Operatori lineari su spazi finito-dimensionali e matrici.

I.1 Operatori lineari e matrici

Consideriamo un operatore lineare $A: \mathbb{C}^N \to \mathbb{C}^N$.
Sia $\{e^{(i)}\}_{i=1}^N$ una base in $\mathbb{C}^N$: in questa base l'operatore A è rappresentato da una ben determinata matrice $\{A_{n,m}\}_{n,m=1}^N$. Infatti:

I.1.1 $$y = Ax \leftrightarrow \sum_i y_i e^{(i)} = A\Big(\sum_j x_j e^{(j)}\Big) = \sum_j x_j A e^{(j)} =$$

$$= \sum_i \sum_j x_j A_{ij} e^{(i)} \rightarrow y_i = \sum_j A_{ij} x_j$$

I.2 Cambiamenti di base

Supponiamo di voler "cambiare base" nel nostro spazio vettoriale $\mathbb{C}^N$, e indichiamo con $\{\eta^{(i)}\}_{i=1}^N$ i vettori della nuova base.
Potremo scrivere:

I.2.1 $$\cancel{e^{(i)} = \sum_{j=1}^N T_{ji}\eta} \quad \eta^{(i)} = \sum_{j=1}^N T_{ji} e^{(j)}$$

La condizione affinché $\{\eta^{(i)}\}$ sia ancora una base in $\mathbb{C}^N$ è che gli N vettori $\eta^{(i)}$ siano linearmente indipendenti, cioè che:

I.2.2 $$\sum_i c_i \eta^{(i)} = 0 \leftrightarrow \{c_i\}_{i=1}^N = 0$$

Deve cioè essere:

I.2.3 $$\sum_i \sum_j c_i T_{ji} e^{(j)} = 0 \leftrightarrow \{c_i\}_{i=1}^N = 0$$

Ma per ipotesi gli $e^{(i)}$ sono una base in $\mathbb{C}^N$, quindi la condizione precedente può essere riformulata

Spazio duale

$$f = \sum_i f^i g_i \qquad g_i(x) = x_i$$

Cambiamento di base: $\quad \eta^{(i)} = T_{ji}\, e^{(j)}$

$$e^{(i)} = \left(T^t\right)^{-1}_{ij} \eta^{(j)}$$

$$g_i(x) = x_i = \sum_j T_{ij}\, \xi_j = \sum_j T_{ij}\, \gamma_{(j)}(\xi)$$

$$f = \sum_{ij} f^i T_{ij}\, \gamma_j = \sum_j \varphi^j \gamma_j \qquad \varphi^j = \sum_i T_{ij} f^i \quad \Rightarrow \varphi = T^t f$$

$$\gamma = T^{-1} g$$

Per effetto di un cambiamento di base in uno spazio lineare, le componenti di un vettore cambiano secondo una trasformazione che è la trasposta dell'inversa della trasf. del cambiamento di base. Le componenti di un funzionale lineare cambiano invece secondo la stessa trasformazione del cambiamento di base.
Per questo i vettori di C^N si dicono controvarianti, i vettori di C^{N*} si dicono covarianti e si ~~indicano~~ chiamano spesso covettori.

Notare che in uno spazio euclideo reale, se ci si limita a considerare il gruppo delle trasformazioni ortogonali, questa differenza di comportamento sparisce: $(T^t)^{-1} = T$.

dicendo che:

I.2.4 $\sum_i T_{ji}\, c_i = 0 \quad \longleftrightarrow \quad \{c_i\}_{i=1}^{N} = 0$

Cioè il sistema lineare omogeneo (I.2.4) deve ammettere solo la soluzione nulla, il che è vero se e solo se la matrice del cambiamento di base $\{T_{ij}\}_{j,i=1}^{N}$ è non singolare.

Vediamo ora come cambiano le componenti di un vettore per effetto di un cambiamento di base

Siano $\{x_i\}_{i=1}^{N}$ le componenti di x secondo la base $\{e^{(i)}\}$ e siano $\{\xi_i\}_{i=1}^{N}$ le componenti dello stesso vettore nella base $\{\eta^{(i)}\}$. Si ha:

$$x = \sum_i x_i e^{(i)} = \sum_j \xi_j \eta^{(j)} = \sum_j \sum_i \xi_j T_{ij}\, e^{(i)}$$

da cui:

I.2.5 $$x_i = \sum_j T_{ij}\, \xi_j$$

Quindi, se T^t è la matrice che fa passare dalla vecchia alla nuova base, T è la matrice che fa passare dalle nuove alle vecchie coordinate e quindi T^{-1} è la matrice che fa passare dalle vecchie alle nuove coordinate. In altre parole, la matrice del cambiamento di coordinate è l'inversa della trasposta della matrice del cambiamento di base.

Se $\mathbb{C}^N$ è uno spazio euclideo, ha senso costruire in esso basi ortonormali.

Se $\{e^{(i)}\}$ e $\{\eta^{(i)}\}$ sono 2 basi ortonormali, si ha:

I.2.6 $T_{ji} = (e^{(j)}, \eta^{(i)}) \quad ; \quad (T^{-1})_{ji} = (\eta^{(j)}, e^{(i)}) = \overline{T_{ij}} = (T^{+})_{ji}$

Quindi:

La matrice di trasformazione per cambiamenti di base ortonormali è una matrice unitaria; le matrici unitarie si indicano generalmente con la lettera U.

La condizione di unitarietà è quindi: $U^{-1} = U^{+}$.

In particolare, se il nostro spazio euclideo è reale (cioè è $\mathbb{R}^N$ dotato di una metrica euclidea) le matrici in gioco sono reali e la condizione di prima (I.2.6) diventa semplicemente $T^{-1} = T^{t}$.

Matrici di questo tipo si dicono ortogonali.

(Tanto le matrici ortogonali che le matrici unitarie costituiscono un <u>gruppo</u>. Infatti la matrice identità I è unitaria (e ortogonale). L'inversa di una matrice unitaria (ortogonale) è ancora una matrice unitaria (ortogonale); il prodotto di due matrici unitarie (ortogonali) è ancora una matrice unitaria (ortogonale).

Dimostriamo quest' ultima affermazione; si ha:

$(U_1 U_2)^{-1} = U_2^{-1} U_1^{-1} = U_2^{+} U_1^{+} = (U_1 U_2)^{+}$, c.v.d.).

Sussiste il <u>teorema</u>:

CNES affinché una matrice T sia unitaria, è che essa conservi il prodotto scalare fra vettori.

Dimostriamolo.

a) Necessità.

Supponiamo Ψ unitaria, allora:

$(\Psi y, \Psi x) = (y, \Psi^{+}\Psi x) = (y, \Psi^{-1}\Psi x) = (y, x)$.

b) Sufficienza.

Sia $x_1 = T\,\xi_1$, $x_2 = T\,\xi_2$, e supponiamo che:

$(x_1, x_2) = (\xi_1, \xi_2)$ $\forall$ coppia di vettor. ξ_1, ξ_2.

Abbiamo:

4

$$(x_1,x_2) = (T\xi_1, T\xi_2) = (\xi_1, T^+T\xi_2)$$

Allora:

$$(\xi_1,(T^+T - I)\xi_2) = 0 \quad \forall\ \xi_1,\xi_2$$

Di conseguenza $(T^+T - I)\xi_2 = 0 \quad \forall\ \xi_2$

e in definitiva: $T^+T = I \iff T^+ = T^{-1}$

Vale anche la condizione più debole:

CNES affinché una matrice T sia unitaria, è che essa conservi le norma dei vettori, cioè che valga: $(T\xi, T\xi) = (\xi,\xi) \quad \forall\ \xi \in \mathbb{C}^N$.

Vediamo ora come cambia, per effetto di un cambiamento di base, la matrice rappresentativa di un operatore lineare.

Sia T^t la matrice non singolare del cambiamento di base $\{e^{(i)}\} \to \{\eta^{(i)}\}$.

Si ha allora, come abbiamo visto:

$$x = T\xi$$

Sia A un operatore lineare su $\mathbb{C}^N$:

$$y = Ax$$

Passando alla nuova base, e indicando con ς_i le componenti di y rispetto alle base $\eta^{(i)}$:

$$T\varsigma = AT\xi \quad , \text{ cioè:}$$

$$\varsigma = T^{-1}AT\xi$$

Quindi, nelle nuove base, la matrice rappresentativa dell'operatore è data dalle matrice:

I.2.7 $\tilde{A} = T^{-1}AT$

Notiamo che somme e prodotto di due operator. si trasformano allo stesso modo:

I.2.8 $\tilde{A}+\tilde{B} = T^{-1}AT + T^{-1}BT = T^{-1}(A+B)T$

I.2.9 $\tilde{A}\tilde{B} = T^{-1}ATT^{-1}BT = T^{-1}(AB)T$

Ne segue in particolare che definendo il commutatore di due operatori lineari mediante le formule:

I.2.10 $C_{A,B} = AB - BA \stackrel{def.}{=} [A,B]$,

~~Risulta~~ risulta:

I.2.11 $[\tilde{A},\tilde{B}] = T^{-1}[A,B]T$

Quindi, se due op. lin. <u>commutano</u> in una base (cioè, se il loro <u>commutatore</u> in una certa base è <u>zero</u>) essi commutano in qualsiasi base.

In altre parole: La proprietà di commutazione di due operatori è una proprietà <u>intrinseca</u>, indipendente dalla base.

Sorge a questo punto la domanda se esistano delle <u>quantità caratteristiche di un operatore che siano invarianti per cambiamenti di base</u>

È immediato verificare che la <u>Traccia</u> (cioè la somma degli elementi diagonali) di una matrice e il <u>determinante</u> di una matrice sono due di questi invarianti.

Infatti si ha:

I.2.12 $T_r\,\tilde{A} = T_r\,T^{-1}AT = T_r\,TT^{-1}A = T_r\,A$

I.2.13 $\det\tilde{A} = \det T^{-1}AT = \det T^{-1}\cdot\det A\cdot\det T = \frac{1}{\det T}\cdot\det A\cdot\det T = \det A$

6

Per dimostrare la (I.2.12) abbiamo fatto uso della proprietà "ciclica" della traccia (che si verifica per calcolo diretto):

I.2.14 $Tr(A_1 A_2 \cdots A_K) = Tr(A_K \cdot A_1 \cdots A_{K-1}) =$
$Tr(A_{K-1} \cdot A_K \cdot A_1 \cdots A_{K-2})$ etc.

~~Sussiste~~ Traccia e determinante non sono tuttavia i soli invarianti di una matrice. Supponiamo infatti di voler risolvere il "problema agli autovalori":

I.2.15 $Ax = \lambda x$

Cioè, di voler determinare per quali valori della variabile complessa λ l'equazione I.2.15 ammette soluzioni non nulle (autosoluzioni). Essi saranno evidentemente i valori di λ che ~~annullano~~ soddisfano l'equazione di grado ~~in~~ N in λ:

I.2.16 $\det(A - \lambda I) = 0$

L'equazione (I.2.16) si dice equazione caratteristica dell'operatore A, e le sue radici si dicono valori caratteristici o autovalori dell'operatore A. Il polinomio di grado N in λ dato da $\det(A - \lambda I)$ si chiama polinomio caratteristico di A, e si indica con $P(\lambda)$.

Sussiste il teorema:

Il polinomio caratteristico di un operatore è invariante per cambiamenti di base.

7

La dimostrazione è semplice.
Si ha infatti:

$$\det(\tilde{A} - \lambda I) = \det(T^{-1} A T - \lambda I) =$$
$$\det(T^{-1} A T - \lambda T^{-1} T) = \det(T^{-1}(A - \lambda I) T) =$$
$$\det(A - \lambda I).$$

Ne segue in particolare che gli zeri del polinomio caratteristico (o i suoi coefficienti), e cioè gli autovalori di A, sono invarianti per cambiamenti di base.

Vogliamo ora far vedere che $T_r A$ e $\det A$ non sono altro che due fra i coefficienti del polinomio caratteristico. Cominciamo da $\det A$, che è più semplice. È immediato verificare che $P(0) = \det A$: $\det A$ è quindi il termine noto del polinomio caratteristico.
Quanto alla traccia, immaginiamo di calcolare $\det(A - \lambda I)$ moltiplicando gli elementi della prima riga per i rispettivi cofattori: è allora chiaro che ~~solo~~ il cofattore del 1º elemento è un polinomio di grado $N-1$ in λ, mentre gli altri sono polinomi di grado $N-2$. Da queste osservazione segue facilmente che il coefficiente di λ^N è $(-1)^N$; quanto al coefficiente di λ^{N-1}, ad esso contribuiscono:
$a_{11} \cdot (-1)^{N-1} + a_{22}(-1)^{N-1} + \cdots a_{NN}(-1)^{N-1}$. Quindi il coefficiente di λ^{N-1} è $(-1)^{N-1}(a_{11} + \cdots + a_{NN}) =$
$(-1)^{N-1} T_r A$.

I.3 Matrici hermitiane e loro proprietà; Matrici normali.

Cominciamo col ricordare alcune definizioni fondamentali.
Dato un operatore lineare A su uno spazio euclideo, si dice aggiunto o hermitiano coniugato di A e si indica con A^+ l'operatore che soddisfa alla relazione

I.3.1 $(x, Ay) = (A^+x, y)$ per ogni coppia $x, y \in \mathbb{C}^N$.

Ciò implica che se $\{A_{ij}\}_{i,j=1}^{N}$ è la matrice rappresentativa di A in una base, $\{(A^+)_{ij}\} = \{\bar{A}_{ji}\}_{i,j=1}^{N}$ è la matrice rappresentative di A^+ nella stessa base. Questa caratterizzazione di A^+ è invariante per cambiamenti di base ~~infatti~~ ortonormali, infatti

$$\tilde{A}^+ = T^{-1}A^+T = T^+A^+T = (T^+AT)^+ = (\tilde{A})^+$$

(o, in altre parole, è invariante per trasf. unitarie).
Sussiste la proprietà:
Gli autovalori di A^+ sono i complessi coniugati degli autovalori di A. Infatti:

I.3.2 $\det\left[A^+ - \lambda I\right] = \det\left[A - \bar{\lambda} I\right]^+ = \overline{\det(A - \bar{\lambda} I)}$

Dalla (I.3.2) segue chiaramente che λ è uno zero dell'equazione caratteristica di A^+ se e solo se $\bar{\lambda}$ è uno zero dell'equazione caratteristica di A.

Sussiste il Teorema fondamentale:

|| Ogni matrice hermitiana può essere diagonalizzata mediante una trasformazione unitaria. ||

In altre parole, data una qualunque matrice hermitiana A, esiste una matrice unitaria U tale che

I.3.3 $\tilde{A} = U^+AU$ è una matrice diagonale.

Per dimostrare questo teorema, premettiamo ~~le~~ due fondamentali proprietà degli operatori hermitiani (valide anche su spazi euclidei finito-dimensionali!).

a) Gli autovalori di un operatore hermitiano sono reali.

b) Autovettori corrispondenti ad autovalori diversi sono ortogonali

Per dimostrare a), consideriamo l'equazione agli autovalori

$$A v = \lambda v$$

Supponiamo che λ_K sia un autovalore di A e che $v^{(K)}$ sia uno dei corrispondenti autovettori. Allora:

I.3.4 $(v^{(K)}, A v^{(K)}) = \lambda_K (v^{(K)}, v^{(K)})$

ma

I.3.5 $(v^{(K)}, A v^{(K)}) = (A v^{(K)}, v^{(K)}) = \overline{\lambda}_K (v^{(K)}, v^{(K)})$

da cui: $\boxed{\lambda_K = \overline{\lambda}_K}$

Dimostriamo ora b) supponendo $\lambda_K \neq \lambda_j$ e indicando con $v^{(K)}, v^{(j)}$ due autovettori corrispondenti. Si ha:

I.3.6 $(v^{(j)}, A v^{(K)}) = \lambda_K (v^{(j)}, v^{(K)})$

I.3.7 $(v^{(j)}, A v^{(K)}) = (A v^{(j)}, v^{(K)}) = \overline{\lambda}_j (v^{(j)}, v^{(K)}) = \lambda_j (v^{(j)}, v^{(K)})$

Quindi, sottraendo:

I.3.8 $(\lambda_K - \lambda_j)(v^{(j)}, v^{(K)}) = 0 \Rightarrow (v^{(j)}, v^{(K)}) = 0$.

Da questo segue intanto che se una matrice hermitiana A ha N autovalori distinti, essa ammette N autovettori ortogonali, quindi linearmente indipendenti, ed è perciò diagonalizzabile.

Infatti, indicando al solito con $e^{(i)}$ la base canonica, tale che: (10)

I.3.9 $A_{ij} = (e^{(i)}, A\, e^{(j)})$

e ricordando le proprietà dei cambiamenti di base ortonormali:

I.3.10 $e^{(i)} = \sum_{\ell} U^{+}_{\ell i}\, v^{(\ell)}$ $\qquad U^{+}_{\ell i} = (v^{(\ell)}, e^{(i)}) = \bar{v}_i^{(\ell)}$

$U_{i\ell} = (e^{(i)}, v^{(\ell)}) = v_i^{(\ell)}$

Si ha:

I.3.11 $A_{ij} = \sum_{\ell}\sum_{k} \overline{U^{+}_{\ell i}}\, (v^{(\ell)}, A\, v^{(k)})\, U^{+}_{kj} =$

$= \sum_{\ell}\sum_{k} U_{i\ell}\, \lambda_\ell\, \delta_{\ell k}\, U^{+}_{kj} = \sum_{k} U_{ik}\, \lambda_k\, (U^{+})_{kj}$

$=$

Da cui:

I.3.12 $A = U\, \tilde{A}\, U^{+}$ $\qquad \tilde{A} = \mathrm{diag}(\lambda_1, \dots, \lambda_N)$

In maniera equivalente la (I.3.12) si scrive:

I.3.13 $A = \sum_{k} \lambda_k\, P^{(k)}$

con $P^{(k)}_{ij} = U_{ik} (U^{+})_{kj} = v_i^{(k)}\, \bar{v}_j^{(k)}$

La rappresentazione I.3.13 si chiama rappresentazione spettrale della matrice A, e le matrici $P^{(k)}$ ~~pr~~ ~~sono~~ rappresentano operatori di proiezione ortogonali.

Per essi valgono le proprietà:

I.3.14a $P^{(k)} P^{(j)} = \delta_{kj}\, P^{(j)}$ $\quad ; \quad \sum_{j=1}^{N} P^{(j)} = I$ (I.3.14b)

Dimostriamole.

(a) $[P^{(k)} P^{(j)}]_{\ell m} = \sum_{s} P^{(k)}_{\ell s}\, P^{(j)}_{sm} = \sum_{s} v_\ell^{(k)}\, \bar{v}_s^{(k)}\, v_s^{(j)}\, \bar{v}_m^{(j)}$

$= v_\ell^{(k)}\, (v^{(k)}, v^{(j)})\, \bar{v}_m^{(j)} = \delta_{kj}\, v_\ell^{(j)}\, \bar{v}_m^{(j)} = \delta_{kj}\, P^{(j)}_{\ell m}$

11

(b) Dimostrare la (I.3.14b) è ovviamente equivalente a dimostrare che

$$\sum_j P^{(j)} x = x \qquad \forall x \in \mathbb{C}^N$$

Ma ciò è immediato:

$$\left[\sum_j P^{(j)} x\right]_k = \sum_j \sum_\ell P^{(j)}_{k\ell} x_\ell = \sum_{j\ell} v_k^{(j)} \bar{v}_\ell^{(j)} x_\ell =$$

$$= \sum_j v_k^{(j)} (v^{(j)}, x)$$

Quindi: $\displaystyle \sum_j P^{(j)} x = \sum_j (v^{(j)}, x) v^{(j)} = x$

(Altra dimostrazione, più diretta ma equivalente:

$$\sum_i P^{(i)}_{jk} = \sum_i v_j^{(i)} \bar{v}_k^{(i)} = \delta_{jk}$$

come conseguenza della completezza dell'insieme dei $v^{(i)}$).

Abbandoniamo ora l'ipotesi che gli autovalori di A siano tutti distinti. Senza fare il caso generale, supponiamo che ve ne sia 1 doppio. Ordiniamo allora gli autovalori in modo da avere $\lambda_1 = \lambda_2 \neq \lambda_3 \neq \cdots \neq \lambda_N$.

Esiste almeno un autovettore, che indicheremo con $v^{(2)}$, associato all'autovalore doppio, che sarà ortogonale a tutti gli altri (i quali ovviamente sono tutti mutuamente ortogonali). Per dimostrare che anche in questo caso A è diagonalizzabile mediante una trasformazione unitaria, occorre e basta trovare un vettore $v^{(1)}$, ortogonale a tutti gli altri, che sia ancora autovettore di A associato all'autovalore λ_1.

Osserviamo anzitutto che, essendo in C^N, esisterà sicuramente un vettore, w, ortogonale a tutti i $v^{(j)}$ $(j=,2,\dots N)$, tale che la base $\{w, v^{(j)}\}$ $(j=2,\dots,N)$ è una base ortonormale in C^N. Consideriamo gli elementi di matrice di A in questa base.
Si ha evidentemente:

$$\tilde{A}_{ij} = (v^{(i)}, A v^{(j)}) \quad (i,j=2,\dots,N) = \lambda_j \delta_{ij}$$

Quanto agli elementi della 1ª riga e della jª $(j\geq 2)$ colonna, abbiamo:

$$\tilde{A}_{1j} = (w, A v^{(j)}) = \lambda_j (w, v^{(j)}) = 0 \quad (j=2,\dots N)$$

E quindi, poiché A è hermitiana (osservazione cruciale) anche $\tilde{A}_{j1} = 0$ $(j=2,\dots,N)$.
Quindi, nella base $\{w, v^{(j)}\}$ A è diagonale.
Per dimostrare che w è autovettore di A corrispondente all'autovalore λ_1, e che quindi, nella suddetta base abbiamo $\tilde{A} = \mathrm{diag}\{\lambda_i\}_{i=1}^{N}$ $(\lambda_1=\lambda_2)$ basta far vedere che $(A-\lambda_1)w$ è il vettore nullo, cioè che esso è ortogonale a tutti i vettori di base. È ovvio intanto che $(w, Aw) = \lambda_1$: basta osservare che il polinomio caratteristico di una matrice è invariante per cambiamenti di base $\left((\tilde{A}_{11}-\lambda)(\lambda_1-\lambda)\cdots(\lambda_N-\lambda) = (\lambda_1-\lambda)^2(\lambda_2-\lambda)\cdots(\lambda_N-\lambda)\right)$ $\Rightarrow \tilde{A}_{11} = (w, Aw) = \lambda_1$
Ma allora:

$$(w, (A-\lambda_1)w) = 0 \quad ; \quad (v^{(j)}, (A-\lambda_1)w) = (\lambda_j - \lambda_1)(v^{(j)}, w) = 0$$

e quindi $(A-\lambda_1)w$ è il vettore nullo.
La proprietà delle matrici hermitiane di essere diagonalizzabili con una trasformazione unitaria si estende a un'altra classe importante di matrici, le matrici <u>normali</u>.

Per procedere a questa estensione, è però necessario premettere il fondamentale risultato sulla diagonalizzabilità simultanea di due matrici ~~normali~~ hermitiane:

CNES affinché due matrici hermitiane siano simultaneamente diagonalizzabili è che esse commutino.

a) Necessità.

Siano A e B due matrici hermitiane che possono essere diagonalizzate con una medesima trasformazione unitaria, cioè tali che:

$$A = U A_d U^+ \quad ; \quad B = U B_d U^+$$

Si ha allora:

$$[A,B] = [UA_dU^+, UB_dU^+] = U[A_d, B_d]U^+ =$$
$$= 0$$ (due matrici diagonali commutano sempre).

b) Sufficienza

Supponiamo che A e B commutino. Potremo sempre assumere che una delle due matrici sia in forma diagonale (i.e.: ci mettiamo nella base degli autovettori di A). Dobbiamo dimostrare che, nella medesima base, anche B viene ad essere diagonale.

Abbiamo allora; ~~in questa base~~

$$[A,B] = [A_d, \tilde{B}] = 0$$

Cioè: $(\lambda_i - \lambda_j)\tilde{B}_{ij} = 0 \qquad (i \neq j)$

È allora immediato che $\tilde{B}_{ij} = 0$ per $i \neq j$ non appena $\lambda_i \neq \lambda_j$.

Quindi il teorema, nel caso in cui tutti gli autovalori di A siano distinti, è già dimostrato.

Nel caso in cui λ_1 abbia molteplicità p_1, ... λ_s molteplicità p_s $\left(\sum_{i=1}^{s} p_i = N\right)$, quello che fin qui possiamo affermare è che la matrice $\tilde{B}$ è una matrice "diagonale a blocchi" (costituita per l'esattezza da s blocchi di dimensione $p_j \times p_j$ $(j = 1, \ldots, s)$, <u>ognuno di questi blocchi essendo di per sé hermitiano</u>. Allora ognuno di questi blocchi può essere diagonalizzato con una trasformazione unitaria $U^{(j)}$ ristretta al sottospazio p_j-dimensionale $S^{(j)}$ generato dai p_j autovettori di A associati all'autovalore λ_j.
D'altra parte, una trasformazione unitaria ristretta a $S^{(j)}$ lascia A_d invariata, in quanto ~~$A_d|_{S^{(j)}} = \lambda_j I$~~ su $S^{(j)}$ A_d agisce come un multiplo dell'identità: $U^{(j)} \lambda_j I^{(j)} U^{+(j)} = \lambda_j I^{(j)}$.
Di conseguenza A e B sono simultaneamente diagonalizzabili.
Diamo ora la seguente definizione:
Una matrice si dice normale se commuta con la sua hermitiana coniugata:

$$[N, N^+] = 0 .$$

Osserviamo ora che ogni matrice può essere scritta come combinazione lineare di due matrici hermitiane, mediante la cosiddetta rappresentazione cartesiana (analoga a quella valida per i numeri complessi):

$$A = A_1 + i A_2 \qquad A_1 = A_1^+ \; ; \; A_2 = A_2^+$$

Infatti:

$$A^+ = A_1 - i A_2$$

E quindi:

$$A_1 = \frac{1}{2}(A + A^+) \quad ; \quad A_2 = \frac{1}{2i}(A - A^+)$$

Supponiamo A normale $= N$; si ha:

$$0 = [N, N^+] = [N_1 + i N_2, N_1 - i N_2] = 2i [N_2, N_1]$$

Quindi N_1 e N_2 commutano e quindi, essendo hermitiane, sono diagonalizzabili mediante la stessa trasformazione unitaria. Ne segue che anche <u>N è diagonalizzabile con una trasformazione unitaria</u>.

Casi particolari di matrici normali sono ovviamente le matrici unitarie, le cui decomposizione spettrale assume la forma:

$$U = \sum_{j=1}^{N} e^{i\theta_j} P^{(j)}$$

(Infatti, gli autovalori di una matrice unitaria hanno modulo 1 : perché?)

I.4 Diagonalizzabilità di matrici generiche

Abbandoniamo ora l'ipotesi che A sia hermitiana o normale: ~~In questo caso~~ chiediamoci se esiste una condizione sufficiente per la diagonalizzabilità di A. La risposta è affermativa ed è contenuta nel seguente teorema:

"Condizione sufficiente affinché una matrice sia diagonalizzabile è che i suoi autovalori siano tutti distinti".

In effetti, se tale condizione è soddisfatta, una matrice $N \times N$ ha N autovettori linearmente indipendenti. Per vederlo, osserviamo anzitutto che se gli autovalori di A sono tutti distinti (cioè se tutte le radici dell'equazione caratteristica hanno molteplicità 1) il rango delle matrici $A - \lambda_k I$ $(k=1,\dots,N)$ è $N-1$, e quindi l'equazione agli autovalori

I.4.1 $$(A - \lambda_k I)\, v^{(k)} = 0$$

ammette ∞^1 soluzioni: vale a dire i vettori $v^{(k)}$ sono individuati a meno di un fattore di normalizzazione complesso $\neq 0$, in altre parole, gli N sottospazi: ~~$S^{(k)}$~~ ~~sta~~

$$S^{(k)} = \{ v^{(k)} : (A - \lambda_k I)\, v^{(k)} = 0 \}$$

sono unidimensionali.

Per dimostrare che i $v^{(k)}$ sono lin. indipendenti, supponiamo dapprima che A ~~abbia~~ non abbia autovalore nullo.

Allora, la condizione $\sum_i c_i v^{(i)} = 0$ equivale alle condizioni:

$$A^K \sum_i c_i v^{(i)} = 0 \quad (K = 0, \ldots, N-1)$$

cioè:

I.4.2 $$\sum_{i=1}^{N} \lambda_i^K c_i v^{(i)} = 0 \quad (K = 0, \ldots, N-1)$$

Osserviamo che, per ognuna delle componenti dei vettori $v^{(i)}$, le condizioni (I.4.2) costituiscono un sistema di N equazioni nelle N incognite $c_i v_j^{(i)}$ (j fissato), la cui matrice dei coefficienti è:

$$C_{iK} = \lambda_i^K$$

il cui determinante vale, com'è noto (determinante di Van der Monde):

$$\det \{C_{iK}\} = \prod_{\substack{i,K \\ i \neq K}} (\lambda_i - \lambda_K)$$

che è quindi non nullo se e solo se gli autovalori di A sono tutti distinti.

Ne consegue che, essendo gli autovalori di A tutti distinti, il sistema ammette, qualunque sia j, solo la soluzione nulla ($c_1 v_j^{(1)} = 0$, $c_2 v_j^{(2)} = 0, \ldots, c_n v_j^{(n)} = 0$), il che implica evidentemente: $c_1 v^{(1)} = 0, \ldots, c_n v^{(n)} = 0$

Poiché $v^{(1)}, \ldots v^{(n)}$ sono vettori non nulli, l'unica possibilità è che sia $C_1 = C_2 = \cdots = C_n = 0$, il che mostra che i vettori $v^{(K)}$ sono linearmente indipendenti.

Nel caso in cui A abbia un autovalore nullo (e uno solo !, nelle ipotesi fatte) basta ripetere il ragionamento svolto in precedenza, sostituendo alla matrice A la matrice $A - \mu_0 I$ essendo μ_0 un arbitrario numero complesso <u>non nullo</u> e diverso da $\lambda_2, \dots, \lambda_N$.

<u>Osservazione</u>

Ribadiamo che la condizione che gli autovalori siano tutti distinti è una condizione <u>sufficiente</u> ma non necessaria per la diagonalizzabilità di A.
Un controesempio è fornito dalla matrice:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

che pur non essendo né hermitiana né normale e avendo due autovalori uguali ammette 3 autovettori linearmente indipendenti.

Una volta costruiti gli autovettori indipendenti $\{v^{(i)}\}_{i=1}^{N}$ è immediato verificare che la seguente trasformazione:

1.43 $$\tilde{A} = T^{-1} A T$$

essendo $$T_{ij} = v_i^{(j)}$$

porta A nella matrice $\tilde{A}$ <u>diagonale</u>, i cui elementi sono gli autovalori di A

Si ha infatti:

I.4.4. $\tilde{A}_{jk} = \sum_{l,m} (T^{-1})_{jl} \, A_{lm} T_{mk} =$

$$= \sum_{l,m} (T^{-1})_{jl} A_{lm} v_m^{(k)} = \sum_{l,m} (T^{-1})_{jl} \lambda_k v_l^{(k)} =$$

$$= \lambda_k \sum_{lm} (T^{-1})_{jl} T_{lk} = \lambda_k \delta_{jk}$$

Viceversa:

I.4.5 $A = T \tilde{A} T^{-1}$

La formula I.4.5 permette di scrivere una decomposizione spettrale per A, a mezzo di operatori idempotenti, che pur non essendo hermitiani, proiettano su sottospazi linearmente indipendenti, e soddisfano le relazioni:

$$P^{(j)} P^{(k)} = \delta_{jk} P^{(k)}$$

$$\sum_i P^{(i)} = I$$

Infatti, dalla I.4.5:

I.4.6 $A_{jk} = \sum_{l} T_{jl} \lambda_l (T^{-1})_{lk} = \sum_l \lambda_l P_{jk}^{(l)}$

I.4.7 $P_{jk}^{(l)} := T_{jl} (T^{-1})_{lk}$

I.4.8 $(P^{(l)} P^{(m)})_{jk} = \sum_r P_{jr}^{(l)} P_{rk}^{(m)} = \sum_r T_{jl} (T^{-1})_{lr} T_{rm} T^{-1}_{mk}$

$$= \cancel{\Sigma} \; T_{jl} \, \delta_{lm} (T^{-1})_{mk} = \delta_{lm} \, T_{jl} (T^{-1})_{lk} = \delta_{lm} P_{jk}^{(l)}$$

Appendice

Teorema di Cayley-Hamilton

Il teorema di Cayley-Hamilton afferma che ogni matrice soddisfa la sua equazione caratteristica.
In altre parole, se $P(\lambda)$ è il polinomio caratteristico di una matrice:

$$P(\lambda) = \det\left[A - \lambda I\right] \qquad \cancel{(-1)^N \lambda^N} \quad (-1)^N\left[\lambda^N - \mathrm{tr}A\,\lambda^{N-1} \ldots - \det A\right] =$$

$$A^N - \mathrm{tr}A\,A^{N-1} \ldots = 0$$

in modo tale che gli zeri di $P(\lambda)$ forniscono gli autovalori di A, la matrice A soddisfa l'equazione:

$$P(A) = 0$$

(Da cui discende tra l'altro che una matrice $N \times N$ ha al più N potenze indipendenti).

Questo teorema vale qualunque sia A. Noi lo dimostreremo unicamente nel caso di matrici diagonalizzabili.
Se A è diagonalizzabile, vale la rappresentazione spettrale:

$$A = \sum_{i=1}^{N} \lambda_i P^{(i)}$$

Si ha quindi:

$$A^2 = \left(\sum_i \lambda_i P^{(i)}\right)\left(\sum_j \lambda_j P^{(j)}\right) =$$

$$= \sum_i \lambda_i \lambda_j P^{(i)} P^{(j)} = \sum_i \lambda_i \lambda_j \delta_{ij} P^{(j)} = \sum_i \lambda_i^2 P^{(i)}$$

E analogamente (per induzione):

$$A^n = \left(\sum_i \lambda_i P^{(i)}\right)\left(\sum_j \lambda_j^{n-1} P^{(j)}\right) = \sum_{i=1}^{N} \lambda_i^n P^{(i)}$$

Consideriamo ora il polinomio caratteristico $P(\lambda)$: per definizione di autovalore, si ha $P(\lambda_i) = 0$ $(i = 1, \dots, N)$.

Di conseguenza:

$$P(A) := \sum_{n=0}^{N} c_n A^n = \sum_{n=0}^{N} c_n \left(\sum_{i=1}^{N} \lambda_i^n P^{(i)}\right) =$$

$$= \cancel{\sum_{n=0}^{N} \sum_{i=1}^{n}} \quad \sum_{i=1}^{N} \left(\sum_{n=0}^{N} c_n \lambda_i^n\right) P^{(i)} = \sum_{i=1}^{N} P(\lambda_i) P^{(i)} = 0$$

I.5 Funzioni di matrici

Sappiamo finora calcolare polinomi di matrici.
Vogliamo vedere se siamo in grado di definire funzioni di matrici.

Cominciamo con l'ipotizzare che A sia diagonalizzabile. Sia $f(x)$ una funzione della variabile reale x (più in generale, $f(z)$ una funzione della variabile complessa z).
Diamo la seguente definizione:
la funzione $f(A)$ è data dalla formula

$$f(A) := \sum_{i=1}^{N} f(\lambda_i)\, P^{(i)} \tag{1.5.1}$$

E quindi esiste non appena f è definita nei punti $\lambda_1, \dots \lambda_n$ che corrispondono agli autovalori di A.

Esempio
Sia $f(A) = A^{-1}$
Abbiamo:

$$A^{-1} = \sum_{i=1}^{N} \frac{1}{\lambda_i} P^{(i)}$$

che esiste se e solo se A non ha autovalori nulli: infatti, se uno degli autovalori di A è nullo, si ha $\det A = 0$, e di conseguenza A non è invertibile.
Consideriamo ora la funzione:

$$f(A) = (A - \lambda I)^{-1}$$

detta <u>risolvente</u> della matrice A, che si indica con $R(A;\lambda)$.

Risulta:

$$R(A;\lambda) = \sum_{i=1}^{N} \frac{1}{\lambda_i - \lambda} P^{(i)}$$

Quindi $R(A;\lambda)$ esiste ~~non appena~~ $\forall\ \lambda \neq \lambda_i \ (i=1,\dots,N)$

Dalla definizione data in (I.5.1) di funzione di matrice, segue per esempio che ogni matrice unitaria può essere sempre posta nella forma:

$U = e^{iH}$, essendo H una matrice hermitiana.

Infatti:

a) Se H è hermitiana, si ha:

$$e^{iH} = \sum_{j=1}^{N} e^{i\lambda_j} P^{(j)}$$

$$(e^{iH})^{-1} = \sum_{j=1}^{N} \frac{1}{e^{i\lambda_j}} P^{(j)} = \sum_{j=1}^{N} e^{-i\lambda_j} P^{(j)} = e^{-iH} =$$

$$= (e^{iH})^{+} \quad \left(\text{poiché } H \text{ è hermitiana, } \lambda_j = \overline{\lambda_j} \ \underline{e} \ P^{(j)} = [P^{(j)}]^{+}\right)$$

Quindi U è unitaria

b) Viceversa, se U è unitaria, vale la rappresentazione spettrale

$$U = \sum_{j=1}^{N} u_j P^{(j)}, \quad \text{con } |u_j|^2 = 1, \quad P^{(j)} = [P^{(j)}]^{+}$$

Possiamo quindi porre: $u_j = e^{i\lambda_j}$ (λ_j <u>reale</u>), e di conseguenza

definendo
$$H = \sum_{i=1}^{N} \lambda_i P^{(i)}$$

abbiamo $H = H^{+}$, $U = e^{iH}$.

Sempre dalla definizione di funzione di matrice, segue che, per effetto di un cambiamento di base, che porta A nella matrice $\tilde{A} = T^{-1}AT$, si ha:

I.5.2 $f(\tilde{A}) = f(T^{-1}AT) =$

$$= \sum_{i=1}^{N} f(\lambda_i)\, T^{-1} P^{(i)} T = T^{-1}\Big(\sum_i f(\lambda_i) P^{(i)}\Big) T = T^{-1} f(A)\, T$$

In particolare, se T è la matrice le cui colonne sono costituite dalle componenti degli autovettori di A (cioè la matrice che diagonalizza A), ricordando la definizione dei $P^{(i)}$, risulta:

I.5.3 $f(\tilde{A}) = \begin{pmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{pmatrix}$

e quindi:

I.5.4 $f(A) = T \begin{pmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{pmatrix} T^{-1}$

Esiste anche un'altra definizione di funzione di matrice, che mostreremo essere equivalente alla precedente nel caso in cui esse siano entrambe applicabili.

A tale scopo, supponiamo che la funzione di variabile complessa $f(z)$ sia sviluppabile in serie di potenze ~~per~~ per $|z| < R$, sicché si possa scrivere:

$$f(z) = \sum_{n=0}^{\infty} f_n z^n \tag{I.5.3}$$

la convergenza essendo assoluta e uniforme per $|z| < R$.

Possiamo allora essere indotti a <u>definire</u>:

$$f(A) := \sum_{n=0}^{\infty} f_n A^n \tag{I.5.4}$$

Si tratta di vedere quando la (I.5.4) ha senso, cioè quando la serie a 2º membro della (I.5.4) converge.

Vedremo che si possono dare vari criteri di convergenza.

a) Supponiamo di chiedere una sorta di "convergenza puntuale" della serie (I.5.4), cioè una convergenza dei singoli elementi di matrice della serie (si tratta in effetti di una convergenza debole).

Osserviamo allora che

$$|(A^2)_{ij}| \leq \sum_{k=1}^{N} |A_{ik}||A_{kj}| \leq N M^2$$

essendo $M = \max_{i,j} |A_{ij}|$

Da cui, per induzione:

$$|(A^n)_{ij}| \leq N^{n-1} M^n$$

E quindi:

$$\left| \sum_n f_n A_n \right|_{ij} \leq \sum_n |f_n| |(A^n)_{ij}| \leq \frac{1}{N} \sum_n |f_n| (MN)^n$$

che converge non appena $NM < R$ (R è il raggio di convergenza della serie I.5.3)

Quindi, condizione sufficiente affinché la (I.5.4) converga per ogni elemento di matrice è che risulti $NM < R$.

b) Supponiamo ora di dotare $\mathbb{C}^N$ di una norma. Allora lo spazio degli operatori lineari su $\mathbb{C}^N$ diventa uno spazio normato completo, con:

$$\|A\| = \sup_{x \in \mathbb{C}^N} \|Ax\| / \|x\|$$

Possiamo quindi chiederci quando la serie (I.5.4) converge in norma.

Osserviamo allora che:

$$\left\| \sum_{n=0}^{P} f_n A^n \right\| \leq \sum_{n=0}^{P} |f_n| \|A^n\| \leq \sum_{n=0}^{P} |f_n| (\|A\|)^n$$

Di conseguenza, condizione sufficiente affinché la (I.5.4) converga in norma, è che $\|A\| < R$.

Consideriamo alcuni esempi di norme in $\mathbb{C}^N$ e vediamo che cosa possiamo dire per la norma di $\|A\|$.

(i) $\|x\| = \max\limits_{1\le k\le N} |x_k|$

Risulta:

$$\|Ax\| = \max_j |Ax|_j = \max_j |\sum_k A_{jk} x_k| \le \max_j \sum_k |A_{jk}||x_k|$$

$$\le \|x\| \max_j \sum_k |A_{jk}|$$

$$\Rightarrow \|A\| \le \max_j \sum_k |A_{jk}|$$

(ii) $\|x\| = \sum\limits_{i=1}^{N} |x_i|$

Risulta:

$$\|Ax\| = \sum_{i=1}^{N} |(Ax)_i| \le \sum_{i=1}^{N}\sum_{j=1}^{N} |A_{ij}||x_j| \le$$

$$\sum_{i=1}^{N} \max_j |A_{ij}| \left(\sum_j |x_j|\right) = \max_j \sum_{i=1}^{N} |A_{ij}| \; \|x\|$$

$$\Rightarrow \|A\| \le \max_j \sum_{i=1}^{N} |A_{ij}|$$

(iii) $\|x\| = \left[\sum\limits_{i=1}^{N} |x_i|^2\right]^{1/2} = (x,x)^{1/2}$

$$\|Ax\|^2 = (Ax, Ax) = \sum_i \overline{(Ax)}_i (Ax)_i =$$

$$= \sum_{ijk} \overline{A}_{ij}\,\overline{x}_j\, A_{ik} x_k = \sum_i (\overline{A}^{(i)}, \overline{x})(\overline{A}^{(i)}, x)$$

essendo $A^{(i)}$ il vettore di component $A_{i1}, \dots, A_{iN}$.

Applicando la diseg. di Cauchy - Schwarz:

$$(Ax, Ax) = |(Ax, Ax)| \le \sum_i |(A^{(i)}, \overline{x})||(\overline{A}^{(i)}, x)|$$

$$\le \|\overline{x}\|\|x\| \sum_i \|\overline{A}^{(i)}\|\|A^{(i)}\| = \|x\|^2 \sum_i \|A^{(i)}\|^2$$

$$\Rightarrow \frac{\|Ax\|}{\|x\|} \le \sum_{i=1}^{N}\sum_{j=1}^{N} |A_{ij}|^2 \Rightarrow \|A\| \le \left(\sum_{ij} |A_{ij}|^2\right)^{1/2}$$

Anche con la definizione (I.5.4), $f(A)$ si trasforma come A per cambiamenti di base. In effetti, se $\tilde{A} = T^{-1} A T$

$$\sum_{n=0}^{p} f_n \tilde{A}^n = \sum_{n=0}^{p} f_n T^{-1} A^n T = T^{-1} \left(\sum_{n=0}^{p} f_n A^n \right) T$$

da cui, se la ~~co~~ serie converge (in ~~f~~ norma o assolutamente per i singoli elementi di matrice):

$$f(\tilde{A}) = T^{-1} f(A) T$$

Osserviamo infine che, se entrambe le definizioni di funzione di matrice sono applicabili (vale a dire, A è diagonalizzabile e $f(z)$ ~~Co~~ è sviluppabile in serie di potenze per $|z| < R$) si ha:

$$f(A) = \sum_{i=1}^{N} f(\lambda_i) P^{(i)} = (\max_i |\lambda_i| < R) =$$

$$= \sum_{i=1}^{N} \left(\sum_{n=0}^{\infty} f_n \lambda_i^n \right) P^{(i)} = \sum_{n=0}^{\infty} f_n \sum_{i=1}^{N} \lambda_i^n P^{(i)} =$$

$$= \sum_{n=0}^{\infty} f_n A^n$$

(e viceversa)

I.6 Matrici come spazio lineare a N^2 dimensioni. Spazi euclidei di matrici e basi ortonormali. Matrici di Pauli.

È ovvio che le matrici $N \times N$, ~~in~~ cioè l'insieme degli operatori limitati su $\mathbb{C}^N$, costituiscono uno spazio lineare a N^2 dimensioni (la somma e la moltiplicazione per un numero complesso essendo intese come somme degli elementi omologhi e come moltiplicazione di ogni elemento per il numero complesso). La base indotta in questo spazio dalla base naturale in $\mathbb{C}^N$ è la base $e^{(i,j)}$ $(i,j=1,\dots,N)$ i cui elementi sono dati dal "prodotto esterno":

$$e^{(i,j)} = e^{(i)} \otimes e^{(j)}$$

(prodotto della "colonna" $e^{(i)}$ per la "riga" $e^{(j)}$). Questo spazio si può dotare della struttura di uno spazio metrico euclideo, introducendo il prodotto scalare:

I.6.1. $(A,B) = \mathrm{Tr}\, A^{+}B$

Esso soddisfa le proprietà:

$$(A,B) = \overline{(B,A)}$$
$$(\alpha A, B) = \bar{\alpha}(A,B)$$
$$(A, \alpha B) = \alpha (A,B)$$
$$(A,A) \geq 0 \quad (=0 \Leftrightarrow A=0)$$

Ciò permette di definire la norma di una matrice $\|A\| = (A,A)^{1/2}$ e l'angolo fra due matrici (Diseg. di Cauchy-Schwarz).

Si possono quindi definire anche basi ortogonali (o ortonormali) di matrici, indicate di solito con σ_n $(n=0,\dots,N^2-1)$ tali che:

I.6.2 $(\sigma_n, \sigma_m) = \delta_{nm}$

I.6.3 $A = \sum_n a_n \sigma_n \quad ; \quad a_n = \mathrm{Tr}\, \sigma_n^+ A$

le matrici più note di questo tipo sono le matrici di Pauli, che costituiscono una base ortogonale (<u>non ortonormale</u>) ~~di matrici~~ per le matrici 2x2.
Una loro rappresentazione esplicita è la seguente:

I.6.4 $\sigma_0 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \; ; \; \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \; ; \; \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \; ; \; \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

Esse verificano le proprietà:

I.6.5 $\sigma_n = \sigma_n^+ \quad ; \quad \sigma_n^2 = I = \sigma_0 \quad ; \quad \sigma_k \sigma_j = i \sigma_l$ $(k, j, l$ perm. ciclica di $1,2,3)$

Si ha quindi:

I.6.6 $(\sigma_n, \sigma_m) := \mathrm{Tr}\, \sigma_n \sigma_m = 2\, \delta_{n,m} \quad (n,m = 0,1,2,3)$

I.6.7 $A = a_n \sigma_n \quad ; \quad a_n = \frac{1}{2} \mathrm{Tr}\, (\sigma_n A)$

Inoltre, dalle I.6.5 segue, definendo

$\{\sigma_m, \sigma_n\} := \sigma_m \sigma_n + \sigma_n \sigma_m$

$\cancel{\{\sigma_n, \sigma_m\} := \sigma_n \sigma_m + \sigma_m \sigma_n = [\delta_{n,0}\, \sigma_m + \delta_{m,0}\, \sigma_n](2 - \delta_{m,n})}$

I.6.8 ~~cioè~~ $\{\sigma_0, \sigma_m\} = 2\sigma_m \quad ; \quad \{\sigma_i, \sigma_j\} = 0 \quad (i,j = 1,2,3)$

e inoltre:

$$[\sigma_0, \sigma_n] = 0 \quad ; \quad [\sigma_k, \sigma_j] = 2i\,\epsilon_{kjl}\,\sigma_l \qquad (j,k,l = 1,2,3)$$

Ciò implica:

$$AB = \left(a_0\sigma_0 + \sum_{j=1}^{3} a_j\sigma_j\right)\left(b_0\sigma_0 + \sum_{k=1}^{3} b_k\sigma_k\right) =$$

$$= \left[a_0b_0 + \sum_{j=1}^{3} a_j b_j\right]\sigma_0 + i\sum_{jk} \epsilon_{jkl}\, a_j b_k \sigma_l$$

E ponendo: $\vec{a}\cdot\vec{b} = \sum_j a_j b_j \,;\, (\vec{a}\wedge\vec{b})_j = \epsilon_{jkl}\, a_k b_l$

otteniamo:

$$AB = [a_0b_0 + \vec{a}\cdot\vec{b}]\,\sigma_0 + i\,(\vec{a}\wedge\vec{b})\cdot\hat{\sigma}$$

Da cui:

$$[A,B] = 2i\,(\vec{a}\wedge\vec{b})\cdot\hat{\sigma}$$

Come applicazione, consideriamo l'esponenziale di una matrice a traccia nulla.

Si ha allora: $A = \sum_{i=1}^{3} a_i\sigma_i := \vec{a}\cdot\hat{\sigma}$ $\quad (\vec{a} = i\theta\hat{n})$

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!} \qquad (A^0 = I)$$

Ma: $A^2 = (\vec{a}\cdot\hat{\sigma})(\vec{a}\cdot\hat{\sigma}) = \vec{a}\cdot\vec{a}\,\sigma_0 := a^2\sigma_0$

$A^3 = a^2\sigma_0 A = a^2\,\vec{a}\cdot\hat{\sigma}$

Quindi:

$$A^{2k} = a^{2k}\sigma_0 \quad ; \quad A^{2k+1} = a^{2k}\,\vec{a}\cdot\hat{\sigma}$$

E in definitive:

$$\exp(A) = \sum_{k=0}^{\infty} \frac{A^{2k}}{2k!} + \sum_{k=0}^{\infty} \frac{A^{2k+1}}{(2k+1)!} =$$

$$= \sigma_0 \cosh a + \frac{\vec{a}\cdot\hat{\sigma}}{a} \operatorname{senh} a$$

I.7 Sistemi di equazioni differenziali ordinarie a coefficienti costanti e funzioni di matrici.

A) ~~Un sistema di eq~~ Consideriamo il seguente sistema di equazioni differenziali ordinarie:

I.7.1 $$\dot{x}_j = \sum_{K=1}^{N} A_{jK}\, x_K + b_j \qquad (j=1,\dots,N)$$

ovvero:

$$\dot{x} = Ax + b \qquad A = \text{cost.}, \quad b = b(t)$$

Supponiamo di voler risolvere l'associato problema di Cauchy, specificato dalla condizione iniziale:

$$x(0) = \bar{x}$$

È facile vedere che la soluzione è data da:

I.7.2 $$x(t) = e^{At}\left[\int_0^t dt'\, e^{-At'}\, b(t') + \bar{x}\right],$$

non appena si osservi che $\frac{d}{dt} e^{At} = A e^{At} = e^{At} A$,

come si ricava facilmente dallo sviluppo in serie.

Infatti, la (I.7.2) implica chiaramente $x(0) = \bar{x}$, e inoltre da essa si ottiene

$$\dot{x} = A\left[e^{At}\int_0^t dt'\, e^{-At'}\, b(t') + \bar{x}\right] + e^{At}\cdot e^{-At}\, b(t) = Ax + b$$

In particolare, se b è costante e A è invertibile:

$$x(t) = e^{At}\,\bar{x} - A^{-1}\left(I - e^{At}\right) b = e^{At}\left(\bar{x} + A^{-1}b\right) - A^{-1}b$$

Se A è diagonalizzabile:

$$e^{At} = \sum_i e^{\lambda_i t} P^{(i)}$$

e quindi:

13 $$x(t) = \sum_{i=1}^{N} \left[e^{\lambda_i t}\left(P^{(i)}\bar{x} + \frac{1}{\lambda_i} P^{(i)} b\right) - \sum_i \frac{1}{\lambda_i} P^{(i)} b\right]$$

Da cui:

$$P^{(j)} x(t) = e^{\lambda_j t} P^{(j)}\bar{x} + \frac{1}{\lambda_j}\left(e^{\lambda_j t} - 1\right) P^{(j)} b$$

che ci dà in funzione del tempo il valore delle componenti del vettore x nella base degli autovettori di A.

(Come cambiano le formule precedenti se uno degli autovettori di A è nullo?)

B) Consideriamo ora il seguente sistema matriciale:

I.7.4 $\dot{X} = AX + XB \quad ; \quad X(0) = \bar{X}$

dove $\bar{X}, A, B$ sono matrici $N \times N$
la matrice $X(t)$, soluzione di (I.7.4), vale:

I.7.5 $X(t) = e^{At}\,\bar{X}\,e^{Bt}$

Infatti:

$$\dot{X} = A e^{At}\bar{X}e^{Bt} + e^{At}\bar{X}e^{Bt}B = AX + XB$$

Caso ~~parti~~ particolare: $B = -A$. Si ha:

$$X(t) = e^{At}\,\bar{X}\,e^{-At} = T\bar{X}T^{-1} \qquad (T = e^{At})$$

Da cui si vede che $X(t)$ è simile a $\bar{X}$: in particolare, ha gli stessi autovalori di $\bar{X}$ (o, in altre parole, gli autovalori di X sono costanti del moto).