Sapienza Università: corso di Algebra 3, 2016-17

1. Il teorema di Cayley-Hamilton

Nel seguito indicheremo con R un anello comutativo con unità.

Data una matrice quadrata $A = (a_{ij}) \in M_{n,n}(R)$ a coefficienti in R possiamo calcolare il determinante mediante la solita formula di somma sulle permutazioni

$$\det(A) = \sum_{\sigma} (-1)^{\sigma} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)},$$

o equivalentemente mediante gli sviluppi di Laplace:

$$\det(A) = \sum_{h=1}^{n} b_{ih} a_{hi} = \sum_{h=1}^{n} a_{jh} b_{hj},$$

dove i b_{ij} sono i coefficienti dell'aggiunta classica di A, ossia

$$b_{ij} = (-1)^{i+j} \cdot \text{determinante dopo cancellazione riga } j \text{ e colonna } i\,.$$

Per motivi puramente combinatorici, se la matrice A ha due righe, o due colonne, uguali il determinante si annulla, e da ciò segue che se $B = (b_{ij})$ è l'aggiunta classica di A si hanno le formule $BA = AB = \det(A)I$, dove I denota la matrice identità.

Ogni matrice a coefficienti polinomi $B(t) \in M_{n,n}(R[t])$ può anche essere pensata come un polinomio a coefficienti matrici, ossia possiamo scrivere

$$B(t) = \sum_{i=0}^{N} B_i t^i, \qquad B_i \in M_{n,n}(R).$$

La sostituzione dell'indeterminata t con una qualunque matrice $A \in M_{n,n}(R)$, definisce due applicazione di R-moduli:

$$R[t] \to M_{n,n}(R), \qquad p(t) = \sum a_i t^i \mapsto p(A) = \sum a_i A^i,$$

$$\varphi_A \colon M_{n,n}(R[t]) \to M_{n,n}(R), \qquad \varphi_A\left(\sum_i B_i t^i\right) = \sum_i B_i A^i\,.$$

Notiamo in particolare che per ogni polinomio $h(t) \in R[t]$ si ha $\varphi_A(h(t)I) = h(A)$.

Lemma 1. *Siano $B(t), C(t) \in M_{n,n}(R[t])$. Per ogni matrice $A \in M_{n,n}(R)$ tale che $AC(t) = C(t)A$ vale la formula:*

$$\varphi_A(B(t)C(t)) = \varphi_A(B(t))\,\varphi_A(C(t))\,.$$

Dimostrazione. Se

$$B = \sum_i B_i t^i, \qquad C = \sum_j C_j t^j\,,$$

la condizione $AC(t) = C(t)A$ equivale a $AC_j = C_j A$ per ogni indice j. Allora si ha anche $A^i C_j = C_j A^i$ per ogni i, j e quindi

$$\varphi_A(B(t)C(t)) = \varphi_A(\sum_{i,j} B_i C_j t^{i+j}) = \sum_{i,j} B_i C_j A^{i+j},$$

$$\varphi_A(B(t))\varphi_A(C(t)) = (\sum_i B_i A^i)(\sum_j C_j A^j) = \sum_{i,j} B_i A^i C_j A^j = \sum_{i,j} B_i C_j A^{i+j}.$$

□

Lemma 2. *Siano $A \in M_{n,n}(R)$ e $B(t) \in M_{n,n}(R[t])$ tali che $B(t)(A - tI) = h(t)I$ con $h(t) \in R[t]$. Allora vale $h(A) = 0$.*

Dimostrazione. Siccome A commuta con $A-tI$ e $\varphi_A(A-tI) = A - IA = 0$, dal Lemma 1 segue immediatamente che

$$h(A) = \varphi_A(h(t)I) = \varphi_A(B(t)(A-tI)) = \varphi_A(B(t))\varphi_A(A-tI) = 0\,.$$

□

Teorema 3 (Cayley-Hamilton)**.** *Siano $A \in M_{n,n}(R)$, $p_A(t) = \det(A-tI) \in R[t]$ il suo polinomio caratteristico. Allora $p_A(A) = 0$.*

Dimostrazione. Indichiamo con $B(t) \in M_{n,n}(R[t])$ l'aggiunta classica di $A-tI$. Siccome $B(t)(A-tI) = p_A(t)I$, il tutto segue immediatamente dal Lemma 2. □

Da ora in poi supponiamo che R sia un dominio a fattorizzazione unica; allora anche $R[t]$ è un dominio a fattorizzazione unica.

Lemma 4. *Dato un dominio a fattorizzazione unica R ed una matrice $A \in M_{n,n}(R)$, vi è un unico polinomio monico (coefficiente direttore $= 1$) $q_A(t) \in R[t]$ tale che se $p(t) \in R[t]$ e $p(A) = 0$, allora $q_A(t)$ divide $p(t)$.*

Dimostrazione. Sia $\mathcal{P} \subset R[t]$ l'insieme dei polinomi $p(t)$ non nulli e tali che $q(A) = 0$. Per Cayley-Hamilton tale insieme è non vuoto e contiene almeno un polinomio monico (ad esempio il polinomio caretteristico moltiplicato per $(-1)^n$). Sia $q(t) \in \mathcal{P}$ di grado minimo: a meno di dividere $q(t)$ per il massimo comune divisore dei suoi coefficienti possiamo supporre $q(t)$ primitivo. Dimostriamo che il coefficiente direttore di $q(t)$ è invertibile, da ciò segue l'esistenza in $\mathcal{P}$ di un polinomio monico di grado minimo. Se

$$q(t) = a_0 + a_1 t + \cdots + a_d t^d, \qquad d \leq n,$$

allora, $a_d p_A(t) - (-1)^n t^{n-d} q(t)$ ha grado $< n$. Ripetendo tale procedura altre $n-d$ volte, troviamo due polinomi $r(t), h(t)$ tali che

$$a_d^{n-d+1}(-1)^n p_A(t) = h(t)q(t) + r(t), \qquad \deg(r(t)) < d\,.$$

Siccome $q(A) = p_A(A) = 0$, se segue che $r(A) = 0$ e quindi che $r(t) = 0$ essendo $q(t)$ di grado minimo tra quelli che annullano A. Dunque $q(t)$ divide $a_d^{n-d+1} p_A(t)$ e siccome $q(t)$ è primitivo, deve dividere $p_A(t)$. Adesso il coefficiente direttore di $p_A(t)$ è invertibile e quindi anche a_d è invertibile. Dividendo $q(t)$ per a_d si ottiene un polinomio monico.

Sia $p(t)$ un qualunque polinomio tale che $p(A) = 0$. L'algoritmo della divisione Euclidea permette di scrivere $p(t) = k(t)q(t) + u(t)$ con il grado di $u(t)$ minore di s. Siccome $u(A) = 0$ deve essere necessariamente $u(t) = 0$ e quindi $q(t)$ divide $p(t)$. Se ne esistessero due, diciamo $q(t)$ e $\tilde{q}(t)$, ciascuno divide l'altro ed essendo entrambi monici devono coincidere. □

Definizione 5. Il polinomio $q_A(t)$ introdotto nel lemma precedente si chiama **polinomio minimo** di A.

Abbiamo già osservato che il polinomio minimo divide il polinomio caratteristico; il prossimo teorema ci dice qual è il quoziente.

Teorema 6. *Siano R un dominio a fattorizzazione unica, $A \in M_{n,n}(R)$, $p_A(t) = \det(A-tI) \in R[t]$ il suo polinomio caratteristico, $q_A(t)$ il suo polinomio minimo e $B(t) \in M_{n,n}(R[t])$ l'aggiunta classica di $A-tI$. Allora il quoziente $p_A(t)/q_A(t)$ è uguale al massimo comune divisore dei coefficienti di $B(t)$.*

Dimostrazione. Se

$$q_A(t) = a_0 + a_1 t + \cdots + a_d t^d, \qquad a_d = 1\,,$$

poiché tI commuta con $A - tI$ possiamo scrivere:

$$\begin{aligned}
0 &= \sum_{i=0}^{d} a_i A^i = \sum_{i=0}^{d} a_i ((A - tI) + tI)^i \\
&= \sum_{i=0}^{d} a_i \sum_{j=0}^{i} \binom{i}{j} (tI)^{i-j} (A - tI)^j = \sum_{j=0}^{d} (A - tI)^j \sum_{i=j}^{d} a_i \binom{i}{j} t^{i-j} \\
&= \sum_{i=0}^{d} a_i \binom{i}{0} t^i + (A - tI) \sum_{j=1}^{d} (A - tI)^{j-1} \sum_{i=j}^{d} a_i \binom{i}{j} t^{i-j} \\
&= q_A(t) I - (A - tI) D(t), \qquad D(t) \in M_{n,n}(R[t]) \, .
\end{aligned}$$

Moltiplicando prima per $B(t)$ e dividendo poi per $q_A(t)$ la relazione $q_A(t)I = (A - tI)D(t)$ si ottiene

$$q_A(t)B(t) = B(t)(A - tI)D(t) = p_A(t)D(t), \qquad B(t) = \frac{p_A(t)}{q_A(t)} D(t) \, ,$$

e questo dimostra che $p_A(t)/q_A(t)$ divide tutti i coefficienti di $B(t)$.

Viceversa, se $s(t) \in R[t]$ è il massimo comune divisore dei coefficienti di $B(t)$ si ha $B(t) = s(t)C(t)$, con $C(t) \in M_{n,n}(R[t])$, e quindi

$$s(t)C(t)(A - tI) = p_A(t) I$$

da cui $s(t)$ divide $p_A(t)$. Se $p_A(t)/s(t) = h(t)$ otteniamo $C(t)(A - tI) = h(t)I$; per il Lemma 2 si ha $h(A) = 0$, quindi $q_A(t)$ divide $h(t)$ e di conseguenza $s(t)$ divide $p_A(t)/q_A(t)$. □