Complementi sulla diagonalizzazione. (A.A. 2010-11)

Ph. Ellia

Printed:
24-01-2011

Indice

I

Il polinomio minimo: definizione.

1. Definizione.

Notazioni 1.1: *Nel seguito denoteremo con E un k spazio vettoriale di dimensione n e con $f : E \to E$ un endomorfismo di E. Noteremo con $P_f(X)$ il polinomio caratteristico di f.*

Se $P(X) = a_nX^n + ... + a_1X + a_0 \in k[X]$, allora $P(f) = a_nf^n + ... + a_1f + a_0Id \in End(E)$. In questo modo si ottiene un'applicazione: $\varphi : k[X] \to End(E)$. Si verifica che φ è un morfismo d'anelli. Quindi: $(PQ)(f) = P(f) \circ Q(f)$ (per verificarlo iniziare con $P(X) = X^m, Q(X) = X^t$ e concludere per linearità).

Fatto **importante**: siccome $k[X]$ è commutativo, abbiamo:
$P(f) \circ Q(f) = Q(f) \circ P(f), \forall P, Q \in k[X]$.

Siccome $dim(End(E)) = n^2$, $Id, f, f^2, ..., f^{n^2}$ sono legati: $a_0Id + ... + a_{n^2}f^{n^2} = 0$, ossia $Q(f) = 0$ con $a_0 + ... + a_{n^2}X^{n^2} = Q(X) \neq 0$. Questo mostra che l'applicazione φ non è iniettiva. Sia $J = Ker(\varphi)$. Il sottinsieme J di $k[X]$ è un *ideale*, cioè: se $P, Q \in J$ allora $P + Q \in J$ e: $\forall P \in J, \forall T \in k[X], TP \in J$.
Un altro fatto **cruciale** è:

Lemma 1.2: *Ogni ideale, J, di $k[X]$ può essere generato da un unico polinomio: $\exists T \in k[X]$ t.c. $J = (T) = \{PT \mid P \in k[X]\}$.*

Dim: Il modo più semplice per vederlo è usare la divisione euclidea. Ricordiamo che se P, M sono due polinomi con grado(P)$\geq$ grado(M), allora esistono Q, R tali che: $P = QM + R$ e grado(R)$<$ grado(M). Detto ciò sia $M \in J$ di *grado minimo*. Se $P \in J$, dividiamo per M: $P = QM + R$. Abbiamo: $R = P - QM \in J$. Siccome grado(R)$<$grado(M), per minimalità del grado di M, l'unica possibilità è: $R = 0$. Quindi $P = QM$. □

Osservazione 1.3: *Un anello (integro) in cui ogni ideale può essere generato da un unico elemento si chiama un PID (principal ideal domain).*
Ogni anello euclideo (cioè provvisto di una divisione euclideo) è un PID. Per esempio $\mathbb{Z}$ è un PID. Ci sono dei PID che non sono euclidei.

Quindi ogni ideale $J \subset k[X]$ può scriversi $J = (M)$ dove M è un polinomio di grado minimo in J. Il polinomio non è univocamente determinato, se $a \in k$, $a \neq 0$, allora $J = (aM)$. In particolare possiamo sempre trovare a tale che aM sia *monico* cioè il coefficiente della più alta potenza di X in aM sia uguale a 1: $aM = X^t + a_{t-1}X^{t-1} + ... + a_0$. Il generatore monico di J è univocamente determinato. Nel seguito quando scriveremo $J = (P)$ si assumerà sempre che P sia il generatore monico.

Definizione 1.4: *Il polinomio minimo, M_f, di f è il generatore monico dell'ideale $J = Ker(\varphi)$.*

Osservazione 1.5: *(i) Il polinomio minimo è il polinomio monico, P, di grado più piccolo tale che $P(f) = 0$.*
(ii) Se P verifica $P(f) = 0$, allora $M_f|P$ (M_f divide P).

Nel seguito cercheremo di studiare le relazioni tra il polinomio minimo M_f e il polinomio caratteristico P_f.

2. Esempi ed esercizi.

Esempio 2.1: Il polinomio minimo di un endomorfismo non nullo ha sempre grado almeno uno.
Inoltre: M_f ha grado uno $\Leftrightarrow$ $f = \alpha Id$ per qualche $\alpha \in k, \alpha \neq 0$ (cioè f è un'omotetia).
Infatti se $M_f(X) = X - \alpha$, da $M_f(f) = 0$ segue $(f - \alpha Id)(v) = 0, \forall v$, cioè: $f(v) = \alpha v, \forall v$. Viceversa se $f(v) = \alpha v, \forall v$, allora $M(f) = 0$ con $M(X) = (X - \alpha)$ e, per minimalità: $M = M_f$.

Esempio 2.2: Sia $f : E \to E$, con $dim(E) = 2$ e $P_f(X) = (X - \alpha)(X - \beta)$.
Se $\alpha \neq \beta$ allora f è diagonalizzabile e $M_f(X) = P_f(X)$. Infatti siccome esistono $v, w \in E$ non nulli con $f(v) = \alpha v$, $f(w) = \beta w$, vediamo che f non è un'omotetia. Dall'esempio precedente segue che $grado(M_f) \geq 2$. Sia $B = (v, w)$ una base di autovettori. Abbiamo: $(f - \alpha Id) \circ (f - \beta Id)(v) = (f - \beta Id) \circ (f - \alpha Id)(v) = (f - \beta Id)(f(v) - \alpha v) = 0$. Nello stesso modo: $(f - \alpha Id) \circ (f - \beta Id)(w) = 0$, quindi $(f - \alpha Id) \circ (f - \beta Id) = 0$. Si conclude per minimalità.

Esercizio 2.1: *Sia $f : E \to E$, con $dim(E) = 2$ e $P_f(X) = (X - \alpha)^2$, $f \neq \alpha Id$. Mostrare che: $M_f = P_f$.*

II

Il teorema di Cayley-Hamilton.

1. Enunciato del teorema di Cayley-Hamilton.

Scopo di questa sezione è dimostrare il:

Teorema 1.1: *[Cayley-Hamilton]*
Sia E un k spazio vettoriale e sia $f : E \to E$ un endomorfismo. Allora $P_f(f) = 0$. Equivalentemente: se $A \in \mathcal{M}_n(k)$, allora A è radice del suo polinomio caratteristico: $P_A(A) = 0$.

Osservazione 1.2: *Questo risultato è molto importante per il seguito.*
Dalla definizione di M_f segue che $M_f | P_f$, cioè: $P_f = M_f.Q$. Vedremo più avanti un risultato più preciso: il polinomio minimo e il polinomio caratteristico hanno gli stessi fattori irriducibili. In particolare, se k è algebricamente chiuso, P_f e M_f hanno le stesse radici (ma eventualmente con molteplicità diverse).
Nel seguito daremo due dimostrazioni diverse del teorema di Cayley-Hamilton.

2. Prima dimostrazione del teorema di Cayley-Hamilton.

Prima di iniziare la dimostrazione facciamo alcuni richiami sulla matrice complementare di una matrice $M \in \mathcal{M}_n(k)$. La matrice complementare, M^c, è la trasposta della matrice dei cofattori di M e vale la relazione:

$$M.M^c = M^c.M = det(M).I_n \quad (*)$$

In realtà la definizione di M^c e la relazione $(*)$ valgono per matrici a coefficienti in un anello commutativo $\mathcal{A}$ ([**E**], II.14, Oss.16.1). Nel seguito sarà $\mathcal{A} = k[X]$.
Per definizione: $P_f(X) = det(XI_n - A)$ dove A è la matrice di f rispetto ad una base qualsiasi. Poniamo $M := XI_n - A \in \mathcal{M}_n(k[X])$. I coefficienti della matrice M sono dei polinomi di grado al più uno: $M = (\delta_{ij}X - a_{ij})$.
Da $(*)$ segue che:

$$M.M^c = P_f(X).I_n \quad (**)$$

I coefficienti della matrice M^c sono i minori di ordine $(n-1)$ di M, quindi sono polinomi di grado al più $(n-1)$. Pertanto M^c è un polinomio di grado (al più) $(n-1)$ a coefficienti in $\mathcal{M}_n(k)$. Facciamo un esempio per chiarire quest'ultima

affermazione. Sia $P = \begin{pmatrix} X^2 - X + 3 & X - 1 \\ 3X^2 + X + 2 & -X^2 + 3 \end{pmatrix}$, allora

$$P = X^2 \begin{pmatrix} 1 & 0 \\ 3 & -1 \end{pmatrix} + X \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 3 & -1 \\ 2 & 3 \end{pmatrix}$$

Ovviamente questa scrittura di P come polinomio in X a coefficienti in $\mathcal{M}_n(k)$ è univocamente determinata.

Dimostrazione del teorema di cayley-Hamilton. Con le notazioni precedenti abbiamo: $M^c = B_1X^{n-1} + ... + B_{n-1}X + B_n$ dove $B_i \in \mathcal{M}_n(k)$. La relazione $(**)$ si scrive:

$$(XI_n - A)(B_1X^{n-1} + ... + B_{n-1}X + B_n) = P_f(X).I_n \ (+)$$

Poniamo $P_f(X) = X^n + a_1X^{n-1} + ... + a_n$ $(a_i \in k, \forall i)$, allora svolgendo in $(+)$ viene:

$$X^nB_1 + X^{n-1}(B_2 - AB_1) + ... + X(B_n - AB_{n-1}) - AB_n = X^nI_n + X^{n-1}a_1I_n + ... + a_nI_n$$

Uguagliando i coefficienti di X^i otteniamo:
$B_1 = I_n$
$B_2 - AB_1 = a_1I_n$
.....
$B_n - AB_{n-1} = a_{n-1}I_n$
$-AB_n = a_nI_n$
Moltiplicando la prima equazione per A^n, la seconda per A^{n-1}, ecc... fino alla penultima per A, otteniamo:
$B_1A^n = A^n$
$B_2A^{n-1} - A^nB_1 = a_1A^{n-1}$
.....
$B_nA - A^2B_{n-1} = a_{n-1}A$
$-AB_n = a_nI_n$
Sommando membro a membro otteniamo:

$$A^n + a_1A^{n-1} + ... + a_nI_n = P_f(A) = 0$$

(In effetti i termini dei primi membri si cancellano tutti.) Quindi la matrice di f verifica il polinomio caratteristico, in termini di endomorfismi: $P_f(f) = 0$. Il teorema è dimostrato. □

3. Matrici triangolari e teorema di Cayley-Hamilton.

Iniziamo con un risultato generale sotto l'ipotesi che il polinomio caratteristico abbia tutte le sue radici nel campo k.

Proposizione 3.1: *Sia E un k spazio vettoriale di dimensione n e sia $f : E \to E$ un endomorfismo di E. Si suppone che il polinomio caratteristico di f abbia tutte le sue radici in k: $P_f(X) = (X - \lambda_1)(X - \lambda_2)..(X - \lambda_n)$ (i λ_i non necessariamente distinti). Allora esiste una base B di E tale che $mat(f; B, B)$ sia triangolare superiore, con $\lambda_1, ..., \lambda_n$ sulla diagonale.*

DIM. La dimostrazione è per induzione su $n = dim(E)$. Se $n = 1$ non c'è nulla da dimostrare. Supponiamo il risultato vero per $n - 1$. Sia $e_1 \neq 0$ tale che $f(e_1) = \lambda_1 e_1$ e sia $B_1 = (e_1, e'_2, .., e'n)$ una base contenente e_1. Abbiamo:

$$mat(f; B_1, B_1) = \begin{pmatrix} \lambda_1 & a_{12}...a_{1n} \\ 0 & \\ . & C \\ 0 & \end{pmatrix}$$

Sia $E' =< e'_2, ..., e'_n >$ e sia $f' : E' \to E \to E \to E'$ la composta $f' = p \circ f \circ i$ dove $i : E' \to E$ è l'inclusione e dove $p : E \to E'$ è la proiezione. Allora $B' = (e'_2, .., e'_n)$ è una base di E' e $mat(f'; B', B') = C$. Inoltre calcolando il polinomio caratteristico usando $maf(f; B_1, B_1)$ vediamo che $P_{f'}(X) = (X - \lambda_2)...(X - \lambda_n)$. Quindi f' soddisfa l'ipotesi di induzione ed esiste una base $B'_2 = (v_2, ..., v_n)$ di E' tale che $mat(f'; B'_2, B'_2)$ sia triangolare superiore. I vettori $e_1, v_2, ..., v_n$ sono linearmente indipendenti, infatti da $\alpha e_1 + \alpha_2 v_2 + ... + \alpha_n v_n = 0$ segue $\alpha e_1 \in E' =< e'_2, ..., e'_n >$ e questo implica $\alpha = 0$. Segue poi che $\alpha_2 = ... = \alpha_n = 0$. Nella base $B = (e_1, v_2, ..., v_n)$ la matrice di f è triangolare superiormente:

$$mat(f; B, B) = \begin{pmatrix} \lambda_1 & b_{12} & ...b_{1n} \\ 0 & \lambda_2 & .... \\ ... & .... & .... \\ 0... & 0 & \lambda_n \end{pmatrix}$$

Gli elementi sulla diagonale sono necessariamente $\lambda_1, ..., \lambda_n$ come si vede calcolando il polinomio caratteristico. □

Corollario 3.2: *Sia $f : E \to E$ un endomorfismo del k spazio vettoriale E. Se $P_f(X)$ ha tutte le sue radici in k (ipotesi senz'altro verificata se k è algebricamente chiuso), allora $P_f(f) = 0$.*

DIM. Se $P_f(X)$ ha tutte le sue radici in k, dalla Prop. Proposizione 3.1, esiste una base B di E tale che $mat(f; B, B)$ sia triangolare superiore (con gli autovalori λ_i sulla diagonale): $mat(f; B, B) = \begin{pmatrix} \lambda_1 & b_{12} & ...b_{1n} \\ 0 & \lambda_2 & .... \\ ... & .... & .... \\ 0... & 0 & \lambda_n \end{pmatrix}$.

Poniamo $f_i = (\lambda_i.Id - f)$ e $F_i = f_1 \circ f_2 \circ ... \circ f_i$, $1 \leq i \leq n$. Osserviamo che $f_i \circ f_j = f_j \circ f_i$. Infatti se $P_i(X) = (\lambda_i - X)$, allora $P_i(f) = f_i$. Siccome: $P_i(X).P_j(X) =$

$P_j(X).P_i(X)$, abbiamo $f_i \circ f_j = f_j \circ f_i$.
Mostriamo, per induzione su t, che $F_t(e_i) = 0$ se $i \leq t$.
Abbiamo $F_1(e_1) = (\lambda_1 Id - f)(e_1) = \lambda_1 e_1 - f(e_1) = 0$ (perchè $f(e_1) = \lambda_1 e_1$ per costruzione della base B).
Supponiamo l'enunciato vero per $t-1$ e dimostriamolo per t: se $j < t$, $F_t(e_j) = (F_{t-1} \circ f_t)(e_j) = (f_t \circ F_{t-1})(e_j) = f_t(F_{t-1}(e_j)) = 0$ (perchè $F_{t-1}(e_j) = 0$ per ipotesi di induzione). Siccome $f_t(e_t) = (\lambda_t.Id - f)(e_t) = \sum_{i=1}^{i=t-1} c_i e_i$, $F_t(e_t) = (F_{t-1} \circ f_t)(e_t) = \sum_{i=1}^{i=t-1} c_i F_{t-1}(e_i) = 0$; e questo conclude la dimostrazione per induzione.
In particolare $F_n(e_i) = 0, \forall i, 1 \leq i \leq n$, cioè $F_n = 0$. Quindi $F_n = (\lambda_1.Id - f) \circ (\lambda_2.Id - f) \circ ... \circ (\lambda_n.Id - f) = P_f(f) = 0$ e la proposizione è dimostrata. □

Osservazione 3.3: *Il corollario precedente dimostra il teorema di Cayley-Hamilton sotto l'ipotesi che $P_f(X)$ abbia tutte le sue radici in k. Osserviamo che questa ipotesi è senz'altro verificata se k è algebricamente chiuso. Quindi il teorema di Cayley-Hamilton è dimostrato per k algebricamente chiuso.*
Per passare al caso generale (k non necessariamente algebricamente chiuso) useremo un risultato assai profondo di algebra:

Teorema 3.4: *Ogni campo k è un sottocampo di un campo algebricamente chiuso. Più precisamente esiste un unico (modulo isomorfismo) "più piccolo" campo algebricamente chiuso contenente k, questo campo, notato $\overline{k}$, è la chiusura algebrica di k.*

Per esempio la chiusura algebrica di $\mathbb{R}$ è $\mathbb{C}$; ma la chiusura algebrica di $\mathbb{Q}$, $\overline{\mathbb{Q}}$, non è $\mathbb{C}$, infatti l'estensione di campi $\mathbb{Q}/\mathbb{C}$ non è algebrica (esistono numeri trascendenti). Abbiamo $\mathbb{Q} \subset \overline{\mathbb{Q}} \subset \mathbb{C}$ e tutte le inclusioni sono strette.
Si rimanda a un buon testo di algebra per una dimostrazione del teorema.

Seconda dimostrazione del teorema di cayley-Hamilton. Sia f un endomorfismo del k spazio vettoriale E. Sia $A = mat(f; B, B)$ la matrice di f rispetto ad una qualsiasi base B di E. Abbiamo $k \subset \overline{k}$, con $\overline{k}$ algebricamente chiuso (Teorema 3.4). Quindi $A \in \mathcal{M}(k) \subset \mathcal{M}_n(\overline{k})$ e possiamo considerare A come la matrice di un endomorfismo di $\overline{k}^n$ ($n = dim(E)$). Abbiamo $P_A(X) \in k[X] \subset \overline{k}[X]$; adesso $P_A(X)$, visto come polinomio in $\overline{k}$, ha tutte le sue radici in $\overline{k}$. Quindi (cf Corollario 3.2): $P_A(A) = 0$. Siccome $P_A(X) = X^n + a_1 X^{n-1} + ... + a_n$ con $a_i \in k, \forall i$, la relazione $P_A(A) = 0$ è verificata anche in $\mathcal{M}_n(k)$, pertanto $P_f(f) = 0$. □

III

Polinomio minimo e polinomio caratteristico.

1. Primi risultati.

Proposizione 1.1: *Sia $f : E \to E$ un endomorfismo del k spazio vettoriale E. Il polinomio minimo di f divide il polinomio caratteristico di f: $M_f | P_f$.*

DIM. Segue dal teorema di Cayley-Hamilton ($P_f(f) = 0$) e dalla definizione del polinomio minimo (cf I.Sezione 1; Osservazione 1.5 (ii)). □

Osservazione 1.2: *Dalla proposizione precedente: $M_f.Q = P_f$. A priori questo* non *implica che se λ è radice di P_f allora λ è anche radice di M_f (λ potrebbe essere radice solo di Q). In realtà come vedremo adesso M_f e P_f hanno le stesse radici.*

Proposizione 1.3: *Sia $f : E \to E$ un endomorfismo del k spazio vettoriale E. Un elemento $\lambda \in k$ è radice di P_f (cioè λ è un autovalore) se e solo se λ è radice di M_f.*

DIM. (a) Sia λ un autovalore e sia $v \in E, v \neq 0$ tale che $f(v) = \lambda v$. Sia $M_f(X) = X^t + c_1 X^{t-1} + ... + c_t$. Siccome $M_f(f) = 0$, abbiamo $M_f(f)(v) = 0$. Quindi: $(f^t + c_1 f^{t-1} + ... + c_t Id)(v) = 0$. Tenendo conto che $f^i(v) = \lambda^i.v$, viene: $0 = \lambda^t.v + c_1\lambda^{t-1}.v + ... + c_t.v = (\lambda^t + \lambda^{t-1}c_1 + ... + c_t).v$. Siccome $v \neq 0$, segue che: $\lambda^t + \lambda^{t-1}c_1 + ... + c_t = M_f(\lambda) = 0$.
(b) Viceversa se λ è radice di M_f, allora $(X-\lambda)|M_f(X)$. Siccome $M_f|P_f$, $(X-\lambda)|P_f$ e λ è radice di P_f. □

Corollario 1.4: *Sia $f : E \to E$ un endomorfismo del k spazio vettoriale E. Si suppone che $P_f(X)$ abbia tutte le sue radici in k (ipotesi verificata se k è algebricamente chiuso):*

$$P_f(X) = (X - \lambda_1)^{\alpha_1}.(X - \lambda_2)^{\alpha_2}...(X - \lambda_r)^{\alpha_r}$$

con $\lambda_i \neq \lambda_j$ se $i \neq j$ e $\alpha_1 + ... + \alpha_r = dim(E)$. Allora:

$$M_f(X) = (X - \lambda_1)^{\beta_1}.(X - \lambda_2)^{\beta_2}...(X - \lambda_r)^{\beta_r} \ \ con: 1 \leq \beta_i \leq \alpha_i, \forall i.$$

DIM. Se $Q(X)$ è un fattore irriducibile di $M_f(X)$, allora Q è un fattore irriducibile di $P_f(X)$ (perchè $M_f|P_f$), quindi Q ha grado uno. Dalla Proposizione 1.3

segue che: $M_f(X) = (X-\lambda_1)^{\beta_1}.(X-\lambda_2)^{\beta_2}...(X-\lambda_r)^{\beta_r}$ con $\beta_i \geq 1, \forall i$. Siccome $M_f|P_f$, $\beta_i \leq \alpha_i$, $\forall i$. □

Corollario 1.5: *Sia $f : E \to E$ un endomorfismo del k spazio vettoriale E. Si suppone che $P_f(X)$ abbia tutte le sue radici in k (ipotesi verificata se k è algebricamente chiuso):*

$$P_f(X) = (X-\lambda_1)^{\alpha_1}.(X-\lambda_2)^{\alpha_2}...(X-\lambda_r)^{\alpha_r}$$

con $\lambda_i \neq \lambda_j$ se $i \neq j$ e $\alpha_1 + ... + \alpha_r = dim(E)$.
Se f è diagonalizzabile:

$$M_f(X) = (X-\lambda_1).(X-\lambda_2)...(X-\lambda_r).$$

In particolare M_f non ha radici multiple.

DIM. Sia $B = (e_{1_1}, ..., e_{1_{\alpha_1}}; e_{2_1}, ..., e_{2_{\alpha_2}}; ...; e_{r_1}, ..., e_{r_{\alpha_r}})$ una base di autovettori ($f(e_{i_j}) = \lambda_i e_{i_j}$). Sia $F = (f-\lambda_1.Id) \circ (f-\lambda_2.Id) \circ ...(f-\lambda_r.Id)$. Abbiamo $F(e_{i_j}) = (f-\lambda_1.Id) \circ ... \circ (f-\lambda_i.Id) \circ ...(f-\lambda_r.Id)(e_{i_j}) = (f-\lambda_1.Id) \circ ... \circ (f-\lambda_r.Id) \circ ((f-\lambda_i.Id)(e_{i_j})) = 0$. Quindi $F = 0$. Segue che $M_f|(X-\lambda_1)...(X-\lambda_r)$. Si conclude con il Corollario 1.4. □

Osservazione 1.6: *La dimostrazione del Corollario 1.5 usa il Corollario 1.4 e quindi il teorema di Cayley-Hamilton. In realtà si può dimostrare il Corollario 1.5 senza Cayley-Hamilton: come nella dimostrazione precedente si vede che $F = 0$, quindi $M_f|(X-\lambda_1)...(X-\lambda_r)$. Segue che ogni fattore irriducibile di M_f è della forma $(X-\lambda_i)$. Si conclude con la Proposizione 1.3.*
Si può dimostrare in modo "elementare" (cioè senza usare Cayley-Hamilton) una versione più debole del Corollario 1.4: con le ipotesi del Corollario 1.4, $M_f(X) = (X-\lambda_1)^{\beta_1}.(X-\lambda_2)^{\beta_2}...(X-\lambda_r)^{\beta_r}$ con $: 1 \leq \beta_i, \forall i$ (cf Esercizio 3.1). A questo punto l'asserzione $\beta_i \leq \alpha_i, \forall i$ è ovviamente equivalente al teorema di Cayley-Hamilton.

2. Il caso reale.

Dalla Proposizione 1.3 segue che P_f e M_f hanno le stesse radici. Vedremo più avanti un risultato più forte: P_f e M_f hanno gli stessi fattori irriducibili.
Grazie ad un semplice lemma di algebra si può dimostrare facilmente questo risultato se $k = \mathbb{R}$.

Lemma 2.1: *Sia $P(X) \in \mathbb{R}[X]$ un polinomio non costante. La fattorizzazione di P in elementi irriducibili è della forma: $P(X) = Q_1^{\alpha_1}...Q_r^{\alpha_r}$ con $1 \leq grado(Q_i) \leq 2, \forall i$. Inoltre se Q_i ha grado due, Q_i non ha radici reali.*
In altri termini un polinomio irriducibile a coefficienti reali ha grado al più due.

DIM. Sia $Q(X) \in \mathbb{R}[X]$ un polinomio irriducibile di grado almeno due. In particolare Q non ha radici reali. Consideriamo $Q(X) \in \mathbb{C}[X]$. Siccome $\mathbb{C}$ è algebricamente chiuso, esiste $z \in \mathbb{C}$ tale che $Q(z) = 0$. Abbiamo $\overline{Q(z)} = 0$, ma siccome Q è a coefficienti reali: $\overline{Q(z)} = Q(\overline{z})$ quindi $\overline{z}$ (il coniugato di z) è radice di Q. Quindi $(X - z)(X - \overline{z})|Q(X)$. Siccome $(X - z)(X - \overline{z}) = X^2 - (z + \overline{z})X + z\overline{z} \in \mathbb{R}[X]$, per irriducibilità abbiamo: $Q(X) = c(X^2 - (z + \overline{z})X + z\overline{z})$ dove $c \in \mathbb{R}$. □

Corollario 2.2: *Sia $f : E \to E$ un endomorfismo del $\mathbb{R}$ spazio vettoriale E. Allora il polinomio caratteristico di f è della forma:*

$$P_f(X) = Q_1(X)^{\alpha_1}...Q_r(X)^{\alpha_r},\ 1 \leq grado(Q_i) \leq 2, \forall i$$

e il polinomio minimo è della forma:

$$M_f(X) = Q_1(X)^{\beta_1}...Q_r(X)^{\beta_r},\ 1 \leq \beta_i \leq \alpha_i, \forall i.$$

In particolare P_f e M_f hanno gli stessi fattori irriducibili.

DIM. Per il lemma Lemma 2.1 P_f e M_f fattorizzano come prodotti di polinomi di grado al più due. Si conclude considerando $P_f(X), M_f(X) \in \mathbb{C}[X]$ e usando il Corollario 1.4. □

Osservazione 2.3: *Il Corollario 2.2 fornisce un algoritmo per calcolare M_f:*
(i) Si procede a fattorizzare P_f: $P_f(X) = Q_1(X)^{\alpha_1}...Q_r(X)^{\alpha_r},\ 1 \leq grado(Q_i) \leq 2, \forall i$
(ii) Si tratta poi di trovare $\beta_i, ..., \beta_r$, $1 \leq \beta_i \leq \alpha_i, \forall i$, minimi tali che: $Q_1^{\beta_1}(f) \circ ... \circ Q_r^{\beta_r}(f) = 0$.

3. Esempi ed esercizi

Esercizio 3.1: *Sia $f : E \to E$ un endomorfismo del k spazio vettoriale E. Dimostrare* senza usare il teorema di Cayley-Hamilton *che: $M_f(\lambda) = 0 \Rightarrow P_f(\lambda) = 0$.*

Esercizio 3.2: *Sia $f : E \to E$ un endomorfismo del $\mathbb{R}$ spazio vettoriale E con $dim(E) = 3$. Si assume $P_f(X) = (X - \alpha)^2(X - \beta)$ $(\alpha \neq \beta)$. Mostrare che se $M_f(X) = (X - \alpha)(X - \beta)$, allora f è diagonalizzabile.*

Esercizio 3.3: *Determinare il polinomio minimo delle seguenti matrici ($\alpha \in k$):*

$$A = \begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}, B = \begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{pmatrix}$$

Esercizio 3.4: [Endomorfismo di uno spazio vettoriale reale di dimensione tre.] *Sia $f : E \to E$ un endomorfismo del $\mathbb{R}$ spazio vettoriale E con $dim(E) = 3$. Per il polinomio caratteristico di f abbiamo le seguenti possibilità:*
(a) P_f ha un'unica radice reale: $P_f(X) = (X - \alpha)Q(X)$ con Q di grado due, senza

radici reali
(b) P_f ha tutte le sue radici in $\mathbb{R}$ e:
(b1) P_f ha una radice tripla: $P_f(X) = (X-\alpha)^3$
(b2) P_f ha una radice doppia: $P_f(X) = (X-\alpha)^2(X-\beta)$ $(\alpha \neq \beta)$
(b3) P_f ha tre radici distinte: $P_f(X) = (X-\alpha)(X-\beta)(X-\delta)$, con α, β, δ due a due distinti.
Cosa si può dire di M_f in ogni caso?
Concludere che: f è diagonalizzabile $\Leftrightarrow$ M_f ha tutte le sue radici in $\mathbb{R}$ e M_f non ha radici multiple. (N.B. Vedremo che questo è un fatto generale.)

IV

Polinomio minimo e diagonalizzazione.

Scopo di questo capitolo è dimostrare il seguente:

Teorema: *Sia $f : E \to E$ un endomorfismo del k spazio vettoriale E. Allora: f è diagonalizzabile $\Leftrightarrow$ M_f ha tutte le sue radici in k e M_f non ha radici multiple.*

1. Proiettori.

Definizione 1.1: *Un endomorfismo $h : E \to E$ del k spazio vettoriale E tale che $h^2 = h$ viene chiamato* proiettore.

Proposizione 1.2: *Sia $h : E \to E$ un proiettore, allora:*
(i) $Im(h) = \{x \in E \,|\, h(x) = x\}$, $Ker(h) = Im(Id - h)$
(ii) $E = Ker(h) \oplus Im(h)$
(iii) h è diagonalizzabile con autovalori 1 e 0, e si ha: $E_h(1) = Im(h)$ e $E_h(0) = Ker(h)$.

Dim. (i) Se $y \in Im(h)$, allora $y = h(x)$ e $h(y) = h^2(x) = h(x)$, quindi $y = h(y)$. E' chiaro che $\{x \,|\, h(x) = x\} \subset Im(h)$.
Se $y \in Im(Id-h)$, $y = x-h(x)$; quindi $h(y) = h(x)-h^2(x) = 0$ ($h^2 = h$). Pertanto $y \in Ker(h)$. Viceversa se $h(y) = 0$, allora $y = y - h(y) \in Im(Id - h)$.
(ii) Sia $x \in E$, abbiamo $x = (x - h(x)) + h(x)$. Per (i): $x - h(x) \in Ker(h)$. Quindi questo mostra: $E = Ker(h) + Im(h)$. Mostriamo che la somma è diretta: se $x \in Ker(h) \cap Im(h)$, allora $x = h(y)$ ($x \in Im(h)$) e $0 = h(x) = h^2(y)$. Ma $h^2(y) = h(y) = x$, quindi $x = 0$.
(iii) Da (i) ogni vettore non nullo di $Im(h)$ è un autovettore per l'autovalore 1 (e viceversa), quindi $Im(h) = E_h(1)$. E' chiaro che $E_h(0) = Ker(h)$, si conclude con (ii). □

Osservazione 1.3: *(i) Sia $p : \mathbb{R}^3 \to \mathbb{R}^3 : (x, y, z) \to (x, y, 0)$. L'endomorfismo p è la proiezione sul piano $< x, y >$, parallelamente all'asse $< z >$. Chiaramente $Im(p) =< x, y >$ e $Ker(p) =< z >$, inoltre: $\mathbb{R}^3 =< x, y > \oplus < z >$. Finalmente $p^2 = p$ e p è un proiettore. Questa situazione è tipica.*
(ii) Sia h un proiettore di E. Sia $(e_1, ..., e_r)$ una base di $Im(h)$ e $(e_{r+1}, ..., e_n)$ una base di $Ker(h)$. Dalla Proposizione 1.2 $B = (e_1, ..., e_n)$ è una base di E e

$mat(h;B,B) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. *Quindi* h *è la proiezione su* $< e_1,...,e_r >= Im(h)$ *parallelamente a* $< e_{r+1},...,e_n >= Ker(h)$.

Lemma 1.4: *Sia* h *un proiettore di* E, $h \neq Id$. *Allora:*
(i) $P_h(X) = (X-1)^r X^{n-r}$ $(r = rgo(h), n = dim(E))$ *e* $M_h(X) = X(X-1)$.
(ii) Anche $Id-h$ *è un proiettore: è la proiezione su* $Ker(h)$ *parallelamente a* $Im(h)$.
(iii) Siano $c_1, c_2 \in k$ *e* $f := c_1 h + c_2(Id-h)$, *allora* f *è diagonalizzabile.*

DIM. (i) Siccome h è diagonalizzabile con autovalori 1 e 0 (Proposizione 1.2), si ha subito: $P_h(X) = (X-1)^r X^{n-r}$.
Siccome $h \circ (h - Id) = 0$ (Proposizione 1.2, (i)), $M_h(X) | X(X-1)$, si conclude per minimalità.
(ii) Abbiamo: $(Id-h)^2 = Id^2 - 2h + h^2 = Id - 2h + h = Id - h$. Se $B = (e_1,...,e_r,e_{r+1},...,e_n)$ è una base di autovettori di h (con $(e_1,...,e_r)$ base di $Im(h)$), allora: $mat(h;B,B) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ e $mat(Id-h;B,B) = \begin{pmatrix} 0 & 0 \\ 0 & I_{n-r} \end{pmatrix}$, quindi $Id-h$ è la proiezione su $Ker(h)$ parallelamente a $Im(h)$.
(iii) Con le notazioni precedenti: $mat(f;B,B) = \begin{pmatrix} c_1 I_r & 0 \\ 0 & c_2 I_{n-r} \end{pmatrix}$. □

Sia $f : E \to E$ un endomorfismo diagonalizzabile con autovalori distinti $\lambda_1,...,\lambda_r$. Quindi $E = E_1 \oplus ... \oplus E_r$ dove E_i è l'autospazio relativo a λ_i. Ogni $x \in E$ si scrive in modo unico: $x = x_1 + ... + x_r$ con $x_i \in E_i$.
Sia $p_i : E \to E : x \to x_i$ la proiezione su E_i parallelamente a $\oplus_{j \neq i} E_j$. Abbiamo $Im(p_i) = E_i$ e $Ker(p_i) = \oplus_{j \neq i} E_j$.
Inoltre: $Id = p_1 + ... + p_r$ e $p_i \circ p_j = 0$ se $i \neq j$ $(Im(p_j) \subset Ker(p_i))$ $(*)$.
Queste due condizioni rispecchiano il fatto che E è somma diretta degli autospazi E_i. Finalmente: $f = \lambda_1 p_1 + ... + \lambda_r p_r$.
Il prossimo risultato (punto chiave nella dimostrazione del teorema principale) mostra che questa situazione è caratteristica degli endomorfismi diagonalizzabili: f è diagonalizzabile se e solo se f è combinazione lineare di proiettori che soddisfano alle due condizioni $(*)$:

Teorema 1.5: *Siano* $h_1,...,h_t$ *dei proiettori del* k *spazio vettoriale* E. *Si assume:*
(i) $Id = h_1 + ... + h_t$
(ii) $h_i \circ h_j = 0$ *se* $i \neq j$
Se $c_1,...,c_t$ *sono degli elementi di* k *e se* $f = c_1 h_1 + ... + c_t h_t$, *allora* f *è diagonalizzabile.*

DIM. (a) Mostriamo che, sotto le nostre ipotesi, abbiamo:

$$E = Im(h_1) \oplus ... \oplus Im(h_t) \quad (*).$$

Sia $x \in E$, da (i): $x = h_1(x) + ... + h_t(x)$ quindi $E = Im(h_1) + ... + Im(h_t)$. Mostriamo adesso che la somma è diretta. Sia $y \in Im(h_i)$ con $y = \sum_j z_j$ ($z_j \in Im(h_j), j \neq i$). Abbiamo: $h_i(y) = y$ perchè h_i è un proiettore. D'altra parte $z_j = h_j(y_j)$ perchè $z_j \in Im(h_j)$. Quindi: $y = h_i(y) = h_i(\sum_j h_j(y_j))$. Siccome $h_i \circ h_j = 0$ se $i \neq j$ per (ii), viene $y = 0$. Quindi $(*)$ è dimostrato.
(b) Se $x \in Im(h_i)$ allora $f(x) = c_i x$ $(**)$.
Infatti: se $x = h_i(y)$, $f(x) = (c_1h_1 + ... + c_th_t)(h_i(y))$. Siccome $h_j \circ h_i = 0$ se $j \neq i$ (ii), abbiamo: $f(x) = c_ih_i^2(y) = c_ih_i(y) = c_ix$.
(c) Se B_i è una base di $Im(h_i)$ allora $B = (B_1, ..., B_t)$ è una base di E per (a). Se $e_{i_j} \in B_i$, per (b), $f(e_{i_j}) = c_ie_{i_j}$, quindi B è una base di autovettori di f e f è diagonalizzabile. □

2. Polinomi di interpolazione di Lagrange.

Stando a quanto precede, dobbiamo, sotto opportune ipotesi, ricavare dei proiettori che soddisfano le condizioni $(*)$ e scrivere f come una loro combinazione lineare. Per questo useremo i polinomi di interpolazione di Lagrange.

Siano $c_1, ..., c_t$ degli elementi di k due a due distinti ($c_i \neq c_j$ se $i \neq j$). Si pone

$$P_i(X) = \frac{\Pi_{j \neq i}(X - c_j)}{\Pi_{j \neq i}(c_i - c_j)}, \quad 1 \leq i \leq t.$$

I polinomi P_i hanno grado al più $t - 1$ e verificano: $P_i(c_j) = \delta_{ij}$.

Definizione 2.1: *I polinomi P_i si chiamano i polinomi di interpolazione di Lagrange rispetto agli elementi $c_1, ..., c_t$ di k.*

Proposizione 2.2: *(i) I polinomi $P_1, ..., P_t$ sono una base dello spazio vettoriale $E = k[X]_{\leq (t-1)}$ dei polinomi di grado al più $(t - 1)$.*
(ii) Se P è un polinomio di grado al più $(t - 1)$, allora:

$$P = P(c_1).P_1 + ... + P(c_t).P_t.$$

(iii) In particolare:

$$1 = P_1 + ... + P_t$$
$$X = c_1P_1 + ... + c_tP_t$$
$$.....$$
$$X^k = c_1^k.P_1 + ... + c_t^k.P_t \quad (k \leq t - 1).$$

Dim. (i) e (ii) Sia $\sum_i \lambda_i P_i = 0$. Allora: $0 = \sum_i \lambda_i P_i(c_j) = \lambda_j$. Pertanto $\lambda_i = 0, \forall i$ e i P_i sono linearmente indipendenti.
Sia $P \in E$, consideriamo $Q = P - \sum_i P(c_i)P_i$. Abbiamo $grado(Q) < t$ e $Q(c_j) = P(c_j) - \sum_i P(c_i)P_i(c_j) = 0$. Il polinomio Q, di grado $< t$, ha t radici distinte,

quindi $Q = 0$, cioè: $P = \sum_i P(c_i)P_i$. Questo è (ii) e dimostra che i (P_i) generano E e quindi sono una base di E.
(iii) Segue da (ii). □

Osservazione 2.3: *Il polinomio* $P(x) = \sum_{i=1}^{t} b_i P_i(x)$ *è l'unico polinomio di grado* $\leq t-1$ *il cui grafico passa per i* t *punti* $(c_1, b_1), ..., (c_t, b_t)$ *del piano. (Infatti* $P(c_k) = \sum_i b_i P_i(c_k) = \sum_i b_i \delta_{ik} = b_k$ *e se* F *è un altro polinomio di grado* $\leq t-1$ *con* $F(c_i) = b_i$*, allora* $(P - F)$ *ha grado* $\leq t-1$ *e* t *radici, quindi* $P = F$*.)*

3. Un criterio di diagonalizzazione.

Possiamo adesso dimostrare il risultato principale:

Teorema 3.1: *Sia* $f : E \to E$ *un endomorfismo del* k *spazio vettoriale* E*. Allora* f *è diagonalizzabile se e soltanto se il polinomio minimo* M_f *ha tutte le sue radici in* k *e se* M_f *non ha radici multiple.*

Dim. Abbiamo già visto (III, Corollario 1.5) che se f è diagonalizzabile allora M_f ha tutte le sue radici in k e non ha radici multiple.
Supponiamo quindi $M_f(X) = (X-\lambda_1)...(X-\lambda_r)$, $\lambda_i \neq \lambda_j$ se $i \neq j$. Se $r = 1$ allora $f = \lambda_1 . Id$, possiamo quindi assumere $r > 1$. Siano:

$$P_i(X) = \frac{\Pi_{j \neq i}(X - \lambda_j)}{\Pi_{j \neq i}(\lambda_i - \lambda_j)} \quad 1 \leq i \leq r$$

i polinomi di interpolazione di Lagrange. Poniamo: $h_i := P_i(f)$.
Siccome: $1 = P_1 + ... + P_r$ (Proposizione 2.2,(iii)), viene: $Id = h_1 + ... + h_r$, (+).
Siccome M_f ha solo radici semplici, $M_f | P_i P_j$ se $i \neq j$. Quindi, visto che $M_f(f) = 0$, abbiamo: $h_i \circ h_j = 0$ se $i \neq j$ (++).
Siccome: $X = \lambda_1 P_1 + ... + \lambda_r P_r$ (Proposizione 2.2, (iii)), abbiamo: $f = \lambda_1 h_1 + ... + \lambda_r h_r$.
Mostriamo adesso che gli h_i sono dei proiettori: $h_i = h_i \circ Id = h_i \circ (h_1 + ... + h_r) = h_i^2$.
Si conclude con il Teorema 1.5. □

Osservazione 3.2: *(i) Se* k *è algebricamente chiuso abbiamo:* f *è diagonalizzabile* $\Leftrightarrow$ M_f *non ha radici multiple.*
(ii) L'ipotesi M_f *non ha radici multiple è essenziale per affermare che* $M_f | P_i P_j$ *e quindi per avere* $h_i \circ h_j = 0$ *se* $i \neq j$*; questa condizione a sua volta è essenziale per affermare che gli* h_i *sono dei proiettori (e che* E *è somma diretta degli autospazi, cf dimostrazione del Teorema 1.5).*
(iii) Osserviamo inoltre che, nella dimostrazione del Teorema 3.1, $h_i \neq 0$ *perchè altrimenti si avrebbe* $M_f | P_i$ *(assurdo per ragioni di grado).*

4. Esempi ed esercizi.

Esercizio 4.1: *Sia $P(X) \in k[X]$, allora: P ha una radice multipla $\Leftrightarrow$ P e P' (la derivata di P) hanno una radice in comune.*

Esercizio 4.2: *Sia $A \in \mathcal{M}_n(k)$ tale che $A^t = I_n$ per un qualche $t \geq 1$. Se k è algebricamente chiuso, allora A è diagonalizzabile.*

Esercizio 4.3: *Sia $A \in \mathcal{M}_n(k)$ tale che: $A^2 = -I_n$.*
(i) Se k è algebricamente chiuso, A è diagonalizzabile.
(ii) Se $k = \mathbb{R}$, A non è diagonalizzabile.

Esercizio 4.4: *Rifare III. Esercizio 3.4 usando il Teorema 3.1.*

Bibliografia

[E] Ellia, Ph. *Appunti di Geometria I*. Pitagora Ed.