Soluzioni di equazioni polinomiali zero dimensionali

Giorgio Ottaviani

versione 20.5.2015

Indice

1 Richiami sulla diagonalizzazione

Ricordiamo, dal corso di Geometria I, le definizioni e i concetti principali sulle matrici diagonalizzabili. Una matrice A $n \times n$ è diagonalizzabile quando ammette una base di autovettori. Con una formulazione equivalente, A è diagonalizzabile se e solo se esiste G invertibile tale che $G^{-1}AG$ è diagonale. Infatti le colonne di G formano la base richiesta di autovettori.

Una matrice a coefficienti reali è diagonalizzabile (sui reali) se e solo se tutti gli autovalori sono reali (cioè tutte le radici del polinomio caratteristico sono reali) e la molteplicità algebrica di ogni autovalore è uguale alla molteplicità geometrica.

Una matrice a coefficienti complessi è diagonalizzabile se e solo se la molteplicità algebrica di ogni autovalore è uguale alla molteplicità geometrica.

Siccome gli autovalori sono calcolabili solo in modo approssimato, questo algoritmo non è effettivo in aritmetica esatta. Ci proponiamo in questa sezione di trovare un algoritmo effettivo, per verificare la diagonalizzabilità di una matrice, basato sulla nozione di polinomio minimo.

Una matrice con autovalori distinti è diagonalizzabile sui complessi, le matrici (complesse) con autovalori distinti formano un aperto (denso) nell'insieme di tutte le matrici complesse.

Per ogni polinomio $p(x) \in K[x]$ (siamo interessati ai casi $K = \mathbb{R}$ oppure $K = \mathbb{C}$) e ogni matrice A a coefficienti in K ha senso considerare la matrice $p(A)$. Se $p(x) = \sum_{i=0}^{d} a_i x^i$ allora $p(A) = \sum_{i=0}^{d} a_i A^i$, con la convenzione che A^0 è la matrice identità.

Lemma 1.1 *Per ogni matrice invertibile G vale $p(G^{-1}AG) = G^{-1}p(A)G$.*

Dimostrazione

$$p(G^{-1}AG) = \sum_{i=0}^{d} a_i (G^{-1}AG)^i = \sum_{i=0}^{d} a_i G^{-1}A^iG = G^{-1}(\sum_{i=0}^{d} a_i A^i)G = G^{-1}p(A)G$$

□

Notiamo che per ogni polinomio p, q vale $p(A)q(A) = q(A)p(A) = pq(A)$.

Teorema 1.2 (Hamilton-Cayley) *Ogni matrice a coefficienti reali (o complessi) è radice del polinomio caratteristico $p_A(t) = \det(A - tI)$, cioè $p_A(A) = 0$.*

Dimostrazione Conviene dimostrare il risultato direttamente per le matrici complesse. Il risultato è immediato per le matrici diagonali, che hanno sulla diagonale gli autovalori, proprio perché gli autovalori sono le radici del polinomio caratteristico. Sia A diagonalizzabile. Allora esiste G tale che $G^{-1}AG = D$ è diagonale. Sappiamo che $p_D(t) = p_A(t)$. Allora dal lemma 1.1 $p_A(A) = p_A(GDG^{-1}) = Gp_A(D)G^{-1} = Gp_D(D)G^{-1} = 0$ e quindi il risultato è vero per le matrici diagonalizzabili.

Adesso la funzione $p_A(A)$ ha coefficienti che sono polinomi nei coefficienti di A, che valgono zero su un aperto denso , pertanto tali polinomi sono identicamente nulli e quindi $p_A(A) = 0 \ \forall A$. □

Osservazione Il procedimento di limite utilizzato nella dimostrazione del Teorema di Hamilton-Cayley può essere omesso, dimostrando a priori l'esistenza di una base che triangolarizza A.

Fissata A, consideriamo il morfismo

$$K[x] \to M_n$$

$$p(x) \mapsto p(A)$$

Il nucleo di questo morfismo è un ideale di $K[x]$, il quale, come tutti gli ideali di $K[x]$, è principale. Il generatore di questo ideale (normalizzato in modo che sia monico) si dice il *polinomio minimo* di A.

Il teorema di Hamilton-Cayley afferma che il polinomio caratteristico appartiene al nucleo appena definito, quindi il polinomio minimo divide il polinomio caratteristico.

Notiamo anche che A e $G^{-1}AG$ hanno lo stesso polinomio minimo.

Lemma 1.3 *Se v è autovettore di A con autovalore λ e $p(x) \in K[x]$ è un polinomio allora*

$$p(A)v = p(\lambda)v$$

cioè v è autovettore di $p(A)$ con autovalore $p(\lambda)$.

Dimostrazione Da $Av = \lambda v$ segue $A^k v = \lambda^k v$ per ogni $k \geq 0$. Quindi, se $p(x) = \sum_{i=0}^{d} a_i x^i$ abbiamo $p(A)v = \sum_{i=0}^{d} a_i A^i v = \sum_{i=0}^{d} a_i \lambda^i v = p(\lambda)v$ □

Proposizione 1.4 *Ogni autovalore di A è radice del polinomio minimo di A.*

Dimostrazione Sia p il polinomio minimo di A, pertanto $p(A) = 0$. Sia λ un autovalore di A con autovettore v. Dal lemma precedente $0 = p(A)v = p(\lambda)v$ da cui $p(\lambda) = 0$. □

Esercizio 1.5 *Fissata A, per ogni $v \in K^n$ definiamo p_v come il generatore (monico) dell'ideale nucleo del morfismo*

$$K[x] \to K^n$$

$$p(x) \mapsto p(A)v$$

(i) Provare che p_v divide il polinomio minimo di A.

(ii) Provare che per ogni base $v_1, \ldots, v_n$, il minimo comune multiplo di p_{v_i} è il polinomio minimo di A.

Il polinomio minimo è ben definito per ogni applicazione lineare $A\colon K^n \to K^n$, nella trattazione seguente adotteremo questo punto di vista. Spesso denotiamo $V = K^n$.

Esercizio 1.6 *(i) Fissata A, sia $W \subset V$ un sottospazio A-invariante. Provare che il polinomio minimo di A ristretta a W divide il polinomio minimo di A.*

(ii) Sia $W_1 \oplus W_2 = V$ una decomposizione in due sottospazi A-invarianti. Per $i = 1, 2$, sia p_i il polinomio minimo di A ristretta a W_i. Provare che il polinomio minimo di A coincide con il minimo comune multiplo tra p_1 e p_2. Generalizzare l'enunciato a una somma di un numero finito di sottospazi.

Se $A\colon V \to V$ è un endomorfismo, e λ è un autovalore di A, abbiamo la catena di inclusioni

$$\ldots \subseteq \ker(A - \lambda I)^n \subseteq \ker(A - \lambda I)^{n+1} \subseteq \ldots$$

Per motivi dimensionali, la catena precedente diventa stazionaria per n sufficientemente grande. Inoltre si può verificare (esercizio) che se $\ker(A-\lambda I)^n = \ker(A-\lambda I)^{n+1}$ allora $\ker(A - \lambda I)^n = \ker(A - \lambda I)^m$ per ogni $m \geq n$.

Definizione 1.7 *Denotiamo $V_\lambda^\infty = \cup_n \ker(A - \lambda I)^n$, che è uguale a $\ker(A - \lambda I)^m$, per qualche m sufficientemente grande. Vedremo che si può sempre sempre scegliere $m \leq$ molteplicità algebrica di λ.*

Prima della Prop. 1.9 premettiamo il seguente (provvisorio)

Lemma 1.8 $\dim V_\lambda^\infty \leq$ *molteplicità algebrica di* λ.

Dimostrazione V_λ^∞ è un sottospazio A-invariante. Il polinomio minimo di A ristretto a V_λ^∞ ha la forma $(x-\lambda)^m$ per qualche m, pertanto per la Proposizione 1.4 il polinomio caratteristico di A ristretto a V_λ^∞ è uguale a $(x-\lambda)^d$ dove $d = \dim V_\lambda^\infty$. Segue che il polinomio caratteristico di A è divisibile per $(x-\lambda)^d$ da cui la tesi. □

Proposizione 1.9 *Sia* $A\colon V \to V$ *è un endomorfismo. Se* $K = \mathbb{C}$ *abbiamo* $V = \oplus_\lambda V_\lambda^\infty$ *dove la somma è estesa a tutti gli autovalori. Inoltre* $\dim V_\lambda^\infty$ *è uguale alla molteplicità algebrica di* λ. V_λ^∞ *si dice autospazio generalizzato.*

Dimostrazione Sia $f(x) = \prod_{i=1}^k (x-\lambda_i)^{n_i}$ il polinomio minimo di A, fattorizzato nella chiusura algebrica. Poniamo $p_j(x) = \prod_{i\neq j}(x-\lambda_i)^{n_i}$, che sono primi tra loro. Pertanto esistono $a_j(x)$ tali che $\sum_{j=1}^k a_j(x)p_j(x) = 1$. Applicando i polinomi alla matrice A si ottiene $\sum_{j=1}^k a_j(A)p_j(A) = I$. Notiamo che $\operatorname{Im} p_j(A) \subset V_{\lambda_j}^\infty$ perché $(A-\lambda_j I)^{n_j} p_j(A) = f(A) = 0$. Per ogni vettore $v \in V$ abbiamo $v = \sum_{j=1}^k p_j(A)a_j(A)v$, dove $p_j(A)a_j(A)v \in V_{\lambda_j}^\infty$ per quanto appena visto. Segue che V si decompone nella somma dell'enunciato. Per provare che la somma è diretta, per il Lemma 1.8, la dimensione della somma dei $V_{\lambda_j}^\infty$ è minore o uguale della somma delle molteplicità algebriche, che è pari a $\dim V$, la dimensione dello spazio ambiente. Pertanto nel Lemma 1.8 deve valere l'uguaglianza. Inoltre la dimensione della somma dei $V_{\lambda_j}^\infty$ è uguale alla somma delle dimensioni dei $V_{\lambda_j}^\infty$, e questo accade precisamente quando la somma è diretta. □

Corollario 1.10 *Sia* n_i *il più piccolo intero tale che* $\ker(A-\lambda_i I)^{n_i} = \ker(A-\lambda_i I)^{n_i+1}$, *allora il polinomio minimo di* A *è* $\prod_{i=1}^k (x-\lambda_i)^{n_i}$

Dimostrazione E' sufficiente provare che il polinomio minimo di A ristretto a ogni V_λ^∞ è uguale a $(x-\lambda_i)^{n_i}$, che è evidente dalla definizione. □

Teorema 1.11 *Una matrice è diagonalizzabile su* K *se e solo se il suo polinomio minimo ha tutte le radici in* K *di molteplicità uno (si spezza come prodotto di fattori lineari distinti).*

Dimostrazione Se A è diagonalizzabile ha lo stesso polinomio minimo di una matrice diagonale D. Se D ha sulla diagonale gli elementi distinti $d_1, \dots d_k$ (nel senso che alcuni di questi valori possono apparire più volte sulla diagonale) allora $\prod_{i=1}^k (x-d_i)$ è il polinomio minimo di D. Viceversa, se il polinomio minimo di A ha tutte le radici di molteplicità uno allora dal Corollario 1.10 segue che $\ker(A-\lambda_i I) = V_{\lambda_i}^\infty$ e dalla Prop. 1.9 segue che A ha una base di autovettori. □

Proposizione 1.12 *Se* λ_i *per* $i = 1, \dots, k$ *sono tutti gli autovalori (complessi) di una matrice* A, *ripetuti con la loro molteplicità* n_i, *allora, per ogni polinomio* $h(x)$, $h(\lambda_i)$ *per* $i = 1, \dots, k$ *sono tutti gli autovalori (complessi) della matrice* $h(A)$, *ripetuti* n_i *volte.*

Dimostrazione La decomposizione $V = \oplus_{\lambda_i} V^\infty_{\lambda_i}$ della Prop. 1.9 è $h(A)$-invariante. Sia n_i la dimensione di $V^\infty_{\lambda_i}$.

Se $v_i \in V^\infty_{\lambda_i}$ abbiamo $(A - \lambda_i)^{n_i} v_i = 0$. Siccome $(x - \lambda_i)$ divide $h(x) - h(\lambda_i)$, segue $(h(A) - h(\lambda_i)I)^{n_i} v_i = 0$, da cui il polinomio minimo di $h(A)$ ristretto a $V^\infty_{\lambda_i}$ ha la forma $(x - h(\lambda_i))^m$ per qualche m, e per la prop. 1.4 $h(\lambda_i)$ è l'unico autovalore di $h(A)$ ristretto a $V^\infty_{\lambda_i}$. Segue che il polinomio caratteristico di $h(A)$ ristretto a $V^\infty_{\lambda_i}$ è $(x - h(\lambda_i))^{d_i}$. □

Algoritmo per il calcolo del polinomio minimo Si scelga il primo valore d tale che $A^0, A^1, A^2, \ldots, A^d$ sono dipendenti. Allora esiste un unico vettore (a meno di costanti) $(a_0, \ldots, a_d)$ tale che $\sum_{i=0}^d a_i A^i = 0$ (se ci fossero due tali vettori, entrambi dovrebbero avere $a_d \neq 0$, ed una loro combinazione lineare darebbe una relazione di dipendenza tra $A^0, A^1, \ldots, A^{d-1}$).

Il polinomio $p(t) = \sum_{i=0}^d \frac{a_i}{a_d} t^i$ è il polinomio minimo di A.

Algoritmo per la verifica della diagonalizzabilità di una matrice, sui complessi Sia data una matrice A. Si costruisce p polinomio minimo di A con l'algoritmo precedente.

Si calcola MCD (p, p') con l'algoritmo euclideo, A è diagonalizzabile se e solo se MCD $(p, p') = 1$.

Spiegazione: p ha fattori ripetuti se e solo se MCD $(p, p') = 1$ e questo equivale alla diagonalizzabilità per il teorema precedente.

Esercizio 1.13 *Sia A matrice $n \times n$ il cui polinomio minimo ha grado n e che ha tutti gli autovalori in K (una tale A si dice regolare). Provare che esiste $v \in K^n$ tale che $v, Av, A^2v, \ldots, A^{n-1}v$ sono indipendenti.*

Suggerimento: provare prima il caso in cui il polinomio minimo ha la forma $(x - \lambda)^n$. In generale, se $p = \prod p_i^{n_i}$ con p_i primi tra loro, si può provare con una tecnica simile a quella usata nella dimostrazione del Teorema 1.11 che $K^n = \oplus_i \ker p_i(A)$ e ci si può ricondurre al caso precedente.

Esercizio 1.14 *Sia A una matrice regolare (definita nell'esercizio precedente). Provare che il centralizzante di A ha dimensione n ed è generato da $\{A^0, A^1, \ldots, A^{n-1}\}$.*

Suggerimento: Il polinomio minimo di A è dato da $p_1^{n_1} \ldots p_k^{n_k}$ con p_i polinomi di grado uno distinti. Posto $V_i = \ker p_i(A)$ si può provare (vedi dimostrazione del Teorema 1.11) che $K^n = \oplus_i V_i$. Se B commuta con A allora ogni V_i è B-invariante. Quindi per dimostrare l'asserto ci si può ricondurre al caso in cui il polinomio minimo di A ha la forma p^n con p di grado uno. In questo caso l'esercizio precedente mostra che abbiamo un unico blocco di Jordan. La conoscenza della forma di Jordan è utile ma non indispensabile per la comprensione di queste note. Il lettore che conosce la forma di Jordan può osservare che le matrici compagne, che introdurremo nella sezione 2, sono regolari e quindi hanno un unico blocco di Jordan per ogni autovalore. Applicando l'esercizio precedente ad $A - I$, otteniamo che esiste $v \in K^n$ tale che $v, (A - I)v, (A - I)^2 v, \ldots, (A - I)^{n-1} v$ sono indipendenti. Un calcolo esplicito, scrivendo la matrice di A rispetto a questa base, mostra che la condizione $AB = BA$ impone $n^2 - n$ condizioni indipendenti su B. Pertanto lo spazio vettoriale generato da $\{A^0, A^1, \ldots, A^{n-1}\}$, che è sempre contenuto

nel centralizzante di A, deve coincidere col centralizzante di A perché ha la stessa dimensione n.

Gli ultimi tre risultati di questa sezione saranno utilizzati a partire dalla sezione 5 e possono essere omessi da chi è interessato soltanto ai polinomi in una variabile.

Teorema 1.15 *Diagonalizzazione simultanea*

(i) Siano $A, B\colon V \to V$ due endomorfismi diagonalizzabili. Allora esiste una base di autovettori comuni a A e B se e solo se $AB = BA$.

(ii) Siano $A_1, \dots A_n\colon V \to V$ endomorfismi diagonalizzabili. Allora esiste una base di autovettori comuni a A_i per $i = 1, \dots, n$ se e solo se $A_iA_j = A_jA_i$ per ogni i, j.

Dimostrazione (i) Sia $\{v_1, \dots v_n\}$ una base comune di autovettori. Allora $Av_i = \lambda_i v_i$ e $Bv_i = \mu_i v_i$. Quindi $ABv_i = \lambda_i\mu_i v_i = BAv_i$. Quindi AB e BA assumono gli stessi valori su una base di V e pertanto sono uguali. Viceversa siano $V_{\lambda_i} = \ker(A - \lambda_i I)$ gli autospazi di A. Se $v \in V_{\lambda_i}$ allora $A(Bv) = B(Av) = B(\lambda_i v) = \lambda_i(Bv)$, da cui $Bv \in V_{\lambda_i}$. Quindi $B(V_{\lambda_i}) \subseteq V_{\lambda_i}$, cioè gli autospazi di A sono B-invarianti. Pertanto il polinomio caratteristico di B si fattorizza come prodotto dei polinomi caratteristici di $B_{|V_{\lambda_i}}$ e quindi gli autovalori di $B_{|V_{\lambda_i}}$ sono tutti in K. Inoltre il polinomio minimo di $B_{|V_{\lambda_i}}$ divide il polinomio minimo di B, siccome quest'ultimo per ipotesi non ha fattori ripetuti, neanche il polinomio minimo di $B_{|V_{\lambda_i}}$ ha fattori ripetuti e pertanto $B_{|V_{\lambda_i}}$ è diagonalizzabile. Mettendo insieme le basi di autovettori per $B_{|V_{\lambda_i}}$, si ottiene una base di autovettori comuni a A e B.

(ii) Se abbiamo una base comune di autovettori l'argomento è lo stesso del punto (i). Viceversa, possiamo ragionare per induzione su n. Se V_i sono gli autospazi di A_n, abbiamo come nel punto (i) che $A_j(V_i) \subseteq V_i$ per ogni i, j. Il ragionamento del punto (i) mostra che, per ogni i, gli $n-1$ endomorfismi $A_{j|V_i}$ per $j = 1, \dots, n-1$ commutano a due a due e sono diagonalizzabili e quindi hanno una base di autovettori comuni su V_i. Mettendo insieme queste basi di autovettori comuni, si ottiene una base di autovettori comuni a tutti gli A_i.

□

Proposizione 1.16 *Siano A_i per $i = 1, \dots, n$ endomorfismi tali che $A_iA_j = A_jA_i$ per ogni i, j. Sia A_1 diagonalizzabile con autovalori distinti. Allora ogni A_i è diagonalizzabile e tutti gli A_i hanno una base comune di autovettori.*

Dimostrazione Per ipotesi, gli autospazi di A_1 $V_{\lambda_i} = \ker(A_1 - \lambda_i I)$ hanno dimensione uno. Come nella dimostrazione precedente abbiamo $A_j(V_{\lambda_i}) \subset V_{\lambda_i}$, e questo vuol dire che i generatori di V_{λ_i} sono autovettori anche per A_j. Pertanto A_j è diagonalizzabile e gli autovettori trovati sono comuni a tutti gli A_j. □

Proposizione 1.17 *Siano A_i per $i = 1, \dots, n$ endomorfismi tali che $A_iA_j = A_jA_i$ per ogni i, j. Esiste un autovettore (complesso) comune a A_i.*

Dimostrazione Sia λ_1 un autovalore di A_1. Allora $V_1 = \ker(A_1 - \lambda_1 I)$ è A_2-invariante, pertanto A_2 ha un autovettore su V_1 con autovalore λ_2 e $V_2 = \bigcap_{i=1}^{2} \ker(A_i - \lambda_i I) \neq 0$.

Analogamente, V_2 è A_3-invariante, per cui esiste λ_3 tale che $V_3 = \bigcap_{i=1}^{3} \ker(A_i - \lambda_i I) \neq 0$. Proseguendo in questo modo, si trova $V_n = \bigcap_{i=1}^{n} \ker(A_i - \lambda_i I) \neq 0$, e ogni vettore non nullo in V_n è un autovettore comune a A_i. □

Dalla Proposizionme precedente segue che un sottospazio di endomorfismi che commutano può essere ridotto (simultaneamente) a forma triangolare.

2 Polinomi in una variabile e matrici compagne

In questa sezione ricordiamo come il calcolo delle radici di un polinomio può essere ricondotto al calcolo degli autovalori di una matrice, detta matrice compagna. Questo è interessante perché esistono molti algoritmi per calcolare numericamente gli autovalori di una matrice.

La struttura algebrica che nasce da questa problema ha delle generalizzazioni nel caso di più variabili, che studieremo successivamente a partire dalla sezione 4.

Sia $f(x) = \sum_{i=0}^{d} a_i x^i$ un polinomio di grado d, che possiamo supporre monico, cioè $a_d = 1$. Chiamiamo R l'anello quoziente $K[x]/(f(x))$ che ha dimensione d ed è generato dalle classi $[1], [x], \ldots, [x^{d-1}]$. Infatti x^d può essere scritto come combinazione delle potenze precedenti modulo f, $[x^d] = -\sum_{i=0}^{d-1} a_i [x^i]$ e così anche le potenze successive, ad esempio

$[x^{d+1}] = -\sum_{i=0}^{d-1} a_i [x^{i+1}] = -\sum_{i=0}^{d-2} a_i [x^{i+1}] + a_{d-1} \sum_{i=0}^{d-1} a_i [x^i]$

Definizione 2.1 *La moltiplicazione per x induce un'applicazione lineare*

$$R \xrightarrow{M_x} R$$

$$[g] \mapsto [gx]$$

la matrice di M_x rispetto alla base $[1], [x], \ldots, [x^{d-1}]$ si dice matrice compagna di $f(x)$.

Analogamente la moltiplicazione per $h(x)$ induce un'applicazione lineare

$$R \xrightarrow{M_{h(x)}} R$$

$$[g] \mapsto [gh]$$

Proposizione 2.2 *$\forall\ h(x), k(x) \in K[x]$ vale che*

(i) $M_{h(x)} + M_{k(x)} = M_{h(x)+k(x)}$

(ii) $M_{ah(x)} = aM_{h(x)} \quad \forall a \in K$

(iii) $M_{h(x)} \cdot M_{k(x)} = M_{k(x)} \cdot M_{h(x)} = M_{h(x)k(x)}$

(iv) $M_{h(k(x))} = h(M_{k(x)})$, *in particolare* $M_{h(x)} = h(M_x)$.

Dimostrazione (i), (ii) e (iii) sono immediate dalla definizione. Notiamo che la commutatività in (iii) segue dalla commutatività dei polinomi. Per provare (iv) supponiamo dapprima $h = x^i$. Allora $M_{(k(x))^i} = \left(M_{k(x)}\right)^i$ come diretta applicazione di (iii).

In generale, se $h(x) = \sum a_i x^i$ allora utilizzando anche (i)-(iii)

$$M_{h(k(x))} = M_{\sum a_i (k(x))^i} = \sum M_{a_i (k(x))^i} = \sum a_i M_{(k(x))^i} = \sum a_i \left(M_{(k(x))}\right)^i = h(M_{k(x)})$$

□

Proposizione 2.3 *La matrice compagna di $f(x)$ è*

$$\begin{bmatrix} 0 & 0 & \dots & -a_0 \\ 1 & 0 & \dots & -a_1 \\ 0 & 1 & \dots & -a_2 \\ \vdots & & & \vdots \\ 0 & 0 & 1 & -a_{d-1} \end{bmatrix}$$

Dimostrazione É un calcolo immediato dalle espressioni precedenti. □

Teorema 2.4 *(i) A meno di scalari moltiplicativi, $f(x)$ è il polinomio minimo di M_x.*
(ii) A meno di scalari moltiplicativi, $f(x)$ è il polinomio caratteristico di M_x.
(iii) M_x è diagonalizzabile se e solo se $f(x)$ ha radici distinte.
(iv) $f(x_0) = 0$ se e solo se x_0 è un autovalore di M_x.
(v) $tr(M_{h(x)}) = \sum_{i=1}^n h(\lambda_i)$ dove $\lambda_1, \dots, \lambda_n$ sono le radici di $f(x)$
(vi) $\det(M_{h(x)}) = \prod_{i=1}^n h(\lambda_i)$, in particolare $\det(M_{h(x)})$ si annulla se e solo se f e h hanno una radice in comune, e costituisce un metodo alternativo (in generale più economico) di calcolo del risultante $Res(f, h)$.

Dimostrazione Per ogni polinomio $p(x)$ abbiamo che $p(M_x) = M_{p(x)}$, che è nulla esattamente quando $p(x)$ è un multiplo di $f(x)$ (basta applicare $M_{p(x)}$ alla classe di 1). Pertanto $f(x)$ è il polinomio minimo di M_x, e siccome divide il polinomio caratteristico, che ha lo stesso grado, segue che polinomio minimo e polinomio caratteristico coincidono, a meno del segno. (iii) è conseguenza di (i) e del Teorema 1.11. (iv) è immediata da (ii). (v) e (vi) seguono dalla Prop. 1.12.

□

Osservazione L'importanza del Teorema 2.4 risiede nel fatto che otteniamo le radici del polinomio f dal calcolo degli autovalori della matrice compagna, che può essere effettuato ad esempio col metodo QR.

3 La forma di Killing e il numero delle radici reali di un polinomio

3.1 Il teorema cinese dei resti e l'interpolazione polinomiale

Consideriamo la fattorizzazione di un intero $n = \prod_{i=1}^k p_i^{m_i}$ con p_i primi distinti. La forma classica del teorema cinese dei resti, su $\mathbb{Z}$, descrive l'isomorfismo

$$q \colon \mathbb{Z}/n \to \oplus_{i=1}^k \mathbb{Z}/(p_i^{m_i})$$

L'estensione naturale ai polinomi è data dal seguente

Teorema 3.1 *Sia $f(x) = \prod_{i=1}^{k} f_i$ con f_i primi tra loro a due a due. Sia $g_i = \frac{f}{f_i}$, siano b_i, r_i tali che $b_i g_i + r_i f_i = 1$ (ricavabili con l'algoritmo euclideo, vedi l'osservazione 3.2). L'applicazione naturale*

$$q\colon K[x]/(f(x)) \to \oplus_{i=1}^{k} K[x]/(f_i)$$

definita da $q(h) = (h, \ldots, h)$ è un isomorfismo con inversa data da

$$\tilde{q}\colon \oplus_{i=1}^{k} K[x]/(f_i) \to K[x]/(f(x))$$

definita da $\tilde{q}(a_1, \ldots, a_k) = \sum_{i=1}^{k} b_i g_i a_i$.

Dimostrazione $b_i g_i$ modulo f_j è uguale a 1 se $i = j$, è uguale a 0 se $i \neq j$. Pertanto $\sum_{i=1}^{k} b_i g_i$ modulo f_j è uguale a 1 $\forall j$, e quindi $\sum_{i=1}^{k} b_i g_i$ è uguale a 1 modulo f, mentre $\sum_{i=1}^{k} b_i g_i a_i$ modulo f_j è uguale ad a_j $\forall j$. Da queste condizioni si ricava che, componendo $q\tilde{q}$ e $\tilde{q}q$, si ottiene in entrambi i casi l'identità. □

Osservazione 3.2 *Con Macaulay2, b_i puo' essere trovato tramite il comando*

```
quotientRemainder(matrix{{1}},matrix{{g_i, f_i}})
```

Notiamo che per calcolare questi quozienti è necessario trovare la base di Groebner dell'ideale (g_i, f_i), che è uguale a 1, e che essenzialmente questo calcolo corrisponde all'algoritmo euclideo del calcolo del MCD tra g_i e f_i.

La seguente applicazione all'interpolazione polinomiale è interessante.

Corollario 3.3 (Interpolazione di Hermite) *Siano $c_1, \ldots, c_k \in \mathbb{R}$ punti distinti. Assegniamo, rispetto a ogni punto c_i, uno sviluppo di Taylor $a_i(x)$ di grado $< m_i$, questo equivale a dare $a_i(x) = \sum_{j=0}^{m_i-1} \frac{\alpha_{j,i}}{j!}(x - c_i)^j$ dove $\alpha_{j,i} = a_i^{(j)}(c_i)$ è la derivata j-esima di a_i valutata in c_i.*

Esiste un polinomio $H(x)$ di grado $< \sum_{i=1}^{k} m_i$, tale che $H^{(j)}(c_i) = \alpha_{j,i}$ per $0 \leq j < m_i$, definito da $H = \sum_{i=1}^{k} b_i g_i a_i$ modulo $f(x) = \prod_{i=1}^{k}(x - c_i)^{m_i}$ (cioè $H\%ideal(f)$) dove , posto $f_i = (x - c_i)^{m_i}$, $g_i = \prod_{j \neq i}(x - c_j)^{m_j}$, b_i viene ottenuto dall'algoritmo euclideo come nel Teorema 3.1, cioè dalla condizione $b_i g_i + r_i f_i = 1$.

Dimostrazione La condizione $H(x) = a_i(x)$ modulo $(x - c_i)^{m_i}$, equivale a $H^{(j)}(c_i) = a_i^{(j)}(c_i)$. Il risultato segue allora dal Teor. 3.1. □

Osservazione 3.4 *Il polinomio $H(x) = \sum_{i=1}^{k} b_i(x) g_i(x) a_i(x)$ ha le derivate richieste nei punti c_i. Considerare la sua forma normale rispetto a f serve soltanto per abbassare il grado fino a renderlo $< \sum_{i=1}^{k} m_i = \deg f$.*

Per $m_i = 1$ abbiamo come caso particolare l' interpolazione di Lagrange, dove si ricava un polinomio di grado $k - 1$ cha assume k valori assegnati.

Corollario 3.5 (Interpolazione di Lagrange) *Siano $c_1, \ldots, c_k \in \mathbb{R}$ punti distinti. Assegniamo, rispetto a ogni punto c_i, un valore $a_i \in K$.*

Esiste un polinomio $H(x)$ di grado $< k$, tale che $H(c_i) = a_i$, definito da $H = \sum_{i=1}^{k} b_i g_i a_i$ dove $g_i = \prod_{j \neq i}(x - c_j)$, $b_i = \frac{1}{\prod_{j \neq i}(c_i - c_j)}$.

Dimostrazione Rispetto all'interpolazione di Hermite, si può scegliere b_i come lo scalare nell'enunciato, infatti $b_i g_i = \frac{\prod_{j\neq i}(x-c_j)}{\prod_{j\neq i}(c_i-c_j)}$ vale 1 su c_i e vale 0 su c_j per $j \neq i$. Il grado di $\sum_{i=1}^{k} b_i g_i a_i$ risulta immediatamente $< k$. □

L'interpolazione di Hermite è utile nello sviluppo di una funzione razionale in fratti semplici, necessaria per la sua integrazione. Infatti, con le notazioni precedenti, dopo aver trovato un'espressione $1 = \sum_{i=1}^{k} b_i g_i$, dividendo per $f(x)$ si ottiene $\frac{1}{f(x)} = \sum_{i=1}^{k} \frac{b_i(x)}{(x-c_i)^{m_i}}$ da cui per un polinomio g

$$\frac{g(x)}{f(x)} = \sum_{i=1}^{k} \frac{g(x)b_i(x)}{(x-c_i)^{m_i}}$$

Calcolando lo sviluppo di Taylor dei numeratori $g(x)b_i(x)$ rispetto al punto c_i, si ottiene l'espressione che puo' essere facilmente integrata.

Esercizio 3.6 *Si trovi, con l'ausilio di Macaulay2, un polinomio di grado* 8 *tale che* $H(2) = \alpha$, $H'(2) = \beta$, $H(3) = 5$, $H'(3) = 7$, $H''(3) = 11$, $H(4) = 13$, $H'(4) = 17$, $H''(4) = 19$, $H'''(4) = 21$, *al variare di* α, β.

3.2 La forma traccia di Killing

Sia $f(x) \in K[x]$ un polinomio. Per ogni $a, b \in K[x]/(f(x))$ è definita la forma (bilineare) B di Killing

$$\begin{array}{ccccc} K[x]/(f(x)) & \times & K[x]/(f(x)) & \longrightarrow & K \\ a & & b & \mapsto & B(a,b) := tr(M_{ab}) \end{array}$$

dove M_{ab} è l'applicazione lineare $K[x]/(f(x)) \xrightarrow{M_{ab}} K[x]/(f(x))$ data dalla moltiplicazione per ab. Notiamo che $M_a M_b = M_b M_a = M_{ab}$.

É associata la forma quadratica $a \mapsto B(a,a) = tr(M_{a^2})$.

Proposizione 3.7 *Sia* $n = \deg f$. *La matrice della forma di Killing nella base* $\{1, x, \dots x^{n-1}\}$ *ha la forma*

$$\begin{bmatrix} s_0 & s_1 & s_2 & \dots & s_{n-1} \\ s_1 & s_2 & s_3 & \dots & s_n \\ s_2 & s_3 & \dots & & s_{n+1} \\ \vdots & & & & \vdots \\ s_{n-1} & s_n & \dots & & s_{2n-2} \end{bmatrix}$$

dove s_i *è la* i*-esima somma di potenze nelle radici* $x_1, \dots x_n$ *di* f, *cioè* $s_i = \sum_{j=1}^{n} x_j^i$ *e prende il nome di Bezoutiante. Numerando le righe e le colonne da* 0 *a* $n-1$, *il coefficiente di posto* (i,j) *è* s_{i+j}.

Dimostrazione Numeriamo le righe e le colonne da 0 a $n-1$. Allora il coefficiente di posto (i,j) è $tr(M_{x^{i+j}})$. Per il Teorema 2.4 (iv) gli autovalori di M_x sono le radici $x_1, \dots, x_n$ di $f(x)$ e quindi gli autovalori di $M_{x^{i+j}} = (M_x)^{i+j}$ sono $x_1^{i+j}, \dots, x_n^{i+j}$ (per il Teor. 2.4 (v)) e la loro somma coincide con s_{i+j}. □

Proposizione 3.8 *Il determinante del Bezoutiante è il discriminante di f, a meno di costanti moltiplicative.*

Dimostrazione Definiamo la matrice di Vandermonde

$$V = \begin{bmatrix} 1 & x_1 & \dots & x_1^{n-1} \\ 1 & x_2 & \dots & x_2^{n-1} \\ \vdots & & & \vdots \\ 1 & x_n & \dots & x_n^{n-1} \end{bmatrix}$$

Un facile calcolo mostra che $V^t \cdot V$ coincide con la matrice Bezoutiante. Inoltre è noto che $\det V = \prod_{i<j}(x_i - x_j)$.

Segue che $\det(V^t \cdot V) = (\det V)^2 = \prod_{i<j}(x_i - x_j)^2$ che coincide con il discriminante di f a meno di costanti moltiplicative. □

Notiamo dalla dimostrazione precedente che se f ha n radici reali e distinte allora il suo Bezoutiante ha la forma $V^t \cdot V$ con V nondegenere e quindi è definito positivo.

Il Teorema di Sylvester 3.10 generalizza questa osservazione ed è il risultato fondamentale di queste note. Lega il numero di radici reali di $f(x)$ con la segnatura della forma di Killing associata a $f(x)$. Prima di enunciarlo, abbiamo

Teorema 3.9 *I sottospazi di $K[x]/(f(x))$ isomorfi a $K[x]/(p_i(x)^{n_i})$, identificati tramite l'isomorfismo dato dal Teorema cinese 3.1, sono ortogonali a due a due rispetto alla forma di Killing B.*

Dimostrazione Due sottospazi sono generati rispettivamente da $b_i(x)g_i(x)$ e da $b_j(x)g_j(x)$ Abbiamo che $f(x)$ divide $g_i(x)g_j(x)$ e quindi l'applicazione lineare $M_{b_i(x)g_i(x)}M_{b_j(x)g_j(x)}$ è nulla in $K[x]/(f(x))$, pertanto la sua traccia è nulla. Questo prova che $\forall a \in K[x]/(p_i(x)^{n_i}), b \in K[x]/(p_j(x)^{n_j})$, vale $B(a,b) = 0$, come volevamo. □

Osservazione Distinguiamo due casi.

1. Gli autospazi generalizzati della matrice compagna M_x corrispondono agli addendi $\mathbb{R}[x]/(p_i(x)^{n_i})$ quando $p_i(x) = x - c_i$, cioè corrispondono alle radici c_i di f. Se la molteplicità della radice c_i è 1 allora l'autospazio generalizzato coincide con l'autospazio relativo a c_i. Infatti $M_x(1) = x = c_i$ (l'ultima uguaglianza modulo $(x - c_i)$.

2. Quando $p_i(x)$ è un polinomio di secondo grado con una coppia di radici complesse coniugate $\{\alpha_i, \overline{\alpha_i}\}$, $\mathbb{R}[x]/(p_i(x)^{n_i})$ è M_x-invariante, ed è somma dei due autospazi generalizzati corrispondenti alla coppia di radici su $\mathbb{C}$.

In entrambi i casi $\mathbb{R}[x]/(p_i(x)^{n_i})$ è un *anello locale*, con unico ideale massimale generato dalla classe di $p_i(x)$. Il quoziente rispetto all'ideale massimale è isomorfo a $\mathbb{R}$ nel primo caso ed a $\mathbb{C}$ nel secondo caso. Gli elementi in $\mathbb{R}[x]/(p_i(x)^{n_i})$ possono essere riguardati come sviluppi di Taylor in $x = c_i$ nel primo caso e come una coppia di sviluppi di Taylor coniugati rispetto a $x = \alpha_i$ e $x = \overline{\alpha_i}$ nel secondo caso. In particolare gli elementi invertibili di $\mathbb{R}[x]/(p_i(x)^{n_i})$ corrispondono nel primo caso alle classi dei

polinomi $q(x) \in \mathbb{R}[x]$ tali che $q(c_i) \neq 0$, mentre nel secondo caso alle classi dei polinomi $q(x) \in \mathbb{R}[x]$ tali che $q(\alpha_i) \neq 0$.

$\mathbb{R}[x]/(p_i(x)^{n_i})$ è uno spazio vettoriale su $\mathbb{R}$ di dimensione n_i nel primo caso e di dimensione $2n_i$ nel secondo caso.

3.3 Il numero di radici reali

Teorema 3.10 *[**Sylvester**] Sia B la matrice Bezoutiante di f, cioè la matrice della forma di Killing associata.*

(i) f ha n radici reali e distinte se e solo se B è definita positiva.

(ii) Il rango di B è il numero di radici (reali o complesse) distinte di f.

(iii) il numero di radici reali (distinte) di f è uguale al numero di autovalori positivi di B meno il numero di autovalori negativi di B.

Dimostrazione Il Teorema 3.9 permette di ricondurre il calcolo della segnatura della forma di Killing su $K[x]/(f(x))$ a quello della segnatura su ogni addendo $K[x]/(p_i(x)^{n_i})$. Tutte le applicazioni lineari della forma

$$M_{p(x)} \colon K[x]/(p(x)^n) \to K[x]/(p(x)^n)$$

sono nilpotenti, quindi hanno tutti gli autovalori nulli e la traccia nulla.

Nel caso in cui $K = \mathbb{R}$ abbiamo due casi da distinguere, dove $p(x) = x - c$ (radice reale) oppure dove $p(x) = x^2 + ax + b$ con $a^2 - 4b < 0$ (due radici complesse coniugate).

Nel primo caso, l'anello $K[x]/((x-c)^n)$ ha la base $1, x-c, (x-c)^2, \ldots, (x-c)^{n-1}$ e per quanto visto la matrice della forma di Killing ha tutti gli elementi nulli escluso quello in alto a sinistra, dove vale n, che è la traccia dell'identità. Pertanto il suo rango è 1 e la sua segnatura è 1, cioè ha un solo autovalore positivo e nessun autovalore negativo.

Nel secondo caso, l'anello $K[x]/((x^2+ax+b)^n)$ ha la base $1, x, (x^2+ax+b), x(x^2+ax+b), (x^2+ax+b)^2 \ldots, x(x^2+ax+b)^{n-1}$ e per quanto visto la matrice della forma di Killing ha tutti gli elementi nulli escluso quelli del blocco 2×2 in alto a sinistra, che corrisponde a

$$\begin{bmatrix} tr(1) & tr(M_x) \\ tr(M_x) & tr(M_{x^2}) \end{bmatrix}$$

Adesso $tr(1) = 2n$. Gli autovalori di M_x sono le radici di $(x^2 + ax + b)^n$. Chiamo $\alpha, \overline{\alpha}$ le due radici di $x^2 + ax + b = 0$, da cui $tr(M_x) = n(\alpha + \overline{\alpha})$, $tr(M_{x^2}) = n(\alpha^2 + \overline{\alpha}^2)$.

Quindi la segnatura della forma di Killing su $K[x]/((x^2 + ax + b)^n)$ equivale alla segnatura della matrice

$$n \begin{bmatrix} 2 & \alpha + \overline{\alpha} \\ \alpha + \overline{\alpha} & \alpha^2 + \overline{\alpha}^2 \end{bmatrix}$$

.

Dividendo la matrice per n (questo non modifica la segnatura), il determinante è $2(\alpha^2 + \overline{\alpha}^2) - (\alpha + \overline{\alpha})^2 = (\alpha - \overline{\alpha})^2 = -4(Im\alpha)^2 < 0$, da cui abbiamo un autovalore positivo e un autovalore negativo. Inoltre il rango è 2.

Sommando i contributi di tutti gli addendi, si ottiene la segnatura della forma di Killing come nel teorema di Sylvester. □

La segnatura di una forma quadratica può essere determinata facilmente, ispezionando il polinomio caratteristico. Infatti è noto che tutti gli autovalori di una matrice simmetrica reale sono reali, e si può applicare la

Teorema 3.11 (Regola di Cartesio sui segni) *Sia $p(x)$ un polinomio a coefficienti reali con tutte le radici reali. Il numero delle radici positive è uguale al numero delle variazioni di segno tra due suoi coefficienti non nulli consecutivi. Radici multiple sono contate quanto la loro molteplicità.*

Per una dimostrazione si veda [Aba]. Nel teorema 3.12 vedremo un modo di contare le radici positive di un polinomio anche senza l'ipotesi che tutte le radici siano reali.

Riguardo la regola di Cartesio, siccome la molteplicità della radice nulla è uguale all'esponente del più piccolo monomio che appare, possono essere calcolate esattamente il numero delle radici positive, nulle e negative (queste ultime per differenza).

Notiamo che M_x è simmetrico rispetto alla forma traccia di Killing, nel senso che $B(M_x a, b) = tr(M_{xab}) = B(a, M_x b)$. Attenzione, perché non è lecito applicare il teorema spettrale a meno che la forma non sia definita positiva, e infatti se la forma è definita positiva sappiamo, grazie al Teor. 2.4 e al Teor. 3.10 che f ha n radici reali distinte e che M_x è diagonalizzabile, in accordo col teorema spettrale.

Per ogni $a, b \in K[x]/(f(x))$, $h \in K[x]$ è definita la forma (bilineare) B_h (generalizzazione della forma di Killing)

$$\begin{array}{ccccc} K[x]/(f(x)) & \times & K[x]/(f(x)) & \longrightarrow & K \\ a & & b & \mapsto & tr(M_{hab}) \end{array}$$

É associata la forma quadratica $a \mapsto tr(M_{ha^2})$. La funzione $h(x)$ va pensata come una sorta di "funzione test", scegliendo $h(x)$ opportune di gradi 1 o 2 si possono avere informazioni rilevanti sulla localizzazione delle radici di f secondo la seguente proposizione, che generalizza il teorema di Sylvester.

Teorema 3.12 *Per un polinomio reale $h(x)$ di grado n, sia B_h la matrice della forma definita da $B_h(a, b) = tr(M_{hab})$ per ogni $a, b \in K[x]/(f(x))$. Notiamo che, rispetto alle notazioni precedenti, $B_1 = B$.*

(i) f ha n radici reali distinte p tali che $h(p) > 0$ se e solo se B_h è definita positiva.

(ii) Il rango di B_h è il numero di radici (reali o complesse) distinte p di f tali che $h(p) \neq 0$.

(iii) il numero di radici reali (distinte) di f tali che $h(p) > 0$ meno il numero di radici reali (distinte) di f tali che $h(p) < 0$ è uguale alla segnatura di B_h.

Inoltre supponiamo che $h(p) \neq 0 \; \forall p \in V(I)$ non reale, ipotesi soddisfatta se $\deg h \leq 1$ *oppure se h è primo con f.*

(iv) il numero di radici reali (distinte) di f tali che $h(p) > 0$ è uguale al numero di autovalori positivi di B_h meno il numero di autovalori negativi di B.

(v) il numero di radici reali (distinte) di f tali che $h(p) < 0$ è uguale al numero di autovalori negativi di B_h meno il numero di autovalori negativi di B.

Dimostrazione Ancora riconduciamo il calcolo della segnatura di B_h su $K[x]/(f(x))$ a quello della segnatura su ogni addendo $K[x]/(p_i(x)^{n_i})$, che sono ancora ortogonali, per lo stesso ragionamento fatto per la forma di Killing B.

Nel caso in cui $K = \mathbb{R}$ abbiamo due casi da distinguere, dove $p(x) = x - c$ (radice reale) oppure dove $p(x) = x^2 + ax + b$ con $a^2 - 4b < 0$ (due radici complesse coniugate).

Nel primo caso, l'anello $K[x]/((x-c)^n)$ ha la base $1, x-c, (x-c)^2, \ldots, (x-c)^{n-1}$ e per quanto visto la matrice della forma B_h, ristretta al sottospazio $K[x]/((x-c)^n)$, ha tutti gli elementi nulli escluso quello in alto a sinistra, dove vale $\mathrm{tr}(M_h) = \mathrm{tr}\,(h(M_x))$, per la Prop. 2.2 (iv). Siccome M_x ha il solo autovalore c, con molteplicità algebrica n, la traccia di $h(M_x)$ vale $nh(c)$. Pertanto, il rango di B_h, ristretta al sottospazio $K[x]/((x-c)^n)$, vale 1 se $h(c) \neq 0$ e vale 0 se $h(c) = 0$, la sua segnatura è 1 se $h(c) > 0$ mentre è -1 se $h(c) < 0$.

Nel secondo caso, l'anello $K[x]/((x^2+ax+b)^n)$ ha la base $1, x, (x^2+ax+b), x(x^2+ax+b), (x^2+ax+b)^3 \ldots, x(x^2+ax+b)^{n-1}$ e per quanto visto la matrice della forma B_h ha tutti gli elementi nulli escluso quelli del blocco 2×2 in alto a sinistra, che corrisponde a

$$\begin{bmatrix} tr(M_h) & tr(M_{xh}) \\ tr(M_{xh}) & tr(M_{x^2h}) \end{bmatrix}$$

Infatti la matrice 2×2 può essere scritta come

$$\begin{bmatrix} h(\alpha)+h(\overline{\alpha}) & h(\alpha)\alpha + h(\overline{\alpha})\overline{\alpha} \\ h(\alpha)\alpha + h(\overline{\alpha})\overline{\alpha} & h(\alpha)\alpha^2 + h(\overline{\alpha})\overline{\alpha}^2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \alpha & \overline{\alpha} \end{bmatrix} \cdot \begin{bmatrix} h(\alpha) & 0 \\ 0 & h(\overline{\alpha}) \end{bmatrix} \cdot \begin{bmatrix} 1 & \alpha \\ 1 & \overline{\alpha} \end{bmatrix} = U^t D U \tag{1}$$

e ha quindi rango due se $h(\alpha) \neq 0$ e α ha parte immaginaria non nulla. Quando questa condizione è soddisfatta, la sua segnatura è $(1,1)$ perché il suo determinante è negativo in quanto vale $(\det U)^2 \det D$ ed abbiamo $(\det U)^2 = (-2i \cdot Im\alpha)^2 < 0$, $\det D = h(\alpha)h(\overline{\alpha}) = h(\alpha)\overline{h(\alpha)} = |h(\alpha)|^2 > 0$.

□

Nel caso $h(x) = x$, il Teorema 3.12 permette di calcolare le radici positive. Nel caso $h(x) = x - a$, il Teorema 3.12 permette di calcolare le radici $> a$.

Osservazione 3.13 *E' possibile calcolare il numero di radici contenute in un qualunque intervallo. Infatti il numero di radici nell'intervallo $(a, b]$ si ottiene sottraendo dal numero di radici $> a$ quelle che sono $> b$.*

La dimostrazione della proposizione seguente è formalmente identica a quella della Prop. 3.7 e viene omessa.

Proposizione 3.14 *Sia $n = \deg f$. La matrice della forma B_x nella base $\{1, x, \ldots x^{n-1}\}$ è*

$$B_x = \begin{bmatrix} s_1 & s_2 & s_3 & \ldots & s_n \\ s_2 & s_3 & s_4 & \ldots & s_{n+1} \\ s_3 & s_4 & \ldots & & s_{n+2} \\ \vdots & & & & \vdots \\ s_n & s_{n+1} & \ldots & & s_{2n-1} \end{bmatrix}$$

Numerando le righe e le colonne da 0 *a* $n-1$*, il coefficiente di posto* (i,j) *di* B' *è* s_{i+j+1}.

Osservazione 3.15 *La matrice di* B_h *nella base standard* $\{1, x, \ldots, x^{n-1}\}$ *si può calcolare nel modo seguente. Se* $h(x) = \sum_k a_k x^k$ *allora al posto* (i,j) *(con la solita numerazione da* 0 *a* $n-1$*) abbiamo* $\sum_k a_k s_{i+j+k}$ *(dove* s_i *è la* i*-esima somma di potenze nelle radici di* f*), calcolabile facilmente con Macaulay 2 come* $tr(h(M_x)M_x^{i+j}) = tr(M_{h(x)x^{i+j}})$*, dove* M_x *è la matrice compagna.*

3.4 Criteri effettivi

Criterio per calcolare il numero di radici reali di un polinomio Sia dato un polinomio $f(x) \in \mathbb{R}[x]$. Dalla traccia della matrice compagna e delle sue potenze si calcola la matrice Bezoutiante. Dalla regola di Cartesio si può calcolare, mediante il polinomio caratteristico della matrice Bezoutiante B, il numero di autovalori positivi e il numero di autovalori negativi di B.

Allora, per il teorema di Sylvester, il numero di radici reali (distinte) è dato dal numero di autovalori positivi di B meno il numero di autovalori negativi di B.

Calcolo effettivo delle molteplicità delle radici. Per conoscere effettivamente le molteplicità di ciascuna radice di $f(x)$, si può calcolare $f_{rid} = \frac{f}{MCD(f,f')}$ e poi continuare induttivamente a valutare le radici di $f_2 = f/f_{rid} = MCD(f,f')$, $f_3 = f_2/f_{2rid}$, e così via. Se chiamiamo d_i il numero di radici distinte di f_i , calcolabile mediante il Teor. 3.10 (ii), allora il numero di radici di molteplicità i è uguale a $d_i - d_{i+1}$. Analogamente, se chiamiamo r_i il numero di radici reali distinte di f_i , calcolabile mediante il Teor. 3.10 (iii), allora il numero di radici reali di molteplicità i è uguale a $r_i - r_{i+1}$. Analogamente si possono trovare le molteplicità delle radici in un qualunque intervallo,

Criterio per stabilire se una matrice reale è diagonalizzabile su $\mathbb{R}$ Sia data A matrice reale. Si calcola il polinomio minimo di A, si veda l'algoritmo prima dell'eserc. 1.13. Si calcola la matrice Bezoutiante B del polinomio , mediante il metodo esposto nell'Osservazione 3.15 con $h = 1$. A è diagonalizzabile se e solo se B è definita positiva. Questo segue da (i) del Teor. 3.10 e dal Teor. 1.11. Ricordiamo che B è definita positiva se e solo se tutti i suoi minori principali sono positivi.

4 Decomposizione primaria di ideali zero-dimensionali e molteplicità

4.1 Matrici compagne in più variabili

Sia $I = (f_1, \ldots, f_k)$ un ideale zero dimensionale di $K[x_1, \ldots, x_n]$ dove $K = \mathbb{R}$ oppure $\mathbb{C}$. Questo significa che il sistema $f_1 = \ldots = f_k = 0$ ha un numero finito di soluzioni pari a d (contate con la relativa molteplicità), il quoziente $K[x_1, \ldots, x_n]/I$ ha dimensione d ed è generato dalle classi $[x^\alpha]$ di monomi non contenuti in $LT(I)$. Per ogni $f \in K[x_1, \ldots, x_n]/I$, calcolando con Macaulay2 $f\%I$, si ottiene l'espressione di f in questa base.

La moltiplicazione per x_i induce un'applicazione lineare

$$R \xrightarrow{M_{x_i}} R$$

$$[g] \mapsto [gx_i]$$

le matrici di M_{x_i} rispetto alla base $\{[x^\alpha] | x^\alpha \notin LT(I)\}$ si dicono *matrici compagne* di I. Nel caso di una variabile, la diagonalizzazione di M_x giocava un ruolo fondamentale. Nel caso di più variabili, è la diagonalizzazione simultanea delle M_{x_i} che entra in gioco.

Proposizione 4.1 *$\forall h(x), k(x) \in K[x_1, \ldots, x_n]$ vale che*
(i) $M_{h(x)} + M_{k(x)} = M_{h(x)+k(x)}$
(ii) $M_{ah(x)} = aM_{h(x)} \quad \forall a \in K$
(iii) $M_{h(x)} \cdot M_{k(x)} = M_{k(x)} \cdot M_{h(x)} = M_{h(x)k(x)}$
(iv) $M_{h(x_1,\ldots,x_n)} = h(M_{x_1}, \ldots, M_{x_n})$.

Dimostrazione (i), (ii) e (iii) sono immediate dalla definizione. Per provare (iv) supponiamo dapprima $h = x_j^i$. Allora $M_{x_j^i} = \left(M_{x_j}\right)^i$ come diretta applicazione di (iii).

In generale, se $h(x) = \sum a_\alpha x^\alpha$ allora, utilizzando anche (i)-(iii),

$$M_{h(x_1,\ldots,x_n)} = M_{\sum a_\alpha x^\alpha} = \sum a_\alpha M_{x^\alpha} = \sum a_\alpha \left(M_{x_1}\right)^{\alpha_1} \ldots \left(M_{x_n}\right)^{\alpha_n} = h(M_{x_1}, \ldots, M_{x_n})$$

□

Lemma 4.2 *Se $M_{x_i} v = \lambda_i v$ e $p \in K[x_1, \ldots, x_n]$ allora*

$$p(M_{x_1}, \ldots, M_{x_n})v = p(\lambda_1, \ldots, \lambda_n)v$$

Dimostrazione É l'analogo multidimensionale del Lemma 1.3.

4.2 Decomposizione primaria e definizione di molteplicità

Sia $I \subset \mathbb{C}[x_1, \ldots, x_n]$ e sia $V(I) = \{p_1, \ldots, p_k\} \subset \mathbb{C}^n$, cioè supponiamo che le soluzioni complesse del sistema definito da I siano un numero finito. Un tale ideale I è detto zero-dimensionale, equivalentemente il polinomio di Hilbert di I è costante, e questo può essere verificato in modo effettivo tramite Macaulay2.

In questa sezione assegneremo una molteplicità a ogni punto $p_i \in V(I)$, analogamente a quanto avviene per le radici di un polinomio in una variabile, in modo che la somma delle molteplicità di tutte le soluzioni sia uguale al grado del polinomio di Hilbert di I (si veda il Corollario 4.12).

Sia M_i l'ideale massimale dei polinomi che si annullano in $p_i = ((p_i)_1, \ldots, (p_i)_n)$. L'ideale M_i è generato dai polinomi $x_j - (p_i)_j$. Con un piccolo abuso di notazione, indicheremo con M_i anche la sua immagine nel quoziente $\mathbb{C}[x_1, \ldots, x_n]/I$, che è ancora un ideale massimale.

Lemma 4.3 *(i)* $V(M_1 \cap \ldots \cap M_k) = p_1 \cup \ldots \cup p_k$.

(ii) $\sqrt{I} = M_1 \cap \ldots \cap M_k$.

Questo significa che $g \in \sqrt{I}$ *se e solo se* $g(p_i) = 0$ *per* $i = 1, \ldots, k$. *In particolare, per ogni elemento del quoziente* $g \in \mathbb{C}[x_1, \ldots, x_n]/I$, *la valutazione* $g(p_i) \in \mathbb{C}$ *è ben definita e non dipende dal rappresentante.*

Dimostrazione (i) è elementare e segue dalla Proposizione 6.3 di [Introd]. Per provare (ii), se $f \in \sqrt{I}$ allora esiste $m > 0$ tale che $f^m(p_i) = 0$, da cui $f(p_i) = 0$ e quindi $f \in M_1 \cap \ldots \cap M_k$. Viceversa, per il teorema degli zeri, $\sqrt{I} = I(V(I)) = I(p_1 \cup \ldots \cup p_k) = I(V(M_1 \cap \ldots \cap M_k)) = \sqrt{M_1 \cap \ldots \cap M_k} \supset M_1 \cap \ldots \cap M_k$.

□

Lemma 4.4 *Dato* $V(I) = \{p_1, \ldots p_k\} \subset \mathbb{C}^n$, *esiste un polinomio* $h(x)$ *tale che* $h(p_i)$ *sono distinti (si veda la Figura 1). Se* I *è generato da polinomi a coefficienti reali, allora* $h(x)$ *puo' essere scelto a coefficienti reali. In tale caso, per ogni coppia di punti complessi coniugati* $\{p, \overline{p}\}$, *abbiamo* $h(\overline{p}) = \overline{h(p)}$.

Dimostrazione Il prodotto scalare euclideo si può estendere (algebricamente) a $\mathbb{C}^n$ ponendo $(z_1, \ldots, z_n) \cdot (w_1, \ldots, w_n) = \sum_{i=1}^n z_i w_i$, $\forall (z_1, \ldots, z_n), (w_1, \ldots, w_n) \in \mathbb{C}^n$. E' sufficiente scegliere un vettore $H = (h_1, \ldots, h_n)$ tale che il prodotto scalare euclideo $H \cdot (p_i - p_j) \neq 0$ $\forall i \neq j$. Questo è possibile perché $(p_i - p_j)$ sono un numero finito di vettori. Allora $h(x) = \sum_{i=1}^n h_i x_i$ soddisfa la condizione richiesta. □

Lemma 4.5 *Un elemento del quoziente* $g \in \mathbb{C}[x_1, \ldots, x_n]/I$ *è invertibile se e solo se* $g(p_i) \neq 0$ $\forall i$.

Dimostrazione Se g è invertibile, segue immediatamente che $g(p_i) \neq 0$. Viceversa, supponiamo $g(p_i) \neq 0$ $\forall i$. Per il lemma 4.4, esiste $h(x)$ tale che $h(p_i)$ sono distinti. Definiamo $g'(x) = \sum_{i=1}^k \frac{1}{g(p_i)} \prod_{j \neq i} \frac{h(x) - h(p_j)}{h(p_i) - h(p_j)}$, che soddisfa le uguaglianze $g(p_i)g'(p_i) = 1$ $\forall i$. Per il Lemma 4.3 (ii) abbiamo $1 - gg' \in \sqrt{I}$, da cui esiste $m > 0$ tale che $(1 - gg')^m \in I$. Espandendo la potenza m-esima e raccogliendo i termini che contengono g, si trova $\tilde{g}$ tale che $1 - \tilde{g}g \in I$, da cui g è invertibile nel quoziente , come volevamo. □

Lemma 4.6 *Sia* $V(I) = \{p_1, \ldots p_k\} \subset \mathbb{C}^n$, *e sia* $h(x) \in \mathbb{C}[x_1, \ldots, x_n]$. *Gli autovalori di* $M_{h(x)} \colon K[x_1, \ldots, x_n]/I \to K[x_1, \ldots, x_n]/I$ *coincidono con i valori* $h(p_i) \in \mathbb{C}$.

Dimostrazione Sia λ l' autovalore di $M_{h(x)}$ corrispondente all'autovettore $v(x)$. Allora $(h(x) - \lambda)\, v(x) \in I$. Affermiamo che $h(p_i) = \lambda$ per qualche i. Se per assurdo $h(p_i) - \lambda \neq 0$ $\forall i$, allora $h(x) - \lambda$ è invertibile per il Lemma 4.5. Quindi $v(x) \in I$, che è una contraddizione perché gli autovettori sono non nulli.

Viceversa, proviamo che $h(p_i)$ è un autovalore di $M_{h(x)}$. Sia $q(t)$ il polinomio minimo di $M_{h(x)}$. Allora $0 = q(M_{h(x)}) = M_{q(h(x))}$. Quindi $q(h(x)) \in I$, da cui, valutando in p_i, $q(h(p_i)) = 0$, pertanto $h(p_i)$ è un autovalore per la Prop. 1.4. □

Ricordiamo che un ideale J si dice *primario* se $fg \in J$ implica $f \in J$ oppure $g^m \in J$ per qualche $m > 0$. Segue immediatamente dalla definizione che il radicale di un ideale primario è primo.

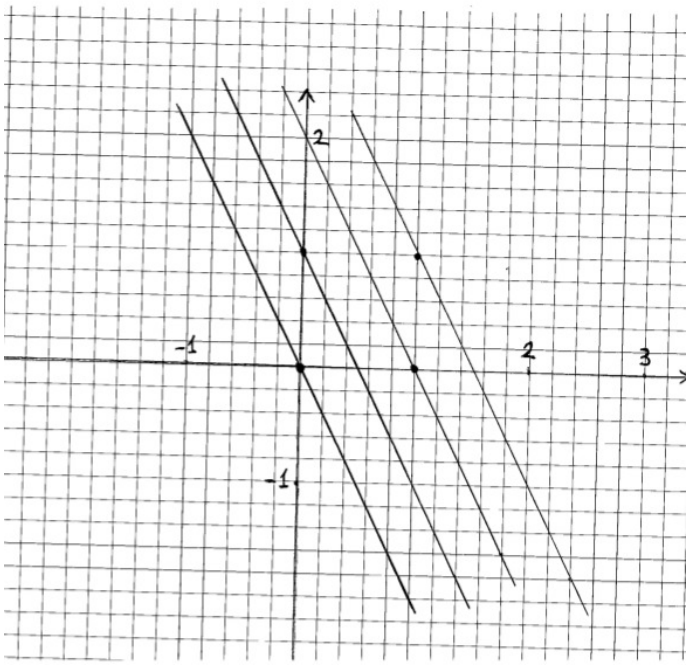

Figura 1: In figura i quattro punti $p_1 = (0,0)$, $p_2 = (1,0)$, $p_3 = (0,1)$, $p_4 = (1,1)$ corrispondono a $V(I)$ dove $I = (x(x-1), y(y-1))$. Posto $h(x,y) = 2x+y$, il fascio di rette parallele $h(x,y) = \lambda$ incontra $V(I)$ per i 4 valori $\lambda = h(p_i)$. I quattro autovalori di $M_{h(x)}$ sono $h(p_i)$. Ciascun punto ha molteplicità 1. In questo caso, M_x e M_y non hanno autovalori distinti (quali sono?).

Teorema 4.7 *[Decomposizione primaria] Sia $V(I) = \{p_1, \dots, p_k\}$. Sia $h(x) \in \mathbb{C}[x_1, \dots, x_n]$ tale che $h(p_i)$ siano distinti (si veda il Lemma 4.4).*

Considero per $i = 1, \dots, k$ le applicazioni lineari

$$M_{h(x)-h(p_i)} \colon \mathbb{C}[x_1, \dots, x_n]/I \to \mathbb{C}[x_1, \dots, x_n]/I$$

(i) Posto $A_i := \ker \left[M_{h(x)-h(p_i)}\right]^\infty$ (sono gli autospazi generalizzati di $M_{h(x)}$ per il Lemma 4.6, sono sottoalgebre e anche ideali di $\mathbb{C}[x_1, \dots, x_n]$), abbiamo la decomposizione diretta di sottoalgebre

$$\mathbb{C}[x_1, \dots, x_n]/I = \oplus_{i=1}^k A_i. \tag{2}$$

Se $v(x) \in A_i$ allora $v(p_j) = 0$ $\forall j \neq i$. Ogni sottoalgebra A_i ha un elemento unità e_i, che soddisfa le proprietà $e_i^2 = e_i$, $e_i e_j = 0$ per $i \neq j$. Inoltre, valutando in p_j, $e_i(p_j) = \delta_{ij}$. Gli elementi $g \in A_i$ sono invertibili in A_i se e solo se $g(p_i) \neq 0$.

(ii)Posto $J_i = \oplus_{j \neq i} A_j$, ideale di $\mathbb{C}[x_1, \dots, x_n]/I$, la sua retroimmagine $\tilde{J}_i \subset \mathbb{C}[x_1, \dots, x_n]$ è un ideale primario, tale che $\sqrt{\tilde{J}_i} = M_i$, $A_i = \mathbb{C}[x_1, \dots, x_n]/(\tilde{J}_i)$,

$$\cap_{i=1}^n \tilde{J}_i = I. \tag{3}$$

L'intersezione (3) si dice la decomposizione primaria di I. Notiamo che e_i corrisponde alla classe di 1 modulo $\tilde{J}_i$.

Dimostrazione

(i) La somma diretta segue dalla Prop. 1.9 e dal Lemma 4.6, come somma di spazi vettoriali. E' facile verificare dalla definizione che A_i è un ideale. Se $v(x) \in A_i$ allora esiste n_i tale che $(h(x) - h(p_i))^{n_i} v(x) \in I$, da cui valutando per $x = p_j$ segue $v(p_j) = 0$. L'unità e_i di ogni sottoalgebra A_i proviene dalla decomposizione in somma diretta $1 = \sum_{i=1}^{k} e_i$, risolubile dividendo 1 per i generatori di ciascuna A_i (aggiungendo eventualmente i generatori di I, si può applicare il comando *"quotientRemainder"* di M2. L'elemento e_i funge da unità in A_i perché, preso $a_i \in A_i \subset \mathbb{C}[x_1, \ldots, x_n]/I$, moltiplicando per 1 abbiamo

$$a_i = 1 \cdot a_i = \sum_{j=1}^{k} e_j a_i = e_i a_i.$$

Se $i \neq j$, abbiamo $e_i e_j \in A_i \cap A_j = 0$, da cui $1 = 1^2 = \sum_{i=1}^{k} e_i^2$ e per l'unicità della decomposizione $e_i^2 = e_i$. L'affermazione sull'invertibilità segue applicando il Lemma 4.5 al caso in cui $V(I)$ contiene un solo punto ($k = 1$).

(ii) Per come è definito l'ideale $\tilde{J}_i$, abbiamo $\mathbb{C}[x_1, \ldots, x_n]/\tilde{J}_i \simeq A_i$. Notiamo che $\cap_{i=1}^{n} J_i = 0$, da cui prendendo le retroimmagini $\cap_{i=1}^{n} \tilde{J}_i = I$. Valutando gli elementi unità e_i abbiamo $p_i = V(\tilde{J}_i)$, dal Nullstellensatz segue che $\sqrt{\tilde{J}_i} = M_i$. Rimane da vedere che $\tilde{J}_i$ è primario. Se $fg \in \tilde{J}_i$ allora $f(p_i)g(p_i) = 0$, da cui $f(p_i) = 0$ oppure $g(p_i) = 0$ e quindi $f^m \in \tilde{J}_i$ oppure $g^m \in \tilde{J}_i$ come volevamo. □

Osservazione 4.8 *Gli anelli A_i hanno come unico ideale massimale M_i e sono quindi anelli locali. La decomposizione (2) spiega l'origine del termine locale. Ogni A_i corrisponde a localizzare in un punto p_i , cioè la classe di un polinomio in A_i è influenzata soltanto dal comportamento vicino as p_i e può essere ricostruita da un opportuno sviluppo di Taylor nel punto p_i (rispetto ai monomi $\notin \tilde{J}_i$). Per approfondimenti sugli anelli locali si può vedere il cap. 4 di [CLO2].*

Definizione 4.9 $\dim A_i$ *si dice* <u>molteplicità</u> *di p_i in I, la indicheremo con m_{p_i}, non dipende dal polinomio $h(x)$ scelto nel Teorema 4.7.*

Per provare che la molteplicità è ben definita e non dipende da $h(x)$, prendiamo un altro polinomio $h'(x)$ che assume valori distinti sui punti p_i. Si osserva che A_i è $h'(x)$-invariante. Siccome $A_i = \mathbb{C}[x_1, \ldots, x_n]/(\tilde{J}_i)$ e $V(\tilde{J}_i) = V(\sqrt{\tilde{J}_i}) = \{p_i\}$ dal Teor. 4.7 (ii), segue che l'unico autovalore di $M_{h'(x)}$ su A_i è $h'(p_i)$ per il Lemma 4.6. Pertanto l' autospazio generalizzato A_i di $M_{h(x)}$ relativo all'autovalore $h(p_i)$ è contenuto nell' autospazio generalizzato A_i' di $M_{h'(x)}$ relativo all'autovalore $h'(p_i)$. Sia la somma dei A_i che quella dei A_i' sono entrambe dirette, quindi vale l'uguaglianza $A_i = A_i'$.

Come conseguenza di questo ragionamento enunciamo esplicitamente la

Proposizione 4.10 *(i) Sia $h(x)$ un polinomio che assume valori distinti sui punti p_i. Gli ideali A_i sono gli autospazi generalizzati di $M_{h(x)}$ e l'unico autovalore di $M_{h(x)}$ su A_i è $h(p_i)$.*

(ii) Per ogni polinomio $k(x)$, l'unico autovalore di $M_{k(x)}$ su A_i è $k(p_i)$ (segue dal Lemma 4.6 applicato al caso $I = \tilde{J}_i$).

(iii) La sottoalgebra A_i del teorema 4.7 (i) dipende solo da I.

(iv) La decomposizione primaria $I = \cap_{i=1}^{n} \tilde{J}_i$ del Teorema 4.7 (ii) è unica.

Esercizio 4.11 *Modificando la figura 1, consideriamo $I = \big(x^2(x-1), y(y-1)\big)$. Provare che, con le notazioni della figura 1, $V(I) = \{p_1, p_2, p_3, p_4\}$, la molteplicità di p_1, p_3 è 2, la molteplicità di p_2, p_4 è 1.*

Corollario 4.12 *La somma delle molteplicità di ciascun p_i in I è uguale alla dimensione di $\mathbb{C}[x_1, \ldots, x_n]/I$. Questo valore è uguale al polinomio di Hilbert di I, che ha grado zero ed è costante.*

Corollario 4.13 *Vale $I = \sqrt{I}$ se e solo se tutti i punti hanno molteplicità 1 in I.*

Dimostrazione Basta confrontare

$$K[x_1, \ldots, x_n]/I = \oplus_{i=1}^{k} K[x_1, \ldots, x_n]/\tilde{J}_i$$

con

$$K[x_1, \ldots, x_n]/\sqrt{I} = \oplus_{i=1}^{k} K[x_1, \ldots, x_n]/M_i,$$

dove nella seconda somma gli addendi hanno dimensione 1, si veda il Lemma 4.3 (ii).□

La decomposizione primaria del Teorema 4.7 si generalizza a ideali qualunque di $K[x_1, \ldots, x_n]$, dove $V(I)$ può avere dimensione positiva, secondo un risultato classico di Lasker-Noether. In questi casi, trovare una decomposizione esplicita è più difficile, e non è necessariamente unica, come nel caso zero-dimensionale. Per dettagli si può consultare il cap. 4 §7 di [CLO1].

4.3 Calcolo effettivo della molteplicità di ogni soluzione

Per ogni polinomio h, la molteplicità algebrica dell'autovalore λ per

$$M_h \colon K[x_1, \ldots x_n]/I \to K[x_1, \ldots x_n]/I$$

è uguale a $\sum_{\{p|h(p)=\lambda\}} m_p$, infatti $K[x_1, \ldots x_n]/I = \oplus A_i$, ogni A_i è M_h- invariante e l'unico autovalore di M_h su A_i è proprio $h(p_i)$ (si veda il Lemma 4.6).

Pertanto, scegliendo una forma lineare h generale che prende valori distinti su ogni punto di $V(I)$, le molteplicità m_{p_i} possono essere calcolate come le molteplicità algebriche degli autovalori di M_h. A_i sono gli autospazi generalizzati per M_h, come definiti nella Prop. 1.9. Con probabilità uno, una forma lineare h scelta casualmente (random) soddisfa questa condizione e permette di calcolare effettivamente le molteplicità. Purtroppo, la scelta random non garantisce a priori di trovare il polinomio h richiesto.

Un criterio sufficiente per verificare se h prende valori distinti su $V(I)$, senza ancora conoscere $V(I)$, è verificare se M_h è regolare, cioè se il suo polinomio minimo e caratteristico coincidono (a meno del segno). Purtroppo, il criterio è solo sufficiente. Ad esempio consideriamo $K[x, y]/(x^2, y^2)$ che corrisponde al punto $(0, 0)$ con molteplicità

4. La matrice M_x ha un autospazio di dimensione 2 generato da x, xy, e ci sono due blocchi di Jordan di ordine 2. La matrice M_y ha un autospazio di dimensione 2 generato da y, xy, e ci sono due blocchi di Jordan di ordine 2. La matrice $M_{x+y} = M_x + M_y$ ha un autospazio di dimensione 2 generato da $xy, x - y$, e ci sono un blocco di Jordan di ordine 1 e un blocco di Jordan di ordine 3, e questo è il comportamento per $M_{\alpha x+\beta y}$ generale.

Una tecnica che garantisce il calcolo effettivo delle molteplicità m_{p_i} è di calcolare le molteplicità degli autovalori di M_{x_1}, isolando ciascun autovalore in un intervallo, seguendo l'Osservazione 3.13 (si veda il Teorema 6.1). Per ciascuno di questi intervalli, si calcolano le molteplicità degli autovalori di M_{x_2}, isolandoli in intervalli rispetto a x_2, e così via. Questa procedura è laboriosa ma ha sempre successo.

5 Sistemi zero dimensionali in più variabili

Sia $K = \mathbb{R}$ oppure $K = \mathbb{C}$. Consideriamo $f_i \in K[x_1, \ldots, x_n]$ per $i = 1, \ldots, k$. Sia $I = (f_1, \ldots, f_k)$ l'ideale generato da questi polinomi.

Teorema 5.1 (Stickelberger) *Sia I un ideale zero dimensionale e siano $M_{x_i} : \mathbb{C}[x_1, \ldots, x_n]/I \to \mathbb{C}[x_1, \ldots, x_n]/I$ le applicazioni lineari (compagne) indotte dalla moltiplicazione per x_i. Esiste un autovettore v comune a M_{x_i} con autovalori λ_i, cioè $M_{x_i} v = \lambda_i v$, se e solo se $(\lambda_1, \ldots, \lambda_n) \in V(I)$.*

Dimostrazione

Sia v un autovettore tale che $M_{x_i} v = \lambda_i v \; \forall i$. Se $f \in I$, ricordiamo che $M_{f(x_1,\ldots,x_n)} = 0$, quindi $0 = M_{f(x_1,\ldots,x_n)} v = f(M_{x_1}, \ldots, M_{x_n}) v = f(\lambda_1, \ldots, \lambda_n) v$, l'ultima uguaglianza per il Lemma 4.2, da cui $f(\lambda_1, \ldots, \lambda_n) = 0$.

Viceversa, dobbiamo provare che le coordinate di ogni $p_i \in V(I)$ sono autovalori di un autovettore comune delle matrici M_{x_j}. Decomponiamo $\mathbb{C}[x_1, \ldots, x_n]/I = \oplus_{i=1}^{k} A_i$ secondo il Teorema 4.7. A_i è M_{x_j}-invariante per $j = 1, \ldots, n$ e $M_{x_1}, \ldots M_{x_n}$ commutano. Pertanto per la Prop. 1.17 esiste un autovettore comune per M_{x_j}, il cui autovalore è la coordinata j-esima di p_i, come volevamo. □ □

Corollario 5.2 *Nelle ipotesi del teorema di Stickelberger, il polinomio monico generatore dell'ideale di eliminazione $I \cap \mathbb{C}[x_i]$ coincide con il polinomio minimo di M_{x_i} (sostituendo $x = x_i$).*

Dimostrazione Sia $p(x)$ il polinomio generatore di $I \cap \mathbb{C}[x_i]$ e sia $h(x)$ il polinomio minimo di M_{x_i}. Siccome $p(M_{x_i}) = M_{p(x_i)}$ è l'applicazione nulla da $K[x_1, \ldots, x_n]/I$ in sé stesso, perché $p \in I$, segue che h divide p. Viceversa $h(M_{x_i}) = M_{h(x_i)}$ è l'applicazione nulla, quindi applicata ad 1 mostra che $h(x_i) \in I$, da cui p divide h. □

Teorema 5.3 *Sia $I \subseteq \mathbb{C}[x_1, \ldots, x_n]$ un ideale zero-dimensionale. Le seguenti condizioni sono equivalenti*

(i) M_{x_i} sono diagonalizzabili

(ii) M_{x_i} sono diagonalizzabili simultaneamente (cioè con una base comune di autovettori).

(iii) $V(I)$ ha punti distinti, cioé I è radicale. (eserc. 12 pag. 61 di [CLO2])

Dimostrazione (i) e (ii) sono equivalenti per il Teor. 1.15.

(iii)$\Longrightarrow$(i) Se $V(I)$ ha punti distinti, scegliamo per il Lemma 4.4 una combinazione lineare $h = \sum_i a_i x_i$ che assume valori distinti su $V(I)$. Allora dal Lemma 4.6 M_h ha d autovalori distinti e quindi è diagonalizzabile. Per la Prop. 1.16 otteniamo che M_{x_i} (che commutano con M_h) sono tutte diagonalizzabili.

(ii)$\Longrightarrow$(iii) Se M_{x_i} sono diagonalizzabili, allora ogni elemento di A_j è un autovettore per M_{x_i} con autovalore $(p_j)_i$. In particolare $e_j(x_i - (p_j)_i) \in \tilde{J}_j$ per ogni i, e dall'invertibilità di e_j (modulo $\tilde{J}_j$) segue $(x_i - (p_j)_i) \in \tilde{J}_j$ per ogni i. Abbiamo che i generatori di M_j appartengono a $\tilde{J}_j$ e quindi $M_j = \tilde{J}_j$, da cui la molteplicità di p_j è 1.

□

Proposizione 5.4 *Siano $x^{\alpha(1)}, \dots x^{\alpha(m)}$ i monomi non in $LT(I)$ che generano $K[x_1, \dots, x_n]/I$ Per ogni punto p di $V(I)$ e ogni polinomio h, il vettore $p^{\alpha(1)}, \dots p^{\alpha(m)}$ (ottenuto calcolando i monomi in p) è un autovettore di M_h^t con autovalore $h(p)$.*

Dimostrazione Siano m_{ij} i coefficienti di M_h. Abbiamo $[x^{\alpha(j)}h] = M_h([x^{\alpha(j)}]) = \sum_{i=1}^m m_{ij}[x^{\alpha(i)}]$.

Valutando in p otteniamo $p^{\alpha(j)}h(p) = \sum_{i=1}^m m_{ij}p^{\alpha(i)}$, che equivale alla tesi. □

Seconda dimostrazione. Diamo una dimostrazione più elegante, senza usare le coordinate. La valutazione in p, $ev(p) \in (K[x_1, \dots, x_n]/I)^\vee$ è definita da $ev(p)(h) = h(p)$. Ricordiamo (vedi [Aba, 8C.2]) che se $A\colon V \to W$ è una applicazione lineare, la sua trasposta $A^t\colon W^\vee \to V^\vee$ è definita da $A^t(w^*)(v) = w^*(A(v))\ \forall w^* \in W^\vee, v \in V$. La notazione è giustificata dal fatto che la matrice di A^t, calcolata nelle basi duali, coincide con la trasposta della matrice di A. Facciamo vedere che $ev(p)$ è autovettore della trasposta M_h^t, con autovalore $h(p)$. Infatti, $\forall b \in K[x_1, \dots, x_n]/I$

$$M_h^t(ev(p))(b) = ev(p)(M_h(b)) = ev(p)(hb) = h(p)b(p) = h(p)ev(p)(b)$$

da cui

$$M_h^t(ev(p)) = h(p)ev(p)$$

che equivale alla tesi. □

La Proposizione 5.4 è utile per il calcolo delle soluzioni di un sistema polinomiale, soprattutto quando tra i monomi non in $LT(I)$ appaiono i generatori x_i.

6 La forma traccia in più variabili e il numero di soluzioni reali

In questa sezione, I è un ideale zero-dimensionale di $\mathbb{R}[x_1, \dots, x_n]$. Le sue soluzioni in $\mathbb{C}^n$ si dividono in punti reali e in coppie di punti complessi coniugati. I può essere visto

come ideale in $\mathbb{C}[x_1,\dots,x_n]$ (generato da polinomi reali), e il suo quoziente R si spezza, come nel Teorema 4.7, nella somma diretta di A_i, dove alcuni A_i corrispondono ai punti reali, mentre altre coppie A_j, $\overline{A_j}$ corrispondono a coppie di punti complessi coniugati. Il coniugio agisce su $\mathbb{C}[x_1,\dots,x_n]$, coniugando i coefficienti di ogni polinomio, ed è un morfismo di anelli, che lascia invarianti le sottoalgebre A_i corrispondenti ai punti reali e scambia tra loro le sottoalgebre coniugate A_j e $\overline{A_j}$. Ne segue che in corrispondenza dei punti reali le unità $e_i \in A_i$ sono reali, mentre coniugando l'unità $e_j \in A_j$ relativa a una coppia coniugata si trova l'unità $\overline{e_j} \in \overline{A_j}$. Notiamo che la somma $A_j \oplus \overline{A_j}$ è una sottoalgebra con unità $e_j + \overline{e_j}$, che è sempre un anello locale. Questo permette di decomporre sui reali $\dim \mathbb{R}[x_1,\dots,x_n]/I$, che diventa somma delle sottoalgebre A_i generate dalle unità reali e_i corrispondenti ai punti reali e dalle sottoalgebre generate da $e_j + \overline{e_j}$ nel caso di coppie di punti coniugati. Il campo residuo di ciascuna sottoalgebra (vista come anello locale) è $\mathbb{R}$ nel primo caso e $\mathbb{C}$ nel secondo caso. In alternativa, scelto $h \in \mathbb{R}[x_1,\dots,x_n]$ che assume valori distinti su $V_{\mathbb{C}}(I)$, la sottoalgebra relativa a p nel primo caso è $\ker M^{\infty}_{h(x)-h(p)}$, mentre la sottoalgebra relativa alla coppia $\{p,\overline{p}\}$ nel secondo caso è $\ker M^{\infty}_{(h(x)-h(p))(h(x)-h(\overline{p}))}$.

La forma traccia è definita analogamente al caso unidimensionale, cioè

$$B_h(a,b) = Tr(M_{hab})$$

per ogni $a, b \in R$. La decomposizione $\oplus_i A_i$ è ortogonale rispetto a B_h (basta calcolarla sulle unità di ogni sottoalgebra).

Il seguente teorema generalizza il criterio di Sylvester al caso multidimensionale e permette di calcolare, in aritmetica esatta, il numero di punti reali e di coppie di punti complessi coniugati in $V(I)$, oltre ad altre informazioni che, per particolari h, permettono di localizzare le radici (ad esempio studiando a quale ottante appartengono).

Teorema 6.1 (Hermite) *(i) Sia* $\dim \mathbb{R}[x_1,\dots,x_n]/I = m$*, sia* $h \in \mathbb{R}[x_1,\dots,x_n]$*. La varietà* $V_{\mathbb{R}}(I)$ *consiste di* m *punti distinti* p *tali che* $h(p) > 0$ *se e solo se* B_h *è definita positiva. In particolare* $V_{\mathbb{R}}(I)$ *consiste di* m *punti distinti se e solo se* B_1 *è definita positiva.*

(ii) Il rango di B_h *è il numero di punti distinti* $p \in V_{\mathbb{C}}(I)$ *tali che* $h(p) \neq 0$*. In particolare il rango di* B_1 *è il numero di punti distinti in* $V(I)$*.*

(iii) il numero di punti reali (distinti) $p \in V_{\mathbb{R}}(I)$ *tali che* $h(p) > 0$ *meno il numero di punti reali (distinti)* $p \in V_{\mathbb{R}}(I)$ *tali che* $h(p) < 0$ *è uguale alla segnatura di* B_h*. In particolare il numero dei punti reali (distinti)* $p \in V_{\mathbb{R}}(I)$ *è uguale alla segnatura di* B_1*.*

Inoltre supponiamo che $h(p) \neq 0\ \forall p \in V(I)$ *non reale, ipotesi soddisfatta se* $\deg h \leq 1$*.*

(iv) il numero di punti reali (distinti) $p \in V_{\mathbb{R}}(I)$ *tali che* $h(p) > 0$ *è uguale al numero di autovalori positivi di* B_h *meno il numero di autovalori negativi di* B_1*.*

(v) il numero di punti reali (distinti) $p \in V_{\mathbb{R}}(I)$ *tali che* $h(p) < 0$ *è uguale al numero di autovalori negativi di* B_h *meno il numero di autovalori negativi di* B_1*.*

Dimostrazione Sviluppiamo i punti salienti della dimostrazione, che è analoga a quella del Teorema 3.12. Per ogni polinomio h, la matrice M_h si decompone su ciascuna

sottoalgebra A_i, inoltre la traccia di M_h su A_i è uguale a $m_{p_i}h(p_i)$. Infatti l'unico autovalore di M_h su A_i è dato da $h(p_i)$ per il Lemma 4.6.

Per calcolare la forma traccia, osserviamo che A_i ha come ideale massimale l'ideale M_i dei polinomi che si annullano in p_i. Pertanto si può scegliere come base di A_i i polinomi e_if_j (vedi il Teor. 4.7 (i)) per $j = 0, \dots m_{p_i} - 1$ dove $f_0 = 1$ e $f_j(p_i) = 0$ per $j \geq 1$. Siccome $(x - p_i)^{\alpha_k}$ formano una base dell'anello dei polinomi, al variare di $\alpha_k \in \mathbb{Z}^n_{\geq 0}$, (e questo per ogni p_i), si può anche scegliere $f_j = (x - p_i)^{\alpha_j}$ per convenienti α_j.

Pertanto calcolando la matrice di B_h rispetto a questa base, per una radice reale rimane solo il contributo di $tr(M_{he_ie_i})$ che vale $h(p_i)$ e tutti gli altri elementi hanno la forma $tr(M_{he_if_{j_1}f_{j_2}})$ dove uno tra j_1 e j_2 è positivo, e quindi $he_if_{j_1}f_{j_2}$ vale zero in p_i a per quanto visto $tr(M_{he_if_{j_1}f_{j_2}}) = 0$. Il rango di B_h ristretta a A_i vale 1.

Nel caso di una coppia di radici complesse coniugate $\{p_i, \overline{p_i}\}$ allora possiamo considerare la base $e_if_j + \overline{e_if_j}$, $\frac{1}{\sqrt{-1}}\left(e_if_j - \overline{e_if_j}\right)$, dove e_i e f_j sono gli stessi polinomi visti nel caso complesso relativi a p_i ed i loro coniugati $\overline{e_i}$, $\overline{f_j}$ vanno intesi come i polinomi con le stessi monomi ed i coefficienti coniugati. In particolare $\overline{e_i}$ è l'unità della sottoalgebra relativa a $\overline{p_i}$, (si veda l'esempio 6.2).

La matrice di B_h rispetto a questa base è nulla tranne il blocco 2×2 in alto a sinistra che è

$$m_{p_i}\begin{bmatrix} h(p_i) + h(\overline{p_i}) & \frac{1}{\sqrt{-1}}(h(p_i) - h(\overline{p_i})) \\ \frac{1}{\sqrt{-1}}(h(p_i) - h(\overline{p_i})) & -(h(p_i) + h(\overline{p_i})) \end{bmatrix} = m_{p_i}U^tDU$$

dove $U = \begin{bmatrix} 1 & -\sqrt{-1} \\ 1 & \sqrt{-1} \end{bmatrix}$, $D = \begin{bmatrix} h(p_i) & 0 \\ 0 & h(\overline{p_i}) \end{bmatrix}$

Siccome $\det U = 2\sqrt{-1}$, $detD = |h(p_i)|^2$ segue $\det(U^tDU) = (\det U)^2 \det D = -4|h(p_i)|^2$, quindi B_h ha rango 2 se $h(p_i) \neq 0$ e rango zero altrimenti, mentre se il rango è 2 allora $\det U^tDU < 0$ e la segnatura di B_h vale zero. Notiamo che la segnatura è zero in ogni caso. □

Esempio 6.2 *Nel caso* $\mathbb{C}[x]/\left((x-a)(x-\overline{a})\right) = \mathbb{C}[x]/(x-a) \oplus \mathbb{C}[x]/(x-\overline{a})$ *l'unità della sottoalgebra* $\mathbb{C}[x]/(x-a)$ *è* $e = \frac{x-\overline{a}}{a-\overline{a}}$ *mentre l'unità dell'altra sottoalgebra* $\mathbb{C}[x]/(x-\overline{a})$ *è* $\overline{e} = \frac{x-a}{\overline{a}-a} = 1 - e$. *In questo caso* $\frac{1}{\sqrt{-1}}(e - \overline{e}) = -\frac{x-Re(a)}{Im(a)}$ *ed abbiamo* $\frac{1}{\sqrt{-1}}(e-\overline{e})(a) + \frac{1}{\sqrt{-1}}(e-\overline{e})(\overline{a}) = 0$ *e più in generale* $\frac{1}{\sqrt{-1}}(e-\overline{e})(a)h(a) + \frac{1}{\sqrt{-1}}(e-\overline{e})(\overline{a})h(\overline{a}) = \frac{1}{\sqrt{-1}}(h(a) - h(\overline{a}))$. *Inoltre* $\left(\frac{1}{\sqrt{-1}}(e-\overline{e})\right)^2 = -1$.

Riferimenti bibliografici

[Aba] M. Abate, *Geometria*, McGraw-Hill, 1996

[CLO1] D. Cox, J. Little, D. O'Shea, *Ideal, Varieties and Algorithms*, Springer, 1992

[CLO2] D. Cox, J. Little, D. O'Shea, *Using Algebraic Geometry*, Springer, 1998

[EM] M. Elkadi, B. Mourrain, *Introduction à la résolution de systèmes polynomiaux*, Mathématiques et Applications 59, Springer, Berlin, 2007

[Introd] G. Ottaviani, *Introduzione alle varietà algebriche, un punto di vista costruttivo*, note reperibili sulla pagina web

[Stu] B. Sturmfels, *Solving systems of polynomial equations*, CBMS Regional Conference Series in Mathematics, 97, AMS, 2002