la dimostrazione.

Proposizione 15.3.2. *Sia* $g = f - \lambda\,\mathrm{id}_V$.

1. Se $m_g(\lambda) = m_a(\lambda)$*, nella catena (*) della Proposizione 15.3.1 sono tutte uguaglianze.*

2. Se $m_g(\lambda) < m_a(\lambda)$*, allora* $\mathrm{Aut}(\lambda) = \ker g$ *è contenuto strettamente in* $\ker g^2$*, e* $\dim\ker g^2 \leq m_a(\lambda)$.

3. Nella catena (*) i sottospazi hanno tutti dimensione $\leq m_a(\lambda)$*; le inclusioni sono strette finchè si arriva a un sottospazio* $\ker g^k$ *di dimensione pari a* $m_a(\lambda)$*, e allora la catena si stabilizza.*

Poniamo $a = m_a(\lambda)$. Si usa la seguente notazione:

$$\begin{array}{ccccccccc}
\mathrm{Aut}_1(\lambda) & \subsetneq & \mathrm{Aut}_2(\lambda) & \subsetneq & \cdots & \subsetneq & \mathrm{Aut}_k(\lambda) & = \cdots = & \mathrm{Aut}_a(\lambda) \\
\| & & \| & & & & \| & & \| \\
\ker g & \subsetneq & \ker g^2 & \subsetneq & \cdots & \subsetneq & \ker g^k & = \cdots = & \ker g^a \\
\| & & & & & & & & \\
\mathrm{Aut}(\lambda) & & & & & & & &
\end{array}$$

I sottospazi che compaiono nella catena sono detti **autospazi generalizzati dell'autovalore** λ**.** La catena si stabilizza certamente dopo a passi ma potrebbe stabilizzarsi anche prima.

Consideriamo ora un endomorfismo $f : V \to V$ tale che $p_f(x)$ ha tutte le sue radici in $\mathbb{K}$:

$$p_f(x) = (\lambda_1 - x)^{a_1}(\lambda_2 - x)^{a_2} \cdots (\lambda_p - x)^{a_p}$$

con $a_i = m_a(\lambda_i)$ per ogni indice i, dunque f è triangolarizzabile. Si ha $n = a_1 + a_2 + \cdots + a_p$. Per ogni autovalore, consideriamo il relativo autospazio generalizzato $\mathrm{Aut}_{a_i}(\lambda_i)$ di dimensione a_i.

Teorema 15.3.3. *Sia* f *un endomorfismo triangolarizzabile di* V*. Allora*

$$V = \mathrm{Aut}_{a_1}(\lambda_1) \oplus \mathrm{Aut}_{a_2}(\lambda_2) \oplus \cdots \oplus \mathrm{Aut}_{a_p}(\lambda_p).$$

Dimostrazione. Omessa. □

Questo teorema servirà a costruire una forma canonica per la matrice di f.

15.4 Forma canonica di Jordan

In questa sezione chiameremo blocco di Jordan una matrice J come nella definizione 15.1.1 triangolare inferiore. Osserviamo che se J è la matrice di un endomorfismo f rispetto a una base $\mathcal{B} = (v_1, \dots, v_n)$, si ha:

$$\begin{array}{ccccc} f(v_1) = \lambda v_1 + v_2 & \iff & (f - \lambda \operatorname{id}_V)(v_1) = & g(v_1) = & v_2 \\ f(v_2) = \lambda v_2 + v_3 & \iff & (f - \lambda \operatorname{id}_V)(v_2) = & g(v_2) = & g^2(v_1) = v_3 \\ \vdots & & \vdots & & \\ f(v_{n-1}) = \lambda v_{n-1} + v_n & \iff & (f - \lambda \operatorname{id}_V)(v_{n-1}) = & g(v_{n-1}) = & g^{n-1}(v_1) = v_n \\ f(v_n) = \lambda v_n & \iff & v_n \text{ è un autovettore di } \lambda, & g(v_n) = 0. & \end{array}$$

Questo è il significato della condizione $M_{\mathcal{B}}(f) = J$.

Il seguente teorema ci dice che, se f ha tutti gli autovalori in $\mathbb{K}$, c'è una base di V rispetto a cui la matrice di f è composta da blocchi di Jordan.

Teorema 15.4.1 (Forma canonica o normale di Jordan)**.** *Sia $f : V \to V$ un endomorfismo tale che*

$$p_f(x) = (\lambda_1 - x)^{a_1}(\lambda_2 - x)^{a_2} \cdots (\lambda_p - x)^{a_p}$$

con $\dim V = n = a_1 + a_2 + \cdots + a_p$. Allora esiste una base $\mathcal{B}$ di V, detta base di Jordan, unione di basi degli autospazi generalizzati $\operatorname{Aut}_{a_i}(\lambda_i)$, $i = 1, \dots, p$, tale che

$$M_{\mathcal{B}}(f) = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & & 0 \\ \cdots & & \ddots & \\ 0 & 0 & & J_r \end{pmatrix},$$

con $r \geq p$, matrice diagonale a blocchi, dove ogni blocco J_i è un blocco di Jordan relativo a uno degli autovalori. Inoltre:

1. la forma canonica di Jordan è unica a meno di permutazioni dei blocchi;

2. il numero di blocchi di ogni autovalore λ_i è uguale a $m_g(\lambda_i)$;

3. preso un autovalore λ, e posto $g = f - \lambda \operatorname{id}_V$, la massima dimensione di un blocco di Jordan di f relativo a λ è il minimo esponente k per cui $\ker g^k$ ha dimensione $m_a(\lambda)$.

Dimostrazione. Omessa. □

Se f ha un solo autovalore λ, con $m_a(\lambda) = n$, la massima dimensione di un blocco di Jordan è il minimo k tale che $\dim \ker g^k = n$. Se f corrisponde alla matrice A e g alla matrice

$B = A - \lambda \mathbb{I}_n$, allora $\dim \ker g^k = n$ se e solo se $\ker g^k = V$ se e solo se $g^k = 0$ se e solo se $B^k = 0$.

Quindi la massima dimensione di un blocco di Jordan è il minimo k tale che $B^k = 0$. In questo caso B è una matrice **nilpotente**, cioè con una potenza nulla, e il minimo k per cui ciò accade è detto **indice di nilpotenza** di B.