Signal Processing 152 (2018) 373–383

Contents lists available at ScienceDirect

Signal Processing

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

Model-driven online parameter adjustment for zero-attracting LMS☆

Danqi Jin[a], Jie Chen[a,*], Cédric Richard[b], Jingdong Chen[a]

[a] Centre of Intelligent Acoustics and Immersive Communications, School of Marine Science and Technology, Northwestern Polytechinical University, Xi'an, China

[b] Université de Nice - Sophia Antipolis, CNRS, France

ARTICLE INFO

Article history:
Received 11 September 2017
Revised 2 June 2018
Accepted 19 June 2018
Available online 20 June 2018

Keywords:
Sparse system identification
Transient behavior model
Variable parameter strategy
Adaptive algorithms
ZA-LMS
RZA-LMS
Complex-valued signal

ABSTRACT

Zero-attracting least-mean-square (ZA-LMS) algorithm has been widely used for online sparse system identification. Similarly to most adaptive filtering algorithms and sparsity-inducing regularization techniques, ZA-LMS appears to face a trade-off between convergence speed and steady-state performance, and between sparsity level and estimation bias. It is therefore important, but not trivial, to optimally set the algorithm parameters. To address this issue, a variable-parameter ZA-LMS algorithm is proposed in this paper, based on a model of the stochastic transient behavior of the ZA-LMS. By minimizing the excess mean-square error (EMSE) at each iteration on the basis of a white input assumption, we obtain closed-form expression of the step-size and regularization parameter. To improve the performance, we introduce the same strategy for the reweighted ZA-LMS (RZA-LMS). Simulation results illustrate the effectiveness of the proposed algorithms and highlight their performance through comparisons with state-of-the-art algorithms, in the case of white and correlated inputs.

1. Introduction

Adaptive filtering methods are powerful tools for online system identification [1,2]. Within the myriad of algorithms proposed in the literature, the least-mean-square (LMS) algorithm has been widely used since it is robust and provides reasonably good performance with low computational complexity. Several applications have recently shown the need for online sparse identification techniques. A driving force behind the development of such algorithms is, for instance, the channel estimation problem because, although the number of coefficients of the impulse response can be large, only a few of them may have significant values. It is therefore important to endow the conventional LMS algorithm with the ability to provide enhanced performance for such scenarios.

In recent years, several algorithms based on the LMS were proposed to promote the sparsity of the estimate. The proportionate normalized LMS (PNLMS) [3] and its variant called improved PNLMS (IPNLMS) [4] update each filter coefficient independently by adjusting the adaptation step-size in proportion to the estimated filter coefficient. Another family of sparsity-inducing algorithms is motivated by the compressive sensing theory, which provides a unified framework for estimating sparse signals [5,6]. In place of the ℓ_0-norm, which provides an exact count of the nonzero coefficients but leads to NP-hard optimization problems (nondeterministic polynomial-time solvable decision problems), other sparsity-inducing norms can be used as a surrogate to overcome this difficulty [7]. The use of the ℓ_1-norm is a popular choice [8]. For instance, the authors in [9] consider an ℓ_1-norm regularizer, and introduce the zero-attracting LMS and the reweighted zero-attracting LMS for sparse system identification. It is shown that the ZA-LMS and the RZA-LMS perform better than the LMS in sparse scenarios. However adjusting the algorithm parameters, including the step size and the regularization parameter, remains a tricky task. On the one hand, as for usual adaptive algorithms, the step-size plays a crucial role to control the trade-off between the convergence speed and the asymptotic performance. A small step-size leads to slower convergence but improved asymptotic performance, while a large step-size leads to faster convergence but at the cost of a higher power of the residual error, or even instability of the algorithm [1,2]. On the other hand, the regularization parameter controls the trade-off between the sparsity of the estimate and the estimation bias. A large regularization parameter associated with the ℓ_1-norm strongly promotes the sparsity of the solution. This however causes a larger bias of the non-zero parameter vector entries. Reweighted ℓ_1-regularization allows to reduce this bias. However, an improper value of the regularization parameter

☆ The work of Jie Chen was supported in part by NSFC grants 61671382 and 61811530283, and 111 project (B18041). The work of Cédric Richard was supported in part by ANR and in part by DGA under Grant ANR-13-ASTR-0030 (ODISSEE Project). The work of Jingdong Chen was supported in part by NSFC grant 61425005.

* Corresponding author.

E-mail addresses: danqijin@mail.nwpu.edu.cn (D. Jin), dr.jie.chen@ieee.org (J. Chen), cedric.richard@unice.fr (C. Richard), jingdongchen@ieee.org (J. Chen).

https://doi.org/10.1016/j.sigpro.2018.06.020
0165-1684/

may even worsen the estimation performance. Though techniques such as regularization path and cross validation help characterize the influence of this parameter [10], they are inappropriate for online learning settings.

Variable parameter strategies provide simple but efficient solutions for optimizing the trade-off between fast convergence and low misadjustment [11]. For LMS, several variable step-size strategies have been proposed in the literature to address this issue. In most cases, the step-size adapts over time depending on the estimation error. Related works include [11–13]. A variable step-size version of the PNLMS, called NPVSS-IPNLMS, is proposed in [11]. It combines the IPNLMS and a variable step-size NLMS (VSS-NLMS) strategy [14]. However, while achieving a lower misadjustment, the convergence speed of NPVSS-IPNLMS slows down significantly after an initial phase. The zero-attracting variable step-size LMS (ZA-VSSLMS) and the reweighted zero-attracting variable step-size LMS (RZA-VSSLMS) introduced in [12] use the variable step-size strategy reported in [15]. A significant improvement in the convergence rate as well as in the misadjustment error can be observed. Another variable step-size RZA-LMS strategy based on a nonlinear relationship between the step-size and the power of the noise-free prior error, called VSS-RZA-LMS, is considered in [13]. Nevertheless, the misadjustment improvement appears to be limited. It is worth noting that some extra parameters are introduced into all these algorithms, but setting their proper values is a nontrivial task, similar to the selection of an appropriate step size.

Motivated by our recent work [16], where a new model is derived for the transient behavior of the ZA-LMS algorithm, we propose in this paper to design a variable-parameter ZA-LMS (VP-ZA-LMS) algorithm where the step-size and the regularization parameter are both adjusted in an online manner. Unlike heuristic strategies considered in the literature, our method is based on an optimization step that minimizes the EMSE at each iteration. Indeed, it turns out to be a quadratic function of the step-size and the regularization parameter when considering the transient model in [16] under a white input assumption. This yields closed-form expressions of the step-size and regularization parameter at each iteration, leading to a faster convergence as well as a lower misadjustment. To further improve the performance, we apply this strategy to the RZA-LMS, leading to a variable-parameter RZA-LMS (VP-RZA-LMS) algorithm. Simulation results illustrate the enhanced performance of our algorithms compared with ZA-LMS, RZA-LMS and other variable step-size algorithms used in sparse system identification applications. We summarize the contributions of this work as follows:

1. Compared to the existing literatures, this work is the first one that derives a variable-parameter strategy based on a theoretical model of the filter performance. The proposed algorithm jointly adjust the step-size and regularization parameter in some optimal sense.
2. Unlike existing works on ZA-LMS that focus on the real-valued data case, we derive an extension to complex-valued systems.
3. While working well for ZA-LMS/RZA-LMS, the proposed framework can be extended to several other adaptive filters having similar structure, such as the LMS with ℓ_0-norm penalty, the group ZA-LMS, etc.

Before proceeding, note that this work and [16] are both related to the transient behavior of the ZA-LMS algorithm but they address different issues. The analysis in [16] focuses on how deriving an accurate model for the transient behavior of ZA-LMS. The current work uses an approximate model that allows us to automatically adjust the algorithm parameters in an online way.

The rest of this paper is organized as follows. Section 2 reviews the ZA-LMS and RZA-LMS algorithms. The VP-ZA-LMS and VP-RZA-LMS algorithms are derived in Sections 3 and 4, respectively. In Section 5, computer simulations are performed to validate the proposed algorithms and to show their superior performance. Section 6 concludes the paper.

Notation. Normal font x denotes scalars. Boldface small letters $\boldsymbol{x}$ denote column vectors. All vectors are column vectors. Boldface capital letters $\boldsymbol{X}$ denote matrices. The superscript $(\cdot)^\top$ denotes the transpose of a matrix or a vector. The inverse of a square matrix is denoted by $(\cdot)^{-1}$. All-zero vector and all-one vector of length N are denoted by $\boldsymbol{0}_N$ and $\boldsymbol{1}_N$, respectively. The Gaussian distribution with mean μ and variance σ^2 is denoted by $\mathcal{N}(\mu, \sigma^2)$. The operator $\text{sgn}\{\cdot\}$ takes the sign of the entries of its argument. The operator $\text{tr}\{\cdot\}$ takes the trace of its matrix argument. The operator $|\cdot|$ takes the absolute value of the entries of its argument. The mathematical expectation is denoted by $\mathbb{E}\{\cdot\}$. The operators $\max\{\cdot, \cdot\}$ and $\min\{\cdot, \cdot\}$ take the maximum and minimum value of their arguments, respectively.

2. System model and zero-attracting LMS

2.1. System model and zero-attracting LMS

To be consistent with ZA-LMS/RZA-LMS framework, and for the sake of simplicity, we start by deriving our parameter adjustment strategies in the case of real-valued signals. In Appendix C, we extend ZA-LMS and RZA-LMS to complex-valued data, and then derive the associated parameter adjustment strategies in a concise manner. Consider an unknown system with input-output relation characterized by the linear model

$$y_n = \boldsymbol{x}_n^\top \boldsymbol{w}^\star + z_n \tag{1}$$

with $\boldsymbol{w}^\star \in \mathbb{R}^L$ denoting an unknown parameter vector, and $\boldsymbol{x}_n \in \mathbb{R}^L$ a regression vector with a positive definite covariance matrix $\boldsymbol{R}_x = \mathbb{E}\{\boldsymbol{x}_n \boldsymbol{x}_n^\top\} > 0$ at instant n. The regression vector $\boldsymbol{x}_n$ and the output signal y_n are assumed to be zero mean. The error signal z_n is assumed to be stationary, independent and identically distributed (i.i.d.), with zero mean and variance σ_z^2, and independent of any other signal. Let $J(\boldsymbol{w})$ denote the mean-square-error (MSE) cost, namely,

$$J(\boldsymbol{w}) = \frac{1}{2}\,\mathbb{E}\left\{[y_n - \boldsymbol{w}^\top \boldsymbol{x}_n]^2\right\}. \tag{2}$$

It is clear from (1) that $J(\boldsymbol{w})$ is minimized at $\boldsymbol{w}^\star$.

The problem considered in this paper is to estimate the unknown parameter vector $\boldsymbol{w}^\star$, which is assumed to be sparse [3,17,18]. This problem can be addressed by minimizing the following regularized MSE cost:

$$\begin{aligned} \boldsymbol{w}_{\text{ZA}}^o &= \arg\min_{\boldsymbol{w}} J_{\text{ZA}}(\boldsymbol{w}) \\ &\text{with } J_{\text{ZA}}(\boldsymbol{w}) = \frac{1}{2}\,\mathbb{E}\left\{[y_n - \boldsymbol{w}^\top \boldsymbol{x}_n]^2\right\} + \lambda \|\boldsymbol{w}\|_1, \end{aligned} \tag{3}$$

where the ℓ_1-norm term, defined as $\|\boldsymbol{w}\|_1 = \sum_{i=1}^L |w_i|$, is used to promote the sparsity of the estimate, and $\lambda \geq 0$ is the regularization parameter. A subgradient of $J_{\text{ZA}}(\boldsymbol{w})$ in problem (3) is given by:

$$\partial J_{\text{ZA}}(\boldsymbol{w}) = \boldsymbol{R}_x \boldsymbol{w} - \boldsymbol{p}_{xy} + \lambda\,\text{sgn}\{\boldsymbol{w}\} \tag{4}$$

where $\boldsymbol{p}_{xy} = \mathbb{E}\{\boldsymbol{x}_n y_n\}$ is the correlation vector between $\boldsymbol{x}_n$ and y_n. Using the instantaneous approximations $\boldsymbol{R}_x \approx \boldsymbol{x}_n \boldsymbol{x}_n^\top$ and $\boldsymbol{p}_{xy} \approx \boldsymbol{x}_n y_n$, the subgradient iteration leads to the ZA-LMS algorithm as derived in [9]:

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \mu\, e_n \boldsymbol{x}_n - \rho\,\text{sgn}\{\boldsymbol{w}_n\}, \tag{5}$$

where e_n is the estimation error given by:

$$e_n = y_n - \boldsymbol{w}_n^\top \boldsymbol{x}_n, \tag{6}$$

μ is a positive step-size, and $\rho = \mu\lambda$.