A new gradient based variable step-size LMS algorithm

This item was submitted to Loughborough University’s Institutional Repository by the/an author.


Additional Information:

• This is a conference paper [© IEEE]. It is also available at: http://ieeexplore.ieee.org/ Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Metadata Record: https://dspace.lboro.ac.uk/2134/5613

Version: Published

Publisher: © IEEE

Please cite the published version.
This item was submitted to Loughborough’s Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/
A NEW GRADIENT BASED VARIABLE STEP-SIZE LMS ALGORITHM

Yonggang Zhang               Ning Li               Jonathon A. Chambers
School of Engineering, Cardiff University, Cardiff UK CF24 3AA
Email: zhangy15@cf.ac.uk

Abstract

A new gradient-based variable step-size LMS algorithm (VSSLMS) is proposed in this paper. The step size of the proposed algorithm is proportional to the squared norm of the smoothed gradient vector, which is a proximity measure of the adaptive process. In comparison with the existing methods, the proposed VSSLMS algorithm has improved convergence properties. Furthermore, the parameter choice of the proposed algorithm can be easily determined according to the analysis. Simulation results show the good properties of the proposed algorithm and support the performance analysis.

1. Introduction

The LMS algorithm has been extensively used in many applications as a consequence of its simplicity and robustness [1][2]. Many VSSLMS algorithms have been proposed to improve the performance of the LMS algorithm. An important class of VSSLMS algorithms is the one in which the step size is updated using the gradient vector [3][4][5][6][7]. In our opinion, all of these algorithms utilize two properties of the gradient vector:

1. The norm of the gradient vector will be large initially and converge to a small value, potentially zero, at steady state.
2. The polarity of the gradient vector will generally be consistent during the early stages of the adaptive process and change frequently after the system converges.

Methods proposed in [4][5] utilize property 1, whereas that proposed in [3] utilizes property 2. Techniques introduced in [6][7] utilize both of the properties. Motivated by the second property of the gradient vector described above, we propose a new VSSLMS algorithm. In this algorithm, the gradient vector is smoothed by using a first order filter to reduce the disturbance of the noise signal. Then the step size is controlled to be proportional to the squared norm of the smoothed gradient vector. Since Mathews' method in [6] provides an important theoretical support of all gradient based methods, we compare the proposed algorithm with this method and Wee-Peng's method [7] which is a modified version of Mathews' method by simulations.

The remainder of the paper is organized as follows: in the next section, we will introduce the proposed VSSLMS algorithm. In Section 3 an approximate analysis for the proposed algorithm is presented. In Section 4 two simulations are given to confirm the analysis and the good properties of the proposed algorithm. Section 5 summarizes this paper.

2. Algorithm formulation

For the convenience of description, we formulate the LMS algorithm within the context of system identification model. In this case the desired signal \(d(n)\) can be formulated as follows:

\[
d(n) = x^T(n)w_{opt} + t(n)
\]

where \(w_{opt}\) is the unknown filter which we wish to find, \(x(n)\) is the adaptive filter input vector, \(t(n)\) is the noise signal, \(n\) denotes the discrete time index and \((\cdot)^T\) denotes the vector transpose operator. The output error of the system \(e(n)\) is the difference between the desired signal and the output of the adaptive filter:

\[
e(n) = d(n) - x^T(n)w(n)
\]

where \(w(n)\) is the vector of the adaptive filter weights. The proposed VSSLMS algorithm can be formulated as follows:

\[
\bar{g}(n) = \beta\bar{g}(n-1) + (1 - \beta)e(n)x(n)
\]

\[
\mu_{VSSLMS}(n) = P\|\bar{g}(n)\|_2^2
\]
\[
\mathbf{w}(n+1) = \mathbf{w}(n) + \mu_{\text{VSSLMS}}(n) e(n) x(n)
\] (5)

where \(\mathbf{g}(n)\) is the smoothed gradient vector, \(\beta\) is the smoothing parameter which is set to be very close to unity to apply sufficient time smoothing, \(P\) is a positive constant and can be chosen easily according to the analysis in the next section, \(\|\|_2^2\) denotes the Euclidian distance, and \(\mu_{\text{VSSLMS}}(n)\) is the time varying step size.

The motivation of the proposed algorithm is as follows: To develop a VSSLMS algorithm, the most important thing is to measure the proximity of the adaptive process to the desired solution. An ideal measure of the adaptive process is the mean square deviation (MSD), which can be defined as \(\frac{1}{2} e^2(n)\). According to the formulation in [8], with a stationary input signal, the squared norm of the smoothed gradient vector, which is formulated in (3), can track the variation of the MSD, thus it is a good measure of the proximity of the adaptive process, and suitable to control the step size. As will be shown by our simulations, the proposed algorithm performs well in both low and high noise conditions.

Next we will perform a steady state analysis for the proposed algorithm.

3. Algorithm analysis

In this section we will give an approximate performance analysis for the proposed VSSLMS algorithm. For the convenience of analysis we make several assumptions:

\(A1\). The input signal \(x(n)\) is a zero-mean white stationary Gaussian signal. The noise signal \(t(n)\) is a zero-mean stationary signal and independent with the input signal \(x(n)\).

\(A2\). At steady state the excess error is much smaller compared with the noise signal, and therefore the error signal \(e(n)\) is approximately equal to the noise signal \(t(n)\).

Also throughout all analysis we assume that the length of the adaptive filter \(\mathbf{w}(n)\) is L-point and equal to that of the optimal filter \(\mathbf{w}_{\text{opt}}\). Assumption \(A1\) is a general assumption for the analysis of the VSSLMS algorithm [6]. Assumption \(A2\) is only true when the step size is very small. Using these assumptions gives insight into the algorithm and provides a guide for the parameter choice of the algorithm.

Since the squared norm of the smoothed gradient vector \(\|\mathbf{g}(n)\|_2^2\) is the key term for the proposed algorithm, we will give a steady state performance analysis for this term first. From (3) we have

\[
\mathbf{g}(n) = (1-\beta) \sum_{i=1}^{n} \beta^{n-i} \mathbf{g}(i)
\]

(6)

assuming \(\mathbf{g}(0) = 0\) and denoting \(\mathbf{g}(i) = e(i)x(i)\). The expected performance of the squared norm of the smoothed gradient vector can then be obtained

\[
E\{\|\mathbf{g}(n)\|_2^2\} = (1-\beta)^2 \sum_{i=1}^{n} \sum_{j=1}^{n} C(ij)
\]

(7)

where \(C(ij)\) is defined as

\[
C(ij) = E\{\beta^{n-i} \mathbf{g}^T(i) \beta^{n-j} \mathbf{g}(j)\}
\]

(8)

When \(n\) approaches infinity, the term \(C(ij)\) in (8) approaches to zero if \(i\) or \(j\) is finite. So when we calculate \(E\{\|\mathbf{g}(n)\|_2^2\}\), the term \(C(ij)\) can be ignored when \(i\) or \(j\) is not infinite. The following analysis will only consider this term at steady state, i.e., \(i\) and \(j\) are both steady state time indexes.

At first we consider \(C(ij)\) when \(i = j\). From assumption \(A2\) we have

\[
e(i) \approx t(i)
\]

(9)

With (9) the gradient vector \(\mathbf{g}(i)\) can also be approximately written as

\[
\mathbf{g}(i) \approx t(i)x(i)
\]

(10)

Substituting this formulation into (8) we obtain

\[
C(ii) \approx E\{\beta^{2n-2i} X^T(i)x(i)t^2(i)\}
\]

(11)

With assumption \(A1\) equation (11) becomes

\[
C(ii) \approx \beta^{2n-2i} L\sigma_n^2 \sigma_t^2
\]

(12)

where \(\sigma_n^2\) and \(\sigma_t^2\) are the variance of the noise signal and the input signal respectively.

When \(i \neq j\) similar derivation can be performed which yields

\[
C(ij) \approx 0 \quad (i \neq j)
\]

(13)

Substituting (12) and (13) into (7) we have

\[
E\{\|\mathbf{g}(n)\|_2^2\} \approx \frac{(1-\beta)}{(1+\beta)} L\sigma_n^2 \sigma_t^2
\]

(14)

With equation (14) the averaged steady state step size value can then be obtained from (4)
As described by equation (16) in [9], the steady state excess mean square error (EMSE) of the LMS algorithm can also be formulated as

$$J_{ex,\text{LMS}}(\infty) = \frac{\mu_{\text{LMS}} L \sigma_x^2 \sigma_t^2}{2 - \mu_{\text{LMS}} L \sigma_x^2}$$  \hspace{1cm} (16)

If we assume that at steady state the step size of the proposed algorithm is very small, and $\mu_{\text{VSSLMS}}(\infty)L\sigma_x^2 << 2$, the EMSE of the propose algorithm can then be formulated as

$$J_{ex,\text{VSSLMS}}(\infty) \approx \frac{1}{2} E\{\mu_{\text{VSSLMS}}(\infty)\} L \sigma_x^2 \sigma_t^2$$  \hspace{1cm} (17)

Substituting (15) into (17) we obtain the steady state EMSE for the proposed VSSLMS algorithm:

$$J_{ex,\text{VSSLMS}}(\infty) \approx \frac{P(1-\beta)L \sigma_x^2 \sigma_t^2}{2(1+\beta)}$$  \hspace{1cm} (18)

Now let’s consider the choice of the parameter $P$. To choose this parameter, we first need to determine a desired $J_{ex,\text{VSSLMS}}(\infty)$ according to the application. With this desired $J_{ex,\text{VSSLMS}}(\infty)$, $P$ can be determined directly according to (18):

$$P \approx \frac{2 J_{ex,\text{VSSLMS}}(\infty)(1+\beta)}{(1-\beta)L \sigma_x^4 \sigma_t^4}$$  \hspace{1cm} (19)

Next we will perform two simulations to confirm the analysis and show the advantages of the proposed algorithm.

4. Simulations

Two simulations will be performed in this section to demonstrate the advantages of the proposed algorithm. In both simulations the proposed algorithm will be compared with Mathews’ method and Wee-Peng’s method. The system identification model is assumed in both simulations.

The performance comparison is performed on the basis of the following performance measures: 1. The time evolution of the step size. 2. The EMSE which is defined as $E\{(e(n)-x(n))^2\}$. The results of both simulations are obtained by 100 Monte Carlo trials.

In the first simulation, we consider identifying a five-point FIR filter $w_{opt} = [0.1 \ 0.3 \ 0.5 \ 0.3 \ 0.1]$ [6]. The input signal $x(n)$ is a pseudo random, zero-mean, unit variance Gaussian process. The noise signal is also a pseudo random, zero-mean, Gaussian process and scaled to make the SNR 20dB. The initial step sizes and adaptive filter vectors of all algorithms are set to be zero. The parameter $\rho$ for Mathew’s method [6] is set to 0.0005. The smoothing parameter $\beta$ for both Wee-Peng’s method [7] and the proposed method is set to 0.99. The parameter $\rho$ for Wee-Peng’s method is set to 0.0002. The parameter $P$ for the proposed algorithm is set to 5. The evolutions of the step sizes and the evolutions of the EMSE curves are shown in Fig. 1.

![Fig. 1 Simulation results when SNR=20dB](image-url)
parameter $\rho$ for Wee-Peng’s method [7] is set to 0.000002. The parameter $P$ for the proposed algorithm is set to 0.05. The evolutions of the step sizes and the evolutions of the EMSE curves are shown in Fig. 2. The theoretical values are also plotted. Similar results are obtained. Both simulations confirm the analysis and show the advantages of the proposed algorithm. Since the theoretical values are very close to the simulation results, we can conclude that with a given desired steady state EMSE, the parameter $P$ can be chosen properly according to (19).

![Fig. 2 Simulation results when SNR=0dB](image)

**5. Conclusion**

In this paper we proposed a new gradient based VSSLMS algorithm. According to our analysis and simulations, the performance of this algorithm is better than classical existing methods. The control of the step size of the proposed algorithm is more reasonable for the adaptive process, and a fast convergence rate is obtained.

**References**