



UNIVERSITÀ DI PISA

---

Dipartimento di Matematica  
Corso di Laurea Triennale in Matematica

Tesi di Laurea Triennale

INDEX SUM THEOREM E  
 $\ell$ -FICAZIONI DI  
MATRICI POLINOMIALI

*Relatore:*  
Prof. Dario Andrea Bini

*Candidato:*  
Gian Maria Negri Porzio

---

Anno accademico 2014–2015



# Indice

<b>1</b>	<b>Introduzione</b>	<b>1</b>
1.1	Nozioni preliminari . . . . .	1
1.2	Problemi generalizzati agli autovalori . . . . .	2
<b>2</b>	<b>Teoria sui polinomi di matrici</b>	<b>5</b>
2.1	Struttura spettrale . . . . .	5
2.2	Polinomi singolari e sottospazi di $\mathbb{F}(\lambda)^n$ . . . . .	8
2.3	Relazioni di equivalenza . . . . .	9
2.4	Dimensione delle linearizzazioni . . . . .	13
<b>3</b>	<b>Matrici companion</b>	<b>21</b>
3.1	Matrici companion di Frobenius . . . . .	22
3.2	Matrici companion di grado $\ell > 1$ . . . . .	28
3.3	Altri tipi di $\ell$ -ficazioni . . . . .	33
3.3.1	$\mathcal{L}$ -ficazione secolare . . . . .	35
<b>4</b>	<b>Index sum theorem</b>	<b>39</b>
4.1	Dimostrazione dell'index sum theorem . . . . .	39
4.2	Caratterizzazione dei gradi delle $\ell$ -ficazioni . . . . .	42
	<b>Bibliografia</b>	<b>45</b>



# Sommario

I problemi generali agli autovalori e i polinomi  $P(\lambda) = \sum_{i=0}^k \lambda^i A_i$ , con  $A_i$  matrici quadrate, emergono in numerosi ambiti applicativi, che spaziano dalla risoluzione di equazioni differenziali alla teoria dei grafi, fino allo studio dei sistemi biologici. Il metodo principale per approcciare questi problemi è chiamato *linearizzazione*: il polinomio iniziale  $P(\lambda)$  viene trasformato nell'equivalente di primo grado  $\lambda C_1 + C_0$ , per il quale già dagli anni '70 conosciamo algoritmi efficienti per trovarne gli autovalori.

La letteratura sull'argomento continua a costruire nuovi tipi di linearizzazioni che conservino la struttura del problema originale, come la simmetria o l'Hermitianità. Recentemente è stato dimostrato come questo non sia sempre possibile: De Terán, Dopico e Mackey hanno generalizzato l'*index sum theorem*, mettendo in relazione gli autovalori dei polinomi di matrici con il loro rango e le loro dimensioni. Il risultato sopra citato ne è una diretta conseguenza. Negli stessi articoli è stato introdotto un metodo più generale per affrontare i problemi agli autovalori, chiamato  $\ell$ -ficazione, in cui un polinomio di matrici  $P(\lambda)$  di grado  $k$  viene trasformato in un polinomio equivalente  $Q(\lambda)$  di grado  $\ell < k$ .

In questa tesi ci siamo proposti di riassumere codesti risultati in modo chiaro e conciso. Benché gli argomenti trattati permettano di seguire numerosi spunti applicativi, abbiamo optato per una prospettiva prettamente teorica. I risultati qui mostrati prediligeranno dunque una maggiore generalità e visione d'insieme, piuttosto che un approfondimento delle applicazioni numeriche. Tuttavia si pensa che in un futuro prossimo si possano introdurre nuovi algoritmi che sfruttino questi teoremi e siano vantaggiosi sotto l'aspetto computazionale e numerico.



# Capitolo 1

## Introduzione

In questi brevi paragrafi mostreremo perché oggi numerosi articoli di letteratura studino i problemi generalizzati agli autovalori, sia da un punto di vista applicativo che pratico. Introduciamo la notazione utilizzata e le definizioni di base in maniera informale, in modo da sottolineare i collegamenti con l'algebra lineare del primo anno di università, per poi esporre una solida struttura teorica nel capitolo 2. Nel capitolo 3 presenteremo invece alcune companion per polinomi di matrici, mentre nel capitolo 4 dimostreremo l'*index sum theorem* e le sue principali conseguenze.

### 1.1 Nozioni preliminari

Iniziamo ripassando brevemente alcune nozioni di base e introducendo le notazioni che utilizzeremo durante questo lavoro. Indicheremo con  $\mathbb{R}$ ,  $\mathbb{C}$ ,  $\mathbb{F}$ ,  $\overline{\mathbb{F}}$  il campo dei numeri reali, quello complesso e un campo generico con la sua chiusura algebrica;  $\mathbb{R}^n$ ,  $\mathbb{C}^n$ ,  $\mathbb{F}^n$  saranno i rispettivi spazi vettoriali di dimensione  $n$ ;  $\mathbb{F}^{m \times n}$  rappresenterà l'anello delle matrici  $m \times n$  a coefficienti in  $\mathbb{F}$ . Diamo per scontato le definizioni di *autovalore*, *autovettore* e *spettro* di una matrice  $A \in \mathbb{F}^{n \times n}$ . Alcune notazioni, che riportiamo di seguito, meritano un'attenzione particolare, in quanto saranno costantemente usate.

**Definizione 1.1.1** (Anello di polinomi e campo delle frazioni). Identifichiamo con  $\mathbb{F}[\lambda]$  l'anello dei polinomi a coefficienti in  $\mathbb{F}$  e con  $\mathbb{F}(\lambda)$  il campo delle funzioni razionali; rispettando le convenzioni ormai usuali,  $\mathbb{F}[\lambda]^m$  rappresenterà il modulo libero di dimensione  $m$  su  $\mathbb{F}[\lambda]$ , mentre  $\mathbb{F}[\lambda]^{m \times n}$  saranno gli operatori lineari da  $\mathbb{F}[\lambda]^n$  a  $\mathbb{F}[\lambda]^m$ . Manterremo le analoghe notazioni per gli spazi vettoriali su  $\mathbb{F}(\lambda)$ .

Lo studio delle proprietà spettrali delle matrici ha numerose applicazioni pratiche, dalla meccanica alla biologia, dai sistemi dinamici alla teoria dei grafi, solo per citare alcuni esempi. Nasce così *il problema agli autovalori* (*standard eigenvalues problem*, *SEP*): data una matrice  $A \in \mathbb{F}^{n \times n}$ , trovare  $\theta$  e

$x$  tali che  $(A - \theta I)x = 0$ . È evidente che questo problema, già per  $n \geq 5$ , non può essere risolto per via algebrica. Negli anni '50 del secolo scorso nacquero diverse tecniche numeriche per affrontarlo: oggi gli algoritmi più utilizzati si dividono in due sottoclassi:

- *Decomposizione QR*: i più diffusi anche da un punto di vista software, hanno un costo computazionale cubico rispetto alla dimensione del problema. Quando si introduce un nuovo algoritmo si è soliti paragonarlo ai metodi QR, sia sotto l'aspetto della velocità che della stabilità numerica [8] [18].
- *Sottospazi di Krylov*: se siamo interessati solo ad una parte dello spettro e stiamo lavorando con matrici sparse di grosse dimensioni.

## 1.2 Problemi generalizzati agli autovalori

In effetti in molti contesti dobbiamo risolvere problemi “più complessi” di un SEP. Prima di introdurre altre generalizzazioni del problema agli autovalori e le relative applicazioni, necessitiamo di una definizione.

**Definizione 1.2.1.** (Polinomio di matrici) Siano  $A_0, \dots, A_k \in \mathbb{F}^{m \times n}$ , con  $A_k \neq 0$ . Un polinomio di matrici di grado  $k$  è un'espressione algebrica della forma:

$$P(\lambda) = \sum_{i=0}^k \lambda^i A_i.$$

In alternativa possiamo vedere  $P(\lambda)$  come una matrice appartenente a  $\mathbb{F}[\lambda]^{m \times n}$ . In questa tesi i due approcci avranno un peso equivalente e non sarà raro riferirsi all'oggetto  $P(\lambda)$  prima come polinomio di matrici, poi come matrice di polinomi. Spesso lasceremo sottinteso “di matrici” o “di polinomi” per non appesantire eccessivamente il discorso.

La teoria sviluppata servirà per tre grandi classi di problemi agli autovalori:

**Problema lineare agli autovalori (LEP)** . Date  $A, B \in \mathbb{F}^{n \times n}$  trovare le coppie  $(\theta_i, v_i) \in \mathbb{F} \times \mathbb{F}^n$  tali che:

$$(A + \theta_i B)v_i = 0$$

**Problema quadratico agli autovalori (QEP)**. Dato un polinomio di secondo grado  $\lambda^2 A + \lambda B + C \in \mathbb{F}[\lambda]^{n \times n}$  di secondo grado, trovare le coppie  $(\theta_i, v_i) \in \mathbb{F} \times \mathbb{F}^n$  tali che:

$$(\theta_i^2 A + \theta_i B + C)v_i = 0$$



**Problema polinomiale agli autovalori (PEP).** Dato un polinomio di matrici  $\in \mathbb{F}[\lambda]^{n \times n}$  di grado  $k$  trovare le coppie  $(\theta_i, v_i) \in \mathbb{F} \times \mathbb{F}^n$  tali che:

$$P(\theta_i)v_i = 0$$

Come per il SEP, chiameremo  $\theta_i, v_i$  *autovalore* e *autovettore*. Spesso nelle applicazioni numeriche è necessario studiare solo gli autovalori o un sottoinsieme di questi. Ad esempio, se  $P(\lambda)$  modellizzasse l'evoluzione di un sistema dinamico, saremmo interessati solo alla parte reale degli autovalori, oppure vorremmo trovare solamente gli autovalori nella sfera unitaria, per studiarne la stabilità [19].

La tecnica standard per risolvere un LEP è una diretta conseguenza del prossimo teorema [17].

**Teorema 1.2.2.** (*Fattorizzazione QZ*) Siano  $A, B \in \mathbb{C}^{n \times n}$ . Allora esistono  $Q, Z \in \mathbb{C}^{n \times n}$  unitarie tali che

$$QTZ = A \quad QSZ = B$$

con  $T, S$  matrici triangolari superiori.

Consideriamo dunque il problema  $P(\lambda) = A - \lambda B$  e le matrici  $Q, Z$  come nel teorema 1.2.2. Vale allora l'uguaglianza  $\det(A - \lambda B) = \det(T - \lambda S)$ ; dunque gli autovalori sono dati da  $\theta_j = s_{jj}/t_{jj}$ . Se fosse  $t_{jj} = 0$  per qualche  $j$ , il problema lineare introdurrebbe *autovalori all'infinito* [17]. Se si presentasse la forma indeterminata  $0/0$  diremmo che il problema è *mal posto* ("ill posed", in inglese). Infatti piccoli errori di approssimazione, causati anche dall'aritmetica di macchina, potrebbero trasformare un autovalore finito in uno infinito e viceversa. Questo caso non è stato particolarmente studiato, in quanto nelle applicazioni numeriche si richiede sempre che il problema sia ben posto. Esistono numerosi algoritmi, raggruppati tutti nella classe dei *metodi QZ*, per ricavare le matrici  $Q$  e  $Z$  date  $A$  e  $B$ . Il capostipite è stato introdotto agli inizi degli anni '70 da Moler e Stewart in [17].

Per quanto riguarda che i problemi quadratici e quelli polinomiali, è chiaro che i primi siano una sottoclasse dei secondi; tuttavia meritano un'attenzione particolare per motivi storici e per i numerosi contesti dove li incontriamo, come possiamo vedere nel prossimo esempio.

**Esempio 1.2.3** (Modello per una corda di chitarra). È noto che le soluzioni di un oscillatore armonico smorzato

$$\begin{cases} m\ddot{x} = -kx - c\dot{x} \\ x(0) = 0 \\ \dot{x}(0) = v \end{cases}$$

(con  $c^2 - 4mk \neq 0$ ) siano nella forma  $x(t) = a_1 e^{\lambda_1 t} + a_2 e^{\lambda_2 t}$ , con  $m\lambda_i^2 + c\lambda_i + k = 0$  e  $a_i$  costanti arbitrarie.<sup>1</sup> Nella figura 1.1 possiamo osservare un modello

<sup>1</sup>Se  $c^2 - 4mk = 0$  e dunque  $\lambda_1 = \lambda_2 = \lambda$  si può dimostrare che  $e^{\lambda t}$  e  $te^{\lambda t}$  formano una base dello spazio delle soluzioni.

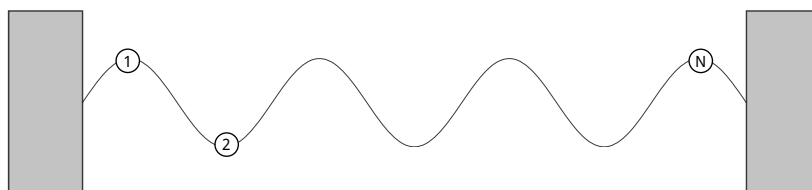


Figura 1.1: Una corda di chitarra può essere modellizzata da  $N$  masse discrete oscillanti.

oscillatorio per la corda di una chitarra:  $N$  masse  $m_i$  di coordinate  $(x_i, y_i)$  sono collegate da  $N - 1$  oscillatori armonici smorzati. Possiamo dunque scrivere :

$$\begin{cases} m_i \ddot{x}_i = -k_{i+1}(x_i - x_{i+1}) - k_i(x_i - x_{i-1}) - c_i \dot{x}_i & \forall i \in \{1, \dots, N\} \\ m_i \ddot{y}_i = -k_{i+1}(y_i - y_{i+1}) - k_i(y_i - y_{i-1}) - c_i \dot{y}_i & \forall i \in \{1, \dots, N\} \end{cases}$$

e definendo le matrici  $M = \text{diag}(m_1 \dots m_N)$ ,  $C = \text{diag}(c_1 \dots c_N)$  e

$$K = \begin{bmatrix} \ddots & & & & & & & & & & \\ \ddots & \ddots & & & & & & & & & \\ \ddots & \ddots & \ddots & & & & & & & & \\ & & -k_i & k_i + k_{i+1} & -k_{i+1} & & & & & & \\ & & & \ddots & & \ddots & \ddots & & & & \\ & & & & & & & \ddots & \ddots & & \\ & & & & & & & & & \ddots & \ddots \end{bmatrix}$$

il nostro sistema differenziale diventa

$$M\ddot{x} + C\dot{x} + Kx = 0, \quad M\ddot{y} + C\dot{y} + Ky = 0.$$

Consideriamo solo il vettore  $x$  delle ascisse e supponiamo, per semplicità, che gli autovalori  $\theta_1, \dots, \theta_{2N}$  siano distinti e chiamiamo  $D = \text{diag}(\theta_1, \dots, \theta_{2N})$ . Allora una soluzione può essere scritta nella forma  $x(t) = xe^{Dt}a$ , con  $a = [a_1, \dots, a_{2N}]^T$  vettore di costanti arbitrarie.

Questo è un esempio “giocattolo” di un *problema quadratico agli autovalori (QEP)*. Nella vita reale i QEP possono avere notevoli ripercussioni. Il 7 Novembre 1940 il Tacoma Narrow Bridge collassò a causa degli effetti risonanti dei forti venti che spiravano nel canyon, a meno di un anno dalla sua apertura. Più recentemente, durante l'inaugurazione avvenuta nel giugno del 2000, il London Millenium Footbridge iniziò a dondolare pericolosamente. Anche in questo caso la struttura era entrata in risonanza con il marciare ritmico dei primi pedoni che la attraversavano. Le frequenze naturali di risonanza furono trovate tramite lo studio di un QEP [19]. Sempre in [19] sono presentate numerose altre applicazioni relative, ad esempio, alla fluidodinamica o alla trasmissione dei segnali.

## Capitolo 2

# Teoria sui polinomi di matrici

In questo capitolo esporremo formalmente la teoria sulle equivalenze spettrali tra matrici di polinomi  $P(\lambda) \in \mathbb{F}[\lambda]^{m \times n}$ . Il risultato principale di questa tesi, l'Index Sum Theorem, sarà esposto nel capitolo 4 e fornirà delle restrizioni su queste equivalenze. Sottolineiamo la generalità dei risultati che mostreremo: non chiediamo ipotesi particolari sul campo  $\mathbb{F}$  e le matrici possono anche essere rettangolari, mentre in molti lavori ci si sofferma sul caso quadrato. Ancora una volta dobbiamo iniziare con alcune definizioni.

### 2.1 Struttura spettrale

Consideriamo un polinomio di matrici:

$$P(\lambda) = \sum_{i=0}^d \lambda^i A_i \quad A_i \in \mathbb{F}^{m \times n}$$

Riprendendo la definizione 1.2.1, diremo che  $P(\lambda)$  è *regolare* se è invertibile come matrice di  $\mathbb{F}(\lambda)^{m \times n}$ ; un polinomio non regolare è *singolare*. Se  $A_d \neq 0_{m \times n}$  (dove  $0_{m \times n}$  è la matrice degli zeri di dimensione  $m \times n$ ) allora il *grado* di  $P(\lambda)$ , che indichiamo con  $\deg(P)$ , sarà  $d$ . Introduciamo con la prossima definizione una generalizzazione del grado di un polinomio matriciale. Questa, oltre a fornirci maggiore versatilità, permette di comprendere meglio la natura dei polinomi singolari e di enunciare numerosi teoremi con maggiore chiarezza.

**Definizione 2.1.1** (Pseudo grado). Sia  $P(\lambda) = \sum_{i=0}^k \lambda^i A_i$  un polinomio di matrici. Chiamiamo *pseudo grado* di  $P$  il numero  $\text{gr } P = k$ .

Dunque non si chiede che  $A_k \neq 0_{m \times n}$ . Il prossimo esempio chiarirà la differenza tra grado e pseudo grado.

**Esempio 2.1.2.** Sia  $A \in \mathbb{F}^{m \times n}$  e siano  $P(\lambda) = \lambda I + A$ ,  $Q(\lambda) = \lambda^2 0 + \lambda I + A$ . Allora  $\deg P = \text{gr } P = \deg Q = 1$ , mentre  $\text{gr } Q = 2$ .

Come possiamo osservare, è necessario assegnare in anticipo lo pseudo grado ad un polinomio. Nel caso non venisse fatto lo si suppone uguale al grado  $\deg(P)$ . Questa estensione del concetto di grado è stata introdotta per la prima volta in [10] da Gohberg, Kaashoek e Lancaster, e in seguito da De Terán, Dopico e Van Dooren dal quale abbiamo ripreso i nomi grado e pseudo grado, e la notazione  $\deg P$ ,  $\text{gr } P$  [5].<sup>1</sup>

Dobbiamo dare ora una buona definizione di autovalori e autovettori. Infatti vogliamo capire come estendere questi concetti quando studiamo matrici rettangolari: la definizione che si basa sul determinante perde infatti ogni significato. Ci viene in aiuto la Forma di Smith, per la quale è necessario definire il concetto di *matrice unimodulare*. La dimostrazione è omessa, ma può essere trovata su un buon manuale di algebra.

**Definizione 2.1.3** (Matrice unimodulare). Una matrice di polinomi  $E(\lambda) \in \mathbb{F}[\lambda]^{m \times m}$  è unimodulare se  $\det E(\lambda)$  è un elemento di  $\mathbb{F}$  diverso da 0. Equivalientemente se  $E(\lambda)^{-1}$  è ancora una matrice di polinomi.

**Teorema 2.1.4** (Forma di Smith). Sia  $P(\lambda)$  una matrice di polinomi  $m \times n$  sul campo  $\mathbb{F}$ . Esistono  $r \in \mathbb{N}$ ,  $E(\lambda)$  e  $F(\lambda)$  matrici unimodulari di dimensioni rispettivamente  $m \times m$  e  $n \times n$  tali che:

$$E(\lambda)P(\lambda)F(\lambda) = \text{diag}(d_1(\lambda), \dots, d_{\min\{m,n\}}(\lambda)) =: D(\lambda)$$

con

1.  $d_i(\lambda) \in \mathbb{F}[\lambda]$ ,  $\forall i$
2.  $d_i(\lambda)$  monici e  $d_i \mid d_{i+1}$ ,  $\forall i \leq r$
3.  $d_i(\lambda) = 0$ ,  $\forall i \quad r < i \leq \min\{m, n\}$

I polinomi non nulli  $d_i(\lambda)$  sono detti polinomi invarianti di  $P(\lambda)$ . Inoltre la matrice  $D(\lambda)$  è unica e il numero  $r$  è il rango di  $P(\lambda)$ , vista come matrice a coefficienti in  $\mathbb{F}(\lambda)$ .

**Osservazione 2.1.5.** L'unicità della matrice  $D(\lambda)$  del teorema 2.1.4 ci assicura che la forma di Smith è insensibile all'estensione di campi. Ne segue che le prossime definizioni sono ben poste.

**Definizione 2.1.6** (Sequenze delle molteplicità parziali). Sia  $P(\lambda)$  una matrice di polinomi  $m \times n$  di rango  $r$ . Per ogni  $\lambda_0 \in \overline{\mathbb{F}}$  consideriamo la decomposizione dei polinomi invarianti di  $P(\lambda)$ :

$$d_i(\lambda) = (\lambda - \lambda_0)^{a_i} p_i(\lambda) \quad a_i \geq 0, \quad p_i(\lambda_0) \neq 0$$

Per come è stata definita la forma di Smith, vale  $a_1 \leq a_2 \leq \dots \leq a_r$ . I fattori nella forma  $(\lambda - \lambda_0)^{a_i}$  sono detti *divisori elementari* di  $\lambda_0$ . La sequenza  $(a_1, a_2, \dots, a_r)$  è chiamata *sequenza delle molteplicità parziali* di  $P(\lambda)$  in  $\lambda_0$ .

<sup>1</sup> Dall'inglese *degree* e *grade*.

Ora possiamo estendere il concetto di autovalore a matrici rettangolari, in quanto il determinante di una matrice rettangolare perde significato.

**Definizione 2.1.7** (Autovalori e loro molteplicità). Data  $P(\lambda)$  come sopra e  $\alpha(\lambda) = (a_1(\lambda), a_2(\lambda), \dots, a_r(\lambda))$  la sequenza delle molteplicità parziali in funzione di  $\lambda \in \overline{\mathbb{F}}$ . Diciamo che  $\lambda_0$  è un *autovalore* finito per  $P(\lambda)$  se  $\alpha(\lambda_0)$  non è la sequenza nulla. Chiamiamo:

- *molteplicità geometrica* di  $\lambda_0$  il numero di  $a_i \neq 0$  in  $\alpha(\lambda_0)$ .
- *molteplicità algebrica* di  $\lambda_0$  la somma  $\sum_{i=1}^r a_i(\lambda_0)$ .

A livello di notazione sarà molto utile la prossima definizione.

**Definizione 2.1.8** (Molteplicità algebrica totale). Sia  $P(\lambda)$  una matrice di polinomi come nel teorema 2.1.4 e siano  $d_1, d_2, \dots, d_r$  i suoi polinomi invarianti. Chiamiamo *molteplicità algebrica totale*  $\delta_{\text{fin}}(P)$  la somma dei gradi dei suoi polinomi invarianti, ovvero:

$$\delta_{\text{fin}}(P) := \sum_{i=1}^r \deg d_i(\lambda)$$

$\delta_{\text{fin}}$  potrebbe essere definito anche come la somma delle molteplicità algebriche di tutti gli autovalori finiti: si preferisce tuttavia la formulazione precedente, in quanto non necessita di passare alla chiusura algebrica  $\overline{\mathbb{F}}$ . Ci riferiremo agli elementi delle precedenti definizioni con la locuzione introdotta in 2.1.9

**Definizione 2.1.9** (Struttura spettrale o di Jordan finita). L'insieme dei polinomi invarianti, delle sequenze delle molteplicità parziali e degli autovalori è chiamata *struttura spettrale o di Jordan finita*.

I polinomi di matrici possono avere anche *autovalori all'infinito*, come abbiamo visto nell'introduzione per il problema lineare agli autovalori. Prima di introdurla più formalmente, definiamo la *matrice rovesciata* di  $P(\lambda)$ .

**Definizione 2.1.10** (*J-rovesciata*). Sia  $P(\lambda)$  una matrice di polinomi di grado  $d \geq 0$ . Per ogni  $j \geq d$  definiamo la *j-rovesciata*

$$\text{rev}_j P(\lambda) = \lambda^j P \left( \frac{1}{\lambda} \right)$$

Se  $j = d$ , per semplicità di notazione, poniamo  $\text{rev}_j P = \text{rev } P$ .

**Definizione 2.1.11** (Struttura spettrale all'infinito). Sia  $P(\lambda)$  di grado  $d$  e pseudo grado  $g$ . La struttura spettrale all'infinito di  $P$  è definita dalla struttura spettrale di  $\text{rev}_g P$  in 0;  $\lambda_0 = \infty$  è *autovalore* di  $P$  se e solo se 0 è un autovalore di  $\text{rev}_g(P)$ . Le *molteplicità algebriche e geometriche* all'infinito sono definite analogamente. Se  $(a_1, a_2, \dots, a_r)$  è la sequenza delle molteplicità parziali di  $\text{rev}_g(P)$  in 0, chiamiamo  $\delta_\infty = a_1 + a_2 + \dots + a_r$ .

Ci riferiremo all'insieme di questi elementi con la locuzione *struttura infinita di Jordan*. Chiariamo con un esempio come lo pseudo grado influenzi la struttura all'infinito.

**Esempio 2.1.12.** Sia  $A \in \mathbb{F}^{n \times n}$  una matrice invertibile. Il polinomio  $P(\lambda) = A$  non ha alcun autovalore, né finito né infinito, in quanto  $\text{rev}_0 P(\lambda) = A$ . Se invece  $Q(\lambda) = \lambda 0_n + A$ , allora  $Q(\lambda)$  non ha alcun autovalore finito, ma  $\text{rev}_1 Q(\lambda) = \lambda A + 0_n$  possiede l'autovalore nullo: è facile convincersi che la sequenza delle molteplicità parziali in 0 sia  $(1, 1, \dots, 1)$ . Dunque  $Q(\lambda)$  ha l'autovalore all'infinito e  $\delta_\infty = n$ .

Nella prossima sezione esploreremo la struttura delle matrici polinomiali singolari.

## 2.2 Polinomi singolari e sottospazi di $\mathbb{F}(\lambda)^n$

Quando un articolo sui problemi agli autovalori si concentra sulle applicazioni numeriche, vengono considerati principalmente polinomi di matrici *regolari*: quest'ipotesi viene assunta nella prima pagina e mai più considerata. Tuttavia siamo interessati anche ai polinomi singolari, in modo da sviluppare una teoria quanto più omogenea possibile. In questa sezione sarà utile vedere una matrice di polinomi

$$P(\lambda) = \sum_{i=0}^k \lambda^i A_i \quad A_i \in \mathbb{F}^{m \times n}$$

come un'applicazione lineare (in una coppia di basi predefinita):

$$\phi: \mathbb{F}(\lambda)^n \mapsto \mathbb{F}(\lambda)^m.$$

Perciò è naturale estendere il concetto di autospazio nullo e nucleo per matrici di scalari  $A \in \mathbb{F}^{m \times n}$  a matrici di polinomi.

**Definizione 2.2.1** (Autospazi nulli). Sia  $P(\lambda) \in \mathbb{F}[\lambda]^{m \times n}$ . Identifichiamo con:

$$\begin{aligned} \mathcal{N}_r(P) &= \{x(\lambda) \in \mathbb{F}(\lambda)^{n \times 1}: P(\lambda)x(\lambda) = 0\} \\ \mathcal{N}_l(P) &= \{y(\lambda)^T \in \mathbb{F}(\lambda)^{1 \times m}: y(\lambda)^T P(\lambda) = 0\} \end{aligned}$$

gli autospazi nulli destro e sinistro di  $P$ .

Consideriamo ora un sottospazio vettoriale  $\mathcal{V}$  di  $\mathbb{F}(\lambda)^n$ . È sempre possibile trovare una base formata da vettori polinomiali: è sufficiente infatti prendere una base qualunque  $(v_1(\lambda), \dots, v_k(\lambda))$  e moltiplicare ciascun  $v_i$  per il minimo comune multiplo dei suoi denominatori.

**Definizione 2.2.2** (Grado di un vettore e ordine di una base). Dato  $v(\lambda)$  vettore polinomiale in  $\mathbb{F}(\lambda)^n$ , chiamiamo *grado* di  $v(\lambda)$  il massimo dei gradi delle sue componenti.

Sia  $\mathcal{V}$  un sottospazio vettoriale di  $\mathbb{F}(\lambda)^n$  e  $\mathcal{B} = (v_1(\lambda), \dots, v_k(\lambda))$  una sua base. L'*ordine* di  $\mathcal{B}$  è la somma dei gradi dei vettori che la compongono.

Ora possiamo introdurre:

**Definizione 2.2.3** (Base minimale). Sia  $\mathcal{V}$  un sottospazio di  $\mathbb{F}(\lambda)^n$ . Una base minimale di  $\mathcal{V}$  è una sua base polinomiale di minimo ordine tra tutte le possibili basi.

Sia  $\mathcal{B} = (v_1(\lambda), \dots, v_k(\lambda))$  una base minimale di  $\mathcal{V}$  con  $\deg v_i(\lambda) \leq \deg v_j(\lambda)$  se  $i < j$ . Si può dimostrare che la lista ordinata dei gradi  $(\deg v_i(\lambda))_{1 \leq i \leq k}$  dipende solamente da  $\mathcal{V}$ : sono chiamati *indici minimali* [7]. Noi saremo interessati solamente a  $\mathcal{V} = \mathcal{N}_r$  o  $\mathcal{V} = \mathcal{N}_l$ , quindi daremo una definizione più restrittiva di *indici minimali*:

**Definizione 2.2.4** (Indici minimali). Sia  $P(\lambda)$  come nelle definizioni precedenti. Siano rispettivamente  $(y_1(\lambda)^T, \dots, y_q(\lambda)^T)$  e  $(x_1(\lambda), \dots, x_p(\lambda))$  basi minimali di  $\mathcal{N}_l(P)$  e  $\mathcal{N}_r(P)$ , ordinate in modo che  $\deg y_1(\lambda) \leq \deg y_2(\lambda) \leq \dots \leq \deg y_q(\lambda)$  e, analogamente,  $\deg x_1(\lambda) \leq \deg x_2(\lambda) \leq \dots \leq \deg x_p(\lambda)$ . Chiamiamo  $\eta_i = \deg y_i(\lambda)$  e  $\varepsilon_i = \deg x_i(\lambda)$  gli *indici minimali sinistri e destri* di  $P(\lambda)$  e definiamo:

$$\mu(P) := \sum_{i=1}^p \varepsilon_i + \sum_{i=1}^q \eta_i.$$

Se  $P(\lambda)$  è regolare poniamo  $\mu(P) = 0$ . Ci riferiremo all'insieme degli *indici minimali*, sia destri che sinistri, con la perifrasi *struttura singolare* di  $P(\lambda)$ .

## 2.3 Relazioni di equivalenza

In questa sezione introdurremo diverse relazioni d'equivalenza in  $\mathbb{F}[\lambda]^{m \times n}$ , alcune divenute classiche in letteratura, altre relativamente recenti. Potremmo così dare una buona definizione di *linearizzazione*, il metodo principe per risolvere i problemi polinomiali agli autovalori.

**Definizione 2.3.1** (Equivalenza unimodulare). Siano  $P(\lambda), Q(\lambda) \in \mathbb{F}[\lambda]^{m \times n}$ . Si dicono *unimodularmente equivalenti*, e si scrive  $P(\lambda) \sim Q(\lambda)$ , se esistono due matrici unimodulari  $E(\lambda) \in \mathbb{F}[\lambda]^{m \times m}$  e  $F(\lambda) \in \mathbb{F}[\lambda]^{n \times n}$  tali che  $E(\lambda)P(\lambda)F(\lambda) = Q(\lambda)$ .

**Definizione 2.3.2** (Equivalenza stretta). Siano  $P(\lambda), Q(\lambda) \in \mathbb{F}[\lambda]^{m \times n}$ . Si dicono *strettamente equivalenti*, e si scrive  $P(\lambda) \cong Q(\lambda)$ , se esistono due matrici  $E \in \mathbb{F}^{m \times m}$ ,  $F \in \mathbb{F}^{n \times n}$  tali che  $E \cdot P(\lambda) \cdot F = Q(\lambda)$ .

Per definizione vale  $P(\lambda) \sim Q(\lambda)$  se e solo  $P(\lambda)$  e  $Q(\lambda)$  hanno la stessa forma di Smith (si veda 2.1.4). Dunque un insieme di invarianti completo è costituito dalla dimensione, dal rango e dai divisori elementari finiti. L'equivalenza stretta invece preserva anche la struttura spettrale all'infinito e la struttura singolare. Tuttavia questi non sono un sistema di invarianti completo. Le matrici

$$P(\lambda) = \begin{bmatrix} \lambda^2 & 0 \\ 0 & 1 \end{bmatrix} \quad Q(\lambda) = \begin{bmatrix} \lambda^2 & \lambda \\ 0 & 1 \end{bmatrix}$$

hanno la stessa struttura spettrale e singolare (quest'ultima è vuota), ma  $P(\lambda) \not\sim Q(\lambda)$ , in quanto  $P(\lambda)$  non ha alcun termine di primo grado. L'esistenza di una forma canonica per l'equivalenza stretta rimane tuttora un problema aperto: l'unico risultato disponibile, la *forma canonica di Kronecker*, richiede come ipotesi aggiuntiva che  $\mathbb{F}$  sia algebricamente chiuso e si applica solamente a matrici di polinomi lineari, dette anche *matrici pencil*. La riportiamo sotto forma di teorema, in quanto sarà necessaria per alcune dimostrazioni successive. Si può trovare una dimostrazione nell'articolo originale [13] e in [9]. Ricordiamo la definizione di somma diretta di matrici.

**Definizione 2.3.3** (Somma diretta). Sia  $\{A_i(\lambda)\}_{1 \leq i \leq k}$  una famiglia finita di polinomi di matrici  $m_i \times n_i$ . La somma diretta di  $\{A_i(\lambda)\}_{1 \leq i \leq k}$ , indicata con  $\bigoplus_{i=1}^k A_i(\lambda)$ , è la matrice diagonale a blocchi:

$$\bigoplus_{i=1}^k A_i(\lambda) = \text{diag}(A_1(\lambda), \dots, A_k(\lambda))$$

**Teorema 2.3.4** (Forma canonica di Kronecker). *Sia  $L(\lambda)$  una matrice pencil a coefficienti in un campo  $\mathbb{F}$  algebricamente chiuso. Allora  $P(\lambda)$  è strettamente equivalente ad una matrice che è somma diretta di un numero variabile dei blocchi seguenti:*

1.  $(\lambda - \lambda_0)I_k + N_k$ , dove  $N_k$  è la matrice nilpotente  $k \times k$ :

$$N_k := \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}.$$

2.  $I_m + \lambda N_m$ .
3.  $S_d(\lambda)$  o  $S_d^T(\lambda)$ , con  $d \geq 0$ , dove  $S_d(\lambda)$  è la matrice  $d \times (d+1)$

$$S_d(\lambda) := \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \end{bmatrix}$$



4.  $0_{p \times q}$ , con  $p, q \geq 0$ .

Si noti come ciascuno dei blocchi della forma canonica di Kronecker rappresenti una particolare proprietà della matrice  $L(\lambda)$ . I blocchi 1 e 2 rispecchiano la struttura spettrale tramite i divisori elementari finiti e infiniti; i blocchi 3 e 4 invece la struttura singolare, con gli indici minimali e il rango.

De Terán, Dopico e Mackey hanno introdotto due nuove relazioni di equivalenza che generalizzano quelle classiche [5].

**Definizione 2.3.5** (Equivalenza spettrale). Siano  $P$  e  $Q$  due matrici di polinomi di dimensione qualunque tali che  $\text{gr } P = g$  e  $\text{gr } Q = h$ .

- $P$  e  $Q$  si dicono *unimodularmente equivalenti in senso esteso* se esistono  $r, s \in \mathbb{N}$  tali che  $\text{diag}(P, I_r) \sim \text{diag}(Q, I_s)$ . Scriveremo  $P \smile Q$ .
- $P$  e  $Q$  si dicono *equivalenti in senso spettrale* se  $P \smile Q$  e  $\text{rev}_g P \smile \text{rev}_h Q$ . Scriveremo  $P \asymp Q$ .

Diamo ora la definizione formale di linearizzazione e, più in generale, di  $\ell$ -ficazione

**Definizione 2.3.6** (Linearizzazioni e  $\ell$ -ficazioni). Sia  $P(\lambda) \in \mathbb{F}[\lambda]^{m \times n}$  di pseudo grado  $\text{gr } P(\lambda) = g$ . Una matrice  $Q(\lambda)$  di pseudo grado  $\text{gr } Q(\lambda) = \ell$  è una  $\ell$ -ficazione di  $P(\lambda)$  se  $Q(\lambda) \smile P(\lambda)$ ; è una  $\ell$ -ficazione forte se  $Q(\lambda) \asymp P(\lambda)$ . Una  $\ell$ -ficazione con  $\ell = 1$  è detta *linearizzazione*.

Questa definizione di linearizzazione estende la definizione che ha preso piede in questi anni, riportata qui sotto.

**Definizione 2.3.7** (Linearizzazione classica). Sia  $P(\lambda) \in \mathbb{F}[\lambda]^{n \times n}$  un matrice di polinomi di grado  $k$ . Diciamo che una matrice  $L(\lambda)$  di taglia  $kn \times kn$  è una linearizzazione classica per  $P(\lambda)$  se vale:

$$L(\lambda) \sim \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_s \end{bmatrix}$$

con  $s = n(k - 1)$ . Diremo inoltre che  $L(\lambda)$  è una linearizzazione classica forte se è si verifica anche:

$$\text{rev}_1 L(\lambda) \sim \begin{bmatrix} \text{rev}_k P(\lambda) & 0 \\ 0 & I_s \end{bmatrix}$$

sempre con  $s = n(k - 1)$ .

La libertà offertaci da parametri  $r$  e  $s$  dell'equivalenza unimodulare ci consente persino di costruire linearizzazioni  $L(\lambda)$  di taglia minore rispetto alla matrice di partenza  $P(\lambda)$ , come mostrerà il prossimo esempio. Tuttavia i teoremi 2.4.3 e 2.4.4 ci forniranno forti limitazioni sulle possibili dimensioni delle linearizzazioni. Queste coincideranno con i risultati già noti per polinomi di matrici regolari.

**Esempio 2.3.8.** Sia  $P(\lambda) = \begin{bmatrix} \lambda & \lambda^2 \\ 0 & 1 \end{bmatrix}$  e  $L(\lambda) = [\lambda]$ , vista come matrice  $1 \times 1$ .

Allora vale

$$P(\lambda) \sim \begin{bmatrix} \lambda & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} L(\lambda) & 0 \\ 0 & I_1 \end{bmatrix}$$

da cui  $P(\lambda) \sim L(\lambda)$ : perciò  $L(\lambda)$  è una linearizzazione secondo 2.3.6, ma non secondo le equivalenze classiche.

Siamo interessati a caratterizzare queste equivalenze tramite le proprietà spettrali e le proprietà singolari delle matrici considerate. Iniziamo con un lemma.

**Lemma 2.3.9.** *Siano  $P(\lambda) \in \mathbb{F}[\lambda]^{m \times n}$  e  $Q(\lambda) \in \mathbb{F}[\lambda]^{p \times q}$  tali che  $P(\lambda) \sim Q(\lambda)$ . Allora  $m - p = n - q$ .*

*Dimostrazione.* È una conseguenza immediata della definizione. Infatti  $P(\lambda) \sim Q(\lambda)$  se e solo se esistono  $r, s \geq 0$  tali che  $\text{diag}(P(\lambda), I_r) \sim \text{diag}(Q(\lambda), I_s)$ . Dunque  $m + r = p + s$  e  $n + r = q + s$ , da cui la tesi.  $\square$

Dobbiamo poi appoggiarci ad un teorema dimostrato da Tisseur e Zaballa [20]. In effetti il risultato provato è più generale di quello qui riportato, ma quest'ultimo sarà più che sufficiente per il nostro scopo.

**Proposizione 2.3.10.** *Siano  $P, Q$  due matrici in  $\mathbb{F}[\lambda]^{m \times n}$  di pseudo grado rispettivamente  $g$  e  $h$  con gli stessi divisori elementari finiti e infiniti. Allora  $\text{rev}_g P$  e  $\text{rev}_h Q$  hanno gli stessi divisori elementari finiti.*

Possiamo ora dimostrare il

**Teorema 2.3.11** (Caratterizzazione di  $P \sim Q$  e  $P \asymp Q$ ). *Sia  $P(\lambda) \in \mathbb{F}[\lambda]^{m \times n}$  di pseudo grado  $g$  e  $Q(\lambda) \in \mathbb{F}[\lambda]^{p \times q}$  di pseudo grado  $h$ . Consideriamo le seguenti condizioni:*

(a)  $\dim \mathcal{N}_l(P) = \dim \mathcal{N}_l(Q)$  e  $\dim \mathcal{N}_r(P) = \dim \mathcal{N}_r(Q)$ .

(b)  $P$  e  $Q$  hanno gli stessi divisori elementari finiti.

(c)  $P$  e  $Q$  hanno gli stessi divisori elementari infiniti.

Allora:

1.  $P \sim Q$  se e solo se sono verificate (a) e (b).

2.  $P \asymp Q$  se e solo se sono verificate (a), (b), (c).

*Dimostrazione.*  $\boxed{1 \Rightarrow}$  Per definizione esistono  $r, s \geq 0$  tali che  $\text{diag}(P, I_r) \sim \text{diag}(Q, I_s)$ , ovvero  $\text{diag}(P, I_r)$  e  $\text{diag}(Q, I_s)$  hanno la stessa forma di Smith e hanno perciò gli stessi divisori elementari finiti. Inoltre:

$$\mathcal{N}_l(P) = \mathcal{N}_l(\text{diag}(P, I_r)) = \mathcal{N}_l(\text{diag}(Q, I_s)) = \mathcal{N}_l(Q)$$

e analogamente per i sottospazi destri, in quanto l'identità non influisce sulla struttura singolare dei polinomi di matrici.

$\boxed{1 \Leftarrow}$  Dall'uguaglianza sulle dimensioni dei nuclei ricaviamo che

$$m - \text{rank } P = p - \text{rank } Q, \quad n - \text{rank } P = q - \text{rank } Q$$

da cui segue

$$m - p = n - q = \text{rank } P - \text{rank } Q.$$

Esistono dunque  $r, s \geq 0$  tali che  $\text{diag}(P, I_r)$  e  $\text{diag}(Q, I_s)$  hanno la stessa dimensione. La condizione (a) ora implica che  $\text{diag}(P, I_r)$  e  $\text{diag}(Q, I_s)$  hanno lo stesso rango; da (b) deduciamo che anche i loro divisori elementari finiti coincidono (l'identità non influisce su quest'ipotesi). Perciò hanno la stessa forma di Smith, quindi  $P \simeq Q$ .

$\boxed{2 \Rightarrow}$  Supponiamo  $P \simeq Q$ . Allora per definizione  $P \simeq Q$ , dunque, per il punto  $\boxed{1}$ , valgono (a) e (b); analogamente  $\text{rev}_g P \simeq \text{rev}_h Q$ , quindi  $\text{rev}_g P$  e  $\text{rev}_h Q$  hanno gli stessi divisori elementari finiti, ovvero anche (c) è verificata.

$\boxed{2 \Leftarrow}$  Dal punto  $\boxed{1}$  segue che  $P \simeq Q$ . Dobbiamo dimostrare che  $\text{rev}_g P \simeq \text{rev}_h Q$ . Per la Proposizione 2.3.10 sappiamo che  $\text{rev}_g P$  e  $\text{rev}_h Q$  hanno gli stessi divisori elementari, ovvero che vale (b). Mostriamo ora che è verificata la condizione (a): questo ci condurrà alla tesi. Osserviamo che un polinomio  $P$  e il suo rovesciato hanno lo stesso rango, in quanto la taglia dei minori invertibili non è influenzata dall'operatore  $\text{rev}$ . Di conseguenza

$$\dim \mathcal{N}_l(\text{rev}_g P) = n - \text{rank}(\text{rev}_g P) = n - \text{rank}(P) = \dim \mathcal{N}_l(P)$$

Poiché  $P \simeq Q$

$$\dim \mathcal{N}_l(\text{rev}_g P) = \dim \mathcal{N}_l(P) = \dim \mathcal{N}_l(Q) = \dim \mathcal{N}_l(\text{rev}_h Q)$$

e analogamente si dimostrano le uguaglianze per i nuclei destri.  $\square$

Questo teorema giustifica il nome di equivalenza spettrale per  $P \simeq Q$ . Notiamo tuttavia che viene preservata anche una parte della struttura singolare. Nelle applicazioni numeriche, in cui i polinomi di matrici sono regolari, la condizione sui nuclei viene banalmente verificata, in quanto coincidono con l'insieme vuoto. Dunque la linearizzazione conserva gli autovalori finiti; se necessitiamo di informazioni su quello infinito dobbiamo trovarne una forte. Nella prossima sezione analizzeremo gli elementi invarianti per le tre relazioni d'equivalenza e, di conseguenza, mostreremo quali sono le possibili dimensioni di una linearizzazione.

## 2.4 Dimensione delle linearizzazioni

Abbiamo dimostrato in 2.3.11 che l'equivalenza  $P \simeq Q$  conserva la struttura spettrale della matrice  $P$ . Questo è in generale falso per matrici unimodularmente equivalenti, come possiamo osservare nel prossimo esempio.

**Esempio 2.4.1.** Consideriamo la matrice (già in forma di Smith)

$$P(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$

La sequenza delle molteplicità all'infinito è  $(0, 1, 1)$  e  $\delta_\infty = 2$ . Applichiamo ora una trasformazione unimodulare:

$$P(\lambda) \sim U(\lambda) \cdot P(\lambda) \cdot I_3 = \begin{bmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{bmatrix} = \begin{bmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{bmatrix} =: Q(\lambda)$$

Definendo

$$E(\lambda) := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ -\lambda & 1 & 0 \end{bmatrix}, \quad F(\lambda) := \begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & \lambda \\ 0 & 1 & 0 \end{bmatrix}$$

risulta

$$E(\lambda) \cdot \text{rev } Q(\lambda) \cdot F(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda^2 \end{bmatrix}.$$

Dunque  $\delta_\infty = 2$  è rimasta invariata, ma la sequenza delle molteplicità è diventata  $\alpha_\infty = (0, 0, 2)$ . In effetti Lancaster e Psarrakos avevano dimostrato che una linearizzazione unimodulare di una matrice  $P(\lambda)$  regolare conserva solamente  $\delta_\infty$ , ma  $\alpha_\infty$  può variare a piacere [14]. La situazione peggiora ulteriormente quando consideriamo matrici polinomiali *singolari*. Infatti la struttura spettrale all'infinito può mescolarsi con quella singolare. Riassumiamo questi comportamenti nel teorema 2.4.2, in cui per semplicità analizziamo polinomi di (pseudo) grado uno. Per comodità viene utilizzato l'*index sum theorem* 4.1.4, ma quest'ultimo sarà dimostrato in maniera indipendente, quindi non vi è circolarità.

**Teorema 2.4.2.** *Siano  $L_1(\lambda), L_2(\lambda) \in \mathbb{F}[\lambda]^{m \times n}$  di pseudo grado 1. Consideriamo le seguenti proprietà:*

- (a)  $L_1(\lambda)$  e  $L_2(\lambda)$  hanno gli stessi divisori elementari finiti.
- (b)  $\text{rank } L_1(\lambda) = \text{rank } L_2(\lambda)$ .
- (c)  $L_1(\lambda)$  e  $L_2(\lambda)$  hanno lo stesso numero di indici minimali sinistri e destri.
- (d)  $\delta_\infty(L_1) + \mu(L_1) = \delta_\infty(L_2) + \mu(L_2)$

Allora sono equivalenti:

1.  $L_1(\lambda) \sim L_2(\lambda)$ .

2.  $L_1(\lambda)$  e  $L_2(\lambda)$  hanno la stessa forma di Smith.
3.  $L_1(\lambda)$  e  $L_2(\lambda)$  soddisfano (a) e (b).
4.  $L_1(\lambda)$  e  $L_2(\lambda)$  soddisfano (a) e (c).
5.  $L_1(\lambda)$  e  $L_2(\lambda)$  soddisfano (a) e (d).

*Dimostrazione.* Abbiamo già visto in precedenza l'equivalenza tra gli enunciati 1, 2 e 3.  $\boxed{3 \Leftrightarrow 4}$  Segue dalla caratterizzazione degli indici minimali:

$$\begin{aligned} \#\{\text{Indici minimali sinistri di } L_i(\lambda)\} &= \dim \mathcal{N}_l(L_i) = m - \text{rank}(L_i) \\ \#\{\text{Indici minimali destri di } L_i(\lambda)\} &= \dim \mathcal{N}_r(L_i) = n - \text{rank}(L_i) \end{aligned}$$

$\boxed{3 \Leftrightarrow 5}$  Sfruttiamo infine l'uguaglianza del teorema 4.1.3

$$\text{rank } L = \delta_{\text{fin}} + \delta_{\infty} + \mu(L)$$

e abbiamo dimostrato la tesi.  $\square$

Il teorema 2.4.2 afferma che l'equivalenza unimodulare conserva, oltre ai divisori elementari finiti, solamente la quantità  $\delta_{\infty} + \mu$ : non ci dice nulla sui singoli addendi. È in effetti possibile trovare coppie di matrici per cui la struttura spettrale all'infinito si converta in struttura singolare. Gli stessi ragionamenti valgono per la relazione  $P \smile Q$ . Vediamo ora come tutto questo influenzi le dimensioni delle linearizzazioni.

**Teorema 2.4.3** (Possibili dimensioni delle linearizzazioni). *Sia  $P(\lambda) \in \mathbb{F}[\lambda]^{m \times n}$  una matrice di rango  $r$ . Siano rispettivamente  $p := m - r$  e  $q := n - r$  le dimensioni di  $\mathcal{N}_l(P)$  e  $\mathcal{N}_r(P)$ . Allora:*

1. *Esiste una linearizzazione di  $P(\lambda)$  di dimensione  $s_1 \times s_2$ , ovvero una matrice  $L(\lambda)$   $s_1 \times s_2$  tale che  $L(\lambda) \smile P(\lambda)$  se e solo se*

$$\begin{aligned} s_1 &\geq \delta_{\text{fin}}(P) + q, \\ s_2 &\geq \delta_{\text{fin}}(P) + p, \quad \text{con } s_1 - s_2 = q - p = m - n. \end{aligned}$$

*In particolare si ottiene la linearizzazione di taglia minima quando valgono le uguaglianze.*

2. *Per qualsiasi scelta degli indici minimali sinistri e destri  $0 \leq \eta_1 \leq \dots \leq \eta_q$ ,  $0 \leq \varepsilon_1 \leq \dots \leq \varepsilon_p$  e per ogni lista (anche vuota) delle molteplicità parziali all'infinito  $0 \leq t_1 \leq \dots \leq t_l$ , esiste una linearizzazione  $\tilde{L}(\lambda)$  di  $P(\lambda)$  avente codesta struttura spettrale e singolare. La dimensione di  $\tilde{L}(\lambda)$  è  $s_1 \times s_2$ , con*

$$\begin{aligned} s_1 &= \delta_{\text{fin}}(P) + q + \tilde{\mu} + \tilde{\delta}_{\infty}, \\ s_2 &= \delta_{\text{fin}}(P) + p + \tilde{\mu} + \tilde{\delta}_{\infty} \end{aligned}$$

dove  $\tilde{\mu} := \sum \eta_i + \sum \varepsilon_i$  e  $\tilde{\delta}_{\infty} := \sum t_i$ .

*Dimostrazione.*  $\boxed{1 \Rightarrow}$  Se  $L \sim P$  sappiamo da 2.3.9 che  $s_1 - m = s_2 - n$ , che è l'ultima condizione. Consideriamo  $L(\lambda)$  come matrice a coefficienti nella chiusura algebrica  $\overline{\mathbb{F}}$  e la sua forma canonica di Kronecker  $K(\lambda)$ ; abbiamo visto nell'Osservazione 2.1.5 che la forma di Smith è insensibile all'estensione di campi, dunque, per il teorema 2.4.2,  $K$  ha gli stessi divisori elementari finiti di  $P$ . Perciò in  $K$  vi sono dei blocchi che corrispondono a ciascun divisore elementare, che occupano in totale  $\delta_{\text{fin}}(P)$  righe e  $\delta_{\text{fin}}(P)$  colonne. Utilizzando sempre 2.4.2 deduciamo che  $K$  ha lo stesso numero di indici minimali sinistri e destri di  $P$ , dunque deve essere composta da almeno altre  $q$  righe e  $p$  colonne. Poiché l'equivalenza stretta conserva le dimensioni, otteniamo che  $L$  ha almeno  $\delta_{\text{fin}}(P) + q$  righe e  $\delta_{\text{fin}}(P) + p$  colonne.

$\boxed{1 \Leftarrow}$  Dato un polinomio scalare monico  $p(\lambda) = \lambda^k + a_{k-1}\lambda^{k-1} + \dots + a_0 \in \mathbb{F}[\lambda]$ , ricordiamo che la matrice companion di Frobenius associata è:

$$C_p(\lambda) := \lambda I_k + \begin{bmatrix} a_{k-1} & a_{k-2} & \cdots & a_0 \\ -1 & 0 & \cdots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & & -1 & 0 \end{bmatrix}$$

Sappiamo che  $C_p \sim \text{diag}(p(\lambda), I_{k-1})$ . Siano ora  $d_j(\lambda), \dots, d_r(\lambda)$  i polinomi invarianti non banali (ovvero di grado strettamente positivo) di  $P(\lambda)$ . Risulta che la somma diretta delle companion associate

$$F(\lambda) := C_{d_j} \oplus \dots \oplus C_{d_r}$$

ha gli stessi divisori elementari di  $P(\lambda)$  e  $F(\lambda)$  ha dimensione  $\delta_{\text{fin}}(P) \times \delta_{\text{fin}}(P)$ , in quanto  $\sum_{k=j}^r \deg d_k = \delta_{\text{fin}}(P)$ . Non ci resta che considerare

$$L(\lambda) := F(\lambda) \oplus I_\alpha \oplus 0_{p \times q},$$

dove  $\alpha \geq 0$  è un parametro libero. Il blocco degli zeri  $0_{p \times q}$  invece fa in modo che i nuclei sinistri e destri di  $L(\lambda)$  abbiano la stessa dimensione di quelli di  $P(\lambda)$ : quindi, per il teorema (2.4.2), otteniamo  $P(\lambda) \sim L(\lambda)$ .

$\boxed{2 \Leftrightarrow}$  Richiameremo ancora la forma canonica di Kronecker. In 2.3.4 avevamo definito

$$S_d(\lambda) := \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \end{bmatrix}_{d \times (d+1)}$$

che serviva a rappresentare gli indici minimali. Chiamiamo dunque

$$S(\lambda) := S_{\varepsilon_1}(\lambda) \oplus \dots \oplus S_{\varepsilon_p}(\lambda) \oplus S_{\eta_1}^T(\lambda) \oplus \dots \oplus S_{\eta_q}^T(\lambda)$$

e

$$\Omega(\lambda) := I_{\delta_\infty} + \lambda N$$

con  $N = N_{t_1} \oplus \cdots \oplus N_{t_k}$  e  $N_{t_i}$  matrice nilpotente di taglia  $t_i \times t_i$ . Le dimensioni di  $S(\lambda)$  e  $N(\lambda)$  sono rispettivamente  $(q + \tilde{\mu}) \times (p + \tilde{\mu})$  e  $\tilde{\delta}_\infty \times \tilde{\delta}_\infty$ . Dico che

$$\tilde{L}(\lambda) := F(\lambda) \oplus S(\lambda) \oplus \Omega(\lambda)$$

è la linearizzazione che cerchiamo, dove  $F(\lambda)$  è definita come sopra. Infatti per costruzione vale  $\tilde{L}(\lambda) \in \mathbb{F}[\lambda]^{s_1 \times s_2}$  ed è un polinomio di primo grado. Infine, grazie a  $F(\lambda)$  e a  $S(\lambda)$ , ha gli stessi divisori elementari finiti e gli stessi indici minimali. Dunque, per il teorema 2.4.2, vale  $P(\lambda) \sim \tilde{L}(\lambda)$ .  $\square$

Notiamo come la dimostrazione sia costruttiva. Il punto  $\boxed{2}$  ci permette dunque di creare degli esempi di matrici equivalenti in cui la struttura spettrale all'infinito diventa struttura singolare, come avevamo anticipato nel paragrafo precedente. Vale un teorema analogo per le linearizzazioni forti.

**Teorema 2.4.4** (Possibili dimensioni delle linearizzazioni forti). *Sia  $P(\lambda) \in \mathbb{F}[\lambda]^{m \times n}$  una matrice di rango  $r$ . Siano rispettivamente  $p := m - r$  e  $q := n - r$  le dimensioni di  $\mathcal{N}_l(P)$  e  $\mathcal{N}_r(P)$ . Se  $P(\lambda)$  è singolare, ovvero  $p + q > 0$ , allora:*

1. *Esiste una forte linearizzazione di  $P(\lambda)$  di dimensione  $s_1 \times s_2$ , ovvero una matrice  $L(\lambda) \in \mathbb{F}[\lambda]^{s_1 \times s_2}$  tale che  $L(\lambda) \asymp P(\lambda)$  se e solo se*

$$\begin{aligned} s_1 &\geq \delta_{\text{fin}}(P) + \delta_\infty(P) + q, \\ s_2 &\geq \delta_{\text{fin}}(P) + \delta_\infty(P) + p, \quad \text{con } s_1 - s_2 = q - p = m - n. \end{aligned}$$

*In particolare si ottiene la linearizzazione di taglia minima quando valgono le uguaglianze.*

2. *Per qualsiasi scelta degli indici minimali sinistri e destri  $0 \leq \eta_1 \leq \cdots \leq \eta_q$ ,  $0 \leq \varepsilon_1 \leq \cdots \leq \varepsilon_p$ , esiste una forte linearizzazione  $\tilde{L}(\lambda)$  di  $P(\lambda)$  avente codesta struttura spettrale e singolare. La dimensione di  $\tilde{L}(\lambda)$  è  $s_1 \times s_2$ , con*

$$\begin{aligned} s_1 &= \delta_{\text{fin}}(P) + \delta_\infty(P) + q + \tilde{\mu}, \\ s_2 &= \delta_{\text{fin}}(P) + \delta_\infty(P) + p + \tilde{\mu} \end{aligned}$$

dove  $\tilde{\mu} := \sum \eta_i + \sum \varepsilon_i$ .

*Viceversa, se  $P(\lambda)$  è regolare, allora esiste una linearizzazione forte  $L(\lambda)$  per  $P(\lambda)$  di dimensione  $s \times s$  se e sole se*

$$s = \delta_{\text{fin}} + \delta_\infty.$$

*Traccia.* Come si può intuire, la dimostrazione di questo teorema non si discosta molta da quella del precedente 2.4.3.  $\boxed{1 \Rightarrow}$  Se supponiamo  $P \asymp L$ , allora vengono conservati anche i divisori elementari infiniti (teorema 2.3.11);

quindi devono esistere in  $K$  dei blocchi relativi all'autovalore infinito che occupano  $\delta_\infty(P)$  righe e colonne.  $\boxed{1 \Leftarrow}$  Quest'implicazione segue direttamente dal punto  $\boxed{2}$  di 2.4.3, in quanto avevamo dimostrato l'esistenza di una linearizzazione per qualsiasi valore  $\tilde{\delta}_\infty$ : basta dunque porre  $\tilde{\delta}_\infty = \delta_\infty$  e  $P(\lambda) \asymp L(\lambda)$  segue da 2.3.11.

$\boxed{2 \Leftrightarrow}$  La costruzione è identica a quella del punto  $\boxed{2}$  del teorema 2.4.3, da cui riprendiamo la notazione. È sufficiente considerare  $\tilde{L}(\lambda) = F(\lambda) \oplus S(\lambda) \oplus \Omega(\lambda)$ , in cui l'unica differenza risiede nel fatto che la sequenza delle molteplicità all'infinito (e quindi  $\Omega(\lambda)$ ) è fissata. Sempre per il teorema 2.3.11, vale  $P(\lambda) \asymp \tilde{L}(\lambda)$ .

Sia infine  $P(\lambda)$  una matrice regolare e  $L(\lambda)$  una sua forte linearizzazione. Allora  $L(\lambda)$  è regolare, con  $\delta_{\text{fin}}(L) = \delta_{\text{fin}}(P)$  e  $\delta_\infty(L) = \delta_\infty(P)$  (stiamo sempre richiamando 2.3.11). Ripercorriamo ora i passi di 2.4.3: vediamo  $L(\lambda)$  come una matrice su  $\overline{\mathbb{F}}$ , dunque la sua forma canonica di Kronecker  $K$  (quindi  $L$  stessa) ha taglia  $s \times s$ , con  $s = \delta_{\text{fin}}(L) + \delta_\infty(L)$ .  $\square$

Le definizioni di forti linearizzazioni qui riportate, introdotte in [5], non sono uniche in letteratura. Ad esempio in [15] una linearizzazione forte di un polinomio regolare  $P(\lambda) \in F[\lambda]^{n \times n}$  di grado  $k$  è una matrice  $L(\lambda)$  pencil di taglia  $kn \times kn$  con la medesima struttura spettrale. Come dimostrato dai teoremi 2.3.11 e 2.4.4, le due equivalenze coincidono sulla classe delle matrici regolari; tuttavia quest'ultima non si estende alle matrici singolari, come possiamo osservare dal prossimo esempio.

**Esempio 2.4.5.** Sia  $P(\lambda) \in \mathbb{F}^{2 \times 2}$  la matrice singolare  $\text{diag}(\lambda^2, 0)$  e siano  $L_1(\lambda)$  e  $L_2(\lambda)$  le due matrici pencil

$$L_1(\lambda) := \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad L_2(\lambda) := \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & 1 & \lambda \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Allora  $P(\lambda)$ ,  $L_1(\lambda)$  e  $L_2(\lambda)$  hanno gli stessi divisori elementari finiti e infiniti, ovvero  $\lambda^2$  e  $L_1(\lambda)$  e  $L_2(\lambda)$  hanno la taglia richiesta  $kn \times kn$ ; tuttavia solo  $L_2(\lambda)$  è una linearizzazione forte nel senso delle Definizioni 2.3.6 e 2.3.7. In effetti  $L_1(\lambda)$  non è nemmeno una linearizzazione per  $P(\lambda)$ . Infatti  $(P(\lambda), I_2)$  ha rango 3, mentre  $L_1(\lambda)$  ha rango 2, dunque non può esistere una trasformazione unimodulare che porti una nell'altra.

Per amore di chiarezza confrontiamo le principali equivalenze qui considerate.

- *L'equivalenza stretta*  $\cong$  (Def 2.3.2) è troppo rigida. Preserva sì tutta la struttura spettrale e singolare, ma non permette cambiamenti né di dimensione né di grado. È quindi poco utile per le  $\ell$ -ficazioni.



- *L'equivalenza unimodulare*  $\sim$  (Def 2.3.1) è troppo lasca, in quanto preserva sì la struttura di Jordan finita, ma può far perdere tutte le informazioni sulla quella singolare e su quella di Jordan infinita. Ci consente tuttavia di cambiare il grado e la dimensione del polinomio di matrice di partenza.
- *L'equivalenza spettrale*  $\asymp$  (Def 2.3.5) sarà quella che approfondiremo. Il grado e la dimensione non sono invarianti, permettendoci dunque di costruire  $\ell$ -ficazioni, mentre la struttura di Jordan, finita e infinita, viene conservata. Per polinomi singolari rimane irrisolto il problema sulla struttura singolare, su cui l'equivalenza spettrale non ha alcun controllo.

Nel prossimo capitolo mostreremo come costruire esplicitamente forti linearizzazioni, e in generale  $\ell$ -ficazioni, per polinomi di matrici  $P(\lambda) \in \mathbb{F}[\lambda]^{m \times n}$ .



## Capitolo 3

# Matrici companion

Fino adesso abbiamo analizzato linearizzazioni e  $\ell$ -ficazioni da un punto di vista meramente teorico. Ora vogliamo esporre qualche metodo concreto per costruirle. Nei teoremi 2.4.3 e 2.4.4 abbiamo dimostrato che esistono delle limitazioni sulla taglia delle linearizzazioni; vale un risultato analogo per i gradi, che dimostreremo nel teorema 4.2.4: dato  $P(\lambda)$  di pseudo grado  $k$ , non è infatti possibile costruire una forte  $\ell$ -ficazione per ogni  $\ell \in \mathbb{N}$ . Iniziamo con una definizione per le matrici companion. Riprende e generalizza quelle che si possono trovare su ogni testo di analisi numerica [9].

**Definizione 3.0.1** (Template). Sia  $\mathbb{F}[\lambda]_g^{m \times n}$  lo spazio dei polinomi di matrici di grado  $g$ . Identifichiamo  $\mathbb{F}[\lambda]_g^{m \times n}$  con  $\mathbb{F}^{(g+1)mn}$ , vedendo ogni entrata della matrice come una variabile indipendente. Un *template* è una funzione

$$\begin{aligned} \Psi : \quad \mathbb{F}[\lambda]_g^{m \times n} &\rightarrow \mathbb{F}[\lambda]_\ell^{p \times q}. \\ (x_1, \dots, x_{(g+1)mn}) &\mapsto C(\lambda) = \sum_{i=0}^{\ell} \lambda^i X_i, \end{aligned}$$

in cui ogni entrata di ogni coefficiente  $X_i$  è una funzione a valori scalari delle variabili  $x_j$ , dove le immagini possono solo essere di due tipi: o una costante  $\alpha \in \mathbb{F}$  o un multiplo di una sola variabile, ovvero una funzione  $\beta x_k$  per  $1 \leq k \leq (g+1)mn$ .

**Definizione 3.0.2** (Companion di grado  $\ell$ ). Una *companion di grado  $\ell$* , o brevemente una  $\ell$ -*companion*, per matrici  $P(\lambda) \in \mathbb{F}[\lambda]_g^{m \times n}$  è un template

$$\Psi : \mathbb{F}[\lambda]_g^{m \times n} \rightarrow \mathbb{F}[\lambda]_\ell^{p \times q},$$

tale che  $\Psi(P(\lambda))$  è una forte  $\ell$ -ficazione di  $P(\lambda)$  per ogni  $P(\lambda) \in \mathbb{F}[\lambda]_g^{m \times n}$ .

Qualora la  $\ell$ -ficazione fosse una linearizzazione, elideremo la  $\ell$  nel termine companion, sempre per non appesantire troppo la notazione. Analizzeremo in primo luogo le prime, e probabilmente più importanti, matrici companion, quelle di Frobenius. Mostriamo poi altre e più recenti  $\ell$ -ficazioni nella sezione 3.3.

**Osservazione 3.0.3.** Spesso le matrici  $\ell$ -companion possono avere ulteriori proprietà, utili quando dobbiamo effettivamente calcolare la struttura spettrale e singolare nelle applicazioni numeriche. Ad esempio:

- Gli autovettori, gli indici minimali e le basi minimali si conservano oppure possono essere recuperati con un basso costo computazionale.
- La struttura della matrice di partenza, come l'hermitianità o la simmetria, si preserva, comportando ancora un vantaggio a livello di calcolo esplicito.

### 3.1 Matrici companion di Frobenius

Consideriamo  $P(\lambda) = \lambda^k A_k + \lambda^{k-1} A_{k-1} + \cdots + A_0$ , con  $A_i \in \mathbb{F}^{m \times n}$ . Siano

$$X_1 := \begin{bmatrix} A_k & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix}, \quad X_0 := \begin{bmatrix} A_{k-1} & A_{k-2} & \cdots & A_0 \\ -I_n & 0 & \cdots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & & -I_n & 0 \end{bmatrix}$$

matrici di taglia  $s_1 \times t_1$ , con  $s_1 = m + n(k-1)$  e  $t_1 = nk$ . Allora definiamo  $C_1(\lambda) := \lambda X_1 + X_0$  la *prima companion di Frobenius*. Chiamiamo invece *seconda companion di Frobenius* la matrice  $C_2(\lambda) := \lambda Y_1 + Y_0$ , con

$$Y_1 = \begin{bmatrix} A_k & & & \\ & I_m & & \\ & & \ddots & \\ & & & I_m \end{bmatrix}, \quad Y_0 = \begin{bmatrix} A_{k-1} & -I_m & & 0 \\ A_{k-2} & 0 & \ddots & \\ \vdots & \vdots & \ddots & -I_m \\ A_0 & 0 & \cdots & 0 \end{bmatrix}$$

di taglia  $s_2 \times t_2$ , dove  $s_2 = km$  e  $t_2 = n + (k-1)m$ .

Se ci restringiamo a polinomi di matrici regolari, la scelta delle prima o della seconda companion di Frobenius è puramente stilistica. Tuttavia  $C_1(\lambda)$  e  $C_2(\lambda)$  godono di proprietà leggermente diverse, come dimostreremo nei teoremi 3.1.2 e 3.1.3. Riportiamo la dimostrazione di [5], che a sua volta è un sottocaso di un risultato più generale esposto in [4]. Nelle prossime dimostrazioni utilizzeremo spesso gli *shift di Horner*, che qui definiamo.

**Definizione 3.1.1** (Shift di Horner). Sia  $P(\lambda)$  un polinomio di matrici  $m \times n$  di pseudo grado  $k$ . Chiamiamo

$$P_i(\lambda) = \lambda^i A_k + \lambda^{i-1} A_{k-1} + \cdots + \lambda A_{k-i+1} + A_{k-i}$$

l' $i$ -esimo *shift di Horner*.

**Teorema 3.1.2.** Sia  $P(\lambda) = \sum_{i=0}^k \lambda^i A_i$  un polinomio di matrici appartenenti a  $\mathbb{F}^{m \times n}$  di pseudo grado  $k \geq 2$  e sia  $C_1(\lambda)$  la prima companion di Frobenius. Allora valgono i seguenti risultati:

1.  $C_1(\lambda) \asymp P(\lambda)$ .
2. (a) Supponiamo che  $(z_1(\lambda), \dots, z_p(\lambda))$  sia una base minimale destra di  $C_1(\lambda)$ , con i vettori partizionati in blocchi conformi ai blocchi colonna di  $C_1(P)$ ; sia inoltre  $x_j(\lambda)$  il  $k$ -esimo blocco  $n \times 1$  di  $z_j(\lambda)$ , con  $j = 1, \dots, p$ . Allora  $(x_1(\lambda), \dots, x_p(\lambda))$  è una base minimale destra di  $P(\lambda)$ .  
 (b) Supponiamo che  $(w_1(\lambda)^T, \dots, w_q(\lambda)^T)$  sia una base minimale sinistra di  $C_1(\lambda)$ , con i vettori partizionati in blocchi conformi ai blocchi riga di  $C_1(P)$ ; sia inoltre  $y_j(\lambda)^T$  il primo blocco  $1 \times m$  di  $w_j(\lambda)^T$ , con  $j = 1, \dots, q$ . Allora  $(y_1(\lambda)^T, \dots, y_q(\lambda)^T)$  è una base minimale sinistra di  $P(\lambda)$ .
3. (a) Se  $0 \leq \varepsilon_1 \leq \dots \leq \varepsilon_p$  sono gli indici minimali destri di  $P(\lambda)$ , allora

$$\varepsilon_1 + k - 1 \leq \dots \leq \varepsilon_p + k - 1$$

sono gli indici minimali destri di  $C_1(\lambda)$ .

- (b) Se  $0 \leq \eta_1 \leq \dots \leq \eta_q$  sono gli indici minimali sinistri di  $P(\lambda)$ , allora

$$\eta_1 \leq \dots \leq \eta_q$$

sono gli indici minimali sinistri di  $C_1(\lambda)$ .

*Dimostrazione.* 1 Per dimostrare che  $P(\lambda)$  e  $C_1(\lambda)$  sono equivalenti in senso spettrale, dobbiamo trovare le due coppie di matrici unimodulari che trasformano  $\text{diag}(P(\lambda), I_s)$  in  $C(\lambda)$  e  $\text{diag}(\text{rev}_k P(\lambda), I_{s'})$  in  $\text{rev}_1 C(\lambda)$  per qualche  $s, s' \in \mathbb{N}$ . Il procedimento utilizzato in tutte queste dimostrazioni consiste nell'applicare operazioni elementari per blocchi di righe e di colonne, componendo infine le matrici che le rappresentano. Per motivi di spazio riporteremo solamente le matrici unimodulari risultanti.

Sia  $P_i(\lambda)$  l' $i$ -esimo *shift di Horner*. Si noti che:

$$\begin{cases} P_0(\lambda) = A_k \\ P_k(\lambda) = P(\lambda) \\ P_{i+1}(\lambda) - \lambda P_i(\lambda) = A_{k-i+1} \quad \text{per } i = 0, \dots, k-1 \end{cases} \quad (3.1)$$

Siano ora:

$$S(\lambda) := \begin{bmatrix} I_m & P_1(\lambda) & P_2(\lambda) & \cdots & P_{k-1}(\lambda) \\ 0 & I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I_n & 0 \\ 0 & \cdots & \cdots & 0 & I_n \end{bmatrix}$$

e

$$R(\lambda) := \begin{bmatrix} \lambda^{k-1}I_n & -I_n & -\lambda I_n & \cdots & -\lambda^{k-2}I_n \\ \lambda^{k-2}I_n & 0 & -I_n & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -\lambda I_n \\ \lambda I_n & 0 & \cdots & 0 & -I_n \\ I_n & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

È immediato verificare che  $S(\lambda)$  è una matrice unimodulare  $s_1 \times s_1$ , con  $s_1 = m + n(k-1)$ ;  $R(\lambda)$  ha taglia  $t_1 \times t_1$ , con  $t_1 = kn$ ; per assicurarsi che sia unimodulare è sufficiente sviluppare per blocchi sull'ultima riga. Tramite un rapido conto e ricordandosi delle uguaglianze di (3.1) risulta:

$$S(\lambda)C_1(\lambda) = \begin{bmatrix} 0 & 0 & 0 & \cdots & P(\lambda) \\ -I_n & \lambda I_n & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & -I_n & -\lambda I_n & 0 \\ 0 & \cdots & 0 & -I_n & -\lambda I_n \end{bmatrix},$$

e quindi

$$S(\lambda)C_1(\lambda)R(\lambda) = \begin{bmatrix} P(\lambda) & & & & \\ & I_n & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & I_n \end{bmatrix},$$

da cui  $P(\lambda) \sim C_1(\lambda)$ . Non ci resta che dimostrare che  $\text{rev}_k P(\lambda) \sim \text{rev}_1 C_1(\lambda)$ . I passaggi sono esattamente gli stessi. Definiamo

$$\tilde{S}(\lambda) := \begin{bmatrix} I_m & -\tilde{P}_{k-2}(\lambda) & -\tilde{P}_{k-3}(\lambda) & \cdots & -\tilde{P}_0(\lambda) \\ 0 & I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I_n & 0 \\ 0 & \cdots & \cdots & 0 & I_n \end{bmatrix},$$

dove  $\tilde{P}_i(\lambda) = \lambda A_i + \lambda^2 A_{i-1} + \cdots + \lambda^{i+1} A_0$  per  $i = 0, 1, \dots, k-2$ , e

$$\tilde{R}(\lambda) := \begin{bmatrix} I_n & 0 & \cdots & \cdots & 0 \\ \lambda I_n & I_n & 0 & \cdots & 0 \\ \lambda^2 I_n & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \lambda I_n & I_n & 0 \\ \lambda^{k-1} I_n & \cdots & \lambda^2 I_n & \lambda I_n & I_n \end{bmatrix}.$$

Ora è ancora più evidente che  $\tilde{S}(\lambda)$  e  $\tilde{R}(\lambda)$  sono matrici unimodulari. Valgono inoltre le relazioni

$$\begin{cases} \tilde{P}_0(\lambda) = \lambda A_0 \\ \lambda \tilde{P}_{k-2}(\lambda) + \lambda A_{k-1} + A_k = \text{rev}_k P(\lambda) \\ \tilde{P}_{i+1}(\lambda) - \lambda \tilde{P}_i(\lambda) = \lambda A_i + 1 \end{cases} \quad \text{per } i = 0, \dots, k-3. \quad (3.2)$$

Dopo un rapido conto, sfruttando le uguaglianze in (3.2), otteniamo:

$$\tilde{S}(\lambda) \cdot \text{rev}_1 C(\lambda) = \begin{bmatrix} \text{rev}_k P(\lambda) & 0 & \cdots & 0 & 0 \\ -\lambda I_n & I_n & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & -\lambda I_n & I_n & 0 \\ 0 & \cdots & 0 & -\lambda I_n & I_n \end{bmatrix},$$

da cui

$$\tilde{S}(\lambda) \cdot \text{rev}_1 C(\lambda) \cdot \tilde{R}(\lambda) = \begin{bmatrix} \text{rev}_k P(\lambda) & & & & \\ & I_n & & & \\ & & \ddots & & \\ & & & I_n & \\ & & & & I_n \end{bmatrix},$$

indi la tesi.

**2a** e **3a** Notiamo che la struttura degli ultimi  $k-1$  blocchi riga implica che qualsiasi vettore polinomiale  $z(\lambda) \in \mathcal{N}_r(C_1)$  deve essere nella forma:

$$z(\lambda)^T = [\lambda^{k-1}x(\lambda) \quad \cdots \quad \lambda x(\lambda) \quad x(\lambda)] \quad (3.3)$$

per qualche  $x(\lambda) \in \mathbb{F}[\lambda]^n$ . Dunque, dal primo blocco riga di  $C_1(\lambda)$ , ricaviamo che  $z(\lambda) \in \mathcal{N}_r(C_1)$  se e solo se  $x(\lambda) \in \mathcal{N}_r(P)$ . Inoltre la struttura di  $z(\lambda)$  in (3.3) ci assicura che un insieme di vettori  $\{z_1(\lambda), \dots, z_j(\lambda)\}$  è linearmente indipendente se e solo se lo è il corrispondente  $\{x_1(\lambda), \dots, x_j(\lambda)\}$ . Questo induce una corrispondenza bigettiva tra le basi di  $\mathcal{N}_r(C_1)$  e quelle di  $\mathcal{N}_r(P)$ . Sempre da (3.3), se  $z(\lambda)$  e  $x(\lambda)$  sono vettori non nulli, abbiamo che  $\deg z(\lambda) = \deg x(\lambda) + k - 1$  e di conseguenza, per una coppia di basi corrispondenti, l'ordine (definizione 2.2.2) differisce esattamente di  $p(k-1)$ , dove  $p = \dim \mathcal{N}_r(P) = \mathcal{N}_r(C_1)$ . Concludiamo quindi che  $(z_1(\lambda), \dots, z_p(\lambda)) \in \mathcal{N}_r(C_1)$  è una base minimale destra di  $\mathcal{N}_r(C_1)$  se e soltanto se  $(x_1(\lambda), \dots, x_p(\lambda))$  lo è di  $\mathcal{N}_r(P)$ . Questo dimostra i punti (2a) e (3a).

**2b** e **3b** Ricordiamo che vale

$$C_1(\lambda) = S^{-1}(\lambda) \cdot \text{diag}(P(\lambda), I_n, \dots, I_n) \cdot R^{-1}(\lambda).$$

Per semplicità chiamiamo  $T(\lambda) = S^{-1} \cdot \text{diag}(P(\lambda), I_n, \dots, I_n)$ . Poiché  $S^{-1}(\lambda)$  e  $R^{-1}(\lambda)$  sono matrici unimodulari, risulta  $\mathcal{N}_l(C_1) = \mathcal{N}_l(T)$ . Ma

$$T(\lambda) = \begin{bmatrix} P(\lambda) & -P_1(\lambda) & -P_2(\lambda) & \cdots & -P_{k-1}(\lambda) \\ 0 & I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I_n & 0 \\ 0 & \cdots & \cdots & 0 & I_n \end{bmatrix},$$

dunque possiamo inferire la struttura dei vettori appartenenti a  $\mathcal{N}_l(C_1)$ :  $w(\lambda)^T \in \mathcal{N}_l(C_1)$  se e soltanto se è nella forma:

$$w(\lambda)^T = [y(\lambda)^T \quad y(\lambda)^T P_1(\lambda) \quad \cdots \quad y(\lambda)^T P_{k-1}(\lambda)], \quad (3.4)$$

per qualche  $y(\lambda)^T \in \mathcal{N}_l(P) \subseteq \mathbb{F}[\lambda]^{1 \times m}$ . Ripetendo il ragionamento del punto precedente, scopriamo che esiste una corrispondenza bigettiva tra le basi polinomiali di  $\mathcal{N}_l(C_1)$  e quelle di  $\mathcal{N}_l(P)$ . Dobbiamo ancora comprendere la connessione tra i gradi. A prima vista sembrerebbe che il grado di un vettore  $w(\lambda)^T$  possa essere maggiore di quello del vettore  $y(\lambda)^T$  corrispondente, per la presenza del blocco  $y(\lambda)^T P_{k-1}(\lambda)$ . Mostriamo invece che vale l'uguaglianza:

$$\deg w(\lambda)^T = \deg y(\lambda)^T. \quad (3.5)$$

Da questo seguirà la corrispondenza tra le basi minimali di  $\mathcal{N}_l(P)$  e di  $\mathcal{N}_l(C_1)$  e tra i rispettivi ordini, completando la dimostrazione degli ultimi due punti. Sia ora  $\hat{P}_j(\lambda) = \lambda^j A_j + \cdots + \lambda A_1 + A_0$  il troncamento al grado  $j$ -esimo di  $P(\lambda)$  e  $P_j(\lambda)$  il  $j$ -esimo shift di Horner, come definito in 3.1.1. Un rapido conto mostra che

$$P(\lambda) = \lambda^{k-i} P_i(\lambda) + \hat{P}_{k-i-1}(\lambda) = P(\lambda), \quad 1 \leq i \leq k-1.$$

Sia ora  $y(\lambda)^T \in \mathcal{N}_l(P)$  un qualsiasi vettore polinomiale per cui  $y(\lambda)^T P_i(\lambda) \neq 0$ . Allora:

$$\lambda^{k-i} y(\lambda)^T P_i(\lambda) = -y(\lambda)^T \hat{P}_{k-i-1}(\lambda)$$

e, considerandone solo i gradi:

$$\begin{aligned} k-i + \deg(y(\lambda)^T P_i(\lambda)) &= \deg(y(\lambda)^T \hat{P}_{k-i-1}(\lambda)) \\ &\leq (k-i-1) + \deg y(\lambda)^T, \end{aligned}$$

da cui  $\deg(y(\lambda)^T P_i(\lambda)) \leq \deg y(\lambda)^T$ . Perciò, per ogni vettore polinomiale  $y(\lambda)^T \in \mathcal{N}_l(P)$ , vale o  $y(\lambda)^T P_i(\lambda) = 0$  o  $\deg(y(\lambda)^T P_i(\lambda)) \leq \deg y(\lambda)^T$  per  $1 \leq i \leq k-1$  e in entrambi i casi l'uguaglianza (3.5) è verificata. Questo conclude la dimostrazione.  $\square$

Come anticipato vale un analogo teorema per la seconda companion di Frobenius. Riportiamo solo l'enunciato della proposizione, in quanto la dimostrazione ricalca le orme di quella appena vista.



**Teorema 3.1.3.** *Sia  $P(\lambda) = \sum_{i=0}^k \lambda^i A_i$  un polinomio di matrici appartenenti a  $\mathbb{F}^{m \times n}$  di pseudo grado  $k \geq 2$  e sia  $C_2(\lambda)$  la seconda companion di Frobenius. Allora valgono i seguenti risultati:*

1.  $C_2(\lambda) \asymp P(\lambda)$ .
2. (a) *Supponiamo che  $(z_1(\lambda), \dots, z_p(\lambda))$  sia una base minimale destra di  $C_2(\lambda)$ , con i vettori partizionati in blocchi conformi ai blocchi colonna di  $C_2(P)$ ; sia inoltre  $x_j(\lambda)$  il primo blocco  $n \times 1$  di  $z_j(\lambda)$ , con  $j = 1, \dots, p$ . Allora  $(x_1(\lambda), \dots, x_p(\lambda))$  è una base minimale destra di  $P(\lambda)$ .*  
 (b) *Supponiamo che  $(w_1(\lambda)^T, \dots, w_q(\lambda)^T)$  sia una base minimale sinistra di  $C_2(\lambda)$ , con i vettori partizionati in blocchi conformi ai blocchi riga di  $C_2(P)$ ; sia inoltre  $y_j(\lambda)^T$  il  $k$ -esimo blocco  $1 \times m$  di  $w_j(\lambda)^T$ , con  $j = 1, \dots, q$ . Allora  $(y_1(\lambda)^T, \dots, y_q(\lambda)^T)$  è una base minimale sinistra di  $P(\lambda)$ .*
3. (a) *Se  $0 \leq \varepsilon_1 \leq \dots \leq \varepsilon_p$  sono gli indici minimali destri di  $P(\lambda)$ , allora*

$$\varepsilon_1 \leq \dots \leq \varepsilon_p$$

*sono gli indici minimali destri di  $C_2(\lambda)$ .*

- (b) *Se  $0 \leq \eta_1 \leq \dots \leq \eta_q$  sono gli indici minimali sinistri di  $P(\lambda)$ , allora*

$$\eta_1 + k - 1 \leq \dots \leq \eta_q + k - 1$$

*sono gli indici minimali sinistri di  $C_2(\lambda)$ .*

I teoremi 3.1.2 e 3.1.3 manifestano le differenze tra le companion di Frobenius nel caso singolare: nella prima sono gli indici minimali destri ad essere traslati uniformemente di  $k - 1$ , dove  $k$  è lo pseudo grado di  $P(\lambda)$ , mentre quelli sinistri rimangono invariati; nella seconda i due ruoli si invertono.

**Osservazione 3.1.4.**  $C_1(\lambda)$  e  $C_2(\lambda)$  sono degli esempi molto più concreti di quello costruito in 2.4.5, che mostrano come due matrici possano essere spettralmente equivalenti (in quanto forti linearizzazioni di uno stesso polinomio) nonostante abbiano dimensioni diverse. Questa è un'ulteriore conferma di come le linearizzazioni di matrici singolari non abbiano una taglia fissata.

De Terán, Dopico e Mackey sono stati i primi a dimostrare questi risultati in un contesto così generale. Nella prossima sezione vedremo alcune (forti)  $\ell$ -ficcioni.

### 3.2 Matrici companion di grado $\ell > 1$

Per mostrare queste generalizzazioni, abbiamo bisogno di alcuni preliminari. Sia  $P(\lambda) = \lambda^k A_k + \dots + A_0$  il consueto polinomio di matrici di pseudo grado  $k$  a coefficienti in  $\mathbb{F}^{m \times n}$ . Come anticipato, mostreremo delle forti  $\ell$ -ficcazioni di  $P(\lambda)$ , a patto che  $\ell$  divida  $k$ . Ponendo  $k = \ell s$ , possiamo introdurre:

$$\begin{aligned} B_1(\lambda) &= \lambda^\ell A_\ell + \lambda^{\ell-1} A_{\ell-1} + \dots + A_0, \\ B_j(\lambda) &= \lambda^\ell A_{\ell j} + \lambda^{\ell-1} A_{\ell(j-1)} + \dots + \lambda A_{\ell(j-1)+1}, \quad \text{per } j = 2, \dots, s. \end{aligned} \quad (3.6)$$

Si noti come  $B_1(\lambda)$  sia formato da  $\ell + 1$  addendi, di cui uno di grado 0, mentre  $B_j(\lambda)$  per  $j \neq 1$  abbiano solo  $\ell$  termini di grado non nullo. Con un rapido conto si verifica che vale l'uguaglianza:

$$P(\lambda) = \lambda^{\ell(s-1)} B_s(\lambda) + \lambda^{\ell(s-2)} B_{s-1}(\lambda) + \dots + \lambda^\ell B_2(\lambda) + B_1(\lambda). \quad (3.7)$$

Utilizzando i polinomi  $B_1(\lambda), \dots, B_s(\lambda)$  possiamo introdurre la prossima definizione.

**Definizione 3.2.1** ( $\ell$ -Companion tipo Frobenius). Sia  $P(\lambda)$  un polinomio di matrici in  $\mathbb{F}[\lambda]^{m \times n}$  e  $B_1(\lambda), \dots, B_s(\lambda)$  come in (3.6). Allora chiamiamo

$$C_1^\ell(\lambda) := \begin{bmatrix} B_s(\lambda) & B_{s-1}(\lambda) & B_{s-2}(\lambda) & \cdots & B_1(\lambda) \\ -I_n & \lambda^\ell I_n & 0 & \cdots & 0 \\ & -I_n & \lambda^\ell I_n & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ 0 & & & -I_n & \lambda^\ell I_n \end{bmatrix} \in \mathbb{F}[\lambda]^{(m+(s-1)n) \times sn} \quad (3.8)$$

*prima  $\ell$ -companion tipo Frobenius e*

$$C_2^\ell(\lambda) := \begin{bmatrix} B_s(\lambda) & -I_m & & & 0 \\ B_{s-1}(\lambda) & \lambda^\ell I_m & -I_m & & \\ B_{s-2}(\lambda) & 0 & \lambda^\ell I_m & \ddots & \\ \vdots & \vdots & \ddots & \ddots & -I_m \\ B_1(\lambda) & 0 & \cdots & 0 & \lambda^\ell I_m \end{bmatrix} \in \mathbb{F}[\lambda]^{sm \times (n+(s-1)m)} \quad (3.9)$$

*seconda  $\ell$ -companion tipo Frobenius.*

Si noti che, ponendo  $\ell = 1$  e  $s = k$ , le  $C_1^1(\lambda)$  e  $C_2^1(\lambda)$  non coincidono esattamente con  $C_1(\lambda)$  e  $C_2(\lambda)$ : infatti nelle seconde i termini di primo grado si trovano solo sulla diagonale, mentre nelle prime  $B_2(\lambda), \dots, B_k(\lambda)$  hanno tutte grado non nullo. Il fatto che 3.2.1 sia una buona definizione è oggetto dei teoremi 3.2.2 e 3.2.3.

**Teorema 3.2.2.** *Sia  $P(\lambda) = \sum_{i=0}^k \lambda^i A_i \in \mathbb{F}[\lambda]^{m \times n}$  di pseudo grado  $k \geq 2$ . Sia  $1 \leq \ell < k$  un divisore di  $k$  tale che  $k = \ell s$  e sia  $C_1^\ell(\lambda)$  la prima  $\ell$ -companion tipo Frobenius. Allora valgono i seguenti risultati:*

1.  $C_1^\ell(\lambda) \simeq P(\lambda)$ .
2. (a) *Supponiamo che  $(z_1(\lambda), \dots, z_p(\lambda))$  sia una base minimale destra di  $C_1^\ell(\lambda)$ , con i vettori partizionati in blocchi conformi ai blocchi colonna di  $C_1^\ell(P)$ ; sia inoltre  $x_j(\lambda)$  il blocco  $s$ -esimo  $n \times 1$  di  $z_j(\lambda)$ , con  $j = 1, \dots, p$ . Allora  $(x_1(\lambda), \dots, x_p(\lambda))$  è una base minimale destra di  $P(\lambda)$ .*  
 (b) *Supponiamo che  $(w_1(\lambda)^T, \dots, w_q(\lambda)^T)$  sia una base minimale sinistra di  $C_1^\ell(\lambda)$ , con i vettori partizionati in blocchi conformi ai blocchi riga di  $C_1^\ell(P)$ ; sia inoltre  $y_j(\lambda)^T$  il primo blocco  $1 \times m$  di  $w_j(\lambda)^T$ , con  $j = 1, \dots, q$ . Allora  $(y_1(\lambda)^T, \dots, y_q(\lambda)^T)$  è una base minimale sinistra di  $P(\lambda)$ .*
3. (a) *Se  $0 \leq \varepsilon_1 \leq \dots \leq \varepsilon_p$  sono gli indici minimali destri di  $P(\lambda)$ , allora*

$$\varepsilon_1 + k - \ell \leq \dots \leq \varepsilon_p + k - \ell$$

*sono gli indici minimali destri di  $C_1^\ell(\lambda)$ .*

- (b) *Se  $0 \leq \eta_1 \leq \dots \leq \eta_q$  sono gli indici minimali sinistri di  $P(\lambda)$ , allora*

$$\eta_1 \leq \dots \leq \eta_q$$

*sono gli indici minimali sinistri di  $C_1^\ell(\lambda)$ .*

*Dimostrazione.* Come si può intuire, la dimostrazione è molto simile a quella del teorema 3.1.2. 1 Iniziamo definendo i polinomi:

$$\begin{aligned} Q_0(\lambda) &:= B_s(\lambda), \\ Q_j(\lambda) &:= \lambda^j Q_{j-1}(\lambda) + B_{s-j}(\lambda), \quad \text{per } j = 1, \dots, s-1, \end{aligned} \tag{3.10}$$

che ricoprono il medesimo ruolo degli *shift di Horner* nella dimostrazione di 3.1.2. Vale l'uguaglianza:

$$Q_{s-1}(\lambda) = P(\lambda),$$

che è una semplice riscrittura di (3.7). Consideriamo le matrici unimodulari

$$S_\ell(\lambda) := \begin{bmatrix} I_m & Q_0(\lambda) & Q_1(\lambda) & \cdots & Q_{s-2}(\lambda) \\ 0 & I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I_n & 0 \\ 0 & \cdots & \cdots & 0 & I_n \end{bmatrix}$$

e

$$R_\ell(\lambda) := \begin{bmatrix} \lambda^{(s-1)\ell} I_n & -I_n & -\lambda^\ell I_n & \cdots & -\lambda^{(s-2)\ell} I_n \\ \lambda^{(s-2)\ell} I_n & 0 & -I_n & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -\lambda^\ell I_n \\ \lambda^\ell I_n & 0 & \cdots & 0 & -I_n \\ I_n & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

rispettivamente di taglia  $(m + (s-1)n)$  e  $sn$ . Un conto simile a quello svolto in 3.1.2 conduce all'identità

$$S_\ell(\lambda) \cdot C_1^\ell(\lambda) \cdot R_\ell(\lambda) = \text{diag}(P(\lambda), I_n, \dots, I_n),$$

da cui  $C_1^\ell \sim P(\lambda)$ . Il procedimento è per dimostrare  $\text{rev}_\ell C_1^\ell(\lambda) \sim \text{rev}_k P(\lambda)$  è leggermente diverso. Dobbiamo prima di tutto osservare che la matrice  $\text{rev}_\ell C_1^\ell(\lambda)$  è della forma

$$\begin{bmatrix} \text{rev}_\ell B_s(\lambda) & \text{rev}_\ell B_{s-1}(\lambda) & \text{rev}_\ell B_{s-2}(\lambda) & \cdots & \text{rev}_\ell B_1(\lambda) \\ -\lambda^\ell I_n & I_n & 0 & \cdots & 0 \\ & -\lambda^\ell I_n & I_n & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ 0 & & & -\lambda^\ell I_n & I_n \end{bmatrix}$$

e definire

$$\begin{aligned} \tilde{Q}_0 &:= \text{rev}_\ell B_1(\lambda) \\ \tilde{Q}_j &:= \lambda^\ell \tilde{Q}_{j-1}(\lambda) + \text{rev}_\ell B_{j+1}(\lambda), \quad j = 1, \dots, s-1, \end{aligned} \tag{3.11}$$

affinché

$$\tilde{Q}_{s-1}(\lambda) = \text{rev}_k P(\lambda).$$

Questa volta le matrici unimodulari che compiono la trasformazione sono:

$$\tilde{S}_\ell := \begin{bmatrix} I_m & -\tilde{Q}_{s-2}(\lambda) & -\tilde{Q}_{s-3}(\lambda) & \cdots & -\tilde{Q}_0(\lambda) \\ 0 & I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I_n & 0 \\ 0 & \cdots & \cdots & 0 & I_n \end{bmatrix},$$

e

$$\tilde{R}_\ell := \begin{bmatrix} I_n & 0 & \cdots & \cdots & 0 \\ \lambda^\ell I_n & I_n & 0 & \cdots & 0 \\ \lambda^{2\ell} I_n & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \lambda I_n & I_n & 0 \\ \lambda^{(s-1)\ell} I_n & \cdots & \lambda^{2\ell} I_n & \lambda^\ell I_n & I_n \end{bmatrix},$$

rispettivamente di taglia  $(m + (s - 1)n)$  e  $sn$ . Sfruttando le uguaglianze appena definite otteniamo

$$\tilde{S}_\ell(\lambda) \cdot \text{rev}_\ell C_1^\ell(\lambda) \cdot \tilde{R}_\ell(\lambda) = \text{diag}(\text{rev}_k P(\lambda), I_n, \dots, I_n),$$

per cui  $C_1^\ell(\lambda) \simeq P(\lambda)$ , come volevamo.

**2a** e **3a** La struttura a blocchi di  $C_1^\ell(\lambda)$  implica che  $z(\lambda) \in \mathcal{N}_r(C_1^\ell)$  se e solo se è nella forma

$$z(\lambda)^T = [\lambda^{(s-1)\ell} x(\lambda) \quad \dots \quad \lambda^\ell x(\lambda) \quad x(\lambda)] \quad (3.12)$$

per qualche  $x(\lambda) \in \mathcal{N}_r(P)$ ; la forma di  $z(\lambda)$  ci assicura che è un vettore polinomiale se e solo se lo è  $x(\lambda)$ ; inoltre un insieme di vettori  $\{z_1(\lambda), \dots, z_j(\lambda)\}$  è linearmente indipendente se e solo se lo è la corrispondente  $(x_1(\lambda), \dots, x_j(\lambda))$ . Tutto questo induce una corrispondenza bigettiva tra le basi di  $\mathcal{N}_r(C_1^\ell)$  e quelle di  $\mathcal{N}_r(P)$ . Sempre da (3.12), se  $z(\lambda)$  e  $x(\lambda)$  sono vettori non nulli, abbiamo che  $\deg z(\lambda) = \deg x(\lambda) + \ell(s - 1) = \deg x(\lambda) + k - \ell$ . Da questo momento la dimostrazione prosegue come nel teorema 3.1.2, sostituendo semplicemente  $k - \ell$  a  $k - 1$ .

**2b** e **3b** La dimostrazione ricalca nuovamente 3.1.2. Sia

$$T_\ell(\lambda) := S_\ell^{-1}(\lambda) \text{diag}(P(\lambda), I_n, \dots, I_n).$$

Ricordando l'uguaglianza:

$$C_1^\ell(\lambda) = S_\ell^{-1}(\lambda) \cdot \text{diag}(P(\lambda), I_n, \dots, I_n) \cdot R_\ell^{-1}(\lambda),$$

risulta  $\mathcal{N}_l(C_1^\ell) = \mathcal{N}_l(T_\ell)$ . Un conto esplicito rivela la struttura di  $T_\ell(\lambda)$ :

$$T_\ell = \begin{bmatrix} P(\lambda) & -Q_0(\lambda) & -Q_1(\lambda) & \dots & Q_{s-2}(\lambda) \\ 0 & I_n & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & I_n & 0 \\ 0 & \dots & \dots & 0 & I_n \end{bmatrix}.$$

Da questo otteniamo la caratterizzazione dei vettori appartenenti a  $\mathcal{N}_l(C_1^\ell)$ :  $w(\lambda)^T \in \mathcal{N}_l(C_1^\ell)$  se e soltanto se è nella forma:

$$w(\lambda)^T = [y(\lambda)^T \quad y(\lambda)^T Q_0(\lambda) \quad \dots \quad y(\lambda)^T Q_{s-2}(\lambda)], \quad (3.13)$$

per qualche  $y(\lambda)^T \in \mathcal{N}_l(P) \subseteq \mathbb{F}[\lambda]^{1 \times m}$ . Se proviamo che  $\deg w(\lambda) = \deg y(\lambda)$  per ogni  $w(\lambda)^T \in \mathcal{N}_l(C_1^\ell)$ , possiamo concludere utilizzando gli stessi argomenti del teorema 3.1.2. È evidente che  $\deg w(\lambda) \leq \deg y(\lambda)$ , dunque proviamo l'altra disuguaglianza. Chiamiamo

$$\hat{Q}_j(\lambda) := \lambda^{\ell(s-(j+2))} B_{s-(j+1)}(\lambda) + \dots + \lambda^\ell B_2(\lambda) + B_1(\lambda)$$

il troncamento di  $P(\lambda)$  di grado  $\ell(s - (j + 2)) + \ell$ . Dalle identità (3.7) e (3.11) otteniamo:

$$\lambda^{\ell(s-(j+1))}Q_j(\lambda) + \hat{Q}_j(\lambda) = P(\lambda), \quad j = 0, 1, \dots, s - 2.$$

Perciò, se  $y(\lambda)^T \in \mathcal{N}_\ell(P)$  e  $y(\lambda)^T Q_j(\lambda) \neq 0$ , vale:

$$\lambda^{\ell(s-(j+1))}y(\lambda)^T Q_j(\lambda) = -y(\lambda)^T \hat{Q}_j(\lambda)$$

Prendendo i gradi da entrambi i membri e ricordando che  $\deg B_{s-(j+1)}(\lambda) = \ell$ :

$$\begin{aligned} \ell(s - (j + 1)) + \deg(y(\lambda)^T Q_j(\lambda)) &= \deg(y(\lambda)^T \hat{Q}_j(\lambda)) \\ &\leq \deg y(\lambda) + \ell(s - (j + 2)) + \ell, \end{aligned}$$

da cui  $\deg(y(\lambda)^T Q_j(\lambda)) \leq \deg y(\lambda)$  per  $j = 0, 1, \dots, s - 2$ . Quindi, per ogni vettore polinomiale  $y(\lambda)^T \in \mathcal{N}_\ell(P)$ , vale o  $\deg(y(\lambda)^T Q_j(\lambda)) \leq \deg y(\lambda)^T$  o  $y(\lambda)^T Q_j(\lambda) = 0$  per  $0 \leq j \leq s - 2$  e in entrambi casi  $\deg w(\lambda) = \deg y(\lambda)$  è verificata. Questo conclude la dimostrazione.  $\square$

Così come fatto per il teorema 3.1.3, anche per 3.2.3 riportiamo solo l'enunciato. Infatti le differenze con il teorema 3.2.2 rispecchiano quelle tra 3.1.2 e 3.1.3.

**Teorema 3.2.3.** *Sia  $P(\lambda) = \sum_{i=0}^k \lambda^i A_i \in \mathbb{F}[\lambda]^{m \times n}$  di pseudo grado  $k \geq 2$ . Sia  $1 \leq \ell < k$  un divisore di  $k$  tale che  $k = \ell s$  e sia  $C_2^\ell(\lambda)$  definita in 3.2.1. Allora valgono i seguenti risultati:*

$$C_2^\ell(\lambda) \simeq P(\lambda).$$

1. *Supponiamo che  $(z_1(\lambda), \dots, z_p(\lambda))$  sia una base minimale destra di  $C_2^\ell(\lambda)$ , con i vettori partizionati in blocchi conformi ai blocchi colonna di  $C_2^\ell(\lambda)$ ; sia inoltre  $x_j(\lambda)$  il primo blocco  $n \times 1$  di  $z_j(\lambda)$ , con  $j = 1, \dots, p$ . Allora  $(x_1(\lambda), \dots, x_p(\lambda))$  è una base minimale destra di  $P(\lambda)$ .*
2. *Supponiamo che  $(w_1(\lambda)^T, \dots, w_q(\lambda)^T)$  sia una base minimale sinistra di  $C_2^\ell(\lambda)$ , con i vettori partizionati in blocchi conformi ai blocchi riga di  $C_2^\ell(\lambda)$ ; sia inoltre  $y_j(\lambda)^T$  il blocco  $s$ -esimo  $1 \times m$  di  $w_j(\lambda)^T$ , con  $j = 1, \dots, q$ . Allora  $(y_1(\lambda)^T, \dots, y_q(\lambda)^T)$  è una base minimale sinistra di  $P(\lambda)$ .*
1. *Se  $0 \leq \varepsilon_1 \leq \dots \leq \varepsilon_p$  sono gli indici minimali destri di  $P(\lambda)$ , allora*

$$\varepsilon_1 \leq \dots \leq \varepsilon_p$$

*sono gli indici minimali destri di  $C_2^\ell(\lambda)$ .*

2. *Se  $0 \leq \eta_1 \leq \dots \leq \eta_q$  sono gli indici minimali sinistri di  $P(\lambda)$ , allora*

$$\eta_1 + k - \ell \leq \dots \leq \eta_q + k - \ell$$

*sono gli indici minimali sinistri di  $C_2^\ell(\lambda)$ .*

Riportiamo ora un semplice corollario, che vale la pena sottolineare. In letteratura una 2-ficcazione viene chiamata *quadrattizzazione*.<sup>1</sup>

**Corollario 3.2.4.** *Ogni polinomio matriciale di pseudo grado pari ammette una forte quadrattizzazione.*

Dal punto di vista applicativo la quadrattizzazione è sicuramente la più importante  $\ell$ -ficcazione dopo le linearizzazioni. Infatti esistono metodi diretti per affrontare un problema quadratico agli autovalori senza passare al suo linearizzato. Citiamo solamente l'algoritmo di Ehrlich-Aberth o il metodo di Bernoulli, che si basa sulla teoria dei solutori [11].

Si potrebbe pensare di introdurre nuovi algoritmi per risolvere un problema polinomiale agli autovalori  $P(\lambda)$  e affiancarli a quelli oggi utilizzati, ovvero:

- *metodo diretto*: agire direttamente su  $P(\lambda)$  senza nessuna trasformazione intermedia, come l'iterazione di Ehrlich-Aberth.
- *$\ell$ -ficcazione più metodo diretto*: trasformare  $P(\lambda)$  in un problema equivalente tramite una  $\ell$ -ficcazione (ad esempio  $\ell = 2$ ) e applicare un metodo diretto su quest'ultimo.
- *linearizzazione*: linearizzare  $P(\lambda)$  e sfruttare l'algoritmo QZ.

Al momento in letteratura non sono stati confrontati a livello numerico gli ultimi due approcci, soprattutto perché la  $\ell$ -ficcazione è un concetto recente [5]. Sarebbe interessante studiare come variano i numeri di condizionamento e il tempo di calcolo, magari per specifiche classi di matrici strutturate.

### 3.3 Altri tipi di $\ell$ -ficcazioni

In questa sezione affronteremo altre  $\ell$ -ficcazioni che differiscono da quelle appena presentate. Nel primo caso non sarà soddisfatta l'ipotesi di equivalenza spettrale, mentre nel secondo la costruzione sarà leggermente più articolata. Un vantaggio di una semplice  $\ell$ -ficcazione  $W^\ell(\lambda)$  rispetto alla companion tipo Frobenius è l'assenza del vincolo sul grado:  $\ell$  non deve necessariamente dividere  $k$ . È vero, perdiamo informazioni sulla struttura spettrale all'infinito, ma se queste per qualche motivo non ci sono utili (come in molte applicazioni numeriche) possiamo propendere per  $W^\ell(\lambda)$ .

---

<sup>1</sup>In inglese *quadrattification* oppure *quadrattization*, ma per la traduzione italiana abbiamo optato per la seconda forma

Sia quindi  $P(\lambda) = \lambda^k A_k + \cdots + A_0$  un polinomio di pseudo grado  $k$  e sia  $\ell < k$ . Definiamo:

$$W_1^\ell(\lambda) := \begin{bmatrix} P_\ell(\lambda) & A_{k-\ell-1} & \cdots & A_1 & A_0 \\ -I_n & \lambda I_n & & 0 & 0 \\ & \ddots & \ddots & \vdots & \vdots \\ & & -I_n & \lambda I_n & 0 \\ 0 & & & -I_n & \lambda I_n \end{bmatrix} \in \mathbb{F}[\lambda]^{(m+n(k-\ell)) \times n(k-\ell+1)} \quad (3.14)$$

e

$$W_2^\ell(\lambda) := \begin{bmatrix} P_\ell(\lambda) & -I_m & & & 0 \\ A_{k-\ell-1} & \lambda I_m & -I_m & & \\ \vdots & & \ddots & \ddots & \\ A_1 & 0 & \cdots & \lambda I_m & -I_m \\ A_0 & 0 & \cdots & 0 & \lambda I_m \end{bmatrix} \in \mathbb{F}[\lambda]^{m(k-\ell+1) \times (n+m(k-\ell))}, \quad (3.15)$$

dove  $P_\ell(\lambda)$  è lo shift di Horner  $\ell$ -esimo (definizione 3.1.1). Si noti che  $W_1^1(\lambda) = C_1(\lambda)$  e  $W_2^1(\lambda) = C_2(\lambda)$ , dunque sarebbero proprio  $W_1^\ell(\lambda)$  e  $W_2^\ell(\lambda)$  le generalizzazioni delle companion di Frobenius. Tuttavia, a livello di “notazione”, sono state preferite  $C_1^\ell(\lambda)$  e  $C_2^\ell(\lambda)$  in quanto sono forti  $\ell$ -ficazioni.

Riportiamo gli enunciati dei teoremi che espongono le proprietà di queste matrici, in analogia con 3.2.2 e 3.2.3. Si può trovare lo schema della dimostrazione di 3.3.1 in [5].

**Teorema 3.3.1.** *Sia  $P(\lambda) = \sum_{i=0}^k \lambda^i A_i \in \mathbb{F}[\lambda]^{m \times n}$  un polinomio matriciale di pseudo grado  $k \geq 2$ . Siano  $1 \leq \ell < k$  e  $W_1^\ell(\lambda)$  definita in (3.14). Allora valgono i seguenti risultati:*

1.  $W_1^\ell(\lambda) \sim P(\lambda)$ .
2. (a) *Supponiamo che  $(z_1(\lambda), \dots, z_p(\lambda))$  sia una base minimale destra di  $W_1^\ell(\lambda)$ , con i vettori partizionati in blocchi conformi ai blocchi colonna di  $W_1^\ell(P)$ ; sia inoltre  $x_j(\lambda)$  il  $(k - \ell + 1)$ -esimo blocco  $n \times 1$  di  $z_j(\lambda)$ , con  $j = 1, \dots, p$ . Allora  $(x_1(\lambda), \dots, x_p(\lambda))$  è una base minimale destra di  $P(\lambda)$ .*  
 (b) *Supponiamo che  $(w_1(\lambda)^T, \dots, w_q(\lambda)^T)$  sia una base minimale sinistra di  $W_1^\ell(\lambda)$ , con i vettori partizionati in blocchi conformi ai blocchi riga di  $W_1^\ell(P)$ ; sia inoltre  $y_j(\lambda)^T$  il primo blocco  $1 \times m$  di  $w_j(\lambda)^T$ , con  $j = 1, \dots, q$ . Allora  $(y_1(\lambda)^T, \dots, y_q(\lambda)^T)$  è una base minimale sinistra di  $P(\lambda)$ .*
3. (a) *Se  $0 \leq \varepsilon_1 \leq \cdots \leq \varepsilon_p$  sono gli indici minimali destri di  $P(\lambda)$ , allora*

$$\varepsilon_1 + k - \ell \leq \cdots \leq \varepsilon_p + k - \ell$$

*sono gli indici minimali destri di  $W_1^\ell(\lambda)$ .*



(b) Se  $0 \leq \eta_1 \leq \dots \leq \eta_q$  sono gli indici minimali sinistri di  $P(\lambda)$ , allora

$$\eta_1 \leq \dots \leq \eta_q$$

sono ancora gli indici minimali sinistri di  $W_1^\ell(\lambda)$ .

**Teorema 3.3.2.** Sia  $P(\lambda) = \sum_{i=0}^k \lambda^i A_i \in \mathbb{F}[\lambda]^{m \times n}$  di pseudo grado  $k \geq 2$ . Siano  $1 \leq \ell < k$  e  $W_2^\ell(\lambda)$  definita in (3.15). Allora valgono i seguenti risultati:

$W_2^\ell(\lambda) \cup P(\lambda)$ .

1. Supponiamo che  $(z_1(\lambda), \dots, z_p(\lambda))$  sia una base minimale destra di  $W_2^\ell(\lambda)$ , con i vettori partizionati in blocchi conformi ai blocchi colonna di  $W_2^\ell(\lambda)$ ; sia inoltre  $x_j(\lambda)$  il primo blocco  $n \times 1$  di  $z_j(\lambda)$ , con  $j = 1, \dots, p$ . Allora  $(x_1(\lambda), \dots, x_p(\lambda))$  è una base minimale destra di  $P(\lambda)$ .
2. Supponiamo che  $(w_1(\lambda)^T, \dots, w_q(\lambda)^T)$  sia una base minimale sinistra di  $W_2^\ell(\lambda)$ , con i vettori partizionati in blocchi conformi ai blocchi riga di  $C_2(P)$ ; sia inoltre  $y_j(\lambda)^T$  il blocco  $(k - \ell + 1)$ -esimo  $1 \times m$  di  $w_j(\lambda)^T$ , con  $j = 1, \dots, q$ . Allora  $(y_1(\lambda)^T, \dots, y_q(\lambda)^T)$  è una base minimale sinistra di  $P(\lambda)$ .

1. Se  $0 \leq \varepsilon_1 \leq \dots \leq \varepsilon_p$  sono gli indici minimali destri di  $P(\lambda)$ , allora

$$\varepsilon_1 \leq \dots \leq \varepsilon_p$$

sono ancora gli indici minimali destri di  $W_2^\ell(\lambda)$ .

2. Se  $0 \leq \eta_1 \leq \dots \leq \eta_q$  sono gli indici minimali sinistri di  $P(\lambda)$ , allora

$$\eta_1 + k - \ell \leq \dots \leq \eta_q + k - \ell$$

sono ancora gli indici minimali sinistri di  $W_2^\ell(\lambda)$ .

### 3.3.1 $\mathcal{L}$ -ficzione secolare

Questa seconda  $\ell$ -ficzione può essere utilizzata solamente su matrici *quadrate* ed è stata introdotta per la prima volta in [2]. L'idea è rappresentare  $P(\lambda) = \sum_{i=0}^k \lambda^i A_i$  in una base di polinomi di matrici della forma  $C_i(\lambda) = \prod_{j=1, j \neq i}^q B_j(\lambda)$ , per  $j = 1, \dots, q$ , dove  $B_i(\lambda)$  sono polinomi di matrici di grado  $d_i$  con le seguenti proprietà:

- *commutatività*:  $B_i B_j = B_j B_i$  per ogni coppia  $(i, j)$ .
- *coprimalità*: L'equazione  $B_i \alpha + B_j \beta = I_m$  ammette sempre almeno una soluzione  $(\alpha, \beta) = (\alpha, \beta)_{i,j}$ , dove  $\alpha$  e  $\beta$  sono polinomi di matrici.

La  $\ell$ -ficzione dipende dalla scelta della base  $C_i(\lambda)$  e risulta  $\ell \geq \lceil \frac{k}{q} \rceil$ . Vediamo un esempio.

Siano  $b_i(\lambda)$  polinomi reali monici coprimi di grado  $d_i$ , tali che  $k = d_1 + \dots + d_q$  e sia  $s \in \mathbb{R}$  tale che  $\theta b_q(\xi) + s \neq 0$  per ogni autovalore  $\theta$  e per ogni radice  $\xi$  dei polinomi  $b_i(\lambda)$ , con  $i = 1, \dots, q-1$ . Allora la  $\ell$ -ficazione  $L(\lambda)$  è data da:

$$\begin{aligned} L(\lambda) &= D(\lambda) + (e \otimes I)[W_1, \dots, W_q], \\ D(\lambda) &= \text{diag}(B_1(\lambda), \dots, B_q(\lambda)), \quad e = [1, \dots, 1] \in \mathbb{R}^q, \\ B_i(\lambda) &= b_i(\lambda)I_m, \quad \text{for } i = 1, \dots, q-1, \\ B_q(\lambda) &= b_q(\lambda)A_k + sI_m, \\ W_i(\lambda) &\in C[\lambda]^{m \times m}, \quad \deg W_i(\lambda) < \deg d_i. \end{aligned}$$

Il polinomio  $L(\lambda)$  ha grado  $\ell = \max\{d_1, \dots, d_q\} \geq \lceil \frac{k}{q} \rceil$  e taglia  $mq \times mq$ ; le matrici  $W_i(\lambda)$  sono tali da verificare l'identità

$$P(\lambda) = B(\lambda) + \sum_{i=1}^q W_i C_i(\lambda), \quad B(\lambda) = \prod_{i=1}^q B_i(\lambda)$$

La dimostrazione dell'esistenza e unicità delle  $W_i(\lambda)$  può essere trovata sempre in [2].

Supponiamo ora che i  $b_i(\lambda) = \lambda - \beta_i$  siano polinomi lineari. Allora  $L(\lambda)$  diventa la matrice pencil:

$$\lambda \begin{bmatrix} I_m & & & & \\ & \ddots & & & \\ & & I_m & & \\ & & & & A_k \end{bmatrix} - \begin{bmatrix} \beta_1 I_m & & & & \\ & \ddots & & & \\ & & \beta_{k-1} I_m & & \\ & & & & \beta_k A_k - s I_m \end{bmatrix} + \begin{bmatrix} I_m \\ I_m \\ \vdots \\ I_m \end{bmatrix} [W_1 \cdots W_k] \quad (3.16)$$

con le matrici  $W_i(\lambda)$ :

$$W_i(\lambda) = \begin{cases} \frac{P(\beta_i)}{\prod_{j=1, j \neq i}^{k-1} (\beta_i - \beta_j)} ((\beta_1 - \beta_k)A_k + sI_m)^{-1} & \text{for } i < k, \\ \frac{P(\beta_k)}{\prod_{j=1}^{k-1} (\beta_k - \beta_j)} - sI_m - s \sum_{j=1}^{k-1} \frac{W_j}{\beta_k - \beta_j} & \text{altrimenti.} \end{cases}$$

Nel caso  $m = 1$ , ovvero quando  $P(\lambda) = p(\lambda)$  è un polinomio scalare, risulta  $\det L(\lambda) = \prod_{i=1}^k (\lambda - \beta_i) (\sum_{i=1}^k \frac{W_i}{\lambda - \beta_i} + 1)$ , dunque il problema agli autovalori può essere riformulato come la ricerca delle radici dell'*equazione secolare*. Per tale ragione la matrice  $L(\lambda)$  è stata chiamata  *$\ell$ -ficazione secolare* [2]. Gli autori dimostrano che la  $\ell$ -ficazione secolare è una forte  $\ell$ -ficazione e mostrano come sia possibile ricavare gli autovettori originali.

Si noti come venga sfruttata la stessa rappresentazione della linearizzazione di Lagrange presentata in [1]. Tuttavia quest'ultima introduce degli autovalori

---

all'infinito, in quanto è sì una forte linearizzazione di  $P(\lambda)$ , ma con pseudo grado  $k + 2$ . Sperimentalmente viene mostrato che una scelta oculata dei punti  $\beta_i$  migliora notevolmente (parliamo di alcuni ordini di grandezza) i numeri di condizionamento della pencil  $L(\lambda)$  rispetto alle due companion di Frobenius  $C_1(\lambda)$  e  $C_2(\lambda)$ . Come selezionare i punti  $\beta_i$  e analisi molto più dettagliate possono essere trovate in [2].



# Capitolo 4

## Index sum theorem

Abbiamo già utilizzato l'Index sum theorem nel teorema 2.4.2 e, implicitamente, nella sezione 3.2. Non daremo solo delle limitazioni sul grado delle possibili  $\ell$ -ficcioni di un polinomio di matrici  $P(\lambda)$ , ma una relazione fondamentale tra i gradi dei suoi divisori elementari e i suoi indici minimali.

### 4.1 Dimostrazione dell'index sum theorem

Ricordiamo alcune notazioni:

- $\delta_{\text{fin}}(P) :=$  somma dei gradi dei polinomi invarianti di  $P(\lambda)$ .
- $\delta_{\infty}(P) :=$  somma delle molteplicità parziali in 0 di  $\text{rev } P(\lambda)$ .
- $\mu(P) :=$  somma degli indici minimali di  $P(\lambda)$ .

Il prossimo risultato è stato dimostrato in [3] per polinomi di matrici regolari in  $\mathbb{C}$ . La dimostrazione, identica ma valida per ogni campo  $\mathbb{F}$ , è stata riportata anche in [5].

**Lemma 4.1.1** (Index sum theorem per polinomi regolari). *Sia  $P(\lambda) = \sum_{i=0}^k \lambda^i A_i \in \mathbb{F}[\lambda]^{m \times m}$  un polinomio di matrici regolare su un campo arbitrario  $\mathbb{F}$ . Allora:*

$$\delta_{\text{fin}}(P) + \delta_{\infty}(P) = km \tag{4.1}$$

*Dimostrazione.* Considerando la forma di Smith è possibile osservare che  $\delta_{\text{fin}}(P)$  è uguale a  $d := \deg(\det P(\lambda)) \leq km$ . Per definizione  $\delta_{\infty}(P)$  corrisponde alla molteplicità di  $\lambda$  come fattore di  $\det(\text{rev}_k P(\lambda))$ . Stiamo infatti sfruttando l'ipotesi di regolarità: se  $P(\lambda)$  è regolare, anche il suo rovesciato lo è. Poniamo

$$\det P(\lambda) = a_d \lambda^d + a_{d-1} \lambda^{d-1} + \cdots + a_1 \lambda + a_0,$$

con  $a_d \neq 0$  e calcoliamo

$$\begin{aligned} \det(\text{rev}_k P(\lambda)) &= \det \left[ \lambda^k P \left( \frac{1}{\lambda} \right) \right] \\ &= \lambda^{km} \det P \left( \frac{1}{\lambda} \right) \\ &= \lambda^{km} (a_d \lambda^{-d} + a_{d-1} \lambda^{-d+1} + \cdots + a_1 \lambda + a_0) \\ &= \lambda^{km-d} (a_d + a_{d-1} \lambda + \cdots + a_1 \lambda^{d-1} + a_0 \lambda^d) \end{aligned}$$

Dunque  $\delta_\infty(P) = km - d$ , da cui  $\delta_{\text{fin}}(P) + \delta_\infty(P) = km$ , come desiderato.  $\square$

Per dimostrare il risultato generale useremo ancora una volta la forma di Kronecker. Ci serve un risultato un po' più debole, da utilizzare in un campo qualsiasi, senza l'ipotesi sulla chiusura algebrica.

**Lemma 4.1.2** (Forma di Kronecker parziale). *Sia  $L(\lambda)$  una matrice pencil  $m \times n$  su un campo arbitrario  $\mathbb{F}$ . Allora  $L(\lambda)$  è strettamente equivalente ad una matrice a blocchi nella forma*

$$L(\lambda) \cong \begin{bmatrix} R(\lambda) & \mathbf{0} \\ \mathbf{0} & S(\lambda) \end{bmatrix},$$

dove  $R(\lambda)$  è una matrice pencil regolare  $r \times r$ , mentre  $S(\lambda)$  ha taglia  $(m - r) \times (n - r)$  ed è somma diretta di blocchi della forma  $0_{k \times l}$ ,  $S_d(\lambda)$  e  $S_d(\lambda)^T$ , dove  $S_d(\lambda)$  è la matrice singolare  $d \times (d + 1)$

$$S_d(\lambda) := \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \end{bmatrix}$$

definita nel teorema 2.3.4.

*Dimostrazione.* La dimostrazione fornita in [9, Ch XII, 35-37] sulla forma di Kronecker inizia provando che ogni matrice pencil può essere ridotta tramite equivalenza stretta in questa forma. L'ipotesi sulla completezza del campo viene utilizzata solamente nella seconda parte, quando  $R(\lambda)$  viene ulteriormente scomposto nei blocchi regolari del teorema 2.3.4.  $\square$

**Lemma 4.1.3** (Index sum theorem per matrici pencil). *Sia  $L(\lambda)$  una matrice di pseudo grado uguale a 1 su un qualsiasi campo  $\mathbb{F}$ . Allora vale l'uguaglianza:*

$$\delta_{\text{fin}}(L) + \delta_\infty(L) + \mu(L) = \text{rank } L(\lambda).$$

*Dimostrazione.* Poiché l'equivalenza stretta conserva sia la struttura spettrale che quella singolare, possiamo supporre che  $L(\lambda)$  sia nella forma del lemma 4.1.2. Poniamo  $\mu := \mu(L) = \mu(S)$ . Se osserviamo la struttura dei blocchi di

$S(\lambda)$  ci accorgiamo che  $\text{rank } S(\lambda) = \mu$ . Infatti i blocchi nulli non influiscono queste due quantità, mentre i blocchi  $S_{d_i}(\lambda)$  hanno rango  $d_i$ , per ogni  $d_i$  indice minimale di  $L(\lambda)$ . Dunque:

$$\begin{aligned}\text{rank } L(\lambda) &= \text{rank } R(\lambda) + \text{rank } S(\lambda) = \delta_{\text{fin}} + \delta_{\infty} + \text{rank } S(\lambda) \\ &= \delta_{\text{fin}}(L) + \delta_{\infty}(L) + \mu(L),\end{aligned}$$

dove nella seconda uguaglianza abbiamo utilizzato il lemma 4.1.1. .  $\square$

Possiamo ora generalizzare il risultato per qualsiasi polinomio di matrici.

**Teorema 4.1.4** (Index sum theorem). *Sia  $P(\lambda) \in \mathbb{F}[\lambda]^{m \times n}$  un polinomio di matrici di pseudo grado  $\text{gr } P(\lambda) = k$ . Allora vale l'uguaglianza:*

$$\delta_{\text{fin}}(P) + \delta_{\infty}(P) + \mu(P) = \text{rank } P(\lambda) \cdot \text{gr } P(\lambda).$$

*Dimostrazione.* La tesi seguirà calcolando il rango di  $C_1(\lambda)$  in due modi diversi. Sia  $p = \dim \mathcal{N}_r(P)$  il numero degli indici minimali destri di  $P(\lambda)$ . Utilizzando il lemma 4.1.3 e la caratterizzazione degli indici minimali del teorema 3.1.2 possiamo scrivere:

$$\text{rank } C_1(\lambda) = \delta_{\text{fin}}(C_1) + \delta_{\infty}(C_1) + \mu(C_1) = \delta_{\text{fin}}(P) + \delta_{\infty}(P) + \mu(P) + p(k-1)$$

Il secondo metodo sfrutta semplicemente il fatto di essere una forte linearizzazione di  $P(\lambda)$ , ovvero  $C_1(\lambda) \sim \text{diag}(P, I_{n(k-1)})$ . Ne deriva

$$\text{rank } C_1(\lambda) = \text{rank } P(\lambda) + n(k-1) = r + n(k-1),$$

con  $r = \text{rank } P(\lambda)$ . Eguagliando i due membri otteniamo

$$\delta_{\text{fin}}(P) + \delta_{\infty}(P) + \mu(P) + p(k-1) = r + n(k-1),$$

da cui

$$\begin{aligned}\delta_{\text{fin}}(P) + \delta_{\infty}(P) + \mu(P) &= r + (n-p)(k-1) \\ &= r + r(k-1) = rk,\end{aligned}$$

come volevasi dimostrare.  $\square$

Oltre che in [5], l'index sum theorem era stato dimostrato in numerose altre occasioni, benché non in questa generalità. Altri sottocasi erano stati provati indipendentemente da Johansson, Kågström e Van Dooren nel 2013 [12].

## 4.2 Caratterizzazione dei gradi delle $\ell$ -ficzioni

Grazie all'index sum theorem possiamo finalmente dimostrare quelle restrizioni sui gradi delle  $\ell$ -ficzioni a cui avevamo accennato nella sezione 3.2. Iniziamo con una definizione che renderà più chiare le prossime dimostrazioni.

**Definizione 4.2.1** (Total index shift). Siano  $P$  e  $Q$  due matrici equivalenti in senso spettrale, ovvero  $P \asymp Q$ , appartenenti a  $\mathbb{F}[\lambda]^{m \times n}$ . Allora definiamo *total index shift* da  $P$  a  $Q$  la differenza:

$$\mathcal{S}(P, Q) = \mu(Q) - \mu(P).$$

Possiamo caratterizzare  $\mathcal{S}(P, Q)$  tramite l'index sum theorem, come si vedrà nel prossimo lemma.

**Lemma 4.2.2.** *Siano  $P$  e  $Q$  due matrici equivalenti in senso spettrale, rispettivamente di pseudo grado  $k$  e  $\ell$ . Allora il total index shift da  $P$  a  $Q$  vale:*

$$\mathcal{S}(P, Q) = \ell \operatorname{rank} Q - k \operatorname{rank} P.$$

*Dimostrazione.* Grazie all'index sum theorem sappiamo che

$$\begin{aligned} \mu(P) &= k \operatorname{rank} P - \delta_{\text{fin}}(P) - \delta_{\infty}(P) \\ \mu(Q) &= \ell \operatorname{rank} Q - \delta_{\text{fin}}(Q) - \delta_{\infty}(Q) \end{aligned}$$

Tuttavia  $P(\lambda) \asymp Q(\lambda)$ , perciò  $\delta_{\text{fin}}(P) = \delta_{\text{fin}}(Q)$  e  $\delta_{\infty}(P) = \delta_{\infty}(Q)$  (teorema 2.3.11), da cui la tesi segue immediatamente.  $\square$

Ora possiamo fornire una risposta, purtroppo non esaustiva, sulle taglie possibili che può assumere una  $\ell$ -ficzione. Il seguente teorema è infatti valido solo per polinomi di matrici *regolari*. Si basa sul *Regular realization theorem*, di cui riportiamo l'enunciato [16, teorema 9.4].

**Teorema 4.2.3** (Regular realization theorem). *Sia  $\mathbb{F}$  un campo algebricamente chiuso e sia  $\mathcal{L}$  una lista di divisori elementari, che includa un numero qualsiasi di ripetizioni e comprenda quelli all'infinito. Allora sono equivalenti i seguenti enunciati:*

1.  $\delta_{\text{fin}}(\mathcal{L}) + \delta_{\infty}(\mathcal{L}) = km$ .
2. *Esiste un polinomio matriciale  $P(\lambda)$  di taglia  $m \times m$  e pseudo grado  $k$  tale che la sua struttura spettrale sia definita dalla lista  $\mathcal{L}$ . Inoltre  $P(\lambda)$  può essere preso in forma triangolare superiore.*

**Teorema 4.2.4** (Restrizioni sulle taglie delle  $\ell$ -ficzioni). *Sia  $P(\lambda)$  un polinomio di matrici regolare di grado  $k$  a coefficienti in  $\mathbb{F}^{m \times m}$ . Dato un numero naturale  $1 \leq \ell \leq k$ :*



1. Se  $Q$  è una forte  $\ell$ -ficzione di  $P$ , allora  $\ell$  è un divisore di  $km$  e  $Q$  ha taglia  $s = km/\ell$ .
2. Supponiamo inoltre che  $\mathbb{F}$  sia algebricamente chiuso: se  $\ell \leq k$  divide  $km$ , allora esiste una  $\ell$ -ficzione forte  $Q$  di  $P$ .

*Dimostrazione.* 1 Assumiamo che  $Q(\lambda)$  sia una forte  $\ell$ -ficzione di  $P(\lambda)$  di taglia  $s \times s$ . Per il teorema 2.3.11 anche  $Q$  è regolare, dunque anche il suo rango coincide con la sua dimensione. Sempre per regolarità osserviamo che  $\mathcal{S}(P, Q) = 0$ . Dal corollario 4.2.2 segue  $\ell s = km$ , ovvero la nostra tesi.

2 Supponiamo allora che  $\mathcal{L}$  sia la lista dei divisori corrispondenti al polinomio regolare  $P(\lambda)$  di dimensione  $m \times m$  e grado  $k$ . Sia  $T := \delta_{\text{fin}}(\mathcal{L}) + \delta_{\infty}(\mathcal{L})$ ; per il lemma 4.1.1 risulta  $T = km$ . Supponiamo inoltre che  $\ell \leq k$  divida  $km$ , ovvero  $T = km = s\ell$ . Per il teorema 4.2.3 esiste  $Q(\lambda)$  di dimensione  $s \times s$  e grado  $\ell$  con gli stessi divisori elementari di  $P(\lambda)$ . Allora  $P(\lambda) \asymp Q(\lambda)$  per la caratterizzazione dell'equivalenza spettrale 2.3.11. □

**Osservazione 4.2.5.** È stato congetturato che non sia necessaria l'ipotesi sulla chiusura algebrica del campo  $\mathbb{F}$  per la seconda parte della tesi. In effetti in [6] è stato dimostrato lo stesso risultato per *campi infiniti*. Se il risultato sia vero o meno su campi finiti rimane un problema aperto.

Ulteriori conseguenze e approfondimenti riguardanti l'index sum theorem possono essere trovate in [5]. Vengono studiate le proprietà del total index shift sulle  $\ell$ -ficzioni di matrici regolari per essere poi applicate alla ricerca di  $\ell$ -companion strutturate per polinomi simmetrici, antisimmetrici e palindromi. Infine viene dimostrato che *non esiste alcuna 1-companion per polinomi simmetrici o Hermitiani di pseudo grado pari*.



# Bibliografía

- [1] A. Amiraslani, R. M. Corless e P. Lancaster. «Linearization of matrix polynomials expressed in polynomial bases». In: *IMA J. Numer. Anal.* 29.1 (2009), pp. 141–157. ISSN: 0272-4979. DOI: 10.1093/imanum/drm051. URL: <http://dx.doi.org/10.1093/imanum/drm051>.
- [2] Dario A. Bini e Leonardo Robol. *On a class of matrix pencils equivalent to a given matrix polynomial*. 2014. arXiv: arXiv:1406.1025.
- [3] Fernando De Terán e Froilán M. Dopico. «Sharp lower bounds for the dimension of linearizations of matrix polynomials». In: *Electron. J. Linear Algebra* 17 (2008), pp. 518–531. ISSN: 1081-3810.
- [4] Fernando De Terán, Froilán M. Dopico e D. Steven Mackey. «Fiedler companion linearizations for rectangular matrix polynomials». In: *Linear Algebra Appl.* 437.3 (2012), pp. 957–991. ISSN: 0024-3795. DOI: 10.1016/j.laa.2012.03.028. URL: <http://dx.doi.org/10.1016/j.laa.2012.03.028>.
- [5] Fernando De Terán, Froilán M. Dopico e D. Steven Mackey. «Spectral equivalence of matrix polynomials and the index sum theorem». In: *Linear Algebra Appl.* 459 (2014), pp. 264–333. ISSN: 0024-3795. DOI: 10.1016/j.laa.2014.07.007. URL: <http://dx.doi.org/10.1016/j.laa.2014.07.007>.
- [6] Fernando De Terán, Froilán M. Dopico e Paul Van Dooren. «Matrix Polynomials with Completely Prescribed Eigenstructure». In: *SIAM J. Matrix Anal. Appl.* 36.1 (2015), pp. 302–328. ISSN: 0895-4798. DOI: 10.1137/140964138. URL: <http://dx.doi.org/10.1137/140964138>.
- [7] G. David Forney Jr. «Minimal bases of rational vector spaces, with applications to multivariable linear systems». In: *SIAM J. Control* 13 (1975), pp. 493–520. ISSN: 0363-0129.
- [8] J. G. F Francis. «The QR transformation - Parts I and II». In: *Comput. J.* 4 (1961-62), pp. 265–271, 332–345. ISSN: 0010-4620.
- [9] F. R. Gantmacher. *The theory of matrices. Vols. 1, 2*. Translated by K. A. Hirsch. Chelsea Publishing Co., New York, 1959, Vol. 1, x+374 pp. Vol. 2, ix+276.

- [10] I. Gohberg, M. A. Kaashoek e P. Lancaster. «General theory of regular matrix polynomials and band Toeplitz operators». In: *Integral Equations Operator Theory* 11.6 (1988), pp. 776–882. ISSN: 0378-620X. DOI: 10.1007/BF01208810. URL: <http://dx.doi.org/10.1007/BF01208810>.
- [11] Nicholas J. Higham e Hyun-Min Kim. «Numerical analysis of a quadratic matrix equation». In: *IMA J. Numer. Anal.* 20.4 (2000), pp. 499–519. ISSN: 0272-4979. DOI: 10.1093/imanum/20.4.499. URL: <http://dx.doi.org/10.1093/imanum/20.4.499>.
- [12] Stefan Johansson, Bo Kågström e Paul Van Dooren. «Stratification of full rank polynomial matrices». In: *Linear Algebra Appl.* 439.4 (2013), pp. 1062–1090. ISSN: 0024-3795. DOI: 10.1016/j.laa.2012.12.013. URL: <http://dx.doi.org/10.1016/j.laa.2012.12.013>.
- [13] L. Kronecker. «Über die congruente Transformationen der bilinearen Formen». In: *Collected Works I* (1968), pp. 423–483.
- [14] P. Lancaster e P. Psarrakos. *A note on weak and strong linearizations of regular matrix polynomials*. Rapp. tecn. 470. Manchester Centre for Computational Mathematics, 2005.
- [15] Peter Lancaster. «Linearization of regular matrix polynomials». In: *Electron. J. Linear Algebra* 17 (2008), pp. 21–27. ISSN: 1081-3810.
- [16] D. Steven Mackey et al. «Möbius transformations of matrix polynomials». In: *Linear Algebra Appl.* 470 (2015), pp. 120–184. ISSN: 0024-3795. DOI: 10.1016/j.laa.2014.05.013. URL: <http://dx.doi.org/10.1016/j.laa.2014.05.013>.
- [17] C. B. Moler e G. W. Stewart. «An algorithm for generalized matrix eigenvalue problems». In: *SIAM J. Numer. Anal.* 10 (1973). Collection of articles dedicated to the memory of George E. Forsythe, pp. 241–256. ISSN: 0036-1429.
- [18] B. N. Parlett. «The QR algorithm». In: *Computing Sci. Eng.* 2 (2000), pp. 38–42.
- [19] Françoise Tisseur e Karl Meerbergen. «The quadratic eigenvalue problem». In: *SIAM Rev.* 43.2 (2001), pp. 235–286. ISSN: 0036-1445. DOI: 10.1137/S0036144500381988. URL: <http://dx.doi.org/10.1137/S0036144500381988>.
- [20] Françoise Tisseur e I. Zaballa. «Finite and infinite elementary divisors of matrix polynomials: A global approach». In: *PIPPI* (2012).