

Sensibilidade de Autovalores em Recuperação de dados Exponenciais¹

F.S.V. BAZÁN², Departamento de Matemática Universidade Federal de Santa Catarina, Campus Universitário, 88010-970 Florianópolis, SC, Brasil.

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 provêm 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. Apresentamos uma análise de sensibilidade do problema de autovalor associado a uma classe de algoritmos e concluímos que, sob certas circunstâncias que são freqüentes em muitas aplicações, esses autovalores são *insensíveis* a pequenas perturbações nos dados de entrada. Apresentamos estimativas teóricas para o erro máximo nos autovalores, bem como experiências numéricas que ilustram a teoria.

1. Introdução

A estimativa de parâmetros embutidos em modelos exponenciais a partir de dados inexatos é um problema que tem sido motivo de muita pesquisa na última década. São encontradas aplicações em áreas como ressonância magnética nuclear (NMR), análise modal de estruturas mecânicas, análise de dados em geofísica, entre outras [15, 1, 6, 7, 9, 8, 13]. Uma abordagem usual para estimar os parâmetros desejados é através da computação de um subespaço associado a uma matriz $\tilde{A} = A + E$, de ordem $m \times n$, contendo os dados disponíveis, onde assume-se que A contém os dados livres de incertezas (ruídos), $\text{posto}(A) = r \ll \min\{m, n\}$, isto é, A é de posto incompleto, e E simula incertezas nos dados, as quais podem ser de natureza diversa (imprecisão de aparelhos que coletam os dados, por exemplo). A hipótese comum nesses modelos é que o subespaço coluna da matriz que contém os dados livres de incertezas, denotado por \mathcal{S}_m , é gerado pelas colunas de uma matriz de Vandermonde W_m , de ordem $m \times r$, $m \gg r$, cujos elementos na posição (i, j) são números complexos da forma z_j^{i-1} , $z_j \neq z_i$, $|z_j| \leq 1$, e os parâmetros de interesse estão embutidos nesses z_j .

Sejam V_m uma matriz com colunas ortonormais que geram \mathcal{S}_m , A^\dagger a submatriz de V_m formada por suas $m - 1$ primeiras linhas, e A^\dagger a submatriz de V_m formada

¹Este trabalho foi desenvolvido com apoio do CNPq, Contrato 300487/94 - O(NV)

²fermin@mtm.ufsc.br

por suas $m - 1$ últimas linhas. Então, prova-se que existe uma matriz T , de ordem $r \times r$, tal que

$$A^\dagger T = A^\downarrow, \quad (1.1)$$

com o espectro de T satisfazendo $\lambda(T) = \{z_1, \dots, z_r\}$ [1]. Este resultado mostra que, se a matriz de dados for livre de ruídos, basta calcular uma matriz ortonormal que gera \mathcal{S}_m , o que pode ser feito via uma decomposição em valores singulares (SVD) de A , ou outra decomposição tal como QR, por exemplo. Os z_j são extraídos via computação dos autovalores de $T = (A^\dagger)^\dagger A^\downarrow$, onde \dagger denota pseudoinversão. Na prática porém, é disponível unicamente a matriz \tilde{A} . Assim, o problema consiste em construir aproximações \tilde{z}_j para os z_j , a partir dos autovalores de uma matriz \tilde{T} , a qual é construída como acima, sendo utilizada uma base ortonormal \tilde{V}_m computada a partir de \tilde{A} . Nessas circunstâncias, um problema de interesse é analisar a variação máxima possível dos autovalores z_j quando extraídos a partir de \tilde{T} . As contribuições neste contexto são escassas, e as poucas existentes se restringem à análise de um caso especial, onde são assumidas algumas propriedades estatísticas da matriz E . Duas exceções onde a matriz E é tratada algebraicamente, ou seja, onde E é vista apenas como uma perturbação, são o trabalho De Groen e De Moor [6], e outro recentemente desenvolvido por Bazán [3]. Nesses trabalhos, o que conta na análise de perturbação dos autovalores z_j é o tamanho da perturbação dado pela norma $\|E\|$ e não a natureza da mesma.

Neste trabalho, analisamos o problema da sensibilidade dos autovalores da matriz T devido à perturbações nos dados e deduzimos um limitante superior para o menor número de condição do problema de autovalor obtendo, com isso, melhores estimativas para o erro máximo nos autovalores z_j comparando com aqueles apresentados recentemente em Bazán [3]. Apresentamos como conclusões mais importantes as seguintes:

- A sensibilidade do problema de autovalor à perturbações nos dados é governada pelo número de condição $k_2(W_m)$.
- Se \tilde{z}_j é a estimativa para z_j obtida a partir de \tilde{T} , então, desde que m seja suficientemente grande, temos que

$$|z_l - \tilde{z}_l| = \mathcal{O}(\text{sen } \Theta),$$

onde Θ é o ângulo entre o subespaço exato \mathcal{S}_m e um subespaço aproximado $\tilde{\mathcal{S}}_m$, obtido a partir de \tilde{A}_m .

Este trabalho é organizado como segue. Na Seção 2, discutimos várias medidas da sensibilidade do problema e deduzimos limitantes para as mesmas. Nossas estimativas para o erro máximo nos autovalores vem na Seção 3, onde nosso resultado principal é que o erro é da ordem do seno do ângulo entre os subespaços \mathcal{S}_m e $\tilde{\mathcal{S}}_m$. Na Seção 4, apresentamos uma estimativa para o seno do ângulo entre subespaços, para o caso especial em que os dados são armazenados em matrizes de Hankel. Para ilustrar a teoria desenvolvida, na Seção 5 apresentamos resultados numéricos, onde analisamos um sinal proveniente de uma aplicação em análise modal de estruturas mecânicas.

2. Estimativas para a Sensibilidade do Problema

Nesta seção, demonstramos que o problema de autovalor associado à matriz T , definida em (1.1), é governada pelo condicionamento de uma matriz proveniente de um “scaling” das colunas da matriz de Vandermonde W_m descrita na seção anterior. Este resultado é utilizado posteriormente para estimar o número de condição de interesse, aproveitando um fato conhecido que relaciona a sensibilidade de autovalores com uma medida de *quase normalidade* da matriz envolvida [11].

Começaremos com a seguinte definição.

Definição 1 *Seja $G = X\Lambda X^{-1}$ uma decomposição autovalor autovetor de uma matriz G de ordem n . O número de condição de Jordan κ_2 , do problema de autovalor associado à G , é definido por*

$$\kappa_2 = \min_{D \in \mathcal{D}} \{\|XD\| \|D^{-1}X^{-1}\|\}, \quad (2.1)$$

onde $\|\cdot\|$ denota a norma espectral de matriz X e \mathcal{D} , o conjunto de todas as matrizes diagonais inversíveis. Quando é utilizada a norma de Frobenius, o número de condição de Jordan é denotado por κ_F .

É conhecido que se os autovalores λ_j de G são simples, então

$$\kappa_F = |s_1|^{-1} + \dots + |s_n|^{-1}, \quad (2.2)$$

onde $|s_j|^{-1}$, o número de condição associado ao autovalor λ_j , é definido por

$$s_j = \frac{y_j^* x_j}{\|y_j\| \|x_j\|}, \quad (2.3)$$

com x_j e y_j autovetores associados ao autovalor λ_j de G à direita e à esquerda, respectivamente. O símbolo $*$ denota conjugação complexa para vetores e transposição conjugada para matrizes.

Teorema 1 *Assuma que T satisfaz (1.1). Então, para todo $m \geq r + 1$,*

$$T = (V_m^* W_m) Z (V_m^* W_m)^{-1}, \quad (2.4)$$

onde V_m é uma matriz com colunas ortonormais, as quais geram S_m . Além disso,

$$\|T\|_F^2 = r + \|\hat{f}_m\|^2 - \|p_1\|^2, \quad (2.5)$$

onde p_1 é a primeira coluna da matriz de projeção sobre S_m , $\mathcal{P} = V_m V_m^*$ e \hat{f}_m é a solução de norma mínima do sistema de equações lineares

$$W_m^* f = Z^{*m} e,$$

onde $e = (1, \dots, 1)^T$ e $Z = \text{diag}(z_1, \dots, z_r)$.

Demonstração: Como as colunas de W_m geram \mathcal{S}_m , existe a matriz não singular F , de ordem $r \times r$, tal que

$$V_m = W_m F. \quad (2.6)$$

Considere V_m e W_m reescritas como

$$V_m = \begin{bmatrix} A^\dagger \\ a^* \end{bmatrix} = \begin{bmatrix} b^* \\ A^\dagger \end{bmatrix} \quad \text{e} \quad W_m = \begin{bmatrix} W_{m-1} \\ e^* Z^{m-1} \end{bmatrix} = \begin{bmatrix} e^* \\ W_{m-1} Z \end{bmatrix}.$$

Usando esta relação juntamente com (2.6), vem $A^\dagger = W_{m-1} Z F$, $A^\dagger = W_{m-1} F$ e $A^{\dagger\dagger} = F^{-1} W_{m-1}^\dagger$. Isto implica que

$$T = A^{\dagger\dagger} A^\dagger = F^{-1} W_{m-1}^\dagger W_{m-1} Z F = F^{-1} Z F,$$

pois $W_{m-1}^\dagger W_{m-1} = I$. A primeira parte do teorema resulta do fato de que $F^{-1} = V_s^* W_m$ e $F = W_m^\dagger V_s = (V_s^* W_m)^{-1}$. Para provar a segunda parte, considere a matriz companheira C , de ordem $m \times m$, definida por

$$C = \begin{bmatrix} e_2^* \\ \vdots \\ e_m^* \\ \widehat{f}_m^* \end{bmatrix}, \quad (2.7)$$

onde e_j é o j -ésimo vetor canônico em \mathbb{R}^m . Segue imediatamente que esta matriz satisfaz

$$C W_m = W_m Z, \quad (2.8)$$

e, com isso, prova-se que T pode ser reescrita como

$$T = V_m^* C V_m.$$

Usando esta relação, (2.8), e o fato que $V_m V_m^* W_m = W_m$, pois as colunas de W_m pertencem ao subspaço \mathcal{S}_m , segue que

$$T^* T = I + x x^* - y y^*, \quad (2.9)$$

onde $x = V_m^* \widehat{f}_m$ e $y = V_m^* e_1$. A relação (2.5), segue de aplicar o traço na relação acima. ■

Este teorema mostra que o número de condição κ_2 do problema de autovalor associado a T depende do condicionamento de uma matriz obtida por escalamento diagonal à direita da matriz de Vandermonde W_m . Usando (2.5) e certos resultados clássicos que relacionam os números de condição $|s_j|^{-1}$, κ_2 , κ_F , e uma medida de quase normalidade de uma matriz, obtemos o seguinte teorema:

Teorema 2 *Seja T como acima e defina o número D_m por*

$$D_m^2 = \|T\|_F^2 - \sum_{j=1}^r |z_j|^2. \quad (2.10)$$

Defina também $\delta_j = \min_{j \neq i} |z_i - z_j|$, $1 \leq i, j \leq r$. Então, o número de condição de Jordan κ_2 associado ao problema de autovalor da matriz T satisfaz

$$\kappa_2 \leq \frac{1}{2} \left[\eta - r + 2 + \sqrt{(\eta - r + 2)^2 - 4} \right], \quad (2.11)$$

onde

$$\eta \leq \sum_{j=1}^r \left(1 + \frac{D_m^2}{(r-1)\delta_j^2} \right)^{(r-1)/2} \quad (2.12)$$

Demonstração: Primeiro observamos que para toda matriz G de ordem r cujos autovalores são simples, temos que [11]

$$r - 2 + \kappa_2 + \kappa_2^{-1} \leq \kappa_F. \quad (2.13)$$

Resolvendo esta desigualdade para κ_2 , vem

$$\kappa_2 \leq \frac{1}{2} \left[\kappa_F - r + 2 + \sqrt{(\kappa_F - r + 2)^2 - 4} \right]. \quad (2.14)$$

Uma outra observação fundamental para nossa prova é que os números s_j satisfazem (ver [11])

$$|s_j|^{-1} \leq \left(1 + \frac{D_m^2}{(r-1)\delta_j^2} \right)^{(r-1)/2}.$$

Usando esta desigualdade em (2.2), obtemos

$$\kappa_F \leq \sum_{j=1}^r \left(1 + \frac{D_m^2}{(r-1)\delta_j^2} \right)^{(r-1)/2}.$$

A prova do teorema termina substituindo este limitante para κ_F em (2.13). ■

Deste teorema, conclui-se que se $D_m^2 \approx 0$, então $\kappa_F \approx r$ (ver (2.14)), a menos que os δ_j sejam muito pequenos em comparação com D_m . Com isso, teríamos que $\kappa_2 \approx 1$. Assim, o comportamento do número de condição κ_2 depende da separação dos z_j e da “grandeza” de D_m . O número D_m fornece uma medida de quanto uma matriz difere de ser normal: para matrizes normais, $D_m = 0$ e o problema de autovalor é perfeitamente condicionado (ou seja $\kappa_2 = 1$).

Um aspecto interessante em conexão com nosso problema é que utilizando (2.10) e Teorema 1, o número D_m pode ser estimado como

$$0 \leq D_m^2 \leq r + \|\hat{f}_m\|^2 - \sum_{j=1}^r |z_j|^2. \quad (2.15)$$

Esta é uma relação que pode ser monitorada facilmente, pois para m suficientemente grande, $\|\hat{f}_m\|^2 \approx 0$ sempre que $|z_j| < 1$, $j = 1 : r$, ou sempre que $|z_j| = 1$, $j = 1 : r$. Uma discussão detalhada sobre o comportamento de $\|\hat{f}_m\|^2$ como uma função de

m , pode ser encontrada em Bazan [2] (ver Teorema 2). Conseqüentemente, sempre que m é suficientemente grande e $|z_j| \approx 1$, usando (2.15) devemos obter $D_m^2 \approx 0$.

Da análise acima concluímos que em todas as aplicações nas quais $|z_j| \approx 1$, o que é muito freqüente em análise modal e NMR por exemplo, a condição para obtermos um problema de autovalor bem condicionado é que m seja suficientemente grande, a menos que os z_j sejam extremamente próximos uns aos outros.

3. Limitante para o Erro nos Autovalores

Como os teoremas acima descrevem sob quais condições o problema de autovalor associado à T é bem condicionado, passamos agora ao cálculo de estimativas para o erro dos autovalores. Para fazermos isso, vamos utilizar o resultado clássico da teoria de perturbação de autovalores, que afirma que se \tilde{z}_j é o autovalor de \tilde{T} que aproxima o autovalor z_j de T , onde $\tilde{T} = T + \mathcal{E}$ é uma aproximação para T , então

$$|\tilde{z}_j - z_j| \leq \kappa_2 \|\mathcal{E}\|, \quad j = 1 : r. \quad (3.1)$$

Na análise a seguir, \tilde{V}_m denota uma matriz de ordem $m \times r$ cujas colunas formam uma base ortonormal para um subespaço \tilde{S}_m , calculado através de algum método numérico a partir da matriz de observações $\tilde{A} = A + E$. A matriz \tilde{T} é calculada como $\tilde{T} = (\tilde{A}^\dagger)^\dagger \tilde{A}^\dagger$, onde \tilde{A}^\dagger e \tilde{A}^\downarrow são construídas a partir de \tilde{V}_m , similarmente como A^\dagger e A^\downarrow foram calculados a partir de V_m (ver Seção 1).

Sejam $\tilde{A}^\dagger = A^\dagger + \Delta_{A^\dagger}$, $\tilde{A}^\downarrow = A^\downarrow + \Delta_{A^\downarrow}$ e assumamos que $\text{posto}(\tilde{A}^\dagger) = r$. Temos então que

$$\begin{aligned} \tilde{T} - T &= \tilde{A}^{\dagger\dagger} \tilde{A}^\downarrow - T \\ &= \tilde{A}^{\dagger\dagger} (A^\downarrow + \Delta_{A^\downarrow}) - T \\ &= \tilde{A}^{\dagger\dagger} (A^\dagger T + \Delta_{A^\dagger}) - T \\ &= \tilde{A}^{\dagger\dagger} (\tilde{A}^\dagger T - \Delta_{A^\dagger} T + \Delta_{A^\downarrow}) - T \\ &= \tilde{A}^{\dagger\dagger} (-\Delta_{A^\dagger} T + \Delta_{A^\downarrow}). \end{aligned} \quad (3.2)$$

A última igualdade segue do fato que $\tilde{A}^{\dagger\dagger} \tilde{A}^\dagger = I$ pois, por hipótese, $\text{posto}(\tilde{A}^\dagger) = r$. Tomando norma em ambos os lados de (3.2), obtemos

$$\|\mathcal{E}\| \leq \|\tilde{A}^{\dagger\dagger}\| \|T\| (\|\Delta_{A^\dagger}\| + \|\Delta_{A^\downarrow}\|), \quad (3.3)$$

onde utilizamos a propriedade $\|T\| = \|V_s^* C V_s\| \geq 1$ [2].

O próximo passo na nossa análise é calcular estimativas para $\|\Delta_{A^\dagger}\|$ e $\|\Delta_{A^\downarrow}\|$. Para isto, enfatizamos que nenhuma estimativa interessante pode ser obtida a menos que V_m seja escolhido suficientemente próxima de \tilde{V}_m . Para realizarmos esta escolha, devemos buscar uma matrix unitária X , de ordem r , de modo que ela seja uma solução para o problema

$$\min \|VX - \tilde{V}_s\|_F,$$

com V sendo qualquer matriz de ordem $m \times r$ cujas colunas formam uma base ortonormal para o subespaço \mathcal{S}_m . Para resolver este problema, defina $G = \tilde{V}^*V$ e seja $G = \tilde{P}\tilde{\Sigma}\tilde{Q}^*$ sua decomposição em valores singulares. Prova-se que a matriz unitária que resolve este problema é $X = \tilde{Q}\tilde{P}^*$ (ver, por exemplo, Golub and Van Loan [5], página 601).

Usando a matrix X acima e escolhendo $V_m = V\tilde{Q}\tilde{P}^*$, temos

$$\begin{aligned}\|V_m - \tilde{V}_m\|^2 &= \|(V\tilde{Q}\tilde{P}^* - \tilde{V}_m)^*(V\tilde{Q}\tilde{P}^* - \tilde{V}_m)\| \\ &= \|2I - \tilde{P}\tilde{Q}^*V^*\tilde{V}_m - \tilde{V}_m^*V\tilde{Q}\tilde{P}^*\| \\ &= 2\|(I - \tilde{P}\tilde{\Sigma}\tilde{P}^*)\| \\ &= 2\|I - \tilde{\Sigma}\|\end{aligned}$$

Agora, como os cosenos dos ângulos canônicos entre os subespaços \mathcal{S}_M e $\tilde{\mathcal{S}}_M$ são os valores singulares de G (contidos em $\tilde{\Sigma}$) (ver, por exemplo, Teorema 2.4 em Stewart [12]), a relação acima implica que

$$\|\Delta_{V_m}\| = \|V_m - \tilde{V}_m\| = 2\text{sen } \frac{\Theta}{2}, \quad (3.4)$$

onde Θ é o maior ângulo canônico entre \mathcal{S}_m e $\tilde{\mathcal{S}}_m$. Agora observe que

$$\|\Delta_{A^\dagger}\| \leq \|\Delta_{V_m}\| \quad \text{e} \quad \|\Delta_{A^\dagger}\| \leq \|\Delta_{V_m}\|.$$

Substituindo estas desigualdades em (3.3) obtemos o seguinte teorema.

Teorema 3

$$\|\mathcal{E}\| \leq 4\|\tilde{A}^\dagger\| \|T\| \text{sen } \frac{\Theta}{2}. \quad (3.5)$$

Usando esta desigualdade em (3.1), obtemos a seguinte estimativa para o erro nos autovalores z_j .

Corolário 1

$$|z_j - \tilde{z}_j| \leq 4\kappa_2\|\tilde{A}^\dagger\| \|T\| \text{sen } \frac{\Theta}{2}, \quad j = 1 : r. \quad (3.6)$$

Para concluir nossa análise, vamos discutir o comportamento das quantidades $\|\tilde{A}^\dagger\|$ e $\|T\|$ que aparecem em (3.6). Primeiro, observe que

$$\|T\| = \|V_s^*CV_s\| \leq \sqrt{1 + \|\hat{f}_m\|^2},$$

visto que C é uma matriz companheira (ver (2.7)). Este resultado também pode ser deduzido calculando os valores singulares de T a partir de (2.9), maiores detalhes podem ser encontrados em Bazán [2]. Isto implica que para m suficientemente grande, $\|T\| \approx 1$, pois nessas condições $\|\hat{f}_m\|^2 \approx 0$. Agora, como as colunas de \tilde{A}^\dagger são quase ortonormais, pode-se provar que

$$\|\tilde{A}^\dagger\| = \frac{1}{1 - \|\hat{x}\|^2},$$

onde \tilde{x} é a última linha da matrix \tilde{V}_m . Desse fato, segue que $\|\tilde{x}\|^2$ não pode ser muito próximo de 1, e com isso $\|\tilde{A}^\dagger\|$ não pode ser muito maior do que 1. Desta discussão podemos deduzir que se m é suficientemente grande e os autovalores z_j não são extremamente próximos uns aos outros (para garantir um número de condição κ_2 moderado, por causa do Corolário 1), o erro máximo nos autovalores pode ser reescrito na forma

$$|z_j - \tilde{z}_j| = O(\text{sen}(\Theta)), \quad j = 1 : r. \quad (3.7)$$

Disso concluímos que se o erro $E = \tilde{A} - A$ nos dados de entrada, é suficientemente pequeno para garantir um ângulo Θ pequeno, os z_j são próximos do círculo unitário, mas não extremamente próximos uns aos outros, e m é suficientemente grande, então os autovalores z_j são relativamente *insensíveis* a pequenas perturbações nos dados de entrada.

4. Estimando $\text{sen}(\Theta)$

Visto que o erro máximo nos autovalores depende fortemente do ângulo entre os subespaços \mathcal{S}_m e $\tilde{\mathcal{S}}_m$ (ver (3.7)), vamos analisar um limitante para $\text{sen}(\Theta)$ com uma função dos autovalores z_j , e, em particular, da dimensão da matrix de dados \tilde{A} . A seguir, para simplicidade, assumiremos que a matrix de dados $\tilde{A} = A + E$ é de ordem $m \times m$ e denotada por \tilde{A}_m . Assumiremos também que o erro nos dados de entrada satisfazem $\|E\| \ll \sigma_r(A)$, onde $\sigma_r(A)$ denota o menor valor singular não nulo da matrix A_m . Aqui, vale a pena lembrar que a matrix A_m , que contém os dados exatos, é de posto incompleto e $\text{posto}(A_m) = r$. Nessas condições, pode-se provar que uma aproximação de primeira ordem para $\text{sen}(\Theta)$ é dada por [9]

$$\text{sen}(\Theta) \leq \frac{\|E\|}{\sigma_r(A_m)}. \quad (4.1)$$

Esta estimativa mostra que o tamanho de $\sigma_r(A_m)$ desempenha um fator fundamental para garantirmos $\tilde{\mathcal{S}}_m \approx \mathcal{S}_m$: quanto maior o tamanho de $\sigma_r(A_m)$ em relação a $\|E\|$, melhor será a aproximação. Por outro lado, visto que

$$\sigma_r(A_{m+1}) \geq \sigma_r(A_m), \quad \text{para } m \geq r, \quad (4.2)$$

o que segue da propriedade do entrelaçamento de valores singulares (ver Golub e Van Loan [5]), segue que devemos utilizar uma matrix de dados com a maior dimensão possível.

A seguir deduziremos uma estimativa (4.1) para o caso especial onde a matrix de dados é da forma

$$A_m = \begin{bmatrix} h_0 & h_1 & \cdots & h_{m-1} \\ h_1 & h_2 & \cdots & h_m \\ \vdots & \vdots & \cdots & \vdots \\ h_{m-1} & h_m & \cdots & h_{2m-1} \end{bmatrix}, \quad (4.3)$$

onde

$$h_k = b_1 z_1^k + b_2 z_2^k + \cdots + b_r z_r^k, \quad k \geq 0, \quad (4.4)$$

com $b_j \in \mathbb{C}$, $j = 1 : r$. Matrizes deste tipo são muito freqüentes em NMR e análise modal de estruturas mecânicas. Os parâmetros procurados nessas áreas são freqüências naturais, amortecimentos, fases, etc, e vêm descritos nos autovalores z_j , os quais são da forma

$$z_j = e^{(\alpha_j + i\omega_j)\Delta t}.$$

Estes z_j freqüentemente satisfazem $|z_j| \approx 1$, pois a constante Δt , chamada *taxa de amostragem*, usualmente é muito pequena.

É bem conhecido que a matriz de Hankel pode ser fatorada como [1]

$$A_m = W_m B W_m^T, \quad B = \text{diag}(b_1, \dots, b_r),$$

e que $\text{posto}(A_m) = r$, para todo $m \geq r$. Usando esta equação, segue imediatamente que

$$\sigma_r(A_m) \geq \beta \sigma_r(W_m)^2, \quad \text{onde } \beta = \min_j \{|b_j|\}, \quad j = 1 : r. \quad (4.5)$$

Seja

$$D^{-1} = \text{diag}(\|W_m e_1\|, \dots, \|W_m e_r\|),$$

e defina $\overline{W}_m = W_m D$. É claro que todas as colunas de \overline{W}_m têm norma um. Agora, como

$$\frac{\sigma_1(W_m)}{\sigma_r(W_m)} = k_2(W_m) = k_2(V_m^* W_m) \quad (4.6)$$

e

$$k_2(V_m^* W_m) = k_2[(V_m^* W_m D) D^{-1}] \leq k_2(V_m^* W_m D) k_2(D^{-1}),$$

levando em conta que todas as colunas de $V_m^* W_m D = V_m^* \overline{W}_m$ foram escaladas para ter norma um, e visto que essa matriz é uma matriz de autovetores da matriz T , pelo Teorema 1, usando um resultado bem conhecido sobre condicionamento de problemas de autovalor (ver, por exemplo, Teorema 2 em Demmel [4]), segue que

$$k_2(V_m^* W_m) \leq \sqrt{r} \kappa_2 \frac{\sqrt{1 + |z_{max}|^2 + \cdots + |z_{max}|^{2m-1}}}{\sqrt{1 + |z_{min}|^2 + \cdots + |z_{min}|^{2m-1}}},$$

onde κ_2 é o menor número de condição do problema de autovalor associado à matriz T em (1.1), $k_2(\cdot)$ é o número de condição da matriz (\cdot) na norma espectral, $|z_{max}| = \max_j |z_j|$, $j = 1 : r$, e $|z_{min}| = \min_j |z_j|$, $j = 1 : r$.

Agora, visto que $\sigma_1(W_m) \geq \|W_m e_j\|$, $j = 1 : r$, da relação acima e (4.6), vem

$$\sigma_r(W_m)^{-1} \leq \frac{\sqrt{r} \kappa_2}{\sqrt{1 + |z_{min}|^2 + \cdots + |z_{min}|^{2m-1}}}.$$

O seguinte teorema resulta de substituir esta desigualdade em (4.5).

Teorema 4

$$\sigma_r(A_m) \geq \beta \frac{1 + |z_{min}|^2 + \cdots + |z_{min}|^{2m-1}}{r \kappa_2^2}. \quad (4.7)$$

Deste resultado, concluímos que se $|z_j| \approx 1$ e m é suficientemente grande, então o ângulo θ entre os subespaços \mathcal{S}_m e $\tilde{\mathcal{S}}_m$ pode ser estimado aproximadamente como

$$\text{sen}(\Theta) \leq K \frac{\|E\|}{m}, \quad (4.8)$$

onde K é uma constante positiva relativamente pequena. Ou seja, para m suficientemente grande, $\text{sen}(\Theta) \approx 0$ e, com isso, o erro nos autovalores deve satisfazer $|z_j - \tilde{z}_j| \approx 0$, isto é, os autovalores devem ser pouco sensíveis a pequenas perturbações nos dados de entrada.

5. Resultados Numéricos

Para ilustrar a teoria desenvolvida nas seções anteriores, realizamos um conjunto de experimentos numéricos, com vários sinais sintetizados a partir de dados experimentais provenientes de análise modal e NMR. Apresentamos os resultados da análise de um sinal que descreve a função resposta ao impulso unitário correspondente a um ponto num protótipo de um prédio de três andares. O sinal foi coletado no Laboratório de Vibrações e Acústica da Universidade Federal de Santa Catarina (UFSC). Informações detalhadas podem ser encontradas em [1].

O sinal escolhido é um sinal real da forma (4.4) e, como isso, os autovalores z_j , os quais são mostrados na Tabela 1, aparecem em pares complexos conjugados. Assim, o posto da matriz de Hankel A_m , contendo os dados sem ruídos, é $\text{posto}(A_m) = 10$ para $m \geq 10$. A tabela também contém os “pesos” b_j , os módulos $|z_j|$, bem como as separações entre os autovalores.

j	z_j	$ z_j $	b_j	δ_j^2
1	$0.9699 \pm 0.2248i$	0.9956	$-0.1366 \pm 0.2490i$	0.0042
2	$0.9532 \pm 0.2931i$	0.9972	$0.7294 \pm 0.5743i$	0.0049
3	$0.9844 \pm 0.1619i$	0.9976	$-0.3162 \pm 0.0844i$	0.0031
4	$0.9921 \pm 0.1055i$	0.9977	$1.3284 \pm 0.6265i$	0.0023
5	$0.9972 \pm 0.0585i$	0.9989	$-0.0591 \pm 0.1958i$	0.0023

Tabela 1: Parâmetros de um sinal sintetizado de uma estrutura mecânica real.

O objetivo do experimento numérico é ilustrar o comportamento do número de condição κ_2 (que governa a sensibilidade do problema de autovalor) e do seno do ângulo entre os subespaços $\tilde{\mathcal{S}}_m$ e \mathcal{S}_m , ambos como funções da dimensão m . Para ilustrar a primeira parte, calculamos o limitante (2.14) para κ_2 , onde κ_F é computado utilizando a relação (2.2) (ver Smith [11]).

Os números de condição $|s_j|^{-1}$ foram obtidos utilizando os autovetores à direita e à esquerda da matriz $T = (V_m^* W_m) Z (V_m^* W_m)^{-1}$, onde as matrizes V_m são obtidas da SVD da matriz de dados A_m , para $m \geq 10$.

O comportamento do limitante (2.14) para κ_2 , que pode ser apreciado na Figura 2, mostra claramente que o problema de autovalor torna-se bem condicionado quando

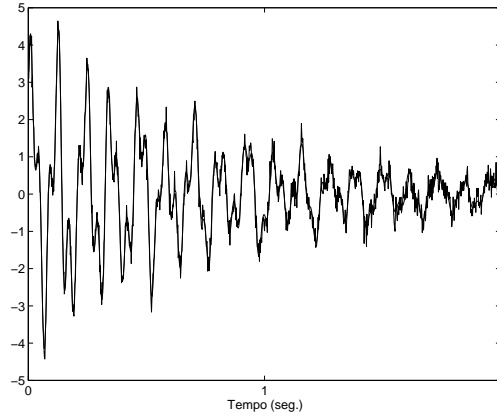


Figura 1: 1024 amostras de um sinal contaminado por ruído Gaussiano.

m é suficientemente grande. Este exemplo é ideal para mostrar que se os autovalores z_j não forem bem separados o problema pode ser muito mal-condicionado se a dimensão não for suficientemente grande (ver as separações δ_j na Tabela 1): para $m = 15$ o limitante é 1.3662×10^8 , enquanto que para $m = 56$, o limitante vale 538.3146. Aumentando a dimensão o limitante melhora mais ainda: se $m = 130$, o limitante é aproximadamente 1.8502. O valor $k_2(W_m)$ neste caso é 1.6563, o que indica que κ_2 pode ser ainda.

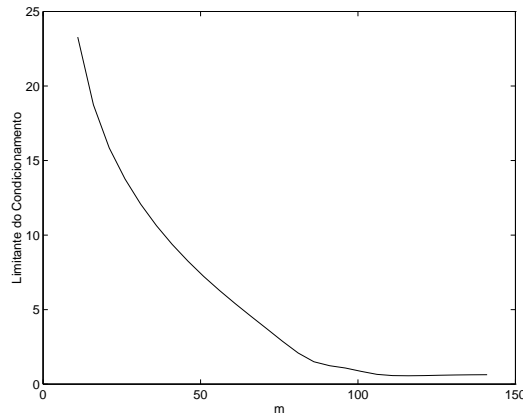


Figura 2: Limitante superior (2.14) para κ_2 como função de m .

Para ilustrar o papel que desempenha o seno do ângulo $\text{sen}(\Theta)$ entre os subespaços \mathcal{S}_m e $\tilde{\mathcal{S}}_m$ na estimativa do erro $\|\mathcal{E}\| = \|T - \tilde{T}\|$, primeiro perturbamos o sinal h_j com ruído Gaussiano e desvio padrão $\sigma = 0.25$, e a seguir calculamos a SVD das matrizes de Hankel \tilde{A}_m e A_m , onde \tilde{A}_m contém os dados “disponíveis”

$\tilde{h}_k = h_k + \epsilon_k$. O sinal “perturbado” \tilde{h} satisfaz $\|\tilde{h} - h\|/\|h\| \approx 0.10$ e pode ser apreciado na Figura 1. O seno do ângulo desejado foi calculado usando a definição (ver Golub e Van Voan [5])

$$\text{sen}(\Theta) = \|V_m V_m^* - \tilde{V}_m \tilde{V}_m^*\|,$$

onde as matrizes V_m e \tilde{V}_m foram obtidas dos vetores singulares associados com os 10 maiores valores singulares de A_m e \tilde{A}_m , respectivamente. Os resultados são apresentados na Figura 3, onde, para fins de comparação, também são incluídas as quantidades $2\text{sen}(\frac{\Theta}{2})$. A conclusão do experimento é que o seno do ângulo entre os subespaços envolvidos também melhora quando m é suficientemente grande, como previsto na teoria (ver estimativa (4.8)). Para ilustração,

$$\text{se } m = 140, \text{ sen}(\Theta) = 0.0488.$$

O erro nos autovalores para $m = 140$, apresentado na Tabela 2, ilustra o resultado teórico que os autovalores são pouco sensíveis a pequenas perturbações nos dados de entrada.

j	1	2	3	4	5
$ z_j - \tilde{z}_j \times 10^3$	0.1212	0.2415	0.6335	0.1620	0.3749

Tabela 2: Erro nos autovalores z_j para $m = 140$.

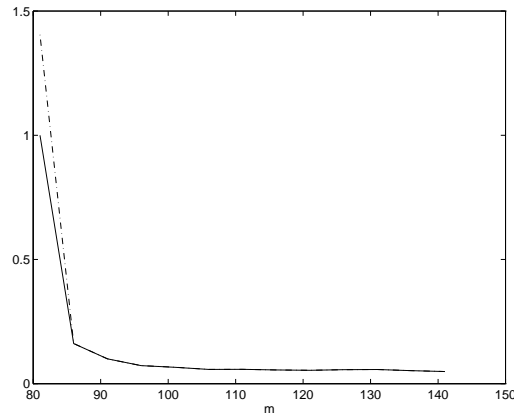


Figura 3: $\text{sen}(\Theta)$ (linha sólida) e $2\text{sen}(\frac{\Theta}{2})$ (linha pontilhada) como funções de m .

6. Conclusões

Apresentamos uma análise da sensibilidade de um problema de autovalor associado a métodos de recuperação de exponenciais, que mostra condições sob as quais as

exponenciais tornam-se insensíveis a pequenas perturbações nos dados de entrada. A conclusão foi que em todas as aplicações nas quais são analisadas 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. Esta conclusão decorreu de limitantes para o condicionamento do problema de autovalor associado, bem como da teoria clássica de perturbação de autovalores. A teoria desenvolvida foi ilustrada com um exemplo numérico onde foi analisado um sinal sintetizado de dados experimentais na área de análise modal de estruturas mecânicas.

Referências

- [1] F.S.V. Bazán, C. Bavastri, An Optimized Pseudo-Inverse Algorithm (OPIA) For Multi-Input Multi-Output Modal Parameter Identification, *Mechanical Systems and Signal Processings*, **10** (1996), 365-380.
- [2] F.S.V. Bazán, Conditioning of Rectangular Vandermonde Matrices with nodes in the Unit Disk, *SIAM J. Matrix Analysis and Applications*, **21**(2) (2000), 679-693.
- [3] F.S.V. Bazán, Sensitivity Eigenanalysis of Single Shift-Invariant Subspace-Based Methods, *Signal Processing*, **80** (2000), 89-100.
- [4] J. Demmel, The Condition Number of Equivalence Transformations que block Diagonalize Matrix Pencils, *SIAM J. Numer. Anal.*, **20**(3), June 1983.
- [5] G.H. Golub, C.F. Van Loan, "Matrix Computations", Johns Hopkins University Press, Baltimore, third edition, 1996.
- [6] P. De Groen, B. de Moor, The fit of a sum of exponentials to noisy data, *Journal of Computational and Applied Mathematics*, **20** (1987), 175-187.
- [7] Y. Hua, T.P. Sarkar, On SVD for estimating generalized eigenvalues of singular matrix pencil in noise, *IEEE Trans. Signal Processing*, **39** (1991), 892-900.
- [8] S. Kung, K.S. Arun, D. Bhaskar Rao, State-space and singular value decomposition-based approximation methods for the harmonical retrieval problem, *Opt. soc. Amer.*, **73** (1983), 1799-1811.
- [9] F. Li, Y. Lu, Bias analysis for Esprit-type algorithms, *IEEE Trans. Antenn. and Propag.*, **42**(3) (1994), 418-423.
- [10] A. Ruhe, On the Closeness of eigenvalues and singular values for almost normal matrices, *Linear algebra and Its applications*, **11** (1975), 87-94.
- [11] R.A. Smith, The Condition numbers of the matrix eigenvalue problem, *Numerische Mathematik*, **10** (1967), 232-240.

- [12] G.W. Stewart, Error and perturbation bounds for subspaces associated with certain eigenvalue problems, *SIAM Review*, **32** (1973), 727-764.
- [13] R. Vaccaro, A second perturbation expansion for the SVD, *SIAM Journal on Matrix Analysis and Applications*, **15**(2) (1994), 661-671.
- [14] A. Van Der Venn, E.F. Deprettere, A. Swindlehurst, Subspace-based signal analysis using singular value decomposition, *Proceedings of the IEEE*, **81**(9) (1993), 1277-1309.
- [15] S. Van Huffel, Enhanced resolution based on minimum variance estimation and exponential data modeling, *Signal Processing*, **33** (1993), 333-355.