



UNIVERSIDADE DA BEIRA INTERIOR
Ciências

Algoritmos para Solventes de Polinómios Matriciais

Fernando António Carvalho Marcos

Tese para obtenção do Grau de Doutor em
Matemática Aplicada
(3º ciclo de estudos)

Orientador: Prof. Doutor Edgar Pereira
Co-orientador: Prof. Doutor Paulo Rebelo

Covilhã, Outubro de 2012

Dedicatória

À Carme, à Rita e ao Jordi... sem eles, não teria valido a pena.

Agradecimentos

Ao Prof. Edgar Pereira pela orientação, acompanhamento e apoio no decorrer e sobretudo na parte final do trabalho;

ao Prof. Paulo Rebelo pela supervisão e apoio;

à Universidade da Beira Interior pela possibilidade de desenvolver este trabalho;

ao Instituto Politécnico da Guarda pelo apoio e condições oferecidas;

à Carme, à Rita e ao Jordi pela paciência e carinho.

Resumo

Os polinômios matriciais têm um papel importante na teoria das equações diferenciais matriciais resultantes de formulações matemáticas cada vez mais exigentes. Precisamente, uma abordagem para o problema de cálculo numérico de soluções de equações diferenciais matriciais é feita através da computação de solventes do polinômio matricial associado $P(X)$ (Lancaster [25], pag. 525). O primeiro trabalho que se conhece neste âmbito está presente em Dennis [9], que deu origem ao desenvolvimento da teoria algébrica dos polinômios matriciais Dennis [10] e [11] e onde são apresentados Algoritmos de cálculo de solventes. Chama-se à atenção do leitor para Dennis [11], pág. 524, onde são definidos dois métodos iterativos que permitem o cálculo de solventes.

O primeiro, citado como método de Traub, permite a computação do solvente dominante, isto é, do solvente cujos valores próprios são maiores, em módulo, do que os valores próprios de qualquer outro solvente. O segundo Algoritmo é uma versão matricial do método de Bernoulli, que consiste basicamente no método da Potência aplicado à matriz companheira de $P(X)$. Após Dennis [11] vários trabalhos consideraram este método (Lancaster [22], Tsai [38], Higham [19], Pereira [34]). O método de Newton clássico também foi adaptado ao contexto dos polinômios matriciais, primeiro à equação quadrática (Davis [5], Kratz [21], Higham [18], Long [26]) e posteriormente para polinômios de grau m qualquer. Recentemente foi também objeto de estudo em Higham [19] e Pereira [33].

Todos os métodos referidos são desenvolvidos com base na álgebra matricial, isto é, com base na equação $P(X) = 0_{n \times n}$ em $\mathbf{C}^{n \times n}$. No presente trabalho é desenvolvido um método do ponto fixo considerando as entradas do polinômio matricial $P(X)$, reduzindo o problema ao nível escalar tentando com isso evitar os problemas de cálculo derivados da álgebra matricial sobretudo quando estes envolvem a inversa de uma matriz.

É também apresentada uma versão vetorial do método de Newton para polinômios matriciais. Na sequência da ideia desenvolvida em (Marcos [27], pág. 357), onde a equação matricial $P(X) = 0_{n \times n}$ é trabalhada ao nível escalar, é também considerado o método de Newton aplicado à equação formada por $n \times n$ equações polinomiais. O objetivo é evitar a derivada de Fréchet e a resolução da respetiva equação matricial de Sylvester em cada iteração, tal como acontece no método definido em (Higham [18], pág. 4).

De acordo com Dennis [9], pág. 80, se X é um solvente do polinômio matricial $P(X)$ então X um bloco valor próprio da matriz companheira, $CV = VX$ no caso mónico, ou bloco valor próprio do feixe companheiro, $C_1V = C_2VX$ no caso não mónico. Relativamente a métodos iterativos com aplicação no cálculo de blocos valores próprios, pelo que se conseguiu apurar, existe apenas o método da Potência definido em (Dennis [9], pág. 83) e aplicado apenas a polinômios mónicos.

Assim, é apresentado o Método Newton Vetorial para Blocos Valores Próprios definido para blocos valores próprios da matriz companheira e posteriormente adaptado ao cálculo de blocos valores próprios do feixe companheiro ou de um feixe genérico $(A, B) = \lambda B - A$ qualquer.

Por último generalizou-se esta formulação para o cálculo de feixes próprios,

$$(\lambda B - A)V = W(\lambda Y - X).$$

resultando no Método de Newton Vetorial para Feixes Próprios, definido para um feixe genérico $(A, B) = \lambda B - A$.

Palavras-chave

Polinómio matricial, λ -Matriz, feixe $\lambda B - A$, produto de Kronecker, matriz companheira, feixe companheiro, forma canónica de Jordan, bloco de Jordan, forma canónica de Weierstrass, solventes, blocos valores próprios e blocos vetores próprios, feixe próprio, derivada de Fréchet, problema de valores próprios quadrático, matriz bloco de Vandermonde.

Abstract

Matrix polynomials play an important role in the theory of matrix differential equations. An important approach in searching numerical solutions to matrix differential equation is through the computation of solvents of the associated matrix polynomial $P(X)$ (Lancaster [25], page 525). The first work we know in numerical analysis dealing with matrix polynomials is in Dennis [9], which gave origin to Dennis [10] and [11] where an algebraic theory was developed and some algorithms were presented. We refer the reader to Dennis [11], page 524, where two iterative methods to find solvents are defined.

The first is a generalization of a scalar algorithm and its purpose is the computation of a dominant solvent, that is a solvent with the eigenvalues greater, in modulus, than the eigenvalues of any other solvent. We will refer this as the Traub method.

The second algorithm is a matrix version of the Bernoulli's algorithm, which is essentially a block matrix power method applied to a block companion matrix of $P(X)$. Since Dennis [10], several works have considered this method (Lancaster [22], Tsai [38], Higham [19], Pereira [34]). The classical Newton's method also had been generalized to matrix polynomials, first to the quadratic equation (Davis [5], Kratz [21], Higham [18], Long [26]) and then to a general degree m . Also this method have been studied in the last years by Higham [19] and Pereira [33].

All methods referred above are based in matrix arithmetics, that is, solving the matrix equation $P(X) = 0_{n \times n}$ in $\mathbf{C}^{n \times n}$. Here we will develop a fixed point method considering the matrix elementwise, so the computations will be carried at the scalar level, our attempt is to avoid the complications of matrix manipulations specially when dealing with the inverse of a matrix.

We present here a vectorial version for the Newton's method for matrix polynomials. This follows the approach of (Marcos [27]), which is to treat again a polynomial matrix equation at a scalar level, for these we consider the polynomial equation by $n \times n$ scalar multivariate complex polynomials. The goal here is to avoid the use of the Fréchet derivative which in the respective algorithm it is needed to solve a generalized Sylvester equation at each step (see Higham [18], page 4).

In Dennis [9], page 80, it is referred that if X is a solvent of $P(X)$ then it is a block eigenvalue of the companion matrix, $\mathcal{C}V = VX$ in the monic case, or a block eigenvalue of the companion pencil, $\mathcal{C}_1V = \mathcal{C}_2VX$ in the nonmonic case.

Another approach in searching numerical solutions to matrix differential equation is through the eigenvalues and generalized eigenvectors of the companion matrix or pencil. Related to block eigenvalue calculus, as much as we could, we just find out the block matrix power method (Dennis [9], page 83) and with application only to monic polynomials.

It is therefore presented the Vectorial Newton method to block eigenvalue defined for the monic case, or a block eigenvalue of the companion pencil, in the nonmonic case. Finally we generalize this formulation for calculating eigenpencils,

$$(\lambda B - A)V = W(\lambda Y - X).$$

resulting in the Vectorial Newton Method for eigenpencils.

Keywords

Matrix polynomials, λ -Matrix, $\lambda B - A$ pencil, Kronecker product, companion matrix, companion pencil, Jordan canonical form, Jordan block, Weierstrass canonical form, solvents, block eigenvalues and block eigenvectors, eigenpencil, Fréchet derivative, quadratic eigenvalue problem, block Vandermonde matrix.

Conteúdo

1	Introdução	1
1.1	Notação	1
1.2	Enquadramento	1
1.3	Objetivos	6
1.4	Estrutura	7
2	Resultados Preliminares	9
2.1	Definições	9
2.2	Formas Canônicas	15
2.3	Valores e vetores próprios da λ -matriz $P(\lambda)$	17
2.3.1	Valores próprios finitos	18
2.3.9	Valor próprio infinito	21
2.4	Solventes do polinômio matricial $P(X)$	26
3	Cálculo de Solventes do Polinômio matricial $P(X)$	31
3.1	Métodos Iterativos	31
3.2	Método do Ponto Fixo	34
3.2.3	Algoritmo	35
3.2.4	Convergência	36
3.2.8	Exemplos	37
3.3	Método de Newton Vetorial	39
3.3.3	Algoritmo Principal	42
3.3.4	Convergência	43
3.3.11	Algoritmos com Procura Unidimensional	49
3.3.12	Exemplos	51
4	Cálculo de Blocos Valores Próprios	55
4.1	Método Newton Vetorial para Blocos Valores Próprios	55
4.1.2	Algoritmo do Método	56
4.1.3	Convergência	56
4.1.7	Exemplos	60
4.1.10	Aplicação à matriz \mathcal{C}	64
4.1.12	Aplicação ao Feixe $\lambda B - A$	66
4.2	Cálculo de Feixes Próprios	73
4.2.2	Método Newton Vetorial para Feixes Próprios	75
4.2.9	Feixe Próprio do Feixe Companheiro $\lambda \mathcal{C}_2 - \mathcal{C}_1$	85
5	Conclusões	91
5.1	Trabalho Futuro	92
	Bibliografia	93
A	Anexos	97
A.1	Algoritmos dos métodos referidos	97
	Índice Remissivo	103

Lista de Tabelas

3.1	Evolução do raio espectral das aproximações no método do Ponto Fixo	38
3.2	Comparação do Algoritmo 3.3.1 com os Algoritmos 2.1, 3.1, 3.2 e 3.3.	49
3.3	Comparação do Algoritmo 3.3.1 com os Algoritmos 2.1, 3.1, 3.2 e 3.3.	49
3.4	Comparação do Algoritmo 3.3.1 com os Algoritmos 2.1, 3.1, 3.2 e 3.3.	49
3.5	Comparação do Algoritmo NV com os Algoritmos NV 3.2 e NV 3.3.	50
3.6	Comparação do Algoritmo NV com os Algoritmos NV 3.2 e NV 3.3.	51
3.7	Comparação do Algoritmo NV com os Algoritmos NV 3.2 e NV 3.3.	51
3.8	Comparação dos Algoritmos NV, NV 3.2 e NV 3.3 com os Algoritmos 3.2 e 3.3. . .	51
3.9	Comparação dos Algoritmos 3.3.1, 3.3.2 e 3.3.3	53
3.10	Comparação dos Algoritmos 3.3.1, 3.3.2 e 3.3.3	53

Lista de Algoritmos

3.2.1 PF	35
3.3.1 NV	42
3.3.2 (NV 3.2)	50
3.3.3 (NV 3.3)	50
4.1.1 (NVBVP)	56
4.1.2 (NVBVP - Polinômios não mónicos)	68
4.2.1 (NVFP)	76
A.1.1MN	97
A.1.2MB	97
A.1.3MT	97
A.1.4(MP)	98

Lista de Acrônimos

MN	Método de Newton
MB	Método de Bernoulli
MT	Método de Traub
MP	Método da Potência
PF	Método do Ponto Fixo
NV	Método de Newton Vetorial
NVBVP	Método de Newton Vetorial para Blocos Valores Próprios
NVFP	Método de Newton Vetorial para Feixes Próprios
PVPQ	Problema dos Valores Próprios Quadrático
CPU	Unidade central de processamento ou processador

Capítulo 1

Introdução

Neste capítulo faz-se o enquadramento dos polinomiais matriciais e/ou λ -matrizes no âmbito da Álgebra Linear Matricial e é exposta a forma como esse enquadramento serviu de base para a motivação e delineação dos objetivos gerais deste trabalho. Faz-se ainda a introdução conceptual relativa aos polinomiais matriciais e/ou λ -matrizes e a sua relação e aplicação com as equações matriciais, quer algébricas, quer diferenciais. Neste âmbito ainda, são apontados diversos métodos de cálculo efetivo, quer de solventes, quer de blocos valores próprios ou pares próprios.

1.1 Notação

A notação encontra-se discriminada no Glossário, pág. 99. Salienta-se que, sempre que se revelar óbvio, poderão por exemplo omitir-se os índices que identificam as dimensões das matrizes consideradas,

$$\begin{aligned} I &= I_n, \\ 0 &= 0_{m \times n}, \\ J(\lambda) &= J_k(\lambda), \\ &\dots \end{aligned}$$

A notação $x = \text{vec}(X) \in \mathbb{C}^{mn}$ (ver Horn [20], pág. 243, Grassó [29] pág. 58), dependendo do contexto, pode designar tanto a matriz coluna em $x \in \mathbb{C}^{mn \times 1}$ como o vetor ou ponto $x \in \mathbb{C}^{mn}$ definido pelas mn coordenadas (x_1, \dots, x_{mn}) . Por uma questão de simplicidade, usa-se $\text{vec}(V, X)$, onde V e X são matrizes de diferentes dimensões, $V \in \mathbb{C}^{m \times n}$, $X \in \mathbb{C}^{n \times n}$, tendo como significado o vetor $(v, x) = (\text{vec}(V)^T, \text{vec}(X)^T)^T \in \mathbb{C}^{mn+n^2}$.

Da mesma forma, quando perfeitamente identificados m e n , a notação $\text{vec}^{-1}(v, x)$, deve ser entendida como o par $(V, X) = (\text{vec}^{-1}(v), \text{vec}^{-1}(x))$ com $V \in \mathbb{C}^{m \times n}$ e $X \in \mathbb{C}^{n \times n}$.

A norma matricial usada é a norma de Frobenius ($\|A\|_2$) por ser aquela que se relaciona mais diretamente com a função vec , $\|\text{vec}(A)\| = \|A\|_2$, omitindo-se o índice sempre que for irrelevante a norma utilizada.

Consideram-se apenas λ -matrizes $P(\lambda)$ regulares, isto é, λ -matrizes cujo determinante $\det(P(\lambda))$ é um polinómio não identicamente nulo. Desta forma, a escolha da letra K para a Forma Canónica de Weierstrass de um feixe (A, B) , notação $K_{(A,B)}$, prende-se com o fato de a Forma Canónica de Kronecker (ver Gantmacher [12], pág. 29, Dooren [6], pág. 111), no caso de um feixe regular, ser também designada por Forma Canónica de Weierstrass.

1.2 Enquadramento

Com o desenvolvimento em áreas científicas tão díspares como análise dinâmica de sistemas mecânicos e acústicos, circuitos elétricos ou eletrónicos, teoria de controlo, mecânica dos fluidos, (ver

Tisseur [36], Gohberg [15], Lancaster [23], Pereira [33]), e também o desenvolvimento informático ao nível da capacidade de processamento e cálculo, ficaram reunidas as condições para a aplicabilidade do cálculo numérico, de áreas da matemática que, até há pouco, por serem impraticáveis relativamente ao cálculo efetivo, estavam ligadas exclusivamente ao formalismo puro, como é o caso da generalidade das equações matriciais. O desenvolvimento de novas formulações matemáticas cada vez mais exigentes, que em diversas áreas correspondem a modelos matriciais, aliado à capacidade de processamento que os novos computadores vieram permitir, dá origem ao desenvolvimento de inúmeros trabalhos na área (inclusive a presente dissertação). Assim, trata-se de um campo em franca expansão quer ao nível de cálculo efetivo quer ao nível dos conceitos matemáticos envolvidos, que abrangem, desde as equações algébricas matriciais (que podem ser convertidas em equações unilaterais através do produto de Kronecker, ver Hernandez [14], pág. 3, Horn [20], pág. 239), como por exemplo a equação de Sylvester,

$$AX - XB = C,$$

a equação de Lyapunov,

$$A^T X + XA = -C,$$

ou a equação de Riccati,

$$XCX - XD - AX + B = 0,$$

onde A, B, C são matrizes complexas de ordem n e X é a matriz incógnita, às equações diferenciais matriciais, como sistemas de equações diferenciais lineares

$$P \left(\frac{d}{dt} \right) x = A_0 x^{(m)}(t) + A_1 x^{(m-1)}(t) + \dots + A_m x(t) = f(t),$$

com A_0, A_1, \dots, A_m matrizes complexas de dimensão $n \times n$ e $f : \mathbb{R} \rightarrow \mathbb{R}^n$, onde $x(t)$ é a função incógnita, de classe $\mathbf{C}^m(\mathbb{R})$ e $x^{(j)} := \frac{d^j x}{dt^j}$.

A solução $x(t)$ deste tipo de sistemas, do ponto de vista teórico, está bem definida mas levanta uma série de dificuldades do ponto de vista do cálculo efetivo. Se bem que na generalidade, as equações diferenciais matriciais lineares mais estudadas, afetas às diversas áreas das engenharias, são de ordem dois, a teoria matemática envolvida está definida para uma ordem qualquer m . Portanto, em termos de cálculo efetivo, o caso geral da equação diferencial matricial de ordem m , é um campo pouco explorado, ao contrário do caso de ordem 2, associado ao problema conhecido como “O problema de valores próprios quadrático” (ver Tisseur [36]).

Este problema, ($m = 2$), consiste na determinação de valores e vetores próprios de uma equação quadrática cujos coeficientes, neste caso, são matrizes reais ou complexas de dimensão $n \times n$.

No caso escalar ($n = 1$), como é bem sabido, a solução da equação diferencial de ordem 2 é definida em função das raízes da equação caraterística. A generalização deste conceito para $n > 1$, relaciona ainda a equação diferencial homogénea

$$P \left(\frac{d}{dt} \right) x = A_0 x''(t) + A_1 x'(t) + A_2 x(t) = 0,$$

com $A_0, A_1, A_2 \in \mathbb{C}^{n \times n}$, e as raízes do “polinómio caraterístico”, com as devidas adaptações ao caso matricial, que se traduz no chamado Problema dos Valores Próprios Quadrático (PVPQ), o qual consiste na obtenção de escalares $\lambda \in \mathbb{C}$ (valores próprios), tais que

$$\det(\lambda^2 A_0 + \lambda A_1 + A_2) = 0,$$

e vetores não nulos $v \in \mathbb{C}^n$ (vetores próprios) que verifiquem a equação

$$(\lambda^2 A_0 + \lambda A_1 + A_2) v = 0.$$

A seguir é apresentado um exemplo elucidativo de um PVPQ onde se percebem as semelhanças e diferenças existentes entre a resolução de um problema deste tipo e o simples cálculo de valores e vetores próprios de uma matriz A (ou, de acordo com formulação agora usada, da λ -matriz $\lambda I - A$).

Exemplo 1.2.1. Sejam as matrizes complexas,

$$A_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_1 = \begin{bmatrix} -4 & 1 \\ 0 & -3 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 4 & -2 \\ 0 & 2 \end{bmatrix}.$$

Tem-se que,

$$\lambda^2 A_0 + \lambda A_1 + A_2 = \begin{bmatrix} \lambda^2 - 4\lambda + 4 & 2 - \lambda \\ 0 & \lambda^2 - 3\lambda + 2 \end{bmatrix},$$

e

$$\det(\lambda^2 A_0 + \lambda A_1 + A_2) = (\lambda^2 - 4\lambda + 4)(\lambda^2 - 3\lambda + 2) = (\lambda - 1)(\lambda - 2)^3,$$

pelo que se obtém o valor próprio $\lambda = 1$ e o valor próprio $\lambda = 2$ com multiplicidade 3. Assim, considerando $\lambda = 1$ tem-se

$$(1^2 A_0 + 1 A_1 + A_2) \vec{v} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_1 - v_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

para qualquer vetor \vec{v} com $v_2 = v_1$;

Portanto, para $\lambda = 1$, tem-se o vetor próprio

$$\begin{bmatrix} v_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Repetindo o processo em relação ao valor próprio $\lambda = 2$, tem-se

$$(2^2 A_0 + 2 A_1 + A_2) \vec{v} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

para qualquer vetor $\vec{v} \neq 0$;

Este problema (PVPQ) consiste numa dupla generalização do cálculo da solução de uma equação diferencial linear, quer em termos da ordem da equação diferencial m quer em termos da dimensão das matrizes envolvidas n .

Para uma dimensão n fixa, o problema consiste na generalização da noção de valor e vetor próprio de uma matriz A , ou valor e vetor próprio do polinómio $P(\lambda) = \lambda I - A$, para polinómios de grau maior ou igual a 2 na variável λ ,

$$P(\lambda) = \lambda^m A_0 + \lambda^{m-1} A_1 + \dots + A_m.$$

Por outro lado, à semelhança da linearização usada numa equação diferencial linear (a transformação desta num sistema de equações diferenciais de ordem 1) uma linearização de $P(\lambda)$ corresponde à redução deste polinómio ao caso linear ($\lambda I - A$ ou $\lambda B - A$). Com efeito, tem-se que a λ -matriz

(ver Lancaster [25], pág. 491, Gohberg [15], pág. 13)

$$\begin{bmatrix} P(\lambda) & 0_{n \times (m-1)n} \\ 0_{(m-1)n \times n} & I_{(m-1)n} \end{bmatrix},$$

é equivalente a $\lambda I - C$, onde C é a matriz companheira (ver Gohberg [15], pág. 491) no caso mónico ou equivalente ao feixe companheiro (ver Gohberg [15], pág. 186, Tisseur [36], pág. 253) no caso não mónico, como será visto na secção 2.2, Observação 2.2.3.

Atendendo à transitividade da relação de equivalência entre λ -matrizes, uma linearização do polinómio $P(\lambda)$ é qualquer λ -matriz da forma $\lambda I_{mn} - C$, com $C \in \mathbb{K}^{mn \times mn}$, equivalente a $\lambda I - C$, no caso mónico ou equivalente ao feixe companheiro no caso não mónico (ver Gohberg [15], pág. 187).

Assim, se o polinómio $P(\lambda)$ for mónico, i.e. $A_0 = I_n$, a equação diferencial homogénea correspondente,

$$x^{(m)}(t) + A_1 x^{(m-1)}(t) + \dots + A_m x(t) = 0,$$

pode ser associada a uma linearização da forma $\lambda I - C$, onde C é a matriz companheira. A solução do sistema $x(t) = \begin{bmatrix} I_n & 0_n & \dots & 0_n \end{bmatrix} y(t)$, calculada através desta linearização $y'(t) = C y(t)$, é dada por

$$x(t) = \begin{bmatrix} I_n & 0_n & \dots & 0_n \end{bmatrix} e^{Ct} w,$$

onde $w \in \mathbb{C}^{mn}$, e^{Ct} representa a exponencial da matriz companheira, Ct e w o vetor com as condições iniciais $x(0)$.

Observação 1.2.2. Em Lancaster [25], pág. 308 (ou Gantmacher [12], pág. 113), define-se a função de uma matriz $A \in \mathbb{K}^{n \times n}$, $f(A)$ desde que a função $f : \mathbb{K} \rightarrow \mathbb{K}$ esteja definida no espectro de A . Isto é, sempre que existam os valores $f(\lambda_k), f'(\lambda_k), \dots, f^{(m_k-1)}(\lambda_k)$, para todo o λ_k valor próprio da matriz A de multiplicidade m_k , a função fica definida por $f(A) := p(A) = \sum_{k=1}^s \sum_{j=0}^{m_k-1} f^{(j)}(\lambda_k) \varphi_{k,j}(A)$, onde $\varphi_{k,j}(\lambda)$ constituem um conjunto fundamental de polinómios interpoladores. A exponencial de uma matriz $A \in \mathbb{K}^{n \times n}$ é a matriz perfeitamente definida pela soma da série

$$e^A = \sum_{n \geq 0} \frac{1}{n!} A^n,$$

e que pode ser também calculada através de $e^A = \sum_{k=1}^s \sum_{j=0}^{m_k-1} e^{\lambda_k} \varphi_{k,j}(A)$. Em Lancaster [25], pág. 310, é também demonstrado que dadas duas matrizes semelhantes A, B com $A = PBP^{-1}$, então $e^A = Pe^B P^{-1}$.

Contudo a matriz exponencial da matriz companheira e^{Ct} quando a dimensão n é “grande” não se apresenta como uma verdadeira alternativa em termos de cálculo. As alternativas disponíveis consistem em transformar a matriz companheira, através de transformações de semelhança, numa matriz J mais simples do ponto de vista do cálculo da respetiva matriz exponencial e^{Jt} . A escolha óbvia é a forma canónica de Jordan $C = S J_C S^{-1}$.

A caracterização da matriz companheira é, deste modo, obtida utilizando a forma canónica de Jordan da matriz companheira, $C = S J_C S^{-1}$, onde a matriz de semelhança S fica definida à custa dos vetores próprios (eventualmente, vetores próprios generalizados) da λ -Matriz $P(\lambda)$ e a matriz J_C é a forma canónica de Jordan de C .

Neste caso, dado $\{(V_1, J_1), \dots, (V_l, J_l)\}$ um sistema completo de pares próprios de $P(\lambda)$ (ver Pereira

[30], pág. 21, Gohberg [15], pág. 44), a matriz companheira pode ser representada por $C = SJ_C S^{-1}$, onde $J_C = \text{Diag}(J_1, \dots, J_l)$ e $S = [V_j J_j^i]_{\substack{i=0, \dots, m-1, \\ j=1, 2, \dots, l}}$, pelo que, nestas condições, a solução da equação diferencial $x(t) = P S e^{J_C t} S^{-1} w$ poderá caracterizar-se na forma, (ver Gohberg [15], pág. 221)

$$x(t) = V_1 e^{J_1 t} z_1 + V_2 e^{J_2 t} z_2 + \dots + V_l e^{J_l t} z_l,$$

onde $z_i \in \mathbb{C}^{k_i}$ são vetores arbitrários com $k_i = \dim(J_i)$ para $i = 1, \dots, l$.

Uma abordagem alternativa, consiste em caracterizar a matriz companheira em termos de solventes do polinômio matricial $P(X) = A_0 X^m + A_1 X^{m-1} + \dots + A_m$, onde X representa a matriz incógnita em $\mathbb{C}^{n \times n}$. Em Pereira [30], faz-se o estudo relativamente à existência e contabilização do número de solventes de polinômios matriciais mónicos, com base na relação estabelecida entre as duas alternativas.

De forma idêntica, dado $\{X_1, X_2, \dots, X_m\}$ um conjunto completo de solventes de $P(X)$, (ver Lancaster [25], pág. 524, Dennis [9], pág. 16), a matriz companheira pode ser representada por $C = \mathcal{V} D \mathcal{V}^{-1}$, com \mathcal{V} a matriz bloco de Vandermonde (Lancaster [25], pag. 524) dos solventes considerados, $\mathcal{V}(X_1, \dots, X_m)$ e $D = \text{Diag}(X_1, X_2, \dots, X_m)$.

Assim, a solução da equação diferencial $x(t) = P \mathcal{V} e^{D t} \mathcal{V}^{-1} w$ admite a forma,

$$x(t) = e^{X_1 t} z_1 + e^{X_2 t} z_2 + \dots + e^{X_m t} z_m,$$

com $z_1, z_2, \dots, z_m \in \mathbb{C}^n$.

Se a matriz A_0 da equação $P\left(\frac{d}{dt}\right)x = 0$ for singular então $\det(P(\lambda))$ tem grau $p < mn$ a que correspondem apenas p valores próprios finitos. O valor próprio infinito de $P(\lambda)$ corresponde ao valor próprio nulo da λ -matriz $\lambda^m P(1/\lambda)$ (ver Gohberg [15], pág. 184, Tisseur [36], pág. 250). Para efeito de cálculo da solução da equação homogênea interessam apenas os pares próprios relacionados com os valores próprios finitos. No caso de a equação completa, $P\left(\frac{d}{dt}\right)x = f(t)$, a determinação da respetiva solução, é feita com base no “resolvente” de $P(\lambda)$, $P(\lambda)^{-1}$.

Nesta situação, pode ainda considerar-se a linearização de $P(\lambda)$ do tipo, $\lambda \mathcal{C}_2 - \mathcal{C}_1$. Agrupando os pares próprios associados aos diversos valores próprios finitos em (V_F, J_F) e designando o par próprio associado ao valor próprio infinito por (V_∞, J_∞) (ver Gohberg [15], pág. 188), então a matriz de semelhança, generaliza-se para a forma, dada em (2.3.12), com

$$\lambda \mathcal{C}_2 - \mathcal{C}_1 = S_m K_{(c_1, c_2)} S^{-1},$$

onde a matriz $K_{(c_1, c_2)}$ representada a Forma Canónica de Weierstrass (ver Gantmacher [12], pág. 29, Dooren [7], pág. 111) do feixe $\lambda \mathcal{C}_2 - \mathcal{C}_1$. A solução geral da equação diferencial completa está definida por (ver Gohberg [15], pág. 219, Lancaster [25], pág. 334, 339, 493, 510),

$$x(t) = V_F e^{J_F t} w + \int_{t_0}^t V_F e^{-J_F(s-t)} Z_F f(s) ds - \sum_{i=0}^{v-1} V_\infty J_\infty^i Z_\infty f^{(i)}(t),$$

onde w é um vetor arbitrário de \mathbb{C}^p , e v tal que $J_\infty^v = 0$.

Relativamente à resolução total ou parcial de um sistema de equações diferenciais está implícita a distinção entre os métodos numéricos usados, uma vez que, do ponto de vista analítico, a solução $x(t)$ pode ser obtida, quer a partir dos solventes do polinômio matricial $P(X)$, quer a partir dos valores e vetores próprios de $P(\lambda)$ ou dos blocos valores próprios da matriz companheira, o que será estudado nas seções 2.4 e Capítulo 4.

Desta forma, alguns métodos que permitem a computação de solventes são descritos a seguir. O método Newton, que corresponde a uma adaptação do método Newton clássico (ver Higham [18], pág. 306, Kratz [21], pág. 359, Pereira [33], pág. 2) ao contexto dos polinômios matriciais, o método de Bernoulli, versão matricial (ver Lancaster [22], pág. 213, Tsai [38], pág. 683, Higham [19], pág. 509), que é essencialmente o método da Potência aplicado à matriz companheira \mathcal{C} e o método de Traub (assim designado por se tratar de uma generalização do método de Traub escalar, ver Traub [37], pág. 138) que permite obter uma aproximação do solvente dominante de $P(X)$ (ver Dennis [9], [10], [11], pág. 44, pág. 844 e pág. 524 respetivamente).

Da mesma forma, consideram-se alguns métodos iterativos, que permitem a computação de valores e vetores próprios quer da λ -matriz associada (ver Tisseur [36], pág. 271), quer dos blocos valores próprios da linearização $\lambda I - \mathcal{C}$ (ver Dennis [9], pág. 83). Os métodos diretos abordam o problema a partir de uma das suas linearizações possíveis $\lambda B - A$, através da redução das matrizes A e B à forma de Schur generalizada complexa, $A = QSZ^*$ e $B = QTZ^*$ onde Q e Z são matrizes unitárias e S e T são matrizes triangulares superiores (ou no caso real, Q e Z são matrizes ortogonais e S e T são matrizes quase-triangulares superiores, se pelo menos alguma das raízes for complexa).

As variantes deste tipo de métodos resultam apenas da linearização usada, uma vez que a forma de Schur não admite generalização ao caso não linear (ver Tisseur [36], pág. 266), impossibilitando, por isso, a sua aplicação ao problema original $P(\lambda)$.

Se $P(\lambda)$ for regular e $\lambda B - A$ uma linearização, considerando as matrizes triangulares superiores, $S = [s_{i,j}]_{\substack{i=0,\dots,n \\ j=1,\dots,n}}$ e $T = [t_{i,j}]_{\substack{i=0,\dots,n \\ j=1,\dots,n}}$, então (ver Tisseur [36], pág. 267)

$$\Lambda(P) = \{\lambda_i = \frac{s_{ii}}{t_{ii}}, i = 1, \dots, 2n\},$$

onde $\lambda_i = \frac{s_{ii}}{t_{ii}} := \infty$ se $t_{ii} = 0$.

Para sistemas de grande dimensão $n \gg 1$, deixa de ser viável a utilização dos métodos diretos referidos, porque ao considerarem uma linearização $\lambda B - A$ (porque se baseiam na forma de Schur) a dimensão do problema é ainda incrementada (de n passa para $n \times m$).

Os métodos iterativos que consideram o problema original $P(\lambda)$ são, basicamente, variantes do método de Newton enquanto que os métodos desenvolvidos a partir da linearização do problema baseiam-se no método de projeção em subespaços de Krylov de $P(\lambda)$, nomeadamente os métodos de Arnoldi e de Lanczos (ver Tisseur [36], pág. 272, Datta [4], pág. 446).

Relativamente a métodos iterativos com aplicação no cálculo de blocos valores próprios da equação $AV = VX$, pelo que se conseguiu apurar através da consulta da diversa documentação e bibliografia existente, identificou-se apenas o método da Potência (ver Dennis [9], pág. 86).

1.3 Objetivos

Nesta secção são traçados os objetivos gerais deste trabalho e a motivação para o desenvolvimento dos métodos de cálculo referidos na introdução. Como foi visto, a caracterização da matriz companheira \mathcal{C} (ver Lancaster [25], pág. 491, Gohberg [15], pág. 13) de um polinômio matricial $P(X)$, ou de uma forma mais geral, do feixe companheiro $\lambda \mathcal{C}_2 - \mathcal{C}_1$ (ver Gohberg [15], pág. 187), pode ser equacionada em termos de solventes de $P(X)$ ou ainda em termos de blocos valores próprios, isto é, soluções da equação $\mathcal{C}_1 V = \mathcal{C}_2 V X$, onde o bloco vetor próprio associado V tem característica máxima. Portanto afigura-se como algo essencial nessa caracterização o cálculo numérico de solventes ou de blocos valores próprios. Relativamente ao cálculo de solventes, após a análise dos diversos métodos referidos, considerou-se a ideia de abordar o problema em termos das entradas

das matrizes envolvidas, com o objetivo de evitar a álgebra matricial. Para esse objetivo foi fundamental o estudo do método dos produtos de Kronecker (ver Hernandez [14], pág. 28 e apêndice pág. 3), usado na análise e determinação de soluções de diversas equações algébricas matriciais, como por exemplo a já citada equação de Sylvester $AX - XB = C$. A ideia que está na base deste método que permite transformar o polinômio matricial numa equação vetorial equivalente, à qual será possível aplicar os métodos iterativos próprios dos sistemas de equações não lineares. Desta maneira, são apresentadas generalizações do método do Ponto Fixo (ver Marcos [27], pág. 356) e do método de Newton (Marcos [28]).

A formulação que resultou da agregação destes dois procedimentos, a transformação da equação polinomial matricial original numa equação vetorial equivalente e a posterior aplicação dos métodos clássicos, permitiu construir uma adaptação destes métodos iterativos ao contexto dos polinômios matriciais numa forma vetorial.

Relativamente ao método do Ponto Fixo (ver Marcos [27], pág. 356), apresentado na secção 3.2, a sua conceção consiste em resolver cada equação $vec(P(X))_l = 0$, do sistema $vec(P(X)) = 0$, em ordem à maior potência de x_l , coordenada com o mesmo índice do vetor $x = vec(X)$, resultando na equação $x = f(x)$ e no Algoritmo 3.2.1. A convergência é estabelecida com base no Teorema do Ponto Fixo de Schauder e na respetiva estabilidade assintótica (ver Shih [35], pág. 144), também conhecida por conjectura de Belitskiĭ e Lyubich.

A construção do chamado método de Newton Vetorial (Marcos [28]), apresentada na secção 3.3, resulta da aplicação do método de Newton clássico à equação vetorial $F(x) = vec(P(X)) = 0$, onde a derivada é formulada em termos de matriz jacobiana evitando-se o uso da derivada de Frechét e a resolução da equação de Sylvester em cada iteração, tal como acontece no método de Newton Matricial definido em Higham [18], pág. 306. O método de Newton Vetorial está definido através do Algoritmo 3.3.1 e a sua convergência fica estabelecida com o Teorema de Kantorovich, válido para uma função entre espaços de Banach.

Como foi mencionado na parte final da secção 1.2, após pesquisa da diversa documentação existente, foi apenas confirmado o método da Potência (ver Dennis [9], pág. 83) como a única alternativa para o cálculo de blocos valores próprios. Para além disso, este método está apenas definido para um feixe do tipo $(A, I) = \lambda I - A$, isto é blocos valores próprios soluções da equação $AV = VX$. É na sequência desta lacuna que se adaptou a mesma formulação ao cálculo de blocos valores próprios de $C_1V = C_2VX$. Primeiramente, a adaptação foi feita relativamente a um feixe do tipo $\lambda C_2 - C_1$, no capítulo 4, onde é apresentada uma aplicação do Método Newton Vetorial para Blocos Valores Próprios para um feixe genérico $(A, B) = \lambda B - A$, Algoritmo 4.1.1.

Por último generalizou-se ainda esta formulação para o cálculo de feixes próprios,

$$(\lambda B - A)V = W(\lambda Y - X),$$

através da construção do método de Newton Vetorial para Feixes Próprios, definido para um feixe genérico $(A, B) = \lambda B - A$ através do Algoritmo 4.2.1.

1.4 Estrutura

Este trabalho encontra-se estruturado da seguinte forma.

No capítulo 2, secções 2.1 e 2.2, são apresentadas as definições e os resultados fundamentais de Álgebra Linear Matricial que permitem a caracterização da matriz companheira ou feixe companheiro, quer em função dos solventes do respetivo polinômio matricial $P(X)$, quer em função dos valores e vetores próprios da λ -matriz associada. Os métodos numéricos mencionados nos capítulos 3 e 4

são, por este motivo, classificados segundo a caracterização escolhida.

A parte original deste trabalho é desenvolvida nos capítulos 3 e 4, seções 3.2, 3.3, 4.1 e 4.2. Os métodos aqui desenvolvidos têm na sua essência a transformação da equação polinomial matricial original numa equação vetorial equivalente, através dos resultados relativos ao produto de Kronecker (ver Hernandez [14], pág. 3 do apêndice), evitando-se a álgebra matricial e beneficiando da possibilidade de aplicação dos métodos clássicos à nova equação vetorial.

No capítulo 3, seção 3.2, faz-se a construção do método do Ponto Fixo (PF), onde a coordenada índice l da equação vetorial $vec(P(X)) = 0$ é resolvida em ordem à maior potência de x_l , coordenada com o mesmo índice de $x = vec(X)$, obtendo-se a equação $x = f(x)$ e resultando no Algoritmo 3.2.1.

Na seção 3.3, faz-se a construção do chamado método de Newton Vetorial (NV). Como referido, este método resulta da aplicação do método de Newton-Raphson clássico à equação vetorial $vec(P(X)) = 0$ (ao contrário do método de Newton definido em Higham [18], pág. 306, e Kratz [21], pág. 359, onde se considera a equação original $P(X) = 0$) evitando-se, deste modo, o uso da derivada de Frechét e a consequente resolução da equação de Sylvester.

No capítulo 4, adapta-se o método NV ao cálculo de blocos valores próprios X e blocos vetores próprios V , soluções da equação matricial $AV = BVX$, resultando no Método Newton Vetorial para Blocos Valores Próprios (NVBVP), Algoritmo 4.1.1.

Por último, na seção 4.2, generaliza-se o método anterior, NVBVP, para o cálculo de feixes próprios $\lambda Y - X$ de um feixe $\lambda B - A$, isto é

$$(\lambda B - A)V = W(\lambda Y - X),$$

onde os vetores V, W têm característica máxima. O método assim obtido, NVFP está definido para um feixe genérico $(A, B) = \lambda B - A$ através do Algoritmo 4.2.1.

Capítulo 2

Resultados Preliminares

Neste capítulo são apresentados resultados e definições próprios da teoria dos polinómios matriciais e/ou λ -matrizes no âmbito da Álgebra Linear Matricial com base nos diversas obras consultadas (Dennis [9], Lancaster [25], Gohberg [15], Tisseur [36], Hernandez [14]).

Neste sentido, na secção 2.1, são expostos os conceitos e definições considerados essenciais, quer relativamente aos polinómios matriciais, quer em relação às λ -matrizes, e realçadas as afinidades existentes entre os dois conceitos.

Na secção 2.2 são formalmente apresentadas as diversas formas canónicas já referidas na introdução, matriz companheira, feixe companheiro, forma canónica de Jordan e forma canónica de Weierstrass. Na secção 2.3, estabelece-se a relação entre cadeias de vetores próprios generalizados (ou de uma forma global, de pares próprios) e formas canónicas e respetivas matrizes de equivalência. Esta relação é estabelecida de forma fracionada, subsecções 2.3.1 e 2.3.9, fazendo-se a distinção entre valores próprios finitos e infinito.

Por último, na secção 2.4 institui-se a relação existente entre solventes do polinómio e cadeias de Jordan da λ -matriz e, mais geralmente, com os blocos vetores próprios e de que forma essa relação se manifesta em termos de formas canónicas e respetivas matrizes de equivalência.

2.1 Definições

Seguidamente são registadas formalmente as definições, consideradas essenciais, de polinómio, λ -matriz, solvente, vetores próprios e vetores próprios generalizados, cadeia de Jordan (ver Tisseur [36], pág. 251, Gohberg [15], pág. 24, Lancaster [25], pág. 504), função vetor e produto de Kronecker (ver Horn [20], pág. 243, Grassó [29] pág. 58).

Na parte final desta secção é também apresentado um exemplo ilustrativo contendo o cálculo detalhado de vetores próprios generalizados e das correspondentes cadeias de Jordan.

Definição 2.1.1. (Lancaster [25], pág. 488 e 520) Sejam $A_0, A_1, \dots, A_m \in \mathbb{C}^{n \times n}$, com A_0 uma matriz não nula,

- os polinómios matriciais associados às matrizes

$$P(X) = A_0 X^m + A_1 X^{m-1} + \dots + A_m, \quad (2.1.1)$$

e

$$\hat{P}(X) = X^m A_0 + X^{m-1} A_1 + \dots + A_m. \quad (2.1.2)$$

de grau m na variável $X \in \mathbb{C}^{n \times n}$;

- e a λ -matriz associada

$$P(\lambda) = \lambda^m A_0 + \lambda^{m-1} A_1 + \dots + A_m, \quad (2.1.3)$$

que representa uma matriz com entradas em $\mathbb{C}[\lambda]$ de grau menor ou igual m .

Os polinômios P e \hat{P} definidos em (2.1.1) e (2.1.2) dizem-se mônicos se $A_0 = I_n$.

Definição 2.1.2. (Gantmacher [13], pág. 145, Gantmacher [12], pág. 24) Chama-se feixe de matrizes ou simplesmente feixe a uma λ -matriz de grau 1,

$$\lambda A_0 + A_1, \quad (2.1.4)$$

ficando este identificado pelo par $(-A_1, A_0)$.

Por exemplo, o feixe $\lambda I - A$ é o par (A, I) e o feixe $\lambda \mathcal{C}_2 - \mathcal{C}_1$ fica identificado pelo par $(\mathcal{C}_1, \mathcal{C}_2)$.

Observação 2.1.3. Uma λ -matriz $P(\lambda)$ diz-se regular se o polinômio $\det(P(\lambda))$ não é identicamente nulo. Caso contrário, se $\det(P(\lambda)) \equiv 0$, diz-se singular. No decorrer deste trabalho, se nada for dito em contrário, consideram-se λ -matrizes regulares.

Definição 2.1.4. (Lancaster [25], pág. 520) Uma matriz $X \in \mathbb{C}^{n \times n}$, solução de $P(X) = 0$, onde 0 é a matriz nula de $\mathbb{C}^{n \times n}$, é designada por solvente de $P(X)$. Isto é, a matriz X verifica

$$A_0 X^m + A_1 X^{m-1} + \dots + A_m = 0. \quad (2.1.5)$$

Atendendo à não comutatividade das matrizes em (2.1.5), X é também designando por solvente à direita e, desse ponto de vista, uma matriz $R \in \mathbb{C}^{n \times n}$ diz-se um solvente à esquerda se

$$R^m A_0 + R^{m-1} A_1 + \dots + A_m = 0. \quad (2.1.6)$$

Isto é, R diz-se um solvente à esquerda se for um solvente do polinômio $\hat{P}(X)$. Evidentemente que se a matriz R comutar com todas as matrizes coeficientes ($RA_i = A_i R, \forall i = 0, 1, \dots, m$) então R é simultaneamente um solvente à direita e à esquerda de $P(X)$.

Além da questão da comutatividade a designação direita e esquerda está relacionada com a fatorização que se obtém para a λ -matriz associada $P(\lambda)$. Por exemplo, para $m = 2$, dado X um solvente à direita de $P(X) = A_0 X^2 + A_1 X + A_2$ o problema PVPQ pode caracterizar-se através da fatorização (Tisseur [36], pág. 252)

$$\lambda^2 A_0 + \lambda A_1 + A_2 = (\lambda A_0 + A_1 + A_0 X)(\lambda I - X),$$

obtendo-se uma fatorização idêntica para um solvente à esquerda R , mas com o fator $(\lambda I - R)$ situado precisamente à esquerda do produto. Em Grassó [29], pág. 88, estão definidos resultados relativos à fatorização de λ -matrizes de grau m qualquer, onde são identificadas condições necessárias e suficientes para que uma λ -matriz admita uma fatorização linear

$$P(\lambda) = (\lambda A_0 - X_1)(\lambda I - X_2) \dots (\lambda I - X_m),$$

para um conjunto completo de solventes de $P(X)$.

Ainda relativamente a uma λ -matriz regular $P(\lambda)$ considera-se a seguinte Definição.

Definição 2.1.5. (ver [15], pág. 24) Um escalar $\lambda_0 \in \mathbb{C}$, tal que $\det(P(\lambda_0)) = 0$, diz-se um valor próprio de $P(\lambda)$. Dado $\lambda_0 \in \mathbb{C}$ um valor próprio de $P(\lambda)$, um vetor $v \in \mathbb{C}^n \setminus \{0\}$, diz-se um vetor próprio à direita de $P(\lambda)$ se verifica

$$P(\lambda_0)v = 0.$$

O vetor v diz-se um vetor próprio à esquerda de $P(\lambda)$ se verifica $v^* P(\lambda_0) = 0$, com $v^* = \bar{v}^T$.

Duas λ -matrizes $P(\lambda)$ e $Q(\lambda)$ dizem-se equivalentes (ver Gantmacher [13], pág. 133) se existirem λ -matrizes $E(\lambda)$ e $F(\lambda)$ com determinantes não nulos e independentes de λ tais que

$$P(\lambda) = E(\lambda)Q(\lambda)F(\lambda). \quad (2.1.7)$$

Sempre que $P(\lambda)$ e $Q(\lambda)$ forem λ -matrizes equivalentes usa-se a seguinte notação:

$$P(\lambda) \sim Q(\lambda). \quad (2.1.8)$$

Observação 2.1.6. No caso de as λ -matrizes consideradas serem regulares e de grau $m = 1$ então as λ -matrizes $E(\lambda)$ e $F(\lambda)$ referidas podem ser substituídas (ver Gantmacher [13], pág. 145) por matrizes constantes (corresponde à noção de λ -matrizes estritamente equivalentes).

O determinante da λ -matriz $P(\lambda)$ terá grau mn se A_0 for não singular e, nesta situação, $P(\lambda)$ é necessariamente regular e terá mn valores próprios em \mathbb{C} simples ou múltiplos (neste caso, contabilizando as multiplicidades).

Observação 2.1.7. Se A_0 for singular então $\det(P(\lambda))$ terá um grau p inferior a mn a que correspondem p valores próprios finitos. Nesta situação, considera-se o valor próprio infinito de $P(\lambda)$ corresponde ao valor próprio nulo da λ -matriz $Q(\lambda) = \lambda^m P(1/\lambda)$.

Com efeito (ver Gohberg [15], pag. 185 e 388, Dooren [7], pág. 546), a função $F(\lambda) = \lambda^{-m} P(\lambda) = A_0 + \lambda^{-1} A_1 + \dots + \lambda^{-m} A_m$ é analítica em ∞ pois $F(1/\lambda) = \lambda^m P(1/\lambda)$ é analítica em 0. Se o valor $\lambda = \infty \in \overline{\mathbb{C}}$ é um valor próprio de $F(\lambda)$, uma cadeia de Jordan de $F(\lambda)$ associada a $\lambda = \infty$ é também uma cadeia de Jordan de $F(1/\lambda)$ associada ao valor próprio $\lambda = 0$. Ou, de forma equivalente, a λ -matriz $Q(\lambda) = F(1/\lambda) = \lambda^m P(1/\lambda) = A_0 + \lambda A_1 + \dots + \lambda^m A_m$ admite o valor próprio nulo com multiplicidade algébrica $mn - p$.

Definição 2.1.8. (ver [15], pág. 24) O conjunto formado por todos os valores próprios (finitos ou infinitos) de $P(\lambda)$ é designado por espectro de P e representa-se por

$$\Lambda(P) = \{\lambda \in \mathbb{C} : \det(P(\lambda)) = 0\}.$$

Além da multiplicidade algébrica de um valor próprio λ_i , importa também determinar a multiplicidade geométrica que designa a dimensão do núcleo da aplicação linear $P(\lambda_i) : \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{n \times n}$, $\ker(P(\lambda_i))$. Um valor próprio diz-se simples (ver Tisseur [36], pág. 251) se ambas as multiplicidades forem iguais a 1. Diz-se semi-simples se as multiplicidades forem iguais entre si e maiores do que 1 e defectivo caso contrário. Evidentemente, se a multiplicidade algébrica de λ_i for superior a n , este valor próprio é necessariamente defectivo uma vez que a multiplicidade geométrica é no máximo n .

Paralelamente a esta terminologia, Wilkinson [39], pág. 13, designa por matriz não derogatória uma matriz que todos os seus valores próprios tenham multiplicidade geométrica 1. Em termos de polinômios matriciais, uma matriz A é não derogatória se todos os valores próprios da λ -matriz $\lambda I - A$ tiverem multiplicidade geométrica 1. Ainda segundo Wilkinson [39], pág. 15, uma matriz A é semelhante à matriz companheira do seu polinômio caraterístico se e só se A for uma matriz não derogatória.

De uma forma geral e independentemente da terminologia, se λ_i for um valor próprio defectivo com multiplicidade algébrica m_i , por definição, o número p_i de vetores próprios ¹ associados a λ_i será

¹Vetores próprios linearmente independentes, bem entendido, cujo número corresponde à $\dim(\ker(P(\lambda_i)))$.

inferior a m_i . Isto é, relativamente a um valor próprio defectivo λ_i , é necessário alargar o conceito de vetor próprio de modo a obter-se um número de vetores associados igual à multiplicidade algébrica do valor próprio λ_i . Assim, diz-se que $v_2 \in \mathbb{C}^n$ é um vetor próprio generalizado se, para algum vetor próprio $v_1 \neq 0$, se tem

$$P(\lambda_i)v_2 + P'(\lambda_i)v_1 = 0, \quad (2.1.9)$$

e, de uma forma geral:

Definição 2.1.9. (ver [15], pág. 25) O conjunto formado pelos l vetores $v_1, v_2, \dots, v_l \in \mathbb{C}^n$ diz-se uma cadeia de vetores próprios generalizados (cadeia de Jordan) de comprimento l de $P(\lambda)$ relativa ao valor próprio λ_i se verifica:

$$\begin{aligned} P(\lambda_i)v_1 &= 0 \\ P(\lambda_i)v_2 + P'(\lambda_i)v_1 &= 0 \\ P(\lambda_i)v_3 + P'(\lambda_i)v_2 + \frac{1}{2}P''(\lambda_i)v_1 &= 0 \\ &\vdots = \vdots \\ P(\lambda_i)v_l + P'(\lambda_i)v_{l-1} + \dots + \frac{1}{(l-2)!}P^{(l-2)}(\lambda_i)v_2 + \frac{1}{(l-1)!}P^{(l-1)}(\lambda_i)v_1 &= 0, \end{aligned} \quad (2.1.10)$$

onde $v_1 \neq 0$ é um vetor próprio de $P(\lambda_i)$.

Observação 2.1.10. O valor l é também designado por multiplicidade parcial relativa ao valor próprio considerado λ_i . A soma das multiplicidades parciais é igual à multiplicidade algébrica do valor próprio λ_i (ver (Pereira [30], pág. 20, Lancaster [25], pág. 504). Isto é, se λ_i tiver multiplicidade algébrica m_i e tiver s_i cadeias de Jordan (s_i a multiplicidade geométrica de λ_i) de comprimento $m_{i,j}$, $j = 1, \dots, s_i$, então $\sum_{j=1}^{s_i} m_{i,j} = m_i$.

Observação 2.1.11. Os vetores próprios generalizados (para $m > 1$) não têm que ser necessariamente linearmente independentes, podendo mesmo acontecer que uma cadeia de Jordan contenha vetores nulos (ver Tisseur [36], pág. 251, Gohberg [15], pág. 24, Lancaster [25], pág. 504).

Para realçar esta ideia, apresenta-se de seguida um exemplo que se julga ser elucidativo pois é detalhado o cálculo dos vetores próprios generalizados, relativamente a cada valor próprio, e faz-se a identificação das correspondentes cadeias de Jordan.

Exemplo 2.1.12. Seja a λ -matriz,

$$P(\lambda) = \begin{bmatrix} \lambda^2 - 4\lambda + 4 & 2 - \lambda \\ 0 & \lambda^2 - 4\lambda + 4 \end{bmatrix},$$

com $\det(P(\lambda)) = \lambda^4 - 8\lambda^3 + 24\lambda^2 - 32\lambda + 16 = (\lambda - 2)^4$.

$P(\lambda)$ tem um único valor próprio $\lambda = 2$ de multiplicidade 4. Assim,

$$\begin{aligned} P(\lambda) &= \begin{bmatrix} \lambda^2 - 4\lambda + 4 & 2 - \lambda \\ 0 & \lambda^2 - 4\lambda + 4 \end{bmatrix} \longrightarrow P(2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ P'(\lambda) &= \begin{bmatrix} 2\lambda - 4 & -1 \\ 0 & 2\lambda - 4 \end{bmatrix} \longrightarrow P'(2) = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \\ P''(\lambda) &= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \longrightarrow P''(2) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}. \end{aligned}$$

Tem-se $P(2)\vec{v} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, para qualquer vetor $\vec{v} \neq \vec{0}$.

$$P(2)\vec{u} + P'(2)\vec{v} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -v_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

obriga a que $v_2 = 0$.

$$\begin{aligned} P(2)\vec{w} + P'(2)\vec{u} + P''(2)\vec{v} &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} -u_2 + v_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \end{aligned}$$

implica que $u_2 = v_1$.

Considerando

$$\begin{aligned} P(2)\vec{z} + P'(2)\vec{w} + P''(2)\vec{u} &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ v_1 \end{bmatrix} \\ &= \begin{bmatrix} -w_2 + u_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \end{aligned}$$

então $v_1 = 0$, o que não pode acontecer pois implicaria que $\vec{v} = \vec{0}$.

Portanto para $v_2 = 0$ tem-se uma cadeia de Jordan com três vetores próprios generalizados

$$\begin{bmatrix} v_1 & u_1 & w_1 \\ 0 & v_1 & w_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Para $v_2 \neq 0$, tem-se uma cadeia de Jordan apenas com um vetor próprio,

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Observação 2.1.13. No exemplo anterior, 2.1.12, fica também clara a noção de multiplicidade parcial (ver Observação 2.1.10) de um valor próprio. Nesta situação, $P(\lambda)$ tinha um único valor próprio $\lambda = 2$ de multiplicidade algébrica 4 e multiplicidade geométrica 2 (portanto defectivo), tendo-se obtido duas cadeias de Jordan de comprimentos 3 e 1. Assim, a soma das multiplicidades parciais, $3 + 1$ é igual à multiplicidade algébrica 4.

Por outro lado, foi considerado um vetor nulo na primeira cadeia de Jordan calculada.

Os resultados que generalizam o conceito de divisão de polinômios para polinômios matriciais e λ -matrizes, (ver Dennis [10], pág. 832, Gohberg [15], pág. 85), estipulam que, X é um solvente de $P(X)$ se e só se existe uma fatorização do tipo

$$P(\lambda) = Q(\lambda)(\lambda I - X), \quad (2.1.11)$$

onde $Q(\lambda)$ é um polinômio matricial de grau $m - 1$.

A equação (2.1.11) indica que os valores próprios de um solvente de $P(X)$ são ainda valores próprios

da λ -Matriz, $P(\lambda)$.

Para $m = 2$, se X for um solvente de $P(X) = A_0X^2 + A_1X + A_2$, então

$$P(\lambda) = (\lambda A_0 + A_1 + A_0X)(\lambda I - X). \quad (2.1.12)$$

No caso $m = 2$, equação (2.1.12) indica ainda que os valores próprios da λ -Matriz, $P(\lambda)$ ficam determinados pelos valores próprios do solvente X juntamente com os do feixe $(-A_1 - A_0X, A_0)$, ambos problemas lineares, $\det(\lambda I - X) = 0$ e $\det(\lambda A_0 + A_1 + A_0X) = 0$.

Na teoria das equações matriciais lineares, entre as quais se destacam a equação de Silvester

$$AX - BX = C,$$

e a equação de Stein

$$X - CXD = E,$$

são introduzidas as seguintes definições (ver Horn [20], pág. 243, Grassó [29] pág. 58), cuja aplicação é recorrente durante este trabalho.

Definição 2.1.14. (ver Horn [20], pág. 243) Para $X \in \mathbb{C}^{m \times n}$, $X = [x_{i,j}]_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$, $vec(X)$ representa o vetor $x \in \mathbb{C}^{mn}$ definido por

$$vec(X) = (x_1^T, \dots, x_n^T)^T,$$

onde $x_1, \dots, x_n \in \mathbb{C}^m$ são as colunas de X .

Definição 2.1.15. (ver Horn [20], pág. 243) Para $X \in \mathbb{C}^{m \times n}$, $Y \in \mathbb{C}^{p \times q}$, com $X = [x_{i,j}]_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$ e $Y = [y_{i,j}]_{\substack{i=1,\dots,p \\ j=1,\dots,q}}$, a matriz

$$X \otimes Y = \begin{bmatrix} x_{1,1}Y & x_{1,2}Y & \cdots & x_{1,n}Y \\ x_{2,1}Y & x_{2,2}Y & \cdots & x_{2,n}Y \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,1}Y & x_{m,2}Y & \cdots & x_{m,n}Y \end{bmatrix},$$

representa o produto de Kronecker, $X \otimes Y \in \mathbb{C}^{mp \times nq}$, das matrizes X e Y .

Observação 2.1.16. Relativamente à função $vec : \mathbb{C}^{m \times n} \rightarrow \mathbb{C}^{mn}$, a notação $x = vec(X) \in \mathbb{C}^{mn}$, dependendo do contexto, pode designar tanto a matriz coluna em $x \in \mathbb{C}^{mn \times 1}$ como o vetor ou ponto $x \in \mathbb{C}^{mn}$ definido pelas mn coordenadas (x_1, \dots, x_{mn}) .

Observação 2.1.17. Fixo o valor de n , está definida a aplicação inversa $vec^{-1} : \mathbb{C}^{mn} \rightarrow \mathbb{C}^{m \times n}$, onde $X = vec^{-1}(x) \in \mathbb{C}^{m \times n}$, representa a matriz definida por

$$\begin{bmatrix} x_1 & x_{m+1} & \cdots & x_{mn-m+1} \\ x_2 & x_{m+2} & \cdots & x_{mn-m+2} \\ \vdots & \vdots & \ddots & \vdots \\ x_m & x_{2m} & \cdots & x_{mn} \end{bmatrix}.$$

Observação 2.1.18. Utiliza-se, por uma questão de simplicidade, a notação $vec(V, X)$, onde V e X são matrizes de diferentes dimensões, $V \in \mathbb{C}^{m \times n}$, $X \in \mathbb{C}^{n \times n}$, tendo como significado o vetor $(v, x) = (vec(V)^T, vec(X)^T)^T \in \mathbb{C}^{mn+n^2}$. Da mesma forma, quando perfeitamente identificados

m e n , a notação $vec^{-1}(v, x)$, deve ser entendida como o par $(V, X) = (vec^{-1}(v), vec^{-1}(x))$ com $V \in \mathbb{C}^{m \times n}$ e $X \in \mathbb{C}^{n \times n}$.

Um dos possíveis métodos diretos usados na resolução de equações matriciais, o chamado o método dos produtos de Kronecker (ver Hernandez [14], pág. 100, Grassó [29], pág. 59), tem como base o seguinte resultado que está na base da formulação apresentada nas secções 3.2, 3.3, 4.1 e 4.2.2.

Lema 2.1.19. (ver Horn [20], pág. 254) Sejam A, B, X matrizes, com dimensões apropriadas, para as quais o produto AXB está definido. Então

$$vec(AXB) = (B^T \otimes A)vec(X).$$

2.2 Formas Canónicas

Como referido na introdução, uma linearização de uma λ -matriz de grau m , $P(\lambda)$, corresponde à redução desta ao caso linear, $\lambda I - A$ ou $\lambda B - A$. Tem-se que (ver Lancaster [25], pág. 491, Gohberg [15], pág. 13), a λ -matriz

$$\begin{bmatrix} P(\lambda) & 0_{n \times (m-1)n} \\ 0_{(m-1)n \times n} & I_{(m-1)n} \end{bmatrix},$$

é equivalente a $\lambda I - C$, onde C é a matriz companheira de $P(\lambda)$ (ver Gohberg [15], pág. 491), no caso mónico, ou equivalente ao feixe companheiro de $P(\lambda)$, $\lambda C_2 - C_1$ (ver Gohberg [15], pág. 186, Tisseur [36], pág. 253), no caso não mónico.

Atendendo à transitividade da relação de equivalência entre λ -matrizes, uma linearização do polinómio $P(\lambda)$ é qualquer λ -matriz da forma $\lambda I_{mn} - C$, com $C \in \mathbb{K}^{mn \times mn}$, equivalente a $\lambda I - C$, no caso mónico ou equivalente ao feixe companheiro no caso não mónico (ver Gohberg [15], pág. 187).

No entanto, as linearizações consideradas neste trabalho e descritas a seguir, consistem em considerar $A = C$, no caso mónico, ou $A = C_1$ e $B = C_2$, no caso não mónico. São ainda expostas as formas canónicas já referidas na introdução, forma canónica de Jordan e forma canónica de Weierstrass.

Definição 2.2.1. (ver Lancaster [25], pág. 491) Seja $P(X)$ o polinómio mónico,

$$P(X) = X^m + A_1 X^{m-1} + \dots + A_m, \quad (2.2.1)$$

com $A_i \in \mathbb{C}^{n \times n}$, $i = 1, \dots, m$. Chama-se matriz companheira de $P(X)$ (e de $P(\lambda)$) à matriz $C \in \mathbb{C}^{nm \times nm}$ definida por

$$C = \begin{bmatrix} 0_n & I_n & \cdots & 0_n \\ 0_n & 0_n & \ddots & \vdots \\ \vdots & \vdots & \ddots & I_n \\ -A_m & -A_{m-1} & \cdots & -A_1 \end{bmatrix}. \quad (2.2.2)$$

Teorema 2.2.2. (Gohberg [15], pág. 491) Seja $P(\lambda)$ a λ -matriz associada a um polinómio matricial mónico $P(X)$ e C a respetiva matriz companheira.

Então $\lambda I - C$ é uma linearização da λ -matriz $P(\lambda)$. Isto é, existem λ -matrizes $E(\lambda)$ e $F(\lambda)$ com determinantes não nulos e independentes de λ tais que

$$\begin{bmatrix} P(\lambda) & 0 \\ 0 & I \end{bmatrix} = E(\lambda)(\lambda I - C)F(\lambda),$$

ou simplesmente,

$$(\lambda I - C) \sim \begin{bmatrix} P(\lambda) & 0 \\ 0 & I \end{bmatrix}. \quad (2.2.3)$$

Em particular,

$$\det(\lambda I - C) = \det(P(\lambda)), \quad (2.2.4)$$

pelo que os valores próprios de $\lambda I - C$ e $P(\lambda)$ coincidem e têm a mesma multiplicidade algébrica. Na realidade a equivalência 2.2.3 obriga a que, para cada valor próprio, as multiplicidades parciais (ver Observação 2.1.10) também sejam coincidentes.

Observação 2.2.3. Se o polinômio $P(X)$ for não mónico,

$$P(X) = A_0 X^m + A_1 X^{m-1} + \dots + A_m. \quad (2.2.5)$$

a linearização toma a forma de $\lambda C_2 - C_1$, (ver Gohberg [15], pág. 186, Tisseur [36], pág. 253). Ou seja,

$$\begin{bmatrix} P(\lambda) & 0 \\ 0 & I \end{bmatrix} = E(\lambda)(\lambda C_2 - C_1)F(\lambda), \quad (2.2.6)$$

com

$$C_2 = \begin{bmatrix} I_n & 0 & \dots & 0 \\ 0 & I_n & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & A_0 \end{bmatrix} \quad \text{e} \quad C_1 = \begin{bmatrix} 0 & I_n & \dots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & I_n \\ -A_m & -A_{m-1} & \dots & -A_1 \end{bmatrix}.$$

À λ -matriz, de grau 1,

$$\lambda C_2 - C_1 \quad (2.2.7)$$

definida em (2.2.6), chama-se (ver Gohberg [15], pág. 186, Tisseur [36], pág. 253) feixe companheiro de $P(\lambda)$.

Sempre que $A_0 = I$ então

$$C_1 = C \quad \text{e} \quad C_2 = I,$$

pelo que, esta linearização coincide com a anterior (2.2.3) quando o polinômio $P(X)$ for mónico.

Tal como referido na introdução, são também enunciadas a seguir as respetivas formas canónicas e matrizes de equivalência, quer da matriz companheira, quer do feixe companheiro, e estabelece-se a relação destas com as cadeias de vetores próprios generalizados.

Definição 2.2.4. (Gohberg [15], pág. 47) Um bloco de Jordan $J_k(\lambda_i)$ é uma matriz triangular superior de dimensão $k \times k$ com λ_i na diagonal principal e da forma:

$$J_k(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{bmatrix}. \quad (2.2.8)$$

Teorema 2.2.5. (ver Horn [20], pág. 126) Uma matriz $A \in \mathbb{C}^{n \times n}$ é semelhante a uma única² matriz, soma direta de blocos de Jordan, da forma:

$$J = \begin{bmatrix} J_{k_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{k_2}(\lambda_2) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & J_{k_l}(\lambda_l) \end{bmatrix}, \quad (2.2.9)$$

onde $J_{k_i}(\lambda_i)$ são bloco de Jordan, com $\sum_i k_i = n$. Isto é, existe uma matriz $S \in \mathbb{C}^{n \times n}$ não singular tal que $A = SJS^{-1}$.

Dada a forma única da matriz (2.2.9), esta designa-se por forma canónica de Jordan de A e representa-se por

$$J_A = \text{Diag}(J_{k_1}(\lambda_1), J_{k_2}(\lambda_2), \dots, J_{k_l}(\lambda_l)).$$

O equivalente à forma canónica de Jordan de uma matriz A para um feixe de matrizes regular $(A, B) = \lambda B - A$, com $A, B \in \mathbb{C}^{n \times n}$, é a forma canónica de Weierstrass³ (ver Gantmacher [12], pág. 29, Dooren [7], pág. 111) definida por

$$K_{(A,B)} = \begin{bmatrix} \lambda I - J & 0 \\ 0 & \lambda N - I \end{bmatrix}, \quad (2.2.10)$$

com J e N matrizes de Jordan do tipo

$$J = \text{Diag}(J_{k_1}(\lambda_1), J_{k_2}(\lambda_2), \dots, J_{k_l}(\lambda_l)),$$

onde os valores próprios λ_i não têm que ser distintos, e

$$N = \text{Diag}(J_{j_1}(0), J_{j_2}(0), \dots, J_{j_p}(0)),$$

onde N é uma matriz nilpotente ($N^L \equiv 0$, onde L é a multiplicidade algébrica do valor próprio infinito, $L = \sum_l j_l$) e $\sum_i k_i + \sum_l j_l = n$.

Teorema 2.2.6. (Lancaster [24], pág. 26) Se $\lambda B - A$ for um feixe regular, então existem matrizes $P, Q \in \mathbb{C}^{n \times n}$ não singulares tais que

$$P(\lambda B - A)Q = \begin{bmatrix} \lambda I - J & 0 \\ 0 & \lambda N - I \end{bmatrix},$$

onde a matriz J contém os valores próprios finitos de $\lambda B - A$ e N contém apenas o valor próprio nulo de $\lambda A - B$, que corresponde ao valor próprio infinito de $\lambda B - A$.

2.3 Valores e vetores próprios da λ -matriz $P(\lambda)$

De acordo com o Teorema 2.2.2, os valores próprios da λ -Matriz, $P(\lambda)$ e da matriz companheira \mathcal{C} , ou de uma forma geral, do feixe companheiro 2.2.6, assim como as respetivas multiplicidades, coincidem. Este resultado corresponde a uma linearização do problema de cálculo de valores e

²A menos de uma permutação na ordem de disposição dos blocos de Jordan.

³A letra K usada tem a ver com o fato de a forma canónica de Kronecker, no caso de um feixe regular, ser também designada por forma canónica de Weierstrass.

vetores próprios da λ -matriz (não é única mas, de certa forma, é entendida como a linearização canônica do problema, ver Gohberg [15], pág.186). A λ -Matriz, $P(\lambda)$ terá mn valores próprios (contabilizando as multiplicidades e, no caso de A_0 ser singular, considerando o valor próprio infinito).

Na subseções seguintes, estabelece-se a relação entre cadeias de vetores próprios generalizados e formas canônicas ou linearizações do problema, matriz companheira ou feixe companheiro, de forma fracionada fazendo-se a distinção entre valores próprios finitos e infinito.

2.3.1 Valores próprios finitos

Uma λ -Matriz de grau m terá mn valores próprios finitos (contabilizando as multiplicidades) sempre que o coeficiente A_0 , do termo de grau m , seja uma matriz não singular e, nesta situação, pode transformar-se numa λ -Matriz mónica (multiplicando eventualmente todos os seus coeficientes por A_0^{-1}),

$$P(\lambda) = \lambda^m + \lambda^{m-1}A_1 + \dots + A_m. \quad (2.3.1)$$

Com vista ao estabelecimento da relação entre cadeias de vetores próprios generalizados ou pares próprios e formas canônicas (matriz companheira ou a linearização correspondente $\lambda I - C$), consideram-se os seguintes resultados.

Teorema 2.3.2. (Gohberg [15], pág. 40) Considere-se a λ -Matriz (2.3.1) e $V = \begin{bmatrix} v_1 & \dots & v_k \end{bmatrix}$ uma matriz $n \times k$, onde os vetores v_1, \dots, v_k formam uma cadeia de Jordan associado ao valor próprio λ_i e J um bloco de Jordan de dimensão $k \times k$ com λ_i na diagonal principal,

$$J = \begin{bmatrix} \lambda_i & 1 & 0 \\ & \ddots & 1 \\ 0 & & \lambda_i \end{bmatrix}. \quad (2.3.2)$$

Então,

$$VJ^m + A_1VJ^{m-1} + \dots + A_mV = 0. \quad (2.3.3)$$

Fixado λ , um valor próprio de $P(\lambda)$, no caso de a multiplicidade geométrica ser $p > 1$, podem agrupar-se as várias cadeias de Jordan associadas V_1, \dots, V_p , com V_l de dimensão $n \times k_l$, $l = 1, \dots, p$, considerando o vetor V e a matriz J definidas por

$$V = \begin{bmatrix} V_1 & \dots & V_p \end{bmatrix} \text{ e } J = \text{Diag}(J_1, \dots, J_p) = \begin{bmatrix} J_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & J_p \end{bmatrix}, \quad (2.3.4)$$

onde cada J_l é um bloco de Jordan de dimensão $k_l \times k_l$, $l = 1, \dots, p$, com λ na diagonal principal.

Definição 2.3.3. (ver Pereira [30], pág. 16) Seja $P(\lambda)$ a λ -Matriz dada por (2.3.1) e V e J as matrizes dadas por (2.3.4) de dimensões $n \times k$ e $k \times k$ respetivamente com $(\sum k_l \leq k$, k a multiplicidade algébrica de λ). Nestas condições, diz-se que o par (V, J) é um par próprio de $P(\lambda)$, associado ao valor próprio λ .

Observação 2.3.4. O par próprio (V, J) não é único, da mesma forma que a cadeia de Jordan também depende da escolha dos vetores próprios (uma vez que o sistema correspondente é indeter-

minado). Para além deste aspeto, o par próprio (V, J) depende também da ordenação considerada para as diversas cadeias de Jordan em (2.3.4).

O par próprio (V, J) satisfaz ainda o Teorema 2.3.2 (ver Gohberg [15], pág.40).

Definição 2.3.5. (ver Pereira [30], pág. 18) Seja $P(\lambda)$ uma λ -Matriz mónica e sejam $(V_1, J_1), \dots, (V_l, J_l)$ pares próprios de $P(\lambda)$ associados aos valores próprios $(\lambda_1, \dots, \lambda_l)$. Se

$$\text{Diag}(J_1, \dots, J_l) = J_{\mathcal{C}}, \quad (2.3.5)$$

onde $J_{\mathcal{C}}$ representa a forma canónica de Jordan da matriz companheira \mathcal{C} , diz-se que $(V_1, J_1), \dots, (V_l, J_l)$ é um sistema completo de pares próprios de $P(\lambda)$.

Teorema 2.3.6. (ver Pereira [30], pág. 21, Gohberg [15], pág. 44) Seja $P(X)$ um polinómio matricial mónico e \mathcal{C} a matriz companheira associada e $J_{\mathcal{C}}$ a respetiva forma canónica de Jordan, com

$$\mathcal{C} = SJ_{\mathcal{C}}S^{-1}. \quad (2.3.6)$$

Então a matriz de semelhança S é da forma

$$S = \begin{bmatrix} V_1 & V_2 & \cdots & V_l \\ V_1 J_1 & V_2 J_2 & \cdots & V_l J_l \\ \vdots & & & \vdots \\ V_1 J_1^{m-1} & V_2 J_2^{m-1} & \cdots & V_l J_l^{m-1} \end{bmatrix}, \quad (2.3.7)$$

onde $(V_1, J_1), \dots, (V_l, J_l)$ é um sistema completo de pares próprios de $P(\lambda)$.

Apresenta-se a demonstração deste resultado porque está na base da relação entre blocos valores próprios e solventes. Assim, basta observar que,

$$\begin{aligned} \mathcal{C}S &= \begin{bmatrix} 0 & I_n & & 0 \\ 0 & 0 & I_n & \\ \vdots & & \ddots & \\ -A_m & -A_{m-1} & \cdots & -A_1 \end{bmatrix} \begin{bmatrix} V_1 & V_2 & \cdots & V_l \\ V_1 J_1 & V_2 J_2 & \cdots & V_l J_l \\ \vdots & & & \vdots \\ V_1 J_1^{m-1} & V_2 J_2^{m-1} & \cdots & V_l J_l^{m-1} \end{bmatrix} \\ &= \begin{bmatrix} V_1 J_1 & V_2 J_2 & \cdots & V_l J_l \\ V_1 J_1^2 & V_2 J_2^2 & \cdots & V_l J_l^2 \\ \vdots & & & \vdots \\ -\sum_{i=1}^m A_{m-i+1} V_1 J_1^{i-1} & -\sum_{i=1}^m A_{m-i+1} V_2 J_2^{i-1} & \cdots & -\sum_{i=1}^m A_{m-i+1} V_l J_l^{i-1} \end{bmatrix}, \end{aligned}$$

$$\begin{aligned} SJ_{\mathcal{C}} &= \begin{bmatrix} V_1 & V_2 & \cdots & V_l \\ V_1 J_1 & V_2 J_2 & \cdots & V_l J_l \\ \vdots & & & \vdots \\ V_1 J_1^{m-1} & V_2 J_2^{m-1} & \cdots & V_l J_l^{m-1} \end{bmatrix} \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & J_l \end{bmatrix} \\ &= \begin{bmatrix} V_1 J_1 & V_2 J_2 & \cdots & V_l J_l \\ V_1 J_1^2 & V_2 J_2^2 & \cdots & V_l J_l^2 \\ \vdots & & & \vdots \\ V_1 J_1^m & V_2 J_2^m & \cdots & V_l J_l^m \end{bmatrix}. \end{aligned}$$

onde a igualdade da última linha resulta de (2.3.3) aplicada a cada par próprio (V_i, J_i) .

Observação 2.3.7. Da mesma forma, se $\lambda_1, \dots, \lambda_l$ são todos os valores próprios da λ -Matriz, podem ainda agrupar-se os diversos pares próprios $(V_1, J_1), \dots, (V_l, J_l)$ associados em matrizes $V = \begin{bmatrix} V_1 & \dots & V_l \end{bmatrix}$ e $J = \text{Diag}(J_1, \dots, J_l)$, designado-se o par assim construído (V, J) por par de Jordan (ver Gohberg [15], pág.40) que satisfaz ainda o Teorema 2.3.2.

Com o objetivo de clarificar os resultados desta subsecção é apresentado de seguida um exemplo onde são calculados os pares próprios e identificadas as respetivas matrizes de semelhança.

Exemplo 2.3.8. Seja o polinómio matricial,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} X^2 + \begin{bmatrix} -4 & 1 \\ 0 & -3 \end{bmatrix} X + \begin{bmatrix} 4 & -2 \\ 0 & 2 \end{bmatrix},$$

cuja λ -matriz associada é,

$$P(\lambda) = \begin{bmatrix} \lambda^2 - 4\lambda + 4 & 2 - \lambda \\ 0 & \lambda^2 - 3\lambda + 2 \end{bmatrix}.$$

Tem-se que $\det(P(\lambda)) = (\lambda^2 - 4\lambda + 4)(\lambda^2 - 3\lambda + 2) = (\lambda - 1)(\lambda - 2)^3$ pelo que, $P(\lambda)$ admite o valor próprio $\lambda = 1$ e o valor próprio $\lambda = 2$ com multiplicidade 3. Assim, considerando

$$P(\lambda) = \begin{bmatrix} \lambda^2 - 4\lambda + 4 & 2 - \lambda \\ 0 & \lambda^2 - 3\lambda + 2 \end{bmatrix}, \quad P'(\lambda) = \begin{bmatrix} 2\lambda - 4 & -1 \\ 0 & 2\lambda - 3 \end{bmatrix}, \quad \text{e} \quad P''(\lambda) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix},$$

tem-se:

$$P(1)\vec{v} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_1 - v_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \text{para qualquer vetor } \vec{v} \text{ com } v_2 = v_1;$$

$$P(1)\vec{u} + P'(1)\vec{v} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} -2 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} u_1 - u_2 - 2v_1 + v_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

o que não pode acontecer pois implicaria que $v_2 = v_1 = 0$.

Portanto, para $\lambda = 1$, tem-se uma cadeia de Jordan com um vetor próprio

$$\begin{bmatrix} v_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Em relação ao valor próprio $\lambda = 2$, tem-se:

$$P(2)\vec{v} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \text{para qualquer vetor } \vec{v} \neq 0;$$

$$P(2)\vec{u} + P'(2)\vec{v} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_2 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

implica que $v_2 = 0$ e \vec{u} qualquer;

$$\begin{aligned} P(2)\vec{w} + P'(2)\vec{u} + P''(2)\vec{v} &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} u_2 + v_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \end{aligned}$$

implicaria $u_2 = 0$ e $v_1 = -u_2 = 0$ o que não pode suceder.

Portanto, para $\lambda = 2$, considerando $v_2 = 0$, obtêm-se uma cadeia de Jordan com dois vetores próprios generalizados

$$\begin{bmatrix} v_1 & u_1 \\ 0 & u_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Para $v_2 \neq 0$ a cadeia de Jordan contém apenas um vetor próprio,

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Assim, obtêm-se os seguintes pares próprios,

$$V_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad J_1 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \quad V_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{e} \quad J_2 = \begin{bmatrix} 1 \end{bmatrix},$$

ou os seguintes pares de Jordan,

$$V = \begin{bmatrix} V_1 & V_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \quad \text{e} \quad J = J_{\mathcal{C}} = \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Tem-se que,

$$\mathcal{C} = SJ_{\mathcal{C}}S^{-1} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -4 & 2 & 4 & -1 \\ 0 & -2 & 0 & 3 \end{bmatrix},$$

de acordo com a equação (2.3.6), onde

$$S = \begin{bmatrix} V_1 & V_2 \\ V_1 J_1 & V_2 J_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 2 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

Verifica-se ainda que os pares próprios $(V_1, J_1), (V_2, J_2)$ constituem um sistema completo de pares próprios de $P(\lambda)$.

2.3.9 Valor próprio infinito

Tal como referido, estabelece-se agora a relação entre cadeias de vetores próprios generalizados associados ao valor próprio infinito e respetivo feixe companheiro $\lambda\mathcal{C}_1 - \mathcal{C}_2$ que corresponde de fato a uma linearização do problema. Assim, no caso de uma λ -Matriz não mónica,

$$P(\lambda) = \lambda^m A_0 + \lambda^{m-1} A_1 + \dots + A_m, \tag{2.3.8}$$

em que o coeficiente A_0 é singular, esta terá mn valores próprios contabilizando as multiplicidades e considerando necessariamente o valor próprio infinito.

Isto é, como a matriz A_0 é singular então $\det(P(\lambda))$ tem grau $p < mn$ a que correspondem apenas

p valores próprios finitos. Neste caso, como $\det(A_0) = 0$, a λ -matriz

$$\lambda^m P(1/\lambda) = A_0 + \lambda A_1 + \cdots + \lambda^m A_m, \quad (2.3.9)$$

terá o valor próprio nulo com multiplicidade algébrica $mn - p$.

Como dito na introdução, o valor próprio infinito de $P(\lambda)$ corresponde ao valor próprio nulo da λ -matriz $\lambda^m P(1/\lambda)$, (ver Gohberg [15], pág. 185). Deste ponto de vista, agrupando as diversas cadeias de Jordan relativas ao valor próprio $\lambda = 0$ da λ -matriz (2.3.9), $V_i = \begin{bmatrix} v_{i,1} & \cdots & v_{i,k_i} \end{bmatrix}$, com

$$J_i = \begin{bmatrix} 0 & 1 & 0 \\ & \ddots & 1 \\ 0 & & 0 \end{bmatrix}_{k_i \times k_i},$$

onde $i = 1, \dots, q$ e $\sum_i k_i = mn - p$, então o par próprio,

$$V = \begin{bmatrix} V_1 & \cdots & V_q \end{bmatrix} \text{ e } J = \text{Diag}(J_1, \dots, J_q), \quad (2.3.10)$$

de forma idêntica à usada em (2.3.4), ainda verifica a igualdade

$$A_0 V + A_1 V J + \cdots + A_m V J^m = 0. \quad (2.3.11)$$

Tendo em conta a construção usada em (2.3.4) e (2.3.10) e agrupando os pares próprios associados aos diversos valores próprios finitos em (V_F, J_F) e designando o par próprio associado ao valor próprio infinito por (V_∞, J_∞) , a matriz de semelhança definida em (2.3.7), generaliza-se (ver Gohberg [15], pág. 188 e Cohen [3], pág. 167), para uma matriz com a forma

$$S = \begin{bmatrix} V_F & V_\infty J_\infty^{m-1} \\ V_F J_F & V_\infty J_\infty^{m-2} \\ \vdots & \vdots \\ V_F J_F^{m-1} & V_\infty \end{bmatrix}, \quad (2.3.12)$$

com

$$\lambda \mathcal{C}_2 - \mathcal{C}_1 = S_m K_{(\mathcal{C}_1, \mathcal{C}_2)} S^{-1},$$

onde a matriz S_m está definida por,

$$S_m = \begin{bmatrix} V_F & V_\infty J_\infty^{m-2} \\ \vdots & \vdots \\ V_F J_F^{m-2} & V_\infty \\ A_0 V_F J_F^{m-1} & -\sum_{i=1}^m A_i V_\infty J_\infty^{i-1} \end{bmatrix}.$$

Com efeito,

$$\begin{aligned}
 (\lambda\mathcal{C}_2 - \mathcal{C}_1)S &= \begin{bmatrix} \lambda I_n & -I_n & & 0 \\ 0 & \lambda I_n & -I_n & \\ \vdots & & \ddots & \\ A_m & A_{m-1} & \cdots & \lambda A_0 + A_1 \end{bmatrix} \begin{bmatrix} V_F & V_\infty J_\infty^{m-1} \\ V_F J_F & V_\infty J_\infty^{m-2} \\ \vdots & \vdots \\ V_F J_F^{m-1} & V_\infty \end{bmatrix} \\
 &= \begin{bmatrix} \lambda V_F - V_F J_F & \lambda V_\infty J_\infty^{m-1} - V_\infty J_\infty^{m-2} \\ \lambda V_F J_F - V_F J_F^2 & \lambda V_\infty J_\infty^{m-2} - V_\infty J_\infty^{m-3} \\ \vdots & \vdots \\ \lambda A_0 V_F J_F^{m-1} + \sum_{i=1}^m A_{m-i+1} V_F J_F^{i-1} & \lambda A_0 V_\infty + \sum_{i=1}^m A_i V_\infty J_\infty^{i-1} \end{bmatrix},
 \end{aligned}$$

e

$$\begin{aligned}
 S_m K_{(\mathcal{C}_1, \mathcal{C}_2)} &= \begin{bmatrix} V_F & V_\infty J_\infty^{m-2} \\ V_F J_F & V_\infty J_\infty^{m-3} \\ \vdots & \vdots \\ W_1 & W_2 \end{bmatrix} \begin{bmatrix} \lambda I - J_F & \\ & \lambda J_\infty - I \end{bmatrix} \\
 &= \begin{bmatrix} \lambda V_F - V_F J_F & \lambda V_\infty J_\infty^{m-1} - V_\infty J_\infty^{m-2} \\ \lambda V_F J_F - V_F J_F^2 & \lambda V_\infty J_\infty^{m-2} - V_\infty J_\infty^{m-3} \\ \vdots & \vdots \\ \lambda W_1 - W_1 J_F & \lambda W_2 J_\infty - W_2 \end{bmatrix},
 \end{aligned}$$

onde a igualdade da última linha das matrizes, $(\lambda\mathcal{C}_2 - \mathcal{C}_1)S$ e $S_m K_{(\mathcal{C}_1, \mathcal{C}_2)}$, resulta das equações (2.3.3) e (2.3.11):

- $W_1 = A_0 V_F J_F^{m-1}$ e $-W_1 J_F = -A_0 V_F J_F^m = \sum_{i=1}^m A_{m-i+1} V_F J_F^{i-1}$;
- $W_2 = -\sum_{i=1}^m A_i V_\infty J_\infty^{i-1}$ e $\lambda W_2 J_\infty = -\lambda \sum_{i=1}^m A_i V_\infty J_\infty^i = \lambda A_0 V_\infty$.

Uma vez mais, recorre-se a um exemplo ilustrativo dos conceitos apresentados e onde são calculados os pares próprios, finito e infinito, e identificadas as respetivas matrizes de semelhança.

Exemplo 2.3.10. Seja a λ -matriz,

$$P(\lambda) = \begin{bmatrix} -2 & \lambda^2 + 1 \\ -1 & \lambda \end{bmatrix}.$$

Tem-se que $\det(P(\lambda)) = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2$ pelo que, $P(\lambda)$ admite apenas o valor próprio finito $\lambda = 1$ com multiplicidade 2 e, nestas condições, o valor próprio $\lambda = \infty$ terá multiplicidade 2.

Considerando

$$P(\lambda) = \begin{bmatrix} -2 & \lambda^2 + 1 \\ -1 & \lambda \end{bmatrix}, \quad P'(\lambda) = \begin{bmatrix} 0 & 2\lambda \\ 0 & 1 \end{bmatrix}, \quad \text{e} \quad P''(\lambda) = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix},$$

tem-se:

$$P(1)\vec{v} = \begin{bmatrix} -2 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -2v_1 + 2v_2 \\ -v_1 + v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

para qualquer vetor \vec{v} com $v_2 = v_1$;

$$P(1)\vec{u} + P'(1)\vec{v} = \begin{bmatrix} -2 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} -2u_1 + 2u_2 + 2v_1 \\ -u_1 + u_2 + v_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

implica $u_1 = v_1 + u_2$;

Se

$$\begin{aligned} P(1)\vec{w} + P'(1)\vec{u} + P''(1)\vec{v} &= \begin{bmatrix} -2 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 + u_2 \\ u_2 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\ &= \begin{bmatrix} -2w_1 + 2w_2 + 2v_1 + 2u_2 + v_1 \\ -w_1 + w_2 + v_1 + u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \end{aligned}$$

implicaria $v_1 = 0$, o que não pode suceder.

Portanto, para $\lambda = 1$, tem-se uma cadeia de Jordan com dois vetores próprios generalizados

$$\begin{bmatrix} v_1 & v_1 + u_2 \\ v_1 & u_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}.$$

Assim,

$$V_F = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \quad \text{e} \quad J_F = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Considerando

$$Q(\lambda) = \lambda^2 P(1/\lambda) = \begin{bmatrix} -2\lambda^2 & 1 + \lambda^2 \\ -\lambda^2 & \lambda \end{bmatrix}, \quad Q'(\lambda) = \begin{bmatrix} -4\lambda & 2\lambda \\ -2\lambda & 1 \end{bmatrix}, \quad \text{e} \quad Q''(\lambda) = \begin{bmatrix} -4 & 2 \\ -2 & 0 \end{bmatrix},$$

tem-se que $\det(Q(\lambda)) = \lambda^4 - 2\lambda^3 + \lambda^2 = \lambda^2(\lambda - 1)^2$ pelo que, $Q(\lambda)$ admite o valor próprio nulo, $\lambda = 0$ com multiplicidade 2 que corresponde ao valor próprio infinito, $\lambda = \infty$, de $P(\lambda)$ com a mesma multiplicidade.

Assim,

$$Q(0)\vec{v} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \text{para qualquer vetor } \vec{v} \text{ tal que } v_2 = 0.$$

$$Q(0)\vec{u} + Q'(0)\vec{v} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ 0 \end{bmatrix} = \begin{bmatrix} u_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

implica $u_2 = 0$.

$$\begin{aligned} Q(0)\vec{w} + Q'(0)\vec{u} + Q''(0)\vec{v} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ 0 \end{bmatrix} + \begin{bmatrix} -2 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} w_2 - 2v_1 \\ -v_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \end{aligned}$$

implicaria $v_1 = 0$, o que não pode suceder.

Portanto, para $\lambda = \infty$, tem-se uma cadeia de Jordan com dois vetores próprios generalizados

$$\begin{bmatrix} v_1 & u_1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

Assim,

$$V_\infty = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \text{ e } J_\infty = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Tem-se que

$$S = \begin{bmatrix} V_F & V_\infty J_\infty \\ V_F J_F & V_\infty \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 3 & 1 & 1 \\ 1 & 2 & 0 & 0 \end{bmatrix}, \text{ com } S^{-1} = \begin{bmatrix} 0 & 2 & 0 & -1 \\ 0 & -1 & 0 & 1 \\ -1 & 1 & 1 & -1 \\ 1 & 0 & 0 & -1 \end{bmatrix},$$

e

$$S_m = \begin{bmatrix} V_F & V_\infty \\ A_0 V_F J_F & -\sum_{i=1}^2 A_i V_\infty J_\infty^{i-1} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

$$\begin{aligned} S_m K_{(\mathcal{C}_1, \mathcal{C}_2)} S^{-1} &= \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda - 1 & -1 & 0 & 0 \\ 0 & \lambda - 1 & 0 & 0 \\ 0 & 0 & -1 & \lambda \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 2 & 0 & -1 \\ 0 & -1 & 0 & 1 \\ -1 & 1 & 1 & -1 \\ 1 & 0 & 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} \lambda & 0 & -1 & 0 \\ 0 & \lambda & 0 & -1 \\ -2 & 1 & 0 & \lambda \\ -1 & 0 & 0 & 1 \end{bmatrix} = \lambda \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 2 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix} \\ &= \lambda \mathcal{C}_2 - \mathcal{C}_1. \end{aligned}$$

Observação 2.3.11. Segundo Dooren [7], pág. 552, da mesma forma que a multiplicidade algébrica de um valor próprio finito λ_0 está relacionada com o número de soluções linearmente independentes associadas à função $e^{\lambda_0 t}$ a multiplicidade algébrica do valor próprio infinito $\lambda = \infty$ está relacionada com o número de soluções associadas à função delta de Dirac, $\delta(t)$. Com efeito, considerando a forma canónica de Weierstrass do exemplo anterior,

$$K_{(\mathcal{C}_1, \mathcal{C}_2)} = \begin{bmatrix} \lambda - 1 & -1 & 0 & 0 \\ 0 & \lambda - 1 & 0 & 0 \\ 0 & 0 & -1 & \lambda \\ 0 & 0 & 0 & -1 \end{bmatrix},$$

o sistema de equações diferenciais associado (tal como referido na introdução), com $\lambda = \frac{d}{dt}$ e $Y(t)$ o vetor incógnita tal que $Y(0) = \begin{bmatrix} y_1(0) & y_2(0) & y_3(0) & y_4(0) \end{bmatrix}^T$, é definido por

$$\begin{bmatrix} \frac{d}{dt} - 1 & -1 & 0 & 0 \\ 0 & \frac{d}{dt} - 1 & 0 & 0 \\ 0 & 0 & -1 & \frac{d}{dt} \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \\ y_4(t) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (2.3.13)$$

O sistema (2.3.13) pode decompor-se em dois sistemas independentes

$$\begin{cases} y_1'(t) - y_1(t) - y_2(t) = 0 \\ y_2'(t) - y_2(t) = 0 \end{cases} \quad \text{e} \quad \begin{cases} -y_3(t) + y_4'(t) = 0 \\ -y_4(t) = 0 \end{cases},$$

e, relativamente a este último (aquele que está relacionado com o valor próprio infinito), obtém-se a seguinte solução:

$$\begin{cases} y_3(t) = -\delta(t)y_4(0) \\ y_4(t) = 0 \end{cases}.$$

Evidentemente que se a multiplicidade geométrica do valor próprio infinito $\lambda = \infty$ for um a correspondente solução é a função nula (tal como sucede com a coordenada $y_4(t)$).

2.4 Solventes do polinómio matricial $P(X)$

Dado um polinómio,

$$P(X) = X^m + A_1X^{m-1} + \dots + A_m, \quad (2.4.1)$$

a relação existente entre solventes do polinómio (2.4.1) e cadeias de Jordan da λ -matriz (2.3.1) e, mais geralmente, os blocos vetores próprios (como será visto no seguimento da Definição 2.4.8), é estabelecida pelo seguinte resultado:

Teorema 2.4.1. (Lancaster [25], pág. 521 e Dennis [9], pág. 79) X é um solvente de $P(X)$ se e só se

$$CV = VX, \quad (2.4.2)$$

onde C é a matriz companheira de $P(X)$ e V representa a matriz coluna de blocos

$$V = \begin{bmatrix} I_n \\ X \\ \vdots \\ X^{m-1} \end{bmatrix}. \quad (2.4.3)$$

Observação 2.4.2. Note-se que a matriz V definida em (2.4.3) tem necessariamente característica máxima n uma vez que o primeiro bloco de V é a matriz identidade I_n .

Definição 2.4.3. (Pereira [30], pág. 13) Se $X_1, \dots, X_m \in \mathbb{C}^{n \times n}$, o bloco de Vandermonde é a matriz de ordem mn

$$\mathcal{V}(X_1, \dots, X_m) = \begin{bmatrix} I_n & I_n & \dots & I_n \\ X_1 & X_2 & \dots & X_m \\ \vdots & \vdots & \ddots & \vdots \\ X_1^{m-1} & X_2^{m-1} & \dots & X_m^{m-1} \end{bmatrix}. \quad (2.4.4)$$

Teorema 2.4.4. (Lancaster [25], pag. 524) Sejam $X_1, X_2, \dots, X_m \in \mathbb{C}^{n \times n}$, m solventes do polinômio matricial $P(X)$, da forma (2.4.1). Se o bloco de Vandermonde $\mathcal{V} = \mathcal{V}(X_1, X_2, \dots, X_m)$ é regular então

$$\mathcal{C} = \mathcal{V} \text{Diag}(X_1, X_2, \dots, X_m) \mathcal{V}^{-1}, \quad (2.4.5)$$

e diz-se que os m solventes formam um conjunto completo de solventes de $P(X)$.

Portanto, nas condições do Teorema 2.4.4, os valores próprios dos solventes percorrem todo o espectro $\Lambda(P)$, isto é, a união dos espectros dos solventes X_1, X_2, \dots, X_m é igual ao espectro da matriz companheira. Esta condição é necessária mas não suficiente para se constituir um conjunto completo de solventes de $P(X)$.

Observação 2.4.5. Na realidade, para se construir um conjunto completo é suficiente (e necessário, ver Pereira [30], pág. 111) que os valores próprios e respectivas multiplicidades parciais dos m solventes X_1, X_2, \dots, X_m e da matriz companheira coincidam.

Dos m solventes distingue-se aquele que possui os valores próprios de maior módulo, com diversas aplicações, nomeadamente no método de Bernoulli, Algoritmo A.1.2, e método de Traub, Algoritmo A.1.3.

Definição 2.4.6. (Dennis [11], pág.524) Dadas m solventes X_1, X_2, \dots, X_m do polinômio matricial $P(X)$, diz-se que X_j é o solvente dominante se os seus valores próprios são maiores, em módulo, do que os valores próprios de todos os outros solventes.

Observação 2.4.7. Se for considerado um conjunto completo de solventes X_1, X_2, \dots, X_m então, segundo a Observação 2.4.5, os valores próprios e respectivas multiplicidades parciais dos m solventes e da matriz companheira coincidem, tal como no caso de ser considerado um sistema completo de pares próprios $(V_1, J_1), \dots, (V_l, J_l)$ (ver Definição 2.3.5). Cada par próprio (Definição 2.3.3, secção 2.3) verificam a equação $\mathcal{C}S_i = S_i J_i, \forall i = 1, \dots, l$, atendendo a (2.3.6) e considerando

$$S_i = \begin{bmatrix} V_i \\ V_i J_i \\ \vdots \\ V_i J_i^{m-1} \end{bmatrix},$$

mas não têm todos necessariamente a mesma dimensão (ver Exemplos 2.1.12, 2.3.8 e 2.3.10) pois dependem das multiplicidades algébrica e geométrica do valor próprio associado. Pelo contrário, segundo o Teorema 2.4.1, os solventes do polinômio X_1, X_2, \dots, X_m verificam ainda a equação $\mathcal{C}V = VX$, com V dado por (2.4.3), e têm uma dimensão fixa n .

Com base na Observação anterior (2.4.7), pode optar-se em contrapartida pela seguinte formulação mais genérica.

Definição 2.4.8. (Dennis [9], pág. 72) Seja A uma matriz quadrada, $A \in \mathbb{C}^{m \times m}$. Uma matriz quadrada $X \in \mathbb{C}^{n \times n}$, com $n < m$, diz-se um bloco valor próprio de A se existir um bloco vetor $V \in \mathbb{C}^{m \times n}$ de característica máxima tal que

$$AV = VX. \quad (2.4.6)$$

Nestas condições V diz-se um bloco vetor próprio de A associado a X .

Observação 2.4.9. De acordo com a Definição 2.4.8, o Teorema 2.4.1 corresponde a afirmar que X é um solvente de $P(X)$ se e só se for um bloco valor próprio da matriz companheira $CV = VX$, onde o bloco vetor próprio V é a matriz coluna dada por (2.4.3).

Por outro lado (ver Dennis [9], pág. 85), dado X um bloco valor próprio da matriz companheira $CV = VX$ de dimensão $n \times n$, se o bloco $V_1 \in \mathbb{C}^{n \times n}$ de $V \in \mathbb{C}^{m \times n}$, for não singular, então

$$Y = V_1 X (V_1)^{-1}$$

é um solvente do polinômio $P(X)$.

Se X é um bloco valor próprio da matriz companheira $CV = VX$, com V o seu bloco vetor próprio associado, então todos os valores próprios de X são ainda valores próprios de C e, por sua vez, da λ -matriz $P(\lambda)$.

Ainda em Dennis [9], pág. 74, estão definidos alguns resultados relativos ao conceito de blocos valores próprios.

Ao contrário de solventes, que poderão não existir (nem sempre se poderá contar com a hipótese de o bloco V_1 , de um bloco vetor próprio V , ser não singular) a matriz C admitirá sempre um bloco valor próprio. Por exemplo, seja $V = [v_1, \dots, v_k]$ uma cadeia de Jordan associada ao valor próprio λ_i de C com J um bloco de Jordan de dimensão $k \times k$ e λ_i na diagonal principal verifica, conforme o Teorema 2.3.2,

$$VJ^m + A_1 VJ^{m-1} + \dots + A_m V = 0.$$

Assim, se a matriz

$$S' = \begin{bmatrix} V \\ VJ \\ \vdots \\ VJ^{m-1} \end{bmatrix},$$

tiver característica máxima, tem-se $CS' = S'J$, e portanto J é um bloco valor próprio de C com S' o bloco vetor próprio associado.

A Definição seguinte é uma adaptação do conceito de sistema completo de pares próprios, Definição 2.3.5 para blocos valores próprios com uma dimensão fixa n (ver Teorema 3.3.4 em Pereira [30], pág. 113).

Definição 2.4.10. Um conjunto de blocos valores próprios X_1, \dots, X_m de uma matriz particionada $A \in \mathbb{C}^{mn \times mn}$ diz-se conjunto completo de blocos valores próprios se existir um conjunto de blocos vetores próprios associados V_1, \dots, V_m tal que a matriz $\begin{bmatrix} V_1 & \dots & V_m \end{bmatrix}$ tem caraterística máxima e

$$A \begin{bmatrix} V_1 & \dots & V_m \end{bmatrix} = \begin{bmatrix} V_1 & \dots & V_m \end{bmatrix} \text{Diag}(X_1, \dots, X_m), \quad (2.4.7)$$

onde a matriz $\text{Diag}(X_1, \dots, X_m)$ tem a mesma dimensão que A .

Poderá mesmo suceder que um dado polinómio $P(X)$ não tenha solventes mas que, por outro lado, a respetiva matriz companheira \mathcal{C} admita um conjunto completo de blocos valores próprios de ordem n (ver exemplo apresentado em Pereira [30], pág. 121).

Capítulo 3

Cálculo de Solventes do Polinômio matricial $P(X)$

Neste capítulo será apresentada e desenvolvida parte do trabalho original desta dissertação.

Como referido, a caracterização da matriz companheira pode ser feita ao utilizarem-se valores e vetores próprios da λ -Matriz $P(\lambda)$, Teorema 2.3.6, ou através dos solventes do polinômio matricial $P(X)$, Teorema 2.4.4. Por outro lado, se existirem, um conjunto completo de blocos valores próprios ou um conjunto completo de solventes, pode ser obtido pelo método da potência associado a uma deflação (Block Wielandt Deflation ou Block Hotelling Deflation (ver Pereira [34], pág. 1179), ou quando se trate do feixe companheiro (ver Pereira [32], pág. 2915)).

Assim, o trabalho exposto neste capítulo consiste em definir métodos alternativos de cálculo de solventes. Esses métodos, aqui desenvolvidos, têm na sua génese a transformação da equação polinomial matricial original numa equação vetorial equivalente, através dos resultados relativos ao produto de Kronecker (ver Hernandez [14], pág. 3 do apêndice), evitando-se a álgebra matricial e beneficiando da possibilidade de aplicação dos métodos clássicos à nova equação vetorial.

Deste ponto de vista, começa-se por fazer o enquadramento referindo alguns métodos iterativos conhecidos que permitem o cálculo de solventes. Este enquadramento permite, em simultâneo, ter um termo de comparação para os novos métodos agora apresentados. No decurso deste capítulo são apresentados diversos exemplos onde se comparam as prestações e eficiências dos métodos aqui desenvolvidos com o método Newton, que corresponde à adaptação do método clássico ao contexto dos polinômios matriciais, (ver Kratz [21], pág. 359, Higham [18], pág. 4, Pereira [33], pág. 3), com o método de Bernoulli, versão matricial (ver Lancaster [22], Tsai [38], Higham [19], pág. 508), que é essencialmente o método da Potência aplicado à matriz companheira \mathcal{C} e também com o método de Traub, assim designado por ser uma generalização do método de Traub escalar (ver Traub [37], pág. 138) e que permite obter-se uma aproximação do solvente dominante de $P(X)$.

3.1 Métodos Iterativos

Inicia-se o enquadramento referido, começando-se por descrever de forma sintética a essência do método de Newton antes citado. Para esse efeito, é fundamental considerar a noção de derivada adaptada ao contexto dos polinômios matriciais que é expressa na seguinte Definição.

Definição 3.1.1. (Kratz [21]) Dado um polinômio matricial de grau m ,

$$P(X) = A_0X^m + A_1X^{m-1} + \dots + A_m,$$

a derivada de Fréchet de $P(X)$, no ponto X aplicada a $Q \in \mathbb{C}^{n \times n}$, está definida por

$$P'(X)[H] = \sum_{1 \leq i \leq m} \left(\sum_{1 \leq j \leq i} A_{m-i} X^{i-j} H X^{j-1} \right). \quad (3.1.1)$$

Assim, a adaptação do método Newton clássico ao contexto dos polinômios matriciais (ver Kratz [21], pág. 359, Higham [18], pág. 4, Pereira [33], pág. 3) resulta da expansão de $P(X)$ em X , uma aproximação do solvente de $P(X)$, na direção de Q ,

$$P(X + Q) = P(X) + P'(X)[Q] + \frac{1}{2}D_{P'(X)[Q]}(X)[Q] + \cdots,$$

desprezando os termos de ordem maior ou igual a 2 em Q .

Dada uma aproximação inicial $X^{(0)}$ do solvente de $P(X)$, para cada $k = 1, 2, \dots$, a aproximação seguinte é calculada resolvendo a equação

$$P'(X^{(k)})[Q^{(k)}] = -P(X^{(k)}),$$

em ordem a $Q^{(k)}$ e considerando $X^{(k+1)} = X^{(k)} + Q^{(k)}$.

Em cada iteração a matriz $Q^{(k)}$ pode ser estimada por diversos processos, quer calculando a inversa de $P'(X^{(k)})$, se esta for regular, quer usando a inversa generalizada de $P'(X^{(k)})$, em cada ponto $X^{(k)}$ ou no ponto inicial $X^{(0)}$, dependendo das condições, sobretudo a dimensão do próprio polinômio matricial. Em todo o caso, por exemplo para $m = 2$, o método de Newton (ver Anexo A, ponto 6 do Algoritmo A.1.1), implica a resolução da respetiva equação matricial de Silvester

$$(X^{(k)} + A_1)Q^{(k)} + Q^{(k)}X^{(k)} = -P(X^{(k)}), \quad (3.1.2)$$

em cada iteração k .

Quando as matrizes X e Q comutam, então a derivada de Fréchet $P'(X)[Q]$ corresponde à derivada usual de P , isto é, $P'(X)[Q] = P'(X)Q = \sum_{1 \leq l \leq m} l A_{m-l} X^{l-1} Q$.

O método de Bernoulli, (ver Dennis [9] e [11], pág. 69 e 529), permite calcular uma sucessão de matrizes, $X^{(0)}, X^{(1)}, \dots, X^{(k)}, \dots$ a partir de matrizes iniciais $X^{(0)} = X^{(1)} = \dots = X^{(m-2)} = 0$ e $X^{(m-1)} = I_n$, onde a matriz $X^{(k+1)}$ fica definida por $X^{(k+1)} = -A_1 X^{(k)} - \dots - A_m X^{(k-m+1)}$.

Isto é, o método pode ser definido através da equação

$$\begin{bmatrix} X^{(k-m+2)} \\ \vdots \\ X^{(k)} \\ X^{(k+1)} \end{bmatrix} = \begin{bmatrix} 0_n & I_n & \dots & 0_n \\ \vdots & 0_n & \ddots & \vdots \\ 0_n & \dots & \ddots & I_n \\ -A_m & -A_{m-1} & \dots & -A_1 \end{bmatrix} \begin{bmatrix} X^{(k-m+1)} \\ \vdots \\ X^{(k-1)} \\ X^{(k)} \end{bmatrix},$$

que, como foi dito, corresponde ao método da Potência aplicado à matriz companheira \mathcal{C} (ver Lancaster [22], Tsai [38], Higham [19], pág. 508).

Se $P(X)$ tem um conjunto completo de solventes X_1, X_2, \dots, X_m , com X_1 um solvente dominante, Definição 2.4.6, e a matriz bloco de Vandermonde $\mathcal{V}(X_2, \dots, X_m)$ é não singular, então o produto $X^{(k)}(X^{(k-1)})^{-1}$, onde $X^{(k-1)}, X^{(k)}$ são dois termos consecutivos da sucessão obtida através do respetivo Algoritmo (ver Algoritmo A.1.2, Anexo A) converge para o solvente dominante X_1 .

Isto é,

$$\lim_{k \rightarrow \infty} X^{(k)}(X^{(k-1)})^{-1} = X_1.$$

O método de Traub permite calcular uma sucessão de matrizes, $X^{(0)}, X^{(1)}, \dots, X^{(k)}, \dots$, com $X^{(k+1)} = \varphi_l(X^{(k)})$, para $k = 0, 1, 2, \dots$, onde as funções φ_l são construídas com base nos Teoremas da divisão (ver Dennis [10], pág. 832) e interpolação (ver Dennis [9], pág. 29) da álgebra dos polinômios matriciais.

Com efeito, (ver Dennis [11], pág. 524) é construída uma sucessão de polinômios matriciais

$$G_0(X), G_1(X), \dots, G_j(X), \dots,$$

com $G_0(X) = I_n$ e $G_{j+1}(X) = XG_j(X) - \Gamma_1^j P(X)$, onde Γ_1^j é a matriz coeficiente da potência X^{m-1} do polinômio $G_j(X)$, isto é, $G_j(X) = \sum_{i=1, \dots, m} \Gamma_i^j X^{m-i}$. Numa primeira fase, o método de Traub procura uma aproximação (em l) dentro do espectro de convergência das iterações do ponto fixo, e só depois, calcula iterações em k com vista à obtenção de uma aproximação do solvente dominante X_1 .

Nas condições de convergência (condições idênticas às do método de Bernoulli, $P(X)$ tem um conjunto completo de solventes X_1, X_2, \dots, X_m , onde X_1 é o solvente dominante e $\mathcal{V}(X_2, \dots, X_m)$ é não singular), as funções $\varphi_l(X) = G_l(X)G_{l-1}^{-1}(X)$ estão bem definidas e, para um valor de l suficientemente grande, $l = 1, 2, \dots$, a sucessão $X^{(k+1)} = \varphi_l(X^{(k)})$ obtida através do respetivo Algoritmo (ver Algoritmo A.1.3, Anexo A), converge para o solvente dominante X_1 .

Observação 3.1.2. Quando se considera $G_0(X) = I_n = 0X^{m-1} + 0X^{m-2} + \dots + I_n X^0$, isto é, $\Gamma_1^0 = \Gamma_2^0 = \dots = \Gamma_{m-1}^0 = 0$ e $\Gamma_m^0 = I_n$, então

$$\begin{aligned} G_1(X) &= XG_0(X) - \Gamma_1^0 P(X) \\ &= XI_n - 0P(X) \\ &= X \\ &= 0X^{m-1} + 0X^{m-2} + \dots + I_n X^1 + 0X^0, \end{aligned}$$

ou seja, o coeficiente Γ_{m-1}^1 do polinômio $G_1(X)$ é a matriz identidade, I_n , e os restantes coeficientes são nulos, $\Gamma_1^1 = \Gamma_2^1 = \dots = \Gamma_{m-2}^1 = 0$ e $\Gamma_m^1 = 0$.

Repetindo este processo para $l = 1, 2, 3, \dots, m-1$, tem-se

$$\begin{aligned} G_2(X) &= XG_1(X) - \Gamma_1^1 P(X) = XX - 0P(X) = X^2 \\ &= 0X^{m-1} + 0X^{m-2} + \dots + 0X^3 + I_n X^2 + 0X^1 + 0X^0 \\ G_3(X) &= XG_2(X) - \Gamma_1^2 P(X) = XX^2 - 0P(X) = X^3 \\ &= 0X^{m-1} + 0X^{m-2} + \dots + I_n X^3 + 0X^2 + 0X^1 + 0X^0 \\ &\vdots \\ G_{m-1}(X) &= XG_{m-2}(X) - \Gamma_1^{m-2} P(X) = XX^{m-2} - 0P(X) = X^{m-1} \\ &= I_n X^{m-1} + 0X^{m-2} + \dots + 0X^3 + 0X^2 + 0X^1 + 0X^0, \end{aligned}$$

e

$$G_m(X) = XG_{m-1}(X) - \Gamma_1^{m-1} P(X) = XX^{m-1} - I_n P(X),$$

isto é,

$$P(X) = XX^{m-1} - G_m(X). \quad (3.1.3)$$

Resolvendo equação $XX^{m-1} - G_m(X) = 0$ em ordem ao primeiro fator X , através da inversa de X^{m-1} , obtêm-se a fórmula do método de Traub,

$$X = \underbrace{G_m(X)(X^{m-1})^{-1}}_{\varphi_m(X)}, \quad (3.1.4)$$

onde,

$$\varphi_m(X_1) = X_1 \quad (3.1.5)$$

e

$$\|\varphi_m(X) - X_1\| \leq c\sigma^{m-1}\|X - X_1\|,$$

com $\sigma < 1$ e X suficientemente próximo do solvente dominante X_1 (ver Dennis [11], pág. 528).

De fato, atendendo à equação (3.1.5), o método de Traub é um método do tipo ponto fixo em $\mathbb{C}^{n \times n}$.

3.2 Método do Ponto Fixo

Com o objetivo de apresentar um método do Ponto Fixo, no qual não seja necessário o cálculo da inversa da matriz X^{m-1} a cada passo como no método de Traub são desenvolvidas nesta secção os conceitos básicos que irão permitir a construção do algoritmo. Assim, com base na Observação 3.1.2, tenta-se agora resolver a equação $X^m - G_m(X) = 0$ em ordem a X , não através da inversa de X^{m-1} , mas através de “ $X = [G_m(X)]^{\frac{1}{m}}$ ” em termos das entradas do polinômio.

O polinômio matricial $P(X)$ pode ser escrito como uma função,

$$\begin{array}{ccc} P : \mathbb{C}^{n \times n} & \longrightarrow & \mathbb{C}^{n \times n} \\ X & \longrightarrow & P(X) \end{array} . \quad (3.2.1)$$

Assim, as entradas do polinômio matricial $P(X)$,

$$p_{i,j} := p(x_{1,1}, x_{1,2}, \dots, x_{i,j}, \dots, x_{n,n})_{i,j}, \quad (3.2.2)$$

com $i, j = 1, 2, \dots, n$, são funções analíticas nas variáveis $x_{i,j}$. Trata-se de resolver cada função $p_{i,j}$ em ordem a $x_{i,j}$.

Considerando $P(X)$ como a função definida em (3.2.1), pode compor-se esta com a função *vec*, ver Definição 2.1.14, de modo a obter-se

$$\begin{array}{ccc} f : U \subseteq \mathbb{C}^{n^2} & \longrightarrow & \mathbb{C}^{n^2} \\ x = \text{vec}(X) & \longrightarrow & f(x) \end{array} , \quad (3.2.3)$$

onde as coordenadas $f_l(x)$, $l = 1, 2, \dots, n^2$, são obtidas resolvendo o respetivo polinômio $p_{i,j}$, com $l = (j-1)n + i$, em ordem à maior potência da variável $x_{i,j}$.

Observação 3.2.1. Intuitivamente, ao considerar-se $f_l(x) = (x_{i,j}^m - p_{i,j})^{\frac{1}{m}}$, as derivadas parciais serão definidas por,

$$\frac{\partial f_l(x)}{\partial x_{i,j}} = \frac{1}{m} \frac{\frac{\partial}{\partial x_{i,j}}(x_{i,j}^m - p_{i,j})}{[f_l(x)]^{m-1}},$$

e, portanto, quanto maior for o valor de m considerado (e o valor de $|f_l(x)| > 1$), menor será o valor do raio espectral da matriz Jacobiana que, sendo menor do que 1, garante a convergência do método.

Pode verificar-se que a maior potência da variável $x_{i,j}$ depende do grau do polinômio m e dos índices i e j . Assim, se $i = j$ a maior potência é m , e, caso contrário, a maior potência é m_1 , onde

$$m_1 = \begin{cases} \frac{m}{2} & , \text{ se } m \text{ par} \\ \frac{m+1}{2} & , \text{ se } m \text{ ímpar} \end{cases} . \quad (3.2.4)$$

Uma vez que o elemento $x_{i,j}$ da matriz X é a coordenada $(j-1)n+i$ do vetor $vec(X)$, para cada $i, j = 1, \dots, n$, então a função f pode ser definida por

$$f_{(j-1)n+i}(x) = \begin{cases} (x_{i,j}^m - p_{i,j})^{\frac{1}{m}} & , \text{ if } i = j \\ \left(\frac{x_{i,j}^{m_1} \cdot q_{i,j} - p_{i,j}}{q_{i,j}} \right)^{\frac{1}{m_1}} & , \text{ if } i \neq j \end{cases} , \quad (3.2.5)$$

onde, se m for ímpar,

$$q_{i,j} = x_{j,i}^{m_1-1} \left((m_1 \cdot (x_{i,i} + x_{j,j}) + (a_1)_{i,i}) + (m_1 - 1) x_{j,i}^{m_1-2} \sum_{1 \leq k \leq n} (x_{j,l} x_{l,i} - x_{j,i} x_{l,i} - x_{j,j} x_{j,i}) \right),$$

e, se m é par,

$$q_{i,j} = x_{j,i}^{m-m_1}.$$

Como consequência direta desta construção tem-se:

Observação 3.2.2. Se x é um ponto fixo de f , então $X = vec^{-1}(x)$ é um solvente de $P(X)$.

Por construção: Se x é um ponto fixo de f , $x_l = f_l(x) \Rightarrow x_l^{m_1} = f_l(x)^{m_1}$, $l = 1, \dots, n^2$. Então $x_l^{m_1} - f_l(x)^{m_1} = 0$ e, considerando (3.2.2),

$$p_{i,j} = (x_1, x_2, \dots, x_{n^2}) = 0, \quad (3.2.6)$$

para cada $i, j = 1, \dots, n$.

Isto é, para $X = vec^{-1}(x)$, $vec(P(X)) = 0$ e $P(X) = 0$.

De uma forma geral, o inverso não se verifica dependendo da determinação considerada para a função $\sqrt[m]{\square}$. Por exemplo, para $m = 2$ se X for um solvente de $P(X)$, para $i = j$, as coordenadas de $vec(X)$ podem ser $\pm f_{(i-1)n+i}(s)$.

3.2.3 Algoritmo

O método pode ser aplicado seguindo os seguintes passos:

Algoritmo 3.2.1 PF

- 1: Dada uma aproximação inicial, $x^{(0)} = vec(X^{(0)})$, e uma margem de erro $\varepsilon > 0$.
 - 2: $x^{(1)} = f(x^{(0)})$.
 - 3: para $k = 1, \dots$, fazer
 - 4: se $\|x^{(k)} - x^{(k-1)}\|_2 < \varepsilon$ então
 - 5: o método para;
 - 6: caso contrário
 - 7: $x^{(k+1)} = f(x^{(k)})$
 - 8: $X = vec^{-1}(x^{(k)})$ é uma aproximação do solvente.
-

3.2.4 Convergência

A convergência do Algoritmo 3.2.1 baseia-se no Teorema do ponto fixo de Schauder e na respectiva estabilidade assintótica (Teorema 3.2.5) também conhecido pela conjectura de Belitskiĭ e Lyubich (ver Shih [35], pág. 144), resultados apresentados de seguida.

Teorema 3.2.5. (Shih [35], pág. 144) Seja \mathcal{X} um espaço de Banach, Ω um subconjunto aberto de \mathcal{X} e $F : \Omega \subseteq \mathcal{X} \rightarrow \mathcal{X}$ uma função contínua, compacta e com derivada de Fréchet em Ω . Se $D \subset \Omega$ for um subconjunto aberto, convexo e não vazio tal que $F(\overline{D}) \subset \overline{D}$ and $\sup_{x \in \overline{D}} \rho(F'(x)) = r < 1$. então existe um único ponto fixo de F em \overline{D} , isto é $x \in \overline{D}$ tal que $F(x) = x$.

O seguinte Teorema é uma adaptação do resultado anterior e estabelece as condições necessárias para assegurar a convergência do Algoritmo 3.2.1 (PF). Mediante a exigência de a função f ser de classe C^1 numa vizinhança de um ponto fixo s , $V_\varepsilon(s) = \{x \in \mathbb{C}^{n^2} : \|x - s\| < \varepsilon\}$, e de a norma da respectiva matriz jacobiana verificar $\|J_f(s)\|_\infty = r < 1$ — portanto o seu raio espectral $\rho(J_f(s))$, sendo o ínfimo das normas $\|J_f(s)\|_p$ em \mathbb{C}^{n^2} , também é menor do que 1 — assegura-se que a função é uma contração em $V_\varepsilon(s)$ ($\|f(x_1) - f(x_2)\|_\infty < r\|x_1 - x_2\|_\infty$, com $r < 1$, para $x_1, x_2 \in V_\varepsilon(s)$) e que portanto a sucessão $x^{(k+1)} = f(x^{(k)})$ converge para s independentemente da aproximação inicial considerada $x^{(0)} \in V_\varepsilon(s)$.

Teorema 3.2.6. (Shih [35], pág. 144) Seja f a função definida em (3.2.5). Se i) existe um ponto s , tal que $s = f(s)$ e $V_\varepsilon(s) = \{x \in \mathbb{C}^{n^2} : \|x - s\| < \varepsilon\}$, com

$$f : V_\varepsilon(s) \subseteq \mathbb{C}^{n^2} \rightarrow \mathbb{C}^{n^2}$$

contínua e diferenciável em $V_\varepsilon(s)$,

ii) $\sup_{x \in V_\varepsilon(s)} \|J_f(x)\|_\infty = r < 1$.

Então, para qualquer aproximação inicial $x^{(0)} \in V_\varepsilon(s)$ a sucessão $x^{(k+1)} = f(x^{(k)}) \in V_\varepsilon(s)$ é convergente e converge para s , a única solução de $f(x) = x$ em $V_\varepsilon(s)$.

Demonstração. Usando a alínea i) tem-se que a função f é contínua e diferenciável com $f'(x) = J_f(x)$ em $V_\varepsilon(s)$. Por outro lado,

$$\sup_{x \in V_\varepsilon(s)} \|J_f(x)\|_\infty < 1 \Rightarrow \sup_{x \in V_\varepsilon(s)} \rho(f'(x)) < 1.$$

Ao considerar $x \in \overline{V_\varepsilon(s)}$, onde $\overline{V_\varepsilon(s)}$ um subconjunto convexo não vazio e compacto de \mathbb{C}^{n^2} , então $\|f(x) - s\|_\infty < r\|x - s\|_\infty \leq \varepsilon$, ou seja, $f(x) \in \overline{V_\varepsilon(s)}$.

Assim, $f(\overline{V_\varepsilon(s)}) \subset \overline{V_\varepsilon(s)}$ pelo que pode considerar-se $f : \overline{V_\varepsilon(s)} \rightarrow \overline{V_\varepsilon(s)}$ onde $\overline{V_\varepsilon(s)}$ é um subconjunto convexo não vazio e compacto do espaço de Banach \mathbb{C}^{n^2} .

Então, pelo Teorema do ponto fixo de Schauder (Teorema 3.2.5), existe um único $\bar{x} \in \overline{V_\varepsilon(s)}$ tal que $f(\bar{x}) = \bar{x}$. Logo $\bar{x} = s$.

Usando a alínea ii), então para qualquer $x^{(0)} \in V_\varepsilon(s)$, a sucessão $x^{(k+1)} = f(x^{(k)})$ pertence a $V_\varepsilon(s)$ e

$$\|x^{(k)} - s\|_\infty < \|f(x^{(k-1)}) - f(s)\|_\infty < r\|x^{(k-1)} - s\|_\infty < r^k\|x^{(0)} - s\|_\infty < r^k\varepsilon \xrightarrow[k \rightarrow +\infty]{} 0.$$

Portanto $x^{(k)} \xrightarrow[k \rightarrow +\infty]{} s$. ■

O Teorema anterior assegura a convergência da sucessão $x^{(k)}$, gerada pelos pontos de 2 a 7 do Algoritmo 3.2.1, para s um ponto fixo de f . Este resultado aliado à Observação 3.2.2 assegura que

no ponto 8 do mesmo Algoritmo se obtém realmente uma aproximação de um solvente de $P(X)$, isto é:

Corolário 3.2.7. A sucessão $vec^{-1}(x^{(k)})$ converge para um solvente de $P(X)$.

3.2.8 Exemplos

Seguidamente, são apresentados exemplos onde se compara o comportamento deste método, Algoritmo 3.2.1, com o métodos de Traub, Algoritmo A.1.3, e de Newton, Algoritmo A.1.1. No primeiro exemplo é considerado um polinômio matricial para o qual o método do Ponto Fixo (PF) converge e o método de Newton não converge. No segundo exemplo, é considerado um polinômio matricial que, como será visto, não se encontra nas condições de convergência do método de Traub, mantendo-se a convergência do método PF.

No exemplo seguinte calcula-se o raio espectral da matriz Jacobiana avaliada no ponto fixo s , $\rho(J_f(s))$, em vez de $\|J_f(s)\|_\infty$. Na realidade, por uma questão de simplicidade, na condição presente na alínea ii) do Teorema 3.2.6 considerou-se a norma infinito mas essa condição poderia ser substituída por $\sup_{x \in V_\varepsilon(s)} \rho(J_f(x)) = r < 1$.

Neste caso ficaria ainda assegurada a convergência do método PF para alguma norma matricial $\|\cdot\|_p$ tal que

$$\sup_{x \in V_\varepsilon(s)} \|J_f(x)\|_p < r + \delta < 1.$$

Exemplo 3.2.9. Seja $P(X)$ o polinômio matricial

$$P(X) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} X^2 + \begin{bmatrix} -0.15 & -0.075 & -0.045 \\ 0.01 & -0.355 & 0.085 \\ -0.05 & -0.085 & 0.125 \end{bmatrix} X + \begin{bmatrix} 6.1333 & -9.46667 & 5.3333 \\ -2.7333 & 33.0333 & -11.8333 \\ 3.5333 & 10.5667 & 1.8333 \end{bmatrix}.$$

Considerando

$$X^{(0)} = 48.05.I_3, \text{ com } \rho(J_f(x^{(0)})) = 0.0724,$$

após 14 iterações do método 3.2.1 PF, obtém-se a aproximação para um solvente de $P(X)$,

$$X^{(14)} = \begin{bmatrix} 0.07677 + 2.2711i & 0.074741 - 1.30772i & 0.00881 + 0.81865i \\ -0.005145 - 0.16994i & 0.1774 + 5.9222i & -0.04302 - 1.50722i \\ 0.02435 + 0.92074i & 0.02937 + 1.51598i & 0.06103 + 1.8335i \end{bmatrix},$$

com $\|P(X^{(14)})\|_\infty = 0.00328$ e $\rho(J_f(x^{(14)})) = 0.409693$.

Relativamente à aproximação inicial

$$X^{(0)} = A_2 = \begin{bmatrix} 6.1333 & -9.46667 & 5.3333 \\ -2.7333 & 33.0333 & -11.8333 \\ 3.5333 & 10.5667 & 1.8333 \end{bmatrix},$$

esta não verifica o critério de convergência pois $\rho(J_f(x^{(0)})) = 6.531 > 1$, no entanto, o método 3.2.1 PF converge em 22 iterações para

$$X^{(22)} = \begin{bmatrix} 0.0749956 + 2.27125i & -0.0378964 - 1.30585i & 0.0227793 + 0.818851i \\ -0.0051091 - 0.164025i & 0.177403 + 5.92228i & -0.0424566 - 1.50716i \\ 0.0249551 + 0.920662i & 0.0425912 + 1.51625i & 0.0623993 + 1.83337i \end{bmatrix}$$

com $\|P(X^{(22)})\|_\infty = 0.00276$ e $\rho(J_f(x^{(22)})) = 0.41106$.

Observação 3.2.10. Neste exemplo, o Algoritmo A.1.1 MN claramente não converge. Para a mesma aproximação inicial, $X^{(0)} = 48.05.I_3$, obtém-se uma sucessão $X_0, X_1, \dots, X_{4000}$, com

$$X_{4000} = \begin{bmatrix} 78.5719 & -431.678 & 196.776 \\ 147.124 & -806.574 & 367.412 \\ 291.594 & -1598.28 & 727.987 \end{bmatrix},$$

onde $\|P(X_{4000})\|_2 = 111.834$. De fato, uma vez que os coeficientes do polinômio matricial são reais e considerando uma aproximação inicial real, todas as iterações do método de Newton serão reais o que invalida a sua convergência para um solvente com coeficientes complexos (para uma aproximação inicial complexa o método de Newton já converge).

Observação 3.2.11. Tendo em conta os valores da Tabela 3.1, observa-se que após a sexta aproximação a distância destas ao solvente ($\|X^{(k)} - X^{(14)}\|_\infty$) diminui e o seu raio espectral ($\rho(J_f(x^{(k)}))$) mantém-se menor do que 1 o que sugere que a partir desta iteração o exemplo já se encontra nas condições de convergência (Teorema 3.2.6) em $V_\varepsilon(s)$, com $\varepsilon \approx 3.83926$.

k	$\ X^{(k)} - X^{(14)}\ _\infty$	$\rho(J_f(x^{(k)}))$
0	49.7494	0.0723773
1	3.7463	0.291798
2	2.71409	0.585448
3	11.9497	1.80783
4	13.4275	0.788064
5	12.9891	0.384312
6	3.83926	0.416268
7	0.416216	0.421291
8	0.148047	0.420923
9	0.0619636	0.420028
10	0.0252052	0.41313
11	0.0109852	0.408552
12	0.00433661	0.408914
13	0.00152436	0.409414
14	0.	0.409693

Tabela 3.1: Evolução do raio espectral das aproximações no método do Ponto Fixo

No exemplo apresentado a seguir aplica-se o método do Ponto Fixo (PF) quer considerando uma aproximação inicial $X^{(0)}$ “próxima” de um dos 5 solventes do polinômio $P(X) = X^2 + A_1X + A_2$ quer considerando $X^{(0)} = -A_1$, sem a preocupação de validar a condição ii) do Teorema 3.2.6 (que de uma forma geral é difícil de estabelecer).

Exemplo 3.2.12. Considere-se o polinômio

$$P(X) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} X^2 + \begin{bmatrix} -5 & 0 \\ -34.667 & -4 \end{bmatrix} X + \begin{bmatrix} 4 & 0 \\ 34.667 & 104 \end{bmatrix}.$$

A equação matricial $P(X) = 0$ tem 5 solventes,

$$X_1 = \begin{bmatrix} 1 & 0 \\ 0 & 2 + 10i \end{bmatrix}, X_2 = \begin{bmatrix} 1 & 0 \\ 0 & 2 - 10i \end{bmatrix}, X_3 = \begin{bmatrix} 1 & 3 \\ 0 & 4 \end{bmatrix},$$

$$X_4 = \begin{bmatrix} 4 & 0 \\ 2 - 10i & 2 + 10i \end{bmatrix}, \quad X_5 = \begin{bmatrix} 4 & 0 \\ 2 + 10i & 2 - 10i \end{bmatrix}.$$

Considerando

$$X^{(0)} = \begin{bmatrix} 4.02 & 0.2 \\ 2.02 - 10i & 2.02 + 10i \end{bmatrix}, \quad \text{com } \|P(X^{(0)})\|_2 = 0.91307,$$

o método do Ponto Fixo, Algoritmo 3.2.1 PF, converge em 21 iterações para uma aproximação

$$X^{(21)} = \begin{bmatrix} 4 + 2.05755 \times 10^{-6}i & 0 \\ 2.00001 - 10i & 2. + 10i \end{bmatrix}, \quad \text{com } \|P(X^{(21)})\|_2 < 0.5 \times 10^{-4}.$$

Para a aproximação inicial $X^{(0)} = -A_1$, que satisfaz as condições de convergência, o método do Ponto Fixo, Algoritmo 3.2.1 PF, converge em 23 iterações para uma aproximação $X^{(23)}$ de X_5 , com $\|P(X^{(23)})\|_2 = 0.000349$.

$$\begin{aligned} X^{(0)} &= \begin{bmatrix} 5 & 0 \\ 34.667 & 4 \end{bmatrix}, \quad \text{com } \|P(X^{(0)})\|_2 = 109.663, \\ &\vdots \\ X^{(23)} &= \begin{bmatrix} 4.00002 & 0 \\ 2.00004 - 10.0001i & 2. + 10.i \end{bmatrix}, \quad \text{com } \|P(X^{(23)})\|_2 = 0.000349. \end{aligned}$$

Observação 3.2.13. Neste caso, o polinómio não está nas condições de convergência do método de Traub, Algoritmo A.1.3, pois não existe um solvente dominante. Para a mesma aproximação inicial $X^{(0)} = -A_1$, obtém-se

$$X^{(30)} = \begin{bmatrix} 4. & 0 \\ -4.03662 \times 10^{13} & 0.345929 \end{bmatrix}, \quad \text{com } \|P(X^{(30)})\|_2 = 1.25764 \times 10^{15}.$$

3.3 Método de Newton Vetorial

Nesta secção serão apresentados três Algoritmos construídos a partir do chamado método de Newton Vetorial. O Algoritmo principal que resulta da aplicação do método de Newton-Rapshson clássico à equação vetorial $F(x) = \text{vec}(P(X)) = 0$ e tratando-se apenas de uma nova versão deste método, é possível acelerar a sua convergência através da através da otimização do acréscimo $h^{(k)}$ (adaptação dos métodos 3.2 e 3.3 de definidos em Jian-hui [26], pág. 647, 650 e 651, a este contexto) obtendo-se os dois Algoritmos com Procura Unidimensional.

No primeiro destes Algoritmos é introduzido um novo parâmetro $\varepsilon_0 > 0$ e, em cada iteração, sempre que o valor da função F em $x^{(k)} + h^{(k)}$ for superior a ε_0 procede-se à otimização do acréscimo $h^{(k)}$ minimizando o valor de $\|F(x^{(k)} + sh^{(k)})\|_2$ em função de $s \in [0, 2]$. No segundo

Algoritmo, para além da otimização já referida, considera-se ainda, sempre que o valor da função F em $x^{(k)} + h^{(k)}$ seja inferior a ε_0 , um termo da sucessão intermédio $x^{(k,1)}$. A solução $h^{(k,1)}$ de $J_F(x^{(k)})h^{(k,1)} = -F(x^{(k,1)})$ pode ser obtida a partir da resolução já considerada na determinação de $h^{(k)}$ uma vez que a matriz $J_F(x^{(k)})$ é comum a ambos os sistemas.

Na sequência da construção do método do Ponto Fixo (PF), Algoritmo 3.2.3, mantendo a mesma ideia relativamente à análise do polinómio matricial em termos das suas entradas, procedeu-se à adaptação do método de Newton a este contexto.

Como foi visto em (3.2.1), o polinómio matricial $P(X)$ pode ser associado à função

$$\begin{aligned} F : \mathbb{C}^{n^2} &\longrightarrow \mathbb{C}^{n^2} \\ x = \text{vec}(X) &\longrightarrow F(x) = \text{vec}(P(X)) \end{aligned} \quad (3.3.1)$$

através da composição

$$\begin{array}{ccc} \mathbb{C}^{n \times n} & \xrightarrow{P} & \mathbb{C}^{n \times n} \\ x & \xrightarrow{\quad} & P(X) \\ \text{vec}^{-1} \uparrow & & \downarrow \text{vec} \\ \mathbb{C}^{n^2} & \xrightarrow{F} & \mathbb{C}^{n^2} \\ & & F(x) \end{array} ,$$

e onde cada coordenada $F_l(x)$, $l = 1, 2, \dots, n^2$ com $l = (j - 1)n + i$, corresponde à coordenada $i = \text{mod}(l - 1, n) + 1, j = \text{quoc}(l - 1, n) + 1$ de $P(X)$. Deste ponto de vista, aplicando o método de Newton à função $F : \mathbb{C}^{n^2} \longrightarrow \mathbb{C}^{n^2}$, obtém-se

$$x^{(k+1)} = x^{(k)} + Q_k, \quad k = 1, 2, \dots$$

com Q_k solução do sistema

$$J_F(x^{(k)})Q_k = -F(x^{(k)}), \quad (3.3.2)$$

onde J_F é a matriz Jacobiana de F .

Dado um polinómio matricial de grau $m = 2$,

$$P(X) = A_0X^2 + A_1X + A_2,$$

com $A_k = [a_{i,j}^k]_{i,j=1,\dots,n}$, $k = 1, 2$, a função F definida atrás pode ser escrita na forma

$$F_l = \sum_{p=1}^n a_{i,p}^0 \left(\sum_{k=1}^n x_{p,k} x_{k,j} \right) + \sum_{k=1}^n a_{i,k}^1 x_{k,j} + a_{i,j}^2, \quad (3.3.3)$$

$l = (j - 1)n + i$, e a respetiva matriz Jacobiana definida por

$$J_F = I_n \otimes A_0X + X^T \otimes A_0 + I_n \otimes A_1, \quad (3.3.4)$$

onde \otimes representa o produto de Kronecker, Definição 2.1.15.

Caraterizando o sistema $P(X) = 0$ em termos de zeros da função analítica F com matriz Jacobiana definida por (3.3.4) para $m = 2$ e, mais geralmente, (3.3.6), obtém-se um sistema linear com a mesma dimensão (3.3.2) evitando-se o cálculo da derivada de Fréchet bem como a resolução da respetiva equação matricial de Sylvester (3.1.2).

Observação 3.3.1. Tratando-se de uma abordagem do problema $P(X) = 0$ diferente da considerada nos artigos Higham [18] e Jian-hui [26], pág. 310 e 649, a matriz Jacobiana J_F é idêntica às matrizes P e F^* definidas apenas para $m = 2$. Em Bin Zhou [41], pág. 328, está definido o “Kronecker matrix polynomial” $A_{\otimes}(F)$, para qualquer m mas não do ponto de vista da derivada de Fréchet que conduz neste caso à matriz J_F .

De uma maneira geral, para um polinômio de grau m , a função F definida por

$$F = \text{vec}(P(X)), \quad (3.3.5)$$

tem uma matriz Jacobiana que se pode representar na forma

$$J_F = \sum_{i=1}^m \left(\sum_{j=1}^i (X^{j-1})^T \otimes A_{m-i} X^{i-j} \right). \quad (3.3.6)$$

Das seguintes considerações pode ver-se como a igualdade (3.3.6) pode ser obtida.

— A derivada de Fréchet do polinômio $P(X) = X^m + A_1 X^{m-1} + \dots + A_m$ está definida por

$$P'(X)[H] = \sum_{1 \leq i \leq m} \left(\sum_{1 \leq j \leq i} A_{m-i} X^{i-j} H X^{j-1} \right).$$

— Pelo lema 2.1.19,

$$\text{vec}(A_{m-i} X^{i-j} H X^{j-1}) = ((X^{j-1})^T \otimes A_{m-i} X^{i-j}) \text{vec}(H).$$

— O operador $\text{vec} : \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{n^2}$ comuta com a operação derivação D . Ou, dito de outra forma,

$$D_{\text{vec}}(X)[H] = \text{vec}(H), \quad H \in \mathbb{C}^{n \times n}.$$

Assim, considerando (3.3.5),

$$F(x) = \text{vec}(P(X)), \quad (3.3.7)$$

com $x = \text{vec}(X)$, e derivando ambos os membros de (3.3.7) no ponto $X \in \mathbb{C}^{n \times n}$ e na direção de $H \in \mathbb{C}^{n \times n}$, tem-se

$$\begin{aligned} D_{F \circ \text{vec}}(X)[H] &= D_{\text{vec} \circ P}(X)[H] \\ D_F(\text{vec}(X))D_{\text{vec}}(X)[H] &= D_{\text{vec}}(P(X))D_P(X)[H] \\ D_F(x)[\text{vec}(H)] &= \text{vec}(D_P(X)[H]) \\ J_F(x)\text{vec}(H) &= \text{vec}(P'(X)[H]), \end{aligned}$$

onde

$$\begin{aligned} \text{vec}(P'(X)[H]) &= \text{vec} \left(\sum_{1 \leq i \leq m} \sum_{1 \leq j \leq i} A_{m-i} X^{i-j} H X^{j-1} \right) \\ &= \left(\sum_{1 \leq i \leq m} \sum_{1 \leq j \leq i} ((X^{j-1})^T \otimes A_{m-i} X^{i-j}) \right) \text{vec}(H). \end{aligned}$$

Esquemáticamente, considerando a composição já vista em (3.3.1) e derivando as funções

$(vec \circ P)(X)$ e $(F \circ vec)(X)$ no ponto $X \in \mathbb{C}^{n \times n}$ e na direção de $H \in \mathbb{C}^{n \times n}$, com $x = vec(X)$ e $h = vec(H)$, tem-se

$$\begin{array}{ccc}
 \mathbb{C}^{n \times n} & \xrightarrow{P} & \mathbb{C}^{n \times n} \\
 x & \xrightarrow{\quad} & P(x) \\
 \downarrow vec & & \downarrow vec \\
 \mathbb{C}^{n^2} & \xrightarrow{F} & \mathbb{C}^{n^2} \\
 & & F(x)
 \end{array}
 \quad \xrightarrow{D} \quad
 \begin{array}{ccc}
 \mathbb{C}^{n \times n} & \xrightarrow{P'(X)} & \mathbb{C}^{n \times n} \\
 H & \xrightarrow{\quad} & P'(X)[H] \\
 \downarrow vec & & \downarrow vec \\
 \mathbb{C}^{n^2} & \xrightarrow{F'(x)} & \mathbb{C}^{n^2} \\
 & & F'(x)h
 \end{array}
 .$$

De acordo com as considerações feitas e considerando ainda a composição já vista em (3.3.1) é imediato o seguinte resultado.

Observação 3.3.2. Se x é um zero de F , então $X = vec^{-1}(x)$ é um solvente de $P(X)$.

Por construção, se x é um zero de F então, por construção (3.3.5) e considerando $X = vec^{-1}(x)$, $vec(P(X)) = F(vec(X)) = F(x) = 0$, pelo que, $P(X) = 0$.

Uma vez definida a função F , (3.3.5), e a derivada respetiva, (3.3.6), o Método de Newton Vetorial, (NV) é agora deduzido através da expansão de $F(x)$ em x na direção de h , onde $X = vec^{-1}(x)$ é uma aproximação do solvente de $P(X)$ e $H = vec^{-1}(h)$,

$$F(x+h) = F(x) + F'(x)(h) + \frac{1}{2}F''(x)(h) + \dots,$$

desprezando os termos de ordem maior ou igual a 2 em h e admitindo que $F(x+h) = 0$ para um vetor h a determinar. Isto é,

$$F(x) + F'(x)(h) \approx 0. \tag{3.3.8}$$

3.3.3 Algoritmo Principal

Como referido, o Algoritmo principal resulta da aplicação do método de Newton-Raphson clássico à equação vetorial $F(x) = vec(P(X)) = 0$. Isto é, resolvendo (3.3.8) em ordem a h obtém-se uma nova aproximação $x+h$.

Assim, dada uma aproximação inicial $x^{(0)}$, para cada $k = 1, 2, \dots$, a aproximação seguinte é calculada resolvendo a equação

$$F'_{(x^{(k)})}(h^{(k)}) = -F(x^{(k)}),$$

em ordem a $h^{(k)}$ de modo a que $F(x^{(k)} + h^{(k)}) \approx 0$, isto é, considerando $x^{(k+1)} = x^{(k)} + h^{(k)}$.

Nestas condições, o Algoritmo do Método de Newton Vetorial pode ser expresso usando os seguintes passos:

Algoritmo 3.3.1 NV

- 1: É dada uma aproximação inicial $x^{(0)}$ e uma margem de erro $\varepsilon > 0$.
 - 2: para $k = 0, 1, 2, \dots$ fazer
 - 3: se $\|F(x^{(k)})\|_2 < \varepsilon$ então
 - 4: o método pára.
 - 5: caso contrário
 - 6: $h^{(k)}$ solução de $J_F(x^{(k)})h^{(k)} = -F(x^{(k)})$;
 - 7: $x^{(k+1)} = x^{(k)} + h^{(k)}$;
 - 8: $X = vec^{-1}(x^{(k+1)})$ é uma aproximação do solvente.
-

A convergência deste método (NV, Algoritmo 3.3.1), atendendo a que corresponde ao método de Newton “clássico” aplicado à função $F : \mathbb{C}^{n^2} \rightarrow \mathbb{C}^{n^2}$, satisfaz os critérios de convergência dos Teoremas de Kantorovich de Smale ou de Nesterov-Nemirovskii (ver Gutierrez [17] e [16], pág. 9 e 135 e Argyros [1] e [2], pág. 154 e 12).

3.3.4 Convergência

O primeiro resultado referido (Teorema de Kantorovich), válido para uma função entre espaços Banach $F : \mathcal{X} \rightarrow \mathcal{Y}$, baseia-se na convergência do método de Newton escalar relativamente ao polinômio $p(t) = \frac{\gamma}{2}t^2 - t + \beta$, com $\gamma, \beta \in \mathbb{R}^+$ e cujas raízes

$$t^* = \frac{1 - \sqrt{1 - 2\beta\gamma}}{\gamma} \text{ e } t^{**} = \frac{1 + \sqrt{1 - 2\beta\gamma}}{\gamma},$$

são reais e positivas, com $t^* < t^{**}$, desde que $2\beta\gamma < 1$.

Com efeito, no intervalo $I = \left[0, \frac{1}{\gamma}\right]$, tem-se $t^* \in I$ e

$$\begin{aligned} p'(t) &= \gamma t - 1 < 0, \quad \forall t \in I; \\ p''(t) &= \gamma > 0, \quad \forall t \in I; \\ p(0)p''(t) &> 0, \quad \forall t \in I. \end{aligned}$$

Portanto, para $t_0 = 0$ a sucessão obtida através do método de Newton,

$$t_{k+1} = t_k - \frac{p(t_k)}{p'(t_k)}, \tag{3.3.9}$$

é crescente e converge para t^* .

Observação 3.3.5. O método de Newton definido por (3.3.9) nas condições referidas é convergente e tem ordem de convergência 2:

$$\frac{t^* - t_{k+1}}{(t^* - t_k)^2} = -\frac{p''(\xi)}{p'(t_k)} = \frac{\gamma}{1 - \gamma t_k}.$$

Além disso, são válidos os seguintes resultados, necessários na demonstração do critério de convergência:

Lema 3.3.6. Considerando $\varphi_k = \frac{\theta^{2^k}}{1 - \theta^{2^k}}$, com $k = 0, 1, 2, \dots$, onde $\theta = \frac{t^*}{t^{**}} < 1$, então

$$\frac{\varphi_k^2}{2\varphi_k + 1} = \varphi_{k+1}. \tag{3.3.10}$$

Demonstração. Com efeito, tem-se

$$\begin{aligned}
 \frac{\varphi_k^2}{2\varphi_k + 1} &= \frac{\frac{(\theta^{2^k})^2}{(1-\theta^{2^k})^2}}{2\frac{\theta^{2^k}}{1-\theta^{2^k}} + 1} \\
 &= \frac{\frac{(\theta^{2^k})^2}{(1-\theta^{2^k})^2}}{\frac{\theta^{2^k}+1}{1-\theta^{2^k}}} \\
 &= \frac{(\theta^{2^k})^2}{(1-\theta^{2^k})^2} \times \frac{1-\theta^{2^k}}{\theta^{2^k}+1} \\
 &= \frac{\theta^{2 \times 2^k}}{(1-\theta^{2^k})(\theta^{2^k}+1)} \\
 &= \frac{\theta^{2^{k+1}}}{1-(\theta^{2^k})^2} = \frac{\theta^{2^{k+1}}}{1-\theta^{2^{k+1}}} \\
 &= \varphi_{k+1}.
 \end{aligned}$$

■

Lema 3.3.7. Seja t_k a sucessão definida por (3.3.9). Então

$$t^* - t_k = (t^{**} - t^*)\varphi_k. \quad (3.3.11)$$

Demonstração. A prova é feita por indução. Para $k = 0$,

$$t^* - t_0 = t^* - 0 = (t^{**} - t^*)\frac{t^*}{t^{**} - t^*} = (t^{**} - t^*)\frac{\theta}{1 - \theta} = (t^{**} - t^*)\varphi_0.$$

Admitindo que o resultado é válido para k , isto é,

$$t^* - t_k = (t^{**} - t^*)\varphi_k,$$

então, considerando t_{k+1} , tem-se

$$t^* - t_{k+1} = t^* - t_k + \frac{p(t_k)}{p'(t_k)}.$$

Como t^* e t^{**} são as raízes do polinômio $p(t)$, este fatoriza-se na forma $p(t) = \frac{\gamma}{2}(t - t^*)(t - t^{**})$, simplificando o quociente

$$\begin{aligned}
 \frac{p(t_k)}{p'(t_k)} &= \frac{\frac{\gamma}{2}(t_k - t^*)(t_k - t^{**})}{\frac{\gamma}{2}(2t_k - t^* - t^{**})} \\
 &= \frac{(t_k - t^*)(t_k - t^{**})}{(2t_k - t^* - t^{**})} \\
 &= \frac{(t_k - t^*)(t_k - t^{**})}{2(t_k - t^*) + (t^* - t^{**})},
 \end{aligned}$$

obtendo-se

$$\begin{aligned}
 t^* - t_{k+1} &= t^* - t_k + \frac{(t_k - t^*)(t_k - t^{**})}{2(t_k - t^*) + (t^* - t^{**})} \\
 &= \frac{-(t^* - t_k)^2}{2(t_k - t^*) + (t^* - t^{**})} \\
 &= \frac{-(t^{**} - t^*)^2 \varphi_k^2}{-2(t^{**} - t^*)\varphi_k + t^* - t^{**}} \\
 &= \frac{-(t^{**} - t^*)^2 \varphi_k^2}{-(t^{**} - t^*)(2\varphi_k + 1)} \\
 &= \frac{(t^{**} - t^*)\varphi_k^2}{(2\varphi_k + 1)} \\
 &= (t^{**} - t^*)\varphi_{k+1}.
 \end{aligned}$$

■

Lema 3.3.8. (Lema de Banach sobre Operadores Invertíveis, ver Driver [8], pág.95, Datta [4], pág. 22) Se T é um operador linear e contínuo definido de \mathcal{X} em \mathcal{X} , então existe T^{-1} se e só se existir um operador linear P de \mathcal{X} em \mathcal{X} , limitado¹ e invertível tal que $\|I - PT\| < 1$. Nestas condições,

$$\|T^{-1}\| \leq \frac{\|P\|}{1 - \|I - PT\|}.$$

De seguida apresenta-se uma versão do Teorema de Kantorovich baseada no desenvolvimento feito por Gutierrez [17], pág. 11, em que a demonstração foi adaptada ao contexto agora apresentado.

Teorema 3.3.9. (Kantorovich) Se a função $F : \mathbb{C}^{n^2} \rightarrow \mathbb{C}^{n^2}$ é diferenciável em \mathbb{C}^{n^2} e verifica as seguintes condições em $B_r(x^{(0)}) = \{x \in \mathbb{C}^{n^2} : \|x - x^{(0)}\| \leq r\}$:

1. A matriz $J_F(x^{(0)}) = J$ é não singular (existe $J_F^{-1}(x^{(0)}) = J^{-1}$);
2. $\|J^{-1}(J_F(x) - J_F(y))\| \leq \gamma\|x - y\|$, $x, y \in B_r(x^{(0)})$, $\gamma > 0$;
3. $\|J^{-1}F(x^{(0)})\| \leq \beta$;
4. $\beta\gamma < 1/2$;
5. $t^* = \frac{1 - \sqrt{1 - 2\beta\gamma}}{\gamma} \leq r$,

então

- i) $x^{(k+1)} = x^{(k)} - J_F^{-1}(x^{(k)})F(x^{(k)})$, está bem definido para cada $k = 1, 2, \dots$ e converge para uma solução s de $F(x) = 0$;
- ii) A solução $s \in \overline{B_{t^*}(x^{(0)})}$ é única em $B_{t^{**}}(x^{(0)})$, onde $t^{**} = \frac{1 + \sqrt{1 - 2\beta\gamma}}{\gamma}$;
- iii) A sucessão $x^{(k+1)}$, definida em i), tem ordem de convergência quadrática e

$$\|x^{(k)} - s\| < (t^{**} - t^*) \frac{\theta^{2^k}}{1 - \theta^{2^k}}, \text{ onde } \theta = \frac{t^*}{t^{**}} < 1.$$

¹ P um operador linear e contínuo entre espaços de Banach \mathcal{X}, \mathcal{Y} é limitado em \mathcal{X} , i.é., $\|P(x)\| \leq c\|x\|$.

Demonstração. Para cada $x \in \overline{B_{t^*}(x^{(0)})}$,

$$\|I - J^{-1}J_F(x)\| = \|J^{-1}J_F(x^{(0)}) - J^{-1}J_F(x)\| \leq \gamma\|x^{(0)} - x\| \leq \gamma t^* = 1 - \sqrt{1 - 2\beta\gamma} < 1,$$

logo, pelo Lema 3.3.8, o operador $J_F(x)$ é invertível para todo $x \in \overline{B_{t^*}(x^{(0)})}$ e

$$\|J_F^{-1}(x)\| \leq \frac{\|J^{-1}\|}{1 - \|I - J^{-1}J_F(x)\|}, \text{ ou } \|J_F^{-1}(x)J\| \leq \frac{1}{1 - \|I - J^{-1}J_F(x)\|} \leq \frac{1}{1 - \gamma t^*}.$$

Assim, $x^{(k+1)} = x^{(k)} - J_F^{-1}(x^{(k)})F(x^{(k)})$, está bem definido para cada $x^{(k)} \in \overline{B_{t^*}(x^{(0)})}$.

Por outro lado, $\|x^{(1)} - x^{(0)}\| \leq t_1 - t_0$ pois

$$\|x^{(1)} - x^{(0)}\| = \|J^{-1}F(x^{(0)})\| \leq \beta = -\frac{p(0)}{p'(0)} = t_1 - t_0,$$

e, admitindo que $\|x^{(k)} - x^{(k-1)}\| \leq t_k - t_{k-1}$, então

$$\|x^{(k+1)} - x^{(k)}\| = \|J_F^{-1}(x^{(k)})F(x^{(k)})\| \leq \|J_F^{-1}(x^{(k)})J\| \|J^{-1}F(x^{(k)})\| \quad (3.3.12)$$

onde,

$$\begin{aligned} J^{-1}F(x^{(k)}) &= J^{-1}[F(x^{(k)}) - F(x^{(k-1)}) - J_F(x^{(k-1)})(x^{(k)} - x^{(k-1)})] \\ &= \int_0^1 J^{-1} \left[F'(x^{(k)} - t(x^{(k)} - x^{(k-1)})) - F'(x^{(k-1)}) \right] (x^{(k)} - x^{(k-1)}) dt, \end{aligned}$$

e portanto,

$$\begin{aligned} \|J^{-1}F(x^{(k)})\| &\leq \int_0^1 \|J^{-1} [J_F(x^{(k)} - t(x^{(k)} - x^{(k-1)})) - J_F(x^{(k-1)})]\| \|x^{(k)} - x^{(k-1)}\| dt \\ &\leq \int_0^1 \gamma \|x^{(k)} - t(x^{(k)} - x^{(k-1)}) - x^{(k-1)}\| \|x^{(k)} - x^{(k-1)}\| dt \\ &\leq \int_0^1 \gamma(1-t) \|x^{(k)} - x^{(k-1)}\|^2 dt \\ &\leq \frac{1}{2} \gamma \|x^{(k)} - x^{(k-1)}\|^2. \end{aligned}$$

Voltando a (3.3.12)

$$\begin{aligned} \|x^{(k+1)} - x^{(k)}\| &\leq \|J_F^{-1}(x^{(k)})J\| \|J^{-1}F(x^{(k)})\| \\ &\leq \frac{1}{1 - \|I - J^{-1}J_F(x^{(k)})\|} \frac{1}{2} \gamma \|x^{(k)} - x^{(k-1)}\|^2 \\ &\leq \frac{\gamma}{2} \frac{1}{1 - \gamma \|x^{(k)} - x^{(0)}\|} \|x^{(k)} - x^{(k-1)}\|^2, \end{aligned}$$

onde a última desigualdade resulta do fato de

$$\begin{aligned} \|I - J^{-1}J_F(x^{(k)})\| \leq \gamma \|x^{(k)} - x^{(0)}\| &\Rightarrow -\|I - J^{-1}J_F(x^{(k)})\| \geq -\gamma \|x^{(k)} - x^{(0)}\| \\ &\Rightarrow 1 - \|I - J^{-1}J_F(x^{(k)})\| \geq 1 - \gamma \|x^{(k)} - x^{(0)}\| \\ &\Rightarrow \frac{1}{1 - \|I - J^{-1}J_F(x^{(k)})\|} \leq \frac{1}{1 - \gamma \|x^{(k)} - x^{(0)}\|}. \end{aligned}$$

Atendendo a que $\|x^{(k)} - x^{(0)}\| \leq t_k$ pois,

$$\|x^{(k)} - x^{(0)}\| \leq (t_k - t_{k-1}) + (t_{k-1} - t_{k-2}) + \cdots + (t_1 - t_0) = t_k,$$

e tendo em conta que $\beta = t_k - \gamma t_k t_{k-1} + \frac{\gamma}{2} t_{k-1}^2 + t_{k-1}$ então pode concluir-se,

$$\|x^{(k+1)} - x^{(k)}\| \leq \frac{\gamma (t_k - t_{k-1})^2}{2 - 1 - \gamma t_k} = -\frac{p(t_k)}{p'(t_k)} = t_{k+1} - t_k. \quad (3.3.13)$$

Assim, $\|x^{(k)} - x^{(k-1)}\| \leq t_k - t_{k-1}$, $\forall k = 1, 2, \dots$, e atendendo a que a sucessão t_k é convergente (Observação 3.3.5) e, portanto, uma sucessão de Cauchy, então por (3.3.13), x_k é também uma sucessão de Cauchy,

$$\|x^{(k+p)} - x^{(k)}\| \leq (t_{k+p} - t_{k+p-1}) + (t_{k+p-1} - t_{k+p-2}) + \cdots + (t_{k+1} - t_k) = t_{k+p} - t_k,$$

logo convergente ($x_k \rightarrow s \in \overline{B_{t^*}(x^{(0)})}$, $t_k < t^*$, $\forall k$).

Para além disso, tem-se que

$$\|J^{-1}F(x^{(k)})\| \leq \frac{\gamma}{2} \|x^{(k)} - x^{(k-1)}\|^2 \leq \frac{\gamma}{2} (t_k - t_{k-1})^2, \quad (3.3.14)$$

pelo que, passando ambos os membros de (3.3.14) ao limite, tem-se

$$F(s) = 0,$$

o que prova i).

Para provar a unicidade da solução s em $B_{t^{**}}(x^{(0)})$, considere-se outra raiz da equação $F(x) = 0$, $s' \in B_{t^{**}}(x^{(0)})$. Assim,

$$\begin{aligned} 0 = J^{-1}(F(s) - F(s')) &= J^{-1} \int_0^1 F'(s - t(s - s'))(s - s') dt \\ &= \left(\int_0^1 J^{-1} F'(s - t(s - s')) dt \right) (s - s') \\ &= M(s - s'), \end{aligned}$$

e

$$\begin{aligned} \|M - I\| &= \left\| \int_0^1 J^{-1} [F'(s - t(s - s')) - F'(x^{(0)})] dt \right\| \\ &\leq \int_0^1 \|J^{-1} [F'(s - t(s - s')) - F'(x^{(0)})]\| dt \\ &\leq \int_0^1 \gamma \|s - t(s - s') - x^{(0)}\| dt \\ &\leq \int_0^1 \gamma \|s - x^{(0)} - t(s - x^{(0)}) - t(x^{(0)} - s')\| dt \\ &\leq \int_0^1 \gamma \|(1-t)(s - x^{(0)}) + t(s' - x^{(0)})\| dt \\ &< \frac{\gamma}{2} (t^* + t^{**}) = 1, \end{aligned}$$

logo M é invertível (Lema 3.3.8, onde $\|I - MI\| = \|I - M\| < 1$, com I um operador linear limitado

e invertível) e portanto $s = s'$, o que prova ii).

Relativamente à alínea iii), para $p \geq 1$, $\|x^{(k+p)} - x^{(k)}\| \leq t_{k+p} - t_k$ e passando ao limite com $p \rightarrow +\infty$, tem-se ainda

$$\|s - x^{(k)}\| \leq t^* - t_k, \quad (3.3.15)$$

onde, pelo lema 3.3.7,

$$t^* - t_k = (t^{**} - t^*) \frac{\theta^{2^k}}{1 - \theta^{2^k}}, \text{ com } \theta = \frac{t^*}{t^{**}} < 1.$$

■

O Teorema anterior assegura que, nas condições de convergência (Teorema 3.3.9), a sucessão x_k é uma sucessão de Cauchy e converge para uma solução s da equação $F(x) = 0$. Segundo a Observação 3.3.2, considerando $S = \text{vec}^{-1}(s)$, tem-se que $P(S) = 0$, pelo que fica também assegurado que, no ponto 8 do Algoritmo 3.3.1, se obtém realmente uma aproximação de um solvente de $P(X)$, situação descrita através do seguinte Corolário.

Corolário 3.3.10. A sucessão $X^{(k)} = \text{vec}^{-1}(x^{(k)})$ converge para um solvente de $P(X)$.

Assegurada a convergência do método de Newton Vetorial, fez-se posteriormente a análise do comportamento do respetivo Algoritmo 3.3.1, quando comparado com outros Algoritmos, nomeadamente os Algoritmos 2.1, 3.1, 3.2 e 3.3, definidos em Jian-hui [26].

Nesta comparação usou-se o polinómio matricial,

$$P(X) = \begin{bmatrix} 17.6 & 1.28 & 2.89 \\ 1.28 & 0.84 & 0.413 \\ 2.89 & 0.413 & 0.725 \end{bmatrix} X^2 + \begin{bmatrix} 7.66 & 2.45 & 2.1 \\ 0.23 & 1.04 & 0.223 \\ 0.6 & 0.756 & 0.658 \end{bmatrix} X + \begin{bmatrix} 121 & 18.9 & 15.9 \\ 0 & 2.7 & 0.145 \\ 11.9 & 3.64 & 15.5 \end{bmatrix}. \quad (3.3.16)$$

Tendo em conta a Observação 3.2.10 consideram-se aproximações iniciais complexas para assegurar que os métodos convergem, mesmo no caso de o solvente ser complexo.

Usando o Algoritmo 3.3.1 (NV), com a aproximação inicial

$$x^{(0)} = \text{vec}(iA_2) = \text{vec} \left(\begin{bmatrix} 121i & 18.9i & 15.9i \\ 0 & 2.7i & 0.145i \\ 11.9i & 3.64i & 15.5i \end{bmatrix} \right),$$

obtém-se após 9 iterações,

$$X^{(9)} = \text{vec}^{-1}(x^{(9)}) = \begin{bmatrix} -0.365507 + 3.20705i & 0.00526813 + 0.19849i & 0.0502906 - 0.728978i \\ 0.226552 - 2.05575i & -0.568877 + 1.39304i & 0.245173 - 2.21197i \\ 1.00784 - 2.36984i & -0.0508553 + 0.106218i & -0.755884 + 8.08455i \end{bmatrix},$$

com $\|F(x^{(9)})\|_2 = 2.61022 \times 10^{-7}$.

Nas tabelas 3.2, 3.3 e 3.4 apresentam-se os resultados obtidos através do Algoritmo 3.3.1 (NV) e dos Algoritmos 2.1, 3.1, 3.2 e 3.3, definidos em Jian-hui [26], pág. 647, 650 e 651, quer em termos de iterações necessárias (k), para se obter uma aproximação do solvente com uma determinada margem de erro (ε), quer em termos de tempo de CPU gasto nesse cálculo (t — medido em

segundos através do comando `Timing[]` do Mathematica).

Assim, usando o Algoritmo 3.3.1, com a aproximação inicial $x^{(0)} = \text{vec}(iI_3)$, obtêm-se para $k = 7$,

$$X^{(7)} = \text{vec}^{-1}(x^{(7)}) = \begin{bmatrix} -0.365507 + 3.20705i & 0.00526813 + 0.19849i & 0.0502906 - 0.728978i \\ 0.226552 - 2.05575i & -0.568877 + 1.39304i & 0.245173 - 2.21197i \\ 1.00784 - 2.36984i & -0.0508553 + 0.106218i & -0.755884 + 8.08455i \end{bmatrix},$$

com $\|F(x^{(7)})\|_2 = 3.27627 \times 10^{-10}$. O tempo decorrido para o cálculo de $X^{(7)}$ foi 0.125s.

$\varepsilon = 10^{-9}$		NV	2.1	3.1	3.2	3.3
$\varepsilon_0 = 10^{-1}$	k	7	7	6	6	6
$X^{(0)} = iI_3$	t	0.125s	0.172s	1.406s	.766s	1.047s

Tabela 3.2: Comparação do Algoritmo 3.3.1 com os Algoritmos 2.1, 3.1, 3.2 e 3.3.

$\varepsilon = 10^{-9}$		NV	2.1	3.1	3.2	3.3
$\varepsilon_0 = 10^{-1}$	k	7	7	5	5	5
$X^{(0)} = 10iI_3$	t	0.11s	0.171s	1.188s	0.562s	0.782s

Tabela 3.3: Comparação do Algoritmo 3.3.1 com os Algoritmos 2.1, 3.1, 3.2 e 3.3.

$\varepsilon = 10^{-9}$		NV	2.1	3.1	3.2	3.3
$\varepsilon_0 = 10^{-1}$	k	20	20	7	7	6
$X^{(0)} = 10^5 iI_3$	t	0.25s	0.484s	1.969s	1.203s	1.531s

Tabela 3.4: Comparação do Algoritmo 3.3.1 com os Algoritmos 2.1, 3.1, 3.2 e 3.3.

Destaca-se o fato de que o presente Algoritmo é o mais “econômico” em termos de tempo gasto no cálculo da aproximação (ver tabela (3.4)): para calcular 20 iterações, o método (NV) necessita apenas 1/6 do tempo daquele que usa o método (3.3) a calcular $X^{(6)}$.

Necessita, contudo, mais iterações que os Algoritmos 3.1, 3.2 e 3.3 mas, como se trata apenas de uma nova versão do método de Newton, a sua convergência é também suscetível de ser acelerada: fazendo uma adaptação dos métodos 3.2 e 3.3 ao contexto do Método de Newton Vetorial (NV), obtêm-se os Algoritmos 3.3.2 e 3.3.3 definidos a seguir.

3.3.11 Algoritmos com Procura Unidimensional

Basicamente, a Procura Unidimensional permite acelerar a convergência do Método Newton Vetorial (NV) através da otimização do acréscimo $h^{(k)}$. A ideia consiste em admitir que a solução do sistema $J_F(x^{(k)})h^{(k)} = -F(x^{(k)})$ determina a melhor direção $h^{(k)}$ para encontrar a solução da equação $F(x) = 0$, mas que o seu comprimento $\|h^{(k)}\|_2$ pode não ser o mais adequado.

Assim, no Algoritmo seguinte (NV 3.2), é introduzido um novo parâmetro $\varepsilon_0 > 0$ e, em cada iteração, sempre que o valor de $\|F(x^{(k)} + h^{(k)})\|_2$ for superior a ε_0 , procede-se à otimização do acréscimo $h^{(k)}$ minimizando o valor de $\|F(x^{(k)} + sh^{(k)})\|_2$ em função de $s \in [0, 2]$. Isto é, nestas condições o termo $x^{(k+1)}$ é calculado usando $x^{(k+1)} = x^{(k)} + th^{(k)}$, onde t é o mínimo da expressão $\|F(x^{(k)} + sh^{(k)})\|_2$ para $s \in [0, 2]$.

Algoritmo 3.3.2 (NV 3.2)

- 1: É dada uma aproximação inicial $x^{(0)}$, $\varepsilon_0 > 0$ e uma margem de erro $\varepsilon > 0$.
 - 2: para $k = 0, 1, 2, \dots$ fazer
 - 3: se $\|F(x^{(k)})\|_2 < \varepsilon$ então
 - 4: o método pára.
 - 5: caso contrário
 - 6: $h^{(k)}$ solução de $J_F(x^{(k)})h^{(k)} = -F(x^{(k)})$;
 - 7: se $\|F(x^{(k)} + h^{(k)})\|_2 \leq \varepsilon_0$ então
 - 8: $x^{(k+1)} = x^{(k)} + h^{(k)}$;
 - 9: caso contrário
 - 10: $t = \min_{s \in [0,2]} \|F(x^{(k)} + sh^{(k)})\|_2$;
 - 11: $x^{(k+1)} = x^{(k)} + th^{(k)}$;
 - 12: $X = \text{vec}^{-1}(x^{(k+1)})$ é uma aproximação do solvente.
-

Enquanto que, no Algoritmo anterior apenas se altera o método de Newton Vetorial (NV quando $\|F(x^{(k)} + h^{(k)})\|_2 > \varepsilon_0$, mantendo a versão original se $\|F(x^{(k)} + h^{(k)})\|_2 \leq \varepsilon_0$, no próximo Algoritmo (NV 3.3) são também introduzidas alterações de cálculo, mesmo no caso de o valor $\|F(x^{(k)} + h^{(k)})\|_2$ ser inferior a ε_0 .

Isto é, uma vez determinado o acréscimo $h^{(k)}$, solução do sistema $J_F(x^{(k)})h^{(k)} = -F(x^{(k)})$, para além da otimização já introduzida no Algoritmo NV 3.2 considera-se ainda, sempre que o valor $\|F(x^{(k)} + h^{(k)})\|_2$ seja inferior a ε_0 , um termo da sucessão intermédio $x^{(k,1)}$ (ponto 8 do Algoritmo NV 3.3) e o sistema $J_F(x^{(k)})h^{(k,1)} = -F(x^{(k,1)})$. A solução $h^{(k,1)}$ pode ser obtida a partir da resolução já considerada na determinação de $h^{(k)}$ uma vez que a matriz do sistema é comum a ambos os sistemas.

Algoritmo 3.3.3 (NV 3.3)

- 1: É dada uma aproximação inicial $x^{(0)}$, $\varepsilon_0 > 0$ e uma margem de erro $\varepsilon > 0$.
 - 2: para $k = 0, 1, 2, \dots$ fazer
 - 3: se $\|F(x^{(k)})\|_2 < \varepsilon$ então
 - 4: o método pára.
 - 5: caso contrário
 - 6: $h^{(k)}$ solução de $J_F(x^{(k)})h^{(k)} = -F(x^{(k)})$;
 - 7: se $\|F(x^{(k)} + h^{(k)})\|_2 \leq \varepsilon_0$ então
 - 8: $x^{(k,1)} = x^{(k)} + h^{(k)}$;
 - 9: $h^{(k,1)}$ solução de $J_F(x^{(k)})h^{(k,1)} = -F(x^{(k,1)})$;
 - 10: $x^{(k+1)} = x^{(k,1)} + h^{(k,1)}$;
 - 11: caso contrário
 - 12: $t = \min_{s \in [0,2]} \|F(x^{(k)} + sh^{(k)})\|_2$;
 - 13: $x^{(k+1)} = x^{(k)} + th^{(k)}$;
 - 14: $X = \text{vec}^{-1}(x^{(k+1)})$ é uma aproximação do solvente.
-

Voltando novamente a considerar o polinómio matricial introduzido em (3.3.16) e as aproximações iniciais $x^{(0)} = \text{vec}(iI_3)$, $X^{(0)} = 10iI_3$ e $X^{(0)} = 10^5iI_3$, os Algoritmos agora definidos, NV 3.2 e NV 3.3, apresentam os seguintes resultados expressos nas tabelas 3.5, 3.6 e 3.7,

$\varepsilon = 10^{-9}$		NV	NV3.2	NV3.3
$\varepsilon_0 = 10^{-1}$	k	7	6	6
$X^{(0)} = iI_3$	t	0.203s	0.687s	0.75s

Tabela 3.5: Comparação do Algoritmo NV com os Algoritmos NV 3.2 e NV 3.3.

Destaca-se o fato de os Algoritmos NV 3.2 e NV 3.3 terem um comportamento idêntico ao dos

$\varepsilon = 10^{-9}$		<i>NV</i>	<i>NV3.2</i>	<i>NV3.3</i>
$\varepsilon_0 = 10^{-1}$		<i>k</i>	7	5
$X^{(0)} = 10iI_3$		<i>t</i>	0.187s	0.562s

Tabela 3.6: Comparação do Algoritmo NV com os Algoritmos NV 3.2 e NV 3.3.

$\varepsilon = 10^{-9}$		<i>NV</i>	<i>NV3.2</i>	<i>NV3.3</i>
$\varepsilon_0 = 10^{-1}$		<i>k</i>	20	6
$X^{(0)} = 10^5iI_3$		<i>t</i>	0.36s	0.797s

Tabela 3.7: Comparação do Algoritmo NV com os Algoritmos NV 3.2 e NV 3.3.

métodos (3.2) e (3.3), em termos de iterações necessárias e uma redução de tempo gasto no cálculo da aproximação (ver tabela (3.8)).

$\varepsilon = 10^{-9}$		<i>NV</i>	<i>NV3.2</i>	<i>NV3.3</i>	<i>3.2</i>	<i>3.3</i>
$\varepsilon_0 = 10^{-1}$		<i>k</i>	20	6	6	7
$X^{(0)} = 10^5iI_3$		<i>t</i>	0.36s	0.797s	0.718s	1.203s

Tabela 3.8: Comparação dos Algoritmos NV, NV 3.2 e NV 3.3 com os Algoritmos 3.2 e 3.3.

A seguir são apresentados exemplos onde se faz a comparação, para além dos Algoritmos NV 3.2 e NV 3.3 e os Algoritmos 3.2 e 3.3, do Algoritmo principal NV com o Algoritmo A.1.1, Anexo A.

3.3.12 Exemplos

No exemplo seguinte faz-se a comparação do Algoritmo principal do método de Newton Vetorial, Algoritmo NV, com o Algoritmo A.1.1, Anexo A, relativamente a um caso extremo onde o Algoritmo A.1.1 não converge, devido ao mau condicionamento das primeiras iterações na equação de Silvester.

Exemplo 3.3.13. Considere-se o polinómio matricial

$$P(X) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} X^2 + \begin{bmatrix} -0.15 & 0.075 \\ 0.01 & -0.355 \end{bmatrix} X + \begin{bmatrix} 6.13333 & -9.46667 \\ -2.73333 & 33.0333 \end{bmatrix}.$$

Para a aproximação inicial

$$x^{(0)} = \text{vec} \left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \right), \text{ com } \|F(x^{(0)})\|_2 = 34.6392,$$

obtém-se, para $k = 127$,

$$X^{(127)} = \text{vec}^{-1}(x^{(127)}) = \begin{bmatrix} -259.901 & 2646.62 \\ -25.539 & 260.048 \end{bmatrix},$$

com $\|F(x^{(127)})\|_2 = 1.89842 \times 10^{-6}$.

Neste caso, usando o método de Newton com derivada de Fréchet definido na secção (A.1.1), para a mesma aproximação inicial $X^{(0)} = \text{vec}^{-1}(x^{(0)})$, obtém-se

$$X^{(200)} = \begin{bmatrix} -1500.84 & -513.912 \\ 4393.14 & 1504.32 \end{bmatrix},$$

com $\|P(X^{(200)})\|_2 = 1333.86$, tendo esta discrepância origem no mau condicionamento das primeiras iterações na equação de Silvester uma vez que, para outras aproximações iniciais, ambos os métodos conduzem a resultados semelhantes: para $X^{(0)} = A_0$, obtém-se, para ambos os métodos,

$$X^{(36)} = \text{vec}^{-1}(x^{(36)}) = \begin{bmatrix} -259.901 & 2646.62 \\ -25.539 & 260.048 \end{bmatrix},$$

com $\|P(X^{(36)})\|_2 = 3.05771 \times 10^{-7}$.

Relativamente aos métodos 2.1, 3.1, 3.2 e 3.3, definidos em Long [26], pág. 650, bem como aos Algoritmos 3.3.2 e 3.3.3, agora definidos, para a aproximação inicial

$$x^{(0)} = \text{vec} \left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \right), \text{ com } \|F(x^{(0)})\|_2 = 34.6392,$$

tão pouco convergem pois, para cada iteração k , o valor de t que minimiza a expressão

$$t = \min_{s \in [0,2]} \|P(X_k - sE_k)\| \approx 0.$$

Considerando $X^{(0)} = A_0$, os métodos referidos conduzem, ainda pelo mesmo motivo, a resultados semelhantes (divergentes), apenas para $X^{(0)} = iA_0$ se obtém a convergência destes, mas para um outro solvente,

$$X^{(5)} = \begin{bmatrix} 0.0758396 + 2.39461i, -0.0380847 - 1.16685i \\ -0.00721708 - 0.337056i, 0.17666 + 5.71038i \end{bmatrix}.$$

A seguir considera-se um exemplo onde se faz a comparação do Algoritmo principal do método de Newton Vetorial, Algoritmo 3.3.1, com os Algoritmos 3.3.2 e 3.3.3, para a mesma aproximação inicial e fixando o número de iterações.

Exemplo 3.3.14. Considere-se o polinômio matricial

$$P(X) = \begin{bmatrix} 17.6 & 1.28i & 2.89 \\ 1.28 & 0.84 & 0.413 \\ 2.89 & 0.413 & 0.725 \end{bmatrix} X^4 + \begin{bmatrix} 7.66 & 2.45 & 2.1 \\ 0.23 & 1.04 & 0.223 \\ 0.6 & 0.756 & 0.658 \end{bmatrix} X^3 + \begin{bmatrix} 121. & 18.9 & 15.9 \\ 0. & 2.7 & 0.145 \\ 11.9 & 3.64 & 15.5 \end{bmatrix} X^2 + \\ + \begin{bmatrix} -36 & -348 & -2 \\ -174 & -558 & -0.2 \\ -2 & -0.2 & -1 \end{bmatrix} X + \begin{bmatrix} -20 & -50 & -10 \\ -30 & -39.19 & -1 \\ -10 & -1 & -50 \end{bmatrix},$$

e a aproximação inicial

$$x^{(0)} = \text{vec} \left(\begin{bmatrix} 100i & 0 & 0 \\ 0 & 100i & 0 \\ 0 & 0 & 100i \end{bmatrix} \right), \text{ com } \|F(x^{(0)})\|_2 = 1.82 \times 10^9.$$

Relativamente aos Algoritmos 3.3.1, 3.3.2 e 3.3.3, fixando o número de iterações em $k = 11$, obtêm-se as seguintes aproximações do solvente de $P(X)$:

– Algoritmo 3.3.1

$$X^{(11)} = \text{vec}^{-1}(x^{(11)}) = \begin{bmatrix} -0.342175 + 5.59429i & -0.536225 + 2.55359i & -0.173195 - 0.504011i \\ -1.6514 + 1.36415i & -5.28056 + 9.16863i & -0.434563 - 0.61365i \\ 2.02363 - 5.29662i & 4.54403 - 12.4746i & 0.490706 + 7.82669i \end{bmatrix},$$

com $\|F(x^{(11)})\|_2 = 5360.47$;

– Algoritmo 3.3.2

$$X^{(11)} = \text{vec}^{-1}(x^{(11)}) = \begin{bmatrix} 0.0607777 + 3.70645i & -0.340695 + 2.64189i & -0.0970939 - 0.794576i \\ -1.63195 + 2.01579i & -5.30708 + 9.17079i & -0.427085 - 0.500204i \\ 0.641935 - 5.59857i & 4.10972 - 12.467i & 0.120044 + 7.71921i \end{bmatrix},$$

com $\|F(x^{(11)})\|_2 = 3.1864 \times 10^{-10}$;

– Algoritmo 3.3.3

$$X^{(11)} = \text{vec}^{-1}(x^{(11)}) = \begin{bmatrix} 0.0607777 + 3.70645i & -0.340695 + 2.64189i & -0.0970939 - 0.794576i \\ -1.63195 + 2.01579i & -5.30708 + 9.17079i & -0.427085 - 0.500204i \\ 0.641935 - 5.59857i & 4.10972 - 12.467i & 0.120044 + 7.71921i \end{bmatrix},$$

com $\|F(x^{(11)})\|_2 = 1.71621 \times 10^{-11}$.

Na tabela (3.9) apresentam-se os resultados obtidos com os Algoritmos 3.3.1, 3.3.2 e 3.3.3 em termos do tempo de CPU t , gasto no cálculo da aproximação $x^{(11)}$ (medido em *segundos* através do comando `Timing[]` do Mathematica) e do erro $\|F(x^{(11)})\|_2$.

$k = 11$		<i>NV</i>	<i>NV3.2</i>	<i>NV3.3</i>
$\varepsilon_0 = 10^{-1}$	t	3.766s	15.875s	17.485s
$X^{(0)} = 10^5 i I_3$	$\ F(x^{(11)})\ _2$	5360.47	3.19×10^{-10}	1.72×10^{-11}

Tabela 3.9: Comparação dos Algoritmos 3.3.1, 3.3.2 e 3.3.3

Como seria de esperar, em 11 iterações os Algoritmos 3.3.2 e 3.3.3 conseguem calcular uma aproximação muito boa (com $\|F(x^{(11)})\|_2 < \frac{1}{2} \times 10^{-9}$) enquanto que o Algoritmo principal, 3.3.1, não apresenta um resultado aceitável ($\|F(x^{(11)})\|_2 > 5000$). Para se obter uma aproximação com $\|F(x^{(k)})\|_2 < \frac{1}{2} \times 10^{-9}$, são necessárias 18 iterações do Algoritmo principal. Apesar de convergirem mais depressa, os Algoritmos 3.3.2 e 3.3.3, consomem mais tempo de CPU nesses cálculos (quase o quádruplo do tempo gasto pelo Algoritmo principal).

Considerando $X^{(0)} = I_3$, os métodos referidos conduzem a resultados semelhantes, tal como consta na tabela (3.10).

$k = 5$		<i>NV</i>	<i>NV3.2</i>	<i>NV3.3</i>
$\varepsilon_0 = 10^{-1}$	t	1.89s	6.812s	7.469s
$X^{(0)} = I_3$	$\ F(x^{(11)})\ _2$	0.006056	3.804×10^{-7}	3.988×10^{-11}

Tabela 3.10: Comparação dos Algoritmos 3.3.1, 3.3.2 e 3.3.3

Neste caso, para se obter uma aproximação com $\|F(x^{(k)})\|_2 < \frac{1}{2} \times 10^{-9}$ são necessárias apenas 17 iterações com Algoritmo principal.

Capítulo 4

Cálculo de Blocos Valores Próprios

De acordo com a Observação 2.4.9, dado X um bloco valor próprio da matriz companheira $\mathcal{C}V = VX$ de dimensão $n \times n$, se o bloco $V_1 \in \mathbb{C}^{n \times n}$ de $V \in \mathbb{C}^{m \times n}$, for não singular, então

$$Y = V_1 X (V_1)^{-1}$$

é um solvente do polinómio $P(X)$. Sendo o assunto principal deste trabalho o cálculo de solventes, um bloco valor próprio nas condições referidas é também, deste ponto de vista, uma alternativa de cálculo de solventes.

De uma forma geral, uma das maneiras de se calcular um bloco valor próprio e um bloco vetor próprio (Definição 2.4.8) de uma matriz $A \in \mathbb{C}^{m \times m}$ é através da resolução da equação

$$AV = VX, \quad (4.0.1)$$

onde $V \in \mathbb{C}^{m \times n}$ é uma matriz de característica máxima e $X \in \mathbb{C}^{n \times n}$, que corresponde a um sistema não linear de $m \times n$ equações ($AV \in \mathbb{C}^{m \times n}$) com $n^2 + m \times n$ variáveis,

$$X = [x_{i,j}]_{i,j=1,\dots,n} \quad \text{e} \quad V = [v_{l,j}]_{\substack{j=1,\dots,n \\ l=1,\dots,m}}$$

Nestas condições, tendo em conta a informação espectral da matriz A (ou, se for esse o caso, da λ -Matriz $P(\lambda)$, tal como referido atrás, Definição 2.3.3), a resolução do sistema (4.0.1) só é possível fixando ou o bloco valor próprio X ou o bloco vetor próprio V .

Deste ponto de vista, um dos métodos iterativos com aplicação no cálculo de blocos valores próprios da equação (4.0.1) é o método da Potência (ver Dennis [9], pág. 83), Algoritmo A.1.4.

Na secção seguinte é feita a construção de um método alternativo ao método da Potência designado por método Newton Vetorial para Blocos Valores Próprios, Algoritmo 4.1.1, e que será posteriormente generalizado para Blocos Valores Próprios de um feixe (A, B) .

4.1 Método Newton Vetorial para Blocos Valores Próprios

Primeiramente, relativamente à equação (4.0.1), consideram-se bloco vetores normalizados na forma,

$$\bar{V} = \begin{bmatrix} I_n \\ V_2 \end{bmatrix}, \quad (4.1.1)$$

com $V_2 \in \mathbb{C}^{(m-n) \times n}$.

Assim, a equação correspondente,

$$A\bar{V} - \bar{V}X = 0, \quad (4.1.2)$$

é um sistema não linear com $m \times n$ equações ($A\bar{V} \in \mathbb{C}^{m \times n}$) e $m \times n$ variáveis, uma vez que, n^2

variáveis de \bar{V} estão fixas ($V_1 = I_n$).

Observação 4.1.1. Considerando bloco vetores normalizados (tal como em (4.1.1)) em certa medida, corresponde a considerar um caso particular de blocos valores próprios não englobando portanto todas as possibilidades que constam na Definição 2.4.8 (um bloco vetor próprio tem n linhas linearmente independentes mas não têm que ser necessariamente as primeiras). Por exemplo, o bloco vetor próprio

$$V = \begin{bmatrix} 2 & 0 \\ 1 & 0 \\ 2 & 2 \\ 0 & 1 \end{bmatrix},$$

tem característica 2 mas as duas primeiras linhas são linearmente dependentes.

Nas condições consideradas na equação (4.1.2), adaptando o método de Newton Vetorial, secção 3.3, a este problema, obtém-se o seguinte Algoritmo:

4.1.2 Algoritmo do Método

Após as adaptações efetuadas ao Algoritmo 3.3.1, NV, que resultam apenas da função $F := \text{vec}(A\bar{V} - \bar{V}X)$ agora considerada e cuja construção será desenvolvida na secção 4.1.3, considera-se o seguinte Algoritmo.

Algoritmo 4.1.1 (NVBVP)

- 1: Dados $\varepsilon > 0$, A , $\bar{V}^{(0)}$ e $X^{(0)}$.
 - 2: para $k = 0, 1, \dots$ fazer
 - 3: $z^{(k)} = \text{vec}(A\bar{V}^{(k)} - \bar{V}^{(k)}X^{(k)})$;
 - 4: se $\|z^{(k)}\|_2 < \varepsilon$ então
 - 5: o método pára.
 - 6: caso contrário
 - 7: $(\tilde{h}^{(k)}, \lambda^{(k)})$, solução de $(I_n \otimes A - (X^{(k)})^T \otimes I_m) \tilde{h}^{(k)} - (I_n \otimes \bar{V}^{(k)}) \lambda^{(k)} = -z^{(k)}$;
 - 8: $(\tilde{H}^{(k)}, \Lambda^{(k)}) = \text{vec}^{-1}(\tilde{h}^{(k)}, \lambda^{(k)})$;
 - 9: $(\bar{V}^{(k+1)}, X^{(k+1)}) = (\bar{V}^{(k)}, X^{(k)}) + (\tilde{H}^{(k)}, \Lambda^{(k)})$
 - 10: $\bar{V} = \bar{V}^{(k)}$, $X = X^{(k)}$ tais que $A\bar{V} = \bar{V}X$.
-

A convergência deste método pode ser obtida pelo Teorema 3.3.9 aplicado à função F referida (definida na equação (4.1.6)), com as adaptações resultantes da construção que se apresenta a seguir.

4.1.3 Convergência

No sentido de serem estabelecidas as condições de convergência do Algoritmo anterior, começa-se por descrever essa construção, fazendo-se posteriormente a adaptação do Teorema 3.3.9.

Observação 4.1.4. Considerando a projeção $p_1 : \mathbb{C}^{m \times n} \rightarrow \mathbb{C}^{(m-n) \times n}$ definida por

$$p_1 \left(\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \right) = V_2,$$

a restrição desta ao conjunto $\{I_n\} \times \mathbb{C}^{(m-n) \times n}$ é uma bijeção sobre $\mathbb{C}^{(m-n) \times n}$ com inversa p_1^{-1}

definida por $p_1^{-1}(V_2) = \bar{V}$. Na realidade é um difeomorfismo e

$$Dp_1^{-1}(V_2) : \mathbb{C}^{(m-n) \times n} \longleftrightarrow \{0_{n \times n}\} \times \mathbb{C}^{(m-n) \times n},$$

onde $Dp_1^{-1}(V_2)[H_2] = \tilde{H} \in \{0_{n \times n}\} \times \mathbb{C}^{(m-n) \times n}$.

Deste ponto de vista, pode entender-se a função $G(\bar{V}, X) = A\bar{V} - \bar{V}X$ como definida em $\mathbb{C}^{(m-n) \times n} \times \mathbb{C}^{n \times n}$, uma vez que n^2 variáveis de \bar{V} estão fixas ($V_1 = I_n$).

Isto é,

$$\begin{array}{ccc} \mathbb{C}^{(m-n) \times n} \times \mathbb{C}^{n \times n} & \xrightarrow{L} & \{I_n\} \times \mathbb{C}^{(m-n) \times n} \times \mathbb{C}^{n \times n} & \xrightarrow{G} & \mathbb{C}^{m \times n} \\ (V_2, X) & \longrightarrow & \underbrace{(I_n, V_2, X)}_{=\bar{V}} & \rightarrow & A\bar{V} - \bar{V}X \end{array}, \quad (4.1.3)$$

onde $L = (p_1^{-1} \times id)$, id a função identidade em $\mathbb{C}^{n \times n}$.

Designando a função composta $G \circ L$ por $\tilde{F} = G \circ L$, tem-se

$$\tilde{F}(V_2, X) = G \circ L(V_2, X) = G(\underbrace{p_1^{-1}[V_2]}_{=\bar{V}}, id(X)) = G(\bar{V}, X), \quad (4.1.4)$$

e, derivando ambos os membros de (4.1.4), no ponto (V_2, X) e na direção de (H_2, Λ) , tem-se

$$\tilde{F}'(V_2, X)[H_2, \Lambda] = G'(\bar{V}, X)[\tilde{H}, \Lambda] \quad (4.1.5)$$

onde $\tilde{H} \in \mathbb{C}^{m \times n}$ é uma matriz da forma

$$\tilde{H} = \begin{bmatrix} 0 \\ H_2 \end{bmatrix},$$

e $\Lambda \in \mathbb{C}^{n \times n}$.

Considerando a composição $F = vec \circ \tilde{F} \circ vec^{-1}$, como no diagrama seguinte,

$$\begin{array}{ccc} \mathbb{C}^{(m-n) \times n} \times \mathbb{C}^{n \times n} & \xrightarrow{\tilde{F}} & \mathbb{C}^{m \times n} \\ \uparrow vec^{-1} & & \downarrow vec \\ \mathbb{C}^{(m-n)n} \times \mathbb{C}^{n^2} & \xrightarrow{F} & \mathbb{C}^{mn} \end{array}$$

$(v_2, x) \xrightarrow{\tilde{F}} A\bar{V} - \bar{V}X$
 $(v_2, x) \xrightarrow{vec^{-1}} \mathbb{C}^{(m-n)n} \times \mathbb{C}^{n^2}$
 $vec(A\bar{V} - \bar{V}X) \xrightarrow{vec} \mathbb{C}^{mn}$

onde $v_2 = vec(V_2)$, $x = vec(X)$, as soluções da equação (4.1.2) ficam caracterizadas em função dos zeros de F ,

$$F : \begin{array}{ccc} \mathbb{C}^{mn} & \longrightarrow & \mathbb{C}^{mn} \\ (v_2, x) & \longrightarrow & vec(A\bar{V} - \bar{V}X). \end{array}$$

Dito de outra forma,

$$F(vec(V_2, X)) = vec(\tilde{F}(V_2, X)) = vec(A\bar{V} - \bar{V}X), \quad (4.1.6)$$

isto é, o diagrama,

$$\begin{array}{ccc} \mathbb{C}^{(m-n) \times n} \times \mathbb{C}^{n \times n} & \xrightarrow{\tilde{F}} & \mathbb{C}^{m \times n} \\ \downarrow \text{vec} & & \downarrow \text{vec} \\ \mathbb{C}^{(m-n)n} \times \mathbb{C}^{n^2} & \xrightarrow{F} & \mathbb{C}^{mn} \end{array},$$

é comutativo.

Esquemáticamente, derivando ambos os membros de (4.1.6) em (V_2, X) na direcção de (H_2, Λ) , tem-se,

$$\begin{array}{ccc} \mathbb{C}^{(m-n) \times n} \times \mathbb{C}^{n \times n} & \xrightarrow{\tilde{F}'(V_2, X)} & \mathbb{C}^{m \times n} \\ \downarrow \text{vec} & & \downarrow \text{vec} \\ \mathbb{C}^{(m-n)n} \times \mathbb{C}^{n^2} & \xrightarrow{F'(v_2, x)} & \mathbb{C}^{mn} \end{array},$$

com

$$v_2 = \text{vec}(V_2), x = \text{vec}(X), h_2 = \text{vec}(H_2), \tilde{h} = \text{vec}(\tilde{H}), \bar{v} = \text{vec}(\bar{V}) \text{ e } \lambda = \text{vec}(\Lambda).$$

Atendendo ao lema 2.1.19, obtém-se

$$\begin{aligned} (F \circ \text{vec})'_{(V_2, X)}[(H_2, \Lambda)] &= \text{vec} \left(\tilde{F}'(V_2, X)[H_2, \Lambda] \right) \\ F'_{\text{vec}(V_2, X)}[\text{vec}(H_2, \Lambda)] &= \text{vec} \left(G'(\bar{V}, X)[\tilde{H}, \Lambda] \right) \\ F'_{(v_2, x)}[h_2, \lambda] &= \text{vec} \left(A\tilde{H} - \tilde{H}X - \bar{V}\Lambda \right) \\ J_{F(v_2, x)}(h_2, \lambda) &= (I_n \otimes A - X^T \otimes I_m) \text{vec}(\tilde{H}) - (I_n \otimes \bar{V}) \text{vec}(\Lambda). \end{aligned}$$

Isto é,

$$J_{F(v_2, x)}(h_2, \lambda) = (I_n \otimes A - X^T \otimes I_m) \tilde{h} - (I_n \otimes \bar{V}) \lambda. \quad (4.1.7)$$

Esta construção, proporciona a aplicação do método de Newton à função $F : \mathbb{C}^{mn} \rightarrow \mathbb{C}^{mn}$,

$$\left(v_2^{(k+1)}, x^{(k+1)} \right) = \left(v_2^{(k)}, x^{(k)} \right) + \left(h_2^{(k)}, \lambda^{(k)} \right), \quad k = 1, 2, \dots \quad (4.1.8)$$

onde $(h_2^{(k)}, \lambda^{(k)})$, é solução do sistema,

$$J_{F(v_2^{(k)}, x^{(k)})}(h_2^{(k)}, \lambda^{(k)}) = -F(v_2^{(k)}, x^{(k)}). \quad (4.1.9)$$

Atendendo às igualdades (4.1.6) e (4.1.7) e considerando $\bar{v} = \text{vec}(\bar{V})$, segue-se que

$$\left(\bar{v}^{(k+1)}, x^{(k+1)} \right) = \left(\bar{v}^{(k)}, x^{(k)} \right) + \left(\tilde{h}^{(k)}, \lambda^{(k)} \right), \quad k = 1, 2, \dots \quad (4.1.10)$$

onde $(\tilde{h}^{(k)}, \lambda^{(k)})$ é solução do sistema

$$\left(I_n \otimes A - \left(X^{(k)} \right)^T \otimes I_m \right) \tilde{h}^{(k)} - \left(I_n \otimes \bar{V}^{(k)} \right) \lambda^{(k)} = -\text{vec} \left(A \bar{V}^{(k)} - \bar{V}^{(k)} X^{(k)} \right). \quad (4.1.11)$$

O Algoritmo 4.1.1 resulta da aplicação do método de Newton à função F , definida em (4.1.6).

A função $F : \mathbb{C}^{(m-n)n} \times \mathbb{C}^{n^2} \rightarrow \mathbb{C}^{(m-n)n} \times \mathbb{C}^{n^2}$, é diferenciável em $\mathbb{C}^{(m-n)n} \times \mathbb{C}^{n^2}$. Se, em

$$B_r(v_2^{(0)}, x^{(0)}) = \{(v_2, x) \in \mathbb{C}^{(m-n)n} \times \mathbb{C}^{n^2} : \|(v_2, x) - (v_2^{(0)}, x^{(0)})\|_2 \leq r\},$$

são satisfeitas as condições, de 1. a 4., do Teorema 3.3.9, então a sucessão

$$(v_2^{(k+1)}, x^{(k+1)}) = (v_2^{(k)}, x^{(k)}) - J_F^{-1}(v_2^{(k)}, x^{(k)}) F(v_2^{(k)}, x^{(k)}), \quad (4.1.12)$$

está bem definida para cada $k = 1, 2, \dots$ e converge para uma solução (v_2, x) de $F(v_2, x) = 0$.

Os próximos resultados visam precisamente assegurar que, nas condições de convergência do Teorema 3.3.9, o Algoritmo 4.1.1 está bem definido e gera uma sucessão que converge para um bloco valor próprio e bloco vetor próprio da matriz A .

De seguida faz-se a transposição do resultado em termos dos zeros da função F para blocos valores próprios e blocos vetores próprios da matriz A , soluções da equação $A\bar{V} = \bar{V}X$.

Observação 4.1.5. Se $(v_2, x) \in \mathbb{C}^{(m-n)n} \times \mathbb{C}^{n^2}$ é um zero de F , então $A\bar{V} = \bar{V}X$, onde,

$$\bar{V} = \begin{bmatrix} I_n \\ \text{vec}^{-1}(v_2) \end{bmatrix}, \quad X = \begin{bmatrix} \text{vec}^{-1}(x) \end{bmatrix}.$$

(Por construção) Seja $(V_2, X) = \text{vec}^{-1}(v_2, x) \in \mathbb{C}^{(m-n) \times n} \times \mathbb{C}^{n \times n}$. Se (v_2, x) é um zero de F , então, por construção (4.1.6), $\tilde{F}(V_2, X) = 0$, e tendo em conta (4.1.4), então $\tilde{F}(V_2, X) = G(\underbrace{p_1^{-1}[V_2]}_{=\bar{V}}, id(X)) = G(\bar{V}, X)$, isto é, $\text{vec}(A\bar{V} - \bar{V}X) = 0$.

A seguir é feita a adaptação do Teorema 3.3.9 à função F agora considerada, de modo a garantir a convergência do Algoritmo 4.1.1 para um zero de F .

Teorema 4.1.6. Se a função F , para uma aproximação inicial $(v_2^{(0)}, x^{(0)}) \in \mathbb{C}^{(m-n)n} \times \mathbb{C}^{n^2}$, satisfaz as condições de convergência do Teorema 3.3.9 (Gutierrez [17], pág. 9), então a sucessão calculada nos pontos 3) e 4) do Algoritmo 4.1.1 está bem definida e converge para (\bar{v}, x) , onde (\bar{V}, X) é uma solução da equação $A\bar{V} = \bar{V}X$.

Demonstração. Nas condições do Teorema e atendendo à igualdade (4.1.7), a solução do sistema considerado em 3),

$$\left(I_n \otimes A - \left(X^{(k)} \right)^T \otimes I_m \right) \tilde{h}^{(k)} - \left(I_n \otimes \bar{V}^{(k)} \right) \lambda^{(k)} = -z^{(k)},$$

está definida por $(\tilde{h}^{(k)}, \lambda^{(k)})$, onde $(h_2^{(k)}, \lambda^{(k)}) = -J_F^{-1}(v_2^{(k)}, x^{(k)}) F(v_2^{(k)}, x^{(k)})$ e as restantes coordenadas de $\tilde{h}^{(k)}$ são nulas. Portanto, para cada $k = 1, 2, \dots$, a sucessão $(\bar{v}^{(k+1)}, x^{(k+1)}) = (\bar{v}^{(k)}, x^{(k)}) + (\tilde{h}^{(k)}, \lambda^{(k)})$ está bem definida.

Por outro lado, a sucessão $(v_2^{(k+1)}, x^{(k+1)}) = (v_2^{(k)}, x^{(k)}) + (h_2^{(k)}, \lambda^{(k)})$ converge para $(v_2, x) \in \mathbb{C}^{(m-n)n} \times \mathbb{C}^{n^2}$ um zero da função F e, por conseguinte, a sucessão $(\bar{v}^{(k+1)}, x^{(k+1)})$ converge para

(\bar{v}, x) . Pela Observação 4.1.5, $A\bar{V} = \bar{V}X$. ■

Este resultado aliado à Observação 4.1.5 garante que o Algoritmo 4.1.1 permite o cálculo de um bloco valor próprio e bloco vetor próprio da matriz A .

Assim, na secção seguinte, são apresentados exemplos elucidativos da convergência do Algoritmo 4.1.1 para um bloco valor próprio e bloco vetor próprio da matriz A

4.1.7 Exemplos

Nos exemplos seguintes consideram-se aproximações iniciais do tipo,

$$X^{(0)} = \alpha I_n, \quad (4.1.13)$$

um múltiplo da matriz identidade em $\mathbb{C}^{n \times n}$, e

$$\bar{V}^{(0)} = \begin{bmatrix} I_n \\ \beta mn U \end{bmatrix}, \quad (4.1.14)$$

onde a matriz U tem todas as entradas iguais a 1 em $\mathbb{C}^{(m-n) \times n}$, m e n são as dimensões da matriz A e do bloco valor próprio X , respetivamente, $\beta = 1.13mn$ e α igual ao raio espectral¹ da matriz A .

Relativamente a cada exemplo, faz-se ainda a comparação deste método, Algoritmo 4.1.1, com o método da Potência (Algoritmo A.1.4). São também apresentados os espectros, quer da matriz A , quer do bloco valor próprio X calculado através do Algoritmo 4.1.1, verificando-se $\Lambda(X) \subset \Lambda(A)$, uma vez que é uma condição necessária (mas não suficiente, ver Teorema 1 em Pereira [31], pág. 47) para que X seja um bloco valor próprio de A .

Neste primeiro exemplo são ainda apresentados de forma detalhada os cálculos que permitem a determinação da iteração seguinte, nomeadamente os pontos 7, 8 e 9 do respetivo Algoritmo 4.1.1.

Exemplo 4.1.8. Considere-se a matriz

$$A = \begin{bmatrix} -22 & -86 & -6 & 50 & 8 & -6 \\ 43 & 107 & -25 & -81 & 3 & 17 \\ -434 & -1330 & 80 & 800 & 40 & -100 \\ 665 & 1561 & -400 & -1120 & 50 & 190 \\ -5180 & -14140 & 1826 & 8770 & 100 & -1140 \\ 7070 & 16030 & -4385 & -11329 & 570 & 1810 \end{bmatrix},$$

e, para $n = 2$, as aproximações iniciais, com $\beta = 1.13mn$ e $\alpha = 512$,

$$\bar{V}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 13.56 & 13.56 \\ 13.56 & 13.56 \\ 13.56 & 13.56 \\ 13.56 & 13.56 \end{bmatrix} \quad \text{e} \quad X^{(0)} = \begin{bmatrix} 512 & 0 \\ 0 & 512 \end{bmatrix}.$$

¹Sem que constitua um critério rigoroso que assegure à partida a convergência, nos exemplos considerados esta escolha revelou-se eficiente quando comparada com outras aproximações iniciais para as quais o método diverge.

Tem-se que, o sistema (4.1.11), no ponto $vec(V^{(0)}, X^{(0)})$, está definido por

$$\begin{bmatrix} -6H_{3,1} + 50H_{4,1} + 8H_{5,1} - 6H_{6,1} - \Lambda_{1,1} \\ -25H_{3,1} - 81H_{4,1} + 3H_{5,1} + 17H_{6,1} - \lambda_{2,1} \\ -432H_{3,1} + 800H_{4,1} + 40H_{5,1} - 100H_{6,1} - 13.56\Lambda_{1,1} - 13.56\Lambda_{2,1} \\ -400H_{3,1} - 1632H_{4,1} + 50H_{5,1} + 190H_{6,1} - 13.56\Lambda_{1,1} - 13.56\Lambda_{2,1} \\ 1826H_{3,1} + 8770H_{4,1} - 412H_{5,1} - 1140H_{6,1} - 13.56\Lambda_{1,1} - 13.56\Lambda_{2,1} \\ -4385H_{3,1} - 11329H_{4,1} + 570H_{5,1} + 1298H_{6,1} - 13.56\Lambda_{1,1} - 13.56\Lambda_{2,1} \\ -6H_{3,2} + 50H_{4,2} + 8H_{5,2} - 6H_{6,2} - \Lambda_{1,2} \\ -25H_{3,2} - 81H_{4,2} + 3H_{5,2} + 17H_{6,2} - \Lambda_{2,2} \\ -432H_{3,2} + 800H_{4,2} + 40H_{5,2} - 100H_{6,2} - 13.56\Lambda_{1,2} - 13.56\Lambda_{2,2} \\ -400H_{3,2} - 1632H_{4,2} + 50H_{5,2} + 190H_{6,2} - 13.56\Lambda_{1,2} - 13.56\Lambda_{2,2} \\ 1826H_{3,2} + 8770H_{4,2} - 412H_{5,2} - 1140H_{6,2} - 13.56\Lambda_{1,2} - 13.56\Lambda_{2,2} \\ -4385H_{3,2} - 11329H_{4,2} + 570H_{5,2} + 1298H_{6,2} - 13.56\Lambda_{1,2} - 13.56\Lambda_{2,2} \end{bmatrix} = \begin{bmatrix} -89.76 \\ 1123.16 \\ -3742.48 \\ 23634.5 \\ -117457 \\ 180682 \\ -537.76 \\ 1571.16 \\ -2846.48 \\ 22738.5 \\ -108497 \\ 171722 \end{bmatrix},$$

cuja solução é o vetor

$$vec(H_2, \Lambda) = vec \left(\begin{bmatrix} H_{3,1} & H_{3,2} \\ H_{4,1} & H_{4,2} \\ H_{5,1} & H_{5,2} \\ H_{6,1} & H_{6,2} \end{bmatrix}, \begin{bmatrix} \Lambda_{1,1} & \Lambda_{1,2} \\ \Lambda_{2,1} & \Lambda_{2,2} \end{bmatrix} \right) = \begin{bmatrix} H_{3,1} \\ H_{4,1} \\ H_{5,1} \\ H_{6,1} \\ H_{3,2} \\ H_{4,2} \\ H_{5,2} \\ H_{6,2} \\ \Lambda_{1,1} \\ \Lambda_{2,1} \\ \Lambda_{1,2} \\ \Lambda_{2,2} \end{bmatrix} = \begin{bmatrix} -51.5668 \\ 27.9737 \\ -307.224 \\ 345.281 \\ -59.5668 \\ 35.9737 \\ -371.224 \\ 409.281 \\ -2731.63 \\ 2848.24 \\ -2731.63 \\ 2848.24 \end{bmatrix}.$$

Assim a aproximação seguinte está definida por

$$\bar{V}^{(1)} = \bar{V}^{(0)} + \tilde{H}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 13.56 & 13.56 \\ 13.56 & 13.56 \\ 13.56 & 13.56 \\ 13.56 & 13.56 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ H_{3,1} & H_{3,2} \\ H_{4,1} & H_{4,2} \\ H_{5,1} & H_{5,2} \\ H_{6,1} & H_{6,2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -38.0068 & -46.0068 \\ 41.5337 & 49.5337 \\ -293.664 & -357.664 \\ 358.841 & 422.841 \end{bmatrix},$$

e

$$X^{(1)} = X^{(0)} + \Lambda^{(0)} = \begin{bmatrix} 512 & 0 \\ 0 & 512 \end{bmatrix} + \begin{bmatrix} \Lambda_{1,1} & \Lambda_{1,2} \\ \Lambda_{2,1} & \Lambda_{2,2} \end{bmatrix} = \begin{bmatrix} -2219.63 & -2731.63 \\ 2848.24 & 3360.24 \end{bmatrix}.$$

Após 7 iterações obtém-se uma aproximação do bloco vetor próprio e bloco valor próprio com $\|AV^{(7)} - V^{(7)}X^{(7)}\|_2 < 10^{-5}$,

$$\bar{V}^{(7)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 6 & -2 \\ 1 & 9 \\ 34 & -30 \\ 15 & 79 \end{bmatrix}, X^{(7)} = \begin{bmatrix} 174 & -338 \\ 169 & 681 \end{bmatrix},$$

onde o bloco $X^{(7)}$ é dominante pois $\Lambda(X^{(7)}) = \{512, 343\}$ e $\Lambda(A) = \{512, 343, 64, 27, 8, 1\}$.

Se forem consideradas as aproximações iniciais, com $\beta = mn$ e $\alpha = 1$,

$$\bar{V}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 12 & 12 \\ 12 & 12 \\ 12 & 12 \\ 12 & 12 \end{bmatrix} \text{ e } X^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

após 5 iterações obtém-se uma aproximação do bloco vetor próprio e bloco valor próprio com $\|AV^{(5)} - V^{(5)}X^{(5)}\|_2 < 10^{-5}$,

$$\bar{V}^{(5)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & -2. \\ 1. & 3. \\ -2. & -6. \\ 3. & 7. \end{bmatrix}, X^{(7)} = \begin{bmatrix} -6. & -14. \\ 7. & 15. \end{bmatrix},$$

onde $\Lambda(X^{(5)}) = \{8, 1\}$.

Para outras aproximações iniciais obtém-se diferentes blocos valores próprios: considerando $\beta mn = -10000$ e $\alpha = 1$, obtém-se um bloco valor próprio $X^{(10)}$ com $\Lambda(X^{(10)}) = \{64, 1\}$.

Relativamente a este exemplo e para a aproximação inicial

$$\bar{V}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 13.56 & 13.56 \\ 13.56 & 13.56 \\ 13.56 & 13.56 \\ 13.56 & 13.56 \end{bmatrix},$$

o método da Potência (Algoritmo A.1.4) converge em 11 iterações para o mesmo bloco vetor próprio da matriz A . Quanto ao tempo dispendido no cálculo, o método 4.1.1 apesar de convergir em menos iterações consome mais tempo de CPU, (0.078s) contra (0.016s) do método da Potência.

O próximo exemplo corresponde a um caso em que o método da Potência não converge (na realidade, a matriz A considerada não tem um bloco valor próprio dominante (Definição equivalente a 2.4.6), condição necessária para assegurar a sua convergência (ver Dennis [9], pág. 84) enquanto que o método de Newton Vetorial para Blocos Valores Próprios converge em 8 iterações.

Exemplo 4.1.9. Seja

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.835461 - 0.450355i & 0.31773 + 0.0283688i & -5.71915 + 0.382979i & 0 & 0 & 0 & 0 & 0 & 0 \\ -1.50922 - 0.0070922i & 0.55461 + 0.567376i & -13.583 - 1.34043i & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.635461 + 0.0496454i & 0.31773 + 0.0283688i & -5.91915 - 0.117021i & 0 & 0 & 0 & 0 & 0 & 0 \\ & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 \\ & 2.67021 + 4.50355i & -0.0851064 - 0.283688i & 1.53191 - 3.82979i & 0 & 0 & 0 & 0 & 0 \\ & 0.404255 + 0.070922i & 2.29787 - 5.67376i & 3.6383 + 13.4043i & 0 & 0 & 0 & 0 & 0 \\ & 0.170213 - 0.496454i & -0.0851064 - 0.283688i & 4.03191 + 1.17021i & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Considerando as aproximações iniciais, com $\beta = 1.13mn$ e $\alpha = 5.38516$,

$$\bar{V}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 13.56 & 13.56 \\ 13.56 & 13.56 \\ 13.56 & 13.56 \\ 13.56 & 13.56 \end{bmatrix} \quad \text{e} \quad X^{(0)} = \begin{bmatrix} 5.38516 & 0 \\ 0 & 5.38516 \end{bmatrix},$$

após 8 iterações obtém-se uma aproximação do bloco vetor próprio e do bloco valor próprio com $\|AV^{(8)} - V^{(8)}X^{(8)}\|_2 < 10^{-5}$,

$$\bar{V}^{(8)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -0.111111 & 0.0555556 \\ 0.1 & 0 \\ 0 & 0.1 \\ -0.0111111 & 0.00555556 \end{bmatrix}, \quad X^{(8)} = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix},$$

onde $\Lambda(X^{(8)}) = \{0.5, 0.5\}$ e $\Lambda(A) = \{2 - 5i, 2 - 5i, 2 + 5i, 2 + 5i, 0.5, 0.5\}$.

Relativamente ao método da Potência (Algoritmo A.1.4), para a mesma aproximação

$$V^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 13.56 & 13.56 \\ 13.56 & 13.56 \\ 13.56 & 13.56 \\ 13.56 & 13.56 \end{bmatrix},$$

após 600 iterações obtém-se

$$V^{(600)} = \begin{bmatrix} 1.08556 - 0.0841213i & 0.348377 - 0.250739i \\ 1.30348 - 0.197716i & -0.62027 + 0.0572379i \\ 1 & 0 \\ 0 & 1 \\ -0.88636 + 0.890544i & 1.32035 + 0.141066i \\ -0.118344 - 0.0519071i & 0.60991 - 0.248082i \end{bmatrix},$$

$$X^{(600)} = \begin{bmatrix} -0.591718 - 0.259536i & 3.04955 - 1.24041i \\ -6.29622 + 0.487904i & 1.97941 + 1.45429i \end{bmatrix},$$

onde

$$\Lambda(X^{(600)}) = \{2 + 5i, -0.612306 - 3.80525I\} \not\subseteq \Lambda(A) = \{2 - 5i, 2 - 5i, 2 + 5i, 2 + 5i, 0.5, 0.5\}.$$

4.1.10 Aplicação à matriz \mathcal{C}

Considere-se um polinômio matricial mônico

$$P(X) = X^m + A_1X^{m-1} + \cdots + A_m, \quad (4.1.15)$$

de grau m na variável $X \in \mathbb{C}^{n \times n}$. A matriz companheira é da forma

$$\mathcal{C} = \begin{bmatrix} 0 & I_n & \cdots & 0 \\ 0 & 0 & I_n & \vdots \\ \vdots & \vdots & \ddots & I_n \\ -A_m & -A_{m-1} & \cdots & -A_1 \end{bmatrix}, \quad (4.1.16)$$

que corresponde a uma matriz particionada em m^2 blocos $C_{i,j} \in \mathbb{C}^{n \times n}$, $i, j = 1, \dots, m$, onde $C_{i,i+1} = I_n$, $i = 1, \dots, m-1$, $C_{m,j} = -A_{m-j+1}$, $j = 1, \dots, m$ e as restantes bloco matrizes são nulas.

A equação

$$\mathcal{C}V = VX, \quad (4.1.17)$$

onde $X \in \mathbb{C}^{n \times n}$ e

$$V = \begin{bmatrix} I_n \\ X \\ \vdots \\ X^{m-1} \end{bmatrix}, \quad (4.1.18)$$

é condição necessária e suficiente (Teorema 2.4.1) para que uma matriz X seja um solvente de $P(X) = 0$.

Isto é, X é um solvente de $P(X)$ se e só se X for um bloco valor próprio da matriz companheira onde V , o vetor definido em (4.1.18), é o seu bloco vetor próprio associado. Assim, é possível determinar um solvente de $P(X)$ através do cálculo de um bloco valor próprio da matriz companheira \mathcal{C} , desde

que o bloco vetor próprio associado seja normalizado na forma \bar{V} .

A exemplo do que pode ser feito com o método da Potência, em que se usam os blocos valores próprios para obter solventes, o método Newton Vetorial para Blocos Valores Próprios, quando aplicado à matriz companheira \mathcal{C} , é ainda um método viável no cálculo de solventes de $P(X)$.

De seguida é apresentado um exemplo de utilização do método Newton Vetorial para Blocos Valores Próprios, aplicado à matriz companheira \mathcal{C} de um polinômio de grau 5, com o objetivo de exibir as potencialidades deste método quando aplicado no cálculo de um solvente.

Utilizam-se ainda aproximações iniciais do tipo,

$$X^{(0)} = \alpha I_n, \tag{4.1.19}$$

um múltiplo da matriz identidade em $\mathbb{C}^{n \times n}$, e

$$\bar{V}^{(0)} = \begin{bmatrix} I_n \\ \beta mnU \end{bmatrix}, \tag{4.1.20}$$

onde a matriz U tem todas as entradas iguais a 1 em $\mathbb{C}^{(m-n) \times n}$, n é a dimensão do bloco valor próprio e mn é a dimensão da matriz companheira considerada.

Exemplo 4.1.11. Considere-se o polinômio matricial mônico

$$P(X) = X^5 + A_1X^4 + A_2X^3 + A_3X^2 + A_4X + A_5,$$

com $A_i \in \mathbb{C}^{2 \times 2}$ e seja

$$\mathcal{C} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ -1950 & -5790 & 1006 & 5390 & 100 & -1700 & -120 & 220 & 20 & -10 \\ 2895 & 6735 & -2695 & -7079 & 850 & 2650 & -110 & -450 & 5 & 35 \end{bmatrix},$$

a sua matriz companheira.

Aplicando, para $n = 2$, o método 4.1.1 e considerando $\alpha = 10$ (igual ao raio espectral da matriz \mathcal{C}) e $\beta = 1.3$,

$$\bar{V}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 13 & 13 \\ 13 & 13 \\ 13 & 13 \\ 13 & 13 \\ 13 & 13 \\ 13 & 13 \\ 13 & 13 \\ 13 & 13 \end{bmatrix} \quad \text{e} \quad X^{(0)} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix},$$

obtém-se, após 8 iterações, uma aproximação do bloco vetor próprio e do bloco valor próprio com $\|\mathcal{C}\bar{V}^{(8)} - \bar{V}^{(8)}X^{(8)}\|_2 < 10^{-5}$,

$$\bar{V}^{(8)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 8 & -2 \\ 1 & 11 \\ 62 & -38 \\ 19 & 119 \\ 458 & -542 \\ 271 & 1271 \\ 3122 & -6878 \\ 3439 & 13439 \end{bmatrix} = \begin{bmatrix} I_n \\ X^{(8)} \\ \vdots \\ (X^{(8)})^4 \end{bmatrix}, \quad X^{(8)} = \begin{bmatrix} 8 & -2 \\ 1 & 11 \end{bmatrix},$$

De acordo com Teorema 2.4.1, tem-se

$$P\left(\begin{bmatrix} 8 & -2 \\ 1 & 11 \end{bmatrix}\right) = 0.$$

Além disso, o bloco (e solvente) $X^{(8)}$ é dominante (Definição 2.4.6) pois $\Lambda(X^{(8)}) = \{10, 9\}$ e

$$\Lambda(\mathcal{C}) = \{10, 9, 8, 7, 6, 5, 4, 3, 2, 1\}.$$

4.1.12 Aplicação ao Feixe $\lambda B - A$

Como já referido, um dos possíveis métodos iterativos com aplicação no cálculo de blocos valores próprios da equação $AV = VX$ é o método da Potência, Algoritmo A.1.4. Acontece que este método está definido apenas para um feixe do tipo $(A, I) = \lambda I - A$. Ou seja, relativamente aos polinómios matriciais, este método, Algoritmo A.1.4, apenas está definido para polinómios matriciais mónicos em que a linearização é da forma $\lambda I - \mathcal{C}$.

Se o polinómio $P(X)$ for não mónico,

$$P(X) = A_0X^m + A_1X^{m-1} + \dots + A_m. \quad (4.1.21)$$

pode considerar-se a linearização (2.2.6) (feixe companheiro de $P(\lambda)$),

$$\lambda \mathcal{C}_2 - \mathcal{C}_1.$$

Estende-se a Definição 2.4.8, de bloco valor próprio e bloco vetor próprio de uma matriz A , como soluções da equação $AV - VX = 0$, para bloco valor próprio e bloco vetor próprio de um feixe (A, B) , como sendo as soluções da equação,

$$AV - BVX = 0. \quad (4.1.22)$$

Se $B = I$, esta Definição coincide com a Definição 2.4.8. Isto é, os blocos valores próprios e blocos vetores próprios de uma matriz A são os blocos valores próprios e blocos vetores próprios do feixe (A, I) .

Adaptando a função F , definida em (4.1.4), ao contexto atual e repetindo todo o processo com $G(V, X) = AV - BVX$, tem-se

$$\tilde{F}(V_2, X) = G(\underbrace{p_1^{-1}[V_2]}_{=\bar{V}}, id(X)) = G(\bar{V}, X) = A\bar{V} - B\bar{V}X, \quad (4.1.23)$$

e, derivando ambos os membros de (4.1.23),

$$\tilde{F}'(V_2, X)[H_2, \Lambda] = G'(\bar{V}, X)[\tilde{H}, \Lambda] = A\tilde{H} - B\tilde{H}X - B\bar{V}\Lambda, \quad (4.1.24)$$

onde $\tilde{H} \in \mathbb{C}^{m \times n}$ é uma matriz da forma

$$\tilde{H} = \begin{bmatrix} 0 \\ H_2 \end{bmatrix},$$

e $\Lambda \in \mathbb{C}^{n \times n}$.

Aplicando a função $vec()$ a ambos os membros de (4.1.23), caracterizam-se as soluções da equação (4.1.22) em termos dos zeros função

$$F : \quad \mathbb{C}^{mn^2} \quad \longrightarrow \quad \mathbb{C}^{mn^2} \\ vec(V_2, X) \quad \rightarrow \quad vec(A\bar{V} - B\bar{V}X) = vec(G(\bar{V}, X)) \quad ,$$

com

$$J_{\tilde{F}}(v_2^{(k)}, x^{(k)}) \left(h_2^{(k)}, \lambda^{(k)} \right) = \left(I_n \otimes A - \left(X^{(k)} \right)^T \otimes B \right) \tilde{h}^{(k)} - \left(I_n \otimes B\bar{V}^{(k)} \right) \lambda^{(k)}. \quad (4.1.25)$$

Feita a adaptação da função F , pode agora apresentar-se uma versão do Algoritmo 4.1.1 para a equação (4.1.22).

 Algoritmo 4.1.2 (NVBVP - Polinômios não mônicos)

- 1: Dados $\varepsilon > 0$, A , B , $\bar{V}^{(0)}$ e $X^{(0)}$.
 - 2: para $k = 0, 1, \dots$ fazer
 - 3: $z^{(k)} = \text{vec}(A\bar{V}^{(k)} - B\bar{V}^{(k)}X^{(k)})$;
 - 4: se $\|z^{(k)}\|_2 < \varepsilon$ então
 - 5: o método pára.
 - 6: caso contrário
 - 7: $(\tilde{h}^{(k)}, \lambda^{(k)})$, solução de $(I_n \otimes A - (X^{(k)})^T B)\tilde{h}^{(k)} - (I_n \otimes B\bar{V}^{(k)})\lambda^{(k)} = -z^{(k)}$;
 - 8: $(\tilde{H}^{(k)}, \Lambda^{(k)}) = \text{vec}^{-1}(\tilde{h}^{(k)}, \lambda^{(k)})$;
 - 9: $(\bar{V}^{(k+1)}, X^{(k+1)}) = (\bar{V}^{(k)}, X^{(k)}) + (\tilde{H}^{(k)}, \Lambda^{(k)})$
 - 10: $\bar{V} = \bar{V}^{(k)}$, $X = X^{(k)}$ tais que $A\bar{V} = B\bar{V}X$.
-

O Algoritmo 4.1.2, definido para um feixe genérico (A, B) é também aplicável à linearização $\lambda\mathcal{C}_2 - \mathcal{C}_1$ de um polinômio não mônico (4.1.21).

A seguir são apresentados exemplos onde se utiliza o Algoritmo 4.1.2 quer a um feixe genérico (A, B) quer ao feixe companheiro $\lambda\mathcal{C}_2 - \mathcal{C}_1$ de um dado polinômio.

Consideram-se aproximações iniciais do tipo,

$$X^{(0)} = \alpha I_n, \quad (4.1.26)$$

um múltiplo da matriz identidade em $\mathbb{C}^{n \times n}$, e

$$\bar{V}^{(0)} = \begin{bmatrix} I_n \\ \beta mnU \end{bmatrix}, \quad (4.1.27)$$

onde a matriz U tem todas as entradas iguais a 1 em $\mathbb{C}^{(m-n) \times n}$, n é a dimensão do bloco valor próprio e mn é a dimensão das matrizes consideradas.

São ainda apresentados os espectros, quer do feixe dado (A, B) , quer do bloco valor próprio X calculado através do Algoritmo 4.1.2, verificando-se $\Lambda(X, Y) \subset \Lambda(A, B)$, uma vez que é uma condição necessária (mas não suficiente, ver Pereira [31] e [32], pág. 74 e pág. 2914) para que X seja um bloco valor próprio de (A, B) .

Neste primeiro exemplo, o Algoritmo 4.1.2 é aplicado ao feixe companheiro $\lambda\mathcal{C}_2 - \mathcal{C}_1$ de um polinômio de grau 5 em que a matriz A_0 é singular.

Exemplo 4.1.13. Seja o polinômio matricial não mônico

$$P(X) = \begin{bmatrix} 2 & -10 \\ 4 & -20 \end{bmatrix} X^5 + \begin{bmatrix} -20 & 10 \\ -5 & -35 \end{bmatrix} X^4 + \begin{bmatrix} 120 & -220 \\ 110 & 450 \end{bmatrix} X^3 + \\ + \begin{bmatrix} -100 & 1700 \\ -850 & -2650 \end{bmatrix} X^2 + \begin{bmatrix} -1006 & -5390 \\ 2695 & 7079 \end{bmatrix} X + \begin{bmatrix} 1950 & 5790 \\ -2895 & -6735 \end{bmatrix},$$

e

$$\lambda\mathcal{C}_2 - \mathcal{C}_1,$$

a sua linearização, com

$$\mathcal{C}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & -10 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & -20 & 0 \end{bmatrix}$$

e

$$\mathcal{C}_1 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ -1950 & -5790 & 1006 & 5390 & 100 & -1700 & -120 & 220 & 20 & -10 \\ 2895 & 6735 & -2695 & -7079 & 850 & 2650 & -110 & -450 & 5 & 35 \end{bmatrix}.$$

Aplicando o Algoritmo 4.1.2, com $\beta = 1.13mn$ e $\alpha = 10$,

$$\bar{V}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 22.6 & 22.6 \\ 22.6 & 22.6 \\ 22.6 & 22.6 \\ 22.6 & 22.6 \\ 22.6 & 22.6 \\ 22.6 & 22.6 \\ 22.6 & 22.6 \\ 22.6 & 22.6 \end{bmatrix} \quad \text{e} \quad X^{(0)} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix},$$

obtem-se após 12 iterações uma aproximação do bloco vetor próprio e do bloco valor próprio com $\|\mathcal{C}_1 \bar{V}^{(12)} - \mathcal{C}_2 \bar{V}^{(12)} X^{(12)}\|_2 < 10^{-5}$,

$$\bar{V}^{(12)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1.89157 & 1.96289 \\ 0.199601 & 1.34104 \\ 3.96983 & 6.34527 \\ 0.645231 & 2.19018 \\ 8.77574 & 16.3016 \\ 1.65766 & 4.20364 \\ 19.8537 & 39.0869 \\ 3.97463 & 8.89105 \end{bmatrix} \text{ e } X^{(12)} = \begin{bmatrix} 1.89157 & 1.96289 \\ 0.199601 & 1.34104 \end{bmatrix},$$

com

$$\Lambda(X^{(12)}) = \{2.30009, 0.932517\}$$

e

$$\Lambda(P) = \{\infty, -17.4325, 3.60583 + 1.81647i, 3.60583 - 1.81647i, -0.309128 + 3.29155i, -0.309128 - 3.29155i, \underline{2.30009}, 1.24075 + 0.854468i, 1.24075 - 0.854468i, \underline{0.932517}\}.$$

Tem-se ainda que $X^{(12)}$ é um solvente do polinômio $P(X)$ pois

$$P\left(\begin{bmatrix} 1.89157 & 1.96289 \\ 0.199601 & 1.34104 \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

o que confirma novamente o resultado do Teorema 2.4.1 relativamente a blocos valores próprios do feixe companheiro $\lambda\mathcal{C}_2 - \mathcal{C}_1$.

Chama-se à atenção para o fato de, no exemplo anterior 4.1.13, apesar de a matriz A_0 ser singular, o Algoritmo 4.1.2 converge. No exemplo seguinte considera-se um polinômio em que a matriz A_0 , para além de ser singular, tem uma coluna nula.

Exemplo 4.1.14. Seja o polinômio matricial não mónico

$$P(X) = \begin{bmatrix} 0 & -10 \\ 0 & -20 \end{bmatrix} X^4 + \begin{bmatrix} -2 & 1 \\ -5 & -35 \end{bmatrix} X^3 + \begin{bmatrix} 12 & -22i \\ 11 & 45 \end{bmatrix} X^2 + \begin{bmatrix} -10 & 17 \\ -85 & -26.5 \end{bmatrix} X + \begin{bmatrix} -1006 & -53.9 \\ 2695 & 707.9 \end{bmatrix},$$

e

$$\lambda\mathcal{C}_1 - \mathcal{C}_2,$$

a sua linearização, com

$$\mathcal{C}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -10 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -20 \end{bmatrix} \text{ e } \mathcal{C}_1 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1006 & 53.9 & 10 & -17 & -12 & 22i & 2 & -1 \\ -2695 & -707.9 & 85 & 26.5 & -11 & -45 & 5 & 35 \end{bmatrix}.$$

Aplicando o Algoritmo 4.1.2, com $\beta = 1.13mn$ e $\alpha = 10$,

$$\bar{V}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 18.08 & 18.08 \\ 18.08 & 18.08 \\ 18.08 & 18.08 \\ 18.08 & 18.08 \\ 18.08 & 18.08 \\ 18.08 & 18.08 \end{bmatrix} \text{ e } X^{(0)} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix},$$

obtem-se após 7 iterações uma aproximação dos bloco vetor próprio e bloco valor próprio com $\|\mathcal{C}_1 \bar{V}^{(7)} - \mathcal{C}_2 \bar{V}^{(7)} X^{(7)}\|_2 < 10^{-5}$,

$$\bar{V}^{(7)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 13.4868 - 0.122121i & 1.6654 + 0.488237i \\ -0.137183 - 0.000316732i & 1.83784 - 0.153651i \\ 181.651 - 3.36155i & 25.6563 + 7.0228i \\ -2.10237 + 0.0329774i & 3.12572 - 0.632277i \\ 2445.97 - 68.4915i & 352.394 + 92.0551i \\ -28.7793 + 0.787252i & 2.13003 - 2.61383i \end{bmatrix},$$

e

$$X^{(7)} = \begin{bmatrix} 13.4868 - 0.122121i & 1.6654 + 0.488237i \\ -0.137183 - 0.000316732i & 1.83784 - 0.153651i \end{bmatrix},$$

com

$$\Lambda(X^{(7)}) = \{13.4672 - 0.127882i, 1.85748 - 0.14789i\}$$

e

$$\Lambda(P) = \{\infty, -9.39951 + 16.7307i, -8.97873 - 16.618i, \underline{13.4672 - 0.127882i}, -2.05344 + 0.214975i, \underline{1.85748 - 0.14789i}, -0.0451072 - 1.74532i, -0.347861 + 1.69348i\}.$$

Tem-se ainda que $X^{(12)}$ sendo um bloco valor próprio do feixe companheiro $\lambda\mathcal{C}_2 - \mathcal{C}_1$ é também um solvente do polinômio $P(X)$ segundo o resultado do Teorema 2.4.1 o que se verifica pois,

$$P\left(\begin{bmatrix} 13.4868 - 0.122121i & 1.6654 + 0.488237i \\ -0.137183 - 0.000316732i & 1.83784 - 0.153651i \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

A seguir é apresentado um exemplo onde se utiliza o Algoritmo 4.1.2 aplicado a um feixe genérico (A, B) , onde $A, B \in \mathbb{C}^{10 \times 10}$.

Exemplo 4.1.15. Considere-se o feixe (A, B) , com

$$A = \begin{bmatrix} 10.06 & 53.9 & 1. & -17 & 0 & 1 & 1 & 10.06 & 53.9 & 0 \\ -3.5 & -707.9 & 8.5 & 26.5 & 0 & 0 & -17 & -3.5 & -707.9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 26.5 & -5 & 1 & 0 \\ 10 & -1.7 & -3.5 & -1.95 & -5.79 & 1.006 & 1.006 & 53.9 & 10 & 0 \\ 85 & 26.5 & -5 & 28.95 & 6.735 & -26.95 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 10.06 & 53.9 & 1. & -17 & -3.5 & 2.75 & 3.5 \\ 53.9 & 10 & -17 & -3.5 & -707.9 & 8.5 & 26.5 & -5 & -11.5 & 3 \\ -707.9 & 8.5 & 26.5 & -5 & 1 & 0 & 0 & 0 & 0 & 1 \\ -1.95 & -5.79 & 1.006 & 53.9 & 10 & -1.7 & -3.5 & 2.75 & 3.5 & -1.25 \\ 28.95 & 6.735 & -26.95 & -707.9 & 85 & 26.5 & -5 & -11.5 & 3 & 7.5 \end{bmatrix}$$

e

$$B = \begin{bmatrix} -8.5 & -17.5 & -9.75 & -28.95 & 0 & 5 & 5 & 50.3 & 269.5 & 0 \\ 132.5 & -25 & 50 & -8.5 & -17.5 & -9.75 & -85 & -17.5 & -3539.5 & 0 \\ 0 & 0 & 425 & 132.5 & -25 & 269.5 & 5. & -85 & -17.5 & 0 \\ 50 & -8.5 & -17.5 & -9.75 & -28.95 & -3539.5 & 42.5 & 132.5 & -25 & 0 \\ 425 & 132.5 & -25 & 144.75 & 33.675 & -134.75 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 50.3 & 269.5 & 5. & -85 & -17.5 & 13.75 & 0 \\ 269.5 & 50 & -85 & -17.5 & -3539.5 & 42.5 & 132.5 & -25 & -57.5 & -85 \\ -3539.5 & 42.5 & 132.5 & -25 & 5 & 0 & 0 & 0 & 0 & 132.5 \\ 0 & 0 & -9.75 & -28.95 & -3539.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 144.75 & 33.675 & -134.75 & 0 & 0 & 0 & 0 & -17.5 \end{bmatrix}$$

Aplicando o método de Newton vetorial (Algoritmo 4.1.2) com vista a determinar uma aproximação de um bloco valor próprio $X \in \mathbb{C}^{5 \times 5}$ do feixe (A, B) , equação $AV = BVX$, com $\beta = mn$ e $\alpha = -500$,

$$\bar{V}^{(0)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 50 & 50 & 50 & 50 & 50 \\ 50 & 50 & 50 & 50 & 50 \\ 50 & 50 & 50 & 50 & 50 \\ 50 & 50 & 50 & 50 & 50 \\ 50 & 50 & 50 & 50 & 50 \end{bmatrix} \text{ e } X^{(0)} = \begin{bmatrix} -500 & 0 & 0 & 0 & 0 \\ 0 & -500 & 0 & 0 & 0 \\ 0 & 0 & -500 & 0 & 0 \\ 0 & 0 & 0 & -500 & 0 \\ 0 & 0 & 0 & 0 & -500 \end{bmatrix},$$

obtem-se após 17 iterações uma aproximação dos bloco vetor próprio e bloco valor próprio com $\|AV^{(17)} - BV^{(17)}X^{(17)}\| < 10^{-5}$,

$$\bar{V}^{(17)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 4.47619 & 0.858355 & -0.193582 & 0.96915 & 2.61045 \\ 7.64241 & -0.796121 & 0.0482937 & -0.119137 & 16.5432 \\ -1.67735 & 0.0147786 & -0.223734 & 0.0617392 & -1.20449 \\ -1.87017 & -0.929085 & -0.0032418 & -1.06489 & -0.781837 \\ 61.002 & -3.56103 & -1.20883 & -2.53589 & 61.1784 \end{bmatrix}$$

e

$$X^{(17)} = \begin{bmatrix} 0.123459 & 0.00748767 & -0.0191711 & -0.483788 & 0.255922 \\ 22.9467 & 0.77147 & -2.11311 & -43.8363 & 24.3675 \\ 6.57105 & 0.160109 & -0.571606 & -11.9761 & 7.57909 \\ -20.3329 & -0.708471 & 1.91586 & 40.0359 & -22.0694 \\ 0.183158 & 0.00626411 & -0.0146748 & -0.30924 & 0.197733 \end{bmatrix},$$

com

$$\Lambda(X) = \{40.6567, -0.134035, 0.0453263, -0.00552832 + 0.0120668i, -0.00552832 - 0.0120668i\}$$

sendo o espectro de (A, B)

$$\Lambda(A, B) = \{40.6567, 0.221016, 0.2, 0.2, 0.2, 0.199485, -0.134035, 0.0453263, -0.00552832 - 0.0120668i, -0.00552832 + 0.0120668i\}.$$

4.2 Cálculo de Feixes Próprios

Após a adaptação do Método Newton Vetorial, Algoritmo 3.3.1, para o cálculo de blocos valores próprios de um feixe (A, I_n) , Algoritmo 4.1.1, secção 4.1.2, e posteriormente, para o cálculo de blocos valores próprios de um feixe genérico (A, B) , Algoritmo 4.1.2, secção 4.1.12, pretende-se agora construir uma nova versão do Método Newton Vetorial para o cálculo de feixes próprios de um feixe (A, B) .

Salienta-se que se Y for um bloco valor próprio do feixe (B, A) (isto é, verifica $BV = AVY$ onde V tem característica máxima) então o valor próprio nulo da matriz nilpotente Y corresponde ao valor próprio infinito de (A, B) .

De uma forma geral (ver Pereira [32], pág. 2913), um feixe regular (A, B) , com $A, B \in \mathbb{C}^{n \times n}$, admite matrizes $V \in \mathbb{C}^{n \times p}$ e $W \in \mathbb{C}^{n \times r}$ de característica máxima, com $p + r \leq n$, tais que

$$AV = BVX \quad \text{e} \quad BW = AWY, \tag{4.2.1}$$

se e só se todos os valores próprios de X são valores próprios finitos de (A, B) e o valor próprio nulo da matriz nilpotente Y corresponde ao valor próprio infinito de (A, B) , tendo as matrizes BV e AW característica máxima também.

Para além disso, se $p + r = n$ então a matriz $(\lambda I_p - X) \oplus (\lambda Y - I_r)$ é equivalente ao feixe $(A, B) = \lambda B - A$ e se X e Y estiverem na forma canónica de Jordan (isto é, $X = J_X$ e $Y = J_Y$, equação (2.2.9)), então a matriz $(\lambda I_p - X) \oplus (\lambda Y - I_r)$ corresponde à forma canónica de Weierstrass do feixe (A, B) , equação (2.2.10).

Nestas condições ainda, as matrizes $J = J_X$ e $N = J_Y$ são designadas por forma canónica finita e forma canónica infinita, respetivamente, do feixe (A, B) .

Por outro lado, segundo Xian [40], pág. 1036, para $s_0 \notin \Lambda(A, B) = \{\lambda \in \mathbb{C} : \det(\lambda B - A) = 0\}$, se a matriz $\Omega = (s_0 B - A)^{-1} B$ tem a seguinte forma de Jordan

$$J = \begin{bmatrix} J(s_1) & & & 0 \\ & J(s_2) & & \\ & & \ddots & \\ & & & J(s_k) \\ 0 & & & & J(0) \end{bmatrix},$$

então o feixe $\lambda B - A$ tem a seguinte forma canónica finita e infinita

$$J_f = \begin{bmatrix} J(s_0 - 1/s_1) & & & 0 \\ & J(s_0 - 1/s_2) & & \\ & & \ddots & \\ & & & J(s_0 - 1/s_k) \\ 0 & & & & \end{bmatrix} \text{ e } J_\infty = \begin{bmatrix} J(0) \end{bmatrix}.$$

Com efeito, se $\lambda \in \Lambda(A, B)$ com $AV = \lambda BV$, então

$$(s_0 B - A)V = s_0 BV - AV = s_0 BV - \lambda BV = (s_0 - \lambda)BV.$$

Isto é, $(s_0 - \lambda) \in \Lambda((s_0 B - A), B)$.

Da mesma forma,

$$V = (s_0 - \lambda)(s_0 B - A)^{-1} BV,$$

ou seja, $(s_0 - \lambda) \in \Lambda(I, \Omega)$. Como para cada par (A, B) se tem ainda $\Lambda(A, B) = 1/\Lambda(B, A)$ então

$$1/(s_0 - \lambda) \in \Lambda(\Omega, I) = \Lambda(\Omega).$$

Esta relação é ainda válida se, em vez de valores próprios, forem considerados blocos valores próprios. Se Λ é um bloco valor próprio de Ω , i. é., $\Omega V = I_m V \Lambda$ então

$$(s_0 B - A)^{-1} BV = V \Lambda \Leftrightarrow BV = (s_0 B - A) V \Lambda,$$

Λ é um bloco valor próprio de $(B, (s_0 B - A))$,

$$BV \Lambda^{-1} = (s_0 B - A) V,$$

Λ^{-1} é um bloco valor próprio de $((s_0 B - A), B)$,

$$BV \Lambda^{-1} = s_0 BV - AV \Leftrightarrow BV \Lambda^{-1} - BV s_0 I_n = -AV \Leftrightarrow BV (s_0 I_n - \Lambda^{-1}) = AV,$$

$(s_0 I_n - \Lambda^{-1})$ um bloco valor próprio de (A, B) .

Deste ponto de vista, na Definição seguinte é adaptada o conceito de bloco valor próprio ao feixe $\lambda B - A$, considerando blocos valores próprios $\lambda Y - X$ tais que $(\lambda B - A)V = W(\lambda Y - X)$.

Definição 4.2.1. (Pereira [32], pág. 2914) Seja (A, B) um feixe onde A, B são matrizes quadradas de ordem $m \times m$ com entradas em \mathbb{C} . O feixe $(X, Y) = \lambda Y - X$, onde X, Y são matrizes quadradas de ordem $n \times n$, com $n < m$, diz-se um feixe próprio de $\lambda B - A$ se existirem bloco vetores $V, W \in \mathbb{C}^{m \times n}$ de característica máxima n tais que

$$(\lambda B - A)V = W(\lambda Y - X). \quad (4.2.2)$$

Note-se que, neste caso, o processo de cálculo de um feixe próprio passa ainda pela resolução de um sistema não linear com $m \times n$ equações $((\lambda B - A)V \in \mathbb{C}^{m \times n})$ e $2m \times n + 2n^2$ variáveis, que é apenas possível, ou reduzindo o número de variáveis ou aumentando o número de equações.

A construção de um método iterativo para o cálculo de feixes próprios é feita na secção seguinte com base nesta redução do número de variáveis e aumentando simultaneamente o número de equações consideradas.

4.2.2 Método Newton Vetorial para Feixes Próprios

Pretende-se, como já referido, construir uma nova versão do método Newton Vetorial para o cálculo de feixes próprios de um feixe (A, B) . Recorrendo novamente à ideia usada na secção 4.1, se em (4.2.2) forem considerados bloco vetores V e W normalizados da forma

$$\bar{V} = \begin{bmatrix} I_n \\ V_2 \end{bmatrix}, \bar{W} = \begin{bmatrix} I_n \\ W_2 \end{bmatrix}, \quad (4.2.3)$$

então a equação

$$(\lambda B - A)\bar{V} - \bar{W}(\lambda Y - X) = 0, \quad (4.2.4)$$

traduz um sistema não linear de $m \times n$ equações, $(\lambda B - A)\bar{V} \in \mathbb{C}^{m \times n}$, com $2m \times n$ variáveis, uma vez que, estão fixas $2n^2$ variáveis, de \bar{V} e \bar{W} ($\bar{V}_1 = I_n$ e $\bar{W}_1 = I_n$).

Para além disso, uma vez que o sistema (4.2.4) é independente do parâmetro λ , este pode ser decomposto em

$$A\bar{V} - \bar{W}X = 0 \wedge B\bar{V} - \bar{W}Y = 0. \quad (4.2.5)$$

que resulta num número de equações, $2m \times n$, igual ao número de incógnitas, e que pode ser comparável com um duplo sistema do tipo bloco valor próprio, definido em (4.1.22).

Trata-se agora de generalizar o Método do Algoritmo 4.1.2 e adaptá-lo a esta situação. A seguir apresenta-se uma nova versão do Algoritmo do método Newton Vetorial ajustado ao cálculo de feixes próprios que, basicamente, consiste em duplicar o número de equações, como apontado em (4.2.5), e em reduzir o número de variáveis, através da utilização de vetores V e W normalizados na forma (4.2.3).

4.2.2.1 Algoritmo do Método

Como dito atrás, o sistema (4.2.4) é independente do parâmetro λ e, por esse motivo, pode ser decomposto em dois sistemas, equação (4.2.5), idênticos aos usados no cálculo de blocos valores próprios.

Assim, o Algoritmo seguinte resulta da adaptação feita ao método NVBVP 4.1.2, nomeadamente nos pontos 3, 7, 8 e 9.

 Algoritmo 4.2.1 (NVFP)

- 1: Dados $\varepsilon > 0$, A , B , $\overline{V}^{(0)}$, $X^{(0)}$, $\overline{W}^{(0)}$ e $Y^{(0)}$.
- 2: para $k = 0, 1, \dots$ fazer
- 3: $z^{(k)} = \text{vec}(A\overline{V}^{(k)} - \overline{W}^{(k)}X^{(k)}, B\overline{V}^{(k)} - \overline{W}^{(k)}Y^{(k)})$;
- 4: se $\|z^{(k)}\| < \varepsilon$ então
- 5: o método pára.
- 6: caso contrário
- 7: $(\tilde{s}^{(k)}, \lambda^{(k)}, \tilde{t}^{(k)}, \omega^{(k)})$, solução do sistema linear

$$\begin{pmatrix} I_n \otimes A\tilde{s}^{(k)} - X^T \otimes I_m \tilde{t}^{(k)} - I_n \otimes \overline{W}\lambda^{(k)} \\ I_n \otimes B\tilde{s}^{(k)} - Y^T \otimes I_m \tilde{t}^{(k)} - I_n \otimes \overline{W}\omega^{(k)} \end{pmatrix} = -z^{(k)};$$

- 8: $(\tilde{S}^{(k)}, \Lambda^{(k)}, \tilde{T}^{(k)}, \Omega^{(k)}) = \text{vec}^{-1}(\tilde{s}^{(k)}, \lambda^{(k)}, \tilde{t}^{(k)}, \omega^{(k)})$;
 - 9: $(\overline{V}^{(k+1)}, X^{(k+1)}, \overline{W}^{(k+1)}, Y^{(k+1)}) = (\overline{V}^{(k)}, X^{(k)}, \overline{W}^{(k)}, Y^{(k)}) + (\tilde{S}^{(k)}, \Lambda^{(k)}, \tilde{T}^{(k)}, \Omega^{(k)})$
 - 10: $\overline{V} = \overline{V}^{(k)}$, $\overline{W} = \overline{W}^{(k)}$, $X = X^{(k)}$ e $Y = Y^{(k)}$ tais que $(\lambda B - A)\overline{V} = \overline{W}(\lambda Y - X)$.
-

Atendendo à Observação 4.1.4 e considerando a função $G(\overline{V}, X, \overline{W}, Y) = (A\overline{V} - \overline{W}X, B\overline{V} - \overline{W}Y)$, esta pode entender-se como definida em $\mathbb{C}^{(m-n) \times n} \times \mathbb{C}^{n \times n} \times \mathbb{C}^{(m-n) \times n} \times \mathbb{C}^{n \times n}$ uma vez que, estão fixas $2n^2$ variáveis de \overline{V} e \overline{W} . Isto é,

$$\begin{array}{ccc} (\mathbb{C}^{(m-n) \times n} \times \mathbb{C}^{n \times n})^2 & \xrightarrow{L} & (\{I_m\} \times \mathbb{C}^{(m-n) \times n} \times \mathbb{C}^{n \times n})^2, \\ & \searrow \tilde{F} & \downarrow G \\ & & \mathbb{C}^{m \times n} \times \mathbb{C}^{m \times n} \end{array}$$

onde $L = (p_1^{-1} \times id \times p_1^{-1} \times id)$,

$$\begin{array}{ccc} (V_2, X, W_2, Y) & \xrightarrow{L} & (\overline{V}, X, \overline{W}, Y), \\ & \searrow \tilde{F} & \downarrow G \\ & & (A\overline{V} - \overline{W}X, B\overline{V} - \overline{W}Y) \end{array}$$

pele que, considerando a função composta $\tilde{F} = G \circ L$, tem-se

$$\begin{aligned} \tilde{F}(V_2, X, W_2, Y) &= G \circ L(V_2, X, W_2, Y) \\ &= G(\underbrace{p_1^{-1}(V_2)}_{=\overline{V}}, id(X), \underbrace{p_1^{-1}(W_2)}_{=\overline{W}}, id(Y)) \\ &= G(\overline{V}, X, \overline{W}, Y) \\ &= (A\overline{V} - \overline{W}X, B\overline{V} - \overline{W}Y), \end{aligned}$$

ou seja,

$$\tilde{F}(V_2, X, W_2, Y) = (A\overline{V} - \overline{W}X, B\overline{V} - \overline{W}Y). \quad (4.2.6)$$

Derivando ambos os membros de (4.2.6) no ponto (V_2, X, W_2, Y) segundo a direção de $(S_2, \Lambda, T_2, \Omega) \in \mathbb{C}^{(m-n) \times n} \times \mathbb{C}^{n \times n} \times \mathbb{C}^{(m-n) \times n} \times \mathbb{C}^{n \times n}$ (ver Observação 4.1.4) tem-se

$$\begin{aligned} \tilde{F}'(V_2, X, W_2, Y)[S_2, \Lambda, T_2, \Omega] &= G'(\bar{V}, X, \bar{W}, Y)[\tilde{S}, \Lambda, \tilde{T}, \Omega] \\ &= (A\tilde{S} - \tilde{T}X - \bar{W}\Lambda, B\tilde{S} - \tilde{T}Y - \bar{W}\Omega), \end{aligned} \quad (4.2.7)$$

onde $\tilde{S}, \tilde{T} \in \mathbb{C}^{m \times n}$ representam matrizes da forma (ver Observação 4.1.4),

$$\tilde{S} = \begin{bmatrix} 0 \\ S_2 \end{bmatrix}, \quad \tilde{T} = \begin{bmatrix} 0 \\ T_2 \end{bmatrix},$$

e $\Lambda, \Omega \in \mathbb{C}^{n \times n}$.

Aplicando a função $vec()$ a ambos os membros de (4.2.6) então as soluções da equação (4.2.5) podem caracterizar-se em termos dos zeros função

$$F : \quad \mathbb{C}^{2mn} \quad \longrightarrow \quad \mathbb{C}^{2mn} \\ vec(V_2, X, W_2, Y) \quad \longrightarrow \quad vec(A\bar{V} - \bar{W}X, B\bar{V} - \bar{W}Y), \quad (4.2.8)$$

cuja derivada, lema 2.1.19, avaliada em (V_2, X, W_2, Y) segundo a direção de $(S_2, \Lambda, T_2, \Omega)$, corresponde a

$$\begin{aligned} vec\left(\tilde{F}'(V_2, X, W_2, Y)[S_2, \Lambda, T_2, \Omega]\right) &= vec\left(G'(\bar{V}, X, \bar{W}, Y)[\tilde{S}, \Lambda, \tilde{T}, \Omega]\right) \\ J_{F(vec(V_2, X, W_2, Y))}(vec(S_2, \Lambda, T_2, \Omega)) &= vec\left(\begin{array}{cc} A\tilde{S} - \tilde{T}X - \bar{W}\Lambda & B\tilde{S} - \tilde{T}Y - \bar{W}\Omega \end{array}\right), \end{aligned}$$

onde

$$vec\left(\begin{array}{cc} A\tilde{S} - \tilde{T}X - \bar{W}\Lambda & B\tilde{S} - \tilde{T}Y - \bar{W}\Omega \end{array}\right) = \begin{pmatrix} I_n \otimes A vec(\tilde{S}) - X^T \otimes I_m vec(\tilde{T}) - I_n \otimes \bar{W} vec(\Lambda) \\ I_n \otimes B vec(\tilde{S}) - Y^T \otimes I_m vec(\tilde{T}) - I_n \otimes \bar{W} vec(\Omega) \end{pmatrix}.$$

Ou seja, a derivada de $vec \circ \tilde{F}$ no ponto (V_2, X, W_2, Y) na direção do vetor $(S_2, \Lambda, T_2, \Omega)$ é igual ao produto da matriz matriz jacobiana de F avaliada no ponto $(v_2, x, w_2, y) = vec(V_2, X, W_2, Y)$ pelo vetor $(s_2, \lambda, t_2, \omega) = vec(S_2, \Lambda, T_2, \Omega)$ que está definido por

$$J_{F(v_2, x, w_2, y)}(s_2, \lambda, t_2, \omega) = \begin{pmatrix} I_n \otimes A\tilde{s} - X^T \otimes I_m \tilde{t} - I_n \otimes \bar{W}\lambda \\ I_n \otimes B\tilde{s} - Y^T \otimes I_m \tilde{t} - I_n \otimes \bar{W}\omega \end{pmatrix}. \quad (4.2.9)$$

Usando esta construção, pode agora aplicar-se o método de Newton à função $F : \mathbb{C}^{2nm} \longrightarrow \mathbb{C}^{2nm}$,

$$\left(v_2^{(k+1)}, x^{(k+1)}, w_2^{(k+1)}, y^{(k+1)}\right) = \left(v_2^{(k)}, x^{(k)}, w_2^{(k)}, y^{(k)}\right) + \left(s_2^{(k)}, \lambda^{(k)}, t_2^{(k)}, \omega^{(k)}\right), \quad k = 0, 1, 2, \dots$$

onde $(\tilde{s}^{(k)}, \lambda^{(k)}, \tilde{t}^{(k)}, \omega^{(k)})$ é solução do sistema

$$J_{F(v_2^{(k)}, x^{(k)}, w_2^{(k)}, y^{(k)})} \left(s_2^{(k)}, \lambda^{(k)}, t_2^{(k)}, \omega^{(k)}\right) = -F \left(v_2^{(k)}, x^{(k)}, w_2^{(k)}, y^{(k)}\right), \quad (4.2.10)$$

e corresponde ao acréscimo da iteração k do Método de Newton, equação (4.2.9).

4.2.2.2 Convergência

A convergência deste método (NVFP definido no Algoritmo 4.2.1), enquadra-se ainda no Teorema 3.3.9 aplicado à função F definida em (4.2.8), com os ajustamentos resultantes da sua construção, equações (4.2.6), (4.2.7), (4.2.9) e (4.2.10).

À semelhança das considerações feitas na secção 4.1.3 relativamente à convergência do Algoritmo 4.1.1, também são agora apresentados Teoremas que visam assegurar que, nas condições de convergência do Teorema 3.3.9, o Algoritmo 4.2.1 está bem definido e gera uma sucessão que converge para um feixe próprio do feixe (A, B) .

Assim, faz-se a adaptação do resultado, definido em termos dos zeros da função F , para feixes próprios do feixe (A, B) , soluções da equação $(\lambda B - A)V = W(\lambda Y - X)$.

Observação 4.2.3. Se $(v_2, x, w_2, y) \in (\mathbb{C}^{(m-n) \times n} \times \mathbb{C}^{n \times n})^2$ é um zero de F , então $(\lambda B - A)\bar{V} = \bar{W}(\lambda Y - X)$, onde, $X = \text{vec}^{-1}(x)$, $Y = \text{vec}^{-1}(y)$ e

$$\bar{V} = \begin{bmatrix} I_n \\ \text{vec}^{-1}(v_2) \end{bmatrix}, \bar{W} = \begin{bmatrix} I_n \\ \text{vec}^{-1}(w_2) \end{bmatrix}.$$

Seja $(V_2, X, W_2, Y) = \text{vec}^{-1}(v_2, x, w_2, y) \in (\mathbb{C}^{(m-n) \times n} \times \mathbb{C}^{n \times n})^2$.

Se (v_2, x, w_2, y) é um zero de F , então, por construção (4.2.6), $\tilde{F}(V_2, X, W_2, Y) = (0, 0)$, ou seja, $(A\bar{V} - \bar{W}X, B\bar{V} - \bar{W}Y) = (0, 0)$. Isto é,

$$\lambda(B\bar{V} - \bar{W}Y) + (A\bar{V} - \bar{W}X) = \lambda 0 + 0,$$

ou

$$(\lambda B - A)\bar{V} - \bar{W}(\lambda Y - X) = 0.$$

A seguir são adaptadas as hipóteses do Teorema 3.3.9 de modo a garantir a convergência do Algoritmo 4.2.1 para um zero da função F , e que, segundo a Observação 4.2.3, se traduz também na convergência da sucessão $(X^{(k)}, Y^{(k)})$ para um feixe próprio (X, Y) do feixe (A, B) .

Teorema 4.2.4. Se a função F , definida em (4.2.8), para uma aproximação inicial $(\bar{v}^{(0)}, x^{(0)}, \bar{w}^{(0)}, y^{(0)}) \in (\mathbb{C}^{(m-n)n} \times \mathbb{C}^{n^2})^2$, verifica as condições de convergência do Teorema 3.3.9 (Gutierrez [17], pág. 9), então a sucessão calculada nos pontos 3) e 4) do Algoritmo 4.2.1 está bem definida e converge para (\bar{v}, x, \bar{w}, y) , onde $(\bar{V}, X, \bar{W}, Y) = \text{vec}^{-1}(\bar{v}, x, \bar{w}, y)$ é uma solução da equação

$$(\lambda B - A)\bar{V} = \bar{W}(\lambda Y - X).$$

Demonstração. Nas condições do Teorema e atendendo à igualdade (4.2.9), a solução do sistema considerado em 3),

$$\begin{pmatrix} I_n \otimes A\tilde{s}^{(k)} - X^T \otimes I_m \tilde{t}^{(k)} - I_n \otimes \bar{W}\lambda^{(k)} \\ I_n \otimes B\tilde{s}^{(k)} - Y^T \otimes I_m \tilde{t}^{(k)} - I_n \otimes \bar{W}\omega^{(k)} \end{pmatrix} = -z^{(k)},$$

está definida por $(\tilde{s}^{(k)}, \lambda^{(k)}, \tilde{t}^{(k)}, \omega^{(k)})$, onde

$$(s_2^{(k)}, \lambda^{(k)}, t_2^{(k)}, \omega^{(k)}) = -J_F^{-1}(v_2^{(k)}, x^{(k)}, w_2^{(k)}, y^{(k)})F(v_2^{(k)}, x^{(k)}, w_2^{(k)}, y^{(k)}),$$

e as restantes coordenadas de $\tilde{s}^{(k)}$ e $\tilde{t}^{(k)}$ são nulas. Portanto, para cada $k = 1, 2, \dots$, a sucessão

$$\left(\bar{v}^{(k+1)}, x^{(k+1)}, w^{(k+1)}, y^{(k+1)}\right) = \left(\bar{v}^{(k)}, x^{(k)}, w^{(k)}, y^{(k)}\right) + \left(\tilde{s}^{(k)}, \lambda^{(k)}, \tilde{t}^{(k)}, \omega^{(k)}\right)$$

está bem definida.

Por outro lado, a sucessão

$$\left(v_2^{(k+1)}, x^{(k+1)}, w_2^{(k+1)}, y^{(k+1)}\right) = \left(v_2^{(k)}, x^{(k)}, w_2^{(k)}, y^{(k)}\right) + \left(s_2^{(k)}, \lambda^{(k)}, t_2^{(k)}, \omega^{(k)}\right)$$

converge para $(v_2, x, w_2, y) \in (\mathbb{C}^{(m-n)n} \times \mathbb{C}^{n^2})^2$ um zero da função F e, por conseguinte, segundo a Observação 4.2.3, a sucessão $(\bar{v}^{(k+1)}, x^{(k+1)}, \bar{w}^{(k+1)}, y^{(k+1)})$ converge para (\bar{v}, x, \bar{w}, y) tal que $(\lambda B - A)\bar{V} = \bar{W}(\lambda Y - X)$. ■

Assegurada a convergência do Algoritmo 4.2.1, através da Observação 4.2.3 e Teorema 4.2.4, são apresentados de seguida, secção 4.2.4.1, exemplos deste método.

4.2.4.1 Exemplos

Dadas as restrições introduzidas em (4.2.3), nos exemplos seguintes consideram-se aproximações iniciais $X^{(0)}, \bar{V}^{(0)}, Y^{(0)}, \bar{W}^{(0)}$, com

$$X^{(0)} = \alpha_1 I_n, \quad Y^{(0)} = \alpha_2 I_n,$$

múltiplos da matriz identidade e

$$\bar{V}^{(0)} = \begin{bmatrix} I_n \\ \beta mn U \end{bmatrix} = \pm \bar{W}^{(0)},$$

onde $U \in \mathbb{C}^{(m-n) \times n}$ é a matriz com entradas iguais a 1.

São ainda apresentados os espectros, quer do feixe dado (A, B) , quer do feixe próprio (X, Y) calculado através do Algoritmo 4.2.1, verificando-se $\Lambda(X, Y) \subset \Lambda(A, B)$, uma vez que é uma condição necessária (mas não suficiente, ver Pereira [32], pág. 2914 e Observação 2.4.5) para que (X, Y) seja um feixe próprio de (A, B) .

Neste primeiro exemplo são ainda expostos de forma detalhada os cálculos que permitem a determinação da iteração seguinte, nomeadamente os pontos 3, 7, 8 e 9 do Algoritmo 4.2.1

Exemplo 4.2.5. Considere-se o feixe

$$\lambda B - A = \begin{bmatrix} 3\lambda - 1 & 2 - 3\lambda & \lambda - 2 \\ -\lambda & -1 & 0 \\ 11\lambda - 18 & 14 - 3\lambda & \lambda - 2 \end{bmatrix},$$

$A, B \in \mathbb{C}^{3 \times 3}$, isto é $m = 3$, relativamente ao qual se irá determinar uma aproximação de um feixe próprio de dimensão $n = 2$. Consideram-se as aproximações iniciais, com $\beta = 1.13$, $\alpha_1 = -20$ e $\alpha_2 = 4$,

$$\bar{V}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 6.78 & 6.78 \end{bmatrix}, \quad X^{(0)} = \begin{bmatrix} -20 & 0 \\ 0 & -20 \end{bmatrix}, \quad \bar{W}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -6.78 & -6.78 \end{bmatrix} \text{ e } Y^{(0)} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}.$$

Tem-se que

$$J_{F(v_2^{(k)}, x^{(k)}, w_2^{(k)}, y^{(k)})} \left(s_2^{(k)}, \lambda^{(k)}, t_2^{(k)}, \omega^{(k)} \right) = \begin{bmatrix} I_2 \otimes A\tilde{s}^{(0)} - X^T \otimes I_3 \tilde{t}^{(0)} - I_2 \otimes \overline{W}\lambda^{(0)} \\ I_2 \otimes B\tilde{s}^{(0)} - Y^T \otimes I_3 \tilde{t}^{(0)} - I_2 \otimes \overline{W}\omega^{(0)} \end{bmatrix}$$

$$= \begin{bmatrix} 2s_{3,1} - \lambda_{1,1} \\ -\lambda_{2,1} \\ 2s_{3,1} + 20.t_{3,1} + 6.78\lambda_{1,1} + 6.78\lambda_{2,1} \\ 2s_{3,2} - \lambda_{1,2} \\ -\lambda_{2,2} \\ 2s_{3,2} + 20.t_{3,2} + 6.78\lambda_{1,2} + 6.78\lambda_{2,2}s_{3,1} - \omega_{1,1} \\ -\omega_{2,1} \\ s_{3,1} - 4t_{3,1} + 6.78\omega_{1,1} + 6.78\omega_{2,1} \\ s_{3,2} - \omega_{1,2} \\ -\omega_{2,2} \\ s_{3,2} - 4t_{3,2} + 6.78\omega_{1,2} + 6.78\omega_{2,2} \end{bmatrix},$$

$$z^{(0)} = \text{vec} \left(A\overline{V}^{(0)} - \overline{W}^{(0)} X^{(0)}, B\overline{V}^{(0)} - \overline{W}^{(0)} Y^{(0)} \right) = \begin{bmatrix} 34.56 \\ 0 \\ -104.04 \\ 11.56 \\ 21 \\ -136.04 \\ 5.78 \\ -1 \\ 44.9 \\ 3.78 \\ -4 \\ 30.9 \end{bmatrix}$$

e

$$\left(\tilde{s}^{(k)}, \lambda^{(k)}, \tilde{t}^{(k)}, \omega^{(k)} \right) = \begin{bmatrix} -9.48988 \\ -4.25558 \\ 15.5802 \\ 0 \\ 3.04884 \\ 21 \\ 0.869286 \\ -0.925 \\ -3.70988 \\ -1 \\ -0.475578 \\ -4 \end{bmatrix}.$$

Logo,

$$\left(\bar{v}^{(1)}, x^{(1)}, \bar{w}^{(1)}, y^{(1)}\right) = \begin{bmatrix} 6.78 \\ 6.78 \\ -20. \\ 0 \\ 0 \\ -20 \\ -6.78 \\ -6.78 \\ 4 \\ 0 \\ 0 \\ 4 \end{bmatrix} + \begin{bmatrix} -9.48988 \\ -4.25558 \\ 15.5802 \\ 0 \\ 3.04884 \\ 21 \\ 0.869286 \\ -0.925 \\ -3.70988 \\ -1 \\ -0.475578 \\ -4 \end{bmatrix} = \begin{bmatrix} -2.70988 \\ 2.52442 \\ -4.41976 \\ 0 \\ 3.04884 \\ 1 \\ -5.91071 \\ -7.705 \\ 0.290121 \\ -1 \\ -0.475578 \\ 0 \end{bmatrix},$$

isto é,

$$\bar{V}^{(1)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -2.70988 & 2.52442 \end{bmatrix}, \quad X^{(1)} = \begin{bmatrix} -4.41976 & 3.04884 \\ 0 & 1 \end{bmatrix},$$

$$\bar{W}^{(1)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -5.91071 & -7.705 \end{bmatrix} \text{ e } Y^{(1)} = \begin{bmatrix} 0.290121 & -0.475578 \\ -1 & 0 \end{bmatrix}.$$

Após 4 iterações obtém-se uma aproximação do feixe próprio com

$$\|z^{(4)}\| = \|\text{vec}(A\bar{V}^{(4)} - \bar{W}^{(4)}X^{(4)}, B\bar{V}^{(4)} - \bar{W}^{(4)}Y^{(4)})\|_2 < 10^{-5},$$

$$\bar{V}^{(4)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -4.92 & 3 \end{bmatrix}, \quad X^{(4)} = \begin{bmatrix} -8.84 & 4 \\ 0 & 1 \end{bmatrix},$$

$$\bar{W}^{(4)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -0.923077 & -4.30769 \end{bmatrix} \text{ e } Y^{(4)} = \begin{bmatrix} -1.92 & 0 \\ -1 & 0 \end{bmatrix}.$$

Ou seja, obtém-se o feixe próprio,

$$\lambda Y - X = \begin{bmatrix} 8.84 - 1.92\lambda & -4 \\ -1\lambda & -1 \end{bmatrix}$$

com $\Lambda(X, Y) = \{-4.25, \infty\} \subset \Lambda(A, B) = \{-4.25, 2, \infty\}$.

A forma canônica de Weierstrass do feixe (A, B) é a λ -matriz

$$K_{(A,B)} = \begin{bmatrix} \lambda I - J_F & 0 \\ 0 & \lambda J_\infty - I \end{bmatrix} = \begin{bmatrix} \lambda - 2 & 0 & 0 \\ 0 & \lambda + 4.25 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

A seguir apresenta-se um exemplo de um feixe (A, B) , com $A, B \in \mathbb{C}^{4 \times 4}$, relativamente ao qual se irá determinar uma aproximação de um feixe próprio de dimensão $n = 3$. São ainda indicados os espectros, quer do feixe dado (A, B) quer do do feixe próprio (X, Y) calculado, verificando-se

$\Lambda(X, Y) \subset \Lambda(A, B)$.

Exemplo 4.2.6. Considere-se o feixe $\lambda B - A$ dado por

$$\lambda B - A = \begin{bmatrix} 3\lambda - 1 & 2 - 3\lambda & \lambda - 2 & 2\lambda - 1 \\ -\lambda & -1 & 0 & 1 - \lambda \\ 11\lambda - 18 & 14 - 3\lambda & \lambda - 2 & 6\lambda - 10 \\ 3 - 2\lambda & -2 & 0 & 2 - \lambda \end{bmatrix},$$

e as seguintes aproximações iniciais, com $\beta = 1.13$ e $\alpha_1 = 6$ e $\alpha_2 = -30$,

$$\bar{V}^{(0)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 13.56 & 13.56 & 13.56 \end{bmatrix}, \quad \bar{W}^{(0)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -13.56 & -13.56 & -13.56 \end{bmatrix},$$

$$X^{(0)} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix} \text{ e } Y^{(0)} = \begin{bmatrix} -30 & 0 & 0 \\ 0 & -30 & 0 \\ 0 & 0 & -30 \end{bmatrix}.$$

Após 8 iterações obtém-se uma aproximação do feixe próprio com

$$\|z^{(8)}\| = \|vec(A\bar{V}^{(8)} - \bar{W}^{(8)}X^{(8)}, B\bar{V}^{(8)} - \bar{W}^{(8)}Y^{(8)})\|_2 < 10^{-5},$$

$$\bar{V}^{(8)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2.85 & 1.65 & -0.55 \end{bmatrix}, \quad \bar{W}^{(8)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1. & 4.66667 & 0.833333 \end{bmatrix},$$

$$X^{(8)} = \begin{bmatrix} -1.85 & -0.35 & 1.45 \\ 2.85 & -0.65 & 0.55 \\ -10.5 & 2.5 & -3.5 \end{bmatrix} \text{ e } Y^{(8)} = \begin{bmatrix} -2.7 & 0.3 & -0.1 \\ 1.85 & -1.65 & 0.55 \\ -6.1 & 6.9 & -2.3 \end{bmatrix},$$

ou seja

$$\lambda Y - X = \begin{bmatrix} 1.85 - 2.7\lambda & 0.3\lambda + 0.35 & -0.1\lambda - 1.45 \\ 1.85\lambda - 2.85 & 0.65 - 1.65\lambda & 0.55\lambda - 0.55 \\ 10.5 - 6.1\lambda & 6.9\lambda - 2.5 & 3.5 - 2.3\lambda \end{bmatrix},$$

com

$$\Lambda(X, Y) = \{0.875 - 1.21835i, 0.875 + 1.21835i, \infty\} \subset \Lambda(A, B) = \{0.875 - 1.21835i, 0.875 + 1.21835i, 2, \infty\}.$$

A forma canônica de Weierstrass de $\lambda B - A$ é a matriz

$$K_{(A,B)} = \begin{bmatrix} \lambda I - J_F & 0 \\ 0 & \lambda J_\infty - I \end{bmatrix} = \begin{bmatrix} \lambda - 2. & 0 & 0 & 0 \\ 0 & \lambda - (0.875 + 1.21835i) & 0 & 0 \\ 0 & 0 & \lambda - (0.875 - 1.21835i) & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

Mantendo a margem de erro usada até agora, $\|z^{(k)}\|_2 < 10^{-5}$, se forem consideradas matrizes de maior dimensão o número de iterações necessárias, k , aumenta mas mantém-se claramente abaixo das 50 iterações, valor que se julga ser aceitável. Neste sentido, a seguir apresentam-se exemplos de

feixes (A, B) , onde $A, B \in \mathbb{C}^{m \times m}$ para $m = 7$ e $m = 8$, relativamente aos quais se irá determinar uma aproximação de um feixe próprio de dimensão $n = 2$.

Em cada caso, tal como nos exemplos anteriores, são ainda apresentados os espectros, quer do feixe dado (A, B) quer do do feixe próprio (X, Y) calculado.

Exemplo 4.2.7. Para $m = 7$, considere-se o feixe

$$\lambda B - A = \begin{bmatrix} \lambda & 0 & 2 & 0 & -3 & 0 & 0 \\ 0 & -\lambda - 3 & 0 & -\lambda & 2\lambda + 2 & 0 & 0 \\ 0 & 0 & \lambda - 1 & 2\lambda - 2 & 0 & 0 & \lambda - 1 \\ \lambda - 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 3\lambda - 3 & 0 & 0 & 0 & 3 \\ 0 & 0 & -\lambda & -1 & 0 & 0 & 0 \end{bmatrix},$$

e as aproximações iniciais com $\beta = 1.13$ e $\alpha_1 = 4$ e $\alpha_2 = -20$,

$$\bar{V}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 15.82 & 15.82 \\ 15.82 & 15.82 \\ 15.82 & 15.82 \\ 15.82 & 15.82 \\ 15.82 & 15.82 \end{bmatrix}, \quad \bar{W}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -15.82 & -15.82 \\ -15.82 & -15.82 \\ -15.82 & -15.82 \\ -15.82 & -15.82 \\ -15.82 & -15.82 \end{bmatrix}$$

$$X^{(0)} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}, \quad Y^{(0)} = \begin{bmatrix} -20 & 0 \\ 0 & -20 \end{bmatrix}.$$

Após 19 iterações, obtém-se uma aproximação do feixe próprio com

$$\|z^{(19)}\| = \|\text{vec}(A\bar{V}^{(19)} - \bar{W}^{(19)}X^{(19)}, B\bar{V}^{(19)} - \bar{W}^{(19)}Y^{(19)})\|_2 < 10^{-5},$$

$$\bar{V}^{(19)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -0.0000118138 & 0 \\ 0.095251 & -0.285714 \\ 0.214288 & 0.357143 \\ 0.142823 & -0.428571 \\ -0.190478 & 0.571429 \end{bmatrix}, \quad \bar{W}^{(19)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0.0000140016 & -0.000006 \\ 1.11111 & -0.333332 \\ 0.111108 & -0.333332 \\ 0.296246 & -0.888865 \\ 0.0493953 & -0.148154 \end{bmatrix},$$

$$X^{(19)} = \begin{bmatrix} 0.642888 & 1.07143 \\ -0.428576 & 2.28571 \end{bmatrix}, \quad Y^{(19)} = \begin{bmatrix} 1 & 0 \\ 0.333325 & 0 \end{bmatrix},$$

onde $\Lambda(X^{(19)}, Y^{(19)}) = \{1.00004, \infty\}$ e $\Lambda(A, B) = \{0, 0.666667, 1, 1, \infty\}$.

Exemplo 4.2.8. Para $m = 8$, considere-se o feixe $\lambda B - A$ dado por

$$\lambda B - A = \begin{bmatrix} 3\lambda - 1 & 2 - 3\lambda & \lambda - 2 & 2\lambda - 1 & 1 - 2\lambda & \lambda - 3 & 3\lambda - 4 & 1 \\ -\lambda & 3\lambda - 1 & 0 & 1 - \lambda & \lambda & 2 - 2\lambda & 1 - 2\lambda & \lambda \\ 11\lambda - 18 & 14 - 5\lambda & \lambda - 2 & 6\lambda - 10 & 12 - 6\lambda & 4 - 4\lambda & 2\lambda - 5 & 3 \\ 3 - 2\lambda & -\lambda - 2 & 0 & 2 - \lambda & -1 & 3\lambda - 1 & \lambda + 1 & 1 - 2\lambda \\ 10 - 5\lambda & 5\lambda - 7 & 2 - \lambda & 5 - 2\lambda & 4\lambda - 7 & -\lambda & 4 - 3\lambda & 2\lambda - 1 \\ \lambda - 3 & -2\lambda & 0 & -1 & 1 - \lambda & 2\lambda - 1 & 2\lambda - 2 & -\lambda - 1 \\ \lambda - 1 & 3 - 3\lambda & \lambda - 2 & \lambda - 1 & 3 - 2\lambda & 3\lambda - 4 & 3\lambda - 4 & 3 - 2\lambda \\ 3\lambda - 11 & 5 - \lambda & 0 & \lambda - 4 & 6 - 2\lambda & \lambda + 1 & 2\lambda - 5 & 1 - \lambda \end{bmatrix}.$$

Utilizando as aproximações iniciais com $\beta = 1.13mn$ e $\alpha_1 = 4$ e $\alpha_2 = -20$,

$$\bar{V}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 18.08 & 18.08 \\ 18.08 & 18.08 \\ 18.08 & 18.08 \\ 18.08 & 18.08 \\ 18.08 & 18.08 \\ 18.08 & 18.08 \end{bmatrix}, \quad \bar{W}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -18.08 & -18.08 \\ -18.08 & -18.08 \\ -18.08 & -18.08 \\ -18.08 & -18.08 \\ -18.08 & -18.08 \\ -18.08 & -18.08 \end{bmatrix}$$

$$X^{(0)} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \text{ e } Y^{(0)} = \begin{bmatrix} -20 & 0 \\ 0 & -20 \end{bmatrix},$$

após 22 iterações obtém-se uma aproximação do feixe próprio

$$\|z^{(22)}\| = \|\text{vec}(A\bar{V}^{(22)} - \bar{W}^{(22)}X^{(22)}, B\bar{V}^{(22)} - \bar{W}^{(22)}Y^{(22)})\|_2 < 10^{-5},$$

$$\bar{V}^{(22)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & -0.0798679 \\ 1 & 0.0798679 \\ -1 & -0.0798679 \\ 0 & -0.920132 \end{bmatrix}, \quad \bar{W}^{(22)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 5.76033 & 6.76033 \\ -6.26033 & -6.26033 \\ -5.76033 & -4.76033 \\ 0 & -1 \\ -6.26033 & -6.26033 \\ 0 & 1 \end{bmatrix},$$

$$X^{(22)} = \begin{bmatrix} 1 & -0.0798679 \\ -1 & -0.0798679 \end{bmatrix} \text{ e } Y^{(22)} = \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix},$$

onde $\Lambda(X^{(22)}, Y^{(22)}) = \{\infty\}$ e $\Lambda(A, B) = \{1, 2, \infty\}$.

No caso de o feixe considerado ser o feixe companheiro $\lambda C_2 - C_1$ a aplicação do Algoritmo 4.2.2.1 é naturalmente viável mas dadas as características muito particulares deste tipo de feixe, na secção seguinte 4.2.9, é-lhe dedicada uma atenção especial.

4.2.9 Feixe Próprio do Feixe Companheiro $\lambda\mathcal{C}_2 - \mathcal{C}_1$

As caraterísticas particulares que irão considerar-se nesta secção prendem-se essencialmente com as dimensões, quer do feixe companheiro $\lambda\mathcal{C}_2 - \mathcal{C}_1$, quer do feixe próprio $\lambda Y - X$.

Isto é, pretende-se agora particularizar o Algoritmo 4.2.1 para o cálculo de feixes próprios de dimensão $n \times n$, do feixe companheiro $\lambda\mathcal{C}_2 - \mathcal{C}_1$ de dimensão $mn \times mn$, correspondente à linearização de um polinómio não mónico (2.2.6),

$$\lambda\mathcal{C}_2 - \mathcal{C}_1 = \begin{bmatrix} \lambda I_n & -I_n & \cdots & 0 \\ 0 & \lambda I_n & -I_n & \vdots \\ \vdots & \vdots & \lambda I_n & -I_n \\ A_m & A_{m-1} & \cdots & \lambda A_0 + A_1 \end{bmatrix}.$$

Segundo a Definição 4.2.1, $\lambda Y - X$ é um feixe próprio de dimensão $n \times n$ do feixe companheiro $\lambda\mathcal{C}_2 - \mathcal{C}_1$ se existirem vetores $V, W \in \mathbb{C}^{nm \times n}$ de caraterística máxima tais que

$$(\lambda\mathcal{C}_2 - \mathcal{C}_1)V = W(\lambda Y - X). \quad (4.2.11)$$

Considerando-se ainda a normalização escolhida em (4.2.3) e, uma vez que a dimensão do feixe próprio n divide a dimensão do feixe companheiro mn , os vetores $V, W \in \mathbb{C}^{nm \times n}$ podem ser particionados em m blocos de dimensão $n \times n$ tal como se apresenta a seguir

$$V = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_{m-1} \\ V_m \end{bmatrix} \quad \text{e} \quad W = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_{m-1} \\ W_m \end{bmatrix}, \quad (4.2.12)$$

onde $V_1 = W_1 = I_n$.

Partindo deste pressuposto (4.2.12) e não tendo sido possível encontrar qualquer resultado que generalizasse o Teorema 2.4.1 (onde se relaciona solventes com bloco valores propios da matriz companheira, ver Observação 2.4.9), é apresentado e demonstrado o seguinte Teorema que estabelece uma relação entre solventes de um polinómio matricial e feixes próprios do respetivo feixe companheiro.

Teorema 4.2.10. A matrix $X \in \mathbb{C}^{n \times n}$ é um solvente do polinómio

$$P(X) = A_0 X^m + A_1 X^{m-1} + \cdots + A_{m-1} X + A_m$$

se e só se o feixe $\lambda I - X$ for um feixe próprio do respetivo feixe companheiro $\lambda\mathcal{C}_2 - \mathcal{C}_1$ onde os vetores V e W referidos são da forma

$$V = \begin{bmatrix} I_n \\ X \\ \vdots \\ X^{m-2} \\ X^{m-1} \end{bmatrix} \quad \text{e} \quad W = \begin{bmatrix} I_n \\ X \\ \vdots \\ X^{m-2} \\ A_0 X^{m-1} \end{bmatrix}. \quad (4.2.13)$$

Demonstração. De acordo com a caracterização considerada em (4.2.12) para os vetores V e W , a equação (4.2.11), definida em termos dos m blocos de dimensão $n \times n$, é equivalente a

$$\begin{bmatrix} \lambda V_1 - V_2 \\ \lambda V_2 - V_3 \\ \vdots \\ \lambda V_{m-1} - V_m \\ A_m V_1 + A_{m-1} V_2 + \cdots + (\lambda A_0 + A_1) V_m \end{bmatrix} = \begin{bmatrix} W_1(\lambda Y - X) \\ W_2(\lambda Y - X) \\ \vdots \\ W_{m-1}(\lambda Y - X) \\ W_m(\lambda Y - X) \end{bmatrix}.$$

Isto é,

$$\begin{aligned} V_1 &= W_1 Y \quad \text{e} \quad V_2 = W_1 X; \\ V_2 &= W_2 Y \quad \text{e} \quad V_3 = W_2 X; \\ &\vdots \qquad \qquad \qquad \vdots \\ V_{m-1} &= W_{m-1} Y \quad \text{e} \quad V_m = W_{m-1} X; \\ A_0 V_m &= W_m Y \quad \text{e} \quad A_m V_1 + A_{m-1} V_2 + \cdots + A_1 V_m = -W_m X. \end{aligned}$$

Considerando a normalização (4.2.3), tem-se que $V_1 = W_1 = I_n$ o que também obriga a que $Y = I_n$. Nestas condições, tem-se

$$V_j = W_j \quad \text{e} \quad V_{j+1} = W_j X, \quad \forall j = 1, \dots, m-1,$$

e $A_0 V_m = W_m$, pelo que

$$V = \begin{bmatrix} I_n \\ X \\ \vdots \\ X^{m-2} \\ X^{m-1} \end{bmatrix}, \quad W = \begin{bmatrix} I_n \\ X \\ \vdots \\ X^{m-2} \\ A_0 X^{m-1} \end{bmatrix} \quad \text{e} \quad Y = \begin{bmatrix} I_n \end{bmatrix}. \quad (4.2.14)$$

Assim, a equação (4.2.11) depende apenas da matriz X e é equivalente a

$$A_m I_n + A_{m-1} X + \cdots + A_1 X^{m-1} + A_0 X^m = 0,$$

isto é, X é um solvente de $P(X)$. ■

Uma vez mais, tendo em conta o Teorema 4.2.10 e a exemplo do que pode ser feito com o método da Potência, em que se usam os blocos valores próprios para obter solventes, o método (NVFP), Algoritmo 4.2.1, quando aplicado ao feixe companheiro $\lambda \mathcal{C}_2 - \mathcal{C}_1$, é ainda um método viável no cálculo de solventes de $P(X)$.

Com o intuito de clarificar esta ideia e mostrar as potencialidades deste método quando aplicado no cálculo de um solvente, é apresentado o exemplo seguinte que resulta da linearização de um polinômio não mónico de grau 5.

Exemplo 4.2.11. Considere-se o polinômio matricial

$$P(x) = \begin{bmatrix} 2 & -10 \\ 4 & -20 \end{bmatrix} X^5 + \begin{bmatrix} -20 & 10 \\ -5 & -35 \end{bmatrix} X^4 + \begin{bmatrix} 120 & -220 \\ 110 & 450 \end{bmatrix} X^3 + \begin{bmatrix} -100 & 1700 \\ -850 & -2650 \end{bmatrix} X^2 + \\ + \begin{bmatrix} -1006 & -5390 \\ 2695 & 7079 \end{bmatrix} X + \begin{bmatrix} 1950 & 5790 \\ -2895 & -6735 \end{bmatrix},$$

e

$$\lambda \mathcal{C}_2 - \mathcal{C}_1 = \begin{bmatrix} \lambda & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2\lambda & -10\lambda & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4\lambda & -20\lambda & 0 \end{bmatrix} - \\ - \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -1950 & -5790 & 1006 & 5390 & 100 & -1700 & -120 & 220 & 20 & -10 & 0 \\ 2895 & 6735 & -2695 & -7079 & 850 & 2650 & -110 & -450 & 5 & 35 & 0 \end{bmatrix},$$

a linearização de $P(X)$ definida em (2.2.6), com $m = 5$, $n = 2$ e $\mathcal{C}_1, \mathcal{C}_2 \in \mathbb{C}^{10 \times 10}$.

Aplicando o Algoritmo (4.2.1) considerando $X^{(0)} = Y^{(0)} = I_2$ e

$$\bar{V}^{(0)} = \bar{W}^{(0)} = \begin{bmatrix} I_2 \\ mn \times nU \end{bmatrix},$$

onde onde $U \in \mathbb{C}^{(10-2) \times 2}$ é a matriz com entradas iguais a 1, obtém-se a seguinte sucessão

$$X^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ e } Y^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \\ \bar{V}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 20 & 20 \\ 20 & 20 \\ 20 & 20 \\ 20 & 20 \\ 20 & 20 \\ 20 & 20 \\ 20 & 20 \\ 20 & 20 \\ 20 & 20 \end{bmatrix} \text{ e } \bar{W}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 20 & 20 \\ 20 & 20 \\ 20 & 20 \\ 20 & 20 \\ 20 & 20 \\ 20 & 20 \\ 20 & 20 \\ 20 & 20 \\ 20 & 20 \end{bmatrix}, \\ X^{(1)} = \begin{bmatrix} 0.676745 & -0.628528 \\ 0.322967 & 1.62621 \end{bmatrix} \text{ e } Y^{(1)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\bar{V}^{(1)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0.676745 & -0.628528 \\ 0.322967 & 1.62621 \\ 0.670995 & -0.674987 \\ 0.317218 & 1.57975 \\ 0.665245 & -0.721445 \\ 0.311468 & 1.53329 \\ 0.659495 & -0.767904 \\ 0.305718 & 1.48683 \end{bmatrix} \text{ e } \bar{W}^{(1)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0.676745 & -0.628528 \\ 0.322967 & 1.62621 \\ 0.670995 & -0.674987 \\ 0.317218 & 1.57975 \\ 0.665245 & -0.721445 \\ 0.311468 & 1.53329 \\ -1.73819 & -16.4041 \\ -3.47637 & -32.8082 \end{bmatrix},$$

⋮

$$X^{(7)} = \begin{bmatrix} 1.89157 & 1.96289 \\ 0.199601 & 1.34104 \end{bmatrix} \text{ e } Y^{(7)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\bar{V}^{(7)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1.89157 & 1.96289 \\ 0.199601 & 1.34104 \\ 3.96983 & 6.34527 \\ 0.645231 & 2.19018 \\ 8.77574 & 16.3016 \\ 1.65766 & 4.20364 \\ 19.8537 & 39.0869 \\ 3.97463 & 8.89105 \end{bmatrix} \text{ e } \bar{W}^{(7)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1.89157 & 1.96289 \\ 0.199601 & 1.34104 \\ 3.96983 & 6.34527 \\ 0.645231 & 2.19018 \\ 8.77574 & 16.3016 \\ 1.65766 & 4.20364 \\ -0.0388348 & -10.7367 \\ -0.0776697 & -21.4733 \end{bmatrix},$$

com

$$\|z^{(7)}\| = \|\text{vec}(\mathcal{C}_1 \bar{V}^{(7)} - \bar{W}^{(7)} X^{(7)}, \mathcal{C}_2 \bar{V}^{(7)} - \bar{W}^{(7)} Y^{(7)})\|_2 < 10^{-5}.$$

A sucessão de matrizes obtida, $X^{(k)}, Y^{(k)}, \bar{V}^{(k)}$ e $\bar{W}^{(k)}$, satisfaz as condições (4.2.14) para cada iteração $k = 1, 2, \dots, k, \dots$, isto é,

$$\bar{V}^{(k)} = \begin{bmatrix} I_n \\ X^{(k)} \\ \vdots \\ (X^{(k)})^{m-2} \\ (X^{(k)})^{m-1} \end{bmatrix}, \bar{W}^{(k)} = \begin{bmatrix} I_n \\ X^{(k)} \\ \vdots \\ (X^{(k)})^{m-2} \\ A_0(X^{(k)})^{m-1} \end{bmatrix} \text{ e } Y^{(k)} = \begin{bmatrix} I_n \end{bmatrix}.$$

Tem-se ainda que a matriz,

$$X = X^{(7)} = \begin{bmatrix} 1.89157 & 1.96289 \\ 0.199601 & 1.34104 \end{bmatrix},$$

de acordo com o Teorema 4.2.10, é um solvente do polinômio $P(X)$ uma vez que o feixe

$$\lambda I - X = \begin{bmatrix} \lambda - 1.89157 & -1.96289 \\ -0.199601 & \lambda - 1.34104 \end{bmatrix},$$

é um feixe próprio do respetivo feixe companheiro, $\lambda\mathcal{C}_2 - \mathcal{C}_1$, com $\bar{V}, \bar{W} \in \mathbb{C}^{10 \times 2}$ dados por,

$$\bar{V} = \begin{bmatrix} I_n \\ X \\ \vdots \\ X^{m-2} \\ X^{m-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1.89157 & 1.96289 \\ 0.199601 & 1.34104 \\ 3.96983 & 6.34527 \\ 0.645231 & 2.19018 \\ 8.77574 & 16.3016 \\ 1.65766 & 4.20364 \\ 19.8537 & 39.0869 \\ 3.97463 & 8.89105 \end{bmatrix} \quad \text{e} \quad \bar{W} = \begin{bmatrix} I_n \\ X \\ \vdots \\ X^{m-2} \\ A_0 X^{m-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1.89157 & 1.96289 \\ 0.199601 & 1.34104 \\ 3.96983 & 6.34527 \\ 0.645231 & 2.19018 \\ 8.77574 & 16.3016 \\ 1.65766 & 4.20364 \\ -0.0388348 & -10.7367 \\ -0.0776697 & -21.4733 \end{bmatrix}.$$

Para outra aproximação inicial, $X^{(0)} = Y^{(0)} = 10I_2$ e

$$\bar{V}^{(0)} = \bar{W}^{(0)} = \begin{bmatrix} I_2 \\ 1.13mn \times nU \end{bmatrix},$$

onde $U \in \mathbb{C}^{(10-2) \times 2}$ é a matriz com entradas iguais a 1, a sucessão obtida converge ainda em 10 iterações para o mesmo resultado.

Comparando este resultado com aquele obtido no Exemplo 4.1.13, secção 4.1.12, observa-se que para a mesma aproximação inicial o método atual converge mais depressa (10 iterações) do que o método usado então, Algoritmo 4.1.2, que convergiu para o mesmo valor próprio X em 12 iterações. Quando comparados em termos do tempo (medido em *seg.* através do comando `Timing []` do Mathematica) gasto no cálculo das aproximações, $X^{(10)}$ e $X^{(12)}$, observa-se que o método atual consome mais tempo de CPU (0.406 s em vez dos 0.125 s usados pelo Algoritmo 4.1.2).

Da mesma forma, utilizando-se o método de Newton Vetorial (NV), Algoritmo 3.3.1 definido na secção 3.3.3, para a aproximação inicial $x^{(0)} = \text{vec}(10I_2)$ equivalente, obtém-se ainda a convergência deste mas necessitando de mais iterações (13 iterações) e usando mais tempo de CPU (0.454 s).

Apesar de, neste exemplo, o método Newton Vetorial (NV) ter sido menos vantajoso do que os Algoritmos 3.3.1, 4.1.2 e 4.2.1, uma vez que necessita mais iterações e mais tempo de CPU para calcular a mesma aproximação, só um estudo mais aprofundado (explorando estes resultados para polinómios de maior dimensão e maior grau) permitiria uma conclusão mais fundamentada o que, por outro lado, obrigaria a dispor de outros meios informáticos.

Capítulo 5

Conclusões

Dentro dos objetivos inicialmente propostos, foi possível alcançar:

Na primeira parte implementaram-se os seguintes métodos para polinômios matriciais.

- Método do Ponto Fixo, seção 3.2, cuja concepção consiste em resolver cada equação $\text{vec}(P(X))_l = 0$, do sistema $\text{vec}(P(X)) = 0$, em ordem à maior potência de x_l , coordenada com o mesmo índice do vetor $x = \text{vec}(X)$, resultando na equação $x = f(x)$ e no Algoritmo 3.2.1. A convergência é estabelecida com base no Teorema do ponto fixo entre espaços de Banach de Schauder e na respetiva estabilidade assintótica (ver Shih [35], pág. 144), também conhecida por conjectura de Belitskiĭ e Lyubich;
- Método de Newton Vetorial, seção 3.3, resulta da aplicação do método de Newton clássico à equação vetorial $F(x) = \text{vec}(P(X)) = 0$, onde a derivada é formulada em termos de matriz jacobiana evitando-se o uso da derivada de Frechét e a resolução da equação de Sylvester em cada iteração, tal como acontece no método definido em Higham [18], pág. 4. O Método de Newton Vetorial está definido através do Algoritmo 3.3.1 e a sua convergência fica estabelecida com o Teorema de Kantorovich, válido para uma função entre espaços de Banach.

Como foi mencionado na introdução, relativamente a métodos iterativos com aplicação no cálculo de blocos valores próprios, após consulta da diversa documentação e bibliografia existente, apenas se conseguiu encontrar o método da Potência (Dennis [9], pág. 83), estando este definido unicamente para um feixe do tipo $(A, I) = \lambda I - A$. Foi na sequência desta lacuna que se adaptou a mesma formulação ao cálculo de blocos valores próprios $\mathcal{C}_1 V = \mathcal{C}_2 V X$ do feixe $(\mathcal{C}_1, \mathcal{C}_2)$ ou, de uma forma geral, de um feixe qualquer (A, B) .

Assim, numa segunda parte, a construção usada na concepção do Algoritmo 3.3.1, foi generalizada para o contexto quer dos blocos valores próprios quer dos feixes próprios:

- Método Newton Vetorial para Blocos Valores Próprios definido para um feixe genérico $(A, B) = \lambda B - A$, Algoritmo 4.1.1;
- Método de Newton Vetorial para Feixes Próprios, definido para um feixe genérico $(A, B) = \lambda B - A$ através do Algoritmo 4.2.1.

Contudo, relativamente a cada método apresentado, não foi possível definir um critério de escolha para a aproximação inicial (por ex. a partir das matrizes A_i coeficientes do polinómio $P(X)$, ou a partir das matrizes A ou B de um feixe (A, B)). De uma forma geral, seguindo o critério usado em [21] e [28], entre outros, essa escolha consistiu em considerar matrizes escalares ou múltiplos de alguma matriz coeficiente A_i , no caso dos Métodos do Ponto Fixo e Newton Vetorial, ou múltiplos da matriz com entradas iguais a 1 no caso dos Métodos Newton Vetorial para Blocos Valores Próprios e Newton Vetorial para Feixes Próprios, mas sem que se conseguisse estabelecer um critério rigoroso que assegurasse à partida a convergência destes (exceção feita às aproximações iniciais consideradas nos Teoremas de convergência, difíceis de obter do ponto de vista prático uma vez que é necessário conhecer previamente a solução exata).

Um conjunto completo de blocos valores próprios (ou um conjunto completo de solventes) pode ser obtido com a utilização de uma deflação (Block Wielandt Deflation ou Block Hotelling Deflation). Sempre que o polinômio tenha valores próprios infinitos, há que determinar blocos valores próprios quer relativamente à equação $BV = AVX$ quer em relação a $AV = BVX$. Ou seja há que resolver dois sistemas e identificar $2nm \times m + 2m^2$ variáveis.

5.1 Trabalho Futuro

A formulação usada na construção dos método Método do Ponto Fixo e de Newton Vetorial, resultante da transformação da equação polinomial matricial original numa equação vetorial equivalente com base no método dos produtos de Kronecker, é extensível aos restantes métodos iterativos próprios dos sistemas de equações não lineares, nomeadamente o método de Broyden que consiste na generalização do método da secante para sistemas de equações não lineares.

Ainda no que se refere ao método de Newton Vetorial, sempre que, nalguma iteração, a matriz $J_F(x^{(k)})$ for singular então $\text{rank}(J) < n^2$ e o método é obrigado a parar.

Uma alternativa consiste na utilização da inversa generalizada J^\dagger de $J_F(x^{(k)})$: se o sistema $J_F(x^{(k)})Q_k = -F(x^{(k)})$ for impossível, $J^\dagger F(x^{(k)})$ representa o mínimo de $\|J_F(x^{(k)})Q + F(x^{(k)})\|$; se sistema $J_F(x^{(k)})Q_k = -F(x^{(k)})$ for indeterminado, neste caso, $J^\dagger F(x^{(k)})$ representa o mínimo de $\|Q\|$ quando sujeito à condição $J_F(x^{(k)})Q + F(x^{(k)}) = 0$.

Segundo a Definição 2.4.8, um bloco vetor próprio $V \in \mathbb{C}^{m \times n}$ deverá ter característica máxima n . Na conceção do método Newton Vetorial para Blocos Valores Próprios, definido no capítulo 4, Algoritmo 4.1.1, consideraram-se blocos vetores normalizados da forma $\bar{V} = [I_n \quad V_2]^T$. Isto é, partiu-se do princípio que as primeiras n linhas do bloco são linearmente independentes o que, em certa medida, corresponde a considerar um caso particular de blocos valores próprios não englobando portanto todas as possibilidades (ver Observação 4.1.1, o fato de a matriz V ter característica máxima implica que tem n linhas linearmente independentes mas não têm que ser necessariamente as primeiras n linhas). Permanece a convicção que, num trabalho futuro, é possível alargar esta construção para um bloco vetor próprio qualquer, ou seja, um bloco vetor próprio em que as n linhas linearmente independentes não sejam necessariamente as primeiras.

Da mesma forma, na construção do método de Newton Vetorial para Feixes Próprios, secção 4.2, Algoritmo 4.2.1, onde foram considerados vetores V e W normalizados na forma $\bar{V} = [I_n \quad V_2]^T$ e $\bar{W} = [I_m \quad W_2]^T$, pensa-se também ser viável alargar este método para qualquer bloco vetor próprio.

Bibliografia

- [1] Ioannis K. Argyros: A new Kantorovich-Type Theorem for Newton's method, *Applicationes Mathematicae*, 26,2, 151-157 (1999). 43
- [2] Ioannis K. Argyros: On the Convergence of a Modified Newton Method for Solving Equations, *Journal of Mathematics*, Punjab University, 41, 11-21 (2009). 43
- [3] N. Cohen: Spectral analysis of regular matrix polynomials, *Integral Equations and Operator Theory*, Birkhäuser Verlag, Basel Vol.6 ISBN 978-0-898716-85-6, (1983). 22
- [4] B. Nath Datta: Numerical Linear Algebra and Applications, SIAM, ISBN 978-0-898716-85-6, (2009). 6, 45
- [5] G. J. Davis: Numerical solution of a quadratic matrix equation, *SIAM Journal of Scientific Computing*, 2, 164-175, (1981). vii, xi
- [6] P. V. Dooren and P. Dewilde: The Generalized Eigenstructure Problem in Linear System Theory, *IEEE Transactions on Automatic Control*, AC-26, 1 111-129 (1981). 1
- [7] P. V. Dooren and P. Dewilde: The Eigenstructure of an Arbitrary Polynomial Matrix, Computational Aspects, *Linear Algebra and its Applications*, Elsevier Science Publishing Co. Inc., 50 545-579 (1983). 5, 11, 17, 25
- [8] B. K. Driver: *Analysis Tools with Applications*, Springer. (2003). 45
- [9] E. Dennis, J. F. Traub and R. P. Weber: On the Matrix Polynomial, Lambda-Matrix and Block Eigenvalue Problems, Computer Science Department, Technical Report, Cornell University, Ithaca, New York and Carnegie-Mellon University, Pittsburgh, Pennsylvania, (1971). vii, xi, 5, 6, 7, 9, 26, 28, 32, 55, 62, 91, 97, 98
- [10] J. E. Dennis, J. F. Traub and R. P. Weber: The algebraic theory of matrix polynomials, *SIAM Journal of Numerical Analysis*, 13, 831-845, (1976). vii, xi, 6, 13, 32
- [11] J. E. Dennis, J. F. Traub and R. P. Weber: Algorithms for solvents of matrix polynomials, *SIAM Journal of Numerical Analysis*, 15, 523-533, (1978). vii, xi, 6, 27, 32, 33, 34, 97
- [12] F. Gantmacher: *The Theory of Matrices*, Vol I, Chelsea, New York. (1960). 1, 4, 5, 10, 17
- [13] F. Gantmacher: *The Theory of Matrices*, Vol II, Chelsea, New York. (1960). 10, 11
- [14] V. Hernandez Garcia: Resolucion de Ecuaciones Matriciales Algebraicas y Diferenciales Mediante la Eliminacion del Caracter no Unilateral, Facultad de Ciencias Matematicas, Universidad de Valencia (1979). 2, 7, 8, 9, 15, 31
- [15] I. Gohberg, P. Lancaster and L. Rodman: *Matrix Polynomials*, Academic Press, New York, (1982). 2, 4, 5, 6, 9, 10, 11, 12, 13, 15, 16, 18, 19, 20, 22
- [16] J. M. Gutiérrez: A new semilocal convergence theorem for Newton's method, *Journal of Computational and Applied Mathematics*, Elsevier, 79, 131-145 (1997). 43
- [17] J. M. Gutiérrez: Newton-Kantorovich theory and some of its variations, *Basque Center for Applied Mathematics*, (2009). 43, 45, 59, 78

- [18] N. J. Higham and H. M. Kim: Solving a quadratic matrix equation by Newton's method with exact line searchers, *SIAM Journal of Matrix Analysis and Applications*, 23, 303-316, (2001). vii, xi, 6, 7, 8, 31, 41, 91
- [19] N. J. Higham and H. M. Kim: Numerical analysis of a quadratic matrix equation, *IMA Journal of Numerical Analysis*, 20, 499-519, (2000). vii, xi, 6, 31, 32
- [20] R. A. Horn, C. R. Johnson: *Topics in matrix analysis*, Cambridge University Press, (1991). 1, 2, 9, 14, 15, 17
- [21] W. Kratz and E. Stichel: Numerical solution of matrix polynomial equations by Newton's method, *IMA Journal of Numerical Analysis*, 7, 355-369, (1987). vii, xi, 6, 8, 31, 91, 97
- [22] P. Lancaster: A fundamental theorem on lambda matrices with applications - II. Difference equations with constant coefficients, *Linear Algebra and its applications*, 18, 213-222, (1977). vii, xi, 6, 31, 32
- [23] P. Lancaster: *Lambda-Matrices and Vibrating Systems*, Pergamon Press, New York, (1966). 2
- [24] P. Lancaster, L. Rodman: *The Algebraic Riccati Equation*, Oxford University Press, Oxford, (1995). 17
- [25] P. Lancaster, M. Tismenetsky: *The Theory of Matrices*, 2nd edition, Academic Press, New York, (1985). vii, xi, 4, 5, 6, 9, 10, 12, 15, 26, 27
- [26] Jian-hui Long, Xi-yan Hu e Lei Zhang: Improved Newton's method with exact line searches to solve quadratic matrix equation, *Journal of Computational and Applied Mathematics*, 645, 645-654, (2008). vii, xi, 39, 41, 48, 52
- [27] F. Marcos, E. Pereira: A Fixed Point Method to Compute Solvents of Matrix Polynomials, *Math. Bohem.* 135, No. 4, 355-362 (2010). vii, xi, 7
- [28] F. Marcos, E. Pereira, P. Rebelo: A Scalar Newton's method to Compute Solvents of Matrix Polynomials, Preprint, (2012). 7, 91
- [29] T. Grassó Matoses: *Sistemas Diferenciales Lineales Implícitos y Factorización de Matrices Polinomiales*, Facultad de Matematicas, Universidad de Valencia (1991). 1, 9, 10, 14, 15
- [30] E. Pereira: Solventes de Polinômios Matriciais Mônicos, Departamento de Matemática e Informática, Universidade da Beira Interior, Covilhã, (2000). 5, 12, 18, 19, 27, 28, 29
- [31] E. Pereira: Block Eigenvalues and Solutions of Differential Matrix Equations, *Mathematical Notes, Miskolc*, 4, 45-51 (2003). 60, 68
- [32] E. Pereira and C. Rosa: Deflation Method for a Regular Matrix Pencils, *Applied Mathematics and Computation*, Elsevier, 218, 2913-2920 (2011). 31, 68, 73, 75, 79
- [33] E. Pereira, R. Serodio and J. Vitória: Newton's method for matrix polynomials, *Differential Equations and Applications*, Nova Science Publishers, Inc., New York, 5, 107-112, (2007). vii, xi, 2, 6, 31
- [34] E. Pereira and J. Vitória: Deflation of block eigenvalues of block partitioned matrices with an application to matrix polynomials of commuting matrices, *Computers and Mathematics with Applications*, 42, 1177-1188, (2001). vii, xi, 31

- [35] M. Shih and J. Wu: Asymptotic Stability in the Schauder fixed point theorem, *Studia Mathematica*, 2, 143-148, (1998). 7, 36, 91
- [36] F. Tisseur and K. Meerbergen: The quadratic eigenvalue problem, *SIAM Review*, 436, 235-286 (1988). 2, 4, 5, 6, 9, 10, 11, 12, 15, 16
- [37] J. F. Traub: The calculation of zeros of polynomials and analytic functions, *Mathematical Aspects of Computer Science*, Providence, R. I., 19 138-152. (1967). 6, 31
- [38] J.S.H. Tsai, L.S. Shieh and T.T.C. Shen: Block Power method for computing solvents and spectral factors of matrix polynomials, *Computers and Mathematics with Applications*, 16, 683-699 (1988). vii, xi, 6, 31, 32
- [39] J. H. Wilkinson: *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford University Press. (1965) 11
- [40] X. Zhang: Calculating the eigenstructure of a regular matrix pencil - an approach based on the Weierstrass form, *International Mathematical Forum Harbin Institute of Technology*, PO Box 416, Harbin, 150001, P. R. China, (2006). 74
- [41] Bin Zhou, Zhao-Yan Li, Guang-Ren Duan e Yong Wang: Solutions to a family of matrix equations by using the Kronecker matrix polynomials, *Applied Mathematics and Computation*, 212, 327-336, (2009). 41

Apêndice A

Anexos

A.1 Algoritmos dos métodos referidos

O Algoritmo do método de Newton (ver Kratz [21], pág. 359),

Algoritmo A.1.1 MN

- 1: É dada uma aproximação inicial $X^{(0)}$ e uma margem de erro $\varepsilon > 0$.
 - 2: para $k = 1, 2, \dots$ fazer
 - 3: se $\|P(X^{(k)})\| < \varepsilon$ então
 - 4: o método pára.
 - 5: caso contrário
 - 6: $Q^{(k)}$ solução de $P'(X^{(k)})[Q^{(k)}] = -P(X^{(k)})$;
 - 7: $X^{(k+1)} = X^{(k)} + Q^{(k)}$;
-

O Algoritmo do método Bernoulli, (ver Dennis [9] e [11], pág. 69 e 529),

Algoritmo A.1.2 MB

- 1: Dadas as matrizes iniciais $X^{(0)} = X^{(1)} = \dots = X^{(m-2)} = 0$ e $X^{(m-1)} = I_n$,
 - 2: para $k = m - 1, m, \dots$ definem-se as matrizes X_k por fazer
 - 3: $S = X^{(k)}(X^{(k-1)})^{-1}$;
 - 4: se $\|P(S)\| < \varepsilon$ então
 - 5: o método pára.
 - 6: caso contrário
 - 7: $X^{(k+1)} = -A_1X^{(k)} - \dots - A_mX^{(k-m+1)}$.
 - 8: $X = S$
-

O Algoritmo do método de Traub (ver Dennis [11], pág. 524),

Algoritmo A.1.3 MT

- 1: Dada uma aproximação inicial $X^{(0)}$, l e uma margem de erro $\varepsilon > 0$.
 - 2: $G_0(X) = I_n = \Gamma_1^0 X^{m-1} + \dots + \Gamma_m^0 X^0$
 - 3: para $j = 0, \dots, l - 1$, fazer
 - 4: $G_j(X) = \sum_{i=1, \dots, m} \Gamma_i^j X^{m-i}$;
 - 5: $G_{j+1}(X) = XG_j(X) - \Gamma_1^j P(X)$;
 - 6: $\varphi_l(X) = G_l(X)G_{l-1}^{-1}(X)$;
 - 7: para $k = 0, \dots$, fazer
 - 8: se $\|P(X^{(k)})\| < \varepsilon$ então
 - 9: o método pára.
 - 10: caso contrário
 - 11: $X^{(k+1)} = \varphi_l(X^{(k)})$.
-

O Algoritmo do método da Potência (ver Dennis [9], pág. 83),

Algoritmo A.1.4 (MP)

- 1: Dados $A, X^{(0)}, V^{(0)}$, e $\varepsilon > 0$.
 - 2: para $k = 0, 1, \dots$ fazer
 - 3: se $\|V^{(k)}X^{(k)} - AV^{(k)}\| < \varepsilon$ então
 - 4: o método pára.
 - 5: caso contrário
 - 6: $W^{(k+1)} = AV^{(k)}$;
 - 7: $V^{(k+1)} = W^{(k+1)} ((W^{(k+1)})_j)^{-1}$;
 - 8: $X^{(k+1)} = (AV^{(k+1)})_j$
 - 9: $V = V^{(k)}$ e $X = (AV^{(k)})_j$ tais que $AV = VX$.
-

Glossário

\mathbb{R}	Corpo dos números reais.
\mathbb{C}	Corpo dos números complexos.
\mathbb{K}	Corpo \mathbb{R} ou \mathbb{C} .
$\mathbb{C}^{n \times m}$	Espaço vetorial das matrizes $n \times m$ com entradas complexas.
I_n	Matriz identidade em $\mathbb{C}^{n \times n}$.
$0_{m \times n}$	Matriz nula em $\mathbb{C}^{m \times n}$.
$X \otimes Y$	Produto de Kronecker das matrizes X e Y .
$vec(X)$	Função vetor de uma matriz $X \in \mathbb{C}^{m \times n}$.
\mathcal{C}	Matriz companheira.
$P(X)$	Polinômio matricial.
$D_f(X)[H]$	(Ou $= f'(X)[H]$) Derivada de Frechét de f em X na direção de $H \in \mathbb{C}^{m \times n}$.
$f'_{(x)}(h)$	Derivada de f em x na direção de $h \in \mathbb{C}^n$.
$P(\lambda)$	λ -Matriz.
(A, B)	Feixe $\lambda B - A$.
$\lambda\mathcal{C}_2 - \mathcal{C}_1$	Feixe companheiro.
$J_k(\lambda_i)$	Um bloco de Jordan de dimensão $k \times k$ com λ_i na diagonal principal.
J_A	Forma canônica de Jordan da matriz A .
$K_{(A,B)}$	Forma canônica de Weierstrass do feixe (A, B) .
$\mathcal{V}(X_1, \dots, X_m)$	Matriz bloco de Vandermonde dos solventes X_1, \dots, X_m .
$\text{Diag}(J_1, \dots, J_l)$	Matriz bloco diagonal com as matrizes J_1, \dots, J_l na diagonal principal.
$J_1 \oplus J_2$	Matriz soma direta das matrizes J_1 e J_2 (o mesmo que $\text{Diag}(J_1, J_2)$).
$\Lambda(P)$	Espectro da λ -Matriz $P(\lambda)$.
$P(\lambda) \sim Q(\lambda)$	λ -Matrizes equivalentes. A λ -Matriz $P(\lambda)$ é equivalente a $Q(\lambda)$
id	Função identidade $id : \mathbb{K} \rightarrow \mathbb{K}$.
$mod(m, n)$	O resto da divisão de m por n , $m, n \in \mathbb{N}$.
$quoc(m, n)$	O quociente da divisão de m por n , $m, n \in \mathbb{N}$.

Índice Remissivo

- λ -matriz, 1, 9
 - associada, 10
 - cadeia de Jordan da, 9, 26
 - definição, 9
 - determinante, 11
 - equivalente, 11
 - fatorização, 13
 - regular, 10
 - valor próprio de, 7, 10, 22, 28
 - valor próprio infinito, 11
- Bloco valor próprio, 1, 5–8, 28, 55, 56, 61–63, 66, 67, 69, 71–75, 91, 92
 - conjunto completo, 31
 - da matriz companheira, 28, 55, 64
 - definição, 28
- Bloco vetor próprio, 28, 55, 61–64, 66, 67, 69, 71, 72, 92
 - definição, 28
- Companheira
 - bloco valor próprio da matriz, 28, 55, 64
 - blocos valores próprios da matriz, 5
 - definição, matriz, 15
 - forma canônica de Jordan da matriz, 19
 - matriz, 4–7, 15, 17–19, 26, 31, 32, 64, 65
- Espectro, 27, 73
 - definição, 11
- Feixe, 7, 66, 91
 - bloco valor próprio, 75, 85
 - bloco vetor próprio, 67
 - companheiro, 6, 7, 16–18, 21, 31
 - companheiro, definição, 67
 - definição, 10
 - forma canônica, 74
 - forma Canônica de Weierstrass, 5
 - próprio, 7, 8, 85, 91
 - próprio, definição, 75
 - regular, 17
 - valor próprio do, 14
- Feixe Companheiro, 86
- Fréchet
 - derivada, 31, 32, 36, 40, 41, 52
- Jordan
 - bloco de, 16–18, 28
 - cadeia de, 9, 11–13, 18, 22, 26, 28
 - cadeia de, multiplicidade parcial, 12, 13
 - forma canônica de, 4, 17, 19, 74
- Lyapunov
 - equação de, 2
- Newton
 - método classico, 6, 31, 43
 - método classico, algoritmo, 97
 - método vetorial, 7, 8, 39, 49, 56, 58, 59, 91, 92
 - método vetorial para blocos valores próprios, 7, 8, 62, 65, 91, 92
 - método vetorial para blocos valores próprios, algoritmo, 56
 - método vetorial para blocos valores próprios, convergência, 56
 - método vetorial para feixes próprios, 7, 8, 75, 91, 92
 - método vetorial para feixes próprios, algoritmo, 76
 - método vetorial para feixes próprios, convergência, 78
 - método vetorial, algoritmo, 42
 - método vetorial, convergência, 43
- Polinômio matricial, 6, 7, 13, 32, 34, 40
 - definição, 9
 - derivada de Fréchet, 31
 - grau 2, 40
 - mônico, 10, 15, 19, 64
 - solvente dominante, 27
 - solventes, 5, 7, 27, 31
- Problema dos valores próprios quadrático, 2
- Produto de Kronecker, 8, 15, 31, 40
 - definição, 14
 - lema, 15
 - método dos, 7, 92
- Riccati
 - equação de, 2

Sylvester

equação de, vii, 2, 7, 8, 14, 32, 40, 51, 52,
91

Solvente, 1, 5–7, 9, 10, 13, 14, 26, 28, 31, 35, 37,
42, 48, 55

à direita, 10

à esquerda, 10

com coeficientes complexos, 38

conjunto completo de, 5, 10, 27, 31–33, 92

definição, 10

dominante, 6, 31–33, 39

dominante, definição, 27

e bloco valor próprio, 28, 64

existência e contabilização de, 5

relação com cadeia de Jordan, 26

valores próprios do, 13, 14

Valores próprios, 2, 11, 17–19, 21

da λ -Matriz, 13, 17

da matriz companheira, 28

de maior valor absoluto, 27

do bloco valor próprio, 28

do solvente, 13, 14, 27

finitos, 5, 11, 17, 22

infinito, 5, 92

problema dos, 2

Vandermonde

definição, matriz bloco de, 27

matriz bloco de, 5

não singular, 27

Vandermonde, matriz bloco de, 32

Vetores próprios, 2, 3, 5–7, 18

blocos, 8, 9, 26, 31, 67

generalizados, 9, 12

número de, 11

Weierstrass

forma canônica de, 5, 17, 81, 82

