Signal Processing 142 (2018) 27–35

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Stochastic analysis of soft limiters in the LMS algorithm for stationary white Gaussian inputs—A unified theory[☆]

Neil J. Bershad[a,*], Eweda Eweda[b], Jose C.M. Bermudez[c]

[a] Department of Electrical Engineering and Computer Science, University of California, Irvine, CA 92660, USA
[b] C4 Advanced Solutions, Abu Dhabi, United Arab Emirates
[c] Department of Electrical and Electronic Engineering, Federal University of Santa Catarina, Trindade, Florianópolis, SC, 88040-970, Brazil.

ARTICLE INFO

Article history:
Received 24 April 2017
Revised 1 June 2017
Accepted 6 June 2017
Available online 24 June 2017

Keywords:
Adaptive filters
Analysis
LMS type algorithms
Stochastic algorithms

ABSTRACT

The effects of saturation-type nonlinearities on the input and the error in the weight update equation for LMS adaptation are investigated for a stationary white Gaussian data model for system identification. Nonlinear recursions are derived for the transient and steady-state weight first and second moments that include the effect of soft limiters on both the input and the error driving the algorithm. By varying a single parameter of the soft limiter, a general theory is presented that is applicable to LMS, soft limiting of the input, error or both and sign–sign LMS.

1. Introduction

The late 1980s and early 1990s saw a great interest in the behavior of the LMS algorithm implemented with physical hardware that introduced nonlinearities into the algorithm when configured as an echo canceller [1-21]. Many of these analyses were based on approximating the statistics of the weight update recursion nonlinearities as conditional Gaussian expectations, conditioned on the present weight error vector. The nonlinearities were constrained to be odd symmetric, and the data sequence was white. Furthermore, other constraints were placed on the statistics of the data sequence and the echo sequence (independent of each other, the echo sequence being zero mean white Gaussian).

The effect of an arbitrary nonlinear operation on the data input to the weight update equation was investigated in [13] for a Gaussian input model. The effect of a soft limiter, modeled as a scaled error function, was studied in [18]. The effect of hard limiters upon the input, error and input-error product was studied in [19-21,45]. However, by varying one parameter of the soft limiter, a general theory is presented here that is applicable to LMS, soft limiters of the input, error or both and sign–sign LMS. The unified theory allows one to simply and precisely study the tradeoff effects of the degree of saturation and the algorithm step-size for the modified LMS algorithm.

More recently approaches based on energy conservation arguments have been employed for deriving recursions for the first and second moment behavior of adaptive algorithms with error and data nonlinearities [22-27]. The case of data nonlinearities was analyzed in [24] for nonlinearities that can be modeled by the product of a full rank data nonlinearity matrix and the input regressor. The case of error nonlinearities was studied in [25,26]. These analyses were then extended in [27] to consider both error and data nonlinearities, the latter still limited to the model used in [24]. This model is not amenable to the study of adaptive filters with soft limiters applied to the data. These results required the assumption of long filters and a Gaussian assumption on the a priori error.

A number of recent papers [28-30] have also investigated this problem from a control systems viewpoint for non-Gaussian additive noise. The non-Gaussian noise is modeled as a two component Gaussian mixture [31-34]. This type of additive non-Gaussian noise can be studied within the framework of this paper by considering conditional expectations for each component separately as was done in [35].

To the best of our knowledge, the case of soft limiters on both the error and the input has not been analyzed previously in the literature. Thus, the unified theory yields an important new result by itself. The use of simultaneous input and error saturation (sign–sign LMS) is suggested in the 32 kbit/s ADPCM speech coder in the ITU-T recommendation G.726 [36—p. 10]. Furthermore,

☆ This work was supported in part by CNPq under grant No. 307071/2013-8.
* Corresponding author.
E-mail addresses: bershad@ece.uci.edu (N.J. Bershad), eweda@ieee.org (E. Eweda), bermudez.jc@gmail.com (J.C.M. Bermudez).

http://dx.doi.org/10.1016/j.sigpro.2017.06.005
0165-1684/

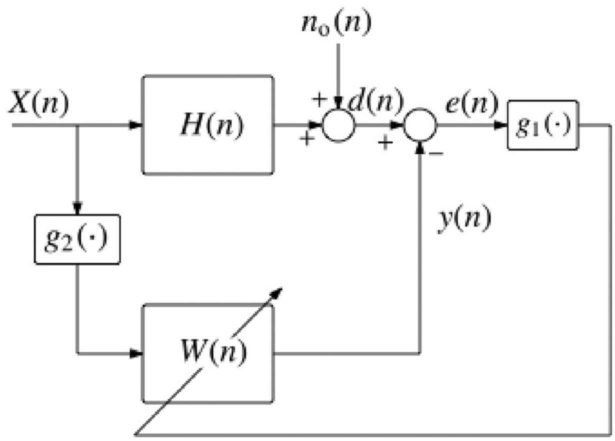

Fig. 1. Adaptive plant identification.

these types of applications are discussed in coding texts such as [37].

The paper is organized as follows. Section 2 defines the problem and the statistical assumptions used to solve the problem. Section 3 derives generic recursions for the mean and second moment behavior of the adaptive weights for arbitrary nonlinearities. These recursions involve unknown expectations of nonlinear functions of conditional Gaussian variates. This evaluation is one of the novelties of this paper. Section 4 specializes these recursions to the soft limiting of both the input and error for identifying a fixed channel. Section 5 extends these results to the identification of a Markov channel. Section 6 compares the developed theory with Monte Carlo simulations. Section 7 presents the conclusions. Capital letters denote vectors or matrices, and small letters denote scalar variables.

2. Problem definition and statistical assumptions

2.1. System identification and the Markov channel model

This paper will study the system identification model given in Fig. 1. All signals and systems are real. The N-dimensional input vector to the adaptive filter tap weights is given by $X(n) = [x(n), x(n-1), \ldots\ldots x(n-N+1)]^T$, where the superscript T means transpose.

Assumption 1. The observation noise $n_o(n)$ is zero-mean white Gaussian with variance σ_o^2 and independent of X(n).

Assumption 2. The sequence {x(n)} is also zero-mean white Gaussian and stationary with power σ_x^2.

The unknown channel is modeled as a linear time varying system whose impulse response is given by the standard random walk model, a particular case of a first-order Markov process [39—Section 14.1]

$$H(n+1) = H(n) + Q(n) \tag{1}$$

where Q(n) is a white Gaussian vector with zero mean and covariance matrix $E[Q(n)Q^T(n)] = \sigma_q^2 I$, where I is the identity matrix.

Assumption 3. The vector sequence Q(n) is independent of both X(n) and $n_o(n)$.

This model is the so-called random walk approximation to the first order Markov model [38]. The random walk model (1) is not realistic. However, it allows a feasible tracking analysis that provides important insights into the ability of the adaptive algorithm to track channel variations.

2.2. Independence theory and the performance measure

Assumption 4. The adaptive filter weights at time n, W(n), are statistically independent of the input vector X(n) [38].

This assumption is denoted the Independence Theory (IT) of adaptive filtering. The use of this assumption considerably simplifies the stochastic analysis of the adaptive filter. The IT assumption has been shown to lead to very accurate models in a wide variety of adaptive filter applications.

Define the weight deviation vector $V(n) = W(n) - H(n)$ and the weight deviation covariance matrix $K_{VV}(n) = E[V(n)V^T(n)]$. Then the mean square deviation (MSD) is given by [39]

$$MSD(n) = E\left[V(n)^T V(n)\right] = \text{trace}[K_{VV}(n)] \tag{2}$$

where trace[B] is the trace of the matrix B. The IT assumption is needed when evaluating the recursions for the mean weight and $K_{VV}(n)$ as will be shown shortly.

2.3. LMS algorithm

The algorithm for changing the weights of the LMS adaptive filter is given by

$$W(n+1) = W(n) + \mu e(n) X(n) \tag{3}$$

where

$$e(n) = H^T(n)X(n) + n_o(n) - W^T(n)X(n) \tag{4}$$

and μ is the step-size.

2.4. Nonlinear LMS algorithm

The algorithm for changing the weights of the nonlinear LMS adaptive algorithm studied here first is given by

$$W(n+1) = W(n) + \mu g_1[e(n)] G_2[X(n)] \tag{5}$$

where $G_2^T[X(n)] = [g_2[x(n)], g_2[x(n-1)], \ldots\ldots g_2[x(n-N+1)]]$, and $g_1[\cdot]$, and $g_2[\cdot]$ are bounded nonlinear odd functions.

3. Stochastic recursions for the mean and second moment behavior of the adaptive weights

For simplicity, the system identification problem is first analyzed for the fixed channel with $H(n) = W_o$. It is easy to extend these results to the Markov channel. Subtracting (1) from (5) yields

$$V(n+1) = V(n) + \mu\, g_1[e(n)] G_2[X(n)] \tag{6}$$

where $e(n) = d(n) - W^T(n)X(n) = n_o(n) - V^T(n)X(n)$, and W_o is the weight vector of the unknown system. Using IT,

$$E\left\{e^2(n)\right\} = \sigma_o^2 + \sigma_x^2\, \text{trace}\,[K_{VV}(n)]. \tag{7}$$

Averaging the ith weight error yields

$$E\{v_i(n+1)\} = E\{v_i(n)\} + \mu\, E\{g_1[e(n)] g_2[x(n-i+1)]\} \tag{8}$$

for i = 1 to N and $v_i(n)$ is the ith component of V(n).

Post-multiplying (6) by its transpose and averaging yields

$$\begin{aligned} K_{VV}(n+1) = {} & K_{VV}(n) + \mu E\left[g_1[e(n)] G_2[X(n)] V^T(n)\right] \\ & + \mu E\left[g_1[e(n)] V(n) G_2^T[X(n)]\right] \\ & + \mu^2 E\left\{g_1^2[e(n)] G_2[X(n)] G_2^T[X(n)]\right\} \end{aligned} \tag{9}$$

The recursions for the diagonal terms are

$$\begin{aligned} K_{VV}(n+1)_{i,i} = {} & K_{VV}(n)_{i,i} + 2\mu\{E[g_1[e(n)] g_2[x(n-i+1)] v_i(n)]\} \\ & + \mu^2 E\left\{g_1^2[e(n)] g_2^2[x(n-i+1)]\right\} \end{aligned} \tag{10}$$