

"Teoria dei sistemi" - Appendice 2

Spazi vettoriali e algebra matriciale

Gli spazi vettoriali	2
Definizione di campo e di spazio vettoriale	2
Vettori linearmente indipendenti e linearmente dipendenti	3
Dimensione di uno spazio vettoriale	3
Norma in uno spazio vettoriale	5
<i>Norme definibili in \hat{A}^n</i>	6
<i>Esempio</i>	6
<i>Osservazione</i>	7
Base di uno spazio vettoriale	7
<i>Cambiamento di base</i>	8
<i>Base canonica di uno spazio vettoriale</i>	10
Algebra matriciale	10
Rango di una matrice	10
<i>Disuguaglianza di Sylvester</i>	11
Range di una matrice	12
Autovalori e autovettori di una matrice.....	12
Proprietà: autovalori e determinante	16
Diagonalizzazione di una matrice	18
Forma canonica di Jordan	20
<i>Esempio</i>	23
<i>Esempio</i>	23
<i>Esempio</i>	25
<i>Rango degli autovettori generalizzati</i>	26
Teorema di Cayley-Hamilton	28
<i>Conseguenza: riduzione di polinomi</i>	29
<i>Esempio</i>	29
Polinomio minimo di una matrice	30
<i>Esempio</i>	31
<i>Osservazione</i>	31
Definizioni varie.....	32
Richiami sui sistemi lineari	32

Gli spazi vettoriali

DEFINIZIONE DI CAMPO E DI SPAZIO VETTORIALE

Sia dato un insieme K di elementi. Si dice che questo insieme K è un “**campo**” se e solo se sono verificate le seguenti condizioni:

- in primo luogo, per gli elementi di tale insieme devono essere definite le operazioni di “*somma*” e di “*prodotto*”;
- in secondo luogo, l’operazione di “*somma*” deve godere delle seguenti proprietà:
 - * commutativa
 - * associativa
 - * distributiva della somma rispetto al prodotto
 - * esistenza dell’elemento neutro
 - * esistenza dell’elemento simmetrico;
- in terzo luogo, l’operazione di “*prodotto*” deve godere delle seguenti proprietà:
 - * commutativa
 - * associativa
 - * distributiva del prodotto rispetto alla somma
 - * esistenza dell’elemento neutro
 - * esistenza dell’elemento simmetrico.

Siano dati ora un campo K e un generico insieme di elementi V . Si dice che questo insieme V è uno “**spazio vettoriale definito sul campo K** ” se e solo se sono verificate le seguenti condizioni:

- in primo luogo, per gli elementi di tale insieme devono essere definite le operazioni di “*somma*” e di “*prodotto per uno scalare $\hat{I} K$* ”;
- in secondo luogo, l’operazione di “*somma*” tra due elementi di V deve godere delle seguenti proprietà:
 - * commutativa
 - * associativa
 - * distributiva della somma rispetto al prodotto
 - * esistenza dell’elemento neutro
 - * esistenza dell’elemento simmetrico;

- in terzo luogo, l'operazione di "prodotto" tra un elemento di V e uno di K deve godere delle seguenti proprietà:
 - * commutativa
 - * associativa
 - * distributiva del prodotto rispetto alla somma
 - * esistenza dell'elemento neutro
 - * esistenza dell'elemento simmetrico.

Citiamo velocemente alcuni esempi importanti e ricorrenti di spazi vettoriali:

- è certamente uno spazio vettoriale, definito sul campo \mathfrak{R} dei numeri reali, l'insieme \mathfrak{R}^n , ossia l'insieme delle n -ple di numeri reali;
- è anche uno spazio vettoriale, definito sul campo C dei numeri complessi, l'insieme C^2 .

VETTORI LINEARMENTE INDIPENDENTI E LINEARMENTE DIPENDENTI

Consideriamo un generico spazio vettoriale V e prendiamo una generica n -pla (x_1, x_2, \dots, x_n) di suoi elementi (ricordiamo che gli elementi di uno spazio vettoriale si dicono "**vettori**"); consideriamo inoltre n coefficienti reali k_1, k_2, \dots, k_n ; sulla base di questi coefficienti, operiamo una combinazione lineare degli elementi di V : otteniamo in tal modo l'elemento

$$y = k_1 x_1 + k_2 x_2 + \dots + k_n x_n$$

Questo elemento può o meno essere l'elemento nullo, a seconda sia del valore dei coefficienti k_i sia del valore degli elementi di V . Sussistono allora le seguenti due definizioni:

- se $\exists (k_1, k_2, \dots, k_n) \in \mathfrak{R}^n$ con $(k_1, k_2, \dots, k_n) \neq (0, 0, \dots, 0)$ e tale che $y = 0$, si dice che gli elementi x_1, x_2, \dots, x_n sono **linearmente dipendenti**;
- se non esiste alcuna $(k_1, k_2, \dots, k_n) \in \mathfrak{R}^n$ con $(k_1, k_2, \dots, k_n) \neq (0, 0, \dots, 0)$ e tale che $y = 0$, si dice che gli elementi x_1, x_2, \dots, x_n sono **linearmente indipendenti**.

In termini meno formali, dati gli elementi (x_1, x_2, \dots, x_n) dello spazio vettoriale V , se è possibile farne almeno una combinazione lineare (mediante coefficienti non tutti nulli) che dia l'elemento nullo, essi sono linearmente dipendenti, mentre invece, se non è possibile trovare tale combinazione lineare, essi sono linearmente indipendenti.

DIMENSIONE DI UNO SPAZIO VETTORIALE

Al concetto di vettori linearmente indipendenti è strettamente legato il concetto di "*dimensione*" di uno spazio vettoriale. Supponiamo di avere un generico spazio vettoriale V : ci chiediamo qual è il numero massimo di vettori linearmente indipendenti che possiamo trovare in tale spazio.

Si intuisce subito che il minimo numero di vettori linearmente indipendenti, di un qualsiasi spazio vettoriale, è 1: infatti, un qualsiasi vettore di V , che non sia il vettore nullo, potrà senz'altro essere

moltiplicato per un coefficiente non nullo senza dare, come risultato del prodotto, il vettore nullo. Vediamo allora di rispondere alla domanda in alcuni casi particolari.

Abbiamo in precedenza detto che l'insieme \mathfrak{R} dei numeri reali è uno spazio vettoriale: se prendiamo due qualsiasi elementi x_1 e x_2 di tale spazio, cioè appunto due numeri reali, è chiaro che essi non sono linearmente indipendenti, in quanto esisterà senz'altro una coppia (k_1, k_2) di coefficienti tali che $k_1 x_1 + k_2 x_2 = 0$. Deduciamo, quindi, che, nello spazio vettoriale \mathfrak{R} , è possibile trovare 1 solo vettore linearmente indipendente.

Passiamo ad \mathfrak{R}^2 : prendiamo due qualsiasi vettori di questo spazio vettoriale, ossia due coppie di numeri reali del tipo (a_1, b_1) e (a_2, b_2) ; consideriamo una loro combinazione lineare del tipo

$$k_1(a_1, b_1) + k_2(a_2, b_2)$$

Affinché questa combinazione lineare dia il vettore nullo, deve risultare

$$k_1(a_1, b_1) + k_2(a_2, b_2) = (0, 0)$$

ossia deve risultare

$$\begin{cases} k_1 a_1 + k_2 a_2 = 0 \\ k_1 b_1 + k_2 b_2 = 0 \end{cases}$$

Questo è un sistema, lineare omogeneo nelle incognite k_1 e k_2 , che non ammette soluzioni eccetto quella banale, dal che deduciamo che i 2 vettori sono senz'altro linearmente indipendenti. Deduciamo, quindi, che in \mathfrak{R}^2 è possibile trovare almeno 2 vettori linearmente indipendenti. Vediamo allora se è possibile trovarne 3. Consideriamo perciò 3 diversi vettori di \mathfrak{R}^2 : affinché siano linearmente indipendenti, non devono esistere 3 coefficienti (k_1, k_2, k_3) , ovviamente diversi da 0, tali da soddisfare la relazione

$$k_1(a_1, b_1) + k_2(a_2, b_2) + k_3(a_3, b_3) = (0, 0, 0)$$

Questa relazione equivale al sistema

$$\begin{cases} k_1 a_1 + k_2 a_2 + k_3 a_3 = 0 \\ k_1 b_1 + k_2 b_2 + k_3 b_3 = 0 \end{cases}$$

Questo sistema, lineare omogeneo nelle incognite k_1, k_2 e k_3 , ammette infinite soluzioni, dal che si deduce che non è possibile trovare 3 vettori linearmente indipendenti in \mathfrak{R}^2 . Possiamo dunque concludere che il numero massimo di vettori linearmente indipendenti che possiamo trovare in \mathfrak{R}^2 è 2.

Si intuisce, dunque, il secondo principio fondamentale: *dato lo spazio vettoriale \hat{A}^n , in esso è possibile trovare al più n vettori linearmente indipendenti.*

Possiamo a questo punto dare la seguente definizione:

Def. Si definisce "**dimensione**" di uno spazio vettoriale il numero massimo di vettori linearmente indipendenti che si possono trovare in tale spazio

In base a quanto appurato prima, la dimensione dello spazio vettoriale \mathfrak{R}^n è proprio n . In modo abbastanza analogo, si può verificare che la dimensione dello spazio vettoriale \mathbb{C}^n (insieme delle n -uple di numeri complessi) è $2n$, oppure che la dimensione dello spazio vettoriale delle funzioni sinusoidali, di frequenza assegnata, è 2 , oppure anche che la dimensione dello spazio vettoriale dei polinomi di grado $\leq n$ è $n+1$.

Questi sono tutti casi in cui la dimensione dello spazio vettoriale è un numero finito e si parla perciò di “**spazi vettoriali a dimensione finita**”; al contrario, ci sono anche spazi vettoriali la cui dimensione è ∞ e si parla ovviamente di “**spazi vettoriali a dimensione infinita**”. Un esempio ne è lo spazio vettoriale costituito dall’insieme delle funzioni, reali di variabile reale, definite in un arbitrario intervallo $[a,b]$: è chiaro che non c’è limite al numero di elementi, appartenenti a tale insieme, che noi possiamo combinare senza ottenere l’elemento nullo, per cui la dimensione è ∞ .

NORMA IN UNO SPAZIO VETTORIALE

Consideriamo adesso un generico spazio vettoriale S . Si definisce “**norma**” in tale spazio vettoriale una funzione avente le seguenti caratteristiche:

- in primo luogo, essa è definita nel modo seguente:

$$\|\bullet\| : S \longrightarrow \mathfrak{R}^+$$

Si tratta cioè di una funzione che ha come dominio lo spazio vettoriale in questione e come codominio l’insieme dei numeri reali positivi incluso lo 0;

- in secondo luogo, essa gode di tre proprietà fondamentali:
 - la prima proprietà dice che, preso un vettore $v \in S$, se risulta $\|v\|=0$, deve necessariamente essere $v=0$; in altre parole, *l’unico caso in cui la norma fornisce valore nullo è quello in cui il vettore cui è applicata coincide con il vettore nullo*;
 - la seconda proprietà prende il nome di “**disuguaglianza triangolare**” e afferma che, presi due qualsiasi vettori $\begin{matrix} v_1 \\ v_2 \end{matrix} \in S$, risulta

$$\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$$

- la terza ed ultima proprietà riguarda il “**prodotto per uno scalare**” e dice che, preso un vettore $v \in S$ ed uno scalare $k \in \mathfrak{R}$, risulta

$$\|kv\| \leq |k| \cdot \|v\|$$

Norme definibili in \hat{A}^n

Dato lo spazio vettoriale S , esistono molteplici funzioni che soddisfano a queste condizioni e rientrano perciò nella classe delle norme definite in tale spazio. Per esempio, consideriamo lo spazio vettoriale \mathfrak{R}^n ; considerato un generico elemento $x = (x_1, x_2, \dots, x_n)$ di tale spazio, è possibile definire le seguenti 3 norme:

- **norma 1:** $\|x\|_1 = \sum_{i=1}^n |v_i|$
- **norma 2** (o "norma Euclidea"): $\|x\|_2 = \sqrt{\sum_{i=1}^n v_i^2}$
- **norma ∞ :** $\|x\|_\infty = \max_{i=1, \dots, n} |v_i|$

E' facile verificare che tutte e tre queste norme rispondono ai requisiti prima enunciati.

Esempio

Consideriamo l'insieme delle n-ples di funzioni reali, definite su un sottoinsieme T di \mathfrak{R} , limitate:

$$S = \left\{ (f_1, f_2, \dots, f_n) \left| \begin{array}{l} f_k : T \longrightarrow \mathfrak{R} \\ |f_k(\bullet)| \leq M \\ k = 1, \dots, n \end{array} \right. \right\}$$

Questo insieme è uno spazio vettoriale di dimensione ∞ . Vogliamo definire una norma per questo spazio vettoriale.

Consideriamo, a questo scopo, un generico elemento $f(\bullet) = (f_1(\bullet), f_2(\bullet), \dots, f_n(\bullet))$ appartenente ad S ; preso un generico valore $t \in T$, consideriamo $f(t) = (f_1(t), f_2(t), \dots, f_n(t))$: con questa simbologia, $f(t)$ rappresenta una n-ple di numeri reali se t viene fissato, mentre rappresenta una funzione di t , ad n componenti, se t rimane generico.

Definiamo allora la seguente funzione:

$$\|f(t)\|_1 = \sum_{k=1}^n f_k(t)$$

Questa è una funzione di t che prende il nome di "funzione norma". Sulla base di questa funzione, possiamo definire la seguente norma su S :

$$\|f(\bullet)\| = \sup_{t \in T} \|f(t)\|_1$$

Ovviamente, questa è solo una possibile norma che noi possiamo definire su S usando la funzione $\|f(t)\|_1$ e comunque si possono fissare altri tipi di norme su S senza utilizzare tale funzione.

Osservazione

Supponiamo di avere 2 distinti spazi vettoriali V e W e supponiamo che su ciascuno di essi sia stata definita una certa norma: indichiamo rispettivamente con $\|\bullet\|_V$ e $\|\bullet\|_W$ tali norme. Consideriamo successivamente una funzione fatta nel modo seguente:

$$\begin{aligned} g &: V \longrightarrow W \\ v \in V &\longrightarrow w = g(v) \end{aligned}$$

Questa funzione trasforma, dunque, in un certo modo, i vettori di V in vettori di W . Per questa funzione è possibile dare la seguente definizione: *dato $v_0 \in V$, si dice che g è "continua in v_0 " se e solo se è verificata la condizione*

$$\lim_{\|v-v_0\|_V} \|g(v) - g(v_0)\|_W = 0$$

E' chiaro che il verificarsi o meno di questa condizione è strettamente legato a come sono state definite le due norme $\|\bullet\|_V$ e $\|\bullet\|_W$: per esempio, è possibile che quella condizione valga solo per due specifiche norme oppure anche che valga per qualsiasi definizione di tali norme. Sussiste allora un importante risultato:

Teorema - *Se V e W sono spazi vettoriali di dimensione finita, la scelta della norma non ha alcuna influenza sulla proprietà di continuità*

In altre parole, se V e W sono spazi vettoriali di dimensione finita, la proprietà $\lim_{\|v-v_0\|_V} \|g(v) - g(v_0)\|_W = 0$ o è valida per qualsiasi norma oppure non è valida per alcuna norma.

Se, invece, V e W sono spazi vettoriali di dimensione infinita, allora subentra la necessità di specificare il tipo di norme considerate.

BASE DI UNO SPAZIO VETTORIALE

Consideriamo sempre un generico spazio vettoriale V e supponiamo che esso abbia dimensione n : in base a quanto detto in precedenza, questo significa che noi siamo in grado di trovare n (e non più di n) vettori linearmente indipendenti di questo spazio vettoriale; l'insieme di questi vettori costituisce una "base" dello spazio vettoriale. La definizione generale di "base" di uno spazio vettoriale è dunque la seguente:

Def. *Si definisce "base" di uno spazio vettoriale V un insieme di vettori, appartenenti a V , linearmente indipendenti ed in numero pari alla dimensione di V*

E' importante sottolineare che i vettori presi, oltre ad essere linearmente indipendenti, devono essere in numero pari alla dimensione dello spazio vettoriale in esame: in altre parole, se lo spazio vettoriale ha dimensione n , un insieme di $m < n$ vettori, anche se sono linearmente indipendenti, non costituisce una base di questo spazio vettoriale.

Il concetto di base di uno spazio vettoriale è importante in quanto sussiste il seguente risultato: *dato uno spazio vettoriale V ed una sua qualsiasi base, ogni vettore $v \in V$ è esprimibile mediante un'unica combinazione lineare dei vettori che compongono la base.*

Scritto in formule, questo risultato dice quanto segue: fissata una base $B = \{x_1, x_2, \dots, x_n\}$ di V , risulta

$$\forall v \in V: \exists (k_1, k_2, \dots, k_n) \in \mathcal{R}^n \text{ tale che } v = k_1 x_1 + k_2 x_2 + \dots + k_n x_n$$

In base a questa definizione, i coefficienti k_1, k_2, \dots, k_n che forniscono la combinazione lineare prendono il nome di “**coordinate del vettore $v \in V$ rispetto alla base fissata**”. È importante sottolineare, in base sempre al risultato di prima, che *una volta fissata la base, per ciascun vettore v la n -pla di coordinate è unica; viceversa, cambiando la base, la n -pla può cambiare anch'essa.*

In altre parole, possiamo affermare che, una volta fissata la base, esiste una corrispondenza biunivoca tra il vettore $v \in V$ prescelto e la n -pla di coefficienti k_1, k_2, \dots, k_n che lo definiscono rispetto a quella base.

Cambiamento di base

È chiaro che uno spazio vettoriale possiede ∞ basi e quindi ciascun vettore di tale spazio possiede ∞ diverse rappresentazioni, una per ogni base. Viene allora da chiedersi se esista o meno un legame tra le coordinate di uno stesso vettore prese rispetto a basi diverse. Effettivamente questo legame esiste e vediamo perciò di individuarlo.

Fissato il generico vettore $v \in V$ e la generica base $B_X = \{x_1, x_2, \dots, x_n\}$ di V , supponiamo che sia (k_1, k_2, \dots, k_n) la n -pla di coordinate che definiscono v rispetto a tale base:

$$v = k_1 x_1 + k_2 x_2 + \dots + k_n x_n$$

In forma più compatta, possiamo usare la seguente notazione:

$$[v]_{B_X} = \begin{bmatrix} k_1 \\ k_2 \\ \dots \\ k_n \end{bmatrix}$$

Adesso supponiamo che $B_Y = \{y_1, y_2, \dots, y_n\}$ sia un'altra base dello spazio vettoriale V : supponiamo allora che sia (c_1, c_2, \dots, c_n) la n -pla di coordinate che definiscono v rispetto a tale base, per cui possiamo scrivere, in modo analogo a prima, che

$$[v]_{B_Y} = \begin{bmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{bmatrix}$$

La base B_Y è composta da n vettori appartenenti a V e, quindi, ciascuno di questi vettori si potrà esprimere mediante una (unica) combinazione lineare degli elementi della base B_X : avremo perciò

$$[y_1]_{B_X} = \begin{bmatrix} \xi_{11} \\ \xi_{21} \\ \dots \\ \xi_{n1} \end{bmatrix} \quad [y_2]_{B_X} = \begin{bmatrix} \xi_{12} \\ \xi_{22} \\ \dots \\ \xi_{n2} \end{bmatrix} \quad \dots \quad [y_n]_{B_X} = \begin{bmatrix} \xi_{1n} \\ \xi_{2n} \\ \dots \\ \xi_{nn} \end{bmatrix}$$

Considerando allora che $v = c_1 y_1 + c_2 y_2 + \dots + c_n y_n$, possiamo sostituire le relazioni appena trovate e scrivere che

$$v = c_1 (\xi_{11} x_1 + \xi_{21} x_2 + \dots + \xi_{n1} x_n) + c_2 (\xi_{12} x_1 + \xi_{22} x_2 + \dots + \xi_{n2} x_n) + \dots + c_n (\xi_{1n} x_1 + \xi_{2n} x_2 + \dots + \xi_{nn} x_n)$$

In questa relazione, abbiamo nuovamente il vettore v espresso in funzione dei vettori della base B_X ; spostando il modo opportuno i vari termini, troviamo quanto segue:

$$v = (\xi_{11} c_1 + \xi_{12} c_2 + \dots + \xi_{1n} c_n) x_1 + (\xi_{21} c_1 + \xi_{22} c_2 + \dots + \xi_{2n} c_n) x_2 + \dots + (\xi_{n1} c_1 + \xi_{n2} c_2 + \dots + \xi_{nn} c_n) x_n$$

D'altra parte, avevamo detto che l'espressione di v , rispetto alla base B_X , era $v = k_1 x_1 + k_2 x_2 + \dots + k_n x_n$; dovendo essere unica la rappresentazione di v rispetto alla base B_X , è necessario che i coefficienti delle due combinazioni trovate siano uguali, ossia è necessario che sia verificato il seguente sistema:

$$\begin{cases} (\xi_{11} c_1 + \xi_{12} c_2 + \dots + \xi_{1n} c_n) = k_1 \\ (\xi_{21} c_1 + \xi_{22} c_2 + \dots + \xi_{2n} c_n) = k_2 \\ \dots \\ (\xi_{n1} c_1 + \xi_{n2} c_2 + \dots + \xi_{nn} c_n) = k_n \end{cases}$$

Scritto in forma matriciale, questo sistema è fatto nel modo seguente:

$$\begin{bmatrix} \xi_{11} & \xi_{12} & \dots & \xi_{1n} \\ \xi_{21} & \xi_{22} & \dots & \xi_{2n} \\ \dots & \dots & \dots & \dots \\ \xi_{n1} & \xi_{n2} & \dots & \xi_{nn} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{bmatrix} = \begin{bmatrix} k_1 \\ k_2 \\ \dots \\ k_n \end{bmatrix}$$

Si tratta, dunque, di un sistema di n equazioni nelle incognite (c_1, c_2, \dots, c_n) , che sono le coordinate del vettore v rispetto alla base B_Y . Si tratta inoltre di un sistema compatibile che ammette 1 ed 1 sola soluzione: infatti, la matrice dei coefficienti ha come colonne i vettori che costituiscono B_Y e quindi si tratta di colonne linearmente indipendenti, il che garantisce che tale matrice non sia singolare (cioè abbia determinante non nullo).

Indicata allora con T la matrice dei coefficienti, possiamo concludere che

$$\boxed{[v]_{B_Y} = T^{-1}[v]_{B_X}}$$

Questa equazione matriciale, che corrisponde a n equazioni scalari, consente dunque di passare dalle coordinate di v rispetto alla base B_X alle coordinate di v rispetto a B_Y. Ovviamente, questo valore per qualsiasi vettore v ∈ V e quali che siano le due basi scelte.

Base canonica di uno spazio vettoriale

Consideriamo, come spazio vettoriale a dimensione finita, lo spazio ℝⁿ costituito dalle n-ple di numeri reali. Sappiamo che questo spazio ha dimensione n, per cui una sua base sarà costituita da n vettori linearmente indipendenti. E' facile verificare che una base di ℝⁿ è la seguente:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ \dots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \\ 1 \end{bmatrix}$$

Si tratta cioè di n vettori aventi tutti gli elementi nulli eccetto uno, di valore pari ad 1. Usando questi vettori come colonne di una matrice, si ottiene la matrice seguente:

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

Questa è la matrice identità: sapendo che il determinante di questa matrice è pari ad 1, abbiamo la conferma che quei vettori formano una base. La particolare natura di questi vettori conferisce a questa base il nome di "**base canonica**" o "**base naturale**".

Algebra matriciale

RANGO DI UNA MATRICE

Consideriamo un generica matrice A di dimensione m*n: si definisce "**rango**" di questa matrice il numero di righe (o di colonne) linearmente indipendenti presenti nella matrice.

E' subito chiaro che il rango di una matrice, solitamente indicato con la lettera greca ρ, è unico e, soprattutto, non può superare né numero di righe né il numero di colonne della matrice. Per esempio, supponiamo che A sia una matrice di dimensione 5*4: sicuramente, non potremo avere più di 5 righe linearmente indipendenti, per cui ρ(A) ≤ 5; d'altra parte, non potremo avere più di 4 colonne linearmente indipendenti, per cui sarà ρ ≤ 4.

In generale, se A è una matrice $n \times m$, possiamo certamente scrivere che

$$\rho(A) \leq \min(n, m)$$

Un altro modo di definire il rango di una matrice è il seguente:

Def. Si definisce "**rango**" di una matrice $n \times m$ l'ordine massimo dei minori non nulli individuabili nella matrice stessa, ossia l'ordine della sottomatrice (quadrata) più grande avente determinante diverso da 0

Questa definizione dice quanto segue: supponiamo di avere una matrice A di ordine $n \times m$ con $n < m$; ciò significa, in base a quanto prima, che $\rho(A) \leq n$; allora verifichiamo se, nella matrice, esiste almeno un minore (cioè una sottomatrice quadrata) di ordine n che abbia determinante $\neq 0$; se questo minore esiste, allora $\rho(A) = n$, altrimenti sarà $\rho(A) \leq n-1$; supponiamo allora che il minore ricercato non esista, per cui il massimo valore di $\rho(A)$ che possiamo trovare è $n-1$: se esiste un minore di ordine $n-1$ che abbia determinante $\neq 0$, allora $\rho(A) = n-1$, altrimenti sarà $\rho(A) \leq n-2$. Seguendo questo procedimento, si arriva a determinare $\rho(A)$.

In base a quanto detto, si capisce subito una cosa: l'unica possibilità per cui la matrice in esame abbia rango nullo è che tutti i suoi elementi sia nulli. Infatti, se c'è almeno 1 elemento non nullo nella matrice, esso potrà essere considerato come un minore, di ordine 1, con determinante non nullo, per cui il rango della matrice sarà al minimo 1.

Concludiamo ricordando che, se A è una matrice quadrata di ordine n e se il suo rango coincide con n (il che significa che A è non singolare), allora si dice che A è una matrice "**a rango pieno**".

Disuguaglianza di Sylvester

Supponiamo di avere due distinte matrici A e B , di dimensioni rispettivamente $m \times n$ ed $n \times p$; supponiamo anche che $r_A = \rho(A)$ e $r_B = \rho(B)$; le dimensioni delle due matrici indicano che possiamo effettuare il prodotto righe per colonne AB : indicata allora con C la matrice risultante, ci chiediamo quanto valga il suo rango. A questa domanda risponde, sia pure in modo parziale, la cosiddetta "**disuguaglianza di Sylvester**": indicato con r_C il rango di $C = AB$, tale disuguaglianza afferma che

$$r_A + r_B - n \leq r_C \leq \min(r_A, r_B)$$

Si osservi che n è la cosiddetta "dimensione comune" delle due matrici, ossia il numero di colonne della prima o il numero di righe della seconda.

In pratica, quindi, la disuguaglianza di Sylvester indica il valore minimo ed il valore massimo tra i quali può variare r_C : spesso tali valori coincidono, per cui anche r_C coincide con essi, ed è in questi casi che si manifesta la principale utilità di tale disuguaglianza.

Un caso particolare di applicazione della disuguaglianza di Sylvester è il seguente: sia A una matrice quadrata di ordine n ; supponiamo che si tratti di una matrice a rango pieno, il che significa che $\rho(A) = n$; consideriamo la matrice A^2 , ottenuta facendo il prodotto righe per colonne di A per se stessa: applicando la disuguaglianza di Sylvester per A^2 , abbiamo che

$$n + n - n \leq \rho(A^2) \leq n$$

dal che si deduce che $\rho(A^2) = n$, ossia che anche A^2 è una matrice a rango pieno. E' intuitivo capire che questa proprietà vale quale che sia l'esponente di A , per cui possiamo concludere, in generale, che, se A è una matrice quadrata di ordine n , a rango pieno, anche la matrice A^k , per " $k \in \mathbb{N}$ ", è una matrice a rango pieno.

RANGE DI UNA MATRICE

Supponiamo di avere due spazi vettoriali, che indichiamo con V e W ; prendiamo due vettori $x \in V$ e $y \in W$ e supponiamo che esista una matrice A , di dimensioni opportune, tale che $y = Ax$. In pratica, quindi, la matrice A serve a "trasformare" vettori dello spazio vettoriale V in vettori dello spazio vettoriale W . E' chiaro che, fissata la matrice A , non è detto che ciascun vettore $x \in V$ sia trasformabile in un vettore $y \in W$, ossia non è detto che, preso un qualsiasi $x \in V$, il corrispondente vettore $y = Ax$ appartenga a W ; viceversa, non è detto che, fissato un qualsiasi $y \in W$ e fissata la matrice A , esista un vettore $x \in V$ tale $y = Ax$. Possiamo allora dare la seguente definizione:

Def. Si definisce "**range della matrice A** " l'insieme di tutti i vettori $y \in W$ tali che esista almeno un vettore $x \in V$ che soddisfi alla relazione $y = Ax$

In formule possiamo dunque scrivere che

$$\text{Range}(A) = R(A) = \{y \in W \mid \exists x \in V \text{ tale che } y = Ax\}$$

Possiamo subito fare una serie di considerazioni su questa definizione:

- la prima considerazione è che il range di A è un sottoinsieme di W , visto che contiene solo vettori di W ; è possibile che $R(A) = W$, il che significa che tutti i vettori di W hanno un corrispondente vettore $x \in V$ rispetto alla matrice A , ma, più in generale, risulta che $R(A) \subset W$, ossia risulta che solo una parte dei vettori di W godono di questa proprietà;
- la seconda considerazione è che $R(A)$ è un sottospazio vettoriale, per il quale è quindi possibile definire la "dimensione" (ossia il massimo numero di vettori linearmente indipendenti individuabili in tale sottospazio); si verifica allora che $\dim[R(A)] = \rho(A)$, ossia che la dimensione di $R(A)$ coincide con il rango della matrice A .

AUTOVALORI E AUTOVETTORI DI UNA MATRICE

Supponiamo di avere una matrice A quadrata di ordine n ed uno spazio vettoriale V , di dimensione pari anch'essa ad n ; è possibile che esista un vettore $x \in V$ tale che il nuovo vettore $y = Ax$ appartenga a sua volta a V . In particolare, è possibile che, una volta fissata la matrice A , sia possibile trovare uno o più vettori $x \in V$ tali che

$$Ax = \lambda x$$

dove λ è uno scalare fissato. Se esistono uno scalare λ e uno o più vettori $x \in V$ che soddisfano quella proprietà, si dice che λ è un “**autovalore di A**” e che i suddetti vettori sono “**autovettori di A associati a λ** ”.

Supponiamo dunque che λ sia un autovalore della matrice A : abbiamo detto che possono esistere uno o più autovettori associati a λ , ossia uno o più vettori $x \in V$ che soddisfano la condizione $Ax = \lambda x$. Questi autovettori definiscono perciò un sottospazio di V e a questo sottospazio si dà il nome di “**autospazio associato a λ** ”. La dimensione di tale autospazio è pari al numero di autovettori distinti associati a λ .

Ci si può porre, a questo punto, il seguente problema: come si fa’ a determinare tutti gli autovalori di una matrice A e tutti gli autovettori ad essi associati? Basta fare il seguente ragionamento: la relazione $Ax = \lambda x$ può anche essere scritta nella forma

$$(A - \lambda I)x = 0$$

Questa è una equazione matriciale omogenea rappresentativa di un sistema di n equazioni scalari omogenee. Fissato λ , l’unica possibilità perché questo sistema ammetta almeno 1 soluzione (cioè un autovettore) diversa da quella banale (cioè il vettore nullo) è che il determinante della matrice dei coefficienti $A - \lambda I$ sia $= 0$. Dato che A ed I sono entrambe matrici quadrate di ordine n , anche la matrice $A - \lambda I$ è quadrata di ordine n : di conseguenza, chiedere che $\det(A - \lambda I) = 0$ equivale a chiedere che la matrice $A - \lambda I$ abbia rango minore di n .

Possiamo dunque affermare che se λ è un autovalore di A , deve risultare $\det(A - \lambda I) = 0$ affinché esista almeno un autovettore associato a λ . Ma vale allora anche il viceversa, nel senso che possiamo affermare che, *se $\det(A - \lambda I) = 0$, allora λ è un autovalore di A* .

Abbiamo dunque trovato il modo di individuare gli autovalori della matrice A : basta trovare tutti i valori λ che soddisfano la relazione $\det(A - \lambda I) = 0$. L’insieme di tutti gli autovalori della matrice A prende il nome di “**spettro**” della matrice stessa.

L’altro problema è quello di determinare gli autovettori associati al generico autovalore λ . Ma la soluzione è stata indicata prima: basta infatti risolvere il sistema

$$(A - \lambda I)x = 0$$

Essendo λ un autovalore, la matrice dei coefficienti di questo sistema è singolare, il che significa che il sistema è compatibile e ammette ∞ soluzioni diverse da quella banale. Queste soluzioni corrispondono dunque a ∞ autovettori che noi possiamo associare all’autovalore λ : è chiaro, però, che a noi non interessano tutti questi ∞ autovettori, ma solo quelli linearmente indipendenti. Il numero di vettori linearmente indipendenti che possiamo tirar fuori da quel sistema prende il nome di “**molteplicità geometrica**” dell’autovalore λ , in quanto rappresenta il numero di autovettori linearmente indipendenti associati a λ stesso.

N.B. E’ facile verificare che questa molteplicità geometrica è pari alla cosiddetta “**nullità**” della matrice $A - \lambda I$, ossia alla differenza $n - \rho(A)$ tra l’ordine della matrice ed il suo rango.

Tornando agli autovalori, abbiamo detto che, fissata A quadrata di ordine n , tutti e soli i suoi autovalori si ottengono imponendo la condizione $\det(A - \lambda I) = 0$. E' intuitivo rendersi conto che la quantità $p(\lambda) = \det(A - \lambda I)$ non è altro che un polinomio nell'incognita λ : esso prende il nome di "**polinomio caratteristico della matrice A** ". Essendo $A - \lambda I$ una matrice quadrata di ordine n , è chiaro che $p(\lambda)$ sarà un polinomio, in λ , di grado al più n , per cui lo possiamo scrivere nella forma seguente:

$$p(\lambda) = \lambda^n + a_1 \lambda^{n-1} + a_2 \lambda^{n-2} + \dots + a_{n-2} \lambda^2 + a_{n-1} \lambda + a_n$$

Se $p(\lambda)$ è di grado n , è chiaro che l'equazione $p(\lambda) = \det(A - \lambda I) = 0$ (detta "**equazione caratteristica della matrice A** ") ammette n soluzioni, ossia n autovalori della matrice A : possiamo perciò affermare che, *data A matrice quadrata di ordine n , essa possiede n autovalori.*

Come sono fatti questi autovalori? Trattandosi delle radici di una equazione di n° grado, le situazioni possibili sono le seguenti:

- il caso più semplice è quello in cui l'equazione caratteristica ammette n soluzioni reali e distinte, il che significa che abbiamo n autovalori reali e distinti;
- l'altra possibilità è che ci siano due o più coppie di autovalori reali e coincidenti, il che significa che avremo meno di n autovalori reali e distinti;
- l'ultima possibilità è che ci siano, oltre ad autovalori reali (distinti o coincidenti), anche coppie di autovalori complessi coniugati.

Possiamo allora definire il concetto di "*molteplicità algebrica*" del generico autovalore λ : l'autovalore λ , in quanto soluzione dell'equazione caratteristica, ha un suo "ordine", nel senso che si può trattare di una radice di ordine 1 (cioè una radice "*semplice*") oppure di una radice di ordine 2 (cioè una radice "*doppia*") e così via a seconda del numero di derivate di $p(\lambda)$ che è in grado di annullare. Allora, l'ordine di λ rappresenta anche la sua "**molteplicità algebrica**": potremo perciò avere autovalori semplici, doppi, tripli e così via.

Esiste una ovvia relazione tra la molteplicità geometrica e la molteplicità algebrica di un autovalore λ_i : indicate infatti, rispettivamente, con g_i e m_i tali molteplicità, risulta che

$$g_i \leq m_i$$

In termini concreti, questa relazione dice che *ad autovalori distinti sono sempre associati autovettori linearmente indipendenti (e quindi distinti).*

Dimostrazione

Supponiamo che la matrice A in esame sia quadrata di ordine n ; siano $\lambda_1, \lambda_2, \dots, \lambda_m$, con $m \leq n$, gli autovalori distinti di A e siano $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_m$ i corrispondenti autovettori. Per dimostrare che questi vettori sono linearmente indipendenti dobbiamo far vedere che l'unica loro combinazione lineare che dà il vettore nullo è quella ottenuta con coefficienti tutti nulli: in formule, dobbiamo cioè far vedere che, se è verificata la relazione

$$\sum_{k=1}^m \alpha_k \bar{x}_k = \vec{0}$$

essa implica necessariamente che i coefficienti α_k siano tutti nulli.

Per prima cosa, se vale quella relazione, allora abbiamo anche che

$$A \sum_{k=1}^m \alpha_k \vec{x}_k = \sum_{k=1}^m A \alpha_k \vec{x}_k = \vec{0}$$

D'altra parte, la matrice A e il generico autovettore \vec{x}_k sono legati dalla relazione $A\vec{x}_k = \lambda_k \vec{x}_k$

$$\sum_{k=1}^m \alpha_k \lambda_k \vec{x}_k = \vec{0}$$

In modo del tutto analogo, abbiamo anche che

$$A^2 \sum_{k=1}^m \alpha_k \vec{x}_k = \sum_{k=1}^m A^2 \alpha_k \vec{x}_k = \sum_{k=1}^m \alpha_k \lambda_k^2 \vec{x}_k = \vec{0}$$

e così via, fino a

$$A^{m-1} \sum_{k=1}^m \alpha_k \vec{x}_k = \sum_{k=1}^m A^{m-1} \alpha_k \vec{x}_k = \sum_{k=1}^m \alpha_k \lambda_k^{m-1} \vec{x}_k = \vec{0}$$

Mettendo insieme le equazioni trovate, abbiamo quanto segue:

$$\begin{cases} \sum_{k=1}^m \alpha_k \lambda_k \vec{x}_k = \alpha_1 \lambda_1 \vec{x}_1 + \alpha_2 \lambda_2 \vec{x}_2 + \dots + \alpha_m \lambda_m \vec{x}_m = \vec{0} \\ \sum_{k=1}^m \alpha_k \lambda_k^2 \vec{x}_k = \alpha_1 \lambda_1^2 \vec{x}_1 + \alpha_2 \lambda_2^2 \vec{x}_2 + \dots + \alpha_m \lambda_m^2 \vec{x}_m = \vec{0} \\ \dots \\ \sum_{k=1}^m \alpha_k \lambda_k^{m-1} \vec{x}_k = \alpha_1 \lambda_1^{m-1} \vec{x}_1 + \alpha_2 \lambda_2^{m-1} \vec{x}_2 + \dots + \alpha_m \lambda_m^{m-1} \vec{x}_m = \vec{0} \end{cases}$$

Tutte queste equazioni possono essere poste in forma matriciale nel modo seguente:

$$[\alpha_1 \vec{x}_1 \mid \alpha_2 \vec{x}_2 \mid \alpha_3 \vec{x}_3 \mid \dots \mid \alpha_m \vec{x}_m] \underbrace{\begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{m-1} \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{m-1} \\ 1 & \lambda_3 & \lambda_3^2 & \dots & \lambda_3^{m-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & \lambda_m & \lambda_m^2 & \dots & \lambda_m^{m-1} \end{bmatrix}}_V = \vec{0}$$

A questo punto, è possibile dimostrare che la matrice indicata con V (che prende il nome di "matrice di Vandermonde", è non singolare in conseguenza del fatto che gli autovalori

$\lambda_1, \lambda_2, \dots, \lambda_m$ sono per ipotesi tutti distinti. Possiamo allora moltiplicare a destra per V^{-1} , ottenendo chiaramente

$$[\alpha_1 \bar{x}_1 \mid \alpha_2 \bar{x}_2 \mid \alpha_3 \bar{x}_3 \mid \dots \mid \alpha_m \bar{x}_m] = \vec{0}$$

Questa relazione dice che

$$\begin{cases} \alpha_1 \bar{x}_1 = \vec{0} \\ \alpha_2 \bar{x}_2 = \vec{0} \\ \dots \\ \alpha_m \bar{x}_m = \vec{0} \end{cases}$$

Ma, per ipotesi, i vettori $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_m$, essendo degli autovettori, sono tutti diversi dal vettore nullo, per cui quelle m relazioni implicano necessariamente che $\alpha_k = 0$ per $\forall k=1, \dots, m$, come volevamo dimostrare.

PROPRIETÀ: AUTOVALORI E DETERMINANTE

Sia A una generica matrice quadrata di ordine n : se $\lambda_1, \lambda_2, \dots, \lambda_n$ sono gli n autovalori (non necessariamente distinti tra di loro) della matrice stessa, il suo polinomio caratteristico si può scrivere nella forma

$$p(\lambda) = \det[A - \lambda I] = (\lambda - \lambda_1)(\lambda - \lambda_2) + \dots + (\lambda - \lambda_n)$$

Se, in quella espressione, poniamo $\lambda=0$, otteniamo

$$p(0) = \det A = (-1)^n \lambda_1 \lambda_2 \dots \lambda_n$$

Abbiamo cioè trovato che *il prodotto degli autovalori di una matrice A è pari al determinante della matrice stessa, a meno di un termine -1 nel caso in cui n sia dispari.*

D'altra parte, il polinomio caratteristico è anche esprimibile nella forma

$$p(\lambda) = \lambda^n + a_1 \lambda^{n-1} + a_2 \lambda^{n-2} + \dots + a_{n-2} \lambda^2 + a_{n-1} \lambda + a_n$$

Se poniamo qui $\lambda=0$, otteniamo evidentemente $p(0) = a_n$, dal che deduciamo, mettendo insieme con quanto visto prima, che

$$p(0) = \det A = a_n = (-1)^n \lambda_1 \lambda_2 \dots \lambda_n$$

Sempre da questa proprietà si deduce immediatamente che *la matrice A è una matrice singolare (cioè con determinante nullo) se e solo se almeno uno dei suoi autovalori è $=0$.*

Ancora, se andassimo a sviluppare i prodotti che compaiono nella formula

$$p(\lambda) = \det[A - \lambda I] = (\lambda - \lambda_1)(\lambda - \lambda_2) + \dots + (\lambda - \lambda_n)$$

potremmo ricavare l'espressione di tutti i coefficienti a_1, a_2, \dots, a_n del polinomio caratteristico: tali espressioni sono

$$\begin{aligned} a_1 &= -\sum_{k=1}^n \lambda_k \\ a_2 &= \sum_{i=1}^n \sum_{k=1}^n \lambda_k \lambda_i \\ &\dots \\ a_n &= (-1)^n \lambda_1 \lambda_2 \dots \lambda_n \end{aligned}$$

Infine, rifacendosi alle proprietà enunciate in questo e nei paragrafi precedenti, è possibile ricavare le seguenti relazioni:

$$\begin{aligned} a_1 &= -(\text{traccia}F) \\ a_2 &= -\frac{1}{2}(a_1 \text{traccia}F + \text{traccia}F^2) \\ a_3 &= -\frac{1}{3}(a_2 \text{traccia}F + a_1 \text{traccia}F^2 + \text{traccia}F^3) \\ &\dots \\ a_n &= -\frac{1}{n}(a_{n-1} \text{traccia}F + a_{n-2} \text{traccia}F^2 + \dots + \text{traccia}F^n) \end{aligned}$$

dove ricordiamo che la “traccia” di una matrice è la somma degli elementi situati sulla diagonale principale della matrice stessa.

Queste relazioni consentono dunque un metodo di calcolo relativamente rapido dei coefficienti di $p(\lambda)$. Oltre a questo, è di particolare importanza la prima relazione: infatti, se $a_1 = -(\text{traccia}F)$ e

avendo in precedenza detto che $a_1 = -\sum_{k=1}^n \lambda_k$, deduciamo che

$$\boxed{\text{traccia}F = \sum_{k=1}^n \lambda_k}$$

ossia che la somma degli elementi diagonali di una matrice (ovviamente quadrata) è pari alla somma degli autovalori della matrice stessa, anche nel caso in cui la matrice non sia triangolare (nel qual caso gli elementi diagonali coincidono con gli autovalori).

DIAGONALIZZAZIONE DI UNA MATRICE

Sia data ancora una matrice A quadrata di ordine n ; possiamo fare, in generale, l'ipotesi che $A \in \mathbb{C}^{n \times n}$, il che significa che gli elementi di questa matrice possono anche essere di tipo complesso. Supponiamo inoltre che la matrice abbia n autovalori distinti $\lambda_1, \lambda_2, \dots, \lambda_n$ e che $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n$ siano i corrispondenti autovettori. Questi autovettori, in base al teorema prima dimostrato, sono senz'altro linearmente indipendenti: allora, se noi usiamo questi autovettori come colonne per costruire la matrice

$$T = [\bar{x}_1 \mid \bar{x}_2 \mid \bar{x}_3 \mid \dots \mid \bar{x}_n]$$

siamo certi che si tratta di una matrice non singolare. Moltiplichiamo allora a sinistra questa matrice per A : otteniamo evidentemente che

$$AT = A[\bar{x}_1 \mid \bar{x}_2 \mid \bar{x}_3 \mid \dots \mid \bar{x}_n] = [A\bar{x}_1 \mid A\bar{x}_2 \mid A\bar{x}_3 \mid \dots \mid A\bar{x}_n]$$

Considerando che la matrice A e il generico autovettore \bar{x}_k sono legati dalla relazione $A\bar{x}_k = \lambda_k \bar{x}_k$, quella equivale anche a

$$AT = [\lambda_1 \bar{x}_1 \mid \lambda_2 \bar{x}_2 \mid \lambda_3 \bar{x}_3 \mid \dots \mid \lambda_n \bar{x}_n]$$

A questo punto, se consideriamo la matrice

$$J = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \lambda_n \end{bmatrix}$$

è intuitivo accorgersi che la matrice AT è esprimibile come

$$AT = TJ$$

D'altra parte, abbiamo detto che la matrice T è non singolare, per cui possiamo pre-moltiplicare ambo i membri di quella relazione per T^{-1} : otteniamo

$$\boxed{J = T^{-1}AT}$$

Per come è stata costruita, la matrice J è una "**matrice simile**" alla matrice A . La definizione completa è la seguente:

Def. Due assegnate matrici $A \in \mathbb{C}^{n \times n}$ e $B \in \mathbb{C}^{n \times n}$ si dicono "**simili**" se esiste una matrice non singolare $T \in \mathbb{C}^{n \times n}$ tale che $B = T^{-1}AT$.

Le matrici simili godono di alcune proprietà fondamentali:

- la prima è che due matrici simili hanno lo stesso determinante;
- la seconda è che la relazione di “similitudine” è una relazione di uguaglianza, il che significa che essa gode delle proprietà riflessiva, simmetrica e transitiva;
- la terza è che due matrici simili hanno gli stessi autovalori.

E' evidente che la matrice J trovata prima è una “particolare” matrice simile ad A : infatti, si tratta di una matrice diagonale i cui elementi diagonali coincidono con gli autovalori della matrice A . La possibilità di trovare, data $A \in \mathbb{C}^{n \times n}$, la corrispondente matrice diagonale J è garantita dal fatto che A possiede n autovettori linearmente indipendenti; è possibile infatti dimostrare il seguente risultato:

Teorema - Una matrice $A \in \mathbb{C}^{n \times n}$ può essere trasformata, per similitudine, in una matrice diagonale J se e solo se possiede n autovettori linearmente indipendenti

Abbiamo prima visto che, se la matrice A , quadrata di ordine n , possiede n autovalori distinti (si dice che si tratta di una matrice “**semplice**”), allora anche i corrispondenti n autovettori sono linearmente indipendenti e quindi, in base al teorema appena citato, è possibile senz'altro trovare la matrice J . Vediamo allora cosa succede quando gli autovalori di A non sono tutti distinti.

Abbiamo prima introdotto il concetto di “**molteplicità algebrica**” di un autovalore generico λ di A : dato il polinomio caratteristico $p(\lambda) = \det(A - \lambda I)$ della matrice A , se λ_k è una sua radice di ordine m_k , allora proprio m_k rappresenta la molteplicità algebrica di λ_k . E' chiaro, allora, che, se $\lambda_1, \lambda_2, \dots, \lambda_p$ sono gli autovalori distinti di A e m_1, m_2, \dots, m_p sono le rispettive molteplicità algebrica, il polinomio caratteristico è esprimibile nella forma

$$p(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_p)^{m_p}$$

Abbiamo anche introdotto il concetto di “**molteplicità geometrica**” di un autovalore generico λ_k di A : si tratta del numero massimo g_k di autovettori linearmente indipendenti associati all'autovalore λ_k , ossia del numero massimo di vettori \vec{x}_k linearmente indipendenti che soddisfano il sistema

$$(A - \lambda_k I)\vec{x}_k = 0$$

N.B. Facciamo osservare che, per come è stata appena definita, la molteplicità geometrica del generico λ_k corrisponde alla cosiddetta “**degenerazione**” della matrice $(A - \lambda_k I)$, ossia alla dimensione del **nucleo** di tale matrice, che ricordiamo essere uno spazio vettoriale definito come

$$\text{Nucleo}(A - \lambda_k I) = \{ \vec{x} \in \mathbb{R}^n \mid (A - \lambda_k I)\vec{x} = 0 \}$$

Quando la matrice A non ammette n autovalori distinti (dove n è la dimensione della matrice), il caso più semplice che si può presentare è quello in cui la molteplicità algebrica di ciascun autovalore coincide con quella geometrica (si parla, in questo caso, di matrice “**semi-semplice**”): in questo caso, risulta evidentemente

$$\sum_{k=1}^p m_k = \sum_{k=1}^p g_k = n$$

per cui possiamo comunque ottenere dagli autovalori un numero di autovettori indipendenti pari ad n da usare come colonne per costruire la matrice di trasformazione T .

Il caso più svantaggioso è invece quello in cui, dati gli autovalori distinti $\lambda_1, \lambda_2, \dots, \lambda_p$, con molteplicità algebriche m_1, m_2, \dots, m_p e molteplicità geometriche g_1, g_2, \dots, g_p , esiste almeno uno tra questi autovettori avente molteplicità geometrica strettamente minore di quella algebrica: in questo caso, siamo certi che

$$\sum_{k=1}^p g_k < \sum_{k=1}^p m_k = n$$

il che significa che non possiamo in alcun modo trovare n autovettori linearmente indipendenti con i quali costruire la matrice T che consenta poi di arrivare alla matrice J . Possiamo perciò concludere che, se la matrice A , quadrata di ordine n , avente $p (< n)$ autovalori distinti, possiede almeno un autovalore la cui molteplicità geometrica è strettamente minore della molteplicità algebrica, la matrice A non è diagonalizzabile.

In questo caso, la matrice A si dice "**difettosa**" ed è comunque possibile trasformare A , mediante opportune trasformazioni, in una matrice ad essa simile: questa matrice prende il nome di "**forma canonica di Jordan**" ed ha la caratteristica di essere "diagonale a blocchi" anziché semplicemente diagonale.

FORMA CANONICA DI JORDAN

In questo paragrafo faremo dunque vedere come si costruisce la forma canonica di Jordan di una data matrice.

Sia dunque A una matrice quadrata di ordine n (ad elementi, in generale, complessi). Siano $\lambda_1, \lambda_2, \dots, \lambda_p$ gli autovalori distinti di A , con molteplicità algebriche m_1, m_2, \dots, m_p e molteplicità geometriche g_1, g_2, \dots, g_p . Facciamo l'ipotesi che almeno uno tra questi autovettori abbia molteplicità geometrica strettamente minore di quella algebrica, il che, come detto prima, ci impedisce di diagonalizzare A in quanto non possiamo trovare n autovettori linearmente indipendenti. Consideriamo il generico autovalore λ_i : se g_i è la sua molteplicità geometrica, noi possiamo risolvere il sistema

$$(A - \lambda_i I) \vec{x}_{ij} = 0$$

in modo da trovare g_i vettori $\vec{x}_{i1}, \vec{x}_{i2}, \dots, \vec{x}_{ig_i}$ che corrispondono agli autovettori linearmente indipendenti associati a λ_i . Consideriamo l'autovettore \vec{x}_{i1} : ci chiediamo se esiste un vettore non nullo $\vec{t}_{i1}^{(1)}$ che soddisfa la relazione

$$(A - \lambda_i I) \vec{t}_{i1}^{(1)} = \vec{x}_{ij}$$

Una volta trovato questo vettore $\vec{t}_{i1}^{(1)}$, andiamo a cercare un altro vettore non nullo $\vec{t}_{i2}^{(1)}$ che soddisfa la relazione

$$(A - \lambda_i I)\vec{t}_{i2}^{(1)} = \vec{t}_{i1}^{(1)}$$

Trovato $\vec{t}_{i2}^{(1)}$, andiamo a cercare un altro vettore non nullo $\vec{t}_{i3}^{(1)}$ che soddisfa la relazione

$$(A - \lambda_i I)\vec{t}_{i3}^{(1)} = \vec{t}_{i2}^{(1)}$$

e così via, fino a trovare un vettore non nullo $\vec{t}_{ik}^{(1)}$ che soddisfa la relazione

$$(A - \lambda_i I)\vec{t}_{ik}^{(1)} = \vec{t}_{i(k-1)}^{(1)}$$

In questo modo, in corrispondenza dell'autovettore \vec{x}_{i1} , avremo trovato un certo numero n_1 di vettori $\vec{t}_{i1}^{(1)}, \vec{t}_{i2}^{(1)}, \dots, \vec{t}_{in_1}^{(1)}$.

Adesso ripetiamo il ragionamento in corrispondenza dell'autovettore \vec{x}_{i2} : troveremo in questo caso n_2 vettori $\vec{t}_{i1}^{(2)}, \vec{t}_{i2}^{(2)}, \dots, \vec{t}_{in_2}^{(2)}$ e così via fino all'autovettore \vec{x}_{ig_i} , in corrispondenza del quale troveremo n_{g_i} vettori $\vec{t}_{i1}^{(g_i)}, \vec{t}_{i2}^{(g_i)}, \dots, \vec{t}_{in_{g_i}}^{(g_i)}$.

Questi vettori

$$\vec{t}_{i1}^{(1)}, \vec{t}_{i2}^{(1)}, \dots, \vec{t}_{in_1}^{(1)}, \vec{t}_{i1}^{(2)}, \vec{t}_{i2}^{(2)}, \dots, \vec{t}_{in_2}^{(2)}, \dots, \vec{t}_{i1}^{(g_i)}, \vec{t}_{i2}^{(g_i)}, \dots, \vec{t}_{in_{g_i}}^{(g_i)}$$

prendono il nome di “**autovettori generalizzati**” associati all'autovalore λ_i . E' possibile dimostrare che questi autovettori generalizzati sono linearmente indipendenti dagli autovettori $\vec{x}_{i1}, \vec{x}_{i2}, \dots, \vec{x}_{ig_i}$ associati all'autovalore λ_i .

Trovando un numero opportuno di autovettori generalizzati associati a tutti i p autovalori distinti della matrice A , è facile verificare che l'insieme costituito da tali autovettori generalizzati e dagli autovettori propriamente detti è costituito da n vettori linearmente indipendenti: questo comporta che questi vettori si possano usare come colonne per costruire una matrice T che risulti non singolare. Utilizzando questa matrice, si può verificare (in modo neanche troppo difficile) che sussiste ancora la relazione

$$\boxed{A = T^{-1}JT}$$

dove la matrice J , che prende appunto il nome di “**forma canonica di Jordan**” equivalente ad A , è una matrice molto particolare: in primo luogo, si tratta di una *matrice diagonale a blocchi* del tipo

$$J = \begin{bmatrix} J_1 & 0 & 0 & \dots & 0 \\ 0 & J_2 & 0 & \dots & 0 \\ 0 & 0 & J_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & J_p \end{bmatrix}$$

Come si vede, il numero di blocchi J_i è pari al numero di autovalori distinti della matrice A . Inoltre, ciascun blocco J_i , che prende il nome di "**blocco di Jordan associato a λ_i** ", è a sua volta diagonale a blocchi del tipo

$$J_i = \begin{bmatrix} J_{i1} & 0 & 0 & \dots & 0 \\ 0 & J_{i2} & 0 & \dots & 0 \\ 0 & 0 & J_{i3} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & J_{ig_i} \end{bmatrix}$$

Il generico blocco J_{ij} associato all'autovalore λ_i prende il nome di "**miniblocco di Jordan**" associato a λ_i : il numero di miniblocchi di Jordan associati a λ_i è uguale alla molteplicità geometrica q_i di λ_i , ossia al numero di autovettori linearmente indipendenti associati a λ_i , ossia anche alla cosiddetta "*degenerazione*" della matrice $\lambda_i I - A$, che corrisponde alla dimensione del nucleo di tale matrice.

Infine, ciascun miniblocco di Jordan è fatto nel modo seguente:

$$J_{ij} = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ 0 & 0 & \lambda_i & \dots & \dots \\ \dots & \dots & \dots & \dots & 1 \\ 0 & 0 & \dots & 0 & \lambda_i \end{bmatrix}_{(n_j+1) \times (n_j+1)}$$

Si tratta cioè di una matrice quadrata, di dimensione pari a $n_j + 1$, avente l'autovalore λ_i sulla diagonale, la sopradiagonale costituita da elementi tutti unitari e tutti gli altri elementi nulli.

In generale, non è semplicissimo calcolare il valore della dimensione n_j del generico miniblocco J_{ij} . Sarà esaminato in seguito il criterio di calcolo di questi parametri. Per il momento, facciamo solo osservare che il numero degli 1 sulla sopradiagonale della matrice J è pari alla dimensione (n) della matrice diminuita della somma delle molteplicità geometriche dei vari autovalori.

E' facile verificare che J è una matrice quadrata di dimensione n : infatti

$$\begin{aligned} \dim J &= \sum_{i=1}^p \dim J_i = \sum_{i=1}^p \sum_{j=1}^{q_i} \dim J_{ij} = \sum_{i=1}^p \sum_{j=1}^{q_i} (n_j + 1) = \sum_{i=1}^p \left[\sum_{j=1}^{q_i} n_j + \sum_{j=1}^{q_i} 1 \right] = \\ &= \sum_{i=1}^p m_i = n \end{aligned}$$

Esempio

Consideriamo la matrice

$$A = \begin{bmatrix} 8 & -8 & -2 \\ 4 & -3 & -2 \\ 3 & -4 & 1 \end{bmatrix}$$

Vogliamo determinare la forma canonica di Jordan di questa matrice.

La prima cosa da fare è calcolare gli autovalori della matrice: si verifica facilmente che il polinomio caratteristico di A è

$$p(\lambda) = (\lambda - 1)(\lambda - 2)(\lambda - 3)$$

il che significa che gli autovalori di A sono $\lambda_1=1$, $\lambda_2=2$ e $\lambda_3=3$, tutti di molteplicità 1. Il fatto che la matrice ammetta tutti autovalori distinti ci dice che si tratta di una matrice diagonalizzabile, per cui la matrice J ha la forma seguente:

$$J = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Ad ogni autovalore è associato un solo blocco di Jordan e, poiché tutti gli autovalori hanno molteplicità geometrica pari ad 1, il numero di 1 sulla sopradiagonale pari a $3-(1+1+1)$, cioè nessuno.

Esempio

Consideriamo la matrice

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

Vogliamo determinare la matrice T che riduce A nella forma canonica di Jordan.

La prima cosa da fare è calcolare gli autovalori della matrice: si verifica facilmente che il polinomio caratteristico di A è $p(\lambda) = (\lambda - 2)^3$, il che significa che A ammette 1 solo autovalore ($\lambda=2$) di molteplicità algebrica 3. Per costruire la matrice T, abbiamo due possibilità: la prima è che $\lambda=2$ abbia molteplicità geometrica anch'essa pari a 3, nel qual caso potremo trovare 3 autovettori linearmente indipendenti e quindi otterremo la matrice J in forma diagonale; la seconda è, invece, che la molteplicità geometrica di $\lambda=2$ risulti minore di quella algebrica, nel qual caso dovremo trovare uno o due autovettori generalizzati associati a λ .

Gli autovettori associato a $\lambda=2$ si ottengono risolvendo il sistema

$$(A - 2I)x = 0$$

(ossia trovando il Nucleo della matrice $A-2I$). In forma esplicita, questo sistema è

$$\begin{cases} x_2 + x_3 = 0 \\ x_2 + x_3 = 0 \\ -x_2 - x_3 = 0 \end{cases}$$

Le tre equazioni sono linearmente dipendenti, il che significa che questo sistema ha rango 1 e quindi ammette ∞^2 soluzioni (ossia che $\text{Nucleo}(A - 2I)$ è uno spazio vettoriale di dimensione 2), ossia che l'autovalore $\lambda=2$ ha molteplicità geometrica pari a 2: le soluzioni di quel sistema sono tutti i vettori del tipo

$$\begin{bmatrix} \alpha \\ \beta \\ -\beta \end{bmatrix}$$

Due vettori linearmente indipendenti di questo tipo sono

$$\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

per cui questi sono i due autovettori che possiamo associare a $\lambda=2$.

A questo punto, per trovare la matrice di trasformazione T , abbiamo bisogno di trovare un autovettore generalizzato: deve trattarsi di un vettore \vec{t} che soddisfi il sistema

$$(A - 2I)\vec{t} = \begin{bmatrix} \alpha \\ \beta \\ -\beta \end{bmatrix}$$

In forma esplicita, abbiamo che

$$\begin{cases} t_2 + t_3 = \alpha \\ t_2 + t_3 = \beta \\ -t_2 - t_3 = -\beta \end{cases}$$

Questo sistema risulta evidentemente compatibile solo se $\alpha=\beta$: allora, preso $\alpha=\beta=1$, il sistema diventa

$$\begin{cases} t_2 + t_3 = 1 \\ t_2 + t_3 = 1 \\ -t_2 - t_3 = -1 \end{cases}$$

ed ammette chiaramente ∞^2 soluzioni, che sono del tipo

$$\begin{bmatrix} t_1 \\ t_2 \\ 1 - t_2 \end{bmatrix}$$

Dobbiamo scegliere i valori di t_1 e t_2 in modo che il corrispondente autovettore generalizzato sia linearmente indipendenti dai due autovettori ottenuti prima: possiamo allora prendere $t_1=0$ e $t_2=1$, in modo da ottenere il vettore

$$\vec{t} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

In conclusione, la matrice di trasformazione richiesta è

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

A titolo di verifica, andiamo ora a trovare la matrice J : essendo $A = T^{-1}JT$, abbiamo che

$$J = TAT^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 0 & 3 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix}^{-1} = \dots = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Era ovvio che si ottenesse questo risultato: infatti, avendo A un solo autovalore, la matrice J si riduce ad un solo blocco; tale blocco, a sua volta, contiene 2 miniblocchi, tanti quanto è la molteplicità geometrica dell'autovalore: un miniblocco $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$, di dimensione 2, e poi un altro blocco $[2]$ che tiene conto del fatto che è stato necessario trovare un autovettore generalizzato associato all'autovalore in esame. Il numero di 1 sulla sopradiagonale è pari a $3-2=1$, dove 3 è la dimensione di A , mentre 2 è la molteplicità geometrica dell'unico autovalore di A .

Esempio

Consideriamo la matrice

$$A = \begin{bmatrix} -1 & 1 & -1 \\ 0 & 1 & -4 \\ 0 & 1 & -3 \end{bmatrix}$$

Vogliamo sempre determinare la forma canonica di Jordan di questa matrice.

Ci andiamo subito a calcolare gli autovalori di A : il polinomio caratteristico di A è $p(\lambda) = (\lambda + 1)^3$, il che significa che A ammette 1 solo autovalore ($\lambda = -1$) di molteplicità algebrica 3. Questo ci consente di dire immediatamente che la matrice J sarà formata da un solo blocco di Jordan di dimensione 3 avente sulla diagonale proprio $\lambda = -1$:

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

Adesso dobbiamo stabilire il numero di miniblocchi associati a $\lambda=-1$, ossia la molteplicità geometrica di $\lambda=-1$: tale molteplicità geometrica coincide con il rango del sistema

$$(A + I)x = 0$$

In forma esplicita, questo sistema è

$$\begin{cases} x_2 - x_3 = 0 \\ 2x_2 - 4x_3 = 0 \\ x_2 - x_3 = 0 \end{cases}$$

La prima e la terza equazione sono uguali e indipendenti dalla seconda: deduciamo che questo sistema ha rango 2 e quindi ammette ∞^1 soluzioni (ossia che $\text{Nucleo}(A + I)$ è uno spazio vettoriale di dimensione 1), ossia che l'autovalore $\lambda=-1$ ha molteplicità geometrica pari a 1. Avremo perciò 1 solo miniblocco di Jordan.

Resta da determinare il numero di 1 sulla sopradiagonale di J: esso sarà pari alla differenza tra la dimensione (3) della matrice A e la molteplicità geometrica (1) di $\lambda=-1$, per cui avremo 2 elementi unitari sulla diagonale.

In conclusione, la matrice J è fatta nel modo seguente:

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

Rango degli autovettori generalizzati

Abbiamo introdotto, nei paragrafi precedenti, il concetto di "autovettore generalizzato" associato ad un autovalore λ della matrice considerata. A questi autovettori generalizzati, nonché agli autovettori "tradizionali", è possibile associare il concetto di "rango". Vediamo di che si tratta.

Sia sempre A la generica matrice quadrata di ordine n della quale vogliamo trovare la forma di Jordan equivalente; siano $\lambda_1, \lambda_2, \dots, \lambda_p$ gli autovalori distinti di A, con molteplicità algebriche m_1, m_2, \dots, m_p e molteplicità geometriche g_1, g_2, \dots, g_p . Consideriamo il primo autovalore λ_1 : dire che g_1 è la sua molteplicità geometrica equivale a dire che il sistema

$$(A - \lambda_1 I)\bar{x} = 0$$

ha rango pari a g_1 , ossia ammette g_1 soluzioni $\bar{x}_{11}, \bar{x}_{12}, \dots, \bar{x}_{1g_1}$ che rappresentano gli autovettori linearmente indipendenti associati all'autovalore λ_1 : si dice allora che questi autovettori sono "**autovettori generalizzati di rango 1**" associati a λ_1 .

In modo del tutto analogo possiamo procedere per gli altri autovalori, per cui, in totale, possiamo trovare $g_1 + g_2 + \dots + g_p$ autovettori generalizzati di rango 1:

$$\begin{aligned} \lambda_1 &\xrightarrow{\text{autovettori generalizzati di rango 1}} \bar{x}_{11}, \bar{x}_{12}, \dots, \bar{x}_{1q_1} \\ \lambda_2 &\xrightarrow{\text{autovettori generalizzati di rango 1}} \bar{x}_{21}, \bar{x}_{22}, \dots, \bar{x}_{2q_2} \\ &\dots \\ \lambda_p &\xrightarrow{\text{autovettori generalizzati di rango 1}} \bar{x}_{p1}, \bar{x}_{p2}, \dots, \bar{x}_{pq_p} \end{aligned}$$

Sia adesso \bar{x}_{ik} un generico autovettore associato all'altrettanto generico autovalore λ_i : abbiamo visto in precedenza che, partendo da \bar{x}_{ik} , ha senso ricercare i suoi autovettori generalizzati applicando le formule

$$\begin{aligned} (A - \lambda_i I) \bar{t}_{i1}^{(1)} &= \bar{x}_{ik} \longrightarrow \bar{t}_{i1}^{(1)} \\ (A - \lambda_i I) \bar{t}_{i2}^{(1)} &= \bar{t}_{i1}^{(1)} \longrightarrow \bar{t}_{i2}^{(1)} \\ (A - \lambda_i I) \bar{t}_{i3}^{(1)} &= \bar{t}_{i2}^{(1)} \longrightarrow \bar{t}_{i3}^{(1)} \\ &\dots \end{aligned}$$

Questi n_k vettori $\bar{t}_{i1}^{(1)}, \bar{t}_{i2}^{(1)}, \dots, \bar{t}_{in_k}^{(1)}$ sono gli “*autovettori generalizzati*” associati a \bar{x}_{ik} ed a λ_i : si dice allora che $\bar{t}_{i1}^{(1)}$ è l'autovettore generalizzato di rango 2, che $\bar{t}_{i2}^{(1)}$ è l'autovettore generalizzato di rango 3 e così via fino a $\bar{t}_{in_k}^{(1)}$ che è l'autovettore generalizzato di rango n_k-1 .

L'insieme di vettori $\bar{x}_{ik}, \bar{t}_{i1}^{(1)}, \bar{t}_{i2}^{(1)}, \dots, \bar{t}_{in_k}^{(1)}$ prende il nome di “**catena**” associata all'autovalore λ_i e all'autovettore \bar{x}_{ik} . E' chiaro che il discorso può essere ripetuto pari pari per gli altri autovettori associati a λ_i e questo vale per tutti gli autovalori della matrice A considerata:

$$\begin{aligned} &\text{catena associata a } \lambda_1 \text{ e } \bar{x}_{11} \longrightarrow \bar{x}_{11}, \bar{t}_{11}^{(1)}, \bar{t}_{12}^{(1)}, \dots, \bar{t}_{1n_1}^{(1)} \\ &\text{catena associata a } \lambda_1 \text{ e } \bar{x}_{12} \longrightarrow \bar{x}_{12}, \bar{t}_{11}^{(2)}, \bar{t}_{12}^{(2)}, \dots, \bar{t}_{1n_2}^{(2)} \\ &\dots \\ &\text{catena associata a } \lambda_1 \text{ e } \bar{x}_{1g_1} \longrightarrow \bar{x}_{1g_1}, \bar{t}_{11}^{(g_1)}, \bar{t}_{12}^{(g_1)}, \dots, \bar{t}_{1n_{g_1}}^{(g_1)} \\ &\text{catena associata a } \lambda_2 \text{ e } \bar{x}_{21} \longrightarrow \bar{x}_{21}, \bar{t}_{21}^{(1)}, \bar{t}_{22}^{(1)}, \dots, \bar{t}_{2n_1}^{(1)} \\ &\text{catena associata a } \lambda_2 \text{ e } \bar{x}_{22} \longrightarrow \bar{x}_{22}, \bar{t}_{21}^{(2)}, \bar{t}_{22}^{(2)}, \dots, \bar{t}_{2n_2}^{(2)} \\ &\dots \\ &\text{catena associata a } \lambda_2 \text{ e } \bar{x}_{2g_2} \longrightarrow \bar{x}_{2g_2}, \bar{t}_{21}^{(g_2)}, \bar{t}_{22}^{(g_2)}, \dots, \bar{t}_{2n_{g_2}}^{(g_2)} \\ &\dots \\ &\text{catena associata a } \lambda_p \text{ e } \bar{x}_{p1} \longrightarrow \bar{x}_{p1}, \bar{t}_{p1}^{(1)}, \bar{t}_{p2}^{(1)}, \dots, \bar{t}_{pn_1}^{(1)} \\ &\text{catena associata a } \lambda_p \text{ e } \bar{x}_{p2} \longrightarrow \bar{x}_{p2}, \bar{t}_{p1}^{(2)}, \bar{t}_{p2}^{(2)}, \dots, \bar{t}_{pn_2}^{(2)} \\ &\dots \\ &\text{catena associata a } \lambda_p \text{ e } \bar{x}_{pg_p} \longrightarrow \bar{x}_{pg_p}, \bar{t}_{p1}^{(g_p)}, \bar{t}_{p2}^{(g_p)}, \dots, \bar{t}_{pn_{g_p}}^{(g_p)} \end{aligned}$$

Abbiamo dunque, per ciascun autovalore, un numero di catene di vettori pari alla molteplicità geometrica dell'autovalore stesso. Con questo schema si evidenziano gli autovettori generalizzati di rango 1 (vettori della 1° colonna), quelli di rango 2 (vettori della 2° colonna), quelli di rango 3 (vettori della 3° colonna) e così via.

Per concludere, l'ultima cosa da osservare è la seguente: è facile verificare che le relazioni

$$\begin{aligned} (A - \lambda_i I) \vec{x}_{ik} &= 0 \\ (A - \lambda_i I) \vec{t}_{i1}^{(1)} &= \vec{x}_{ik} \\ (A - \lambda_i I) \vec{t}_{i2}^{(1)} &= \vec{t}_{i1}^{(1)} \\ &\dots \\ (A - \lambda_i I) \vec{t}_{in_i}^{(1)} &= \vec{t}_{i,n_i-1}^{(1)} \end{aligned}$$

sono del tutto equivalenti alle relazioni

$$\begin{aligned} (A - \lambda_i I) \vec{x}_{ik} &= 0 \\ (A - \lambda_i I)^2 \vec{t}_{i1}^{(1)} &= 0 \quad \text{con} \quad (A - \lambda_i I) \vec{t}_{i1}^{(1)} \neq 0 \\ (A - \lambda_i I)^3 \vec{t}_{i2}^{(1)} &= 0 \quad \text{con} \quad (A - \lambda_i I)^2 \vec{t}_{i2}^{(1)} \neq 0 \\ &\dots \\ (A - \lambda_i I)^{n_i} \vec{t}_{in_i}^{(1)} &= 0 \quad \text{con} \quad (A - \lambda_i I)^{n_i-1} \vec{t}_{i,n_i-1}^{(1)} \neq 0 \end{aligned}$$

Queste relazioni sono utili perché consentono di stabilire quanti autovettori generalizzati è possibile associare al generico autovettore \vec{x}_{ik} : infatti, la ricerca degli autovalori va terminata quando si arriva ad un certo valore dell'esponente a tale che la matrice $(A - \lambda_i I)^a$ presenti degenerazione pari a m_i , ossia pari alla molteplicità algebrica di λ_i .

TEOREMA DI CAYLEY-HAMILTON

Sia sempre A una generica matrice quadrata di ordine n . Siano inoltre $q(x)$ un polinomio, nella variabile x , del tipo seguente:

$$q(x) = q_m x^m + q_{m-1} x^{m-1} + q_{m-2} x^{m-2} + \dots + q_2 x^2 + q_1 x + q_0$$

Allora, si definisce "**polinomio di matrice**" un polinomio la cui variabile indipendente, anziché essere una variabile scalare, è una matrice:

$$Q(A) = q_m A^m + q_{m-1} A^{m-1} + q_{m-2} A^{m-2} + \dots + q_2 A^2 + q_1 A + q_0$$

E' chiaro, da quest'ultima definizione, che un polinomio di matrice è a sua volta una matrice.

Il concetto di polinomio di matrice serve ad introdurre il "**teorema di Cayley-Hamilton**", che ha il seguente enunciato:

Teorema - Data una matrice $A \in \mathbb{C}^{n \times n}$ e dato il suo polinomio caratteristico $p(\lambda) = \lambda^n + a_1 \lambda^{n-1} + \dots + a_{n-1} \lambda + a_n$, la matrice A soddisfa la sua equazione caratteristica, ossia sussiste la relazione

$$p(A) = A^n + a_1 A^{n-1} + \dots + a_{n-1} A + a_n I = 0$$

Conseguenza: riduzione di polinomi

Tra le innumerevoli conseguenze del teorema di Cayley-Hamilton, c'è la seguente.

Sia A una matrice quadrata di ordine n e sia $p(\lambda) = \lambda^n + a_1 \lambda^{n-1} + \dots + a_{n-1} \lambda + a_n$ il suo polinomio caratteristico. Abbiamo appena visto che, in base al teorema di Cayley-Hamilton, sussiste la relazione

$$A^n + a_1 A^{n-1} + \dots + a_{n-1} A + a_n I = 0$$

da cui consegue che

$$A^n = -a_1 A^{n-1} - \dots - a_{n-1} A - a_n I$$

Se consideriamo ora la potenza $(n+1)$ sima della matrice A , possiamo scrivere che

$$A^{n+1} = A \cdot A^n = -a_1 A^n - a_2 A^{n-1} - \dots - a_{n-1} A^2 - a_n A$$

e lo stesso avremmo per le potenze successive. In altre parole, abbiamo trovato che le potenze della matrice A , a partire da quella di ordine n , si possono ricavare come combinazioni lineari delle potenze precedenti usando, come coefficienti, quelli del polinomio caratteristico di A .

Questo consente di ridurre ogni polinomio di matrice di grado $m \geq n$ ad un polinomio di grado $n-1$. Per chiarirci, facciamo un esempio.

Esempio

Sia dato il polinomio di matrice

$$Q(A) = A^4 + A^3 + A^2 + A + I$$

dove A è la matrice

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

Il polinomio $Q(A)$ è di grado 4: allora, essendo A una matrice di ordine $n=2$, possiamo ridurre $Q(A)$ ad un polinomio di ordine $n-1=1$. Per fare questo, dobbiamo esprimere A^4 , A^3 ed A^2 in funzione di A e, a questo scopo, ci serve il polinomio caratteristico di A (del quale ci interessano i coefficienti):

$$p(\lambda) = \det(\lambda I - A) = \det \begin{bmatrix} \lambda & -1 \\ 2 & \lambda + 3 \end{bmatrix} = \lambda^2 + 3\lambda + 2$$

Noto questo polinomio caratteristico, possiamo scrivere, in base al teorema di Cayley-Hamilton, che

$$p(A) = A^2 + 3A + 2I = 0 \longrightarrow A^2 = -3A - 2I$$

da cui quindi ricaviamo che

$$A^3 = -3A^2 - 2A = -3(-3A - 2I) - 2A = 7A + 6I$$

$$A^4 = A \cdot A^3 = 7A^2 + 6A = 7(-3A - 2I) + 6A = -15A - 14I$$

Andando infine a sostituire nell'espressione di $Q(A)$, otteniamo quanto segue:

$$\begin{aligned} Q(A) &= A^4 + A^3 + A^2 + A + I = -15A - 14I + 7A + 6I - 3A - 2I + A + I = \\ &= -10A - 9I \end{aligned}$$

Abbiamo dunque concluso che $Q(A) = -10A - 9I$ e questa è ovviamente una espressione che consente di calcolare in modo molto più rapido quanto valga $Q(A)$, in quanto basta sostituire l'espressione di A senza dover calcolare le potenze di tale matrice.

POLINOMIO MINIMO DI UNA MATRICE

Sia sempre A una matrice quadrata di ordine n e sia $p(\lambda) = \lambda^n + a_1\lambda^{n-1} + \dots + a_{n-1}\lambda + a_n$ il suo polinomio caratteristico. Abbiamo visto che, in base al teorema di Cayley-Hamilton, risulta $p(A) = 0$. E' intuitivo pensare al fatto che il polinomio caratteristico di A non sia l'unico polinomio che si annulla quando l'argomento è la matrice A . In altre parole, se consideriamo l'insieme

$$\{P'(A) \mid P'(A) = 0\}$$

questo insieme comprende un numero infinito di elementi, ossia appunto un numero infinito di polinomi che, calcolati in A , si annullano. Tra questi polinomi gode di particolare importanza il cosiddetto "**polinomio minimo**", il quale gode di due fondamentali proprietà (oltre a quella, evidentemente, di annullarsi in A):

- in primo luogo, è il polinomio di grado minimo tra quelli che si annullano in A ;
- in secondo luogo, è un polinomio monico, il che significa che il coefficiente della potenza di grado massimo è unitario.

Vediamo come si costruisce il polinomio minimo di una assegnata matrice A quadrata di ordine n .

In primo luogo, supponiamo che la matrice A abbia $p(\leq n)$ autovalori distinti, che indichiamo con $\lambda_1, \lambda_2, \dots, \lambda_p$; in base a quanto visto a proposito della forma canonica di Jordan, sappiamo che a ciascun autovalore λ_k è associato un blocco di Jordan J_k e, a sua volta, tale blocco è costituito da un numero di miniblocchi J_{ki} (che ricordiamo hanno l'autovalore sulla diagonale principale e la

sopradiagonale formata da elementi unitari) pari alla molteplicità geometrica di λ_k . Questi blocchi J_{k_i} hanno ciascuno una propria dimensione: indichiamo allora con v_k il massimo ordine dei blocchi J_{k_i} associati a λ_k (si tratta cioè della dimensione del blocco J_{k_i} di dimensione maggiore). Così facendo, in corrispondenza dei p autovalori distinti, avremo p dimensioni v_1, v_2, \dots, v_p , che prendono il nome di “**indici**” degli autovalori $\lambda_1, \lambda_2, \dots, \lambda_p$: si dimostra, allora, che il polinomio minimo associato alla matrice A è

$$m(\lambda) = (\lambda - \lambda_1)^{v_1} (\lambda - \lambda_2)^{v_2} \dots (\lambda - \lambda_p)^{v_p}$$

Facendo un confronto con il polinomio caratteristico $p(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_p)^{m_p}$, ci accorgiamo che l’espressione è la stessa, con la differenza che, nel polinomio caratteristico, gli esponenti dei termini $(\lambda - \lambda_k)$ sono le rispettive molteplicità algebriche.

Esempio

Consideriamo la matrice

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

Vogliamo calcolare il suo polinomio minimo.

La prima cosa che si osserva è che A si trova già in forma di Jordan: si deduce, allora, immediatamente, che essa ha un solo autovalore ($\lambda = -2$) di molteplicità algebrica $m=3$ e di molteplicità geometrica $g=2$. Il fatto che $g=2$ ci dice che all’autovalore $\lambda = -2$ sono associati 2 blocchi, di cui uno di dimensione 2 e l’altro di dimensione 1. Deduciamo che $v = 2$ e quindi che il polinomio minimo della matrice è

$$m(\lambda) = (\lambda + 2)^2$$

mentre il polinomio caratteristico è ovviamente $p(\lambda) = (\lambda + 2)^3$.

A titolo di parziale verifica dei calcoli, potremmo verificare che sia $m(\lambda)$ sia $p(\lambda)$ si annullano in $\lambda = -2$.

Osservazione

Questo esempio aiuta a comprendere una importante proprietà: se la matrice $\lambda I - A$ (detta “**matrice caratteristica associata ad A** ”), calcolata in corrispondenza di tutti gli autovalori di A , ha sempre degenerazione semplice, ossia ha sempre il Nucleo di dimensione 0 (il che significa che a ciascun autovalore di A corrisponde un solo blocco di Jordan), allora il polinomio caratteristico e quello minimo coincidono; viceversa, se esiste almeno un autovalore al quale sono associati almeno due blocchi di Jordan, allora i due polinomi sono senz’altro distinti.

DEFINIZIONI VARIE

Sia data una matrice A quadrata di dimensione n:

Def. Si definisce "**minore fondamentale**" estratto da A una qualsiasi sottomatrice quadrata M di A che goda delle proprietà per cui $\det M \neq 0$ e $\rho(M) = \rho(A)$

Data una stessa matrice A, è possibile estrarre da essa più di un minore fondamentale.

.....

RICHIAMI SUI SISTEMI LINEARI

Sia dato un sistema lineare di m equazioni di 1° grado in n incognite:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

Sia S l'insieme di tutte le eventuali soluzioni di questo sistema.

Questo sistema si definisce "**incompatibile**" se l'insieme S è un insieme vuoto, ossia se il sistema non ammette alcuna soluzione. Al contrario, se S contiene almeno un elemento, allora il sistema si definisce "**compatibile**": in particolare, se S contiene un numero finito di elementi, allora si dice che il sistema è "determinato"; se, invece, l'insieme S contiene un numero infinito di elementi, allora si dice che il sistema è "indeterminato".

Il sistema risulta compatibile se e solo se la matrice dei coefficienti A ha lo stesso rango della matrice ottenuta da A aggiungendo la colonna dei termini noti. Se, poi, il rango di tali matrici è pari al numero n di incognite (e di equazioni), allora il sistema ammette 1 sola soluzione.

Se, nel sistema il numero n di incognite è pari al numero di equazioni e la matrice dei coefficienti (che risulta quadrata di dimensione n) ha rango anch'esso pari ad n, allora il sistema si definisce "sistema di Kramer", in quanto è compatibile ed ammette soluzione unica.

Un caso particolare si ha quando il sistema è omogeneo, ossia quando sono nulli tutti i termini noti: in questo caso, intanto il sistema è senz'altro compatibile; inoltre, se il rango della matrice dei coefficienti è pari al numero delle incognite, allora il sistema ammette come unica soluzione la soluzione banale; viceversa, se il rango p della matrice dei coefficienti è minore del numero n di incognite, allora il sistema ammette ∞^{n-p} soluzioni.

Autore: **SANDRO PETRIZZELLI**
e-mail: sandry@iol.it
sito personale: <http://users.iol.it/sandry>
succursale: <http://digilander.iol.it/sandry1>