

Capitolo G47:

Similarità, diagonalizzabilità, autovalori e autovettori

Contenuti delle sezioni

- a. Matrici equivalenti -ME e matrici simili p.1 b. Trasformazioni e matrici diagonalizzabili p.3
 c. Autovettori e autovalori p.4 d. Polinomio caratteristico e spettro di una matrice diagonalizzabile p.8
 e. Matrici ortogonali e matrici normali p.11

G47:0.01 In questo capitolo, riprendiamo la proprietà di similarità fra matrici e fra trasformazioni lineari per introdurre le nozioni di matrici e operatori diagonalizzabili e successivamente discutere le prime proprietà degli autovalori e degli autovettori di una matrice quadrata e di un operatore lineare entro uno spazio vettoriale $\mathbb{R}^{\times d}$ munito del prodotto scalare canonico.

Le pagine che seguono fanno riferimento ad un unico campo \mathbb{F} che per le questioni generali può considerarsi qualsiasi. Per alcuni risultati sui sistemi di autovettori e autovalori risulta opportuno servirsi del campo sui numeri complessi, ma non procederemo molto in questa direzione.

G47:a. Matrici equivalenti -ME e matrici simili

G47:a.01 Ricordiamo che due matrici A e B sul campo \mathbb{F} dello stesso profilo $e \times d$ sono state chiamate equivalenti-ME sse si trovano due matrici invertibili $P \in \mathbf{Mat}_e$ e $Q \in \mathbf{Mat}_d$ tali che sia $B = P A Q$. Prescindendo dalle specifiche matrici P e Q , denotiamo con \sim_{ME} la relazione che collega matrici dello stesso profilo ed equivalenti -ME. Essa è effettivamente una relazione di equivalenza: infatti \sim_{ME} è riflessiva ($A = \hat{\mathbf{1}}_e A \hat{\mathbf{1}}_d$), simmetrica ($B = P A Q \implies A = P^{-1} B Q^{-1}$) e transitiva ($B = P A Q, C = R B S$ con $R, P \in \mathbf{MatInv}_d$ e $Q, S \in \mathbf{MatInv}_e$ implicano $C = (R P) A (Q S)$ con $R P \in \mathbf{MatInv}_e$ e $Q S \in \mathbf{MatInv}_d$).

Una classe di equivalenza -ME si può interpretare come insieme delle rappresentazioni di una unica trasformazione lineare relative a diverse coppie di basi, la prima componente base per lo spazio dominio e la seconda componente base per lo spazio codominio.

Lo studio dell'equivalenza -ME fra matrici ha mostrato che ogni matrice di profilo $e \times d$ è equivalente-ME ad una matrice che presenta una forma a blocchi riducibile mediante permutazioni di righe e colonne

$$\begin{bmatrix} \hat{\mathbf{1}}_r & \mathbf{0}_{r,d-r} \\ \mathbf{0}_{e-r,r} & \mathbf{0}_{e-r,d-r} \end{bmatrix}.$$

Questa matrice rappresenta una trasformazione lineare L di uno spazio \mathbf{V} a d dimensioni in uno spazio \mathbf{W} ad e dimensioni rispetto a due basi ordinate \mathcal{B} di \mathbf{V} e \mathcal{C} di \mathbf{W} ; la L ha rango r e nullità $d - r$ e la base $\mathcal{B} = \langle \mathbf{b}_1, \dots, \mathbf{b}_r, \mathbf{b}_{r+1}, \dots, \mathbf{b}_d \rangle$ ha gli ultimi $d - r$ vettori che costituiscono una base per $\ker(L)$, mentre i primi r costituiscono una base di un sottospazio complementare di tale nucleo; a sua volta la

base $\mathcal{C} = \langle \mathbf{c}_1, \dots, \mathbf{c}_r, \mathbf{c}_{r+1}, \dots, \mathbf{c}_e \rangle$ ha i primi r vettori che costituiscono una base per $\text{cod}(L)$, i rimanenti costituendo una base di un sottospazio complementare di $\text{cod}(L)$.

Riferendoci a queste basi la L è la trasformazione che manda il generico vettore $\sum_{i=1}^d \mathbf{b}_i v_i$ di \mathbf{V} nel vettore $\sum_{i=1}^r \mathbf{c}_i v_i$ di \mathbf{W} .

Per le applicazioni di una $L \in \{ \mathbf{V} \xrightarrow{\mathbb{F}lin} \mathbf{W} \}$ l'equivalenza -ME interessa in quanto si collega alla possibilità di individuare una base per $\mathbf{V} = \text{dom}(L)$ e una base per $\text{cod}(L)$ relativamente alle quali la trasformazione L risulta particolarmente maneggevole.

G47:a.02 Limitiamoci ora agli endomorfismi lineari, cioè alle trasformazioni lineari di uno spazio in se, e alle loro rappresentazioni riferite ad un'unica base.

Due matrici quadrate A ed B di uno stesso ordine d si dicono **matrici similari** sse si trova una matrice P dello stesso ordine d invertibile tale che sia $B = P A P^{-1}$. Questa P la chiamiamo **matrice similarizzante** per $\langle A, B \rangle$; chiaramente P^{-1} è la matrice similarizzante per la coppia riflessa $\langle B, A \rangle$. Serve osservare che per una qualsiasi coppia di matrici similari $\langle A, B \rangle$ la matrice similarizzante P non è univocamente determinata, perchè lo è anche ogni αP per ogni $\alpha \in \mathbb{F}_{nz}$.

Prescindendo dalle specifiche P che collegano due matrici similari A e B , per denotare tale similarità scriviamo $A \approx B$.

Per maggiore chiarezza presentiamo anche una particolarizzazione delle considerazioni all'inizio di :a.01.

(1) Prop.: La similarità è una relazione di equivalenza fra matrici quadrate dello stesso ordine.

Dim.: Consideriamo matrici il cui ordine denotiamo con d . La riflessività della relazione è ottenuta dalla $\hat{\mathbf{1}}_d A \hat{\mathbf{1}}_d^{-1} = A$; la simmetria discende dal fatto che $P A P^{-1} = B \iff A = P^{-1} B P$; la transitività della relazione si ottiene con la seguente implicazione

$$B = P A P^{-1}, C = Q B Q^{-1} \implies C = Q P A P^{-1} Q^{-1} = (Q P) A (Q P)^{-1} \blacksquare$$

Due matrici similari sono sicuramente equivalenti -ME, mentre vi sono molte coppie di matrici equivalenti -ME che non sono similari. In altre parole, nell'ambito delle matrici di un dato ordine d , la similarità è un'equivalenza decisamente più fine della equivalenza -ME.

In particolare sono matrici di ordine d similari soltanto con sé stesse la matrice $\mathbf{0}_{d,d}$, per la quale per ogni matrice invertibile P si ha $P \mathbf{0}_{d,d} P^{-1} = \mathbf{0}_{d,d}$, e la matrice unità $\hat{\mathbf{1}}_d$, per la quale per ogni P invertibile si ha $P \hat{\mathbf{1}}_d P^{-1} = P P^{-1} = \hat{\mathbf{1}}_d$; più in generale, ad ogni $\lambda \in \mathbb{F}$ corrisponde la classe di similarità $\{\hat{\lambda}\}$.

Si osserva che, mentre per ogni $\alpha \in \mathbb{F}_{nz}$ si ha $A \sim_{ME} B \implies \alpha A \sim_{ME} \alpha B$, viceversa per ogni $\alpha \in (\mathbb{F} \setminus \{0, 1\})$ accade che $A \approx B \implies \alpha A \not\approx B$.

G47:a.03 Nella classe di similarità di una matrice diagonale non costante D si trovano evidentemente, le matrici diagonali D' ottenute permutando le entrate sulla diagonale principale. La matrice P similarizzante di una tale coppia $\langle D, D' \rangle$ è una matrice permutativa (che ricordiamo ha come inversa un'altra matrice permutativa). Ad esempio si trova la relazione di similarità

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 11 & 0 & 0 \\ 0 & 22 & 0 \\ 0 & 0 & 33 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 22 & 0 & 0 \\ 0 & 33 & 0 \\ 0 & 0 & 11 \end{bmatrix}.$$

Sono molte le classi di similarità con infiniti elementi. Ad esempio in $\text{Mat}_{2,\mathbb{R}}$ appartengono alla stessa classe di similarità le matrici che esprimono le riflessioni rispetto alle rette per l'origine, rette che si

possono utilmente caratterizzare con le equazioni $x \cos \phi + y \sin \phi = 0$, per tutti gli angoli convessi $\phi \in [0, \pi)$.

G47:a.04 (1) Prop.: Se $B = P A P^{-1}$ ed A è una matrice invertibile, allora anche B è invertibile.

Dim.: Si ottiene la matrice inversa della $P A P^{-1}$ con lo sviluppo $(P A P^{-1})^{-1} = (P^{-1})^{-1} A^{-1} P^{-1} = P A^{-1} P^{-1}$ e si verifica che questa è l'inversa della B :

$$(P A^{-1} P^{-1}) (P A P^{-1}) = P A^{-1} A P^{-1} = \hat{\mathbf{1}}_d \quad \blacksquare$$

Quindi la classe delle matrici invertibili è unione di classi di similarità.

(2) Prop.: Per ogni scalare $\lambda \in \mathbb{F}$, se $A - \lambda \hat{\mathbf{1}}_d$ e B sono matrici simili, allora A è simile alla $B + \lambda \hat{\mathbf{1}}_d$.

Dim.: Se $B = P(A - \lambda \hat{\mathbf{1}}_d) P^{-1}$, allora $B = P A P^{-1} - \lambda \hat{\mathbf{1}}_d$ e $P A P^{-1} = B + \lambda \hat{\mathbf{1}}_d \quad \blacksquare$

(3) Prop.: La similarità rispetta la moltiplicazione per uno scalare, cioè per ogni $\alpha \in \mathbb{F}$ si ha

$$B = P A P^{-1} \implies \alpha B = P(\alpha A) P^{-1} \quad \blacksquare$$

In generale invece la similarità non rispetta la somma e il prodotto di matrici. Possiamo enunciare solo la proprietà che segue.

(4) Prop.: Consideriamo $A, B, C, D \in \mathbf{Mat}_d$ e $P \in \mathbf{MatInv}_d$ tali che sia $C = P A P^{-1}$, $D = P B P^{-1}$:

- (i) $C + D = P(A + B)P^{-1}$,
- (ii) $C \cdot D = P(A \cdot B)P^{-1} \quad \blacksquare$

(5) Prop.: Se la matrice P è la similarizzante della coppia di matrici $\langle A, B \rangle$, allora sono in relazione di similarità grazie alla stessa P anche le potenze A^m e B^m per m intero naturale qualsiasi e, più in generale, le matrici $p(A)$ e $p(B)$ costruite con un qualsiasi polinomio in una variabile sul campo $\mathbb{F} p(x)$.

Dim.: Se $B = P A P^{-1}$, ricordando la (4) (ii), si trova $B^m = (P A P^{-1})^m = P A^m P^{-1}$.

Consideriamo il polinomio $p(x) \in \mathbb{F}[x]$ avente la forma $p(x) = \alpha_0 + \alpha_1 x + \dots + \alpha_n x^n$; tenendo conto della similarità fra le potenze e della (3), per ciascuno degli addendi dei polinomi $p(A)$ e $p(B)$, cioè per $j = 0, 1, \dots, n$, si ha

$$\alpha_j B^j = \alpha_j (P A P^{-1})^j = \alpha_j P A^j P^{-1} = P \alpha_j A^j P^{-1}.$$

Per la somma di questi monomi basta ricordare la (4)(i) \blacksquare

G47:b. Trasformazioni e matrici diagonalizzabili

G47:b.01 Le matrici diagonali sono matrici facilmente interpretabili e maneggevoli. In particolare consideriamo in $\mathbb{R} \times \mathbb{R}$ e in $\mathbb{R}^{\times 3}$ le matrici

$$\begin{bmatrix} 0.5 & 0 \\ 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

La prima dimezza le ascisse e raddoppia le ordinate; il suo effetto si ricava dal fatto di trasformare il quadrato con i vertici in $\langle 0, 0 \rangle$, $\langle 1, 0 \rangle$, $\langle 1, 1 \rangle$ e $\langle 0, 2 \rangle$ nel rettangolo con i vertici in $\langle 0, 0 \rangle$, $\langle 0.5, 0 \rangle$, $\langle 0.5, 2 \rangle$ e $\langle 0, 2 \rangle$, o anche dal trasformare la circonferenza $\text{Circ}(\mathbf{0}_2, 1)$ nell'ellisse con centro in $\mathbf{0}_2$, avente semiasse maggiore verticale di lunghezza 2 e avente semiasse minore orizzontale di lunghezza 0.5.

La seconda consiste nel raddoppio delle ascisse e nella riflessione rispetto ad Oy e alla triplicazione delle ordinate.

La terza non cambia le coordinate x , raddoppia le y e triplica le z ; essa trasforma la sfera $x^2 + y^2 + z^2 = 1$ avente centro nell'origine e raggio 1 nell'ellissoide avente centro nell'origine, come lunghezze dei semiassi, risp., 1, 2 e 3.

La quarta trasforma le figure tridimensionali in figure del piano $z = 0$, non modificando le coordinate x e raddoppiando le y .

G47:b.02 Per il seguito risulta comodo denotare l'insieme delle matrici diagonali di ordine d con $\mathbf{MatDiag}_d$ ed individuare una matrice diagonale di ordine d che come entrate sulla diagonale principale presenta nell'ordine gli scalari $\lambda_1, \lambda_2, \dots, \lambda_d$ con la notazione

$$\text{diagmat}(\lambda_1, \lambda_2, \dots, \lambda_d) := \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_d \end{bmatrix}.$$

Questa notazione consente di esprimere facilmente il prodotto di due matrici diagonali come:

$$\text{diagmat}(\lambda_1, \lambda_2, \dots, \lambda_d) \cdot \text{diagmat}(\mu_1, \mu_2, \dots, \mu_d) = \text{diagmat}(\lambda_1\mu_1, \lambda_2\mu_2, \dots, \lambda_d\mu_d).$$

Si trova la seguente proprietà per le combinazioni lineari di matrici diagonali

$$\alpha \cdot \text{diagmat}(\lambda_1, \lambda_2, \dots, \lambda_d) + \beta \cdot \text{diagmat}(\mu_1, \mu_2, \dots, \mu_d) = \text{diagmat}(\alpha\lambda_1 + \beta\mu_1, \alpha\lambda_2 + \beta\mu_2, \dots, \alpha\lambda_d + \beta\mu_d).$$

Questa proprietà consente di considerare l'insieme delle matrici diagonali di un dato ordine come uno spazio vettoriale dello stesso ordine. Inoltre per ogni polinomio nella variabile x , $P(x) = \sum_{j=1}^m p_j x^j$, si trova che il polinomio in una matrice diagonale è la matrice diagonale costituita dai valori del polinomio nelle rispettive entrate diagonali

$$P(\text{diagmat}(\lambda_1, \lambda_2, \dots, \lambda_d)) = \text{diagmat}(P(\lambda_1), P(\lambda_2), \dots, P(\lambda_d)).$$

G47:b.03 Un endomorfismo lineare del genere $\{\mathbf{V} \xrightarrow{\mathbb{F}lin} \mathbf{V}\}$ si dice **operatore diagonalizzabile** o **endomorfismo diagonalizzabile** nel campo \mathbb{F} sse si trova una base per \mathbf{V} nella quale è rappresentato da una matrice diagonale.

Una matrice di \mathbf{Mat}_d si dice **matrice diagonalizzabile** sse è simile ad una matrice diagonale, ovvero sse rappresenta un endomorfismo diagonalizzabile.

Un esempio di matrice diagonalizzabile è il seguente

$$\begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}.$$

Non tutte le matrici sono diagonalizzabili: in particolare ogni matrice di ordine d strettamente triangolare inferiore o superiore, essendo nilpotente, cioè dotata di una potenza nulla, non può essere simile ad una matrice diagonale diversa dalla $\mathbf{0}_{d,d}$, matrice che ha tutte le potenze diagonali.

Occorre segnalare che talune matrici non sono diagonalizzabili in un dato campo, ma lo sono in un campo più esteso; in particolare vi sono matrici sui reali che non sono diagonalizzabili in \mathbb{R} , ma lo sono nel campo dei numeri complessi, grazie alla maggiore versatilità di quest'ultimo.

Dati i vantaggi operativi che offrono le matrici diagonali, risulta chiaramente utile essere in grado di stabilire se una data matrice è diagonalizzabile e nel caso positivo essere in grado di individuare una sua matrice simile diagonale, cioè essere in grado di procedere alla sua **diagonalizzazione**.

G47:b.04 (1) Prop.: Un endomorfismo lineare L dello spazio V a d dimensioni è diagonalizzabile sse esistono una base ordinata di V $\mathcal{B} = \langle |1\rangle, |2\rangle, \dots, |d\rangle \rangle$ e una d -upla di scalari $\langle \lambda_1, \lambda_2, \dots, \lambda_d \rangle$ tali che valgono le equazioni

$$L|1\rangle = \lambda_1|1\rangle, \quad L|2\rangle = \lambda_2|2\rangle, \quad \dots, \quad L|d\rangle = \lambda_d|d\rangle.$$

Dim.: Per definizione un endomorfismo è diagonalizzabile sse esiste una base nella quale è rappresentato da una matrice diagonale e i vettori di tale base sono trasformati dalla matrice in vettori loro proporzionali ■

G47:c. Autovettori e autovalori

G47:c.01 Consideriamo uno spazio V sul campo \mathbb{F} a d dimensioni e un operatore $L \in \{V \mapsto_{\mathbb{F}lin} V\}$. Si dice che l'operatore L possiede un **autovettore** $\mathbf{v} \in V \setminus \mathbf{0}_d$ relativo ad un **autovalore** $\lambda \in \mathbb{F}$ sse si ha

$$L(\mathbf{v}) = \lambda \mathbf{v}.$$

Si osserva che tutti i vettori proporzionali ad un autovettore di un operatore sono anch'essi autovettori dello stesso operatore relativi allo stesso autovalore: infatti

$$\forall \alpha \in \mathbb{F}_{nz} : L(\alpha \mathbf{v}) = \lambda \alpha \mathbf{v} \iff L(\alpha \mathbf{v}) = \lambda \alpha \mathbf{v}.$$

Si possono quindi prendere in considerazione **raggi di autovettori**. Nell'ambito di un tale insieme di vettori talora risulta opportuno scegliere come rappresentativo un autovettore di norma 1; questa scelta non è univoca nè per gli spazi sui reali (invece di un autovettore normalizzato \mathbf{v} si potrebbe considerare $-\mathbf{v}$), nè per gli spazi sui complessi (invece di un autovettore normalizzato \mathbf{v} si potrebbero considerare tutti i vettori della forma $e^{i\theta} \mathbf{v}$ per qualsiasi $\theta \in [0, 2\pi)$).

Consideriamo un operatore che in una base ordinata $\langle \mathbf{b}_1, \dots, \mathbf{b}_d \rangle$ di V è rappresentato da una matrice diagonale $\text{diagmat}(\lambda_1, \dots, \lambda_d)$: i successivi vettori della base sono suoi autovettori relativi, risp., agli autovalori $\lambda_1, \dots, \lambda_d$.

Le definizioni precedenti si estendono alle matrici quadrate in modo prevedibile, cioè considerandole rappresentazioni di operatori relative a determinate basi di riferimento. Si dice che la matrice A di ordine d possiede un **autovettore** (colonna) $\mathbf{v} \in \mathbb{F}^{\times d} \setminus \{\mathbf{0}_d\}$ relativo ad un **autovalore** $\lambda \in \mathbb{F}$ sse avviene che $A\mathbf{v} = \lambda \mathbf{v}$.

Si considerino ad esempio le matrici

$$A = \begin{bmatrix} 5/2 & 1/2 \\ 1/2 & 5/2 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 11/6 & 1/6 \\ 1/3 & 5/3 \end{bmatrix}$$

Si verifica facilmente che $\langle 1, 1 \rangle$ è autovettore della A con autovalore 3 e che $\langle -1, 1 \rangle$ è autovettore della A con autovalore 2; a sua volta per la B si verifica che $\langle 1, 1 \rangle$ è autovettore con autovalore 2 e che $\langle 1, -2 \rangle$ è autovettore relativo all'autovalore 1.5.

G47:c.02 Osserviamo l'opportunità di non considerare tra i possibili autovettori il vettore nullo $\mathbf{0}_d$; infatti tale vettore sarebbe autovettore relativo all'autovalore 0 per ogni operatore lineare su $\mathbb{F}^{\times d}$ e questo fatto non porterebbe ad alcuna distinzione fra gli operatori.

Inoltre per gli spazi di una sola dimensione gli autovettori e gli autovalori risultano inutili: nello spazio \mathbb{F} ogni operatore è la moltiplicazione per uno scalare λ , ogni vettore diverso da 0 è suo autovettore

ed ha come autovalore λ . Gli autovettori, come vedremo, servono a distinguere i diversi raggi di uno spazio per quanto concerne l'azione di alcuni operatori, e questa distinzione in una dimensione non presenta alcun interesse. Già in due dimensioni invece lo studio di autovettori e autovalori risulta significativo e concretamente utile.

Osserviamo esplicitamente che stabilire se un operatore L ha come autovalore lo 0 può risultare importante. Questa precisazione equivale alla determinazione del nucleo di L , in quanto $L(\mathbf{z}) = \mathbf{z} \implies \mathbf{z} \in \ker(L)$.

Segnaliamo che ai termini italiani autovettore ed autovalore corrispondono i termini inglesi *eigenvector* e *eigenvalue*, a loro volta derivati dai termini tedeschi *eigenvektor* ed *eigenwert*. Talora si usano anche i termini **vettore caratteristico** e **valore caratteristico**, nonchè i termini **vettore proprio** e **valore proprio**.

L'insieme degli autovalori di un operatore lineare L si chiama **spettro** di L e si denota con $\text{spec}(L)$; corrispondentemente l'insieme degli autovalori di una matrice quadrata A si chiama **spettro** di A e si denota con $\text{spec}(A)$. Il termine spettro deriva dal fatto che in [[spettroscopia]] lo spettro di un sistema atomico o molecolare \mathcal{S} , cioè l'insieme delle radiazioni che il sistema \mathcal{S} emette o assorbe, si spiega facendo riferimento all'insieme degli autovalori dell'[[operatore hamiltoniano]] che governa la dinamica di tale \mathcal{S} .

G47:c.03 Cerchiamo gli autovettori e gli autovalori della matrice $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \\ -1 & 2 & 3 \end{bmatrix}$. Per individuare uno scalare λ tale che esista un vettore $\mathbf{v} = \langle v_1, v_2, v_3 \rangle$ per il quale sia $A\mathbf{v} = \lambda\mathbf{v}$, riscriviamo l'equazione matriciale nella forma $(A - \lambda\hat{\mathbf{1}}_3)\mathbf{v} = \mathbf{0}_3$:

$$\begin{bmatrix} 1-\lambda & 2 & 1 \\ 2 & 0-\lambda & -2 \\ -1 & 2 & 3-\lambda \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Questo sistema lineare omogeneo ha soluzione diversa da $\mathbf{0}_3$ solo se la matrice dei coefficienti ha determinante nullo; quindi chiediamo

$$0 = |A - \lambda\hat{\mathbf{1}}_3| = \begin{vmatrix} 1-\lambda & 2 & 1 \\ 2 & 0-\lambda & -2 \\ -1 & 2 & 3-\lambda \end{vmatrix} = \lambda^3 - 4\lambda^2 + 4\lambda = \lambda(\lambda - 2)^2.$$

Gli autovalori sono quindi $\lambda_1 = 0$ e $\lambda_2 = 2$, ossia $\text{spec}(A) = \{0, 2\}$. Nel caso $\lambda_1 = 0$ si ha il sistema lineare omogeneo

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \\ -1 & 2 & 3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{che implica} \quad \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} \alpha \\ -\alpha \\ \alpha \end{bmatrix} \quad \text{per un qualsiasi } \alpha \in \mathbb{R}_{nz}$$

Il non annullamento di α scende dalla richiesta non annullamento degli autovettori.

Analogamente nel caso $\lambda_2 = 2$ si ha il sistema

$$\begin{bmatrix} -1 & 2 & 1 \\ 2 & -2 & -2 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{che implica} \quad \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} \beta \\ 0 \\ \beta \end{bmatrix} \quad \text{per un qualsiasi } \beta \in \mathbb{R}_{nz}$$

Anche il non annullamento di β segue dalla richiesta non annullamento degli autovettori.

G47:c.04 Il semplice esempio precedente suggerisce di cercare interi sottospazi di autovettori; questi possono avere 2 o più dimensioni. In effetti insieme a due autovettori relativi a uno stesso autovalore, sono autovettori relativi a tale autovalore tutte le loro combinazioni lineari:

$$L\mathbf{v} = \lambda\mathbf{v}, L\mathbf{w} = \lambda\mathbf{w} \implies \forall \langle \alpha, \beta \rangle \in \mathbb{R} \times \mathbb{R} : L(\alpha\mathbf{v} + \beta\mathbf{w}) = \lambda(\alpha\mathbf{v} + \beta\mathbf{w}).$$

Quindi ad un autovalore di una matrice quadrata o di un operatore lineare risulta associato un intero sottospazio-nz di autovettori.

Diciamo quindi **autospatio** relativo a un autovalore λ di un operatore lineare o di una matrice quadrata il sottospazio-nz determinato dagli autovettori relativi all'autovalore λ .

G47:c.05 Semplici esempi di operatori diagonalizzabili dotati di un autospatio a più dimensioni si ottengono dai proiettori su sottospazi sac v. G40:b.02). Ad esempio, per lo spazio \mathbb{R}^4 la matrice

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

esprime il proiettore sul piano delle combinazioni lineari dei versori canonici \mathbf{e}_1 ed \mathbf{e}_3 ; chiaramente essa presenta l'autovalore 1 e il suo autospatio è il sottospazio-nz sotteso dai suddetti versori, cioè $\text{span}_{nz}(\mathbf{e}_1, \mathbf{e}_3) = SAC_{1010}$.

Più in generale una matrice diagonale con gruppi di entrate sulla diagonale principale coincidenti presenta autospatzi multidimensionali. Ad esempio in \mathbb{C}^7 la matrice di ordine 7

$$\text{diagmat} \left(2, 1 - i, 2, \sqrt{5}, 1 - i, \sqrt{5}, \sqrt{5} \right)$$

presenta 3 autospatzi: $\text{span}_{nz}(\mathbf{e}_1, \mathbf{e}_3)$ bidimensionale relativo all'autovalore 2, $\text{span}_{nz}(\mathbf{e}_2, \mathbf{e}_5)$ bidimensionale relativo all'autovalore $1 - i$ e $\text{span}_{nz}(\mathbf{e}_4, \mathbf{e}_6, \mathbf{e}_7)$ tridimensionale relativo all'autovalore $\sqrt{5}$.

G47:c.06 Una visione più completa degli autospatzi di un operatore o di una matrice diagonalizzabile scende dalla proposizione che segue.

(1) Prop.: Ogni insieme di autovettori di un operatore lineare L (o di una matrice) relativi ad autovalori distinti è costituita da vettori linearmente indipendenti.

Dim.: Per l'operatore L agente su uno spazio a d dimensioni e per $i = 1, \dots, k \leq d$ si abbia $A \mathbf{v}_i = \lambda_i \mathbf{v}_i$ con $i \neq j \implies \lambda_i \neq \lambda_j$.

Procediamo per induzione dimostrando che per $h = 1, \dots, k$ sono linearmente indipendenti gli autovettori $\mathbf{v}_1, \dots, \mathbf{v}_h$.

Ovviamente il solo $\{\mathbf{v}_1\}$ costituisce un insieme di vettori linearmente indipendenti.

Facciamo l'ipotesi induttiva di avere dimostrata l'indipendenza degli autovettori $\mathbf{v}_1, \dots, \mathbf{v}_{h-1}$, cioè di poter affermare

$$\sum_{i=1}^{h-1} \alpha_i \mathbf{v}_i = \mathbf{0}_d \implies \alpha_1 = \dots = \alpha_{h-1} = 0.$$

Per il passo h -esimo si abbia una combinazione lineare di h autovettori che si annulla $\sum_{i=1}^h \alpha_i \mathbf{v}_i = \mathbf{0}_d$; ad entrambi i membri di questa uguaglianza applichiamo prima l'operatore L poi la moltiplicazione per λ_h , ottenendo $L \sum_{i=1}^h \alpha_i \mathbf{v}_i = \sum_{i=1}^h \lambda_i \alpha_i \mathbf{v}_i = \mathbf{0}_d$ e $\lambda_h \sum_{i=1}^h \alpha_i \mathbf{v}_i = \sum_{i=1}^h \lambda_h \alpha_i \mathbf{v}_i = \mathbf{0}_d$.

Sottraiamo membro a membro ed otteniamo $\sum_{i=1}^h (\lambda_h - \lambda_i) \alpha_i \mathbf{v}_i = \sum_{i=1}^{h-1} (\lambda_h - \lambda_i) \alpha_i \mathbf{v}_i = \mathbf{0}_d$.

A questo punto possiamo applicare l'ipotesi induttiva ed otteniamo per $i = 1, \dots, h - 1$: $(\lambda_h - \lambda_i) \alpha_i = 0$; per la supposta diversità degli autovalori si conclude che deve essere $\alpha_1 = \dots = \alpha_{h-1} = 0$; in conseguenza di queste relazioni di annullamento, la prima equazione conduce alla $\alpha_h \mathbf{v}_h = \mathbf{0}_d$ e quindi alla $\alpha_h = 0$ e in definitiva alla indipendenza lineare degli h autovettori $\mathbf{v}_1, \dots, \mathbf{v}_h$.

Per l'arbitrarietà di h l'enunciato risulta dimostrato per induzione ■

G47:c.07 A questo punto si riesce a stabilire un criterio sufficiente a garantire la diagonalizzabilità di una qualsiasi matrice quadrata.

(1) Coroll.: Una matrice di ordine d che possiede d autovalori diversi è diagonalizzabile.

Dim.: Denotiamo con A una matrice con i requisiti dell'enunciato; i suoi autovettori, grazie a :c.06, sono linearmente indipendenti, cioè costituiscono una base di $\mathbb{F}^{\times d}$. Denotiamo questi vettori con $\mathbf{b}_1, \dots, \mathbf{b}_d$; ciascun \mathbf{b}_i si trova a partire dal sistema delle equazioni lineari $(A - \lambda_i \hat{\mathbf{1}}_d)\mathbf{b}_i = \mathbf{0}_d$ per $i = 1, 2, \dots, d$. La matrice A si può considerare come la rappresentazione nella base dei versori canonici $|1\rangle = \mathbf{e}_1, \dots, |d\rangle = \mathbf{e}_d$ di un operatore lineare L e si può scrivere $A = [j, i = 1, \dots, d : a_{j,i}] = [j, i = 1, \dots, d : \langle j|L|i\rangle]$.

Si tratta di trovare la rappresentazione di L nella base $\{\mathbf{b}_1, \dots, \mathbf{b}_d\}$ cioè di trovare

$$D = [k, h = 1, \dots, d : d_{k,h}] = [k, h = 1, \dots, d : \langle \mathbf{b}_k|L|\mathbf{b}_h\rangle].$$

Chiaramente per $h, k = 1, \dots, d : d_{k,h} = \langle \mathbf{b}_k|L|\mathbf{b}_h\rangle = \delta_{k,h} \lambda_k \|\mathbf{b}_k\|^2$.

La diagonalizzazione della A si ottiene per il tramite della matrice P le cui successive righe forniscono le componenti nella base canonica dei successivi vettori $\mathbf{b}_1, \dots, \mathbf{b}_d$. Infatti per $k, h = 1, \dots, d$ si ha :

$$\langle \mathbf{b}_k|L|\mathbf{b}_h\rangle = \langle \mathbf{b}_k|\hat{\mathbf{1}}_d L \hat{\mathbf{1}}_d|\mathbf{b}_h\rangle = \sum_{j,i=1}^d \langle \mathbf{b}_k|\mathbf{e}_j\rangle \langle \mathbf{e}_j|L|\mathbf{e}_i\rangle \langle \mathbf{e}_i|\mathbf{b}_h\rangle = \sum_{j,i=1}^d \langle \mathbf{b}_k|\mathbf{e}_j\rangle a_{j,i} \langle \mathbf{e}_i|\mathbf{b}_h\rangle.$$

Quindi, definita la matrice $P := [\langle \mathbf{b}_k|\mathbf{e}_j\rangle : k, j = 1, \dots, d]$, la matrice diagonale richiesta si ottiene come $D = P A P^{-1}$ ■

Va ribadito che l'avere autovalori tutti diversi è condizione sufficiente e non necessaria per la diagonalizzabilità. Ad esempio la matrice $\text{diagmat}(2, 1 - i, 2, \sqrt{5}, 1 - i, \sqrt{5}, \sqrt{5})$ incontrata in :c.04 costituisce un esempio di matrice di ordine 7 che è diagonalizzabile (diagonale) pur presentando solo 3 autovalori diversi.

G47:c.08 Nell'ambito degli spazi vettoriali e delle matrici sui reali vi sono operatori e corrispondenti matrici quadrate che non possiedono alcun autovettore. Gli esempi più evidenti sono quelli costituiti dalle rotazioni proprie diverse dall'identità e dalla simmetria centrale. Nel caso semplice delle rotazioni del piano $\mathbb{R} \times \mathbb{R}$, abbiamo le matrici

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad \text{con } 0 < \theta < \pi \quad \text{oppure} \quad \pi < \theta < 2\pi .$$

Per trovare un autovalore reale di tale matrice si dovrebbe individuare una soluzione reale dell'equazione

$$\begin{vmatrix} \cos \theta - \lambda & -\sin \theta \\ \sin \theta & \cos \theta - \lambda \end{vmatrix} = 0 \quad \text{ovvero} \quad (\cos \theta - \lambda)^2 + \sin^2 \theta = \lambda^2 - 2 \cos \theta \lambda + 1 = 0 .$$

Per le soluzioni di questa equazione si trova $\lambda = \frac{2 \cos \theta \pm \sqrt{4 \cos^2 \theta - 4}}{2} = \cos \theta \pm \sqrt{-\sin^2 \theta}$; quindi non si ha alcuna soluzione reale.

Si osserva che per gli spazi vettoriali sui reali si può definire autovettore di un operatore lineare L ogni vettore sul quale L non esercita alcun cambiamento di direzione diverso dal passaggio alla direzione oposta (*grosso modo*, che non esercita alcuna rotazione efficace).

G47:c.09 (1) Prop.: Consideriamo un operatore lineare L sullo spazio $\mathbb{F}^{\times d}$ dotato di un autovettore \mathbf{v} e più precisamente sia $L\mathbf{v} = \lambda\mathbf{v}$. Valgono le seguenti relazioni

- (i) $L^2\mathbf{v} = \lambda^2\mathbf{v}$ e più in generale $\forall m = 2, 3, \dots : (L^m)\mathbf{v} = (\lambda^m)\mathbf{v}$.
- (ii) $\forall \alpha \in \mathbb{F} : (\alpha L)\mathbf{v} = (\alpha\lambda)\mathbf{v}$.
- (iii) $\forall \alpha \in \mathbb{F} : (L + \alpha\hat{\mathbf{1}}_d)\mathbf{v} = (\lambda + \alpha)\mathbf{v}$.
- (iv) Se $p(x)$ è un polinomio nella variabile x , allora $p(L)\mathbf{v} = p(\lambda)\mathbf{v}$.
- (v) Se L è invertibile, allora $\forall m = 2, 3, \dots : (L^{-1})^m\mathbf{v} = \lambda^{-m}\mathbf{v}$ ■

G47:d. Polinomio caratteristico e spettro di una matrice diagonalizzabile

G47:d.01 Consideriamo $A \in \mathbf{Mat}_d$; si dice **polinomio caratteristico** di A nella variabile x il determinante della matrice $A - x\hat{\mathbf{1}}_d$, che denotiamo con:

$$\mathbf{charpol}(A, x) := \det(A - x\hat{\mathbf{1}}_d) .$$

Ad esempio per le matrici degli ordini 2 e 3:

$$\mathbf{charpol} \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}, x \right) = \begin{vmatrix} a-x & b \\ c & d-x \end{vmatrix} = x^2 - (a+d)x + ad - bc ,$$

$$\mathbf{charpol} \left(\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}, x \right) = \begin{vmatrix} a_{1,1}-x & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2}-x & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3}-x \end{vmatrix} .$$

$$-x^3 + x^2(a_{1,1} + a_{2,2} + a_{3,3}) - x \left(\begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} + \begin{vmatrix} a_{1,1} & a_{1,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} + \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} \right) + \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} .$$

Va segnalato che su alcuni testi del polinomio caratteristico di una matrice A si dà la definizione leggermente diversa $\det(x\hat{\mathbf{1}}_d - A)$, espressione che per noi fornisce $(-1)^d \mathbf{charpol}(A, x)$.

G47:d.02 Diamo ora una formula per il polinomio caratteristico di una qualsiasi matrice A di un qualsivoglia ordine d .

$$(1) \quad \mathbf{charpol}(A, x) = \begin{vmatrix} a_{1,1}-x & a_{1,2} & \dots & a_{1,d} \\ a_{2,1} & a_{2,2}-x & \dots & a_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d,1} & a_{d,2} & \dots & a_{d,d}-x \end{vmatrix} = (-x)^d + \sum_{j=1}^{d-1} (-x)^j S_{d,j} + \det(A) ,$$

dove $S_{d,j}$ denota la somma su tutti i sottoinsiemi I di $\{1, 2, \dots, d\}$ formati da $d - j$ elementi (indici interi) dei determinanti delle sottomatrici ottenute dalla A eliminando le righe e le colonne con gli indici in I ; questa definizione si formalizza scrivendo

$$(2) \quad S_{d,j} := \sum_{I \in \mathbf{Subset}_{d-j}(1,2,\dots,d)} \det \left(A_{\mathbb{E}_{I,I}} \right) .$$

L'espressione a secondo membro si spiega con le seguenti considerazioni. Il determinante $|A - x\hat{\mathbf{1}}_d|$ va visto come somma di addendi provenienti dai $d!$ termini dello sviluppo del determinante a loro volta sviluppati per eliminare fattori della forma $(a_{i,i} - x)$; il monomio $S_{d,j}(-x)^j$ del polinomio è ottenuto dagli addendi nei quali entrano j fattori derivati da j entrate della forma $(a_{i,i} - x)$; a seconda della provenienza, questi addendi si ripartiscono in corrispondenza con i sottoinsiemi di indici \bar{I} che forniscono i fattori $(-x)^j$; questi j fattori sono da moltiplicare per i contributi delle righe e delle colonne della A appartenenti ad $I = \{1, 2, \dots, d\} \setminus \bar{I}$, contributi espressi da $\det \left(A_{\mathbb{E}_{I,I}} \right)$.

Queste considerazioni in particolare portano ad affermare che

$$S_{d,d-1} = \sum_{i=1}^d a_{i,i} = \text{Tr}(A) \quad S_{d,1} = \sum_{i=1}^d \det \left(A_{\mathbb{E}_{i,i}} \right) \quad S_{d,0} = \det(A)$$

G47:d.03 (1) Prop.: Il polinomio caratteristico di una matrice non cambia quando la matrice viene sostituita da una similare.

Dim.: Si considerino le matrici $P \in \mathbf{MatInv}_d$, $A \in \mathbf{Mat}_d$, e la matrice similare di quest'ultima $B := PAP^{-1}$. Per i polinomi caratteristici si ha

$$\begin{aligned} \mathbf{charpol}(B) &= \det(B - x\hat{\mathbf{1}}_d) = \det(PAP^{-1} - P(x\hat{\mathbf{1}}_d)P^{-1}) = \det(P(A - x\hat{\mathbf{1}}_d)P^{-1}) \\ &= \det(P) \det(A - x\hat{\mathbf{1}}_d) \det(P^{-1}) = \det(A - x\hat{\mathbf{1}}_d) = \mathbf{charpol}(A) \blacksquare \end{aligned}$$

L'invarianza per similarità di una matrice rende lecito definire come **polinomio caratteristico** dell'operatore L nella variabile x e denotare con $\mathbf{charpol}(L, x)$ il polinomio caratteristico nella variabile x di una qualsiasi delle matrici A che rappresentano l'operatore lineare L .

Si dice **equazione caratteristica** nella variabile x di una matrice quadrata A di ordine d l'equazione polinomiale di grado d nella variabile x $\mathbf{charpol}(A, x) = 0$.

Si dice **equazione caratteristica** nella variabile x di un operatore lineare L in uno spazio con d dimensioni l'equazione polinomiale di grado d nella variabile x : $\mathbf{charpol}(L, x) = 0$, ovvero l'equazione caratteristica nella variabile x di una qualsiasi A delle matrici di ordine d che rappresentano la L .

G47:d.04 Vedremo tra poco l'importanza del polinomio caratteristico di una matrice o di un operatore e delle sue radici, cioè delle soluzioni dell'equazione caratteristica di una matrice o di un operatore. Il calcolo del polinomio caratteristico di una matrice con il crescere del suo ordine diventa sempre più oneroso; l'espressione generale :d.02(1) spesso non è utilizzabile efficientemente. Le soluzioni tramite espressioni delle equazioni algebriche generali risultano agevoli fino al II grado, sono date da formule con radicali piuttosto pesanti per il III e il IV grado e, per il [[teorema di Ruffini-Abel]], risulta impossibile trattarle con espressioni generali che si servono di radicali per i gradi dal V in su.

Inoltre risulta determinante il campo degli scalari sul quale si considerano definiti lo spazio vettoriale e le matrici da trattare. Per questi problemi risulta importante potersi collocare nel campo dei numeri complessi: infatti in questo campo vale il teorema fondamentale dell'algebra che afferma che ogni polinomio di grado d in tale campo possiede d zeri, pur di contarli con la dovuta molteplicità.

Per molti problemi applicativi si impone la necessità di ricorrere a calcoli approssimati. Spesso conviene cercare se il calcolo del polinomio caratteristico e delle sue radici possa essere semplificato con considerazioni specifiche, abbandonando il piano delle soluzioni generali. Possono allora essere utili considerazioni sulle simmetrie della matrice e sulla possibilità di sostituirla con una matrice similare.

G47:d.05 Una semplificazione evidente è fattibile per la diagonalizzazione delle matrici triangolari superiori o inferiori a blocchi.

(1) Prop.: I polinomi caratteristici delle matrici di ordine d triangolare a blocchi si fattorizzano con le formule seguenti:

$$\begin{aligned} \mathbf{charpol} \left(\begin{bmatrix} A_1 & B \\ \mathbf{0}_{d-k,k} & A_2 \end{bmatrix}, x \right) &= \begin{vmatrix} A_1 - x\hat{\mathbf{1}}_k & B \\ \mathbf{0}_{d-k,k} & A_2 - x\hat{\mathbf{1}}_{d-k} \end{vmatrix} = |A_1 - x\hat{\mathbf{1}}_k| \cdot |A_2 - x\hat{\mathbf{1}}_{d-k}|, \\ \mathbf{charpol} \left(\begin{bmatrix} A_1 & \mathbf{0}_{k,d-k} \\ C & A_2 \end{bmatrix}, x \right) &= \begin{vmatrix} A_1 - x\hat{\mathbf{1}}_k & \mathbf{0}_{k,d-k} \\ C & A_2 - x\hat{\mathbf{1}}_{d-k} \end{vmatrix} = |A_1 - x\hat{\mathbf{1}}_k| \cdot |A_2 - x\hat{\mathbf{1}}_{d-k}|. \end{aligned}$$

Dim.: Sono conseguenze delle formule di fattorizzazione del determinante delle matrici triangolari a blocchi (v. G42:d) \blacksquare

Formule analoghe si hanno per matrici triangolari a blocchi per sistemi che presentano $k \cdot k$ blocchi con $k > 2$.

G47:d.06 (1) Prop.: Una matrice di ordine d il cui polinomio caratteristico possiede d radici diverse è diagonalizzabile.

Dim.: Se le diverse radici dell'equazione caratteristica della A si denotano con $\lambda_1, \dots, \lambda_d$, allora si può scrivere

$$\text{charpol}(A, x) = (\lambda_1 - x) \cdots (\lambda_d - x) .$$

Grazie a :c.06, i d autovettori relativi agli autovalori diversi costituiscono una base per lo spazio $\mathbb{F}^{\times d}$ e grazie a :c.07 essi consentono di individuare una matrice simile alla A che è diagonale ■

G47:d.07 (1) Prop.: Una matrice quadrata sul campo complesso ha almeno un autovettore.

Dim.: Discende dal teorema fondamentale dell'algebra ■

G47:d.08 Teorema di Cayley-Hamilton Ogni matrice quadrata costituisce uno zero per il proprio polinomio caratteristico, ovvero $\text{charpol}(A, A) = 0$.

G47:e. Matrici ortogonali e matrici normali

G47:e.01 Ricordiamo (v. G41:f) che si dice **matrice ortogonale** una matrice quadrata A tale che, denotando con d il suo ordine, si trova che

$$A \cdot A^\top = A^\top \cdot A = \hat{\mathbf{1}}_d .$$

Una matrice ortogonale A quindi è una matrice invertibile con l'inversa uguale alla trasposta: $A^{-1} = A^\top$. Denotiamo con **MatOrt_d** l'insieme delle matrici ortogonali di ordine d .

Esempi ovvi di matrici ortogonali sono le matrici identità dei vari ordini.

Sono ortogonali anche tutte le matrici permutative: accade infatti che per una tale matrice P si ha $P^{-1} = P^\top$.

(1) Prop.: Le matrici ortogonali di ordine 2 sono le matrici della forma

$$\begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} .$$

Dim.:

Anche per le matrici ortogonali di ordine superiore a 2 si trovano espressioni parametriche, ma si tratta di espressioni sensibilmente più complesse. Tuttavia alcune delle matrici ortogonali di ordine 3 si ottengono con semplici ampliamenti delle precedenti matrici:

$$\begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} , \quad \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} , \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{bmatrix} .$$

G47:e.02 Il precedente procedimento per ottenere altre matrici ortogonali a partire da alcune matrici può essere esteso grazie agli enunciati algebrici che seguono.

(1) Prop.: Il prodotto di due matrici ortogonali è una matrice ortogonale.

Dim.: Consideriamo le matrici $A, B \in \text{MatOrt}_d$; per il loro prodotto si ha $AB(AB)^\top = A(BB^\top)A^\top = A\hat{\mathbf{1}}_d A^\top = AA^\top = \hat{\mathbf{1}}_d$ e $(AB)^\top \cdot AB = B^\top(A^\top A)B = B^\top \hat{\mathbf{1}}_d B^\top = BB^\top = \hat{\mathbf{1}}_d$ ■

(2) Prop.: Se $A \in \text{MatOrt}_d$ e P è una matrice permutativa di ordine d , allora anche PAP^\top è ortogonale.

Dim.: Segue subito dalla (1) e dal fatto che P e P^\top sono matrici ortogonali ■

(3) Prop.: Se $A \in \mathbf{MatOrt}_d$ e $B \in \mathbf{MatOrt}_h$ per d ed h interi positivi, risulta ortogonale anche la matrice quadrata a blocchi avente A e B come sottomatrici diagonali

$$E = \begin{bmatrix} A & \mathbf{0}_{d,h} \\ \mathbf{0}_{h,d} & B \end{bmatrix} \in \mathbf{Mat}_{d+h} .$$

Dim.: Basta considerare le espressioni a blocchi dei prodotti costituenti le condizioni di ortogonalità.

$$E \cdot E^T = \begin{bmatrix} A & \mathbf{0}_{d,h} \\ \mathbf{0}_{h,d} & B \end{bmatrix} \cdot \begin{bmatrix} A^T & \mathbf{0}_{h,d} \\ \mathbf{0}_{d,h} & B^T \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{1}}_d & \mathbf{0}_{h,d} \\ \mathbf{0}_{d,h} & \hat{\mathbf{1}}_h \end{bmatrix} = \hat{\mathbf{1}}_{d+h}$$

$$E^T E = \begin{bmatrix} A^T & \mathbf{0}_{h,d} \\ \mathbf{0}_{d,h} & B^T \end{bmatrix} \cdot \begin{bmatrix} A & \mathbf{0}_{d,h} \\ \mathbf{0}_{h,d} & B \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{1}}_d & \mathbf{0}_{h,d} \\ \mathbf{0}_{d,h} & \hat{\mathbf{1}}_h \end{bmatrix} = \hat{\mathbf{1}}_{d+h} \quad \blacksquare$$

G47:e.03 Riformuliamo le uguaglianze che esprimono le condizioni di ortogonalità mediante i vettori forniti dalle righe e dalle colonne della matrice.

$$A \cdot A^T = \begin{bmatrix} A_{1,*} \\ \vdots \\ A_{d,*} \end{bmatrix} \cdot [A_{1,*}^T \quad \dots \quad A_{d,*}^T] = [A_{i,*} \cdot A_{j,*} \mid i, j = 1, 2, \dots, d] = \hat{\mathbf{1}}_d$$

$$A^T \cdot A = \begin{bmatrix} A_{1,*} \\ \vdots \\ A_{d,*} \end{bmatrix} \cdot [A_{1,*}^T \quad \dots \quad A_{d,*}^T] = [A_{i,*} \cdot A_{j,*} \mid i, j = 1, 2, \dots, d] = \hat{\mathbf{1}}_d$$

Queste uguaglianze sono in effetti del tutto equivalenti alle seguenti:

$$\forall j, k = 1, 2, \dots, d : \sum_{i=1}^d a_{i,j} a_{i,k} = \delta_{j,k}$$

$$\forall i, h = 1, 2, \dots, d : \sum_{j=1}^d a_{i,j} a_{h,j} = \delta_{i,h}$$

Le prime dicono che le colonne di A costituiscono un sistema di vettori ortonormali per lo spazio $\mathbb{F}^{d \times d}$ munito del prodotto interno canonico; le seconde dicono che anche le righe di A costituiscono un sistema di vettori ortonormali per tale spazio con prodotto interno.

G47:e.04 Si dice **matrice normale** sui reali ogni matrice quadrata sui reali N che commuta con la propria trasposta, cioè tale che

$$N \cdot N^T = N^T \cdot N .$$

(1) Prop.: Le matrici simmetriche, le antisimmetriche e le ortogonali sono matrici normali.

Dim.: Se $S^T = S$, allora $[S, S^T] = [S, S] = \mathbf{0}$.

Se $A^T = -A$, allora $[A, A^T] = [A, -A] = -[A, A] = -\mathbf{0}$.

Se $A^T = A^{-1}$, allora $[A, A^T] = AA^{-1} - A^{-1}A = \hat{\mathbf{1}} - \hat{\mathbf{1}} = \mathbf{0}$. \blacksquare

Vi sono però anche matrici normali che non sono nè simmetriche, nè antisimmetriche, nè ortogonali; tra le matrici di questo tipo vi sono

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} , \quad \begin{bmatrix} \alpha & \beta & 0 \\ 0 & \alpha & \beta \\ \beta & 0 & \alpha \end{bmatrix} \quad \forall \alpha, \beta \in \mathbb{R} .$$

In genere la somma e il prodotto di due matrici normali non è normale. Si trova tuttavia questo fatto.

(2) Prop.: Se A e B sono due matrici normali di ordine d che commutano, allora $A + B$ e $A \cdot B$ sono matrici normali.

Dim.: ■

Le varie componenti di questo testo sono accessibili in <http://www.mi.imati.cnr.it/~alberto>