§2 Operatori diagonalizzabili

Abbiamo visto che la matrice associata a un operatore dipende dalla base; ci si può chiedere se, dato un operatore, esista una base in cui la matrice che lo rappresenta sia diagonale. Se questo accade diremo che l'operatore è diagonalizzabile. Come vedremo non sempre un operatore è diagonalizzabile, e, anche quando lo è, non necessariamente lo è in una base di vettori reali (un operatore che sia diagonalizzabile sui complessi si dirà semisemplice). In questo paragrafo studieremo in dettaglio il caso degli operatori diagonalizzabili; vedremo più avanti, nel §4, come trattare gli operatori che non siano diagonalizzabili.

Definizione 2.1 (Operatore diagonalizzabile) *Dato uno spazio vettoriale reale E, un operatore lineare $T \in L(E)$ si dice* diagonalizzabile *se esiste una base $\{e_1, \ldots, e_n\}$ in cui l'operatore è* diagonale*, i.e. in cui T è rappresentato da una matrice diagonale $A_{ij} = \lambda_i \delta_{ij}$.*

Definizione 2.2 (Autovalori e autovettori) *Dato un operatore $T \in L(E)$, si definisce* polinomio caratteristico *di T il polinomio (di grado n in λ)*

$$p_n(\lambda) := \det (T - \lambda \mathbb{1}) . \tag{2.1}$$

Le radici $\lambda_1, \ldots, \lambda_n$ del polinomio caratteristico si chiamano autovalori *di T. L'insieme degli autovalori di T, che sarà indicato con $\Sigma(T) = \{\lambda_1, \ldots, \lambda_n\}$, prende il nome di* spettro *dell'operatore T. Dato un autovalore λ_i, $i = 1, \ldots, n$, si definisce* autovettore *associato all'autovalore λ_i un qualsiasi vettore non nullo v_i che sia soluzione dell'equazione*

$$(T - \lambda_i \mathbb{1})\, v_i = 0. \tag{2.2}$$

Un operatore $T \in L(E)$ avrà sempre n autovalori, se $n = \dim(E)$, poiché n sono le radici del polinomio caratteristico – come conseguenza del teorema fondamentale dell'algebra (cfr. gli esercizi 29 e 30). In generale gli autovalori di T sono numeri complessi, anche se gli elementi della matrice che rappresenta T sono reali. Inoltre, alcuni autovalori possono essere coincidenti, dal momento che un polinomio di grado n può avere radici coincidenti. Poiché ogni autovalore è una radice del polinomio caratteristico, e quindi $\det(T - \lambda_i \mathbb{1}) = 0$, l'equazione (2.2) ammette sempre una soluzione non nulla (cfr. l'esercizio 31); tuttavia dalla (2.2) possiamo concludere solo che a ogni autovalore λ_i corrisponde un autovettore v_i, i.e. che il numero di autovettori non può essere minore del numero di autovalori distinti. Vale tuttavia il seguente risultato.

Teorema 2.3 *Se un operatore $T \in L(E)$ ha autovalori reali distinti, allora è diagonalizzabile e la base in cui è rappresentato da una matrice diagonale è la base degli autovettori.*

Dimostrazione. Se $\lambda_1, \ldots, \lambda_n$ sono gli autovalori di T e $v_1, \ldots, v_n$ sono gli autovettori corrispondenti, dobbiamo dimostrare che $\{v_1, \ldots, v_n\}$ costituisce una base in E. Infatti in tal caso, nella base degli autovettori, si ha

$$(Tv_j)_i = \lambda_j (v_j)_i = \lambda_j \delta_{ij}, \tag{2.3}$$

come segue dalla (1.14) e dalla (2.2). La (2.3) mostra allora che nella base $\{v_1, \dots, v_n\}$ l'operatore T è diagonale.

Supponiamo per assurdo che $\{v_1, \dots, v_n\}$ non costituisca una base. Sarà allora possibile esprimere (almeno) uno dei vettori, per esempio v_n, come combinazione lineare di m degli altri linearmente indipendenti, con $m \leq n-1$; riordinando eventualmente i vettori, possiamo supporre che essi siano i primi m. Si avrà

$$v_n = \sum_{i=1}^{m} t_i v_i,$$

con i coefficienti t_i non tutti nulli, i.e. $(t_1, \dots, t_m) \neq (0, \dots, 0)$, e quindi, per la (2.2),

$$0 = (T - \lambda_n \mathbb{1})\, v_n = (T - \lambda_n \mathbb{1}) \sum_{i=1}^{m} t_i v_i = \sum_{i=1}^{m} (T - \lambda_n \mathbb{1})\, t_i v_i = \sum_{i=1}^{m} (\lambda_i - \lambda_n)\, t_i v_i, \tag{2.4}$$

dove si è usata la linearità di T. Per ipotesi $\lambda_i \neq \lambda_j$ per ogni $i \neq j$, quindi l'ultima formula in (2.4) esprime una combinazione lineare a coefficienti non tutti nulli dei vettori $v_1, \dots, v_m$. Poiché si è supposto che tali vettori siano linearmente indipendenti, tale combinazione lineare non può essere nulla. Si è perciò arrivati a una contraddizione. ■

Corollario 2.4 *Se l'operatore $T \in L(E)$ ha autovalori reali distinti $\lambda_1, \dots, \lambda_n$ e A è la matrice che lo rappresenta in una base qualsiasi $\{e_1, \dots, e_n\}$, allora esiste una matrice Q tale che $D := QAQ^{-1}$, che rappresenta T nella base degli autovettori, è la matrice diagonale*

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Dimostrazione. Che D rappresenti T nella base degli autovettori $\{v_1, \dots, v_n\}$ segue dalla (2.3); basta allora definire Q come la matrice che fa passare dal sistema di coordinate nella base $\{e_1, \dots, e_n\}$ al sistema di coordinate nella base $\{v_1, \dots, v_n\}$ e applicare il lemma 1.42. ■

Un discorso analogo vale qualora gli autovalori di un operatore T definito in uno spazio vettoriale reale siano sempre distinti, ma complessi (e non più necessariamente reali). Tuttavia in tal caso occorre considerare uno spazio vettoriale esteso, in cui siano compresi anche vettori complessi e l'operazione di moltiplicazione sia definita per scalari complessi.

Definizione 2.5 (Complessificazione di uno spazio vettoriale) *Sia E uno spazio vettoriale reale. Definiamo la sua* complessificazione *come lo spazio vettoriale complesso*

$$E_C = \Big\{ z : z = \sum_i \lambda_i x_i, \quad x_i \in E, \quad \lambda_i \in \mathbb{C} \Big\},$$

i.e. come lo spazio vettoriale i cui vettori si possono scrivere come combinazione lineare dei vettori di E a coefficienti complessi.

Sia σ l'operatore di *coniugazione complessa*. Se F è uno spazio vettoriale complesso, con una sua base qualsiasi, e $z = (z_1, \ldots, z_n) \in F$, allora

$$\sigma(z) = \sigma(z_1, \ldots, z_n) = (\bar{z}_1, \ldots, \bar{z}_n) =: \bar{z}, \tag{2.5}$$

dove $\bar{z}_i$ è il complesso coniugato di z_i (cfr. pag. 8). Per costruzione σ è un'*involuzione* in E, i.e. $\sigma^2 := \sigma \circ \sigma = \mathbb{1}$.

Definizione 2.6 (Decomplessificazione di uno spazio vettoriale) *Sia F uno spazio vettoriale complesso tale che $\sigma(F) \subset F$. Definiamo la sua* decomplessificazione *come lo spazio vettoriale reale*

$$F_R = \{z \in F : \sigma(z) = z\}.$$

i.e. come l'insieme dei vettori reali in F.

Osservazione 2.7 La condizione $\sigma(F) \subset F$ è essenziale per poter definire la decomplessificazione di uno spazio vettoriale complesso. Infatti se $z = x + iy \in F$, dove i è l'unità immaginaria, allora solo se anche $\bar{z} = x - iy$ appartiene a F si ha

$$x = \frac{z + \bar{z}}{2} \in F,$$

in quanto combinazione lineare di vettori dello spazio vettoriale F, così che F_R costituisce un sottoinsieme di F che ha ancora struttura di spazio vettoriale.

Osservazione 2.8 Ovviamente la decomplessificazione della complessificazione di uno spazio vettoriale reale E è lo spazio vettoriale E stesso, i.e. $E_{CR} = E$.

Definizione 2.9 (Complessificazione di un operatore) *Sia $T \in L(E)$ e sia E_C la complessificazione di E. Definiamo la* complessificazione *$T_C \in L(E_C)$ dell'operatore T nel modo seguente: se $z \in E_C$, e quindi*

$$z = \sum_i \lambda_i x_i, \tag{2.6}$$

allora

$$T_C z = \sum_i \lambda_i T x_i.$$

Proposizione 2.10 *Sia E uno spazio vettoriale reale ed E_C la sua complessificazione. Se $Q \in L(E_C)$ esiste un operatore $T \in L(E)$ tale che $Q = T_C$ se e solo se*

$$\sigma \circ Q = Q \circ \sigma,$$

dove σ è la coniugazione complessa $\sigma \colon E_C \to E_C$.

Dimostrazione. Se $Q = T_C$, per qualche $T \in L(E)$, allora se z è della forma (2.6) si ha

$$\begin{aligned} Q \circ \sigma(z) &= Q\sigma(z) = Q\sigma\Big(\sum_i \lambda_i x_i\Big) = Q\Big(\sum_i \bar\lambda_i x_i\Big) = T_C\Big(\sum_i \bar\lambda_i x_i\Big) = \sum_i \bar\lambda_i T x_i \\ &= \sigma\Big(\sum_i \lambda_i T x_i\Big) = \sigma(Qz) = \sigma \circ Q(z) \qquad \Longrightarrow \qquad Q \circ \sigma = \sigma \circ Q. \end{aligned}$$

Viceversa, se $Q \circ \sigma = \sigma \circ Q$, si ha, per $x \in E$,

$$\sigma(Qx) = Q\sigma(x) = Qx,$$

così che

$$Qx \in \{y \in E_C : \sigma(y) = y\} = E_{CR} = E,$$

i.e. $QE \subset E$. Pertanto $T = Q|E$, la *restrizione* di Q a E, è un operatore lineare in $L(E)$, così che possiamo porre $Q = T_C$. ■

Osservazione 2.11 Il significato della proposizione 2.10 è il seguente. Dato un operatore $T \in L(E)$ è sempre possibile considerarne la complessificazione T_C, mentre dato un operatore $Q \in L(E_C)$ non è detto che esista un operatore $T \in L(E)$ di cui Q sia la complessificazione. In generale questo non è vero: una condizione necessaria e sufficiente perché questo sia possibile è che Q commuti con l'operatore di coniugazione complessa σ.

Lemma 2.12 *Sia $T \in L(E)$ e T_C la sua complessificazione. Allora T e T_C hanno gli stessi autovalori.*

Dimostrazione. Basta notare che T e T_C hanno lo stesso polinomio caratteristico (2.1): infatti, data una base, T e T_C sono rappresentati dalla stessa matrice (l'unica differenza è che il primo agisce sui vettori di E e l'altro sui vettori di E_C). Poiché gli autovalori sono le radici del polinomio caratteristico, T e T_C devono avere gli stessi autovalori. ■

Teorema 2.13 *Se un operatore $T \in L(E)$ ha autovalori complessi distinti $\lambda_1, \ldots, \lambda_n$, allora $T_C \in L(E_C)$ è in forma diagonale nella base degli autovettori $\{v_1, \ldots, v_n\}$.*

Dimostrazione. La dimostrazione è identica a quella del teorema 2.3. L'unica accortezza è che dobbiamo considerare la complessificazione T_C di T: allora $T_C \in L(E_C)$ ha autovalori distinti e quindi è in forma diagonale nella base degli autovettori (complessi). ■

Definizione 2.14 *Un operatore $T \in L(E)$ si dice* semisemplice *se esiste una base $\{e_1, \ldots, e_n\}$ in E_C in cui l'operatore T_C è* diagonale.

Osservazione 2.15 Si noti che un operatore $T \in L(E)$ è semisemplice se la sua complessificazione $T_C \in L(E_C)$ è diagonalizzabile in E_C. In particolare, un operatore lineare semisemplice T i cui autovalori sono reali è anche diagonalizzabile in E.

Proposizione 2.16 *Sia T un operatore lineare in uno spazio vettoriale reale E di dimensione n. I suoi autovalori si possono dividere in due classi, la prima costituita da s autovalori reali $\lambda_1, \ldots, \lambda_s$ e la seconda costituita da r coppie di autovalori complessi coniugati $\mu_1, \bar{\mu}_1, \ldots, \mu_r, \bar{\mu}_r$, con $s + 2r = n$. Gli autovettori corrispondenti a una coppia di autovalori complessi coniugati sono a loro volta complessi coniugati.*

Dimostrazione. Basta dimostrare che lo spettro di T è preservato dall'operazione di coniugazione σ introdotto in (2.5). Sia T_C la complessificazione di T, e siano μ un autovalore complesso di T_C (cfr. il lemma 2.12) e φ l'autovettore associato a μ. Notiamo innanzitutto che essendo μ complesso deve essere φ complesso. Si ha inoltre

$$T_C \bar{\varphi} = T_C \sigma(\varphi) = \sigma(T_C \varphi) = \sigma(\mu \varphi) = \bar{\mu} \bar{\varphi},$$

dove si è usata la proposizione 2.10. Ne segue che $\bar{\mu}$ è un autovalore di T e $\bar{\varphi}$ è l'autovettore associato a $\bar{\mu}$, per il lemma 2.12. ∎

Dato un operatore $T \in L(E)$ con autovalori distinti, in parte reali e in parte complessi (non reali), ordinati come nella proposizione 2.16, possiamo decomporre lo spazio vettoriale E nella somma diretta di due sottospazi $E_a \oplus E_b$, generati, rispettivamente, dagli autovettori associati agli autovalori reali e dagli autovettori associati agli autovalori complessi. Corrispondentemente risulterà determinata una decomposizione dell'operatore T nella somma diretta di due operatori, $T = T_a \oplus T_b$. Infatti $TE_a \subset E_a$ e $TE_b \subset E_b$, poiché l'azione di T_C su un autovettore è semplicemente la moltiplicazione per il corrispondente autovalore, e quindi combinazioni lineari di autovettori vengono trasformati da T_C in combinazioni lineari degli stessi autovettori. Possiamo allora definire $T_a = T|E_a$ e $T_b = T|E_b$; in particolare se consideriamo la restrizione di T al sottospazio vettoriale generato dagli autovettori associati a una coppia di autovalori complessi coniugati, otteniamo un operatore invariante.

In conclusione, possiamo scrivere $E_a = E_{a1} \oplus \ldots \oplus E_{ar}$ e, analogamente, $E_b = E_{b1} \oplus \ldots \oplus E_{bs}$, dove E_{ai} è il sottospazio generato dall'autovettore v_i associato all'autovalore λ_i, con $i = 1, \ldots, r$, ed E_{bi} è il sottospazio generato dagli autovettori φ_i e $\bar{\varphi}_i$ associati agli autovalori μ_i e $\bar{\mu}_i$, con $i = 1, \ldots, s$. Si ha $T_a = T_{a1} \oplus \ldots \oplus T_{ar}$ e $T_b = T_{b1} \oplus \ldots \oplus T_{bs}$. L'operatore T_{ai} agisce (sui vettori in E_{ai}) come un operatore di moltiplicazione per lo scalare λ_i, mentre l'operatore T_{bi} agisce (sui vettori in E_{bi}) come una matrice 2×2, la cui forma è data dal lemma seguente.

Lemma 2.17 *Sia E uno spazio vettoriale reale bidimensionale e sia $T \in L(E)$ un operatore lineare in E con autovalori non reali $\mu, \bar{\mu}$, dove $\mu = a + ib$. Esiste allora una base $\{v, u\}$ in cui la matrice A che rappresenta T ha la forma*

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \tag{2.7}$$

e u, v sono tali che $\varphi = u + iv$ è l'autovettore associato a μ.

Dimostrazione. Dati gli autovalori $\mu, \bar{\mu}$, siano $\varphi, \bar{\varphi}$ gli autovettori associati. Sia E_C la complessificazione di E. Si ha allora $\varphi, \bar{\varphi} \in E_C$. Poniamo $\varphi = u + iv$. Poiché

$$u = \frac{\varphi + \bar{\varphi}}{2}, \qquad v = \frac{\varphi - \bar{\varphi}}{2i},$$

risulta $u, v \in E$. Si ha inoltre

$$T_C \varphi = \mu\varphi = (a + ib)(u + iv) = (au - bv) + i(bu + av). \tag{2.8}$$

D'altra parte

$$T_C \varphi = T_C (u + iv) = Tu + iTv, \tag{2.9}$$

così che, eguagliando parte reale e parte immaginaria delle (2.8) e (2.9), otteniamo

$$Tv = bu + av, \qquad Tu = -bv + au.$$

Poiché $\varphi, \bar{\varphi}$ sono linearmente indipendenti, lo sono anche v, u. In conclusione $\{v, u\}$ è una base di E. In tale base l'operatore T è rappresentato dalla matrice di elementi

$$A_{11} = (Tv)_1 = a, \qquad A_{12} = (Tu)_1 = -b, \qquad A_{21} = (Tv)_2 = b, \qquad A_{22} = (Tu)_2 = a,$$

(cfr. la (1.14)), da cui segue l'asserto. ∎

Osservazione 2.18 Dato un operatore T definito in uno spazio vettoriale reale con autovalori complessi coniugati $\mu, \bar{\mu}$, con $\mu = a + ib$, $b \neq 0$, sia $\{\varphi, \bar{\varphi}\}$ la base degli autovettori, con $\varphi = u + iv$. La matrice P che esprime i vettori della base $\{\varphi, \bar{\varphi}\}$ in funzione della base $\{v, u\}$ è data da

$$\begin{pmatrix} \varphi \\ \bar{\varphi} \end{pmatrix} = P \begin{pmatrix} v \\ u \end{pmatrix}, \qquad P = \begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix},$$

così che, ricordando il lemma 1.41,

$$Q^{-1} = P^T = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}, \qquad Q = \frac{1}{2i} \begin{pmatrix} 1 & i \\ -1 & i \end{pmatrix},$$

Se $y = (y_1, y_2)$ è il sistema di coordinate nella base $\{\varphi, \bar{\varphi}\}$ e $z = (z_1, z_2)$ è il sistema di coordinate nella base $\{v, u\}$, si ha

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = Q \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \frac{1}{2i} \begin{pmatrix} 1 & i \\ -1 & i \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix},$$

e

$$D := \begin{pmatrix} \mu & 0 \\ 0 & \bar{\mu} \end{pmatrix} = Q \begin{pmatrix} a & -b \\ b & a \end{pmatrix} Q^{-1} = \frac{1}{2i} \begin{pmatrix} 1 & i \\ -1 & i \end{pmatrix} \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}. \tag{2.10}$$

La (2.10) permette di esprimere la matrice A del lemma 2.17 in termini della matrice diagonale D che rappresenta l'operatore T nella base degli autovettori (complessi), i.e. $A = Q^{-1}DQ$.