

Análise de Estabilidade do LMS segundo a Teoria de Lyapunov

Celso de Sousa Júnior, Romis Ribeiro de Faissol Attux, Ricardo Suyama e João Marcos Travassos Romano

Resumo—Este trabalho apresenta uma análise de estabilidade do algoritmo LMS utilizando a abordagem de Lyapunov para sistemas estocásticos e determinísticos. São obtidos limitantes para o passo de adaptação do algoritmo em ambos os casos sem que se recorra necessariamente a hipóteses simplificadoras como a teoria da independência, o que abre perspectivas interessantes em termos de estudos sobre fatores como velocidade de convergência e *misadjustment*.

Palavras-Chave—LMS, Estabilidade no sentido de Lyapunov.

Abstract—This work presents a stability analysis of the LMS algorithm based on deterministic and stochastic Lyapunov approaches. In both cases it is shown to be possible to obtain expressions for the maximum value of the adaptation step-size without necessarily resorting to simplifying hypotheses such as the independence theory. This reveals interesting perspectives of studies concerning aspects like convergence rate and *misadjustment*.

Keywords—LMS, Lyapunov Stability.

I. INTRODUÇÃO

Os primeiros estudos acerca da estabilidade do algoritmo LMS foram iniciados há algumas décadas e, classicamente, conduzidos sob a égide de hipóteses ou estratégias como: teoria da independência [1], simplificação do modelo para sistemas determinísticos e estacionários [2], teoria de Lyapunov com uso de Gramianos de controlabilidade [3] e uso da transformada z [4].

Um ponto crucial em análises desse tipo é buscar critérios para escolha do passo de adaptação, critérios estes que serão tão mais eficientes quanto mais plausíveis forem as hipóteses por eles requeridas. Uma possibilidade interessante nesse sentido parece ser o uso da teoria de Lyapunov para sistemas estocásticos, a qual ainda permanece pouco explorada no âmbito de ferramentas para filtragem adaptativa. O uso dessa teoria pode permitir análises em domínios envolvendo, por exemplo, sistemas não-lineares, variantes no tempo, estocásticos e não-estacionários. Ademais, a análise de Lyapunov permite que um estudo sistemático seja levado a termo sem que se recorra necessariamente a hipóteses simplificadoras usuais como aquelas decorrentes da teoria da independência [1].

C. S. Júnior, R. R. F. Attux, R. Suyama e J. M. T. Romano Laboratório de Processamento de Sinais para Comunicações, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, Brasil, E-mails: celso_de_sousa_junior@yahoo.com.br, attux@dca.fee.unicamp.br, rsuyama@dmo.fee.unicamp.br e romano@dmo.fee.unicamp.br. Este trabalho foi parcialmente financiado pelo CNPq e FAPESP.

A principal contribuição deste trabalho é apresentar uma análise do LMS que aponte limites que devem ser obedecidos pelo passo de adaptação para que o algoritmo seja estável no sentido de Lyapunov, e que seja baseada num conjunto de hipóteses tão parcimonioso quanto possível. O estudo será conduzido em consonância com uma formulação que abrange problemas representativos como equalização de canal e identificação de sistemas.

II. SÍMBOLOS E VARIÁVEIS

$|\cdot|$: valor absoluto;
 $E\{\cdot\}$: operador esperança;
 a : variável escalar;
 \mathbf{a} : vetor ou matriz;
 s_n : sinal transmitido ou sinal de referência;
 \mathbf{w}_n : pesos do filtro transversal no instante n ;
 $\hat{\mathbf{w}}$: estimativa de \mathbf{w} ;
 μ : passo de adaptação para algoritmos LMS;
 $\mathcal{L}^2(\mathbb{R})$: $\int |f(t)|^2 dt < +\infty$: funções de energia finita;
 $\mathcal{L}^p(\mathbb{R})$: espaço com norma $(\int_0^\infty |f(t)|^p dt)^{\frac{1}{p}} < +\infty$;
 \mathbb{N} : conjunto dos números naturais.

III. TEORIA DE LYAPUNOV PARA ESTABILIDADE

Nesta seção, são apresentados alguns conceitos sobre a teoria de Lyapunov para os casos determinístico e estocástico, bem como as hipóteses e restrições consideradas neste trabalho. Uma descrição detalhada dessa teoria pode ser encontrada em referências como [5][6].

A teoria de Lyapunov para análise de estabilidade pode ser considerada à luz de uma interpretação baseada na evolução do uso da energia total de um sistema dinâmico. É possível que o sistema dispenda energia ao longo do tempo e que essa mesma energia não seja repostada [7]. Nesse caso, o estado do sistema pode ser direcionado lentamente para algum ponto ou região em que tenderá a ficar encerrado indefinidamente, o que nos remete, no caso de Lyapunov, à noção de estabilidade.

Se considerarmos que V é uma função genérica de energia do sistema dinâmico, então a explanação acima implica

$$V_n \geq V_{n+1} \geq \dots V_{n+N}.$$

Se for considerado um sistema estocástico, interessante-mente, a abordagem intuitiva permanece, de certa forma, aplicável, desde que sejam consideradas todas as possíveis realizações [7]. Conseqüentemente, temos

$$E\{V_n\} \geq E\{V_{n+1}\} \geq \dots E\{V_{n+N}\}.$$

Teorema 1: Dado um sistema dinâmico **determinístico**, o ponto de equilíbrio \underline{z} é estável no sentido de Lyapunov se existe um função $V(\mathbf{z}_n)$ tal que [8][6]

- 1) Existe uma região fechada e limitada em torno do ponto de equilíbrio \underline{z} ;
- 2) $V(\mathbf{z}_n)$ é uma função contínua e definida positiva, com $V(\underline{z}) = 0$; Adicionalmente, V tem uma função limitante superior e inferior, ou seja, $\alpha(\|\mathbf{z}_n\|^2) \leq V(\mathbf{z}_n) \leq \beta(\|\mathbf{z}_n\|^2)$;
- 3) $\Delta V(\mathbf{z}_n) = V(\mathbf{z}_{n+1}) - V(\mathbf{z}_n) \leq 0$;
- 4) Por outro lado, se $\Delta V(\mathbf{z}_n) = V(\mathbf{z}_{n+1}) - V(\mathbf{z}_n) < 0$ para todo $\mathbf{z}_n \neq 0$, então tem-se estabilidade assintótica.

Para o caso estocástico, seja a tripla $\{\Omega, \mathfrak{R}, P\}$ que caracteriza um espaço de probabilidade, onde Ω é o espaço amostral, \mathfrak{R} é a família de eventos e P é a função probabilística. Considere ainda \mathbf{Z} um espaço de dimensão n de vetores randômicos. $\mathcal{L}_p(\Omega)$ é o conjunto de elementos \mathbf{z} de \mathbf{Z} que satisfazem $E\{\|\mathbf{z}(\omega)\|^p\} < \infty$ [9].

Considere o sistema estocástico

$$\mathbf{z}_{n+1}(\omega) = f_n(\mathbf{z}_n(\omega), \omega), \omega \in \Omega, n \in \mathbb{N}. \quad (1)$$

Definição 1: A solução do sistema descrito pela equação (1) é estável em $\mathcal{L}_p(\Omega)$, $1 < p < \infty$, com respeito a $\mathcal{A} = \{\mathbf{z}, \mathbf{z} \in \mathcal{L}_2(\Omega), \|\mathbf{z}\|_2 \leq H\}$ para todo $\epsilon > 0$, $n_0 \in \mathbb{N}$ se existe $\delta > 0$ tal que $\mathbf{z}_0(\omega) \in \mathcal{A} \cap \mathcal{L}_p(\Omega)$ e $\|\mathbf{z}_0(\omega)\|_p < \delta$ então $\|\mathbf{z}_n(n_0, x_0, \omega)\|_p < \epsilon$ para $n \geq n_0$. [9].

Definição 2: A solução do sistema dado pela equação (1) é uniformemente estável em $\mathcal{L}_p(\Omega)$, $1 < p < \infty$, com relação ao subconjunto \mathcal{A} se a definição 1 é válida e δ não depende de n_0 [9].

Teorema 2: O ponto de equilíbrio de um sistema dinâmico **estocástico** é uniformemente estável em $\mathcal{L}_2(\Omega)$ com respeito a $\mathcal{A} = \{\mathbf{z}, \mathbf{z} \in \mathcal{L}_2(\Omega), \|\mathbf{z}\|_2 \leq H\}$ se [9]:

- 1) A função V é contínua e definida positiva. V tem limitante superior e inferior α e β , sendo ambas funções de classe K^1 . Ou seja, $\alpha(\|\mathbf{z}_n\|^2) \leq V_n(\mathbf{z}_n) \leq \beta(\|\mathbf{z}_n\|^2)$;
- 2) $E\{V_{n+1}(\mathbf{z}_{n+1}(n_0, \mathbf{z}_0, \omega))\} \leq E\{V_n(\mathbf{z}_n(n_0, \mathbf{z}_0, \omega))\}$ para $n_0 \in \mathbb{N}$, $n > n_0$ e $\|\mathbf{z}_n\|_2 \geq H$.

Nota 1: No teorema 2, a desigualdade $\alpha(\|\mathbf{z}_n\|^2) \leq V_n(\mathbf{z}_n)$ corresponde a chamar $V_n(\mathbf{z}_n)$ de função definida positiva [5].

Nota 2: No teorema 2, a desigualdade $V_n(\mathbf{z}_n) \leq \beta(\|\mathbf{z}_n\|^2)$ é relevante, pois garante o comportamento de V em torno do ponto de equilíbrio [5]. Essa é uma exigência adicional nos sistemas não-autônomos ² para provar estabilidade uniforme ³. Caso contrário, a função V fica vulnerável a picos dependentes do tempo onde o

¹Uma função α é dita de classe K quando $\alpha(0) = 0$ e $\alpha(\mathbf{z})$ é crescente para $\mathbf{z} > 0$

²Sistemas autônomos são aqueles que cujo o comportamento é invariante com relação ao tempo [10].

³No caso de sistemas autônomos, o conceito de estabilidade é mais intuitivo, uma vez que necessário apenas V positiva definida e $V_{n+1}(\mathbf{z}_{n+1}) - V_n(\mathbf{z}_n) < 0$. Contudo, o mesmo não se aplica em sistemas não-autônomos.

sistema foi inicializado, o que descaracteriza a propriedade de uniformidade.

Nota 3: Para efeito de desenvolvimento desse trabalho, consideramos o passo de adaptação μ é fixo.

Hipótese 1: Consideraremos sinais reais.

Hipótese 2: Os sinais em análise são estacionários e ergódicos.

A teoria de Lyapunov para sistemas estocásticos pode ser apresentada em outras formas, nas quais, por exemplo, é considerado o comportamento do processo estocástico como um semi-martingale [7] ou mesmo como sistema markoviano [11]. No entanto essas abordagens não abrangem todos os casos de sistemas dinâmicos. Neste trabalho consideramos o tratamento abrangente introduzido por [9].

O sistema estocástico poderia ser aproximado por uma representação (sistema) determinístico com ruído aditivo [10]. Contudo, essa aproximação não cobre os casos gerais, pois nem sempre a característica estocástica de um sistema pode ser simplesmente representada como um termo aditivo.

IV. CRITÉRIO DE WIENER E ALGORITMO LMS

Tomemos como exemplo o problema de equalização de canais de comunicação, ilustrado na figura 1. Em geral, o canal de comunicação distorce o sinal transmitido, levando a degradações na taxa efetiva de transmissão de dados. Uma maneira usual de se combater tais efeitos consiste em empregar um filtro, denominado *equalizador*, a fim de reverter ou minimizar as distorções introduzidas pelo canal.

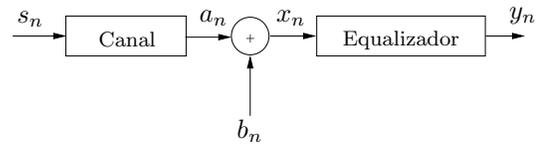


Fig. 1. Modelo para sistema de comunicação.

Idealmente, a saída do equalizador deve ser uma cópia fiel da mensagem transmitida a menos de um possível atraso d , ou seja, $y_n = s_{n-d}$. Portanto, é natural que o ajuste dos parâmetros do filtro seja guiado pela idéia de se obter na saída do equalizador um sinal tão próximo quanto possível de s_{n-d} , idéia que dá origem ao critério de Wiener [1]. Matematicamente, os pesos ótimos do equalizador são tais que minimizam a seguinte função custo

$$J_{Wiener} = E\{(s_{n-d} - y_n)^2\} \quad (2)$$

$$= E\{(e_n)^2\} \quad (3)$$

que quantifica a noção de distância através do erro quadrático médio entre a saída do equalizador e o sinal transmitido.

Considerando que o equalizador possui a estrutura de um filtro linear com resposta ao impulso finita (FIR), sua saída será expressa por $y_n = \mathbf{x}_n^T \mathbf{w}$, onde $\mathbf{w} =$

$[w_0, w_1, \dots, w_L]^T$ denota um vetor contendo a resposta ao impulso do equalizador, e $\mathbf{x} = [x_n, x_{n-1}, \dots, x_{n-L}]^T$ representa o vetor de entrada do filtro, que contém amostras do sinal recebido.

Dessa forma, podemos obter a versão ótima do vetor parâmetros, ou *solução de Wiener*, buscando o ponto no qual o gradiente de (2) se anula. A solução pode ser obtida analiticamente, e é dada por

$$\mathbf{w}_W = \mathbf{R}^{-1}\mathbf{p}, \quad (4)$$

onde $\mathbf{R} = E\{\mathbf{x}_n\mathbf{x}_n^T\}$ denota a matriz de autocorrelação do sinal x_n e $\mathbf{p} = E\{\mathbf{x}_n s_{n-d}\}$ é o vetor de correlação cruzada entre o sinal recebido e o sinal desejado.

O algoritmo LMS foi proposto com a finalidade de se obter iterativamente a solução de Wiener (4). A expressão de atualização dos pesos apresenta a seguinte forma [1]

$$\hat{\mathbf{w}}_{n+1} = \hat{\mathbf{w}}_n + \mu e_n \mathbf{x}_n, \quad (5)$$

Nessa expressão, tem um papel crucial o passo de adaptação, de cuja escolha depende a estabilidade e o desempenho do algoritmo.

Algumas definições serão importantes para a análise da estabilidade do LMS. Primeiramente, definamos o erro entre o vetor ótimo e o vetor de parâmetros estimados, i.e.,

$$\tilde{\mathbf{w}}_n = \hat{\mathbf{w}}_n - \mathbf{w}_W. \quad (6)$$

A partir de (6), podemos reescrever (5) como:

$$\tilde{\mathbf{w}}_{n+1} = \hat{\mathbf{w}}_n + \mu e_n \mathbf{x}_n - \mathbf{w}_W = \tilde{\mathbf{w}}_n + \mu e_n \mathbf{x}_n. \quad (7)$$

Denotando por v_n o erro associado à solução de Wiener, i.e., $\mathbf{x}_n^T \mathbf{w}_W + v_n = s_{n-d}$ podemos reescrever o sinal de erro e_n como

$$\begin{aligned} e_n &= s_{n-d} - \mathbf{x}_n^T \hat{\mathbf{w}}_n = \mathbf{x}_n^T \mathbf{w}_W + v_n - \mathbf{x}_n^T \hat{\mathbf{w}}_n \\ &= v_n - \mathbf{x}_n^T \tilde{\mathbf{w}}_n. \end{aligned} \quad (8)$$

É importante ressaltar que, uma vez que a solução de Wiener equivale ao ponto no qual o gradiente de (2) se anula, verifica-se que

$$\frac{\partial J_{Wiener}}{\partial \mathbf{w}} = -2E\{e_n \mathbf{x}_n\} = 0. \quad (9)$$

Este resultado indica que o sinal de erro associado à solução de Wiener, o qual denotaremos por v_n , e o sinal de entrada do filtro serão necessariamente descorrelacionados entre si. Esta observação será importante para o desenvolvimento da análise de estabilidade.

Note ainda que a representação do erro em (8) se assemelha à forma obtida no problema de identificação de sistemas, onde o erro depende da distância entre o vetor de coeficientes da planta e do modelo ($\tilde{\mathbf{w}}$), e possivelmente de um ruído de medida (v_n) descorrelacionado do sinal de entrada [2]. Conseqüentemente, substituindo (8) em (7), temos

$$\tilde{\mathbf{w}}_{n+1} = \tilde{\mathbf{w}}_n + \mu v_n \mathbf{x}_n - \mu \mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n. \quad (10)$$

V. ANÁLISE DO ALGORITMO LMS UTILIZANDO ABORDAGEM DE LYAPUNOV DETERMINÍSTICA

Seja a candidata a função de Lyapunov ⁴

$$V_n = \tilde{\mathbf{w}}_n^T \tilde{\mathbf{w}}_n, \quad (11)$$

Nota 4: A abordagem determinística é comumente considerada na literatura pois seria uma forma particular do caso estocástico e pode ser considerada como a limitante para o pior caso. Alguns trabalhos que utilizam uma função de Lyapunov determinística mesmo em problemas estocásticos são [12][3][13].

Nota 5: Na construção da equação (11) é considerado que todos os termos da equação (5) são determinísticos. Conseqüentemente, a teoria de Lyapunov para sistemas determinísticos é aplicável.

A diferença temporal da candidata à função de Lyapunov é expressa como (vide teorema 1)

$$\Delta V = V_{n+1} - V_n = \tilde{\mathbf{w}}_{n+1}^T \tilde{\mathbf{w}}_{n+1} - \tilde{\mathbf{w}}_n^T \tilde{\mathbf{w}}_n. \quad (12)$$

Na análise de (12), é preciso considerar a decomposição de $\tilde{\mathbf{w}}_{n+1}^T \tilde{\mathbf{w}}_{n+1}$. Da equação (10), temos

$$\begin{aligned} \tilde{\mathbf{w}}_{n+1}^T \tilde{\mathbf{w}}_{n+1} &= \tilde{\mathbf{w}}_n^T \tilde{\mathbf{w}}_n + \mu \tilde{\mathbf{w}}_n^T \mathbf{x}_n v_n \\ &\quad - \mu \tilde{\mathbf{w}}_n^T \mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n + \mu \mathbf{x}_n^T \tilde{\mathbf{w}}_n v_n \\ &\quad + \mu^2 \mathbf{x}_n^T \mathbf{x}_n v_n^2 - \mu^2 v_n \mathbf{x}_n^T \mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n \\ &\quad - \mu \tilde{\mathbf{w}}_n^T \mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n - \mu^2 v_n \mathbf{x}_n^T \mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n \\ &\quad + \mu^2 (\mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n)^T (\mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n). \end{aligned} \quad (13)$$

considerando a hipótese 1 (i.e $\mathbf{x}_n^T \tilde{\mathbf{w}}_n = \tilde{\mathbf{w}}_n^T \mathbf{x}_n$),

$$\begin{aligned} \tilde{\mathbf{w}}_{n+1}^T \tilde{\mathbf{w}}_{n+1} &= \tilde{\mathbf{w}}_n^T \tilde{\mathbf{w}}_n + 2\mu \tilde{\mathbf{w}}_n^T \mathbf{x}_n v_n \\ &\quad - 2\mu \tilde{\mathbf{w}}_n^T \mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n + \mu^2 \mathbf{x}_n^T \mathbf{x}_n v_n^2 \\ &\quad - 2\mu^2 v_n \mathbf{x}_n^T \mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n \\ &\quad + \mu^2 (\mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n)^T (\mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n). \end{aligned} \quad (14)$$

Portanto a expressão (12) pode ser reescrita como

$$\begin{aligned} \Delta V &= +2\mu \tilde{\mathbf{w}}_n^T \mathbf{x}_n v_n - 2\mu \tilde{\mathbf{w}}_n^T \mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n \\ &\quad + \mu^2 \mathbf{x}_n^T \mathbf{x}_n v_n^2 - 2\mu^2 v_n \mathbf{x}_n^T \mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n \\ &\quad + \mu^2 (\mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n)^T (\mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n). \end{aligned} \quad (15)$$

Para provar estabilidade em torno do ponto de equilíbrio devemos ter $\Delta V \stackrel{?}{\leq} 0$, ou seja

$$\begin{aligned} &\mu [\mathbf{x}_n^T \mathbf{x}_n v_n^2 - 2v_n \mathbf{x}_n^T \mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n \\ &\quad + (\mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n)^T (\mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n)] \stackrel{?}{\leq} \\ &\quad - 2\tilde{\mathbf{w}}_n^T \mathbf{x}_n v_n + 2\tilde{\mathbf{w}}_n^T \mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n. \end{aligned} \quad (16)$$

Conseqüentemente, o passo de adaptação deve ser limitado por

$$\mu < \frac{2\tilde{\mathbf{w}}_n^T \mathbf{x}_n (\mathbf{x}_n^T \tilde{\mathbf{w}}_n - v_n)}{\mathbf{x}_n^T \mathbf{x}_n (\mathbf{x}_n^T \tilde{\mathbf{w}}_n - v_n)^T (\mathbf{x}_n^T \tilde{\mathbf{w}}_n - v_n)} = \theta. \quad (17)$$

⁴Veja que a condição 2 do teorema 1 é obedecida para esta função.

Em casos próximos a condições de equalização/identificação perfeitas, podemos considerar v_n como sendo valor desprezível. Uma outra linha de raciocínio é lembrar v_n é um ruído de medida, o qual não temos controle durante o projeto ou pode ser negligenciado. Nessa caso a expressão (17) para um limitante superior do passo de adaptação fica ainda mais simplificada e adquire um formulação prática

$$\mu < \frac{2}{\text{tr}\{\mathbf{x}_n \mathbf{x}_n^T\}}. \quad (18)$$

Nota-se que essa expressão simplificada é similar à obtida através da análise clássica do LMS [1], o que indica o potencial de que haja elos de grande relevância entre a abordagem de Lyapunov e a metodologia dependente da idéia de convergência no sentido do erro quadrático.

Na figura 2 apresentamos o resultado correspondente à equação (17) no contexto do problema de equalização de um canal composto por dois elementos com um filtro com dois coeficientes. Nesse ensaio, considera-se um canal de fase mínima (com função de transferência $h(z) = 1 + 0,6z^{-1}$) excitado por sinal de entrada binário (+1 e -1) e i.i.d. O atraso de equalização é considerado nulo. Cada ponto da superfície foi obtido através da média de várias realizações do sinal de entrada considerado o mesmo vetor de pesos. Veja que a superfície apresenta uma forma convexa nas proximidade do mínimo de Wiener. Quanto mais próximo ao mínimo, menor deve ser o valor do passo de adaptação para que se dê um passo minimizante no que se refere ao critério de Lyapunov. Note que a obtenção dessa superfície significa que conseguimos obter um guia para escolha do passo de adaptação para o problema abordado, o que era a nossa meta fundamental.

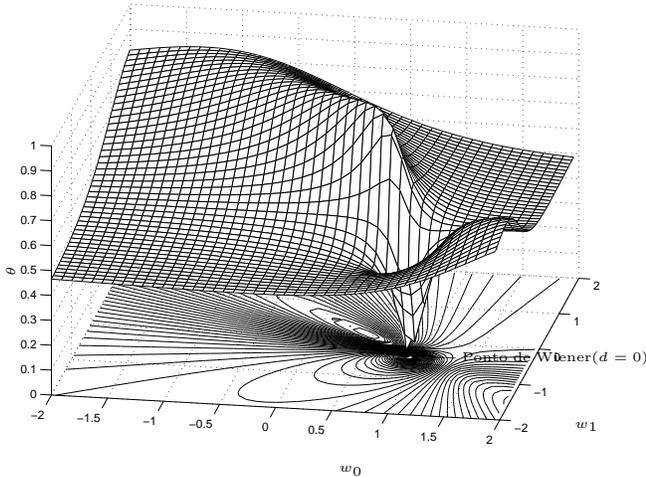


Fig. 2. Superfície que simboliza o passo de adaptação para equação (17).

VI. ANÁLISE DO ALGORITMO LMS UTILIZANDO ABORDAGEM DE LYAPUNOV ESTOCÁSTICA

No caso estocástico, é necessário considerar uma versão da candidata à função de Lyapunov que utiliza o operador

esperança, ou seja

$$V_n = E\{\tilde{\mathbf{w}}_n^T \tilde{\mathbf{w}}_n\}. \quad (19)$$

Conseqüentemente,

$$\Delta V = E\{\tilde{\mathbf{w}}_{n+1}^T \tilde{\mathbf{w}}_{n+1}\} - E\{\tilde{\mathbf{w}}_n^T \tilde{\mathbf{w}}_n\}. \quad (20)$$

Utilizando a expressão (14), obtemos (vide teorema 2)

$$\begin{aligned} \Delta V &= E\{\tilde{\mathbf{w}}_n^T \tilde{\mathbf{w}}_n\} + E\{2\mu \tilde{\mathbf{w}}_n^T \mathbf{x}_n v_n\} \\ &\quad - E\{2\mu \tilde{\mathbf{w}}_n^T \mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n\} + E\{\mu^2 \mathbf{x}_n^T \mathbf{x}_n v_n^T v_n\} \\ &\quad - E\{2\mu^2 v_n \mathbf{x}_n^T \mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n\} \\ &\quad + E\{\mu^2 (\mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n)^T (\mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n)\} \\ &\quad - E\{\tilde{\mathbf{w}}_n^T \tilde{\mathbf{w}}_n\}, \end{aligned} \quad (21)$$

$$\mu \leq \frac{(2E\{\tilde{\mathbf{w}}_n^T \mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n\} - 2E\{\tilde{\mathbf{w}}_n^T \mathbf{x}_n v_n\})}{\left(E\{\mathbf{x}_n^T \mathbf{x}_n v_n^T v_n\} - E\{2v_n \mathbf{x}_n^T \mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n\} + E\{(\mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n)^T (\mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n)\} \right)} = \theta. \quad (22)$$

Para casos que o valor de v_n é desprezível em comparação com os outros termos, a equação (22) pode ser reduzida a

$$\mu \leq \frac{2E\{\tilde{\mathbf{w}}_n^T \mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n\}}{E\{(\mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n)^T (\mathbf{x}_n \mathbf{x}_n^T \tilde{\mathbf{w}}_n)\}} = \theta. \quad (23)$$

A figura 3 apresenta uma exemplo de superfície de nível θ da equação (22). As condições para esta simulação são as mesmas da figura 2.

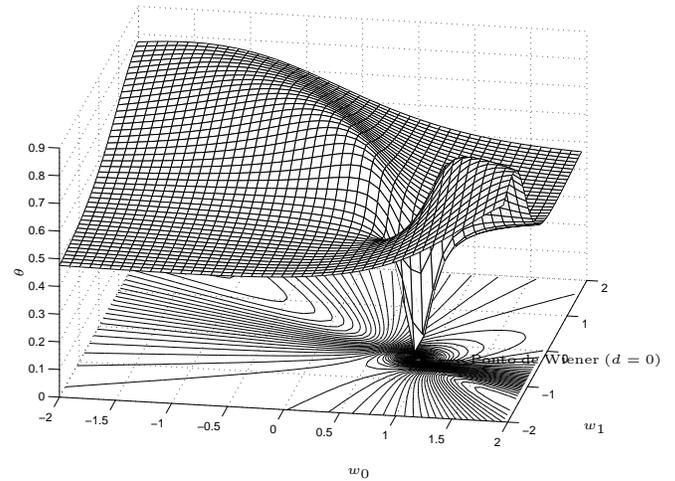


Fig. 3. Superfície que simboliza o passo de adaptação para equação (22).

A superfície referente à equação (23) é representada na figura 4.

É possível perceber que a figura 4 apresenta uma notável semelhança em sua forma geral com a figura 2, o que

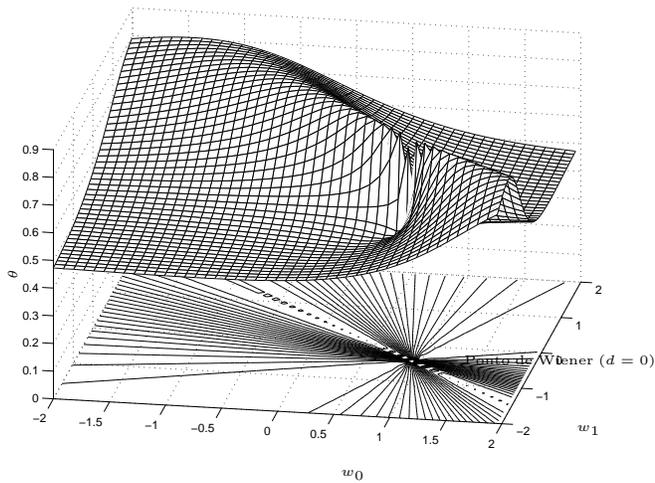


Fig. 4. Superfície que simboliza o passo de adaptação para equação (23).

revela uma conexão entre a abordagem estocástica e a abordagem determinística quando diversas realizações são levadas em conta. Ainda mais interessante que isso, no entanto, é refletir sobre o limitante obtido através da expressão mostrada na figura 4. Antes de mais nada, é preciso ter em mente que a figura foi gerada para um caso de equalização em que a solução de Wiener não era capaz de obter um erro nulo. Isso faz com que seja esperada a existência de flutuações do vetor de pesos em torno da solução de Wiener, o que nos conduz à noção fundamental de *misadjustment* [1]. A existência dessas flutuações, de certo modo, justifica o "pico invertido" que se verifica na vizinhança do ótimo, pois, para, em termos simples, obter uma tendência de confinamento do vetor numa região cada vez menor em torno da solução de Wiener, passa a ser necessário empregar passos de adaptação cada vez mais reduzidos. Quando o vetor é igual ao ótimo, a única forma de mantê-lo nessa posição é zerar o passo para evitar flutuações residuais. Essa linha de raciocínio abre perspectivas interessantes em termos de análise de *misadjustment* vs. passo, idéias estas que pretendemos investigar na seqüência deste primeiro esforço. Para dar um pouco mais de suporte a essas idéias, mostramos, na figura 3, uma curva em que o erro residual da solução de Wiener é considerado nulo: nesse caso, não se verifica a existência do "pico invertido", pois mesmo um passo não-nulo é incapaz de perturbar o equilíbrio associado ao ótimo. Esse seria o caso típico em problemas ideais de identificação, por exemplo.

Outro ponto interessante é que as curvas obtidas parecem estar razoavelmente limitadas por um valor máximo, fato que pode indicar, a partir de uma análise mais rigorosa, a viabilidade de se estabelecerem resultados relacionando valores do passo de adaptação que levam, em algum sentido estocástico, à convergência para o ótimo a partir de regiões iniciais diversas.

VII. CONCLUSÕES

Dois formulações para determinar o valor limite para o passo de adaptação do algoritmo LMS foram apresentadas. Foi possível obter limitantes para o passo de adaptação, tanto na abordagem determinística, quanto na abordagem estocástica, sem recorrer necessariamente a hipóteses simplificadoras como a teoria da independência. Também foram indicadas condições em que parece haver uma relação estreita entre a análise proposta e resultados clássicos acerca da estabilidade do LMS, bem como certas possibilidades que podem levar a metodologias de escolha do passo em termos, por exemplo, do nível de *misadjustment* tolerado. Vale ainda ressaltar que são raros os trabalhos nessa área que utilizam a abordagem de Lyapunov estocástica devido à complexidade desta formulação.

Por fim, indicamos que ainda há muito trabalho a fazer no que se refere à análise e mais detalhada interpretação das expressões obtidas, bem como das conexões que elas parecem indicar com a teoria clássica e com aspectos extremamente relevantes como *misadjustment* e convergência. Também existe a perspectiva de buscar a aplicação do ferramental de Lyapunov na análise de outros algoritmos supervisionados e de algoritmos cegos.

REFERÊNCIAS

- [1] S. Haykin. *Adaptive Filter Theory*. Prentice-Hall, Upper Saddle River, 3rd edition, 1996.
- [2] V. Nascimento e A. Sayed. An unbiased and cost-effective leaky-lms filter. *Asilomar Conference on Signals, Systems and Computers, Proceedings of Asilomar Conference on Signals, Systems and Computers*, 2:1078–1082, Oct 1996.
- [3] V. Solo. The stability of lms. *IEEE Transactions on Signal Processing*, 45 (12):3017–3026, Dec 1997.
- [4] J. C. M. Bermudez e N. N. Bershad. Non-wiener behavior of the filtered lms algorithm. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 46 (8):1110–1113, Aug 1999.
- [5] J.-J. E. Slotine e W. Li. *Applied Nonlinear Control*. Prentice Hall, New Jersey, 1991.
- [6] M. Vidyasagar. *Nonlinear System Analysis*. Prentice-Hall Inc, New Jersey, 2nd edition, 1993.
- [7] M. Aoki. *Optimization of Stochastic Systems*. Academic Press, New York, 1967.
- [8] G. C Goodwin e K. S. Sin. *Adaptive Filtering, Prediction and Control*. Prentice-Hall, Upper Saddle River, 1984.
- [9] T. Morozan. Stability of stochastic discrete systems. *Journal of Mathematical Analysis and Applications*, 23:1–9, 1968.
- [10] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, New Jersey, 2nd edition, 1996.
- [11] R. Z. Has'minskii. *Stochastic Stability of Differential Equations*. Sijthoff & Noordhoff, Maryland, 1980.
- [12] P. I. Hubscher J. C. M. Bermudez e V. H. Nascimento. A mean-square stability analysis of the least mean fourth adaptive algorithm. *IEEE Transactions on Signal Processing*, 55 (8):4018–4028, Aug 2007.
- [13] R. R. Bitmead e B. D. O Anderson. Lyapunov techniques for the exponential stability of linear difference equations with random coefficients. *IEEE Transactions on Automatic Control*, AC-25 (4):782–787, Aug 1980.