

Universidade Federal de Santa Catarina
Curso de Pós-Graduação em Matemática e
Computação Científica

Problemas de Identificação Paramétrica

Maria Inez Cardoso
Orientador: Prof. Dr. Fermín S. V. Bazán

Florianópolis
Março de 2000

Universidade Federal de Santa Catarina
Curso de Pós-Graduação em Matemática e
Computação Científica

Problemas de Identificação Paramétrica

Dissertação apresentada ao Curso de Pós-Graduação em Matemática e Computação Científica, do Centro de Ciências Físicas e Matemáticas da Universidade Federal de Santa Catarina, para a obtenção do grau de Mestre em Matemática, com Área de Concentração em Álgebra Linear Computacional.

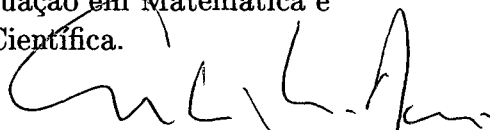
Maria Inez Cardoso
Florianópolis
Março de 2000

Problemas de Identificação Paramétrica

por

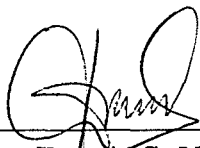
Maria Inez Cardoso

Esta Dissertação foi julgada para a obtenção do Título de “Mestre”,
Área de Concentração em Álgebra Linear Computacional, e aprovada em sua forma
final pelo Curso de Pós-Graduação em Matemática e
Computação Científica.

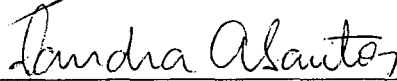


Celso Melchíades Dória
Coordenador

Comissão Examinadora



Prof. Dr. Fermán S. V. Bazán (UFSC-Orientador)



Prof. Dr. Sandra Augusta Santos (UNICAMP)



Prof. Dr. Mário Cesar Zambaldi (UFSC)



Prof. Dr. Antônio G. Leitão (UFSC)

Florianópolis, 30 de março de 2000.

A Deus
À minha mãe
A Daniel

Agradecimentos

Primeiramente gostaria de agradecer a minha Família, minha mãe, Dominicia, a quem dedico este trabalho, pois fez de mim, a principal razão da sua vida. Meu Pai, que tenho certeza, sempre olha por mim. Meu noivo, Daniel, pelo seu amor e pelo encorajamento nas horas difíceis. E meu padrinho, José Carlos, que é um irmão pra mim.

Ao meu orientador, Prof. Fermín S. V. Bazán pelos 5 anos de paciência, por me ensinar a pesquisar e a persistir, sua dedicação à pesquisa deve ser reconhecida.

Gostaria de agradecer também, aos colegas da “PG”, pelas longas horas de estudo e pelas horas de alegria nos estudos de “computação científica”. Em especial, Andresa e Claiton, que me acompanham desde a graduação, Dirceu, meu companheiro de sala, Rafael, pela ajuda com o Latex, Fábio, pelas discussões sobre Álgebra Linear e Oswaldo, que apesar muitas vezes não deixar ninguém estudar direito, sempre nos trouxe alegria.

Aos professores Carmen Suzane C. Gimenez, que fez despertar o gosto pela matemática, Neri Terezinha B. Carvalho, pelo incentivo na troca de curso, Lício H. Bezerra, pela motivação para entrar no PET, José Luiz Rosas Pinho e Ruy Exel Filho, pelas ótimas aulas de Cálculo e Análise Funcional, Mário César Zambaldi, pela iniciação na Álgebra Linear, Mírian Buss Gonçalves, pelo incentivo e encorajamento. E os demais professores que contribuíram com a minha formação. Agradeço também a Elisa, secretária da PG, que além de sempre nos atender prontamente, é uma boa companhia de almoço.

A Capes, pelo financiamento de meus estudos.

Finalmente, gostaria de agradecer a Deus, por ter me dado forças, energia e inteligência, e por colocar todas estas pessoas e oportunidades na minha vida.

Sumário

Resumo	1
Abstract	2
Introdução	3
1 Conceitos Fundamentais da Álgebra Linear	6
1.1 Notação	6
1.2 Autovalores e Autovetores	8
1.2.1 Condicionamento dos Autovalores	10
1.3 Decomposição em Valores Singulares (SVD)	12
1.3.1 Sensibilidade dos Valores Singulares	14
1.4 Pseudo Inversa	14
1.5 Problemas de Quadrados Mínimos Lineares	15
1.5.1 Problemas de Quadrados Mínimos Totais	16
1.6 Ângulo entre Subespaços	18
2 O Problema da Recuperação de Exponenciais: Apresentação e Propriedades	20
2.1 Recuperação de exponenciais e o Método de Prony	20
2.2 Algumas Propriedades	26
3 Métodos de Identificação Paramétrica	28
3.1 Técnicas de Predição Linear	28
3.2 Métodos de Subespaço	30
3.3 O Problema de Identificação em Situações Reais	34
4 Sensibilidade dos Autovalores do Sinal	35
4.1 Autovalores do Sinal	35
4.2 Condicionamento de Matrizes de Vandermonde	36
4.2.1 Estimativas Preliminares	36
4.2.2 Limitantes para $\kappa_2(W_M)$	39

4.3	Erro nos autovalores	41
4.4	Resultados Numéricos	45
5	Computação Rápida do Subespaço Sinal	49
5.1	Multiplicação Rápida Matriz-Vetor	49
5.2	O Procedimento de Rayleigh-Ritz	53
5.3	Iteração em Subespaços	55
5.4	Iteração em Subespaços com Aceleração de Ritz	56
5.5	Método de Lanczos	57
5.5.1	Método de Lanczos Básico	58
5.5.2	Método de Lanczos com Reortogonalização Completa	61
5.5.3	Lanczos com Semiortogonalização	62
5.5.4	Método de Lanczos com Recomeços Implícitos	62
5.6	Experiências Numéricas	74
6	Experimentos Numéricos	77
6.1	Sinal com 5 exponenciais proveniente de NMR	78
6.2	Sinal com 11 exponenciais proveniente de MRS	79
6.3	Sinal da Análise Modal experimental	82
	Conclusões	85
	Rereferências Bibliográficas	89

Resumo

Em muitas áreas das ciências aplicadas nos deparamos com o problema de estimar um conjunto de parâmetros a partir de dados obtidos experimentalmente, onde assume-se que os dados provém de modelos exponenciais. Esses dados são armazenados em matrizes, as quais servem como ponto de partida para uma série de algoritmos que extraem os parâmetros de interesse, via solução de um problema de autovalor. O objetivo deste trabalho é apresentar o problema de recuperação de exponenciais, as dificuldades associadas ao problema, e as diferentes maneiras de abordá-lo, bem como, experiências numéricas as quais ilustram a performance dos métodos apresentados.

Abstract

In many areas of applied sciences we face the problem of estimating a set of parameters from a given set of data points, which are assumed to satisfy exponential models. These data are stored in matrices, which serve as a startpoint to a lot of algorithms which extract the parameters of interest through the solution of an eigenvalue problem. The objective of this work is to present the problem of exponential recovering, as well as several algorithms used in solving this important problem. Particular emphasis is given to algorithms based on the fast computation of the so-called signal subspace, which are shown to yield accurate estimates of the wanted parameters. Numerical results illustrating the performance of the methods are presented, including the recovering of parameters from time domain NMR (Nuclear Magnetic resonance) data.

Introdução

O problema de recuperação de parâmetros embutidos em exponenciais complexas tem sido motivo de pesquisa desde aproximadamente o século XVIII. A primeira referência do problema vem de 1795, quando R. de Prony deparou-se com o problema ao realizar algumas experiências químicas. A ele também pode ser creditada a primeira solução. Apesar de possuir mais de 200 anos, o método desenvolvido por Prony tem servido de motivação para o desenvolvimento de outros métodos, entre os quais se destaca o método de Tufts e Kumaresan [22].

Atualmente, o problema de se estimar um conjunto de parâmetros associados a sinais descritos como uma combinação linear de exponenciais complexas, a partir de um conjunto de amostras do sinal, é motivo de pesquisa em diversas áreas, como processamento digital de voz, processamento de sinais, análise modal de estruturas mecânicas, ressonância magnética nuclear, etc., [1, 7, 11, 21, 17, 39]. Por exemplo, em análise modal de estruturas mecânicas, a determinação dos parâmetros fornece informações sobre as amplitudes, frequências naturais de vibração e amortecimentos da estrutura em estudo [1]; em ressonância magnética nuclear, fornecem informações como razões de concentração de metabolismo, valor do pH de tecidos, etc., as quais auxiliam no diagnóstico médico [8].

Em aplicações práticas, as amostras disponíveis do sinal são versões perturbadas do sinal exato. As perturbações são chamadas de ruídos e podem ser resultado de imprecisões numéricas, interferências térmicas, elétricas, atmosféricas, etc. Com isso, a determinação dos parâmetros desejados torna-se mais complicada.

A idéia básica das abordagens existentes para o problema é armazenar as amostras do sinal numa matriz, a qual possui uma estrutura predefinida, *Toeplitz* ou *Hankel*, que é chamada *matriz de observação*. Quando o sinal é livre de ruídos, ou seja, quando não apresenta nenhuma perturbação, o número de colunas linearmente independentes dessa matriz fornece informação sobre o número de componentes do sinal. Os parâmetros procurados podem ser extraídos a partir do subespaço gerado pelas colunas (linhas) linearmente independentes dessa matriz, pela solução de um problema de autovalor. Este subespaço é chamado *subespaço sinal*, e pode ser estimado por meio dos vetores singulares à direita ou à esquerda, associados aos valores singulares não nulos da matriz de observação.

Na prática porém, como as amostras do sinal são perturbadas, a matriz de observação se comporta como sendo de posto completo, ou seja, com todas as colunas linearmente independentes. Dessa forma, o seu espaço coluna (linha) possui informações provenientes do sinal e informações provenientes do ruído. Assim, a primeira dificuldade é separar este subespaço em dois, um subespaço associado ao sinal, e outro associado ao ruído, chamado *subespaço ruído*. Se o nível de ruído nos dados não for muito alto, o subespaço associado ao sinal é dominante, e o subespaço associado ao ruído é subordinado: o subespaço sinal está associado aos maiores valores singulares da matriz de observação, enquanto o subespaço ruído está associado aos menores. Uma forma de realizar essa separação é olhar para um “salto” entre os valores singulares. Porém, essa não é uma tarefa muito simples, especialmente quando o nível de ruído nos dados é alto ou quando o sinal analisado tem exponenciais muito próximas. Existem várias técnicas baseadas na estimativa do subespaço sinal, chamadas *Métodos de Subespaço*, tais como: o método de Kung [24], o método HSVD [21], o método OPIA [1], o método de Zeiger e Mac-Ewen [48], e outros [41].

Uma outra forma de abordar o problema é pela das Técnicas de Predição Linear [3, 7, 8, 19, 22, 23, 41], as quais são, de alguma forma, generalizações do método de Prony, em que os parâmetros são estimados a partir das raízes de certos polinômios.

Em ambas as classes de métodos, faz-se necessária uma estimativa do número de componentes do sinal. Como essa estimativa não é simples, assumiremos que o número de exponenciais é conhecido previamente.

O objetivo deste trabalho é apresentar ao leitor o problema de recuperação de exponenciais, as dificuldades associadas ao problema, e as diferentes maneiras de abordá-lo.

O trabalho está organizado da seguinte maneira: no primeiro capítulo, é feita uma breve revisão de alguns conceitos da Álgebra Linear, os quais serão utilizados no decorrer do trabalho. O capítulo 2 é dedicado à apresentação detalhada do problema da recuperação de exponenciais, bem como às características e propriedades dos problemas que serão estudados. No capítulo 3, são descritos alguns métodos utilizados para a determinação dos parâmetros, as Técnicas de Predição Linear e os Métodos de Subespaço; é mostrado que a estimativa dos parâmetros está associada à solução de um problema de autovalor. No capítulo 4 é feita uma análise da sensibilidade dos autovalores do sinal, em relação a perturbações nos dados; são apresentados resultados que mostram que, sob certas condições, esses autovalores são insensíveis a pequenas perturbações. No capítulo 5, são apresentados métodos numéricos para o cálculo parcial dos autovalores de uma matriz, com o objetivo de calcular o subespaço sinal de maneira rápida, introduzindo uma multiplicação rápida, que utiliza a Transformada Rápida de Fourier e que leva em conta a estrutura da

matriz. Para finalizar, no capítulo 6, são apresentados os resultados de alguns experimentos numéricos, comparando dois métodos apresentados no capítulo 3. Todos os experimentos numéricos apresentados foram feitos utilizando o software Matlab. São apresentadas, ainda, algumas conclusões finais.

Capítulo 1

Conceitos Fundamentais da Álgebra Linear

Este capítulo inicial destina-se a fazer uma breve revisão de alguns conceitos fundamentais de álgebra linear. Inicialmente, é apresentada uma descrição das principais notações que são utilizadas. A seguir, são definidos os conceitos de autovalor e autovetor de uma matriz, apresentadas algumas das suas principais propriedades e alguns resultados clássicos da teoria de perturbação de autovalores. Na seção 1.3, é apresentada uma das mais importantes decomposições de matrizes, a Decomposição em Valores Singulares (SVD), bem como alguns resultados sobre a sensibilidade dos valores e vetores singulares a perturbações na matriz. Utilizando a SVD, na seção 1.4, é definida a pseudo-inversa de uma matriz. Na seção seguinte, são apresentados e discutidos os problemas de quadrados mínimos lineares e os problemas de quadrados mínimos totais. Um outro conceito apresentado é o de ângulo entre subespaços, na seção 1.6.

1.1 Notação

Inicialmente, é descrita a notação dos conjuntos de números, vetores e matrizes. O conjunto dos números reais é denotado por \mathbb{R} , e o conjunto dos números complexos por \mathbb{C} .

Para o espaço de todos os vetores coluna, com n componentes reais, é usada a notação \mathbb{R}^n , e no caso das n componentes serem complexas, é usado \mathbb{C}^n . O conjunto de todas as matrizes $m \times n$, onde todas as componentes são reais, é escrito como $\mathbb{R}^{m \times n}$, quando todas as componentes são complexas, é escrito como $\mathbb{C}^{m \times n}$.

Em quase todo o texto, usam-se letras minúsculas para vetores, letras gregas minúsculas para escalares, e letras maiúsculas e gregas maiúsculas para matrizes.

O elemento de uma matriz A , que se encontra na linha i e na coluna

j , é denotado por a_{ij} , algumas vezes será usada a notação a_j , para designar a coluna j de A (ou a linha j de A).

A matriz nula, o vetor nulo e o escalar zero são denotados por 0 . A matriz identidade é escrita como I , ou I_n se for necessário especificar a sua ordem. O vetor formado só de uns é chamado de e , e o i -ésimo vetor canônico, isto é, o vetor cuja i -ésima componente é 1 e as demais são zeros, de e_i .

A transposta da matriz A , é denotada por A^T , sua conjugada transposta por A^* , isto é, $A^* = \bar{A}^T$. A inversa denotada por A^{-1} , a inversa da transposta por A^{-T} , e a inversa da conjugada transposta por A^{-*} .

O espaço coluna de $A \in \mathbb{C}^{m \times n}$ é chamado de

$$\mathcal{R}(A) = \{Ax : x \in \mathbb{C}^n\}$$

e o espaço nulo de

$$\mathcal{N}(A) = \{x : Ax = 0\}.$$

O posto de A é dado por, $\text{posto}(A) = \dim(\mathcal{R}(A))$, onde $\dim(\mathcal{X})$ denota a dimensão do espaço \mathcal{X} . O determinante de A é escrito como $\det(A)$, e o traço como $\text{tr}(A)$.

A norma-2 de um vetor $x \in \mathbb{C}^n$, também chamada de norma Euclidiana, é escrita como $\|x\|_2$, e definida como a raiz quadrada positiva de:

$$\sum_{i=1}^n |x_i|^2 = x^* x. \quad (1.1)$$

A norma Frobenius de uma matriz $A \in \mathbb{C}^{m \times n}$, $\|A\|_F$ é definida por:

$$\|A\|_F = \left(\sum_{j=1}^n \sum_{i=1}^m |a_{ij}|^2 \right)^{1/2}, \quad (1.2)$$

e a norma-2 de A é definida por

$$\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}. \quad (1.3)$$

Em alguns cálculos, para efeito de simplificação da notação, a norma-2 será denotada somente por $\|\cdot\|$.

Existe ainda a norma p , dada por

$$\|A\|_p = \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}, \quad (1.4)$$

onde $\|x\|_p = (|x_1|^p + \dots + |x_n|^p)^{\frac{1}{p}}$.

Seja $\{x_1, x_2, \dots, x_n\}$ um conjunto de vetores, o espaço gerado por estes vetores é denotado por $\text{span}\{x_1, x_2, \dots, x_n\}$.

A seguir, são definidos alguns tipos especiais de matrizes:

Definição 1.1. Dizemos que uma matriz P é uma matriz de projeção ortogonal se $P = P^*$ e $P^2 = P$.

Definição 1.2. - Seja $A \in \mathbb{C}^{m \times n}$

- Simétrica (hermitiana), se $A^T = A$ ($A^* = A$).
- Positiva definida (semipositiva definida), se é hermitiana e se $x^*Ax > (\geq) 0$ para todo $x \neq 0$.
- Unitária, ou ortogonal no caso real, se $A^*A = AA^* = I$.
- Normal, se $A^*A = AA^*$.
- Triangular superior, se é quadrada e se $i > j \Rightarrow a_{ij} = 0$, isto é, se os elementos abaixo da diagonal principal são iguais a zero.
- Triangular inferior, se é quadrada e se $i < j \Rightarrow a_{ij} = 0$, isto é, se os elementos acima da diagonal principal são iguais a zero.
- Diagonal, se é triangular superior e inferior, ou seja, se os elementos fora da diagonal principal são todos nulos. Se $\{\kappa_1, \kappa_2, \dots, \kappa_p\}$ números reais ou complexos, então $\text{rmdiag}(\kappa_1, \kappa_2, \dots, \kappa_p)$, denota uma matriz diagonal formada por estes números.
- Hessenberg superior, se $a_{ij} = 0$ para qualquer par i, j tal que $i > j + 1$.
- Matriz de Hankel, se as anti-diagonais são constantes, isto é, $a_{ij} = f_{i+j}$ para algum vetor f .
- Matriz de Toeplitz, se as diagonais são constantes, isto é, $a_{ij} = f_{j-i}$ para algum vetor f .

1.2 Autovalores e Autovetores

Informalmente falando, um autovetor é um vetor que não muda a sua direção, quando multiplicado por A , e seu autovalor é o tamanho da sua expansão ou contração neste processo.

Definição 1.3. Seja $A \in \mathbb{C}^{n \times n}$. O par (x, λ) é chamado um auto-par da matriz A , se $x \neq 0$ e

$$Ax = \lambda x. \tag{1.5}$$

O vetor x é chamado *autovetor* à direita de A , e λ o *autovalor* associado. Se $y^*A = \lambda y^*$, y é chamado autovetor à esquerda de A , associado ao autovalor λ . O conjunto de todos os autovalores de A é denotado por $\lambda(A)$. A equação (1.5) pode ser rescrita como:

$$(A - \lambda I)x = 0, \quad x \neq 0$$

o que ocorre se e somente se $(A - \lambda I)$ é singular, ou equivalentemente, se e somente se:

$$\det(A - \lambda I) = 0 \tag{1.6}$$

A expressão (1.6) é chamada *polinômio característico* de A , e possui grau n em λ , suas raízes são os autovalores de A . Conseqüentemente, a matriz A possui exatamente n autovalores.

Diz-se que um autovalor λ possui multiplicidade algébrica μ , se ele é uma raiz com multiplicidade μ do polinômio característico. Se a multiplicidade algébrica de um autovalor for 1, ele é chamado *simples*, caso contrário é chamado *múltiplo*.

Um autovalor λ possui multiplicidade geométrica γ , se o número máximo de autovetores associados a λ é γ , ou seja, a dimensão de $\mathcal{N}(A - \lambda I)$ é γ . Seja $A \in \mathbb{C}^{n \times n}$, a seguir, são apresentadas algumas propriedades dos autovalores e autovetores de A :

P1 . A e A^T tem os mesmos autovalores.

P2 . Se λ é autovalor de A com autovetor v e A é não singular, então λ^{-1} é autovalor de A^{-1} com autovetor v .

P3 . Se A é uma matriz triangular de blocos, digamos:

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ & A_{22} & \dots & A_{2n} \\ & & \ddots & \\ & & & A_{nn} \end{pmatrix}.$$

Então o conjunto dos autovalores de A , contando a multiplicidade, é igual a união dos conjuntos de autovalores de A_{11}, \dots, A_{nn} .

P4 . A soma dos autovalores é igual ao traço da matriz.

P5 . O produto dos autovalores é igual ao determinante da matriz.

P6 . Sejam v_1, \dots, v_k autovetores de A associados a distintos autovalores $\lambda_1, \dots, \lambda_k$. Então v_1, \dots, v_k são *LI*.

- P7 . Se A tem n autovalores distintos, então A pode ser diagonalizada, ou seja, existe uma matriz X não singular e uma matriz D diagonal tal que $A = X^{-1}DX$.
- P8 . Os autovalores de A^k são $\lambda_1^k, \dots, \lambda_n^k$, a k -ésima potência dos autovalores de A .
- P9 . Seja $\rho \in \mathbb{C}$, se v é um autovetor de A com autovalor λ , então v é também um autovetor de $A - \rho I$ com autovalor $\lambda - \rho$.
- P10 . Os autovalores de uma matriz simétrica ou de uma matriz hermitiana são reais.
- P11 . Os autovetores de uma matriz simétrica real ou de uma matriz hermitiana, correspondentes a autovalores distintos são ortogonais entre si.
- P12 . Uma matriz simétrica real pode ser fatorada em $A = Q\Lambda Q^T$, com autovalores em Λ e respectivos autovetores ortogonais em Q .
- P13 . Se A é real, então seu polinômio característico é real, e os seus autovalores complexos devem ocorrer em pares conjugados complexos.
- P14 . Os autovalores de uma matriz triangular são os elementos da sua diagonal.

1.2.1 Condicionamento dos Autovalores

Nesta subseção são apresentados alguns resultados clássicos da teoria de perturbação de autovalores, os quais fornecem medidas da sensibilidade de qualquer problema de autovalor.

Definição 1.4. *Seja $A \in \mathbb{C}^{n \times n}$ e seja $A = X\Lambda X^{-1}$ uma decomposição autovalor-autovetor de A . Os números de condição de Jordan do problema de autovalor associado a A , são definidos por*

$$\kappa_2 = \min_{D \in \mathcal{D}} \{\|XD\|_2 \|D^{-1}X^{-1}\|_2\}, \quad \kappa_F = \min_{D \in \mathcal{D}} \{\|XD\|_F \|D^{-1}X^{-1}\|_F\}, \quad (1.7)$$

onde \mathcal{D} é o conjunto de todas as matrizes diagonais inversíveis.

Sejam u_j e v_j autovetores à esquerda e à direita de A associados ao autovalor λ_j , define-se

$$s_j = \frac{u_j^* v_j}{\|u_j\|_2 \cdot \|v_j\|_2}. \quad (1.8)$$

Se λ_j é um autovalor simples, então x e y são únicos e $|s_j|$ é unicamente determinado. O número $|s_j|^{-1}$ fornece uma medida da sensibilidade do autovalor λ_j , a pequenas

perturbações nas entradas da matriz A , [45], e é chamado número de condição de λ_j .

O Teorema 1.5 fornece uma outra forma de calcular κ_F , quando os autovalores de A são simples.

Teorema 1.5. *Se todos os autovalores de A são simples, então*

$$\kappa_F = |s_1|^{-1} + |s_2|^{-1} + \cdots + |s_n|^{-1}. \quad (1.9)$$

Demonstração:

A demonstração é encontrada em [33].

O próximo teorema fornece algumas relações envolvendo o número de condição $|s_j|^{-1}$.

Teorema 1.6. *Suponha que todos os autovalores de A são simples. Então*

$$1 \leq |s_j|^{-1} \leq \left[1 + \frac{D^2(A)}{(n-1)\delta_j^2} \right]^{\frac{n-1}{2}}, \quad 1 \leq j \leq n, \quad (1.10)$$

onde $\delta_j = \min_{\substack{1 \leq k \leq n \\ k \neq j}} |\lambda_j - \lambda_k|$ e $D(A)^2 = \|A\|_F^2 - \sum_{j=1}^n |\lambda_j|^2$. Além disso o número de condição $\kappa_F(X) = \|X\|_F \|X^{-1}\|_F$, satisfaz

$$\sum_{j=1}^n \frac{1}{|s_j|} \leq \kappa_F(X) \leq n \cdot \max_j \frac{1}{|s_j|} \cdot \frac{\max_j \|e_j^* X^{-1}\|_2}{\min_j \|e_j^* X^{-1}\|_2}. \quad (1.11)$$

Demonstração:

A desigualdade (1.10) e a desigualdade à esquerda de (1.11) seguem do Teorema 5 e do Teorema 3, de Smith [33], respectivamente. Para provar a desigualdade à direita de (1.11), defina os vetores $u_j^* = e_j^* X^{-1}$ e $v_j = X e_j$, os quais são autovetores à esquerda e à direita de A , associados ao autovalor λ_j . Usando a definição de $\kappa_F(X)$ e (1.8), segue que

$$\kappa_F^2(X) = \sum_{j=1}^n \|X e_j\|_2^2 \sum_{j=1}^n \|e_j^* X^{-1}\|_2^2 = \sum_{j=1}^n \frac{1}{|s_j|^2} \cdot \frac{1}{\|e_j^* X^{-1}\|_2^2} \sum_{j=1}^n \|e_j^* X^{-1}\|_2^2.$$

Basta agora, limitar os termos na soma. □

O número $D(A)^2$ conhecido como *medida de não normalidade de A* (*departure from normality*), fornece uma medida de quão perto uma matriz está de ser uma matriz normal. Se A for uma matriz normal, $D(A)^2 = 0$. Além disto, os autovalores de uma matriz normal são perfeitamente bem condicionados.

É apresentado a seguir, um teorema bastante conhecido, o Teorema de Bauer-Fike, cuja demonstração pode ser encontrada em [15, pág 321]. O teorema

fornece uma estimativa do erro entre os autovalores da matriz A , e os autovalores da matriz perturbada $A + E$, ou seja, apresenta uma estimativa do comportamento dos autovalores em função da perturbação nas entradas da matriz.

Teorema 1.7. *Se $\tilde{\lambda}$ é um autovalor de $A + E \in \mathbb{C}^{n \times n}$ e $X^{-1}AX = \Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$, então*

$$\min_{\lambda \in \lambda(A)} |\tilde{\lambda} - \lambda| \leq \kappa_p(X) \|E\|_p,$$

onde $\|\cdot\|_p$, denota a norma p definida em (1.4).

1.3 Decomposição em Valores Singulares (SVD)

Nesta seção é apresentada uma ferramenta da álgebra linear computacional muito poderosa, que facilita o trabalho com problemas de posto incompleto, isto é, problemas que envolvem matrizes $A \in \mathbb{C}^{m \times n}$, cujo $\text{posto}(A) < \min\{m, n\}$. A SVD é de fundamental importância no desenvolvimento deste trabalho.

Teorema 1.8. (Decomposição em Valores Singulares (SVD)) *Seja $A \in \mathbb{C}^{m \times n}$, uma matriz com posto d . Então existem matrizes unitárias $U \in \mathbb{C}^{m \times m}$ e $V \in \mathbb{C}^{n \times n}$, e uma matriz diagonal $\Sigma \in \mathbb{C}^{m \times n}$, tais que:*

$$A = U\Sigma V^* \tag{1.12}$$

onde a matriz Σ é da forma:

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \dots & & \\ & & & \dots & \\ & & & & \sigma_p & \\ & & & & & \dots \\ & & & & & & \dots \\ & & & & & & & \dots \\ & & & & & & & & \dots \\ & & & & & & & & & \dots \\ & & & & & & & & & & \dots \end{bmatrix}$$

com $p = \min\{m, n\}$ e $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_d > \sigma_{d+1} = \dots = \sigma_n = 0$.

Demonstração:

A demonstração deste teorema pode ser encontrada em [15].

Os números σ_i são chamados valores singulares de A , enquanto os vetores u_i e v_i são chamados vetores singulares à esquerda e à direita de A , respectivamente. Será usado $\sigma_i(A)$, ou simplesmente σ_i , para denotar o i -ésimo valor singular da matriz A .

Geometricamente falando, a SVD de A mostra que existem duas bases ortogonais, a saber, as colunas de U e V , tais que a aplicação A , de \mathbb{R}^n em \mathbb{R}^m , é representada por uma matriz diagonal Σ , cujas entradas são números reais.

A partir das relações $A^*A = V\Sigma^2V^*$ e $AA^* = U\Sigma^2U$, verifica-se que a SVD está fortemente relacionada com as decomposições autovalor-autovetor das matrizes hermitianas A^*A e AA^* .

A equação (1.12) pode ser rescrita como:

$$A = [U_1 \ U_2] \Sigma \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix} \quad (1.13)$$

onde U_1 e V_1 são as d primeiras colunas das matrizes U e V respectivamente, U_2 e V_2 são as $m - d$ e $n - d$ últimas colunas das matrizes U e V respectivamente.

Em termos destas matrizes, a SVD de A pode também ser escrita da seguinte maneira:

$$\begin{aligned} A &= U_1 \Sigma_1 V_1^* \\ &= \sum_{i=1}^d u_i \sigma_i v_i^* \end{aligned} \quad (1.14)$$

onde $\Sigma_1 = \text{diag}(\sigma_1, \dots, \sigma_d)$. A decomposição (1.14) é chamada *SVD truncada* ou *SVD reduzida*. Assim, a matriz A pode ser escrita como um produto de matrizes de posto d , ou como a soma de d matrizes de posto 1.

Tem-se ainda que:

$$\begin{aligned} \mathcal{N}(A) &= \text{span}\{v_{d+1}, \dots, v_n\} \\ \mathcal{R}(A) &= \text{span}\{u_1, \dots, u_d\} \\ \mathcal{N}(A^*) &= \text{span}\{u_{d+1}, \dots, u_m\} \\ \mathcal{R}(A^*) &= \text{span}\{v_1, \dots, v_d\}. \end{aligned}$$

A SVD também fornece informação sobre a norma-2 e a norma Frobenius de A

$$\begin{aligned} \|A\|_F^2 &= \sigma_1^2 + \dots + \sigma_p^2, \quad p = \min\{m, n\} \\ \|A\|_2 &= \sigma_1 \\ \min_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} &= \sigma_n \quad (m \geq n) \end{aligned}$$

Abaixo, são listadas as principais projeções ortogonais associadas com a SVD. Suponha que A satisfaz as condições do teorema 1.8, e as matrizes U e V são particionadas como em (1.13), então:

$$V_1 V_1^* = \text{projeção sobre } \mathcal{N}(A)^\perp = \mathcal{R}(A^*)$$

$$V_2 V_2^* = \text{projeção sobre } \mathcal{N}(A)$$

$$U_1 U_1^* = \text{projeção sobre } \mathcal{R}(A)$$

$$U_2 U_2^* = \text{projeção sobre } \mathcal{R}(A)^\perp = \mathcal{N}(A^*)$$

1.3.1 Sensibilidade dos Valores Singulares

É apresentado a seguir um resultado clássico da teoria de perturbação de valores singulares, o qual afirma que os valores singulares são bem condicionados em relação a perturbações:

Teorema 1.9. *Seja A e $\tilde{A} = A + E \in \mathbb{C}^{m \times n}$, $m \geq n$ com valores singulares $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$ e $\tilde{\sigma}_1 \geq \tilde{\sigma}_2 \geq \dots \geq \tilde{\sigma}_n$, respectivamente. Então:*

$$\begin{aligned} |\sigma_i - \tilde{\sigma}_i| &\leq \|E\|_2 \\ \sum_{i=1}^n |\sigma_i - \tilde{\sigma}_i| &\leq \|E\|_F^2 \end{aligned}$$

Deste teorema decorre que os $n - d$ valores singulares da matriz \tilde{A} são da ordem de $\|E\|_2$, uma vez que $\sigma_{d+1} = \dots = \sigma_n = 0$. E desta forma a matriz \tilde{A} se comporta como sendo de posto completo.

Quanto aos efeitos da perturbação nos vetores singulares, eles também são estimados na ordem de $\|E\|_2$, veja [34].

1.4 Pseudo Inversa

Dada uma matriz $A \in \mathbb{C}^{m \times n}$, a pseudo-inversa de A , também conhecida como inversa de Moore-Penrose, é definida por meio das seguintes condições:

$$\begin{aligned} (1) \quad AA^\dagger A &= A & (2) \quad A^\dagger AA^\dagger &= A^\dagger \\ (3) \quad (AA^\dagger)^* &= AA^\dagger & (4) \quad (A^\dagger A)^* &= A^\dagger A \end{aligned}$$

Como $AA^\dagger = U_1 U_1^*$, tem-se que AA^\dagger é a projeção ortogonal sobre $\mathcal{R}(A)$, e $A^\dagger A = V_1 V_1^*$ é a projeção ortogonal sobre $\mathcal{R}(A^*)$.

Em particular, a pseudo-inversa pode ser definida pela SVD truncada de A :

$$A^\dagger = V_1 \Sigma_1^{-1} U_1^* \tag{1.15}$$

onde U_1 , V_1 , e Σ_1 são dadas como em (1.13). Observe que se $m = n$, e se $\text{posto}(A) = m$, então $A^\dagger = A^{-1}$.

Define-se também o *número de condicionamento* de A , por

$$\kappa_p = \|A\|_p \|A^\dagger\|_p, \tag{1.16}$$

o qual fornece uma medida da sensibilidade do problema $Ax = b$, a perturbações em A , onde $x \in \mathbb{C}^n$, e $b \in \mathbb{C}^m$.

1.5 Problemas de Quadrados Mínimos Lineares

O Problema de Quadrados Mínimos Lineares (PQML) pode ser formulado da seguinte maneira: Dados um vetor $b \in \mathbb{C}^m$ e uma matriz $A \in \mathbb{C}^{m \times n}$, $m > n$, encontrar um vetor $x \in \mathbb{C}^n$, tal que Ax é a “melhor aproximação” para b , isto é, resolver:

$$\min_{x \in \mathbb{C}^n} \|Ax - b\|_2 \quad (1.17)$$

O vetor x é chamado solução de quadrados mínimos, e o vetor $r = b - Ax$, é chamado vetor de resíduos.

O próximo resultado caracteriza o conjunto de todas as soluções do PQML (1.17).

Teorema 1.10. *Seja $\mathcal{S} = \{x \in \mathbb{C}^n / \|Ax - b\|_2 = \min\}$ o conjunto de todas as soluções de (1.17). Então, $x \in \mathcal{S}$ se e somente se ocorre a seguinte condição de ortogonalidade*

$$A^*(b - Ax) = 0 \quad (1.18)$$

Demonstração:

A demonstração deste teorema pode ser encontrada em [9, pág. 5].

O Teorema acima afirma que o vetor de resíduos de uma solução de quadrados mínimos, $r = b - Ax$, está em $\mathcal{N}(A^*)$. conseqüentemente, qualquer solução do PQML decompõe, de maneira única o vetor b em duas componentes ortogonais

$$b = Ax + r, \quad Ax \in \mathcal{R}(A), \quad r \in \mathcal{N}(A^*).$$

Esta propriedade é ilustrada geometricamente na figura 1.1.

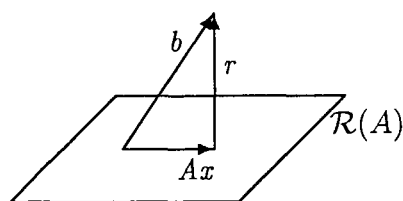


Figura 1.1: Interpretação Geométrica do PQML

Da equação (1.18) segue que a solução de quadrados mínimos satisfaz às equações normais

$$A^*Ax = A^*b. \quad (1.19)$$

A matriz $A^*A \in \mathbb{C}^{n \times n}$ é hermitiana e semipositiva definida, sendo positiva definida se e somente se $\text{posto}(A) = n$. Assim, se $\text{posto}(A) = n$, então o problema de quadrados mínimos lineares possui solução única, assim x e o correspondente vetor de resíduos r , são dados por

$$x = (A^*A)^{-1}A^*b, \quad r = b - A(A^*A)^{-1}A^*b.$$

Se $\text{posto}(A) < n$, então A possui um espaço nulo não trivial, e a solução de quadrados mínimos não é única, porém, entre todas as soluções existe uma única que minimiza $\|x\|_2$, veja o próximo teorema.

A SVD é uma poderosa ferramenta para resolver PQML, pois as matrizes ortogonais que transformam A em uma matriz diagonal (1.12) não alteram a norma-2 dos vetores. É apresentado a seguir um resultado que se aplica tanto a sistemas sobre-determinados (mais equações do que variáveis) como a sub-determinados (mais variáveis do que equações).

Teorema 1.11. *Considere o problema de quadrados mínimos lineares geral*

$$\min_{x \in \mathcal{S}} \|x\|_2, \quad \mathcal{S} = \{x \in \mathbb{C}^n / \|b - Ax\|_2 = \min\},$$

onde $A \in \mathbb{C}^{m \times n}$ e $\text{posto}(A) = d \leq \min\{m, n\}$. Este problema sempre possui solução única, a qual pode ser escrita em termos da SVD truncada de A como

$$x^\dagger = V_1 \begin{pmatrix} \Sigma_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} U_1^* b, \quad (1.20)$$

onde V_1 , U_1 e Σ_1^{-1} , são dados como em (1.13).

Demonstração:

A demonstração pode ser encontrada em [9].

Escreve-se (1.20) como $x^\dagger = A^\dagger b$, e a solução x^\dagger é chamada pseudo-solução.

1.5.1 Problemas de Quadrados Mínimos Totais

Em muitas aplicações existem problemas onde a matriz A não é precisamente conhecida. Por exemplo, A pode ser disponível apenas por medidas, ou pode ser uma aproximação de um operador. Assim, é necessário desenvolver métodos que levem em conta os erros tanto em A como em b . Um destes métodos é o de Quadrados Mínimos Totais (TLS).

Definição 1.12. *Dado um sistema de equações, onde ocorrem erros na matriz A e no vetor b ,*

$$(A + E)x = b + r. \quad (1.21)$$

As estimativas para x no sistema (1.21), são encontradas através da solução do problema de quadrados mínimos totais,

$$\min_{E,r} \|(E, r)\|_F \quad (1.22)$$

sujeito a

$$(A + E)x = b + r.$$

Observe que a restrição em (1.22), implica em $b + r \in \mathcal{R}(A + E)$. Se é encontrado um minimizador (E, r) para o problema (1.22), então qualquer x que satisfaz $(A + E)x = b + r$, resolve o problema de quadrados mínimos totais. Observe também, que esta restrição pode ser escrita como

$$(\tilde{A}, \tilde{b}) \begin{pmatrix} x \\ -1 \end{pmatrix} = 0, \quad \tilde{A} = A + E, \quad \tilde{b} = b + r.$$

O que mostra que a solução de quadrados mínimos totais faz com que a matriz $(A + E, b + r)$ seja de posto incompleto, e que $(x, -1)^*$ seja um vetor singular à direita, associado a um valor singular nulo. A solução de um problema TLS pode ser obtida através da SVD, de fato: seja

$$(A, b) = \hat{U}\hat{\Sigma}\hat{V}^*, \quad \Sigma = \text{diag}(\hat{\sigma}_1, \dots, \hat{\sigma}_{n+1}),$$

onde $\hat{U}^*\hat{U} = I_m$, $\hat{V}^*\hat{V} = I_{n+1}$ e $\hat{\sigma}_1 \geq \dots \geq \hat{\sigma}_{n+1} \geq 0$. Pelo teorema 1.2.9, [9], os valores singulares de A ,

$$\sigma_1 \geq \dots \geq \sigma_n \geq 0$$

entrelaçam os valores singulares de (A, b) , i.e.

$$\hat{\sigma}_1 \geq \sigma_1 \geq \hat{\sigma}_2 \geq \dots \geq \hat{\sigma}_n \geq \sigma_n \geq \hat{\sigma}_{n+1}.$$

Assuma que $\text{posto}(A) = n$, ou seja que $\sigma_n > 0$. Se $\hat{\sigma}_{n+1} = 0$, então segue que $b \in \mathcal{R}(A)$. Neste caso, o sistema original, $Ax = b$, é compatível, e pode-se tomar $(E, r) = 0$. Se $\hat{\sigma}_{n+1} > 0$, então $b \notin \mathcal{R}(A)$ e do teorema de Eckart-Young-Mirsky, [20, pag], segue que

$$\min_{\text{posto}(A+E, b+r) < n+1} \|(E, r)\|_F = \hat{\sigma}_{n+1}.$$

Se ocorrer

$$\hat{\sigma}_k > \hat{\sigma}_{k+1} = \dots = \hat{\sigma}_{n+1}, \quad k \leq n,$$

então o mínimo é atingido para qualquer perturbação de posto 1 da forma

$$(E, r) = -(A, b)\hat{v}\hat{v}^*, \quad \hat{v} \in \mathcal{S} = \text{span}[\hat{v}_{k+1}, \dots, \hat{v}_{n+1}],$$

onde $\hat{v}_{k+1}, \dots, \hat{v}_{n+1}$ são os vetores singulares à direita, correspondentes a $\hat{\sigma}_{k+1}, \dots, \hat{\sigma}_{n+1}$. Se $e_{n+1} = (0, \dots, 0, 1)^*$ é ortogonal a \mathcal{S} , então o problema TLS não possui solução.

Assuma que é possível encontrar um vetor $\hat{v} \in \mathcal{S}$, cuja $(n + 1)$ -ésima componente, γ é não nula. Então, com

$$\hat{v} = \begin{pmatrix} z \\ \gamma \end{pmatrix} = -\gamma \begin{pmatrix} x \\ -1 \end{pmatrix}, \quad x = -\gamma^{-1}z \quad (1.23)$$

tem-se

$$-\gamma(A + E, b + r) \begin{pmatrix} x \\ -1 \end{pmatrix} = (A + E, b + r)\hat{v} = (A, b)(I - \hat{v}\hat{v}^*)\hat{v} = 0.$$

Portanto $(A + E)x = b + r$, o que mostra que x é solução do problema de quadrados mínimos totais. Usando (1.23), a perturbação mínima pode ser escrita como

$$(E, r) = -\gamma^2(A, b) \begin{pmatrix} x \\ -1 \end{pmatrix} (x^*, -1) = \gamma^2(\hat{r}x^*, -\hat{r})$$

onde

$$\hat{r} = b - Ax, \quad \gamma^2 = (1 + \|x\|_2^2)^{-1}.$$

O próximo teorema fornece uma condição suficiente para que o problema (1.22) possua solução única.

Teorema 1.13. *Sejam $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$ os valores singulares de A . Se $\sigma_n > \hat{\sigma}_{n+1}$, então o problema de quadrados mínimos totais possui solução única.*

Demonstração: A demonstração pode ser encontrada em [9] ou [20].

1.6 Ângulo entre Subespaços

Sejam \mathcal{S} e \mathcal{X} dois subespaços em \mathbb{C}^m cujas dimensões satisfazem

$$p = \dim(\mathcal{S}) \geq \dim(\mathcal{X}) = q \geq 1.$$

Os ângulos principais entre \mathcal{S} e \mathcal{X} , $\theta_1, \dots, \theta_q \in [0, \pi/2]$ são definidos recursivamente por

$$\cos(\theta_k) = \max_{u \in \mathcal{S}} \max_{v \in \mathcal{X}} u^*v = u_k^*v_k$$

sujeito a:

$$\begin{aligned} \|u\| &= \|v\| = 1 \\ u^*u_i &= 0 \quad i = 1, \dots, k-1 \\ v^*v_i &= 0 \quad i = 1, \dots, k-1 \end{aligned}$$

Observe que os ângulos principais satisfazem $0 \leq \theta_1 \leq \dots \leq \theta_q \leq \pi/2$. Os vetores $\{u_1, \dots, u_q\}$ e $\{v_1, \dots, v_q\}$ são chamados *vetores principais* entre os subespaços \mathcal{S} e \mathcal{X} .

O maior ângulo principal está relacionado com a noção de distância entre subespaços que possuem a mesma dimensão. Se $p = q$, então

$$\text{dist}(\mathcal{S}, \mathcal{X}) = \sqrt{1 - \cos(\theta_p)^2} = \text{sen}(\theta_p).$$

A relação entre os ângulos principais e a SVD é dada no teorema abaixo:

Teorema 1.14. *Suponha que $Q_S \in \mathbb{C}^{m \times p}$ e $Q_X \in \mathbb{C}^{m \times q}$ são bases ortogonais dos subespaços \mathcal{S} e \mathcal{X} . Considere a SVD*

$$M = Q_S^* Q_X = Y C Z^*, \quad C = \text{diag}(\sigma_1, \dots, \sigma_p)$$

onde $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p$, $Y^* Y = Z^* Z = I_q$. Então os ângulos e vetores principais são dados por

$$\cos \theta_k = \sigma_k, \quad U = Q_S Y, \quad V = Q_X Z.$$

Demonstração:

A demonstração pode ser encontrada em [9].

Capítulo 2

O Problema da Recuperação de Exponenciais: Apresentação e Propriedades

Neste capítulo é apresentado o problema de recuperação de exponenciais, o qual é o problema central estudado neste trabalho. É apresentado um dos métodos mais antigos usado para resolvê-lo, o método de Prony. Sua apresentação é puramente descritiva e serve de motivação para o estudo de outros métodos, uma vez que o método de Prony possui algumas deficiências, sendo a principal delas a sensibilidade a perturbações nos dados. Finalizando o capítulo, são descritas algumas propriedades intrínsecas do problema.

2.1 Recuperação de exponenciais e o Método de Prony

A recuperação de exponenciais é um problema muito antigo, e foi abordado inicialmente pelo francês R. de Prony [31], em 1795. Hoje, é muito freqüente em diversas áreas como, comunicações, localização de objetos por radar, identificação paramétrica de sistemas dinâmicos, análise modal de estruturas mecânicas, ressonância magnética nuclear, análise de compartimentos, diagnósticos médicos, etc, [21, 1, 7, 11, 17].

Seja h_j , $j = 1, 2, \dots$, um conjunto de amostras de um sinal discreto no tempo, modelado como a soma de d exponenciais complexas:

$$h_j = \sum_{l=1}^d r_l z_l^{j-1} = \sum_{l=1}^d r_l e^{(\alpha_l + i\omega_l)\Delta t j} = \sum_{l=1}^d r_l e^{s_l \Delta t j}, \quad \text{para } j = 1, 2, \dots \quad (2.1)$$

onde os coeficientes r_l são números complexos, $i^2 = -1$, $\alpha_l \leq 0$, $\omega_l \in \mathbb{R}$, e Δt a taxa

de amostragem. Sinais deste tipo são geralmente associados à resposta impulsiva de sistemas invariantes no tempo, e o significado dos parâmetros envolvidos depende do sistema em estudo ou da origem do sinal. Para citar um exemplo, se h_j é a resposta livre ou a resposta ao impulso unitário de um sistema mecânico, os α_l fornecem informação sobre o amortecimento do sistema, os ω_l descrevem as frequências naturais de vibração, e os r_l medem o grau de participação de cada modo de vibrar do sistema.

O problema consiste em: dada uma seqüência finita de amostras h_1, h_2, \dots, h_L , onde L é o numero de amostras, determinar o menor inteiro d , os coeficientes complexos r_l e distintos z_l , que satisfaçam 2.1.

Observe que o modelo (2.1) pode ser formulado usando uma matriz de Vandermonde, da seguinte maneira:

$$V_L r = \mathbf{h}_L, \quad (2.2)$$

onde

$$V_L = \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & \dots & z_d \\ z_1^2 & z_2^2 & \dots & z_d^2 \\ \vdots & \vdots & \dots & \vdots \\ z_1^{L-1} & z_2^{L-1} & \dots & z_d^{L-1} \end{bmatrix},$$

$$r = [r_1, r_2, \dots, r_d]^T \text{ e } \mathbf{h}_L = [h_1, \dots, h_L]^T.$$

A idéia chave de R. de Prony foi usar o fato que o sinal h_k satisfaz uma equação de diferenças. De fato, seja $P(z)$ um polinômio de grau d , cujos zeros são os z_l , $l = 1, \dots, d$

$$P(z) = \prod_{l=1}^d (z - z_l) = \sum_{i=0}^d \gamma_i z^i \quad (2.3)$$

onde $\gamma_d = 1$. Então fazendo o somatório $\sum_{j=0}^d \gamma_j h_{j+i}$, para $1 \leq i \leq L - d$, e usando a equação (2.1), tem-se:

$$\begin{aligned} \sum_{j=0}^d \gamma_j h_{j+i} &= \sum_{j=0}^d \gamma_j \sum_{l=1}^d r_l z_l^{j+i} = \sum_{j=0}^d \gamma_j \sum_{l=1}^d r_l z_l^j z_l^i \\ &= \sum_{l=1}^d r_l z_l^i \sum_{j=0}^d \gamma_j z_l^j = \sum_{l=1}^d r_l z_l^i P(z_l) = 0, \end{aligned}$$

pois os z_l são raízes de $P(z)$. Assim, o sinal h_j satisfaz à equação de diferenças

$$\sum_{j=0}^d \gamma_j h_{j+i} = 0.$$

Reescrevendo esta equação obtém-se:

$$h_{d+i} = - \sum_{j=0}^{d-1} \gamma_j h_{j+i},$$

que pode ser expresso pelo sistema:

$$\begin{bmatrix} h_1 & h_2 & \cdots & h_d \\ h_2 & h_3 & \cdots & h_{d+1} \\ \vdots & \vdots & \cdots & \vdots \\ h_{L-d} & h_{L-d+1} & \cdots & h_{L-1} \end{bmatrix} \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_{d-1} \end{bmatrix} = - \begin{bmatrix} h_{d+1} \\ h_{d+2} \\ \vdots \\ h_L \end{bmatrix}, \quad (2.4)$$

o qual é chamado sistema de Yule-Walker. Curiosamente, Prony apresentou esta equação matricial mais de um século antes de Yule (1927) [47] e Walker (1931) [44]. Ele assumiu que d era par e tomou $L = 2d$, fazendo com que a matriz de (2.4) fosse quadrada. Quando o número de amostras era ímpar, ele forçava um dos z 's a ser igual a um, e eliminava o coeficiente associado a este z de (2.4). O método de Prony pode ser resumido da seguinte maneira:

1. Resolva o sistema quadrado, (2.4)
2. Determine as raízes do polinômio $P(z)$ (2.3)
3. Encontre os r_j resolvendo o sistema (2.2)

Como estes sinais são provenientes de problemas práticos, eles possuem uma certa quantidade de informação não desejada, chamada de *ruído*, a qual dificulta a estimativa dos parâmetros. Assim, o objetivo de qualquer método de recuperação de exponenciais é, a partir do sinal contaminado por ruídos, encontrar aproximações para os parâmetros do sinal livre de ruídos. Quando é considerada a presença de ruído ou de outro tipo de erro nos dados, os passos 1 e 3 acima podem ser resolvidos no sentido de quadrados mínimos. Porém, para isto é necessário um conhecimento prévio do número de exponenciais.

A qualidade dos parâmetros estimados pelo método de Prony depende do grau de condicionamento da matriz do sistema, o qual pode ser calculado através de (1.16), e do nível de ruído nos dados, ou seja, do grau de perturbação nas amostras no sinal. Quando o sistema (2.4) é resolvido por meio de pseudo-inversão, usando o teorema de decomposição em valores singulares (*SVD*), são obtidos melhores resultados [23].

Para terminar esta apresentação, é importante observar que os parâmetros embutidos no sinal são muito sensíveis a pequenas perturbações nos dados. Intuitivamente, isto decorre do fato de sinais compostos por exponenciais contendo parâmetros muito diferentes, fornecerem curvas, que à primeira vista são muito

próximas umas das outras. Um exemplo que ilustra esta situação é mostrado na Figura (2.1), onde são apresentados os gráficos de dois sinais, $s_1(t)$ (com 3 exponenciais) e $s_2(t)$ (com 2 exponenciais) definidos por

$$\begin{aligned} s_1(t) &= 0.305e^{-t/0.633} + 2.202e^{-t/0.225} \\ s_2(t) &= 0.0951e^{-t} + 0.8607e^{-t/0.333} + 1.557e^{-t/2} \end{aligned} \quad (2.5)$$

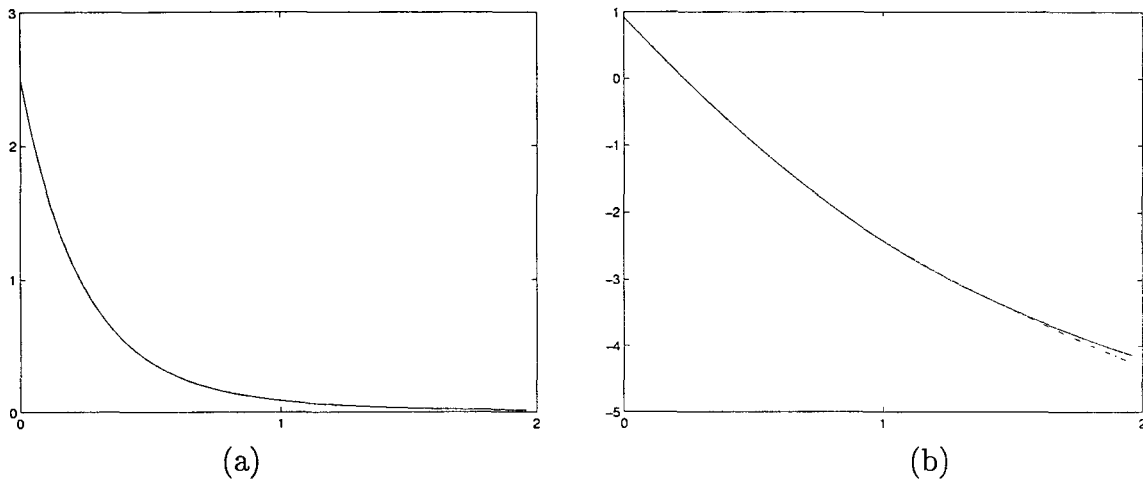


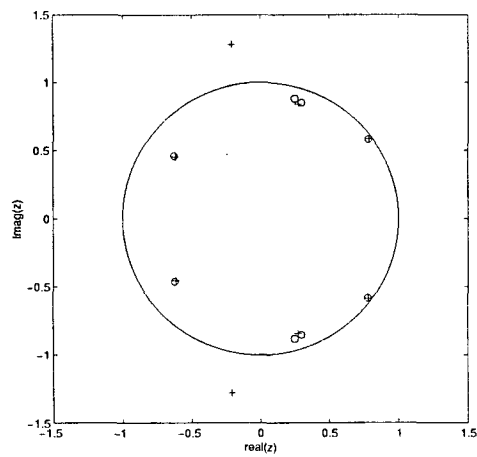
Figura 2.1: (a): Sinais $s_1(t)$ e $s_2(t)$. (b): Mesmos sinais de (a) em escala semilogarítmica

Este exemplo motiva a seguinte questão: É possível construir um sinal $s_3(t)$, com 4 ou mais componentes, “próximo” de $s_1(t)$? A resposta é sim, e o processo é simples. Basta perturbar levemente $s_1(t)$, e ajustar os dados usando um número arbitrário de exponenciais (maior ou igual a quatro). Este fato é ilustrado com outro exemplo, onde é considerado um sinal real, com 8 componentes ($d = 8$), cujos parâmetros são dados na Tabela 2.1.

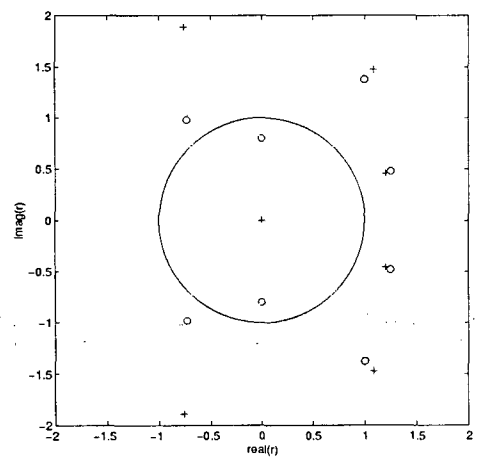
real(z_l)	imag(z_l)	real(r_l)	imag(r_l)
0.2500	± 0.8800	0	± 0.8000
0.3000	± 0.8500	-0.7300	± 0.9800
0.7800	± 0.5800	1.2500	± 0.4800
-0.6250	± 0.4600	1.0000	± 1.3750

Tabela 2.1: Parâmetros de um sinal real com 8 componentes.

A sensibilidade do problema é ilustrada na Figura 2.2, onde são apresentados os resultados do uso do método de Prony, em um conjunto de dados obtidos da soma de uma perturbação da ordem de 3%, com o sinal exato (i.e., $\|\epsilon\|/\|h\| \approx 0.03$, onde ϵ e h denotam vetores contendo as amostras do ruído e do

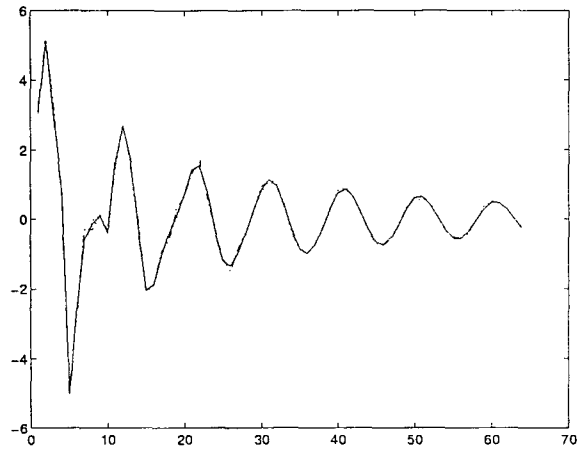


(a)

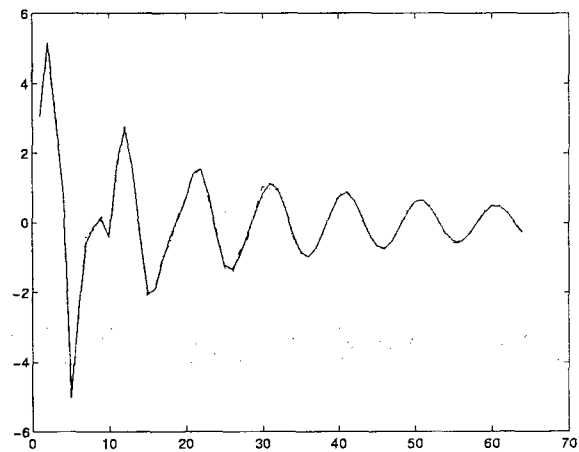


(b)

Figura 2.2: Estimativa dos parâmetros r_l e z_l para o sinal perturbado, pelo do método de Prony ('o': valores exatos, '+' :valores aproximados). Em (a) z_l , em (b) r_l .



(a)



(b)

Figura 2.3: Linha contínua: sinal exato, linha pontilhada: sinal perturbado, linha tracejada : sinal reconstruído usando as 64 primeiras amostras do sinal com 14 e 20 componentes ((a)-(b)).

sinal, respectivamente; o ruído é Gaussiano com média zero. Ao todo foram consideradas 64 amostras, porém para aplicar o método de Prony, foram utilizadas somente as 16 primeiras. Observe neste gráfico, como os parâmetros z 's mais próximos uns dos outros são mais sensíveis a perturbações.

Finalmente, é ilustrado como o conjunto de “dados” disponíveis pode ser expresso como uma soma arbitrária de exponenciais. Isto é apresentado na Figura 2.3, onde é ilustrado como é possível ter uma “boa” reconstrução do sinal, com um número arbitrário de termos.

Para contornar estes problemas, surgiram algumas modificações do método de Prony. Uma das mais conhecidas é o método de Kumaresan e Tufts [23], conhecido como Método KT. O método está baseado na determinação dos zeros do polinômio

$$\mathcal{P}(z) = c_N + c_{N-1}z + \dots + c_1z^{N-1} - z^N, \quad N > d$$

e na decomposição em valores singulares truncada (1.13). Desta maneira são encontrados $N - d$ zeros que não apresentam nenhum significado, os quais são chamados *espúrios*. Torna-se necessário então, um critério para separar os zeros desejados, dos espúrios. Kumaresan propôs um critério para fazer esta separação, o qual envolve a solução de um problema de quadrados mínimos da ordem $\binom{L}{d}$, o que não é interessante para problemas onde se necessita de respostas imediatas.

Em [22], é apresentada uma forma de estimar o número de exponenciais, baseada na seguinte idéia: Escolha $\hat{d} = 1$, aplique o método de Prony, e calcule o erro entre o sinal e o sinal reconstruído, chame este erro de E_1 . Escolha agora $\hat{d} = 2$, aplique o método de Prony e calcule E_2 . Repita este processo até que a razão entre o decréscimo do erro e o crescimento dos valores de \hat{d} seja pequena. O inteiro i para o qual E_i mostra a mais significativa mudança na razão de decréscimo é escolhido como o número de exponenciais.

2.2 Algumas Propriedades

Como foi comentado na introdução do trabalho, as abordagens existentes para o problema utilizam uma matriz que contém as amostras do sinal (matriz de observação). A seguir são descritas algumas propriedades do problema que provém dessa matriz.

Considere a matriz de observação

$$H(l) = \begin{bmatrix} h_l & h_{l+1} & \dots & h_{l+N-1} \\ h_{l+1} & h_{l+2} & \dots & h_{l+N} \\ \vdots & \vdots & \dots & \vdots \\ h_{l+M-1} & h_{l+M} & \dots & h_{l+M+N-2} \end{bmatrix}_{M \times N}, \quad (2.6)$$

onde $h_l = h(l\Delta t)$, $l \geq 0$ e Δt é intervalo de amostragem. Quando $l = 0$, será usada a notação H .

Como os sinais são modelados por funções do tipo

$$\begin{aligned} h(t) &= r_1 e^{s_1 t} + r_2 e^{s_2 t} + \dots + r_d e^{s_d t}, \quad t \geq 0 \\ &= r_1 e^{(\alpha_1 + i\omega_1)t} + r_2 e^{(\alpha_2 + i\omega_2)t} + \dots + r_d e^{(\alpha_d + i\omega_d)t}, \end{aligned} \quad (2.7)$$

a matriz de observação pode ser fatorada como

$$H(l) = V_M Z^l R W \quad (2.8)$$

onde $Z = \text{diag}(z_1, \dots, z_d)$, com $z_j = e^{s_j \Delta t}$, $j = 1, 2, \dots, d$, $R = \text{diag}(r_1, \dots, r_d)$, V_M uma matriz de Vandermonde, $M \times d$ descrita por:

$$V_M = \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & \dots & z_d \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{M-1} & z_2^{M-1} & \dots & z_d^{M-1} \end{bmatrix}_{M \times d} \quad (2.9)$$

e W é a transposta da submatriz de V_M , formada pelas suas N primeiras linhas.

Uma conseqüência da decomposição (2.8) é que para todo $l \geq 0$,

$$\text{posto}(H(l)) = d,$$

sempre que $M \geq N \geq d$ e $s_i \neq s_j$, $i \neq j$ [1]. Segue também, que o espaço coluna de $H(l)$, denotado por \mathcal{S}_M , é gerado pelos vetores coluna da matriz V_M e chamado de *subespaço sinal*. O subespaço linha é gerado pelas colunas de W^* e também é chamado de subespaço sinal.

A matriz $H(l)$ possui uma propriedade chamada *invariância a deslocamentos*, que significa que o espaço coluna da matriz deslocada $H(l+1)$ está contido no espaço coluna da matriz $H(l)$.

Como na prática as amostras disponíveis são da forma $\tilde{h}_l = h_l + \epsilon_l$, onde os ϵ_l denotam o ruído nas amostras, a matriz de Hankel, $\tilde{H}(l) = H(l) + E$, se comporta como sendo de posto completo.

Capítulo 3

Métodos de Identificação Paramétrica

Neste capítulo, são abordados dois tipos de métodos: as chamadas técnicas de predição linear, e os métodos de subespaço.

Apresenta-se a derivação detalhada dos métodos, bem como uma breve discussão das dificuldades encontradas quando o problema é abordado usando sinais contaminados por ruídos.

3.1 Técnicas de Predição Linear

A idéia principal das técnicas de predição linear é prever amostras futuras do sinal, usando uma combinação de amostras passadas. Especificamente, se h_l , $l \geq 0$ são amostras do sinal (2.1), então um modelo de predição linear de ordem N ($N \geq d$) assume que a amostra futura h_{l+N} tem a forma:

$$c_1 h_l + c_2 h_{l+1} + \cdots + c_N h_{l+N-1} = h_{l+N}, \quad l \geq 0. \quad (3.1)$$

Nesta formulação, os coeficientes c 's, são chamados *parâmetros preditores*.

Um resultado importante relacionado ao modelo de predição linear é o fato dos parâmetros que entram de maneira não linear no sinal, (α_j, ω_j) , poderem ser extraídos a partir do conhecimento dos coeficientes c_j . Para ver isto, será mostrado que os parâmetros $s_j = \alpha_j + i\omega_j$ podem ser extraídos a partir das raízes z_j ; $1 \leq j \leq N$ do polinômio:

$$P_p(t) = c_1 + c_2 t + \cdots + c_N t^{N-1} - t^N,$$

o qual é chamado *polinômio preditor progressivo*, onde os parâmetros c_j são calculados

a partir da solução do sistema:

$$\begin{bmatrix} h_l & h_{l+1} & \cdots & h_{l+N-1} \\ h_{l+1} & h_{l+2} & \cdots & h_{l+N} \\ \vdots & \vdots & \cdots & \vdots \\ h_{l+M-1} & h_{l+M} & \cdots & h_{l+M+N-2} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} = \begin{bmatrix} h_{l+N} \\ h_{l+N-1} \\ \vdots \\ h_{l+M+N-1} \end{bmatrix}. \quad (3.2)$$

Lema 3.1. *Se $N \geq d$, então as raízes z_j são autovalores da matriz companheira \mathcal{C} dada por*

$$\mathcal{C} = \begin{bmatrix} 0 & 0 & \cdots & 0 & c_1 \\ 1 & 0 & \cdots & 0 & c_2 \\ 0 & 1 & \cdots & 0 & c_3 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & c_N \end{bmatrix}_{N \times N},$$

onde $\mathbf{c} = [c_1, c_2, \dots, c_N]^T$ é obtido pela solução de (3.2).

Demonstração:

Esta afirmação é facilmente verificada através do cálculo do polinômio característico de \mathcal{C} , que é exatamente $P_p(t)$. De fato, como a matriz \mathcal{C} satisfaz a relação

$$H(l+1) = H(l)\mathcal{C},$$

usando a decomposição $H(l) = V_M Z^l R W$, veja (2.8), obtém-se que $W\mathcal{C} = ZW$, isto é, as linhas de W são autovetores à esquerda de \mathcal{C} , associados aos autovalores z_1, \dots, z_d . \square

Do lema 3.1, segue que d zeros do polinômio $P_p(t)$ são da forma $z_l = e^{s_l \Delta t}$, $l = 1, \dots, d$, os quais são chamados *zeros do sinal*.

Analogamente, um modelo de predição na ordem inversa, é da forma:

$$d_1 h_l + d_2 h_{l+1} + \cdots + d_N h_{l+N-1} = h_{l-1}, \quad l \geq 1. \quad (3.3)$$

Os parâmetros s_j são obtidos agora dos zeros de polinômios preditores regressivos:

$$P_r(t) = d_N + d_{N-1}t + \cdots + d_1 t^{N-1} - t^N,$$

cujos coeficientes são obtidos pela solução de $H(l+1)\mathbf{d} = \mathbf{h}_l$, com $\mathbf{h}_l = [h_l, \dots, h_{l+M-1}]^*$. Neste caso, o problema de calcular os zeros do polinômio $P_r(t)$ é equivalente ao problema de calcular os autovalores da matriz companheira \mathcal{D} , dada por:

$$\mathcal{D} = \begin{bmatrix} d_1 & 1 & 0 & \cdots & 0 \\ d_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ d_{N-1} & 0 & 0 & \cdots & 1 \\ d_N & 0 & 0 & \cdots & 0 \end{bmatrix}_{N \times N}.$$

Com um raciocínio análogo ao anterior, obtém-se que se $N \geq d$, então d autovalores de \mathcal{D} , são $z_j^{-1} = e^{-s_j \Delta t}$, e portanto, d zeros de $P_r(t)$ são da forma $z_i^{-1} = e^{-s_i \Delta t}$.

Em ambos os polinômios, existem $N - d$ zeros sem nenhum significado físico e que decorrem do uso de um polinômio com ordem maior do que o necessário, visto que d é desconhecido a priori. É preciso então, separar os d zeros não desejados, porém esta separação não é simples, uma vez que os zeros não desejados dependem da escolha dos coeficientes c_i (d_i), feita dentre as infinitas soluções do sistema (3.2). De fato, o referido sistema apesar de ser consistente (o lado direito pertence ao espaço coluna da matriz $H(l)$), sua matriz $H(l)$ é de posto incompleto ($\text{posto}(H(l)) = d$).

As técnicas de predição linear consistem, então, em extrair os parâmetros desejados a partir dos zeros dos polinômios preditores, o que deve ser feito por meio de algum critério de separação, pois a ordem do polinômio preditor é maior do que o número de parâmetros desejados.

No caso do polinômio progressivo, se os coeficientes c_i são a solução de norma mínima do problema (3.2), isto é se $c = H(l)^\dagger \mathbf{h}_{l+N}$, onde $c = [c_1, \dots, c_N]^*$ e $\mathbf{h}_{l+N} = [h_{l+N}, \dots, h_{l+M+N-1}]^*$, então os zeros espúrios (não desejados) possuem módulo menor do que 1, [2]. Desta forma, a separação entre os zeros do sinal e os zeros não desejados não é uma tarefa simples.

Já no caso do polinômio regressivo, se os coeficientes d_i são a solução de norma mínima do sistema $H(l+1)\mathbf{d} = \mathbf{h}_l$, com $\mathbf{h}_l = [h_l, \dots, h_{l+M-1}]^*$, isto é, se $\mathbf{d} = H(l+1)^\dagger \mathbf{h}_l$, então os zeros do sinal estão localizados em $z = e^{-s_l \Delta t}$, $l = 1, \dots, d$, e neste caso possuem módulo maior do que 1, enquanto os $N - d$ zeros não desejados possuem módulo menor do que 1, ou seja, estão localizados no interior do círculo unitário [2]. Neste caso, é mais simples fazer a separação, porém quando o nível de ruídos for muito alto, esta separação torna-se complicada. Em [3], é apresentado um método para calcular os coeficientes d_j , pelo método dos Gradientes Conjugados. Um outro método é o de Kumaresan e Tufts [23], onde os coeficientes são estimados a partir da SVD truncada.

3.2 Métodos de Subespaço

Os métodos do tipo subespaço podem ser divididos em duas classes, [41]: Técnicas de Ajuste de Subespaço [42] e Métodos invariantes a deslocamentos (*Single Shift Invariant* - SSI), [24, 1, 21, 19, 41].

Os métodos de ajuste de subespaço utilizam técnicas de otimização para resolver um problema de otimização não linear separável, onde algumas das incógnitas entram linearmente no problema e outras não linearmente. Se não for dado um bom chute inicial, o método pode não convergir, ou encontrar mínimos

locais que não correspondem à solução procurada. Devido a estas dificuldades, estas técnicas não serão abordadas neste trabalho.

A seguir são considerados os métodos invariantes a deslocamentos, os quais exploram a propriedade da invariância a deslocamentos, descrita na seção 2.2. Seja V_M a matriz definida como em (2.8), cujas colunas geram o espaço sinal de $H(l)$, se \mathcal{A} é uma matriz formada de todas as linhas da matriz V_M , exceto da última, e \mathcal{B} é formada por de todas as linhas de V_M , exceto da primeira, então $\mathcal{A}Z = \mathcal{B}$. A propriedade da invariância a deslocamentos da matriz de Hankel assegura que se \mathcal{V} é outra matriz, cujas colunas geram o espaço coluna de $H(l)$, e se A e B são obtidas a partir de \mathcal{V} , assim como \mathcal{A} e \mathcal{B} foram obtidas a partir de V_M , então existe uma matriz T $d \times d$, chamada *matriz de transição*, tal que:

$$AT = B \quad (3.4)$$

Lema 3.2. *Os autovalores de T são $z_l = e^{s_l \Delta t}$, $l = 1, \dots, d$.*

Demonstração:

Como as colunas de V_M e as colunas de \mathcal{V} geram o subespaço sinal \mathcal{S}_M , existe uma matriz $F \in \mathbb{C}^{d \times d}$, não singular, tal que

$$\mathcal{V} = V_M F. \quad (3.5)$$

Decompondo \mathcal{V} e V_M como

$$\mathcal{V} = \begin{bmatrix} A \\ x^* \end{bmatrix} = \begin{bmatrix} y^* \\ B \end{bmatrix}, \quad V_M = \begin{bmatrix} V_{M-1} \\ e^* Z^{M-1} \end{bmatrix} = \begin{bmatrix} e^* \\ V_{M-1} Z \end{bmatrix},$$

onde x^* é a última linha de \mathcal{V} e y^* a primeira linha de \mathcal{V} , usando esta decomposição e usando (3.5), segue que $B = V_{M-1} Z F$ e $A = V_{M-1} F$. Substituindo A e B em (3.4), tem-se

$$(V_{M-1} F) T = V_{M-1} Z F,$$

ou seja,

$$T = F^{-1} V_{M-1}^\dagger V_{M-1} Z F = F^{-1} Z F,$$

pois $V_{M-1}^\dagger V_{M-1} = I$. Logo, os autovalores de T são os z_l , $l = 1, 2, \dots, d$. \square

O mesmo raciocínio pode ser feito usando uma matriz $\mathcal{U} \in \mathbb{C}^{N \times d}$, cujas colunas geram o espaço linha de $H(l)$. Neste caso, a equação (3.4) envolve uma matriz de transição semelhante a Z^* .

As diferentes escolhas de \mathcal{V} e diferentes formas de resolver (3.4) resultam em vários métodos do tipo subespaço. A seguir, são apresentados alguns destes métodos.

Método de Kung

O primeiro é um método bastante conhecido, o método de Kung [24], o qual foi criado em 1978, e é conhecido em Ressonância Magnética Nuclear, como HSVD [21]. O método está baseado na SVD de $H(l)$, a matriz \mathcal{V} é a matriz dos vetores singulares à esquerda associados aos valores singulares não nulos. Neste método a solução de (3.4) é dada no sentido dos quadrados mínimos lineares, isto é, a matriz de transição é calculada via pseudo-inversão. Assim:

$$T_K = A^\dagger B. \quad (3.6)$$

Na prática, a pseudo-inversa é eficientemente calculada sem ser necessário fazer nenhuma inversão de matriz. De fato, suponha que a SVD de $H(l)$ é dada por:

$$H(l) = U\Sigma V^* = [U_1 \ U_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix} = U_1 \Sigma_1 V_1^*, \quad (3.7)$$

onde U_1 , e V_1 são as matrizes dos vetores singulares associados aos valores singulares não nulos, Σ_1 é uma matriz diagonal, cujos elementos da diagonal são $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_d > 0$.

Escrevendo,

$$U_1 = \begin{bmatrix} A \\ u^* \end{bmatrix},$$

onde u^* é a última linha da matriz U_1 , e observando que $I = U_1^* U_1 = uu^* + AA^*$, obtém-se:

$$A^\dagger = (A^* A)^{-1} A^* = (I - uu^*)^{-1} A^* = \left(I + \frac{uu^*}{1 - u^* u} \right) A^*. \quad (3.8)$$

Desta forma, é possível calcular a matriz de transição da equação (3.6), usando a relação (3.8).

Método HTLS

Se a equação (3.4) for resolvida no sentido dos quadrados mínimos totais, obtém-se o algoritmo chamado HTLS, o qual extrai os autovalores do sinal da matriz:

$$T_S = -\hat{V}_{12}(\hat{V}_{22})^{-1}$$

onde \hat{V}_{12} , \hat{V}_{22} , são matrizes $d \times d$, obtidas da SVD da matriz aumentada

$$[B \ A] = \hat{U} \hat{\Sigma} \hat{V}^*, \text{ com } \hat{V} = \begin{bmatrix} \hat{V}_{11} & \hat{V}_{12} \\ \hat{V}_{21} & \hat{V}_{22} \end{bmatrix}.$$

Este método foi desenvolvido por S. Van Huffel, [21], e é muito utilizado em problemas de ressonância magnética nuclear.

Considere agora a matriz $H(l)$ e a matriz $H(l+1)$. Usando (2.8) tem-se que as matrizes $H(l)$ e $H(l+1)$ podem ser fatoradas em:

$$H(l) = \bar{V}\bar{W}, \quad H(l+1) = \bar{V}Z\bar{W}$$

onde $\bar{V} = V_M Z^l$ possui dimensão $M \times d$ e $\bar{W} = RW$ possui dimensão $d \times N$. Como \bar{V} e \bar{W} são de posto completo, tem-se

$$Z = \bar{V}^\dagger H(l+1) \bar{W}^\dagger.$$

Aplicando uma outra decomposição de posto completo em $H(l)$, $H(l) = FG$, obtém-se uma matriz T , semelhante a Z :

$$T = F^\dagger H(l+1) G^\dagger. \quad (3.9)$$

Das diversas escolhas para as matrizes F e G , são obtidos diversos métodos para determinar o subespaço sinal. A seguir, são apresentados dois métodos baseados na decomposição (3.7).

Método de Zeiger e Mac-Ewen

Primeiramente é apresentado o método de Zeiger e Mac-Ewen [48], o qual faz as seguintes escolhas para as matrizes F e G :

$$F = U_1 \Sigma_1^{1/2}, \quad G = \Sigma_1^{1/2} V_1^*.$$

Substituindo estas escolhas em (3.9), obtém-se a seguinte matriz de transição T :

$$T_Z = \Sigma_1^{-1/2} U_1^* H(l+1) V_1 \Sigma_1^{-1/2}. \quad (3.10)$$

Método OPIA

Uma outra escolha das matrizes F e G leva ao método chamado **OPIA** (Optimized Pseudo Inverse Algorithm), desenvolvido por Bazán [1], o método não envolve $H(l+1)$ nos cálculos, e é obtido a partir de (3.10). Inicialmente observe que todas as colunas de $H(l+1)$, exceto a última, são exatamente as últimas colunas de $H(l)$, portanto $H(l+1) = H(l)\mathcal{C}$, onde \mathcal{C} é uma matriz companheira. A matriz de transição é calculada usando esta relação em (3.10), com $H(l)$ como em (3.7), obtém-se:

$$T_O = V_1^* \mathcal{C} V_1,$$

a qual é semelhante à matriz T_Z , pois $T_Z = \Sigma^{1/2} T_O \Sigma^{-1/2}$.

Observação 3.3. *Uma vantagem dos métodos de subespaço em relação às técnicas de predição linear é a dimensão do problema de autovalor a ser resolvido, que no primeiro caso é $d \times d$, e no segundo caso é $N \times N$, $N \gg d$.*

3.3 O Problema de Identificação em Situações Reais

Os métodos descritos neste capítulo foram desenvolvidos a partir do sinal exato, isto é, em situação ideal, onde não há presença de ruídos. Porém, quando estes métodos são aplicados em problemas práticos, como já foi citado anteriormente, os sinais em questão são perturbados por incertezas de diversas naturezas, tais como; erros nos instrumentos de medição, variação de temperatura, erros de arredondamento, etc. Ou seja, na prática, o que se possui são amostras do tipo: $\tilde{h}_j = h_l + \epsilon_j$, e se utiliza uma matriz perturbada $\tilde{H}(l) = H(l) + E$.

Surgem então algumas dificuldades. A primeira delas é a detecção do número de exponenciais, pois $\tilde{H}(l)$ se comporta como sendo de posto completo. Se o posto não for estimado corretamente, os coeficientes \tilde{c}_j , podem diferir enormemente dos c_j , uma vez que a pseudo inversa é sensível ao posto e a perturbações nos dados. Analogamente, o mesmo problema ocorre com a estimativa dos coeficientes d_j . No caso dos métodos de subespaço, se o posto estimado for menor do que o posto real, o subespaço computado $\tilde{\mathcal{S}}_M$ não irá aproximar corretamente o subespaço sinal exato \mathcal{S}_M , e se o posto for superestimado, serão introduzidas em $\tilde{\mathcal{S}}_M$ informações associadas ao ruído.

Uma segunda dificuldade é a sensibilidade do problema de autovalor, uma vez que em ambas as classes de métodos apresentadas (Predição Linear e Métodos de Subespaço) o problema central a ser resolvido é um problema de autovalor. Surge assim, uma pergunta natural: O que ocorre com os autovalores calculados nestas condições? Ou seja: Qual a sensibilidade dos z 's em função do erro E ?

Como as técnicas de predição linear estão relacionadas com o cálculo de raízes de polinômios, elas são mais sensíveis a perturbações nos dados, ver [19]. Por esta razão, estas técnicas não serão abordadas neste trabalho.

O próximo capítulo analisa a sensibilidade dos autovalores associados aos métodos de subespaço, que utilizam a propriedade da invariância a deslocamentos.

Capítulo 4

Sensibilidade dos Autovalores do Sinal

Quando os sinais são contaminados por ruídos, a matriz de transição T utilizada nos métodos de subespaço é estimada a partir de um sistema não compatível. Desta forma é importante saber como se comportam os autovalores do sinal nestas condições. Neste capítulo, é feita uma análise da sensibilidade dos autovalores, para os métodos que utilizam a propriedade da invariância a deslocamentos; assume-se que o número de exponenciais, d , é conhecido previamente. É mostrado que a sensibilidade dos autovalores do sinal é governada pelo número de condição $\kappa(V_M)$, bem como são feitas estimativas para o erro nos autovalores.

4.1 Autovalores do Sinal

Seja $V_S \in \mathbb{C}^{M \times d}$ uma matriz com colunas ortonormais, as quais geram \mathcal{S}_M , o subespaço sinal (o subespaço associado ao sinal livre de ruídos). O próximo teorema fornece informação sobre o condicionamento do problema de autovalor, associado com os métodos de subespaço do tipo SSI.

Teorema 4.1. *Suponha que $T \in \mathbb{C}^{d \times d}$ satisfaz a equação da invariância a deslocamentos (3.4), com A , e B obtidas a partir de V_S . Então, para $M \geq d + 1$*

$$T = (V_S^* V_M) Z (V_S^* V_M)^{-1},$$

onde V_M , é a matriz de Vandermonde dada em (2.9).

Demonstração:

Como as colunas de V_M também geram o subespaço \mathcal{S}_M , existe uma matriz $F \in \mathbb{C}^{d \times d}$, não singular, tal que

$$V_S = V_M F. \tag{4.1}$$

Decompondo V_S e V_M como

$$V_S = \begin{bmatrix} A \\ x^* \end{bmatrix} = \begin{bmatrix} y^* \\ B \end{bmatrix}, \quad V_M = \begin{bmatrix} V_{M-1} \\ e^* Z^{M-1} \end{bmatrix} = \begin{bmatrix} e^* \\ V_{M-1} Z \end{bmatrix},$$

onde x^* é a última linha de \mathcal{V} e y^* a primeira linha de \mathcal{V} . Usando esta decomposição e usando (4.1), segue que $B = V_{M-1} Z F$, $A = V_{M-1} F$ e $A^\dagger = F^{-1} V_{M-1}^\dagger$. Substituindo A^\dagger e B em (3.4), tem-se

$$T = F^{-1} V_{M-1}^\dagger V_{M-1} Z F = F^{-1} Z F,$$

pois $V_{M-1}^\dagger V_{M-1} = I$. De (4.1) segue que, $F^{-1} = V_S^* V_M$ e $F = V_M^\dagger V_S = (V_S^* V_M)^{-1}$, logo

$$T = (V_S^* V_M) Z (V_S^* V_M)^{-1}.$$

□

É interessante observar que $\kappa_2(V_S^* V_M) = \kappa_2(V_M)$. De fato, note que

$$(V_S^* V_M)^* (V_S^* V_M) = V_M^* V_S V_S^* V_M = V_M^* V_M,$$

pois $V_S V_S^*$ é o operador projeção ortogonal em \mathcal{S}_M e as colunas de $V_M \in \mathcal{S}_M$. Isto é, ambas as matrizes $V_S^* V_M$ e V_M possuem os mesmos valores singulares.

Assim, o Teorema 4.1, mostra que a sensibilidade dos autovalores do sinal é governada por $\kappa_2(V_M)$. Desta forma, torna-se necessário uma análise do comportamento de $\kappa_2(V_M)$.

4.2 Condicionamento de Matrizes de Vandermonde

O objetivo aqui é obter limitantes para o número de condição $\kappa(V_M)$. Os resultados obtidos decorrem de analisar matrizes de Vandermonde do tipo $W_M \in \mathbb{C}^{d \times M}$:

$$W_M = \begin{bmatrix} 1 & z_1 & z_1^2 & \cdots & z_1^{M-1} \\ 1 & z_2 & z_2^2 & \cdots & z_2^{M-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & z_d & z_d^2 & \cdots & z_d^{M-1} \end{bmatrix}. \quad (4.2)$$

4.2.1 Estimativas Preliminares

O próximo teorema fornece informações sobre o comportamento de $\|W_M^\dagger\|$ como uma função de M .

Teorema 4.2. *Seja W_M uma matriz de Vandermonde, descrita como em (4.2), com z_j^{i-1} , $z_j \neq z_i$, $|z_j| < 1$. Então para $M > d$, $\|W_M^\dagger\|$ decresce monotonamente com M , e este decréscimo é estrito ao longo da subsequência de inteiros $\{d, 2d, \dots\}$. Além disso, se $M = p \cdot d$, $p \geq 1$, então*

$$\|W_M^\dagger\| \leq \frac{\|W_d^{-1}\|}{\sqrt{1 + \beta^{2d} + \beta^{4d} + \dots + \beta^{2(p-1)d}}} \quad (4.3)$$

onde $\beta = \min|z_j|$, $j = 1, \dots, d$.

Demonstração:

A demonstração pode ser encontrada em [5].

Como $\|W_M^\dagger\|$ decresce monotonamente com M , este primeiro resultado leva a esperar que $\kappa_2(W_M)$ melhore à medida que M cresce, apesar do limitante (4.3) poder ser muito pessimista se W_d for muito mal condicionada. Apesar disto, ele pode ser muito útil para fins teóricos, como é visto a seguir.

Teorema 4.3. *Seja \hat{f}_M a solução de norma mínima do sistema indeterminado*

$$W_M f = Z^M e,$$

onde $Z = \text{diag}(z_1, z_2, \dots, z_d)$, e é o vetor $e = [1, 1, \dots, 1]^T \in \mathbb{R}^d$. Suponha que todos os z_j satisfazem $|z_j| = 1$, ou $|z_j| < 1$, então sempre que $M \rightarrow \infty$ tem-se

$$\|\hat{f}_M\| \rightarrow 0.$$

Demonstração:

Observe que $\|\hat{f}_M\| = \|W_M^\dagger Z^M e\| \leq \|W_M^\dagger\| \|Z^M e\| \leq \|W_M^\dagger\| \sqrt{d} \alpha^M$ onde $\alpha = \max|z_j|$, $j = 1, 2, \dots, d$.

Usando o teorema 4.2 obtém-se

$$\|\hat{f}_M\| \leq \frac{\|W_d^{-1}\|}{\sqrt{1 + \beta^{2d} + \beta^{4d} + \dots + \beta^{2(N/d-1)d}}} \sqrt{d} \alpha^M$$

e tomando limite quando $M \rightarrow \infty$, obtém-se $\|\hat{f}_M\| \rightarrow 0$. □

O vetor \hat{f}_M descrito acima é conhecido em problemas de predição linear como *vetor de parâmetros preditores* (em [2] podem ser encontrados mais detalhes). Neste contexto, ele prediz a última coluna de W_{M+1} . Este fato produz a seguinte relação:

$$Z W_M = W_M \mathcal{F}_M \quad (4.4)$$

onde \mathcal{F}_M é uma matriz companheira $M \times M$, da forma

$$\mathcal{F}_M = [e_2, e_3, \dots, e_{M-1}, \hat{f}_M],$$

onde e_i , é o i -ésimo vetor canônico em \mathbb{R}^M .

Defina F_M por

$$F_M = \mathcal{W}_M^* \mathcal{F}_M \mathcal{W}_M, \quad (4.5)$$

onde

$$\mathcal{W}_M = W_M^* (W_M W_M^*)^{-1/2}.$$

Observe que F_M é bem definida, pois $W_M W_M^*$ é positiva definida. Obtém-se então, o seguinte teorema:

Teorema 4.4. *A matriz F_M descrita em (4.5) possui uma decomposição autovalor autovetor dada por:*

$$F_M = Q_M Z Q_M^{-1}, \quad (4.6)$$

onde $Q_M = (W_M W_M^*)^{-1/2}$. Além disso, esta matriz de autovetores satisfaz $\kappa_2(Q_M) = \kappa_2(W_M)$.

Demonstração:

De (4.5) e de (4.4) tem-se

$$\begin{aligned} F_M &= \mathcal{W}_M^* \mathcal{F}_M \mathcal{W}_M = (W_M W_M^*)^{-1/2} W_M \mathcal{F}_M W_M^* (W_M W_M^*)^{-1/2} \\ &= (W_M W_M^*)^{-1/2} Z (W_M W_M^*) (W_M W_M^*)^{-1/2} \\ &= (W_M W_M^*)^{-1/2} Z (W_M W_M^*)^{1/2} \end{aligned}$$

Falta agora demonstrar que, $\kappa_2(Q_M) = \kappa_2(W_M)$. De fato,

$$\begin{aligned} \kappa_2(Q_M) &= \|Q_M\| \|Q_M^\dagger\| \\ &= \|(W_M W_M^*)^{-1/2}\| \|(W_M W_M^*)^{1/2}\| \\ &= \|W_M^{-1}\| \|W_M\| = \kappa_2(W_M). \end{aligned}$$

□

Muitos dos resultados aqui apresentados sobre limitantes de $\kappa_2(W_M)$, dependem dos autovalores e dos valores singulares de F_M . O próximo teorema caracteriza o espectro singular de F_M .

Teorema 4.5. *Seja F_M definida como em (4.5). Então seu espectro singular é descrito por*

$$\begin{aligned} \sigma_1^2(F_M) &= \frac{2 + \|\hat{f}_M\|^2 - \|p_1\|^2 + \sqrt{(\|\hat{f}_M\|^2 + \|p_1\|^2)^2 - 4|f_1|}}{2} \\ \sigma_j^2(F_M) &= 1, \quad j = 2, \dots, d-1 \\ \sigma_d^2(F_M) &= \frac{2 + \|\hat{f}_M\|^2 - \|p_1\|^2 - \sqrt{(\|\hat{f}_M\|^2 + \|p_1\|^2)^2 - 4|f_1|}}{2} \end{aligned} \quad (4.7)$$

onde p_1 é a primeira coluna de \mathcal{P}_M , o projetor ortogonal sobre $\mathcal{R}(W_M^*)$, e f_1 a primeira componente do vetor \hat{f}_M , introduzido no Teorema 4.3

Demonstração:

A demonstração pode ser encontrada em [5].

4.2.2 Limitantes para $\kappa_2(W_M)$

O condicionamento dos autovalores a perturbações na matriz pode ser estimado como uma função da proximidade entre os valores singulares e o valor absoluto dos autovalores, veja [33]. Assim, quanto mais próximos os valores singulares do valor absoluto dos autovalores, mais bem condicionado é o problema de autovalor.

Os resultados sobre limitantes para $\kappa_2(W_M)$ são dados no teorema a seguir.

Teorema 4.6. *Seja W_M uma matriz de Vandermonde definida como em (4.2), com z 's no disco unitário. Defina $\alpha = \max_j |z_j|$, $\beta = \min_j |z_j|$ e $\delta = \min_{\substack{j \neq k \\ j, k}} |z_j - z_k|$. Defina*

$$D_M^2 = D(F_M) = \|F_M\|_F^2 - (|z_1|^2 + \dots + |z_d|^2).$$

Então, para $M > d \geq 2$, o número de condição associado à norma-2 de W_M satisfaz

$$\frac{\sigma_1(F_M)}{\alpha} \leq \kappa_2(W_M) \leq \frac{1}{2} \left(\eta + \sqrt{\eta^2 - 4} \right), \quad (4.8)$$

onde $\eta = \rho - d + 2$,

$$\rho = d \left[1 + \frac{D_M^2}{(d-1)\delta^2} \right]^{\frac{d-1}{2}} \phi_M(\alpha, \beta), \quad (4.9)$$

$$\phi_M(\alpha, \beta) = \sqrt{\frac{1 + \alpha^2 + \alpha^4 + \dots + \alpha^{2(N-1)}}{1 + \beta^2 + \beta^4 + \dots + \beta^{2(N-1)}}}. \quad (4.10)$$

Demonstração:

Para mostrar a desigualdade à esquerda, basta tomar a norma-2 em ambos os lados de (4.6) e usar o teorema 4.4.

Para provar a desigualdade à direita, observe que $\|e_j^* Q_M^{-1}\| = \|e_j^* (W_M W_M^*)^{1/2}\| = \|e_j^* W_M\|$, e usando (1.11) e (1.10), tem-se que

$$\kappa_F(W_M) \leq \rho, \quad (4.11)$$

com ρ como em (4.9).

Será usado agora o Teorema 1, de Smith [33], o qual afirma que, para uma matriz $X \in \mathbb{C}^{d \times d}$ não singular

$$d - 2 + \kappa_2(X) + \kappa_2^{-1}(X) \leq \kappa_F(X).$$

Mas, $d - 2 + \kappa_2(X) + \kappa_2^{-1}(X) \leq \kappa_F(X) \Leftrightarrow [\kappa_2(X)]^2 - \nu[\kappa_2(X)] + 1 \leq 0$, onde $\nu = \kappa_F(X) - d + 2$. Resolvendo esta inequação do segundo grau obtém-se:

$$\kappa_2(X) \leq \frac{1}{2} \left[\nu + \sqrt{\nu^2 - 4} \right].$$

Adaptando agora ao problema em questão, obtém-se

$$\begin{aligned} \kappa_2(W_M) &\leq \frac{1}{2} \left[\kappa_F(W_M) - d + 2 + \sqrt{(\kappa_F(W_M) - d + 2)^2 - 4} \right] \leq \\ &\leq \frac{1}{2} \left[\rho - d + 2 + \sqrt{(\rho - d + 2)^2 - 4} \right] = \\ &= \frac{1}{2} \left[\eta + \sqrt{\eta^2 - 4} \right] \end{aligned}$$

□

A seguir, o limitante superior em (4.8) é analisado como função de M . Observe que o limitante depende de três fatores: da separação entre os z_j no disco unitário, do número D_M , e de $\phi_M(\alpha, \beta)$. Porém, como a contribuição de $\phi_M(\alpha, \beta)$ não é muito expressiva, uma vez que os z_j estão no disco unitário, a qualidade do limitante depende basicamente da razão $D_M^2/(d-1)\delta^2$. Portanto, quando D_M^2 for da mesma magnitude de $(d-1)\delta^2$, e d não for muito grande, tem-se limitantes moderados. O próximo lema mostra que isto é possível.

Lema 4.7. *Seja D_M^2 como no teorema 4.8. Então para cada $M \geq d$*

$$(d-1) + \frac{\prod_{j=1}^d |z_j|^2}{1 + \|\hat{f}_M\|^2} - \sum_{j=1}^d |z_j|^2 \leq D_M^2 \leq (d-1) + \|\hat{f}_M\|^2 + \prod_{j=1}^d |z_j|^2 - \sum_{j=1}^d |z_j|^2, \quad (4.12)$$

e portanto

$$D_\infty^2 = \lim_{M \rightarrow \infty} D_M^2 = (d-1) + \prod_{j=1}^d |z_j|^2 - \sum_{j=1}^d |z_j|^2.$$

Demonstração:

Usando o fato que o produto dos valores singulares de uma matriz quadrada é igual ao produto do valor absoluto dos correspondentes autovalores, e usando o teorema 4.5, tem-se

$$\sigma_1^2(F_M) \sigma_d^2(F_M) = \prod_{j=1}^d |z_j|^2.$$

conseqüentemente, como $\|F_M\|_F^2 = \sigma_1^2(F_M) + \dots + \sigma_d^2(F_M) = \sigma_1^2(F_M) + \sigma_d^2(F_M) + (d-2)$, usando a definição de D_M^2 , obtém-se:

$$D_M^2 = (d-2) + \sigma_1^2(F_M) + \frac{\prod_{j=1}^d |z_j|^2}{\sigma_1^2(F_M)} - \sum_{j=1}^d |z_j|^2. \quad (4.13)$$

Mas, como

$$1 \leq \sigma_1^2(F_M) \leq 1 + \|\hat{f}_M\|^2, \quad (4.14)$$

pelo teorema 4.5, aplicando (4.14) em (4.13), obtém-se a desigualdade desejada.

Para obter a igualdade envolvendo D_∞^2 , basta tomar $\lim_{M \rightarrow \infty}$ em ambos os lados de (4.12), e usar o teorema 4.3. \square

O lema 4.7 mostra que o comportamento de D_M^2 como função de M , depende da velocidade com que $\|\hat{f}_M\|^2$ converge para zero. Quando M cresce, para M grande, o tamanho de D_M^2 irá depender do tamanho dos z_j . Desta forma, pode-se concluir que, sempre que M for suficientemente grande e $|z_j| \approx 1$, o número D_M^2 será pequeno. conseqüentemente, a não ser que os autovalores sejam muito próximos uns dos outros, a condição $D_M^2 < (d-1)\delta^2$ deve ser satisfeita, assegurando assim, limitantes pequenos para $\kappa_2(W_M)$.

4.3 Erro nos autovalores

Nesta seção serão apresentadas estimativas para o erro $|z_l - \tilde{z}_l|$, $l = 1, \dots, d$, onde \tilde{z}_l , são os autovalores calculados a partir de uma matriz de transição aproximada \tilde{T} . Deseja-se mostrar que, sob certas circunstâncias, os autovalores são insensíveis a pequenas perturbações nos dados. Da teoria clássica de perturbação de autovalores, veja a relação (1.7), e do teorema (4.1) é possível afirmar que

$$|z_l - \tilde{z}_l| \leq \|T - \tilde{T}\| \kappa(V_M), \quad 1 \leq l \leq d. \quad (4.15)$$

Como o comportamento de $\kappa_2(V_M)$ já foi estudado anteriormente, falta estudar agora o comportamento de $\|T - \tilde{T}\|$. O erro $\|T - \tilde{T}\|$, será analisado somente nos casos onde \tilde{T} é estimada através de técnicas de quadrados mínimos, como descrito em (3.6).

Sejam V_S e \tilde{V}_S matrizes $M \times d$, com colunas ortonormais, que geram o subespaço sinal exato \mathcal{S}_M e o subespaço sinal aproximado $\tilde{\mathcal{S}}_M$, respectivamente. Decomponha \tilde{V}_S e V_S como

$$V_S = \begin{bmatrix} A \\ x^* \end{bmatrix} = \begin{bmatrix} y^* \\ B \end{bmatrix}, \quad \tilde{V}_S \begin{bmatrix} \tilde{A} \\ \tilde{x}^* \end{bmatrix} = \begin{bmatrix} \tilde{y}^* \\ \tilde{B} \end{bmatrix},$$

onde x^* e y^* representam a primeira e a última linha de V_S respectivamente, e \tilde{x}^* e \tilde{y}^* a primeira e a última linha de \tilde{V}_S respectivamente.

Observe que, para sinais livres de ruídos, a matriz T é a única solução do sistema (3.4). Sejam $\tilde{A} = A + \Delta_A$, $\tilde{B} = B + \Delta_B$ e assumamos que \tilde{A} possui posto igual a d . Então o erro em T é:

$$\begin{aligned}
\tilde{T} - T &= \tilde{A}^\dagger \tilde{B} - T \\
&= \tilde{A}^\dagger (B + \Delta_B) - T \\
&= \tilde{A}^\dagger (AT + \Delta_B) - T \\
&= \tilde{A}^\dagger (\tilde{A}T - \Delta_A T + \Delta_B) - T \\
&= \tilde{A}^\dagger (-\Delta_A T + \Delta_B).
\end{aligned} \tag{4.16}$$

A última desigualdade segue do fato de $\tilde{A}^\dagger \tilde{A} = I$, pois foi assumido que $\text{posto}(\tilde{A}) = d$. Tomando a norma 2 em ambos os lados de (4.16), tem-se:

$$\|\tilde{T} - T\| \leq \|\tilde{A}^\dagger\| \|T\| (\|\Delta_A\| + \|\Delta_B\|), \tag{4.17}$$

uma vez que $\|T\| \geq 1$ (ver [4]). Observe que, para obter qualquer estimativa de erro significativa, V_S deve ser escolhida o mais próximo possível de \tilde{V}_S . Isto é sempre possível, e pode ser feito escolhendo $V_S = VX$, onde V é qualquer matriz $M \times d$, cujas colunas são ortonormais e geram o subespaço \mathcal{S}_M , e $X \in \mathbb{C}^{d \times d}$ é uma matriz unitária, a qual resolve o problema de Procrustes Ortogonal:

$$\min \|VX - \tilde{V}_S\|_F$$

Defina $G = \tilde{V}_S^* V$, e considere a sua decomposição em valores singulares, $G = \bar{P} \bar{\Sigma} \bar{Q}^*$. Então, a matriz unitária que resolve o problema de Procrustes é $X = \bar{Q} \bar{P}^*$, maiores detalhes sobre a solução deste problema podem ser encontrados em [15]. Usando esta escolha de X , tem-se $V_S = V \bar{Q} \bar{P}^*$. E obtém-se que

$$\begin{aligned}
\|V_S - \tilde{V}_S\|^2 &= \|(V \bar{Q} \bar{P}^* - \tilde{V}_S)^* (V \bar{Q} \bar{P}^* - \tilde{V}_S)\| \\
&= \|2I - \bar{P} \bar{Q}^* V^* \tilde{V}_S - \tilde{V}_S^* V \bar{Q} \bar{P}^*\| \\
&= \|2I - \bar{P} \bar{Q}^* G^* - G \bar{Q} \bar{P}^*\| \\
&= 2\|I - \bar{P} \bar{\Sigma} \bar{P}^*\|
\end{aligned}$$

Da seção 1.6, capítulo 1, tem-se que os cosenos dos ângulos canônicos entre os subespaços \mathcal{S}_M e $\tilde{\mathcal{S}}_M$ são os valores singulares de $\tilde{V}_S^* V$, os quais estão contidos em $\bar{\Sigma}$. Usando esta observação, a desigualdade acima implica que

$$\|\Delta V_S\| = \|V_S - \tilde{V}_S\| = 2 \sin \frac{\theta}{2}, \tag{4.18}$$

onde θ é o ângulo entre os subespaços \mathcal{S}_M e $\tilde{\mathcal{S}}_M$ (maior ângulo canônico entre \mathcal{S}_M e $\tilde{\mathcal{S}}_M$, veja seção 1.6). Como $\|\Delta_A\| \leq \|\Delta V_S\|$ e $\|\Delta_B\| \leq \|\Delta V_S\|$, usando estas desigualdades em (4.18), obtém-se o próximo lema:

Lema 4.8.

$$\|T - \tilde{T}\| \leq 4 \|\tilde{A}^\dagger\| \|T\| \sin \frac{\theta}{2} \quad (4.19)$$

É analisado a seguir o comportamento de $\|\tilde{A}^\dagger\|$ e $\|T\|$, como função de M , o número de linhas da matriz $\tilde{H}(l)$.

Seja C uma matriz companheira, $M \times M$, tal que

$$CV_M = V_M Z,$$

onde V_M , é uma matriz de Vandermonde, descrita como em (2.9) e C é dada por

$$C = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ c_1 & c_2 & \cdots & c_{M-1} & c_M \end{bmatrix}_{M \times M},$$

com $c = [c_1, \dots, c_M]^*$ obtido pela solução de $cV_M = e^*Z^M$. Então, como $V_S V_S^* V_M = V_M$, pois $V_S V_S^*$ é o operador projeção ortogonal em \mathcal{S}_M e as colunas de V_M pertencem a este espaço, a equação acima pode ser rescrita como

$$CV_S V_S^* V_M = V_S V_S^* V_M Z.$$

Usando o fato de $V_S^* V_S = I$ e o teorema 4.1, tem-se

$$V_S^* C V_S = (V_S^* V_M) Z (V_S^* V_M)^{-1} = T.$$

$$\Rightarrow \|T\| = \|V_S^* C V_S\| \leq \sqrt{1 + \|\hat{f}_M\|^2}.$$

A desigualdade decorre do Teorema 2 de [4]. Agora, como já foi visto que $\|\hat{f}_M\|$ decresce para zero, à medida que M cresce, então para M suficientemente grande, tem-se $\|T\| \approx 1$. Observe a seguir que $\|\tilde{A}^\dagger\| = 1/(1 - \|\tilde{x}\|^2) = 1/(1 - \|\tilde{p}_M\|^2)$, onde \tilde{p}_M é a última coluna do operador projeção ortogonal no subespaço $\tilde{\mathcal{S}}_M$. Isto sugere que se $\|\tilde{p}_M\|$ se comportar aproximadamente como $\|p_M\|$, a qual decresce com M e permanece próxima de 0, para M suficientemente grande, [4], então $\|\tilde{A}^\dagger\|$ não deve ser muito maior do que 1. Desta forma, conclui-se que a estimativa dada no lema 4.8, depende fortemente de M , apesar da proximidade de \mathcal{S}_M e $\tilde{\mathcal{S}}_M$, medida pelo ângulo entre os subespaços também ser importante. Isto significa que para valores de M grandes, o limitante (4.19) fica próximo de $2 \sin \theta$, a menos que \mathcal{S}_M esteja muito distante de $\tilde{\mathcal{S}}_M$. Apesar deste resultado não provar que os erros nas entradas da matriz de dados não são propagados no cálculo de \tilde{T} , ele assegura pequenos erros em T , desde que M seja suficientemente grande e o ângulo entre subespaço suficientemente pequeno.

Substituindo (4.19) em (4.15) obtém-se o próximo teorema

Teorema 4.9. *O erro nos autovalores satisfaz*

$$|z_l - \tilde{z}_l| \leq 2 \|\tilde{A}^\dagger\| \|T\| (\eta + \sqrt{\eta^2 - 4}) \operatorname{sen} \frac{\theta}{2}, \quad (4.20)$$

onde η é dado no teorema 4.6.

Como já foi discutido na seção anterior, para valores de M suficientemente grandes, η torna-se uma constante moderada, o comportamento de $\|\tilde{A}^\dagger\|$ e $\|T\|$ como função de M assegura que, se \tilde{V}_S é suficientemente próxima de V_S , no sentido da norma Frobenius, então o limitante para o erro nos autovalores do sinal pode ser considerado como o produto de $\operatorname{sen} \theta$ por uma constante moderada, isto é

$$|z_l - \tilde{z}_l| = \mathcal{O}(\operatorname{sen} \theta), \quad 1 \leq l \leq d. \quad (4.21)$$

Será assumido agora que \tilde{V}_S é estimada pela SVD da matriz de dados $\tilde{H}(l) = H(l) + E$. O próximo corolário fornece um limitante que depende do tamanho de $\|E\|$ e de $\sigma_d(H(l))$, o menor valor singular não nulo da matriz de dados exata.

Corolário 4.10. *Assuma que uma base para $\tilde{\mathcal{S}}_M$ é calculada pela SVD de $\tilde{H}(l)$. Então, assumindo também $\|E\| \ll \sigma_d(H(l))$, a seguinte estimativa de primeira ordem, para o limitante de $\|T - \tilde{T}\|$ em (4.19) ocorre*

$$\|T - \tilde{T}\| \leq 2 \|\tilde{A}^\dagger\| \|T\| \frac{\|E\|}{\sigma_d(H(l))}. \quad (4.22)$$

conseqüentemente, o limitante do erro nos autovalores do sinal (4.20), torna-se

$$|z_l - \tilde{z}_l| \leq \|\tilde{A}^\dagger\| \|T\| (\eta + \sqrt{\eta^2 - 4}) \frac{\|E\|}{\sigma_d(H(l))}, \quad 1 \leq l \leq d. \quad (4.23)$$

Demonstração:

Sejam a SVD de $H(l)$ e de $\tilde{H}(l)$ respectivamente:

$$H(l) = [U_1 \ U_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix} \quad \text{e} \quad \tilde{H}(l) = [\tilde{U}_1 \ \tilde{U}_2] \begin{bmatrix} \tilde{\Sigma}_1 & 0 \\ 0 & \tilde{\Sigma}_2 \end{bmatrix} \begin{bmatrix} \tilde{V}_1^* \\ \tilde{V}_2^* \end{bmatrix},$$

onde U_1 , \tilde{U}_1 , V_1 e \tilde{V}_1 possuem d colunas. Se $\|E\| \ll \sigma_d(H(l))$, então uma base ortonormal para o subespaço sinal aproximado (gerado pelas colunas de \tilde{U}_1), é formada pelas colunas de uma matriz $M \times d$, digamos \hat{U}_1 , tal que

$$\hat{U}_1 = (U_1 + U_2 P)(I + P^* P)^{-1/2}, \quad (4.24)$$

onde P é uma matriz de ordem $(M - d) \times d$, cuja norma é da ordem de $\|E\|$ (em [40] podem ser encontradas maiores explicações), e $(I + P^* P)^{-1/2}$ é a inversa da

raiz quadrada da matriz positiva definida $(I + P^*P)$. Embora não seja possível encontrar a matriz P , a sua norma $\|P\|$ é interpretada como a tangente do maior ângulo canônico Θ , veja Stewart [35]. Por outro lado, como

$$U_1^* \hat{U}_1 = \hat{U}_1^* U_1 = (I + P^*P)^{-1/2}, \quad (4.25)$$

e, como os valores singulares desta matriz hermitiana são os cossenos dos ângulos canônicos entre \mathcal{S}_M e $\hat{\mathcal{S}}_M$, após algumas manipulações algébricas segue que, $\|U_1 - \hat{U}_1\|_F = 2 \operatorname{sen} \frac{\theta}{2}$. Isto assegura que U_1 é a matriz mais próxima de \hat{U}_1 , no sentido que $\|U_1 - \hat{U}_1\|_F$ é minimizada. Finalmente, como

$$2 \operatorname{sen} \frac{\theta}{2} \leq \tan \Theta = \|P\| \leq \frac{\|E\|}{\sigma_d(H(l))},$$

onde a última desigualdade vale até a primeira ordem de aproximação (veja Vaccaro [40]). As desigualdades (4.22) e (4.23) seguem da substituição desta desigualdade em (4.19) e (4.20), respectivamente. \square

Observe que a desigualdade (4.22) mostra que irão ocorrer reduções nos erros de entrada, sempre que $2 \|\tilde{A}^\dagger\| \|T\| < \sigma_d(H(l))$, um fato que é frequentemente observado se M for suficientemente grande. Por outro lado, esta desigualdade sugere escolher as dimensões da matriz de dados, de forma que o limitante a ser minimizado garanta um ângulo entre subespaços pequeno. Não existe nenhum resultado teórico sobre o comportamento de $\sigma_d(H(l))$ como uma função de M e N , mas evidências empíricas sugerem escolher a matriz de dados o mais próximo possível de uma matriz quadrada. Desta forma, $\sigma_d(H(l))$ pode atingir um valor máximo, [4], e neste caso o limitante pode ser minimizado.

4.4 Resultados Numéricos

Os limitantes, obtidos teoricamente nas seções anteriores, são ilustrados nesta seção, por meio de experimentos numéricos. São apresentados os resultados da análise de um sinal típico de Ressonância Magnética Nuclear, extraído de [21].

O sinal é modelado por (2.7) e possui 5 exponenciais. Os parâmetros que descrevem o sinal, bem como a separação entre os z 's, são dados na tabela 4.1. Os parâmetros z_l , $l = 1, \dots, 5$, satisfazem $|z_l| \approx 1$, e a separação entre eles é satisfatória. O intervalo de amostragem foi $\Delta t = 0.0001s$.

Os experimentos numéricos estão divididos em duas partes. Na primeira parte é ilustrado o comportamento dos limitantes para $\|W_M^\dagger\|$ e $\kappa_2(W_M)$, com o intuito de mostrar que eles se tornam cada vez menores, à medida que M se torna suficientemente grande. Para tanto, é usada uma matriz de Vandermonde de ordem $5 \times M$.

l	r_l	z_l	$ z_l $	δ_l^2
1	$5.8921 + 1.5788z$	$0.6342 - 0.7463z$	0.9794	0.1786
2	$9.5627 + 2.5623z$	$0.8858 - 0.4067z$	0.9747	0.0644
3	$5.7956 + 1.5529z$	$0.9663 - 0.1661z$	0.9805	0.0644
4	$2.7046 + 0.7247z$	$0.9642 + 0.2174z$	0.9884	0.0100
5	$16.4207 + 4.3999z$	$0.8811 + 0.2729z$	0.9224	0.0100

Tabela 4.1: Parâmetros do sinal usado nos testes

O número D_M é calculado usando-se (4.13) e ilustrado na figura 4.2. Na figura 4.1, é ilustrado o rápido decréscimo de $\|\hat{f}_M\|^2$ para 0 à medida que M cresce.

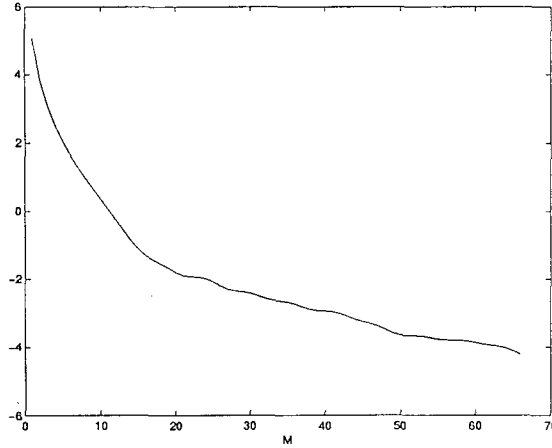


Figura 4.1: Comportamento de $\|\hat{f}_M\|^2$.

Na figura 4.3 pode-se observar o comportamento do limitante (4.8), o qual mostra que $\kappa_2(W_M)$ torna-se cada vez menor, à medida que M torna-se suficientemente grande. A desigualdade que garante valores moderados para este limitante, $D_M^2 < (d-1)\delta^2$, é rapidamente satisfeita, uma vez que, valor de $(d-1)\delta^2$ é 0.0399, e $D_{50}^2 = 0.0396$, $D_{70}^2 = 0.0328$ $D_{100}^2 = 0.0308$.

A seguir, é ilustrado o comportamento dos limitantes para o erro nos autovalores, $|z_l - \tilde{z}_l|$. Foram calculados o limitante (4.15), o erro $|z_l - \tilde{z}_l|$, $\sin \frac{\theta}{2}$, $2 \sin \frac{\theta}{2}$, $\|E\|/\sigma_d(H)$ e $\|T - \tilde{T}\|$. Os \tilde{z}_l foram calculados usando (3.4), as matrizes \tilde{A} e \tilde{B} foram calculadas a partir da SVD da matriz $\tilde{H}(l) = H(l) + E$, cujas entradas são amostras do sinal perturbado $\tilde{h} = h + \epsilon$, onde o ruído ϵ é da ordem de 3%, isto é, $\|\epsilon\|/\|h\| \approx 0.03$. Os resultados apresentados são a média dos valores obtidos para 100 experimentos com ruídos diferentes, porém da mesma ordem.

A figura 4.4 mostra os comportamentos de $\sin \theta$, $2 \sin \frac{\theta}{2}$, $\|E\|/\sigma_d(H)$. Pode-se perceber que estes valores decrescem, à medida que a matriz $\tilde{H}(l)$ se torna uma matriz quadrada.

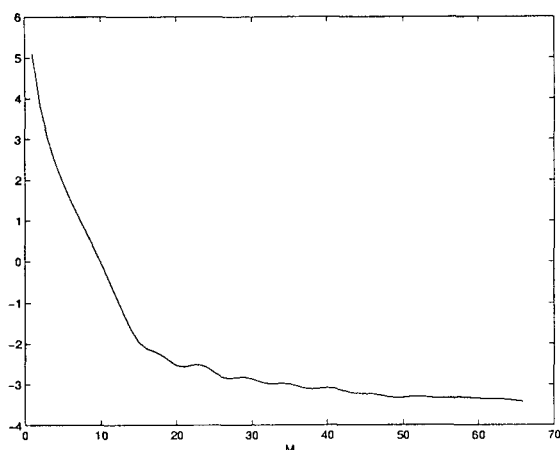


Figura 4.2: Comportamento de D_M^2 .

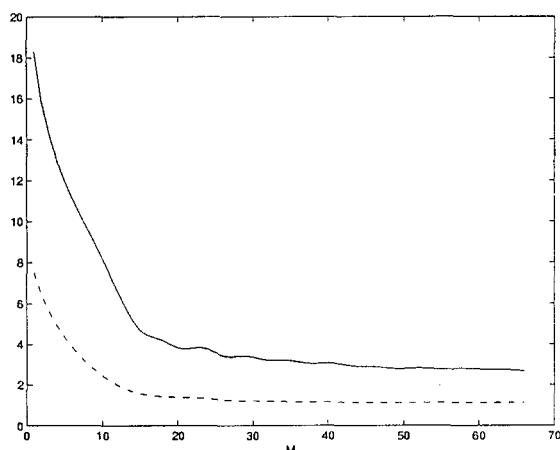


Figura 4.3: Limitante superior para $\kappa_2(W_M)$, obtido no teorema 4.8: linha sólida, e $\kappa_2(W_M)$: linha pontilhada, em escala logarítmica.

O limitante para o erro nos autovalores (4.15), o erro $\|T - \tilde{T}\|$ e o erro propriamente dito, nos autovalores z_2 , z_4 e z_5 são apresentados na figura 4.5, onde se pode novamente perceber que o erro nos autovalores diminui, à medida que a matriz se torna quadrada.

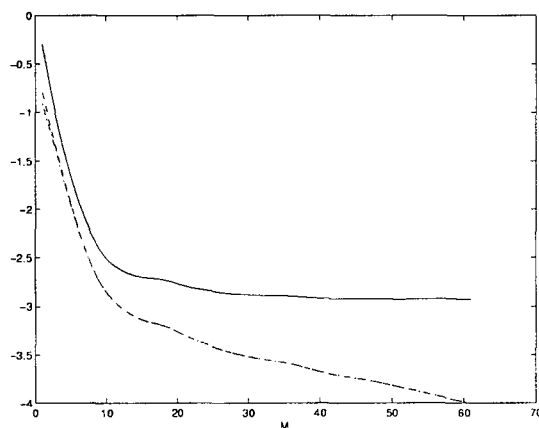


Figura 4.4: $\|E\|/\sigma_d(H(l))$: linha contínua, $\text{sen } \theta$: linha tracejada, $2 \text{sen } \frac{\theta}{2}$: linha tracejada e pontilhada, em escala logarítmica.

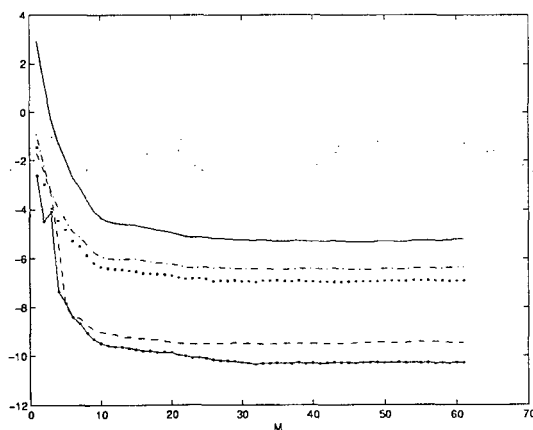


Figura 4.5: Limitante (4.15): linha contínua, $\|T - \tilde{T}\|$: linha tracejada e pontilhada, e $|z_l - \tilde{z}_l|$ para $l = 2, 4, 5$: linhas pontilhada, tracejada e contínua-pontilhada, respectivamente.

Capítulo 5

Computação Rápida do Subespaço Sinal

Neste capítulo, são apresentados algoritmos para calcular o subespaço sinal, os quais utilizam iterações de Lanczos e exploram multiplicações rápidas via FFT (Fast Fourier Transform) [13]. Primeiramente, é apresentado o Procedimento de Rayleigh Ritz, seus aspectos teóricos e computacionais. A seguir, é introduzida a Iteração em Subespaços, e mostrada uma combinação destes dois métodos, a Iteração em Subespaços com aceleração de Rayleigh-Ritz. Finalmente, é apresentado um dos mais eficientes métodos para calcular alguns dos autovalores de uma matriz simétrica, o método de Lanczos, bem como suas variações.

É introduzida uma escolha inicial dos processos iterativos, que garante convergência em no máximo $d + 1$ iterações, quando o sinal for livre de ruído, são mostrados resultados numéricos que ilustram a validade desta propriedade, ainda no caso de dados com ruído, desde que a separação entre $\tilde{\sigma}_d(\tilde{H}(l))$ e $\tilde{\sigma}_{d+1}(\tilde{H}(l))$ seja bem notória. São apresentados também, resultados teóricos sobre valores singulares da matriz Hankel sem ruído, os quais explicam quando pode ser obtida a separação acima descrita, em função do tamanho do ruído nos dados.

5.1 Multiplicação Rápida Matriz-Vetor

Ao se resolver um problema matemático numericamente, além de se desejar encontrar a solução que mais se aproxima da solução real, deseja-se também, que esta solução seja encontrada no menor tempo computacional possível. Assim, foi estudada uma forma de acelerar a multiplicação de uma matriz por um vetor. A multiplicação de uma matriz de Hankel densa $H_{M \times N}$ por um vetor tem MN multiplicações. Porém, aproveitando a sua estrutura Hankel, e usando a Transformada rápida de Fourier (*FFT*), é possível diminuir esta estimativa para $O((M + N) \log_2(M + N))$. Esta idéia é inicialmente apresentada para matrizes de Toeplitz, sendo estendida para

as matrizes de Hankel.

Seja $\mathcal{T} \in \mathbb{C}^{M \times N}$, uma matriz de Toeplitz dada por:

$$\mathcal{T} = \begin{bmatrix} h_M & h_{M+1} & \cdots & h_{M+N-1} \\ h_{M-1} & h_M & \cdots & h_{M+N-2} \\ \vdots & \vdots & \cdots & \vdots \\ h_1 & h_2 & \cdots & h_N \end{bmatrix}. \quad (5.1)$$

A FFT não pode ser usada diretamente para calcular o produto da matriz \mathcal{T} , por um vetor. Para usar a *FFT*, as principais informações da matriz \mathcal{T} são extraídas, e armazenadas em uma matriz circulante, isto é, uma matriz do tipo:

$$C = \begin{bmatrix} c_0 & c_{N-1} & c_{N-2} & \cdots & c_1 \\ c_1 & c_0 & c_{N-1} & \cdots & c_2 \\ c_2 & c_1 & c_0 & \cdots & c_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{N-1} & c_{N-2} & c_{N-3} & \cdots & c_0 \end{bmatrix}_{N \times N}$$

Observe que a matriz $C \in \mathbb{C}^{N \times N}$ é um tipo de matriz de Toeplitz, onde cada coluna é obtida a partir da coluna anterior, deslocando o elemento c_{ij} para a posição c_{i+1j+1} , sendo que o elemento c_{Nj} vai para a posição c_{1j+1} . Assim, é fácil perceber que a matriz C fica bem definida a partir de sua primeira coluna e pode ser escrita como:

$$C = (c \ Rc \ R^2c \ \dots \ R^{N-1}c),$$

onde $c = (c_0 \ c_1 \ \dots \ c_{N-1})^T$ e $R = (e_2 \ e_3 \ e_4 \ \dots \ e_N \ e_1)$.

A partir desta estrutura, podem ser deduzidas várias propriedades das matrizes circulantes, veja [12]. O teorema seguinte relaciona a transformada discreta de Fourier com matrizes circulantes, e é a base do processo de multiplicação rápida.

Teorema 5.1. *Se C é uma matriz circulante, então $C = F_N^* \Delta F_N$, onde F_N é matriz de Fourier:*

$$F_N = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & w^1 & w^2 & \cdots & w^{N-1} \\ 1 & w^2 & w^4 & \cdots & w^{2(N-1)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & w^{N-1} & w^{2(N-1)} & \cdots & w^{(N-1)(N-1)} \end{bmatrix},$$

$w = \exp(-2\pi i/N)$, $i = \sqrt{-1}$, e $\Delta = \text{diag}(F_N c)$, onde c , denota a primeira coluna da matriz C .

Demonstração:

A demonstração pode ser encontrada em [13].

Quando um vetor é multiplicado por F_N , o resultado obtido é a sua *Transformada Discreta de Fourier*. Então, do Teorema acima, tem-se que os autovalores de C são a transformada discreta de Fourier da sua primeira coluna. Da decomposição de C tem-se:

$$y = Cx = F_N^* \text{diag}(F_N c) F_N x = F_N^* ((F_N c) .* (F_N x)), \quad (5.2)$$

onde $(.*)$ representa a multiplicação ponto a ponto. Portanto, para efetuar a multiplicação Cx , calcula-se $F_N x$ e $F_N c$, usando a *FFT*. Além disso, o uso da *FFT* é ainda melhor se a dimensão do vetor for uma potência de 2, e neste caso possui $O(N \log_2 N)$ multiplicações complexas. Observe ainda que, se for necessário calcular este produto para vários vetores, é preciso calcular $F_N c$ somente uma vez, o que reduz o número de *FFT*'s a serem calculadas.

Uma matriz de Toeplitz \mathcal{T} , $M \times N$, fica completamente definida pela sua primeira linha e primeira coluna. Suponha que \mathcal{T} é definida como em (5.1), então \mathcal{T} é descrita por:

$$t_c = \begin{bmatrix} h_M \\ h_{M-1} \\ \vdots \\ h_1 \end{bmatrix} \in \mathbb{C}^M, \quad t_l = \begin{bmatrix} h_{M+N-1} \\ h_{M+N-2} \\ \vdots \\ h_{M+1} \end{bmatrix} \in \mathbb{C}^{N-1}. \quad (5.3)$$

Com estes vetores, pode-se formar então a primeira coluna da matriz circulante:

$$c = \begin{bmatrix} t_c \\ 0 \\ t_l \end{bmatrix} \in \mathbb{C}^n, \quad (5.4)$$

onde 0 representa um vetor de zeros, cuja dimensão faz com que a dimensão de c seja a potência de 2 maior e mais próxima de $M + N - 1$. O exemplo a seguir ilustra este fato:

Exemplo 5.2. *Seja \mathcal{T} uma matriz de Toeplitz, 3×3*

$$\begin{bmatrix} h_3 & h_4 & h_5 \\ h_2 & h_3 & h_4 \\ h_1 & h_2 & h_3 \end{bmatrix}$$

a qual é descrita por:

$$t_c = \begin{bmatrix} h_3 \\ h_2 \\ h_1 \end{bmatrix}, \quad t_l = \begin{bmatrix} h_5 \\ h_4 \end{bmatrix},$$

então o vetor c é dado por:

$$c = (h_3 \ h_2 \ h_1 \ 0 \ 0 \ 0 \ h_5 \ h_4)^T.$$

Ou seja,

$$C = \begin{bmatrix} h_3 & h_4 & h_5 & 0 & 0 & 0 & h_1 & h_2 \\ h_2 & h_3 & h_4 & h_5 & 0 & 0 & 0 & h_1 \\ h_1 & h_2 & h_3 & h_4 & h_5 & 0 & 0 & 0 \\ 0 & h_1 & h_2 & h_3 & h_4 & h_5 & 0 & 0 \\ 0 & 0 & h_1 & h_2 & h_3 & h_4 & h_5 & 0 \\ 0 & 0 & 0 & h_1 & h_2 & h_3 & h_4 & h_5 \\ h_5 & 0 & 0 & 0 & h_1 & h_2 & h_3 & h_4 \\ h_4 & h_5 & 0 & 0 & 0 & h_1 & h_2 & h_3 \end{bmatrix}$$

Do exemplo verifica-se que a matriz $\mathcal{T}_{3 \times 3}$, aparece na parte superior esquerda da matriz C . O mesmo ocorre para uma matriz de Toeplitz com dimensão $M \times N$, ou seja:

$$C = \begin{bmatrix} \mathcal{T} & * \\ * & * \end{bmatrix}.$$

Então, para calcular o produto $y = \mathcal{T}x$, deve-se calcular o produto $C\bar{x}$, usando o Teorema 5.1, veja (5.2), com C a matriz circulante cuja primeira coluna é o vetor c obtido de (5.3) e (5.4) e \bar{x} um vetor da forma

$$\bar{x} = \begin{pmatrix} x \\ 0 \end{pmatrix}$$

Tem-se então o seguinte algoritmo:

Algoritmo 5.1. *Multiplicação Rápida.*

Dados T como em (5.1), e c como em (5.4).

1. Calcule $g = F_n c$
2. Calcule $z = F_n x$
3. Calcule $w = g \cdot z$
4. Calcule $y = F_n^* w$

Observação 5.3. Observe que, introduzindo $J = [e_n \ e_{n-1} \ \cdots \ e_1]$ então

$$JC = \begin{bmatrix} * & * \\ H & * \end{bmatrix},$$

onde H é uma matriz de Hankel $M \times N$, descrita como em (2.6). E para obter o produto $y = Hx$, basta, no algoritmo 5.1, tomar as M primeiras componentes de y na ordem inversa, isto é: $\bar{y} = [y_M \ y_{M-1} \ \cdots \ y_1]^T$. Desta forma é possível aplicar a multiplicação rápida também para as matrizes com estrutura Hankel.

5.2 O Procedimento de Rayleigh-Ritz

Seja $A = H^*H$, $A \in C^{N \times N}$, onde H denota a matriz $H(l)$ definida em (2.6), e $Q \in C^{N \times d}$ uma matriz ortonormal, cujas d primeiras colunas são aproximações dos d autovetores associados aos d autovalores desejados. A idéia do Procedimento de Rayleigh-Ritz é aproximar os d autovalores de A , pelos autovalores de $B = Q^*AQ$, a qual possui dimensão $d \times d$. Os autovalores de B são chamados valores de Ritz. Esta aproximação é feita da seguinte maneira:

Algoritmo 5.2. *Procedimento de Rayleigh Ritz*

*Aplicado a $A = H^*H$. Dada $Q \in C^{N \times d}$, $Q^*Q = I$.*

Para $k=1, 2, \dots$ faça

- 1. $AQ = H^*(HQ)$*

- 2. $B = Q^*(AQ)$*

- 3. Calcule os autovetores e autovalores de B , $Bg_i = \tau_i g_i$ e os vetores de Ritz*

$$y_i = Qg_i, \quad i = 1, \dots, d$$

- 4. Calcule os resíduos $s_i = Ay_i - \tau_i y_i$, $i = 1, \dots, d$*

Observe que se $d = 1$, então no passo 2 tem-se que $B = Q^*AQ$ é exatamente o Quociente de Rayleigh Ritz, $\rho(q, A)$. Assim, para $d > 1$ tem-se uma generalização do Quociente de Rayleigh Ritz. Além disso, como B é hermitiana e positiva definida, o passo 3 pode ser calculado de maneira rápida e com baixo tempo computacional. Para aproveitar a estrutura da matriz H , e a multiplicação rápida, a multiplicação do passo 1 é dividida em duas etapas. Primeiramente é calculado HQ e a seguir $H^*(HQ)$. As novas aproximações dos d autovalores e autovetores desejados são dadas pelos valores de Ritz e pelos vetores de Ritz, respectivamente. Parlett [29, pág. 234] demonstra três maneiras nas quais estas aproximações são ótimas.

A primeira pode ser vista como um corolário do Teorema do Minimax (veja, por exemplo, [15]).

$$\lambda_j(A) = \min_{\mathcal{F}^j \subset C^n} \max_{f \in \mathcal{F}^j} \frac{f^* A f}{f^* f} \quad (f \neq 0), \quad j = 1, 2, \dots, n$$

onde \mathcal{F}^j é um subespaço de C^n , com dimensão j .

Seja $\mathcal{Q}^d = \text{span}(Q)$, e \mathcal{G}^j um subespaço de \mathcal{Q}^d , com dimensão j . Assim, uma definição natural da melhor aproximação para λ_j , no subespaço \mathcal{Q}^d é:

$$\sigma_j(A) = \min_{\mathcal{G}^j \subset \mathcal{Q}^d} \max_{g \in \mathcal{G}^j} \frac{g^* A g}{g^* g} \quad (g \neq 0), \quad j = 1, 2, \dots, d.$$

A segunda abordagem, define uma matriz de resíduos $R(C) = AQ - AC$ e mostra que $B = Q^*AQ$ minimiza o resíduo, isto é, $\|R(B)\| < \|R(C)\|$.

De um terceiro modo, Parlett verifica que os valores e vetores de Ritz são os autovalores e autovetores da projeção de A em \mathcal{Q}^d , ou seja, da matriz que gera em \mathcal{Q}^d , o subespaço mais próximo de $\text{span}(A)$.

O valor mínimo da norma da matriz de resíduos, $R(C)$, pode ainda ser visto como uma medida de quão perto \mathcal{Q}^d está de se tornar um subespaço invariante de A . Se \mathcal{Q}^d for um subespaço invariante de A , então o produto Aq_i , onde q_i é uma coluna de Q , será uma combinação linear das colunas de Q , isto é, $Aq_i = Qc$. Assim, existe uma matriz C , tal que $R(C) = 0$. Se \mathcal{Q}^d não é um subespaço invariante de A , então não existe nenhuma matriz C , tal que $\|R(C)\| = 0$, porém, $B = Q^*AQ$ continua minimizando $R(C)$.

Seja agora $X = (x_1, x_2, \dots, x_d)$ uma base ortonormal de \mathcal{Q}^d e Δ uma matriz diagonal, $\Delta = \text{diag}(\delta_1, \delta_2, \dots, \delta_d)$, então $\|AX - X\Delta\|$ é minimizada quando $x_i = y_i$ e $\delta_i = \tau_i$, $i = 1, \dots, d$. De fato: Seja $Y = (y_1, \dots, y_d)$, $\Phi = \text{diag}(\tau_1, \dots, \tau_d)$ e $G = (g_1, \dots, g_d)$

$$\begin{aligned} \|AY - Y\Phi\| &= \|AQG - QG\Phi\| = \|AQ - QG\Phi G^*\| = \|AQ - QBGG^*\| = \\ &= \|AQ - QB\| = \|R(B)\|. \end{aligned}$$

Mesmo quando $X \neq Y$ ou $\Delta \neq \Phi$, ainda é possível expressar X na base Q , $X = QD$, onde $D^*D = DD^* = I$, pois elas geram o mesmo subespaço. Obtém-se então

$$\|AX - X\Delta\| = \|AQD - QD\Delta\| = \|AQ - QD\Delta D^*\| > \|R(B)\|. \quad (5.5)$$

Assim, quando \mathcal{Q}^d é um subespaço invariante de A , (5.5), juntamente com o fato de $B = Q^*AQ$ satisfazer $R(B) = 0$, significam que, os valores e vetores de Ritz são exatamente os autovalores e autovetores de A .

Quando os autovalores de A são bem separados, uma vez calculados os resíduos \mathbf{s}_i , pode-se facilmente obter limitantes para os valores de Ritz. Estes limitantes podem ser obtidos a partir do teorema abaixo, cuja demonstração pode ser encontrada em [13].

Teorema 5.4. *Seja $Q \in C^{N \times d}$, $Q^*Q = I$, e seja (τ_i, y_i) , $i = 1, \dots, d$, os valores e vetores de Ritz de A , com resíduos $\mathbf{s}_i = Ay_i - \tau_i y_i$. Então, cada intervalo*

$$[\tau_i - \|\mathbf{s}_i\|, \tau_i + \|\mathbf{s}_i\|]$$

possui um autovalor de A .

Como em cada intervalo correspondente aos valores de Ritz, existe um autovalor de A , se todos estes intervalos forem disjuntos, teremos d autovalores de A . Porém, se alguns deles se interceptarem, podemos ter dois valores de Ritz convergindo para o mesmo autovalor de A . No teorema abaixo, tem-se um limitante adicional para valores de Ritz em intervalos que se interceptam, sua demonstração pode ser encontrada em [29, cap. 11].

Teorema 5.5. *Seja $Q \in C^{N \times d}$, $Q^*Q = I$, e seja (τ_i, y_i) , $i = 1, \dots, d$, os valores e vetores de Ritz de A , com resíduos $s_i = Ay_i - \tau_i y_i$. Então, existem d autovalores de A , λ'_i , $i = 1, \dots, d$, tais que*

$$|\tau_i - \lambda'_i| \leq \|S\|$$

onde, $S = (s_1, \dots, s_d) = AY - Y\Phi$.

A determinação de limitantes para os vetores de Ritz é um pouco mais complicada, pois, autovetores associados a autovalores com multiplicidade algébrica $\neq 1$, não são unicamente determinados. Qualquer combinação de autovetores correspondentes ao mesmo autovalor, é um autovetor. Portanto, não é possível determinar uma estimativa de erro para estes autovetores. Porém, este fato é irrelevante, visto que o que interessa é o subespaço gerado por estes autovetores. Da mesma forma, vetores de Ritz correspondentes a autovalores muito próximos, tendem a ser sensíveis e darem estimativas ruins dos correspondentes autovetores, porém, o subespaço por eles gerado pode ser uma boa aproximação do subespaço associado a estes autovalores muito próximos.

5.3 Iteração em Subespaços

A primeira versão da Iteração em Subespaços surgiu em 1957, e foi introduzida por Bauer, com o nome de *Treppeniteration* (iteração escada). A Iteração em Subespaços pode ser vista como uma generalização do método das Potências, só que, ao invés de iterar um vetor, a iteração em subespaços itera d vetores simultaneamente, e obtém aproximações para os d autovetores de A , associados aos d maiores autovalores. Infelizmente, esta generalização apresenta um problema. Suponha uma matriz inicial $Q^{(0)} = (q_1^{(0)}, q_2^{(0)}, \dots, q_d^{(0)})$, $Q^{(0)H}Q = I$ e multiplique A iterativamente. Então, após j iterações obtém-se

$$Q^j = A^j Q^{(0)} = (A^j q_1^{(0)}, A^j q_2^{(0)}, \dots, A^j q_d^{(0)}).$$

Pela teoria do método das Potências, [43, cap. 4], pode-se deduzir que todas as colunas de Q^j irão convergir para o mesmo autovetor de A , o qual está associado ao maior autovalor. Este problema surge pois, as colunas de Q^j tornam-se linearmente dependentes. Para evitar isto, basta manter a ortogonalidade entre as

colunas de Q^j , através de um passo de ortogonalização em cada iteração. A seguir, é apresentado o algoritmo da Iteração em Subespaços.

Algoritmo 5.3. *Iteração em Subespaços*

Seja $A = H^*H$. Dada $Q^{(0)} \in C^{N \times d}$, $(Q^{(0)})^*Q^{(0)} = I$.

Para $j = 1, 2, \dots$ faça:

1. $Z^{(j)} = AQ^{(j-1)} = H^*(HQ^{(j-1)})$
2. Ortonormalize $Z^{(j)} = X^{(j)}R^{(j)}$, onde $R^{(j)}$ é uma matriz triangular superior e $X^{(j)}$ é unitária.
3. Faça $Q^{(j)} = X^{(j)}$
4. Teste a convergência de cada coluna

Assim como no método das Potências, a convergência da Iteração em Subespaços depende do *gap* entre os autovalores, isto é, da distância entre o último autovalor desejado, e o primeiro não desejado. Assim, se este *gap* for pequeno, para obter uma convergência mais rápida, basta aumentar o número de autovalores desejados, de forma que este *gap* fique maior. O passo 1 é feito em duas etapas, para aproveitar a multiplicação rápida.

5.4 Iteração em Subespaços com Aceleração de Ritz

Uma maneira de acelerar a convergência da iteração em subespaços, é usar o Procedimento de Rayleigh Ritz em cada iteração. Após o passo 2 do algoritmo 5.3, os vetores de Ritz são calculados a partir da matriz X^j , e as colunas da nova Q^j serão estes vetores. Esta idéia foi criada por Rutishauser, na década de 60, e é chamada Iteração em Subespaços com Aceleração de Ritz (SIR), [38]. Como as aproximações ótimas dos autovetores são calculadas em cada iteração, as componentes nas direções destes vetores são ampliadas, o que aumenta a razão de convergência.

Algoritmo 5.4. *Iteração em Subespaços com Aceleração de Ritz.*

Seja $A = H^*H$. Dada uma matriz inicial $Q^{(0)} \in C^{N \times d}$.

Para $j = 1, 2, \dots$ faça:

1. $Z^{(j)} = HQ^{(j-1)}$

2. Calcule a *svd* reduzida de $Z^{(j)}$: $[X_1, \Sigma_1, Y_1^*] = \text{svd}(Z^{(j)})$, onde $X_1 \in C^{M \times d}$, $\Sigma_1 \in R^{d \times d}$, $Y_1 \in C^{d \times d}$
3. $W^{(j)} = H^* X_1$
4. Calcule a *svd* reduzida de $W^{(j)}$: $[X_2, \Sigma_2, Y_2^*] = \text{svd}(W^{(j)})$, onde $X_2 \in C^{M \times d}$, $\Sigma_2 \in R^{d \times d}$, $Y_2 \in C^{d \times d}$
5. $Q^{(j)} = X_2$

Observe que o algoritmo 5.4, não deixa explícito passos da Iteração em Subespaços, nem passos da Aceleração de Ritz. Contudo, se for considerado que a *SVD* é usada somente para calcular uma base ortogonal de uma matriz, então os passos 1,2 e 3 equivalem ao passo 2 do algoritmo 5.3. O passo 4 calcula uma base ortogonal para H^*HQ , e esta base é usada como a nova Q . Desta forma, o algoritmo 5.4 é equivalente ao algoritmo 5.3, porém, com um passo extra de ortogonalização, o passo 2.

Para ficarem explícitos os passos da Aceleração de Ritz, é necessário levar em conta os efeitos dos cálculos da *SVD*. Seja $B := Q^*H^*HQ = Y_1 \Sigma_1^2 Y_1^*$ a correspondente decomposição autovalor para a *SVD* de HQ , calculada no passo 2, assim o quadrado dos elementos de Σ_1 são os valores de Ritz de $A = H^*H$, e as colunas de QY_1 seus vetores de Ritz. Como os vetores de Ritz são boas aproximações dos autovalores de A , eles também são boas aproximações dos vetores singulares à direita de H , os u_i 's. Pela relação $u_i = \tau_i^{-1} H v_i$, é fácil verificar que as colunas de X_1 são aproximações dos vetores singulares à esquerda de H . Analogamente, verifica-se que as colunas de X_2 são aproximações dos vetores singulares à direita de H . Ou seja, o algoritmo acima apresenta dois passos da Aceleração de Ritz.

5.5 Método de Lanczos

O método de Lanczos foi criado em 1950 por C. Lanczos, era utilizado para encontrar uma matriz tridiagonal T , unitariamente semelhante a uma matriz simétrica A , $T = Q^*AQ$, onde $Q = (q_1, q_2, \dots, q_n)$ e $Q^*Q = I$. Em aritmética exata, isto pode ser feito em n passos, porém, em aritmética finita aparecem problemas numéricos, decorrentes da perda de ortogonalidade das colunas de Q . Por esta razão, durante alguns anos o método de Lanczos foi substituído por métodos mais estáveis, como os métodos de Givens [14] e Householder [18].

O interesse pelo método de Lanczos só retornou em 1971, quando Paige [27] fez uma análise de erro detalhada do método, mostrando sua eficiência no cálculo de somente uma parte do espectro de uma matriz simétrica. A partir daí, devido à sua simplicidade, pois somente utiliza multiplicações do tipo matriz-vetor, e também

devido à qualidade das suas aproximações, o método de Lanczos tem sido fonte de diversas pesquisas, especialmente no sentido de evitar a perda de ortogonalidade das colunas de Q .

Inicialmente, o método é descrito na sua forma básica, a seguir são apresentadas outras versões, as quais evitam a perda de ortogonalidade de diferentes maneiras.

5.5.1 Método de Lanczos Básico

O método de Lanczos está baseado no fato que de é possível, em um número finito de passos, transformar uma matriz hermitiana A , em uma matriz tridiagonal T , por meio de transformações unitárias, [15].

$$T = Q^* A Q. \quad (5.6)$$

Ainda mais, se $Q^* A Q = T$, onde T é uma matriz tridiagonal real, com os elementos das diagonais secundárias todos positivos, então T e $Q = (q_1, q_2, \dots, q_n)$ são unicamente determinados por A e q_1 . A demonstração deste fato pode ser encontrada em [29, cap. 7]. A hipótese de os elementos das diagonais secundárias serem positivos não é uma restrição essencial.

Equacionando as j primeiras colunas de (5.6) obtém-se

$$A Q_j = Q_j T_j + r_j e_j^T, \quad (5.7)$$

onde $r_j = \beta_j q_{j+1}$ e $Q_j = (q_1, \dots, q_j)$, $Q_j^* Q_j = I$, T_j é o menor $j \times j$ de T

$$T_j = \begin{pmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{j-1} \\ & & & \beta_{j-1} & \alpha_j \end{pmatrix},$$

com $\alpha_i = q_i^* A q_i$ e $\beta_i = \|r_i\|$. Os vetores q_i são chamados vetores de Lanczos.

Similarmente, a equação (5.7) pode ser rescrita como

$$r_j = \beta_j q_{j+1} = A q_j - \alpha_j q_j - \beta_{j-1} q_{j-1}. \quad (5.8)$$

A partir desta equação é deduzida a recursão de Lanczos, a qual, dado um vetor inicial q_1 , gera a cada iteração, um novo vetor de Lanczos, e uma nova linha e coluna são adicionadas à matriz T . Os autovalores da matriz T convergem para os autovalores de A .

Algoritmo 5.5. *Método de Lanczos Básico.*

Seja $A = H^*H$. Dados $r_0 \neq 0$, $\beta_0 = \|r_0\|$.

Para $j = 1, 2, \dots$ faça:

1. $q_j := r_{j-1}/\beta_{j-1}$

2. $u_j := Aq_j = H^*(Hq_j)$

3. $r_j := u_j - \beta_{j-1}q_{j-1}$

4. $\alpha_j := q_j^*r_j$

5. $r_j := r_j - \alpha_jq_j$

6. $\beta_j := \|r_j\|$

7. Se desejar, calcule os valores de Ritz τ_i e vetores de Ritz y_i . Se a quantidade desejada de valores e vetores de Ritz convergiu, então pare.

Quando é completado um ciclo, isto é, quando são completados os passos de 1 até 6, tem-se um passo de Lanczos. Este procedimento gera vetores de Lanczos até que $\beta_j = 0$, o que deve ocorrer para algum $j \leq n$. Em particular, múltiplos autovalores de A , forçam $\beta_j = 0$, para algum $j < n$ (veja [45]). Os vetores de Lanczos formam uma base ortogonal no subespaço gerado por $(q_1, Aq_1, \dots, A^j q_1)$ e os autovetores aproximados estão neste subespaço. Portanto, o método não é capaz de detectar nenhum autovetor ortogonal a q_1 , o que significa que somente é possível determinar um autovetor, do subespaço correspondente a um autovalor múltiplo.

Da equação (5.8) tem-se que r_j é ortogonal a q_{j-1} e q_j . Pode-se mostrar, [29], que r_j é ortogonal a todos os vetores de Lanczos (q_1, \dots, q_j) . Os vetores $\alpha_j q_j$ e $\beta_{j-1} q_{j-1}$ são as projeções ortogonais de Aq_j em q_j e q_{j-1} . Isto significa que o próximo vetor de Lanczos é obtido da ortogonalização de Aq_j em relação à q_j e q_{j-1} e os vetores q_i , $1 \leq i \leq j$, formam uma base ortogonal no subespaço de Krylov

$$\mathcal{K}_j(A, q_1) = \text{span}\{q_1, Aq_1, \dots, A^{j-1}q_1\}. \quad (5.9)$$

Usando a ortogonalidade dos vetores de Lanczos e a relação (5.7), verifica-se que $T_j = Q_j^* A Q_j$, isto é, T_j é a projeção de A em $\text{span}(Q_j)$. Assim, no passo 7 do algoritmo 5.5, os valores de Ritz são calculados como os autovalores de T_j

$$T_j g_i^{(j)} = \tau_i^{(j)} g_i^{(j)}, \quad i = 1, 2, \dots, j \quad (5.10)$$

e os vetores de Ritz são definidos por

$$y_i^{(j)} = Q_j g_i^{(j)}, \quad i = 1, 2, \dots, j. \quad (5.11)$$

Os teoremas 5.4 e 5.5 mostram que a norma $\|Ay_i^{(j)} - \tau_i^{(j)}y_i^{(j)}\|$ é uma boa estimativa da precisão dos valores de Ritz. A princípio, é possível calcular τ_i e y_i a partir de T_j em cada iteração do algoritmo de Lanczos, porém, felizmente, é possível calcular $\|Ay_i^{(j)} - \tau_i^{(j)}y_i^{(j)}\|$ sem calcular y_i . Usando (5.7), (5.10), (5.11) e o fato de $\|q_{j+1}\| = 1$ tem-se:

$$\begin{aligned} \|s_i^{(j)}\| &= \|Ay_i^{(j)} - \tau_i^{(j)}y_i^{(j)}\| = \|AQ_jg_i^{(j)} - \tau_i^{(j)}Q_jg_i^{(j)}\| \\ &= \|AQ_jg_i^{(j)} - Q_jT_jg_i^{(j)}\| = \|\beta_jq_{j+1}e_j^{(T)}g_i^{(j)}\| \\ &= \beta_j|e_j^{(T)}g_i^{(j)}| =: \beta_{ij}. \end{aligned} \quad (5.12)$$

Assim, a convergência dos valores de Ritz para os autovalores de A , pode ser verificada pelos primeiros elementos dos autovetores normalizados de T_j , evitando o cálculo dos vetores de Ritz. Este resultado explica porque alguns valores de Ritz podem ser precisos, mesmo quando β_j não é muito pequeno.

Como já foi citado anteriormente, em aritmética finita, muitas das propriedades do procedimento de Lanczos são perdidas. Nas primeiras iterações, os resultados não diferem muito do processo exato. Até que um novo vetor de Lanczos é calculado, e este numericamente deixa de ser ortogonal aos anteriores. Algumas iterações mais tarde, Q_j deixou de ter posto completo, isto é, suas colunas são linearmente dependentes. A partir daí, o método começa a gerar cópias de vetores de Ritz já calculados, pois os novos vetores de Lanczos possuem componentes nas direções dos vetores de Ritz que já convergiram. Paige ([27],[28]) fez uma análise detalhada do desempenho do método em aritmética finita, e concluiu que a ortogonalidade entre os q 's é perdida quando surge um β_{ij} muito pequeno. Na prática, o processo nunca terminará com $\beta_j = 0$, mas irá continuar calculando mais e mais cópias dos mesmos vetores de Ritz, como consequência dos erros de arredondamento. Em resumo, de um lado, a presença de um β_{ij} pequeno indica que foi encontrada uma boa aproximação de pelo menos um autovalor de A , e de outro lado, isto também indica que as dificuldades numéricas começaram.

Uma maneira de resolver este problema é fazer uma reortogonalização completa, isto é, manter cada novo vetor de Lanczos ortogonal aos anteriores. Só que neste caso, além do trabalho extra para a reortogonalização, é necessário manter todos os vetores de Lanczos armazenados. Esta idéia é chamada de Lanczos com reortogonalização completa. E no outro extremo, tem-se o método de Lanczos sem reortogonalização, porém, com a exigência de uma análise mais detalhada das aproximações obtidas. Entre estes dois estão os métodos que fazem uma reortogonalização parcial, ou reortogonalização seletiva, isto é, reortogonalizam somente quando há necessidade. Nas próximas subseções vamos dar mais detalhes sobre alguns destes métodos.

5.5.2 Método de Lanczos com Reortogonalização Completa

A maneira mais direta de resolver o problema da perda de ortogonalidade, proposta pelo próprio Lanczos, é fazer a reortogonalização de cada novo vetor de Lanczos em relação aos anteriores, obtendo assim, o seguinte procedimento

Algoritmo 5.6. *Método de Lanczos com Reortogonalização Completa (LANCR).*
Seja $A = H^*H$. Dados $r_0 \neq 0$, $\beta_0 = \|r_0\|$.

Para $j = 1, 2, \dots$ faça:

1. $q_j := r_{j-1}/\beta_{j-1}$

2. $u_j := Aq_j = H^*(Hq_j)$

3. $r_j := u_j - \beta_{j-1}q_{j-1}$

4. $\alpha_j := q_j^*r_j$

5. $r_j := r_j - \alpha_jq_j$

5.5 $r_j := r_j - q_\nu(q_\nu^*r_j)$, $\nu = j, j-1, \dots, 1$

6. $\beta_j := \|r_j\|$

7. Se desejar, calcule os valores de Ritz τ_i e vetores de Ritz y_i . Se a quantidade desejada de valores e vetores de Ritz convergiu, então pare.

A desvantagem da reortogonalização completa é ter que manter armazenado todos os vetores de Lanczos, uma vez que r_j é explicitamente ortogonalizado em relação à q_j e q_{j-1} . Por outro lado, este algoritmo tem a vantagem de não calcular nenhum vetor de Ritz redundante, ou seja, não calcula mais de uma vez o mesmo vetor de Ritz. E como vetores de Ritz redundantes requerem um número maior de iterações para convergir, o algoritmo 5.6 requer um número mínimo de iterações para convergir.

Quando são desejados somente alguns dos maiores autovalores, as desvantagens deste algoritmo não são tão sérias. Normalmente os valores de Ritz correspondentes a estes autovalores convergem primeiro, e não são necessárias muitas iterações. Isto depende naturalmente do espectro em questão.

Como cita Parlett, [29, pág. 303], a reortogonalização não pode por si só garantir a ortogonalidade numérica dos vetores de Lanczos. Quando o corrente vetor de Lanczos tem uma componente considerável na direção de um dos outros vetores, então pode haver algum tipo de cancelamento no passo 5.5. Segundo Parlett, um passo extra de reortogonalização é suficiente para tornar estes dois vetores numericamente ortogonais.

Na próxima subseção, é visto que a reortogonalização completa é mais do que é necessário, e para diminuir a quantidade de trabalho, a reortogonalização será feita somente quando for preciso.

5.5.3 Lanczos com Semiortogonalização

As desvantagens da reortogonalização completa com relação ao trabalho extra e armazenamento, podem diminuir a eficiência quando os vetores de Lanczos possuem muitas componentes, ou quando o problema requer um grande número de iterações. Isto levou muitas pessoas a estudarem uma forma de transformar o método de Lanczos Básico, em um método mais estável que o LANCR, veja Paige [27, 28].

Uma idéia é reortogonalizar somente quando for necessário, e somente em relação àqueles autovetores que perderam a ortogonalidade. O primeiro algoritmo deste tipo, foi feito em 1979 por Parlett e Scott [30]. Eles publicaram um trabalho com um método chamado: Lanczos com Ortogonalização Seletiva (LANSO), no qual a ortogonalidade é mantida através da ortogonalização dos novos vetores de Lanczos em relação aos vetores de Ritz que já convergiram. Em 1984, Simon [32] escreveu um artigo, no qual mostrou que, mantendo os vetores num nível de ortogonalidade de \sqrt{u} , isto é, $|q_i^* q_{j+1}| \leq \sqrt{u}$, $1 \leq i < j$, é suficiente para obter boas aproximações dos autovalores e autovetores, onde u denota a precisão da máquina. Ele usou a noção de semiortogonalidade, quando os vetores de Lanczos satisfazem este nível de ortogonalidade. Esta análise levou a um novo algoritmo chamado Lanczos com Reortogonalização Parcial (LANPR).

5.5.4 Método de Lanczos com Recomeços Implícitos

Em problemas onde não existe um gap distinto entre os autovalores desejados e os não desejados, normalmente são necessários muitos passos de Lanczos para convergir. Assim a reortogonalização completa, apresenta uma desvantagem, pois ela exige o armazenamento de todos os vetores de Lanczos. Uma forma de contornar este problema é recomeçar as iterações após um número finito de passos de Lanczos, mantendo assim, a quantidade de vetores a ser armazenada fixa, e a reortogonalização feita somente quando for necessário. Calvetti, Reichel e Sorensen desenvolveram um método de Lanczos com recomeços implícitos (IRL), [10], o qual é uma adaptação do método de Arnoldi com recomeços implícitos, [25], para o caso Hermitiano. Este método força o vetor inicial a estar num subespaço invariante da matriz, por meio de repetidas filtragens do vetor inicial, usando um filtro polinomial, e recomeçando as iterações implicitamente.

Este estudo está baseado no artigo de Calvetti, Reichel e Sorensen, [10], porém, está restrito ao cálculo dos d maiores autovalores. Inicialmente, é discutido

como o vetor inicial deve ser escolhido, para que o método convirja em d iterações, e produza um subespaço invariante correspondente aos d autovetores desejados. A seguir, é descrito o algoritmo (IRL), como uma mistura do método de Lanczos, com o Algoritmo QR com deslocamento implícito.

O Vetor Inicial

Quando se deseja aproximar d autovalores e autovetores, usando o método de Lanczos, a situação ideal é encontrar um subespaço invariante após exatamente d iterações, isto é, quando r_d da equação (5.7), se anula. Como já foi visto no início da seção 5.5.1, a iteração de Lanczos fica completamente definida pelo seu vetor inicial, assim, q_1 determina se a iteração irá convergir em d iterações. Surge então uma pergunta: Quais condições o vetor inicial deve satisfazer, para que se obtenha um subespaço invariante após d iterações? O teorema abaixo afirma que se q_1 for uma combinação linear de d autovetores, então a recursão terminará em d passos, com $r_d = 0$. Além disto, a equação (5.8), força os vetores de Lanczos a estarem no subespaço gerado por estes d autovetores. Assim, uma boa escolha de vetor inicial, é como tal combinação.

Teorema 5.6. *Seja $A_{N \times N}$ uma matriz simétrica, e seja $AQ_d - Q_dT_d = r_d e_d^T$, com T simétrica não reduzida, o passo d da fatoração de Lanczos de A . Então $r_d = 0$ se e somente se $q_1 = V_d x$, onde*

$$AV_d = V_d \Lambda_d \quad (5.13)$$

com $V_d^* V_d = I_d$ e Λ_d uma matriz diagonal de ordem d .

Demonstração:

(\Rightarrow) Suponha que $r_d = 0$, então, $AQ_d = Q_dT_d$. Seja $T_d G_d = G_d \Lambda_d$ a decomposição autovalor-autovetor de T_d , $G_d^* G_d = I_d$, chamando $V_d = Q_d G_d$, então $AQ_d G_d = Q_d G_d \Lambda_d$, isto é, $AV_d = V_d \Lambda_d$, além disso, $\text{posto}(V_d) = d$, e $q_1 = Q_d e_1 = Q_d G_d G_d^* e_1 = V_d x$, com $x = G_d^* e_1$.

(\Leftarrow) Suponha agora que $AV_d = V_d \Lambda_d$, $\text{posto}(V_d) = d$, e $q_1 = V_d x$. Então, para qualquer inteiro $m > 0$, $A^m V_d = V_d \Lambda_d^m$ e assim, para todo m ,

$$A^m q_1 = A^m V_d x = V_d \Lambda_d^m x \in \mathcal{R}(V_d).$$

Conseqüentemente, $\dim(\mathcal{K}^{d+1}(A, q_1)) \leq \text{posto}(V_d) = d$. Como T_d é não reduzida, usando uma pequena modificação do Teorema 7.4.3 de Golub e Van Loan [15], tem-se que, $\dim(\mathcal{K}^{d+1}(A, q_1)) = d$, e então $r_d = 0$. \square

O Algoritmo IRL

Sejam d e p , números pré-especificados fixos, onde, d é o número de autovalores desejados, e p o número de passos extras que serão realizados. Após $d + p$ passos de Lanczos, tem-se uma matriz $(d + p) \times (d + p)$, hermitiana e tridiagonal T_{d+p} , e uma matriz $Q_{d+p} \in C^{N \times (d+p)}$ com colunas ortonormais, tais que

$$AQ_{d+p} = Q_{d+p}T_{d+p} + r_{d+p}e_{d+p}^T \quad (5.14)$$

onde $r_{d+p} = \beta_{d+p}q_{d+p+1}$. Por simplicidade, vamos chamar $Q = Q_{d+p}$, $T = T_{d+p}$ e $r = r_{d+p}$, e aplicar o algoritmo QR com shift, a esta fatoração truncada de A . Seja μ_1 um shift, calcule a decomposição QR de $T - \mu_1 I$, $T - \mu_1 I = X_1 R_1$, onde $X_1, R_1 \in C^{(d+p) \times (d+p)}$, $X_1 X_1^* = I$ e R_1 é uma matriz triangular superior.

Então somando e subtraindo $\mu_1 Q$ na equação (5.14), e usando o fato de $T - \mu_1 I = X_1 R_1$, obtém-se:

$$\begin{aligned} AQ - \mu_1 Q - QT + \mu_1 Q &= r e_{d+p}^T \\ (A - \mu_1 I)Q - Q(T - \mu_1 I) &= r e_{d+p}^T \\ (A - \mu_1 I)Q - QX_1 R_1 &= r e_{d+p}^T, \end{aligned} \quad (5.15)$$

multiplicando a equação (5.15) por X_1 e rearranjando os termos

$$\begin{aligned} (A - \mu_1 I)(QX_1) - (QX_1)(R_1 X_1) &= r e_{d+p}^T X_1 \\ A(QX_1) - (QX_1)(R_1 X_1 + \mu_1 I) &= r e_{d+p}^T X_1. \end{aligned}$$

Seja $Q^+ = QX_1$ e $T^+ = R_1 X_1 + \mu_1 I = X_1^* T X_1$. Então T^+ é simétrica, e como T é tridiagonal, X_1 é Hessenberg superior, assim, T^+ também é tridiagonal. Explicitando a primeira coluna de ambos os lados da equação (5.15) tem-se

$$(A - \mu_1 I)q_1 = QX_1 R_1 e_1 = Q^+ e_1 \rho_{11} = q_1^+ \rho_{11}$$

onde $q_1 = Qe_1$, $\rho_{11} = e_1^T R_1 e_1$ e $q_1^+ = Q^+ e_1$. Tem-se assim, uma relação entre q_1 e q_1^+ .

Aplicando p shifts e usando a equação (5.15), obtém-se:

$$\begin{aligned} AQ^+ &= Q^+ T^+ + r e_{d+p}^T \hat{X} \\ &= (Q^+ \quad q_{d+p+1}) \begin{pmatrix} T^+ \\ \beta_{d+p} e_{d+p}^T \hat{X} \end{pmatrix}, \end{aligned} \quad (5.16)$$

onde $q_{d+p+1} = r/\beta_{d+p}$, $Q^+ = Q\hat{X}$, $T^+ = \hat{X}^*T\hat{X}$ e $\hat{X} = X_1X_2\dots X_p$, X_i é a matriz ortogonal associada ao shift i .

Observe que, como T^+ permanece tridiagonal a cada shift, as matrizes X_i , são Hessenberg superiores, e os primeiros $d - 1$ elementos da linha $d + p$ em \hat{X} serão iguais a zero, o que significa que:

$$\beta_{d+p}e_{d+p}^T\hat{X} = (0, 0, \dots, \tilde{\beta}_{d+p}, b^*), \quad \text{onde } \tilde{\beta}_{d+p} = e_d^T\beta_{d+p}.$$

Particionando as matrizes,

$$Q^+ = (Q_d^+, \hat{Q}_p^+) \quad T^+ = \begin{pmatrix} T_d^+ & \hat{\beta}_d e_d e_1^T \\ \hat{\beta}_d e_1 e_d & \hat{T}_p \end{pmatrix},$$

com $\hat{\beta} = e_{d+1}^T T^+ e_d$ e substituindo em (5.16), tem-se:

$$A(Q_d^+, \hat{Q}_p^+) = (Q_d^+, \hat{Q}_p^+, q_{d+p+1}) \begin{pmatrix} T_d^+ & \hat{\beta}_d e_d e_1^T \\ \hat{\beta}_d e_1 e_d & \hat{T}_p \\ \tilde{\beta}_{d+p} e_d^T & b^* \end{pmatrix}. \quad (5.17)$$

Igualando as primeiras d colunas da equação 5.17, obtém-se:

$$AQ_d^+ = Q_d^+ T_d^+ + e_d^+ e_d^T = (Q_d^+, q_{d+1}^+) \begin{pmatrix} T_d^+ \\ \beta_d^+ e_d^T \end{pmatrix},$$

onde $q_{d+1}^+ = \frac{1}{\beta_d^+} r_d^+$, $r_d^+ = (\hat{Q}_p e_1 \hat{\beta}_d + q_{d+p+1} \tilde{\beta}_{d+p})$ e $\beta_d^+ = \|r_d^+\|$. Como $(Q_d^+)^* Q_d^+ = I_d$ e $(Q_d^+)^* q_{d+1}^+ = 0$, então esta é outra decomposição de Lanczos de A . O novo vetor inicial pode ser escrito como $q_1^+ = \Psi_p(A)q_1$, onde Ψ_p é um polinômio mônico de grau p , cujas raízes são os p shifts, $\mu_1, \mu_2, \dots, \mu_p$. Recomeçando a partir desta fatoração, podemos calcular p passos extras de Lanczos, o que nos remete à equação (5.14). Então, aplicando alternadamente p passos de Lanczos e p shifts, obtém-se o método de Lanczos com Recomeços Implícitos. A cada iteração obtém-se aproximações dos autovetores, a partir de um subespaço de *Krylov*, de dimensão $d + p$, com um custo de p multiplicações do tipo matriz-vetor, ao invés de $d + p$.

A escolha dos shifts

Os p passos QR aplicados desempenham um papel de filtragem do vetor inicial, fazendo com que este esteja num subespaço invariante apropriado, $\text{span}(V_d)$. O lema abaixo, extraído de [10], sugere que para eliminar o conjunto de autovalores não desejados de T_{d+p} , deve-se escolher os p shifts μ_j como sendo os autovalores que estão na parte não desejada do espectro. Segundo Calvetti [10], estes shifts são chamados de *shifts exatos*.

Lema 5.7. *Seja $\lambda(T_{d+p}) = \{\tau_1, \dots, \tau_d\} \cup \{\mu_1, \dots, \mu_p\}$ uma partição disjunta do espectro de T_{d+p} e seja*

$$T_{d+p}^+ = \hat{X}^* T_{d+p} \hat{X} \quad (5.18)$$

onde $\hat{X} = X_1 X_2 \dots X_p$, e X_j é implicitamente determinado pelo shift μ_j . Seja β_j o elemento da posição $(j+1)$ da subdiagonal de T_{d+p}^+ . Se $\beta_j \neq 0$ para $1 \leq j < d$, então $\beta_d = 0$ e

$$T_{d+p}^+ = \begin{pmatrix} T_d^+ & 0 \\ 0 & \hat{T}_p \end{pmatrix},$$

onde $\lambda(T_d^+) = \{\tau_1, \dots, \tau_d\}$ e $\lambda(\hat{T}_p) = \{\mu_1, \dots, \mu_p\}$. Além disso,

$$q_1^+ = Q_{d+p} \hat{X} e_1 = \sum_{j=1}^d \eta_j y_j,$$

onde $\eta_j \in C$ e y_j é o vetor de Ritz associado ao valor de Ritz τ_j , isto é, $y_i = Q_{d+p} g_j$ com $T_{d+p} g_j = g_j \tau_j$ para $1 \leq j \leq d$.

Demonstração:

Para simplificar a demonstração, vamos chamar $T = T_{d+p}$ e $Q = Q_{d+p}$. Seja $x_1^{(i)} = X_1 X_2 \dots X_i e_1$ e $\rho_{11}^{(i)} = e_1^T$. Quando o primeiro shift é aplicado, obtém-se $T - \mu_1 I = X_1 R_1$. Explicitando a primeira coluna tem-se: $\frac{1}{\rho_{11}^{(1)}}(T - \mu_1 I) = x_1^{(1)}$. Similarmente, aplicando p shifts chega-se a:

$$x_1 = \hat{X} e_1 = \frac{1}{\rho_{11}^{(1)} \dots \rho_{11}^{(p)}} (T - \mu_p I) \dots (T - \mu_1 I) e_1 =: \psi(T) e_1,$$

onde $\psi(\lambda) = \frac{1}{\nu} \prod_{i=1}^p (\lambda - \mu_i)$ e $\nu = \rho_{11}^{(1)} \dots \rho_{11}^{(p)}$.

Seja $T = G \Lambda G^*$ a decomposição autovalor - autovetor de T , onde $G^* G = I$ e $\Lambda = \text{diag}(\tau_1, \dots, \tau_d, \mu_1, \dots, \mu_p)$. Assim,

$$(T - \mu_1 I) = G(\Lambda - \mu_1 I)G^*, \dots, (T - \mu_p I) = G(\Lambda - \mu_p I)G^*,$$

logo

$$\begin{aligned} (T - \mu_1 I)(T - \mu_2 I) \dots (T - \mu_p I) &= G(\Lambda - \mu_1 I)(\Lambda - \mu_2 I) \dots (\Lambda - \mu_p I)G^* = \\ &= G \begin{pmatrix} \tilde{\tau}_1 & & & & \\ & \ddots & & & \\ & & \tilde{\tau}_d & & \\ & & & 0 & \\ & & & & 0 \end{pmatrix} G^* \\ &= G \tilde{\Lambda} G^*. \end{aligned}$$

Suponha que $e_1 = \sum_{j=1}^{d+p} \varphi_j g_j = G\bar{\varphi}$, onde $\bar{\varphi} = [\varphi_1, \dots, \varphi_{d+p}]^T$. Então $x_1 = \psi(T)e_1 = G\tilde{\Lambda}G^*G\bar{\varphi} = G\tilde{\Lambda}\bar{\varphi} = G_d\bar{\varphi}_d$, com $G_d = [g_1, \dots, g_d]$, g_i , $1 \leq i \leq d$ são autovetores associados à parte desejada do espectro, e $\bar{\varphi}_d = [\varphi_1, \dots, \varphi_d]^T$. Isto mostra que as componentes de e_1 na direção dos autovetores associados a μ_i , $1 \leq i \leq p$ são anuladas, então x_1 pode ser escrito como $x_1 = \sum_{j=1}^d \eta_j g_j$. Equacionando as primeiras j colunas da relação (5.18), tem-se uma decomposição de Lanczos para a matriz T , com vetor inicial x_1 . Usando uma pequena modificação do Teorema 7.4.3 de Golub e Van Loan [15], segue que $\beta_d = 0$, portanto $T_{d+p}^+ = \begin{pmatrix} T_d^+ & 0 \\ 0 & \hat{T}_p \end{pmatrix}$. Além disso,

$$q_1^+ = Q\hat{X}e_1 = Qx_1 = \sum_{j=1}^d \eta_j Qg_j = \sum_{j=1}^d \eta_j y_j.$$

□

O lema 5.7 mostra que a escolha dos shifts exatos é matematicamente equivalente a recomençar as iterações de Lanczos com um vetor inicial que seja uma combinação linear dos vetores de Ritz de A , correspondentes aos d autovalores desejados, isto é, como uma aproximação de uma combinação linear dos d autovetores. O método IRL pode ser usado para aproximar qualquer parte do espectro de A . Contudo, o lema 5.7 não apresenta nenhuma informação sobre a razão de convergência, e infelizmente sob certas condições, esta pode ser muito lenta. Por exemplo, o cálculo dos menores autovalores, requer muitas iterações até convergir, pois o método precisa trabalhar muito até vencer a influência dos maiores autovalores, os quais aparecem como shifts em muitas iterações. Em [10], encontra-se uma discussão sobre outros tipos de shifts.

A seguir, é apresentado um esboço do algoritmo IRL, para aproximar os d maiores autovalores. É assumido que

$$\tau_1 \geq \tau_2 \geq \dots \geq \tau_{d+p}$$

Algoritmo 5.7. *Método de Lanczos com Recomeços Implícitos e Shifts Exatos (IRL-ES).*

Seja $A = H^*H$. Dados $r_0 \neq 0$, $\beta_0 = \|r_0\|$.

1. Faça $d + p$ passos usando o algoritmo 5.5 (reortogonalizando quando for necessário)

Repita

2. Calcule a decomposição autovalor-autovetor de T_{d+p} , e teste a convergência de $\beta_{(d+p)i}$

3. Se a quantidade desejada de pares de Ritz convergiu, calcule os vetores de Ritz e pare.
4. Aplique o algoritmo QR com shifts $\mu = \tau_j$, $j = d + 1, \dots, d + p$
- 5 Faça p passos de Lanczos, começando de (5.18), e volte para o passo 2 acima.

Como os pares de Ritz possuem diferentes razões de convergência, uma forma de aumentar a convergência do método seria remover os vetores de Ritz que já convergiram, armazenar os de interesse e, manter os vetores de Lanczos ortogonais a estes. Com isso, a dimensão do espaço de Krylov usado é reduzida, e como este espaço é ortogonal aos vetores de Ritz que já convergiram, as próximas iterações do método irão estimar a parte restante desejada do espectro.

O problema de remover aproximações de autovalores e autovetores que já convergiram, é muito estudado, e existem muitas técnicas de deflação. Lehoucq e Sorensen [26], desenvolveram regras de deflação para o método IRL, baseadas em técnicas do método QR. Basicamente tem-se dois tipos de deflação. Se o valor de Ritz pertence à parte desejada do espectro, é necessário manter o correspondente vetor de Ritz na próxima fatoração de Lanczos. Este tipo de deflação é chamado *Locking*. E se o valor de Ritz pertence à parte não desejada do espectro, é preciso removê-lo das próximas iterações, e este é chamado de *Purging*.

Convergência

Será feita agora uma análise da convergência do IRL-ES, considerando o caso em que se deseja aproximar apenas os d maiores autovalores. A discussão está baseada no trabalho de Sorensen [37], onde primeiramente é mostrado que à medida que ocorrem as iterações, os autovalores de T_d convergem para certos limites. Então é mostrado que estes limites são os autovalores de A e, finalmente, que são os d maiores.

Inicialmente são apresentados dois lemas, que serão usados posteriormente.

Lema 5.8. *Seja $M = \begin{pmatrix} T & \beta e_d \\ \beta e_d^T & \alpha \end{pmatrix}$, uma matriz simétrica tridiagonal. Então as raízes da equação*

$$\beta^2 e_d^T (T - \lambda I)^{-1} e_d = \alpha - \lambda \quad (5.19)$$

são autovalores de M .

Demonstração:

A demonstração segue do cálculo do polinômio característico de M . \square

Seja $T = G\Lambda G^*$, a decomposição autovalor-autovetor de T , como $(T - \lambda I)^{-1} = G(\Lambda - \lambda I)^{-1}G^*$, substituindo na equação (5.19) tem-se:

$$\beta^2 \sum_{j=1}^{\alpha} \frac{\eta_j^2}{(\tau_j - \lambda)} = \alpha - \lambda. \quad (5.20)$$

onde $(\eta_1, \eta_2, \dots, \eta_d) = e_d^T G$ e $\Lambda = \text{diag}(\tau_1, \dots, \tau_d)$. Da equação (5.20), pode-se concluir que nenhum autovalor de M é igual a τ_i , $1 \leq i \leq d$. Usando a propriedade do *interlacing*, veja Golub [15, pág. 396], tem-se:

$$\hat{\mu} < \tau_d < \tilde{\tau}_d < \tau_{d-1} < \dots < \tau_1 < \tilde{\tau}_1 \quad (5.21)$$

onde $\{\tilde{\tau}_1, \dots, \tilde{\tau}_d, \hat{\mu}\}$ são autovalores de M .

Observe também que, se os elementos da subdiagonal de T são todos não nulos, então pela discussão que segue o algoritmo 5.5, os τ_j são distintos, e nenhum dos η_j é zero.

O próximo lema afirma que se q_1 está em um subespaço de dimensão menor do que d , então algum β_j deve ser zero, e deve ocorrer a deflação.

Lema 5.9. *Suponha que $AQ = QT + re_k^T$ é uma decomposição de Lanczos de A , e seja β_j o j -ésimo elemento da subdiagonal de T . Se $q_1 = v\gamma + w\zeta$, com $\gamma^2 + \zeta^2 = 1$, $\|v\| = \|w\| = 1$, $v^*w = 0$ e $v = \sum_{j=1}^i \gamma_j \zeta_j$, $Av_j = \lambda_j v_j$, onde $\{\lambda_j\}$ é um conjunto arbitrário de i autovalores de A , então*

$$\prod_{j=1}^i \beta_j \leq \zeta \left\| \prod_{j=1}^i (A - \lambda_j I) \right\| \quad (5.22)$$

Demonstração:

A demonstração pode ser encontrada em [37].

Com estes dois lemas em mãos, é possível começar a mostrar que o IRL-ES converge. Serão escolhidos os p menores autovalores de T_{d+p} .

Seja l o número da iteração. Então $q_1^{(l)}$ representa o vetor inicial na iteração l , e após p passos de Lanczos tem-se:

$$AQ_{d+p}^{(l)} - Q_{d+p}^{(l)} T_{d+p}^{(l)} = r_{d+p}^{(l)} e_{d+p}^T.$$

Seja $T_{d+p}^{(l)} = \begin{pmatrix} T_d^{(l)} & \beta_d^{(l)} e_d e_1^T \\ \beta_d^{(l)} e_1 e_d^T & \hat{T}^{(l)} \end{pmatrix}$, com autovalores $\tau_{1,l+1} > \dots > \tau_{d,l+1} > \mu_{1,l+1} > \dots > \mu_{p,l+1}$.

Aplicando p shifts tem-se:

$$\hat{X}^{(l)H} T^{(l)} \hat{X}^{(l)} = \begin{pmatrix} T_d^{(l+1)} & 0 \\ 0 & \hat{T}_p \end{pmatrix}, \quad Q_d^{(l+1)} = (Q_{d+p}^{(l)} \hat{X}^{(l)}) \begin{pmatrix} I_d \\ 0 \end{pmatrix},$$

onde $\hat{X}^{(l)} = X_1^{(l)} \dots X_p^{(l)}$ são construídos aplicando os shifts $\mu_{1,l+1}, \dots, \mu_{p,l+1}$.

O próximo lema mostra que os autovalores τ_{jl} de $H_d^{(l)}$ convergem.

Lema 5.10. *Cada $\{\tau_{jl}, l = 1, 2, \dots\}$ é uma seqüência convergente para cada $j = 1, 2, \dots, d$.*

Demonstração:

Como $T_{d+p}^{(l)}$ é obtida através de p abordagens sucessivas de $T_d^{(l)}$, segue após p aplicações do lema 5.9 que

$$\tau_{jl} < \tau_{j,l+1} \quad \text{para } j = 1, \dots, d.$$

Como cada τ_{jl} é um quociente de Rayleigh com relação à A , segue que $\lambda_n \leq \tau_{jl} \leq \lambda_1$, para todo j, l (veja Strang [36, pág. 266]). Ou seja, a seqüência $\{\tau_{jl}\}$ é limitada, e como é crescente, então converge. \square

A seguir, é mostrado que o limite para o qual as seqüências $\{\tau_{jl}\}_{l=1}^{\infty}$ convergem, são autovalores de A .

Lema 5.11. *Seja $T_d^{(l)} = G^{(l)} \Lambda^{(l)} G^{(l)H}$ e $(\eta_1^l, \eta_2^l, \dots, \eta_d^l) = e_d^T G^{(l)}$. Assuma que τ_j são disjuntos, onde $\tau_{jl} \rightarrow \tau_j$. então*

$$\beta_d^l \eta_j^l \rightarrow 0, \quad \text{quando } l \rightarrow \infty, \quad j = 1, \dots, d$$

e em conseqüência

$$\|AQ^{(l)}g_j^{(l)} - Q^{(l)}g_j^{(l)}\tau_{jl}\| = |\beta^{(l)}\eta_j^l| \rightarrow 0,$$

onde $g_j^l = G^{(l)}e_j$ para $j = 1, \dots, d$.

Demonstração:

Considere a submatriz $(d+1) \times (d+1)$ principal de $T_{d+p}^{(l)}$

$$M = \begin{pmatrix} T_d^{(l)} & \beta_d^{(l)} e_d \\ \beta_d^{(l)} e_d^T & \alpha^{(l)} \end{pmatrix}.$$

Do lema 5.8 segue que os d maiores autovalores $\hat{\tau}_{jl}$ de $M^{(l)}$ satisfazem

$$\tau_{jl} < \hat{\tau}_{jl} < \tau_{j,l+1},$$

onde a última desigualdade surge de $p-1$ aplicações adicionais do lema 5.8.

Além disso, por manipulações algébricas da equação (5.20) obtém-se:

$$(\beta\eta_j)^2 = (\tau_j - \lambda) \left[\frac{(\alpha - \lambda) - \beta^2 \sum_{i=j+1}^d \frac{\eta_i^2}{(\tau_i - \lambda)}}{1 + \sum_{i=1}^{j-1} \frac{\eta_i^2 (\tau_j - \lambda)}{\eta_j^2 (\tau_i - \lambda)}} \right] \quad (5.23)$$

para qualquer raiz λ . Substituindo da matriz $T_d^{(l)}$ as quantidades apropriadas indexadas por l , e fazendo $\lambda = \hat{\tau}_{jl}$ tem-se:

$$(\beta_d^{(l)} \eta_j^{(l)})^2 < \left| (\tau_{jl} - \hat{\tau}_{jl}) \left| (\alpha^{(l)} - \hat{\tau}_{jl}) - \beta_d^{(l)^2} \sum_{i=j+1}^d \frac{\eta_j^{(l)^2}}{(\tau_{il} - \hat{\tau}_{jl})} \right| \right|.$$

Observe que, como $\hat{\tau}_{jl} > \tau_{jl} > \tau_{il}$ para $i = 1, 2, \dots, j-1$, então o denominador na equação 5.23 é maior do que 1.

A hipótese de que os limites τ_j são distintos implica que

$$\left| (\alpha^{(l)} - \hat{\tau}_{jl}) - \beta_d^{(l)^2} \sum_{i=j+1}^d \frac{\eta_j^{(l)^2}}{(\tau_{il} - \hat{\tau}_{jl})} \right|,$$

tem limites finitos para cada j . Assim, para l suficientemente grande, existe uma constante positiva K , tal que

$$(\beta_d^{(l)} \eta_j^{(l)})^2 < K |\tau_{jl} - \hat{\tau}_{jl}| < K |\tau_{jl} - \tau_{j,l+1}| \rightarrow 0, \text{ quando } l \rightarrow \infty,$$

pois cada seqüência $\{\tau_{jl}, l = 1, 2, \dots\}$ é uma seqüência crescente convergente, para cada $j = 1, 2, \dots, d$. \square

Uma vez mostrado que os τ_{jl} convergem para os autovalores de A , o próximo passo é mostrar que estes são os d maiores. Observe que o vetor inicial $q_1^{(l)}$ pode ser escrito como:

$$q_1^{(l)} = \frac{\Psi_l(A)q_1}{\|\Psi_l(A)q_1\|},$$

onde $\psi_i(\lambda) = \nu_i \prod_{k=1}^p (\lambda - \mu_{k,i})$, $\Psi_l(\lambda) = \prod_{i=1}^l \psi_i(\lambda)$ e q_1 é o vetor inicial original.

Teorema 5.12. *Suponha que o vetor inicial q_1 satisfaz $v_j^* q_1 = \gamma_j \neq 0$ para $j = 1, \dots, d$, onde v_j é o autovetor de A associado ao autovalor λ_j , com os autovalores de A listados em ordem decrescente. Seja $\beta_i^{(l)}$ o elemento da posição i da subdiagonal de $T_d^{(l)}$, e assuma que $\beta_i^{(l)} > \epsilon > 0$ para todo i, l . Então as seqüências*

$$\tau_{jl} \rightarrow \tau_j = \lambda_j \text{ quando } l \rightarrow \infty$$

Demonstração:

A hipótese de $\beta_i^{(l)} > \epsilon > 0$ assegura que a separação dos τ_{jl} é uniforme para todo l , de forma que os limites τ_j são distintos. Isto implica que cada τ_j é um autovalor de A . Além disso, essa hipótese implica uma limitação inferior uniforme de $|\eta_j^{(l)}|$ e então, $\beta_d^{(l)} \rightarrow 0$. Estas observações seguem do lema 5.8.

Seja $\hat{p}_d^{(l)}(\lambda) = \prod_{i=1}^d (\lambda - \tau_{il})$ o polinômio característico de $T_d^{(l)}$ e seja $\hat{p}_d(\lambda) = \prod_{i=1}^d (\lambda - \tau_i)$ o polinômio limite de $\hat{p}_d^{(l)}$. Então:

$$\|\hat{p}_d^{(l)}(A)q_1^{(l)}\| = \prod_{j=1}^d \beta_d^{(l)} \rightarrow 0$$

e então

$$\hat{p}_d^{(l)}(\lambda_j)\gamma_j^{(l)} = v_j^* \hat{p}_d^{(l)}(A)q_1^{(l)} \rightarrow 0$$

Assim,

$$\hat{p}_d(\lambda_j) = 0 \text{ ou } \gamma_j^{(l)} \rightarrow 0$$

para todo $j = 1, \dots, n$. Isto significa que $n - d$ dos coeficientes da expansão de $\gamma_j^{(l)}$, tendem para 0, quando $l \rightarrow \infty$. Além disso, como cada seqüência $\{\tau_{jl}, l = 1, 2, \dots\}$ é convergente para $j = 1, \dots, d$, tem-se que os d coeficientes da expansão, correspondentes aos autovalores τ_j , devem ser todos limitados inferiormente por 0, pela hipótese $\beta_i^{(l)} > \epsilon > 0$ para todo $j = 1, 2, \dots, d - 1$ e l .

Suponha agora que $\lambda_{j1} = \tau_1 < \lambda_1$. Então o coeficiente da expansão

$$\gamma_{j1}^{(l)} = v_{j1}^* q_1^{(l)} = v_{j1}^* \frac{\Psi_l(A)q_1}{\|\Psi_l(A)q_1\|} = \frac{\Psi_l(\tau_1)v_{j1}^*q_1}{\|\sum_{i=1}^n \gamma_i \Psi_l(\lambda_i)v_i\|} = \frac{\gamma_{j1}\Psi_l(\tau_1)}{\sqrt{\sum_{i=1}^n \gamma_i^2 \Psi_l^2(\lambda_i)}},$$

onde v_{j1} representa o autovetor de A associado ao autovalor τ_1 , e $q_1 = \sum_{i=1}^n \gamma_i v_i$ com $v_i^* v_j = 0$ quando $i \neq j$.

Assim,

$$(\gamma_{j1}^{(l)})^2 = \frac{(\gamma_{j1}\Psi_l(\tau_1)/\Psi_l(\lambda_1))^2}{\gamma_1^2 + \sum_{i=2}^n \gamma_i^2 \Psi_l^2(\lambda_i)/\Psi_l^2(\lambda_1)} \leq \left(\frac{\gamma_{j1}\Psi_l(\tau_1)}{\gamma_1\Psi_l(\lambda_1)}\right)^2,$$

onde os γ_j são os coeficientes da expansão de q_1 . Agora, as raízes μ_{il} dos polinômios filtros satisfazem, $\lambda_n \leq \mu_{il} < \tau_{1l} < \lambda_1$, de forma que:

$$\frac{\Psi_l(\tau_1)}{\Psi_l(\lambda_1)} = \prod_{k=1}^l \left(\prod_{i=1}^p \left(\frac{\tau_1 - \mu_{ik}}{\lambda_1 - \mu_{ik}} \right) \right) \leq \left(\frac{\tau_1 - \lambda_n}{\lambda_1 - \lambda_n} \right)^{lp} \rightarrow 0,$$

pois, $(\tau_1 - \lambda_n)/(\lambda_1 - \lambda_n) < 1$. O que é uma contradição. Concluimos então que $\tau_1 = \lambda_1$.

Um argumento similar pode ser feito para cada j , nos casos $\tau_j < \lambda_j$, e isso conclui a demonstração. \square

Até este ponto, foi analisada a convergência do IRL, porém não foi dito nada a respeito da taxa de convergência. A seguir, são apresentados alguns resultados que descrevem condições sob as quais obtém-se convergência rápida.

Lema 5.13. *Seja $A = B + \tau I$ uma matriz hermitiana $M \times M$, onde B é hermitiana semipositiva definida, $\text{posto}(B) = d$, e $\tau > 0$. Assuma que os d maiores autovalores de B são distintos. Se q_0 é não ortogonal aos d autovetores associados aos d maiores autovalores, e não ortogonal a no mínimo um autovetor associado a um autovalor não desejado, então $\dim(\mathcal{K}^{d+1}(A, q_0)) = d+1$, ou seja, o método de Lanczos converge em no máximo $d + 1$ passos.*

Demonstração:

A prova deste lema decorre da propriedade: $\mathcal{K}^m(A, q_o) = \mathcal{K}^m(A - \tau I, q_o) = \mathcal{K}^m(B, q_o)$ e $\dim(\mathcal{K}^m(B, q_o) = d + 1$, para todo $m > d$ [46]. \square

Outro resultado importante em relação à convergência do método de Lanczos, aplicado ao cálculo dos d maiores valores singulares, é que a taxa de convergência depende da separação entre σ_d e σ_{d+1} , (ver Teorema 6.4.1 em [16]). Quanto maior a separação, maior é a taxa de convergência. Considerando que no problema $\tilde{H} = H + E$ e $\text{posto}(H) = d$, a teoria de perturbação de valores singulares, garante que $\tilde{\sigma}_{d+1}(\tilde{H}) \leq \|E\|$. Assim, uma melhor separação entre $\tilde{\sigma}_d(\tilde{H})$ e $\tilde{\sigma}_{d+1}(\tilde{H})$ ocorrerá quando $\|E\| < \sigma_d(H)$ [15]. Isto mostra que o tamanho de $\sigma_d(H)$, como uma função da dimensão do problema e das características do sinal, desempenha um papel importante na convergência do método. O seguinte teorema fornece uma estimativa para $\sigma_d(H)$, no caso em que são utilizadas matrizes quadradas.

Teorema 5.14. *Seja H a matriz de Hankel $M \times M$, descrita em (2.6) com $j = 0$. Defina δ como no Teorema 4.6. Então um limitante inferior para $\sigma_d(H)$ satisfaz*

$$\sigma_d(H) \geq K \left[1 + \frac{D_M^2}{(d-1)\delta^2} \right]^{1-d} (1 - e^{-2\alpha\Delta t M}) \quad (5.24)$$

onde

$$K = \frac{1}{\rho} \frac{1}{(1 - e^{-2\alpha\Delta t})}, \quad \alpha = \min |\text{real}(s_l)|, \quad \rho = \min |r_l|,$$

onde D_M^2 é dada no teorema 4.6.

Demonstração:

Observe primeiramente que a matriz de Hankel pode ser rescrita como (veja a decomposição (2.8))

$$H = V_M R V_M^T,$$

onde V_M é uma matriz de Vandermonde $M \times d$, descrita como em (2.9), e $R = \text{diag}(\tau_1, \dots, \tau_d)$. A seguir, usando propriedades de valores singulares segue que

$$\sigma_d(H) \geq \sigma_d(R) \sigma_d(V_M)^2. \quad (5.25)$$

A demonstração termina aplicando o Teorema 4.6. \square

Como já foi visto anteriormente, no lema 4.7, $D_M^2 \approx 0$ quando M é suficientemente grande e $|z_l| \approx 1$, o que é muito freqüente em aplicações envolvendo processamento de sinais. Além disso, em [4], é analisado o comportamento do limitante dado em (5.25), para matrizes de Hankel retangulares, e a conclusão apresentada é que a estimativa será melhor quando a matriz de Hankel for quadrada.

O teorema 5.14 sugere ainda que $\sigma_d(H)$ deve crescer com a dimensão de H , e que a estimativa melhora quando os z_l ficam perto do círculo unitário no plano complexo, porém, não muito próximos uns dos outros. Portanto, deve-se utilizar matrizes de Hankel quadradas, com a maior dimensão possível, a fim de que $\sigma_d(H)$ possa vencer o tamanho da perturbação $\|E\|$, possibilitando assim, uma boa separação entre $\tilde{\sigma}_d(\tilde{H})$ e $\tilde{\sigma}_{d+1}(\tilde{H})$.

Dos resultados obtidos do Lema 5.13 e Teorema 5.14, deduz-se que para obter convergência rápida do método de Lanczos, no cálculo dos d maiores valores singulares da matriz de Hankel \tilde{H} e correspondente subespaço invariante, deve-se ter $\|E\| < \sigma_d(H)$, a dimensão da matriz de Hankel deve ser a maior possível, e o vetor inicial q_0 deve ter fortes componentes no subespaço de interesse. Quando isto é satisfeito, pode-se esperar que a convergência seja rápida. Na seção seguinte é apresentado um exemplo numérico que mostra que se $\|E\| \ll \sigma_d(H)$, então o método de Lanczos converge em pouco mais de d iterações, desde que vetor inicial seja escolhido adequadamente.

Finalizando esta seção, é apresentada uma escolha de vetor inicial para o método de Lanczos, a qual mostra resultados melhores em relação à velocidade de convergência, comparada com uma escolha aleatória. O vetor

$$q_0 = \tilde{H}^* \tilde{b}, \quad (5.26)$$

onde $\tilde{b} = [\tilde{h}_{j-1}, \tilde{h}_j, \dots, \tilde{h}_{j+M}]^*$, é proposto como vetor inicial. Para justificar esta escolha, considere a decomposição em valores singulares (SVD) de \tilde{H} :

$$\tilde{H} = \tilde{U} \tilde{\Sigma} \tilde{V}^* = \tilde{\sigma}_1 \tilde{u}_1 \tilde{v}_1^* + \tilde{\sigma}_2 \tilde{u}_2 \tilde{v}_2^* + \dots + \tilde{\sigma}_N \tilde{u}_N \tilde{v}_N^*.$$

Usando esta decomposição, segue imediatamente que

$$\tilde{v}_j^* \tilde{H}^* \tilde{b} = \tilde{\sigma}_j \tilde{u}_j^* \tilde{b}, \quad j = 1, 2, \dots, N.$$

Este resultado mostra que, se $\|E\| \ll \sigma_d$ e $\|\tilde{b} - b\| \ll \|b\|$, onde $b = [h_{j-1}, \dots, h_{j+M}]^*$, então as componentes mais fortes do vetor q_0 estão associadas com os d primeiros vetores singulares, os quais deseja-se aproximar.

5.6 Experiências Numéricas

Para observar o desempenho do IRL em relação à escolha do vetor inicial q_0 , à escolha de p e ao uso da multiplicação rápida, são apresentados resultados de alguns experimentos numéricos realizados com um sinal proveniente de simulações em Ressonância Magnética Nuclear, [39]. Para contar o número de operações de

ponto flutuante, foi usada a função *flops* do Matlab. O sinal foi construído pela seguinte modelagem:

$$\tilde{h}_j = \sum_{k=1}^d r_k e^{i\phi_k} e^{(\alpha_k + i\omega_k)j\Delta t} + \epsilon_j, \quad j = 1, 2, \dots, 512,$$

onde os parâmetros r_l , α_l , ω_l e ϕ_l representam a amplitude real, o fator de decaimento, a frequência angular e a fase, respectivamente. Os ϵ_j correspondem ao ruído, o qual, em nossos testes foi gerado usando o comando *randn* do Matlab com diferentes valores de desvio padrão ($\sigma = 5, 10, 15$). O sinal mostrado possui 11 exponenciais, e os parâmetros r_l , α_l , ω_l e ϕ_l são dados na tabela 5.1.

l	1	2	3	4	5	6	7	8	9	10	11
r_l	75	150	75	150	150	150	150	150	1400	60	500
α_l	-50	-50	-50	-50	-50	-50	-50	-25	-286	-25	-200
$\omega_l \setminus (2\pi)$	-86	-70	-54	152	168	292	308	360	440	490	530
ϕ_l	135	135	135	135	135	135	135	135	135	135	135

Tabela 5.1: Parâmetros do sinal, onde r_l é dada em unidades arbitrárias, α_l e ω_l em Hz e ϕ_l em graus.

A partir do sinal, foram geradas matrizes de Hankel de ordem 256×256 , de acordo com (2.6). Na figura 5.1, mostramos os 15 maiores valores singulares de \tilde{H} , para dois níveis de ruído correspondentes a $\sigma = 5$ e $\sigma = 15$, bem como os 11 valores singulares não nulos da matriz de Hankel H . Observe que para $\sigma = 5$, existe um gap bem notório entre $\tilde{\sigma}_{11}$ e $\tilde{\sigma}_{12}$. Quando o nível de ruído é aumentado para $\sigma = 15$, este gap diminui, tornando o problema de calcular os 11 maiores valores e correspondentes vetores singulares mais complicado. Além disso, se o ruído introduzido for muito alto, ou seja se $\|E\| > \sigma_{11}$, este último valor singular pode não ser recuperado.

σ	p	Recomeços q_0 aleatório	Recomeços $q_0 = \tilde{H}^* \tilde{b}$	razão1	razão2
5	5	3	0	3.2476	4.5251
10	7	5	0	5.6751	4.2278
15	10	8	8	1.0057	2.2862
15	11	7	0	9.7371	3.6875

Tabela 5.2: Comparação do número de operações de ponto flutuante para o IRL-ES, p denota o Número de passos Lanczos adicionais.

Para observar o desempenho do método em relação à escolha do vetor inicial, foram realizados testes numéricos com q_0 aleatório e com $q_0 = \tilde{H}^* \tilde{b}$, para vários valores de p . Os resultados dos experimentos são mostrados na Tabela 5.2, onde razão1 denota o quociente entre o número de flops gastos quando o método é

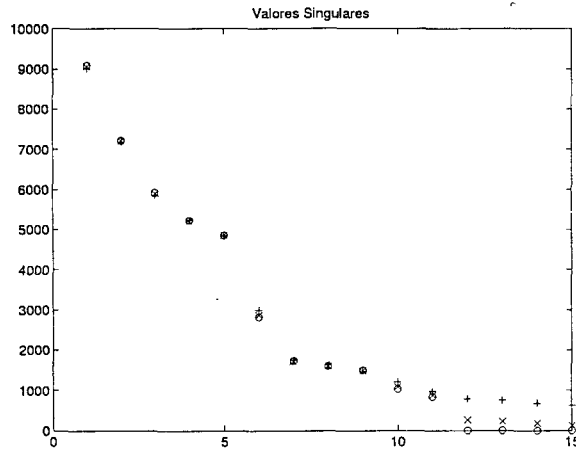


Figura 5.1: Em 'o' os 11 valores singulares não nulos de H , em 'x' e '+' os 15 maiores valores singulares de \tilde{H} para $\sigma = 5$ e $\sigma = 15$, respectivamente.

inicializado com q_0 aleatório e o número de flops gastos quando $q_0 = \tilde{H}^* \tilde{b}$. Analisando os resultados numéricos dessa tabela, concluímos que se o nível de ruído é alto, o valor de p deve ser tomado um pouco maior do que d , caso contrário, o valor de p é em geral menor do que d . Além disso, verificamos que a convergência é mais rápida quando o vetor inicial é escolhido como $q_0 = \tilde{H}^* \tilde{b}$. Os valores apresentados nas duas últimas colunas, representam a razão entre o número de flops do IRL-ES sem o uso da multiplicação rápida, e com o uso da multiplicação rápida, a coluna razão1 refere-se ao IRL-ES com escolha inicial aleatória, e a coluna razão2 refere-se ao IRL-ES com escolha inicial $q_0 = \tilde{H}^* \tilde{b}$. Foram realizados os mesmos testes para outros sinais, e os resultados obtidos foram similares.

Capítulo 6

Experimentos Numéricos

Neste capítulo, são apresentados os resultados numéricos da estimativa dos parâmetros α_l e ω_l . São ilustrados os resultados do uso de dois métodos, o método de Kung e o método HTLS, os quais foram descritos na seção 3.2. Para isto, os dois primeiros sinais utilizados foram perturbados com ruído Gaussiano, e a partir destes sinais perturbados, foram construídas matrizes de Hankel, com dimensão 256×256 , este processo foi repetido 100 vezes para cada sinal. O terceiro sinal é um sinal real, e provém de experimentos práticos, para este sinal são apresentadas somente estimativas dos parâmetros α_l e ω_l .

Para facilitar a compreensão das tabelas, foi criada a seguinte notação:

- N_{sr} : Relação entre o ruído e o sinal, definida por

$$N_{sr} = \frac{\|\epsilon\|}{\|h\|}$$

- E_α : Erro relativo para o fator de decaimento estimado, definido por

$$E_\alpha = \frac{|\alpha - \tilde{\alpha}|}{|\alpha|},$$

- E_ω : Erro relativo das frequências estimadas, dado por

$$E_\omega = \frac{|\omega - \tilde{\omega}|}{|\omega|},$$

6.1 Sinal com 5 exponenciais proveniente de NMR

O primeiro sinal utilizado possui 5 exponenciais, é um sinal complexo típico de ressonância magnética nuclear, foi extraído de [21] e seus parâmetros são dados na tabela 4.1. Na figura 6.1, pode-se observar que a separação entre os 5 maiores valores singulares e os demais, é bastante notória, com isto a identificação não é tão complicada. Na mesma figura, é apresentado o espectro de Fourier do sinal exato e do sinal perturbado. Nas tabelas 6.1 e 6.2, têm-se as estimativas para α_l e ω_l respectivamente, bem como, a medida do erro relativo das estimativas correspondentes aos 100 experimentos. Pode-se perceber que o método HTLS produz melhores estimativas, porém as diferenças não são tão acentuadas.

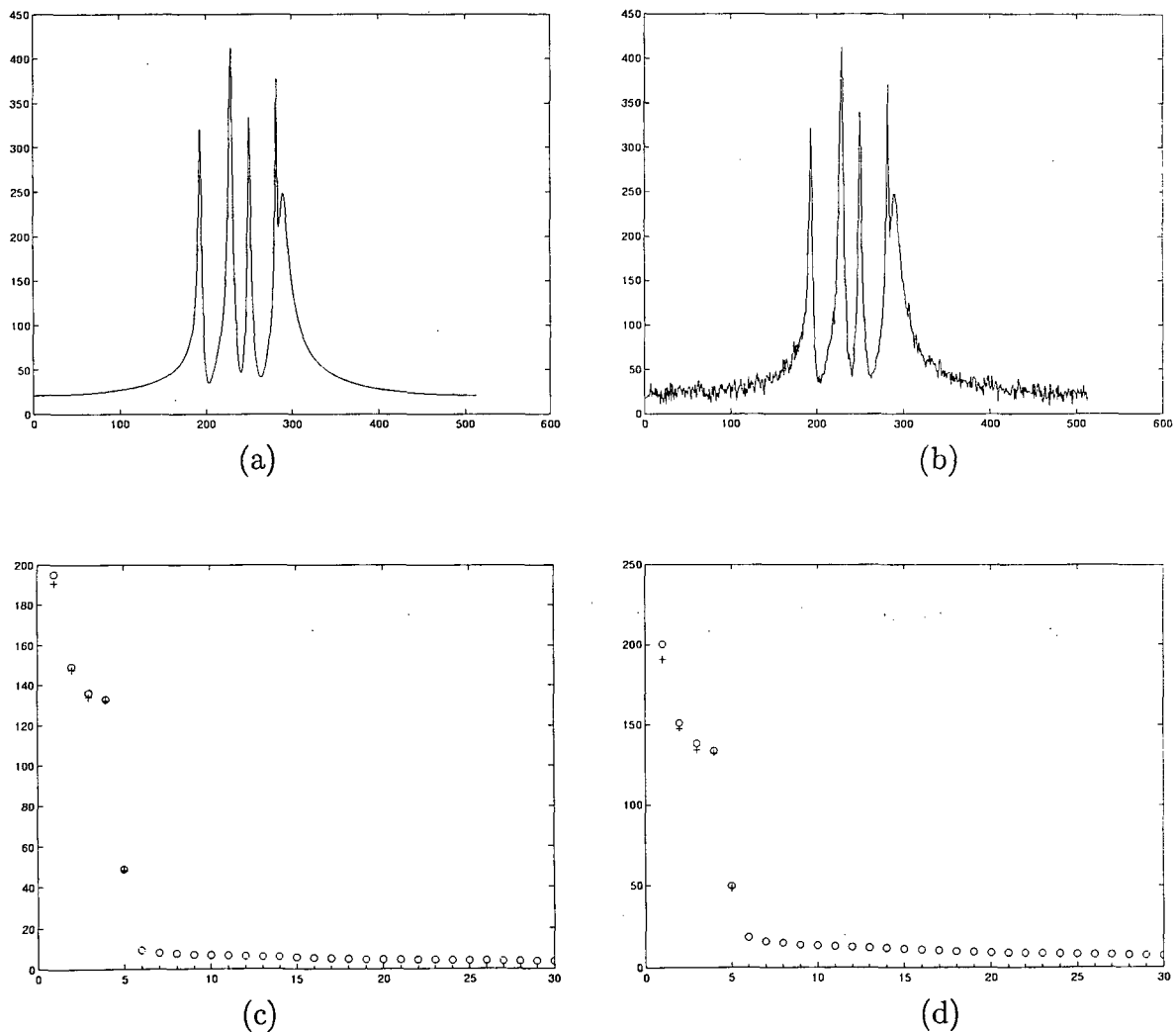


Figura 6.1: (a): Espectro de Fourier do primeiro sinal exato, (b): Espectro de Fourier do primeiro sinal perturbado, $Nsr = 10\%$. Trinta primeiros valores singulares de H (em '+') e de \tilde{H} (em 'o'), para o primeiro sinal - (c): $Nsr = 10\%$. (c): $Nsr = 20\%$.

			Kung		HTLS	
N_{sr}	l	α_l	$\tilde{\alpha}_l$	E_α	$\tilde{\alpha}_l$	E_α
10	1	208	209.43	0.0069	209.24	0.0060
	2	256	256.76	0.0030	256.67	0.0026
	3	197	197.53	0.0027	197.39	0.0020
	4	117	118.86	0.0159	118.07	0.0092
	5	808	804.22	0.0047	801.92	0.0075
20	1	208	213.00	0.0254	212.26	0.0216
	2	256	257.62	0.0139	257.23	0.0106
	3	197	196.25	0.0029	195.60	0.0055
	4	117	125.25	0.0397	121.71	0.0227
	5	808	802.59	0.0067	792.94	0.0186

Tabela 6.1: Estimativas dos fatores de decaimento para o primeiro sinal: $N_{sr} = 10\%$ e 20% .

			Kung		HTLS	
N_{sr}	l	ω_l	$\tilde{\omega}_l$	E_ω	$\tilde{\omega}_l$	E_ω
10	1	-1379	-1378.38	0.0004	-1378.37	0.0005
	2	-685	-685.55	0.0008	-685.54	0.0008
	3	-271	-270.58	0.0015	-270.57	0.0016
	4	353	354.06	0.0030	354.22	0.0035
	5	478	475.47	0.0053	475.30	0.0056
20	1	-1379	-1377.87	0.0008	-1377.85	0.0008
	2	-685	-686.13	0.0017	-686.11	0.0016
	3	-271	-270.27	0.0027	-270.23	0.0028
	4	353	355.10	0.0060	355.88	0.0082
	5	478	474.68	0.0069	473.84	0.0087

Tabela 6.2: Estimativas das frequências para o primeiro sinal: $N_{sr} = 10\%$ e 20% .

6.2 Sinal com 11 exponenciais proveniente de MRS

O segundo sinal apresentado foi derivado a partir do espectro de um sinal 31^P *in vivo* do tecido de um cérebro humano [39]. Seus parâmetros são dados na tabela 5.1. Este é um sinal mais sensível do que o primeiro, pois seu espectro apresenta picos muito próximos, veja Figura 6.3. Em termos da matriz de observação \tilde{H} , isto significa que a norma do ruído $\|E\| = \|\tilde{H} - H\|$, torna-se muito próxima do

			Kung		HTLS	
Nsr	l	α_l	$\tilde{\alpha}_l$	E_α	$\tilde{\alpha}_l$	E_α
3	1	50	50.38	0.0077	50.76	0.0152
	2	50	50.36	0.0073	50.90	0.0182
	3	50	50.43	0.0086	50.68	0.0138
	4	50	50.96	0.0193	49.57	0.0085
	5	50	49.19	0.0161	49.65	0.0069
	6	50	50.70	0.0141	49.10	0.0180
	7	50	49.88	0.0023	49.70	0.0059
	8	25	24.88	0.0045	24.91	0.0033
	9	285.7	284.95	0.0026	284.87	0.0029
	10	25	24.76	0.0094	25.18	0.0073
	11	200	201.31	0.0066	199.76	0.0012

Tabela 6.3: Estimativas dos fatores de decaimento para o segundo sinal: $Nsr = 3\%$.

menor valor singular exato $\sigma_{11}(H)$, ainda para baixos níveis de ruído, o que impede que exista uma boa separação entre os valores singulares $\tilde{\sigma}_j$, veja Figura 6.2.

São apresentadas as estimativas dos parâmetros α_l e ω_l para dois níveis de ruídos $Nsr = 3\%$ e $Nsr = 6\%$, tabelas 6.3, 6.4 e 6.5,6.6, respectivamente.

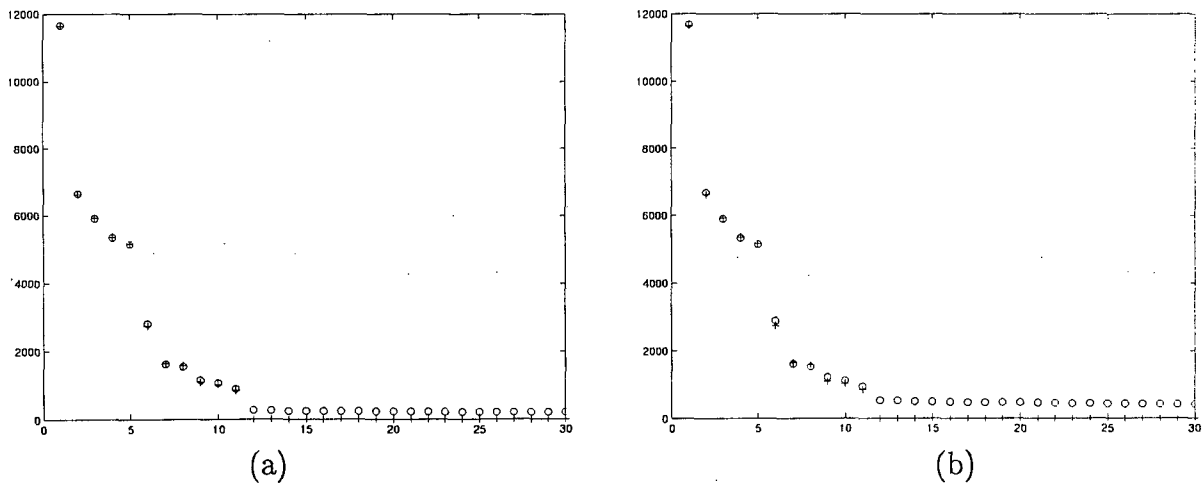


Figura 6.2: 30 primeiros valores singulares de H (em '+') e de \tilde{H} (em 'o'), para o segundo sinal - (a): $Nsr = 3\%$. (b): $Nsr = 6\%$.

		Kung			HTLS	
Nsr	l	ω_l	$\tilde{\omega}_l$	E_ω	$\tilde{\omega}_l$	E_ω
3	1	-86	-85.96	0.0004	-85.94	0.0006
	2	-70	-69.92	0.0010	-69.93	0.0009
	3	-54	-54.02	0.0005	-54.04	0.0008
	4	152	151.99	0.0000	152.00	0.0001
	5	168	168.00	0.0000	167.99	0.0001
	6	292	291.74	0.0009	291.75	0.0008
	7	308	308.17	0.0006	308.16	0.0005
	8	512	440.05	0.1405	360.01	0.0000
	9	440	490.00	0.1136	440.17	0.0004
	10	490	512.02	0.0450	490.02	0.0001
	11	530	529.48	0.0010	529.50	0.0009

Tabela 6.4: Estimativas das frequências para o segundo sinal: $Nsr = 3\%$.

			Kung		HTLS	
Nsr	l	α_l	$\tilde{\alpha}_l$	E_α	$\tilde{\alpha}_l$	E_α
6	1	50	52.42	0.0485	53.01	0.0603
	2	50	52.26	0.0452	52.01	0.0402
	3	50	51.40	0.0281	51.60	0.0321
	4	50	51.87	0.0376	48.66	0.0267
	5	50	49.73	0.0052	46.56	0.0687
	6	50	48.68	0.0264	48.30	0.0339
	7	50	48.64	0.0271	48.25	0.0349
	8	25	25.12	0.0051	24.80	0.0077
	9	285.7	283.12	0.0090	282.57	0.0109
	10	25	24.74	0.0104	25.72	0.0292
	11	200	202.89	0.0145	199.0991	0.0045

Tabela 6.5: Estimativas dos fatores de decaimento para o segundo sinal: $Nsr = 6\%$.

			Kung		HTLS	
Nsr	l	ω_l	$\tilde{\omega}_l$	\bar{E}_ω	$\tilde{\omega}_l$	\bar{E}_ω
6	1	-86	-86.02	0.0003	-85.96	0.0004
	2	-70	-70.07	0.0011	-70.03	0.0005
	3	-54	-53.92	0.0014	-54.01	0.0003
	4	152	151.90	0.0006	151.93	0.0004
	5	168	168.12	0.0008	168.08	0.0005
	6	292	291.39	0.0021	291.44	0.0019
	7	308	308.49	0.0016	308.42	0.0014
	8	512	440.32	0.1400	360.03	0.0001
	9	440	490.03	0.1137	440.62	0.0014
	10	490	512.07	0.0451	490.07	0.0001
	11	530	528.89	0.0021	528.76	0.0023

Tabela 6.6: Estimativas das frequências para o segundo sinal: $Nsr = 6\%$.

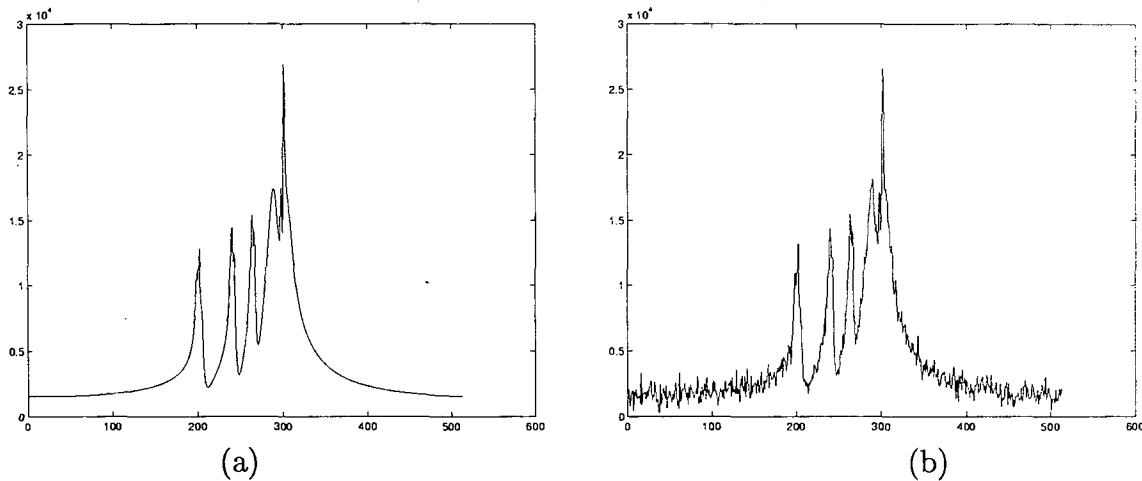


Figura 6.3: (a): Espectro de Fourier do segundo sinal exato, (b): Espectro de Fourier do segundo sinal perturbado, $Nsr = 3\%$.

6.3 Sinal da Análise Modal experimental

O sinal analisado é um sinal real, composto por exponenciais complexas que vêm em pares conjugados. Ele foi obtido no Laboratório de Vibrações e Acústica da UFSC, e corresponde à resposta ao impulso de um protótipo de um prédio de três andares [1]. São disponíveis 1024 amostras, com $\Delta t = 1/512$. Como o número de componentes do sinal não pode ser detectado através do espectro singular da matriz de observação, veja Tabela 6.7, o método HTLS foi aplicado usando dimensões do “subespaço sinal” $d = 8, 10, \text{ e } 12$. Os resultados do experimento são mostrados na Tabela 6.8. Nessa tabela, o símbolo # indica a presença de autovalores reais, os quais foram descartados, pois não apresentam nenhuma contribuição no sis-

tema vibratório. Os $\xi_k = \alpha_k/|s_k|$ são os fatores de amortecimento modal, enquanto que os f_k são as frequências naturais de vibração, medidas em Hertz. Os resultados obtidos concordam com resultados publicados em [1], onde foram apresentados resultados da análise modal global da mesma estrutura, obtidos através de uma técnica do tipo multi-input multi-output. A qualidade dos parâmetros obtidos é ilustrada na Figura 6.4, onde são apresentados o sinal original e o reconstruído usando 10 componentes.

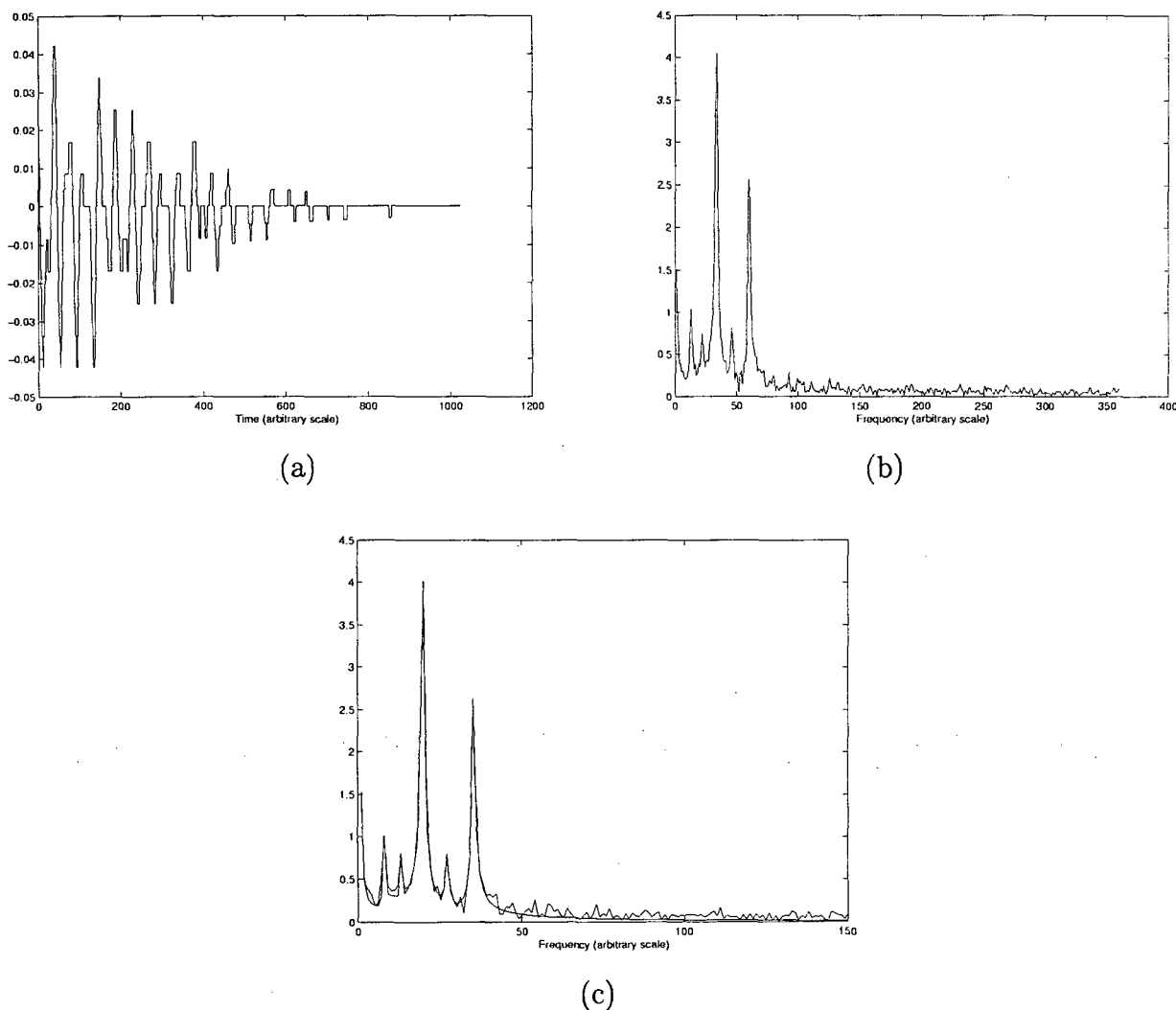


Figura 6.4: (a): Sinal Experimental. (b): Espectro de Fourier do sinal. (c): Reconstrução do sinal usando os parâmetros calculados.

k	$\tilde{\sigma}_k$	k	$\tilde{\sigma}_k$
1	2.2822	8	0.4241
2	2.2275	9	0.4208
3	1.4403	10	0.2343
4	1.4117	11	0.2340
5	0.7036	12	0.2163
6	0.4643	13	0.2142
7	0.4600	14	0.1564

Tabela 6.7: Valores singulares da matriz de observação para $N = 360$

k	HTLS ($d = 8$)		HTLS ($d = 10$)		HTLS ($d = 12$)	
	ξ_k	f_k (Hz)	ξ_k	f_k (Hz)	ξ_k	f_k (Hz)
1	0.0358	4.6876	0.0543	4.6802	0.0500	4.6412
2	#	0	#	0	#	0
3					0.0007	11.4961
4	0.0186	13.4313	0.0193	13.4007	0.0194	13.4166
5			0.0144	18.5458	0.0139	18.5663
6	0.0105	24.2784	0.0101	24.2915	0.0101	24.2940

Tabela 6.8: Parâmetros obtidos a partir do sinal experimental

Conclusões

O problema estudado é um problema difícil em diversas situações, como já foi comentado, principalmente devido à dificuldade em separar as contribuições do sinal, daquelas associadas ao ruído.

Através dos limitantes obtidos para o condicionamento do problema de autovalor associado, e da teoria clássica de perturbação de autovalores, pode-se concluir que a análise da sensibilidade do problema de autovalor associado a métodos de recuperação de exponenciais, baseados em técnicas de subespaço, mostrou condições sob as quais as exponenciais tornam-se insensíveis a pequenas perturbações nos dados de entrada. Em todas as aplicações nas quais são analisados sinais pouco amortecidos, as exponenciais envolvidas tornam-se pouco sensíveis a perturbações nos dados, sempre que a dimensão das matrizes utilizadas na análise seja suficientemente grande.

A transformação do problema de valores singulares num problema de autovalores, permitiu um cálculo parcial dos valores singulares, o que é muito relevante nos problemas apresentados neste trabalho, pois os métodos tradicionais para o cálculo de valores singulares calculam todos os valores singulares e não levam em conta a estrutura da matriz H . O uso da multiplicação rápida permitiu explorar a estrutura da matriz H , tornando mais rápida a multiplicação matriz-vetor. Por outro lado, foi proposta uma escolha inicial para o processo de Lanczos, que aumenta a taxa de convergência, em relação à escolha aleatória, isto foi verificado através de exemplos.

Um problema que deve ser estudado é a determinação do número de exponenciais a partir do sinal perturbado. Foram feitas algumas tentativas a partir da análise dos coeficientes do vetor $\tilde{H}^* \tilde{b}$, escrito como uma combinação dos vetores singulares à direita da matriz de Hankel (veja (5.26)), e observou-se em vários testes que o número de componentes do sinal podia ser detectado observando-se o número de coeficientes destacados em módulo. Porém, acredita-se que este assunto necessita de uma análise mais cuidadosa, o que poderá ser abordado futuramente.

Referências Bibliográficas

- [1] F. S. V. BAZÁN E C. BAVASTRI, An Optimized Pseudo-Inverse Algorithm (OPIA) For Multi-Input Multi-Output Modal Parameter Identification, *Mechanical Systems and Signal Processings*, 10 (1996), pp. 365–380.
- [2] F. S. V. BAZÁN E L. H. BEZERRA, On Zero Location of Predictor Polynomials, *Numer. Linear Algebra With Applications*, 4(6) (1997), pp. 459–468.
- [3] F. S. V. BAZÁN E PH. L. TOINT E M. C. ZAMBALDI, A Conjugate-Gradients Based Method for Harmonic Retrieval Problems that Does Not Use Explicit Signal Subspace Computation, technical report 16, Department of Mathematics, FUNDP, Namur, Belgium, November 1997.
- [4] F. S. V. BAZÁN E PH. L. TOINT, Singular values of predictor matrices and signal eigenvalue bounds, technical report 5, Department of Mathematics, FUNDP, Namur, Belgium, February 1998.
- [5] F. S. V. BAZÁN, Conditioning of Rectangular Vandermonde Matrices with nodes in the Unit Disk, *SIAM Matrix Analysis and Applications*, Vol 21(2) pp. 679-693, 2000.
- [6] F. S. V. BAZÁN, Sensitivity Eigenanalysis of Single Shift-Invariant Subspace-Based Methods, *Signal Processing*, 80 pp. 89-100, 2000.
- [7] R. DE BEER E D. VAN ORMONDT, Analysis of NMR Data Using Time Domain Fitting Procedure. *NMR Basic Principles and Progress*, vol 26, 1992.
- [8] R. DE BEER, c59 Quantitative In Vivo NMR (lecture notes). University of Technology Delft, Department of Applied Physics, The Netherlands, 1994
- [9] Å. BJÖRCK, *Numerical Methods for Least Squares Problems*. SIAM, Philadelphia, 1996.
- [10] D. CALVETTI, L. REICHEL E D. C. SORENSEN, An Implicitly Restarted Method for Large Symmetric Eigenvalue Problems. *Electronic Transactions on Numerical Analysis*, 2:1-21, 1994.

- [11] R. CARIÈRE E R. L. MOSES, High Resolution Radar Target Modeling Using a Modified Peony Estimator. *IEEE Transaction on Antenas e Propagation*, vol 40.1, 1992.
- [12] P. J. DAVIS, *Circulant Matrices*. John Wiley & Sons, 1979.
- [13] E. LINDSTRÖM, *Singular Value Computations for Toeplitz Matrices*. Lic. Thesis, Linköping University, ISBN 91-7871-719-1, May 1996.
- [14] W. GIVENS, Computations of Plane Unitary Rotations Transforming a General Matriz to Triangular Form. *SIAM J. Appl. Math.*, 6:26-50,1958.
- [15] G. H. GOLUB E C. F. VAN LOAN, *Matrix Computations*, Johns Hopkins University Press, Baltimore, third edition, 1996.
- [16] P. C. HANSEN, *Rank-Deficient and Discrete Ill-Posed Problems*. SIAM, Philadelphia, 1997.
- [17] P. S. R. HANSEN, *Signal Subspce Methos for Speech Enchancement*. PH.D Thesis, Department of Mathematical Modeling, Technical University of Denarc, 1997.
- [18] A. S. HOUSEHOLDER, Unitary Triangularization of a Nonsymmetric Matrix. *Jornal of the ACM*, 5:339-342,1958.
- [19] Y. HUA AND T. P. SARKAR, Matrix Pencil Method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Trans. On Acoust. Speech and Signal Processings*, ASSP - 38(5):814-824, May 1990.
- [20] S. VAN HUFFEL E J. VANDEWALLE, *The Total Least Squares Problem - Computational Aspects and Analysis*. SIAM, Philadelphia, 1991.
- [21] S. VAN HUFFEL, H. CHEN, C. DECANNIERE E P. VAN HECKE, Algorithm for time-domain NMR data fitting based on Total Least Squares. *Jornal of Magnetic Resonance A*, 110:228-237, 1994.
- [22] R. KUMARESAN E D. W. TUFTS, Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise. *IEEE Trans. Acoust. Speech and Sig. Proc.*, 30:833-840, 1982.
- [23] R. KUMARESAN, D. W. TUFTS E L. L. SCHARF, A Prony Method for noisy data: Choosing the signal components and selecting the order in Exponential Signal Models. *Proceedings of the IEEE*, vol 72:230-233, February, 1984.

- [24] S.Y. KUNG, *A New Identification and Model Reduction Algorithm via Singular Value Decomposition. Twelfth Asilomar Conf. on Circuits, Systems and Comp.*, (Asilomar,CA), 705-714, Nov. 1978.
- [25] R. B. LEHOUCQ, *Analysis and Implementation of an Implicitly Restarted Arnoldi Iteration*. PhD Thesis, Rice University, Houston, Texas, 1995.
- [26] R. B. LEHOUCQ E D. C. SORENSEN, Deflation techniques for an Implicitly Restarted Arnoldi Iteration. *SIAM J. Matrix Anal. Appl.*, 17(4):789-821, 1996.
- [27] C. PAIGE, *The Computation of Eigenvalues and Eigenvectors of Very Large Sparce Matrices*. PhD Thesis, University of London, 1971.
- [28] C. PAIGE, Error Analysis of the Lanczos Algorithm for Tridiagonalizing a Symmetric Matrix. *J. Inst. Math Applic.*, 18:341-349, 1976.
- [29] B. N. PARLETT, *The Symmetric Eigenvalue Problem*. SIAM, Philadelphia, 1997.
- [30] B. N. PARLETT E D. S. SCOTT, The Lanczos Algorithm with Selective Orthogonalization. *Math. Comp.*,33(145):217-238,1979.
- [31] R. DE PRONY, Essai expérimental et analytique sur les loi de la dilatabilité et sur celles de la force expansive de la vapeur de l'eau et de la capeur de l'alkool, à différentes températures. *J. de l'Ecole Polytechnique*, 1:24-76, 1795.
- [32] H. D. SIMON, Analysis of the Symmetric Lanczos Algorithm with reorthogonalization Methods. *Linear Algebra Appl.*, 61:101-131,1984.
- [33] R. A. SMITH, The Condition Numbers of the Matrix Eigenvalue Problem. *Numerische Mathematik*, 10, 132:240, 1967.
- [34] G. W. STEWART, Perturbation Theory for the Singular Value Decomposition. *SVD and Signal Processing II: Algorithms, Analysis and Applications*, 99-109, Elsevier, 1991.
- [35] G. W. STEWART, Error and perturbations bounds for subspaces associated with certain eigenvalue problems. *SIAM Review*. 32:727-764, 1973.
- [36] G. STRANG, *Linear Algebra and its Aplications*. Academic Press, 1992.
- [37] D. C. SORENSEN, Implicit Aplication of Polynomial Filters in a k-step Arnoldi Method. *SIAM J. Matrix Appl.*, 13:357-385, 1992.

- [38] H. RUTISHAUSER, Simultaneous Iteration Method for Real Symmetric Matrices. *Numer. Math.* 16, 205-223, 1970.
- [39] J. TOTZ, A. V. D. BOOGAART, S. V. HUFFEL, D.GRAVERON-DEMILLY, I. DOLOGKOU, R. HEIDLER AND D. MICHEL, The Use of Continuous Regularization in the Automated Analysis of MRS Time-Domain Data. *Journal of Magnetic Resonance.* 124, 400-409, 1997.
- [40] R. J. VACCARO, A second pertubation expansion for the SVD. *SIAM Journal on Matrix Analysis and Applications*, 15(2):661-671, Abril, 1994.
- [41] A. VAM DER VEEN, E. F. DEPRETTERE E A. LEE SWINDLEHURST, Subspace-based Signal Analysis using Singular Value Decomposition. *Proceedings of the IEEE*, 81(9):1277-1309, Setembro, 1993.
- [42] M. VIBERG E B. OTTERSTEN, Sensor Array Procecing Based on Subspace Fitting, *IEEE Trans. Signal Processing*, vol 39, pp. 1110-1121, 1991.
- [43] D. S. WATKINS, *Fundamentals of Matrix Computations*. John Wiley & Sons, New York, 1191.
- [44] G. WALKER, On periodicity in series of related terms. *Proc. Royal Soc. London Ser. A*, 131A:518-532, 1931.
- [45] J. H. WILKINSON, *The Algebraic Eigenvalue Problem*. Oxford University Press, Oxford, Great Britain, 1965.
- [46] G. XU AND T. KAILATH, Fast Estimation of Principal Eigenspace using Lanczos Algorithm, *SIAM Journal on Matrix Analysis and Applications*, 15:974-994, 1994 .
- [47] G. U. YULE, On a method of investigating periodices in disturbed series, with special reference to Wolfer's sinspot numbers. *Philos. Trans. Roy. Soc. London ser. A*, 226A:267:298, 1927.
- [48] H. P. ZEIGER E A. J. MAC-EWEN, Aproximate Linear Realization of Given Dimension via Ho's Algorithm, *IEEE, Trans. Aut. Control AC-19*, pp 153, 1974.