Geometria

18 giugno 2015

Capitolo 1

Diagonalizzazione

1.1 Autovalori e autovettori

Definizione 1.1.1. *Sia V uno spazio vettoriale su un campo K, e T un endomorfismo su V. Si dice* autovalore *di un endomorfismo $T \in \text{End}(V)$ uno scalare $\lambda \in K$ per il quale esiste un vettore $v \in V \setminus \{0\}$ tale che $T(v) = \lambda v$. Un tale vettore v si dice* autovettore *di T associato all'autovalore λ.*

Ad esempio, l'applicazione

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

individua una rotazione di un vettore di $\mathbb{R}^2$. Se α è diverso da multipli interi di π, allora A non ammette alcun autovalore; se invece $\alpha = k\pi$ la matrice A individua una rotazione di 0 oppure π, cioè una dilatazione di v, eventualmente con un cambiamento di verso del vettore, quindi qualunque vettore di $\mathbb{R}^2$ è un autovettore di A.

Teorema 1.1.2. Sia V uno spazio vettoriale di dimensione finita. Il determinante della matrice associata a un endomorfismo $T \in \text{End}(V)$ non dipende dalla scelta della base per rappresentarla.

Dimostrazione. Sia $\dim V = m$. Fissata una medesima base $\mathcal{B} = \{e_i\}_{i=1}^m$ di V sia in arrivo che in partenza, sia A_T la matrice associata a T nella base data, definita quindi come $(A_T)_{ij} = T(e_i)_j$. Scegliendo una base differente $\mathcal{B} = \{e_i'\}_{i=1}^m$ per V, essa si può sempre ottenere applicando una matrice C alla precedente base $\mathcal{B}$, quindi

$$e_j' = \sum_{i=1}^m C_{ij} e_i,$$

e detta A_T' la matrice che rappresenta T nella nuova base, si ha $A_T' = C^{-1} A_T C$. Il determinante di A_T, sotto questa nuova base, è dunque $\det A_T' = \det(C^{-1} A_T C) = \det C^{-1} \det A_T \det C = (\det C)^{-1} \det A_T \det C = \det A_T$, per il teorema di Binet e la commutatività di K. □

Possiamo quindi definire il *determinante di un endomorfismo*, sapendo che sarà lo stesso qualunque matrice si scelga per rappresentarlo.

Teorema 1.1.3. Sia V uno spazio vettoriale di dimensione finita su un campo K, $T \in \text{End}(V)$ e $\lambda \in K$. Le seguenti affermazioni si equivalgono:

1. λ è un autovalore di T;

2. l'operatore lineare $T - \lambda I$ non è iniettivo;

3. $\det(T - \lambda I) = 0$.

Dimostrazione. Se λ è un autovalore, allore esiste un vettore $v \in V$ non nullo per cui $T(v) = \lambda v = (\lambda I)(v)$ quindi $T(v) - (\lambda I)(v) = (T - \lambda I)(v) = 0_V$. Ma ciò significa che $\ker(T - \lambda I)$ non contiene soltanto 0_V, dunque non è iniettivo. Viceversa, se $\ker(T - \lambda I) \neq \{0_V\}$ vuol dire che esiste un $v \in V$ per cui $(T - \lambda I)(v) = 0_V$, cioè risalendo i passaggi precedenti $T(v) = \lambda v$.

Se $T - \lambda I$ non è iniettivo, non può essere di conseguenza invertibile, quindi il suo determinante deve essere nullo. Viceversa, se il determinante è nullo allora non è invertibile, vale a dire $T - \lambda I$ non è iniettivo oppure non è suriettivo. In spazi di dimensione finita, però, se un endomorfismo è iniettivo è automaticamente suriettivo (in virtù del primo teorema dell'isomorfismo **??**), dunque le due affermazioni sono equivalenti: T non è né iniettivo né suriettivo. □

Definizione 1.1.4. *Si definisce* poliniomio caratteristico *dell'applicazione* $T \in \operatorname{End}(V)$ *il polinomio* $\chi_T(x) = \det(T - xI)$.

È possibile definire il polinomio caratteristico di un endomorfismo, equivalentemente, come $\det(A - xI)$ dive A rappresenta T in qualche base $\mathcal{B}$. Ma se $\tilde{\mathcal{B}}$ è un'altra base e L la matrice di cambiamento di base tra le due, allora la matrice che rappresenta T in $\tilde{\mathcal{B}}$ è $\tilde{A} = L^{-1}AL$. Il polinomio caratteristico di questa nuova matrice è

$$\chi_{\tilde{A}}(x) = \det(\tilde{A} - xI) = \det(L^{-1}AL - xI) = \det(L^{-1}AL - L^{-1}xIL) = \det(L^{-1})\det(A - xI)\det L = \chi_A(x)$$

quindi matrici simili hanno il medesimo polinomio caratteristico. La definizione data tramite l'endomorfismo è quindi ben posta, in quanto le matrici rappresentanti un endomorfismo rispetto a basi differenti sono tutte simili.[1]

Le radici di $\chi_T(x)$, per l'ultimo punto del teorema precedente, sono chiaramente gli autovalori di T. Il grado di questo polinomio inoltre equivale alla dimensione dello spazio vettoriale V. Si può alternativamente definire il polinomio caratteristico come $\chi_T(x) = \det(xI - T)$: gli autovalori trovati come radici non variano, perché per le proprietà del determinante vale $\det(xI - T) = (-1)^m \det(T - xI)$, dove $m = \dim V$, quindi le radici sono le stesse per entrambe le definizioni.

Rifacendosi al primo esempio

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

il polinomio caratteristico di A è

$$\begin{aligned} \chi_A(x) &= \begin{vmatrix} \cos\alpha - \lambda & -\sin\alpha \\ \sin\alpha & \cos\alpha - \lambda \end{vmatrix} = \\ &= (\cos\alpha - \lambda)^2 + \sin^2\alpha = \\ &= \lambda^2 - 2\cos\alpha\lambda + 1, \end{aligned}$$

il cui discriminante è $-4\sin^2\alpha$, che quindi non è negativo solo se $\alpha = k\pi$ $(k \in \mathbb{Z})$. Soltanto per questi valori A ammette dunque autovalori.

Se v è un autovettore associato a λ, anche i suoi multipli per uno scalare sono a loro volta autovettori: infatti se $T(v) = \lambda v$ e $k \in K$ segue per la linearità di T che $T(kv) = kT(v) = k\lambda v$, cioè qualsiasi multiplo kv è un autovettore di T associato a λ.

Definizione 1.1.5. *L'insieme V_λ degli autovettori associati ad un unico autovalore λ è detto* autospazio *di T:*

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

L'autospazio V_λ non si riduce mai allo zero, poiché gli autovettori sono per definizione non nulli, ed è anche un sottospazio vettoriale di V. Se $v, w \in V_\lambda$ e $h, k \in K$, si ha

$$T(hv + kw) = hT(v) + kT(w) = h\lambda v + k\lambda w = \lambda(hv + kw),$$

quindi $hv + kw$ è ancora un autovettore associato a λ, cioè appartiene a V_λ. La somma di due autovettori di T associati a due autovalori distinti però *non* è ancora necessariamente un autovettore di T.

[1] Questo ovviamente vale purché le basi di partenza e di arrivo coincidono!

Teorema 1.1.6. Siano V uno spazio vettoriale di dimensione finita su un campo K, T un endomorfismo in V. Siano $v_1, v_2, \dots, v_k$ autovettori di T associati rispettivamente agli autovalori $\lambda_1, \lambda_2 \dots, \lambda_k$. Se questi autovalori sono tutti distinti, allora i $v_1, v_2, \dots, v_k$ sono linearmente indipendenti.

Dimostrazione. Dimostriamolo per induzione rispetto a k. Se $k = 1$, un elemento soltanto $v_1 \in V$ è linearmente indipendente perché essendo un autovettore non è mai nullo, quindi il teorema è subito dimostrato. Sia ora $k > 1$. Consideriamo una combinazione lineare

$$a_1v_1 + a_2v_2 + \cdots + a_kv_k = 0, \tag{a}$$

e si dimostra che $a_1 = \cdots = a_k = 0_K$. Moltiplicando la (a) per λ_1 si ottiene

$$a_1\lambda_1v_1 + a_2\lambda_1v_2 + \cdots + a_k\lambda_1v_k = 0, \tag{b}$$

e applicando T sempre alla (a) si ottiene un'altra combinazione

$$a_1\lambda_1v_1 + \cdots + a_k\lambda_kv_k = 0. \tag{c}$$

Sottraendo la (b) a quest'ultima risulta

$$a_2(\lambda_2 - \lambda_1)v_2 + a_3(\lambda_3 - \lambda_1)v_3 + \cdots + a_k(\lambda_k - \lambda_1)v_k = 0.$$

Dato che $\lambda_i \neq \lambda_1$ per ogni $i > 1$, non può che essere $a_2 = a_3 = \cdots = a_k = 0_K$, a cui segue nella (a) che $a_1v_1 = 0_V$, da cui $a_1 = 0$. Poiché dunque $a_1 = a_2 = \cdots = a_k = 0_K$, l'insieme $\{v_1, v_2 \dots, v_k\}$ è linearmente indipendente. □

Definizione 1.1.7. *Sia $T \in \mathrm{End}(V)$, con $\dim V < +\infty$, e λ un autovalore di T. Si chiama* molteplicità geometrica *di λ, e si indica con γ_λ, la dimensione dell'autospazio V_λ; si chiama* molteplicità algebrica*, e si indica con α_λ, la sua molteplicità come radice del polinomio caratteristico di T.*

Teorema 1.1.8. Sia λ un autovalore di $T \in \mathrm{End}(V)$, con V di dimensione finita. Vale la relazione

$$1 \leq \gamma_\lambda \leq \alpha_\lambda \leq \dim V.$$

Dimostrazione. Se λ è un autovalore, esiste un autospazio ad esso associato non vuoto, che ha quindi una dimensione non nulla; inoltre esiste una radice di $\chi_T(x)$, ed essa ha quindi una molteplicità non nulla. Allora $\alpha_\lambda, \gamma_\lambda \geq 1$.

L'autospazio è un sottospazio vettoriale di V, quindi $\gamma_\lambda \dim V_\lambda \leq \dim V$, e il grado di $\chi_T(x)$ (che equivale alla dimensione di V) non può essere superato dalla somma delle molteplicità delle radici per il teorema **??** dunque $\alpha_\lambda \leq \dim V$. Si può individuare una base dell'autospazio V_λ composta da γ_λ elementi, che sono autovettori associati a λ. Nell'autospazio V_λ l'endomorfismo T si comporta come un multiplo dell'identità, precisamente λI_n (dove n è la dimensione dell'autospazio, cioè è γ_λ), dato che sono tutti autovettori con autovalore λ. Nella base scelta, dunque, T è rappresentato da

$$A = \begin{bmatrix} \lambda I_n & 0 \\ 0 & M \end{bmatrix}$$

con $n = \gamma_\lambda$ come già visto, e M è una matrice qualunque (è il blocco corrispondente all'azione di T su V meno l'autospazio V_λ). Il suo polinomio caratteristico è $\chi_T(x) = \det(A - xI) = \det(\lambda I_n - xI_n)\det(M - xI_{\dim V - n}) = (\lambda - x)^n g(x)$ dove $g(x) = \det(M - xI_{\dim V-n})$ è un generico polinomio di grado $\dim V - n$. Allora $\lambda - x$ divide $\chi_T(x)$ almeno γ_λ volte, vale a dire che la molteplicità della radice λ non è minore di γ_λ, cioè $\alpha_\lambda \geq \gamma_\lambda$. □

1.2 Diagonalizzabilità

Definizione 1.2.1. *Sia $T \in \mathrm{End}(V)$, di dimensione finita. T si dice diagonalizzabile se esiste una base di V costituita soltanto da autovettori di V.*

In questa base di autovettori, la matrice che rappresenta T è diagonale, e sappiamo che matrici di un medesimo endomorfismo associate a differenti basi sono simili. Dunque equivalentemente si può dire che una matrice $M \in \mathrm{Mat}(n, K)$ è diagonalizzabile se $\exists P \in GL(n, K)$ tale per cui il coniugio $P^{-1}MP$ sia una matrice diagonale. Se chiamiamo $D = P^{-1}PM$ la matrice diagonale e $\mathcal{B} = \{b_1, \dots, b_n\}$ la base in cui rappresenta l'endomorfismo, allora P è la matrice per cambiare la base da quella precedente di M a quella in forma diagonale, cioè $P = (b_1|\cdots|b_n)$. Allora se $P^{-1}MP = D$ si ha $MP = PD$, ossia

$$(Mb_1|\cdots|Mb_n) = (b_1|\cdots|b_n)D = (d_1b_1|\cdots|d_nb_n)$$

cioè $Mb_i = d_ib_i$: i vettori b_i della base (che di conseguenza non possono essere nulli) sono quindi gli autovettori di M.

Teorema 1.2.2. Siano $T \in \mathrm{End}(V)$ e $\lambda_1, \dots, \lambda_k$ i suoi autovalori distinti, di molteplicità geometrica $\gamma_1, \dots, \gamma_k$ e algebrica $\alpha_1, \dots, \alpha_k$. Le seguenti affermazioni sono equivalenti:

- T è diagonalizzabile;
- il polinomio caratteristico è $\chi_T(x) = (\lambda_1 - x)^{\alpha_1} \dots (\lambda_k - x)^{\alpha_k}$, con $\alpha_i = \gamma_i$ per ogni $i \in \{1, \dots, k\}$;
- $\sum_{i=1}^{k} \gamma_i = \dim V$;

Dimostrazione. Se T è diagonalizzabile allora esiste una base $\mathcal{B} = \{e_i\}_{i=1}^{k}$ di autovettori dell'endomorfismo. Possiamo riordinarli in modo che i primi γ_1 siano gli autovettori relativi a λ_1, quelli da $\gamma_1 + 1$ a $\gamma_1 + \gamma_2$ siano gli autovettori relativi a λ_2 e così via, fino a esaurirli, ossia

$$\begin{gathered} e_1, \dots, e_{\gamma_1} \in V_1, \\ e_{\gamma_1+1}, \dots, e_{\gamma_1+\gamma_2} \in V_2, \\ \vdots \\ e_{\gamma_1+\cdots+\gamma_{k-1}}, \dots, e_{\gamma_1+\cdots+\gamma_{k-1}+\gamma_k} \in V_k \end{gathered}$$

dove V_j è l'autospazio dell'autovalore λ_j. In tale base, T è rappresentato da una matrice D diagonale, poiché $T(e_i)$ moltiplica e_i per il rispettivo autovalore: in ogni autospazio V_i l'endomorfismo agisce dunque come $\lambda_i I_{\gamma_i}$, ossia come un multiplo dell'identità, quindi

$$D = \begin{bmatrix} \lambda_1 I_{\gamma_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{\gamma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k I_{\gamma_k} \end{bmatrix}. \tag{1.2.1}$$

scrivendo D come matrice a blocchi. Il polinomio caratteristico di D, quindi anche di T, è ovviamente

$$\chi_T(x) = (\lambda_1 - x)^{\gamma_1}(\lambda_2 - x)^{\gamma_2} \dots (\lambda_k - x)^{\gamma_k} = \prod_{i=1}^{k} (\lambda_i - x)^{\gamma_i} \tag{1.2.2}$$

Sappiamo però che ogni λ_i è radice di χ_T con molteplicità α_i, dunque $\forall i \in \{1, \dots, k\}$ risulta $(\lambda_i - x)^{\alpha_i} | \chi_T$. Allora χ_T ammette una fattorizzazione unica in termini di polinomi irriducibili tra cui figurano certamente i fattori $(\lambda_i - x)^{\alpha_i}$, perciò

$$\chi_T(x) = (\lambda_1 - x)^{\alpha_1}(\lambda_2 - x)^{\alpha_2} \dots (\lambda_k - x)^{\alpha_k} g(x) = g(x) \prod_{i=1}^{k} (\lambda_i - x)^{\alpha_i}$$

dove g è il prodotto di altri polinomi irriducibili $K[x]$ diversi ovviamente dai $\lambda_i - x$. Per l'unicità della fattorizzazione cioè i fattori irriducibili delle due "versioni" devono essere uguali: deve allora verificarsi che $\alpha_i = \gamma_i \ \forall i \in \{1, \ldots, k\}$. Di conseguenza $\deg g = 0$, vale a dire $g = 1$. Quindi

$$\chi_T(x) = \prod_{i=1}^{k} (\lambda_i - x)^{\alpha_i}.$$

Poiché la somma delle molteplicità algebriche, dato che i fattori $(\lambda_i - x)$ sono gli unici presenti in χ_T, dà il grado di χ_T, segue immediatamente se $\gamma_i = \alpha_i$ che

$$\dim V = \sum_{i=1}^{k} \alpha_i = \sum_{i=1}^{k} \gamma_i. \tag{1.2.3}$$

In ogni autospazio V_i troviamo γ_i vettori linearmente indipendenti (tutti autovettori di T), che formano una base del sottospazio. Riunendo le basi di tutti gli autospazi, otteniamo un insieme $\mathcal{I}$ linearmente indipendente in V, in base al teorema **??** in quanto gli autospazi sono linearmente indipendenti per il teorema 1.1.6. Per il punto precedente, $\sum_{i=1}^{k} \gamma_i = \dim V$ quindi il numero di vettori in $\mathcal{I}$ è proprio $\dim V$. Per il teorema **??** segue dunque che $\mathcal{I}$ è una base di V, e ciò prova che T è diagonalizzabile. □

Segue immediatamente il seguente corollario.

Corollario 1.2.3. Sia $T \in \mathrm{End}(V)$, con $\dim V = m < +\infty$. Se T ha m autovalori distinti, allora T è diagonalizzabile.

Dimostrazione. La dimostrazione è immediata, dato che esistono $m = \dim V$ autovettori linearmente indipendenti (per il teorema 1.1.6, dato che sono associati ad autovalori distinti), dunque essi formano una base di V, perciò T è diagonalizzabile. □

Teorema 1.2.4. Sia $T \in \mathrm{End}(V)$ con V di dimensione finita. Dati $\lambda_1, \ldots, \lambda_k$ autovalori distinti con i corrispondenti autospazi $V_{\lambda_1}, \ldots, V_{\lambda_k}$ si ha che T è diagonalizzabile se V può essere scritto come somma diretta degli autospazi, ossia $V = \bigoplus_{i=1}^{k} V_{\lambda i}$.

Dimostrazione. La dimostrazione è analoga alla seconda parte di quella svolta nel teorema **??**. □

1.3 Polinomio minimo

Sia V uno spazio vettoriale sul campo K e $T \in \mathrm{End}(V)$. Diciamo che un polinomio $f \in K[x]$ *annulla* T se $f(T) = 0$ (l'endomorfismo nullo). Considerando $T \in \mathrm{End}(V)$ ho che la dimensione della matrice associata, se la dimensione di V è m ed è finita, è m^2, ora consideriamo $f(T)$ come il polinomio che annulla T, si ha quindi $f(T) = 0$, quindi $T^0 = I, T = T^1, T^2 = T \circ T, \ldots, T^k = T \circ T^{k-1}$, possiamo riscrivere con queste considerazioni prendendo la matrice associata a V rispetto a L:

$$f(T) = a_{m^2} T^{m^2} + a_{m^2-1} T^{m^2-1} + \cdots + a_0 I,$$

ora le varie potenze presenti in T sono $m^2 + 1$ quindi il sistema deve essere linearmente dipendente. Esistono allora opportuni a_i per cui il sistema ammette una soluzione. Concludiamo che esiste sempre un polinomio, al massimo di grado m^2 che annulla l'operatore T.

Chiamiamo I_T, con $T \in \mathrm{End}(V)$ e V spazio di dimensione finita e sul campo K, l'insieme

$$I_T = \{p \in K[x] \colon p(T) = 0\}$$

che non è mai uguale al solo $0 \in \mathrm{End}(V)$. Si ha che per $p, q \in I_T$ la loro somma sta ancora in I_T, perché $(p + q)(T) = p(T) + q(T) = 0 + 0 = 0$, e analogamente $(\lambda p)(T) = \lambda p(T) = \lambda 0 = 0$, dunque I_T è un ideale. Essendo $K[x]$ un dominio a ideali principali, quindi, I_T ammette un (unico) generatore monico.

Definizione 1.3.1. *Sia $T \in \mathrm{End}(V)$, con V spazio di dimensione finita e sul campo K, si definisce* polinomio minimo *di T, e si indica com $m_T(x)$, il generatore monico dell'ideale I_T dei polinomi che annullano T.*

Poiché $T^n \in \mathrm{End}(V)$ per ogni $n \in \mathbb{N}_0$, ogni polinomio $f \in K[x]$ è tale che $f(T) \in \mathrm{End}(V)$. In virtù dell'isomorfismo tra matrici quadrate ed endomorfismi, se T è rappresentato da A allora l'endomorfismo $f(T)$ è rappresentato da $f(A)$.

È inoltre facile vedere che due matrici simili hanno lo stesso polinomio minimo: infatti se $C = B^{-1}AB$, allora

$$\begin{aligned} C^0 &= I = B^{-1}IB \\ C &= B^{-1}AB \\ C^2 &= B^{-1}ABB^{-1}AB = B^{-1}A^2B \\ &\dots \\ C^n &= B^{-1}A^nB \end{aligned}$$

come si verifica facilmente per induzione. Si possono raggruppare allora tutti i termini B^{-1} e B, per cui per qualsiasi polinomio si ha $f(C) = B^{-1}f(A)B$.

Teorema 1.3.2. Sia $T \in \mathrm{End}(V)$, con V spazio vettoriale di dimensione finita sul campo K. Il polinomio minimo e il polinomio caratteristico di T hanno le stesse radici.

Dimostrazione. Sia $\lambda \in K$ una radice del polinomio minimo, ossia $m_T(\lambda) = 0$. Per il teorema di Ruffini **??** allora $(x-\lambda)|m_T$, dunque possiamo scrivere il polinomio minimo come $m_T(x) = (x-\lambda)q(x)$ per un certo $q \in K[x]$. Inoltre $\deg q < \deg m_T$, perciò $q \notin I_T$ non essendo un multiplo di m_T: di conseguenza $q \neq 0$. Esiste allora $v \in V$ tale che $q(T)(v) \neq 0$: sia $w = q(T)(v)$. Poiché per come è definito il polinomio minimo $m_T(T) = 0$, si ha dunque

$$0 = m_T(T)(v) = \big[(T - \lambda I)q(T)\big](v) = (T - \lambda I)(w)$$

cioè w è un autovettore di T con autovalore λ: ma allora $\chi_T(\lambda) = 0$.

Viceversa, sia ora $\lambda \in K$ tale che $\chi_T(\lambda) = 0$. Allora λ è un autovalore di T, dunque esiste un $v \in V$ per il quale $T(v) = \lambda v$, di conseguenza $m_T(T)(v) = m_T(\lambda)(v)$.[2] Dunque $0 = m_T(T)(v) = m_T(\lambda)(v)$. Poiché $v \neq 0$, deve necessariamente essere $m_T(\lambda) = 0$, quindi λ è una radice del polinomio minimo. □

Prendiamo ad esempio un endomorfismo T di $\mathbb{R}^3$ che sia la proiezione nel sottospazio $W = \{(x_1, x_2, x_3)^T \in \mathbb{R}^3 : 2x_1 + 3x_2 - x_3 = 0\}$. Rispetto alla base canonica, la matrice che lo rappresenta è

$$A = \frac{1}{14}\begin{pmatrix} 10 & -6 & 2 \\ -6 & 5 & 3 \\ 2 & 3 & 13 \end{pmatrix}$$

e il suo polinomio caratteristico è $\chi_T(x) = \det(A - xI_3) = -x^3 + 2x^2 - x = -x(x-1)^2$. Allo stesso tempo, poiché è una proiezione, sappiamo che $T^2 - T = 0$, ossia $T^2 - T = 0$. Il polinomio $x^2 - x$ dunque annulla T, ed è anche il suo polinomio minimo: esso ha 0 e 1 come radici, che sono anche quelle di χ_T, seppur con molteplicità diversa.

1.4 Sottospazi invarianti

Definizione 1.4.1. *Sia V uno spazio vettoriale sul campo K, e W un suo sottospazio. Dato $T \in \mathrm{End}(V)$, diremo che W è* T-invariante *se $T(W) \subseteq W$, ossia se $\forall w \in W$ si ha $T(w) \in W$.*

[2] Ciò vale per un generico $f \in K[x]$: ad esempio, se $f(x) = ax^2 + bx + c$, allora $f(T)(v) = aT^2(v) + bT(v) + cI(v) = aT(\lambda v) + b\lambda v + cv = a\lambda^2 v + b\lambda v + cv = f(\lambda)(v)$.

La proprietà principale di un sottospazio W che sia T-invariante è che è possibile restringere l'applicazione in tale insieme, ossia è possibile definire $T|_W : W \to W$. L'intero spazio V e $\{0\}$ sono, banalmente, sottospazi invarianti di qualsiasi applicazione lineare.

Ecco alcuni esempi di sottospazi invarianti.

- Per alcune applicazioni, i sottospazi banali sono gli unici sottospazi invarianti: un facile esempio è una rotazione in $\mathbb{R}^2$ di un angolo diverso da $k\pi$, per $k \in \mathbb{Z}$.

- Per ogni spazio V e $T \in \operatorname{End}(V)$, $\operatorname{Ker} T$ e $\operatorname{Im} T$ sono T-invarianti. Se infatti $v \in \operatorname{Ker} T$, allora $T(v) = 0$ e $0 \in \operatorname{Ker} T$, come in tutti i sottospazio vettoriali. Analogamente $T(v) \in \operatorname{Im} T$, banalmente, qualsiasi sia $v \in V$, dunque anche per $v \in \operatorname{Im} T$.

- Un autospazio V_λ associato a un autovalore λ di un $T \in \operatorname{End}(V)$ è T-invariante, poiché per ogni $v \in V_\lambda$ si ha $T(v) = \lambda v \in V_\lambda$.

- Se $T, S \in \operatorname{End}(V)$ commutano, se $a \in \operatorname{Ker} T$ allora $T\big(S(a)\big) = S\big(T(a)\big) = S(0) = 0$ quindi $S(a) \in \operatorname{Ker} T$, cioè $\operatorname{Ker} T$ è S-invariante. Vale anche per $\operatorname{Im} T$?

Sia $W \leq V$ sottospazio dello spazio vettoriale V sul campo K. Dato $T \in \operatorname{End}(V)$, posto $v \in V$ definiamo l'insieme

$$S_T(v, W) = \{g \in K[x] : g(T)(v) \in W\}$$

ossia, fissati $v \in V$ e il sottospazio W, l'insieme dei polinomi di $K[x]$ tali che, valutati in T, portano v in W. Il polinomio minimo appartiene a questo insieme: risulta infatti $m_T(T)(v) = 0(v) = 0 \in W$, per ogni $v \in V$ e $W \leq V$.

Lemma 1.4.2. Sia V uno spazio vettoriale sul campo K e $T \in \operatorname{End}(V)$. Se $W \leq V$ è T-invariante, allora è anche $g(T)$-invariante per ogni $g \in K[x]$, ossia

$$S_T(v, W) = \{g \in K[x] : g(T)(v) \in W\}$$

è un ideale di $K[x]$ per ogni $v \in V$.

Dimostrazione. Se W è T-invariante, allora per ogni $w \in W$ si ha $T(w) \in W$. Prendiamo un generico polinomio di secondo grado $g(x) = ax^2 + bx + c$ e, valutato in T, applichiamo l'endomorfismo che ne risulta a w:

$$g(T)(w) = aT^2(w) + bT(w) + cw. \tag{1.4.1}$$

Chiaramente $cw \in W$, e dato che $T(w) \in W$ allora anche $T^2(w) = T\big(T(w)\big) \in W$, dunque $g(T)(w) \in W$. Lo stesso ragionamento si applica a polinomi di grado qualunque, poiché si ha $T^n(w) \in W$ per ogni $n \in N$. Allora W è $g(T)$-invariante.

Mostriamo quindi che $S_T(v, W)$ è un ideale. Siano $f, g \in S_T(v, W)$ e $h \in K[x]$: risulta, per $v \in V$,

$$[(f + g)(T)](v) = [f(T) + g(T)](v) = f(T)(v) + g(T)(v) \tag{1.4.2}$$

che appartiene a W, poiché esso è un sottospazio e $f(T)(v), g(T)(v) \in W$ dato che $f, g \in S_T(v, W)$, dunque anche $f + g$ è nell'insieme. Inoltre

$$[(hf)(T)](v) = h(T)[f(T)(v)],$$

ma $f(T)(v) \in W$, e dato che come mostrato prima W è $p(T)$-invariante per ogni $p \in K[x]$, lo è anche per $h(T)$, perciò $h(T)[f(T)(v)] \in W$, cioè $hf \in S_T(v, W)$. Ciò prova che $S_T(v, W)$ è un ideale di $K[x]$. □

Siano V uno spazio vettoriale sul campo K, $T \in \operatorname{End}(V)$, e W un sottospazio T-invariante di V. Dato l'insieme $S_T(\alpha, W)$ come definito in precedenza, chiamiamo polinomio T-*conducente*, di α in W, il generatore monico di $S_T(\alpha, W)$.

Lemma 1.4.3. Sia $T \in \mathrm{End}(V)$ con V spazio vettoriale, di dimensione finita, sul campo K. Se il suo polinomio minimo è della forma $m_T(x) = (x-\lambda_1)^{r_1} \dots (x-\lambda_k)^{r_k}$, con $\lambda_i \in K$, e se $W \leq V$ (con $W \neq V$) è T-invariante, allora $\exists v \in V \setminus W$ tale che $(T-\lambda I)(v) \in W$ per qualche autovalore λ di T.

Dimostrazione. Sia $\beta \in V \setminus W$ e g il polinomio T-conducente di β in W: ciò implica che $g(T)(\beta) \in W$. Nell'ideale $(g) \subset K[x]$ si trova anche il polinomio minimo, quindi $g|m_T$. Deve risultare $g \neq 1$, poichè altrimenti $g(T)(\beta) = \beta \notin W$: ma g è il polinomio T-conducente di β in W quindi per definizione deve essere $g(T)(\beta) \in W$. Quindi $\deg g > 1$: per l'ipotesi sulla fattorizzazione del polinomio minimo, si ha

$$g(x) = (x-\lambda_1)^{b_1} \cdots (x-\lambda_i)^{b_i},$$

con $b_i \leq r_i$ e almeno un b_i maggiore di 1. Poniamo $b_j > 1$ per un $j \in \{1, \dots, k\}$, corrispondente a λ_j: per il teorema di Ruffini **??** possiamo riscrivere g come $g(x) = (x-\lambda_j)h(x)$, per un certo $h \in K[x]$. Sia ora $\alpha := h(T)(\beta)$: risulta

$$(T-\lambda_j I)(\alpha) = (T-\lambda_j I)h(T)(\beta) = g(T)(\beta)$$

e $g(T)(\beta)$ appartiene a W per come è definito g. Ora, $h \neq 0$ perché è il prodotto di fattori $(x-\lambda_i)^{b_i}$, certamente non nulli, e poiché divide g non può appartenere a S_T, perciò $h(T)$ non "porta" β in W. Dunque $\alpha = h(T)(\beta) \notin W$, e certamente $\alpha \in V$. Dal teorema 1.3.2 inoltre sappiamo che le radici λ_i sono anche radici del polinomio caratteristico di T, dunque sono i suoi autovalori. Questo prova che esiste $\alpha \in V \setminus W$ e un autovalore λ_j per i quali $(T-\lambda_j I)(\alpha) \in W$ come nella tesi. □

Teorema 1.4.4. Sia V uno spazio vettoriale di dimensione finita sul campo K e sia $T \in \mathrm{End}(V)$. T è diagonalizzabile se e solo se $m_T(x) = (x-\lambda_1)\cdots(x-\lambda_k)$, con distinti $\lambda_i \in K$.

Dimostrazione. Sia T diagonalizzabile. Presi gli autovalori distinti λ_i di T, consideriamo gli endomorfismi $(T-\lambda_1 I), \dots, (T-\lambda_k I)$; sia inoltre α un autovettore di T. Chiaramente $(T-\lambda_j I)\alpha = 0$, se α è autovettore associato all'autovalore λ', secondo il teorema 1.1.3. Se componiamo tutti gli operatori $T-\lambda_i I$, se essi commutassero troveremmo ancora 0 applicandoli ad α: basterebbe scambiare l'ordine fino ad avere $T-\lambda' I$ applicato a α, che fa zero, da cui tutto il prodotto è zero. Verifichiamo questo fatto. Dati a, b autovalori di T, risulta

$$(T-aI)(T-bI) = T^2 - aT - bT - abI = T^2 - bI - aI - baI = (T-bI)(T-aI) \tag{1.4.3}$$

dato che a, b sono in un campo. Allora

$$(T-\lambda_1 I)\cdots(T-\lambda_k I)(\alpha) = (T-\lambda_1)\cdots(T-\lambda_k)(T-\lambda' I)(\alpha) = (T-\lambda_1)\cdots(T-\lambda_k)(0) = 0. \tag{1.4.4}$$

Definiamo dunque $p(x) = (T-\lambda_1)\cdots(T-\lambda_k)(T-\lambda_j)$. Per la 1.4.4 allora $p(T)\alpha = 0$ per ogni autovettore α di T. Ma T è diagonalizzabile, perciò esiste una base di autovettori $\{\alpha_i\}_{i=1}^m$ di V (con $m = \dim V$), per ognuno dei quali si ha

$$p(T)(\alpha_j) = \prod_{i=1}^{k}(T-\lambda_i I)(\alpha_j) = (T-\lambda_1 I)\cdots(T-\lambda_k I)(T-\lambda_j I)(\alpha_j) = 0. \tag{1.4.5}$$

Ogni $v \in V$ si può scrivere come $\sum_{i=1}^m c_i\alpha_i$, per opportuni coefficienti $c_i \in K$, ma allora per linearità

$$p(T)(v) = p(T)\Big(\sum_{i=1}^m c_i\alpha_i\Big) = \sum_{i=1}^m c_i p(T)(\alpha_i) = 0 \tag{1.4.6}$$

dunque $p(T)$ è l'operatore nullo, vale a dire $p \in (m_T)$. Allora $m_T|p$: ma m_T ha come radici tutti gli autovalori $\lambda_1, \dots, \lambda_k$ per il teorema 1.3.2, dunque $p = m_T$ e ciò prova la tesi.

Sia ora $m_T = \prod_{i=1}^k (x-\lambda_i)$ e $W \leq V$ il sottospazio generato dagli autovettori di T: la sua base è composta da autovettori di T. Possiamo assumere $W \neq V$, perché altrimenti V ammetterebbe una base di autovettori di T, che sarebbe dunque subito diagonalizzabile (la dimostrazione finirebbe qui). Essendo la somma di autospazi di T, W è T-invariante. Se $W \neq V$, allora esiste un $\alpha \in V \setminus W$ tale per cui $\beta := (T - \lambda_j I)(\alpha) \in W$: la sua esistenza è garantita dal lemma 1.4.3. Vogliamo mostrare che un tale α non può esistere senza contraddire le ipotesi, e di conseguenza non può essere che $W \neq V$ (altrimenti un tale α esisterebbe sempre). Sempre per l'invarianza di W, si ha $f(T)(\beta) \in W$ per ogni $f \in K[x]$, dal lemma 1.4.2. Definiamo

$$q_j(x) = \prod_{\substack{i=1 \\ i \neq j}} (x - \lambda_i),$$

tale che $m_T(x) = (x - \lambda_j) q_j(x)$. Consideriamo $q_j(x) - q_j(\lambda_j)$: questo polinomio ha evidentemente λ_j come radice, dunque è divisibile per $x - \lambda_j$, cioè $q_j(x) - q_j(\lambda_j) = h(x)(x - \lambda_j)$ per un certo $h \in K[x]$. Valutiamo questo polinomio in T, e applichiamo l'operatore che ne risulta a α:

$$[q_j(T) - q_j(\lambda_j)](\alpha) = [h(T)(T - \lambda_j I)](\alpha) = h(T)(\beta). \tag{1.4.7}$$

Sappiamo anche che $m_T(T) = 0$ e $m_T(x) = (x - \lambda_j) q_j(x)$, perciò

$$0 = m_T(T)(\alpha) = [(T - \lambda_j I) q_j(T)](\alpha) \tag{1.4.8}$$

Ora, se $q_j(T)(\alpha) = 0$ automaticamente appartiene a W; altrimenti, si ha comunque $(T - \lambda_j I) q_j(T)(\alpha) = 0$ quindi per il teorema 1.1.3 $q_j(T)(\alpha)$ è un autovettore di T, quindi anche in questo caso appartiene a W. Poiché dalla 1.4.7 risulta $q_j(\lambda_j)\alpha = q_j(T)(\alpha) - h(T)(\beta)$, e gli addendi del secondo membro appartengono a W, si ha $q_j(\lambda_j)\alpha \in W$.[3] Ma $\alpha \notin W$, per come è stato scelto, di conseguenza deve essere $q_j(\lambda_j) = 0$ (lo zero di K, poiché è uno scalare). Ciò però significa che m_T ha λ_j come radice *doppia*, che contraddice l'ipotesi su m_T avente solo radici semplici. Per questo non può esistere $\alpha \in V \setminus W$, cioè $V \setminus W = \varnothing$, dato che un tale elemento si può sempre trovare se W non è l'intero V: allora $W = V$, cioè T è diagonalizzabile. □

1.5 Triangolarizzazione

Definizione 1.5.1. *Sia $T \in \mathrm{End}(V)$, con V spazio vettoriale sul campo K, T si dice* triangolabile *(o triangolarizzabile) se esiste una base $\{e_i\}_{i=1}^n \in V$ tale che la matrice associata a T rispetto a quella base è in forma triangolare (v. definizione* **??***).*

Il teorema di Cayley-Hamilton mostra un importante risultato: per ogni endomorfismo T di uno spazio V di dimensione finita,

$$\chi_T(T) = 0.$$

La dimostrazione generale è però difficile; qui vediamo il teorema solo per endomorfismi triangolabili.

Teorema 1.5.2 (di Cayley-Hamilton)**.** Sia $T \in \mathrm{End}(V)$, con V spazio vettoriale di dimensione finita su un campo K. Se T è triangolabile, allora $\chi_T(T) = 0$.

Dimostrazione. Se T è triangolabile allora esiste una base $\{e_i\}_{i=1}^m$ di V, tale che la matrice A che corrisponde a T in tale base è della forma

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ 0 & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{mm} \end{pmatrix}.$$

Definiamo dunque m spazi $W_1 = \langle\{e_1\}\rangle$, $W_2 = \langle\{e_1, e_2\}\rangle$, …, $W_{m-1} = \langle\{e_1, \dots, e_{m-1}\}\rangle$, $W_m = V$. Per il fatto che T è triangolare su questa base, tutti questi W_i sono T-invarianti.[4] Dimostriamo il

[3] Riguardo alla notazione: $q_j(\lambda_j)$ è un polinomio di $K[x]$ valutato in $\lambda_j \in K$, dunque è uno scalare di K: perciò usiamo la notazione moltiplicativa $q_j(\lambda_j)\alpha$ anziché trattarlo come un operatore e scrivere $q_j(\lambda_j)(\alpha)$ come per i termini restanti.

[4] Questo fatto si può facilmente vedere moltiplicando la matrice A per i vettori di base dei W_i.

teorema per induzione su m. Per $m = 1$ si ha $\chi_T(x) = A - xI = (a_{11}) - xI$, perciò valutato in A risulta $\chi_T(A) = (a_{11}) - (a_{11}) = 0$. Assumiamo ora che sia vero $m - 1$. Poichè A è triangolare, lo è anche $A - xI$, quindi il suo determinante si calcola facilmente come

$$\chi_T(x) = (a_{11} - x)(a_{22} - x) \cdots (a_{mm} - x).$$

Consideriamo il sottospazio W_{m-1} e restringiamo l'applicazione T a tale sottospazio: sia $\overline{T} = T|_{W_{m-1}}$, con $\overline{T}\colon W_{m-1} \to W_{m-1}$. La matrice che rappresenta $\overline{T}$ nella base $\{e_i\}_{i=1}^{m-1}$ di W_{m-1} è triangolare, perciò $\chi_{\overline{T}}(x) = (a_{11} - x) \cdots (a_{m-1,m-1} - x)$. Allora

$$\chi_T(x) = \chi_{\overline{T}}(x)(a_{mm} - x),$$

e se valutiamo tale polinomio in T e lo applichiamo agli elementi e_i della base, risulta

$$\begin{aligned} \chi_T(T)(e_i) &= [\chi_{\overline{T}}(T)(a_{mm}I - T)](e_i) = \\ &= \chi_{\overline{T}}(T)[a_{mm}e_i - T(e_i)] = \\ &= a_{mm}[\chi_{\overline{T}}(T)](e_i) - \chi_{\overline{T}}(T)T(e_i). \end{aligned}$$

Ma $\chi_{\overline{T}}(T)(v) = 0$ per ogni $v \in W_1, \ldots, W_{m-1}$ per l'ipotesi di induzione, e sia e_i che $T(e_i)$ appartengono a uno di essi, per la T-invarianza dei sottospazi $W_1, \ldots, W_{m-1}$: allora

$$\chi_T(T)(e_i) = 0 \tag{a}$$

per ogni $i \in \{1, \ldots, m-1\}$. Per l'ultimo vettore e_m vale invece $T(e_m) = a_{1m}e_1 + \cdots + a_{mm}e_m$, da cui:

$$\begin{aligned} \chi_T(T)(e_m) &= \chi_{\overline{T}}(T)(T - a_{mm}I)(e_m) = \\ &= \chi_{\overline{T}}(T)\Big(\sum_{i=1}^{m} a_{im}e_i - a_{mm}e_m\Big) = \\ &= \sum_{i=1}^{m-1} a_{im}\chi_{\overline{T}}(T)(e_i) = 0 \end{aligned}$$

in virtù della (a). Dunque $\chi_T(T)(e_i) = 0$ per ogni $i \in \{1, \ldots, m\}$, ossia per ogni vettore della base: ma allora $\chi_T(T)(v) = 0$ per ogni $v \in V$, ossia $\chi_T(T)$ è l'endomorfismo nullo. Ciò prova che $\chi_T(T) = 0$ per ogni T triangolabile. □

Corollario 1.5.3. Dato uno spazio V di dimensione finita e $T \in \mathrm{End}(V)$, se T è triangolabile allora $m_T | \chi_T$.

Dimostrazione. Se T è triangolabile, per il teorema precedente, $\chi_T(T) = 0$ dunque il polinomio caratteristico appartiene all'ideale (m_T), vale a dire $m_T | \chi_T$. □