Capitolo 9

Diagonalizzazione

Ci occuperemo ora del problema di trovare un rappresentante particolare nella classe di similitudine di una data matrice A di $M_n(\mathbb{K})$. La prima domanda che ci poniamo è quando la classe di A contiene una matrice diagonale.

9.1 Autovalori, autovettori e autospazi

Definizione 9.1.1 (Autovalori e autovettori)**.**

1. Sia V un $\mathbb{K}$-spazio vettoriale e sia $f : V \to V$ un endomorfismo di V. Uno scalare $\lambda \in \mathbb{K}$ è un **autovalore di** f se esiste un vettore $v \in V$ **non nullo** tale che $f(v) = \lambda v$. In tal caso v è detto **autovettore di** f **di autovalore** λ**.**

2. Sia $A \in M_n(\mathbb{K})$ una matrice quadrata. Uno scalare λ è un **autovalore di** A se esiste un vettore $v \in \mathbb{K}^n$ **non nullo** tale che $A \cdot v = \lambda v$. In tal caso v è detto **autovettore di** A **di autovalore** λ**.**

Dalla definizione si che λ, v sono rispettivamente un autovalore e un autovettore di A se e solo se lo sono per l'endomorfismo $L_A : \mathbb{K}^n \to \mathbb{K}^n$.

Notiamo che **gli autovettori sono vettori non nulli**, altrimenti la nozione di autovalore avrebbe poco senso. Infatti, per ogni $\lambda \in \mathbb{K}$, indicato con 0_V il vettore nullo di V, si ha $f(0_V) = 0_V = \lambda 0_V$.

Se richiediamo, invece, che un autovettore sia $v \neq 0_V$, l'autovalore ad esso associato è unico. Se per assurdo si avesse

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

avremmo $(\lambda - \mu)v = 0_V$, ed essendo $v \neq 0_V$, si ha necessariamente $\lambda - \mu = 0$.

Osservazione 36. Sia V uno spazio vettoriale di dimensione finita. Osserviamo che λ è un autovalore di f se e solo se λ è un autovalore della matrice $M = M^{\mathcal{B}}_{\mathcal{B}}(f)$, per ogni base $\mathcal{B}$ di V. Infatti se $a_1, \ldots, a_n$ sono le coordinate di v rispetto a $\mathcal{B}$, si ha $f(v) = \lambda v$ se e solo se

$$M \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \lambda \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

Definizione 9.1.2 (Autospazi)**.** Sia $f : V \to V$ un endomorfismo, $\lambda \in K$. L'insieme $\mathrm{Aut}(\lambda) = \{v \in V \mid f(v) = \lambda v\} = \{0_V\} \cup \{\text{autovettori di } \lambda\}$ è detto **autospazio di** λ.

Analoga definizione nel caso di una matrice $A \in M(n \times n, K)$. L'autospazio di λ è $\mathrm{Aut}(\lambda) = \{x \in K^n \mid Ax = \lambda x\} = \{0_{K^n}\} \cup \{\text{autovettori di } \lambda\}$.

Osserviamo che $\mathrm{Aut}(\lambda)$ è un sottospazio vettoriale di V, risp. di K^n, e $\mathrm{Aut}(\lambda) \neq (0)$ se e solo se λ è autovalore. Infatti $0 \in \mathrm{Aut}(\lambda)$; inoltre se $v, w \in \mathrm{Aut}(\lambda)$ si ha $f(v) = \lambda v$, $f(w) = \lambda w$, dunque, presi due scalari $a, b \in \mathbb{K}$, per la linearità di f si ha $f(av + bw) = af(v) + bf(w) = a\lambda v + b\lambda w = \lambda(av + bw)$, perciò $av + bw \in \mathrm{Aut}(\lambda)$.

Osservazione 37. L'autospazio di uno scalare λ si può anche esprimere come nucleo di un endomorfismo. Infatti $\mathrm{Aut}(\lambda) = \ker(f - \lambda\,\mathrm{id}_V)$, dove

$$\begin{array}{rccl} f - \lambda\,\mathrm{id}_V : & V & \to & V \\ & v & \to & (f - \lambda\,\mathrm{id}_V)(v) = f(v) - \lambda v. \end{array}$$

Analogamente per le matrici: $v \in \mathrm{Aut}(\lambda)$, autospazio di una matrice A, se e solo se $v \in \ker(L_A - \lambda\,\mathrm{id}_{\mathbb{K}^n}) = \ker(L_A - \lambda L_{\mathbb{I}_n}) = \ker(L_{A-\lambda\mathbb{I}_n})$. Quindi gli autovettori di λ sono le soluzioni non nulle del sistema lineare $(A - \lambda\mathbb{I}_n) \cdot X = 0$.

Definizione 9.1.3. La **molteplicità geometrica** dello scalare λ, denotata $m_g(\lambda)$, è la dimensione del relativo autospazio: $m_g(\lambda) = \dim \mathrm{Aut}(\lambda)$.

Se λ è un autovalore, dalla definizione segue immediatamente che $m_g(\lambda) \geq 1$.

Esempi 9.1.4. Sia $A = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ e sia $v = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$.

Si ha $Av = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 2b_1 \\ 0 \end{pmatrix}$, $\lambda v = \begin{pmatrix} \lambda b_1 \\ \lambda b_2 \end{pmatrix}$. Dunque $Av = \lambda v$ se e solo se $\begin{cases} 2b_1 = \lambda b_1 \\ 0 = \lambda b_2 \end{cases}$,

cioè se e solo se $\begin{cases} (2-\lambda)b_1 = 0 \\ \lambda b_2 = 0 \end{cases}$. Questo accade, con $v \neq 0$, se e solo se $\lambda = 2$ e $v = \begin{pmatrix} b_1 \\ 0 \end{pmatrix}$, oppure $\lambda = 0$ e $v = \begin{pmatrix} 0 \\ b_2 \end{pmatrix}$. In particolare, gli autovalori di A sono $\lambda_1 = 2$ e $\lambda_2 = 0$.

Per $\lambda_1 = 2$, i relativi autovettori sono tutti i vettori non nulli con $b_2 = 0$; quindi $Aut(2) = \mathrm{Span}(e_1) \subset \mathbb{K}^2$.

Se consideriamo $\lambda_2 = 0$, i relativi autovettori sono tutti i vettori non nulli con $b_1 = 0$; quindi $Aut(0) = \mathrm{Span}(e_2) \subset \mathbb{K}^2$.

In altre parole, la restrizione di L_A alla retta $\mathrm{Span}(e_1)$ è la moltiplicazione per 2; mentre la restrizione alla retta $\mathrm{Span}(e_2)$ è la mappa nulla.

Teorema 9.1.5. *Autovettori relativi ad autovalori distinti sono linearmente indipendenti.*

Ovvero, se $\lambda_1, \dots, \lambda_k$ sono autovalori di f, con $\lambda_i \neq \lambda_j$ per ogni $i \neq j$, e se $v_1, \dots, v_k$ sono autovettori relativi rispettivamente a $\lambda_1, \dots, \lambda_k$, allora $v_1, \dots, v_k$ sono linearmente indipendenti.

Dimostrazione. Per induzione su k. Se $k = 1$, il teorema è vero perché ogni autovettore è non nullo.

Sia $k \geq 2$ e supponiamo che la tesi del teorema sia vera per $k - 1$. Consideriamo una combinazione lineare di $v_1, \dots, v_k$ che dia il vettore nullo:

$$\mu_1 v_1 + \cdots + \mu_k v_k = 0. \tag{9.1}$$

Allora $f(\mu_1 v_1 + \cdots + \mu_k v_k) = f(0_V) = 0_V$; inoltre, per la linearità di f si ha

$$f(\mu_1 v_1 + \cdots + \mu_k v_k) = \mu_1 f(v_1) + \cdots + \mu_k f(v_k) \overset{\text{autovettori}}{=} \mu_1(\lambda_1 v_1) + \cdots + \mu_k(\lambda_k v_k) = \lambda_1(\mu_1 v_1) + \cdots + \lambda_k(\mu_k v_k).$$

Per usare l'ipotesi induttiva sostituiamo in (9.1) a $\mu_k v_k$ l'espressione $-\mu_1 v_1 - \cdots - \mu_{k-1} v_{k-1}$ e otteniamo:

$$\lambda_1 \mu_1 v_1 + \cdots + \lambda_{k-1} \mu_{k-1} v_{k-1} - \lambda_k(\mu_1 v_1 + \cdots + \mu_{k-1} v_{k-1}) = 0$$

da cui

$$(\lambda_1 \mu_1 - \lambda_k \mu_1) v_1 + \cdots + (\lambda_{k-1} \mu_{k-1} - \lambda_k \mu_{k-1}) v_{k-1} = 0.$$

Per ipotesi induttiva $v_1, \dots, v_{k-1}$ sono linearmente indipendenti, quindi otteniamo

$$\begin{cases} \lambda_1 \mu_1 - \lambda_k \mu_1 = (\lambda_1 - \lambda_k) \mu_1 = 0 \\ \dots \\ (\lambda_{k-1} - \lambda_k) \mu_{k-1} = 0 \end{cases}$$

Essendo gli autovalori tutti distinti $\lambda_1 - \lambda_k \neq 0, \dots, \lambda_{k-1} - \lambda_k \neq 0$, e quindi $\mu_1 = \cdots = \mu_{k-1} = 0$. Allora la (9.1) diventa $\mu_k v_k = 0$, ma $v_k \neq 0$, e concludiamo che anche $\mu_k = 0$. $\square$

Corollario 9.1.6. *Se f è un endomorfismo di V con $\dim V = n$, può avere al più n autovalori distinti. Analogamente, una matrice $n \times n$ può avere al più di n autovalori distinti.*

Osserviamo che $0 \in \mathbb{K}$ può essere un autovalore di f (si veda l'esempio 9.1.4); ciò accade se e solo se esiste un vettore non nullo v tale che $f(v) = 0v = 0_V$, cioè se e solo se f non è iniettiva. In tal caso $\mathrm{Aut}(0) = \ker(f)$. Analogamente per una matrice A, lo scalare 0 è un autovalore se e solo se L_A non è iniettiva, ovvero $\ker L_A \neq 0$ ovvero $\mathrm{rg}(A) < n$. In tal caso $\mathrm{Aut}(0) = \ker L_A$.

9.2 Ricerca degli autovalori e polinomio caratteristico

Se $f : V \to V$ è un endomorfismo di V, spazio vettoriale di dimensione finita, abbiamo visto che λ è un autovalore di f se e solo se $\mathrm{Aut}(\lambda) = \ker(f - \lambda\,\mathrm{id}_V)$ è un sottospazio vettoriale non nullo; ciò succede se e solo se $f - \lambda\,\mathrm{id}_V$ non è un isomorfismo, ovvero se e solo se il determinante

$$\det(f - \lambda\,\mathrm{id}_V) = 0.$$

Se $\mathcal{B}$ è una base di V, si ha

$$\det(f - \lambda\,\mathrm{id}_V) = \det(M^{\mathcal{B}}_{\mathcal{B}}(f - \lambda\,\mathrm{id}_V)) = \det(M^{\mathcal{B}}_{\mathcal{B}}(f) - \lambda M^{\mathcal{B}}_{\mathcal{B}}(\mathrm{id}_V)) = \det(M^{\mathcal{B}}_{\mathcal{B}}(f) - \lambda \mathbb{I}_n).$$

Ponendo $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$ concludiamo che

$$\lambda \text{ è autovalore di } f \text{ se e solo se } \det(A - \lambda\mathbb{I}_n) = 0. \tag{9.2}$$

In particolare, se A è una matrice quadrata

$$\lambda \text{ è autovalore di } A \text{ se e solo se } \det(A - \lambda\mathbb{I}_n) = 0. \tag{9.3}$$

Osservazione 38. Osserviamo che se A e A' sono matrici simili,

$$\det(A - \lambda\mathbb{I}_n) = \det(A' - \lambda\mathbb{I}_n).$$

Infatti, se $A' = S \cdot A \cdot S^{-1}$, si ha

$$\det(S \cdot A \cdot S^{-1}) = \det(S \cdot A \cdot S^{-1} - \lambda S \cdot \mathbb{I}_n \cdot S^{-1}) = \det(S \cdot (A - \lambda\mathbb{I}_n) \cdot S^{-1}) = \det(A - \lambda\mathbb{I}_n),$$

per il Corollario 8.8.3 del Teorema di Binet.

Sia x un'indeterminata, e consideriamo la matrice $A - x\mathbb{I}_n$. Scritta per esteso è del tipo:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} - x \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix} = \begin{pmatrix} a_{11} - x & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - x & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - x \end{pmatrix} = (a_{ij} - x\delta_{ij})_{i,j=1,\dots,n}.$$

Sviluppiamo il determinante di $A - x\mathbb{I}_n$ con la formula di Leibniz (Teorema 8.1):

$$\det(A - x\mathbb{I}_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)(a_{1\sigma(1)} - x\delta_{1\sigma(1)}) \cdots (a_{n\sigma(n)} - x\delta_{n\sigma(n)}).$$

Analizziamo i vari addendi:

- per $\sigma = \mathrm{id}_{I_n}$ abbiamo $(a_{11} - x)(a_{22} - x) \cdots (a_{nn} - x)$; sviluppando troviamo $(-1)^n x^n + (-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn})x^{n-1}$ + termini di grado più basso;
- per ogni altra permutazione σ troviamo un prodotto di n fattori di cui al più $n-2$ sono sulla diagonale principale di $A - x\mathbb{I}_n$, in quanto c'è esattamente un fattore per ogni riga e uno per ogni colonna; quindi tale prodotto è un polinomio in x di grado $\leq n - 2$.

In definitiva povveromo scrivere

$$\det(A - x\mathbb{I}_n) = (-1)^n x^n + (-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn})x^{n-1} + \alpha_{n-2}x^{n-2} + \cdots + \alpha_1 x + \det(A); \quad (9.4)$$

l'ultimo termine è il termine noto che si ottiene ponendo $x = 0$, e i coefficienti $\alpha_1, \ldots, \alpha_{n-2}$ sono scalari. Abbiamo quindi ottenuto un polinomio di grado n in x, a coefficienti in $\mathbb{K}$.

Definizione 9.2.1. Il polinomio $\det(A - x\mathbb{I}_n) = \det(f - xE_n)$ è detto **polinomio caratteristico** di f, e denotato $p_f(x)$.

Per l'Osservazione 38, il polinomio $p_f(x)$ dipende solo da f, non dalla base scelta nè dalla relativa matrice $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$. In particolare, **tutti i coefficienti di $p_f(x)$ dipendono soltanto da f, ed anche il suo termine noto che è det(f).** Inoltre **anche le radici del polinomio caratteristico dipendono solo da** f, cioè gli elementi λ di $\mathbb{K}$ tali che $p_f(\lambda) = 0$.

Il coefficiente di $(-1)^{n-1}x^{n-1}$ è $a_{11} + a_{22} + \cdots + a_{nn}$, ed è detto **traccia di** f; è la somma degli elementi della diagonale principale della matrice A.

Da questa discussione abbiamo ottenuto:

Proposizione 9.2.2. *Sia $f : V \to V$ un endomorfismo di un $\mathbb{K}$-spazio vettoriale di dimensione finita. Gli autovalori di f sono le radici in $\mathbb{K}$ del polinomio caratteristico $p_f(x)$, ovvero sono le soluzioni in $\mathbb{K}$ dell'equazione algebrica $p_f(x) = 0$.*

Se anzichè partire dall'endomorfismo f si parte da una matrice, il polinomio caratteristico di A è $p_A(x) = \det(A - x\mathbb{I}_n)$.

Osservazione 39. Se λ è un autovalore di f, allora è una radice di $p_f(x)$, e per il teorema di Ruffini $p_f(x)$ è divisibile per $x-\lambda$, cioè esiste un polinomio $q_1(x)$ di grado $n-1$ tale che

$$p_f(x) = (x-\lambda)q(x).$$

Ora povveromo considerare $q_1(\lambda)$: se è $\neq 0$ abbiamo finito, altrimenti applichiamo di nuovo il teorema di Ruffini e otteniamo che esiste un secondo polinomio $q_2(x)$ tale che $q_1(x) = (x-\lambda)q_2(x)$, e dunque $p_f(x) = (x-\lambda)^2 q_2(x)$. Povveromo ripetere il procedimento, al massimo per n volte, fino ad arrivare ad esprimere il polinomio caratteristico nella forma

$$p_f(x) = (x-\lambda)^m q(x), \quad q(\lambda) \neq 0.$$

Definizione 9.2.3. Tale numero m è detto **molteplicità algebrica dell'autovalore** λ e denotato con $m_a(\lambda)$. È il massimo esponente di $x-\lambda$ che divide $p_f(x)$.

9.3 Esempi I

Esempi 9.3.1. Riprendiamo l'Esempio 9.1.4. Il polinomio caratteristico della matrice $A = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ è $p_A(x) = \mid A - xE_2 \mid = \det \begin{pmatrix} 2-x & 0 \\ 0 & -x \end{pmatrix} = -(2-x)x = (x-2)x$. Tale polinomio è già fattorizzato in due fattori della forma $x-\lambda$, da cui si vede che vi sono due autovalori, e precisamente 0 e 2, ciascuno con molteplicità algebrica 1. Ciò conferma quanto trovato in precedenza con un ragionamento diretto.

Esempi 9.3.2. Sia $V = \mathbb{R}^2$; consideriamo la matrice

$$R_\alpha := \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

dove $\alpha \in \mathbb{R}$. L'endomorfismo L_{R_α} di $\mathbb{R}^2$ è la **rotazione di angolo** α. In tale rotazione i vettori della base canonica vengono mandati rispettivamente nei vettori $w_1 = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$, e $w_2 = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}$ (sono le colonne di R_α). Il polinomio caratteristico è

$$p_{R_\alpha}(x) = \det \begin{pmatrix} \cos\alpha - x & -\sin\alpha \\ \sin\alpha & \cos\alpha - x \end{pmatrix} = x^2 - 2\cos\alpha\, x + 1.$$

Per trovare gli autovalori dobbiamo dunque risolvere l'equazione di secondo grado

$$x^2 - 2\cos\alpha\, x + 1 = 0. \tag{9.5}$$

Dal momento che $\frac{\Delta}{4} = \cos^2\alpha - 1 = -\sin^2\alpha$ è sempre ≤ 0, povveromo distinguere due casi:

1. $\sin^2 \alpha \neq 0$: in tal caso l'equazione (9.5) non ha soluzioni reali, dunque la rotazione non ha autovalori nè autovettori; ciò si verifica se α non è un multiplo intero di π;

2. $\sin^2 \alpha = 0$; in tal caso $\alpha = 2k\pi$ oppure $\alpha = \pi + 2k\pi$, $k \in \mathbb{Z}$. Se $\alpha = 2k\pi$, si ha $R_0 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{I}_2$; quindi $p_{R_0}(x) = (x-1)^2$, dunque c'è il solo autovalore $\lambda = 1$ con $m_a(1) = 2$. Inoltre, $R_0 - \lambda \mathbb{I}_2$ è la matrice nulla, dunque $\mathrm{Aut}(1) = \mathbb{R}^2$, e $m_g(1) = 2$.

   Se invece $\alpha = \pi + 2k\pi$, si ha $R_\pi = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\mathbb{I}_2$; quindi $p_{R_\pi}(x) = (x+1)^2$. In questo caso c'è il solo autovalore $\lambda = -1$ con molteplità $m_a(-1) = 2$. Anche questa volta $R_\pi - \lambda \mathbb{I}_2 = 0$ dunque $\mathrm{Aut}(-1) = \mathbb{R}^2$ e $m_g(-1) = 2$.

Se cambiamo campo e lavoriamo su $\mathbb{C}$ anzichè su $\mathbb{R}$, la matrice R_α definisce un endomorfismo denotato ancora L_{R_α} di $\mathbb{C}^2$. Questa volta se α non è un multipo intero di π, l'equazione (9.5) ha due soluzioni complesse coniugate: ${}_1\lambda_2 = \cos\alpha \pm \sqrt{-\sin^2\alpha} = \cos\alpha \pm i \sin\alpha$. Quindi abbiamo due autovalori e due autospazi dell'endomorfismo complesso L_{R_α}.

9.4 Diagonalizzazione

Definizione 9.4.1. 1. Un endomorfismo $f : V \to V$ con $\dim V = n$, è detto **diagonalizzabile** o **semplice** se esiste una base $\mathcal{B} = (v_1, \dots, v_n)$ di V formata da autovettori di f. Se $\lambda_1, \dots, \lambda_n$ sono i loro autovalori, **non necessariamente distinti**, si ha $f(v_i) = \lambda_i v$ per ogni $i = 1, \dots, n$. Allora

$$M_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \tag{9.6}$$

   è una matrice diagonale, che contiene sulla diagonale principale gli autovalori $\lambda_1, \dots, \lambda_n$. In altre parole, esiste una base rispetto a cui la matrice di f è diagonale. Una tale base è detta **diagonalizzante**.

2. Una matrice quadrata A è **diagonalizzabile** se lo è L_A, cioè se A è simile a una matrice diagonale, ovvero esistono una matrice invertibile S e una matrice diagonale D tali che $D = S^{-1}AS$.

Osserviamo esplicitamente che in tal caso sulla diagonale principale di D compaiono gli autovalori di A, eventualmente nulli e con ripetizioni.

Teorema 9.4.2. *Sia $f : V \to V$ un endomorfismo, $\dim V = n$, e sia λ un autovalore di f. Allora $m_a(\lambda) \geq m_g(\lambda)$.*

Dimostrazione. Fissiamo una base $(v_1, \dots, v_m)$ di $\mathrm{Aut}(\lambda)$, dove $m = m_g(\lambda)$. La prolunghiamo a una base $\mathcal{B}$ di V: $(v_1, \dots, v_m, \dots, v_n)$. Sia $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$. Si ha $f(v_1) = \lambda v_1, \dots, f(v_m) = \lambda v_m$, perciò A è una matrice a blocchi della forma:

$$A = \left(\begin{array}{ccc} \lambda \mathbb{I}_m & | & B \\ --- & | & --- \\ 0 & | & A' \end{array} \right)$$

dove B, A' sono matrici opportune, e A' è quadrata di ordine $n - m$. Allora il polinomio caratteristico di f è

$$p_f(x) = \mid A - x\mathbb{I}_n \mid = \det \begin{pmatrix} (\lambda - x)\mathbb{I}_m & | & B \\ --- & | & --- \\ 0 & | & A' - x\mathbb{I}_{n-m} \end{pmatrix} = (\lambda - x)^m \mid A' - x\mathbb{I}_{n-m} \mid = (\lambda - x)^m q(x).$$

Dunque $m_g(\lambda) = m \leq m_a(\lambda)$, e vale la disuguaglianza stretta se e solo se $q(\lambda) = 0$. $\square$

Esempi 9.4.3. Sia $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Si ha $p_A(x) = \det \begin{pmatrix} -x & 1 \\ 0 & -x \end{pmatrix} = x^2$. Dunque l'unico autovalore è $\lambda = 0$ con $m_a(0) = 2$. Inoltre, si ha $\mathrm{Aut}(0) = \ker L_A = \mathrm{Span}(e_1)$. Quindi $1 = m_g(0) < m_a(0) = 2$.

Osserviamo che A non è diagonalizzabile, altrimenti sarebbe simile alla matrice nulla, perché 0 è il suo unico autovalore; ma la classe di similitudine della matrice nulla chiaramente non contiene matrici non nulle.

Osservazione 40. Se $f : V \to V$ con $\dim V = n$ ha n **autovalori distinti** $\lambda_1, \dots, \lambda_n$, per il Teorema 9.1.5 n autovettori corrispondenti a tali autovalori sono linearmente indipendenti, perciò formano una base di V, dunque f **è diagonalizzabile**. In questo caso il polinomio caratteristico è della forma $p_f(x) = (-1)^n(x - \lambda_1) \cdots (x - \lambda_n) = (\lambda_1 - x) \cdots (\lambda_n - x)$, e tutte le molteplicità algebriche risultano uguali a 1. Per il Teorema 9.4.2 anche le molteplicità geometriche sono dunque uguali a 1. Gli autospazi hanno tutti dimensione 1. La matrice A risulta simile alla matrice

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \tag{9.7}$$

e a ogni matrice diagonale in cui gli elementi della diagonale principale compaiono in un ordine diverso.

Per esempio, da quanto visto nell'esempio 9.3.2, segue che sul campo complesso $\mathbb{C}$, la matrice di una rotazione

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

è simile alla matrice

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

dove $\lambda_1 = \cos\alpha + i\sin\alpha = e^{i\alpha}$, $\lambda_2 = \cos\alpha - i\sin\alpha = e^{-i\alpha}$ (abbiamo usato la notazione esponenziale per i numeri complessi).

Osservazione 41. Sia $f : V \to V$ un endomorfismo con $p_f(x) = (\lambda - x)^n$; allora si ha $m_a(\lambda) = n$ e f ha il solo autovalore λ. Dunque f è diagonalizzabile se e solo se V ha una base di autovettori tutti di autovalore λ, rispetto a cui la matrice di f è $\lambda\mathbb{I}_n$; inoltre $\mathrm{Aut}(\lambda) = V$ e $m_g(\lambda) = n$.

In questo caso $f = \lambda\,\mathrm{id}_V$, ovvero $f(v) = \lambda v$ per ogni vettore $v \in V$. Un tale endomorfismo è detto **omotetia di rapporto** λ. È rappresentato da $\lambda\mathbb{I}_n$ rispetto a qualunque base, e in effetti se una matrice è simile a $\lambda\mathbb{I}_n$ coincide con $\lambda\mathbb{I}_n$, infatti $S \cdot (\lambda\mathbb{I}_n) \cdot S^{-1} = \lambda(S \cdot \mathbb{I}_n \cdot S^{-1}) = \lambda\mathbb{I}_n$.

Enunciamo finalmente un teorema che ci dà condizioni necessarie e sufficienti per la diagonalizzabilità di un endomorfismo.

Teorema 9.4.4. *Sia $f : V \to V$ un endomorfismo di un $\mathbb{K}$-spazio vettoriale V di dimensione finita n. Le seguenti condizioni sono equivalenti:*

1) f è diagonalizzabile;

2) il polinomio caratteristico $p_f(x)$ si fattorizza in $\mathbb{K}[x]$ come prodotto di n fattori lineari **non necessariamente distinti***:*

$$p_f(x) = (\lambda_1 - x)(\lambda_2 - x)\cdots(\lambda_n - x),$$

e per ogni autovalore λ_i si ha

$$m_a(\lambda_i) = m_g(\lambda_i);$$

3) detti $\mu_1, \dots, \mu_k$ gli autovalori distinti di f, si ha

$$V = \mathrm{Aut}(\mu_1) \oplus \mathrm{Aut}(\mu_2) \oplus \cdots \oplus \mathrm{Aut}(\mu_k). \tag{9.8}$$

Dimostrazione. 1) $\to$ 2) Sia f diagonalizzabile, allora esiste una base $\mathcal{B}$ di autovettori di f e $M_{\mathcal{B}}^{\mathcal{B}}(f)$ ha la forma (9.7). Quindi $p_f(x) = \mid M_{\mathcal{B}}(f) - x\mathbb{I}_n \mid = (\lambda_1 - x)(\lambda_2 - x)\cdots(\lambda_n - x)$. Questo prova la prima asserzione.

Consideriamo $m_a(\lambda_i)$: è uguale al numero di volte che λ_i compare sulla diagonale principale di $M_{\mathcal{B}}^{\mathcal{B}}(f)$, ovvero al numero di vettori della base $\mathcal{B}$ che hanno λ_i come autovettore. D'altra parte $m_g(\lambda_i) = \dim \mathrm{Aut}(\lambda_i) = n - \mathrm{rg}(M_{\mathcal{B}}^{\mathcal{B}}(f) - \lambda_i \mathbb{I}_n)$, che quindi coincide con $m_a(\lambda_i)$.

2) $\to$ 3) Mettiamo in evidenza nel polinomio caratteristico le molteplicità algebriche dei vari autovettori: $p_f(x) = (\mu_1 - x)^{m_1} \cdots (\mu_k - x)^{m_k}$, dove m_i è la molteplicità algebrica di μ_i. Osserviamo che, visto che $p_f(x)$ ha grado n, si ha $m_1 + m_2 + \cdots + m_k = n$. Per ipotesi, per ogni i si ha $m_i = m_a(\mu_i) = m_g(\mu_i)$. Perciò

$$\dim \mathrm{Aut}(\mu_1) + \cdots + \dim \mathrm{Aut}(\mu_k) = n = \dim V.$$

Ci rimane da dimostrare che la somma $\mathrm{Aut}(\mu_1) + \mathrm{Aut}(\mu_2) + \cdots + \mathrm{Aut}(\mu_k)$ è diretta. Ricordiamo che una somma di sottospazi $W_1, \ldots, W_r$ è diretta se ogni elemento di $W_1 + \cdots + W_r$ ha un'unica espressione $w_1 + \cdots + w_r$ come somma di vettori presi nei vari sottospazi considerati (vedere Sezione 5.4). Sia dunque $w_1 + \cdots + w_k = w_1' + \cdots + w_k'$, con $w_i, w_i' \in \mathrm{Aut}(\mu_i)$. Allora

$$(w_1 - w_1') + (w_2 - w_2') + \cdots + (w_k - w_k') = 0, \tag{9.9}$$

e ogni addendo $w_i - w_i' \in \mathrm{Aut}(\mu_i)$, perché un autospazio è sottospazio vettoriale. Alcuni addendi possono risultare nulli; se ce ne fossero di non nulli, la (9.9) sarebbe una combinazione lineare nulla non banale (con coefficienti uguali a 1) di autovettori relativi ad autovalori distinti: ma questo contraddice il Teorema 9.1.5, quindi in realtà gli addendi sono tutti nulli e per ogni i si ha $w_i = w_i'$, cioè l'unicità dell'espressione, e la somma è diretta.

3) $\to$ 1) Se vale la (9.8), si può costruire una base di V facendo l'unione di basi dei sottospazi $\mathrm{Aut}(\mu_1), \ldots, \mathrm{Aut}(\mu_k)$ (Proposizione **??**), si ha così una base di autovettori, quindi l'endomorfismo f è diagonalizzabile. □

Esempi 9.4.5. Sia $A \in M_3(\mathbb{R})$ la seguente matrice:

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

Ci chiediamo se A è diagonalizzabile. Consideriamo il suo polinomio caratteristico

$$p_A(x) = \det \begin{pmatrix} -x & -1 & 1 \\ -3 & -2-x & 3 \\ -2 & -2 & 3-x \end{pmatrix} = -x^3 + x^2 + x - 1.$$

Osserviamo che è possibile fattorizzare il polinomio caratteristico nel prodotto di fattori lineari; infatti

$$-x^3 + x^2 + x - 1 = -x^2(x-1) + (x-1) = -(x^2-1)(x-1) = -(x-1)^2(x+1).$$

Gli autovalori sono dunque

- $\lambda_1 = 1$ con $m_a(1) = 2$;
- $\lambda_2 = -1$ con $m_a(-1) = 1$ e quindi anche $m_g(-1) = 1$.

Dunque A è diagonalizzabile se e solo se $m_g(1) = 2$. Ricordiamo che $\mathrm{Aut}(1) = \ker(L_{A-\mathbb{I}_3})$, dove

$$A - \mathbb{I}_3 = \begin{pmatrix} -1 & -1 & 1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{pmatrix}.$$

Tale matrice ha rango 1, perché le sue righe sono tutte e tre proporzionali, dunque il nucleo ha dimensione 2, quindi $m_g(1) = 2$ e A è diagonalizzabile, e precisamente è simile alla matrice diagonale

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

Vogliamo ora trovare una base $\mathcal{B} = (v_1, v_2, v_3)$ diagonalizzante per A, quindi tale che $M^{\mathcal{B}}_{\mathcal{B}}(L_A) = D$. Dobbiamo prendere $v_1, v_2 \in \mathrm{Aut}(1)$, $v_3 \in \mathrm{Aut}(-1)$.

Il primo autospazio $\mathrm{Aut}(1)$ è lo spazio delle soluzioni dell'equazione omogenea $x_1 + x_2 - x_3 = 0$, corrispondente a una riga di $A - \mathbb{I}_3$. Osserviamo che x_2, x_3 sono entrambe variabili libere, dunque una generica soluzione è del tipo $(-s_2 + s_3, s_2, s_3)$. Due soluzioni linearmente indipendenti si ottengono ponendo $(s_2, s_3) = (0, 1)$ oppure $(1, 0)$. Otteniamo così

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

Il secondo autospazio è dato da $\mathrm{Aut}(-1) = \ker(L_{A+\mathbb{I}_3})$, dove

$$A + \mathbb{I}_3 = \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ -2 & -2 & 4 \end{pmatrix}.$$

Operiamo con trasformazioni elementari per trasformarla in una matrice a scala, verifichiamo così che ha rango 2 e troviamo una base per lo spazio delle soluzioni:

$$\mathrm{Aut}(-1) = \mathrm{Span}\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}.$$

Abbiamo dunque trovato una base che diagonalizzante $\mathcal{B} = \{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}\} = \{v_1, v_2, v_3\}$.

Vogliamo infine determinare una matrice invertibile S e la sua inversa S^{-1}, tali che $D = S \cdot A \cdot S^{-1}$. La matrice S^{-1} si può scrivere immediatamente perché $S^{-1} = M^{\mathcal{B}}_{\mathcal{E}}(\mathrm{id}_{\mathbb{R}^3})$ e dunque $S^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 3 \\ 1 & 0 & 2 \end{pmatrix}$. Calcoliamo S come inversa di S^{-1} usando l'algoritmo di Gauss; si otiiene

$$S = -\frac{1}{2}\begin{pmatrix} 2 & 2 & -4 \\ 3 & 1 & -3 \\ -1 & -1 & 1 \end{pmatrix}.$$

Se vogliamo fare la verifica, osserviamo che $D = SAS^{-1}$ se e solo se $DS = SA$.

Se prendiamo invece come base diagonalizzante $\mathcal{B}' = (v_3, v_1, v_2)$ troveremo anzichè D la matrice diagonale $D' = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Attenzione! Per trasformazioni elementari sulle righe di una matrice A, il polinomio caratteristico cambia, quindi non è possibile trasformare preliminarmente A in una matrice a gradini per semplificare il calcolo del polinomio caratteristico.

9.5 Esempi II

Esempi 9.5.1. Consideriamo l'applicazione $\rho : \mathbb{R}^2 \to \mathbb{R}^2$ definita da

$$\rho\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Chiaramente ρ è un endomorfismo, perché è del tipo L_A, con $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. È la **riflessione rispetto all'asse** x che è detta asse della riflessione. L'endomorfismo ρ è già espresso nella forma L_A con A diagonale, e vediamo che ha due autovalori reali distinti $1, -1$. L'autospazio $\mathrm{Aut}(1)$ è l'asse della riflessione $\mathrm{Span}(e_1)$, mentre $\mathrm{Aut}(-1) = \mathrm{Span}(e_2)$.

Ora ruotiamo di $\alpha/2$ l'asse della riflessione, e otteniamo l'applicazione ρ_r, riflessione di asse r. Per scrivere delle equazioni per ρ_r, osserviamo che, se $\mathcal{B} = \{v_1, v_2\}$ è la base ottenuta dalla base canonica con la rotazione di $\alpha/2$, la matrice di ρ_r rispetto a $\mathcal{B}$ è $M^{\mathcal{B}}_{\mathcal{B}}(\rho_r) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Allora

$$M^{\mathcal{E}}_{\mathcal{E}}(\rho_r) = M^{\mathcal{B}}_{\mathcal{E}}(\mathrm{id}_{\mathbb{R}^2}) M^{\mathcal{B}}_{\mathcal{B}}(\rho_r) M^{\mathcal{E}}_{\mathcal{B}}(\mathrm{id}_{\mathbb{R}^2}),$$

dove

$$M^{\mathcal{B}}_{\mathcal{E}}(\mathrm{id}_{\mathbb{R}^2}) = \begin{pmatrix} \cos\frac{\alpha}{2} & -\sin\frac{\alpha}{2} \\ \sin\frac{\alpha}{2} & \cos\frac{\alpha}{2} \end{pmatrix}$$

è la matrice della rotazione di $\frac{\alpha}{2}$ e

$$M^{\mathcal{E}}_{\mathcal{B}}(\mathrm{id}_{\mathbb{R}^2}) = \begin{pmatrix} \cos\frac{\alpha}{2} & \sin\frac{\alpha}{2} \\ -\sin\frac{\alpha}{2} & \cos\frac{\alpha}{2} \end{pmatrix}$$

è la matrice della trasformazione inversa, che è la rotazione di $-\frac{\alpha}{2}$. Eseguendo il prodotto otteniamo

$$M^{\mathcal{E}}_{\mathcal{E}}(\rho_r) = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} =: S_\alpha.$$

Abbiamo trovato S_α, che è la matrice, rispetto alla base canonica, della riflessione di asse r che forma l'angolo di $\frac{\alpha}{2}$ con l'asse x. È una matrice a traccia nulla simile ad A. Il polinomio caratteristico è

$$p_{S_\alpha}(x) = \det \begin{pmatrix} \cos\alpha - x & \sin\alpha \\ \sin\alpha & -\cos\alpha - x \end{pmatrix} = x^2 - 1 = (x-1)(x+1).$$

Dunque per ogni α vi sono i due autovalori $1, -1$ con molteplicità algebrica e geometrica uguali a 1. Questo era aspettato perché S_α è simile ad A,e quindi hanno lo stesso polinomio caratteristico. Dunque tutti gli endomorfismi ρ_r sono diagonalizzabili. I due autospazi sono: per $\lambda = 1$, la retta r di equazione $(\cos\alpha - 1)x + \sin\alpha y = 0$, per $\lambda = -1$ la retta di equazione $(\cos\alpha + 1)x + \sin\alpha y = 0$.

Esempi 9.5.2. Sia $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$; il suo polinomio caratteristico è $P_A(x) = \det \begin{pmatrix} -x & 1 \\ -1 & -x \end{pmatrix} = x^2 + 1$. La matrice A non è diagonalizzabile su $\mathbb{R}$, perché $x^2 + 1$ non ha radici reali, ma lo è su $\mathbb{C}$. Infatti in $\mathbb{C}[x]$ si ha $x^2 + 1 = (x+i)(x-i)$, quindi sul campo complesso A è simile alla matrice diagonale $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$. L'autospazio $\mathrm{Aut}(i) = \ker L \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix}$, che ha equazione

$x_1 + ix_2 = 0$. Una soluzione è $\begin{pmatrix} i \\ -1 \end{pmatrix}$. Analogamente $\mathrm{Aut}(-i) = \mathrm{Span}(\begin{pmatrix} i \\ 1 \end{pmatrix}$. Una base diagonalizzante è $\mathcal{B} = \left\{\left(\begin{pmatrix} i \\ -1 \end{pmatrix}, \begin{pmatrix} i \\ 1 \end{pmatrix} \right)\right\}$. Si ha $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = S^{-1}AS$, con $S = \begin{pmatrix} i & i \\ -1 & 1 \end{pmatrix}$. Inoltre det$S = 2i$, dunque $S^{-1} = \frac{1}{2i}\begin{pmatrix} 1 & -i \\ i & i \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} i & 1 \\ i & -1 \end{pmatrix}$.

Esempi 9.5.3. Sia J la matrice triangolare

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

Il suo polinomio caratteristico è $p_J(x) = (2-x)^3$, prodotto di fattori lineari in $\mathbb{R}[x]$ e in $\mathbb{Q}[x]$. L'unico autovalore è $\lambda = 2$ con $m_a(2) = 3$. L'autospazio $\mathrm{Aut}(2)$ ha dimensione $3 - \mathrm{rg}(J - 2\mathbb{I}_3)$. Abbiamo

$$J - 2\mathbb{I}_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

che ha rango 2, perciò $m_g(2) = 1$, quindi J non è diagonalizzabile su nessun campo.

9.6 Teorema fondamentale dell'Algebra

Abbiamo visto esempi di endomorfismi diagonalizzabili, ed esempi di endomorfismi che non lo sono, perché non vale una o l'altra delle due condizioni necessarie e sufficienti del Teorema 9.4.4, punto 2. Negli esempi 9.3.2 e 9.5.2 sul campo reale, il polinomio caratteristico non si fattorizza in fattori lineari in $\mathbb{R}[x]$. Nell'esempio 9.5.3, pur fattorizzandosi $p_J(x)$ in fattori lineari su qualunque campo, non vale la condizione di uguaglianza delle due molteplicità algebrica e geometrica.

La condizione che il polinomio caratteristico $p_f(x)$ si fattorizzi in $\mathbb{K}[x]$ come prodotto di n fattori lineari si esprime dicendo che $p_f(x)$ ha tutte le sue radici in $\mathbb{K}$.

A tale proposito, enunciamo senza dimostrazione l'importante teorema:

Teorema 9.6.1 (Teorema fondamentale dell'Algebra)**.** *Sia $p(x) \in \mathbb{C}[x]$ un polinomio di grado $n \geq 1$ a coefficienti complessi. Allora $p(x)$ ha almeno una radice λ in $\mathbb{C}$.*

Dunque $p(x)$ è divisibile per $x - \lambda$, cioè $p(x) = (x-\lambda)p_1(x)$, con $p_1(x)$ di grado $n-1$; il polinomio $p_1(x)$ si trova eseguendo l'algoritmo di divisione tra polinomi. Ora, se $n - 1 \geq 1$, si può ripetere il procedimento, fino a concludere che $p(x)$ è prodotto di n fattori lineari in $\mathbb{C}[x]$, dunque $p(x)$ ha tutte le sue n radici in $\mathbb{C}$, purchè si conti ogni radice tante volte quant'è la sua molteplicità algebrica.

Tuttavia, anche se tutte le radici di qualunque polinomio $p(x)$ esistono in $\mathbb{C}$, non è sempre possibile calcolarle. Per polinomi di grado $n = 2$, esiste la ben nota formula risolutiva che esprime le soluzioni dell'equazione $p(x) = 0$ in funzione dei coefficienti di p, mediante un'estrazione di radice quadrata. Per gradi $n = 3, 4$ pure esistono della formule risolutive "per radicali", dette **formule di Cardano**, che esprimono le radici di un polinomio $p(x)$ in funzione dei coefficienti ricorrendo all'estrazione di radici terze e quarte. Invece, per gradi $n \geq 5$, è stato dimostrato che **non esiste una formula risolutiva per radicali.** Questo risultato, parzialmente dimostrato da Paolo Ruffini nei primi anni dell'800, fu ottenuto dal matematico norvegese Niels Abel nel 1826.

Vi sono tuttavia delle formule risolutive che fanno uso di altre funzioni, anzichè di radicali; nella pratica, però, dovendo risolvere equazioni specifiche, si ricorre a metodi numerici che forniscono le soluzioni approssimate, con il grado di approssimazione desiderato.

Se si lavora invece su campi $\mathbb{Z}_p$, che hanno un numero finito di elementi, le soluzioni di un'equazione si possono cercare, con un numero finito di tentativi, sostituendo via via all'incognita i vari elementi del campo.