Università degli Studi di Roma Tor Vergata

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea Magistrale in Matematica

Tesi di laurea magistrale

Studio della convergenza del metodo delle potenze

Candidato:
Sara Loss
Matricola 0245827

Relatore:
Carmine Di Fiore

Anno Accademico 2017–2018

Alla mia famiglia che ha creduto in me.

Indice

Introduzione

Data una qualsiasi matrice, non sempre è necessario conoscere l'insieme di tutti i suoi autovalori , spesso è possibile limitarsi alla ricerca di quelli estremi, ovvero quelli di modulo massimo e/o di modulo minimo.
In questa tesi esporremo il *metodo delle potenze*, che è un classico metodo iterativo per approssimare l'autovalore di modulo massimo di una matrice, meglio detto autovalore dominante, e il corrispondente autovettore.

Nel Capitolo 1 iniziamo ad introdurre alcune definizioni ed osservazioni che saranno utili nei capitoli successivi. In particolare, tratteremo nozioni riguardanti la norma vettoriale, le matrici, gli autovalori ed i corrispondenti autovettori.

Nel Capitolo 2 introdurremo il *metodo delle potenze*. Inizieremo studiando tale metodo nel caso più semplice, ovvero quello di una matrice diagonalizzabile (in particolare definita positiva) e successivamente, passeremo prima a studiare il caso in cui la matrice sia generica, per poi concludere con due casi particolari ovvero, il metodo delle potenze applicato ad una matrice non negativa ed irriducibile e ad una non negativa, irriducibile e stocastica per colonne.

Nel Capitolo 3 verrà affrontato un problema riguardante la monotonia della successione φ_k definita nell'algoritmo del metodo delle potenze, visto nel capitolo precedente, nel caso particolare di una matrice non negativa ed irriducibile.

Capitolo 1

Nozioni utili

1.1 Norme vettoriali e prodotto scalare

Introduciamo alcune definizioni riguardanti le norme vettoriali ed il prodotto scalare.

Definizione 1.1. Una *norma vettoriale* è un'applicazione $||\cdot|| : \mathbb{C}^n \longrightarrow \mathbb{R}^+ \cup \{0\}$ tale che valgono le seguenti proprietà:

1. $||\mathbf{x}|| \geq \mathbf{0}\ \forall \mathbf{x} \in \mathbf{C}^n$ e $||\mathbf{x}|| = \mathbf{0}$ se e solo se $\mathbf{x} = \mathbf{0}$.
2. $||\alpha\mathbf{x}|| = |\alpha|||\mathbf{x}||\ \forall \alpha \in \mathbb{C}, \forall \mathbf{x} \in \mathbb{C}^n$.
3. $||\mathbf{x}+\mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||\ \forall \mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ (*disuguaglianza triangolare*).

Esempi importanti di norme vettoriali sono:

1. $||\mathbf{x}||_1 = \sum_{j=1}^{n} |x_j|$ (*norma-1*).
2. $||\mathbf{x}||_2 = \left(\sum_{j=1}^{n} |x_j|^2\right)^{\frac{1}{2}}$ (*norma Euclidea o norma-2*).
3. $||\mathbf{x}||_\infty = \max_j |x_j|$ (*norma infinito*).

Tutte queste norme sono casi particolari della *norma-p*, definita come:

$$||\mathbf{x}||_p = \left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}}.$$

Definizione 1.2. Il *prodotto scalare* su $\mathbb{C}^n$ è un'applicazione $\langle .,. \rangle : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}$ che alla coppia $(\mathbf{x}, \mathbf{y})$ associa lo scalare $\langle \mathbf{x}, \mathbf{y} \rangle$ tale che:

1. $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0 \ \forall \mathbf{x} \in \mathbb{C}^n$ e $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ se e solo se $\mathbf{x} = \mathbf{0}$.
2. $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle} \ \forall \mathbf{x}, \mathbf{y} \in \mathbb{C}^n$.
3. $\langle \mathbf{x}, \lambda\mathbf{y} \rangle = \lambda\langle \mathbf{x}, \mathbf{y} \rangle \ \forall \mathbf{x} \in \mathbb{C}^n, \forall \lambda \in \mathbb{C}$.
4. $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle \ \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{C}^n$.

Osservazione 1.1. (Disuguaglianza di Cauchy-Schwarz) Per ogni $\mathbf{x}$ e $\mathbf{y}$ vettori in $\mathbb{C}^n$, il loro prodotto scalare soddisfa la disuguaglianza di Cauchy-Schwarz, cioè:

$$|\langle \mathbf{x}, \mathbf{y} \rangle|^2 \leq ||\mathbf{x}||_2 \cdot ||\mathbf{y}||_2$$

(nota: vale l'uguaglianza se e solo se i vettori $\mathbf{x}$ ed $\mathbf{y}$ sono linearmente dipendenti).

1.2 Matrici

Introduciamo alcune definizioni riguardanti le matrici, che ci saranno utili.

Definizione 1.3. Sia $\mathbf{A}$ una matrice quadrata di ordine n. Diremo che $\mathbf{A}$ è *diagonalizzabile* se e solo se esiste $\mathbf{P}$, matrice invertibile, tale che:

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$$

dove $\mathbf{D}$ è diagonale.
La matrice $\mathbf{P}$ si dirà la matrice diagonalizzante di $\mathbf{A}$.

Definizione 1.4. Sia $\mathbf{U}$ una matrice complessa e quadrata di ordine n. Diremo che $\mathbf{U}$ è *unitaria* se soddisfa la condizione:

$$\mathbf{U}^H\mathbf{U} = \mathbf{U}\mathbf{U}^H = \mathbf{I}$$

dove $\mathbf{I}$ è la matrice identità e $\mathbf{U}^H$ è la trasposta coniugata di $\mathbf{U}$.

Esempio 1.1. Le matrici di permutazione sono esempi di matrici unitarie. Ricordiamo che $\mathbf{P} \in \mathbb{R}^{n \times n}$ si dice matrice di permutazione se è ottenuta dalla matrice identica operando su di essa una permutazione delle righe o delle colonne. Sia $\{i_1, i_2, \ldots, i_n\}$ una permutazione degli indici $\{1, 2, \ldots, n\}$ allora, ogni matrice di permutazione si può scrivere come:

$$\mathbf{P} = \begin{bmatrix} \mathbf{e}_{i_1} & \mathbf{e}_{i_2} & \ldots & \mathbf{e}_{i_n} \end{bmatrix}$$

dove $\mathbf{e}_k \in \mathbb{C}^n$ e:

$$(e_k)_r = \begin{cases} 0 & \text{se } r \neq k \\ 1 & \text{se } r = k \end{cases}$$

Si osserva che:

$$\begin{aligned}\mathbf{P}^T\mathbf{P} &= \begin{bmatrix} \mathbf{e}_{i_1}^T \\ \mathbf{e}_{i_2}^T \\ \vdots \\ \mathbf{e}_{i_n}^T \end{bmatrix} \cdot \begin{bmatrix} \mathbf{e}_{i_1} & \mathbf{e}_{i_2} & \dots & \mathbf{e}_{i_n} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{e}_{i_1}^T\mathbf{e}_{i_1} & \mathbf{e}_{i_1}^T\mathbf{e}_{i_2} & \dots & \dots & \mathbf{e}_{i_1}^T\mathbf{e}_{i_n} \\ \mathbf{e}_{i_2}^T\mathbf{e}_{i_1} & & & & \\ & & & & \\ \mathbf{e}_{i_n}^T\mathbf{e}_{i_1} & & & & \mathbf{e}_{i_n}^T\mathbf{e}_{i_n} \end{bmatrix}\end{aligned}$$

ed in particolare:

$$(\mathbf{P}^T\mathbf{P})_{r,s} = \mathbf{e}_{i_r}^T\mathbf{e}_{i_s} = \begin{cases} 0 & \text{se } r \neq s \\ 1 & \text{se } r = s \end{cases}$$

da cui ne segue che:

$$\mathbf{P}^T\mathbf{P} = \mathbf{I}$$

quindi, le matrici di permutazione sono unitarie.

Definizione 1.5. Una matrice quadrata $\mathbf{A}$ di ordine n si dice *diagonale* se gli unici elementi che possono essere non nulli sono quelli sulla diagonale principale. Ossia, se:

$$A_{i,j} = 0,\ i \neq j,\ \forall i, j = 1, \dots, n.$$

Definizione 1.6. Sia $\mathbf{A}$ una matrice quadrata di ordine n. Diremo che $\mathbf{A}$ è *diagonalizzabile mediante matrice unitaria* se esiste una matrice unitaria $\mathbf{U}$, tale che:

$$\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \mathbf{D}$$

dove $\mathbf{D}$ è diagonale.

Definizione 1.7. Due matrici $\mathbf{A}$, $\mathbf{B}$ di ordine n si dicono *simili* se esiste $\mathbf{P}$, matrice invertibile, con la proprietà che:

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}.$$

Quindi, una matrice è diagonalizzabile se è simile ad una matrice diagonale.

Definizione 1.8. Una matrice quadrata $\mathbf{A}$ si dice *triangolare* se ha tutti gli elementi nulli sopra o sotto la diagonale principale.
In particolare, a seconda che gli elementi nulli siano sopra o sotto la diagonale, la matrice viene chiamata rispettivamente *triangolare inferiore* e *triangolare superiore*.

Definizione 1.9. Una matrice quadrata $\mathbf{A}$ di ordine n e complessa si dice *normale* se:

$$\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}.$$

Lemma 1.1. *Una matrice normale e triangolare è diagonale.*

Dimostrazione. Consideriamo la matrice triangolare a blocchi:

$$\begin{bmatrix} \mathbf{a} & \mathbf{b}^H \\ 0 & \mathbf{C} \end{bmatrix}$$

dove $\mathbf{C}$ è una matrice triangolare.
Poiché per ipotesi è normale, allora:

$$\begin{bmatrix} \bar{\mathbf{a}} & 0 \\ \mathbf{b} & \mathbf{C}^H \end{bmatrix} \cdot \begin{bmatrix} \mathbf{a} & \mathbf{b}^H \\ 0 & \mathbf{C} \end{bmatrix} = \begin{bmatrix} \mathbf{a} & \mathbf{b}^H \\ 0 & \mathbf{C} \end{bmatrix} \cdot \begin{bmatrix} \bar{\mathbf{a}} & 0 \\ \mathbf{b} & \mathbf{C}^H \end{bmatrix}.$$

Da cui, sviluppando i calcoli:

$$\begin{bmatrix} \bar{\mathbf{a}}\mathbf{a} & \bar{\mathbf{a}}\mathbf{b}^H \\ \mathbf{a}\mathbf{b} & \mathbf{b}\mathbf{b}^H + \mathbf{C}^H\mathbf{C} \end{bmatrix} = \begin{bmatrix} \mathbf{a}\bar{\mathbf{a}} + \mathbf{b}^H\mathbf{b} & \mathbf{b}^H\mathbf{C}^H \\ \mathbf{C}\mathbf{b} & \mathbf{C}\mathbf{C}^H \end{bmatrix}$$

quindi, sono uguali se e solo se:

$$\begin{cases} \bar{\mathbf{a}}\mathbf{a} = \mathbf{a}\bar{\mathbf{a}} + \mathbf{b}^H\mathbf{b} \\ \bar{\mathbf{a}}\mathbf{b}^H = \mathbf{b}^H\mathbf{C}^H \\ \mathbf{a}\mathbf{b} = \mathbf{C}\mathbf{b} \\ \mathbf{b}\mathbf{b}^H + \mathbf{C}^H\mathbf{C} = \mathbf{C}\mathbf{C}^H \end{cases}$$

da queste, otteniamo che:

$$(||\mathbf{b}||_2)^2 = 0$$

e quindi:

$$\mathbf{b} = \mathbf{0}.$$

Ma se $\mathbf{b} = \mathbf{0}$, allora:
$$\mathbf{C}^H\mathbf{C} = \mathbf{C}\mathbf{C}^H$$
ovvero, $\mathbf{C}$ è normale.
Ma poiché per ipotesi $\mathbf{C}$, che ha ordine $(n-1) \times (n-1)$, è triangolare allora è diagonale.
Da questo, ne segue che l'intera matrice quadrata di ordine n è diagonale.

$\square$

Teorema 1.1. *(Decomposizione di Schur) Ogni matrice $\mathbf{A} \in \mathbb{C}^{n \times n}$ si può triangolarizzare con una trasformazione per similitudine data da una matrice unitaria. Ovvero, esistono una matrice triangolare superiore $\mathbf{T}$ ed una matrice unitaria $\mathbf{U}$ tali che:*

$$\mathbf{U}^H\mathbf{A}\mathbf{U} = \mathbf{T}.$$

Dimostrazione. Procediamo per induzione sulla dimensione n della matrice $\mathbf{A}$.
Sia $n = 1$.
In tal caso, la matrice $\mathbf{A}$ è un numero complesso e la decomposizione di Schur vale con $\mathbf{T} = \mathbf{A}$ e $\mathbf{U} = 1$.
In generale, supponiamo ora che la decomposizione valga per dimensione $n-1$ e vogliamo dimostrare che ciò è vero anche se la dimensione della matrice è pari ad n.
Consideriamo quindi $\mathbf{A} \in \mathbb{C}^{n \times n}$ e denotiamo con λ ed $\mathbf{x}$ rispettivamente un autovalore ed un autovettore di $\mathbf{A}$, cioè $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ (vedi Definizione 1.17.). Supponiamo che $\mathbf{x}$ sia normalizzato in modo che $\mathbf{x}^H\mathbf{x} = 1$ e consideriamo una base ortonormale formata dai vettori $\{\mathbf{y}_2, \dots, \mathbf{y}_n\}$ dello spazio ortogonale ad $\mathbf{x}$.
Costruiamo la matrice $\mathbf{Q}$ le cui colonne sono costituite dai vettori $\{\mathbf{x}, \mathbf{y}_2, \dots, \mathbf{y}_n\}$.
In particolare, la matrice appena costruita è unitaria, ovvero:

$$\mathbf{Q}^H\mathbf{Q} = \mathbf{I}$$

ed inoltre si ha che:
$$\mathbf{Q}\mathbf{e}_1 = \mathbf{x}$$
da cui $\mathbf{e}_1 = \mathbf{Q}^H\mathbf{x}$, dove $\mathbf{e}_1 = (1, 0, \dots, 0)^T$.
Inoltre vale:
$$\mathbf{Q}^H\mathbf{A}\mathbf{Q} = \begin{bmatrix} \lambda & \mathbf{u}^T \\ 0 & \mathbf{A}_{n-1} \end{bmatrix}$$
dove $\mathbf{A}_{n-1} \in \mathbb{C}^{(n-1)\times(n-1)}$.
Per verificare quest'ultima relazione, basta considerare $\mathbf{Q}^H\mathbf{A}\mathbf{Q}\mathbf{e}_1$, che è la

prima colonna della matrice $\mathbf{Q}^H\mathbf{A}\mathbf{Q}$.
Allora, vale:

$$\mathbf{Q}^H\mathbf{A}\mathbf{Q}\mathbf{e}_1 = \mathbf{Q}^H\mathbf{A}\mathbf{x} = \mathbf{Q}^H\lambda\mathbf{x} = \lambda\mathbf{Q}^H\mathbf{x} = \lambda\mathbf{e}_1$$

come richiesto.
Inoltre, per l'ipotesi induttiva, si ha che:

$$\mathbf{A}_{n-1} = \mathbf{U}_{n-1}\mathbf{T}_{n-1}\mathbf{U}_{n-1}^H$$

da cui posto:

$$\mathbf{U}_n = \mathbf{Q}\begin{bmatrix} 1 & 0 \\ & \mathbf{U}_{n-1} \end{bmatrix}$$

si ottiene:

$$\mathbf{U}_n^H\mathbf{A}\mathbf{U}_n = \begin{bmatrix} \lambda & \mathbf{u}^T\mathbf{U}_{n-1}^H \\ 0 & \mathbf{T}_{n-1} \end{bmatrix}$$

da cui la tesi, essendo $\mathbf{U}_n^H\mathbf{A}\mathbf{U}_n$ una matrice triangolare superiore.

□

Osservazione 1.2. Un'osservazione utile è che la decomposizione di Schur della matrice $\mathbf{A}$ non è unica. Infatti, per qualsiasi scelta di $\mathbf{P}$ matrice di permutazione e $\mathbf{D}$ matrice diagonale unitaria, moltiplicando $\mathbf{A}$ a sinistra per $\mathbf{P}diag\{e^{-i\theta}\}\mathbf{U}_n^H$ ed a destra per $\mathbf{U}_n\mathbf{P}diag\{e^{i\theta}\}\mathbf{P}^T$, cioè:

$$(\mathbf{P}diag\{e^{-i\theta}\}\mathbf{U}_n^H)\mathbf{A}(\mathbf{U}_n\mathbf{P}diag\{e^{i\theta}\}\mathbf{P}^T)$$

si ottiene sempre una matrice triangolare superiore sulla cui diagonale vi sono gli autovalori di $\mathbf{A}$.

La decomposizione di Schur ha delle conseguenze interessanti, una prima conseguenza riguarda in particolare le matrici hermitiane.
Infatti, se $\mathbf{U}^H\mathbf{A}\mathbf{U} = T$ è la decomposizione di Schur della matrice hermitiana $\mathbf{A}$ allora, la proprietà $\mathbf{A} = \mathbf{A}^H$ implica:

$$\mathbf{T}^H = \mathbf{U}^H\mathbf{A}^H\mathbf{U} = \mathbf{U}^H\mathbf{A}\mathbf{U} = \mathbf{T}.$$

Cioè $\mathbf{T}$ è hermitiana, e quindi essendo triangolare, è diagonale.
Ovvero si ottiene che, le matrici hermitiane sono diagonalizzabili con una trasformazione per similitudine unitaria. In particolare, gli elementi diagonali della matrice $\mathbf{T}$ sono tali che $t_{i,i} = \overline{t_{i,i}}$ e quindi sono reali.
Se invece la matrice $\mathbf{A}$ è anti-hermitiana, cioè se $\mathbf{A} = -\mathbf{A}^H$ procedendo analogamente a quanto detto sopra, si ottiene che anche $\mathbf{T}$ è una matrice anti-hermitiana ed in particolare, è una matrice diagonale con elementi tali che $t_{i,i} = -\overline{t_{i,i}}$.

Un risultato più generale è espresso dal seguente teorema.

Teorema 1.2. *Una matrice* $\mathbf{A} \in \mathbb{C}^{n \times n}$ *è normale, cioè* $\mathbf{A}^H\mathbf{A} = \mathbf{A}\mathbf{A}^H$, *se e solo se è diagonalizzabile mediante matrice unitaria.*

Dimostrazione. Dimostriamo la doppia implicazione.
Si supponga dapprima che $\mathbf{A}$ sia normale. Per la decomposizione di Schur (Teorema 1.1.) esiste una matrice unitaria $\mathbf{U}$ tale che:

$$\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^H$$

con $\mathbf{T}$ matrice triangolare superiore.
Poiché $\mathbf{A}$ è normale, si ottiene:

$$(\mathbf{U}\mathbf{T}\mathbf{U}^H)^H(\mathbf{U}\mathbf{T}\mathbf{U}^H) = (\mathbf{U}\mathbf{T}\mathbf{U}^H)(\mathbf{U}\mathbf{T}\mathbf{U}^H)^H$$

$$(\mathbf{U}\mathbf{T}^H\mathbf{U}^H)(\mathbf{U}\mathbf{T}\mathbf{U}^H) = (\mathbf{U}\mathbf{T}\mathbf{U}^H)(\mathbf{U}\mathbf{T}^H\mathbf{U}^H)$$

$$(\mathbf{U}\mathbf{T}^H)(\mathbf{U}^H\mathbf{U})(\mathbf{T}\mathbf{U}^H) = (\mathbf{U}\mathbf{T})(\mathbf{U}^H\mathbf{U})(\mathbf{T}^H\mathbf{U}^H)$$

Da ciò:

$$\mathbf{U}\mathbf{T}^H\mathbf{T}\mathbf{U}^H = \mathbf{U}\mathbf{T}\mathbf{T}^H\mathbf{U}^H \Rightarrow \mathbf{T}^H\mathbf{T} = \mathbf{T}\mathbf{T}^H$$

ovvero, $\mathbf{T}$ è normale e triangolare. Ma questo, per il Lemma 1.1., implica che $\mathbf{T}$ è una matrice diagonale.
Viceversa, supponiamo ora che $\mathbf{A}$ sia diagonalizzabile mediante una matrice unitaria e vogliamo dimostrare che $\mathbf{A}$ è normale.
Per ipotesi esiste una matrice unitaria $\mathbf{U}$ tale che:

$$\mathbf{U}^H\mathbf{A}\mathbf{U} = \mathbf{D}$$

dove $\mathbf{D}$ è una matrice diagonale.
Da questo si ha che:

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^H.$$

Calcoliamoci ora le quantità $\mathbf{A}^H\mathbf{A}$ e $\mathbf{A}\mathbf{A}^H$:

$$\mathbf{A}^H\mathbf{A} = (\mathbf{U}\mathbf{D}\mathbf{U}^H)^H(\mathbf{U}\mathbf{D}\mathbf{U}^H) = (\mathbf{U}\mathbf{D}^H\mathbf{U}^H)(\mathbf{U}\mathbf{D}\mathbf{U}^H) = \mathbf{U}\mathbf{D}^H\mathbf{D}\mathbf{U}^H$$

$$\mathbf{A}\mathbf{A}^H = (\mathbf{U}\mathbf{D}\mathbf{U}^H)(\mathbf{U}\mathbf{D}\mathbf{U}^H)^H = (\mathbf{U}\mathbf{D}\mathbf{U}^H)(\mathbf{U}\mathbf{D}^H\mathbf{U}^H) = \mathbf{U}\mathbf{D}\mathbf{D}^H\mathbf{U}^H.$$

E quindi, si ha che $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$ se e solo se:

$$\mathbf{D}\mathbf{D}^H = \mathbf{D}^H\mathbf{D}$$

ma questo è sempre vero.

□

Osservazione 1.3. Una matrice unitaria $\mathbf{A}$ è in particolare una matrice normale e quindi la $\mathbf{T}$ della decomposizione di Schur è diagonale. Inoltre, dal fatto che $\mathbf{A}^H\mathbf{A} = \mathbf{I}$, segue che:

$$|t_{i,i}|^2 = 1.$$

Osservazione 1.4. Se esiste una matrice unitaria $\mathbf{U}$ che diagonalizza $\mathbf{A}$ allora ogni altra $\mathbf{UDP}$ diagonalizza $\mathbf{A}$, dove $\mathbf{D} = diag(e^{i\theta_k},\ k = 1, \ldots, n))$ e $\mathbf{P}$ di permutazione. (Vedi anche l'Osservazione 1.2.)

Infatti, se $\mathbf{U}$ diagonalizza la matrice $\mathbf{A}$, per definizione si ha:

$$\mathbf{U}^H\mathbf{A}\mathbf{U} = \mathbf{\Lambda_A}$$

dove $\mathbf{\Lambda_A}$ è matrice diagonale che ha gli autovalori di $\mathbf{A}$ sulla diagonale principale.
Moltiplichiamo ora ambo i membri a sinistra per $\mathbf{D}^H$ e a destra per $\mathbf{D}$, ottenendo:

$$(\mathbf{D}^H\mathbf{U}^H)\mathbf{A}(\mathbf{UD}) = (\mathbf{UD})^H\mathbf{A}(\mathbf{UD}) = \mathbf{D}^H\mathbf{\Lambda_A}\mathbf{D}$$

da cui:

$$\begin{bmatrix} e^{-i\theta_1} & 0 & \ldots & 0 \\ 0 & e^{-i\theta_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & e^{-i\theta_n} \end{bmatrix} \mathbf{\Lambda_A} \begin{bmatrix} e^{i\theta_1} & 0 & \ldots & 0 \\ 0 & e^{i\theta_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & e^{i\theta_n} \end{bmatrix} = \mathbf{\Lambda_A}.$$

Definizione 1.10. Dato uno scalare λ, chiamiamo *blocco di Jordan* di ordine t con autovalore λ la matrice di ordine t avente tutti gli elementi appartenenti alla diagonale principale uguali a λ, tutti gli elementi della diagonale subito superiore a quella principale uguali ad 1 e tutti gli altri nulli.
(nota: si dimostra che ogni matrice $\mathbf{A} \in \mathbb{C}^{n\times n}$ è simile ad una matrice diagonale a blocchi dove i blocchi diagonali sono di Jordan).

Definizione 1.11. Una matrice $\mathbf{A}$ si dice *non negativa* se tutti i suoi elementi sono maggiori o uguali a zero.

Definizione 1.12. Una matrice $\mathbf{A}$ si dice *riducibile* se esiste una matrice di permutazione $\mathbf{P}$ tale che $\mathbf{PAP}^{-1} = \mathbf{PAP}^T$ è triangolare a blocchi, cioè:

$$\mathbf{PAP}^{-1} = \begin{bmatrix} \mathbf{B} & 0 \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$$

dove $\mathbf{B}$ e $\mathbf{D}$ sono matrici quadrate.
Una matrice $\mathbf{A}$ che non è riducibile si dice *irriducibile*.

Definizione 1.13. Una matrice quadrata $\mathbf{A} = \{a_{i,j}\}$ di ordine n ad elementi non negativi si dice *stocastica per colonne* (*righe*) se la somma degli elementi su ogni colonna (o su ogni riga) è uguale ad 1, cioè:

$$a_{i,j} \geq 0 \text{ e } \sum_{i=1}^{n} a_{i,j} = 1,\ j = 1, \dots, n$$

o analogamente sommando sulle righe.

Definizione 1.14. Una matrice quadrata $\mathbf{A}$ di ordine n e complessa si dice *hermitiana* se coincide con la trasposta coniugata, ovvero:

$$\mathbf{A} = \mathbf{A}^H$$

Definizione 1.15. Una matrice quadrata $\mathbf{A}$ di ordine n si dice *definita positiva* se:

$$\mathbf{z}^H \mathbf{A} \mathbf{z} > \mathbf{0},\ \forall \mathbf{z} \in \mathbb{C}^n \text{ e } \mathbf{z} \neq \mathbf{0}$$

Osservazione 1.5. $\mathbf{A}$ definita positiva $\Rightarrow$ $\mathbf{A}$ hermitiana, $\mathbf{A}$ con elementi diagonali positivi ed $\mathbf{A}$ diagonalizzabile con autovalori positivi.

Osservazione 1.6. Data una matrice $\mathbf{A}$ definita positiva, esiste una matrice $\mathbf{B}$ definita positiva tale che:

$$\mathbf{B}^2 = \mathbf{A}.$$

Infatti, $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^H$ con $D_{i,i}$ positivi, quindi:

$$\mathbf{A} = (\mathbf{U}\sqrt{\mathbf{D}}\mathbf{U}^H)(\mathbf{U}\sqrt{\mathbf{D}}\mathbf{U}^H)$$

dove $\sqrt{\mathbf{D}}$ è la matrice diagonale dove gli elementi sulla diagonale sono le radici degli autovalori.

Definizione 1.16. Una *matrice di Toeplitz* è una matrice in cui gli elementi su ogni diagonale discendente da sinistra a destra sono uguali tra loro.
Per esempio, una matrice quadrata $\mathbf{A}$ di ordine n è detta di Toeplitz se:

$$\mathbf{A} = \begin{bmatrix} a_0 & a_{-1} & a_{-2} & \dots & \dots & a_{-n+1} \\ a_1 & a_0 & a_{-1} & \ddots & & \vdots \\ \vdots & a_1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & a_{-1} & a_{-2} \\ \vdots & & & \ddots & a_0 & a_{-1} \\ a_{n-1} & \dots & \dots & \dots & a_1 & a_0 \end{bmatrix} = (a_{i-j})_{i,j=1}^{n}.$$

1.3 Autovalori ed autovettori

Introduciamo alcune definizioni ed osservazioni riguardanti gli autovalori ed autovettori di una matrice.

Definizione 1.17. Sia $\mathbf{A}$ una matrice quadrata di ordine n; $\lambda \in \mathbb{C}$ è detto *autovalore* di $\mathbf{A}$ se esiste $\mathbf{x} \in \mathbb{C}^n$, $\mathbf{x} \neq \mathbf{0}$, tale che:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

Dove il vettore $\mathbf{x}$ è detto *autovettore* associato all'autovalore λ.

In particolare se λ è autovalore di $\mathbf{A}$ allora è soluzione dell'equazione caratteristica:

$$p_{\mathbf{A}}(\lambda) = det(\lambda\mathbf{I} - \mathbf{A}) = 0$$

dove $\mathbf{I}$ è la matrice identità e $p_{\mathbf{A}}(\lambda)$ è detto *polinomio caratteristico* di $\mathbf{A}$.

Definizione 1.18. Data una generica matrice $\mathbf{A}$ viene chiamato $\sigma(\mathbf{A})$, lo *spettro* di $\mathbf{A}$, l'insieme degli autovalori della matrice, cioè:

$$\sigma(\mathbf{A}) = \{\lambda \in \mathbb{C} : \exists\mathbf{v} \neq \mathbf{0} \text{ tale che } \mathbf{A}\mathbf{v} = \lambda\mathbf{v}\}$$

e si definisce $\rho(\mathbf{A})$ il suo *raggio spettrale*:

$$\rho(\mathbf{A}) = \max_{\lambda \in \sigma(\mathbf{A})} |\lambda|$$

Definizione 1.19. Sia $\mathbf{A}$ una matrice quadrata di ordine n e $\lambda \in \mathbb{C}$ autovalore di $\mathbf{A}$. Si dice *molteplicità algebrica* di λ (e si indica con $m_a(\lambda)$) la molteplicità che λ ha come radice del polinomio caratteristico.
Si dice invece *molteplicità geometrica* di λ (e si indica con $m_g(\lambda)$) la dimensione dell'autospazio relativo a λ, ovvero il numero degli autovettori linearmente indipendenti relativi all'autovalore λ.

Lemma 1.2. *Siano $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\lambda \in \mathbb{C}$ autovalore di $\mathbf{A}$ ed $\mathbf{x} \in \mathbb{C}^n$ corrispettivo autovettore, cioè $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. Allora ne segue che:*

$$p(\mathbf{A})\mathbf{x} = p(\lambda)\mathbf{x}$$

per ogni polinomio p.

Osservazione 1.7. Una matrice quadrata $\mathbf{A}$ di ordine n è diagonalizzabile se e solo se la molteplicità algebrica e geometrica coincidono per ogni autovalore di $\mathbf{A}$. Inoltre, se $\mathbf{A}$ è diagonalizzabile, gli elementi sulla diagonale di $\mathbf{D}$ (vedi Definizione 1.3.) sono gli autovalori di $\mathbf{A}$[1] e le colonne di $\mathbf{P}$ sono gli autovettori corrispondenti.

[1] $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} \Rightarrow p_{\mathbf{A}}(\lambda) = p_{\mathbf{D}}(\lambda)$

Osservazione 1.8. Autovettori relativi ad autovalori distinti sono indipendenti. In particolare, nel caso in cui la matrice è normale, sono ortogonali.

Osservazione 1.9. Se $\mathbf{A}$ è una matrice quadrata di ordine n e definita positiva, allora i suoi autovalori sono tutti strettamente positivi. Inoltre, il massimo autovalore coincide proprio con il raggio spettrale $\rho(\mathbf{A})$ (nota: $\rho(\mathbf{A})$ può avere molteplicità algebrica maggiore di 1).

Capitolo 2

Metodo delle Potenze

2.1 Caso A diagonalizzabile

In questa sezione dimostriamo un teorema in cui si analizza la convergenza delle potenze di una matrice A diagonalizzabile (vedi Teorema 2.1.).

Il metodo delle potenze è un metodo iterativo per il calcolo approssimato dell'autovalore di modulo massimo di una matrice e di un autovettore ad esso associato.

Sia $\mathbf{A} \in \mathbb{C}^{n\times n}$ diagonalizzabile con autovalori $\lambda_1, \lambda_2, \ldots, \lambda_n$. Supponiamo, inoltre, che gli autovalori siano ordinati in modo tale che:

$$|\lambda_1| > |\lambda_j|, \ \forall \lambda_j \neq \lambda_1$$

(nota: molteplicità geometrica ed algebrica di λ_1 coincidono).
Con tali ipotesi λ_1 è detto autovalore dominante per la matrice $\mathbf{A}$.

Siccome per ipotesi $\mathbf{A}$ è diagonalizzabile, esiste una base di autovettori $\{\mathbf{x}_1, ..., \mathbf{x}_n\}$ dove $\mathbf{x}_i$ è l'autovettore associato all'autovalore λ_i, cioè $\mathbf{A}\mathbf{x}_i = \lambda_i\mathbf{x}_i$, $i = 1, ..., n$. Sia $\mathbf{v} \in \mathbb{C}^n$, allora questo può essere scritto come:

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{x}_i$$

Osserviamo che, per le identità $\mathbf{A}^k\mathbf{x}_i = \lambda_i^k\mathbf{x}_i$ (vedi Lemma 1.2.), l'azione di $\mathbf{A}^k$ su $\mathbf{v}$ ammette la seguente rappresentazione:

$$\begin{aligned}\mathbf{A}^k\mathbf{v} = \mathbf{A}^k \sum_{i=1}^{n} \alpha_i \mathbf{x}_i &= \sum_{i=1}^{n} \alpha_i \mathbf{A}^k \mathbf{x}_i = \sum_{i=1}^{n} \alpha_i \lambda_i^k \mathbf{x}_i \\ &= \sum_{i:\lambda_i=\lambda_1} \alpha_i \lambda_1^k \mathbf{x}_i + \sum_{i:\lambda_i\neq\lambda_1} \alpha_i \lambda_i^k \mathbf{x}_i.\end{aligned}$$

Dividiamo adesso ambo i membri per λ_1^k e si ha:

$$\frac{1}{\lambda_1^k}\mathbf{A}^k\mathbf{v} = \sum_{i:\lambda_i=\lambda_1} \alpha_i \mathbf{x}_i + \sum_{i:\lambda_i\neq\lambda_1} \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^k \mathbf{x}_i$$

e passando al limite per $k \to \infty$ si ottiene che:

$$\frac{1}{\lambda_1^k}\mathbf{A}^k\mathbf{v} \to \sum_{i:\lambda_i=\lambda_1} \alpha_i \mathbf{x}_i \tag{2.1}$$

e quindi $\frac{1}{\lambda_1^k}\mathbf{A}^k\mathbf{v}$ tende a disporsi nella direzione dell'autovettore $\mathbf{x} = \sum_{i:\lambda_i=\lambda_1} \alpha_i \mathbf{x}_i$ associato all'autovalore λ_1. Dalla (2.1) otteniamo:

$$\frac{||\mathbf{A}^k\mathbf{v}||}{|\lambda_1|^k} \xrightarrow[k\to\infty]{} ||\mathbf{x}|| \tag{2.2}$$

Considerando sia la (2.1) che la (2.2) si ha una successione convergente all'autovettore $\mathbf{x}$ normalizzato:

$$\frac{|\lambda_1|^k}{\lambda_1^k}\frac{\mathbf{A}^k\mathbf{v}}{||\mathbf{A}^k\mathbf{v}||} \xrightarrow[k\to\infty]{} \frac{\mathbf{x}}{||\mathbf{x}||}. \tag{2.3}$$

La (2.1) implica pure che, per ogni vettore $\mathbf{u} \in \mathbb{C}^n$:

$$\frac{1}{\lambda_1^k}(\mathbf{u}^H\mathbf{A}^k\mathbf{v}) \xrightarrow[k\to\infty]{} \mathbf{u}^H\mathbf{x}$$

allora sarà anche vero che:

$$\frac{1}{\lambda_1^{k+1}}(\mathbf{u}^H\mathbf{A}^{k+1}\mathbf{v}) \xrightarrow[k\to\infty]{} \mathbf{u}^H\mathbf{x} \tag{2.4}$$

e quindi, se $\mathbf{u}^H\mathbf{x} \neq 0$:

$$\frac{\mathbf{u}^H\mathbf{A}^{k+1}\mathbf{v}}{\mathbf{u}^H\mathbf{A}^k\mathbf{v}} \xrightarrow[k\to\infty]{} \lambda_1 \tag{2.5}$$

Da cui possiamo osservare che, con il metodo delle potenze, è possibile calcolare anche l'autovalore dominante.

Enunciamo di conseguenza il seguente teorema:

Teorema 2.1. *Sia* $\mathbf{A} \in \mathbb{C}^{n\times n}$ *diagonalizzabile, siano* λ_i*,* $i = 1, 2, \ldots, n$*, i suoi autovalori e* $\mathbf{x}_i$ *autovettori linearmente indipendenti corrispondenti ai* λ_i*.*

Supponiamo $|\lambda_1| > |\lambda_j|$, $\forall \lambda_j \neq \lambda_1$ *(osservazione:* $m_a(\lambda_1) = m_g(\lambda_1)$*).*
Sia $\mathbf{v} \in \mathbb{C}^n$, $||\mathbf{v}|| = 1$, *tale che nella rappresentazione di* $\mathbf{v}$, $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{x}_i$, *il vettore* $\mathbf{x} = \sum_{i:\lambda_i=\lambda_1} \alpha_i \mathbf{x}_i$ *sia non nullo (nota: ciò implica* $\mathbf{A}^k\mathbf{v}$ *non nullo,* $\forall k = 1, 2, 3, ...$*).*
Sia anche $\mathbf{u} \in \mathbb{C}^n$, *tale che* $\mathbf{u}^H\mathbf{x} \neq \mathbf{0}$ *(nota:* $\frac{1}{\lambda_1^k}(\mathbf{u}^H\mathbf{A}^k\mathbf{v}) \xrightarrow[k\to\infty]{} \mathbf{u}^H\mathbf{x}$ *e quindi* $\exists k^* \in \mathbb{N}$ *per cui sono non nulli* $\forall k > k^*$*). Allora:*

$$\frac{|\lambda_1|^k}{\lambda_1^k} \frac{\mathbf{A}^k\mathbf{v}}{||\mathbf{A}^k\mathbf{v}||} \xrightarrow[k\to\infty]{} \frac{\mathbf{x}}{||\mathbf{x}||} \tag{2.6}$$

$$\frac{\mathbf{u}^H\mathbf{A}^{k+1}\mathbf{v}}{\mathbf{u}^H\mathbf{A}^k\mathbf{v}} \xrightarrow[k\to\infty]{} \lambda_1 \tag{2.7}$$

*dove la successione in (*2.6*) è ben definita* $\forall k$ *ed* $\exists k^* \in \mathbb{N}$ *tale che la successione (*2.7*) è ben definita* $\forall k > k^*$.
*Le quantità in (*2.6*) e (*2.7*) si possono calcolare tramite il seguente algoritmo:*

$\mathbf{v}_0 := \mathbf{v}$. *Per k = 1, 2, ...*

$$\begin{cases} \mathbf{a}_k := \mathbf{A}\mathbf{v}_{k-1} \\ \varphi_k := \dfrac{\mathbf{u}^H\mathbf{a}_k}{\mathbf{u}^H\mathbf{v}_{k-1}}, \quad \mathbf{u}^H\mathbf{v}_{k-1} \neq 0 \ \forall k \\ \mathbf{v}_k := \dfrac{\mathbf{a}_k}{||\mathbf{a}_k||} \end{cases} \tag{2.8}$$

Infatti, $\varphi_k = \frac{\mathbf{u}^H\mathbf{A}^k\mathbf{v}}{\mathbf{u}^H\mathbf{A}^{k-1}\mathbf{v}}$ *(e quindi* $\varphi_k \xrightarrow[k\to\infty]{} \lambda_1$*) e* $\mathbf{v}_k = \frac{\mathbf{A}^k\mathbf{v}}{||\mathbf{A}^k\mathbf{v}||}$ $\forall k = 1, 2, ...$ *(e quindi* $\frac{|\lambda_1|^k}{\lambda_1^k}\mathbf{v}_k \xrightarrow[k\to\infty]{} \frac{\mathbf{x}}{||\mathbf{x}||}$*).*

Dimostrazione. Rimangono da dimostrare solamente le affermazioni:

1. $\mathbf{A}^k\mathbf{v}$ non nullo, $\forall k = 1, 2, 3, ...$
2. $\varphi_k = \dfrac{\mathbf{u}^H\mathbf{A}^k\mathbf{v}}{\mathbf{u}^H\mathbf{A}^{k-1}\mathbf{v}}$ e $\mathbf{v}_k = \dfrac{\mathbf{A}^k\mathbf{v}}{||\mathbf{A}^k\mathbf{v}||}$ $\forall k = 1, 2, ...$

Punto 1:
Supponiamo per assurdo che $\mathbf{A}^k\mathbf{v}$ sia nullo, allora anche la sua rappresentazione data da:

$$\mathbf{A}^k\mathbf{v} = \sum_{i:\lambda_1=\lambda_i} \alpha_i \lambda_1^k \mathbf{x}_i + \sum_{i:\lambda_i\neq\lambda_1} \alpha_i \lambda_i^k \mathbf{x}_i$$

sarà nulla.
Ma poiché per ipotesi gli autovettori $\mathbf{x}_i$ sono linearmente indipendenti, questo implica che gli scalari α_i, con $i : \lambda_i = \lambda_1$, siano tutti uguali a 0 (perché $\lambda_1 \neq 0$). Da questo ne segue però che anche il vettore $\mathbf{x}$ sarebbe nullo, che è assurdo.

Punto 2:
Procediamo per induzione.
Sia $k = 1$ e vogliamo dimostrare che:

$$\mathbf{v}_1 = \frac{\mathbf{Av}}{||\mathbf{Av}||}$$

$$\varphi_1 = \frac{\mathbf{u}^H\mathbf{Av}}{\mathbf{u}^H\mathbf{v}}.$$

Utilizzando le definizioni dell'algoritmo (2.8) otteniamo:

$$\mathbf{v}_1 = \frac{\mathbf{a}_1}{||\mathbf{a}_1||} = \frac{\mathbf{Av}_0}{||\mathbf{Av}_0||} = \frac{\mathbf{Av}}{||\mathbf{Av}||}$$

$$\varphi_1 = \frac{\mathbf{u}^H\mathbf{a}_1}{\mathbf{u}^H\mathbf{v}_0} = \frac{\mathbf{u}^H\mathbf{Av}_0}{\mathbf{u}^H\mathbf{v}_0} = \frac{\mathbf{u}^H\mathbf{Av}}{\mathbf{u}^H\mathbf{v}}.$$

Ora supponiamo che la tesi è vera per $k \leq n$, in particolare:

$$\varphi_n = \frac{\mathbf{u}^H\mathbf{A}^n\mathbf{v}}{\mathbf{u}^H\mathbf{A}^{n-1}\mathbf{v}}$$

$$\mathbf{v}_n = \frac{\mathbf{A}^n\mathbf{v}}{||\mathbf{A}^n\mathbf{v}||}$$

Dimostriamola per $k = n + 1$, cioè vogliamo dimostrare che:

$$\varphi_{n+1} = \frac{\mathbf{u}^H\mathbf{A}^{n+1}\mathbf{v}}{\mathbf{u}^H\mathbf{A}^n\mathbf{v}}$$

$$\mathbf{v}_{n+1} = \frac{\mathbf{A}^{n+1}\mathbf{v}}{||\mathbf{A}^{n+1}\mathbf{v}||}.$$

Utilizzando le definizioni dell'algoritmo (2.8) e ciò che abbiamo supposto vero nell'ipotesi induttiva, otteniamo:

$$\varphi_{n+1} = \frac{\mathbf{u}^H\mathbf{a}_{n+1}}{\mathbf{u}^H\mathbf{v}_n} = \frac{\mathbf{u}^H\mathbf{Av}_n}{\mathbf{u}^H\mathbf{v}_n} = \frac{\mathbf{u}^H\mathbf{A}(\mathbf{A}^n\mathbf{v})}{\mathbf{u}^H\mathbf{A}^n\mathbf{v}} = \frac{\mathbf{u}^H\mathbf{A}^{n+1}\mathbf{v}}{\mathbf{u}^H\mathbf{A}^n\mathbf{v}}$$

$$\mathbf{v}_{n+1} = \frac{\mathbf{a}_{n+1}}{||\mathbf{a}_{n+1}||} = \frac{\mathbf{Av}_n}{||\mathbf{Av}_n||} = \frac{\mathbf{A}^{n+1}\mathbf{v}}{||\mathbf{A}^{n+1}\mathbf{v}||}$$

la tesi è quindi verificata.

$\square$

Osservazione 2.1. La velocità di convergenza del metodo va come:

$$\Big(\max_{j:\lambda_j\neq\lambda_1}\Big(\frac{|\lambda_j|}{|\lambda_1|}\Big)\Big)^k.$$

Osservazione 2.2. Il costo per ogni passo è dato dal prodotto della matrice $\mathbf{A}$ per un vettore. In generale quindi è n^2, ma nel caso in cui $\mathbf{A}$ ha qualche particolare struttura, tale costo si può ridurre. Ad esempio, se $\mathbf{A}$ è di Toeplitz (vedi Definizione 1.16.) allora $\mathbf{A}$ può essere immersa in una matrice circolante di ordine $2n$ e quindi, il prodotto di $\mathbf{A}$ per un vettore è calcolabile con $O(n \log n)$ operazioni aritmetiche (vedi [5]).

2.1.1 A definita positiva

Come conseguenza del Teorema 2.1., scegliendo $||.|| = ||.||_2$ e $\mathbf{u} = \mathbf{v}$, si può enunciare il seguente risultato nel caso particolare in cui la matrice sia definita positiva.

Corollario 2.1. *Sia* $\mathbf{A} \in \mathbb{C}^{n\times n}$ *definita positiva, siano* $0 < \lambda_n \leq \cdots \leq \lambda_{t+1} < \lambda_t = \cdots = \lambda_1$ *i suoi autovalori e* $\mathbf{x}_i$ *autovettori corrispondenti ai* λ_i *per* $i = 1, 2, \ldots$.
Sia $\mathbf{v} \in \mathbb{C}^n$, $||\mathbf{v}||_2 = 1$, *tale che nella rappresentazione di* $\mathbf{v}$, $\mathbf{v} = \sum_{i\leq t} \alpha_i \mathbf{x}_i + \sum_{i>t} \alpha_i \mathbf{x}_i$, *il vettore* $\mathbf{x} = \sum_{i\leq t} \alpha_i \mathbf{x}_i$ *sia non nullo (nota: ciò implica che* $\mathbf{A}^k\mathbf{v} \neq \mathbf{0}$, $\mathbf{v}^H\mathbf{x} = (||\mathbf{x}||_2)^2 \neq 0$ *(vedi Osservazione 1.8.)).*
Sia anche $\mathbf{u} = \mathbf{v} \in \mathbb{C}^n$, *(nota:* $\mathbf{v}^H\mathbf{x} \neq 0$ *e* $0 < \frac{1}{\lambda_1^k}(\mathbf{v}^H\mathbf{A}^k\mathbf{v}) \xrightarrow[k\to\infty]{} \mathbf{v}^H\mathbf{x}$ $\forall k \in \mathbb{N}$*). Allora:*

$$\frac{\mathbf{A}^k\mathbf{v}}{||\mathbf{A}^k\mathbf{v}||_2} = \frac{|\lambda_1|^k}{\lambda_1^k}\frac{\mathbf{A}^k\mathbf{v}}{||\mathbf{A}^k\mathbf{v}||_2} \xrightarrow[k\to\infty]{} \frac{\mathbf{x}}{||\mathbf{x}||_2} \tag{2.9}$$

$$\frac{\mathbf{v}^H\mathbf{A}^{k+1}\mathbf{v}}{\mathbf{v}^H\mathbf{A}^k\mathbf{v}} \xrightarrow[k\to\infty]{} \lambda_1 \tag{2.10}$$

dove le successioni in (2.9) e (2.10) sono ben definite $\forall k \in \mathbb{N}$.
Le quantità in (2.9) e (2.10) si possono calcolare tramite il seguente algoritmo:

$$\mathbf{v}_0 := \mathbf{v}. \ \textit{Per } k = 1, 2, \ldots$$

$$\begin{cases} \mathbf{a}_k := \mathbf{A}\mathbf{v}_{k-1} \\ \varphi_k := \dfrac{\mathbf{v}^H\mathbf{a}_k}{\mathbf{v}^H\mathbf{v}_{k-1}}, \quad \mathbf{v}^H\mathbf{v}_{k-1} \neq 0 \ \forall k \\ \mathbf{v}_k := \dfrac{\mathbf{a}_k}{||\mathbf{a}_k||_2} \end{cases} \tag{2.11}$$

Infatti, $\varphi_k = \dfrac{\mathbf{v}^H\mathbf{A}^k\mathbf{v}}{\mathbf{v}^H\mathbf{A}^{k-1}\mathbf{v}}$ *(e quindi* $\varphi_k \xrightarrow[k\to\infty]{} \lambda_1$*) e* $\mathbf{v}_k = \dfrac{\mathbf{A}^k\mathbf{v}}{||\mathbf{A}^k\mathbf{v}||_2}$ $\forall k = 1, 2, ...$ *(e quindi* $\mathbf{v}_k \xrightarrow[k\to\infty]{} \dfrac{\mathbf{x}}{||\mathbf{x}||_2}$*). Inoltre:*

$$\varphi_k \leq \varphi_{k+1} \leq \lambda_1 = \rho(\mathbf{A}) \ \forall k = 1, 2, \dots . \tag{2.12}$$

cioè la successione φ_k *converge in modo monotono a* $\lambda_1 = \rho(\mathbf{A})$*.*

Dimostrazione. Rimane solo da dimostrare che la successione φ_k ottenuta in (2.11) è crescente.
Utilizzando la definizione di φ_k in (2.11) si ha che $\varphi_k \leq \varphi_{k+1}$ se e solo se:

$$\frac{\mathbf{v}^H\mathbf{A}^k\mathbf{v}}{\mathbf{v}^H\mathbf{A}^{k-1}\mathbf{v}} \leq \frac{\mathbf{v}^H\mathbf{A}^{k+1}\mathbf{v}}{\mathbf{v}^H\mathbf{A}^k\mathbf{v}}$$

cioè se e solo se moltiplicando il membro di destra per $\mathbf{v}^H\mathbf{A}^{k-1}\mathbf{v}$ e quello di sinistra per $\mathbf{v}^H\mathbf{A}^k\mathbf{v}$:

$$(\mathbf{v}^H\mathbf{A}^k\mathbf{v})^2 \leq (\mathbf{v}^H\mathbf{A}^{k-1}\mathbf{v})(\mathbf{v}^H\mathbf{A}^{k+1}\mathbf{v}). \tag{2.13}$$

Ma la (2.13) risulta verificata perché equivale alla seguente disuguaglianza di Cauchy-Schwarz:

$$\left(\mathbf{v}^H\mathbf{A}^{\frac{k-1}{2}}\mathbf{A}^{\frac{k+1}{2}}\mathbf{v}\right)^2 \leq \left(||\mathbf{A}^{\frac{k-1}{2}}\mathbf{v}||_2\right)^2\left(||\mathbf{A}^{\frac{k+1}{2}}\mathbf{v}||_2\right)^2$$

(nota: è possibile considerare la radice quadrata della matrice $\mathbf{A}$ in quanto quest'ultima per ipotesi è definita positiva (vedi Osservazione 1.6.)).

□

Abbiamo verificato quindi che, nel caso in cui $\mathbf{A}$ è matrice definita positiva, si genera una successione φ_k che tende in modo monotono, in questo caso crescente, all'autovalore dominante di $\mathbf{A}$.

Osservazione 2.3. Se si volesse invece calcolare l'autovalore minimo, che chiamiamo λ_{min}, della matrice $\mathbf{A}$ definita positiva, basta applicare il Corollario precedente alla sua inversa, ovvero ad $\mathbf{A}^{-1}$.
In particolare, in tal caso, si produce una successione φ_k che tende in modo crescente a $\dfrac{1}{\lambda_{min}}$, ovvero, una successione $\dfrac{1}{\varphi_k}$ che tende in modo decrescente a λ_{min} di $\mathbf{A}$. Il costo per ogni passo è dato dal prodotto della matrice $\mathbf{A}^{-1}$ per un vettore, ovvero risolvere un sistema lineare con $\mathbf{A}$ matrice dei coefficienti. In generale quindi è $O(n^3)$. Però, se $\mathbf{A}$ ad esempio è di Toeplitz allora il costo si riduce a $O(n^2)$(vedi [2]).

2.2 Caso A generica

Sia $\mathbf{A} \in \mathbb{C}^{n \times n}$ generica e sia $\mathbf{X} \in \mathbb{C}^{n \times s}$ una matrice del tipo:

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \dots & \mathbf{x}_s \end{bmatrix}$$

tale che:

$$\mathbf{AX} = \mathbf{X} \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{bmatrix} \tag{2.14}$$

ovvero:

$$\begin{aligned} \mathbf{Ax}_1 &= \lambda \mathbf{x}_1 \\ \mathbf{Ax}_2 &= \mathbf{x}_1 + \lambda \mathbf{x}_2 \\ \mathbf{Ax}_3 &= \mathbf{x}_2 + \lambda \mathbf{x}_3 \\ &\vdots \\ \mathbf{Ax}_s &= \mathbf{x}_{s-1} + \lambda \mathbf{x}_s \end{aligned}$$

Poiché ci servirà successivamente, vogliamo ora ottenere una rappresentazione di $\mathbf{A}^k \mathbf{x}_j$ con: $1 \leq j \leq s$.

Per $j = 1$ si ha immediatamente che:

$$\mathbf{A}^k \mathbf{x}_1 = \lambda^k \mathbf{x}_1$$

Vediamo il caso in cui $j = 2$:

$$\begin{aligned} \mathbf{A}^2 \mathbf{x}_2 = \ & \mathbf{A}(\mathbf{Ax}_2) = \mathbf{A}(\mathbf{x}_1 + \lambda \mathbf{x}_2) = \mathbf{Ax}_1 + \lambda \mathbf{Ax}_2 \\ & = \lambda \mathbf{x}_1 + \lambda \mathbf{x}_1 + \lambda^2 \mathbf{x}_2 = 2\lambda \mathbf{x}_1 + \lambda^2 \mathbf{x}_2 \end{aligned}$$

$$\begin{aligned} \mathbf{A}^3 \mathbf{x}_2 = \ & \mathbf{A}(\mathbf{A}^2 \mathbf{x}_2) = 2\lambda \mathbf{Ax}_1 + \lambda^2 \mathbf{Ax}_2 = 2\lambda^2 \mathbf{x}_1 + \lambda^2 \mathbf{x}_1 + \lambda^3 \mathbf{x}_2 \\ & = 3\lambda^2 \mathbf{x}_1 + \lambda^3 \mathbf{x}_2 \end{aligned}$$

$$\begin{aligned} \mathbf{A}^4 \mathbf{x}_2 = \ & \mathbf{A}(\mathbf{A}^3 \mathbf{x}_2) = \mathbf{A}(3\lambda^2 \mathbf{x}_1 + \lambda^3 \mathbf{x}_2) = 3\lambda^2 \mathbf{Ax}_1 + \lambda^3 \mathbf{Ax}_2 \\ & = 3\lambda^3 \mathbf{x}_1 + \lambda^3 \mathbf{x}_1 + \lambda^4 \mathbf{x}_2 = 4\lambda^3 \mathbf{x}_1 + \lambda^4 \mathbf{x}_2 \end{aligned}$$

Dalle precedenti uguaglianze possiamo congetturare:

$$\mathbf{A}^k\mathbf{x}_2 = k\lambda^{k-1}\mathbf{x}_1 + \lambda^k\mathbf{x}_2,\ k \geq 1 \tag{2.15}$$

(nota: (2.15) è vera anche per $k = 0$ se $\lambda \neq 0$).

Dimostrazione. Dimostriamo la congettura procedendo per induzione.
Sia $k = 1$, allora per quanto già visto precedentemente:

$$\mathbf{A}\mathbf{x}_2 = \mathbf{x}_1 + \lambda\mathbf{x}_2.$$

Supponiamo ora che sia vera per $k \leq n$, in particolare:

$$\mathbf{A}^n\mathbf{x}_2 = n\lambda^{n-1}\mathbf{x}_1 + \lambda^n\mathbf{x}_2.$$

Dimostriamola per $k = n + 1$:

$$\begin{aligned}\mathbf{A}^{n+1}\mathbf{x}_2 =&\ \mathbf{A}(\mathbf{A}^n\mathbf{x}_2) = \mathbf{A}(n\lambda^{n-1}\mathbf{x}_1 + \lambda^n\mathbf{x}_2)\\ &= n\lambda^{n-1}\mathbf{A}\mathbf{x}_1 + \lambda^n\mathbf{A}\mathbf{x}_2\\ &= n\lambda^{n-1}(\lambda\mathbf{x}_1) + \lambda^n(\mathbf{x}_1 + \lambda\mathbf{x}_2)\\ &= n\lambda^n\mathbf{x}_1 + \lambda^n\mathbf{x}_1 + \lambda^{n+1}\mathbf{x}_2\\ &= (n+1)\lambda^n\mathbf{x}_1 + \lambda^{n+1}\mathbf{x}_2\end{aligned}$$

la congettura è quindi verificata.

□

Caso $j = 3$:

$$\begin{aligned}\mathbf{A}^2\mathbf{x}_3 =&\ \mathbf{A}(\mathbf{A}\mathbf{x}_3) = \mathbf{A}(\mathbf{x}_2 + \lambda\mathbf{x}_3)\\ &= \mathbf{A}\mathbf{x}_2 + \lambda\mathbf{A}\mathbf{x}_3 = \mathbf{x}_1 + \lambda\mathbf{x}_2 + \lambda\mathbf{x}_2 + \lambda^2\mathbf{x}_3\\ &= \mathbf{x}_1 + 2\lambda\mathbf{x}_2 + \lambda^2\mathbf{x}_3\end{aligned}$$

$$\begin{aligned}\mathbf{A}^3\mathbf{x}_3 =&\ \mathbf{A}(\mathbf{A}^2\mathbf{x}_3) = \mathbf{A}\mathbf{x}_1 + 2\lambda\mathbf{A}\mathbf{x}_2 + \lambda^2\mathbf{A}\mathbf{x}_3\\ &= \lambda\mathbf{x}_1 + 2\lambda(\mathbf{x}_1 + \lambda\mathbf{x}_2) + \lambda^2(\mathbf{x}_2 + \lambda\mathbf{x}_3)\\ &= \lambda\mathbf{x}_1 + 2\lambda\mathbf{x}_1 + 2\lambda^2\mathbf{x}_2 + \lambda^2\mathbf{x}_2 + \lambda^3\mathbf{x}_3\\ &= 3\lambda\mathbf{x}_1 + 3\lambda^2\mathbf{x}_2 + \lambda^3\mathbf{x}_3\end{aligned}$$

$$\begin{aligned}\mathbf{A}^4\mathbf{x}_3 =&\ \mathbf{A}(\mathbf{A}^3\mathbf{x}_3) = \mathbf{A}(3\lambda\mathbf{x}_1 + 3\lambda^2\mathbf{x}_2 + \lambda^3\mathbf{x}_3)\\ &= 3\lambda\mathbf{A}\mathbf{x}_1 + 3\lambda^2\mathbf{A}\mathbf{x}_2 + \lambda^3\mathbf{A}\mathbf{x}_3\\ &= 3\lambda(\lambda\mathbf{x}_1) + 3\lambda^2(\mathbf{x}_1 + \lambda\mathbf{x}_2) + \lambda^3(\mathbf{x}_2 + \lambda\mathbf{x}_3)\\ &= 3\lambda^2\mathbf{x}_1 + 3\lambda^2\mathbf{x}_1 + 3\lambda^3\mathbf{x}_2 + \lambda^3\mathbf{x}_2 + \lambda^4\mathbf{x}_3\\ &= 6\lambda^2\mathbf{x}_1 + 4\lambda^3\mathbf{x}_2 + \lambda^4\mathbf{x}_3\end{aligned}$$

Da quanto ottenuto possiamo congetturare che:

$$\mathbf{A}^k\mathbf{x}_3 = \frac{1}{2}k(k-1)\lambda^{k-2}\mathbf{x}_1 + k\lambda^{k-1}\mathbf{x}_2 + \lambda^k\mathbf{x}_3,\ k \geq 2 \tag{2.16}$$

(nota: (2.16) è vera anche per $k = 0, 1$ se $\lambda \neq 0$).

Dimostrazione. Procediamo a dimostrare la congettura per induzione.
Per $k = 2$ otteniamo che:

$$\mathbf{A}^2\mathbf{x}_3 = \mathbf{x}_1 + 2\lambda\mathbf{x}_2 + \lambda^2\mathbf{x}_3$$

e ciò è vero per quanto visto precedentemente.
Supponiamo che sia vera per $k \leq n$, in particolare:

$$\mathbf{A}^n\mathbf{x}_3 = \frac{1}{2}n(n-1)\lambda^{n-2}\mathbf{x}_1 + n\lambda^{n-1}\mathbf{x}_2 + \lambda^n\mathbf{x}_3.$$

Dimostro ora che la congettura è vera per $k = n + 1$.

$$\begin{aligned}
\mathbf{A}^{n+1}\mathbf{x}_3 =\ & \mathbf{A}(\mathbf{A}^n\mathbf{x}_3) = \mathbf{A}\left[\frac{1}{2}n(n-1)\lambda^{n-2}\mathbf{x}_1 + n\lambda^{n-1}\mathbf{x}_2 + \lambda^n\mathbf{x}_3\right] \\
=\ & \frac{1}{2}n(n-1)\lambda^{n-2}\mathbf{A}\mathbf{x}_1 + n\lambda^{n-1}\mathbf{A}\mathbf{x}_2 + \lambda^n\mathbf{A}\mathbf{x}_3 \\
=\ & \frac{1}{2}n(n-1)\lambda^{n-1}\mathbf{x}_1 + n\lambda^{n-1}(\mathbf{x}_1 + \lambda\mathbf{x}_2) + \lambda^n(\mathbf{x}_2 + \lambda\mathbf{x}_3) \\
=\ & \frac{1}{2}n(n-1)\lambda^{n-1}\mathbf{x}_1 + n\lambda^{n-1}\mathbf{x}_1 + n\lambda^n\mathbf{x}_2 + \lambda^n\mathbf{x}_2 + \lambda^{n+1}\mathbf{x}_3 \\
=\ & \frac{1}{2}n(n+1)\lambda^{n-1}\mathbf{x}_1 + (n+1)\lambda^n\mathbf{x}_2 + \lambda^{n+1}\mathbf{x}_3
\end{aligned}$$

La congettura è quindi verificata.

□

Da ciò che abbiamo verificato precedentemente, possiamo quindi congetturare il caso per j generico in $\{1, 2, \ldots, s\}$:

$$\begin{aligned}
\mathbf{A}^k\mathbf{x}_j =\ & \binom{k}{j-1}\lambda^{k-j+1}\mathbf{x}_1 + \binom{k}{j-2}\lambda^{k-j+2}\mathbf{x}_2 + ...+ \\
& + \frac{k}{2}(k-1)\lambda^{k-2}\mathbf{x}_{j-2} + k\lambda^{k-1}\mathbf{x}_{j-1} + \lambda^k\mathbf{x}_j,\ k \geq j-1
\end{aligned} \tag{2.17}$$

(nota: (2.17) è vera anche per $k = 0, 1, \ldots, j-2$ se $\lambda \neq 0$).

Dimostrazione. Nell'ipotesi (2.14) più precisamente dimostriamo questo:

$$\mathbf{A}^k\mathbf{x}_j = \begin{cases} \binom{k}{k}\lambda^0\mathbf{x}_{j-k} + \binom{k}{k-1}\lambda\mathbf{x}_{j-k+1} + \cdots + \binom{k}{1}\lambda^{k-1}\mathbf{x}_{j-1} + \binom{k}{0}\lambda^k\mathbf{x}_j & \text{se } k < j-1 \\ \binom{k}{j-1}\lambda^{k-j+1}\mathbf{x}_1 + \binom{k}{j-2}\lambda^{k-j+2}\mathbf{x}_2 + \cdots + \binom{k}{1}\lambda^{k-1}\mathbf{x}_{j-1} + \binom{k}{0}\lambda^k\mathbf{x}_j & \text{se } k \geq j-1 \end{cases} \tag{2.18}$$

Le formule del sistema (2.18) sono intuitive esaminando i casi in cui j è piccola, oppure come segue, procediamo a dimostrarle per induzione su j. Abbiamo che: $\mathbf{A}^k\mathbf{x}_1 = \lambda^k\mathbf{x}_1$ per $k \geq 0$, da cui ne segue che la (2.18) è vera per $j = 1$. Inoltre:

$$\mathbf{A}^k\mathbf{x}_2 = \begin{cases} \mathbf{x}_2 & \text{se } k = 0 \\ k\lambda^{k-1}\mathbf{x}_1 + \lambda^k\mathbf{x}_2 & \text{se } k \geq 1 \end{cases}$$

quindi la (2.18) è vera per $j = 2$.

Supponiamo che la tesi sia vera per j, si deve dimostrare ora che sia vera anche per $j+1$, $j = 1, \ldots, s-1$, cioè:

$$\mathbf{A}^k\mathbf{x}_{j+1} = \begin{cases} \binom{k}{k}\lambda^0\mathbf{x}_{j+1-k} + \binom{k}{k-1}\lambda\mathbf{x}_{j-k+2} + \cdots + \binom{k}{1}\lambda^{k-1}\mathbf{x}_j + \binom{k}{0}\lambda^k\mathbf{x}_{j+1} & \text{se } k < j \\ \binom{k}{j}\lambda^{k-j}\mathbf{x}_1 + \binom{k}{j-1}\lambda^{k-j+1}\mathbf{x}_2 + \cdots + \binom{k}{1}\lambda^{k-1}\mathbf{x}_j + \lambda^k\mathbf{x}_{j+1} & \text{se } k \geq j \end{cases}$$

ovvero che:

$$(\mathbf{A}^k - \lambda^k\mathbf{I})\mathbf{x}_{j+1} = \begin{cases} \binom{k}{k}\lambda^0\mathbf{x}_{j+1-k} + \binom{k}{k-1}\lambda\mathbf{x}_{j-k+2} + \cdots + \binom{k}{1}\lambda^{k-1}\mathbf{x}_j & \text{se } k < j \\ \binom{k}{j}\lambda^{k-j}\mathbf{x}_1 + \binom{k}{j-1}\lambda^{k-j+1}\mathbf{x}_2 + \cdots + \binom{k}{1}\lambda^{k-1}\mathbf{x}_j & \text{se } k \geq j \end{cases}.$$

Si nota innanzitutto che:

$$\begin{aligned} &(\mathbf{A}^{k-1} + \lambda\mathbf{A}^{k-2} + \lambda^2\mathbf{A}^{k-3} + \cdots + \lambda^{k-2}\mathbf{A} + \lambda^{k-1}\mathbf{I})(\mathbf{A} - \lambda\mathbf{I}) \\ &= \mathbf{A}^k - \lambda\mathbf{A}^{k-1} + \lambda\mathbf{A}^{k-1} - \lambda^2\mathbf{A}^{k-2} + \lambda^2\mathbf{A}^{k-2} - \lambda^3\mathbf{A}^{k-3} + \\ &\quad + \cdots + \lambda^{k-2}\mathbf{A}^2 - \lambda^{k-1}\mathbf{A} + \lambda^{k-1}\mathbf{A} - \lambda^k\mathbf{I} = \mathbf{A}^k - \lambda^k\mathbf{I} \end{aligned}$$

e $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x}_{j+1} = \mathbf{x}_j$ per $j = 1, 2, \ldots, s-1$.

Da questo ne segue che:

$$\begin{aligned} (\mathbf{A}^k - \lambda^k\mathbf{I})\mathbf{x}_{j+1} &= (\mathbf{A}^{k-1} + \lambda\mathbf{A}^{k-2} + \lambda^2\mathbf{A}^{k-3} + \cdots + \lambda^{k-2}\mathbf{A} + \lambda^{k-1}\mathbf{I})\mathbf{x}_j \\ &= \mathbf{A}^{k-1}\mathbf{x}_j + \lambda\mathbf{A}^{k-2}\mathbf{x}_j + \lambda^2\mathbf{A}^{k-3}\mathbf{x}_j + \cdots + \lambda^{k-2}\mathbf{A}\mathbf{x}_j + \lambda^{k-1}\mathbf{x}_j. \end{aligned}$$

Ora, per quest'ultimo risultato ottenuto, nei due casi distinti si ha:

Primo caso: $k < j$

In tal caso, per tutti gli addendi utilizziamo la prima equazione della (2.18), ottenendo:

$$\begin{aligned}
(\mathbf{A}^k - \lambda^k\mathbf{I})\mathbf{x}_{j+1} &= \binom{k-1}{k-1}\lambda^0\mathbf{x}_{j-k+1} + \binom{k-1}{k-2}\lambda\mathbf{x}_{j-k+2} + \cdots + \binom{k-1}{0}\lambda^{k-1}\mathbf{x}_j \\
&+ \lambda\left[\binom{k-2}{k-2}\lambda^0\mathbf{x}_{j-k+2} + \binom{k-2}{k-3}\lambda\mathbf{x}_{j-k+3} + \dots\right. \\
&+ \left.\binom{k-2}{0}\lambda^{k-2}\mathbf{x}_j\right] + \lambda^2\left[\binom{k-3}{k-3}\lambda^0\mathbf{x}_{j-k+3} + \binom{k-3}{k-4}\lambda\mathbf{x}_{j-k+4}\right. \\
&+ \left.\binom{k-3}{0}\lambda^{k-3}\mathbf{x}_j\right] + \cdots + \lambda^{k-2}\left[\binom{1}{1}\lambda^0\mathbf{x}_{j-1} + \binom{1}{0}\lambda\mathbf{x}_j\right] \\
&+ \lambda^{k-1}\binom{0}{0}\lambda^0\mathbf{x}_j \\
&= \binom{k-1}{k-1}\mathbf{x}_{j-k+1} + \lambda\mathbf{x}_{j-k+2}\left[\binom{k-2}{k-2} + \binom{k-1}{k-2}\right] + \\
&+ \lambda^2\mathbf{x}_{j-k+3}\left[\binom{k-3}{k-3} + \binom{k-2}{k-3} + \binom{k-1}{k-3}\right] + \dots \\
&+ \lambda^{k-2}\mathbf{x}_{j-1}\left[\binom{1}{1} + \binom{2}{1}\cdots + \binom{k-1}{1}\right] + \\
&+ \lambda^{k-1}\mathbf{x}_j\left[\binom{0}{0} + \binom{1}{0} + \cdots + \binom{k-1}{0}\right] \\
&= \binom{k}{k}\mathbf{x}_{j-k+1} + \lambda\mathbf{x}_{j-k+2}\binom{k}{k-1} + \lambda^2\mathbf{x}_{j-k+3}\binom{k}{k-2} + \dots \\
&+ \lambda^{k-2}\mathbf{x}_{j-1}\binom{k}{2} + \lambda^{k-1}\mathbf{x}_j\binom{k}{1}
\end{aligned}$$

(nota: nell'ultimo passaggio abbiamo usato il fatto che: $\binom{k}{s} = \sum_{j=s-1}^{k-1} \binom{j}{s-1}$).

Secondo caso: $k \geq j$

In questo caso, per i primi $k-j+1$ addenti utilizziamo la seconda equazione della (2.18) e per i rimanenti successivi, la prima. Ottenendo:

$$\begin{aligned}
(\mathbf{A}^k - \lambda^k\mathbf{I})\mathbf{x}_{j+1} &= \binom{k-1}{j-1}\lambda^{k-j}\mathbf{x}_1 + \binom{k-1}{j-2}\lambda^{k-j+1}\mathbf{x}_2 + \cdots + \binom{k-1}{2}\lambda^{k-3}\mathbf{x}_{j-2} + \\
&+ \binom{k-1}{1}\lambda^{k-2}\mathbf{x}_{j-1} + \lambda^{k-1}\mathbf{x}_j + \lambda\left[\binom{k-2}{j-1}\lambda^{k-j-1}\mathbf{x}_1 + \right. \\
&+ \binom{k-2}{j-2}\lambda^{k-j}\mathbf{x}_2 + \cdots + \binom{k-2}{2}\lambda^{k-4}\mathbf{x}_{j-2} + \binom{k-2}{1}\lambda^{k-3}\mathbf{x}_{j-1} +
\end{aligned}$$

$$
\begin{aligned}
&+\lambda^{k-2}\mathbf{x}_j\Big] + \cdots + \lambda^{k-j}\left[\binom{j-1}{j-1}\lambda^0\mathbf{x}_1 + \binom{j-1}{j-2}\lambda\mathbf{x}_2 + \cdots +\right.\\
&\left.+\binom{j-1}{2}\lambda^{j-3}\mathbf{x}_{j-2} + \binom{j-1}{1}\lambda^{j-2}\mathbf{x}_{j-1} + \lambda^{j-1}\mathbf{x}_j\right]+\\
&+\lambda^{k-j+1}\left[\binom{j-2}{j-2}\lambda^0\mathbf{x}_2 + \binom{j-2}{j-3}\lambda\mathbf{x}_3 + \cdots + \binom{j-2}{1}\lambda^{j-3}\mathbf{x}_{j-1}+\right.\\
&\left.+\binom{j-2}{0}\lambda^{j-2}\mathbf{x}_j\right] + \lambda^{k-j+2}\left[\binom{j-3}{j-3}\lambda^0\mathbf{x}_3 + \binom{j-3}{j-4}\lambda\mathbf{x}_4+\right.\\
&\left.+\cdots+\binom{j-3}{1}\lambda^{j-4}\mathbf{x}_{j-1} + \binom{j-3}{0}\lambda^{j-3}\mathbf{x}_j\right] + \cdots +\\
&+\lambda^{k-3}\left[\binom{2}{2}\lambda^0\mathbf{x}_{j-2} + \binom{2}{1}\lambda\mathbf{x}_{j-1} + \binom{2}{0}\lambda^2\mathbf{x}_j\right]+\\
&+\lambda^{k-2}\left[\binom{1}{1}\lambda^0\mathbf{x}_{j-1} + \binom{1}{0}\lambda\mathbf{x}_j\right] + \lambda^{k-1}\left[\binom{0}{0}\lambda^0\mathbf{x}_j\right]\\
&=\lambda^{k-1}\mathbf{x}_j\left[\binom{0}{0} + \binom{1}{0} + \cdots + \binom{j-3}{0} + \binom{j-2}{0} + k-j+1\right]+\\
&+\lambda^{k-2}\mathbf{x}_{j-1}\left[\binom{1}{1} + \binom{2}{1} + \cdots + \binom{j-3}{1} + \binom{j-2}{1} + \binom{j-1}{1}\right]+\\
&+\cdots+\binom{k-2}{1} + \binom{k-1}{1}\Big] + \lambda^{k-3}\mathbf{x}_{j-2}\left[\binom{2}{2} + \binom{3}{2} + \cdots +\right.\\
&\left.+\binom{j-3}{2} + \binom{j-2}{2} + \binom{j-1}{2} + \cdots + \binom{k-2}{2} + \binom{k-1}{2}\right]+\\
&+\cdots+\lambda^{k-j+1}\mathbf{x}_2\left[\binom{j-2}{j-2} + \binom{j-1}{j-2} + \cdots + \binom{k-2}{j-2}+\right.\\
&\left.+\binom{k-1}{j-2}\right] + \lambda^{k-j}\mathbf{x}_1\left[\binom{j-1}{j-1} + \cdots + \binom{k-2}{j-1} + \binom{k-1}{j-1}\right]\\
&=\lambda^{k-1}\mathbf{x}_j\binom{k}{1} + \lambda^{k-2}\mathbf{x}_{j-1}\binom{k}{2} + \lambda^{k-3}\mathbf{x}_{j-2}\binom{k}{3} + \cdots +\\
&\lambda^{k-j+1}\mathbf{x}_2\binom{k}{j-1} + \lambda^{k-j}\mathbf{x}_1\binom{k}{j}.
\end{aligned}
$$

Quindi in entrambi i casi, la tesi risulta verificata.

□

Dal risultato appena dimostrato, possiamo enunciare il seguente corollario.

Corollario 2.2. *i) Nell'ipotesi* (2.14) *vale* (2.18) *per* $j \in \{1, 2, \ldots, s\}$.
ii) Sia $k < s - 1$. *Sappiamo allora che vale la seconda rappresentazione in*

(2.18) *di* $\mathbf{A}^k\mathbf{x}_j$ *per* $j = 1, 2, \ldots, k+1$ *mentre vale la prima rappresentazione per* $j = k+2, \ldots, s$*, cioè:*

$$\mathbf{A}^k \begin{bmatrix} \\ \mathbf{x}_1 & \ldots & \mathbf{x}_k & \mathbf{x}_{k+1} & \mathbf{x}_{k+2} & \ldots & \mathbf{x}_s \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \mathbf{x}_1 & \ldots & \mathbf{x}_k & \mathbf{x}_{k+1} & \mathbf{x}_{k+2} & \ldots & \mathbf{x}_s \\ \\ \end{bmatrix}$$

$$\cdot \begin{bmatrix} \binom{k}{0}\lambda^k & \ldots & \binom{k}{k-1}\lambda & \binom{k}{k} & 0 & \ldots & 0 \\ 0 & \ddots & \vdots & \binom{k}{k-1}\lambda & \binom{k}{k} & \ddots & \vdots \\ \vdots & \ddots & \binom{k}{0}\lambda^k & \vdots & \ddots & \ddots & 0 \\ \vdots & & 0 & \binom{k}{0}\lambda^k & & \ddots & \binom{k}{k} \\ \vdots & & & 0 & \ddots & & \binom{k}{k-1}\lambda \\ \vdots & & & & \ddots & \ddots & \vdots \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 & \binom{k}{0}\lambda^k \end{bmatrix}.$$

Sia invece $k \geq s-1$*. Sappiamo che vale la seconda rappresentazione in* (2.18) *di* $\mathbf{A}^k\mathbf{x}_j$ *per* $j = 1, 2, \ldots, s$ *(cioè per tutte le colonne* $\mathbf{x}_1, \ldots, \mathbf{x}_s$ *relative al blocco di Jordan di* λ*), cioè:*

$$\mathbf{A}^k \begin{bmatrix} \\ \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_j & \ldots & \mathbf{x}_s \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_j & \ldots & \mathbf{x}_s \\ \\ \end{bmatrix}$$

$$\cdot \begin{bmatrix} \binom{k}{0}\lambda^k & \binom{k}{1}\lambda^{k-1} & \ldots & \binom{k}{j-1}\lambda^{k-j+1} & \ldots & \binom{k}{s-1}\lambda^{k-s+1} \\ 0 & \binom{k}{0}\lambda^k & \binom{k}{1}\lambda^{k-1} & \ldots & \ldots & \binom{k}{s-2}\lambda^{k-s+2} \\ \vdots & 0 & \binom{k}{0}\lambda^k & \ddots & & \vdots \\ \vdots & & 0 & \binom{k}{0}\lambda^k & \ddots & \vdots \\ \vdots & & & 0 & \ddots & \binom{k}{1}\lambda^{k-1} \\ 0 & \ldots & \ldots & \ldots & 0 & \binom{k}{0}\lambda^k \end{bmatrix}.$$

$\square$

Sia $\mathbf{J}$ la forma canonica di Jordan di $\mathbf{A}$ allora, $\exists \mathbf{X} \in \mathbb{C}^{n\times n}$ non singolare: $\mathbf{AX} = \mathbf{XJ}$. Siano $\mathbf{x}_{i+1}, \ldots, \mathbf{x}_{i+s}$ le colonne di $\mathbf{X}$ corrispondenti al generico blocco di $\mathbf{J}$ di ordine s di $\mathbf{A}$. Dalla (2.18) possiamo dire che, per k abbastanza grande, $k \geq s-1$, è vera la seguente relazione:

$$\begin{aligned}\mathbf{A}^k\mathbf{x}_{i+j} = &\binom{k}{j-1}\lambda^{k-j+1}\mathbf{x}_{1+i} + \binom{k}{j-2}\lambda^{k-j+2}\mathbf{x}_{2+i}\ldots + \frac{k}{2}(k-1)\lambda^{k-2}\mathbf{x}_{j-2+i} \\ &+ k\lambda^{k-1}\mathbf{x}_{j-1+i} + \lambda^k\mathbf{x}_{j+i} = \sum_{r=1}^{j}\mathbf{x}_{i+r}\lambda^{k+r-j}\binom{k}{j-r}, \ j=1,\ldots,s.\end{aligned} \tag{2.19}$$

Per $k \geq n-1$ la (2.19) è vera per le colonne di $\mathbf{X}$ corrispondenti ad ogni blocco di Jordan di $\mathbf{A}$.

Dato un autovalore λ generico, per comodità chiamiamo g la sua molteplicità geometrica (cioè il numero di blocchi di Jordan ad esso associati) ed a quella algebrica.
Assumiamo che il generico blocco di Jordan associato a λ occupi le posizioni i_l+1 e i_l+s_l quindi, s_l è la dimensione del blocco mentre la somma degli s_l per $l=1,\ldots,g$ è uguale ad a.
Sia ora $\mathbf{v} \in \mathbb{C}^n$, $||\mathbf{v}||=1$, allora $\mathbf{v}$ si può rappresentare in questo modo:

$$\mathbf{v} = \cdots + \sum_{l=1}^{g}\sum_{j=1}^{s_l}\alpha_{i_l+j}\mathbf{x}_{i_l+j} + \cdots + \mathbf{x} \tag{2.20}$$

dove $\mathbf{x} = \sum_{m=1}^{t}\alpha_m\mathbf{x}_m$ e $t = m_a(\lambda_1) = m_g(\lambda_1)$.
Si suppone che un autovalore dominante di $\mathbf{A}$, λ_1, abbia molteplicità geometrica ed algebrica coincidenti e che occupi la parte in alto della forma canonica di Jordan di $\mathbf{A}$ ($J_{k,k} = \lambda_1, \ k=1,\ldots,t$).
Moltiplichiamo ora entrambi i membri della (2.20) per $\mathbf{A}^k$, ottenendo quindi:

$$\begin{aligned}\mathbf{A}^k\mathbf{v} = \ &\mathbf{A}^k\mathbf{x} + \mathbf{A}^k(\cdots + \sum_{l=1}^{g}\sum_{j=1}^{s_l}\alpha_{i_l+l}\mathbf{x}_{i_l+j} + \ldots) \\ &= \lambda_1^k\mathbf{x} + (\cdots + \sum_{l=1}^{g}\sum_{j=1}^{s_l}\alpha_{i_l+j}\mathbf{A}^k\mathbf{x}_{i_l+j} + \ldots)\end{aligned}$$

Utilizzando poi la relazione (2.19) su ciò che abbiamo appena ottenuto, si ha:

$$\mathbf{A}^k\mathbf{v} = \lambda_1^k\mathbf{x} + \cdots + \sum_{l=1}^{g}\sum_{j=1}^{s_l}\alpha_{i_l+j}\sum_{r=1}^{j}\mathbf{x}_{i_l+r}\lambda^{k+r-j}\binom{k}{j-r} + \ldots \tag{2.21}$$

Ora dividiamo entrambi i membri per λ_1^k:

$$\frac{\mathbf{A}^k\mathbf{v}}{\lambda_1^k} = \mathbf{x} + \cdots + \sum_{l=1}^{g}\sum_{j=1}^{s_l}\alpha_{i_l+j}\sum_{r=1}^{j}\mathbf{x}_{i_l+r}\frac{\lambda^{k+r-j}}{\lambda_1^k}\binom{k}{j-r} + \dots \tag{2.22}$$

Infine, passando al limite per $k \to \infty$ (nota: $\left(\frac{\lambda}{\lambda_1}\right)^k \lambda^{r-j} \xrightarrow[k\to\infty]{} 0$ e $\binom{k}{j-r} \xrightarrow[k\to\infty]{}$ $+\infty$, dove il primo fra i due converge più velocemente), si ottiene:

$$\frac{1}{\lambda_1^k}(\mathbf{A}^k\mathbf{v}) \xrightarrow[k\to\infty]{} \mathbf{x} \tag{2.23}$$

Dalla (2.23) otteniamo:

$$\frac{||\mathbf{A}^k\mathbf{v}||}{|\lambda_1|^k} \xrightarrow[k\to\infty]{} ||\mathbf{x}|| \tag{2.24}$$

Considerando sia la (2.23) che la (2.24) si ha una successione di vettori approssimante $\mathbf{x}$ normalizzato:

$$\frac{|\lambda_1|^k}{\lambda_1^k}\frac{\mathbf{A}^k\mathbf{v}}{||\mathbf{A}^k\mathbf{v}||} \xrightarrow[k\to\infty]{} \frac{\mathbf{x}}{||\mathbf{x}||} \tag{2.25}$$

Inoltre la (2.22) implica pure che, $\forall$ vettore $\mathbf{u} \in \mathbb{C}^n$:

$$\frac{1}{\lambda_1^k}(\mathbf{u}^H\mathbf{A}^k\mathbf{v}) \xrightarrow[k\to\infty]{} \mathbf{u}^H\mathbf{x}$$

Allora sarà anche vero che:

$$\frac{1}{\lambda_1^{k+1}}(\mathbf{u}^H\mathbf{A}^{k+1}\mathbf{v}) \xrightarrow[k\to\infty]{} \mathbf{u}^H\mathbf{x} \tag{2.26}$$

e quindi, $\mathbf{x}\mathbf{u}^H\mathbf{x} \neq 0$:

$$\frac{\mathbf{u}^H\mathbf{A}^{k+1}\mathbf{v}}{\mathbf{u}^H\mathbf{A}^k\mathbf{v}} \xrightarrow[k\to\infty]{} \lambda_1 \tag{2.27}$$

Da cui possiamo osservare che con il metodo delle potenze è possibile calcolare anche l'autovalore dominante.

Teorema 2.2. *Sia* $\mathbf{A} \in \mathbb{C}^{n\times n}$ *generica. Allora, ne segue che esiste* $\mathbf{X} \in \mathbb{C}^{n\times n}$ *non singolare (cioè* $det\mathbf{X} \neq 0$*) tale che:*

$$\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 & & & & & & \\ 0 & \ddots & \ddots & \vdots & & & & & & \\ \vdots & \ddots & \ddots & 0 & & & & & & \\ 0 & \dots & 0 & \lambda_1 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & \lambda & 1 & & & \\ & & & & & & \ddots & \ddots & & \\ & & & & & & & \ddots & 1 & \\ & & & & & & & & \lambda & \\ & & & & & & & & & \ddots \end{bmatrix} =: \mathbf{J}$$

dove $\mathbf{J}$ *è una matrice diagonale a blocchi (nota: ci possono essere più blocchi diagonali con lo stesso* λ*), in particolare si ha che:*

$$\mathbf{A}\begin{bmatrix} \\ \mathbf{x}_{i_l+1} & \dots & \dots & \mathbf{x}_{i_l+s_l} \\ \\ \end{bmatrix} =$$

$$\begin{bmatrix} \\ \mathbf{x}_{i_l+1} & \dots & \dots & \mathbf{x}_{i_l+s_l} \\ \\ \end{bmatrix} \cdot \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{bmatrix} \in \mathbb{C}^{n \times s_l}$$

con s_l*:* $1 \leq s_l < n$*, e* $\lambda \in \sigma(\mathbf{A})$*.*
Supponiamo $|\lambda_1| > |\lambda_j|$*,* $\forall \lambda_j \neq \lambda_1$ *e* $m_a(\lambda_1) = m_g(\lambda_1)$*.*
Sia $\mathbf{v} \in \mathbb{C}^n$*,* $||\mathbf{v}|| = 1$*, tale che nella rappresentazione di* $\mathbf{v}$*,* $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{x}_i$*, il vettore* $\mathbf{x} = \sum_{m=1}^{t} \alpha_m \mathbf{x}_m$ *sia non nullo (nota: ciò implica* $\mathbf{A}^k\mathbf{v}$ *non nullo,* $\forall k = 1, 2, 3, ...$*).*
Sia anche $\mathbf{u} \in \mathbb{C}^n$*, tale che* $\mathbf{u}^H\mathbf{x} \neq \mathbf{0}$ *(nota:* $\frac{1}{\lambda_1^k}(\mathbf{u}^H\mathbf{A}^k\mathbf{v}) \xrightarrow[k\to\infty]{} \mathbf{u}^H\mathbf{x}$ *e quindi* $\exists k^* \in \mathbb{N}$ *per cui sono non nulli* $\forall k > k^*$*). Allora:*

$$\frac{|\lambda_1|^k}{\lambda_1^k}\frac{\mathbf{A}^k\mathbf{v}}{||\mathbf{A}^k\mathbf{v}||} \xrightarrow[k\to\infty]{} \frac{\mathbf{x}}{||\mathbf{x}||} \tag{2.28}$$

$$\frac{\mathbf{u}^H\mathbf{A}^{k+1}\mathbf{v}}{\mathbf{u}^H\mathbf{A}^k\mathbf{v}} \xrightarrow[k\to\infty]{} \lambda_1 \tag{2.29}$$

dove la successione in (2.28) *è ben definita ed* $\exists k^* \in \mathbb{N}$ *tale che la successione* (2.29) *è ben definita* $\forall k > k^*$.
Le quantità in (2.28) *e* (2.29) *si possono calcolare tramite il seguente algoritmo:*

$$\mathbf{a}_0 := \mathbf{v},\ \mathbf{v}_0 := \frac{\mathbf{v}}{||\mathbf{v}||}.\ \textit{Per } k = 1,\ 2,\ \ldots$$

$$\begin{cases} \mathbf{a}_k := \mathbf{A}\mathbf{v}_{k-1} \\ \varphi_k := \dfrac{\mathbf{u}^H \mathbf{a}_k}{\mathbf{u}^H \mathbf{v}_{k-1}}, \quad \mathbf{u}^H \mathbf{v}_{k-1} \neq 0\ \forall k \\ \mathbf{v}_k := \dfrac{\mathbf{a}_k}{||\mathbf{a}_k||} \end{cases} \tag{2.30}$$

Infatti, $\varphi_k = \dfrac{\mathbf{u}^H \mathbf{A}^k \mathbf{v}}{\mathbf{u}^H \mathbf{A}^{k-1} \mathbf{v}}$ *(e quindi* $\varphi_k \xrightarrow[k\to\infty]{} \lambda_1$*) e* $\mathbf{v}_k = \dfrac{\mathbf{A}^k \mathbf{v}}{||\mathbf{A}^k \mathbf{v}||}$ $\forall k = 1, 2, ...$ *(e quindi* $\dfrac{|\lambda_1|^k}{\lambda_1^k} \mathbf{v}_k \xrightarrow[k\to\infty]{} \dfrac{\mathbf{x}}{||\mathbf{x}||}$*).*

Osservazione 2.4. La velocità di convergenza del metodo va come:

$$\Big(\max_{j:\lambda_j \neq \lambda_1} \Big(\frac{|\lambda_j|}{|\lambda_1|} \Big) \Big)^k p(k)$$

dove il grado del polinomio p è uguale all'ordine del blocco di Jordan di ordine massimo relativo all'autovalore che realizza il massimo.

Osservazione 2.5. Il costo per ogni passo è dato dal prodotto della matrice $\mathbf{A}$ per un vettore. In generale quindi è n^2.

2.3 Caso A non negativa

2.3.1 Cenni sulla teoria di Perron-Frobenius

Diamo adesso un risultato che descrive le proprietà dell'autovalore massimo e del corrispondente autovettore di una matrice quadrata, non negativa ed irriducibile.

Teorema 2.3 (Teorema di Perron-Frobenius [4])**.** *Sia* $\mathbf{A}$ *matrice quadrata di ordine* n*, non negativa ed irriducibile. Allora:*

1. $\rho(\mathbf{A})$ *è un autovalore di* $\mathbf{A}$*.*
2. $\rho(\mathbf{A}) > 0$*.*
3. *Esiste un autovettore corrispondente all'autovalore* $\rho(\mathbf{A})$ *con componenti positive.*
4. $\rho(\mathbf{A})$ *è autovalore di* $\mathbf{A}$ *semplice, cioè la sua molteplicità algebrica è uguale ad 1.*
5. *L'autovalore di modulo massimo, visto come funzione* $\rho(\mathbf{A})$ *della matrice* $\mathbf{A}$*, è una funzione strettamente crescente in ognuno dei suoi elementi: cioè, se per ogni* $\mathbf{B} \in \mathbb{C}^{n\times n}$ *tale che* $|\mathbf{B}| \geq \mathbf{A}$ *(ovvero, la disuguaglianza vale elemento per elemento) e* $\mathbf{B} \neq \mathbf{A}$*, allora* $\rho(\mathbf{B}) > \rho(\mathbf{A})$*.*

Se quindi $\mathbf{A}$ è matrice non negativa ed irriducibile allora, dal teorema appena enunciato, segue che $\mathbf{A}$ possiede un unico autovettore $\mathbf{z}$ positivo e stocastico (cioè $||\mathbf{z}||_1 = 1$) rispetto all'autovalore $\rho(\mathbf{A})$. Tale autovettore viene chiamato *autovettore di Perron* di $\mathbf{A}$.
La coppia $(\rho(\mathbf{A}), \mathbf{z})$ è quindi detta *coppia di Perron* di una matrice $\mathbf{A}$ non negativa ed irriducibile.

2.3.2 A non negativa ed irriducibile

Sia $\mathbf{A}$ matrice quadrata di ordine n, non negativa ed irriducibile allora, per il Teorema di Perron-Frobenius (vedi Teorema 2.3.), si ha che:

$$\exists!\mathbf{z} > \mathbf{0}, ||\mathbf{z}||_1 = 1 \text{ tale che } \mathbf{Az} = \rho(\mathbf{A})\mathbf{z},$$
$$\rho(\mathbf{A}) > 0,\ m_a(\rho(\mathbf{A})) = m_g(\rho(\mathbf{A})) = 1.$$

Grazie a questo risultato si può ottenere una versione del Teorema 2.2. nel caso in cui $\mathbf{A}$ sia matrice non negativa ed irriducibile, in cui l'unica ipotesi è $\lambda_1 = \rho(\mathbf{A})$ dominante (per avere convergenza).
Come $||.||$ si sceglie $||.||_1$ e come $\mathbf{u}$ si sceglie $\mathbf{u} = \mathbf{e}$.

Teorema 2.4. *Sia $\mathbf{A} \in \mathbb{C}^{n\times n}$ non negativa ed irriducibile. Allora, ne segue che esiste $\mathbf{X} \in \mathbb{C}^{n\times n}$ non singolare (cioè $det\mathbf{X} \neq 0$) tale che:*

$$\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \begin{bmatrix} \rho(\mathbf{A}) & 0 & \dots & \dots & \dots & 0 \\ 0 & \ddots & 0 & & & \vdots \\ \vdots & \ddots & \lambda & 1 & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & \dots & 0 & \lambda \end{bmatrix} =: \mathbf{J}$$

dove $\mathbf{J}$ è una matrice diagonale a blocchi (nota: ci possono essere più blocchi diagonali con lo stesso λ), in particolare si ha che:

$$\mathbf{A}\begin{bmatrix} & & & \\ \mathbf{x}_{i_l+1} & \dots & \dots & \mathbf{x}_{i_l+s_l} \\ & & & \end{bmatrix} =$$

$$\begin{bmatrix} & & & \\ \mathbf{x}_{i_l+1} & \dots & \dots & \mathbf{x}_{i_l+s_l} \\ & & & \end{bmatrix} \cdot \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{bmatrix} \in \mathbb{C}^{n\times s_l}$$

con s_l: $1 \leq s_l < n$, e $\lambda \in \sigma(\mathbf{A})$.
Supponiamo $\lambda_1 := \rho(\mathbf{A}) > |\lambda_j|$, $\forall j \geq 2$ (nota: $m_a(\lambda_1) = m_g(\lambda_1) = 1$).
Siano $\mathbf{v} > \mathbf{0}$, $||\mathbf{v}||_1 = 1$ e α, α_i tali che: $\mathbf{v} = \alpha\mathbf{z} + \sum_{i=2}^{n} \alpha_i\mathbf{x}_i$ (nota: $\mathbf{X}\mathbf{e}_1 = \mathbf{x}_1 = \mathbf{z}$).
Allora il vettore $\mathbf{x} = \alpha\mathbf{z}$ è non nullo perchè $\alpha > 0$ (nota: ciò implica $\mathbf{A}^k\mathbf{v}$ non nullo, $\forall k = 1, 2, 3, ...$).
Sia $\mathbf{e} = [1, \dots, 1]^T$, allora $\mathbf{e}^T\mathbf{x} > \mathbf{0}$ (nota: ciò implica che $\frac{1}{\lambda_1^k}(\mathbf{e}^T\mathbf{A}^k\mathbf{v}) \xrightarrow[k\to\infty]{} \mathbf{e}^T\mathbf{x}$ ed è positivo $\forall k \in \mathbb{N}$). Allora:

$$\frac{\mathbf{A}^k\mathbf{v}}{||\mathbf{A}^k\mathbf{v}||_1} = \frac{|\lambda_1|^k}{\lambda_1^k}\frac{\mathbf{A}^k\mathbf{v}}{||\mathbf{A}^k\mathbf{v}||_1} \xrightarrow[k\to\infty]{} \frac{\mathbf{x}}{||\mathbf{x}||_1} = \frac{\alpha\mathbf{z}}{||\alpha\mathbf{z}||_1} = \mathbf{z} \tag{2.31}$$

$$\frac{\mathbf{e}^T\mathbf{A}^{k+1}\mathbf{v}}{\mathbf{e}^T\mathbf{A}^k\mathbf{v}} \xrightarrow[k\to\infty]{} \lambda_1 \tag{2.32}$$

dove le successioni in (2.31) *e* (2.32) *sono ben definite* $\forall k \in \mathbb{N}$.
Le quantità (2.31) *e* (2.32) *si possono calcolare tramite il seguente algoritmo:*

$\mathbf{v}_0 := \mathbf{v} > \mathbf{0}$. *Per* k = *1, 2, ...*

$$\begin{cases} \mathbf{a}_k := \mathbf{A}\mathbf{v}_{k-1}, & \textit{nota: } \mathbf{a_k} > 0 \\ \varphi_k := \dfrac{\mathbf{e}^T\mathbf{a}_k}{\mathbf{e}^T\mathbf{v}_{k-1}}, & \textit{nota: } \mathbf{e}^T\mathbf{v}_{k-1} > 0\ \forall k \\ \mathbf{v}_k := \dfrac{\mathbf{a}_k}{||\mathbf{a}_k||_1}, & \end{cases} \tag{2.33}$$

Infatti, $\varphi_k = \dfrac{\mathbf{e}^T\mathbf{A}^k\mathbf{v}}{\mathbf{e}^T\mathbf{A}^{k-1}\mathbf{v}}$ *(e quindi* $\varphi_k \xrightarrow[k\to\infty]{} \lambda_1$*) e* $\mathbf{v}_k = \dfrac{\mathbf{A}^k\mathbf{v}}{||\mathbf{A}^k\mathbf{v}||_1}$ $\forall k = 1, 2, ...$ *(e quindi* $\mathbf{v}_k \xrightarrow[k\to\infty]{} \dfrac{\mathbf{x}}{||\mathbf{x}||_1} = \dfrac{\alpha\mathbf{z}}{||\alpha\mathbf{z}||_1} = \mathbf{z}$*).*

Dimostrazione. Sia:

$$\mathbf{X} = \begin{bmatrix} & & & & \\ \mathbf{z} & \mathbf{x}_2 & \dots & \dots & \mathbf{x}_n \\ & & & & \end{bmatrix}, \mathbf{X}^{-1}\mathbf{A} = \mathbf{J}\mathbf{X}^{-1}$$

allora, ne segue che:

$$\mathbf{e}_1^T\mathbf{X}^{-1}\mathbf{A} = \mathbf{e}_1^T\mathbf{J}\mathbf{X}^{-1}$$

cioè:

$$(\mathbf{e}_1^T\mathbf{X}^{-1})\mathbf{A} = \rho(\mathbf{A})(\mathbf{e}_1^T\mathbf{X}^{-1})$$

da cui si ha che esiste un $\beta \in \mathbb{C}$ tale che:

$$\mathbf{e}_1^T\mathbf{X}^{-1} = \beta\mathbf{y}^T \text{ con } \mathbf{y}\text{: } \mathbf{y} > \mathbf{0},\ ||\mathbf{y}||_1 = 1 \text{ e } \mathbf{y}^T\mathbf{A} = \rho(\mathbf{A})\mathbf{y}^T$$

(perché $m_g(\rho(\mathbf{A})) = 1$ e $\mathbf{y}$ è univocamente definito per il teorema di Perron-Frobenius applicato ad $\mathbf{A}^T$).
Ma:

$$1 = (\mathbf{e}_1^T\mathbf{X}^{-1})(\mathbf{X}\mathbf{e}_1) = \beta\mathbf{y}^T\mathbf{z} \Rightarrow \beta = \frac{1}{\mathbf{y}^T\mathbf{z}} \text{ ed anche } \mathbf{e}_1^T\mathbf{X}^{-1} = \frac{1}{\mathbf{y}^T\mathbf{z}}\mathbf{y}^T.$$

Inoltre per $i > 1$ si ha che:

$$0 = (\mathbf{e}_1^T\mathbf{X}^{-1})(\mathbf{X}\mathbf{e}_i) = \frac{1}{\mathbf{y}^T\mathbf{z}}\mathbf{y}^T\mathbf{x}_i$$

da cui:

$$\mathbf{y}^T\mathbf{x}_i = 0 \text{ se } i > 1.$$

Ora, data la rappresentazione di $\mathbf{v}$, $\mathbf{v} = \alpha\mathbf{z} + \sum_{i=2}^{n} \alpha_i\mathbf{x}_i$, moltiplicando ambo i membri per $\mathbf{y}^T$, si ottiene:

$$\mathbf{y}^T\mathbf{v} = \alpha\mathbf{y}^T\mathbf{z} + \sum_{i=2}^{n} \alpha_i\mathbf{y}^T\mathbf{x}_i = \alpha\mathbf{y}^T\mathbf{z} \Rightarrow \alpha = \frac{\mathbf{y}^T\mathbf{v}}{\mathbf{y}^T\mathbf{z}}. \tag{2.34}$$

Quindi, la rappresentazione del vettore $\mathbf{v}$ è:

$$\mathbf{v} = \left(\frac{\mathbf{y}^T\mathbf{v}}{\mathbf{y}^T\mathbf{z}}\right)\mathbf{z} + \sum_{i=2}^{n} \alpha_i\mathbf{x}_i.$$

Osserviamo che $\mathbf{x} = \left(\frac{\mathbf{y}^T\mathbf{v}}{\mathbf{y}^T\mathbf{z}}\right)\mathbf{z}$ è non nullo se $\mathbf{y}^T\mathbf{v} \neq 0$, ma poichè $\mathbf{v} > \mathbf{0}$ si ha che $\mathbf{y}^T\mathbf{v} > 0$, da cui otteniamo che $\mathbf{x} > \mathbf{0}$.
Ed infine si ha che: $\alpha > 0$ e $\frac{\mathbf{x}}{||\mathbf{x}||_1} = \mathbf{z}$.

□

2.3.3 A non negativa, irriducibile e stocastica per colonne

Sia $\mathbf{A}$ matrice quadrata di ordine n, non negativa, irriducibile e stocastica per colonne allora $\rho(\mathbf{A}) = 1$ e, per il Teorema di Perron-Frobenius (vedi Teorema 2.3.), si ha che:

$$\exists!\mathbf{z} > \mathbf{0}, ||\mathbf{z}||_1 = 1 \text{ tale che } \mathbf{A}\mathbf{z} = \mathbf{z} \text{ e } m_a(1) = m_g(1) = 1.$$

Grazie a questo risultato si può ottenere una versione del Teorema 2.2. nel caso in cui $\mathbf{A}$ sia matrice non negativa, irriducibile e stocastica per colonne.

Teorema 2.5. *Sia* $\mathbf{A} \in \mathbb{C}^{n\times n}$ *non negativa, irriducibile e stocastica per colonne. Allora, ne segue che esiste* $\mathbf{X} \in \mathbb{C}^{n\times n}$ *non singolare (cioè* $det\mathbf{X} \neq 0$*) tale che:*

$$\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \begin{bmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & \ddots & 0 & & & \vdots \\ \vdots & \ddots & \lambda & 1 & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & \dots & 0 & \lambda \end{bmatrix} =: \mathbf{J}$$

dove $\mathbf{J}$ *è una matrice diagonale a blocchi (nota: ci possono essere più blocchi diagonali con lo stesso* λ*), in particolare si ha che:*

$$\mathbf{A} \begin{bmatrix} \\ \mathbf{x}_{i_l+1} & \dots & \dots & \mathbf{x}_{i_l+s_l} \\ \\ \end{bmatrix} =$$

$$\begin{bmatrix} \\ \mathbf{x}_{i_l+1} & \dots & \dots & \mathbf{x}_{i_l+s_l} \\ \\ \end{bmatrix} \cdot \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{bmatrix} \in \mathbb{C}^{n \times s_l}$$

con s_l*:* $1 \leq s_l < n$*, e* $\lambda \in \sigma(\mathbf{A})$*.*
Supponiamo $\lambda_1 := \rho(\mathbf{A}) = 1 > |\lambda_j|$*,* $\forall j \geq 2$ *(nota:* $m_a(\lambda_1) = m_g(\lambda_1) = 1$*).*
Sia $\mathbf{v} > \mathbf{0}$*,* $||\mathbf{v}||_1 = 1$ *e* α, α_i *tali che:* $\mathbf{v} = \alpha\mathbf{z} + \sum_{i=2}^{n} \alpha_i \mathbf{x}_i$ *(nota:* $\mathbf{X}\mathbf{e}_1 = \mathbf{z}$*).*
Allora il vettore $\mathbf{x} = \alpha\mathbf{z}$ *è non nullo perché* $\alpha = 1$ *(nota:* $\mathbf{A}^k\mathbf{v} > 0$ $\forall k = 1, 2, 3, ...$*).*
Sia $\mathbf{e} = [1, \dots, 1]^T$*, allora* $\mathbf{e}^T\mathbf{x} > \mathbf{0}$ *(nota: ciò implica che* $\frac{1}{\lambda_1^k}(\mathbf{e}^T\mathbf{A}^k\mathbf{v}) = \mathbf{e}^T\mathbf{v} = 1$ $\forall k \in \mathbb{N}$*). Allora:*

$$\mathbf{A}^k\mathbf{v} = \frac{\mathbf{A}^k\mathbf{v}}{||\mathbf{A}^k\mathbf{v}||_1} = \frac{|\lambda_1|^k}{\lambda_1^k} \frac{\mathbf{A}^k\mathbf{v}}{||\mathbf{A}^k\mathbf{v}||_1} \xrightarrow[k\to\infty]{} \frac{\mathbf{x}}{||\mathbf{x}||_1} = \mathbf{z} \tag{2.35}$$

$$1 = \frac{\mathbf{e}^T\mathbf{A}^{k+1}\mathbf{v}}{\mathbf{e}^T\mathbf{A}^k\mathbf{v}} \xrightarrow[k\to\infty]{} \lambda_1 = 1 \tag{2.36}$$

dove le successioni in (2.35) *sono ben definite* $\forall k \in \mathbb{N}$.
Le quantità in (2.35) *e* (2.36) *si possono calcolare tramite il seguente algoritmo:*

$$\mathbf{v}_0 := \mathbf{v} > \mathbf{0}.\ \textit{Per } k = 1,\ 2,\ \ldots$$

$$\begin{cases} \mathbf{a}_k := \mathbf{A}\mathbf{v}_{k-1} > \mathbf{0} \\ \varphi_k := \dfrac{\mathbf{e}^T\mathbf{a}_k}{\mathbf{e}^T\mathbf{v}_{k-1}} = 1 \\ \mathbf{v}_k := \dfrac{\mathbf{a}_k}{||\mathbf{a}_k||_1} = \mathbf{a}_k \end{cases} \tag{2.37}$$

(nota: $\mathbf{a}_k > \mathbf{0}$ *e* $\mathbf{e}^T\mathbf{a}_k = ||\mathbf{a}_k||_1 = 1$*) e quindi l'algoritmo si può riscrivere come:*

$$\mathbf{v}_0 > \mathbf{0},\ ||\mathbf{v}_0||_1 = 1$$
$$\mathbf{v}_k = \mathbf{A}\mathbf{v}_{k-1},\ k = 1, 2, \ldots$$

Allora $\varphi_k = \dfrac{\mathbf{e}^T\mathbf{A}^k\mathbf{v}}{\mathbf{e}^T\mathbf{A}^{k-1}\mathbf{v}} = 1$ *(e quindi* $\varphi_k \xrightarrow[k\to\infty]{} \lambda_1 = 1$*) e* $\mathbf{v}_k = \dfrac{\mathbf{A}^k\mathbf{v}}{||\mathbf{A}^k\mathbf{v}||_1} = \mathbf{A}^k\mathbf{v}\ \forall k = 1, 2, \ldots$ *(e quindi* $\mathbf{v}_k = \dfrac{|\lambda_1|^k}{\lambda_1^k}\mathbf{v}_k \xrightarrow[k\to\infty]{} \dfrac{\mathbf{x}}{||\mathbf{x}||_1} = \mathbf{z}$*).*

Dimostrazione. Dall'epressione di α in (2.34) ottenuta nel teorema del caso di $\mathbf{A}$ non negativa ed irriducibile, segue che:

$$\alpha = \frac{\mathbf{y}^T\mathbf{v}}{\mathbf{y}^T\mathbf{z}} = \frac{\mathbf{e}^T\mathbf{v}}{\mathbf{e}^T\mathbf{z}} = \frac{1}{1} = 1.$$

□

Osservazione 2.6. Un'applicazione importante del precedente teorema è il calcolo del vettore Pagerank $\mathbf{p}$, definito come l'unico vettore, con componenti positive, di norma 1, uguale ad 1, tale che:

$$\mathbf{p} = \mathbf{G}\mathbf{p}$$

dove $\mathbf{G}$ è la matrice di Google.
Si ricorda che la componenti i-esima di $\mathbf{p}$ risulta uguale all'importanza della pagina i del web (o più in generale, al vertice i di un grafo orientato), (vedi [6]).

Capitolo 3

Problema

In questo capitolo affronteremo il seguente problema.
Abbiamo potuto verificare che nel caso in cui $\mathbf{A}$ sia una matrice definita positiva la successione φ_k, generata dal metodo delle potenze, converge in modo monotono all'autovalore dominante $\rho(\mathbf{A})$(vedi Corollario 2.1.). Ci si può chiedere se un risultato analogo si può ottenere anche per la successione di numeri reali φ_k generata nel caso in cui $\mathbf{A}$ sia una matrice non negativa ed irriducibile.
Ovvero, data φ_k come in (2.33), sappiamo che φ_k tende a $\rho(\mathbf{A})$, ma quale è il suo comportamento? Cioè, si può dimostrare che $\varphi_k \searrow \rho(\mathbf{A})$ o $\varphi_k \nearrow \rho(\mathbf{A})$ almeno sotto alcune ipotesi su $\mathbf{A}$?

In particolare, studiamo la disuguaglianza: $\varphi_k \geq \varphi_{k+1}$, dove ricordiamo che:

$$\varphi_k = \frac{\mathbf{e}^T\mathbf{A}^k\mathbf{v}}{\mathbf{e}^T\mathbf{A}^{k-1}\mathbf{v}}.$$

Supponiamo $k = 1$, allora si ha:

$$\varphi_1 \geq \varphi_2 \Leftrightarrow \mathbf{e}^T\mathbf{A}\mathbf{v} \geq \frac{\mathbf{e}^T\mathbf{A}^2\mathbf{v}}{\mathbf{e}^T\mathbf{A}\mathbf{v}}$$

(si ricorda che $\mathbf{e}^T\mathbf{v} = ||\mathbf{v}||_1 = 1$).
Moltiplicando ora ambo i membri per $\mathbf{e}^T\mathbf{A}\mathbf{v}$, si ha:

$$\left(\mathbf{e}^T\mathbf{A}\mathbf{v}\right)^2 \geq \mathbf{e}^T\mathbf{A}^2\mathbf{v}$$

ovvero, scritta in forma estesa:

$$\left(\begin{bmatrix}1 & \dots & 1\end{bmatrix}\begin{bmatrix}\mathbf{A}\end{bmatrix}\begin{bmatrix}v_1\\ \vdots \\ v_n\end{bmatrix}\right)^2 \geq \begin{bmatrix}1 & \dots & 1\end{bmatrix}\begin{bmatrix}\mathbf{A}^2\end{bmatrix}\begin{bmatrix}v_1\\ \vdots \\ v_n\end{bmatrix}. \tag{3.1}$$

Consideriamo il caso in cui $n = 2$.
Per questo si ha:

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \, \mathbf{v} = \begin{bmatrix} v \\ w \end{bmatrix}$$

quindi, sostituendo tali $\mathbf{A}$ e $\mathbf{v}$ in (3.1) si ottiene:

$$\begin{aligned}(\mathbf{v}(a+c) + \mathbf{w}(b+d))^2 &\geq \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} a^2+bc & b(a+d) \\ c(a+d) & d^2+bc \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} \\ &= va^2 + wd^2 + (v+w)bc + (a+d)(vc+wb).\end{aligned}$$

Per semplificare i calcoli possiamo scegliere $v = w = \frac{1}{2}$ (ricorda: $||\mathbf{v}||_1 = v + w = 1$).
Sostituiamo tali valori e si ha:

$$\frac{(a+b+c+d)^2}{4} \geq \frac{a^2+d^2}{2} + bc + \frac{(a+d)(b+c)}{2}$$

in cui, svolgendo i calcoli necessari:

$$\frac{(a^2+b^2+c^2+d^2+2ad+2bc)}{4} - \frac{(a^2+d^2)}{2} - bc \geq 0$$

e questo se e solo se:

$$\frac{(b-c)^2-(a-d)^2}{4} \geq 0$$

$\left(\text{nota: } (b-c)^2 - (a-d)^2 = det(\mathbf{A} - \mathbf{JAJ}) \text{ dove } \mathbf{J} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right)$.

Riassumendo si ha il seguente risultato:

Proposizione 3.1. $\varphi_1 - \varphi_2 \geq 0$ *se e solo se:*

$$\frac{1}{4}det(\mathbf{A} - \mathbf{JAJ}) = \frac{1}{4}[(b-c)^2 - (a-d)^2] = (\varphi_1 - \varphi_2)(\mathbf{e}^T\mathbf{A}\mathbf{v})(\mathbf{e}^T\mathbf{A}^0\mathbf{v}) \geq 0.$$

Consideriamo il caso $k = 2$, si ha che:

$$\varphi_2 \geq \varphi_3 \Leftrightarrow \frac{\mathbf{e}^T\mathbf{A}^2\mathbf{v}}{\mathbf{e}^T\mathbf{A}\mathbf{v}} \geq \frac{\mathbf{e}^T\mathbf{A}^3\mathbf{v}}{\mathbf{e}^T\mathbf{A}^2\mathbf{v}} \tag{3.2}$$

moltiplichiamo ora il membro di sinistra per $\mathbf{e}^T\mathbf{A}^2\mathbf{v}$ e quello di destra per $\mathbf{e}^T\mathbf{A}\mathbf{v}$, si ha:

$$\left(\mathbf{e}^T\mathbf{A}^2\mathbf{v}\right)^2 \geq (\mathbf{e}^T\mathbf{A}^3\mathbf{v})(\mathbf{e}^T\mathbf{A}\mathbf{v}).$$

Consideriamo, anche qui, $n = 2$ e $\mathbf{v}$ uniforme quindi analogamente al caso precedente:

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \ \mathbf{v} = \begin{bmatrix} v \\ w \end{bmatrix} \text{ con } v = w = \frac{1}{2}.$$

Inoltre, poichè:

$$\mathbf{A}^2 = \begin{bmatrix} a^2+bc & b(a+d) \\ c(a+d) & d^2+bc \end{bmatrix}$$

$$\mathbf{A}^3 = \mathbf{A}\mathbf{A}^2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} a^2+bc & b(a+d) \\ c(a+d) & d^2+bc \end{bmatrix}$$

$$= \begin{bmatrix} a^3+2abc+bcd & a^2b+b^2c+abd+bd^2 \\ a^2b+b^2c+abd+bd^2 & abc+2bcd+d^3 \end{bmatrix}$$

allora:

$$\mathbf{e}^T\mathbf{A}\mathbf{v} = \frac{a+b+c+d}{2}$$

$$(\mathbf{e}^T\mathbf{A}^2\mathbf{v})^2 = \left(\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} a^2+bc & b(a+d) \\ c(a+d) & d^2+bc \end{bmatrix} \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} \right)^2$$

$$= \left[\frac{a^2+d^2}{2} + bc + \frac{(a+d)(b+c)}{2} \right]^2$$

$$\mathbf{e}^T\mathbf{A}^3\mathbf{v} = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} a^3+2abc+bcd & a^2b+b^2c+abd+bd^2 \\ a^2b+b^2c+abd+bd^2 & abc+2bcd+d^3 \end{bmatrix} \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$$

$$= \frac{a^3+d^3}{2} + \frac{3abc}{2} + \frac{3bcd}{2} + \frac{acd+abd}{2} + \\ + \frac{a^2c+d^2c+b^2c+a^2b+b^2c+bd^2}{2}.$$

Sostituendo i precedenti risultati in (3.2), otteniamo che:

$$\left[\frac{a^2+d^2}{2} + bc + \frac{(a+d)(b+c)}{2} \right]^2 \geq \left(\frac{a^3+d^3}{2} + \frac{3abc}{2} + \frac{3bcd}{2} + \frac{acd+abd}{2} + \right. \\ \left. + \frac{a^2c+d^2c+b^2c+a^2b+b^2c+bd^2}{2} \right) \left(\frac{a+b+c+d}{2} \right)$$

da cui, dopo gli opportuni calcoli e semplificazioni si ha che:

$$\frac{1}{4}(ad-bc)[(b-c)^2-(a-d)^2] \geq 0.$$

Riassumendo si ha il seguente risultato:

Proposizione 3.2. $\varphi_2 - \varphi_3 \geq 0$ *se e solo se:*

$$\frac{1}{4}det(\mathbf{A})det(\mathbf{A}-\mathbf{JAJ}) = \frac{1}{4}(ad-bc)[(b-c)^2-(a-d)^2] \\ = (\varphi_2-\varphi_3)(\mathbf{e}^T\mathbf{A}\mathbf{v})(\mathbf{e}^T\mathbf{A}^2\mathbf{v}) \geq 0.$$

Dai risultati ottenuti precedentemente, possiamo congetturare il seguente risultato nel caso $n = 2$.

Proposizione 3.3. *(Congettura) Per $n = 2$, $\mathbf{v} = \frac{1}{2}\mathbf{e}$ e $\forall k \geq 1$, vale la seguente uguaglianza:*

$$\begin{aligned}\frac{1}{4}det(\mathbf{A})^k det(\mathbf{A} - \mathbf{JAJ}) &= \frac{1}{4}(ad - bc)^{k-1}[(b-c)^2 - (a-d)^2] \\ &= (\varphi_k - \varphi_{k+1})(\mathbf{e}^T\mathbf{A}^{k-1}\mathbf{v})(\mathbf{e}^T\mathbf{A}^k\mathbf{v}) \\ &= \left(\frac{\mathbf{e}^T\mathbf{A}^k\mathbf{v}}{\mathbf{e}^T\mathbf{A}^{k-1}\mathbf{v}} - \frac{\mathbf{e}^T\mathbf{A}^{k+1}\mathbf{v}}{\mathbf{e}^T\mathbf{A}^k\mathbf{v}}\right)(\mathbf{e}^T\mathbf{A}^{k-1}\mathbf{v})(\mathbf{e}^T\mathbf{A}^k\mathbf{v}) \\ &= (\mathbf{e}^T\mathbf{A}^k\mathbf{v})^2 - (\mathbf{e}^T\mathbf{A}^{k+1}\mathbf{v})(\mathbf{e}^T\mathbf{A}^{k-1}\mathbf{v}).\end{aligned}$$

Quindi, $\varphi_k - \varphi_{k+1} \geq 0$ se e solo se:

$$\frac{1}{4}det(\mathbf{A})^k det(\mathbf{A} - \mathbf{JAJ}) = \frac{1}{4}(ad - bc)^{k-1}[(b-c)^2 - (a-d)^2] \geq 0.$$

Da cui, se $ad - bc > 0$, la successione φ_k è monotona (nota: in particolare se $[(b-c)^2 - (a-d)^2] > 0$ è decrescente, mentre se $[(b-c)^2 - (a-d)^2] < 0$ è crescente), altrimenti è alterna.

Si è dimostrato che tale risultato vale per le seguenti matrici:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \text{ per } k = 1, 2, 3, 4, 5$$

$$\mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \text{ per } k = 1, 2, 3, 4$$

$$\mathbf{C} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \text{ per } k = 1, 2, 3, 4, 5$$

In particolare, verifichiamo con due esempi quanto detto per **A** e **C** con $k = 4$.

Esempio 3.1. Sia $k = 4$ e considero la matrice:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

poichè in tal caso $det(\mathbf{A}) = 0$, si vuole far vedere che:

$$(e^T\mathbf{A}^4\mathbf{v})^2 - (e^T\mathbf{A}^5\mathbf{v})(e^T\mathbf{A}^3\mathbf{v}) = 0.$$

Per cui, considerando:

$$\mathbf{A}^2 = \mathbf{A}\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

$$\mathbf{A}^3 = \mathbf{A}^2\mathbf{A} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix}$$

$$\mathbf{A}^4 = \mathbf{A}^3\mathbf{A} = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 8 \\ 8 & 8 \end{bmatrix}$$

$$\mathbf{A}^5 = \mathbf{A}^4\mathbf{A} = \begin{bmatrix} 8 & 8 \\ 8 & 8 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 16 & 16 \\ 16 & 16 \end{bmatrix}$$

allora:

$$\mathbf{e}^T\mathbf{A}^4\mathbf{v} = \begin{bmatrix} 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 8 & 8 \\ 8 & 8 \end{bmatrix} \cdot \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 16 & 16 \end{bmatrix} \cdot \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = 16$$

$$\mathbf{e}^T\mathbf{A}^5\mathbf{v} = \begin{bmatrix} 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 16 & 16 \\ 16 & 16 \end{bmatrix} \cdot \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 32 & 32 \end{bmatrix} \cdot \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = 32$$

$$\mathbf{e}^T\mathbf{A}^3\mathbf{v} = \begin{bmatrix} 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix} \cdot \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 8 & 8 \end{bmatrix} \cdot \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = 8$$

da cui si ottiene che:

$$(e^T\mathbf{A}^4\mathbf{v})^2 - (e^T\mathbf{A}^5\mathbf{v})(e^T\mathbf{A}^3\mathbf{v}) = (16)^2 - (32 \cdot 8) = 0.$$

Esempio 3.2. Sia $k = 4$ e considero la matrice:

$$\mathbf{C} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

in tal caso vogliamo dimostrare che:

$$\frac{1}{4}(ad - bc)^3[(b - c)^2 - (a - d)^2] = (\mathbf{e}^T\mathbf{C}^4\mathbf{v})^2 - (\mathbf{e}^T\mathbf{C}^5\mathbf{v})(\mathbf{e}^T\mathbf{C}^3\mathbf{v})$$

dove gli elementi della matrice, nel nostro caso, sono $a = 0, b = c = d = 1$. Per cui, consideriamo:

$$\mathbf{C}^2 = \mathbf{C}\mathbf{C} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\mathbf{C}^3 = \mathbf{C}^2\mathbf{C} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$$

$$\mathbf{C}^4 = \mathbf{C}^3\mathbf{C} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$$

$$\mathbf{C}^5 = \mathbf{C}^4\mathbf{C} = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ 5 & 8 \end{bmatrix}$$

allora:

$$\mathbf{e}^T\mathbf{C}^4\mathbf{v} = \begin{bmatrix} 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix} \cdot \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 5 & 8 \end{bmatrix} \cdot \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = \frac{13}{2}$$

$$\mathbf{e}^T\mathbf{C}^5\mathbf{v} = \begin{bmatrix} 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 & 5 \\ 5 & 8 \end{bmatrix} \cdot \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 8 & 13 \end{bmatrix} \cdot \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = \frac{21}{2}$$

$$\mathbf{e}^T\mathbf{C}^3\mathbf{v} = \begin{bmatrix} 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 3 & 5 \end{bmatrix} \cdot \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = 4$$

da cui si ottiene che:

$$(\mathbf{e}^T\mathbf{C}^4\mathbf{v})^2 - (\mathbf{e}^T\mathbf{C}^5\mathbf{v})(\mathbf{e}^T\mathbf{C}^3\mathbf{v}) = \left(\frac{13}{2}\right)^2 - \left(\frac{21}{2}\right)4 = \frac{1}{4}.$$

Inoltre, osserviamo che:

$$\frac{1}{4}(ad - bc)^3[(b-c)^2 - (a-d)^2] = \frac{1}{4}[(-1)(-1)] = \frac{1}{4}$$

da cui quindi segue:

$$\frac{1}{4}(ad - bc)^3[(b-c)^2 - (a-d)^2] = \frac{1}{4} = (\mathbf{e}^T\mathbf{C}^4\mathbf{v})^2 - (\mathbf{e}^T\mathbf{C}^5\mathbf{v})(\mathbf{e}^T\mathbf{C}^3\mathbf{v})$$

che è proprio il risultato che volevamo ottenere.

□

Nell'ultimo esempio si osserva che, per $k = 1$:

$$det\left(\begin{bmatrix} \mathbf{e}^T\mathbf{C}\mathbf{e} & \mathbf{e}^T\mathbf{C}^2\mathbf{e} \\ \mathbf{e}^T\mathbf{e} & \mathbf{e}^T\mathbf{C}\mathbf{e} \end{bmatrix}\right) = det(\mathbf{C})det\left(\begin{bmatrix} \mathbf{e}^T\mathbf{e} & \mathbf{e}^T\mathbf{C}\mathbf{e} \\ \mathbf{e}^T\mathbf{C}^{-1}\mathbf{e} & \mathbf{e}^T\mathbf{e} \end{bmatrix}\right).$$

Allora, si può pensare che questa relazione fra due determinanti successivi sia vera per ogni matrice quadrata di ordine 2:

Lemma 3.1. *(Congettura) Per $n = 2$ si ha che:*

$$\begin{aligned} det\left(\begin{bmatrix} \mathbf{e}^T\mathbf{A}^k\mathbf{e} & \mathbf{e}^T\mathbf{A}^{k+1}\mathbf{e} \\ \mathbf{e}^T\mathbf{A}^{k-1}\mathbf{e} & \mathbf{e}^T\mathbf{A}^k\mathbf{e} \end{bmatrix}\right) &= (\mathbf{e}^T\mathbf{A}^k\mathbf{e})^2 - (\mathbf{e}^T\mathbf{A}^{k+1}\mathbf{e})(\mathbf{e}^T\mathbf{A}^{k-1}\mathbf{e}) \\ &= (ad - bc)[(\mathbf{e}^T\mathbf{A}^{k-1}\mathbf{e})^2 - (\mathbf{e}^T\mathbf{A}^k\mathbf{e})(\mathbf{e}^T\mathbf{A}^{k-2}\mathbf{e})] \qquad (3.3) \\ &= det(\mathbf{A})det\left(\begin{bmatrix} \mathbf{e}^T\mathbf{A}^{k-1}\mathbf{e} & \mathbf{e}^T\mathbf{A}^k\mathbf{e} \\ \mathbf{e}^T\mathbf{A}^{k-2}\mathbf{e} & \mathbf{e}^T\mathbf{A}^{k-1}\mathbf{e} \end{bmatrix}\right). \end{aligned}$$

Ora, possiamo osservare che se valesse l'identità in (3.3) si potrebbe facilmente dimostrare la Proposizione 3.3. nel modo riportato in seguito.

Dimostrazione. Ciò che vogliamo dimostrare è che, per $\mathbf{v} = \frac{1}{2}\mathbf{e}$ e $\forall k \geq 1$, vale la seguente uguaglianza:

$$\begin{aligned}\frac{1}{4}(ad-bc)^{k-1}[(b-c)^2-(a-d)^2] &= (\mathbf{e}^T\mathbf{A}^k\mathbf{v})^2 - (\mathbf{e}^T\mathbf{A}^{k+1}\mathbf{v})(\mathbf{e}^T\mathbf{A}^{k-1}\mathbf{v})\\ &= \frac{1}{4}[(\mathbf{e}^T\mathbf{A}^k\mathbf{e})^2 - (\mathbf{e}^T\mathbf{A}^{k+1}\mathbf{e})(\mathbf{e}^T\mathbf{A}^{k-1}\mathbf{e})].\end{aligned}$$

Procediamo per induzione.
Sia $k = 1$, allora si ha che:

$$(b-c)^2-(a-d)^2 = (\mathbf{e}^T\mathbf{A}\mathbf{e})^2 - \mathbf{e}^T\mathbf{A}^2\mathbf{e}$$

che è verificato per quanto abbiamo visto precedentemente nel caso in cui $k = 1$.
Supponiamo che l'ipotesi sia vera per $k \leq m$, in particolare:

$$(ad-bc)^{m-1}[(b-c)^2-(a-d)^2] = (\mathbf{e}^T\mathbf{A}^m\mathbf{e})^2 - (\mathbf{e}^T\mathbf{A}^{m+1}\mathbf{e})(\mathbf{e}^T\mathbf{A}^{m-1}\mathbf{e}).$$

Dimostriamo che è vero il caso per $k = m + 1$, ovvero vogliamo dimostrare che:

$$(ad-bc)^{m}[(b-c)^2-(a-d)^2] = (\mathbf{e}^T\mathbf{A}^{m+1}\mathbf{e})^2 - (\mathbf{e}^T\mathbf{A}^{m+2}\mathbf{e})(\mathbf{e}^T\mathbf{A}^{m}\mathbf{e}).$$

Utilizzando ora il Lemma 3.1. e l'ipotesi induttiva, si ottiene:

$$\begin{aligned}(\mathbf{e}^T\mathbf{A}^{m+1}\mathbf{e})^2 - (\mathbf{e}^T\mathbf{A}^{m+2}\mathbf{e})(\mathbf{e}^T\mathbf{A}^{m}\mathbf{e}) &= (ad-bc)[(\mathbf{e}^T\mathbf{A}^m\mathbf{e})^2 - (\mathbf{e}^T\mathbf{A}^{m+1}\mathbf{e})(\mathbf{e}^T\mathbf{A}^{m-1}\mathbf{e})]\\ &= (ad-bc)\{(ad-bc)^{m-1}[(b-c)^2-(a-d)^2]\}\\ &= (ad-bc)^m[(b-c)^2-(a-d)^2]\end{aligned}$$

che è proprio ciò che volevamo dimostrare e quindi, la tesi è verificata.

□

Osservazione 3.1. Data una matrice quadrata $\mathbf{A}$ di ordine n, invertibile, gli si può sempre associare la seguente matrice di Toeplitz di ordine infinito:

$$
\begin{bmatrix}
\ddots & \ddots & & & \ddots & & & & & \\
\ddots & \ddots & \ddots & & & \ddots & & & & \\
 & \ddots & \ddots & \ddots & & & \mathbf{e}^T\mathbf{A}^k\mathbf{e} & & & \\
 & & \ddots & \mathbf{e}^T\mathbf{e} & \mathbf{e}^T\mathbf{A}\mathbf{e} & & & \mathbf{e}^T\mathbf{A}^k\mathbf{e} & & \\
\ddots & & & \mathbf{e}^T\mathbf{A}^{-1}\mathbf{e} & \mathbf{e}^T\mathbf{e} & \mathbf{e}^T\mathbf{A}\mathbf{e} & & & \ddots & \\
 & \ddots & & & \mathbf{e}^T\mathbf{A}^{-1}\mathbf{e} & \mathbf{e}^T\mathbf{e} & \ddots & & & \ddots \\
 & & \mathbf{e}^T\mathbf{A}^{-k}\mathbf{e} & & & \ddots & \ddots & \ddots & & \\
 & & & \mathbf{e}^T\mathbf{A}^{-k}\mathbf{e} & & & \ddots & \ddots & \ddots & \\
 & & & & \ddots & & & \ddots & \ddots & \ddots \\
 & & & & & \ddots & & & \ddots & \ddots
\end{bmatrix}.
$$

Osserviamo che, quando l'ordine di $\mathbf{A}$ è pari a 2, le sottomatrici coinvolte nell'identità (3.3) del Lemma 3.1. sono sottomatrici successive 2×2 della matrice di Toeplitz precedente.

$\square$

Ci si può chiedere come diventerebbe la Proposizione 3.3. nel caso in cui $n = 3$.
In particolare, ci chiediamo se può valere ancora un'identità del tipo:

$$[det(\mathbf{A})]^{k-1} det(\mathbf{A} - \mathbf{JAJ}) = (\mathbf{e}^T\mathbf{A}^k\mathbf{e})^2 - (\mathbf{e}^T\mathbf{A}^{k+1}\mathbf{e})(\mathbf{e}^T\mathbf{A}^{k-1}\mathbf{e}) \tag{3.4}$$

per un'opportuna scelta della matrice $\mathbf{J}$ di ordine 3. Proviamo a verificare, in particolare, se l'identità in (3.4) con:

$$\mathbf{J} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

vale per il seguente esempio.

Esempio 3.3. Consideriamo la matrice quadrata $\mathbf{A}$ di ordine 3 data da:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

e sia $k = 2$. Allora, vogliamo far vedere che vale:

$$det(\mathbf{A})det(\mathbf{A}-\mathbf{JAJ}) = (\mathbf{e}^T\mathbf{A}^2\mathbf{e})^2 - (\mathbf{e}^T\mathbf{A}^3\mathbf{e})(\mathbf{e}^T\mathbf{A}\mathbf{e}).$$

Per cui, consideriamo:

$$\mathbf{A}^2 = \mathbf{AA} = \begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix} \cdot \begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix} = \begin{bmatrix}3&3&3\\3&3&3\\3&3&3\end{bmatrix}$$

$$\mathbf{A}^3 = \mathbf{AA}^2 = \begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix} \cdot \begin{bmatrix}3&3&3\\3&3&3\\3&3&3\end{bmatrix} = \begin{bmatrix}9&9&9\\9&9&9\\9&9&9\end{bmatrix}$$

allora:

$$\mathbf{e}^T\mathbf{A}^2\mathbf{e} = \begin{bmatrix}1&1&1\end{bmatrix} \cdot \begin{bmatrix}3&3&3\\3&3&3\\3&3&3\end{bmatrix} \cdot \begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}9&9&9\end{bmatrix} \cdot \begin{bmatrix}1\\1\\1\end{bmatrix} = 27$$

$$\mathbf{e}^T\mathbf{A}^3\mathbf{e} = \begin{bmatrix}1&1&1\end{bmatrix} \cdot \begin{bmatrix}9&9&9\\9&9&9\\9&9&9\end{bmatrix} \cdot \begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}27&27&27\end{bmatrix} \cdot \begin{bmatrix}1\\1\\1\end{bmatrix} = 81$$

$$\mathbf{e}^T\mathbf{A}\mathbf{e} = \begin{bmatrix}1&1&1\end{bmatrix} \cdot \begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix} \cdot \begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}3&3&3\end{bmatrix} \cdot \begin{bmatrix}1\\1\\1\end{bmatrix} = 9$$

da cui si ottiene:

$$(\mathbf{e}^T\mathbf{A}^2\mathbf{e})^2 - (\mathbf{e}^T\mathbf{A}^3\mathbf{e})(\mathbf{e}^T\mathbf{A}\mathbf{e}) = (27)^2 - (81)(9) = 729 - 729 = 0.$$

Inoltre, poiché in tal caso il $det(\mathbf{A}) = 0$, si ha:

$$det(\mathbf{A})det(\mathbf{A}-\mathbf{JAJ}) = 0$$

da cui segue che:

$$det(\mathbf{A})det(\mathbf{A}-\mathbf{JAJ}) = (\mathbf{e}^T\mathbf{A}^2\mathbf{e})^2 - (\mathbf{e}^T\mathbf{A}^3\mathbf{e})(\mathbf{e}^T\mathbf{A}\mathbf{e})$$

che è proprio il risultato che volevamo ottenere. Quindi, in tal caso specifico l'identità in (3.4) è valida.

Osservazione 3.2. Nel caso $n = 3$, con $\mathbf{J}$ matrice data da:

$$\mathbf{J} = \begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}$$

si può osservare che:

$$det(\mathbf{A}-\mathbf{JAJ}) = 0.$$

In realtà ciò è vero per ogni n dispari.

Dimostrazione. Moltiplico $\mathbf{A} - \mathbf{JAJ}$ sia a destra che a sinistra per $\mathbf{J}$, ottenendo:

$$\mathbf{J}(\mathbf{A} - \mathbf{JAJ})\mathbf{J} = (\mathbf{JA} - \mathbf{AJ})\mathbf{J} = \mathbf{JAJ} - \mathbf{A} \tag{3.5}$$

(nota: $\mathbf{J}^2 = \mathbf{I}$).
Per comodità chiamiamo $\mathbf{M} := \mathbf{A} - \mathbf{JAJ}$.
Dalla (3.5) otteniamo:

$$\mathbf{JMJ} = -\mathbf{M}$$

da cui ne segue che:

$$det(\mathbf{JMJ}) = det(-\mathbf{M}).$$

Utilizzando ora il Teorema di Binet[1] si ha:

$$det(\mathbf{J})det(\mathbf{M})det(\mathbf{J}) = det(-\mathbf{M}) = (-1)^n det(\mathbf{M})$$

ed applicando ulteriormente Binet al primo membro otteniamo:

$$det(\mathbf{J}^2)det(\mathbf{M}) = (-1)^n det(\mathbf{M})$$

da cui:

$$det(\mathbf{M}) = (-1)^n det(\mathbf{M}).$$

In conclusione, se n è pari otteniamo semplicemente l'identità:

$$det(\mathbf{M}) = det(\mathbf{M})$$

mentre se n è dispari otteniamo:

$$det(\mathbf{M}) = -det(\mathbf{M}) \Rightarrow det(\mathbf{M}) = 0$$

$\square$

L'identità in (3.4) viene però contraddetta con:

$$\mathbf{J} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

dal seguente esempio:

---

[1]Se $\mathbf{A}$ e $\mathbf{B}$ sono due matrici quadrate, con lo stesso numero di righe, in un campo $\mathbb{K}$ allora: $det(\mathbf{AB}) = det(\mathbf{A})det(\mathbf{B})$

Esempio 3.4. Sia $k = 2$ e sia $\mathbf{A}$ matrice data da:

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

Dall'osservazione precedente, poiché $n = 3$, occorrerebbe solo verificare che:

$$(\mathbf{e}^T\mathbf{A}^2\mathbf{e})^2 - (\mathbf{e}^T\mathbf{A}^3\mathbf{e})(\mathbf{e}^T\mathbf{A}\mathbf{e}) = 0 \tag{3.6}$$

Poiché:

$$\mathbf{A}^2 = \mathbf{A}\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \\ 0 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \\ 0 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 1 \\ 4 & 5 & 2 \\ 4 & 4 & 3 \end{bmatrix}$$

$$\mathbf{A}^3 = \mathbf{A}\mathbf{A}^2 = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \\ 0 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 & 2 & 1 \\ 4 & 5 & 2 \\ 4 & 4 & 3 \end{bmatrix} = \begin{bmatrix} 7 & 7 & 3 \\ 14 & 13 & 7 \\ 12 & 14 & 7 \end{bmatrix}$$

si ha:

$$\mathbf{e}^T\mathbf{A}^2\mathbf{e} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 & 2 & 1 \\ 4 & 5 & 2 \\ 4 & 4 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 11 & 11 & 6 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 28$$

$$\mathbf{e}^T\mathbf{A}^3\mathbf{e} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 7 & 7 & 3 \\ 14 & 13 & 7 \\ 12 & 14 & 7 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 33 & 34 & 17 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 84$$

$$\mathbf{e}^T\mathbf{A}\mathbf{e} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \\ 0 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 & 4 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 9$$

da cui si ottiene:

$$(\mathbf{e}^T\mathbf{A}^2\mathbf{e})^2 - (\mathbf{e}^T\mathbf{A}^3\mathbf{e})(\mathbf{e}^T\mathbf{A}\mathbf{e}) = (28)^2 - (84)(9) = 784 - 756 = 28.$$

Da questo ne segue che l'identità (3.6) non è verificata.

Il precedente esempio ci dimostra che l'identità in (3.4) non è vera per la nostra scelta di $\mathbf{J}$, ma potrebbe essere vera per un'altra scelta di $\mathbf{J}$ che per $n = 2$ diventi:

$$\mathbf{J} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Rimane quindi aperta la seguente congettura:

Proposizione 3.4. *(Congettura) Data* $\mathbf{A}$ *non negativa ed irriducibile, se* $det(\mathbf{A}) > 0$ *allora* φ_k *converge in modo monotono a* $\rho(\mathbf{A})$ *mentre, se* $det(\mathbf{A}) < 0$ *converge a* $\rho(\mathbf{A})$ *da sopra e da sotto in modo alternato.*

Bibliografia

[1] Dario A. Bini, Università di Pisa, *La forma normale di Schur*, http://pagine.dm.unipi.it/bini/Didattica/AnaNum/testi/Dispense/schur.pdf, (30 Ottobre 2013).

[2] D. Bini, M. Capovani, O. Menchi, *Metodi numerici per l'algebra lineare*, Zanichelli, (1998).

[3] Gene H. Golub, Charles F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, (2013).

[4] Richard S. Varga, *Matrix Iterative Analysis*, Springer, (1999).

[5] Philip J. Davis, *Circulant Matrices*, American Mathematical Soc., (2013).

[6] P. Berkhin, *A Survey on Pagerank Computing*, Internet Math, (2005).

Ringraziamenti

Alla fine di questo intenso percorso universitario è finalmente giunto il giorno della mia laurea, scrivere queste frasi di ringraziamento è un tocco finale della mia tesi. Vorrei quindi spendere qualche parola per le persone che mi hanno aiutato in questo periodo.

Prima di tutto, vorrei ringraziare il prof. Carmine Di Fiore, relatore di questa tesi, che oltre all'aiuto fornitomi nell'apprendimento e nella conoscenza più approfondita di alcuni argomenti riguardanti l'Analisi Numerica, è sempre stato cortese e disponibile durante tutto il periodo della stesura di questo lavoro. Senza il suo supporto questa tesi non avrebbe preso vita.

Un ringraziamento enorme va alla mia famiglia, che con i loro consigli ed il loro sostegno, sia morale che economico, mi hanno permesso di seguire questo percorso fino al giorno della mia laurea, contribuendo alla mia formazione personale.

Infine, ma non meno importanti, vorrei ringraziare tutti i miei compagni di corso con cui ho condiviso questa esperienza, sia durante le fatiche che l'hanno caratterizzata sia durante i momenti di gioia e di soddisfazione.

Grazie a tutti.

Sara Loss
Roma, 22 Maggio 2019.