
MÉTODOS ITERATIVOS LINEARES PROJETADOS ATRAVÉS DA TEORIA DE CONTROLE E SUAS APLICAÇÕES

Oumar Diene*
oumar@ufrj.br

Amit Bhaya*
amit@nacad.ufrj.br

*Programa de Engenharia Elétrica, COPPE,
Núcleo de Atendimento em Computação de Alto Desempenho, NACAD,
Universidade Federal do Rio de Janeiro
C.P. 68504, RJ 21945-970, Brasil

RESUMO

É mostrado como projetar uma família de novos métodos iterativos lineares através de uma abordagem pela teoria de controle. Escolhas apropriadas de funções de Liapunov resultam em controladores da família PI/PD com ganhos variantes no tempo. Os novos métodos iterativos são aplicados à determinação dos pesos ótimos de filtros adaptativos. Simulações são feitas para mostrar que o desempenho dos novos algoritmos para filtros adaptativos é comparável ao método RLS.

PALAVRAS-CHAVE: Função de Liapunov, método iterativo, estabilidade, controlador PI/PD, gradiente conjugado, convergência, filtros adaptativos.

ABSTRACT

It is shown how to built a family of new iterative methods using a feedback control viewpoint together with appropriate choices of control Liapunov function, leading to a family of time varying PI/PD controllers. The new iterative methods are used to determine the optimal coefficients of adaptive filters. Simulations are performed in order to show that the new methods have the same performance as the RLS algorithm.

KEYWORDS: Control Liapunov function, iterative meth-

ods, stability, PI/PD controller, conjugate gradient algorithm, convergence, adaptive filters.

1 INTRODUÇÃO

O método do Gradiente Conjugado (CG) é considerado um dos melhores métodos iterativos para a classe de matrizes simétricas e definidas positivas (spd) (Saad e VanDerVorst (2000), VanDerVorst (2000), Saad (1996)). O método CG clássico é um método de direções conjugadas (Hestenes, 1980) que utiliza a ortogonalidade dos resíduos gerados pelo algoritmo para determinar os parâmetros necessários para a convergência. Em aplicações tais como filtragem adaptativa, a atualização das matrizes de correlação e de correlação cruzada a cada amostra acarreta a perda da ortogonalidade dos resíduos do algoritmo de atualização dos pesos do filtro. Esta perda de ortogonalidade resulta em uma perda de convergência do algoritmo e em um aumento do erro médio quadrático da saída do filtro. Neste trabalho o método CG é apresentado e analisado como um sistema dinâmico do tipo planta + controlador em realimentação unitária. Esta abordagem permite determinar os parâmetros do método, independentemente da ortogonalidade do resíduo, através de uma *control Liapunov function* (CLF), bem como projetar novos algoritmos. Esta técnica já foi utilizada em Bhaya e Kaszkurewicz (2004) para a análise/projeto dos métodos tipo Krylov, e consiste em uma escolha adequada de uma *Control Liapunov Function* (CLF). Estes autores mostraram, através de uma CLF, que o algoritmo do gradiente conjugado (CG)

Artigo submetido em 03/10/2005
1a. Revisão em 21/09/2006
Aceito sob recomendação do Editor Associado
Prof. Liu Hsu

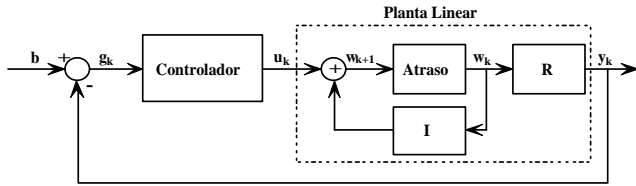


Figura 1: Diagrama de blocos ilustrativo do problema de regulação equivalente à resolução de (1) por métodos iterativos. O problema é: projetar um controlador tal que $\lim_{k \rightarrow \infty} \mathbf{g}_k = \mathbf{0} \Leftrightarrow \lim_{k \rightarrow \infty} \mathbf{R}\mathbf{w}_k = \mathbf{b}$.

pode ser interpretado como um sistema linear dinâmico sujeito a um tipo de controle proporcional e derivativo. Esta função permite determinar os valores ótimos dos parâmetros que tornam o processo assintoticamente estável. Mostra-se através de várias simulações que os métodos determinados desta forma apresentam um desempenho comparável ao do método RLS com um custo e complexidade computacional menor.

2 FORMULAÇÃO DO PROBLEMA

Considere o sistema linear

$$\mathbf{R}\mathbf{w} = \mathbf{b}. \quad (1)$$

Os métodos iterativos de resolução do sistema (1) podem ser descritos pela seguinte equação

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \mathbf{u}_k \quad (2)$$

onde \mathbf{w}_k é uma aproximação da solução de (1) e \mathbf{u}_k é uma correção calculado conforme o algoritmo utilizado, para gerar a cada instante uma aproximação mais precisa \mathbf{w}_{k+1} da solução de (1) (Saad, 1996), e levar o resíduo, definido como $\mathbf{g}_k = \mathbf{b} - \mathbf{R}\mathbf{w}_k$, a zero em um número finito de passos. O problema da resolução do sistema (1) por métodos iterativos pode ser representado por um problema de regulação de sinal, conforme mostra a figura 1. O problema da resolução de (1) por métodos iterativos consiste em projetar um controle \mathbf{u}_k para que o erro de regime seja nulo. Isto é $\lim_{k \rightarrow \infty} \mathbf{g}_k = \mathbf{0} \Leftrightarrow \lim_{k \rightarrow \infty} \mathbf{R}\mathbf{w}_k = \mathbf{b}$. O sistema discreto (3) representa a dinâmica deste problema de regulação:

$$\begin{cases} \mathbf{w}_{k+1} = \mathbf{w}_k + \mathbf{u}_k \\ \mathbf{y}_k = \mathbf{R}\mathbf{w}_k \\ \mathbf{g}_k = \mathbf{b} - \mathbf{y}_k \\ \mathbf{u}_k = f(\mathbf{g}_k) \end{cases} \quad (3)$$

onde $f(\cdot)$ representa o controle adotado.

3 PROJETO DE CONTROLADORES

Na teoria de controle, tanto clássica como moderna, existem várias técnicas de resolução do problema de regulação de sinal (3) equivalente ao problema de resolução do sistema (1) que, na realidade, corespondem a maneiras diferentes de escolher o controlador $f(\cdot)$ em (3). Porém, conforme mostrado em Diene (2004), a maioria das técnicas consagradas na literatura (alocação de pólos, realimentação de estados) resulta em um controlador que usa explicitamente a inversa \mathbf{R}^{-1} da matriz do sistema. Neste trabalho o problema de regulação de sinal é resolvido por controladores do tipo proporcional derivativo e/ou proporcional integral que não utilizam a inversa de \mathbf{R} explicitamente. Para determinar os parâmetros necessários para estabilidade do sistema (equivalente a convergência do algoritmo de resolução de (1)), é usada a mesma técnica aplicada em Bhaya e Kaszkurewicz (2004) para analisar os métodos iterativos do tipo Krylov.

3.1 Controlador proporcional e derivativo (PD)

O sistema (3), representado na forma de diagrama de blocos na figura 1, leva a interpretação de um método iterativo como um sistema de controle realimentado (planta $\{\mathbf{I}, \mathbf{I}, \mathbf{R}, \mathbf{0}\}$ com controlador em configuração de realimentação unitária). O controlador proporcional e derivativo (derivada do estado) é tal que $\mathbf{u}_k = \alpha_k \mathbf{g}_k + \beta_k \Delta \mathbf{w}_k$, sendo que $\Delta \mathbf{w}_k = \mathbf{w}_k - \mathbf{w}_{k-1}$. Portanto, definindo

$$\mathbf{p}_k := \mathbf{g}_k + \gamma_k (\mathbf{w}_k - \mathbf{w}_{k-1}) \quad (4)$$

o sistema (3) resulta em

$$\begin{cases} \mathbf{u}_k = \alpha_k \mathbf{p}_k \\ \mathbf{w}_{k+1} = \mathbf{w}_k + \alpha_k \mathbf{p}_k \\ \mathbf{g}_{k+1} = \mathbf{g}_k - \alpha_k \mathbf{R}\mathbf{p}_k \\ \beta_k = \alpha_k \gamma_k \\ \mathbf{p}_{k+1} = \mathbf{g}_{k+1} + \beta_k \mathbf{p}_k, \end{cases} \quad (5)$$

onde α_k é o ganho proporcional e β_k o ganho derivativo. A figura 2 mostra o diagrama de blocos da planta linear com o controlador discreto proporcional e derivativo em malha fechada.

Note que, considerando os parâmetros α_k e β_k como variáveis de controle e tomando \mathbf{g}_k e \mathbf{p}_k como as variáveis de estado, o sistema a ser controlado é do tipo bilinear. Isto é, o sistema é linear no controle se o estado é mantido constante, e linear no estado se o controle é mantido constante.

$$\begin{cases} \mathbf{g}_{k+1} = \mathbf{g}_k - \alpha_k \mathbf{R}\mathbf{p}_k \\ \mathbf{p}_{k+1} = \mathbf{g}_{k+1} + \beta_k \mathbf{p}_k \end{cases} \quad (6)$$

O objetivo de controle é escolher os escalares α_k e β_k para levar os vetores de estado \mathbf{g}_k e \mathbf{p}_k a zero. Uma vez que o parâmetro α_k não é identicamente nulo, pode-se ver facilmente

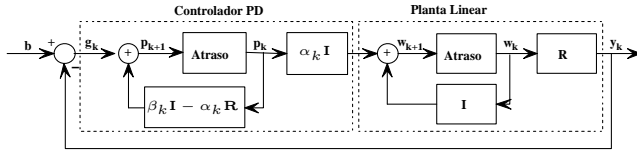


Figura 2: Diagrama de blocos da Planta Linear com um controlador PD. Os parâmetros α_k e β_k do controlador são calculados conforme as equações (11) e (16) para matrizes \mathbf{R} spd.

que o equilíbrio do sistema bilinear (6) se localiza na origem $\mathbf{g}_k = \mathbf{p}_k = \mathbf{0}$ para todo k . A escolha dos parâmetros α_k e β_k é feita em duas etapas. Na primeira, a partir de uma CLF, determina-se o valor de α_k que leva ao decrescimento de alguma norma de \mathbf{g}_k . Na segunda, uma outra função de Liapunov orienta a determinação de β_k que resulta no decrescimento de alguma norma de \mathbf{p}_k . A conclusão é que ambos os vetores \mathbf{g}_k e \mathbf{p}_k convergem para zero, como desejado.

Supondo que a matriz \mathbf{R} é simétrica positiva definida, a CLF escolhida para a determinação de α_k é dada pela equação (7). Note que a escolha da CLF é arbitrária e que cada escolha diferente de CLF resulta em um algoritmo diferente.

$$\mathbf{V}_g(k) = \langle \mathbf{g}_k, \mathbf{R}^{-1} \mathbf{g}_k \rangle. \quad (7)$$

Pela teoria de Liapunov (Slotine e Li (1991) e Khalil (2002)) o equilíbrio de (6) será assintoticamente estável se

$$\begin{aligned} \Delta \mathbf{V}_g &= \mathbf{V}_g(k+1) - \mathbf{V}_g(k) \\ &= \langle \mathbf{g}_{k+1}, \mathbf{R}^{-1} \mathbf{g}_{k+1} \rangle - \langle \mathbf{g}_k, \mathbf{R}^{-1} \mathbf{g}_k \rangle < 0. \end{aligned} \quad (8)$$

Substituindo o resíduo de (6) em (8) e simplificando resulta:

$$\Delta \mathbf{V}_g = -2\alpha_k \langle \mathbf{g}_k, \mathbf{p}_k \rangle + \alpha_k^2 \langle \mathbf{p}_k, \mathbf{R} \mathbf{p}_k \rangle. \quad (9)$$

O valor de α_k que minimiza $\Delta \mathbf{V}_g$ é determinado através de sua derivada fazendo:

$$\frac{\partial \Delta \mathbf{V}_g}{\partial \alpha_k} = -2 \langle \mathbf{g}_k, \mathbf{p}_k \rangle + 2\alpha_k \langle \mathbf{p}_k, \mathbf{R} \mathbf{p}_k \rangle = 0. \quad (10)$$

Portanto,

$$\alpha_k = \frac{\langle \mathbf{g}_k, \mathbf{p}_k \rangle}{\langle \mathbf{p}_k, \mathbf{R} \mathbf{p}_k \rangle}. \quad (11)$$

Substituindo (11) em (9) resulta:

$$\Delta \mathbf{V}_g = -\frac{\langle \mathbf{g}_k, \mathbf{p}_k \rangle^2}{\langle \mathbf{p}_k, \mathbf{R} \mathbf{p}_k \rangle} < 0. \quad (12)$$

Prosseguindo na determinação dos parâmetros estabilizantes do sistema (6), escolha-se a CLF

$$\mathbf{V}_p(k) = \langle \mathbf{p}_k, \mathbf{R} \mathbf{p}_k \rangle, \quad (13)$$

para orientar o cálculo de β_k . Portanto tem-se

$$\begin{aligned} \mathbf{V}_p(k+1) &= \langle \mathbf{p}_{k+1}, \mathbf{R} \mathbf{p}_{k+1} \rangle \\ &= \langle \mathbf{g}_{k+1} + \beta_k \mathbf{p}_k, \mathbf{R} (\mathbf{g}_{k+1} + \beta_k \mathbf{p}_k) \rangle \\ &= \langle \mathbf{g}_{k+1}, \mathbf{R} \mathbf{g}_{k+1} \rangle + 2\beta_k \langle \mathbf{g}_{k+1}, \mathbf{R} \mathbf{p}_k \rangle + \\ &\quad + (\beta_k^2 - 1) \langle \mathbf{p}_k, \mathbf{R} \mathbf{p}_k \rangle + \langle \mathbf{p}_k, \mathbf{R} \mathbf{p}_k \rangle, \end{aligned}$$

o que implica que:

$$\Delta \mathbf{V}_p = \langle \mathbf{g}_{k+1}, \mathbf{R} \mathbf{g}_{k+1} \rangle + 2\beta_k \langle \mathbf{g}_{k+1}, \mathbf{R} \mathbf{p}_k \rangle + (\beta_k^2 - 1) \langle \mathbf{p}_k, \mathbf{R} \mathbf{p}_k \rangle. \quad (14)$$

Seguindo o mesmo raciocínio de função ótima de Liapunov, calcula-se

$$\frac{\partial \Delta \mathbf{V}_p}{\partial \beta_k} = 2 \langle \mathbf{g}_{k+1}, \mathbf{R} \mathbf{p}_k \rangle + 2\beta_k \langle \mathbf{p}_k, \mathbf{R} \mathbf{p}_k \rangle \quad (15)$$

assim,

$$\beta_k = -\frac{\langle \mathbf{g}_{k+1}, \mathbf{R} \mathbf{p}_k \rangle}{\langle \mathbf{p}_k, \mathbf{R} \mathbf{p}_k \rangle} \quad (16)$$

é uma escolha ótima (torna $\partial \Delta \mathbf{V}_p / \partial \beta_k$ o mais negativo possível) que resulta em

$$\|\mathbf{p}_{k+1}\|_{\mathbf{R}}^2 = \|\mathbf{g}_{k+1}\|_{\mathbf{R}}^2 - \frac{\langle \mathbf{g}_{k+1}, \mathbf{R} \mathbf{p}_k \rangle^2}{\langle \mathbf{p}_k, \mathbf{R} \mathbf{p}_k \rangle} < \|\mathbf{g}_{k+1}\|_{\mathbf{R}}^2. \quad (17)$$

De (12) e pela equivalência de normas, pode-se concluir que \mathbf{g}_k decresce em qualquer norma induzida. Portanto, (17) implica que \mathbf{p}_{k+1} decresce na norma induzida pela matriz \mathbf{R} .

Note que o pseudocódigo do método iterativo decorrente do controlador PD, mostrado no algoritmo 1, corresponde ao método do gradiente conjugado (CG). Vale ressaltar que o método utilizado para a derivação dos parâmetros α_k e β_k do método CG, apresentada neste trabalho, permite ter uma noção da robustez do método CG. A análise por uma CLF, ao contrário da análise padrão (por exemplo Saad (1996)), não utiliza nenhuma informação sobre as propriedades do resíduo \mathbf{g}_k ou do vetor \mathbf{p}_k (como a ortogonalidade do resíduo \mathbf{g}_k , isto é $\mathbf{g}_{k+1}^T \mathbf{g}_k = 0$). Isto indica que, em uma implementação de precisão finita, mesmo que as propriedades como a ortogonalidade sejam perdidas, a escolha dos parâmetros α_k e β_k conforme as equações (11) e (16) garante o decrescimento das normas de \mathbf{g}_k e de \mathbf{p}_k ponderadas respectivamente pelas matrizes \mathbf{R}^{-1} e \mathbf{R} . Na presença de ortogonalidade, pode-se mostrar que os parâmetros α_k e β_k dadas pelas equações (11) e (16) são iguais aos parâmetros comumente encontrados na literatura (Hestenes e Stiefel, 1952).

Algoritmo 1

Calcular $\mathbf{g}_0 = \mathbf{b} - \mathbf{R} \mathbf{w}_0$, $\mathbf{p}_0 = \mathbf{g}_0$

Para $k = 0, 1, \dots$ até convergência

$$\alpha_k = \frac{\langle \mathbf{g}_k, \mathbf{p}_k \rangle}{\langle \mathbf{p}_k, \mathbf{R} \mathbf{p}_k \rangle}$$

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \alpha_k \mathbf{p}_k$$

$$\mathbf{g}_{k+1} = \mathbf{g}_k - \alpha_k \mathbf{R} \mathbf{p}_k$$

$$\beta_k = -\frac{\langle \mathbf{g}_{k+1}, \mathbf{R} \mathbf{p}_k \rangle}{\langle \mathbf{p}_k, \mathbf{R} \mathbf{p}_k \rangle}$$

$$\mathbf{p}_{k+1} = \mathbf{g}_{k+1} + \beta_k \mathbf{p}_k$$

Fim

Para garantir a ortogonalidade dos resíduos do método CG padrão, é necessário inicializar os vetores tais que $\mathbf{p}_0 = \mathbf{g}_0$ (Hestenes e Stiefel, 1952). Para ilustrar o problema gerado pela perda de ortogonalidade dos resíduos do método CG padrão, implementa-se o algoritmo 1 e o método CG padrão (veja, por exemplo, Saad (1996)) e compara-se o desempenho dos algoritmos quando $\mathbf{p}_0 = c\mathbf{g}_0$, c uma constante real. Neste exemplo, as matrizes são escolhidas tais que:

$$\mathbf{R} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{w}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Neste caso, note que $\mathbf{g}_0 = \mathbf{b}$, a tabela 1 apresenta o desempenho dos métodos. O critério de parada é dado pela condição $\|\mathbf{g}_k\| \leq 10^{-3}$, e o número máximo de iterações é igual a 50. Note que em todos os casos da tabela 1 ocorre uma perda de ortogonalidade dos resíduos dos algoritmos, exceto quando $\mathbf{p}_0 = \mathbf{g}_0$. Portanto, para a implementação padrão do CG, pode-se ver que, na ausência de ortogonalidade, quanto maior a razão $\|\mathbf{p}_0\|/\|\mathbf{g}_0\|$ menor será a taxa de convergência, e quando $\|\mathbf{p}_0\| \rightarrow 0$, o algoritmo perde a convergência. Em outras palavras, na ausência de ortogonalidade dos resíduos do método CG padrão, a escolha de um passo inicial α_0 resulta em uma convergência lenta do algoritmo enquanto a escolha de α_0 grande provoca a divergência do método. Por outro lado, o algoritmo com parâmetros determinados via CLF apresenta uma robustez no sentido da estabilidade de Liapunov.

3.2 Controlador proporcional e integral (PI)

A forma geral do controlador discreto proporcional e integral é $\mathbf{u}_k = \alpha_k \mathbf{g}_k + \beta_k \sum_{i=0}^k \mathbf{g}_i$, onde o primeiro termo é denominado *termo proporcional* e o segundo, *termo integral*. Isto pode ser reescrito na forma da seguinte equação de estados

$$\begin{cases} \mathbf{q}_{k+1} = \mathbf{q}_k + \mathbf{g}_k \\ \mathbf{u}_k = k_i \mathbf{q}_k \end{cases} \quad (18)$$

Introduzindo este novo estado ‘integrador’ em (3), o sistema resulta em

$$\begin{cases} \mathbf{w}_{k+1} = \mathbf{w}_k + \alpha_k \mathbf{p}_k \\ \mathbf{g}_{k+1} = \mathbf{g}_k - \alpha_k \mathbf{R} \mathbf{p}_k \\ \mathbf{q}_{k+1} = \mathbf{q}_k + \mathbf{g}_k \\ \mathbf{p}_{k+1} = \mathbf{g}_{k+1} + \gamma_k \mathbf{q}_{k+1} \end{cases} \quad (19)$$

onde α_k é o ganho proporcional, \mathbf{q}_k representa o estado do integrador, β_k é um ganho introduzido para que o estado do integrador seja nulo em regime permanente, \mathbf{p}_k é um estado auxiliar e γ_k multiplicado por α_k representa o ganho integral, conforme mostra a figura 3.

A análise do sistema com controlador PI é semelhante à do sistema com controlador PD. O sistema bilinear considerado

Cond. Ini.	Método	Iterações	$\ \mathbf{g}_k\ $
$\mathbf{p}_0 = \mathbf{g}_0$	CG CLF	2	0
	CG	2	0
$\mathbf{p}_0 = 2\mathbf{g}_0$	CG CLF	2	0
	CG	13	5.98×10^{-4}
$\mathbf{p}_0 = 5\mathbf{g}_0$	CG CLF	2	0
	CG	36	9.43×10^{-4}
$\mathbf{p}_0 = 20\mathbf{g}_0$	CG CLF	2	0
	CG	50	1.67×10^{-2}
$\mathbf{p}_0 = 0.5\mathbf{g}_0$	CG CLF	2	0
	CG	50	1.76
$\mathbf{p}_0 = 0.2\mathbf{g}_0$	CG CLF	2	1.57×10^{-16}
	CG	50	1.03×10^{20}

Tabela 1: Comparação do desempenho do método CG utilizando parâmetros determinados via CLF com o método CG assumindo ortogonalidade. A perturbação é gerada na inicialização do vetor \mathbf{p}_0 . Neste caso \mathbf{p}_0 é escolhido aleatoriamente (normalmente $\mathbf{p}_0 = \mathbf{g}_0$). O algoritmo com parâmetros determinados via CLF apresenta uma robustez no sentido da estabilidade de Liapunov enquanto o algoritmo tradicional perde a convergência. O número máximo de iterações é 50 e o critério de parada é $\|\mathbf{g}_k\| \leq 10^{-3}$.

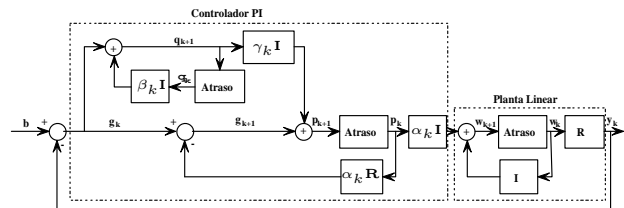


Figura 3: Diagrama de blocos da Planta Linear com um controlador PI. Os parâmetros α_k , β_k e γ_k do controlador são calculados conforme as equações (21), (22) e (23) para \mathbf{R} spd.

neste caso é dado por:

$$\begin{cases} \mathbf{g}_{k+1} = \mathbf{g}_k - \alpha_k \mathbf{R} \mathbf{p}_k \\ \mathbf{q}_{k+1} = \mathbf{q}_k + \mathbf{g}_k \\ \mathbf{p}_{k+1} = \mathbf{g}_{k+1} + \gamma_k \mathbf{q}_{k+1} \end{cases} \quad (20)$$

Utilizando a mesma CLF, $\mathbf{V}_g(k)$ dada pela equação (7), pode-se mostrar facilmente que as escolhas

$$\alpha_k = \frac{\langle \mathbf{g}_k, \mathbf{p}_k \rangle}{\langle \mathbf{p}_k, \mathbf{R} \mathbf{p}_k \rangle} \quad (21)$$

$$\beta_k = -\frac{\langle \mathbf{g}_k, \mathbf{R} \mathbf{q}_k \rangle}{\langle \mathbf{q}_k, \mathbf{R} \mathbf{q}_k \rangle} \quad (22)$$

$$\gamma_k = -\frac{\langle \mathbf{g}_{k+1}, \mathbf{R} \mathbf{q}_{k+1} \rangle}{\langle \mathbf{q}_{k+1}, \mathbf{R} \mathbf{q}_{k+1} \rangle} \quad (23)$$

resultam em

$$\Delta V_{\mathbf{g}} = -\frac{\langle \mathbf{g}_k, \mathbf{p}_k \rangle^2}{\langle \mathbf{p}_k, \mathbf{R}\mathbf{p}_k \rangle} < 0 \quad (24)$$

$$\begin{aligned} \|\mathbf{q}_{k+1}\|_{\mathbf{R}}^2 &= \|\mathbf{g}_k\|_{\mathbf{R}}^2 - \frac{\langle \mathbf{g}_k, \mathbf{R}\mathbf{q}_k \rangle^2}{\langle \mathbf{q}_k, \mathbf{R}\mathbf{q}_k \rangle} \\ &< \|\mathbf{g}_k\|_{\mathbf{R}}^2 \end{aligned} \quad (25)$$

$$\begin{aligned} \|\mathbf{p}_{k+1}\|_{\mathbf{R}}^2 &= \|\mathbf{g}_{k+1}\|_{\mathbf{R}}^2 - \frac{\langle \mathbf{g}_{k+1}, \mathbf{R}\mathbf{q}_{k+1} \rangle^2}{\langle \mathbf{q}_{k+1}, \mathbf{R}\mathbf{q}_{k+1} \rangle} \\ &< \|\mathbf{g}_{k+1}\|_{\mathbf{R}}^2 \end{aligned} \quad (26)$$

respectivamente.

O método de projeto de controlador Proporcional Integral (PI) proposto resulta no algoritmo 2, apresentado na forma de pseudocódigo. Note que os parâmetros α_k , β_k e γ_k do método são também determinados através de CLF's sem utilizar nenhuma propriedade dos vetores \mathbf{g}_k , \mathbf{p}_k e \mathbf{q}_k . Isto sugere que, em aplicações nas quais não se pode garantir a propriedade básica $\mathbf{p}_0 = \mathbf{g}_0$ (que, por sua vez garante a ortogonalidade dos resíduos), seria vantajoso utilizar os métodos variantes CG-CLF e PI-CLF, ao invés de utilizar o CG padrão, que poderia divergir.

Algoritmo 2

Calcular $\mathbf{r}_0 = \mathbf{b} - \mathbf{R}\mathbf{w}_0$, $\mathbf{p}_0 = \mathbf{g}_0$, $\mathbf{q}_0 \neq \mathbf{g}_0$

Para $k = 0, 1, \dots$, até convergência

$$\alpha_k = \frac{\langle \mathbf{g}_k, \mathbf{p}_k \rangle}{\langle \mathbf{p}_k, \mathbf{R}\mathbf{p}_k \rangle}$$

$$\beta_k = -\frac{\langle \mathbf{g}_k, \mathbf{R}\mathbf{q}_k \rangle}{\langle \mathbf{q}_k, \mathbf{R}\mathbf{q}_k \rangle}$$

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \alpha_k \mathbf{p}_k$$

$$\mathbf{g}_{k+1} = \mathbf{g}_k - \alpha_k \mathbf{R}\mathbf{p}_k$$

$$\mathbf{q}_{k+1} = \mathbf{g}_k + \beta_k \mathbf{q}_k$$

$$\gamma_k = -\frac{\langle \mathbf{g}_{k+1}, \mathbf{R}\mathbf{q}_{k+1} \rangle}{\langle \mathbf{q}_{k+1}, \mathbf{R}\mathbf{q}_{k+1} \rangle}$$

$$\mathbf{p}_{k+1} = \mathbf{g}_{k+1} + \gamma_k \mathbf{q}_{k+1}$$

Fim

O algoritmo 2 apresenta um produto matriz/vetor a mais que o algoritmo 1. Porém, o pseudocódigo pode ser reorganizado notando que $\beta_k = \gamma_{k-1}$, portanto $\mathbf{q}_{k+1} = \mathbf{g}_k + \gamma_{k-1} \mathbf{q}_k = \mathbf{p}_k$. Isto implica que $\mathbf{p}_{k+1} = \mathbf{g}_{k+1} + \gamma_k \mathbf{p}_k$ e $\gamma_k = -\frac{\langle \mathbf{g}_{k+1}, \mathbf{R}\mathbf{p}_k \rangle}{\langle \mathbf{p}_k, \mathbf{R}\mathbf{p}_k \rangle}$. Neste caso, o algoritmo 2 corresponde ao método do gradiente conjugado (algoritmo 1), a menos da inicialização. Em geral, os algoritmos 2 e 1 apresentam o mesmo desempenho porém, podem existir casos em que essa diferença na inicialização resulte em uma diferença no desempenho dos dois algoritmos. Vale ressaltar que o método CG pode então ser interpretado como um sistema dinâmico constituído pela planta linear $\{\mathbf{I}, \mathbf{I}, \mathbf{R}, \mathbf{0}\}$ com um controlador proporcional e integral variante no tempo em realimentação unitária.

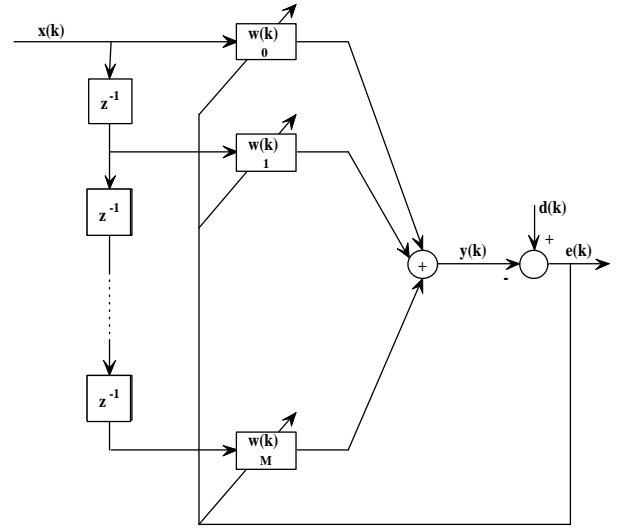


Figura 4: Filtro adaptativo de resposta finita ao impulso.

4 APLICAÇÃO EM FILTRAGEM ADAPTATIVA

O problema de filtragem adaptativa consiste basicamente em determinar o valor atual de um sinal variante no tempo, a partir de valores passados deste mesmo, corrompidos por um ruído (também variante no tempo). A realização mais básica de um filtro adaptativo é obtido através da forma direta de filtro digital de resposta finita ao impulso (FIR), como ilustrado na figura 4, cuja saída é dada por

$$y(k) = \sum_{i=0}^M w_i(k)x(k-i) = \mathbf{w}^T(k)\mathbf{x}(k) \quad (27)$$

onde $\mathbf{x}(k) = [x(k) \ x(k-1) \ \dots \ x(k-M)]^T$ e $\mathbf{w}(k) = [w_0(k) \ w_1(k) \ \dots \ w_M(k)]^T$ são os vetores de entrada e de pesos respectivamente, e M é a dimensão do tap do filtro (número de atrasadores). Uma das funções objetivo mais utilizadas para ajustar os pesos de um filtro adaptativo é o erro médio quadrático definido como

$$\begin{aligned} F(e(k)) &= \xi(k) = E[e^2(k)] \\ &= E[d^2(k) - 2d(k)y(k) + y^2(k)] \end{aligned} \quad (28)$$

onde $e(k) = d(k) - y(k)$, $y(k)$ é a saída do filtro adaptativo e $d(k)$ é o sinal desejado (Haykin, 1991). Neste caso, a função objetivo pode ser reescrita como

$$\begin{aligned} E[e^2(k)] &= \xi(k) \\ &= E[d^2(k) - 2d(k)\mathbf{w}^T(k)\mathbf{x}(k) + \mathbf{w}^T(k)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{w}(k)] \\ &= E[d^2(k)] - 2E[d(k)\mathbf{w}^T(k)\mathbf{x}(k)] + E[\mathbf{w}^T(k)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{w}(k)] \end{aligned} \quad (29)$$

Portanto, para um filtro com coeficientes fixos, a função do erro médio quadrático é dada por

$$\xi = E[d^2(k)] - 2\mathbf{w}^T \mathbf{b} + \mathbf{w}^T \mathbf{R} \mathbf{w} \quad (30)$$

onde $\mathbf{b} = E[d(k)\mathbf{x}(k)]$ é o vetor de correlação cruzada entre a resposta desejada e o sinal de entrada, e $\mathbf{R} = E[\mathbf{x}(k)\mathbf{x}^T(k)]$ é a matriz de autocorrelação do sinal de entrada. Como a função objetivo ξ é uma função quadrática do vetor de pesos \mathbf{w} , o problema de minimizar ξ é equivalente ao problema de achar \mathbf{w} tal que

$$\mathbf{g}_w = \nabla \xi = \frac{\partial \xi}{\partial \mathbf{w}} = -2\mathbf{b} + 2\mathbf{R}\mathbf{w} = \mathbf{0}. \quad (31)$$

Pela equação (31) e notando que $\frac{\partial^2 \xi}{\partial \mathbf{w}^2} = \mathbf{R} > 0$, pode-se dizer que o vetor \mathbf{w}^* tal que $\mathbf{R}\mathbf{w}^* = \mathbf{b}$ é o argumento que minimiza a função objetivo ξ . Em outras palavras, \mathbf{w}^* é solução de

$$\min_{\mathbf{w}} \{\xi = E[d^2(k)] - 2\mathbf{w}^T \mathbf{b} + \mathbf{w}^T \mathbf{R} \mathbf{w}\}. \quad (32)$$

Para resolver o problema da equação (32), métodos iterativos tais como o método CG podem ser utilizados, como mostrado em Boray e Srinath (1992) e Chang e Willson (2000). Em Boray e Srinath (1992) as matrizes \mathbf{R} e \mathbf{b} são estimadas utilizando uma janela deslizante de dados. Neste caso o gradiente da função objetivo é estimado por uma média de n_w valores passados conforme a seguinte equação

$$\mathbf{g}_w(k) = \left(\frac{2}{n_w} \right) \sum_{j=k-n_w+1}^k \{\mathbf{w}^T(j)\mathbf{x}(k) - d(k)\}\mathbf{x}(k). \quad (33)$$

Note que para cada amostra k serão necessários $\min(M, n_w)$ iterações para a convergência do algoritmo CG. Portanto, a escolha de n_w , a janela sobre qual a média do gradiente da função custo é avaliada, tem um impacto direto sobre a convergência dos pesos do filtro. Desta forma, um aumento na dimensão do *tap* e/ou dos dados de entrada causa diretamente um aumento no custo computacional do método, podendo ultrapassar o custo computacional da inversão da matriz \mathbf{R} . No trabalho de Chang e Willson (2000), uma modificação é feita utilizando uma janela de dados exponencialmente decrescente. Esta modificação permite realizar apenas uma iteração a cada amostragem, porém para manter a convergência é necessário a introdução de um fator η , determinado experimentalmente, no cálculo do parâmetro α_k , e do método Polak-Ribiere para o cálculo do parâmetro β_k . Neste trabalho mostra-se através de simulações que, utilizando os métodos desenvolvidos por CLF's, não é necessário a introdução de nenhum fator empírico e/ou de nenhum outro método não linear para garantir a convergência dos métodos em ambientes variantes no tempo.

Utilizando a janela de dados exponencialmente decrescente, proposta por Chang e Willson (2000), as funções de autocorrelação e de correlação cruzada são dadas por

$$\mathbf{R}_{k+1} = \lambda_f \mathbf{R}_k + \mathbf{x}_{k+1} \mathbf{x}_{k+1}^T \quad (34)$$

$$\mathbf{b}_{k+1} = \lambda_f \mathbf{b}_k + d_{k+1} \mathbf{x}_{k+1} \quad (35)$$

onde λ_f é o fator de esquecimento. Para um processamento amostra por amostra, pode-se achar uma formula recursiva do vetor de resíduo utilizando as equações (2), (34) e (35):

$$\begin{aligned} \mathbf{g}_{k+1} &= \mathbf{b}_{k+1} - \mathbf{R}_{k+1} \mathbf{w}_{k+1} \\ &= \lambda_f \mathbf{g}_k - \alpha_k \mathbf{R}_{k+1} \mathbf{p}_k + \\ &\quad + \mathbf{x}_{k+1} (d_{k+1} - \mathbf{x}_{k+1}^T \mathbf{w}_k). \end{aligned} \quad (36)$$

Introduzindo as equações (34), (35) e (36) nos algoritmos 1 e 2, obtêm-se as versões para filtragem adaptativa dos algoritmos CG e PI, apresentadas nos algoritmos 3 e 4.

Algoritmo 3 Método CG-CLF para filtros adaptativos

Calcular $\mathbf{R}_0 = \mathbf{x}_0 \mathbf{x}_0^T$, $\mathbf{b}_0 = d_0 \mathbf{x}_0$

Calcular $\mathbf{g}_0 = \mathbf{b}_0 - \mathbf{R}_0 \mathbf{w}_0$, $\mathbf{p}_0 = \mathbf{g}_0$

Para $k = 0, 1, \dots, N-1$

$$\mathbf{R}_{k+1} = \lambda_f \mathbf{R}_k + \mathbf{x}_{k+1} \mathbf{x}_{k+1}^T$$

$$\alpha_k = \frac{\langle \mathbf{g}_k, \mathbf{p}_k \rangle}{\langle \mathbf{p}_k, \mathbf{R}_{k+1} \mathbf{p}_k \rangle}$$

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \alpha_k \mathbf{p}_k$$

$$\mathbf{g}_{k+1} = \lambda_f \mathbf{g}_k - \alpha_k \mathbf{R}_{k+1} \mathbf{p}_k + \mathbf{x}_{k+1} (d_{k+1} - \mathbf{x}_{k+1}^T \mathbf{w}_k)$$

$$\beta_k = -\frac{\langle \mathbf{g}_{k+1}, \mathbf{R}_{k+1} \mathbf{p}_k \rangle}{\langle \mathbf{p}_k, \mathbf{R}_{k+1} \mathbf{p}_k \rangle}$$

$$\mathbf{p}_{k+1} = \mathbf{g}_{k+1} + \beta_k \mathbf{p}_k$$

Fim

Algoritmo 4 Método PI para filtros adaptativos

Calcular $\mathbf{R}_0 = \mathbf{x}_0 \mathbf{x}_0^T$, $\mathbf{b}_0 = d_0 \mathbf{x}_0$

Calcular $\mathbf{g}_0 = \mathbf{b}_0 - \mathbf{R}_0 \mathbf{w}_0$, $\mathbf{p}_0 = \mathbf{g}_0$, $\mathbf{q}_0 \neq \mathbf{g}_0$

Para $k = 0, 1, \dots, N-1$

$$\beta_k = -\frac{\langle \mathbf{g}_k, \mathbf{R}_k \mathbf{q}_k \rangle}{\langle \mathbf{q}_k, \mathbf{R}_k \mathbf{q}_k \rangle}$$

$$\mathbf{R}_{k+1} = \lambda_f \mathbf{R}_k + \mathbf{x}_{k+1} \mathbf{x}_{k+1}^T$$

$$\alpha_k = \frac{\langle \mathbf{g}_k, \mathbf{p}_k \rangle}{\langle \mathbf{p}_k, \mathbf{R}_{k+1} \mathbf{p}_k \rangle}$$

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \alpha_k \mathbf{p}_k$$

$$\mathbf{g}_{k+1} = \lambda_f \mathbf{g}_k - \alpha_k \mathbf{R}_{k+1} \mathbf{p}_k + \mathbf{x}_{k+1} (d_{k+1} - \mathbf{x}_{k+1}^T \mathbf{w}_k)$$

$$\mathbf{q}_{k+1} = \mathbf{g}_k + \beta_k \mathbf{q}_k$$

$$\gamma_{k+1} = -\frac{\langle \mathbf{g}_{k+1}, \mathbf{R}_{k+1} \mathbf{q}_{k+1} \rangle}{\langle \mathbf{q}_{k+1}, \mathbf{R}_{k+1} \mathbf{q}_{k+1} \rangle}$$

$$\mathbf{p}_{k+1} = \mathbf{g}_{k+1} + \gamma_{k+1} \mathbf{q}_{k+1}$$

Fim

Note que o algoritmo 3 não introduz nenhum custo computacional adicional em relação ao algoritmo de Chang e Willson (2000) enquanto o algoritmo 4 envolve a computação adicional de um produto matriz/vetor.

5 SIMULAÇÕES

Várias simulações são feitas utilizando as três configurações básicas (Sayed, 2003) e (Haykin, 1991): Equalização, Identificação de sistema e Predição linear. Note que, em todos os casos simulados neste artigo, a diferença entre as variações do método CG (algoritmos 1 e 2) não resulta em uma

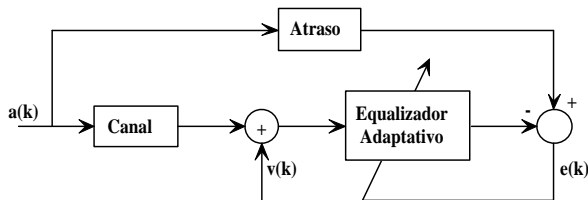


Figura 5: Diagrama de blocos do equalizador adaptativo.

diferença significativa no desempenho dos algoritmos 3 e 4. Portanto, apenas o desempenho do algoritmo 3 será apresentado.

5.1 Equalizador adaptativo

A figura 5 mostra o diagrama de blocos do equalizador adaptativo transversal utilizado para equalizar a distorção introduzida por um canal de banda limitada. O exemplo é tirado de Haykin (1991). A entrada do sistema é uma seqüência aleatória $\{a(k)\}$ com valores ± 1 , portanto de média nula. A seqüência é distorcida por um canal de resposta em frequência dada por

$$\begin{aligned} h(k) &= \frac{1}{2}[1 + \cos(2\pi(k-2)/W)], & k = 1, 2, 3 \\ &= 0, & k > 3 \end{aligned} \quad (37)$$

onde W controla a amplitude da distorção no canal. Um aumento em W causa um aumento da distorção do canal e resulta no aumento do espalhamento ρ dos autovalores da matriz de autocorrelação \mathbf{R} das entradas do equalizador. O canal é corrompido por um vetor de ruído branco aditivo de média nula e variância σ^2 . O equalizador é composto por 11 taps. Como a resposta ao impulso do canal é simétrica em relação a $k = 2$, os pesos dos taps do equalizador são simétricos em relação ao sexto tap. Desta forma, a entrada $\{a(k)\}$ do canal é atrasada por 7 amostras para gerar a resposta desejada do equalizador. A experiência é feita com dois valores de W , 2.9 e 3.5, que resultam em espalhamentos ρ de \mathbf{R} iguais a 6.07 e 46.821 respectivamente. Nas figuras 6, 7, 8, e 9 mostra-se as *curvas de aprendizado*, gráfico do erro instantâneo quadrático $e^2(k) = (d_k - y_k)^2$ (onde d_k é o sinal desejado e y_k é a saída do equalizador) em função da iteração k , dos equalizadores dos algoritmos 3, Polak-Ribiere (Chang e Willson, 2000), CG padrão (Boray e Srinath, 1992) e RLS (Haykin, 1991) respectivamente. Os parâmetros escolhidos neste exemplo são tais que $\lambda_f = \eta = 0.99$ e $\sigma^2 = 0.001$ e o *ensemble* é constituído por 200 amostras independentes. A tabela 2 apresenta os valores de regime para quais converge o erro médio quadrático dos equalizadores em função de W após aproximadamente 40 iterações. Nota-se que os dois métodos propostos neste trabalho apresentam um desempenho

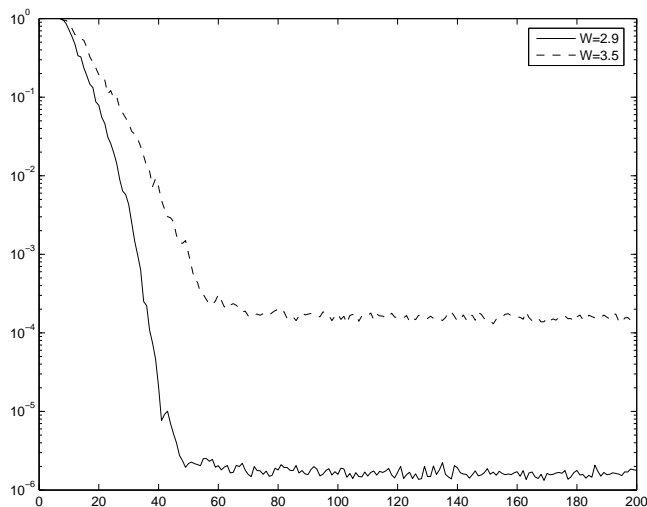


Figura 6: Curva de aprendizado do equalizador adaptativo utilizando o método CG determinado por CLF's. O erro médio quadrático converge para valores de regime iguais a 1.74×10^{-6} e 1.68×10^{-4} para $W = 2.9$ e $W = 3.5$ respectivamente. Note que a taxa de convergência não depende de quanto o sinal de entrada é perturbado.

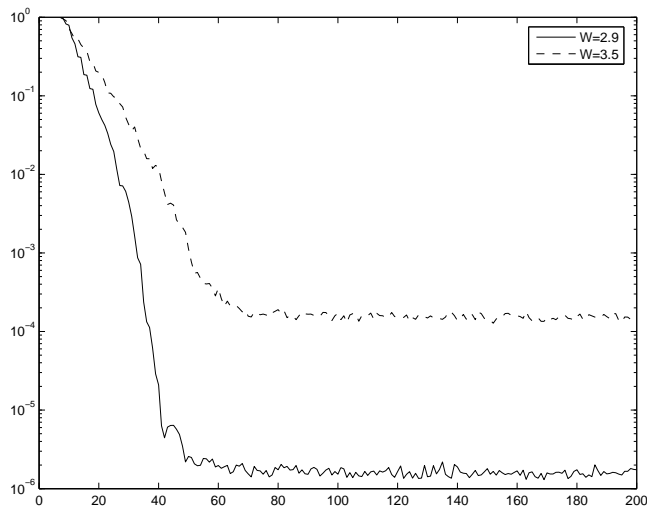


Figura 7: Curva de aprendizado do equalizador adaptativo utilizando o método CG Polak Ribiere de Chang e Willson (2000). O erro médio quadrático converge para valores de regime iguais a 1.69×10^{-6} e 1.63×10^{-4} para $W = 2.9$ e $W = 3.5$ respectivamente. Note que a taxa de convergência não depende da amplitude da perturbação.

comparável ao método de Chang e Willson (2000) e ao método RLS. No caso do método CG padrão, como é realizado apenas uma iteração para cada par $(\mathbf{R}_k, \mathbf{b}_k)$, seria necessário

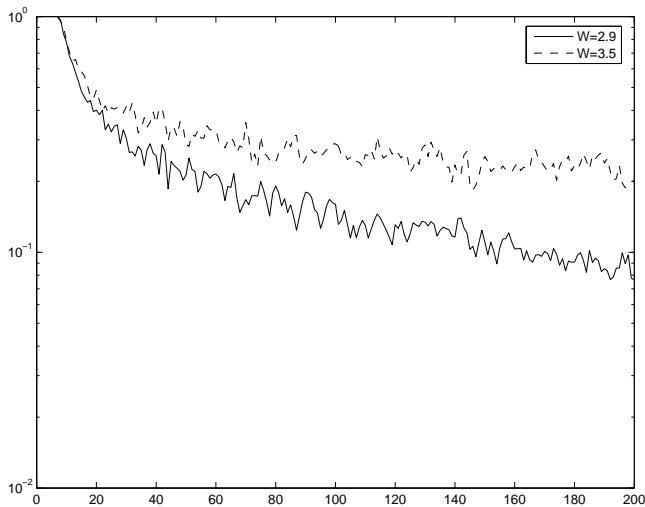


Figura 8: Curva de aprendizado do equalizador adaptativo utilizando o método CG padrão de Boray e Srinath (1992). O erro médio quadrático converge para valores de regime iguais a 1.36×10^{-1} e 2.62×10^{-1} para $W = 2.9$ e $W = 3.5$ respectivamente. Note que a taxa de convergência não depende da amplitude da perturbação.

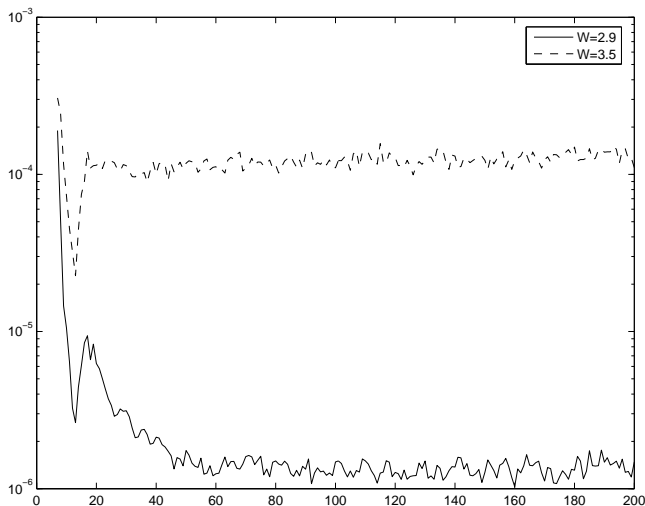


Figura 9: Curva de aprendizado do equalizador adaptativo utilizando o método RLS Haykin (1991). O erro médio quadrático converge para valores de regime iguais a 1.40×10^{-6} e 1.20×10^{-4} para $W = 2.9$ e $W = 3.5$ respectivamente. Note que a taxa de convergência não depende da amplitude da perturbação.

fazer $\mathbf{p}_k = \mathbf{g}_k$ em cada iteração (isto corresponde a reiniciar o método CG tal que $\mathbf{p}_0 = \mathbf{g}_0$ a cada iteração k) para garantir a ortogonalidade dos resíduos gerados. Portanto, no-

Equalizador	$W = 2.9$	$W = 3.5$
CG-CLF	1.74×10^{-6}	1.68×10^{-4}
SD-CLF	1.82×10^{-6}	3.70×10^{-3}
Polak-Ribière	1.69×10^{-6}	1.63×10^{-4}
CG	1.36×10^{-1}	2.62×10^{-1}
RLS	1.40×10^{-6}	1.20×10^{-4}

Tabela 2: Erro médio quadrático dos equalizadores em função de W . Os métodos propostos neste trabalho apresentam um desempenho comparável ao método de Chang e Willson (2000) e ao método RLS.

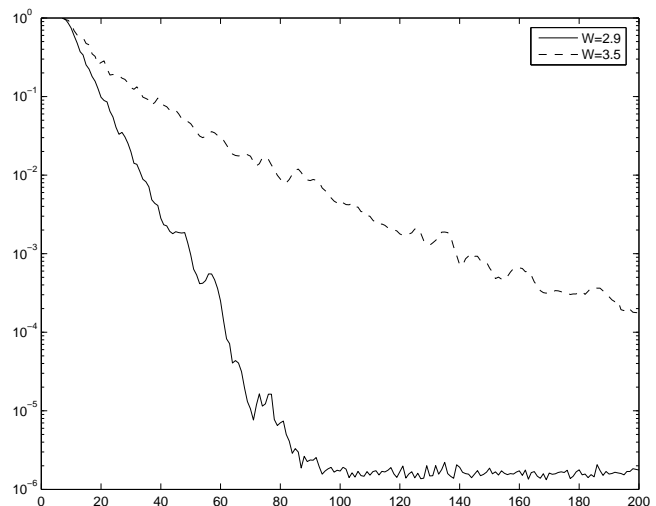


Figura 10: Curva de aprendizado do equalizador adaptativo utilizando o método SD. O erro médio quadrático converge para valores de regime iguais a 1.82×10^{-6} e 3.70×10^{-3} para $W = 2.9$ e $W = 3.5$ respectivamente. Note que a taxa de convergência não depende de quanto o sinal de entrada é perturbado.

tando que $\mathbf{p}_k = \mathbf{g}_k + \beta_{k-1} \mathbf{p}_{k-1}$, pode-se dizer que ocorre uma perda de ortogonalidade dos resíduos ($\mathbf{g}_{k+1}^T \mathbf{g}_k \neq 0$) do método CG padrão (Hestenes e Stiefel, 1952) utilizado em Boray e Srinath (1992), causando um desempenho pior do método.

A partir desta constatação, propõe-se uma outra modificação do método CG no algoritmo 5. A modificação consiste em substituir o ultimo passo do método CG por $\mathbf{p}_{k+1} = \mathbf{g}_{k+1}$, isto equivale a substituir no algoritmo todo o vetor \mathbf{p}_k por \mathbf{g}_k . Vale ressaltar que esta nova modificação do método CG resulta no método *Steepest Descent* (SD). A figura 10 e a tabela 2 apresentam respectivamente a curva de aprendizagem e os valores de regime do método SD. Pode-se notar que o SD-CLF e os outros métodos convergem para o mesmo valor

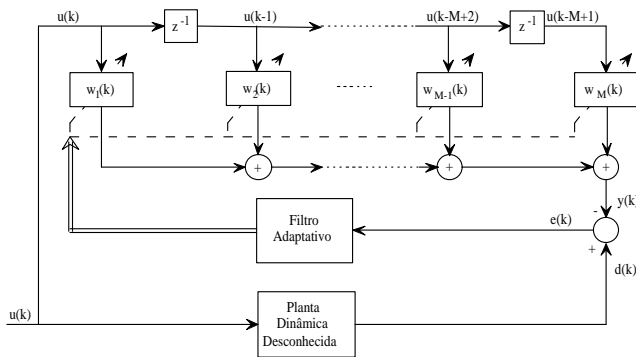


Figura 11: Identificação de uma planta dinâmica desconhecida por filtro adaptativo com *tap* atrasador. Os dados de entrada u e os pesos do *tap* w são reais (Haykin, 1991).

de erro médio quadrático em regime, porém a taxa de convergência do método SD-CLF é menor e é inversamente proporcional ao espalhamento ρ dos autovalores da matriz de autocorrelação (quanto maior ρ , menor será a taxa de convergência). Vale ressaltar que, mesmo que tenha uma taxa de convergência menor que o método RLS em um ambiente variante no tempo, o método SD-CLF pode ser utilizado em casos onde o método LMS (que possui uma taxa de convergência comparável, mas com erro médio quadrático de regime maior e custo computacional maior) é aplicável.

Algoritmo 5 Método SD para filtros adaptativos

Calcular $\mathbf{R}_0 = \mathbf{x}_0 \mathbf{x}_0^T$, $\mathbf{b}_0 = d_0 \mathbf{x}_0$

Calcular $\mathbf{g}_0 = \mathbf{b}_0 - \mathbf{R}_0 \mathbf{w}_0$

Para $k = 0, 1, \dots, N - 1$

$$\alpha_k = \frac{\langle \mathbf{g}_k, \mathbf{g}_k \rangle}{\langle \mathbf{g}_k, \mathbf{R}_{k+1} \mathbf{g}_k \rangle}$$

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \alpha_k \mathbf{g}_k$$

$$\mathbf{R}_{k+1} = \lambda_f \mathbf{R}_k + \mathbf{x}_{k+1} \mathbf{x}_{k+1}^T$$

$$\mathbf{g}_{k+1} = \lambda_f \mathbf{g}_k - \alpha_k \mathbf{R}_{k+1} \mathbf{g}_k + \mathbf{x}_{k+1} (d_{k+1} - \mathbf{x}_{k+1}^T \mathbf{w}_k)$$

Fim

Também note que o algoritmo 5 tem menos passos e um custo computacional menor que os algoritmos de Chang e Willson (2000), Boray e Srinath (1992) e RLS.

5.2 Identificação de Sistema

Na configuração de identificação de sistema considera-se que a planta desconhecida é um filtro de resposta ao impulso finita (FIR) de ordem 20 e o sinal de entrada é um ruído branco Gaussiano de variância $\sigma^2 = 1$. A figura 11 mostra o diagrama de blocos da configuração de identificação de sistema. As figuras 12, 13, 14, 15, e 16 mostram as curvas de aprendizado dos filtros adaptativos utilizando respectiva-

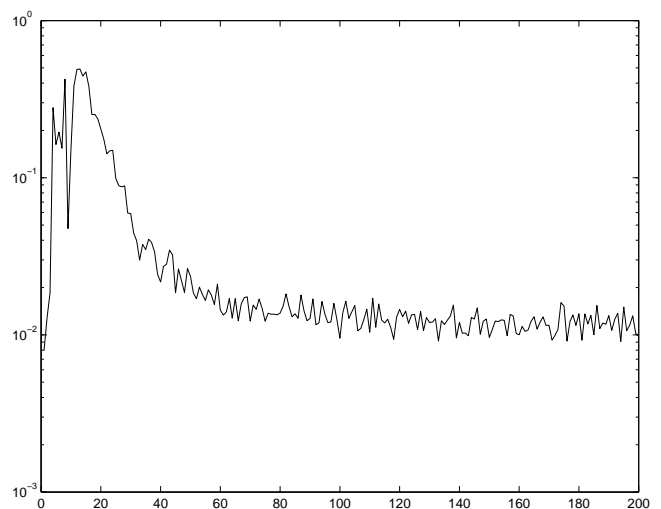


Figura 12: Curva de aprendizado do filtro adaptativo na configuração de identificação de sistema utilizando o método CG determinado por CLF's. O erro médio quadrático converge para um valor de regime igual a 1.35×10^{-2} .

mente os métodos dos algoritmos 3, 5, Polak-Ribiere (Chang e Willson, 2000), CG padrão (Boray e Srinath, 1992), e RLS (Haykin, 1991) na configuração de identificação de sistema. Neste caso o *ensemble* é constituído por 100 amostras independentes. A tabela 3 apresenta os valores de regime para quais converge o erro médio quadrático dos filtros adaptativos após aproximadamente 40 iterações. Como no caso do equalizador adaptativo, nota-se que os métodos propostos neste trabalho e o método Polak-Ribiere de Chang e Willson (2000) apresentam um desempenho comparável ao método RLS. Por outro lado a perda de convergência do método CG padrão de Boray e Srinath (1992), devido à perda de ortogonalidade dos resíduos, é mais acentuada na configuração de identificação de sistema.

Filtro Adaptativo	Erro Médio Quadrático
CG-CLF	1.35×10^{-2}
SD-CLF	1.35×10^{-2}
Polak-Ribière	1.36×10^{-2}
CG	5.05×10^{-1}
RLS	6.90×10^{-3}

Tabela 3: Erro médio quadrático dos filtros adaptativos na configuração de identificação de sistema. Os métodos propostos neste trabalho apresentam um desempenho comparável ao método de Chang e Willson (2000) e ao método RLS.

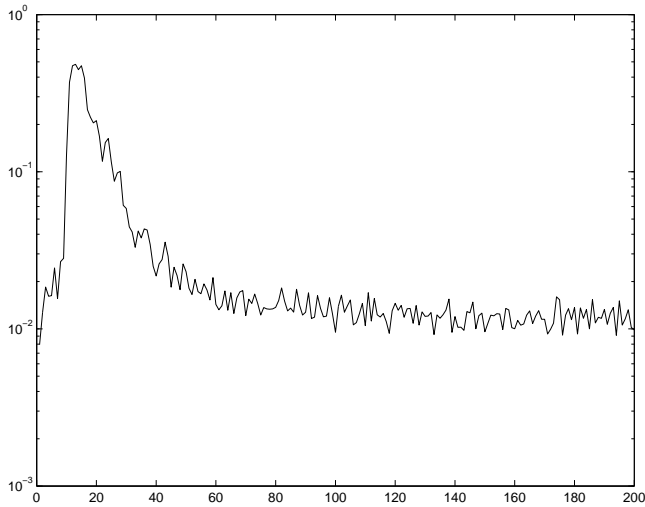


Figura 13: Curva de aprendizado do filtro adaptativo na configuração de identificação de sistema utilizando o método SD determinado por CLF's. O erro médio quadrático converge para um valor de regime igual a 1.35×10^{-2} .

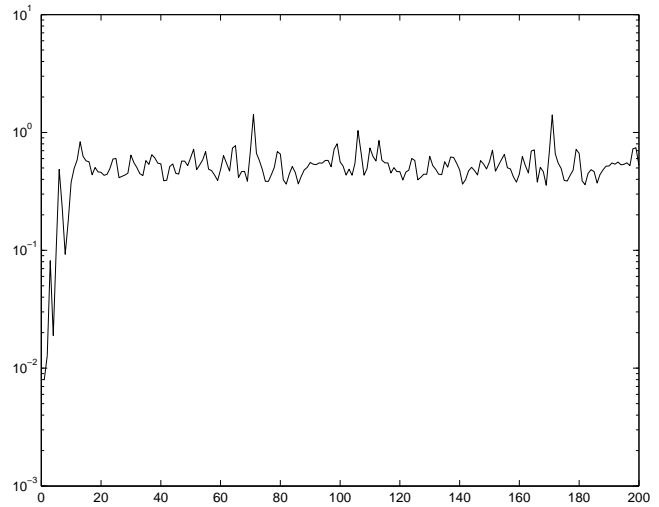


Figura 15: Curva de aprendizado do filtro adaptativo na configuração de identificação de sistema utilizando o método CG padrão de Boray e Srinath (1992). O erro médio quadrático converge para um valor de regime igual a 5.05×10^{-1} .

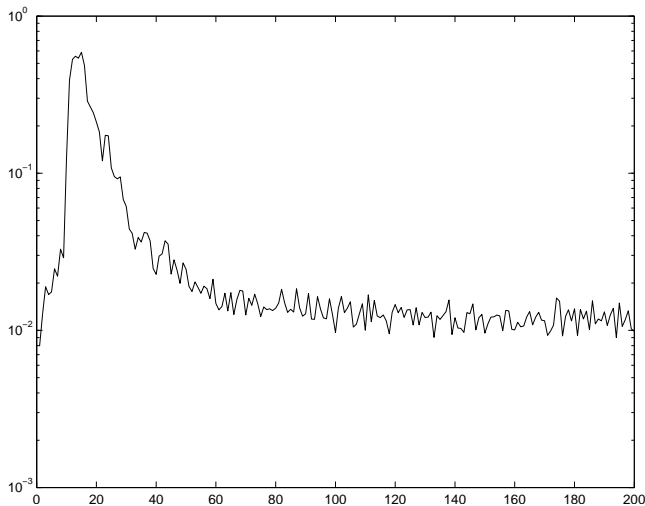


Figura 14: Curva de aprendizado do filtro adaptativo na configuração de identificação de sistema utilizando o método CG Polak Ribiere de Chang e Willson (2000). O erro médio quadrático converge para um valor de regime igual a 1.36×10^{-2} .

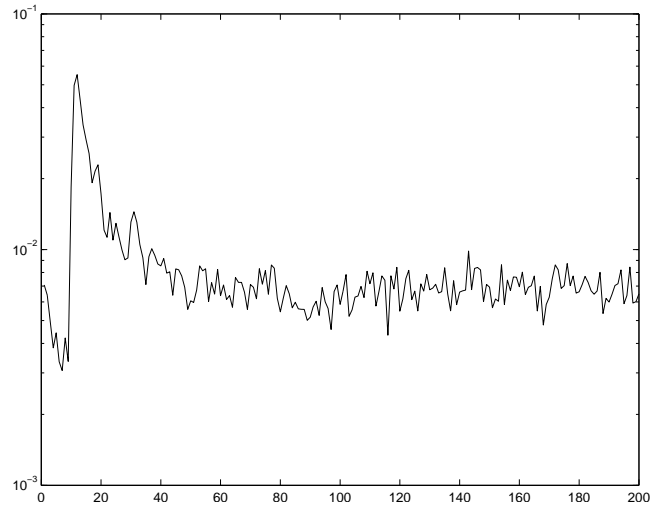


Figura 16: Curva de aprendizado do filtro adaptativo na configuração de identificação de sistema utilizando o método RLS (Haykin, 1991). O erro médio quadrático converge para um valor de regime igual a 6.90×10^{-3} .

5.3 Predição Linear

A terceira configuração de filtro adaptativo simulada neste trabalho é o cenário de predição linear. A figura 17 mostra a configuração do filtro adaptativo de 2 taps no cenário de predição linear. As entradas dos taps $u(k-1)$ e $u(k-2)$ são extraídas do processo autoregressivo (AR) de valores reais

dada pela seguinte equação

$$u(k) + a_1u(k-1) + a_2u(k-2) = v(k) \quad (38)$$

onde $v(k)$ é um ruído branco de média nula e variância σ_v^2 . Os valores de a_1 e a_2 são determinados pelo espalhamento dos autovalores da matriz de autocorrelação dos dados de entrada. Para uma tentativa do experimento, são obtidas 500

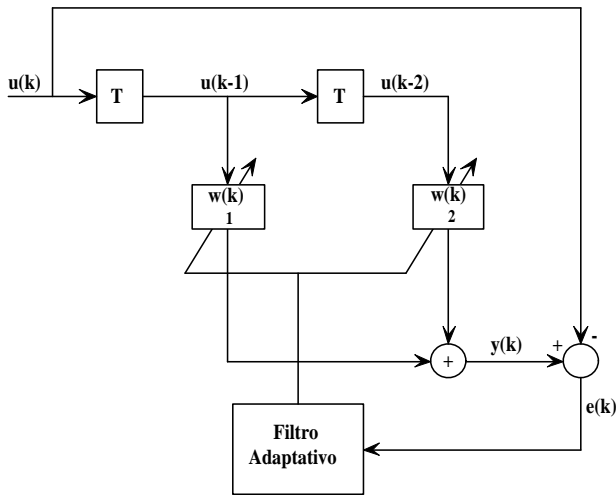


Figura 17: Diagrama de blocos de um filtro adaptativo na configuração de predição linear.

amostras do ruído branco $\{v(i)\}$ a partir de um gerador de números aleatórios de média nula e variância ajustável. O *ensemble* é constituído por 200 tentativas independentes e o experimento é feito para valores de espalhamento $\rho = 3$ e $\rho = 10$. As figuras 18, 19, 20, 21, e 22 apresentam as curvas de aprendizagem dos filtros adaptativos no cenário de predição linear. A tabela 4 apresenta os valores de regime para quais converge o erro quadrático dos preditores lineares. Neste caso também, nota-se que os métodos propostos neste artigo apresentam propriedades de convergência comparáveis aos métodos Polak-Ribière de Chang e Willson (2000) e RLS. Vale ressaltar que a convergência lenta do método SD-CLF, que depende do espalhamento dos autovalores, observado no caso do equalizador adaptativo, não ocorre. Para o método CG padrão, a perda de ortogonalidade resulta em uma perda de convergência para pequenos valores de espalhamento. A tabela 5 compara a porcentagem de desajuste definida como

$$M = \frac{J_{\infty} - J_{min}}{J_{min}}, \quad (39)$$

onde J_{∞} é o valor esperado do erro médio quadrático em regime e J_{min} é o erro médio quadrático mínimo (Boray e Srinath, 1992).

6 CUSTO COMPUTACIONAL

Para determinar o custo computacional dos métodos considerados neste artigo, determina-se o número de operações básicas (adição, subtração, multiplicação e divisão) de ponto flutuante (FLOP) realizadas. A tabela 7 apresenta a comparação do custo computacional dos novos métodos propostos

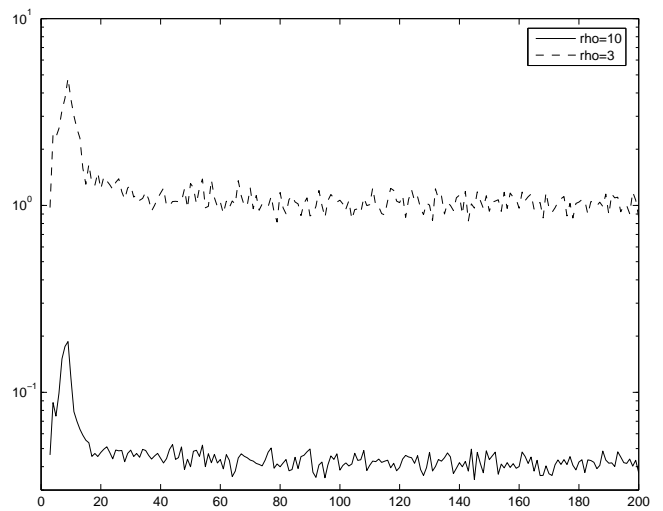


Figura 18: Curva de aprendizado do filtro adaptativo na configuração de predição linear utilizando o método CG determinado por CLF's. O erro médio quadrático converge para valores de regime iguais a 4.27×10^{-2} e 1.02 para $\rho = 10$ e $\rho = 3$ respectivamente.

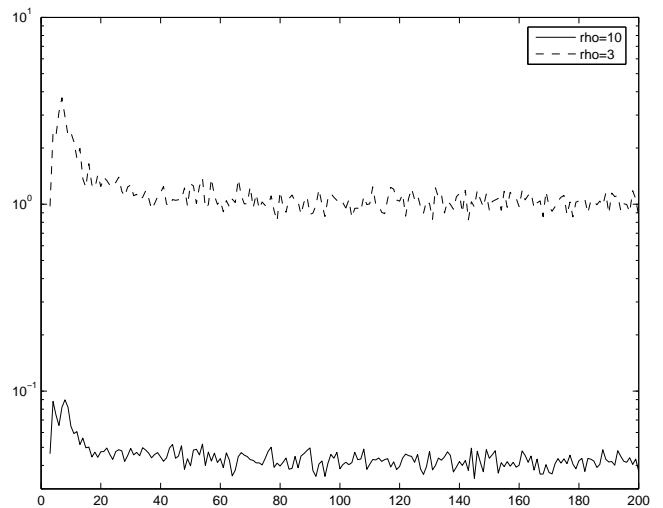


Figura 19: Curva de aprendizado do filtro adaptativo na configuração de predição linear utilizando o método SD determinado por CLF's. O erro médio quadrático converge para valores de regime iguais a 4.27×10^{-2} e 1.02 para $\rho = 10$ e $\rho = 3$ respectivamente.

neste artigo com os métodos RLS (algoritmo 6), CG padrão de Boray e Srinath (1992) e Polak-Ribière de Chang e Willson (2000).

Como já dito anteriormente, o método SD-CLF é o método

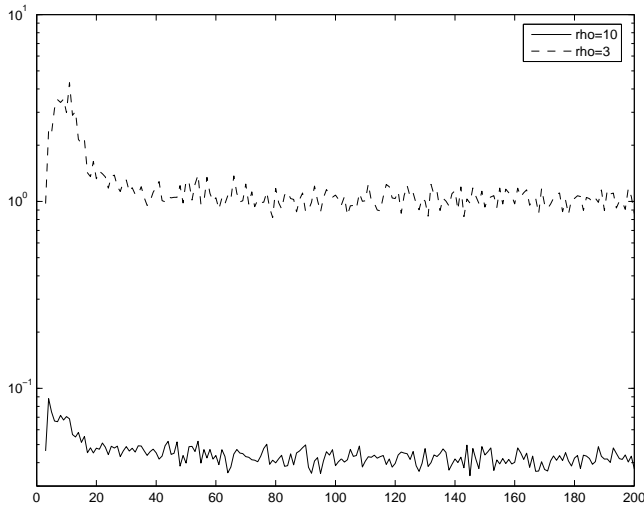


Figura 20: Curva de aprendizado do filtro adaptativo na configuração de predição linear utilizando o método CG Polak Ribiere de Chang e Willson (2000). O erro médio quadrático converge para valores de regime iguais a 4.28×10^{-2} e 1.02 para $\rho = 10$ e $\rho = 3$ respectivamente.

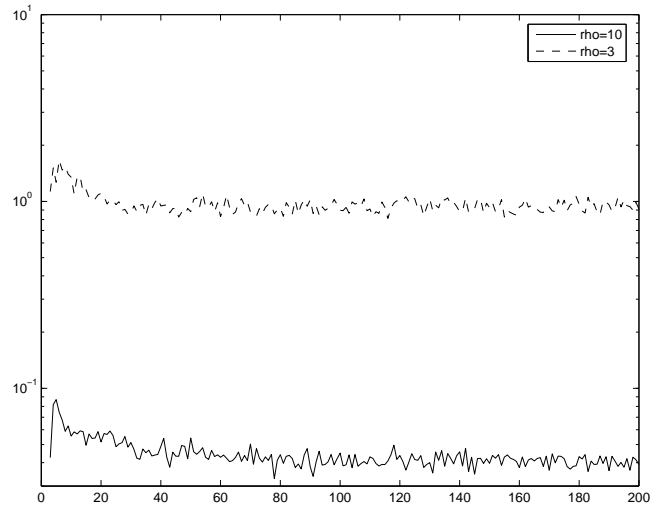


Figura 22: Curva de aprendizado do filtro adaptativo na configuração de predição linear utilizando o método RLS (Haykin, 1991). O erro médio quadrático converge para valores de regime iguais a 4.08×10^{-2} e 0.95 para $\rho = 10$ e $\rho = 3$ respectivamente.

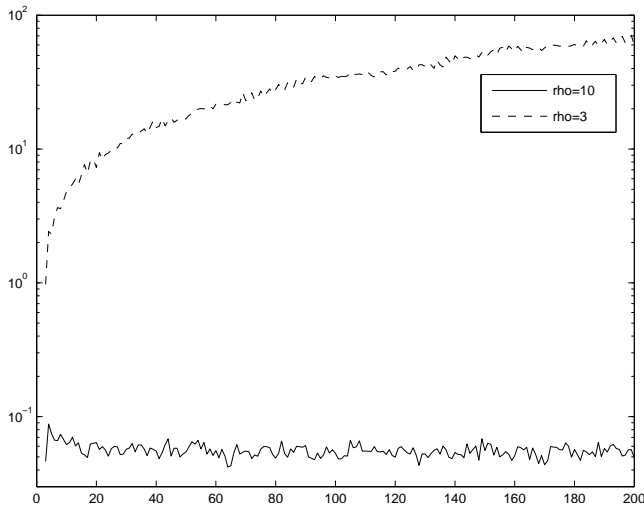


Figura 21: Curva de aprendizado do filtro adaptativo na configuração de predição linear utilizando o método CG padrão de Boray e Srinath (1992). O erro médio quadrático converge para valores de regime iguais a 5.55×10^{-2} e 80.38 para $\rho = 10$ e $\rho = 3$ respectivamente.

Preditor linear	$\rho = 10$	$\rho = 3$
CG-CLF	4.27×10^{-2}	1.02
SD-CLF	4.27×10^{-2}	1.02
Polak-Ribière	4.28×10^{-2}	1.02
CG	5.55×10^{-2}	80.38
RLS	4.08×10^{-2}	0.95

Tabela 4: Erro médio quadrático dos filtros adaptativos na configuração de predição linear. Os métodos propostos neste trabalho apresentam um desempenho comparável ao método de Chang e Willson (2000) e ao método RLS.

ρ	J_{\min}	CG-CLF	SD-CLF	P.R.	CG
3	0.0731	12.95	12.95	12.95	1098.59
10	0.0322	0.33	0.33	0.33	0.72

Tabela 5: Porcentagem de Desajuste $M(\%)$

(2000) e RLS, por realizarem mais operações com vetores e/ou matrizes, apresentam um custo computacional maior que os métodos propostos neste artigo.

que possui menos etapas e portanto tem custo computacional menor. O método CG padrão apresenta um custo computacional menor que o método CG-CLF, com uma diferença correspondente ao custo do cálculo de um produto interno ($2M - 1$). Os métodos Polak-Ribière de Chang e Willson

7 CONCLUSÃO

Neste artigo, os métodos iterativos de resolução de sistemas lineares de equações são analisados do ponto de vista da teoria de controle. O problema da resolução da equação li-

Método	FLOP
CG-CLF	$5M^2 + 17M - 2$
SD-CLF	$5M^2 + 13M - 2$
Polak-Ribière	$5M^2 + 20M - 2$
RLS	$6M^2 + 6M + 2$
CG	$5M^2 + 15M - 1$

Tabela 6: Comparação do custo computacional dos métodos.

Algoritmo 6 Método RLS para filtros adaptativos

$$\mathbf{w}_0 = \mathbf{0}$$

$$\text{Calcular } \mathbf{P}_0 = \delta^{-1} \mathbf{I}$$

Para $k = 0, 1, \dots, N - 1$

$$\mathbf{v}_{k+1} = \frac{\lambda^{-1} \mathbf{P}_k \mathbf{x}_{k+1}}{1 + \lambda^{-1} \mathbf{x}_{k+1}^T \mathbf{P}_k \mathbf{x}_{k+1}}$$

$$\alpha_{k+1} = d_{k+1} - \mathbf{w}_k^T \mathbf{x}_{k+1}$$

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \mathbf{v}_{k+1} \alpha_{k+1}$$

$$\mathbf{P}_{k+1} = \lambda^{-1} \mathbf{P}_k - \lambda^{-1} \mathbf{v}_{k+1} \mathbf{x}_{k+1}^T \mathbf{P}_k$$

Fim

near $\mathbf{R}\mathbf{w} = \mathbf{b}$ é transformado em um problema de regulação de sinal e, o processo iterativo é representado por um sistema dinâmico constituído pela planta linear $\{\mathbf{I}, \mathbf{I}, \mathbf{R}, \mathbf{0}\}$ e um controlador em malha fechada. Esta formulação permite desenvolver métodos iterativos robustos utilizando técnicas de controle do tipo PI/PD com parâmetros variantes no tempo e funções de Liapunov para determinar esses parâmetros. Vale ressaltar que, com a ordenação adequada das etapas e na presença de ortogonalidade dos resíduos, o método PI corresponde ao método CG padrão com uma inicialização diferente.

Estes métodos iterativos são aplicados na determinação dos pesos ótimos de filtros adaptativos em cenários de equalização, identificação de sistemas e predição linear como já feito em Boray e Srinath (1992) e Chang e Willson (2000) com o método CG. Chang e Willson (2000) analisam um sistema linear variante no tempo, entretanto utilizam a formula Polak-Ribière, desenvolvida para funções não lineares gerais, para determinar um dos parâmetros necessários para a convergência. Dentro desta perspectiva, este artigo propõe uma análise que utiliza a estrutura linear do sistema variante no tempo ($\mathbf{R}\mathbf{w} = \mathbf{b}$) e a estrutura bilinear específica dos sistemas que surgem na análise do método CG aplicado a esta situação. Mostra-se através de várias simulações que os métodos propostos neste artigo apresentam um desempenho comparável aos métodos RLS e Chang e Willson (2000). Uma das vantagens dos métodos propostos, em relação aos dois supracitados, é o fato de determinar os pesos ótimos dos filtros sem determinar explicitamente a inversa da matriz de autocorrelação \mathbf{R} e sem utilizar técnicas não lineares e/ou parâmetros determinados empiricamente. A prova sistemática da convergência dos métodos iterativos para sistemas lineares va-

riantes no tempo sob hipóteses adequadas é um tópico sob investigações.

AGRADECIMENTOS

Esta pesquisa foi parcialmente financiada pelos projetos 141566/2004-3, 551863/2002-1 do CNPq, e também contou com o apoio das agências CAPES e FAPERJ.

REFERÊNCIAS

- Bhaya, A. e Kaszkurewicz, E. (2004). Análise e projeto de métodos iterativos vistos como sistemas dinâmicos com controle realimentado, *CBA Gramado: RS*.
- Boray, G. K. e Srinath, M. D. (1992). Conjugate gradient techniques for adaptive filtering, *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications* **39**: 1–10.
- Chang, P. S. e Willson, A. N. (2000). Analysis of conjugate gradient algorithms for adaptive filtering, *IEEE Transactions on Signal Processing* **48**: 409–418.
- Diene, O. (2004). *Métodos iterativos lineares: Análise e projeto através de ferramentas da teoria de controle*, Master's thesis, Universidade Federal do Rio de Janeiro.
- Haykin, S. (1991). *Adaptive Filter Theory*, Prentice Hall, Englewood Cliffs, New Jersey 07632.
- Hestenes, M. R. (1980). *Conjugate Direction Methods in Optimization*, Springer-Verlag, New York.
- Hestenes, M. R. e Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems, *Journal of Research of the National Bureau of Standards* **49**: 409–436.
- Khalil, H. K. (2002). *Nonlinear Systems*, Prentice Hall, Upper Saddle River, New Jersey 07458.
- Saad, Y. (1996). *Iterative methods for sparse linear systems*, PWS Publishing Company, 20 Park Plaza Boston, MA 02116.
- Saad, Y. e VanDerVorst, H. A. (2000). Iterative solution of linear systems in the 20th century, *J. Computational and Applied Mathematics* **122**: 1–33.
- Sayed, A. H. (2003). *Fundamentals of Adaptive Filtering*, John Wiley and Sons, 111 River Street, Hoboken, New Jersey 07030.
- Slotine, J. J. E. e Li, W. (1991). *Applied nonlinear control*, Prentice Hall, Englewood Cliffs, New Jersey.
- VanDerVorst, H. A. (2000). Krylov subspace iteration, *Computing in Science and Engineering* **2**: 32–37.