Digital Signal Processing 33 (2014) 50–59

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Transient and steady-state MSE analysis of the IMPNLMS algorithm

Diego B. Haddad [b,*], Mariane R. Petraglia [a]

[a] Federal University of Rio de Janeiro, EPOLI, COPPE, CP 68504, CEP 21945-970, Rio de Janeiro, Brazil
[b] CEFET, Nova Iguaçu, and Federal University of Rio de Janeiro, COPPE, Rio de Janeiro, Brazil

ARTICLE INFO

Article history:
Available online 9 July 2014

Keywords:
Adaptive filters
Convergence analysis
Sparse impulse response identification
Proportionate normalized least mean squares algorithm

ABSTRACT

Several techniques have been proposed in the literature to accelerate the convergence of adaptive algorithms for the identification of sparse impulse responses (i.e., with energy concentrated in a few coefficients). Among these techniques, the improved μ-law proportionate normalized least mean squares (IMPNLMS) algorithm is one of the most effective. This paper presents an accurate transient analysis of this algorithm and derives an estimate of its steady-state MSE, without requiring the assumption of white Gaussian input signals.

1. Introduction

This paper focuses on the class of supervised adaptive algorithms that seek the identification of sparse impulse responses, common in various acoustic, chemical and seismic processes, as well as in wireless communications channels [1,2]. Here, a sequence is considered sparse if most of its elements are close to zero, which implies a concept of sparsity weaker than that usually employed in numerical analysis [3].

There is no mathematical difficulty in applying supervised algorithms to identify such responses, but some practical problems may arise, the main one being a slow convergence. In this paper we investigate the convergence behavior of algorithms of the PNLMS family (which follow the paradigm proposed by [4] to accelerate the convergence of the identification method), analyzing the performance evolution of one of its most successful algorithms (the so-called IMPNLMS [5]) and obtaining an estimate of its mean-square error in steady state.

If the adaptive filter has length L, we can define the input vector $\mathbf{x}_k$ in terms of the input signal $x(k)$ as

$$\mathbf{x}_k = \begin{bmatrix} x(k) & x(k-1) & \cdots & x(k-L+1) \end{bmatrix}^T. \tag{1}$$

Although the input signal can be colored, the vast majority of the transient analysis of PNLMS-type algorithms found in the literature assume that it is white [4,6–8]. Violation of this assumption results in substantial differences in the convergence of the algorithms.

In the k-th iteration, the adaptive weight vector $\widehat{\mathbf{h}}_k$ is expressed as[1]

$$\widehat{\mathbf{h}}_k = \begin{bmatrix} \widehat{h}_k(0) & \widehat{h}_k(1) & \cdots & \widehat{h}_k(L-1) \end{bmatrix}, \tag{2}$$

and the optimal response to be provided by the adaptive algorithm is $d(k) = \mathbf{h}\mathbf{x}_k$. In this work, we assume that the length of the filter impulse response $\mathbf{h}$ to be identified is equal to or less than L and that the uncertainty on the desired response measurement can be modeled by an additive white Gaussian noise $\nu(k)$, which is independent of the input signal and has variance σ_ν^2. The unpredictability of this noise prevents its removal, and hence the adaptive system has only access to $\widehat{d}(k) = d(k) + \nu(k)$. The adaptive algorithm should change the parameters to minimize a cost function dependent on the measured error defined as

$$e(k) = \widehat{d}(k) - y(k) = \big[\mathbf{h} - \widehat{\mathbf{h}}_k\big]\mathbf{x}_k + \nu(k), \tag{3}$$

where $y(k) = \widehat{\mathbf{h}}_k \mathbf{x}(k)$ is the adaptive filter output signal at instant k. Fig. 1 illustrates the structure of a typical supervised adaptive identification algorithm.

The process of minimizing the cost function determines the characteristics of the adaptive learning system. Among the various cost functions found in the literature, the most popular is the squared error $e^2(k)$ [9–11], which can be interpreted as an instantaneous estimate of the mean square error (MSE). The corresponding adaptation algorithm, known as Least-Mean-Square (LMS) algorithm, employs the steepest descent optimization method. Its normalized version (the NLMS algorithm), accelerates the convergence rate by varying the learning factor along the iterations,

* Corresponding author.
E-mail addresses: diego@pads.ufrj.br (D.B. Haddad), mariane@pads.ufrj.br (M.R. Petraglia).

[1] The weight vectors are defined as row vectors, while all other vectors are column vectors.

http://dx.doi.org/10.1016/j.dsp.2014.07.002

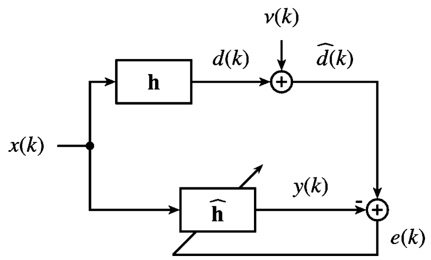

Fig. 1. Structure of a supervised adaptive identification algorithm.

avoiding, through the normalization, that eventually vectors $\mathbf{x}_k$ with high modulus amplify the noise in the adaptive process. The update equation of the NLMS algorithm is

$$\widehat{\mathbf{h}}_{k+1} = \widehat{\mathbf{h}}_k + \beta \frac{\mathbf{x}_k e(k)}{\delta + \mathbf{x}_k^T \mathbf{x}_k}, \tag{4}$$

where δ is a constant slightly larger than zero to avoid divisions by zero and β is the step-size or learning factor.

Recently several adaptive algorithms have been proposed specifically for sparse response identification systems. Such algorithms aim at overcoming the slow convergence of the NLMS in sparse configurations [1,12] through an uneven distribution of updating energy over the coefficients of $\widehat{\mathbf{h}}_k$, with larger learning factors assigned to the coefficients of greater magnitude. This strategy can be interpreted as a cooperation established by a central resource administrator, which gives more prominent updates for the coefficients of greater magnitude.

In this context, the PNLMS algorithm increases the factor β corresponding to $\widehat{h}_k(n)$ proportionally to its magnitude. The elements of $\widehat{\mathbf{h}}_k$ that are farther away from zero will have larger updates than those of smaller magnitude. The algorithm also performs a regularization for small amplitude signals [4].

2. Adaptive algorithms for the identification of sparse responses

In this section we briefly present some of the main contributions in adaptive identification of sparse impulse responses.

It is very common to use the NLMS for the adaptation of high-order adaptive filters, such as in echo cancellation [4]. One of the first proposed alternatives consisted of using filters with fewer adaptive coefficients than the length of the impulse response, by only updating the subsets of coefficients that corresponded to the dispersive regions [13,14]. One of the great advantages of this strategy lies in the substantial reduction in computational cost.

Another possibility of accelerating the adaptation convergence in the context of sparse impulse responses, when the NLMS has a sub-optimal performance, is the distribution of the learning factor β through the coefficients, as explained above. The first such proposal was the PNLMS (Proportionate Normalized Least-Mean-Squares) algorithm, derived for echo cancellation [4]. All the algorithms studied in this paper are derived from the PNLMS, and for that reason we say that they belong to the family of the PNLMS algorithms.

For sparse systems, the PNLMS algorithm presents faster initial convergence than does the NLMS. However, the convergence rate is dramatically reduced after the initial period and is slower than that of the NLMS for non-sparse impulse responses [5]. For this reason, the PNLMS++ algorithm [15] adopts switching between PNLMS and NLMS algorithms in order to reduce this degradation in non-sparse configurations. In [16], an approximation of the optimal step-size control factors is proposed in order to circumvent this drawback of the PNLMS algorithm. Instead of adjusting the adaptation step-size proportionally to the magnitude of the estimated filter coefficient, the resulting algorithm employs the logarithm of these magnitudes. In order to reduce the computational cost, the logarithmic function is approximated by a piecewise linear function, leading to the μ-law proportionate NLMS (MPNLMS) algorithm.

Table 1
IMPNLMS algorithm.

Initialization (typical values)
$\delta = 0.01,\ \epsilon = 0.001,\ \beta = 0.25, \lambda = 0.1$
$\xi(-1) = 0.96$
$\widehat{\mathbf{h}}_0 = [\,\widehat{h}_0(0)\ \widehat{h}_0(1)\ \cdots\ \widehat{h}_0(L-1)\,] = \mathbf{0}$
Processing and adaptation
For $\quad k = 0, 1, 2, \cdots$
$\quad \mathbf{x}_k = [\,x(k)\ x(k-1)\ \cdots\ x(k-L+1)\,]^T$
$\quad y(k) = \widehat{\mathbf{h}}_k \mathbf{x}_k$
$\quad e(k) = \widehat{d}(k) - y(k)$
$\quad \xi_{\widehat{\mathbf{h}}_k} = \frac{L}{L-\sqrt{L}}\left(1 - \frac{\sum_{j=0}^{L-1} \lvert\widehat{h}_k(j)\rvert}{\sqrt{L \sum_{j=0}^{L-1} \widehat{h}_k^2(j)}}\right)$
$\quad \xi(k) = (1-\lambda)\xi(k-1) + \lambda \xi_{\widehat{\mathbf{h}}_k}$
$\quad \alpha(k) = 2\xi(k) - 1$
$\quad$ For $\quad i = 0, 1, \cdots, L-1$
$\quad\quad g_k(i) = \frac{1-\alpha(k)}{2L} + \frac{(1+\alpha(k))F(\lvert\widehat{h}_k(i)\rvert)}{2\sum_{j=0}^{L-1} F(\lvert\widehat{h}_k(j)\rvert)+\epsilon}$
$\quad$ End For
$\quad \mathbf{\Gamma}_k = \mathrm{diag}\{g_k(0), \cdots, g_k(L-1)\}$
$\quad \widehat{\mathbf{h}}_{k+1} = \widehat{\mathbf{h}}_k + \beta \frac{\mathbf{x}_k^T \mathbf{\Gamma}_k e(k)}{\mathbf{x}_k^T \mathbf{\Gamma}_k \mathbf{x}_k + \delta}$
End For

The above approaches have the disadvantage of requiring sparseness of the impulse response to be identified for fast convergence, which is not always the case. In [5] such problem is mitigated by employing a measure of the sparseness of the system impulse response, resulting in the improved MPNLMS (IMPNLMS) algorithm shown in Table 1. The function $\xi_{\widehat{\mathbf{h}}_k}$ estimates the degree of sparseness based on the available estimated impulse response at each iteration. Such function assumes values in the interval $[0, 1]$, approaching 1 when the impulse response is sparse and 0 when it is dispersive. The conversion of $\xi_{\widehat{\mathbf{h}}}$ to the domain of the parameter $\alpha(k)$ was arbitrated by simulations [5]. The piecewise linear function

$$F\left(\lvert\widehat{h}_k(n)\rvert\right) = \begin{cases} 400\lvert\widehat{h}_k(n)\rvert, & \lvert\widehat{h}_k(n)\rvert < 0.005 \\ 8.51\lvert\widehat{h}_k(n)\rvert + 1.96, & \text{otherwise} \end{cases}, \tag{5}$$

which approximates the logarithmic function [16], is adopted in the update of the step-size control factors $g_k(i)$.

Among others, alternative strategies (not explored in this paper) for the identification of sparse responses consist of using an approximation of l_0-norm or l_1-norm of the weight vector to obtain a more accurate sparseness measure [17–20] and the use of Krylov subspace [6].

3. Transient analysis of the IMPNLMS algorithm

The theoretical estimation of the mean square error convergence of an adaptive algorithm eliminates the need of Monte Carlo averaging, among other advantages already well acknowledged in the literature. In this section, we derive recursive equations that describe in a reasonably accurate form the evolution of the MSE along the iterations.

In all analyses of PNLMS-type algorithms found in the literature, it is assumed that the input signal is white [7,8,21]. The violation of this hypothesis makes the algorithm convergence much slower, which disagrees with the analytical results. Therefore, in the benefit of generality, our analysis imposes no constraint on the input signal. We focus on the analysis of the IMPNLMS.

The equations of interest here are (see Table 1):

$$\xi_{\widehat{\mathbf{h}}_k} = \frac{L}{L-\sqrt{L}}\left(1 - \frac{\sum_{j=0}^{L-1} \lvert\widehat{h}_k(j)\rvert}{\sqrt{L \sum_{j=0}^{L-1} \widehat{h}_k^2(j)}}\right), \tag{6}$$