31 Marzo 2023

Sull'esponenziale di matrice

Francesca Acquistapace

In queste note vogliamo descrivere, data una matrice quadrata A, alcune proprietà della matrice esponenziale e^A. Questa nozione gioca nei sistemi di equazioni differenziali lineari a coefficienti costanti un ruolo del tutto analogo a quello della funzione esponenziale nel caso di una sola equazione.

Si può infatti dimostrare, cfr. ad esempio Paolo Acquistapace: Appunti di Analisi Matematica 2, pag 86 (https://people.dm.unipi.it/acquistp/analisi2.pdf), che dato un sistema di equazioni differenziali lineari a coefficienti costanti, la funzione $e^{tA}X_0$ è soluzione del problema di Cauchy

$$\begin{cases} \dot{x}_i = \sum a_{i,j}x_j(t), \\ X(0) = X_0 \end{cases}$$

dove $X(t) = (x_1(t), \ldots, x_n(t))^T$ e A è la matrice dei coefficienti del sistema.

1 Esponenziale di matrice.

Possiamo definire l'esponenziale di una matrice quadrata A di ordine n a coefficienti reali o complessi in analogia alla serie di potenze che converge a e^x

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

dove si intende $0! = 1$, sostituendo alla x la matrice A con l'ovvia convenzione che $A^0 = I$.

Definizione 1.1.

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

Osserviamo che la definizione è ben posta in quanto si ha

Proposizione 1.2. *La serie di e^A converge in norma per ogni matrice A.*

Prova. Lo spazio delle matrici $n \times n$ a coefficienti reali è isomorfo a $\mathbb{R}^{n^2}$, quindi possiamo definire la *norma* di una matrice come

$$\|A\| = \left(\sum_{i,j} |a_{ij}|^2 \right)^{\frac{1}{2}},$$

quindi

$$\|e^A\| \leqslant \sum_n \frac{\|A\|^n}{n!} = e^{\|A\|} < \infty$$

$\square$

Proposizione 1.3. *Se A e B sono due matrici tali che $AB = BA$ allora si ha $e^{A+B} = e^A e^B$.*

In particolare e^A è sempre invertibile e la sua inversa è e^{-A}

Prova.

$$e^{A+B} = \sum_{n=0}^{\infty} \frac{(A+B)^n}{n!} = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{j=0}^{n} \binom{n}{j} A^j B^{n-j} \right) = \sum_{n=0}^{\infty} \sum_{j=0}^{n} \frac{1}{n!} \frac{n!}{j!(n-j)!} A^j B^{n-j} =$$

$$= \sum_{n=0}^{\infty} \left(\sum_{j+k=n} \frac{A^j}{j!} \frac{B^k}{k!} \right) = e^A \cdot e^B$$

$\square$

Vediamo ora alcuni casi particolari.

- Se A è una matrice diagonale, cioè

$$A = \begin{pmatrix} a_1 & 0 & 0 & \dots \\ 0 & a_2 & 0 & \dots \\ 0 & 0 & \ddots & \dots \\ 0 & 0 & \dots & a_n \end{pmatrix},$$

  allora

$$A^k = \begin{pmatrix} a_1^k & 0 & 0 & \dots \\ 0 & a_2^k & 0 & \dots \\ 0 & 0 & \ddots & \dots \\ 0 & 0 & \dots & a_n^k \end{pmatrix}$$

  e quindi e^A è diagonale con sulla diagonale $e^{a_1}, e^{a_2}, \dots, e^{a_n}$.

- Se A è nilpotente, allora e^A è un polinomio di grado $< n$ in A.
- Se A è triangolare superiore, allora $A = D + N$ con D matrice diagonale e N matrice nilpotente, e se $D = aI$, allora D e N commutano e $e^A = e^a I \cdot p(N)$ dove $p(N)$ è un polinomio in N.
- e^{tA} è derivabile termine a termine e la sua derivata è Ae^{tA}

Osservazione 1.4. . Se A e B sono simili, anche e^A e e^B sono simili. Infatti se $A = M^{-1}BM$, $A^k = M^{-1}B^kM$, quindi $e^A = M^{-1}e^BM$.

Pertanto se A è diagonalizzabile, cioè nella sua classe di similitudine c'è una matrice diagonale, anche e^A è diagonalizzabile. Se A non è diagonalizzabile ci servirebbe qualcosa che somigliasse ad una forma canonica che caratterizzasse la classe di similitudine di A. Questa forma canonica esiste ed è sostanzialmente unica, come vedremo nel prossimo paragrafo.

Osservazione 1.5. . Ricordiamo come è fatta la matrice M che dà la similitudine. Possiamo pensare A come matrice associata ad un endomorfismo $T : V \to V$ rispetto ad una base $\mathcal{B} = (v_1, \dots, v_n)$. La matrice simile $M^{-1}AM$ rappresenta T rispetto ad una base diversa $\mathcal{B}' = (w_1, \dots, w_n)$. Ogni base fornisce un isomorfismo con $\mathbb{R}^n$, $\phi_{\mathcal{B}} : V \to \mathbb{R}^n$ e l'isomorfismo manda v_j nel vettore e_j, $j = 1, \dots, n$. Analogamente $\phi_{\mathcal{B}'}(w_j) = e_j$, $j = 1, \dots, n$. La matrice M è l'isomorfismo $\phi_{\mathcal{B}} \cdot \phi_{\mathcal{B}'}^{-1}$, quindi $M^j = M(e_j) = \phi_{\mathcal{B}}(w_j) =$ coordinate di w_j nella base $\mathcal{B}$. Analogamente la $j-$esima colonna di M^{-1} contiene le coordinate di v_j nella base $\mathcal{B}'$. Chiamando $\mathcal{B}$ base vecchia e $\mathcal{B}'$ base nuova possiamo dire: M ha per colonne le coordinate dei vettori della base nuova rispetto alla base vecchia, mentre la sua inversa $M^{-1} = \phi_{\mathcal{B}'} \cdot \phi_{\mathcal{B}}^{-1}$ ha per colonne le coordinate dei vettori della base vecchia rispetto alla base nuova.

Esempio. Vediamo quanto detto su un esempio concreto semplice.

Partiamo dal sistema di equazioni differenziali

$$\begin{cases} \dot{x} = x + 2y \\ \dot{y} = 2x + y \end{cases}$$

e indichiamo con A la matrice del sistema $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$.

La soluzione generale, senza condizioni iniziali, è data da $\begin{pmatrix} x \\ y \end{pmatrix} = e^{At} \begin{pmatrix} a \\ b \end{pmatrix}$

Osserviamo che la matrice A è diagonalizzabile: ha due autovalori distinti $3, -1$. Pertanto in una base di autovettori (ad esempio $v_1 = e_1 + e_2, v_2 = e_1 - e_2$) si ha $\widetilde{A} = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$ e quindi $e^{\tilde{A}t} = \begin{pmatrix} e^{3t} & 0 \\ 0 & e^{-t} \end{pmatrix}$.

Detta C la matrice del cambiamento di base si ha

$$e^{At} = Ce^{\tilde{A}t}C^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} e^{3t} & 0 \\ 0 & e^{-t} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} e^{3t}+e^{-t} & e^{3t}-e^{-t} \\ e^{3t}-e^{-t} & e^{3t}+e^{-t} \end{pmatrix}.$$

Quindi si ha

$$\begin{pmatrix} x \\ y \end{pmatrix} = e^{At}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a(e^{3t}+e^{-t}) + b(e^{3t}-e^{-t}) \\ a(e^{3t}-e^{-t}) + b(e^{3t}+e^{-t}) \end{pmatrix}$$

da cui si ottiene per la soluzione generale

$$\begin{cases} x = (a+b)e^{3t} + (a-b)e^{-t} \\ y = (a+b)e^{3t} + (-a+b)e^{-t} \end{cases}$$

2 Forma canonica di Jordan.

Per semplificare la trattazione in questo paragrafo penseremo le matrici a coefficienti complessi.

Sia $A \in \mathcal{M}(n, \mathbb{C})$ e siano $\lambda_1, \dots, \lambda_k$ i suoi autovalori con molteplicità algebriche rispettivamente $m_1, \dots, m_k$; indicati con $V_{\lambda_1}, \dots, V_{\lambda_k}$ gli autospazi di A, in generale risulta $\dim V_{\lambda_i} \leqslant m_i$.

Se A è diagonalizzabile, $\mathbb{C}^n$ è somma diretta degli autospazi V_{λ_i} di A,

$$\mathbb{C}^n = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$$

e in questo caso si ha $\dim V_{\lambda_i} = m_i$ per ogni i. Se A non è diagonalizzabile, $\mathbb{C}^n$ si può ancora decomporre in somma diretta di k spazi invarianti $W_1, \dots, W_k$ tali che

- $A_{|W_i}$ ha solo l'autovalore λ_i
- $\dim W_i = m_i$

Il primo passo è la separazione degli autovalori che è l'argomento del seguente teorema.

Teorema 2.1. *Sia A una matrice con autovalori $\lambda_1, \dots, \lambda_k$ di molteplicità algebrica $m_1, \dots, m_k$. Allora si ha una decomposizione*

$$C^n = W_1 \oplus \cdots \oplus W_k$$

dove

1. *W_j è invariante per A,* $\quad j = 1, \dots, k,$

2. *la restrizione di A a W_j ha solo l'autovalore λ_j.*

3. *dim $W_j = m_j$.*

La prova di questo teorema è conseguenza del cosiddetto Lemma di Fitting.

Lemma 2.2. *Sia V uno spazio vettoriale di dimensione n e $T : V \to V$ un endomorfismo. Esiste h con $0 \leqslant h \leqslant n$ tale che $V = \operatorname{Ker} T^h \oplus \operatorname{Im} T^h$.*

Prova. Consideriamo le due successioni di sottospazi di V

$$\operatorname{Ker} T \subset \operatorname{Ker} T^2 \subset \cdots \subset \operatorname{Ker} T^i \subset \cdots$$

$$\operatorname{Im} T \supset \operatorname{Im} T^2 \supset \cdots \supset \operatorname{Im} T^i \supset \cdots$$

e naturalmente si ha dim $\operatorname{Ker} T^i$+dim $\operatorname{Im} T^i = n$. Se $\operatorname{Ker} T^h = \operatorname{Ker} T^{h+1}$ e $v \in \operatorname{Ker} T^{h+2}$ allora $T(v) \in \operatorname{Ker} T^{h+1} = \operatorname{Ker} T^h$. Quindi $v \in \operatorname{Ker} T^{h+1}$. Pertanto $\operatorname{Ker} T^{h+2} = \operatorname{Ker} T^{h+1} = \operatorname{Ker} T^h$. Iterando il ragionamento si ha che il nucleo delle potenze di T si stabilizza da h in poi. Per il teorema della dimensione anche la catena delle immagini si stabilizza allo stesso livello h. Inoltre si ha $\operatorname{Ker} T^h \cap \operatorname{Im} T^h = \{0\}$. Infatti $T^h : \operatorname{Im} T^h \to \operatorname{Im} T^{2h}$ è un isomorfismo e quindi $\operatorname{Im} T^h$ interseca $\operatorname{Ker} T^h$ solo in 0. Quindi si ha $V = \operatorname{Ker} T^h \oplus \operatorname{Im} T^h$. □

Dal lemma segue immediatamente il teorema.

Prova. Per induzione su k. Per $k = 1$ c' è un solo autovalore λ_1 e $\mathbb{C}^n = W_1$ soddisfa le tre condizioni.

Supponiamo il teorema vero per endomorfismi con $k-1$ autovalori e proviamolo per l'endomorfismo A che ha k autovalori.

Applichiamo il Lemma di Fitting a $S = A - \lambda_1 I : \mathbb{C}^n \to \mathbb{C}^n$. Troviamo h tale che $\mathbb{C}^n = \operatorname{Ker} S^h \oplus \operatorname{Im} S^h$. Osserviamo che entrambi gli addendi sono invarianti per S e quindi sono anche invarianti per A. Dunque $A : \operatorname{Im} S^h \to \operatorname{Im} S^h$ e il polinomio caratteristico della restrizione di A a $\operatorname{Im} S^h$ è un fattore del polinomio caratteristico di A che non ha più la radice λ_1: infatti tutti gli autovettori di autovalore λ_1 sono in $\operatorname{Ker} S \subset \operatorname{Ker} S^h$ e quest'ultimo interseca $\operatorname{Im} S^h$ solo in 0. Quindi la restrizione di A verifica l'ipotesi induttiva e dunque $\operatorname{Im} S^h = V_2' \oplus \cdots \oplus V_k'$. Ma se $v \in V_j'$ si ha $(A - \lambda_j I)^h(v) = (A_{|_{\operatorname{Im} S^h}} - \lambda_j I)^h(v) = 0$, quindi l'autospazio generalizzato V_j' della restrizione di A coincide con l'autospazio generalizzato W_j di A. Dunque $\mathbb{C}^n = W_1 \oplus ImS^h = W_1 \oplus W_2 \oplus \cdots \oplus W_k$ e il teorema è provato. □

Abbiamo separato gli autovalori di A.

Definizione 2.3. *Gli addendi di $\mathbb{C}^n = \sum_{j=1}^k W_j$ si chiamano* autospazi generalizzati *di A. Si ha $W_j = \operatorname{Ker}(A - \lambda_j I)^{h_j}$, dove h_j è l'esponente che compare nel lemma di Fitting e la dimensione di W_j è la molteplicità algebrica dell'autovalore λ_j.*

Il secondo passo è la decomposizione in blocchi di Jordan della restrizione a W_j dell'applicazione nilpotente $A - \lambda_j I$.

Sia W un autospazio generalizzato di A e sia λ il suo autovalore.

Definizione 2.4. *Un blocco di Jordan relativo a λ di ordine r è la matrice $r \times r$*

$$J_r(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ & & & \dots & & \\ 0 & 0 & \dots & 0 & \lambda & 1 \\ 0 & 0 & \dots & 0 & 0 & \lambda \end{pmatrix}$$

Il teorema di Jordan per W è il seguente.

Teorema 2.5. *Esiste una base di W tale che la matrice associata alla restrizione di A a W rispetto a questa base è composta da blocchi di Jordan relativi a λ e tale decomposizione è unica a meno dell'ordine dei blocchi.*

Osservazione 2.6. Se $A_{|W} : W \to W$ è diagonalizzabile, allora i blocchi di Jordan sono tutti di ordine 1 e W è l'autospazio di A relativo all'autovalore λ.

Prova. $W = \operatorname{Ker}(A - \lambda I)^h$. Chiamiamo $S = A - \lambda I$. $W = \operatorname{Ker} S^h \supset \operatorname{Ker} S^{h-1}$. Quindi $\operatorname{Ker} S^h = \operatorname{Ker} S^{h-1} \oplus V_h$ dove V_h è un supplementare di $\operatorname{Ker} S^{h-1}$ ed è formato da vettori v tali che $S^h(v) = 0$, ma $S^{h-1}(v) \neq 0$. Sia $r_h = \dim V_h$. A sua volta $\operatorname{Ker} S^{h-1} = \operatorname{Ker} S^{h-2} \oplus V_{h-1}, S(V_h) \subset V_{h-1}$, dove se $v \in V_{h-1}$, $S^{h-1}(v) = 0$ ma $S^{h-2}(v) \neq 0$. Andando avanti avremo: $\operatorname{Ker} S^{j+1} = \operatorname{Ker} S^j \oplus V_{j+1}, V_{j+1} \supset S(V_{j+2})$, $\dim V_{j+1} = r_{j+1}$ e i suoi vettori vanno a zero dopo $j+1$ passi e non dopo j passi. L'ultimo passo è $\operatorname{Ker} S^2 = \operatorname{Ker} S \oplus V_2, V_2 \supset S(V_3)$, $\dim V_2 = r_2$.

Consideriamo ora una base B_h di V_h. È chiaro che $S(B_h) \subset V_{h-1}$, $S^2(B_h) \subset V_{h-2}$, $S^{h-1}(B_h) \subset \operatorname{Ker} S$. Per costruzione i vettori di $B_h \cup S(B_h) \cup S^2(B_h) \cup \dots \cup S^{h-1}(B_h)$ sono indipendenti, perché per ogni j $\operatorname{span}(S^j(B_h)) \cap \operatorname{Ker} S^{j-1} = 0$.

Ogni elemento di B_h genera un blocco di Jordan $J_h(0)$ di ordine h. Infatti sia $v \in B_h$, allora $S^h(v) = 0$ quindi $S^h(v) = S(v_1), v_1 \in \operatorname{Ker} S, v_1 = S(v_2), v_2 \in \operatorname{Ker} S^2, v_2 = S(v_3), v_3 \in \operatorname{Ker} S^3, \dots, v_{h-1} \in \operatorname{Ker} S^{h-1}, v_{h-1} = S(v)$ e la matrice di S ristretta a $\operatorname{span}(v_1, \dots v_{h-1}, v)$ rispetto alla base $(v_1, \dots v_{h-1}, v)$ è proprio il blocco di Jordan $J_h(0)$.

Abbiamo così r_h blocchi di Jordan di ordine h.

Dentro V_{h-1} ci sono gli r_h vettori indipendenti di $S(B_h)$. Completiamo a base $S(B_h) \cup B_{h-1}$. Come prima, gli elementi di B_{h-1} generano $r_{h-1} - r_h$ blocchi di Jordan di ordine $h-1$. Naturalmente se $r_{h-1} - r_h = 0$ tali blocchi non ci sono e si prosegue. Dentro V_j ci sono i vettori indipendenti di $S^{h-j}(B_h) \cup S^{h-j-1}(B_{h-1}) \cup \dots \cup S(B_{j+1})$. Completiamo a base con B_j. Gli elementi di B_j danno blocchi di Jordan di ordine j. Quando arriviamo a $\operatorname{Ker} S$ dentro ci troviamo tutti i vettori relativi alle prime colonne dei blocchi di Jordan già costruiti, che sono vettori indipendenti. Se essi non generano $\operatorname{Ker} S$ completiamo a base con B_1 che dà luogo a blocchi di Jordan di ordine 1.

La forma di Jordan J_S per $S : W \to W$ è completa. Per ottenere quella della restrizione di A a W basta prendere $\lambda I + J_S$.

Siamo arrivati alla conclusione.

Facendo questo lavoro per tutti gli autospazi generalizzati si ottiene che A è simile ad una matrice che ha diagonalmente $\lambda_j I_{m_j} + J_{S_j}, j = 1, \ldots, k$ dove $S_j = A_{|W_j} - \lambda_j I_{m_j}$ e J_{S_j} è la forma di Jordan di S_j. □

Abbiamo provato sostanzialmente il seguente enunciato.

Teorema 2.7. *Ogni matrice a coefficenti complessi è simile ad una matrice a blocchi di Jordan e tale matrice è unica a meno dell'ordine dei blocchi. Inoltre due matrici sono simili se e solo se hanno la stessa forma di Jordan*

Prova. Che due matrici che hanno la stessa forma di Jordan siano simili è chiaro. Per il viceversa occorre fare un piccolo conto che dimostrerà che la forma di Jordan dipende solo dalla matrice A. Vogliamo calcolare per ogni $j, 1 \leqslant j \leqslant h$ il numero dei blocchi di Jordan di ordine j nell'autospazio generalizzato W_i relativo all'autovalore λ_i. Denotiamo con $S = (A - \lambda_i I)_{|W_i}$. Quindi $W_i = \operatorname{Ker} S^h$. Definiamo $k_l = \dim \operatorname{Ker} S^l$. I blocchi di Jordan di ordine h sono dunque $k_h - k_{h-1} = \dim V_h$. I blocchi di Jordan di ordine $h-1$ sono $k_{h-1} - k_{h-2} - (k_h - k_{h-1}) = 2k_{h-1} - h_k - k_{h-2}$ perché in $S(V_h)$ ci sono i penultimi vettori relativi ai blocchi di ordine h. I blocchi di Jordan di ordine j sono $k_j - k_{j-1} - (k_{j+1} - k_j) = 2k_j - k_{j+1} - k_{j-1}$ perché tutti i blocchi di Jordan di ordine maggiore di j hanno un elemento in V_{j+1}. Con lo stesso ragionamento, i blocchi di Jordan di ordine 2 sono $2k_2 - k_3 - k_1$, mentre i blocchi di ordine 1 sono $2k_1 - k_2$.

Gli interi k_l sono dati da $n - rk(A - \lambda_i I)^l$, quindi dipendono solo da A. Inoltre, se A e B sono matrici simili, esse sono matrici associate allo stesso endomorfismo di $\mathbb{C}^n$ e quindi $rk(A - \lambda_i I)^l = rk(B - \lambda_i I)^l$ per ogni l maggiore o uguale a 1 e minore o uguale a h. Pertanto la forma di Jordan di A è la stessa di quella di B. □

Dunque la forma di Jordan ci fornisce un sistema completo di invarianti per le matrici complesse e per le matrici reali che hanno solo autovalori reali.

Osservazione 2.8. Per un dato autospazio generalizzato W_j, il numero dei blocchi di Jordan, compresi quelli di ordine 1, ci dà la dimensione $m_g(\lambda_j)$ dell'autospazio V_{λ_j}. Dall'ordine dei vari blocchi si possono dedurre le dimensioni dei $\operatorname{Ker} S_j^l, l = 1, \ldots, h_j - 1$, mentre dim $W_j = m_j$. Si può anche calcolare il polinomio minimo di A: esso sarà $(x - \lambda_1)^{h_1}(x - \lambda_2)^{h_2} \ldots (x - \lambda_k)^{h_k}$ dove h_j è il massimo ordine dei blocchi di Jordan di J_{S_j}.

Per finire sottoliniamo ancora che tutto quello che abbiamo fatto non dipende dalla scelta delle basi $B_h, \ldots, B_1$, ma solo dalle dimensioni degli spazi costruiti che dipendono a loro volta dalle dimensioni dei nuclei $\mathrm{Ker}\, S_j^l$, cioè in definitiva da A, come abbiamo visto nella prova dell'ultimo teorema.

3 Forma reale del Teorema di Jordan

Sia A una matrice reale. Se gli autovalori di A sono tutti reali, vale per A il Teorema 2.5, cioè A è simile ad una matrice a blocchi di Jordan, come abbiamo già osservato. Supponiamo ora che invece A abbia una coppia (o più coppie) di autovalori complessi coniugati, $\lambda = a + ib$, $\overline{\lambda} = a - ib$. Per fissare le idee, supponiamo $b > 0$.

A ha una forma di Jordan complessa e due autospazi generalizzati con la loro forma di Jordan W_λ e $W_{\overline{\lambda}}$. Poiché A è reale, i due autospazi generalizzati hanno la stessa dimensione $m_a(\lambda) = m_a(\overline{\lambda})$, gli stessi blocchi e $m_g(\lambda) = m_g(\overline{\lambda})$,

Per ottenere una forma di Jordan reale il trucco è semplice: scelta una base per qualcosa in W_λ, si sceglie la base coniugata in $W_{\overline{\lambda}}$ per lo stesso qualcosa. Fatto ciò, si sceglie opportunamente una base reale per la somma diretta dei due sottospazi generalizzati che ci darà la forma di Jordan reale.

Sia $(z_1, \ldots, z_m)$ una base di Jordan di W_λ (quindi $(\overline{z_1}, \ldots, \overline{z_m})$ è una base di Jordan per $W_{\overline{\lambda}}$). Consideriamo i vettori reali, per $j = 1, \ldots, m$:

$$x_j = \frac{z_j + \overline{z_j}}{2}$$
$$y_j = \frac{z_j - \overline{z_j}}{2i}$$

Questi $2m$ vettori sono indipendenti e quindi costituiscono una base di $W_\lambda \oplus W_{\overline{\lambda}}$. Calcoliamo $A(x_j)$ e $A(y_j)$, ricordando che $A(z_j) = \lambda z_j$ se z_j è un autovettore, altrimenti $A(z_j) = \lambda z_j + z_{j-1}$ per come sono fatti i blocchi di Jordan.

Quindi, se z_j è autovettore si ha:

$$A(x_j) = A\left(\frac{z_j + \overline{z_j}}{2}\right) = \frac{\lambda z_j + \overline{\lambda}\overline{z_j}}{2} = Re(\lambda z_j) = Re(\lambda)x_j - Im(\lambda)y_j,$$
$$A(y_j) = A\left(\frac{z_j - \overline{z_j}}{2i}\right) = \frac{\lambda z_j - \overline{\lambda}\overline{z_j}}{2i} = Im(\lambda z_j) = Im(\lambda)x_j + Re(\lambda)y_j.$$

Se z_j non è autovettore, allora

$$A(x_j) = Re(\lambda)x_j - Im(\lambda)y_j + x_{j-1}$$

$$A(y_j) = Im(\lambda)x_j + Re(\lambda)y_j + y_{j-1}.$$

Questo significa che per ogni coppia di blocchi $J_r(\lambda)$ e $J_r(\overline{\lambda})$ nella forma di Jordan complessa $J(A)$, nella forma reale $J_{\mathbb{R}}(A)$ al posto di questi due blocchi ci sarà un blocco $2r \times 2r$ della forma

$$\begin{pmatrix} H_\lambda & I & & \dots & & \\ & H_\lambda & I & & \dots & \\ & & & \dots & & \\ & & \dots & & H_\lambda & I \\ & & \dots & & & H_\lambda \end{pmatrix}$$

dove

$$H_\lambda = \begin{pmatrix} Re(\lambda) & Im(\lambda) \\ -Im(\lambda) & Re(\lambda) \end{pmatrix} \qquad \text{e} \qquad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Riassumiamo la procedura.

1. Si separano gli autovalori reali e complessi di A e si definiscono i relativi autospazi generalizzati.

2. Per ogni autovalore reale α, si sceglie una base di Jordan di W_α come indicato nella prova del Teorema 2.5.

3. Se λ e $\overline{\lambda}$ sono autovalori complessi coniugati di A con $Im(\lambda) > 0$, si sceglie una base di Jordan in W_λ e la base coniugata in $W_{\overline{\lambda}}$.

4. Si definisce la base reale di $W_\lambda \oplus W_{\overline{\lambda}}$, $x_j = Re(z_j), y_j = Im(z_j), j = 1, \dots, m$, $m = m_a(\lambda)$.

5. Si sostituiscono i blocchi $J_r(\lambda)$ e $J_r(\overline{\lambda})$ con il blocco $2r \times 2r$ che ha le matrici 2×2 H_λ sulla diagonale e con le matrici I_2 al posto degli 1 sulla sopra diagonale. Si fa questo lavoro per tutti i blocchi e per tutte le coppie di autovalori complessi coniugati di A.

4 Conclusione per e^A.

Torniamo all'esponenziale di matrici. Se J è una matrice di Jordan che ha blocchi $J_1, \dots J_s$ si ha che J^p ha blocchi $J_1^p, \dots J_s^p$.

Quindi e^J è una matrice a blocchi che sono dati da $e^{J_1}, \dots, e^{J_s}$. Inoltre $J_i = \lambda_i I + N_i$ dove N_i è nilpotente, dunque $e^{J_i} = e^{\lambda_i} I_{m_i}$ per un polinomio in N_i.

In conclusione e^A, dove A non è diagonalizzabile, è simile alla matrice e^J dove J è la forma canonica di Jordan della classe di similitudine di A.

Se A è una matrice reale con qualche autovalore complesso, nella forma di Jordan di A ci sono i blocchi di Jordan $J_1, \dots J_s$ relativi agli autovalori reali e i blocchi $J_1^{\mathbb{R}}, \dots, J_t^{\mathbb{R}}$. Anche in questo caso la potenza p-esima di $J_{\mathbb{R}}(A)$ è a blocchi che sono le potenze p-esime dei blocchi iniziali. Poiché se scriviamo i blocchi reali come somma della diagonale con le matrici H_λ e della matrice nilpotente che contiene gli I_2, i due addendi commutano, come è facile verificare, si avrà ancora che e^A è simile a $e^{J_{\mathbb{R}}(A)}$, matrice che ha una decomposizione simile al caso complesso con gli ovvii cambiamenti per i blocchi reali, nel senso che il blocco $e^{J^{\mathbb{R}}(\lambda)}$ è il prodotto di e^H per un polinomio nell'altro addendo, dove H è la matrice con i blocchi 2×2 di matrici H_λ.

5 Osservazioni ed esercizi.

Nel seguito useremo le seguenti definizioni e notazioni.

Diremo che una matrice quadrata A di ordine n è *triangolare superiore* se per ogni $i < j$ si ha $a_{i,j} = 0$ e che A è *strettamente triangolare superiore* se anche gli elementi sulla diagonale principale sono nulli, cioè se per ogni i si ha $a_{i,i} = 0$.

Indicheremo per $k = 1, 2, \dots$ con D_k la k-esima parallela sopra la diagonale principale $D(= D_0)$, cioè la diagonale i cui elementi sono $a_{1,1+k}, \dots, a_{i,i+k}, \dots, a_{n-k,n}$.

Esercizio 1. Sia V uno spazio vettoriale di dimensione n e $f : V \to V$ un suo endomorfismo. Sia v un vettore tale che $f^n(v) = 0$ e $f^{n-1}(v) \neq 0$. Provare che i vettori $v, f(v), f^2(v), \dots, f^{n-1}(v)$ sono indipendenti.

Scrivere la matrice associata ad f nella base $(f^{n-1}(v), f^{n-2}(v), \dots, f(v), v)$.

Esercizio 2. Sia $f : V \to V$ un endomorfismo di uno spazio vettoriale V la cui matrice associata nella base $\{v_1, \dots, v_n\}$ sia una matrice strettamente triangolare superiore A.

-)Provare che $A^n = 0$ cioè $f^n = 0$.

Supponiamo inoltre che la matrice strettamente triangolare superiore A abbia gli elementi su D_1, la prima parallela alla diagonale principale, tutti diversi da zero.

-) Provare che allora si ha $f^{n-1} \neq 0$,

cioè che l'indice di nullità [1] di f è esattamente n.

Suggerimenti.

-) La prima affermazione può essere facilmente provata considerando i sottospazi $V = V_n = \text{span} < v_1, \dots, v_n >, \dots, V_k = \text{span} < v_1, \dots, v_k >, \dots, V_0 = 0$ *ed*

[1] *Indice di nullità* di f il più piccolo intero k per cui $f^k = 0$ e $f^{k-1} \neq 0$

osservando che $f(V_k) \subset V_{k-1}$*; pertanto* $f^n(V) = f^n(V_n) \subset f^{n-1}(V_{n-1}) \subset \ldots \subset f(V_1) = V_0 = 0$.

-) Per la seconda si osservi che, con le notazioni appena introdotte, $f(v_j) = 0 \cdot v_j + a_{j-1,j}v_{j-1} + w$ *con* $w \in V_{j-2}$ *e quindi si ha*

$$f^n(v_n) = f^{n-1}f(v_n) = f^{n-1}(a_{n-1,n}v_{n-1} + w) = f^{n-1}(a_{n-1,n}v_{n-1})$$

poiché, essendo $w \in V_{n-2}$*, si ha* $f^{n-1}(w) = 0$*. Quindi*

$$\begin{aligned} f^{n-1}(a_{n-1,n}v_{n-1}) = f^{n-2}f(a_{n-1,n}v_{n-1}) = f^{n-2}(a_{n-1,n}f(v_{n-1})) = \\ = f^{n-2}(a_{n-1,n}a_{n-2,n-1}v_{n-2} + w') \end{aligned}$$

con $w' \in V_{n-3}$.

Procedendo in questo modo si perviene a

$$f^{n-1}(v_n) = a_{n-1,n} \cdot a_{n-2,n-1} \cdots a_{1,2}v_1$$

che è un vettore non nullo per l'ipotesi fatta sugli elementi nella parallela D_1 *alla diagonale principale.*

Pertanto per l'endomorfismo f *abbiamo trovato un vettore* v*, il vettore* v_n*, per cui si ha* $f^n(v) = 0, f^{n-1}(v) \neq 0$ *e quindi (vedi es. 1) i vettori* $\{v, f(v), f^2(v), \ldots, f^{n-1}(v)\}$ *formano una base di Jordan per* f *la cui matrice associata è pertanto un blocco di Jordan di ordine* n.

Esercizio 3. Sia J la matrice del blocco di Jordan di ordine n relativo all'autovalore $\lambda \neq 0$, cioè

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ldots & 0 \\ & & \ldots & & \\ 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & \ldots & 0 & \lambda \end{pmatrix}$$

Calcolare la forma di Jordan delle matrici J^k.

Suggerimento. Detto $J_n(0)$ *il blocco di Jordan di ordine* n *relativo all'autovalore* 0*, si ha* $J_n(\lambda) = \lambda I_n + J_n(0)$.

Poiché λI *e* $J_n(0)$ *commutano si ha che*

$$J_n(\lambda)^k = (\lambda I + J_n(0))^k = \lambda^k I + k\lambda^{k-1}J_n(0) + \ldots$$

pertanto la matrice $J_n(\lambda)^k$ *è di questo tipo*

$$J_n(\lambda)^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} & 0 & \dots & 0 \\ 0 & \lambda^k & k\lambda^{k-1} & \dots & 0 \\ & & \dots & & \\ 0 & 0 & \dots & \lambda^k & k\lambda^{k-1} \\ 0 & 0 & \dots & 0 & \lambda^k \end{pmatrix}$$

quindi ha il solo autovalore λ^k e ha la diagonale D_1 sopra la diagonale principale con tutti elementi $\neq 0$. Pertanto (vedi esercizi precedenti) la matrice di Jordan di tale matrice è formata di un solo blocco di ordine massimo.

Diverso è il caso se l'autovalore è zero (vedi esercizio seguente)

Esercizio 4. Sia J la matrice del blocco di Jordan di ordine n relativo all'autovalore 0, cioè

$$J_n(\lambda) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

Calcolare la forma di Jordan delle matrici J^k.

Suggerimento Vediamo un caso particolare con $n = 11$. In altri termini, detto f l'endomorfismo relativo alla matrice $J_{11}(0)$ abbiamo una base di Jordan costituita da 11 elementi, del tipo $\{v, f(v), f^2(v), \dots, f^{10}(v)\}$ con $f^{11}(v) = 0$. Consideriamo il caso $k = 4$.

È immediato verificare che

$$\{v, f^4(v), f^8(v)\} \quad \{f(v), f^5(v), f^9(v)\} \quad \{f^2(v), f^6(v), f^{10}(v)\}$$

danno luogo a 3 blocchi di Jordan di ordine 3 e *$\{f^3(v), f^7(v)\}$ un blocco di odine 2.*

Schematicamente rappresentiamo la situazione raggruppando i vettori della base in questo modo

$$\begin{matrix} v & f(v) & f^2(v) & f^3(v) \\ f^4(v) & f^5(v) & f^6(v) & f^7(v) \\ f^8(v) & f^9(v) & f^{10}(v) & \end{matrix}$$

Ragionando in maniera analoga in termini più generali si vede che se $n = qk + r$, la matrice di Jordan di J^k sarà composta da r blocchi di ordine $q + 1$ e da $k - r$ blocchi di ordine q. Si osservi tra l'altro che $n = r(q + 1) + (k - r)q = r + kq$.

$$\begin{matrix} v & f(v) & f^2(v) & \dots & \dots & \dots & f^{k-1} \\ f^k(v) & f^{k+1}(v) & f^{k+2}(v) & \dots & \dots & \dots & f^{2k-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ f^{(q-1)k}(v) & f^{(q-1)k+1}(v) & f^{(q-1)k+2}(v) & \dots & \dots & \dots & f^{qk-1} \\ f^{qk}(v) & f^{qk+1}(v) & f^{qk+2}(v) & \dots & f^{qk+r-1}(v) & & \end{matrix}$$

Nel paragrafo 3 abbiamo mostrato come ricavare dalla forma complessa di Jordan quella reale. Ma se si è interessati a sapere solo se due matrici sono simili avvalendosi della loro forma di Jordan tale passaggio non è necessario come mostra l'esercizio seguente.

Esercizio 5. Siano A e B due matrici quadrate di ordine n a coefficienti reali. Mostrare che esse sono simili come matrici reali se e solo se lo sono pensate come matrici complesse.

Suggerimento. Dire che A e B sono simili come matrici reali significa che esiste una matrice reale invertibile M tale che $M^{-1}AM = B$ cioè $MB = AM$.

Se due matrici sono simili su $\mathbb{R}$ evidentemente sono simili su $\mathbb{C}$.

Per mostrare il viceversa occorre mostrare che se esiste una matrice invertibile a coefficienti complessi M tale che $AM = MB$ ne esiste una invertibile a coefficienti reali che verifica la stessa relazione.

Scriviamo $M = M_1 + iM_2$. Essendo A e B reali abbiamo che $AM_1 = M_1B$ e $AM_2 = M_2B$ e una delle matrici M_1, M_2 andrebbe bene se fosse invertibile.

Consideriamo le matrici $M_1 + xM_2$: il loro determinante è una funzione polinomiale reale $p(x)$ di grado n che, se non identicamente nulla, ammette almeno un x reale per cui è diversa da 0. Ma il fatto che la matrice $M = M_1 + iM_2$ sia invertibile ci assicura che $p(i) \neq 0$ e quindi che la funzione polinomiale $p(x)$ non è identicamente nulla.

Esercizio 6. Sia A una matrice quadrata di rango 1. Calcolare e^A.

Suggerimento. Se A ha rango 1 ha un nucleo di dimensione $n - 1$. A può essere diagonalizzabile, nel qual caso ha un autovalore $\lambda \neq 0$ e quindi anche e^A è diagonalizzabile. Se invece A ha solo l'autovalore 0 è facile provare che $A^2 = 0$, quindi $e^A = I + A$.

Esercizio 7. Sia A una matrice reale 6×6. Si supponga che il polinomio caratteristico di A abbia una radice complessa $\alpha = a + ib$ di molteplicità algebrica 1 e una radice complessa $\beta = c + id$ di molteplicità algebrica 2 e di molteplicità geometrica 1. Trovare la forma di Jordan reale di A.