

Algoritmos Rápidos para Estimar o Subespaço Sinal em Problemas de Posto Incompleto

M.I. CARDOSO¹, F.S.V. BAZÁN², Departamento de Matemática, Universidade Federal de Santa Catarina, Campus Universitário, 88010-970 Florianópolis, SC, Brasil.

Resumo. Apresentamos uma versão modificada do método de Lanczos para calcular o subespaço invariante associado aos maiores valores singulares de matrizes de Hankel de posto incompleto. O método utiliza recomeços implícitos para acelerar a convergência e efetua produtos matriz vezes vetor de maneira rápida. São apresentados resultados numéricos para ilustrar a potencialidade do método.

1. Introdução

A estimativa dos d -maiores valores singulares e correspondentes vetores singulares de uma matriz de Hankel H , com dimensão $M \times N$, e cujas entradas são amostras de sinais discretos, tem atraído muita atenção em diversas áreas, por exemplo, ressonância magnética nuclear, processamento de imagens, análise modal de estruturas mecânicas, etc. O subespaço gerado por esses vetores singulares é chamado *subespaço sinal*. Os sinais em estudo são descritos por:

$$h(t) = r_1 e^{s_1 t} + r_2 e^{s_2 t} + \dots + r_d e^{s_d t}, \quad t \geq 0,$$

com $r_i, s_i \in \mathbf{C}$, $\text{real}(s_i) < 0$, $i = 1, \dots, d$. As matrizes de Hankel são do tipo:

$$H = \begin{bmatrix} h_j & h_{j+1} & \dots & h_{j+N-2} & h_{j+N-1} \\ h_{j+1} & h_{j+2} & \dots & h_{j+N-1} & h_{j+N} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ h_{j+M-1} & h_{j+M} & \dots & h_{j+M+N-3} & h_{j+M+N-2} \end{bmatrix}_{M \times N}, \quad (1.1)$$

onde $h_j = h(j\Delta t)$, $j \geq 0$. Prova-se que $\text{posto}(H) = d$, [1]. Na prática, as amostras disponíveis são da forma $\tilde{h}_j = h_j + \epsilon_j$, onde os ϵ_j denotam o ruído nas amostras e assim, a matriz de Hankel $\tilde{H} = H + E$ se comporta como sendo de posto completo.

O algoritmo mais conhecido para calcular valores e vetores singulares é o algoritmo proposto por Golub e Reinsch [6]. Porém, como este algoritmo calcula todos

¹inez@mtm.ufsc.br

²fermin@mtm.ufsc.br

os valores singulares, ele pode ser muito demorado e, portanto, inconveniente em nosso caso, pois em geral $d \ll \min\{M, N\}$. Isso motiva o estudo de algoritmos iterativos para calcular decomposições espectrais parciais, as quais utilizam unicamente multiplicações matriz vezes vetor. Alguns algoritmos deste tipo são analisados em [5], onde os valores e vetores singulares de interesse são computados a partir do problema de autovalor associado a $\tilde{H}^* \tilde{H}$, onde $*$ denota a Hermitiana de \tilde{H} .

Neste trabalho, apresentamos algoritmos para calcular o subespaço sinal, os quais utilizam iterações de Lanczos e exploram multiplicações rápidas via FFT (Fast Fourier Transform) [5]. Introduzimos uma escolha inicial dos processos iterativos que garante convergência em no máximo $d + 1$ iterações quando o sinal for livre de ruído, mostrando resultados numéricos que ilustram a validade desta propriedade, ainda no caso de dados com ruído, desde que a separação entre $\tilde{\sigma}_d(\tilde{H})$ e $\tilde{\sigma}_{d+1}(\tilde{H})$ seja bem notória. Apresentamos, 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.

2. Multiplicação Rápida

Explorando a estrutura da matriz de Hankel e usando a Transformada Rápida de Fourier (*FFT*), é possível diminuir o número de operações necessárias para efetuar multiplicações do tipo matriz-vetor. Sendo $H \in \mathbf{C}^{M \times N}$, através do uso da *FFT*, as conhecidas MN multiplicações complexas necessárias para realizar essas multiplicações podem ser reduzidos para $O((M + N) \log_2(M + N))$. Para mostrar como isto pode ser feito, vamos introduzir uma matriz circulante, denotada por C ,

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

O seguinte teorema relaciona a transformada discreta de Fourier com matrizes circulantes, e é a base do processo de multiplicação rápida.

Teorema 1 *Se C é uma matriz circulante, então $C = F_N^* \Delta F_N$, onde F_N é matriz de Fourier*

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

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

Usando o teorema acima, obtemos uma forma de realizar o produto $y = Cx$ usando duas transformadas de Fourier e uma transformada inversa de Fourier, as quais podem ser feitas eficientemente via o algoritmo da FFT. O uso da FFT é ainda melhor se a dimensão do vetor for uma potência de 2. Neste caso, o produto matriz vezes vetor utiliza $O(N \log_2 N)$ multiplicações complexas. Observe ainda que, se for necessário calcularmos este produto para vários vetores, só precisamos calcular $F_N c$ uma vez, o que reduz o número de FFT a serem calculadas.

Considere agora a matriz circulante C definida por

$$C = \begin{bmatrix} h_M & h_{M+1} & h_{M+2} & \dots \\ h_{M-1} & h_M & h_{M+1} & \dots \\ \vdots & \vdots & \vdots & \dots \\ h_1 & h_2 & h_3 & \dots \\ 0 & h_1 & h_2 & \dots \\ \vdots & \vdots & \vdots & \dots \\ h_{M+N-1} & 0 & 0 & \dots \\ h_{M+N-2} & h_{M+N-1} & 0 & \dots \\ \vdots & \vdots & \vdots & \dots \\ h_{M+1} & h_{M+2} & h_{M+3} & \dots \end{bmatrix}$$

e observe que se introduzimos $J = [e_M \ e_{M-1} \ \dots \ e_1]$, então

$$JC = \begin{bmatrix} \times & \times \\ H & \times \end{bmatrix}.$$

Para obter o produto $y = Hx$, basta fazer $z = C\bar{x}$ usando a FFT de acordo com o teorema acima com $\bar{x} = [x \ 0]^T$ e, a seguir, tomar as M primeiras componentes de z , na ordem inversa, isto é, $y = [z_M \ z_{M-1} \ \dots \ z_1]^T$.

Algoritmo 1 *Multiplicação Rápida.*

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

onde (\cdot) denota a multiplicação ponto a ponto.

3. Método de Lanczos com Recomeços Implícitos

Existem várias técnicas para aproximar uma parte do espectro de uma matriz simétrica. Muitas delas, baseadas na projeção sobre subespaços de dimensão pequena, normalmente subespaços de Krylov, dentre as quais destaca-se o método de

Lanczos. A seguir, apresentamos uma modificação do método de Lanczos, conhecida como o método de Lanczos com recomeços implícitos, onde os produtos matriz vezes vetor são feitos através do algoritmo da multiplicação rápida descrito anteriormente.

O método clássico de Lanczos está baseado no fato de que toda matriz Hermitiana, $A \in \mathbf{C}^{N \times N}$, pode ser transformada numa matriz tridiagonal T , num número finito de passos, através de transformações unitárias [6]:

$$T = Q^* A Q. \quad (3.2)$$

Equacionando as j primeiras colunas de (3.2) obtemos

$$A Q_j = Q_j T_j + r_j e_j^T, \quad (3.3)$$

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

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

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 (3.3) pode ser rescrita como

$$r_j = \beta_j q_{j+1} = A q_j - \alpha_j q_j - \beta_{j-1} q_{j-1}. \quad (3.4)$$

A partir de (3.4) 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_j . Se τ_l, g_l , $1 \leq l \leq j$, são um par autovalor-autovetor de T_j , então os τ_l , chamados *valores de Ritz*, aproximam os maiores autovalores de A , enquanto que os vetores $y_l = Q_j g_l$, chamados *vetores de Ritz*, aproximam os correspondentes autovetores [6].

Algoritmo 2 *Método de Lanczos Básico.*

Seja $A = \tilde{H}^* \tilde{H}$. 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. $u_j := A q_j = H^*(H q_j)$
3. $r_j := u_j - \beta_{j-1} q_{j-1}$
4. $\alpha_j := q_j^* r_j$
5. $r_j := r_j - \alpha_j q_j$
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.

Uma deficiência do método é que, em aritmética finita, os vetores q_j perdem a ortogonalidade. Com isso, são produzidas cópias de vetores de Ritz já calculados, tornando o processo lento. Uma maneira de resolver este problema, é fazer uma reortogonalização completa. 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. 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. Uma maneira de não deixar o número de vetores de Lanczos ficar muito grande, é recomençar o processo após um número fixo de iterações, com um vetor inicial escolhido de maneira a aproveitar a estrutura da matriz A . Neste caso, a reortogonalização é possível, pois somente um número fixo de vetores de Lanczos é mantido na memória.

Em problemas onde não existe uma boa separação 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 recomençar as iterações após um número finito de passos, mantendo a quantidade de vetores a ser armazenada fixa e a reortogonalização é feita somente quando for necessário. Calvetti, Reichel e Sorensen desenvolveram um método de Lanczos com recomeços implícitos (IRL) [4], o qual é uma adaptação do método de Arnoldi com recomeços [8] para o caso Hermitiano. Este método força o vetor inicial a estar num subespaço invariante da matriz através de repetidas filtragens do vetor inicial e recomençando as iterações implicitamente.

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 (3.4) 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 2 *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, passo d da fatoração de Lanczos de A . Então, $r_d = 0$ se, e somente se, $q_1 = V_d x$, onde*

$$AV_d = V_d \Lambda_d, \quad (3.5)$$

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

Demonstração. (\Rightarrow) Suponha que $r_d = 0$. Então, $AQ_d = Q_d T_d$. 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 $AV_d = V_d \Lambda_d$, $\text{rank}(V_d) = d$ 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 ,

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

Consequentemente, $\dim(\mathcal{K}^{d+1}(A, q_1)) \leq \text{posto}(V_d) = d$. Como T_d é não reduzida, usando uma pequena modificação do Teorema 7.4.3 de Golub e Van Loan [6], temos que $\dim(\mathcal{K}^{d+1}(A, q_1)) = d$ e, então, $r_d = 0$. \square

Fixando o número de passos Lanczos em $d + p$, onde p em geral não é muito maior do que d , sendo menor em alguns casos, obtem-se:

$$AQ_{d+p} = Q_{d+p}T_{d+p} + \beta_{d+p}Q_{d+p}e_{d+p}^T. \quad (3.6)$$

O *IRL* aplica o algoritmo *QR* com shifts a esta decomposição truncada de A da seguinte maneira: Seja μ_1 um shift, calcule a decomposição *QR* de $T_{d+p} - \mu_1 I = X_1 R_1$, onde $X_1, R_1 \in \mathbf{C}^{(d+p) \times (d+p)}$, $X_1^* X_1 = I_{d+p}$ e R_1 é triangular superior. Partindo agora de $T_{d+p}^+ = X_1^* T_{d+p} X_1$, aplique novamente o algoritmo *QR* com shift μ_2 . Desta forma, após p shifts obtém-se $AQ_{d+p}^+ = Q_{d+p}^+ T_{d+p}^+ + \beta_{d+p} q_{d+p+1} e_{d+p}^T \hat{X}$, onde $Q_{d+p}^+ = Q_{d+p} \hat{X}$, $T_{d+p}^+ = \hat{X}^* T_{d+p} \hat{X}$ e $\hat{X} = X_1 X_2 \cdots X_p$. Analisando as d primeiras colunas desta decomposição obtém-se: $AQ_d^+ = Q_d^+ T_d^+ + \beta_d^+ q_{d+1}^+ e_d^T$. Partindo agora desta nova decomposição de Lanczos e calculando p passos de Lanczos, voltamos a mesma situação da equação (3.6). Portanto, o *IRL* consiste em aplicar alternadamente p passos Lanczos e p 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 [4], 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 [4], estes shifts são chamados de *shifts exatos*.

Lema 1 *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*

$$T_{d+p}^+ = \hat{X}^* T_{d+p} \hat{X}, \quad (3.7)$$

onde $\hat{X} = X_1 X_2 \cdots X_p$ 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

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

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 $\eta_j \in \mathbf{C}$ 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 $1 \leq j \leq d$.

A demonstração do Lema 1 pode ser encontrada em [5]. O lema mostra que a escolha dos shifts exatos, é matematicamente equivalente a recomeçar as iterações com

um vetor inicial, que é 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 lema, porém, não apresenta nenhuma informação sobre a razão de convergência. Em [4], encontra-se uma discussão sobre outro tipo de shifts. Antes de discutir a convergência do método apresentamos o algoritmo IRL para aproximar os d maiores autovalores. Assume-se que $\tau_1 \geq \tau_2 \geq \dots \geq \tau_{d+p}$.

Algoritmo 3 *Método de Lanczos Recomeços Implícitos e Shifts Exatos (IRL-ES).* Seja $A = \tilde{H}^* \tilde{H}$. Dados $r_0 \neq 0$, $\beta_0 = \|r_0\|$:

1. Faça $d+p$ passos usando o Algoritmo 2 (reortogonalizando quando 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 (3.7), e volte para o passo 2 acima.

Nossa discussão sobre a convergência dos valores de Ritz está baseada no trabalho de Sorensen [10], onde é mostrado que à medida em que ocorrem as iterações, os autovalores de T_d convergem para certos limites. Partindo disto, mostramos que estes limites são os autovalores de A e, ainda mais, que são os d maiores. Seja l o número da iteração. Assim, $q_1^{(l)}$ 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 $T_{d+p}^{(l)} = \begin{bmatrix} T_d^{(l)} & \beta_d^{(l)} e_d e_1^T \\ \beta_d^{(l)} e_1 e_d^T & \tilde{T}^{(l)} \end{bmatrix}$ com autovalores $\tau_{1,l+1} > \dots > \tau_{d,l+1} > \mu_{1,l+1} > \dots > \mu_{p,l+1}$.

Aplicando p shifts, temos

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

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

Para mostrar que o limite para o qual as sequências $\{\tau_j^{(l)}\}_{l=1}^\infty$ convergem são autovalores de A , vamos precisar de um lema auxiliar, o qual é enunciado a seguir:

Lema 2 *Seja $M = \begin{bmatrix} T & \beta e_d \\ \beta e_d^T & \alpha \end{bmatrix}$ uma matriz simétrica tridiagonal. Então, as raízes da equação*

$$\beta^2 e_d^T (T - \lambda I)^{-1} e_d = \alpha - \lambda \tag{3.8}$$

são autovalores de M .

Demonstração. A demonstração segue do cálculo do polinômio característico de M , ou veja [9]. \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 (3.8) tem-se:

$$\beta^2 \sum_{j=1}^d \frac{\eta_j^2}{(\tau_j - \lambda)} = \alpha - \lambda, \quad (3.9)$$

onde $(\eta_1, \eta_2, \dots, \eta_d) = e_d^T G$ e $\Lambda = \text{diag}(\tau_1, \dots, \tau_d)$. Da equação (3.9), pode-se concluir que nenhum autovalor de M é igual a τ_i , $1 \leq i \leq d$. Usando a propriedade do interlacing, veja Ggolub [6, pag. 396], tem-se:

$$\hat{\mu} < \tau_d < \tilde{\tau}_d < \tau_{d-1} < \dots < \tau_1 < \tilde{\tau}_1 \quad (3.10)$$

onde $\{\tilde{\tau}_1, \dots, \tilde{\tau}_d, \hat{\mu}\}$ são autovalores de M . Se os elementos da subdiagonal de T são todos não nulos, então os τ_j são distintos (veja [9], capítulo 7) e nenhum dos η_j é zero.

Lema 3 *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 $\tau_{j,l} \rightarrow \tau_j$. Então,*

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

e, como uma consequência,

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

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

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

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

Do Lema 2, segue que os d maiores autovalores $\hat{\tau}_{j,l}$ de $M^{(l)}$ satisfazem

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

onde a última desigualdade surge de $p-1$ aplicações adicionais do Lema 2. Além disso, através de manipulações algébricas da equação (3.9), obtém-se:

$$(\beta\eta_j)^2 = (\tau_j - \lambda) \left[\frac{(\alpha - \lambda) - \beta^2 \sum_{i=j+1}^d \frac{\eta_i^2}{(\tau_i - \lambda)}}{1 + \sum_{i=1}^{j-1} \frac{\eta_i^2 (\tau_j - \lambda)}{\eta_j^2 (\tau_i - \lambda)}} \right], \quad (3.11)$$

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

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

Observe que, como $\hat{\tau}_{jl} > \tau_{jl} > \tau_{il}$ para $i = 1, 2, \dots, j-1$, então o denominador na equação (3.11) 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

Como os τ_{jl} convergem para os autovalores de A , falta 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 3 *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 assumamos 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 em uma limitação inferior uniforme de $|\eta_j^{(l)}|$ e então, $\beta_d^{(l)} \rightarrow 0$. Estas observações seguem do Lema 2.

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, portanto

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

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 $j = 1, \dots, d$, temos 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 $\lambda_{j-1} = \tau_1 < \lambda_1$. Então, o coeficiente da expansão é

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

onde v_{j-1} 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,

$$(\gamma_{j-1}^{(l)})^2 = \frac{(\gamma_{j-1} \Psi_l(\tau_1) / \Psi_l(\lambda_1))^2}{\gamma_1^2 + \sum_{i=2}^n \gamma_i^2 \Psi_l^2(\lambda_i) / \Psi_l^2(\lambda_1)} \leq \left(\frac{\gamma_{j-1} \Psi_l(\tau_1)}{\gamma_1 \Psi_l(\lambda_1)} \right)^2,$$

onde os γ_j são os coeficientes da expansão de q_1 . Agora, as raízes μ_{il} dos polinômios filtros satisfazem $\lambda_n \leq \mu_{il} < \tau_{1l} \leq \tau_1 < \lambda_1$, 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. Concluímos, 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, analisamos a convergência do IRL, porém não foi dito nada a respeito da taxa de convergência. A seguir, apresentamos alguns resultados que descrevem condições sob os quais obtém-se convergência rápida.

Lema 4 *Seja $A = B + \tau I$ uma matriz Hermitiana $M \times M$, onde B é Hermitiana semi-positiva definida, $\text{posto}(B) = d$ e $\tau > 0$. Se, os d autovetores associados aos d maiores autovalores são distintos, e se q_0 não é ortogonal a nenhum destes d autovetores e 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: $\mathcal{K}^m(A, q_0) = \mathcal{K}^m(A - \tau I, q_0) = \mathcal{K}^m(B, q_0)$ e $\dim(\mathcal{K}^m(B, q_0)) = d + 1$, para todo $m > d$ [12]. \square

Outro resultado importante em relação a convergência do método de Lanczos aplicado ao cálculo dos d maiores valores singulares, é que a taxa de convergência depende da separação entre $\tilde{\sigma}_d$ e $\tilde{\sigma}_{d+1}$ (ver Teorema 6.4.1 em [7]): quanto maior a separação, maior é a taxa de convergência. Considerando que, em nosso caso, $\tilde{H} = H + E$ e $\text{posto}(H) = d$, e usando teoria de perturbação de valores singulares, segue que $\tilde{\sigma}_{d+1} \leq \|E\|$ e que uma melhor separação entre $\tilde{\sigma}_d$ e $\tilde{\sigma}_{d+1}$ ocorre quando $\|E\| < \sigma_d$ [6]. Isto mostra que o tamanho do σ_d , 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 σ_d no caso em que são utilizadas matrizes quadradas.

Teorema 4 *Seja H a matriz de Hankel $M \times M$, descrita em (1.1) com $j = 0$. Defina $z_l = e^{s_l \Delta t}$ e $\delta = \min_{l \neq k} |z_l - z_k|$, $1 \leq l, k \leq d$. Então, um limitante inferior para σ_d satisfaz*

$$\sigma_d \geq K \left[1 + \frac{D_M^2}{(d-1)\delta^2} \right]^{1-d} (1 - e^{-2\alpha \Delta t M}), \quad (3.12)$$

onde

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

com D_M^2 uma constante que depende do tamanho dos z_l e M .

Demonstração. Primeiro observamos que a matriz de Hankel pode ser rescrita como

$$H = WRW^T,$$

onde W é uma matriz de Vandermonde $M \times d$, com entradas $W(i, j) = z_j^{i-1}$ e $R = \text{diag}(r_1, \dots, r_d)$. A seguir, usando propriedades de valores singulares, segue que

$$\sigma_d \geq \sigma_d(R)\sigma_d(W)^2. \quad (3.13)$$

A demonstração termina aplicando o Teorema 3.2 de Bazan [3] □.

Em [3], pode ser encontrada uma expressão para D_M , onde é mostrado que $D_M^2 \approx 0$ quando M é suficientemente grande e $|z_l| \approx 1$, o que é muito freqüente em aplicações envolvendo processamento de sinais. Em [2], é analisado o comportamento do limitante dado em 3.13 para matrizes de Hankel retangulares. O resultado dessa análise é que a estimativa será melhor quando a matriz de Hankel for quadrada.

O teorema sugere ainda que σ_d 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, devemos utilizar matrizes de Hankel quadradas com a maior dimensão possível a fim de que σ_d possa vencer o tamanho da perturbação $\|E\|$, possibilitando, assim, uma boa separação entre $\tilde{\sigma}_d$ e $\tilde{\sigma}_{d+1}$.

Dos resultados obtidos do Lema 4 e Teorema 4, deduzimos 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 $\|E\| < \sigma_d$, 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 apresentaremos um exemplo numérico que mostra que se $\|E\| \ll \sigma_d$, então o método de Lanczos converge em pouco mais de d iterações, desde que o vetor inicial seja escolhido adequadamente.

Finalizando esta seção, apresentamos uma escolha de vetor inicial para o método de Lanczos, a qual mostra resultados melhores em relação a velocidade de convergência comparado com uma escolha aleatória. Propomos como vetor inicial o vetor $q_0 = \tilde{H}^* \tilde{b}$, onde $\tilde{b} = [\tilde{h}_{j-1}, \tilde{h}_j, \dots, \tilde{h}_{j+M}]^*$. Para explicar o motivo desta 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 $\|\tilde{b} - b\| \ll \|b\|$, então as componentes mais fortes do vetor q_0 estão associadas com os d primeiros vetores singulares, os quais queremos aproximar.

4. Experimentos Numéricos

Para observar o desempenho do IRL em relação a escolha do vetor inicial q_0 , a escolha de p e do uso da multiplicação rápida, apresentamos resultados de alguns experimentos numéricos realizados com um sinal proveniente de simulações em Ressonância Magnética Nuclear [11]. O método foi programado em Matlab, e para medir o tempo computacional gasto, foi usada a função *flops* do Matlab. O sinal foi construído através da seguinte modelagem:

$$\tilde{h}_j = \sum_{k=1}^d a_k e^{i\phi_k} e^{(\alpha_k + i\omega_k)j\Delta t} + \epsilon_j, \quad j = 1, 2, \dots, 512, \quad (4.14)$$

l	1	2	3	4	5	6	7	8	9	10	11
a_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 1: Parâmetros do sinal, onde a_l é dada em unidades arbitrárias, α_l e ω_l em Hz e ϕ_l em graus.

onde os parâmetros a_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 a_l , α_l , ω_l e ϕ_l são dados na Tabela 1.

A partir do sinal, foram geradas matrizes de Hankel de ordem 256×256 , de acordo com (1.1). Na Figura 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.

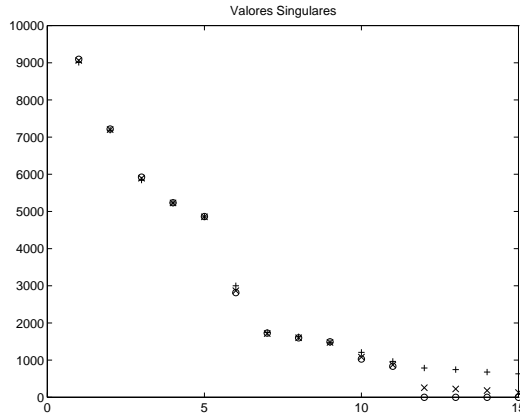


Figura 1: Em 'o' os 11 valores singulares não nulos de H , em 'x' e '+' os 15 maiores valores singulares de \tilde{H} para $\sigma = 5$ e $\sigma = 15$, respectivamente.

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

Para observar a performance do método em relação a 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 2, onde razão1 denota o quociente entre o número de flops gastos quando o método é 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 $q_0 = \tilde{H}^* \tilde{b}$. A tabela apresenta também, na terceira coluna, a razão entre o número de flops do IRL-ES sem o uso da multiplicação rápida e com o uso da multiplicação rápida. Foram realizados os mesmos testes para outros sinais e os resultados obtidos foram similares.

5. Conclusões

A transformação do problema de valores singulares num problema de autovalores nos permitiu o cálculo parcial dos valores singulares, o que é muito relevante no nosso caso, 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 nos 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, a qual aumenta a taxa de convergência em relação à escolha aleatória.

Referências

- [1] F.S.V. Bazán and 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 and Ph.L. Toint, On the 2-Condition Number of Infinite Hankel Matrices of Finite Rank, “technical report 10”, Department of Mathematics, FUNDP, Namur, Belgium, February 1998.
- [3] F.S.V. Bazán, Conditioning of Rectangular Vandermonde Matrices with nodes in the Unit Disk, *SIAM Journal on Matrix Analysis and Applications*, a aparecer.
- [4] D. Calvetti, L. Reichel and D.C. Sorensen, An Implicitly Restarted Method for Large Symmetric Eigenvalue Problems, *Electronic Transactions on Numerical Analysis*, **2** (1994), 1-21.
- [5] E. Lundström, Singular Value Computations for Toeplitz Matrices, Lic. Thesis, Linköping University, ISBN 91-7871-719-1, May 1996.
- [6] G.H. Golub and C.F. Van Loan, “Matrix Computations”, Johns Hopkins University Press, Baltimore, third edition, 1996.
- [7] P.C. Hansen, “Rank-Deficient and Discrete Ill-Posed Problems”, SIAM, Philadelphia, 1997.

- [8] R.B. Lehoucq, Analysis and Implementation of an Implicitly Restarted Arnoldi Iteration. “PhD Thesis”, Rice University, Houston, Texas, 1995.
- [9] B.N. Parlett, “The Symmetric Eigenvalue Problem”, SIAM, Philadelphia, 1997.
- [10] D.C. Sorensen, Implicit Application of Polynomial Filters in a k-step Arnoldi Method, *SIAM Journal on Matrix Analysis and Applications*, **13** (1992), 357-385.
- [11] 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** (1997), 400-409.
- [12] G. Xu and T. Kailath, Fast Estimation of Principal Eigenspace using Lanczos Algorithm, *SIAM Journal on Matrix Analysis and Applications*, **15** (1994), 974-994.

