Appunti per Geometria e Algebra Computazionale 3-6. Calcolo effettivo delle molteplicità . Teorema di Stickelberger. Forma traccia in più variabili.

Corso di Laurea in Matematica, Università di Firenze, 2019/20

Giorgio Ottaviani

2 giugno 2020

Calcolo effettivo della molteplicità, I

Il numero di radici con una data molteplicità può essere calcolato dal radicale di un ideale (algoritmo di Krick-Logar) ed iterando il comando $\mathrm{quotient}(\mathrm{I}, \mathrm{radical\ I})$. Questo metodo è complesso, sia in teoria che in pratica. Accenniamo a una tecnica alternativa che ha successo con probabilità uno ed è più semplice.
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 $h(p_i)$. 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. Una forma lineare h scelta 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.

Calcolo effettivo della molteplicità, II

Un criterio sufficiente per verificare se h prende valori distinti su $V(I)$, senza 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.

Calcolo effettivo della molteplicità, III

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, con le tecniche viste usando la bezoutiante in una variabile. 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.

Il teorema di Stickelberger

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 (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 del teorema di Stickelberger

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
$0 = M_{f(x_1,\ldots,x_n)} v = f(M_{x_1}, \ldots, M_{x_n}) v = f(\lambda_1, \ldots, \lambda_n) v$, 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 di Decomposizione Primaria. A_i è M_{x_j}-invariante per $j = 1, \ldots, n$ e $M_{x_1}, \ldots M_{x_n}$ commutano. Pertanto (triangolarizzazione simultanea) esiste un autovettore comune per gli endomorfismi $(M_{x_j})_{|A_i}$, per $j = 1, \ldots, n$, il cui autovalore relativo a $(M_{x_j})_{|A_i}$ è la coordinata j-esima di p_i, per i risultati visti in GAC3-5, come volevamo. □

L'ideale di eliminazione di I in ciascuna variabile.

Corollario

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 sè 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. □

Un ideale zero-dimensionale è radicale se e solo se le matrici compagne M_{x_i} sono diagonalizzabili.

Teorema

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, cioè I è radicale. (eserc. 12 pag. 61 di [?])

Un ideale zero-dimensionale è radicale se e solo se le matrici compagne M_{x_i} sono diagonalizzabili, la dimostrazione

Dimostrazione.

(i) e (ii) sono equivalenti perché commutano.
(iii)$\Longrightarrow$(i) Se $V(I)$ ha punti distinti, scegliamo una combinazione lineare $h = \sum_i a_i x_i$ che assume valori distinti su $V(I)$. Allora dal Lemma visto in GAC3-5 M_h ha d autovalori distinti e quindi è diagonalizzabile. Per la Prop. vista sulla diagonalizzazione simultanea 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.

Calcolo delle soluzioni mediante gli autovettori della trasposta di M_h

Proposizione

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

Dimostrazione.

La valutazione in p, $ev(p) \in (K[x_1, \ldots, x_n]/I)^\vee$ è definita da $ev(p)(h) = h(p)$. Ricordiamo 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 $A^t(w^*)(v) = w^*(A(v))\ \forall w^* \in W^\vee, v \in V$. Facciamo vedere che $ev(p)$ è autovettore della trasposta M_h^t, con autovalore $h(p)$. Infatti, $\forall b \in K[x_1, \ldots, x_n]/I\ M_h^t(ev(p))(b) = ev(p)(M_h(b)) = ev(p)(hb) = h(p)b(p) = h(p)ev(p)(b)$ da cui $M_h^t(ev(p)) = h(p)ev(p)$ che equivale alla tesi. □

Calcolo effettivo delle soluzioni

La Proposizione precedente 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. Ad esempio l'ideale $((x+y)^4 + 2xy^2, x^2 + y^2 - x) \subset \mathbb{Q}[x,y]$ ha una base di $\mathbb{Q}[x,y]/LT(I)$ data dagli otto monomi $\{1, x, xy, xy^2, xy^3, y, y^2, y^3\}$. Notiamo che nelle posizioni $0, 1, 5$ appaiono rispettivamente $\{1, x, y\}$. Allora per ogni autovettore $v = \{v_0, \ldots, v_7\}$ di M_h^t, le espressioni $x = v_1/v_0$, $y = v_5/v_0$ forniscono le coordinate di un punto $p \in V(I)$, tale che l'autovalore corrispondente a v è proprio $h(p)$.

Coppie di sottoalgebre coniugate, I

Consideriamo I ideale zero-dimensionale di $\mathbb{R}[x_1, \ldots, x_n]$. 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 $\mathbb{C}[x_1, \ldots, x_n]$ (generato da polinomi reali), e il suo quoziente R si spezza, come nel Teorema di decomposizione primaria, 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 $\mathbb{C}[x_1, \ldots, x_n]$, 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 $\overline{e_j} \in \overline{A_j}$.

Coppie di sottoalgebre coniugate, II

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 $\dim \mathbb{R}[x_1, \ldots, x_n]/I$, 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 $h \in \mathbb{R}[x_1, \ldots, x_n]$ 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 in più variabili e il numero di soluzioni reali

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 (Hermite)

(i) Sia $\dim \mathbb{R}[x_1, \ldots, x_n]/I = m$, *sia* $h \in \mathbb{R}[x_1, \ldots, x_n]$. *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.

Il Teorema di Hermite, II

Teorema (Hermite)

Inoltre supponiamo che $h(p) \neq 0 \ \forall p \in V(I)$ *non reale, ipotesi soddisfatta se* h *assume valori distinti su* $V_{\mathbb{C}}(I)$.
(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 del Teorema di Hermite, I

Dimostrazione.

La dimostrazione è analoga a quella in una variabile. 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 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 $e_i f_j$ 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 $tr(M_{he_i e_i})$ che vale $h(p_i)$ e tutti gli altri elementi hanno la forma $tr(M_{he_i f_{j_1} f_{j_2}})$ dove uno tra j_1 e j_2 è positivo, e quindi $he_i f_{j_1} f_{j_2}$ vale zero in p_i . Il rango di B_h ristretta a A_i vale 1. □

Dimostrazione.

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 $e_i f_j + \overline{e_i f_j}$, $\frac{1}{\sqrt{-1}}\left(e_i f_j - \overline{e_i f_j}\right)$, 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 $\overline{p_i}$. La matrice di B_h rispetto a questa base è nulla tranne il blocco 2×2 in alto a sinistra, corrispondente agli elementi della base $\{e_i + \overline{e_i}, \frac{1}{\sqrt{-1}}(e_i - \overline{e_i})\}$, che è

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

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

Dimostrazione del Teorema di Hermite, III

Dimostrazione.

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

Un esempio conclusivo

Esempio

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* $\left(\frac{1}{\sqrt{-1}}(e-\overline{e})\right)^2 = -1$.