December 30, 2018

CENNI SUL DISCRIMINANTE

Fissiamo un campo $\mathbf{K}$ *algebricamente chiuso* (per esempio $\mathbf{K} = \mathbb{C}$). Per ogni $n \geq 2$, indichiamo con

$$\mathcal{P}_n = \mathcal{P}_n(\mathbf{K})$$

il sottoinsieme di $\mathbf{K}[t]$ formato dai *polinomi monici di grado* n. Un tale polinomio è della forma

$$p(t) = a_0 + a_1 t + a_2 t^2 + \dots a_{n-1} t^{n-1} + t^n$$

e quindi $\mathcal{P}_n$ può essere indentificato con una copia di $\mathbf{K}^n$ con coordinate $a = (a_0, \dots, a_{n-1})$. Indichiamo con

$$\Delta_n$$

il sottoinsieme di $\mathcal{P}_n$ formato dai polinomi che hanno almeno una radice multipla, cioè di molteplicità $m > 1$. Δ_n è chiamato il *(luogo) discriminante* di $\mathcal{P}_n$. $\mathcal{P}_n \setminus \Delta_n$ consiste dei polinomi che hanno tutte le radici semplici. Vogliamo analizzare alcuni aspetti della struttura del discriminante.

1. Aspetti geometrico-algebrici

Vogliamo esibire un polinomio

$$\delta_n(T_0, \dots T_{n-1}) \in \mathbf{K}[T_0, \dots, T_{n-1}]$$

detto il *polinomio discriminante di* $\mathcal{P}_n$, tale che Δ_n sia il suo luogo di zeri:

$$\Delta_n = \{a \in \mathcal{P}_n |\ \delta_n(a) = 0\}\ .$$

Indichiamo con $\mathcal{R}_n = \mathcal{R}_n(\mathbf{K})$ un'altra copia di $\mathbf{K}^n$ con coordinate $\mu = (\mu_1, \dots, \mu_n)$. È definita l'applicazione

$$\phi = \phi_n : \mathcal{R}_n \to \mathcal{P}_n,\ \phi(\mu) := (t - \mu_1) \cdots (t - \mu_n) = a_0(\mu) + a_1(\mu) t + \dots a_{n-1}(\mu) t^{n-1} + t^n\ .$$

Essa è *surgettiva* perché essendo $\mathbf{K}$ algebricamente chiuso ogni $p(t) \in \mathcal{P}_n$ è completamente fattorizzabile su $\mathbf{K}$. Non è iniettiva perché la fattorizzazione non comporta alcun ordinamento privilegiato delle radici. Precisamente ogni $a_j(\mu)$, $j = 0, \dots, n-1$, è una *funzione polinomiale omogenea e simmetrica di grado* j dove "simmetrica" significa che per ogni permutazione σ su l'insieme degli indici $I := \{1, 2, \dots, n\}$,

$$a_j(\mu) = a_j(\sigma(\mu)),\ \sigma(\mu) := (\mu_{\sigma(1)}, \dots, \mu_{\sigma(n)})\ .$$

La funzione $a_j(\mu)$ è esplicitamente definita come segue; per ogni sottoinsieme $J = \{i_1, \dots i_j\}$ di $I := \{1, 2, \dots, n\}$ con j elementi, si considera il monomio

$$m_J(\mu) := \mu_{i_1} \cdots \mu_{i_j}$$

allora

$$a_j(\mu) = (-1)^{n-j} \sum_J m_J(\mu)\ .$$

Per ogni sottoinsieme con due elementi di I, $H = \{r, s\} \subset I$, $r \neq s$, sia P_H il sottospazio vettoriale di $\mathcal{R}_n$ di dimensione $n-1$ (P_H è quindi detto un "iperpiano" di $\mathcal{R}_n$) definito dall'equazione lineare

$$p_H(\mu) = 0,\ p_H(\mu) = \mu_r - \mu_s\ ;$$

l'equazione è ben definita a meno di un segno perché gli elementi di H non sono ordinati, mentre il suo luogo di zeri P_H dipende solo dal sottoinsieme non ordinato H. Definiamo allora l'unione finita di iperpiani di $\mathcal{R}_n$

$$\tilde{\Delta}_n = \cup_H P_H\ .$$

È chiaro che

$$\Delta_n = \phi(\tilde{\Delta}_n) = \cup_H \phi(P_H)\ .$$

Inoltre, definita la funzione polinomiale

$$\tilde{\delta}_n(\mu) = \prod_H p_H(\mu) \in \mathbf{K}[\mu]$$

si ha che

$$\tilde{\Delta}_n = \{\tilde{\delta}_n(\mu) = 0\} \ .$$

Vogliamo far vedere che $\tilde{\delta}_n$ "discende" su un desiderato polinomio δ_n.

Per ogni polinomio $p(t)$ definiamo la sua *derivata* $Dp(t) \in \mathbf{K}[t]$ per mezzo della abituale formula (che ha senso su ogni campo $\mathbf{K}$). La derivata del prodotto di due polinomi verifica l'abituale formula di Leibniz.

Il polinomio $p(t) \in \Delta_n$ se e solo se è della forma

$$p(t) = (t-\lambda)^2 q(t)$$

per qualche radice $\lambda \in \mathbf{K}$, quindi la derivata è della forma

$$Dp(t) = (t-\lambda)b(t)$$

quindi abbiamo immediatamente che affinché $p(t) \in \Delta_n$ è necessario che il polinomio è la sua derivata abbiano una radice in comune. Si mostra che questa condizione è anche sufficiente. Ci sono due modi per sfruttare questa osservazione al fine di determinare un'equazione polinomiale per Δ_n.

(i) Definiamo la funzione polinomiale su $\mathcal{P}_n \times \mathcal{R}_n$ definita da

$$\hat{\delta}_n(p(t), \mu) = \prod_{i=1}^n Dp(\mu_i)$$

Allora $p(t) \in \Delta_n$ se e solo se

$$p(t) = \phi(\mu) \text{ e } \hat{\delta}_n(p(t), \mu) = 0 \ .$$

L'insieme delle soluzioni di questo sistema di due equazioni polinomiali su $\mathcal{P}_n \times \mathcal{R}_n$ si proietta rispettivamente su Δ_n e $\tilde{\Delta}_n$. Se effettuiamo la sostituzione data dalla prima equazione, otteniamo $\hat{\delta}(\phi(\mu), \mu) = 0$ che è un' equazione polinomiale per $\tilde{\Delta}_n$ che è sostanzialmente uguale a quella già messa in evidenza: $\tilde{\delta}(\mu) = 0$. D'altra parte, il polinomio $\hat{\delta}(p(t), \mu) = \hat{\delta}(a, \mu)$ può essere considerato come un polinomio nelle variabili μ con coefficienti polinomiali nelle variabili a; in quanto tale è evidentemente simmetrico nelle μ nel senso detto in precedenza; allora per un fatto noto (non banale) si può esprimere come un polinomio avente come variabili i polinomi simmetrici basici $a_j(\mu)$ che intervengono nella definizione dell'applicazione ϕ. Ne segue che $\hat{\delta}_n(\phi(\mu), \mu)$ può essere interpretato anche come un polinomio $\delta_n(a)$ tale che

$$\Delta_n = \{\delta_n(a) = 0\}$$

come desiderato.

La definizione del polinomio δ_n è piuttosto implicita. Esplicitiamo la discussione per qualche n piccolo.

$n = 2$: $p(t) = a_0 + a_1 t + t^2$, $Dp(t) = a_1 + 2t^2$,

$$\hat{\delta}_2(p, \mu) = (a_1 + 2\mu_1)(a_1 + 2\mu_2) = a_1^2 + 2a_1(\mu_1 + \mu_2) + 4\mu_1\mu_2 = a_1^2 - 2a_1^2 + 4a_0$$

da cui

$$\delta_2(a_0, a_1) = 4a_0 - a_1^2 \ .$$

Si osservi che

$$\delta_2(a_0, a_1) = \det \begin{pmatrix} 1 & a_1 & a_0 \\ 2 & a_1 & 0 \\ 0 & 2 & a_1 \end{pmatrix}$$

$n = 3$: $p(t) = a_0 + a_1 t + a_2 t^2 + t^3$, $Dp(t) = a_1 + 2a_2 t + 3t^2 := a_1 - b(t)$,

$$\hat{\delta}_3(p, \mu) = \prod_{i=1}^3 (a_1 + 2a_2\mu_i + 3\mu_i^2) = \prod_{i=1}^3 (a_1 - b(\mu_1)) = \phi(b(\mu))(a_1),\ b(\mu) = (b(\mu_1), \dots, b(\mu_n)) \ .$$

Si tratta ora di esplicitare i polinomi simmetrici basici nelle $b(\mu)$ per mezzo dei polinomi simmetrici basici nelle μ. Per esempio

$$b(\mu_1) + b(\mu_2) + b(\mu_3) = -(2a_2(\mu_1 + \mu_2 + \mu_3) + 3(\mu_1^2 + \mu_2^2 + \mu_3^2))$$
$$\mu_1^2 + \mu_2^2 + \mu_3^2 = (\mu_1 + \mu_2 + \mu_3)^2 - 2(\mu_1\mu_2 + \mu_1\mu_3 + \mu_2\mu_3)$$

e così via.

(ii) L'espressione determinantale di δ_2 è un caso particolare della seconda via. Per ogni $p(t) \in \mathcal{P}_n$, definiamo la *matrice di Sylvester* $S(p(t)) = S(a)$. Questa è una matrice di formato $(2n-1) \times (2n-1)$ costruita per mezzo dei coefficienti di $p(t)$ e di $Dp(t)$ come suggerito qui sotto nel caso $n = 3$ e sopra per $n = 2$:

$$S(a_0, a_1, a_2) := \begin{pmatrix} 1 & a_2 & a_1 & a_0 & 0 \\ 0 & 1 & a_2 & a_1 & a_0 \\ 3 & 2a_2 & a_1 & 0 & 0 \\ 0 & 3 & 2a_2 & a_1 & 0 \\ 0 & 0 & 3 & 2a_2 & a_1 \end{pmatrix}$$

Affermiamo che $p(t)$ *e* $Dp(t)$ *hanno una radice in comune se e soltanto se* $\det S(p(t)) = 0$. Infatti i due polinomi hanno una radice in comune se e solo se esistono due polinomi non nulli $u(t)$ e $b(t)$ di grado rispettivamente non superiore a $n-2$ e $n-1$ tali che

$$u(t)p(t) + b(t)Dp(t) = 0 \ .$$

Questa equazione può essere interpretata come un sistema lineare omogeneo dove le incognite sono i coefficienti di $u(t)$ e $b(t)$ e per cui $S(p(t))$ è proprio la matrice dei coefficienti. Dunque abbiamo trovato un'altra equazione polinomiale per Δ_n:

$$\Delta_n = \{\det S(a) = 0\} \ .$$

Si può dimostrare che come nel caso $n = 2$

$$\det S(a) = \delta_n(a)$$

da cui risulta per esempio che si tratta di un polinomio nelle a a *coefficienti interi*.

1.1. **Specializzazione del discriminante.** Sia $\mathbf{F} \subset \mathbf{K}$ un sotto-campo non necessariamente algebricamente chiuso (per esempio $\mathbb{R} \subset \mathbb{C}$). $\mathcal{R}_n(\mathbf{F}) \subset \mathcal{R}_n(\mathbf{K})$, $\mathcal{P}_n(\mathbf{F}) \subset \mathcal{P}_n(\mathbf{K})$. L'applicazione ϕ si restrige su $\mathcal{R}_n(\mathbf{F})$ e ha come immagine il sottoinsieme $\mathcal{P}_n^f(\mathbf{F})$ di $\mathcal{P}_n(\mathbf{F})$ formato dai polinomi completamente fattorizzabili su $\mathbf{F}$. $\Delta(\mathcal{P}_n^f(\mathbf{F}))$ è il luogo dei polinomi in $\mathcal{P}_n^f(\mathbf{F})$ che hanno almeno una radice multipla (necessariamente) in $\mathbf{F}$; è l'immagine tramite la restrizione di ϕ della famiglia di iperpiani $\tilde{\Delta}_n(\mathbf{F})$ in $\mathcal{R}_n(\mathbf{F})$ definiti dalla specializzazione su $\mathbf{F}$ delle equazioni lineari che definivano $\tilde{\Delta}_n$ in $\mathcal{R}_n = \mathcal{R}_n(\mathbf{K})$. Poiché il polinomio δ_n definito sopra è a coefficienti interi, esso può essere considerato a coefficienti in $\mathbf{F}$ e possiamo definire

$$\Delta_n(\mathbf{F}) = \{a \in \mathbf{F}; \delta_n(a) = 0\}$$

che coincide con l'intersezione $\Delta_n \cap \mathcal{P}_n(\mathbf{F})$. È chiaro che $\Delta(\mathcal{P}_n^f(\mathbf{F})) \subset \Delta_n(\mathbf{F})$, ma il secondo è più grande perché consiste dei polinomi in $\mathcal{P}_n(\mathbf{F})$ che hanno almeno una radice multipla in $\mathbf{K}$.

1.2. **Il discriminante di $M(n, \mathbf{K})$.** Consideriamo l'applicazione polinomiale

$$\chi_n : M(n, \mathbf{K}) \to \mathcal{P}_n(\mathbf{K}),\ \chi(A) = (-1)^n \det(A - tI_n) := (-1)^n p_A(t) \ .$$

Allora

$$\Delta_n^* := \chi^{-1}(\Delta_n)$$

è detto il *(luogo) discriminante* di $M(n, \mathbf{K})$ ed è definito dall'equazione polinomiale

$$\Delta_n^* = \{\delta_n(\chi(A)) = 0\} \ .$$

L'insieme complementare

$$M(n, \mathbf{K}) \setminus \Delta_n^*$$

è formato dalle matrici il cui polinomio caratteristico ha solo radici semplici e quindi sono *diagonalizzabili*.

2. Aspetti topologici

Faremo qui riferimento ad alcune nozioni solo parzialmente note al lettore e che saranno sviluppate nel corso di Geometria 2.

Poniamo $\mathbf{K} = \mathbb{C}$. Ogni $\mathbb{C}^m$ è munito della distanza euclidea

$$d(z, w) = \sqrt{\sum_{i=1}^{m} (z_i - w_i)\overline{(z_i - w_i)}}$$

e quindi della topologia indotta. Tutte le applicazioni polinomiali considerate in precedenza sono continue. Poiché $\{0\}$ è chiuso in $\mathbb{C}$, ne segue che Δ_n, $\tilde{\Delta}_n$ e Δ_n^* sono chiusi in $\mathcal{P}_n$, in $\mathcal{R}_n$ e in $M(n, \mathbb{C})$ rispettivamente. I loro rispettivi complementari sono insiemi aperti. È immediato verificare che $\mathcal{R}_n \setminus \tilde{\Delta}_n$ è *denso*. Con sforzo di poco maggiore vediamo che anche $\mathcal{P}_n \setminus \Delta_n$ e $M(n, \mathbb{C}) \setminus \Delta_n^*$ sono densi. Nell'ultimo caso vediamo quindi che l'insieme delle matrici con polinomio caratteristico con tutte le radici semplici è un aperto denso: "*genericamente una matrice complessa è diagonalizzabile*". Le matrici non diagonalizzabili sono contenute in Δ_n^* che è "magro" essendo il luogo di zeri di un polinomio. Sia infatti in generale $X = \{z \in \mathbb{C}^m;\ p(z) = 0\}$ un sottoinsieme di $\mathbb{C}^m$ definito da una equazione polinomiale. Se $m = 1$ allora X consiste in un numero finito di punti (e non c'è modo migliore di essere 'magri'). Se $m = 2$, l'intersezione di X con qualsiasi retta affine 'verticale' di equazione $z_1 = c$ è un insieme finito di punti, le radici del polinomio in una variabile $p(c, z_2)$. Se per esempio c è generico per cui le radici di questo polinomio sono semplici, questo resta vero su un intorno abbastanza piccolo U di c in $\mathbb{C}_{z_1}$ e al di sopra di U troviamo in X un numero finito di fogli disgiunti che si proiettano in modo omeomorfo su U. Le stesse considerazioni si applicano per m arbitrario. Sviluppandole, si può mostrare che non solo $\mathbb{C}^m \setminus X$ è un *aperto denso* ma che è anche *connesso* (il caso base è $m = 1$ dove è chiaro che un numero finito di punti non sconnette $\mathbb{C}$).

2.1. **Il caso reale.** Si osserva immediatamente che $\mathcal{R}_n(\mathbb{R}) \setminus \tilde{\Delta}_n(\mathbb{R})$ è un *aperto non connesso* di $\mathcal{R}_n(\mathbb{R})$; il caso base è $n = 2$ dove consideriamo $\mathbb{R}^2$ privato della diagonale. Le componenti connesse sono determinate da opportuni sistemi di disequazioni lineari strette. Anche $\mathcal{P}_n(\mathbb{R}) \setminus \Delta_n(\mathbb{R})$ è un aperto non connesso di $\mathcal{P}_n(\mathbb{R})$. Vogliamo argomentare che esso ha lo stesso numero di componenti connesse di $\mathcal{P}_n(\mathbb{R}) \setminus \Delta(\mathcal{P}_n^f(\mathbb{R}))$. Infatti $\Delta_n(\mathbb{R}) \setminus \Delta(\mathcal{P}_n^t(\mathbb{R}))$ consiste dei polinomi a coefficienti reali che ammettono almeno una radice complessa non reale multipla. È una proprietà ben nota dei polinomi reali che se $p(\alpha) = 0$ per qualche radice $\alpha \in \mathbb{C}$, $\alpha \neq \bar{\alpha}$, di molteplicità m, allora $p(\bar{\alpha}) = 0$ e con la stessa molteplicità. Questo implica che $\Delta_n(\mathbb{R}) \setminus \Delta(\mathcal{P}_n^t(\mathbb{R}))$ è contenuto nell'intersezione con $\mathcal{P}_n(\mathbb{R})$ dell'immagine tramite ϕ in $\mathcal{P}_n(\mathbb{C})$ di una famiglia di intersezioni di due iperpiani in $\mathcal{R}_n(\mathbb{C})$ della forma $P_H \cap \overline{P_H}$ che abbia dimensione $n-2$; la sua intersezione con $\mathcal{R}_n(\mathbb{R})$ non sconnette quest'ultimo. Elaborando questa osservazione ci rendiamo conto che $\Delta_n(\mathbb{R}) \setminus \Delta(\mathcal{P}_n^t(\mathbb{R}))$ non può ulteriormente sconnettere alcuna componente connessa di $\mathcal{P}_n(\mathbb{R}) \setminus \Delta(\mathcal{P}_n^f(\mathbb{R}))$. Il caso $n = 2$ è ben noto fin dalle scuole medie: in questo caso $\Delta(\mathcal{P}_2^f(\mathbb{R})) = \Delta_2(\mathbb{R})$ ed è costituito dai polinomi con una radice reale doppia; $\delta_2 < 0$ è la componente connessa del complementare formata dai polinomi reali con due radici reali distinte; $\delta_2 > 0$ è la componente dei polinomi reali con due radici complesse non reali coniugate.