

MODELO ANALÍTICO PARA O MOMENTO DE PRIMEIRA ORDEM DO ALGORITMO LEAKY DELAYED LMS

ORLANDO J. TOBIAS E RUI SEARA

LINSE: Circuitos e Processamento de Sinais
Departamento de Engenharia Elétrica
Universidade Federal de Santa Catarina
Campus Universitário, 88040-900 - Florianópolis SC - Brasil
Tel: (0xx48) 331-9504, Fax: (0xx48) 331-9091, orlando@linse.ufsc.br

RESUMO

Este trabalho apresenta um novo modelo para descrever o comportamento médio dos coeficientes do algoritmo *Leaky Delayed LMS* (LDLMS). Esse novo modelo não utiliza a teoria da independência. Nele, novas suposições de análise são consideradas. Assim, é possível se estudar a convergência do algoritmo quando atrasos diferentes estão presentes, o que não poderia ser efetuado através dos atuais modelos existentes. As simulações realizadas mostram que o comportamento teórico do modelo proposto apresenta um casamento muito bom com os resultados obtidos via simulações Monte Carlo.

1. INTRODUÇÃO

O algoritmo *Delayed LMS* (DLMS) pode ser considerado em muitas aplicações práticas [1-3], sendo encontrados, na literatura, muitos trabalhos dedicados a esse assunto [1-7]. Tais trabalhos abordam o problema desses algoritmos sob duas principais condições. A primeira é a de que um atraso D é colocado no caminho do erro e a análise estocástica do algoritmo é realizada invocando a teoria da independência. A segunda é a de que, para garantir a estabilidade do algoritmo, o sinal de referência $\mathbf{X}(n)$ deve ser atrasado no mesmo número de amostras D . Na prática, entretanto, tal condição é muito difícil de ser satisfeita, assim uma condição mais geral deve ser considerada, ou seja, $\mathbf{X}(n - \hat{D})$, sendo $\hat{D} \neq D$. Essa última condição não pode ser estudada à luz da teoria da independência. Recentemente, em [8], foi apresentada uma análise do algoritmo DLMS clássico não considerando tal teoria. Dessa forma, foi possível conceber uma modelagem para atrasos diferentes e foi explicado teoricamente porque, em situações de estimação imperfeita de D , o algoritmo DLMS diverge para sinal de entrada ruído branco e converge para sinal colorido. Neste trabalho, propomos demonstrar que, através da introdução de um fator de perdas, é possível se ter atrasos diferentes, mantendo ainda a convergência do algoritmo

para qualquer tipo de sinal de entrada (ruído branco e colorido).

A nova abordagem para a análise do algoritmo DLMS com perdas é baseada nas seguintes hipóteses:

- i) o atraso D se encontra na malha de adaptação, representando um caso mais geral do que o atraso no caminho do erro (ver Seção 2.4);
- ii) o sinal de referência pode experimentar um atraso \hat{D} , diferente do atraso existente na malha de adaptação;
- iii) diferentemente das análises anteriores, não é invocada a teoria da independência para determinar o valor esperado dos coeficientes do filtro adaptativo. Entretanto, uma nova hipótese de análise é proposta, mostrando-se adequada em relação ao comportamento do modelo e ao desenvolvimento matemático envolvido em sua derivação.

Portanto, na literatura, não se encontra disponível nenhuma análise do algoritmo em questão sob as hipóteses acima mencionadas.

2. ANÁLISE

A Fig. 1 ilustra o diagrama em blocos do algoritmo DLMS. Neste diagrama, o vetor de coeficientes adaptativo é denotado por $\mathbf{W}(n) = [w_0(n), w_1(n), \dots, w_{N-1}(n)]^T$; o sinal de referência e sua versão atrasada são dados por $\mathbf{X}(n) = [x(n), x(n-1), \dots, x(n-N+1)]^T$ e $\mathbf{X}(n - \hat{D})$, respectivamente; D denota o atraso real; \hat{D} , sua estimação. Os sinais: desejado, saída do filtro adaptativo e sinal de erro são denotados por $d(n)$, $y(n)$, e $e(n)$, respectivamente. Também assumimos que $x(n)$ e $d(n)$ são processos com média zero e estacionários em sentido amplo. O ruído de medição $z(n)$ é uma seqüência com média zero, variância σ_z^2 e independente de qualquer outro sinal no sistema.

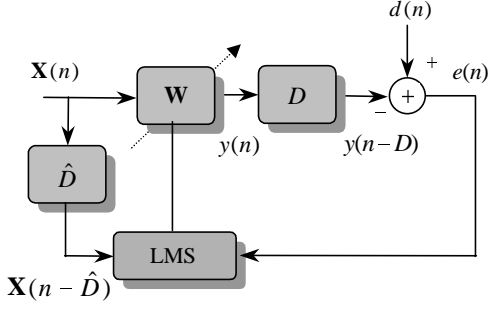


Figura 1. Diagrama em blocos do algoritmo DLMS.

Como hipótese simplificativa, assumimos que as correlações entre os vetores de entrada são mais importantes do que as correlações entre o sinal de entrada e o vetor de coeficientes. Dessa forma, podemos escrever a seguinte aproximação:

$$\begin{aligned} E[\mathbf{X}(n-i)\mathbf{X}^T(n-j)\mathbf{W}(n-k)] &\cong \quad \forall i, j, k \quad (1) \\ E[\mathbf{X}(n-i)\mathbf{X}^T(n-j)]E[\mathbf{W}(n-k)] \end{aligned}$$

Em [9], são apresentados os resultados de exaustivas simulações para o algoritmo LMS filtrado, verificando-se tal aproximação. É importante observar que a hipótese simplificativa apresentada em (1) não corresponde à conhecida clássica teoria da independência, a qual estabelece que o valor esperado $E[\mathbf{X}(n-i)\mathbf{X}^T(n-j)]$ seja zero para $i \neq j$ [10,11].

2.1 Vetor de Coeficientes Ótimos

A função custo que o algoritmo *Leaky* DLMS (LDLMS) minimiza é:

$$J = E[e^2(n)] + \gamma \mathbf{W}^T \mathbf{W}, \quad (2)$$

onde γ representa o fator de perdas. Note que a equação (2) é a função custo clássica penalizada por um termo que é proporcional à norma quadrática de \mathbf{W} [7]. Para um dado vetor \mathbf{W} fixo e assumindo-se $z(n) = 0$, o sinal de erro, obtido da Fig. 1, pode ser expresso por:

$$e(n) = d(n) - y(n-D) = d(n) - \mathbf{X}^T(n-D)\mathbf{W}. \quad (3)$$

A fim de obtermos a função custo (para a determinação dos coeficientes ótimos), calculamos o quadrado de $e(n)$, tomamos o seu valor esperado, substituindo-o em (2). Assim,

$$J = E[d^2(n)] - 2\mathbf{P}_D \mathbf{W} + \mathbf{W}^T \mathbf{R}_0 \mathbf{W} + \gamma \mathbf{W}^T \mathbf{W}, \quad (4)$$

onde $\mathbf{R}_0 = E[\mathbf{X}(n-D)\mathbf{X}^T(n-D)]$ representa a matriz de autocorrelação do vetor de entrada atrasado; o vetor $\mathbf{P}_D = E[d(n)\mathbf{X}^T(n-D)]$ denota a correlação cruzada entre o vetor $d(n)$ e o vetor de entrada atrasado. Finalmente,

minimizando-se (4) em relação ao vetor \mathbf{W} , obtemos a seguinte expressão para os coeficientes ótimos:

$$\mathbf{W}_{ot} = (\mathbf{R}_0 + \gamma \mathbf{I})^{-1} \mathbf{P}_D, \quad (5)$$

com \mathbf{I} denotando a matriz identidade.

2.2 Equação de Atualização dos Coeficientes

De acordo com o algoritmo do gradiente estocástico, a equação de atualização dos coeficientes é obtida pela minimização da versão instantânea da equação (2), resultando na seguinte expressão recursiva:

$$\begin{aligned} \mathbf{W}(n+1) &= \mathbf{W}(n) - \frac{\mu}{2} \nabla \hat{J}(n) \\ &= (1 - \mu \gamma) \mathbf{W}(n) + \mu e(n) \mathbf{X}(n - \hat{D}). \end{aligned} \quad (6)$$

Note que, em (6), é utilizada uma estimativa do gradiente, sendo este uma função de $\mathbf{X}(n - \hat{D})$ ao invés de $\mathbf{X}(n - D)$. Essa última representa o caso ideal, sendo utilizada nas análises anteriormente publicadas [1-7]. Dessa forma, a situação aqui tratada representa um caso mais realista. Assim, substituindo-se $e(n)$, obtido a partir da Fig. 1, em (6), temos:

$$\begin{aligned} \mathbf{W}(n+1) &= v \mathbf{W}(n) + \mu (d(n) \\ &\quad - \mathbf{X}^T(n-D)\mathbf{W}(n-D) + z(n)) \mathbf{X}(n - \hat{D}), \end{aligned} \quad (7)$$

onde $v = 1 - \mu \gamma$.

2.3 Modelo para o Valor Esperado do Vetor de Coeficientes

Tomando-se o valor esperado em ambos os lados da Eq. (7), obtemos:

$$E[\mathbf{W}(n+1)] = v E[\mathbf{W}(n)] + \mu (\mathbf{P}_D - \mathbf{R}_{D-\hat{D}} E[\mathbf{W}(n-D)]), \quad (8)$$

onde $\mathbf{R}_{D-\hat{D}} = E[\mathbf{X}(n - \hat{D})\mathbf{X}^T(n-D)]$ e $\mathbf{P}_D = E[d(n)\mathbf{X}(n - \hat{D})]$. O termo envolvendo $E[z(n)]$ é zero pelas características assumidas para $\{z(n)\}$. É importante notar que a Eq. (8) é obtida a partir de (1) para o caso em que $i = \hat{D}$ e $j = k = D$. Se agora assumimos que o algoritmo converge, o valor em regime permanente para o vetor de coeficientes adaptativo, denotado por \mathbf{W}_∞ , pode ser determinado a partir de (8), fazendo-se $\lim_{n \rightarrow \infty} E[\mathbf{W}(n+1)] = \lim_{n \rightarrow \infty} E[\mathbf{W}(n)] = \mathbf{W}_\infty$. Desse modo, obtemos a seguinte expressão:

$$\mathbf{W}_\infty = (\mathbf{R}_{D-\hat{D}} + \gamma \mathbf{I})^{-1} \mathbf{P}_D. \quad (9)$$

Observe que (9) se reduz à expressão (5) quando $\hat{D} = D$. Por outro lado, é muito importante salientar que a Eq. (9)

não poderia ter sido obtida utilizando-se a hipótese clássica de independência. Assim, se considerarmos a hipótese de independência, por exemplo, para um caso de identificação de sistemas, no qual $d(n) = \mathbf{X}^T(n)\mathbf{W}^o$ com \mathbf{W}^o denotando a resposta ao impulso da planta, obtemos os valores esperados $E[\mathbf{X}(n-\hat{D})\mathbf{X}^T(n-D)]$ e $E[d(n)\mathbf{X}(n-\hat{D})]$ iguais a zero para $\hat{D} \neq D$. Isso transforma a Eq. (8) em $E[\mathbf{W}(n+1)] = \nu E[\mathbf{W}(n)]$, levando a um resultado claramente incorreto. Além do mais, com base na hipótese clássica, as expressões (5) e (9) ficam sem significado, visto que $\mathbf{P}_D = E[d(n)\mathbf{X}^T(n-D)]$ e $\mathbf{P}_{\hat{D}} = E[d(n)\mathbf{X}^T(n-\hat{D})]$ são iguais a zero para quaisquer $D \neq 0$ e $\hat{D} \neq 0$, respectivamente.

2.4 Transformação dos Modelos Prévios no Modelo da Fig. 1

Os resultados obtidos a partir da Fig. 1 podem ser estendidos para descrever os casos em que o atraso encontra-se no caminho do erro [1-5]. Note que a Fig. 1 pode também modelar essa situação se deslocarmos o atraso através do ponto de soma como mostrado na Fig. 2. Dessa forma, o modelo aqui descrito pode ser transformado naquele apresentado em [1-5] pela substituição de $d(n)$ por $d(n-D)$.

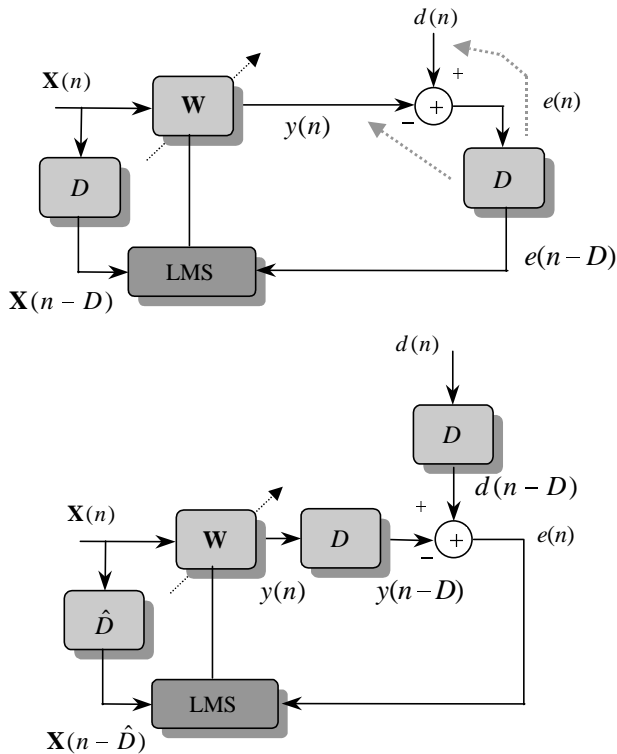


Figura 2. Transformação do modelo proposto no modelo apresentado em [1-5].

3. RESULTADOS EXPERIMENTAIS

Nesta seção, são apresentados exemplos que verificam a validade das equações (5), (8) e (9), servindo também para justificar os benefícios decorrentes da introdução de um fator de perdas no algoritmo DLMS. Para tal fim, vamos considerar, como exemplo, um esquema de identificação de sistemas, com a planta dada por: $\mathbf{W}^o = [1; 0,8; 0,6; 0,1; -0,2]^T$. O valor utilizado para a taxa de aprendizagem é $\mu = 0,001$. O sinal de entrada é um ruído branco com variância $\sigma_x^2 = 1$. A variância do ruído de medição é $\sigma_z^2 = 0,001$.

A Fig. 3 ilustra os resultados de simulações Monte Carlo (média de 100 rodadas independentes) e os resultados obtidos pelo modelo proposto (Eq. (8)), para os seguintes casos: $\{\hat{D} = 0, D = 1\}$ com $\gamma = 1,5$ (Fig. 3(a)), $\{\hat{D} = 2, D = 1\}$ com $\gamma = 1,5$ (Fig. 3(b)), e finalmente, $\{\hat{D} = D = 1\}$ com $\gamma = 1,5$ e $\gamma = 0$ (Fig. 3(c)). Nesse último caso, as curvas para $\gamma = 0$ são mostradas, a fim de se observar o efeito do fator de perdas γ sob o comportamento dos coeficientes do filtro adaptativo.

A partir da Fig. 3, podemos comprovar o bom casamento entre o modelo proposto e os resultados de simulação. Também nessa mesma figura, podemos verificar o valor em regime permanente obtido pela Eq. (9). Assim, para $\{\hat{D} = 0, D = 1\}$ e $\{\hat{D} = 2, D = 1\}$, obtemos $\mathbf{W}_\infty = [0,66; 0,08; 0,34; -0,16; -0,02]^T$ (Fig. 3(a)) e $\mathbf{W}_\infty = [0,29; 0,15; -0,13; 0; 0]^T$ (Fig. 3(b)), respectivamente.

A Fig. 3(c) ilustra o caso em que uma perfeita estimação do atraso é obtida ($\hat{D} = D = 1$). Para esse caso, através da expressão (5), obtemos $\mathbf{W}_{opt} = [0,32; 0,24; 0,04; -0,08; 0]^T$ e $\mathbf{W}_{opt} = [0,8; 0,6; 0,1; -0,2; 0]^T$ com $\gamma = 1,5$ e $\gamma = 0$, respectivamente.

4. CONCLUSÕES

Neste trabalho, é apresentado um novo modelo para descrever o comportamento médio dos coeficientes de um filtro adaptativo que utiliza o algoritmo LDLMS. Na derivação do modelo, foram utilizadas novas hipóteses de análise, ao invés da hipótese clássica, utilizada tradicionalmente para o algoritmo LMS, representada pela hipótese de independência. Assim, o modelo proposto permite avaliar e estudar certas condições de operação que, com o modelo existente, não seriam possíveis. O novo modelo é mais geral, permitindo representar também todos os casos descritos pelo modelo já existente. O bom casamento entre os resultados apresentados por simulações

Monte Carlo e o modelo proposto atestam que as novas hipóteses simplificativas são bastante adequadas.

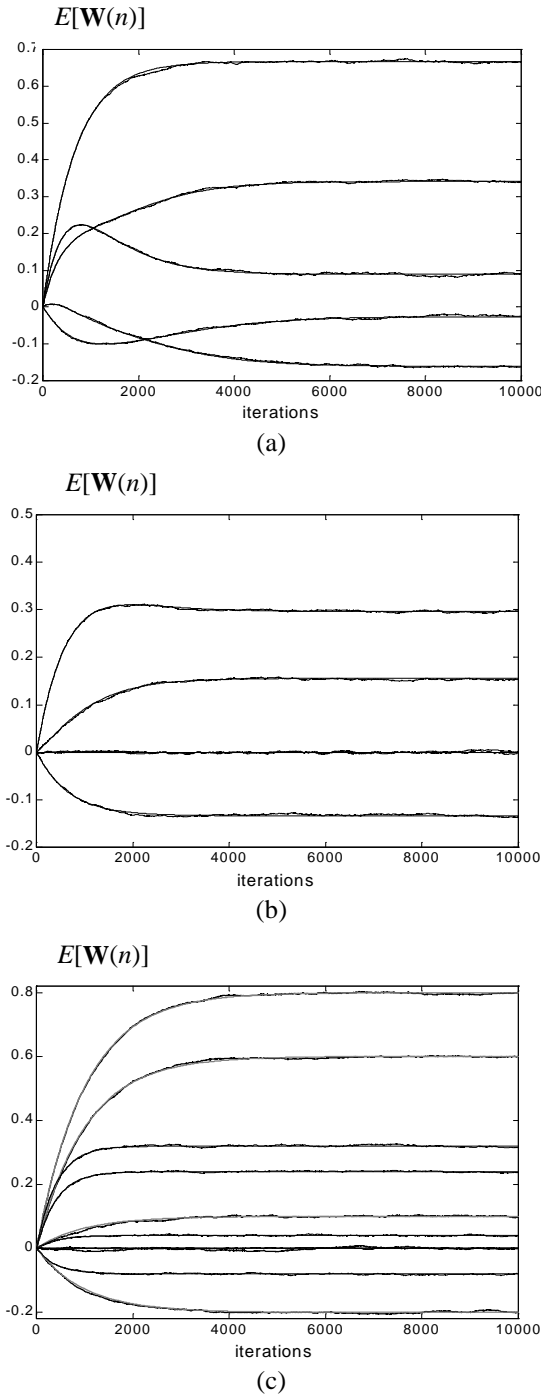


Figura 3. Evolução de $E[\mathbf{W}(n)]$ para: (a) $\hat{D} = 0$ e $D = 1$ com $\gamma = 1,5$; (b) $\hat{D} = 2$ e $D = 1$ com $\gamma = 1,5$; (c) $\hat{D} = D = 1$ quadrados: $\gamma = 1,5$ e círculo: $\gamma = 0$. Ragged curves: simulações Monte Carlo (média de 100 realizações). (—): modelo proposto (Eq. (8)).

REFERÊNCIAS

- [1] G. Long, F. Ling e J. Proakis, "The LMS with delayed coefficient adaptation," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 1397-1405, Sept. 1989.
- [2] G. Long, F. Ling e J. Proakis, "Corrections to The LMS with delayed coefficient adaptation," *IEEE Trans. on Signal Processing*, vol. 40, pp. 230-232, Jan. 1992.
- [3] P. Kabal, "The stability of adaptive minimum mean square error equalizers using delayed adjustment," *IEEE Trans. on Comm.*, vol. 31, pp. 430-432, Mar. 1983.
- [4] E. Bjarnason, "Analysis of the filtered-X LMS algorithm," *IEEE Trans. on Speech and Audio Processing*, vol. 3, pp. 504-514, Nov. 1995.
- [5] M. Rupp e R. Frenzel, "Analysis of LMS and NLMS algorithms with delayed coefficient update under the presence of spherically invariant processes," *IEEE Trans. on Signal Processing*, vol. 42, pp. 668-672, Mar. 1994.
- [6] T. Ernst e A. Kaelin, "Analysis of the LMS algorithm with delayed coefficient update," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, pp. II.1247-II.1250, 1995.
- [7] F. Laichi, T. Aboulnasr e W. Steenaart, "Effect of delay on the performance of the leaky LMS adaptive algorithm," *IEEE Trans. on Signal Processing*, vol. 45, pp. 811-813, Mar. 1997.
- [8] O. J. Tobias, J. C. M. Bermudez e N. J. Bershad, "Stochastic analysis of the Delayed LMS algorithm for a new model," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Istanbul, Turkey, pp. 404-407, 2000.
- [9] O. J. Tobias, J. C. M. Bermudez e N. J. Bershad, "Mean weigh behavior of the Filtered-X LMS algorithm," *IEEE Trans. on Signal Processing*, vol. 48, pp. 1061-1075, Apr. 2000.
- [10] J. E. Mazo, "On the independence theory of equalizer convergence," *Bell Syst. Tech. J.*, vol. 58, pp. 963-993, May-Jun. 1979.
- [11] S. Haykin, *Adaptive Filter Theory*, Third edition, Prentice Hall, 1996.