

CAPITOLO 5

SPAZI VETTORIALI NORMATI

In questo capitolo i concetti fondamentali dell'algebra lineare vengono ripresi e riformulati con una notazione più funzionale agli scopi del calcolo numerico. Vengono inoltre ripresi ed ampliati i concetti relativi alle norme dei vettori e delle matrici in vista della loro utilizzazione nell'analisi dell'errore di ogni procedimento numerico atto ad approssimare la soluzione dei vari problemi che verranno proposti nel corso.

1-NOZIONI PRELIMINARI.

Consideriamo lo spazio vettoriale $R^n, (C^n)$ sul corpo $R (C)$, cioè le n -uple ordinate di numeri reali (complessi) (x_1, x_2, \dots, x_n) con l'operazione di somma:

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1+y_1, x_2+y_2, \dots, x_n+y_n)$$

e l'operazione esterna

$$\lambda(x_1, x_2, \dots, x_n) = (\lambda x_1, \lambda x_2, \dots, \lambda x_n) \quad \lambda \in R (C).$$

Sia $\mathcal{L}(R^n, R^n)$ l'insieme delle applicazioni lineari di R^n in se. Esso è a sua volta spazio vettoriale su R ed ogni applicazione $A \in \mathcal{L}(R^n, R^n)$ è rappresentata da una matrice $n \times n$ con coefficienti a_{ij} , che indicheremo ancora con $A=(a_{ij})$. Analogamente, l'insieme $\mathcal{L}(C^n, C^n)$ è lo spazio delle matrici a coefficienti complessi. La somma ed il prodotto esterno rispetto al quale le matrici formano spazio vettoriale sono $A+B=(a_{ij}+b_{ij})$ ed $\alpha A=(\alpha a_{ij})$. Per le matrici si definisce anche una moltiplicazione AB che corrisponde alla applicazione composta $A \circ B$. Tale operazione non è commutativa.

In C^n si può definire una forma lineare, detta prodotto scalare,

$$\langle x, y \rangle := \bar{x}_1 y_1 + \bar{x}_2 y_2 + \dots + \bar{x}_n y_n$$

che gode delle proprietà:

$$1- \langle x, x \rangle \geq 0, \quad \langle x, x \rangle = 0 \Leftrightarrow x=0$$

$$2- \langle x, y \rangle = \overline{\langle y, x \rangle}$$

$$3- \langle \alpha x, y \rangle = \bar{\alpha} \langle x, y \rangle \quad \alpha \in \mathbb{C}$$

$$4- \langle x+y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$$

Se ora interpretiamo i vettori come "vettori colonna", cioè come matrici $n \times 1$ appartenenti a $\mathcal{L}(\mathbb{C}^1, \mathbb{C}^n)$, si vede che $\langle x, y \rangle$ altro non è che il prodotto matriciale di x^H per y , dove

$$x^H := (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$$

è detto "trasposto coniugato" di x . Naturalmente, con questa convenzione, il generico vettore x , che abbiamo deciso di interpretare come vettore colonna, andrà rappresentato con

$$x = (x_1, x_2, \dots, x_n) \quad \text{oppure, più convenientemente, con } x = (x_1, x_2, \dots, x_n)^T.$$

Nella precedente notazione matriciale, anche il prodotto scalare può essere rappresentato in maniera più compatta nella seguente forma:

$$\forall x, y \in \mathbb{R}^n \quad \langle x, y \rangle = x^T y$$

$$\forall x, y \in \mathbb{C}^n \quad \langle x, y \rangle = x^H y.$$

Di conseguenza, le proprietà del prodotto scalare possono essere riscritte nel seguente modo:

$$1- x^H x \geq 0 \quad x^H x = 0 \Leftrightarrow x \equiv 0$$

$$2- x^H y = \overline{y^H x} = (y^H x)^H$$

$$3- (\alpha x)^H y = \bar{\alpha} x^H y$$

$$4- (x+y)^H z = x^H z + y^H z$$

Si osservi inoltre che la notazione colonna rende formalmente corretta la scrittura Ax che viene interpretata come "moltiplicazione" della matrice A (di dimensione $n \times n$) con la matrice x (di dimensione $n \times 1$). Analogamente, la moltiplicazione a sinistra va scritta nella forma $x^H A$.

Ricordiamo ora alcune definizioni e notazioni:

Due vettori x ed y si dicono **ortogonali** ($x \perp y$) se $x^H y = 0$

Sia $A \in \mathcal{L}(R^n, R^n)$, $A = (a_{ij})$, $a_{ij} \in R$

La matrice $A^T = (a_{ji})$ è detta matrice **trasposta** di A ;

A è detta **simmetrica** se $A = A^T$;

A è detta **ortogonale** se $A A^T = A^T A = I$, cioè se le sue colonne e le sue righe sono a due a due ortogonali in R^n .

La matrice A è detta **regolare**, o **non singolare**, se esiste una matrice, detta **inversa** di A ed indicata con A^{-1} , tale che $A A^{-1} = A^{-1} A = I$

A è detta **involutoria** se $A = A^{-1}$ e quindi $A^2 = I$.

Sia ora $A \in \mathcal{L}(C^n, C^n)$, $A = (a_{ij})$, $a_{ij} \in C$

La matrice $A^H = (\bar{a}_{ji})$ è detta matrice **trasposta coniugata** di A ;

A è detta **hermitiana** se $A = A^H$;

A è detta **unitaria** se $A A^H = A^H A = I$, cioè se le sue colonne e le sue righe sono a due a due ortogonali in C^n .

Le nozioni di matrice regolare ed involutoria rimangono le stesse del caso reale.

Salvo esplicita dichiarazione contraria, d'ora in poi considereremo sempre matrici e vettori complessi e tutte le definizioni e proprietà che dimostreremo varranno in particolare nel caso reale.

La matrice A è detta **semidefinita positiva** se $x^H A x \geq 0 \quad \forall x \neq 0$,
ed è detta **definita positiva** se $x^H A x > 0 \quad \forall x \neq 0$.

Definizione. Il numero $\lambda \in C$ è detto **autovalore** di A se \exists un vettore $x \neq 0$, detto **autovettore** o **autosoluzione** associata a λ , tale che $Ax = \lambda x$. L'insieme degli autovalori di A è indicato con $S(A)$ ed è chiamato **spettro** di A .

Poichè un sistema lineare omogeneo, $Bx=0$, ammette soluzioni non nulle se e solo se la matrice B è singolare, gli autovalori sono dati dalle radici della seguente **equazione caratteristica** associata

$$\det(A-\lambda I)=0.$$

Il termine $\det(A-\lambda I)$ è un polinomio di grado n a coefficienti reali o complessi e quindi possiede n radici complesse, il cui prodotto è il termine noto del polinomio, e quindi è il $\det(A)$, e la cui somma è il coefficiente del termine di grado $n-1$, e quindi coincide con la somma degli elementi diagonali di A , che è detta **traccia** di A . In breve:

$$\prod \lambda_i = \det(A)$$

$$\sum \lambda_i = \text{tr}(A) := \sum a_{ii}$$

Qualcuna delle radici può essere multipla e la molteplicità della radice è detta **molteplicità algebrica** dell'autovalore. L'insieme degli autovalori è detto brevemente **spettro** di A ed è indicato con $S(A)$, mentre l'insieme degli autovettori è detto **autospazio** di A .

Valgono le seguenti proposizioni le cui dimostrazioni sono immediate.

PROPOSIZIONE 1. $(Ax)^H = x^H A^H$ e quindi, più in generale, $(AB)^H = B^H A^H$

PROPOSIZIONE 2. $(x^H A y)^H = (A y)^H x = y^H A^H x$

PROPOSIZIONE 3. A hermitiana $\Rightarrow x^H A x$ è reale.

PROPOSIZIONE 4. A def. pos. $\Rightarrow |A| \neq 0$.

Dim. Se A fosse singolare esisterebbe $x \neq 0$ tale che $Ax=0$ ed anche $x^H A x=0$.

PROPOSIZIONE 5. A def. pos. $\Rightarrow a_{ii} > 0$; A semi-def. pos. $\Rightarrow a_{ii} \geq 0 \quad \forall i$.

Dim. Si prenda il vettore canonico $e_i = (0,0,\dots,1,\dots,0)$ e si osservi che $Ae_i = (a_{1i}, a_{2i}, \dots, a_{ni})$ e che $e_i^H Ae_i = a_{ii}$

PROPOSIZIONE 6. A hermitiana \Rightarrow autovalori reali.

Dim. Sia $Ax = \lambda x$. Allora $x^H Ax = x^H(\lambda x) = \lambda x^H x$. Passando ai coniugati

$$(x^H Ax)^H = (\lambda x^H x)^H$$

$$x^H A^H x = \bar{\lambda} x^H x$$

$$x^H Ax = \bar{\lambda} x^H x \quad \text{da cui segue } \bar{\lambda} = \lambda$$

PROPOSIZIONE 7. Per ogni autovalore λ , A def. pos. $\Rightarrow \lambda > 0$ e A semidef. pos. $\Rightarrow \lambda \geq 0$

Dim. Sia $Ax = \lambda x$. Allora $x^H Ax = \lambda x^H x$ da cui segue la tesi

PROPOSIZIONE 8. $A^H = A^{-1} \Rightarrow x^H y = (Ax)^H Ay$

In altre parole: le matrici unitarie conservano il prodotto scalare. Infatti

$$(Ax)^H Ay = x^H A^H Ay = x^H A^{-1} Ay = x^H y.$$

PROPOSIZIONE 9 (Disuguaglianza di Schwartz): $|x^H y|^2 \leq x^H x \cdot y^H y$

o equivalentemente:

$$|x^H y| \leq \sqrt{x^H x} \sqrt{y^H y}$$

$$\left| \sum \bar{x}_i y_i \right| \leq \sqrt{\sum |x_i|^2} \sqrt{\sum |y_i|^2}.$$

PROPOSIZIONE 10: Una matrice A che goda di 2 delle tre proprietà:

A è hermitiana

A è unitaria

A è involutoria,

gode anche della terza.

PROPOSIZIONE 11: $\lambda \in S(A) \Rightarrow \lambda^k \in S(A^k)$ per ogni intero $k \geq 0$ e, per matrici non singolari, anche per $k < 0$.

TEOREMA 1.1 (DI GERSCHGORIN). *Data la matrice $A = (a_{ij})$, ogni autovalore di A è contenuto nella riunione dei dischi $D_i = \{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{i \neq j} |a_{ij}| \}$ $i = 1, \dots, n$. Inoltre se k dischi D_i sono disgiunti dagli altri, allora k autovalori appartengono all'insieme U , riunione di quei k dischi.*

Dim: Sia λ un autovalore arbitrario di A ed x il corrispondente autovettore. La generica riga k -esima della relazione $Ax=\lambda x$ è:

$$\sum_j a_{kj}x_j = \lambda x_k.$$

Da essa si ha:

$$\sum_{j \neq k} a_{kj}x_j = (\lambda - a_{kk})x_k$$

$$|\lambda - a_{kk}| |x_k| \leq \sum_{j \neq k} |a_{kj}| |x_j|.$$

Sia x_i la componente di modulo massimo di x , cioè tale che $|x_j| \leq |x_i|$ per ogni j . Per l'indice i si ha allora

$$|\lambda - a_{ii}| |x_i| \leq \sum_{j \neq i} |a_{ij}| |x_j|$$

$$|\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}| \frac{|x_j|}{|x_i|} \leq \sum_{j \neq i} |a_{ij}|$$

cioè $\lambda \in D_i$. Non potendo sapere qual'è l'indice i , possiamo concludere che $\lambda \in \cup_i D_i$. La prima parte del teorema è così dimostrata. Per quanto riguarda la seconda parte, sia:

$$A(t) = D + t(A - D) \quad t \in [0, 1]$$

e si osservi che:

$$A(0) = D \quad \text{ed} \quad A(1) = A.$$

Mentre t cresce da 0 ad 1 la matrice $A(t)$ passa con continuità dalla matrice D alla matrice A , e così pure gli autovalori $\lambda_i(t)$ delle matrici $A(t)$ percorrono con continuità delle traiettorie da $\lambda_i(0)$ a $\lambda_i(1)$ (la continuità di ciascuna traiettoria è dovuta al fatto che le radici di un polinomio sono funzioni continue dei suoi coefficienti). Così pure i dischi $D_i(t)$ della matrice $A(t)$, che sono centrati negli elementi diagonali a_{ii} per ogni t , hanno raggio crescente al crescere di t in $[0,1]$. Da un lato, per $t=0$, essi si riducono all'elemento diagonale stesso a_{ii} , dall'altro, per $t=1$, essi sono i dischi D_i della matrice A .

Se D_p è uno dei dischi inclusi in U , anche $D_p(t)$ è incluso in U per ogni t . Per la prima parte del teorema la traiettoria $\lambda_p(t)$, uscente da $a_{pp}=\lambda_p(0)$ deve rimanere in U per ogni valore di t , altrimenti verrebbe a mancare la continuità. In particolare per $t=1$, l'autovalore $\lambda_p=\lambda_p(1)$ appartiene ad U . Ciò è vero per ciascuno dei k dischi che compongono U , quindi U include almeno k autovalori. D'altra parte, per quanto abbiamo appena visto, i rimanenti $n-k$ autovalori devono stare nell'insieme V , costituito dalla riunione dei rimanenti dischi, che risulta a sua volta disgiunto da U .

Consideriamo ora alcune matrici di forma particolare che interverranno spesso in seguito.

Matrici di permutazione: Sono matrici che hanno un coefficiente uguale ad 1 su ciascuna riga e ciascuna colonna e tutti gli altri coefficienti uguali a 0, ad esempio la seguente matrice 5×5 :

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Per una matrice di permutazione P che abbia il coefficiente p_{ij} uguale ad 1, il prodotto PA provoca, sulla matrice A , il trasferimento della riga j -esima nella i -esima. Il prodotto AP provoca invece lo stesso trasferimento sulle colonne. Poichè ogni riga ed

E' universalmente adottata la notazione L (Lower) per le matrici triangolari inferiori e U (Upper) per quelle triangolari superiori. Si osservi che, in riferimento alla matrice Q di riflessione, QLQ e QUQ sono, rispettivamente, triangolari superiore e inferiore. Si osservi infine che il prodotto di matrici triangolari L od U è ancora una matrice dello stesso tipo.

Matrici diagonali. Sono dette matrici diagonali le matrici con elementi non nulli solo sulla diagonale:

$$A = \begin{pmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \lambda_n \end{pmatrix}$$

Esse saranno rappresentate, più semplicemente, nel seguente modo:

$$A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

Il passaggio dalla matrice diagonale al vettore degli elementi diagonali della matrice si ottiene moltiplicando A per il **vettore unitario** $u=(1,1,\dots,1)$: $Au=(\lambda_1,\lambda_2,\dots,\lambda_n)^T$.

Gli autovalori di una matrice diagonale o triangolare sono gli elementi della diagonale principale.

Definizione: Le matrici A e B si dicono simili ($A \approx B$) se esiste una matrice non singolare P tale che $B=P^{-1}AP$.

E' immediato osservare che se λ è autovalore di A con autovettore x, allora λ è autovalore della matrice B con autovettore $P^{-1}x$. Infatti: $Ax=\lambda x \Leftrightarrow P^{-1}Ax=\lambda P^{-1}x \Leftrightarrow P^{-1}APP^{-1}x=\lambda P^{-1}x \Leftrightarrow BP^{-1}x=\lambda P^{-1}x$.

Definizione. Data una matrice A, si chiama raggio spettrale di A, e si indica con $\rho(A)$, il numero $\max_i |\lambda_i|$.

2-NORME DI VETTORI E DI MATRICI.

Indichiamo in generale con V uno spazio vettoriale .

Definizione: *Sullo spazio V si definisce **norma** ogni applicazione $\| \cdot \|$ a valori reali non negativi che gode delle proprietà:*

$$\|x\| \geq 0; \quad \|x\| = 0 \Leftrightarrow x=0 \text{ (elemento nullo di } V)$$

$$\|\alpha x\| = |\alpha| \|x\|$$

$$\|x+y\| \leq \|x\| + \|y\| \quad \text{per ogni } x \text{ ed } y \text{ di } V, \text{ e per ogni scalare } \alpha.$$

In generale si possono definire più norme sullo stesso spazio vettoriale. Consideriamo dapprima lo spazio C^n e definiamo alcune possibili norme:

$$\|x\|_1 = \sum |x_i|$$

$$\|x\|_2 = \sqrt{\sum |x_i|^2} \quad \text{è la norma Euclidea} \quad \|x\|_2^2 = x^H x$$

$$\|x\|_p = \sqrt[p]{\sum |x_i|^p}$$

$$\|x\|_\infty = \max_i |x_i|$$

$$\|x\|_A = \|Ax\| \quad |A| \neq 0$$

$\|x\|_{ell} = \sqrt{x^H A x}$ con $A^H=A$, A def.pos. E' detta norma ellittica ed è ricavata dal prodotto scalare: $\langle x, y \rangle_A := x^H A y$.

Si osservi che $\lim_{p \rightarrow \infty} \|x\|_p = \|x\|_\infty$ infatti $\sqrt[p]{\sum |x_i|^p} \geq \|x\|_\infty$ ed inoltre $\sqrt[p]{\sum |x_i|^p} \leq \sqrt[p]{n \max_i |x_i|^p} = \sqrt[p]{n} \|x\|_\infty \rightarrow \|x\|_\infty$ per $p \rightarrow \infty$.

D'ora in poi, nel considerare lo spazio vettoriale R^n , C^n , o in generale uno spazio V , lo penseremo dotato di una norma e lo indicheremo con $(V, \| \cdot \|)$. Ricordiamo che negli spazi vettoriali di dimensione finita, quali quelli che stiamo considerando, tutte le norme sono equivalenti, cioè per ogni coppia di norme $\| \cdot \|$, $\| \cdot \|^\wedge$ esistono due costanti m , M tali che

$$m \|x\| \leq \|x\|^\wedge \leq M \|x\| \quad \forall x \in V.$$

Il lettore verifichi che valgono le seguenti relazioni di equivalenza

$$\|x\|_{\infty} \leq \|x\|_2 \leq \sqrt{n} \|x\|_{\infty}$$

$$\|x\|_{\infty} \leq \|x\|_1 \leq n \|x\|_{\infty}$$

$$\frac{1}{\sqrt{n}} \|x\|_1 \leq \|x\|_2 \leq \|x\|_1$$

(la dimostrazione della terza richiede l'impiego della disuguaglianza di Cauchy-Schwarz).

Una immediata conseguenza dell'equivalenza delle norme sta nel fatto che, se una successione converge in una norma, converge in ogni altra norma.

Poichè ogni matrice A di dimensione $n \times m$ può essere vista come un vettore ad nm componenti, anche le matrici possono essere normate. Si avrà, per esempio:

$$\|A\|_{\text{Euclidea}} = \sqrt{\sum_{ij} |a_{ij}|^2}$$

$$\|A\|_{\text{max}} = \max_{ij} |a_{ij}|.$$

Un altro modo di definire delle norme per le matrici, è quello di considerare le matrici come elementi dello spazio $\mathcal{L}(C^n, C^n)$, cioè come applicazioni lineari dello spazio vettoriale normato $(C^n, \|\cdot\|)$ in se. Abbiamo già detto che $\mathcal{L}(C^n, C^n)$ è, a sua volta, uno spazio vettoriale. La norma definita sui vettori di C^n induce in modo naturale una norma sugli elementi di $\mathcal{L}(C^n, C^n)$, cioè sulle matrici. A tale scopo consideriamo l'applicazione di $\mathcal{L}(C^n, C^m)$ in \mathbb{R} così definita:

$$\|A\| := \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \quad A \in \mathcal{L}(C^n, C^m)$$

Essa gode delle seguenti proprietà (che dimostreremo tra poco):

1. $\|A\| \geq 0$, $\|A\| = 0 \Leftrightarrow A=0$ (matrice nulla)
2. $\|\alpha A\| = |\alpha| \|A\|$
3. $\|A+B\| \leq \|A\| + \|B\|$
4. $\|AB\| \leq \|A\| \|B\|$

e viene detta **norma naturale** o **norma indotta** dalle norme di vettore. In particolare, se C^n è dotato della norma $\| \cdot \|_i$, scriveremo

$$\| A \|_i = \max_{x \neq 0} \frac{\| Ax \|_i}{\| x \|_i}$$

Osserviamo subito che per ogni norma naturale si ha $\| I \| = 1$.

Prima di dimostrare che valgono le quattro proprietà delle norme, osserviamo che la norma indotta si può esprimere in altro modo:

$$\| A \| = \max_{x \neq 0} \frac{\| Ax \|}{\| x \|} = \max_{x \neq 0} \left\| \frac{Ax}{\| x \|} \right\| = \max_{\| x \| = 1} \| Ax \|$$

In seguito sarà utile ricordare che esiste un $x : \| x \| = 1$ e $\| Ax \| = \| A \|$. Inoltre è ovvia la seguente proprietà delle norme naturali:

Proprietà : Per ogni norma naturale di matrice si ha:

$$\| Ax \| \leq \| A \| \| x \| \quad \forall A \quad \text{e} \quad \forall x.$$

Verifichiamo ora che l'applicazione $A \rightarrow \| A \|$ verifica le 4 proprietà della norma precedentemente enunciate.

1. Parte della dimostrazione è ovvia; dimostriamo solo che $\| A \| = 0 \Rightarrow A=0$. Poichè $\| A \| = 0$ allora $\| Ax \| = 0$ per ogni $\| x \| = 1$. D'altra parte $\| Ax \| = 0 \Rightarrow Ax=0$ e, per l'arbitrarietà di x nella sfera unitaria, si ha $A=0$.

2. Dimostrazione ovvia.

3. Sia x tale che $\| x \| = 1$ e $\| (A+B)x \| = \| A+B \|$. Si ha $\| (A+B)x \| = \| Ax+Bx \| \leq \| Ax \| + \| Bx \| \leq \| A \| \| x \| + \| B \| \| x \| = \| A \| + \| B \|$.

4. Dimostrazione analoga.

La definizione di norma naturale non consente, in generale, di valutare facilmente la norma di una matrice in termini dei suoi coefficienti. Ciò risulta possibile in alcuni casi particolari.

TEOREMA 1.2. Per ogni $A \in \mathcal{L}(C^n, C^n)$, $\|A\|_\infty = \max_i \sum_j |a_{ij}|$.

TEOREMA 1.3. Per ogni $A \in \mathcal{L}(C^n, C^n)$, $\|A\|_1 = \max_j \sum_i |a_{ij}|$

TEOREMA 1.4. $\|A\|_2 = \sqrt{\rho(A^H A)} = \sqrt{\rho(A A^H)}$

E' immediato osservare che $\rho(A)$ è una minorazione per tutte le norme naturali di A . Infatti, detto λ l'autovalore di modulo massimo ed x il corrispondente autovettore normalizzato in una qualunque norma ($\|x\|=1$), si ha:

$$Ax = \lambda x \Rightarrow \|Ax\| = |\lambda| \|x\| = |\lambda| \Rightarrow |\lambda| \leq \|A\| \|x\| = \|A\|.$$

da cui si ricava $\rho(A) \leq \|A\|$. Per qualche matrice può accadere che il raggio spettrale coincida con qualche norma. Per esempio se A è hermitiana, dal teorema 1.14 segue che $\rho(A) = \|A\|_2$. Per una matrice arbitraria A , vale il seguente importante risultato

TEOREMA 1.5. Per ogni $A \in \mathcal{L}(C^n, C^n)$,

$$\rho(A) = \inf \|A\|,$$

dove l'estremo inferiore è inteso rispetto a tutte le norme naturali.

Per lo studio dei metodi iterativi, di cui ci occuperemo nel prossimo capitolo, è importante il concetto di *matrice convergente*, cioè di matrice la cui successione di potenze A^m tende alla matrice nulla, per m che tende a infinito.

Il seguente teorema caratterizza le matrici convergenti.

TEOREMA 1.6. Per ogni matrice A le seguenti tre affermazioni sono equivalenti :

1. $\rho(A) < 1$
2. $\|A^m\| \rightarrow 0$,
3. $(I - A)^{-1} = I + A + A^2 + \dots + A^n + \dots$ (*serie di Neumann*).

Dim. (1. \Leftrightarrow 2.) Se $\rho(A) < 1$, esiste una norma tale che $\|A\| < 1$. Quindi $\|A^m\| \leq \|A\|^m \rightarrow 0$. Viceversa, poiché $\rho^m(A) = \rho(A^m) \leq \|A^m\|$ ed inoltre $\|A^m\| \rightarrow 0$, allora $\rho^m(A) \rightarrow 0$ e quindi $\rho(A) < 1$.

(2. \Leftrightarrow 3.) Poiché la matrice è convergente, il raggio spettrale è minore di uno e quindi $\lambda = 1$ non è autovalore di A . Di conseguenza $(I - A)$ è invertibile. Per ogni intero m si ha

$$(I - A)(I + A + A^2 + \dots + A^m) = I - A^{m+1}$$

$$I + A + A^2 + \dots + A^m = (I - A)^{-1}(I - A^{m+1}).$$

Poiché la matrice A è convergente, si può concludere, per la continuità della norma, che A^{m+1} tende alla matrice nulla e quindi il termine a destra dell'ultima uguaglianza tende a $(I - A)^{-1}$. Di conseguenza anche il termine a sinistra è convergente e converge allo stesso limite. Viceversa se la serie di Neumann converge allora il termine generale della serie deve essere infinitesimo.

Si osservi l'analogia tra la serie di Neumann e la serie geometrica $1 + x + \dots + x^n + \dots$ che converge a $1/(1-x)$ per ogni x reale (o complesso) di modulo minore di uno. Ciò non è tanto sorprendente se si pensa che l'insieme delle matrici quadrate, con le usuali operazioni di somma e moltiplicazione, godono di quasi tutte le proprietà dei numeri reali o complessi. La sola proprietà mancante è la commutatività del prodotto, ma nella serie di Neumann i prodotti sono fatti solo sulla stessa matrice A .

Concludiamo questa parte introduttiva con una importante proprietà del raggio spettrale.

TEOREMA 1.7. Per ogni matrice A e per ogni norma naturale si ha:

$$\lim_{k \rightarrow \infty} \|A^k\|^{1/k} = \rho(A).$$