Appunti per il Corso di Matematica 2

Marco A. Garuti

1 marzo 2005

1 Cenni sulla teoria di Jordan

Definizione 1 *Diremo* blocco di Jordan *di dimensione n e di autovalore λ la matrice $J_{n,\lambda} \in M_n(\mathbb{R})$ definita dalle proprietà seguenti:*

a) Gli elementi sulla diagonale sono tutti uguali a λ;

b) Gli elementi di posto $(i, i+1)$ sono tutti uguali ad 1, per $i = 1, \ldots, n-1$;

c) Tutti gli altri elementi sono nulli, in particolare $J_{n,\lambda}$ è triangolare superiore.

Come esempio abbiamo quindi i seguenti blocchi di dimensione 1, 2, 3:

$$J_{1,\lambda} = (\lambda); \qquad J_{2,\lambda} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}; \qquad J_{3,\lambda} = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Osserviamo innanzitutto che $p_{J_{n,\lambda}}(t) = (\lambda - t)^n$. Inoltre il rango di $(J_{n,\lambda} - \lambda I_n)$ è $n-1$ quindi l'autospazio V_λ per la matrice $J_{n,\lambda}$ ha dimensione 1.
Se L è un endomorfismo di $\mathbb{R}^n$, dire che la matrice di L in una base $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ è $J_{n,\lambda}$, vuol dire che:

$$L(\mathbf{v}_1) = \lambda \mathbf{v}_1, \quad L(\mathbf{v}_2) = \mathbf{v}_1 + \lambda \mathbf{v}_2, \quad \ldots \quad L(\mathbf{v}_n) = \mathbf{v}_{n-1} + \lambda \mathbf{v}_n.$$

Quindi $\mathbf{v}_1$ è un autovettore di autovalore λ mentre $\mathbf{v}_2$ soddisfa la relazione $L(\mathbf{v}_2) - \lambda \mathbf{v}_2 = \mathbf{v}_1$ e così via. Sia A la matrice di L rispetto alla base canonica di $\mathbb{R}^n$: come sappiamo le coordinate del vettore $\mathbf{v}_1$ saranno soluzione del sistema omogeneo:

$$(A - \lambda I_n)X = 0.$$

Supponiamo che queste coordinate siano $(a_1, \ldots, a_n)$: le coordinate del vettore $\mathbf{v}_2$ sono allora soluzione del sistema *non omogeneo*:

$$(A - \lambda I_n)X = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Supponiamo di averle calcolate e di aver scelto la soluzione $(b_1, \ldots, b_n)$: le coordinate del vettore $\mathbf{v}_3$ sono allora soluzione del sistema non omogeneo

$$(A - \lambda I_n)X = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

e così via ...

Definizione 2 *Una* matrice di Jordan *è una matrice quadrata a blocchi tale che*

a) i blocchi lungo la diagonale sono blocchi di Jordan;

b) i blocchi esterni alla diagonale sono matrici nulle.

In particolare, una matrice di Jordan è triangolare superiore ed i suoi elementi non nulli si trovano o sulla diagonale o sopra un elemento della diagonale cioè se $a_{i,j} \neq 0$ allora $j = i$ oppure $j = i + 1$ (e in questo secondo caso $a_{i,i+1} = 1$).
Ovviamente, un blocco di Jordan è una matrice di Jordan con un solo blocco lungo la diagonale. Una matrice diagonale $n \times n$ è una matrice di Jordan con n blocchi 1×1 lungo la diagonale. Ecco degli esempi di matrici di Jordan con due blocchi lungo la diagonale

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}; \quad \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}; \quad \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 1 \\ 0 & 0 & \mu \end{pmatrix}; \quad \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \mu & 1 \\ 0 & 0 & 0 & \mu \end{pmatrix}$$

ed ancora degli esempi di matrici di Jordan con tre blocchi lungo la diagonale:

$$\begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}; \quad \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & \nu \end{pmatrix}; \quad \begin{pmatrix} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & \mu & 1 & 0 \\ 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & \nu \end{pmatrix}; \quad \begin{pmatrix} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & \nu \end{pmatrix}.$$

Data una qualunque matrice quadrata $A \in M_n(\mathbb{R})$ ci siamo posti il problema di trovare una matrice simile ad A che abbia la forma *più semplice possibile*. La soluzione migliore si ha quando A è diagonalizzabile, cioè simile ad una matrice diagonale. Come abbiamo visto, A è diagonalizzabile se e solo se sono soddisfatte due condizioni:

a) il polinomio caratteristico $p_A(t)$ è completamente riducibile;

b) le molteplicità algebrica e geometrica di ogni autovalore coincidono.

Cosa possiamo dire se queste condizioni non sono soddisfatte? Se la prima condizione non è verificata, possiamo sempre decidere di spostare il problema da $\mathbb{R}$ a $\mathbb{C}$: per il teorema fondamentale dell'algebra, ogni polinomio $p(t) \in \mathbb{C}[t]$ è completamente riducibile.
E se invece è la seconda condizione a non essere soddisfatta? Come abbiamo visto, gli esempi più semplici sono dati da matrici che sono blocchi di Jordan di dimensione almeno 2. Più in generale, lo stesso vale per una matrice di Jordan che non sia diagonale. D'altra parte non possiamo certo sperare di trovare una matrice ancora più semplice! Per il seguente teorema, di cui omettiamo la dimostrazione, possiamo sempre ricondurci a matrici di Jordan.

Teorema 1 *Sia A una matrice quadrata $n \times n$ il cui polinomio caratteristico sia completamente riducibile. Allora A è simile ad una matrice di Jordan. Inoltre, se J è una matrice di Jordan simile ad A e se λ è un autovalore di A, allora per i blocchi di autovalore λ in J abbiamo che:*

a) il numero di tali blocchi è pari alla molteplicità geometrica di λ;

b) la somma delle dimensioni di tali blocchi è pari alla molteplicità algebrica di λ.

Osserviamo che, se per un autovalore λ la molteplicità geometrica $m_g(\lambda)$ coincide con la molteplicità algebrica $m_a(\lambda)$, nella J i blocchi relativi a λ saranno $m_a(\lambda) = m_g(\lambda)$, necessariamente tutti di dimensione uno. In particolare, se questo vale per ogni λ, la matrice J è diagonale, come del resto risulta dal criterio di diagonalizzazione.

2 Diagonalizzazione delle matrici simmetriche reali

Ricordiamo che una matrice $H \in M_n(\mathbb{R})$ è detta *ortogonale* se $H^T H = I_n$. In altre parole, H è invertibile ed $H^{-1} = H^T$.

Definizione 3 *Diremo che una matrice $A \in M_n(\mathbb{R})$ è* ortogonalmente diagonalizzabile *se esistono una matrice diagonale D ed una matrice ortogonale H tali che $A = HDH^{-1} = HDH^T$.*

Equivalentemente, A è diagonalizzabile ed ammette una base *ortonormale* di autovettori. Per verificare che una matrice è ortogonalmente diagonalizzabile, basta verificare che è diagonalizzabile e che gli autospazi sono tra loro ortogonali: infatti col procedimento di Gram-Schmidt, possiamo costruire una base ortonormale per ciascun autospazio e, se questi autospazi sono tra loro ortogonali, l'unione di tutte queste basi, che è una base di $\mathbb{R}^n$ perché A è diagonalizzabile, è anche ortonormale.
Osserviamo innanzitutto che se A è ortogonalmente diagonalizzabile, A è necessariamente una matrice simmetrica. Infatti, ricordando che $(MN)^T = N^T M^T$:

$$A = HDH^T \quad \Rightarrow \quad A^T = \left(HDH^T\right)^T = HD^T H^T = HDH^T = A$$

In effetti, vale anche il viceversa, come afferma il seguente teorema:

Teorema 2 *Se $A \in M_n(\mathbb{R})$ è simmetrica, è ortogonalmente diagonalizzabile.*

Dobbiamo dimostrare tre cose:

a) Il polinomio caratteristico $p_A(t)$ è completamente riducibile.

b) Gli autospazi di A sono tra loro ortogonali.

c) A è diagonalizzabile.

(a) Sia $\lambda \in \mathbb{C}$ una radice di $p_A(t)$, quindi λ è un autovalore della matrice A pensata come matrice a coefficienti complessi: esiste quindi un vettore non nullo $Z^T = (z_1, \dots, z_n) \in \mathbb{C}^n$ tale che $AZ = \lambda Z$. Data una matrice $M = (m_{i,j})$ a coefficienti complessi, conveniamo di porre $\bar{M} = (\bar{m}_{i,j})$ la matrice ottenuta da M prendendo il coniugato di tutti i coefficienti. In particolare, preso $\bar{Z}^T = (\bar{z}_1, \dots, \bar{z}_n)$, osserviamo che

$$Z^T \bar{Z} = (z_1 \dots z_n) \begin{pmatrix} \bar{z}_1 \\ \vdots \\ \bar{z}_n \end{pmatrix} = z_1\bar{z}_1 + \cdots + z_n\bar{z}_n = |z_1|^2 + \cdots + |z_n|^2$$

è un numero reale, positivo perché almeno uno degli z_i è diverso da zero. Chiamiamo $c = Z^T \bar{Z}$. Allora, poiché $A^T = A$ (perché A è simmetrica) ed $\bar{A} = A$ (perché A è a coefficienti reali), abbiamo

$$\lambda c = (\lambda Z)^T \bar{Z} = (AZ)^T \bar{Z} = Z^T A \bar{Z} = Z^T \overline{(AZ)} = Z^T \overline{(\lambda Z)} = \bar{\lambda} Z^T \bar{Z} = \bar{\lambda} c.$$

Quindi $c(\lambda - \bar{\lambda}) = 0$ e siccome $c \neq 0$, deve essere $\lambda = \bar{\lambda}$, cioè $\lambda \in \mathbb{R}$.
Per il teorema fondamentale dell'algebra, $p_A(t)$ è completamente riducibile su $\mathbb{C}$ e, per quanto appena visto, tutte le sue radici sono reali: quindi $p_A(t)$ è completamente riducibile su $\mathbb{R}$.

(b) Siano $\lambda \neq \mu$ due autovalori di A e siano $X^T = (x_1, \dots, x_n)$ ed $Y^T = (y_1, \dots, y_n)$ autovettori di autovalore rispettivamente λ e μ, cioè $AX = \lambda X$ ed $AY = \mu Y$. Allora

$$\begin{aligned} \lambda(x_1, \dots, x_n) \bullet (y_1, \dots, y_n) &= (\lambda X^T)Y \\ &= (AX)^T Y \\ &= X^T A Y \\ &= X^T(\mu Y) \\ &= \mu(x_1, \dots, x_n) \bullet (y_1, \dots, y_n). \end{aligned}$$

Quindi $(\lambda - \mu)(x_1, \dots, x_n) \bullet (y_1, \dots, y_n) = 0$ e, siccome $\lambda \neq \mu$ deve essere

$$(x_1, \dots, x_n) \bullet (y_1, \dots, y_n) = 0.$$

Quindi autovettori relativi ad autovalori distinti sono ortogonali. Pertanto gli autospazi di A sono a due a due ortogonali.

(c) Dimostriamo che A è diagonalizzabile per induzione su n. Se $n = 1$ non c'è nulla da dimostrare. Supponiamo quindi che ogni matrice simmetrica $(n-1)\times(n-1)$ sia diagonalizzabile. Sia $\mathbf{v}$ un autovettore di A di autovalore λ e sia $W =< \mathbf{v} >^{\perp}$. Quindi W è un sottospazio di dimensione $n-1$ ed abbiamo la decomposizione

$$\mathbb{R}^n =< \mathbf{v} > \oplus W.$$

Sia L l'endomorfismo di $\mathbb{R}^n$ definito da A. Osserviamo che il trasformato $L(\mathbf{w})$ di un vettore $\mathbf{w} \in W$ mediante L appartiene ancora a W, cioè è ortogonale a $\mathbf{v}$. Infatti, indicando con $X^T = (x_1, \dots, x_n)$ ed $Y^T = (y_1, \dots, y_n)$ le coordinate di $\mathbf{v}$ e $\mathbf{w}$ rispettivamente, allora

$$\mathbf{v} \bullet L(\mathbf{w}) = X^T(AY) = (AX)^T Y = \lambda X^T Y = \lambda(\mathbf{v} \bullet \mathbf{w}) = 0.$$

Quindi l'endomorfismo L manda il sottospazio $W =< \mathbf{v} >^{\perp}$ in se stesso. Scegliamo allora la base ortonormale $\mathbf{v}_1 = \frac{\mathbf{v}}{\|\mathbf{v}\|}$ di $< \mathbf{v} >$ ed una base ortonormale $\{\mathbf{w}_2, \dots, \mathbf{w}_n\}$ di W (possiamo sempre trovarne una grazie al teorema di Gram-Schmidt). Unendole, otteniamo una base ortonormale di $\mathbb{R}^n$. Quindi la matrice H del cambiamento di base in $\mathbb{R}^n$ dalla base $\{\mathbf{v}_1, \mathbf{w}_2, \dots, \mathbf{w}_n\}$ alla base canonica è una matrice ortogonale. La matrice $H^T AH$, che è ancora una matrice simmetrica, rappresenta l'endomorfismo L nella base $\{\mathbf{v}_1, \mathbf{w}_2, \dots, \mathbf{w}_n\}$.
Poiché $L(\mathbf{v}_1) = \lambda \mathbf{v}_1$ ed L manda il sottospazio W in sé stesso, questa matrice è della forma

$$H^T AH = \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 0 & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & b_{n2} & \dots & b_{nn} \end{pmatrix}. \quad \text{Poniamo} \quad B = \begin{pmatrix} b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \vdots \\ b_{n2} & \dots & b_{nn} \end{pmatrix}.$$

Poiché $H^T AH$ è simmetrica, anche B è una matrice simmetrica. Ma abbiamo supposto che ogni matrice simmetrica $(n-1) \times (n-1)$ sia diagonalizzabile. Quindi B è diagonalizzabile e di conseguenza anche $H^T AH$, che è simile ad A. Pertanto anche A è diagonalizzabile.