GEOMETRIA 1
terza parte

Gilberto Bini - Cristina Turrini

2017/2018

index

Siano V e W spazi vettoriali f.g. sul campo $\mathbb{K}$ con $n = \dim(V), m = \dim(W)$ e sia $f : V \to W$ un'applicazione lineare.

PROBLEMA - Esistono basi $\mathcal{B}$ di V e $\mathcal{C}$ di W tali che la matrice rappresentativa di f in queste basi sia "particolamente semplice", ossia di elementi
$a_{ij} = 1$ se $i = j = 1, \ldots k$, $a_{ij} = 0$ in tutti gli altri casi, ovvero sia della forma

$$A = \begin{pmatrix} I_k & O \\ O & O \end{pmatrix}.$$

Anzitutto, perché ciò sia possibile è necessario che sia $k = \dim(Im(f))$ (il rango della matrice rappresentativa coincide con la dimensione dell'immagine dell'applicazione).

Se $k = \dim(Im(f))$ la risposta è SÌ.

Per il teorema nullità + rango, si ha $\dim(\ker(f)) = n - k$.
Sia inoltre $\{\underline{v}_{k+1}, \ldots, \underline{v}_n\}$ una base di $\ker(f)$ (se $n > k$) e la si completi ad una base $\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_k, \underline{v}_{k+1}, \ldots, \underline{v}_n\}$ di V.

Allora $\{f(\underline{v}_1),\dots,f(\underline{v}_k)\}$ è una base di $Im(f)$. Si completi tale base ad una base $\mathcal{C}=\{f(\underline{v}_1),\dots,f(\underline{v}_k),\underline{w}_{k+1},\dots\underline{w}_m\}$ di W.

La matrice rappresentativa di f rispetto a tali basi ha la forma richiesta.

Ad esempio, per $f:\mathbb{R}^3\to\mathbb{R}^2$ definita da

$$f\begin{pmatrix} x\\ y\\ z\end{pmatrix}=\begin{pmatrix} 2x-y\\ z\end{pmatrix},$$

si può prendere $\mathcal{B}=\{\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix},\begin{pmatrix}1\\2\\0\end{pmatrix}\}$ e $\mathcal{C}=\{\begin{pmatrix}-1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\}$.

index

Endomorfismi diagonalizzabili

Siano ora V f.g. con $\dim(V) = n$ e $f : V \to V$ un endomorfismo.

PROBLEMA - Esiste una base $\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_n\}$ di V tale che la matrice A rappresentativa di f rispetto a tale base (*sia in dominio che in codominio*) sia diagonale

$$A = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \quad ?$$

Se la risposta è affermativa l'endomorfismo f viene detto *diagonalizzabile* e la base $\mathcal{B}$ viene detta *diagonalizzante*.

OSSERVAZIONE - I vettori di una base diagonalizzante verificano:

$$f(\underline{v}_j) = \lambda_j \underline{v}_j, \quad \forall j = 1, \ldots, n$$

Un vettore non nullo $\underline{v} \in V, \quad \underline{v} \neq \underline{0}$ viene detto *autovettore* per f se esiste $\lambda \in \mathbb{K}$ tale che $f(\underline{v}) = \lambda \underline{v}$.

Lo scalare λ (che è univocamente associato a $\underline{v}$) viene detto *autovalore* relativo all'autovettore $\underline{v}$.

Una immediata conseguenza delle considerazioni fatte sopra è il

TEOREMA - Un endomorfismo f è diagonalizzabile se e solo se esiste una base di V interamente costituita da autovettori di f.

Sia λ è un autovalore di f. Consideriamo l'insieme $A_\lambda(f)$ degli autovettori di f relativi a λ.

L' insieme

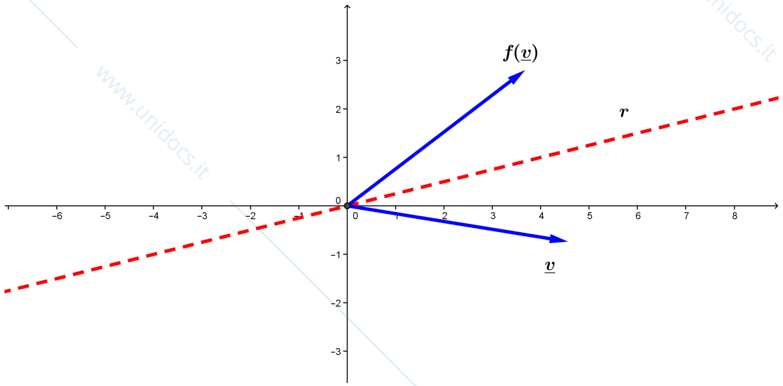
$$V_\lambda(f) = A_\lambda(f) \cup \{\underline{0}\}$$

è un sottospazio di V (verificarlo) detto *autospazio* relativo all'autovalore λ.

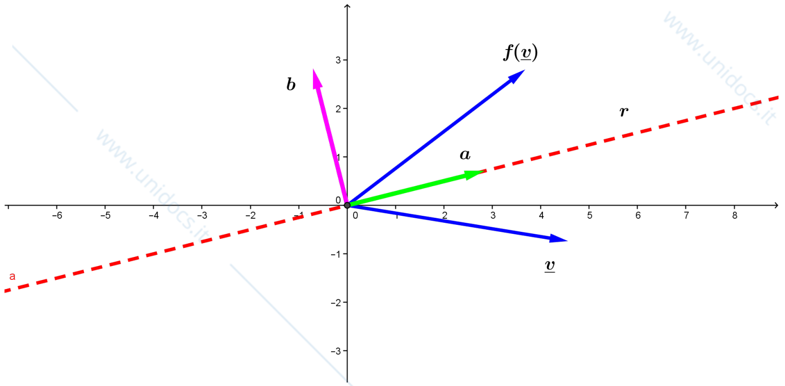
Qualche esempio nel caso di $Vect_O(\mathbb{R}^2)$

OSSERVAZIONE - Un autovettore, nel caso dei vettori geometrici, è un vettore trasformato in un vettore parallelo.

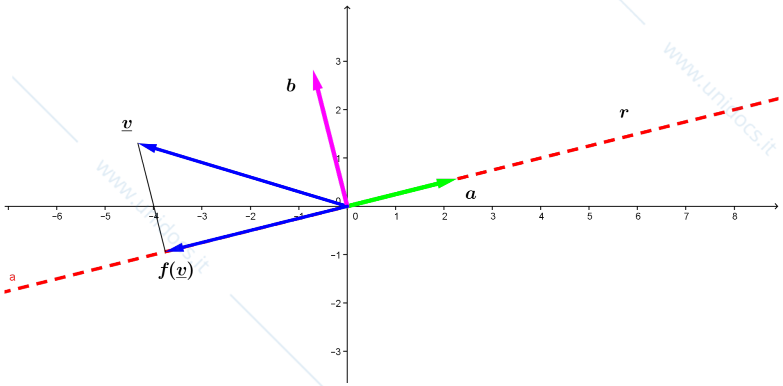
Riflessione rispetto alla retta r.



Nella riflessione rispetto alla retta r gli autovettori sono i vettori di r (con autovalore 1) e i vettori ortogonali a r (con autovalore -1).

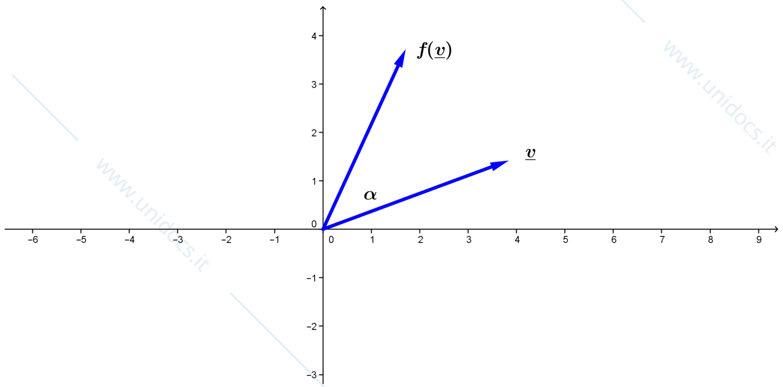


Proiezione ortogonale sulla retta r.



Nella proiezione ortogonale sulla retta r gli autovettori sono i vettori di r (con autovalore 1) e i vettori ortogonali a r (con autovalore 0).

Rotazione di un angolo α attorno O.



Se α non è congruo a 0 o a π (mod. 2π), la rotazione di un angolo α attorno O non ammette autovettori.

Matrici diagonalizzabili

Le nozioni di diagonalizzabilità, autovalori, autovettori introdotte per gli endomorfismi si trasferiscono alle matrici quadrate:

- una matrice quadrata $n \times n$ A è diagonalizzabile se lo è l'endomorfismo $L_A : \mathbb{K}^n \to \mathbb{K}^n$;
- un autovettore di A è un vettore non nullo $\underline{x} \in \underline{K}^n$ tale che $A \cdot \underline{x} = \lambda \underline{x}$;
- lo scalare λ viene detto autovalore della matrice A.

Ricordando la nozione di matrici simili introdotta nella seconda parte di queste note, si ha (verificarlo):

Una matrice quadrate $n \times n$ A è diagonalizzabile se e solo se è simile a una matrice diagonale.

Sia f un endomorfismo di V f.g., con $\dim(V) = n$.

Problema: ricerca (se esiste) di una base di autovettori.

TEOREMA - Se $\lambda_1, \ldots, \lambda_k$ sono autovettori di f distinti tra loro, e $\underline{v}_1, \ldots, \underline{v}_k$ sono autovettori relativi a $\lambda_1, \ldots, \lambda_k$ (rispett.), allora i vettori $\underline{v}_1, \ldots, \underline{v}_k$ sono linearmente indipendenti.

Dimostrazione
Per induzione su k.

- Se $k = 1$, $\underline{v}_1$ è l.i. in quanto non nullo.
- Supponendo vero il risultato nel caso di $k-1$ autovalori, dimostriamolo nel caso di k autovalori.

Supponiamo che sia

$$(*) \qquad a_1\underline{v}_1 + a_2\underline{v}_2 + \cdots + a_k\underline{v}_k = \underline{0} \qquad a_1, \ldots, a_k \in \mathbb{K}.$$

Applicando f a entrambe i membri di $(*)$ si ottiene

$$(\circ) \qquad a_1\lambda_1\underline{v}_1 + a_2\lambda_2\underline{v}_2 + \cdots + a_k\lambda_k\underline{v}_k = \underline{0}.$$

Moltiplicando entrambe i membri di $(*)$ per λ_k si ottiene

$$(\circ\circ) \qquad a_1\lambda_k\underline{v}_1 + a_2\lambda_k\underline{v}_2 + \cdots + a_k\lambda_k\underline{v}_k = \underline{0}.$$

Sottraendo membro a membro $(\circ\circ)$ da $(\circ)$ si ottiene

$$a_1(\lambda_1 - \lambda_k)\underline{v}_1 + a_2(\lambda_2 - \lambda_k)\underline{v}_2 + \cdots + a_{k-1}(\lambda_{k-1} - \lambda_k)\underline{v}_{k-1} = \underline{0}.$$

Per l'ipotesi di induzione allora deve essere:

$$a_1(\lambda_1 - \lambda_k) = a_2(\lambda_2 - \lambda_k) = \dots a_{k-1}(\lambda_{k-1} - \lambda_k) = 0,$$

e quindi, trattandosi di autovalori distinti tra loro,

$$a_1 = a_2 = \cdots = a_{k-1} = 0 \quad \Rightarrow \quad a_k = 0.$$

COROLLARIO - Se f ha $n = \dim(V)$ autovalori distinti, allora è diagonalizzabile.

index

Sia f un endomorfismo di V ($\dim(V) = n$).

OSSERVAZIONE - $\lambda \in \mathbb{K}$ è un autovalore di f se e solo se esiste $\underline{v} \in \ker(f - \lambda id_V), \underline{v} \neq \underline{0}$.

OSSERVAZIONE - Se $\lambda \in \mathbb{K}$ è un autovalore di f allora si ha

$$V_\lambda(f) = \ker(f - \lambda id_V).$$

In particolare, se $\lambda = 0$ è un autovalore per f, allora $V_0(f) = \ker(f)$.

Sia ora $\mathcal{B}$ una base di V e $A = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ la matrice rappresentativa di f rispetto alla base $\mathcal{B}$.

$\lambda \in \mathbb{K}$ è un autovalore di f se e solo se $f - \lambda id_V$ non è un isomorfismo se e solo se $A - \lambda I_n$ non ha rango massimo se e solo se $\det(A - \lambda I_n) = 0$

Il polinomio

$$P_A(t) = \det(A - tI_n) = \det \begin{pmatrix} a_{11} - t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - t & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - t \end{pmatrix}$$

viene detto *polinomio caratteristico* di A.

OSSERVAZIONE - Se $A = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ e $B = \mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(f)$ sono matrici rappresentative dello stesso endomorfismo rispetto a basi diverse, allora

$$P_A(t) = P_B(t).$$

Infatti

$$\det(B - tI) = \det(C^{-1}AC - tI) = \det(C^{-1}AC - tC^{-1}IC) =$$

$$\det(C^{-1}(A - tI)C) = \det(C^{-1})\det(A - tI)\det(C) = \det(A - tI).$$

In particolare, per $t = 0$, si ha anche $\det(B) = \det(A)$.

Per questo motivo il polinomio $\det(A - tI_n)$ viene anche detto *polinomio caratteristico di f* e denotato con $P_f(t)$ e il determinante di A viene anche detto *determinante di f* e denotato con $\det(f)$.

OSSERVAZIONE - Il polinomio caratteristico $P_f(t)$

- ha grado n,
- ha coefficiente direttore $(-1)^n$,
- ha termine noto $P_f(0) = \det(f)$,
- le sue radici in $\mathbb{K}$ sono gli autovalori di f.

OSSERVAZIONE - Se $\mathbb{K} = \mathbb{C}$, tutte le radici di $P_f(t) \in \mathbb{C}[t]$ sono in $\mathbb{K}$ e pertanto sono autovalori di f. Se invece $\mathbb{K} = \mathbb{R}$, allora solo le radici reali di $P_f(t)$ sono autovalori.

Ricerca degli autovalori e autovettori

Sia f un endomorfismo di V ($\dim(V) = n$). Per cercare autovalori e autovettori di f;

- Si considera una base $\mathcal{B}$ di V e si costruisce la matrice $A = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ rappresentativa di f rispetto alla base $\mathcal{B}$.
- Si calcola il polinomio caratteristico $P_A(t)$ e si determinano le sue radici $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ che sono gli autovalori di f.
- Per ciascuno degli autovalori λ_i si risolve il sistema lineare $(A - \lambda_i I)\underline{x} = \underline{0}$
- Le soluzioni non nulle x del sistema $(A - \lambda_i I)\underline{x} = \underline{0}$ sono le coordinate, nella base $\mathcal{B}$ degli autovettori relativi a λ_i.

Esempi

- $(n=2)$
  $\det\begin{pmatrix} a_{11}-t & a_{12} \\ a_{21} & a_{22}-t \end{pmatrix} = t^2 - (a_{11}+a_{22})t + \det(A)$
- $(n=3)$
  $\det\begin{pmatrix} a_{11}-t & a_{12} & a_{13} \\ a_{21} & a_{22}-t & a_{23} \\ a_{31} & a_{32} & a_{33}-t \end{pmatrix} = -t^3 + (a_{11}+a_{22}+a_{33})t^2 -$
  $(\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \det\begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix} + \det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix})t + \det(A).$
- In generale
  $P_A(t) = (-1)^n t^n + (-1)^{n-1}\sigma_1 t^{n-1} + \cdots + (-1)^{n-i}\sigma_i t^{n-i} + \cdots - t\sigma_{n-1} + \sigma_n$,
  ove σ_i è la somma dei minori *principali* (ossia aventi come diagonale parte della diagonale di A) di A.

  In particolare $\sigma_1 = a_{11} + a_{22} + \cdots + a_{nn}$ viene detta *traccia* di A e $\sigma_n = \det(A)$.

index

Ricordo che una radice $\alpha \in \mathbb{K}$ di un polinomio $p(t) \in \mathbb{K}[t]$ si dice avere *molteplicità* $m > 0$ se $p(t) = (t-\alpha)^m q(t)$, con $q(\alpha) \neq 0$, ovvero m è il massimo degli l tali che $(t-\alpha)^l$ sia un fattore di $p(t)$.

Abbiamo visto che un autovalore λ di f è necessariamente una radice in $\mathbb{K}$ del polinomio caratteristico $P_f(t)$ di f.

Si dice *molteplicità algebrica* $m_a(\lambda)$ dell'autovalore λ la sua molteplicità come radice del polinomio $P_f(t)$.

Se λ è un autovalore di f, l'autospazio $V_\lambda(f)$ non è lo spazio nullo. Si definisce *molteplicità geometrica* $m_g(\lambda)$ dell'autovalore λ la dimensione dell'autospazio $V_\lambda(f)$.

TEOREMA - Siano V uno spazio vettoriale f.g. sul campo $\mathbb{K}$, $f : V \to V$ un endomorfismo e $\lambda \in \mathbb{K}$ un autovalore di f. Si ha

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

Dimostrazione
La relazione $1 \leq m_g(\lambda)$ segue dal fatto che, essendo λ un autovalore, si ha $\dim(V_\lambda(f)) > 0$.

Consideriamo una base $\{\underline{v}_1, \ldots, \underline{v}_{m_g(\lambda)}\}$ di $V_\lambda(f)$ e completiamola a una base $\{\underline{v}_1, \ldots, \underline{v}_{m_g(\lambda)}, \underline{w}_{m_g(\lambda)+1} \cdots \underline{w}_n\}$ di V.

In tale base f è rappresentato da una matrice della forma

$$\begin{pmatrix} \lambda & 0 & \ldots & 0 & \star & \ldots & \star \\ 0 & \lambda & \ldots & 0 & \star & \ldots & \star \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \lambda & \star & \ldots & \star \\ 0 & 0 & \ldots & 0 & \star & \ldots & \star \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & o & \star & \ldots & \star \end{pmatrix}$$

Il polinomio caratteristico di f allora risulta

$$P_f(t) = \begin{pmatrix} \lambda - t & 0 & \dots & 0 & \star & \dots & \star \\ 0 & \lambda - t & \dots & 0 & \star & \dots & \star \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda - t & \star & \dots & \star \\ 0 & 0 & \dots & 0 & \star - t & \dots & \star \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & o & \star & \dots & \star - t \end{pmatrix} = (\lambda - t)^{m_g(\lambda)} q(t)$$

(segue iteratamente dallo sviluppo di Laplace del determinante secondo la prima colonna).

Pertanto si ha $m_g(\lambda) \leq m_a(\lambda)$.

TEOREMA - Siano V uno spazio vettoriale f.g. sul campo $\mathbb{K}$ e $f : V \to V$ un endomorfismo f è diagonalizzabile se e solo se

i) tutte le radici di $P_f(t)$ sono in $\mathbb{K}$;

ii) per ogni autovalore λ di f si ha $m_g(\lambda) = m_a(\lambda)$.

OSSERVAZIONI
1) Se $\mathbb{K} = \mathbb{C}$ la condizione i) è sempre verificata.

2) Se $m_a(\lambda) = 1$, allora $m_g(\lambda) = 1$, quindi la condizione ii) è verificata.

3) In generale, per calcolare $m_g(\lambda)$:

$$m_g(\lambda) = \dim(\ker(f - \lambda id_V)) = n - car(A - \lambda I).$$

ESEMPI

- $$A = \begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & \cos(\theta) \end{pmatrix} \qquad \mathbb{K} = \mathbb{R}.$$

  $$P_A(t) = (cos(\theta) - t)^2 + sin^2(\theta) = t^2 - 2cos(\theta)t + 1$$

  che ha discriminante $\Delta = 4(cos^2(\theta) - 1)$, quindi se $\theta \neq 0, \pi$ non vale la i).

- $$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \qquad \mathbb{K} = \mathbb{R}.$$

  $$P_A(t) = (1 - t)^3$$

  L'unica radice è $\lambda = 1$, quindi vale la i), inoltre $m_a(1) = 3$.

  $car(A - I) = 2$, quindi $m_g(1) = 1$, per cui non vale la ii).

- $$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \mathbb{K} = \mathbb{R}.$$

  $$P_A(t) = -t^2(t-2)$$

  Le radici sono $0, 2$, quindi vale la $i)$, inoltre $m_a(0) = 2, m_a(2) = 1$.

  Ovviamente $m_a(2) = 1 = m_g(2)$.
  $car(A) = 1$, quindi $m_g(0) = 3 - car(A) = 2 = m_a(0)$, per cui vale anche la $ii)$ e la matrice è diagonalizzabile.

Diagonalizzazione simultanea

TEOREMA - Siano A e B due matrici diagonalizzabili in $Mat_n(\mathbb{K})$.
Allora esiste una matrice M invertibile tale che $D_1 = M^{-1}AM$ e $D_2 = M^{-1}BM$ siano diagonali se e solo se A e B commutano, cioè $AB = BA$.

Dimostrazione - Se esiste M, si ha

$$AB = MD_1D_2M^{-1} = MD_2D_1M^{-1} = BA.$$

Supponiamo che A e B siano diagonalizzabili e che commutino. Per quest'ultima condizione, l'applicazione lineare L_B è un endomorfismo di ciascun autospazio di A (verificarlo).
Dato che L_B è diagonalizzabile, ciascun autospazio di A si scompone in autospazi di B.
Per ciascun autospazio di A prendiamo una base fatta da autovettori di L_B ristretta a tale autospazio. Si genera una base che per costruzione diagonalizza A e B simultaneamente.

index

Sia $(A, +, \cdot)$ un anello commutativo.

Si dice *ideale* di A un sottoinsieme non vuoto $I \subset A$ tale che

- $\forall i_1, i_2 \in I$ si ha $i_1 - i_2 \in I$;
- $\forall j \in I, \forall a \in A$ si ha $a \cdot j \in I$.

Ad esempio, nel caso degli interi $(\mathbb{Z}, +, \cdot)$, fissato un intero n, l'insieme $I = n\mathbb{Z}$ dei multipli di n è un ideale.

Un ideale $I \subset A$ si dice *principale* se esiste $\bar{j} \in I$ tale che, $\forall j \in I$, esiste $a \in A$ tale che sia $j = a \cdot \bar{j}$.

Un elemento $\bar{j}$ come sopra si dice *generatore* di I.

Nel caso dell'anello $\mathbb{K}[x]$ dei polinomi a coefficienti in un campo $\mathbb{K}$ tutti gli ideali sono principali.

TEOREMA - Sia $I \subset \mathbb{K}[x]$ un ideale. Esiste un $g(x) \in I$ tale che $\forall f(x) \in I$ si abbia $f(x) = q(x)g(x)$, per qualche $q(x) \in \mathbb{K}[x]$.

Dimostrazione
Se $I = \{0\}$, il risultato è ovvio. Possiamo quindi supporre $I \neq \{0\}$.

Sia $g(x)$ un polinomio di grado minimo tra i polinomi non nulli di I. Proviamo che $g(x)$ è un generatore di I.

Sia $f(x) \in I$. L'algoritmo della divisione di polinomi garantisce l'esistenza (e unicità) di due polinomi $q(x), r(x)$ (rispettivamente *quoziente* e *resto* della divisione di $f(x)$ per $g(x)$) tali che sia

- $f(x) = q(x)g(x) + r(x)$;
- o $r(x) \equiv 0$, o $\deg(r(x)) < \deg(g(x))$.

Allora $r(x) = f(x) - g(x)q(x) \in I$.

Non può allora essere $\deg(r(x)) < \deg(g(x))$ (perché $g(x)$ ha il grado minimo tra i polinomi di I), per cui è necessariamente $r(x) \equiv 0$.

Pertanto si ha $f(x) = g(x)q(x)$.

OSSERVAZIONE - Il generatore di un ideale di I non è unico, però lo è il generatore *monico*, ossia che ha 1 come coefficiente direttore.

Infatti, se g_1, g_2 fossero due diversi generatori monici di I allora $g_1 - g_2$ apparterrebbe a I ed avrebbe grado minore di quello di g_1 e g_2.

index

Sia V uno spazio vettoriale di dimensione n su $\mathbb{K}$.

Siano poi $A \in Mat_n(\mathbb{K})$ e f un endomorfismo di V.

Dato un polinomio $p(t) \in K[t]$:

$$p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_k t^k.$$

Si definisce $p(A)$ (rispett. $p(f)$) la matrice (rispett. il polinomio) così ottenuta/o

$$p(A) = a_0 \cdot I + a_1 \cdot A + a_2 \cdot A^2 + \cdots + a_k \cdot A^k,$$

(ove I è la matrice identica di ordine n e le operazioni sono l'usuale somma e prodotto riga per colonna di matrici, ossia $A^h = A \cdot A \cdot \ldots A$, riga per colonna)

e

$$p(f) = a_0 \cdot id_V + a_1 \cdot f + a_2 \cdot f^2 + \cdots + a_k \cdot f^k,$$

(ove id_v è l'endomorfismo identico e le operazioni sono l'usuale somma e composizione di endomorfismi, ossia $f^h = f \circ f \cdots \circ f$.)

ESEMPIO - Siano
$A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \in Mat_2(\mathbb{R})$, $\quad f : R^2 \to R^2$, $\quad f(\begin{pmatrix} x \\ y \end{pmatrix}) = \begin{pmatrix} 2x \\ 0 \end{pmatrix}$ e
$p(t) = t^2 - 1$.

Risulta

$$p(A) = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 3 \\ 0 & 3 \end{pmatrix}$$

$$p(f)(\begin{pmatrix} x \\ y \end{pmatrix}) = f(f(\begin{pmatrix} x \\ y \end{pmatrix})) - \begin{pmatrix} x \\ y \end{pmatrix} = f(\begin{pmatrix} 2x \\ 0 \end{pmatrix}) - \begin{pmatrix} x \\ y \end{pmatrix} =$$

$$\begin{pmatrix} 4x \\ 0 \end{pmatrix} - \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x \\ -y \end{pmatrix}.$$

OSSERVAZIONE - Se $p, q \in \mathbb{K}[t]$, e $\lambda \in \mathbb{K}$, si ha

$$(p+q)(A) = p(A) + q(A), \qquad (p+q)(f) = p(f) + q(f);$$

$$(p \cdot q)(A) = p(A)q(A), \qquad (p \cdot q)(f) = p(f)q(f);$$

$$(\lambda p)(A) = \lambda p(A), \qquad (\lambda p)(f) = \lambda p(f).$$

Si dice che il *polinomio* $p(t)$ *si annulla su* A (rispett. su f) se $p(A) = O \in Mat_n(K)$ (rispett. $p(f) = 0 \in End(V)$).

Ad esempio, per $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ e $p(t) = (1-t)(2-t) = 2 - 3t + t^2$ (polinomio caratteristico di A), si ha

$$p(A) = 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - 3\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} =$$

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} - \begin{pmatrix} 3 & 3 \\ 0 & 6 \end{pmatrix} + \begin{pmatrix} 1 & 3 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Siano $A \in Mat_n(\mathbb{K})$ e f un endomorfismo di V.

L'insieme

$$\mathcal{J}(A) = \{p(t) \in \mathbb{K}[t] \quad | \quad p(A) = O\}$$

è un ideale di $\mathbb{K}[t]$ (verificarlo).

Lo stesso accade per

$$\mathcal{J}(f) = \{p(t) \in \mathbb{K}[t] \quad | \quad p(f) = 0\}$$

Il polinomio monico generatore di $\mathcal{J}(A)$ (rispettivamente $\mathcal{J}(f)$) viene detto *polinomio minimo* di A (rispett. f). Denoteremo il polinomio minimo di A (rispett. f) con $\mu_A(t)$, (rispett. $\mu_f(t)$).

OSSERVAZIONE - Se A e B sono matrici simili, allora $\mathcal{J}(A) = \mathcal{J}(B)$.
In effetti, per ogni polinomio $p(t)$ risulta $p(A) = O$ se e solo se $p(B) = O$.
Se è $B = M^{-1} \cdot A \cdot M$, allora

$$B^n = (M^{-1} \cdot A \cdot M) \cdot (M^{-1} \cdot A \cdot M) \cdot \dots \cdot (M^{-1} \cdot A \cdot M) \cdot (M^{-1} \cdot A \cdot M) = M^{-1} \cdot A^n \cdot M.$$

Quindi, se $p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_k t^k$, allora

$$p(B) = a_0 I + a_1 B + a_2 B^2 + \cdots + a_k B^k =$$

$$a_0 M^{-1} \cdot I \cdot M + a_1 M^{-1} \cdot A \cdot M + a_2 M^{-1} \cdot A^2 \cdot M + \cdots + a_k M^{-1} \cdot A^k \cdot M = M^{-1} \cdot p(A) \cdot M.$$

Pertanto, se $p(A)$ è la matrice nulla allora anche $p(B)$ lo è (e, per la simmetria della relazione, vale anche il viceversa).

CONSEGUENZA - Matrici simili hanno lo stesso polinomio minimo.

index

Sia $A \in Mat_n(\mathbb{K})$, e

$$P_A(t) = \det(A - tI) = a_0 + a_1 t + \cdots + a_{n-1} t^{n-1} + a_n t^n$$

il suo polinomio caratteristico.

TEOREMA (di Cayley-Hamilton) - Per ogni $A \in Mat_n(\mathbb{K})$, si ha

$$P_A(A) = a_0 I + a_1 A + \cdots + a_{n-1} A^{n-1} + a_n A^n = O,$$

ossia $P_A(t) \in \mathcal{J}(A)$.

Dimostrazione
Ricordo che, data una matrice quadrata M si ha

$$M \cdot^t (cof(M)) = \det M \cdot I.$$

Ponendo $M = A - tI$ e $B =^t (cof(M))$ si ha

$$(A - tI) \cdot B = p_A(t) I = (a_0 + a_1 t + \cdots + a_{n-1} t^{n-1} + a_n t^n) I.$$

Gli elementi della matrice B sono cofattori di $A - tI$ e pertanto sono polinomi di grado $n - 1$ in t. Si può allora scrivere

$$B = B_0 + B_1 t + \cdots + B_{n-1} t^{n-1},$$

dove le B_j sono matrici di costanti.
Nell'equazione di polinomi di matrici

$$(A - tI) \cdot (B_0 + B_1 t + B_2 t^2 + \cdots + B_{n-1} t^{n-1}) = (a_0 + a_1 t + \cdots + a_{n-1} t^{n-1} + a_n t^n) I$$

ordiniamo i due membri secondo diverse potenze di t :

$$A \cdot B_0 + (A \cdot B_1 - B_0) t + (A \cdot B_2 - B_1) t^2 + \cdots + (A \cdot B_{n-1} - B_{n-2}) t^{n-1} - B_{n-1} t^n =$$

$$a_0 I + a_1 I t + \cdots + a_{n-1} I t^{n-1} + a_n I t^n.$$

Uguagliando le matrici coefficienti delle diverse potenze di t si ottengono le relazioni

$$\begin{aligned} A \cdot B_0 &= a_0 I \\ A \cdot B_1 - B_0 &= a_1 I \\ A \cdot B_2 - B_1 &= a_2 I \\ &\vdots \\ A \cdot B_{n-1} - B_{n-2} &= a_{n-1} I \\ -B_{n-1} &= a_n I. \end{aligned}$$

Moltiplicando a sinistra tali relazioni rispettivamente per $I, A, A^2, \ldots, A^{n-1}, A^n$ e sommando otteniamo:

$$A \cdot B_0 + A \cdot (A \cdot B_1 - B_0) + A^2 \cdot (A \cdot B_2 - B_1) + \ldots A^{n-1} \cdot (A \cdot B_{n-1} - B_{n-2}) + A^n \cdot (-B_{n-1})$$

$$= a_0 I + a_1 A + \cdots + a_{n-1} A^{n-1} + a_n A^n$$

ossia

$$O = a_0 I + a_1 A + \cdots + a_{n-1} A^{n-1} + a_n A^n = P_A(A).$$

Il teorema di Cayley-Hamilton vale ovviamente anche per un endomorfismo f di uno spazio vettoriale.

TEOREMA (di Cayley-Hamilton) - Ogni endomorfismo f annulla il suo polinomio caratteristico, ossia $P_f(t) \in \mathcal{J}(f)$.

COROLLARIO - Il polinomio minimo (di una matrice o di un endomorfismo) divide il polinomio caratteristico.

Si può dimostrare che (v. dopo):

- le radici in $\mathbb{K}$ del polinomio minimo sono le stesse del polinomio caratteristico (eventualmente con molteplicità diversa);
- un endomorfismo f è diagonalizzabile se e solo se il polinomio minimo di f è della forma

$$\mu_f(t) = (t - \lambda_1)(t - \lambda_2)\dots(t - \lambda_h),$$

  con $\lambda_1, \lambda_2, \dots \lambda_h \in \mathbb{K}$ tutti distinti tra loro.

ESEMPIO 1

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad P_A(t) = -t^3;$$

Si ha $\mu_A = t^2$ infatti $A^2 = O$, ma $A \neq O$.

ESEMPIO 2

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad P_A(t) = (1-t)(2-t)^2;$$

Si ha $\mu_A = (1-t)(2-t)$ infatti

$$(I-A)\cdot(2I-A) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

ma $I - A \neq O$ e $2I - A \neq O$.

Sia A una matrice quadrata $n \times n$ a coefficienti in $\mathbb{K}$.

Il teorema di Cayley-Hamilton garantisce che le radici (in $\mathbb{K}$) del polinomio minimo $\mu_A(t)$ siano autovalori di A (il polinomio minimo infatti divide il polinomio caratteristico).

Mostriamo viceversa che ogni autovalore di A è anche radice di $\mu_A(t)$.

Sia λ un autovalore di A e sia $\mathbf{v} \neq \mathbf{0}$ un autovettore relativo a λ. Si ha:

$$A\mathbf{v} = \lambda\mathbf{v}, \quad A^2\mathbf{v} = A\lambda\mathbf{v} = \lambda^2\mathbf{v}, \quad \dots, \quad A^h\mathbf{v} = \lambda^h\mathbf{v}, \dots$$

quindi, per ogni polinomio $p(t) = a_0 + a_1 t + \cdots + a^h t^h$ si ha

$$p(A)\mathbf{v} = (a_0 I + a_1 A + \cdots + a^h A^h)\mathbf{v} = (a_0 + a_1\lambda + \cdots + a^h\lambda^h)\mathbf{v} = p(\lambda)\mathbf{v}.$$

Quindi anche $\mu_A(A) \cdot \mathbf{v} = \mu_A(\lambda)\mathbf{v}$. Ma sappiamo che $\mu(A) = O$, quindi $\mu_A(\lambda)\mathbf{v} = O$, per cui, essendo $\mathbf{v} \neq O$, deve essere $\mu_A(\lambda) = 0$.

Vogliamo ora dimostrare che

- una matrice A è diagonalizzabile se e solo se il suo polinomio minimo è della forma
$$\mu_A(t) = (t-\lambda_1)(t-\lambda_2)\dots(t-\lambda_h),$$
con $\lambda_1, \lambda_2, \dots \lambda_h \in \mathbb{K}$ tutti distinti tra loro.

Se A è diagonalizzabile, A è simile a una matrice diagonale

$$B = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda_1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & \lambda_k & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_k \end{pmatrix}$$

e per la matrice B una verifica diretta mostra che

$$(A-\lambda_1 I)\cdot(A-\lambda_2 I)\cdots(A-\lambda_k I) = O.$$

Quindi il polinomio $p(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_k)$ si annulla su A. Sappiamo che il polinomio minimo $\mu_A(t)$ ha $\lambda_1, \lambda_2, dots \lambda_k$ come radici e quindi è un multiplo di $p(t)$. Per definizione di polinomio minimo allora si ha $\mu_A(t) = p(t)$.

Per il viceversa premettiamo un

LEMMA - Siano U, V, W spazi vettoriali su $\mathbb{K}$, e $\phi : U \to V$, $\psi : V \to W$ applicazioni lineari. Allora

$$\dim(\ker(\psi \circ \phi)) \leq \dim(\ker(\phi)) + \dim(\ker(\psi)).$$

Dimostrazione - Poniamo $K = \ker(\psi \circ \phi)$. Si ha $K = \phi^{-1}(\ker(\psi)$. Consideriamo la restrizione θ di ϕ a K:

$$\theta = \phi_{|K} : K \to \ker(\psi).$$

Si ha $\ker(\theta)) \subseteq \ker(\phi)$, infatti, se $\mathbf{u} \in \ker(\theta))$ allora $\theta(\mathbf{v}) = \mathbf{0}$ per cui $\phi(\mathbf{v}) = \mathbf{0}$.

Quindi:

$$\dim(\ker(\psi \circ \phi)) = \dim(K) =$$

$$\dim(\ker(\phi)) + \dim(Im(\theta)) \leq \dim(\ker(\phi)) + \dim(\ker(\psi)),$$

ossia si ha la tesi.

Per induzione si ha anche

$$(*) \dim(\ker(\phi_1 \circ \phi_2 \circ \cdots \circ \phi_k)) \leq \dim(\ker(\phi_1)) + \ldots \dim(\ker(\phi_k)).$$

Supponiamo ora che sia

$$\mu_A(t) = (t - \lambda_1)(t - \lambda_2) \ldots (t - \lambda_h),$$

e mostriamo che A è diagonalizzabile.

Il fatto che $\mu_A(A) = O$ ci dice che

$$(A - \lambda_1 I) \cdot (A - \lambda_2 I) \cdots (A - \lambda_k I) = O,$$

ossia che

$$(L_A - \lambda_1 id) \circ (L_A - \lambda_2 id) \circ (L_A - \lambda_k id)$$

è l'operatore nullo.

Per il Lemma allora si ha

$$n = \dim(V) = \dim(\ker((L_A - \lambda_1 id) \circ (L_A - \lambda_2 id) \circ \cdots \circ (L_A - \lambda_k id)))$$

$$\leq \dim(\ker((L_A - \lambda_1 id)) + \dim(\ker(L_A - \lambda_2 id)) + \cdots + \dim(\ker((L_A - \lambda_k id))$$

$$= \dim(\ker((L_A - \lambda_1 id)) \oplus \ker(L_A - \lambda_2 id)) \oplus \cdots \oplus \ker((L_A - \lambda_k id))$$

$$= \dim(V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}) \leq n$$

(perchè gli autospazi relativi ad autovalori diversi hanno in comune solo il vettore nullo).
Di conseguenza nella formula precedente tutti i $\leq$ sono $=$, e pertanto si ha anche

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k},$$

ossia esiste una base di autovettori e la matrice A è diagonalizzabile

La matrice dell'esempio 1 è una matrice nilpotente: una matrice quadrata A non nulla si dice *nilpotente* se esiste un intero $k > 1$ tale che A^k sia la matrice nulla. Analoga definizione si ha nel caso di un operatore f (nilpotente: $f^k = O$).

Il polinomio minimo di una matrice nilpotente è della forma $t^h, \quad h > 1$.

Una matrice quadrata A non nulla e diversa dalla matrice identica si dice *idempotente* se $A^2 = A$. Analoga definizione si ha nel caso di un operatore f (idempotente: $f^2 = f$).

Il polinomio minimo di una matrice idempotente è della forma $t^2 - t$.

Un esempio di operatore idempotente $f : R^n \to R^n$ è una proiezione

$$f\begin{pmatrix} x_1 \\ \vdots \\ x_h \\ x_{h+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_h \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Il teorema di Cayley-Hamilton può essere usato anche per calcolare l' inversa di una matrice,o le potenze di una matrice.

ESEMPIO

Data $A = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}$ si ha $p_A(t) = (1-t)^2 - 4 = t^2 - 2t - 3$ quindi $A^2 - 2A - 3I = O$. Allora $A \cdot (A - 2I) = 3I$, quindi

$$A^{-1} = \frac{1}{3}(A - 2I) = \frac{1}{3}\begin{pmatrix} -1 & 1 \\ 4 & -1 \end{pmatrix}.$$

Inoltre $A^2 = 2A + 3I$ per cui, ad esempio,

$$A^5 = A \cdot (A^2)^2 = A \cdot (2A+3I)^2 = A \cdot (4A^2+12A+9I) = A \cdot (4(2A+3I)+12A+9I)$$

$$= A \cdot (20A + 21I) = 20A^2 + 21A = 20(2A + 3I) + 21A = 61A + 60I =$$

$$\begin{pmatrix} 121 & 244 \\ 61 & 121 \end{pmatrix}.$$