

DESENVOLVIMENTO DE ALGORITMOS ADAPTATIVOS USANDO A
DESCRIÇÃO DE DISTÚRBO MÍNIMO COM RESTRIÇÕES

Fabiano Tondello Castoldi

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO
DOS PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE
EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Aprovada por:

Prof. Marcello Luiz Rodrigues de Campos, Ph.D.

Prof. Luiz Wagner Pereira Biscainho, D.Sc.

Prof. José Antonio Apolinário Júnior., D.Sc.

RIO DE JANEIRO, RJ - BRASIL

NOVEMBRO DE 2007

CASTOLDI, FABIANO TONDELLO

Desenvolvimento de Algoritmos
Adaptativos Usando a Descrição de
Distúrbio Mínimo com Restrições [Rio
de Janeiro] 2007

XI, 96 p., 29,7 cm (COPPE/UFRJ,
M.Sc., Engenharia Elétrica, 2007)

Dissertação - Universidade Federal do
Rio de Janeiro, COPPE

1.Filtros Adaptativos 2.Otimização
3.Descrição por Distúrbio Mínimo
4.Barreira Logarítmica

I.COPPE/UFRJ II.Título (série)

Agradecimentos

Dedico este trabalho a meu pai Agenor e minha mãe Teresinha, por sempre me apoiarem, incentivando-me a prosseguir com meus estudos. A meu irmão Marcelo pela ajuda e pelo companheirismo durante toda essa caminhada. A minha irmã Gilvana, pelo suporte e pelos conselhos, mesmo distante. A meu cunhado Márcio e minha cunhada Márcia por estarem presentes me incentivando e ajudando de uma maneira ou outra.

Ao meu orientador Marcello L. R. de Campos pela confiança, pela paciência, pelos ensinamentos e pelas diversas vezes que me ajudou durante todo o desenvolvimento dessa tese. Aos demais professores pelos conhecimentos passados.

Aos amigos, que, sempre presentes, influenciaram direta ou indiretamente a conclusão dessa tese.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

DESENVOLVIMENTO DE ALGORITMOS ADAPTATIVOS USANDO A
DESCRIÇÃO DE DISTÚRBO MÍNIMO COM RESTRIÇÕES

Fabiano Tondello Castoldi

Novembro/2007

Orientador: Marcello Luiz Rodrigues de Campos

Programa: Engenharia Elétrica

As aplicações de algoritmos adaptativos são importantes em diversas áreas, como telecomunicações, sistemas de controle e outras. O modo como a função objetivo de um algoritmo adaptativo é apresentada pode oferecer informações importantes sobre o desempenho ou o comportamento do algoritmo.

Os algoritmos adaptativos são normalmente apresentados e desenvolvidos baseados em uma aproximação estocástica do vetor gradiente e matriz Hessiana, ou na minimização determinística dos erros quadráticos *a posteriori* de saída. Algoritmos de gradiente-descendente, como o algoritmo LMS (*Least Mean Squares*) e o algoritmo QN (quasi-Newton), estão no primeiro grupo, enquanto que o algoritmo RLS (*Recursive Least Squares*) se enquadra no segundo grupo. Esta é somente a maneira como normalmente os algoritmos podem ser apresentados e analisados, porém o algoritmo RLS pode ser visto como um algoritmo de aproximação estocástica e o algoritmo LMS também minimiza uma função objetivo determinística. Entretanto, algumas descrições determinísticas minimizadas por alguns algoritmos oferecem pouca informação sobre o comportamento do algoritmo.

Este trabalho propõe uma maneira alternativa para o desenvolvimento de algoritmos adaptativos, através da descrição da função-objetivo determinística como um problema de otimização da norma quadrática do vetor de coeficientes sujeita a restrições. Essa descrição alternativa, chamada de distúrbio mínimo, é estudada neste trabalho como uma ferramenta para o desenvolvimento de algoritmos adaptativos diversos, e resulta em uma função-objetivo determinística simples e que pode ser modificada para obtenção de novos algoritmos ou reformulação de algoritmos conhecidos.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

DEVELOPMENT OF ADAPTATION ALGORITHMS USING MINIMUM
DISTURBANCE DESCRIPTION WITH RESTRICTIONS

Fabiano Tondello Castoldi

November/2007

Advisor: Marcello Luiz Rodrigues de Campos

Department: Electrical Engineering

The applications of adaptive algorithms are important in many areas, such as telecommunication and control systems. How the objective function of an adaptive algorithm is presented can provide relevant information about the algorithm's performance and behavior.

Adaptation algorithms are usually presented and developed based either on a stochastic approximation of the gradient vector and the Hessian matrix, or on a deterministic minimization of square *a posteriori* output errors. Gradient-descent algorithms, such as the LMS (*Least Mean-Squares*) algorithm and the QN (quasi-Newton) algorithm, are usually placed in the first group, whereas RLS (*Recursive Least Squares*) algorithm is placed in the second group. These are only the way of presenting and analyzing the algorithms. However the RLS algorithm can also be seen as a stochastic approximation algorithm, and the LMS algorithm also does minimize a deterministic objective function. On the other hand, some description of deterministic functions minimized by some algorithms offer very limited insight on their behavior.

In this work an alternative way to develop adaptation algorithms based on the description of a deterministic objective function as an optimization problem of the quadratic norm of the coefficient vector subjected to restrictions is studied. This alternative description, called minimum disturbance, is studied in this work as a tool for the development of adaptation algorithms, resulting in a simple deterministic objective function that can be modified to yield new algorithms or reformulating known algorithms.

Sumário

1	Introdução	1
1.1	Introdução	1
1.2	Filtro de Wiener	5
1.3	O Algoritmo RLS	6
1.4	Método de Otimização	7
1.5	Objetivos e Contribuições desta Tese	8
1.5.1	Contribuições desta Tese	9
2	Filtragem Adaptativa como Problema de Otimização	11
2.1	Multiplicadores de Lagrange	12
2.2	Método de Pontos Interiores	13
2.3	Desenvolvimento de Algoritmos Adaptativos pelo Método da Barreira Logarítmica	16
2.3.1	Algoritmo RLS	17
2.3.2	Algoritmo QN	19
2.4	Desenvolvimento de Algoritmos Adaptativos Usando a Descrição de Distúrbio Mínimo	22
2.4.1	Algoritmo RLS	24
2.4.2	Algoritmo QN	25
2.5	Contribuições do Capítulo	27
3	Novos Algoritmos	28
3.1	Algoritmo 1	28
3.2	Algoritmo 2	31
3.3	Algoritmo 3	32

3.4	Algoritmo 4	33
3.5	Algoritmo 5	35
3.6	Algoritmo 6	38
3.7	Contribuições do Capítulo	40
4	Simulações	42
4.1	Iniciação da Matriz de Autocorrelação	43
4.2	Influência do Passo no Algoritmo 1	47
4.3	Experimento I: Ruído Branco como Sinal de Entrada	49
4.3.1	Experimento Ia: SNR = 80dB	50
4.3.2	Experimento Ib: SNR = 40dB	51
4.3.3	Experimento Ic: SNR = 20dB	53
4.4	Experimento II: Ruído Colorido como Sinal de Entrada	53
4.4.1	Experimento IIa: SNR = 80dB	54
4.4.2	Experimento IIb: SNR = 40dB	56
4.4.3	Experimento IIc: SNR = 20dB	57
4.5	Experimento III: Sinal de Entrada Senoidal	57
4.5.1	Experimento IIIa: SNR = 80dB	58
4.5.2	Experimento IIIb: SNR = 40dB	60
4.5.3	Experimento IIIc: SNR = 20dB	62
4.6	Experimento IV: Sinal de Entrada Senoidal com Ruído	62
4.6.1	Experimento IVa: SNR = 80dB	63
4.6.2	Experimento IVb: SNR = 40dB	64
4.6.3	Experimento IVc: SNR = 20dB	66
4.7	Experimento V: Mudança Abrupta dos Coeficientes	67
4.7.1	Experimento Va: Sinal de Entrada Ruído Branco	67
4.7.2	Experimento Vb: Sinal de Entrada Ruído Colorido	69
4.7.3	Experimento Vc: Sinal de Entrada Senoidal	70
4.7.4	Experimento Vd: Sinal de Entrada Senoidal com Ruído	72
4.8	Simulações Comparativas	74
4.8.1	Simulações Comparativas I: Coeficientes Fixos e Sinal de Entrada Ruído Branco	75

4.8.2	Simulações Comparativas II: Coeficientes Fixos e Sinal de Entrada Ruído Colorido	76
4.8.3	Simulações Comparativas III: Coeficientes Fixos e Sinal de Entrada Senoidal	77
4.8.4	Simulações Comparativas IV: Coeficientes Fixos e Sinal de Entrada Senoidal Distorcido	79
4.8.5	Simulações Comparativas V: Variação Abrupta dos Coeficientes e Sinal de Entrada Ruído Branco	80
4.8.6	Simulações Comparativas VI: Variação Abrupta dos Coeficientes e Sinal de Entrada Ruído Colorido	81
4.8.7	Simulações Comparativas VII: Variação Abrupta dos Coeficientes e Sinal de Entrada Senoidal Distorcido	81
5	Considerações Finais	85
	Referências Bibliográficas	90
A	Pseudo-Código dos Algoritmos Desenvolvidos	92
A.1	Algoritmo 1	92
A.2	Algoritmo 2	93
A.3	Algoritmo 3	93
A.4	Algoritmo 4	94
A.5	Algoritmo 5	95
A.6	Algoritmo 6	95

Lista de Figuras

1.1	Diagrama de um sistema adaptativo.	2
1.2	Filtragem adaptativa para identificação de sistema.	3
1.3	Filtragem adaptativa para equalização de canal.	3
1.4	Filtragem adaptativa para predição de sinal.	4
1.5	Filtragem adaptativa para melhoramento de sinal.	4
2.1	Função indicadora e sua aproximação pela barreira logarítmica. . . .	15
3.1	Comportamento de $\mu(n)$ para sinal de entrada senoidal	37
3.2	Comportamento de $\mu(n)$ para sinal de entrada ruído branco	38
4.1	Comportamento do passo do algoritmo 5 para entrada senoidal. . . .	44
4.2	Comportamento do MSE do algoritmo 5 para entrada senoidal. . . .	45
4.3	Comportamento do passo do algoritmo 5 para entrada WGN.	47
4.4	Comportamento do MSE do algoritmo 5 para entrada WGN.	48
4.5	Comportamento do MSE do algoritmo RLS para entrada WGN.	48
4.6	Comportamento do MSE do algoritmo 1 para diferentes valores do passo do algoritmo.	49
4.7	MSE dos algoritmos 1–3 para sinal de entrada ruído branco e SNR=80dB.	50
4.8	MSE dos algoritmos 4–6 para sinal de entrada ruído branco e SNR=80dB.	51
4.9	MSE dos algoritmos 1–3 para sinal de entrada ruído branco e SNR=40dB.	52
4.10	MSE dos algoritmos 4–6 para sinal de entrada ruído branco e SNR=40dB.	52
4.11	MSE dos algoritmos 1–3 para sinal de entrada ruído branco e SNR=20dB.	53
4.12	MSE dos algoritmos 4–6 para sinal de entrada ruído branco e SNR=20dB.	54
4.13	MSE dos algoritmos 1–3 para sinal de entrada ruído colorido e SNR=80dB.	55
4.14	MSE dos algoritmos 4–6 para sinal de entrada ruído colorido e SNR=80dB.	55
4.15	MSE dos algoritmos 1–3 para sinal de entrada ruído colorido e SNR=40dB.	56

4.16	MSE dos algoritmos 4–6 para sinal de entrada ruído colorido e SNR=40dB.	56
4.17	MSE dos algoritmos 1–3 para sinal de entrada ruído colorido e SNR=20dB.	58
4.18	MSE dos algoritmos 4–6 para sinal de entrada ruído colorido e SNR=20dB.	58
4.19	MSE dos algoritmos 1–3 para sinal de entrada senoidal e SNR=80dB.	59
4.20	MSE dos algoritmo 2 para sinal de entrada senoidal e SNR=80dB. . .	60
4.21	MSE dos algoritmos 4–6 para sinal de entrada senoidal e SNR=80dB.	60
4.22	MSE dos algoritmos 1 e 3 para sinal de entrada senoidal e SNR=40dB.	61
4.23	MSE dos algoritmos 5 e 6 para sinal de entrada senoidal e SNR=40dB.	61
4.24	MSE dos algoritmos 1 e 3 para sinal de entrada senoidal e SNR=20dB.	62
4.25	MSE dos algoritmos 5 e 6 para sinal de entrada senoidal e SNR=20dB.	63
4.26	MSE dos algoritmos 1–3 para sinal de entrada senoidal distorcida e SNR=80dB.	63
4.27	MSE dos algoritmos 4–6 para sinal de entrada senoidal distorcida e SNR=80dB.	64
4.28	MSE dos algoritmos 1–3 para sinal de entrada senoidal distorcida e SNR=40dB.	65
4.29	MSE dos algoritmos 4–6 para sinal de entrada senoidal distorcida e SNR=40dB.	65
4.30	MSE dos algoritmos 1–3 para sinal de entrada senoidal distorcida e SNR=20dB.	66
4.31	MSE dos algoritmos 4–6 para sinal de entrada senoidal distorcida e SNR=20dB.	67
4.32	MSE dos algoritmos 1–3 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada WGN.	68
4.33	MSE dos algoritmos 4–6 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada WGN.	69
4.34	MSE dos algoritmos 1–3 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada correlacionado.	70
4.35	MSE dos algoritmos 4–6 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada correlacionado.	70
4.36	MSE dos algoritmos 1 e 3 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada senoidal.	71

4.37	MSE dos algoritmos 4–6 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada senoidal.	72
4.38	Passos dos algoritmos 5 e 6 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada senoidal.	72
4.39	MSE dos algoritmos 1–3 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada senoidal distorcida.	73
4.40	MSE dos algoritmos 4–6 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada senoidal distorcida.	73
4.41	MSE dos algoritmos 1, 6, RLS e QN para sistema com coeficientes fixos e sinal de entrada WGN.	76
4.42	MSE dos algoritmos 1, 6, RLS e QN para sistema com coeficientes fixos e sinal de entrada ruído colorido.	77
4.43	MSE dos algoritmos 1, 6, RLS e QN para sistema com coeficientes fixos e sinal de entrada senoidal.	78
4.44	MSE dos algoritmos 1, 6, RLS e QN para sistema com coeficientes fixos e sinal de entrada senoidal.	78
4.45	MSE dos algoritmos 1, 6, RLS e QN para sistema com coeficientes fixos e sinal de entrada senoidal distorcido.	79
4.46	MSE dos algoritmos 1, 6, RLS e QN para sistema com variação abrupta dos coeficientes e sinal de entrada WGN.	80
4.47	MSE dos algoritmos 1, 6, RLS e QN para sistema com variação abrupta dos coeficientes e sinal de entrada ruído colorido.	82
4.48	MSE dos algoritmos 1, 6, RLS e QN para sistema com variação abrupta dos coeficientes e sinal de entrada senoidal distorcido.	83

Capítulo 1

Introdução

1.1 Introdução

A filtragem adaptativa é uma área de conhecimento da engenharia na qual se tenta estimar os coeficientes de um filtro para manipular informações contidas no sinal de entrada de modo a conseguir uma resposta desejada em sua saída. Diferentemente dos filtros convencionais, que possuem seus coeficientes fixos, os filtros adaptativos têm seus coeficientes variáveis de modo a acompanhar mudanças no sistema, encontrar os valores ótimos de operação do filtro ou eliminar ruídos de canal, entre outros benefícios que os filtros convencionais não fornecem ou são de difícil implementação.

Filtros adaptativos normalmente não possuem especificações sobre o sistema, fazendo-se necessário um algoritmo adaptativo para reger o comportamento dos seus coeficientes, além de algumas informações relevantes, podendo ser sobre o comportamento do sinal de entrada ou o conhecimento prévio do sinal desejado, entre outras. Quando não se dispõe de informações sobre o ambiente em que se aplica o filtro adaptativo, como em modelagem inversa de canal em sistemas de telecomunicações, é comum a utilização de sinais de treinamento: envio de uma seqüência de dados conhecidos pelo receptor, de modo a fazer com que o filtro adaptativo encontre seus valores ótimos de operação antes de processar os sinais de interesse.

Uma maneira simples de interpretar a filtragem adaptativa digital é dada pela Figura 1.1. O sistema fornece o sinal de entrada do filtro, $u(n)$, que é ponderado pelos coeficientes do filtro adaptativo, gerando a saída $y(n)$. A saída é subtraída do

sinal desejado, $d(n)$, resultando num erro, $e(n)$, que normalmente é utilizado como um parâmetro de entrada do algoritmo adaptativo. O algoritmo ajusta os valores dos coeficientes do filtro, tornando-o, em princípio, cada vez mais próximo ao filtro com os coeficientes ótimos.

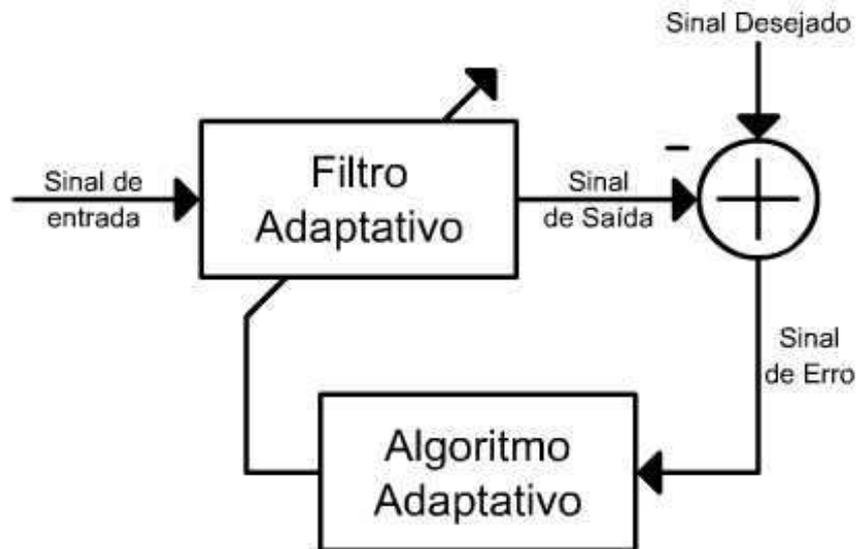


Figura 1.1: Diagrama de um sistema adaptativo.

As aplicações de filtros adaptativos são importantes em diversas áreas, tais como telecomunicações, sistemas de controle e outras. A necessidade de melhores e mais rápidas respostas em sistemas desconhecidos, ou com coeficientes variantes no tempo, tornam a área de processamento de sinais com uso de filtros adaptativos um importante campo de pesquisa.

Apesar das diferentes aplicações para os filtros adaptativos, estes possuem uma característica em comum: um vetor de entrada e um sinal de referência são utilizados para calcular uma estimativa do erro, utilizada para controlar os valores do conjunto de coeficientes.

A maneira como se extrai o sinal de referência é a diferença principal de como a filtragem adaptativa é utilizada em um sistema. Dessa maneira, podem-se distinguir quatro classes básicas de aplicações de filtragem adaptativa:

I. Identificação: Nesse tipo de aplicação o filtro adaptativo é utilizado para fornecer os parâmetros de um modelo que melhor represente, segundo algum critério pré-estabelecido, um sistema desconhecido. Na identificação de sistemas, o sinal de referência é a saída do sistema desconhecido, enquanto que o sinal de entrada para

o filtro adaptativo é o mesmo utilizado na entrada do sistema a ser identificado. A Figura 1.2 ilustra o funcionamento do filtro adaptativo para identificação de sistemas.

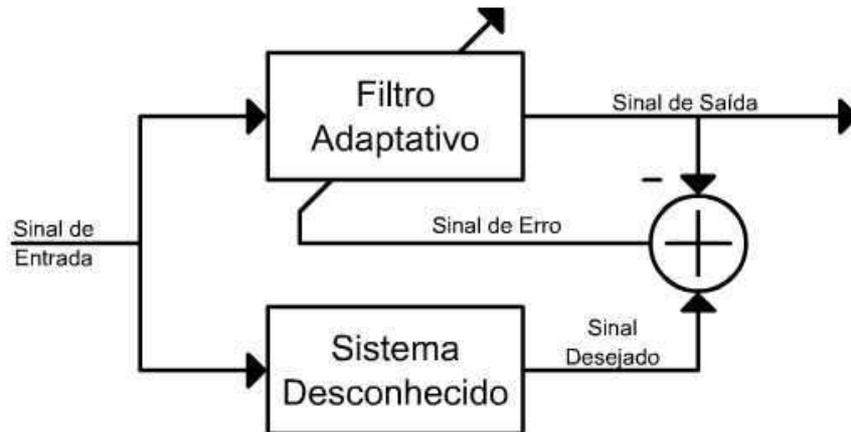


Figura 1.2: Filtragem adaptativa para identificação de sistema.

II. Modelagem Inversa: Também conhecida como equalização de canal, consiste na utilização do sinal que foi transmitido e distorcido por um canal como entrada do filtro adaptativo, e uma versão atrasada do sinal original é usada como sinal de referência. Em sistemas de comunicação, a versão atrasada do sinal de entrada está disponível no receptor na forma de sinal de treinamento. A função do filtro adaptativo é encontrar o melhor modelo inverso para reduzir ou anular a distorção provocada pelo canal. O diagrama em blocos do equalizador de canal é apresentado na Figura 1.3.

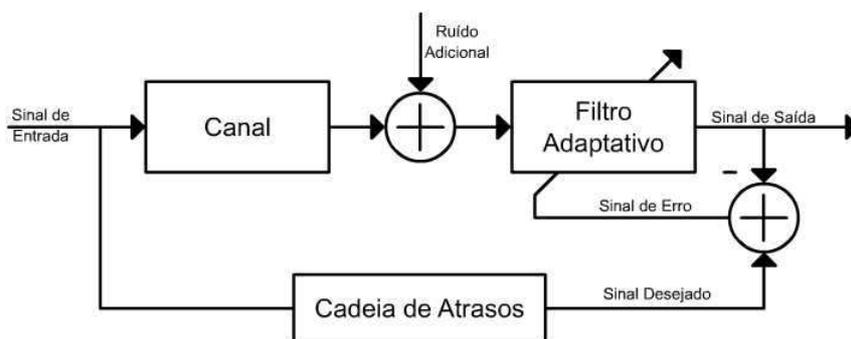


Figura 1.3: Filtragem adaptativa para equalização de canal.

III. Predição: A função do filtro adaptativo, neste caso, é fornecer a melhor predição do valor de uma seqüência. O sinal de referência é uma versão adiantada

(ou eventualmente atrasada) do sinal de entrada do filtro adaptativo, como ilustra a Figura 1.4. Os valores passados do sinal fornecem a entrada aplicada ao filtro adaptativo, e dependendo da aplicação de interesse, a saída do filtro adaptativo ou o erro estimado (predito) pode servir como saída do sistema. No caso de a saída do filtro ser a resposta do sistema, diz-se que o filtro adaptativo opera como um preditor, enquanto que no caso em que se usa o erro estimado como saída do sistema diz-se que o filtro adaptativo opera como um preditor de erro.

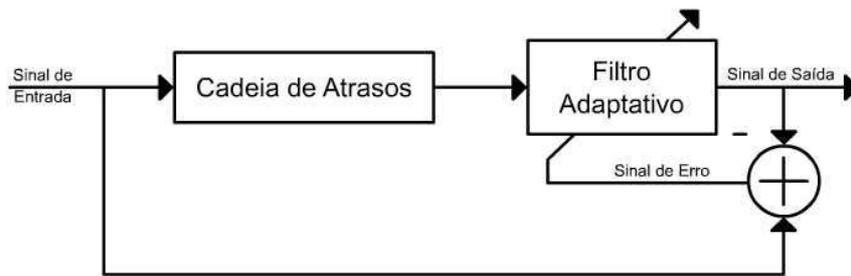


Figura 1.4: Filtragem adaptativa para predição de sinal.

IV. Cancelamento de Interferência: O filtro adaptativo pode ser usado para cancelar interferência desconhecida contida em um sinal. De uma maneira geral, o funcionamento do cancelamento de interferência se baseia na utilização de um sinal corrompido por interferência como sinal de referência, e de um sinal correlacionado com a interferência como entrada do filtro adaptativo. Após a convergência do sistema, o sinal de erro da saída será uma versão melhorada do sinal corrompido. A Figura 1.5 ilustra o funcionamento desse sistema.

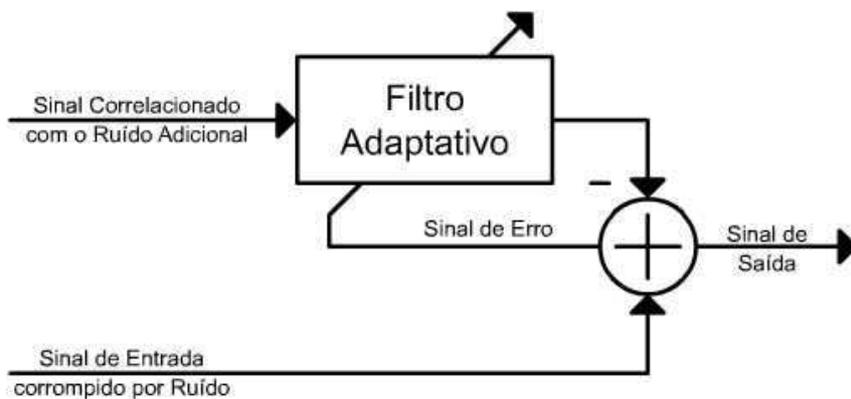


Figura 1.5: Filtragem adaptativa para melhoramento de sinal.

O critério para adaptação dos parâmetros do filtro influencia na complexidade

e qualidade da resposta ao problema, portanto precisa ser escolhido cuidadosamente. Usualmente, utiliza-se como critério em filtros adaptativos a minimização da função-objetivo dada pelo valor quadrático médio de um sinal de erro (MSE - *Mean Square Error*), dado pela diferença entre algum sinal de referência e a saída atual do filtro. O problema de minimização do MSE pode ser resolvido através da solução do sistema de equações de Wiener, de onde obtém-se a resposta ótima. O MSE é definido como

$$MSE = E \left\{ [d(n) - \mathbf{u}^T(n)\mathbf{w}]^2 \right\}, \quad (1.1)$$

onde $E \{ \cdot \}$ é o operador valor esperado.

O uso do MSE como função custo de um problema de filtragem adaptativa é comum, e algoritmos como o RLS (*Recursive Least Squares*) e o algoritmo LMS (*Least Mean Squares*) são exemplos de algoritmos que buscam o mínimo desta função-objetivo. A maneira como a função-objetivo é utilizada para derivar o algoritmo, as formas de aproximar as estatísticas dos sinais envolvidos e o método do desenvolvimento do algoritmo, entre outros detalhes, definem o algoritmo e determinar seu comportamento para diferentes tipos de sinais de entrada.

Existem vários métodos para se derivar algoritmos de filtragem adaptativa, que podem se basear em conceitos estocásticos ou determinísticos, ou até mesmo na formulação matemática de um sistema em um problema de otimização. Apesar da grande diversidade dos algoritmos iterativos que podem resultar da solução de um problema utilizando como função-objetivo o MSE, a maioria leva a uma resposta que tem relação direta com a resposta dada pelo filtro de Wiener.

1.2 Filtro de Wiener

A solução para o problema de minimização do MSE pode ser obtida sob uma visão estocástica. Esse método é utilizado no filtro de Wiener para encontrar o vetor de coeficientes ótimos, $\mathbf{w} = \mathbf{w}_o$, que minimizam a função-objetivo do MSE.

Após a diferenciação da função custo do MSE em relação aos elementos de \mathbf{w} e igualando a zero o resultado, obtemos

$$E \left\{ \mathbf{u}(n) [d(n) - \mathbf{u}^T(n)\mathbf{w}_o] \right\} = 0. \quad (1.2)$$

Pelo fato de a Equação (1.1) ser quadrática em relação a \mathbf{w} , a solução única é garantida se a segunda derivada for positiva. A equação acima é freqüentemente chamada de princípio da ortogonalidade [1], pois indica que para o conjunto ótimo de coeficientes do filtro \mathbf{w}_o , o erro, dado por $d(n) - \mathbf{u}^T(n)\mathbf{w}_o$, é ortogonal às M amostras do sinal de entrada, onde $\mathbf{u}^T(n) = [\mathbf{u}(n) \quad \mathbf{u}(n-1) \quad \cdots \quad \mathbf{u}(n-M+1)]$.

As equações de Wiener-Hopf levam à solução do problema de minimização da função custo do MSE, que em sua formulação matricial pode ser expressa em termos da matriz de autocorrelação do sinal de entrada, \mathbf{R} , e do vetor de correlação cruzada do sinal de entrada com o sinal desejado, \mathbf{p} . A solução é dada por

$$\mathbf{R}\mathbf{w}_o = \mathbf{p}, \quad (1.3)$$

onde

$$\mathbf{R} = E [\mathbf{u}(n)\mathbf{u}^T(n)] \quad (1.4)$$

e

$$\mathbf{p} = E [d(n)\mathbf{u}(n)]. \quad (1.5)$$

Para a solução das equações de Wiener-Hopf, assume-se que a matriz de autocorrelação do sinal de entrada, \mathbf{R} , seja definida positiva. Deste modo pode-se pré-multiplicar a Equação (1.3), em ambos os lados, por \mathbf{R}^{-1} , a inversa da matriz de autocorrelação. Como resposta, obtém-se a equação que descreve o filtro de Wiener,

$$\mathbf{w}_o = \mathbf{R}^{-1}\mathbf{p}. \quad (1.6)$$

1.3 O Algoritmo RLS

Dadas as seqüências $u(i)$ e $d(i)$, $i = 0, 1, \dots, n$, um método para solucionar o problema de otimização da função-objetivo é utilizar uma aproximação do valor esperado. Isso é realizado através de um somatório de modo que à medida que o número de iterações realizadas se torne muito elevado, ou seja, $n \rightarrow \infty$, a resposta do sistema seja, na média, uma boa aproximação para o filtro de Wiener.

O algoritmo RLS é freqüentemente visto de modo determinístico para a solução do problema de minimização da função-objetivo MSE. O desenvolvimento da solução deste problema, em sua forma mais simples, é dado por

$$\xi^d(n) = \sum_{i=1}^n \lambda^{n-i} e^2(i) = \sum_{i=1}^n \lambda^{n-i} [d(i) - \mathbf{x}^T(i)\mathbf{w}(n)]^2 \quad (1.7)$$

onde λ é o fator de esquecimento, usado para controlar a influência que as informações passadas têm sobre o desempenho do algoritmo e o sobrescrito $(\cdot)^d$ indica que a função-objetivo $\xi^d(n)$ foi definida segundo uma abordagem determinística do problema.

Diferenciando-se $\xi^d(n)$ em relação a $\mathbf{w}(n)$ e igualando-se a zero para encontrar a solução ótima, obtém-se

$$-\sum_{i=1}^n \lambda^{n-i} \mathbf{u}(i)\mathbf{u}^T(i)\mathbf{w}(n) + \sum_{i=1}^n \lambda^{n-i} \mathbf{u}(i)d(i) = \mathbf{0}, \quad (1.8)$$

onde $\mathbf{0}$ é um vetor de zeros. A expressão resultante para o vetor com os coeficientes ótimos é dada por

$$\mathbf{w}(n) = \left[\sum_{i=1}^n \lambda^{n-i} \mathbf{u}(i)\mathbf{u}^T(i) \right]^{-1} \sum_{i=1}^n \lambda^{n-i} \mathbf{u}(i)d(i) = \mathbf{R}_D^{-1}(n)\mathbf{p}_D(n). \quad (1.9)$$

Temos, como pode ser identificado comparando-se as Equações (1.9) e (1.6), uma relação direta entre a solução determinística do problema de mínimos quadrados e a solução das equações de Wiener-Hopf, se as aproximações para a matriz de autocorrelação do sinal de entrada e do vetor de correlação cruzada do sinal de entrada e do sinal de referência, respectivamente \mathbf{R}_D e \mathbf{p}_D , forem válidas. Pode-se comprovar que a solução determinística tende para a solução de Wiener se os sinais envolvidos são ergódicos [2].

1.4 Método de Otimização

A interpretação do problema do sistema adaptativo como um problema de otimização já é conhecido na literatura, porém seu uso é pouco explorado. A formulação de um sistema adaptativo como um problema de otimização pode ser de-

safiadora, entretanto as ferramentas para solução deste tipo de problema levam a excelentes resultados. O método de Newton é normalmente utilizado para resolver problemas de otimização sem restrições, onde a procura da solução é realizada em sentido descendente, com convergência extremamente rápida, podendo ser encontrada em apenas um passo. A busca pelo método de Newton é dada por:

$$x_{n+1} = x_n - t [\nabla^2 f(x_n)]^{-1} \nabla f(x_n) \quad (1.10)$$

Na expressão acima, t é o passo por procura em linha *backtracking*, com $x_n \in \text{dom} f$, e $f(x_n)$ é a função que se deseja minimizar. Se $t = 1$ a equação acima é usualmente chamada de método de Newton puro. A expressão multiplicada por t é chamada de passo de Newton, o qual é realizado em sentido descendente, a menos que x_n seja um ponto de ótimo local, quando será igual a zero [3].

O método de Newton, apesar de rápido, envolve um esforço computacional muito intenso devido ao uso da matriz Hessiana, $\nabla^2 f(x_n)$, e sua inversa, podendo ainda requerer deduções complexas tanto do gradiente quanto da Hessiana da função-objetivo. Existem outras alternativas para resolver problemas de otimização sem restrições, chamadas de métodos quasi-Newton. Esses métodos reduzem o esforço computacional para determinar a direção de procura do algoritmo simplificando o cálculo da Hessiana por uma aproximação, porém ainda compartilham algumas vantagens dos métodos Newton, como a rápida velocidade de convergência perto do ponto ótimo [3].

1.5 Objetivos e Contribuições desta Tese

A formulação de um sistema adaptativo como um problema de otimização oferece uma nova visão sobre a função-objetivo dos algoritmos adaptativos, além de possibilitar a inclusão de restrições para incluir alguma característica desejada, como por exemplo tornar o algoritmo mais robusto. Esta tese tem como objetivo apresentar um método para interpretação de algoritmos adaptativos, tratando-os como problemas de otimização. Em seguida, apresenta novos algoritmos derivados a partir desta formulação.

Primeiramente realizou-se um estudo dos algoritmos adaptativos existentes,

verificando seus pontos positivos e negativos. O algoritmo RLS destaca-se como padrão comparativo por ser um algoritmo clássico e por sua rápida convergência. Entretanto, o algoritmo RLS pode se tornar instável se sua matriz de autocorrelação deixar de ser positiva definida.

Outro algoritmo que recebe atenção nesta tese é baseado no método quasi-Newton. Este algoritmo possui uma matriz de autocorrelação do sinal de entrada que permanece sempre positiva definida, mesmo que o sinal de entrada não seja persistentemente excitante [4] e independente do sistema em que ele é aplicado. Porém, o custo deste benefício é o algoritmo ser mais complexo e normalizado, dificultando a formulação de métodos rápidos para sua implementação [4].

Apresentam-se no Capítulo 2 as idéias desenvolvidas até chegar-se à metodologia utilizada para formular as funções-objetivo de algoritmos adaptativos como problemas de otimização com descrição de distúrbio mínimo. O algoritmo RLS e um algoritmo baseado no método quasi-Newton são reformulados e desenvolvidos sob a visão de problemas de otimização.

No Capítulo 3 são apresentados outros algoritmos desenvolvidos utilizando-se a descrição por distúrbio mínimo proposto no Capítulo 2, demonstrando o potencial dessa nova ferramenta para o desenvolvimento de algoritmos adaptativos. Estuda-se, também, uma maneira para encontrar o passo ótimo para os algoritmos desenvolvidos, levando em consideração as restrições a eles aplicadas. Nesse capítulo os algoritmos são analisados através de simulações, apresentando-se seus pontos positivos e negativos.

Simulações comparativas são realizadas no Capítulo 4, onde podem-se analisar comportamento e desempenho dos novos algoritmos em relação ao algoritmo RLS e ao algoritmo quasi-Newton (QN) apresentados na primeira parte desta tese.

1.5.1 Contribuições desta Tese

Esta tese estuda o desenvolvimento de algoritmos adaptativos sob a ótica de otimização de um funcional. Este novo método possibilita novas interpretações do comportamento do algoritmo em relação a sua função-objetivo. Ela pode até mesmo não apresentar nenhuma interpretação física em sua forma original, como será visto no Capítulo 2 para um método quasi-Newton.

O método propõe uma modificação na função-objetivo, de modo que esta seja descrita como a função de norma mínima, ou de distúrbio mínimo, sujeita a restrições que são funções do erro *a posteriori* de saída. Quando restrições de desigualdade são utilizadas, é sugerido o uso do método da barreira logarítmica para simplificar a solução dos problemas.

Alguns benefícios obtidos pela formulação do problema de filtragem adaptativa através desse método podem estar na vantagem de se deduzir facilmente diversos algoritmos variando-se apenas poucos termos da função-objetivo, ou seja, modificando-se as restrições aplicadas ao problema de norma mínima. Outro benefício é a dissociação das funções-objetivo em relação ao somatório, facilitando a diferenciação da função-objetivo em relação ao fator de convergência, $\mu(n)$.

Capítulo 2

Filtragem Adaptativa como Problema de Otimização

No âmbito desta tese, um problema de otimização é um problema matemático onde se deseja encontrar a melhor dentre todas as soluções possíveis, dadas certas restrições ou limitações. Normalmente, se deseja encontrar o valor máximo (por exemplo, o lucro em investimentos) ou mínimo (por exemplo, o erro quadrático médio) de uma função-objetivo que representa algum tipo de sistema. A forma matemática usual dos problemas de otimização é dada por

$$\begin{aligned} \min_x \quad & f_0(x) \\ \text{sujeito a} \quad & f_i(x) < 0, \quad i = 1, \dots, m \\ & h_i(x) = 0, \quad i = 1, \dots, p, \end{aligned} \tag{2.1}$$

onde x é a variável de otimização do problema; as funções f_0 e f_i são, respectivamente, a função-objetivo e as m funções de restrições de desigualdade; e h_i são as p funções de restrições de igualdade. Em um problema de minimização, como o da Equação (2.1), um ponto x_0 é considerado ótimo, ou também chamado de solução do sistema, resultando no menor valor objetivo, isto é, o menor valor dentro do conjunto de soluções possíveis, dentre todos os valores que satisfazem as restrições [3].

Normalmente os problemas de otimização são divididos em famílias ou classes, dependendo das formas particulares da função-objetivo e de suas restrições.

Os problemas de otimização convexa têm a forma da Equação (2.1), e as funções $f_0, \dots, f_m : \mathbf{R}^n \rightarrow \mathbf{R}$ convexas, ou seja, satisfazendo

$$f_i(\alpha x + \beta y) \leq \alpha f_i(x) + \beta f_i(y) \quad (2.2)$$

para todo $x, y \in \mathbf{R}^n$ e todo $\alpha, \beta \in \mathbf{R}$ com $\alpha + \beta = 1$, $\alpha \geq 0$, $\beta \geq 0$.

A solução de problemas de otimização convexa é realizada dentro de uma região chamada de região de trabalho, que, sendo convexa, garante que o problema terá solução independente do método utilizado na busca. Não existe uma forma analítica geral para solucionar esse tipo de problema, porém há métodos muito efetivos para resolvê-los. A utilização de otimização convexa é, ao menos conceitualmente, parecida com os problemas de mínimos quadrados ou programação linear. Caso se consiga formular um problema adaptativo de um sistema como um problema de otimização convexa, então pode-se resolvê-lo exatamente como no caso de problemas de mínimos quadráticos. Entretanto, transformar um problema qualquer na forma de equações de um problema convexo pode ser uma atividade complexa e desafiadora [3].

2.1 Multiplicadores de Lagrange

Para a solução de problemas de otimização com restrições, uma ferramenta comumente usada são os multiplicadores de Lagrange. Entretanto existem outros métodos, como os baseados em pontos interiores, que apresentam excelentes resultados, reduzindo a complexidade na solução de um problema de otimização convexa.

Os multiplicadores de Lagrange são utilizados para a formulação da função dual de Lagrange, a qual transforma um problema com restrições em outro sem, utilizando as restrições da Equação (2.1) na função-objetivo como um somatório ponderado das funções de restrições [3]. Deste modo, a função Lagrangiana associada ao problema da Equação (2.1) é definida como

$$L(x, \lambda, \nu) = f_0 + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \quad (2.3)$$

onde λ_i são os multiplicadores de Lagrange associados com as restrições de desi-

igualdade e, similarmente, ν_i são os multiplicadores de Lagrange associados com as restrições de igualdade.

A função dual de Lagrange, ou somente função dual, $g(\lambda, \nu)$, é definida como o mínimo valor do Lagrangeano sobre x

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu). \quad (2.4)$$

A função dual fornece os limites inferiores do valor ótimo de x^* para o problema da Equação (2.1); pode-se demonstrar que essa formulação pode ser utilizada para se achar o valor ótimo de x^* [3].

2.2 Método de Pontos Interiores

Outros métodos para a solução de problemas de otimização convexa que merecem destaque são os métodos de pontos interiores. Estes métodos resolvem problemas de otimização convexa com restrições de igualdade e desigualdade lineares modificando o problema e reduzindo-o a uma seqüência de problemas com restrições lineares de igualdade. Os métodos de pontos interiores trabalham muito bem na prática, podendo resolver facilmente problemas com centenas de variáveis e milhares de restrições em um computador comum, em pouco tempo.

Dentre os métodos de pontos interiores, o método da barreira será utilizado aqui como ferramenta no auxílio do desenvolvimento de algoritmos, devido à sua simplicidade e aos bons resultados que apresenta. Neste método, as restrições de desigualdade são incluídas diretamente na função-objetivo, reduzindo a um novo problema de otimização com apenas restrições de igualdade, se houver.

Para a utilização deste método no auxílio à solução de problemas de otimização convexa, primeiramente devemos apresentar o problema na forma:

$$\begin{aligned} \min_x \quad & f_0(x) \\ \text{sujeito a} \quad & f_i(x) \leq 0, \quad i = 1, \dots, m \end{aligned} \quad (2.5)$$

onde $f_0, \dots, f_m : \mathbf{R}^n \rightarrow \mathbf{R}$ são funções convexas duplamente diferenciáveis.

O próximo passo é reescrever este problema, fazendo as restrições de desi-

gualdade explícitas na função-objetivo:

$$\min_x \quad f_0(x) + \sum_{i=1}^m I_-(f_i(x)) \quad (2.6)$$

onde $I_- : \mathbf{R} \rightarrow \mathbf{R}$ é a função indicadora dada pela equação:

$$I_-(u) = \begin{cases} 0 & u \leq 0 \\ \infty & u > 0 \end{cases} \quad (2.7)$$

A barreira se impõe no lugar da restrição, limitando a área de busca do algoritmo adaptativo dentro da superfície da função-objetivo a ser minimizada “corta” pela restrição imposta. Utilizando essa notação o problema descrito pela Equação (2.6) não possui mais restrições de desigualdade, mas sua função-objetivo ainda não é diferenciável, dificultando sua solução. Neste ponto, apresenta-se uma nova função diferenciável, que aproxima a Equação (2.7), dada por

$$\widehat{I}_-(u) = -(1/t)\log(-u), \quad \text{dom } \widehat{I}_- = -\mathbf{R}_{++} \quad (2.8)$$

onde $t > 0$ é o parâmetro que ajusta a precisão da aproximação. Esta nova formulação da função indicadora é chamada de barreira logarítmica. Assim como $I_-(u)$, $\widehat{I}_-(u)$ é convexa e não decrescente, e tende ao infinito para $u > 0$. A figura 2.1 apresenta a comparação entre a função indicadora e sua aproximação (barreira logarítmica), para alguns valores de t . Observa-se que à medida que t aumenta, a aproximação torna-se mais precisa [3].

Com essa nova função indicadora, o problema de otimização apresenta uma função-objetivo diferenciável e convexa:

$$\min_x \quad f_0(x) + \sum_{i=1}^m -(1/t)\log[-f_i(x)] \quad (2.9)$$

Entretanto, como essa função é apenas uma aproximação, a qualidade da solução do problema a ser minimizado depende do comportamento da mesma. Como o próprio gráfico sugere, quanto mais alto o valor de t mais parecido com a barreira ideal é a aproximação, e melhor é a resposta para o problema. Por outro lado, quando o parâmetro t é grande, a Hessiana varia rapidamente próximo ao limite

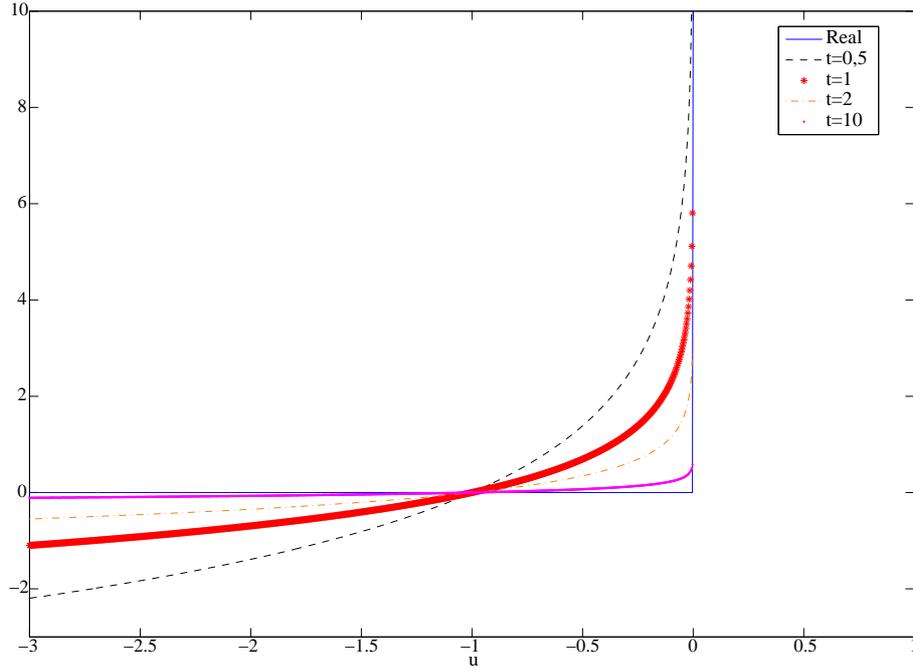


Figura 2.1: Função indicadora e sua aproximação pela barreira logarítmica.

do conjunto realizável. Uma maneira para eliminar esse obstáculo é resolver uma seqüência de problemas iguais ao da Equação (2.9), aumentando o valor de t a cada etapa, e utilizando a resposta anterior como ponto de partida para a próxima etapa.

O gradiente e a Hessiana para este novo problema são dados por

$$\nabla \xi(x) = \nabla f_0(x) + \frac{1}{t} \sum_{i=1}^m \frac{1}{-f_i(x)} \nabla f_i(x) \quad (2.10)$$

e

$$\nabla^2 \xi(x) = \nabla^2 f_0(x) + \frac{1}{t} \left[\sum_{i=1}^m \frac{1}{-f_i(x)^2} \nabla f_i(x) \nabla f_i(x)^T + \sum_{i=1}^m \frac{1}{-f_i(x)} \nabla^2 f_i(x) \right], \quad (2.11)$$

respectivamente.

2.3 Desenvolvimento de Algoritmos Adaptativos pelo Método da Barreira Logarítmica

Uma alternativa à derivação comumente realizada no desenvolvimento de algoritmos adaptativos pode ser obtida transformando-se um problema de filtragem adaptativa em um problema de otimização convexa com restrições. Uma vez que o problema de otimização esteja matematicamente formulado, pode-se resolvê-lo utilizando os multiplicadores de Lagrange ou o método da barreira logarítmica.

A utilização de restrições pode levar a diversos algoritmos adaptativos diferentes, partindo de uma mesma função-objetivo. Desse modo, desenvolveu-se uma primeira aproximação para se encontrar um novo método de tratar algoritmos adaptativos como problemas de otimização, utilizando na Equação (2.9), como função-objetivo, a norma quadrática do vetor de coeficientes ponderada pela iniciação da matriz de autocorrelação. Como restrição se utiliza uma exponencial negativa da função-objetivo do método adaptativo a ser tratado, mais precisamente

$$\begin{aligned} \min_{w(n)} \quad & \mathbf{w}(n)^T \mathbf{R}(0) \mathbf{w}(n) \\ \text{s.t.} \quad & -\exp[-f_1(x)] \leq 0, \end{aligned} \tag{2.12}$$

onde $f_1(x)$ é a função-objetivo do algoritmo adaptativo. A escolha da exponencial da função-objetivo determinística do algoritmo adaptativo, consegue resultar em um problema de otimização com a mesma função-objetivo devido a forma como a exponencial cancela a barreira logarítmica. Essa solução resultante para a derivação do mesmo algoritmo é viável pois foi levada em consideração diretamente a forma como a barreira logarítmica é interpretada.

Essa é uma maneira puramente didática para a formulação de problemas de filtragem adaptativa, uma vez que a exponencial atua de modo a cancelar a barreira logarítmica. Entretanto, utilizando essa formulação o desenvolvimento do algoritmo adaptativo passa a ser um problema de otimização com restrições que gera o mesmo algoritmo que seria obtido derivando-se diretamente a função-objetivo. Uma vantagem dessa nova formulação é que podemos incluir restrições para impor um comportamento desejado ao algoritmo. Pode-se, inclusive, utilizar restrições que

não envolvam o vetor de coeficientes do algoritmo, não influenciando na dedução do algoritmo em si, para impor limitações a outras variáveis utilizadas no algoritmo como, por exemplo, garantir explicitamente que a matriz de autocorrelação seja sempre positiva definida.

A seguir, desenvolvem-se dois algoritmos conhecidos através da transformação dos mesmos em problemas de otimização convexa e utilizando o método de barreira. Estes desenvolvimentos têm como objetivo chegar ao mesmo algoritmo conhecido, partindo da formulação do problema de acordo com a Equação (2.12).

2.3.1 Algoritmo RLS

Dentre os algoritmos adaptativos mais conhecidos, destacam-se os algoritmos LMS e RLS. O primeiro tem como principais características sua simplicidade e imunidade a ruído de medida. Por sua vez, o algoritmo RLS apresenta rápida convergência, mesmo quando o espalhamento dos autovalores da matriz de autocorrelação do sinal de entrada é grande. Entretanto, o algoritmo RLS pode apresentar alguns problemas de estabilidade.

O algoritmo RLS necessita da iniciação da matriz de autocorrelação do sinal de entrada, a qual requer um cuidado especial, uma vez que interfere diretamente no comportamento do algoritmo. Este tipo de algoritmo é apresentado usualmente em sua forma determinística, para minimização do erro quadrático médio *a posteriori*:

$$\xi(n) = \sum_{i=1}^n \lambda^{n-i} [d(i) - \mathbf{u}^T(i)\mathbf{w}(n)]^2 \quad (2.13)$$

A solução do problema da Equação (2.13) pode ser obtida derivando-se esta função-objetivo em relação ao vetor de coeficientes, $\mathbf{w}(n)$, e igualando-se o resultado a zero para se obter a resposta. Essa solução resulta na forma convencional como o algoritmo RLS é apresentado:

$$e(n) = d(n) - \mathbf{w}^T(n-1)\mathbf{u}(n) \quad (2.14)$$

$$\mathbf{R}^{-1}(n) = \frac{1}{\lambda} \left[\mathbf{R}^{-1}(n-1) - \frac{\mathbf{R}^{-1}(n-1)\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{R}^{-1}(n-1)}{\lambda + \mathbf{u}^T(n)\mathbf{R}^{-1}(n-1)\mathbf{u}(n)} \right] \quad (2.15)$$

$$\mathbf{w}(n) = \mathbf{w}(n-1) + e(n)\mathbf{R}^{-1}(n)\mathbf{u}(n), \quad (2.16)$$

onde $d(i)$ e $\mathbf{u}(i)$ são os sinais de referência e de entrada, respectivamente, e $\mathbf{w}(n)$ é o vetor de coeficientes. O parâmetro $\lambda \in (0, 1]$ é propositalmente chamado de fator de esquecimento devido ao seu controle sobre a ponderação das informações passadas na estimação de $\mathbf{R}(n)$ e $\mathbf{p}(n)$, e tem grande influência no desempenho do algoritmo, especialmente em ambiente não-estacionário [4].

Apesar da apresentação determinística do algoritmo RLS, sua solução no instante n é uma aproximação estocástica recursiva da solução de Wiener. No algoritmo RLS são feitas aproximações estocásticas da matriz de autocorrelação do sinal de entrada e do vetor de correlação-cruzada do sinal de referência com o vetor do sinal de entrada. Essas aproximações são obtidas pela média temporal dos dados observados, ponderando-se os dados passados com fatores de decréscimo exponencial [2].

A Equação (2.13) é um problema de otimização que pode ser tratado com o método de barreira apresentado anteriormente. Exatamente o mesmo algoritmo descrito pelas Equações (2.14)–(2.16) pode ser encontrado se modificarmos a função-objetivo de modo a descrever o seguinte problema de otimização convexa:

$$\begin{aligned} \min_{\mathbf{w}(n)} \quad & \mathbf{w}(n)^T \mathbf{R}(0) \mathbf{w}(n) \\ \text{s.t.} \quad & -\exp\left(-\left\{\sum_{i=1}^n \lambda^{n-i} [d(i) - \mathbf{u}^T(i)\mathbf{w}(n)]^2\right\}\right) \leq 0 \end{aligned} \quad (2.17)$$

Aplicando o método da barreira logarítmica, no qual usaremos $t = 1$ por comodidade para trabalhar com a equação, resulta em:

$$\xi(n) = \mathbf{w}(n)^T \mathbf{R}(0) \mathbf{w}(n) - \log \left[\exp\left(-\left\{\sum_{i=1}^n \lambda^{n-i} [d(i) - \mathbf{u}^T(i)\mathbf{w}(n)]^2\right\}\right) \right] \quad (2.18)$$

Calculando o gradiente em relação ao vetor de coeficientes nesta expressão e igualando-o a zero, pode-se obter a solução de Wiener aproximada do RLS.

2.3.2 Algoritmo QN

Existem na literatura alguns tipos de algoritmos de filtragem adaptativa quasi-Newton, *e.g.* [5], [6]. Apesar desses algoritmos serem classificados como do tipo quasi-Newton, eles são na verdade algoritmos do tipo Newton ou baseados no método de Newton. Para serem conhecidos como algoritmos quasi-Newton, na classificação de Fletcher [4], o algoritmo deve possuir a propriedade de hereditariedade, dada por

$$\mathbf{R}^{-1}(n)\gamma(\mathbf{n}) = \delta(\mathbf{n}), \quad (2.19)$$

onde $\gamma(n)$ e $\delta(n)$ são dados, respectivamente, por

$$\gamma(n) = \mathbf{g}(n) - \mathbf{g}(n-1) \quad (2.20)$$

e

$$\delta(n) = \mathbf{w}(n) - \mathbf{w}(n-1), \quad (2.21)$$

sendo $\mathbf{g}(n)$ e $\mathbf{g}(n-1)$ os gradientes calculados nos pontos $\mathbf{w}(n)$ e $\mathbf{w}(n-1)$, respectivamente. Como $\mathbf{g}(n)$, no contexto de filtros adaptativos, não é conhecido durante a iteração n , já que para isso seria necessário o conhecimento de $\mathbf{u}(n+1)$ e $d(n+1)$, em seu lugar é utilizado $\mathbf{g}'(n)$, um gradiente no ponto $\mathbf{w}(n)$ calculado com uso de $\mathbf{u}(n)$ e $d(n)$. Deste modo, temos

$$\mathbf{g}(n-1) = -2e(n)\mathbf{u}(n)\mathbf{g}'(n) = -2e_p(n)\mathbf{u}(n), \quad (2.22)$$

onde $e(n)$ e $e_p(n)$ são os erros *a priori* e *a posteriori*, respectivamente, dados por

$$e(n) = d(n) - \mathbf{w}^T(n-1)\mathbf{u}(n)e_p(n) = d(n) - \mathbf{w}^T(n)\mathbf{u}(n) \quad (2.23)$$

Dentre os algoritmos que se enquadram na definição de Fletcher, aborda-se nesta tese o algoritmo proposto em [4]. Este algoritmo mostra-se estável mesmo em situações nas quais o algoritmo RLS torna-se instável, alcançando velocidade de convergência similar.

O desenvolvimento do algoritmo é baseado na aproximação quasi-Newton

padrão de posto um, no conceito do erro *a posteriori* e na busca exata em linha. O resultado é um algoritmo que não precisa de parâmetro de ajuste, e apenas a utilização de um valor inicial na matriz de autocorrelação de modo a garantir que as mudanças no vetor de coeficientes estejam sempre na direção descendente. Este algoritmo pode ser descrito como:

Passo 1: Calcular a direção de atualização do vetor de coeficientes $\mathbf{w}(n-1)$:

$$\mathbf{h}(n) = -\mathbf{R}_{QN}^{-1}(n)\mathbf{g}(n-1) \quad (2.24)$$

onde $\mathbf{g}(n-1)$ é o gradiente avaliado no ponto $\mathbf{w}(n-1)$ e $\mathbf{R}_{QN}(n-1)$ é uma estimativa da matriz de autocorrelação do sinal de entrada.

Passo 2: Calcular o valor do fator de convergência (tamanho do passo), $\mu(n)$, realizar a procura em linha em $\mathbf{h}(n)$ e encontrar

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu(n)\mathbf{h}(n). \quad (2.25)$$

Passo 3: Atualizar $\mathbf{R}_{QN}^{-1}(n-1)$ e obter $\mathbf{R}_{QN}^{-1}(n)$ como uma melhor estimativa da matriz de autocorrelação do sinal de entrada.

Utilizando essa definição, apresenta-se em [4] o desenvolvimento deste algoritmo. Destaca-se a forma como a inversa da matriz de autocorrelação do sinal de entrada é atualizada, dada pela equação abaixo.

$$\mathbf{R}_{QN}^{-1}(n) = \mathbf{R}_{QN}^{-1}(n-1) + \frac{[\delta(n) - \mathbf{R}_{QN}^{-1}(n-1)\gamma(n)] [\delta(n) - \mathbf{R}_{QN}^{-1}(n-1)\gamma(n)]^T}{[\mathbf{u}^T(n)\mathbf{R}_{QN}^{-1}(n-1)\mathbf{u}(n)]^T \gamma(n)} \quad (2.26)$$

Desta forma, garante-se a condição de quasi-Newton, no conceito de Fletcher, independentemente de como os vetores-gradiente são calculados.

O fator de convergência é dado por:

$$\mu(n) = \frac{1}{2\mathbf{u}^T(n)\mathbf{R}_{QN}^{-1}(n-1)\mathbf{u}(n)}. \quad (2.27)$$

Utilizando esses dados e realizando algumas manipulações matemáticas, simplifica-

se a atualização da inversa da Hessiana para:

$$\mathbf{R}_{QN}^{-1}(n) = \mathbf{R}_{QN}^{-1}(n-1) + [\mu(n) - 1] \frac{\mathbf{R}_{QN}^{-1}(n-1)\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{R}_{QN}^{-1}(n-1)}{2\mathbf{u}^T(n)\mathbf{R}_{QN}^{-1}(n-1)\mathbf{u}(n)}. \quad (2.28)$$

Assim, esse algoritmo quasi-Newton pode ser resumido como segue.

Dados $\mathbf{u}(n)$ e $d(n)$, calcule:

$$e(n) = d(n) - \mathbf{w}^T(n-1)\mathbf{u}(n) \quad (2.29)$$

$$\mathbf{t}(n) = \mathbf{R}_{QN}^{-1}(n-1)\mathbf{u}(n) \quad (2.30)$$

$$\tau(n) = \mathbf{u}^T(n)\mathbf{t}(n) \quad (2.31)$$

$$\mu(n) = \frac{1}{2\tau(n)} \quad (2.32)$$

$$\mathbf{R}_{QN}^{-1}(n) = \mathbf{R}_{QN}^{-1}(n-1) + \frac{[\mu(n) - 1]}{\tau(n)} \mathbf{t}(n)\mathbf{t}^T(n) \quad (2.33)$$

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \frac{e(n)}{\tau(n)} \mathbf{t}(n). \quad (2.34)$$

Realizando o processo inverso, utilizando o lema da inversão de matriz, pode-se obter a equação da matriz de autocorrelação do sinal de entrada.

$$\mathbf{R}_{QN}(n) = \mathbf{R}_{QN}(n-1) + 2[1 - \mu(n)]\mathbf{u}(n)\mathbf{u}^T(n) \quad (2.35)$$

Através de algumas manipulações matemáticas, é fácil obter a estimativa do vetor de correlação cruzada entre o vetor de entrada e o sinal de referência.

$$\mathbf{p}_{QN}(n) = \mathbf{p}_{QN}(n-1) + 2[d(n) - \mu(n)\mathbf{u}^T(n)\mathbf{w}(n-1)]\mathbf{u}(n) \quad (2.36)$$

De posse de \mathbf{R}_{QN} e \mathbf{p}_{QN} , pode-se substituí-los na equação do filtro de Wiener

para encontrar a seguinte função-objetivo do algoritmo quasi-Newton [4]:

$$\begin{aligned} \xi(n) = & \sum_{i=1}^n [1 - \mu(i)] \left[\frac{d(i) - \mu(i) \mathbf{u}^T(i) \mathbf{w}(i-1)}{1 - \mu(i)} - \mathbf{u}^T(i) \mathbf{w}(n) \right]^2 \\ & + \frac{1}{2} \mathbf{w}^T(n) \mathbf{R}_{QN}(0) \mathbf{w}(n) \end{aligned} \quad (2.37)$$

Esta função-objetivo compartilha pouca informação física sobre o comportamento do algoritmo em si, e é relativamente complexa. Para buscar algumas informações adicionais sobre o comportamento da função-objetivo do algoritmo quasi-Newton, utiliza-se o método de barreira, de modo que este produza ao final o mesmo algoritmo quasi-Newton. O problema de otimização convexa, neste caso, torna-se:

$$\begin{aligned} \min_{\mathbf{w}(n)} \quad & \frac{1}{2} \mathbf{w}(n)^T \mathbf{R}_{QN}(0) \mathbf{w}(n) \\ \text{s.t.} \quad & -\exp \left(- \left\{ \sum_{i=1}^n [1 - \mu(i)] \left[\frac{d(i) - \mu(i) \mathbf{u}^T(i) \mathbf{w}(i-1)}{1 - \mu(i)} - \mathbf{u}^T(i) \mathbf{w}(n) \right]^2 \right\} \right) \leq 0 \end{aligned} \quad (2.38)$$

Aplicando o método de barreira a esta função-objetivo e calculando o gradiente em relação ao vetor de coeficientes, $\mathbf{w}(n)$, obtém-se:

$$\nabla \xi(n) = \mathbf{R}_{QN}(0) \mathbf{w}(n) + 2 \sum_{i=1}^n [d(i) - \mu(i) \mathbf{u}^T(i) \mathbf{w}(i-1) + \mathbf{u}^T(i) \mathbf{w}(n)] \mathbf{u}(i) = 0 \quad (2.39)$$

$$\mathbf{R}_{QN}(n) \mathbf{w}(n) - \mathbf{p}_{QN}(n) = 0, \quad (2.40)$$

onde $\mathbf{R}_{QN}(n)$ e $\mathbf{p}_{QN}(n)$ são dados pelas Equações (2.35) e (2.36), respectivamente. A dedução a partir deste ponto segue igual ao desenvolvimento estocástico realizado na dedução do algoritmo RLS, observando-se a diferença nas atualizações da matriz de autocorrelação e do vetor de correlação cruzada.

2.4 Desenvolvimento de Algoritmos Adaptativos Usando a Descrição de Distúrbio Mínimo

O desenvolvimento de algoritmos usando o método da barreira logarítmica foi a primeira tentativa de prover uma nova interpretação em relação à função-

objetivo de sistemas adaptativos. Entretanto, esse método, apresentado na seção anterior, fornece pouca informação adicional sobre o comportamento do algoritmo. Talvez a melhor informação que se possa obter usando esse método é a dependência do algoritmo em relação à iniciação da matriz de autocorrelação. Além disso, o problema de otimização ainda possui a vantagem de poder agregar qualquer restrição desejada ao algoritmo.

Uma outra metodologia foi desenvolvida para prover melhor interpretação em relação à função-objetivo do problema de otimização. Este método utiliza como função-objetivo uma descrição de distúrbio mínimo dada pela norma quadrática do distúrbio entre os coeficientes das iterações n e $n - 1$, ponderada por uma matriz definida positiva \mathbf{A} , como apresentado em [7]. A descrição por distúrbio mínimo já foi utilizada antes para derivar e analisar algoritmos, e.g. [8]; porém, nesta tese, será utilizada de uma maneira ligeiramente diferente. Matematicamente, a função de distúrbio mínimo é dada por

$$\min_{\mathbf{w}(n)} \|\mathbf{w}(n) - \mathbf{w}(n-1)\|_{\mathbf{A}}^2, \quad (2.41)$$

onde a norma ponderada é dada por:

$$\|\mathbf{w}(n) - \mathbf{w}(n-1)\|_{\mathbf{A}}^2 = [\mathbf{w}(n) - \mathbf{w}(n-1)]^T \mathbf{A} [\mathbf{w}(n) - \mathbf{w}(n-1)]. \quad (2.42)$$

A escolha do distúrbio mínimo como função-objetivo a ser otimizada é boa opção pois age de forma a evitar modificações bruscas nos coeficientes durante a convergência, bem como minimiza a diferença do erro entre o vetor de coeficientes e os coeficientes ótimos. A matriz que pondera a descrição por distúrbio mínimo age de forma a modelar a superfície a ser minimizada, podendo tornar a procura pelo ótimo mais rápida ou lenta, de acordo com a escolha realizada. Utilizando como restrição funções de igualdade e desigualdade, e ponderando tanto as restrições quanto a função-objetivo da Equação (2.41), podemos chegar a diversos algoritmos adaptativos. Os algoritmos RLS e QN, derivados nas seções anteriores deste capítulo, serão reapresentados como problemas de otimização sob esta ótica alternativa. O uso dessa abordagem oferece uma nova interpretação do comportamento da nova função-objetivo, principalmente no caso do algoritmo QN.

2.4.1 Algoritmo RLS

O algoritmo RLS pode ser derivado a partir de um problema de otimização com descrição de distúrbio mínimo, obtendo-se exatamente o mesmo algoritmo descrito pelas Equações (2.14)–(2.16). A função-objetivo utiliza o método proposto de descrição de distúrbio mínimo, Equação (2.41), ponderada por um “fator de convergência”, o qual pode ser feito equivalente ao fator de esquecimento λ para levar ao mesmo algoritmo RLS. A função-objetivo ainda leva em consideração o quadrado do erro *a posteriori* instantâneo.

Para a derivação do algoritmo RLS, a matriz \mathbf{A} que pondera a norma do distúrbio dos coeficientes é igual à matriz de autocorrelação do sinal de entrada, $\mathbf{R}(n-1)$. Essa nomenclatura da matriz é apenas para facilitar a visualização da relação entre a matriz de autocorrelação encontrada no desenvolvimento do algoritmo RLS pelo modo convencional com o método proposto, como será demonstrado na dedução do algoritmo realizada a seguir.

$$\min_{\mathbf{w}(n)} \lambda \|\mathbf{w}(n) - \mathbf{w}(n-1)\|_{\mathbf{R}(n-1)}^2 + [d(n) - \mathbf{u}^T(n)\mathbf{w}(n)]^2. \quad (2.43)$$

A diferenciação da Equação (2.43) em relação aos coeficientes do filtro adaptativo, igualando-se o resultado a zero para resolver o problema de minimização, resulta em

$$2\lambda\mathbf{R}(n-1) [\mathbf{w}(n) - \mathbf{w}(n-1)] - 2d(n)\mathbf{u}(n) + 2\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{w}(n) = 0 \quad (2.44)$$

$$[\lambda\mathbf{R}(n-1) + \mathbf{u}(n)\mathbf{u}^T(n)] \mathbf{w}(n) = \lambda\mathbf{R}(n-1)\mathbf{w}(n-1) + d(n)\mathbf{u}(n) \quad (2.45)$$

$$\mathbf{R}(n)\mathbf{w}(n) = [\mathbf{R}(n) - \mathbf{u}(n)\mathbf{u}^T(n)] \mathbf{w}(n-1) \quad (2.46)$$

$$\mathbf{w}(n) = \mathbf{w}(n-1) + e(n)\mathbf{R}^{-1}(n)\mathbf{u}(n), \quad (2.47)$$

onde $e(n)$ é o erro *a priori*, e a matriz de autocorrelação é obtida no desenvolvimento do algoritmo pelos elementos dentro dos colchetes no lado esquerdo da Equação

(2.45),

$$\mathbf{R}(n) = \lambda \mathbf{R}(n-1) + \mathbf{u}(n)\mathbf{u}^T(n). \quad (2.48)$$

Desse modo, a inversa da matriz de autocorrelação do sinal de entrada pode ser encontrada utilizando-se o lema de inversão de matrizes, resultando na mesma inversa da matriz do algoritmo RLS, ou seja,

$$\mathbf{R}^{-1}(n) = \frac{1}{\lambda} \left[\mathbf{R}^{-1}(n-1) - \frac{\mathbf{R}^{-1}(n-1)\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{R}^{-1}(n-1)}{\lambda + \mathbf{u}^T(n)\mathbf{R}^{-1}(n-1)\mathbf{u}(n)} \right]. \quad (2.49)$$

2.4.2 Algoritmo QN

O método de descrição por distúrbio mínimo favorece a criação de novos algoritmos, bem como a reformulação da função-objetivo de algoritmos conhecidos. Assim como no caso do algoritmo RLS, o algoritmo QN descrito pelas Equações (2.29)–(2.34) pode ser deduzido a partir de um problema de otimização com descrição por distúrbio mínimo. Para isso, a matriz \mathbf{A} novamente é igual à matriz de autocorrelação do sinal de entrada, $\mathbf{R}_{QN}(n-1)$, porém neste caso a matriz precisa ser igual àquela obtida no algoritmo QN, para se respeitar a restrição $\mathbf{u}^T(n)\mathbf{R}_{QN}^{-1}(n)\mathbf{u}(n) = 1/2$, que fornece suas boas propriedades numéricas. O problema a ser otimizado é dado por

$$\begin{aligned} \min_{\mathbf{w}(n)} \quad & \|\mathbf{w}(n) - \mathbf{w}(n-1)\|_{\mathbf{R}_{QN}(n-1)}^2 \\ \text{s.t.} \quad & \begin{cases} d(n) = \mathbf{w}^T(n)\mathbf{u}(n) \\ \mathbf{u}^T(n)\mathbf{R}_{QN}^{-1}(n)\mathbf{u}(n) = 1/2. \end{cases} \end{aligned} \quad (2.50)$$

Utilizando os multiplicadores de Lagrange para resolver a equação acima, tem-se

$$\xi(n) = \|\mathbf{w}(n) - \mathbf{w}(n-1)\|_{\mathbf{R}_{QN}(n-1)}^2 + \nu [d(n) - \mathbf{w}^T(n)\mathbf{u}(n)]. \quad (2.51)$$

Derivando a Equação (2.51) em relação aos coeficientes do filtro adaptativo

e igualando o resultado a zero, obtém-se

$$\nabla \xi(n) = 2\mathbf{R}_{QN}(n-1) [\mathbf{w}(n) - \mathbf{w}(n-1)] + \nu \mathbf{u}(n) = 0 \quad (2.52)$$

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \frac{\nu \mathbf{R}_{QN}^{-1}(n-1) \mathbf{u}(n)}{2} \quad (2.53)$$

Substituindo a Equação (2.53) na restrição, descobre-se quem é o multiplicador de Lagrange ν .

$$d(n) = \mathbf{u}^T(n) \left[\mathbf{w}(n-1) + \frac{\nu \mathbf{R}_{QN}^{-1}(n-1) \mathbf{u}(n)}{2} \right] \quad (2.54)$$

$$\frac{\nu \mathbf{u}^T \mathbf{R}_{QN}^{-1}(n-1) \mathbf{u}(n)}{2} = d(n) - \mathbf{u}^T(n) \mathbf{w}(n-1) \quad (2.55)$$

$$\nu = \frac{2e(n)}{\tau(n)} \quad (2.56)$$

onde $\tau(n)$ é dado como na Equação (2.31).

Agora, substituindo a equação (2.56) em (2.53) para obter a solução do problema de otimização resulta em

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \frac{e(n) \mathbf{R}_{QN}^{-1}(n-1) \mathbf{u}(n)}{\tau(n)}, \quad (2.57)$$

que é a mesma equação de atualização do algoritmo QN. Como pode ser observado, a restrição imposta sobre a matriz de autocorrelação do sinal de entrada não interfere na dedução do algoritmo em si, entretanto, a matriz que satisfaz essa restrição é dada pela Equação (2.28). Desse modo, o mesmo algoritmo QN apresentado em [4] é obtido pela descrição de distúrbio mínimo, Equação (2.50), a que possui uma função-objetivo mais atraente.

2.5 Contribuições do Capítulo

O Capítulo 2 traz algumas contribuições interessantes à interpretação dos algoritmos adaptativos. O método de descrição pela barreira logarítmica deixa claro que algoritmos baseados na função de mínimos quadráticos (*least squares*) têm dependência da iniciação da matriz de autocorrelação. Essa iniciação pode polarizar a resposta do algoritmo, bem como desempenhar um papel importante no seu desempenho. Esse fato é brevemente discutido no Capítulo 4, durante as simulações.

Apesar de o desenvolvimento de algoritmos adaptativos utilizando a função objetivo de distúrbio mínimo já ter sido explorado anteriormente para o desenvolvimento do algoritmo LMS [8], o estudo realizado nesta tese segue um propósito diferente: primeiramente, por utilizar a função-objetivo de distúrbio mínimo no contexto de problema de otimização com restrições a serem agregadas; e além disso, por aproveitá-la como uma função-objetivo simples que permite desenvolver diversos tipos de algoritmos adaptativos. A descrição através da norma do distúrbio do vetor de coeficientes ponderada por uma matriz positiva foi utilizada para desenvolver dois algoritmos da literatura, o algoritmo RLS e o algoritmo QN. O próximo capítulo estudará algumas possibilidades que essa generalização da função-objetivo pode oferecer para o desenvolvimento de algoritmos adaptativos.

Capítulo 3

Novos Algoritmos

A metodologia para o desenvolvimento de algoritmos adaptativos como problemas de otimização apresenta vantagens tanto para o desenvolvimento de algoritmos novos, para o estabelecimento de uma nova forma de interpretação de algoritmos conhecidos. A descrição por distúrbio mínimo com restrições tem como principais vantagens oferecer uma nova ótica para interpretação da função-objetivo de algoritmos adaptativos e a dissociação do passo do algoritmo em relação a informações passadas. Para demonstrar esse ponto de vista, desenvolvem-se neste capítulo alguns algoritmos novos utilizando a abordagem do capítulo anterior e, sempre que necessário aplica-se a barreira logarítmica para resolver problemas com restrições de desigualdade.

3.1 Algoritmo 1

Uma outra forma possível e mais simples da função-objetivo de mínimos quadrados recursiva propõe a minimização recursiva de uma função-objetivo baseada na soma dos quadrados dos erros *a posteriori*:

$$\xi(n) = \sum_{i=1}^n \mu(i) [d(i) - \mathbf{u}^T(i)\mathbf{w}(n)]^2. \quad (3.1)$$

O gradiente dessa função-objetivo pode ser calculado em relação aos coefici-

entes do filtro adaptativo; e igualado-O a zero para encontrar a solução, obtêm-se

$$\nabla \xi(n) = - \sum_{i=1}^n \mu(i) d(i) \mathbf{u}(i) + \sum_{i=1}^n \mu(i) \mathbf{u}(i) \mathbf{u}^T(i) \mathbf{w}(n) = 0. \quad (3.2)$$

Podem-se escrever recursivamente as aproximações estocásticas baseadas em amostras da matriz de autocorrelação do sinal de entrada, $\widehat{\mathbf{R}}(n)$, e do vetor de correlação cruzada do sinal de entrada com o sinal de referência, $\widehat{\mathbf{p}}(n)$, respectivamente, como

$$\widehat{\mathbf{R}}(n) = \sum_{i=1}^n \mu(i) \mathbf{u}(i) \mathbf{u}^T(i) = \widehat{\mathbf{R}}(n-1) + \mu(n) \mathbf{u}(n) \mathbf{u}^T(n) \quad (3.3)$$

e

$$\widehat{\mathbf{p}}(n) = \sum_{i=1}^n \mu(i) d(i) \mathbf{u}(i) = \widehat{\mathbf{p}}(n-1) + \mu(n) d(n) \mathbf{u}(n). \quad (3.4)$$

Desse modo, obtêm-se o conjunto de equações que resolvem o problema de otimização, que possuem a forma da equação de Wiener-Hopf [1], e a atualização dos coeficientes do filtro adaptativo é dada por

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu(n) e(n) \widehat{\mathbf{R}}^{-1}(n) \mathbf{u}(n). \quad (3.5)$$

A inversa da matriz de autocorrelação é encontrada aplicando-se o lema de inversão de matrizes [1] na Equação (3.3), o que resulta em

$$\widehat{\mathbf{R}}^{-1}(n) = \widehat{\mathbf{R}}^{-1}(n-1) - \frac{\widehat{\mathbf{R}}^{-1}(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \widehat{\mathbf{R}}^{-1}(n-1)}{\frac{1}{\mu(n)} + \mathbf{u}^T(n) \widehat{\mathbf{R}}^{-1}(n-1) \mathbf{u}(n)}. \quad (3.6)$$

A utilização de uma função-objetivo baseada na descrição por distúrbio mínimo para se obter o mesmo algoritmo, conduz ao seguinte problema de otimização:

$$\min_{\mathbf{w}(n)} \|\mathbf{w}(n) - \mathbf{w}(n-1)\|_{\mathbf{R}_{A1}(n-1)}^2 + \mu(n) [d(n) - \mathbf{u}^T(n) \mathbf{w}(n)]^2, \quad (3.7)$$

onde o subscrito $A1$ identifica o algoritmo 1.

Resolvendo o problema da Equação (3.7), tem-se

$$2\mathbf{R}_{A1}(n-1) [\mathbf{w}(n)^T - \mathbf{w}(n-1)] + 2\mu(n) [-d(n) + \mathbf{u}^T(n) \mathbf{w}(n)] \mathbf{u}(n) = 0 \quad (3.8)$$

$$[\mathbf{R}_{A1}(n-1) + \mu(n)\mathbf{u}(n)\mathbf{u}^T(n)] \mathbf{w}(n) = \mathbf{R}_{A1}(n-1)\mathbf{w}(n-1) + \mu(n)d(n)\mathbf{u}(n) \quad (3.9)$$

$$\mathbf{R}_{A1}(n)\mathbf{w}(n) = [\mathbf{R}_{A1}(n) - \mu(n)\mathbf{u}(n)\mathbf{u}^T(n)] \mathbf{w}(n-1) + \mu(n)d(n)\mathbf{u}(n) \quad (3.10)$$

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu(n) [d(n) - \mathbf{w}^T(n-1)\mathbf{u}(n)] \mathbf{R}_{A1}^{-1}(n)\mathbf{u}(n) \quad (3.11)$$

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu(n)e(n)\mathbf{R}_{A1}^{-1}(n)\mathbf{u}(n). \quad (3.12)$$

Pode-se perceber pelo elemento entre colchetes no lado esquerdo da Equação (3.9) a equivalência entre a matriz de autocorrelação do sinal de entrada utilizada na atualização dos coeficientes do filtro adaptativo da Equação (3.9) e a Equação (3.3). Desse modo, obtém-se o mesmo algoritmo derivado da Equação (3.1), partindo-se de uma função-objetivo baseada na descrição por distúrbio mínimo, Equação (3.7), com mesma inversa da matriz de autocorrelação.

Quando $\mu(n) = 1$ este algoritmo é igual ao algoritmo RLS convencional. Entretanto, apesar de ser uma maneira geral de apresentar a função-objetivo de mínimos quadráticos recursiva, a estratégia utilizada para desenvolver este algoritmo tem uma habilidade limitada de esquecer informações passadas, não realizando a ponderação exponencial do algoritmo RLS convencional. No presente algoritmo, a cada instante n , o coeficiente $\mu(n)$ age somente sobre o erro *a posteriori* do momento atual, como pode ser verificado pela Equação (3.7).

O algoritmo 1 tem sua função-objetivo semelhante à do algoritmo BEACON (*Bounding Ellipsoidal Adaptive Constrained*) *least-squares*, o qual utiliza algumas restrições relativas a filtros *set-membership*, que atualizam os parâmetros do filtro apenas dentro da região realizável (*feasibility set*). O algoritmo BEACON utiliza técnicas de algoritmos OBE (*optimal bounding ellipsoid*) para realizar a verificação do conjunto realizável, necessária para definir as especificações de desempenho de filtros *set-membership*. Dessa forma, o algoritmo BEACON compartilha algumas características desejáveis de algoritmos OBE; porém, realiza atualizações com menor frequência [9]. A descrição por distúrbio mínimo pode ser utilizada para desenvolver o algoritmo BEACON, partindo-se da mesma função-objetivo do algoritmo 1 e incluindo-se restrições referentes à atualização do algoritmo.

3.2 Algoritmo 2

Levando em consideração as funções-objetivo obtidas pela descrição de distúrbio mínimo dos algoritmos recursivos baseados em mínimos quadrados, uma modificação natural da função-objetivo é torná-la uma combinação convexa entre o distúrbio mínimo dos coeficientes e o erro *a posteriori*, ou seja,

$$\xi(n) = [1 - \mu(n)] \|\mathbf{w}(n) - \mathbf{w}(n-1)\|_{\mathbf{R}_{A2}(n-1)}^2 + \mu(n) [d(n) - \mathbf{u}^T(n)\mathbf{w}(n)]^2, \quad (3.13)$$

onde o subscrito $A2$ se refere ao algoritmo 2.

Calculando-se a derivada da Equação (3.13) com relação a $\mathbf{w}(n)$ e igualando-a a zero, obtém-se

$$\begin{aligned} \nabla \xi(n) = & 2[1 - \mu(n)] \mathbf{R}_{A2}(n-1) [\mathbf{w}(n) - \mathbf{w}(n-1)] \\ & + 2\mu(n) [-d(n) + \mathbf{u}^T(n)\mathbf{w}(n)] \mathbf{u}(n) = 0. \end{aligned} \quad (3.14)$$

Pode-se resolver a equação em função dos coeficientes do filtro adaptativo, o que resulta resulta em,

$$\begin{aligned} \{[1 - \mu(n)] \mathbf{R}_{A2}(n-1) + \mu(n)\mathbf{u}(n)\mathbf{u}^T(n)\} \mathbf{w}(n) = & [1 - \mu(n)] \mathbf{R}_{A2}(n-1)\mathbf{w}(n-1) \\ & + \mu(n)d(n)\mathbf{u}(n). \end{aligned} \quad (3.15)$$

Com isso, pode-se escrever recursivamente a matriz de autocorrelação do sinal de entrada como

$$\mathbf{R}_{A2}(n) = [1 - \mu(n)] \mathbf{R}_{A2}(n-1) + \mu(n)\mathbf{u}(n)\mathbf{u}^T(n), \quad (3.16)$$

obter a atualização recursiva do vetor de coeficientes como

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu(n)e(n)\mathbf{R}_{A2}^{-1}(n)\mathbf{u}(n) \quad (3.17)$$

e escrever recursivamente o vetor de correlação cruzada do sinal de entrada com o

sinal de referência como

$$\mathbf{p}_{A2}(n) = [1 - \mu(n)] \mathbf{p}_{A2}(n-1) + \mu(n)d(n)\mathbf{u}(n). \quad (3.18)$$

Um ponto interessante a ressaltar é que quando $1 - \mu(n) = \lambda < 1$ este algoritmo, o qual já é conhecido com o nome de LMS-Newton em [10], pode se tornar equivalente ao algoritmo RLS. De fato, o mesmo algoritmo pode ser obtido minimizando-se

$$\xi(n) = \sum_{i=1}^n \prod_{j=i+1}^n [1 - \mu(j)] \mu(i) [d(i) - \mathbf{u}^T(n)\mathbf{w}(n)]^2 \quad (3.19)$$

com relação ao vetor de coeficientes $\mathbf{w}(n)$.

3.3 Algoritmo 3

Outra maneira de se utilizar a descrição por distúrbio mínimo é incluir as restrições explicitamente no problema de otimização. Um algoritmo que pode ser derivado desse modo é o algoritmo de média quadrática mínima normalizada (NLMS), onde a matriz \mathbf{A} que pondera a norma de distúrbio mínimo dos coeficientes do filtro adaptativo é simplesmente a matriz identidade, e é utilizada uma restrição que força erro *a posteriori* igual a zero.

Esse problema é parecido ao formulado para o caso do algoritmo quasi-Newton, com a diferença da matriz que pondera a norma de distúrbio mínimo e da restrição imposta a essa matriz para garantir que seja positiva definida sempre. Desse modo, o problema a ser minimizado para o algoritmo 3 é apenas:

$$\begin{aligned} \min_{\mathbf{w}(n)} \quad & \|\mathbf{w}(n) - \mathbf{w}(n-1)\|^2 \\ \text{s.t.} \quad & d(n) - \mathbf{w}^T(n)\mathbf{u}(n) = 0 \end{aligned} \quad (3.20)$$

Para encontrar a solução do problema de otimização acima, utilizam-se os multiplicadores de Lagrange, obtendo-se a seguinte função-objetivo modificada:

$$\xi(n) = \|\mathbf{w}(n) - \mathbf{w}(n-1)\|^2 + \nu [d(n) - \mathbf{w}^T(n)\mathbf{u}(n)] \quad (3.21)$$

Derivando esta função em relação a $\mathbf{w}(n)$, tem-se

$$2[\mathbf{w}(n) - \mathbf{w}(n-1)] + \nu \mathbf{u}(n) = 0 \quad (3.22)$$

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \frac{\nu \mathbf{u}(n)}{2} \quad (3.23)$$

Utilizando a Equação (3.23) na restrição para encontrar o valor de ν , obtém-se

$$d(n) - \left[\mathbf{w}(n-1) + \frac{\nu \mathbf{u}(n)}{2} \right] \mathbf{u}(n) = 0 \quad (3.24)$$

$$\nu = \frac{2e(n)}{\|\mathbf{u}(n)\|^2}. \quad (3.25)$$

Finalmente, a solução do problema de otimização em (3.20) é encontrada substituindo-se (3.25) em (3.23):

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \frac{e(n)\mathbf{u}(n)}{\|\mathbf{u}(n)\|^2}. \quad (3.26)$$

A utilização da matriz identidade ponderando o distúrbio mínimo dos coeficientes pode ser uma escolha tão boa quanto qualquer outra matriz definida positiva, considerando que não se tenha conhecimento sobre as estatísticas do sinal de entrada.

3.4 Algoritmo 4

Uma alternativa para o caso da função objetivo do algoritmo quasi-Newton é utilizar uma versão *leakage* desse algoritmo, a qual preserva as boas propriedades numéricas do algoritmo QN original, proposto em [4], enquanto tenta reduzir a influência da normalização imposta pela restrição que impõe erro *a posteriori* zero. Para tal, propõe-se o seguinte problema de otimização:

$$\begin{aligned} \min_{\mathbf{w}(n)} \quad & \|\mathbf{w}(n) - \mathbf{w}(n-1)\|_{\mathbf{R}_{LQN}(n-1)}^2 + [d(n) - \mathbf{w}^T(n)\mathbf{u}(n)]^2 \\ \text{s.t.} \quad & \mathbf{u}^T(n)\mathbf{R}_{LQN}^{-1}(n)\mathbf{u}(n) = 1/2, \end{aligned} \quad (3.27)$$

onde o subscrito *LQN* identifica o algoritmo 4, sendo a referência a denominação *Leakage Quasi-Newton*.

A solução deste problema em relação aos coeficientes resulta na mesma equação de atualização do vetor de coeficientes do filtro adaptativo e da matriz de autocorrelação dadas pelo algoritmo RLS, Equações (2.47) e (2.49), respectivamente, repetidas aqui por conveniência:

$$\mathbf{w}(n) = \mathbf{w}(n-1) + e(n)\mathbf{R}_{RLS}^{-1}(n)\mathbf{u}(n) \quad (3.28)$$

$$\mathbf{R}_{LQN}^{-1}(n) = \frac{1}{\mu(n)} \left[\mathbf{R}_{LQN}^{-1}(n-1) - \frac{\mathbf{R}_{LQN}^{-1}(n-1)\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{R}_{LQN}^{-1}(n-1)}{\mu(n) + \mathbf{u}^T(n)\mathbf{R}_{LQN}^{-1}(n-1)\mathbf{u}(n)} \right]. \quad (3.29)$$

A restrição imposta no problema não possui relação com o vetor de coeficientes, $\mathbf{w}(n)$, resultando no mesmo algoritmo, mas sim com o passo do algoritmo, $\mu(n)$, que é atualizado a cada iteração de modo a respeitar a restrição. A aplicação da equação de atualização (3.29) na restrição $\mathbf{u}^T(n)\mathbf{R}_{LQN}^{-1}(n)\mathbf{u}(n) = 1/2$ resulta em

$$\mathbf{u}^T(n) \left\{ \frac{1}{\mu(n)} \left[\mathbf{R}_{LQN}^{-1}(n-1) - \frac{\mathbf{R}_{LQN}^{-1}(n-1)\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{R}_{LQN}^{-1}(n-1)}{\mu(n) + \mathbf{u}^T(n)\mathbf{R}_{LQN}^{-1}(n-1)\mathbf{u}(n)} \right] \right\} \mathbf{u}(n) = \frac{1}{2} \quad (3.30)$$

ou, equivalentemente,

$$\frac{1}{\mu(n)} \left[\frac{\tau(n)(\mu(n) + \tau(n)) - \tau^2(n)}{\mu(n) + \tau(n)} \right] = \frac{1}{2} \quad (3.31)$$

onde $\tau(n)$ é calculado da mesma forma que no algoritmo QN,

$$\tau(n) = \mathbf{u}^T(n)\mathbf{R}_{LQN}^{-1}(n-1)\mathbf{u}(n). \quad (3.32)$$

Resolvendo a Equação (3.31) em relação a $\mu(n)$ para encontrar a equação de atualização do passo do algoritmo, obtém-se

$$\mu(n) = \tau(n) \quad (3.33)$$

3.5 Algoritmo 5

Na busca de um algoritmo que aproveite alguns dos pontos positivos dos algoritmos quasi-Newton e RLS, desenvolve-se o algoritmo 5. O algoritmo quasi-Newton possui uma matriz de autocorrelação do sinal de entrada que uma vez iniciada positiva semi-definida permanecerá neste estado. Entretanto, possui uma função-objetivo complexa, o que torna sua interpretação física difícil e aumenta a complexidade computacional envolvida. O algoritmo RLS, em contrapartida, apresenta uma função-objetivo simples e com clara interpretação física, mas sua matriz de autocorrelação pode deixar de ser semi-definida positiva, levando o algoritmo a divergir.

O novo algoritmo surge, então, como uma parceria entre esses dois algoritmos. Formulou-se uma função-objetivo que tivesse a matriz de autocorrelação do sinal de entrada parecida com a do algoritmo quasi-Newton apresentado no capítulo anterior, ou seja, uma matriz Hessiana que, uma vez iniciada semi-positiva definida, permanecesse nesse estado independentemente do sinal de entrada. Por outro lado, buscou-se na nova função-objetivo uma forma mais simples de interpretar seu comportamento. Assim sendo, utilizando como função-objetivo uma versão do erro médio quadrático ponderado por um coeficiente variável, obteve-se

$$\xi(n) = \sum_{i=1}^n [1 - \mu(i)] [d(i) - \mathbf{u}^T(i)\mathbf{w}(n)]^2. \quad (3.34)$$

Pode-se observar a correspondência da equação acima com a função-objetivo do algoritmo RLS; porém, com essa escolha obtém-se uma matriz de autocorrelação que compartilha a característica do algoritmo quasi-Newton. Esse algoritmo, assim como os algoritmos RLS e quasi-Newton, requer o uso de um valor inicial para a matriz de autocorrelação do sinal de entrada, $\mathbf{R}(0)$. O efeito dessa iniciação será tratado brevemente no próximo capítulo.

Reestruturando a função custo deste novo algoritmo de modo a obter sua descrição por distúrbio mínimo, obtém-se o seguinte problema de minimização:

$$\min_{\mathbf{w}(n)} \|\mathbf{w}(n) - \mathbf{w}(n-1)\|_{\mathbf{R}_{A5}(n-1)}^2 + [1 - \mu(n)] [d(n) - \mathbf{u}^T(n)\mathbf{w}(n)]^2, \quad (3.35)$$

onde o subscrito $A5$ se refere ao algoritmo 5

Para resolver o problema da Equação (3.35), calcula-se o gradiente em relação aos coeficientes e iguala-se o resultado a zero, o que resulta em:

$$2\mathbf{R}_{A5}(n-1) [\mathbf{w}(n) - \mathbf{w}(n-1)] + 2[1 - \mu(n)] [-d(n) + \mathbf{w}^T(n)\mathbf{u}(n)] \mathbf{u}(n) = 0 \quad (3.36)$$

$$\begin{aligned} \{\mathbf{R}_{A5}(n-1) + [1 - \mu(n)] \mathbf{u}(n)\mathbf{u}^T(n)\} \mathbf{w}(n) &= \mathbf{R}_{A5}(n-1)\mathbf{w}(n-1) \\ &+ [1 - \mu(n)] d(n)\mathbf{u}(n) \end{aligned} \quad (3.37)$$

$$\mathbf{R}_{A5}(n)\mathbf{w}(n) = \mathbf{R}_{A5}(n)\mathbf{w}(n-1) + [1 - \mu(n)] [d(n) - \mathbf{w}^T(n)\mathbf{u}(n)] \mathbf{u}(n) \quad (3.38)$$

$$\mathbf{w}(n) = \mathbf{w}(n-1) + [1 - \mu(n)] e(n)\mathbf{R}_{A5}^{-1}(n)\mathbf{u}(n), \quad (3.39)$$

onde $e(n)$ é o erro *a priori* e

$$\mathbf{R}_{A5}(n) = \mathbf{R}_{A5}(n-1) + [1 - \mu(n)] \mathbf{u}(n)\mathbf{u}^T(n). \quad (3.40)$$

Como se pode observar pela Equação (3.40), a matriz de autocorrelação do sinal de entrada deste algoritmo apresenta similaridade com a do algoritmo quasi-Newton apresentado anteriormente. Como o escalar $\mu(n)$ foi obtido no algoritmo QN a partir da mesma matriz de autocorrelação, utilizada no presente algoritmo, lá se obtiveram ótimos resultados, utiliza-se aqui a mesma atualização do passo $\mu(n)$, ou seja,

$$\mu(n) = \frac{1}{2\mathbf{u}^T(n)\mathbf{R}^{-1}(n-1)\mathbf{u}(n)}. \quad (3.41)$$

A atualização recursiva da inversa da matriz de autocorrelação do sinal de entrada pode ser encontrada aplicando-se o lema da inversão de matrizes à Equação (3.40), obtendo-se

$$\mathbf{R}^{-1}(n) = \mathbf{R}^{-1}(n-1) + \frac{[1 - \mu(n)] \mathbf{R}^{-1}(n-1)\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{R}^{-1}(n-1)}{[1 + \mu(n)] 2\mathbf{u}^T(n)\mathbf{R}^{-1}(n-1)\mathbf{u}(n)}. \quad (3.42)$$

É importante apresentar o comportamento do passo do algoritmo, $\mu(n)$, para compreendermos sua influência na atualização da matriz de autocorrelação. Esse coeficiente depende diretamente do tipo de entrada do sistema adaptativo, como

pode ser visto pelas Figuras 3.1 e 3.2.

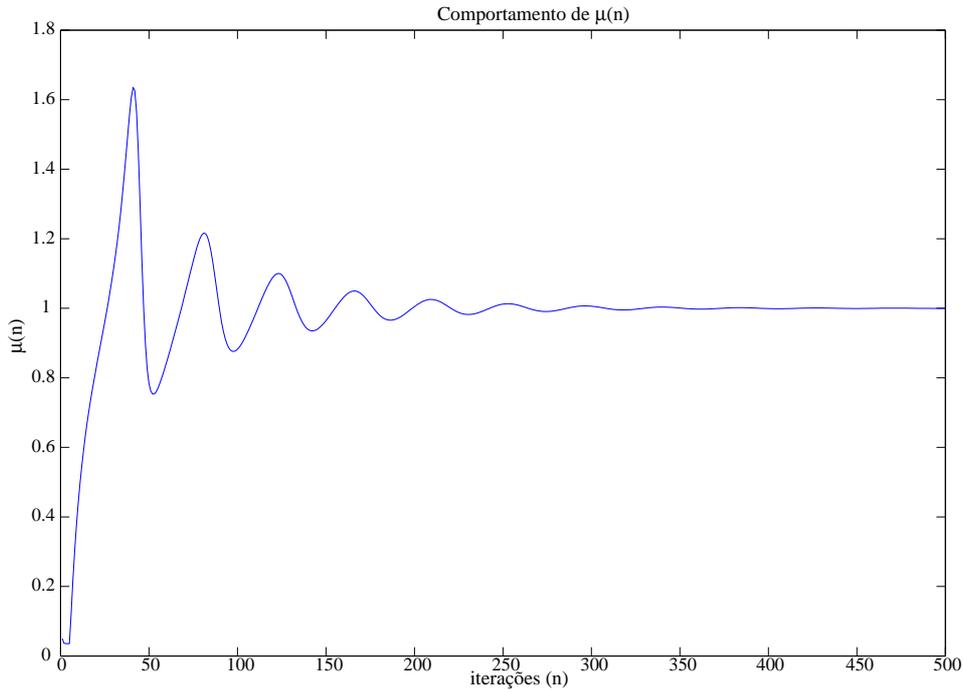


Figura 3.1: Comportamento de $\mu(n)$ para sinal de entrada senoidal

Para sinais de entrada senoidais observa-se que $\mu(n)$ tende para 1, o que faz o sistema parar de atualizar a matriz de autocorrelação e também o vetor de pesos. Esse comportamento é importante para o caso de entrada senoidal, pois evita a divergência do algoritmo.

No caso de a entrada ser ruído branco gaussiano, WGN (*White Gaussian Noise*), o passo $\mu(n)$ tem um comportamento aleatório e, se não for restringido, leva o algoritmo a divergir. Existem alguns métodos para evitar esse problema, dentre os quais permitir que o algoritmo atualize os vetores de pesos apenas quando $\mu(n)$ for menor que 1 foi o método que apresentou melhor resultado, sendo então o utilizado nas simulações.

Desse modo, o algoritmo proposto realiza uma verificação no valor de $\mu(n)$ após sua atualização. Se o valor estiver dentro do limite, ou seja, for menor do que 1, a atualização dos coeficientes é realizada; em caso contrário, apenas a matriz de autocorrelação do sinal de entrada é atualizada. Esse procedimento visa a melhorar o condicionamento da matriz de autocorrelação, ao mesmo tempo que garante que

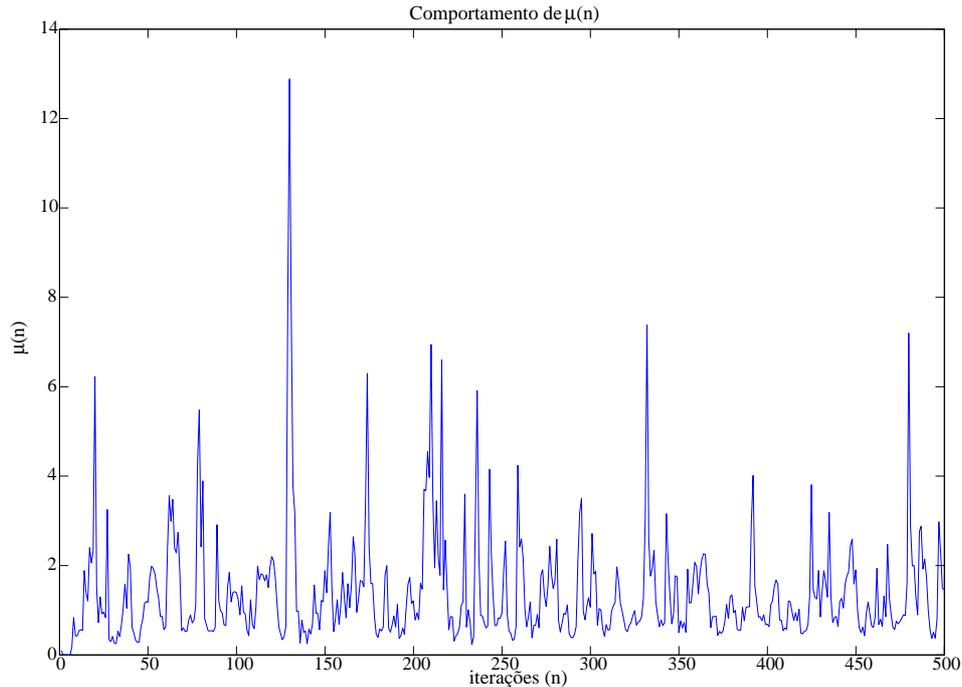


Figura 3.2: Comportamento de $\mu(n)$ para sinal de entrada ruído branco

o algoritmo mantenha o erro quadrático médio dos coeficientes decrescente.

Com o uso desta restrição, o algoritmo converge independentemente do tipo de entrada do sistema. Apesar de incluir esse elemento, nenhum comprometimento na velocidade de convergência do algoritmo foi percebida. Observa-se que o uso de valores altos na iniciação da inversa da matriz de autocorrelação aceleram a convergência do algoritmo, fato que será demonstrado no próximo capítulo.

3.6 Algoritmo 6

A garantia de a matriz de autocorrelação do sinal de entrada ser sempre definida positiva pode ser imposta de forma explícita nas funções-objetivo dos algoritmos através de restrições. Podemos utilizar o algoritmo 1, por exemplo, que possui uma função-objetivo bastante simples, e fazer com que o passo do algoritmo seja atualizado a cada iteração de forma a garantir que a inversa da matriz de autocorrelação seja sempre definida positiva. Além disso, pode-se impor uma restrição para garantir que a matriz não aumente indenidamente, evitando que um de seus

autovalores tenda a zero e desse modo, garantindo que sempre existirá uma inversa da matriz de autocorrelação. A função-objetivo com as restrições pode ser descrita como

$$\begin{aligned} \min_{w(n)} \quad & \|\mathbf{w}(n) - \mathbf{w}(n-1)\|_{\mathbf{R}_{A6}(n-1)}^2 + \mu(n) [d(n) - \mathbf{u}^T(n)\mathbf{w}(n)]^2 \quad (3.43) \\ \text{s.t.} \quad & \begin{cases} \mathbf{u}^T(n)\mathbf{R}_{A6}^{-1}(n)\mathbf{u}(n) \geq \alpha \\ \mathbf{u}^T(n)\mathbf{R}_{A6}^{-1}(n)\mathbf{u}(n) \leq \beta \\ \mu(n) \geq 0 \\ \alpha \leq \beta. \end{cases} \end{aligned}$$

O subscrito $A6$ se refere ao algoritmo 6.

Restringir o passo do algoritmo a ser maior ou igual a zero garante que tanto a matriz de autocorrelação quanto o vetor de coeficientes do filtro adaptativo sejam atualizados na mesma direção. As deduções seguem as mesmas apresentadas no algoritmo 1, resultando nas seguintes equações de atualização recursiva dos coeficientes do filtro adaptativo e da inversa da matriz de autocorrelação, respectivamente:

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu(n)e(n)\mathbf{R}_{A6}^{-1}(n)\mathbf{u}(n) \quad (3.44)$$

$$\mathbf{R}_{A6}^{-1}(n) = \mathbf{R}_{A6}^{-1}(n-1) - \frac{\mathbf{R}_{A6}^{-1}(n-1)\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{R}_{A6}^{-1}(n-1)}{\frac{1}{\mu(n)} + \mathbf{u}^T(n)\mathbf{R}_{A6}^{-1}(n-1)\mathbf{u}(n)}. \quad (3.45)$$

Para completar esse algoritmo, falta encontrar a equação que fornecerá o passo ótimo a cada iteração, capaz de garantir que a matriz seja definida positiva e esteja dentro dos limites $\alpha \leq \mathbf{u}^T(n)\mathbf{R}_{A6}^{-1}(n)\mathbf{u}(n) \leq \beta$. Para isso, pode-se resolver em separado o problema imposto pela restrição para simplificar a obtenção do passo do algoritmo, através da formulação

$$\begin{aligned} \min_{\mu(n)} \quad & \mathbf{u}^T(n)\mathbf{R}_{A6}^{-1}(n)\mathbf{u}(n) \geq \alpha \\ & \mathbf{u}^T(n)\mathbf{R}_{A6}^{-1}(n)\mathbf{u}(n) \leq \beta \\ & \mu(n) \geq 0. \end{aligned}$$

Utilizando o método da barreira logarítmica, pode-se escrever:

$$\min_{\mu(n)} -\frac{1}{t_1} \log [\mathbf{u}^T(n)\mathbf{R}_{A6}^{-1}(n)\mathbf{u}(n) - \alpha] - \frac{1}{t_2} \log [-\mathbf{u}^T(n)\mathbf{R}_{A6}^{-1}(n)\mathbf{u}(n) + \beta]. \quad (3.46)$$

Resolvendo o problema acima, e fazendo algumas manipulações matemáticas, chega-se à função para atualização de $\mu(n)$:

$$\mu(n) = \frac{\tau(n)(t_1 t_2) - (\beta t_2 + \alpha t_1)}{\tau(n)(\beta t_2 + \alpha t_1)}, \quad (3.47)$$

onde α e β são constantes escolhidas pelo usuário para determinar os limites para a restrição, t_1 e t_2 são constantes que controlam a precisão da barreira e

$$\tau(n) = \mathbf{u}^T(n)\mathbf{R}_{A6}^{-1}(n-1)\mathbf{u}(n). \quad (3.48)$$

A restrição $\mu(n) \geq 0$ é imposta ao algoritmo na forma de um teste, assim como feito no algoritmo 5.

3.7 Contribuições do Capítulo

Este capítulo apresentou diversos algoritmos novos, cada um com características próprias, especialmente quanto à forma de aproxima a matriz de auto-correlação do sinal de entrada. Por exemplo, a função objetivo apresentada no algoritmo 1 é uma alternativa para o desenvolvimento do algoritmo BEACON [9], no qual o algoritmo de mínimos quadrados utiliza restrições para reduzir a necessidade de atualização do algoritmo.

Dessa forma, o Capítulo 3 demonstra de maneira prática o potencial da descrição por distúrbio mínimo para o desenvolvimento de diversos algoritmos que compartilham características em sua função objetivo. A formulação uso das funções-objetivo como problemas de otimização permite a inclusão de restrições para melhorar o desempenho de um algoritmo ou até mesmo fazer um algoritmo compartilhar características de outro. O algoritmo 4, por exemplo, é uma alternativa para relaxar a restrição de erro *a posteriori* igual a zero imposta no algoritmo quasi-Newton, que resulta em um desajuste de 3dB fixo, reduzindo assim esse desajuste [7].

A otimização do passo do algoritmo é abordada nesse capítulo, sendo empregada na atualização do passo no algoritmo 6, dadas as restrições em relação à matriz de autocorrelação do sinal de entrada. Apresenta-se, também, uma maneira de utilizar o método de barreira para ajudar a resolver o problema de otimização com restrições de desigualdades. Apesar de a otimização ter sido simplificada para o passo do algoritmo satisfazer apenas as restrições, esse algoritmo demonstra como as restrições que não levam em consideração o vetor de coeficientes podem ser utilizadas para adicionar outras características a algoritmos adaptativos. Ressalta-se, também, que as funções-objetivo dos algoritmos desenvolvidos através da descrição por distúrbio mínimo estão dissociadas do somatório, o que pode facilitar a otimização da função-objetivo, tanto em relação ao vetor de coeficientes, $\mathbf{w}(n)$, quanto para o passo do algoritmo, $\mu(n)$.

O Capítulo 4 complementarará o estudo realizado neste capítulo através de diversas simulações. Essas simulações visam a demonstrar a convergência dos algoritmos, extraindo as qualidades e defeitos de cada algoritmo frente a diversas situações de sinal de entrada, ruído de medida e variação nos coeficientes do sistema a ser identificado.

Capítulo 4

Simulações

Os algoritmos desenvolvidos no Capítulo 3 podem ser utilizados com qualquer aplicação de sistemas adaptativos. Para demonstrar a utilidade e convergência dos algoritmos, neste capítulo são apresentadas simulações dos algoritmos sob variações do sistema no qual estes são utilizados. O desempenho dos algoritmos é verificado sob diversas situações para avaliar sua capacidade de convergência. Os algoritmos serão referenciados neste capítulo com os mesmos nomes com que foram apresentados no capítulo anterior.

Aplicam-se os algoritmos na identificação de um sistema desconhecido, e verifica-se o comportamento do erro quadrático médio associado. Os algoritmos são expostos a entradas diferentes como ruído branco, ruído colorido, senóide e senóide com ruído. Verificam-se os desempenhos dos algoritmos sob circunstâncias de ruído adicional, não-correlacionado com o sinal de entrada, que resulte em razões sinal-ruído (SNR) de 20dB, 40dB e 80dB. A capacidade de acompanhar mudanças no sistema, como uma súbita modificação dos coeficientes ótimos após a convergência dos algoritmos também é considerada; isto é feito apenas para o caso de razão sinal-ruído de 80dB.

Em cada experimento, o sistema a ser identificado possui 10 coeficientes gerados aleatoriamente. O sinal de entrada é considerado com potência unitária e filtro do sistema a ser identificado é normalizado para oferecer ganho unitário. Realiza-se a média de 200 simulações consecutivas com o mesmo sistema para se obter o gráfico de convergência do MSE. Cada simulação é realizada até a convergência do algoritmo, quando está for atingível.

A Tabela 4.1 apresenta os valores e equações para o passo dos algoritmos, $\mu(n)$, que foram utilizados nas simulações dos Experimentos I a VI. Nos algoritmos com passo fixo, os valores de $\mu(n)$ foram escolhidos para que os algoritmos obtivessem o mesmo desajuste.

Tabela 4.1: Passo dos Algoritmos

Algoritmo	Passo do Algoritmo
Algoritmo 1	$\mu(n) = 100$
Algoritmo 2	$\mu(n) = 0,01$
Algoritmo 3	$\mu(n) = 0,3$
Algoritmo 4	$\mu(n) = \tau(n)$
Algoritmo 5	$\mu(n) = \frac{1}{2\mathbf{u}^T(n)\mathbf{R}^{-1}(n-1)\mathbf{u}(n)}$
Algoritmo 6	$\mu(n) = \frac{\tau(n)(t_1 t_2) - (\beta t_2 + \alpha t_1)}{\tau(n)(\beta t_2 + \alpha t_1)}$

Verifica-se o comportamento dos algoritmos frente aos diferentes casos analisados, de modo a comprovar sua convergência e obter suas qualidades e defeitos. Após os experimentos iniciais de convergência, analisa-se o desempenho dos algoritmos que apresentaram melhores resultados em comparação aos algoritmos RLS e QN.

4.1 Iniciação da Matriz de Autocorrelação

Antes de começar os experimentos, faz-se necessária uma análise inicial do comportamento do passo, $\mu(n)$, do algoritmo 5 para se entender o motivo de utilizar um valor inicial diferente da identidade, na inversa da matriz de autocorrelação, $\mathbf{R}^{-1}(0)$. A iniciação da matriz de autocorrelação com a matriz identidade é geralmente utilizada durante o processo de iniciação do filtro adaptativo, pois considera-se que não se tem informações sobre o sinal de entrada. Para demonstrar o efeito que um valor diferente na iniciação da matriz de autocorrelação tem sobre o comportamento do algoritmo 5 na Figura 4.1 compara-se desempenho do passo do algoritmo para dois valores diferentes de valor inicial de $\mathbf{R}^{-1}(0)$ sendo utilizado a matriz identidade multiplicada por um fator δ , $\delta = 1$ e $\delta = 1000$, para um sinal de entrada senoidal. O algoritmo é utilizado na identificação de um filtro FIR com 10 coeficientes, com uma razão sinal-ruído (SNR) de 80dB.

Como pode-se observar, um valor menor na iniciação de $\mathbf{R}^{-1}(0)$ pode fazer com que o passo tenha valor maior que 1 durante o início do processo de adaptação

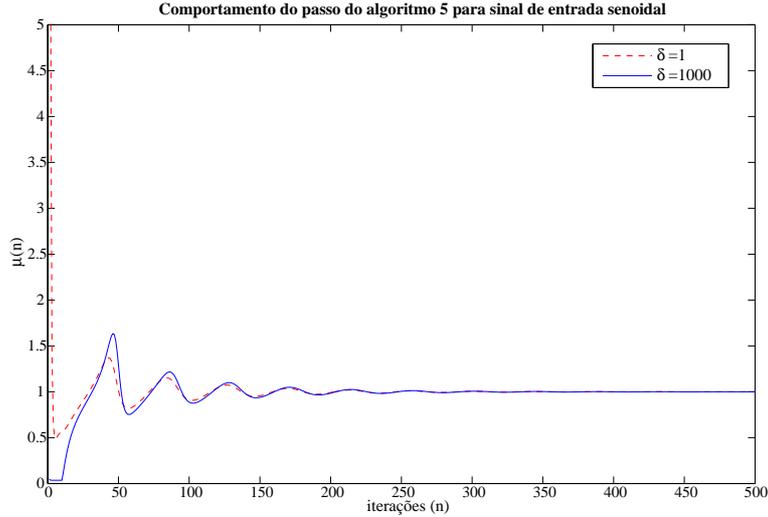


Figura 4.1: Comportamento do passo do algoritmo 5 para entrada senoidal.

do algoritmo, levando os coeficientes do filtro adaptativo a serem atualizados no sentido reverso. Obviamente, a restrição *hard* imposta na atualização dos coeficientes do filtro adaptativo impede que os mesmo recebam uma adaptação incorreta, mas também impede que sejam adaptados enquanto o passo não reduzir seu valor abaixo de 1. Além disso, para entradas senoidais, o passo do algoritmo tende para 1, condição na qual a atualização da inversa da matriz de autocorrelação é anulada, bem como a atualização dos coeficientes. Esse fato pode levar a valores elevados de desajuste do MSE, caso os coeficientes do filtro ainda não tenham convergido para seus valores ótimos.

Por outro lado, quando se aumenta o valor inicial da inversa da matriz de autocorrelação do sinal de entrada, o passo do algoritmo pode assumir valores inferiores a 1 desde o início da adaptação do algoritmo e assim, atualizando os coeficientes do filtro adaptativo durante mais tempo, possibilita a convergência dos coeficientes. Ressalta-se que os coeficientes ótimos para entradas senoidais podem não necessariamente ser os coeficientes da planta a ser identificada, uma vez que o sinal de entrada não persistentemente excitante faz com que a matriz de autocorrelação não tenha posto completo, permitindo infinitas soluções.

Outro motivo para a convergência ser alcançada com o aumento da iniciação da inversa da matriz de autocorrelação é que esta acelera a velocidade de convergência do algoritmo, fazendo com que ele consiga convergir antes que seu passo

estabilize em 1. Na Figura 4.2 tem-se a comparação entre os erros quadráticos médios encontrados na simulação do algoritmo 5 para a iniciação de $R^{-1}(0)$ com fatores $\delta = 1$ e $\delta = 1000$, com sinal de entrada senoidal com potência unitária. Neste experimento, realizou-se a média de 200 simulações para se obter uma curva de convergência mais clara. Como pode ser observado, com um valor maior da iniciação da inversa da matriz de autocorrelação o algoritmo consegue convergir para o mínimo MSE possível, 10^{-8} neste caso.

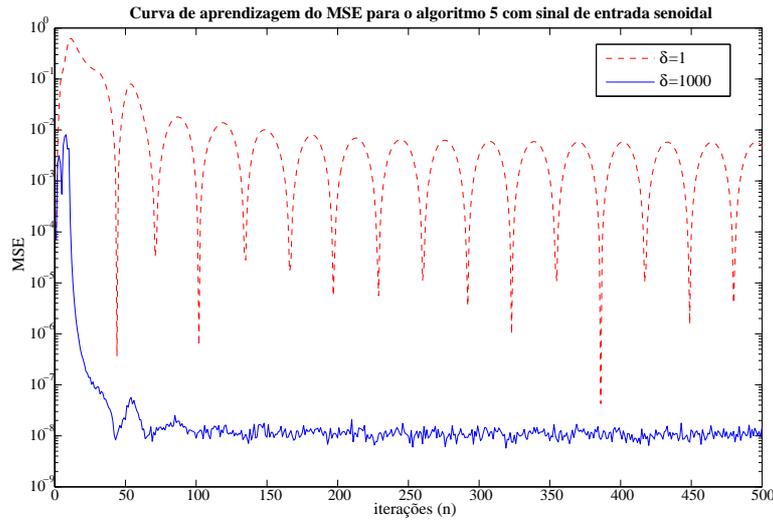


Figura 4.2: Comportamento do MSE do algoritmo 5 para entrada senoidal.

Como foi observado no Capítulo 3, pela formulação dos algoritmos através de problemas de otimização com restrição pelo método de barreira, os algoritmos que utilizam a inversa da matriz de autocorrelação na atualização dos coeficientes do filtro adaptativo dependem da iniciação desta matriz, que em última análise pode polarizar a resposta do erro quadrático médio associado. A redução na influência da iniciação da matriz de autocorrelação corresponde a um aumento no valor inicial da inversa desta matriz, o que, portanto é um procedimento justificável.

Em [2], o efeito da iniciação da matriz de autocorrelação para o algoritmo RLS é brevemente tratado. Destaca-se em [2] o fato de o algoritmo RLS convergir para a solução ótima em $N + 1$ iterações em um ambiente sem ruído, onde N é o número de coeficientes do filtro adaptativo, e que nesse ponto a iniciação da matriz pode ser negligenciada. Isso se deve ao fato de que após a iteração $N + 1$, o vetor do sinal de entrada já possui informação suficiente para permitir ao algoritmo adaptativo

identificar os coeficientes de um sistema desconhecido, ou seja, a cadeia de atrasos já está completa com informações do sinal de entrada.

Além de uma possível redução na polarização da resposta, o aumento no valor inicial da matriz de autocorrelação também faz os algoritmos convergirem mais rapidamente. Essa é uma característica apresentada por praticamente todos os algoritmos que utilizam a inversa da matriz de autocorrelação desenvolvidos nesta tese, com exceção do algoritmo 4 (*Leakage quasi-Newton*) que é imune a essas variações.

A análise do motivo para que δ elevado leve à melhora na velocidade de convergência dos algoritmos está além dos objetivos desta tese; porém, uma possibilidade é que a iniciação de $\mathbf{R}^{-1}(0)$ possa estar exercendo influência no passo do algoritmo adaptativo. Com um valor alto nessa iniciação, teríamos aumentado o valor do passo e conseqüentemente acelerado a convergência do algoritmo. Por outro lado, sabe-se que em muitos algoritmos o aumento do passo resulta em maior desajuste do MSE; porém, a matriz de autocorrelação realiza o desvanecimento de informações passadas, o mesmo valendo para o valor de sua iniciação, o que propiciaria uma aceleração na convergência do algoritmo sem aumento no desajuste do erro quadrático médio.

O valor de $\mu(n)$, para o caso de o sinal de entrada ser ruído branco gaussiano (WGN) nos algoritmos com passo variável, não converge para nenhum valor específico; porém, este tipo de entrada também apresenta convergência mais rápida com o aumento do valor inicial de $\mathbf{R}^{-1}(0)$. Na Figura 4.3 observa-se o comportamento do passo do algoritmo 5 para sinal de entrada WGN, e na Figura 4.4 comprova-se que o aumento do valor do fator δ acelera a convergência do algoritmo.

Deve-se ressaltar que, para este tipo de sinal de entrada, o algoritmo 5 consegue convergir para o mínimo MSE, mesmo que a iniciação utilize um fator multiplicativo de $\delta = 1$, porém, a velocidade é consideravelmente mais lenta.

O aumento da velocidade de convergência do algoritmo em função do aumento no valor de $\mathbf{R}^{-1}(0)$ não é uma particularidade somente dos algoritmos desenvolvidos no Capítulo 3. O mesmo comportamento pode ser observado nos algoritmos RLS e Quasi-Newton, como mostra a Figura 4.5.

Desse modo, como o algoritmo 5 necessita de uma iniciação maior da inversa da matriz de autocorrelação do sinal de entrada para convergir em entradas senoidais

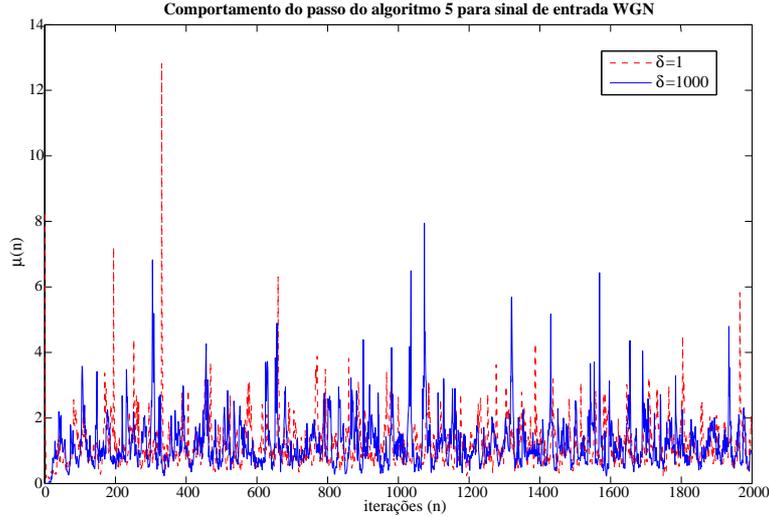


Figura 4.3: Comportamento do passo do algoritmo 5 para entrada WGN.

e pelo fato de que essa modificação não altera o desajuste dos algoritmos, o fator $\delta = 1000$ é utilizado nos experimentos realizados nesta tese.

4.2 Influência do Passo no Algoritmo 1

O algoritmo 1 apresenta uma característica interessante quanto ao passo utilizado: o valor do passo não interfere no valor do desajuste do erro médio quadrático. Pode-se observar na Figura 4.6, que corresponde a simulações com entrada WGN, o efeito do passo do algoritmo na velocidade de convergência, sem, entretanto, alterar o valor mínimo do MSE quando o algoritmo alcança o regime permanente. Para esta simulação, utilizou-se o fator multiplicativo da iniciação da inversa da matriz de autocorrelação $\delta = 1$ para tornar o efeito do passo mais evidente na velocidade de convergência. Pode-se observar que o mesmo valor de erro médio quadrático é alcançado, independentemente do passo do algoritmo.

Uma análise aprofundada sobre o comportamento dos algoritmos não está no escopo desta tese; o fato de o passo do algoritmo 1 não alterar o valor do desajuste do erro médio quadrático necessita de uma investigação posterior. Entretanto, uma provável razão seria o fato de o passo do algoritmo ser equivalente a uma constante multiplicando a matriz de autocorrelação do sinal de entrada e também o vetor de correlação cruzada do sinal de entrada com o sinal de referência, resultando na

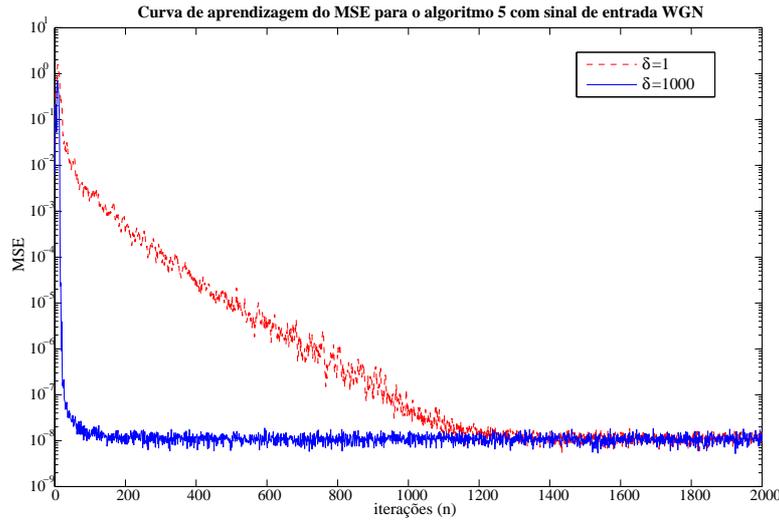


Figura 4.4: Comportamento do MSE do algoritmo 5 para entrada WGN.

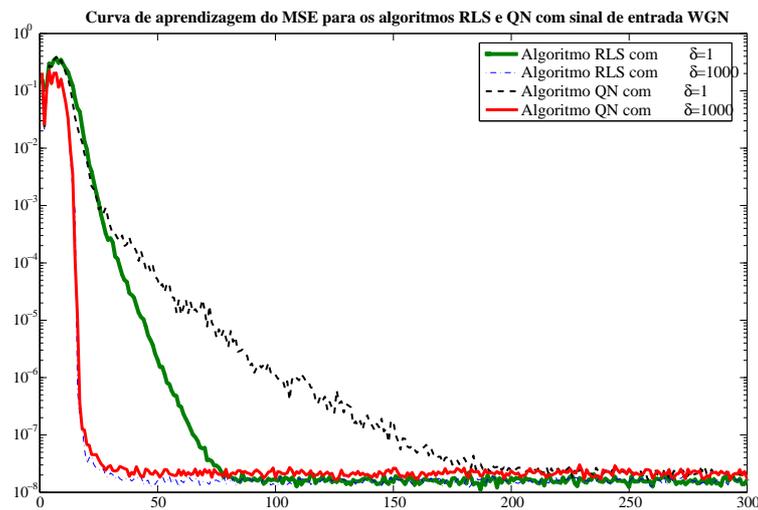


Figura 4.5: Comportamento do MSE do algoritmo RLS para entrada WGN.

solução de Wiener e anulando sua influência no desajuste em relação aos coeficientes do filtro.

Diferentemente, quando se utiliza o passo variável para este algoritmo, como se faz no algoritmo 6, o desajuste é controlado pelas restrições impostas, possivelmente por não podemos mais expressar o passo como uma constante multiplicando a matriz de autocorrelação e o vetor de correlação cruzada. Esse efeito está presente no algoritmo BEACON onde, normalmente, se utilizam restrições conhecidas como *Set-Membership* [9] para reduzir a quantidade de atualizações em relação o nível do

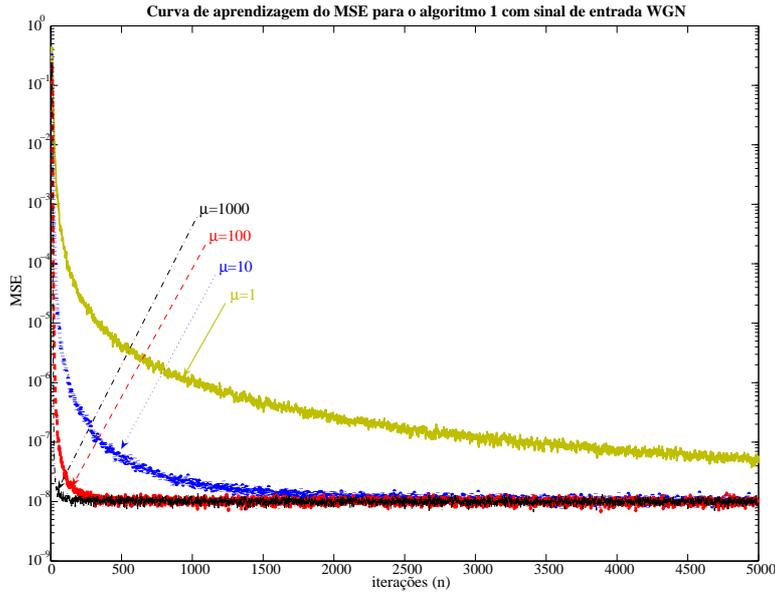


Figura 4.6: Comportamento do MSE do algoritmo 1 para diferentes valores do passo do algoritmo.

erro durante a adaptação.

4.3 Experimento I: Ruído Branco como Sinal de Entrada

A partir desta seção, serão realizados experimentos para comprovar a convergência dos algoritmos desenvolvidos nesta tese. Primeiramente, será utilizado como sinal de entrada um ruído branco, o qual, por simplicidade será gerado com distribuição gaussiana. O experimento foi dividido em três seções, em que serão utilizadas razões sinal-ruído diferentes.

Tanto para este experimento quanto para os demais, os gráficos da convergência do erro médio quadrático serão separados entre algoritmos com passo fixo e com passo variável, pois o objetivo desses experimentos é apenas comprovar a convergência dos algoritmos frente a diferentes sistemas; desse modo, podem-se tecer mais facilmente considerações relevantes sobre os algoritmos e suas famílias.

4.3.1 Experimento Ia: SNR = 80dB

Primeiramente, utilizaremos uma razão sinal-ruído de 80dB, caso extremamente otimista de um sistema praticamente sem interferência de ruídos. A Figura 4.7 apresenta a curva de aprendizagem dos algoritmos 1, 2 e 3 para este experimento.

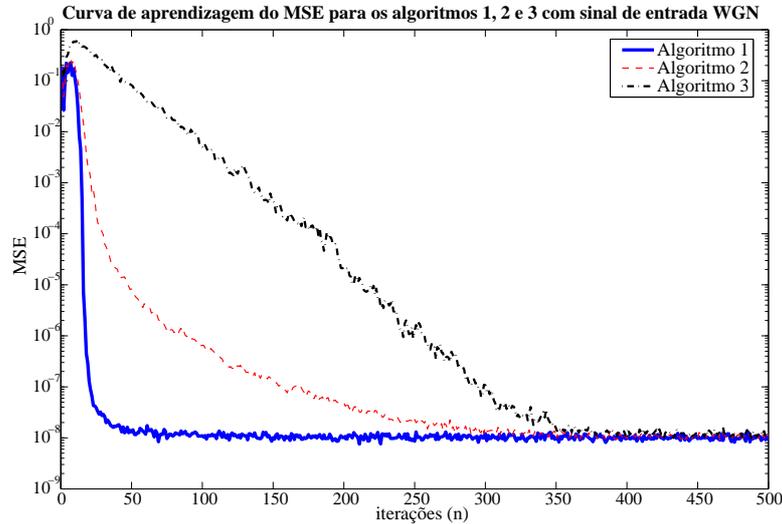


Figura 4.7: MSE dos algoritmos 1–3 para sinal de entrada ruído branco e SNR=80dB.

Como pode ser observado, os algoritmos convergem sem problemas para o valor mínimo permitido de MSE. Destaca-se que o algoritmo 3 é um algoritmo que não utiliza informações passadas do sinal de entrada, e por isso sua velocidade de convergência é inferior, ao mesmo tempo que é mais simples e computacionalmente menos complexo. A velocidade de convergência dos algoritmos é acelerada consideravelmente pela utilização de um valor elevado para iniciação da matriz inversa de autocorrelação do sinal de entrada.

Na Figura 4.8, tem-se a curva de aprendizagem do MSE para os algoritmos 4, 5 e 6, que possuem passo variável. O algoritmo 6 possui quatro constantes definidas pelo usuário que influenciam seu desempenho pois controlam como as restrições devem se comportar e, conseqüentemente, o comportamento do passo do algoritmo. Dependendo de como se configuram os valores dessas constantes, o algoritmo pode ter uma convergência mais rápida ou mais lenta. Os valores utilizados foram escolhidos para obter uma velocidade de convergência próxima à do algoritmo 5 e manter um desajuste comparável ao dos algoritmos 4 e 5; foram eles $\alpha = 0,001$, $\beta = 1$,

$t_1 = 10$ e $t_2 = 1$.

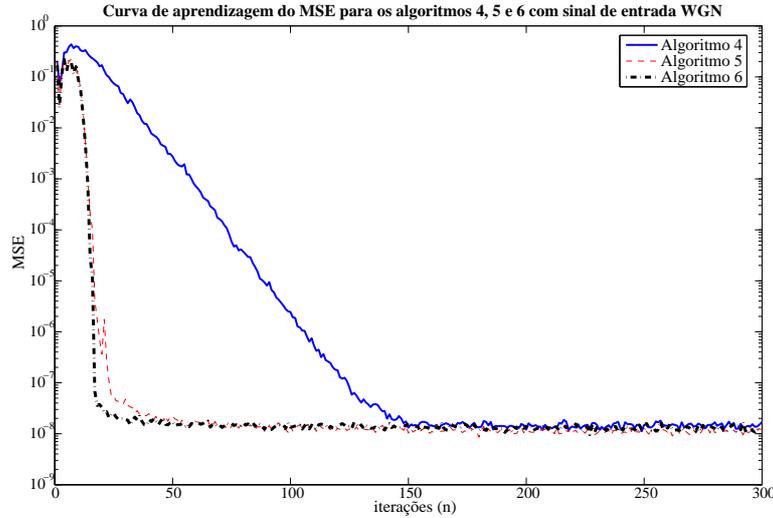


Figura 4.8: MSE dos algoritmos 4–6 para sinal de entrada ruído branco e SNR=80dB.

Observa-se pela Figura 4.8, que os algoritmos convergem para o mesmo valor de MSE, sendo que os algoritmos 5 e 6 possuem velocidades de convergência semelhantes, e superiores à do algoritmo 4.

4.3.2 Experimento Ib: SNR = 40dB

Para este experimento, o sistema é o mesmo, uma planta com 10 coeficientes a ser identificada com sinal de entrada WGN; porém, o ruído adicionado ao sistema agora é moderado, resultando numa razão sinal-ruído igual a 40dB. Apresenta-se o comportamento do erro quadrático médio nas Figuras 4.9 e 4.10, para os algoritmos 1–3 e 4–6, respectivamente.

Os algoritmos convergem para o mínimo MSE possível, levando em consideração os desajustes gerados por cada algoritmo em relação ao valor de seu passo, demonstrando que, apesar da degradação da resposta do MSE, os algoritmos tentam encontrar o valor ótimo dos coeficientes. O algoritmo 2 apresenta um comportamento interessante, considerando-se os parâmetros utilizados nesse experimento: $\mu(n) = 0,01$ e iniciação da matriz com multiplicador $\delta = 1000$. Esse algoritmo possui uma convergência rápida até $MSE \approx 10^{-4}$ onde forma uma espécie de “joelho”, reduzindo sua velocidade a partir desse ponto até o regime permanente. No algo-

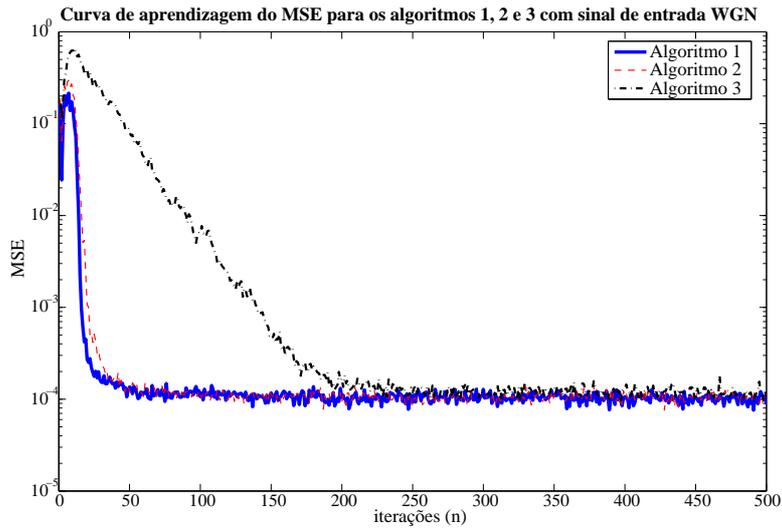


Figura 4.9: MSE dos algoritmos 1–3 para sinal de entrada ruído branco e SNR=40dB.

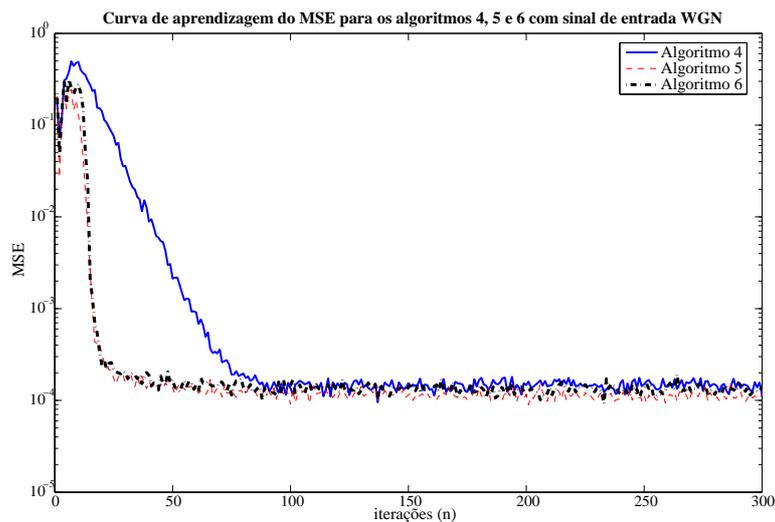


Figura 4.10: MSE dos algoritmos 4–6 para sinal de entrada ruído branco e SNR=40dB.

ritmo 2, assim como no algoritmo 5, a iniciação da matriz de autocorrelação possui forte influência na convergência, possivelmente devido ao fato de ter sido escolhido um valor baixo para o passo do algoritmo, o que aumenta a influência de informações passadas. Os demais algoritmos têm desempenho semelhante ao apresentado para SNR de 80dB.

4.3.3 Experimento Ic: SNR = 20dB

Para finalizar o experimento I, expõe-se o sistema a um ruído tal que a SNR seja igual a 20dB. As Figuras 4.11 e 4.12 mostram a curva de aprendizagem do MSE dos algoritmos. Como pode ser verificado, os algoritmos conseguem convergir sem problemas para o mínimo MSE possível. O desempenho dos algoritmos permanece praticamente o mesmo, considerando apenas o valor maior de MSE para o qual os algoritmos convergem, o que reduz o número de iterações necessárias para os algoritmos 3 e 4 atingirem o regime permanente.

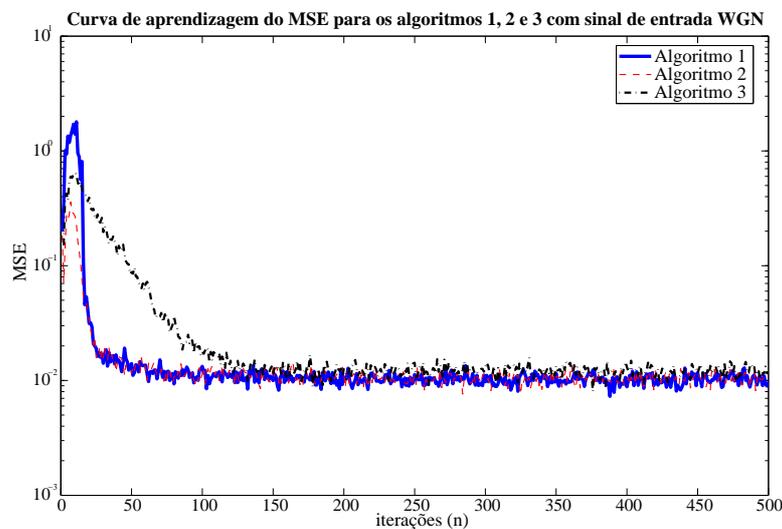


Figura 4.11: MSE dos algoritmos 1–3 para sinal de entrada ruído branco e SNR=20dB.

4.4 Experimento II: Ruído Colorido como Sinal de Entrada

Algoritmos baseados no algoritmo LMS, como por exemplo o próprio LMS convencional e o algoritmo 3 derivado nesta tese que é o algoritmo LMS normalizado, normalmente apresentam problemas de convergência quando o sinal de entrada é altamente correlacionado. Entretanto, normalmente esse fato não é observado em algoritmos recursivos baseados na função-objetivo *Least-Squares*, pois o uso da matriz de autocorrelação dá a estes algoritmos um certo grau de imunidade ao espalhamento

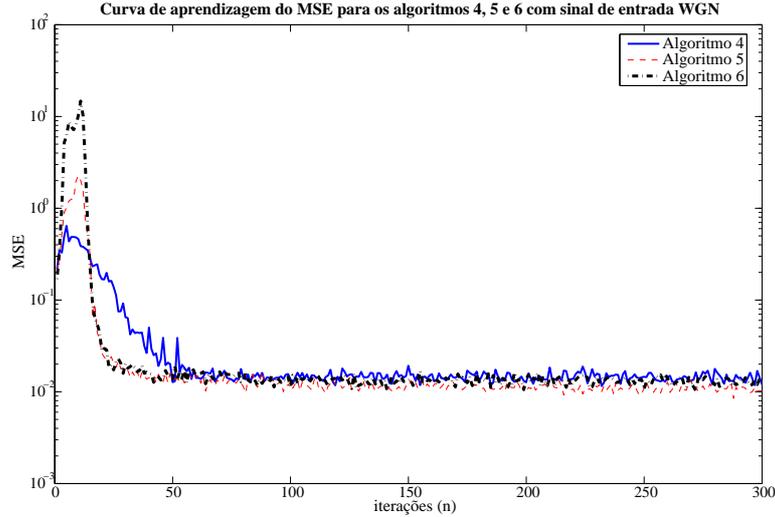


Figura 4.12: MSE dos algoritmos 4–6 para sinal de entrada ruído branco e SNR=20dB.

dos autovalores da matriz de autocorrelação do sinal de entrada.

Neste experimento, verifica-se o desempenho dos algoritmos obtidos no capítulo anterior quando o sinal de entrada é correlacionado. Para isso, utilizamos um sinal de entrada WGN que é tornado correlacionado através de um filtro passa-baixas com frequência de corte em $0,35w_s$, sendo w_s a frequência de amostragem do sinal de entrada. Com essa filtragem do sinal de entrada, obtém-se um espalhamento dos autovalores, $\lambda_{\text{máx}}/\lambda_{\text{mín}}$, de aproximadamente 450 para o sinal de entrada dos sistemas desconhecido e adaptativo. São utilizadas três razões sinal-ruído diferentes para avaliar o desempenho dos algoritmos.

4.4.1 Experimento IIa: SNR = 80dB

O comportamento dos algoritmos quando expostos a um sinal de entrada correlacionado, em um ambiente praticamente livre de interferência, com $SNR = 80dB$, pode ser observado pelas Figuras 4.13 e 4.14.

Na Figura 4.13, observa-se que os algoritmos que utilizam a matriz de autocorrelação para atualização dos coeficientes do filtro são praticamente imunes ao espalhamento do sinal de entrada. O algoritmo 3, por sua vez, tem seu desempenho comprometido pelo aumento do espalhamento do sinal de entrada, sendo que o algoritmo ainda estava convergindo após o limite de iterações utilizado nesta simulação.

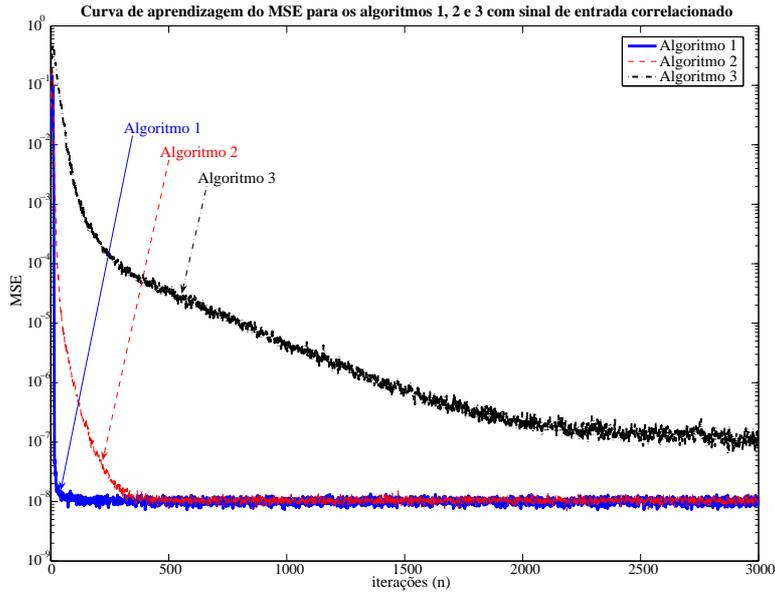


Figura 4.13: MSE dos algoritmos 1–3 para sinal de entrada ruído colorido e SNR=80dB.

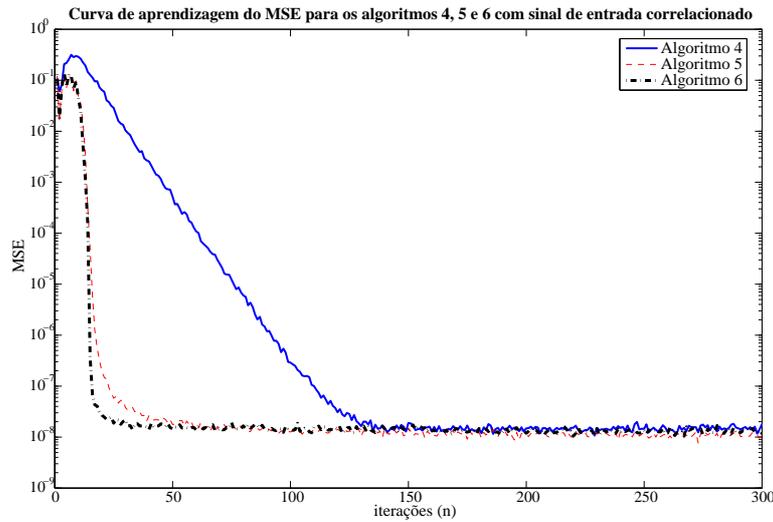


Figura 4.14: MSE dos algoritmos 4–6 para sinal de entrada ruído colorido e SNR=80dB.

Os algoritmos 4, 5 e 6 também têm suas características de convergência praticamente independentes do espalhamento do sinal de entrada, como se vê na Figura 4.14, sendo que o algoritmo 6 apresenta uma pequena redução em sua velocidade de convergência.

4.4.2 Experimento IIb: SNR = 40dB

Nesta seção, modifica-se o valor do SNR para 40dB e verifica-se a influência que o aumento no ruído adicional possui sobre o desempenho dos algoritmos. Nas Figuras 4.15 e 4.16 apresenta-se a curva de aprendizagem do MSE para os algoritmos.

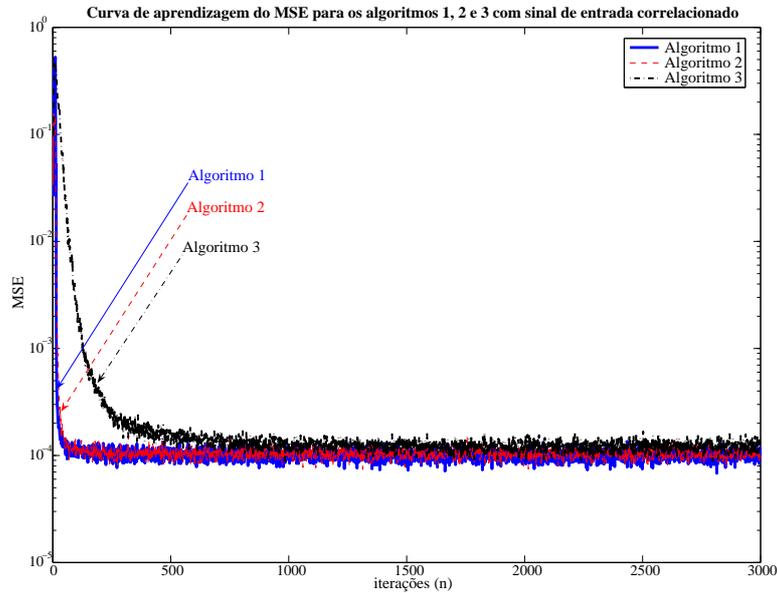


Figura 4.15: MSE dos algoritmos 1–3 para sinal de entrada ruído colorido e SNR=40dB.

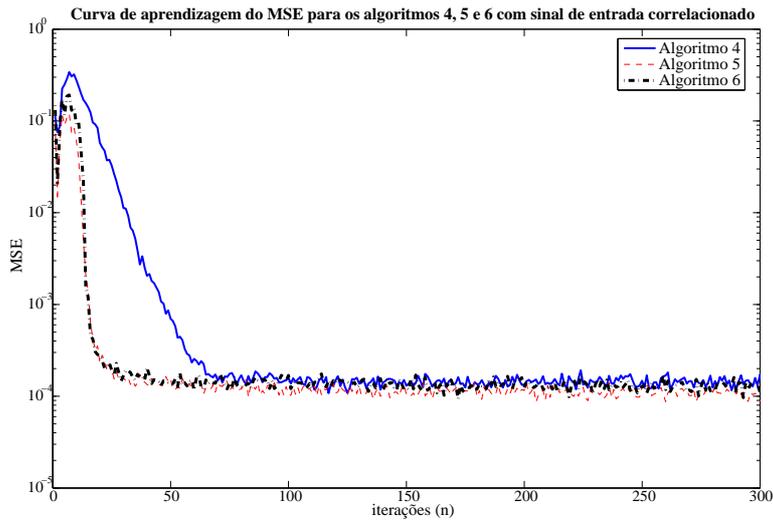


Figura 4.16: MSE dos algoritmos 4–6 para sinal de entrada ruído colorido e SNR=40dB.

Observando as Figuras 4.15 e 4.16, nota-se que a redução da razão sinal-ruído modifica o valor mínimo da convergência, como esperado, mas não altera a velocidade de convergência no geral. Um detalhe a ser ressaltado é que o uso da iniciação da inversa da matriz de autocorrelação com um multiplicador δ faz com que o “joelho” do gráfico de convergência dos algoritmos adaptativos seja deslocado para um valor em dB maior ou menor, dependendo da escolha de δ , como discutido no início deste capítulo. Por esse motivo, a convergência para valores menores de SNR permite que os algoritmos (inclusive o 3, que não usa a matriz de autocorrelação na atualização dos coeficientes mas tem um “joelho” na curva de convergência perto de 40dB) convirjam mais rapidamente.

4.4.3 Experimento IIc: SNR = 20dB

Finalizando o experimento, avalia-se o desempenho dos algoritmos quando expostos a um ruído de medida tal que a SNR seja de 20dB. Para esse nível de ruído, o “joelho” da curva de convergência dos algoritmos é praticamente o mesmo para todos, já que o fator multiplicativo utilizado acelera a convergência da maioria dos algoritmos até que $MSE \approx 10^{-4}$. Desse modo, praticamente todos os algoritmos apresentam velocidades de convergência semelhantes, com exceção do algoritmo 3, como pode ser observado nas Figuras 4.17 e 4.18. Para os algoritmos 4, 5 e 6, é difícil diferenciar o desempenho entre eles, sendo que o algoritmo 4 leva um pouco mais de tempo para convergir.

4.5 Experimento III: Sinal de Entrada Senoidal

Para avaliarmos o desempenho dos algoritmos com sinais de entrada não-persistentemente excitantes de ordem N , onde N é a ordem do filtro, neste experimento utilizam-se os algoritmos adaptativos para identificar um sistema desconhecido quando a entrada de ambos os filtros, desconhecido e adaptativo, é um sinal senoidal. A amostragem da onda senoidal é 20 vezes mais rápida que sua frequência. Os experimentos são divididos em três partes, de acordo com a razão sinal-ruído utilizada no experimento.

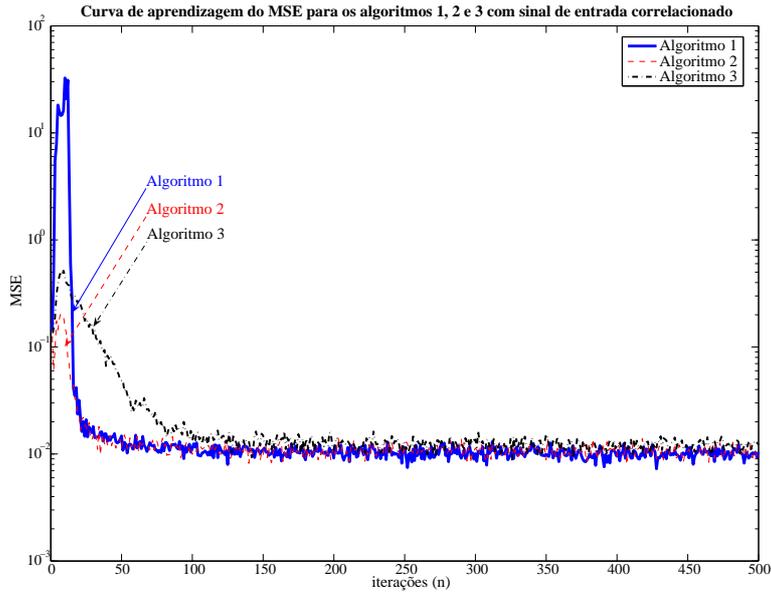


Figura 4.17: MSE dos algoritmos 1–3 para sinal de entrada ruído colorido e SNR=20dB.

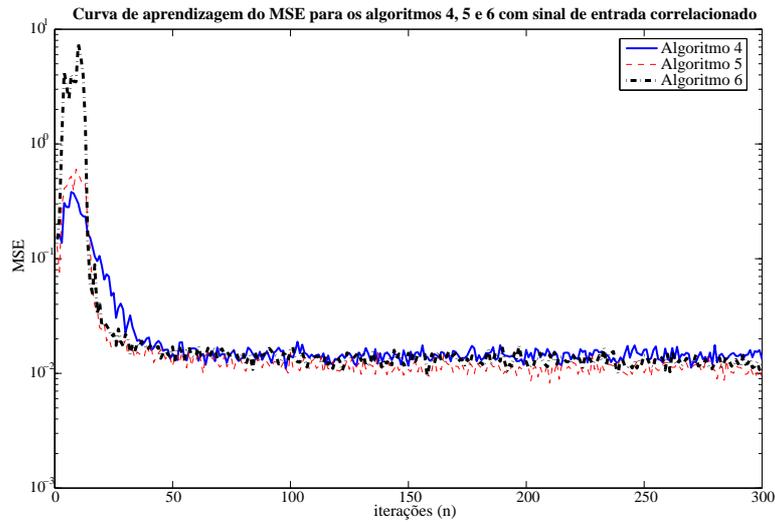


Figura 4.18: MSE dos algoritmos 4–6 para sinal de entrada ruído colorido e SNR=20dB.

4.5.1 Experimento IIIa: SNR = 80dB

Neste experimento utiliza-se uma razão sinal-ruído de 80dB, correspondente a um sistema praticamente livre de ruídos, apenas para que se imponha um limite ao erro mínimo de convergência em regime. A Figura 4.19 mostra a convergência

do erro quadrático médio para os algoritmos 1, 2 e 3.

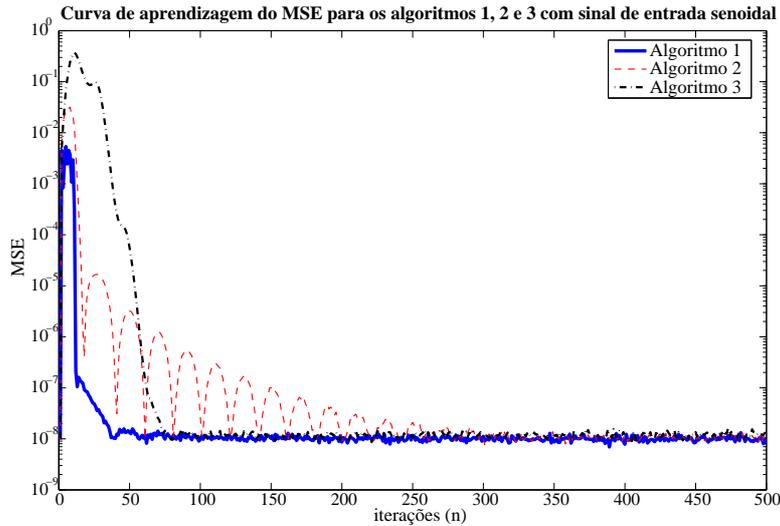


Figura 4.19: MSE dos algoritmos 1–3 para sinal de entrada senoidal e SNR=80dB.

Observa-se que o algoritmo 3 (NLMS) tem boa resposta para esse tipo de entrada, conseguindo convergir antes até que o algoritmo 2. Apesar da convergência aparente mostrada pela Figura 4.19, espera-se que o algoritmo 2 tenha comportamento parecido com o do algoritmo RLS, por suas matrizes de autocorrelação possuírem características semelhantes. Deste modo, é esperado que o algoritmo diverja quando exposto a sinais não-persistentemente excitantes, como é o caso do sinal de entrada senoidal. A Figura 4.20 apresenta um número maior de iterações da simulação para o algoritmo 2, mostrando a divergência do algoritmo. O algoritmo 1 consegue convergir para este tipo de sinal, constituindo-se numa boa opção, considerando sua rápida convergência, a independência do desajuste em relação ao passo do algoritmo e sua capacidade de convergir com sinais não-persistentemente excitantes.

A curva de convergência do erro quadrático médio para os algoritmos 4, 5 e 6 pode ser observada na Figura 4.21, onde nota-se que o algoritmo 4 mostra claros sinais de divergência para este tipo de entrada. Quanto aos outros dois algoritmos, observa-se que eles começam a convergir após o vetor do sinal de entrada estar completo de informações, ou seja, a linha de atraso estar completa com informação do sinal de entrada, sendo que o algoritmo 6 apresenta velocidade de convergência superior à do algoritmo 5. A convergência para esse tipo de sinal de entrada, não-

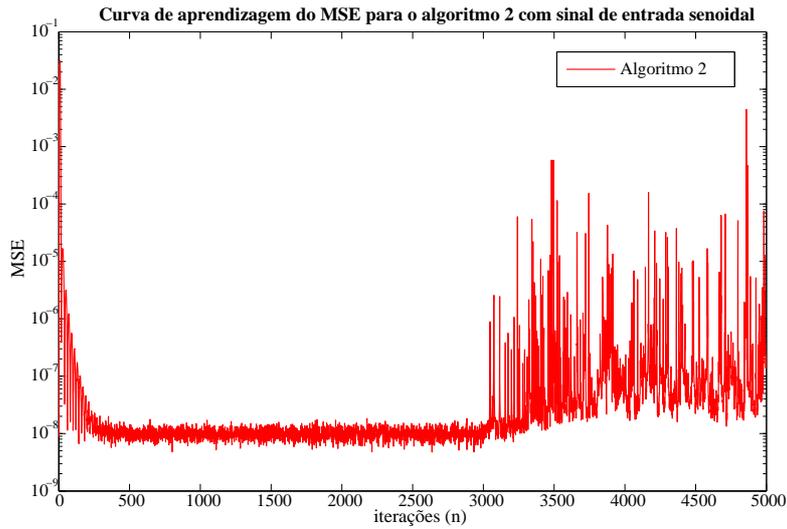


Figura 4.20: MSE dos algoritmo 2 para sinal de entrada senoidal e SNR=80dB.

persistentemente excitante, é também observada no algoritmo quasi-Newton [4].

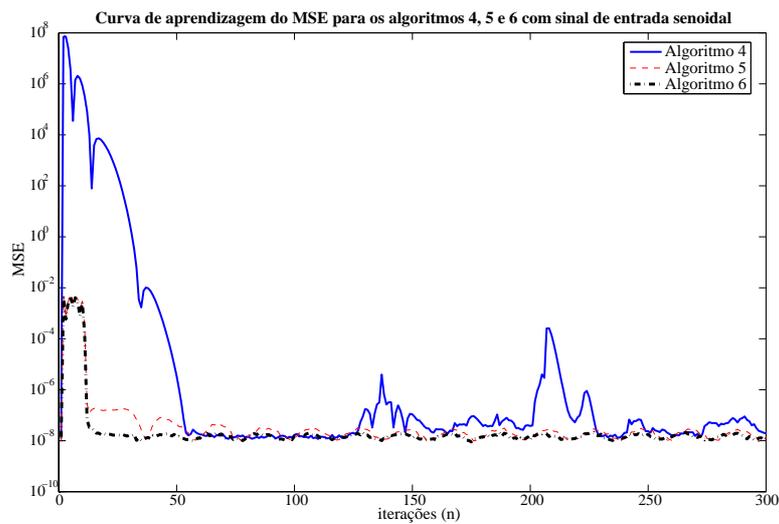


Figura 4.21: MSE dos algoritmos 4–6 para sinal de entrada senoidal e SNR=80dB.

4.5.2 Experimento IIIb: SNR = 40dB

Verifica-se, neste experimento, o comportamento dos algoritmos quando a entrada é senoidal e um ruído de medida com potência moderada, correspondente a SNR igual a 40dB, é aplicado ao sistema. As Figuras 4.22 e 4.23 mostram as curvas de aprendizagem para os algoritmos 1 e 3, e 5 e 6, respectivamente. Os algoritmos

2 e 4 foram omitidos pois ambos divergem quando aplicados a entradas com sinais sem persistência de excitação.

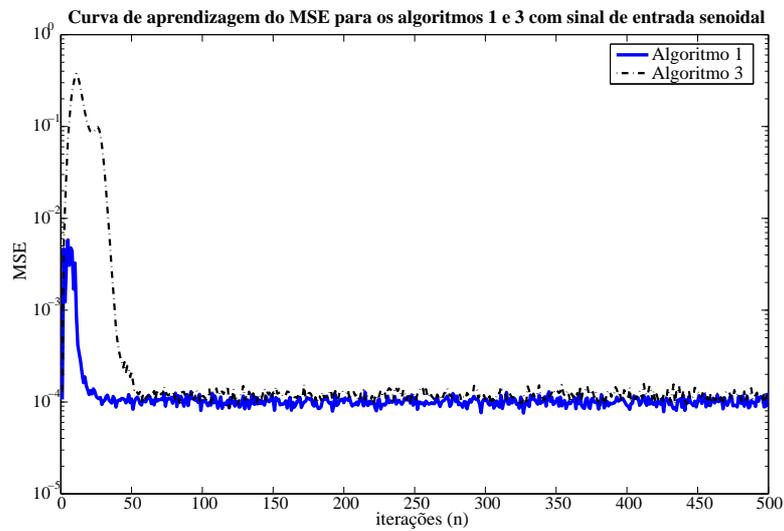


Figura 4.22: MSE dos algoritmos 1 e 3 para sinal de entrada senoidal e SNR=40dB.

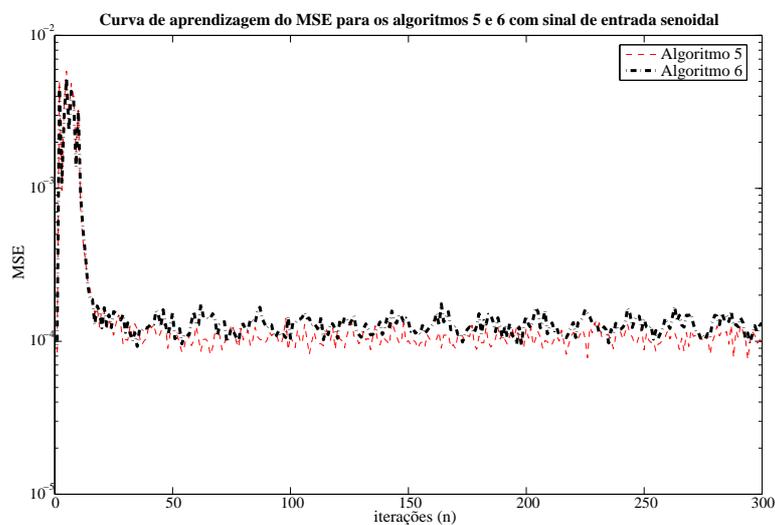


Figura 4.23: MSE dos algoritmos 5 e 6 para sinal de entrada senoidal e SNR=40dB.

Como pode ser observado pelas figuras, a maioria dos algoritmos convergem para o mínimo MSE, sendo que o algoritmo 6 apresenta um desajuste ligeiramente superior ao do algoritmo 5. Entretanto, modificando-se as constantes do algoritmo 6, pode-se reduzir o desajuste do erro quadrático médio.

4.5.3 Experimento IIIc: SNR = 20dB

Para finalizarmos a análise do impacto de entradas senoidais nos algoritmos derivados no capítulo anterior, a Figura 4.24 apresenta a convergência do erro quadrático médio dos algoritmos 1 e 3 com uma razão sinal-ruído de 20dB.

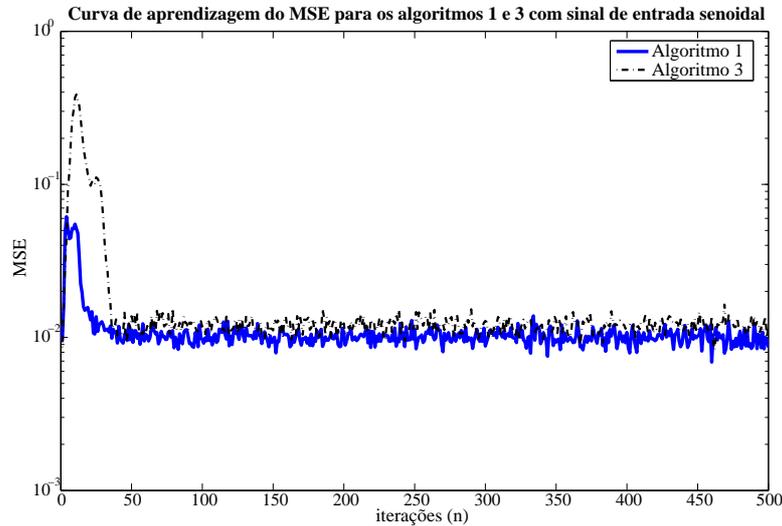


Figura 4.24: MSE dos algoritmos 1 e 3 para sinal de entrada senoidal e SNR=20dB.

Os algoritmos 1 e 3 conseguem convergir normalmente para esse sistema. Podemos considerar, assim, que esses algoritmos são robustos em relação à interferência de ruídos no sistema, e que não têm problemas de convergência devido a entradas não-persistentemente excitantes. O mesmo pode ser dito para os algoritmos 5 e 6, como pode ser observado na Figura 4.25.

4.6 Experimento IV: Sinal de Entrada Senoidal com Ruído

Neste experimento foi utilizado um ruído branco com variância de 0,01 para corromper aditivamente o sinal de entrada senoidal, gerado conforme o experimento anterior. O comportamento do erro médio quadrático dos algoritmos é avaliado em três situações de ruídos de medida.

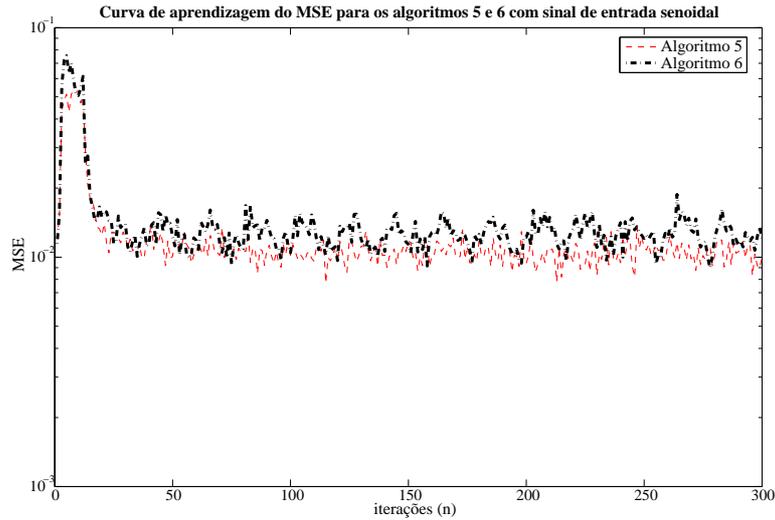


Figura 4.25: MSE dos algoritmos 5 e 6 para sinal de entrada senoidal e SNR=20dB.

4.6.1 Experimento IVa: SNR = 80dB

Primeiramente avalia-se o desempenho dos algoritmos em um ambiente praticamente livre de ruídos, representado por uma razão sinal-ruído de 80dB. Nas Figuras 4.26 e 4.27 observam-se as curvas de aprendizagem do MSE para os algoritmos 1, 2 e 3 e os algoritmos 4, 5 e 6, respectivamente.

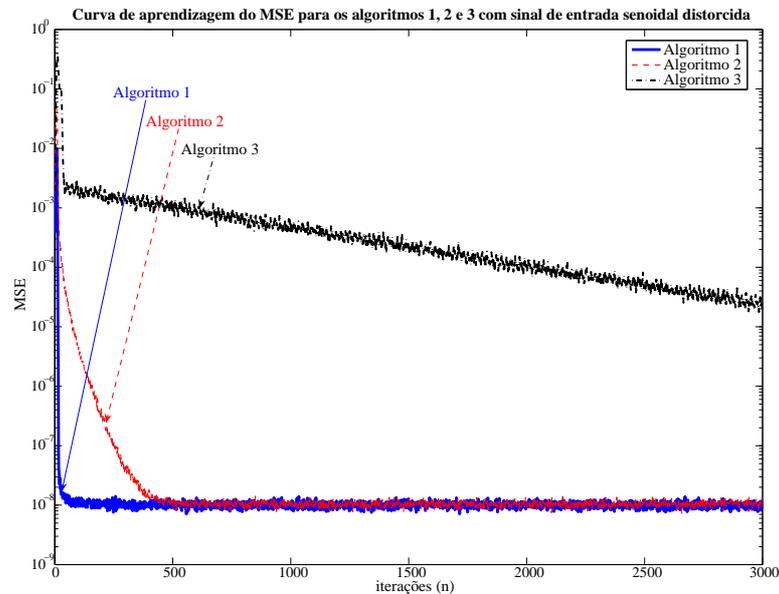


Figura 4.26: MSE dos algoritmos 1–3 para sinal de entrada senoidal distorcida e SNR=80dB.

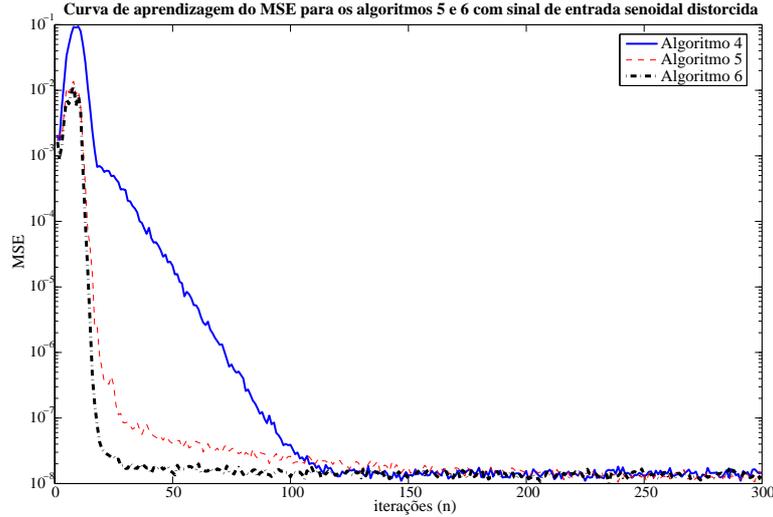


Figura 4.27: MSE dos algoritmos 4–6 para sinal de entrada senoidal distorcida e SNR=80dB.

Observa-se que os algoritmos 2 e 4 que antes divergiam para entradas não persistentemente excitantes, agora convergem. Isso se deve ao ruído que distorce a forma de onda senoidal, fazendo com que o sinal agora seja persistentemente excitante. Já o algoritmo 3 necessita de muito mais iterações para alcançar a convergência, pois o sinal de entrada possui alta correlação do sinal de entrada. A dependência do algoritmo 5 em relação à correlação do sinal de entrada fica mais evidente nesse experimento, levando o algoritmo a necessitar de mais tempo para convergir. O algoritmo 6, por sua vez, mantém um desempenho semelhante ao apresentado para sinais persistentemente excitantes, praticamente independentemente do grau de correlação do sinal de entrada.

4.6.2 Experimento IVb: SNR = 40dB

Para o caso de o ruído de medida do sistema levar a uma SNR igual a 40dB, observa-se que o comportamento dos algoritmos segue padrões semelhantes aos do experimento da Seção 4.6.1: foram resistentes a ruídos moderados para sinais de entrada senoidais distorcidos. As Figuras 4.28 e 4.29 mostram a curva de aprendizagem do MSE dos algoritmos para este caso.

Observa-se na Figura 4.28 que os algoritmos 1 e 2 possuem praticamente a mesma velocidade de convergência, devido principalmente ao fato de a inicialização

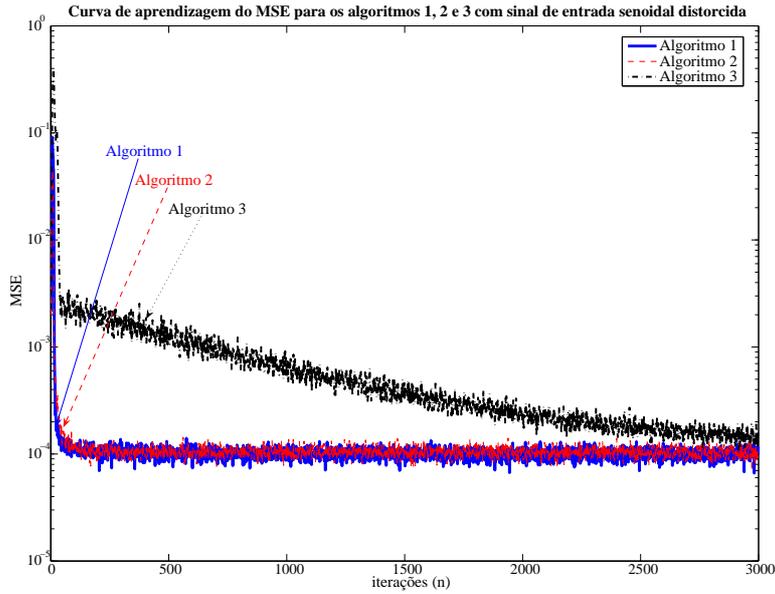


Figura 4.28: MSE dos algoritmos 1–3 para sinal de entrada senoidal distorcida e SNR=40dB.

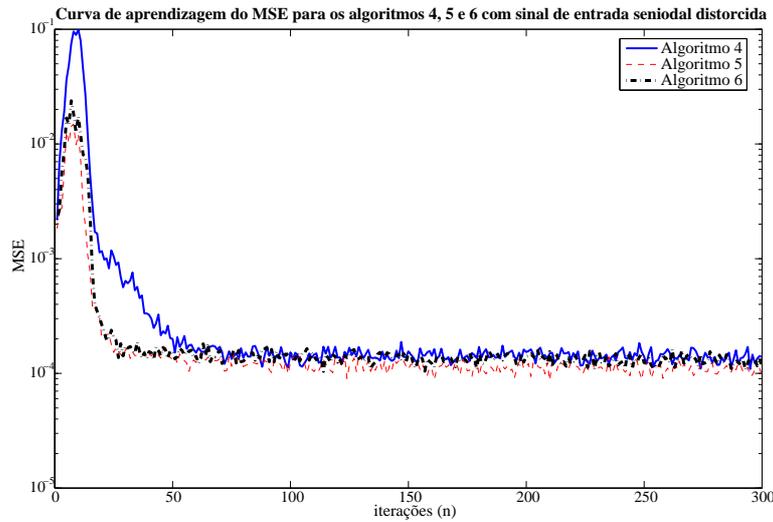


Figura 4.29: MSE dos algoritmos 4–6 para sinal de entrada senoidal distorcida e SNR=40dB.

da inversa da matriz de autocorrelação com um multiplicador $\delta = 1000$ fazer estes algoritmos terem convergência mais rápida até $MSE \approx 10^{-4}$, onde forma-se um “joelho”. O mesmo efeito pode ser observado nos outros algoritmos que utilizam a inversa da matriz de autocorrelação.

Já pela Figura 4.29, pode-se notar que o algoritmo 4 ainda sofre influência

das características de uma entrada senoidal, apresentando um desajuste maior e uma convergência mais ruidosa. Os algoritmos 5 e 6 têm desempenho semelhante ao já apresentado anteriormente para sinais de entrada correlacionados.

4.6.3 Experimento IVc: SNR = 20dB

Para finalizar, o experimento com sinal de entrada senoidal distorcido por ruído branco, avalia-se o desempenho dos algoritmos quando submetidos a um ruído de medida que resulta numa SNR de 20dB. A Figura 4.30 ilustra o desempenho dos algoritmos 1, 2 e 3 para este experimento. Como pode ser observado nestas figuras e nas anteriores a este experimento, com valores de razão sinal-ruído, as convergências iniciais dos algoritmos apresentam características similares às apresentadas anteriormente nas simulações para este sinal de entrada.

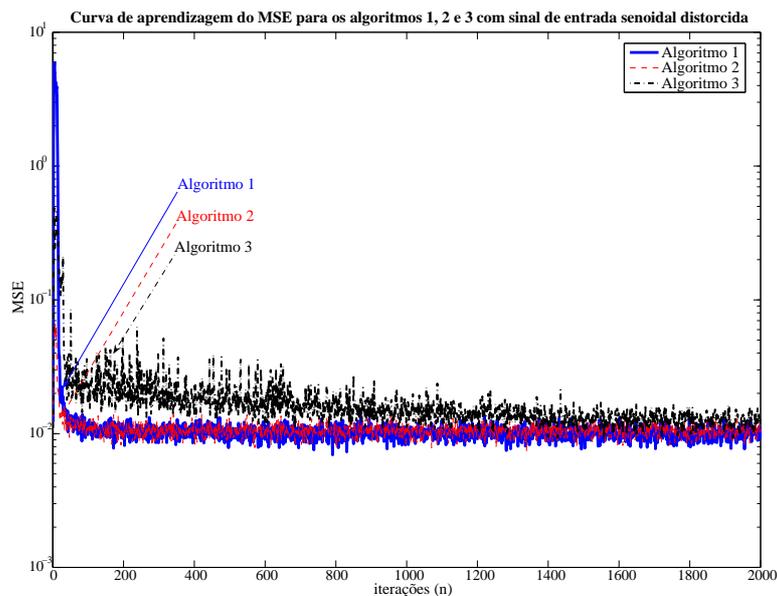


Figura 4.30: MSE dos algoritmos 1–3 para sinal de entrada senoidal distorcida e SNR=20dB.

O desempenho dos algoritmos 4, 5 e 6 pode ser visualizado na Figura 4.31. Os algoritmos 5 e 6 apresentam velocidades de convergência semelhantes, com características semelhantes à convergência para sinal de entrada senoidal puro. Já o algoritmo 4 leva mais tempo para convergir, além de ter seu comportamento mais ruidoso.

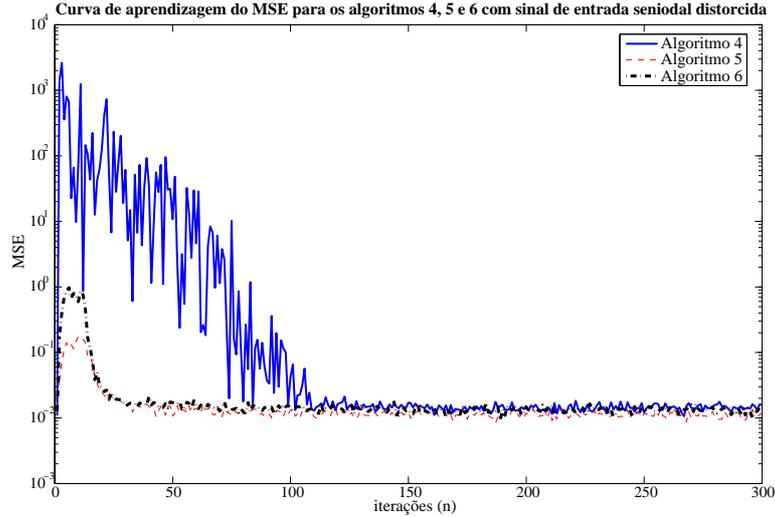


Figura 4.31: MSE dos algoritmos 4–6 para sinal de entrada senoidal distorcida e SNR=20dB.

4.7 Experimento V: Mudança Abrupta dos Coeficientes

Neste experimento, avalia-se a capacidade dos algoritmos em situações onde os coeficientes do sistema desconhecido têm seus valores modificados após a convergência dos algoritmos, exigindo uma readaptação dos coeficientes dos filtros adaptativos. A ordem dos filtros adaptativos e do sistema desconhecido são iguais (10 coeficientes), e analisa-se o desempenho dos algoritmos através da curva de aprendizagem do erro médio quadrático. A mudança dos coeficientes é realizada na iteração $n = 500$, quando todos os algoritmos já alcançaram a convergência. É considerado um ambiente praticamente livre de interferência (SNR = 80dB), e aplicam-se os mesmos tipos de entrada utilizados nas outras seções: ruído gaussiano com espalhamento dos autovalores da matriz de autocorrelação do sinal de entrada igual a 1 (ruído branco) e aproximadamente 450 (ruído colorido pelo mesmo filtro passa-baixas utilizado na Seção 4.4), sinal de entrada senoidal e senoidal distorcido.

4.7.1 Experimento Va: Sinal de Entrada Ruído Branco

A curva de aprendizado dos algoritmos 1, 2 e 3 para este experimento com sinal de entrada do tipo WGN é apresentado na Figura 4.32, onde observa-se que

o melhor desempenho foi do algoritmo 3 (NLMS), e que o algoritmo 1 não possui grande capacidade de esquecer informações passadas, tornando a convergência após a mudança dos coeficientes ótimos acentuadamente mais lenta. O algoritmo 2 tem um desempenho relativamente melhor que o algoritmo 1 porém as informações passadas ainda têm grande efeito sobre o algoritmo. A comparação com o algoritmo 3 é relativamente complicada, uma vez que esse algoritmo não utiliza a matriz de autocorrelação do sinal de entrada, considerando apenas as informações mais recentes no vetor do sinal de entrada.

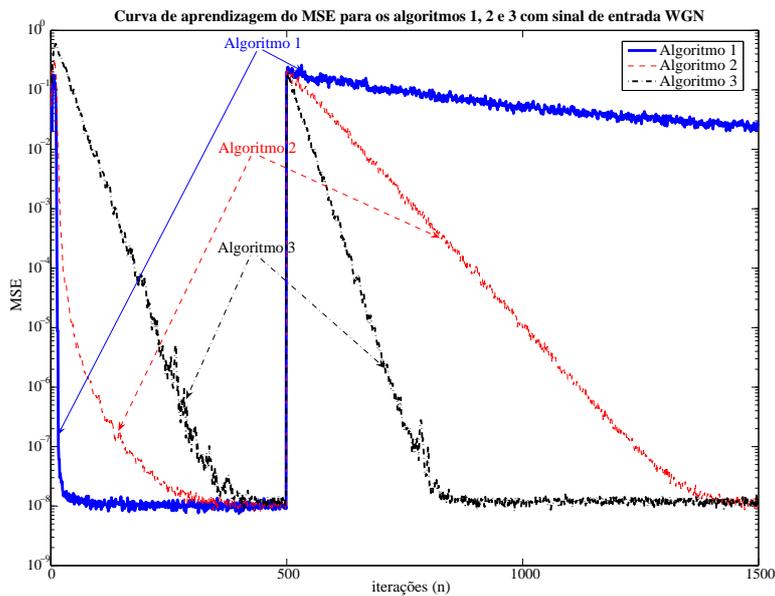


Figura 4.32: MSE dos algoritmos 1–3 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada WGN.

Na Figura 4.33 observa-se o comportamento da convergência do MSE dos algoritmos com passo variável. Como visto na figura, o algoritmo 4 consegue acompanhar melhor a mudança do sistema, o qual deve ter desempenho parecido com o algoritmo RLS, pois ambos foram derivados a partir da mesma função-objetivo. O algoritmo 6 não consegue convergir após a mudança dos coeficientes, devido a o passo do algoritmo fixar praticamente em zero após a convergência inicial, parando de atualizar tanto a matriz de autocorrelação quanto o vetor de coeficientes do filtro adaptativo.

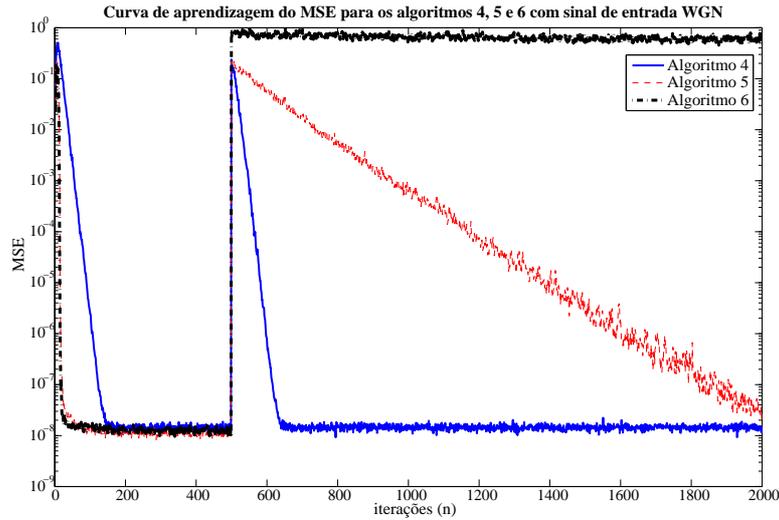


Figura 4.33: MSE dos algoritmos 4–6 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada WGN.

4.7.2 Experimento Vb: Sinal de Entrada Ruído Colorido

Os algoritmos que utilizam a matriz de autocorrelação demonstraram, nas outras seções deste capítulo, que um sinal de entrada correlacionado não exerce muita interferência em seus desempenhos. Verifica-se agora qual a influência que o uso deste tipo de sinal de entrada tem sobre os algoritmos quando o sistema a ser identificado altera o valor de seus coeficientes após a convergência dos filtros adaptativos. As Figuras 4.34 e 4.35 ilustram o comportamento do MSE para os algoritmos neste experimento.

Como pode ser observado nas figuras, os algoritmos não tem queda de desempenho devido a o sinal de entrada ser correlacionado. O algoritmo 3, como observado na Seção 4.4, converge lentamente quando o sinal de entrada é correlacionado: quanto maior a correlação pior o desempenho do algoritmo, que não consegue convergir mesmo antes da troca dos coeficientes. No algoritmo 5 observa-se a influência do sinal correlacionado quando há a troca de coeficientes do sistema. O algoritmo 1 continua com convergência lenta após a mudança dos coeficientes e o algoritmo 6 não consegue readaptar o filtro.

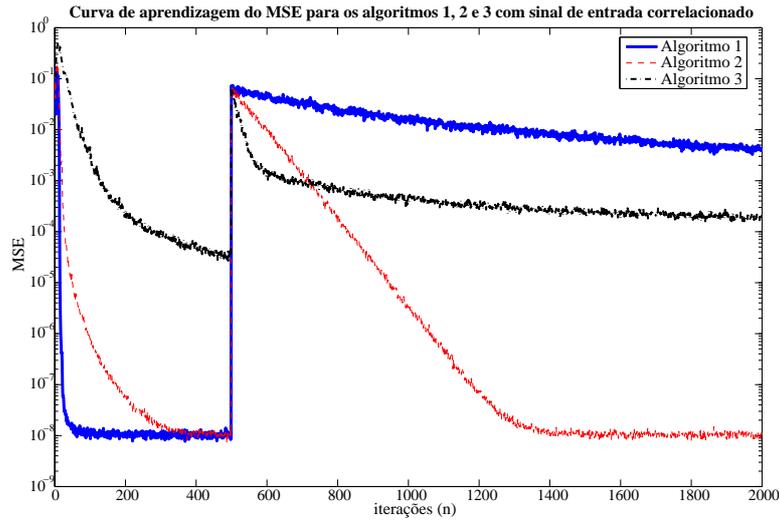


Figura 4.34: MSE dos algoritmos 1–3 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada correlacionado.

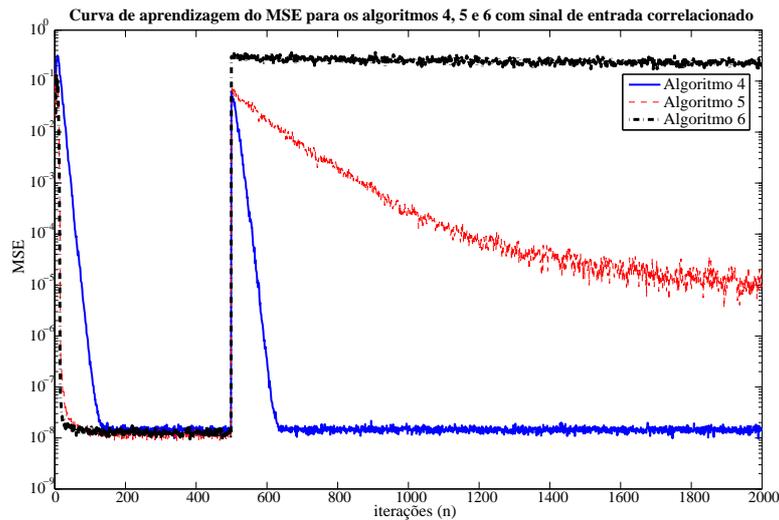


Figura 4.35: MSE dos algoritmos 4–6 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada correlacionado.

4.7.3 Experimento Vc: Sinal de Entrada Senoidal

O uso de entrada senoidal faz com que alguns algoritmos não consigam convergir. Este é o caso do algoritmo RLS e outros dois derivados no capítulo anterior e avaliados nas simulações da Seção 4.5. Verifica-se, neste experimento, o comportamento do erro médio quadrático dos algoritmos quando se utiliza sinal de entrada não persistentemente excitante e com modificação abrupta dos coeficientes do sistema

desconhecido após a convergência dos algoritmos. O desempenho dos algoritmos 1 e 3 é ilustrado na Figura 4.36. O algoritmo 1 apresenta o mesmo comportamento, sendo um algoritmo robusto em relação ao sinal de entrada, porém não consegue bom desempenho quando há a troca abrupta do sistema desconhecido. E o algoritmo 3, o algoritmo NLMS, apresenta um ótimo desempenho, para este tipo de entrada.

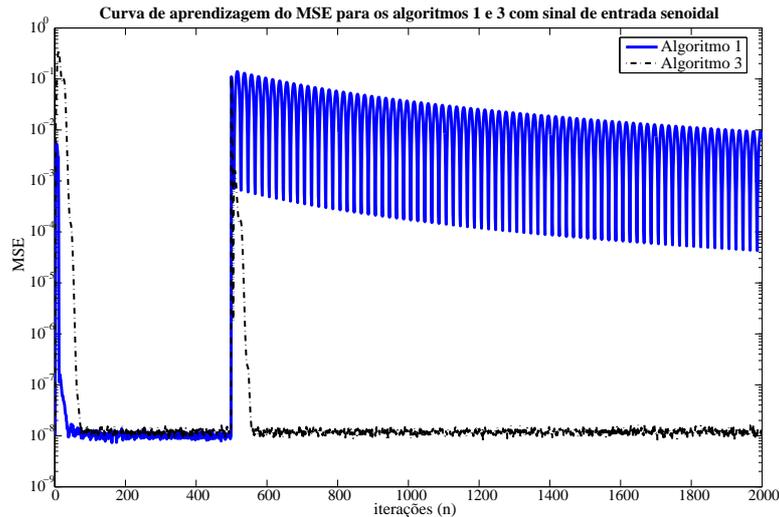


Figura 4.36: MSE dos algoritmos 1 e 3 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada senoidal.

O desempenho dos algoritmos 5 e 6 para este experimento pode ser observado na Figura 4.37. Para esse tipo de entrada, o passo do algoritmo 5 tende para 1 e desse modo pára de atualizar a matriz de autocorrelação e os coeficientes do filtro. Isto, por sua vez, faz com que o algoritmo mantenha os valores ótimos anteriores e não os atualize, apesar da mudança no vetor de coeficientes ótimos. O algoritmo 6 continua sem conseguir readaptar os coeficientes do filtro adaptativo após a modificação no sistema desconhecido.

Os passos dos algoritmos 5 e 6 tendem para um e zero, respectivamente, fazendo com que os algoritmos não consigam readaptar os coeficientes do filtro. Isso pode ser observado pela Figura 4.38.

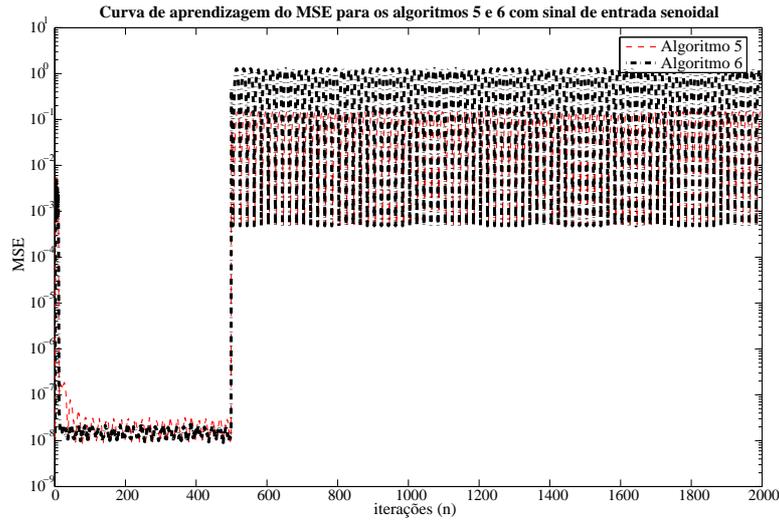


Figura 4.37: MSE dos algoritmos 4–6 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada senoidal.

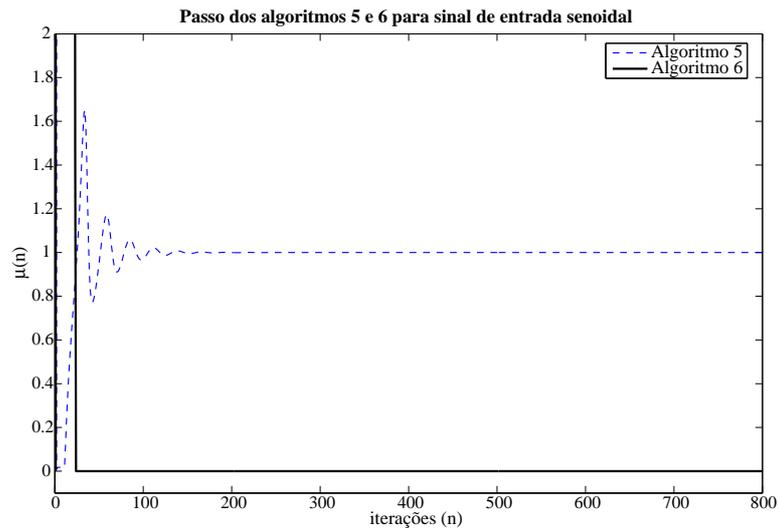


Figura 4.38: Passos dos algoritmos 5 e 6 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada senoidal.

4.7.4 Experimento Vd: Sinal de Entrada Senoidal com Ruído

O uso de um sinal de entrada senoidal distorcido por um ruído apresentou resultados interessantes. Enquanto alguns algoritmos mantiveram seus desempenhos semelhantes aos obtidos para entrada senoidal pura, outros tiveram comportamento semelhante ao apresentado com entrada WGN. Isso se deve ao nível de ruído inserido na forma de onda senoidal do sinal de entrada: quanto maior a distorção, mais

próximo do obtido para um sinal de ruído aleatório será o comportamento observado do sistema. De forma contrária, quanto menor o ruído que distorce a senóide, mais próximo será o desempenho dos algoritmos se assemelha daquel obtido com sinal de entrada senoidal. O mesmo efeito pode ser observado quando existe a troca abrupta dos coeficientes do sistema desconhecido. Observa-se nas Figuras 4.39 e 4.40 a convergência do erro médio quadrático dos algoritmos.

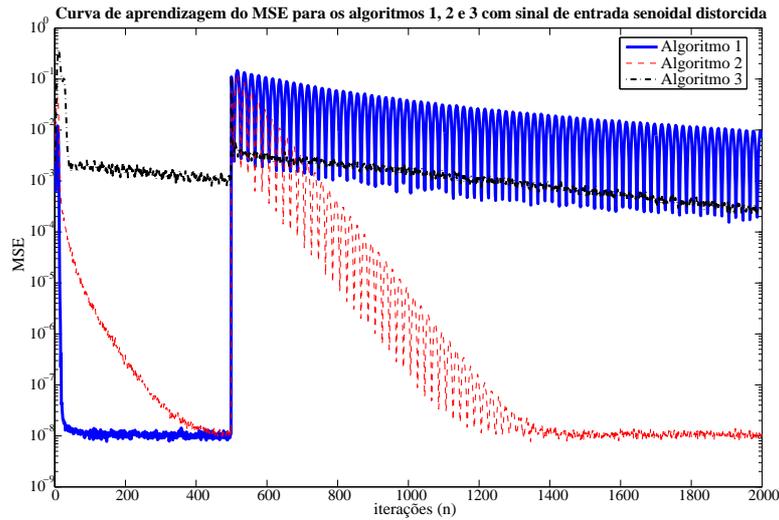


Figura 4.39: MSE dos algoritmos 1–3 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada senoidal distorcida.

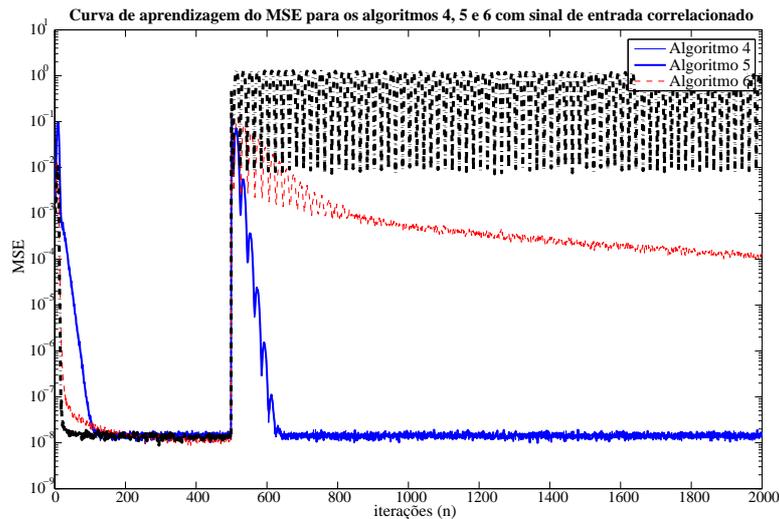


Figura 4.40: MSE dos algoritmos 4–6 para mudança abrupta dos coeficientes do sistema desconhecido para sinal de entrada senoidal distorcida.

O algoritmo 3 tem um desempenho parecido como o apresentado com sinal de entrada correlacionado. Os algoritmos 2 e 5 conseguem convergir para esta entrada, uma vez que o primeiro não está mais exposto a uma entrada sem excitação persistente, e no segundo o passo do algoritmo não converge para 1. O algoritmo 4 consegue convergir do mesmo modo que ocorreu na Seção 4.6; podemos, assim, concluir que este algoritmo é robusto para sistemas com variação abrupta dos coeficientes, desde que sua entrada seja persistentemente excitante. O algoritmo 6 não consegue readaptar os coeficientes, como observado em todas as simulações de variação abrupta dos coeficientes e, assim como o algoritmo 1, não deve ser considerado nesse caso.

4.8 Simulações Comparativas

Nesta seção, apresentam-se simulações comparativas entre o algoritmo RLS, o algoritmo quasi-Newton apresentado em [4] e alguns dos algoritmos desenvolvidos no Capítulo 3, os que obtiveram melhor desempenho nas simulações de convergência. Os algoritmos escolhidos para as simulações comparativas foram escolhidos de acordo com o tipo de teste realizado. O algoritmo 1, dos algoritmos com passo fixo, e o algoritmo 6, dos algoritmos com passo variável, são utilizados nas simulações I a IV, onde os coeficientes do sistema desconhecido são fixos durante toda a simulação e se aplicam os quatro tipos de sinais de entrada avaliados nas simulações de convergência: ruído branco, ruído colorido, senóide e senóide com ruído. Para as comparações com variação abrupta dos coeficientes, entre os algoritmos com passo fixo foi escolhido o algoritmo 2, que utiliza a matriz de autocorrelação em relação ao algoritmo 3, que é o algoritmo NLMS já conhecido; dos algoritmos com passo variável, escolheu-se o algoritmo 4, devido ao seu melhor desempenho nesse tipo de situação.

As simulações serão realizadas aplicando os algoritmos adaptativos à identificação de um sistema desconhecido com 10 coeficientes. A razão sinal-ruído será de 80dB, de modo a interpretar a velocidade de convergência dos algoritmos, uma vez que pelas simulações de convergência observou-se que os algoritmos convergem independentemente do ruído adicionado ao sistema.

Os algoritmos são simulados até alcançarem a convergência, e é realizada a

média de 200 simulações para se obter a curva de convergência do erro quadrático médio dos algoritmos. Para o algoritmo 1 foi utilizado um passo $\mu(n) = 100$, o mesmo utilizado nas simulações anteriores; já o algoritmo RLS tem seu fator de esquecimento $\lambda = 0,98$, de modo a se atingir desajuste semelhante ao do algoritmo 1. O algoritmo 6 tem suas constantes definidas para se obter desajuste próximo ao do algoritmo 1 e boa velocidade de convergência: $\alpha = 0,00001$, $\beta = 0,01$, $t_1 = 100$ e $t_2 = 1$. O algoritmo QN não precisa de nenhuma constante definida pelo usuário, porém a iniciação da matriz de autocorrelação é feita do mesmo modo que nos demais algoritmos para uma comparação justa. O valor do fator de multiplicação da inversa da matriz de autocorrelação do sinal de entrada para todos os algoritmos é $\delta = 1$, pois é o valor com que se encontra nas simulações realizadas em artigos. O algoritmo QN possui um desajuste, devido à imposição da normalização do erro *a posteriori* igual a zero, de aproximadamente 3dB.

Nas simulações com variação abrupta dos coeficientes do sistema a ser identificado, o valor do passo do algoritmo 2 foi definido para obter um desajuste similar ao do algoritmo 4, $\mu(n) = 0,1$. O algoritmo RLS tem seu fator de esquecimento ajustado para $\lambda = 0,90$, para obter o mesmo desajuste que os algoritmos 2 e 4. As simulações com sinal de entrada senoidal foram omitidas, pois os únicos algoritmos que conseguiram convergir para este tipo de entrada e com variação abrupta dos coeficientes do sistema a ser identificado foram o algoritmo QN e o algoritmo 3; porém, o algoritmo 3 é o algoritmo LMS normalizado, já conhecido na literatura, o que dispensa os testes.

4.8.1 Simulações Comparativas I: Coeficientes Fixos e Sinal de Entrada Ruído Branco

Considera-se inicialmente o uso de ruído branco como sinal de entrada, utilizando-se uma distribuição Gaussiana por conveniência. A Figura 4.41 ilustra a curva de aprendizagem do MSE para os algoritmos simulados. Como pode ser observado, os algoritmos 1 e 6 tiveram desempenhos próximos, sendo que o algoritmo 6 consegue convergir mais rapidamente. Por sua vez, o algoritmo QN consegue convergir mais rapidamente que o algoritmo RLS, que tem uma velocidade de convergência inferior para poder garantir um excesso de MSE menor. Entretanto, o algoritmo QN tem um

desajuste de 3dB em relação ao mínimo MSE, como pode ser observado na figura. Esse desajuste pode ser considerado desprezível em situações com razão sinal-ruído elevada.

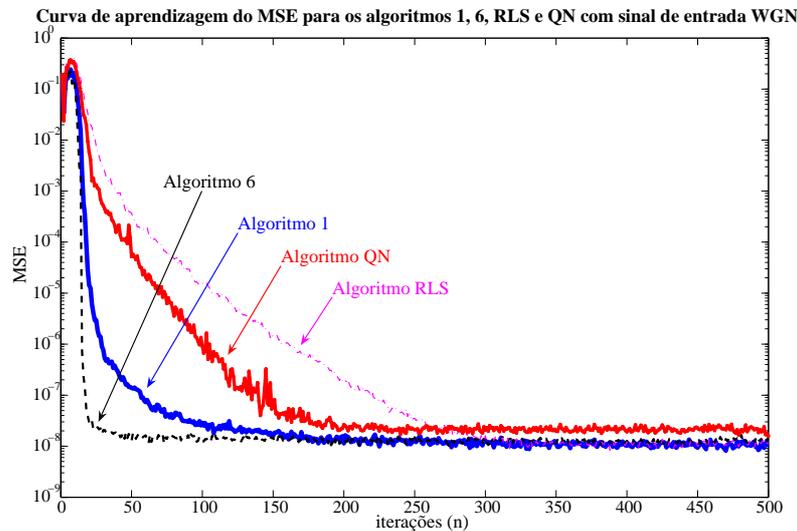


Figura 4.41: MSE dos algoritmos 1, 6, RLS e QN para sistema com coeficientes fixos e sinal de entrada WGN.

Destaca-se que o algoritmo 1 pode ter sua velocidade de convergência melhorada com o aumento no passo do algoritmo, sem causar aumento no desajuste. Em contrapartida, o algoritmo 6 possui diversas variáveis que podem ser alteradas, modificando a velocidade de convergência e o desajuste mínimo porém, com a barreira mantendo a matriz de autocorrelação limitada, evitando que o acréscimo de elementos a cada nova iteração leve a matriz a deixar de ser definida positiva.

4.8.2 Simulações Comparativas II: Coeficientes Fixos e Sinal de Entrada Ruído Colorido

A correlação do sinal de entrada não tem grande interferência no desempenho da maioria dos algoritmos recursivos baseados na função-objetivo de mínimos quadrados desenvolvidos nesta tese, como já foi comprovado nas simulações anteriores. Na Figura 4.42 apresenta-se o desempenho dos algoritmos quando o sinal de entrada é correlacionado por um filtro passa-baixas com frequência de corte em $0,35w_s$. Nota-se pela figura que o desempenho dos algoritmos é similar ao obtido utilizando-se sinal de entrada WGN, com exceção do algoritmo quasi-Newton, que

tem sua velocidade de convergência comprometida, que resulta inferior, inclusive, ao algoritmo RLS. Novamente, os algoritmos 1 e 6 destacam-se pelas velocidades de convergência próximas, ao mesmo tempo obtendo melhor desempenho em relação ao menor valor de MSE, em relação ao algoritmo QN.

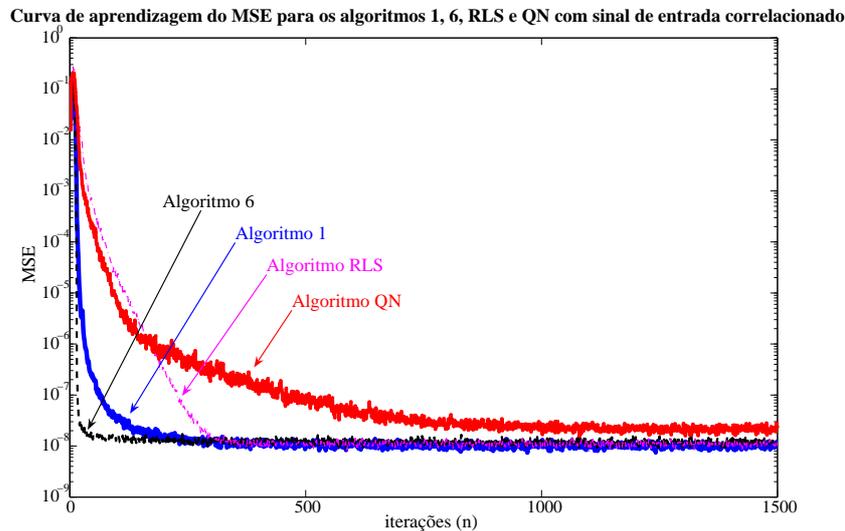


Figura 4.42: MSE dos algoritmos 1, 6, RLS e QN para sistema com coeficientes fixos e sinal de entrada ruído colorido.

4.8.3 Simulações Comparativas III: Coeficientes Fixos e Sinal de Entrada Senoidal

Nesta simulação utiliza-se como sinal de entrada dos sistemas adaptativos e desconhecido uma senóide, para a qual sabe-se que o algoritmo RLS apresenta problemas de convergência. Esse fato pode ser observado pela Figura 4.43, onde observa-se que os outros algoritmos conseguem convergir, enquanto que o algoritmo RLS mesmo convergindo inicialmente, termina por divergir em função de a matriz de autocorrelação do sinal de entrada deixar de ser definida positiva.

Na Figura 4.44, mostram-se apenas as 500 primeiras iterações dos algoritmos, onde pode-se melhor observar seus desempenhos iniciais. O algoritmo 6 continua tendo o melhor desempenho em relação à velocidade de convergência, seguido do algoritmo 1. O algoritmo RLS consegue convergir rapidamente para esse tipo de entrada, mas, como vimos, esse tipo de sinal leva o algoritmo à divergência. Por fim,

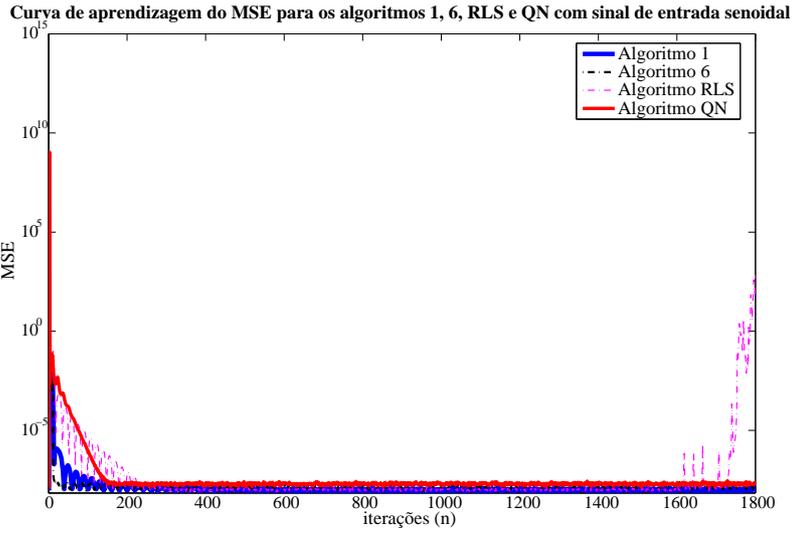


Figura 4.43: MSE dos algoritmos 1, 6, RLS e QN para sistema com coeficientes fixos e sinal de entrada senoidal.

o algoritmo QN converge aproximadamente na mesma velocidade que o algoritmo RLS, sem ter problemas de divergir, mas com desajuste de 3dB.

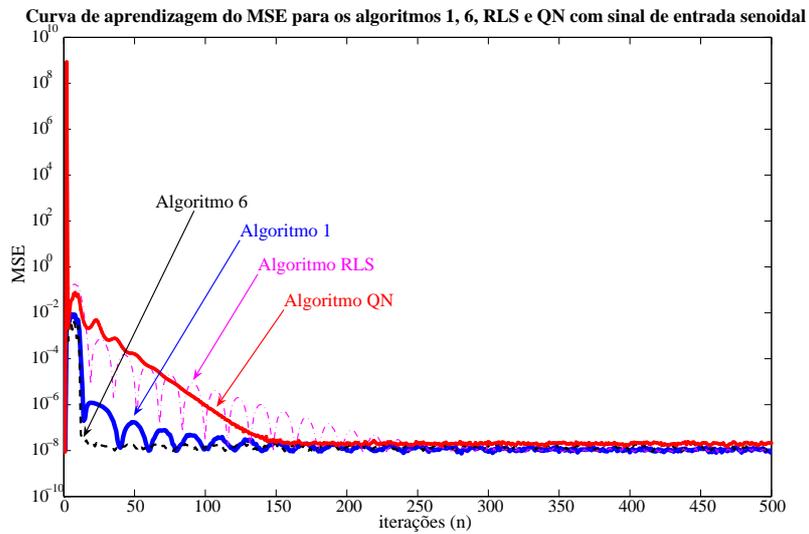


Figura 4.44: MSE dos algoritmos 1, 6, RLS e QN para sistema com coeficientes fixos e sinal de entrada senoidal.

4.8.4 Simulações Comparativas IV: Coeficientes Fixos e Sinal de Entrada Senoidal Distorcido

Para sistemas desconhecidos com coeficientes fixos e sinal de entrada senoidal distorcido por ruído branco, os algoritmos tendem a apresentar desempenho similar aos seus desempenhos com sinais de entrada senoidal ou ruído branco, dependendo do nível de distorção inserida no sinal senoidal. Para esta simulação utilizou-se o mesmo nível de distorção das simulações de convergência, ou seja, um ruído branco com variância de 0,01 é adicionado ao sinal senoidal de entrada dos sistemas.

Observa-se na Figura 4.45 o comportamento do erro quadrático médio dos algoritmos, onde destaca-se que o algoritmo RLS não diverge. O algoritmo 6 continua com a melhor velocidade de convergência entre os algoritmos da simulação, devido ao alto valor inicial do passo do algoritmo. O algoritmo 1 sofre uma perda de velocidade devido a o sinal de entrada ser muito próximo de uma entrada que não é permanentemente excitante. O algoritmo QN tem seu desempenho comprometido, mostrando a dependência que o algoritmo tem em relação ao sinal de entrada.

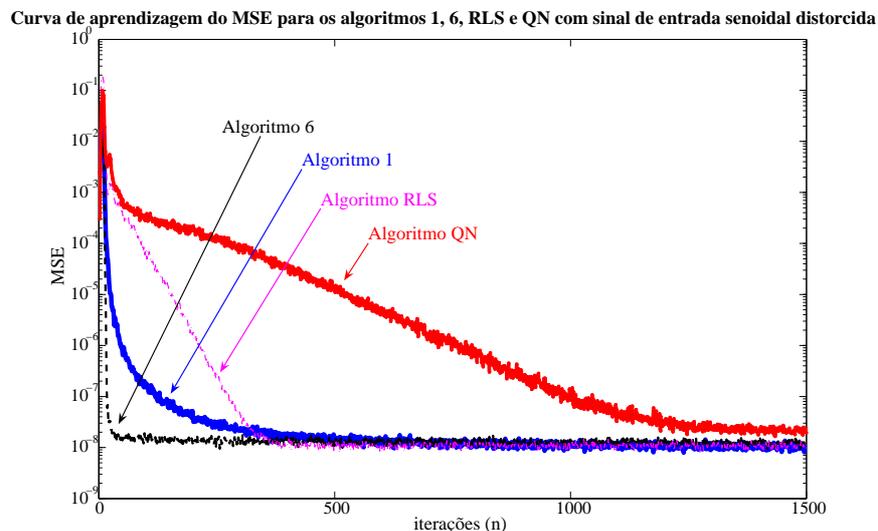


Figura 4.45: MSE dos algoritmos 1, 6, RLS e QN para sistema com coeficientes fixos e sinal de entrada senoidal distorcido.

4.8.5 Simulações Comparativas V: Variação Abrupta dos Coeficientes e Sinal de Entrada Ruído Branco

Analisa-se agora casos em que após os algoritmos atingirem a convergência os coeficientes do sistema a ser identificado são modificados. Para os casos de variação abrupta dos coeficientes, os algoritmos 1 e 6 foram substituídos pelos algoritmos 2 e 4, respectivamente. A escolha dos algoritmos para comparação já foi explicada anteriormente.

A Figura 4.46 ilustra a curva de aprendizagem do MSE para este caso, quando o sinal de entrada dos sistemas desconhecido e adaptativo é gerado por uma distribuição Gaussiana com variância igual a 1. O desajuste dos algoritmos é superior ao dos algoritmos nas simulações comparativas I a IV, devido ao ajuste dos algoritmos ao mínimo MSE alcançado pelo algoritmo 4. Com isso, o algoritmo RLS atinge maior velocidade de convergência, devido a o passo utilizado poder relaxar o fator de esquecimento.

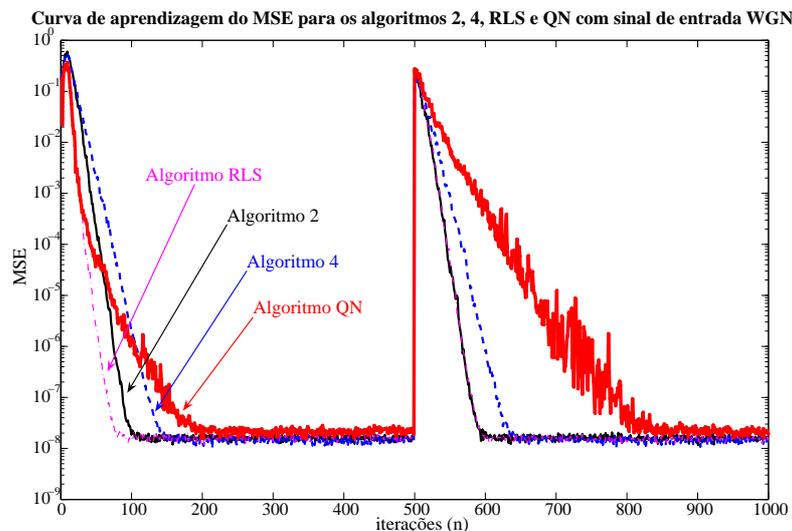


Figura 4.46: MSE dos algoritmos 1, 6, RLS e QN para sistema com variação abrupta dos coeficientes e sinal de entrada WGN.

Como pode ser observado pela figura, o algoritmo RLS tem a melhor velocidade de convergência, seguido do algoritmo 2. Porém, quando os algoritmos precisam readaptar os coeficientes, o algoritmo 2 consegue uma velocidade de convergência equivalente à do algoritmo RLS. O algoritmo 4 continua possuindo boa velocidade de convergência, enquanto que o algoritmo quasi-Newton leva mais iterações

para conseguir reajustar os coeficientes do filtro adaptativo. Mesmo com o desajuste superior ao dos algoritmos 1 e 6, o algoritmo 4 ainda consegue menor valor de desajuste que o algoritmo QN.

4.8.6 Simulações Comparativas VI: Variação Abrupta dos Coeficientes e Sinal de Entrada Ruído Colorido

Avalia-se o desempenho dos algoritmos 2 e 4 em comparação aos algoritmos RLS e QN quando há a variação dos coeficientes do sistema desconhecido após a convergência dos algoritmos, utilizando um sinal de entrada correlacionado. O sinal de entrada é gerado por uma distribuição Gaussiana e passado pelo mesmo filtro utilizado nas seções em que utilizou-se sinal de entrada correlacionado, resultando em um espalhamento dos autovalores da matriz de autocorrelação do sinal de entrada, $\lambda_{\text{máx}}/\lambda_{\text{mín}}$, de aproximadamente 450. Diferentemente da simulação anterior, os coeficientes do filtro a ser identificado são modificados na iteração 1000, devido ao fato de a velocidade de convergência do algoritmo QN ser consideravelmente lenta para esse tipo de sinal de entrada.

Observando-se a Figura 4.47, nota-se que os algoritmos têm desempenhos independentes da correlação do sinal de entrada, excluindo-se o algoritmo QN, que tem sua convergência comprometida. Novamente o algoritmo RLS teve melhor desempenho na convergência inicial, e mesma velocidade de readaptação dos coeficientes que o algoritmo 2. O algoritmo 4 continua com boa velocidade de convergência e velocidade equivalente de readaptação dos coeficientes do filtro adaptativo, enquanto que o algoritmo QN tem tanto a velocidade de convergência inicial quanto a velocidade para readaptação dos coeficientes do filtro adaptativo reduzidas pela correlação do sinal de entrada.

4.8.7 Simulações Comparativas VII: Variação Abrupta dos Coeficientes e Sinal de Entrada Senoidal Distorcido

Finalizando as simulações, considera-se o caso em que os coeficientes do sistema desconhecido são modificados após a convergência dos algoritmos, e o sinal de entrada utilizado é senoidal distorcido por um ruído branco com variância de 0,01.

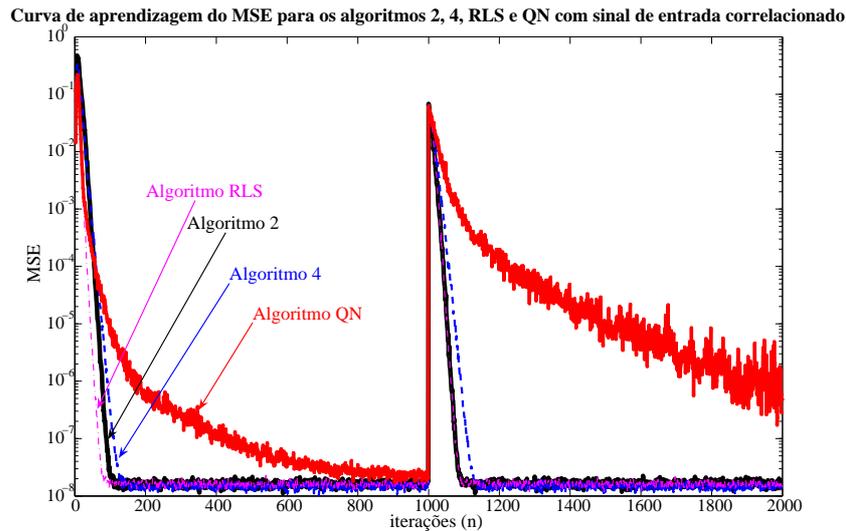


Figura 4.47: MSE dos algoritmos 1, 6, RLS e QN para sistema com variação abrupta dos coeficientes e sinal de entrada ruído colorido.

O algoritmo QN nesse caso tem sua velocidade de convergência inferior à do caso do sinal de entrada correlacionado da simulação anterior, Subseção 4.8.6.

A Figura 4.48 mostra que o desempenho dos algoritmos, para este caso continua o mesmo visto nas simulações anteriores, com o algoritmo RLS conseguindo melhor velocidade de convergência inicial e o algoritmo 2 obtendo convergência na readaptação dos coeficientes do filtro adaptativo similar à do algoritmo RLS. O algoritmo 4 mantém o desempenho visto anteriormente, tendo velocidades similares na convergência inicial e na readaptação dos coeficientes do filtro adaptativo.

Após todas as simulações realizadas nesse capítulo, alguns considerações podem ser feitas sobre os algoritmos. O algoritmo 1 é aparentemente um algoritmo robusto quanto ao sinal de entrada, não conseguindo, entretanto, adaptar os coeficientes do filtro quando o sistema muda suas características. A função-objetivo do algoritmo é conhecida e utilizada em algoritmos BEACON, onde se utilizam restrições no algoritmo para reduzir a necessidade de adaptação de acordo com um limiar. Apesar de o passo do algoritmo não influenciar no valor de desajuste do algoritmo, que é nulo, sua matriz inversa de autocorrelação do sinal de entrada aumenta indefinidamente. Para conter esse problema, foi desenvolvido o algoritmo 6, em que é proposto um passo variável de modo a manter a matriz dentro de um limiar. Observou-se que o algoritmo mantém as características positivas de rápida

Curva de aprendizagem do MSE para os algoritmos 2, 4, RLS e QN com sinal de entrada senoidal distorcida

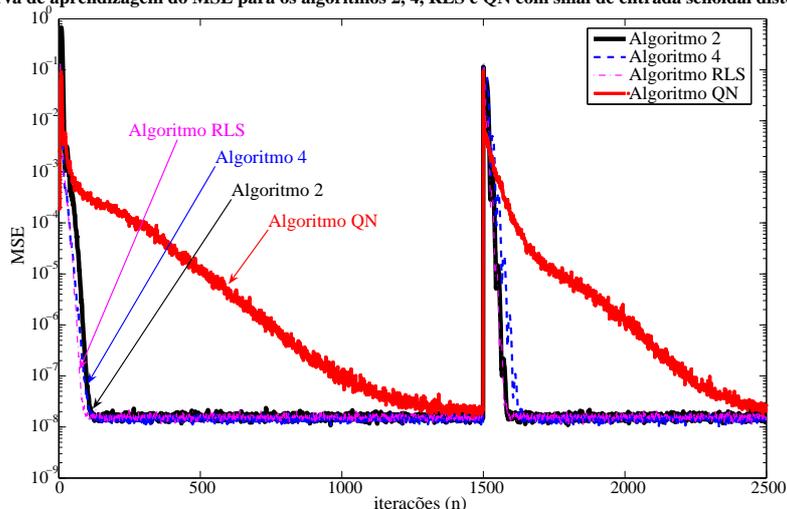


Figura 4.48: MSE dos algoritmos 1, 6, RLS e QN para sistema com variação abrupta dos coeficientes e sinal de entrada senoidal distorcido.

convergência do algoritmo 1, porém com um leve acréscimo no desajuste.

O algoritmo 2 surge como uma opção para situações em que é necessária a readaptação dos coeficientes do filtro. Esse algoritmo tem uma função-objetivo convexa, tendo seu desempenho entre o do algoritmo 1 e o do algoritmo RLS, dependendo da escolha do passo. O algoritmo 3 foi utilizado para comparação do desempenho dos algoritmos com sinal de entrada correlacionado, por ser o algoritmo NLMS conhecido da literatura. O algoritmo 4 surge como uma opção para entradas persistentemente excitantes, porém sem necessidade de escolha do valor do passo do algoritmo. Seu desempenho é praticamente o mesmo independentemente do sinal de entrada. O algoritmo 5, apesar de ter um ótimo desempenho, comparável ao do algoritmo 6, depende diretamente da iniciação da matriz de autocorrelação. O efeito dessa iniciação não apenas sobre o algoritmo 5, mas em algoritmos adaptativos em geral, deve ser estudado, podendo acelerar a convergência inclusive do algoritmo RLS. A Tabela 4.2 resume as características de cada algoritmo desenvolvido nesta tese, com exceção do algoritmo 3, que tem suas características já exploradas na literatura.

Tabela 4.2: Características dos Algoritmos Desenvolvidos

Algoritmo	Características Positivas	Características Negativas
Algoritmo 1	<ul style="list-style-type: none"> - Desajuste mínimo (nulo) e independente do valor do passo do algoritmo (passo constante). - Convergência para sinais não-persistentemente excitantes. - Robustez e rápida convergência, independente das características do sinal de entrada. 	<ul style="list-style-type: none"> - Aumento indefinido da matriz de autocorrelação do sinal de entrada pode levar o algoritmo à divergência com o número de iterações tendendo ao infinito. - Não consegue acompanhar mudanças nos coeficientes do sistema.
Algoritmo 2	<ul style="list-style-type: none"> - Capacidade de acompanhar mudanças que podem ocorrer no sistema. - Velocidade de convergência e desajuste comparáveis aos do algoritmo RLS. 	<ul style="list-style-type: none"> - Diverge para sinais de entrada não-persistentemente excitantes.
Algoritmo 4	<ul style="list-style-type: none"> - Não necessita de ajuste de parâmetros pelo usuário. - Capacidade de acompanhar mudanças que podem ocorrer no sistema. - Alternativa ao algoritmo QN. 	<ul style="list-style-type: none"> - Diverge para sinais de entrada não-persistentemente excitantes. - Desajuste fixo, apesar de ser menor que o do algoritmo QN.
Algoritmo 5	<ul style="list-style-type: none"> - Convergência para sinais não-persistentemente excitantes. - Restrição na matriz de autocorrelação do sinal de entrada, garantido que seja sempre positiva definida. 	<ul style="list-style-type: none"> - Dependência da iniciação da matriz de correlação do sinal de entrada. - Correlação do sinal de entrada afeta o desempenho do algoritmo. - Não consegue acompanhar mudanças nos coeficientes do sistema.
Algoritmo 6	<ul style="list-style-type: none"> - Robustez e rápida convergência, independente das características do sinal de entrada. - Convergência para sinais não-persistentemente excitantes. - Restrição na matriz de autocorrelação do sinal de entrada, garantido que seja sempre positiva definida. 	<ul style="list-style-type: none"> - Muitos parâmetros de ajuste pelo usuário, com grande relevância no desempenho do algoritmo. - Não consegue acompanhar mudanças nos coeficientes do sistema.

Capítulo 5

Considerações Finais

Algoritmos adaptativos vêm sendo utilizados atualmente em diversas áreas da engenharia, constituindo uma importante área de estudo na engenharia elétrica. Os algoritmos adaptativos são desenvolvidos a partir de funções objetivo para se obter características desejadas; entretanto, essas funções normalmente oferecem pouca informação sobre o desempenho do algoritmo. O objetivo desta tese foi apresentar uma forma de apresentar funções-objetivo de algoritmos adaptativos numa forma simples e que possa oferecer alguma informação, seja ela nova ou que realce características conhecidas, através da interpretação do comportamento da própria função-objetivo.

A formulação das funções-objetivo como problemas de otimização oferece a possibilidade de incluir restrições nos problemas. Essas restrições possibilitam, de forma explícita, inserir características desejadas nos algoritmos, como por exemplo a restrição, explorada nesta tese, de manter a matriz de autocorrelação positiva definida. O método apresentado, nesta tese, para o desenvolvimento de algoritmos adaptativos utilizou a descrição da função-objetivo como o mínimo da norma da diferença entre os vetores de coeficientes nos instantes n e $n - 1$. A função-objetivo leva ainda em consideração um elemento relativo ao erro medido na saída do algoritmo (sinal de referência menos sinal de saída do filtro adaptativo), para restringir o comportamento do algoritmo adaptativo. O uso do distúrbio mínimo dos coeficientes é um bom parâmetro para a função objetivo, uma vez que variações bruscas nos coeficientes podem levar o algoritmo a não seguir o sentido descendente da norma da diferença dos coeficientes e, conseqüentemente, fazer o algoritmo divergir.

No Capítulo 2 avaliou-se o uso do método de barreira como um meio de resol-

ver problemas de otimização convexa, sendo necessário, para tal, o equacionamento da função-objetivo de um algoritmo adaptativo como um problema de otimização convexa. A metodologia abordada mostrou-se pouco útil para o desenvolvimento de algoritmos com restrições, uma vez que as funções-objetivo permaneciam as mesmas embora aparecendo como restrições da função-objetivo da norma do vetor de coeficientes, $\|\mathbf{w}(n)\|$, ponderado pela iniciação da matriz de autocorrelação, $\mathbf{R}(0)$.

Apesar de o problema de otimização formulado pelo método de barreira não oferecer interpretações novas sobre o comportamento dos algoritmos adaptativos, a função-objetivo desenvolvida mostra claramente que a iniciação da matriz de autocorrelação tem influência no desempenho do algoritmo, como foi observado nas simulações do Capítulo 4. A iniciação da matriz de autocorrelação pode acelerar a velocidade de convergência de algoritmos baseados no método recursivo de mínimos quadrados e, ainda, reduzir o excesso de MSE em certos algoritmos com pouca capacidade de esquecimento de informações passadas. Nesse sentido, quanto menor o valor de iniciação da matriz de autocorrelação do sinal de entrada, menor interferência terá sobre o algoritmo no decorrer das iterações.

Em seguida, formularam-se os problemas de otimização dos algoritmos adaptativos a partir de uma descrição por distúrbio mínimo. Essa metodologia apresenta novas possibilidades para o desenvolvimento de novos algoritmos e a reformulação de algoritmos clássicos. Os problemas de minimização dos algoritmos RLS e QN foram formulados como problemas de otimização descritos na forma de distúrbio mínimo, levando em consideração também uma função do erro *a posteriori*.

Dentre as vantagens obtidas pelo uso dessa metodologia, pode-se citar a obtenção de uma função-objetivo simples, com possibilidade de inclusão de restrições para melhorar o desempenho dos algoritmos e que pode ser utilizada para descrever a função-objetivo de diversos algoritmos baseados no método de mínimos quadrados. Um grande destaque foi o algoritmo quasi-Newton, que através da descrição por distúrbio mínimo apresenta uma função-objetivo mais atraente e com fácil extração de informação sobre o comportamento do algoritmo. Entretanto, talvez o aspecto mais importante da metodologia de descrição por distúrbio mínimo apresentada é oferecer uma função-objetivo alternativa para algoritmos baseados em mínimos quadrados que não inclui n versões do vetor de coeficientes. Apesar de esse detalhe

poder passar despercebido, a função-objetivo determinística alternativa permite a otimização em relação ao passo dos algoritmos, $\mu(n)$, sem a dificuldade imposta pela dependência implícita que o passo tem das n versões de $\mathbf{w}(n)$.

O método de descrição por distúrbio mínimo dos coeficientes permite a fácil criação e derivação de novos algoritmos. No Capítulo 3, desenvolveram-se alguns deles para demonstrar a versatilidade dessa ferramenta. Cada algoritmo possui características próprias, que podem ser melhoradas com a inclusão de restrições apropriadas. A solução dos problemas de otimização é relativamente simples, uma vez que as funções-objetivo dos algoritmos são simplificadas pela descrição por distúrbio mínimo, que é uma função convexa: utiliza-se apenas do fato de que igualando-se o primeiro gradiente da função-objetivo a zero pode-se obter o ponto de mínimo.

As restrições de igualdade utilizadas nas descrições das funções-objetivo dos algoritmos podem ser resolvidas através dos multiplicadores de Lagrange, enquanto que o método de barreira é uma ótima ferramenta para a solução de problemas de otimização com restrições de desigualdade. Os multiplicadores de Lagrange foram utilizados no desenvolvimento dos algoritmos 3 e QN, pela descrição de distúrbio mínimo, enquanto que o método de barreira foi utilizado para solucionar com simplicidade o problema do passo ótimo do algoritmo 6. Como discutido no Capítulo 4, a relação entre a precisão das barreiras utilizadas e os limites impostos às restrições, tanto superior quanto inferior, influencia no comportamento do algoritmo 6, tornando o ajuste fino no desempenho do algoritmo uma tarefa um pouco mais complicada para o usuário. Apesar disso, o algoritmo 6 foi o que se mostrou mais robusto em relação aos cenários sem variação abrupta dos coeficientes em que foram testados os algoritmos. Os algoritmos desenvolvidos possuem complexidade computacional igual a complexidade do algoritmo RLS, exceto o algoritmo 3 (NLMS) que tem complexidade computacional inferior, $\mathcal{O}(N)$, onde N é a ordem do filtro adaptativo. A complexidade computacional dos algoritmos é da ordem $\mathcal{O}(N^2)$, sendo que os algoritmos que tem atualização do passo possuem um pequeno acréscimo em sua complexidade da ordem $\mathcal{O}(N)$.

No Capítulo 4 constata-se o desempenho dos algoritmos frente a diversas situações que permitem observar as qualidades e defeitos dos algoritmos, e quando eles podem ser melhor aplicados. Verifica-se a influência da iniciação da matriz

de autocorrelação, a qual concede melhor velocidade de convergência sem, entretanto, aumentar o desajuste do algoritmo, com o uso de um valor pequeno em sua iniciação. Esse fato deve ser analisado com mais cuidado, entretanto um possível motivo para esse fenômeno é o esquecimento normal que a matriz de autocorrelação sofre em relação aos dados adicionados no decorrer das iterações, não interferindo no desajuste. Em contrapartida, o valor inicial baixo na matriz de autocorrelação equivale a um valor elevado em sua inversa, provocando um valor grande no passo do algoritmo nas iterações iniciais. Com o conhecimento desse fato, as simulações dos algoritmos foram realizadas com um fator multiplicativo na inversa da matriz de autocorrelação, aumentando seu valor inicial, em todos os algoritmos recursivos, e conseqüentemente acelerando suas convergências.

Ao final das simulações do Capítulo 4, observou-se que o algoritmo 6 apresentou melhor desempenho geral, considerando a velocidade de convergência, o mínimo desajuste do erro quadrático médio e o fato de impedir que a matriz de autocorrelação aumente indefinidamente. Já nos casos em que é necessária a readaptação dos coeficientes do filtro adaptativo, o algoritmo 2 aparece como uma boa opção. Foram realizadas simulações comparativas dos algoritmos 1 e 6 com dois algoritmos-chave da literatura que foram abordados durante a tese, algoritmos RLS e quasi-Newton, para os quatro tipos de sinais de entrada utilizados para analisar os algoritmos. Os algoritmos 1 e 6 foram escolhidos para as comparações devido ao seu melhor desempenho, sendo que ambos possuem funções-objetivo similares, diferenciando-se apenas quanto ao passo do algoritmo, já que o algoritmo 1 utiliza passo fixo e o algoritmo 6 tem seu passo variável e otimizado de acordo com as restrições impostas no algoritmo.

Observam-se algumas vantagens dos algoritmos 1 e 6 em relação ao algoritmo RLS, e também sobre o algoritmo QN, ressaltadas pelo fato deste possuir um desajuste de MSE de 3 dB devido à normalização do erro *a posteriori*. Desse modo, pode-se indicar o uso do algoritmo 6, assim como apresentado nessa tese, em aplicações onde não há modificação no sistema a ser identificado. Outro fator importante é que a normalização do algoritmo QN impede a derivação de algoritmos rápidos; entretanto, o algoritmo 6 tem suas restrições relaxadas, oferecendo a possibilidade de desenvolver algoritmos rápidos baseados em sua função-objetivo.

Nos casos de variação abrupta dos coeficientes do sistema desconhecido, o algoritmo 2 surge como uma ótima opção ao algoritmo RLS, que teve o melhor desempenho entre os algoritmos. A relação convexa do algoritmo 2, que possibilita ao algoritmo ter características parecidas com o algoritmo RLS e o algoritmo 1, deve ser melhor explorada, podendo, inclusive, utilizar um passo variável que possibilite uma velocidade de convergência similar à do algoritmo RLS ao mesmo tempo que reduz o valor do erro médio quadrático após a convergência. Já o algoritmo 4 é uma opção para filtros adaptativos com passo variável, sem necessitar de nenhum ajuste por parte do usuário, possuindo um desajuste fixo como o do algoritmo QN, porém menor.

Esta tese oferece várias contribuições que apontam opções para futuros trabalhos, seja na pesquisa de algoritmos rápidos para os algoritmos nela desenvolvidos, seja para gerar novos algoritmos. A inclusão de restrições nos algoritmos pode melhorar seus desempenhos em aplicações específicas, como por exemplo em sistemas adaptativos para antenas inteligentes, ou até mesmo estudar alternativas sobre como utilizar o algoritmo 6 quando necessário a readaptação dos coeficientes. Outra opção pode ser o estudo da convergência através de modelos matemáticos, buscando-se uma forma generalizada para algoritmos desenvolvidos a partir da descrição por distúrbio mínimo.

A formulação de outros algoritmos adaptativos da literatura através da descrição por distúrbio mínimo oferece um bom foco de pesquisa; pode-se, por exemplo, tentar encontrar uma função-objetivo simplificada para algoritmos como o apresentado em [11], de modo a reduzir o esforço computacional necessário à inversão da matriz Hessiana ali utilizada. Por fim, o uso da barreira logarítmica pode ser melhor explorado no desenvolvimento de algoritmos, ainda mais se for possível encontrar formas fechadas simples da Hessiana que possibilitem o uso do lema da inversão de matrizes para redução do esforço computacional.

Referências Bibliográficas

- [1] HAYKIN, S., *Adaptive Filter Theory*. 2 ed. Englewood Cliffs, New Jersey, Prentice Hall, 1991.
- [2] DINIZ, P. S. R., *Adaptive Filtering: Algorithms and Practical Implementation*. Norwell, Massachusetts, Kluwer Academic Publishers, 2002.
- [3] BOYD, S., VANDENBERGHE, L., *Convex Optimization*. New York, New York, Cambridge University Press, 2004.
- [4] de CAMPOS, M. L. R., ANTONIOU, A., “A new quasi-Newton adaptive filtering algorithm”. In: *IEEE Transactions on Circuits and Systems - II: Analog and Digital Signal Processing*, v. 44, pp. 924–934, November 1997.
- [5] MARSHALL, D. F., JENKINS, W. K., “A fast quasi-Newton adaptive filtering algorithm”. In: *IEEE Transactions of Signal Processing*, v. 40, pp. 1652–1662, July 1992.
- [6] ZOU, Y., CHAN, S.-C., “A robust quasi-Newton adaptive filtering algorithm for impulse noise suppression”. In: *IEEE International Symposium on Circuits and Systems*, v. 2, pp. 677–680, May 2001.
- [7] CASTOLDI, F. T., CAMPOS, M. L. R. D., “A leakage quasi-Newton adaptation algorithm”. In: *XXV Simpósio Brasileiro de Telecomunicações*, Setembro 2007.
- [8] SAYED, A. H., *Fundamentals of Adaptive Filtering*. New York, New York, John Wiley & Sons, 2003.

- [9] NAGARAJ, S., GOLLAMUDI, S., KAPOOR, S., *et al.*, “BEACON: a adaptive set-membership filtering technique with sparse updates”. In: *IEEE Transactions on Signal Processing*, v. 47, pp. 2928–2941, November 1999.
- [10] DINIZ, P. S. R., de CAMPOS, M. L. R., ANTONIOU, A., “Analysis of LMS-Newton adaptive filtering algorithms with variable convergence factor”. In: *IEEE Transactions on Signal Processing*, v. 43, pp. 617–627, 1995.
- [11] AFKHAMIE, K. H., LUO, Z.-Q., WONG, K. M., “Adaptive linear filtering using interior point optimization techniques”. In: *IEEE Transactions on Signal Processing*, v. 48, pp. 1637–1647, June 2000.
- [12] FLETCHER, R., *Practical Methods of Optimization*. 2 ed. New York, New York, Wiley, 1987.

Apêndice A

Pseudo-Código dos Algoritmos Desenvolvidos

A.1 Algoritmo 1

Iniciação

$$\mathbf{R}^{-1}(-1) = \delta \mathbf{I},$$

onde \mathbf{I} é a matriz identidade com dimensão igual ao número de coeficientes do filtro adaptativo e δ é uma variável de iniciação para a matriz de autocorrelação, como discutido no Capítulo 4. E escolha um valor para o passo do algoritmo, μ .

Para $n \geq 0$ faça

$$y(n) = \mathbf{w}(n-1)^T \mathbf{x}(n)$$

$$e(n) = d(n) - y(n)$$

$$\tau(n) = \mathbf{x}(n)^T \mathbf{R}^{-1}(n-1) \mathbf{x}(n)$$

$$\mathbf{R}^{-1}(n) = \mathbf{R}^{-1}(n-1) - \frac{\mathbf{R}^{-1}(n-1) \mathbf{x}(n) \mathbf{x}(n)^T \mathbf{R}^{-1}(n-1)}{\frac{1}{\mu} + \tau(n)}$$

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \frac{e(n) \mathbf{R}^{-1}(n-1) \mathbf{x}(n)}{\frac{1}{\mu} + \tau(n)},$$

onde $\mathbf{w}(n)$ é o vetor de coeficientes do filtro adaptativo, $\mathbf{x}(n)$ é o vetor de sinal de entrada do filtro adaptativo, $y(n)$ é a saída do sistema adaptativo, $e(n)$ é o erro *a priori*, $d(n)$ é o sinal de referência e $\mathbf{R}^{-1}(n)$ é a inversa da matriz de autocorrelação.

O operador $(\cdot)^T$ é o operador de transposição.

A.2 Algoritmo 2

Iniciação

$$\mathbf{R}^{-1}(-1) = \delta \mathbf{I},$$

onde \mathbf{I} é a matriz identidade com dimensão igual ao número de coeficientes do filtro adaptativo e δ é uma variável de iniciação para a matriz de autocorrelação, como discutido no Capítulo 4. E escolha um valor para o passo do algoritmo, μ , entre 0 e 1.

Para $n \geq 0$ faça

$$y(n) = \mathbf{w}(n-1)^T \mathbf{x}(n)$$

$$e(n) = d(n) - y(n)$$

$$\tau(n) = \mathbf{x}(n)^T \mathbf{R}^{-1}(n-1) \mathbf{x}(n)$$

$$\mathbf{R}^{-1}(n) = \frac{1}{1-\mu} \left[\mathbf{R}^{-1}(n-1) - \frac{\mathbf{R}^{-1}(n-1) \mathbf{x}(n) \mathbf{x}(n)^T \mathbf{R}^{-1}(n-1)}{\frac{1-\mu}{\mu} + \tau(n)} \right]$$

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \frac{e(n) \mathbf{R}^{-1}(n-1) \mathbf{x}(n)}{\frac{1-\mu}{\mu} + \tau(n)},$$

onde $\mathbf{w}(n)$ é o vetor de coeficientes do filtro adaptativo, $\mathbf{x}(n)$ é o vetor de sinal de entrada do filtro adaptativo, $y(n)$ é a saída do sistema adaptativo, $e(n)$ é o erro *a priori*, $d(n)$ é o sinal de referência e $\mathbf{R}^{-1}(n)$ é a inversa da matriz de autocorrelação. O operador $(\cdot)^T$ é o operador de transposição.

A.3 Algoritmo 3

Iniciação

Escolha um valor para o passo do algoritmo, μ , entre 0 e 2.

Para $n \geq 0$ faça

$$\begin{aligned} y(n) &= \mathbf{w}(n-1)^T \mathbf{x}(n) \\ e(n) &= d(n) - y(n) \\ \mathbf{w}(n) &= \mathbf{w}(n-1) + \mu \frac{e(n) \mathbf{x}(n)}{\mathbf{x}(n)^T \mathbf{x}(n)}, \end{aligned}$$

onde $\mathbf{w}(n)$ é o vetor de coeficientes do filtro adaptativo, $\mathbf{x}(n)$ é o vetor de sinal de entrada do filtro adaptativo, $y(n)$ é a saída do sistema adaptativo, $e(n)$ é o erro *a priori* e $d(n)$ é o sinal de referência. O operador $(\cdot)^T$ é o operador de transposição.

A.4 Algoritmo 4

Iniciação

$$\mathbf{R}^{-1}(-1) = \delta \mathbf{I},$$

onde \mathbf{I} é a matriz identidade com dimensão igual ao número de coeficientes do filtro adaptativo e δ é uma variável de iniciação para a matriz de autocorrelação, como discutido no Capítulo 4.

Para $n \geq 0$ faça

$$\begin{aligned} y(n) &= \mathbf{w}(n-1)^T \mathbf{x}(n) \\ e(n) &= d(n) - y(n) \\ \tau(n) &= \mathbf{x}(n)^T \mathbf{R}^{-1}(n-1) \mathbf{x}(n) \\ \mu(n) &= \tau(n) \\ \mathbf{R}^{-1}(n) &= \frac{1}{\mu(n)} \left[\mathbf{R}^{-1}(n-1) - \frac{\mathbf{R}^{-1}(n-1) \mathbf{x}(n) \mathbf{x}(n)^T \mathbf{R}^{-1}(n-1)}{\mu(n) + \tau(n)} \right] \\ \mathbf{w}(n) &= \mathbf{w}(n-1) + \frac{e(n) \mathbf{R}^{-1}(n-1) \mathbf{x}(n)}{\mu(n) + \tau(n)}, \end{aligned}$$

onde $\mathbf{w}(n)$ é o vetor de coeficientes do filtro adaptativo, $\mathbf{x}(n)$ é o vetor de sinal de entrada do filtro adaptativo, $y(n)$ é a saída do sistema adaptativo, $e(n)$ é o erro *a priori*, $d(n)$ é o sinal de referência e $\mathbf{R}^{-1}(n)$ é a inversa da matriz de autocorrelação. O operador $(\cdot)^T$ é o operador de transposição.

A.5 Algoritmo 5

Iniciação

$$\mathbf{R}^{-1}(-1) = \delta \mathbf{I},$$

onde \mathbf{I} é a matriz identidade com dimensão igual ao número de coeficientes do filtro adaptativo e δ é uma variável de iniciação para a matriz de autocorrelação, como discutido no Capítulo 4.

Para $n \geq 0$ faça

$$y(n) = \mathbf{w}(n-1)^T \mathbf{x}(n)$$

$$e(n) = d(n) - y(n)$$

$$\tau(n) = \mathbf{x}(n)^T \mathbf{R}^{-1}(n-1) \mathbf{x}(n)$$

$$\mu(n) = \frac{1}{2\tau(n)}$$

$$\mathbf{R}^{-1}(n) = \mathbf{R}^{-1}(n-1) - \frac{1 - \mu(n)}{1 + \mu(n)} \frac{\mathbf{R}^{-1}(n-1) \mathbf{x}(n) \mathbf{x}(n)^T \mathbf{R}^{-1}(n-1)}{\tau(n)}$$

se $\mu(n) \leq 1$, faça

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \frac{e(n) \mathbf{R}^{-1}(n-1) \mathbf{x}(n)}{\mu + \tau(n)}$$

onde $\mathbf{w}(n)$ é o vetor de coeficientes do filtro adaptativo, $\mathbf{x}(n)$ é o vetor de sinal de entrada do filtro adaptativo, $y(n)$ é a saída do sistema adaptativo, $e(n)$ é o erro *a priori*, $d(n)$ é o sinal de referência e $\mathbf{R}^{-1}(n)$ é a inversa da matriz de autocorrelação. O operador $(\cdot)^T$ é o operador de transposição.

A.6 Algoritmo 6

Iniciação

$$\mathbf{R}^{-1}(-1) = \delta \mathbf{I},$$

onde \mathbf{I} é a matriz identidade com dimensão igual ao número de coeficientes do filtro adaptativo e δ é uma variável de iniciação para a matriz de autocorrelação, como discutido no Capítulo 4. Escolha os valores para a precisão das barreiras, t_1 e t_2 , e para os valores das restrições α e β , onde $\alpha \leq \beta$.

Para $n \geq 0$ faça

$$y(n) = \mathbf{w}(n-1)^T \mathbf{x}(n)$$

$$e(n) = d(n) - y(n)$$

$$\tau(n) = \mathbf{x}(n)^T \mathbf{R}^{-1}(n-1) \mathbf{x}(n)$$

$$\mu(n) = \frac{\tau(n)(t_1 + t_2) - (\beta t_2 + \alpha t_1)}{\tau(n)(\beta t_2 + \alpha t_1)}$$

$$\text{se } \mu(n) \leq 0, \mu(n) = 0$$

$$\mathbf{R}^{-1}(n) = \mathbf{R}^{-1}(n-1) - \frac{\mu(n) \mathbf{R}^{-1}(n-1) \mathbf{x}(n) \mathbf{x}(n)^T \mathbf{R}^{-1}(n-1)}{1 + \mu(n) + \tau(n)}$$

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \frac{e(n) \mathbf{R}^{-1}(n-1) \mathbf{x}(n)}{\frac{1}{\mu} + \tau(n)}$$

onde $\mathbf{w}(n)$ é o vetor de coeficientes do filtro adaptativo, $\mathbf{x}(n)$ é o vetor de sinal de entrada do filtro adaptativo, $y(n)$ é a saída do sistema adaptativo, $e(n)$ é o erro *a priori*, $d(n)$ é o sinal de referência e $\mathbf{R}^{-1}(n)$ é a inversa da matriz de autocorrelação. O operador $(\cdot)^T$ é o operador de transposição.