A GLOBALLY CONVERGENT MODIFIED OE IIR ADAPTIVE FILTER FOR SUFFICIENT MODELING

K. Mayyas
Department of Electrical Engg.
Jordan University of Science and Technology
Irbid, Jordan 221 10
mayyas@just.edu.jo

T. Aboulnasr
School of Information Technology and Engg.
University of Ottawa
Ottawa, Ontario, Canada K1N 6N5
aboulnas@eng.uottawa.ca

Abstract

A modification to the OE IIR system structure is proposed to ensure global convergence for sufficient modeling of the unknown system. The proposed structure is effectively equivalent to whitening the input signal before being applied to the original OE setup. This guarantees the unimodality of the error surface for sufficient modeling. An adaptive update scheme for the new structure is derived based on the least mean square (LMS) technique. Examples are provided to demonstrate the effectiveness of the proposed structure under different conditions.

1 Introduction

Though adaptive IIR filters require fewer coefficients to estimate the unknown system, the nonlinearity of the OE performance function results in error surface that can have several local minima. Gradient search techniques are commonly used in the OE formulation and are known to easily converge to a local minimum when initialized in its neighborhood . The failure to reach the global minimum leads to suboptimal solutions of the adaptive filtering problem. This fact is a major limitation restricting the commercial use of the OE IIR adaptive filters.

Some remedies were proposed but the problem is far from being completely resolved [1]. Those techniques basically revolve around using the equation-error formulation (EE) (it has a quadratic MSE function but leads to a biased estimate of the adaptive filter coefficients) first and then gradually shifting to the OE structure. Performance depends on how the decision to switch from EE to OE formulation is made. Convergence to the global minimum is thus not generally guaranteed and sometimes stability problems result when the equation-error biased global minimum falls outside the stability region.

Sufficient conditions for the unimodality of the error surface exist. It has been shown that the error surface of a sufficient order IIR filter, i.e. with enough poles and zeros to model the unknown system, has a unique minimum when the input is white [2] (this is shown to be always true for first and second order IIR filter, and true with a mild condition for higher orders [3]). When the input signal is correlated, local minima may appear [4].

In this paper, we address the problem of local minima in OE IIR filters from an algorithm-independent perspective. Given that a unique optimum results for white input sufficient order, we propose first a structure that incorporates a whitener to transform the input signal to the unknown system and adaptive IIR filter from a correlated one to a white one. The proposed structure is practically limited, and therefore, a modified setup is presented to closely approximate the original one for slow adaptation. The least mean square (LMS) update equations are derived for the coefficients of the simplified structure.

The proposed simplified setup can also be viewed as a filtered-X IIR setup with the filter being an adaptive whitener of the input signal instead of a fixed correlating filter in the standard FIR Filtered-X structure [7].

2 The proposed modified OE formulation

The unknown system to be identified here is assumed to be a general ARMA system. The output of the adaptive filter in Fig.1 is given by

$$\begin{aligned}\hat{d}(n) &= \sum_{j=1}^{\hat{M}} a_j(n)\hat{d}(n-j) + \sum_{j=0}^{\hat{N}-1} b_j(n)x(n-j)\\ &= A(q^{-1},n)\hat{d}(n) + B(q^{-1},n)x(n)\\ &= \mathbf{A}^{\mathbf{T}}(n)\hat{\mathbf{D}}(n) + \mathbf{B}^{\mathbf{T}}(n)\mathbf{X}(n)\\ &= \Theta^{\mathbf{T}}(\mathrm{n})\Psi(n) \qquad (1)\end{aligned}$$

where $A(q^{-1},n) = \sum_{j=1}^{\hat{M}} a_j(n)q^{-j}$ and $B(q^{-1},n) = \sum_{j=0}^{\hat{N}-1} b_j(n)q^{-j}$ where q^{-j} is the delay operator, $a_j(n)$ and $b_j(n)$ are the coefficients of the adaptive IIR filter computed at time (n), $\Theta(n) =$

$[\mathbf{A^T}(n),\ \mathbf{B^T}(n)]^T$, $\Psi(n) = [\hat{\mathbf{D}}^\mathbf{T}(n),\ \mathbf{X^T}(n)]^T$, $\mathbf{A}(n) = [a_1(n), a_2(n), \ldots, a_{\hat{M}}(n)]^T$, $\mathbf{B}(n) = [b_0(n), b_2(n), \ldots, b_{\hat{N}-1}(n)]^T$, $\hat{\mathbf{D}}(n) = [\hat{d}(n), \hat{d}(n-2), \ldots, \hat{d}(n-\hat{M})]^T$, and $\mathbf{X}(n) = [x(n), x(n-1), \ldots, x(n-\hat{N}+1)]^T$. The output-error (OE) in Fig.1 is defined by

$$e(n) = d(n) - \hat{d}(n) \qquad (2)$$

where $d(n)$ is the output of the unknown system and is given by $d(n) = \sum_{j=1}^{M} a_j^* d(n-j) + \sum_{j=0}^{N-1} b_j^* x(n-j)$ where a_j^* and b_j^* are the coefficients of the unknown ARMA system. The coefficients of a_j^* and b_j^* are to be estimated. We assume that the adaptive filter appropriately models the unknown system, i.e., $n^* = min(\hat{M} - M, \hat{N} - N) \geq 0$. This is commonly known as the "sufficient order" case of the OE setup. A fundamental characteristic of the OE formulation is the possible existence of local minima. Stearn [2] conjectured that for $n^* \geq 0$ and with a white input signal, the error surface has a unique global minimum. Fan and Nayeri [3] showed that Stearn's conjecture holds only for first ($\hat{M} = 1$) and second ($\hat{M}=2$) order IIR filters . They showed that for higher order ($\hat{M} > 2$), sufficient modeling adaptive IIR filters, and a white input signal $x(n)$, the unimodality of the error surface is guaranteed by an additional constraint presented by Soderstrom and Stoica [5], namely $\hat{N} - M \geq 0$. Clearly, the nature of the input signal is a vital factor in determining the unimodality/multimodality of the error surface. A second order example was constructed in [4] to show that with a sufficient order case and a colored input, the error surface can be multimodal. Given the earlier results on white input, sufficient order, it follows that if the colored input signal is transformed into a white one before being applied to the setup, then the performance surface will have a unique minimum.

Based on the above discussion, a straightforward solution is to insert a "decorrelator" block as shown in Fig.2. The output of the decorrelator is given by $x_f(n) = Q(q^{-1}, n)x(n)$, $Q(q^{-1}, n) = 1 - \sum_{i=1}^{K} c_i(n) q^{-i}$ and the coefficients $c_1(n), c_2(n), \ldots, c_K(n)$ are adaptively adjusted to minimize the mean squared error $E\{x_f^2(n)\}$. Assuming the adaptive decorrelator has converged to the set of optimal coefficients that produce the minimum of $E\{x_f^2(n)\}$, and if the decorrelator filter is of sufficient order to whiten $x(n)$, then $x_f(n)$, the input to the adaptive IIR filter and the unknown system, will consist of white samples. The performance surface will be re-shaped to a unimodal one. However, the structure in Fig.2 is practically very limited since the decorrelator alters the input to the *unknown system* as well as the adaptive filter. Also, it does not retain the input/output relationship of the original structure of Fig.1. Note that at steady state, the adaptive IIR filter and the decorrelator coefficients converge and can be treated as time-invariant systems. Under such conditions, the structure in Fig.2 is equivalent to the structure in Fig.3. For slow adaptation, Fig.3 setup closely approximates the one in Fig.2. Upon convergence, the two structures asymptotically behave the same.

Analogous to the setup in Fig.2, the objective function in Fig.3 structure is to minimize the MSE, $E\{e_f^2(n)\}$ where

$$e_f(n) = d_f(n) - \hat{d}_f(n) \qquad (3)$$

and $\hat{d}_f(n) = Q^*(q^{-1})(\ A(q^{-1}, n)\hat{d}(n) + B(q^{-1}, n)x(n))$ To simplify the adaptation mechanism, we assume slow adaptation of the adaptive IIR filter coefficients such that

$$\hat{d}_f(n) = A(q^{-1}, n)\hat{d}_f(n) + B(q^{-1}, n)x_f(n) \qquad (4)$$

where $x_f(n) = Q^*(q^{-1})x(n)$ and $\hat{d}_f(n) = Q^*(q^{-1})\hat{d}(n)$. Note that we assumed that the decorrelator has converged to the optimal filter $E\{Q(q^{-1}, \infty\} = Q^*(q^{-1})$. The least mean square (LMS) method is used to minimize asymptotically the mean squared error (MSE) of the OE in Eq.(3), i.e., $E\{e_f^2(n)\}$, relative to $\Theta(n)$. Thus, the resulting approximate stochastic update is

$$\begin{aligned} \Theta(n+1) &= \Theta(n) - \mu \frac{\partial e_f^2(n)}{\partial \Theta(n)} \\ &= \Theta(n) + 2\mu e_f(n) \frac{\partial \hat{d}_f(n)}{\Theta(n)} \qquad (5) \end{aligned}$$

The step size μ is usually small in the IIR filtering application, then $\Theta(n) \approx \Theta(n-1) \ldots \approx \Theta(n-M)$, $\frac{\partial \hat{d}_f(n-j)}{\partial \Theta(n)} \approx \frac{\partial \hat{d}_f(n-j)}{\partial \Theta(n-j)}$, $j = 1,\ 2,\ \ldots, \hat{M}$ and

$$\frac{\partial \hat{d}_f(n)}{\partial \Theta(n)} \approx \Psi_\mathbf{f}(n) + A(q^{-1}, n) \frac{\partial \hat{d}_f(n)}{\partial \Theta(n)} \qquad (6)$$

where $\Psi_\mathbf{f}(n) = [\hat{d}_f(n-1), \hat{d}_f(n-2), \ldots, \hat{d}_f(n-\hat{M}), x_f(n), x_f(n-1), \ldots, x_f(n-\hat{N}+1)]^T$. Substituting Eq.(6) in Eq.(5) gives the update recursion of the OE LMS algorithm of the structure of Fig.3,

$$\Theta(n+1) = \Theta(n) + 2\mu \Psi_\mathbf{f}^\mathbf{f}(n) e_f(n) \qquad (7)$$

where $\Psi_\mathbf{f}^\mathbf{f}(n) = \frac{1}{(1-A(q^{-1},n))} \Psi_\mathbf{f}(n)$. A more practical approximation is to obtain the components of $\Psi_\mathbf{f}^\mathbf{f}(n)$ from the following equations $x_f^f(n-i) = \frac{1}{(1-A(q^{-1},n-i))} x_f(n-i)$ $\hat{d}_f^f(n-j) = \frac{1}{(1-A(q^{-1},n-j))} \hat{d}_f(n-j)$ where $i = 0, 1, \ldots, \hat{N}-1$, and $j = 1, 2, \ldots, \hat{M}$. The LMS algorithm is also used to update the linear predictor coefficients. Taking the gradient of $x_f^2(n)$ with respect to $c_i(n), i = 1, 2, \ldots, K$, yields the LMS adaptation equation of the decorrelator coefficients:

$$\mathbf{C}(n+1) = \mathbf{C}(n) + 2\mu_p x_f(n)\mathbf{X_P}(n) \quad (8)$$

where $\mathbf{X_P}(n) = [x(n-1), x(n-2), \ldots, x(n-K)]^T$ and $\mathbf{C}(n) = [c_1(n), c_2(n), \ldots, c_K(n)]^T$.

The coefficients of the adaptive IIR filter of Fig.1 $\Theta(n)$ are adjusted to minimize the square error of Eq.(2). Following the same derivation procedure of the LMS IIR filter of Fig.3, we get $\Theta(n+1) = \Theta(n) + 2\mu_{LMS}\Psi^{\mathbf{f}}(n)e(n)$, where $\Psi^{\mathbf{f}}(n) = \frac{1}{(1-A(q^{-1},n))}\Psi(n)$.

The nonquadratic nature of the resultant MSE function complicates the convergence of the LMS method leading the algorithm to converge very slowly. Following the approach in [6, p.26-29], an RLS version of the proposed algorithm can be derived by minimizing the objective functions $\epsilon_{LS}(n) = \sum_{i=1}^{n} \lambda^{n-i}(d_f(i) - \hat{d}_f(i/n+1)\ )^2$ and $\epsilon_{LSp}(n) = \sum_{i=1}^{n} \lambda^{n-i}(x(i) - \mathbf{C^T}(n+1)\mathbf{X_P}(i)\ )^2$ where $\hat{d}_f(i/n+1) = Q^*(q^{-1})(\ A(q^{-1}, n+1)\hat{d}(i/n+1) + B(q^{-1}, n+1)x(i)\ )$ and λ is an exponential weighting parameter, $0 < \lambda \leq 1$, $\hat{d}(i/n+1)$ is the output of the adaptive filter at time i using the coefficient vector $\Theta(n+1)$.

3 Simulations

In the following two examples, the desired signal has an additional zero-mean white Gaussian noise of 0.0001 variance. Results are obtained by averaging over 50 ensemble average.

Example 1

In the first example, the system to be modeled is $H(z) = \frac{1}{1-1.4z^{-1}+0.49z^{-2}}$ and the adaptive filter transfer function is $H(z,n) = \frac{b_0(n)}{1-a_1(n)z^{-1}-a_2(n)z^{-2}}$. The input signal is obtained by passing a zero-mean uncorrelated Gaussian signal with unity variance through the colouring filter $(1-0.7z^{-1})^2(1+0.7z^{-1})^2$ (SNR=43 dB). The resulting error surface has two minima The global minimum is located at $[b_0, a_1, a_2] = [1, 1.4, -0.49]$, and the local minimum is approximately at $[-0.22, -1.35, -0.49]$. Both algorithms of Fig.2 and Fig.3 are used with the same step size $\mu = 0.002$. The whitening algorithm step size is set to $\mu_p = 0.01$. The initial point for both algorithms is given by $[b_0(0), a_1(0), a_2(0)] = [0, -1.4, -0.5]$ which was chosen near the local minimum. It is found that a decorrelator of length 20 is sufficient to whiten the input signal. Results are shown in Fig.4 (a). The structure in Fig.3 is seen to converge to the global minimum whereas that of Fig.1 converges to the local minimum. It is important to examine the influence of the whitener order on the performance of the proposed algorithm. The above example is repeated with $K = 5$. The step size used with the whitener is $\mu_p = 0.01$. Fig.4 (b) demonstrates the insufficiency of the whitener order to generate a unimodal error surface. The local minimum still exists and, as expected, the algorithm converges to it.

Example 2

To demonstrate the applicability of the proposed structure in Fig.3 for higher order filters, we consider a third order IIR filter, $H(z) = \frac{0.5-0.4z^{-1}-1.83z^{-2}-z^{-4}}{1-0.5z^{-3}}$ and that of the adaptive filter is $H(z,n) = \frac{b_0(n)+b_1(n)z^{-1}+b_2(n)z^{-2}+b_3(n)z^{-3}+b_4(n)z^{-4}}{1-a_1(n)z^{-1}-a_2(n)z^{-2}-a_3(n)z^{-3}}$. Note that $\hat{N} = 5$ and $M = 3$. Accordingly, $\hat{N} - M > 0$. The transfer function of the coloring filter is $(1-0.7z^{-1})^2(1+0.7z^{-2})^2$ ($SNR = 43dB$). All coefficients are initialized to zero. Results when applying the LMS and RLS update algorithm of the structure in Fig.3 are shown on Fig.5. The order of the decorrelator is $K = 20$. The step sizes were $\mu = 0.0008$ and $\mu_p = 0.01$, and the forgetting factor is $\lambda = 0.99$. Fig.5 demonstrates the convergence of the proposed structure in Fig.3 to the global minimum for higher order cases without problems.

4 Conclusion

We proposed a new OE IIR update structure that basically whitens the input signal. The new structure guarantees that the adaptive updating algorithm operates on a unimodal surface, provided sufficient whitening of the input signal, without altering the input to the adaptive filter and unknown system. Performance of the proposed structure was confirmed through simulations.

References

[1] J. B. Kenney, and C. E. Rohrs, "The Composite Regressor Algorithm for IIR Adaptive Systems", *IEEE Trans. Signal Processing*, Vol. 41, no. 2, pp. 617-628, Feb. 1993.

[2] S. D. Stearns, " Error Surfaces of Recursive adaptive Filters", *IEEE Trans. Acoust., Speech, and Signal Processing*, Vol. ASSP-29, No. 3, June 1981.

[3] H. Fan, and M. Nayeri, "On Error Surfaces of Sufficient Order Adaptive IIR Filters: Proofs and Counterexamples to a Unimodality Conjecture", *IEEE Trans. Acoust., Speech, and Signal Processing*, Vol. 37, no. 9, pp. 1436-1442, Sept. 1989.

[4] T. Soderstrom, "On the Uniqueness of a Maximum Likelihood Identification", Automatica, Vol. 11, pp. 193-197, 1975.

[5] T. Soderstrom, and P. Stoica, "Some Properties of the Output Error Method," *Automatica*, Vol. 18, pp. 93-99, Jan. 1982.

[6] L. Ljung and T. Soderstrom, *Theory and Practice of Recursive Identification*, MIT Press, 1983.

[7] M.Mboup et al, "LMS Coupled Adaptive Prediction and System Identification: A Statistical Model and Transient Mean Analysis, " IEEE Trans. on Signal Processing, Vol. 42, no 10, 1994.

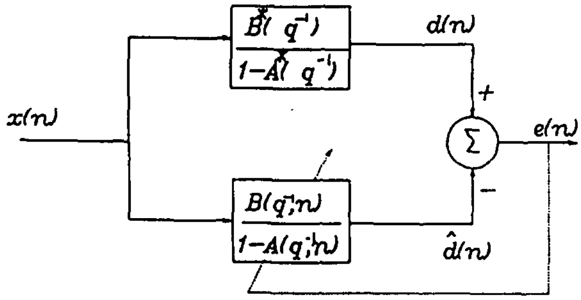

Fig.1 Output-Error IIR adaptive formulation model.

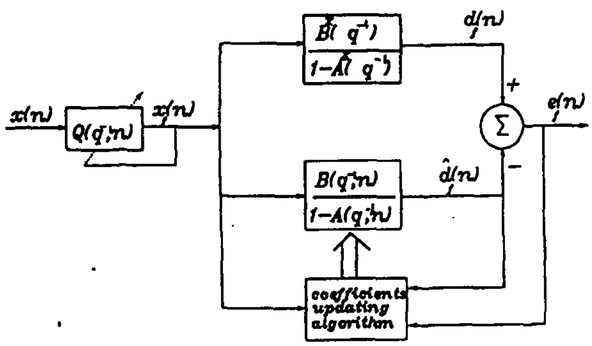

Fig.2 The proposed new structure for OE IIR adaptive filtering.

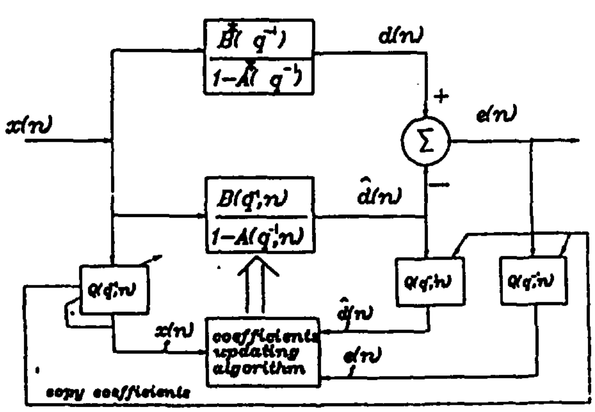

Fig.3 Approximate of the new structure for OE IIR adaptive filtering.

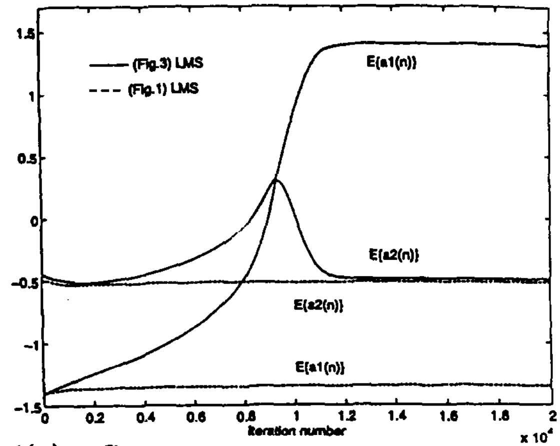

Fig.4(a) Comparison of ensemble average of $\mathbf{A}(n)$ between the structures of Fig.1 and Fig.3, using the LMS algorithm and with $K = 20$, for the first example.

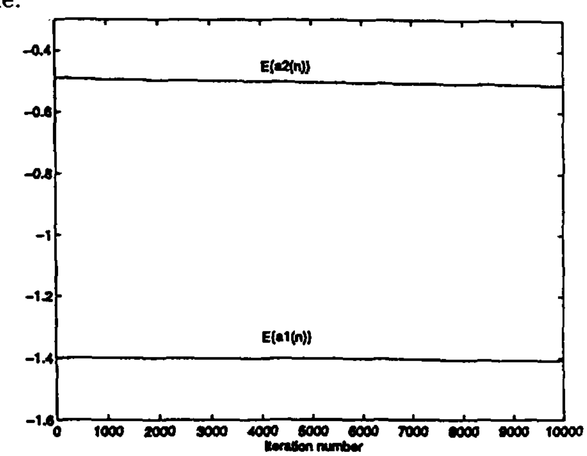

Fig. 4(b) Ensemble average of $\mathbf{A}(n)$ of the structures of Fig.3, using the LMS algorithm and with $K = 5$, for the first example.

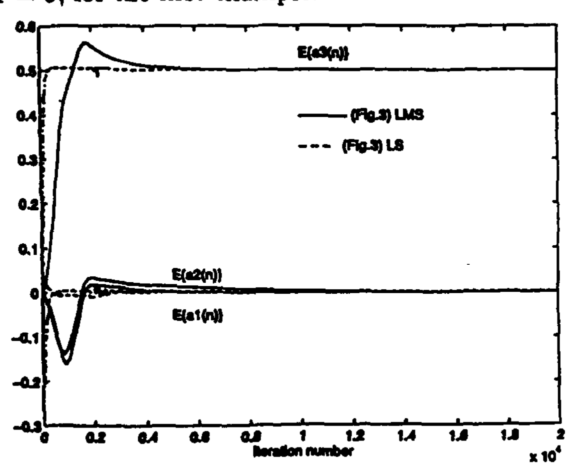

Fig.5 Ensemble average of $\mathbf{A}(n)$ of the structures of Fig.3, using the LMS and RLS algorithms, for the second example.