Appunti di teoria delle matrici

Una matrice $n \times n$ A ammette sempre n autovalori $\lambda_1, \lambda_2, ...\lambda_n$ distinti o coincidenti (reali o complessi), che sono soluzioni dell'equazione di grado n in λ (equazione caratteristica)

$$\det\|A - \lambda 1\| = 0. \tag{1}$$

La struttura del primo membro della (1), che prende il nome di polinomio caratteristico, è

$$\begin{aligned}\det\|A - \lambda 1\| =& \lambda_1\lambda_2...\lambda_n - (\lambda_1\lambda_2...\lambda_{n-1} + \lambda_1\lambda_2...\lambda_{n-2}\lambda_n + ... + \lambda_2\lambda_3...\lambda_n)\lambda + \\ &+ ... + (-1)^{n-1}(\lambda_1 + \lambda_2 + ... + \lambda_n)\lambda^{n-1} + (-1)^n\lambda^n = P(\lambda).\end{aligned} \tag{2}$$

Si chiama autovettore di A, corrispondente all'autovalore λ_i, il vettore colonna a n somponenti $x^{(i)}$, che soddisfa l'equazione

$$Ax^{(i)} = \lambda_i x^{(i)}. \tag{3}$$

Si dimostra che per ogni autovettore distinto di A esiste almeno un autovettore e che autovettori corrispondenti ad autovalori distinti sono <u>linearmente indipendenti</u>.

Data una matrice non singolare S (cioè che ammette inversa S^{-1}, ovvero tale che $\det S \neq 0$) si chiama trasformazione di similitudine la seguente trasformazione

$$A \to S^{-1}AS. \tag{4}$$

L'equazione caratteristica è invariante sotto trasformazione di similitudine:

$$\det\|S^{-1}AS - \lambda 1\| = \det\|S^{-1}(A - \lambda 1)S\| = \det S^{-1}\det\|A - \lambda 1\| \; \det S = \det\|A - \lambda 1\|,$$

dove si è usato il fatto che

$$\det\|S^{-1}S\| = \det S^{-1}\det S = 1. \tag{5}$$

Ne segue che i coefficienti del polinomio caratteristico $P(\lambda)$, vedi l'eq.(2), sono invarianti per trasformazione di similitudine (e quindi lo sono anche gli autovalori λ_i). In particolare il primo coefficiente è il determinante di A:

$$\det A = \lambda_1 \lambda_2 ... \lambda_n, \tag{6}$$

e il penultimo coefficiente è la traccia di A (somma degli elementi diagonali):

$$\mathrm{Tr} A = \sum_{i=1}^{i=n} A_{ii} = \lambda_1 + \lambda_2 + ... + \lambda_n .$$

Ricordiamo infine il <u>teorema di Cayley-Hamilton</u>: lèquazione caratteristica è soddisfatta dalla matrice A. Cioè

$$P(A) = 0, \tag{7}$$

dove P è il polinomio caratteristico della matrice A.

<u>Matrici semisemplici</u>

Una matrice $n \times n$ A si chiama semisemplice se ha n autovettori linearmente indipendenti. Se in particolare A ha n autovalori distinti allora è semisemplice.

L'importanza di questa definizione deriva dal fatto che se A è semisemplice allora si può porre sotto forma diagonale (cioè con solo gli elementi diagonali diversi da zero) mediante una trasformazione di similitudine. Purtroppo non esiste un criterio semplice per stabilire se una matrice è o non è semisemplice; esiste tuttavia un algoritmo che permette di stabilire se una matrice è o non è semisemplice (diagonalizzabile) senza bisogno di conoscere gli autovalori, basato sul polinomio minimo, che non discutiamo (vedi M.Abate, "Geometria", McGraw-Hill, pagg. 357 e seguenti). Vedremo però più sotto che alcuni tipi di matrici (normali e in particolare hermitiane, unitarie ecc.) facilmente riconoscibili sono semisemplici.

Sia A semisemplice e siano $\lambda_1, \lambda_2, ..., \lambda_n$ i suoi autovalori (eventualmente coincidenti a gruppi, nel qual caso si dice che c'è degenerazione). L'equazione agli autovalori si scriverà:

$$Ax^{(i)} = \lambda_i x^{(i)}, \quad (i = 1, 2, ..., n), \tag{8}$$

dove per ipotesi i vettori $x^{(i)}$ sono linearmente indipendenti e quindi costituiscono una base per lo spazio dei vettori colonna a n componenti.

Consideriamo la matrice $n \times n$

$$S = (x^{(1)}, x^{(2)}, ..., x^{(n)}) \tag{9}$$

(le cui colonne sono costituite dai vettori $x^{(i)}$). S è non singolare poichè gli $x^{(i)}$ sono linearmente indipendenti, da cui segue det$S \neq 0$.

Per calcolare S^{-1} procediamo nel modo seguente. Dato un vettore w

$$w = \begin{pmatrix} w_1 \\ w_2 \\ . \\ . \\ w_n \end{pmatrix},$$

si definisce il suo coniugato hermitiano $w^\dagger$ come il vettore riga:

$$w^\dagger = (\bar{w}_1, \bar{w}_2, ..., \bar{w}_n).$$

Si definiscono allora vettori reciproci (duali) degli autovettori $x^{(i)}$ i vettori $w^{(i)}$ tali che

$$w^{(i)\dagger} x^{(j)} = \delta_{ij}. \tag{10}$$

I $w^{(i)}$ sono <u>linearmente indipendenti</u> perchè tali sono gli $x^{(i)}$ (se infatti non lo fossero esisterebbero delle costanti c_i tali che $\sum_{i=1}^n c_i w^{(i)} = 0$. Moltiplicando per $\bar{c}_i$ la (9) e sommando su i si avrebbe $\sum_{i=1}^n \bar{c}_i \delta_{ij} = \bar{c}_j = 0$, contro l'ipotesi che non tutte le c_i siano nulle).

Risulta allora

$$S^{-1} = \begin{pmatrix} w^{(1)\dagger} \\ w^{(2)\dagger} \\ . \\ . \\ w^{(n)\dagger} \end{pmatrix}, \tag{11}$$

(dove le righe sono formate dai vettori riga $w^{(i)\dagger}$) come si verifica facilmente (è facile la verifica di $S^{-1}S = 1$ mediante la (10), mentre per $SS^{-1} = 1$ occorre l'uso della completezza della base $x^{(i)}$).

Se ora si calcola la matrice (forma diagonale di A)

$$\Lambda = S^{-1}AS = S^{-1}A(x^{(1)}, x^{(2)}, ..., x^{(n)}) = S^{-1}(\lambda_1 x^{(1)}, \lambda_2 x^{(2)}, ..., \lambda_n x^{(n)}) =$$

$$\begin{pmatrix} \lambda_1 w^{(1)\dagger} x^{(1)}, \lambda_2 w^{(1)\dagger} x^{(2)}, ... \\ \lambda_1 w^{(2)\dagger} x^{(1)}, \lambda_2 w^{(2)\dagger} x^{(2)}, ... \\ ... \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & 0 & ... & 0 \\ 0 & \lambda_2 & 0 & ... & 0 \\ 0 & 0 & \lambda_3 & ... & 0 \\ . & . & . & ... & . \\ 0 & 0 & 0 & ... & \lambda_n \end{pmatrix}, \tag{12}$$

si ha che Λ è diagonale.

Osserviamo che la trasformazione S non è unica, se non altro perchè l'ordine dei λ_i è arbitrario.

Si definisce coniugata hermitiana di A la matrice $A^\dagger = \bar{A}^T$, dove A^T è la matrice trasposta di A. In componenti

$$(A^\dagger)_{ij} = \bar{A}_{ji}. \tag{13}$$

Si può dimostrare che i vettori reciproci $w^{(i)}$ sono autovettori di $A^\dagger$, corrispondenti agli autovalori $\bar{\lambda}_i$:

$$A^\dagger w^{(i)} = \bar{\lambda}_i w^{(i)}.$$

Matrici normali

Una matrice $n \times n$ N si dice normale se per essa vale:

$$NN^\dagger = N^\dagger N. \tag{14}$$

Si dimostra che una matrice normale è semisemplice e che i suoi autovettori si possono scegliere in modo che siano a coppie ortogonali ($x^{(i)\dagger} x^{(j)} = 0$ se $i \neq j$).

Poichè dalla (3) segue che i vettori sono definiti a meno di un fattore arbitrario, i vettori $x^{(i)}$ si potranno scegliere in modo che valga l'ortonormalità:

$$x^{(i)\dagger}x^{(i)} = \delta_{ij} \tag{15}$$

(quindi i vettori reciproci coincidono in questo caso con i vettori $x^{(i)}$).

Ne segue che per la matrice S che diagonalizza la N vale

$$S^{\dagger}S = 1. \tag{16}$$

Una matrice U che soddisfa la

$$UU^{\dagger} = 1 \tag{17}$$

si dice <u>unitaria</u> (caso particolare di matrice normale). Si dimostra che gli autovalori di U sono di <u>modulo 1</u> e che autovettori corrispondenti ad autovalori distinti sono ortogonali.

Una matrice H tale che

$$H = H^{\dagger} \tag{18}$$

si dice <u>hermitiana</u> (altro caso particolare di matrice normale). Si dimostra che gli autovalori di una matrice hermitiana sono <u>reali</u> e che autovettori corrispondenti ad autovalori distinti sono ortogonali.

Se in particolare U è reale si ha $U^{\dagger} = U^T$; in tal caso si ha una matrice, che si può indicare con O, che si chiama <u>ortogonale</u>

$$O^T O = 1. \tag{19}$$

Se H è reale si ha $H = H^T$; in tal caso si ha una matrice simmetrica S:

$$S = S^T. \tag{20}$$

Questi tipi di matrici sono tutti casi particolari di matrici normali e quindi semisemplici. Quindi diagonalizzabili.

Esempio di diagonalizzazione di una matrice semisemplice

Consideriamo la matrice $3x3$:

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

Poichè $A = A^T$ sappiamo già che A è semisemplice e quindi è diagonalizzabile.

L'equazione caratteristica

$$\det\|A - \lambda 1\| = \begin{vmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{vmatrix} = -\lambda^3 + 2\lambda = 0$$

[n.b. per la (7) si avrà $-A^3 + 2A = 0$]
ha per soluzioni

$$\lambda_1 = \sqrt{2}, \quad \lambda_2 = 0, \quad \lambda_3 = -\sqrt{2},$$

cioè tre autovalori distinti e reali. L'autovettore corrispondente all'autovalore λ_1 sia $x^{(1)}$:

$$Ax^{(1)} = \lambda_1 x^{(1)}.$$

Posto

$$x^{(1)} = \begin{pmatrix} a \\ b \\ c \end{pmatrix},$$

si trovano le equazioni

$$b = \sqrt{2}a, \quad a + c = \sqrt{2}b, \quad b = \sqrt{2}c,$$

e quindi

$$x^{(1)} = a \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}$$

con a indeterminato.

Analogamente si trova

$$x^{(2)} = a' \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad x^{(3)} = a'' \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}.$$

Questi tre autovettori sono linearmente indipendenti e tra loro ortogonali (A è simmetrica reale e quindi hermitiana).

Se si normalizzano all'unità, cioè se si richiede $x^{(i)\dagger} x^{(i)} = 1$, $i = 1, 2, 3$, si deve scegliere

$$a = \frac{1}{2}, \quad a' = \frac{1}{\sqrt{2}}, \quad a'' = \frac{1}{2},$$

cioè

$$x^{(1)} = \frac{1}{2} \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}, \quad x^{(2)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad x^{(3)} = \frac{1}{2} \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix},$$

col che si ha

$$x^{(i)\dagger} x^{(j)} = \delta_{ij}, \quad (i = 1, 2, 3),$$

cioè la base dei vettori reciproci coincide con la base $x^{(i)}$.

Posto allora

$$S = (x^{(1)}, \quad x^{(2)}, \quad x^{(3)}) = \begin{pmatrix} \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{\sqrt{2}} & \frac{1}{2} \end{pmatrix},$$

si ha

$$S^{-1} = \begin{pmatrix} x^{(1)\dagger} \\ x^{(2)\dagger} \\ x^{(3)\dagger} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{\sqrt{2}} & \frac{1}{2} \end{pmatrix},$$

e quindi

$$\Lambda = S^{-1} A S = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix},$$

come si verifica facilmente.

Esempio di diagonalizzazione di una matrice semisemplice non normale

Sia

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad (AA^\dagger \neq A^\dagger A),$$

$$\det\|A - \lambda 1\| = -\lambda(1-\lambda)^2 = 0.$$

Quindi gli autovalori sono

$$\lambda_1 = \lambda_2 = 1, \quad \lambda_3 = 0.$$

All'autovalore 1 corrisponde l'equazione agli autovalori

$$Ax = x,$$

e posto

$$x = \begin{pmatrix} a \\ b \\ c \end{pmatrix},$$

si hanno le equazioni

$$a = a, \quad c = b, \quad c = c.$$

Perciò x è della forma

$$x = \begin{pmatrix} a \\ b \\ b \end{pmatrix}.$$

Si possono perciò scegliere due autovettori linearmente indipendenti

$$x^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad x^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

All'autovalore 0 corrisponde l'autovettore (normalizzato a 1)

$$x^{(3)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Si verifica che $x^{(2)}$ e $x^{(3)}$, pur corrispondendo ad autovalori diversi, non sono tra loro ortogonali (la matrice non è normale).

Si possono calcolare i vettori reciproci, che risultano

$$w^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad w^{(2)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad w^{(3)} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix},$$

tali che

$$w^{(i)\dagger} x^{(j)} = \delta_{ij}.$$

La matrice S di diagonalizzazione è

$$S = (x^{(1)}, \quad x^{(2)}, \quad x^{(3)}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

e

$$S^{-1} = \begin{pmatrix} w^{(1)} & \\ w^{(2)} & \\ w^{(3)} & \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

Infine

$$\Lambda = S^{-1} A S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Esempio di matrice non semisemplice

Sia

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

$$\det\|A - \lambda 1\| = (1-\lambda)^3 = 0;$$

si ha quindi l'autovalore 1 tre volte degenere:

$$\lambda_1 = \lambda_2 = \lambda_3 = 1.$$

L'equazione agli autovettori

$$Ax = x$$

ha per soluzione un unico autovettore

$$x^{(1)} = a \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

dove si può porre $a = 1$.

Non abbiamo quindi un numero sufficiente di autovettori linearmente indipendenti per costruire la matrice di diagonalizzazione S. La matrice A non è semisemplice e non è quindi diagonalizzabile. Si può però ricondurre la matrice A alla cosiddetta forma canonica di Jordan, mediante una procedura che è applicabile a qualsiasi matrice non semisemplice.

Il metodo è il seguente: per completare lo spazio degli autovettori si costruiscono degli autovettori generalizzati $x^{(2)}$ e $x^{(3)}$ così definiti:

$$\begin{aligned} (A-\lambda)x^{(2)} &= x^{(1)}, \\ (A-\lambda)x^{(3)} &= x^{(2)}, \end{aligned} \tag{21}$$

da cui si ha, con una opportuna scelta delle costanti indeterminate, la base di vettori

$$x^{(1)} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad x^{(2)} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \quad x^{(3)} = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}.$$

I vettori reciproci risultano

$$w^{(1)} = \begin{pmatrix} \frac{3}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \quad w^{(2)} = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}, \quad w^{(3)} = \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix}.$$

Posto

$$S = (x^{(1)}, \quad x^{(2)}, \quad x^{(3)}) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ -1 & -1 & 2 \end{pmatrix}$$

si ha

$$S^{-1} = \begin{pmatrix} w^{(1)\dagger} \\ w^{(2)\dagger} \\ w^{(3)\dagger} \end{pmatrix} = \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}.$$

Se si calcola Λ si ha

$$\Lambda = S^{-1}AS = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

che è la forma canonica di Jordan.

Più in generale la forma canonica di Jordan ottenibile con questo procedimento è una matrice triangolare (elementi sotto la diagonale principale (o sopra) tutti nulli) con gli autovalori (necessariamente) sulla diagonale principale, raccolti a gruppi di degenerazione. Ogni gruppo di degenerazione, per il quale non si abbiano autovettori linearmente indipendenti in numero sufficiente, è bordato superiormente (o inferiormente) da degli 1, uno per ogni autovettore generalizzato che è stato necessario introdurre. Ogni autovettore genera degli autovettori generalizzati, mediante formule analoghe alle (21), che formano un insieme detta catena. L'insieme delle catene per ogni autovalore degenere, con grado di degenerazione m_λ, forniscono in modo automatico tanti autovettori generalizzati quanti sono necessari per sottendere il sottospazio di degenerazione corrispondente all'autovalore λ. Nel caso dell'esempio si ha una sola catena formata da un autovettore e da due autovettori generalizzati.

Matrici definite positive

Una matrice $n \times n$ hermitiana A si dice definita positiva se per ogni vettore x non nullo, di componeneti $x_i, \quad i = 1, 2, ..., n$, si ha

$$\bar{x}_i A_{ij} x_j > 0,$$

o in notazione di matrici

$$x^\dagger A x > 0. \tag{22}$$

Dimostriamo che condizione necessaria e sufficiente perchè valga la (22) è che gli autovalori di A siano positivi

$$\lambda_i > 0, \quad i = 1, 2, ..., n. \tag{23}$$

La condizione è necessaria, infatti scelto x coincidente con un autovettore $x = x^{(i)}$ si ha

$$x^{(i)\dagger} A x^{(i)} = \lambda_i |x^{(i)}|^2 > 0, \tag{24}$$

da cui $\lambda_i > 0$.

La condizione è sufficiente. Infatti ricordiamo che una matrice hermitiana si può diagonalizzare con una matrice unitaria (vedi l'eq. (16)). Sia S una matrice unitaria (ortogonale se A è reale, quindi simmetrica) che diagonalizza A. Se $\lambda_i > 0$ $(i = 1, 2, ..., n)$ scelto un vettore non nullo x definiamo $y = S^{-1}x = S^\dagger x$. Si ha allora

$$x^\dagger A x = (Sy)^\dagger A (Sy) = y^\dagger S^\dagger A S y = y^\dagger S^{-1} A S y = y^\dagger \Lambda y, \tag{25}$$

dove Λ è la forma diagonale di A:

$$\Lambda_{ij} = \lambda_i \delta_{ij}.$$

Si ha perciò

$$x^\dagger A x = \bar{y}_i \lambda_i y_i = \sum_{i=1}^{n} \lambda_i |y_i|^2 > 0, \tag{26}$$

e quindi

$$x^\dagger A x > 0,$$

qualunque sia x.

Una matrice hermitiana si dice poi semidefinita positiva se è

$$x^\dagger A x \geq 0 \tag{27}$$

per un qualsiasi vettore x non nullo. In questo caso si ha che qualche autovalore di A è nullo.

* * * * * * * * *