VOLUME-1 ISSUE-3
ISSN : 2799-18

INVESTIGATION ON LEAST-MEAN-SQUARE ALGORITHM

Mrs. P.S.Indrani
Malla Reddy Engineering College (Autonomous) Department of Electronics & Communication Engineering

ABSTRACT

The Non-Negative Least-Mean-Square (NNLMS) algorithm and its variants have been proposed for online estimation under non-negativity constraints. The transient behavior of the NNLMS, Normalized NNLMS, Exponential NNLMS and Sign-Sign NNLMS algorithms have been studied in the literature. In this letter, we derive closed-form expressions for the steady-state excess mean-square error (EMSE) for the four algorithms. Simulation results illustrate the accuracy of the theoretical results. This work complements the understanding of the behavior of these algorithms**.** This algorithm builds on a fixed-point iteration strategy driven by the Karush–Kuhn–Tucker conditions. It was shown to provide low variance estimates, but it however suffers from unbalanced convergence rates of these estimates. In this paper, we address this problem by introducing a variant of the NNLMS algorithm. We provide a theoretical analysis of its behavior in terms of transient learning curve, steady-state and tracking performance. We also introduce an extension of the algorithm for online sparse system identification. Monte-Carlo simulations are conducted to illustrate the performance of the algorithm and to validate the theoretical results.

1. INTRODUCTION

NON-NEGATIVITY is one important constraint that can be imposed on parameters to estimate. It is often imposed to avoid physically unreasonable solutions and to comply with natural physical characteristics. Non-negativity constraints appear, for example, in deconvolution problems [1]–[3], image processing [4], [5], audio processing [6], remote sensing [7]–[9], and neuroscience. The Non-Negative Least-Mean-Square algorithm (NNLMS) and its three variants, namely, Normalized NNLMS, Exponential NNLMS and Sign-Sign NNLMS were proposed to adaptively find solutions of a typical Wiener filtering problem under non-negativity constraints. The transient behavior of these algorithms has been studied. Analytical recursive models have been derived for the mean and mean-square behaviors of the adaptive weights.

This paper complements the work by deriving closed form expressions for the steady-state excess mean square error of each of these algorithms. These expressions cannot be directly obtained from the transient recursions derived in because the weight updates include nonlinearities on the adaptive weights. Moreover, they cannot be derived following the conventional energy-conservation relations.

Hence, new analyses are required to understand the steady-state behavior of these algorithms. In this paper, we derive accurate models for the steady-state behaviors of NNLMS and its variants using a common analysis framework, with clear physical interpretation of each term in the expressions. Simulations are conducted to validate the theoretical results. This work therefore complements the understanding of the behavior of these algorithms, and introduces a new methodology for the study of the steady-state performance of adaptive algorithms. We recommend that readers refer to for a more detailed understanding of the algorithms and their transient behavior. Readers may also refer to the associated report for some detailed calculation steps.

Online learning aims at determining a mapping from a dataset to the corresponding labels when

VOLUME-1 ISSUE-3
ISSN : 2799-18

the data are available in a sequential fashion. In particular, algorithms such as the Least-Mean-Square (LMS) and the Recursive Least-Square (RLS) algorithms minimize the mean square-error cost function in an online manner based on input/ output measurement sequences [1,2]. In practice, rather than leaving the parameters to be estimated totally free and relying on data, it is often desirable to introduce some constraints on the parameter space. These constraints are usually introduced to impose some specific structures, or to incorporate prior knowledge, so as to improve the estimation accuracy and the interpretability of results in learning systems [3,4]. The nonnegativity constraint is one of the most frequently used constraints among several popular ones [5]. It can be imposed to avoid physically unreasonable solutions and to comply with physical characteristics. For example, quantities such as intensities [6,7], chemical concentrations [8], and material fractions of abundance [9] must naturally fulfill nonnegativity constraints.

Nonnegativity constraints may also enhance the physical interpretability of some analyzed results. For instance, Nonnegative Matrix Factorization leads to more meaningful image decompositions than Principle Component Analysis (PCA). PCA and neural networks can also be conducted subject to nonnegativity constraints in order to enhance result interpretability.

Finally, there are important problems in signal processing that can be cast as optimization problems under nonnegativity constraints. Other applications of learning systems related to nonnegativity constraints can be found. Nonnegativity constrained problems can be solved in a batch mode via active set methods gradient projection methods and multiplicative methods to cite a few. Online system identification methods subject to nonnegativity constraints can also be of great interest in applications that require to adaptively identify a dynamic system. An LMS-type algorithm, called the non-negative least-mean-square (NNLMS) algorithm, was proposed to address the least-mean-square problem under nonnegativity constraints. It was derived based on a stochastic gradient descent approach combined with a fixed-point iteration strategy that ensures convergence toward a solution satisfying the Karush–Kuhn–Tucker (KKT) conditions. In several variants of the NNLMS were derived to improve its convergence behavior in some sense.

2. MOTIVATING FACTS AND THE ALGORITHM

2.1. Motivation

The weight update in (6) corresponds to the classical stochastic gradient LMS update with its ith component scaled by $\alpha i(n)$. The mean value of the update vector $\boldsymbol{D\alpha}(n)\boldsymbol{x}(n)e(n)$ is thus no longer in the direction of the gradient of $J(\boldsymbol{\alpha})$, as is the case for LMS. On the other hand, it is exactly this scaling by $\alpha i(n)$ that enables the corrections $\alpha i(n)xi(n)e(n)$ to reduce gradually to zero for coefficients $\alpha i(n)$ tending to zero, which leads to low-variance estimates for these coefficients.1 If a coefficient $\alpha k(n)$ that approaches zero turns negative due to the stochastic update, its negative sign will induce a change αk(n)xk(n)e(n) in the kth weight component that is contrary to what would indicate the stochastic gradient, and thus towards zero.

The presence of the factor αi(n) in the update αi(n)xi(n)e(n) of the ith coefficient leads to different convergence rates for coefficients of different values. This is particularly problematic for the coefficients in the active set as they approach zero. Because of the factor αi(n), the convergence of these coefficients eventually stalls due to insignificant correction along their axes. Though the algorithm leads to a very low steady-state error for these coefficients, this happens only after a long convergence process. In addition, the dispersion of coefficient values introduces difficulties for step size selection and coefficient initialization, since each estimated coefficient acts as a different directional gain for the same step size. In order to address these problems, it is of interest to derive a variant of the NNLMS algorithm that satisfies the following requirements:

VOLUME-1 ISSUE-3
ISSN : 2799-18

! The coefficients should converge to the fixed point satisfying (4), so that it still solves the nonnegativity constrained problem (2).

! The sensitivity of the algorithm (6) to the spread of the coefficient values at each iteration should be reduced, yielding more balanced convergence rates and steady-state weight errors than the original algorithm (6).

! The performance improvement should be achieved without introducing significant computational burden.

2.2. The inversely proportional NNLMS algorithm

The Exponential NNLMS replaces the gradient scaling αi(n) with αiγ(n) = sgn(αi(n))|αi(n)|γ, where γ = γ /γ 1 2 with γ1 and γ2 being two odd numbers such that γ > γ > 0 2 1 . This variant mitigates the aforementioned drawbacks of NNLMS to some extent, but introduces additional computational burden. In addition, the stability of the algorithm is still affected by the weight dynamics since αiγ(n) is unbounded.

$$f_i(\boldsymbol{\alpha}(n)) = f_i(n) = \frac{1}{|\alpha_i(n)| + \epsilon}$$

Computational complexity:

A comparison of the computational complexities in the implementation of the NNLMS, Exponential NNLMS and IP-NNLMS algorithms needs to consider only the weight update term, since this term is what distinguishes the three algorithms. We consider their real-time implementation using N coefficients and m-bit integer operations. Also, because there exists a variety of multiplication algorithms, let us denote by (m) the complexity of the chosen algorithm to multiply two m-bit integers. The NNLMS update (6) sets the weighting function wi(n) to αi(n), and has complexity. The Exponential NNLMS [27] sets f (α) = α γ− i I 1 in (5), which leads to w(n) = |α (n)|α γ(n) i i i . Evaluating αiγ(n) has complexity of □(□(m)log m) because it uses exponential and logarithm elementary functions. Then, IP-NNLMS addresses the two important NNLMS limitations described above without a significant increase in computational complexity. Fig. 1 presents the computation time required for the calculation of the weight updates of NNLMS, IP-NNLMS and Exponential NNLMS for 106 iterations on a laptop with Matlab as a function of the number of filter coefficients N. This experiment shows that complexity of these algorithms increases linearly with N, with a factor that is significantly larger for Exponential NNLMS.

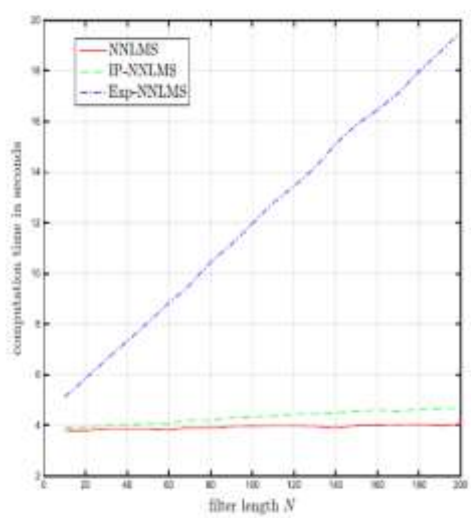

Fig. 2. Computation time in seconds of NNLMS, IP-NNLMS and Exponential NNLMS for 106 iterations as a function of the number of filter coefficients N.

3. EXPERIMENT VALIDATION

In this section, we present examples to illustrate the correspondence between theoretical steady-state EMSE and simulated results for NNLMS and its variants. Consider an unknown system of order and weights defined by

$$\boldsymbol{\alpha}^* = [0.8, 0.6, 0.5, -0.05, 0.4, -0.04, 0.3, -
0.2, -0.02, 0.1, -0.01, 0, 0, 0]^\top,$$

VOLUME-1 ISSUE-3
ISSN : 2799-18

where negative coefficients were explicitly included to activate the non-negativity constraint.

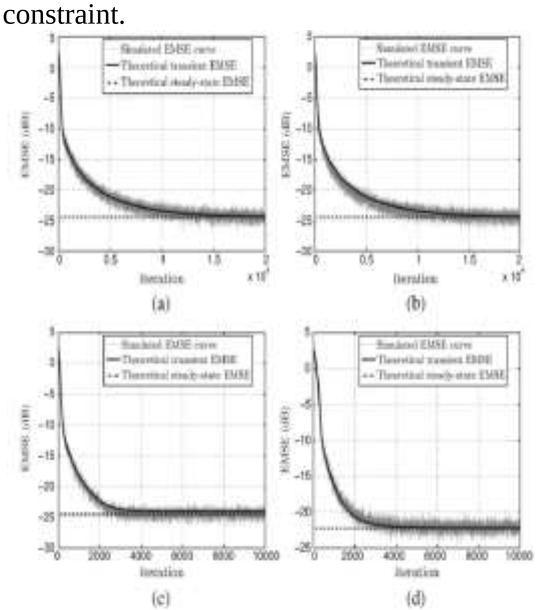

Fig. 1. Steady-state EMSE model validation for NNLMS and its variants (a) Original NNLMS (b) Normalized NNLMS (c) Exponential NNLMS (d) Sign-Sign NNLMS.

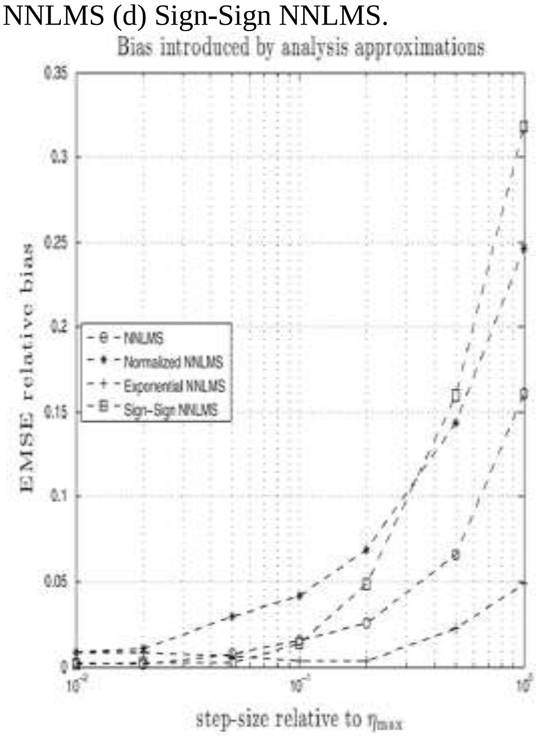

Fig. 2. Bias introduced by the assumptions made in the analysis. The bias is calculated as the relative difference of the EMSE obtained from simulations and predicted by the models

CONCLUSION

In this letter, we derived closed-form expressions for the steady-state excess mean-square errors of the Non-Negative LMS algorithm and its variants. Experiments illustrated the accuracy of the derived results. Future work may include the derivation of other useful variants of NNLMS and the study of their stochastic performance. Future work may include the derivation of other useful variants of NNLMS and the study of their stochastic performance.

REFERENCES

[1] M. D. Plumbley, "Algorithms for nonnegative independent component analysis," *IEEE Trans. Neural Netw.*, vol. 14, no. 3, pp. 534–543, Mar. 2003.

[2] S. Moussaoui, D. Brie, A. Mohammad-Djafari, and C. Carteret, "Separation of non-negative mixture of non-negative sources using a bayesian approach and MCMC sampling," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4133–4145, Nov. 2006.

[3] Y. Lin and D. D. Lee, "Bayesian regularization and nonnegative deconvolution for room impulse response estimation," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 839–847, Mar. 2006.

[4] F. Benvenuto, R. Zanella, L. Zanni, and M. Bertero, "Nonnegative least-squares image deblurring: Improved gradient projection approaches," *Inv. Probl.*, vol. 26, no. 1, p. 025004, Feb. 2010.

[5] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.

[6] A. Cont and S. Dubinov, "Realtime multiple pitch and multiple-instrument recognition for music signals using sparse non-negative constraints," in *Proc. 10th Int. Conf.Digital Audio Effects (DAFx-07)*, Bordeaux, France, Sep. 2007, pp. 85–92.

VOLUME-1 ISSUE-3
ISSN : 2799-18

[7] J. Chen, C. Richard, and P. Honeine, "Nonlinear unmixing of hyperspectral data based on a linear-mixture/nonlinear-fluctuation model," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 480–492, 2013.

[8] J. Chen, C. Richard, H. Lantéri, C. Theys, and P. Honeine, "A gradient based method for fully constrained least-squares unmixing of hyperspectral images," in *Proc. IEEE Statistical Signal Processing Workshop (SSP)*, Nice, France, Jun. 2011, pp. 301–304.

[9] P. Honeine and C. Richard, "Geometric unmixing of large hyperspectral images: A barycentric coordinate approach," *IEEE Trans. Geosci. Remote Sensing*, vol. 50, no. 6, pp. 2185–2195, 2012.

[10] A. Cichocki, R. Zdunek, and A. H. Phan, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*. Hoboken, NJ, USA: Wiley, 2009.