

UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE ENGENHARIA ELÉTRICA

DEPARTAMENTO DE COMUNICAÇÕES

DETECÇÃO DE FREQUÊNCIAS ATRAVÉS DE PREDIÇÃO LINEAR

Por: **Ernesto José Batista Antunes**
Engenheiro Eletricista (UFPE - 1988)

Orientador : Prof. Dr. Amauri Lopes †
Professor MS4 do Departamento de Comunicações da
Faculdade de Engenharia Elétrica da UNICAMP

Dissertação submetida à Faculdade de Engenharia Elétrica da UNICAMP como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica.

Campinas, maio de 1992



Este exemplar corresponde ao original final da tese defendida por **Ernesto José Batista Antunes**

Julgadora em 05/06/92

Amauri Lopes
Orientador

Comissão

SUMÁRIO

Este trabalho descreve os aspectos teóricos e práticos da utilização de predição linear para estimar as frequências de sinais senoidais. A predição linear baseada no critério dos mínimos quadrados é empregada nos contextos forward, backward e forward-backward, e é apresentado todo o formalismo matemático necessário a tais implementações. Simulações em computador evidenciam as principais características das técnicas de estimação, bem como ilustram aspectos ligados aos seus desempenhos em função dos vários parâmetros envolvidos no processo. O método baseado na predição forward-backward e na decomposição da matriz de correlação em valores singulares, denominado FBLP Modificado, é desenvolvido e analisado em detalhes, enriquecendo os resultados da literatura. Por fim, é proposto um novo algoritmo que propicia uma melhora significativa no desempenho deste método quando aplicado na detecção de senóides reais.

AGRADECIMENTOS

Gostaria de agradecer a todos os que de alguma forma tenham me ajudado na realização deste trabalho, seja através de discussões, críticas, sugestões ou colocações pessoais.

Agradeço ao meu orientador Prof. Amauri Lopes não só pela sua imensa competência profissional, mas também pelo incentivo e amizade dispensados, sem os quais esta tarefa tornar-se-ia bastante árdua.

A Celi, Márcia, Gisleine, Ieda e Ademilde por suas valiosas colaborações profissionais e paciência às nossas solicitações.

Aos amigos Paulo, Tereza, Adrian, Francisco e Leonardo pelo companheirismo e intercâmbio de conhecimentos técnicos durante os cursos e a tese.

Ao CNPq e CPqD/TELEBRAS pelo apoio financeiro e a UNICAMP pela infraestrutura do seu curso de pós-graduação, sem os quais o desenvolvimento deste trabalho seria impossível.

A Margarida, pelo companheirismo, incentivo e paciência no dia a dia.

aos meus pais

para Guida

ÍNDICE

CAPÍTULO 1 : INTRODUÇÃO	1
1.1 Detectar Senóides - Um Problema Clássico	1
1.2 Os Subespaços de Sinal e de Ruído	2
1.3 Predição Linear para Estimar Frequências	2
1.4 Objetivos	3
1.5 Resumo dos Capítulos Seguintes	4
CAPÍTULO 2 : PREDIÇÃO LINEAR DETERMINÍSTICA	5
2.1 Introdução	6
2.2 Predição Linear Segundo o Critério dos Mínimos Quadrados	6
2.3 O Método de Predição Linear Forward - FLP	8
2.4 O Método de Predição Linear Backward - BLP	14
2.5 O Método de Predição Linear Forward-Backward - FBLP	17
2.6 O Teorema da Decomposição via Valor Singular - SVD	21
2.7 Definição Geral de Pseudo-Inversa de uma Matriz	26
2.8 Resolução da Equação Normal Determinística	29
2.9 A Solução de Norma Mínima	37
2.10 Outra Formulação para a Solução de Norma Mínima	43
2.11 Conclusões	46
CAPÍTULO 3 :DETECÇÃO DE SINAIS SENOIDAIS EM RUÍDO BRANCO	47
3.1 Introdução	47
3.2 Os Zeros do Filtro de Erro de Predição - Caso sem Ruído	48
3.3 Estimação das Frequências de Exponenciais Não-Amortecidas - Caso sem Ruído	64
3.4 Estimação das Frequências de Exponenciais Não-Amortecidas - Caso com Ruído	67
3.5 Conclusões	79

CAPÍTULO 4 : O MÉTODO FBLP MODIFICADO	81
4.1 Introdução	81
4.2 Melhorando o Método FBLP	82
4.3 O Método FBLP Modificado	84
4.4 Comparações Quantitativas	89
4.5 O Caso Kumaresan-Prony - KP	92
4.6 Variação de Alguns Parâmetros	93
4.7 Estimativa do Número de Exponenciais	95
4.8 Uma Aplicação Típica: Detecção de Ângulo de Incidência Vertical em Radar	97
4.9 Conclusões	103
CAPÍTULO 5 : REDUZINDO A SNR DE LIMIAR	105
5.1 Introdução	105
5.2 Uma Tentativa para Diminuir a SNR de Limiar	106
5.3 Detecção de Senóides Reais	111
5.4 Conclusões	122
CAPÍTULO 6 : CONCLUSÃO	123
6.1 Resumo	123
6.2 Conclusões e Contribuições	124
6.3 Sugestões para Proseguimento deste Trabalho	125
APÊNDICE A : OS AUTOVALORES DE UMA MATRIZ HERMITIANA	126
APÊNDICE B : OS AUTOVETORES DE UMA MATRIZ HERMITIANA	129
REFERÊNCIAS	

CAPÍTULO 1

INTRODUÇÃO

1.1 DETECTAR SENÓIDES - UM PROBLEMA CLÁSSICO

A detecção de senóides contaminadas por ruído branco tem despertado interesse especial em processamento digital de sinais. O problema básico consiste em estimar as frequências e, eventualmente, as potências das senóides, utilizando o menor número possível de amostras do sinal. Estas amostras variam linearmente com as amplitudes das senóides, e de forma não-linear com suas frequências. O interesse principal recai sobre a estimativa das frequências. Uma vez que as mesmas tenham sido determinadas de forma suficientemente precisa, as amplitudes podem ser obtidas por um ajuste linear apropriado dos mínimos quadrados para o conjunto de amostras disponíveis [1]. A detecção de senóides contaminadas por ruído branco tem diversas aplicações, tais como: sonar, radar, economia, geofísica, sinalização telefônica, modelagem do sinal de voz, astronomia, acústica subaquática, estimação de respostas impulsivas racionais, sismologia, análise de sinais biomédicos [14]–[20].

A determinação das frequências, não exige estimação da densidade espectral de potência do sinal, mas apenas a construção de funções que apresentem picos elevados nas localizações das frequências das senóides. Tais funções não possibilitam a recuperação da sequência de autocorrelação do sinal através da transformada de Fourier inversa, nem tão pouco preservam a potência do sinal [2].

Caso a separação entre as frequências das senóides seja maior que o recíproco do intervalo de observação do sinal e a relação sinal-ruído (SNR^1) seja suficientemente elevada, métodos baseados na transformada de Fourier, tal como o periodograma, ou métodos paramétricos de análise espectral [2], estimam as

¹do inglês Signal Noise Ratio

frequências de forma precisa [17], [21]. Porém, para um pequeno número de amostras disponíveis, frequências muito próximas e SNR pequena, tais métodos apresentam deficiências intoleráveis.

1.2 OS SUBESPAÇOS DE SINAL E DE RUÍDO

No caso de senóides contaminadas por ruído branco, a matriz de correlação verdadeira, quando decomposta em termo de seus autovalores e autovetores, pode ser dividida em dois subespaços vetoriais: *o subespaço de sinal e o subespaço de ruído*. O subespaço de sinal contém todas as componentes relativas à porção de sinal das amostras, a qual é constituída unicamente pela soma das senóides, e algumas componentes relativas ao ruído branco. O subespaço de ruído contém as componentes restantes do ruído branco [2]. Construir uma nova matriz de correlação por desprezar os autovalores e autovetores do subespaço de ruído, produz efeitos semelhantes àqueles proporcionados por um aumento da SNR. Podemos obter também vetores de sinal, denominados s_i , contendo a informação de frequência da i -ésima senóide. Demonstra-se que tais vetores são ortogonais aos autovetores do subespaço de ruído [2].

Porém, a sequência de autocorrelação geralmente não é conhecida, o que nos obriga a gerar estimativas da mesma. Neste caso, as propriedades dos subespaços de sinal e de ruído são apenas aproximadamente verdadeiras. Ainda assim, o uso adequado destes subespaços aproximados possibilita a elaboração de estimadores de frequência bastante eficientes. Tais estimadores podem ser divididos em duas classes: *Estimadores de Frequência do Subespaço de Sinal e Estimadores de Frequência do Subespaço de Ruído*. Os estimadores do subespaço de sinal caracterizam-se por desprezar a parte da matriz de autocorrelação estimada relativa ao subespaço de ruído, obtendo uma nova matriz de posto reduzido. A partir disto, a utilização adequada de técnicas de análise espectral fornecerá estimativas mais precisas das frequências das senóides. Os estimadores do subespaço de ruído baseiam-se na ortogonalidade dos vetores de sinal s_i e os autovetores do subespaço de ruído. Constituem exemplos de tal classe de estimadores, o método MUSIC e a Decomposição Harmônica de Pisarenko [2].

1.3 PREDIÇÃO LINEAR PARA ESTIMAR FREQUÊNCIAS

A teoria de filtragem de Wiener baseia-se no conhecimento das características estatísticas de primeira e segunda ordem do sinal. Já que dispomos

de apenas um conjunto finito de amostras, somente médias temporais tornam-se possíveis, originando a predição linear segundo o critério dos mínimos quadrados.

A predição linear segundo tal critério pode ser utilizada para estimar frequências de senóides em meio a ruído branco, através dos zeros do polinômio formado pelos coeficientes do filtro de erro de predição para as amostras disponíveis [11]–[13]. Tufts e Kumaresan [7], originalmente, propuseram um novo estimador de frequência baseado no subespaço de sinal. A técnica consiste em utilizar o método de predição linear forward-backward(FBLP) [12]–[13] para estimar as frequências, desprezando os autovalores-autovetores do subespaço de ruído da matriz de correlação estimada. Tal método, chamado de *Método FBLP Modificado*, mostra-se bastante eficiente para estimar senóides com frequências bem próximas, com SNR pequena e dispondo de um pequeno número de amostras do sinal. Seu desempenho é próximo ao do método da máxima-verossimilhança (ML) e sua complexidade computacional é menor. Tudo isto credencia o método FBLP Modificado como uma técnica de alta resolução competitiva com o método ML.

1.4 OBJETIVOS

Este trabalho propõe-se a realizar um estudo detalhado da predição linear aplicada à estimativa de frequências e da utilização do subespaço de sinal para melhorar o desempenho de tal estimador. Destaque especial foi dado ao método FBLP Modificado, o qual representou um marco importantíssimo na tentativa de solucionar este problema. A qualidade de seu desempenho, aliada à relativa simplicidade computacional, fazem com que a avaliação de novos métodos se dê através da comparação com o método FBLP Modificado, seja em desempenho ou em complexidade computacional.

Tudo isto nos motivou a realizar este trabalho, onde o método FBLP Modificado foi desenvolvido, simulado e comparado a outros métodos. A dedução detalhada do mesmo foi motivada pelo forte desejo em conhecer perfeitamente as razões que o tornam tão eficiente e também criar uma base para novos avanços.

O método tem sua origem em predição linear segundo o critério dos mínimos quadrados. Por isto desenvolvemos esta teoria, abordando de forma mais detalhada e exemplificada o assunto já explorado em literatura especializada. Apresentamos também de forma detalhada, a teoria envolvida em métodos de predição linear para estimadores de frequência de alta resolução, que utilizam as características dos subespaços de sinal e de ruído da matriz de correlação estimada.

Baseado nestas teorias, desenvolvemos um conjunto de programas na

linguagem C, necessários para a implementação dos métodos de detecção estudados. Exceto para algumas poucas exceções, claramente anunciadas no texto, todas as simulações apresentadas neste trabalho são frutos da aplicação destes programas.

O conhecimento propiciado por tais realizações levou-nos ao desenvolvimento de um novo algoritmo para a detecção de senóides reais, que permite melhorar o desempenho do método FBLP Modificado nesta aplicação.

1.5 RESUMO DOS CAPÍTULOS SEGUINTEs

No *capítulo 2*, desenvolveremos toda a teoria de predição linear segundo o critério dos mínimos quadrados. As predições serão realizadas nos contextos forward (FLP), backward (BLP) e forward-backward (FBLP).

No *capítulo 3* demonstraremos como a predição linear pode ser utilizada para detectar frequências de senóides em meio a ruído branco. Com isto, toda a teoria desenvolvida no capítulo 2 poderá ser aplicada para este fim, objetivo principal do nosso trabalho, originando o método FBLP para estimar frequências. Simulações ilustrativas comporão o capítulo.

O *capítulo 4* apresentará o método FBLP Modificado, o qual resulta da aplicação do conceito de subespaço de sinal e de ruído ao método FBLP. Este método se mostrará mais eficiente do que o anterior, tanto na resolução quanto na variância das estimativas. Os aspectos que influenciam estes parâmetros serão apresentados, evidenciando assim os motivos pelos quais a desconsideração do subespaço de ruído melhora de forma acentuada o desempenho. Diversas simulações permitirão avaliar o desempenho e a dependência deste com alguns parâmetros. Por fim apresentaremos uma aplicação típica do método FBLP Modificado.

O *capítulo 5* apresentará alguns comentários e resultados atingidos com algumas especulações objetivando a redução da SNR de limiar do Método FBLP Modificado, bem como enfocaremos o caso de detecção de senóides reais. Neste contexto, desenvolveremos, originalmente, um algoritmo que apresenta melhor desempenho que o obtido com a aplicação exclusiva do Método FBLP Modificado.

No *capítulo 6* apresentaremos as conclusões, destacando as principais contribuições deste trabalho.

CAPÍTULO 2

PREDIÇÃO LINEAR DETERMINÍSTICA

CONTEÚDO	Pag.
2.1 Introdução.....	6
2.2 Predição Linear	
Segundo o Critério dos Mínimos Quadrados.....	6
2.3 O Método de Predição Linear Forward - FLP.....	8
2.3.1 O Filtro de Erro de Predição Forward.....	8
2.3.2 A Equação Normal Determinística.....	9
2.3.3 A Equação Normal Determinística e as Funções de Correlação.....	11
2.3.4 Propriedades da Matriz de Correlação Determinística.....	13
2.4 O Método de Predição Linear Backward - BLP.....	14
2.4.1 O Filtro de Erro de Predição Backward.....	14
2.4.2 A Equação Normal Determinística.....	15
2.5 O Método de Predição Linear Forward-Backward - FBLP....	17
2.6 O Teorema da Decomposição via Valor Singular - SVD.....	21
2.7 Definição Geral de Pseudo-Inversa de uma Matriz.....	26
2.8 Resolução da Equação Normal Determinística.....	29
2.9 A Solução de Norma Mínima.....	37
2.10 Outra Formulação para a Solução de Norma Mínima.....	43
2.11 Conclusões.....	46

2.1 INTRODUÇÃO

Neste capítulo desenvolveremos a teoria de predição linear de um sinal para a situação onde se dispõe apenas de um conjunto finito de amostras do mesmo. Em função disto, as médias estatísticas da teoria de filtragem de Wiener serão substituídas por médias temporais. Utilizaremos o método de otimização de um preditor linear baseado na minimização dos erros quadráticos, configurando a *Predição Linear Segundo o Critério dos Mínimos Quadrados*, o qual pode ser interpretada como a versão determinística da predição linear segundo os critérios de Wiener.

O critério dos mínimos quadrados será desenvolvido no contexto das predições forward (progressiva), backward (regressiva) e forward-backward (progressiva-regressiva), e o projeto do preditor nos conduzirá à resolução de uma equação denominada *Equação Normal Determinística*. O conjunto de possíveis soluções para esta equação será analisado, onde constataremos que a mesma sempre possui solução.

Dentre as possíveis soluções da equação normal determinística, atenção especial será dada à *Solução de Norma Mínima*. A fim de obtermos esta solução, o *Teorema da Decomposição via Valor Singular* será enunciado e demonstrado, bem como definiremos o conceito de pseudo-inversa de uma matriz.

2.2 PREDIÇÃO LINEAR SEGUNDO O CRITÉRIO DOS MÍNIMOS QUADRADOS

Seja um conjunto de amostras de um dado sinal com valores complexos $\{u(1), u(2), \dots, u(N)\}$, observados nos instantes t_1, t_2, \dots, t_N , uniformemente espaçados. Seja também um filtro transversal, com $(L+1)$ fatores de ganho, na forma:

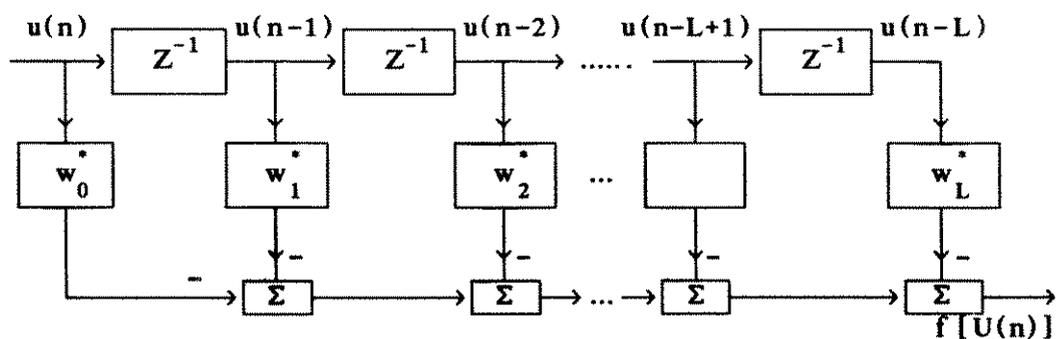


Fig. 2.1 - Filtro de Erro de Predição Linear Genérico

Onde: $U(n) = \{u(n), u(n-1), \dots, u(n-L)\}$

Suponhamos agora que o i 'ésimo fator de ganho w_i^* seja igual a -1 , e que objetivamos minimizar : $\epsilon_L(w_0, w_1, \dots, w_L) = \sum_n |f[U(n)]|^2$. O ideal, segundo este objetivo, seria que todos os $f[U(n)]$ fossem iguais a zero, obtendo-se assim o menor ϵ_L possível, ou seja, o ideal seria que:

$$- w_0^* \cdot u(n) - w_1^* \cdot u(n-1) - \dots + u(n-i) - \dots - w_L^* \cdot u(n-L) = f[U(n)] = 0$$

portanto:

$$u(n-i) = w_0^* \cdot u(n) + w_1^* \cdot u(n-1) + \dots + w_{(n-i+1)}^* \cdot u(n-i+1) \\ + w_{(n-i-1)}^* \cdot u(n-i-1) + \dots + w_L^* \cdot u(n-L)$$

Este resultado pode ser interpretado como a tentativa de prever o valor da amostra $u(n-i)$ a partir de uma combinação linear das demais amostras presentes no filtro. Daí o nome Predição Linear, enquanto que o filtro da fig. 2.1 é conhecido como o *Filtro de Erro de Predição Linear*, pois $f[U(n)]$ representa este erro. Na prática, geralmente, não conseguimos ϵ_L igual a zero, mas o filtro é ajustado de modo a torná-lo o mínimo possível.

O problema de predição linear segundo o critério dos mínimos quadrados, consiste em obter um mínimo para $\epsilon_L(w_0, w_1, \dots, w_L) = \sum_n |f[U(n)]|^2$, sujeito à condição que $w_i^* = -1$, ou seja, obter $w_0, w_1, \dots, w_i = -1, \dots, w_L$ tal que ϵ_L seja o menor possível, onde $f[U(n)]$ é uma função das amostras observadas. Dado que o filtro é composto por L elementos atrasadores, dizemos que o filtro tem ordem L .

Foi omitido acima os limites para o somatório em ϵ_L . Escolhas distintas para estes limites originam diferentes expressões para ϵ_L e, por consequência, para os ganhos do filtro, dando origem aos seguintes métodos [1]:

a) Método da Autocorrelação - Assumimos que os dados anteriores ao instante de tempo t_1 e posteriores a t_N são zero. Portanto, neste caso, são considerados os erros produzidos quando o filtro não está completamente cheio de amostras, ou seja, os limites para o somatório em ϵ_L serão 1 e $(N+L)$.

b) Método do Pré-janelamento - Assumimos que os dados anteriores ao instante de tempo t_1 são zero, mas nenhuma suposição é feita acerca dos dados posteriores a t_N . Portanto, neste caso não se leva em conta os erros produzidos após a última amostra $u(N)$ ter adentrado ao filtro, ou seja, os limites do somatório em ϵ_L serão 1 e N .

c) Método do Pós-janelamento - Assumimos que os dados posteriores ao instante de tempo t_N são zero, mas nenhuma suposição é feita acerca dos dados anteriores ao instante t_1 . Portanto, neste caso, não se leva em conta os erros produzidos antes que a primeira amostra $u(1)$ alcance a derivação do ganho w_L^* , ou seja, $n = L+1, \dots, N+L$.

É sensato, porém, não fazer qualquer suposição acerca dos dados fora do intervalo t_1, \dots, t_N , já que só dispomos dos dados neste intervalo. Assim, serão levadas em conta apenas as situações em que o filtro de erro de predição linear estiver cheio, originando assim o quarto método, denominado *Método da Covariância*, ou seja:

$$\epsilon_L(w_0, w_1, \dots, w_{i=-1}, \dots, w_L) = \sum_{n=L+1}^N |f[U(n)]|^2 \Big|_{w_i^* = -1}$$

Obter w_k , $k = 0, \dots, L$, e $k \neq i$, segundo função de custo ϵ_L , é conhecido como *O Método de Predição Linear Segundo o Critério dos Mínimos Quadrados*. Este método pode ser visto como a contrapartida determinística da teoria de predição linear clássica, uma vez que os filtros de predição linear resultantes da teoria de filtragem de Wiener são derivados de médias estatísticas. O resultado é que apenas um filtro ótimo é especificado para qualquer conjunto de amostras de um sinal, o qual é suposto estacionário no sentido amplo. Por outro lado, o método dos mínimos quadrados é uma alternativa para as situações onde se dispõe apenas de um conjunto finito de amostras do sinal a ser filtrado. Neste método, o conjunto de ganhos do filtro dependem diretamente do conjunto de amostras disponíveis. Logo, cada conjunto gerará um filtro de erro de predição diferente.

Daqui por diante, o critério para o projeto de filtros de predição linear será sempre o dos mínimos quadrados. Portanto, a frase "*segundo o critério dos mínimos quadrados*" será, por conveniência, omitida na maioria das vezes, mesmo que, a rigor, o correto seja utilizá-la.

2.3 O MÉTODO DE PREDIÇÃO LINEAR FORWARD - FLP

2.3.1 O FILTRO DE ERRO DE PREDIÇÃO FORWARD

Supondo, na fig. 2.1, que $w_0^* = -1$ e denominando $f[U(n)]$ de $f_L(n)$, obtemos o esquema da fig. 2.2.

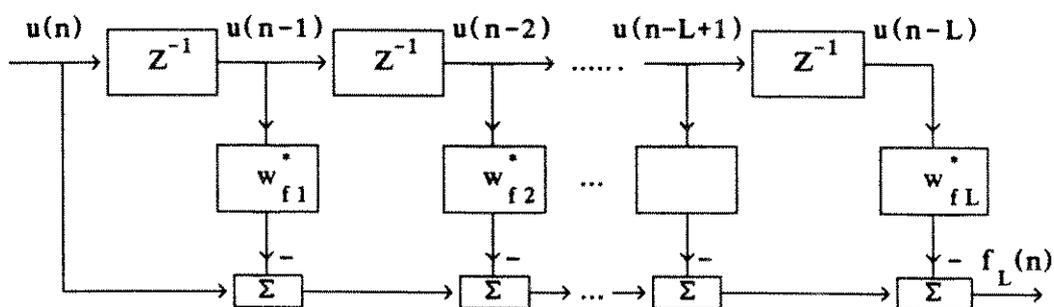


Fig. 2.2 - Filtro de Erro de Predição Linear Forward

As entradas $\{u(n-1), u(n-2), \dots, u(n-L)\}$ das derivações são utilizadas na predição linear de $u(n)$. Portanto, a predição de $u(n)$ se baseia em amostras anteriores a ela, o que justifica a denominação Predição Forward (sem atraso). Definimos $f_L(n)$, chamado de *Erro de Predição Forward*, como:

$$f_L(n) = u(n) - \sum_{k=1}^L w_{fk}^* \cdot u(n-k) \quad n = L+1, \dots, N \quad (2.1)$$

onde o subscrito L denota a ordem do filtro e o subscrito f, o tipo de predição a ser realizada.

O filtro da fig. 2.2 é denominado *Filtro de Erro de Predição Linear Forward* e é obtido pela minimização da soma dos erros quadráticos de predição, definida como:

$$\epsilon_f(w_{f1}, w_{f2}, \dots, w_{fL}) = \sum_{n=L+1}^N |f_L(n)|^2 \Big|_{w_{f0}^* = -1} \quad (2.2)$$

Denominamos este método de projeto como *Método de Predição Linear Forward - FLP¹*.

2.3.2 A EQUAÇÃO NORMAL DETERMINÍSTICA

A eq. (2.2) pode ser interpretada como a *Energia do Erro de Predição Forward*. O problema aqui é encontrar w_{fi} , $i = 1, \dots, L$, com $w_{f0} = -1$, tal que esta energia seja a menor possível. São definidos a seguir, alguns vetores e matrizes necessários ao desenvolvimento da solução deste problema.

1) Vetor de ganhos do filtro transversal - w_f : dimensões $L \times 1$.

$$w_f^t = [w_{f1} \ w_{f2} \ \dots \ w_{fL}] \quad (2.3)$$

2) Vetor de entrada - \mathbf{u} : dimensões $L \times 1$.

Contém as entradas das derivações do filtro.

$$\mathbf{u}^T(n-1) = [u(n-1) \ u(n-2) \ \dots \ u(n-L)] \quad (2.4)$$

Portanto, podemos expressar $f_L(n)$, o erro de predição forward, como:

$$f_L(n) = u(n) - \mathbf{w}_f^H \cdot \mathbf{u}(n-1) \quad n = L+1, \dots, N \quad (2.5)$$

onde: o sobre-escrito H denota o transposto conjugado (Hermitiano).

3) Vetor erro de estimação forward - \mathcal{F} : dimensões $(N-L) \times 1$.

$$\mathcal{F}^H = [f_L(L+1) \ f_L(L+2) \ \dots \ f_L(N)] \quad (2.6)$$

4) Vetor resposta desejada forward - \mathbf{b}_f : dimensões $(N-L) \times 1$.

$$\mathbf{b}_f^H = [u(L+1) \ u(L+2) \ \dots \ u(N)] \quad (2.7)$$

Podemos reescrever a eq. (2.6) como:

$$\begin{aligned} \mathcal{F}^H &= \mathbf{b}_f^H - \mathbf{w}_f^H \cdot [u(L) \ u(L+1) \ \dots \ u(N-1)] \\ \mathcal{F}^H &= \mathbf{b}_f^H - \mathbf{w}_f^H \cdot \mathbf{A}_f^H \end{aligned} \quad (2.8)$$

onde: $\mathbf{A}_f^H = [u(L) \ u(L+1) \ \dots \ u(N-1)]$ (2.9)
 \mathbf{A}_f tem dimensões $(N-L) \times L$.

Tomando o Hermitiano de ambos os lados de (2.8), obtemos:

$$\mathcal{F} = \mathbf{b}_f - \mathbf{A}_f \cdot \mathbf{w}_f \quad (2.10)$$

Expressando \mathbf{A}_f^H na sua forma expandida:

$$\mathbf{A}_f^H = \begin{bmatrix} u(L) & u(L+1) & \dots & u(N-1) \\ u(L-1) & u(L) & \dots & u(N-2) \\ \vdots & \vdots & \vdots & \vdots \\ u(1) & u(2) & \dots & u(N-L) \end{bmatrix} \quad (2.11)$$

Podemos observar que A_f é completamente definida pelas amostras disponíveis, $u(1), u(2), \dots, u(N)$.

De posse dessas informações, a eq. (2.2) pode ser reescrita como:

$$\epsilon_f(\mathbf{w}_f) = \mathcal{F}^H \mathcal{F} \quad (2.12)$$

Substituído a eq. (2.8) em (2.12), obtemos:

$$\begin{aligned} \epsilon_f(\mathbf{w}_f) &= (\mathbf{b}_f^H - \mathbf{w}_f^H \cdot \mathbf{A}_f^H) \cdot (\mathbf{b}_f - \mathbf{A}_f \cdot \mathbf{w}_f) \\ &= \mathbf{b}_f^H \cdot \mathbf{b}_f - \mathbf{b}_f^H \cdot \mathbf{A}_f \cdot \mathbf{w}_f - \mathbf{w}_f^H \cdot \mathbf{A}_f^H \cdot \mathbf{b}_f + \mathbf{w}_f^H \cdot \mathbf{A}_f^H \cdot \mathbf{A}_f \cdot \mathbf{w}_f \end{aligned} \quad (2.13)$$

Diferenciando a eq. (2.13) em relação a \mathbf{w}_f [1-pg 105 a 107], obtemos:

$$\frac{\partial \epsilon_f(\mathbf{w}_f)}{\partial \mathbf{w}_f} = - 2 \cdot \mathbf{A}_f^H \cdot \mathbf{b}_f + 2 \cdot \mathbf{A}_f^H \cdot \mathbf{A}_f \cdot \mathbf{w}_f \quad (2.14)$$

Igualando a eq. (2.14) a zero para encontrar o vetor \mathbf{w}_f que minimiza a energia do erro de predição forward ϵ_f , obtemos:

$$\begin{aligned} - 2 \cdot \mathbf{A}_f^H \cdot \mathbf{b}_f + 2 \cdot \mathbf{A}_f^H \cdot \mathbf{A}_f \cdot \mathbf{w}_f &= 0 \\ \mathbf{A}_f^H \cdot \mathbf{A}_f \cdot \mathbf{w}_f &= \mathbf{A}_f^H \cdot \mathbf{b}_f \end{aligned} \quad (2.15)$$

A equação acima (2.15) é chamada de *Equação Normal Determinística* para o problema dos mínimos quadrados. A resolução desta equação, que permite obter \mathbf{w}_f ótimo, será discutida mais adiante.

A fim de obter o valor mínimo da energia do erro de predição forward, correspondendo ao preditor ótimo, basta tomar a expressão para ϵ_f , dada por (2.13), e substituir \mathbf{w}_f pelo seu valor ótimo, dado pela eq.(2.15), resultando em:

$$\begin{aligned} \epsilon_{fmin}(\mathbf{w}_f) &= \mathbf{b}_f^H \cdot \mathbf{b}_f - \mathbf{b}_f^H \cdot \mathbf{A}_f \cdot \mathbf{w}_f - \mathbf{w}_f^H \cdot \mathbf{A}_f^H \cdot \mathbf{b}_f + \mathbf{w}_f^H \cdot \mathbf{A}_f^H \cdot \mathbf{b}_f \\ \epsilon_{fmin}(\mathbf{w}_f) &= \mathbf{b}_f^H \cdot \mathbf{b}_f - \mathbf{b}_f^H \cdot \mathbf{A}_f \cdot \mathbf{w}_f \end{aligned} \quad (2.16)$$

2.3.3 A EQUAÇÃO NORMAL DETERMINÍSTICA E AS FUNÇÕES DE CORRELAÇÃO

Por conveniência, transcrevemos a equação normal determinística.

$$\mathbf{A}_f^H \cdot \mathbf{A}_f \cdot \mathbf{w}_f = \mathbf{A}_f^H \cdot \mathbf{b}_f \tag{2.17}$$

Observando (2.17), podemos concluir que:

1) O vetor $\mathbf{A}_f^H \cdot \mathbf{b}_f$ será:

$$\mathbf{A}_f^H \cdot \mathbf{b}_f = \begin{bmatrix} u(L) & u(L+1) & \dots & u(N-1) \\ u(L-1) & u(L) & \dots & u(N-2) \\ \vdots & \vdots & \vdots & \vdots \\ u(1) & u(2) & \dots & u(N-L) \end{bmatrix} \cdot \begin{bmatrix} u^*(L+1) \\ u^*(L+2) \\ \vdots \\ u^*(N) \end{bmatrix}$$

$$\mathbf{A}_f^H \cdot \mathbf{b}_f = (N - L) \cdot \begin{bmatrix} \hat{r}(-1) \\ \hat{r}(-2) \\ \vdots \\ \hat{r}(-L) \end{bmatrix} = (N - L) \cdot \hat{\mathbf{r}}_f \tag{2.18}$$

Onde:
$$\hat{r}(i) = \frac{1}{N - L} \cdot \sum_{k=L}^{N-1} u(k+i+1) \cdot u^*(k+1) \quad ; -L \leq i \leq -1 \tag{2.19}$$

ou seja, $\hat{\mathbf{r}}_f$ é um vetor de correlação estimada para $u(n)$, formado pelas estimativas da sequência de correlação $\hat{r}(i)$, $-L \leq i \leq -1$. No limite, quando $N \rightarrow \infty$, e considerando o processo ergódico, $\hat{r}(i) \rightarrow r(i)$, onde $r(i)$ é a verdadeira sequência de correlação do processo $u(n)$ [2].

2) A Matriz $\mathbf{A}_f^H \cdot \mathbf{A}_f$ será:

$$\begin{aligned} \mathbf{A}_f^H \cdot \mathbf{A}_f &= \begin{bmatrix} u(L) & u(L+1) & \dots & u(N-1) \\ u(L-1) & u(L) & \dots & u(N-2) \\ \vdots & \vdots & \vdots & \vdots \\ u(1) & u(2) & \dots & u(N-L) \end{bmatrix} \cdot \begin{bmatrix} u^*(L) & u^*(L-1) & \dots & u^*(1) \\ u^*(L+1) & u^*(L) & \dots & u^*(2) \\ \vdots & \vdots & \vdots & \vdots \\ u^*(N-1) & u^*(N-2) & \dots & u^*(N-L) \end{bmatrix} \\ &= (N - L) \cdot \begin{bmatrix} \hat{r}(0,0) & \hat{r}(0,-1) & \dots & \hat{r}(0,-L+1) \\ \hat{r}(-1,0) & \hat{r}(-1,-1) & \dots & \hat{r}(-1,-L+1) \\ \vdots & \vdots & \vdots & \vdots \\ \hat{r}(-L+1,0) & \hat{r}(-L+1,-1) & \dots & \hat{r}(-L+1,-L+1) \end{bmatrix} = (N - L) \cdot \hat{\mathbf{R}}_f \end{aligned} \tag{2.20}$$

Onde:
$$\hat{r}(i,j) = \frac{1}{N-L} \cdot \sum_{K=L}^{N-1} u(k+i) \cdot u^*(k+j), (-L+1) \leq i, j \leq 0 \quad (2.21)$$

$\hat{\mathbf{R}}_f$ é uma matriz de correlação estimada para $u(n)$ e é chamada de matriz de correlação determinística. No limite, quando $N \rightarrow \infty$, e considerando o processo ergódico, $\hat{r}(i,j) \rightarrow r(i-j)$, onde $r(i-j)$ é a verdadeira sequência de correlação para $u(n)$, no deslocamento $(i-j)$ [2].

Usando as relações desenvolvidas em (2.18) e (2.20), a equação normal em (2.17) pode ser reescrita da seguinte forma:

$$\hat{\mathbf{R}}_f \cdot \mathbf{w}_f = \hat{\mathbf{r}}_f \quad (2.22)$$

A aplicação da teoria de filtragem ótima de Wiener ao problema de predição linear leva à equação normal dada por:

$$\mathbf{R} \cdot \mathbf{w}_f = \mathbf{r} \quad (2.23)$$

onde se emprega médias estatísticas [1].

Fica evidente, observando os resultados dos limites analisados em (2.19) e (2.21), que a equação (2.22) é o resultado da equação normal clássica de (2.23), quando substituímos a matriz e o vetor de correlação, \mathbf{R} e \mathbf{r} , por estimativas polarizadas [2] dos mesmos.

Seguindo a mesma linha de raciocínio, a expressão para o erro mínimo de predição forward, dado pela eq. (2.16), pode ser reescrita como:

$$\epsilon_{fmin}(\mathbf{w}_f) = \mathbf{b}_f^H \cdot \mathbf{b}_f - (N-L) \cdot \hat{\mathbf{r}}_f^H \cdot \mathbf{w}_f \quad (2.24)$$

2.3.4 PROPRIEDADES DA MATRIZ DE CORRELAÇÃO DETERMINÍSTICA

Algumas propriedades da matriz de correlação determinística $\hat{\mathbf{R}}_f$, as quais serão úteis mais adiante, serão agora enunciadas. As demonstrações podem ser encontradas em [1].

Propriedade 2.1: A matriz de correlação determinística $\hat{\mathbf{R}}_f$ é Hermitiana.

Propriedade 2.2: A matriz de correlação determinística $\hat{\mathbf{R}}_f$ é definida não-negativa, isto é, $\mathbf{x}^H \cdot \hat{\mathbf{R}}_f \cdot \mathbf{x} \geq 0$, para qualquer vetor $\mathbf{x}_{(L \times 1)}$.

O fato de $\hat{\mathbf{R}}_f$ ser definida não-negativa significa que o seu determinante e todos os menores principais são não-negativos. Se a condição acima é satisfeita com o sinal de desigualdade, seu determinante e todos os menores principais são positivos. Neste caso, $\hat{\mathbf{R}}_f$ é não-singular e $\hat{\mathbf{R}}_f^{-1}$ existe [3].

Propriedade 2.3: Os autovalores da matriz de correlação determinística $\hat{\mathbf{R}}_f$ são todos reais e não negativos.

Propriedade 2.4: A matriz de correlação determinística $\hat{\mathbf{R}}_f$ é o produto de duas matrizes retangulares de Toeplitz.

Esta propriedade é importante para a solução da equação normal determinística (2.22). O Algoritmo de Levinson-Durbin [1] resolveria a eq. (2.22) caso $\hat{\mathbf{R}}_f$ fosse Hermitiana, Toeplitz, e $\hat{\mathbf{R}}_f^{-1}$ existisse. Entretanto, o prejuízo não é total, uma vez que Marple [2] desenvolveu um algoritmo que explora a característica de $\hat{\mathbf{R}}_f$ descrita na propriedade 2.4. Embora este algoritmo seja mais complexo computacionalmente que o de Levinson-Durbin, ainda assim é mais eficiente do que tentar resolver (2.22) através de outros métodos que invertem a matriz $\hat{\mathbf{R}}_f$.

2.4 O MÉTODO DE PREDIÇÃO LINEAR BACKWARD - BLP

2.4.1 O FILTRO DE ERRO DE PREDIÇÃO BACKWARD

Supondo, na fig. 2.1, que $w_L^* = -1$ e denominando $f[U(n)]$ de $b_L(n)$, resulta o esquema da fig. 2.3.

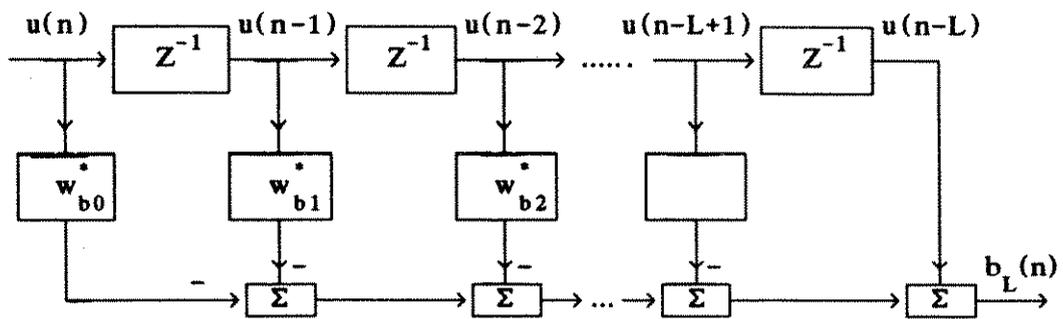


Fig.2.3 - Filtro de Erro de Predição Linear Backward

Utilizamos as entradas $(u(n), u(n-1), \dots, u(n-L+1))$ das derivações para fazer uma predição linear de $u(n-L)$. Esta predição é realizada através de amostras posteriores a $u(n-L)$, o que motiva a denominação de Predição Backward (regressiva). Definimos $b_L(n)$, o *Erro de Predição Backward*, como:

$$b_L(n) = u(n-L) - \sum_{k=0}^{L-1} w_{bk}^* u(n-k) \quad n = L+1, \dots, N \quad (2.25)$$

Por questão de conveniência das deduções matemáticas que se seguirão, o erro de predição expresso por (2.25), embora se refira à amostra $u(n-L)$, é aqui representado por $b_L(n)$.

O filtro da fig. 2.3 é denominado *Filtro de Erro de Predição Linear Backward* e é projetado minimizando-se a soma dos erros quadráticos de predição, definida como:

$$\epsilon_b(w_{b0}, w_{b1}, \dots, w_{b(L-1)}) = \sum_{n=L+1}^N |b_L(n)|^2 \Big|_{w_{bL}^* = -1} \quad (2.26)$$

Este método de projeto é denominado *Método de Predição Linear Backward - BLP²*.

Para um dado conjunto de dados $(u(1), \dots, u(N))$, os métodos FLP e BLP fornecem filtros inteiramente diferentes, pois este conjunto de amostras é utilizado de forma distinta em cada um deles, como demonstram as expressões (2.1) e (2.25).

2.4.2 A EQUAÇÃO NORMAL DETERMINÍSTICA

A eq. (2.26) pode ser interpretada como a *Energia do Erro de Predição Backward* e o problema está em encontrar $w_{bi}, i = 0, \dots, L-1$ com $w_{bL} = -1$, tal que esta energia seja a menor possível. Assim, da mesma forma que no caso forward, definimos:

1) Vetor de ganhos do filtro transversal - w_b : dimensões $L \times 1$.

$$w_b^t = [w_{b(L-1)}^* \ w_{b(L-2)}^* \ \dots \ w_{b0}^*] \quad (2.27)$$

² do inglês Backward Linear Prediction

2) Vetor de entrada - u : dimensões $L \times 1$.

Contém as entradas das derivações do filtro.

$$u^t(n-L+1) = [u^*(n-L+1) \ u^*(n-L+2) \ \dots \ u^*(n)] \quad (2.28)$$

Portanto, podemos expressar $b_L^*(n)$, complexo conjugado do erro de predição backward, como:

$$b_L^*(n) = u^*(n-L) - w_b^H \cdot u(n-L+1) \quad n = L+1, \dots, N \quad (2.29)$$

3) Vetor erro de estimação backward - \mathcal{B} : dimensões $(N-L) \times 1$.

$$\mathcal{B}^H = [b_L^*(L+1) \ b_L^*(L+2) \ \dots \ b_L^*(N)] \quad (2.30)$$

4) Vetor resposta desejada backward - b_b : dimensões $(N-L) \times 1$.

$$b_b^H = [u^*(1) \ u^*(2) \ \dots \ u^*(N-L)] \quad (2.31)$$

Podemos reescrever a eq. (2.30) como:

$$\begin{aligned} \mathcal{B}^H &= b_b^H - w_b^H \cdot [u(2) \ u(3) \ \dots \ u(N-L+1)] \\ \mathcal{B}^H &= b_b^H - w_b^H \cdot A_b^H \end{aligned} \quad (2.32)$$

onde: $A_b^H = [u(2) \ u(3) \ \dots \ u(N-L+1)] \quad (2.33)$
 A_b tem dimensões $(N-L) \times L$.

Tomando o Hermitiano de ambos os lados de (2.32), obtemos:

$$\mathcal{B} = b_b - A_b \cdot w_b \quad (2.34)$$

É conveniente expressar A_b^H na sua forma expandida.

$$A_b^H = \begin{bmatrix} u^*(2) & u^*(3) & \dots & u^*(N-L+1) \\ u^*(3) & u^*(4) & \dots & u^*(N-L+2) \\ \vdots & \vdots & \vdots & \vdots \\ u^*(L+1) & u^*(L+2) & \dots & u^*(N) \end{bmatrix} \quad (2.35)$$

A eq. (2.26) pode agora ser reescrita como:

$$\epsilon_b(\mathbf{w}_b) = \mathcal{B}^H \cdot \mathcal{B} \quad (2.36)$$

Notamos, portanto, que as equações para os filtros forward e backward são idênticas na forma, embora os vetores envolvidos sejam distintos. Logo, a repetição do desenvolvimento realizado no ítem anterior permite afirmar que a equação normal determinística para o filtro backward será:

$$\mathbf{A}_b^H \cdot \mathbf{A}_b \cdot \mathbf{w}_b = \mathbf{A}_b^H \cdot \mathbf{b}_b \quad (2.37)$$

A expressão da mínima energia do erro de predição backward, será portanto:

$$\epsilon_{b\min}(\mathbf{w}_b) = \mathbf{b}_b^H \cdot \mathbf{b}_b - \mathbf{b}_b^H \cdot \mathbf{A}_b \cdot \mathbf{w}_b \quad (2.38)$$

Vale salientar que \mathbf{w}_b , na equação (2.38), é o vetor ótimo segundo o critério do mínimos quadrados, dado por (2.37).

Tudo o que foi deduzido no contexto da predição forward, envolvendo a matriz e o vetor de correlação, é válido também para o caso backward.

2.5 O MÉTODO DA PREDIÇÃO LINEAR FORWARD-BACKWARD - FBLP

Podemos gerar ainda um terceiro método de projeto de preditores, combinando os métodos forward e backward. O critério de otimização consiste em minimizar a energia do erro de predição forward somada à energia do erro de predição backward, ou seja, os coeficientes dos filtros de erro de predição serão tais que minimizem a energia do erro de predição forward-backward dada por:

$$\epsilon_{fb}(\mathbf{w}_{fb}) = \sum_{k=L+1}^N [|f_L(k)|^2 + |b_L(k)|^2]$$

$$\epsilon_{fb}(\mathbf{w}_{fb}) = \epsilon_f(\mathbf{w}_f) + \epsilon_b(\mathbf{w}_b) \quad (2.39)$$

Este método é denominado de *Método de Predição Linear Forward-Backward - FBLP*³.

³ do Inglês Forward Backward Linear Prediction

Repetimos, por conveniência, as equações normais determinísticas para os preditores forward e backward.

a) Forward

$$\mathbf{A}_f^H \cdot \mathbf{A}_f \cdot \mathbf{w}_f = \mathbf{A}_f^H \cdot \mathbf{b}_f \tag{2.40}$$

ou: $\hat{\mathbf{R}}_f \cdot \mathbf{w}_f = \hat{\mathbf{r}}_f \tag{2.41}$

b) Backward

$$\mathbf{A}_b^H \cdot \mathbf{A}_b \cdot \mathbf{w}_b = \mathbf{A}_b^H \cdot \mathbf{b}_b \tag{2.42}$$

ou: $\hat{\mathbf{R}}_b \cdot \mathbf{w}_b = \hat{\mathbf{r}}_b \tag{2.43}$

Admitindo a ergodicidade, no limite quando $N \rightarrow \infty$, observamos que:

$$\hat{\mathbf{R}}_f \rightarrow \mathbf{R} \quad \hat{\mathbf{r}}_f \rightarrow \mathbf{r} \tag{2.44}$$

$$\hat{\mathbf{R}}_b \rightarrow \mathbf{R} \quad \hat{\mathbf{r}}_b \rightarrow \mathbf{r}$$

ou: $\hat{\mathbf{R}}_f = \hat{\mathbf{R}}_b \quad \text{e} \quad \hat{\mathbf{r}}_f = \hat{\mathbf{r}}_b \tag{2.45}$

Onde \mathbf{R} e \mathbf{r} são a matriz e o vetor de correlação, respectivamente. Consequentemente (2.41) e (2.43) tornam-se:

$$\mathbf{R} \cdot \mathbf{w}_f = \mathbf{r} \tag{2.46}$$

e: $\mathbf{R} \cdot \mathbf{w}_b = \mathbf{r}$

o que implica em: $\mathbf{w}_f = \mathbf{w}_b \tag{2.47}$

ou seja:

$$[w_{f1} \ w_{f2} \ \dots \ w_{fL}] = [w_{b(L-1)}^* \ w_{b(L-2)}^* \ \dots \ w_{b0}^*] \tag{2.48}$$

Observando (2.48), constatamos que, quando $N \rightarrow \infty$, os coeficientes do filtro de erro de predição backward são os complexos conjugados dos coeficientes do forward escritos na ordem inversa.

O método FBLP consiste em projetar os filtros forward e backward de modo a minimizar a soma das energias dos erros de predição forward e backward, fornecida pela eq. (2.39). Mas, minimizar (2.39) significa obter :

$$\frac{\partial \epsilon_f(\mathbf{w}_f)}{\partial \mathbf{w}_f} + \frac{\partial \epsilon_b(\mathbf{w}_b)}{\partial \mathbf{w}_b} = 0 \tag{2.49}$$

$$- 2.A_f^H . \mathbf{b}_f + 2.A_f^H . A_f . \mathbf{w}_f - 2.A_b^H . \mathbf{b}_b + 2.A_b^H . A_b . \mathbf{w}_b = 0 \tag{2.50}$$

Embora as soluções \mathbf{w}_f e \mathbf{w}_b de (2.50) sejam, no caso geral, distintas, quando N é finito, no método FBLP assumimos a validade da eq. (2.48), mesmo para N finito.

A figura 2.4 mostra os filtros de erro de predição forward e backward a serem projetados pelo método FBLP.

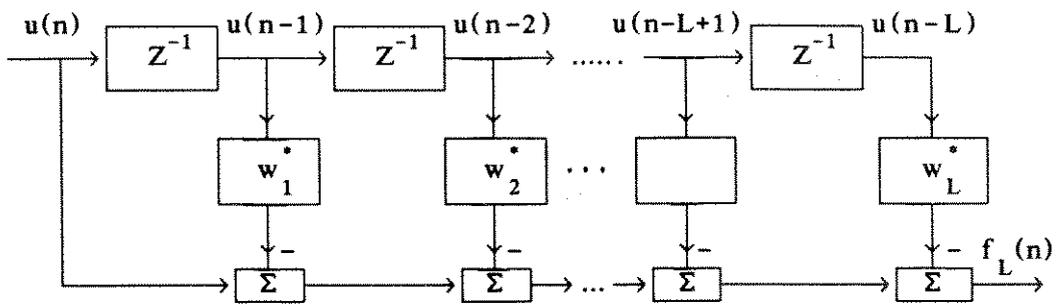


Fig. 2.4a - Filtro de Erro de Predição Linear Forward

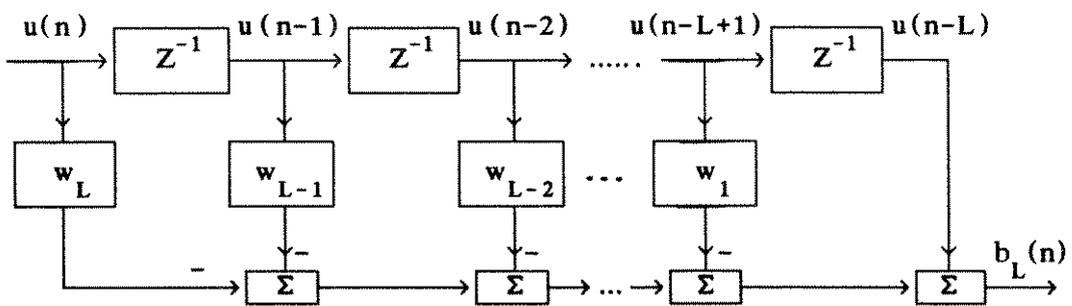


Fig. 2.4b - Filtro de Erro de Predição Linear Backward

Os subscritos para \mathbf{w} , referentes à forward e backward, foram retirados, por conveniência de notação.

Assim, temos um caso de otimização conjunta de \mathbf{w}_f e \mathbf{w}_b , sujeito à restrição $\mathbf{w}_f = \mathbf{w}_b = \mathbf{w}$. A eq. (2.50) torna-se então:

$$- 2.A_f^H .b_f + 2.A_f^H .A_f .w - 2.A_b^H .b_b + 2.A_b^H .A_b .w = 0 \quad (2.51)$$

$$A_f^H .A_f .w + A_b^H .A_b .w = A_f^H .b_f + A_b^H .b_b \quad (2.52)$$

$$[A_f^H \quad | \quad A_b^H] . \begin{bmatrix} A_f \\ A_b \end{bmatrix} .w = [A_f^H \quad | \quad A_b^H] . \begin{bmatrix} b_f \\ b_b \end{bmatrix}$$

$$A^H .A .w = A^H .b \quad (2.53)$$

onde:

. A é a matriz de dados forward-backward, de dimensões 2(N-L)xL, dada por:

$$A^H = [A_f^H \quad | \quad A_b^H] =$$

$$\begin{bmatrix} u(L) & u(L+1) & \dots & u(N-1) & u^*(2) & u^*(3) & \dots & u^*(N-L+1) \\ u(L-1) & u(L) & \dots & u(N-2) & u^*(3) & u^*(4) & \dots & u^*(N-L+2) \\ \vdots & \vdots \\ u(1) & u(2) & \dots & u(N-L) & u^*(L+1) & u^*(L+2) & \dots & u^*(N) \end{bmatrix} \quad (2.54)$$

. b é o vetor resposta desejada forward-backward, de dimensões 2(N-L)x1, dado por:

$$b^H = [b_f^H \quad | \quad b_b^H] =$$

$$[u(L+1), u(L+2), \dots, u(N) \quad u^*(1), u^*(2), \dots, u^*(N-L)] \quad (2.55)$$

Onde: N → Número de amostras disponível do sinal

L → Ordem do filtro

As equações (2.53), (2.54) e (2.55) determinam o w ótimo segundo o método FBLP, isto é, a solução representa uma otimização conjunta sujeita à restrição $w_f = w_b$.

As colunas da metade esquerda de A^H representam os conjuntos de entradas das derivações, utilizadas para obter a predição linear forward, das respectivas respostas desejadas contidas na metade esquerda de b^H . As colunas da metade direita de A^H representam os complexos conjugados dos conjuntos de entradas das derivações, utilizadas para obter a predição linear backward, dos complexos conjugados das respectivas respostas desejadas, contidas na metade direita de b^H .

Por analogia, a energia mínima do erro de predição linear forward-backward

será (ver eq. (2.16) e (2.38)):

$$\epsilon_{\min}(\mathbf{w}) = \mathbf{b}^H \cdot \mathbf{b} - \mathbf{b}^H \cdot \mathbf{A} \cdot \mathbf{w} \quad (2.56)$$

onde \mathbf{w} é o vetor ótimo dado por (2.53).

Tudo o que foi deduzido no contexto das predições forward e backward, envolvendo a matriz e o vetor de correlação, é válido também para o caso forward-backward.

A motivação para o estudo do método FBLP vem da expectativa de resultados promissores para a estimativa dos ganhos dos filtros. Tal expectativa resulta de uma utilização mais completa das amostras de sinal disponíveis, propiciada pela aglutinação das equações dos métodos iniciais.

Para solucionar as equações normais determinísticas, dadas por (2.53), (2.37) e (2.17), é necessário o conhecimento de alguns teoremas básicos de álgebra matricial, bem como de resolução de sistemas de equações. Este estudo será feito logo abaixo, para posteriormente, analisarmos as possíveis soluções destas equações.

2.6 O TEOREMA DA DECOMPOSIÇÃO VIA VALOR SINGULAR - SVD

TEOREMA 2.1: Se $\mathbf{A}_{(m \times n)}$ tem posto k , então é possível escrever:

$$\mathbf{U}^H \cdot \mathbf{A} \cdot \mathbf{V} = \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

e:

$$\mathbf{A} = \mathbf{U} \cdot \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \mathbf{V}^H$$

onde: $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_k)$.

. $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k > 0$, são chamados *valores singulares de A*.

. σ_i^2 , $i = 1, 2, \dots, k$, são os autovalores não nulos de $\mathbf{A}^H \cdot \mathbf{A}$ e de $\mathbf{A} \cdot \mathbf{A}^H$.

. $\mathbf{U} \rightarrow$ Matriz unitária, de dimensões $m \times m$, onde suas colunas são os autovetores de $\mathbf{A} \cdot \mathbf{A}^H$.

. $\mathbf{V} \rightarrow$ Matriz unitária, de dimensões $n \times n$, onde suas colunas são os autovetores de $\mathbf{A}^H \cdot \mathbf{A}$.

. As primeiras k colunas de \mathbf{U} correspondem aos autovetores de $\mathbf{A} \cdot \mathbf{A}^H$ associados aos autovalores σ_i^2 , $i = 1, \dots, k$, e as $(m - k)$ colunas restantes correspondem aos autovetores de $\mathbf{A} \cdot \mathbf{A}^H$ associados aos autovalores nulos.

- . As primeiras k colunas de V correspondem aos autovetores de $A^H.A$ associados aos autovalores σ_i^2 , $i = 1, \dots, k$, e as $(n - k)$ colunas restantes correspondem aos autovetores de $A^H.A$ associados aos autovalores nulos.
- . Tomando U e V nas formas particionadas por colunas:

$$U = [u_1 \ u_2 \ \dots \ u_m]$$

$$V = [v_1 \ v_2 \ \dots \ v_n]$$

U e V guardam a seguinte relação:

$$u_i = \sigma_i^{-1} . A . v_i$$

$$i = 1, 2, \dots, k$$

$$v_i = \sigma_i^{-1} . A^H . u_i$$

DEMONSTRAÇÃO: Sendo $\text{posto}(A^H.A) = \text{posto}(A.A^H) = \text{posto}(A)$, e este por sua vez, igual a k, é evidente que a matriz $A^H.A$ tem exatamente $(n - k)$ autovalores nulos, e $A.A^H$ tem $(m - k)$ autovalores nulos. Em ambos os casos, os k autovalores restantes são reais e positivos (ver apêndice A).

Podemos denominar os autovalores não nulos de $A^H.A$ como σ_i^2 , $i = 1, 2, \dots, k$, já que todos são positivos, e ordená-los na forma $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k > 0$.

É sempre possível escolher um conjunto ortonormal de autovetores v_1, v_2, \dots, v_k para $A^H.A$, associados respectivamente com os autovalores $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$ (ver apêndice B). Definimos agora:

$$u_i = \sigma_i^{-1} . A . v_i \tag{2.57a}$$

$$i = 1, 2, \dots, k$$

$$A . v_i = \sigma_i . u_i \tag{2.57b}$$

Demonstraremos agora que, $u_i, i = 1, 2, \dots, k$, formam um conjunto ortonormal de autovetores para $A.A^H$, associados aos autovalores $\sigma_i^2, i = 1, 2, \dots, k$.

Tomando, inicialmente, o produto $A.A^H.u_i$ e substituindo u_i pela expressão (2.57a), obtemos:

$$A.A^H.u_i = A.A^H.(\sigma_i^{-1} . A . v_i) = \sigma_i^{-1} . A . (A^H . A . v_i) \tag{2.58}$$

$$i = 1, 2, \dots, k$$

Já que $\sigma_i^2, i = 1, 2, \dots, k$, são autovalores de $A^H.A$, associados a $v_i, i = 1, 2, \dots, k$, então:

$$\mathbf{A}^H \cdot \mathbf{A} \cdot \mathbf{v}_i = \sigma_i^2 \cdot \mathbf{v}_i \quad i = 1, 2, \dots, k \quad (2.59)$$

Substituindo (2.59) em (2.58):

$$\mathbf{A} \cdot \mathbf{A}^H \cdot \mathbf{u}_i = \sigma_i^{-1} \cdot \mathbf{A} \cdot \sigma_i^2 \cdot \mathbf{v}_i = \sigma_i^2 \cdot (\sigma_i^{-1} \cdot \mathbf{A} \cdot \mathbf{v}_i) \quad i = 1, 2, \dots, k \quad (2.60)$$

Substituindo (2.57a) em (2.60), vem:

$$\mathbf{A} \cdot \mathbf{A}^H \cdot \mathbf{u}_i = \sigma_i^2 \cdot \mathbf{u}_i \quad i = 1, 2, \dots, k \quad (2.61)$$

E está provado que $\mathbf{u}_i, i = 1, 2, \dots, k$, são autovetores de $\mathbf{A} \cdot \mathbf{A}^H$ associados a $\sigma_i^2, i = 1, 2, \dots, k$, que são os k autovalores não nulos.

Para mostrar que são ortogonais, tomamos o seguinte produto interno:

$$\begin{aligned} (\mathbf{u}_i, \mathbf{u}_j) &= \mathbf{u}_i^H \cdot \mathbf{u}_j = (\sigma_i^{-1} \cdot \mathbf{A} \cdot \mathbf{v}_i, \sigma_j^{-1} \cdot \mathbf{A} \cdot \mathbf{v}_j) = \sigma_i^{-1} \cdot \sigma_j^{-1} (\mathbf{v}_i, \mathbf{A}^H \cdot \mathbf{A} \cdot \mathbf{v}_j) \\ &= \sigma_i^{-1} \cdot \sigma_j^{-1} \cdot (\mathbf{v}_i, \sigma_j^2 \cdot \mathbf{v}_j) = \sigma_i^{-1} \cdot \sigma_j \cdot (\mathbf{v}_i, \mathbf{v}_j) \end{aligned}$$

$$\therefore (\mathbf{u}_i, \mathbf{u}_j) = 0 \quad \forall i \neq j \quad \text{pois} \quad (\mathbf{v}_i, \mathbf{v}_j) = 0 \quad \forall i \neq j$$

$$(\mathbf{u}_i, \mathbf{u}_i) = 1 \quad \forall i = j \quad \text{pois} \quad \sigma_i^{-1} \cdot \sigma_i \cdot (\mathbf{v}_i, \mathbf{v}_i) = 1$$

Fica provado, desta forma que, $\mathbf{u}_i, i = 1, 2, \dots, k$, são ortonormais.

Utilizando (2.57a), temos:

$$\mathbf{A}^H \cdot \mathbf{u}_i = \sigma_i^{-1} \cdot \mathbf{A}^H \cdot \mathbf{A} \cdot \mathbf{v}_i \quad i = 1, 2, \dots, k$$

substituindo (2.59) na expressão acima:

$$\mathbf{A}^H \cdot \mathbf{u}_i = \sigma_i \cdot \mathbf{v}_i \quad i = 1, 2, \dots, k$$

$$\mathbf{v}_i = \sigma_i^{-1} \cdot \mathbf{A}^H \cdot \mathbf{u}_i$$

que, juntamente com (2.57a), complementa a relação entre \mathbf{v}_i e $\mathbf{u}_i, i = 1, 2, \dots, k$.

Já que $\mathbf{A}^H \cdot \mathbf{A}$ e $\mathbf{A} \cdot \mathbf{A}^H$ têm k autovalores não nulos, $(n - k)$ e $(m - k)$ autovalores nulos, respectivamente, e como foram exibidos k autovalores não nulos (que são os σ_i^2), para cada uma das matrizes, então estes autovalores são os únicos não nulos. Portanto, é possível estender os conjuntos $\{\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_k\}$ e $\{\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_k\}$

para conjuntos ortonormais $U = (u_1 \ u_2 \ \dots \ u_m)$ e $V = (v_1 \ v_2 \ \dots \ v_n)$ (ver apêndice B), onde:

$$A \cdot A^H \cdot u_i = 0 \quad i = K+1, K+2, \dots, m \quad (2.62)$$

$$A^H \cdot A \cdot v_i = 0 \quad i = k+1, k+2, \dots, n \quad (2.63)$$

A eq. (2.63) implica em (ver apêndice A):

$$A \cdot v_i = 0 \quad i = k+1, k+2, \dots, n \quad (2.64)$$

Demonstrando a parte final do teorema, segue que:

$$\begin{aligned}
 U^H \cdot A \cdot V &= \begin{bmatrix} u_1^H \\ u_2^H \\ \vdots \\ u_m^H \end{bmatrix} \cdot A \cdot [v_1 \ v_2 \ \dots \ v_n] \\
 &= \begin{bmatrix} u_1^H \\ u_2^H \\ \vdots \\ u_m^H \end{bmatrix} \cdot [A \cdot v_1 \ A \cdot v_2 \ \dots \ A \cdot v_k \ A \cdot v_{k+1} \ \dots \ A \cdot v_n]
 \end{aligned}$$

Utilizando (2.64), vem:

$$U^H \cdot A \cdot V = \begin{bmatrix} u_1^H \\ u_2^H \\ \vdots \\ u_m^H \end{bmatrix} \cdot [A \cdot v_1 \ A \cdot v_2 \ \dots \ A \cdot v_k \ 0 \ 0 \ \dots \ 0]$$

Utilizando (2.57b):

$$U^H \cdot A \cdot V = \begin{bmatrix} \sigma_1 \cdot u_1^H \cdot u_1 & \sigma_2 \cdot u_1^H \cdot u_2 & \dots & \sigma_k \cdot u_1^H \cdot u_k & 0 & \dots & 0 \\ \sigma_1 \cdot u_2^H \cdot u_1 & \sigma_2 \cdot u_2^H \cdot u_2 & \dots & \sigma_k \cdot u_2^H \cdot u_k & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \sigma_1 \cdot u_k^H \cdot u_1 & \sigma_2 \cdot u_k^H \cdot u_2 & \dots & \sigma_k \cdot u_k^H \cdot u_k & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \sigma_1 \cdot u_m^H \cdot u_1 & \sigma_2 \cdot u_m^H \cdot u_2 & \dots & \sigma_k \cdot u_m^H \cdot u_k & 0 & \dots & 0 \end{bmatrix}$$

como os vetores $u_i, i = 1, 2, \dots, m$, são ortonormais, obtemos:

$$U^H \cdot A \cdot V = \begin{bmatrix} \sigma_1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \sigma_k & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix}$$

Portanto:

$$U^H \cdot A \cdot V = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \tag{2.65}$$

Onde U, V e Σ satisfazem o enunciado do teorema. Como U e V são matrizes unitárias, é imediato que:

$$A = U \cdot \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \cdot V^H \tag{2.66}$$

A equação (2.65) é a expressão matemática da *Decomposição via Valor Singular - SVD*⁴.

É importante notar que a SVD não é necessariamente única, pois é possível escolher vários conjuntos ortonormais U e V. Nenhuma restrição é feita, também, quantos aos valores de m e n.

⁴do Inglês Singular Value Decomposition

2.7 DEFINIÇÃO GERAL DE PSEUDO-INVERSA DE UMA MATRIZ

Definimos a pseudo-inversa de uma matriz A como:

$$A^{\#} = V \cdot \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} \cdot U^H \tag{2.67}$$

Onde: $\Sigma^{-1} = \text{diag}(\sigma_1^{-1}, \sigma_2^{-1}, \dots, \sigma_k^{-1})$, e $\sigma_i, i = 1, 2, \dots, k$, são os valores singulares de A . k é o posto de A .

- $U \rightarrow$ Matriz unitária, de dimensões $m \times m$, onde suas colunas são os autovetores de $A \cdot A^H$.
- $V \rightarrow$ Matriz unitária, de dimensões $n \times n$, onde suas colunas são os autovetores de $A^H \cdot A$.

De (2.66), é imediato que:

$$A^H = V \cdot \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \cdot U^H \tag{2.68}$$

Cabe aqui, citar algumas igualdades envolvendo autovalores e autovetores, que, a posteriori, tornar-se-ão necessárias. Lembrando que, para uma matriz $B_{(p \times p)}$ qualquer [1]:

se: $B \cdot x_i = \lambda_i \cdot x_i \quad i = 1, 2, \dots, p$

então: $B^n \cdot x_i = \lambda_i^n \cdot x_i \quad i = 1, 2, \dots, p$

É imediato deduzirmos que:

1) autovetores de $(A^H \cdot A)^H \cdot (A^H \cdot A) =$ autovetores de $(A^H \cdot A) \cdot (A^H \cdot A)^H =$
autovetores de $(A^H \cdot A) \cdot (A^H \cdot A) =$ autovetores de $(A^H \cdot A) = V$

2) autovalores de $(A^H \cdot A)^H \cdot (A^H \cdot A) =$ autovalores de $(A^H \cdot A) \cdot (A^H \cdot A)^H =$
= autovalores de $(A^H \cdot A) \cdot (A^H \cdot A) =$
[autovalores de $(A^H \cdot A)^2 = \{ \sigma_1^4, \sigma_2^4, \dots, \sigma_k^4 \}$]

3) autovetores de $(A \cdot A^H)^H \cdot (A \cdot A^H) =$ autovetores de $(A \cdot A^H) \cdot (A \cdot A^H)^H =$
autovetores de $(A \cdot A^H) \cdot (A \cdot A^H) =$ autovetores de $(A \cdot A^H) = U$

4) autovalores de $(\mathbf{A}\mathbf{A}^H)(\mathbf{A}\mathbf{A}^H)^H =$ autovalores de $(\mathbf{A}\mathbf{A}^H)^H(\mathbf{A}\mathbf{A}^H) =$
 $=$ autovalores de $(\mathbf{A}\mathbf{A}^H)(\mathbf{A}\mathbf{A}^H) =$
 [autovalores de $(\mathbf{A}\mathbf{A}^H)^2 = \{ \sigma_1^4, \sigma_2^4, \dots, \sigma_k^4 \}$]

Agora, utilizando as igualdades acima e a teoria de SVD, chegamos às seguintes expressões:

$$(\mathbf{A}^H\mathbf{A})^\# = \mathbf{V} \cdot \begin{bmatrix} \Sigma^{-2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \mathbf{V}^H \tag{2.69}$$

$$(\mathbf{A}\mathbf{A}^H)^\# = \mathbf{U} \cdot \begin{bmatrix} \Sigma^{-2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \mathbf{U}^H \tag{2.70}$$

Utilizando (2.68) e (2.69), obtemos o produto:

$$\begin{aligned} (\mathbf{A}^H\mathbf{A})^\# \cdot \mathbf{A}^H &= \mathbf{V} \cdot \begin{bmatrix} \Sigma^{-2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \mathbf{V}^H \cdot \mathbf{V} \cdot \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \mathbf{U}^H \\ &= \mathbf{V} \cdot \begin{bmatrix} \Sigma^{-2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \mathbf{U}^H \\ &= \mathbf{V} \cdot \begin{bmatrix} \Sigma^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \mathbf{U}^H = \mathbf{A}^\# \end{aligned}$$

Portanto: $\mathbf{A}^\# = (\mathbf{A}^H\mathbf{A})^\# \cdot \mathbf{A}^H \tag{2.71}$

Utilizando (2.68) e (2.70), obtemos:

$$\begin{aligned} \mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^\# &= \mathbf{V} \cdot \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \mathbf{U}^H \cdot \mathbf{U} \cdot \begin{bmatrix} \Sigma^{-2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \mathbf{U}^H \\ &= \mathbf{V} \cdot \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \begin{bmatrix} \Sigma^{-2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \mathbf{U}^H \\ &= \mathbf{V} \cdot \begin{bmatrix} \Sigma^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \mathbf{U}^H = \mathbf{A}^\# \end{aligned}$$

Portanto: $\mathbf{A}^\# = \mathbf{A}^H \cdot (\mathbf{A}\mathbf{A}^H)^\# \tag{2.72}$

(2.71) e (2.72) constituem, portanto, expressões alternativas para A^* .

Caso uma dada matriz seja quadrada e não-singular, sua pseudo-inversa, ou, inversa generalizada, é a sua própria matriz inversa. Esta afirmação tem demonstração imediata: se A é quadrada e não singular, o posto de A é cheio e a expressão:

$$A^* = V \cdot \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} \cdot U^H$$

dá lugar a:

$$A^* = V \cdot \Sigma^{-1} \cdot U^H \tag{2.73}$$

Lembrando que, pela mesma razão, podemos escrever:

$$A = U \cdot \Sigma \cdot V^H \tag{2.74}$$

Então:

$$A \cdot A^* = U \cdot \Sigma \cdot V^H \cdot V \cdot \Sigma^{-1} \cdot U^H = I \tag{2.75}$$

$$A^* \cdot A = V \cdot \Sigma^{-1} \cdot U^H \cdot U \cdot \Sigma \cdot V^H = I \tag{2.76}$$

Portanto:

$$A^* = A^{-1} \tag{2.77}$$

ou seja, existe a inversa (pois o posto é cheio) e ela é igual à pseudo-inversa. Concluimos que a inversa de uma matriz é um caso particular da sua pseudo-inversa, quando a matriz é quadrada e não-singular.

Supondo que a matriz $A_{(m \times n)}$ possua posto k cheio, a sua pseudo-inversa toma formas particulares. Dois casos devem ser analisados.

1) $m \geq n$ e $k = n$

$A^H \cdot A$ é quadrada, de dimensões $n \times n$ e tem posto cheio, pois:
 posto($A^H \cdot A$) = posto(A) = n . Baseado em (2.77):

$$(A^H \cdot A)^* = (A^H \cdot A)^{-1}$$

Reescrevendo (2.71), resulta:

$$A^* = (A^H \cdot A)^{-1} \cdot A^H \tag{2.78}$$

2) $m \leq n$ e $k = m$

$A \cdot A^H$ é quadrada, de dimensões $m \times m$ e tem posto cheio, pois:

$\text{posto}(\mathbf{A}\mathbf{A}^H) = \text{posto}(\mathbf{A}) = m$. Baseado em (2.77):

$$(\mathbf{A}\mathbf{A}^H)^\# = (\mathbf{A}\mathbf{A}^H)^{-1}$$

Reescrevendo (2.72), resulta:

$$\mathbf{A}^\# = \mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1} \tag{2.79}$$

Neste caso, para calcular a pseudo-inversa de \mathbf{A} , devemos optar, entre (2.78) e (2.79), por aquela expressão que for mais simples e/ou adequada, observando as dimensões m e n de \mathbf{A} .

2.8 RESOLUÇÃO DA EQUAÇÃO NORMAL DETERMINÍSTICA

Antes de analisar as possíveis soluções das equações normais determinísticas, dadas por (2.53), (2.37) e (2.17), algumas definições e teoremas tornam-se necessários.

DEFINIÇÃO 2.1: Uma matriz quadrada \mathbf{A} que satisfaz $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$ é chamada *matriz normal*.

Toda matriz Hermitiana é normal, pois:

$$\mathbf{A}\mathbf{A}^H = \mathbf{A}\mathbf{A} = \mathbf{A}^H\mathbf{A}$$

TEOREMA 2.2: Seja uma matriz $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_n]$, na sua forma particionada por colunas, formando um conjunto ortogonal em um espaço V . Se $\mathbf{v} = \alpha_1.\mathbf{v}_1 + \alpha_2.\mathbf{v}_2 + \dots + \alpha_n.\mathbf{v}_n$, então:

$$\alpha_i(\mathbf{v}_i, \mathbf{v}_i) = (\mathbf{v}_i, \mathbf{v}) \quad ; \quad i = 1, 2, \dots, n$$

DEMONSTRAÇÃO: Considerando o produto interno de \mathbf{v}_i por \mathbf{v} :

$$\begin{aligned} (\mathbf{v}_i, \mathbf{v}) &= (\mathbf{v}_i, \alpha_1.\mathbf{v}_1 + \alpha_2.\mathbf{v}_2 + \dots + \alpha_n.\mathbf{v}_n) \\ &= \alpha_1.(\mathbf{v}_i, \mathbf{v}_1) + \alpha_2.(\mathbf{v}_i, \mathbf{v}_2) + \dots + \alpha_n.(\mathbf{v}_i, \mathbf{v}_n) \\ &= \alpha_i.(\mathbf{v}_i, \mathbf{v}_i) \end{aligned}$$

TEOREMA 2.3: Considerando as n equações lineares, com n incógnitas, descritas por $(A - \lambda.I).x = b$, onde A é uma matriz $n \times n$ normal, λ é um escalar dado e b é um vetor $n \times 1$, verifica-se uma das seguintes alternativas:

a) λ não é um dos autovalores λ_j de A . Neste caso, a solução é única e pode ser escrita como:

$$x = \sum_{j=1}^n \frac{(x_j, b)}{\lambda_j - \lambda} \cdot x_j$$

Onde: $X = [x_1 \ x_2 \ \dots \ x_n]$ é uma base ortonormal constituída por autovetores de A .

b) $\lambda = \lambda_j$ para algum autovalor λ_j de A . Neste caso existe solução x se e somente se b é ortogonal a todos os autovetores x_j associados com o autovalor λ_j . Se x qualquer é solução, então podem ser encontradas infinitas soluções, adicionando a x uma combinação linear arbitrária dos autovetores associados a λ_j .

DEMONSTRAÇÃO: Sendo $[x_1 \ x_2 \ \dots \ x_n]$ uma base ortonormal, constituída por autovetores de A (isto é sempre possível, ver apêndice B), podemos escrever:

$$b = b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_n \cdot x_n \tag{2.80}$$

$$x = \alpha_1 \cdot x_1 + \alpha_2 \cdot x_2 + \dots + \alpha_n \cdot x_n \tag{2.81}$$

Pelo teorema 2.2:

$$b_1 \cdot (x_1, x_1) = b_1 = (x_1, b)$$

$$b_2 \cdot (x_2, x_2) = b_2 = (x_2, b)$$

e assim por diante.

Portanto, a eq. (2.80) torna-se:

$$b = (x_1, b) \cdot x_1 + (x_2, b) \cdot x_2 + \dots + (x_n, b) \cdot x_n \tag{2.82}$$

Ao substituir (2.81) e (2.82) na equação $(A - \lambda.I).x = b$, obtemos:

$$(A - \lambda.I) \cdot \sum_{j=1}^n \alpha_j \cdot x_j = \sum_{j=1}^n (x_j, b) \cdot x_j \tag{2.83}$$

$$\sum_{j=1}^n (\alpha_j \cdot A \cdot x_j - \alpha_j \cdot \lambda \cdot x_j) = \sum_{j=1}^n (x_j, b) \cdot x_j \quad (2.84)$$

Como $x_j, j = 1, 2, \dots, n$ são autovetores de A:

$$A \cdot x_j = \lambda_j \cdot x_j \quad j = 1, 2, \dots, n \quad (2.85)$$

(2.85) em (2.84) fornece:

$$\begin{aligned} \sum_{j=1}^n \alpha_j \cdot (\lambda_j - \lambda) \cdot x_j &= \sum_{j=1}^n (x_j, b) \cdot x_j \\ \sum_{j=1}^n [\alpha_j \cdot (\lambda_j - \lambda) - (x_j, b)] \cdot x_j &= 0 \end{aligned} \quad (2.86)$$

Como $x_j, j = 1, 2, \dots, n$, são ortogonais, cada componente não pode ser expressa através de uma combinação linear das demais. Logo, para que (2.86) tenha solução, é necessário que:

$$\alpha_j \cdot (\lambda_j - \lambda) = (x_j, b) \quad j = 1, 2, \dots, n \quad (2.87)$$

Dois casos devem ser observados:

1) $\lambda \neq \lambda_j, j = 1, 2, \dots, n$

Então a solução será única e igual a:

$$\alpha_j = \frac{(x_j, b)}{\lambda_j - \lambda}$$

E a solução do sistema de equações será:

$$x = \sum_{j=1}^n \alpha_j \cdot x_j \Rightarrow x = \sum_{j=1}^n \frac{(x_j, b)}{\lambda_j - \lambda} \cdot x_j \quad (2.88)$$

2) Suponha que A tenha um autovalor de multiplicidade s. Sejam $\lambda_k = \lambda_{k+1} = \dots = \lambda_{k+s-1}$ estes autovalores, isto é, $\lambda_{k+i}, i = 0, 1, \dots, s-1$, e k é um valor particular de j da eq (2.87). Suponha também que $\lambda = \lambda_{k+i}, i = 0, 1, \dots, s-1$. Então, por (2.87):

$$\alpha_{k+i} \cdot 0 = (x_{k+i}, b) \quad i = 0, 1, \dots, s-1 \quad (2.89a)$$

Portanto: $(\mathbf{x}_{k+i}, \mathbf{b}) = 0 \quad i = 0, 1, \dots, s-1 \quad (2.89b)$

O que indica que, para o sistema ter solução, \mathbf{b} deve ser perpendicular a todos os autovetores \mathbf{x}_{k+i} associados a λ_{k+i} . Se λ_{k+i} tem multiplicidade s , e, portanto, possui s autovetores ortogonais associados a ele, \mathbf{b} tem que ser ortogonal a estes s autovetores. Caso contrário, não será possível encontrar α_{k+i} que satisfaça (2.89) e o sistema não terá solução.

Supondo que (2.89) seja satisfeita, as soluções para o sistema tornam-se:

$$\alpha_j = \frac{(\mathbf{x}_j, \mathbf{b})}{\lambda_j - \lambda} \quad j = 1, 2, \dots, n \text{ e } j \neq (k+i), i = 0, 1, \dots, s-1$$

e

$$\alpha_j = \text{qualquer} \quad j = (k+i), i = 0, 1, \dots, s-1$$

pois se (2.89) for satisfeita quando $\lambda_j = \lambda_{k+i} = \lambda$, então:

$$\alpha_{k+i} \cdot 0 = (\mathbf{x}_{k+i}, \mathbf{b}) \quad \therefore \alpha_{k+i} \cdot 0 = 0$$

e α_{k+i} pode ser qualquer para $i = 0, 1, \dots, s-1$

Portanto, as soluções tornam-se:

$$\mathbf{x} = \sum_{\substack{j=1 \\ \lambda_j \neq \lambda}}^n \frac{(\mathbf{x}_j, \mathbf{b})}{\lambda_j - \lambda} \cdot \mathbf{x}_j + \sum_{\substack{j=1 \\ \lambda_j = \lambda}}^n \alpha_j \cdot \mathbf{x}_j \quad , \text{supondo que } (\mathbf{x}_j, \mathbf{b}) = 0 \quad (2.90)$$

quando $\lambda_j = \lambda$
 α_j qualquer

Se α_j pode ser qualquer todas as vezes que $\lambda_j = \lambda$, então o sistema terá infinitas soluções. O desenvolvimento abaixo explicita estas soluções.

Denominando os autovetores ortogonais associados a $\lambda_{k+i} = \lambda$ de $\mathbf{x}_{k+i}, i = 0, 1, \dots, s-1$, então:

$$\mathbf{A} \cdot \mathbf{x}_{k+i} = \lambda_{k+i} \cdot \mathbf{x}_{k+i} \quad i = 0, 1, \dots, s-1 \quad (2.91)$$

Suponhamos que (2.89) seja satisfeita, ou seja:

$$(\mathbf{x}_{k+i}, \mathbf{b}) = 0 \quad i = 0, 1, \dots, s-1$$

Tomando $(\mathbf{A} - \lambda \cdot \mathbf{I}) \cdot \mathbf{x} = \mathbf{b}$, e lembrando que $\lambda_{k+i} = \lambda, i = 0, 1, \dots, s-1$, temos:

$$(A - \lambda_{k+i}.I).x = b \quad i = 0,1,\dots,s-1 \quad (2.92)$$

onde x , dado pela equação (2.90), é solução para este sistema de equações. Agora, tomando:

$$x' = x + \sum_{i=0}^{s-1} \beta_i \cdot x_{k+i} \quad (2.93)$$

$$\begin{aligned} \text{Então: } (A - \lambda_{k+i}.I).x' &= (A - \lambda_{k+i}.I).(x + \sum_{i=0}^{s-1} \beta_i \cdot x_{k+i}) \\ &= (A - \lambda_{k+i}.I).x + \sum_{i=0}^{s-1} \beta_i \cdot A \cdot x_{k+i} - \sum_{i=0}^{s-1} \beta_i \cdot \lambda_{k+i} \cdot x_{k+i} \end{aligned}$$

Utilizando (2.91), segue que:

$$\begin{aligned} (A - \lambda_{k+i}.I).x' &= \\ (A - \lambda_{k+i}.I).x + \sum_{i=0}^{s-1} \beta_i \cdot \lambda_{k+i} \cdot x_{k+i} - \sum_{i=0}^{s-1} \beta_i \cdot \lambda_{k+i} \cdot x_{k+i} &= \\ (A - \lambda_{k+i}.I).x' = (A - \lambda_{k+i}.I).x = (A - \lambda.I).x' = b \end{aligned}$$

Portanto x' também é solução dos sistema de equações $(A - \lambda.I).x = b$. Esta afirmação indica que, adicionando a x uma combinação linear qualquer dos autovetores associados a $\lambda_{k+i}, i = 0,1,\dots,s-1$, como dado por (2.93), obtemos uma nova solução para o sistema de equações. Neste caso, infinitas soluções podem ser encontradas, o que conclui a prova do teorema.

Resumindo, para o sistema de equações $(A - \lambda.I).x = b$, três situações distintas podem ocorrer:

- 1) $\lambda \neq \lambda_j, j = 1,2,\dots,n \Rightarrow$ o sistema tem solução única
- 2) $\lambda = \lambda_{k+i}, i = 0,1,\dots,s-1$ e (2.89) é satisfeita para $i = 0,1,\dots,s-1 \Rightarrow$ o sistema tem infinitas soluções
- 3) $\lambda = \lambda_{k+i}, i = 0,1,\dots,s-1$ e (2.89) não é satisfeita para $i = 1,2,\dots,s-1 \Rightarrow$ o sistema não tem solução

Finalmente, agora existem condições para analisarmos as possíveis soluções das equações normais determinísticas. Para efeito de simplicidade de notação tomaremos uma equação normal genérica, representando (2.53), (2.37) e (2.17), fornecida por:

$$\mathbf{A}^H \cdot \mathbf{A} \cdot \mathbf{w} = \mathbf{A}^H \cdot \mathbf{b} \tag{2.94}$$

É evidente que todas as considerações que faremos, daqui por diante, para (2.94), são válidas para (2.53), (2.37) e (2.17).

Esta equação recai em um caso particular da equação do teorema 2.3. É imediato particularizar o teorema para o sistema de (2.94).

TEOREMA 2.4: Considerando as n equações lineares com n incógnitas, descritas por $\mathbf{A}^H \cdot \mathbf{A} \cdot \mathbf{w} = \mathbf{A}^H \cdot \mathbf{b}$, onde $\mathbf{A}^H \cdot \mathbf{A}$ é uma matriz $n \times n$ Hermitiana, verifica-se uma das seguintes alternativas:

a) 0 não é um dos autovalores λ_j de $\mathbf{A}^H \cdot \mathbf{A}$, e a solução única pode ser escrita como:

$$\mathbf{w} = \sum_{j=1}^n \frac{\mathbf{v}_j^H \cdot \mathbf{A}^H \cdot \mathbf{b}}{\lambda_j} \cdot \mathbf{v}_j = \sum_{j=1}^n \frac{\mathbf{v}_j}{\lambda_j} \cdot \mathbf{v}_j^H \cdot \mathbf{A}^H \cdot \mathbf{b} \tag{2.95}$$

Onde: $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_n]$ é uma base ortonormal constituída por autovetores de $\mathbf{A}^H \cdot \mathbf{A}$.

b) $\lambda_j = 0$ é algum autovalor de $\mathbf{A}^H \cdot \mathbf{A}$, e existe solução \mathbf{w} se e somente se $\mathbf{A}^H \cdot \mathbf{b}$ é ortogonal a todos os autovetores \mathbf{v}_j associados com o autovalor 0. Se \mathbf{w} qualquer é solução, então podem ser encontradas infinitas soluções, adicionando a \mathbf{w} uma combinação linear arbitrária dos autovetores associados a 0.

É importante ter em mente os resultados obtidos no teorema da decomposição do valor singular, e pseudo-inversa da matriz \mathbf{A} . É imediato perceber que a matriz \mathbf{V} do teorema 2.4 é a mesma matriz \mathbf{V} obtida na SVD.

Para analisar as soluções de (2.94), consideramos:

a) $\mathbf{A}^H \cdot \mathbf{A}$ não possui autovalores nulos

Pelo ítem (a) do teorema 2.4, se 0 não é autovalor de $\mathbf{A}^H \cdot \mathbf{A}$, (2.94) tem solução única. Também, a ausência deste autovalor implica que o posto de $\mathbf{A}^H \cdot \mathbf{A}$ é

cheio (ver apêndice A) e, portanto, igual a n . Se o posto é cheio, a matriz é não-singular e possui inversa [3]. Logo, a solução única será:

$$w = (A^H \cdot A)^{-1} \cdot A^H \cdot b \tag{2.96}$$

De imediato, pode parecer que (2.96) e (2.95) são diferentes. Porém, como $A^H \cdot A$ é quadrada e não-singular:

$$(A^H \cdot A)^{-1} = (A^H \cdot A)^{\#} \tag{2.97}$$

Neste caso, a expressão para a pseudo-inversa de $A^H \cdot A$, dada pela eq. (2.69), torna-se:

$$(A^H \cdot A)^{\#} = V \cdot \Sigma^{-2} \cdot V^H \tag{2.98}$$

Portanto:

$$(A^H \cdot A)^{-1} = V \cdot \Sigma^{-2} \cdot V^H = [v_1 \ v_2 \ \dots \ v_n] \cdot \Sigma^{-2} \cdot \begin{bmatrix} v_1^H \\ v_2^H \\ \vdots \\ v_n^H \end{bmatrix}$$

$$(A^H \cdot A)^{-1} = \sum_{j=1}^n \frac{v_j \cdot v_j^H}{\sigma_j^2} = \sum_{j=1}^n \frac{v_j \cdot v_j^H}{\lambda_j} \tag{2.99}$$

onde: $\lambda_j = \sigma_j^2, j = 1, 2, \dots, n$, são os autovalores de $A^H \cdot A$.

Substituindo (2.99) em (2.96), obtemos:

$$w = \sum_{j=1}^n \frac{v_j}{\lambda_j} \cdot v_j^H \cdot A^H \cdot b$$

Que é exatamente a eq. (2.95). Portanto (2.95) e (2.96) se equivalem e representam a solução quando (2.94) possui solução única.

Passamos a analisar agora o outro caso:

b) $A^H.A$ possui algum autovalor nulo

$A^H.A$ é singular e $(A^H.A)^{-1}$ não existe, o que é coerente com o fato de (2.94) não possuir solução única. Neste caso, pelo item (b) do teorema 2.4, só existirá solução (infinitas) caso $A^H.b$ seja ortogonal a todos os autovetores v_j associados com o autovalor 0. Pela SVD da matriz A , sabemos que os autovetores de $A^H.A$ associados com o autovalor 0 são: $v_{j,j} = k+1, k+2, \dots, n$. Fazendo o teste de ortogonalidade:

$$A = U \cdot \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \cdot V^H$$

$$A^H = V \cdot \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}^t \cdot U^H$$

$$v_i^H(A^H.b) = v_i^H \cdot V \cdot \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}^t \cdot U^H.b$$

$$= v_i^H \cdot [v_1 \ v_2 \ \dots \ v_n] \cdot \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}^t \cdot U^H.b$$

Para $i = k+1, k+2, \dots, n$:

$$v_i^H(A^H.b) = [0 \ 0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0 \ 0] \cdot \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}^t \cdot U^H.b = 0 \cdot U^H.b = 0$$

pois o elemento 1, do primeiro vetor acima, encontra-se na posição $i = k+1, k+2, \dots, n$.

Portanto: $v_i^H(A^H.b) = 0 \quad i = k+1, k+2, \dots, n$

ou seja, autovetores associados a autovalores nulos são ortogonais a $A^H.b$, o que significa que neste caso, (2.94) possui infinitas soluções. Já que o caso a) ou o b) ocorrem necessariamente, enunciemos o seguinte importante teorema.

O TEOREMA DA UNICIDADE: O sistema $A^H.A.w = A^H.b$ sempre tem solução. A solução será única se o posto de $A_{(m \times n)}$ for igual a n . Caso o posto de A seja diferente de n , o sistema terá infinitas soluções.

2.9 A SOLUÇÃO DE NORMA MÍNIMA

Caso o posto de A não seja n , o problema linear dos mínimos quadrados, definido pela equação normal determinística, tem infinitas soluções. Uma das soluções terá um especial interesse: aquela que fornece o vetor de ganhos w de norma mínima, ou seja, aquela na qual : $\|w\|^2 = w^H \cdot w$ é o mínimo possível. Esta solução será dada por:

$$w = A^* \cdot b \tag{2.100}$$

onde A^* é a pseudo-inversa de A .

A eq. (2.100) é conhecida como *A Solução de Norma Mínima* para a equação normal determinística. Seguimos agora com a demonstração destes fatos.

Inicialmente, manipulamos a eq. (2.56) da energia mínima do erro de predição.

$$\epsilon_{\min} = b^H \cdot b - b^H \cdot A \cdot w$$

$$\epsilon_{\min} = b^H \cdot (b - A \cdot w)$$

onde w são as possíveis soluções de (2.94), ou seja, é o w ótimo segundo o critério dos mínimos quadrados.

Utilizando o fato de que, no teorema da SVD, U e V são unitárias:

$$\epsilon_{\min} = b^H \cdot U \cdot U^H \cdot (b - A \cdot V \cdot V^H \cdot w) \tag{2.101}$$

$$= b^H \cdot U \cdot (U^H \cdot b - U^H \cdot A \cdot V \cdot V^H \cdot w)$$

Particionamos U e V da seguinte forma:

$$U = [U_1 \ U_2] \tag{2.102}$$

$$V = [V_1 \ V_2] \tag{2.103}$$

onde:

$$U_1 = [u_1 \ u_2 \ \dots \ u_k]$$

$$U_2 = [u_{k+1} \ u_{k+2} \ \dots \ u_m]$$

k é o posto de A

$$\begin{aligned} V_1 &= [v_1 \ v_2 \ \dots \ v_k] \\ V_2 &= [v_{k+1} \ v_{k+2} \ \dots \ v_n] \end{aligned}$$

Definimos agora:

$$z = V^H \cdot w = \begin{bmatrix} V_1^H \\ V_2^H \end{bmatrix} \cdot w = \begin{bmatrix} V_1^H \cdot w \\ V_2^H \cdot w \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \tag{2.104}$$

onde: z_1 é um vetor $k \times 1$
 z_2 é um vetor $(n-k) \times 1$

$$c = U^H \cdot b = \begin{bmatrix} U_1^H \\ U_2^H \end{bmatrix} \cdot b = \begin{bmatrix} U_1^H \cdot b \\ U_2^H \cdot b \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tag{2.105}$$

onde: c_1 é um vetor $k \times 1$
 c_2 é um vetor $(m-k) \times 1$

Substituindo (2.65), (2.104) e (2.105) em (2.101), temos:

$$\begin{aligned} \epsilon_{\min} &= b^H \cdot U \cdot \left\{ \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} - \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \right\} \\ \epsilon_{\min} &= b^H \cdot U \cdot \left\{ \begin{bmatrix} c_1 - \Sigma \cdot z_1 \\ c_2 \end{bmatrix} \right\} \end{aligned} \tag{2.106}$$

Vamos agora demonstrar que $c_1 = \Sigma \cdot z_1$ e que, portanto, $\epsilon_{\min} = b^H \cdot U \cdot \begin{bmatrix} 0 \\ c_2 \end{bmatrix}$. De (2.66) é

imediato que:

$$A \cdot V = U \cdot \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \tag{2.107}$$

A eq.(2.104) fornece: $w = V \cdot z$

Portanto (2.94) torna-se:

$$\mathbf{A}^H \cdot \mathbf{A} \cdot \mathbf{V} \cdot \mathbf{z} = \mathbf{A}^H \cdot \mathbf{b}$$

Substituindo (2.107) na expressão acima obtemos:

$$\mathbf{A}^H \cdot \mathbf{U} \cdot \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \mathbf{z} = \mathbf{A}^H \cdot \mathbf{b} \quad (2.108)$$

Por outro lado, (2.105) possibilita escrevermos:

$$\mathbf{b} = \mathbf{U} \cdot \mathbf{c}$$

$$\mathbf{A}^H \cdot \mathbf{b} = \mathbf{A}^H \cdot \mathbf{U} \cdot \mathbf{c} \quad (2.109)$$

Comparando (2.108) e (2.109) temos:

$$\begin{aligned} \mathbf{A}^H \cdot \mathbf{U} \cdot \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \mathbf{z} &= \mathbf{A}^H \cdot \mathbf{U} \cdot \mathbf{c} \\ \mathbf{A}^H \cdot \mathbf{U} \cdot \left\{ \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} - \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \right\} &= \mathbf{0} \\ \mathbf{A}^H \cdot \mathbf{U} \cdot \begin{bmatrix} \Sigma \cdot z_1 - c_1 \\ -c_2 \end{bmatrix} &= \mathbf{0} \end{aligned} \quad (2.110)$$

mas (2.68) fornece:

$$\mathbf{A}^H \cdot \mathbf{U} = \mathbf{V} \cdot \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (2.111)$$

Substituindo (2.111) em (2.110):

$$\mathbf{V} \cdot \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \begin{bmatrix} \Sigma \cdot z_1 - c_1 \\ -c_2 \end{bmatrix} = \mathbf{0}$$

substituindo (2.103) na expressão acima:

$$[\mathbf{V}_1 \ \mathbf{V}_2] \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \begin{bmatrix} \Sigma \cdot \mathbf{z}_1 - \mathbf{c}_1 \\ -\mathbf{c}_2 \end{bmatrix} = \mathbf{0}$$

$$[\mathbf{V}_1 \cdot \Sigma \ \mathbf{0}] \cdot \begin{bmatrix} \Sigma \cdot \mathbf{z}_1 - \mathbf{c}_1 \\ -\mathbf{c}_2 \end{bmatrix} = \mathbf{0}$$

$$\mathbf{V}_1 \cdot \Sigma^2 \cdot \mathbf{z}_1 - \mathbf{V}_1 \cdot \Sigma \cdot \mathbf{c}_1 = \mathbf{0}$$

mas: $\mathbf{V}_1^H \cdot \mathbf{V}_1 = \mathbf{I}$

portanto: $\Sigma^2 \cdot \mathbf{z}_1 = \Sigma \cdot \mathbf{c}_1$

$$\mathbf{c}_1 = \Sigma \cdot \mathbf{z}_1 \tag{2.112a}$$

e: $\mathbf{z}_1 = \Sigma^{-1} \cdot \mathbf{c}_1$ (2.112b)

Substituindo \mathbf{c}_1 em (2.112b), podemos escrever \mathbf{z}_1 como:

$$\mathbf{z}_1 = \Sigma^{-1} \cdot \mathbf{U}_1^H \cdot \mathbf{b} \tag{2.113}$$

Os resultados de (2.112) e (2.113) foram obtidos para \mathbf{w} ótimo segundo o critério dos mínimos quadrados, ou seja, estas igualdades decorrem da solução ótima e, portanto, estão de acordo com a condição de erro mínimo.

Usando (2.112a), a eq. (2.106) pode ser reescrita da seguinte forma:

$$\boldsymbol{\epsilon}_{\min} = \mathbf{b}^H \cdot \mathbf{U} \cdot \begin{bmatrix} \mathbf{0} \\ \mathbf{c}_2 \end{bmatrix} \tag{2.114}$$

evidenciando que $\boldsymbol{\epsilon}_{\min}$ depende somente de \mathbf{c}_2 . Concluimos ainda que:

1) \mathbf{c}_1 , \mathbf{z}_1 e \mathbf{c}_2 estão univocamente determinados por (2.105) e (2.113), só restando definir o vetor \mathbf{z}_2 .

2) Observando (2.105) e (2.114) constatamos que, quando o posto $k = m$, U_2 , c_2 e portanto ϵ_{\min} são nulos. Logo, o sinal será totalmente predizível.

3) ϵ_{\min} independe de z_2 . Portanto, para qualquer z_2 escolhido, ϵ_{\min} dado por (2.114), ainda será mínimo segundo o critério dos mínimos quadrados. Por (2.104):

$$w = V \cdot \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \tag{2.115}$$

Observando (2.104), (2.115) e a conclusão 1, acima, constatamos que, quando o posto $k = n$, ratificando o teorema da unicidade, um único vetor w é obtido como solução. Caso isto não ocorra, z_2 pode ser qualquer e cada escolha fornecerá um w distinto, porém ótimo. Escolhendo $z_2 = 0$, e substituindo (2.112b), (2.115) torna-se:

$$w = V \cdot \begin{bmatrix} \Sigma^{-1} \cdot c_1 \\ 0 \end{bmatrix}$$

$$w = V \cdot \begin{bmatrix} \Sigma^{-1} & .0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

$$w = V \cdot \begin{bmatrix} \Sigma^{-1} & .0 \\ 0 & 0 \end{bmatrix} \cdot U^H \cdot b \tag{2.116}$$

Substituindo a definição de pseudo-inversa da matriz A , dado pela eq. (2.67), na equação acima:

$$w = A^* \cdot b \tag{2.117}$$

Fica então provado que (2.117) é uma solução para a equação normal determinística.

Tomando agora a norma euclidiana de w , utilizando (2.115) e (2.112b) temos:

$$w = V \cdot \begin{bmatrix} \Sigma^{-1} \cdot c_1 \\ z_2 \end{bmatrix}$$

$$\|w\|^2 = \|v\|^2 \cdot \left\| \begin{bmatrix} \Sigma^{-1} \cdot c_1 \\ z_2 \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} \Sigma^{-1} \cdot c_1 \\ z_2 \end{bmatrix} \right\|^2 \quad (2.118)$$

$$= [(\Sigma^{-1} \cdot c_1)^H \quad z_2^H] \cdot \begin{bmatrix} \Sigma^{-1} \cdot c_1 \\ z_2 \end{bmatrix}$$

Portanto: $\|w\|^2 = \|\Sigma^{-1} \cdot c_1\|^2 + \|z_2\|^2 \quad (2.119)$

Logo, é claro que a solução com $z_2 = 0$ é a de norma mínima. Fica mostrado assim que o vetor solução w , dado por (2.117), possui a menor norma euclidiana dentre todas as possíveis soluções.

Concluimos que $w = A^\# \cdot b$ é o único vetor solução para o problema em questão, que simultaneamente:

- 1) Minimiza a energia do erro de predição.
- 2) Possui a menor norma euclidiana dentre todas as possíveis soluções.

$w = A^\# \cdot b$ é conhecida como *A Solução de Norma Mínima*. Os motivos para o tratamento diferenciado dado a esta solução, serão justificados posteriormente.

Caso o sistema tenha solução única, é imediato provar que o vetor dado por $w = A^\# \cdot b$ (que no caso de infinitas soluções é denominado solução de norma mínima) é a própria solução única. Para isto, devemos observar que a equação (2.96) estabelece uma expressão para a solução única. Aplicando (2.97) em (2.96), obtemos:

$$W = (A^H \cdot A)^\# \cdot A^H \cdot b$$

Agora a expressão (2.71) permite escrever:

$$W = A^\# \cdot b$$

Evidenciando que, neste caso, calcular $w = A^\# \cdot b$ é equivalente a calcular a solução única.

2.10 OUTRA FORMULAÇÃO PARA A SOLUÇÃO DE NORMA MÍNIMA

A solução de norma mínima é fornecida por:

$$w = A^* . b$$

Substituindo (2.67):

$$w = V . \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} . U^H . b$$

Substituindo (2.102) e (2.103):

$$w = [V_1 \ V_2] . \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} . \begin{bmatrix} U_1^H \\ U_2^H \end{bmatrix} . b$$

$$w = [V_1 . \Sigma^{-1} \ 0] . \begin{bmatrix} U_1^H \\ U_2^H \end{bmatrix} . b$$

$$w = V_1 . \Sigma^{-1} . U_1^H . b \tag{2.120}$$

Percebemos, portanto, que w só depende dos k (posto de A) valores singulares não nulos(Σ), os seus correspondentes autovetores de $A^H . A(V_1)$ e de $A . A^H(U_1)$, e do vetor respostas desejadas(b).

Através de (2.107) chegamos a:

$$U_1 = A . V_1 . \Sigma^{-1} \Rightarrow U_1^H = \Sigma^{-1} . V_1^H . A^H \tag{2.121}$$

Substituindo (2.121) em (2.120) temos:

$$w = V_1 . \Sigma^{-1} . \Sigma^{-1} . V_1^H . A^H . b \tag{2.122}$$

$$w = [v_1 \ v_2 \ \dots \ v_k] . \begin{bmatrix} \sigma_1^{-2} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \sigma_k^{-2} \end{bmatrix} . \begin{bmatrix} v_1^H \\ \vdots \\ v_k^H \end{bmatrix} . A^H . b$$

$$\begin{aligned}
 \mathbf{w} &= \begin{bmatrix} \frac{\mathbf{v}_1}{\sigma_1^2} & \frac{\mathbf{v}_2}{\sigma_2^2} & \dots & \frac{\mathbf{v}_k}{\sigma_k^2} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{v}_1^H \\ \vdots \\ \mathbf{v}_k^H \end{bmatrix} \cdot \mathbf{A}^H \cdot \mathbf{b} \\
 \mathbf{w} &= \left[\frac{\mathbf{v}_1 \cdot \mathbf{v}_1^H}{\sigma_1^2} + \frac{\mathbf{v}_2 \cdot \mathbf{v}_2^H}{\sigma_2^2} + \dots + \frac{\mathbf{v}_k \cdot \mathbf{v}_k^H}{\sigma_k^2} \right] \cdot \mathbf{A}^H \cdot \mathbf{b} \\
 \mathbf{w} &= \sum_{i=1}^k \frac{\mathbf{v}_i}{\sigma_i^2} \cdot \mathbf{v}_i^H \cdot \mathbf{A}^H \cdot \mathbf{b} \tag{2.123}
 \end{aligned}$$

$\sigma_i^2, i = 1, 2, \dots, k$, são os valores singulares de \mathbf{A} elevados ao quadrado, ou equivalentemente, são os autovalores não-nulos de $\mathbf{A}^H \cdot \mathbf{A}$ e de $\mathbf{A} \cdot \mathbf{A}^H$. Denominando $\sigma_i^2 = \lambda_i$, temos:

$$\mathbf{w} = \sum_{i=1}^k \frac{\mathbf{v}_i}{\lambda_i} \cdot \mathbf{v}_i^H \cdot \mathbf{A}^H \cdot \mathbf{b} \tag{2.124}$$

A expressão (2.124) fornece a solução de norma mínima de \mathbf{w} em função dos k autovalores (λ_i) não nulos de $\mathbf{A}^H \cdot \mathbf{A}$ e seus respectivos autovetores (\mathbf{v}_i). Quando $k = n$, temos a situação de solução única, e (2.124) torna-se igual a (2.95).

Uma expressão alternativa a (2.124) pode ser obtida lembrando que o teorema da SVD assegura que:

$$\mathbf{v}_i = \sigma_i^{-1} \cdot \mathbf{A}^H \cdot \mathbf{u}_i \quad i = 1, 2, \dots, k$$

Escrevendo na forma matricial, temos:

$$\mathbf{V}_1 = \mathbf{A}^H \cdot \mathbf{U}_1 \cdot \Sigma^{-1} \tag{2.125}$$

obtendo, desta forma, \mathbf{V}_1 em função de \mathbf{U}_1 . Substituindo (2.125) em (2.120):

$$\mathbf{w} = \mathbf{A}^H \cdot \mathbf{U}_1 \cdot \Sigma^{-1} \cdot \Sigma^{-1} \cdot \mathbf{U}_1^H \cdot \mathbf{b}$$

$$\mathbf{w} = \mathbf{A}^H \cdot \mathbf{U}_1 \cdot \Sigma^{-2} \cdot \mathbf{U}_1^H \cdot \mathbf{b}$$

Seguindo o procedimento anterior, obtemos:

$$\mathbf{w} = \sum_{i=1}^k \frac{\mathbf{u}_i^H \cdot \mathbf{b}}{\sigma_i^2} \cdot \mathbf{A}^H \mathbf{u}_i \quad (2.126)$$

e:

$$\mathbf{w} = \sum_{i=1}^k \frac{\mathbf{u}_i^H \cdot \mathbf{b}}{\lambda_i} \cdot \mathbf{A}^H \mathbf{u}_i \quad (2.127)$$

A expressão (2.127) permite o cálculo da solução de norma mínima em função dos k autovalores (λ_i) não-nulos de $\mathbf{A} \cdot \mathbf{A}^H$ e seus respectivos autovetores (\mathbf{u}_i).

Obtivemos, portanto, as expressões alternativas (2.124) e (2.127), objetivando o cálculo da solução de norma mínima para a equação normal determinística. A eq. (2.124) envolve k autovalores e os respectivos autovetores de $\mathbf{A}^H \cdot \mathbf{A}$. A eq. (2.127) envolve k autovalores e os respectivos autovetores de $\mathbf{A} \cdot \mathbf{A}^H$. Sendo \mathbf{A} uma matriz de dimensões $m \times n$, é conveniente utilizar (2.124) quando $m > n$, de modo a minimizar o esforço computacional. Isto porque $\mathbf{v}_i, i = 1, 2, \dots, k$, têm dimensão $n \times 1$, enquanto que $\mathbf{u}_i, i = 1, 2, \dots, k$, têm dimensão $m \times 1$. Caso $m < n$, é óbvio que (2.127) deve ser utilizada.

Para obter os autovalores e autovetores envolvidos nestas equações devemos utilizar a SVD em detrimento à rotinas padrões para cálculo de auto-valores e auto-vetores de matrizes. O motivo para tal é que os algoritmos baseados em SVD são numericamente estáveis, o mesmo não ocorrendo para as outras rotinas [5].

Portanto, calcular a solução \mathbf{w} de norma mínima para a equação normal determinística requer o seguinte procedimento:

1) $m \geq n$

1.1) Calculamos SVD da matriz \mathbf{A} , obtendo os valores singulares $\sigma_1, \sigma_2, \dots, \sigma_k$ e os vetores singulares direitos $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ associados.

1.2) Utilizamos a fórmula (2.123) para o cálculo de \mathbf{w} .

2) $m < n$

2.1) Calculamos SVD da matriz \mathbf{A} , obtendo os valores singulares $\sigma_1, \sigma_2, \dots, \sigma_k$ e os vetores singulares esquerdos $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ associados.

2.2) Utilizamos a fórmula (2.126) para o cálculo de \mathbf{w} .

2.11 CONCLUSÕES

A predição linear segundo o critério dos mínimos quadrados foi desenvolvida nos contextos forward, backward e forward-backward. Nos três casos, o projeto do preditor recai na mesma equação: A Equação Normal Determinística, embora os vetores e matrizes envolvidos sejam distintos para cada caso. Obtivemos, também, as expressões para o valor mínimo da energia do erro de predição, alcançado quando o vetor de ganhos do preditor opera no seu valor ótimo.

Pelo teorema da unicidade, ficou provado que a equação normal determinística genérica, válida para os casos FLP, BLP e FBLP, sempre possui solução. Dependendo do posto da matriz de dados A , teremos uma única solução ou infinitas soluções.

Interesse especial foi dado à solução de norma mínima: $w = A^* \cdot b$, a qual foi atingida através do estudo do teorema da decomposição via valor singular e da pseudo-inversa A^* . Caso A possua posto cheio, o cálculo de A^* dispensa SVD, podendo ser efetuado através de inversão e produtos matriciais.

Para efeito operacional, foram desenvolvidas expressões alternativas para o cálculo da solução de norma mínima. Uma delas envolve autovalores e autovetores de $A^H \cdot A$, enquanto que a outra, autovalores e autovetores de $A \cdot A^H$. O motivo do destaque dado a esta solução será elucidado no próximo capítulo.

CAPÍTULO 3

DETECÇÃO DE SINAIS SENOIDAIS EM RUÍDO BRANCO

CONTEÚDO	Pag.
3.1 Introdução.....	47
3.2 Os Zeros do Filtro de Erro de Predição - Caso sem Ruído	48
3.2.1 Os Zeros de Sinal.....	48
3.2.2 Os Zeros Estranhos.....	59
3.3 Estimação das Frequências de Exponenciais	
Não-Amortecidas - Caso sem Ruído.....	64
3.3.1 Introdução.....	64
3.3.2 O Método FBLP para Estimar Frequências.....	64
3.4 Estimação das Frequências de Exponenciais	
Não-Amortecidas - Caso com Ruído.....	67
3.4.1 Introdução.....	67
3.4.2 O Subespaço de Sinal e o Subespaço de Ruído.....	68
3.4.3 O Método FBLP para Estimar Frequências.....	70
3.4.4 Limitações do Método FBLP.....	77
3.5 Conclusões.....	79

3.1 INTRODUÇÃO

O objetivo deste capítulo é aplicar a teoria de predição linear desenvolvida no capítulo 2, para estimar as frequências de sinais senoidais contaminados por ruído. Os coeficientes do filtro de erro de predição projetado com base nas amostras conhecidas do sinal, serão utilizados para estimar estas frequências. O método de predição linear a ser aplicado será o FBLP, originando o *Método FBLP para Estimação de Frequências*.

Dado que uma senóide pode ser representada pela soma de duas exponenciais complexas, o sinal aqui será composto pela soma de M exponenciais complexas contaminadas por ruído branco. Também, suporemos que existem N amostras deste sinal disponíveis para a tarefa de estimação das frequências.

Inicialmente analisaremos o caso sem ruído. Demonstraremos que neste caso os zeros da função de transferência do preditor têm propriedades poderosas que permitem detectar as frequências de forma precisa. Estas propriedades não mais se verificam quando o ruído está presente, ocorrendo uma degradação significativa do desempenho do método.

Analisaremos a matriz de correlação estimada $\hat{\mathbf{R}}$ em função de seus autovalores e autovetores possibilitando a subdivisão dos mesmos em dois subespaços: o de sinal e o de ruído. Esta análise permitirá mostrar que os elementos que mais contribuem para a degradação do método FBLP são os autovalores-autovetores do subespaço de ruído.

Serão tecidas ainda considerações envolvendo a ordem do preditor e o compromisso entre resolução e estabilidade do método. Simulações ilustrativas em computador complementarão o capítulo.

3.2 OS ZEROS DO FILTRO DE ERRO DE PREDIÇÃO - CASO SEM RUÍDO

3.2.1 OS ZEROS DE SINAL

Consideremos N amostras de um sinal, composto apenas pela soma de M exponenciais complexas:

$$u(n) = \sum_{k=1}^M \{a_k \cdot \exp[n \cdot s_k]\} \quad ; n = 1, 2, \dots, N \quad (3.1)$$

onde: a_k e s_k são números complexos desconhecidos, tais que:

$$a_k = \alpha_k \cdot \exp[j \cdot \phi_k]$$

$$s_k = \beta_k + j \cdot \omega_k$$

α_k → amplitude das exponenciais

ϕ_k → fase das exponenciais

$\beta_k \rightarrow$ fator de amortecimento das exponenciais

$\omega_k \rightarrow$ frequência das exponenciais; $\omega_k = 2\pi f_k$, onde f_k é a frequência normalizada em relação à frequência de amostragem

$a_k \neq 0, k = 1, 2, \dots, M$ e os $s_k, k = 1, 2, \dots, M$ são distintos entre si.

Projetamos, então, um preditor linear forward com ordem L , descrito na secção 2.3, para o sinal da eq. (3.1). A expressão para os erros correspondentes, dada por (2.8), é transcrita por conveniência:

$$\mathcal{F}^H = \mathbf{b}_f^H - \mathbf{w}_f^H \mathbf{A}_f^H \tag{3.2}$$

onde as expressões (2.3), (2.7) e (2.11) apresentam \mathbf{w}_f , \mathbf{b}_f e \mathbf{A}_f , respectivamente, nas suas formas expandidas.

Consideremos agora um polinômio $W_f(z)$, cujos coeficientes são as componentes de \mathbf{w}_f dispostos na seguinte forma:

$$W_f(z) = 1 + \sum_{k=1}^L (-w_{fk}) \cdot z^{-k}$$

TEOREMA 3.1: Se L satisfaz a desigualdade $M \leq L \leq (N - M)$, e se o vetor de coeficientes forward \mathbf{w}_f satisfaz $\mathbf{A}_f \mathbf{w}_f = \mathbf{b}_f$, então $W_f(z)$ tem M de seus L zeros em $\exp[s_k], k = 1, 2, \dots, M$.

DEMONSTRAÇÃO: Considerando os seguintes M vetores de dimensões $1 \times L$:

$$\mathbf{f}_k = [\exp[-s_k] \quad \exp[-2s_k] \quad \dots \quad \exp[-Ls_k]] ; k = 1, 2, \dots, M$$

a l ésima linha de \mathbf{A}_f , denotada por \mathbf{l}_l , pode ser escrita através da seguinte combinação linear dos \mathbf{f}_k 's:

$$\mathbf{l}_l = \sum_{k=1}^M a_k \cdot \exp[s_k \cdot (L+1)] \cdot \mathbf{f}_k$$

Supondo que os $s_k, k = 1, 2, \dots, M$, são distintos e $L \geq M$, então os $\mathbf{f}_k, k = 1, 2, \dots, M$, são todos linearmente independentes e os M vetores \mathbf{f}_k constituem uma base para as linhas de \mathbf{A}_f . Portanto, o posto de \mathbf{A}_f será no máximo M .

Vamos mostrar agora que, sob certas condições, o posto de \mathbf{A}_f é exatamente M . Sendo \mathbf{A}_f de dimensões $(N - L) \times L$, para que o seu posto seja M é necessário que:

i) $(N - L) \geq M$

ii) $L \geq M$

Portanto, tomaremos estas condições necessárias como hipóteses para demonstrar que o posto de A_f é M . Tomando como base a expressão acima para as linhas de A_f , podemos escrever esta matriz da seguinte forma:

$$A_f = \begin{bmatrix} a_1^* \cdot \exp[s_1^* \cdot (L+1)] & a_2^* \cdot \exp[s_2^* \cdot (L+1)] & \dots & a_M^* \cdot \exp[s_M^* \cdot (L+1)] \\ a_1^* \cdot \exp[s_1^* \cdot (L+2)] & a_2^* \cdot \exp[s_2^* \cdot (L+2)] & \dots & a_M^* \cdot \exp[s_M^* \cdot (L+2)] \\ \vdots & \vdots & \dots & \vdots \\ a_1^* \cdot \exp[s_1^* \cdot (N)] & a_2^* \cdot \exp[s_2^* \cdot (N)] & \dots & a_M^* \cdot \exp[s_M^* \cdot (N)] \end{bmatrix} \cdot \begin{bmatrix} f_1^* \\ f_2^* \\ \vdots \\ f_M^* \end{bmatrix} \tag{3.3}$$

Definimos, neste momento, os seguintes M vetores, de dimensões $1 \times (N - L)$:

$$e_i = [\exp[s_i \cdot (L+1)] \quad \exp[s_i \cdot (L+2)] \quad \dots \quad \exp[s_i \cdot (N)]]$$

$i = 1, 2, \dots, M$

Portanto (3.3) torna-se:

$$A_f = [a_1^* \cdot e_1^H \quad a_2^* \cdot e_2^H \quad \dots \quad a_M^* \cdot e_M^H] \cdot \begin{bmatrix} f_1^* \\ f_2^* \\ \vdots \\ f_M^* \end{bmatrix}$$

$$A_f = \sum_{i=1}^M a_i^* \cdot e_i^H \cdot f_i^* \tag{3.4}$$

$$A_f = E^H \cdot A \cdot F^* \tag{3.5}$$

onde: $E^H = [e_1^H \quad e_2^H \quad \dots \quad e_M^H]$ - dimensões $(N-L) \times M$

$A = \text{diag} [a_1 \quad a_2 \quad \dots \quad a_M]$ - dimensões $M \times M$

$F = [f_1 \quad f_2 \quad \dots \quad f_M]^t$ - dimensões $M \times L$

Tomamos, agora, o seguinte produto:

$$A_f \cdot x = y \tag{3.6}$$

onde: x é qualquer vetor $L \times 1$
 y é qualquer vetor $(N-L) \times 1$

Substituindo (3.5) em (3.6):

$$E^H \cdot A \cdot F^* \cdot x = y \tag{3.7}$$

Definimos o vetor $c_{(M \times 1)}$ como:

$$c = A \cdot F^* \cdot x \tag{3.8}$$

Portanto, de (3.7) e (3.8):

$$E^H \cdot c = y$$

$$y = \sum_{i=1}^M c_i \cdot e_i^H \tag{3.9}$$

O posto de A_f é igual ao número de linhas ou colunas linearmente independentes. Observando a eq. (3.6) vemos que y é uma combinação linear das colunas de A_f com fatores de ponderação definidos por x . Portanto, o número de vetores linearmente independentes que somados resultam y , fornece o posto de A_f .

Supondo, pela condição necessária i), que $(N - L) \geq M$, os $e_i^H, i = 1, 2, \dots, M$ são todos linearmente independentes. De posse desta informação e observando (3.9) deduzimos que o posto de A_f será, de fato, no máximo igual a M , pois y é composto pela soma de M vetores linearmente independentes (e_i^H), ponderados pelos respectivos c_i 's. Logo, o posto será M se $c_i \neq 0, i = 1, 2, \dots, M$. Observando a eq. (3.8) vemos que, se $A \cdot F^*$ tiver posto cheio igual a M , c será uma combinação linear de M vetores linearmente independentes, e como c tem dimensões $(M \times 1)$, então $c_i \neq 0, i = 1, 2, \dots, M$. Calculemos, pois, o posto de $A \cdot F^*$:

$$A \cdot F^* = \begin{bmatrix} a_1 & 0 & 0 & \dots & 0 \\ 0 & a_2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & \dots & \dots & 0 & a_M \end{bmatrix} \cdot \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_M \end{bmatrix} = \begin{bmatrix} a_1 \cdot f_1 \\ a_2 \cdot f_2 \\ \vdots \\ a_M \cdot f_M \end{bmatrix} \tag{3.10}$$

Já que, pela condição necessária ii), $L \geq M$, então $f_i, i = 1, 2, \dots, M$, são linearmente

independentes. Sendo $a_i \neq 0, i = 1, 2, \dots, M$, de (3.10) concluímos que o posto de $A \cdot F$ é M . Portanto:

$$\text{posto}(A \cdot F) = M \Rightarrow c_i \neq 0, i = 1, 2, \dots, M \Rightarrow \text{posto}(A_f) = M \quad (3.11)$$

Fica mostrado, portanto, que as condições necessárias para que o posto de A_f seja M são também suficientes. Combinando as condições necessárias e suficientes i) e ii) teremos:

$$M \leq L \leq (N - M) \quad (3.12)$$

Temos como hipótese inicial deste teorema que $A_f \cdot w_f = b_f$. Para que isto ocorra, o vetor erro de predição \mathcal{F} em (3.2) deve ser nulo, o que significa que uma determinada amostra de $u(n)$ é totalmente predizível através de L amostras anteriores. Substituindo $\mathcal{F} = 0$ e tomando o hermitiano em ambos os lados de (3.2), obtemos:

$$A_f \cdot w_f = b_f \quad (3.13)$$

Veremos, a seguir, que condições garantem que o sistema de equações de (3.13) tenha solução. Da teoria de soluções de equações temos que [3]:

- a) Se o posto de $[b_f \ ; \ A_f]$ é maior que o posto de A_f , o sistema de equações de (3.13) é inconsistente e, por isso, não tem solução.
- b) Se o posto de $[b_f \ ; \ A_f]$ é igual ao posto de A_f , e este é igual ao número de incógnitas L , o sistema de (3.13) tem solução única.
- c) Se o posto de $[b_f \ ; \ A_f]$ é igual ao posto de A_f , e este é menor que o número de incógnitas L , o sistema de (3.13) tem infinitas soluções.

As linhas de $[b_f \ ; \ A_f]$ podem ser escritas como uma combinação linear dos seguintes M vetores de dimensões $1 \times (L+1)$:

$$h_k = [1 \ \exp[-s_k] \ \exp[-2.s_k] \ \dots \ \exp[-L.s_k]] \quad ; k = 1, 2, \dots, M$$

supondo que os $s_k, k = 1, 2, \dots, M$, são distintos e $L \geq M$, os $h_k, k = 1, 2, \dots, M$, são todos linearmente independentes e, de forma análoga a A_f , o posto de $[b_f \ ; \ A_f]$

será no máximo M . Portanto, para garantir que o sistema de (3.13) tenha solução, basta que $\text{posto}[A_f]$ seja M , pois desta forma, já que $\text{posto}[b_f \mid A_f]$ é no máximo M :

$$\text{posto}[b_f \mid A_f] = \text{posto}[A_f] = M$$

Para que isto ocorra, é necessário e suficiente que a desigualdade em (3.12) seja satisfeita. A solução só será única se $L = M$, caso contrário existirão infinitas soluções.

Até aqui demonstramos que se L satisfaz a desigualdade de (3.12), o posto de A_f é M e é possível obter w_f tal que $A_f \cdot w_f = b_f$, ou seja, o erro mínimo de predição será sempre nulo. A seguir mostraremos que se estas condições se verificam, o polinômio $W_f(z)$ tem M de seus zeros em $\exp[s_k]$, $k = 1, 2, \dots, M$. Dando prosseguimento à demonstração, ao observarmos a expressão (2.7) para b_f , podemos escrever:

$$b_f = \sum_{k=1}^M a_k^* \cdot e_k^H \tag{3.14}$$

Substituindo (3.5) e (3.14) em (3.13), obtemos:

$$\begin{aligned} E^H \cdot A^* \cdot F^* \cdot w_f - \sum_{k=1}^M a_k^* \cdot e_k^H &= 0 \\ E^H \cdot d - \sum_{k=1}^M a_k^* \cdot e_k^H &= 0 \\ \sum_{k=1}^M d_k \cdot e_k^H - \sum_{k=1}^M a_k^* \cdot e_k^H &= 0 \\ \sum_{k=1}^M (d_k - a_k^*) \cdot e_k^H &= 0 \end{aligned} \tag{3.15}$$

onde: $A^* \cdot F^* \cdot w_f = d$

Como, por hipótese $(N - L) \geq M$, os e_k , $k = 1, 2, \dots, M$, são todos linearmente independentes, e um vetor não pode ser escrito como uma combinação linear dos demais. Portanto, (3.15) só possuirá a solução trivial:

$$d_k - a_k^* = 0 \quad k = 1, 2, \dots, M$$

Então, podemos escrever:

$$A_f \cdot F \cdot w_f - \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix} = 0$$

$$\begin{bmatrix} a_1 \cdot f_1 \cdot w_f \\ a_2 \cdot f_2 \cdot w_f \\ \vdots \\ a_M \cdot f_M \cdot w_f \end{bmatrix} - \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix} = 0$$

$$a_k \cdot f_k \cdot w_f - a_k = 0 \quad k = 1, 2, \dots, M$$

$$1 - f_k \cdot w_f = 0 \quad k = 1, 2, \dots, M$$

pois, por hipótese, $a_k \neq 0, k = 1, 2, \dots, M$.

$$[1 \quad \exp[-s_k] \quad \dots \quad \exp[-L \cdot s_k]] \cdot \begin{bmatrix} 1 \\ -w_{f1} \\ \vdots \\ -w_{fL} \end{bmatrix} = 0 \quad ; k = 1, 2, \dots, M$$

$$1 + (-w_{f1}) \cdot \exp[-s_k] + \dots + (-w_{fL}) \cdot \exp[-L \cdot s_k] = 0 \quad ; k = 1, 2, \dots, M$$

$$W_f(z) = 1 + \sum_{i=1}^L (-w_{fi}) \cdot z^{-i} \Big|_{z=\exp[s_k]} = 0$$

$$k = 1, 2, \dots, M \tag{3.16}$$

Portanto $z = \exp[s_k], k = 1, 2, \dots, M$, são raízes de $W_f(z)$ e o teorema 3.1 está demonstrado.

Para que (3.16) seja sempre verdadeira, o posto de A_f deve ser M , e para que isto ocorra (3.12) deve se verificar. Se $M < L \leq (N - M)$ então existem infinitos w_f tais que $A_f \cdot w_f = b_f$. Porém, para todos eles vale o fato de que $\exp[s_k], k = 1, 2, \dots, M$, são raízes de $W_f(z)$. Isto porque a ordem de $W_f(z)$ é $L > M$ e a distinção entre as infinitas soluções se dá através dos demais $(L - M)$ zeros, não tratados aqui.

Estudaremos agora o mesmo problema anterior, porém utilizando a predição linear backward, descrita na secção 2.4, para o sinal da eq. (3.1). A expressão para os erros de predição backward é dada por (2.32), transcrita por conveniência:

$$g^H = b_b^H - w_b^H \cdot A_b^H \tag{3.17}$$

onde as expressões (2.27), (2.31) e (2.35) apresentam w_b , b_b e A_b , respectivamente, nas suas formas expandidas.

Consideremos agora um polinômio $W_b(z)$, cujos coeficientes são as componentes de w_b , dispostos na seguinte forma:

$$W_b(z) = 1 + \sum_{k=1}^L (-w_{b(L-k)}) \cdot z^{-k}$$

TEOREMA 3.2: Se L satisfaz a desigualdade $M \leq L \leq (N - M)$, e se o vetor de coeficientes backward w_b satisfaz $A_b \cdot w_b = b_b$, então $W_b(z)$ tem M de seus L zeros em $\exp[-s_k^*], k = 1, 2, \dots, M$.

DEMONSTRAÇÃO: Considerando os seguintes M vetores de dimensões $1 \times L$:

$$f_k = [\exp[s_k^*] \quad \exp[2 \cdot s_k^*] \quad \dots \quad \exp[L \cdot s_k^*]] ; k = 1, 2, \dots, M$$

A i ésima linha de A_b , denotada por l_i , pode ser escrita através da seguinte combinação linear dos f_k 's:

$$l_i = \sum_{k=1}^M a_k \cdot \exp[s_k \cdot i] \cdot f_k^*$$

Supondo que os $s_k, k = 1, 2, \dots, M$, são distintos e $L \geq M$, então os $f_k, k = 1, 2, \dots, M$, são todos linearmente independentes e os M vetores f_k constituem uma base para as linhas de A_b . Portanto, o posto de A_b será no máximo M .

Podemos escrever A_b na seguinte forma:

$$A_b = \begin{bmatrix} a_1 \cdot \exp[s_1] & a_2 \cdot \exp[s_2] & \dots & a_M \cdot \exp[s_M] \\ a_1 \cdot \exp[2 \cdot s_1] & a_2 \cdot \exp[2 \cdot s_2] & \dots & a_M \cdot \exp[2 \cdot s_M] \\ \vdots & \vdots & \dots & \vdots \\ a_1 \cdot \exp[(N-L) \cdot s_1] & a_2 \cdot \exp[(N-L) \cdot s_2] & \dots & a_M \cdot \exp[(N-L) \cdot s_M] \end{bmatrix} \cdot \begin{bmatrix} f_1^* \\ f_2^* \\ \vdots \\ f_M^* \end{bmatrix} \tag{3.18}$$

Definimos os seguintes M vetores, de dimensões 1x(N - L):

$$e_i = [\exp[s_1^*] \quad \exp[2.s_1^*] \quad \dots \quad \exp[(N-L).s_1^*]]$$

$$i = 1, 2, \dots, M$$

Portanto (3.18) torna-se:

$$A_b = [a_1.e_1^H \quad a_2.e_2^H \quad \dots \quad a_M.e_M^H] \cdot \begin{bmatrix} f_1^* \\ f_2^* \\ \vdots \\ f_M^* \end{bmatrix}$$

$$A_b = \sum_{i=1}^M a_i.e_i^H.f_i^* \tag{3.19}$$

$$A_b = E^H . A^* . F^* \tag{3.20}$$

onde: $E^H = [e_1^H \quad e_2^H \quad \dots \quad e_M^H]$ - dimensões (N-L)xM

$A = \text{diag} [a_1^* \quad a_2^* \quad \dots \quad a_M^*]$ - dimensões MxM

$F = [f_1 \quad f_2 \quad \dots \quad f_M]^t$ - dimensões Mx(L)

As equações (3.20) e (3.5) são idênticas na forma, embora as matrizes envolvidas sejam distintas. Como consequência, o procedimento é idêntico ao do teorema 3.1. Caso $M \leq L \leq (N - M)$, A_b tem posto M, e prosseguindo a demonstração chegamos a:

$$E^H . A^* . F^* . w_b - \sum_{k=1}^M a_k . e_k^H = 0$$

$$\sum_{k=1}^M (d_k - a_k) . e_k^H = 0 \tag{3.21}$$

onde: $A^* . F^* . w_b = d$

Como, por hipótese $(N - L) \geq M$, os $e_k, k = 1, 2, \dots, M$, são todos linearmente independentes, e um vetor não pode ser escrito como uma combinação linear dos demais. Portanto, (3.21) só possuirá a solução trivial:

$$d_k - a_k = 0 \quad k = 1, 2, \dots, M$$

Então, podemos escrever:

$$\begin{bmatrix} a_1 \cdot f_1 \cdot w_b \\ a_2 \cdot f_2 \cdot w_b \\ \vdots \\ a_M \cdot f_M \cdot w_b \end{bmatrix} - \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix} = 0$$

$$1 - f_k \cdot w_b = 0 \quad k = 1, 2, \dots, M$$

$$\begin{bmatrix} 1 & \exp[s_k] & \dots & \exp[L \cdot s_k] \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -w_{b(L-1)} \\ \vdots \\ -w_{b0} \end{bmatrix} = 0 \quad ; k = 1, 2, \dots, M$$

$$1 + (-w_{b(L-1)}) \cdot \exp[s_k] + \dots + (-w_{b0}) \cdot \exp[L \cdot s_k] = 0 \quad k = 1, 2, \dots, M$$

$$W_b(z) = 1 + \sum_{i=1}^L (-w_{b(L-i)}) \cdot z^{-i} \Big|_{z=\exp[-s_k]} = 0 \quad k = 1, 2, \dots, M \quad (3.22)$$

Portanto $z = \exp[-s_k], k = 1, 2, \dots, M$, são raízes de $W_b(z)$. De forma análoga ao caso forward, para que (3.22) seja sempre verdadeira é necessário e suficiente que:

$$M \leq L \leq (N - M)$$

As demais observações são válidas aqui também e o teorema 3.2 está demonstrado.

Os M zeros, $z = \exp[s_k], k = 1, 2, \dots, M$, para $W_f(z)$, e $z = \exp[-s_k], k = 1, 2, \dots, M$, para $W_b(z)$ são definidos como *Zeros de Sinal* destes polinômios, denominação esta motivada pelo fato de que através deles obtemos os parâmetros desconhecidos $s_k, k = 1, 2, \dots, M$, para as exponenciais complexas da eq. (3.1). Tais zeros de $W_f(z)$ e de $W_b(z)$ guardam entre si a relação de recíprocos

conjugados com respeito à circunferência de raio unitário do plano z .

Se a sequência $u(n)$ for composta por exponenciais não amortecidas, isto é, s_k são imaginários puros, então os M zeros de sinal de $W_f(z)$ e $W_b(z)$ serão iguais, pois $\exp[s_k] = \exp[-s_k^*]$. Portanto, se $W_f(z)$ e $W_b(z)$ possuem os mesmos zeros de sinal, podemos combinar os teoremas 3.1 e 3.2 de modo a tratar o caso forward-backward da secção 2.5. Para isto, será necessário considerar as soluções particulares onde $w_f = w_b$.

TEOREMA 3.3: Seja $s_k, k = 1, 2, \dots, M$, imaginários puros, A uma matriz $2.(N - L) \times L$ e b um vetor $2.(N - L) \times 1$ nas formas:

$$A = \begin{bmatrix} A_f \\ A_b \end{bmatrix}$$

$$b = \begin{bmatrix} b_f \\ b_b \end{bmatrix}$$

e: $w_f = w_b = w$

Se L satisfaz a desigualdade $M \leq L \leq (N - M/2)$, e se o vetor de coeficientes forward-backward w , satisfaz $A.w = b$, então o polinômio $W(z) = 1 + \sum_{i=1}^L (-w_i^*) z^{-i}$, tem M de seus zeros na circunferência de raio unitário em $\exp[s_k] = \exp[-s_k^*], k = 1, 2, \dots, M$. Caso $(M/2)$ não seja inteiro, este valor é arredondado para o maior inteiro mais próximo.

DEMONSTRAÇÃO: A demonstração segue os mesmos passos das anteriores, porém agora a matriz A é $2.(N - L) \times L$, e analogamente:

$$2.(N - L) \geq M$$

$$L \leq (N - M/2)$$

portanto:

$$M \leq L \leq (N - M/2) \tag{3.23}$$

3.2.2 OS ZEROS ESTRANHOS

O polinômio $W(z)(W_f(z)$ ou $W_b(z))$, do sub-ítem acima, possui L zeros. M deles foram caracterizados como os *zeros de sinal*, ao passo que os outros $(L - M)$ zeros são aqui denominados de *zeros estranhos* do polinômio $W(z)(W_f(z)$ ou $W_b(z))$. Como visto anteriormente, a equação $A.w = b(A_f.w_f = b_f$ ou $A_b.w_b = b_b)$ pode apresentar infinitas soluções para $w(w_f$ ou $w_b)$ e conseqüentemente para $W(z)(W_f(z)$ ou $W_b(z))$, bastando para isto que $M < L \leq (N - M/2)(M < L \leq (N - M))$. Pretendemos utilizar os zeros de sinal de $W(z)(W_f(z)$ ou $W_b(z))$ para estimar os parâmetros s_k , e assim sendo, nem todas as soluções de $W(z)(W_f(z)$ ou $W_b(z))$ nos interessam, mas apenas aquelas que permitam distinguir os zeros de sinal dos zeros estranhos. Isto porque embora os zeros de sinal possuam posições definidas, nada podemos afirmar ainda em relação à localização dos zeros estranhos no plano z . No próximo teorema será especificado o vetor $w(w_f$ ou $w_b)$ que possui as características desejadas que permitem a distinção entre os zeros.

No restante deste capítulo, trataremos apenas do caso forward-backward, embora os resultados possam ser aplicados aos casos separados de forma imediata.

TEOREMA 3.4: Os $(L - M)$ zeros estranhos do polinômio $W(z) = 1 + \sum_{i=1}^L (-w_i^*) . z^{-i}$ localizam-se no interior do círculo de raio unitário, quando os M zeros de sinal ocupam posições quaisquer porém fixas e o vetor de coeficientes w é escolhido tal que: $1 + |w_1|^2 + |w_2|^2 + \dots + |w_L|^2$ é um mínimo.

DEMONSTRAÇÃO: O polinômio $W(z)$ pode ser fatorado da seguinte forma:

$$W(z) = W_1(z) . W_2(z) \tag{3.24}$$

onde: $W_1(z) = 1 + b_1 . z^{-1} + b_2 . z^{-2} + \dots + b_M . z^{-M}$

$$W_2(z) = 1 + c_1 . z^{-1} + c_2 . z^{-2} + \dots + c_{(L-M)} . z^{-(L-M)}$$

O polinômio $W_1(z)$ é tal que as suas raízes representam os M zeros de sinal de $W(z)$. O interesse aqui recai nos zeros de $W_2(z)$, que são os $(L - M)$ zeros estranhos, quando $w^H . w$ é o mínimo possível sob a restrição que $W_1(z)$ possua os zeros de sinal arbitrários porém fixos. Desde que o polinômio $W(z)$ é o produto de $W_1(z)$ e $W_2(z)$, os seus coeficientes são relacionados pela operação de convolução:

$$w_n = \sum_{k=-\infty}^{\infty} c_k \cdot b_{n-k}$$

onde: $w_0 = b_0 = c_0 = 1$
 $b_i = 0 ; i > M \text{ e } i < 0$
 $c_i = 0 ; i > (L - M) \text{ e } i < 0$

Podemos imaginar que a convolução acima é a descrição da atuação de um filtro de erro de predição linear, onde c_i são os coeficientes deste filtro, b_i é a sequência de entrada e w_n é a sequência de erro de predição que desejamos minimizar. Portanto, minimizar $w^H \cdot w$ sujeito a $W_1(z)$ fixo, torna-se equivalente a um problema de predição linear e os limites empregados na operação de convolução acima implicam que a escolha dos coeficientes c_k que minimizam w_n , será feita segundo o método da autocorrelação [1]. A equação normal correspondente é do tipo:

$$R \cdot c = r \tag{3.25}$$

onde R e r são compostos por elementos do tipo:

$$r(i) = \frac{1}{M + 1} \cdot \sum_{k=0}^M b_k^* \cdot b_{k+i} \tag{3.26}$$

e R é Toeplitz e Hermitiana. c é o vetor de coeficientes:

$$c^t = [1 \ c_1 \ c_2 \ \dots \ c_{(L-M)}]$$

A resolução de (3.25) minimiza $w^H \cdot w$ sujeito a $W_1(z)$ fixo e define o polinômio $W_2(z)$. É sabido que tal filtro de erro de predição é de fase mínima [1 - pg. 150] e, portanto, possui os seus zeros no interior do círculo de raio unitário. Com isto o teorema 3.4 está demonstrado.

Os teoremas anteriores permitem concluir que se o sinal é composto por exponenciais não-amortecidas, se L satisfaz a desigualdade da eq. (3.23) e se $A \cdot w = b$, existe um $W(z)$ particular onde os M zeros de sinal localizam-se na circunferência de raio unitário em $z = \exp[s_k] = \exp[-s_k^*], k = 1, 2, \dots, M$. Além disto, os $(L - M)$ zeros estranhos localizam-se no interior do círculo de raio unitário, tornando possível distinguir os zeros de sinal dos zeros estranhos.

A expressão $A \cdot w = b$ pode ser pré-multiplicada, em ambos os lados, por A^H resultando em:

$$A^H \cdot A \cdot w = A^H \cdot b \tag{3.27}$$

onde: A e b são dados por (2.54) e (2.55)

A eq. (3.27) já é bem conhecida: é a equação normal determinística. Admitindo que $M \leq L \leq (N - M/2)$, minimizar $w^H \cdot w$ sujeito a $A \cdot w = b$ é equivalente a tornar $w^H \cdot w$ um mínimo sujeito à (3.27). Isto porque foi demonstrado que sempre existe w que torna o erro mínimo nulo. Logo, quando o erro mínimo é nulo, toda solução de (3.27) é também solução de $A \cdot w = b$ e vice-versa. Mas a solução de (3.27) que minimiza $w^H \cdot w$ é a solução de norma mínima para a equação normal determinística, a qual é dada por:

$$w = A^{\#} \cdot b$$

onde: $A^{\#}$ é a pseudo-inversa da matriz A

Fica assim explicado o motivo pelo qual se dispensou atenção especial a tal solução no capítulo anterior.

Supondo o caso de exponenciais amortecidas, sabemos, pelo teorema 3.1, que os M zeros de sinal de $W_f(z)$ estão em $z = \exp[s_k]$, $k = 1, 2, \dots, M$; e pelo teorema 3.2, que os M zeros de sinal de $W_b(z)$ estão em $z = \exp[-s_k^*]$, $k = 1, 2, \dots, M$. Porém, a princípio, nenhuma ligação existe entre os zeros estranhos destes polinômios. Sabemos apenas que, no caso da solução de norma mínima, eles se encontram no interior do círculo de raio unitário. No entanto, o próximo teorema estabelece uma ligação entre os zeros estranhos de $W_f(z)$ e $W_b(z)$.

TEOREMA 3.5: Seja w_f e w_b satisfazendo as restrições dos teoremas 3.1, 3.2 e 3.4. Então, os $(L - M)$ zeros estranhos dos polinômios $W_f(z)$ e $W_b(z)$ são idênticos.

DEMONSTRAÇÃO: Reescrevendo a equação (3.24) para os casos do teorema 3.1 e do teorema 3.2, obtemos:

$$W_f(z) = W_{f1}(z) \cdot W_{f2}(z)$$

$$W_b(z) = W_{b1}(z) \cdot W_{b2}(z)$$

Sabemos que os M zeros de sinal de $W_{f1}(z)$ são as reflexões dos M zeros de $W_{b1}(z)$ em torno da circunferência de raio unitário em direção às suas posições recíprocas. Supondo que os coeficientes de $W_{f1}(z)$ são $[1 \ b_{f1} \ b_{f2} \ \dots \ b_{fM}]$, devido a esta propriedade especial dos zeros de sinal, os coeficientes de $W_{b1}(z)$ são os coeficientes de $W_{f1}(z)$ escritos na forma reversa, complexos conjugados e divididos

por b_{fM}^* , ou seja, os coeficientes de $W_{b1}(z)$ são $[b_{fM}^* b_{f(M-1)}^* \dots 1]/b_{fM}^*$ [11].

Os $(L - M)$ zeros estranhos dependem do vetor c solução de (3.25). Este, por sua vez, depende dos $r(i)$, dados por (3.26), quando substituirmos b_k por b_{fk} (ou b_{bk}) e b_{k+1} por $b_{f(K+i)}$ (ou $b_{b(K+i)}$). Os valores $r(i)$ não são afetados quando escrevemos os b_{fk} na sua ordem inversa e conjugada e, ao aplicarmos o fator de escala $1/b_{fM}^*$, este aparecerá em ambos os lados de (3.25). Portanto, o valor de c é o mesmo para os casos forward e backward.

Concluimos que, tomando as soluções de norma mínima, os $(L - M)$ zeros estranhos de $W_f(z)$ e $W_b(z)$ são idênticos, enquanto que os zeros de sinal se situam em posições recíprocas em relação à circunferência de raio unitário.

Espalhamento dos Zeros Estranhos

Caso w tenha norma mínima, demonstramos no teorema 3.4, que os $(L - M)$ zeros estranhos de $W(z)$ localizam-se no interior do círculo de raio unitário. Daremos, agora, uma explicação qualitativa para mostrar que estes zeros estão distribuídos de forma *aproximadamente uniforme* no interior do círculo de raio unitário.

Para isto, vamos retomar a imagem de um filtro de erro de predição com coeficientes c_n atuando sobre o sinal fictício cujas amostras são os coeficientes b_n , ilustrada na fig. 3.1.

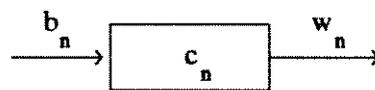


Fig. 3.1 - Geração de w_n

A otimização dos coeficientes do filtro de erro de predição faz com que ocorra uma tendência de "branqueamento" do sinal na sua saída [1], isto é, a característica espectral do sinal tende a ser constante com a frequência. Portanto, haverá uma tendência de alocação de seus zeros em torno da circunferência de raio unitário e próximos a ela, nas regiões onde o espectro de b_n possui concentração de energia, isto é, nas regiões onde o espectro de b_n não possui zeros.

Os únicos zeros do espectro de b_n são os zeros de sinal. Logo, os zeros estranhos tenderão a se distribuir de forma uniforme no interior da c.r.u. nas regiões onde não existam zeros de sinal. Eles estarão tanto mais próximos da circunferência quanto maior o valor de $(L - M)$.

Simulações

Um sinal do tipo:

$$u(n) = \exp[s_1.n] + \exp[s_2.n] + \exp[s_3.n] \quad n = 1, 2, \dots, 25$$

onde:

$$s_1 = 0.1 + j.2.\pi.(0.35)$$

$$s_2 = j.2.\pi.(0.5)$$

$$s_3 = -0.3 + j.2.\pi.(0.65)$$

é utilizado para as simulações, objetivando a exemplificação da teoria desenvolvida até então. $\exp[s_2]$ tem módulo unitário, pois o seu fator de amortecimento é zero, localizando-se, assim, na c.r.u.. $\exp[s_1]$ localiza-se no exterior do círculo de raio unitário pois o seu fator de amortecimento é positivo, enquanto que $\exp[s_3]$ localiza-se no interior do círculo de raio unitário. Os ângulos para as posições de $\exp[s_1]$, $\exp[s_2]$ e $\exp[s_3]$ são, respectivamente, $0.7.\pi$, π e $1.3.\pi$, medidos no sentido anti-horário a partir da origem do plano z.

A fig. 3.2a mostra os zeros de sinal e os zeros estranhos para $W_f(z)$, quando $L = 20$. A fig. 3.2b mostra o mesmo para $W_b(z)$. Os zeros de sinal estão indicados por setas. Notamos que os zeros estranhos estão uniformemente distribuídos no interior do círculo de raio unitário, enquanto que os de sinal são refletidos em torno da c.r.u..

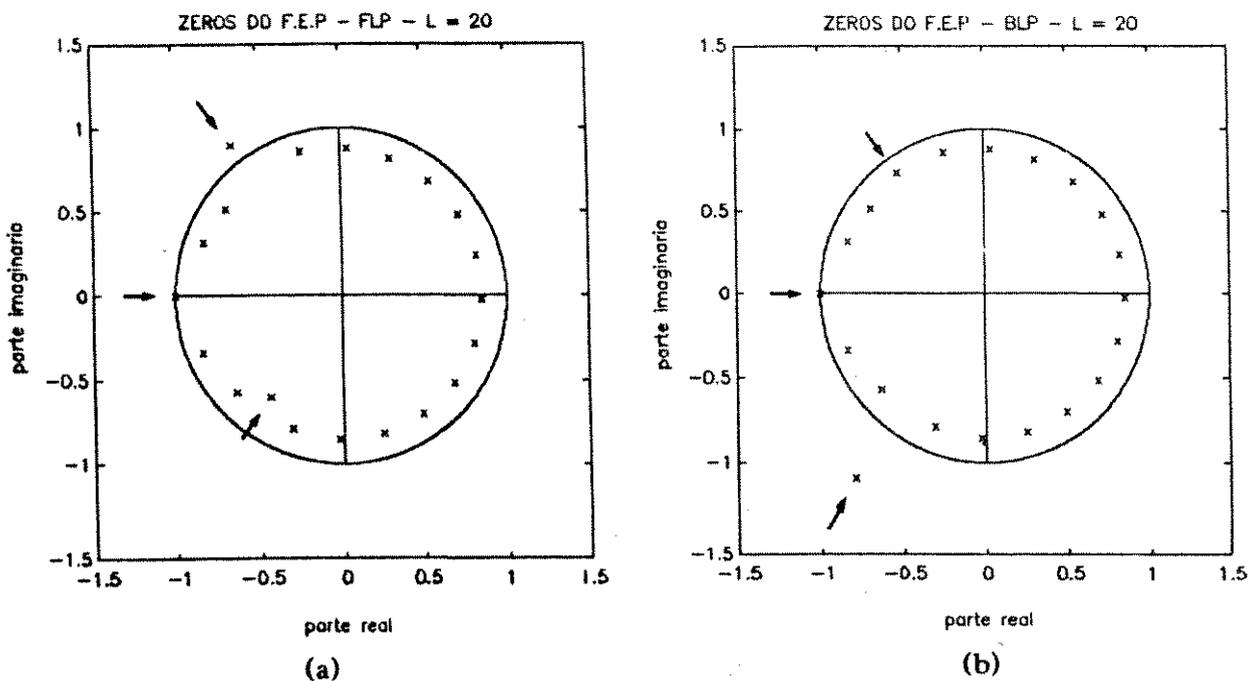


Fig. 3.2 - Zeros do Filtro de Erro de Predição - Caso sem ruído - $L = 20$ - (a) Forward (b) Backward

3.3 ESTIMAÇÃO DAS FREQUÊNCIAS DE EXPONENCIAIS NÃO-AMORTECIDAS - CASO SEM RUÍDO

3.3.1 INTRODUÇÃO

Suponhamos agora que os parâmetros s_k da eq. (3.1) sejam imaginários, ou seja, as exponenciais são não-amortecidas. Desejamos estimar as frequências $\omega_k, k = 1, 2, \dots, M$, das exponenciais com base nas N amostras de $u(n)$.

De acordo com o teorema 3.3, os M zeros de sinal do filtro de erro de predição forward-backward se localizarão nas posições $\exp[j\omega_k], k = 1, 2, \dots, M$, do plano z , ao passo que, de acordo com o teorema 3.4, os zeros estranhos estarão uniformemente distribuídos no interior da c.r.u.. Portanto, neste caso, o preditor forward-backward pode ser utilizado para a detecção das frequências. Embora os três métodos possam ser aplicados, o FBLP é o mais interessante, pois oferece uma gama maior de valores possíveis para L . Veremos mais adiante, com detalhes, que a ordem L do filtro está diretamente ligada à capacidade de resolver duas frequências muito próximas. Quanto maior o L , maior esta capacidade. Portanto, o método FBLP possui uma maior resolução para a detecção de frequências em relação aos outros dois métodos.

Além disto, sendo o método FBLP uma combinação dos métodos FLP e BLP, ele consegue extrair mais informação das N amostras disponíveis. Isto pode ser constatado ao observarmos as expressões para as matrizes de dados dos casos forward, backward e forward-backward, fornecidas por (2.11), (2.35) e (2.54), respectivamente. Sendo a matriz de dados forward-backward a união das matrizes forward e backward, cada elemento da matriz de correlação estimada $\hat{R} = (N - L) \cdot A^H \cdot A$ possui um número maior de produtos das amostras de $u(n)$ disponíveis. É certo que no caso forward-backward alguns destes produtos ocorrem de forma repetida, mas mesmo assim, o número de produtos inéditos é maior do que nos casos separados. Isto possibilita estimativas mais fiéis e mais estáveis para a matriz de correlação.

Por todos estes motivos, daqui por diante, o método a ser utilizado para detecção de frequências será o FBLP.

3.3.2 O MÉTODO FBLP PARA ESTIMAR FREQUÊNCIAS

O seguinte procedimento pode ser utilizado para estimar as frequências da eq. (3.1):

1) Formar a matriz de dados A e o vetor de respostas desejadas b , forward-backward, dados por (2.54) e (2.55).

2) Encontrar a solução de norma mínima, $w = A^* . b$, para a equação normal determinística, $A^H . A . w = A^H . b$. Aqui há duas situações a serem consideradas:

2.1) $2.(N - L) \geq L$, $k = M$

a) Calcular a SVD de A

b) Utilizar a expressão (2.123), com o posto de $A(k) = M$.

$$w = \sum_{i=1}^M \frac{v_i}{\sigma_i^2} \cdot v_i^H . A^H . b$$

2.2) $2.(N - L) \leq L$, $k = M$

a) Calcular a SVD de A

b) Utilizar a expressão (2.126), com o posto de $A(k) = M$

$$w = \sum_{i=1}^M \frac{u_i^H . b}{\sigma_i^2} \cdot A^H . u_i$$

Este procedimento é aquele descrito no final do capítulo 2, fazendo m , o número de linhas da matriz A , igual a $2.(N - L)$, e o número de colunas n , igual a L . Além disso, como mostrado, o posto de A é $k = M$. A solução só será única se $k = L = M$.

3) Encontrar os zeros do polinômio $W(z)$, dado por:

$$W(z) = 1 + \sum_{k=1}^L (-w_k^*) . z^{-k}$$

onde w_k é a k 'ésima componente de w obtido no passo 2.

4) Os M zeros que se localizarem na c.r.u. determinam $\omega_k, k = 1, 2, \dots, M$.

No sub-ítem 2, ao realizarmos a SVD da matriz A , encontraremos M valores singulares maiores que zero, e os restantes dos $\min[2.(N - L), L]$ valores singulares serão todos iguais a zero. Isto se deve ao fato de que $\text{posto}(A) = M$.

Simulações

Objetivando realizar simulações que ilustrem a aplicação do método acima, utilizaremos o seguinte conjunto de amostras de dados:

$$u(n) = \exp[j.(2.\Pi.f_1.n + \phi_1)] + \exp[j.(2.\Pi.f_2.n + \phi_2)]$$

$$; n = 1, 2, \dots, 25 \quad (3.28)$$

onde:

$$f_1 = 0.5 \quad \phi_1 = -\Pi$$

$$f_2 = 0.52 \quad \phi_2 = -(0.79).\Pi$$

Dado que estamos utilizando duas exponenciais e que dispomos de 25 amostras de $u(n)$, então $M = 2$ e $N = 25$. Além disto escolhemos $L = 18$. A tabela 3.1 mostra os valores singulares para A , encontrados pela rotina SVD. Constatamos que apenas dois valores singulares são diferentes de zero. Isto se deve ao fato do posto de A ser igual ao número de exponenciais $M = 2$. A fig. 3.3a mostra os zeros de $W(z)$ no plano z . As posições angulares corretas das frequências, que neste caso coincidem com os zeros de sinal, são indicadas por setas. A estimação das frequências torna-se imediata na medida que é óbvio distinguirmos os 2 zeros de sinal, que estão na c.r.u., dos $(L - M) = 16$ zeros estranhos, que encontram-se no interior do círculo de raio unitário. A fig. 3.3b mostra o espectro amostra do sinal, o qual é definido como:

$$E(\omega) = \frac{1}{|1 + \sum_{i=1}^L (-w_i^*) e^{-j.\omega}|^2} \quad (3.29)$$

Picos se pronunciarão em $E(\omega)$, quando $\omega = \omega_1$ ou $\omega = \omega_2$.

SNR = ∞	
vs[1]	= 14.400537
vs[2]	= 8.795547
vs[3]	= 0.000001
vs[4]	= 0.000000
vs[5]	= 0.000000
vs[6]	= 0.000000
vs[7]	= 0.000000
vs[8]	= 0.000000
vs[9]	= 0.000000
vs[10]	= 0.000000
vs[11]	= 0.000000
vs[12]	= 0.000000
vs[13]	= 0.000000
vs[14]	= 0.000000
vs[15]	= 0.000000
vs[16]	= 0.000000
vs[17]	= 0.000000
vs[18]	= 0.000000

Tab. 3.1 - Valores Singulares da Matriz de Dados -
Caso sem ruído - L = 18

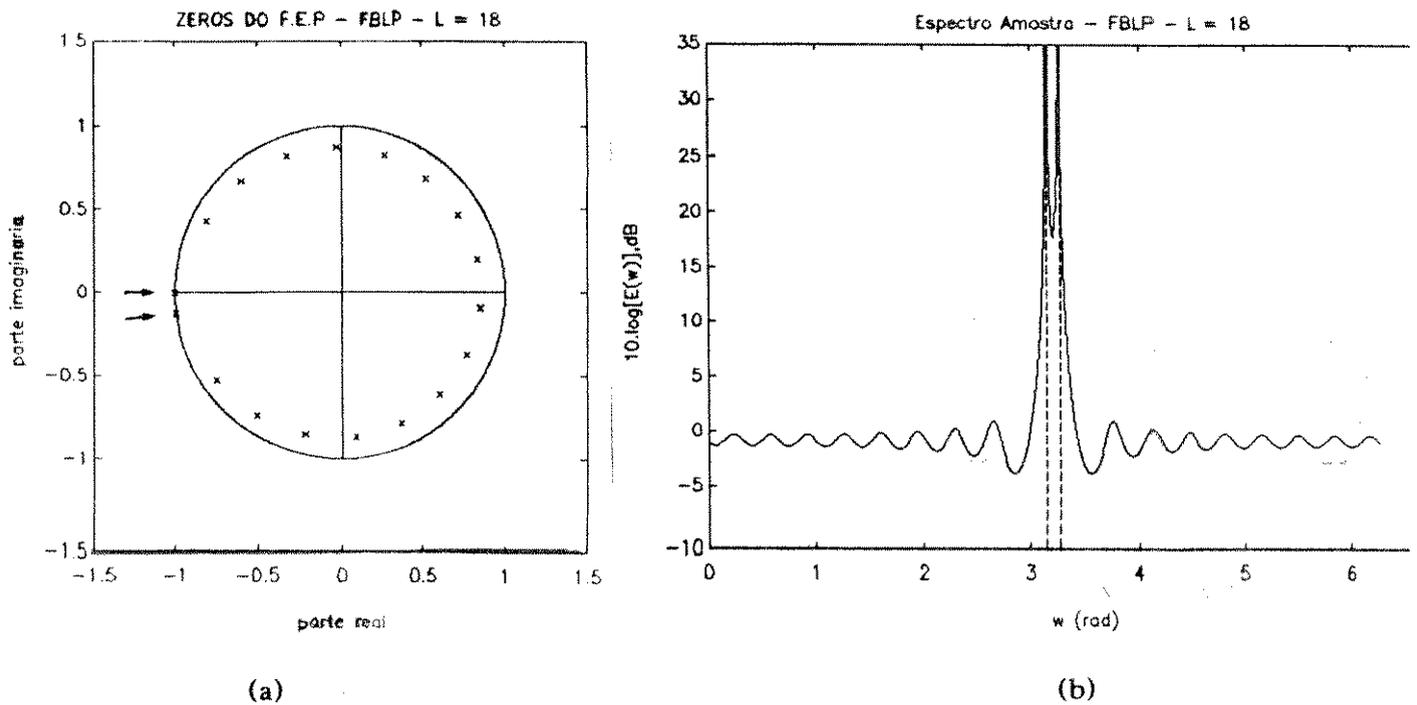


Fig. 3.3 - Método FBLP - Caso sem ruído - L = 18 -
 (a) Zeros do F.E.P. (b) Espectro Amostra

3.4 ESTIMAÇÃO DAS FREQUÊNCIAS DE EXPONENCIAIS NÃO-AMORTECIDAS - CASO COM RUÍDO

3.4.1 INTRODUÇÃO

Toda a teoria até aqui desenvolvida tratou de exponenciais complexas sem ruído. Agora consideraremos o sinal do tipo:

$$u(n) = \sum_{k=1}^M a_k \exp[j\omega_k \cdot n] + y(n) \quad n = 1, 2, \dots, N \quad (3.30)$$

onde: $y(n)$ é um ruído branco complexo com média zero e variância σ_y^2 , definida da seguinte forma:

$$\sigma_y^2 = \sigma_{yr}^2 + \sigma_{yi}^2$$

onde: $\sigma_{yr}^2 \rightarrow$ variância da parte real de $y(n)$
 $\sigma_{yi}^2 \rightarrow$ variância da parte imaginária de $y(n)$

Quando as amostras encontram-se contaminadas pelo ruído, todo o comportamento deduzido nas secções anteriores se modifica. A matriz de dados A não mais possui posto M , e a distribuição espacial dos zeros de $W(z)$ não é mais como antes. Isto deteriora em demasia o método FBLP, principalmente para relação sinal-ruído pequena, como será discutido a seguir.

3.4.2 O SUBESPAÇO DE SINAL E O SUBESPAÇO DE RUÍDO

No caso do sinal contaminado por ruído, a independência entre as amostras de ruído branco faz com que o posto da matriz A seja cheio e igual a: $\min[2.(N - L), L]$. Portanto, agora não mais existirão apenas M valores singulares não-nulos. Já que o posto é cheio, todos os $\min[2.(N - L), L]$ valores singulares serão reais e maiores que zero (ver apêndice A).

Sem ruído, M valores singulares não nulos estão associados a M autovalores de $A^H.A$ e $A.A^H$: o valor singular ao quadrado fornece o autovalor. Estes M autovalores estão associados a M autovetores de $A^H.A$, que são $v_i, i = 1, 2, \dots, M$. Estes mesmos autovalores não-nulos, também estão associados a M autovetores de $A.A^H$, dados por $u_i, i = 1, 2, \dots, M$. Os outros $(L - M)$ autovetores de $A^H.A$, e $(2.(N - L) - M)$ autovetores de $A.A^H$ estão associados a autovalores nulos.

No caso ruidoso, os $(\min[2.(N - L), L] - M)$ valores singulares que anteriormente eram nulos, tornam-se positivos. Além disto, à medida que a relação sinal-ruído decresce, tais valores singulares aproximam-se mais e mais dos M valores singulares que não eram nulos. Portanto, percebemos uma associação direta entre os M maiores valores singulares, mais seus autovalores e autovetores, com a parte de sinal de $u(n)$; e entre os $(\min[2.(N - L), L] - M)$ valores singulares restantes, mais seus autovalores e autovetores, com a parte de ruído de $u(n)$.

Podemos, portanto, dividir o espaço do sinal $u(n)$ em dois subespaços:

1) O Subespaço de Sinal - Associado aos M maiores valores singulares de A , M autovalores e autovetores de $A^H.A$ e $A.A^H$. Diz respeito à parte de sinal de $u(n) \rightarrow$

$$\sum_{k=1}^M a_k \exp[j\omega_k n].$$

2) O Subespaço de Ruído - Associado aos $(\min[2.(N - L), L] - M)$ menores valores singulares de A , $(L - M)$ autovalores e autovetores de $A^H.A$ e $(2.(N - L) - M)$ autovalores e autovetores de $A.A^H$. Diz respeito à parte de ruído de $u(n) \rightarrow y(n)$.

À medida que a SNR decresce, mais e mais os valores singulares ruidosos aproximam-se dos valores singulares de sinal. Quando a SNR torna-se menor ou igual a

1, os valores singulares ruidosos tornam-se maiores que os de sinal. Isto é previsível, pois neste caso a potência do ruído é maior que a de sinal, tornando inclusive ambígua qual a parcela de sinal de $u(n)$. Esta situação confusa reflete-se nos valores singulares.

A tabela 3.2 apresenta algumas simulações mostrando o comportamento descrito acima para os valores singulares da matriz A . Foi utilizado $N = 25$, $L = 16$, $M = 2$ e diversos valores para a relação sinal-ruído. Os valores singulares são sempre ordenados em ordem decrescente, sendo portanto imperceptível quando os do subespaço de ruído tornam-se maiores que os do subespaço de sinal para SNR muito pequenas. Porém, a desordem provocada na magnitude dos valores singulares é perceptível.

SNR = -30 dB	SNR = -10 dB	SNR = 0 dB
$v_s[1] = 296.269836$	$v_s[1] = 34.968746$	$v_s[1] = 16.218092$
$v_s[2] = 245.526321$	$v_s[2] = 21.059191$	$v_s[2] = 11.908287$
$v_s[3] = 194.881516$	$v_s[3] = 19.051193$	$v_s[3] = 8.558061$
$v_s[4] = 181.534012$	$v_s[4] = 16.572426$	$v_s[4] = 6.190883$
$v_s[5] = 149.273468$	$v_s[5] = 15.959529$	$v_s[5] = 5.583709$
$v_s[6] = 118.564430$	$v_s[6] = 14.659857$	$v_s[6] = 5.346431$
$v_s[7] = 100.814491$	$v_s[7] = 14.316841$	$v_s[7] = 4.556769$
$v_s[8] = 94.716217$	$v_s[8] = 13.393435$	$v_s[8] = 4.363620$
$v_s[9] = 79.989990$	$v_s[9] = 11.745172$	$v_s[9] = 4.028704$
$v_s[10] = 65.812569$	$v_s[10] = 10.109732$	$v_s[10] = 3.484564$
$v_s[11] = 60.632576$	$v_s[11] = 8.918715$	$v_s[11] = 2.710432$
$v_s[12] = 46.269432$	$v_s[12] = 7.098941$	$v_s[12] = 2.371058$
$v_s[13] = 34.063324$	$v_s[13] = 5.811961$	$v_s[13] = 2.015867$
$v_s[14] = 22.343048$	$v_s[14] = 4.850129$	$v_s[14] = 1.494554$
$v_s[15] = 14.755050$	$v_s[15] = 3.911629$	$v_s[15] = 1.082530$
$v_s[16] = 9.555789$	$v_s[16] = 2.537366$	$v_s[16] = 0.869409$
SNR = 10 dB	SNR = 30 dB	SNR = ∞
$v_s[1] = 16.215933$	$v_s[1] = 15.789410$	$v_s[1] = 15.756425$
$v_s[2] = 7.525684$	$v_s[2] = 8.387375$	$v_s[2] = 8.381574$
$v_s[3] = 2.712126$	$v_s[3] = 0.186623$	$v_s[3] = 0.000001$
$v_s[4] = 2.413261$	$v_s[4] = 0.161882$	$v_s[4] = 0.000000$
$v_s[5] = 2.311778$	$v_s[5] = 0.144090$	$v_s[5] = 0.000000$
$v_s[6] = 1.846060$	$v_s[6] = 0.135992$	$v_s[6] = 0.000000$
$v_s[7] = 1.605766$	$v_s[7] = 0.123777$	$v_s[7] = 0.000000$
$v_s[8] = 1.364724$	$v_s[8] = 0.103496$	$v_s[8] = 0.000000$
$v_s[9] = 1.196760$	$v_s[9] = 0.078011$	$v_s[9] = 0.000000$
$v_s[10] = 1.106709$	$v_s[10] = 0.070613$	$v_s[10] = 0.000000$
$v_s[11] = 0.829739$	$v_s[11] = 0.049773$	$v_s[11] = 0.000000$
$v_s[12] = 0.609888$	$v_s[12] = 0.040648$	$v_s[12] = 0.000000$
$v_s[13] = 0.432652$	$v_s[13] = 0.036060$	$v_s[13] = 0.000000$
$v_s[14] = 0.338053$	$v_s[14] = 0.027443$	$v_s[14] = 0.000000$
$v_s[15] = 0.138582$	$v_s[15] = 0.021676$	$v_s[15] = 0.000000$
$v_s[16] = 0.104834$	$v_s[16] = 0.006337$	$v_s[16] = 0.000000$

Tab. 3.2 - Valores Singulares da Matriz de Dados - $L = 16$

Os Zeros do Filtro de Erro de Predição

O posto de A é agora $k = \min[2.(N - L),L]$. Portanto, os teoremas demonstrados na secção 3.2 não mais se verificam. A distribuição dos zeros de sinal e dos zeros estranhos de $W(z)$ não possuem mais a simetria que possibilitava a detecção das frequências das exponenciais. Apenas nos casos onde a SNR é alta, os zeros de $W(z)$ fornecerão boas estimativas para os ω_k , pois neste caso o sinal fica mais próximo de sua versão sem ruído, e os teoremas da secção 3.2 tornam-se aproximadamente verdadeiros.

3.4.3 O MÉTODO FBLP PARA ESTIMAR FREQUÊNCIAS

A versão do método FBLP para estimar frequências de exponenciais em meio a ruído branco torna-se:

1) Formar a matriz de dados A e o vetor de respostas desejadas b , forward-backward, dados por (2.54) e (2.55).

2) Encontrar a solução de norma mínima, $w = A^* . b$. Da mesma forma que no ítem 3.3.2, é necessário aqui levar em conta duas situações distintas: 2.1) $2.(N - L) \geq L$ e 2.2) $2.(N - L) \leq L$. Cada uma delas apresenta duas opções de cálculo do vetor w .

2.1) $2.(N - L) \geq L$

1ª opção

a) Calcular a SVD de A

b) Utilizar a expressão (2.123), com o posto k de A igual a L .

$$w = \sum_{i=1}^L \frac{v_i}{\sigma_i^2} \cdot v_i^H . A^H . b \tag{3.31}$$

2ª opção

Utilizamos (2.78) e:

$$w = A^* . b$$

$$w = (A^H . A)^{-1} . A^H . b \tag{3.32}$$

pois o posto de $A^H . A$ será cheio e igual a L . Logo $A^H . A$ é não-singular e existe a inversa.

2.2) $2.(N - L) \leq L$

1ª opção

a) Calcular a SVD de A

b) Utilizar a expressão (2.126), com o posto k de A igual a $2.(N - L)$

$$w = \sum_{i=1}^{2.(N-L)} \frac{u_i^H \cdot b}{\sigma_i^2} \cdot A^H \cdot u_i \quad (3.33)$$

2ª opção

Utilizamos (2.79) e:

$$w = A^* \cdot b$$

$$w = A^H \cdot (A \cdot A^H)^{-1} \cdot b \quad (3.34)$$

3) Encontrar os zeros do polinômio $W(z)$, dado por:

$$W(z) = 1 + \sum_{k=1}^L (-w_k^*) \cdot z^{-k}$$

4) Os M zeros mais próximos da c.r.u. são tomados como estimativas dos $\omega_k, k = 1, 2, \dots, M$.

Vale salientar que, para o caso de exponenciais em ruído, o teorema da unicidade garante que no caso 2.1 existe uma única solução. De fato, (3.32) comprova esta afirmação. Por outro lado, no caso 2.2 existem infinitas soluções e calculamos a de norma mínima.

Como mostramos no capítulo 2, a solução de norma mínima pode ser calculada por uma das seguintes expressões:

$$w = \sum_{i=1}^k \frac{v_i}{\sigma_i^2} \cdot v_i^H \cdot A^H \cdot b \quad (3.35)$$

$$w = \sum_{i=1}^k \frac{u_i^H \cdot b}{\sigma_i^2} \cdot A^H \cdot u_i \quad (3.36)$$

onde: $k = \min[2.(N - L), L] = \text{posto de } A$

(3.35) e (3.36) são equivalentes e fornecem w de norma mínima. O passo 2, acima,

poderia ser substituído por (3.35) ou (3.36); a sub-divisão em 2.1 e 2.2 objetivou apenas as formas mais eficientes de cálculo computacional. Além da matriz A e do vetor b , (3.35) envolve autovalores e autovetores de $A^H.A$, enquanto que (3.36) envolve autovalores e autovetores de $A.A^H$.

Para efeito de simplificação, daqui por diante, vamos nos referir apenas à eq. (3.35), mas é evidente que todas as considerações tecidas também serão válidas para (3.36).

Como visto na eq. (2.20), podemos denominar $A^H.A$ da seguinte forma:

$$A^H.A = (N - L).\hat{R}$$

onde: \hat{R} é a matriz de correlação estimada polarizada para a sequência $u(n)$.

Como a constante multiplicativa $(N - L)$ em nada afeta o comportamento do estimador FBLP, daqui por diante assumiremos:

$$\hat{R} = A^H.A$$

Resolução X Estabilidade

Quanto maior L , maior o número de zeros e mais perfeita será a tarefa de branquear o sinal erro na saída, pois os zeros podem se arranjar mais próximos à c.r.u. e mais próximos entre si. Com isto, aumenta a capacidade de encontrarmos zeros devidamente alocados às exponenciais complexas na tentativa de anulá-las, ou seja, maior grau de liberdade é dado aos zeros do preditor. Intuitivamente, o valor de L deveria ser tão grande quanto possível, pois assim a resolução do estimador seria a maior possível. Entretanto, o número de produtos $(2.(N - L))$, que somados formam os elementos da matriz de correlação estimada \hat{R} , diminui à medida que L cresce, tornando esta matriz, e por consequência, seus autovalores e autovetores, mais instáveis; com isto, a variância (instabilidade) do estimador aumenta. Portanto, existe um compromisso entre *resolução e estabilidade* do estimador.

Pequenos valores de L possibilitam uma grande estabilidade, mas em compensação, devido ao empobrecimento da resolução, frequências muito próximas podem não ser distinguidas através do método FBLP. Grandes valores de L possibilitam, potencialmente, resolução alta, mas a estabilidade do estimador fica comprometida, provocando picos espectrais espúrios no espectro amostra.

Ao analisarmos a expressão (3.35) vemos que ela é constituída por:

- a) Os M maiores autovalores e M autovetores da matriz de correlação estimada \hat{R} , correspondendo ao subespaço de sinal.
- b) $(k - M)$ autovalores e $(k - M)$ autovetores da matriz de correlação estimada \hat{R} , correspondentes ao subespaço de ruído. Estes autovalores e autovetores ruidosos são bastante instáveis, contribuindo assim de forma marcante para o aumento da instabilidade do estimador. É importante observar em (3.35) e (3.36) que o cálculo de w depende dos valores inversos dos pequenos autovalores ruidosos. Esta inversão amplifica a instabilidade provocada por estes autovalores.

Portanto, no que diz respeito à instabilidade do método FBLP, duas conclusões, de importância vital, são atingidas:

- 1) Existe um compromisso entre *resolução e estabilidade* do estimador, relacionando o grau de liberdade do preditor ao número de produtos realizados para estimar os elementos de \hat{R} . Este compromisso está diretamente ligado ao valor de L .
- 2) Os autovalores e autovetores ruidosos também provocam instabilidade. Este efeito, em conjunto com o do item 1, definem o grau de instabilidade do estimador. Como o número de autovalores-autovetores ruidosos $(k - M)$ também depende de L , escolher o valor adequado para este parâmetro torna-se crucial para alcançar o melhor desempenho do estimador de frequência.

Simulações

Ao introduzirmos o ruído branco $y(n)$, o sinal $u(n)$ torna-se parte estocástico e parte predizível, o que faz com que os zeros do filtro de erro de predição, assim como as estimativas, assumam natureza estocástica. A fim de analisarmos o desempenho de um estimador, dois parâmetros tornam-se necessários: polarização e variância [2]. Para visualisarmos o comportamento destes parâmetros, foram realizadas simulações envolvendo o seguinte sinal:

$$u(n) = \exp[j.(2.\pi.f_1.n + \phi_1)] + \exp[j.(2.\pi.f_2.n + \phi_2)] + y(n)$$

$n = 1, 2, \dots, 25$

onde:

$f_1 = 0.5$	$\phi_1 = -\pi$
$f_2 = 0.52$	$\phi_2 = -(0.79).\pi$

Utilizamos $N = 25$, $M = 2$, $SNR = 10$ dB. SNR é definida como $10.\log \frac{1}{2.\sigma^2}$, onde σ^2 é a variância de cada uma das partes, real e imaginária, do ruído branco. 30

experimentos foram realizados para cada valor de L . Cada experimento utilizou as mesmas amostras de sinal, mudando apenas o conjunto de 25 amostras do ruído $y(n)$, porém mantendo a SNR exigida. Foram utilizados os seguintes valores para L : 4, 8, 12, 16, 18, 20, 22, 24. As figs. 3.4a a 3.4h mostram, para cada valor de L e 30 experimentos superpostos, os zeros do filtro de erro de predição e o espectro amostra. As posições angulares corretas das frequências estão indicadas por setas.

As seguintes observações podem ser feitas a partir destas figuras:

- 1) $L = 4$ é um valor muito pequeno, e o estimador não consegue distinguir as duas frequências.
- 2) À medida que L cresce, um menor número de produtos $2.(N - L)$ vão sendo realizados para a estimativa \hat{R} e um maior número de autovalores e autovetores ruidosos participam do cálculo de w , aumentando assim a variância dos zeros estranhos. Os zeros de sinal, porém, ficam melhor resolvidos.
- 3) Em $L = 8$, as duas frequências começam a ser identificadas.
- 4) A partir de $L = 12$ a variância é tão grande, que frequentemente os zeros de estranhos ficam mais próximos da c.r.u. do que os próprios zeros de sinal. Isto provoca picos espectrais espúrios, os quais podem ser observados nos espectros amostras, tornando assim o método FBLP inviável.
- 5) Para $L \geq 20$ a situação começa a melhorar. Observando a eq. (3.35) constatamos que, para $L > 2.N/3$, o posto $k = 2.(N - L)$, da matriz de correlação estimada diminui com o crescimento do L . Com isto, o número de autovalores e autovetores ruidosos que contribuem para o cálculo de w , também diminui. Por exemplo, para $L = 22$, apenas 4 autovetores ruidosos contribuem para formar w . Isto provoca uma diminuição na variância, que tende a compensar a degradação provocada pelo menor número de produtos realizados para compor os elementos de \hat{R} . $L = (N - M/2) = 24$ é o melhor caso, pois nenhum autovalor-autovetor ruidoso participa do cálculo de w . Este caso é conhecido como *Kumaresan-Prony* e será abordado com maiores detalhes no próximo capítulo.
- 6) Em todos os casos, a polarização foi pequena quando comparada à variância e para grandes valores de L torna-se progressivamente desprezível [7].

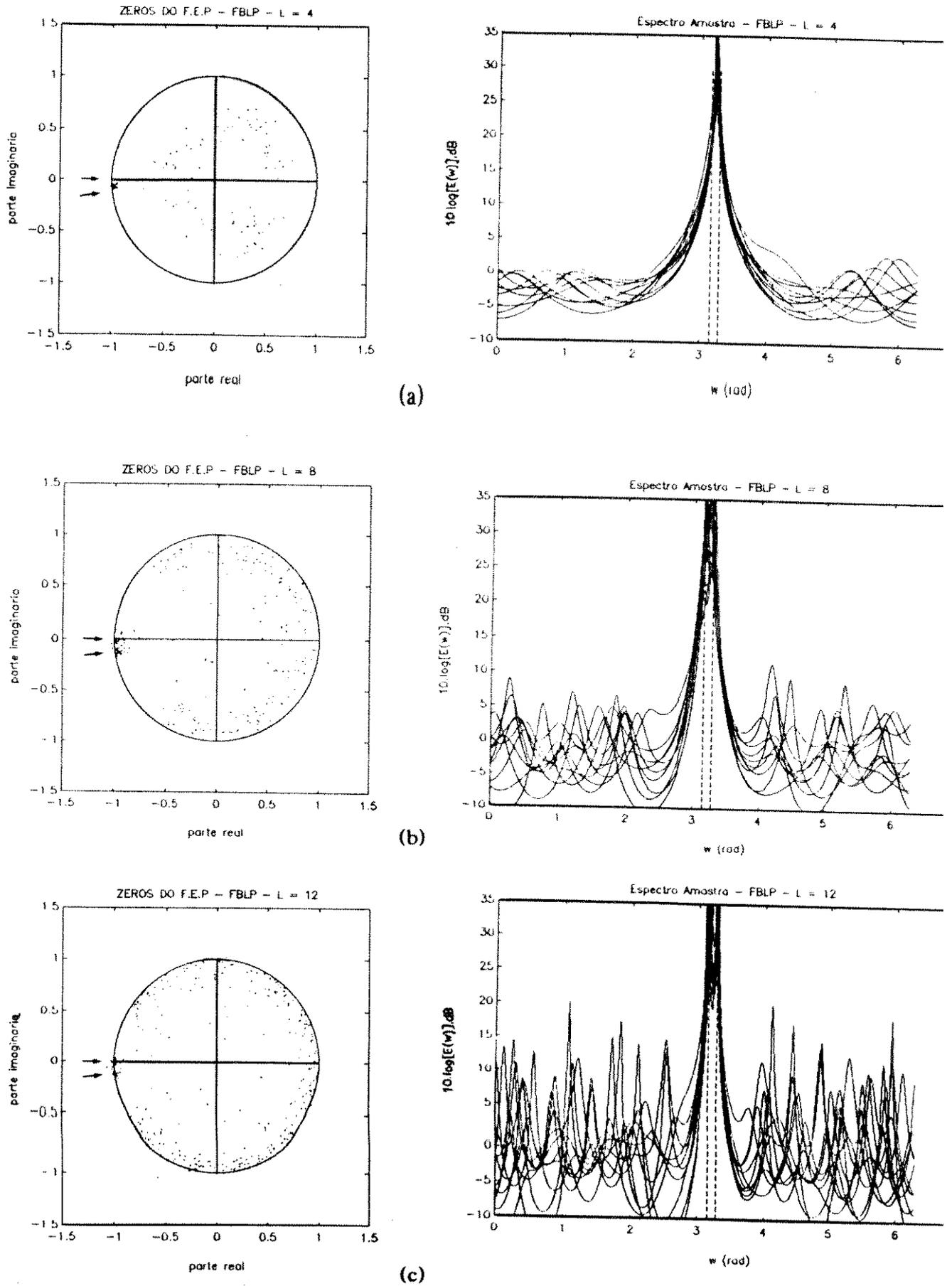
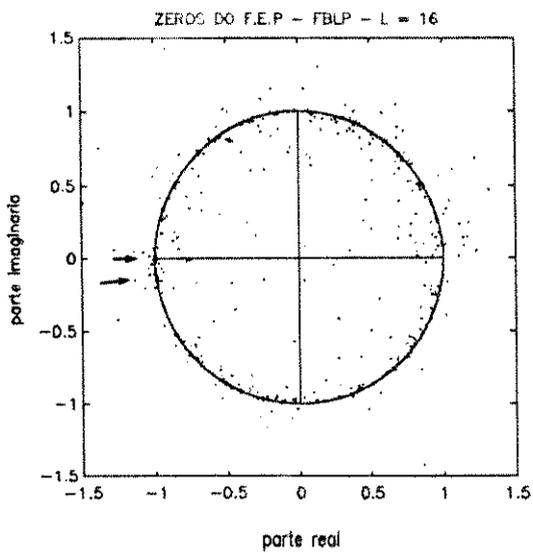
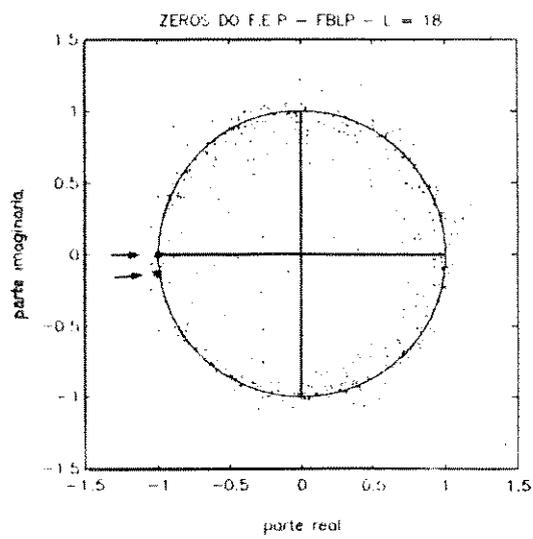
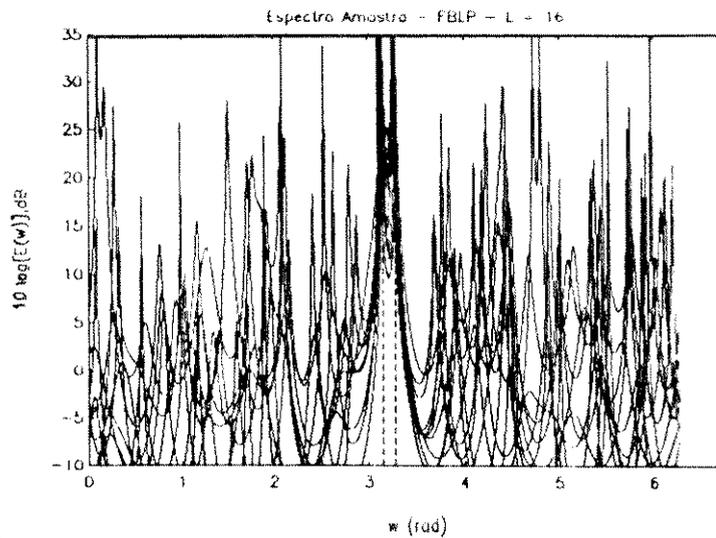


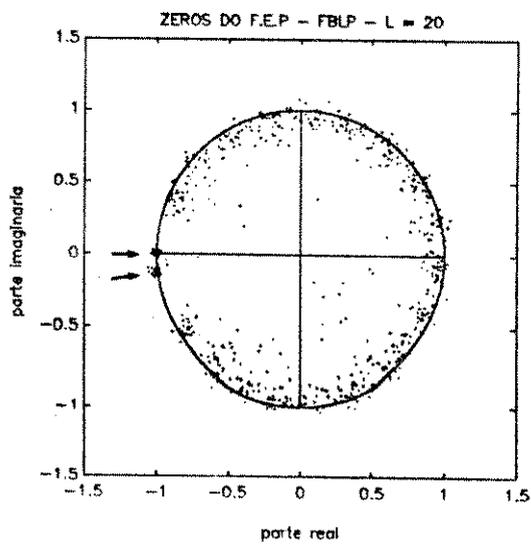
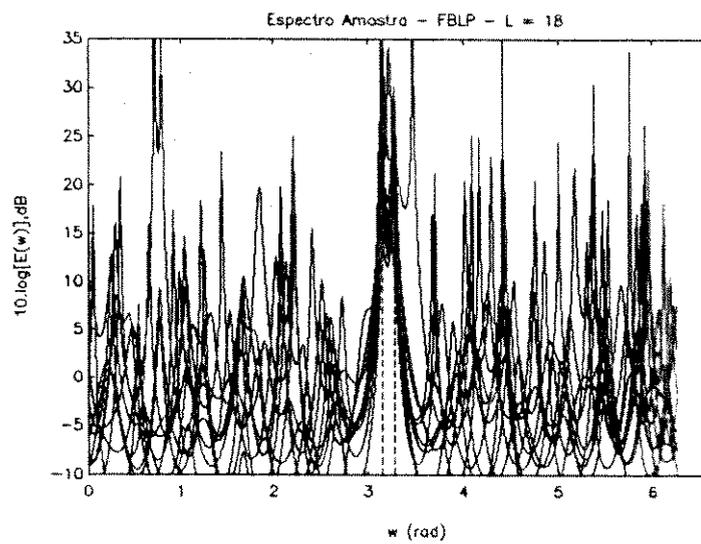
Fig. 3.4 - Método FBLP - Zeros do F.E.P. e Espectro Amostra - SNR = 10 dB - 30 experimentos (contínua)



(d)



(e)



(f)

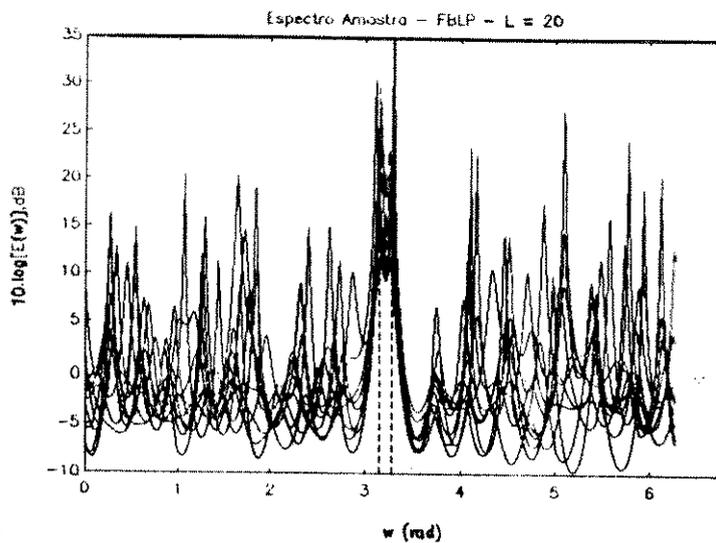


Fig. 3.4 - Método FBLP - Zeros do F.E.P. e Espectro Amostra - SNR = 10 dB - 30 experimentos (contínua)

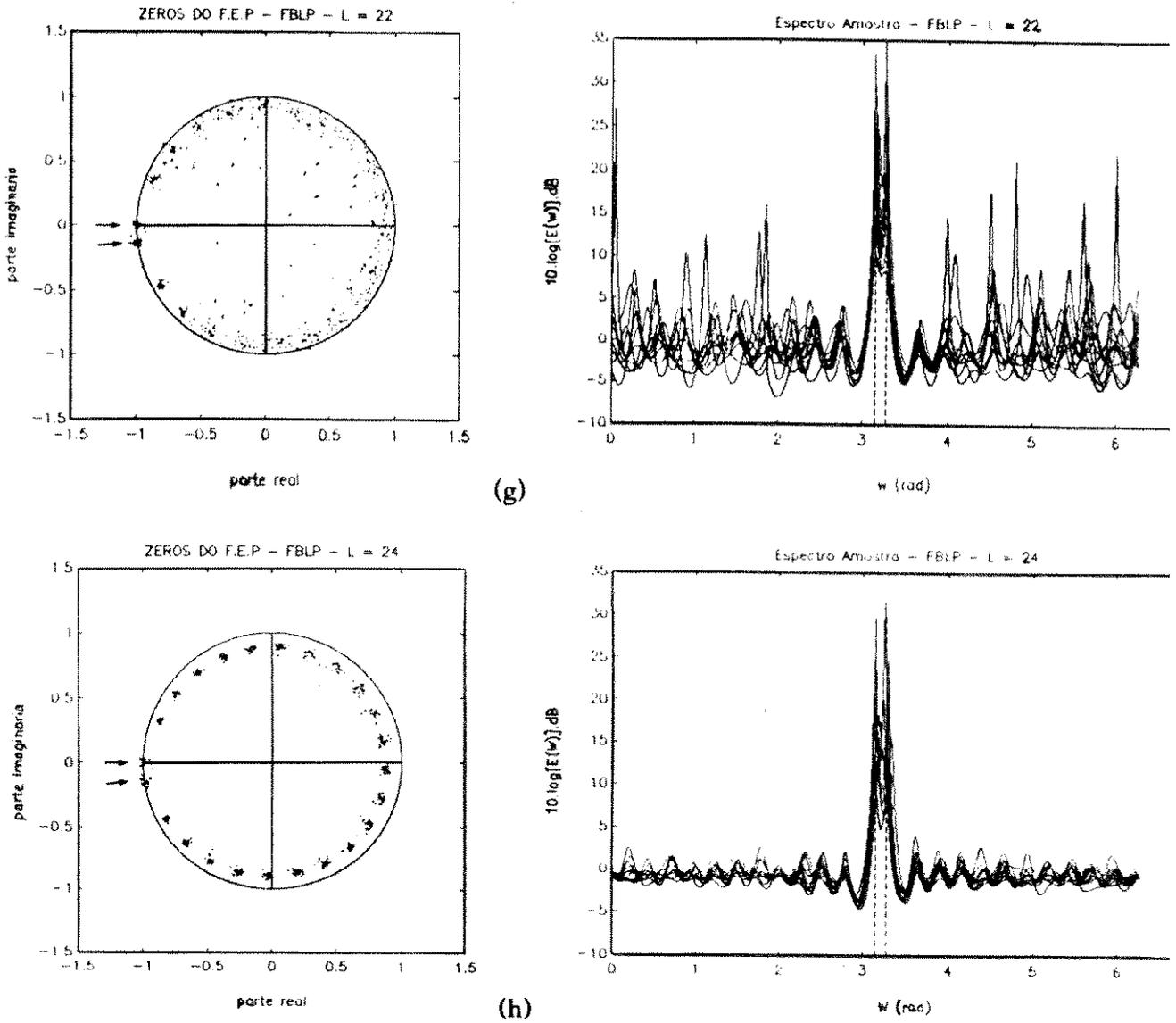


Fig. 3.4 - Método FBLP - Zeros do F.E.P. e Espectro Amostra - SNR = 10 dB - 30 experimentos

- (a) L = 4 (b) L = 8 (c) L = 12 (d) L = 16
 (e) L = 18 (f) L = 20 (g) L = 22 (h) L = 24

3.4.4 Limitações do Método FBLP

Para frequências separadas por uma distância menor que o recíproco do intervalo de observação do sinal e valores pequenos da relação sinal-ruído, como no exemplo anterior, vimos, através das simulações, que o método FBLP torna-se inviável.

A fim de avaliar o desempenho de alguns estimadores de frequência não polarizados de uma maneira quantitativa, traçamos na fig. 3.5 a curva de variância da estimativa em função da relação sinal-ruído(SNR). Cada ponto do gráfico é o

resultado de 500 experimentos independentes, cada qual empregando 25 amostras do sinal do exemplo anterior. Para cada experimento, os ângulos dos dois zeros do polinômio $W(z)$ mais próximos à c.r.u. foram seleccionados, e o menor deles, medido no sentido anti-horário a partir da origem do plano z , foi tomado como a estimativa $\hat{\omega}_1$ da frequência $\omega_1 = 2\pi(0.5)$. A seguir, calculamos a variância da estimativa da frequência f_1 , que para estimadores não polarizados é igual ao erro quadrático médio, fornecida pela seguinte expressão:

$$\text{var}(\hat{f}_1) = \left(\sum_{k=1}^{500} \left[\frac{\hat{\omega}_{1,k}}{2\pi} - 0.5 \right]^2 \right) / 500$$

onde: $\hat{\omega}_{1,k}$ denota a estimativa de ω_1 do k 'ésimo experimento

O valor $10.\log[1/\text{var}(\hat{f}_1)]$ é então traçado na fig. 3.5 em função da SNR(dB). O limite inferior para a variância de um estimador não polarizado é conhecido como *O Limite Inferior de Cramer-Rao(CR)*. Este limite foi calculado [4] para o intervalo de SNR e o sinal em questão. Os estimadores baseados no método de máxima-verossimilhança - ML^1 são os que mais se aproximam do limite inferior CR [1 - pg. 543]. As curvas para os estimadores ML e FBLP ($L = 12$ e $L = 18$) foram reproduzidas a partir de [7].

Observando a fig. 3.5 e as simulações do ítem anterior, chegamos às seguintes conclusões para o método FBLP:

1) Para valores suficientemente elevados da SNR, o desempenho do método FBLP é alguns decibéis mais pobre do que o limite CR, o qual, porém, é alcançado pelo método ML.

2) O método FBLP apresenta um limiar, que é alcançado para valores relativamente elevados da SNR, a partir do qual o desempenho do método deteriora-se rapidamente. Este limiar também ocorre no método ML, porém para valores da SNR bem menores. O limiar do método FBLP é causado por um dos seguintes motivos:

a) Quando a ordem L do preditor é muito pequena.

A resolução do método é deficiente e não é possível distinguir duas frequências muito próximas. Observado através do espectro amostra, este efeito provoca a superposição de picos espectrais próximos.

b) Quando a ordem L do preditor é uma grande fração do comprimento N do conjunto de

¹do inglês Maximum Likelihood

amostras do sinal.

Vimos que o crescimento de L provoca um aumento da variância da matriz de correlação estimada, e conseqüentemente dos seus autovalores e autovetores, que participam da formação do vetor w . Este efeito, somado à presença dos autovalores-autovetores ruidosos, aumenta a instabilidade do posicionamento dos zeros ruidosos, permitindo aproximações perigosas junto à c.r.u.. Estas aproximações fazem com que existam eventuais zeros estranhos mais próximos da c.r.u. que as próprias estimativas dos zeros de sinal. Nestes casos, os zeros estranhos são tomados como as estimativas das frequências. Logo, um grande erro é cometido e o limiar é alcançado. Este efeito provoca um aumento da amplitude dos picos espectrais espúrios, tornando-os indistinguíveis dos picos espectrais devido aos zeros de sinal.

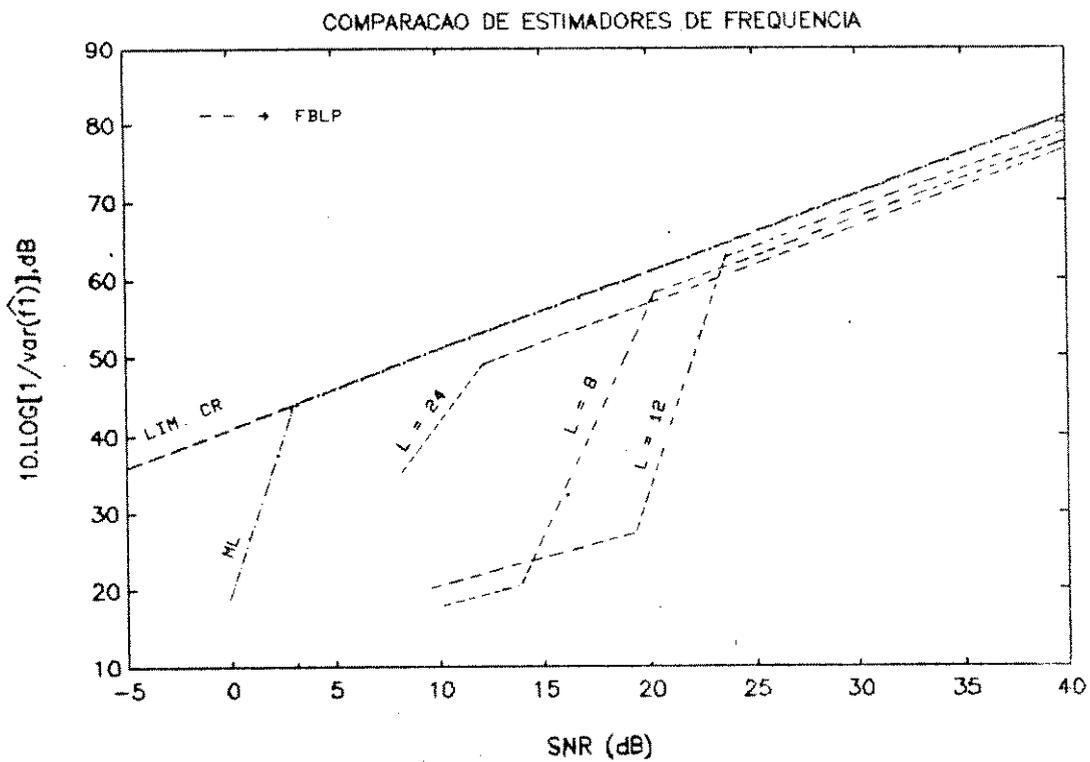


Fig. 3.5 - Comparação de desempenho entre os Métodos FBLP e ML

3.5 CONCLUSÕES

A teoria desenvolvida neste capítulo mostrou que os zeros do polinômio formado pelos coeficientes do filtro de erro de predição do sinal, possuem propriedades poderosas que permitem detectar as frequências das exponenciais. A

Predição forward-backward deve ser utilizada dando origem ao *Método Forward-Backward para Estimação de Frequências*.

Quando o ruído está presente, as propriedades dos zeros do polinômio $W(z)$ não mais se verificam e o método deteriora-se demasiadamente. Realizando uma análise através de autovalores e autovetores, constatamos que o sinal pode ser dividido em dois subespaços: o de sinal e o de ruído. Os elementos que compõem o subespaço de ruído são os grandes responsáveis pela queda no desempenho do método FBLP, principalmente no que diz respeito à instabilidade do estimador. Verificamos também que existe um compromisso entre resolução e estabilidade e que este compromisso depende da ordem do preditor. Por fim, devido à grande queda de desempenho provocada pelos autovalores-autovetores do subespaço ruidoso, constatamos que o método FBLP possui grandes limitações, principalmente para valores da SNR pequenos e frequências de exponenciais muito próximas.

Referindo-nos à equação (3.35), observamos que $k = \min[2.(N - L), L]$ autovalores e autovetores da matriz de correlação estimada \hat{R} contribuem para o cálculo de w . Os autovalores do subespaço de sinal, aqueles M maiores autovalores, geralmente apresentam uma magnitude bem maior do que o grupo de autovalores de ruído. Além disso, estes autovalores, e consequentemente os autovetores, são mais robustos à presença do ruído.

Os autovetores do subespaço de ruído de \hat{R} têm uma contribuição significativa para w , pois se não participavam do cálculo no caso sem ruído, agora eles são amplificados pelos recíprocos dos pequenos autovalores a eles associados. Isto resulta em consideráveis flutuações dos coeficientes w do filtro de erro de predição, resultando, eventualmente, em zeros que provocarão picos espectrais espúrios. Estes efeitos são mais acentuados para grandes valores do posto de \hat{R} e pequenos valores da relação sinal-ruído, pois mais autovalores-autovetores do subespaço de ruído participam do cálculo de w .

Esta situação se modifica significativamente quando $L = (N - M/2)$, pois então apenas os autovetores de sinal são empregados no cálculo de w . Este fato, aliado à grande resolução proporcionada pelo valor de L , produz resultados interessantes. Este caso é conhecido como Kumaresan-Prony e será analisado com mais detalhes no próximo capítulo.

Tomando como base estas observações, Kumaresan e Tufts [7] propuseram um novo método para estimar as frequências das exponenciais, o qual é objeto do próximo capítulo.

CAPÍTULO 4

O MÉTODO FBLP MODIFICADO

CONTEÚDO	Pag.
4.1 Introdução.....	81
4.2 Melhorando o Método FBLP.....	82
4.3 O Método FBLP Modificado.....	84
4.4 Comparações Quantitativas.....	89
4.5 O Caso Kumaresan-Prony - KP.....	92
4.6 Variação de Alguns Parâmetros.....	93
4.7 Estimativa do Número de Exponenciais.....	95
4.8 Uma Aplicação Típica: Detecção de Ângulo de Incidência Vertical em Radar.....	97
4.9 Conclusões.....	103

4.1 INTRODUÇÃO

No final do capítulo 3, foi visto que o método FBLP não fornece resultados satisfatórios quando as frequências das exponenciais estão muito próximas e a relação sinal-ruído é pequena. Um dos motivos para este desempenho fraco são os efeitos produzidos pelos autovalores e autovetores do subespaço de ruído, os quais causam instabilidade nos valores estimados para os zeros da função de transferência do filtro de erro de predição.

Este capítulo apresenta um novo método onde se obtém uma versão de posto reduzido da matriz A , através da supressão dos autovalores e autovetores do subespaço de ruído. Esta versão de posto reduzido se aproxima da matriz ideal do caso sem ruído, produzindo efeitos semelhantes a um aumento da relação sinal-ruído. Chamaremos o novo método de *Método FBLP Modificado*.

Simulações serão realizadas, bem como comparações com os métodos FBLP e

ML. Quando L é feito igual ao seu limite superior ($N - M/2$), obteremos um caso de especial interesse, devido à sua simplicidade computacional e desempenho razoável: é o caso *Kumaresan-Prony*.

Analisaremos o desempenho do Método FBLP Modificado quanto à variação de alguns parâmetros tais como: número de amostras disponíveis, ordem do preditor, relação sinal-ruído e posicionamento das frequências.

Até aqui, admitimos o número de exponenciais M como conhecido. Caso isto não ocorra, deverão ser realizadas estimativas deste parâmetro. Neste capítulo serão discutidos alguns procedimentos para estas estimativas. Por fim, apresentaremos a aplicação do método na detecção de ângulo de incidência vertical em radar.

4.2 MELHORANDO O MÉTODO FBLP

O método FBLP sofre danos severos devido à presença dos $(k - M)$ autovalores-autovetores do subespaço de ruído de $\hat{\mathbf{R}}$. O ideal seria que conseguíssemos descartar a parte ruidosa do sinal, de maneira que os zeros de $W(z)$ permitissem detectar, de forma precisa, as frequências das exponenciais complexas. Obter esta SNR infinita não é possível, mas podemos adotar medidas cujos efeitos se parecem com aqueles produzidos por um aumento da SNR do sinal dado por:

$$u(n) = \sum_{k=1}^M a_k \cdot \exp[j(\omega_k \cdot n + \phi_k)] + y(n) \quad n = 1, 2, \dots, N \quad (4.1)$$

onde $y(n)$ é um ruído branco.

Descreveremos, agora, os dois passos que determinam o novo método:

1) Cálculo de uma nova matriz de correlação estimada $\tilde{\mathbf{R}}$

Da álgebra matricial, sabemos que a autodecomposição da matriz de correlação estimada $\hat{\mathbf{R}}$ é fornecida por [1 - pg. 59]:

$$\hat{\mathbf{R}} = \mathbf{A}^H \cdot \mathbf{A} = \sum_{l=1}^k \sigma_l^2 \cdot \mathbf{v}_l \cdot \mathbf{v}_l^H \quad (4.2)$$

onde: k é o posto de $\hat{\mathbf{R}}$
 $\sigma_l^2, l = 1, 2, \dots, k$, são os autovalores não nulos de $\hat{\mathbf{R}}$
 $\mathbf{v}_l, l = 1, 2, \dots, k$, são os autovetores ortonormais de $\hat{\mathbf{R}}$
 associados a $\sigma_l^2, l = 1, 2, \dots, k$

A proposta aqui apresentada foi originalmente desenvolvida por Tufts e Kumaresan [7] e consiste em utilizar uma nova matriz de correlação dada pela seguinte expressão:

$$\tilde{\mathbf{R}} = \tilde{\mathbf{A}}^H \tilde{\mathbf{A}} = \sum_{i=1}^M \sigma_i^2 \mathbf{v}_i \mathbf{v}_i^H \quad (4.3)$$

onde estão presentes apenas os M maiores autovalores de $\hat{\mathbf{R}}$ e seus autovetores correspondentes.

Isto equivale a substituir a matriz A, por uma versão de posto reduzido(M) da mesma. Decorrente disto, as equações (3.35) e (3.36) transformam-se em:

$$\mathbf{w} = \sum_{i=1}^M \frac{\mathbf{v}_i}{\sigma_i^2} \cdot \mathbf{v}_i^H \cdot \mathbf{A}^H \cdot \mathbf{b} \quad (4.4)$$

$$\mathbf{w} = \sum_{i=1}^M \frac{\mathbf{u}_i^H \cdot \mathbf{b}}{\sigma_i^2} \cdot \mathbf{A}^H \cdot \mathbf{u}_i \quad (4.5)$$

Desta maneira descartamos das eqs. (3.35) e (3.36) os (k - M) autovalores-autovetores do subespaço de ruído.

A linha de raciocínio que utilizamos aqui foi que os M autovalores-autovetores do subespaço de sinal são consideravelmente mais robustos às perturbações ruidosas do que os (k - M) do subespaço ruidoso. No caso ideal, no qual o sinal u(n) não possui ruído, a matriz de correlação tem M autovalores não nulos e (L - M) autovalores nulos. Portanto, no caso prático de existência de ruído, ao retermos os M maiores autovalores e os respectivos autovetores, associados ao subespaço de sinal da matriz de correlação estimada $\hat{\mathbf{R}}$, e desprezarmos os restantes, relacionados ao subespaço de ruído, estamos objetivando nos aproximar do caso ideal sem ruído. De fato, vimos no cap. 3 que, para níveis suficientemente elevados da SNR, os M valores singulares (e por consequência autovalores e autovetores) do subespaço de sinal permanecem próximos às suas contrapartidas não ruidosas.

Tufts e Kumaresan [7] demonstraram que, para valores não muito pequenos da SNR, é mais provável que a matriz $\tilde{\mathbf{R}}$ esteja mais próxima da versão sem ruído do que $\hat{\mathbf{R}}$. Isto torna evidente que para níveis razoáveis de perturbação (SNR suficientemente grandes) os M autovalores-autovetores do subespaço de sinal de $\hat{\mathbf{R}}$ permanecem próximos às suas contrapartidas não ruidosas e, portanto, o uso de $\tilde{\mathbf{R}}$ e $\tilde{\mathbf{A}}$ em lugar de $\hat{\mathbf{R}}$ e A nos aproxima da versão sem ruído do vetor de ganhos do preditor w.

Este procedimento faz com que os zeros do filtro de erro de predição apresentem, de forma aproximada, o mesmo comportamento do caso sem ruído, fornecendo assim, melhores condições para a detecção precisa das frequências.

Obviamente, o valor de w , dado por (4.4) ou (4.5), não mais minimizará a energia do erro de predição forward-backward. Porém, isto não é relevante, pois, de fato, o que pretendemos é obter w o mais próximo possível de sua versão não ruidosa.

Descreveremos, a seguir, o segundo passo para o novo método proposto.

2) Crescimento da ordem L do preditor

Como vimos no capítulo anterior, a estabilidade do estimador está diretamente ligada ao valor de L . O número de produtos disponíveis para formar os elementos de $\hat{\mathbf{R}}$ e o número dos autovalores-autovetores ruidosos instáveis que participam do cálculo de w , determinam esta estabilidade. Uma vez que removemos o efeito indesejável dos autovalores-autovetores do subespaço de ruído, um destes fatores foi eliminado. Portanto, o novo método fornecerá resultados mais estáveis. Além disto, o crescimento de L na faixa $M < L \leq (2.N/3)$ não mais implicará em um aumento no número de autovalores-autovetores ruidosos que participam do cálculo de w , como acontece com o método FBLP. Portanto, L poderá atingir valores maiores sem provocar grandes danos à variância, aumentando assim a capacidade de resolução. Concluimos, portanto, que este novo método deve ser mais estável e possuir uma capacidade de resolução potencialmente maior que o FBLP.

Porém, mesmo para este novo método existe um limite prático para L , pois um dos fatores responsáveis pela instabilidade ainda permanece, perdurando o compromisso de resolução \times estabilidade. Se por um lado o crescimento de L aumenta o grau de liberdade dos zeros do preditor, por outro lado, um número menor de produtos são realizados, aumentando a instabilidade de $\hat{\mathbf{R}}$, e conseqüentemente de $\tilde{\mathbf{R}}$ e w . Portanto, o crescimento do desempenho com L é comprometido quando se atinge valores elevados para L . Tufts e Kumaresan [7] definiram, experimentalmente, que $L \approx (3.N/4)$ é o valor mais adequado para a ordem do preditor.

4.3 O MÉTODO FBLP MODIFICADO

Este novo método, definido por (4.4) ou (4.5), é chamado de *Método FBLP Modificado*. Vale salientar que, as matrizes $\tilde{\mathbf{A}}$ e $\tilde{\mathbf{R}}$ possuem apenas efeito didático, não sendo computadas na prática, onde o procedimento para o cálculo de w é o descrito a seguir.

1) Formar a matriz de dados A e o vetor de respostas desejadas b , forward-backward, dados por (2.54) e (2.55).

2) Calcular w

2.1) $2.(N - L) \geq L$

a) Calcular a SVD de A e identificar os M maiores valores singulares e os autovalores-autovetores associados de $A^H.A$.

b) Utilizar (4.4) para calcular $w = \tilde{A}^\# . b$

$$w = \sum_{i=1}^M \frac{v_i}{\sigma_i^2} \cdot v_i^H . A^H . b \quad (4.4)$$

2.2) $2.(N - L) < L$

a) Calcular a SVD de A e identificar os M maiores valores singulares e os autovalores-autovetores associados de $A.A^H$.

b) Utilizar (4.5) para calcular $w = \tilde{A}^\# . b$

$$w = \sum_{i=1}^M \frac{u_i^H . b}{\sigma_i^2} \cdot A^H . u_i \quad (4.5)$$

3) Encontrar os zeros do polinômio $W(z)$, dado por:

$$W(z) = 1 + \sum_{k=1}^L (-w_k^*) . z^{-k}$$

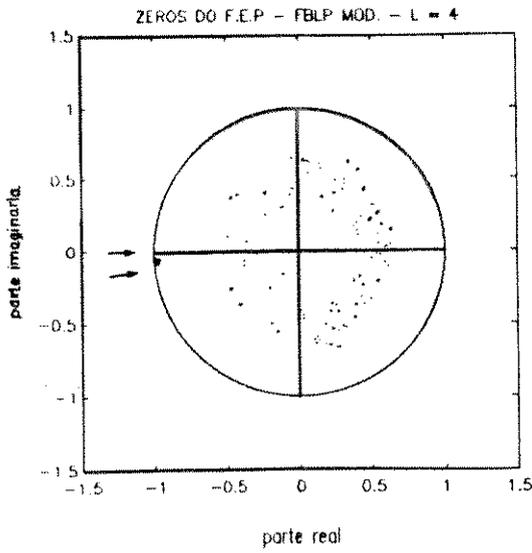
4) Os M zeros que se localizarem mais próximos à c.r.u. determinam $\omega_k, k = 1, 2, \dots, M$.

O método FBLP possui uma segunda opção para o cálculo de w , envolvendo matrizes inversas e produtos matriciais. Excetuando-se os casos em que $L = M$ e $L = (N - M/2)$, isto não é possível para o FBLP Modificado, pois agora $k = M \neq \min[2.(N - L), L]$.

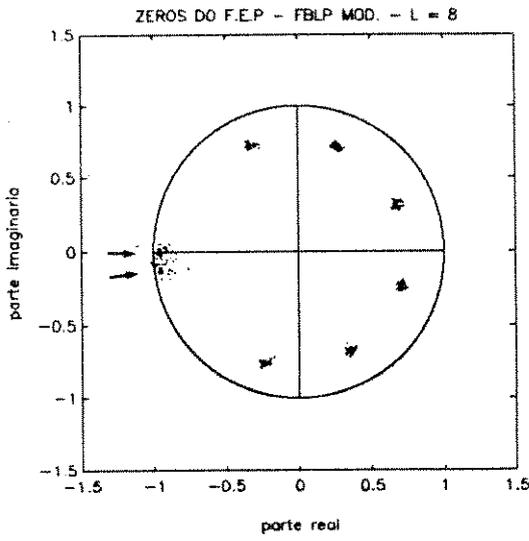
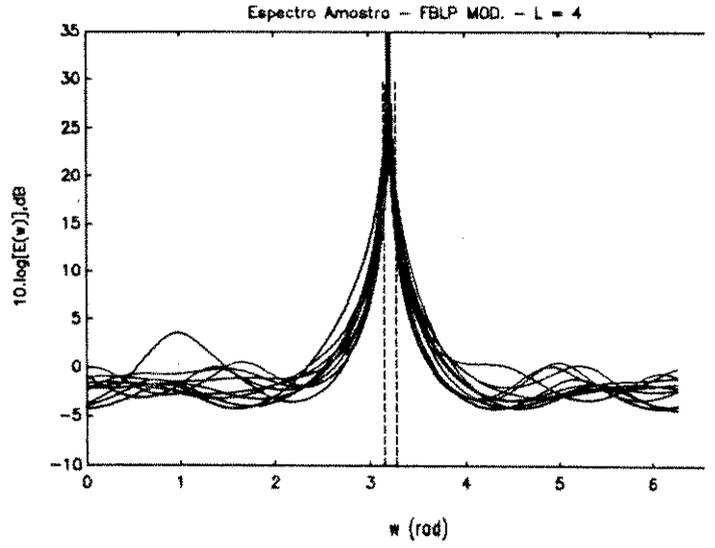
Simulações

As mesmas simulações da secção 3.4.3 são agora realizadas para o método FBLP Modificado, fornecendo as figs. 4.1a a 4.1h.

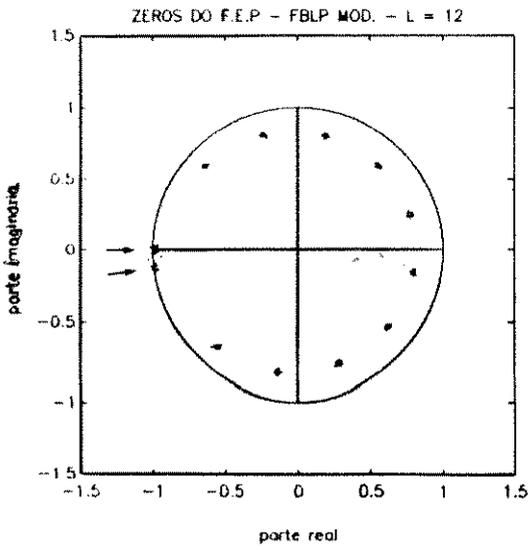
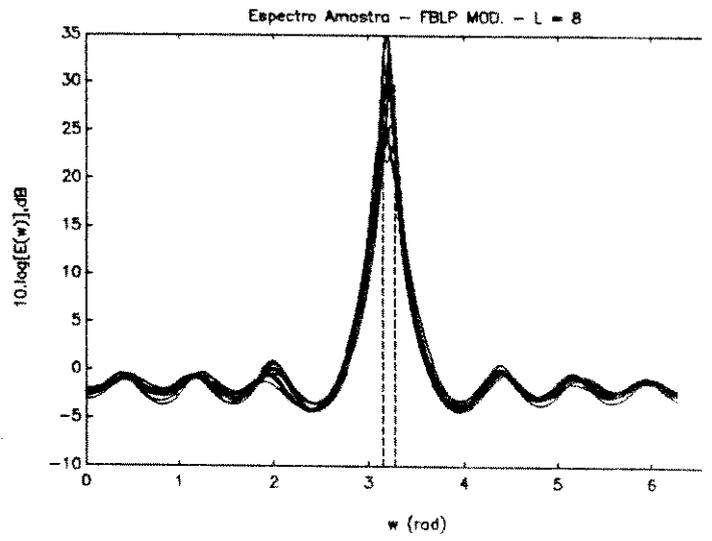
Observando estas figuras, simultaneamente com as figs. 3.4a a 3.4h, chegamos às seguintes conclusões:



(a)



(b)



(c)

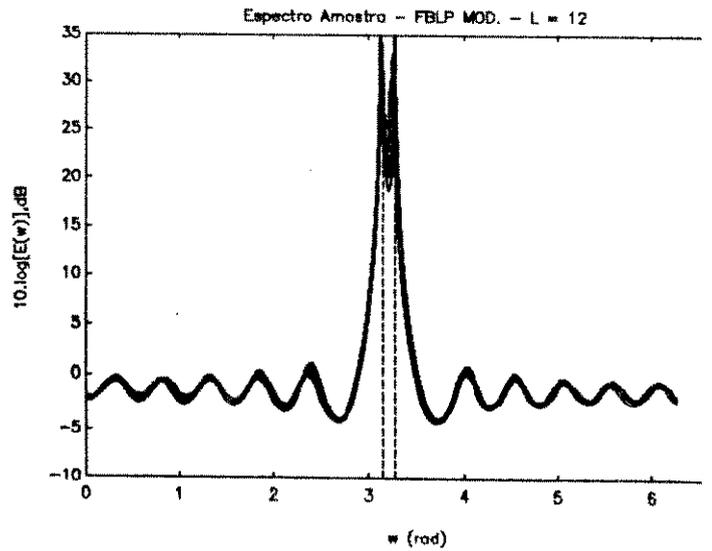
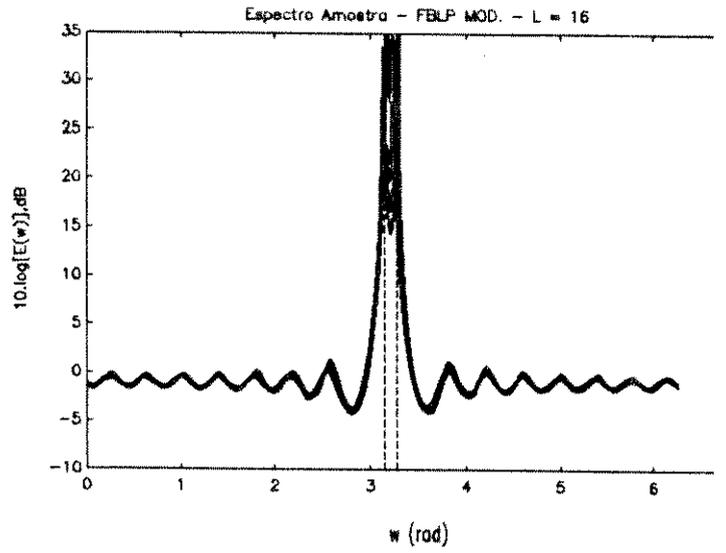
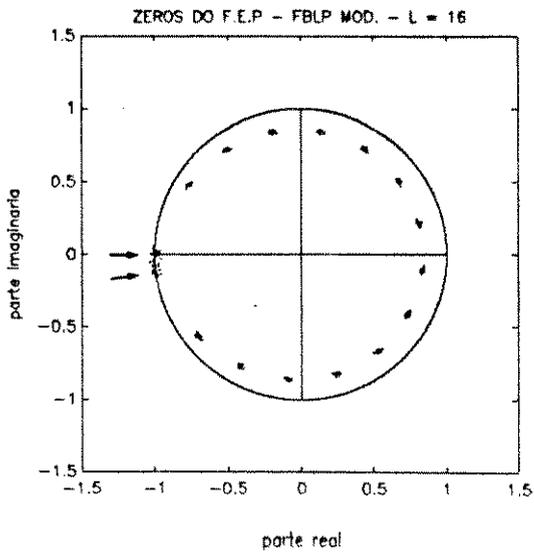
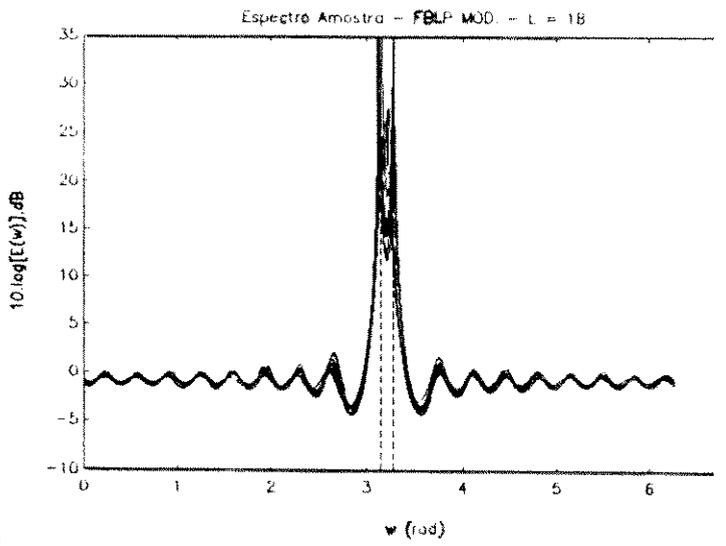
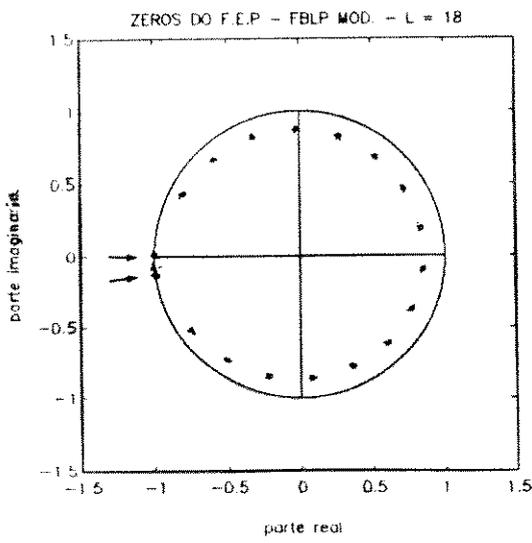


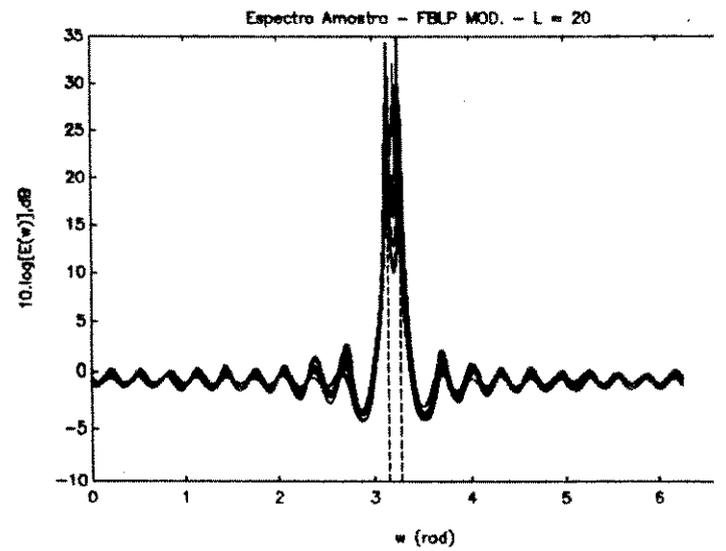
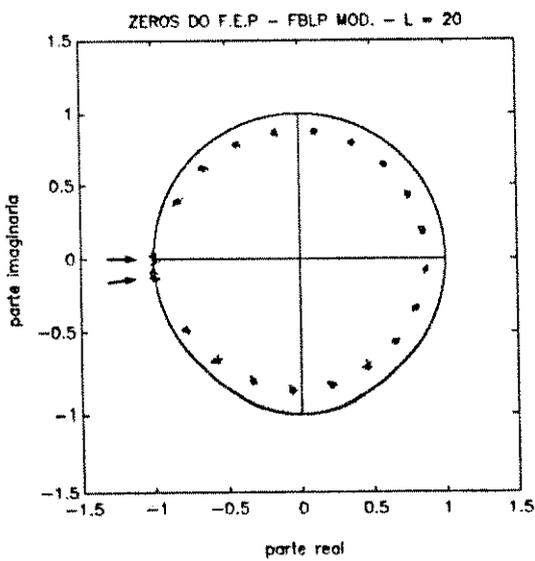
Fig. 4.1 - Método FBLP Mod. - Zeros do F.E.P. e Espectro Amostra - SNR = 10 dB - 30 experimentos (continua)



(d)



(e)



(f)

Fig. 4.1 - Método FBLP Mod. - Zeros do F.E.P. e Espectro Amostra - SNR = 10 dB - 30 experimentos (continua)

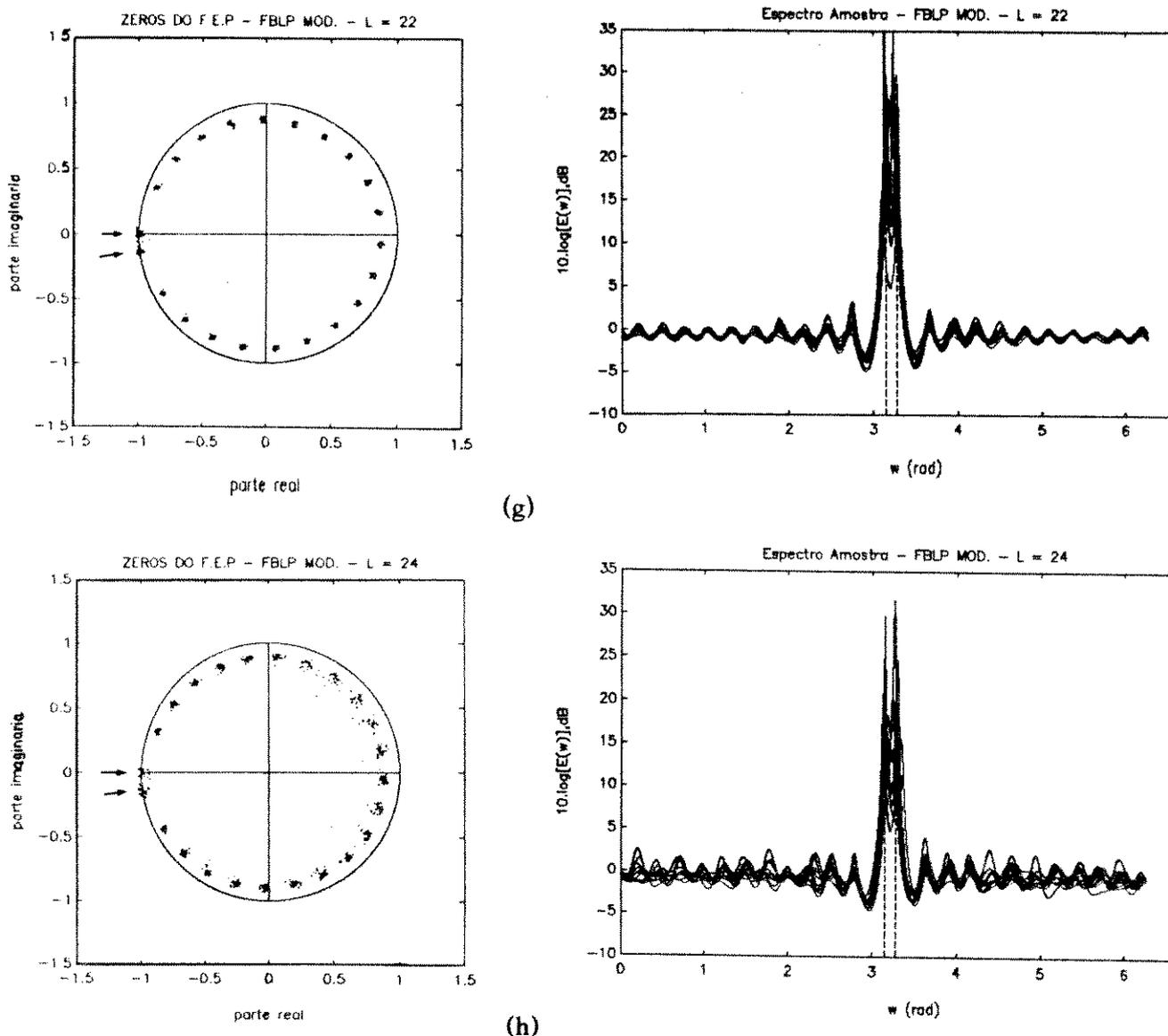


Fig. 4.1 - Método FBLP Mod. - Zeros do F.E.P. e Espectro Amostra - SNR = 10 dB - 30 experimentos

- (a) L = 4 (b) L = 8 (c) L = 12 (d) L = 16
 (e) L = 18 (f) L = 20 (g) L = 22 (h) L = 24

1) O espalhamento dos grupos de zeros do método FBLP Modificado é bem menor que o do método FBLP, refletindo, desta forma, uma maior estabilidade do FBLP Modificado.

2) A partir de $L = 4$, à medida que a ordem cresce, o espalhamento dos grupos de zeros de sinal e estranhos decresce, alcançando um mínimo em $L = 18 (\approx 3.N/4)$. Isto ocorre devido ao crescimento da resolução sem provocar grandes danos à estabilidade. A partir de $L = 20$, os grupos de zeros tornam-se mais espalhados, devido à predominância do efeito de instabilidade produzido pelo decréscimo do número de produtos realizados para compor \tilde{R} . Isto confirma a escolha de $3.N/4$ como o valor ótimo para a ordem do preditor no método FBLP Modificado.

Tanto na fig. 3.4 como na 4.1, os espectros amostras ratificam todas as observações feitas até aqui. Concluímos então que, para este caso com duas frequências próximas e $SNR = 10$ dB, o método FBLP Modificado com $L = 18$ fornece excelentes estimativas.

3) Notamos que o comportamento de ascendente ($L = 4$ a 8)- descendente ($L = 12$ a 18)- ascendente ($L = 20$ a 24) da qualidade das estimativas no método FBLP, é substituído por ascendente ($L = 4$ a 18)-descendente ($L = 20$ a 24) no método FBLP Modificado. Isto se deve à ausência dos autovalores-autovetores instáveis do subspaço ruidoso no cálculo de w , e por consequência, nos zeros de $W(z)$.

4.4 COMPARAÇÕES QUANTITATIVAS

A fig. 4.2 mostra o gráfico de SNR x variância, descrito no cap. 3 (fig. 3.5), para os estimadores FBLP, FBLP Modificado, ML e para o limite CR, a fim de obtermos uma comparação quantitativa entre os diversos estimadores. O desempenho do método FBLP para os valores $L = 8(N/3)$ e $L = 12(N/2)$, e do FBLP Modificado para $L = 8, 12, 18$ e 24 são traçados nesta figura. A fig. 4.3 mostra as SNR 's de limiar para os métodos FBLP e FBLP Modificado em diversos valores da ordem L do preditor, bem como apresenta o limiar para o método ML. A fig. 4.4 mostra a variância da estimativa da frequência f_1 , em 15, 25 e 35 dB, para o método FBLP Modificado em diversos valores de L , bem como apresenta o limite de CR para tais SNR 's.

Algumas conclusões podem ser obtidas pela observação destas figuras. O efeito de limiar no método FBLP ocorre para valores relativamente elevados da SNR : ≈ 23.5 dB para $L = 12$ e ≈ 20 dB para $L = 8$. Já no método FBLP Modificado, este limiar ocorre para valores bem inferiores da SNR : ≈ 9.4 dB para $L = 12$ e ≈ 12 dB para $L = 8$. Notamos também que, para $L = 12$, acima do limiar, o desempenho do método FBLP Modificado se aproxima mais do limite CR que os dois casos FBLP.

Para $L = 18$ e $L = 20$, acima do limiar, o desempenho do método FBLP Modificado torna-se praticamente igual ao limite CR. O limiar para $L = 14, 16$ e 18 ocorre para valores muito pequenos da SNR : cerca de 6.8 dB. Para valores de L próximos a 18, em SNR 's maiores que a de limiar, as variâncias das estimativas estão bem próximas ao limite de CR. Podemos constatar que $L = 18(3.N/4)$ representa o melhor compromisso quando consideramos conjuntamente a SNR de limiar e a variância acima do limiar. Também, $L = 16$ é uma excelente alternativa.

Mesmo abaixo do limiar, as frequências podem eventualmente ser estimadas corretamente. A fig. 4.5, onde utilizamos $L = 18$ e $SNR = 0$ dB, sugere a existência

de conjuntos de zeros nos quais os de sinal são os mais próximos da c.r.u.. Porém, o problema nestes níveis de SNR é a eventualidade deste evento. O comportamento observado na figura 4.2, abaixo do limiar, é o desempenho médio decorrente de situações onde os zeros de sinal são identificados corretamente, juntamente com aquelas onde são confundidos com zeros estranhos.

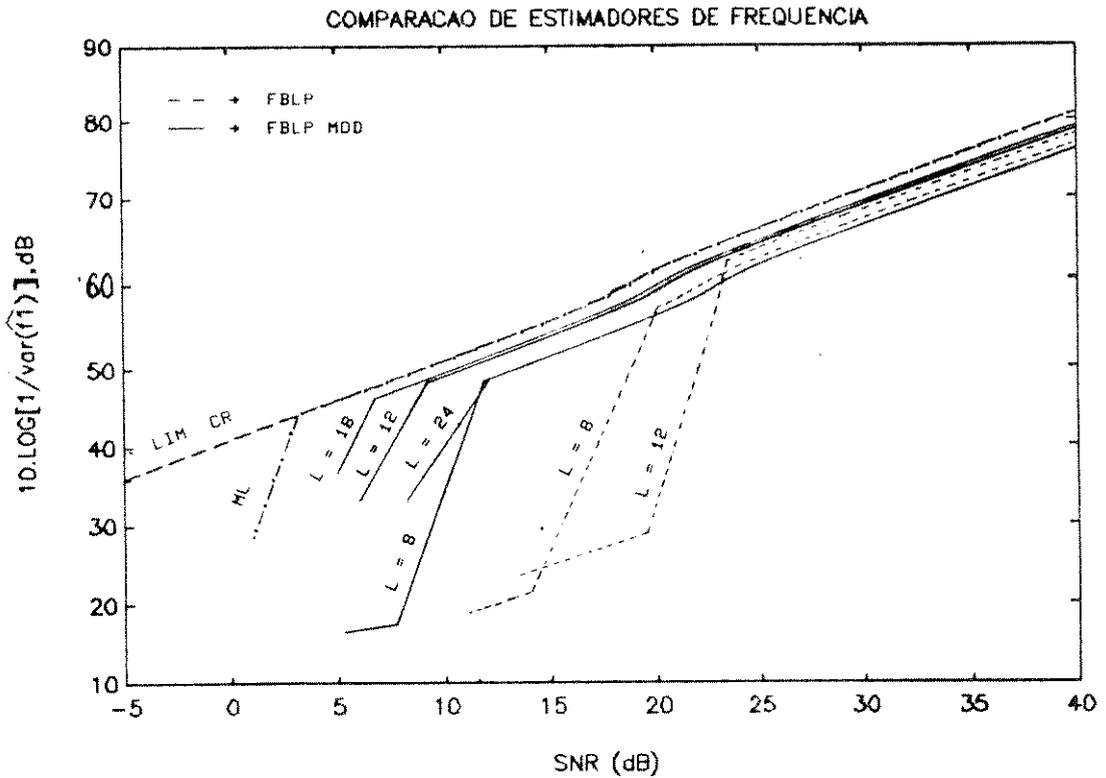


Fig. 4.2 - Comparação de desempenho entre os Métodos FBLP, FBLP Mod. e ML

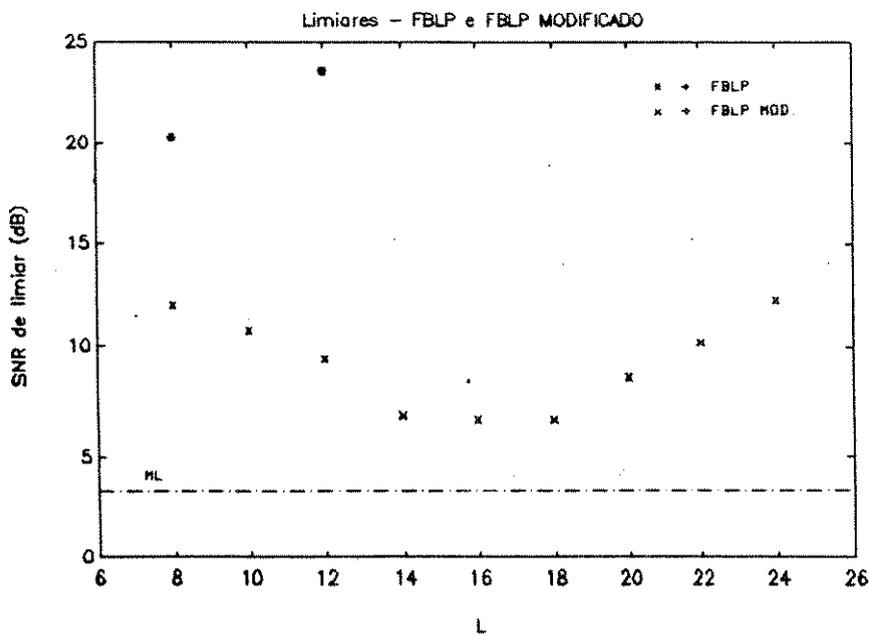


Fig. 4.3 - SNR de limiar x L - ML, FBLP e FBLP Mod.

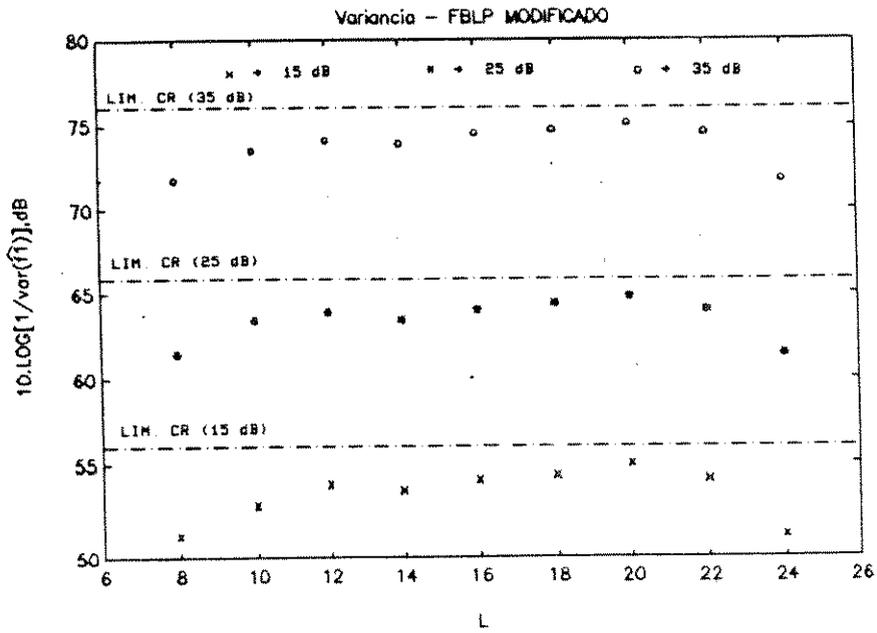


Fig. 4.4 - Variância das estimativas x L - SNR = 15, 25, 35 dB

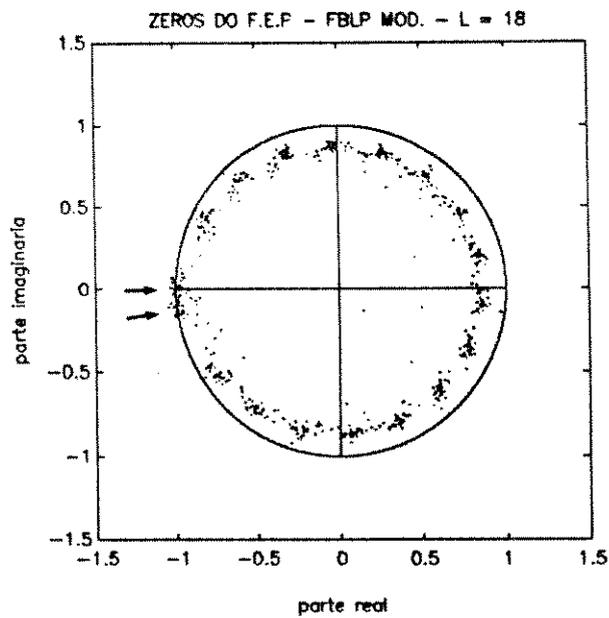


Fig. 4.5 - Método FBLP Mod. - Zeros do F.E.P. - SNR = 0 dB - L = 18 - 30 experimentos

Constatamos que o método FBLP Modificado tem um desempenho bem próximo ao ML, onde o limiar ocorre em cerca de 3 dB, contra 7 dB do FBLP Modificado para $L = 18$. Portanto, o método proposto está apenas a alguns decibéis, em termos de limiar, abaixo do método ML e oferece uma simplicidade computacional bem maior.

4.5 O CASO KUMARESAN-PRONY - KP

Sabemos que, existem limites para a ordem do preditor:

$$M \leq L \leq (N - M/2)$$

Para L fora destes limites, os zeros de $W(z)$ não detectam corretamente as frequências, mesmo no caso sem ruído.

Quando L iguala-se ao seu limite superior: $L = (N - M/2)$, ocorre uma situação bastante interessante. As dimensões da matriz A tornam-se:

$$2.(N - L) \times L$$

isto é:

$$M \times L$$

Como $L \geq M$, o posto de A e, conseqüentemente, o de $\hat{R} = A^H \cdot A$ é $k = M$. Com isto, a matriz de correlação estimada \hat{R} , e sua versão de posto reduzido \tilde{R} são iguais, tornando redundante o primeiro passo para o método FBLP Modificado, apresentado na secção 4.2.

Nesta situação os autovalores-autovetores do subespaço ruidoso de \hat{R} são automaticamente eliminados do cálculo de w , tornando assim os métodos FBLP e FBLP Modificado idênticos. Este fato é de fácil constatação operacional, uma vez que a expressão (3.35) iguala-se à (4.4), e a expressão (3.36) iguala-se à (4.5). Torna-se assim indiferente aplicar o algoritmo da secção 3.4.3 ou o da secção 4.3. Notamos que o caso 2.2 se verifica para ambos algoritmos. Aqui, a SVD de A não deve ser realizada, pois a 2ª opção do item 2.2 da secção 3.4.3 deve ser aplicado, por uma questão de simplicidade computacional. Transcrevemos, por conveniência, a eq. (3.34):

$$w = A^H \cdot (A \cdot A^H)^{-1} \cdot b \quad (4.6)$$

Estamos diante de uma expressão onde apenas produtos matriciais e a inversão de uma matriz $M \times M$ são necessários para o cálculo de w . Como, geralmente, o número de exponenciais M é pequeno, (4.6) torna-se mais eficiente computacionalmente que a utilização da SVD e a aplicação de (4.4) ou (4.5). Este caso é chamado de *Caso Kumaresan-Prony(KP)*.

Dado que os métodos FBLP e FBLP Modificado são idênticos, não é de se estranhar que as figs. 3.4h e 4.1h sejam iguais. Embora nenhum autovalor-autovetor

ruidoso participe do cálculo de w , este valor elevado para L deteriora a estabilidade do método, pois apenas $2.(N - L) = M$ elementos são mediados para formar os elementos de \hat{R} . Porém, utilizamos o preditor com o máximo grau de liberdade possível e resultados bastante satisfatórios ainda são alcançados, como constatamos nas figs. 3.4h, 4.1h e 4.2. Nas figs. 3.4a a 3.4h, observamos que, para valores pequenos da SNR(10 dB), o caso KP não produz picos espectrais espúrios como acontece com outros valores de L . Na fig. 4.2, vemos que, para valores elevados da SNR, o desempenho do caso KP é ligeiramente inferior aos casos FBLP, mas a extensão do limiar é evidente. Também, devido à sua grande simplicidade computacional e desempenho razoável, em algumas aplicações pode ser mais vantajoso aplicar o caso KP, em detrimento, por exemplo, do método FBLP Modificado utilizando $L = 3.N/4$.

4.6 VARIAÇÃO DE ALGUNS PARÂMETROS

Avaliaremos agora o desempenho do método FBLP Modificado quanto à variação de alguns parâmetros importantes na estimativa das frequências.

Inicialmente variaremos o número N de amostras disponíveis. Para o mesmo conjunto de frequências das simulações anteriores e $SNR = 20$ dB, tomamos $N = 15$ e $N = 35$. Em cada situação fizemos $L = (3.N/4)$ e $L = (N - M/2)(KP)$, resultando em $L = 11$ e $L = 14$ para $N = 15$, e $L = 26$ e $L = 34$ para $N = 35$. A fig. 4.6 mostra o resultado das simulações descritas acima bem como para $N = 25$, onde utilizamos $L = 18$ e $L = 24$. O limite CR para cada ponto também é apresentado.

A figura 4.6 permite algumas observações:

- . É claramente perceptível a diminuição da variância absoluta com o crescimento do N , tanto para $L = (3.N/4)$ quanto para o caso KP. Para $L = (3.N/4)$, o crescimento do N possibilita uma maior aproximação dos pontos em relação ao limite CR.
- . A degradação do desempenho provocada pela redução do número de amostras disponíveis é percebida principalmente em $N = 15$ para $L = 14(KP)$, onde a discrepância em relação ao limite CR permite afirmar que o limiar já foi atingido, apesar da SNR ser de 20 dB.

Voltemos agora ao caso $N = 25$, com o objetivo de analisar o desempenho na estimativa de outros valores de frequências. A fig. 4.7 mostra os zeros do filtro de erro de predição para 30 experimentos, $L = 18$, $SNR = 10$ dB e $f_1 = 0.125$ e $f_2 = 0.145$.

A fig. 4.8 mostra a análise comparativa do desempenho do método em estimar as frequências 0.125 e 0.5, para $L = 18$ e $L = 24$.

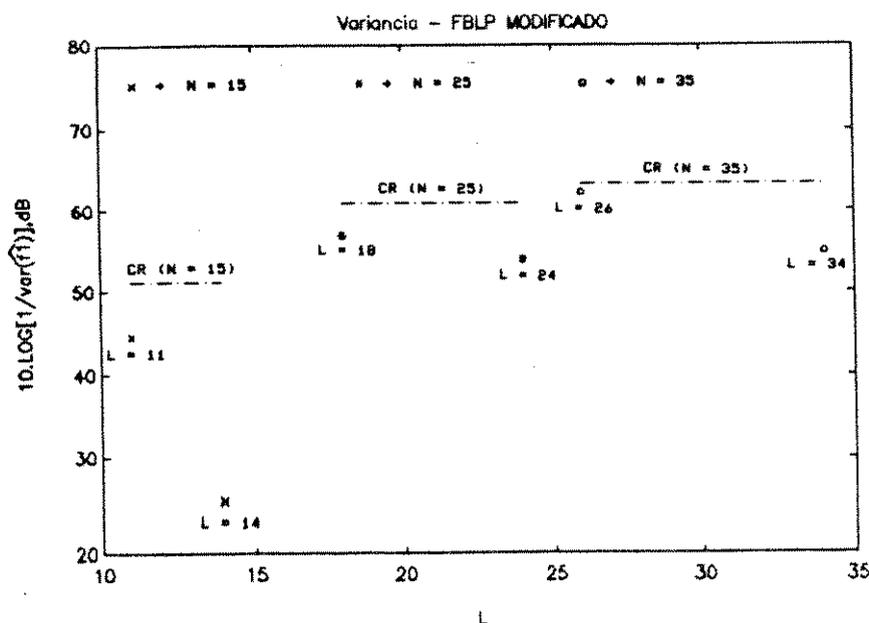


Fig. 4.6 - Variância x L - N = 15, 25, 35 - L = 3.N/4, KP - SNR = 20 dB

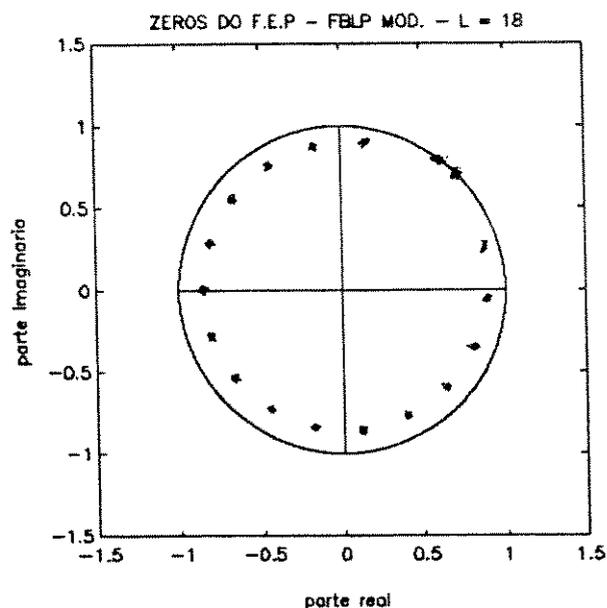


Fig. 4.7 - Método FBLP Mod. - Zeros do F.E.P. - SNR = 10 dB - L = 18 - 30 experimentos - frequências = 0.125, 0.145

Observando a fig. 4.8 percebemos que:

- . Acima dos limiares os desempenhos são idênticos, mostrando que, nesta região, os mesmos independem da localização das frequências a serem estimadas.
- . Os limiares para $f_1 = 0.125$ são alcançados em SNR's um pouco menores do que para

$f_1 = 0.5$. Para explicar este fato é necessário considerar que, em muitas situações, apenas um zero estranho está mais próximo à c.r.u. que um dos zeros de sinal. Portanto, uma das estimativas será definida por um dos zeros de sinal enquanto que a outra será definida pelo zero estranho, sendo esta última completamente equivocada. Como estamos estimando f_1 , caso a estimativa determinada pelo zero estranho esteja situada em um ângulo maior do que a determinada pelo zero de sinal, tomaremos esta última como a estimativa de f_1 . Caso contrário tomaremos a estimativa determinada pelo zero estranho como aquela de f_1 e erros bem maiores serão cometidos. Existem menos zeros estranhos entre a origem do plano Z e $f_1 = 0.125$ do que entre a origem e $f_1 = 0.5$. Logo, sempre que ocorrer a situação descrita acima, haverá uma maior probabilidade da estimativa ser aquela do zero de sinal para $f_1 = 0.125$ do que para $f_1 = 0.5$, o que justifica o limiar em uma SNR menor para o primeiro caso.

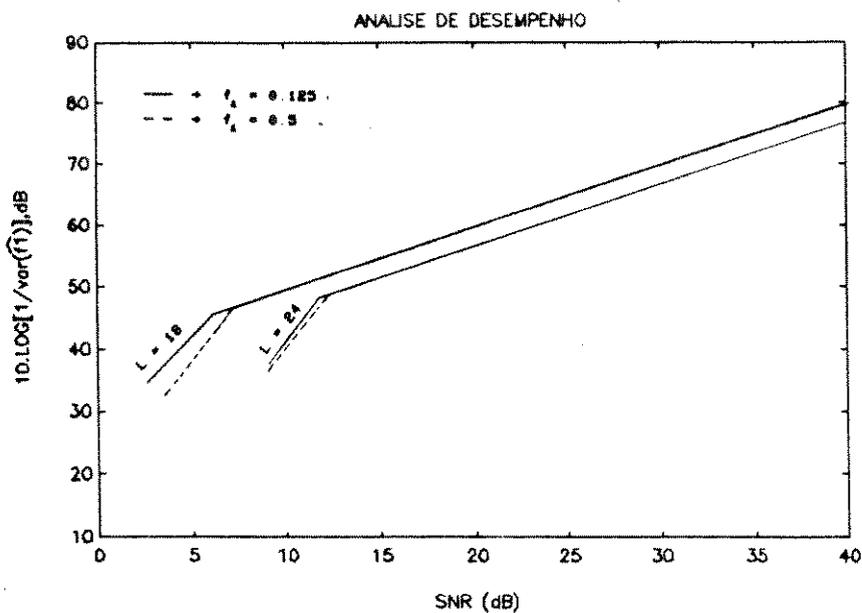


Fig. 4.8 - Análise de desempenho - $f_1 = 0.125, 0.5$

4.7 ESTIMATIVA DO NÚMERO DE EXPONENCIAIS

Até então assumimos que o número de exponenciais M é conhecido, porém esta nem sempre é a situação real. Nestes casos, a princípio, não é possível determinar, com certeza, este parâmetro. Alguns trabalhos, baseados em métodos estatísticos, foram realizados por Schimdt [8], Wax e Kailath [9] e Wax [10].

Uma opção seria observar a magnitude dos valores singulares da matriz A . Para uma relação sinal-ruído não muito pequena, os M valores singulares relacionados ao subespaço de sinal apresentam magnitude bem mais elevada do que os demais.

Portanto, uma queda abrupta na magnitude dos valores singulares deve indicar um valor correto para M . Porém, as magnitudes dos valores singulares sofrem influência de L , N , da SNR e do espaçamento entre as frequências. Pequenos valores para N e SNR, e frequências muito próximas podem tornar as magnitudes próximas entre si, impossibilitando a correta estimativa do número de exponenciais superpostas. Por outro lado, esta homogeneidade das magnitudes dos valores singulares caracteriza uma situação onde existem grandes dificuldades para estimar frequências muito próximas [7].

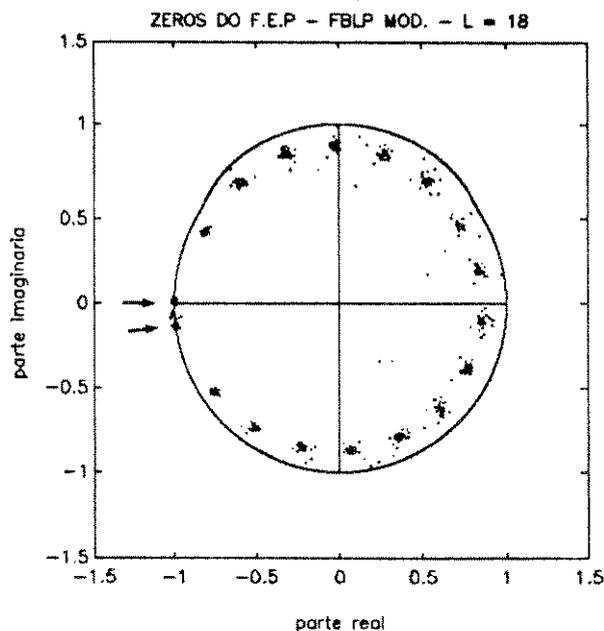


Fig. 4.9 - Método FBLP Mod. - Zeros do F.E.P. - SNR = 10 dB -
 $L = 18$ - 30 experimentos - $M = 3$ (equivocadamente)

A fig. 4.9 mostra o caso em que, equivocadamente estimamos o valor de M como sendo 3 (ao invés de 2). O efeito deste erro é incluir um autovalor-autovetor do subespaço ruidoso no cálculo de w .

A fig. 4.9 deve ser comparada com a 4.1e, onde o valor de M foi estimado corretamente. Realmente, a instabilidade introduzida pelo autovalor-autovetor ruidoso é evidente. Porém, constatamos que, na maioria dos experimentos, os zeros da fig. 4.9 estão bastante próximos de suas posições na fig. 4.1e; apenas eventualmente os mesmos se aproximam perigosamente da c.r.u. de modo a provocar picos espectrais espúrios. Portanto, o método FBLP Modificado operando acima do limiar e com $L = (3.N/4)$, oferece boas chances para a correta identificação do número de exponenciais presentes.

4.8 UMA APLICAÇÃO TÍPICA: DETECÇÃO DE ÂNGULO DE INCIDÊNCIA VERTICAL EM RADAR

Um dos problemas principais em radar é a estimativa dos ângulos de incidência vertical das ondas sobre a antena receptora. Tais ângulos são gerados pela reflexão direta de ondas planas em alvos de interesse, pela presença de multipercurso ou por interferências indesejáveis [22]. Denominaremos tais fatores de fontes da onda plana incidente.

Quando desejamos melhorar a resolução na estimativa de ângulos de incidência, em sistemas de radar tradicionais, aumentamos as dimensões físicas da antena receptora. Alternativamente, podemos utilizar um arranjo linear de antenas com a aplicação de técnicas apropriadas de processamento de sinais. Esta solução torna-se atraente devido ao avanço permanente das técnicas de processamento de sinais bem como à redução dos custos de hardware necessário para implementação de tais técnicas.

Nesta secção estudaremos um sistema de radar onde uma antena transmissora é utilizada para emitir ondas planas para a região onde localizam-se os alvos de interesse e um arranjo linear de antenas é utilizado para analisar o sinal de eco recebido. A partir disto, serão estimados os ângulos de incidência vertical das ondas planas que correspondem às posições das fontes no espaço.

Modelamento do Problema

Consideremos um arranjo linear composto por N antenas, uniformemente espaçadas na direção vertical e com M ondas planas de direções distintas incidindo no arranjo. Assumiremos que todas estas ondas planas são de faixa estreita e com a mesma frequência de portadora f_c . A fig. 4.10 ilustra a configuração do arranjo linear com uma única onda incidindo em uma das antenas, vindo de uma direção definida pelo vetor $-\nu_m$. O módulo de ν_m será o número de onda da m 'ésima onda na direção de propagação, isto é, $|\nu_m| = 1/\lambda$, onde λ é o comprimento de onda.

Vamos considerar o sinal produzido pela m 'ésima onda incidente na n 'ésima antena no instante t , tomando como referência a antena central do arranjo. Da teoria de propagação de ondas planas eletromagnéticas no espaço, este sinal pode ser descrito através da seguinte expressão:

$$s(n,m,t) = A_m \cdot \cos \left[2\pi f_c t + 2\pi \left(n - \frac{N+1}{2} \right) k_m \cdot d + \alpha_m \right] \quad (4.7)$$

- onde:
- $n = 1, 2, \dots, N$
 - $m = 1, 2, \dots, M$
 - f_c = frequência da portadora
 - t = tempo
 - k_m = número de onda da m 'ésima onda plana incidente
 - z = vetor unitário na direção vertical (direção do arranjo)
 - A_m = amplitude de $s(n, m, t)$ medida no centro do arranjo, onde $n = (N+1)/2$
 - α_m = fase de $s(n, m, t)$ medida no centro do arranjo
 - d = espaçamento entre antenas

Seja θ_m o ângulo de incidência da m 'ésima onda plana, medido em relação à reta normal ao arranjo, como ilustrado na fig. 4.10. O número de onda k_m toma então a seguinte forma:

$$k_m = v_m \cdot z = \frac{1}{\lambda} \cdot \text{sen}[\theta_m] \tag{4.8}$$

Consequentemente, (4.7) pode ser reescrita da seguinte forma:

$$s(n, m, t) = A_m \cdot \text{cos} \left[2\pi f_c t + \frac{2\pi d}{\lambda} \left(n - \frac{N+1}{2} \right) \cdot \text{sen}[\theta_m] + \alpha_m \right] \tag{4.9}$$

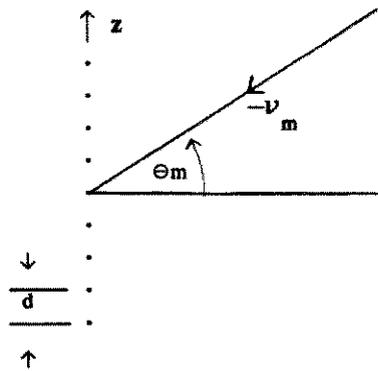


Fig. 4.10 - Arranjo linear de antenas com uma única onda incidindo em uma das antenas

Estaremos interessados nas amostras de $s(n, m, t)$ ao longo do arranjo em um dado instante de tempo. Dado que o instante de tempo influenciará apenas na fase comum a todas as ondas, podemos simplificar a representação de $s(n, m, t)$ adotando a representação para a amplitude complexa deste sinal, definida por:

$$s(n,m) = A_m \cdot \exp\left[j \cdot \frac{2 \cdot \pi \cdot d}{\lambda} \cdot \left(n - \frac{(N+1)}{2}\right) \cdot \text{sen}[\theta_m] + j \cdot \alpha_m \right] \quad (4.10)$$

Definiremos agora o ângulo de fase elétrica ϕ_m entre duas antenas adjacentes do arranjo, como:

$$\phi_m = \frac{2 \cdot \pi \cdot d}{\lambda} \cdot \text{sen}[\theta_m] \quad (4.11a)$$

Se conhecermos ϕ_m poderemos inferir o ângulo de incidência θ_m :

$$\theta_m = \text{sen}^{-1}\left[\frac{\lambda \cdot \phi_m}{2 \cdot \pi \cdot d} \right] \quad (4.11b)$$

Definiremos também a_m , denominado de amplitude complexa do sinal $s(n,m,t)$ medida no centro do arranjo como:

$$a_m = s(n,m) \Big|_{n=(N+1)/2} = A_m \cdot \exp[j \cdot \alpha_m] \quad (4.12)$$

Substituindo (4.11) e (4.12) em (4.10) chegamos a:

$$s(n,m) = a_m \cdot \exp\left[j \cdot \left(n - \frac{(N+1)}{2}\right) \cdot \phi_m \right] \quad (4.13)$$

A equação (4.13) define a amplitude complexa do sinal recebido pela n 'ésima antena devido a m 'ésima onda plana. Somando a contribuição produzida pelo conjunto das M ondas planas, teremos o sinal $s(n)$ recebido na n 'ésima antena devido às M ondas planas.

O sinal **medido** na saída de qualquer antena do arranjo **difere** do sinal **teórico** até aqui tratado por uma componente aditiva produzida por ruído. Na prática, este ruído é modelado como branco, gaussiano, complexo e de média zero. Podemos portanto modelar o sinal medido na saída da n 'ésima antena através da expressão:

$$u(n) = s(n) + y(n) \quad ; n = 1, 2, \dots, N \quad (4.14)$$

onde: $y(n)$ = ruído branco, complexo, gaussiano, de média zero

$$\text{ainda: } u(n) = \sum_{k=1}^M a_k \cdot \exp\left[j \cdot \phi_k \cdot \left(n - \frac{N+1}{2}\right) \right] + y(n) \quad n = 1, 2, \dots, N \quad (4.15)$$

Comparando (4.15) com (3.30), é imediato perceber que as expressões são idênticas do

ponto de vista matemático, exceto pelo deslocamento $(N+1)/2$. A tabela 4.1 resume as relações entre os parâmetros destas equações.

exp. (3.30)	\longleftrightarrow	exp. (4.15)
$a_k = \text{amp. complexa}$	\longleftrightarrow	$a_k = \text{amp. complexa no centro do arranjo}$
$\omega_k = \text{frequência das exponenciais}$	\longleftrightarrow	$\phi_k = \text{ângulo de fase elétrica (relacionado ao ângulo de incidência)}$
$y(n) = \text{ruído branco}$	\longleftrightarrow	$y(n) = \text{ruído branco}$
$M = \text{núm. de exponenciais complexas}$	\longleftrightarrow	$M = \text{núm. de ondas planas incidentes (possíveis alvos)}$
$N = \text{núm. de amostras disponíveis (dim. de tempo)}$	\longleftrightarrow	$N = \text{núm. de antenas do arranjo (dim. de espaço)}$

Tab 4.1 - Relação entre os parâmetros das exps. (3.30) e (4.15)

Portanto, é imediato notar que todo o ferramental do método FBLP Modificado, utilizado para estimar frequências de exponenciais, poderá ser empregado na estimação de ângulo de incidência em radar, caracterizando uma aplicação típica do estudo realizado até então. Isto é constatado ao percebermos a associação entre o ângulo de fase elétrica ϕ_m e a frequência da exponencial ω_k e observarmos a relação direta com o ângulo de incidência dada pela expressão (4.11b).

É importante perceber que a dimensão de tempo n da equação (3.30) dá lugar a uma dimensão de espaço em (4.15), onde a primeira amostra em (3.30) equivale à primeira antena em (4.15), e assim por diante. Em (3.30) temos N amostras de $u(n)$ observadas nos instantes $\Delta t, 2\Delta t, \dots, N\Delta t$; em (4.15) temos os sinais $u(n)$ observados nas antenas $1, 2, \dots, N$ em um determinado instante t_0 , tornando clara a associação entre número de amostras e número de antenas. Se o número de antenas N for ímpar, o elemento central $n = (N+1)/2$ corresponde à antena do centro do arranjo. Por outro lado, quando N for par, o centro do arranjo é um ponto fictício.

A Transformada de Fourier

Consideremos o sinal $u(n)$, $n = 1, 2, \dots, N$, gerado na forma descrita pela equação (4.15). Por analogia com a análise de Fourier de sequências temporais, definimos o espectro de ângulo desta sequência espacial como:

$$U(\phi) = \sum_{n=1}^N u(n) \cdot \exp[-j(n - \frac{N+1}{2}) \cdot \phi] \quad (4.16)$$

Analogamente ao caso temporal, dado o espectro de ângulo $U(\phi)$, podemos estimar as direções das fontes das ondas planas calculando os picos espectrais, isto é, os valores de ϕ para os quais a magnitude do espectro de ângulo $U(\phi)$ alcança um máximo.

Consideremos inicialmente o caso de uma única onda plana incidente no arranjo, na ausência de ruído, dado pela eq. (4.13).

$$s(n,m) = a_m \cdot \exp[j \cdot (n - \frac{N+1}{2}) \cdot \phi_m] \quad n = 1, 2, \dots, N$$

Obtendo o espectro de ângulo deste sinal temos:

$$S_m(\phi) = a_m \cdot \sum_{n=1}^N \exp[-j \cdot (n - \frac{N+1}{2}) \cdot (\phi - \phi_m)] \quad (4.17)$$

$$S_m(\phi) = a_m \cdot W(\phi - \phi_m) \quad (4.18)$$

onde:
$$W(\phi) = \frac{\text{sen}[N \cdot \phi/2]}{\text{sen}[\phi/2]} \quad (4.19)$$

A fig. 4.11 mostra o gráfico de $W(\phi)$ em função do ângulo de fase elétrica ϕ , evidenciando o espectro, bastante conhecido, associado à janela retangular.

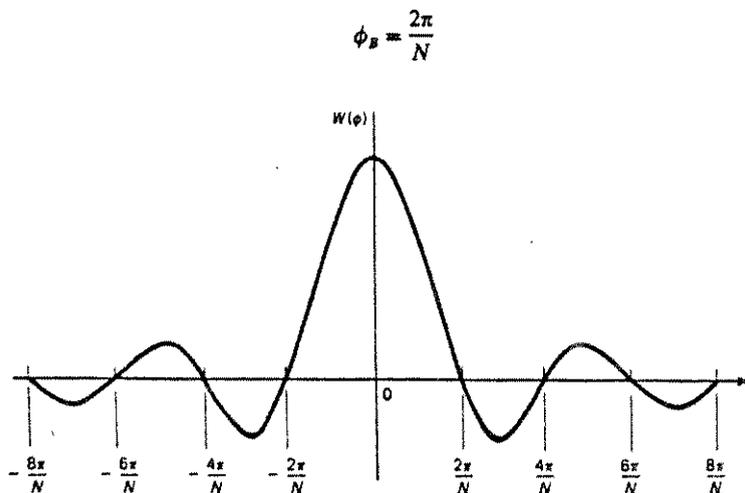
O gráfico de $W(\phi)$ apresenta um lóbulo principal e diversos lóbulos laterais. O primeiro nulo ocorre no ângulo:

$$\phi_B = \frac{2 \cdot \pi}{N}$$

ϕ_B é então a separação angular entre o pico do lóbulo principal e o primeiro nulo no espectro $W(\phi)$, e é denominado de largura de feixe padrão.

Quando M fontes estão presentes em diferentes pontos do espaço, podemos utilizar o princípio da superposição, extendendo o resultado de (4.18), e obtendo o espectro total na seguinte forma:

$$S(\phi) = \sum_{m=1}^M a_m \cdot W(\phi - \phi_m) \quad (4.20)$$

Fig. 4.11 - $W(\phi) \times \phi$ (ângulo de fase elétrica)

A observação de (4.20) e da fig. 4.11 permite afirmar que quando as fontes estão separadas por ϕ_B ou mais, a superposição dos espectros de $W(\phi - \phi_m)$ de (4.20) se dá de forma tal que não ocorre interferência significativa entre seus lóbulos principais [2]. Logo, a magnitude do espectro de ângulo apresentará picos distintos, cada qual localizado no ângulo que determina a direção da fonte correspondente ou nas vizinhanças do mesmo. Caso duas fontes estejam separadas por um ângulo menor que ϕ_B , ocorre interferência entre lóbulos principais associados e a análise baseada em $S(\phi)$ não permite resolver estas duas fontes, isto é, não permite sequer distinguir a presença de duas fontes. A interferência dará origem a um único pico indicando uma única fonte.

Com isto concluímos que quanto menor o ϕ_B maior será a capacidade em resolver fontes que estejam próximas entre si. Para diminuir o ϕ_B é necessário aumentar o número N de antenas do arranjo.

Outro problema a enfrentar na atual conjuntura é a magnitude dos lóbulos laterais. Os mesmos interferem sobre os lóbulos principais das demais fontes, provocando desvios no posicionamento dos picos espectrais. Além disto, sua combinação com os demais lóbulos laterais pode levar à indicação de fontes falsas. É sabido que [2] para diminuir o efeito (magnitude) destes lóbulos seria necessário realizarmos um janelamento espacial dos dados na saída do arranjo. Isto equivale a multiplicarmos o sinal na saída de cada antena por um ganho distinto, o qual até o presente momento foi constante e igual a 1, mostrando a presença implícita de uma janela retangular. Existe uma infinidade de janelas disponíveis na literatura [2]

para a realização de tal propósito. A utilização deste janelamento porém, tem o efeito negativo de alargar o lóbulo principal, comprometendo desta forma a resolução.

Todas as observações feitas até aqui sobre o espectro de ângulo já são bem conhecidas da teoria de estimação do espectro de frequência [2]. Nosso propósito foi, com base na tabela 4.1, mostrar a analogia perfeita que existe entre estimação do espectro de frequência e estimação do ângulo de incidência vertical em radar (estimação do espectro de ângulo), procurando evidenciar que todos os resultados conhecidos para o domínio da frequência, são aplicáveis diretamente ao domínio do ângulo em radar.

Pelo já exposto, fica evidente que métodos baseados no espectro de ângulo não são capazes de resolver fontes separadas por um ângulo menor que ϕ_B , sendo plenamente justificável a aplicação do método FBLP Modificado para alcançar tal objetivo.

4.9 CONCLUSÕES

Os métodos baseados em predição linear para estimar frequências de exponenciais complexas em meio a ruído branco, tais como o FLP, BLP e FBLP, apresentam um desempenho pobre para valores pequenos da relação sinal-ruído.

O método FBLP Modificado consegue atingir desempenhos significativamente superiores ao empregar uma aproximação de posto reduzido para a matriz de correlação estimada, baseando-se na decomposição da mesma em autovalores e autovetores.

Duas opções foram apresentadas a fim de suavizar a carga computacional do método FBLP Modificado. Através delas sempre calculamos autovalores-autovetores e efetuamos produtos matriciais com as menores dimensões possíveis. Uma boa alternativa seria utilizar grandes valores para L (próximos a $(N - M/2)$) e calcular w por (4.5), assegurando resultados satisfatórios com uma grande simplicidade computacional. O caso KP é o mais simples computacionalmente, principalmente quando o número de exponenciais M é pequeno. Sua utilização é aconselhável sempre que seu desempenho for satisfatório para uma determinada aplicação.

O desempenho do método sofre forte influência de alguns parâmetros, tais como número de amostras disponíveis, ordem do preditor e SNR, enquanto que a influência provocada pelo posicionamento das frequências é pequena.

Através de simulações, demonstramos que o desempenho do método FBLP Modificado para detecção de frequências, praticamente se iguala ao limite superior CR para um grande intervalo de valores da SNR. Também, o efeito de limiar ocorre

para valores muito pequenos da SNR. Tudo isto habilita o método FBLP Modificado como um substituto viável ao método ML para detecção de frequências.

Detecção de ângulo de incidência vertical em radar possui uma analogia perfeita com detecção de frequências de senóides, constituindo-se em uma aplicação típica do Método FBLP Modificado.

CAPÍTULO 5

REDUZINDO A SNR DE LIMIAR

CONTEÚDO	Pag.
5.1 Introdução.....	105
5.2 Uma Tentativa para Diminuir a SNR de Limiar.....	106
5.3 Detecção de Senóides Reais.....	111
5.3.1 Os Zeros do Filtro de Erro de Predição.....	111
5.3.2 Um Algoritmo para Detectar Senóides Reais.....	114
5.3.3 Comparações Quantitativas.....	119
5.3.4 Análise de Desempenho.....	120
5.4 Conclusões.....	122

5.1 INTRODUÇÃO

Neste capítulo, apresentaremos os resultados de alguns estudos especulativos que objetivaram desenvolver uma estratégia suplementar para uma melhor distinção entre os zeros de sinal e os estranhos do filtro de erro de predição (f.e.p.).

Esta estratégia se baseia em um pré-processamento nas amostras disponíveis com a finalidade de introduzir um fator de amortecimento positivo nas exponenciais complexas. Sobre este sinal pré-processado aplicaremos o método FLP Modificado, objetivando uma maior probabilidade em distinguir corretamente os zeros de sinal dos zeros estranhos. Em seguida aplicaremos o método FBLP Modificado, tendo porém o conhecimento da localização aproximada dos zeros de sinal, fornecida pelo procedimento descrito acima. Realizaremos simulações exemplificando tal procedimento, assim como um estudo mostrando a influência de diversos parâmetros em tal método.

Após isto, enfocaremos de forma especial o caso de detecção de senóides reais. Apresentaremos as características dos zeros do f.e.p. de tal caso, mostrando que para cada zero, seja de sinal ou estranho, existe um correspondente zero complexo conjugado.

Baseando-nos na diferença de comportamento dos zeros de sinal e dos zeros estranhos quanto à proximidade da c.r.u., proporemos um algoritmo para detectar senóides reais. Este é capaz de realizar uma distinção mais robusta entre os zeros de sinal e os zeros estranhos e, portanto, reduzir a SNR de limiar alcançada pelo método FBLP Modificado. Consegue também diminuir a variância das estimativas em quaisquer SNR através de uma operação de mediação. Comparações quantitativas mostrarão o nível de desempenho propiciado pelo novo algoritmo. Analisaremos também este desempenho quanto à variação da ordem L do preditor e da SNR.

5.2 UMA TENTATIVA PARA DIMINUIR A SNR DE LIMIAR

É conhecido que, em situação ideal, os zeros de sinal localizam-se sobre a c.r.u., nas frequências das senóides, enquanto que os zeros estranhos permanecem no interior do círculo de raio unitário. Eventualmente, em SNR's relativamente baixas, por motivos já discutidos em capítulos anteriores, algum (ou alguns) zero estranho posiciona-se mais próximo à c.r.u. do que os próprios zeros de sinal. Neste caso, o zero estranho é tomado com a estimativa de uma das frequências e o limiar é alcançado.

Caso houvesse uma forma alternativa de distinguir os zeros de sinal dos zeros estranhos, mesmo os zeros de sinal que não estivessem tão próximos à c.r.u. forneceriam melhores estimativas do que zeros estranhos espúrios muito próximos à c.r.u., e o limiar deixaria de existir.

Se fosse possível realizar um pré-processamento nas amostras disponíveis do sinal em conjunto com a posterior utilização de um dos métodos de predição linear, de modo a expulsar os zeros de sinal para o exterior do círculo de raio unitário, mantendo porém a posição dos zeros estranhos, aumentaríamos a capacidade de distinção. A partir disto, ao identificar os zeros que estivessem no exterior do círculo de raio unitário, definiríamos regiões preferenciais na c.r.u. onde as frequências a estimar (os zeros de sinal) possivelmente se encontrariam. Aplicaríamos então o método FBLP Modificado sobre as amostras originais do sinal buscando os zeros mais próximos à c.r.u. somente nas regiões definidas no passo anterior. Com isto, estaríamos eliminando o risco de tomar como estimativa das frequências, zeros estranhos muito próximos à c.r.u. que estivessem fora das tais regiões.

Do teorema 3.1 sabemos que no caso sem ruído e com exponenciais que apresentem fator de amortecimento positivo, a predição linear forward posiciona os zeros de sinal no exterior do círculo de raio unitário, como desejamos. Na presença de ruído a aplicação do método FLP Modificado produz este comportamento de forma aproximada. Partindo de exponenciais não-amortecidas, o seguinte procedimento torna-se, portanto, necessário:

. Gerar um sinal $u'(n)$ a partir das amostras disponíveis de $u(n)$, da seguinte forma:

$$u(n) = \sum_{k=1}^M a_k \cdot \exp[j\omega_k \cdot n] + y(n) \quad (5.1)$$

$$u'(n) = \exp[\alpha \cdot n] \cdot [u(n)] \quad ; \quad \alpha > 0 \quad (5.2)$$

$$u'(n) = \sum_{k=1}^M a_k \cdot \exp[\alpha \cdot n + j\omega_k \cdot n] + \exp[\alpha \cdot n] \cdot y(n) \quad (5.3)$$

$$u'(n) = \sum_{k=1}^M a_k \cdot \exp[(\alpha + j\omega_k) \cdot n] + y'(n) \quad (5.4)$$

Agora todas as exponenciais apresentam um fator de amortecimento α . Daremos prosseguimento à proposta de extender o limiar através dos seguintes passos:

- . Aplicar o método FLP Modificado sobre $u'(n)$.
- . Definir as regiões do plano Z onde os zeros estejam mais afastados do centro do círculo de raio unitário.
- . Aplicar o método FBLP Modificado sobre $u(n)$ seleccionando os M zeros mais próximos à c.r.u. somente nas regiões definidas no ítem anterior.

Mostraremos agora um caso onde um zero estranho produzido pelo método FBLP Modificado se encontra mais próximo à c.r.u. do que um dos M zeros de sinal, levando a uma estimativa desastrosa das frequências. Porém, a utilização do procedimento descrito acima possibilita que os zeros de sinal determinem as estimativas das frequências.

Utilizamos o mesmo sinal das secções 3.4.3 e 4.3. A fig. 5.1 mostra um experimento particular resultante da aplicação do caso KP com uma SNR = 10 dB. O zero estranho muito próximo à c.r.u. está indicado por uma seta. A aplicação do método FLP Modificado com $L = 10$ e $\alpha = 0.2$ gera os dois zeros mais afastados do centro do círculo de raio unitário em posições vizinhas. Logo, é sensato admitir que

trata-se de duas frequências próximas entre si, definindo apenas a região indicada na figura para os zeros de sinal. A definição desta região permite excluir o zero estranho das estimativas fornecidas pelo caso KP. Com isto torna-se possível identificar os verdadeiros zeros de sinal, fornecidos por este caso e utilizá-los para determinar as duas frequências.

Este procedimento apresenta uma complexidade computacional adicional, pois além de aplicarmos o método FBLP Modificado, ainda realizamos um pré-processamento nas amostras do sinal bem como aplicamos o método FLP Modificado. Esta desvantagem pode ser minorada pela utilização do caso KP.

Infelizmente, entretanto, este procedimento não se mostra eficiente em propiciar uma redução da SNR de limiar. É certo que para alguns experimentos, como aquele da fig. 5.1, obtivemos sucesso. Porém, em outros são definidas faixas de frequência equivocadas, não só em experimentos similares ao da fig. 5.1, mas também em alguns onde o FBLP Modificado já indicava os verdadeiros zeros de sinal como a estimativa das frequências, transformando assim estimativas de boa qualidade em desastrosas.

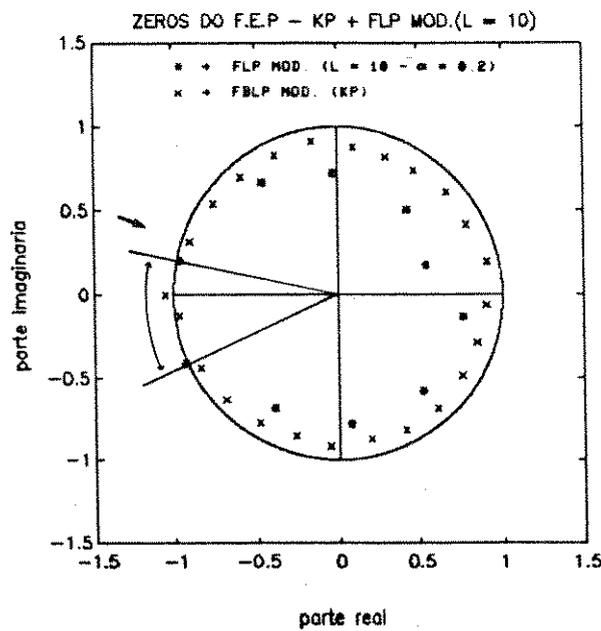


Fig. 5.1 - KP + FLP Mod.(L = 10, $\alpha = 0.2$) - Zeros do F.E.P. - SNR = 10 dB

Identificaremos agora algumas causas que possivelmente contribuem para tal insucesso:

1) O desempenho do método FLP Modificado é inferior ao do FBLP Modificado. Assim, mesmo para SNR's relativamente altas, são definidas faixas equivocadas.

2) O ruído $y'(n)$, da eq. (5.4), não é mais branco. Já que as características geométricas dos zeros do f.e.p. baseiam-se na propriedade branca do ruído, o desempenho do método FLP Modificado aplicado sobre $u'(n)$ deteriora-se consideravelmente.

3) O ideal seria que o fator de amortecimento α fosse o maior possível, pois quanto maior este parâmetro mais distante do círculo de raio unitário, em direção ao seu exterior, se localizariam os zeros de sinal, aumentando a probabilidade das faixas serem selecionadas corretamente. Porém, quanto maior o α maior será o grau de correlação das amostras do ruído, deteriorando bastante o desempenho do método. A fig. 5.2 ilustra a relação do α com o desempenho. A fig. 5.2a mostra o mesmo que a fig. 5.1 para o método FLP Modificado, utilizando $\alpha = 0.1$ em 30 experimentos. A fig. 5.2b utiliza $\alpha = 0.3$.

A desordem observada nos zeros para $\alpha = 0.3$ é claramente perceptível.

Portanto, quando o método FBLP Modificado atingir o limiar, o FLP Modificado também o terá atingido, definindo faixas equivocadas para os zeros de sinal. A fig 5.3 ilustra esta situação mostrando 50 experimentos do método FLP Modificado aplicado sobre $u'(n)$, onde utilizamos $\alpha = 0.2$ e $L = 10$ em uma SNR = 5 dB. Em tal SNR, o método FBLP Modificado está abaixo do limiar mesmo para $L = 18(3.N/4)$. A aplicação do método FLP Modificado neste caso não conseguirá estender o limiar, pois na maioria das vezes serão definidas faixas equivocadas.

Testamos ainda a utilização do método BLP Modificado empregando $\alpha < 0$, cujo comportamento é similar ao FLP Modificado com $\alpha > 0$. Variamos também o valor da ordem L do preditor e do fator de amortecimento α , objetivando otimizar o desempenho do método FLP Modificado. Porém, não se conseguiu atingir um desempenho que assegurasse a redução da SNR de limiar de forma robusta e a estratégia foi abandonada.

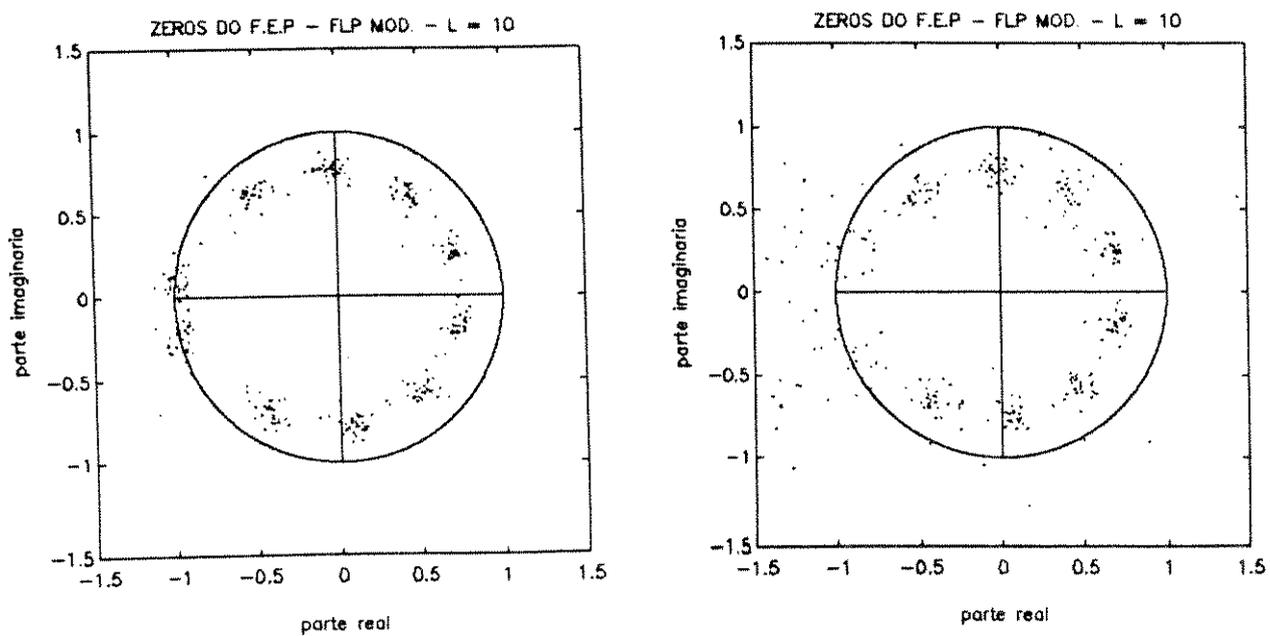


Fig. 5.2 - FLP Mod.(L = 10) - SNR = 10 dB - 30 experimentos
 (a) $\alpha = 0.1$ (b) $\alpha = 0.3$

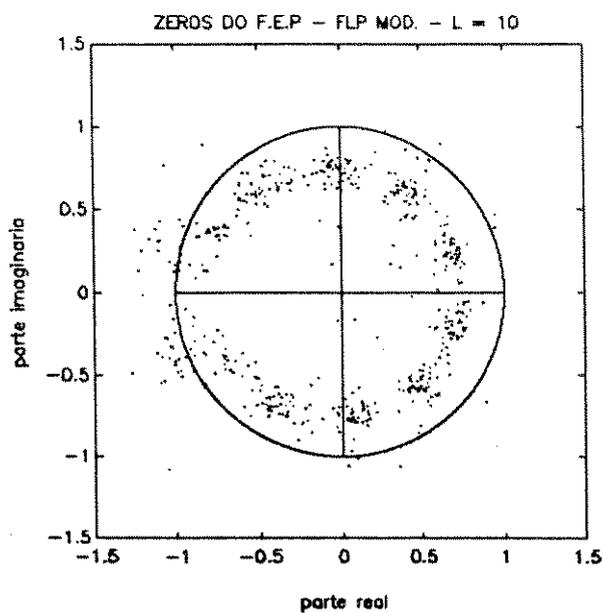


Fig. 5.3 - FLP Mod.(L = 10, $\alpha = 0.2$) -
 SNR = 5 dB - 50 experimentos

5.3 DETECÇÃO DE SENÓIDES REAIS

5.3.1 OS ZEROS DO FILTRO DE ERRO DE PREDIÇÃO

Senóides reais constituem um tipo particular do sinal até então em questão, onde as frequências das exponenciais ocorrem em pares complexos conjugados. Neste caso teremos sempre M par e apenas $M/2$ senóides reais estarão presentes, resultando em um sinal na seguinte forma:

$$u(n) = a_1 \cdot \exp[j \cdot \omega_1 \cdot n] + a_1 \cdot \exp[-j \cdot \omega_1 \cdot n] + \\ a_2 \cdot \exp[j \cdot \omega_2 \cdot n] + a_2 \cdot \exp[-j \cdot \omega_2 \cdot n] + \dots + \\ a_{M/2} \cdot \exp[j \cdot \omega_{M/2} \cdot n] + a_{M/2} \cdot \exp[-j \cdot \omega_{M/2} \cdot n] + y(n) \\ n = 1, 2, \dots, N$$

Objetivando diminuir a SNR de limiar, poderíamos explorar a natureza complexa conjugada dos zeros de sinal para distingui-los dos zeros estranhos. Isto porém não é possível, pois os zeros estranhos também localizam-se em posições complexas conjugadas. Tal comportamento pode ser observado na fig. 5.4, onde utilizamos um sinal do tipo:

$$u(n) = \exp[j \cdot \pi/4 \cdot n] + \exp[-j \cdot \pi/4 \cdot n] + y(n) \quad n = 1, 2, \dots, 25 \quad (5.5)$$

e o método FBLP Modificado foi empregado com $L = 20$ e $SNR = 10$ dB em dez experimentos distintos.

Imaginando uma ação que rompesse a simetria complexa dos zeros estranhos, poderíamos inserir no sinal $u(n)$ em (5.5), duas exponenciais espúrias situadas nos 3º e 4º quadrantes do plano Z. A princípio, tal procedimento implicaria em um aumento no número M de exponenciais, ocasionando uma degradação no desempenho do método FBLP Modificado. Além disso, próximo às frequências espúrias, existiriam faixas nas quais os verdadeiros zeros de sinal sofreriam forte influência daqueles espúrios inseridos artificialmente.

Através de simulações constatamos também que a influência das frequências espúrias só é sentida próximo às mesmas. À medida que nos afastamos delas, os zeros estranhos tendem novamente a ser complexos conjugados. A fig. 5.5 mostra o mesmo da fig. 5.4 com a inserção de duas frequências espúrias, simétricas em torno de $3 \cdot \pi/2$, para $L = 18$, exemplificando o comportamento descrito acima.

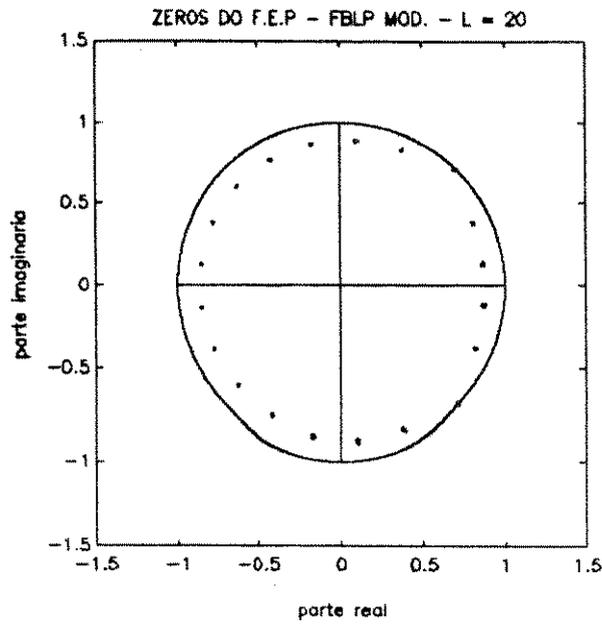


Fig. 5.4 - Método FBLP Mod. - Zeros do F.E.P. - SNR = 10 dB -
 L = 20 - 10 experimentos - $\omega_1 = \pi/4$, $\omega_2 = -\pi/4$

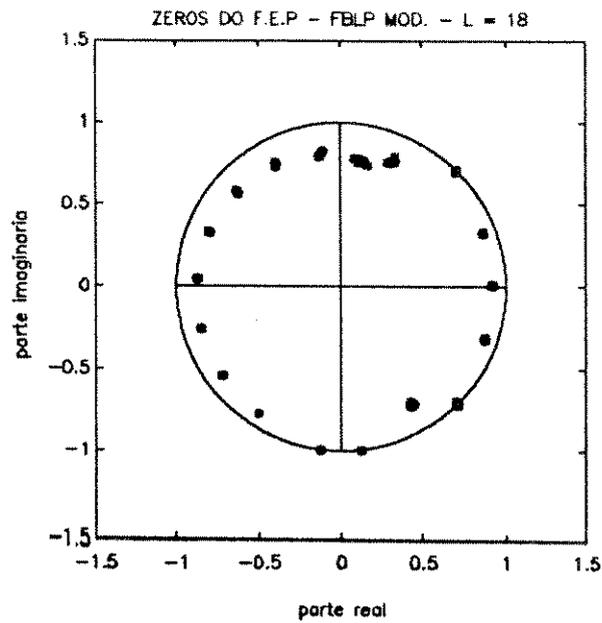


Fig. 5.5 - Método FBLP Mod. - Zeros do F.E.P. - SNR = 10 dB -
 L = 18 - 10 experimentos - 2 frequências espúrias

Como uma nova alternativa, poderíamos inserir aproximadamente $[(L - M)/2]$ zeros espúrios nos 3º e 4º quadrantes, de forma que os zeros estranhos em tais quadrantes seriam substituídos por estes zeros espúrios e, portanto, ocupariam posições pré-determinadas. Porém, tomando como base as discussões anteriores, dois problemas deverão surgir:

- 1) Aumento significativo no número de exponenciais a serem detectadas, comprometendo assim a capacidade do método em distinguir duas frequências muito próximas.
- 2) Aumento no número de faixas de frequências onde a presença dos verdadeiros zeros de sinal seria problemática.

A fim de caracterizar tais limitações, geramos a fig. 5.6 utilizando um sinal similar ao da expressão (5.5), porém com a frequência da senóide real em 0.01. Foram inseridas frequências espúrias normalizadas correspondentes aos ângulos -30° , -60° , -90° , -120° e -150° , com $L = 18$ e $SNR = 10$ dB para dois experimentos. A posição correta dos zeros de sinal estão indicadas por setas.

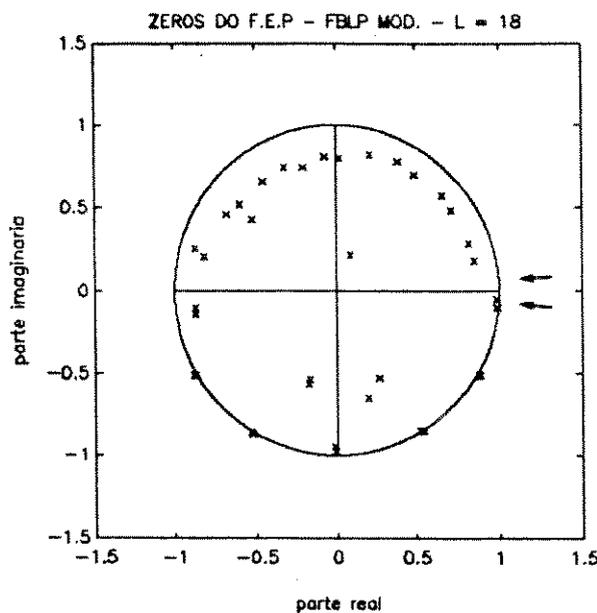


Fig. 5.6 - Método FBLP Mod. - Zeros do F.E.P. - SNR = 10 dB - L = 18 - 2 experimentos - 5 frequências espúrias - freq = 0.01

Observando a fig. 5.6, percebemos que realmente os zeros estranhos não mais se encontram em posições complexas conjugadas, porém um dos zeros de sinal também não se encontra próximo à c.r.u., ratificando o problema caracterizado no ítem 1 acima.

A fig. 5.7 é a mesma simulação da fig. 5.6, porém com as frequências espúrias localizadas em -25° , -65° , -115° e -155° . Para este posicionamento particular dos zeros de sinal e dos zeros espúrios, em alguns experimentos, o método alcança bons resultados. Neste caso, os zeros de sinal posicionam-se mais próximos à c.r.u., possibilitando estimativas mais precisas para as frequências.

Porém, simulações demonstraram que o sucesso é condicionado pelo posicionamento dos zeros. Concluímos então que tal estratégia não oferece robustez para distinguir zeros de sinal de zeros estranhos de senóides reais.

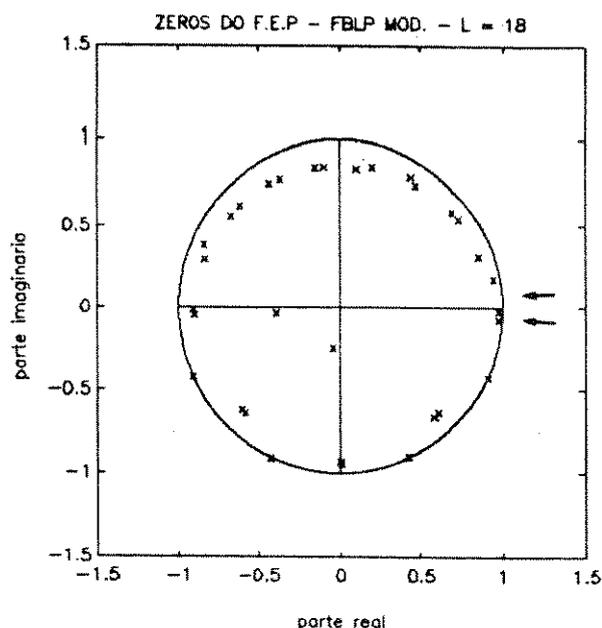


Fig. 5.7 - Método FBLP Mod. - Zeros do F.E.P. - SNR = 10 dB - L = 18 - 2 experimentos - 4 frequências espúrias - freq = 0.01

5.3.2 UM ALGORITMO PARA DETECTAR SENÓIDES REAIS

As inúmeras simulações realizadas ao longo do estudo anterior permitiram observar que no caso de $M/2$ senóides reais com SNR's pequenas, o deslocamento de um zero estranho para posições muito próximas à c.r.u. é completamente aleatório. Portanto o seu complexo conjugado não necessariamente se encontrará também muito próximo à c.r.u., ao contrário dos zeros de sinal, onde cada elemento do par complexo conjugado tende a manter a proximidade em relação à c.r.u.. A partir desta diferença de comportamento, pudemos divisar uma nova forma de interpretar os zeros do f.e.p. do método FBLP Modificado, objetivando a distinção entre os zeros de sinal e os zeros estranhos. Este procedimento novo leva em consideração não só a proximidade dos zeros em relação à c.r.u. mas também a natureza complexa conjugada dos zeros de sinal.

O Algoritmo

A fim de analisarmos a proximidade dos zeros em relação à c.r.u., classificamos os mesmos partindo do mais próximo para o mais afastado. Ao invés de simplesmente selecionar os M zeros mais próximos à c.r.u., proporemos agora um procedimento alternativo.

Imaginemos que apenas um zero estranho esteja entre os M zeros mais próximos, pois como discutido acima, o seu complexo conjugado não necessariamente estará tão próximo à c.r.u.. Portanto, os (M + 1) zeros mais próximos à c.r.u. serão constituídos por M/2 pares de zeros de sinal complexos conjugados e um zero estranho, já que os zeros de sinal tendem a manter a mesma aproximação. Levando em conta esta diferença de comportamento, selecionaremos os M/2 pares de zeros complexos conjugados mais próximos à c.r.u.. Este procedimento, na situação descrita acima, eliminará o zero estranho muito próximo à c.r.u. das possíveis estimativas. É óbvio que ao procurarmos pelos pares, os mesmos não estarão em posições exatamente complexas conjugadas, sendo necessário considerar um ângulo de tolerância, o qual denominaremos de $\Delta\theta$.

Após encontrarmos os M/2 primeiros pares complexos conjugados, não nos limitamos a tomar um ou outro elemento de um determinado par como estimativa de uma das frequências das senóides reais. Ao contrário, levamos em consideração os dois zeros tomando como estimativa a média aritmética entre o ângulo de um dos zeros e o ângulo do conjugado do outro. Isto possibilitou uma redução na variância das estimativas.

O algoritmo pode ser formalizado através dos seguintes passos:

- 1) Obter os zeros do f.e.p. fornecidos pelo método FBLP Modificado.
- 2) Classificar os L zeros do f.e.p. na ordem crescente em termos da distância em relação à c.r.u.. Tomando o i-ésimo zero desta classificação e definindo como $z(i)$ sua posição no plano Z, construímos o vetor z cujas componentes são os $z(i)$ na ordem ditada pela classificação.
- 3) Pesquisar, a partir do início de z , até encontrar M/2 situações nas quais:

$$\text{ângulo}[\overline{z(j)}] - \Delta\theta \leq \text{ângulo}[z(i)] \leq \text{ângulo}[\overline{z(j)}] + \Delta\theta$$

$$i = 2, 3, \dots, L \quad \text{e} \quad j = 1, 2, \dots, i-1$$

com isto, as M/2 frequências das senóides reais estarão estimadas.

4) Supondo um par selecionado $z(i_1)$ e $z(j_1)$, então a estimativa \hat{f}_1 de uma das frequências será:

$$\hat{f}_1 = \frac{1}{2 \cdot \pi} \cdot \frac{\text{ângulo}[z(i_1)] + \text{ângulo}[\overline{z(j_1)}]}{2}$$

Através dos passos 1, 2 e 3 encontraremos os pares complexos conjugados mais próximos à c.r.u. e um ou mais zeros estranhos isolados muito próximos são eliminados. Com isto conseguimos reduzir a SNR onde ocorre o fenômeno de limiar. O passo 4 possibilita menores variâncias para SNR's quaisquer.

O $\Delta\theta$ é utilizado para permitir uma tolerância no posicionamento dos pares complexos conjugados. Não deve ser muito pequeno pois pares complexos conjugados poderiam ser ignorados. $\Delta\theta$ muito grande também nos causaria problemas, pois poderíamos ser levados a selecionar pares complexos conjugados equivocadamente. Utilizando simulações e considerações acerca da geometria dos zeros do f.e.p., chegamos à conclusão que $\Delta\theta = \pi/L$ constitui uma excelente escolha. A fim de discutir tal escolha, apresentamos na fig. 5.8, 30 experimentos para duas senóides reais nas frequências normalizadas de $f_1 = 0.125$ e $f_2 = 0.145$, em uma SNR = 10 dB e $L = 18$ para o método FBLP Modificado. O número de amostras disponíveis para cada experimento é $N = 25$.

A priori, $\Delta\theta = \pi/L$ pode parecer elevado. Porém, através da fig. 5.8 vemos que entre os grupos de zeros estranhos existe uma separação angular satisfatória, a qual garante que o emprego de $\Delta\theta = \pi/L$ não levará a equívocos quando da associação entre um grupo de zeros estranhos e o seu verdadeiro grupo complexo conjugado. O mesmo ocorrerá em relação aos grupos de zeros de sinal, pois existem lacunas consideráveis em torno dos mesmos.

A fim de ilustrar a atuação do algoritmo, a fig. 5.9 mostra um experimento utilizando os mesmos parâmetros da fig. 5.8 porém a uma SNR de 6 dB. O zero estranho indicado por uma seta está mais próximo à c.r.u. do que um dos 4 zeros de sinal e provoca um erro grosseiro na estimação da menor frequência.

Os numerais ao lado dos zeros indicam suas posições no vetor \mathbf{z} do algoritmo, ou seja, o número 1 está mais próximo que o 2, e assim por diante. A aplicação do algoritmo elimina tal zero estranho como uma das estimativas das frequências, pois os pares de zeros complexos conjugados mais próximos à c.r.u. são 1-2 e 3-5, dado que o complexo conjugado do zero estranho 4 ocupa apenas a 16^a posição no vetor \mathbf{z} . Assim, se antes, este experimento contribuía para que 6 dB estivesse abaixo do limiar, agora não mais contribui.

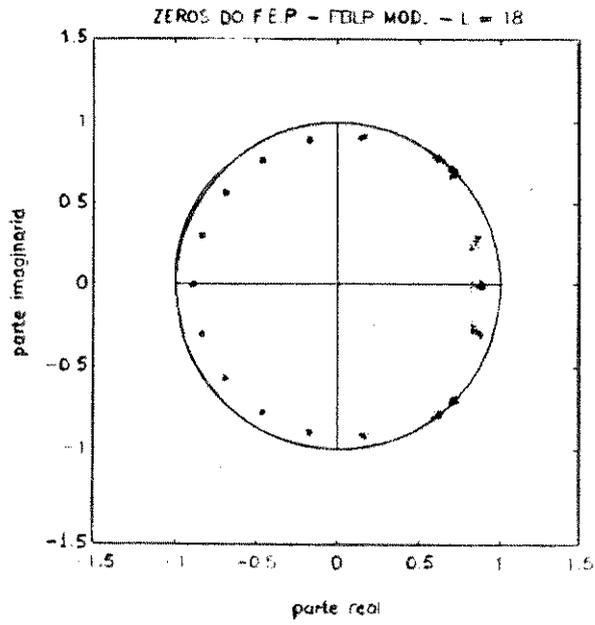


Fig. 5.8 - Método FBLP Mod. - Zeros do F.E.P. - SNR = 10 dB - L = 18 - 30 experimentos - Sen. reais em 0.125 e 0.145

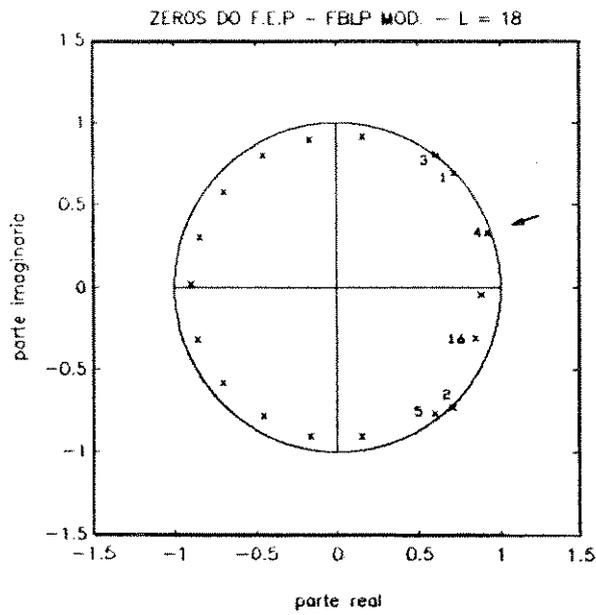


Fig. 5.9 - Método FBLP Mod. - Zeros do F.E.P. - SNR = 6 dB - L = 18 - Sen. reais em 0.125 e 0.145

Uma situação de interesse especial é a mostrada na fig. 5.10, que mostra um outro experimento semelhante àquele descrito na fig. 5.9, porém para uma SNR de 4 dB.

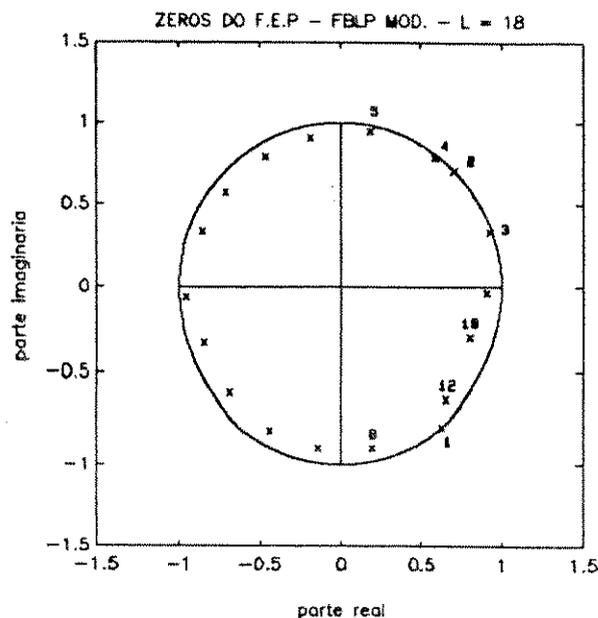


Fig. 5.10 - Método FBLP Mod. - Zeros do F.E.P. - SNR = 4 dB -
L = 18 - Sen. reais em 0.125 e 0.145

Devido à proximidade entre os zeros de sinal e à existência do ângulo de tolerância $\Delta\theta$, o algoritmo toma como os dois primeiros pares complexos conjugados: 1-2 e 1-4. Neste caso particular, onde um único zero, o 1 no caso, participa de mais de um par complexo conjugado, o algoritmo toma o seguinte procedimento:

- a) Verifica qual dos dois pares tem uma simetria complexa conjugada mais perfeita.
 - Neste caso, o par 1-4 satisfaz tal requisito.
- b) Elimina o zero comum no outro par, e procura no vetor z por outro zero que seja complexo conjugado ao zero remanescente.
 - Neste caso, elimina o zero 1 do par 1-2 e procura no vetor z por outro zero que seja complexo conjugado ao zero 2. Encontrará o zero 12, formando o novo par 2-12.

Esta situação geralmente só ocorre entre os zeros de sinal, devido à sua proximidade e a existência da tolerância $\Delta\theta$. Neste caso, embora a busca dos pares mais próximos à c.r.u. conduza, em princípio, à escolha dos pares 1-4 e 5-8 como as estimativas das frequências, o algoritmo consegue selecionar corretamente os pares 1-4 e 2-12. Isto se deve ao fato de explorarmos a proximidade entre os zeros de sinal. Ao formarmos inicialmente os pares 1-2 e 1-4 e escolhermos o 1-4 como o correto, é como se admitíssemos que, devido à ocorrência de tal situação, o zero 2

deva ser um zero de sinal e portanto devemos buscar um conjugado para ele, encontrando o zero 12 e eliminando portanto o par 5-8.

5.3.3 COMPARAÇÕES QUANTITATIVAS

Para evidenciar os avanços produzidos pelo algoritmo, tanto na redução da SNR de limiar quanto na redução da variância das estimativas, realizamos as simulações da fig. 5.11 com o mesmo sinal da figs. 5.8, 5.9 e 5.10. A figura mostra o gráfico de variância das estimativas de $0.125(f_1) \times \text{SNR}$, similar ao da fig. 4.2. Utilizamos $L = 18(3.N/4)$ e $L = 23(KP)$ e para cada valor de L , três curvas foram traçadas, caracterizando uma das seguintes situações:

Caso 1 - Aplicação do método FBLP Modificado para $M = 4$, tomando a menor frequência como a estimativa de f_1 . Equivale aos passos 1 e 2 do algoritmo, com a posterior seleção dos 4 zeros mais próximos à c.r.u..

Caso 2 - Aplicação dos passos 1,2 e 3 do algoritmo.

Caso 3 - Aplicação dos passos 1,2,3 e 4 do algoritmo.

Observando a fig. 5.11, alguns fatos podem ser constatados:

- . Para ambos valores de L , o caso 3, que corresponde à aplicação do nosso algoritmo de modo completo, alcança o melhor desempenho, tanto em SNR de limiar quanto em variância das estimativas.
- . A aplicação do caso 2 (passos 1,2 e 3) diminui a SNR de limiar em relação ao caso 1. O caso 3 possibilita um deslocamento em toda a curva 2, produzindo uma diminuição na variância para todas as SNR's.
- . Mesmo abaixo do limiar do caso 3 (ou 2) o desempenho ainda é melhor do que no caso 1. Isto significa que mesmo nesta região o algoritmo consegue identificar alguns (mas não todos) experimentos onde zeros estranhos foram tomados como estimativas no caso 1.

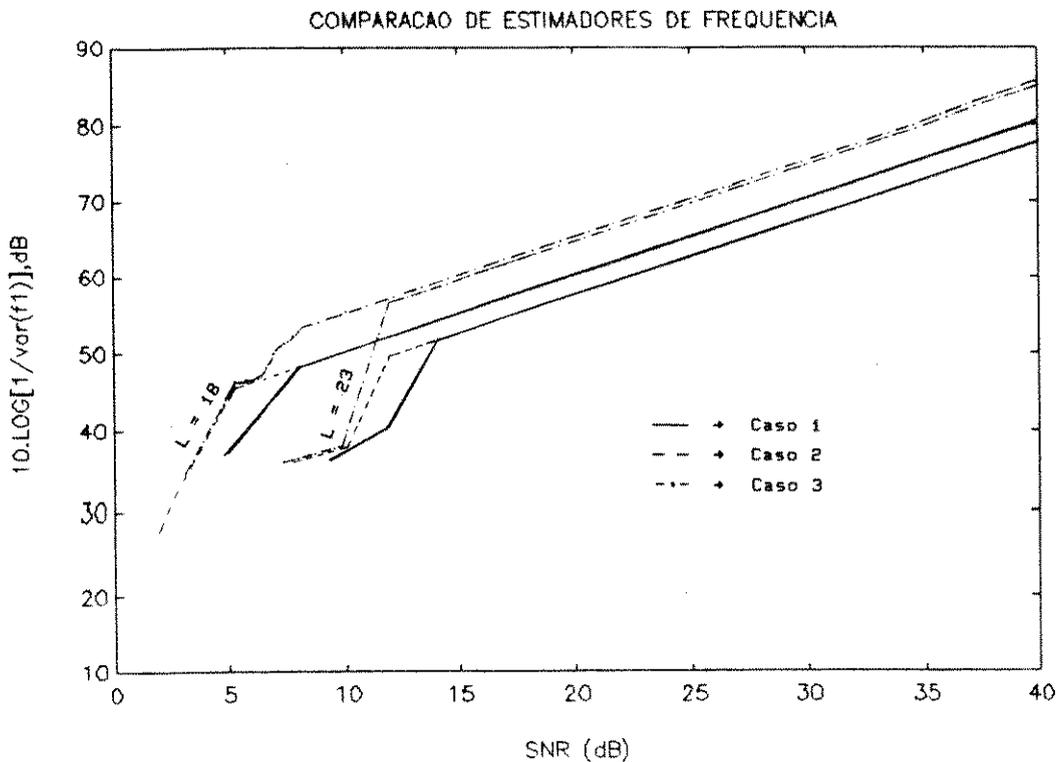


Fig. 5.11 - Comparação de desempenho entre os casos 1,2 e 3

5.3.4 ANÁLISE DE DESEMPENHO

A fim de avaliarmos o comportamento do desempenho do algoritmo com a variação de alguns parâmetros, tais como a ordem L e a SNR, foram traçadas curvas similares àquelas da secção 4.4.

A fig. 5.12 mostra o desempenho do algoritmo para L = 8,12,16,18,23. A fig. 5.13 mostra as SNR's de limiar para os casos 1 e 3 em diversos valores da ordem L do preditor. A fig. 5.14 mostra a variância da estimativa de f_1 em 15, 25 e 35 dB para diversos valores de L. Nesta figura são apresentados pontos duplos em L = 18 e L = 23 e para cada SNR: o de menor valor no eixo das ordenadas corresponde ao caso 1 enquanto que o de maior valor, ao caso 3.

Estas figuras permitem colocar algumas conclusões. L = 18(3.N/4) apresenta a menor SNR de limiar e a mesma se situa em cerca de 6 dB, o que caracteriza um excelente desempenho. A menor variância, em SNR's maiores que a de limiar, é atingida com L = 22. Entretanto, os desempenhos para L = 18 e L = 20 são muito próximos deste máximo, como podemos constatar na fig. 5.14. Podemos afirmar então que L = 18(3.N/4) representa o melhor compromisso quando se considera simultaneamente a SNR de limiar e variância após o limiar.

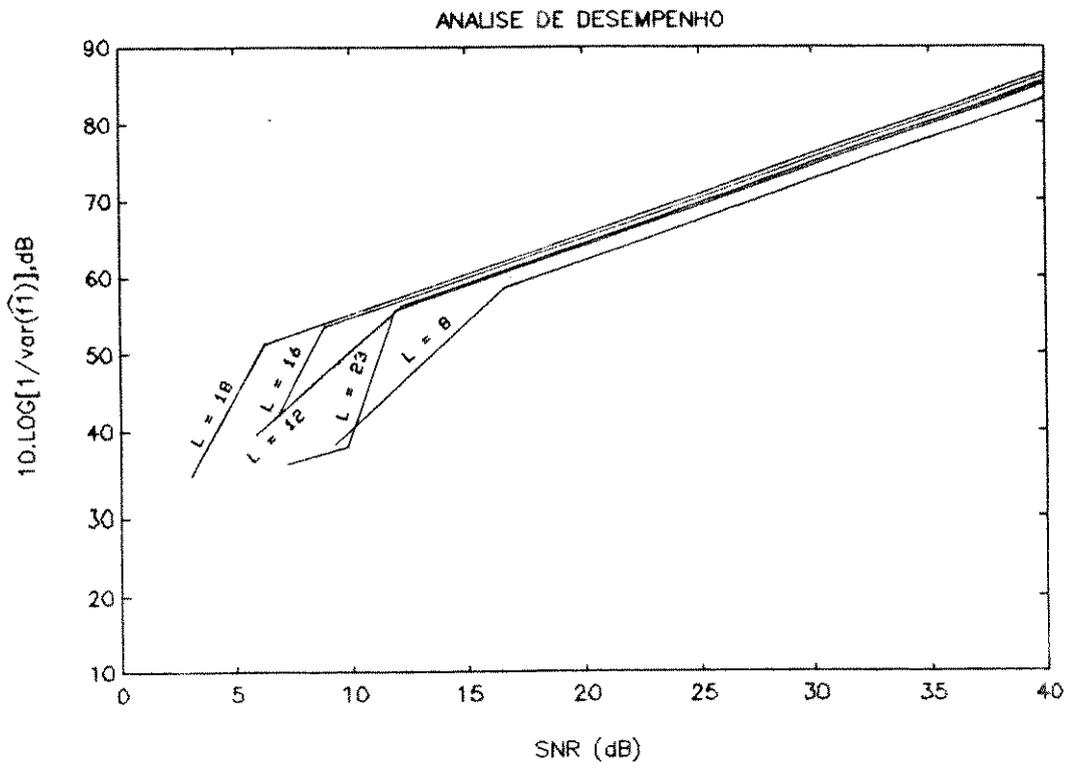


Fig. 5.12 - Análise de desempenho do algoritmo quanto à variação do L

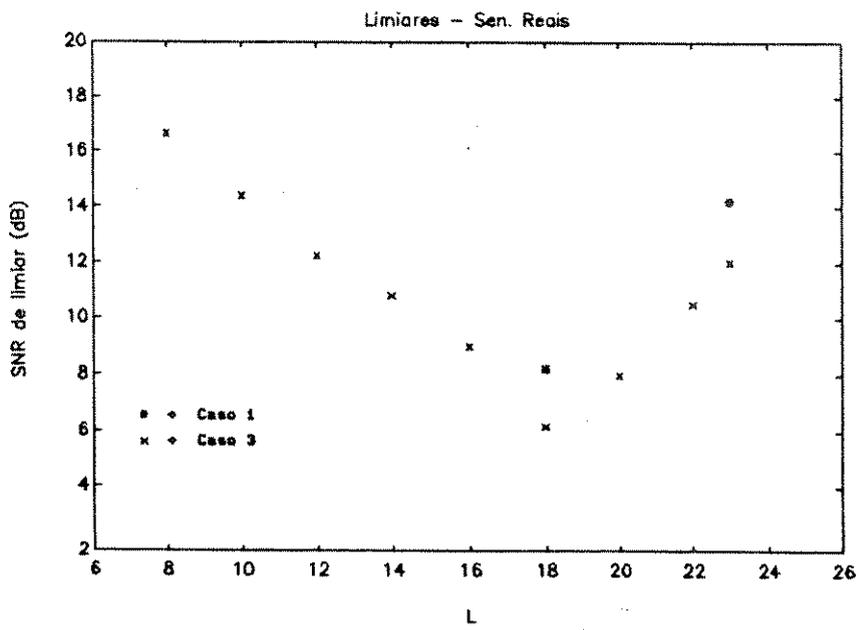


Fig. 5.13 - SNR de limiar X L - casos 1 e 3

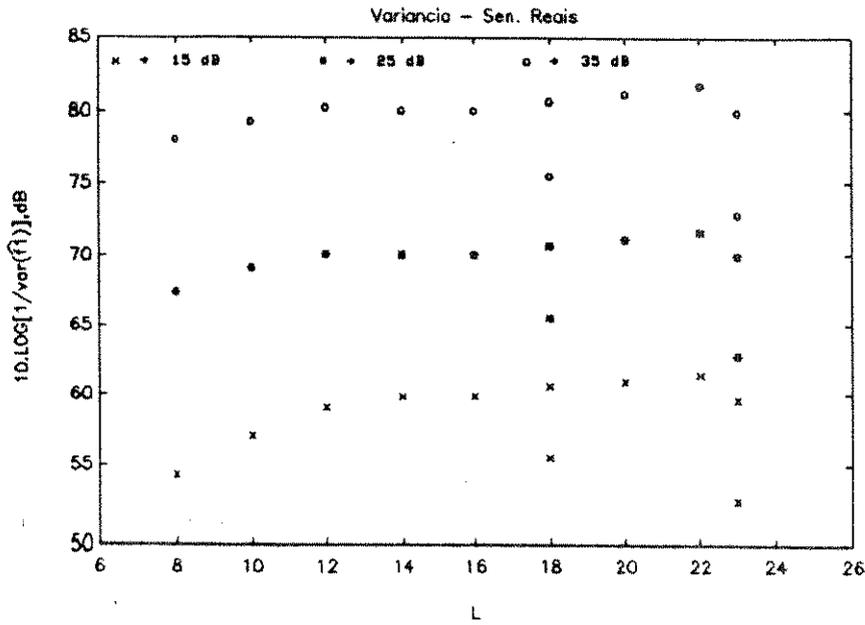


Fig. 5.14 - Variância das estimativas x L - SNR = 15, 25, 35 dB - casos 1 e 3

5.4 CONCLUSÕES

Quando o método FBLP Modificado atinge o limiar, métodos como o FLP Modificado e o BLP Modificado também o atingem. Por isso, procedimentos baseados nestes últimos para distinguir os zeros de sinal dos zeros estranhos, não funcionam satisfatoriamente.

No caso da detecção de senóides reais, propusemos um algoritmo novo que realiza uma distinção mais robusta entre os zeros de sinal e os zeros estranhos que àquela oferecida pelo Método FBLP Modificado. Com isto, conseguimos reduzir a SNR de limiar e a variância das estimativas e as simulações credenciaram $L = (3.N/4)$ como o valor da ordem do preditor que fornece o melhor desempenho.

O nosso algoritmo consegue extrair mais informação das amostras do sinal, propiciando resultados significativamente melhores que àqueles obtidos pela aplicação exclusiva do Método FBLP Modificado.

CAPÍTULO 6

CONCLUSÃO

6.1 RESUMO

Este trabalho apresentou a teoria de predição linear segundo o critério dos mínimos quadrados nos contextos forward, backward e forward-backward. Foi dada atenção especial à solução da equação normal determinística, demonstrando-se que a mesma sempre possui solução. Estudamos o teorema da decomposição via valor singular objetivando o cálculo da pseudo-inversa da matriz de dados e a solução de norma mínima da equação normal.

Demonstramos como os zeros de um polinômio formado pelos coeficientes do filtro de erro de predição, projetado segundo o método FBLP, são utilizados para estimar as frequências das senóides.

Desenvolvemos o método FBLP Modificado, utilizando o conceito de subespaço de sinal e subespaço de ruído. Analisamos seu desempenho quanto à variação de alguns parâmetros e apresentamos um exemplo prático de aplicação do método.

O capítulo 5 apresentou os resultados de especulações sobre a tentativa infrutífera de diminuir a SNR de limiar do Método FBLP Modificado. Estes estudos criaram subsídios para a elaboração de um algoritmo novo para a detecção de senóides reais.

6.2 CONCLUSÕES E CONTRIBUIÇÕES

O método FBLP mostra-se ineficiente quando dispomos de um número reduzido de amostras do sinal, as frequências estão mais próximas que o recíproco do intervalo de observação do sinal e a relação sinal-ruído é pequena. Os zeros estranhos frequentemente se situam mais próximos da c.r.u. que os zeros de sinal, provocando picos espectrais espúrios e, conseqüentemente, o efeito de limiar na curva de desempenho. Verificamos também que existe um compromisso entre estabilidade e resolução das estimativas e que tal compromisso depende do valor da ordem L do preditor. Quanto maior o L , maior será a resolução e menor a estabilidade do estimador. As causas destes comportamentos puderam ser identificadas ao dividirmos a matriz de correlação das amostras do sinal em seus subespaços de sinal e de ruído. O subespaço de ruído é o principal responsável pela queda no desempenho do método FBLP. Retendo apenas os autovalores e autovetores do subespaço de sinal, um novo método, denominado FBLP Modificado, foi desenvolvido. Tal método mostrou-se eficiente para detectar frequências, com desempenho praticamente igual ao do método Maximum-Likelihood e bastante robusto para diversas situações.

Extrapolamos as simulações realizadas por Tufts e Kumaresan afim de caracterizarmos este método de forma mais detalhada e observarmos o comportamento das estimativas com a variação dos seguintes parâmetros: número de amostras disponíveis, ordem do preditor, relação sinal-ruído e posicionamento das frequências. Confirmamos o fato conhecido que o melhor valor para a ordem do preditor é $3/4$ do número de amostras, porém demonstramos que este valor é na verdade o melhor compromisso quando se deseja a menor variância nas estimativas e a menor SNR de limiar. Enfocamos de forma especial o caso Kumaresan-Prony, pois apesar de seu desempenho ser um pouco inferior ao ótimo, sua simplicidade computacional relativa o torna extremamente atraente.

Os estudos por nós realizados sobre o comportamento dos zeros de sinal e os zeros estranhos, no caso de senóides reais, evidenciaram uma diferença significativa de comportamento entre os pares de zeros complexos conjugados de sinal e os estranhos. Esta diferença diz respeito à proximidade de cada par em relação à c.r.u. e permitiu desenvolvermos um algoritmo inédito para a detecção de senóides reais. Tal algoritmo oferece desempenho significativamente superior àquele obtido pela aplicação exclusiva do método FBLP Modificado.

Por fim, o desenvolvimento detalhado de toda a teoria envolvida, em conjunto com o pacote de programas utilizado nas simulações, credenciam o nosso trabalho como uma base sólida para as pesquisas futuras na FEE, que objetivarão novos avanços nesta técnica.

6.3 SUGESTÕES PARA PROSSEGUIMENTO DESTE TRABALHO

Durante todo o trabalho, as análises de desempenho levaram em conta apenas a estimação de uma das M frequências, no caso a menor delas. Um estudo que considerasse o desempenho na estimativa simultânea das M frequências poderia trazer novas informações. Também, se faz necessária a avaliação do desempenho do Método FBLP Modificado em função de M .

No algoritmo desenvolvido para detectar senóides reais, poderia ser feita uma análise mais detalhada do desempenho quanto à variação do ângulo de tolerância $\Delta\theta$, bem como de todos os parâmetros simulados e sugeridos para o Método FBLP Modificado.

Uma questão pouco explorada na literatura especializada, são as consequências que um ruído não-branco produziria sobre o desempenho do método FBLP Modificado. Também, ao longo de todo o trabalho ficou bastante claro a importância da precisão nos cálculos associados à decomposição via valor singular e à extração de zeros dos polinômios. Estudos envolvendo esta questão de precisão se relacionam com aqueles visando rotinas mais ágeis e mais apropriadas ao desenvolvimento de circuitos especializados que permitam uma aplicação mais ampla do método FBLP Modificado e da própria técnica de eliminação de componentes do subespaço de ruído.

APÊNDICE A

OS AUTOVALORES DE UMA MATRIZ HERMITIANA

DEFINIÇÃO A.1: Seja A uma matriz $m \times n$ que define uma transformação linear de um espaço vetorial V em um espaço vetorial W . O núcleo de A é o subespaço linear V_0 de V que consiste em todas as soluções da equação $A.v = 0$.

TEOREMA A.1: Seja a matriz $A_{(m \times n)}$, então:

$$\text{posto}(A) = \text{posto}(A^H \cdot A) = \text{posto}(A \cdot A^H)$$

DEMONSTRAÇÃO: Denotando $\nu(A)$ como a dimensão do núcleo da matriz A , sabemos que [3 - pg. 126]:

$$\nu(A) = n - k$$

$$\nu(A^H) = m - k$$

onde: k é o posto de A

Se $A^H \cdot A \cdot x = 0$, para algum x , então:

$$(x, A^H \cdot A \cdot x) = x^H \cdot A^H \cdot A \cdot x = (A \cdot x, A \cdot x) = |A \cdot x|^2 = 0$$

Portanto: $A^H \cdot A \cdot x = 0 \Rightarrow |A \cdot x|^2 = 0 \Rightarrow A \cdot x = 0$

É óbvio também que: $A \cdot x = 0 \Rightarrow A^H \cdot A \cdot x = 0$

Concluimos que: $A^H \cdot A \cdot x = 0 \Leftrightarrow A \cdot x = 0$ (A.1)

Implicando em:

$$\begin{aligned} \nu(A^H \cdot A) = \nu(A) &\Rightarrow n - \text{posto}(A^H \cdot A) = n - k \Rightarrow \\ \text{posto}(A^H \cdot A) = k &\Rightarrow \text{posto}(A^H \cdot A) = \text{posto}(A) \end{aligned} \tag{A.2}$$

Tomando agora: $A \cdot A^H \cdot x = 0$, para algum x , temos:

$$(x, A \cdot A^H \cdot x) = x^H \cdot A \cdot A^H \cdot x = (A^H \cdot x, A^H \cdot x) = \|A^H \cdot x\|^2 = 0$$

Portanto: $A \cdot A^H \cdot x = 0 \Rightarrow \|A^H \cdot x\|^2 = 0 \Rightarrow A^H \cdot x = 0$

É óbvio também que: $A^H \cdot x = 0 \Rightarrow A \cdot A^H \cdot x = 0$

Concluimos que: $A \cdot A^H \cdot x = 0 \Leftrightarrow A^H \cdot x = 0$ (A.3)

Implicando em:

$$\begin{aligned} \nu(A \cdot A^H) = \nu(A^H) &\Rightarrow m - \text{posto}(A \cdot A^H) = m - k \Rightarrow \\ \text{posto}(A \cdot A^H) = k &\Rightarrow \text{posto}(A \cdot A^H) = \text{posto}(A) \end{aligned} \tag{A.4}$$

TEOREMA A.2: Uma matriz Hermitiana do tipo $(A^H \cdot A)_{(n \times n)}$, de posto k , possui $(n - k)$ autovalores zero, enquanto que os demais (k no caso) são reais e positivos.

DEMONSTRAÇÃO: Se $A^H \cdot A \cdot x = 0$, o número de vetores x linearmente independentes, para os quais isto acontece, representa a dimensão do núcleo de $A^H \cdot A$. Sabemos também que:

$$\nu(A^H \cdot A) = n - k$$

Portanto, se existem $(n - k)$ autovetores linearmente independentes associados com o autovalor zero, a multiplicidade deste autovalor é $(n - k)$.

Considerando as auto-equações para os autovalores não nulos de $A^H \cdot A$:

$$A^H \cdot A \cdot x = \lambda \cdot x \quad \text{para } x \neq 0, \text{ e } \lambda \neq 0 \tag{A.5}$$

Podemos escrever: $\lambda \cdot \|x\|^2 = \lambda \cdot (x, x) = \lambda \cdot (x, \frac{A^H \cdot A \cdot x}{\lambda})$
 $= (x, A^H \cdot A \cdot x) = x^H \cdot A^H \cdot A \cdot x = (A \cdot x, A \cdot x) > 0$

Pois: $(A.x, A.x) = 0 \Leftrightarrow A.x = 0$

e pela eq. (A.1), $A.x = 0 \Leftrightarrow A^H.A.x = 0$, que por hipótese não ocorre. Portanto:

$$\lambda.\|x\|^2 > 0 \Rightarrow \lambda > 0$$

Fica assim provado que os k restantes autovalores são reais e positivos.

APÊNDICE B

OS AUTOVETORES DE UMA MATRIZ HERMITIANA

DEFINIÇÃO B.1: Uma matriz quadrada A que satisfaz $A.A^H = A^H.A$ é chamada *matriz normal*.

Toda matriz Hermitiana é normal, pois:

$$A.A^H = A.A = A^H.A$$

DEFINIÇÃO B.2: Uma matriz P que satisfaz $P^H.P = P.P^H = I$ é chamada *matriz unitária*.

Portanto, toda matriz unitária é não-singular e, neste caso, $P^H = P^{-1}$.

DEFINIÇÃO B.3: Se existe uma matriz unitária P tal que $P^H.A.P = B$, então dizemos que a matriz B é *equivalente unitária a A* , e que B foi obtida de A por meio de uma *transformação unitária*.

TEOREMA B.1: Matrizes equivalentes unitárias têm a mesma equação característica, e conseqüentemente, os mesmos autovalores.

DEMONSTRAÇÃO: Supondo B equivalente unitária a A , como $\text{DET}(P^H).\text{DET}(P) = \text{DET}(P^H.P) = \text{DET}(I) = 1$, então:

$$\begin{aligned}\text{DET}(B - \lambda.I) &= \text{DET}(P^H.(A - \lambda.I).P) \\ &= \text{DET}(P^H).\text{DET}(A - \lambda.I).\text{DET}(P) = \text{DET}(A - \lambda.I)\end{aligned}$$

Portanto, as equações características, e conseqüentemente os autovalores para A e B são idênticos.

TEOREMA B.2: Qualquer matriz quadrada $A_{(n \times n)}$, pode ser reduzida, por uma transformação unitária $P^H.A.P$, a uma matriz T , triangular superior, com os autovalores de A sobre a diagonal de T . Chamamos T uma *Forma Canônica de Schur* para A e a decomposição $A = P.T.P^H$ é chamada uma *Decomposição de Schur* de A .

DEMONSTRAÇÃO: Seja λ_1 um autovalor de A , com um autovetor associado x_1 normalizado, de maneira que $(x_1, x_1) = 1$. Podemos escolher vetores w_2, w_3, \dots, w_n tais que $x_1, w_2, w_3, \dots, w_n$ formam uma base ortonormal, ou seja, tais que:

$$Q = [x_1 \ w_2 \ w_3 \ \dots \ w_n] = [x_1 \ W] \quad \text{é unitária.}$$

Como $(w_i, x_1) = 0, i = 2, 3, \dots, n$, então:

$$W^H . x_1 = 0$$

Como x_1 é autovetor de A , com autovalor associado λ_1 :

$$A . x_1 = \lambda_1 . x_1$$

$$\begin{aligned} Q^H . A . Q &= \begin{bmatrix} x_1^H \\ W^H \end{bmatrix} . A . [x_1 \ W] = \begin{bmatrix} x_1^H \\ W^H \end{bmatrix} . [\lambda_1 . x_1 \ A . W] \\ &= \begin{bmatrix} \lambda_1 & x_1^H . A . W \\ 0 & W^H . A . W \end{bmatrix} = \begin{bmatrix} \lambda_1 & b^H \\ 0 & C \end{bmatrix} \end{aligned} \tag{B.1}$$

onde: $b^H = x_1^H . A . W$ (B.2)

$$C = W^H . A . W \tag{B.3}$$

Prosseguimos por indução, para provar o que diz respeito à triangularização. O teorema é verdadeiro para $n = 2$, pois C torna-se um escalar e a matriz da expressão (B.1) fica sob a forma exigida, triangular superior.

Suponhamos que A seja $n \times n$, e que o teorema é verdadeiro para $(n - 1)$. Então: $C = W^H . A . W$ é quadrada, de ordem $(n - 1)$, e já que foi suposto que o teorema é válido para $(n - 1)$, existe uma matriz unitária V , de ordem $(n - 1)$, tal que $V^H . C . V$ é triangular superior. Então, a matriz:

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

é unitária, pois:

$$U^H \cdot U = \begin{bmatrix} 1 & 0 \\ 0 & V^H \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & V \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & V^H \cdot V \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & I \end{bmatrix} = I \quad (B.4)$$

Prosseguindo ademonstração, tomamos agora o seguinte produto matricial:

$$\begin{aligned} U^H \cdot (Q^H \cdot A \cdot Q) \cdot U &= \begin{bmatrix} 1 & 0 \\ 0 & V^H \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & b^H \\ 0 & C \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & V \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & b^H \\ 0 & V^H \cdot C \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & V \end{bmatrix} = \begin{bmatrix} \lambda_1 & b^H \cdot V \\ 0 & V^H \cdot C \cdot V \end{bmatrix} \end{aligned} \quad (B.5)$$

Então, por (B.5), verificamos que: $U^H \cdot Q^H \cdot A \cdot Q \cdot U = (Q \cdot U)^H \cdot A \cdot (Q \cdot U)$ é triangular superior, pois $V^H \cdot C \cdot V$ é triangular superior, por hipótese. Ainda:

$$(Q \cdot U)^H \cdot (Q \cdot U) = U^H \cdot Q^H \cdot Q \cdot U = I = Q \cdot U \cdot U^H \cdot Q^H = (Q \cdot U) \cdot (Q \cdot U)^H \quad (B.6)$$

pois Q e U são unitárias.

(B.6) implica que: $P = Q \cdot U$ é unitária. Podemos, portanto, reescrever (B.5) como:

$$P^H \cdot A \cdot P = T \quad (B.7)$$

Onde T é triangular superior. O teorema torna-se então verdadeiro para uma matriz de dimensão n. Dado que o teorema é verdadeiro para n = 2, e sendo verdadeiro para (n - 1) é verdadeiro para n, então a triangularização está provada por indução.

Pelo teorema B.1, os autovalores de A e T são idênticos, e como os autovalores de uma matriz triangular estão sobre a sua diagonal principal, é imediato afirmar que, os autovalores de A estão sobre a diagonal de T. Logo, o teorema está demonstrado.

TEOREMA B.3: Uma matriz $A_{(n \times n)}$ tem um conjunto linearmente independente de n autovetores se e somente se existe uma matriz não-singular $P_{(n \times n)}$ e uma matriz diagonal Λ para a qual:

$$A = P \cdot \Lambda \cdot P^{-1} \tag{B.8}$$

$$\Lambda = P^{-1} \cdot A \cdot P \tag{B.9}$$

As colunas de P podem ser tomadas como autovetores de A associados respectivamente com os autovalores λ_i , onde λ_i é o (i,i) -ésimo elemento de Λ .

DEMONSTRAÇÃO:

(\Rightarrow)

Se A tem um conjunto linearmente independente de n autovetores p_1, p_2, \dots, p_n associados com os autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$, então a matriz:

$$P = [p_1 \ p_2 \ \dots \ p_n] \quad \text{é não-singular pois tem posto } n.$$

Escrevendo as auto-equações para A :

$$A \cdot p_i = \lambda_i \cdot p_i \quad i = 1, 2, \dots, n$$

Reescrevendo as equações acima matricialmente:

$$A \cdot P = [A \cdot p_1 \ A \cdot p_2 \ \dots \ A \cdot p_n] = [\lambda_1 \cdot p_1 \ \lambda_2 \cdot p_2 \ \dots \ \lambda_n \cdot p_n] = P \cdot \Lambda$$

onde:

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

Como P é não-singular:

$$A \cdot P = P \cdot \Lambda \Rightarrow A = P \cdot \Lambda \cdot P^{-1} \quad \text{e} \quad \Lambda = P^{-1} \cdot A \cdot P$$

(\Leftarrow)

Se $A = P \Lambda P^{-1}$ e $\Lambda = P^{-1} \cdot A \cdot P$

Então: $A \cdot P = P \cdot \Lambda$, que devido as estruturas de P e Λ , possibilita escrever:

$$A \cdot p_i = \lambda_i \cdot p_i \quad i = 1, 2, \dots, n$$

Portanto λ_i são autovalores de A , e p_i os respectivos autovetores. Como P é não-singular, os autovetores p_1, p_2, \dots, p_n , formam um conjunto linearmente independente, e a demonstração está concluída.

TEOREMA B.4: Uma matriz $A_{(n \times n)}$ é normal se e somente se A pode ser reduzida, por uma transformação unitária, a uma forma canônica de Schur diagonal: $D = P^H \cdot A \cdot P$ (onde P é unitária e D é diagonal). Os autovalores de A estarão sobre a diagonal de D .

DEMONSTRAÇÃO:

(\Leftarrow)

Suponhamos que exista P unitária tal que $P^H \cdot A \cdot P = D$, onde D é uma matriz diagonal, então $A = P \cdot D \cdot P^H$, e:

$$A^H \cdot A = P \cdot D^H \cdot D \cdot P^H$$

$$A \cdot A^H = P \cdot D \cdot D^H \cdot P^H$$

mas: $D \cdot D^H = D^H \cdot D$, pois D é diagonal. Portanto $A^H \cdot A = A \cdot A^H$, e A é normal.

(\Rightarrow)

Agora, suponhamos A uma matriz normal. O teorema B.2 afirma que, existe uma matriz unitária Q tal que $Q^H \cdot A \cdot Q = T$, onde T é triangular superior. Como A é normal:

$$T^H \cdot T = Q^H \cdot A^H \cdot A \cdot Q = Q^H \cdot A \cdot A^H \cdot Q = T \cdot T^H \Rightarrow T \text{ é normal}$$

Como T é triangular superior, $t_{ij} = 0$, para $i > j$, e o (i,i) -ésimo elemento de $T \cdot T^H$ é:

$$\sum_{j=1}^n |t_{ij}|^2$$

Enquanto que o (i,i) -ésimo elemento de $T^H.T$ é:

$$\sum_{j=1}^i |t_{ji}|^2$$

Como $T.T^H = T^H.T$:

$$\begin{aligned} \sum_{j=1}^n |t_{ij}|^2 &= \sum_{j=1}^i |t_{ji}|^2 \\ \sum_{j=i+1}^n |t_{ij}|^2 &= \sum_{j=1}^{i-1} |t_{ji}|^2 \end{aligned} \tag{B.10}$$

fazendo $i = 1$:

$$\sum_{j=2}^n |t_{1j}|^2 = 0$$

de forma que: $t_{12} = t_{13} = \dots = t_{1n} = 0$ (B.11)

Tomando $i = 2$ em (B.10):

$$\sum_{j=3}^n |t_{2j}|^2 = |t_{12}|^2 = 0$$

pois, por (B.11), $t_{12} = 0$. Portanto: $t_{23} = t_{24} = \dots = t_{2n} = 0$. Continuando este procedimento, é imediato perceber que:

$$t_{ij} = 0 \text{ para } j > i \Rightarrow T \text{ é diagonal}$$

Pelo teorema B.2, os elementos da diagonal são os autovalores de A , e a prova do teorema está concluída.

TEOREMA B.5: Uma matriz $A_{(n \times n)}$ é normal se e somente se tem um conjunto de n autovetores linearmente independentes, e pode ser escrita como $P^H.A.P = D$, onde D é uma matriz diagonal com os elementos da sua diagonal principal denotados por d_{ii} e as colunas de P são os autovetores linearmente independentes, associados aos autovalores d_{ii} .

DEMONSTRAÇÃO: Pelo teorema B.4, A é normal se e somente se existe:

$$P^H \cdot A \cdot P = D \quad (\text{B.12})$$

Como P é unitária, $P^H = P^{-1}$, e (B.12) torna-se:

$$P^{-1} \cdot A \cdot P = D \quad (\text{B.13})$$

$$A = P \cdot D \cdot P^{-1} \quad (\text{B.14})$$

Pelo teorema B.3, fica provado que A tem um conjunto de n autovetores linearmente independentes e as colunas de P podem ser tomadas como autovetores de A , associados, respectivamente, com os autovalores d_{ii} , que são os elementos da diagonal principal de D .

TEOREMA B.6: Autovetores associados com autovalores distintos de uma matriz normal são ortogonais.

DEMONSTRAÇÃO: As auto-equações para $A_{(n \times n)}$ são:

$$A \cdot x_i = \lambda_i \cdot x_i \quad i = 1, 2, \dots, n$$

Tomando o produto:

$$\begin{aligned} \lambda_j \cdot (x_i, x_j) &= \lambda_j \cdot (x_i, \frac{A \cdot x_j}{\lambda_j}) = (x_i, A \cdot x_j) = x_i^H \cdot A \cdot x_j \\ &= (A^H \cdot x_i)^H \cdot x_j = (A^H \cdot x_i, x_j) \end{aligned}$$

Como: $A^H \cdot x_i = \lambda_i^* \cdot x_i, i = 1, 2, \dots, n$, [3] temos:

$$(A^H \cdot x_i, x_j) = (\lambda_i^* \cdot x_i, x_j) = \lambda_i^* \cdot (x_i, x_j)$$

$$\text{Portanto:} \quad \lambda_j \cdot (x_i, x_j) = \lambda_i^* \cdot (x_i, x_j) \quad (\text{B.15})$$

Supondo $\lambda_i \neq \lambda_j$, para (B.15) tornar-se verdadeira, é necessário que:

$$(x_i, x_j) = 0$$

Portanto, x_i e x_j são ortogonais, quando $\lambda_i \neq \lambda_j$.

TEOREMA B.7: Uma matriz $A_{(n \times n)}$ é normal se e somente se A tem um conjunto linearmente independente de n autovetores que podem ser escolhidos de maneira a formar um conjunto ortonormal. Além disso, no caso de uma matriz normal, um autovalor de multiplicidade s tem associado um conjunto ortonormal de s autovetores.

DEMONSTRAÇÃO:

(\Rightarrow)

Seja $A_{(n \times n)}$, uma matriz normal. Pelo teorema B.5:

$$P^H \cdot A \cdot P = D$$

$$A \cdot P = P \cdot D \Rightarrow A \cdot [p_1 \ p_2 \ \dots \ p_n] = [p_1 \ p_2 \ \dots \ p_n] \cdot \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

$$A \cdot p_i = \lambda_i \cdot p_i \quad i = 1, 2, \dots, n$$

Portanto, a base ortonormal pode ser escolhida como sendo p_1, p_2, \dots, p_n , pois sabemos que P é unitária.

(\Leftarrow)

Seja p_1, p_2, \dots, p_n autovetores ortonormais de A . Portanto, $P = [p_1 \ p_2 \ \dots \ p_n]$ é unitária e:

$$A \cdot p_i = \lambda_i \cdot p_i \quad i = 1, 2, \dots, n$$

$$A \cdot P = \Lambda \cdot P$$

onde: Λ é uma matriz diagonal contendo os λ_i .

$$P^H \cdot A \cdot P = \Lambda$$

Pelo teorema B.5 fica provado que A é normal.

Concluimos que A tem n autovetores ortonormais. "Podem ser escolhidos",

citado no enunciado do teorema, se faz necessário, pois os autovetores associados com autovalores distintos de matrizes normais, são automaticamente ortogonais (teorema B.6), mas os autovetores associados com os autovalores de multiplicidade s não são obrigatoriamente ortogonais entre si. O teorema B.5 afirma apenas que são linearmente independentes. Porém, pelo teorema B.7 sempre é possível escolhê-los ortogonais entre si. Supondo que $A_{(n \times n)}$ possui p autovalores distintos, cada um com multiplicidade s_i , $i = 1, 2, \dots, p$, é óbvio que $\sum_{i=1}^p s_i = n$. Já que autovetores associados a autovalores distintos são automaticamente ortogonais, cada um dos p grupos de s_i autovetores já possui os seus s_i autovetores ortogonais a todos os autovetores dos outros grupos. Portanto, basta garantir a ortogonalidade dentro do seu próprio grupo, e a ortogonalidade estará garantida para todos os n autovetores.

REFERÊNCIAS E BIBLIOGRAFIA

- [1] Haykin, S.; *"Adaptive Filter Theory"*, Prentice-Hall, Englewood Cliffs, New Jersey, 1986.
- [2] Marple, S. L., Jr.; *"Digital Spectral Analysis"*, Prentice-Hall, Englewood Cliffs, New Jersey, 1987.
- [3] Noble, B. and Daniel, J. W.; *"Álgebra Linear Aplicada"*, Prentice-Hall do Brasil, Rio de Janeiro, RJ, 1986.
- [4] Kay, S. M.; *"Modern Spectral Estimation: Theory and Application"*, Prentice-Hall, Englewood Cliffs, New Jersey, 1988.
- [5] Klema, V. C. and Laub, A. J.; *"The Singular Value Decomposition: Its Computation and Some Applications"*, IEEE Trans. Automatic Control, vol. AC-25, pp. 164-176, 1980.
- [6] Lang, S. W. and McClellan, J. H.; *"Frequency Estimation with Maximum Entropy Spectral Estimators"*, IEEE Trans. Acoust., Speech., and Signal Processing, vol. ASSP-28, pp. 716-724, 1980.
- [7] Tufts, D. W. and Kumaresan, R.; *"Estimation of Frequencies of Multiple Sinusoids: Making Linear Prediction Perform Like Maximum Likelihood"*, Proc. IEEE, vol. 70, pp. 975-989, 1982.
- [8] Schmidt, R. O.; *"A Signal Subspace Approach to Multiple Emitter Location and Spectral Estimation"*, Ph. D. Dissertation, Stanford University, Stanford, Calif., 1981.

- [9] Wax, M. and Kailath, T.; "*Detection of Signals by Information Theoretic Criteria*", IEEE Trans. Acoustics, Speech, and Signal Processing, vol. ASSP-33, pp. 387-392, 1985.
- [10] Wax, M.; "*Detection and Estimation of Superimposed Signals*", Ph. D. Dissertation, Stanford University, Stanford, Calif., 1985.
- [11] Kumaresan, R.; "*On Zeros of the Linear Prediction-Error for Deterministic Signals*", IEEE Trans. Acoustics, Speech, and Signal Processing, vol. ASSP-31, pp. 217-220, 1983.
- [12] Ulrych, T. J. and Clayton, R. W.; "*Times Series Modelling and Maximum Entropy*", Phys. Earth Planet Interiors, vol. 12, pp. 188-200, 1976.
- [13] Nuttall, A. H. ; "*Spectral Analysis of a Univariate Process with Bad Data Points via Maximum Entropy and Linear Predictive Techniques*", Naval Underwater System Center(NUSC) Scientific and Engineering Studies, Spectral Estimation, NUSC, New London, Conn, 1976.
- [14] Parthasarathy, S. and Tufts, D. W.; "*Signal Modeling by Exponential Segments and Application in Voiced Speech Analysis*", Proc. IEEE ICASSP, vol. 2, pp 645-648, 1987.
- [15] Kay, S. M.; "*Frequency Estimation by Principal Component AR Spectral Estimation Method without Eigendecomposition*", IEEE Trans. on Acoustics, Speech, and Signal Processing, vol. 36, NO. 1, 1988.
- [16] Rao, S. S. and Raghavan, K. R.; "*Estimating the Frequencies of Sinusoids at Low Signal to Noise Ratios Using the Principal Component Method in Tandem with The Burg Algorithm*", Proc. IEEE ICASSP, pp 2210-2213, 1989.
- [17] Clergeot, H., Tressens, S. and Ouamri, A.; "*Performance of High Resolution Frequencies Estimation Methods Compared to the Cramer-Rao Bounds*", IEEE Trans. on Acoustics, Speech, and Signal Processing, vol. 37, NO. 11, 1989.
- [18] Wigren, T. and Nehorai, A.; "*Asymptotic Cramér-Rao Bounds for Estimation of the Parameters of Damped Sine Waves in Noise*", IEEE Trans. on Acoustics, Speech, and Signal Processing, vol. 34, NO. 4, 1991.

Referências e Bibliografia

- [19] Matausek, M. R., Stanković, S. S. and Radović, D. V.; *"Iterative Inverse Filtering Approach to the Estimation of Frequencies of Noisy Sinusoids"*, IEEE Trans. on Acoustics, Speech, and Signal Processing, vol. ASSP-31, NO. 6, 1983.
- [20] Porat, B. and Friedlander, B.; *"A Modification of the Kumaresan-Tufts Method for Estimating Rational Impulse Responses"*, IEEE Trans. on Acoustics, Speech, and Signal Processing, vol. ASSP-34, NO. 5, 1986.
- [21] Kay, S.; *"Accurate Frequency Estimation at Low Signal-to-Noise Ratio"*, IEEE Trans. on Acoustics, Speech, and Signal Processing, vol. ASSP-32, NO. 3, 1984.
- [22] Antunes, E. J. B. and Lopes, A.; *"Frequencies Estimation of Real Sinusoids"*, International Symposium on Signals, Systems and Electronics , ISSSE'92, Paris, France, 1-4 september 1992.
- [23] Antunes, E. J. B. e Lopes, A.; *"Estimação de Frequências de Senóides Reais"*, V Simpósio Brasileiro de Microondas / X Simpósio Brasileiro de Telecomunicações, TELEMÓ'92, Brasília, Brasil, 21-24 de julho de 1992.
- [24] Makhoul, J.; *"Linear Prediction: A Tutorial Review"*, Proc. IEEE, vol. 64, Feb. 1976, Reprinted from Proc. IEEE, vol. 63, April 1975.
- [25] Lawson, C.L. and Hanson, R.J.; *"Solving Least Squares Problems"*, Prentice-Hall, Englewood Cliffs, New Jersey, 1974.