Signal Processing 144 (2018) 201–213

Contents lists available at ScienceDirect

Signal Processing

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

Boxed-constraint least mean square algorithm and its performance analysis

Wenyuan Wang, Haiquan Zhao*

Key Laboratory of Magnetic Suspension Technology and Maglev Vehicle, Ministry of Education, and the School of Electrical Engineering, Southwest Jiaotong University, Chengdu, 610031, China

ARTICLE INFO

Article history:
Received 3 March 2017
Revised 9 September 2017
Accepted 4 October 2017
Available online 13 October 2017

Keywords:
Boxed-constraint
Karush-Kuhn-Tucker condition
Fixed-point iteration
Least mean square

ABSTRACT

In this paper, a novel adaptive filter algorithm, called boxed-constraint least mean square (BXCLMS) algorithm, is proposed for identifying the boxed-constrained system where the parameter to estimate is limited in a range from lower bound to upper bound. The proposed algorithm is derived by using the Karush-Kuhn-Tucker (KKT) conditions and fixed-point iteration algorithm. In addition, the stochastic behavior analysis of proposed algorithm is performed in terms of mean and mean square performance. Finally, simulations are carried out to demonstrate the performance of BXCLMS algorithm and verify the correctness of the analytical results.

1. Introduction

Online system identification is widely used in lots of the practical applications over the past decades [1,2]. Its object is to seek a mapping between a dataset and its corresponding labels. As the most common used algorithms in online system identification, the recursive least square (RLS) and the least mean square (LMS) algorithms are derived by minimizing the online object cost function which is collected from input or output measurement data [2]. In practical application, the parameters to be estimated are usually forced some constraint due to the inherent physical feature or other reason [3]. As a popular constraint, the nonnegativity constraint has been attracted much attention in refraining from physically unreasonable solutions, such as chemometrician's toolbox [4], material fractions of abundance [5] and so on. To conserve intrinsic feature of solutions corresponding to chemical concentrations, pixel intensities and other applications of science and engineering, respecting the nonnegativity is significative, because it can prevent absurd and unexpected consequents [6].

As one of the most frequently used constraint, the nonnegativity constraint has been studied in references [7–17]. Online system identification method, which is under nonnegativity constraints, has been also applied to adaptively identify the system which is dynamic. Specifically, the nonnegative LMS (NNLMS) algorithm has been presented by Chen in [18] to solve the LMS optimal issue which is subject to nonnegativity constraints. The NNLMS algorithm is obtained by using a gradient descent method and a fixed-point iteration algorithm. In addition, the Karush-Kuhn-Tucker (KKT) conditions [19] are also used to derive this algorithm. In [20–22], the variants of NNLMS algorithm were proposed to enhance the performance.

Generally, the inequality constraint is also usually applied to the signal processing and communications [23–27]. For example, in the modern sampling theory, constrained reconstruction solves the following optimal problem [27]

$$x_{CLS} = \arg\min_{x\in\mathcal{G}} \|Sx - c\|^2 \tag{1}$$

where $\mathcal{G} = \{x : \|Lx\| \leq \rho, x \in \mathcal{W}\}$. Particularly, the boxed constraint which means the parameters are subject to a certain boxed range is widely used in estimation problems of practical importance. In the real-time actuator optimization, the forces f should be limited in the boxed constraint range [23]. In optimal power generation and distribution, generator powers g_i and line power flows p_i should also be limited in the boxed constraint ranges [23]. In [28], the authors studied overlapping community detection through bounded nonnegative matrix tri-factorization. Moreover, the boxed-constraint Kalman filtering have been proposed and used in real application [33]. Dual MIMU pedestrian navigation are proposed by using inequality constraint Kalman filtering [34]. In [36], an improved extended Kalman filter with inequality constraints for gas turbine engine health monitoring was proposed. In [31], Li et al. proposed an auxiliary particle filtering algorithm with inequality constraints. Truncation nonlinear filters for state estimation with nonlinear inequality constraints were proposed by [32].

* Corresponding author.
E-mail addresses: wenyuanwang@my.swjtu.edu.cn (W. Wang), hqzhao_swjtu@126.com (H. Zhao).

https://doi.org/10.1016/j.sigpro.2017.10.006

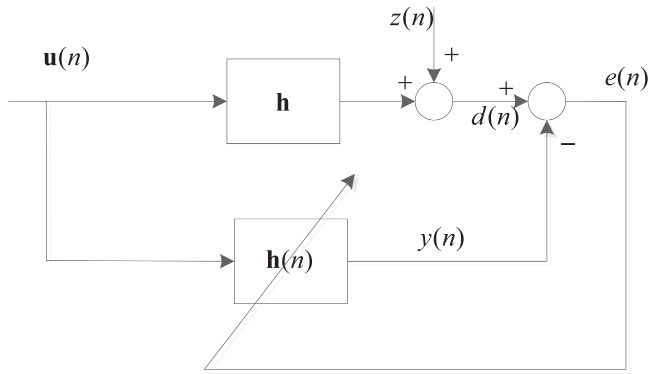

Fig. 1. Diagram of system identification.

In [35], Yang et al. formulated the constrained H∞ filtering problem for the dynamical systems perturbed by bounded disturbances with interval state constraints. In adaptive filter theory, Nascimento proposed RLS adaptive filter with inequality constraint where the coefficients of adaptive are boxed constraint [30].

In this paper, motived by the RLS adaptive filter with boxed-constraint [30], the boxed-constraint least mean square (BXCLMS) algorithm is derived to identify a particular system in which the parameters to be estimated are under boxed-constraints, which has lower computation complexity than RLS adaptive filter. The proposed algorithm is obtained by using the KKT conditions which are established for any convex cost function. In addition to the derivation of the calculus, we give the performance analysis of the proposed algorithm including the mean performance analysis and the mean square performance analysis. Finally, extensive simulations are carried out to verify the analysis results presented. It is worth emphasizing that the proposed method of this paper can be easily extended to the other variants LMS algorithm, which is our future work.

The paper is organized as follows. In Section 2, the boxed-constraint least mean square algorithm is provided. In Section 3, the perform analysis of proposed algorithm is conducted. Section 4 conducts the simulations. Finally, the conclusion is drawn in Section 5.

2. Boxed-constraint least mean square algorithm

Considering the estimation of a linear plant $\mathbf{h} = [h_1, h_2, ..., h_N]^T$, as shown in Fig. 1, the observed system output is given by

$$d(n) = \mathbf{h}^T\mathbf{u}(n) + z(n) \tag{2}$$

where $z(n)$ is the additive white Gaussian noise with power σ_z^2, and $\mathbf{u}(n) = [u(n), u(n-1), ..., u(n-N+1)]^T$ represents a column vector with the zero mean Gaussian signal $u(n)$.

In parameter estimation of some system, such as power system and overlapping community detection, boxed-constraint is usually necessary. Consequently, this type optimum problem can be formulated as follows,

$$\begin{aligned} &\mathbf{h}^o = \arg\min_{\mathbf{h}} J(\mathbf{h}) \\ &\text{subject to} \quad h_i \geq \eta \\ &\qquad\qquad\quad h_i \leq \xi, \quad i = 1, 2, ..., N \end{aligned} \tag{3}$$

where ξ is the upper bound of the estimate $\mathbf{h}$ of the system, η is the lower bound of the estimate $\mathbf{h}$ of the system, $\eta < \xi$ holds, h_i is the ith entry of $\mathbf{h}$, $J(\mathbf{h})$ is convex cost function, and $\mathbf{h}^o = [h_1^o, h_2^o, ..., h_N^o]^T$ is the optimal solution.

To solve above optimum problem, the KKT condition is considered as follows [19],

$$\nabla_{\mathbf{h}} L(\mathbf{h}^o,\ \boldsymbol{\lambda}_1^o,\ \boldsymbol{\lambda}_2^o) = 0 \tag{4}$$

$$\left(h_i^o - \eta\right)\left[\boldsymbol{\lambda}_1^o\right]_i = 0, \quad i = 1, 2, ..., N, \tag{5}$$

$$\left(\xi - h_i^o\right)\left[\boldsymbol{\lambda}_2^o\right]_i = 0, \quad i = 1, 2, ..., N, \tag{6}$$

where $\nabla_{\mathbf{h}}$ denotes the gradient operator in regard to $\mathbf{h}$, $\boldsymbol{\lambda}_1^o$ and $\boldsymbol{\lambda}_2^o$ are the vectors of optimum Lagrange multipliers, and $L(\mathbf{h},\ \boldsymbol{\lambda}_1,\ \boldsymbol{\lambda}_2) = J(\mathbf{h}) - \boldsymbol{\lambda}_1\mathbf{h} + \boldsymbol{\lambda}_2\mathbf{h}$ means the Lagrange function with the vectors of Lagrange multipliers $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$. Then, we can get the following formula

$$\nabla_{\mathbf{h}} L(\mathbf{h},\ \boldsymbol{\lambda}_1,\ \boldsymbol{\lambda}_2) = \nabla_{\mathbf{h}} J(\mathbf{h}) - \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2. \tag{7}$$

Replacing $\mathbf{h}, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2$ by $\mathbf{h}^o$, $\boldsymbol{\lambda}_1^o, \boldsymbol{\lambda}_2^o$, respectively, we have

$$\nabla_{\mathbf{h}} L(\mathbf{h}^o, \boldsymbol{\lambda}_1^o,\ \boldsymbol{\lambda}_2^o) = \nabla_{\mathbf{h}} J(\mathbf{h}^o) - \ \boldsymbol{\lambda}_1^o + \boldsymbol{\lambda}_2^o = 0. \tag{8}$$

Plugging back into Eq. (6) yields

$$\left(\xi - h_i^o\right)[\nabla_{\mathbf{h}} J(\mathbf{h}^o) - \ \boldsymbol{\lambda}_1^o]_i = 0. \tag{9}$$

Arranging (9) leads to the following formula

$$\left(\xi - h_i^o\right)[\nabla_{\mathbf{h}} J(\mathbf{h}^o)]_i - \ [\boldsymbol{\lambda}_1^o]_i\xi + [\boldsymbol{\lambda}_1^o]_i h_i^o = 0. \tag{10}$$

Substituting (5) into (10) yields

$$\left(\xi - h_i^o\right)[\nabla_{\mathbf{h}} J(\mathbf{h}^o)]_i = [\boldsymbol{\lambda}_1^o]_i(\xi - \eta). \tag{11}$$

Multiplying both sides of (11) by $(h_i^o - \eta)$ and using (5), we have

$$\left(h_i^o - \eta\right)\left(\xi - h_i^o\right)[\nabla_{\mathbf{h}} J(\mathbf{h}^o)]_i = 0. \tag{12}$$

Using the fix-point iteration method to solve this optimum problem [18], the component-wise gradient descent algorithm is expressed as

$$\begin{aligned} h_i(n+1) = h_i(n) + \mu_i f_i(\mathbf{h}(n))(h_i(n) - \eta)(\xi - h_i(n)) \\ [-\nabla_{\mathbf{h}} J(\mathbf{h}(n))]_i, \end{aligned} \tag{13}$$

where μ_i is the step size, $f_i(\mathbf{h}(n))$ is the positive function with respect to the $\mathbf{h}(n)$.

Consider the cost function

$$J(\mathbf{h}) = E\left[\left|d(n) - \mathbf{h}^T\mathbf{u}(n)\right|^2\right] \tag{14}$$

According to [2], assuming that $\mathbf{p}$ represents the cross-correlation vector between the input vector $\mathbf{u}(n)$ and the desired response d(n) and $\mathbf{R}_u$ denotes autocorrelation matrix of $\mathbf{u}(n)$, the gradient $\nabla_{\mathbf{h}} J(\mathbf{h}(n))$ can be expressed as

$$\nabla_{\mathbf{h}} J(\mathbf{h}) = 2(\mathbf{R}_u\mathbf{h} - \mathbf{p}) \tag{15}$$

Combining (13) and (15), the iteration update of proposed algorithm is obtained as

$$\begin{aligned} h_i(n+1) = h_i(n) + 2\mu_i f_i(h(n))(h_i(n) - \eta)(\xi - h_i(n)) \\ [\mathbf{R}_u\mathbf{h} - \mathbf{p}]_i \end{aligned} \tag{16}$$

Then, setting $f_i(h(n))$ to 0.5 for all i and replacing the correlation matrix and the cross-correlation vector with their instantaneous values $\mathbf{u}(n)\mathbf{u}^T(n)$ and $\mathbf{u}(n)d(n)$, respectively, we obtain

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \mu e(n)\mathbf{D}_u(n)[\xi\mathbf{1} - \mathbf{h}(n)] \circ [\mathbf{h}(n) - \eta\mathbf{1}] \tag{17}$$

where $\mathbf{1}$ represents the column vector whose entries are all one, $\mathbf{D}_u(n)$ denotes a diagonal matrix whose entries are given by vector $\mathbf{u}(n)$, $\circ$ means the Hadamard product operation and $e(n)$ is the estimated error which is defined as

$$e(n) = d(n) - \mathbf{h}^T(n)\mathbf{u}(n) \tag{18}$$

Remark. Compared with the update equation of conventional LMS algorithm, the update Eq. (17) has the extra multiplying factor $[\xi\mathbf{1} - \mathbf{h}(n)] \circ [\mathbf{h}(n) - \eta\mathbf{1}]$. Then, the update term of the ith entry of $\mathbf{h}(n)$ has the extra factor $[\xi - h_i(n)] \circ [h_i(n) - \eta]$ compared to