Alma Mater Studiorum · Università di Bologna

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

UN APPROCCIO GEOMETRICO AI SISTEMI DI EQUAZIONI DIFFERENZIALI

Tesi di Laurea in Analisi Matematica

Relatore:
Chiar.mo Prof.
Angelo Favini

Presentata da:
Sara Calandrini

I Sessione
Anno Accademico 2010/11

Introduzione

Con questa tesi voglio trattare in modo inusuale lo studio dei sistemi di equazioni differenziali lineari del primo ordine a coefficenti costanti, ovvero vedremo l'applicazione delle funzioni di matrici e dei fasci di matrici all'integrazione di un sistema di equazioni differenziali lineari a coefficenti costanti. Come primo caso tratteremo un sistema di equazioni differenziali lineari omogenee del primo ordine a coefficenti costanti

$$\frac{dx}{dt} = Ax$$

dove $x_1, \ldots, x_n$ sono le componenti dell'incognita x. Come secondo caso tratteremo un sistema di equazioni differenziali lineari non omogenee a coefficenti costanti

$$\frac{dx}{dt} = Ax + f(t)$$

dove $f_1(t), \ldots, f_n(t)$ sono funzioni continue e le incognite sono $x_1, \ldots, x_n$; ed infine ci occuperemo di un sistema di equazioni differenziali del primo ordine a coefficenti costanti

$$Ax + B\frac{dx}{dt} = f(t) \quad ,$$

dove $x_1, \ldots, x_n$ sono le incognite mentre A e B sono matrici $m \times n$ e B é possibilmente non invertibile.

Per fare ció avremo bisogno di ricordare importanti nozioni geometriche.

La tesi é suddivisa in quattro capitoli.

Il primo é un capitolo introduttivo che ci servirá per trattare, nel secondo, i primi due tipi di sistemi di equazioni differenziali introdotti sopra. É dedicato alle nozioni geometriche di base, come quella di funzione di matrice,

di polinomio di interpolazione di Lagrange-Sylvester, di matrice costituente, ecc..., che ci accompagneranno per tutto il percorso.
Il secondo capitolo tratta dei primi due tipi di equazioni differenziali enunciati sopra e mostra due esempi, uno per ciascun tipo di sistema.
Il terzo capitolo, come il primo, é un capitolo introduttivo che ci aprirá la strada all'ultimo tipo di sistema di equazioni differenziali. Verranno trattate, quindi, nozioni geometriche come i fasci di matrici regolari e singolari e la forma canonica di un fascio di matrici.
Il quarto capitolo, invece, affronta l'ultimo tipo di sistema di equazioni differenziali visto sopra utilizzando gli strumenti geometrici fornitici dal capitolo precedente.
In tutta questa trattazione mi sono basata principalmente su due libri, entrambi scritti dall'autore F.R.Gantmacher, intitolati ″Théorie des matrices, Tome 1, Théorie generale‶ e ″Théorie des matrices, Tome 2, Questions spéciales et applications‶.

Indice

Capitolo 1

In questo capitolo daremo delle nozioni geometriche di base per poter, nei capitoli successivi, trattare l'integrazione di sistemi di equazioni differenziali a coefficenti costanti.
Sia $A = ||a_{ik}||_1^n$ una matrice quadrata e $f(\lambda)$ una funzione di argomento scalare λ. Cercheremo di dare un significato a $f(A)$, cioé di estendere la funzione $f(\lambda)$ a dei valori matriciali dell'argomento ed illustreremo 3 metodi per calcolare nella pratica $f(A)$. Prima di fare ció daremo una serie di definizioni e proprietá che ci serviranno per perseguire questi nostri due obiettivi.

Definizione 1.0.1. *Un polinomio scalare $f(\lambda)$ é chiamato polinomio annullatore della matrice quadrata A se $f(A) = 0$. Il polinomio annullatore $\psi(\lambda)$ di grado piú basso e di cui il coefficente di grado piú alto é uguale a 1 é chiamato polinomio minimo di A.*

Il polinomio caratteristico é, chiaramente, un polinomo annullatore di A ma, in generale, esso é diverso dal polimonio minimo. Due proprietá importanti del polinomio minimo (che non dimostreremo) sono:

- Tutti i polinomi annullatori di una matrice sono divisibili senza resto per il polinomio minimo.
- Il polinomio minimo di una data matrice A é unico.

Definizione 1.0.2. *Consideriamo la matrice A e consideriamo il suo polinomio caratteristico $\Delta(\lambda) = |\lambda I - A|$. La matrice*

$$B(\lambda) = ||b_{ik}(\lambda)||_1^n,$$

dove $b_{ik}(\lambda)$ é il complemento algebrico dell'elemento $\lambda\delta_{ik} - a_{ik}$ di $\Delta(\lambda)$, é chiamata la matrice aggiunta di A.

Questa definizione implica le indentitá seguenti in λ:

$$(\lambda I - A)B(\lambda) = \Delta(\lambda)I, \tag{1.1}$$

$$B(\lambda)(\lambda I - A) = \Delta(\lambda)I. \tag{1.2}$$

Definizione 1.0.3. *Consideriamo la matrice A. Sia $\lambda I - A$ la sua matrice caratteristica e sia $\Delta(\lambda)$ il suo polinomio caratteristico. Indichiamo con $D_{n-1}(\lambda)$ il massimo comun divisore di tutti i minori di ordine $n-1$ della matrice caratteristica, cioé di tutti gli elementi della matrice $B(\lambda) = ||b_{ik}||_1^n$. Allora*

$$B(\lambda) = D_{n-1}(\lambda)C(\lambda), \tag{1.3}$$

dove $C(\lambda)$ é una matrice polinomiale detta matrice aggiunta ridotta di $\lambda I - A$.

Da (1.1) e (1.3) segue che

$$\Delta(\lambda)I = (\lambda I - A)C(\lambda)D_{n-1}(\lambda). \tag{1.4}$$

Dato che $\Delta(\lambda)$ é divisibile senza resto per $D_{n-1}(\lambda)$[1] vale:

$$\frac{\Delta(\lambda)}{D_{n-1}(\lambda)} = \psi(\lambda) \tag{1.5}$$

dove $\psi(\lambda)$ é il polinomio minimo. Sostituendo (1.5) in (1.4) otteniamo:

$$\psi(\lambda)I = (\lambda I - A)C(\lambda).$$

Ora passiamo al primo dei nostri due problemi: dare una definizione appropriata di $f(A)$.

[1] Se no esisterebbe un polinomio annullatore di grado inferiore a quello del polinomio minimo.

1.0.1 Funzione di una matrice

Nel caso piú semplice in cui $f(\lambda)$ é un polinomio in λ, cioé:
$f(\lambda) = \gamma_0\lambda^n + \gamma_1\lambda^{n-1} + ... + \gamma_n$ abbiamo che:

$$f(A) = \gamma_0 A^n + \gamma_1 A^{n-1} + ... + \gamma_n I.$$

Da questo caso particolare passiamo al caso piú generale. Consideriamo

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_s)^{m_s} \tag{1.6}$$

il polinomio minimo di A, dove $\lambda_1, \lambda_2, ..., \lambda_s$ sono le radici distinte del polinomio caratteristico. Il grado di questo polinomio é

$$m = \sum_{k=1}^{n} m_k.$$

Siano $g(\lambda)$ e $h(\lambda)$ due polinomi tali che $g(A) = h(A)$.
Allora la differenza $d(\lambda) = g(\lambda) - h(\lambda)$ é un polinomio annullatore di A e quindi é divisibile senza resto per $\psi(\lambda)$.
Scriviamo questo fatto nel modo seguente:

$$g(\lambda) \equiv h(\lambda) \quad [mod\ \psi(\lambda)]. \tag{1.7}$$

Da (1.6) abbiamo: $d(\lambda_k) = 0, d'(\lambda_k) = 0, \ldots, d^{(m_k-1)}(\lambda_k) = 0 \quad (k = 1, 2, \cdots, s)$, cioé:

$$g(\lambda_k) = h(\lambda_k), g'(\lambda_k) = h'(\lambda_k), \ldots, g^{(m_k-1)}(\lambda_k) = h^{(m_k-1)}(\lambda_k) \quad (k = 1, 2, \ldots, s). \tag{1.8}$$

Gli m numeri

$$f(\lambda_k), f'(\lambda_k), \ldots, f^{(m_k-1)}(\lambda_k) \quad (k = 1, 2, \ldots, s) \tag{1.9}$$

sono chiamati i *valori della funzione* $f(\lambda)$ *sullo spettro della matrice* A e l'insieme di tutti questi valori é indicato con $f(\Lambda_A)$. Se per una funzione $f(\lambda)$ i valori (1.9) esistono, diremo che *la funzione* $f(\lambda)$ *é definita sullo spettro della matrice* A. L'equazione (1.8) mostra che i polinomi $g(\lambda)$ e $h(\lambda)$ hanno

gli stessi valori sullo spettro di A per cui $g(\Lambda_A) = h(\Lambda_A)$.
Cosí, data una matrice A, i valori del polinomio $g(\lambda)$ sullo spettro di A determinano completamente la matrice $g(A)$. Noi postuleremo che, nel caso generale, la definizione di $f(A)$ segua lo stesso principio: i valori della funzione $f(\lambda)$ sullo spettro della matrice A devono determinare $f(A)$ completamente, cioé tutte le funzioni $f(\lambda)$ aventi gli stessi valori sullo spettro di A devono avere lo stesso valore matriciale $f(A)$.
Quindi abbiamo la definizione seguente:

Definizione 1.0.4. *Se la funzione $f(\lambda)$ é definita sullo spettro della matrice A, allora*

$$f(A) = g(A)$$

dove $g(\lambda)$ é un polinomio arbitrario che prende, sullo spettro di A, gli stessi valori di $f(\lambda)$:

$$f(\Lambda_A) = g(\Lambda_A).$$

Tra i polinomi a coefficenti complessi che prendono sullo spettro di A gli stessi valori di $f(\lambda)$, c'é uno e un solo polinomio $r(\lambda)$ di grado inferiore a m (dove con m indichiamo ancora il grado del polinomio minimo). Il polinomio $r(\lambda)$ é determinato in modo unico dalle condizioni di interpolazione:

$$r(\lambda_k) = f(\lambda_k), r'(\lambda_k) = f'(\lambda_k), \ldots, r^{(m_k-1)} = f^{(m_k-1)}(\lambda_k) \quad (k = 1, 2, \ldots, s). \tag{1.10}$$

Il polinomio $r(\lambda)$ é chiamato *polinomio di interpolazione di Lagrange-Sylvester* per $f(\lambda)$ sullo spettro di A. La definizione (1.0.4) puó essere quindi formulata anche nel seguente nuovo modo:

Definizione 1.0.5. *Siano $f(\lambda)$ una funzione definita sullo spettro di una matrice A e $r(\lambda)$ il polinomio d'interpolazione di Lagrange-Sylvester corrispondente, allora*

$$f(A) = r(A).$$

A questo punto siamo pronti per trattare il calcolo effettivo di $f(A)$. Come abbiamo giá preannunciato, vedremo tre diversi metodi per farlo.
Consideriamo il primo.

1.0.2 Polinomio d'interpolazione di Lagrange-Sylvester

Calcoliamo nella pratica $f(A)$ sfruttando la definizione (1.0.5) appena data. Abbiamo tre casi:

1) Consideriamo il caso in cui l'equazione caratteristica $|\lambda I - A| = 0$ non ha radici multiple. Le radici di questa equazione sono indicate con $\lambda_1, \lambda_2, \ldots, \lambda_n$. Allora:

$$\psi(\lambda) = |\lambda I - A| = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

perché, in questo caso, il polinomio caratteristico coincide con quello minimo e la condizione (1.10) puó essere scritta

$$r(\lambda_k) = f(\lambda_k) \quad (k = 1, 2, \ldots, n).$$

In questo caso $r(\lambda)$ é il polinomio di interpolazione ordinario di Lagrange per la funzione $f(\lambda)$ nei punti $\lambda_1, \lambda_2, \ldots, \lambda_n$:

$$r(\lambda) = \sum_{k=1}^{n} \frac{(\lambda - \lambda_1) \cdots (\lambda - \lambda_{k-1})(\lambda - \lambda_{k+1}) \ldots (\lambda - \lambda_n)}{(\lambda_k - \lambda_1) \cdots (\lambda_k - \lambda_{k-1})(\lambda_k - \lambda_{k-1}) \ldots (\lambda_k - \lambda_n)} f(\lambda_k).$$

Secondo la definizione (1.0.5)

$$f(A) = r(A) = \sum_{k=1}^{n} \frac{(A - \lambda_1 I) \cdots (A - \lambda_{k-1} I)(A - \lambda_{k+1} I) \ldots (A - \lambda_n I)}{(\lambda_k - \lambda_1) \cdots (\lambda_k - \lambda_{k-1})(\lambda_k - \lambda_{k-1}) \ldots (\lambda_k - \lambda_n)} f(\lambda_k).$$

2) Supponiamo ora che il polinomio caratteristico abbia radici multiple, ma che il polinomio minimo non ne abbia, ovvero abbia solo radici semplici, per cui $\psi(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_m)$.

In questo caso, come nel caso precedente, tutti gli esponenti m_k in (1.6) sono uguali a 1 e l'equazione (1.10) prende la forma

$$r(\lambda_k) = f(\lambda_k) \quad (k = 1, 2, \ldots, m).$$

$r(\lambda)$ é ancora il polinomio di interpolazione ordinario di Lagrange e

$$f(A) = \sum_{k=1}^{m} \frac{(A - \lambda_1 I) \cdots (A - \lambda_{k-1} I)(A - \lambda_{k+1} I) \ldots (A - \lambda_m I)}{(\lambda_k - \lambda_1) \cdots (\lambda_k - \lambda_{k-1})(\lambda_k - \lambda_{k-1}) \ldots (\lambda_k - \lambda_m)} f(\lambda_k).$$

3) Consideriamo ora il caso generale

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_s)^{m_s} \quad (m_1 + m_2 + \ldots + m_s = m).$$

Rappresentiamo la funzione razionale $\dfrac{r(\lambda)}{\psi(\lambda)}$, dove il grado di $r(\lambda)$ é minore del grado di $\psi(\lambda)$, come una somma di frazioni semplici:

$$\frac{r(\lambda)}{\psi(\lambda)} = \sum_{k=1}^{s} \left[\frac{\alpha_{k1}}{(\lambda - \lambda_k)^{m_k}} + \frac{\alpha_{k2}}{(\lambda - \lambda_k)^{m_k - 1}} + \ldots + \frac{\alpha_{k,m_k}}{\lambda - \lambda_k} \right], \qquad (1.11)$$

dove α_{kj} $(j = 1, 2, \ldots, m_k\ ;\ k = 1, 2, \ldots, s)$ sono delle costanti.
Per determinare i numeratori α_{kj} delle frazioni semplici, moltiplichiamo i due membri della (1.11) per $(\lambda - \lambda_k)^{m_k}$ e poniamo:

$$\psi^k(\lambda) = \frac{\psi(\lambda)}{(\lambda - \lambda_k)^{m_k}}.$$

Cosí otteniamo:

$$\frac{r(\lambda)}{\psi^k(\lambda)} = \alpha_{k1} + \alpha_{k2}(\lambda - \lambda_k) + \cdots + \alpha_{km_k}(\lambda - \lambda_k)^{m_k - 1} + (\lambda - \lambda_k)^{m_k} \varrho(\lambda)$$
$$(k = 1, 2, \ldots, s), \quad (1.12)$$

dove $\varrho(\lambda)$ é una funzione razionale regolare per $\lambda = \lambda_k$[2]. Da cui, per lo sviluppo della formula di Taylor,

$$\alpha_{k1} = \left[\frac{r(\lambda)}{\psi^k(\lambda)} \right]_{\lambda = \lambda_k},$$

$$\alpha_{k2} = \left[\frac{r(\lambda)}{\psi^k(\lambda)} \right]'_{\lambda = \lambda_k} = r(\lambda_k) \left[\frac{1}{\psi^k(\lambda)} \right]'_{\lambda = \lambda_k} + r'(\lambda_k) \frac{1}{\psi^k(\lambda_k)}, \ldots (k = 1, 2, \ldots, s). (1.13)$$

Le formule (1.13) mostrano che le costanti α_{kj} possono essere espresse in funzione dei valori del polinomio $r(\lambda)$ sullo spettro di A. Questi valori sono noti: essi sono uguali ai valori corrispondenti della funzione $f(\lambda)$ e delle sue derivate. Di conseguenza,

$$\alpha_{k1} = \frac{f(\lambda_k)}{\psi^k(\lambda_k)},\ \alpha_{k2} = f(\lambda_k) \left[\frac{1}{\psi^k(\lambda)} \right]'_{\lambda = \lambda_k} + f'(\lambda_k) \frac{1}{\psi^k(\lambda_k)}, \ldots (k = 1, 2, \ldots, s). (1.14)$$

[2] Ovvero questa funzione non tende all'infinito per $\lambda \to \lambda_k$.

Le formule (1.14) possono essere scritte come segue:

$$\alpha_{kj} = \frac{1}{(j-1)!}\left[\frac{f(\lambda)}{\psi^k(\lambda)}\right]^{(j-1)}_{\lambda=\lambda_k} (j = 1, 2, \ldots, m_k;\ \ k = 1, 2, \ldots, s). \tag{1.15}$$

Una volta che tutte le α_{kj} sono ottenute, possiamo determinare $r(\lambda)$ con l'aiuto della formula seguente, ottenuta moltiplicando i due membri di (1.11) per $\psi(\lambda)$,

$$r(\lambda) = \sum_{k=1}^{n}[\alpha_{k1} + \alpha_{k2}(\lambda - \lambda_k) + \ldots + \alpha_{k,m_k}(\lambda - \lambda_k)^{m_k-1}]\psi^k(\lambda). \tag{1.16}$$

Ricordiamo che in questa formula l'espressione tra parentesi é uguale alla somma dei primi m_k termini dello sviluppo di Taylor di $f(\lambda)$ in potenze di $(\lambda - \lambda_k)$. Ora trattiamo il secondo metodo.

1.0.3 Matrici costituenti

Ritorniamo ora alla formula (1.16) per $r(\lambda)$. Sostituendo in (1.16) le espressioni (1.14) dei coefficienti α e, combinando i termini che contengono uno e un solo valore delle funzione $f(\lambda)$ o una delle sue derivate, possiamo rappresentare $r(\lambda)$ nella forma

$$r(\lambda) = \sum_{k=1}^{s}[f(\lambda_k)\varphi_{k1}(\lambda) + f'(\lambda_k)\varphi_{k2}(\lambda) + \ldots + f^{(m_k-1)}(\lambda_k)\varphi_{k,m_k}(\lambda)], \tag{1.17}$$

dove

$$\varphi_{kj}(\lambda) \quad (j = 1, 2, \ldots, m_k;\ k = 1, 2, \ldots, s)$$

sono dei polinomi in λ di grado inferiore a m. Questi polinomi sono completamente determinati una volta che $\psi(\lambda)$ é data e non dipendono dalla scelta della funzione $f(\lambda)$. Il numero di questi polinomi é uguale al numero dei valori della funzione $f(\lambda)$ sullo spettro di A, cioé m.

Le funzioni $\varphi_{kj}(\lambda)$ rappresentano il polinomio di interpolazione di Lagrange-Sylvester per una funzione i cui i valori sullo spettro di A sono tutti nulli, a eccezione di $f^{(j-1)}(\lambda_k)$ che é uguale 1.

Si dimostra che (noi non lo faremo) tutti i polinomi

$$\varphi_{kj}(\lambda) \quad (j = 1, 2, \ldots, m_k;\ k = 1, 2, \ldots, s)$$

sono linearmente indipendenti.
Da (1.17) risulta la *formula fondamentale per $f(A)$*:

$$f(A) = \sum_{k=1}^{s} [f(\lambda_k)Z_{k1} + f'(\lambda_k)Z_{k2} + \ldots + f^{(m_k-1)}(\lambda_k)Z_{km_k}] \tag{1.18}$$

dove

$$Z_{kj} = \varphi_{kj}(A) \quad (j = 1, 2, \ldots, m_k; \;\; k = 1, 2, \ldots, s). \tag{1.19}$$

Le matrici Z_{kj} sono completamente determinate quando A é data e non dipendono dalla scelta della funzione $f(\lambda)$. Nel membro di destra di (1.18), ritroviamo i valori della funzione $f(\lambda)$ sullo spettro della matrice A.
Le matrici Z_{kj} sono chiamate *matrici costituenti* della matrice data A.

Proposizione 1.0.1. *Le matrici costituenti sono linearmente indipendenti.*

Dimostrazione. Supponiamo che

$$\sum_{k=1}^{s} \sum_{j=1}^{m_k} c_{kj} Z_{kj} = 0.$$

Poniamo

$$\chi(\lambda) = \sum_{k=1}^{s} \sum_{j=1}^{m_k} c_{kj} \varphi_{kj} \tag{1.20}$$

quindi, per (1.19),

$$\chi(A) = 0. \tag{1.21}$$

Poiché da (1.20) emerge che il grado di $\chi(\lambda)$ é inferiore a m, segue da (1.21) che: $\chi(\lambda) = 0$[3]. Ma, poiché le m funzioni $\varphi_{kj}(\lambda)$ sono linearmente indipendenti, (1.20) implica che: $c_{kj} = 0 \quad (j = 1, 2, \ldots, m_k; \; k = 1, 2, \ldots, s)$. □

[3]Se no esiterebbe un polinomio annullatore di grado inferiore a quello del polinomio minimo.

Dall'indipendenza lineare delle matrici costituenti Z_{kj}, risulta che nessuna di queste matrici puó essere nulla. Ora vediamo come poter determinare le matrici costituenti.

Quando la matrice A é data possiamo porre, nella formula fondamentale, $f(\mu) = \frac{1}{\lambda - \mu}$ dove λ é un parametro. Allora otteniamo

$$(\lambda I - A)^{-1} = \frac{C(\lambda)}{\psi(\lambda)} = \sum_{k=1}^{s} \left[\frac{Z_{k1}}{\lambda - \lambda_k} + \frac{1!Z_{k2}}{(\lambda - \lambda_k)^2} + \ldots + \frac{(m_k - 1)!Z_{km_k}}{(\lambda - \lambda_k)^{m_k}} \right] \quad (1.22)$$

dove $C(\lambda)$ é la matrice aggiunta ridotta di $\lambda I - A$. Le matrici $(j-1)!Z_{kj}$ sono i numeratori della frazioni semplici della decomposizione (1.22) e, per analogia con (1.11), esse possono essere espresse con l'aiuto dei valori di $C(\lambda)$ sullo spettro di A attraverso formule simili alle formule (1.13):

$$(m_k - 1)!Z_{km_k} = \frac{C(\lambda_k)}{\psi^k(\lambda_k)}, \quad (m_k - 2)!Z_{k,m_k-1} = \left[\frac{C(\lambda)}{\psi^k(\lambda)} \right]'_{\lambda=\lambda_k}, \ldots \ ,$$

da cui

$$Z_{kj} = \frac{1}{(j-1)!(m_k - j)!} \left[\frac{C(\lambda)}{\psi^k(\lambda)} \right]^{(m_k - j)}_{\lambda=\lambda_k} \quad (j = 1, 2, \ldots, m_k;\ k = 1, 2, \ldots, s). \quad (1.23)$$

Sostituendo in (1.18) le matrici costituenti con le loro espressioni (1.23), possiamo rappresentare la formula fondamentale sotto la forma:

$$f(A) = \sum_{k=1}^{s} \frac{1}{(m_k - 1)!} \left[\frac{C(\lambda)}{\psi^k(\lambda)} f(\lambda) \right]^{(m_k - 1)}_{\lambda=\lambda_k}. \quad (1.24)$$

Ora trattiamo il terzo e ultimo metodo.

1.0.4 Sostituzione di polinomi

Iniziamo con un esempio che ci mostra come determinare $f(A)$ partendo semplicemente dalla formula fondamentale.

Esempio 1.0.1. $A = \left\| \begin{array}{rrr} 2 & -1 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{array} \right\|, \quad \psi(\lambda) = (\lambda - 1)^2(\lambda - 2).$

Allora

$$f(A) = f(1)Z_1 + f'(1)Z_2 + f(2)Z_3. \quad (1.25)$$

Sostituendo in successione $1, \lambda - 1, (\lambda - 1)^2$ a $f(\lambda)$ in (1.20) otteniamo:

$$Z_1 + Z_3 = I = \left\| \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \right\|,$$

$$Z_2 + Z_3 = A - I = \left\| \begin{matrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{matrix} \right\|,$$

$$Z_3 = (A - I)^2 = \left\| \begin{matrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{matrix} \right\|.$$

Cosí possiamo determinare tutte le Z calcolando prima Z_3 nella terza equazione e poi riportandolo nelle prime due. Sostituendo in (1.20), otteniamo l'espressione di $f(A)$.

Questo terzo metodo é forse quello che conviene maggiormente usare nella pratica. In generale esso puó essere descritto come segue:
In (1.18) sostituiamo a $f(\lambda)$, in successione, dei polinomi $g_1(\lambda), g_2(\lambda), \ldots, g_m(\lambda)$:

$$g_i(A) = \sum_{k=1}^{n} [g_i(\lambda_k) Z_{k1} + g_i'(\lambda_k) Z_{k2} + \ldots + g_i^{(m_k-1)}(\lambda_k) Z_{k,m_k}] \quad (i = 1, 2, \ldots, m). \tag{1.26}$$

A partire dalle m equazioni (1.26), determiniamo le matrici Z_{kj} e sostituiamo le espressioni cosí ottenute in (1.18). Il risultato dell'eliminazione di Z_{kj} dalle $m + 1$ equazioni (1.26) e (1.18) puó essere scritto sotta la forma

$$\begin{vmatrix} f(A) & f(\lambda_1) & \ldots & f^{(m_1-1)}(\lambda_1) & \ldots & f(\lambda_s) & \ldots & f^{(m_s-1)}(\lambda_s) \\ g_1(A) & g_1(\lambda_1) & \ldots & g_1^{(m_1-1)}(\lambda_1) & \ldots & g_1(\lambda_s) & \ldots & g_1^{(m_s-1)}(\lambda_s) \\ \vdots & \vdots & & & & & & \\ g_m(A) & g_m(\lambda_1) & \ldots & g_m^{(m_1-1)}(\lambda_1) & \ldots & g_m(\lambda_s) & \ldots & g_m^{(m_s-1)}(\lambda_s) \end{vmatrix} = 0.$$

Sviluppando questo determinante secondo la prima colonna, otteniamo l'espressione cercata di $f(A)$. Infatti, calcolando il determinante, abbiamo

$f(A)$ che moltiplica il suo complemento algebrico che non é altro che il determinante

$$\Delta = |g_i^{(j)}(\lambda_k)|$$

(alla i-esima riga di Δ troviamo i valori del polinomio $g_i(\lambda)$ sullo spettro di A, con $i = 1, 2, \ldots, m$). Per determinare $f(A)$, bisogna che $\Delta \neq 0$.
Sará cosí se nessuna combinazione lineare dei polinomi $g_1(\lambda), g_2(\lambda), \ldots, g_m(\lambda)$ é identicamente nulla sullo spettro di A, cioé se nessuna combinazione lineare é divisibile per $\psi(\lambda)$. La condizione $\Delta \neq 0$ é sempre soddisfatta quando i gradi dei polinomi $g_1(\lambda), g_2(\lambda), \ldots, g_m(\lambda)$ sono rispettivamente $0, 1, \ldots, m-1$.

Capitolo 2

In questo capitolo applicheremo le nozioni date nel primo capitolo ai sistemi di equazioni differenziali. In particolare vedremo due diversi tipi di sistemi.

2.0.5 Sistema di equazioni differenziali omogenee

Consideriamo un sistema di equazioni differenziali lineari omogenee del primo ordine a coefficenti costanti:

$$\begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n \quad , \\ \frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n \quad , \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \frac{dx_n}{dt} &= a_{n1}x_1 + a_{n2}x_2 + ... + a_{nn}x_n \quad ; \end{aligned} \tag{2.1}$$

dove t é la variabile indipendente e dove $x_1, x_2, ..., x_n$ sono delle funzioni e $a_{ik}(i, k = 1, 2, ..., n)$ dei numeri complessi. Introduciamo la matrice quadrata $A = ||a_{ik}||_1^n$ dei coefficenti e la matrice colonna $x = (x_1, x_2, ..., x_n)$. Il sistema puó allora essere scritto in forma matriciale:

$$\frac{dx}{dt} = Ax \quad , \tag{2.2}$$

dove $\frac{dx}{dt}$ é la matrice colonna avente per elementi $\frac{dx_1}{dt}, \frac{dx_2}{dt}, ..., \frac{dx_n}{dt}$.
Noi cerchiamo una soluzione del sistema di equazioni differenziali che soddisfa le seguenti condizioni iniziali:

$$x_1|_{t=0} = x_{10}, x_2|_{t=0} = x_{20}, ..., x_n|_{t=0} = x_{n0} \quad , \tag{2.3}$$

o piú brevemente $x|_{t=0} = x_0$. Sviluppiamo la colonna delle incognite x in serie di MacLaurin secondo le potenze di t:

$$x = x_0 + \dot{x}_0 t + \ddot{x}_0 \frac{t^2}{2!} + ... \quad \left(\dot{x}_0 = \left. \frac{dx}{dt} \right|_{t=0}, \ddot{x}_0 = \left. \frac{d^2x}{dt^2} \right|_{t=0}, ... \right). \tag{2.4}$$

Derivando ulteriormente (2.2) otteniamo:

$$\frac{d^2x}{dt^2} = A\frac{dx}{dt} = A^2x, \quad \frac{d^3x}{dt^3} = A\frac{d^2x}{dt^2} = A^3x, ... \tag{2.5}$$

Prendendo $t = 0$ in (2.2) e (2.5) abbiamo:

$$\dot{x}_0 = Ax_0, \quad \ddot{x}_0 = A^2x_0, ...$$

cosí la serie (2.4) puó essere scritta:

$$x = x_0 + tAx_0 + \frac{t^2}{2!}A^2x_0 + ... = e^{At}x_0. \tag{2.6}$$

Per sostituzione diretta in (2.2) vediamo che (2.6) é una soluzione dell'equazione differenziale, infatti:

$$\frac{d(e^{At})}{dt} = \frac{d}{dt}(I + At + \frac{A^2t^2}{2!} + ...) = A + A^2t + \frac{A^3t^2}{2!} + ... = Ae^{At}.$$

Ponendo $t = 0$ in (2.6) otteniamo $x|_{t=0}$. La formula (2.6) ci dá quindi la soluzione del sistema di equazioni differenziali che soddisfa le condizioni iniziali (2.3). Poniamo $f(\lambda) = e^{\lambda t}$. Allora,

$$e^{At} = ||q_{ik}(t)||_1^n = \sum_{k=1}^{s}(Z_{k1} + Z_{k2}t + ... + Z_{km_k}t^{m_k-1})e^{\lambda_k t}. \tag{2.7}$$

La soluzione (2.6) puó allora essere scritta sotto la forma

$$\begin{aligned} x_1 &= q_{11}(t)x_{10} + q_{12}(t)x_{20} + ... + q_{1n}(t)x_{n0} \quad , \\ x_2 &= q_{21}(t)x_{10} + q_{22}(t)x_{20} + ... + q_{2n}(t)x_{n0} \quad , \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n &= q_{n1}(t)x_{10} + q_{n2}(t)x_{20} + ... + q_{nn}(t)x_{n0} \quad ; \end{aligned} \tag{2.8}$$

dove $x_{10}, x_{20}, ..., x_{n0}$ sono della costanti uguali ai valori iniziali delle funzioni incognite $x_1, x_2, ..., x_n$. Cosí l'integrazione del sistema (2.1) di equazioni differenziali si riduce al calcolo degli elementi della matrice e^{At}.
Se $t = t_0$ é il valore iniziale dell'argomento, (2.6) deve essere sostituito dalla formula:

$$x = e^{a(t-t_0)}x_0. \tag{2.9}$$

Esempio 2.0.2.

$$\frac{dx_1}{dt} = 3x_1 - x_2 + x_3 \quad ,$$
$$\frac{dx_2}{dt} = 2x_1 + x_3 \quad ,$$
$$\frac{dx_3}{dt} = x_1 - x_2 + 2x_3 \quad .$$

La matrice dei coefficenti é

$$A = \left\| \begin{matrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{matrix} \right\|.$$

Calcoliamo il polinomio caratteristico:

$$|\lambda E - A| = -|A - \lambda E| = \Delta(\lambda) = - \begin{vmatrix} 3-\lambda & -1 & 1 \\ 2 & -\lambda & 1 \\ 1 & -1 & 2-\lambda \end{vmatrix} = (\lambda - 1)(\lambda - 2)^2.$$

Il massimo comune divisore dei minori di ordine 2 é $D_2(\lambda) = 1$. Il polinomio caratteristico coincide con il polinomio minimo per cui ho:

$$\Psi(\lambda) = \Delta(\lambda) = (\lambda - 1)(\lambda - 2)^2.$$

La formula fondamentale si scrive: $f(A) = f(1)Z_1 + f(2)Z_2 + f'(2)Z_3$.
Ora utilizziamo il terzo metodo, visto nel primo capitolo, per determinare le matrici costituenti Z_1, Z_2, Z_3. Per $f(\lambda)$ consideriamo in successione $1, \lambda - 2, (\lambda - 2)^2$ e, andando a sostituire i tre rispettivi valori di $f(\lambda)$ nella

formula fondamentale, abbiamo:

$$Z_1 + Z_2 = I = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix} \quad ,$$

$$-Z_1 + Z_3 = A - 2I = \begin{Vmatrix} 1 & -1 & 1 \\ 2 & -2 & 1 \\ 1 & -1 & 0 \end{Vmatrix} \quad ,$$

$$Z_1 = (A-2I)^2 = \begin{Vmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{Vmatrix} \quad .$$

Da quí otteniamo Z_1, Z_2, Z_3 e sostituendo nella formula fondamentale ricaviamo $f(A)$:

$$f(A) = f(1) \begin{Vmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{Vmatrix} + f(2) \begin{Vmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{Vmatrix} + f'(2) \begin{Vmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{Vmatrix}.$$

Se sostituiamo $f(\lambda)$ con $e^{\lambda t}$ otteniamo:

$$e^{At} = e^t \begin{Vmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{Vmatrix} + e^{2t} \begin{Vmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{Vmatrix} + te^{2t} \begin{Vmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{Vmatrix} =$$

$$= \begin{Vmatrix} (1+t)e^{2t} & -te^{2t} & te^{2t} \\ -e^t + (1+t)e^{2t} & e^t - te^{2t} & te^{2t} \\ -e^t + e^{2t} & e^t - e^{2t} & e^{2t} \end{Vmatrix}$$

da cui

$$x_1 = C_1(1+t)e^{2t} - C_2 te^{2t} + C_3 te^{2t} \quad ,$$

$$x_2 = C_1[-e^t + (1+t)e^{2t}] + C_2(e^t - te^{2t}) + C_3 te^{2t} \quad ,$$

$$x_3 = C_1(-e^t + e^{2t}) + C_2(e^t - e^{2t}) + C_3 e^{2t} \quad ,$$

dove

$$C_1 = x_{10} \quad , C_2 = x_{20} \quad , C_3 = x_{30}.$$

2.0.6 Sistema di equazioni differenziali non omogenee

Consideriamo ora un sistema di equazioni differenziali lineari non omogenee a coefficenti costanti:

$$\begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n + f_1(t) \quad , \\ \frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n + f_2(t) \quad , \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \frac{dx_n}{dt} &= a_{n1}x_1 + a_{n2}x_2 + ... + a_{nn}x_n + f_n(t) \quad ; \end{aligned} \tag{2.10}$$

dove $f_1(t), f_2(t), ..., f_n(t)$ sono delle funzioni continue nell'intervallo $[t_0, t_1]$. Indicando con $f(t)$ la matrice colonna $(f_1(t), f_2(t), ..., f_n(t))$ e ponendo ancora $A = ||a_{ik}||_1^n$, scriviamo il sistema nella forma:

$$\frac{dx}{dt} = Ax + f(t). \tag{2.11}$$

Sostituiamo x con la nuova colonna z di funzioni incognite collegate a x dalla relazione

$$x = e^{At}z. \tag{2.12}$$

Derivando (2.12) termine a termine otteniamo: $\frac{dx}{dt} = Ae^{At}z + e^{At}\frac{dz}{dt}$ e sostituendo l'espressione di $\frac{dx}{dt}$ della (2.11) ricaviamo:

$$e^{At}\frac{dz}{dt} = f(t), \tag{2.13}$$

da cui:

$$z(t) = c + \int_{t_0}^{t} e^{-A\tau} f(\tau) d\tau. \tag{2.14}$$

Sostituendo l'espressione (2.14) nella (2.12) otteniamo:

$$x = e^{At}\left[c + \int_{t_0}^{t} e^{-A\tau} f(\tau) d\tau\right] = e^{At}c + \int_{t_0}^{t} e^{A(t-\tau)} f(\tau) d\tau \quad , \tag{2.15}$$

dove c é una colonna di elementi costanti arbitrari.
Dando, in (2.15), il valore t_0 all'argomento otteniamo: $c = e^{-At}x_0$, in modo che la (2.15) possa essere scritta sotto la forma

$$x = e^{A(t-t_0)}x_0 + \int_{t_0}^{t} e^{A(t-\tau)}f(\tau)d(\tau). \tag{2.16}$$

Ponendo $e^{At} = ||q_{ij}||_1^n$, possiamo scrivere la soluzione (2.16) sotto la forma estesa:

$$\begin{aligned} x_1 &= q_{11}(t-t_0)x_{10} + ... + q_{1n}(t-t_0)x_{n0} + \int_{t_0}^{t} \left[q_{11}(t-\tau)f_1(\tau) + ... + q_{1n}(t-\tau)f_n(\tau)\right] d(\tau) \quad , \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots , \\ x_n &= q_{n1}(t-t_0)x_{10} + ... + q_{nn}(t-t_0)x_{n0} + \int_{t_0}^{t} \left[q_{n1}(t-\tau)f_1(\tau) + ... + q_{nn}(t-\tau)f_n(\tau)\right] d(\tau) \quad . \end{aligned} \tag{2.17}$$

Esempio 2.0.3. Come esempio consideriamo il movimento di un punto materiale nel vuoto, vicino alla superficie della terra, tenendo conto del movimento della terra. Noi sappiamo che, in questo caso, l'accellerazione del punto relativa alla terra é determinata dalla forza peso (costante) mg e dalla forza d'inerzia di Coriolis $-2m\omega$ x v dove v é la velocitá del punto relativa alla terra e ω é la velocitá angolare costante della terra. Di conseguenza, l'equazione differenziale del movimento del punto é della forma:

$$a = \frac{dv}{dt} = g - 2\omega \times v. \tag{2.18}$$

Definiamo un operatore lineare A nello spazio euclideo di dimensione tre dall'equazione

$$Ax = -2\omega \times x \tag{2.19}$$

e scriviamo, al posto di (2.18),

$$\frac{dv}{dt} = Av + g. \tag{2.20}$$

Confrontando (2.20) e (2.11), segue direttamente da (2.16) che:

$$v = e^{At}v_0 + \int_0^t e^{As}ds \cdot g \qquad (v_0 = v_{t=0}).$$

Integrando termine a termine otteniamo:

$$\int_0^t v\ ds = \int_0^t e^{As} ds \cdot v_0 + \int_0^t ds \int_0^q e^{A\rho} d\rho \cdot g\ ,$$

cosí determiniamo il vettore polare del movimento del punto

$$r = r_0 + \int_0^t e^{As} ds \cdot v_0 + \int_0^t ds \int_0^q e^{A\rho} d\rho \cdot g\ , \tag{2.21}$$

dove $r_0 = r_{t=0}$ e $v_0 = v_{t=0}$.
Sostituendo a e^{At} la serie

$$I + A\frac{t}{1!} + A^2\frac{t^2}{2!} + ...$$

e sostituendo A con la sua espressione data in (2.19), abbiamo:

$$\int_0^t I ds \cdot v_0 + \int_0^t ds \int_0^q I d\rho \cdot g = v_0 \int_0^t ds + g \int_0^t s\ ds = v_0 t + \frac{1}{2} g t^2\ ,$$

$$\int_0^t As\ ds \cdot v_0 + \int_0^t ds \int_0^q A\rho\ d\rho \cdot g = A\big(\frac{1}{2}t^2 v_0\big) + g\int_0^t \frac{A}{2}s^2\ ds = A\big(\frac{1}{2}t^2 v_0\big) + \frac{A}{6}t^3 g =$$
$$= \frac{1}{2}A\big(t^2 v_0 + \frac{1}{3}t^3 g\big) = \frac{1}{2}\big(- 2\omega \times \big(t^2 v_0 + \frac{1}{3}t^3 g\big)\big)\ .$$

Continuando analogamente ricaviamo che

$$r = r_0 + v_0 t + \frac{1}{2}gt^2 - \omega \times \left(v_0 t^2 + \frac{1}{3}gt^3\right) + \omega \times \left[\omega \times \left(\frac{2}{3}v_0 t^3 + \frac{1}{6}gt^4\right)\right] + ...$$

Considerando la velocitá angolare ω piccola (per la terra $\omega \simeq 7,3 \cdot 10^{-5} s^{-1}$), possiamo tralasciare i termini che contengono le potenze di ω superiori o uguali a due; per lo spostamento addizionale del punto dovuto al movimento della terra, otteniamo allora la formula approssimata

$$d \simeq -\omega \times \left(v_0 t^2 + \frac{1}{3}gt^3\right).$$

Ritornando alla soluzione esatta (2.21), calcoliamo e^{At}. Prima di tutto stabiliamo che il polinomio minimale dell'operatore A é della forma

$$\Psi(\lambda) = \lambda(\lambda^2 + 4\omega^2).$$

Infatti, a partire da (2.19), otteniamo:

$$A^2x = 4\,\omega \times (\omega \times x) = 4(\omega x)\omega - 4\,\omega^2 x \quad ,$$

$$A^3x = -2\,\omega \times A^2x = 8\,\omega^2(\omega \times x) \quad .$$

Da ció segue , con (2.19), che gli operatori I, A, A^2 sono linearmente indipendenti e che $A^3 + 4\omega^2 A = 0$.

Il polinomio minimo $\Psi(\lambda)$ ha come radici semplici $0, 2\omega i, -2\omega i$. La formula di interpolazione di Lagrange per e^{At} é della forma

$$1 + \frac{sen2\omega t}{2\omega}\lambda + \frac{1 - cos2\omega t}{4\omega^2}\lambda^2.$$

Allora

$$e^{At} = I + \frac{sen2\omega t}{2\omega}A + \frac{1 - cos2\omega t}{4\omega^2}A^2.$$

Sostituendo questa espressione ad e^{At} in (2.21) e mettendo al posto dell'operatore A la sua espressione data in (2.19), otteniamo:

$$\begin{aligned} r = r_0 + v_0 t + g\frac{t^2}{2} - \omega \times \left(\frac{1 - cos2\omega t}{2\omega^2}v_0 + \frac{2\omega t - sen2\omega t}{4\omega^3}g\right) + \\ + \omega \times \left[\omega \times \left(\frac{2\omega t - sen2\omega t}{2\omega^3}v_0 + \frac{-1 + 2\omega^2 t^2 + cos2\omega t}{4\omega^4}g\right)\right]. \end{aligned} \tag{2.22}$$

Ora consideriamo il caso particolare in cui $v_0 = 0$: sviluppando il triplo prodotto vettoriale otteniamo

$$r = r_0 + g\frac{t^2}{2} + \frac{2\omega t - sen2\omega t}{4\omega^3}(g \times \omega) + \frac{cos2\omega t - 1 + 2\omega^2 t^2}{4\omega^3}(gsen\varphi\omega - \omega g) \quad ,$$

dove φ é la latitudine geografica del punto di cui noi consideriamo il movimento. Il termine

$$\frac{2\omega t - sen2\omega t}{4\omega^3}(g \times \omega)$$

rappresenta lo spostamento verso est, perpendicolare al piano del meridiano, e l'ultimo termine a destra dell'ultima formula ci dice lo spostamento nel piano meridiano perpendicolare all'asse della terra.

Capitolo 3

Questo capitolo sará un capitolo introduttivo che ci permetterá, nel capitolo 4, di trattare il caso di un sistema di m equazioni differenziali del primo ordine a coefficenti costanti e n funzioni incognite:

$$Ax + B\frac{dx}{dt} = f(t) \ ,$$

dove B puó non essere invertibile, questa é la novitá di questo caso.
Come prima cosa diamo una serie di definizioni:

Definizione 3.0.6. *Una matrice polinomiale, o λ-matrice, é una matrice rettangolare $A(\lambda)$ i cui elementi sono dei polinomi in λ:*

$$A(\lambda) = ||a_{ik}(\lambda)|| = ||a_{ik}^{(0)}\lambda^l + a_{ik}^{(1)}\lambda^{l-1} + \ldots + a_{ik}^{(l)}|| \quad \begin{pmatrix} i = 1, 2, \ldots, m; \\ k = 1, 2, \ldots, n; \end{pmatrix}$$

dove l é il grado piú alto dei polinomi $a_{ik}(\lambda)$.

Ponendo

$$A_j = ||a_{ik}^{(j)}|| \quad (i = 1, 2, \ldots, m; \ \ k = 1, 2, \ldots, n; \ \ j = 0, 1, \ldots, l),$$

possiamo rappresentare la matrice polinomiale $A(\lambda)$ sotto la forma di un polinomio matriciale in λ, cioé sotto la forma di un polinomio in λ a coefficenti matriciali: $A(\lambda) = A_0\lambda^l + A_1\lambda^{l-1} + \ldots + A_{l-1}\lambda + A_l$.

Definizione 3.0.7. *Due matrici rettangolari $A(\lambda)$ e $B(\lambda)$ sono dette equivalenti se*

$$B(\lambda) = P(\lambda)A(\lambda)Q(\lambda),$$

dove $P(\lambda)$ e $Q(\lambda)$ sono delle matrici quadrate polinomiali con dei determinanti costanti non nulli.

Ora introduciamo il concetto di *polinomio invariante.*
Sia $A(\lambda)$ una matrice polinomiale di rango r. Indichiamo con $D_j(\lambda)$ il piú grande divisore comune di tutti i minori d'ordine j di $A(\lambda)$ ($j = 1, 2, \dots, r$). Allora ciascun polinomio $D_r(\lambda), D_{r-1}(\lambda), \dots, D_1(\lambda), D_0(\lambda) \equiv 1$ é divisibile per il polinomio precedente. I quozienti corrispondenti li indicheremo con $i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda)$:

$$i_1(\lambda) = \frac{D_r(\lambda)}{D_{r-1}(\lambda)}, \quad i_2(\lambda) = \frac{D_{r-1}(\lambda)}{D_{r-2}(\lambda)}, \quad \dots, i_r(\lambda) = \frac{D_1(\lambda)}{D_0(\lambda)} = D_1(\lambda).$$

Definizione 3.0.8. *I polinomi $i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda)$ definiti sopra sono chiamati i polinomi invarianti della matrice rettangolare $A(\lambda)$.*

Teorema 3.0.2. *Se, in una matrice rettangolare quasi diagonale*

$$C(\lambda) = \begin{pmatrix} A(\lambda) & 0 \\ 0 & B(\lambda) \end{pmatrix}$$

ogni polinomio invariante di $A(\lambda)$ divide ogni polinomio invariante di $B(\lambda)$, l'insieme dei polinomi invarianti di $C(\lambda)$ é l'unione dei polinomi invarianti di $A(\lambda)$ e di $B(\lambda)$.

Per determinare i polinomi invarianti di $C(\lambda)$, nel caso generale di polinomi invarianti arbitrari di $A(\lambda)$ e di $B(\lambda)$, utilizziamo il concetto importante di divisore elementare. Decomponiamo i polinomi invarianti $i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda)$ in fattori irriducibili sul campo di numeri $\mathcal{K}$:

$$\begin{aligned}
&i_1(\lambda) = [\varphi_1(\lambda)]^{c_1}[\varphi_2(\lambda)]^{c_2} \cdots [\varphi_s(\lambda)]^{c_s}, \\
&i_2(\lambda) = [\varphi_1(\lambda)]^{d_1}[\varphi_2(\lambda)]^{d_2} \cdots [\varphi_s(\lambda)]^{d_s}, \quad c_k \geq d_k \geq \dots \geq l_k \geq 0; \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \qquad k = 1, 2, \dots, s, \\
&i_r(\lambda) = [\varphi_1(\lambda)]^{l_1}[\varphi_2(\lambda)]^{l_2} \cdots [\varphi_s(\lambda)]^{l_s}
\end{aligned}$$

dove $\varphi_1(\lambda), \varphi_2(\lambda), \dots, \varphi_s(\lambda)$ sono tutti i fattori distinti su $\mathcal{K}$ che compaiono in $i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda)$.

Definizione 3.0.9. *Tutte le potenze* $[\varphi_1(\lambda)]^{c_1}, \ldots, [\varphi_s(\lambda)]^{l_2}$, *sono chiamate divisori elementari della matrice* $A(\lambda)$ *sul campo* $\mathcal{K}$.

Teorema 3.0.3. *L'insieme dei divisori elementari della matrice quasi diagonale rettangolare*

$$C(\lambda) = \begin{pmatrix} A(\lambda) & 0 \\ 0 & B(\lambda) \end{pmatrix}$$

é ottenuto combinando i divisori elementari di $A(\lambda)$ *con quelli di* $B(\lambda)$.

A questo punto possiamo occuparci dell'obbiettivo principale di questo capitolo: essendo date quattro matrici A, B, A_1, B_1, tutte di dimensioni $m \times n$ a elementi appartenenti ad un campo $\mathcal{K}$, dobbiamo precisare sotto quali condizioni esistono due matrici quadrate e regolari, P e Q, d'ordine rispettivo m e n, tali che

$$PAQ = A_1, \quad PBQ = B_1. \tag{3.1}$$

Introducendo i fasci di matrici $A + \lambda B$ e $A_1 + \lambda B_1$ le due equazioni matriciali (3.1) possono essere sostituite dalla unica equazione

$$P(A + \lambda B)Q = A_1 + B_1. \tag{3.2}$$

Definizione 3.0.10. *Due fasci di matrici rettangolari* $A+\lambda B$ *e* $A_1+\lambda B_1$ *della stessa dimensione* $m\times n$*, legati dall'equazione (3.2) sono detti strettamente equivalenti.*

I fasci $A+\lambda B$ e $A_1+\lambda B_1$ sono equivalenti se una equazione della forma (3.2), nella quale P e Q sono due λ-matrici quadrate con determinanti costanti non nulli, esiste. Per l'equivalenza stretta, bisogna in piú che P e Q non dipendano da λ. Un criterio di equivalenza dei fasci $A + \lambda B$ e $A_1 + \lambda B_1$ consiste nell'uguaglianza dei polinomi invarianti, o dei divisori elementari, dei fasci. Quindi in questo capitolo, stabiliremo un criterio di equivalenza stretta di due fasci di matrici e determineremo, per ciascun fascio, una forma canonica strettamente equivalente.
Tutti i fasci di matrici $A + \lambda B$ di dimensione $m \times n$ sono di due tipi: *regolari* o *singolari*.

Definizione 3.0.11. *Un fascio di matrici $A + \lambda B$ é detto regolare se:*

1. *A e B sono delle matrici quadrate dello stesso ordine n.*

2. *Il determinante $|A + \lambda B|$ non si annulla identicamente.*

In tutti gli altri casi ($m \neq n$ o $m = n$ ma $|A + \lambda B| \equiv 0$) il fascio é detto singolare.

3.0.7 Fasci regolari di matrici

Un criterio di equivalenza stretta di fasci regolari e una forma canonica per tali fasci sono stati stabiliti da Weierstrass nel 1867.
Innanzitutto, nel caso speciale in cui i fasci $A+\lambda B$ e $A_1+\lambda B_1$ siano costituiti da matrici quadrate con $|B| \neq 0$ e $|B_1 \neq 0|$ abbiamo che i due concetti di equivalenza ed equivalenza stretta coincidono. In generale vale

Teorema 3.0.4. *Due fasci di matrici quadrate dello stesso ordine $A + \lambda B$ e $A_1 + \lambda B_1$, nei quali $|B| \neq 0$ e $|B_1| \neq 0$, sono strettamente equivalenti se e solamente se questi fasci hanno gli stessi divisori elelementari su $\mathcal{K}$.*

Se, come ipotesi, mettiamo che i due fasci siano regolari il teorema non é piú vero[1], quindi dobbiamo generalizzarlo. Introduciamo, per questo, il concetto di divisori elementari *infiniti* di un fascio.
Definiamo il fascio $A + \lambda B$ con l'aiuto dei parametri λ, μ: $\mu A + \lambda B$. Determinando il massimo comun divisore $D_k(\lambda, \mu)$ di tutti i minori di ordine k della matrice $\mu A + \lambda B$ con $k = 1, 2, \ldots, n$, otteniamo i polinomi invarianti:

$$i_1(\lambda, \mu) = \frac{D_n(\lambda, \mu)}{D_{n-1}(\lambda, \mu)}, \quad i_2(\lambda, \mu) = \frac{D_{n-1}(\lambda, \mu)}{D_{n-2}(\lambda, \mu)} \ldots;$$

dove tutti i $D_k(\lambda, \mu)$ e i $i_j(\lambda, \mu)$ sono dei polinomi in λ e μ. Sviluppando i polinomi invarianti in potenze di polinomi irriducibili in $\mathcal{K}$, otteniamo i divisori elementari $e_\alpha(\lambda, \mu)$ ($\alpha = 1, 2, \ldots$) del fascio $\mu A + \lambda B$ su $\mathcal{K}$.
Se poniamo $\mu = 1$ in $e_\alpha(\lambda, \mu)$ ritroviamo i divisori elementari $e_\alpha(\lambda)$ del fascio

[1] É infatti possibile avere $|B| = 0$ e anche $|A| = |B| = 0$ in un fascio regolare.

$A + \lambda B$; inversamente, a partire da un divisore elementare $e_\alpha(\lambda)$ di grado q del fascio $A + \lambda B$, otteniamo il divisore elementare corrispondente $e_\alpha(\lambda.\mu)$ attraverso la formula $e_\alpha(\lambda, \mu) = \mu^q e_\alpha(\lambda)$.
Possiamo cosí ottenere tutti i divisori elementari del fascio $\mu A + \lambda B$ con l'eccezione di quelli della forma μ^q. I divisori elementari della forma μ^q esistono se e solamente se $|B| = 0$ e sono chiamati divisori elementari *infiniti* del fascio $A + \lambda B$. Poiché l'equivalenza stretta dei fasci $A + \lambda B$ e $A_1 + \lambda B_1$ implica l'equivalenza stretta dei fasci $\mu A + \lambda B$ e $\mu A_1 + \lambda B_1$, vediamo che, nei fasci strettamente equivalenti $A+\lambda B$ e $A_1+\lambda B_1$, non solo i divisori elementari finiti, ma anche quelli infiniti devono coincidere. Dato che si dimostra anche il viceversa, siamo arrivati alla generalizzazione seguente:

Teorema 3.0.5. *Due fasci regolari $A + \lambda B$ e $A_1 + \lambda B_1$ sono strettamente equivalenti se e solamente se hanno gli stessi divisori elementari finiti e infiniti.*

Per quanto riguarda la forma canonica dei fasci regolari, vale il seguente teorema:

Teorema 3.0.6. *Tutti i fasci regolari $A + \lambda B$ possono essere ridotti ad una forma quasi diagonale canonica (strettamente equivalente)*

$$\{N^{(u_1)}, N^{(u_2)}, \ldots, N^{(u_s)}, J + \lambda I\} \tag{3.3}$$

dove i primi s blocchi diagonali corrispondono ai divisori elementari infiniti $\mu^{n_1}, \mu^{n_2}, \ldots, \mu^{n_s}$ del fascio $A+\lambda B$ e dove la forma normale dell'ultimo blocco diagonale $J + \lambda I$ é determinata in modo unico dai divisori elementari finiti del fascio dato.

3.0.8 Fasci singolari e teorema di riduzione

Consideriamo ora un fascio singolare di matrici $A + \lambda B$ di dimensione $m \times n$ e indichiamo con r il rango del fascio. Dalla singolaritá del fascio segue che $r < n$ o $r < m$. Supponiamo $r < n$. Allora le colonne della λ-matrice $A+\lambda B$

sono linearmente indipendenti, cioé l'equazione

$$(A + \lambda B)x = 0, \tag{3.4}$$

dove x é un vettore colonna incognito, ha una soluzione non nulla. Noi ci limiteremo a esaminare le soluzioni $x(\lambda)$ di (3.4) che sono dei polinomi in λ e, tra queste soluzioni, sceglieremo quella che ha il grado piú piccolo ϵ:

$$x(\lambda) = x_0 - \lambda x_1 + \lambda^2 x_2 - \ldots + (-1)^\epsilon \lambda^\epsilon x_\epsilon \quad (x_\epsilon \neq 0). \tag{3.5}$$

Sostituendo questa soluzione in (3.4) e uguagliando a zero i coefficenti delle potenze di λ, otteniamo

$$Ax_0 = 0,\ Bx_0 - Ax_1 = 0,\ Bx_1 - Ax_2 = 0, \ldots, Bx_{\epsilon-1} - Ax_\epsilon = 0,\ Bx_\epsilon = 0. \tag{3.6}$$

Considerando questo come un sistema di equazioni omogenee lineari per gli elementi delle colonne $x_0, -x_1, x_2, \ldots, (-1)^\epsilon x_\epsilon$, deduciamo che la matrice dei coefficenti del sistema

$$M_\epsilon = M_\epsilon[A + \lambda B] = \begin{Vmatrix} A & 0 & \cdots & 0 \\ B & A & & \vdots \\ 0 & B & \ddots & \vdots \\ \vdots & \vdots & \ddots & A \\ 0 & 0 & \cdots & B \end{Vmatrix} \tag{3.7}$$

é di rango $\rho < (\epsilon + 1)n$. In virtú della proprietá minimale di ϵ, i ranghi $\rho_0, \ldots, \rho_{\epsilon-1}$ delle matrici

$$M_0 = \begin{Vmatrix} A \\ B \end{Vmatrix},\ M_1 = \begin{Vmatrix} A & 0 \\ B & A \\ 0 & B \end{Vmatrix},\ M_{\epsilon-1} = \begin{Vmatrix} A & 0 & \cdots & 0 \\ B & A & & \vdots \\ \vdots & \vdots & \cdots & A \\ 0 & & \cdots & B \end{Vmatrix} \tag{3.8}$$

soddisfano le equazioni $\rho_0 = n,\ \rho_1 = 2n,\ \cdots, \rho_{\epsilon-1} = \epsilon n$.
Quindi il numero ϵ é il piú piccolo valore dell'indice k per il quale si ha il segno $<$ nella relazione $\rho_k \leq (k + 1)n$.
Cosí possiamo formulare il *teorema fondamentale di riduzione* (che non dimostreremo):

Teorema 3.0.7. *Se l'equazione (3.4) ha una soluzione di grado minimo ϵ e se $\epsilon > 0$, il fascio dato $A + \lambda B$ é strettamente equivalente ad un fascio della forma*

$$\left\| \begin{matrix} L_\epsilon & 0 \\ 0 & \widehat{A} + \lambda \widehat{B} \end{matrix} \right\|,$$

dove

$$L_\epsilon = \left\| \begin{matrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \end{matrix} \right\|$$

é di dimensione $(\epsilon + 1 \times \epsilon)$ e dove $\widehat{A} + \lambda \widehat{B}$ é un fascio di matrici per il quale l'equazione $(\widehat{A} + \lambda \widehat{B})\widehat{x} = 0$ non ha soluzioni di grado inferiore a ϵ.

3.0.9 Forma canonica di un fascio singolare di matrici

Sia $A + \lambda B$ un fascio singolare arbitrario di matrici di dimensione $m \times n$. Innanzitutto supponiamo che non ci sia dipendenza lineare tra le colonne o le righe del fascio. Sia $r < n$, dove r é il rango del fascio, in modo che le colonne di $A + \lambda B$ sono linearmente dipendenti. Cosí l'equazione $(A + \lambda B)x = 0$ ha una soluzione non nulla di grado minimo $\epsilon_1 > 0$. Di conseguenza, in virtú del teorema di riduzione, il fascio dato puó essere messo sotto la forma

$$\begin{pmatrix} L_{\epsilon_1} & 0 \\ 0 & A_1 + \lambda B_1 \end{pmatrix}$$

dove l'equazione $(A_1 + \lambda B_1)x^{(1)} = 0$ non ha soluzioni $x^{(1)}$ di grado inferiore a ϵ_1. Se questa equazione ha una soluzione non nulla di grado minimo ϵ_2, con $\epsilon \geq \epsilon_2$, applicando il teorema di riduzione al fascio $A_1 + \lambda B_1$, possiamo mettere il fascio dato sotto la forma

$$\begin{pmatrix} L_{\epsilon_1} & 0 & 0 \\ 0 & L_{\epsilon_2} & 0 \\ 0 & 0 & A_2 + \lambda B_2 \end{pmatrix}.$$

Continuando questo processo, possiamo mettere il fascio sotto la seguente forma quasi diagonale

$$\begin{pmatrix} L_{\epsilon_1} & & & & 0 \\ 0 & L_{\epsilon_2} & & & \\ & & \ddots & & \\ 0 & & & L_{\epsilon_p} & \\ 0 & & & & A_p + \lambda B_p \end{pmatrix}, \tag{3.9}$$

dove $0 < \epsilon_1 \leq \epsilon_2 \leq \cdots \leq \epsilon_p$, e l'equazione $(A_p + \lambda B_p)x^{(p)} = 0$ non ha soluzioni non nulle, in modo che le colonne di $A_p + \lambda B_p$ siano linearmente indipendenti. Se le righe di $A_p + \lambda B_p$ sono linearmente dipendenti, il fascio trasposto $A_p^T + \lambda B_p^T$ puó essere messo sotto la forma (3.9) dove, al posto dei numeri $\epsilon_1, \epsilon_2, \ldots, \epsilon_p$, abbiamo i numeri $0 < \eta_1 \leq \eta_2 \leq \cdots \leq \eta_q$. Ma allora il fascio dato $A + \lambda B$ puó essere messo sotto la forma quasi diagonale

$$\begin{pmatrix} L_{\epsilon_1} & & & & & & & & 0 \\ & L_{\epsilon_2} & & & & & & & \\ & & \ddots & & & & & & \\ & & & L_{\epsilon_p} & & & & & \\ & & & & L_{\eta_1}^T & & & & \\ & & & & & L_{\eta_2}^T & & & \\ & & & & & & \ddots & & \\ & & & & & & & L_{\eta_q}^T & \\ 0 & & & & & & & & A_0 + \lambda B_0 \end{pmatrix} \tag{3.10}$$

$$(0 < \epsilon_1 \leq \epsilon_2 \leq \cdots \leq \epsilon_p, \quad 0 < \eta_1 \leq \eta_2 \leq \cdots \leq \eta_q),$$

dove le colonne e le righe di $A_0 + \lambda B_0$ sono linearmente indipendenti in modo che $A_0 + \lambda B_0$ sia un fascio regolare.

Consideriamo ora il caso generale in cui le righe e le colonne del fascio dato possono essere legate da relazioni lineari a coefficenti costanti. Indichiamo rispettivamente con g e h il numero massimo di soluzioni indipendenti costanti delle equazioni $(A + \lambda B)x = 0$ e $(A^T + \lambda B^T)y = 0$.

Si dimostra che il fascio dato prende allora la forma

$$\begin{pmatrix} \overbrace{\ 0\ }^{g}]_h, & 0 \\ 0, & A_0 + \lambda B_0 \end{pmatrix} \tag{3.11}$$

dove non vi é piú dipendenza lineare con coefficenti costanti tra le righe o le colonne del fascio $A_0 + \lambda B_0$. Il fascio $A_0 + \lambda B_0$ puó essere ora rappresentato nella forma (3.10). Cosí, nel caso generale, il fascio $A+\lambda B$ puó sempre essere messo nella forma quasi canonica

$$\{\overbrace{\ 0\ }^{g}]_h, L_{\epsilon_{g+1}}, \dots, L_{\epsilon_p}, L^T_{\eta_{h+1}}, \dots, L^T_{\eta_g}, A_0 + \lambda B_0\}. \tag{3.12}$$

La scelta degli inidici di ϵ e η é legata al fatto che conviene prendere

$$\epsilon_1 = \epsilon_2 = \dots = \epsilon_g = 0 \ \ e \ \ \eta_1 = \eta_2 = \dots = \eta_h = 0.$$

Sostituendo, in (3.12), il fascio regolare $A_0 + \lambda B_0$ con la sua forma canonica (3.3), otteniamo la matrice quasi diagonale seguente:

$$\{\overbrace{\ 0\ }^{g}]_h, L_{\epsilon_{g+1}}, \dots, L_{\epsilon_p}, L^T_{\eta_{h+1}}, \dots, L^T_{\eta_g};\ N^{(u_1)}, \dots, N^{(u_s)};\ J + \lambda I\}, \tag{3.13}$$

dove la matrice J é la forma normale di Jordan o forma normale naturale.
In conclusione la matrice (3.13) é la forma canonica del fascio $A + \lambda B$ nel caso piú generale.

3.0.10 Indici minimi di un fascio

Sia $A + \lambda B$ un fascio singolare arbitrario di matrici rettangolari. Siano $x_1(\lambda), x_2(\lambda), \dots, x_k(\lambda)$ le k colonne polinomiali soluzioni dell'equazione

$$(A + \lambda B)x = 0, \tag{3.14}$$

e consideriamo la matrice polinomiale $X = [x_1(\lambda), x_2(\lambda), \dots, x_k(\lambda)]$.
Tra tutte le soluzioni di (3.14), scegliamo una soluzione non nulla $x_1(\lambda)$ di grado almeno ϵ_1. Tra tutte le soluzioni, della stessa equazione, linearmente

indipendenti con $x_1(\lambda)$, scegliamo una soluzione $x_2(\lambda)$ di grado almeno ϵ_2, con $\epsilon_1 < \epsilon_2$. Continuando in questo modo, poiché il numero di soluzioni linearmente indipendenti di (3.14) é sempre al piú n, si vede che il processo ha una fine. Otteniamo cosí una *serie fondamentale di soluzioni* di (3.14)

$$x_1(\lambda), x_2(\lambda), \ldots, x_p(\lambda) \tag{3.15}$$

di grado

$$\epsilon_1 \leq \epsilon_2 \leq \ldots \leq \epsilon_p. \tag{3.16}$$

In generale, una serie fondamentale di soluzioni non é determinata in modo univoco dal fascio $A + \lambda B$. Tuttavia, due serie fondamentali di soluzioni distinte hanno sempre una e una sola serie di gradi $\epsilon_1, \epsilon_2, \ldots, \epsilon_p$.
Tutte le soluzioni $x_k(\lambda)$ della serie fondamentale (3.15) intrecciano una dipendenza lineare di grado ϵ_k tra le colonne di $A + \lambda B$ $(k = 1, 2, \ldots, p)$. I numeri $\epsilon_1, \epsilon_2, \ldots, \epsilon_p$ sono quindi chiamati *indici minimi per le colonne* del fascio $A+\lambda B$. Gli *indici minimi* $\eta_1, \eta_2, \ldots, \eta_q$ *per le righe* del fascio $A+\lambda B$ sono introdotti nello stesso modo. In questo caso, l'equazione $(A+\lambda B)x = 0$ é sostituita dall'equazione $(A^T + \lambda B^T)y = 0$ e i numeri $\eta_1, \eta_2, \ldots, \eta_q$ sono definiti come gli indici minimi per le colonne del fascio trasposto $A^T + \lambda B^T$.
Si puó dimostrare (ma noi non lo faremo) la seguente proposizione:

Proposizione 3.0.8. *Dei fasci strettamente equivalenti hanno gli stessi indici minimi.*

Ora calcoliamo gli indici minimi per la matrice quasi diagonale canonica

$$\{\overbrace{\ 0\ }^{g}]_h, L_{\epsilon_{g+1}}, \ldots, L_{\epsilon_p};\ \ L^T_{\eta_{h+1}}, \ldots, L^T_{\eta_q}, A_0 + \lambda B_0\}, \tag{3.17}$$

dove $A_0 + \lambda B_0$ é un fascio regolare avente la forma normale (3.3).
Innanzi tutto notiamo che il sistema completo degli inidici per le colonne (righe) di una matrice quasi diagonale é ottenuto attraverso l'unione dei sistemi corrispondenti degli indici minimi dei blocchi diagonali individuali. La matrice L_ϵ non ha che un solo indice ϵ per le colonne, e le righe di questa

matrice sono linearmente indipendenti. Analogamente, la matrice L_η^T non ha che un solo indice η per le righe, e le colonne di questa matrice sono linearmente indipendenti. Il fascio regolare $A_0 + \lambda B_0$ non ha indici minimi. Ne segue che la matrice (3.17) ha come indici minimi per la colonne

$$\epsilon_1 = \cdots = \epsilon_g = 0, \quad \epsilon_{g+1}, \cdots, \epsilon_p$$

e per le righe

$$\eta_1 = \cdots = \eta_h = 0, \quad \eta_{h+1}, \cdots, \eta_q.$$

Notiamo, inoltre, che L_ϵ non ha dei divisori elementari poiché, tra i suoi minori di ordine massimo, ce n'é uno uguale a 1 e uno uguale a λ^ϵ. La stessa cosa vale per la matrice trasposta L_ϵ^T. Poiché i divisori elementari di una matrice quasi diagonale sono ottenuti combinando quelli dei blocchi diagonali individuali (teorema 3.0.2), i divisori elementari della λ-matrice (3.17) coincidono con quelli della matrice $A_0 + \lambda B_0$.
La forma canonica del fascio (3.17) é completamente determinata dagli indici minimi $\epsilon_1, \ldots, \epsilon_p, \eta_1, \ldots, \eta_q$ e dai divisori elementari di questo fascio o del fascio equivalente $A_0 + \lambda B_0$. Poiché due fasci aventi una e una sola forma canonica sono strettamente equivalenti, abbiamo dimostrato il teorema seguente, chiamato *teorema di Kronecker*:

Teorema 3.0.9. *Due fasci arbitrari $A+\lambda B$ e $A_1+\lambda B_1$ di matrici rettangolari della stessa dimensione $m\times n$ sono strettamente equivalenti se e solo se hanno gli stessi indici minimi e gli stessi divisori elementari (finiti e infiniti).*

3.0.11 Fasci singolari di forme quadratiche

Supponiamo date le due forme quadratiche complesse

$$A(x,x) = \sum_{i,k=1}^{n} a_{ik}x_i x_k, \quad B(x,x) = \sum_{i,k=1}^{n} b_{ik}x_i x_k; \tag{3.18}$$

esse costituiscono un fascio di forme quadratiche $A(x,x) + \lambda B(x,x)$. Questo fascio di forme corrisponde ad un fascio di matrici simmetriche $A + \lambda B$ (cioé

$A^T = A$, $B^T = B$). Se noi applichiamo alle variabili del fascio di forme $A(x, x) + \lambda B(x, x)$ una trasformazione lineare regolare $x = Tz$ ($|T| \neq 0$), il fascio trasformato di forme

$$\widetilde{A}(z, z) + \lambda \widetilde{B}(z, z)$$

corrisponde al fascio di matrici

$$\widetilde{A} + \lambda \widetilde{B} = T^T (A + \lambda B) T; \tag{3.19}$$

dove T é una matrice quadrata regolare costante di ordine n.

Definizione 3.0.12. *Due fasci di matrici $A + \lambda B$ e $\widetilde{A} + \lambda \widetilde{B}$ legati dalle relazione (3.19) sono detti congruenti.*

La congruenza é un caso particolare dell'equivalenza tra fasci di matrici. Tuttavia, se consideriamo la congruenza tra due fasci di matrici simmetriche, il concetto di congruenza coincide con quello di equivalenza.

Teorema 3.0.10. *Due fasci strettamente equvalenti di matrici complesse, simmetriche sono sempre congruenti.*

Omettiamo la dimostrazione.
Da questo teorema dal teorema di Kronecker deduciamo il seguente corollario:

Corollario 3.0.11. *Due fasci di forme quadratiche*

$$A(x, x) + \lambda B(x, x) \text{ e } \widetilde{A}(z, z) + \lambda \widetilde{B}(z, z)$$

possono essere resi identici l'uno all'altro da una trasformazione $X = Tz$ (con $|T| \neq 0$) se e solamente se i fasci di matrici simmetriche $A + \lambda B$ e $\widetilde{A} + \lambda \widetilde{B}$ hanno gli stessi divisori elementari (finiti e infiniti) e gli stessi indici minimi.

Ora poniamo la questione seguente:
essendo date due forme quadratiche complesse arbitrarie

$$A(x, x) = \sum_{i,k=1}^{n} a_{ik} x_i x_k, \quad B(x, x) = \sum_{i,k=1}^{n} b_{ik} x_i x_k,$$

sotto quali condizioni le due forme possono essere trasformate in somme di quadrati

$$\sum_{i=1}^{n} a_i {z_i}^2 \quad \text{e} \quad \sum_{i=1}^{n} b_i {z_i}^2 \tag{3.20}$$

attraverso una trasformazione regolare delle variabili $X = Tz$ ($|T| \neq 0$)? Supponiamo che le forme quadratiche $A(x,x)$ e $B(x,x)$ posseggano questa proprietá. Allora il fascio di matrici $A+\lambda B$ é congruente al fascio di matrici diagonali

$$\{a_1 + \lambda b_1, a_2 + \lambda b_2, \dots, a_n + \lambda b_n\}. \tag{3.21}$$

Supponiamo che, tra i binomi diagonali $a_i + \lambda b_i$, ce ne siano esattamente r ($r \leq n$) che non sono identicamente nulli. Senza perdita di generalitá, possiamo inoltre supporre che

$$a_1 = b_1 = 0, \dots, a_{n-r} = b_{n-r} = 0, \quad a_i + \lambda b_i \neq 0 \quad (i = n-r+1, \dots, n).$$

E ponendo

$$A_0 + \lambda B_0 = \{a_{n-r+1} + \lambda b_{n-r+1}, \dots, a_n + \lambda b_n\}, \tag{3.22}$$

possiamo rappresentare la matrice (3.21) cosí: $\{\overbrace{\,0\,}^{n-r}, A_0 + \lambda B_0\}$.
Confrontando (3.22) e (3.17) , vediamo che, tutti gli indici minimi sono nulli. Di piú, tutti i divisori elementari sono lineari; perció abbiamo che:

Teorema 3.0.12. *Due forme quadratiche $A(x,x)$ e $B(x,x)$ possono essere ridotte simultaneamente in somme di quadrati (3.20) attraverso una trasformazione delle variabili se e solamente se il fascio di matrici $A+\lambda B$ e i divisori elementari sono lineari e se gli indici minimi sono nulli.*

Supponiamo che il fascio di matrci simmetriche $A+\lambda B$ abbia indici minimi $\epsilon_1 = \dots = \epsilon_g = 0$, $\epsilon_{g+1} \neq 0, \dots, \epsilon_p \neq 0$, divisori elementari infiniti $\mu^{(u_1)}, \mu^{(u_2)}, \dots, \mu^{(u_s)}$ e divisori elementari finiti

$$(\lambda + \lambda_1)^{c_1}, (\lambda + \lambda_2)^{c_2}, \dots, (\lambda + \lambda_t)^{c_t}.$$

Allora, nella forma canonica (3.13), poiché per ipotesi le matrici sono simmetriche abbiamo $g = h,\ p = q,$ e $\epsilon_{g+1} = \eta_{g+1}, \dots, \epsilon_p = \eta_p$.
Sostituiamo, in (3.13), ogni coppia di blocchi diagonali della forma L_ϵ e L_ϵ^T con un solo blocco diagonale

$$\begin{pmatrix} 0 & L_\epsilon^T \\ L_\epsilon & 0 \end{pmatrix},$$

e ogni blocco della forma $N^{(u)}$ con il blocco simmetrico strettamente equivalente

$$\widetilde{N}^{(u)} = V^{(u)} N^{(u)} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & \lambda \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & \lambda & \cdots & 0 & 0 \end{pmatrix} \text{con } V^{(u)} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Di piú, in (3.13), al posto del blocco diagonale regolare $J + \lambda I$ (J é una matrice di Jordan)

$$J + \lambda I = \{(\lambda + \lambda_1)E^{(c_1)} + H^{(c_1)}, \dots, (\lambda + \lambda_t)E^{(c_t)} + H^{(c_t)}\}^2$$

prendiamo il blocco strettamente equivalente $\{Z_{\lambda_1}^{(c_1)}, \dots, Z_{\lambda_t}^{(c_t)}\}$, dove

$$Z_{\lambda_i}^{(c_i)} = V^{(c_i)}[(\lambda + \lambda_i)E^{(c_i)} + H^{(ci)}] = \begin{pmatrix} 0 & \cdots & & \lambda + \lambda_i \\ 0 & \cdots & \lambda + \lambda_i & 1 \\ & & & \\ \lambda + \lambda_i & 1 & \cdots & 0 \end{pmatrix} \tag{3.23}$$

con $i = 1, 2, \dots, t$.
Il fascio $A + \lambda B$ é strettamente equivalente al fascio simmetrico

$$\widetilde{A} + \lambda \widetilde{B} = \left\{0, \begin{pmatrix} 0 & L_{\epsilon_{g+1}}^T \\ L_{\epsilon_{g+1}} & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & L_{\epsilon_p}^T \\ L_{\epsilon_p} & 0 \end{pmatrix}; \widetilde{N}^{(u_1)}, \dots, \widetilde{N}^{(u_s)}, Z_{\lambda_1}^{(c_1)}, \dots, Z_{\lambda_t}^{(c_t)}\right\}. \tag{3.24}$$

[2] H é una matrice avente tutti gli elementi nulli tranne quelli della sovradiagonale che sono uguali a 1.

In conclusione due forme quadratiche a coefficenti complessi $A(x,x)$ e $B(x,x)$ possono essere ridotte simultaneamente in forme canoniche $\widetilde{A}(z,z)$ e $\widetilde{B}(z,z)$ definite da (3.24) attraverso una trasformazione delle variabili

$$x = Tz \quad (|T| \neq 0).$$

Capitolo 4

I risultati ottenuti nel capitolo 3 verranno ora applicati a un sistema di m equazioni differenziali del primo ordine a coefficenti costanti e n funzioni incognite (il caso particolare in cui $m = n$, e dove il sistema é risolubile rispetto alle derivate, é giá stato trattato nel capitolo 2):

$$\sum_{k=1}^{n} a_{ik}x_k + \sum_{k=1}^{n} b_{ik}\frac{dx_k}{dt} = f_i(t) \quad (i = 1, 2, ..., m) \tag{4.1}$$

o in notazione matriciale

$$Ax + B\frac{dx}{dt} = f(t); \tag{4.2}$$

dove $A = ||a_{ik}||, \quad B = ||b_{ik}|| \quad (i = 1, 2, ..., m; \quad k = 1, 2, ..., n)$,
x é la matrice colonna $x = (x_1, x_2, ..., x_n), \quad f = (f_1, f_2, ..., f_n)$.
Introduciamo le nuove funzioni incognite $z_1, z_2, ..., z_n$ che sono collegate alle precedenti $x_1, x_2, ..., x_n$ tramite una trasformazione lineare regolare a coefficenti costanti:

$$x = Qz \quad [z = (z_1, z_2, ..., z_n); |Q| \neq 0]. \tag{4.3}$$

Di piú, al posto delle equazioni (4.1) possiamo considerare m combinazioni indipendenti arbitrarie di queste equazioni, e questo é equivalente alla moltiplicazione a destra delle matrici A, B, f per una matrice quadrata regolare P d'ordine m. Sostituendo Qz a x in (4.2) e moltiplicando, a destra, (4.2) per P, otteniamo

$$\widetilde{A}z + \widetilde{B}\frac{dz}{dt} = \widetilde{f}(t) \quad , \tag{4.4}$$

dove

$$\widetilde{A} = PAQ, \quad \widetilde{B} = PBQ, \quad \widetilde{f} = Pf = (\widetilde{f}_1, \widetilde{f}_2, ..., \widetilde{f}_n). \tag{4.5}$$

Il fascio di matrici $A + \lambda B$ e $\widetilde{A} + \lambda\widetilde{B}$ sono strettamente equivalenti:

$$\widetilde{A} + \lambda\widetilde{B} = P(A + \lambda B)Q. \tag{4.6}$$

Scegliamo le matrici P e Q in modo che il fascio $\widetilde{A}+\lambda\widetilde{B}$ abbia la forma quasi diagonale canonica

$$\widetilde{A} + \lambda\widetilde{B} = \{0, L_{\varepsilon_g+1}, ..., L_{\varepsilon_p}, L'_{\eta_h+1}, ..., L'_{\eta_g}, N^{(u_1)}, ..., N^{(u_s)}, J + \lambda I\}. \tag{4.7}$$

In concordanza con i blocchi diagonali di (4.7), il sistema di equazioni differenziali é diviso in

$$v = p - g + q - h + s + 2$$

sistemi separati della forma:

$$0 \cdot \overset{1}{z} = \overset{1}{\widetilde{f}}, \tag{4.8}$$

$$L_{\varepsilon_g+i}\left(\frac{d}{dt}\right)\overset{1+i}{z} = \overset{1+i}{\widetilde{f}} \quad (i = 1, 2, ..., p - g), \tag{4.9}$$

$$L'_{\eta_h+j}\left(\frac{d}{dt}\right)\overset{p-g+1+j}{z} = \overset{p-g+1+j}{\widetilde{f}} \quad (j = 1, 2, ..., q - h), \tag{4.10}$$

$$N^{(u_k)}\left(\frac{d}{dt}\right)\overset{p-g+q-h+1+k}{z} = \overset{p-g+q-h+1+k}{\widetilde{f}} \quad (k = 1, 2, ..., s), \tag{4.11}$$

$$\left(J + \frac{d}{dt}\right)\overset{v}{z} = \overset{v}{\widetilde{f}}, \tag{4.12}$$

dove

$$z = \left\|\begin{matrix} \overset{1}{z} \\ \overset{2}{z} \\ \vdots \\ \overset{v}{z} \end{matrix}\right\|, \widetilde{f} = \left\|\begin{matrix} \overset{1}{\widetilde{f}} \\ \overset{2}{\widetilde{f}} \\ \vdots \\ \overset{v}{\widetilde{f}} \end{matrix}\right\|, \tag{4.13}$$

$$\overset{1}{z} = (z_1, ..., z_g), \overset{1}{\widetilde{f}} = (\widetilde{f}_1, ..., \widetilde{f}_h), \overset{2}{z} = (z_{g+1}, ...), \overset{2}{\widetilde{f}} = (\widetilde{f}_{h+1}, ...) \; ecc..., \quad (4.14)$$

$$\Lambda\left(\frac{d}{dt}\right) = A + B\frac{d}{dt}, \quad \text{se} \quad \Lambda(\lambda) \equiv A + \lambda B. \quad (4.15)$$

Cosí, nel caso piú generale, l'integrazione del sistema (4.2) si riduce all'integrazione dei sistemi particolari (4.8)-(4.12) dello stesso tipo. In questi sistemi, il fascio di matrici $A + \lambda B$ ha le rispettive forme:

$$0, L_\varepsilon, L'_\eta, N^{(u)}, J + \lambda I.$$

1. Il sistema (4.8) non é possibile se e solamente se $\overset{1}{\widetilde{f}} \equiv 0$, cioé se

$$\widetilde{f}_1 \equiv 0, ..., \widetilde{f}_h \equiv 0.$$

   In questo caso, noi possiamo prendere delle funzioni arbitrarie di t come funzioni incognite $z_1, z_2, ..., z_g$ le quali formano la colonna $\overset{1}{z}$.

2. Il sistema (4.9) é della forma $L_\varepsilon\left(\frac{d}{dt}\right) z = \widetilde{f}$ oppure piú esplicitamente[1],

$$\frac{dz_1}{dt} + z_2 = \widetilde{f}_1(t), \frac{dz_2}{dt} + z_3 = \widetilde{f}_2(t), ..., \frac{dz_\varepsilon}{dt} + z_{\varepsilon+1} = \widetilde{f}_\varepsilon(t). \quad (4.16)$$

   Un tale sistema é sempre risolubile. Se prendiamo per $z_{\varepsilon+1}(t)$ una funzione arbitraria di t, tutte le funzioni incognite restanti $z_\varepsilon, z_{\varepsilon-1}, ..., z_1$ possono essere determinate a partire da (4.16) attraverso delle quadrature successive.

3. Il sistema (4.10) é della forma $L'_\eta\left(\frac{d}{dt}\right) z = \widetilde{f}$ oppure piú esplicitamente

$$\frac{dz_1}{dt} = \widetilde{f}_1(t), \frac{dz_2}{dt} + z_1 = \widetilde{f}_2(t), \dots, \frac{dz_\eta}{dt} + z_{\eta-1} = \widetilde{f}_\eta(t), z_\eta = \widetilde{f}_{\eta+1}(t). \quad (4.17)$$

[1] D'ora in poi modificheremo gli indici di z e $\widetilde{f}$ per semplificare la notazione. Per passare da (4.16) a (4.9), bisogna sostituire ε con ε_i e aggiungere a ciascun inidice di z il numero $g + \varepsilon_{g+1} + ... + \varepsilon_{g+i-1} + i - 1$, e a ciascun indice di $\widetilde{f}$ il numero $h + \varepsilon_{g+1} + ... + \varepsilon_{g+i-1}$.

A partire dalle equazioni (4.17), con l'eccezione della prima, possiamo determinare univocamente $z_\eta, z_{\eta-1}, \ldots, z_1$:

$$z_\eta = \widetilde{f}_{\eta+1}, z_{\eta-1} = \widetilde{f}_\eta - \frac{d\widetilde{f}_{\eta+1}}{dt}, \ldots, z_1 = \widetilde{f}_2 - \frac{d\widetilde{f}_3}{dt} + \ldots + (-1)^{\eta-1}\frac{d^{\eta-1}\widetilde{f}_{\eta+1}}{dt^{\eta-1}}. \quad (4.18)$$

Sostituendo l'espressione di z_1 ottenuta nella prima equazione, abbiamo la seguente condizione di possibilitá:

$$\widetilde{f}_1 - \frac{d\widetilde{f}_2}{dt} + \frac{d^2\widetilde{f}_3}{dt^2} - \ldots + (-1)^\eta \frac{d^\eta \widetilde{f}_{\eta+1}}{dt^\eta} = 0. \quad (4.19)$$

4. Il sistema (4.11) é della forma $N^{(u)}\left(\frac{d}{dt}\right) z = \widetilde{f}$ oppure, piú esplicitamente,

$$\frac{dz_2}{dt} + z_1 = \widetilde{f}_1, \frac{dz_3}{dt} + z_2 = \widetilde{f}_2, \ldots, \frac{dz_u}{dt} + z_{u-1} = \widetilde{f}_{u-1}, z_u = \widetilde{f}_u. \quad (4.20)$$

Possiamo quindi determinare in modo univoco le soluzioni:

$$\begin{aligned} &z_u = \widetilde{f}_u \,, \\ &z_{u-1} = \widetilde{f}_{u-1} - \frac{d\widetilde{f}_u}{dt} \,, \qquad (4.21) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \,, \\ &z_1 = \widetilde{f}_1 - \frac{d\widetilde{f}_2}{dt} + \frac{d^2\widetilde{f}_3}{dt} - \ldots + (-1)^{u-1}\frac{d^{u-1}\widetilde{f}_u}{dt^{u-1}} \,. \end{aligned}$$

5. Il sistema (4.12) é della forma $Jz + \frac{dz}{dt} = \widetilde{f}$. Come abbiamo giá visto nel capitolo 2, la soluzione generale di un tale sistema é della forma

$$z = e^{-Jt}z_0 + \int_0^t e^{-J(t-\tau)f(\tau)d\tau}, \qquad (4.22)$$

dove z_0 é una matrice colonna a elementi arbitrari (i valori iniziali della funzioni incognite per $t = 0$). Il passaggio inverso dal sistema (4.4) al sistema (4.2) si effettua tramite le formule (4.3) e (4.5), secondo le quali ciascuna funzione $x_1, x_2, \ldots, x_n$ é una combinazione lineare delle funzioni $z_1, z_2, \ldots, z_n$ e ciascuna delle funzioni $\widetilde{f}_1(t), \ldots, \widetilde{f}_m(t)$ si esprime linearmente (con dei coefficenti costanti) a partire dalle funzioni $f_1(t), \ldots, f_m(t)$.

Bibliografia

[1] F.R. Gantmacher, Théorie des matrices, Tome 1, Théorie generale, Dunod, 1966

[2] F.R. Gantmacher, Théorie des matrices, Tome 2, Questions spéciales et applications, Dunod, 1966