Bollettino Unione Matematica Italiana

Mauro Picone

Sulla teoria delle matrici nel corpo complesso.

Bollettino dell'Unione Matematica Italiana, Serie 3, Vol. **13** (1958), n.1, p. 1–6.

Zanichelli

<http://www.bdim.eu/item?id=BUMI_1958_3_13_1_1_0>

Bollettino dell'Unione Matematica Italiana, Zanichelli, 1958.

SEZIONE SCIENTIFICA

BREVI NOTE

Sulla teoria delle matrici nel corpo complesso.

Nota di MAURO PICONE (a Roma)

Sunto. - *Si estende il teorema vettoriale di* BESSEL *alle matrici e se ne deduce una limitazione superiore, non restringibile, del rango di un numero caratteristico di una matrice. Si osserva che il teorema di* SYLVESTER *della teoria delle matrici sussiste per un'operazione che sia funzione razionale di un'arbitraria operazione distributiva la quale trasformi in sè uno spazio di elementi, lineare nel corpo complesso.*

Summary. - BESSEL'*s inequality for vectors is extended to matrices, and a bound for the rank of an eigenvalue of a matrix is deduced, which is proved to be unrestrictable.*

SYLVESTER'*s theorem of matrix theory is shown to old true for operations, which be rational functions of an arbitrary distributive operation, changing in itself a space linear on the complex field.*

(da una lettera (*) al prof. SALVATORE CHERUBINO)

. . . .

1. Anche tu, nel tuo aureo libro sul calcolo delle matrici (1), segui l'uso invalso di indicare col simbolo $|a|$ il determinante di una matrice quadrata a, mentre che, se la matrice ha ordine uno, detto simbolo *deve* essere destinato ad indicare il modulo dell'unico elemento a costituente la matrice, la quale ha per determinante a (2). Inoltre se la matrice a è un vettore riga od un vettore colonna, il simbolo stesso *deve* indicarne il modulo. Ora ad una matrice a qualsivoglia, di elementi a_{ij}, con m righe e n colonne, cioè, come dirò, d'ordine (m, n), si può associare il vettore ad mn

(*) Redatta nell'Istituto Nazionale per le Applicazioni del Calcolo.

(1) SALVATORE CHERUBINO, *Calcolo delle matrici* [nelle Monografie matematiche del Consiglio Nazionale delle Ricerche, Edizioni Cremonese, Roma (1957)].

(2) L'inconveniente può essere eliminato, riservando, come tu fai, le lettere maiuscole alla designazione delle matrici.

componenti (reali o complesse, prese in un certo ordine)

$$a_{ij}(i=1, 2, \dots, m; \ j=1, 2, \dots, n),$$

e anche perciò, a mio avviso, é più coerente, indicare con $|a|$, la quantità non negativa (norma della matrice, secondo FROBENIU

(2) $$|a| = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2},$$

e chiamare tale quantità, nulla allora e allora soltanto che lo sia la matrice, il *modulo della matrice.*

Per indicare poi il determinante di una matrice quadrata a, adotterei sempre la notazione

$$\det a.$$

Si dimostra immediatamente che se a e b sono due matrici, rispettivamente di ordine (m, n) e (n, p), sussistono le relazioni

(2) $$|ab| \leq |a| |b|,$$

e, se a e b sono due matrici simili,

(3) $$||a| - |b|| \leq |a+b| \leq |a| + |b|,$$

precisamente come per il prodotto scalare ab e per la somma $a+b$ di due vettori simili a e b. Ne segue subito, com'è noto, che se μ é un *numero caratteristico* per una matrice quadrata a, risulta

(4) $$|\mu| \leq |a|,$$

limitazione questa più ristretta di quella di HIRSCH, indicata a pag. 123 del tuo libro. Se, infatti, per un vettore non nullo x si ha $ax = \mu x$, se ne deduce $|\mu x| = |\mu| |x| = |ax| \leq |a| |x|$, donde la (4), per essere $|x| > 0$.

Pare però opportuno introdurre altri moduli per una matrice a, per i quali sussistano sempre le (2) e le (3), dando a quello definito dalla (1) il nome di *modulo euclideo.*

Chiamerei, per esempio, *modulo a destra* di una matrice a di ordine (m, n) il numero

$$|a|_d = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |a_{ij}|,$$

modulo a sinistra il numero

$$|a|_s = \max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}|.$$

Per un vettore colonna x a n componenti $x_1, x_2, \dots, x_n$, porrei dunque, rispettivamente per i suoi moduli a destra e a sinistra

$$|x|_d = \max_{1 \leq i \leq n} |x_i| \quad , \quad |x|_s = \sum_{i=1}^{n} |x_i| .$$

Per i moduli a destra e a sinistra sussistono sempre le (2) e (3) e se ne deducono, immediatamente come sopra, per un numero caratteristico μ di una matrice quadrata a, le limitazioni ulteriori

$$|\mu| \leq |a|_d \quad , \quad |\mu| \leq |a|_s \text{ (}^3\text{)}.$$

2. Introdotto il modulo euclideo $|a|$ per una matrice a d'ordine (m, n) si può enunciare il seguente teorema di BESSEL.

Se a *è una matrice d'ordine* (m, n), $x^{(1)}, x^{(2)}, \ldots, x^{(p)}$, *con* $p \leq m$, *sono vettori a* m *componenti formanti un sistema ortonormale e* $y^{(1)}, y^{(2)}, \ldots, y^{(q)}$, *con* $q \leq n$, *vettori a* n *componenti, formanti, essi pure, un sistema ortonormale, sussiste la relazione*

$$|a|^2 \geq \sum_{h=1}^{p} \sum_{k=1}^{q} |\bar{x}^{(h)} a y^{(k)}|^2, \tag{5}$$

il segno di eguaglianza verificandosi quando e solo quando esistono pq *numeri* $\gamma_{hk} (h = 1, 2, \ldots, p;\ k = 1, 2, \ldots, q)$, *tali da aversi*

$$a_{ij} = \sum_{h=1}^{p} \sum_{k=1}^{q} \gamma_{hk} x_i^{(h)} \bar{y}_j^{(k)} \qquad (i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n), \tag{6}$$

designando le $x_i^{(h)}$ $(i = 1, 2, \ldots, m)$ *le componenti di* $x^{(h)}$ *e le* $y_j^{(k)}$ $(j = 1, 2, \ldots, n)$ *quelle di* $y^{(k)}$.

Concepita, infatti, la matrice a come il vettore a mn componenti a_{ij}, introdotti i pq vettori

$$u^{(h,k)} \qquad (h = 1, 2, \ldots, p;\ k = 1, 2, \ldots, q),$$

a mn componenti

$$u_{ij}^{(h,k)} = \bar{x}_i^{(h)} y_j^{(k)} \qquad (i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n),$$

i quali formano un sistema ortonormale, per essere

$$u^{(h,k)} u^{(h',k')} = (x^{(h)} x^{(h')})(y^{(k)} y^{(k')}),$$

sussiste la relazione di BESSEL:

$$|a|^2 \geq \sum_{h=1}^{p} \sum_{k=1}^{q} \left| \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \bar{x}_i^{(h)} y_j^{(k)} \right|^2,$$

(3) Già osservate da FROBENIUS.

cioè la (5), il segno di eguaglianza verificandosi solo se il vettore a è una combinazione lineare dei vettori $u^{(h,k)}$, se cioè sussistono le (6).

Il teorema enunciato non è senza corollari. Uno ne è il seguente, che non mi risulta già noto.

Per il rango r *di un numero caratteristico* μ *di una matrice quadrata* a, *cioè per il numero* r *di vettori* x, *linearmente indipendenti, verificanti l'equazione* $\mathrm{ax} = \mu\mathrm{x}$, *sussiste la limitazione*

$$r \leq \frac{|a|^2}{|\mu|^2} \tag{7}$$

il segno di eguaglianza verificandosi quando e solo quando esiste un sistema ortonormale $\mathrm{x}^{(1)}, \mathrm{x}^{(2)}, \dots, \mathrm{x}^{(\mathrm{r})}$ *di* ($\mathrm{r} \leq n$) *vettori per i quali si abbia*

$$a_{ij} = \mu \sum_{h=1}^{r} x_i^{(h)} \bar{x}_j^{(h)} \qquad (i, j = 1, 2, \dots, n). \tag{8}$$

Se infatti, i vettori $x^{(1)}, x^{(2)}, \dots, x^{(r)}$ costituiscono un sistema ortonormale di r ($r \leq n$) soluzioni dell'equazione $ax = \mu x$ si ha, in virtù della (5),

$$|a|^2 \geq \sum_{h=1}^{r} \sum_{k=1}^{r} |\bar{x}^{(h)} a x^{(k)}|^2 = |\mu|^2 \sum_{h=1}^{r} \sum_{k=1}^{r} |\bar{x}^{(h)} x^{(k)}|^2 = r\,|\mu|^2.$$

Il segno di eguaglianza può aver luogo soltanto quando si abbia, a norma della (6), per certi numeri γ_{hk},

$$a_{ij} = \sum_{h=1}^{r} \sum_{k=1}^{r} \gamma_{hk} x_i^{(h)} \bar{x}_j^{(k)}.$$

Le eguaglianze $ax^{(l)} = \mu x^{(l)}$ forniscono allora le seguenti

$$\sum_{h=1}^{r} \gamma_{hl} x_i^{(h)} = \mu x_i^{(l)} \qquad (l = 1, 2, \dots, r;\ i = 1, 2, \dots, n),$$

donde $\gamma_{hl} = 0$, quando $h \neq l$, $\gamma_{hh} = \mu$, e quindi le (8).

Si verifica poi immediatamente che la matrice quadrata a, con gli elementi dati dalla (8), ammette il numero caratteristico μ, di rango r, e l'equazione $ax = \mu x$ le soluzioni $x^{(1)}, x^{(2)}, \dots, x^{(r)}$, avendosi $|a|^2 = r\,|\mu|^2$. Per $r < n$, la matrice stessa ha lo zero per ulteriore numero caratteristico, di rango $n - r$, e per $r = n$, il solo μ.

3. Osserverò, infine, che l'enunciato del teorema di Sylvester, dato a pag. 118 del tuo libro, non è completo, poichè quel teorema afferma che, per ogni polinomio $\varphi(z)$, non soltanto il coinsieme di $\varphi(z)$, relativo all'insieme dei numeri caratteristici di a, è — come mi pare che si limiti a dire quell'enunciato — *contenuto* nell'insieme dei numeri caratteristici di $\varphi(a)$, ma addirittura che il detto coinsieme *coincide* con quest'ultimo. Ciò che, d'altronde, è immediato, perchè se ν è numero caratteristico di $\varphi(a)$, dette $\mu_1, \mu_2, \dots, \mu_s$ le radici dell'equazione $\varphi(\mu) = \nu$, coi rispettivi ordini di molteplicità $k_1, k_2, \dots, k_s$, avendosi, detto γ il coefficiente della più alta potenza di z in $\varphi(z)$ e indicata con e la matrice unità,

$$\varphi(a) - \nu e = \gamma \prod_{i=1}^{s} (a - \mu_i e)^{k_i}, \tag{9}$$

risulta

$$\det [\varphi(a) - \nu e] = \gamma^n \prod_{i=1}^{s} [\det (a - \mu_i e)]^{k_i}.$$

Il teorema di SYLVESTER, vale la pena di osservarlo, sussiste, con pari immediatezza, per la più generale operazione distributiva a, definita per gli elementi x di un qualsivoglia spazio S, lineare nel corpo complesso, la quale abbia coinsieme contenuto in S. Dirò che l'operazione *a ammette un autoelemento*, se esiste un elemento non nullo x di S, che chiamerò un *autoelemento per l'operazione* per il quale ax è nullo; date due operazioni a_2 e a_1 della specie indicata, che l'operazione a_2 è *multipla* della a_1, se esiste un'operazione a della stessa specie per la quale si abbia $a_2 = aa_1$ oppure $a_2 = a_1 a$; che il numero μ, reale o complesso, è *un numero caratteristico per l'operazione* a, se, detta e l'operazione identica, l'operazione $a - \mu e$ ammette un autoelemento. Sussiste il teorema:

Se $a_1, a_2, \dots, a_n$ *sono operazioni della specie detta, ed il loro prodotto* $a_n a_{n-1} \dots a_2 a_1$ *ammette un autoelemento, una almeno di esse deve ammetterlo.*

Infatti, se $a_2 a_1$ ammette un autoelemento x e $a_1 x$ non è nullo, $a_1 x$ è un autoelemento per a_2. Supponiamo, allora, che sia stato dimostrato il teorema per il prodotto $a_{n-1} a_{n-2} \dots a_2 a_1$ e sia x un autoelemento per il prodotto $a_n a_{n-1} \dots a_2 a_1$. Se x è autoelemento per il prodotto $a_{n-1} a_{n-2} \dots a_2 a_1$, esso dovrà esserlo per una delle operazioni a_i, con $i \leq n-1$; nell'altro caso, l'elemento non nullo $(a_{n-1} a_{n-2} \dots a_2 a_1)x$ è un autoelemento per a_n.

Ciò posto, se μ è un numero caratteristico per l'operazione a, lo è anche per l'operazione $\varphi(a)$, poichè $\varphi(a) - \varphi(\mu)e$ è multipla di $a - \mu e$; se ν è un numero caratteristico per $\varphi(a)$, dalla (9) si deduce, in forza del teorema osservato, che una delle radici dell'equazione $\varphi(\mu) = \nu$ è un numero caratteristico per a. Cfr.

GAETANO FICHERA, *Lezioni sulle trasformazioni lineari*, Vol. I (Istituto Matematico dell'Università di Trieste) pp. 240-241, dove lo stesso ragionamento è impiegato per dimostrare il teorema di SYLVESTER in ipotesi superflue per la validità del teorema e del ragionamento, necessarie per le applicazioni fattene.

Ma è immediato anche il seguente teorema più generale.

Se $\varphi(\mathrm{z})$ *e* $\psi(\mathrm{z})$ *sono due polinomi qualsivogliano, nella variabile complessa* z, *e* $\psi(\mathrm{z})$ *non è identicamente nullo, supposta* $\psi(\mathrm{a})$ *dotata di inversa, il coinsieme della funzione razionale* $\varphi(\mathrm{z})/\psi(\mathrm{z})$ *relativo all'insieme dei numeri caratteristici dell'operazione* a, *coincide con quello dei numeri caratteristici dell'operazione*

$$\psi(a)^{-1}\varphi(a) \equiv \varphi(a)\psi(a)^{-1},$$

che indicherò col simbolo

$$\varphi(a)/\psi(a) \equiv \frac{\varphi(a)}{\psi(a)}.$$

DIMOSTRAZIONE. – Osserviamo, anzitutto, che $\varphi(\mu)/\psi(\mu)$ non perde mai di significato al variare di μ nell'insieme I dei numeri caratteristici di a, poichè, per μ in I, $\psi(\mu)$ è numero caratteristico di $\psi(a)$ e quindi è sempre $\psi(\mu) \neq 0$, essendo $\psi(a)$ dotata di inversa. Se μ è in I, si ha poi

$$\varphi(a)x = \varphi(\mu)x, \qquad \psi(a)x = \psi(\mu)x, \quad \text{con un} \quad x \neq 0,$$

e quindi

$$\frac{\varphi(a)}{\psi(a)}\, x = \psi(a)^{-1}\varphi(a)x = \psi(a)^{-1}[\varphi(\mu)x] = \varphi(\mu)[\psi(a)^{-1}x] = \frac{\varphi(\mu)}{\psi(\mu)}\, x,$$
$$\text{con un} \quad x \neq 0,$$

ciò che prova l'appartenenza di $\varphi(\mu)/\psi(\mu)$ all'insieme J dei numeri caratteristici di $\varphi(a)/\psi(a)$. Viceversa, se ν è in J, si ha

$$\frac{\varphi(a)}{\psi(a)}x - \nu x = 0, \qquad \text{con un} \qquad x \neq 0,$$

e quindi $\varphi(a)x - \nu\psi(a)x = 0$, con un $x \neq 0$, L'operazione $\varphi(a) - \nu\psi(a)$ possiede dunque un autoelemento e perciò, dette $\mu_1, \mu_2, \dots, \mu_r$ le radici dell'equazione $\varphi(\mu)/\psi(\mu) = \nu$, almeno una delle operazioni $a - \mu_1 e$, $a - \mu_2 e, \dots$, $a - \mu_r e$ deve ammettere un autoelemento. Se $a - \mu_h e$ è una di queste, troviamo il punto μ_h di I per cui $\varphi(\mu_h)/\psi(\mu_h) = \nu$ è in J.

. . . .