Bollettino Unione Matematica Italiana

Lionello Cantoni

Serie di potenze e logaritmo di una matrice.

Bollettino dell'Unione Matematica Italiana, Serie 3, Vol. **10** (1955), n.3, p. 376–381.

Zanichelli

<http://www.bdim.eu/item?id=BUMI_1955_3_10_3_376_0>

Serie di potenze e logaritmo di una matrice.

Nota di Lionello Cantoni (a Bologna)

Sunto. - *Come al n. 1.*

1. Per suggerimento del Prof. Cherubino, ho risolto qui alcuni semplici problemi di calcolo di matrici: nei n. 2,3 determino la segnatura di una serie di potenze di matrice facendo qualche applicazione (n. 3) del risultato conseguito, relativamente a questioni di diagonalizzabilità. Nei n. 4,5 ho invece esteso un recente teorema di Cherubino sul logaritmo di una matrice [1], dimostrando che se A è una matrice non singolare, il logaritmo principale di A si esprime in modo univoco mediante un polinomio in A il cui grado è pari a quello dell'equazione minima di A diminuito di uno.

1. - Segnatura di una serie di potenze di una matrice e applicazioni.

2. Sia A una matrice di ordine n possedente come radici caratteristiche i numeri $\alpha_1, \alpha_2, \ldots, \alpha_m$ $(m \leq n)$ e sia μ_j la molteplicità di α_j per A $(j = 1, 2, \ldots, m)$ [2]. Sia inoltre:

$$f(z) = a_0 + a_1 z + \ldots = \sum_{i=0}^{\infty} a_i z^i \tag{1}$$

[1] La nozione di logaritmo di una matrice, si deve ad H. Richter che l'ha introdotta nel suo lavoro: *Zum Logarithmus einer Matrix*, Arch. der Math., 2, 360-363 (1949-50). La definizione di questo A. è stata ripresa e semplificata da S. Cherubino nel recente lavoro: *Permutabilità e logaritmi delle matrici*, Rend. Mat. e Appl. (5), 13 (1954) ed è a quest'ultima che mi riferirò sempre nel seguito. Il teorema che ho richiamato nel testo si trova al n. 7, § 3 dell'opera cit.

[2] Mi riferirò esclusivamente, qui e nel seguito alla segnatura (e alla forma canonica) come la dà S. Cherubino nei suoi lavori. Si veda ad es. S. Cherubino, *Segnatura, divisori elementari e forme canoniche di una matrice*, Boll. U. M. I. (2) 4, 1-11, (1941). S. Cherubino, *Sulla riduzione delle matrici a forma canonica*. Note I e II; Atti Acc. Lincei (Rend.) (6), 23, 478-482; 647-653, (1936).

una serie di potenze per cui converge anche la:

$$(2) \qquad f(A) = a_0 I + a_1 A + \dots = \sum_{i=0}^{\infty} a_i A^i \qquad (A^0 = I) \text{ [3]};$$

è ben noto che la matrice $f(A)$ ha per radici caratteristiche i numeri $f(\alpha_j)$ (potendo però essi non essere tutti distinti).

Converrà disporre le α_j in λ gruppi tali che la (1) assuma valori uguali su due numeri appartenenti allo stesso gruppo e valori distinti su numeri appartenenti a gruppi diversi. Siano ad es. $\alpha_1, \alpha_2, \dots, \alpha_{n_1}$ le componenti del primo gruppo e indichiamo con f_1 il valore (comune) che la (1) assume su di essi. Evidentemente, la molteplicità di f_1 per la (2) è data da $\bar{\mu}_1 = \sum_{i=1}^{n_1} \mu_i$. Sia poi H una matrice canonizzante la A e C_1 la componente della forma canonica di A relativa al primo gruppo di radici caratteristiche; poniamo cioè:

$$H^{-1}AH = C = \left[\begin{array}{c|c} C_1 & 0 \\ \hline 0 & C' \end{array}\right] \qquad \text{ove} \qquad C_1 = \left[\begin{array}{c|c|c|c} C_{\alpha_1} & 0 & \dots & 0 \\ \hline 0 & C_{\alpha_2} & \dots & 0 \\ \hline \dots & \dots & \dots & \dots \\ \hline 0 & 0 & \dots & C_{\alpha_{n_1}} \end{array}\right]$$

allora, come è ben noto, si ha:

$$H^{-1}f(A)H = f(C) = \left[\begin{array}{c|c} f(C_1) & 0 \\ \hline 0 & f(C') \end{array}\right]$$

e per determinare la segnatura di f_1 basterà esaminare la nullità delle successive potenze della matrice $f(C_1) - f_1 I_{n_1}$ e poichè quest'ultima è composta dalle matrici $f(C_{\alpha_i}) - f_1 I_{\mu_i}$ $(i = 1, 2, \dots, n_1)$, si ha che la nullità in questione sarà data, per ciascuna potenza, dalla somma delle nullità delle stesse potenze delle singole componenti. Ora, esaminando la nullità ad es. di $f(C_{\alpha_1}) - f_1 I_{\mu_1}$, ed indicando con $(h_1^{(1)}, h_2^{(1)}, \dots, h_{i_1}^{(1)})$ la segnatura di α_1 per A, si vede subito che essa eguaglia $h_{i_1}^{(1)}$ se $f'(\alpha_1) \neq 0$, $h_{i_1}^{(1)} + h_{i_1-1}^{(1)}$ se $f'(\alpha_1) = 0$ $f''(\alpha_1) \neq 0$ e in generale $h_{i_1}^{(1)} + h_{i_1-1}^{(1)} + \dots + h_{i_1-r+1}^{(1)}$ se è $f'(\alpha_1) = f''(\alpha_1) = \dots = f^{(r-1)}(\alpha_1) = 0$; $f^{(r)}(\alpha_1) \neq 0$, e che, nell'ipotesi generale in cui ci siamo posti, la nullità della potenza s-esima della matrice $f(C_{\alpha_1}) - f_1 I_{\mu_1}$, supera la nullità della potenza immediatamente inferiore della quantità:

$$h_{i_1-(s-1)r} + h_{i_1-(s-1)r-1} + \dots + h_{i_1-sr+1}$$

(naturalmente quando $sr \leq i_1$).

(3) Per la definizione e per le condizioni di convergenza di una serie di potenze di una matrice, si veda ad es. C. C. Mac Duffee: *The Theory of matrices*, Berlino, Springer Ed. (1933).

Ne segue subito il teorema seguente:

Sia A *una matrice complessa di ordine* n *ed* $f(z)$ *una serie di potenze della variabile* z, *convergente per* $z = A$. *Raggruppando le radici caratteristiche di* A *in insiemi* $G_1, G_2, \ldots, G_\lambda$ *tali che la* f *assuma lo stesso valore sui punti dello stesso insieme e valori diversi sui punti non appartenenti allo stesso insieme e indicando con* $\alpha_1^{(j)}, \alpha_2^{(j)}, \ldots, \alpha_{\nu_j}^{(j)}$ *le componenti dell'insieme* G_j $(j = 1, 2, \ldots, \lambda)$ *e con* f_j *il relativo valore della* $f(z)$, *si ha che la molteplicità della radice* f_j *per* $f(A)$ *è data della somma delle molteplicità delle* $\alpha_t^{(j)}$, $t = 1, 2, \ldots, \nu_j$ *per* A; *indicando poi con* $(h_1^{(j,t)}, h_2^{(j,t)}, \ldots, h_{i_{(j,t)}}^{(j,t)})$ *la segnatura della radice* $\alpha_{(t)}^{(j)}$ *per* A *e con* $s_{(j,t)}$ *la molteplicità della predetta radice per la derivata prima* $f'(z)$, *di* $f(z)$, (*intendendo di prendere* $s_{(j,t)} = 0$ *se* $f'(\alpha_t^j) \neq 0$), *eseguendo il quoziente dell'indice* $i_{(j,t)}$ *di* α_t^j *per* $s_{(j,t)}$:

$$i_{(j,t)} = (s_{(j,t)} + 1) \cdot q_{(j,t)} + r_{(j,t)}; \quad (r_{(j,t)} \leq s_{(j,t)})$$

e ponendo:

$$\begin{aligned}
a_0^{(j,t)} &= h_1^{(j,t)} + h_2^{(j,t)} + \ldots + h_{r_{(j,t)}}^{(j,t)}, \\
a_1^{(j,t)} &= h_{r_{(j,t)}+1}^{(j,t)} + h_{r_{(j,t)}+2}^{(j,t)} + \ldots + h_{r_{(j,t)}+s_{(j,t)}+1}^{(j,t)} \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
a_{q_{(j,t)}}^{(j,t)} &= h_{i_{(j,t)}-s_{(j,t)}}^{(j,t)} + h_{i_{(j,t)}-s_{(j,t)}+1}^{(j,t)} + \ldots + h_{i_{(j,t)}}^{(j,t)},
\end{aligned}$$

la segnatura di f_j *per* $f(A)$, *scritti gli intieri in ordine non crescente risulta:*

$$\left(\sum_{t=1}^{\nu_j} a_{q_{(j,t)}}^{(j,t)}, \ \sum_{t=1}^{\nu_j} a_{q_{(j,t)}-1}^{(j,t)}, \ldots \right).$$

Ponendo poi $J_{(j,t)} = q_{(j,t)}$, *oppure* $J_{(j,t)} = q_{(j,t)} + 1$, *a seconda che* $i_{(j,t)}$ *è o no divisibile per* $s_{(j,t)} + 1$, *si ha che l'indice di* f_j *è il maggiore dei* ν_j *numeri* $J_{(j,t)}$.

3. Il teorema enunciato al n. precedente, permette di stabilire a quali condizioni deve soddisfare una matrice A affinchè, data una serie di potenze $f(z)$, la matrice $f(A)$ risulti diagonalizzabile (quando esiste).

Se α_s è una radice caratteristica qualunque di A e $r_s - 1$ è la molteplicità di α_s per $f'(z)$, (intendendo al solito, di porre $r_s = 1$ se α_s non annulla $f'(z)$), si ha che l'indice di α_s deve soddisfare alla condizione:

$$i_s \leq r_s \tag{3}$$

e viceversa. Passando attraverso la forma canonica, nella quale è facile tener conto della (3), è quindi possibile costruire tutte le matrici A che soddisfano al problema.

Se in particolare si ha $f(z) = z^r$ (r intiero ≥ 2), si ha evidentemente che nessun numero $\neq 0$ è radice di $f'(z)$, mentre lo zero è radice di ordine $r-1$ per la stessa derivata. Da questo e dalla (3) discende che:

Condizione necessaria e sufficiente affinchè una matrice A *possegga potenze ad esponente intero relativo* $\neq 0$ [1] *diagonalizzabili, è che le radici caratteristiche non nulle di* A *abbiano tutte indice uno.*

Si ha precisamente che le potenze ad esponente intiero diagonalizzabili di A sono tutte e sole quelle il cui esponente non è, in valore assoluto, inferiore all'indice della radice 0 per la matrice considerata.

Si può, al contrario di quanto è stato fatto in questo numero, supporre data la A e ricercare quali sono le serie di potenze $f(z)$ per cui $f(A)$ risulti diagonalizzabile. Se $\alpha_1, \alpha_2, \ldots, \alpha_m$ sono le radici caratteristiche di A aventi per indici rispettivi $i_1, i_2, \ldots, i_m$, dal teorema precedente discende subito che la condizione necessaria e sufficiente affinchè $f(A)$ (quando esiste) risulti diagonalizzabile, è che risulti

$$f^{(r)}(\alpha_s) = 0 \qquad \begin{array}{l} s = 1, 2, \ldots, m \\ r = 1, 2, \ldots, i_s - 1 \end{array}$$

ove si è indicato, come d'abitudine, con $f^{(r)}(z)$ la derivata r-esima della $f(z)$.

2. - Sul logaritmo principale di una matrice.

4. Siano date due matrici A e B di ordine n, non degeneri e simili, esista cioè una matrice S (non degenere e di ordine n) per cui:

$$S^{-1}AS = B. \tag{4}$$

Nelle nostre ipotesi, esistono i logaritmi principali delle matrici A e B e dimostreremo che si ha:

$$S^{-1}\lg(A)S = \lg(B) = \lg(S^{-1}AS) \tag{5}$$

cioè che i logaritmi principali delle due matrici sono fra loro simili come le due matrici stesse.

[1] Si tenga presente che se A^k (k intiero non nullo e A non singolare) è diagonalizzabile, lo è anche A^{-k}.

Infatti, supponiamo dapprima che A e (quindi anche) B siano diagonalizzabili, e indichiamo con D la comune forma canonica. Se H è una matrice che canonizza la A, si ha subito dalla (4), che $S^{-1}H$, canonizza la B e quindi, per la definizione stessa di logaritmo principale si ha:

$$\lg(B) = S^{-1}H \lg(D)H^{-1}S = S^{-1}[H \lg(D)H^{-1}]S = S^{-1} \lg(A)S$$

cioè la (3).

Se invece A e B non sono diagonalizzabili, decomponiamole, assieme alla comune forma canonica C, nelle relative parti principale e complementare [5]. Poniamo cioè

$$A = A_D + A_J; \quad B = B_D + B_J; \quad C = D + J.$$

Si ha subito:

$$B_D = (S^{-1}H)D(S^{-1}H)^{-1} = S^{-1}HDH^{-1}S = S^{-1}A_DS \tag{6}$$

e in modo analogo:

$$B_J = S^{-1}A_JS.$$

Dalle ultime relazioni, si deduce subito che:

$$\lg[I + B_JB_D^{-1}] = S^{-1} \lg[I + A_JA_D^{-1}]S$$

e da questa e dalla (6) segue subito la (3).

5. Sia A una matrice non degenere di ordine n avente radici caratteristiche distinte: $\alpha_1, \alpha_2, \ldots, \alpha_m$ con indici $i_1, i_2, \ldots, i_m$. Indichiamo con $\mu = \sum_{j=0}^{m} i_j$ il grado dell'equazione minima di A e dimostriamo che:

Esiste uno ed un solo polinomio in A, *di grado* $\mu - 1$ *che dia il logaritmo* (*principale*) *di* A.

Infatti se indichiamo con

$$g(z) = \sum_{j=0}^{\mu-1} a_j z^{\mu-1-j}$$

un polinomio di grado $\mu - 1$ in z, si ha che le due equazioni

$$g(A) = \lg(A) \qquad g(C) = \lg(C) \tag{7}$$

sono, per quando dimostrato al precedente numero, equivalenti. Ponendo in evidenza, nella C, le m componenti C_s relative alle

[5] Si veda, per questa decomposizione il § 1 della Nota di S. Cherubino citata in (1).

m radici caratteristiche di A, si ha che la $(7)_2$ è equivalente alle relazioni:

$$(8) \qquad g(C_s) = \lg(C_s) \qquad s = 1, 2, \ldots, m.$$

Per convincersene, basta decomporre ciascuna C_s nelle relative parti principali e complementare ed osservare che la matrice $\lg(C)$ è composta dalle m matrici $\lg(C_s)$. Si ha poi, per definizione stessa del logaritmo:

$$\lg(C_s) = \lg(D_s) + \lg[I + J_s D_s^{-1}] =$$

$$= \left[\begin{array}{c|c|c|c|c}
\lg \alpha_s I_{h_1^{(s)}} & \frac{1}{\alpha_s} I_{h_1^{(s)}} \; \vdots \; 0 & -\frac{1}{2\alpha_s^2} I_{h_1^{(s)}} \;\; 0 & \cdots\cdots & \frac{(-1)^{i_s}}{(i_s-1)\alpha_s^{i_s-1}} I_{h_1^{(s)}} \; \vdots \; 0 \\
\hline
0 & \lg \alpha_s I_{h_1^{(s)}} & \frac{1}{\alpha_s} I_{h_2^{(s)}} \; \vdots \; 0 & \cdots\cdots & \frac{(-1)^{i_s-1}}{(i_s-2)\alpha_s^{i_s-2}} I_{h_2^{(s)}} \; \vdots \; 0 \\
\hline
\cdots\cdots & \cdots\cdots & \cdots\cdots\cdots & \cdots\cdots & \cdots\cdots\cdots\cdots \\
\cdots\cdots & \cdots\cdots & \cdots\cdots\cdots & \cdots\cdots & \cdots\cdots\cdots\cdots \\
\hline
0 & 0 & 0 & \cdots\cdots & \lg \alpha_s I_{h_{i_s}^{(s)}}
\end{array}\right]$$

La (8) si traduce perciò nelle μ equazioni lineari nelle a_j $(j = 0, 1, \ldots, \mu - 1)$:

$$g(\alpha_s) = \lg \alpha_s \qquad s = 1, 2, \ldots, m$$

$$\frac{1}{k!} g^{(k)}(\alpha_s) = \frac{(-1)^{k-1}}{k\alpha_s^k} \qquad k = 1, 2, \ldots, i_s - 1$$

cioè nelle:

$$g(\alpha_s) = \lg \alpha_s \qquad s = 1, 2, \ldots, m$$

$$g^{(k)}(\alpha_s) = \left[\frac{d^k \lg x}{dx^k}\right]_{x=\alpha_s} \qquad k = 1, 2, \ldots, i_s - 1$$

che sono tutte linearmente indipendenti. Infatti il determinante del sistema da esse formato vale, come si prova con un semplice calcolo (oppure notando che esso coincide, a meno di un fattore non nullo con il valore di un notissimo WRONSKIANO [6]):

$$\Delta = \prod_{k=1}^{m-1} \left[\prod_{s=k+1}^{m} (\alpha_s - \alpha_k)^{i_s i_k}\right]$$

ed è quindi diverso da zero poichè gli m numeri α_s sono per ipotesi tutti distinti. Ne segue il teorema.

[6] Si veda ad es. G. TORELLI, *Lezioni di calcolo infinitesimale*, Napoli, Tip. Acc. Sc. Fis. Mat. 1921, § 21, cap. XXI.