

AUTOVALORI E AUTOVETTORI

DISPENSA PER IL CORSO DI ALGEBRA LINEARE

ENRICO GREGORIO

1. INTRODUZIONE

La nozione di autovalore di una matrice quadrata nasce insieme al concetto stesso di matrice o, meglio, al concetto di invariante di una matrice.

Ci sono molti problemi che conducono in modo naturale al concetto di autovalore: ad esempio, ci si può chiedere se una applicazione lineare $f: \mathbb{C}^n \rightarrow \mathbb{C}^n$ ha, rispetto ad qualche base ordinata, una forma particolarmente semplice; la forma più semplice possibile, dal punto di vista della moltiplicazione, che sappiamo corrispondere alla composizione di applicazioni, è sicuramente quella *diagonale*. Infatti, se $A = \text{diag}\{a_1, a_2, \dots, a_n\}$, si ha

$$A^k = \text{diag}\{a_1^k, a_2^k, \dots, a_n^k\}$$

e quindi è molto facile calcolare le *potenze* di f .

Storicamente, il concetto di autovalore è nato in modo un po' curioso: si stava cercando di calcolare le perturbazioni dell'orbita di Urano, che si pensava fossero dovute alla presenza di un altro pianeta, fino a quel momento sconosciuto. I calcoli portarono ad una equazione, che fu detta *equazione secolare* e poi si scoprì che le soluzioni di questa equazione erano legate strettamente ad una matrice. Si scoprì anche Nettuno!

La definizione che si dà oggi è ovviamente diversa.

1.1. Definizione. Sia $f: V \rightarrow V$ un'applicazione lineare dello spazio vettoriale V in sé. Un sottospazio U di V si dice *f-invariante* se

$$f(\mathbf{u}) \in U, \quad \text{per ogni } \mathbf{u} \in U.$$

Di particolare interesse sono i sottospazi f -invarianti di dimensione 1: infatti se $U = \langle \mathbf{u} \rangle$, con $\mathbf{u} \neq \mathbf{0}$, U è f -invariante se e solo se $f(\mathbf{u}) \in \langle \mathbf{u} \rangle$, cioè se e solo se esiste uno scalare λ tale che $f(\mathbf{u}) = \lambda \mathbf{u}$.

Un vettore che abbia questa proprietà si dirà un *autovettore* di f , mentre λ si dirà un *autovalore* di f .

2. AUTOVALORI ED AUTOVETTORI DI MATRICI

Il caso delle matrici è di gran lunga il più importante, diamo nuovamente la definizione in questo caso: come al solito identifichiamo la matrice A con l'applicazione lineare f_A .

2.1. Definizione. Sia A una matrice quadrata $n \times n$. Un vettore $\mathbf{u} \in \mathbb{C}^n$ si dice un *autovettore* della matrice A se:

- $\mathbf{u} \neq \mathbf{0}$ e
- esiste uno scalare $\lambda \in \mathbb{C}$ tale che $A\mathbf{u} = \lambda\mathbf{u}$.

Uno scalare λ si dice un *autovalore* di A se esiste un vettore $\mathbf{u} \in \mathbb{C}^n$ tale che $\mathbf{u} \neq \mathbf{0}$ e $A\mathbf{u} = \lambda\mathbf{u}$. In tal caso si dice che \mathbf{u} è un *autovettore di A relativo all'autovalore λ* .

Il fatto che \mathbf{u} sia un autovettore di A relativo a λ significa che \mathbf{u} è una soluzione non nulla del sistema omogeneo

$$(A - \lambda I_n)\mathbf{x} = \mathbf{0}$$

e perciò che la matrice $A - \lambda I_n$ è singolare. Viceversa, se la matrice $A - \lambda I_n$ è singolare, esiste una soluzione non nulla del sistema e quindi un autovettore di A relativo a λ .

2.2. Teorema. Sia A una matrice $n \times n$. Uno scalare $\lambda \in \mathbb{C}$ è un autovalore di A se e solo se $\det(A - \lambda I_n) = 0$.

Dimostrazione. Sappiamo già che una matrice quadrata B è singolare se e solo se $\det B = 0$. \square

Come esempio, proviamo a calcolare gli autovalori di

$$A = \begin{bmatrix} 1 & -2 \\ -3 & 2 \end{bmatrix}.$$

Se λ è un autovalore, devo avere $\det(A - \lambda I_2) = 0$, cioè

$$0 = \det \begin{bmatrix} 1 - \lambda & -2 \\ -3 & 2 - \lambda \end{bmatrix} = (1 - \lambda)(2 - \lambda) - 1 = 2 - 2\lambda - \lambda + \lambda^2 - 6 = \lambda^2 - 3\lambda - 4$$

e quindi gli unici valori possibili per λ sono -1 e 4 . Di fatto questi sono autovalori, per quanto abbiamo detto prima.

Cerchiamo un autovettore relativo all'autovalore 4 , cioè una soluzione non nulla del sistema lineare $(A - 4I_2)\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} -3 & -2 \\ -3 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2/3 \\ 0 & 0 \end{bmatrix}$$

e quindi un autovettore è $[-2/3 \ 1]^T$. Cerchiamo un autovettore relativo a -1 , cioè una soluzione non nulla del sistema lineare $(A + I_2)\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$$

e quindi un autovettore è $[1 \ 1]^T$.

La discussione precedente mostra che è possibile trovare gli autovalori di una matrice calcolando le radici di un polinomio: questo polinomio è $\det(A - tI)$.

2.3. Definizione. Il polinomio $p_A(t) = \det(A - tI)$ si chiama *polinomio caratteristico* di A .

Scritto per esteso il polinomio è:

$$p_A(t) = \det \begin{bmatrix} a_{11} - t & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} - t & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} - t & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} - t \end{bmatrix}.$$

2.4. Proposizione. Il polinomio caratteristico $p_A(t)$ della matrice quadrata A di ordine n ha le seguenti proprietà:

- (1) $p_A(t)$ ha grado n e il coefficiente di t^n è $(-1)^n$;
- (2) il coefficiente di t^{n-1} è $(-1)^{n-1}(\sum_i a_{ii})$;
- (3) il termine noto è $\det A$.

Dimostrazione. La terza affermazione è ovvia: basta calcolare in 0 il polinomio: il termine noto è proprio $p_A(0) = \det(A - 0I) = \det A$.

Dimostriamo le altre due affermazioni per induzione su n . Il caso di $n = 1$ è ovvio.

Supponiamo l'asserto vero per tutte le matrici di ordine $n - 1$. Sviluppiamo $\det(A - tI)$ rispetto all'ultima riga: allora

$$p_A(t) = a_{n1}B_{n1} + a_{n2}B_{n2} + \dots + a_{n-1,n}B_{n-1,n} + (a_{nn} - t)p_{A'}(t)$$

dove indichiamo con B_{ij} i cofattori della matrice $B = A - tI$ e A' è la matrice che si ottiene da A cancellando l'ultima riga e l'ultima colonna.

Se esaminiamo le matrici che si ottengono da B cancellando una riga ed una colonna, ci accorgiamo che al più un coefficiente di ciascuna riga contiene l'indeterminata t . È facile allora dimostrare che il determinante di queste matrici (dalle quali calcoliamo i cofattori B_{ij}), sono polinomi di grado al più $n - 1$ ed anzi, quelle che ci servono hanno una riga in cui non

comparare l'indeterminata e perciò i loro determinanti hanno grado al più $n - 2$. (esercizio; per induzione).

Ne segue che il coefficiente del termine di grado n in $p_A(t)$ è l'opposto del coefficiente del termine di grado massimo di $p_{A'}(t)$; per l'ipotesi induttiva, questo è $(-1)^{n-1}$.

Ancora per l'ipotesi induttiva, il coefficiente di t^{n-2} in $p_{A'}(t)$ è

$$(-1)^{n-2} \left(\sum_{i=1}^{n-1} a_{ii} \right)$$

ed è quindi facile verificare la tesi. □

La domanda che ci poniamo ora è: *è vero che ogni matrice quadrata ha almeno un autovalore?*

La risposta è: *sì*, in conseguenza del seguente fatto, che accettiamo senza dimostrazioni.

2.5. Teorema. *Ogni polinomio di grado positivo a coefficienti complessi ha almeno una radice.*

Questo teorema, la cui prima dimostrazione è dovuta a Gauss (costituiva la sua tesi di laurea!), giustifica il fatto che preferiamo lavorare con matrici a coefficienti complessi. È ben noto infatti che esistono polinomi a coefficienti reali che non ammettono radici *reali*.

Il noto teorema di Ruffini stabilisce che un numero complesso λ è radice del polinomio $p_A(t)$ se e solo se $p_A(t)$ è divisibile per $t - \lambda$. Poiché il quoziente è ancora un polinomio a coefficienti complessi, possiamo concludere che $p_A(t)$ è prodotto di polinomi di primo grado (in modo unico a meno dell'ordine, come si vedrà rigorosamente nel corso di Algebra). Se raggruppiamo i fattori di primo grado che hanno la stessa radice, otteniamo una decomposizione di $p_A(t)$ nella seguente forma:

$$p_A(t) = c(t - \lambda_1)^{m_1}(t - \lambda_2)^{m_2} \dots (t - \lambda_r)^{m_r}$$

dove c è una costante e $\lambda_1, \lambda_2, \dots, \lambda_r$ sono gli autovalori di A , a due a due distinti, e c è un numero complesso. Ma sappiamo calcolare c : è proprio il coefficiente di t^n , cioè $(-1)^n$. Perciò possiamo scrivere il polinomio caratteristico di A nella seguente forma:

$$p_A(t) = (\lambda_1 - t)^{m_1}(\lambda_2 - t)^{m_2} \dots (\lambda_r - t)^{m_r}$$

che si chiama la *fattorizzazione primaria*. Il numero intero m_i si chiama *molteplicità algebrica* dell'autovalore λ_i . Siccome $p_A(t)$ ha grado n , si ha, evidentemente, $m_1 + m_2 + \dots + m_r = n$. Anche la seguente proposizione segue immediatamente dalla decomposizione primaria.

2.6. Proposizione. *Sia A una matrice $n \times n$ e sia λ un suo autovalore. Allora la molteplicità algebrica di λ è il massimo intero n tale che $(\lambda - t)^n$ divida $p_A(t)$.*

Dalla fattorizzazione primaria di $p_A(t)$ e da facili proprietà dei polinomi si ottiene la seguente proposizione.

2.7. Proposizione. *Sia A una matrice quadrata di ordine n e siano $\lambda_1, \lambda_2, \dots, \lambda_r$ i suoi autovalori distinti, di molteplicità algebrica rispettiva m_1, m_2, \dots, m_r . Allora:*

(1) $m_1 + m_2 + \dots + m_r = n$;

(2) *il coefficiente del termine di grado $n - 1$ in $p_A(t)$ è*

$$(-1)^{n-1}(m_1\lambda_1 + m_2\lambda_2 + \dots + m_r\lambda_r);$$

(3) *il termine noto di $p_A(t)$ è $\lambda_1^{m_1}\lambda_2^{m_2} \dots \lambda_r^{m_r}$.*

Diamo ora alcuni esempi. Prima di tutto osserviamo che A e A^T hanno lo stesso polinomio caratteristico.

2.8. Esempio. *Sia $A = [a_{ij}]$ una matrice triangolare. Allora i suoi autovalori sono i coefficienti sulla diagonale.*

Basta infatti osservare che anche $A - tI$ è triangolare; se è triangolare superiore, si sviluppa il determinante rispetto alla prima colonna. Allora, per induzione, $\det(A - tI) = (a_{11} - t)(a_{22} - t) \dots (a_{nn} - t)$.

2.9. Esempio. Se λ è un autovalore di A e k è un intero positivo, allora λ^k è un autovalore di A^k .

Dimostriamo che, se \mathbf{v} è un autovettore di A relativo a λ , allora \mathbf{v} è un autovettore di A^k relativo a λ^k , cioè che

$$A^k \mathbf{v} = \lambda^k \mathbf{v}.$$

La dimostrazione è per induzione su k ; il caso $k = 1$ è ovvio. Supponiamo vera l'affermazione per $k - 1$. Allora $A^{k-1} \mathbf{v} = \lambda^{k-1} \mathbf{v}$. Quindi

$$A^k \mathbf{v} = A(A^{k-1} \mathbf{v}) = A(\lambda^{k-1} \mathbf{v}) = \lambda^{k-1} (A \mathbf{v}) = \lambda^{k-1} \lambda \mathbf{v} = \lambda^k \mathbf{v}.$$

2.10. Esempio. Se A è una matrice nilpotente, cioè se esiste un intero positivo m tale che $A^m = \mathbb{O}$, allora l'unico autovalore di A è 0 .

È chiaro che l'unico autovalore di \mathbb{O} è 0 . Perciò, per l'esempio precedente, se λ è un autovalore di A , si ha $\lambda^m = 0$.

2.11. Esempio. Esistono matrici reali 2×2 che non hanno autovalori reali. Ogni matrice reale 3×3 ha almeno un autovalore reale.

Se A è una matrice reale, il suo polinomio caratteristico è a coefficienti reali. Se A è 3×3 , il polinomio caratteristico ha grado 3 e quindi, per il Teorema di Bolzano-Weierstrass, ha almeno una radice: la funzione reale $x \mapsto p_A(x)$ assume valori positivi e negativi ed è continua; quindi il suo grafico interseca l'asse delle x .

Se invece consideriamo la matrice

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

il polinomio caratteristico è $p_A(t) = 1 + t^2$, che non ha radici reali, ma solo le due radici complesse i e $-i$.

3. AUTOSPAZI

Sia $\lambda \in \mathbb{C}$ e sia A una matrice complessa $n \times n$; definiamo

$$E_A(\lambda) = \{ \mathbf{v} \in \mathbb{C}^n \mid A \mathbf{v} = \lambda \mathbf{v} \}.$$

In altre parole $E_A(\lambda) = N(A - \lambda I)$; perciò $E_A(\lambda)$ è un sottospazio vettoriale di \mathbb{C}^n , che è non nullo se e solo se λ è un autovalore di A .

Se λ è un autovalore di A , chiamiamo $E_A(\lambda)$ *autospatto di A relativo a λ* . In tal caso $E_A(\lambda)$ consiste di tutti gli autovettori di A relativi a λ ed in più del vettore nullo. La dimensione di $E_A(\lambda)$ si chiama *molteplicità geometrica dell'autovalore λ* .

3.1. Definizione. Siano U_1, U_2, \dots, U_k sottospazi dello spazio vettoriale V . Diremo che essi sono *indipendenti* se, dati $\mathbf{u}_1 \in U_1, \mathbf{u}_2 \in U_2, \dots, \mathbf{u}_k \in U_k$,

$$\text{da } \mathbf{u}_1 + \mathbf{u}_2 + \dots + \mathbf{u}_k = \mathbf{0} \text{ segue } \mathbf{u}_1 = \mathbf{u}_2 = \dots = \mathbf{u}_k = \mathbf{0}.$$

3.2. Esercizio. Siano $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ vettori linearmente indipendenti in V ; allora i sottospazi $\langle \mathbf{v}_1 \rangle, \langle \mathbf{v}_2 \rangle, \dots, \langle \mathbf{v}_k \rangle$ sono indipendenti.

3.3. Teorema. Sia A una matrice $n \times n$ e siano $\lambda_1, \lambda_2, \dots, \lambda_r$ i suoi autovalori distinti. Allora gli autospazi $E_A(\lambda_1), E_A(\lambda_2), \dots, E_A(\lambda_r)$ sono indipendenti.

Dimostrazione. Sia $\mathbf{v}_i \in E_A(\lambda_i)$ ($i = 1, 2, \dots, r$) e supponiamo che

$$\mathbf{v}_1 + \mathbf{v}_2 + \dots + \mathbf{v}_r = \mathbf{0}.$$

Procediamo per assurdo: se esiste una scelta dei vettori \mathbf{v}_i non tutti nulli, ce n'è una con il minimo numero di vettori non nulli. Possiamo supporre che siano i primi s , eventualmente cambiando la numerazione degli autovalori; allora l'ipotesi è che

$$(*) \quad \mathbf{v}_1 + \mathbf{v}_2 + \dots + \mathbf{v}_s = \mathbf{0}$$

ma che le somme di $s - 1$ vettori non nulli, presi ciascuno in un autospatto differente, siano non nulle.

Moltiplichiamo la precedente identità per A : otteniamo

$$A\mathbf{v}_1 + A\mathbf{v}_2 + \cdots + A\mathbf{v}_s = \mathbf{0}$$

e quindi, ricordando la definizione di autospazio,

$$\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_s\mathbf{v}_s = \mathbf{0}.$$

Se invece moltiplichiamo l'identità (*) per λ_s , otteniamo

$$\lambda_s\mathbf{v}_1 + \lambda_s\mathbf{v}_2 + \cdots + \lambda_s\mathbf{v}_s = \mathbf{0}.$$

Se allora eseguiamo la sottrazione, abbiamo

$$(\lambda_s - \lambda_1)\mathbf{v}_1 + (\lambda_s - \lambda_2)\mathbf{v}_2 + \cdots + (\lambda_s - \lambda_{s-1})\mathbf{v}_{s-1} = \mathbf{0}.$$

In altre parole, siamo riusciti ad ottenere una identità dello stesso tipo, con un vettore di meno. Notiamo anche che, ponendo $\mathbf{v}'_i = (\lambda_s - \lambda_i)\mathbf{v}_i$ ($i = 1, 2, \dots, s-1$), abbiamo ancora che $\mathbf{v}'_i \in E_A(\lambda_i)$. Dall'ipotesi che abbiamo fatto segue allora che $\mathbf{v}'_i = \mathbf{0}$. Ma, essendo $\lambda_s - \lambda_i \neq 0$, questo significa $\mathbf{v}_i = \mathbf{0}$ ($i = 1, 2, \dots, s-1$). Ma allora anche $\mathbf{v}_s = \mathbf{0}$: assurdo. \square

3.4. Corollario. *Autovettori relativi ad autovalori distinti sono linearmente indipendenti.*

Dimostrazione. Siano $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$, autovettori della matrice A relativi agli autovalori (distinti) $\lambda_1, \lambda_2, \dots, \lambda_k$. Se

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k = \mathbf{0},$$

per la proposizione precedente è $\alpha_i\mathbf{v}_i = \mathbf{0}$, per $i = 1, 2, \dots, k$. Ma allora, essendo $\mathbf{v}_i \neq \mathbf{0}$, deve essere $\alpha_i = 0$ ($i = 1, 2, \dots, k$). \square

Data una matrice A ed un suo autovalore λ , denoteremo con $m_A(\lambda)$ e con $d_A(\lambda)$ rispettivamente la molteplicità algebrica e la molteplicità geometrica di λ .

3.5. Teorema. *Sia A una matrice $n \times n$ e sia λ un suo autovalore. Allora*

$$1 \leq d_A(\lambda) \leq m_A(\lambda).$$

Dimostrazione. Per semplicità scriviamo $m = m_A(\lambda)$ e $d = d_A(\lambda)$. Fissiamo una base ordinata $\{\mathbf{x}_1; \mathbf{x}_2; \dots; \mathbf{x}_d\}$ di $E_A(\lambda)$ e la completiamo ad una base ordinata \mathcal{B} di \mathbb{C}^n . Allora la matrice B associata all'applicazione lineare $f_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$ rispetto alla base \mathcal{B} su dominio e codominio è

$$\begin{aligned} B &= [C_{\mathcal{B}}(A\mathbf{x}_1) \ C_{\mathcal{B}}(A\mathbf{x}_2) \ \dots \ C_{\mathcal{B}}(A\mathbf{x}_d) \ \dots] = \\ &= [C_{\mathcal{B}}(\lambda\mathbf{x}_1) \ C_{\mathcal{B}}(\lambda\mathbf{x}_2) \ \dots \ C_{\mathcal{B}}(\lambda\mathbf{x}_d) \ \dots] = \\ &= \begin{bmatrix} \lambda I_d & C \\ \mathbb{O} & D \end{bmatrix} \end{aligned}$$

per opportune matrici C e D . È facile, sviluppando il determinante rispetto alla prima colonna d volte, verificare che $p_B(t) = (\lambda - t)^d g(t)$, dove $g(t)$ è un opportuno polinomio di grado $n - d$. Poiché però sappiamo che esiste una matrice S tale che $B = SAS^{-1}$, abbiamo

$$\begin{aligned} \det(B - tI) &= \det(SAS^{-1} - tI) = \det(SAS^{-1} - S(tI)S^{-1}) = \\ &= \det(S(A - tI)S^{-1}) = (\det S)(\det(A - tI))(\det S)^{-1} = p_A(t) \end{aligned}$$

e quindi $p_A(t)$ è divisibile per $(\lambda - t)^d$. Dunque $m \geq d$. \square

4. MATRICI SIMILI

La tecnica usata per finire la dimostrazione dell'ultimo teorema suggerisce una definizione.

4.1. Definizione. Due matrici A e B di ordine n si dicono *simili* se esiste una matrice non-singolare S tale che $B = SAS^{-1}$.

4.2. Esercizio. Dimostrare che la similitudine di matrici è una relazione di equivalenza.

La seguente proposizione allora si dimostra esattamente con la stessa tecnica.

4.3. Proposizione. Se A e B sono matrici simili, allora $p_A(t) = p_B(t)$. Di conseguenza A e B hanno gli stessi autovalori con la stessa molteplicità algebrica.

Siano ora A e B matrici simili, e sia λ un autovalore (di entrambe). Fissiamo una matrice non-singolare S tale che $B = SAS^{-1}$. Se $\mathbf{v} \in E_A(\lambda)$, poniamo $f(\mathbf{v}) = S^{-1}\mathbf{v}$. Allora

$$Bf(\mathbf{v}) = BS^{-1}\mathbf{v} = S^{-1}SBS^{-1}\mathbf{v} = S^{-1}A\mathbf{v} = S^{-1}\lambda\mathbf{v} = \lambda(S^{-1}\mathbf{v}) = \lambda f(\mathbf{v})$$

e perciò $f(\mathbf{v}) \in E_B(\lambda)$. In altre parole abbiamo definito un'applicazione (che è ovviamente lineare) $f: E_A(\lambda) \rightarrow E_B(\lambda)$. È chiaro, con un calcolo analogo, che possiamo definire anche $g: E_B(\lambda) \rightarrow E_A(\lambda)$ ponendo, per $\mathbf{w} \in E_B(\lambda)$, $g(\mathbf{w}) = S\mathbf{w}$.

È allora immediato che la composizione $g \circ f$ è l'identità su $E_A(\lambda)$ e che la composizione $f \circ g$ è l'identità su $E_B(\lambda)$. Quindi i due spazi vettoriali $E_A(\lambda)$ e $E_B(\lambda)$ sono isomorfi e perciò hanno la stessa dimensione. Abbiamo allora dimostrato quanto riassunto nel teorema seguente.

4.4. Teorema. Siano A e B matrici simili. Allora A e B hanno gli stessi autovalori con la stessa molteplicità algebrica e la stessa molteplicità geometrica.

5. MATRICI DIAGONALIZZABILI

Una matrice quadrata A si dice *diagonalizzabile* se è simile ad una matrice diagonale, cioè se esistono una matrice non singolare S ed una matrice diagonale D tali che $D = SAS^{-1}$ (o, è lo stesso, $D = S^{-1}AS$).

In molte applicazioni è utile sapere che una matrice è diagonalizzabile. Ad esempio, se A lo è, possiamo calcolare le potenze di A in modo facile:

$$A^k = SD^kS^{-1}$$

(lo si dimostri per induzione); se $D = \text{diag}(\delta_1, \delta_2, \dots, \delta_n)$, si ha che $D^k = \text{diag}(\delta_1^k, \delta_2^k, \dots, \delta_n^k)$.

5.1. Teorema. Sia A una matrice $n \times n$ a coefficienti complessi e supponiamo che la decomposizione primaria del suo polinomio caratteristico sia

$$p_A(t) = (\lambda_1 - t)^{m_1}(\lambda_2 - t)^{m_2} \dots (\lambda_r - t)^{m_r}.$$

Indichiamo con m_i e con d_i rispettivamente la molteplicità algebrica e geometrica dell'autovalore λ_i . Sono equivalenti le affermazioni seguenti:

- (1) A è diagonalizzabile;
- (2) \mathbb{C}^n ha una base formata di autovettori di A ;
- (3) per ogni autovalore λ_i , la molteplicità geometrica coincide con la molteplicità algebrica: $d_i = m_i$ ($i = 1, 2, \dots, r$).

Dimostrazione. (1) \implies (2) Sia $A = SDS^{-1}$, dove $S = [\mathbf{v}_1 \dots \mathbf{v}_n]$ e $D = \text{diag}(\delta_1, \dots, \delta_n)$. Allora abbiamo $AS = SD$ e questo dice che $A\mathbf{v}_i = \delta_i\mathbf{v}_i$. Allora le colonne di S sono autovettori di A e formano una base perché S è non-singolare, quindi ha rango n .

(2) \implies (1) Sia $\mathcal{B} = \{\mathbf{v}_1; \dots; \mathbf{v}_n\}$ una base di \mathbb{C}^n formata da autovettori di A . Allora $A\mathbf{v}_i = \delta\mathbf{v}_i$ ($i = 1, \dots, n$) e quindi la matrice associata a f_A rispetto alla base \mathcal{B} su dominio e codominio è $D = \text{diag}(\delta_1, \dots, \delta_n)$. Ma allora A e D sono simili.

(2) \implies (3) Raggruppiamo i vettori della base \mathcal{B} formata da autovettori secondo gli autovalori. Supponiamo che il numero degli autovettori della base relativi all'autovalore λ_i sia k_i ($i = 1, \dots, r$). Allora questi vettori sono linearmente indipendenti e stanno in $E_A(\lambda_i)$, perciò $k_i \leq d_i$.

Allora abbiamo

$$n = \sum_{i=1}^r k_i \leq \sum_{i=1}^r d_i \leq \sum_{i=1}^r m_i = n$$

e, essendo $k_i \leq d_i \leq m_i$, per ogni i , si ha

$$d_i = m_i \quad (i = 1, 2, \dots, r).$$

(3) \implies (2) Si fissi una base \mathcal{B}_i di ciascun autospazio $E_A(\lambda_i)$. Se poniamo

$$\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_r$$

otteniamo un insieme linearmente indipendente per il Teorema 3.3, che è formato da

$$\sum_{i=1}^r d_i = \sum_{i=1}^r m_i = n$$

vettori. Quindi \mathcal{B} è una base formata da autovettori di A . □

5.2. Osservazione. Notiamo che se A è diagonalizzabile, ogni matrice diagonale a cui A è simile dovrà avere sulla diagonale gli autovalori di A , con le loro molteplicità.

5.3. Corollario. *Se la matrice A di ordine n ha n autovalori distinti, allora A è diagonalizzabile.*

Dimostrazione. Se $\lambda_1, \lambda_2, \dots, \lambda_n$ sono gli autovalori, abbiamo, per ogni i ,

$$1 \leq d_A(\lambda_i) \leq m_A(\lambda_i) = 1$$

e quindi il teorema permette di concludere. □

5.4. Osservazione. Attenzione: il corollario appena dimostrato dà una condizione *sufficiente* per la diagonalizzabilità. Basta infatti pensare alla matrice identità, che ha un solo autovalore (che è 1), e che è diagonalizzabile, essendo diagonale!

Possiamo dare un teorema analogo per matrici a coefficienti reali. Se A è a coefficienti reali e λ è un autovalore *reale* di A , allora esiste un autovettore di A relativo a λ a coordinate reali. Infatti, se $\mathbf{v} \in \mathbb{C}^n$ ha la proprietà che $\mathbf{v} \neq \mathbf{0}$ e che $A\mathbf{v} = \lambda\mathbf{v}$, possiamo scrivere $\mathbf{v} = \mathbf{a} + i\mathbf{b}$, dove \mathbf{a} e \mathbf{b} hanno coordinate reali. Ora, uguagliando parte reale e parte immaginaria di ogni componente, dobbiamo avere $A\mathbf{a} = \lambda\mathbf{a}$ e $A\mathbf{b} = \lambda\mathbf{b}$; inoltre, almeno uno fra \mathbf{a} e \mathbf{b} deve essere non nullo e quindi abbiamo trovato un autovettore di A relativo a λ a coordinate reali.

Possiamo allora considerare $E_A(\lambda; \mathbb{R}) = \{ \mathbf{v} \in \mathbb{R}^n : A\mathbf{v} = \lambda\mathbf{v} \}$, che è un sottospazio di \mathbb{R}^n . La sua dimensione si può chiamare *molteplicità geometrica reale* dell'autovalore reale λ .

Diremo che A è *diagonalizzabile con una matrice reale* se esistono una matrice S reale non-singolare ed una matrice diagonale reale D tali che $A = SDS^{-1}$. È chiaro dal teorema precedente che una condizione necessaria affinché A sia diagonalizzabile con una matrice reale è che tutti i suoi autovalori siano reali.

5.5. Teorema. *Sia A una matrice $n \times n$ a coefficienti reali e supponiamo che la decomposizione primaria del suo polinomio caratteristico sia*

$$p_A(t) = (\lambda_1 - t)^{m_1} (\lambda_2 - t)^{m_2} \dots (\lambda_r - t)^{m_r},$$

con $\lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{R}$. Indichiamo con m_i e con d_i^g rispettivamente la molteplicità algebrica e geometrica reale dell'autovalore λ_i . Sono equivalenti le affermazioni seguenti:

- (1) A è diagonalizzabile con una matrice reale;
- (2) \mathbb{R}^n ha una base formata di autovettori di A ;
- (3) per ogni autovalore λ_i , la molteplicità geometrica reale coincide con la molteplicità algebrica $d_i^g = m_i$ ($i = 1, 2, \dots, r$).

5.6. Esempio. Consideriamo le matrici

$$A = \begin{bmatrix} i & 1+i \\ 0 & 1+i \end{bmatrix}, \quad B = \begin{bmatrix} \pi & 1 \\ -1 & \pi-2 \end{bmatrix}.$$

Calcoliamo gli autovalori di A :

$$p_A(t) = \det(A - tI_2) = \det \begin{bmatrix} i-t & 1+i \\ 0 & 1+i-t \end{bmatrix} = (i-t)(1+i-t).$$

Perciò A è diagonalizzabile, avendo due autovalori distinti. Per trovare una base di \mathbb{C}^2 formata da autovettori, calcoliamo gli autospazi.

Calcolo di $E_A(i) = N(A - iI_2)$: dobbiamo risolvere il sistema omogeneo $(A - iI_2)\mathbf{x} = \mathbf{0}$; perciò

$$\begin{bmatrix} 0 & 1+i \\ 0 & 1 \end{bmatrix} \xrightarrow{E_2((i+1)^{-1})} \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \xrightarrow{E_{21}(-1)} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

e quindi un autovettore in $E_A(i)$ è $[1 \ 0]^T$.

Calcolo di $E_A(1+i) = N(A - (1+i)I_2)$: il sistema è $(A - (1+i)I_2)\mathbf{x} = \mathbf{0}$; perciò

$$\begin{bmatrix} -1 & 1+i \\ 0 & 0 \end{bmatrix} \xrightarrow{E_1(-1)} \begin{bmatrix} 1 & -1-i \\ 0 & 0 \end{bmatrix}$$

e quindi un autovettore in $E_A(1+i)$ è $[1+i \ 1]^T$.

Con facili calcoli, $p_B(t) = \det(B - tI_2) = (t - \pi + 1)^2$ e quindi B ha un solo autovalore (che è $\pi - 1$) di molteplicità algebrica 2.

Calcolo di $E_B(\pi - 1) = N(B - (\pi - 1)I_2)$: il sistema è $(B - (\pi - 1)I_2)\mathbf{x} = \mathbf{0}$; perciò

$$\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \xrightarrow{E_{12}(1)} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix};$$

poiché la matrice ha rango 1, l'autospazio ha dimensione 1 e quindi la matrice B non è diagonalizzabile.

6. MATRICI UNITARIAMENTE SIMILI

Diamo ora un'altra definizione: due matrici quadrate A e B di ordine n si dicono *unitariamente simili* se esiste una matrice unitaria U tale che

$$B = U^H A U.$$

Ricordiamo che U è unitaria se e solo se $U^H = U^{-1}$, cioè se e solo se le colonne di U formano una base ortonormale di \mathbb{C}^n .

È evidente che due matrici unitariamente simili sono anche simili. Non è vero il viceversa, come si capirà più avanti.

Diremo che A è *diagonalizzabile con una matrice unitaria* se A è unitariamente simile ad una matrice diagonale.

Per matrici reali, l'aggettivo "unitario" viene sostituito con "ortogonale"; perciò si parlerà di matrici *reali* ortogonalmente simili o diagonalizzabili con una matrice ortogonale.

6.1. Proposizione. *Sia A una matrice diagonalizzabile con una matrice unitaria. Allora $A^H A = A A^H$.*

Dimostrazione. Sia $A = U D U^H$, con U unitaria e $D = \text{diag}(\delta_1, \dots, \delta_n)$. Allora

$$A A^H = U D U^H (U D U^H)^H = U D U^H U D^H U^H = U D D^H U^H$$

$$A^H A = (U D U^H)^H U D U^H = U D^H U^H U D U^H = U D^H D U^H$$

e, poiché $D D^H = \text{diag}(|\delta_1|^2, \dots, |\delta_n|^2) = D^H D$, abbiamo l'asserto. \square

Diciamo che una matrice A è *normale* se $A^H A = A A^H$. Ad esempio, la matrice A dell'esempio 5.6 non è normale, quindi non è diagonalizzabile con una matrice unitaria, pur essendo diagonalizzabile. Infatti

$$A^H A = \begin{bmatrix} -i & 0 \\ 1-i & 1-i \end{bmatrix} \begin{bmatrix} i & 1+i \\ 0 & 1+i \end{bmatrix} = \begin{bmatrix} 1 & 1-i \\ 1+i & 4 \end{bmatrix}$$

mentre

$$A A^H = \begin{bmatrix} i & 1+i \\ 0 & 1+i \end{bmatrix} \begin{bmatrix} -i & 0 \\ 1-i & 1-i \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix}.$$

6.2. Definizione. Una matrice quadrata A si dice *normale* se $A^H A = A A^H$.

Il termine “normale” non tragga in inganno: le matrici normali, come vedremo, sono una classe molto particolare. L'aggettivo deriva da “norma”, intesa come “lunghezza” di un vettore.

6.3. Esempio. Sia A una matrice reale 2×2 . Allora A è normale se e solo se A è simmetrica oppure $A = rU$, con U ortogonale e $r \in \mathbb{R}$.

Se infatti scriviamo $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, la condizione di normalità diventa:

$$a^2 + b^2 = a^2 + c^2$$

$$b^2 + d^2 = c^2 + d^2$$

$$ab + cd = ac + bd$$

e quindi, in particolare, $b^2 = c^2$. Nel caso $b = c$, abbiamo che A è simmetrica. Se invece $c = -b$, la terza equazione diventa $b(a - d) = 0$. Se $b = 0$, A è simmetrica. Rimane allora il caso $b \neq 0$ e $a = d$. Allora la matrice è

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} = r \begin{bmatrix} a' & b' \\ -b' & a' \end{bmatrix} = rU$$

dove $r = \sqrt{a^2 + b^2}$, $a' = a/r$ e $b' = b/r$. La matrice U è ovviamente ortogonale.

Per esaminare le matrici normali è utile il seguente teorema dovuto a Schur, che enunciamo separatamente per il caso complesso e quello reale.

6.4. Teorema di Schur (Caso complesso). *Sia A una matrice quadrata complessa. Allora A è unitariamente simile ad una matrice triangolare superiore.*

6.5. Teorema di Schur (Caso reale). *Sia A una matrice quadrata reale. Allora A è ortogonalmente simile ad una matrice triangolare superiore se e solo se tutti gli autovalori di A sono reali.*

Dimostrazione. È chiaro che la condizione che gli autovalori siano reali è necessaria nel caso reale, poiché matrici simili hanno gli stessi autovalori e gli autovalori di una matrice triangolare sono i coefficienti sulla diagonale. La dimostrazione può allora procedere allo stesso modo per i due casi (nel caso reale si suppone dunque che gli autovalori siano reali e, ovviamente, che i vettori considerati siano in \mathbb{R}^n).

Faremo induzione sull'ordine n della matrice. Se $n = 1$ l'asserto è banale: ogni matrice è triangolare.

Supponiamo $n > 1$ e che l'asserto sia vero per $n - 1$. Fissiamo un autovalore λ_1 ed un autovettore \mathbf{x}_1 di norma 1. e completiamo ad una base ortonormale $\mathcal{B} = \{\mathbf{x}_1; \mathbf{x}_2; \dots; \mathbf{x}_n\}$. La matrice

$$U_1 = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \dots \quad \mathbf{x}_n]$$

è unitaria (ortogonale nel caso reale). Allora, tenendo conto che $A\mathbf{x}_1 = \lambda_1\mathbf{x}_1$, si ha

$$U_1^{-1} A U_1 = \begin{bmatrix} \lambda_1 & \mathbf{x}^T \\ \mathbf{0} & C \end{bmatrix} = A'$$

per un opportuno vettore $\mathbf{x} \in \mathbb{C}^{n-1}$ ($\mathbf{x} \in \mathbb{R}^{n-1}$ ed una opportuna matrice C di ordine $n - 1$). Allora A' ha gli stessi autovalori di A e quindi gli autovalori di C sono anche autovalori di A ; basta infatti calcolare il polinomio caratteristico di A' sviluppando rispetto alla prima colonna.

Per l'ipotesi induttiva, C è unitariamente (ortogonalmente) simile ad una matrice triangolare superiore T_1 , cioè esiste una matrice unitaria (ortogonale) V tale che $C = VTV^{-1}$. Poniamo

$$U_2 = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & V \end{bmatrix}.$$

Allora U_2 è unitaria (ortogonale) e quindi tale è anche $U = U_1U_2$. Calcoliamo $U^{-1}AU$:

$$\begin{aligned} U^{-1}AU &= U_2^{-1}U_1^{-1}AU_1U_2 = U_2^{-1} \begin{bmatrix} \lambda_1 & \mathbf{x}^T \\ \mathbf{0} & C \end{bmatrix} U_2 \\ &= \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & V^{-1} \end{bmatrix} \begin{bmatrix} \lambda_1 & \mathbf{x}^T \\ \mathbf{0} & C \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & V \end{bmatrix} = \begin{bmatrix} \lambda_1 & \mathbf{x}^T V \\ \mathbf{0} & T_1 \end{bmatrix} \end{aligned}$$

e l'ultima matrice è triangolare superiore, perché lo è T_1 . □

Il teorema di Schur ha una conseguenza importante, che non è possibile dimostrare direttamente.

6.6. Corollario. *Sia A una matrice con autovalori $\lambda_1, \lambda_2, \dots, \lambda_r$. Per ogni intero positivo k , gli autovalori di A^k sono $\lambda_1^k, \lambda_2^k, \dots, \lambda_r^k$.*

Dimostrazione. Sia $A = UTU^H$, con U unitaria e T triangolare. Allora $A^k = UT^kU^H$ e perciò gli autovalori di A^k sono gli stessi di T^k , che sono proprio le potenze dei coefficienti sulla diagonale di T , i quali, a loro volta, sono proprio gli autovalori di A , ripetuti con le loro molteplicità. □

Notiamo però che i numeri $\lambda_1^k, \lambda_2^k, \dots, \lambda_r^k$ non sono necessariamente distinti.

La conseguenza più importante del teorema di Schur è la caratterizzazione delle matrici diagonalizzabili con una matrice unitaria.

6.7. Lemma. *Sia T una matrice triangolare. Allora T è normale se e solo se T è diagonale.*

Dimostrazione. È chiaro che ogni matrice diagonale è normale.

Sia T triangolare e normale di ordine n ; non è restrittivo supporre che T sia triangolare superiore. Per $n = 1$ l'asserto è ovvio; supponiamolo allora vero per matrici di ordine $n - 1$.

Scriviamo

$$T = \begin{bmatrix} t_{11} & \mathbf{r}^H \\ \mathbf{0} & S \end{bmatrix}$$

dove S è una matrice triangolare superiore di ordine $n - 1$. Se eseguiamo i prodotti TT^H e T^HT e imponiamo la condizione di normalità, troviamo le seguenti uguaglianze:

$$\begin{aligned} |t_{11}|^2 + \mathbf{r}^H \mathbf{r} &= |t_{11}|^2, & \mathbf{r}^H S^H &= (t_{11} \mathbf{r}^H, \\ S \mathbf{r} &= t_{11} \mathbf{r}, & S S^H &= \mathbf{r} \mathbf{r}^H + S^H S. \end{aligned}$$

Ne possiamo ricavare che: $\mathbf{r}^H \mathbf{r} = 0$ e quindi che $\mathbf{r} = \mathbf{0}$; ma allora $S S^H = S^H S$, quindi che S è normale. Per l'ipotesi induttiva, S è diagonale e quindi lo è anche T . □

6.8. Teorema. *Una matrice è diagonalizzabile con una matrice unitaria se e solo se è normale.*

Dimostrazione. Abbiamo già visto la necessità. Supponiamo allora che A sia una matrice normale e dimostriamo che A è diagonalizzabile con una matrice unitaria.

Per il teorema di Schur, A è unitariamente simile ad una matrice triangolare, cioè esistono una matrice triangolare T ed una matrice unitaria U tali che $T = UAU^H$. Allora

$$\begin{aligned} TT^H &= UAU^H(UAU^H)^H = UAU^HUA^HUU^H = UAA^HU^H, \\ T^HT &= (UAU^H)^HUAU^HUA^HUU^HUAU^H = UA^HUU^HUAU^H. \end{aligned}$$

Ma allora T è normale e quindi diagonale. □

Con la stessa tecnica usata nel teorema sulla diagonalizzabilità delle matrici, si dimostra che una matrice A è diagonalizzabile con una matrice unitaria se e solo se esiste una base ortonormale di \mathbb{C}^n formata da autovettori di A . Usando il teorema precedente, possiamo concludere con il seguente corollario.

6.9. Corollario. *Una matrice A di ordine n è normale se e solo se esiste una base ortonormale di \mathbb{C}^n formata da autovettori di A .*

6.10. Esercizio. Sia A una matrice $n \times n$ con autovalori $\lambda_1, \lambda_2, \dots, \lambda_n$ elencati con le loro molteplicità algebriche. Si provi usando la forma triangolare di Schur che, se

$$\sum_{i,j=1}^n |a_{ij}|^2 = \sum_{i=1}^n |\lambda_i|^2,$$

allora A è normale.

6.11. Esercizio. Si dimostri che autovettori relativi ad autovalori distinti di una matrice normale sono ortogonali.

7. IL TEOREMA SPETTRALE

Le matrici normali ammettono un'altra descrizione.

7.1. Teorema spettrale. *Sia A una matrice normale, con autovalori distinti $\lambda_1, \lambda_2, \dots, \lambda_r$. Indichiamo con P_i la matrice di proiezione sull'autospazio $E_A(\lambda_i)$ ($i = 1, 2, \dots, r$). Allora*

$$A = \sum_{i=1}^r \lambda_i P_i$$

e $P_j P_k = \mathbb{O}$, per $j \neq k$.

Dimostrazione. Sappiamo che esiste una base ortonormale di \mathbb{C}^n (dove n è l'ordine di A) formata da autovettori di A . Raggruppiamo i vettori della base ponendo all'inizio quelli relativi a λ_1 , poi quelli relativi a λ_2 e così via. Se $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{d_1}$ sono quelli relativi a λ_1 , sappiamo che

$$P_1 = \sum_{i=1}^{d_1} \mathbf{x}_i \mathbf{x}_i^H$$

e analogamente per le altre matrici di proiezione. È allora facile vedere che $P_j P_k = \mathbb{O}$, tenendo conto dell'esercizio 6.11. Se U indica la matrice che ha come colonne gli elementi della base, abbiamo $A = U D U^H$, con $D = \text{diag}(\delta_1, \delta_2, \dots, \delta_n)$, avente sulla diagonale λ_1 ripetuto secondo la sua molteplicità, λ_2 ripetuto secondo la sua molteplicità e così via.

Ma allora $A = U D U^H$ equivale a

$$A = \sum_{i=1}^n \delta_i \mathbf{x}_i \mathbf{x}_i^H = \sum_{i=1}^{d_1} \lambda_1 \mathbf{x}_i \mathbf{x}_i^H + \sum_{i=d_1+1}^n \delta_i \mathbf{x}_i \mathbf{x}_i^H = \lambda_1 P_1 + \sum_{i=d_1+1}^n \delta_i \mathbf{x}_i \mathbf{x}_i^H$$

e, proseguendo un gruppo di autovalori alla volta, otteniamo la scrittura richiesta. \square

7.2. Esempio. Calcoliamo la *decomposizione spettrale* della matrice

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & -1 & 3 \end{bmatrix}.$$

Notiamo che A è normale in quanto simmetrica reale. Dobbiamo trovare una base ortonormale di \mathbb{C}^3 formata da autovettori di A .

Polinomio caratteristico di A :

$$p_A(t) = \det(A - tI) = \det \begin{bmatrix} 2-t & 0 & 0 \\ 0 & 3-t & -1 \\ 0 & -1 & 3-t \end{bmatrix} = (2-t)^2(4-t)$$

e perciò gli autovalori sono $\lambda_1 = 2$, $\lambda_2 = 4$.

Autospazio $E_A(2)$:

$$A - 2I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \xrightarrow{E_{13}} \begin{bmatrix} 0 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{E_1(-1)} \begin{bmatrix} 0 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{E_{21}(-1)} \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

e quindi una base di $E_A(2)$ è data dai vettori $\mathbf{v}_1 = [1 \ 0 \ 0]^T$ e $\mathbf{v}_2 = [0 \ 1 \ 1]^T$. Poiché questi vettori sono ortogonali, la base ortonormale richiesta è data da

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

Autospazio $E_A(4)$:

$$A - 4I = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{bmatrix} \xrightarrow{E_1(-1/2)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{bmatrix} \xrightarrow{E_2(-1)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{bmatrix} \xrightarrow{E_{32}(1)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

e quindi una base di $E_A(4)$ è data dal vettore $\mathbf{v}_3 = [0 \ 1 \ -1]^T$. La base ortonormale richiesta è data da

$$\mathbf{x}_3 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}.$$

Matrice P_1 :

$$P_1 = \mathbf{x}_1 \mathbf{x}_1^H + \mathbf{x}_2 \mathbf{x}_2^H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{bmatrix}.$$

Matrice P_2 :

$$P_2 = \mathbf{x}_3 \mathbf{x}_3^H = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1/2 & -1/2 \\ 0 & -1/2 & 1/2 \end{bmatrix}.$$

Verifica:

$$\lambda_1 P_1 + \lambda_2 P_2 = 2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{bmatrix} + 4 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1/2 & -1/2 \\ 0 & -1/2 & 1/2 \end{bmatrix} = A.$$