A NON STATIONARY LMS ALGORITHM FOR ADAPTIVE TRACKING OF A MARKOV TIME-VARYING SYSTEM

M. TURKI, M. JAIDANE-SAIDANE

L.S.Télécoms, ENIT, Campus Universitaire, Le Belvédère, Tunis, TUNISIA
Telephone: (216)1514700; E–mail: Jaidane@enit.rnrt.tn

Abstract

We propose in this paper a new adaptive algorithm which is designed to track system represented by a filter which has a P order markovian time evolution. The Non Stationary LMS (NSLMS) algorithm is able to identify the unknown order and parameters of the markov model. An analysis of the performances of the adaptive filter when the input is i.i.d. shows that the NSLMS presents better performances than the classical LMS. In particular, this superiority occurs when the system time evolution is so fast that the tracking with LMS is harmful.

1 Introduction

The most widely used among the gradient based adaptation algorithm is the Least Mean Square (LMS) algorithm. Basically the LMS algorithm is designed to estimate recursively the value of a fixed unknown filter. However, in a non stationary context this algorithm has interesting tracking performances. This steady state property has been extensively analyzed in the literature for random walk variations (see for example [1]).
In fact, the adaptive identification offered by the LMS is blind regarding to the nature of the time evolution model of the real filter.

In order to guarantee better results than those realized by the classic LMS, we propose, in this paper a new algorithm that can identify the markovian time evolution of the real filter, encountered in transmission systems. Contrarily to the Kalman approach, the proposed NSLMS doesn't suppose a prior knowledge of the non stationarity structure and the unknown statistics of the observation noise and the filter noise, [2]. In other respects, the NSLMS constitutes a new approach different from those relative to the gradient algorithms based on the idea of a variable convergence factor, [3], [4].

2 Presentation of the problem

In this paper we are interested in adaptive identification of markovian time varying filters. The classical formulation of such filtering problem is depicted in figure (1).

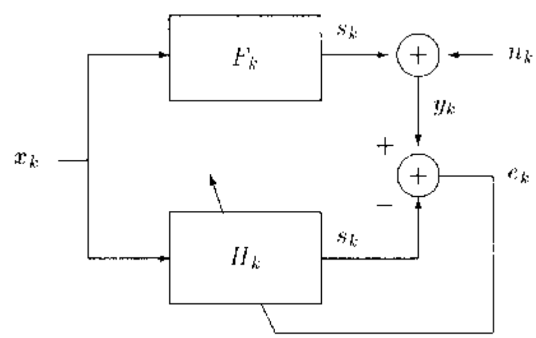

Figure 1: Adaptive identification of time varying filter

The noisy input/output equation of the filter is,

$$y_k = F_k^T X_k + n_k \tag{1}$$

where, $X_k = (x_k, x_{k-1}, \cdots, x_{k-N+1})^T$ is the known stationary input vector and n_k is an unknown i.i.d. observation noise. The filter parameter vector is assumed P order markov time varying,

$$F_k = \sum_{i=1}^{P} a_i F_{k-i} + \Omega_k \tag{2}$$

where $(a_i)_{i=1,P}$ ensure the stability of the filter, and Ω_k called non stationary noise, is an unknown zero−mean, i.i.d. process independent of X_k and n_k. This general model represents different types of non stationarity such as variations of oscillatory nature. In particular, the non stationarity of an ionospheric radio mobile transmission channel is well represented by second order markov model, [5].

The evolution of the parameter vector H_k of the adaptive filter is governed by the estimate error.

$e_k = y_k - H_k^T X_k$, in order to minimize a criterion, such as the mean square error $E(e_k^2)$, for the LMS.
The tracking capacity of the adaptive algorithm is measured by the normalized misadjustement, $M = \lim_{k\to\infty} (E(e_k^2) - P_n)/P_n$, where P_n is the power of n_k.

3 Design of the Non Stationary LMS algorithm

The Non Stationary (NSLMS) algorithm is designed in order to take in to account the prior knowledge of the structure non stationarity model, (2). We keep the structure of the classical LMS, described by $H_{k+1} = H_k + \mu e_k X_k$, and we include the constraints on the nature of the non stationarity as follows:

$$H_{k+1} = \sum_{i=1}^{\widehat{P}} \widehat{a}_i H_{k+1-i} + \mu X_k e_k \qquad (3)$$

where $\widehat{P}$ is an estimation of P, the exact order of the markov model, and $(\widehat{a}_i)_{i=1,\widehat{P}}$ is an adaptive estimation of the unknown parameters of the markov model.
The adaptive estimation of each parameter $\widehat{a}_i$ of the model is made in order to still minimize $E(e_k^2)$. The gradient of e_k^2 is given by,

$$\frac{de_k^2}{da_i}\Big|_{\widehat{a}_i=\widehat{a}_i(k)} = 2e_k \left(H_{k-i}^T X_k + \sum_{i=1}^{\widehat{P}} \widehat{a}_i \frac{dH_{k-i}}{d\widehat{a}_i} \right)_{\widehat{a}_i=\widehat{a}_i(k)} \qquad (4)$$

This complexity is due to recursive nature of the markovian structure (3).
In order to simplify the algorithm, we use an approximation of the true gradient, in such a way that the Non Stationary LMS algorithm is described by,

$$e_k = y_k - H_k^T X_k \qquad (5)$$

$$\widehat{a}_i(k+1) = \widehat{a}_i(k) + \mu_i (H_{k-i}^T X_k) e_k \qquad (6)$$

$$H_{k+1} = \sum_{i=1}^{\widehat{P}} \widehat{a}_i(k+1) H_{k+1-i} + \mu X_k e_k \qquad (7)$$

where $(\mu_i)_{i=1,P} > 0$ is a (small) step size that controls the adaptive identification of a_i.

4 Steady state behaviour of the NSLMS: first order markov case

We consider here a first order markov non stationarity,

$$F_k = aF_{k-1} + \Omega_k$$

The analysis of the two coupled adaptation, (6) (7), is complex. First, we analyze the tracking capacity of the algorithm in the steady state when the time evolution of the adaptive filter is defined by,

$$H_{k+1} = \widehat{a} H_k + \mu X_k e_k \qquad (8)$$

where $|\widehat{a}| < 1$ is a fixed estimation of the parameter a. The purpose of this section is to establish a comparative study between NSLMS and LMS.

4.1 Residual misadjustement

Let, $V_k = H_k - F_k$ denote the filter deviation vector at time k, consequently $e_k = n_k - V_k^T X_k$. It's easy to show that the deviation vector V_k obeys to a linear recurrence,

$$V_{k+1} = \left(\widehat{a} - \mu X_k X_k^T\right) V_k + \mu X_k n_k + (\widehat{a} - a) F_k - \Omega_k \quad (9)$$

Under the independence assumption between V_k and X_k, we can prove that the normalized misadjustement is composed of two parts,

$$M_\nu^{NSLMS} = M_\nu^S + M_\nu^{NS} \qquad (10)$$

where $\nu = \mu N P_x$ is the normalized step size.

- the first part

$$M_\nu^S = \frac{\nu^2}{\nu(2\widehat{a} - \nu) + N(1 - \widehat{a}^2)} \qquad (11)$$

is increasing as ν increases and is related to the associate stationary context [6].
The calculus of this part is elaborated without the independence assumption.

- the second part

$$M_\nu^{NS} = \frac{\delta N \mathcal{S}}{\nu(\widehat{a} - \nu) + N(1 - \widehat{a}^2)} \qquad (12)$$

where $\delta = P_x P_\Omega / P_n$, and

$$\mathcal{S} = \frac{(\widehat{a} - a) + (1 - \widehat{a}a)\left[\widehat{a}\nu + N(1 - \widehat{a}^2)\right]}{(1 - a^2)\left[a\nu + N(1 - a\widehat{a})\right]}$$

is decreasing as ν decreases and is relative to the lag component of the misadjustement.

The above result is made when the sequence $\{x_k\}$ is characterized by $|x_k| = Cte$. The calculus of the misadjustement is also possible under the assumption of the gaussian sequence; the results are slightly different.

4.2 Limitations of the LMS

The misadjustement of the LMS is given by equation (10–12) for $\widehat{a} = 1$. From (10) we deduce that,

$$M_\nu^{LMS} = \frac{1}{2-\nu}\left[\nu + \frac{2\delta N}{(1+a)(a\nu + N(1-a))}\right] \quad (13)$$

The analysis of the above result show the superiority of the NSLMS over the LMS which present some limitations for certain type of variations. In fact the analysis of the theoretical expression of the optimal step size (ν_{opt}), that minimizes M_ν^{LMS}, yields to the following results. They display prominently a new aspect of tracking which we called harmful tracking [7].

- The normalized misadjustement M_ν^{LMS} is increasing for $\nu > 0$, ($\nu_{opt} < 0$), when $a < \frac{N}{N+2}$. This condition shows that the tracking is harmful for all values of the power, P_Ω, of the non stationarity noise. In such case, a positive value of ν is necessary to initialize the process. After convergence of the LMS, we must stop the identification that damages the performances.

- In the case of important values of a, ($a > \frac{N}{N+2}$) the tracking is harmful when $\frac{P_s}{P_n} < \frac{N(1-a)}{2a - N(1-a)}$. P_s is the power of the output of the real filter, ($s_k = F_k^T X_k$). Also here, we must stop the identification because the observation noise is so important to allow the tracking.

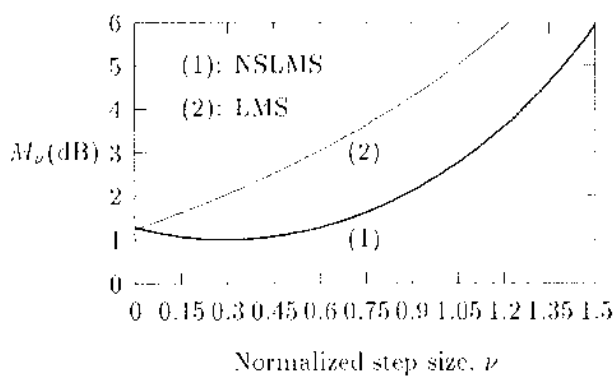

Figure 2: Superiority of NSLMS over LMS ($a = 0.5$, $P_x = 1$ and $\delta = 1.0$)

Figure (2) illustrates this limitation of the LMS. In this figure, we superpose the theoretical results given by LMS and NSLMS when $\delta = 1.0$ and $a = 0.5$. For the LMS the tracking is harmful, however the NSLMS is able not only to track this non stationarity but also to make better than LMS for this markovian variation.

4.3 Optimal steady state behavior of the NSLMS

When we compare M_ν^{NSLMS}, (10) with M_ν^{LMS}(13), it's easy to show that the misadjustement given by NSLMS is less than that given by LMS. However, the choice of $\hat{a} = a$ doesn't guarantee the best performance for all values of ν. In fact, the theoretical study of the function $f(\nu, a, \hat{a}) = M_\nu^{NSLMS} - M_\nu^{LMS}$ shows in particular that the minimum misadjustement that corresponds to the optimal step size is realized for $\hat{a} = a$. The misadjustement is then given by

$$M_{NSLMS}(\hat{a} = a) = \frac{1}{\nu(2-\nu) + N(1-a^2)} \left[\nu^2 + \delta N\right] \tag{14}$$

The implementation of the theoretical result (10), allow us to derive the minimum misadjustement realized by the NSLMS for different values of $\hat{a}$.

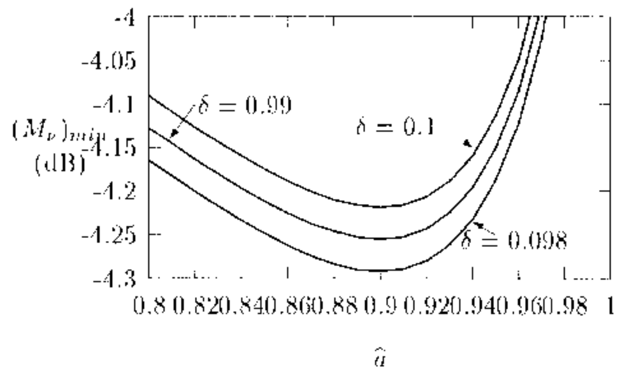

Figure 3: The best performance of the NSLMS correspond to $\hat{a} = a = 0.9$

In figure (3), we plot, for three values of δ, the minimum misadjustement $(M_\nu^{NSLMS})_{min}$ versus $\hat{a}$. Here we choose $a = 0.9$ for F_k time variation. All curves show that the minimum misadjustement corresponds to $\hat{a} = 0.9$.

5 Adaptive behavior of NSLMS: identification of the non stationarity

In this section, we will present the results of several simulations that demonstrate the good properties of the algorithm described by (5 − 7). The simulations are realized for an i.i.d excitation characterized by a power $P_x = 1$..
The real filter used for this purpose, has a length $N = 3$ and is represented by a first markov model, $F_k = aF_{k-1} + \Omega_k$. We consider an adaptive filter of order 1 or 2. The power of the noise Ω_k depends on the value of δ chosen.

5.1 Superiority of NSLMS over LMS

First, we suppose that the order of the non stationarity model is known ($\hat{P} = P = 1$).

Figure (4), where the misadjustement is plotted versus the normalized step size, correspond to $a = 0.8$ and $\delta = 1.5$. Two values of the step size μ_1 are used, 0.01 (fig 4.b) and 0.05 (fig 4.c).

The analysis of the curves (4 a–d) presented, yields to

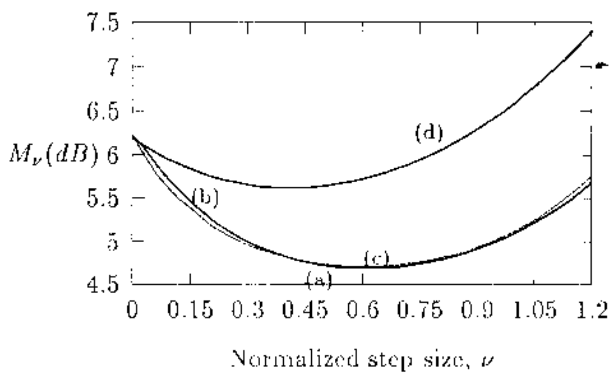

Figure 4: Superiority of NSLMS over LMS ($\delta = 1.5$, $a = 0.8$)

the following results,

– the NSLMS has better performances than the LMS (fig 4–d).

– the performances of the NSLMS don't vary significantly when we change the step size μ_1.

– when we estimate the parameter, the performances of the NSLMS are equivalent to those given when we fix the parameter to its exact value 0.8 (fig 4a). In particular, this result show that convergence of the adaptive parameter $\widehat{a}(k)$ to 0.8 is realized for $\mu_1 = 0.01$ and $\mu_1 = 0.05$.

5.2 The identification of the non stationarity

The NSLMS is able to identify the non stationarity markov model.

For this purpose, we use a first order markov filter, $F_k = aF_{k-1} + \Omega_k$. However, for the adaptive filter, we fix $\widehat{P}$ to 2, in such a way that, $H_{k+1} = \widehat{a}_1(k+1)H(k) + \widehat{a}_2(k+1)H(k-1) + \mu e_k X_k$. The variations of F_k are characterized by $a = 0.9$ and $\delta = 1$.

In figure (5), we show the time evolution of the parameters $\widehat{a}_1(k)$ and $\widehat{a}_2(k)$ versus time, when ν is fixed to its optimal value, 0.23. The convergence of $\widehat{a}_1(k)$ to 0.9, and $\widehat{a}_2$ to 0, is obvious.

6 Conclusion

We present in this paper a new algorithm (NSLMS) that takes into account the time evolution of the real filter which it must track. The superiority of the NSLMS over the LMS is its ability to identify the markovian non stationarity of the filter to reach.

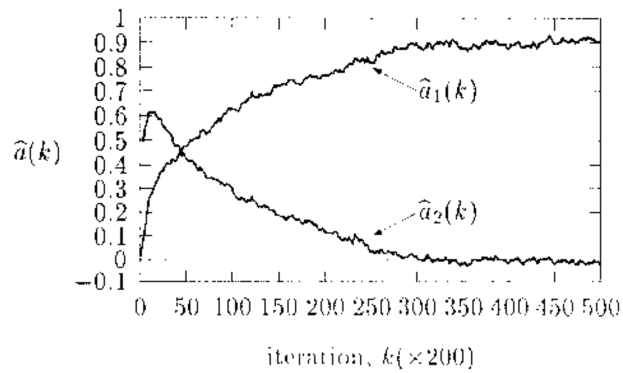

Figure 5: Convergence of the adaptive parameters

References

[1] O. Macchi. Adaptive signal processing, the least mean squares approach with application in transmissions. *Wiely*, 1994.

[2] Lars Lindbom. A simplified kalman estimation of fading mobile radio channels: high performance at lms computational load. *Proceedings of ICASSP*, 3:352–355, 1993.

[3] V. John Mathews and Zhenhua Xie. A stochastic gradient adaptive filter with gradient adaptive step size. *IEEE Trans. on Acoust. Speech and Signal Processing*, ASSP-41(6):2075–2087, June 1993.

[4] Paulo S. R. DinizLars, M.L.R. de Campos, and A. Antoniou. Analysis of lms–newton adaptive filtering algorithms with variable convergence factor. *IEEE Trans. on Acoust. Speech and Signal Processing*, ASSP-43(3):617–621, March 1995.

[5] Robert A. Ziegler and John M. Cioffi. Estimation of time-varying digital radio channels. *IEEE Trans. on vehicular technology*, 41(2):134–151, May 1992.

[6] M. Turki. Poursuite adaptative de non stationnarités markoviennes. *PHD Thesis (preliminary report)*, February 1995.

[7] O. Macchi and M. Turki. The non stationarity degree: can an adaptive filter be worse than no processing ? *ACASP, Grenoble*, 1992.