

DESCRIÇÃO UNIFICADA DOS MÉTODOS MODERNOS PARA ESTIMAÇÃO DOA EM ARRANJO DE SENSORES

Ricardo F. Colares, Carlos A. Alves e Amauri Lopes

Resumo - Este trabalho apresenta os principais métodos para a estimação do parâmetro DOA que surgiram nas últimas décadas. Esses métodos, em geral, baseiam-se em conceitos e propriedades distintas, dificultando uma comparação direta entre os mesmos. A contribuição deste trabalho consiste em apresentar tais métodos através de uma formulação e notação unificada, a qual permite relacioná-los, evidenciando suas características de desempenho e esforço computacional. Além dessa visão geral, tal descrição unificada abre caminhos para o desenvolvimento de novos métodos.

Palavras-chave: Arranjo de sensores, Estimadores para DOA, Máxima verossimilhança, Predição linear, Ortogonalidade entre subespaços .

Abstract - This work presents the most important direction of arrival (DOA) estimation methods proposed in the last two decades. In general, those methods are based on different concepts and properties rendering the comparisons among them a hard task. The aim of this paper is to present those methods in a unified formulation and notation to allow a direct comparison of their performance and computational effort. In addition, the paper also provides insights into the design of new methods.

Keywords: Sensor Arrays, DOA Estimators, Maximum Likelihood, Linear Prediction, Orthogonality of subspaces .

1. INTRODUÇÃO

O processamento do sinal aplicado em um arranjo de sensores tem por objetivo extrair parâmetros úteis relacionados às ondas incidentes no arranjo. Em particular, o parâmetro de nosso interesse determina a direção de chegada (DOA) ou ângulo de chegada (AOA) dessas ondas, e é geralmente dado pelo valor de um ângulo θ que define o azimute [1], [2], [3], [4].

A estimação do parâmetro DOA foi inicialmente utilizada para determinar a posição de objetos em RADAR (Radio Detection And Ranging) e SONAR (Sound Navigation Ranging). Mais recentemente, tem sido aplicada em sistemas de comunicações móveis [4], [5], [6].

As técnicas convencionais baseadas em espectros utilizam a medida da potência de saída do arranjo em função do

ângulo de direção de chegada para a obtenção do parâmetro DOA [1], [2], [3]. Entretanto, para sinais correlacionados ou mesmo coerentes, o desempenho desses métodos torna-se insuficiente [4], [7]. Os métodos paramétricos modernos, por sua vez, não fazem uso de uma função espectral, mas sim de um modelo para o sinal baseado nos parâmetros de interesse. Neste caso, o parâmetro DOA é estimado do modelo que mais se ajusta ao sinal recebido. Embora tais métodos sejam eficientes e robustos, exigem, em geral, grande esforço computacional.

Dentre os métodos paramétricos, o procedimento de estimação baseado no critério da *máxima-verossimilhança* (ML) apresenta uma grande capacidade de resolução e nenhuma dificuldade particular na aplicação em arranjos com qualquer geometria ou para o caso multidimensional. Porém, tal procedimento exige um esforço computacional proibitivo [4], [8], [9]. Assim, novos métodos surgiram para superar esta carga computacional e, ao mesmo tempo, manter níveis próximos de desempenho.

Dentre esses novos procedimentos destacam-se, pela simplicidade computacional, aqueles baseados na predição linear. Os métodos FBLP (Forward-Backward Linear Prediction) Modificado [10] e WTLS-LP (Whitened Total Least Squares-Linear Prediction) são exemplos de sucesso do uso da predição linear. Tais procedimentos realizam a estimação do parâmetro DOA através de um mapeamento adequado entre ângulo espacial associado ao parâmetro DOA e ângulo elétrico associado a uma frequência. As desvantagens desses procedimentos, quando comparados com aqueles baseados no critério ML, consistem na limitação a aplicações com arranjos lineares e uniformes e na baixa capacidade de resolução quando aplicados em sinais fortemente correlacionados e em baixas relações sinal-ruído [10], [11], [12] e [13].

Os procedimentos baseados em subespaços consistem em uma outra alternativa. Esses realizam a estimação do parâmetro DOA através da decomposição do sinal observado em subespaços de sinal e de ruído, bem como da propriedade de ortogonalidade entre os mesmos [14], [15], [16] e [17]. Os métodos MUSIC (Multiple Signal Classification) e NORMA-MÍNIMA fazem uso desses procedimentos. Uma outra variante para os procedimentos baseados em subespaços ("Subspace Fitting-SF") se destaca como competidor do critério ML por apresentar uma capacidade de resolução equivalente [18]. Esse critério consiste em obter uma aproximação para um conjunto de dados ruidosos cujo subespaço de sinal apresente a menor "distância" possível em relação ao subespaço de sinal desejado [4], [19] e [20]. O método ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) pertence a essa classe de procedimentos, na qual destacamos o método WSF (Weighted Subspace Fitting) como a melhor alternativa.

Ricardo F. Colares está na Universidade de Fortaleza (UNIFOR), Carlos A. Alves está na Universidade Estadual Paulista/Ilha Solteira e Amauri Lopes está na Universidade Estadual de Campinas. E-mails: ricardocolares@unifor.br, caa@dee.feis.unesp.br, amauri@decom.fee.unicamp.br.

Editor de Área responsável: Rui Seara. Artigo submetido em 08/2000, revisado em 03/2001 e aceito em 08/2001.

Uma outra forma de contornar o esforço computacional exigido pelo critério ML surge do uso de alternativas para a sua implementação, empregando-se algoritmos numéricos mais eficientes que propiciam desempenhos equivalentes [21] e [22]. Porém, tais algoritmos são geralmente iterativos e alguns sofrem de problemas de convergência [23]. Como exemplos destacamos os métodos IQML (Iterative Quadratic Maximum Likelihood) e MODE (Method Of Direction Estimation).

Nem todos os métodos apresentados aqui são consagrados para a aplicação em estimação do parâmetro DOA. No entanto, além de serem factíveis para tal aplicação, nos permitem uma visão mais ampla das várias alternativas.

Todos esses procedimentos de estimação obtêm suas estimativas através da minimização de uma função custo adequada. Algumas destas funções apresentam uma dependência do tipo exponencial em relação aos parâmetros DOA, onerando sobremaneira o processo de minimização. Entretanto, esta dependência pode ser alterada através de uma reparametrização adequada do problema original. Tal reparametrização possui um papel fundamental na descrição dos métodos de forma unificada [24].

Conforme evidenciado, embora as descrições de todos esses métodos estejam disponíveis na literatura, não se encontram em uma formulação unificada. As comparações entre os diversos métodos existentes na literatura, restringem-se a análises isoladas de desempenhos obtidos via simulação. A análise unificada e a classificação propostas aqui possibilitam uma melhor compreensão das características e do desempenho de cada um dos métodos, explicitam as relações entre os mesmos e evidenciam alternativas inéditas [25], [26].

A comparação entre os desempenhos dos métodos apresentados será realizada apenas de forma qualitativa, a partir de suas formulações unificadas. Uma comparação quantitativa, através de resultados de simulação computacional, está sob investigação e será objeto de trabalho futuro.

Apresentamos agora a organização do artigo. A seção 2 apresenta o modelo de sinal para um arranjo linear uniforme, a descrição dos subespaços de sinal e de ruído associados à matriz de correlação do sinal na saída dos sensores e um resumo da notação e siglas utilizadas no artigo. Este material será empregado em todas as demais seções.

O estimador de máxima verossimilhança aplicado ao problema DOA é abordado na seção 3. Deduzimos o critério ML determinístico e sua expressão em função dos subespaços da matriz de correlação. Deduzimos ainda sua modificação através de um processo de mapeamento dos parâmetros a serem estimados, denominado de reparametrização.

A seção 4 apresenta os principais métodos de estimação DOA baseados no critério ML. São utilizadas as técnicas de reparametrização e de eliminação do subespaço de ruído.

As propriedades da predição linear que permitem seu emprego na estimação DOA são tratadas na seção 5. O conceito de subespaços também é aplicado nesta seção.

Os principais métodos de estimação DOA baseados na predição linear são apresentados na seção 6. Resultam da aplicação de técnicas como a predição forward-backward, dos critérios de otimização LS (Least Squares) e TLS (Total Least Squares) e da restrição a subespaço. Nesta seção apre-

sentamos ainda a comparação destes métodos com aqueles baseados no critério ML da seção 4.

A terceira categoria de estimadores DOA, a estimação totalmente baseada em subespaços, é abordada na seção 7. Ao contrário dos casos anteriores, optamos por dividir esta categoria em três subcategorias e por apresentar os respectivos métodos nas respectivas subseções. Resultaram então o critério da ortogonalidade baseado na matriz de correlação, o critério da ortogonalidade entre subespaços e o critério de “distância” entre subespaços. Em cada caso são apresentadas comparações com os métodos anteriores.

Por fim, a seção 8 apresenta as conclusões gerais do trabalho.

2. ASPECTOS GERAIS

Introduziremos agora a notação e algumas formulações matemáticas que servirão de base para a apresentação dos diversos métodos existentes para a estimação do parâmetro DOA. A apresentação prematura das expressões e resultados a seguir visa evitar a repetição daquelas comuns aos métodos a serem apresentados, bem como facilitar uma comparação futura entre os mesmos.

No final desta seção apresentamos um resumo da notação e das siglas empregadas em todo o artigo.

2.1 MODELO DO SINAL

O vetor das amostras do sinal na saída de um arranjo N -dimensional com geometria arbitrária, exposto a $M < N$ ondas distintas incidentes, tomadas nos instantes $t = k\Delta t$, considerando a hipótese de banda estreita, pode ser escrito como [27]

$$\mathbf{x}(k) = \sum_{m=1}^M \mathbf{a}(\theta_m) s_m(k), \quad \in \mathcal{C}^{N \times 1} \quad (1)$$

onde $s_m(k)$ representa a amplitude da m 'ésima forma de onda incidente. O vetor $\mathbf{a}(\theta_m) = [a_1(\theta_m) \cdots a_N(\theta_m)]^T$ é denominado de vetor da direção angular θ_m , que representa o parâmetro DOA para a m 'ésima onda e $\mathbf{x}(k) = [x_0(k) \cdots x_{N-1}(k)]^T$. $[\cdot]^T$ denota transposto.

No caso de um arranjo linear com elementos uniformemente espaçados, e assumindo que todos os elementos possuem a mesma diretividade, teremos

$$\mathbf{a}(\theta_m) = [1 \quad e^{j\phi_m} \quad \dots \quad e^{j(N-1)\phi_m}]^T, \quad (2)$$

para um sinal composto por exponenciais complexas não-amortecidas. Em (2) $\phi_m = (2\pi/\lambda)d \cos(\theta_m)$, onde d representa a distância entre cada elemento do arranjo e λ o comprimento de onda.

A Figura 1 ilustra um arranjo linear vertical composto por N sensores espaçados por uma distância d , exposto a m 'ésima fonte de sinal com direção θ_m em relação ao eixo vertical.

De uma forma mais compacta e geral, podemos escrever a expressão de saída (1) como

$$\mathbf{y}(k) = \mathbf{A}(\Phi)\mathbf{s}(k) + \mathbf{n}(k) \quad (3)$$

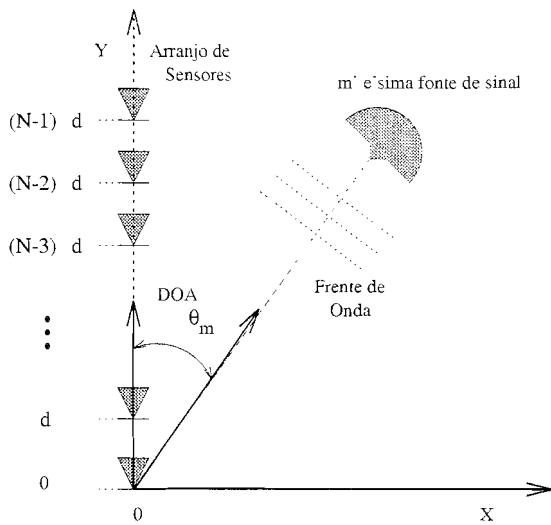


Figura 1. Arranjo linear uniforme

onde $\mathbf{y}(k) = [y_0(k) \cdots y_{N-1}(k)]^T$, $\mathbf{A}(\Phi) = [\mathbf{a}(\phi_1) \cdots \mathbf{a}(\phi_M)] \in \mathbb{C}^{N \times M}$ é uma matriz de Vandermonde, $\Phi = [\phi_1 \cdots \phi_M]^T$, $\mathbf{s}(k) = [s_1(k) \cdots s_M(k)]^T$ e $\mathbf{n}(k)$ representa o ruído aditivo.

Dois diferentes tipos de sinais podem surgir do modelo em (3) [4], [28]. O primeiro deles, denominado de modelo condicional ou determinístico, assume que $\mathbf{s}(k)$ não é aleatório. Isto é, para K conjuntos de amostras coletadas nos instantes $k = 1, 2, \dots, K$, a seqüência de K vetores $\{\mathbf{s}(k)\}_{k=1}^K$ não é alterada em todas as realizações da seqüência de vetores $\{\mathbf{x}(k)\}_{k=1}^K$, e a seqüência de vetores $\{\mathbf{n}(k)\}_{k=1}^K$ varia a cada realização. No segundo modelo, denominado de incondicional ou estocástico, $\mathbf{s}(k)$ é também considerado aleatório.

A distinção entre esses dois modelos será aprofundada na seção a seguir.

2.2 CORRELAÇÃO E SUBESPAÇOS

A matriz de correlação para o sinal na saída do arranjo é calculada a partir de (3) como

$$\begin{aligned} \mathbf{R}(k, i) &= E \{ \mathbf{y}(k) \mathbf{y}^H(i) \} \\ &= \mathbf{A} E \{ \mathbf{s}(k) \mathbf{s}^H(i) \} \mathbf{A}^H + E \{ \mathbf{n}(k) \mathbf{n}^H(i) \}. \end{aligned} \quad (4)$$

com $k, i = 1, 2, \dots, K$, onde o sobrescrito H denota conjugado-transposto, $E \{ \cdot \}$ simboliza a esperança estatística, $E \{ \mathbf{n}(k) \mathbf{n}^H(i) \} = \sigma^2 \mathbf{I} \delta_{k,i}$ é a matriz de correlação do ruído considerado branco, na qual $\delta_{k,i}$ é igual a 1 quando $k = i$ e zero no caso contrário, e $E \{ \mathbf{s}(k) \mathbf{s}^H(i) \}$ é a matriz Hermitiana da correlação das amplitudes complexas do sinal.

Em um modelo genérico consideramos que tanto o módulo como o argumento do sinal $\mathbf{s}(k)$ podem variar com o tempo. Neste caso, a matriz de correlação entre esses sinais será dada por $E \{ \mathbf{s}(k) \mathbf{s}^H(i) \} = \mathbf{P}(k, i)$, onde $\mathbf{P}(k, i)$ é uma matriz Hermitiana qualquer.

Para o caso particular em que consideramos $\mathbf{s}(k)$ como um sinal tipo ruído branco, teremos [28]

$$E \{ \mathbf{s}(k) \mathbf{s}^H(i) \} = \mathbf{P} \delta_{k,i} \quad (5a)$$

e

$$E \{ \mathbf{s}(k) \mathbf{s}^T(i) \} = \mathbf{0}, \quad (5b)$$

para todo k e i . O sinal $\mathbf{s}(k)$ pode ser qualquer sinal aleatório, porém o modelo utilizado em (5) foi escolhido por ser matematicamente tratável.

Nesse modelo também é assumido que o sinal $\mathbf{s}(k)$ e o ruído $\mathbf{n}(i)$ são descorrelacionados para todo k e i .

Levando em conta (5), a correlação em (4) assume a forma

$$\mathbf{R} = \mathbf{R}(k, k) = \mathbf{A} \mathbf{P} \mathbf{A}^H + \sigma^2 \mathbf{I}_{N \times N}. \quad (6)$$

Considerando a matriz \mathbf{P} em (5), podemos estabelecer três situações quanto à correlação cruzada entre os sinais provenientes de M fontes distintas: totalmente descorrelacionados ou independentes, totalmente correlacionados ou coerentes e parcialmente correlacionados.

Para o caso em que os sinais são parcialmente correlacionados, a matriz \mathbf{P} será Hermitiana e $1 < \text{Posto}(\mathbf{P}) \leq M$.

Para o caso em que os sinais são descorrelacionados, a matriz \mathbf{P} será diagonal e, conseqüentemente, $\text{Posto}(\mathbf{P}) = M$. No caso particular de sinais normalizados e com mesma potência, teremos $\mathbf{P} = \mathbf{I}$.

Um outro caso particular com relação ao sinal $\mathbf{s}(k)$, porém de bastante interesse, é quando os sinais são coerentes. Dizemos que dois sinais são coerentes se um é uma versão escalonada e/ou defasada do outro. Isto é, todos os sinais $\{s_m(k)\}$, para $\forall m$, são múltiplos complexos de um sinal comum, $s_0(k)$. Assim $\mathbf{s}(k)$ pode ser expresso como

$$\mathbf{s}(k) = \mathbf{p} s_0(k), \quad (7)$$

na qual \mathbf{p} é um vetor de dimensões $M \times 1$ com elementos complexos quaisquer, porém fixos [29].

Nesse caso podemos notar que a matriz $\mathbf{P} = \rho_0 \mathbf{p} \mathbf{p}^H$, com ρ_0 indicando a potência do sinal $s_0(k)$, possuirá posto unitário.

A coerência entre sinais provenientes de direções distintas é um fenômeno comum resultante do efeito de propagação através de múltiplos percursos [4], [7]. O resultado dessa coerência se apresenta através de uma deficiência no posto da matriz \mathbf{P} .

Consideraremos agora a decomposição da matriz \mathbf{R} em autovetores unitários e autovalores

$$\mathbf{R} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H, \quad (8)$$

na qual $\mathbf{\Lambda}$ é uma matriz diagonal cujos elementos são os autovalores ordenados de forma decrescente.

Considerando a existência de apenas $M < N$ fontes de sinais independentes, em uma situação sem ruído, teremos teoricamente

$$\mathbf{\Lambda} = \text{diag} \{ \lambda_1, \lambda_2, \dots, \lambda_M, \lambda_{M+1}, \dots, \lambda_N \},$$

com $\lambda_{M+1} = \lambda_{M+2} = \dots = \lambda_N = 0$. Com a presença de ruído branco com variância σ^2 teremos

$$\mathbf{\Lambda} = \text{diag} \{ \lambda_1 + \sigma^2, \dots, \lambda_M + \sigma^2, \sigma^2, \dots, \sigma^2 \}.$$

Portanto, indo mais além na decomposição apresentada em (8), podemos reescrever a matriz \mathbf{R} como

$$\mathbf{R} = \mathbf{U}_s \mathbf{\Lambda}_s \mathbf{U}_s^H + \mathbf{U}_n \mathbf{\Lambda}_n \mathbf{U}_n^H, \quad (9)$$

na qual $\Lambda_n = \sigma^2 \mathbf{I}_{(N-M) \times (N-M)}$. As colunas de $[\mathbf{U}_n]_{N \times (N-M)}$ são os autovetores associados aos autovalores em Λ_n , os quais geram subespaço de ruído da matriz \mathbf{R} . Da mesma forma, $\Lambda_s = \text{diag} \{ \lambda_1, \lambda_2, \dots, \lambda_M \}$ e as colunas de $[\mathbf{U}_s]_{N \times M}$ são os autovetores associados aos autovalores em Λ_s , os quais geram o subespaço de sinal de \mathbf{R} .

Observando a equação (6) podemos constatar que qualquer vetor ortogonal às colunas da matriz \mathbf{A} é um autovetor da matriz \mathbf{R} , cujo autovalor é σ^2 . Supondo que \mathbf{A} e \mathbf{P} têm posto cheio igual a M , \mathbf{APA}^H terá posto M . Neste caso, considerando as dimensões envolvidas, existem $N - M$ vetores linearmente independentes com a propriedade acima, os quais, juntamente com seus autovalores, correspondem ao subespaço de ruído. Portanto, todos os autovetores de ruído são ortogonais às colunas da matriz \mathbf{A} , e as colunas de \mathbf{U}_s em (9) devem gerar o mesmo espaço que as colunas da matriz \mathbf{A} , enquanto que as colunas de \mathbf{U}_n geram o seu complemento ortogonal, denominado de espaço nulo de \mathbf{A} . [4], [14].

Os operadores de projeção no subespaço de sinal e subespaço de ruído são então definidos, respectivamente, como [30]

$$\Pi_{\mathbf{A}} = \mathbf{A}\mathbf{A}^\dagger = \mathbf{U}_s\mathbf{U}_s^H \quad (10)$$

e

$$\Pi_{\bar{\mathbf{A}}} = \mathbf{I} - \mathbf{A}\mathbf{A}^\dagger = \mathbf{U}_n\mathbf{U}_n^H, \quad (11)$$

onde $\mathbf{A}^\dagger = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ é a *pseudo-inversa* de Moore Penrose. A segunda igualdade em (10) e (11) pode ser obtida considerando a decomposição em valores singulares da matriz \mathbf{A} .

Para o caso de um número finito de amostras, as esperanças estatísticas presentes em (4) serão substituídas por médias algébricas, dando origem à matriz de correlação estimada

$$\hat{\mathbf{R}} = \frac{1}{K} \sum_{k=1}^K \mathbf{y}(k)\mathbf{y}^H(k). \quad (12)$$

Conseqüentemente, a expressão (9), reescrita para a matriz estimada, será dada por

$$\hat{\mathbf{R}} = \hat{\mathbf{U}}_s \hat{\Lambda}_s \hat{\mathbf{U}}_s^H + \hat{\mathbf{U}}_n \hat{\Lambda}_n \hat{\mathbf{U}}_n^H, \quad (13)$$

onde $\hat{\mathbf{U}}_s$, $\hat{\Lambda}_s$, $\hat{\mathbf{U}}_n$, e $\hat{\Lambda}_n$ são estimativas de suas versões estatísticas.

A seguir apresentaremos os métodos de estimação dos parâmetros DOA mais citados na literatura. Estes métodos serão classificados de acordo com a similaridade de seus modelos, do mais genérico ao mais específico através de uma única notação. Variações dos métodos e mais de uma interpretação dos modelos serão enfatizadas nessa apresentação. Embora a classificação e a seqüência de apresentação divirjam daquelas normalmente encontradas na literatura, nos parecem as mais naturais e surgem como uma nova proposta de relacionar os métodos.

Os métodos serão descritos em três categorias: baseados no critério ML, baseados em predição linear e baseados em subespaços. Essa proposta de classificação favorece as relações que pretendemos estabelecer entre os diversos métodos baseados em conceitos distintos.

2.3 NOTAÇÃO

- $[\cdot]^T$ - denota transposto;
- $[\cdot]^H$ - denota conjugado-transposto;
- $(\cdot)^*$ - denota o complexo conjugado;
- $E\{\cdot\}$ - denota a esperança estatística;
- $\hat{\cdot}$ - sobrescrito denotando estimativa;
- $\|\cdot\|$ - denota a norma-2 de um vetor;
- Tr - denota o traço de uma matriz;
- $\mathbf{a}(\theta_m)$ - vetor da direção angular θ_m ;
- \mathbf{A} - matriz formada pelos vetores da direção angular;
- $\mathbf{A}^\dagger = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ - *pseudo-inversa* de Moore Penrose;
- $b(z)$ - polinômio cujas raízes estimam os ângulos ϕ_m ;
- $\mathbf{b} = [b_1 \dots b_L]^T$ - vetor formado por coeficientes do polinômio $b(z)$;
- $\mathbf{b}' = [b_0 \ b_1 \dots b_L]^T$ - vetor formado pelos coeficientes do polinômio $b(z)$;
- \mathbf{B} - matriz composta pelos elementos de \mathbf{b} ;
- d - distância entre cada elemento do arranjo;
- \mathbf{D} - matriz de dados associada à predição linear;
- \mathbf{D}' - matriz de dados \mathbf{D} aumentada;
- θ_m - direção angular da m 'ésima de onda incidente;
- ϕ_m - ângulo no plano z associado a θ_m ;
- Φ - vetor formado pelos ângulos ϕ_m ;
- φ_m - variável independente nas otimizações para estimar ϕ_m ;
- φ - vetor formado pelas variáveis φ_m ;
- Ψ - conjunto de restrições para otimização de \mathbf{b}' ;
- k - variável de contagem dos snapshots;
- K - número de snapshots;
- L - ordem do preditor linear;
- M - número de ondas incidentes;
- N - número de sensores;
- $n_n(k)$ - amostra do ruído aditivo na saída do n 'ésimo sensor;
- $\mathbf{n}(k)$ - vetor de amostras do ruído aditivo;
- \mathbf{P} - matriz de correlação do sinal;
- $\Pi_{\mathbf{A}}$ - projetor no subespaço de sinal de \mathbf{R} ;
- $\Pi_{\bar{\mathbf{A}}}$ - projetor no subespaço de ruído de \mathbf{R} ;
- \mathbf{R} - matriz de correlação para o sinal mais ruído na saída do arranjo;
- $s_m(k)$ - amplitude da m 'ésima onda incidente;
- $\mathbf{s}(k)$ - vetor das amplitudes das ondas incidentes;
- σ^2 - variância do ruído;
- \mathbf{U} - matriz de autovetores de \mathbf{R} ;
- \mathbf{U}_n - matriz de autovetores de ruído de \mathbf{R} ;
- \mathbf{U}_s - matriz de autovetores de sinal de \mathbf{R} ;
- λ_i - i 'ésimo autovalor de \mathbf{R} ;
- Λ - matriz diagonal com os autovalores de \mathbf{R} ;
- Λ_n - matriz diagonal com os autovalores de ruído de \mathbf{R} ;
- Λ_s - matriz diagonal com os autovalores de sinal de \mathbf{R} ;
- \mathbf{U}_D - matriz dos vetores singulares esquerdos de \mathbf{D} ;
- \mathbf{V}_D - matriz dos vetores singulares direitos de \mathbf{D} ;
- Σ_D - matriz diagonal com os valores singulares de \mathbf{D} ;
- σ_i - valores singulares de \mathbf{D} ;
- $x_n(k)$ - amostra do sinal na saída do n 'ésimo sensor;
- $\mathbf{x}(k)$ - vetor de amostras do sinal na saída do arranjo;
- $y_n(k)$ - amostra do sinal mais ruído na saída do n 'ésimo sensor;

$\mathbf{y}(k)$ - vetor de amostras do sinal mais ruído na saída do arranjo;
 $\mathbf{Y}(k)$ - matriz composta pelos elementos $y_n(k)$;
 \mathbf{W} - matriz de ponderação;

2.4 SIGLAS

AOA - Angle Of Arrival;
 BLP - Backward Linear Prediction;
 DOA - Direction Of Arrival;
 ESPRIT - Estimation of Signal Parameters via Rotational Invariance Techniques;
 FBLP - Forward-Backward Linear Prediction;
 FLP - Forward Linear Prediction;
 IQML - Iterative Quadratic Maximum Likelihood;
 LS - Least Squares;
 ML - Maximum Likelihood;
 MODE - Method Of Direction Estimation;
 MUSIC - Multiple Signal Classification;
 NSF - Noise Subspace Fitting;
 RADAR - Radio Detection And Ranging;
 SF - Subspace Fitting;
 SONAR - SOund NAVigation Ranging;
 SSF - Signal Subspace Fitting;
 TLS - Total Least Squares;
 WSF - Weighted Subspace Fitting;
 WTLS-LP - Whitened Total Least Squares-Linear Prediction.

3. MÁXIMA-VEROSSIMILHANÇA (ML)

O procedimento de estimação através do critério da máxima-verossimilhança será descrito a seguir considerando o sinal $\mathbf{s}(k)$ determinístico ou condicional [4], [28], [31].

3.1 O CRITÉRIO ML DETERMINÍSTICO

Quando o sinal $\mathbf{s}(k)$ é considerado determinístico, uma única medida do vetor $\mathbf{y}(k)$ será uma variável aleatória Gaussiana de dimensão N , com matriz de correlação $\sigma^2 \mathbf{I}$, média $\mathbf{A}(\Phi)\mathbf{s}(k)$ e função de densidade de probabilidade dada por [30], [32]

$$\frac{1}{(\pi\sigma^2)^N} e^{-\|\mathbf{y}(k) - \mathbf{A}(\Phi)\mathbf{s}(k)\|^2/\sigma^2}, \quad (14)$$

na qual $\|\cdot\|$ denota a norma-2 de um vetor. Tomando-se um conjunto de K medidas independentes desse sinal, nos instantes $t = k\Delta t$, para $k = 1, 2, \dots, K$, a função de verossimilhança é obtida como [21]

$$L_{ML} = \prod_{k=1}^K (\pi\sigma^2)^{-N} e^{-\|\mathbf{y}(k) - \mathbf{A}(\Phi)\mathbf{s}(k)\|^2/\sigma^2}, \quad (15)$$

na qual os parâmetros desconhecidos são as direções angulares ϕ_1, \dots, ϕ_M , o vetor do sinal $\mathbf{s}(k)$, e a variância do ruído σ^2 . Os estimadores ML desses parâmetros são aqueles que maximizam o valor da função $L_{ML}(\Phi, \mathbf{s}(k), \sigma^2)$. Por conveniência matemática, as estimativas são definidas como os

argumentos que minimizam a função $l_{ML}(\Phi, \mathbf{s}(k), \sigma^2) = -\log [L_{ML}(\Phi, \mathbf{s}(k), \sigma^2)]$. Normalizando-se em relação a K e ignorando o termo constante, $N \log \pi$, teremos

$$l_{ML}(\Phi, \mathbf{s}(k), \sigma^2) = N \log \sigma^2 + \frac{1}{K\sigma^2} \sum_{k=1}^K \|\mathbf{y}(k) - \mathbf{A}(\Phi)\mathbf{s}(k)\|^2. \quad (16)$$

Através da hipótese de que os parâmetros podem ser estimados de forma separada, demonstra-se que as estimativas dadas por [28]

$$\hat{\sigma}^2 = \frac{1}{N-M} \text{Tr} \left\{ \Pi_{\hat{\mathbf{A}}} \hat{\mathbf{R}} \right\} = \frac{1}{N-M} \text{Tr} \left\{ \hat{\mathbf{\Lambda}}_n \right\} \quad (17)$$

$$\hat{\mathbf{s}}(k) = \mathbf{A}^\dagger \mathbf{y}(k) = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{y}(k),$$

são aquelas que minimizam a expressão (16), onde $\text{Tr} \{ \cdot \}$ denota o traço da matriz [33]. A estimativa $\hat{\mathbf{s}}(k)$ em (17) é aquela que minimiza o erro de estimação quadrático [34]. Substituindo (17) em (16), e observando que

$$\frac{1}{K} \sum_{k=1}^K \|\Pi_{\hat{\mathbf{A}}} \mathbf{y}(k)\|^2 = \text{Tr} \left\{ \Pi_{\hat{\mathbf{A}}} \hat{\mathbf{R}} \right\} \triangleq J_{ML},$$

transformamos o problema original no seguinte problema de minimização

$$\min_{\varphi} J_{ML}(\varphi) = \min_{\varphi} \text{Tr} \left\{ \Pi_{\hat{\mathbf{A}}}(\varphi) \hat{\mathbf{R}} \right\}, \quad (18)$$

com $\varphi = [\varphi_1 \dots \varphi_m]^T$, onde φ_m , $1 \leq m \leq M$, são as variáveis independentes na otimização associadas aos ângulos ϕ_m para $m = 1, \dots, M$.

Como dispomos de apenas de um número finito de amostras, as estimativas para os parâmetros $\{\phi_1, \dots, \phi_M\}$ apresentarão, em geral, um desvio em relação aos seus verdadeiros valores. Entretanto, se o cenário for estacionário, esse desvio convergirá para zero com o aumento do número K de conjuntos de amostras.

Para o cálculo da estimativa em (18) é necessário a solução de uma otimização não-linear M -dimensional. Essa constatação torna a implementação direta do método ML inviável em algumas aplicações diante do esforço computacional exigido.

3.2 O CRITÉRIO ML DETERMINÍSTICO BASEADO EM SUBESPAÇOS

Para permitir uma notação unificada dos métodos de DOA, apresentaremos uma descrição alternativa do estimador ML determinístico baseada em subespaços. Tal alternativa consiste basicamente em escrever a matriz de correlação estimada através de sua versão decomposta em subespaços.

Tomando-se (13) e substituindo em (18) teremos a seguinte função custo a ser minimizada

$$J_{SML}(\varphi) = \text{Tr} \left\{ \Pi_{\hat{\mathbf{A}}}(\varphi) \left(\hat{\mathbf{U}}_s \hat{\mathbf{\Lambda}}_s \hat{\mathbf{U}}_s^H + \hat{\sigma}^2 \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H \right) \right\}, \quad (19)$$

na qual tomamos $\hat{\mathbf{\Lambda}}_n = \hat{\sigma}^2 \mathbf{I}$ como estimador consistente para os autovalores de ruído. O subscrito *SML* em (19) indica o uso do estimador ML baseado em subespaço.

Usando (11) e supondo que os subespaços são estimados corretamente, devemos esperar $\mathbf{U}_n^H \hat{\mathbf{U}}_s = \mathbf{0}$ quando o conjunto de parâmetros estimados $\hat{\Phi}$ é ótimo. Inspirado nesse resultado, a função custo descrita em (19) permite desprezar o termo associado ao subespaço de ruído, levando à seguinte função custo equivalente

$$J_{SML}(\varphi) = Tr \left\{ \Pi_{\hat{\mathbf{A}}}(\varphi) \left(\hat{\mathbf{U}}_s \hat{\Lambda}_s \hat{\mathbf{U}}_s^H \right) \right\}. \quad (20)$$

3.3 O CRITÉRIO ML DETERMINÍSTICO REPARAMETRIZADO

Esse procedimento faz uso de uma reparametrização do operador de projeção no espaço nulo da matriz \mathbf{A} , através da definição de uma particular matriz de *Toeplitz* \mathbf{B} , de dimensões $(N - L) \times N$, dada por

$$\mathbf{B}^H = \begin{bmatrix} b_L & b_{L-1} & \cdots & b_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & b_L & b_{L-1} & \cdots & b_0 \end{bmatrix}. \quad (21)$$

na qual os elementos b_0, b_1, \dots, b_L são os coeficientes de um polinômio $b(z)$ cujas raízes devem estar associadas aos parâmetros $\phi_1, \phi_2, \dots, \phi_M$. Mais precisamente, para $L = M$, o polinômio $b(z)$ deve possuir suas raízes sobre a circunferência de raio unitário (CRU) na posição angular indicada pelos parâmetros $\phi_1, \phi_2, \dots, \phi_M$ do sinal, podendo ser descrito, genericamente, como

$$b(z) = b_0 \prod_{m=1}^M (z - e^{j\phi_m}) \quad (22)$$

Tomando-se as matrizes \mathbf{A} e \mathbf{B}^H e considerando $L = M$, podemos notar que para um sinal livre de ruído

$$\mathbf{B}^H \mathbf{A} = \mathbf{A}^H \mathbf{B} = \mathbf{0}. \quad (23)$$

Para $L = M$, se \mathbf{B}^H tem posto cheio e igual a $(N - M)$ e \mathbf{A} possui posto cheio e igual a M , as linhas de \mathbf{B}^H formarão uma base para o espaço ortogonal às colunas de \mathbf{A} , ou seja, para o *espaço nulo* de \mathbf{A} . Desta forma, podemos escrever

$$\Pi_{\mathbf{B}} = \mathbf{B}(\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H = \Pi_{\hat{\mathbf{A}}} \quad (24)$$

Considerando a função custo em (18), o critério ML pode ser reparametrizado para os coeficientes b_k , e as estimativas dos coeficientes podem ser calculadas minimizando-se a função custo de quarta ordem a seguir

$$J_{ML}(\mathbf{b}') = Tr \left\{ [\mathbf{B}(\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H] \hat{\mathbf{R}} \right\}. \quad (25)$$

onde $\mathbf{b}' = [b_0 \ b_1 \ \dots \ b_L]^T$ com $L = M$.

Os métodos apresentados nas seções 4, 5, 6 e 7 supõem conhecido o número M de fontes de sinal. Quando M não é conhecido, poderá ser estimado utilizando métodos específicos para esse fim [35], [36]. Todos esses métodos são baseados na decomposição em autovalores e autovetores da matriz de correlação das amostras. Portanto, são adequados para serem utilizados conjuntamente com os métodos que fazem uso da mesma decomposição.

4. MÉTODOS BASEADOS NO CRITÉRIO ML

Os métodos apresentados nessa seção fazem uso do critério ML, implementando a busca multidimensional exigida pela sua minimização através de alternativas de menor esforço computacional.

4.1 MÉTODO IQML

O método da iteração quadrática para o critério ML (IQML) procura resolver o problema de minimização do critério ML através da solução de um problema de segunda ordem, utilizando um algoritmo iterativo [21].

O algoritmo IQML procura minimizar (25) considerando, em um primeiro passo, o seguinte problema de segunda ordem

$$\hat{\mathbf{b}}' = \arg \left\{ \min_{\mathbf{b}'} Tr \left\{ \mathbf{B} \mathbf{B}^H \hat{\mathbf{R}} \right\} \right\}, \quad (26)$$

prossequindo com a formação da matriz $\hat{\mathbf{B}}$ e do termo $(\hat{\mathbf{B}}^H \hat{\mathbf{B}})^{-1}$, presente em (25), a partir das estimativas dadas por $\hat{\mathbf{b}}'$ em (26). Em um segundo passo, novas estimativas são obtidas, desta vez a partir de (25) com a substituição do termo $(\mathbf{B}^H \mathbf{B})^{-1}$ por sua estimativa $(\hat{\mathbf{B}}^H \hat{\mathbf{B}})^{-1}$. Esse segundo passo continua iterativamente até atingir a convergência. O método tem como último passo o cálculo das raízes de $b(z)$ para a obtenção das estimativas $\hat{\phi}_m$, para $m = 1, 2, \dots, M$.

Entrando em alguns detalhes da sua implementação, o método é construído através da igualdade proveniente da comutatividade da operação de convolução, apresentada abaixo

$$\mathbf{B}^H \mathbf{y}(k) = \mathbf{Y}(k) \mathbf{b}', \text{ para } k = 1, 2, \dots, K, \quad (27)$$

onde a matriz $\mathbf{Y}(k)$ de dimensões $(N - M) \times (M + 1)$ é definida como

$$\mathbf{Y}(k) \triangleq \begin{bmatrix} y_M(k) & y_{M-1}(k) & \cdots & y_0(k) \\ y_{M+1}(k) & y_M(k) & \cdots & y_1(k) \\ \vdots & \vdots & \vdots & \vdots \\ y_{N-1}(k) & y_{N-2}(k) & \cdots & y_{N-M-1}(k) \end{bmatrix} \quad (28)$$

para $k = 1, 2, \dots, K$, com $N > M$.

Substituindo $\hat{\mathbf{R}}$ da expressão (12) em (25) obtemos

$$J_{IQML}(\mathbf{b}') = Tr \left\{ \frac{1}{K} \sum_{k=1}^K \mathbf{B}(\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \mathbf{y}(k) \mathbf{y}^H(k) \right\}. \quad (29)$$

Utilizando (27), considerando que

$$Tr \left\{ \left(\mathbf{B}(\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \mathbf{y}(k) \right) \mathbf{y}^H(k) \right\} = \mathbf{y}^H(k) \left(\mathbf{B}(\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \mathbf{y}(k) \right),$$

e ainda desconsiderando o termo $1/K$ à esquerda, teremos o seguinte problema de minimização equivalente dado por

$$\hat{\mathbf{b}}' = \arg \left\{ \min_{\mathbf{b}' \in \Psi} \mathbf{b}'^H \left[\sum_{k=1}^K \mathbf{Y}^H(k) (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{Y}(k) \right] \mathbf{b}' \right\}, \quad (30)$$

onde Ψ representa o conjunto adequado de restrições impostas ao parâmetro \mathbf{b}' . Este conjunto é construído através da interseção de duas restrições

$$\Psi = \Psi_0 \cap \Psi_s. \quad (31)$$

A primeira restrição, denominada de "não-trivialidade", é definida como

$$\Psi_0 \triangleq \{\mathbf{b}' \mid \|\mathbf{b}'\| \neq 0\}, \quad (32)$$

A implementação de tal restrição pode ser realizada através de uma relação linear, fazendo algum elemento de \mathbf{b}' diferente de zero, considerando sua parte real e imaginária de forma distinta. Pode ainda ser realizada através de uma relação quadrática na qual impomos um valor diferente de zero para a norma do vetor \mathbf{b}' .

A segunda restrição Ψ_s é escolhida de acordo com o modelo do sinal considerado, e é essencial para a obtenção da solução ótima. Para o modelo de sinal adotado aqui, ou seja, composto pela soma de M exponenciais complexas não-amortecidas, deveremos impor que as raízes do polinômio $b(z)$ estejam localizadas sobre a CRU para o caso ótimo, ou seja

$$\Psi_s \triangleq \{\mathbf{b}' \mid b(z) \text{ possui raízes sobre a CRU}\}.$$

Devido à dificuldade de incorporar essa restrição na minimização em (30), o método IQML impõe apenas a simetria complexa conjugada sobre os coeficientes de $b(z)$, fazendo $b_m = b_{M-m}^*$, $m = 0, \dots, M$. Embora tal restrição não seja necessária nem suficiente para satisfazer Ψ_s , tal implementação apresenta a vantagem de reduzir pela metade o número de parâmetros a serem obtidos. No entanto, não pode ser utilizada juntamente com uma restrição de não trivialidade na qual $b_0 = 1$, pois um polinômio da forma $b(z) = b_0 \prod_{m=1}^M (z - e^{j\phi_m})$ não satisfaz, de uma forma geral, a relação $b_m = b_{M-m}^*$ quando $b_0 = 1$. Neste caso, outra alternativa para implementação da restrição de não-trivialidade deve ser utilizada.

O algoritmo utilizado pelo método IQML para a minimização da função custo em (30), considera a implementação da restrição de simetria complexa conjugada de forma implícita, ou seja, a simetria é incorporada na função custo. Para evitar a solução trivial faz-se $Re(b_0) = 1$, onde $Re(\cdot)$ indica a parte real [21], [37], [38].

Podemos concluir que as duas idéias mais relevantes nesse método são o uso da reparametrização do critério ML e a utilização de uma forma iterativa para resolver um problema de minimização de quarta ordem com restrições a partir da solução de um problema de segunda ordem sem restrições. Como deficiências do método IQML, devemos citar que não existe garantia de convergência para o processo iterativo. Por outro lado, quando ocorre a convergência, está se dá após vários passos de iteração. Por fim, mesmo após a convergência não há garantias de se atingir o mínimo global.

4.2 MÉTODO MODE

O método MODE pode ser interpretado como um aprimoramento das idéias do método IQML, no qual se obtém

melhores estimativas no passo inicial, levando a uma rápida convergência. Além disso, no seu desenvolvimento faz-se uso dos subespaços definidos pela matriz de correlação conforme a versão do critério ML baseada em subespaços, apresentada anteriormente [22].

Mais especificamente, a proposta básica do método MODE é minimizar a função custo

$$J(\mathbf{b}') = Tr_{\mathbf{b}' \approx \mathbf{b}'_o} \left[\mathbf{B}(\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \hat{\mathbf{R}} \right], \quad (33)$$

localmente em torno de \mathbf{b}'_o , onde \mathbf{b}' e \mathbf{b}'_o indicam, respectivamente, o parâmetro desconhecido e seu valor ótimo assintótico em termos de K . Desta forma, o método MODE pretende contornar os problemas de convergência e reduzir as iterações necessárias para alcançar as estimativas finais.

Estamos interessados no comportamento local de $J(\mathbf{b}')$ no qual $\|\hat{\mathbf{b}}' - \mathbf{b}'_o\| \approx 0$ para K grande. Isso implica em que $\|\hat{\mathbf{B}} - \mathbf{B}_o\| \approx 0$ e, conseqüentemente, $\hat{\mathbf{B}}^H \hat{\mathbf{U}}_s = \hat{\mathbf{U}}_s^H \hat{\mathbf{B}} = 0$ para K grande [39]. Partindo-se da decomposição da matriz $\hat{\mathbf{R}}$ em subespaços, conforme realizado em (13), e considerando ainda que $\hat{\mathbf{U}}_s \hat{\mathbf{U}}_s^H + \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H = \mathbf{I}$, podemos então aproximar a função custo (33) localmente como [22]

$$J(\mathbf{b}') \approx Tr_{\mathbf{b}' \approx \mathbf{b}'_o} \left\{ \hat{\mathbf{U}}_s^H \mathbf{\Pi}_B \hat{\mathbf{U}}_s \left(\hat{\mathbf{\Lambda}}_s - \hat{\mathbf{\Lambda}}_n \right) \right\}, \quad (34)$$

com $\hat{\mathbf{\Lambda}}_n \approx \sigma^2 \mathbf{I}$, uma vez que $\hat{\mathbf{\Lambda}}_n$ é um estimador consistente para $\sigma^2 \mathbf{I}$.

Desta forma, o estimador MODE será dado pela minimização da seguinte função custo equivalente a (34)

$$J(\mathbf{b}') = Tr \left\{ \hat{\mathbf{U}}_s^H \mathbf{\Pi}_B \hat{\mathbf{U}}_s \left(\hat{\mathbf{\Lambda}}_s - \hat{\sigma}^2 \mathbf{I} \right) \right\}, \quad (35)$$

com $\mathbf{\Pi}_B = \mathbf{B}(\hat{\mathbf{B}}^H \hat{\mathbf{B}})^{-1} \mathbf{B}^H$ e na qual $\hat{\sigma}^2$ e $\hat{\mathbf{B}}$ são estimativas consistentes de σ^2 e \mathbf{B} , respectivamente.

Uma estimativa consistente de σ^2 pode ser obtida através de

$$\begin{aligned} \hat{\sigma}^2 &= \frac{1}{N-M} \left\{ Tr \left\{ \hat{\mathbf{R}} \right\} - Tr \left\{ \hat{\mathbf{\Lambda}}_s \right\} \right\} \\ &= \frac{1}{N-M} Tr[\hat{\mathbf{\Lambda}}_n]. \end{aligned} \quad (36)$$

As estimativas $\hat{\mathbf{B}}$ e $\hat{\mathbf{b}}'$, por sua vez, podem ser determinados através do problema de minimização quadrática

$$\hat{\mathbf{b}}' = \arg \left\{ \min_{\mathbf{b}' \in \Psi} Tr \left\{ \mathbf{B} \mathbf{B}^H \hat{\mathbf{U}}_s \mathbf{W}_{MODE} \hat{\mathbf{U}}_s^H \right\} \right\}, \quad (37)$$

onde Ψ representa o conjunto adequado de restrições e \mathbf{W}_{MODE} é uma matriz de ponderação escolhida adequadamente que generaliza o método MODE, dando origem à generalização da função custo em (35)

$$J_{MODE}(\mathbf{b}') = Tr \left\{ \hat{\mathbf{U}}_s^H \mathbf{\Pi}_B \hat{\mathbf{U}}_s \mathbf{W}_{MODE} \right\}. \quad (38)$$

A consistência da estimativa obtida através de (38) se justifica pelo fato de a expressão alcançar seu mínimo valor apenas quando $\mathbf{B}^H \hat{\mathbf{U}}_s = 0$ e quando a matriz \mathbf{W}_{MODE} é positiva definida.

Demonstra-se que a escolha ótima para \mathbf{W}_{MODE} , \mathbf{W}_o , sob o ponto de vista de minimização da variância do estimador, é dada por ([4], [22] e [40])

$$\mathbf{W}_o = (\hat{\Lambda}_s - \hat{\sigma}^2 \mathbf{I})^2 \hat{\Lambda}_s^{-1}. \quad (39)$$

A vantagem da utilização de (39) em (37) é que as estimativas se aproximam dos seus valores assintóticos, dando origem a um procedimento direto para obtenção da estimativa $\hat{\mathbf{b}}'$, em dois passos.

Considerando a estimativa dos \bar{r} maiores autovalores de $\hat{\mathbf{R}}$, como $\bar{r} = \min\{M, \text{posto}(\hat{\mathbf{P}})\}$, no primeiro passo resolve-se o problema quadrático

$$\hat{\mathbf{b}}' = \arg \left\{ \min_{\mathbf{b}' \in \Psi} Tr \left\{ \mathbf{B} \mathbf{B}^H \hat{\mathbf{U}}_s \mathbf{W}_o \hat{\mathbf{U}}_s^H \right\} \right\}, \quad (40)$$

do qual obtém-se $\hat{\mathbf{B}}$.

No segundo passo resolve-se o problema quadrático

$$\hat{\mathbf{b}}' = \arg \left\{ \min_{\mathbf{b}' \in \Psi} Tr \left\{ \mathbf{B} (\hat{\mathbf{B}}^H \hat{\mathbf{B}})^{-1} \mathbf{B}^H \hat{\mathbf{U}}_s \mathbf{W}_o \hat{\mathbf{U}}_s^H \right\} \right\}, \quad (41)$$

com $\hat{\mathbf{B}}$ dado pelo primeiro passo.

As estimativas finais para ϕ_m , $m = 1, 2, \dots, M$, são então obtidas através das raízes do polinômio $b(z)$ formado pelos elementos de $\hat{\mathbf{b}}'$ obtidos no último passo.

Comparando as expressões (41) e (40), respectivamente, com (25) e (26), poderemos contrastar os métodos MODE e IQML. Inicialmente, devemos observar que o produto $\hat{\mathbf{U}}_s \mathbf{W}_o \hat{\mathbf{U}}_s^H$ é uma estimativa da parcela de $\hat{\mathbf{R}}$ associada ao subespaço de sinal. Assim, a comparação direta entre as expressões citadas nos permite concluir que os métodos diferem pelo fato do método MODE operar sobre a parcela do subespaço de sinal de $\hat{\mathbf{R}}$, com o objetivo de eliminar os efeitos indesejáveis da presença do ruído. Dessa forma torna-se possível a obtenção de uma melhor estimativa inicial para \mathbf{b}' , que leva, em geral, à estimativa final no segundo passo, evitando assim os passos de iteração presentes no método IQML e seus problemas de convergência.

Podemos ainda notar que a minimização em (40) corresponde ao primeiro passo do método IQML quando $(\hat{\mathbf{B}}^H \hat{\mathbf{B}})^{-1} = \mathbf{I}$ e $\hat{\mathbf{R}} \cong \hat{\mathbf{U}}_s \mathbf{W}_o \hat{\mathbf{U}}_s^H$. Nesse caso, a matriz de correlação é substituída pela sua versão decomposta em subespaços da qual apenas a porção associada ao sinal é considerada. Além disso, é utilizada uma estimativa mais próxima do valor ótimo para os autovalores associados ao subespaço de sinal, dadas por \mathbf{W}_o .

Embora o método MODE apresente uma redução no esforço computacional em relação ao IQML, devido à eliminação das iterações, se faz necessário a obtenção de estimativas adicionais através da decomposição da matriz de correlação. Assim sendo, seu sucesso depende da qualidade dessas estimativas.

O algoritmo de solução para o método MODE também implementa a restrição de simetria complexa conjugada dos coeficientes b_k de forma implícita. O algoritmo MODE sugere ainda uma restrição ao primeiro elemento do vetor \mathbf{b}' , b_0 , da forma

$$Re(b_0) = 1 \text{ ou } Im(b_0) = 1, \quad (42)$$

evitando assim a solução trivial para \mathbf{b}' e a incompatibilidade do uso $b_0 = 1$. A opção $Re(b_0) = 1$, como sugerido no método IQML, não se aplica às situações nas quais a soma dos parâmetros $\{\phi_m\}$ é um múltiplo ímpar de π . Nessas situações pode-se, portanto, utilizar $Im(b_0) = 1$, que por sua vez não se aplica às situações nas quais a soma dos parâmetros $\{\phi_m\}$ é um múltiplo par de π . Assim, o uso de uma ou outra forma, de acordo com a situação, será válida para todas circunstâncias [37], [38].

Para a minimização de (41), o algoritmo MODE sugere a utilização do critério dos mínimos quadrados (LS) implementado através da decomposição QR [22].

A função custo associada ao problema de minimização em (30), encontrado a cada passo de iteração do método IQML, pode ser escrita na forma compacta

$$J_{IQML}(\mathbf{b}') = \left\| (\boldsymbol{\Omega}_{IQML})^{1/2} \mathbf{b}' \right\|^2, \quad (43)$$

na qual

$$\boldsymbol{\Omega}_{IQML} = \sum_{k=1}^K \left\{ \mathbf{Y}^H(k) (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{Y}(k) \right\},$$

que, por sua vez, é semelhante em sua formulação à função custo do método MODE. Desta forma, os procedimentos de solução sugeridos pelos dois algoritmos podem ser aplicados a ambos indistintamente.

Apesar de ser considerado um algoritmo não-iterativo, o algoritmo MODE está naturalmente preparado para ser iterado levando, em geral, a estimativas mais próximas de seus valores ótimos [22]. Por apresentar a melhor precisão assintótica possível e não exigir grande número de iterações, tal método é um forte candidato a melhor método para arranjos lineares uniformes.

5. ESTIMAÇÃO DOS PARÂMETROS DOA VIA PREDIÇÃO LINEAR

O procedimento de estimação dos parâmetros DOA através de predição linear consiste basicamente em se otimizar um filtro de predição para o sinal em sua entrada e estimar os parâmetros DOA a partir das raízes do polinômio formado pelos coeficientes do filtro resultante [11], [12], [13].

Para analisar os conceitos básicos do problema de otimização da predição linear, consideremos um conjunto de N amostras de um sinal no instante k , dado pelo vetor $\mathbf{y}(k)$, como a entrada do filtro transversal com $(L + 1)$ fatores de ganhos, conforme a Figura 2.

O sinal na saída deste filtro será dado por

$$\epsilon_n(k) = \sum_{l=0}^L y_{n-l}(k) b_l, \quad (44)$$

para $n = L, L + 1, \dots, N - 1$. Para que esse filtro funcione como um filtro de erro de predição tomamos $b_0 = -1$. Desta forma, a energia do sinal de erro na saída do filtro, para cada

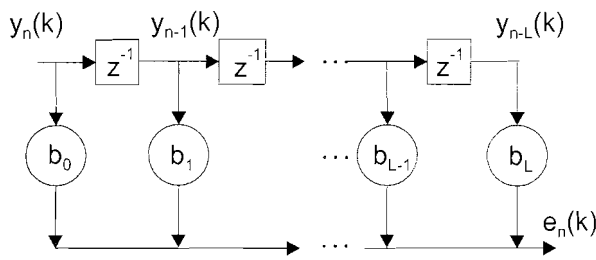


Figura 2. Filtro de erro de predição

valor de k , será dada por

$$E(k) = \sum_{n=L}^{N-1} |e_n(k)|^2 = \sum_{n=L}^{N-1} \left| \sum_{l=1}^L y_{n-l}(k)b_l - y_n(k) \right|^2. \quad (45)$$

Os parâmetros b_l , $l = 1, \dots, L$, são então obtidos para cada experimento k minimizando-se (45).

Mais especificamente, se definirmos a matriz de dados, de dimensões $(N - L) \times L$, no instante k como

$$\mathbf{D}(k) \triangleq \begin{bmatrix} y_{L-1}(k) & y_{L-2}(k) & \dots & y_0(k) \\ y_L(k) & y_{L-1}(k) & \dots & y_1(k) \\ \vdots & \vdots & \vdots & \vdots \\ y_{N-2}(k) & y_{N-3}(k) & \dots & y_{N-L-1}(k) \end{bmatrix}, \quad (46)$$

e o vetor dos valores preditos como

$$\mathbf{d}(k) = [y_L(k) \ \dots \ y_{N-1}(k)]^T,$$

poderemos descrever o processo de filtragem através da forma matricial

$$\mathbf{D}(k)\mathbf{b}(k) - \mathbf{d}(k) = \mathbf{e}(k), \quad (47)$$

na qual

$$\mathbf{b}(k) = [b_1 \ b_2 \ \dots \ b_L]^T$$

$$\mathbf{e}(k) = [e_L(k) \ e_{L+1}(k) \ \dots \ e_{N-1}(k)]^T.$$

A expressão em (47) pode ser reescrita como

$$\begin{aligned} \mathbf{D}(k)\mathbf{b}(k) - \mathbf{d}(k) &= \begin{bmatrix} \mathbf{d}(k); \mathbf{D}(k) \end{bmatrix} \begin{bmatrix} -1 \\ \mathbf{b}(k) \end{bmatrix} \\ &= \mathbf{D}'(k)\mathbf{b}'(k) = \mathbf{e}(k), \end{aligned} \quad (48)$$

na qual $\mathbf{D}'(k)$ e $\mathbf{b}'(k)$ denotam as versões aumentadas, respectivamente, da matriz de dados $\mathbf{D}(k)$ e do vetor dos ganhos do filtro $\mathbf{b}(k)$ para $b_0 = -1$. Podemos observar que $\mathbf{D}'(k)$ é a própria matriz $\mathbf{Y}(k)$ definida em (28) para $L = M$. Tal constatação permite relacionar os métodos baseados na predição linear com os métodos baseados no ML que utilizam a reparametrização dada pela matriz \mathbf{B}^H em (21).

Para um único experimento, substituindo o vetor de erro dado por (48) em (45), teremos a seguinte função custo a ser minimizada nos métodos baseados na predição linear

$$J_{PL}(\mathbf{b}) = \left\{ (\mathbf{b}')^H (\mathbf{D}')^H \mathbf{D}' \mathbf{b}' \right\}, \text{ com } b_0 = -1, \quad (49)$$

na qual \mathbf{b}' é o mesmo vetor utilizado pelos métodos baseados no ML quando $L = M$.

Devemos enfatizar que a minimização proposta fornecerá um vetor de coeficientes ótimos para cada experimento k , de forma independente. No entanto, tal procedimento não consiste na forma mais eficiente de utilização dos dados disponíveis. Podemos notar que nos métodos baseados no critério ML o procedimento de minimização faz uso de vários experimentos para a obtenção de um único vetor \mathbf{b} ótimo. Essa inabilidade de operar sobre mais de um experimento simultaneamente é uma das limitações do uso da predição linear para a estimação de parâmetros DOA, quando comparada com o critério ML. Uma forma de contornar tal deficiência será proposta neste trabalho como mais uma contribuição.

Apresentaremos a seguir as propriedades da predição linear que permitem sua aplicação na estimação dos parâmetros ϕ_m em (3), para $m = 1, 2, \dots, M$. Para tanto, consideraremos o polinômio de ordem L dado por

$$b(z) = b_0 z^L + b_1 z^{L-1} + \dots + b_L, \quad (50)$$

cujos os coeficientes são os ganhos do preditor linear otimizado para o vetor de amostras $\mathbf{y}(k)$ sem ruído, ou seja, $\mathbf{x}(k)$.

Em [8], [10] e [13] está demonstrado que se L satisfaz a desigualdade $M \leq L \leq N$, e se \mathbf{b} satisfaz $\mathbf{D}\mathbf{b} = \mathbf{d}$, então $b(z)$ terá M de seus zeros posicionados sobre a CRU em $z_m = \exp[j\phi_m]$, para $m = 1, \dots, M$. Este resultado pode ser explicado observando que o filtro de erro de predição, otimizado para sinais do tipo exponenciais complexas nas frequências ϕ_m , deverá apresentar zeros de sua função de transferência nas frequências ϕ_m de forma a minimizar o erro na sua saída. Estes zeros são denominados de zeros de sinal [41]. Desta forma, os ângulos indicados pelos zeros de sinal fornecem estimativas dos parâmetros ϕ_m , para $m = 1, \dots, M$.

O posicionamento dos $L - M$ zeros restantes dependerá da solução escolhida, dentre as várias que satisfazem $\mathbf{D}\mathbf{b} = \mathbf{d}$. Para a solução que apresenta a menor norma-2, denominada de norma-mínima, esses zeros estarão uniformemente posicionados no interior da CRU [8], [10] e [13]. Para esta solução, o preditor será denominado de norma-mínima e, conseqüentemente, apresentará fase mínima [10], [13], [41].

Para o caso ruidoso, tal comportamento não é mais válido, uma vez que todos os zeros flutuam em torno de suas posições da situação sem ruído. Conseqüentemente, tais flutuações produzirão uma variância nas estimativas obtidas. Neste caso, a estratégia de identificação dos zeros de sinal consiste em associá-los aos M zeros mais próximos da CRU. Essa estratégia permite a distinção entre os dois conjuntos de zeros em relações sinal-ruído (SNR) moderadas [11]. Devemos notar que para o critério ML reparametrizado tal distinção entre os zeros não existe, pois $L = M$.

Como a minimização da função custo em (49) é um problema de segunda ordem, podemos fazer uso do procedimento LS de minimização, de onde derivamos a solução geral a seguir [10], [11], [12], [13], [42]

$$\hat{\mathbf{b}}_{PL} = (\mathbf{D}^H \mathbf{D})^{-1} \mathbf{D}^H \mathbf{d} = \mathbf{D}^\dagger \mathbf{d}. \quad (51)$$

Reconhecendo o termo $(\mathbf{D}')^H \mathbf{D}'$ em (49) como uma estimativa da matriz de correlação, poderemos contrastar esse

método com os anteriores baseados na matriz de correlação estimada. Para um único experimento ($K = 1$) e $L = M$, podemos constatar que $\mathbf{D}' = \mathbf{Y}$, definida em (28). Assim, poderemos reescrever (49) como

$$J_{PL}(\mathbf{b}) = \begin{bmatrix} b_0 \\ \mathbf{b} \end{bmatrix}^H \mathbf{Y}^H \mathbf{Y} \begin{bmatrix} b_0 \\ \mathbf{b} \end{bmatrix}, \quad (52)$$

com $b_0 = -1$ e $L = M$.

A extensão do método de predição linear para a situação em que dispomos de múltiplos experimentos (K "snapshots") pode ser obtida naturalmente substituindo-se a função custo em (52) por sua generalização

$$J(\mathbf{b}) = \begin{bmatrix} b_0 \\ \mathbf{b} \end{bmatrix}^H \mathbf{K}_y \begin{bmatrix} b_0 \\ \mathbf{b} \end{bmatrix}, \quad \text{com } b_0 = -1, \quad (53)$$

onde a matriz $\mathbf{K}_y \triangleq \frac{1}{K} \sum_{k=1}^K \mathbf{Y}(k)^H \mathbf{Y}(k)$, com dimensões $(L+1) \times (L+1)$, possui um caráter de matriz de correlação.

O uso de múltiplos experimentos para a formação da matriz \mathbf{K}_y permite combater as dificuldades de resolução desse método na presença de sinais coerentes.

Podemos agora comparar (53) com a função custo associada ao critério ML reparametrizado em (25). Concluímos que a minimização da função custo da forma em (53) não apresenta a matriz $(\mathbf{B}^H \mathbf{B})^{-1}$ em (25), a qual é útil na compensação de correlação entre as fontes de sinal. Dessa forma (53) equivale à função custo em (25) quando temos $(\mathbf{B}^H \mathbf{B})^{-1} = \mathbf{I}$. Por outro lado, uma vantagem aparente desse procedimento sobre o ML reparametrizado consiste na possibilidade de utilização de uma ordem $L > M$ para o polinômio $\hat{b}(z)$, possibilitando um maior grau de liberdade para a alocação dos zeros de sinal que, em geral, melhora a qualidade das estimativas obtidas. Além disso, podemos observar que a minimização da função custo em (53) consiste em um problema de otimização de segunda ordem, notadamente mais simples que o enfrentado pelo procedimento de estimação através do critério ML. Essa característica torna a estimação do parâmetro DOA via predição linear uma alternativa atraente. No entanto, sua aplicação se restringe a um cenário no qual os sinais possuem potências semelhantes, são levemente correlacionados e em SNRs moderadas [4], [10].

5.1 PREDIÇÃO LINEAR BASEADA EM SUBESPAÇOS

A predição linear baseada em subespaços consiste na utilização da decomposição da matriz de dados \mathbf{D} em vetores e valores singulares (SVD). Para definir esta decomposição consideraremos que qualquer matriz \mathbf{D} $m \times n$ pode ser fatorada da forma $\mathbf{D} = \mathbf{U}_D \mathbf{\Sigma}_D \mathbf{V}_D^H$, onde \mathbf{U}_D e \mathbf{V}_D são unitárias com dimensões $m \times m$ e $n \times n$, respectivamente, e $\mathbf{\Sigma}_D$ é uma matriz diagonal de dimensões $m \times n$ composta por elementos reais σ_i . Tais elementos são denominados de valores singulares e são ordenados tal que $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p > 0$, com $p \leq \min(m, n)$, onde p é o posto de \mathbf{D} . As colunas de \mathbf{U}_D e \mathbf{V}_D , denotadas por \mathbf{u}_i^D e \mathbf{v}_i^D , são denominadas, respectivamente, de vetores singulares à esquerda e à direita de \mathbf{D} [11], [12], [33].

Podemos agora aplicar estes conceitos sobre a matriz \mathbf{D} em (46), considerando a seguinte partição

$$\mathbf{D} = \begin{bmatrix} \mathbf{U}_1^D & \mathbf{U}_2^D \end{bmatrix} \cdot \begin{bmatrix} \mathbf{\Sigma}_1^D & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_2^D \end{bmatrix} \cdot \begin{bmatrix} \mathbf{V}_1^D & \mathbf{V}_2^D \end{bmatrix}^H, \quad (54)$$

onde \mathbf{U}_1^D é uma partição de \mathbf{U}_D contendo suas p primeiras colunas e \mathbf{U}_2^D suas $(N-L) - p$ últimas colunas; \mathbf{V}_1^D é a partição de \mathbf{V}_D contendo as suas p primeiras colunas e \mathbf{V}_2^D as $L-p$ colunas restantes; $\mathbf{\Sigma}_1^D$ é uma matriz diagonal com dimensões $p \times p$ contendo os p maiores valores singulares de \mathbf{D} e $\mathbf{\Sigma}_2^D$ é uma matriz diagonal com dimensões $(N-L-p) \times (L-p)$, contendo os $\min((N-L-p), (L-p))$ últimos elementos ordenados de $\mathbf{\Sigma}_D$.

No caso de um sinal composto por M exponenciais complexas independentes e sem ruído, o posto de \mathbf{D} será $p = M$ e $\mathbf{\Sigma}_2^D$ será nula. Portanto, existem M valores singulares, não nulos, associados aos vetores singulares de \mathbf{U}_1^D e \mathbf{V}_1^D , denominados, respectivamente, de valores e vetores singulares de sinal.

No caso de um sinal ruidoso, os valores singulares presentes em $\mathbf{\Sigma}_2^D$, anteriormente nulos, tornam-se positivos e com magnitude crescente com a redução da relação sinal-ruído (SNR). Portanto, para um filtro preditor de ordem L , teremos os $L-M$ menores valores singulares de \mathbf{D} em $\mathbf{\Sigma}_2^D$ e seus correspondentes vetores singulares de \mathbf{U}_2^D e \mathbf{V}_2^D , associados exclusivamente à porção ruidosa do sinal, sendo denominados, respectivamente, de valores e vetores singulares de ruído.

Podemos, portanto, dividir o espaço vetorial da matriz de dados \mathbf{D} em dois subespaços: subespaço de sinal e subespaço de ruído. Na verdade, tal denominação é imprecisa visto que o subespaço de sinal, assim definido, também está corrompido por ruído, sendo por essa razão também denominado de subespaço de sinal mais ruído. Tais subespaços não devem ser confundidos com aqueles antes definidos em relação à matriz \mathbf{R} , embora existam relações entre eles. Por esse motivo serão denominados a partir de então de subespaços de sinal e de ruído associados à matriz de dados.

Conhecendo-se estes subespaços é possível diminuir a influência do ruído no cálculo da solução ótima para o preditor através da sua extração. Portanto, a distinção entre esses subespaços forma a base fundamental dos métodos de predição linear baseados em subespaços que serão descritos mais adiante. Por enquanto, vamos apenas expressar a solução em (51) através dos elementos da SVD de \mathbf{D} . Pode ser demonstrado que a pseudo inversa de \mathbf{D} pode ser expressa em função de seus valores e vetores singulares, resultando em [10], [12], [33]

$$\hat{\mathbf{b}}_{PL} = \mathbf{D}^\dagger \mathbf{d} = \sum_{i=1}^p \frac{1}{\sigma_i} \mathbf{u}_i^D (\mathbf{v}_i^D)^H \mathbf{d}, \quad (55)$$

onde $p = \min(N-L, L)$.

6. MÉTODOS BASEADOS NA PREDIÇÃO LINEAR

Existem algumas alternativas para a estrutura do filtro preditor e, conseqüentemente, para a composição da ma-

triz de dados \mathbf{D} . O filtro do tipo progressivo (“forward”), dá origem à predição linear progressiva (FLP), que procura prever as amostras $\{y_n(k)\}_{n=L}^{N-1}$ a partir das amostras anteriores $\{y_{n-l}(k)\}_{n=L}^{N-1}$, para $l = 1, 2, \dots, L$, conforme (44). Já o filtro denominado de regressivo (“backward”), que dá origem à predição linear regressiva (BLP), estima as amostras $\{y_n(k)\}_{n=0}^{N-L-1}$ a partir das amostras posteriores $\{y_{n+l}(k)\}_{n=0}^{N-L-1}$, para $l = 1, 2, \dots, L$. Embora ambos os filtros utilizem as mesmas amostras do sinal disponível, o vetor de parâmetros que minimiza o erro de predição em cada caso é distinto [10], [11].

Combinando-se as formas de predição progressiva e regressiva, obtém-se uma terceira forma denominada de progressiva-regressiva (“forward-backward”), que dá origem à predição linear progressiva-regressiva (FBLP). Nesse caso, o critério de otimização consiste na minimização da soma das energias dos erros de predição progressivo e regressivo, impondo-se como restrição uma relação adequada entre as soluções de ambos os filtros. Esse procedimento é aquele que melhor utiliza as amostras disponíveis do sinal e, conseqüentemente, apresenta melhores resultados [10] e [11]. Como desvantagem, esse método está limitado às aplicações nas quais o sinal é composto por exponenciais complexas não amortecidas. Devemos observar que para o método FBLP, a matriz de dados possui dimensões $2(N-L) \times L$, enquanto que para os métodos FLP e BLP sua dimensão é $(N-L) \times L$ ([10], [11]).

As expressões para o método FBLP são, portanto, as mesmas da predição linear em (51) com a ordem do preditor satisfazendo $M \leq L \leq N - M/2$ e a estrutura da matriz de dados alterada para comportar os dados na forma progressiva e regressiva. Demonstra-se, experimentalmente, que tal método apresenta melhor desempenho quando $L \cong N/3$ ([10]).

6.1 MÉTODO FBLP MODIFICADO

Através da expressão (55) podemos observar que todos os valores e vetores singulares da matriz de dados \mathbf{D} , tanto os de sinal como os de ruído, participam da solução para o método FBLP.

O método FBLP modificado pode ser interpretado como uma versão do método FBLP baseada em subespaços, na qual se elimina a participação dos valores e vetores singulares de ruído na constituição da solução. Para isto, o extremo superior do somatório em (55) é feito igual a M , permitindo apenas a participação dos M valores e vetores singulares do subespaço de sinal, ou seja

$$\hat{\mathbf{b}}_{FB} = \sum_{i=1}^M \frac{1}{\sigma_{FB(i)}} \mathbf{v}_{FB(i)} \left(\mathbf{u}_{FB(i)}^H \mathbf{d}_{FB} \right),$$

com $b_0 = -1$. Dessa forma, tal método permite o aumento da ordem do preditor para valores além daqueles apropriados para método FBLP, melhorando assim sua capacidade de resolução. Este método apresenta seu melhor desempenho quando L é aproximadamente igual a $3N/4$ ([10], [11], [13]).

Podemos notar, ainda, que para o caso em que $L = N - M/2$ tal método é equivalente ao método FBLP devido às

dimensões da matriz \mathbf{D} correspondente. Este caso particular é denominado Kumaresan-Prony (KP) [10], [11], [12], [13], [42].

6.2 MÉTODO TLS-FBLP

O método, ou conjunto de métodos, aqui denominado de TLS-LP consiste em obter a solução para o preditor linear através do procedimento TLS de minimização [12], [42], [43].

O critério TLS, quando aplicado à predição linear, consiste em resolver o seguinte problema de minimização

$$\min_{\mathbf{C}(k), \mathbf{b}'} \|\mathbf{D}'(k) - \mathbf{C}(k)\|_F$$

sujeito a $\mathbf{C}(k)\mathbf{b}' = \mathbf{0}$ com $\|\mathbf{b}'\| \neq 0$, onde $\|\cdot\|_F$ denota a norma de Frobenius [44], [45]. Embora a matriz $\mathbf{C}(k)$ não seja importante para os estimadores descritos na seqüência, fornecemos aqui algumas de suas características. Demonstra-se em [44] e [45] que $\mathbf{C}(k)$, após a otimização, pode ser descrita a partir da decomposição de $\mathbf{D}'(k)$ em valores singulares: altera-se o valor do menor valor singular, anulando-o, e em seguida recompõe-se a matriz, resultando então $\mathbf{C}(k)$. O posto de $\mathbf{C}(k)$ será igual ao posto de $\mathbf{D}'(k)$ reduzido de uma unidade.

Assim como o procedimento LS aplicado sobre a predição linear progressiva, regressiva e progressiva-regressiva dá origem aos métodos FLP, BLP e FBLP, o procedimento TLS aplicado a essas formas de predição dá origem aos respectivos métodos TLS-FLP, TLS-BLP e TLS-FBLP [12], [46], [47].

Pela superioridade apresentada pela predição na forma progressiva-regressiva, apresentaremos o método TLS-FBLP, permitindo assim comparações com o método FBLP Modificado.

O método TLS-FBLP propõe uma modificação na solução de norma mínima apresentada pelo procedimento TLS clássico. Nessa modificação é realizada uma restrição adicional de subespaço sobre a matriz de dados aumentada [12].

Inicialmente consideremos a decomposição da matriz de dados aumentada da forma

$$\begin{aligned} [\mathbf{d} \mid \mathbf{D}] &= \mathbf{D}' = [\mathbf{D}_1 \mid \mathbf{D}_2] \\ &= [\mathbf{U}_1^D \mid \mathbf{U}_2^D] \begin{bmatrix} \Sigma_1^D & \mathbf{0} \\ \mathbf{0} & \Sigma_2^D \end{bmatrix} \begin{bmatrix} (\mathbf{V}_1^D)^H \\ (\mathbf{V}_2^D)^H \end{bmatrix}, \end{aligned} \quad (56)$$

na qual $\mathbf{D}_1 = \mathbf{U}_1^D \Sigma_1^D (\mathbf{V}_1^D)^H$ representa a partição de \mathbf{D}' associada aos seus M maiores valores singulares presentes na matriz (Σ_1^D) de dimensões $M \times M$ e $\mathbf{D}_2 = \mathbf{U}_2^D \Sigma_2^D (\mathbf{V}_2^D)^H$ representa sua porção complementar associada aos $(L - M + 1)$ menores valores singulares presentes na matriz (Σ_2^D) , de dimensões $(2(N-L) - M) \times (L - M + 1)$. Podemos assim, associar essas duas partições aos subespaço de sinal e de ruído da matriz \mathbf{D}' , respectivamente.

A solução de norma mínima fornecida pelo procedimento TLS clássico consiste no vetor singular à direita associado ao menor valor singular da matriz de dados aumentada. Se esse menor valor singular for um valor singular repetido, a solução

será dada, então, pela combinação dos vetores singulares associados aos menores valores singulares repetidos.

Em uma situação prática na qual o sinal é ruidoso, teremos os valores singulares em Σ_2^D distintos. Nesse caso, a solução TLS clássica será dada pelo vetor singular à direita associado ao menor valor singular de \mathbf{D}' .

A modificação sugerida pelo método TLS-FBLP sobre a solução TLS clássica se inspira no comportamento teórico da matriz \mathbf{D}' em que os valores singulares em Σ_2^D são repetidos e iguais ao desvio padrão do ruído branco presente no sinal. Esse resultado nos leva a desconsiderar as discrepâncias entre as magnitudes desses valores singulares.

Portanto, a solução sugerida para esse método será equivalente à solução TLS clássica para o caso em que os $\min[2(N-L) - M, L - M + 1]$ menores valores singulares de \mathbf{D}' ou Σ_2^D são repetidos, mesmo que na prática possuam diferenças em suas magnitudes. Considerando (56), tal solução é expressa por

$$\mathbf{b} = -\frac{\mathbf{y}}{\alpha} = -\frac{\mathbf{V}'_{D_2} \mathbf{P}^*}{\mathbf{P}^H \mathbf{P}}, \quad (57)$$

na qual utilizamos a partição $\mathbf{V}_{D_2} = \begin{bmatrix} \mathbf{P} & (\mathbf{V}'_{D_2})^T \end{bmatrix}^T$ e $(\cdot)^*$ denota o complexo conjugado [45], [46], [47], [48]. Considerando a SVD da matriz de dados aumentada, podemos observar que a solução em (57) é baseada no subespaço de ruído dessa matriz, enquanto que a solução apresentada pelo método FBLP Modificado é baseada no subespaço de sinal da matriz de dados original. Dessa forma, o método TLS-LP também incorpora em sua solução a distinção entre os subespaços.

Além das diferenças na concepção da solução obtida pelos procedimentos LS e TLS, os mesmos também diferem no ponto de vista computacional. O procedimento LS utilizado no método FBLP Modificado requer a obtenção de M vetores singulares da matriz de dados \mathbf{D}_{FBLP} , com dimensões $2(N-L) \times L$, onde M determina o número de parâmetros desconhecidos do sinal. Já o método TLS-FBLP opera sobre a matriz de dados aumentada $[\mathbf{d} \mid \mathbf{D}]$, de dimensões $2(N-L) \times (L+1)$, e requer a obtenção dos vetores singulares dessa matriz associados aos seus $\min[2(N-L) - M, L - M + 1]$ menores valores singulares. Portanto, para ordens elevadas, o procedimento LS será mais eficiente computacionalmente.

6.3 MÉTODO WTLS

O método WTLS utiliza o procedimento TLS de forma distinta dos métodos TLS-LP [46], [47]. O termo “W” em sua denominação denota “Whitened” ou “Weighted” do inglês.

Essa forma de ponderação consiste em impor uma característica de ruído branco ao erro de predição, em concordância com a propriedade “branqueadora” do filtro de erro de predição quando otimizado para os dados de entrada. Esse comportamento apesar de conhecido não foi utilizado pelos métodos baseados em predição linear apresentados anteriormente.

Sendo mais específico, considerando (48), podemos escrever o erro de predição como

$$\mathbf{e} = \mathbf{B}^H \mathbf{y} = \mathbf{B}^H (\mathbf{x} + \mathbf{n}). \quad (58)$$

onde \mathbf{B}^H é dada por (21), $\mathbf{n} = [n_0 \ n_1 \ \dots \ n_{N-1}]^T$ e $\mathbf{y} = [y_0 \ y_1 \ \dots \ y_{N-1}]^T$.

Considerando os dados livres de ruído, teremos o polinômio $b(z)$, de ordem $L \geq M$, com M de suas raízes localizadas sobre a CRU nas posições angulares dada por ϕ_m , $m = 1, 2, \dots, M$. Nessa condição teremos $\mathbf{B}^H \mathbf{x} = 0$ e, conseqüentemente, $\mathbf{e} = \mathbf{B}^H \mathbf{n}$. Através da suposição de que \mathbf{n} é um ruído tipo branco, e sabendo que para o filtro preditor operar em sua condição ótima devemos ter em sua saída um sinal tipo ruído branco, deveríamos esperar que o sinal de erro e também fosse branco. No entanto, $\mathcal{E}\{\mathbf{e}\mathbf{e}^H\} = \sigma^2 \mathbf{B}^H \mathbf{B}$. Portanto, é natural branquearmos o sinal de erro fazendo

$$\mathbf{e}_w = (\mathbf{B}^H \mathbf{B})^{-1/2} \mathbf{B}^H \mathbf{y} = (\mathbf{B}^H \mathbf{B})^{-1/2} \mathbf{D}' \mathbf{b}'. \quad (59)$$

Retornando ao problema original, devemos agora encontrar o vetor de norma-mínima \mathbf{b}' , com $b_0 = -1$ que minimiza a norma do erro em (59). Dessa forma, o método WTLS procura uma solução que minimiza a seguinte função custo

$$\begin{aligned} J_{WTLS}(\mathbf{b}) &= \{(\mathbf{e}_w)^H \mathbf{e}_w\} \\ &= \begin{bmatrix} -1 \\ \mathbf{b} \end{bmatrix}^H (\mathbf{D}')^H (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{D}' \begin{bmatrix} -1 \\ \mathbf{b} \end{bmatrix}. \end{aligned} \quad (60)$$

Podemos reescrever (60) da forma compacta

$$J_{WTLS}(\mathbf{b}') = \|\mathbf{D}(\mathbf{B}) \mathbf{b}'\|^2, \quad (61)$$

na qual $\mathbf{b}' = \begin{bmatrix} -1 & \mathbf{b}^T \end{bmatrix}^T$ e $\mathbf{D}(\mathbf{B}) = (\mathbf{B}^H \mathbf{B})^{-1/2} \mathbf{D}'$. Desde que a matriz $\mathbf{D}(\mathbf{B})$ seja fixa, a minimização de (61) poderá ser obtida através do procedimento TLS, como na fórmula original, ou mesmo do LS. Podemos notar que se fizermos $\mathbf{B}^H \mathbf{B} = \mathbf{I}$ em (61) teremos a mesma função custo proposta para os métodos clássicos baseados na predição linear em (49) [49]. Dessa forma, por apresentar um problema de minimização de quarta ordem, o método WTLS pode ser visto como uma generalização dos métodos baseados em predição linear. Esta generalização incorpora o termo $(\mathbf{B}^H \mathbf{B})^{-1}$ na função custo, corrigindo a deficiência, antes observada, causada pela ausência deste termo na função custo em (52).

Podemos notar que a forma compacta em (61) possui a mesma estrutura que aquela em (43) para o IQML. Dessa forma, a solução para o método WTLS-LP poderá ser implementada de forma semelhante ao algoritmo IQML, considerando $\mathbf{D}(\mathbf{B})$ fixa a cada iteração. A grande diferença entre os dois procedimentos consiste na forma de obtenção das estimativas finais. No método WTLS, por se tratar de um método baseado na predição linear de ordem L , as estimativas para os parâmetros ϕ_m , $m = 1, 2, \dots, M$, são dadas pelas posições angulares dos M zeros de $b(z)$ mais próximos da CRU. Através de simulações constata-se que os $L - M$ zeros restantes estão localizados fora da CRU, facilitando assim a distinção entre esses e os M zeros de sinal [46].

A partir de (60) é possível contrastar a função custo do método WTLS com a do método IQML em (30). Para o caso de um único experimento $K = 1$ e $L = M$, teremos $\mathbf{D}' = \mathbf{Y}$ e nesse caso as funções diferem apenas na restrição imposta sobre \mathbf{b}' , com $b_0 = -1$ no método WTLS. Como

foi enfatizado anteriormente, tal restrição não é utilizada no método IQML por ser incompatível com a restrição de simetria complexo conjugada. No método WTLS, a condição $b_0 = -1$ é oriunda da estrutura de um suposto filtro de erro de predição, que é a base de sua concepção. Portanto, apesar de possuírem formulações semelhantes, as soluções apresentadas pelos dois métodos são distintas, embora sejam ótimas sob cada critério. A formulação semelhante permite ainda afirmar que os métodos WTLS e IQML possuem teoricamente o mesmo potencial de resolução quando aplicados na estimação de frequências, com uma eventual vantagem do método WTLS sobre o IQML por permitir utilizar $L > M$ em suas formulações.

Para permitir uma comparação com os métodos baseados na matriz de correlação, podemos considerar $\hat{\mathbf{R}}_w = \mathbf{D}(\mathbf{B})^H \mathbf{D}(\mathbf{B})$ como uma matriz de correlação associada aos dados da matriz \mathbf{D}' através de uma transformação particular, e reescrever (60) como

$$J_{WTLS}(\mathbf{b}) = \begin{bmatrix} -1 \\ \mathbf{b} \end{bmatrix}^H \hat{\mathbf{R}}_w \begin{bmatrix} -1 \\ \mathbf{b} \end{bmatrix}. \quad (62)$$

A partir de (62) podemos estender a aplicação do método WTLS quando dispomos de mais de um experimento ($K > 1$), tomando a estimativa $\hat{\mathbf{R}}_w$ sobre todos os experimentos.

Concluimos, portanto, que a obtenção da solução para o método WTLS-LP consiste em um problema de minimização de quarta ordem que pode ser resolvido de forma iterativa utilizando o procedimento TLS de minimização. Devido à semelhança de formulação entre os métodos, tal procedimento pode ser então utilizado para obter a solução do método IQML ou do MODE, bem como ao contrário.

7. ESTIMAÇÃO TOTALMENTE BASEADA EM SUBESPAÇOS

Existem procedimentos de estimação baseados em subespaços que não possuem origem nem no conceito de predição nem no critério ML. Tais procedimentos, denominados aqui de estimação totalmente baseada em subespaços, consistem no uso de propriedades dos subespaço de sinal e de ruído da matriz \mathbf{R} .

O primeiro trabalho publicado a fazer uso dos subespaços foi o de Pisarenko em 1973 [15]. Porém, foi com o algoritmo MUSIC ("Multiple Signal Classification") que tal abordagem apresentou maior interesse [4], [50].

Enquanto que o algoritmo MUSIC e o método de Pisarenko se baseiam na ortogonalidade dos subespaços de sinal e de ruído, métodos mais recentes utilizam o critério "subspace fitting", o qual denominamos aqui de "distância" entre os subespaços [4], [14], [18], [19].

Por fim, o método ESPRIT tem como principal característica a exploração de uma propriedade da estrutura do arranjo de sensores. Entretanto, optamos por classificá-lo na categoria dos métodos totalmente baseados em subespaços em virtude da semelhança de sua formulação [51], [52].

7.1 O CRITÉRIO DE ORTOGONALIDADE BASEADO NA MATRIZ DE CORRELAÇÃO

Esse critério considera a dimensão do arranjo de sensores igual ao número de fontes de sinais mais uma unidade ($N = M + 1$), caracterizando assim o subespaço de ruído através do menor autovalor e correspondente autovetor de \mathbf{R} .

Considerando um sinal composto pela soma de M exponenciais complexas não-amortecidas na ausência de ruído, podemos verificar que

$$\mathbf{A}^H \mathbf{b}' = 0, \quad (63)$$

onde \mathbf{A} é definida em (3) e \mathbf{b}' , conforme (25) e (22), é um vetor N -dimensional, cujos elementos são os coeficientes de um polinômio $b(z)$ de ordem M que possui suas raízes sobre a CRU nas posições associadas às frequências das exponenciais.

Conseqüentemente, através das formas decompostas da matriz \mathbf{R} , apresentadas em (6) e (9) temos

$$\mathbf{R} \mathbf{b}' = 0. \quad (64)$$

Considerando agora o caso de um sinal ruidoso, o mesmo raciocínio leva a

$$\mathbf{R} \mathbf{b}' = \sigma^2 \mathbf{b}'. \quad (65)$$

As equações (64) e (65) serão satisfeitas se, e somente, se \mathbf{b}' for autovetor do subespaço de ruído de \mathbf{R} .

7.1.1 MÉTODO DE PISARENKO

O método de Pisarenko consiste na busca do vetor \mathbf{b}' que produz a melhor aproximação para (65) considerando um sinal com ruído e tomando a matriz estimada $\hat{\mathbf{R}}$.

Portanto, para $N - 1 = M$, o critério a ser minimizado nesse método é dado por

$$\min_{\mathbf{b}'} \frac{(\mathbf{b}')^H \hat{\mathbf{R}} (\mathbf{b}')}{\|\mathbf{b}'\|^2} = \sigma_{\min}^2, \quad (66)$$

que, no caso de não haver restrições para a escolha de \mathbf{b}' , fornece como solução o autovetor de $\hat{\mathbf{R}}$ correspondente ao seu menor autovalor. Assim, a partir da relação em (63) podemos obter estimativas para os parâmetros do sinal através das raízes do polinômio $b(z)$ formado pelos elementos do vetor \mathbf{b}' que satisfaz (66).

O método de Pisarenko é ótimo na situação assintótica com relação ao número de experimentos ($K \rightarrow \infty$) e, nesse caso, coincide com o resultado obtido através do critério ML [21].

Para permitir a comparação com os métodos baseados no critério ML, consideraremos a mesma matriz de reparametrização dada por \mathbf{B}^H em (21), com $N = M + 1$ e $L = M$. Nesse caso teremos $\mathbf{B} = \mathbf{b}'$ e poderemos transformar o problema em (66) no seguinte problema equivalente

$$\min_{\mathbf{b}'} \text{Tr} \left\{ \mathbf{B} (\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \hat{\mathbf{R}} \right\}. \quad (67)$$

Nesse caso, o problema enfrentado pelo método de Pisarenko coincide com o problema a ser solucionado pelo método IQML para o caso particular em que $N = M + 1$. Além

dessa limitação, o método de Pisarenko não sugere nenhuma restrição sobre os elementos de \mathbf{b}' , como nos métodos IQML e MODE.

Considerando ainda a relação $\hat{\mathbf{R}} = (\mathbf{D}')^H \mathbf{D}'$ para $N = L + 1$, e substituindo em (66), teremos

$$\min_{\mathbf{b}} \frac{\begin{bmatrix} b_0 \\ \mathbf{b} \end{bmatrix}^H (\mathbf{D}')^H \mathbf{D}' \begin{bmatrix} b_0 \\ \mathbf{b} \end{bmatrix}}{\left\| \begin{bmatrix} b_0 \\ \mathbf{b} \end{bmatrix} \right\|^2} = \min_{\mathbf{b}'} \frac{\|\mathbf{D}'\mathbf{b}'\|^2}{\|\mathbf{b}'\|^2}. \quad (68)$$

Através de (68), com $b_0 = -1$, podemos interpretar o método de Pisarenko como um caso particular da predição linear, quando a ordem do filtro é igual ao número de amostras disponíveis menos a unidade ($L = N - 1$) e, ainda, $L = M$. Nessa comparação constatamos a vantagem da predição linear pela possibilidade de utilizar $L > M$, desde que $N > L$.

7.2 O CRITÉRIO DE ORTOGONALIDADE ENTRE SUBESPAÇOS

Considerando as decomposições da matriz \mathbf{R} em (6) e (9), tal critério usa o fato de que os autovetores associados ao subespaço de ruído, dados pelas colunas de \mathbf{U}_n , são ortogonais aos autovetores associados ao subespaço de sinal, dados pelas colunas de \mathbf{U}_s .

Considerando, ainda, a hipótese de posto cheio para a matriz \mathbf{P} , teremos as colunas de $A(\Phi)$ ortogonais aos autovetores do subespaço de ruído de \mathbf{R} . Podemos, então, escrever a seguinte expressão

$$\mathbf{U}_n^H \mathbf{a}(\phi_m) = \mathbf{0}, \text{ para } m = 1, 2, \dots, M. \quad (69)$$

a qual pode ser utilizada para estimar os parâmetros ϕ_m , para $m = 1, 2, \dots, M$, desde que \mathbf{U}_n^H possa ser adequadamente estimada. Neste caso, tal procedimento considera a relação de ortogonalidade em (69) de forma aproximada, ou seja, $\hat{\mathbf{U}}_n^H \mathbf{a}(\phi_m) \approx \mathbf{0}$.

No contexto de uma otimização baseada nos mínimos-quadrados, esse procedimento consistirá no seguinte problema de minimização

$$\hat{\phi}_m = \arg \left\{ \min_{\varphi} \left\| \hat{\mathbf{U}}_n^H \mathbf{a}(\varphi) \right\| \right\}, \quad (70)$$

para $m = 1, 2, \dots, M$, e onde φ representa a variável independente no processo de minimização.

Um procedimento de estimação que considera a relação em (69) para todas as colunas de $A(\Phi)$ simultaneamente, consiste no seguinte problema de otimização de segunda ordem

$$\hat{\Phi} = \arg \left\{ \min_{\varphi} \left\| \hat{\mathbf{U}}_n^H \mathbf{A}(\varphi) \right\|_F^2 \right\}, \quad (71)$$

o qual exige uma busca M -dimensional dos parâmetros.

Utilizando a relação $\|\mathbf{M}\|_F^2 = \text{Tr} \{ \mathbf{M}\mathbf{M}^H \}$ em (71), teremos o seguinte problema equivalente

$$\hat{\Phi} = \arg \left\{ \min_{\varphi} \text{Tr} \left\{ \hat{\Pi}_{\hat{\mathbf{A}}} \left(\mathbf{A}(\varphi) \mathbf{A}^H(\varphi) \right) \right\} \right\}, \quad (72)$$

no qual utilizamos a igualdade $\hat{\Pi}_{\hat{\mathbf{A}}} = \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H$.

A partir da função custo derivada para o critério ML baseado em subespaço em (20), considerando $\hat{\mathbf{R}}_s = \hat{\mathbf{U}}_s \hat{\Lambda}_s \hat{\mathbf{U}}_s^H$, podemos escrever

$$\hat{\Phi} = \arg \left\{ \min_{\varphi} \text{Tr} \left\{ \Pi_{\hat{\mathbf{A}}}^{\perp}(\varphi) \hat{\mathbf{R}}_s \right\} \right\}. \quad (73)$$

Através da comparação entre (72) e (73), observamos que ambas buscam os parâmetros $\hat{\phi}_m$, $m = 1, 2, \dots, M$, que melhor satisfazem a ortogonalidade entre o subespaço de sinal, representado pelo produto $\mathbf{A}(\varphi) \mathbf{A}^H(\varphi)$ em (72) e por $\hat{\mathbf{R}}_s$ em (73), e o subespaço de ruído, representado pelo projetor ortogonal $\Pi_{\hat{\mathbf{A}}}^{\perp}(\varphi)$. A diferença básica entre as formas de obtenção das estimativas é que (72) é baseada no subespaço de ruído enquanto que o critério ML baseia-se no subespaço de sinal, uma vez que a estimação em (72) se dá em relação a $\mathbf{A}(\Phi)$ e não em relação a $\Pi_{\hat{\mathbf{A}}}^{\perp}$, como no caso do critério ML em (20) ou (73). Por esse motivo, $\Pi_{\hat{\mathbf{A}}}^{\perp}$ é estimado através do subespaço de ruído. No caso de sinais independentes e de mesma potência, teremos $\mathbf{P} = \mathbf{I}$ e, conseqüentemente, teremos $\mathbf{R}_s = \mathbf{A} \mathbf{A}^H$ e neste caso os dois critérios serão assintoticamente equivalentes.

Uma das limitações desse critério é a incapacidade de resolver sinais próximos entre si, sobretudo em baixas relações sinal-ruído [39], [53]. Além disso, devemos enfatizar que as matrizes \mathbf{A} e \mathbf{U}_s geram o mesmo espaço apenas quando \mathbf{P} tem posto cheio. Em um caso geral, quando o posto de \mathbf{P} é igual a $M' < M$, a matriz \mathbf{U}_s gera um subespaço M' -dimensional da matriz \mathbf{A} . Este é o caso de sinais coerentes, onde a propriedade de posto cheio da matriz \mathbf{P} e, conseqüentemente, de $\mathbf{A} \mathbf{P} \mathbf{A}^H$, é violada e o critério falha na obtenção de estimativas consistentes. Nesse caso, a redução no posto de \mathbf{P} dificulta a distinção entre o subespaço de sinal e de ruído e, conseqüentemente, prejudica a obtenção de estimativas adequadas para $\hat{\mathbf{U}}_n$ [54]. Para superar essa desvantagem, uma matriz positiva definida de ponderações pode ser introduzida na função custo a ser minimizada, visando controlar a participação de cada autovetor de $\hat{\mathbf{U}}_n$, no intuito de compensar a correlação existente entre os sinais [16], [53], [55], [56].

Considerando a reparametrização em (24) e fazendo $\hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H = \hat{\Pi}_{\mathbf{B}}$, a estimativa do projetor ortogonal poderá ser obtida a partir de uma estimativa dos coeficientes b_k da predição linear. Além disso, uma outra variação para o problema em (72) consiste na utilização de uma estimativa do subespaço de sinal em substituição à estimativa do subespaço de ruído, na qual consideramos $\hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H = \mathbf{I} - \hat{\mathbf{U}}_s \hat{\mathbf{U}}_s^H$.

7.2.1 MÉTODO MUSIC

Utilizando os conceitos apresentados anteriormente, o método MUSIC tem por objetivo formar uma função tipo espectral definida como

$$P_{MUSIC}(\varphi) = \frac{1}{\|\hat{\mathbf{U}}_n^H \mathbf{a}(\varphi)\|^2}, \quad (74)$$

a qual apresenta picos nas proximidades dos parâmetros ϕ_m do sinal observado, para $m = 1, 2, \dots, M$, possibilitando assim a obtenção de estimativas para esse parâmetro. Tal forma

de obtenção dos parâmetros DOA dá origem ao método MUSIC Espectral, ou método MUSIC baseado no espectro [16] e [17].

Uma forma alternativa e mais vantajosa consiste na busca das raízes de um polinômio associado ao denominador da função $P_{MUSIC}(\varphi)$ em (74), tal que suas raízes estejam relacionadas com os picos da função espectral. Essa forma paramétrica de estimação dá origem ao método root-MUSIC.

Mais especificamente, seja o polinômio

$$b(z) = z^{N-1} \mathbf{b}^H(z) \mathbf{U}_n \mathbf{U}_n^H \mathbf{b}(z), \quad (75)$$

no qual $\mathbf{b}(z) = [1 \ z \ \dots \ z^{N-1}]^T$ com $z = e^{j\varphi}$. Podemos observar que para $z = e^{j\phi_m}$, teremos $\mathbf{b}(z) = \mathbf{a}(\phi_m)$ e, portanto, (80) é equivalente a (74), apresentando picos espectrais em $\varphi = \phi_m$, para $m = 1, 2, \dots, M$. O polinômio resultante é dado, portanto, pela combinação linear dos autovetores associados ao subespaço de ruído da matriz de correlação do sinal.

Podemos observar que para $N - 1 = M$, teremos apenas um autovetor associado ao subespaço de ruído e neste caso este método coincide com o método de Pisarenko [16]. [17].

Uma outra variação para o método MUSIC resulta da minimização direta do denominador de (74), como em (71), dando origem ao problema de minimização

$$\hat{\Phi} = \arg \left\{ \min_{\varphi} \text{Tr} \left\{ \mathbf{A}(\varphi) \mathbf{A}^H(\varphi) \left(\hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H \right) \right\} \right\}, \quad (76)$$

para $m = 1, 2, \dots, M$. Por considerar a obtenção das estimativas $\hat{\phi}_m$, para $m = 1, 2, \dots, M$, de forma simultânea o procedimento de minimização em (76) dá origem ao método denominado de MUSIC multidimensional (MD-MUSIC), notadamente mais complexo que o MUSIC clássico.

Uma extensão do MUSIC, denominada de MUSIC Ponderado, faz uso de uma matriz de ponderação \mathbf{W}_M de dimensões $(N - M) \times (N - M)$, que pondera a influência de cada um dos autovetores de forma distinta, levando à composição da função custo

$$J_{MW}(\varphi) = \text{Tr} \left\{ \mathbf{A}^H(\varphi) \left(\hat{\mathbf{U}}_n \mathbf{W}_M \hat{\mathbf{U}}_n^H \right) \mathbf{A}(\varphi) \right\}. \quad (77)$$

Uma outra variação para o MUSIC Ponderado faz uso da seguinte função custo

$$J_{MV}(\varphi) = \text{Tr} \left\{ \mathbf{A}(\varphi) \mathbf{V}_M \mathbf{A}^H(\varphi) \left(\hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H \right) \right\}. \quad (78)$$

onde \mathbf{V}_M é uma matriz que pondera as colunas (vetores de direção angular) de $\mathbf{A}(\varphi)$.

Embora o método MUSIC original apresente uma menor variância assintótica que as suas versões ponderadas, em situações em que o número de amostras é pequeno, a relação sinal-ruído é baixa e os sinais são correlacionados, podemos obter uma melhor capacidade de resolução escolhendo adequadamente os elementos da matriz \mathbf{W}_M ou \mathbf{V}_M [4].

Para a comparação entre o método MUSIC e os métodos baseados no critério ML, como o IQML e MODE, devemos considerar a reparametrização $\hat{\mathbf{\Pi}}_B = \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H$ e $N - 1 = M$ em (78). Nesse caso, podemos observar que a escolha $\mathbf{V}_M = \mathbf{P}$ nos permite reescrever (78) como

$$J_{MV}(\Phi) = \text{Tr} \left\{ \hat{\mathbf{\Pi}}_B \tilde{\mathbf{R}} \right\}, \quad (79)$$

na qual utilizamos $\mathbf{A}(\Phi) \mathbf{P} \mathbf{A}^H(\Phi) = \mathbf{R} - \mathbf{I}\sigma^2 = \tilde{\mathbf{R}}$ (ver expressão (6)). Desta forma podemos constatar que o método MUSIC Ponderado para uma escolha particular da matriz de ponderações, possui uma formulação semelhante a do método MODE, no qual a matriz de correlação é substituída por uma versão truncada baseada apenas no subespaço de sinal.

Para estabelecer comparações com os métodos baseados na predição linear deveremos considerar o método root-MUSIC em (75). Nesse caso, podemos considerar um preditor de ordem $L = N - 1$ associado ao polinômio $b(z)$ em (75).

O desempenho do método MUSIC cresce assintoticamente com o aumento do número de amostras disponíveis para alta relação sinal-ruído [39] e [53]. Além disso, a sua formulação original em (74) requer um esforço computacional inferior quando comparado com os métodos baseados no critério ML.

7.2.2 MÉTODO DA NORMA-MÍNIMA

O princípio do método da Norma-mínima é procurar estimativas dos parâmetros DOA através da localização dos picos da seguinte função espectral

$$P_{NM}(\varphi) = \frac{1}{\left\| (\mathbf{U}_n \mathbf{v})^H \mathbf{b}(z) \right\|^2}, \quad (80)$$

na qual \mathbf{v} é um vetor unitário tal que $\mathbf{b}' = \mathbf{U}_n \mathbf{v}$ apresenta a menor norma-2 e $\mathbf{b}(z) = [1 \ z \ \dots \ z^{N-1}]^T$ [4], [55].

Uma vez que \mathbf{b}' é dado por uma combinação linear dos autovetores associados ao subespaço de ruído, teremos, conseqüentemente, \mathbf{b}' ortogonal ao subespaço de sinal, gerado pelas colunas de \mathbf{U}_s e, portanto, satisfazendo a relação

$$\mathbf{U}_s^H \mathbf{b}' = 0. \quad (81)$$

Para o caso de dispormos apenas de uma estimativa de \mathbf{U}_s , a solução de (81), considerando o critério LS de otimização, é, portanto, obtida através do seguinte problema de minimização

$$\mathbf{b}' = \arg \left\{ \min \left\| \hat{\mathbf{U}}_s^H \mathbf{b}' \right\| \right\}, \quad (82)$$

tal que a norma $\|\mathbf{b}'\|$ seja mínima e tomando-se $b_0 = -1$ para evitar a solução trivial. Esta escolha permite ainda relacionar esse método com os métodos baseados na predição linear.

Definindo $\mathbf{b}' = [-1 \ \mathbf{b}]^T$, a solução fechada para o vetor \mathbf{b} que minimiza a função custo em (82) tal que a norma $\|\mathbf{b}'\|$ seja mínima, pode ser obtida através do procedimento LS de otimização e é dada por [42], [48]

$$\mathbf{b} = \left(\hat{\mathbf{U}}_s' \right)^\dagger \mathbf{g} = -\frac{\hat{\mathbf{U}}_s' \mathbf{g}}{1 - \mathbf{g}^H \mathbf{g}}, \quad (83)$$

onde $\left(\hat{\mathbf{U}}_s' \right)^\dagger$ designa a pseudo-inversa de $\hat{\mathbf{U}}_s'$, conforme definição em (11), e $\hat{\mathbf{U}}_s'$ se origina da seguinte decomposição para $\hat{\mathbf{U}}_s$

$$\hat{\mathbf{U}}_s = \begin{bmatrix} \mathbf{g}^T \\ \hat{\mathbf{U}}_s' \end{bmatrix}.$$

Após obter o vetor solução \mathbf{b}' , o método da Norma-mínima pode obter as estimativas para o parâmetro ϕ_m do sinal, para

$m = 1, 2, \dots, M$, através dos picos da função em (80) ou, preferencialmente, através das raízes do polinômio $b(z) = (\mathbf{b}')^T \mathbf{b}(z)$.

A solução de norma-mínima (83) faz com que M dos N zeros de $b(z)$ se situem sobre a circunferência de raio unitário nas posições angulares dadas pelos parâmetros ϕ_m do sinal, e ao mesmo tempo reduz a magnitude dos $N - M$ zeros restantes ao valor mínimo possível. Essa característica é a responsável pela capacidade de resolução desse método.

Visto que \mathbf{b}' é dado pela combinação linear dos autovetores do subespaço de ruído, podemos interpretar o método da Norma-mínima como um caso particular do método root-MUSIC em que o vetor solução em (75) é escolhido com a restrição de norma-2 mínima.

A relação com os métodos baseados na predição linear se faz associando o vetor \mathbf{b}' ao vetor dos coeficientes de um filtro preditor de ordem $L = N - 1$.

A partir de (82) podemos construir o seguinte problema equivalente

$$\hat{\mathbf{b}} = \arg \left\{ \min_{\mathbf{b}} \left\{ \begin{bmatrix} -1 \\ \mathbf{b} \end{bmatrix}^H \tilde{\mathbf{R}}_s \begin{bmatrix} -1 \\ \mathbf{b} \end{bmatrix} \right\} \right\}, \quad (84)$$

no qual, dentre todas as possíveis soluções, devemos escolher aquela de menor norma. Também, interpretamos $\tilde{\mathbf{R}}_s = \hat{\mathbf{U}}_s \hat{\mathbf{U}}_s^H$ como uma versão da matriz $\hat{\mathbf{R}}_s$ tal que $\hat{\Lambda}_s = \mathbf{I}$. Tal estimativa é válida para os casos em que as fontes de sinais são descorrelacionadas e possuem mesma potência.

Através de (84) podemos, então, interpretar o método da Norma-mínima como um procedimento de predição linear de ordem $L = N - 1$ na qual a matriz de correlação estimada, dada anteriormente pelo produto $\mathbf{Y}^H \mathbf{Y}$ em (52), é substituída por uma versão truncada baseada apenas no subespaço de sinal, dada por $\tilde{\mathbf{R}}_s$.

7.3 O CRITÉRIO DE “DISTÂNCIA” ENTRE SUBESPAÇOS (“SUBSPACE FITTING”)

Mais recentemente que os procedimentos baseados no critério de ortogonalidade entre os subespaços descritos anteriormente, surgiu uma nova categoria de estimadores baseados no que denominamos aqui de critério de “distância” entre subespaços (“Subspace Fitting”) [4], [9], [18], [20].

Uma medida de “distância” entre o subespaço de interesse e sua aproximação é obtida, geralmente, como um erro de aproximação entre duas matrizes [19], [20]. Apresentaremos, a seguir, essa classe de estimadores através de uma reparametrização adequada do critério ML.

Comparando (6) com (9) e lembrando que $\Lambda_n = \sigma^2 \mathbf{I}$, teremos

$$\mathbf{A} \mathbf{P} \mathbf{A}^H + \sigma^2 \mathbf{I} = \mathbf{U}_s \Lambda_s \mathbf{U}_s^H + \sigma^2 \mathbf{U}_n \mathbf{U}_n^H. \quad (85)$$

Pós-multiplicando por \mathbf{U}_s pela direita e rearranjando os termos, teremos a relação

$$\mathbf{U}_s = \mathbf{A}(\Phi) \mathbf{T}. \quad (86)$$

na qual $\mathbf{T} = \mathbf{P} \mathbf{A}^H \mathbf{U}_s (\Lambda_s - \sigma^2 \mathbf{I})^{-1}$ é uma matriz de posto cheio. Visto que os parâmetros ϕ_m , para $m = 1, \dots, M$, e

\mathbf{T} são desconhecidos, é natural procurar seus valores que satisfazem (86). Se uma estimativa $\hat{\mathbf{U}}_s$ de \mathbf{U}_s for utilizada, tal solução não existirá. Neste caso, deveremos minimizar alguma “distância” entre $\hat{\mathbf{U}}_s$ e $\mathbf{A} \mathbf{T}$. Para esse propósito a norma de Frobenius será utilizada.

As estimativas são, então, obtidas resolvendo o seguinte problema de minimização

$$\left\{ \hat{\Phi}, \hat{\mathbf{T}} \right\} = \arg \left\{ \min_{\varphi, \mathbf{T}} \left\| \hat{\mathbf{U}}_s - \mathbf{A}(\varphi) \mathbf{T} \right\|_F^2 \right\}, \quad (87)$$

para $m = 1, 2, \dots, M$.

Similar ao problema enfrentado para o critério ML, esse é um problema de mínimos quadrados não-linear separável, cuja solução para o parâmetro linear $\hat{\mathbf{T}}$, com \mathbf{A} fixo, é dada por [33]

$$\hat{\mathbf{T}} = \mathbf{A}^\dagger \hat{\mathbf{U}}_s. \quad (88)$$

Substituindo (88) em (87) e utilizando a relação $\|\mathbf{M}\|_F^2 = \text{Tr} \{ \mathbf{M} \mathbf{M}^H \}$, obtemos a seguinte função custo a ser minimizada para essa categoria de estimadores

$$\begin{aligned} J_{SF}(\varphi) &= \text{Tr} \left\{ \left(\mathbf{I} - \mathbf{A} \mathbf{A}^\dagger \right) \hat{\mathbf{U}}_s \hat{\mathbf{U}}_s^H \right\} \\ &= \text{Tr} \left\{ \Pi_{\hat{\mathbf{A}}} \hat{\mathbf{U}}_s \hat{\mathbf{U}}_s^H \right\}. \end{aligned} \quad (89)$$

É interessante constatar que a estimação pela minimização de (89) se dá em relação a $\Pi_{\hat{\mathbf{A}}}$ como no caso do critério ML em (18).

De uma forma geral, reconhecendo a igualdade $\hat{\Pi}_{\mathbf{A}} = \hat{\mathbf{U}}_s \hat{\mathbf{U}}_s^H$, poderemos reescrever (89) como

$$J_{SF}(\varphi) = \text{Tr} \left\{ \Pi_{\hat{\mathbf{A}}} \hat{\Pi}_{\mathbf{A}} \right\}. \quad (90)$$

Uma variação para o problema em (87) pode surgir se utilizarmos estimativas do subespaço de ruído em sua formulação através da relação $\hat{\mathbf{U}}_s \hat{\mathbf{U}}_s^H = \mathbf{I} - \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H$, dando origem à seguinte função custo [4]

$$\begin{aligned} J_{NSF}(\varphi) &= \text{Tr} \left\{ \Pi_{\mathbf{A}} \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H \right\} \\ &= \text{Tr} \left\{ \Pi_{\mathbf{A}} \hat{\Pi}_{\mathbf{A}}^\perp \right\}. \end{aligned} \quad (91)$$

Podemos notar que em (91) a estimação em se dá em relação a $\Pi_{\mathbf{A}}$, a partir de uma estimativa de $\hat{\Pi}_{\hat{\mathbf{A}}} = \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H$ como no caso do critério de ortogonalidade em (72). Dessa forma, podemos interpretar o critério de ortogonalidade como um caso particular do critério de “distância” entre subespaços.

Devemos lembrar que as matrizes \mathbf{A} e \mathbf{U}_s geram o mesmo espaço apenas quando \mathbf{P} tem posto cheio. Caso contrário uma matriz positiva definida de ponderação pode ser adicionada ao problema [4]. A utilização de uma matriz de ponderação permite favorecer a escolha dos autovetores em $\hat{\mathbf{U}}_s$ mais representativos do sinal, além de reduzir a dimensão do problema original.

A adição de uma matriz de ponderação \mathbf{W} no problema em (89) leva à seguinte função custo [4], [20].

$$J_{WSF}(\varphi) = \text{Tr} \left\{ \Pi_{\hat{\mathbf{A}}} \hat{\mathbf{U}}_s \mathbf{W} \hat{\mathbf{U}}_s^H \right\}. \quad (92)$$

Particularmente, escolhendo $\mathbf{W} = \hat{\Lambda}_s$ em (92) podemos relacionar esse critério com o critério ML baseado em

subespaço em (20). Para esse caso podemos constatar a equivalência entre os dois critérios. Portanto, para o caso em que as fontes de sinais possuem potências semelhantes e são descorrelacionadas poderemos ainda considerar $\hat{\Lambda}_s = \mathbf{I}$ e nesse caso os critérios apresentam desempenho equivalentes. Essa constatação faz parte das contribuições desse trabalho e justifica o desempenho equivalente entre os dois critérios.

A função custo em (92) generaliza, portanto, os métodos paramétricos baseados em subespaços, assim como (18) generaliza os métodos baseados no critério ML e (61) generaliza os métodos baseados na predição linear.

Devemos enfatizar que os procedimentos baseados em subespaços também podem ser reparametrizados através da substituição do projetor $\Pi_{\hat{\Lambda}}^{\perp}$ pelo projetor equivalente dado por $\Pi_{\mathbf{B}}$, permitindo assim a sua comparação com os procedimentos baseados na predição linear através de um preditor de ordem $L = M$ e considerando a restrição $b_0 = -1$ ao polinômio $b(z)$.

Os métodos apresentados a seguir fazem uso deste critério e apresentam bom desempenho em situações que envolvem um número finito de amostras e possíveis correlações entre as fontes de sinais, equiparando-se aos melhores métodos.

7.3.1 MÉTODO ESPRIT

O método ESPRIT surgiu antes do conceito de "distância" entre subespaços. Optamos por sua classificação como um método baseado nesse critério devido à semelhança de sua formulação [51], [52], [58].

Este método pode ser interpretado como a utilização de dois arranjos de sensores, lineares e uniformemente espaçados, separados por uma distância igual ao espaçamento entre seus sensores. Na prática não existe a necessidade de utilizar dois arranjos fisicamente separados, pois, através da geometria do arranjo, os elementos podem ser selecionados conforme essa propriedade [4], [51] e [58].

O método usa a estrutura da matriz de *Vandermonde* \mathbf{A} para definir as submatrizes \mathbf{A}_1 e \mathbf{A}_2 através da extração, respectivamente, da última e da primeira linha de \mathbf{A} , isto é

$$\mathbf{A}_{N \times M} = \begin{bmatrix} \mathbf{A}_1 \\ \text{última linha} \\ \mathbf{A}_2 \end{bmatrix} = \begin{bmatrix} \text{primeira linha} \\ \mathbf{A}_2 \end{bmatrix}. \quad (93)$$

Dessa forma, um arranjo linear uniforme com N elementos espaçados entre si por uma distância d pode ser visto como dois arranjos com $N - 1$ elementos. Um dos arranjos é composto pelos primeiros $N - 1$ elementos do arranjo inicial, enquanto que o outro é formado pelos $N - 1$ últimos elementos. Neste contexto, vamos considerar os pares de sensores compostos pelo primeiro e o segundo elemento do arranjo inicial, o segundo e o terceiro elemento e assim sucessivamente.

Se considerarmos os sinais recebidos de uma fonte na direção θ_m pelo n 'ésimo par, descritos aqui, de acordo com (1), como $x_{1,n}(k)$ e $x_{2,n}(k)$, a diferença entre esses dois sinais será dada por

$$x_{2,n}(k) = x_{1,n}(k)e^{j2\pi\Delta_o \cos \theta_m} \quad (94)$$

onde $\Delta_o = d/\lambda_c$, representa a distância entre os dois elementos medida em comprimentos de onda. Conseqüentemente, considerando todos os pares e a presença de ruído, os sinais

fornecidos pelos dois arranjos, de acordo com (3), são dados por

$$\mathbf{y}_1(k) = \mathbf{A}_1 \mathbf{s}(k) + \mathbf{n}_1(k) \quad (95)$$

e

$$\mathbf{y}_2(k) = \mathbf{A}_1 \beta \mathbf{s}(k) + \mathbf{n}_2(k), \quad (96)$$

onde \mathbf{A}_1 é uma matriz de *Vandermonde* correspondente ao primeiro subarranjo, e β uma matriz diagonal com elementos dados por

$$\beta_{mm} = e^{j2\pi\Delta_o \cos \theta_m}, \quad m = 1, 2, \dots, M. \quad (97)$$

Comparando os sinais $\mathbf{y}_1(k)$ e $\mathbf{y}_2(k)$, podemos concluir que a matriz dos vetores de direção para o segundo subarranjo é dada por $\mathbf{A}_2 = \mathbf{A}_1 \beta$.

Considere as matrizes \mathbf{J}_0 e \mathbf{J}_1 com dimensões $(N - 1) \times N$, compostas por "1s" e "0s" e tais que $\mathbf{J}_0 \mathbf{A} = \mathbf{A}_1$ e $\mathbf{J}_1 \mathbf{A} = \mathbf{A}_1 \beta$. Definindo

$$\mathbf{J} \triangleq \begin{bmatrix} \mathbf{J}_0 \\ \mathbf{J}_1 \end{bmatrix} \quad (98)$$

podemos escrever

$$\mathbf{J} \mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_1 \beta \end{bmatrix}. \quad (99)$$

Conforme apresentado na seção 2.1, quando as matrizes \mathbf{P} , definida em (5), e \mathbf{A} têm posto cheio, as colunas de \mathbf{A} e de \mathbf{U}_s (definida em (9)), geram o mesmo subespaço de sinal. Neste caso existe uma matriz \mathbf{T}_A com dimensões $M \times M$, não singular, tal que $\mathbf{U}_s = \mathbf{A} \mathbf{T}_A$. Utilizando a equação (98), podemos então escrever

$$\mathbf{J} \mathbf{U}_s = \mathbf{J} \mathbf{A} \mathbf{T}_A. \quad (100)$$

Definindo

$$\mathbf{J} \mathbf{U}_s \triangleq \begin{bmatrix} \mathbf{U}_{s1} \\ \mathbf{U}_{s2} \end{bmatrix} \quad (101)$$

e usando as expressões (99) e (100), concluímos que

$$\mathbf{U}_{s1} = \mathbf{A}_1 \mathbf{T}_A \quad \text{e} \quad \mathbf{U}_{s2} = \mathbf{A}_1 \beta \mathbf{T}_A. \quad (102)$$

Combinando estas duas últimas relações obtemos

$$\mathbf{U}_{s1} \mathbf{T}_U = \mathbf{U}_{s2}, \quad (103)$$

na qual \mathbf{T}_U é uma matriz de transformação não-singular única, definida como $\mathbf{T}_U = (\mathbf{T}_A)^{-1} \beta \mathbf{T}_A$. Dessa forma as matrizes \mathbf{T}_U e β estão relacionadas por uma transformação de similaridade.

Considerando as estimativas $\hat{\mathbf{U}}_{s1}$ e $\hat{\mathbf{U}}_{s2}$, teremos a relação aproximada

$$\hat{\mathbf{U}}_{s1} \mathbf{T}_U \approx \hat{\mathbf{U}}_{s2}. \quad (104)$$

O método ESPRIT consiste na busca dos elementos de \mathbf{T}_U que melhor satisfaçam (104). Isto é conseguido através da minimização da norma do erro da relação aproximada em (104), ou seja,

$$\hat{\mathbf{T}}_U = \arg \left\{ \min_{\mathbf{T}_U} \left\| \hat{\mathbf{U}}_{s1} \mathbf{T}_U - \hat{\mathbf{U}}_{s2} \right\|_F \right\}. \quad (105)$$

Podemos notar que o problema enfrentado pelo método ESPRIT em (105) possui a mesma formulação para o problema em (87), justificando assim a classificação do método ESPRIT como baseado no critério de “distância” entre subespaços.

A solução para o problema em (105) pode ser obtida através do procedimento LS ou TLS, dando origem aos respectivos métodos LS-ESPRIT e TLS-ESPRIT. Os parâmetros DOA são, então, obtidos a partir dos autovalores de $\hat{\mathbf{T}}_L$.

Devido à flexibilidade na partição do arranjo através da matriz \mathbf{A} , outras variações para o método ESPRIT são possíveis [4], [51], [58].

7.3.2 MÉTODO “SUBSPACE FITTING PONDERADO” - WSF

O método precursor do WSF denominado de SSF (“Signal Subspace Fitting”) consiste na utilização direta do critério de “distância” dado pela função custo em (89) [4], [19].

Um outro método denominado de NSF (“Noise Subspace Fitting”) surge da implementação direta do critério de “distância” considerando uma aproximação baseada no subespaço de ruído, cuja função custo é dada por (91) [4]. Podemos notar, ainda, que esta função custo possui a mesma formulação daquela encontrada para o método MUSIC Ponderado para um caso particular em que $\mathbf{V}_M = (\mathbf{A}^H \mathbf{A})^{-1}$ em (78).

Observa-se que o método NSF, assim como o método SSF, apresentam estimativas pouco confiáveis em um cenário em que as fontes de sinais são coerentes.

O método WSF surge para contornar essas dificuldades através da implementação da generalização para o critério de “distância” entre subespaços derivado em (92), o qual consiste em adicionar uma matriz positiva definida de ponderação \mathbf{W}_{SSF} , levando ao seguinte problema de minimização [4], [20]

$$\hat{\Phi} = \arg \left\{ \min_{\varphi} Tr \left\{ \Pi_{\hat{\mathbf{A}}}(\varphi) \hat{\mathbf{U}}_s \mathbf{W}_{SSF} \hat{\mathbf{U}}_s^H \right\} \right\}. \quad (106)$$

Podemos notar que, para uma escolha adequada da matriz \mathbf{W}_{SSF} , o problema de minimização em (106) é equivalente ao problema encontrado para os métodos baseados no ML, sobretudo o método MODE, se considerarmos a reparametrização $\Pi_{\hat{\mathbf{A}}} = \Pi_{\mathbf{B}}$ [18]. Por esse motivo, as considerações com respeito à melhor escolha para a matriz \mathbf{W}_{SSF} seguem os resultados apresentados para a matriz de ponderações \mathbf{W}_{MODE} em (39). Além disso, quando utilizamos a relação $\Pi_{\hat{\mathbf{A}}} = \Pi_{\mathbf{B}}$ em (106), surge o método denominado de root-WSF. Este consiste na solução do método WSF através da busca das raízes do polinômio $b(z)$ formado pelos elementos da matriz de reparametrização \mathbf{B}^H . Neste caso, a utilização do algoritmo de solução proposto para o método MODE se aplica diretamente.

Embora não tenhamos conhecimento de tal proposta, a consideração de uma matriz de ponderação também pode ser estendida para o método NSF, dando origem ao problema de minimização

$$\hat{\Phi} = \arg \left\{ \min_{\varphi} Tr \left\{ \Pi_{\hat{\mathbf{A}}}(\varphi) \hat{\mathbf{U}}_n \mathbf{W}_{NSF} \hat{\mathbf{U}}_n^H \right\} \right\}. \quad (107)$$

Podemos notar que, para uma escolha adequada da matriz \mathbf{W}_{NSF} , o problema de minimização em (107) é equivalente ao problema encontrado para o método MUSIC Ponderado em (77) com $\mathbf{W}_M = \mathbf{W}_{NSF}$.

Podemos, então, concluir que os métodos baseados em subespaços podem ser descritos através de uma mesma formulação geral e que esta, por sua vez, permite explicitar comparações com os métodos baseados no critério ML. A comparação dessas duas categorias de métodos com os métodos baseados na predição linear é realizada através da reparametrização do problema original na busca das raízes de um polinômio associado a um filtro de erro de predição.

8. CONCLUSÕES

A necessidade de alta resolução nas aplicações atuais de estimação do parâmetro DOA exige métodos com capacidade de estimar parâmetros próximos entre si mesmo em baixas relações sinal-ruído. A correlação eventualmente existente entre as fontes de sinais aliada à pequena quantidade dos dados disponíveis, torna tal problema de estimação complexo e com ampla aplicação. O método da máxima-verossimilhança (ML), apesar de perfeitamente aplicável ao problema, apresenta esforço computacional proibitivo. O grande desafio para qualquer método alternativo é procurar ser computacionalmente viável sem contudo comprometer seu desempenho quando comparado com o do método ML.

Como resultado de uma ampla pesquisa, classificamos e propomos a descrição dos métodos alternativos ao método ML em três categorias. A primeira consiste dos métodos baseados no próprio critério ML, a segunda compreende os métodos baseados na predição linear e a terceira representa os métodos baseados em subespaços.

Apesar desses métodos possuírem origens e hipóteses distintas para as suas concepções, os mesmos foram aqui interpretados e descritos como a busca das estimativas que minimizam critérios semelhantes.

Com relação à primeira categoria de métodos, a diferença entre a implementação direta do critério ML e os métodos alternativos IQML e MODE foi apresentada como um procedimento de reparametrização visando a redução do esforço computacional. Particularmente, o método MODE foi interpretado como um procedimento de estimação direto derivado do algoritmo iterativo IQML, no qual as iterações foram eliminadas através do uso de estimativas auxiliares mais próximas dos seus valores ótimos. Por esse motivo o método MODE não sofre problemas de convergência como o método IQML, além de apresentar menor esforço computacional, embora ambos exijam a minimização de uma função custo de quarta ordem. Além dessas características, o método MODE faz uso de uma versão da matriz de correlação composta apenas pela parte associada ao subespaço de sinal.

Os métodos baseados em predição linear foram apresentados através de sua formulação tradicional utilizando o mesmo polinômio $b(z)$ da reparametrização adotada para os métodos IQML e MODE. Foi também apresentada uma formulação alternativa baseada na distinção entre os subespaços de sinal e de ruído da matriz de dados associada ao problema. Dessa forma foi possível relacioná-los não apenas aos métodos ba-

seados no critério ML, como também aos métodos modernos baseados em subespaços.

Além disso, ficou evidente que a utilização de informações adicionais em sua formulação, como a distinção entre os subespaços utilizada pelo método FBLP Modificado ou a característica de filtro branqueador utilizada no método WTLS, permite a obtenção de melhores resultados sem um aumento considerável do esforço computacional.

A comparação com os métodos baseados no critério ML permitiu constatar que o processo de reparametrização, adotado pelos métodos IQML e MODE, pode ser visto como a utilização de um filtro preditor de ordem igual ao número de parâmetros desconhecidos. Essa relação revelou ainda a menor complexidade relativa dos métodos baseados na predição linear, bem como sua menor capacidade de resolução quando os sinais são correlacionados e possuem baixa relação sinal-ruído. Uma vantagem da predição linear em relação a esses métodos é a possibilidade do aumento da ordem do preditor para melhorar a capacidade de resolução.

Quando à terceira categoria, os métodos MUSIC e Norma-mínima, concebidos anteriormente aos métodos IQML e MODE, puderam ser descritos como alternativas do critério ML baseadas em subespaços. Os métodos Norma-mínima e Pisarenko foram interpretados como casos particulares do método MUSIC. Através dessa descrição geral, foi possível compreender que, apesar da redução do esforço computacional oferecida por esses métodos, através de uma função custo de segunda ordem, os mesmos não foram concebidos para sinais correlacionados, principalmente em baixa relação sinal-ruído.

Quando comparados com os métodos baseados na predição linear, foi possível uma nova interpretação para o método MUSIC que deu origem ao método root-MUSIC.

Foi ainda possível estabelecer uma relação entre os métodos MUSIC, IQML e MODE, na qual o método MUSIC foi identificado como o resultado do primeiro passo do processo de minimização comum aos métodos IQML e MODE, constatando-se assim a superioridade de desempenho desses últimos e o menor esforço computacional do método MUSIC.

Fica portanto evidente que a utilização das informações associadas aos subespaços oferece um aumento significativo no desempenho dos métodos convencionais baseados no critério ML ou na predição linear. O método MUSIC, por exemplo, fornece estimativas com uma precisão equivalente ao da implementação do critério ML determinístico, desde que as fontes de sinais sejam descorrelacionadas. No entanto, tais métodos geralmente exibem uma grande polarização em suas estimativas para amostras finitas, levando a problemas de resolução. Tal dificuldade de resolução é, então, agravada para fontes de sinais correlacionadas [4], [17], [39], [53].

A apresentação dos métodos baseados em subespaços que utilizam um critério de "distância" entre o subespaço de sinal e sua aproximação, como no caso do SSF e sua versão baseada no subespaço de ruído, NSF, permitiu constatar a equivalência desse último com o método MUSIC. Foi ainda possível classificar o método ESPRIT nessa categoria devido à semelhança em suas formulações. Uma generalização para esse critério foi obtida adicionando-se uma matriz de ponderações em sua formulação. O método denominado de

WSF faz uso dessa ponderação e apresenta a mesma função custo do método MODE.

Concluímos que os diversos métodos de estimação de parâmetros encontrados na literatura possuem fortes relações com o conceito básico de predição linear ou com o critério ML, seja através da ordem do filtro utilizado, da forma de utilização dos dados disponíveis ou do uso da distinção entre os subespaços que compõem o sinal observado. Tais variações para um mesmo conceito influenciam diretamente na solução do problema original e, conseqüentemente, das características de resolução, complexidade numérica e de desempenho do estimador obtido. Particularmente, a forma de manipular os subespaços se mostrou de grande importância na obtenção de melhores estimativas.

Este trabalho evidenciou, portanto, a vasta possibilidade de relações entre os métodos existentes. A classificação dos métodos em três categorias e a utilização de uma formulação unificada para descrevê-los se apresentam como suas principais contribuições. Este trabalho apresentou ainda a proposta de extensão do procedimento de predição linear sobre vários experimentos através da estimação de uma matriz de correlações adequada. Uma outra proposta sugere a utilização de uma matriz de ponderações sobre o método NSF, dando origem ao método NSF Ponderado ou WNSF, a exemplo do método WSF. Novas propostas ainda estão sob investigação e serão objeto de trabalhos futuros.

AGRADECIMENTOS

Agradecemos o trabalho paciente e criterioso dos revisores que muito contribuiu para a qualidade e precisão deste artigo.

REFERÊNCIAS

- [1] P. S. Unnikrishna. *Array Signal Processing*, Springer-Verlag, New York, 1989.
- [2] S. Haykin. *Array Signal Processing*, Prentice-Hall, Englewood Cliffs, New Jersey, 1985.
- [3] D. H. Johnson, "The Application of Spectral Estimation Methods to Bearing Estimation Problems" *Proc. of the IEEE*, Vol. 70, No. 9, Sep. 1982.
- [4] H. Krim e M. Viberg, "Two Decades of Array Signal Processing Research", *IEEE Signal Processing Magazine*, Vol. 13, No. 4, Jul., 1996.
- [5] L. C. Godara, "Applications of Antenna Arrays to Mobile Communications, Part I and Part II," *Proc. of the IEEE*, Vol. 85, No. 7, July 1997.
- [6] A. J. Paulraj e C. B. Papadias, "Space-Time Processing for Wireless Communications", *IEEE Signal Processing Magazine*, Nov., 1997.
- [7] J. Sheinvald, M. Wax e A. J. Weiss, "On Maximum-Likelihood Localization of Coherent Signals", *IEEE Trans. on Signal Processing*, Vol. 44, No. 10, Out. 1996.
- [8] S. M. Kay, *Modern Spectral Estimation: Theory and Application*, Prentice-Hall, Englewood Cliffs, New Jersey, 1988.
- [9] P. Stoica and K. C. Sharman, "Maximum Likelihood Methods for Direction-of-Arrival Estimation" *IEEE Trans. on Acoustic, Speech and Signal Processing*, Vol. 38, No. 7, Jul. 1990.
- [10] D. W. Tufts and R. Kumaresan, "Estimation of Frequencies of Multiple Sinusoids: Making Linear Prediction Perform Like Maximum Likelihood" *Proc. of the IEEE*, Vol. 70, No. 9, Set. 1982.

- [11] R. F. Colares e A. Lopes, "Aprimorando Método para Detecção de Frequências Baseado na Norma-nínima", Tese de Mestrado, UNICAMP, Campinas-SP, Set. 1997.
- [12] R. P. Lemos, e A. Lopes, "Mínimos-quadrados Totais e Máxima-verossimilhança em Estimação de Frequências", Tese de Doutorado, UNICAMP, Campinas-SP, Jun. 1997.
- [13] E. J. B. Antunes, "Detecção de Frequências Através de Predição Linear", Tese de Mestrado, UNICAMP, Campinas-SP, Maio 1992.
- [14] M. Jansson, B. Göransson e B. Ottersten, "A Subspace Method for Direction of Arrival Estimation of Uncorrelated Emitter Signals", *IEEE Trans. on Signal Processing*, Vol. 47, No. 4, Abril 1999.
- [15] V. F. Pisarenko, "The Retrieval of Harmonics from a Covariance Function", *Roy. J. Astron. Soc.*, Vol. 33, 1973.
- [16] M. M. Reddi, "Multiple Source Location - A Digital Approach", *IEEE Trans. on Aerospace, Electronic and System*, Vol. 15, No. 1, Jan. 1979.
- [17] J. A. Cadzow, "Multiple Source Location-The Signal Subspace Approach", *IEEE Trans. on Acoustic, Speech and Signal Processing*, Vol. 38, No. 7, Jul. 1990.
- [18] B. Ottersten, M. Viberg e T. Kailath, "Analysis of Subspace Fitting and ML Techniques for Parameter Estimation from Sensor Array Data", *IEEE Trans. on Signal Processing*, Vol. 40, No. 3, Mar. 1992.
- [19] M. Viberg e B. Ottersten, "Sensor Array Processing Based on Subspace Fitting" *IEEE Trans. on Signal Processing*, Vol. 39, No. 5, Maio 1991.
- [20] M. Viberg, B. Ottersten e T. Kailath, "Detection and Estimation in Sensor Array Using Weighted Subspace Fitting" *IEEE Trans. on Signal Processing*, Vol.39, No. 11, Nov. 1991.
- [21] Y. Bresler and A. Macovski, "Exact Maximum Likelihood Parameter Estimation of Superimposed Exponential Signals in Noise" *IEEE Trans. on Acoustic, Speech and Signal Processing*, Vol. 34, No. 10, Out. 1986.
- [22] P. Stoica and K. Sharman, "Novel Eigenanalysis Method for Direction Estimation", *IEE Proc.*, Vol. 137, Part. F, No. 1, Fev. 1990.
- [23] P. Stoica, J. Li e T. Söderström, "On the Inconsistency of IQML", *Signal Processing, Elsevier Science B. V.* No. 56, 1997.
- [24] R. F. Colares e A. Lopes, "Visão Unificada dos Métodos Modernos para DOA Através de Arranjos de Sensores Lineares e Uniformes", Anais do XVII Simpósio Brasileiro de Telecomunicações, ES, Set. 1999.
- [25] R. F. Colares, C. A. Alves e A. Lopes, "Proposta para Aplicação do Método WTLS em DOA", Anais do XVIII Simpósio Brasileiro de Telecomunicações, RS, Set. 2000.
- [26] R. F. Colares e A. Lopes, "Improving DOA Estimation Methods Using a Priori Knowledge About The Sources Location", 42nd Midwest Symposium on Circuits and Systems, New Mexico State University, Novo México, USA, Ago. 1999.
- [27] M. Zatman, "How Narrow is Narrowband", *IEE Proc. on Radar, Sonar and Navig.*, Vol 145, No. 2, Abril 1998.
- [28] P. Stoica and A. Nehorai, "Performance Study of Conditional and unconditional Direction-of-Arrival Estimation", *IEEE Trans. on Acoustic, Speech and Signal Processing*, Vol.38, No. 10, Out. 1990.
- [29] J. Sheinvald, M. Wax e A. J. Weiss, " On Maximum-Likelihood Localization of Coherent Signals", *IEEE Trans. on Signal Processing*, Vol. 44, No. 10, Out. 1996.
- [30] S. M. Kay, *Fundamentals of Statistical Processing, Volume I: Estimation Theory*, Prentice-Hall, Englewood Cliffs, New Jersey, 1993.
- [31] M. Wax, "Detection and Localization of Multiple Sources Via the Stochastic Signals Model", *IEEE Trans. on Signal Processing*, Vol. 39, No. 11, Nov. 1991.
- [32] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3a. ed., New York, NY, MacGraw-Hill, 1991.
- [33] G. H. Golub and C. F. VanLoan, *Matrix Computation*, Johns Hopkins University Press, Baltimore, MD., 2a. ed., 1989.
- [34] H. W. Sorenson, "Least-squares estimation: from Gauss to Kalman", *IEEE Spectrum*, Julho 1970.
- [35] M. Wax e T. Kailath, "Determining the number of sources by information theoretic criteria", *IEEE Trans. on Acoustic, Speech and Signal Processing*, Vol. 33, No. 2, Fev. 1985.
- [36] H. Wang e M. Kaveh, "On the performance of signal subspace processing - Part I: Narrow-band systems ", *IEEE Trans. on Acoustic, Speech and Signal Processing*, Vol. 34, No. 10, Out. 1986.
- [37] C. A. Alves, R. F. Colares e A. Lopes, "Análise de Métodos para DOA em Função de Restrições de Otimização", Anais do XVIII Simpósio Brasileiro de Telecomunicações, RS, Set. 2000.
- [38] C. A. Alves, A. Lopes e R. F. Colares, "Constraints Implementation for IQML and MODE Direction-of-Arrival Estimators", 43rd Midwest Symposium on Circuits and Systems - MWS-CAS 2000, Michigan State University, Michigan, USA, Ago. 2000.
- [39] P. Stoica e A. Nehorai, "MUSIC, Maximum Likelihood and Cramér-Rao Bound", *IEEE Trans. on Acoustic, Speech and Signal Processing*, Vol. 37, No. 5, Maio 1989.
- [40] J. Li, P. Stoica e Z. Liu, "Comparative Study of IQML and MODE Direction-of-Arrival Estimators", *IEEE Trans. on Acoustic, Speech and Signal Processing*, Vol. 46, No. 1, Jan. 1998.
- [41] R. Kumaresan, "On the Zeros of the Linear Prediction-error Filter for Deterministic Signals", *IEEE Trans. on Acoustic, Speech and Signal Processing*, Vol. 71, No. 2, Fev. 1983.
- [42] T. J. Shan, M. Wax, e T. Kailath, "On Spatial Smoothing for Direction of Arrival Estimation of Coherent Signal", *IEEE Trans. on Acoustic, Speech and Signal Processing*, Vol. 33, No. 4, Abril 1985.
- [43] B. De Moor e J. Vandewalle, "A Unifying Theorem for Linear and Total Linear Least Squares", *IEEE Trans. on Automatic Control*, Vol. 35, No. 5, Maio 1990.
- [44] G. H. Golub e C. F. Van Loan, "An Analysis of the Total Least Squares Problem", *SIAM J. Numer. Anal.*, Vol. 17, No. 6, Dez. 1980.
- [45] S. Van Huffel e J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*, SIAM Frontiers in Applied Mathematics, Vol. 9, Philadelphia, 1991.
- [46] Y. Hua e T. K. Sarkar, "On the Total Least Squares Linear Prediction Method for Frequency Estimation", *IEEE Trans. on Acoustic, Speech and Signal Processing*, Vol. 38, No. 12, Dez 1990.
- [47] M. A. Rahman e K.-B. Yu, "Total Least Squares Approach for Frequency Estimation using Linear Prediction", *IEEE Trans. on Acoustic, Speech and Signal Processing*, Vol. 35, No. 10, Out. 1987.
- [48] E. M. Dowling e R. D. DeGroat, "The Equivalence of the Total Least Squares and Minimum Norm Methods", *IEEE Trans. on Signal Processing*, Vol. 39, No. 8, Ago. 1996.
- [49] J. I. Makhoul, "Linear Prediction: A Tutorial Review", *Proc. of IEEE*, Vol. 63, No. 4, Abril 1975.
- [50] R. O. Schmidt, " Multiple Emitter Location and Signal Parameter Estimation", *IEEE Trans. on Ant. and Prop.*, vol. 34, 1986.
- [51] R. Roy e T. Kailath, "ESPRIT - Estimation of Signal Parameters via Rotational Invariance Techniques", *IEEE Trans. on*

- Acoustic, Speech and Signal Processing*, Vol. 37, No. 7, Jul. 1989.
- [52] P. Stoica e T. Söderström. "Statistical Analysis of MUSIC and Subspace Rotational Estimates of Sinusoidal Frequencies", *IEEE Transc. on Acoustic, Speech and Signal Processing*, Vol. 39, No. 8, Ago. 1991.
- [53] P. Stoica e A. Nehorai, "MUSIC. Maximum Likelihood and Cramér-Rao Bound: Further Results and Comparisons", *IEEE Transc. on Acoustic, Speech and Signal Processing*, Vol. 38, No. 12, Dec. 1990.
- [54] H. Krim, P. Forster e J. G. Proakis, "Operator Approach to Performance Analysis of Root-MUSIC and Root Min-Norm", *IEEE Transc. on Acoustic, Speech and Signal Processing*, Vol. 40, No. 7, Jul. 1992.
- [55] M. Kaveh e A. J. Barabell. "The Statistical Performance of the MUSIC and the Minimum-Norm Algorithms in Resolving Plane Waves in Noise", *IEEE Transc. on Acoustic, Speech and Signal Processing*, Vol. 34, No. 4, Abril 1986.
- [56] R. Kumaresan e D. Tufts, "Estimating the Angles of Arrival of Multiple Plane Waves", *IEEE Trans on Aerospace and Electronic Systems*, Vol. 19, No. 1, Jan. 1983.
- [57] V. T. Ermolaev e A. G. Flaksman. "Fast Algorithm for Minimum-Norm Direction-of-Arrival Estimation", *IEEE Trans. on Signal Processing*, Vol. 42, No. 9, Set. 1994.
- [58] A. Paulraj, R. Roy e T. Kailath. "Subspace Rotational Approach to Signal Parameter Estimation", *Proceedings of the IEEE*, No. 7, Jul. 1986.
- Ricardo Fialho Colares** nasceu em Fortaleza, CE, Brasil, em 1969. Graduiu-se em Engenharia Elétrica pela Universidade Federal do Ceará, CE, em 1992. Recebeu os títulos de Mestre e Doutor em Engenharia Elétrica da Universidade Estadual de Campinas em 1997 e 2000, respectivamente. Participou de pesquisas no âmbito do convênio CPqD/UNICAMP em 1998 e 1999. Em 2000 foi professor assistente na Universidade São Francisco, SP, e desde o início de 2001 está na Universidade de Fortaleza (UNIFOR), CE, como Professor Titular e como Coordenador do Curso de Graduação em Engenharia de Telecomunicações. Suas áreas de interesse são processamento digital de sinais, matemática aplicada às telecomunicações e arranjo de antenas.
- Carlos Antonio Alves** graduou-se e obteve o título de Mestre em Engenharia Elétrica pela Universidade Estadual Paulista/UNESP, Ilha Solteira-SP, em 1989 e 1995, respectivamente. Desde 1990 ele está na Faculdade de Engenharia de Ilha Solteira - FEIS/UNESP com Professor Assistente. Atualmente ele desenvolve o programa de doutorado na Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas. Sua área de pesquisa é processamento digital de sinais.
- Amauri Lopes** graduou-se e obteve os títulos de Mestre e Doutor em Engenharia Elétrica pela Universidade Estadual de Campinas em 1972, 1975 e 1982, respectivamente. Ele está na Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas desde 1973, onde exerce atualmente o cargo de professor associado. Suas áreas de interesse são processamento digital de sinais, teoria de circuitos e transmissão de sinais digitais.