Algebretta per la Meccanica Quantistica

P. Decleva

March 26, 2013

Contents

1 Premessa

Queste note contengono un'introduzione minima al linguaggio di Algebra Lineare e alle proprieta' fondamentali adoperate in Meccanica Quantistica. Contengono alcune estensioni a quanto gia' visto nel corso di Matematica II. In particolare

1. Spazi sul campo complesso
2. Proprieta' ulteriori degli operatori lineari
3. Proprieta' ulteriori dell'equazione ad autovalori
4. In particolare legame tra commutativita' di operatori e basi di autovettori comuni.

2 Introduzione

Dal corso di matematica II siamo familiari con

1. I vettori geometrici
$$\vec{v} = \vec{AB}$$
nel piano e nello spazio, e le operazioni di somma di due vettori e di moltiplicazione di un vettore per un numero reale.

2. Gli spazi
$$R^n = \{(x_1, \ldots, x_n)\}$$
insiemi delle n-uple di numeri reali, in cui sono naturalmente definite le operazioni di somma e di moltiplicazione per un numero reale
$$x = (x_1, \ldots, x_n) \qquad y = (y_1, \ldots, y_n)$$
$$x + y = (x_1 + y_1, \ldots, x_n + y_n) \qquad ax = (ax_1, \ldots, ax_n)$$

3. Si e' anche visto che introdotti assi cartesiani nel piano o nello spazio esiste una corrispondenza naturale tra vettori geometrici e spazi R^2 o R^3.

In realta' le proprieta' algebriche, legate cioe' alle proprieta' della somma di due vettori e al prodotto di un vettore per un numero, sono comuni a molti altri oggetti (matrici, funzioni, ecc.), cosicche' e' utile formalizzarle e studiarle in astratto una volta per tutte, prescindendo cioe' dalla natura particolare degli enti che godono di queste proprieta'. In tal modo, per ogni insieme di oggetti su cui siano definite le due operazioni viste, e che soddisfino alle stesse proprieta' fondamentali (assiomi), varranno tutte le relazioni che verranno studiate, senza bisogno di riverificarle in ogni caso particolare.

3 Spazi Vettoriali

Per definire uno spazio vettoriale occorrono due insiemi, un insieme di elementi

$$V = \{x, y, z, \ldots\}$$

detti vettori, e un insieme di numeri (campo)

$$K = \{a, b, c, \ldots\}$$

detti scalari. Negli esempi visti sopra, $K = R$, si e' adoperato il campo dei numeri reali, ma e' immediata l'estensione al campo dei numeri complessi, che e' indispensabile in MQ. Del resto e' facile convincersi che negli spazi C^n, costituiti dalle n-uple di numeri complessi, le due operazioni di somma e prodotto (adesso per un numero complesso), definite come in R^n, soddisfano le stesse proprieta' che nel caso reale. (Chi si sentisse in dubbio ci rifletta finche' si senta a proprio agio).

Di fatto si puo' adoperare un campo numerico qualunque, ma questa generalita' non serve nel seguito, percio' per noi il campo K sara' sempre quello dei numeri complessi C, oppure dei numeri reali R, come sottocaso con minime e ovvie modifiche. Definiamo allora

Uno spazio vettoriale

$$V = \{x, y, z, \ldots\}$$

su un campo K e' un insieme su cui sono ovunque definite due operazioni, somma di due vettori, e moltiplicazione di un vettore per un numero. Questo significa che per ogni coppia di elementi $x, y \in V$ e' ben definito il vettore somma $z = x + y,\ z \in V$, e per ogni coppia $a \in K,\ x \in V$, e' ben definito il vettore prodotto $ax \in V$. Le operazioni devono soddisfare le seguenti proprieta'.

A 1 (Proprieta' associativa) *per ogni* $x, y, z \in V$ *e'*

$$(x + y) + z = x + (y + z)$$

A 2 (Esistenza dell'elemento neutro) *esiste un elemento* $0 \in V$ *tale che,* $\forall x \in V$ *e'*

$$x + 0 = 0 + x = x$$

A 3 (Esistenza dell'inverso) *per ogni* $x \in V$ *esiste un elemento* $-x \in V$ *tale che*

$$x + (-x) = -x + x = 0$$

A 4 (Proprieta' commutativa) *per ogni* $x, y \in V$ *e'*

$$x + y = y + x$$

Per queste proprieta' si dice che V e' un gruppo commutativo (o abeliano) rispetto alla somma. Un secondo gruppo di assiomi riguarda le proprieta' del prodotto per un numero $a \in K$

P 1 (Proprieta' distributiva rispetto alla somma in V) *per ogni* $a \in K, \quad x, y \in V$ *e'*

$$a(x + y) = ax + ay$$

P 2 (Proprieta' distributiva rispetto alla somma in K) *per ogni* $a, b \in K, \quad x \in V$ *e'*

$$(a + b)x = ax + bx$$

P 3 (Proprieta' associativa del prodotto) *per ogni* $a, b \in K, \quad x \in V$ *e'*

$$a(bx) = (ab)x$$

P 4 (Unita' in K) *se* 1 *e' l'unita' in* K*, per ogni* $x \in V$ *e'*

$$1x = x$$

Osservazioni

- Lo 0 di K e lo 0 di V sono due oggetti diversi, il primo e' un numero, il secondo un vettore, e a rigore si dovrebbero usare simboli diversi. Pero' non si usa, ed e' chiaro dal contesto a chi ci si riferisce.

- l'elemento neutro 0 in V e' unico, e cosi' l'inverso $-x$ di ogni elemento
- si scrive in forma abbreviata

$$x + (-y) \equiv x - y$$

- Dagli assiomi seguono le regole di calcolo

$$0x = 0 \quad \text{e} \quad a0 = 0$$

$$ax = 0 \Rightarrow a = 0 \quad \text{oppure} \quad x = 0$$

$$(-1)x = -x \quad \text{e anche} (-a)x = -(ax) = a(-x) \equiv -ax$$

  Ad esempio

$$x = 1x = (1 + 0)x = 1x + 0x = x + 0x$$

  e cancellando x da entrambi i membri si ottiene la prima $0x = 0$. Analogamente le altre.

Esempi

1. K^n e' spazio vettoriale su K, con le solite operazioni, come gia' visto. In particolare, per $n = 1$, K e' spazio vettoriale su se' stesso.
2. Anche C si puo' considerare spazio vettoriale su R
3. Si puo' considerare l'insieme delle successioni

$$K^\infty = \{(a_1, a_2, \ldots, a_n, \ldots)\}$$

   di elementi di K, $a_n \in K$. E' spazio vettoriale con le solite operazioni definite separatamente su ogni componente, come nel caso finito.
4. Cosi' anche l'insieme delle successioni (a_n) quasi tutte nulle $(q.t.n.)$

$$K^\omega = \{(a_1, a_2, \ldots, a_n, \ldots), a_n \ q.t.n.\}$$

   dove (a_n) quasi tutte nulle significa che solo un numero finito di componenti sono diverse da zero. Osserviamo che sia la somma di due successioni $q.t.n.$ che il prodotto per un numero forniscono ancora successioni dello stesso tipo.

5. L'insieme P_n dei polinomi di grado massimo n

$$P_n = \{a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n\}$$

e anche l'insieme P di tutti i polinomi di qualunque grado.

6. L'insieme $K^{m\times n}$ delle matrici $m \times n$ con elementi in K, con le solite operazioni

$$A = (a_{ij}), \quad B = (b_{ij}) \;\; \Rightarrow \;\; A + B = (a_{ij} + b_{ij}), \quad aA = (a\, a_{ij})$$

7. Esempio molto importante: se X e' un insieme qualunque, $X = \{x, y, \ldots\}$, e V uno spazio vettoriale su K, l'insieme di tutte le funzioni di X in V

$$F(X, V) = \{f : X \to V\}$$

e' spazio vettoriale su K, con le definizioni naturali

$$(f + g)(x) = f(x) + g(x) \qquad (af)(x) = af(x)$$

Osserviamo che $f(x), g(x)$ sono vettori di V, e quindi le operazioni di somma e prodotto per un numero sono ben definite su di essi. Si verifica subito che F e' vettoriale su K (perche' eredita la struttura da V). In particolare V puo' ridursi a K stesso.

Questo esempio copre un insieme molto vasto di casi particolari, ad esempio l'insieme di tutte le funzioni $I \to R$, dove I e' un intervallo di R, che puo' anche estendersi a tutto R, l'insieme delle funzioni reali con due derivate continue, le funzioni $R \to C$, etc.

Del resto le n-uple $(a_1, \ldots a_n)$ sono le funzioni

$$\{1, 2, \ldots, n\} \to K$$

e analogamente le successioni, etc.

3.1 combinazioni lineari, dipendenza e indipendenza lineare

Data una famiglia di vettori $(x_i)_{i=1,\ldots,n}$ si dice loro combinazione lineare un'espressione del tipo

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = \sum_{i=1}^{n} a_i x_i$$

dove gli $a_i \in K$ si dicono coefficienti della combinazione lineare. Una combinazione lineare ha sempre un numero finito di addendi, quindi e' perfettamente definita. Nel caso di una famiglia infinita $(x_i)_{i\in I}$, per combinazione lineare degli x_i si intende ancora un'espressione del tipo visto, con i coefficienti a_i quasi tutti nulli, cioe' solo un numero finito di $a_i \neq 0$, cosi' che la somma si riduce sempre a un numero finito di termini.

Una famiglia di vettori (x_i) si dice linearmente dipendente se esiste una combinazione lineare nulla

$$\sum_i a_i x_i = 0$$

con i coefficienti non tutti nulli, cioe' con almeno un $a_i \neq 0$. Se una famiglia non e' linearmente dipendente, si dice che e' linearmente indipendente. Se (x_i) e' linearmente indipendente, la condizione

$$\sum_i a_i x_i = 0 \quad \Rightarrow \quad a_i = 0 \; \forall i$$

implica che tutti gli a_i siano nulli.

Si dice anche che i vettori x_i sono linearmente dipendenti (l.d.) o linearmente indipendenti (l.i.).

Se (x_i) e' l.d. , allora uno degli x_i puo' essere espresso come combinazione lineare dei rimanenti. Infatti esiste almeno un $a_k \neq 0$ tale che

$$\sum_i a_i x_i = 0$$

ovvero

$$\sum_{i\neq k} a_i x_i + a_k x_k = 0$$

$$x_k = -\frac{1}{a_k} \sum_{i\neq k} a_i x_i = -\sum_{i\neq k} \frac{a_i}{a_k} x_i$$

e viceversa, se un vettore e' combinazione lineare di altri, la famiglia ottenuta aggiungendo questo vettore ai precedenti, e' certamente l.d. Si verifica subito che se a una famiglia dipendente si aggiunge un vettore, la famiglia resta l.d., se a una famiglia l.i. si toglie un vettore, questa resta l.i. E piu' in generale, ogni famiglia che contenga una sotto famiglia l.d. e' ancora l.d., viceversa ogni sottofamiglia di una famiglia l.i. e' ancora l.i. Banalmente, ogni famiglia che contenga il vettore nullo e' l.d.

3.2 Sottospazi

Dato V su K, un sottoinsieme S di V si dice sottospazio se e' ancora spazio vettoriale su K con le stesse operazioni ereditate da V, cioe' gli elementi di S soddisfano ancora gli assiomi visti. Condizione necessaria e sufficiente perche' S sia sottospazio e' che sia

$$\forall\, x, y \in S, \;\; a \in K \quad \Rightarrow \quad x + y \in S, \;\; ax \in S$$

ovvero che la somma di due elementi qualunque di S stia ancora in S, e cosi' anche il prodotto di un elemento qualunque di K per un elemento di S. O in generale, che una qualunque combinazione lineare di elementi di S stia ancora in S. Si verifica senza difficolta' che questo basta a soddisfare tutti gli assiomi, perche' tutte le proprieta' che erano vere in V continuano a essere vere in S.

In particolare, data una famiglia arbitraria $X = (x_i)$ di elementi di V, si definisce il sottospazio S generato da X come l'insieme di tutte le combinazioni lineari di elementi di X

$$S = \left\{ \sum_i a_i x_i \right\}$$

che si indica anche con $S = \langle X \rangle$. Si verifica immediatatmente che e' un sottospazio, poiche' la somma di due combinazioni lineari, e il prodotto di una combinazione lineare per un numero, sono ancora una combinazione lineare degli stessi elementi. Questo sottospazio e' il sottospazio minimo tra tutti quelli che contengono gli elementi di X, infatti devono automaticamente contenerne tutte le combinazioni lineari. Si dice anche che X e' una famiglia di generatori per S.

3.3 Basi e componenti

Sia data una famiglia $X = (x_i)$ tale che

1. X e' una famiglia di generatori per tutto lo spazio V

$$\langle X \rangle = V$$

2. X e' linearmente indipendente

Una famiglia che soddisfi a queste due proprieta' si dice una base nello spazio. La proprieta' piu' importante della base e' che permette di esprimere in modo unico ogni elemento dello spazio come combinazione lineare degli elementi di base. Sia $y \in V$ un elemento dello spazio. Poiche' X genera V e' certamente

$$y = \sum_i a_i x_i$$

Vediamo l'unicita' degli a_i. Supponiamo che y si possa esprimere con un'altra combinazione lineare

$$y = \sum_i b_i x_i$$

Allora

$$0 = \sum_i a_i x_i - \sum_i b_i x_i = \sum_i (a_i - b_i) x_i$$

da cui, per l'indipendenza lineare degli x_i, e'

$$(a_i - b_i) = 0 \qquad a_i = b_i$$

I coefficienti a_i si dicono componenti del vettore y nella base $X = (x_i)$ assegnata. Da quanto visto, fissata una base, vi e' una corrispondenza biunivoca tra un vettore y e la famiglia delle sue componenti, che si possono considerare un vettore di C^n (nel caso di una base finita di n vettori). E' utile in tal caso adoperare la notazione matriciale raccogliendo le componenti in un vettore colonna a

$$a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

Si dimostra che:

1. Tutti gli spazi vettoriali ammettono una base
2. Tutte le basi di uno stesso spazio hanno lo stesso numero di elementi. Ovvero possono essere messe in corrispondenza biunivoca. Se questo numero e' finito, n, si dice dimensione dello spazio

   $$dim\ V = n$$

   e rappresenta il massimo numero di vettori linearmente indipendenti. Se invece esistono infiniti vettori linearmente indipendenti, si dice che lo spazio ha dimensione infinita.

3. Dati due sottoinsiemi T e S di V, $T \subset S \subset V$, tali che T sia linearmente indipendente e S un insieme di generatori di V esiste sempre una base X di V tale che
$$T \subset X \subset S$$
Questo significa che
    - Ogni insieme linearmente indipendente si puo' estendere a una base (aggiungendo ulteriori elementi linearmente indipendenti)
    - Da ogni insieme di generatori si puo' estrarre una base (eliminando uno alla volta i vettori linearmente dipendenti dai rimanenti)

Osservazioni:

- Se $dim\ V = n$ ne segue che ogni insieme di m vettori, con $m < n$, non puo' generare lo spazio, e ogni insieme con $m > n$ e' necessariamente linearmente dipendente. Ogni insieme di n vettori linearmente indipendenti costituisce una base.

- Nel caso della dimensione finita, non e' difficile trovare una base, si prendono successivamente vettori linearmente indipendenti, fino a che non ve ne sono altri. In questo caso l'uso di una base per esprimere ogni vettore attraverso le sue componenti e' spesso molto comodo

- le componenti di un vettore dipendono ovviamente dalla base assegnata. Per questo e' piu' comodo adoperare i vettori astratti, piuttosto che le ennuple di componenti, anche se le due formulazioni sono compleatmente equivalenti. E' nei problemi concreti che tipicamente si introduce una base, e si eseguono i calcoli sulle componenti, cioe' sui loro valori numerici.

- Nel caso della dimensione infinita, e' spesso impossibile dare esplicitamente una base, che pertanto perde molta della sua utilita'. La maggior parte degli spazi di funzioni (in particolare quello che si adopera in MQ) sono di questo tipo. Attenzione a non confondere lo sviluppo di un vettore come combinazione lineare degli elementi di base, la somma deve essere finita anche nel caso di una base infinita, ovvero gli a_i q.t.n., con un'espressione del tipo
$$y = \sum_{i=1}^{\infty} a_i x_i$$

che acquista significato solo come limite della successione delle ridotte

$$y = \lim_{n\to\infty} \sum_{i=1}^{n} a_i x_i$$

e presuppone una nozione di convergenza, che esula dalle operazioni algebriche fin qui definite. Naturalmente queste considerazioni diventano molto importanti nel caso della dimensione infinita, dove si verificano fatti nuovi rispetto al caso finito dimensionale. Lo studio di questi fatti e' l'ambito dell'analisi funzionale, che e' argomento molto vasto. A questo livello ignoreremo tali fatti, trattando la dimensione infinita intuitivamente sullo stesso piano di quella finita, e accenneremo alle nuove proprieta' quando necessario.

- Osserviamo che abbiamo definito la base $X = (x_i)$ come una famiglia, anziche' un insieme di vettori. Spesso i due concetti possono confondersi, il fatto talora importante e' che gli elementi di base sono ordinati (dal loro indice). Ovviamente le componenti (a_i) dei vettori dipendono dall'ordinamento della base. Tale ordinamento definisce anche l'orientazione dello spazio, ove questa abbia importanza (ad esempio per definire la chiralita'). In R^3 questo porta alla regola della vite, o della mano destra. Assegnati i 3 vettori di base (x, y, z), questi definiscono un'orientazione (diciamo positiva), mentre se scambiamo ad es. x e y la base (y, x, z) definisce l'orientazione opposta.

Vediamo alcuni esempi (verificare!)

- In R^n o C^n una base e' la cosiddetta base canonica

  $$e_1 = (1, 0, 0, \dots, 0) \;\; e_2 = (0, 1, 0, \dots, 0) \;\; \cdots \;\; e_n = (0, 0, \dots, 0, 1)$$

  ma altre basi sono possibili, ad es. in R^2

  $$f_1 = (1, 1) \;\; f_2 = (1, -1)$$

  $$g_1 = (1, 0) \;\; g_2 = (1, 1)$$

  ecc. Ovvero ogni insieme di n vettori linearmente indipendenti costituisce una base

- la successione infinita di vettori

$$e_i = (0, 0, \ldots, 1_i, 0, \ldots, 0) \quad i = 1, \ldots, \infty$$

costituisce una base in in K^ω, ma non in K^∞

- Cosi' si osservi la differenza tra gli spazi P_n dei polinomi di grado massimo n, P spazio di tutti i polinomi (di qualunque ordine) e $F = \{f(x)\}$, insieme delle funzioni reali definite su $[-1.1]$ tali che ammettano uno sviluppo in serie uniformemente convergente

$$f(x) = \sum_{i=0}^{\infty} a_i x^i$$

Nel primo caso la dimensione dello spazio e' finita, uguale a $n+1$

$$\{1, x, x^2, \ldots, x^n\}$$

nel secondo e' infinita, e ha come elementi tutte le potenze x^n, nel terzo e' ancora infinita, ma molto piu' grande della precedente (e non se ne puo' dare esplicitamente una)

- Nello spazio $M^{m\times n}$ delle matrici ad elementi in K una base e' data dalle matrici $\{E_{ij}\}$

$$E_{ij} = \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 1_{ij} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}$$

che hanno tutti gli elementi nulli tranne che un 1 al posto ij. Da qui segue subito che la dimensione di $M^{m\times n}$ e' mn.

Nel caso delle matrici quadrate $n \times n$, un'altra base e' data dalle matrici simmetriche e antisimmetriche $\{A_{ij}, B_{ij}\}$

$$A_{ij} = E_{ij} + E_{ji} \quad B_{ij} = E_{ij} - E_{ji}$$

4 Prodotto scalare

Definiamo in V una nuova operazione, che associa ad ogni coppia di vettori un numero, detta prodotto scalare, o prodotto interno. La indicheremo con

$$\langle x, y \rangle$$

ed e' quindi un'applicazione $V \times V \to C$ (con R come sottocaso ovvio). Si richiede che il prodotto scalare soddisfi gli assiomi (indicheremo con z^* il complesso coniugato di z, come e' tradizione in MQ, a differenza della tradizione in matematica, che lo indica con $\bar{z}$)

1. $$\langle y, x \rangle = \langle x, y \rangle^*$$

2. $$\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$$
$$\langle x, ay \rangle = a \langle x, y \rangle$$

3. $$\langle x, x \rangle \geq 0 \qquad \langle x, x \rangle = 0 \Rightarrow x = 0$$

A parole: si possono scambiare gli argomenti nel prodotto scalare prendendo il complesso coniugato. Il prodotto scalare e' lineare nel membro di destra (anche qui osserviamo che la tradizione matematica lo definisce lineare nel membro di sinistra). Da (1) e (2) segue subito che a sinistra il prodotto scalare e' antilineare, cioe' un coefficiente moltiplicativo viene portato fuori dal prodotto come complesso coniugato

$$\langle ax, y \rangle = a^* \langle x, y \rangle$$

Infatti

$$\langle ax, y \rangle = \langle y, ax \rangle^* = (a \langle y, x \rangle)^* = a^* \langle y, x \rangle^* = a^* \langle x, y \rangle$$

In definitiva una combinazione lineare all'interno del prodotto scalare si puo' portare fuori come sta a destra, e coi coefficienti complessi coniugati a sinistra

$$\left\langle y, \sum_i a_i x_i \right\rangle = \sum_i a_i \, \langle y, x_i \rangle$$

$$\left\langle \sum_i a_i x_i, y \right\rangle = \sum_i a_i^* \langle x_i, y\rangle$$

Da ultimo il prodotto scalare di un vettore con se stesso (che per (1) e' sempre reale) e' sempre positivo o nullo, ed e' nullo se e solo se il vettore x e' nullo, ovvero l'annullarsi del prodotto scalare implica l'annullarsi del vettore stesso. Nel caso reale scompaiono i segni di complesso coniugato, e il prodotto scalare e' simmetrico nei suoi argomenti. Il prodotto scalare di un vettore con se stesso si interpreta in analogia col caso dei vettori geometrici come quadrato del modulo del vettore stesso.

Esempi

- In R^3 il prodotto scalare e' definito da

$$\langle x, y\rangle \equiv x \cdot y = |x||y| \cos\theta = x_1 y_1 + x_2 y_2 + x_3 y_3$$

- In generale in R^n si puo' generalizzare a

$$\langle x, y\rangle = \sum_{i=1}^{n} a_i b_i$$

  dove (a_i) e (b_i) sono le componenti di x e y. Analogamente in C^n

$$\langle x, y\rangle = \sum_{i=1}^{n} a_i^* b_i$$

  Queste sono le definizioni piu' comuni, osserviamo pero' che altre scelte sono possibili

- Nel caso delle funzioni continue su $[a, b]$,

$$F = \{f : [a, b] \to C\}$$

  e' utile la definizione

$$\langle f, g\rangle = \int_a^b f(x)^* g(x)\, dx$$

  La definizione comporta qualche sottigliezza, ma si estende anche a classi piu' generali di funzioni, purche' l'integrale

$$\langle f, f\rangle = \int f(x)^* f(x)\, dx = \int |f(x)|^2\, dx$$

  sia ben definito.

4.1 Norma

Definiamo ancora la norma di un vettore come

$$\|x\| = \sqrt{\langle x, x\rangle}$$

che interpretiamo come modulo del vettore. Dalla definizione segue subito che e'

1.
$$\|x\| \geq 0 \qquad \|x\| = 0 \Leftrightarrow x = 0$$

2.
$$\|ax\| = |a|\|x\|$$

3.
$$\|x + y\| \leq \|x\| + \|y\|$$

La terza proprieta' detta disuguaglianza triangolare (generalizza la nota proprieta' dei triangoli per cui un lato non puo mai essere maggiore della somma degli altri due) non e' ovvia, e verra' dimostrata piu' avanti.

Ancora un vettore si dira' normalizzato se la sua norma e' uguale a uno

$$\|x\| = 1$$

in generale, dato un vettore x qualsiasi, e' sempre possibile normalizzarlo, cioe' costruire un suo multiplo con norma uno, basta a tale scopo dividerlo per la sua norma

$$x \to x' = \frac{x}{\|x\|} \qquad \|x'\| = 1$$

Osserviamo pero' che, dato x, il vettore normalizzato corrispondente non e' unico. Infatti, moltiplicandolo per un coefficiente a di modulo uno, $|a| = 1$, resta normalizzato. Un coefficiente di modulo uno giace sul cerchio di raggio uno nel piano complesso, e si puo mettere sempre nella forma

$$|a| = 1 \quad \Rightarrow \quad a = e^{i\phi} \quad \phi \in R$$

ed e' detto un fattore di fase, con la fase ϕ reale. In R l'ambiguita' si riduce ad un segno, oviamente i vettori x e $-x$ hanno lo stesso modulo.

4.2 Ortogonalita'

- Due vettori si dicono ortogonali se il loro prodotto scalare e' nullo

$$\langle x, y \rangle = 0$$

e si scrive anche $x \perp y$. La definizione si estende naturalmente a sottoinsiemi S di elementi di V, Diremo che x e ortogonale a S, $x \perp S$, se e' ortogonale a tutti gli elementi di S, cioe'

$$\forall y \in S \quad \Rightarrow \quad \langle x, y \rangle = 0$$

e ancora $S \perp T$ se per ogni $y \in S, z \in T$ e'

$$\langle y, z \rangle = 0$$

Ancora, data una famiglia di vettori (x_i) diremo che sono ortonormali (o.n.), se sono tutti normalizzati e a coppie ortogonali, cioe'

$$\langle x_i, x_j \rangle = \delta_{ij}$$

dove il delta di Kronecker δ_{ij} vale 1 se $i = j$ e vale 0 se $i \neq j$. Questo e' un simbolo molto usato. Se compare in una somma su uno dei suoi indici, il risultato e' di ridurre la somma al solo termine con indice uguale al secondo indice

$$\sum_i a_i x_i \delta_{ij} = a_j x_j$$

- Si vede subito che vettori o.n. sono linearmente indipendenti

$$\sum_i a_i x_i = 0 \quad \Rightarrow \quad a_i = \langle x_i, 0 \rangle = 0$$

Diremo anche proiezione di y su x_i o.n. il vettore

$$a_i x_i = \langle x_i, y \rangle \, x_i$$

che corrisponde geometricamente alla proiezione del vettore y sull'asse definito dal versore x.

- Il caso di una base ortonormale X e' molto importante. Ad esempio e' facile ottenere i coefficienti di sviluppo di un particolare vettore in tale base. Da
$$y = \sum_i a_i x_i$$
prendendo il prodotto scalare di entrambi i membri con x_j a sinistra e'
$$\langle x_j, y\rangle = \langle x_i, \sum_i a_i x_i\rangle = \sum_i a_i \, \langle x_j, x_i\rangle = \sum_k a_i \delta_{ji} = a_j$$
Quindi i coefficienti a_i dello sviluppo (detti anche coefficienti di Fourier) si ottengono dai prodotti scalari
$$a_i = \langle x_i, y\rangle$$

- Data una famiglia (x_i) o.n. (non necessariamente una base) e un vettore qualsiasi y, il vettore
$$y' = y - \sum_i \langle x_i, y\rangle x_i$$
e' ortogonale a tutti gli x_i. Infatti
$$\langle x_j, y'\rangle = \langle x_j, y\rangle - \sum_i \langle x_i, y\rangle \langle x_j, x_i\rangle = 0$$
Il significato geometrico e' evidente. La proiezione di un vettore y su un vettore normalizzato x_i non e' altro che la proiezione geometrica ortogonale di y sull'asse x_i. Togliendo via via queste proiezioni quello che rimane e' un vettore ortogonale a tutti i vettori (assi) x_i.

- Questo permette di affermare la seguente importante proprieta': data una qualunque famiglia $Y = (y_k)$ linearmente indipendente, esiste sempre una famiglia o.n. $X = (x_i)$ ad essa equivalente, tale cioe' che X e Y generino lo stesso sottospazio, ovvero gli x_i si possano scrivere come combinazione lineare degli y_i e viceversa. Questo si puo' vedere attraverso una costruzione esplicita degli x_i, che e' nota come ortogonalizzazione di Schmidt. Consiste nei seguenti passi.

  1. si normalizza il primo vettore
  $$x_1 = \frac{y_1}{\|y_1\|}$$

2. si ortogonalizza y_2 a x_1, e lo si normalizza

$$x'_2 = y_2 - \langle x_1, y_2 \rangle x_1$$

$$x_2 = \frac{x'_2}{\|x'_2\|}$$

3. al passo $(n+1)$-esimo si ortogonalizza il vettore y_{n+1} agli x_i precedentemente ottenuti, e lo si normalizza

$$x'_{n+1} = y_{n+1} - \sum_{i=1}^{n} \langle x_i, y_{n+1} \rangle x_i$$

$$x_{n+1} = \frac{x'_{n+1}}{\|x'_{n+1}\|}$$

Osserviamo che ad ogni passo il vettore ortogonalizzato x'_n non e' mai nullo, perche' altrimenti vorrebbe dire che y_n non e' linearmente indipendente dai precedenti.

- Con lo stesso ragionamento possiamo ottenere l'importante disuguaglianza di Schwarz

$$|\langle x, y \rangle| \leq \|x\| \|y\|$$

Costruiamo il vettore z ortogonalizzando y a x, $z \perp x$

$$z = y - \langle x, y \rangle \frac{x}{\|x\|^2} = y - ax$$

E'

$$0 \leq \langle z, z \rangle = \langle y - ax, z \rangle = \langle y, z \rangle = \langle y, y \rangle - a \langle y, x \rangle$$

$$0 \leq \|y\|^2 - |\langle x, y \rangle|^2 / \|x\|^2$$

da cui la tesi. Osserviamo ancora che il segno di uguaglianza vale se e solo se $z = 0$, ovvero $y = ax$, i due vettori solo paralleli.

- Da qui anche la disugualianza triangolare per la norma

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle =$$

$$\|x\|^2 + \|y\|^2 + 2Re\langle x, y \rangle \leq \|x\|^2 + \|y\|^2 + 2|\langle x, y \rangle| \leq \|x\|^2 + \|y\|^2 + \|x\| \|y\| = (\|x\| + \|y\|)^2$$

da cui la tesi

- In generale data la base X si definisce matrice di sovrapposizione S la matrice dei prodotti scalari tra gli elementi di base

  $$S_{ij} = \langle x_i, x_j \rangle$$

  che nel caso di una base o.n. si riduce alla matrice identica I

  $$S_{ij} = \delta_{ij} \qquad S = I$$

  Attraverso la matrice dei prodotti scalari si puo' esprimere il prodotto scalare tra una coppia qualunque di vettori. Dati

  $$y = \sum_i c_i x_i \qquad z = \sum_i d_i x_i$$

  e'

  $$\langle y, z \rangle = \langle \sum_i c_i x_i, \sum_j d_j x_j \rangle = \sum_{ij} c_i^* S_{ij} d_j$$

  Nel caso della base o.n. questo si riduce a

  $$\langle y, z \rangle = \sum_i c_i^* d_i$$

  che rappresentano le espressioni dei prodotti scalari attraverso le componenti dei vettori dati.

5 Trasformazioni lineari

Dati due spazi lineari V e W, diremo che un'applicazione $f : V \to W$ e' lineare se

$$f(x + y) = f(x) + f(y)$$

$$f(ax) = af(x)$$

Chiamiamo $L(V, W)$ l'insieme di tutte le applicazioni lineari di V in W

$$L(V, W) = \{f : V \to W,\ f \text{ lineare}\}$$

Poiche' date f e g lineari, $f + g$ e af sono ancora lineari, anche $L(V, W)$ e' uno spazio lineare. In particolare saremo interessati al caso di applicazioni di V in se stesso, chiameremo le applicazioni lineari $A : V \to V$ operatori, e scriveremo semplicemente

$$Ax = y$$

invece di $A(x) = y$. Cosi' l'insieme di tutti gli operatori lineari su V, $L(V, V)$ si indichera' semplicemente come $L(V)$

$$L(V) = \{A : V \to V,\ A \text{ lineare}\}$$

che e' ancora uno spazio lineare.

E' importante anche il sottocaso $W = K$, cioe' delle applicazioni lineari che fanno corrispondere un vettore a un numero. Queste si dicono funzionali (lineari).

6 Operatori lineari

Consideriamo piu' in dettaglio gli operatori lineari

$$A(x + y) = Ax + Ay \qquad A(ax) = aAx$$

e quindi in generale, l'operatore A attraversa ogni combinazione lineare

$$A \sum_i c_1 x_i = \sum_i c_1 A x_i$$

Indicheremo generalmente gli operatori lineari con lettere maiuscole. Nel caso di ambiguita' (raramente) useremo un accento circonflesso per indicare un operatore, ad es. $\hat{A}$

Come gia' detto l'insieme di tutti gli operatori lineari su V

$$L(V) = \{A\}$$

e' uno spazio vettoriale, sono ben definiti la somma di due operatori e il prodotto di un numero per un operatore

$$(A + B)x = Ax + Bx \qquad (aA)x = aAx$$

definiamo anche l'operatore 0 e l'operatore identita' 1 o I

$$\hat{0}x = 0 \qquad \hat{1}x \equiv Ix = x$$

col che valgono tutte le regole di calcolo per gli spazi vettoriali.

In $L(V)$ e' pero' possibile definire ancora un'altra operazione, il prodotto di due operatori, $AB = C$, che e' ancora un operatore su V, definito come la composizione delle due applicazioni

$$(AB)x = A(Bx)$$

cioe' prima B agisce su x, dando un vettore $y = Bx$, e poi A agisce su y, dando il vettore finale. E' facile vedere che il prodotto e' associativo

$$(AB)C = A(BC)$$

infatti

$$[(AB)C]x = (AB)(Cx) = A(B(Cx)) = A(BCx) = [A(BC)]x$$

Per gli operatori lineari il prodotto e' anche bilineare, cioe'

$$A(B + C) = AB + AC \qquad (A + B)C = AC + BC$$

$$A(aB) = a(AB) = (aA)B$$

come si verifica facilmente operando su un vettore x arbitrario.

Per gli operatori non vale pero' la proprieta' commutativa del prodotto, in generale

$$AB \neq BA$$

Vediamo qualche esempio. Definiamo A e B attraverso l'azione su due vettori dello spazio come segue

$$Ax_1 = x_2 \qquad Ax_2 = 0$$

$$Bx_1 = 0 \qquad Bx_2 = x_1$$

allora e'

$$ABx_1 = 0 \qquad ABx_2 = x_2$$

$$BAx_1 = x_1 \qquad BAx_2 = 0$$

Come secondo esempio consideriamo in R^2 i due operatori corrispondenti a una rotazione antioraria di 90° e alla riflessione rispetto alla retta $y = x$ (diagonale del primo quadrante). Siano x_1 e x_2 i versori dei due assi coordinati, allora

$$Ax_1 = x_2 \qquad Ax_2 = -x_1$$

$$Bx_1 = x_2 \qquad Bx_2 = x_1$$

e

$$ABx_1 = -x_1 \qquad ABx_2 = x_2$$

$$BAx_1 = x_1 \qquad BAx_2 = -x_2$$

che di nuovo non commutano.

Quindi in definitiva con gli operatori valgono tutte le regole dell'algebra ordinaria, ad eccezione della proprieta' commutativa. Cosi' nello svolgimento di espressioni algebriche occorera' fare empre attenzione a conservare l'ordine dei fattori, ad esempio

$$(aA + bB)^2 = (aA + bB)(aA + bB) = a^2A^2 + ab(AB + BA) + b^2B^2$$

dove

$$(AB + BA) \neq 2AB$$

Un insieme, come $L(V)$, che sia spazio vettoriale e in cui sia anche definita un'operazione di prodotto bilineare si dice un'algebra, per questo $L(V)$ si dice l'algebra degli operatori su V.

Ancora, dato un operatore A, si definisce suo inverso A^{-1} un operatore tale che

$$AA^{-1} = A^{-1}A = I$$

Per un A generico, non e' detto che esista l'inverso, che richiede che A come applicazione sia biunivoca. In tal caso si dice che A e' invertibile. Osserviamo che in generale entrambe le condizioni $AA^{-1} = I$ e $A^{-1}A = I$ sono richieste per l'esitenza dell'inverso. Negli spazi a dimensione finita ne basta una (ciascuna implica l'altra), ma questo non e' piu' vero nel caso della dimensione infinita.

Se A e B sono invertibili, anche il prodotto lo e' ed e'

$$(AB)^{-1} = B^{-1}A^{-1}$$

infatti

$$AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = I$$

e analoga dall'altra parte. Quindi occorre prendere il prodotto degli inversi in ordine inverso.

6.1 Commutatori

La non commutativita' di due operatori generici suggerisce di definire il commutatore di due operatori come

$$[A, B] = AB - BA$$

che gioca un ruolo molto importante. Quindi dire che il commutatore di due operatori e' uguale a zero equivale a dire che i due operatori commutano. Il commutatore di due operatori e' ancora un operatore di $L(V)$, e puo' essere visto come una nuova operazione che associa a due operatori A e B il loro commutatore (detto anche prodotto di Lie). Il commutatore soddisfa le proprieta'

1. antisimmetria
$$[A, B] = -[B, A]$$

2. bilinearita'
$$[A, B + C] = [A, B] + [A, C]$$
$$[A + B, C] = [A, C] + [B, C]$$
$$[aA, B] = a[A, B] = [A, aB]$$

3. commutatore di un prodotto
$$[A, BC] = B[A, C] + [A, B]C$$
$$[AB, C] = A[B, C] + [A, C]B$$
quindi il fattore a sinistra si porta fuori a sinistra, e quello a destra fuori a destra

4. identita' di Jacobi
$$[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0$$
(si osservi la permutazione circolare dei 3 operatori) ovvero
$$[A, [B, C]] - [[A, B], C] = [B, [A, C]]$$
che mostra come per il commutatore come prodotto non valga la proprieta' associativa, sostituita dall'identita' di Jacobi

Tutte queste proprieta' si dimostrano molto semplicemente sviluppando i commutatori secondo la definizione.

6.2 Matrici di operatori

Fissata una base $X = (x_i)$ in V, un generico operatore A e' completamente determinato noto il suo effetto sui soli elementi della base. Sia $\bar{x}_i = Ax_i$ il trasformato del vettore di base x_i. Essendo un elemento di V puo' essere sviluppato nella base data

$$Ax_i = \bar{x}_i = \sum_j c_j x_j$$

Per tener conto del particolare elemento x_i trasformato, i coefficienti di sviluppo devono avere anche un secondo indice i, scriviamo $c_j = A_{ji}$ e in definitiva

$$Ax_i = \sum_j x_j A_{ji}$$

equazione che definisce la matrice $A = A_{ji}$ associata all'operatore A, che indichiamo con la stessa lettera. In generale e' chiaro dal contesto di chi si tratta. Osserviamo anche che nel prodotto di un vettore per un numero, essendo entita' diverse, e' irrilevante scrivere il numero davanti o dopo il vettore, cx e xc sono la stessa cosa. Scriviamo $x_j A_{ji}$ per conservare la regola del prodotto righe per colonne tipico delle matrici, ovvero le somme sugli indici interni contigui, che e' piu' facile da ricordare.

Se e' data la matrice di un operatore, l'azione su un vettore qualunque si ottiene per linearita'. Da

$$y = \sum_i c_i x_i$$

e'

$$Ay = \sum_i c_i Ax_i = \sum_i c_i \sum_j x_j A_{ji} = \sum_j (\sum_i A_{ji} c_i) x_j = \sum_j \bar{c}_j x_j$$

con

$$\bar{c}_j = \sum_i A_{ji} c_i$$

Quindi alla trasformazione

$$Ay = \bar{y}$$

corrisponde la moltiplicazione del vettore colonna c per la matrice A

$$Ac = \bar{c}$$

Osserviamo quindi come l'introduzione di una base associ in modo univoco ad ogni vettore y di V un vettore colonna c in C^n, e ad ogni operatore A la sua matrice A di $C^{n\times n}$. Tale corrispondenza conserva le operazioni algebriche tra vettori, tra operatori, e tra operatori e vettori. Si dice per questo un isomorfismo tra gli spazi relativi, ogni relazione vera tra vettori e operatori astratti corrisponde alla stessa relazione tra componenti, e viceversa. E' quindi solo questione di convenienza (nel caso della dimensione finita) adoperare vettori e operatori astratti, oppure componenti. In generale come gia' detto i primi rendono piu' compatta la formulazione teorica (che oltretutto diventa indipendente dalla base scelta), ma i problemi concreti sono spesso formulati e risolti in termini di componenti.

Vediamo ad esempio le matrici associate agli operatori gia' visti

$$Ax_1 = x_2 = x_1 A_{11} + x_2 A_{21} \quad \Rightarrow \quad A_{11} = 0, \ \ A_{21} = 1$$

$$Ax_2 = 0 = x_1 A_{12} + x_2 A_{22} \quad \Rightarrow \quad A_{12} = 0, \ \ A_{22} = 0$$

da cui

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

e cosi' per

$$Ax_1 = x_2 \qquad Ax_2 = -x_1$$

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

Come si vede, le colonne della matrice sono i coefficienti dei vettori trasformati Ax_i.

Nel caso importante in cui la base X e' o.n., anche gli elementi di matrice di un operatore si ottengono semplicemente attraverso i prodotti scalari

$$\langle x_i, Ax_j\rangle = \langle x_i, \sum_k x_k A_{kj}\rangle = \sum_k \delta_{ik} A_{kj} = A_{ij}$$

quindi

$$A_{ij} = \langle x_i, Ax_j\rangle$$

Per questo in generale, anche nel caso di una coppia qualsiasi di vettori x e y si definisce elemento di matrice di A relativo a questi vettori il prodotto scalare

$$\langle x, Ay\rangle$$

In particolare, dato un vettore $x \neq 0$ definiamo valor medio di un operatore il numero

$$\langle A \rangle = \frac{\langle x, Ax \rangle}{\langle x, x \rangle}$$

che nel caso di x normalizzato si scrive semplicemente

$$\langle A \rangle = \langle x, Ax \rangle$$

Osserviamo che il valor medio $\langle A \rangle$ dipende sia dall'operatore A che dal particolare vettore x considerato. Normalmente quest'ultimo e' sottinteso, se si vuole mettere in evidenza, lo si puo' porre a pedice

$$\langle A \rangle_x$$

Osserviamo che anche nel caso di una base non o.n. si definisce una matrice dei prodotti scalari

$$A^S_{ij} = \langle x_i, Ax_j \rangle$$

che in questo caso pero' non coincide con la matrice dell'operatore definita dalla sua azione sulla base. Si vede facilmente che la relazione e'

$$A^S = SA$$

dove S e' la matrice di sovrapposizione della base data. Se la base e' o.n. e' $S = I$ e si ha l'eguaglianza $A^S = A$.

Normalmente anche la matrice dei prodotti scalari si indica con A, anche nel caso non o.n., il che richiede attenzione.

6.3 Proprieta' legate al prodotto scalare

Il prodotto scalare permette di associare ad ogni operatore A su V un nuovo operatore, detto aggiunto (o hermitiano coniugato) di A, che si indica con A^+, (si legga A croce) definito dalla proprieta'

$$\langle x, A^+ y \rangle = \langle Ax, y \rangle$$

per qualunque coppia di vettori x e y. Anche se la definizione sembra astratta, definisce completamente l'operatore A^+, ad esempio attraverso la sua matrice, come vedremo.

Dalla definizione segue immediatamente che e' anche

$$\langle x, Ay \rangle = \langle A^+ x, y \rangle$$

infatti

$$\langle x, Ay \rangle = \langle Ay, x \rangle^* = \langle y, A^+ x \rangle^* = \langle A^+ x, y \rangle$$

Quindi la regola generale e' che si puo' spostare un operatore A da una parte all'altra del prodotto scalare prendendone l'aggiunto.

Valgono le proprieta'

- A^+ e' lineare
- $$(A^+)^+ = A$$

  Come si vede l'operazione di aggiunto per gli operatori e' analoga alla coniugazione di un numero complesso
- $$(A + B)^+ = A^+ + B^+$$
- $$(aA)^+ = a^* A^+$$
- $$(AB)^+ = B^+ A^+$$

  l'aggiunto di un prodotto si ottiene prendendo il prodotto degli aggiunti in ordine inverso
- $$\hat{0}^+ = \hat{0} \qquad I^+ = I$$
- $$(A^{-1})^+ = (A^+)^{-1}$$

che si possono dimostrare prendendone il prodotto scalare tra una coppia di vettori qualsiasi. Ad esempio

$$\langle x, ABy\rangle = \langle A^+x, By\rangle = \langle B^+A^+x, y\rangle = \langle (AB)^+x, y\rangle$$

e per l'ultima

$$(AA^{-1})^+ = I^+ = I = (A^{-1})^+A^+$$
$$(A^{-1}A)^+ = I^+ = I = A^+(A^{-1})^+$$

mostrano che $(A^{-1})^+$ e' l'inverso di A^+.

In generale data una matrice A_{ij} si definisce

- Matrice coniugata
$$(A^*)_{ij} = A^*_{ij}$$
ottenuta coniugando tutti gli elementi della matrice A

- Matrice trasposta A^T
$$A^T_{ij} = A_{ji}$$
cioe' la matrice trasposta si ottiene ribaltando gli elementi della matrice data attorno alla diagonale principale

- Matrice coniugata hermitiana A^+
$$A^+_{ij} = (A_{ji})^* \qquad A^+ = (A^T)^*$$
cioe' la matrice ottenuta trasponendo la matrice data e coniugando i suoi elementi

Sia X una base e consideriamo la matrice dell'operatore definita dai prodotti scalari

$$A_{ij} = \langle x_i, Ax_j\rangle$$

Per la matrice dell'operatore aggiunto e' allora

$$(A^+)_{ij} = \langle x_i, A^+x_j\rangle = \langle Ax_i, x_j\rangle = \langle x_j, Ax_i\rangle^* = A^*_{ji}$$

Quindi la matrice dei prodotti scalari di A^+ coincide con la matrice coniugata hermitiana della matrice A dell'operatore di partenza. In una base o.n. la matrice dei prodotti scalari coincide con la matrice dell'operatore (definita dalla trasformazione dei vettori di base), quindi (solo in tal caso) anche per essa la matrice dell'aggiunto coincide con la coniugata hermitiana della matrice data.

Nel caso reale la matrice aggiunta coincide semplicemente con la trasposta.

6.3.1 Operatori hermitiani e unitari

Un operatore che coincida con il proprio aggiunto si dice hermitiano

$$A^+ = A$$

Gli operatori hermitiani sono una categoria molto importante. Le loro proprieta' assomigliano a quelle dei numeri reali (che coincidono con il loro coniugato) rispetto ai numeri complessi. Siano $A = A^+$ e $B = B^+$ hermitiani, allora e'

- $$(A + B)^+ = A + B$$

  La somma di due operatori hermitiani e' ancora hermitiana

- $$(aA)^+ = a^* A$$

  L'operatore (aA) e' hermitiano se e solo se a e' reale, $a^* = a$

- $$(AB)^+ = BA$$

  Il prodotto di due operatori hermitiani e' hermitiano se e solo se i due operatori commutano

- Ancora, per A qualunque, e'

  $$(A + A^+)^+ = (A^+ + A) = A + A^+$$

  $$(A^+ A)^+ = A^+ (A^+)^+ = A^+ A$$

  cioe' gli operatori $(A + A^+)$ e $A^+ A$, AA^+ sono sempre hermitiani

Il valor medio di un operatore hermitiano e' sempre reale. Sia x normalizzato, allora

$$\langle A \rangle^* = \langle x, Ax \rangle^* = \langle Ax, x \rangle = \langle x, Ax \rangle = \langle A \rangle$$

Un operatore hermitiano si dice semidefinito positivo se il suo valore medio e' non negativo per qualunque vettore x

$$\langle A \rangle \geq 0$$

Questo e' sempre vero per un operatore della forma A^+A, qualunque sia A

$$\langle A^+A\rangle = \langle x, A^+Ax\rangle = \langle Ax, Ax\rangle = \langle y, y\rangle \geq 0$$

dove si e' posto $Ax = y$

Si definisce operatore unitario un operatore che ammette inverso, e il cui inverso coincide con l'aggiunto

$$U^+ = U^{-1}$$

La proprieta' piu importante degli operatori unitari e' che conservano il prodotto scalare. Dati i trasformati

$$\bar{x} = Ux \qquad \bar{y} = Uy$$

e'

$$\langle \bar{x}, \bar{y}\rangle \ = \langle Ux, Uy\rangle = \langle x, U^+Uy\rangle = \langle x, U^{-1}Ux\rangle = \langle x, y\rangle$$

ovvero i trasformati hanno lo stesso prodotto scalare dei vettori di partenza. In particolare un operatore unitario trasforma una base o.n. in un'altra base o.n.

Il prodotto di due operatori unitari T e U e' ancora unitario

$$(TU)^+ = U^+T^+ = U^{-1}T^{-1} = (TU)^{-1}$$

Le matrici di un operatore hermitiano soddisfano

$$(A^+)_{ij} = A_{ij} = A^*_{ji}$$

ovvero coincidono con la loro coniugata hermitiana, come per l'operatore corrispondente. Matrici che soddisfano questa proprieta' si dicono hermitiane

Le matrici corrispondenti agli operatori unitari soddisfano invece

$$(U^+)_{ij} = U^*_{ji} = U^{-1}_{ij}$$

cioe' il loro inverso coincide con il coniugato hermitiano. Questa proprieta' si scrive anche

$$UU^+ = I \qquad U^+U = I$$

da cui, prendendo le matrici corrispondenti

$$\sum_j U_{ij}U^*_{kj} = \delta_{ik}$$

e

$$\sum_j U^*_{ji} U_{jk} = \delta_{ik}$$

cioe' righe e colonne di una matrice unitaria sono ortonormali.

Nel caso reale le matrici hermitiane si riducono alle matrici simmetriche

$$A^T = A \qquad A_{ij} = A_{ji}$$

e quelle unitarie a

$$O^T = O^{-1}$$

che si dicono ortogonali

7 Notazioni di Dirac

In MQ e' molto comune una notazione introdotta da Dirac, che rappresenta i vettori con il simbolo $| \ \rangle$, detto ket (e' la seconda meta' della parola bra(c)ket, cioe' parentesi, mezza parentesi). Dentro alla mezza parentesi si mette tutto quello che serve a identificare il vettore, ad esempio

$$x_i \equiv |x_i\rangle \equiv |i\rangle$$

che e' comoda ad esempio quando ci siano diversi indici. In realta' la notazione di Dirac fa uso anche dell'altra meta' della parentesi, che definisce il cosiddetto vettore bra $\langle \ |$. Il prodotto scalare si forma mettendo assieme un vettore bra e uno ket

$$\langle x|y\rangle$$

mentre un elemento di matrice assume la forma simmetrica

$$\langle x|A|y\rangle$$

Noi non adopereremo i vettori bra, che di fatto non sono necessari. In algebra lineare sono gli elementi di uno spazio associato a quello dato, detto spazio duale, che non tratteremo.

Vi e' pero' un'estensione della notazione che si rivela molto pratica, e che data una coppia di vettori x e y definisce un operatore lineare, che si indica con

$$|x\rangle\langle y|$$

e che e' definito dalla sua azione su un generico vettore z da

$$|x\rangle\langle y|\ z = |x\rangle\langle y|z\rangle = \langle y, z\rangle\ x$$

cioe' fornisce sempre un multiplo del vettore x, moltiplicato per il prodotto scalare di y e z. La linearita' segue subito da quella del prodotto scalare, e' facile anche verificare che l'aggiunto si ottiene scambiando di posto i due vettori

$$(|x\rangle\langle y|)^{+} = |y\rangle\langle x|$$

basta prenderne il prodotto scalare tra una coppia di vettori generici z e w. Ne segue che in particolare che l'operatore

$$|x\rangle\langle x|$$

e' hermitiano, e se x e' normalizzato, $\langle x, x\rangle = 1$, e' anche idempotente, cioe' coincide col suo quadrato

$$|x\rangle\langle x||x\rangle\langle x| = |x\rangle\langle x, x\rangle\langle x| = |x\rangle\langle x|$$

8 Equazione ad autovalori

1. Dato un operatore A e un vettore x non nullo tale che

$$Ax = ax$$

si dice che a e' autovalore dell'operatore A, e x un autovettore di A relativo all'autovalore a. Quindi per un autovettore x l'azione di A si riduce a moltiplicarlo per il numero a, ne cambia il modulo ma non la direzione.

L'equazione

$$Ax = ax$$

dove a e x sono incognite si dice equazione ad autovalori per A.

2. L'insieme

$$\sigma(A) = \{a\}$$

di tutti gli autovalori di A si dice spettro dell'operatore A.

3. Dati x_1 e x_2 autovettori di A relativi ad uno stesso autovalore a,

$$Ax_1 = ax_1 \qquad Ax_2 = ax_2$$

allora anche ogni loro combinazione lineare e' ancora autovettore relativo ad a

$$A(c_1x_1 + c_2x_2) = c_1Ax_1 + c_2Ax_2 = a(c_1x_1 + c_2x_2)$$

Questo mostra che l'insieme di tutti gli autovettori relativi ad un dato autovalore a e' un sottospazio vettoriale dello spazio V, che diremo autospazio relativo all'autovalore a, V_a

$$V_a = \{x : Ax = ax\}$$

La dimensione di V_a, cioe' il numero massimo di autovettori linearmente indipendenti relativi ad a, si dice grado di degenerazione dell'autovalore a. In particolare se $dimV_a = 1$ l'autovalore si dice non degenere, se e' uguale a due, tre, etc, si dice doppiamente, triplamente degenere, e cosi' via. Se il grado di degenerazione e' n, possiamo sempre scegliere una base o.n. di autovettori

$$Ax_i = ax_i \qquad \langle x_i, x_j\rangle = \delta_{ij}, \quad i,j = 1,\ldots,n$$

perche' una qualunque base in V_a puo' sempre ortogonalizzarsi.

4. Per ogni operatore, il valor medio dell'operatore relativo a un autovettore (che supponiamo normalizzato)

$$\langle A\rangle = \langle x, Ax\rangle = \langle x, ax\rangle = a$$

coincide con l'autovalore corrispondente

5. Autovettori relativi ad autovalori diversi sono sempre linearmente indipendenti. Dati

$$Ax_i = a_ix_i$$

con a_i tutti diversi, supponiamo

$$\sum_{i=1}^{n} c_ix_i = 0$$

E' allora

$$(A - a_1)\sum_{i=1}^{n} c_i x_i = \sum_{i=2}^{n} c_i(a_i - a_1)x_i = 0$$

si e' cioe' annullato il contributo del primo termine, e procedendo con $(A - a_2)$ ecc. si arriva fino a

$$c_n(a_n - a_1)\cdots(a_n - a_{n-1})x_n = 0$$

da cui $c_n = 0$, e procedendo a ritroso, tutti i $c_i = 0$.

6. Consideriamo il caso importante in cui a possiede una base di autovettori (x_i)

   $$Ax_i = a_i x_i$$

   Per ogni operatore che ammetta una base di autovettori, in tale base la matrice dell'operatore e' diagonale, con gli autovalori sulla diagonale

   $$Ax_j = a_j x_j = \sum_i x_i A_{ij}$$

   da cui

   $$A_{ij} = a_i \delta_{ij}$$

   $$A = \mathcal{A} = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & a_n \end{pmatrix}$$

   e viceversa ogni operatore che abbia la matrice diagonale in una base data, ammette tale base come base di autovettori, e gli elementi diagonali sono gli autovalori corrispondenti.

   Quindi un operatore che ammetta una base di autovettori e' completamente definito dai suoi autovalori e autovettori $\{a_i, x_i\}$ e per ogni vettore dello spazio e'

   $$y = \sum_i c_i x_i \qquad Ay = \sum_i a_i c_i x_i$$

   Per questo il problema di trovare una base di autovettori equivale al problema di trovare una base in cui la matrice dell'operatore sia diagonale, e si dice problema di diagonalizzazione dell'operatore (o della matrice corrispondente in una base arbitraria)

7. Consideriamo il caso importante in cui si abbia degenerazione. Ovvero supponiamo data una base di autovettori $(x_i)_{i=1,n}$

$$Ax_i = a_i x_i$$

dove l'indice i si riferisce agli autovettori, ma gli a_i coincidano in parte, ad esempio $a_1 = a_2 = \ldots = a_{n_1}$ etc. Se vogliamo indicizzare gli autovalori distinti con un indice i, avremo allora bisogno di un secondo indice per contare gli autovettori x_{ij} indipendenti relativi allo stesso a_i, con $j = 1, \ldots, n_i$, dove n_i e' il grado di degenerazione dell'autovalore a_i

$$Ax_{ij} = a_i x_{ij} \qquad j = 1, n_i$$

La matrice di A sara' adesso

$$\mathcal{A} = \begin{pmatrix}
a_1 & 0 & \cdots & 0 & & & & & & & & & & 0 \\
0 & a_1 & \cdots & 0 & & & & & & & & & & 0 \\
0 & 0 & \ddots & 0 & & & & & & & & & & 0 \\
0 & 0 & \cdots & a_1 & \cdots & & & & & & & & & 0 \\
0 & & \cdots & & a_2 & 0 & \cdots & 0 & & & & & & 0 \\
0 & & \cdots & & 0 & a_2 & \cdots & 0 & & & & & & 0 \\
0 & & \cdots & & 0 & 0 & \ddots & 0 & & & & & & 0 \\
0 & & \cdots & & 0 & 0 & \cdots & a_2 & & & & & & 0 \\
0 & \cdots & \cdots & & \cdots & \cdots & & & \ddots & & & & & 0 \\
0 & \cdots & \cdots & & \cdots & \cdots & & & & \ddots & & & & 0 \\
0 & \cdots & \cdots & & \cdots & \cdots & & & & & a_n & 0 & \cdots & 0 \\
0 & \cdots & \cdots & & \cdots & \cdots & & & & & 0 & a_n & \cdots & 0 \\
0 & \cdots & \cdots & & \cdots & \cdots & & & & & 0 & 0 & \ddots & 0 \\
0 & \cdots & \cdots & & \cdots & \cdots & & & & & 0 & 0 & & a_n
\end{pmatrix}$$

8.1 Il problema spettrale

Dato un generico operatore A, il problema di determinare tutti i suoi autovalori e autovettori e' detto problema spettrale. E' un problema piuttosto complicato, che e' completamente risolto nel caso di spazi a dimensione finita. Noi pero' ci limiteremo a un sottocaso semplice, anche se di estrema importanza. Inoltre nel caso della dimensione infinita compaiono fatti nuovi, di cui il piu' cospicuo e' l'esistenza dello spettro continuo, cioe' della possibilita' che gli autovalori coprano ad esempio tutto un intervallo reale. La stessa definizione di spettro deve venir opportunamente generalizzata.

In generale un operatore puo' anche non avere nessun autovalore. Ad esempio l'operatore di rotazione di 90° in R^2 evidentemente cambia la direzione (ruota) di tutti i vettori: non ha alcun autovettore. Osserviamo che la situazione e' diversa in R^3, adesso ogni vettore sull'asse di rotazione resta inalterato, ed e' quindi autovettore relativo all'autovalore 1.

Per uno spazio vettoriale di dimensione finita V su C, $dimV = n$, valgono le seguenti proprieta'

1. Ogni operatore A ammette almeno un autovalore (e relativo autovettore)

2. Ogni operatore che commuti col proprio aggiunto

$$[A, A^+] = 0$$

   ammette una base di autovettori o.n. Questi operatori sono detti operatori normali

3. In particolare questo vale per gli operatori hermitiani e unitari. Dunque ogni operatore hermitiano ammette sempre una base di autovettori o.n.

$$Ax_i = a_i x_i \qquad \langle x_i, x_j \rangle = \delta_{ij}, \quad i, j = 1, \ldots, n$$

   e per ogni vettore dello spazio e'

$$y = \sum_i c_i x_i \qquad c_i = \langle x_i, y \rangle$$

4. Vediamo in particolare per gli operatori hermitiani

   (a) tutti gli autovalori di un operatore hermitiano sono reali, $a^* = a$ (come gia' visto per i valori medi)

$$a_i = \langle x_i, Ax_i \rangle = \langle Ax_i, x_i \rangle = \langle x_i, Ax_i \rangle^* = a_i^*$$

   (b) autovettori di un operatore hermitiano relativi ad autovalori diversi sono ortogonali. Siano

$$Ax = ax \qquad Ay = by \qquad a \neq b$$

e consideriamo

$$\langle x, Ay\rangle = \langle x, by\rangle = b\langle x, y\rangle =$$

$$\langle Ax, x\rangle = \langle ax, y\rangle = a\langle x, y\rangle$$

$$(a-b)\langle x, y\rangle = 0 \qquad \langle x, y\rangle = 0$$

e quindi si puo' sempre scegliere una base ortonormale, basta prendere insiemi ortonormali in ciascun autospazio. Analoghe proprieta' valgono per gli operatori unitari: i loro autovalori sono tutti di modulo 1, autovettori relativi ad autovalori diversi sono ortogonali.

8.2 Sottospazi invarianti e somme dirette

Dato un operatore A e un sottospazio S dello spazio V, diciamo che S e' invariante per A se per ogni elemento y di S il trasformato con A e' ancora in S

$$y \in S \quad \Rightarrow \quad Ay \in S$$

cioe' A opera all'interno di S, si scrive anche

$$A(S) \subset S$$

Se $(x_i)_{i=1,\dots,m}$ e' una base in S, e' allora

$$Ax_j = \sum_{i=1}^{m} x_i A_{ij}$$

Pertanto si puo' considerare la sua restrizione a S, cioe' un operatore definito unicamente in S e quindi la sua equazione ad autovalori in S.

Ora completiamo la base di S in modo da ottenere una base di V

$$(x_i)_{i=1,\dots,n} = (x_i)_{i=1,\dots,m} \cup (x_i)_{i=m+1,\dots,n}$$

e definiamo il sottospazio T generato dai vettori $(x_i)_{i=m+1,\dots,n}$. E' evidente che ogni vettore $y \in V$ si scrive in modo unico come somma di un vettore $y_1 \in S$ e un vettore $y_2 \in T$, che chiameremo le sue componenti in S e T

$$y = \sum_{i=1,n} c_i x_i = \sum_{i=1,m} c_i x_i + \sum_{i=m+1,n} c_i x_i = y_1 + y_2$$

Si dice in tal caso che V e' somma diretta dei due sottospazi S e T

$$V = S \oplus T$$

Se ora anche T e' invariante per A, $A(T) \subset T$, anche

$$Ax_j = \sum_{i=m+1,n} x_i A_{ij}$$

e A opera separatamente all'interno dei due sottospazi, si riduce alla somma di due operatori separati, A^1 che opera in S e A^2 in T, e si scrive

$$A = A^1 \oplus A^2$$

Quindi anche l'equazione ad autovalori per A si separa in due equazioni ad autovalori indipendenti, una per A^1 all'interno di S e una per A^2 all'interno di T

$$A^1 y_i^1 = a_i^1 y_i^1$$
$$A^2 y_i^2 = a_i^2 y_i^2$$

e

$$\{a_i\} = \{a_i^1\} \cup \{a_i^2\} \qquad \{y_i\} = \{y_i^1\} \cup \{y_i^2\}$$

saranno gli autovalori di A, e analogamente gli autovettori

In particolare la matrice di A nella base (x_i) ha la struttura a blocchi

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1m} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & 0 & \cdots & 0 \\ A_{m1} & \cdots & A_{mm} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & A_{m+1m+1} & \cdots & A_{m+1n} \\ 0 & \cdots & 0 & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & A_{nm+1} & \cdots & A_{nn} \end{pmatrix} = \begin{pmatrix} A^1 & 0 \\ 0 & A^2 \end{pmatrix}$$

E' evidente che questa proprieta' puo' estendersi a un numero qualunque di sottospazi.

8.3 Equazione ad autovalori e commutativita' di operatori

Supponiamo dati due operatori A e B hermitiani. Valgono allora le seguenti proprieta'

1. Se A e B possiedono una base di autovettori comuni, allora commutano. Infatti sia (x_i) una base tale che

$$Ax_i = a_i x_i \qquad Bx_i = b_i x_i$$

allora

$$(AB - BA)x_i = Ab_i x_i - Ba_i x_i = (b_i a_i - a_i b_i)x_i = 0$$

e quindi

$$[A, B] = 0$$

poiche' annullandosi su una base, si annulla applicato a un qualsiasi vettore

2. Se A e B commutano, cioe'

$$[A, B] = 0$$

allora ogni autospazio V_a di A e' invariante per B. Infatti, sia x un elemento di V_a, vogliamo vedere che anche il vettore Bx appartiene a V_a

$$A(Bx) = B(Ax) = Bax = a(Bx)$$

quindi anche Bx e' ancora autovettore di A relativo allo stesso autovalore

3. Se

$$[A, B] = 0$$

allora esiste una base di autovettori comuni di A e B, cioe' una base (x_i) tale che

$$Ax_i = a_i x_i \qquad Bx_i = b_i x_i$$

Infatti, se si considerano gli autospazi relativi agli autovalori a_i di A, ciascuno di essi e' invariante per B, e genera una decomposizione di tutto lo spazio V is somma diretta

$$V = V_1 \oplus V_2 \oplus \cdots V_m$$

L'equazione ad autovalori per B si separa dunque in un'equazione separata all'interno di ciascun V_i

$$B^i y^i_j = b^i_j y^i_j \qquad y^i_j \in V_i$$

e quindi anche
$$Ay_j^i = a_i y_j^i$$
La riunione di tutti gli $y_j^i \equiv y_{ij}$ fornisce quindi una base (y_i) di autovettori comuni per A e B.

4. Quindi due operatori (hermitiani) commutano se e solo se esiste una base di autovettori comune. E' evidente che questa proprieta' si estende ad un insieme qualunque di operatori che commutino tutti a due a due
$$\{A, B, C, \dots\}$$
$$[A, B] = 0 \quad [A, C] = 0 \quad [B, C] = 0 \quad \dots$$
che quindi possiedono una base di autovettori comuni (x_i)
$$\begin{aligned} A\,x_i &= a_i\,x_i \\ B\,x_i &= b_i\,x_i \\ C\,x_i &= c_i\,x_i \\ &\dots \end{aligned}$$

5. Consideriamo due operatori A e B, che abbiano gli autovettori comuni
$$\begin{aligned} Ax_1 &= a_1 x_1 & \qquad Bx_1 &= b_1 x_1 \\ Ax_2 &= a_2 x_2 & \qquad Bx_2 &= b_2 x_2 \\ Ax_3 &= a_3 x_3 & \qquad Bx_3 &= b_3 x_3 \end{aligned}$$
con $a_1 = a_2$, e $b_2 = b_3$. E' evidente che mentre la specifica dell'autovalore a_3 identifica univocamente (a meno di un fattore di fase inessenziale) il vettore x_3 (unico autovettore relativo ad a_3), la specifica di a_1 non e' sufficiente, perche' qualsiasi combinazione lineare di x_1 e x_2 fornisce lo stesso autovalore. Lo stesso per l'autovalore b_2. Quindi in generale, se c'e' degenerazione, la specifica dei soli autovalori a_i di A non fissa univocamente gli autovettori, e analogamente per B. Ma se consideriamo gli autovettori comuni di A e B, la specifica di una qualsiasi coppia di autovalori (a_i, b_i) identifica univocamente un autovettore. Indichiamo gli autovalori distinti con
$$a_1 = a_2 = a \qquad a_3 = a'$$
$$b_1 = b \qquad b_2 = b_3 = b'$$

allora

$$\begin{array}{ll} (a,b) & \Rightarrow x_1 \\ (a,b') & \Rightarrow x_2 \\ (a',b') & \Rightarrow x_3 \end{array}$$

Diciamo che un insieme di operatori che commutano e' un insieme completo di operatori che commutano (CSCO, complete set of commuting operators) se ad ogni p-upla di autovalori corrisponde un unico autovettore comune

$$(a_1, b_i, c_i, \ldots) \leftrightarrow x_i$$

il che e' come dire che la corrispondenza tra le p-uple $(a, b, c, \ldots)$ e l'indice i che indicizza gli autovettori e' biunivoca. Questo significa che la base di autovettori comuni e' unica (a meno di fattori di fase) e puo' venir indicizzata dagli autovalori stessi $x_{a,b,c,\ldots}$

$$A\, x_{a,b,c,\ldots} = a\, x_{a,b,c,\ldots}$$

$$B\, x_{a,b,c,\ldots} = b\, x_{a,b,c,\ldots}$$

$$C\, x_{a,b,c,\ldots} = c\, x_{a,b,c,\ldots}$$

$$\ldots$$

Questo implica anche che se si ha un ulteriore operatore, Z, che commuta con tutti gli operatori del CSCO, esso ammette la stessa base di autovettori comuni (poiche' e' unica), e i suoi autovalori sono completamente determinati da quelli del CSCO

$$z_i = z(a, b, c, \ldots)$$

sono quindi una funzione di questi. Si dice allora che anche Z e' funzione del CSCO, $Z = Z(A, B, C, \ldots)$

Infine, si assume che gli operatori $A, B, C, \ldots$ che costituiscono il CSCO siano tutti indipendenti, cioe' nessuno sia funzione dei rimanenti. Questo equivale a dire che se si toglie uno qualsiasi degli operatori, l'insieme rimanente non e' piu' un CSCO, ovvero esiste una m-upla di autovalori comuni a cui corrispondono piu' autovettori indipendenti.

6. Nel caso degli autovettori comuni di un CSCO, e' spesso usata la notazione di Dirac per gli autovettori, che sono indicizzati dagli autovalori corrispondenti

$$A\, |abc\ldots\rangle = a\, |abc\ldots\rangle$$

$$B\,|abc\ldots\rangle = b\,|abc\ldots\rangle$$
$$C\,|abc\ldots\rangle = c\,|abc\ldots\rangle$$
$$\ldots$$

8.4 Equazione ad autovalori matriciale

Dato un operatore A, e introdotta una base (x_i) nello spazio, l'equazione ad autovalori per A

$$Ay = ay$$

si traduce nella stessa equazione per le rispettive componenti, la matrice di A e il vettore colonna c

$$Ac = ac$$

Gli autovettori cosi' ottenuti

$$Ac_k = a_k c_k \qquad c_k = (c_{1k}\; c_{2k}\; \ldots c_{nk})^T$$

relativi ad ogni autovalore a_k, rappresentano lo sviluppo dei corrispondenti autovettori y_k nella base di partenza

$$Ay_k = a_k y_k \qquad y_k = \sum_i x_i c_{ik}$$

Se raccogliamo le colonne degli autovettori in una matrice C

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

possiamo riscrivere l'equazione ad autovalori anche nella forma

$$AC = C\mathcal{A}$$

dove $\mathcal{A}$ e' la matrice diagonale degli autovalori

$$\mathcal{A} = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}$$

Infatti

$$(AC)_{ik} = (Ac_k)_i = a_k c_{ik}$$
$$(C\mathcal{A})_{ik} = \sum_j C_{ij} a_j \delta_{jk} = c_{ik} a_k$$

Il problema matriciale, completamente equivalente, e' detto problema ad autovalori algebrico. Possiamo riformulare l'equazione ad autovalori nella forma

$$(A - aI)c = 0$$

che si puo' scrivere per esteso come

$$\begin{cases} (A_{11} - a)c_1 & + & A_{12}c_2 & + & \cdots & + & A_{1n}c_n & = 0 \\ A_{21}c_1 & + & (A_{22} - a)c_2 & + & \cdots & + & A_{2n}c_n & = 0 \\ \cdots & & \cdots & & \cdots & & \cdots & = 0 \\ A_{n1}c_1 & + & A_{n2}c_2 & + & \cdots & + & (A_{nn} - a)c_n & = 0 \end{cases}$$

che rappresenta, noto l'autovalore a, un sistema di equazioni lineari nelle incognite c_i, vettore di componenti dell'autovettore y. Per motivi storici sono anche dette equazioni secolari. Poiche' il sistema e' omogeneo, se c e' un vettore di soluzioni, anche ogni suo multiplo lo e', come deve essere perche' un autovettore e' in ogni vaso determinato a meno di una costante moltiplicativa, che ne fissa la norma. Questo equivale a dire che almeno uno dei c_i puo' essere scelto arbitrariamente. Poiche' in ogni caso il vettore nullo $c = 0$ e' soluzione del sistema (cosiddetta soluzione banale), il sistema avra' soluzioni non nulle se e solo se non e' determinato, cioe' se il determinante della matrice dei coefficienti e' uguale a 0. La condizione di esistenza di autovalori e' allora

$$det(A - aI) = 0$$

Sviluppando il determinante questo conduce as un polinomio di ordine n (tra gli altri termini contiene i prodotti degli elementi sulla diagonale principale) nella variabile a, detto polinomio caratteristico, le cui radici, cioe' le soluzioni dell'equazione caratteristica

$$P_n(a) = 0$$

annullano il determinante e costituiscono quindi gli autovalori a_i della matrice (e quindi dell'operatore) A. Infatti, per ogni a_i l'annullarsi del determinate assicura l'esistenza di almeno una soluzione non nulla (autovettore) del sistema.

Poiche' in C ogni polinomio ha almeno una radice, ne segue che ogni operatore ammette almeno un autovalore e autovettore. Anzi possiamo precisare che esiste certamente almeno un autovettore per ogni radice distinta. Questo non e' piu' vero in R, perche' tutte le radici possono essere complesse, ma un autovettore esiste certamente in R^n con n dispari, perche' in tal caso esiste certamente almeno una radice reale (ad es. $P_n(a)$ diverge per $a \to \pm\infty$ da parti opposte, ed essendo una funzione continua si annulla almeno in un punto dell'asse reale).

Poiche' in ogni modo le radici possono essere ripetute, fino al limite di avere un'unica radice distinta, resta il problema di determinare quanti autovettori possiede l'operatore, e in particolare l'esistenza di una base. Limitiamoci qui al caso degli operatori (e matrici) hermitiane, per il quale e' assicurata l'esistenza di una base di autovettori o.n.

Quindi in linea di principio si possono trovare autovalori e autovettori risolvendo l'equazione caratteristica per trovare gli autovalori, e poi sostituendo questi uno alla volta nelle equazioni secolari, e risolverle ottenendo gli autovalori corrispondenti. Vedremo alcuni esempi.

Merita ricordare che la soluzione numerica del problema algebrico ad autovalori e' un problema centrale in moltissime applicazioni, e in particolare nella MQ. Di fatto l'algoritmo proposto e' utile solo nel caso di matrici molto piccole, e sono stati sviluppati algoritmi molto piu' efficienti per la soluzione, che variano a seconda della classe di problemi, e che sono codificati in programmi di calcolo di pubblico dominio.

8.5 Esempi ed esercizi

- Esempio 1.

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

Gli autovalori sono le radici del polinomio caratteristico

$$|A - aI| = 0$$

$$\begin{vmatrix} 1-a & -1 & 0 \\ -1 & 2-a & -1 \\ 0 & -1 & 1-a \end{vmatrix} = (1-a)[(2-a)(1-a)-1]-(-1)[(-1)(1-a)]$$

$$= (1-a)(2-3a+a^2-1-1) = (1-a)(a^2-3a) = 0$$

che ha radici

$$a_1 = 0$$

$$a_2 = 1$$

$$a_3 = 3$$

Ricaviamo gli autovettori corrispondenti risolvendo il sistema

$$(A - a_k I)c_k = 0$$

dove

$$c_k = \begin{pmatrix} c_{1k} \\ c_{2k} \\ c_{3k} \end{pmatrix}$$

e' l'autovettore relativo all'autovalore a_k. Osserviamo che c_k e' determinato a meno di un coeficiente moltiplicativo arbitrario, che fisseremo con la normalizzazione. Questo vuol dire che possiamo scegliere a piacere una componente, ad esempio c_{1k}, e ricavare le altre in funzione di questa. Per alleggerire la scrittura omettiamo l'indice k.

1. $a_1 = 0$

   $$A - aI = A$$

   abbiamo quindi $Ac = 0$, cioe'

   $$1 \cdot c_1 - 1 \cdot c_2 + 0 \cdot c_3 = 0$$

   $$-1 \cdot c_1 + 2 \cdot c_2 - 1 \cdot c_3 = 0$$

   Per il discorso fatto prima, l'ultima equazione e' pleonastica, non aggiunge nulla di nuovo. Quindi

   $$c_1 - c_2 \quad = 0 \quad \Rightarrow \quad c_2 = c_1$$

   $$-c_1 + 2c_2 - c_3 = 0 \quad \Rightarrow \quad c_3 = c_1$$

   L'autovettore e' quindi

   $$c = \begin{pmatrix} c_1 \\ c_1 \\ c_1 \end{pmatrix} \equiv \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

poiche' possiamo scegliere arbitrariamente c_1. La norma e'

$$\sum_i c_i^2 = 3$$

il vettore normalizzato e' allora

$$c_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

2. $a_2 = 1$

$$(A - 1I)c = 0$$

$$-c_2 = 0 \quad \Rightarrow \quad c_2 = 0$$

$$-c_1 + c_2 - c_3 = 0 \quad \Rightarrow \quad c_3 = -c_1$$

$$c = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad \Rightarrow \quad c_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

3. $a_3 = 3$

$$(A - 3I)c = 0$$

$$-2c_1 - c_2 = 0 \quad \Rightarrow \quad c_2 = -2c_1$$

$$-c_1 - c_2 - c_3 = 0 \quad \Rightarrow \quad c_3 = c_1$$

$$c = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \quad \Rightarrow \quad c_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

Il problema e' cosi' completamente risolto. Verifichiamo per esercizio che i vettori trovati sono effettivamente autovettori relativi agli autovalori corrispondenti. Vediamo per a_3, c_3 :

$$Ac_3 = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1+2 \\ -1-4-1 \\ 2+1 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

come doveva essere.

- Esempio 2.

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

Autovalori:

$$|A - aI| = \begin{vmatrix} 1-a & 1 & 1 \\ 1 & 1-a & 1 \\ 1 & 1 & 1-a \end{vmatrix} = 0$$

$$(1-a)[(1-a)^2 - 1] - 1[(1-a) - 1] + 1[1 - (1-a)] = -a^3 + 3a^2 = 0$$

$$a_1 = 0 \qquad a_2 = 0 \qquad a_3 = 3$$

In questo caso abbiamo degenerazione, due radici sono uguali.

Autovettori. Vediamo prima $a_3 = 3$

$$(A - 3I)c = 0$$

$$-2c_1 + c_2 + c_3 = 0$$

$$c_1 - 2c_2 + c3 = 0$$

Sottraendo membro a membro

$$-3c_1 + 3c_2 = 0 \qquad c_2 = c_1$$

$$-2c_1 + c_2 + c_3 = 0 \qquad c_3 = c_1$$

$$c = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \Rightarrow \quad c_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Consideriamo adesso $a_1, a_2 = 0$

$$Ac = 0$$

$$c_1 + c_2 + c_3 = 0$$

Le altre due righe sono identiche. In questo caso possiamo scegliere arbitrariamente 2 componenti, ad es. c_1 e c_2, e c_3 resta determinato. Avremo pertanto 2 soluzioni indipendenti.

Ad esempio, se scegliamo

$$c_1 = 1, \quad c_2 = 0 \quad \Rightarrow \quad c_3 = -1$$

$$c = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad \Rightarrow \quad c_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

e scegliendo

$$c_1 = 0, \quad c_2 = 1 \quad \Rightarrow \quad c_3 = -1$$

$$c = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \quad \Rightarrow \quad c_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

Quindi in corrispondenza dell'autovalore 0, di molteplicita' 2, otteniamo 2 autovettori linearmente indipendenti. La molteplicita' di a come radice corrisponde al grado di degenerazione dell'autovalore.

Osserviamo come in caso di degenerazione, la scelta della coppia (n nel caso di degenerazione di ordine n) di autovettori linearmente indipendenti sia arbitraria.

Ad esempio scegliendo

$$c_1 = 1, \quad c_2 = 1 \quad \Rightarrow \quad c_3 = -2$$

si ottiene

$$c_1' = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

oppure

$$c_1 = 1, \quad c_2 = -1 \quad \Rightarrow \quad c_3 = 0$$

$$c_2' = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

che non sono altro che la somma e la differenza dei due vettori prima ottenuti. Qualunque coppia di vettori linearmente indipendenti e' accettabile, costituisce una base di dimensione 2 associata all'autovalore $a = 0$ di molteplicita' 2.

Non solo, in questo caso gli autovettori indipendenti non sono necessariamente ortogonali. Ad esempio la prima coppia non lo e', la seconda si'. Anche nel primo caso pero' e' possibile ortogonalizzare gli autovettori, ad esempio ortogonalizzando (alla Schmidt) il secondo autovettore al primo. E'

$$\langle c_1, c_2 \rangle = \frac{1}{2}(1 \cdot 0 + 0 \cdot 1 + (-1) \cdot (-1)) = \frac{1}{2}$$

e

$$\bar{c}_2 = c_2 - \frac{1}{2}c_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 - 1/2 \\ 1 \\ -1 + 1/2 \end{pmatrix}$$

cioe' rinormalizzando

$$\bar{c}_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}$$

- Esempio 3.

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad |A - aI| = (1-a)^2 = 0 \qquad a_1 = a_2 = 1$$

Per $a = 1$ abbiamo $(A - I)c = 0$

$$0c_1 + 1c_2 = 0 \qquad c_2 = 0$$

$$0c_1 + 0c_2 = 0 \qquad \text{identicamente soddisfatta}$$

Quindi l'unico autovettore e'

$$c = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

In questo caso, nonostante la molteplicita' dell'autovalore $a = 1$ sia 2, esiste un solo autovettore relativo. Quindi in questo caso A non ammette una base di autovettori.

Osserviamo che abbiamo assicurato l'esistenza di una base di autovettori per A hermitiano ($A^+ = A$), che nel caso reale si riduce ad A simmetrica ($A^T = A$, e difatti le prime due matrici considerate erano simmetriche. Se la matrice non e' simmetrica (hermitiana), non e' assicurata una base di autovettori. Puo' aversi, oppure no.

- Esempio 4.

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

Si ottiene

$$|A - aI| = a^4 - 1 = 0$$

con le radici

$$a_1 = 1 \quad a_2 = -1 \quad a_3 = i \quad a_4 = -i$$

Osserviamo che se la matrice non e' hermitiana, i suoi autovalori non sono necessariamente reali. In questo caso, nonostante la matrice non sia simmetrica i suoi autovalori sono tutti distinti, e poiche' ad ogni autovalore corrisponde almeno un autovettore, avremo ancora una base di autovettori.

1. $a_1 = 1$

$$(A - 1I)c = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = 0$$

$$-c_1 + c_2 = 0 \qquad c_2 = c_1$$

$$-c_2 + c_3 = 0 \qquad c_3 = c_2 = c_1$$

$$-c_3 + c_4 = 0 \qquad c_4 = c_3 = c_1$$

$$c = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad \Rightarrow \quad c_1 = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

2. $a_2 = -1$

$$(A + 1I)c = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = 0$$

$$c_1 + c_2 = 0 \qquad c_2 = -c_1$$

$$c_2 + c_3 = 0 \qquad c_3 = c_1$$

$$c_3 + c_4 = 0 \qquad c_4 = -c_1$$

$$c_2 = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

3. $a_3 = i$

$$(A - iI)c = \begin{pmatrix} -i & 1 & 0 & 0 \\ 0 & -i & 1 & 0 \\ 0 & 0 & -i & 1 \\ 1 & 0 & 0 & -i \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = 0$$

$$-ic_1 + c_2 = 0 \qquad c_2 = ic_1$$

$$-ic_2 + c_3 = 0 \qquad c_3 = ic_2 = -c_1$$

$$-ic_3 + c_4 = 0 \qquad c_4 = ic_3 = -ic_1$$

$$c_3 = \frac{1}{2}\begin{pmatrix} 1 \\ i \\ -1 \\ -i \end{pmatrix}$$

4. $a_4 = -i$

$$(A + iI)c = \begin{pmatrix} i & 1 & 0 & 0 \\ 0 & i & 1 & 0 \\ 0 & 0 & i & 1 \\ 1 & 0 & 0 & i \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = 0$$

$$ic_1 + c_2 = 0 \qquad c_2 = -ic_1$$

$$ic_2 + c_3 = 0 \qquad c_3 = -ic_2 = -c_1$$

$$ic_3 + c_4 = 0 \qquad c_4 = -ic_3 = ic_1$$

$$c_4 = \frac{1}{2}\begin{pmatrix} 1 \\ i \\ -1 \\ -i \end{pmatrix}$$

Si dice molteplicita' algebrica dell'autovalore a la sua molteplicita' come radice del polinomio caratteristico (contiene cioe' un fattore $(a-a_k)^m$), e molteplicita' geometrica in numero di autovettori linearmente indipendenti corrispondenti. E' sempre

$$1 \leq \text{molt. geometrica} \leq \text{molt. algebrica}$$

Solo quando coincidono si ha una base di autovettori.

Altri esercizi

- Esercizio 1.

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

Si ottiene

$$|A - aI| = -a^3 + 4a = a(4 - a^2) = 0$$

con le radici

$$a_1 = 0 \quad a_2 = 2 \quad a_3 = -2$$

$$c_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad c_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad c_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

- Esercizio 2.

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

Si ottiene

$$|A - aI| = (a - 1)(-a^2 + a + 2) = 0$$

con le radici

$$a_1 = 1 \quad a_2 = -1 \quad a_3 = 2$$

$$c_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad c_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \quad c_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

- Esercizio 3. Vediamo una matrice hermitiana

$$A = \begin{pmatrix} 0 & i & 0 \\ -i & 0 & i \\ 0 & -i & 0 \end{pmatrix}$$

Gli autovalori sono $a_1 = 0,\ a_2 = \sqrt{2},\ a_3 = -\sqrt{2}$

8.6 La matrice simmetrica 2×2

Sia data

$$A = \begin{pmatrix} a_1 & b \\ b & a_2 \end{pmatrix}$$

Conviene per la discussione seguente considerare tale matrice come la rappresentazione dell'operatore

$$A = A^0 + B$$

nella base o.n. (x_1, x_2) degli autovettori di A_0

$$A^0 x_1 = a_1 x_1$$

$$A^0 x_2 = a_2 x_2$$

e

$$\langle x_1, Bx_1 \rangle = \langle x_2, Bx_2 \rangle = 0$$

(quest'ultima condizione in realta' e' irrilevante, in quanto eventuali elementi diagonali di B possono essere sommati ad A^0). Quindi in assenza di B A^0 sarebbe gia' diagonale con gli autovettori (x_1, x_2), la presenza di B modifica autovalori e autovettori, che si ottengono diagonalizzando la matrice dell'operatore completo A. Diremo anche che A^0 e' l'operatore imperturbato, e B la perturbazione, o interazione. E' quindi

$$A^0 = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}$$

diagonale, e

$$B = \begin{pmatrix} 0 & b \\ b & 0 \end{pmatrix}$$

L'equazione caratteristica e'

$$|A - aI| = \begin{vmatrix} a_1 - a & b \\ b & a_2 - a \end{vmatrix} = (a_1 - a)(a_2 - a) - b^2 = 0$$

$$a^2 - (a_1 + a_2) + a_1 a_2 - b^2 = 0$$

$$a_\pm = \frac{a_1 + a_2}{2} \pm \sqrt{(\frac{a_1 - a_2}{2})^2 + b^2} = \frac{a_1 + a_2}{2} \pm D$$

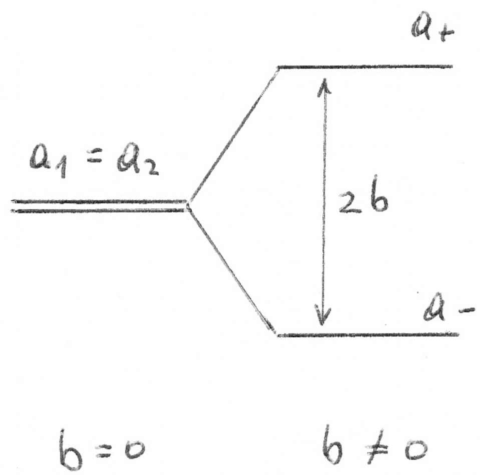

Figure 1: Autovalori del caso degenere

dove si e' posto

$$\Delta = \frac{a_1 - a_2}{2}$$
$$D = \sqrt{\Delta^2 + b^2}$$

- Vediamo prima il caso particolare $\Delta = 0$ (caso degenere, $a_1 = a_2$)

  E'

$$a_+ = a_1 + b \quad \rightarrow \quad -bc_1 + bc_2 = 0 \quad c_1 = c_2 \quad c_+ = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$a_- = a_1 - b \quad \rightarrow \quad bc_1 + bc_2 = 0 \quad c_1 = -c_2 \quad c_- = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

  Si dice caso degenere perche' in assenza del termine extradiagonale, b (cioe' dell'operatore B) i due autovalori a_1 e a_2 della matrice A_0

$$\begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}$$

con i corrispondenti autovettori che sono i due vettori di base canonici11

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

sono degeneri.

In presenza di b (interazione) i livelli si separano, e la separazione e' data da $2b$ (figura 1). Inoltre i corrrispondenti autovettori risultano da una mescolanza $1:1$ degli autovettori iniziali

$$c_+ = \frac{1}{\sqrt{2}}(e_1 + e_2) \qquad c_+ = \frac{1}{\sqrt{2}}(e_1 - e_2)$$

indipendentemente dal valore di b. Quindi, se $a_1 = a_2$, un valore anche piccolissimo di b causa un mescolamento completo degli stati. Osserviamo che in figura si e' assunto $b > 0$. Nel caso $b < 0$ la figura si inverte, a_+ e' il livello inferiore e a_- quello superiore.

- Caso non degenere, $\Delta \neq 0$

$$a_+ = \frac{a_1 + a_2}{2} + D$$

e sostituendo nella seconda riga (e' piu' comodo, si ottiene lo stesso risultato, partendo dalla prima si osservi che $D^2 = \Delta^2 + b^2$ da cui $(\Delta + D)(\Delta - D) = -b^2$)

$$bc_1 + (a_2 - \frac{a_1 + a_2}{2} - D)c_2 = 0$$

$$bc_1 - (\Delta + D)c_2 = 0$$

$$\frac{c_2}{c_1} = \frac{b}{\Delta + D} = t \qquad c_2 = c_1 t \qquad c_+ = \begin{pmatrix} 1 \\ t \end{pmatrix}$$

Se si vuole normalizzare

$$c_1^2 + c_2^2 = 1 + t^2 \qquad c_+ = \frac{1}{\sqrt{1 + t^2}} \begin{pmatrix} 1 \\ t \end{pmatrix}$$

Analogamente per a_- si ottiene

$$\frac{c_1}{c_2} = \frac{\Delta - D}{b} = -\frac{b}{\Delta + D} = -t$$

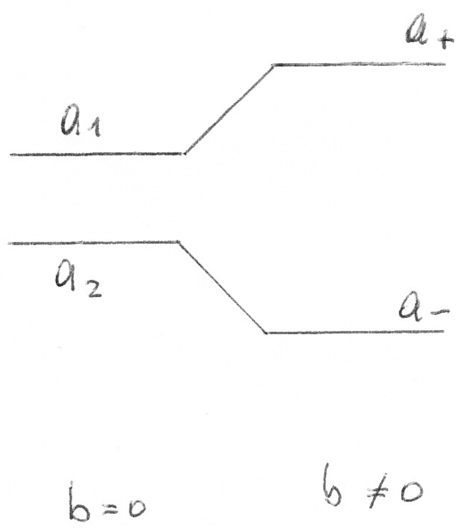

Figure 2: Autovalori del caso non degenere

$$c_- = \frac{1}{\sqrt{1+t^2}} \begin{pmatrix} -t \\ 1 \end{pmatrix}$$

Osserviamo che la condizione di normalizzazione e' automaticamente soddisfatta se poniamo

$$c_1 = \cos\theta \qquad c_2 = \sin\theta \quad \Rightarrow c_1^2 + c_2^2 = 1$$

In tal caso e'

$$\frac{c_2}{c_1} = \frac{\sin\theta}{\cos\theta} = \tan\theta = t$$

Quindi e' $\theta = \arctan t$, e si ha

$$c_+ = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \qquad c_- = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$$

Si sono espressi quindi completamente gli autovalori a_+ e a_- e i rispettivi autovettori c_+ e c_- in termini dei tre elementi di matrice a_1, a_2, b. In questo caso la situazione e' quella indicata in figura 2.

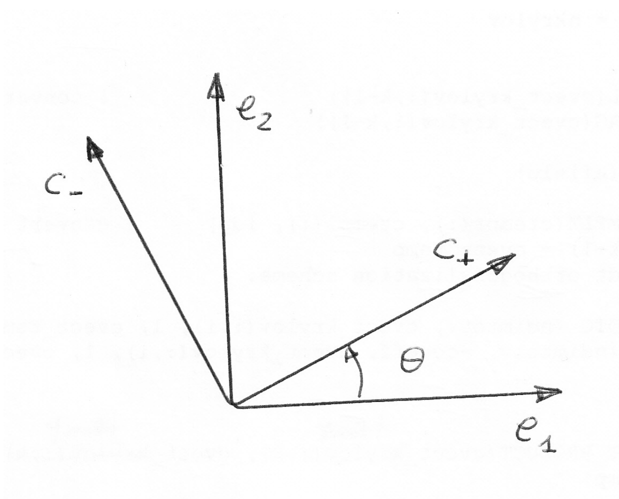

Figure 3: Rotazione dalla base iniziale e_i agli autovettori $c_\pm$

In ogni caso la spaziatura tra gli autovalori "imperturbati" a_1 e a_2, cioe' quelli che la matrice avrebbe per $b = 0$, e gli autovalori finali a_+ e a_- aumenta. Per effetto dell'interazione gli autovalori si divaricano: il piu' alto sale, il piu' basso scende di altrettanto. Osserviamo che e'

$$a_1 + a_2 = a_+ + a_-$$

perche', come sappiamo, la traccia si conserva. In questo caso pero', per

$$b \to 0 \quad \Rightarrow \quad c_+ \to e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad c_- \to e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

quindi per b piccolo si modificano poco, al contrario del caso degenere. Per

$$b \to 0 \quad t \to 0 \quad \cos\theta \to 1 \quad \sin\theta \to 0$$

Se consideriamo $A = A^0 + B$, B "matrice di interazione", che mescola tra loro gli autovettori di A^0

$$A^0 e_1 = a_1 e_1 \qquad A^0 e_2 = a_2 e_2$$

vediamo che l'effetto di B e' di aprire la separazione tra i livelli di A^0, e di "ruotare" gli autovettori di un angolo

$$\theta = \arctan \frac{b}{\Delta + D}$$

Per $\frac{b}{\Delta}$ piccolo il mescolamento e' molto piccolo, per $\Delta \to 0$ e' $\theta = \pi/4$ indipendente da b (figura 3).

- Possiamo considerare di variare a_1, a_2, ovvero Δ, con continuita', con $b > 0$ fissato. Scriviamo

$$a_1 = \frac{a_1 + a_2}{2} + \Delta \qquad a_2 = \frac{a_1 + a_2}{2} - \Delta$$

e consideriamo gli autovalori del problema completo, a_+ e a_- in funzione di Δ, (figura 4). Per $\Delta > 0$ e' $a_1 > a_2$, e per Δ molto grande $a_+ \to a_1$ e $a_- \to a_2$, e anche $c_+ \to e_1$ e $c_- \to e_2$, viceversa per $\Delta < 0$. Per $\Delta \to 0$ invece $a_+ = b$ e $a_- = -b$. Questa, $(2b)$, e' la separazione minima tra i due autovalori a_+ e a_-. Quindi, se consideriamo i due valori diagonali a_1 e a_2 (autovalori di A_0) come funzioni di un parametro s

$$a_1 = as, \qquad a_2 = -as, \qquad \Delta = 2as$$

avremo la situazione come in figura 4: anche se a_1 e a_2 si incrociano, passando per 0, a_+ e a_- restano sempre separati e non si incrociano mai. Questa situazioen e' detta "non crossing rule". Tuttavia la natura degli autovettori cambia bruscamente passando attttraverso la situazione degenere: a destra $c_+ \simeq e_1$, $c_- \simeq e_2$, a sinistra la situazione si capovolge, e ovviamente alla degenerazione il mescolamento e' $1:1$.

- Si puo' considerare piu' in dettaglio il comportamento per b piccolo. Se

$$|b| \ll |\Delta|$$

allora

$$\sqrt{\Delta^2 + b^2} = \Delta\sqrt{1 + \frac{b^2}{\Delta^2}} \simeq \Delta(1 + \frac{1}{2}\frac{b^2}{\Delta^2} + \cdots)$$

da cui

$$a_+ \simeq \frac{a_1 + a_2}{2} + \Delta + \frac{b^2}{2\Delta} = a_1 + \frac{b^2}{2\Delta} = a_1 + \frac{b^2}{a_1 - a_2}$$

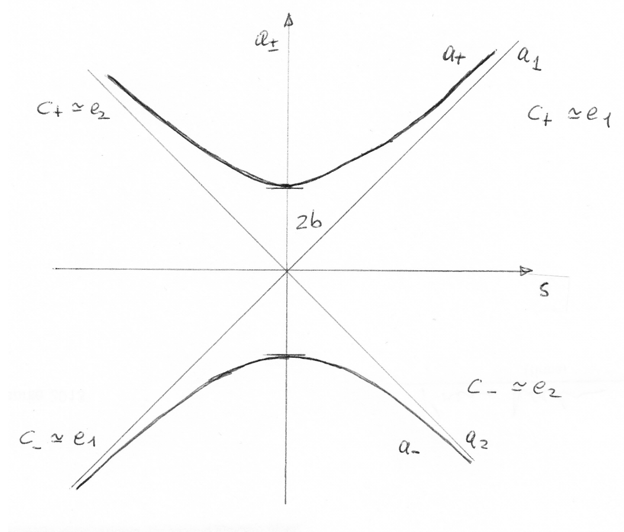

Figure 4: variazione degli autovalori $a_\pm$ al variare degli autovalori imperturbati in funzione di un parametro

$$a_- \simeq \frac{a_1 + a_2}{2} - \Delta - \frac{b^2}{2\Delta} = a_2 - \frac{b^2}{2\Delta} = a_2 + \frac{b^2}{a_2 - a_1}$$

Quindi la correzione all'autovalore a_i e' data dal contributo (detto del 2^o ordine)

$$a_i^{(2)} = \frac{b^2}{a_i - a_j}$$

Analogamente si ha per gli autovettori

$$t = \frac{b}{\Delta + D} \simeq \frac{b}{2\Delta}$$

ovvero

$$\psi_i \simeq \phi_i + \frac{b}{a_i - a_j}\phi_j = c^{(0)}\phi_i + c^{(1)}\phi_j$$

con

$$c^{(0)} \simeq 1 \quad \text{e} \quad c^{(1)} = \frac{b}{a_i - a_j}$$

Osserviamo che l'accuratezza di queste espressioni e' tanto maggiore quanto piu' $|\frac{b}{\Delta}|$ e' piccolo; inoltre la correzione agli autovalori e autovettori e' tanto maggiore quanto piu' e' grande $|b|$ e quanto piu' piccolo e' $|a_i - a_j|$. Lo sviluppo in serie perde significato (non converge) per $|\frac{b}{\Delta}| \geq 1$

8.7 Il caso della base non ortonormale

Consideriamo ancora l'equazione ad autovalori per un operatore A

$$Ay = ay$$

e sia (x_i) una base nello spazio, non ortonormale, con matrice di sovrapposizione

$$S_{ij} = \langle x_i, x_j \rangle$$

e matrice dei prodotti scalari di A

$$A_{ij} = \langle x_i, Ax_j \rangle$$

In ogni caso possiamo sviluppare gli autovettori y

$$y = \sum_i c_j x_j$$

e prendendo il prodotto scalare di $Ay = ay$ a sinistra per x_i

$$\langle x_i, A \sum_j c_j x_j \rangle = \langle x_i, a \sum_j c_j x_j \rangle$$

$$\sum_j A_{ij} c_j = a \sum_j S_{ij} c_j$$

ovvero l'equazione matriciale

$$Ac = aSc$$

dove compare adesso la matrice di sovrapposizione Anche in questo caso questa equazione equivale a un sistema omogeneo

$$(A - aS)c = 0$$

per esteso

$$\begin{cases} (A_{11} - aS_{11})c_1 & + & (A_{12} - aS_{12})c_2 & + & \cdots & + & (A_{1n} - aS_{1n})c_n & = 0 \\ (A_{21} - aS_{21})c_1 & + & (A_{22} - aS_{22})c_2 & + & \cdots & + & (A_{2n} - aS_{2n})c_n & = 0 \\ \cdots & & \cdots & & \cdots & & \cdots & = 0 \\ (A_{n1} - aS_{n1})c_1 & + & (A_{n2} - aS_{n2})c_2 & + & \cdots & + & (A_{nn} - aS_{nn})c_n & = 0 \end{cases}$$

che ha soluzioni non banali se e solo se il determinante della matrice dei coefficienti si annulla

$$det(A - aS) = 0$$

cioe' se a e' una radice del polinomio caratteristico corrispondente

$$P_n(a) = 0$$

e si procede esattamente come nel caso ortonormale. Questo problema, in cui compare la matrice di sovrapposizione, e' noto come problema algebrico agli autovalori generalizzato.

8.7.1 Il problema generalizzato 2×2

Non si perde di generalita' assumendo che i due vettori di base siano normalizzati, in tal caso $S_{11} = S_{22} = 1$ e poniamo $S_{12} = S_{21} = S$. E' allora

$$A - aS = \begin{pmatrix} a_1 - a & b - aS \\ b - aS & a_2 - a \end{pmatrix}$$

e l'equazione caratteristica

$$P_2(a) = (a_1 - a)(a_2 - a) - (b - aS)^2 = 0$$

Consideriamo il caso degenere, $a_1 = a_2$. E' allora

$$(a_1 - a)^2 = (b - aS)^2$$

$$(a_1 - a) = \pm(b - aS)$$

che fornisce le due soluzioni

$$a_+ = \frac{a_1 - b}{1 - S}$$

$$a_- = \frac{a_1 + b}{1 + S}$$

Se assumiamo $b < 0, S > 0$ i livelli si separano come in figura nn: adesso lo spostamento non e' piu' simmetrico, l'abbassamento di a_- rispetto al valore imperturbato a_1 e' minore dell'innalzamento del livello a_+. Questo perche' il denominatore $1 + S$ e' maggiore di quello $1 - S$, e la differenza e' tanto maggiore quanto maggiore e' l'overlap S (ricordiamo che deve essere $S < 1$ per la disuguaglianza di Schwartz)